Universidad Complutense Madrid

PROGRAMACIÓN MATEMÁTICA:

Métodos de optimización

Begoña Vitoriano

bvitoriano@mat.ucm.es

www.mat.ucm.es/~bvitoria

Andrés Ramos

Andres.ramos@iit.icai.upcomillas.es

http://www.iit.upcomillas.es/~aramos

ÍNDICE

I. Optimización lineal

I.1. Introducción

La programación lineal (LP) es la aplicación clásica por excelencia y la más desarrollada de la optimización. En cada momento se están ejecutando miles de aplicaciones basadas en LP. Los modelos de LP son más utilizados que todos los otros tipos de optimización juntos. Abarcan cualquier tipo de actividad humana como economía, finanzas, marketing, organización de la producción, planificación de la operación, selección de procesos, asignación de tareas, etc. Su importancia se debe a la existencia de técnicas potentes, estables y robustas para encontrar el óptimo que han permitido su uso en multitud de aplicaciones.

Sea el siguiente problema de programación lineal genérico en *forma estándar* o *canónica*:

$$\begin{aligned} &\min_{x_j} z = \sum_j c_j x_j \\ &\sum_j a_{ij} x_j = b_i \\ &x_j \geq 0 \end{aligned} \tag{1.1}$$

Las variables reciben también el nombre de *actividades, decisiones* o *columnas*. Las restricciones se denominan también *recursos* o *filas*.

La programación lineal se sustenta en las siguientes hipótesis sobre las ecuaciones y variables que conforman el problema matemático:

- *Proporcionalidad*

  La contribución de cada actividad (variable) x_j al valor de la función objetivo z es proporcional al nivel de la actividad, $c_j x_j$. La contribución de cada actividad x_j al valor de la parte izquierda de cada restricción es proporcional al nivel de la actividad, $a_{ij} x_j$.

- *Aditividad*

  Cada ecuación en un problema LP es la suma de las contribuciones individuales de las respectivas actividades.
- *Divisibilidad*

  Cualquier variable puede tomar cualquier valor, no necesariamente entero, que satisfaga las restricciones incluyendo las de no negatividad.
- *Certidumbre*

  Los parámetros (constantes) de un problema LP se suponen conocidos con certidumbre (pueden ser estimaciones, pero éstas se tratan como valores conocidos).

I.2. Solución gráfica

Supongamos el siguiente problema de programación lineal en un espacio bidimensional, que gráficamente queda representado en la figura adjunta.

$$
\begin{array}{llll}
\max\ z = 3x_1 + 5x_2 & & & \\
x_1 & & \le 4 & \\
 & 2x_2 & \le 12 & \\
3x_1 & +2x_2 & \le 18 & \\
x_1, & x_2 & \ge 0 &
\end{array}
$$

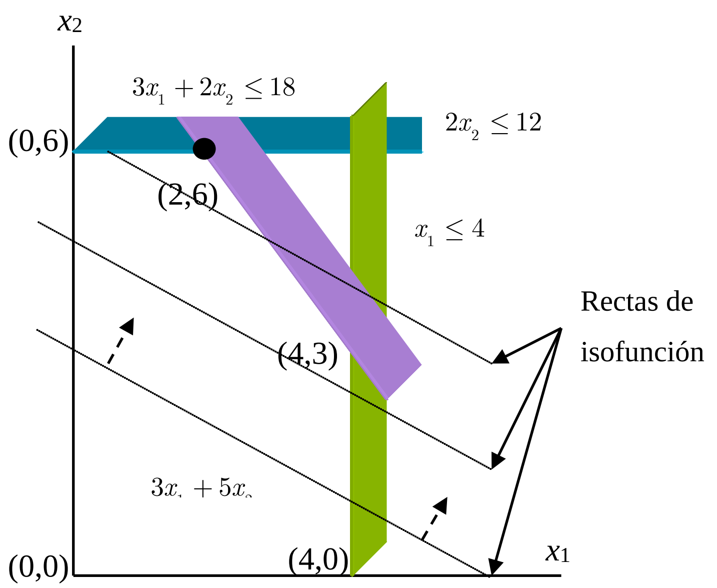


Figura 2.1 Solución gráfica de un problema LP.

La *región factible* está encerrada por las tres rectas que constituyen las restricciones más los dos ejes, que son las rectas correspondientes a la no negatividad de las variables. La solución óptima se halla gráficamente desplazando paralela a sí misma la recta de isofunción objetivo en el sentido de maximización (dirección del gradiente de la función objetivo, c) hasta que se alcanza el último punto de la región factible. Ésta corresponde al punto $(x_1, x_2) = (2, 6)$ con un valor de la función objetivo $\hat{z} = 36$.

Otras posibilidades que pueden aparecer al resolver un problema de programación lineal son las siguientes: solución con múltiples óptimos y solución no acotada. Por otra parte, si la región factible es el conjunto vacío, el problema es infactible.

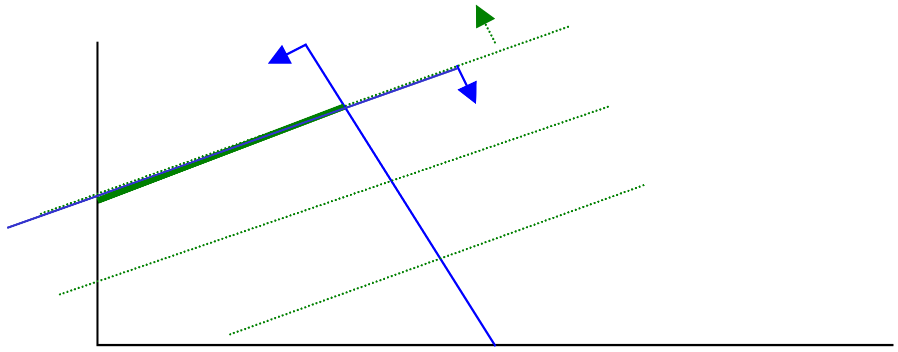


Figura 2.2 Solución gráfica de un problema LP con múltiples óptimos.

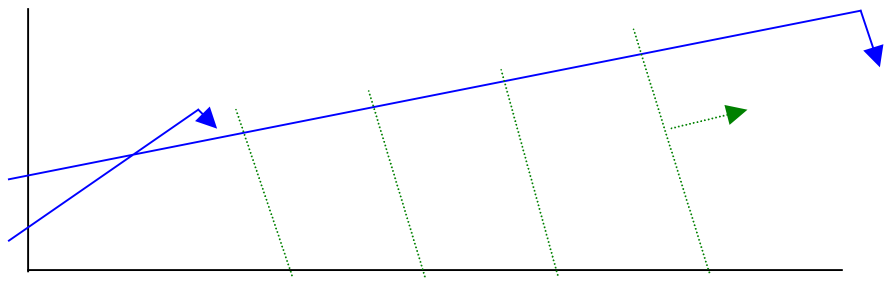


Figura 2.3 Solución gráfica de un problema LP con solución no acotada.

I.3. Geometría de la programación lineal

Hay una serie de conceptos geométricos que hay que definir antes de analizar la geometría de un problema de programación lineal.

Un *hiperplano* está definido por el conjunto de puntos x_j que cumplen una ecuación cualquiera del problema LP, $\sum_j a_{ij} x_j = b_i$. Un hiperplano en dos dimensiones es una recta. El hiperplano divide el espacio en dos semiespacios definidos por $\sum_j a_{ij} x_j \le b_i$ y $\sum_j a_{ij} x_j \ge b_i$.

Un *poliedro* es la región definida por la intersección de un conjunto finito de semiespacios. La región factible de un problema de programación lineal es un poliedro. Un *politopo* es un poliedro no vacío y acotado (por ejemplo, las regiones factibles de las figuras 2.1 y 2.2).

Un *vértice* o *punto extremo* de un poliedro se define como un punto que no puede ponerse como combinación lineal convexa de dos puntos diferentes del conjunto. Se determina por la intersección de n hiperplanos. En dos dimensiones, son dos rectas.

El *contorno* de una región factible contiene las soluciones factibles que están en uno o más hiperplanos. *Arista* es el segmento obtenido por solución de $n-1$ hiperplanos, siendo sus extremos soluciones factibles vértices.

Teorema de existencia de puntos extremos

Sea $S=\{x \,/\, Ax=b, x\geq 0\}\neq\varnothing$, $A\in\mathbb{R}^{m\times n}$, $r(A)=m$. Entonces S tiene al menos un punto extremo.

Teorema de representación de un politopo

Sea $S=\{x \,/\, Ax=b, x\geq 0\}\neq\varnothing$ y acotado, $A\in\mathbb{R}^{m\times n}$, $r(A)=m$. Sean $x_1,\ldots,x_k$ los puntos extremos de S. Entonces $x\in S$ si y sólo si x se puede expresar como una combinación lineal convexa de los puntos extremos del conjunto, es decir, $\exists\lambda_1,\ldots,\lambda_k\geq 0$, tales que $\sum_i \lambda_i = 1$ y $x=\sum_i \lambda_i x_i$. Cualquier punto de un politopo se puede expresar como combinación lineal convexa de los vértices del mismo.

Una aplicación inmediata del resultado anterior es su aplicación al problema de programación lineal, ya que es fácilmente demostrable que buscar en una región factible acotada el punto que optimiza una determinada función objetivo es equivalente a buscar el punto extremo que lo hace. Es decir, entonces la búsqueda del óptimo se limita a optimizar en un número finito de puntos: los vértices o puntos extremos.

Sin embargo, este resultado no es de una aplicación inmediata ya que, por un lado, el número total de vértices o puntos extremos posibles del politopo (no todos ellos son necesariamente factibles) puede ser muy grande (es de $C_{n,m}=\binom{n}{m}=\dfrac{n!}{m!(n-m)!}$ y crece exponencialmente con el tamaño del problema) y, por otro lado, hay que calcular cuáles son los puntos extremos, que si bien están caracterizados y existe el procedimiento para hacerlo puede ser largo y costoso. Además en el caso de que no haya ningún punto factible, no sería detectado hasta la prueba de todos los puntos candidatos a ser extremos ($C_{n,m}$ puntos) y, en el caso de que la solución sea no acotada, con tal procedimiento no se detectaría esta situación.

Por lo tanto, se impone la búsqueda del óptimo de una forma dirigida, es decir, con ciertos criterios que favorezcan esta búsqueda y que además permitan la detección de situaciones como las descritas en el apartado anterior. Estos criterios se plasman en un método o algoritmo de tipo algebraico que se describe en la siguiente sección.

I.4. Método simplex

El método simplex, que debe su origen a George B. Dantzig en 1947 (Dantzig, 1949), es un procedimiento algebraico iterativo para resolver un problema de programación lineal que tiene conceptos geométricos subyacentes. Recorre exclusivamente los vértices (o puntos extremos) de la región factible. La idea geométrica subyacente se describe a continuación. Partiendo de un vértice inicial (luego se describirá el procedimiento para obtenerlo), se va moviendo de un vértice a otro vértice adyacente mejorando la función objetivo. Se selecciona la arista con mayor tasa de mejora de la función objetivo para alcanzar el vértice adyacente, hasta alcanzar un punto en el que no existe ninguna arista incidente por la que mejorar. En el ejemplo anterior, partiendo del vértice (0,0) con función objetivo $\hat{z} = 0$, el método simplex llegaría al óptimo (2,6) con función objetivo $\hat{z} = 36$ después de pasar por el vértice (0,6) con función objetivo $\hat{z} = 30$, tal como se aprecia en la figura. Se comprueba la condición de óptimo porque en dicho vértice las tasas de mejora de la función objetivo hacia vértices adyacentes son negativas.

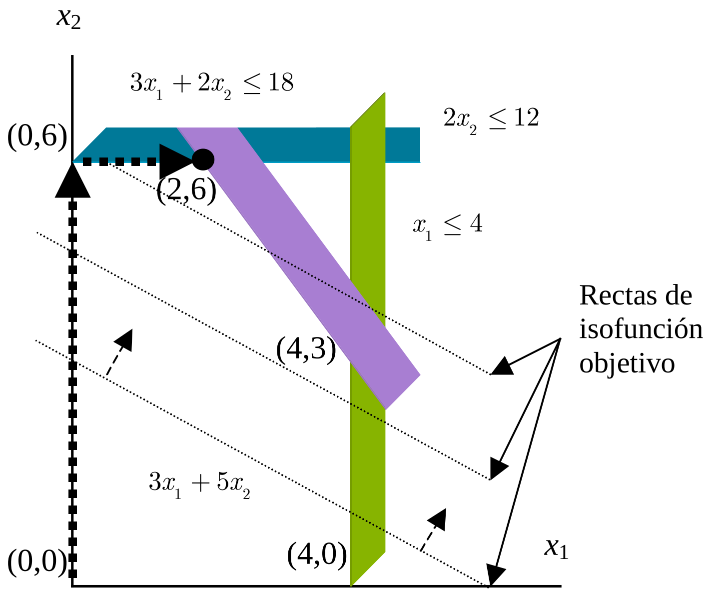


Figura 2.4 Representación gráfica del método simplex.

Un problema de optimización lineal tiene la siguiente *forma estándar*[1]:

$$\begin{aligned} &\min z = c^T x \\ &Ax = b \\ &x \geq 0 \end{aligned} \tag{1.2}$$

donde $x \in \mathbb{R}^n$ es el vector de las variables del problema, $A \in \mathbb{R}^{m \times n}$ es la matriz de restricciones del problema o matriz del lado izquierdo (*left hand side* LHS), $c \in \mathbb{R}^n$ es el vector de los coeficientes de las variables en la función objetivo o vector de costes, $b \in \mathbb{R}^m$ es el vector del lado derecho (*right hand side* RHS) o vector de cotas de las restricciones, siendo $b \geq 0$, y la variable $z \in \mathbb{R}$ es el valor de la función objetivo.

Por consiguiente,

[1] Por convención en la formulación los vectores son columna, su transposición se representa por un superíndice T, las variables se ubican a la izquierda de las ecuaciones y los coeficientes de las variables preceden a éstas.

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix},\ b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix},\ c = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} \text{ y } x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

El número de restricciones *m*, para el problema expresado en su forma estándar, ha de ser estrictamente menor que el número de variables *n*, $m < n$, y el rango de la matriz de restricciones máximo, $r(A) = m$.

La forma estándar se utiliza como base para introducir y desarrollar el método simplex, no es necesario que el problema se exprese de esta forma para ser resuelto por un optimizador. Los códigos comerciales no necesitan que los problemas sean escritos de esta manera.

El ejemplo anterior expresado en la forma estándar aparece como:

$$\begin{array}{llllll} \min\ z = -3x_1 - 5x_2 & & & & & \\ x_1 & & +x_3 & & & = 4 \\ & 2x_2 & & +x_4 & & = 12 \\ 3x_1 & +2x_2 & & & +x_5 & = 18 \\ x_1, & x_2, & x_3, & x_4, & x_5 & \geq 0 \end{array}$$

siendo $A = \begin{bmatrix} 1 & \cdot & 1 & \cdot & \cdot \\ \cdot & 2 & \cdot & 1 & \cdot \\ 3 & 2 & \cdot & \cdot & 1 \end{bmatrix},\ b = \begin{bmatrix} 4 \\ 12 \\ 18 \end{bmatrix},\ c = \begin{bmatrix} -3 \\ -5 \\ \cdot \\ \cdot \\ \cdot \end{bmatrix} \text{ y } x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix}.$

Para convertir un problema de programación lineal cualquiera a su forma estándar se efectúan algunas transformaciones:

Función objetivo

La dirección de maximización se trata como una minimización del valor negativo de la función objetivo.

$$\max\ z \Rightarrow \min -z \qquad\qquad (\max\ z = -\min -z)$$

Restricciones

En las restricciones de tipo ≤ se introduce una *variable de holgura* (*slack*, en inglés) $u_i \geq 0$.

$$\sum_j a_{ij} x_j \leq b_i \Rightarrow \sum_j a_{ij} x_j + u_i = b_i$$

En las restricciones de tipo ≥ se introduce una *variable de exceso o de holgura* $v_i \geq 0$.

$$\sum_j a_{ij} x_j \geq b_i \Rightarrow \sum_j a_{ij} x_j - v_i = b_i$$

Las variables originales x_j se denominan variables *naturales* o *estructurales*.

Variables

Si una variable es negativa y no acotada inferiormente $-\infty \leq x_j \leq 0$, simplemente se sustituye por una variable que tome su valor opuesto $x_j = -y_j$ siendo $0 \leq y_j \leq \infty$.

Si la variable negativa está acotada inferiormente por un valor $L_j < 0$, $L_j \leq x_j \leq 0$, se desplaza en dicha cantidad $x_j = y_j + L_j$ siendo $0 \leq y_j \leq -L_j$. Posteriormente se ve cómo se tratan las variables acotadas.

Si la variable es libre (no acotada inferior ni superiormente, $-\infty \leq x_j \leq \infty$) se sustituye por dos variables que corresponden a su componente positiva y negativa $x_j = x_j^+ - x_j^-$ siendo $0 \leq x_j^+ \leq \infty$ y $0 \leq x_j^- \leq \infty$.

Examinemos los tipos de soluciones del problema de programación lineal:

Solución factible

Aquélla que satisface todas las restricciones.

Cualquier punto del interior o del contorno de la región factible.

Solución infactible

Aquélla que viola al menos una restricción.

Cualquier punto del exterior de la región factible.

Solución básica factible

Una *base* de la matriz de restricciones o matriz del problema es una matriz cuadrada no singular (es decir, tiene inversa) de rango máximo en A, $r(A) = m$. Las variables asociadas a las columnas que forman la base se denominan *variables básicas* y el resto *variables secundarias* (o *no básicas*).

Solución básica factible es aquélla con m variables básicas que pueden tomar valor diferente de 0 y $(n-m)$ variables no básicas que toman valor 0, está asociada a una base de la matriz. Los valores de las variables básicas, si la base es B, se calculan por resolución de un sistema de $m \times m$ ecuaciones, es decir, se obtienen calculando $B^{-1}b$.

Las variables no básicas en un problema de programación lineal con variables acotadas superior e inferiormente se hallan en uno de sus límites mientras que las variables básicas se encuentran estrictamente dentro de los límites. Este caso corresponde a una *solución básica factible no degenerada*. Si alguna de las variables básicas se encuentra en uno de sus límites se denomina *solución básica factible degenerada*.

Toda solución básica factible está asociada a un único punto extremo o vértice y todo punto extremo está asociado a alguna solución básica factible. Sin embargo, puede existir un punto extremo que corresponda a dos bases diferentes en presencia de degeneración.

Dos soluciones factibles vértices son *adyacentes* si la línea que los conecta es una arista, es decir, si todas excepto una de sus variables son iguales. Cada solución factible vértice tiene n soluciones factibles vértices adyacentes. Moverse de una solución básica factible a otra adyacente es cambiar de una variable no básica a básica y al contrario para otra variable.

Solución óptima

Aquella solución básica factible (vértice) con mejor valor de la función objetivo. Pueden existir múltiples soluciones óptimas. Esta situación, que por inspección del problema de dos dimensiones parecería infrecuente, es una situación habitual en la realidad por razón de que el criterio de optimalidad se comprueba numéricamente, es decir, cuando la tasa de mejora de la función objetivo está por debajo de cierta tolerancia muy pequeña pero no cero.

Existe un resultado fundamental en el que se fundamenta la validez del algoritmo del simplex,

Teorema fundamental de la programación lineal

Dado un problema de programación lineal en forma estándar, se tiene que:

- Si admite una solución factible, admite al menos una solución básica factible.
- Si admite una solución óptima finita, admite al menos una solución básica óptima.

El primer apartado es equivalente al teorema geométrico de existencia de puntos extremos y el segundo a la aplicación del problema de representación que asegura que la solución óptima se alcanza en un punto extremo (solución básica). Si hay múltiples óptimos, en una región factible acotada, al menos dos deben ser soluciones factibles vértices adyacentes.

El comportamiento del método simplex en el peor caso es muy pobre. Sin embargo, en la práctica recorre un número muy inferior al número total de vértices posibles. Para el problema anterior el método simplex recorre 3 vértices (incluyendo el inicial) de un número total posible (factibles e infactibles) de $C_{5,2} = \binom{5}{2} = 10$.

Como criterio aproximado se puede estimar que en media el número de iteraciones necesarias para resolver un problema lineal por el método simplex es proporcional al número de restricciones m y el tiempo de solución es proporcional a m^3.

Veamos a continuación cómo se explica y desarrolla el método simplex. Supongamos que se dispone de una solución básica factible $x^T = \begin{bmatrix} x_B & x_N \end{bmatrix}^T$. Sea B la base, $A = \begin{bmatrix} B & N \end{bmatrix}$ y $c^T = \begin{bmatrix} c_B & c_N \end{bmatrix}^T$, siendo

$x_B \in \mathbb{R}^m$ vector de *variables básicas*

$x_N \in \mathbb{R}^{n-m}$ vector de *variables no básicas*

$B \in \mathbb{R}^{m \times m}$ matriz *base*

$N \in \mathbb{R}^{m \times (n-m)}$ matriz no básica

$c_B \in \mathbb{R}^m$ coeficientes de las variables básicas en la función objetivo

$c_N \in \mathbb{R}^{n-m}$ coeficientes de las variables no básicas en la función objetivo

Reformulando el sistema de ecuaciones $Ax = b$ y siempre que B sea no singular

$$Bx_B + Nx_N = b \Leftrightarrow x_B = B^{-1}(b - Nx_N) = B^{-1}b - B^{-1}Nx_N \tag{1.3}$$

La función objetivo se puede expresar entonces como

$$z = c_B^T x_B + c_N^T x_N = c_B^T B^{-1} b - c_B^T B^{-1} N x_N + c_N^T x_N = c_B^T B^{-1} b + \left[c_N^T - c_B^T B^{-1} N\right] x_N \tag{1.4}$$

Lo que se ha hecho es simplemente un cambio de base, es decir, expresar los elementos del problema en función de la base B. En este caso, la solución básica asociada a la base B, viene determinada por los valores de x_B ya que las variables no básicas x_N toman valor 0 (es decir, se resuelve un sistema de ecuaciones lineales para determinar el valor de las variables básicas)

$$x_B = \hat{b} = B^{-1} b \tag{1.5}$$

$$\hat{z} = c_B^T B^{-1} b = c_B^T x_B \tag{1.6}$$

Definamos

$$w^T \equiv c_B^T B^{-1} \tag{1.7}$$

$$Y \equiv B^{-1} N \tag{1.8}$$

$$\hat{c}_N^T \equiv c_N^T - c_B^T B^{-1} N = c_N^T - w^T N = c_N^T - c_B^T Y \tag{1.9}$$

y para cada variable no básica x_j su *coste reducido* $\hat{c}_j$ (o *derivada direccional*) en particular sería

$$\hat{c}_j \equiv c_j - c_B^T B^{-1} a_j = c_j - w^T a_j = c_j - c_B^T y_j = c_j - z_j \tag{1.10}$$

siendo a_j la columna correspondiente a la variable x_j en la matriz A.

Conviene resaltar el significado del vector de costes reducidos $\hat{c}_N$, tal como se desprende de la expresión de la función objetivo. Para ello reformulando la expresión de la función objetivo en función de la base resulta

$$z = c_B^T x_B + \sum_{j \in I_N} \hat{c}_j x_j = \hat{z} + \sum_{j \in I_N} (c_j - z_j) x_j$$

Así, el c*oste reducido* de una variable $\hat{c}_j$ es el cambio en la función objetivo debido a un incremento unitario de una variable no básica. En el óptimo de un problema de minimización todos los costes reducidos de las variables no básicas son positivos o nulos, ya que en tal caso incrementar una variable desde su valor 0 a otro mayor irá en contra de la minimización. Los costes reducidos de las variables básicas son cero. Si el coste reducido de una variable no básica es cero en la solución óptima entonces pueden existir múltiples soluciones óptimas, siendo al menos dos de ellas vértices. En un apartado posterior se menciona analiza expresamente esta posibilidad.

Veamos a continuación cómo se utilizan las expresiones derivadas anteriormente. Si elegimos como variables básicas $x_B = \begin{bmatrix} x_3 \\ x_4 \\ x_5 \end{bmatrix}$ y como no básicas $x_N = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ entonces

$$c_B = \begin{bmatrix} \cdot \\ \cdot \\ \cdot \end{bmatrix},\ c_N = \begin{bmatrix} -3 \\ -5 \end{bmatrix},\ B = \begin{bmatrix} 1 & \cdot & \cdot \\ \cdot & 1 & \cdot \\ \cdot & \cdot & 1 \end{bmatrix} = B^{-1} \text{ y } N = \begin{bmatrix} 1 & \cdot \\ \cdot & 2 \\ 3 & 2 \end{bmatrix}.$$

Los valores que toman dichas variables serán $x_B = \hat{b} = B^{-1}b = \begin{bmatrix} 4 \\ 12 \\ 18 \end{bmatrix}$ y $x_N = \begin{bmatrix} \cdot \\ \cdot \end{bmatrix}$, $w^T = c_B^T B^{-1} = \begin{bmatrix} \cdot & \cdot & \cdot \end{bmatrix}$, el valor de la función objetivo será $\hat{z} = c_B^T B^{-1} b = 0$ y los costes reducidos de las variables no básicas $\hat{c}_N^T = c_N^T - w^T N = \begin{bmatrix} -3 & -5 \end{bmatrix}$.

Si alguna variable no básica incrementara su valor desde 0, $x_N > 0$, el valor que tomarían las variables básicas sería

$$x_B = B^{-1}b - B^{-1}Nx_N = \begin{bmatrix} 4 \\ 12 \\ 18 \end{bmatrix} - \begin{bmatrix} 1 & \cdot \\ \cdot & 2 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

y el de la función objetivo

$$z = c_B^T B^{-1} b + \left[c_N^T - c_B^T B^{-1} N \right] x_N = 0 + \begin{bmatrix} -3 & -5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

De acuerdo con esta ecuación la función objetivo disminuye al incrementar desde 0 cualquiera de las variables no básicas x_1 o x_2 por ser negativos sus costes reducidos $\hat{c}_N^T = \begin{bmatrix} -3 & -5 \end{bmatrix}$, es decir, la base actual no es óptima.

La *prueba de optimalidad* se basa en comprobar si existe algún coste reducido negativo (en un problema de minimización) en la función objetivo. En caso de que no exista, la solución básica actual es óptima. En otro caso, se puede seleccionar una de las variables no básicas con coste reducido negativo para entrar en la base. La variable seleccionada x_t, denominada *variable básica entrante*, será la que tenga coste reducido negativo de mayor valor absoluto.

Para determinar hasta dónde se puede incrementar dicha variable sin violar ninguna restricción de no negatividad de las variables básicas se observan las ecuaciones de las variables básicas

$$x_B = B^{-1}b - B^{-1}Nx_N = \hat{b} - Yx_N = \hat{b} - y_t x_t \tag{1.11}$$

siendo a_t la columna de A correspondiente a x_t e $y_t = B^{-1}a_t$ la *columna pivote*.

Podemos escribir esta ecuación componente a componente

$$(x_B)_i = \hat{b}_i - y_{it}x_t \tag{1.12}$$

Si $y_{it} > 0$ entonces $(x_B)_i$ disminuye cuando la variable básica entrante x_t aumenta, hasta alcanzar el valor 0 para $x_t = \hat{b}_i / y_{it}$, $\hat{b}_i > 0$ por ser el valor de la variable básica.

Para valores $y_{it} \le 0$ la variable $(x_B)_i$ aumenta o permanece igual. Si todos los valores fueran $y_{it} \le 0$ el problema sería *no acotado*, puesto que la variable básica entrante se puede incrementar indefinidamente sin que una variable básica alcance el valor 0. La dirección de no acotamiento para dicha variable no básica x_t es un vector formado por el vector $d_B = -y_t = -B^{-1}a_t$ para las variables básicas, $d_t = 1$ para la variable básica entrante x_t y $d_i = 0$ para el resto de variables no básicas. Esta dirección satisface la condición $Ad = 0$ y $d \ge 0$, es un rayo extremo y pertenece al cono de recesión, que se define como el conjunto de direcciones en las que el problema resulta no acotado. Obsérvese que la detección de problema no acotado se puede hacer con cualquier variable no básica no necesariamente con la de coste reducido negativo de mayor valor absoluto.

La variable x_t puede aumentar su valor hasta $\overline{x}_t$ mientras todas las variables básicas sigan siendo no negativas. Así el máximo valor que puede tomar y en el que además una variable básica se hace 0 y puede salir de la base es

$$\overline{x}_t = \min_{1 \le i \le m} \left\{ \frac{\hat{b}_i}{y_{it}} : y_{it} > 0 \right\} \tag{1.13}$$

El valor mínimo2 de la relación previa identifica la variable básica saliente, es decir, la variable básica que se hace no básica (toma valor 0).

---

[2] Si existe más de una relación que alcance el valor mínimo al mismo tiempo usualmente se toma aquél con mayor coeficiente en la columna pivote.

Después del intercambio de variables la nueva solución básica factible es un nuevo vértice adyacente al anterior con mejor valor de la función objetivo. Para la nueva base se calcula el nuevo valor de la función objetivo y de las variables básicas. El resto de variables, las no básicas, siguen siendo cero.

$$\hat{z} = \hat{z} + \hat{c}_t \overline{x}_t \tag{1.14}$$

$$x_B = \hat{b} - y_t \overline{x}_t \tag{1.15}$$

El *método simplex* consiste en el siguiente algoritmo, que realiza iterativamente este cambio de base mejorando la función objetivo:

- *Inicialización*

  Se supone que se dispone de una matriz base B correspondiente a una *solución básica factible inicial* $x_B = \hat{b} = B^{-1}b \geq 0$ y una función objetivo $\hat{z} = c_B^T B^{-1} b = c_B^T x_B$.

- *Prueba de optimalidad*

  Se calcula el vector de *costes reducidos* $\hat{c}_N^T = c_N^T - c_B^T B^{-1} N = c_N^T - c_B^T Y$. Si todos los costes reducidos son no negativos $\hat{c}_N^T \geq 0$ la base actual es óptima. En caso contrario, se selecciona[3] como *variable básica entrante* x_t una variable con coste reducido estrictamente negativo $\hat{c}_t < 0$.

  La prueba de optimalidad es local pero como los problemas de programación lineal son convexos, la solución óptima será global.

- *Iteración*

  Sea $y_t = B^{-1} a_t$, la *columna pivote* t correspondiente a la variable básica entrante. Se busca la *fila pivote* s que verifica

[3] La variable básica entrante seleccionada es la de coste reducido negativo de mayor valor absoluto. Sin embargo, esta opción no considera el cálculo de la relación mínima (es decir, el incremento que se le va a dar a dicha variable) luego la mejora en la función objetivo puede ser pequeña. Tampoco tiene en cuenta el efecto del escalado de las variables en el problema. En general, no existe manera práctica de predecir qué elección de variable básica entrante puede dar lugar al menor número de iteraciones.

$$\frac{\hat{b}_s}{y_{st}} = \min_{1 \le i \le m} \left\{ \frac{\hat{b}_i}{y_{it}} : y_{it} > 0 \right\}$$

y determina la *variable básica saliente* x_s y el elemento pivote y_{st}.

Si $y_{it} \le 0$ para toda fila i entonces el problema es no acotado.

- *Pivotamiento*

  Actualización de la matriz B^{-1} y del vector de variables básicas y volver al paso 2. Más adelante al presentar la forma tabular se muestra cómo se actualiza la matriz B^{-1} en un número reducido de operaciones.

Sigamos aplicando el método simplex al caso ejemplo anterior. Elegimos x_2 como la variable básica entrante por tener el coste reducido negativo con mayor valor absoluto. Calculamos los elementos de la columna pivote

$$y_t = y_2 = B^{-1} a_2 = \begin{bmatrix} \cdot \\ 2 \\ 2 \end{bmatrix}$$

Las relaciones correspondientes a los dos últimos elementos de y_t, los únicos estrictamente positivos, son

$$\hat{b}_2 / y_{22} = 12/2 = 6 \text{ y } \hat{b}_3 / y_{32} = 18/2 = 9$$

Como la primera relación es la menor, x_4 será la variable básica saliente. Se reemplaza en la base la variable básica x_4 por la variable no básica x_2 y tendremos

$$x_B = \begin{bmatrix} x_3 \\ x_2 \\ x_5 \end{bmatrix} \text{ y } x_N = \begin{bmatrix} x_1 \\ x_4 \end{bmatrix}$$

Entonces $B = \begin{bmatrix} 1 & \cdot & \cdot \\ \cdot & 2 & \cdot \\ \cdot & 2 & 1 \end{bmatrix}$, $B^{-1} = \begin{bmatrix} 1 & \cdot & \cdot \\ \cdot & 1/2 & \cdot \\ \cdot & -1 & 1 \end{bmatrix}$ y $N = \begin{bmatrix} 1 & \cdot \\ \cdot & 1 \\ 3 & \cdot \end{bmatrix}$. Los costes de la variables básicas son $c_B^T = [\cdot \quad -5 \quad \cdot]$ y los de las no básicas $c_N^T = [-3 \quad \cdot]$.

El valor de las variables básicas será ahora

$$x_B = \hat{b} = B^{-1}b = \begin{bmatrix} 1 & \cdot & \cdot \\ \cdot & 1/2 & \cdot \\ \cdot & -1 & 1 \end{bmatrix} \begin{bmatrix} 4 \\ 12 \\ 18 \end{bmatrix} = \begin{bmatrix} 4 \\ 6 \\ 6 \end{bmatrix}$$

y el valor de la función objetivo será $\hat{z} = c_B^T B^{-1} b = -30$.

Los coeficientes

$$w^T = c_B^T B^{-1} = \begin{bmatrix} \cdot & -5 & \cdot \end{bmatrix} \begin{bmatrix} 1 & \cdot & \cdot \\ \cdot & 1/2 & \cdot \\ \cdot & -1 & 1 \end{bmatrix} = \begin{bmatrix} \cdot & -5/2 & \cdot \end{bmatrix}$$

y los costes reducidos

$$\hat{c}_N^T = c_N^T - w^T N = \begin{bmatrix} -3 & \cdot \end{bmatrix} - \begin{bmatrix} \cdot & -5/2 & \cdot \end{bmatrix} \begin{bmatrix} 1 & \cdot \\ \cdot & 1 \\ 3 & \cdot \end{bmatrix} = \begin{bmatrix} -3 & 5/2 \end{bmatrix}.$$

Elegimos x_1 como la variable básica entrante por ser la única con un coste reducido negativo. Calculamos los elementos de la columna pivote

$$y_t = y_1 = B^{-1} a_1 = \begin{bmatrix} 1 & \cdot & \cdot \\ \cdot & 1/2 & \cdot \\ \cdot & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ \cdot \\ 3 \end{bmatrix} = \begin{bmatrix} 1 \\ \cdot \\ 3 \end{bmatrix}$$

Las relaciones correspondientes a los elementos de y_t estrictamente positivos son

$$\hat{b}_1 / y_{11} = 4/1 = 4 \text{ y } \hat{b}_3 / y_{31} = 6/3 = 2$$

Como la segunda relación es la menor x_5 será la variable básica saliente. Se reemplaza en la base la variable básica x_5 por la variable x_1 y tendremos

$$x_B = \begin{bmatrix} x_3 \\ x_2 \\ x_1 \end{bmatrix} \text{ y } x_N = \begin{bmatrix} x_5 \\ x_4 \end{bmatrix}$$

Entonces $B = \begin{bmatrix} 1 & \cdot & 1 \\ \cdot & 2 & \cdot \\ \cdot & 2 & 3 \end{bmatrix}$, $B^{-1} = \begin{bmatrix} 1 & 1/3 & -1/3 \\ \cdot & 1/2 & \cdot \\ \cdot & -1/3 & 1/3 \end{bmatrix}$ y $N = \begin{bmatrix} \cdot & \cdot \\ \cdot & 1 \\ 1 & \cdot \end{bmatrix}$. Los costes de la variables básicas son $c_B^T = \begin{bmatrix} \cdot & -5 & -3 \end{bmatrix}$ y los de las no básicas $c_N^T = \begin{bmatrix} \cdot & \cdot \end{bmatrix}$.

El valor de las variables básicas será ahora

$$x_B = \hat{b} = B^{-1}b = \begin{bmatrix} 1 & 1/3 & -1/3 \\ \cdot & 1/2 & \cdot \\ \cdot & -1/3 & 1/3 \end{bmatrix} \begin{bmatrix} 4 \\ 12 \\ 18 \end{bmatrix} = \begin{bmatrix} 2 \\ 6 \\ 2 \end{bmatrix}$$

y el valor de la función objetivo será $\hat{z} = c_B^T B^{-1} b = -36$.

Los coeficientes

$$w^T = c_B^T B^{-1} = \begin{bmatrix} \cdot & -5 & -3 \end{bmatrix} \begin{bmatrix} 1 & 1/3 & -1/3 \\ \cdot & 1/2 & \cdot \\ \cdot & -1/3 & 1/3 \end{bmatrix} = \begin{bmatrix} \cdot & -3/2 & -1 \end{bmatrix}$$

y los costes reducidos

$$\hat{c}_N^T = c_N^T - w^T N = \begin{bmatrix} \cdot & \cdot \end{bmatrix} - \begin{bmatrix} \cdot & -3/2 & -1 \end{bmatrix} \begin{bmatrix} \cdot & \cdot \\ \cdot & 1 \\ 1 & \cdot \end{bmatrix} = \begin{bmatrix} 1 & 3/2 \end{bmatrix}$$

Como todos los costes reducidos son positivos no se puede mejorar la función objetivo, es decir, se ha alcanzado la solución óptima.

I.4.1. Problema de maximización

Si bien se ha dicho que un problema de maximización puede ser transformado en uno de minimización para ser resuelto, también es cierto que el algoritmo anterior es fácilmente adaptable al caso de maximización, ya que sólo afecta a la dirección de mejora y, por lo tanto, sólo hay que modificar el criterio de optimalidad y el de entrada en la base.

En tal caso, una solución será óptima si todos los costes reducidos son menores o iguales que 0 y, en caso de que no sea así, la variable básica entrante es la que tenga coste reducido estrictamente positivo de mayor valor.

I.4.2. Múltiples óptimos

En algunos problemas no existe un único óptimo, sino que puede haber varios. Esta condición es fácilmente detectable pues para que existan múltiples óptimos al finalizar el algoritmo del simplex debe haber al menos una variable no básica cuyo coste reducido sea igual a 0, de modo que si la variable aumenta su valor la función objetivo no se ve afectada, con lo que puede ser introducida en la base.

Sin embargo, no es una condición suficiente que el coste reducido sea 0, hay que asegurar también que cuando esta variable entra en la base lo hace con valor distinto de 0, pues en otro caso, si sustituye a una variable con valor 0, se produce un cambio de base pero no de punto.

Una vez identificadas todas las soluciones básicas factibles (puntos extremos) óptimas, serán soluciones óptimas todas las combinaciones lineales convexas de éstas. Por ejemplo, si existen dos soluciones básicas óptimas, todo el segmento que une los puntos correspondientes son soluciones óptimas del problema.

En algunos problemas, identificar más de una solución óptima puede ser de vital importancia, especialmente si existe más de un criterio de optimización y éstos han sido jerarquizados (optimización multiobjetivo).

I.4.3. Convergencia del algoritmo

La convergencia del algoritmo está asegurada si no existe degeneración. Sin embargo, en caso de degeneración de varias variables básicas el algoritmo puede ciclar, moviéndose por una secuencia de bases que correspondan al mismo punto extremo sin mejorar la función objetivo. Existen procedimientos, que aquí no se presentan, para evitar esta situación, siendo el más popular de ellos, que no el único, el método lexicográfico.

I.4.4. Variables acotadas superiormente

Las cotas superiores no se deben tratar como restricciones, esto tiene ventajas computacionales. En el método simplex las cotas superiores no se deben sobrepasar cuando la variable básica entrante se incrementa hasta alcanzar un nuevo vértice.

Supongamos una variable acotada superiormente $0 \le x_j \le u_j$, entonces su variable complementaria definida como $y_j = u_j - x_j$ también está acotada por el mismo valor $0 \le y_j \le u_j$. Si $x_j = 0$, x_j es no básica. Si $x_j = u_j$, entonces $y_j = 0$ y, por tanto, y_j será no básica.

En el método simplex, la variable básica saliente es la primera que alcanza su cota 0 al incrementar la variable básica entrante. Ahora esta comprobación se extiende para la variable básica entrante a su cota u_j. Cuando se alcanza la cota superior se cambia la variable x_j a y_j como nueva variable no básica.

I.4.5. Forma tabular

Una forma sencilla de presentar este método y resolver ejemplos pequeños es mediante la *tabla* del simplex. En esta tabla se refleja el problema en función de la base actual en cada iteración. La aportación de esta tabla es que la actualización de los valores para pasar de una base a otra (actualización de la matriz B o B^{-1}) se puede realizar por simple eliminación gaussiana[4], lo que también se denomina *pivotamiento.*

Para el problema original la tabla tiene este aspecto

---

[4] La eliminación gaussiana actualiza la tabla para que en ella aparezca un 1 en el elemento pivote y 0 en el resto de la columna. Para aplicarla seguir los siguientes pasos:

1. Dividir la fila pivote por el elemento pivote y_{st}
2. Restar a las demás filas de la tabla la nueva fila pivote multiplicada por el elemento que se desea pase a ser cero, es decir, el elemento correspondiente de la columna pivote.

Además, se han de hacer 0 los coeficientes de la función objetivo correspondientes a las variables básicas.

Variables básicas	z	x_N	x_B	**Cotas**
$-z$	-1	c_N^T	c_B^T	0
x_B	0	N	B	b

Tabla 2.1 Tabla (A) del simplex en la primera iteración.

siendo c_B^T y c_N^T son los coeficientes de las variables básicas y no básicas actuales en la función objetivo, B las columnas bajo las variables básicas actuales de la matriz original A y N las columnas de las variables no básicas actuales de la matriz original A. Su forma de interpretarla es la siguiente

$$\begin{aligned} -z + c_N^T x_N + c_B^T x_B &= 0 \\ 0z + N x_N + B x_B &= b \end{aligned}$$

En una iteración cualquiera en función de la base B se tiene

Variables básicas	z	x_N	x_B	**Cotas**
$-z$	-1	$c_N^T - c_B^T B^{-1} N$	0	$-c_B^T B^{-1} b$
x_B	0	$B^{-1} N$	I	$B^{-1} b$

Tabla 2.2 Tabla (A) del simplex en una iteración cualquiera.

luego directamente el valor de la función objetivo y de las variables básicas se obtiene de la última columna

$$\begin{aligned} -z &= -c_B^T B^{-1} b \\ x_B &= B^{-1} b \end{aligned}$$

La ecuación de la función objetivo de las nuevas tablas resulta de multiplicar las ecuaciones de las restricciones de las tablas originales respectivas por $w^T = c_B^T B^{-1}$ y restarla de la ecuación de la función objetivo de las tablas originales respectivas. Las ecuaciones de las restricciones de las nuevas tablas resultan de multiplicar las de las tablas originales respectivas por B^{-1}.

Si en la solución inicial la base fuera una matriz identidad y los costes de las variables básicas fueran 0, la tabla original inicial sería la siguiente (por notación se designa con un apóstrofo a lo que es inicial)

Variables básicas	z	x'_N	x'_B	**Cotas**
$-z$	-1	c'^T_N	0	0
x_B	0	N'	I	b

Tabla 2.3 Tabla (B) del simplex en la primera iteración partiendo de una matriz identidad.

y, manteniendo la asignación (ordenación) original por columnas de las variables, la tabla de simplex tiene este aspecto para una iteración cualquiera (siendo N' las columnas de las variables no básicas iniciales en la matriz original A)

Variables básicas	z	x'_N	x'_B	**Cotas**
$-z$	-1	$c'^T_N - c^T_B B^{-1} N'$	$-c^T_B B^{-1}$	$-c^T_B B^{-1} b$
x_B	0	$B^{-1} N'$	B^{-1}	$B^{-1} b$

Tabla 2.4 Tabla (B) del simplex en una iteración cualquiera.

Bajo las columnas de las variables básicas originales inicialmente estaba la matriz identidad I y ahora está la matriz inversa de la base B^{-1}. Luego esta matriz contiene las manipulaciones hechas a las restricciones hasta una iteración cualquiera. c'^T_N son los coeficientes de las variables no básicas iniciales y c^T_B son los coeficientes de las variables básicas de cada iteración.

Observando las tablas 2.1 y 2.2 se puede decir que el método simplex persigue hallar la partición adecuada entre variables básicas y no básicas que cumplen la condición de optimalidad $\hat{c}^T_N = c^T_N - c^T_B B^{-1} N \geq 0$. Observando las tablas 2.3 y 2.4 consistiría en hallar los valores que cumplen la condición de optimalidad del problema, es decir, coeficientes de la función objetivo no negativos en el caso de minimización $\hat{c}^T_N = c'^T_N - c^T_B B^{-1} N' \geq 0$ y $-w^T = -c^T_B B^{-1} \geq 0$.

Veamos con un ejemplo las iteraciones del método simplex en forma tabular. La tabla inicial del problema anterior es:

Variables básicas	z	x_1	x_2	x_3	x_4	x_5	Cotas	Relación
$-z$	–1	–3	–5				0	
x_3		1		1			4	
x_4			2		1		12	12/2=6
x_5		3	2			1	18	18/2=9

La solución básica factible inicial es $(x_1, x_2, x_3, x_4, x_5) = (0, 0, 4, 12, 18)$ y $z = 0$. El recuadro dentro de la tabla identifica la matriz inversa de la base B^{-1}.

Se toma x_2 como variable básica entrante por tener el coste reducido negativo de mayor valor absoluto. Luego su columna es la columna pivote.

Se toma x_4 como variable básica saliente por tener la menor relación. Luego su fila es la fila pivote.

Se actualiza la tabla para anular los elementos de la columna pivote excepto el elemento pivote que se hace 1 (es decir, se anula el coeficiente de la variable básica entrante en la función objetivo y se crea una columna de la matriz identidad en dicha variable) para resolver el sistema de ecuaciones por eliminación gaussiana. Se hacen 0 los coeficientes de la función objetivo correspondientes a las variables básicas.

Variables básicas	z	x_1	x_2	x_3	x_4	x_5	Cotas	Relación
$-z$	–1	–3			5/2		30	
x_3		1		1			4	4/1=4
x_2			1		1/2		6	
x_5		3			–1	1	6	6/3=2

La nueva solución básica factible es $(x_1, x_2, x_3, x_4, x_5) = (0, 6, 4, 0, 6)$ y $z = -30$.

Se toma x_1 como variable básica entrante por tener el único coste reducido negativo. Luego su columna es la columna pivote.

Se toma x_5 como variable básica saliente por tener la menor relación. Luego su fila es la fila pivote.

Se vuelve a actualizar la tabla del simplex.

Variables básicas	z	x_1	x_2	x_3	x_4	x_5	**Cotas**
$-z$	–1				3/2	1	36
x_3				1	1/3	–1/3	2
x_2			1		1/2		6
x_1		1			–1/3	1/3	2

La nueva solución básica factible es $(x_1, x_2, x_3, x_4, x_5) = (2, 6, 2, 0, 0)$ y $\hat{z} = -36$. Ésta es ya la solución óptima por ser todos los costes reducidos, correspondientes a las variables no básicas (x_5, x_4), no negativos $\hat{c}_N^T = \begin{bmatrix} 1 & 3/2 \end{bmatrix}$.

I.4.6. Solución básica factible inicial

El método simplex parte de una solución básica factible y se mueve en cada iteración a otra hasta llegar a la solución óptima. Resuelve este problema de optimización lineal, presentado en su forma estándar

$$\begin{aligned} &\min c^T x \\ &Ax = b \\ &x \geq 0 \end{aligned} \qquad (P)$$

Sin embargo, no siempre es fácil proporcionar una solución básica factible inicial. Como se ha visto, la forma sencilla para hacerlo es partiendo de una matriz identidad en los coeficientes de las variables básicas iniciales.

Para ello en las restricciones de tipo $\leq$ se aprovechan las *variables de holgura*. En las restricciones de tipo $\geq$ se introduce una *variable de exceso* (a veces a ésta también se la denomina de holgura) con signo negativo $v_i \geq 0$ y una *variable artificial* $w_i \geq 0$. Y en las restricciones de tipo $=$ se introduce una *variable artificial* $w_i \geq 0$.

$$\sum_j a_{ij} x_j \leq b_i \Rightarrow \sum_j a_{ij} x_j + u_i = b_i$$

$$\sum_j a_{ij} x_j \geq b_i \Rightarrow \sum_j a_{ij} x_j - v_i + w_i = b_i$$

$$\sum_j a_{ij} x_j = b_i \Rightarrow \sum_j a_{ij} x_j + w_i = b_i$$

Con las variables de holgura y artificiales se parte de una matriz identidad para las variables básicas y el problema de optimización lineal a resolver es

$$\begin{aligned} &\min c^T x \\ &Ax + Ix_a = b \\ &x, x_a \geq 0 \end{aligned} \qquad (P(a))$$

siendo x_a las variables artificiales.

Pero si la solución incluye variables artificiales no es una solución factible del problema original. Por consiguiente, hay que eliminar esas variables aprovechando el propio método simplex, de modo que en su desarrollo estas variables dejen de ser variables básicas. Si al final no se consiguen anular las variables artificiales el problema original será infactible.

Obsérvese que en las hipótesis del método simplex se ha supuesto que la matriz era de rango máximo, sin embargo, esta hipótesis no se ha comprobado nunca. La razón es que al partir de una matriz identidad de rango máximo la hipótesis necesariamente se verifica. Sin embargo, si no se parte de una matriz identidad habría que comprobar que se cumple esta condición. Obsérvese también que aunque el problema original no la cumpliera, al añadir las variables artificiales sí lo hará.

Para lograr anular las variables artificiales se emplean habitualmente dos métodos:

- el *método de la M mayúscula*
- el *método de las dos fases*

El método de la M mayúscula introduce las variables artificiales en la función objetivo original pero penalizadas con un coeficiente M de valor muy elevado. Este coeficiente de penalización puede introducir problemas numéricos por lo que, en general, es preferible el método siguiente.

En el método de las dos fases, la fase I del método simplex tiene como función objetivo la suma de las variables artificiales. Si esta fase I acaba con valor de la función objetivo igual a cero entonces el problema original es factible y se dispone de una solución básica factible siendo no básicas las variables artificiales. La fase II restablece la función objetivo original y aplica de nuevo el método simplex hasta alcanzar la solución óptima a partir de la solución básica factible del final de la fase I.

Supóngase el problema original (P) y el problema una vez que se han introducido las variables artificiales $(P(a))$. El método de las dos fases consta de los siguientes pasos:

Fase I: Resolver el problema

$$\begin{aligned}&\min z' = \sum_a x_a\\&Ax + Ix_a = b\\&x, x_a \geq 0\end{aligned}$$

Existen dos posibles finales de esta fase

- $z' > 0$, entonces el problema (P) no es factible.
- $z' = 0$, entonces el problema (P) es factible y tenemos una solución factible inicial.

Sin embargo, todavía es posible distinguir dos casos en esta situación, antes de pasar a la fase II:

1) No hay variables artificiales en la base se puede pasar a la fase II.

2) Hay variables artificiales en la base con valor 0 (degeneradas). En este caso, la solución obtenida no es una solución básica del problema original ya que incluye variables artificiales. Para eliminarlas de la base, se sustituyen mediante pivotamiento por cualquier variable original no básica (incluyendo de holgura y exceso) cuyo coste reducido sea 0 y cuyo elemento en la fila de la variable artificial sea distinto de 0 (incluso puede ser negativo). Si ninguna variable original tuviera un valor distinto de cero en esa fila (la fila de la variable artificial tiene todo ceros en las variables originales), quiere decir que esa fila es redundante y puede ser eliminada directamente.

Fase II: Eliminar las variables artificiales del problema[5] y retomar la función objetivo original, recalculando los costes reducidos para la solución básica factible obtenida en la fase I.

Veamos con un ejemplo el método de las dos fases.

$$\begin{array}{lll} \min z = 0.4x_1 + 0.5x_2 & & \\ 0.3x_1 & +0.1x_2 & \le 2.7 \\ 0.5x_1 & +0.5x_2 & = 6 \\ 0.6x_1 & +0.4x_2 & \ge 6 \\ x_1, & x_2 & \ge 0 \end{array}$$

Se convierte el problema a la forma estándar y se añaden las variables artificiales pertinentes

$$\begin{array}{lllllll} \min z = 0.4x_1 + 0.5x_2 & & & & & & \\ 0.3x_1 & +0.1x_2 & +x_3 & & & & = 2.7 \\ 0.5x_1 & +0.5x_2 & & +\overline{x}_4 & & & = 6 \\ 0.6x_1 & +0.4x_2 & & & -x_5 & +\overline{x}_6 & = 6 \\ x_1, & x_2, & x_3, & \overline{x}_4, & x_5, & \overline{x}_6 & \ge 0 \end{array}$$

siendo x_3 una variable de holgura, x_5 una variable de exceso y $\overline{x}_4$ y $\overline{x}_6$ variables artificiales.

La fase I del método simplex minimiza la suma de las variables artificiales

$$\begin{array}{lllllll} \min z' = \overline{x}_4 + \overline{x}_6 & & & & & & \\ 0.3x_1 & +0.1x_2 & +x_3 & & & & = 2.7 \\ 0.5x_1 & +0.5x_2 & & +\overline{x}_4 & & & = 6 \\ 0.6x_1 & +0.4x_2 & & & -x_5 & +\overline{x}_6 & = 6 \\ x_1, & x_2, & x_3, & \overline{x}_4, & x_5, & \overline{x}_6 & \ge 0 \end{array}$$

[5] En muchos casos no se eliminan estas variables, sino que se mantienen aunque no se cuente con ellas para entrar en la base, ya que aportan información al final (B^{-1}, ...).

Si esta fase acaba con la función objetivo con valor 0 significa que se ha encontrado una solución básica factible inicial del problema original y se puede seguir con la fase II. En caso contrario el problema es infactible.

Variables básicas	z'	x_1	x_2	x_5	x_3	$\overline{x}_4$	$\overline{x}_6$	**Cotas**
$-z'$	–1					1	1	
x_3		0.3	0.1		1			2.7
$\overline{x}_4$		0.5	0.5			1		6
$\overline{x}_6$		0.6	0.4	–1			1	6

Se transforma la tabla para hacer 0 los coeficientes de las variables básicas

Variables Básicas	z'	x_1	x_2	x_5	x_3	$\overline{x}_4$	$\overline{x}_6$	**Valores**	**Relación**
$-z'$	–1	–1.1	–0.9	1				–12	
x_3		0.3	0.1		1			2.7	9
$\overline{x}_4$		0.5	0.5			1		6	12
$\overline{x}_6$		0.6	0.4	–1			1	6	10

Se elige x_1 como variable básica entrante y entonces x_3 es la variable básica saliente. Se hacen 0 los elementos de la columna pivote excepto el elemento pivote que se hace 1.

Variables Básicas	z'	x_1	x_2	x_5	x_3	$\overline{x}_4$	$\overline{x}_6$	**Valores**	**Relación**
$-z'$	–1		–0.53	1	3.67			–2.1	
x_1		1	0.33		3.33			9	27
$\overline{x}_4$			0.33		–1.67	1		1.5	4.5
$\overline{x}_6$			0.2	–1	–2		1	0.6	3

Se elige x_2 como variable básica entrante y entonces $\overline{x}_6$ es la variable básica saliente. Se hacen 0 los elementos de la columna pivote excepto el elemento pivote que se hace 1.

Variables Básicas	z'	x_1	x_2	x_5	x_3	$\overline{x}_4$	$\overline{x}_6$	**Valores**	**Relación**
$-z'$	–1			–1.67	–1.67		2.67	–0.5	
x_1		1		1.67	6.67		–1.67	8	1.2
$\overline{x}_4$				1.67	1.67	1	–1.67	0.5	0.3
x_2			1	–5	–10		5	3	

Existen dos variables no básicas x_3 y x_5 con igual coeficiente negativo de mayor valor absoluto en la función objetivo. Se puede elegir cualquiera de ellas como variable básica entrante. Arbitrariamente, se elige x_3 y, por tanto, $\overline{x}_4$ es la variable básica saliente. Se hacen 0 los elementos de la columna pivote excepto el elemento pivote que se hace 1.

Variables Básicas	z'	x_1	x_2	x_5	x_3	$\overline{x}_4$	$\overline{x}_6$	**Valores**
$-z'$	–1					1	1	
x_1		1		–5		–4	5	6
x_3				1	1	0.6	–1	0.3
x_2			1	5		6	–5	6

La solución $(x_1, x_2, x_3, \overline{x}_4, x_5, \overline{x}_6) = (6, 6, 0.3, 0, 0, 0)$ es óptima y como $\hat{z}' = 0$ ésta es una solución básica factible inicial para el problema original.

Se plantea entonces la fase II donde se introduce la función objetivo original.

Variables Básicas	z	x_1	x_2	x_5	x_3	$\overline{x}_4$	$\overline{x}_6$	**Valores**
$-z$	–1	0.4	0.5					
x_1		1		–5		–4	5	6
x_3				1	1	0.6	–1	0.3
x_2			1	5		6	–5	6

Se deben hacer 0 los coeficientes de las variables básicas x_1, x_2 y x_3 en la función objetivo

Variables Básicas	z	x_1	x_2	x_5	x_3	$\overline{x}_4$	$\overline{x}_6$	**Valores**	**Relación**

$-z$	–1			–0.5		–1.4	0.5	–5.4	
x_1		1		–5		–4	5	6	
x_3				1	1	0.6	–1	0.3	0.3
x_2			1	5		6	–5	6	1.2

A partir de ahora se ignoran los coeficientes de las variables artificiales para elegir la variable básica entrante para evitar que éstas tomen valor positivo. Por tanto, se elige x_5 como variable básica entrante y entonces x_3 es la variable básica saliente. Se hacen 0 los elementos de la columna pivote excepto el elemento pivote que se hace 1.

Variables Básicas	z	x_1	x_2	x_5	x_3	$\overline{x}_4$	$\overline{x}_6$	**Valores**
$-z$	–1				0.5	–1.1		–5.25
x_1		1			5	–1		7.5
x_5				1	1	0.6	–1	0.3
x_2			1		–5	3		4.5

Como no aparecen costes reducidos negativos (excepto en una variable artificial) la solución $(x_1, x_2, x_3, \overline{x}_4, x_5, \overline{x}_6) = (7.5, 4.5, 0, 0, 0.3, 0)$ es óptima y el valor de la función objetivo es $\hat{z} = 5.25$.

I.4.7. Método simplex revisado

El método simplex en formato tabular efectúa más cálculos de los estrictamente necesarios en cada iteración. De hecho en una iteración cualquiera sólo se necesitan los coeficientes de la variable básica entrante (la columna pivote) no los del resto de variables no básicas. El *método simplex revisado* calcula en cada iteración sólo la información que necesita en dicha iteración. Por consiguiente, requiere menos cálculo y menos almacenamiento.

En forma tabular el método simplex revisado se desarrolla sobre una tabla de dimensiones $(m+1)\times(m+1)$, siendo m el número de restricciones. La tabla del método simplex revisado es la siguiente:

x'_B	**Valores**
$w^T = c_B^T B^{-1}$	$\hat{z} = c_B^T B^{-1} b$
B^{-1}	$\hat{b} = B^{-1} b$

Tabla 2.5 Tabla (B) del simplex revisado en una iteración cualquiera.

En cada iteración se calculan los costes reducidos de las variables no básicas ($\hat{c}_j = c_j - z_j = c_j - w^T a_j$) fuera de la tabla para aplicar el criterio de optimalidad y, en su caso, determinar la variable entrante, x_t. Igualmente, se calcula la columna pivote fuera de la tabla ($y_t = B^{-1} a_t$) para aplicar el criterio de salida y determinar la variable básica saliente ($\frac{\hat{b}_s}{y_{st}} = \min_{1 \le i \le m} \left\{ \frac{\hat{b}_i}{y_{it}} : y_{it} > 0 \right\}$). El pivotamiento se hace sobre la tabla actual, suponiendo que existe una columna adicional (realmente no es añadida) que es la columna pivote.

Obsérvese que todos los elementos necesarios de la iteración anterior están recogidos en esta tabla del simplex revisado y que de una iteración a otra los elementos que se obtienen mediante pivotamiento son los mismos, pero actualizados con la nueva base (consecuencia de la eliminación gaussiana).

I.4.8. Forma producto de la inversa

El mayor esfuerzo computacional del método simplex se lo llevan los cálculos asociados a la inversa de la base B^{-1}. La razón estriba en que, mientras la matriz base B para problemas grandes suele ser cuasivacía, la inversa tiende a ser densa. En primer lugar, este hecho disuade del cálculo explícito de B^{-1} a partir de B en cada iteración. En segundo lugar, B cambia sólo ligeramente en cada iteración luego será posible actualizar tanto B como B^{-1} sin necesidad de recalcularlas en cada paso tal como se hace en el método simplex explicado hasta ahora. Para ello se utiliza el método denominado forma producto de la inversa, que es un avance del método simplex revisado.

Como el cambio en la matriz base B se produce en una columna en cada iteración es fácil actualizar dicha matriz y también la inversa. Si en una iteración cualquiera la columna pivote corresponde a x_t y la fila pivote es la s, la nueva matriz base $\overline{B}$ se puede calcular a partir de la matriz B anterior mediante transformaciones elementales simplemente reemplazando la columna s por la columna pivote y_t. Como $y_t = B^{-1}a_t$ o $By_t = a_t$ entonces

$$\overline{B} = BF \tag{1.16}$$

siendo

$$F = \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & \left[\ y_t\ \right] & & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix}$$

obtenida a partir de la matriz identidad reemplazando la columna s por y_t.

Si denominamos $E = F^{-1}$ la nueva matriz inversa $\overline{B}^{-1}$ se obtendrá a partir de la anterior premultiplicando por la matriz E.

$$\overline{B}^{-1} = EB^{-1} \tag{1.17}$$

donde

$$E = \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & \left[\ \eta\ \right] & & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix} \text{ siendo } \eta = \begin{bmatrix} -y_{1t}/y_{st} \\ \vdots \\ 1/y_{st} \\ \vdots \\ -y_{mt}/y_{st} \end{bmatrix}.$$

La matriz E se llama *matriz elemental* y esta forma de actualizar la matriz inversa se denomina *forma producto*, porque representa la inversa de la base en cualquier iteración k como un producto de matrices elementales E_k de las iteraciones $(k-1, k-2, \dots, 2, 1)$, suponiendo que la matriz base inicial es la matriz identidad.

$$B_k^{-1} = E_{k-1}E_{k-2}\cdots E_2E_1 \tag{1.18}$$

Veamos la actualización de la inversa de la base del ejemplo anterior con la forma producto de la inversa. En la primera iteración la columna pivote es $y_t = \begin{pmatrix} \cdot & 2 & 2 \end{pmatrix}^T$ y el elemento pivote es el segundo, $s = 2$. Luego

$$\eta_1 = \begin{pmatrix} \cdot \\ 1/2 \\ -1 \end{pmatrix},\ E_1 = \begin{pmatrix} 1 & \cdot & \cdot \\ \cdot & 1/2 & \cdot \\ \cdot & -1 & 1 \end{pmatrix} \text{ y } B_2^{-1} = E_1B_1^{-1} = E_1 = \begin{pmatrix} 1 & \cdot & \cdot \\ \cdot & 1/2 & \cdot \\ \cdot & -1 & 1 \end{pmatrix}$$

En la segunda iteración la columna pivote es $y_t = \begin{pmatrix} 1 & \cdot & 3 \end{pmatrix}^T$ y el elemento pivote es el tercero, $s = 3$. Luego

$$\eta_2 = \begin{pmatrix} -1/3 \\ \cdot \\ 1/3 \end{pmatrix},\ E_2 = \begin{pmatrix} 1 & \cdot & -1/3 \\ \cdot & 1 & \cdot \\ \cdot & \cdot & 1/3 \end{pmatrix} \text{ y } B_3^{-1} = E_2E_1 = \begin{pmatrix} 1 & 1/3 & -1/3 \\ \cdot & 1/2 & \cdot \\ \cdot & -1/3 & 1/3 \end{pmatrix}$$

En la implantación informática más que actualizar en cada iteración la matriz se almacenan las transformaciones elementales de manera compacta. Cuando el almacenamiento ocupado es excesivo se elimina y se vuelve a refactorizar la matriz B^{-1} desde el principio.

De hecho, ni siquiera es necesario disponer de la inversa de la matriz base B^{-1} sino utilizarla en productos de dicha matriz por vectores. Puede ser postmultiplicación por un vector como para el cálculo del valor de las variables básicas $\hat{b} = B^{-1}b$ o de la columna pivote $y_t = B^{-1}a_t$ o premultiplicación por un vector como en el caso del cálculo de $w^T = c_B^TB^{-1}$.

La postmultiplicación por un vector se realiza secuencialmente de esta forma

$$B_k^{-1}a = E_{k-1}(E_{k-2}(E_{k-3}(\cdots(E_2(E_1a))\cdots))) \tag{1.19}$$

premultiplicando en primer lugar el vector a por la matriz elemental E_1, el vector resultante premultiplicándolo por E_2 y así sucesivamente. Esta operación se denomina transformación hacia delante (FTRAN) pues recorre hacia alante las matrices elementales

La premultiplicación por un vector se realiza secuencialmente de esta forma

$$c^T B_k^{-1} = ((\cdots(((c^T E_{k-1})E_{k-2})E_{k-3})\cdots)E_2)E_1 \tag{1.20}$$

postmultiplicando en primer lugar el vector c^T por la matriz elemental E_{k-1}, el vector resultante postmultiplicándolo por E_{k-2} y así sucesivamente. Esta operación se denomina transformación hacia atrás (BTRAN) porque recorre hacia atrás las matrices elementales.

Cada una de estas operaciones de producto matriz-vector es muy rápida y simple. Sea E una matriz elemental con un vector η en su columna s. El producto genérico Ea resulta ser

$$Ea = \begin{pmatrix} 1 & & \eta_1 & & \\ & \ddots & \vdots & & \\ & & \eta_s & & \\ & & \vdots & \ddots & \\ & & \eta_m & & 1 \end{pmatrix} \begin{pmatrix} a_1 \\ \vdots \\ a_s \\ \vdots \\ a_m \end{pmatrix} = \begin{pmatrix} a_1 \\ \vdots \\ 0 \\ \vdots \\ a_m \end{pmatrix} + a_s \begin{pmatrix} \eta_1 \\ \vdots \\ \eta_s \\ \vdots \\ \eta_m \end{pmatrix} \tag{1.21}$$

de manera que ni siquiera la matriz elemental E necesita ser formada explícitamente. Se reemplaza el término s del vector a por 0 y a dicho vector se le añade el vector η multiplicado por el escalar a_s.

El producto genérico $c^T E$ resulta ser

$$\begin{aligned} c^T E &= \begin{pmatrix} c_1 & c_2 & \cdots & c_s & \cdots & c_m \end{pmatrix} \begin{pmatrix} 1 & & & \eta_1 & & \\ & 1 & & \vdots & & \\ & & \ddots & \vdots & & \\ & & & \eta_s & & \\ & & & \vdots & \ddots & \\ & & & \eta_m & & 1 \end{pmatrix} \\ &= \begin{pmatrix} c_1 & c_2 & \cdots & c_{s-1} & c^T\eta & c_{s+1} & \cdots & c_m \end{pmatrix} \end{aligned} \tag{1.22}$$

dejando el vector c sin modificación excepto en el término s que es reemplazado por el escalar $c^T\eta$.

I.4.9. Factorización de la matriz base

En cada iteración del método simplex se realiza el cálculo de los costes reducidos de las variables no básicas, del valor de las variables básicas y de la columna pivote en función de la inversa de la base B^{-1}

$$\begin{aligned} w^T &= c_B^T B^{-1} \\ \hat{b} &= B^{-1} b \\ y_t &= B^{-1} a_t \end{aligned} \tag{1.23}$$

Estos cálculos se pueden expresar también como sistemas de ecuaciones en función de la matriz base B

$$\begin{aligned} w^T B &= c_B^T \\ B\hat{b} &= b \\ By_t &= a_t \end{aligned} \tag{1.24}$$

La factorización LU de la matriz base expresa ésta como producto de dos matrices

$$B = LU \tag{1.25}$$

donde L y U son matrices triangular inferior y superior respectivamente. Estas matrices mantienen la condición de ser cuasivacías como la matriz A original. Por consiguiente, se pueden aplicar técnicas específicas de manejo de dichas matrices que requieren poca memoria.

Inicialmente la matriz base B es la identidad. De la misma forma las matrices L y U son también la identidad. La matriz base cambia (se actualiza) al realizar iteraciones del simplex. De la misma forma se hace con las matrices triangulares. Después de un cierto número de iteraciones (por ejemplo, 100) deja de ser eficiente la actualización a partir de matrices elementales. En ese momento se refactoriza la matriz base, es decir, se calculan de nuevo la matrices triangulares L y U correspondientes y se empieza de nuevo el proceso de actualización. En la iteración final también se suele realizar una refactorización.

Con la factorización la resolución de un sistema de ecuaciones se realiza en dos etapas sin necesidad de invertir la matriz. Por ejemplo, para el cálculo del valor de las variables básicas $\hat{b}$

$$\begin{aligned} B\hat{b} &= b \\ LU\hat{b} &= b \end{aligned} \tag{1.26}$$

Si denominamos $\hat{b}_1 = U\hat{b}$ se resuelve primeramente este sistema triangular *hacia delante* para calcular $\hat{b}_1$

$$L\hat{b}_1 = b \tag{1.27}$$

y después se resuelve este otro sistema triangular *hacia atrás* para calcular $\hat{b}$

$$U\hat{b} = \hat{b}_1 \tag{1.28}$$

Este método de resolución de sistemas de ecuaciones es numéricamente más estable que el de forma producto de la inversa.

En cada iteración del método simplex se actualiza la matriz base B mediante una matriz elemental F

$$\overline{B} = BF \tag{1.29}$$

siendo

$$F = \begin{pmatrix} 1 & & & & \\ & \ddots & \left[\begin{matrix} \\ y_t \\ \\ \end{matrix} \right] & & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix}$$

obtenida a partir de la matriz identidad reemplazando la columna s por y_t.

También se pueden calcular las nuevas matrices $\overline{L}$ y $\overline{U}$.

$$\overline{B} = \overline{L}\,\overline{U} \tag{1.30}$$

Luego, en una iteración cualquiera, éstas se pueden obtener como producto de matrices elementales y de las matrices originales de la factorización de manera análoga a como se hace en la forma producto de la inversa.

Alternativamente a la factorización LU se puede utilizar la descomposición de Cholesky

$$B = R^T R \tag{1.31}$$

donde R es una matriz triangular superior.

I.4.10. Estrategias de cálculo de costes reducidos

Para aplicar el criterio de optimalidad en cada iteración es necesario calcular los costes reducidos de las variables no básicas.

$$\hat{c}_N^T \equiv c_N^T - c_B^T B^{-1} N$$

Hasta ahora se suponía que se calculaban los costes reducidos de todas las variables no básicas y se seleccionaba como variable básica entrante a aquélla con coste reducido de menor valor negativo. Existen otras estrategias de cálculo de costes reducidos y de selección de la variable básica entrante (denominadas *pricing strategies* en inglés).

La primera gran división es calcular los costes reducidos de todas las variables no básicas (*full pricing*) o sólo de algunas (*partial pricing*). En el primer caso, se hace mayor esfuerzo computacional en cada iteración y, en principio, menor número de iteraciones del simplex. En el segundo, lo contrario.

Entre las alternativas para la selección de la variable básica entrante t están:

1) La de coste reducido de menor valor negativo $\hat{c}_t$
2) La primera con coste reducido negativo $\hat{c}_t$
3) La de mayor decremento de la función objetivo calculado como $\hat{c}_t \overline{x}_t$, siendo $\hat{c}_t$ el coste reducido de la variable no básica y $\overline{x}_t = \min_{1 \le i \le m} \left\{ \frac{\hat{b}_i}{y_{ij}} : y_{ij} > 0 \right\}$

4) La de mayor mejora de la función objetivo por incremento unitario de la variable no básica calculado como $\hat{c}_t = \min_{1 \le j \le n-m} \frac{\hat{c}_j}{\sqrt{\|y_j\|^2 + 1}}$, siendo $y_j = B^{-1} a_j$. Esta estrategia es la más utilizada habitualmente, en inglés se denomina *steepest-edge*. Tiene la desventaja de que requiere el cálculo de todos los costes reducidos y el esfuerzo computacional adicional asociado al cálculo del denominador.

I.5. Dualidad

Supongamos un problema de programación lineal, denominado *problema primal* (P), expresado en forma estándar como

$$\begin{aligned} &\min z = c^T x \\ &Ax = b \qquad (P) \\ &x \ge 0 \end{aligned} \tag{1.32}$$

Supóngase que se quiere calcular una cota inferior de la función objetivo de este problema antes de resolverlo. Una forma de hacerlo sería operar con las filas de la matriz, multiplicándolas por algún valor $y^T A x = y^T b$ y sumándolas entre ellas de modo que no se superen los coeficientes de la función objetivo $y^T A \le c^T$, en tal caso el lado derecho resultante de tales operaciones sería una cota inferior para el valor de la función objetivo $y^T b \le c^T x$. Desde luego desearemos que esa cota sea lo más alta posible, ya que se podría asegurar que la función objetivo no puede ser menor que ese valor.

Por lo tanto, el problema de encontrar esa cota inferior, sería el problema de encontrar coeficientes y_1 a y_m para las filas de la matriz tales que al multiplicarlos por las filas y sumarlos no superen los coeficientes de la función objetivo, es decir, tales que

$$A^T y \le c$$

y tal que el resultado obtenido en el lado derecho al operar sea máximo. Este nuevo problema planteado a raíz del anterior se denomina *problema dual* (D)

$$\begin{aligned} &\max y_0 = b^T y \\ &A^T y \le c \end{aligned} \quad (D) \tag{1.33}$$

siendo $y \in \mathbb{R}^m$ el vector de variables del problema dual o *variable duales*.

Veamos un ejemplo de este razonamiento. Sea el siguiente problema primal

$$\begin{aligned} &\min 5x_1 - 3x_2 - x_3 + 4x_4 \\ &x_1 - 2x_2 + 3x_3 + x_4 = 7 \\ &12x_1 - 9x_2 - 4x_3 - 4x_4 \ge -3 \\ &8x_1 - 5x_2 + 7x_3 + 9x_4 \le 6 \\ &x_1 \ge 0, x_2, x_3 \text{ libres}, x_4 \le 0 \end{aligned}$$

Queremos obtener una cota inferior de este problema. Definimos y_1, y_2 e y_3 como los pesos de las restricciones por ese orden. La combinación lineal de las restricciones resulta

$$y_1(x_1 - 2x_2 + 3x_3 + x_4) + y_2(12x_1 - 9x_2 - 4x_3 - 4x_4) + y_3(8x_1 - 5x_2 + 7x_3 + 9x_4)$$

reagrupando términos

$$(y_1 + 12y_2 + 8y_3)x_1 + (-2y_1 - 9y_2 - 5y_3)x_2 + (3y_1 - 4y_2 + 7y_3)x_3 + (y_1 - 4y_2 + 9y_3)x_4$$

cada uno de estos coeficientes de x_1, x_2, x_3 y x_4 debe tener una relación con el del problema original. Como $x_1 \ge 0$ para que la función objetivo del nuevo problema sea menor su coeficiente ha de ser menor $y_1 + 12y_2 + 8y_3 \le 5$. Como $x_4 \le 0$ su coeficiente ha de ser mayor $y_1 - 4y_2 + 9y_3 \ge 4$. Como x_2 es libre la relación entre sus coeficientes dependerá del valor que tome. Si $x_2 \ge 0$ entonces $-2y_1 - 9y_2 - 5y_3 \le -3$. Si $x_2 \le 0$, $-2y_1 - 9y_2 - 5y_3 \ge -3$. La única forma de que se puedan cumplir ambas a la vez es $-2y_1 - 9y_2 - 5y_3 = -3$. Análogamente para x_3, $3y_1 - 4y_2 + 7y_3 = -1$. Luego las restricciones que deben cumplir los pesos para que la combinación lineal sea cota inferior son

$$
\begin{aligned}
&y_1 + 12y_2 + 8y_3 \leq 5 \\
&-2y_1 - 9y_2 - 5y_3 = -3 \\
&3y_1 - 4y_2 + 7y_3 = -1 \\
&y_1 - 4y_2 + 9y_3 \geq 4
\end{aligned}
$$

La ecuación $7y_1 - 3y_2 + 6y_3$ ha de tomar el máximo valor posible. Para ello se sustituyen los coeficientes de y_1, y_2 e y_3 en la combinación lineal por sus cotas en las restricciones del primal garantizando que la combinación lineal siga siendo cota inferior.

Como $x_1 - 2x_2 + 3x_3 + x_4 = 7$, y_1 puede tomar cualquier valor, es libre. Como $12x_1 - 9x_2 - 4x_3 - 4x_4 \geq -3$, para que se cumpla $(12x_1 - 9x_2 - 4x_3 - 4x_4)y_2 \geq -3y_2$ necesariamente $y_2 \geq 0$. Como $8x_1 - 5x_2 + 7x_3 + 9x_4 \leq 6$, para que se cumpla $(8x_1 - 5x_2 + 7x_3 + 9x_4)y_3 \geq 6y_3$ necesariamente $y_3 \leq 0$.

Luego el problema, denominado dual, que obtiene la mayor cota inferior del problema primal original es

$$
\begin{aligned}
&\max 7y_1 - 3y_2 + 6y_3 \\
&y_1 + 12y_2 + 8y_3 \leq 5 \\
&-2y_1 - 9y_2 - 5y_3 = -3 \\
&3y_1 - 4y_2 + 7y_3 = -1 \\
&y_1 - 4y_2 + 9y_3 \geq 4 \\
&y_1 \text{ libre}, y_2 \geq 0, y_3 \leq 0
\end{aligned}
$$

Cualquier problema de programación lineal denominado *primal* tiene asociado a él otro problema también de programación lineal llamado *dual*. Tal como se observa en la tabla siguiente, si el problema primal es de maximización con m restricciones y n variables, $(m \times n)$, el problema dual es de minimización con n restricciones y m variables, $(n \times m)$. Se elige este caso particular de optimización para facilitar la presentación de la dualidad pero son generalizables a cualquier problema de optimización lineal.

Primal $(m \times n)$	Dual $(n \times m)$
$\max_x z = c^T x$ $Ax \leq b$ $x \geq 0$	$\min_y y_0 = b^T y$ $A^T y \geq c$ $y \geq 0$
$\max_{x_j} z = \sum_{j=1}^{n} c_j x_j$ $\sum_{j=1}^{n} a_{ij} x_j \leq b_i \quad i = 1,\dots,m$ $x_j \geq 0 \quad j = 1,\dots,n$	$\min_{y_i} y_0 = \sum_{i=1}^{m} b_i y_i$ $\sum_{i=1}^{m} a_{ij} y_i \geq c_j \quad j = 1,\dots,n$ $y_i \geq 0 \quad i = 1,\dots,m$

Para el ejemplo habitual el problema primal y dual son los siguientes

Primal	Dual
$\max z = 3x_1 + 5x_2$ $x_1 \leq 4$ $2x_2 \leq 12$ $3x_1 + 2x_2 \leq 18$ $x_1, x_2 \geq 0$	$\min y_0 = 4y_1 + 12y_2 + 18y_3$ $y_1 + 3y_3 \geq 3$ $2y_2 + 2y_3 \geq 5$ $y_1, y_2, y_3 \geq 0$
$\max z = \begin{bmatrix} 3 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ $\begin{bmatrix} 1 & \cdot \\ \cdot & 2 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \leq \begin{bmatrix} 4 \\ 12 \\ 18 \end{bmatrix}$ $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \geq \begin{bmatrix} 0 \\ 0 \end{bmatrix}$	$\min y_0 = \begin{bmatrix} 4 & 12 & 18 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$ $\begin{bmatrix} 1 & \cdot & 3 \\ \cdot & 2 & 2 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} \geq \begin{bmatrix} 3 \\ 5 \end{bmatrix}$ $\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} \geq \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

A continuación se muestra la tabla completa (vista de izquierda a derecha o viceversa según sea el problema primal) que permite la transformación de un problema primal cualesquiera en su dual y viceversa.

min		max
Variable		**Restricción**
≥ 0	$\Leftrightarrow$	$\leq$
≤ 0	$\Leftrightarrow$	$\geq$
no restringida	$\Leftrightarrow$	=
Restricción		**Variable**
$\geq$	$\Leftrightarrow$	≥ 0
$\leq$	$\Leftrightarrow$	≤ 0
=	$\Leftrightarrow$	no restringida

Tabla 2.6 Tabla de conversión entre primal y dual.

A cada restricción (variable) del primal le corresponde una variable (restricción) del dual. El tipo de restricción (variable) del primal afecta al tipo de variable (restricción) del dual.

Como caso ejemplo general para observar la conversión entre primal y dual se presenta el siguiente problema.

$$
\begin{array}{ll}
\min\limits_{x_1,x_2,x_3} c_1x_1 + c_2x_2 + c_3x_3 & \max\limits_{y_1,y_2,y_3} y_1b_1 + y_2b_2 + y_3b_3 \\
a_{11}x_1 + a_{12}x_2 + a_{13}x_3 \leq b_1 & y_1a_{11} + y_2a_{21} + y_3a_{31} \leq c_1 \\
a_{21}x_1 + a_{22}x_2 + a_{23}x_3 \geq b_2 & y_1a_{12} + y_2a_{22} + y_3a_{32} \geq c_2 \\
a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3 & y_1a_{13} + y_2a_{23} + y_3a_{33} = c_3 \\
x_1 \geq 0, x_2 \leq 0, x_3 \text{ libre} & y_1 \leq 0, y_2 \geq 0, y_3 \text{ libre}
\end{array}
$$

Existe un conjunto de propiedades básicas de dualidad que ligan las soluciones del primal y dual.

Teorema: El dual del problema dual es el problema primal.

Propiedad de la simetría

Para cualquier problema primal y su dual, todas las relaciones entre ellos son simétricas porque el dual de su dual es el primal.

Implícitamente las propiedades que siguen se muestran suponiendo que el *primal* es de *maximización* y el *dual* de *minimización*, pero son completamente generalizables.

Propiedad débil de la dualidad

Si x es una solución factible del primal e y es una solución factible del dual, se verifica que $c^T x \le b^T y$.

Propiedad fuerte de la dualidad

Si $\hat{x}$ es una solución óptima del primal e $\hat{y}$ es una solución óptima del dual, se verifica que $c^T \hat{x} = b^T \hat{y}$.

Propiedad de soluciones básicas complementarias

En cada iteración del método simplex se encuentra una solución básica (vértice) factible del primal y una solución básica complementaria infactible del dual tal que $c^T x = b^T y$ o $z = y_0$. Si x no es óptima para el primal y no es factible para el dual.

Si la solución básica del primal está asociada a la base B, la *solución dual complementaria* tiene la expresión $y^T \equiv c_B^T B^{-1}$. Nótese que son los coeficientes que aparecen en la expresión de los costes reducidos del problema primal (denotados entonces por w^T) y aparecen, cambiados de signo, en la tabla del método simplex bajo las variables básicas iniciales x'_B. Luego, las variables duales son los coeficientes de las variables básicas iniciales en la función objetivo cambiados de signo.

Variables básicas	z	x'_N	x'_B	**Cotas**
$-z$	–1	$c'^T_N - c_B^T B^{-1} N'$	$-c_B^T B^{-1}$	$-c_B^T B^{-1} b$
x_B	0	$B^{-1} N'$	B^{-1}	$B^{-1} b$

Tabla 2.7 Tabla (B) del simplex en una iteración cualquiera.

A partir de esta expresión la comprobación de la propiedad enunciada es inmediata y se explican las relaciones de transformación de un problema a otro.

Propiedad de solución básica óptima complementaria

En la iteración final se encuentra simultáneamente la solución básica óptima $\hat{x}$ del primal y la básica óptima complementaria $\hat{y}$ del dual y se verifica $c^T \hat{x} = b^T \hat{y}$ o $\hat{z} = \hat{y}_0$.

Es decir, en el óptimo del problema primal la solución dual complementaria es también óptima para el problema dual.

Además, esta solución dual puede ser fácilmente identificada en la forma tabular del simplex, corresponde a los coeficientes de las variables básicas iniciales (variables de holgura y artificiales) en la función objetivo. Para las variables de holgura, la solución dual complementaria puede ser vista en los costes reducidos de estas variables, simplemente cambiando el signo. Si la variable fuera de exceso es todavía más directo, ya que no habría que cambiar el signo. Sólo en el caso de las restricciones donde no se hayan introducido variables de holgura o exceso, es decir, restricciones de igualdad (o si se hubieran eliminado las variables artificiales tras la fase I) no es posible ver el valor de las variables duales asociadas a esas restricciones.

Variables básicas	z	x'_N	x'_B	**Cotas**
$-z$	–1	–variables duales de exceso y de holgura	–variables duales originales	$-c_B^T B^{-1} b$
x_B	0	$B^{-1}N'$	B^{-1}	$B^{-1}b$

Propiedad de la complementariedad de holguras

Si una solución óptima del primal tiene una variable de holgura o exceso > 0 en una restricción (es decir, es variable básica > 0), la variable dual asociada a esa restricción tiene valor 0 y, de la misma forma, si una variable original del primal (no las de holgura, exceso o artificiales) en el óptimo tiene valor > 0, su correspondiente restricción del dual no tiene holgura (es decir, su variable asociada del dual es 0). Por el contrario, si una variable de holgura o exceso del primal es 0 en el óptimo, la variable dual asociada a dicha restricción puede tener un valor distinto de 0 (positivo o negativo según corresponda al tipo de restricción y de función objetivo –maximización o minimización–) en el óptimo del problema dual.

Supongamos, los problemas primal y dual ya mencionados

$$\begin{aligned} &\min z = c^T x \\ &Ax = b \qquad (P) \\ &x \geq 0 \end{aligned}$$

$$\begin{aligned} &\max y_0 = b^T y \\ &A^T y \le c \end{aligned} \quad (D)$$

Introducimos variables de holgura en las restricciones del problema dual

$$\begin{aligned} &\max y_0 = b^T y \\ &A^T y + s = c \\ &s \ge 0 \end{aligned}$$

Alternativamente, la propiedad de la complementariedad de holguras para las soluciones óptimas se puede expresar como

$$x_j s_j = 0 \quad j = 1, \ldots, n \tag{1.34}$$

Esta propiedad impide que los valores de x_j y de s_j en las soluciones óptimas de primal y dual sean simultáneamente diferentes de 0. Por otra parte, permite que ambos sean 0.

Sean dos puntos factibles del problema primal y dual, es decir, que cumplen

$$\begin{aligned} Ax = b \qquad & x \ge 0 \\ A^T y + s = c \quad & s \ge 0 \end{aligned} \tag{1.35}$$

Se define el intervalo de dualidad (*duality gap*) como la diferencia

$$c^T x - b^T y = (A^T y + s)^T x - b^T y = y^T A x + s^T x - b^T y = x^T s = \sum_{j=1}^{n} x_j s_j \tag{1.36}$$

Para las respectivas soluciones óptimas del problema primal y dual este intervalo de dualidad es 0 según la propiedad de la complementariedad de holguras.

Teorema de dualidad

Las únicas relaciones posibles entre primal y dual son:

1. Si un problema tiene soluciones óptimas factibles entonces también las tiene el otro. Se pueden aplicar las propiedades débil y fuerte de dualidad.
2. Si un problema tiene soluciones factibles y función objetivo no acotada, el otro problema no tiene soluciones factibles.

3. Si un problema no tiene soluciones factibles, el otro o no tiene soluciones factibles o no tiene función objetivo acotada.

Otra forma de expresar estas mismas relaciones entre primal y dual es el lema de Farkas. Sea el problema en su forma estándar

$$\begin{aligned} &\min c^T x \\ &Ax = b \\ &x \geq 0 \end{aligned} \tag{1.37}$$

Entonces exactamente uno de los dos sistemas de ecuaciones siguientes tiene solución.

Sistema 1: $A^T x = b$ y $x \geq 0$ para algún $x \in \mathbb{R}^n$

Sistema 2: $A^T y \leq 0$ y $b^T y > 0$ para algún $y \in \mathbb{R}^m$

Veamos estas propiedades en el caso ejemplo.

$$\begin{array}{llll} \max\ z = 3x_1 + 5x_2 & & & \\ x_1 & & \leq 4 & \\ & 2x_2 & \leq 12 & (P) \\ 3x_1 & +2x_2 & \leq 18 & \\ x_1, & x_2 & \geq 0 & \end{array}$$

El problema dual de éste es

$$\begin{array}{lllll} \min y_0 = 4y_1 + 12y_2 + 18y_3 & & & & \\ y_1 & & +3y_3 & \geq 3 & \\ & 2y_2 & +2y_3 & \geq 5 & (D) \\ y_1, & y_2, & y_3 & \geq 0 & \end{array}$$

Sea el problema primal anterior en su forma estándar y la tabla inicial correspondiente

$$
\begin{array}{llllll}
\min\ z = -3x_1 - 5x_2 & & & & & \\
x_1 & & +x_3 & & & = 4 \\
 & 2x_2 & & +x_4 & & = 12 \\
3x_1 & +2x_2 & & & +x_5 & = 18 \\
x_1, & x_2, & x_3, & x_4, & x_5 & \geq 0
\end{array}
$$

Variables básicas	z	x_1	x_2	x_3	x_4	x_5	**Cotas**
$-z$	–1	–3	–5				0
x_3		1		1			4
x_4			2		1		12
x_5		3	2			1	18

El problema dual en forma estándar, aunque sin añadir variables artificiales porque no lo queremos resolver por el método simplex tal como se ha explicado, y la tabla inicial correspondiente son

$$
\begin{array}{llllll}
\min y_0 = 4y_1 + 12y_2 + 18y_3 & & & & & \\
y_1 & & +3y_3 & -y_4 & & = 3 \\
 & 2y_2 & +2y_3 & & -y_5 & = 5 \\
y_1, & y_2, & y_3, & y_4, & y_5 & \geq 0
\end{array}
$$

Variables básicas	y_0	y_4	y_5	y_1	y_2	y_3	**Cotas**
$-y_0$	–1			4	12	18	0
$-y_4$		–1		1		3	3
$-y_5$			–1		2	2	5

La *solución básica factible* inicial para el problema primal es $(x_1, x_2, x_3, x_4, x_5) = (0, 0, 4, 12, 18)$, siendo $x_B^T = \begin{bmatrix} x_3 & x_4 & x_5 \end{bmatrix}$ y $x_N^T = \begin{bmatrix} x_1 & x_2 \end{bmatrix}$, y la función objetivo $\hat{z} = 0$. Para el problema dual la *solución básica infactible* correspondiente es $(y_1, y_2, y_3, y_4, y_5) = (0, 0, 0, -3, -5)$, siendo $y_B^T = \begin{bmatrix} y_4 & y_5 \end{bmatrix}$ y $y_N^T = \begin{bmatrix} y_1 & y_2 & y_3 \end{bmatrix}$, con función objetivo $\hat{y}_0 = 0$.

Los valores de las variables del problema dual (variables duales) corresponden a los coeficientes de las variables del problema primal en la función objetivo. Las variables originales del problema dual en las columnas correspondientes a las variables básicas iniciales del problema primal y las variables de exceso del problema dual en las columnas de las variables no básicas iniciales del primal. En cualquier iteración, excepto la última, del método simplex aplicado al problema primal algún coste reducido de las variables no básicas toma valor negativo luego existe una variable dual negativa, luego el problema dual para esa solución básica es infactible.

En el óptimo del problema primal se tienen la siguiente tabla

Variables básicas	z	x_1	x_2	x_3	x_4	x_5	**Cotas**
$-z$	–1				3/2	1	36
x_3				1	1/3	–1/3	2
x_2			1		1/2		6
x_1		1			–1/3	1/3	2

La solución básica óptima del primal es $(x_1, x_2, x_3, x_4, x_5) = (2, 6, 2, 0, 0)$ y $\hat{z} = -36$ para el de minimización, $\hat{z} = 36$ para el original de maximización. Ésta es la solución óptima porque los costes reducidos de las variables no básicas $\hat{c}_N^T = \begin{bmatrix} 3/2 & 1 \end{bmatrix}$ son no negativos. Las variables duales de este problema de minimización son $(y_1, y_2, y_3) = (0, -3/2, -1)$, los costes reducidos de las variables básicas originales cambiados de signo. Las variables duales del problema original de maximización son las opuestas $(y_1, y_2, y_3) = (0, 3/2, 1)$.

Para el problema dual la tabla final es la siguiente

Variables básicas	y_0	y_4	y_5	y_1	y_2	y_3	**Cotas**
$-y_0$	–1	2	6	2			–36
y_3		–1/3		1/3		1	1
y_2		1/3	–1/2	–1/3	1		3/2

La solución básica óptima del dual es $(y_1, y_2, y_3, y_4, y_5) = (0, 3/2, 1, 0, 0)$ e $\hat{y}_0 = 36$. Ésta es la solución óptima porque los costes reducidos de las variables no básicas $\hat{c}_N^T = \begin{bmatrix} 2 & 6 & 2 \end{bmatrix}$ son no negativos. Las variables duales de este problema son

$(x_1, x_2) = (2,6)$, corresponden a los costes reducidos de las variables de exceso y_4 e y_5 sin cambiar de signo.

En el óptimo del problema primal la primera restricción no es activa, por ser su variable de holgura x_3 estrictamente positiva, mientras que las dos últimas restricciones lo son por ser nulas las variables de holgura correspondientes, x_4 y x_5. Efectivamente, en el problema dual la variable y_1 asociada a la primera restricción toma valor 0, mientras que las variables y_2 y y_3 correspondientes a la segunda y tercera toman valor no nulo.

En cuanto a la interpretación de las variables duales, éstas tienen un claro sentido *económico* en el propio problema primal. Obsérvese que dadas las relaciones existentes entre ambos problemas se ha dicho que los valores óptimos de ambos problemas son iguales, es decir, suponiendo un problema primal de maximización se tiene que

$$\max c^T x = \min b^T y$$

o lo que es lo mismo,

$$\max \sum_{j=1}^{n} c_j x_j = \min \sum_{i=1}^{m} b_i y_i$$

Por lo tanto, si se incrementa una unidad la disponibilidad de un recurso b_i el óptimo de la función objetivo se verá incrementado en y_i unidades, es decir, las variables duales representan el valor marginal de cada recurso, también llamado *precio sombra* o *precio justo*, ya que representa la cantidad máxima que se podría llegar a pagar por una unidad más de recurso.

Las variables básicas son estrictamente positivas. Si alguna de las variables de holgura o exceso es básica entonces la variable dual de dicha restricción es 0 (propiedad de la complementariedad de holguras). Económicamente es evidente, ya que si hay holgura de un recurso su precio marginal es cero, no se está dispuesto a pagar por disponer de una unidad más.

Hay que tener precaución con esta interpretación pues todo tiene un límite, es decir, esto será cierto mientras no cambie el valor de las variables duales, lo que se puede asegurar mientras la base actual sea óptima (ya que como hemos visto las variables duales tienen la expresión $y^T = c_B^T B^{-1}$, luego no dependen de b). De ahí que una interpretación completa de una solución requiera también el análisis de los límites en que puede darse esta interpretación. El cálculo de estos límites corresponden a lo que se denomina análisis de sensibilidad que se presenta en la siguiente sección.

Veamos a continuación un ejemplo muy sencillo de planificación de la operación de sistemas de energía eléctrica escrito en GAMS. El problema primal es la minimización de los costes variables de explotación de un sistema para una hora, denominado problema de despacho económico. El problema dual corresponde a la maximización de la función de utilidad de consumidores y generadores.

Se trata de un sistema eléctrico con tres generadores cuya potencia máxima es de 200, 200 y 100 MW y con costes variables de 30, 50 y 70 €/MWh. La demanda a cubrir es de 300 MW.

$$
\begin{array}{llllll}
\min\limits_{P_1,P_2,P_3} c_1P_1 + c_2P_2 + c_3P_3 & & & & & \\
P_1 & +P_2 & +P_3 & \geq d & : \lambda \\
P_1 & & & \leq \overline{P}_1 & : \mu_1 \\
& P_2 & & \leq \overline{P}_2 & : \mu_2 \\
& & P_3 & \leq \overline{P}_3 & : \mu_3 \\
P_1, & P_2, & P_3 & \geq 0 &
\end{array}
\qquad
\begin{array}{llllll}
\max\limits_{\lambda,\mu_1,\mu_2,\mu_3} d\lambda + \overline{P}_1\mu_1 + \overline{P}_2\mu_2 + \overline{P}_3\mu_3 & & & & & \\
\lambda & +\mu_1 & & & \leq c_1 & : P_1 \\
\lambda & & +\mu_2 & & \leq c_2 & : P_2 \\
\lambda & & & +\mu_3 & \leq c_3 & : P_3 \\
\lambda \geq 0 & \mu_1, & \mu_2, & \mu_3 & \leq 0 &
\end{array}
$$

```
set i grupos de generación / g1 * g3 /

parameter
   cv(i)   coste variable  / g1  30, g2  50, g3  70 /
   ptmx(i) potencia máxima / g1 200, g2 200, g3 100 /

scalar dem demanda         / 300 /

positive variables P(i) potencia producida
variable coste

equations
fo         coste variable total
demanda    cobertura de la demanda
prodmax(i) limitación de la potencia producida ;

fo         .. coste =E= sum[i, cv(i) * P(i)] ;
demanda    .. sum[i, P(i)] =G= dem ;
prodmax(i) .. P(i) =L= ptmx(i) ;

model opera / all /

solve opera using lp minimizing coste
```

```
Solution Report     SOLVE opera Using LP From line 23


               S O L V E      S U M M A R Y

     MODEL   opera               OBJECTIVE  coste
     TYPE    LP                  DIRECTION  MINIMIZE
     SOLVER  CPLEX               FROM LINE  23

**** SOLVER STATUS     1 NORMAL COMPLETION
**** MODEL STATUS      1 OPTIMAL
**** OBJECTIVE VALUE            11000.0000

 RESOURCE USAGE, LIMIT          0.530      1000.000
 ITERATION COUNT, LIMIT         0         10000

GAMS/Cplex    May 15, 2003 WIN.CP.CP 21.0 023.025.041.VIS For Cplex 8.1
Cplex 8.1.0, GAMS Link 23

Optimal solution found.
Objective :       11000.000000


                           LOWER          LEVEL          UPPER         MARGINAL

---- EQU fo                  .              .              .             1.0000
---- EQU demanda          300.0000       300.0000        +INF           50.0000

  fo  coste variable total
  demanda  cobertura de la demanda

---- EQU prodmax  limitación de la potencia producida

          LOWER          LEVEL          UPPER         MARGINAL

g1        -INF          200.0000       200.0000       -20.0000
g2        -INF          100.0000       200.0000          .
g3        -INF            .            100.0000          .

---- VAR P  potencia producida

          LOWER          LEVEL          UPPER         MARGINAL

g1          .           200.0000        +INF            .
g2          .           100.0000        +INF            .
g3          .             .             +INF          20.0000

                           LOWER          LEVEL          UPPER         MARGINAL

---- VAR coste             -INF        11000.0000        +INF             .
```

```
set i grupos de generación / g1 * g3 /

parameter
   cv(i)   coste variable  / g1  30, g2  50, g3  70 /
   ptmx(i) potencia máxima / g1 200, g2 200, g3 100 /

scalar dem demanda         / 300 /

positive variable  LA    precio marginal
negative variables MU(i) precio marginal
variable benef

equations
fo         remuneración total
prgrupo(i) limitación del pago al grupo ;
```

```
fo         .. benef =E= dem * LA + sum[i, ptmx(i) * MU(i)] ;
prgrupo(i) .. LA + MU(i) =L= cv(i) ;

model opera / all /

solve opera using lp maximizing benef
```

```
Solution Report     SOLVE opera Using LP From line 22


               S O L V E      S U M M A R Y

     MODEL   opera               OBJECTIVE  benef
     TYPE    LP                  DIRECTION  MAXIMIZE
     SOLVER  CPLEX               FROM LINE  22

**** SOLVER STATUS     1 NORMAL COMPLETION
**** MODEL STATUS      1 OPTIMAL
**** OBJECTIVE VALUE            11000.0000

 RESOURCE USAGE, LIMIT          0.030      1000.000
 ITERATION COUNT, LIMIT         2         10000

GAMS/Cplex    May 15, 2003 WIN.CP.CP 21.0 023.025.041.VIS For Cplex 8.1
Cplex 8.1.0, GAMS Link 23

Optimal solution found.
Objective :       11000.000000


                           LOWER          LEVEL          UPPER         MARGINAL

---- EQU fo                  .              .              .             1.0000

  fo  remuneración total

---- EQU prgrupo  limitación del pago al grupo

          LOWER          LEVEL          UPPER         MARGINAL

g1        -INF           30.0000        30.0000       200.0000
g2        -INF           50.0000        50.0000       100.0000
g3        -INF           50.0000        70.0000         .
                           LOWER          LEVEL          UPPER         MARGINAL

---- VAR LA                  .            50.0000        +INF            .

  LA  precio marginal

---- VAR MU  precio marginal

          LOWER          LEVEL          UPPER         MARGINAL

g1        -INF          -20.0000          .             .
g2        -INF             .              .           100.0000
g3        -INF             .              .           100.0000

                           LOWER          LEVEL          UPPER         MARGINAL

---- VAR benef             -INF       11000.0000        +INF            .
```

El coste variable total resultante es de 11000 €. El grupo 1 produce sus 200 MW y el grupo 2 produce 100 MW y es el marginal. Para este problema, la variable dual de la ecuación de la demanda indica el incremento en costes variables del sistema por incremento marginal de la demanda. Tiene un valor de 50 €/MWh, lo que quiere decir que si la demanda se incrementara en 1 MW el coste variable total se incrementaría en 50 €. Corresponde al coste variable del grupo marginal.

Una aplicación directa de la dualidad se plantea a la hora de resolver un problema ya que, como se ha visto, resolviendo uno u otro se obtiene la misma solución óptima. Entonces, dado un problema primal cualquiera y sabiendo que se puede obtener su dual y resolverlo, ¿cuál de los dos se debe resolver? Sabiendo que el tiempo de resolución crece proporcionalmente a m^3, si $m \ll n$ entonces se resolverá el problema primal y si $m \gg n$ se resolverá el problema dual en estas etapas: (i) se pasa del primal al dual, (ii) se resuelve el dual y (iii) se pasa de nuevo del dual al primal.

I.6. Análisis de sensibilidad

El análisis de sensibilidad o de *postoptimalidad* estudia los efectos sobre la solución óptima debidos a cambios en cualquier parámetro del problema de optimización lineal A, b o c partiendo de la solución óptima ya alcanzada.

Observando la tabla 2.8, tabla del simplex en cualquier iteración, se advierte que las columnas de la tabla bajo las variables básicas iniciales x'_B no dependen de los parámetros del problema original. Entonces, si éstos cambian ($b \to \overline{b}$, $A \to \overline{A}$ o $c'_N \to \overline{c}'_N$) se deberán recalcular para los nuevos valores solamente los bloques de la tabla que se alteran y continuar a partir de ahí con el método simplex.

Variables Básicas	z	x'_N	x'_B	**Valores**
$-z$	–1	$\overline{c}'^T_N - c^T_B B^{-1} \overline{N}'$	$-c^T_B B^{-1}$	$-c^T_B B^{-1} \overline{b}$
x_B	0	$B^{-1} \overline{N}'$	B^{-1}	$B^{-1} \overline{b}$

Tabla 2.8 Tabla del simplex (B) en una iteración cualquiera con cambios en parámetros.

También se pueden determinar los cambios de manera incremental. Si ($b \to b + \Delta b$, $A \to A + \Delta A$ y $c'_N \to c'_N + \Delta c'_N$) la tabla presenta ahora este aspecto

Variables Básicas	z	x'_N	x'_B	Valores
$-z$		$\Delta c_N'^T - c_B^T B^{-1} \Delta N'$		$-c_B^T B^{-1} \Delta b$
x_B		$B^{-1} \Delta N'$		$B^{-1} \Delta b$

Tabla 2.9 Cambios incrementales en la tabla del simplex (B) en una iteración cualquiera con cambios en parámetros.

Luego las modificaciones de los bloques correspondientes se pueden escribir como $\Delta \hat{c}_N^T = \Delta c_N'^T - c_B^T B^{-1} \Delta N'$, $\Delta \hat{z} = w^T \Delta b$, $\Delta \hat{b} = B^{-1} \Delta b$

El procedimiento genérico completo de análisis de sensibilidad tiene estas etapas:

1) introducción de los cambios en los bloques correspondientes de la tabla del simples

2) preparación para adecuarla a una iteración del simplex (función objetivo con coeficientes nulos en variables básicas y matriz identidad en las columnas de dichas variables)

3) prueba de factibilidad (todas las variables básicas no negativas)

4) prueba de optimalidad (costes reducidos de las variables no básicas no negativos en un problema de minimización)

5) nueva iteración del método simplex o del método simplex dual (éste se ve en un siguiente apartado)

En el ejemplo habitual se ha hecho mediante GAMS análisis de sensibilidad con respecto a las cotas de las restricciones y a los coeficientes de la función objetivo y da estos resultados.

$$
\begin{array}{lll}
\max\ z = 3x_1 + 5x_2 & & \\
x_1 & & \leq 4 \\
 & 2x_2 & \leq 12 \\
3x_1 & +2x_2 & \leq 18 \\
x_1, & x_2 & \geq 0
\end{array}
$$

```
positive variables x1, x2
```

```
variable z

equations fo, r1, r2, r3 ;

fo .. z =e= 3 * x1 + 5 * x2 ;
r1 .. x1 =l= 4 ;
r2 .. 2 * x2 =l= 12 ;
r3 .. 3 * x1 + 2 * x2 =l= 18 ;

model caso / all / ;

caso.optfile= 1

solve caso maximazing z using lp
```

```
Solution Report     SOLVE caso Using LP From line 15


               S O L V E      S U M M A R Y

     MODEL   caso                OBJECTIVE  z
     TYPE    LP                  DIRECTION  MAXIMIZE
     SOLVER  CPLEX               FROM LINE  15

**** SOLVER STATUS     1 NORMAL COMPLETION
**** MODEL STATUS      1 OPTIMAL
**** OBJECTIVE VALUE               36.0000

 RESOURCE USAGE, LIMIT          0.010      1000.000
 ITERATION COUNT, LIMIT         0         10000

GAMS/Cplex    May 15, 2003 WIN.CP.CP 21.0 023.025.041.VIS For Cplex 8.1
Cplex 8.1.0, GAMS Link 23


User supplied options:
objrng all
rhsrng all

Optimal solution found.
Objective :          36.000000

EQUATION NAME                                LOWER      CURRENT        UPPER
-------------                                -----      -------        -----
fo                                            -INF            0         +INF
r1                                               2            4         +INF
r2                                               6           12           18
r3                                              12           18           24


VARIABLE NAME                                LOWER      CURRENT        UPPER
-------------                                -----      -------        -----
x1                                              -3            0          4.5
x2                                              -3            0         +INF
z                                                0            1         +INF

                       LOWER     LEVEL     UPPER    MARGINAL

---- EQU fo              .         .         .        1.000
---- EQU r1             -INF      2.000     4.000      .
---- EQU r2             -INF     12.000    12.000     1.500
---- EQU r3             -INF     18.000    18.000     1.000

                       LOWER     LEVEL     UPPER    MARGINAL

---- VAR x1              .        2.000     +INF       .
---- VAR x2              .        6.000     +INF       .
---- VAR z              -INF     36.000     +INF       .
```

A continuación se analizan específicamente algunos cambios: en las cotas de las restricciones, en un coeficiente de una variable en las restricciones o en la función objetivo, introducción de una nueva variable o restricción.

I.6.1. Cambios en cotas de restricciones

Observando la tabla 2.8 se ve que los cambios en las cotas de las restricciones sólo afectan al valor de la función objetivo y a los valores de las variables básicas. Como éstos se pueden hacer negativos se puede perder factibilidad. En este caso, se aplica el método simplex dual que se verá más adelante.

Geométricamente, los cambios en las cotas de las restricciones suponen desplazar paralelas a sí mismas las restricciones. Dado que el punto óptimo viene dado por las restricciones que son activas, al realizar este desplazamiento también se desplaza el punto óptimo. El nuevo punto puede seguir dentro de la región factible, en cuyo caso sigue siendo el óptimo o quedar fuera, con lo que hay que aplicar el método denominado simplex dual para volver a la región factible, es decir, recuperar la factibilidad.

Veamos un ejemplo de cambio de la cota en la segunda restricción del problema. Las cotas pasan de ser $b = \begin{bmatrix} 4 \\ 12 \\ 18 \end{bmatrix}$ a tomar el valor $\overline{b} = \begin{bmatrix} 4 \\ 24 \\ 18 \end{bmatrix}$. La inversa de la matriz base en la iteración final era $B^{-1} = \begin{bmatrix} 1 & 1/3 & -1/3 \\ \cdot & 1/2 & \cdot \\ \cdot & -1/3 & 1/3 \end{bmatrix}$ y las variables duales $y^T = c_B^T B^{-1} = \begin{bmatrix} \cdot & -3/2 & -1 \end{bmatrix}$.

El nuevo valor de la función objetivo será $\hat{z} = c_B^T B^{-1}\overline{b} = y^T\overline{b} = \begin{bmatrix} \cdot & -3/2 & -1 \end{bmatrix} \begin{bmatrix} 4 \\ 24 \\ 18 \end{bmatrix} = -54$ y el de las variables básicas $x_B = \hat{b} = B^{-1}\overline{b} = \begin{bmatrix} 1 & 1/3 & -1/3 \\ \cdot & 1/2 & \cdot \\ \cdot & -1/3 & 1/3 \end{bmatrix} \begin{bmatrix} 4 \\ 24 \\ 18 \end{bmatrix} = \begin{bmatrix} 6 \\ 12 \\ -2 \end{bmatrix}$, que resulta ser infactible por tomar un valor negativo.

Estos cambios se pueden realizar también de forma incremental

$$\hat{z} = \hat{z} + \Delta\hat{z} = \hat{z} + y^T\Delta b = -36 + \begin{bmatrix} \cdot & -3/2 & -1 \end{bmatrix} \begin{bmatrix} \cdot \\ 12 \\ \cdot \end{bmatrix} = -36 - 18 = -54$$

$$x_B = x_B + \Delta x_B = x_B + B^{-1}\Delta b = \begin{bmatrix} 2 \\ 6 \\ 2 \end{bmatrix} + \begin{bmatrix} 1 & 1/3 & -1/3 \\ \cdot & 1/2 & \cdot \\ \cdot & -1/3 & 1/3 \end{bmatrix} \begin{bmatrix} \cdot \\ 12 \\ \cdot \end{bmatrix} = \begin{bmatrix} 2 \\ 6 \\ 2 \end{bmatrix} + \begin{bmatrix} 4 \\ 6 \\ -4 \end{bmatrix} = \begin{bmatrix} 6 \\ 12 \\ -2 \end{bmatrix}$$

Es interesante determinar también el intervalo de valores para los que la solución básica óptima se mantiene factible cambiando únicamente el valor de b_i en el problema.

Continuando con el ejemplo de incremento de la cota de la segunda restricción $\hat{b} = \hat{b} + B^{-1}\Delta b = \begin{bmatrix} 2 \\ 6 \\ 2 \end{bmatrix} + \begin{bmatrix} 1/3 \\ 1/2 \\ -1/3 \end{bmatrix} \Delta b_2 \geq 0$. Por tanto, $\begin{Bmatrix} 2 + \Delta b_2/3 \geq 0 \\ 6 + \Delta b_2/2 \geq 0 \\ 2 - \Delta b_2/3 \geq 0 \end{Bmatrix}$. Es decir, $\Delta b_2 \geq -6$, $\Delta b_2 \geq -12$ y $\Delta b_2 \leq 6$. Luego, $-6 \leq \Delta b_2 \leq 6$ da el intervalo de modificación de la cota de la segunda restricción sin que se pierda factibilidad con la misma solución básica factible.

I.6.2. Cambio en un coeficiente de una variable no básica

Se analiza el cambio en los coeficientes de una variable no básica en la iteración final o bien en la función objetivo $\hat{c}_N^T = \overline{c}_N'^T - c_B^T B^{-1} \overline{N}'$ o bien en las restricciones $B^{-1}\overline{N}'$. Estos cambios no afectan a la factibilidad de la solución pero pueden afectar a la optimalidad. Geométricamente, el cambio de un coeficiente de una variable en la función objetivo, sea básica o no, supone un cambio de dirección de optimización, pero no afecta a la región factible. Así pues, lo único que puede ocurrir es que el punto óptimo actual deje de serlo y haya que iterar para encontrar el nuevo óptimo.

Si $\hat{c}_N^T = \overline{c}_N'^T - c_B^T B^{-1} \overline{N}' \geq 0$ la solución sigue siendo óptima. Se puede determinar el intervalo de valores del coeficiente de la variable no básica en la función objetivo (cambiando solamente éste) para los que se mantiene la optimalidad. En caso contrario la solución no es óptima, ésta será la nueva variable básica entrante y se debe continuar con el método simplex.

De acuerdo con esta expresión el *coste reducido* se puede interpretar como la cantidad mínima en que tiene que reducirse el coeficiente de la variable en la función objetivo para que interese realizar esa actividad (incrementar desde 0), o bien como el máximo decremento permisible de dicho coeficiente para mantener la actual solución básica factible como óptima.

I.6.3. Introducción de una nueva variable

Es un caso particular de la situación anterior. Una variable que no existe es equivalente a una variable no básica en la solución final (igual a 0) a la que se le cambian los coeficientes de la función objetivo y de las restricciones.

Para introducir una nueva variable se calcula el coste reducido de la variable y si éste es menor que 0 será una variable cuyo valor interese aumentar y por lo tanto hay que introducirla en la tabla. Para ello se calcula $B^{-1}a_j$ y se añade como columna de esa variable en la tabla y su coste reducido en la fila de la función objetivo. A continuación se sigue con las iteraciones del simplex.

I.6.4. Cambio en un coeficiente de una variable básica

El cambio en un coeficiente de una variable básica en la iteración final (aunque era una variable no básica en la iteración inicial) en la función objetivo supone un cambio en todos los costes reducidos de las variables no básicas que han de ser actualizados. Una vez hecho esto, la solución actual puede seguir siendo óptima (con diferente valor de la función objetivo) o no, en cuyo caso hay que seguir iterando con el método simplex (pérdida de optimalidad).

El cambio de algún coeficiente de una variable básica en la matriz de restricciones $B^{-1}\overline{N}'$ altera la tabla del simplex de manera que se requiere ponerla de nuevo en la forma adecuada para eliminación gaussiana. Esto puede causar pérdida de factibilidad (resuelto mediante método simplex dual) y/o de optimalidad (resuelto mediante método simplex), e incluso las variables básicas actuales pueden dejar de formar una base, requiriendo la introducción de una variable artificial que la sustituya en la base.

I.6.5. Introducción de nueva restricción

Al introducir una nueva restricción se comprueba si ésta se satisface para la solución óptima. Si lo hace, la solución sigue siendo óptima. Si no, se introduce la restricción en la tabla final del simplex con la variable de holgura, por exceso o artificial, que se toma como variable básica. Se prepara la tabla para eliminación gaussiana y se sigue con el método simplex dual ya que se ha perdido factibilidad.

Es importante apuntar, que dado que la solución óptima actual no verifica la restricción añadida, al introducirla en la tabla y adaptar ésta para la eliminación gaussiana, necesariamente el valor de la variable básica en esa restricción ha de ser negativo, pues en caso contrario, es que la solución actual sí verificaba la restricción.

I.7. Método simplex dual

El método simplex dual es conveniente utilizarlo cuando:

1) Se necesitan introducir muchas variables artificiales para construir una solución básica factible inicial

2) Se hace infactible la solución óptima tras una perturbación en análisis de sensibilidad

Opera en el problema primal como si el método simplex se aplicara simultáneamente al problema dual. Maneja soluciones básicas que satisfacen la condición de optimalidad del primal (la solución dual complementaria es factible), aunque no sean factibles para el primal, y se mueve hacia la solución óptima para conseguir factibilidad del primal, sin perder nunca la factibilidad dual (optimalidad primal). Es decir, mantiene los coeficientes de la ecuación de la función objetivo mayores o iguales que cero y los de las cotas de las restricciones con alguno estrictamente negativo, excepto en el óptimo en que logra la factibilidad primal.

Este método, por lo tanto, se mueve por fuera de la región factible, moviéndose por vértices exteriores (infactibles) con mejor valor de la función objetivo que el óptimo que luego se alcanza y entra en la región factible sólo al alcanzar el óptimo.

A continuación se presenta el procedimiento del método simplex dual

1) Inicialización

Se parte de una solución básica que cumpla el criterio de optimalidad (dual factible), es decir, cuyos costes reducidos sean todos mayores o iguales que 0.

Si la solución es factible (todas las variables básicas tienen valor ≥ 0), la solución es óptima. En caso contrario, iterar.

2) Iteración

a. Determinar la *variable básica saliente* (variable básica con valor negativo) que determina la *fila pivote* s.

Se suele elegir dentro de las negativas la mayor en valor absoluto, aunque esto no garantiza una convergencia más rápida.

b. Determinar la *variable básica entrante*

Se ha de utilizar un criterio de relación semejante al del método simplex primal para asegurar que se mantiene la factibilidad dual. De entre las variables que tienen coeficientes < 0 en la fila pivote se selecciona aquélla que minimice en valor absoluto el cociente entre el coste reducido y su coeficiente en la fila pivote s :

$$\min_{1\le j\le n}\left\{\left|\frac{c_j - z_j}{y_{sj}}\right| : y_{sj} < 0\right\}$$

En el caso de que no haya ningún elemento estrictamente menor que cero en la fila pivote, quiere decir que el problema dual no está acotado y, por lo tanto, el problema primal es infactible.

c. Determinar la nueva solución básica por eliminación gaussiana.

Veamos con un ejemplo el método simplex dual. Sea el siguiente problema de programación lineal

$$\begin{aligned} \min 4x_1 + x_2 \\ 3x_1 + x_2 \ge 3 \\ 4x_1 + 3x_2 \ge 6 \\ x_1 + 2x_2 \ge 4 \\ x_1, x_2 \ge 0 \end{aligned}$$

Como las restricciones son de tipo ≥ habría que introducir muchas variables artificiales y emplear el método de las dos fases para solucionarlo. En su lugar, se transforman las restricciones en tipo ≤ y se aplica el método simplex dual.

$$\begin{aligned} \min 4x_1 + x_2 \\ -3x_1 - x_2 \le -3 \\ -4x_1 - 3x_2 \le -6 \\ -x_1 - 2x_2 \le -4 \\ x_1, x_2 \ge 0 \end{aligned}$$

Una vez transformado en la forma estándar se obtiene el siguiente problema y la siguiente tabla del simplex

$$\begin{aligned} \min 4x_1 + x_2 \\ -3x_1 - x_2 + x_3 = -3 \\ -4x_1 - 3x_2 + x_4 = -6 \\ -x_1 - 2x_2 + x_5 = -4 \\ x_1, x_2, x_3, x_4, x_5 \geq 0 \end{aligned}$$

Variables Básicas	z	x_1	x_2	x_3	x_4	x_5	**Valores**
Relación		–1	–1/3				
$-z$	–1	4	1				
x_3		–3	–1	1			–3
x_4		–4	–3		1		–6
x_5		–1	–2			1	–4

Obsérvese cómo la tabla tiene las características necesarias para utilizar el simplex dual: coeficientes de la función objetivo mayores o iguales que 0 (factibilidad del dual, optimalidad del primal) y alguna cota de las restricciones negativa (no optimalidad del dual, infactibilidad del primal).

Se elige x_4 como variable básica saliente por tener el mayor coeficiente negativo en la función objetivo del dual (mayor valor negativo de cota). Ésta define la fila pivote. Se calcula la relación entre los coeficientes de la función objetivo y los estrictamente negativos de la fila pivote. Se elige como variable básica entrante x_2, aquélla con menor relación en valor absoluto, será la primera que alcance valor 0 al incrementar x_4. Se realizan ahora las transformaciones sobre la tabla para prepararla para la eliminación gaussiana.

Variables Básicas	z	x_1	x_2	x_3	x_4	x_5	**Valores**
Relación		–8/5			–1		
$-z$	–1	8/3			1/3		–2
x_3		–5/3		1	–1/3		–1
x_2		4/3	1		–1/3		2
x_5		5/3			–2/3	1	

Se elige x_3 como variable básica saliente y x_4 como variable básica entrante y se hacen las transformaciones pertinentes.

Variables Básicas		x_1	x_2	x_3	x_4	x_5	**Valores**
$-z$	–1	1		1			–3
x_4		5		–3	1		3
x_2		3	1	–1			3
x_5		5		–2		1	2

Ésta es la solución óptima $(x_1, x_2, x_3, x_4, x_5) = (0,3,0,3,2)$ y la función objetivo $\hat{z} = 3$. Las variables duales son $(y_1, y_2, y_3) = (-1,0,0)$ para las restricciones de $\leq$, para las restricciones originales $\geq$ serán $(y_1, y_2, y_3) = (1,0,0)$.

Veamos el problema dual del problema original y la tabla inicial del simplex para este problema

$$\begin{aligned} \max y_0 &= 3y_1 + 6y_2 + 4y_3 \\ 3y_1 + 4y_2 + y_3 &\leq 4 \\ y_1 + 3y_2 + 2y_3 &\leq 1 \\ y_1, y_2, y_3 &\geq 0 \end{aligned}$$

Variables Básicas	y_0	y_1	y_2	y_3	y_4	y_5	**Valores**	**Relación**
$-y_0$	–1	–3	–6	–4				
y_4		3	4	1	1		4	1
y_5		1	3	2		1	1	1/3

Entra la variable y_2 y sale la variable y_5. Reorganizando la matriz de restricciones queda la siguiente tabla.

Variables Básicas	y_0	y_1	y_2	y_3	y_4	y_5	**Valores**	**Relación**
$-y_0$	–1	–1				2	2	
y_4		5/3		–5/3	1	–4/3	8/3	8/5
y_2		1/3	1	2/3		1/3	1/3	1

Entra la variable y_1 y sale la variable y_2. Reorganizando la matriz de restricciones queda la siguiente tabla, que corresponde a la solución óptima.

Variables Básicas	y_0	y_1	y_2	y_3	y_4	y_5	**Valores**
$-y_0$	–1		3	2		3	3
y_4			–5	–5	1	–3	1
y_1		1	3	2		1	1

La solución óptima es $(y_1, y_2, y_3, y_4, y_5) = (1,0,0,1,0)$ y función objetivo $\hat{y}_0 = 3$ para el problema de maximización. Las variables duales son $(x_1, x_2) = (0,3)$. Es decir, el método simplex dual es equivalente a aplicar el método simplex al problema dual.

I.8. Programación lineal paramétrica

Así como el análisis de sensibilidad determinaba los cambios en la solución óptima del problema original por cambios discretos en los valores de algún parámetro, el análisis paramétrico determina el cambio en la función objetivo en función de cambios continuos en los parámetros. Se presenta a continuación el análisis paramétrico frente a cambios en los coeficientes de la función objetivo y cotas de las restricciones.

I.8.1. Cambios simultáneos en coeficientes de la función objetivo

Suponemos la función objetivo $z = \sum_{j=1}^{n} c_j x_j$. Se desea conocer el cambio en la función objetivo óptima $\hat{z}(\theta) = \sum_{j=1}^{n} \left(c_j + \alpha_j \theta\right) x_j$ al cambiar el parámetro θ desde 0, supuestos conocidos los pesos α_j para cada variable x_j.

1) Resolver hasta optimalidad el problema con $\theta = 0$ por el método simplex.

2) Mediante análisis de sensibilidad introducir $\Delta c_j = \alpha_j \theta$.

3) Modificar la ecuación de la función objetivo para hacer 0 los coeficientes de las variables básicas en la función objetivo. Aplicar la condición de optimalidad (coeficientes de variables no básicas $\geq$ 0) para obtener el valor de θ.

4) Incrementar θ hasta anular un coeficiente de variable no básica. Ésta es la nueva variable básica entrante.

5) Encontrar la nueva solución óptima mediante el algoritmo simplex primal. Ir al paso 3.

Obsérvese que en este tipo de análisis la región factible no cambia, sólo se va modificando la dirección de optimización y para los distintos valores del parámetro θ se va detectando un nuevo punto óptimo. El proceso de análisis lleva a definir zonas del espacio paramétrico, siendo el óptimo en cada una de ellas un punto extremo diferente.

I.8.2. Cambios simultáneos en cotas de las restricciones

Se desea conocer el cambio en la función objetivo óptima $\hat{z}$ al cambiar θ desde 0 en las cotas de las restricciones $b_i + \alpha_i \theta$, supuestos conocidos los pesos α_i. Los cambios en las cotas b_i son equivalentes a cambios en coeficientes de la función objetivo del problema dual.

1) Resolver hasta optimalidad el problema con $\theta = 0$ por el método simplex.

2) Mediante análisis de sensibilidad introducir $\Delta b_i = \alpha_i \theta$.

3) Expresar la condición de optimalidad de las variables del dual (cotas de las restricciones $\geq$ 0).

4) Incrementar θ hasta anular las cotas de las restricciones (la solución se haría infactible). Ésta es la nueva variable básica saliente en el simplex dual.

5) Encontrar la nueva solución óptima mediante el algoritmo simplex dual. Ir al paso 2.

Obsérvese en este caso, que los diferentes valores del parámetro θ determinan diferentes regiones factibles, así el espacio paramétrico es dividido en intervalos de modo que en cada uno de ellos hay una base que es óptima (unas restricciones activas que determinan el punto extremo óptimo). En este caso, el proceso de análisis acaba con un intervalo donde para cualquier valor del parámetro hay una base que es óptima o con un intervalo del parámetro donde el problema resulta no factible.

I.9. Método de punto interior primal-dual

La investigación en métodos de punto interior aplicados a programación lineal nació a consecuencia del teórico mal comportamiento (exponencial) del método simplex en el peor caso (cuando ha de recorrer todos los vértices del contorno de la región factible).

El método simplex requiere un número de iteraciones proporcional al número de ecuaciones. Sin embargo, el número de iteraciones puede aumentar sustancialmente en problemas degenerados o disminuir drásticamente si se dispone de una base inicial. Cada iteración del simplex es poco costosa. Esencialmente es el cálculo de una actualización de la inversa de la matriz base.

Los métodos de punto interior requieren pocas iteraciones para resolver un problema lineal. Sin embargo, estas iteraciones son muy costosas. Esencialmente cada iteración conlleva la resolución de un sistema de ecuaciones lineales, como se verá más adelante.

En 1979 Khachiyan presentó el *método del elipsoide* y en 1984 Karmarkar describió un *método proyectivo*, reclamando ambos un requerimiento de tiempo polinómico. De hecho, los métodos barrera fueron desarrollados originalmente para programación no lineal en los años 60 por Fiacco y McCormick [Fiacco, 1968].

El *método primal-dual*, que se presenta a continuación, es un caso particular de ellos y ha dado muy buenos resultados en la práctica. Se denomina así porque resuelve simultáneamente el problema primal y el dual.

Supongamos un problema lineal en su forma estándar

$$\begin{aligned} \min z &= c^T x \\ Ax &= b \quad (P) \\ x &\geq 0 \end{aligned} \tag{1.38}$$

y su correspondiente problema dual

$$\left.\begin{aligned} \max y_0 &= b^T y \\ A^T y &\leq c \end{aligned}\right. \quad (D) \tag{1.39}$$

siendo $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$, $x \in \mathbb{R}^n$ e $y \in \mathbb{R}^m$.

Añadiendo variables de holgura en las restricciones del dual tenemos

$$\begin{aligned} \max y_0 &= b^T y \\ A^T y + s &= c \quad (D) \\ s &\geq 0 \end{aligned} \tag{1.40}$$

siendo $s \in \mathbb{R}^m$.

Sea $\overline{x}$ una solución factible del problema primal (P) y $(\overline{y}, \overline{s})$ una solución factible del problema dual (D), no necesariamente básicas. Serán óptimos en sus respectivos problemas si cumplen la propiedad de la complementariedad de holguras

$$x_j s_j = 0 \quad j = 1, \dots, n \tag{1.41}$$

La idea principal del método primal-dual es moverse por una secuencia de soluciones estrictamente factibles[6] en ambos problemas tratando de que se verifiquen las condiciones de complementariedad de holguras. Específicamente, se trata de encontrar $x(\mu)$, $y(\mu)$ y $s(\mu)$ para $\mu > 0$ que satisfagan el siguiente sistema de $(m+n)$ ecuaciones lineales (denominadas condiciones de factibilidad) y n ecuaciones no lineales (denominadas condiciones de complementariedad)

$$\begin{aligned} & Ax = b \\ & x \geq 0 \\ & A^T y + s = c \qquad (PD) \\ & s \geq 0 \\ & x_j s_j = \mu \qquad \forall j \end{aligned} \tag{1.42}$$

El valor del parámetro μ se va reduciendo hasta lograr la convergencia. Mientras $\mu > 0$ la condición $x_j s_j = \mu$ implica que $x > 0$ y $s > 0$, es decir, que ambos son puntos estrictamente factibles. Este mismo avance se puede ver como una reducción del intervalo de dualidad (*duality gap*) hasta alcanzar el valor cero en la solución óptima

$$c^T x - b^T y = x^T s = n\mu \tag{1.43}$$

El método primal-dual se puede interpretar como una variante del método de Newton de programación no lineal que resuelve iterativamente un sistema de ecuaciones no lineales donde las direcciones de movimiento corresponden a Δx^k, Δy^k y Δs^k. Para ello se pasa de una solución estrictamente factible $x^k > 0$ y $s^k > 0$ en la

---

[6] Se define una solución como estrictamente factible en el poliedro $Ax = b$, $x \geq 0$ a aquélla que verifica $Ax = b$ y $x > 0$.

iteración k a otra en la iteración $k+1$ que mantenga estricta factibilidad (de ahí el nombre de método de punto interior) $x^{k+1} > 0$ y $s^{k+1} > 0$ y disminuya el intervalo de dualidad $x^{k+1} \leftarrow x^k + \Delta x^k$, $y^{k+1} \leftarrow y^k + \Delta y^k$ y $s^{k+1} \leftarrow s^k + \Delta s^k$

$$\begin{aligned} A\left(x^k + \Delta x^k\right) = b \quad &\rightarrow \quad A\Delta x^k = 0 \\ A^T\left(y^k + \Delta y^k\right) + s^k + \Delta s^k = c \quad &\rightarrow \quad A^T \Delta y^k + \Delta s^k = 0 \end{aligned} \tag{1.44}$$

que es un sistema de $(m+n)$ ecuaciones lineales en función de Δx^k, Δy^k e Δs^k.

Además se deben satisfacer las condiciones de complementariedad de holguras para cada dimensión j

$$\begin{aligned} x_j^{k+1} s_j^{k+1} = (x_j^k + \Delta x_j^k)(s_j^k + \Delta s_j^k) = \mu^{k+1} \\ s_j^k \Delta x_j^k + x_j^k \Delta s_j^k + \Delta x_j^k \Delta s_j^k = \mu^{k+1} - x_j^k s_j^k \end{aligned} \tag{1.45}$$

que es un sistema de n ecuaciones no lineales en función de Δx_j^k y Δs_j^k. Si despreciamos el término $\Delta x_j^k \Delta s_j^k$, ya que se supone que los incrementos de las variables en cada iteración serán pequeños, se obtiene un sistema de ecuaciones lineales

$$s_j^k \Delta x_j^k + x_j^k \Delta s_j^k = \mu^{k+1} - x_j^k s_j^k \tag{1.46}$$

Representemos la condición de complementariedad de holguras en forma matricial en función de $X = \operatorname{diag}(x)$, matriz diagonal donde x_j ocupa la posición *j*, $S = \operatorname{diag}(s)$ y $e = \begin{pmatrix} 1 & \cdots & 1 \end{pmatrix}^T$. Luego $x = Xe$, $s = Se$ y, por consiguiente,

$$XSe = \mu e \tag{1.47}$$

De forma completa, una iteración cualquiera del método primal-dual resuelve este sistema de $(m+2n)$ ecuaciones lineales

$$\begin{aligned} S\Delta x + X\Delta s = \mu e - XSe \\ A\Delta x = 0 \\ A^T \Delta y + \Delta s = 0 \end{aligned} \tag{1.48}$$

Obteniendo Δs en función de Δy de la tercera ecuación y sustituyendo en la primera

$$\begin{gathered} S\Delta x - XA^T\Delta y = \mu e - XSe \\ A\Delta x = 0 \end{gathered} \tag{1.49}$$

Premultiplicando por AS^{-1} y aprovechando la segunda ecuación obtenemos este sistema de m ecuaciones lineales para determinar Δy

$$-AS^{-1}XA^T\Delta y = AS^{-1}(\mu e - XSe) \tag{1.50}$$

Si definimos la matriz diagonal $D \equiv S^{-1}X$ y $v(\mu) = \mu e - XSe$

$$-ADA^T\Delta y = AS^{-1}v(\mu) \tag{1.51}$$

y los otros vectores solución se pueden calcular secuencialmente de esta manera

$$\begin{aligned} \Delta s &= -A^T\Delta y \\ \Delta x &= S^{-1}v(\mu) - D\Delta s \end{aligned} \tag{1.52}$$

El coste computacional más importante de este método está en la resolución del sistema de ecuaciones lineales para el cálculo de la dirección de movimiento Δy. En el caso de columnas densas de A la matriz ADA^T tiene muchos más elementos no nulos que la misma matriz A.

Si en el método simplex tanto el manejo de matrices cuasivacías como la factorización de la inversa son fundamentales para una buena implantación en problemas de gran tamaño estas técnicas resultan críticas en el método de punto interior.

El algoritmo completo parte de una solución estrictamente factible en la iteración k para $x^k > 0$, y^k y $s^k > 0$. Se calculan las direcciones de movimiento Δx^k, Δy^k e Δs^k que definen las nuevas soluciones $x^k + \Delta x^k$, $y^k + \Delta y^k$ y $s^k + \Delta s^k$. Entonces se reduce el valor del parámetro $\mu^{k+1} = \sigma^k \mu^k$, siendo $\sigma^k \in [0,1]$ el parámetro de disminución para la siguiente iteración $k+1$ y se repite el proceso. El parámetro μ se puede disminuir rápidamente (hasta un orden de magnitud en cada iteración) siempre que se mantenga estricta factibilidad de las soluciones y se garantice una disminución de la holgura. Para garantizar la estricta factibilidad hay que limitar los movimientos de las variables de forma que éstas permanezcan dentro de sus cotas

$$\begin{aligned} x(\alpha_P, \mu) &= x + \alpha_P \Delta x \\ y(\alpha_D, \mu) &= y + \alpha_D \Delta y \\ s(\alpha_D, \mu) &= s + \alpha_D \Delta s \end{aligned} \tag{1.53}$$

donde $\alpha_P = \min_{\Delta x_j < 0} \left(-x_j / \Delta x_j \right)$ y $\alpha_D = \min_{\Delta s_j < 0} \left(-s_j / \Delta s_j \right)$ son las longitudes del paso que aseguran $x > 0$ y $s > 0$. Obsérvese que el parámetro α_D es único para y^k y s^k. Una estrategia habitual es tomar una longitud de paso única $\alpha = \min(\alpha_P, \alpha_D)$

$$\begin{aligned} x_j + \alpha \Delta x_j &\geq 0 \\ y_j + \alpha \Delta y_j &\geq 0 \\ s_j + \alpha \Delta s_j &\geq 0 \end{aligned} \tag{1.54}$$

Si además de asegurar estricta factibilidad en la selección de α_P y α_D se satisface la condición de centralidad

$$(x^k, s^k) \in C^k \tag{1.55}$$

siendo $C^k = \left\{ (x,s) : x_j s_j \geq \gamma \mu^k, j = 1, \dots, n \right\}$ para $\gamma \in (0,1)$ estamos en la estrategia *path following* del método primal-dual.

La selección heurística de valores adecuados de la longitud de paso α, parámetro de centrado γ e intervalo de dualidad μ en cada iteración son imprescindibles para obtener en la práctica un comportamiento robusto del algoritmo.

El número de iteraciones es polinomial de orden n o $n^{3/2}$ con el tamaño del problema.

La obtención de un punto inicial estrictamente factible para el problema primal y dual puede ser difícil. Sin embargo, las transformaciones anteriores se pueden modificar para incluir puntos iniciales no factibles.

$$\begin{aligned} A\left(x^k + \Delta x^k\right) = b \qquad &\rightarrow \qquad A\Delta x^k = b - Ax^k \equiv r_P \\ A^T\left(y^k + \Delta y^k\right) + s^k + \Delta s^k = c \qquad &\rightarrow \qquad A^T\Delta y^k + \Delta s^k = c - (A^T y^k + s^k) \equiv r_D \end{aligned} \tag{1.56}$$

o bien

$$\begin{gathered} S\Delta x + X\Delta s = \mu e - XSe \equiv v(\mu) \\ A\Delta x = b - Ax \equiv r_P \\ A^T\Delta y + \Delta s = c - A^T y - s \equiv r_D \end{gathered} \tag{1.57}$$

siendo r_P y r_D precisamente los residuos de las ecuaciones del primal y dual respectivamente. Entonces, el cálculo de los vectores de actualización en el método primal-dual para puntos iniciales infactibles queda de esta manera

$$\begin{aligned} &\Delta y = -(ADA^T)^{-1}\left[AS^{-1}v(\mu) - ADr_D - r_P\right] \\ &\Delta s = -A^T\Delta y + r_D \\ &\Delta x = S^{-1}v(\mu) - D\Delta s \end{aligned} \tag{1.58}$$

En el caso de puntos infactibles se debe dar prioridad en la disminución de la infactibilidad sobre el decremento del intervalo de dualidad μ.

El método primal-dual *predictivo-correctivo* de Mehrotra está diseñado para no ignorar los términos de segundo orden $\Delta x_j \Delta s_j$ que anteriormente se habían despreciado. Esto se realiza en dos etapas. En la *etapa predictiva* o *afín* se predicen (calculan) los valores de Δx y de Δs junto con un valor del parámetro μ relacionado con los valores de x y de s según el sistema de ecuaciones anterior. En la *etapa correctiva* con paso de Newton (i.e., segunda derivada) se utilizan estos valores para resolver el siguiente sistema de ecuaciones lineales

$$\begin{gathered} S\Delta x + X\Delta s = \mu e - XSe - \Delta X\Delta Se \equiv v(\mu) \\ A\Delta x = b - Ax \equiv r_P \\ A^T\Delta y + \Delta s = c - A^T y - s \equiv r_D \end{gathered} \tag{1.59}$$

donde $\Delta X = \operatorname{diag}(\Delta x)$ e $\Delta S = \operatorname{diag}(\Delta s)$ son los valores obtenidos en la etapa predictiva.

Los métodos de punto interior no producen soluciones básicas óptimas. Por ello cuando llegan al óptimo se suele realizar un proceso de permutación (*crossover*) para determinar la solución básica óptima.

Veamos a continuación la resolución de este problema por el método primal-dual.

$$\begin{array}{lcl}
\min\ z = x_1 + x_2 & & \min\ z = x_1 + x_2 \\
-2x_1 + x_2 \le 2 & & -2x_1 + x_2 + x_3 = 2 \\
-x_1 + 2x_2 \le 6 & & -x_1 + 2x_2 + x_4 = 6 \\
x_1 + 2x_2 \le 10 & \text{o expresado en forma estándar} & x_1 + 2x_2 + x_5 = 10 \\
x_1 + 2x_2 \ge 5 & & -x_1 - 2x_2 + x_6 = -5 \\
x_1 - x_2 \le 4 & & x_1 - x_2 + x_7 = 4 \\
x_1, x_2 \ge 0 & & x_1, x_2, x_3, x_4, x_5, x_6, x_7 \ge 0
\end{array}$$

Su problema dual será el siguiente

$$\begin{array}{l}
\max\ y_0 = 2y_1 + 6y_2 + 10y_3 - 5y_4 + 4y_5 \\
-2y_1 - y_2 + y_3 - y_4 + y_5 \le 1 \\
y_1 + 2y_2 + 2y_3 - 2y_4 - y_5 \le 1 \\
y_1 \le 0 \\
y_2 \le 0 \\
y_3 \le 0 \\
y_4 \le 0 \\
y_5 \le 0
\end{array}$$

que añadiendo variables de holgura

$$\begin{array}{l}
\max\ y_0 = 2y_1 + 6y_2 + 10y_3 - 5y_4 + 4y_5 \\
-2y_1 - y_2 + y_3 - y_4 + y_5 + s_1 = 1 \\
y_1 + 2y_2 + 2y_3 - 2y_4 - y_5 + s_2 = 1 \\
y_1 + s_3 = 0 \\
y_2 + s_4 = 0 \\
y_3 + s_5 = 0 \\
y_4 + s_6 = 0 \\
y_5 + s_7 = 0 \\
s_1, s_2, s_3, s_4, s_5, s_6, s_7 \ge 0
\end{array}$$

Luego el sistema de 12 ecuaciones lineales y 7 no lineales a resolver es el siguiente

$$
\begin{aligned}
&-2x_1 + x_2 + x_3 = 2 \\
&-x_1 + 2x_2 + x_4 = 6 \\
&x_1 + 2x_2 + x_5 = 10 \\
&-x_1 - 2x_2 + x_6 = -5 \\
&x_1 - x_2 + x_7 = 4 \\
&-2y_1 - y_2 + y_3 - y_4 + y_5 + s_1 = 1 \\
&y_1 + 2y_2 + 2y_3 - 2y_4 - y_5 + s_2 = 1 \\
&y_1 + s_3 = 0 \\
&y_2 + s_4 = 0 \\
&y_3 + s_5 = 0 \\
&y_4 + s_6 = 0 \\
&y_5 + s_7 = 0 \\
&x_1 s_1 = 0 \\
&x_2 s_2 = 0 \\
&x_3 s_3 = 0 \\
&x_4 y_4 = 0 \\
&x_5 s_5 = 0 \\
&x_6 s_6 = 0 \\
&x_7 s_7 = 0 \\
&x_1, x_2, x_3, x_4, x_5, x_6, x_7 \geq 0 \\
&s_1, s_2, s_3, s_4, s_5, s_6, s_7 \geq 0
\end{aligned}
$$

Lo convertimos en un problema de minimización con restricciones de igualdad

$$
\begin{array}{ccc}
\max\ z = 3x_1 + 5x_2 & & \min -3x_1 - 5x_2 \\
\begin{array}{rrl} x_1 & & \leq 4 \\ & 2x_2 & \leq 12 \\ 3x_1 & +2x_2 & \leq 18 \\ & x_1, x_2 & \geq 0 \end{array} & \text{o bien} & \begin{array}{rrrrrl} x_1 & & +x_3 & & & = 4 \\ & 2x_2 & & +x_4 & & = 12 \\ 3x_1 & +2x_2 & & & +x_5 & = 18 \\ & & & & x_1, x_2, x_3, x_4, x_5 & \geq 0 \end{array}
\end{array}
$$

Su problema dual será el siguiente

$$\begin{array}{lllll} \max 4y_1 & +12y_2 & +18y_3 & \\ y_1 & & +3y_3 & \leq -3 \\ & 2y_2 & +2y_3 & \leq -5 \\ y_1 & & & \leq 0 \\ & y_2 & & \leq 0 \\ & & y_3 & \leq 0 \end{array} \quad \text{o bien} \quad \begin{array}{lllllllll} & & & & & & & \max 4y_1 +12y_2 +18y_3 \\ y_1 & & +3y_3 & +s_1 & & & & & = -3 \\ & 2y_2 & +2y_3 & & +s_2 & & & & = -5 \\ y_1 & & & & & +s_3 & & & = 0 \\ & y_2 & & & & & +s_4 & & = 0 \\ & & y_3 & & & & & +s_5 & = 0 \\ & & & & & & & s_1, s_2, s_3, s_4, s_5 & \geq 0 \end{array}$$

I.10. Referencias

Bazaraa, M.S. and Jarvis, J.J. (1998) *Programación Lineal y Flujo en Redes.* Limusa.

Bertsimas. D. and Tsitsiklis. J.N. (1997) *Introduction to Linear Optimization*. Athena Scientific.

Chvátal. V. (1983) *Linear Programming*. W.H. Freeman and Co.

Dantzig, G. (1949) Programming of Independent Activities II, mathematical model. *Econometrica* (17) pp. 200-211.

Dantzig, G. and Thapa, M. (1997) *Linear Programming. Introduction.* Springer.

Fiacco, A. V. and McCormick, G. P. (1968) *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*. John Wiley and Sons.

Gill, P.E., Murray, W. and Wright, M.H. (1991) *Numerical Linear Algebra and Optimization*. Addison Wesley.

Nash. S.G. and Sofer. A. (1996) *Linear and Nonlinear Programming*. McGraw-Hill.

Ríos Insúa, S. (1996) *Investigación operativa. Programación lineal y aplicaciones* Ed. Centro de estudios Ramón Areces.

Ríos Insúa, S., Ríos Insúa, D., Mateos, A. y Martín, J. (1997) *Programación lineal y aplicaciones. Ejercicios resueltos* Ed. Ra-Ma.

Rivier, M. (1998) *Modelo probabilista de explotación de un sistema eléctrico: contribución a la teoría marginalista*. Tesis doctoral. Universidad Pontificia Comillas.

Vanderbei. R.J. (1996) *Linear Programming: Foundations and Extensions.* Kluwer Academic Pusblishers.

I.11. Biblioteca de problemas

PROBLEMA 1

Dado el problema de programación lineal:

$$\begin{aligned}&\min x_1 + x_2 - 3x_3\\&3x_1 - x_3 = 5\\&x_2 - x_3 = 1\\&x_1, x_2, x_3 \geq 0\end{aligned}$$

Se pide:

1) Dar todas las soluciones básicas factibles.

2) Comparar la función objetivo para todas las soluciones básicas factibles y decir cuál es la mejor de todas ellas.

3) ¿Es el vector $x = (1335, 4001, 4000)$ una solución para el problema anterior? ¿Es mejor solución que la obtenida en el apartado anterior?

4) Comprobar que la base $B = \begin{pmatrix} 3 & \cdot \\ \cdot & 1 \end{pmatrix}$ está asociada a una solución no acotada.

PROBLEMA 2

Resolver los siguientes problemas de programación lineal, utilizando el método de penalizaciones y el método de las dos fases, cuando sea necesario.

1. Discutir el problema geométricamente

$$\begin{aligned}
&\max 3x + 2y \\
&4x + y \leq 16 \\
&x + 4y \leq 16 \\
&5x + 6y \leq 30 \\
&x, y \geq 0
\end{aligned}$$

2. Discutir el problema geométricamente

$$\begin{aligned}
&\max 3y + 8z \\
&y + 2z = 4 \\
&z \geq 3 \\
&y, z \geq 0
\end{aligned}$$

3. Discutir el problema geométricamente

$$\begin{aligned}
&\max 3x + 2y \\
&5x + y \geq 0 \\
&y \geq x \geq 0
\end{aligned}$$

4.

$$\begin{aligned}
&\min 3x + 5y - 4z + 6t \\
&x + 2z - t = 6 \\
&y + 4z + t = 9 \\
&2z - 2t = 3 \\
&x, y, z, t \geq 0
\end{aligned}$$

5.

$$\begin{aligned}
&\min - x_1 + 2x_2 - 3x_3 \\
&x_1 + x_2 + x_3 = 6 \\
&-x_1 + x_2 + 2x_3 = 4 \\
&2x_2 + 3x_3 = 10 \\
&x_3 \leq 2 \\
&x_1, x_2, x_3 \geq 0
\end{aligned}$$

PROBLEMA 3

Dados los siguientes problemas de programación lineal, plantear los correspondientes problemas duales:

1.

$$\begin{aligned}
&\min 3x_1 + 2x_2 - x_3 \\
&x_1 + x_2 - x_3 \geq 5 \\
&-x_1 - x_2 + 7x_3 \geq 4 \\
&x_1, x_2 \geq 0 \\
&x_3 \text{ libre}
\end{aligned}$$

2.

$$\begin{aligned}
&\max 3x_1 + 4x_2 \\
&x_1 - x_2 \geq 7 \\
&x_1 + x_2 \geq 3 \\
&x_1, x_2 \text{ libres}
\end{aligned}$$

3.

$$\begin{aligned}
&\max x_1 + x_2 + x_3 \\
&x_1 - x_2 \geq 5 \\
&x_1 + x_3 = 8 \\
&x_1, x_2 \geq 0
\end{aligned}$$

PROBLEMA 4

Dados los siguientes problemas

$$\text{(i)} \quad \begin{aligned}
&\min x_1 + x_2 + x_3 + x_4 \\
&2x_1 - x_2 + x_3 = 4 \\
&x_1 + 2x_3 - x_4 = 3 \\
&x_1, x_2, x_3 \geq 0
\end{aligned}$$

$$\text{(ii)} \quad \begin{aligned} &\max x_1 + 2x_2 + 3x_3 \\ &x_1 - x_2 + 2x_3 \geq 5 \\ &x_1 - x_3 = 3 \\ &x_1, x_2, x_3 \geq 0 \end{aligned}$$

Se pide:

1) Resolverlos mediante el algoritmo primal.

2) Plantear el dual y resolverlo.

3) Comprobar la relación existente entre ambas soluciones.

PROBLEMA 5

Dado el problema de programación lineal

$$\begin{aligned} &\min x_1 + x_2 + x_3 + x_4 \\ &2x_1 - x_2 + x_3 = 4 \\ &x_1 + 2x_3 - x_4 = 3 \\ &x_1, x_2, x_3 \geq 0 \end{aligned}$$

Se pide: (cada apartado es independiente de los demás)

1) La solución óptima del problema si se añade la restricción $x_1 + x_2 - x_3 = 2$

2) Idem si se añade la restricción $x_1 + x_2 - x_3 = 3$

3) Idem al cambiar el vector $b = \begin{pmatrix} 4 \\ 3 \end{pmatrix}$ por $\overline{b} = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$

4) Añadir x_5 con coste $c_5 = -3$ y $a_5 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$

PROBLEMA 6

Dado el problema

$$\begin{aligned}
&\min c_1x_1 + c_2x_2 + c_3x_3\\
&x_1 + 2x_2 - x_3 \le 5\\
&x_1 + 2x_3 \le 8\\
&x_1, x_2, x_3 \ge 0
\end{aligned}$$

donde

$$\begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} 3 \\ -2 \\ -1 \end{pmatrix} + \begin{pmatrix} -3 \\ 2 \\ 1 \end{pmatrix}\theta \qquad \theta \in [0, \infty]$$

PROBLEMA 7

Dado el problema de programación lineal

$$\begin{aligned}
&\min 3x_1 + 2x_2 + 4x_3\\
&3x_1 + 2x_2 - x_3 \le 5\\
&x_1 + x_3 = 3\\
&x_1, x_2, x_3 \ge 0
\end{aligned}$$

Se pide:

1) La solución óptima.

2) Resolver el problema paramétrico asociado si el término independiente es $b^T(\theta) = (5,3)^T + (4,1)^T\theta$, $\theta \in [0, \infty)$

PROBLEMA 8

Un ama de casa, típico ejemplo de la economía sumergida, hace en sus ratos domésticos libres dos clases de salsa de tomate que vende en jarras de barro al supermercado de al lado. La primera requiere utilizar 3 kg de tomates y 4 tazas de vinagre por jarra de salsa. La segunda requiere 5 kg de tomates y 2 tazas de vinagre. La primera salsa le produce un beneficio neto de 0.4 € por jarra y la segunda 0.5.

El supermercado que evacúa su producción casera hacia los circuitos comerciales (no sabemos con qué beneficio relativo) le impone a nuestra amiga las siguientes condiciones:

- Que produzca como mínimo 3 jarras a la semana.
- Que le compre como máximo 24 kg de tomate y 3 botellas de vinagre a la semana.

Sabiendo que una botella de vinagre equivale a 16 tazas y que el supermercado monopoliza la venta de tomate y vinagre en la región, analizar los precios a los que estaría dispuesta a pagar el tomate y el vinagre nuestra ama de casa a un primo suyo contrabandista, que se los puede proporcionar recurriendo a otro de los variados mecanismos de la economía subterránea.

PROBLEMA 9

Dado el siguiente problema de programación lineal

$$\begin{aligned}
&\max 2x_1 - x_2 - 17x_3 \\
&5x_1 - 2x_2 + 31x_3 \leq -1 \\
&3x_1 - x_2 + 19x_3 \leq 0 \\
&x_1, x_2, x_3 \geq 0
\end{aligned}$$

Se pide:

1) Resolver por el algoritmo dual.
2) Resolver por el algoritmo primal.
3) Obtener, usando postoptimización, la solución para $\overline{b_1} = 1$.

PROBLEMA 10

Dado el siguiente problema de programación lineal:

$$\begin{aligned}
&\max 2x_1 + x_2 + 0.5x_3 \\
&2x_1 + x_2 + 2x_3 \leq 2 \\
&x_1 - x_2 - x_3 \geq 1 \\
&x_1, x_2, x_3 \geq 0
\end{aligned}$$

1) Resolver por el algoritmo primal.
2) Identificar razonadamente, en la tabla óptima anterior, el valor de las variables duales.

3) Verificar las condiciones de holgura complementaria.

4) Hacer una iteración más del algoritmo primal, ¿qué ocurre?

PROBLEMA 11

Dado el siguiente problema y su correspondiente tabla óptima

$$\max 4x + 2y$$
$$x + y \leq 5$$
$$2x + y \leq 7$$
$$x \leq 3$$
$$x, y \geq 0$$

	x	y	s_1	s_2	s_3	
$-z$	0	0	0	–2	0	–14
s_1	0	0	1	–1	1	1
y	0	1	0	1	–2	1
x	1	0	0	0	1	3

Se pide:

1) Resolverlo gráficamente.

2) ¿La solución es única? Si no es así, dar todas las posibles soluciones.

3) Plantear el problema dual.

4) Dar los precios sombra de los recursos.

5) Comprobar las relaciones existentes entre las soluciones primal y dual.

6) Obtener la nueva solución óptima, sin resolver todo el problema de nuevo, si se modifica la primera componente del vector b, cambiando de $b_1 = 5$ a $\overline{b_1} = 4$.

7) Obtener la solución óptima, sin resolver todo el problema de nuevo, si se añade la restricción $5x+3y\leq 15$.

8) Si se parametriza el vector de coeficientes de la función objetivo en la dirección (1,2), es decir, $c=(4,2)+(1,2)\theta$, $\theta\geq 0$, ¿para qué valores del parámetro la solución actual sigue siendo óptima?

PROBLEMA 12

Dado el siguiente problema lineal

$$\begin{aligned}&\max z=2x_1+2x_2\\&x_1+x_2\leq 3\\&x_1+2x_2\leq 4\\&x_1,x_2\geq 0\end{aligned}$$

Se pide:

1) Resolverlo gráficamente.

2) Resolverlo aplicando el método simplex. Razonar la aplicación y el resultado.

3) ¿Qué valores del coeficiente de x_2 en la función objetivo mantienen la base óptima?

4) ¿Cuál es la solución si el lado derecho de la segunda restricción pasa a ser 10? Resolverlo mediante postoptimización.

5) Obtener los precios duales de los dos recursos. ¿Cuál es su significado? ¿cuál es el sentido de su signo?

6) Determinar, mediante postoptimización, la solución óptima si añadimos al problema original la restricción $2x_1+x_2\leq 2$.

PROBLEMA 13

En una fábrica se producen dos tipos de artículos. La semana siguiente se estima que han de producirse al menos 50 artículos en total. Para producirlos se usan dos tipos de materia prima, cuyo coste unitario es de 100 y 200 €, respectivamente. El problema que se plantea para maximizar los beneficios de la empresa y su tabla óptima son los siguientes:

$$\begin{aligned}
&\max 200x + 400y \\
&x + y \geq 50 \\
&x + 2y \leq 80 \\
&x + y \leq 60 \\
&x, y \geq 0
\end{aligned}$$

	x	y	s_1	s_2	s_3	
$-z$	0	0	0	–200	0	–16000
x	1	0	–2	–1	0	20
y	0	1	1	1	0	30
s_3	0	0	1	0	1	10

1) Resolver el problema gráficamente.
2) Hallar todas las posibles soluciones analíticamente.
3) Interpretar todos los elementos del problema y las soluciones.
4) Dar los precios sombra e interpretarlos.
5) Hallar mediante postoptimización la nueva solución si se dispone de 120 unidades de la primera materia en lugar de 80.
6) Estudiar mediante postoptimización qué ocurre si la producción del segundo producto se limita a 10 unidades.

PROBLEMA 14

Dado el siguiente problema de programación lineal:

$$
\begin{aligned}
&\max 2x + y \\
&x + y \geq 2 \\
&x + 4y \leq 16 \\
&3x + y \leq 15 \\
&x, y \geq 0
\end{aligned}
$$

1) Resolver el problema mediante el algoritmo del simplex.

2) Obtener la nueva solución mediante postoptimización al añadir la restricción $x + y \leq 4$.

PROBLEMA 15

En una fábrica se producen dos productos, P1 y P2, cuyos beneficios unitarios son 20 y 70, respectivamente. En la producción se utilizan dos materias primas, disponiendo de 30 y 60 unidades cada una. La demanda total es de al menos 20 productos y ha de ser satisfecha. El planteamiento del problema es:

$$
\begin{aligned}
&\max 20x + 70y \\
&x + 3y \leq 30 \\
&2x + 4y \leq 60 \\
&x + y \geq 20 \\
&x, y \geq 0
\end{aligned}
$$

1) Resolver mediante el algoritmo del simplex.

2) Dar los valores e interpretación de las variables originales, las de holgura, las artificiales (si existen), las duales, de los costes reducidos y de la función objetivo, si la tabla óptima fuera

		s			
		– 25		5	– 650
		1 /2		/2	5
		–			1

		1			0	
		– 1/2		/2	5	1

3) Surge ahora en el mercado la posibilidad de fabricar un tercer producto, P3. Este producto requiere 1 unidad de la primera materia prima (de la que se dispone de 30) y 1 unidad de la segunda, siendo su beneficio unitario de 60. La demanda total sería de 20 unidades igualmente. Sin embargo, este producto requiere del alquiler de un tipo de maquinaria cuyo coste es de 100 unidades monetarias, independientemente del tiempo que sea usada y de la cantidad que se produzca. Además, si se produce alguna cantidad de este producto P3 han de producirse al menos 8 unidades del producto P2. Plantear un problema de programación lineal entera para maximizar el beneficio de la empresa.

PROBLEMA 16

Dado el siguiente problema de programación lineal

$$
\begin{aligned}
&\min 3x_1 + 5x_2 + 7x_3 \\
&x_1 + x_2 + x_3 \geq 300 \\
&x_1 \leq 200 \\
&x_2 \leq 200 \\
&x_3 \leq 100 \\
&x_1, x_2, x_3 \geq 0
\end{aligned}
$$

Se pide:

1) Resolver por el algoritmo primal.

2) Resolver el problema dual por el algoritmo primal.

PROBLEMA 17

Dado el siguiente problema de programación lineal

$$\begin{aligned}
&\min 4x_1 + x_2 \\
&3x_1 + x_2 \geq 3 \\
&4x_1 + 3x_2 \geq 6 \\
&x_1 + 2x_2 \geq 4 \\
&x_1, x_2 \geq 0
\end{aligned}$$

Se pide:

1) Resolver por el algoritmo dual.

2) Resolver el problema dual por el algoritmo primal.

PROBLEMA 18

Convertir el siguiente problema a uno de programación lineal en forma estándar

$$\begin{aligned}
&\min |x| + |y| + |z| \\
&x + y \leq 1 \\
&2x + z = 3
\end{aligned}$$

PROBLEMA 19

Un cierto tipo de funciones poligonales se puede representar como

$$f(x) = \max\left\{c_1^T x + d_1, c_2^T x + d_2, \ldots, c_p^T x + d_p\right\}$$

donde $x, c_1, c_2, \ldots, c_p \in \mathbb{R}^n$ y $d_1, d_2, \ldots, d_p \in \mathbb{R}$

Para tal función se considera el problema

$$\begin{aligned}
&\min f(x) \\
&Ax = b \\
&x = 0
\end{aligned}$$

Mostrar cómo se puede convertir este problema en uno de programación lineal.

PROBLEMA 20

Un fabricante desea producir 90 toneladas de una aleación de modo que el porcentaje del metal A esté entre el 60 y el 70%. En el mercado hay cinco aleaciones, cuya composición y precio se indican a continuación, a partir de las cuales el metalúrgico desea conseguir al precio más barato la aleación deseada.

Aleación	Al1	Al2	Al3	Al4	Al5
%A	10	25	50	75	90
%otros	90	75	50	25	10
Precio/t (€)	25	40	65	55	30

A tal efecto, el metalúrgico modela la situación con el siguiente problema de programación lineal

$$\begin{aligned}
&\min 25x_1 + 40x_2 + 65x_3 + 55x_4 + 30x_5 \\
&x_1 + x_2 + x_3 + x_4 + x_5 = 90 \\
&0.1x_1 + 0.25x_2 + 0.5x_3 + 0.75x_4 + 0.9x_5 \le 0.7(x_1 + x_2 + x_3 + x_4 + x_5) \\
&0.1x_1 + 0.25x_2 + 0.5x_3 + 0.75x_4 + 0.9x_5 \ge 0.6(x_1 + x_2 + x_3 + x_4 + x_5) \\
&x_1, x_2, x_3, x_4, x_5 \ge 0
\end{aligned}$$

que tras agrupar términos en cada restricción, añadir dos variables de holgura (s_2 y s_3 en la segunda y tercera restricción, respectivamente) y resolver (añadiendo una variable artificial en la primera restricción, a_1), proporciona la siguiente tabla óptima

	x_1	x_2	x_3	x_4	x_5	s_2	s_3	a_1	Cotas
$c_j - z_j$	0	14.0625	37.5	25.9375	0	0	6.25	–28.125	–2531.25
x_5	0	0.1875	0.5	0.8125	1	0	–1.25	0.625	56.25
s_2	0	0	0	0	0	1	1	0.1	9
x_1	1	0.8125	0.5	0.1875	0	0	1.25	0.375	33.75

A la vista de los resultados, se pide (cada apartado es independiente de los demás):

1) Determinar mediante postoptimización las cantidades óptimas y el nuevo valor de la función objetivo si se tuvieran que producir 100 toneladas. A la vista de lo obtenido ¿qué conclusión puedes sacar?

2) Si el metalúrgico pudiera comprar una tonelada de aleación en el mercado con los porcentajes deseados del metal A para satisfacer su demanda sin tener que producirlo él, ¿a qué precio máximo la pagaría? ¿por qué? ¿hasta qué cantidad?

3) ¿Qué precios máximos tendrían que tener las aleaciones no utilizadas para que interesara incluirlas en la aleación final?

4) ¿Qué ocurriría si en el mercado aparece una nueva aleación con un 40 % del metal A y a un precio por tonelada de 30 €?

5) ¿Qué ocurriría si en el mercado sólo hay disponible 30 toneladas de la aleación Al1?

PROBLEMA 21

Un artesano fabrica trenes y camiones de juguete muy sencillos a base de tornillos, bloques de plástico y ruedas. Para la próxima semana dispone de 8000, 6000 y 6300 unidades de las citadas componentes. La adquisición de estas componentes y los gastos semanales de operación y amortización del local y utillaje ascienden a 40 €.

La estructura de producto de trenes y camiones se refleja en la siguiente tabla y los beneficios unitarios que reportan son de 1.6 y 1.4 € respectivamente.

	Tornillos	Bloques	Ruedas
Tren	10	15	18
Camión	20	10	6

El objetivo del artesano es el de determinar qué gama de producción le generará la mayor cantidad de ingresos esta semana. Para ello modela la situación con el siguiente modelo lineal

$$\begin{aligned}
&\max z = 1.6x_1 + 1.4x_2 + 0x_3 + 0x_4 + 0x_5\\
&10x_1 + 20x_2 + x_3 = 8000\\
&15x_1 + 10x_2 + x_4 = 6000\\
&18x_1 + 6x_2 + x_5 = 6300\\
&x_1, x_2, x_3, x_4, x_5 \geq 0
\end{aligned}$$

del que se adjunta la última tabla del método simplex de la que se deduce la solución óptima:

	x_1	x_2	x_3	x_4	x_5	Cotas
$c_j - z_j$	0	0	–0.025	–0.09	0	–740
x_5	0	0	0.45	–1.5	1	900
x_2	0	1	0.075	–0.05	0	300
x_1	1	0	–0.05	0.1	0	200

Situémonos sucesivamente e independientemente en los siguientes supuestos y tomemos las decisiones pertinentes utilizando, cuando corresponda, el análisis de postoptimalidad (o sensibilidad):

1) El beneficio unitario de los trenes baja a 1 € y el de camiones aumenta a 2.2 €.

2) Por otro lado existe la posibilidad de adquirir 200 tornillos con un precio unitario incrementado sobre el precio normal de adquisición 0.05 €. ¿Merece la pena adquirirlos? ¿Se adquirirían los 200 tornillos si el citado incremento fuera tan sólo de 0.02 €? ¿Por qué?

3) Parece que una de las razones por las que los juguetes han perdido atractivo es que son un poco pobres de ornamentación. Cabe la posibilidad de añadirles algunas ventanillas como adornos: los trenes necesitan 7 ventanillas por unidad y los camiones 4. El artesano puede adquirir para la próxima semana un máximo de 2100 ventanillas. ¿Cómo afecta esta incidencia a la producción óptima de juguetes?

4) Otro enterado en estos asuntos le sugiere la posibilidad de ampliar la gama de juguetes con la fabricación de aviones. Un avión necesita 8 tornillos, 12 bloques y 3 ruedas y proporciona un beneficio unitario de 1.1 €. ¿Interesa fabricar el nuevo juguete?

PROBLEMA 22

El ingeniero del ICAI Antón Pirulero Artesano, conocido por APA entre sus compañeros de promoción, se dedica al ejercicio libre de la profesión, orientándose a actividades comerciales. En la línea que le ha hecho famoso como empresario-riesgo, se dispone a participar en la feria agro-alimentaria de Vladivostock presentando una colección de tres tipos de piruletas de su invención. Estas piruletas son: A: piruletas de naranja, B: piruletas de limón, C: piruletas de anís. Las piruletas tienen una base común de azúcar y miel, en las cantidades que se indican en la siguiente tabla. Con la venta de las piruletas tiene previsto cubrir en parte los gastos de participación en la feria. Por razones de limitación en el peso autorizado, a la feria sólo puede llevar 220 kg de azúcar y 90 kg de miel y quiere vender al menos 20000 piruletas.

Tipo de piruleta	Azúcar (g/unidad)	Miel (g/unidad)	Beneficio Unitario (€)
A	10	4	0.09
B	9	5	0.12
C	12	6	0.15

La solución óptima de su problema la ha obtenido resolviendo el siguiente modelo lineal

$$\max 0.09A + 0.12B + 0.15C$$
$$10A + 9B + 12C \leq 220000$$
$$4A + 5B + 6C \leq 90000$$
$$A + B + C \geq 20000$$
$$A, B, C \geq 0$$

donde *A*, *B* y *C* son las cantidades respectivas de piruletas de los tipos A, B y C que puede fabricar con los kilos de azúcar y miel que puede trasladar, y cuya última tabla es

Variables básicas	*A*	*B*	*C*	*Holgura* 1	*Holgura* 2	*Holgura* 3	Cotas
c_j–z_j	0	0	0	0	–0.03	–0.03	–2100
Holgura 1	0	–2	0	1	–1	6	10000
C	0	0.5	1	0	0.5	2	5000
A	1	0.5	0	0	–0.5	–3	15000

1) ¿Qué valores y qué interpretación tienen las variables duales del problema?

2) ¿Hay alguna otra solución económicamente equivalente a la obtenida? ¿Por qué? ¿Cuál es esa solución? ¿Cuántas soluciones tiene el problema? ¿Cuál es su estructura general?

3) Antón Pirulensko Artesanov, delegado comercial de APA en Rusia, tras un estudio de mercado entre la población que acaba de abandonar el biberón, le aconseja que lleve al menos 12000 piruletas de limón. ¿Cómo afecta esto a los planes de APA?

4) Nuestro amigo APA, además de empresario emprendedor, patrocina actividades de I+D y así, con el apoyo del IIT y de los profesores de MM, ha obtenido un nuevo y original tipo de piruleta, la pirulica-I, el “no va más” de las piruletas (*the remilk on the bicycle*, como la calificó la famosa revista inglesa de gastronomía “The big pork”), cuya base sigue siendo el azúcar y la miel (12 y 8 gramos respectivamente por unidad) aunque el resto de los ingredientes es un secreto industrial celosamente guardado por todos los que hemos intervenido en su diseño. Si el beneficio unitario de una pirulica-I es de 0.22 €, ¿interesa presentarla y venderla en feria, en orden a un mayor beneficio?

PROBLEMA 23

Se considera el programa lineal

$$\begin{aligned}&\max 10x_1 + 7x_2 + 12x_3\\&x_1 + x_2 + x_3 \le 9\\&3x_1 + 2x_2 + 4x_3 \le 24\\&x_i \ge 0\end{aligned}$$

1) Obtener todos los vértices del poliedro de soluciones factibles, comprobando que el óptimo se obtiene para $x_1 = 6,\ x_2 = 3,\ x_3 = 0$.

2) ¿Qué ocurre si se añade la nueva restricción $x_1 - 2x_2 \ge 1$? Razonar y obtener el resultado a partir de la tabla óptima del simplex del problema original.

PROBLEMA 24

Dado el problema

$$\begin{aligned}&\max\ 4x + y\\&s.a.\ 2x + y \le 9\\&\quad 4x - 3y \le 8\\&\quad x, y \ge 0\end{aligned}$$

Se pide:

1) Mostrar la tabla óptima al resolver mediante el simplex.

2) Determinar el rango de variación del coeficiente en la función objetivo de x para que la solución siga siendo óptima.

3) Plantear el problema dual, determinar el valor de las variables duales en el óptimo, y dar el rango de variación de cada término independiente en el que no varían los valores de las variables duales.

4) Resolver numéricamente y mostrando todas las tablas que sean necesarias, el mismo problema pero con variables enteras (utilizar postoptimización).

PROBLEMA 25

Una empresa produce una determinada bebida a partir de dos ingredientes básicos. El proceso de elaboración de la bebida puede llevarse a cabo de dos formas distintas. Para producir un litro de la bebida, de la forma 1 requerirá 1 litro del primer ingrediente y 2 del segundo, mientras que de la forma 2 requerirá 2 litros del primer ingrediente y 1 del segundo. El proveedor habitual le puede suministrar hasta 900 litros de cada ingrediente a un coste de 3€ por litro el primero de ellos, y de 5€ por litro el segundo. La empresa plantea el siguiente problema para determinar cómo debe ser su producción de la próxima semana para satisfacer una demanda de 500 litros de la bebida con el menor coste posible:

$$\begin{aligned}&\min z = 13x + 11y\\&x + y \geq 500\\&x + 2y \leq 900\\&2x + y \leq 900\\&x, y \geq 0\end{aligned}$$

siendo su correspondiente tabla óptima

	x	y	$s1$	$s2$	$s3$	
$-z$	0	0	15	2	0	–5700
y	0	1	1	1	0	400
$s3$	0	0	3	1	1	300
x	1	0	–2	–1	0	100

Se pide responder a las siguientes cuestiones, independientemente unas de otras, y razonando y calculando a partir de la tabla óptima:

1) ¿Qué precio mínimo debería pedir para cubrir costes por cada litro más que se le solicite de la bebida? Si la demanda se mantiene pero puede acceder a otro proveedor distinto del habitual aunque más caro para que le suministre los ingredientes, ¿qué precio máximo debería pagar por cada litro de los dos ingredientes?

2) ¿Cuánto podría variar el coste del ingrediente 1 sin que cambie la solución actual?

3) ¿Cuál sería la solución si la demanda fuera de 400 litros?

PROBLEMA 26

Un fabricante de piensos produce tres piensos compuestos, X, Y, y Z, a base de dos nutrientes N1 y N2 de los que actualmente tiene existencias por valor de 1050 y 650 kilos respectivamente.

La composición de un kilo de cada pienso viene dada en la siguiente tabla

	X	*Y*	*Z*
N1	0.5	0.6	0.5
N2	0.4	0.2	0.5

Un kilo de X produce un beneficio de 5.9 €, uno de Y 4 y uno de Z 7 €. Un contrato con un cliente obliga a fabricar al menos 300 kilos de X.

El siguiente programa lineal maximiza la cantidad a obtener por la fabricación de los piensos

$$\max G = 5.9X + 4Y + 7Z$$
$$0.5X + 0.6Y + 0.5Z \leq 1050$$
$$0.4X + 0.2Y + 0.5Z \leq 650$$
$$X \geq 300$$

La última tabla del simplex es

	G	*X*	*Y*	*Z*	*H1*	*H2*	*E3*	Cotas
G	1	0	0	0	–3	–11	0	10300
Y	0	0	1	0	2.5	–2.5	0.25	925
Z	0	0	0	1	–1	3	0.7	690
X	0	1	0	0	0	0	–1	300

1) ¿Hay más soluciones al problema? Si así fuera, ¿cuál es la estructura general de todas ellas?

2) ¿En cuánto se reduce el beneficio si las existencias del nutriente N1 se reducen en 200 kilos?

3) ¿Qué sucede si, además del contrato sobre el pienso X, el fabricante de piensos recibe un pedido de 1000 kilos de Y ?

4) Sin tener en cuenta el nuevo contrato y en las condiciones iniciales del problema, ¿le interesa producir un cuarto pienso W que da un beneficio por kilo de 6 € y en cuya composición por kilo entran 0.3 kilos de N1 y 0.35 de N2? Justificar la respuesta.

PROBLEMA 27

Un trapero se dedica a recoger tres tipos de artículos, *X*1, *X*2 y *X*3, que transporta en un vehículo que puede acarrear hasta 300 kg de peso y un volumen máximo de 480 unidades de volumen. Las características de peso, volumen y beneficio generado por cada unidad de esos artículos vienen dadas en la tabla siguiente.

Artículo	Peso (kg)	Volumen (u.v.)	Beneficio (€)
*X*1	2	3	7
*X*2	1	2	6
*X*3	3	1	4

El siguiente problema lineal proporciona el número de artículos de cada tipo que el trapero debe recoger a fin de hacer máximo el beneficio

$$\begin{aligned}
&\max W = 7X_1 + 6X_2 + 4X_3 \\
&2X_1 + X_2 + 3X_3 \le 300 \\
&3X_1 + 2X_2 + X_3 \le 480 \\
&X_i \ge 0
\end{aligned}$$

La última tabla del método simplex de este problema es

	*X*1	*X*2	*X*3	*H*1	*H*2	Cotas
–*W*	–2.2	0	0	–0.4	–2.8	–1464
*X*2	1.4	1	0	–0.2	0.6	228
*X*3	0.2	0	1	0.4	–0.2	24

1) ¿Tiene más soluciones este programa? ¿Por qué? ¿Cuál es el programa dual del dado? ¿Qué valores y dimensiones tienen las variables duales?

2) Si le ofrecen alquilarle por un día un vehículo que pudiera transportar hasta 320 kg de peso y contener 550 unidades de volumen, ¿hasta qué cantidad podría pagar por el alquiler de dicho vehículo?

3) ¿Qué ocurre si una ordenanza municipal prohibiera trasladar más de 200 artículos?

4) El trapero muestra interés por un cuarto artículo *X*4 que pesa 2.5 kg, ocupa 2 unidades de volumen y proporciona un beneficio unitario de 5 €. ¿Le interesa recoger unidades de este artículo?

I.12. Resultados de la biblioteca de problemas

RESULTADO DEL PROBLEMA 1

Sólo una: $B=(a_1,a_2)$, solución (5/3,1,0)

Valor de la función objetivo 8/3

Sí, es solución factible (verifica las restricciones) con función objetivo 1335+4001–12000=–6664, luego es mejor solución.

$B^{-1}=\begin{pmatrix}1/3 & \cdot \\ \cdot & 1\end{pmatrix}$, $y_3=B^{-1}a_3=\begin{pmatrix}-1/3\\ -1\end{pmatrix}$ es un vector con todos sus componentes < 0 y el coste reducido de x_3 es menor que 0, $\hat{c}_3=-5/3$.

RESULTADO DEL PROBLEMA 2

1) $(x,y)=(66/19,40/19)$

2) No hay solución factible.

3) Solución no acotada, dirección (1,1) siendo $x=y$.

4) $(x,y,z,t)=(12/5,0,21/10,3/5)$

5) $(x_1, x_2, x_3) = (2,2,2)$

RESULTADO DEL PROBLEMA 3

1)

$$\max 5w_1 + 4w_2$$
$$w_1 - w_2 \leq 3$$
$$w_1 - w_2 \leq 2$$
$$-w_1 + 7w_2 = -1$$
$$w_1, w_2 \geq 0$$

2)

$$\min 7w_1 + 3w_2$$
$$w_1 + w_2 = 3$$
$$-w_1 + w_2 = 4$$
$$w_1, w_2 \leq 0$$

3)

$$\min 5w_1 + 8w_2$$
$$w_1 + w_2 \geq 1$$
$$-w_1 \geq 1$$
$$w_2 = 1$$
$$w_1 \leq 0$$

RESULTADO DEL PROBLEMA 4

i.a) $(x_1, x_2, x_3, x_4) = (2,0,0,-1)$, $z = 1$

i.b) $(w_1, w_2) = (1,-1)$, $z = 1$

i.c) Se cumplen las condiciones de holgura complementaria.

ii.a) Solución no acotada, dirección (1/3,1,1/3).

ii.b) Dual no factible (observar geométricamente).

ii.c) Obvias.

RESULTADO DEL PROBLEMA 5

1) La solución verifica la restricción, sigue siendo óptima.

2) $(x_1, x_2, x_3, x_4) = (7/3, 2/3, 0, -2/3)$

3) $(x_1, x_2, x_3, x_4) = (3/2, 0, 0, 1/2)$

4) Solución no acotada.

RESULTADO DEL PROBLEMA 6

$0 \leq \theta \leq 1 \;\; (x_1, x_2, x_3) = (0, 9/2, 4) \;\; z = -13 + 13\theta$

$\theta \geq 1 \;\; (x_1, x_2, x_3) = (6, 0, 1) \;\; z = 17 - 17\theta$

RESULTADO DEL PROBLEMA 7

1) $(x_1, x_2, x_3) = (2, 0, 1)$

2) $0 \leq \theta \leq 4$ $(x_1, x_2, x_3) = (2 + 5/4\theta, 0, 1 - 1/4\theta)$ $\theta \geq 4$ $(x_1, x_2, x_3) = (3 + \theta, 0, 0)$

RESULTADO DEL PROBLEMA 8

$$\begin{aligned}
&\max 0.4x_1 + 0.5x_2 \\
&x_1 + x_2 \geq 3 && :y_1 \\
&3x_1 + 5x_2 \leq 24 && :y_2 \\
&4x_1 + 2x_2 \leq 3 \cdot 16 && :y_3 \\
&x_1, x_2 \geq 0
\end{aligned}$$

Por kg de tomate $y_2 = 0.4/3$ €/kg e $y_3 = 0$ €/taza por el vinagre.

RESULTADO DEL PROBLEMA 9

1) $x = (0, 0.5, 0)$

2) $x = (0, 0.5, 0)$

3) $x = (0, 0, 0)$

RESULTADO DEL PROBLEMA 10

1) $x_1 = 1$, $x_2 = 0$, $x_3 = 0$

2) $w_1 = 1$, $w_2 = 0$

3) Comprobar que se verifican, evidente con estas soluciones.

4) Al introducir la segunda de holgura, cambia la base pero no la solución.

RESULTADO DEL PROBLEMA 11

1) No. Las soluciones son $\lambda(2,3)+(1-\lambda)(3,1)$.

2) $$\begin{aligned} &\min 5w_1 + 7w_2 + 3w_3 \\ &w_1 + 2w_2 + w_3 \geq 4 \\ &w_1 + w_2 \geq 2 \\ &w_1, w_2, w_3 \geq 0 \end{aligned}$$

3) $(w_1, w_2, w_3) = (0, 2, 0)$

4) Mismo valor de la función objetivo y condiciones de holgura complementaria.

5) $(x, y) = (3,1)$, sólo cambia la holgura que ahora es 0.

6) $(x, y) = (3,0)$

7) Tan sólo para $\theta = 0$.

RESULTADO DEL PROBLEMA 12

1) Múltiples soluciones, $\lambda(2,1) + (1 - \lambda)(3,0)$ $\lambda \in [0,1]$ $z = 6$.

2) La solución (2,1) sigue siendo óptima para valores entre 2 y 4, y la solución (3,0) sigue siendo óptima para valores menores que 2.

3) Múltiples soluciones, $\lambda(0,3) + (1 - \lambda)(3,0)$ $\lambda \in [0,1]$ $z = 6$.

4) $w_1 = 2$, $w_2 = 0$

5) $(x_1, x_2) = (0,2)$

RESULTADO DEL PROBLEMA 13

b) Múltiples soluciones $(x,y) = \lambda(20,30) + (1 - \lambda)(40,20)\ \ \lambda \in [0,1]\ \ z = 16000$

d) $(w_1, w_2, w_3) = (0, 200, 0)$

e) $(x,y) = (0,60)$, $z = 24000$

f) $(x,y) = (50,10)$, $z = 14000$

RESULTADO DEL PROBLEMA 14

1) $(x,y) = (4,3)$, $z = 11$

2) $(x,y) = (4,0)$, $z = 8$

RESULTADO DEL PROBLEMA 20

1) $$B^{-1}\begin{pmatrix}100\\0\\0\end{pmatrix} = \begin{pmatrix}0.625 & 0 & -1.25\\0.1 & 1 & 1\\0.375 & 0 & 1.25\end{pmatrix}\begin{pmatrix}100\\0\\0\end{pmatrix} = \begin{pmatrix}62.5\\10\\37.5\end{pmatrix}$$

Se puede concluir que son los porcentajes de la Al1 (37.5) y de la Al5 (62.5) sea cuál sea la cantidad a producir.

2) La pagaría a lo sumo a 28.125 que es el valor de la primera variable dual, y para cualquier cantidad hasta la demanda (es decir, hasta las 90 toneladas) ya que nunca se hará negativa ninguna variable básica.

3) Son sus costes menos los costes reducidos, es decir,

Al2: 40–14.0625=25.9375, Al3: 65–37.5=27.5, Al4:55–25.9375=29.0625

4) $$B^{-1}\begin{pmatrix}1\\0.4-0.7\\0.6-0.4\end{pmatrix} = \begin{pmatrix}0.375\\0\\0.625\end{pmatrix},\ 30 - (30\quad 0\quad 25)\begin{pmatrix}0.375\\0\\0.625\end{pmatrix} = 3.125$$

5) Nada, no interesa producir con la nueva.

6) Habría que añadir la restricción: $x_1 \leq 30$, quedando (quitando la artificial):

	x_1	x_2	x_3	x_4	x_5	s_2	s_3	s_4	Cotas
$c_j - z_j$	0	14.0625	37.5	25.9375	0	0	6.25	0	–2531.25
x_5	0	0.1875	0.5	0.8125	1	0	–1.25	0	56.25
s_2	0	0	0	0	0	1	1	0	9
x_1	1	0.8125	0.5	0.1875	0	0	1.25	0	33.75
s_4	1	0	0	0	0	0	0	1	30

Reorganizando (a la última restar la penúltima solamente):

	x_1	x_2	x_3	x_4	x_5	s_2	s_3	s_4	Cotas
$c_j - z_j$	0	14.0625	37.5	25.9375	0	0	6.25	0	–2531.25
x_5	0	0.1875	0.5	0.8125	1	0	–1.25	0	56.25
s_2	0	0	0	0	0	1	1	0	9
x_1	1	0.8125	0.5	0.1875	0	0	1.25	0	33.75
s_4	0	–0.8125	–0.5	–0.1875	0	0	–1.25	1	–3.75

Aplicando el dual entra la tercera de holgura y sale la cuarta

	x_1	x_2	x_3	x_4	x_5	s_2	s_3	s_4	Cotas
$c_j - z_j$	0	10	35	25	0	0	0	5	–2550
x_5	0	1	1	1	1	0	0	–1	60
s_2	0	–0.65	–0.4	–0.15	0	1	0	0.8	6
x_1	1	0	0	0	0	0	0	1	30
s_4	0	0.65	0.4	0.15	0	0	1	–0.8	3

RESULTADO DEL PROBLEMA 21

1)

	x_1	x_2	x_3	x_4	x_5	Cotas
$c_j - z_j$	0	0	–0.115	0.01	0	–860
x_5	0	0	0.45	–1.5	1	900
x_2	0	1	0.075	–0.05	0	300
x_1	1	0	–0.05	0.1	0	200

Cambia la solución actual:

	x_1	x_2	x_3	x_4	x_5	Cotas
$c_j - z_j$	–0.1	0	–0.11	0	0	–880
x_5	15	0	–0.3	0	1	3900
x_2	0.5	1	0.05	0	0	400
x_4	10	0	–0.5	1	0	2000

2) Incrementando en 0.05 no interesa, la variable dual es 0.025. Incrementando 0.02 sí interesa y la nueva solución es

$$B^{-1}\overline{b} = \begin{pmatrix} 0.45 & -1.5 & 1 \\ 0.075 & -0.05 & 0 \\ -0.05 & 0.1 & 0 \end{pmatrix} \begin{pmatrix} 8200 \\ 6000 \\ 6300 \end{pmatrix} = \begin{pmatrix} 990 \\ 315 \\ 190 \end{pmatrix}$$

El nuevo beneficio es $740 + 200(0.025 - 0.020) = 741$.

Hay que añadir la restricción $7x_1 + 4x_2 \leq 2100$ que no es verificada por la solución actual. La tabla que se tendría sería:

	x_1	x_2	x_3	x_4	x_5	x_6	Cotas
$c_j - z_j$	0	0	–0.025	–0.09	0	0	–740
x_5	0	0	0.45	–1.5	1	0	900
x_2	0	1	0.075	–0.05	0	0	300
x_1	1	0	–0.05	0.1	0	0	200
x_6	0	0	0.05	–0.5	0	1	–500

Resultando la siguiente tabla óptima:

	x_1	x_2	x_3	x_4	x_5	x_6	Cotas
$c_j - z_j$	0	0	–0.034	0	0	–0.18	–650
x_5	0	0	0.3	0	1	–3	2400
x_2	0	1	0.07	0	0	–0.1	350
x_1	1	0	–0.04	0	0	0.2	100
x_4	0	0	–0.1	1	0	–2	1000

5. El coste reducido sería: $1.1-(0.025 \quad 0.09 \quad 0)\begin{pmatrix} 8 \\ 12 \\ 3 \end{pmatrix}=-0.18$. No interesa.

RESULTADO DEL PROBLEMA 22

1) $w_1 = 0$ Por un gramo extra de azúcar no se pagaría nada, pues no hay cambio en la función objetivo (hay holgura)

$w_2 = 0.03$ Por un gramo extra de miel pagaría un sobrecoste de 0.03, ya que un aumento de disponibilidad en un gramo de miel, supone un beneficio extra de 0.03 € (lo que varía la función objetivo)

$w_3 = -0.03$ Si el número de piruletas que hay que hacer aumenta en 1 el beneficio disminuye en 0.03 €. Alternativamente, si disminuyera en 1 piruleta la cantidad que APA quiere llevar, lograría un beneficio de 0.03 € mayor.

2)

	A	*B*	*C*	*Holgura* 1	*Holgura* 2	*Exceso* 3	
$c_j - z_j$	0	0	0	0	–0.03	–0.03	–2100
Holgura 1	0	–2	0	1	–1	6	10000
C	0	0.5	1	0	0.5	2	5000
A	1	0.5	0	0	–0.5	–3	15000

	A	*B*	*C*	*Holgura* 1	*Holgura* 2	*Exceso* 3	
$c_j - z_j$	0	0	0	0	–0.03	–0.03	–2100
Holgura 1	0	0	4	1	1	14	30000
B	0	1	2	0	1	4	10000
A	1	0	–1	0	–1	–5	10000

Tiene infinitas soluciones de la forma

$$\begin{pmatrix} A \\ B \\ C \end{pmatrix} = \lambda \begin{pmatrix} 15000 \\ 0 \\ 5000 \end{pmatrix} + (1-\lambda) \begin{pmatrix} 10000 \\ 10000 \\ 0 \end{pmatrix} \quad \lambda \in [0,1]$$

3) Ninguna de las soluciones cumple la nueva restricción, así que hay que añadir $B \geq 12000$, es decir, $B - Exceso_4 = 12000$. Sería:

	A	*B*	*C*	*Holgura* 1	*Holgura* 2	*Exceso* 3	*Exceso* 4	
$c_j - z_j$	0	0	0	0	–0.03	–0.03	0	–2100
Holgura 1	0	–2	0	1	–1	6	0	10000
C	0	0.5	1	0	0.5	2	0	5000
A	1	0.5	0	0	–0.5	–3	0	15000
Exceso 4	0	1	0	0	0	0	–1	12000

Reorganizando:

	A	*B*	*C*	*Holgura* 1	*Holgura* 2	*Exceso* 3	*Exceso* 4	
$c_j - z_j$	0	0	0	0	–0.03	–0.03	0	–2100
Holgura 1	0	–2	0	1	–1	6	0	10000
C	0	0.5	1	0	0.5	2	0	5000
A	1	0.5	0	0	–0.5	–3	0	15000
Exceso 4	0	–1	0	0	0	0	1	–12000

	A	*B*	*C*	*Holgura* 1	*Holgura* 2	*Exceso* 3	*Exceso* 4	
$c_j - z_j$	0	0	0	0	–0.03	–0.03	0	–2100
Holgura 1	0	–2	0	1	–1	6	–2	34000
C	0	0	1	0	0.5	2	0.5	–1000
A	1	0	0	0	–0.5	–3	0.5	9000
B	0	1	0	0	0	0	–1	12000

Dual no acotado, primal no factible.

Con la otra tabla sería:

	A	*B*	*C*	*Holgura* 1	*Holgura* 2	*Exceso* 3	*Exceso 4*	
$c_j - z_j$	0	0	0	0	–0.03	–0.03		–2100
Holgura 1	0	0	4	1	1	14		30000
B	0	1	2	0	1	4		10000
A	1	0	–1	0	–1	–5		10000
Exceso 4	0	1	0	0	0	0	–1	12000

Reorganizando

	A	*B*	*C*	*Holgura* 1	*Holgura* 2	*Exceso* 3	*Exceso* 4	
$c_j - z_j$	0	0	0	0	–0.03	–0.03	0	–2100
Holgura 1	0	0	4	1	1	14	0	30000
B	0	1	2	0	1	4	0	10000
A	1	0	–1	0	–1	–5	0	10000
Exceso 4	0	0	2	0	1	4	1	–2000

Directamente puede verse que es dual no acotado, luego primal no factible.

4)

$$\hat{c}_j = c_j - z_j = c_j - c_B^T B^{-1} a_j = c_j - y^T a_j =$$

$$= 0.22 - (0 \quad 0.15 \quad 0.09)(-2 \quad 2 \quad -1)^T =$$ Interesa.

$$= 0.22 - (0 \quad 0.03 \quad -0.03)(12 \quad 8 \quad 1)^T = 0.01$$

$$B^{-1}a_j = \begin{pmatrix} 1 & -1 & -6 \\ 0 & 0.5 & -2 \\ 0 & -0.5 & 3 \end{pmatrix} \begin{pmatrix} 12 \\ 8 \\ 1 \end{pmatrix} = \begin{pmatrix} -2 \\ 2 \\ -1 \end{pmatrix}$$

	A	*B*	*C*	*Holgura* 1	*Holgura* 2	*Exceso* 3	*NUEVA*	
$c_j - z_j$	0	0	0	0	–0.03	–0.03	0.01	–2100
Holgura 1	0	–2	0	1	–1	6	–2	10000
C	0	0.5	1	0	0.5	2	2	5000
A	1	0.5	0	0	–0.5	–3	–1	15000

	A	*B*	*C*	*Holgura* 1	*Holgura* 2	*Exceso* 3	*NUEVA*	
$c_j - z_j$	0	–0.0025	–0.005	0	–0.0325	–0.0301	0	–2125
Holgura 1	0	–1.5	1	1	–0.5	8	0	15000
NUEVA	0	0.25	0.5	0	0.25	1	1	2500
A	1	0.75	0.5	0	–0.25	–2	0	17500

RESULTADO DEL PROBLEMA 23

a) Base (a4,a5) Punto (0,0,0) holguras (9,24) Z=0

Base (a4,a1) Punto (8,0,0) holguras (1,0) Z=80

Base (a4,a2) Punto (0,12,0) holguras (–3,0) NO FACTIBLE

Base (a4,a3) Punto (0,0,6) holguras (3,0) Z=72

Base (a5,a1) Punto (9,0,0) holguras (0,–3) NO FACTIBLE

Base (a5,a2) Punto (0,9,0) holguras (0,6) Z=63

Base (a5,a3) Punto (0,0,9) holguras (0,–12) NO FACTIBLE

Base (a1,a2) Punto (6,3,0) holguras (0,0) Z=81

Base (a1,a3) Punto (66/7,0,–3/7) holguras (0,0) NO FACTIBLE

Base (a2,a3) Punto (0,6,3) holguras (0,0) Z=78

b) $B^{-1} = \begin{pmatrix} -2 & 1 \\ 3 & -1 \end{pmatrix}$

	x_1	x_2	x_3	x_4	x_5	
	0	0	–1	–1	–3	–81
x_1	1	0	2	–2	1	6
x_2	0	1	–1	3	–1	3

Añadir $x_1 - 2x_2 \geq 1$ $6 - 2 \cdot 3 = 0 < 1$ $x_1 - 2x_2 - x_6 = 1$

	0	0	–1	–1	–3	0	–81
x_1	1	0	2	–2	1	0	6
x_2	0	1	–1	3	–1	0	3
x_6	1	–2	0	0	0	–1	1

Organizando:

	0	0	–1	–1	–3	0	–81
x_1	1	0	2	–2	1	0	6
x_2	0	1	–1	3	–1	0	3
x_6	0	0	4	–8	3	1	–1

Dual:

	0	0	–3/2	0	–27/8	–1/8	–647/8
x_1	1	0	1	0	1/4	–1/4	25/4
x_2	0	1	1/2	3	1/8	3/8	21/8
x_4	0	0	–1/2	1	–3/8	–1/8	1/8

RESULTADO DEL PROBLEMA 24

a)

	X	*Y*	*S*1	*S*2	
Z	4	1	0	0	0
*S*1	2	1	1	0	9
*S*2	4	–3	0	1	8

	X	*Y*	*S*1	*S*2	
Z	0	4	0	–1	–8
S1	0	5/2	1	–1/2	5
X	1	–3/4	0	1/4	2

	X	*Y*	*S*1	*S*2	
Z	0	0	–8/5	–1/5	–16
Y	0	1	2/5	–1/5	2
X	1	0	3/10	1/10	7/2

b)
$$\left.\begin{array}{ll} 0-(2/5+c_x 3/10)<0 & \Rightarrow c_x > -4/3 \\ 0-(-1/5+c_x 1/10)<0 & \Rightarrow c_x > 2 \end{array}\right\} \Rightarrow \quad c_x > 2$$

c) Dual:
$$\begin{array}{l} \min 9w+8t \\ 2w+4t \geq 4 \\ w-3t \geq 1 \\ w,t \geq 0 \end{array}$$
Valor variables duales en el óptimo: $w = 8/5 \quad t = 1/5$

$$B^{-1}b = \begin{pmatrix} \dfrac{2}{5} & \dfrac{-1}{5} \\ \dfrac{3}{10} & \dfrac{1}{10} \end{pmatrix} \begin{pmatrix} 9+inc1 \\ 8 \end{pmatrix} = \begin{pmatrix} \dfrac{18+2inc1-8}{5} \\ \dfrac{27+3inc1+8}{10} \end{pmatrix} \geq 0 \Rightarrow \begin{array}{l} 10+2inc1 \geq 0 \\ 35+3inc1 \geq 0 \end{array} \Rightarrow inc1 \geq -5$$

$$B^{-1}b = \begin{pmatrix} 2/5 & -1/5 \\ 3/10 & 1/10 \end{pmatrix} \begin{pmatrix} 9 \\ 8+inc2 \end{pmatrix} = \begin{pmatrix} \dfrac{18-8-inc2}{5} \\ \dfrac{27+8+inc2}{10} \end{pmatrix} \geq 0 \Rightarrow \begin{array}{l} 10-inc2 \geq 0 \\ 35+inc2 \geq 0 \end{array} \Rightarrow \begin{array}{l} inc2 \geq -35 \\ inc2 \leq 10 \end{array}$$

d) Enteras: Branch and Bound. $\widehat{Z} = -\infty$

Nodo 0: Relajadas Solución: (7/2, 3) *Z*= 16 No entera. Ramificar en $x \leq 3$ y $x \geq 4$.

Nodo 1: Añadir la restricción $x \leq 3$, es decir, $x + S3 = 3$

	x	*y*	*S*1	*S*2	*S*3	
Z	0	0	–8/5	–1/5	0	–16
Y	0	1	2/5	–1/5	0	2
X	1	0	3/10	1/10	0	7/2
*S*3	1	0	0	0	1	3

	x	*y*	*S*1	*S*2	*S*3	
Z	0	0	–8/5	–1/5	0	–16
Y	0	1	2/5	–1/5	0	2
X	1	0	3/10	1/10	0	7/2
*S*3	0	0	–3/10	–1/10	1	–1/2

	x	*y*	*S*1	*S*2	*S*3	
Z	0	0	–1	0	–2	–15
Y	0	1	1	0	–2	3
X	1	0	0	0	–10	3
*S*2	0	0	3	1	–10	5

Solución nodo: (3,4) *Z*=15, entera → podar y $\widehat{Z} = 15$. No podar otros, seguir nodo 0

Nodo 2: Añadir la restricción $x \geq 4$, es decir, $x - S3 = 4$

	x	*y*	*S*1	*S*2	*S*3	
Z	0	0	–8/5	–1/5	0	–16
Y	0	1	2/5	–1/5	0	2
X	1	0	3/10	1/10	0	7/2
*S*3	1	0	0	0	–1	4

	x	*y*	*S*1	*S*2	*S*3	
Z	0	0	–8/5	–1/5	0	–16
Y	0	1	2/5	–1/5	0	2
X	1	0	3/10	1/10	0	7/2
*S*3	0	0	3/10	1/10	1	–1/2

No factible, podar. Todos los nodos explorados, la solución óptima es (3,3), *Z*=15.

RESULTADO DEL PROBLEMA 26

$$\max G = 5.9X + 4Y + 7Z$$
$$0.5X + 0.6Y + 0.5Z \leq 1050$$
$$0.4X + 0.2Y + 0.5Z \leq 650$$
$$X \geq 300$$

	G	*X*	*Y*	*Z*	*H*1	*H*2	*E*3	Cotas
G	1	0	0	0	–3	–11	0	–10300
Y	0	0	1	0	2.5	–2.5	0.25	925
Z	0	0	0	1	–1	3	0.7	690
X	0	1	0	0	0	0	–1	300

1) Sí, nueva tabla:

	G	*X*	*Y*	*Z*	*H*1	*H*2	*E*3	Cotas

G	1	0	0	0	–3	–11	0	–10300
Y	0	0	1	–0.36	2.86	–3.57	0	678.57
*E*3	0	0	0	1.43	–1.43	4.29	1	985.71
X	0	1	0	1.43	–1.43	4.29	–1	1285.71

Estructura: $\lambda\begin{pmatrix}300\\925\\690\end{pmatrix}+(1-\lambda)\begin{pmatrix}1285.71\\678.57\\0\end{pmatrix}\qquad \lambda\in[0,1]$

2) $\quad b=\begin{pmatrix}1050\\650\\300\end{pmatrix}\qquad b'=\begin{pmatrix}1050\\450\\300\end{pmatrix}\qquad B^{-1}b'=\begin{pmatrix}2.5&-2.5&-0.25\\-1&3&-0.7\\0&0&1\end{pmatrix}\begin{pmatrix}1050\\450\\300\end{pmatrix}=\begin{pmatrix}1425\\90\\300\end{pmatrix}\qquad G=8100$

3) Añadir $Y\geq 1000$ que no se cumple:

	X	*Y*	*Z*	*H*1	*H*2	*E*3	*E*4	Cotas		*X*	*Y*	*Z*	*H*1	*H*2	*E*3	*E*4	Cotas
G	0	0	0	–3	–11	0	0	–10300		0	0	0	–3	–11	0	0	–10300
Y	0	1	0	2.5	–2.5	0.25	0	925		0	1	0	2.5	–2.5	0.25	0	925
Z	0	0	1	–1	3	0.7	0	690		0	0	1	–1	3	0.7	0	690
X	1	0	0	0	0	–1	0	300		1	0	0	0	0	–1	0	300
*E*4	0	1	0	0	0	0	–1	1000		0	0	0	2.5	–2.5	0.25	1	–75

	X	*Y*	*Z*	*H*1	*H*2	*E*3	*E*4	Cotas
G	0	0	0	–14	0	–1.1	–4.4	–9970
Y	0	1	0	0	0	0	–1	1000
Z	0	0	1	2	0	1	1.2	600
X	1	0	0	0	0	–1	0	300
*H*2	0	0	0	–1	1	–0.1	–0.4	30

4) Interesa:

$$6-\begin{pmatrix}7&7&5.9\end{pmatrix}\begin{pmatrix}2.5&-2.5&-0.25\\-1&3&-0.7\\0&0&1\end{pmatrix}\begin{pmatrix}0.3\\0.35\\0\end{pmatrix}=6-\begin{pmatrix}7&7&5.9\end{pmatrix}\begin{pmatrix}-0.125\\0.75\\0\end{pmatrix}=1.625$$

RESULTADO DEL PROBLEMA 27

No hay más soluciones, cualquiera en que una variable no básica pase a tener valor positivo tendrá un valor estrictamente menor de la función objetivo que el actual, ya que todos los costes reducidos son distintos de 0.

El dual: $\left.\begin{array}{l} \min 300w_1 + 480w_2 \\ 2w_1 + 3w_2 \geq 7 \\ w_1 + 2w_2 \geq 6 \\ 3w_1 + w_2 \geq 4 \\ w_1, w_2 \geq 0 \end{array}\right\} \quad w_1^* = 0.4€/\text{kg} \quad w_2^* = 2.8€/\text{u.v.}$

$$\overline{b} = \begin{pmatrix} 320 \\ 550 \end{pmatrix} \quad B^{-1}\overline{b} = \begin{pmatrix} -0.2 & 0.6 \\ 0.4 & -0.2 \end{pmatrix}\begin{pmatrix} 320 \\ 550 \end{pmatrix} = \begin{pmatrix} 266 \\ 18 \end{pmatrix} = \begin{pmatrix} X_2 \\ X_3 \end{pmatrix} \quad \overline{W} = 1668$$

$\overline{W} - W = 1668 - 1464 = 204€$

Añadir $X_1 + X_2 + X_3 \leq 200 \rightarrow X_1 + X_2 + X_3 + H_3 = 200$

	*X*1	*X*2	*X*3	*H*1	*H*2	*H*3	Cotas
–*W*	–2.2	0	0	–0.4	–2.8	0	–1464
*X*2	1.4	1	0	–0.2	0.6	0	228
*X*3	0.2	0	1	0.4	–0.2	0	24
*H*3	1	1	1	0	0	1	200

	*X*1	*X*2	*X*3	*H*1	*H*2	*H*3	Cotas
–*W*	–2.2	0	0	–0.4	–2.8	0	–1464
*X*2	1.4	1	0	–0.2	0.6	0	228
*X*3	0.2	0	1	0.4	–0.2	0	24
*H*3	–0.6	0	0	–0.2	–0.4	1	–52

	*X*1	*X*2	*X*3	*H*1	*H*2	*H*3	Cotas
–*W*	–1	0	0	0	–2	–2	–1360
*X*2	2	1	0	0	1	–1	280

*X*3	–1	0	1	0	–1	2	–80
*H*1	3	0	0	1	2	–5	260

	*X*1	*X*2	*X*3	*H*1	*H*2	*H*3	Cotas
–*W*	0	0	–1	0	–1	–4	–1280
*X*2	0	1	2	0	–1	3	120
*X*1	1	0	–1	0	1	–2	80
*H*1	0	0	3	1	–1	1	20

Calculamos su coste reducido: $c_4 - z_4 = 5 - \begin{pmatrix} 0.4 & 2.8 \end{pmatrix} \begin{pmatrix} 2.5 \\ 2 \end{pmatrix} = 5 - 6.6 = -1.6$. No le interesa

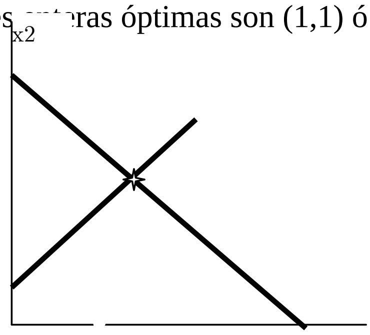
II. Optimización lineal entera mixta

II.1. Resolución por métodos heurísticos

La solución de un problema lineal entero mixto (MIP) no es necesariamente la solución del problema relajado[7] discretizada heurísticamente (redondeada a los valores enteros más próximos), se puede producir una posible pérdida de optimalidad o de factibilidad (es decir, no tiene por qué ser matemáticamente óptima o factible). No obstante, esta discretización heurística puede ser una solución aproximada válida, en el sentido físico del problema modelado no en el matemático, en el caso de que las variables enteras tomen valores muy elevados.

En este ejemplo se obtienen soluciones no factibles al redondear. En él la solución del problema LP es (1.5,2). Las soluciones enteras más próximas son (1,2) ó (2,2) que resultan infactibles. Mientras que las soluciones enteras óptimas son (1,1) ó (2,1).

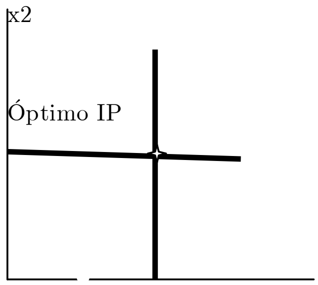
$$
\begin{aligned}
&\max \quad x_2 \\
&-x_1 + x_2 \le 0.5 \\
&x_1 + x_2 \le 3.5 \\
&x_1, x_2 \ge 0 \quad x_i \in \mathbb{Z}
\end{aligned}
$$



En el siguiente ejemplo las soluciones redondeadas son no óptimas. La solución del problema LP es (2,9/5). Las soluciones redondeadas son (2,2), que no es factible, y (2,1). Mientras que la solución entera óptima es (0,2).

[7] Se entiende por problema relajado el problema original pero sin las condiciones de integralidad de las variables.

$$\max \quad x_1 + 5x_2$$
$$x_1 + 10x_2 \leq 20$$
$$x_1 \leq 2$$
$$x_1, x_2 \geq 0 \quad x_i \in \mathbb{Z}$$



A pesar de los contraejemplos mencionados, existen métodos metaheurísticos de búsqueda como alternativa a la programación matemática (algoritmos genéticos, búsqueda heurística, etc.) cuando los problemas son tan complejos que no pueden solucionarse mediante procedimientos que aseguren la solución óptima por falta de tiempo o memoria.

Como criterio general la solución de problemas MIP es más difícil y costosa tanto en tiempo como en memoria de ordenador que la de un problema LP.

II.2. Método de ramificación y acotamiento

En un problema MIP se pierde la convexidad de la región factible, puesto que los puntos de su interior no se pueden poner como combinación lineal convexa de sus puntos extremos. También se pierde la potencia matemática asociada a las variables continuas (derivadas, condiciones de optimalidad, sensibilidades, etc.). Por consiguiente, la solución de un problema MIP es mucho más difícil que la de un problema LP.

La enumeración exhaustiva sería una posibilidad, pero no es un método de solución práctico ya que existen 2^n posibles soluciones cuando todas las variables son binarias.

El método de ramificación y acotamiento (*branch and bound*) recurre a una enumeración implícita (como contrapuesta a la exhaustiva) de las soluciones enteras factibles. Utiliza el principio de divide y vencerás.

Dividir es ramificar, es decir, partir el conjunto de soluciones enteras en subconjuntos disjuntos cada vez menores.

Conquistar es acotar, determinar el valor de la mejor solución del subconjunto (la cota).

En un problema de maximización una *cota inferior* de la solución óptima de un problema MIP es la mayor solución entera factible encontrada hasta el momento.

En un problema de maximización una *cota superior* de la solución óptima de un problema MIP es la solución óptima del problema lineal relajado RMIP o LP.

Eliminar es podar la rama (o nodo) del árbol si la cota indica que no puede contener la solución óptima.

Sea el siguiente problema de programación lineal entera mixta

$$\begin{aligned}
&\max z = \sum_{j=1}^{n} c_j x_j \\
&\sum_{j=1}^{n} a_{ij} x_j \le b_i \quad i = 1,\dots,m \\
&x_j \ge 0 \qquad\qquad j = 1,\dots,n \\
&x_j \in \mathbb{Z} \qquad\qquad j = 1,\dots,I \quad I \le n
\end{aligned}$$

El procedimiento que realiza el método de ramificación y acotamiento es el siguiente:

1) Inicialización de la función objetivo al valor de $\hat{z} = -\infty$ para un problema de maximización (si fuera de minimización sería $\hat{z} = \infty$). $\hat{z}$ es la mejor solución entera encontrada hasta el momento.

Resolver una relajación del problema (habitualmente es la lineal, aunque pueden usarse otras). Éste problema es el nodo raíz.

2) Se aplica al problema el acotamiento o poda y el criterio de optimalidad que se describen a continuación. Si no se puede etiquetar el problema como podado, comienza una iteración completa.

Pasos en cada iteración

RAMIFICACIÓN

Seleccionar uno de los nodos (por ejemplo, el creado más recientemente[8]) entre los no explorados (subproblemas restantes).

Seleccionar una variable entera que tenga valor no entero en la solución óptima del subproblema relajado. Un criterio de selección es elegir la primera en el orden natural[9]. Si $\hat{x}_j$ es el valor óptimo en el problema relajado, entonces se ramifica en dos *ramas*, donde cada una incorpora una de estas restricciones respectivamente $\begin{cases} x_j \le \lfloor \hat{x}_j \rfloor \\ x_j \ge \lfloor \hat{x}_j \rfloor + 1 \end{cases}$, siendo $\lfloor \hat{x}_j \rfloor$ parte entera de $\hat{x}_j$.

La ramificación en problemas con variables binarias consiste en fijar variables, ya que si la variable toma un valor fraccionario será un valor entre 0 y 1 y en cada rama lo que se hace es fijar el valor de esa variable a 0 ó a 1. Esto impide que se pueda ramificar dos veces una variable binaria, pero sí puede ocurrir en variables enteras.

La ramificación añade una restricción, luego el problema será resuelto por análisis de sensibilidad mediante el método simplex dual. El problema primal de la nueva rama resulta infactible, cada rama elimina la solución óptima del subproblema ascendiente. El problema dual es factible pero no óptimo.

ACOTAMIENTO

Para cada uno de los subproblemas (nodos) se obtiene su *cota* z mediante relajación, resolución del problema lineal.

Cuando se aplica a problemas de programación entera pura (PIP o BIP) el método permite redondear las cotas de la función objetivo (hacia abajo o arriba según corresponda) a valores enteros si todos los coeficientes que aparecen en las restricciones o en la función objetivo son enteros.

---

[8] Más adelante se verán otros criterios más elaborados.

[9] Este criterio también puede ser más elaborado.

PODA

Se intentan *podar* nodos y, por tanto, eliminar ramas que saldrían de ellos, con los siguientes criterios de poda para un problema de maximización:

- Se ha obtenido una solución entera. El nodo será descartado, aunque todavía existen dos posibilidades para la función objetivo z. En el caso de maximización:
- $z \leq \hat{z}$, es decir, es peor que la mejor solución entera encontrada hasta este momento $\hat{z}$, con lo cual se abandona esta solución. Se poda la rama.
- $z > \hat{z}$, es decir, es mejor que la mejor solución entera encontrada hasta este momento $\hat{z}$, con lo que ésta será la nueva solución entera, actualizando $\hat{z} = z$.
- La solución no es entera y $z \leq \hat{z}$, con lo que ramificando a partir de aquí no se puede obtener una solución mejor, por lo que se descarta continuar por ella. Se poda la rama.
- El subproblema es infactible (ramificando, es decir, añadiendo restricciones no se recuperará la factibilidad).

3) Criterio de optimalidad

Detener el proceso cuando no existan subproblemas sin analizar. Entonces, la solución entera actual $\hat{z}$ es la óptima.

4) Realizar otra iteración.

Con este procedimiento, se va creando un *árbol de soluciones*, de modo que cada nodo es el problema original con una serie de restricciones añadidas.

Veamos el método de ramificación y acotamiento con un ejemplo. Sea el problema MIP siguiente

$$
\begin{array}{llllll}
\max & z = 4x_1 - 2x_2 + 7x_3 - x_4 & & & & \\
x_1 & & +5x_3 & & \leq 10 \\
x_1 & +x_2 & -x_3 & & \leq 1 \\
6x_1 & -5x_2 & & & \leq 0 \\
-x_1 & & +2x_3 & -2x_4 & \leq 3 \\
x_j \geq 0 & j = 1,\ldots,4 & & & \\
x_j \in \mathbb{Z} & j = 1,\ldots,3 & & &
\end{array}
$$

Se resuelve el problema LP relajado y se obtiene $z = 14.25$ y $(x_1, x_2, x_3, x_4) = (1.25, 1.5, 1.75, 0)$.

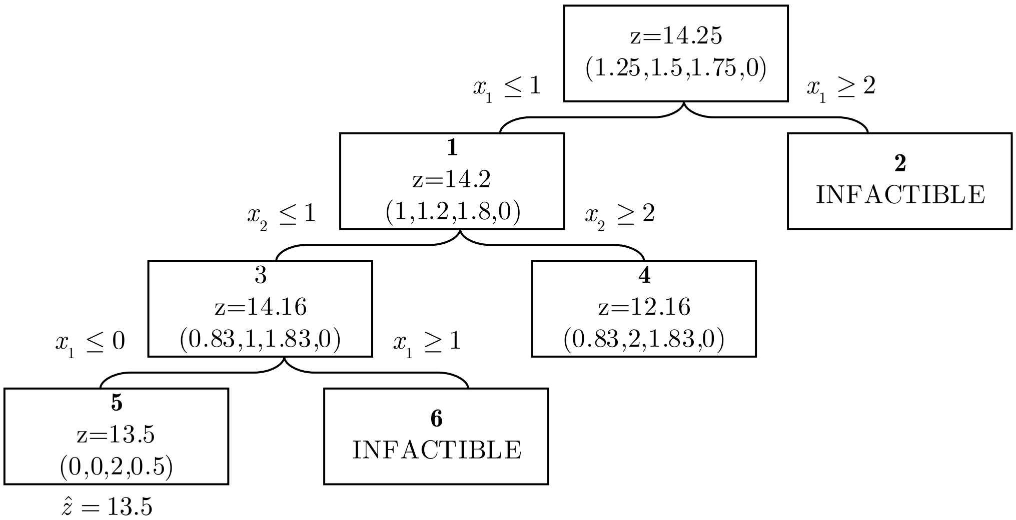
Se ramifica con la primera variable que debiera ser entera y no lo es, x_1. Cada rama corresponde a añadir las restricciones $x_1 \leq 1$ (nodo 1) y $x_1 \geq 2$ (nodo 2), respectivamente. Se resuelve el nodo 1 y se obtiene $z = 14.2$ y $(x_1, x_2, x_3, x_4) = (1, 1.2, 1.8, 0)$. Luego cualquier solución óptima descendiente de este nodo cumplirá $\hat{z} \leq 14.2$. El nodo 2 resulta infactible luego esa rama queda podada.

Partiendo del nodo 1 se ramifica con la primera variable que debiera ser entera y no lo es, x_2, de manera que el nodo 3 tiene las restricciones originales más $x_1 \leq 1$ y $x_2 \leq 1$ y en el nodo 4 se añaden $x_1 \leq 1$ y $x_2 \geq 2$. Resolviendo el nodo 3 se obtiene $z = 14.1\widehat{6}$ y $(x_1, x_2, x_3, x_4) = (0.8\widehat{3}, 1, 1.8\widehat{3}, 0)$. Luego cualquier solución óptima descendiente de esta rama cumplirá $\hat{z} \leq 14.1\widehat{6}$. Al resolver el nodo 4 $z = 12.1\widehat{6}$ y $(x_1, x_2, x_3, x_4) = (0.8\widehat{3}, 2, 1.8\widehat{3}, 0)$. Luego sus descendientes cumplen $\hat{z} \leq 12.1\widehat{6}$.

Se selecciona ramificar el nodo 3 por tener la mayor función objetivo (ser el nodo más prometedor, aunque también este criterio puede ser modificado) y se ramifica con la variable x_1. En el nodo 5 se habrán añadido las restricciones $x_1 \leq 1$, $x_2 \leq 1$ y $x_1 \leq 0$ y su resultado es $z = 13.5$ y $(x_1, x_2, x_3, x_4) = (0, 0, 2, 0.5)$. Resulta la primera solución entera factible del problema MIP, con valor de la función objetivo $\hat{z} = 13.5$. Una vez obtenida una solución entera se intentan podar las ramas no exploradas y se observa que el nodo 4 puede ser descartado porque su función objetivo es menor (en maximización) que la solución entera actual.

Para el nodo 6 se añaden las restricciones $x_1 \leq 1$, $x_2 \leq 1$ y $x_1 \geq 1$ y resulta infactible.

Como no existen ramas sin explorar se ha alcanzado la solución óptima $\hat{z} = 13.5$ y $(x_1, x_2, x_3, x_4) = (0, 0, 2, 0.5)$.



En la implantación del método existen diversas estrategias de búsqueda que no son estrictas y pueden ser probadas en la búsqueda de la solución óptima:

- Selección de la rama a resolver

Se puede llevar a cabo una búsqueda en profundidad, resolviendo siempre el último nodo generado (sin siquiera resolver los dos descendientes que se generan al ramificar, sino sólo uno de ellos). Con esta estrategia de *factibilidad,* se supone que es más rápido encontrar una solución entera inicial con la que comparar las cotas, aunque ésta pueda estar muy alejada del óptimo y tener un mal valor de comparación. Por otra parte, la búsqueda en profundidad es adecuada para utilizar la postoptimización por el método simplex.

Otra opción es seleccionar el nodo más prometedor, es decir, aquél cuya función objetivo sea mayor en un problema de maximización. Esta estrategia se denomina de *optimalidad*.

Existen también otros criterios más sofisticados que intentan explorar o predecir mínimamente en un nodo qué se puede esperar si se ramifica a partir de él.

- Selección de la variable entera a ramificar

Se ha presentado el criterio de selección que elige la encontrada en primer lugar. Sin embargo, otros criterios como elegir la de mayor o menor infactibilidad entera resultan más adecuados en las estrategias previas.

También se puede modificar (relajar) el criterio de parada del método para evitar explorar exhaustivamente el árbol. Es fundamental en la solución de problemas reales ya que puede marcar la diferencia entre acabar un problema, con una solución cuasióptima dentro de una cierta tolerancia conocida, o no solucionar el problema. Para ello se añade en el criterio de poda una cierta tolerancia relativa o absoluta que descarta una solución no entera sólo ligeramente mejor que la solución entera actual, es decir, poco prometedora. El criterio para maximización descarta las ramas con función objetivo z que cumplen siendo α el error de tolerancia relativo (por ejemplo 10^{-3}) o $\hat{z} \leq z \leq \hat{z} + \beta$ siendo el error de tolerancia absoluto, ambas constantes conocidas.

Otra técnica que se utiliza en optimizadores reales es suponer que se dispone inicialmente de una cota válida de la función objetivo (inventada razonablemente u obtenida mediante heurísticos) (denominada *incumbent*) para poder podar rápidamente el árbol. Si esta cota está muy alejada del valor del problema relajado tendrá mucha capacidad de poda pero puede hacer que se pode incluso la rama que contiene la solución óptima. Por otra parte, si es muy próxima al valor del problema relajado su capacidad de poda será muy pequeña.

II.3. Método de planos de corte

Es el primer método que se propuso para resolver problemas de programación entera. Se llama *plano de corte* a toda restricción válidamente deducida de las restricciones del problema.

La idea del procedimiento es ir reduciendo la región factible relajada añadiendo planos de corte que reducen la región factible continua pero no la entera, hasta encontrar una solución entera. El método sería el siguiente:

1) Inicializar, resolviendo la relajación lineal del problema.
2) Si la solución óptima del problema relajado es entera, es la solución del problema entero. Si no, ir al paso 3.
3) Obtener un plano de corte, que la solución óptima continua actual no verifique (se dice, un plano de corte que esté *violado* por la solución actual o que *aísle* a la solución actual).
4) Añadir el plano de corte a las restricciones del problema y reoptimizar. Volver al paso 2.

Naturalmente, la mayor dificultad se encuentra en deducir planos de corte válidos para las soluciones enteras y que aíslen a la solución continua. Existen diversos planos de corte, los primeros y más conocidos son los formulados por Gomory (corte fraccional y corte totalmente entero).

Este método en general no asegura la convergencia, pero para los de Gomory, en particular, manteniendo determinadas normas a la hora de aplicarlos, sí está demostrada, aunque puede ser muy lenta.

Actualmente, la metodología de planos de corte como método de resolución independiente está en desuso por mostrarse menos eficiente que el método anterior, sin embargo, combinado con otros procedimientos está tomando mucho auge, pues puede acelerar enormemente la convergencia de otros algoritmos.

En concreto, el método más actual y que mejores resultados está dando se ve en la siguiente sección y es denominado método de ramificación y corte.

II.4. Método de ramificación y corte

Este método es una combinación de los dos anteriores. Básicamente consiste en aplicar un método de ramificación y acotamiento, introduciendo planos de corte en los nodos del árbol para acelerar la convergencia. De forma general, una iteración del método sería:

1) Elegir un nodo (en el inicio es el nodo raíz que es el problema original relajado).
2) Resolver el problema de ese nodo.
3) Decidir si generar o no planos de corte. En caso afirmativo, obtenerlos y añadirlos al subproblema y resolver.
4) Estudiar las condiciones de poda del método de ramificación y acotamiento. Si se dan, podar esa rama e ir a otro nodo. Si no, ramificar con los criterios del método de ramificación y acotamiento.

Como se ve, la única diferencia con el método de ramificación y acotamiento es que en cada nodo hay que decidir si generar o no planos de corte que añadir al subproblema y reoptimizar.

Aunque la diferencia teórica no es muy grande, en la práctica produce muy buenos resultados. Claro que hay varias cuestiones que resolver, por ejemplo, cuál es el criterio para decidir si generar planos de corte, o por ejemplo, qué planos de corte, ya que hay de muchos tipos. Actualmente, hay diferentes implantaciones de este procedimiento y, en general, particularizados a ciertas clases de problemas, de modo que los planos de corte son específicos para esa clase. Sin embargo, ya existe una implantación general en un código comercial en el que se pueden introducir planos de corte de varios tipos en el proceso (GUB, *cover, cliques, flow covers and Gomory cuts*).

II.5. Preproceso y reformulación

En la resolución de problemas de programación lineal entera han sido fundamentales los avances que se han producido para permitir la resolución de problemas de gran tamaño. Entre ellos cabe destacar: el *preproceso automático*, tanto para problemas continuos como enteros. Como se ha visto en la resolución, una buena formulación de un problema puede ser de vital importancia a la hora de resolverlo. Una medida de la bondad de una formulación en programación entera que se suele utilizar es la diferencia entre el óptimo del problema relajado y el óptimo entero (habitualmente se utiliza el término anglosajón *gap*). Un valor pequeño de esta diferencia, suele implicar un tiempo de resolución reducido, que a la postre es lo que se desea.

A continuación se presentan, brevemente, en dos bloques las principales técnicas de preproceso. El primer bloque son técnicas de tipo general, aplicables a cualquier problema lineal y el segundo bloque es particular para programación entera.

II.5.1. Preproceso general

El objeto de este preproceso es la reducción de las dimensiones del problema.

Eliminación de filas múltiples y dominadas

El objetivo de este método es detectar y eliminar filas repetidas o dominadas por otras y eliminarlas del problema. Una consecuencia indirecta es que en algunos casos se pueden detectar situaciones de infactibilidad con este procedimiento.

La aplicación de la técnica es básicamente por comparación, de modo que si hay una fila en la que los elementos distintos de cero son un subconjunto de los de otra, ambas pueden ser comparadas para ver si es posible eliminar una de ellas o detectar una infactibilidad.

Eliminación de restricciones redundantes

La idea básica es detectar la redundancia por medio de las cotas de las variables. En el proceso también se pueden fijar algunas variables a sus cotas y detectar infactibilidad.

La idea básica es que si una restricción se satisface incluso en la situación más “difícil” la restricción puede ser eliminada.

Por ejemplo, para restricciones de se igualan a su cota superior las variables con coeficientes positivos y a la inferior las demás. Las siguientes restricciones son redundantes para variables con valores entre 0 y 1

$$\begin{aligned} 3x_1 + 2x_2 &\le 6 \\ 3x_1 - 2x_2 &\le 3 \\ 3x_1 - 2x_2 &\ge -3 \end{aligned}$$

Las restricciones suelen ser redundantes a consecuencia del proceso de reforzamiento de cotas que se explica a continuación.

Reforzamiento de cotas

Esta técnica es particularmente importante cuando existen variables enteras de tipo general. La idea es aumentar las cotas inferiores y disminuir las superiores mediante inspección de las variables en las restricciones en que intervienen. Una consecuencia directa es la asignación de valores de algunas variables, cuando se llegan a igualar la cota inferior y la superior, y la detección de infactibilidad cuando se alcanza que la cota inferior sea mayor que la cota superior.

Reducción euclídea

Se utiliza especialmente en programación entera para evitar coeficientes no enteros en las restricciones y la función objetivo. Básicamente consiste en multiplicar la restricción (o función objetivo) por el menor valor que haga que todos los coeficientes sean enteros. Si bien en programación lineal no es de especial relevancia, en programación lineal entera puede ser fundamental la condición de que todos los coeficientes sean enteros, siendo incluso una hipótesis de trabajo en muchas ocasiones.

Asignación de variables

Consiste en identificar variables que pueden fijarse a una de sus cotas ya que los demás valores no pueden dar una solución factible y óptima.

Si un valor de una variable no puede satisfacer una restricción, aun cuando las demás variables tomen sus valores más favorables para intentar cumplirla, la variable debe fijarse al valor opuesto.

La idea básica es que si una variable sólo va en contra de la función objetivo y de las restricciones puede ser fijada a su cota inferior. Si es justo al revés, puede ser fijada a su cota superior.

El procedimiento para variables binarias y restricciones de tipo con cota positiva se describe a continuación. Se identifica la variable con el mayor coeficiente positivo. Si la suma de dicho coeficiente y cualquier coeficiente negativo excede la cota de la restricción, la variable debe fijarse a 0. Aquí se ven algunos ejemplos fáciles de asignación de variables para variables binarias.

$$\left.\begin{array}{llll} 3x_1 & & & \geq 2 \\ 3x_1 & +x_2 & & \geq 2 \\ 5x_1 & +x_2 & -2x_3 & \geq 2 \end{array}\right\} \Rightarrow x_1 = 1$$

$$x_1 + x_2 - 2x_3 \geq 1 \Rightarrow x_3 = 0$$

$$3x_1 + x_2 - 3x_3 \geq 2 \Rightarrow x_1 = 1, x_3 = 0$$

$$3x_1 - 2x_2 \leq -1 \Rightarrow x_1 = 0, x_2 = 1$$

El procedimiento se repetirá para las siguientes variables con el mayor coeficiente en una reacción en cadena.

$$3x_1 + x_2 - 2x_3 \geq 2 \Rightarrow x_1 = 1$$

De esta restricción se deduce que $x_1 = 1$, valor que se utiliza en la siguiente restricción

$$x_1 + x_4 + x_5 \leq 1 \Rightarrow x_4 = 0, x_5 = 0$$

De ésta se deduce que $x_4 = 0$ y $x_5 = 0$, que se utiliza en la siguiente restricción

$$-x_5 + x_6 \leq 0 \Rightarrow x_6 = 0$$

II.5.2. Preproceso mixto 0-1 y reformulación

Estas técnicas son específicas para problemas enteros y, en particular, con variables binarias. Dos formulaciones de un problema entero se dicen *0-1 equivalentes* si tienen las mismas soluciones enteras. Dadas dos formulaciones equivalentes de un problema entero, se dice que una es *más fuerte* que la otra, si la región factible de su relajación lineal está estrictamente contenida en la región factible de la otra. En la formulación más fuerte el intervalo de integralidad (*integrality gap*), es decir, la diferencia entre la función objetivo de la solución entera y la del problema relajado linealmente es menor. La formulación más fuerte ideal es aquélla cuyo intervalo de integralidad es nulo, es decir, cuya solución entera se puede obtener resolviendo un problema lineal. Sin embargo, formulaciones fuertes pueden requerir un número exponencial de restricciones.

El preproceso con variables enteras busca a partir de la formulación del problema encontrar otra formulación equivalente y más fuerte, es decir, que tenga las mismas soluciones enteras, pero menos soluciones continuas.

Difiere fundamentalmente de las técnicas anteriores en que en este preproceso es posible que las dimensiones del problema no sólo no se mantengan sino que aumenten en la búsqueda de esa formulación más fuerte, añadiendo nuevas restricciones al problema. Por esta razón también a veces se incluyen dentro de un concepto más amplio que es la reformulación de problemas.

Técnicas de mejora de coeficientes

El objetivo es reducir la región factible del problema lineal sin eliminar soluciones factibles del problema BIP, variando los coeficientes de las restricciones.

Dentro de estas técnicas hay dos variantes fundamentalmente: la reducción de coeficientes y el aumento de coeficientes.

Para reducir coeficientes la técnica más habitual es la denominada técnica *miope*. El procedimiento para restricciones de tipo $\leq$ como $a_1x_1 + a_2x_2 + \cdots + a_nx_n \leq b$ con variables binarias se presenta a continuación.

1) Calcular S como la suma de los valores a_j positivos

2) Elegir cualquier $a_j \neq 0$ tal que $S < b + |a_j|$

3) Si no existe dicho coeficiente implica que no se puede ajustar más la restricción

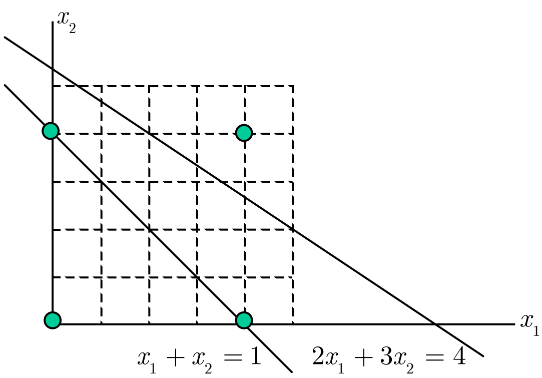
4) Si $a_j > 0$, se calculan $a'_j = S - b$ y $b' = S - a_j$, el nuevo valor del coeficiente es $a_j = a'_j$ y el de la cota $b = b'$

5) Si $a_j < 0$, se calcula $a'_j = S - b$, el nuevo valor del coeficiente es $a_j = a'_j$

6) Ir a 1.

Veamos la restricción $2x_1 + 3x_2 \leq 4$. La suma de los coeficientes positivos es $S = 2 + 3 = 5$. Se toma $a_1 \neq 0$, como $5 < 4 + 2$, los valores de $a'_1 = 5 - 4 = 1$ y $b' = 5 - 2 = 3$, luego $a_1 = 1$, $b = 3$ y, por consiguiente, la restricción queda $x_1 + 3x_2 \leq 3$. Para la nueva restricción la suma es $S = 1 + 3 = 4$. Se toma $a_2 \neq 0$, como $4 < 3 + 3$, los valores de $a'_2 = 4 - 3 = 1$ y $b' = 4 - 3 = 1$, luego $a_2 = 1$ y $b = 1$, entonces la restricción ajustada queda como $x_1 + x_2 \leq 1$.

Con esta nueva restricción se puede ver que el conjunto de puntos enteros factibles no ha cambiado, pero sí la región factible del problema relajado.



Este procedimiento puede ser más efectivo si se aprovecha información no sólo de la propia restricción sino también de otras restricciones que pueda haber en el problema.

En cuanto al aumento o “*lifting*” de coeficientes, consiste en aumentar los coeficientes de las variables en alguna restricción manteniendo invariante su cota. En general, el aumento más habitual consiste en la inclusión de variables que no aparecían en esa restricción. Es especialmente utilizada cuando se generan nuevas restricciones (ver un ejemplo en la sección siguiente) para sacar el mayor partido de ellas.

Generación de nuevas restricciones (planos de corte)

El concepto de plano de corte ya se ha explicado previamente lo que es. El plano de corte es una nueva restricción que reduce la región factible del problema lineal sin eliminar soluciones factibles del problema IP. Disminuye la diferencia entre función objetivo del LP y del MIP. Como ya se ha dicho existen diferentes tipos de planos de corte que pueden generarse en un problema de programación lineal entera mixta. Algunos son específicos para una clase de problemas y otros, aunque surgieron de la misma manera, pueden ser aprovechados en formulaciones generales. Por ejemplo, para el problema de la mochila se desarrollaron los planos de corte de tipo *cubrimiento*, pero dado que casi todos los problemas incluyen alguna restricción tipo mochila, con ellas se pueden generar este tipo de planos que serán válidos para el problema completo.

El procedimiento para generar planos de corte tipo cubrimiento es el siguiente:

1) Considerar cualquier restricción con variables binarias de tipo $\leq$ con todos los coeficientes no negativos (restricción mochila).

2) Encontrar un grupo de variables (cubrimiento minimal) tal que

- Se viola la restricción si las variables del cubrimiento son 1 y el resto son 0
- Se satisface la restricción si una de las variables del cubrimiento se hace 0 en lugar de 1

3) Formación del plano de corte como $\sum_{i\in\omega} x_i \leq N-1$, siendo N el número de variables del cubrimiento ω

Por ejemplo, para la restricción $6x_1 + 3x_2 + 5x_3 + 2x_4 \leq 9$ se pueden encontrar estos planos de corte $x_1 + x_2 + x_4 \leq 2$, y $x_1 + x_3 \leq 1$.

Estas desigualdades a su vez pueden ser reforzadas aumentando coeficientes.

Un plano de corte tipo cubrimiento se puede reforzar incluyendo en la restricción aquellas variables cuyo coeficiente sea mayor o igual que todos los del cubrimiento (este procedimiento ya está automatizado al calcular estos planos de corte). Para el ejemplo anterior, el primer y tercer plano no pueden ser reforzados ya que incluyen la variable de mayor coeficiente, pero el segundo sí, quedando su expresión $x_1 + x_2 + x_3 + x_4 \leq 2$. Esta nueva restricción es más fuerte que la segunda y que la primera, por lo que pueden ser sustituidas ambas por ella. Sin embargo, no es comparable con la tercera ya que los términos independientes son diferentes y los coeficientes de las variables también.

Otro tipo de planos de corte que se estudiaron fueron planos de corte en problemas de flujo con coste fijo, que ahora son utilizados en programación entera mixta.

En general, los cuatro tipos más utilizados (no específicos) son cortes GUB, *cliques*, cubrimientos (*covers*) y flujos (*flow covers*). Además se están retomando los planos de corte de Gomory. Todos estos planos pueden ser utilizados en el preproceso, o en los nodos del árbol de enumeración cuando se desarrolla la técnica de ramificación y corte.

II.6. Referencias

Minoux, M. (1986) *Mathematical Programming. Theory and Algorithms*. John Wiley and Sons

Nemhauser, G.L., Wolsey, L.A. (1988) *Integer and Combinatorial Optimization*. John Wiley and Sons.

Wolsey, L.A. (1998) *Integer Programming*. John Wiley and Sons.

II.7. Biblioteca de problemas

PROBLEMA 1

Resolver gráficamente y mediante el método ramificación y acotamiento los siguientes problemas.

1.

$$\begin{aligned}&\max 3x_1 + 2x_2\\&2x_1 + 2x_2 \leq 9\\&3x_1 + 3x_2 \geq 18\\&x_1, x_2 \geq 0 \text{ enteras}\end{aligned}$$

2.

$$\begin{aligned}&\max 2x_1 + 3x_2\\&5x_1 + 7x_2 \leq 35\\&4x_1 + 9x_2 \leq 36\\&x_1, x_2 \geq 0 \text{ enteras}\end{aligned}$$

3.

$$\begin{aligned}&\max x_1 + x_2\\&2x_1 + 5x_2 \leq 16\\&6x_1 + 5x_2 \leq 27\\&x_1, x_2 \geq 0 \text{ enteras}\end{aligned}$$

4.

$$\begin{aligned}&\min 5x_1 + 4x_2\\&3x_1 + 2x_2 \geq 5\\&2x_1 + 3x_2 \geq 7\\&x_1, x_2 \geq 0, x_2 \in \mathbb{Z}\end{aligned}$$

5.

$$\begin{aligned}&\max 5x_1 + 7x_2\\&2x_1 + x_2 \leq 13\\&5x_1 + 9x_2 \leq 41\\&x_1, x_2 \geq 0,\ x_2 \in \mathbb{Z}\end{aligned}$$

PROBLEMA 2

Resolver los siguientes problemas mediante el método ramificación y acotamiento.

1.
$$\begin{aligned}&\max 18x_1+14x_2+8x_3+4x_4\\&15x_1+12x_2+7x_3+4x_4+x_5\le 37\\&x_j\in\{0,1\}\ j=1,\ldots,5\end{aligned}$$

2.
$$\begin{aligned}&\max x_1+2x_2+5x_3\\&\left|-x_1+10x_2-3x_3\right|\ge 15\\&2x_1+x_2+x_3\le 10\\&x_1,x_2x_3\ge 0\end{aligned}$$

PROBLEMA 3

Dado el siguiente problema

$$\begin{aligned}&\max 3x_1+2x_2+2x_3+2x_4\\&3x_1+x_2+2x_3+x_4\le 5\\&x_1+2x_2+2x_3+x_4\le 3\\&x_1,x_2,x_3,x_4\in\{0,1\}\end{aligned}$$

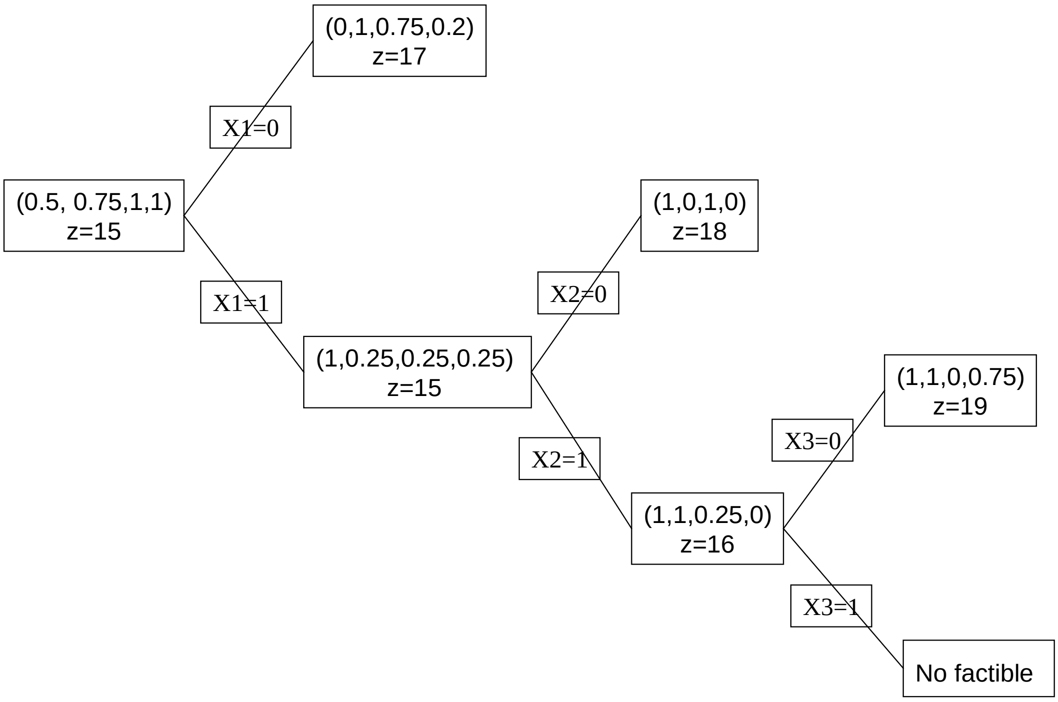
1) Resolver mediante técnicas de ramificación y acotamiento.

2) Obtener para cada desigualdad los cubrimientos minimales y los planos de corte obtenidos con ellos. Reforzar los cortes obtenidos.

3) Dada la solución de la relajación lineal $x^*=(1,0.5,0,1)$ ¿cuáles de las desigualdades anteriores se pueden utilizar para aislar esta solución?

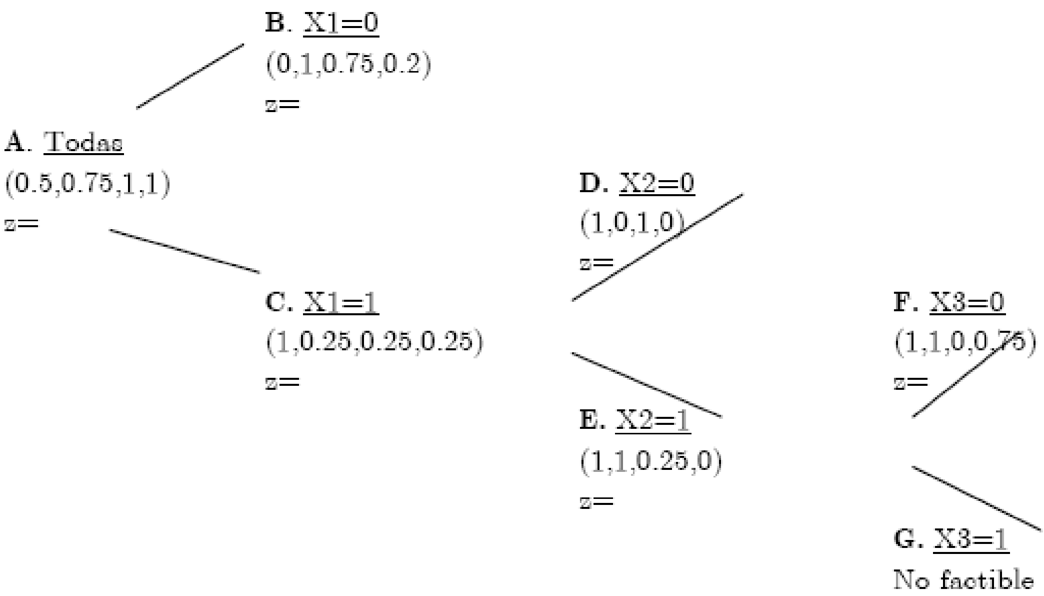
PROBLEMA 4

Dado el siguiente árbol incompleto de soluciones obtenido en un paso intermedio al resolver mediante ramificación y acotamiento un problema de programación lineal entera de minimización en que todas las variables son binarias, decir qué ramas se descarta seguir explorando y por qué, además de decir cuál es el mínimo valor que se puede obtener en el problema llegados a este punto y por qué.



PROBLEMA 5

Dado el siguiente árbol de soluciones obtenido al resolver mediante ramificación y acotación un problema de programación lineal entera de minimización en que todas las variables son binarias, dar valores a las cotas en cada nodo para que sea el árbol final obtenido. Explicar por qué es el árbol final y el orden en qué se han generado los nodos.



II.8. Resultados de la biblioteca de problemas

RESULTADO DEL PROBLEMA 1

1) No factible

2) $z^* = 14$, dos óptimos, $x_1^* = (4,2)$, $x_2^* = (7,0)$

3) $z^* = 4$, tres óptimos, $x_1^* = (2,2)$, $x_2^* = (3,1)$, $x_3^* = (4,0)$

4) $z^* = 10.5$, $x^* = (0.5,2)$

5) $z^* = 37$, dos óptimos, $x_1^* = (6,1)$, $x_2^* = (4.6,2)$

RESULTADO DEL PROBLEMA 2

1) $z^* = 40$, $x^* = (1,1,1,0,0)$

2) $z^* = 50$, $x^* = (0,0,10)$, $\delta^* = 1$

RESULTADO DEL PROBLEMA 3

1) $x_1 = 1$, $x_2 = 5$, objetivo 5

2) restricción 1 $C_{11} = \{1,2,3\}$, $C_{12} = \{1,3,4\}$

3) $x_1 + x_2 + x_3 \leq 2$, $x_1 + x_3 + x_4 \leq 2$

4) restricción 2 $C_{21} = \{2,3\}$, $C_{22} = \{1,2,4\}$, $C_{23} = \{1,3,4\}$

5) $x_2 + x_3 \leq 1$, $x_1 + x_2 + x_4 \leq 2$, $x_1 + x_3 + x_4 \leq 2$

6) $x_1 + x_2 + x_4 \leq 2$

RESULTADO DEL PROBLEMA 4

Nodos sin explorar o ramificar:

No factible: se poda ya que no hay solución factible

z=18: solución entera, luego z*=18

z=19: se poda ya que aunque no es entera, es peor que la mejor encontrada hasta el momento (minimización)

z=17: no es entera, la cota es mejor que la de la solución entera mejor encontrada hasta el momento, es factible: no se puede podar, hay que seguir ramificando.

Resto de nodos ya están explorados.

La mejor solución por el momento es z=18, pero podría llegar a ser 17 explorando el nodo que aún queda activo.

III. Optimización no lineal

De forma general, el problema de programación no lineal (*non linear programming* NLP) se plantea como

$$\begin{aligned} &\min_x f(x) \\ &g_i(x) \le b_i \quad i = 1,\ldots,m \end{aligned} \tag{1.60}$$

donde $x \in \mathbb{R}^n$, $f(x): \mathbb{R}^n \to \mathbb{R}$ y $g_i(x): \mathbb{R}^n \to \mathbb{R}$. Nótese que, en este caso, la no negatividad de las variables no forma parte de la formulación general. Aunque la función objetivo puede ser de maximización o minimización, en lo que sigue, cuando se hable de un problema de programación no lineal se supondrá, sin pérdida de generalidad, que es un problema de minimización.

III.1. Clasificación de problemas de programación no lineal

Los problemas de NLP se presentan de muchas formas distintas y no existe un único algoritmo o método para resolverlos. En su lugar, se han desarrollado algoritmos para unas clases o tipos especiales de NLP. A continuación, se presentan las clases más importantes y en secciones posteriores algunos de los algoritmos desarrollados.

La resolución de estos problemas es, en general, más difícil y computacionalmente más costosa que uno de LP de tamaño equivalente. Mientras que en LP es fácil encontrar optimizadores robustos, en NLP no existen métodos completamente robustos que solucionen cualquier tipo de problema. Por ello, en la formulación de un problema NLP es necesario seguir dos recomendaciones esenciales:

1) escalar adecuadamente las variables y restricciones del problema alrededor de 1, es decir, hacer que tanto las variables originales como las de holgura o exceso de las restricciones tomen en el óptimo valores cercanos a 1 y

2) proporcionar un punto inicial suficientemente cercano al óptimo. Un método para obtener este punto puede ser ejecutar el mismo problema linealizado.

Optimización no restringida

Son problemas de optimización donde no hay restricciones

$$\min_x f(x) \tag{1.61}$$

Para este tipo de problemas en la siguiente sección se verán cuáles son las condiciones que ha de cumplir un punto para que sea óptimo y algunos algoritmos de búsqueda para encontrar tal punto, diferenciando entre problemas de una variable (unidimensionales) o de más variables (multidimensionales) y entre algoritmos que usan diferenciación y otros que no.

Optimización linealmente restringida

Son problemas donde todas las restricciones son lineales, aunque la función objetivo no lo es

$$\begin{aligned} &\min_x f(x) \\ &Ax = b \end{aligned} \tag{1.62}$$

En este caso, el problema se simplifica bastante y existe una extensión del algoritmo simplex para resolverlo además de algunos casos particulares, como el del problema cuadrático, con algoritmos particulares muy eficientes.

Programación cuadrática

Son problemas donde las restricciones son lineales, pero la función objetivo es cuadrática, es decir, que incluye algún término con el cuadrado de alguna variable o con el producto de dos variables

$$\begin{aligned} &\min_x f(x) = \frac{1}{2} x^T Q x - b^T x \\ &Ax = b \end{aligned} \tag{1.63}$$

Se han desarrollado muchos algoritmos para este tipo de problemas, algunos con la hipótesis adicional de que $f(x)$ sea convexa (cóncava para el caso de maximización).

La programación cuadrática es muy importante, ya que surge de forma natural en muchas formulaciones, por ejemplo cuando la función objetivo representa distancias euclídeas (mínimos cuadrados). Además, una de sus utilidades es como problema auxiliar en otros tipos de optimización.

Programación convexa

La programación convexa engloba una amplia clase de problemas, entre los que se encuentran los casos anteriores cuando las funciones son convexas. Sus hipótesis son: $f(x)$ es convexa (cóncava si es maximización) y $g_i(x)$ es convexa, $\forall i = 1,\ldots,m$.

En tal caso, además, se puede asegurar que todo óptimo local es también global.

Programación separable

Es un caso especial de la programación convexa, donde se añade la hipótesis de que las funciones $f(x)$ y $g_i(x)$ son separables.

Una función separable es una función en la que cada término incluye una sola variable, por lo que la función se puede separar en una suma de funciones de las variables individuales

$$f(x) = \sum_{j=1}^{n} f_j(x_j) \tag{1.64}$$

Supone que se cumple la hipótesis de aditividad, pero no la de proporcionalidad de la programación lineal.

Programación no convexa

Incluye todos los problemas de programación no lineal que no cumplen la hipótesis de la programación convexa. En este caso, aun cuando se logre encontrar un óptimo local, no se puede asegurar que sea un óptimo global.

Para este tipo de problemas existen algoritmos, bastante adecuados en algunos casos, para encontrar óptimos locales. Para algunos casos especiales, además, hay algoritmos específicos que permiten resolver los problemas sin gran dificultad.

Programación geométrica

Surge a menudo en problemas de diseño de ingeniería, donde la función objetivo y las restricciones toman la forma

$$g(x) = \sum_{j=1}^{n} c_j P_j(x) \tag{1.65}$$

donde

$$P_j(x) = x_1^{a_{j1}} x_2^{a_{j2}} \ldots x_n^{a_{jn}}, \quad j = 1,\ldots,n \tag{1.66}$$

En tales casos, c_j y a_{jj} representan las constantes físicas y x_j las variables de diseño. Estas funciones, en general, no son ni cóncavas ni convexas, pero, existe un caso particular bien resuelto cuando los coeficientes c_j son estrictamente positivos.

Programación fraccional

Surge cuando la función objetivo se encuentra como una fracción

$$f(x) = \frac{f_1(x)}{f_2(x)} \tag{1.67}$$

Situación que se puede plantear, por ejemplo, al maximizar la razón de la producción entre las horas-hombre empleadas (productividad) o la ganancia entre el capital invertido (tasa de rendimiento), etc.

Existen algunos procedimientos especiales para ciertas formas de las funciones $f_1(x)$ y $f_2(x)$, aunque en general, se intenta transformar el problema en algún otro para el que se disponga de un procedimiento eficiente.

III.2. Expansión en serie de Taylor

Antes de empezar con los métodos de resolución de problemas de optimización vamos a revisar algunos conceptos básicos de cálculo numérico.

La expansión en serie de Taylor de una función nos permite calcular el valor de la función en puntos próximos a otro dado, cuyo valor ya se conoce. Se puede expresar como

$$f(x^* + p) = f(x^*) + \nabla f(x^*)^T p + \frac{1}{2} p^T \nabla^2 f(\xi) p \tag{1.68}$$

donde $x \in \mathbb{R}^n$, $p \in \mathbb{R}^n$ es un vector diferente de 0, ξ es un punto entre x y x^*, $f(x)$ es el valor de la función, $\nabla f(x)$ el *gradiente* y $\nabla^2 f(x)$ o $H_f(x)$ el *hessiano* o matriz hessiana, tal que $h_{ij} = \dfrac{\partial^2 f}{\partial x_i \partial x_j}(x^*)$. Ésta es una matriz simétrica por tener f segundas derivadas continuas.

Veamos el gradiente y el hessiano de una función cuadrática

$$f(x) = \frac{1}{2} x^T Q x - b^T x$$

donde $x \in \mathbb{R}^n$, $b \in \mathbb{R}^n$ y $Q \in \mathbb{R}^{n \times n}$.

El gradiente de una función cuadrática es

$$\nabla f(x) = Qx - b$$

y el hessiano

$$\nabla^2 f(x) = Q$$

Si tenemos la función

$$f(x_1, x_2) = \frac{1}{2} x_1^2 + 2x_1 x_2 + \frac{1}{2} x_2^2 - x_2 + 9,$$

el gradiente es

$$\nabla f(x_1, x_2) = \begin{pmatrix} \dfrac{\partial f}{\partial x_1} \\ \dfrac{\partial f}{\partial x_2} \end{pmatrix} = \begin{pmatrix} x_1 + 2x_2 \\ 2x_1 + x_2 - 1 \end{pmatrix}$$

y el hessiano

$$\nabla^2 f(x_1, x_2) = \begin{pmatrix} \dfrac{\partial^2 f}{\partial x_1^2} & \dfrac{\partial^2 f}{\partial x_1 \partial x_2} \\ \dfrac{\partial^2 f}{\partial x_2 \partial x_1} & \dfrac{\partial^2 f}{\partial x_2^2} \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$$

Si el punto $x^* = (1, -1)$ es un máximo local con valor de la función $f(1, -1) = 9$, el valor de la función en un nuevo punto $x = (1.1, -0.9)$ se puede calcular mediante la expansión en serie de Taylor hasta segundo orden, que es exacta para una función cuadrática, siendo $p = (0.1, 0.1)$

$$f(1.1, -0.9) = 9 + \begin{pmatrix} -1 & 0 \end{pmatrix} \begin{pmatrix} 0.1 \\ 0.1 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 0.1 & 0.1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 0.1 \\ 0.1 \end{pmatrix} = 8.93$$

o, como comprobación, por evaluación directa en el nuevo punto

$$f(1.1, -0.9) = 0.605 - 1.98 + 0.405 + 0.9 + 9 = 8.93$$

III.3. Optimización no restringida

La optimización no restringida es aquélla donde no hay restricciones

$$\min_x f(x) \tag{1.69}$$

En primer lugar se ven las condiciones necesarias y suficientes para que un punto sea mínimo local y para que lo sea global, bajo ciertas hipótesis. A continuación se ven diferentes algoritmos iterativos de cálculo de dicho mínimo.

III.3.1. Condiciones de optimalidad

Condición necesaria de optimalidad

Sea $f:\mathbb{R}^n \to \mathbb{R}$ diferenciable en x^*; si x^* es un *mínimo local*[10] entonces

$$\nabla f(x^*) = 0 \tag{1.70}$$

Ésta se denomina condición de primer orden porque involucra derivadas primeras.

Sin embargo, cualquier punto que satisfaga esta condición es un *punto estacionario*, no necesariamente un mínimo.

Ejemplo: $\nabla f(x_1,x_2) = \begin{pmatrix} x_1 + 2x_2 \\ 2x_1 + x_2 - 1 \end{pmatrix} = 0$ luego $(x_1,x_2) = (2/3,-1/3)$

Sea $f:\mathbb{R}^n \to \mathbb{R}$ dos veces diferenciable en x^*; si x^* es un *mínimo local* entonces $\nabla f(x^*) = 0$ y $H_f(x^*)$ es *semidefinida positiva*.

La condición de que el hessiano en un punto sea una matriz semidefinida positiva se denomina condición de segundo orden porque involucra derivadas segundas. Es equivalente a decir que la función es convexa en dicho punto, cóncava para el caso de maximización.

Una función $f(x)$ continua es *convexa* en el intervalo $[a,b]$ si $\forall u,v \in [a,b]$ y $\forall \lambda \in [0,1]$

$$f\big(\lambda u + (1-\lambda)v\big) \le \lambda f(u) + (1-\lambda) f(v)$$

Se dice que $f(x)$ es *cóncava* en el intervalo $[a,b]$ si $\forall u,v \in [a,b]$ y $\forall \lambda \in [0,1]$

[10] *Mínimo local* es un punto $x^* \in \mathbb{R}^n$ que satisface la condición $f(x^*) \le f(x)$ para todo x tal que $\|x - x^*\| < \varepsilon$ siendo ε un número positivo (típicamente pequeño) cuyo valor puede depender de x^*. Si $f(x^*) < f(x)$ se habla de *mínimo local estricto*.

$$f\left(\lambda u+(1-\lambda)v\right)\geq \lambda f(u)+(1-\lambda)f(v)$$

Una función convexa (cóncava) en $[a,b]$ tiene un valor mínimo (máximo) en dicho intervalo.

Una matriz A es semidefinida positiva si $x^T A x \geq 0$ para cualquier vector no nulo x o bien todos sus autovalores son no negativos o bien si todos los determinantes de orden $1,2,\ldots,n$ (siendo n la dimensión de la matriz) obtenidos añadiendo filas y columnas consecutivas desde el primer elemento son mayores o iguales que cero.

Una matriz A es semidefinida negativa si $x^T A x \leq 0$ para cualquier vector no nulo x o bien todos sus autovalores son no positivos o bien si los determinantes de orden $1,2,\ldots,n$ (siendo n la dimensión de la matriz) obtenidos añadiendo filas y columnas consecutivas desde el primer elemento son, alternativamente, menor que cero, mayor que cero, etc. o bien si la matriz $-A$ (resultado de cambiar el signo de todos los elementos) es semidefinida positiva.

Si una matriz tiene autovalores tanto positivos como negativos se dice que es indefinida.

Ejemplo: $\nabla^2 f(x_1,x_2)=\begin{pmatrix}2 & 1\\ 1 & 2\end{pmatrix}$ es una matriz semidefinida positiva ya que sus determinantes son positivos $|2|=2\geq 0$ y $\begin{vmatrix}2 & 1\\ 1 & 2\end{vmatrix}=3\geq 0$. Sin embargo, la matriz hessiana del ejemplo anterior, $\nabla^2 f(x_1,x_2)=\begin{pmatrix}1 & 2\\ 2 & 1\end{pmatrix}$, no lo es ya que el determinante de orden 2 de la matriz es estrictamente menor que cero.

Condición suficiente de optimalidad

Sea $f:\mathbb{R}^n\to\mathbb{R}$ dos veces diferenciable en x^*; si $\nabla f(x^*)=0$ y $H_f(x^*)$ es *definida positiva*, entonces x^* es *mínimo local estricto.*

Condición necesaria y suficiente de optimalidad

Sea $f:\mathbb{R}^n \to \mathbb{R}$ convexa y diferenciable en x^* (si es dos veces diferenciable el hessiano será semidefinido positivo); x^* es un *mínimo global* si y sólo si $\nabla f(x^*) = 0$.

III.3.2. Funciones unidimensionales

Sea el problema

$$\begin{array}{l} \min_x f(x) \\ x \in [a,b] \end{array} \tag{1.71}$$

siendo f una función convexa, $x \in \mathbb{R}$ y $[a,b]$ el intervalo de búsqueda del mínimo.

En principio, con las condiciones de optimalidad anteriores, un método para encontrar el óptimo de un problema sería resolver el sistema de ecuaciones resultante de igualar el gradiente de la función a 0. Sin embargo, no siempre es posible aplicar este método, algunas veces porque la función no es diferenciable y otras porque no se puede resolver directamente el sistema de ecuaciones, por lo que se presentan a continuación algunos algoritmos de búsqueda diseñados para localizar el mínimo de una función.

Al igual que otros procedimientos de búsqueda, consisten en encontrar una sucesión de puntos o soluciones prueba que conduzcan hacia una solución óptima. Hay diversos algoritmos para hacerlo, siendo la clasificación fundamental entre algoritmos de comparación (que no utilizan derivadas) y algoritmos de interpolación (que sí las usan).

Existen varios algoritmos de comparación: *método de la sección áurea*, *método de Fibonacci*, *método de la bisección*, etc. Entre los cuales se va a presentar el primero. Entre los algoritmos de interpolación se presenta el *método de Newton*.

Método de la sección áurea

La idea fundamental es ir haciendo intervalos cada vez menores donde se encuentre el mínimo, evaluando y comparando los valores de la función en puntos. El algoritmo acaba cuando la longitud del intervalo considerado en una iteración es menor que un determinado nivel de tolerancia.

Es un método que da una forma particular de seleccionar estos puntos, determinando esos intervalos cada vez menores donde se debe encontrar el mínimo buscado.

Sea I_k la longitud del intervalo $[a_k, b_k]$ en la iteración k. Sean λ_k, μ_k tales que $a_k < \lambda_k < \mu_k < b_k$. Entonces se pueden dar dos casos:

- Si $f(\lambda_k) > f(\mu_k)$, en la siguiente iteración se considera $a_{k+1} = \lambda_k$ y $b_{k+1} = b_k$
- Si $f(\lambda_k) \leq f(\mu_k)$, en la siguiente iteración se considera $a_{k+1} = a_k$ y $b_{k+1} = \mu_k$

Además para ahorrar una evaluación de la función f en la iteración $k+1$, se hacen coincidir los valores λ_{k+1} y μ_{k+1} con alguno de los anteriores. Así, en la situación a), se hace $\lambda_{k+1} = \mu_k$; y en la situación b), se hace $\mu_{k+1} = \lambda_k$.

En cualquier caso, es fácil comprobar que $I_k = b_k - a_k = I_{k+1} + I_{k+2}$. El método de la sección áurea además pretende que la reducción de la longitud de los intervalos sea constante, es decir, $I_{k+1} = \alpha I_k$, $0 < \alpha < 1$. Con estas dos condiciones se deduce que el valor de α que hay que manejar es la solución de la ecuación $\alpha^2 - \alpha + 1 = 0$, cuyo valor es $\alpha = \left(\sqrt{5} - 1\right)\big/2 \simeq 0.618$, conocido como número áureo. De esta forma, los valores de los puntos interiores del intervalo serán

$$\lambda_k = a_k + (1 - \alpha) I_k$$
$$\mu_k = a_k + \alpha I_k$$

La convergencia del algoritmo es lineal.

Método de Newton

Los algoritmos de interpolación realizan en cada iteración una aproximación de la función f, en el punto x_k considerado en dicha iteración, por un polinomio de segundo o de tercer grado. Para hacerlo necesita evaluar las primeras derivadas de la función en x_k.

Este método ajusta, en la iteración k, una parábola $q(x)$ a $f(x)$, $f'(x)$, $f''(x)$ y toma x_{k+1} como el vértice de esa parábola

$$q(x_{k+1}) = f(x_k) + f'(x_k)(x_{k+1} - x_k) + \frac{1}{2} f''(x_k)(x_{k+1} - x_k)^2 \tag{1.72}$$

$$q'(x_{k+1}) = 0 \Rightarrow x_{k+1} = x_k - \frac{f'(x_k)}{f''(x_k)} \tag{1.73}$$

El algoritmo converge cuadráticamente bajo ciertas condiciones, pero es muy inestable y suele ser necesario tomar precauciones e incluir protecciones.

III.3.3. Funciones multidimensionales

Sea el problema

$$\min_x f(x) \tag{1.74}$$

siendo f una función convexa y $x \in \mathbb{R}^n$.

En el caso de n variables, el problema de resolver el sistema de ecuaciones no lineales $\nabla f(x) = 0$ plantea mayores dificultades que el de resolver $f'(x) = 0$, por lo que se hace más necesario diseñar procedimientos alternativos de búsqueda del mínimo. Uno en particular es utilizar el método de Newton para resolver el sistema de ecuaciones. De hecho estos algoritmos utilizan aproximaciones similares.

También ahora, la clasificación de los algoritmos iterativos se hace igual que antes por algoritmos sin diferenciabilidad y algoritmos que utilizan diferenciación. Existen varios métodos que no requieren diferenciabilidad: mé*todo de Rosenbrock o de las coordenadas cíclicas*, *método de Hooke y Jeeves* (con paso continuo o paso discreto), *método de Nelder y Mead*, etc. De los cuales se presenta el primero. Existen varios métodos que requieren diferenciabilidad para minimizar funciones de varias variables. Se pueden clasificar en:

- Aquellos que sólo utilizan derivadas primeras: método del máximo descenso y método del gradiente conjugado
- Aquellos que además utilizan segundas derivadas: método de Newton y método de cuasi-Newton (con actualizaciones Broyden-Fletcher-Goldfarb-Shanno BFGS, Davidon-Fletcher-Powell DFP)

Los métodos que utilizan derivadas son procedimientos iterativos con un mecanismo de actualización del punto del tipo

$$x_{k+1} = x_k + \alpha_k p_k \tag{1.75}$$

Como se verá a continuación, las diferencias entre ellos se establecen por la selección de la dirección de búsqueda p_k y de la longitud del paso α_k. Existen condiciones generales o mejoras que se introducen para garantizar que convergen a un mínimo, como elegir el escalar α_k de manera que $f(x_{k+1}) < f(x_k)$ y que la dirección sea descendente $p_k^T \nabla f(x_k) < 0$. Existen dos métodos principales: el de *búsqueda unidimensional* (*line-search methods*) y los de *región de confianza* (*trust-region methods*).

Los métodos de búsqueda unidimensional controlan en las primeras iteraciones que la función objetivo desciende suficientemente. En las últimas iteraciones, próximas a la solución final, no afecta el movimiento de la solución. Se trata de encontrar el valor α_k que nos proporcione el mínimo de esta función mediante este problema de minimización unidimensional

$$\min_{\alpha_k > 0} F(\alpha) \equiv f(x_k + \alpha_k p_k) \tag{1.76}$$

Si el valor óptimo de α_k resulta calculable fácilmente y sin excesivo coste computacional, se utiliza éste. Alternativamente, se pueden imponer dos condiciones adicionales sobre α_k

- que el descenso sea suficiente

$$f(x_k + \alpha_k p_k) \le f(x_k) + \mu \alpha_k p_k^T \nabla f(x_k) \quad 0 < \mu < 1$$

que se denomina condición de Armijo

- que el descenso no sea demasiado pequeño

Por ejemplo, se define α_k como una secuencia 1, ½, ¼, 1/8, etc. Se van utilizando los valores de la secuencia empezando por 1. Si para $\alpha = 1$ se satisface la condición anterior, si no se utiliza el valor siguiente.

La dirección de búsqueda p_k debe cumplir dos condiciones adicionales

- que el descenso sea suficiente

$$-\frac{p^T \nabla f(x_k)}{\|p\| \cdot \|\nabla f(x_k)\|} \geq \varepsilon > 0$$

es decir, que los dos vectores no sean ortogonales

- que la dirección de movimiento esté relacionada con el gradiente

$$\|p\| \geq m \|\nabla f(x_k)\|$$

es decir, que la norma de la dirección de búsqueda no sea demasiado pequeña con respecto al gradiente.

Método de las coordenadas cíclicas

Consiste en partir de un punto x_1 y minimizar la función f en la dirección $d_1 = (1, 0, 0, \ldots, 0)$ (minimización de una función unidimensional); alcanzado el punto x_2 que minimiza la función en esa dirección, se minimiza desde ese punto en la dirección $d_2 = (0, 1, 0, \ldots, 0)$ para determinar el punto x_3 y así sucesivamente hasta llegar al punto x_{n+1} en que se vuelve a minimizar en la dirección d_1.

El proceso se repite hasta alcanzar la precisión deseada. No es un método muy eficiente, los otros métodos aprovechan mejor las direcciones detectadas de mejora, pero es bueno para hacerse una primera idea de lo que son métodos de búsqueda sin hipótesis de diferenciabilidad.

Método del máximo descenso

El método del máximo descenso (*steepest descent*) es el más sencillo de ellos y se basa en que la dirección de máximo descenso desde un punto es la dirección opuesta al gradiente.

$$p_k = -\nabla f(x_k) \tag{1.77}$$

Se puede interpretar como un caso particular del método de Newton (que se verá a continuación) cuando el hessiano se sustituye por la matriz identidad.

Así dado x_k se evalúa $p_k = -\nabla f(x_k)$ (que es una dirección ortogonal a la curva de nivel en ese punto) y se minimiza (minimización en una variable) en esa dirección para determinar el punto x_{k+1} en que se vuelve aplicar el procedimiento. El proceso se repite hasta que el valor del gradiente está por debajo de un cierto umbral.

No requiere el uso de segundas derivadas, ni resolver un sistema de ecuaciones lineales. Por tanto, es poco costoso computacionalmente. Tiene una convergencia más lenta, lineal. En general, no se debe utilizar.

Sea la función cuadrática

$$\min_x \; f(x) = \frac{1}{2} x^T Q x - b^T x$$

donde $Q = \begin{pmatrix} 1 & \cdot & \cdot \\ \cdot & 5 & \cdot \\ \cdot & \cdot & 25 \end{pmatrix}$ y $b = \begin{pmatrix} -1 \\ -1 \\ -1 \end{pmatrix}$

En este caso particular, el óptimo se puede determinar mediante este sistema de ecuaciones lineales $\nabla f(x) = Qx - b = 0$ y, por tanto, $x^* = Q^{-1}b = \begin{pmatrix} -1 \\ -1/5 \\ -1/25 \end{pmatrix}$ y $f(x^*) = -0.62$.

La dirección de máximo descenso es $p_k = -\nabla f(x_k) = -(Qx_k - b)$.

Si en este problema cuadrático se utiliza una búsqueda unidimensional exacta para determinar la longitud el paso el valor óptimo resultante de α_k es

$$\alpha_k = -\frac{\nabla f(x_k)^T p_k}{p_k^T \nabla^2 f(x_k) p_k} \tag{1.78}$$

Supongamos un punto inicial $x_0 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$. Para ese punto $f(x_0) = 0$, $\nabla f(x_0) = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ y si se utiliza la norma 2 del gradiente como medida de la convergencia $\|\nabla f(x_0)\| = \sqrt{1^2 + 1^2 + 1^2} = 1.73$. El valor de $\alpha_0 = 3/31 = 0.097$

La nueva estimación de la solución es $x_1 = x_0 + \alpha_0 p_0 = \begin{pmatrix} -0.097 \\ -0.097 \\ -0.097 \end{pmatrix}$. Para este punto $f(x_1) = -0.145$, $\nabla f(x_1) = \begin{pmatrix} 0.903 \\ 0.516 \\ -1.419 \end{pmatrix}$ y su norma 2 $\|\nabla f(x_1)\| = 1.76$. El valor de $\alpha_1 = 0.059$.

La nueva estimación de la solución es $x_2 = \begin{pmatrix} -0.150 \\ -0.127 \\ -0.013 \end{pmatrix}$. Para este punto $f(x_2) = -0.237$, $\nabla f(x_2) = \begin{pmatrix} 0.850 \\ 0.364 \\ 0.673 \end{pmatrix}$ y su norma 2 $\|\nabla f(x_2)\| = 1.14$.

El proceso continúa hasta que la norma 2 del gradiente se haga suficientemente pequeña (inferior a una cierta tolerancia, 10^{-8} por ejemplo). En este ejemplo se necesitan 216 iteraciones hasta alcanzar esta tolerancia.

La convergencia del método del máximo descenso para una función cuadrática con búsqueda unidimensional exacta es lineal. La relación de mejora entre dos iteraciones consecutivas se puede acotar superiormente de esta manera

$$\frac{f(x_{k+1}) - f(x^*)}{f(x_k) - f(x^*)} \leq \left(\frac{\operatorname{cond}(Q) - 1}{\operatorname{cond}(Q) + 1} \right)^2 \tag{1.79}$$

donde cond(A) el número de condición de la matriz A que se define como $\operatorname{cond}(A) \equiv \|A\| \cdot \|A^{-1}\|$ y $\|A\|_2 = \sqrt{\lambda_{\max}(A^T A)}$, siendo $\lambda_{\max}(A^T A)$ el autovalor máximo de la matriz $A^T A$. Si A es una matriz definida positiva y simétrica $\text{cond}(A) = \lambda_1 / \lambda_n$ siendo λ_1 y λ_n el mayor y menor autovalor respectivamente.

En el ejemplo anterior $\text{cond}(Q) = 25$. Un valor elevado del número de condición indica que la convergencia es muy lenta. Para $\text{cond}(Q) = 100$ este método mejora la solución como mucho del 4 % en cada iteración.

Para funciones no lineales generales la convergencia es también lineal con la misma cota superior, pero donde ahora $Q = \nabla^2 f(x^*)$ es el hessiano de la función en la solución.

Método de Newton

Es más eficiente que el método del máximo descenso y análogo al de una variable, ya que consiste en aproximar por una función cuadrática la función $f(x)$ en el punto considerado en una iteración, hallar el mínimo de esa función cuadrática y volver a aplicar el método

$$q(x_{k+1}) = f(x_k) + \nabla f(x_k)^T (x_{k+1} - x_k) + \frac{1}{2}(x_{k+1} - x_k)^T \nabla^2 f(x_k)(x_{k+1} - x_k) \qquad (1.80)$$

$$\nabla q(x_{k+1}) = 0 \Rightarrow x_{k+1} = x_k - \left[\nabla^2 f(x_k)\right]^{-1} \nabla f(x_k) \qquad (1.81)$$

Obsérvese que en este caso el cálculo de la dirección de movimiento en cada iteración implica la resolución de un sistema de ecuaciones lineales

$$\nabla^2 f(x_k) p_k = -\nabla f(x_k) \qquad (1.82)$$

El algoritmo converge cuadráticamente en la cercanía del óptimo bajo ciertas condiciones (relativamente exigentes) y requiere algunas modificaciones para su correcto y potente funcionamiento. Tiene la desventaja de que en cada iteración se necesita calcular y almacenar el hessiano.

Método de cuasi-Newton

El método de cuasi-Newton es una mejora del método de Newton que disminuye el coste computacional asociado a calcular y almacenar el hessiano y la resolución del sistema de ecuaciones lineales.

Está basado en aproximar el hessiano de la función en cada punto $\nabla^2 f(x_k)$ por otra matriz B_k definida positiva más fácil de calcular. Las diferentes variantes difieren en la elección de B_k y en su actualización en cada iteración.

Las ventajas estriban en utilizar sólo primeras derivadas en la aproximación de B_k y, por consiguiente, la dirección de búsqueda se puede calcular con menor coste computacional. Las desventajas son que la convergencia ya no es cuadrática y además requieren el almacenamiento de una matriz, luego no es apto para problemas de gran tamaño.

El algoritmo del método de cuasi-Newton es el siguiente:

1) Especificación de una solución inicial x_0 y de una aproximación inicial al hessiano B_0. La inicialización de B_k suele ser la matriz identidad $B_0 = I$

2) En cada iteración $k = 0,1,\ldots$ hasta encontrar la solución óptima

- Calcular la dirección de movimiento

$$B_k p_k = -\nabla f(x_k) \tag{1.83}$$

- Hacer una búsqueda unidimensional para determinar

$$x_{k+1} = x_k + \alpha_k p_k \tag{1.84}$$

- Calcular la matriz B_{k+1} mediante las ecuaciones

$$\begin{aligned} s_k &= x_{k+1} - x_k \\ y_k &= \nabla f(x_{k+1}) - \nabla f(x_k) \end{aligned} \tag{1.85}$$

El cálculo de la matriz B_k se basa en la condición de la secante. Para una función cuadrática el hessiano Q satisface de manera exacta esta condición.

$$\begin{gathered}\nabla^2 f(x_{k+1})(x_{k+1}-x_k) \approx \nabla f(x_{k+1}) - \nabla f(x_k) \\ B_{k+1}(x_{k+1}-x_k) \approx \nabla f(x_{k+1}) - \nabla f(x_k) \\ B_{k+1}s_k = y_k\end{gathered} \tag{1.86}$$

Sin embargo, la matriz B_{k+1} no se calcula directamente sino mediante actualización.

Actualizar la aproximación del hessiano

$$B_{k+1} = B_k + \text{actualización} \tag{1.87}$$

Existen diferentes actualizaciones de B_k. La denominada BFGS (Broyden-Fletcher-Goldfarb-Shanno)

$$B_{k+1} = B_k - \frac{(B_k s_k)(B_k s_k)^T}{s_k^T B_k s_k} + \frac{y_k y_k^T}{y_k^T s_k} \tag{1.88}$$

o la denominada DFP (Davidon-Fletcher-Powell)

$$B_{k+1} = B_k - \frac{(B_k s_k)(B_k s_k)^T}{s_k^T B_k s_k} + \frac{y_k y_k^T}{y_k^T s_k} + (s_k^T B_k s_k) u_k u_k^T \tag{1.89}$$

donde $u_k = \dfrac{y_k}{y_k^T s_k} - \dfrac{B_k s_k}{s_k^T B_k s_k}$

En ambas actualizaciones para garantizar que B_{k+1} sigue siendo definida positiva se debe cumplir $y_k^T s_k > 0$ condición que se debe garantizar controlando la búsqueda unidimensional.

III.4. Optimización restringida

Al igual que en el caso de la optimización no restringida, en primer lugar se van a ver las condiciones de optimalidad, que en muchos casos permiten determinar directamente la solución óptima del problema entre los puntos que son soluciones del sistema de ecuaciones que se plantea y, a continuación, algoritmos de búsqueda como alternativa al procedimiento anterior.

III.4.1. Lagrangiano y multiplicadores de Lagrange

Sea el problema de minimización de función objetivo no lineal con restricciones lineales de igualdad (P'')

$$\begin{aligned} &\min_x f(x) \\ &Ax = b \end{aligned} \tag{1.90}$$

donde $x \in \mathbb{R}^n$, $A \in \mathbb{R}^{m \times n}$ y $b \in \mathbb{R}^m$.

Se define la función no lineal denominada *lagrangiano* como

$$L(x,\lambda) = f(x) + \lambda^T (Ax - b) \tag{1.91}$$

donde $\lambda \in \mathbb{R}^m$ se denominan *multiplicadores* o *multiplicadores de Lagrange*. El mínimo del lagrangiano es el mínimo del problema (P'') puesto que en el óptimo $Ax = b$.

El gradiente del lagrangiano con respecto a cada conjunto de variables es

$$\begin{aligned} \nabla_x L(x,\lambda) &= \nabla f(x) + A^T \lambda \\ \nabla_\lambda L(x,\lambda) &= Ax - b \end{aligned} \tag{1.92}$$

Aplicando las condiciones necesarias y suficientes de optimalidad al lagrangiano tenemos

$$\nabla L(x^*, \lambda^*) = 0 \tag{1.93}$$

o, expresado de otra manera, si x^* es un mínimo local debe verificar

$$\nabla f(x^*) = -A^T \lambda^* \tag{1.94}$$

Interpretando esta ecuación se observa que en un mínimo local el gradiente de la función objetivo es una combinación lineal de los gradientes de las restricciones y los multiplicadores de Lagrange son los pesos. Los multiplicadores representan el cambio en la función objetivo para un cambio unitario (marginal) en la cota de cada restricción. En el caso particular de programación lineal éstos recibían el nombre de *variables duales* o *precios sombra*. Con esta formulación del lagrangiano los multiplicadores resultan con signo contrario a las variables duales.

Veamos ahora el problema de minimización de función objetivo no lineal con restricciones no lineales de igualdad y desigualdad (P''')

$$\begin{aligned} &\min_x f(x) \\ &g_i(x) \leq 0 \quad i = 1,\dots,m \\ &h_j(x) = 0 \quad j = 1,\dots,l \\ &x \in X \end{aligned} \tag{1.95}$$

donde X es un conjunto abierto no vacío en $\mathbb{R}^n$, $f:\mathbb{R}^n \to \mathbb{R}$ y $g_i:\mathbb{R}^n \to \mathbb{R}$.

Definimos el lagrangiano como

$$L(x,\lambda,\mu) = f(x) + \sum_{i=1}^{m} \lambda_i g_i(x) + \sum_{j=1}^{l} \mu_j h_j(x) \tag{1.96}$$

donde $\lambda \in \mathbb{R}^m$, $\mu \in \mathbb{R}^l$ son los *multiplicadores de Lagrange*. Esta función es siempre una cota inferior de $f(x)$ para valores factibles de x y valores conocidos de $\lambda \geq 0$ (no negativos) y μ (libre). Por consiguiente, interesa obtener la mayor cota inferior $\max\limits_{\substack{\lambda \geq 0 \\ \mu}} L(x,\lambda,\mu)$ y el problema original (P''') será

$$\begin{aligned} &\min_x \left\{ \max_{\substack{\lambda \geq 0 \\ \mu}} L(x,\lambda,\mu) \right\} \\ &x \in X \end{aligned} \tag{1.97}$$

De esta forma hemos transformado un problema de minimización con restricciones en otro sin restricciones. Este procedimiento se denomina *relajación lagrangiana*.

III.4.2. Condiciones de optimalidad de Karush-Kuhn-Tucker

Estas condiciones están basadas en unas previas, denominadas de Fritz John pero menos útiles para resolver problemas. Fueron enunciadas por Karush en su tesis de máster y publicadas en ruso, pero el hecho de que fueran presentadas en este idioma dificultó su expansión y, unos años después, Kuhn y Tucker (Kuhn, 1951) las enunciaron y demostraron de nuevo (sin conocer el estudio previo en ruso), por lo que durante mucho tiempo se conocieron como las condiciones de Kuhn-Tucker, hasta que posteriormente se incluyó a la primera persona que realmente las publicó.

Históricamente, se enunciaron primero para problemas con restricciones de desigualdad y luego para problemas con restricciones de desigualdad e igualdad, por lo que se va a mantener este orden.

Condiciones necesarias con restricciones de desigualdad

Sea el problema (P)

$$\begin{aligned}&\min_x f(x)\\ &g_i(x) \le 0 \quad i = 1,...,m\\ &x \in X\end{aligned} \tag{1.98}$$

donde X es un conjunto abierto no vacío en $\mathbb{R}^n$, $f:\mathbb{R}^n \to \mathbb{R}$ y $g_i:\mathbb{R}^n \to \mathbb{R}$.

Sea x^* un punto factible (es decir, debe satisfacer las restricciones); $I = \left\{ i / g_i(x^*) = 0 \right\}$; f y $\left\{ g_i, i \in I \right\}$ diferenciables en x^*; $\left\{ g_i, i \notin I \right\}$ continuas en x^* y sean $\left\{ \nabla g_i(x^*) \right\}_{i \in I}$ linealmente independientes. Si x^* es mínimo local de (P) existen unos escalares $\left\{ \lambda_i, i \in I \right\}$ tales que

$$\begin{aligned}&\nabla f(x^*) + \sum_{i \in I} \lambda_i \nabla g_i(x^*) = 0\\ &\lambda_i \ge 0 \quad \forall i \in I\end{aligned} \tag{1.99}$$

si además las funciones $\left\{ g_i, i \notin I \right\}$ son diferenciables en x^* entonces

$$
\begin{aligned}
&\nabla f(x^*) + \sum_{i=1}^{m} \lambda_i \nabla g_i(x^*) = 0 \\
&\lambda_i g_i(x^*) = 0 \quad i = 1,\ldots,m \\
&\lambda_i \geq 0 \quad i = 1,\ldots,m
\end{aligned}
\tag{1.100}
$$

Veamos un ejemplo

$$
\begin{aligned}
&\min_{x,y} \ f(x,y) = 9y - (x-5)^2 \\
&-x^2 + y \leq 0 \\
&-x - y \leq 0 \\
&x - 1 \leq 0
\end{aligned}
$$

Las condiciones de KKT serían este sistema de ecuaciones no lineales

$$
\begin{aligned}
&-2(x-5) - 2x\lambda_1 - \lambda_2 + \lambda_3 = 0 \\
&9 + \lambda_1 - \lambda_2 = 0 \\
&\lambda_1(-x^2 + y) = 0 \\
&\lambda_2(-x - y) = 0 \\
&\lambda_3(x-1) = 0 \\
&\lambda_1, \lambda_2, \lambda_3 \geq 0
\end{aligned}
$$

además se deben considerar todas las restricciones del problema puesto que el punto ha de ser factible.

Vamos a resolverlo analíticamente enumerando todas las combinaciones posibles ($=0$ o $\neq 0$) de valores de los escalares λ_i. Obsérvese que si $\lambda_2 = 0$ entonces $\lambda_1 = -9$ (valor infactible), luego $\lambda_2 \neq 0$ y, por tanto, $-x = y$.

Ahora, consideremos dos casos:

- $\lambda_1 = 0$: entonces $\lambda_2 = 9$ y el sistema que queda es

$$
\begin{aligned}
&-2x + 10 - 9 + \lambda_3 = 0 \\
&\lambda_3(x-1) = 0
\end{aligned}
$$

Ahora se tiene:

- o Si $\lambda_3 = 0$, entonces $x = 1/2$ e $y = -1/2$. Entonces $A = (1/2, -1/2)$ y $\lambda_A = (0, 9, 0)$.

  - Si $x=1$, entonces $y=-1$ y resulta $\lambda_3=1$. Entonces $B=(1,-1)$ y $\lambda_B=\left(0,9,1\right)$.

- $\lambda_1 \neq 0$: entonces se tienen $x^2=y$ y $x=-y$, con lo que $x^2=-x$, que da dos posibilidades
  - $x=0$, que implica $y=0$ y al resolver se tiene $C=(0,0)$ y $\lambda_C=(1,10,0)$
  - $x=-1$, que implica $y=1$ y al resolver se tiene $D=(1,-1)$ y $\lambda_D=\left(-3,6,0\right)$

Así, el punto D no es un punto que cumpla las condiciones de KKT (los multiplicadores no son mayores o iguales que 0). Los otros tres puntos las cumplen, pero gráficamente se puede observar que B y C son mínimos locales (ambos con valor –25 de la función objetivo), A no es nada y no existe mínimo global. De hecho, la función no está acotada.

Si se hubiera deseado también buscar los máximos locales, basta buscar puntos con las mismas condiciones pero en que todos los multiplicadores sean menores o iguales que 0 y se habría obtenido el punto $E=(1,1)$ con $\lambda_E=(-9,0,-26)$ que es un máximo local.

Condiciones suficientes con restricciones de desigualdad

Sea x^* un punto factible del problema (P); $I=\left\{i\,/\,g_i(x^*)=0\right\}$; f y $\left\{g_i, i\in I\right\}$ convexas y diferenciables en x^*. Si existen escalares $\left\{\lambda_i, i\in I\right\}$ tales que

$$\begin{aligned}&\nabla f(x^*)+\sum_{i\in I}\lambda_i\nabla g_i(x^*)=0\\&\lambda_i\geq 0\quad \forall i\in I\end{aligned}\qquad(1.101)$$

entonces x^* es *mínimo global* de (P).

Alternativamente, en lugar de poner la condición de que f y $\left\{g_i, i\in I\right\}$ sean convexas y diferenciables en x^* se puede expresar también como que el *lagrangiano restringido* $L(x)=f(x)+\sum_{i\in I}\lambda_i^* g_i(x)$, siendo λ_i^* los multiplicadores de Lagrange de las restricciones, tenga un hessiano, $\nabla^2 L(x^*) = \nabla^2 f(x^*) + \sum_{i \in I} \lambda_i^* \nabla^2 g_i(x^*)$, que sea una matriz semidefinida positiva en x^*.

Las condiciones suficientes para el caso de maximización se traducen en que la función f sea cóncava en el punto, las restricciones no cambian y los multiplicadores sean menores o iguales que 0.

Condiciones necesarias con restricciones de desigualdad e igualdad

Sea el problema (P')

$$\begin{aligned} & \min_x f(x) \\ & g_i(x) \le 0 \quad i = 1,\dots,m \\ & h_j(x) = 0 \quad j = 1,\dots,l \\ & x \in X \end{aligned} \tag{1.102}$$

donde X es un conjunto abierto no vacío en $\mathbb{R}^n$, $g_i : \mathbb{R}^n \to \mathbb{R}$ y $h_j : \mathbb{R}^n \to \mathbb{R}$.

Sea x^* un punto factible; $I = \{ i / g_i(x^*) = 0 \}$; f y $\{ g_i, i \in I \}$ diferenciables en x^*; $\{ g_i, i \notin I \}$ continuas en x^*; $\{ h_j, j = 1,\dots,l \}$ continuamente diferenciables en x^* y sean $\{ \nabla g_i(x^*), i \in I; \nabla h_j(x^*), j = 1,\dots,l \}$ linealmente independientes. Si x^* es mínimo local de (P') existen unos escalares $\{ \lambda_i, i \in I; \mu_j, j = 1,\ldots,l \}$ tales que

$$\begin{aligned} & \nabla f(x^*) + \sum_{i \in I} \lambda_i \nabla g_i(x^*) + \sum_{j=1}^{l} \mu_j \nabla h_j(x^*) = 0 \\ & \lambda_i \ge 0 \quad \forall i \in I \end{aligned} \tag{1.103}$$

si además las funciones $\{ g_i, i \notin I \}$ son diferenciables en x^* entonces

$$\begin{aligned} & \nabla f(x^*) + \sum_{i=1}^{m} \lambda_i \nabla g_i(x^*) + \sum_{j=1}^{l} \mu_j \nabla h_j(x^*) = 0 \\ & \lambda_i g_i(x^*) = 0 \quad i = 1,\dots,m \\ & \lambda_i \ge 0 \quad i = 1,\dots,m \end{aligned} \tag{1.104}$$

Condiciones suficientes con restricciones de desigualdad e igualdad

Sea x^* un punto factible del problema (P'); $I = \{i / g_i(x^*) = 0\}$; f y $\{g_i, i \in I\}$ convexas y diferenciables en x^*. Si existen unos escalares $\{\lambda_i, i \in I; \mu_j, j = 1, \ldots, l\}$ tales que

$$\begin{aligned} &\nabla f(x^*) + \sum_{i \in I} \lambda_i \nabla g_i(x^*) + \sum_{j=1}^{l} \mu_j \nabla h_j(x^*) = 0 \\ &\lambda_i \geq 0 \quad \forall i \in I \end{aligned} \tag{1.105}$$

de modo que h_j sea convexa en x^* si $\mu_j > 0$ o h_j sea cóncava en x^* si $\mu_j < 0$, entonces x^* es *mínimo global* de (P').

III.4.3. Transformación de un problema no lineal con restricciones lineales de igualdad en uno sin restricciones

Sea el problema de minimización (P''). Supongamos un punto factible $\bar{x}$ y una dirección factible p que debe cumplir $Ap = 0$, entonces cualquier punto factible se puede expresar como

$$x = \bar{x} + p \tag{1.106}$$

Si Z es una matriz $\mathbb{R}^{n \times r}$ (siendo $r \geq n - m$) del espacio nulo de A (espacio definido por el conjunto de vectores ortogonales a las filas de A, representa el conjunto de direcciones factibles de las restricciones) la región factible también se puede expresar como

$$x = \bar{x} + Zv \tag{1.107}$$

donde $v \in \mathbb{R}^r$.

De esta forma el problema se transforma en una minimización de la función $\phi(v)$ que es la *función reducida* (proyectada) de f en la región factible, que es un problema de optimización sin restricciones y de menores dimensiones. Se ha pasado de un problema de n dimensiones a uno de $n-m$.

$$\min_{v} \phi(v) = f(\overline{x} + Zv) \tag{1.108}$$

El gradiente reducido (o proyectado) de f en x es

$$\nabla \phi(v) = Z^T \nabla f(x) \tag{1.109}$$

El hessiano reducido de f en x es

$$\nabla^2 \phi(v) = Z^T \nabla^2 f(x) Z \tag{1.110}$$

Para el cálculo de la matriz de transformación Z existen dos posibles métodos:

1) Método de reducción de variables

$$Ap = \begin{pmatrix} B & N \end{pmatrix} \begin{pmatrix} p_B \\ p_N \end{pmatrix} = Bp_B + Np_N = 0$$

$$p = \begin{pmatrix} p_B \\ p_N \end{pmatrix} = \begin{pmatrix} -B^{-1}N \\ I \end{pmatrix} p_N$$

$$Z = \begin{pmatrix} -B^{-1}N \\ I \end{pmatrix} \tag{1.111}$$

$$x = \overline{x} + p = \overline{x} + Zp_N = \begin{pmatrix} B^{-1}b \\ 0 \end{pmatrix} + \begin{pmatrix} -B^{-1}N \\ I \end{pmatrix} p_N$$

donde $A \in \mathbb{R}^{m \times n}$, $B \in \mathbb{R}^{m \times m}$, $N \in \mathbb{R}^{m \times n-m}$, $Z \in \mathbb{R}^{n \times n-m}$

2) Factorización QR

$$A^T = QR$$

$$Q = \begin{pmatrix} Q_1 & Q_2 \end{pmatrix}$$

$$\begin{aligned} AQ &= R^T \\ AQ_1 &= R_1^T \\ AQ_2 &= 0 \end{aligned}$$

donde $A \in \mathbb{R}^{m \times n}$, $Z \in \mathbb{R}^{n \times n-m}$, $Q \in \mathbb{R}^{n \times n}$, $Q_1 \in \mathbb{R}^{n \times m}$, $Q_2 \in \mathbb{R}^{n \times n-m}$, $R \in \mathbb{R}^{n \times m}$ y $R_1 \in \mathbb{R}^{m \times m}$ (es una submatriz triangular superior de R)

$$Z = Q_2 \tag{1.112}$$

es la base ortogonal ya que $Z^T Z = I$.

Este método tiene la ventaja de que el cálculo de Z es numéricamente estable pero su coste computacional es elevado y la matriz Z es densa aunque la matriz A sea cuasivacía.

Condiciones de optimalidad con restricciones lineales de igualdad

Las condiciones *necesarias*

$$Z^T \nabla f(x^*) = 0 \tag{1.113}$$

$$p^T \nabla^2 f(x^*) p \geq 0 \tag{1.114}$$

siendo $p = Zv$ un vector del espacio nulo de A.

Las condiciones *suficientes* para *mínimo local estricto*

$$Z^T \nabla f(x^*) = 0 \tag{1.115}$$

siendo el hessiano reducido $Z^T \nabla^2 f(x^*) Z$ en el punto *definida* positiva.

Veamos un ejemplo

$$\begin{aligned} &\min f(x) = x_1^2 + 2x_1 + x_2^2 - x_3^2 + 4x_3 \\ &x_1 - x_2 + 2x_3 = 2 \end{aligned}$$

El gradiente y hessiano de la función son

$$\nabla f(x) = \begin{pmatrix} 2x_1 - 2 \\ 2x_2 \\ -2x_3 + 4 \end{pmatrix}, \ \nabla^2 f(x) = \begin{pmatrix} 2 & \cdot & \cdot \\ \cdot & 2 & \cdot \\ \cdot & \cdot & -2 \end{pmatrix}$$

Sea el punto $x^* = \begin{pmatrix} 2.5 \\ -1.5 \\ -1 \end{pmatrix}$ y la matriz de transformación $Z = \begin{pmatrix} 1 & -2 \\ 1 & \cdot \\ \cdot & 1 \end{pmatrix}$. Entonces el nuevo punto y el valor de la función a minimizar queda

$$x = \bar{x} + Zv = \begin{pmatrix} 2.5 \\ -1.5 \\ -1 \end{pmatrix} + \begin{pmatrix} 1 & -2 \\ 1 & \cdot \\ \cdot & 1 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 2.5 + v_1 - 2v_2 \\ -1.5 + v_1 \\ -1 + v_2 \end{pmatrix}$$

$$\begin{aligned} \min_{v_1, v_2} \phi(v) = & \left(2.5 + v_1 - 2v_2\right)^2 + 2\left(2.5 + v_1 - 2v_2\right) + \\ & + \left(-1.5 + v_1\right)^2 - \left(-1 + v_2\right)^2 + 4\left(-1 + v_2\right) \end{aligned}$$

Aplicamos las condiciones necesarias y suficientes de primer y segundo orden a la función reducida

$$Z^T \nabla f(x^*) = \begin{pmatrix} 1 & 1 & \cdot \\ -2 & \cdot & 1 \end{pmatrix} \begin{pmatrix} 3 \\ -3 \\ 6 \end{pmatrix} = \begin{pmatrix} \cdot \\ \cdot \end{pmatrix}$$

$$Z^T \nabla^2 f(x^*) Z = \begin{pmatrix} 1 & 1 & \cdot \\ -2 & \cdot & 1 \end{pmatrix} \begin{pmatrix} 2 & \cdot & \cdot \\ \cdot & 2 & \cdot \\ \cdot & \cdot & -2 \end{pmatrix} \begin{pmatrix} 1 & -2 \\ 1 & \cdot \\ \cdot & 1 \end{pmatrix} = \begin{pmatrix} 2 & 2 & \cdot \\ -4 & \cdot & -2 \end{pmatrix} \begin{pmatrix} 1 & -2 \\ 1 & \cdot \\ \cdot & 1 \end{pmatrix} = \begin{pmatrix} 4 & -4 \\ -4 & 6 \end{pmatrix}$$

El hessiano de la función no es una matriz definida positiva y, sin embargo, el hessiano reducido sí lo es. Luego el punto es un mínimo local estricto.

Condiciones de optimalidad con restricciones lineales de desigualdad

Sea el problema de minimización no lineal con restricciones lineales de desigualdad (P''')

$$\begin{aligned} & \min_x f(x) \\ & Ax \ge b \end{aligned} \tag{1.116}$$

donde $x \in \mathbb{R}^n$, $A \in \mathbb{R}^{m \times n}$ y $b \in \mathbb{R}^m$.

Las condiciones *necesarias* de optimalidad son

$$\begin{aligned} &\nabla f(x^*) = -A^T \lambda^* \\ &\lambda^{*T}(Ax^* - b) = 0 \end{aligned} \tag{1.117}$$

o equivalentemente

$$\begin{aligned} &Z^T \nabla f(x^*) = 0 \\ &\lambda^{*T}(Ax^* - b) = 0 \end{aligned} \tag{1.118}$$

donde $\lambda^* \le 0$. El último conjunto de ecuaciones corresponde a las condiciones de complementariedad de holguras (o bien la restricción se cumple o bien el multiplicador es 0 pero no los dos). Adicionalmente $Z^T \nabla^2 f(x^*) Z$ ha de ser *semidefinida* positiva, siendo Z una matriz del espacio nulo del conjunto de restricciones activas de A.

Las condiciones *suficientes* para que sea un *mínimo local estricto* son

$$\begin{aligned} &Z^T \nabla f(x^*) = 0 \\ &\lambda^{*T}(Ax^* - b) = 0 \end{aligned} \tag{1.119}$$

y además $Z^T \nabla^2 f(x^*) Z$ ha de ser *definida* positiva.

Condiciones de optimalidad con restricciones no lineales de igualdad

Sea el problema (P)

$$\begin{aligned} &\min_x f(x) \\ &g_i(x) = 0 \quad i = 1, \ldots, m \end{aligned} \tag{1.120}$$

donde $x \in \mathbb{R}^n$ y $g_i : \mathbb{R}^n \to \mathbb{R}$. El lagrangiano de este problema es

$$L(x, \lambda) = f(x) + \lambda^T g(x) \tag{1.121}$$

Las condiciones *necesarias* de primer y segundo orden

$$Z(x^*)^T \nabla f(x^*) = 0 \tag{1.122}$$

y además $Z(x^*)^T \nabla_{xx}^2 L(x^*, \lambda^*) Z(x^*)$ ha de ser una matriz *semidefinida* positiva, donde $Z(x^*)$ es una matriz del espacio nulo del jacobiano $\nabla g(x^*)^T$.

Las condiciones *suficientes* de primer y segundo orden para *mínimo local estricto*

$$Z(x^*)^T \nabla f(x^*) = 0 \tag{1.123}$$

y además $Z_+(x^*)^T \nabla^2_{xx} L(x^*, \lambda^*) Z_+(x^*)$ ha de ser *definida* positiva. Z es una matriz del espacio nulo del conjunto de restricciones activas y no degeneradas del jacobiano $\nabla g(x^*)^T$.

Condiciones de optimalidad con restricciones no lineales de desigualdad

Sea el problema (P)

$$\begin{aligned} &\min_x f(x) \\ &g_i(x) \ge 0 \quad i = 1, \dots, m \end{aligned} \tag{1.124}$$

donde $x \in \mathbb{R}^n$ y $g_i : \mathbb{R}^n \to \mathbb{R}$. El lagrangiano de este problema es

$$L(x, \lambda) = f(x) + \lambda^T g(x) \tag{1.125}$$

Las condiciones *necesarias* de primer orden

$$\begin{aligned} &\nabla_x L(x^*, \lambda^*) = 0 \\ &\lambda^{*T} g(x^*) = 0 \end{aligned} \tag{1.126}$$

o equivalentemente

$$\begin{aligned} &Z(x^*)^T \nabla f(x^*) = 0 \\ &\lambda^{*T} g(x^*) = 0 \end{aligned} \tag{1.127}$$

siendo $\lambda^* \le 0$. El último subconjunto de ecuaciones son las condiciones de complementariedad de holguras. Además la condición de segundo orden indica que ha de ser una matriz *semidefinida* positiva, donde $Z(x^*)$ es una matriz del espacio nulo del jacobiano de las restricciones activas de $\nabla g(x^*)^T$.

Las condiciones *suficientes* de optimalidad para *mínimo local estricto*

$$\begin{aligned} &Z(x^*)^T \nabla f(x^*) = 0 \\ &\lambda^{*T} g(x^*) = 0 \end{aligned} \tag{1.128}$$

y además $Z_+(x^*)^T \nabla^2_{xx} L(x^*, \lambda^*) Z_+(x^*)$ ha de ser *definida* positiva, donde $Z_+(x^*)$ es una matriz del espacio nulo del jacobiano de las restricciones activas y no degeneradas de $\nabla g(x^*)^T$.

Veamos un ejemplo

$$\begin{aligned} &\min f(x) = x_1 \\ &(x_1+1)^2 + x_2^2 \geq 1 \\ &x_1^2 + x_2^2 \leq 2 \end{aligned}$$

El lagrangiano, el gradiente y el hessiano son

$$L(x,\lambda) = x_1 - \lambda_1((x_1+1)^2 + x_2^2 - 1) + \lambda_2(x_1^2 + x_2^2 - 2)$$

$$\nabla_x L(x,\lambda) = \begin{pmatrix} 1 - 2\lambda_1(x_1+1) + 2\lambda_2 x_1 \\ -2\lambda_1 x_2 + 2\lambda_2 x_2 \end{pmatrix}$$

$$\nabla^2_{xx} L(x,\lambda) = \begin{pmatrix} -2(\lambda_1+\lambda_2) & \cdot \\ \cdot & -2(\lambda_1+\lambda_2) \end{pmatrix}$$

a) Probar si $x = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ es óptimo.

Como la restricción 2 no es activa $\lambda_2 = 0$. El valor de λ_1 se obtiene al establecer la condición necesaria de primer orden

$$\nabla_x L(x,\lambda) = \begin{pmatrix} 1-2\lambda_1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Luego $\lambda_1 = 0.5$

Si se toma $Z = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ como la matriz base del espacio nulo del jacobiano de las restricciones activas (sólo la primera) $\nabla g(x^*)^T = \begin{pmatrix} 2 & 0 \end{pmatrix}$

$$Z_+(x_*)^T \nabla^2_{xx} L(x_*, \lambda_*) Z_+(x_*) = \begin{pmatrix} \cdot & 1 \end{pmatrix} \begin{pmatrix} -1 & \cdot \\ \cdot & -1 \end{pmatrix} \begin{pmatrix} \cdot \\ 1 \end{pmatrix} = -1$$

luego $x = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ no es óptimo.

b) Probar si $x = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$ es óptimo.

El valor de λ_1 y λ_2 se obtiene al establecer la condición necesaria de primer orden

$$\nabla_x L(x,\lambda) = \begin{pmatrix} 1-2\lambda_2 \\ 2\lambda_1 - 2\lambda_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Luego $\lambda_1 = \lambda_2 = 0.5$

La matriz base del espacio nulo del jacobiano de las restricciones activas es vacía. Luego la condición suficiente de segundo orden se satisface, luego el punto es un mínimo local estricto.

c) Probar si $x = \begin{pmatrix} 0 \\ \sqrt{2} \end{pmatrix}$ es óptimo.

Como la restricción 1 no es activa $\lambda_1 = 0$. El valor de λ_2 se obtiene al establecer la condición necesaria de primer orden

$$\nabla_x L(x,\lambda) = \begin{pmatrix} 1+2\lambda_2 0 \\ 2\lambda_2\sqrt{2} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

El sistema es inconsistente, no se puede obtener el valor de λ_2. Luego no se puede satisfacer la condición necesaria de primer orden.

III.4.4. Métodos de programación no lineal (NLP)

Son procedimientos algorítmicos de búsqueda que se detienen cuando encuentran un óptimo local y que suelen ser aplicados partiendo de distintos puntos iniciales, con el fin de encontrar tantos óptimos locales distintos como sea posible y seleccionar el mejor de ellos. En principio, el óptimo de un problema no lineal no está necesariamente en el contorno de la región factible como sucedía en LP.

Los métodos de programación no lineal se pueden clasificar en dos grandes familias:

- *Métodos factibles*

Mantienen factibilidad al partir de un punto factible y moverse en direcciones factibles. Son una generalización del método simplex de LP. Resuelven una secuencia de subproblemas con un conjunto de restricciones activas que cambia en cada iteración. Los principales inconvenientes son la selección del conjunto de restricciones activas y la dificultad de satisfacer las restricciones. Entre ellos se van a mostrar: *método del máximo descenso reducido*, *método de Newton reducido*, *método de cuasi-Newton reducido*, *método de programación cuadrática secuencial*.

- *Métodos de penalización o de minimización no restringida secuencial SUMT*

Minimizan una función relacionada con el lagrangiano que tiene el mismo mínimo. Sustituyen un problema NLP por una sucesión de problemas de optimización no restringida, cuyas soluciones convergen a la misma solución. En la función objetivo se incluyen unas penalizaciones que miden las violaciones de las restricciones y además unos parámetros que determinan la importancia de cada restricción. Dentro de estos métodos existen dos variantes, una que se mueve por fuera de la región factible (penalización exterior) y otra que se mueve por dentro de la región factible (penalización interior). El *método de penalización* (exterior) penaliza la violación de las restricciones. Se mueve por puntos infactibles. Se aplica para restricciones de igualdad y desigualdad. El *método barrera* (interior) evita que se alcance el contorno de las restricciones. Se mueve por puntos estrictamente factibles. No es válido para restricciones de igualdad. El *método del lagrangiano aumentado* añade una penalización cuadrática al lagrangiano de una función.

Método del máximo descenso reducido

Sea el problema de minimización de función objetivo no lineal con restricciones lineales (P'')

$$\begin{aligned} &\min_x f(x) \\ &Ax = b \end{aligned} \qquad (1.129)$$

donde $x \in \mathbb{R}^n$, $A \in \mathbb{R}^{m \times n}$ y $b \in \mathbb{R}^m$.

El problema se transforma en una minimización de la función reducida $\phi(v)$ en la región factible

$$\min_v \phi(v) = f(\overline{x} + Zv) \tag{1.130}$$

donde $v \in \mathbb{R}^{n-m}$.

Cada iteración determina un nuevo punto

$$x_{k+1} = x_k + p_k \tag{1.131}$$

$$p_k = Zv_k \tag{1.132}$$

donde p_k se obtiene a partir de v_k.

En el método del máximo descenso v_k en la función proyectada es su gradiente

$$v_k = -Z^T \nabla f(x_k) \tag{1.133}$$

$$p_k = Zv_k = -ZZ^T \nabla f(x_k) \tag{1.134}$$

donde p_k depende matemáticamente de la elección de la matriz Z.

Método de Newton reducido

Sea el problema de minimización de función objetivo no lineal con restricciones lineales (P''). En el método de Newton v_k se determina mediante la resolución de un sistema de ecuaciones lineales. p_k no depende de la elección de la matriz de Z, matemáticamente da lugar a la misma dirección, pero numéricamente sí puede afectar de forma drástica

$$\left[Z^T \nabla^2 f(x_k) Z \right] v_k = -Z^T \nabla f(x_k) \tag{1.135}$$

Método de cuasi-Newton reducido

Sea el problema de minimización de función objetivo no lineal con restricciones lineales (P''). En el método de cuasi-Newton v_k se calcula resolviendo un sistema de ecuaciones con la aproximación del hessiano en cada iteración

$$B_k v_k = -Z^T \nabla f(x_k) \tag{1.136}$$

y las matrices auxiliares en la actualización de la matriz B_k

$$s_k = Z^T (x_{k+1} - x_k) \tag{1.137}$$

$$y_k = Z^T \left(\nabla f(x_{k+1}) - \nabla f(x_k) \right) \tag{1.138}$$

Programación cuadrática secuencial

Sea el problema de función objetivo no lineal con restricciones no lineales de igualdad (P)

$$\begin{aligned} &\min_x f(x) \\ &g_i(x) = 0 \quad i = 1, \dots, m \end{aligned} \tag{1.139}$$

donde $x \in \mathbb{R}^n$ y $g_i : \mathbb{R}^n \to \mathbb{R}$. El lagrangiano de este problema es

$$L(x, \lambda) = f(x) + \lambda^T g(x) \tag{1.140}$$

Se aplica el método de Newton a la resolución del sistema de ecuaciones no lineales $\nabla L(x, \lambda) = 0$, que es la condición de optimalidad del lagrangiano, mediante el procedimiento iterativo

$$\begin{pmatrix} x_{k+1} \\ \lambda_{k+1} \end{pmatrix} = \begin{pmatrix} x_k \\ \lambda_k \end{pmatrix} + \begin{pmatrix} p_k \\ \rho_k \end{pmatrix} \tag{1.141}$$

donde p_k y ρ_k se determinan resolviendo este sistema de ecuaciones lineales

$$\nabla^2 L(x_k, \lambda_k) \begin{pmatrix} p_k \\ \rho_k \end{pmatrix} = -\nabla L(x_k, \lambda_k) \tag{1.142}$$

$$\begin{pmatrix} \nabla^2_{xx} L(x_k,\lambda_k) & -\nabla g(x_k) \\ -\nabla g(x_k)^T & 0 \end{pmatrix} \begin{pmatrix} p_k \\ \rho_k \end{pmatrix} = \begin{pmatrix} -\nabla_x L(x_k,\lambda_k) \\ g(x_k) \end{pmatrix} \tag{1.143}$$

Este sistema de ecuaciones representa también las condiciones de optimalidad de primer orden de este problema de optimización cuadrática con restricciones lineales

$$\begin{aligned} &\min_{p_k} \frac{1}{2} p_k^T \left[\nabla^2_{xx} L(x_k,\lambda_k)\right] p_k + p_k^T \left[\nabla_x L(x_k,\lambda_k)\right] \\ &\left[\nabla g(x_k)\right]^T p_k + g(x_k) = 0 \end{aligned} \tag{1.144}$$

siendo ρ_k los multiplicadores de Lagrange de las ecuaciones lineales.

La función objetivo cuadrática resulta ser una aproximación en serie de Taylor del lagrangiano en el punto $\begin{pmatrix} x_k \\ \lambda_k \end{pmatrix}$ y las restricciones resultan una aproximación lineal de las restricciones originales $g(x_k + p)$.

En cada iteración se resuelve el problema cuadrático anterior y se obtiene p_k y ρ_k y se actualizan los valores de x_{k+1} y λ_{k+1}.

Método de penalización exterior

Este procedimiento consiste en incluir las restricciones del problema en la función objetivo penalizando los valores no permitidos en las restricciones. Puede utilizarse para problemas con restricciones de desigualdad y de igualdad. Las iteraciones se pueden mover por puntos exteriores de la región factible.

Sea el problema (P')

$$\begin{aligned} &\min_x f(x) \\ &g_i(x) \le 0 \quad i = 1,\dots,m \\ &h_j(x) = 0 \quad j = 1,\dots,l \end{aligned} \tag{1.145}$$

Se define una función de penalización $P(x)$ de la forma

$$P(x) = \sum_{i=1}^{m} \phi(g_i(x)) + \sum_{j=1}^{l} \gamma(h_j(x)) \tag{1.146}$$

de modo que $\begin{cases} \phi(y) = 0 & y \leq 0 \\ \phi(y) > 0 & y > 0 \end{cases}$ y $\begin{cases} \gamma(y) = 0 & y = 0 \\ \gamma(y) > 0 & y \neq 0 \end{cases}$.

Unas funciones que se utilizan habitualmente son $\phi(y) = (\max\{0, y\})^p$ y $\gamma(y) = |y|^q$ o $\gamma(x) = \frac{1}{2}\sum_{j=1}^{l} h_j(x)^2 = \frac{1}{2}h(x)^T h(x)$

En cada iteración partiendo de un punto x_k se resuelve el problema

$$\min_x \pi(x, \mu_k) = \{f(x) + \mu_k P(x)\} \tag{1.147}$$

siendo μ_k un parámetro de penalización escalar positivo y se obtiene una solución óptima x_{k+1}. Si $\mu_k P(x_{k+1}) < \varepsilon$ (nivel de tolerancia admitido) acaba el proceso; si no, se incrementa monótonamente hacia ∞ la penalización definiendo el multiplicador de la siguiente iteración como $\mu_{k+1} = \beta\mu_k$, donde $\beta > 1$, y se repite el proceso. Conforme aumenta el parámetro de penalización aumenta el efecto de dicho término y los puntos se acercan a las restricciones de igualdad. En el límite este parámetro es el encargado de hacer cumplir las restricciones de igualdad y, por tanto, obtener soluciones factibles.

Para este problema de optimización

$$\begin{aligned} &\min_x f(x) \\ &h_j(x) = 0 \quad j = 1, \ldots, l \end{aligned} \tag{1.148}$$

las condiciones de optimalidad de la función de penalización suponiendo una penalización cuadrática son

$$\nabla f(x(\mu)) + \mu\sum_{j=1}^{l} \nabla h_j(x(\mu)) h_j(x(\mu)) = 0 \tag{1.149}$$

y las condiciones necesarias de optimalidad de primer orden del problema original

$$\begin{aligned} &\nabla f(x(\mu)) + \sum_{j=1}^{l} \lambda_j(\mu) \nabla h_j(x(\mu)) = 0 \\ &\lambda_j(\mu) = \mu h_j(x(\mu)) \quad j = 1, \ldots, l \end{aligned} \tag{1.150}$$

Obsérvese que la condición de complementariedad de holguras en lugar de ser 0 es $1/\mu$.

Método de penalización interior o método barrera

El método de penalización interior es similar al anterior pero plantea funciones que no permiten que la solución esté fuera de la región factible. Se mueven siempre por el interior, por lo que su uso se ha solido limitar a problemas con restricciones de desigualdad.

Sea el problema (P)

$$\begin{aligned} &\min_x f(x) \\ &g_i(x) \leq 0 \quad i = 1,\dots,m \end{aligned} \tag{1.151}$$

Se define una función barrera, función continua en el interior de la región factible que se hace ∞ al acercarse al contorno

$$B(x) = \sum_{i=1}^{m} \varphi(g_i(x)) \tag{1.152}$$

donde $\begin{cases} \varphi(y) \geq 0 & y < 0 \\ \lim_{y \to 0^-} \varphi(y) = \infty \end{cases}$. Lo más habitual es definir $\varphi(y) = -1/y$ o $\varphi(x) = -\sum_{i=1}^{m} \log(g_i(x))$ aunque haya otras.

Partiendo de un punto que cumple las restricciones, el procedimiento en la iteración k consiste en resolver el problema

$$\min_x \beta(x, \mu_k) = \{f(x) + \mu_k B(x)\} \tag{1.153}$$

siendo μ_k un parámetro barrera escalar positivo y se obtiene una solución óptima x_{k+1}. Si $\mu_k B(x_{k+1}) < \varepsilon$ (nivel de tolerancia admitido) acaba el proceso; si no, se decrementa monótonamente hacia 0 el parámetro barrera, definiendo el multiplicador de la siguiente iteración como $\mu_{k+1} = \beta \mu_k$, donde $\beta \in (0,1)$ (un valor habitual es 0.01), y se repite el proceso. Conforme decrece el parámetro barrera disminuye el efecto del término barrera y los puntos se acercan al contorno de la región factible.

La condición de optimalidad para una función barrera definida como la inversa de la función es

$$\nabla f(x) + \sum_{i=1}^{m} \frac{\mu}{g_i(x)} \nabla g_i(x) = 0 \tag{1.154}$$

Las condiciones necesarias de optimalidad de primer orden del problema original son

$$\begin{aligned}
&\nabla f(x(\mu)) + \sum_{i=1}^{m} \lambda_i(\mu) \nabla g_i(x(\mu)) = 0 \\
&\lambda_i(\mu) g_i(x(\mu)) = \mu \quad i = 1,\dots,m \\
&\lambda_i(\mu) \geq 0
\end{aligned} \tag{1.155}$$

Obsérvese que la condición de complementariedad de holguras en lugar de ser 0 es μ.

Método del lagrangiano aumentado

Sea el problema de función objetivo no lineal con restricciones no lineales de igualdad (P)

$$\begin{aligned}
&\min_x f(x) \\
&g_i(x) = 0 \quad i = 1,\dots,m
\end{aligned} \tag{1.156}$$

cuyo óptimo es coincidente con el de este otro

$$\begin{aligned}
&\min_{x,\lambda} L(x,\lambda) = f(x) + \lambda^T g(x) \\
&g_i(x) = 0 \quad i = 1,\dots,m
\end{aligned} \tag{1.157}$$

aplicando ahora un método de penalización exterior con penalización cuadrática a las restricciones de igualdad

$$\min_{x,\lambda,\mu} A(x,\lambda,\mu) = f(x) + \lambda^T g(x) + \frac{1}{2} \mu g(x)^T g(x) \tag{1.158}$$

El proceso iterativo empieza por elegir valores de x_0, λ_0 y μ_0. Se comprueba la condición de optimalidad. Si se verifica se detiene el algoritmo. En caso contrario, se sigue.

Se resuelve el problema no lineal sin restricciones y se calcula x_{k+1}

$$\min_x A(x,\lambda_k,\mu_k) = f(x) + \lambda_k^T g(x) + \frac{1}{2}\mu_k g(x)^T g(x) \tag{1.159}$$

Se actualizan λ_{k+1} y μ_{k+1}, considerando $\beta > 1$

$$\begin{aligned} \mu_{k+1} &= \beta\mu_k \\ \lambda_{k+1} &= \lambda_k + \mu_k g(x_{k+1}) \end{aligned} \tag{1.160}$$

Este método proporciona soluciones factibles al introducir un parámetro de penalización μ_k suficientemente elevado. En el límite este parámetro es el encargado de hacer cumplir las restricciones y, por tanto, obtener soluciones factibles lo mismo que en el método de penalización exterior.

III.4.5. Resumen de métodos de optimización no lineal

Los métodos de programación no lineal transforman el problema general con restricciones en otro sin restricciones. Para ello utilizan tres técnicas principales dependiendo del tipo de restricciones: la proyección (reducción) del problema en la región factible, la formación del lagrangiano o la penalización de las mismas. Una vez obtenido el problema sin restricciones utilizan un procedimiento iterativo para moverse de un punto a otro o alternativamente resuelven iterativamente el sistema de ecuaciones no lineales que representan las condiciones de optimalidad.

Método		
Máximo descenso reducido	$\min_x f(x)$ $Ax = b$	$\min_v \phi(v) = f(\overline{x} + Zv)$ $p_k = Zv_k = -ZZ^T\nabla f(x_k)$
Newton reducido	$\min_x f(x)$ $Ax = b$	$\min_v \phi(v) = f(\overline{x} + Zv)$ $\left[Z^T\nabla^2 f(x_k)Z\right]v_k = -Z^T\nabla f(x_k)$
Cuasi-Newton reducido	$\min_x f(x)$ $Ax = b$	$\min_v \phi(v) = f(\overline{x} + Zv)$ $B_k v_k = -Z^T\nabla f(x_k)$

Programación cuadrática secuencial	$\min_x f(x)$ $g_i(x) = 0$	$L(x,\lambda) = f(x) + \lambda^T g(x)$ $\begin{pmatrix} \nabla^2_{xx} L(x_k,\lambda_k) & -\nabla g(x_k) \\ -\nabla g(x_k)^T & 0 \end{pmatrix} \begin{pmatrix} p_k \\ \rho_k \end{pmatrix} = \begin{pmatrix} -\nabla_x L(x_k,\lambda_k) \\ g(x_k) \end{pmatrix}$ $\min_{p_k} \frac{1}{2} p_k^T \left[\nabla^2_{xx} L(x_k,\lambda_k)\right] p_k + p_k^T \left[\nabla_x L(x_k,\lambda_k)\right]$ $\left[\nabla g(x_k)\right]^T p_k + g(x_k) = 0$
Penalización exterior	$\min_x f(x)$ $g_i(x) \le 0$ $h_j(x) = 0$	$\min_x \pi(x,\mu_k) = \{f(x) + \mu_k P(x)\}$ $P(x) = \sum_{i=1}^m \phi(g_i(x)) + \sum_{j=1}^l \gamma(h_j(x))$
Penalización interior o barrera	$\min_x f(x)$ $g_i(x) \le 0$	$\min_x \beta(x,\mu_k) = \{f(x) + \mu_k B(x)\}$ $B(x) = \sum_{i=1}^m \varphi(g_i(x))$
Lagrangiano aumentado	$\min_x f(x)$ $g_i(x) = 0$	$\min_{x,\lambda} L(x,\lambda) = f(x) + \lambda^T g(x)$ $g_i(x) = 0$ $\min_x A(x,\lambda_k,\mu_k) = f(x) + \lambda_k^T g(x) + \frac{1}{2}\mu_k g(x)^T g(x)$

III.5. Referencias

Bazaraa, M.S., Sherali, H.D. and Shetty, C.M. (1993) *Nonlinear Programming. Theory and Algorithms*. John Wiley and Sons.

Gill, P.E., Murray, W. and Wright, M.H. (1981) *Practical Optimization*. Academic Press.

Kuhn, H.W. and Tucker, A.W. (1951) Nonlinear Programming *Econometrica* (19) pp. 50-51.

Luenberger, D.G. (1984) *Linear and Nonlinear Programming*. Addison Wesley.

Nash. S.G. and Sofer. A. (1996) *Linear and Nonlinear Programming*. McGraw-Hill.

III.6. Biblioteca de problemas

PROBLEMA 1

Un vendedor desea localizar el punto donde situar un kiosco de chucherías en las inmediaciones de dos colegios. Los colegios se hallan situados, en un plano de coordenadas, en los puntos $A=(2,0)$ y $B=(0,2)$. En el punto (0,0) existe una tienda de características semejantes a ésta y, por ordenanza municipal, el emplazamiento de la nueva tienda ha de distar al menos 4 metros de la anterior. Determinar el punto en que debe situarse la nueva tienda para que se minimice la suma de las distancias al cuadrado a los colegios.

PROBLEMA 2

Una compañía eléctrica se enfrenta a una demanda dividida en intervalos de carga punta y valle. Si se cobra un precio de p_1 €/MWh durante el tiempo de carga punta, los clientes pedirán $60-0.5p_1$ MWh de energía. Si se cobra un precio de p_2 €/MWh durante el tiempo de carga valle, los clientes pedirán $40-p_2$ MWh de energía. La compañía tiene que tener la suficiente capacidad para satisfacer la demanda durante los dos intervalos. El coste de producción de la compañía es de 100 €/MWh. Determinar cómo puede la compañía optimizar su margen de contribución (ingresos – costes).

PROBLEMA 3

Una empresa produce y vende dos productos A y B a los precios p_A y p_B y en las cantidades Q_A y Q_B. Las cantidades vendidas se relacionan con los precios mediante las siguientes funciones de demanda $Q_A=60-21p_A+0.1p_B$ y $Q_B=50-25p_B+0.1p_A$.

Los costes de producción por unidad de producto son 0.2 y 0.3 y la capacidad de producción disponible en cada uno es 25 y 50, respectivamente. Además, existe un límite en la capacidad total de producción de la empresa representado por la ecuación $Q_A+2Q_B\le 50$.

Formular y resolver el problema de determinar los valores que maximizan el margen de contribución (beneficios brutos) de la empresa.

PROBLEMA 4

Resolver los siguientes problemas utilizando las condiciones de Karush-Khun-Tucker:

1)
$$\begin{aligned}&\min (x-9/4)^2+(y-2)^2\\&x^2-y\le 0\\&x+y-6\le 0\\&-x\le 0\\&-y\le 0\end{aligned}$$

2)
$$\begin{aligned}&\max (x-13/6)^2+(y-1/6)^2\\&4x^2+3y-10\le 0\\&-2x+y\le 0\\&x+y\ge 0\end{aligned}$$

Probar que $A=(1,2)$, $B=(2,\text{-}2)$, $C=(0,0)$ son máximos globales.

3)
$$\begin{aligned}&\min -3x+y-z^2\\&x+y+z\le 0\\&-x+2y+z^2=0\end{aligned}$$

4)
$$\begin{aligned}&\min x\\&x+y+z=0\\&x^2+y^2+z^2=\alpha^2 \quad \alpha>0\end{aligned}$$

¿Es $\overline{x}=(\frac{-2\alpha}{\sqrt{6}},\frac{\alpha}{\sqrt{6}},\frac{\alpha}{\sqrt{6}})$ mínimo global del problema, siendo los multiplicadores $(v_1,v_2)=(-1/3,\frac{1}{\alpha\sqrt{6}})$?

PROBLEMA 5

Sea el problema de minimizar la distancia al cuadrado del punto (0,3) al conjunto de puntos que satisfacen las restricciones

$$y^2 \leq x-1$$
$$x \leq 2$$

1) Representar gráficamente el dominio definido por las restricciones.
2) Determinar las condiciones de Karush-Kuhn-Tucker que proporcionan los puntos estacionarios o críticos y utilizarlas para determinar si los puntos A(1.54,0.735) y B(2,1) lo son.
3) ¿Alguno de los puntos anteriores es mínimo global del problema? ¿Por qué?

PROBLEMA 6

Se considera el problema

$$\max x^2 + y^2$$
$$(x-4)^2 + (y-4)^2 \leq 12$$
$$x + y \geq 5$$

Responder a las siguientes preguntas.

1) Dar una interpretación geométrica del problema
2) Establecer todas las condiciones de óptimo de Karush-Kuhn-Tucker
3) Probar que el punto $x = y = 4 + \sqrt{6}$ satisface dichas condiciones. ¿Es un máximo global?

PROBLEMA 7

Resolver los siguientes problemas de programación no lineal:

$$\min e^{3x_1+4x_2}$$
$$x_1^2 + x_2^2 = 1$$

$$\min 2(x_1^2 + x_2^2 - 1) - x_1$$
$$x_1^2 + x_2^2 - 1 = 0$$

$$\min \frac{1}{2} x^T Q x$$
$$Ax = b$$

$$Q = \begin{pmatrix} \cdot & -13 & -6 & -3 \\ -13 & 23 & -9 & 3 \\ -6 & -9 & -12 & 1 \\ -3 & 3 & 1 & -1 \end{pmatrix} \quad A = \begin{pmatrix} 2 & 1 & 2 & 1 \\ 1 & 1 & 3 & -1 \end{pmatrix} \quad b = \begin{pmatrix} 3 \\ 2 \end{pmatrix} \text{ partiendo de } x = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\min \frac{1}{2} x_1^2 + \frac{1}{2} x_2^2$$
$$x_1 + x_2 \geq 1$$

$$\min \frac{1}{3} x_1^3 + \frac{1}{3} x_2^3$$
$$x_1 + x_2 \geq 1$$

$$\min \frac{1}{3} x_1^3 + \frac{1}{3} x_2^3$$
$$x_1^2 + x_2^2 \geq 1$$

$$\min x_1^2 + x_2^2$$
$$x_1 + x_2 = 1$$

$$\min x_1^2 + x_2^2$$
$$\frac{x_1^2}{25} + \frac{x_2^2}{9} = 1$$

Resolver por el método de penalización interior (barrera): $\begin{array}{l} \min x_1 + x_2 \\ x_1 \geq 0 \\ x_1 \geq 0 \end{array}$

Resolver partiendo del punto (0,3): $\min(x_1 - 2)^4 + (x_1 - 2x_2)^2$

PROBLEMA 8

Calcular analíticamente

$$\min \frac{3}{2} x^2 - 2xy + y^2 + xz + 3z^2 - 4zw + 2w^2 - x - \frac{4}{3} z$$

donde *x*, *y*, *z* y *w* son variables enteras

PROBLEMA 9

Una furgoneta que tiene 4 m^3 de capacidad tiene que transportar 600 kg de tres productos, A, B y C. Cada kilo de A ocupa 0.08 m^3, cada kilo de B ocupa 0.06 m^3 y uno de C 0.07 m^3.

El coste unitario de un kilo de A es función de la cantidad adquirida y viene dado por

$$c_A(x) = \frac{9x+10}{x}, x \geq 50 \text{ (la cantidad mínima a llevar es de 50 kg),}$$

y lo mismo sucede con el artículo B: el coste unitario de un kilo depende de la cantidad adquirida y viene dado por

$$c_B(x) = \frac{12}{1+0.1x}$$

Un kilo de C cuesta 10 €, independiente de la cantidad adquirida.

Plantear un problema para resolver cuántos kilos de cada artículo han de transportarse para que el coste de transporte sea mínimo, y sus condiciones de optimalidad.

PROBLEMA 10

Estamos a 27 de enero. La demanda de cierto producto para el próximo mes de febrero es de 50 toneladas, siendo para el mes de marzo de 70 toneladas. El beneficio por tonelada producida durante el mes de febrero, teniendo en cuenta los costes de materias primas y operativos, depende del número x_1 de toneladas producidas y es, en euros, $100 - \frac{x_1}{6}$, mientras que durante el mes de marzo es de $60 - \frac{x_2}{6}$, donde x_2 son las toneladas producidas en ese mes. La demanda debe ser satisfecha. Si en el mes de febrero se producen más toneladas de las demandadas, el excedente se puede almacenar con un coste por tonelada de 30 €, mientras que si hay excedentes en el mes de marzo son saldados con una pérdida residual por tonelada de 40 €, ya que la empresa ha de pagar esa cantidad por tonelada para que le retiren sus excedentes de producción.

Elaborar el correspondiente modelo de optimización, establecer las condiciones de KKT asociadas y comprobar que x_1 =90, x_2 =60 es la solución global del problema.

PROBLEMA 11

Sea el modelo de programación lineal

$$\begin{aligned}&\min 4x+3y\\&x+y\geq 5\\&3x+2y\leq 12\\&x,y\geq 0\end{aligned}$$

1) Resolverlo gráficamente
2) Establecer para dicho modelo las condiciones necesarias de óptimo de KKT y resolver
3) Comprobar que los puntos que las cumplen son óptimos globales.
4) Plantear y resolver por el método simplex el problema dual asociado a [1].
5) Interpretar los multiplicadores de Lagrange o KKT en relación con el problema dual

PROBLEMA 12

Un inversor quiere invertir 9000 € en dos negocios A y B de actividades afines. Las rentabilidades R_A y R_B obtenidas en cada uno de ellos son variables aleatorias tales que

$$\begin{gathered}E\left(R_A\right)=0.30 \qquad E\left(R_B\right)=0.20\\ Var\left(R_A\right)=0.15 \qquad Var\left(R_B\right)=0.1\\ Cov(R_A,R_B)=0.05\end{gathered}$$

Se quiere obtener una rentabilidad esperada superior al 24%, pero de forma que se minimice la varianza de dicha rentabilidad. ¿Cómo se han de invertir los 9000 €? Probar que, efectivamente, esa inversión es la mejor posible.

PROBLEMA 13

La función de potencial sobre el sólido definido por la bola elipsoidal

$$x^2 - 2xy + 2y^2 + 2yz + 2z^2 \leq 9$$

y el semiespacio

$$x + y + 4z \geq 8$$

viene dada por

$$W = x^2 + y^2 + 2yz + 4z^2 - x + 5y$$

Comprobar que el potencial mínimo se alcanza en el punto $(1, -1, 2)$. Para ello, probar que dicho punto satisface las condiciones necesarias de KKT y las condiciones suficientes de mínimo global.

III.7. Resultados de la biblioteca de problemas

RESULTADO DEL PROBLEMA 1

$$\begin{aligned}
&\min (x-0)^2 + (y-2)^2 + (x-2)^2 + (y-0)^2 \\
&x^2 + y^2 \geq 4^2
\end{aligned}$$

El mínimo (global por las condiciones suficientes de KKT) se alcanza en el punto $(2\sqrt{2}, 2\sqrt{2})$ (multiplicador $u_1 = 2 - \sqrt{2}\,/\,2$)

RESULTADO DEL PROBLEMA 2

$$\begin{aligned}
&\max (60 - 0.5p_1)(p_1 - 100) + (40 - p_2)(p_2 - 100) \\
&60 - 0.5p_1 \geq 0 \\
&40 - p_2 \geq 0
\end{aligned}$$

El máximo (global por las condiciones suficientes de KKT) se alcanza en los valores $(p_1, p_2) = (110, 40)$ (multiplicadores $u_1 = 0$, $u_2 = 60$), beneficio = 50 €.

RESULTADO DEL PROBLEMA 4

1) Mínimo (global KKT): (3/2,9/4), multiplicadores (1/2,0,0,0)

2) Para A, los multiplicadores son (–5/14,–109/42,0). Función objetivo 85/18

   Para B, los multiplicadores son (–4/13,0,–205/39). Función objetivo 85/18

   Para C, los multiplicadores son (0,–4/3,–5/3). Función objetivo 85/18

   Son candidatos a máximo. La función está acotada y otros puntos de KKT candidatos a máximo son: D(1/2,1) (0,–7/3,0), valor de la función objetivo 125/36; y E(1,–1) (0,0,–7/3), valor de la función objetivo 98/36. Al tener menor valor se puede asegurar que A, B y C son máximos globales.

3) Mínimo: (–115/588, –95/588, 5/14) (multiplicadores $u_1 = 5/3$, $v_1 = -4/3$).

4) Como f es convexa, $v_1 < 0$ y h_1 es cóncava, $v_2 < 0$ y h_2 es convexa, se dan las condiciones suficientes de KKT, con lo que es mínimo global.

RESULTADO DEL PROBLEMA 5

1) Problema original forma para formular las condiciones KKT:

$$\begin{aligned}&\min x^2 + (y-3)^2\\ &-x + y^2 + 1 \le 0\\ &x - 2 \le 0\end{aligned}$$

   Condiciones KKT:

$$\begin{aligned}&2x - u_1 + u_2 = 0\\ &2(y-3) + 2yu_1 = 0\\ &u_1(-x + y^2 + 1) = 0\\ &u_2(x-2) = 0\\ &u_1, u_2 \ge 0\end{aligned}$$

2) ¿Son factibles ambos puntos? Sí (redondeo del tercer decimal)

   Condiciones para primer punto:

   Se da la primera restricción del problema con igualdad pero no la segunda, entonces $u_2 = 0$. Entonces de la primera ecuación se tiene $2x - u_1 = 0$, es decir,

$u_1 = 2x = 2 \cdot 1.54 = 3.08$ y de la segunda ecuación se tiene $u_1 = \dfrac{-2(0.735-3)}{2 \cdot 0.735} = 3.08$ (redondeando, claro). Es candidato a mínimo.

Condiciones para segundo punto:

Se dan ambas restricciones con igualdad, luego ambos multiplicadores pueden ser distintos de cero. De la segunda ecuación se tiene $u_1 = \dfrac{-2(1-3)}{2 \cdot 1} = 2$ y sustituyendo en la primera sería $u_2 = u_1 - 2x = 2 - 2 \cdot 2 = -2$, luego no es candidato a mínimo.

El segundo no puede serlo y el primero lo es ya que todas las funciones son convexas (basta verlo para la objetivo y la primera que es la activa aunque es trivial que la segunda lo es; las matrices hessianas son:

$$H_f = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \text{ y } H_{g_1} = \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}$$

ambas semidefinidas positivas).

RESULTADO DEL PROBLEMA 6

1) Se puede interpretar como el problema de encontrar el punto más distante al origen que está dentro del círculo de centro (4,4) y radio $\sqrt{12}$ y pertenece al semiespacio $x + y \geq 5$.

2) KKT:

$$\begin{aligned} &2x + 2(x-4)u_1 - u_2 = 0 \\ &2y + 2(y-4)u_1 - u_2 = 0 \\ &u_1((x-4)^2 + (y-4)^2 - 12) = 0 \\ &u_2(-x-y+5) = 0 \\ &u_1, u_2 \leq 0 \end{aligned}$$

3) $u_2 = 0$ pues no se da la igualdad en la segunda inecuación, y de las dos primeras se tiene:

$$\left.\begin{array}{r} 2(4+\sqrt{6})+2\sqrt{6}u_1=0 \\ 2(4+\sqrt{6})+2\sqrt{6}u_1=0 \end{array}\right\} \quad u_1=\frac{-4\sqrt{6}-6}{6}<0\,.$$

Luego sí cumple las condiciones (es de maximización).

No puede asegurarse que sea óptimo global por las KKT (la función objetivo es convexa no cóncava). Para asegurarlo había que ver todos los puntos KKT que salen y si es el único se puede asegurar pues la función está acotada (salvo para puntos en los que los gradientes de las restricciones activas sean linealmente dependientes que podrían no salir al resolver el sistema de ecuaciones de KKT). Es el único que sale al resolver.

RESULTADO DEL PROBLEMA 8

Nodo 0: todas relajadas y sin restricciones, resolver gradiente =0:

$$\begin{array}{ll} 3x-2y+z-1=0 & x+z-1=0 \\ -2x+2y=0 \Rightarrow \quad x=y & \\ x+6z-4w-4/3=0 & x+2z-4/3=0 \\ -4z+4w=0 \Rightarrow \quad z=w & \\ (x,y,z,w)=(2/3,2/3,1/3,1/3) & f.ob=-0.5555 \end{array}$$

Se ramifica en $x \leq 0$ y $x \geq 1$.

Nodo 1: Ramificado con $x \leq 0$. Al aplicar las condiciones KKT el sistema es casi igual, excepto la primera restricción que es la derivada respecto a x, y la holgura:

$$\begin{array}{lll} 3x-2y+z-1+u_1=0 & & x+z-1+u_1=0 \\ -2x+2y=0 \Rightarrow \quad x=y & & \\ x+6z-4w-4/3=0 & \text{Multiplicador mayor o igual que 0,} & x+2z-4/3=0 \\ -4z+4w=0 \Rightarrow \quad z=w & & \\ u_1x=0 & & u_1x=0 \end{array}$$

- $u_1=0$ Sería el mismo sistema anterior cuya solución no es factible en este caso $(x>0)$
- $u_1 \neq 0 \Rightarrow x=0 \Rightarrow z=2/3 \Rightarrow u_1=1/3$

Solución: $(x,y,z,w)=(0,0,2/3,2/3)$ $u_1=1/3$ $f.ob=-4/9=-0.444$

Nodo 2: Ramificado con $x\geq 1$, o sea, $-x+1\leq 0$. Al aplicar las condiciones KKT el sistema es casi igual, excepto la primera restricción que es la derivada respecto a x, y la holgura:

$$\begin{aligned}&3x-2y+z-1-u_1=0\\&-2x+2y=0\Rightarrow\quad x=y\\&x+6z-4w-4/3=0\\&-4z+4w=0\Rightarrow\quad z=w\\&u_1(x-1)=0\end{aligned}$$

Multiplicador mayor o igual que 0,

$$\begin{aligned}&x+z-1-u_1=0\\&x+2z-4/3=0\\&u_1(x-1)=0\end{aligned}$$

- $u_1=0$ Sería el mismo sistema del nodo 0 cuya solución no es factible en este caso ($x<1$)
- $u_1\neq 0\Rightarrow x=1\Rightarrow z=1/6\Rightarrow u_1=1/6$

Solución: $(x,y,z,w)=(1,1,1/6,1/6)$ $u_1=1/6$ $f.ob=-0.5277$

No hay poda. Ramificamos este último nodo (más prometedor), en z. Es decir, se ramifica en $z\leq 0$ (nodo 3) y $z\geq 1$ (Nodo 4), además de tener $x\geq 1$ en ambos.

<u>Nodo 3:</u> Ramificado con $z\leq 0$. Al aplicar las condiciones KKT el sistema es casi igual al del nodo 2, excepto la tercera restricción que es la derivada respecto a z, y la holgura:

$$\begin{aligned}&3x-2y+z-1-u_1=0\\&-2x+2y=0\Rightarrow\quad x=y\\&x+6z-4w-4/3+u_2=0\\&-4z+4w=0\Rightarrow\quad z=w\\&u_1(x-1)=0\\&u_2z=0\end{aligned}$$

Multiplicador mayor o igual que 0,

$$\begin{aligned}&x+z-1-u_1=0\\&x+2z-4/3+u_2=0\\&u_1(x-1)=0\\&u_2z=0\end{aligned}$$

- $u_2=0$ Sería el mismo sistema del nodo 2 cuya solución no es factible en este caso ($z>0$)
- $u_2\neq 0\Rightarrow z=0$

$u_1 = 0 \Rightarrow x = 1 \Rightarrow u_2 = 1/3$:

$(x,y,z,w) = (1,1,0,0) \quad (u_1,u_2) = (0,1/3)\ f.ob = -0.5$

$u_1 \neq 0 \Rightarrow x = 1 \Rightarrow u_1 = 0, u_2 = 1/3$ misma solución que la anterior

El nodo 3 es podado, no se ramifica al ser entero, y además es la mejor solución obtenida hasta el momento. Se puede podar el nodo 2, ya que tiene peor cota, pero no el nodo 4.

Nodo 4: Ramificado con $z \geq 1$, o sea, $-z + 1 \leq 0$. Al aplicar las condiciones KKT el sistema es casi igual al del nodo 2, excepto la tercera restricción que es la derivada respecto a z, y la holgura:

$$\begin{aligned}
&3x - 2y + z - 1 - u_1 = 0 \\
&-2x + 2y = 0 \Rightarrow \quad x = y \\
&x + 6z - 4w - 4/3 - u_2 = 0 \\
&-4z + 4w = 0 \Rightarrow \quad z = w \\
&u_1(x-1) = 0 \\
&u_2(z-1) = 0
\end{aligned}$$

Multiplicador mayor o igual que 0,

$$\begin{aligned}
&x + z - 1 - u_1 = 0 \\
&x + 2z - 4/3 - u_2 = 0 \\
&u_1(x-1) = 0 \\
&u_2(z-1) = 0
\end{aligned}$$

- $u_2 = 0$ Sería el mismo sistema del nodo 2 cuya solución no es factible en este caso ($z < 1$)
- $u_2 \neq 0 \Rightarrow z = 1$

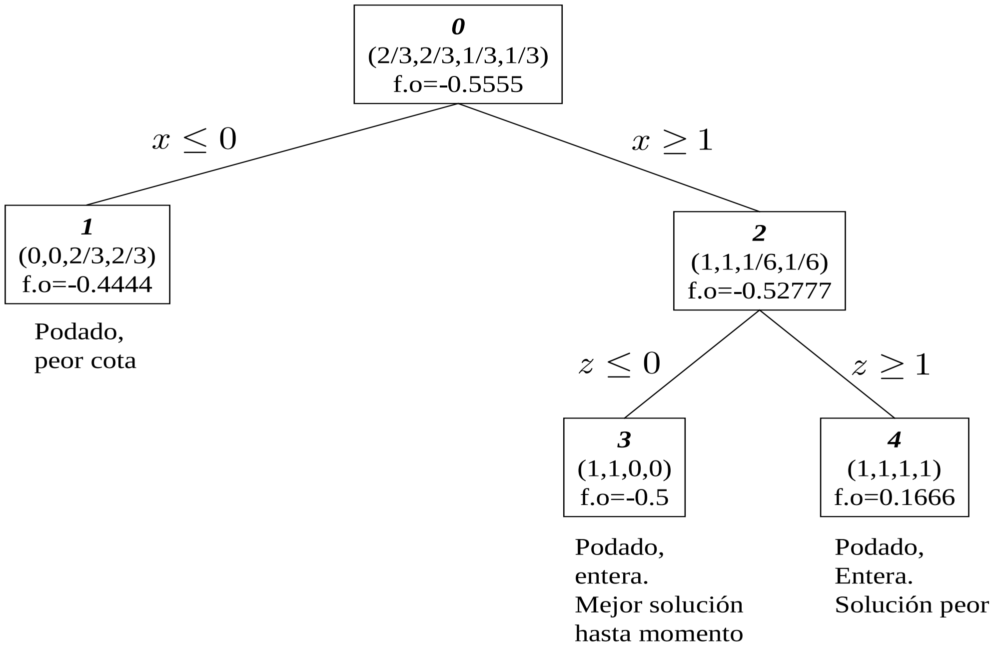
  $u_1 = 0 \Rightarrow x = 0$ No factible.

  $u_1 \neq 0 \Rightarrow x = 1 \Rightarrow u_1 = 1, u_2 = 5/3$

  $(x,y,z,w) = (1,1,1,1) \quad (u_1,u_2) = (1,5/3)\ f.ob = 0.166666$

Nodo podado por ser entera; y cota peor que la ya obtenida hasta el momento.

No quedan nodos sin explorar. FINAL.



RESULTADO DEL PROBLEMA 9

$$\left.\begin{array}{ll}\min & \dfrac{9x+10}{x}x+\dfrac{12}{1+0.1y}y+10z \\ s.a. & x+y+z=600 \\ & 0.08x+0.06y+0.07z\le 4 \\ & x\ge 50 \\ & y\ge 0 \\ & z\ge 0\end{array}\right\}\begin{array}{ll}\min & 9x+10+\dfrac{12y}{1+0.1y}+10z \\ s.a. & x+y+z=600 \\ & 0.08x+0.06y+0.07z\le 4 \\ & x\ge 50 \\ & y\ge 0 \\ & z\ge 0\end{array}$$

$$\text{KKT:}\quad\begin{array}{l} 9+v+0.08u_2-u_3=0 \\ \dfrac{12}{(1+0.1y)^2}+v+u_2-u_4=0 \\ 10+v+0.07u_2-u_5=0 \\ u_2(0.08x+0.06y+0.07z-4)=0 \\ u_3(x-50)=0 \\ u_4y=0 \\ u_5z=0 \\ u_i\ge 0 \quad (x,y,z)\ factible\end{array}$$

2º orden: $H_f = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -2.4/(1+0.1y)^3 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ semidefinida negativa, luego cóncava, luego no se dan las condiciones suficientes KKT. El problema es acotado, luego si sólo saliera un punto que verifique las necesarias sería el mínimo global, pero, si salen más habría que comparar entre ellos cuál es el de menor valor.

RESULTADO DEL PROBLEMA 10

Modelo: $\begin{cases} \max\ (100-\dfrac{x_1}{6})x_1 + (60-\dfrac{x_2}{6})x_2 - 30(x_1-50) - 40(x_1+x_2-120) \\ s.a.\ x_1 \geq 50 \to \qquad 50 - x_1 \leq 0 \\ \quad x_1 + x_2 \geq 120 \to 120 - x_1 - x_2 \leq 0 \\ \quad x_2 \geq 0 \to \qquad -x_2 \leq 0 \end{cases}$

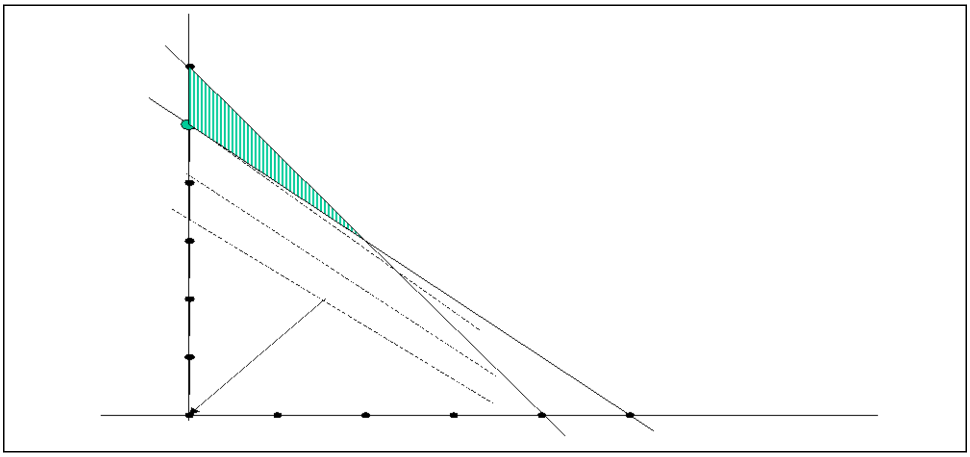
KKT(necesarias):

$$\begin{cases} 100 - \dfrac{2}{6}x_1 - 30 - 40 - \lambda_1 - \lambda_2 = 0 \\ 60 - \dfrac{2}{6}x_2 - 40 - \lambda_2 - \lambda_3 = 0 \\ \lambda_1(50-x_1) = 0 \quad \lambda_2(120-x_1-x_2) = 0 \quad \lambda_3(-x_1) = 0 \quad \lambda_i \leq 0 \end{cases}$$

KKT (suficientes): las restricciones son convexas (por ser lineales) y la función objetivo ¿es cóncava? $H_f(x) = \begin{pmatrix} -2/6 & 0 \\ 0 & -2/6 \end{pmatrix}$ definida negativa, luego es cóncava para todo x. Así pues las condiciones KKT son necesarias y suficientes para encontrar el óptimo global de este problema.

Punto (90,60): para este punto (que es factible ya que cumple todas las restricciones), ninguna restricción es activa, luego $\lambda_i = 0$. La primera ecuación KKT sería entonces 100–180/6–30–40=0 que se cumple, y la segunda sería 60–120/6–40=0 que se cumple. Luego, cumple las condiciones KKT, luego es máximo global del problema.

RESULTADO DEL PROBLEMA 11



1) $$\left.\begin{aligned} \min \quad & 4x+3y \\ & -x-y+5\le 0 \\ & 3x+2y-12\le 0 \\ & -x\le 0 \\ & -y\le 0 \end{aligned}\right\} LP \qquad KKT \quad \begin{cases} 4-\lambda_1+3\lambda_2-\lambda_3 \quad =0 \\ 3-\lambda_1+2\lambda_2 \quad -\lambda_4=0 \\ \lambda_1(x+y-5)=0 \\ \lambda_2(3x+2y-12)=0 \\ \lambda_3 x=0 \\ \lambda_4 y=0 \quad \lambda_i\ge 0 \end{cases}$$

Resolver sistema:

1. $\lambda_3=0$

$$\lambda_4=0 \quad\Rightarrow\quad 4-\lambda_1+3\lambda_2=0 \text{ y } 3-\lambda_1+2\lambda_2=0 \quad\Rightarrow\quad \lambda_2=-1 \text{ NO}$$
$$\lambda_4\ne 0 \quad\Rightarrow\quad y=0 \quad \text{No factible}$$

2. $\lambda_3\ne 0 \quad\Rightarrow\quad x=0$

$$\lambda_4=0$$
$$\lambda_1=0 \quad\Rightarrow\quad \lambda_2=-3/2 \text{ NO}$$
$$\lambda_1\ne 0 \quad\Rightarrow\quad x+y-5=0 \quad\Rightarrow\quad y=5 \quad\Rightarrow\quad \lambda_2=0 \quad\Rightarrow \lambda_1=3 \quad\Rightarrow\quad \lambda_3=1$$
$$Punto \quad (0,5) \quad Multiplicadores\ (3,0,1,0)$$
$$\lambda_4\ne 0 \quad\Rightarrow\quad y=0 \quad \text{No factible}$$

2) Todo lineal, se da la convexidad de todas las funciones y por lo tanto las condiciones suficientes de KKT.

3) Dual

$$\left.\begin{array}{ll}\max & 5w_1+12w_2\\ & w_1+3w_2\le 4\\ & w_1+2w_2\le 3\\ & w_1\ge 0, w_2\le 0\end{array}\right\}\quad t_2=-w_2\quad \left\{\begin{array}{l}\max \quad 5w_1-12t_2\\ w_1-3t_2\le 4\\ w_1-2t_2\le 3\\ w_1\ge 0, t_2\ge 0\end{array}\right. \quad\Rightarrow\quad \begin{array}{l}\max \quad 5w_1-12t_2\\ w_1-3t_2+S1=4\\ w_1-2t_2+S2=3\\ w_1,t_2,S1,S2\ge 0\end{array}$$

	5	–12	0	0	0
S1	1	–3	1	0	4
S2	1	–2	0	1	3

	0	–2	0	–5	–15
S1	0	–1	1	–1	1
W1	1	–2	0	1	3

4) Fácil ver que las condiciones KKT corresponden al mismo sistema, y a las condiciones de holgura complementaria del óptimo del dual.

RESULTADO DEL PROBLEMA 13

Se trata de minimizar la función

$$\begin{aligned}&\min W = x^2+y^2+2yz+4z^2-x+5y\\&x^2-2xy+2y^2+2yz+2z^2\le 9\\&x+y+4z\ge 8\end{aligned}$$

que puesta en forma estándar

$$\begin{aligned}&\min W = x^2+y^2+2yz+4z^2-x+5y\\&x^2-2xy+2y^2+2yz+2z^2-9\le 0\\&-x-y-4z+8\le 0\end{aligned}$$

El lagrangiano será

$$\begin{aligned}&L(x,y,z,\lambda_1,\lambda_2)=x^2+y^2+2yz+4z^2-x+5y+\\&+\lambda_1(x^2-2xy+2y^2+2yz+2z^2-9)+\lambda_2(-x-y-4z+8)\end{aligned}$$

Las condiciones de KKT son las siguientes

$$
\begin{aligned}
&2x - 1 + \lambda_1(2x - 2y) + \lambda_2(-1) = 0 \\
&2y + 2z + 5 + \lambda_1(-2x + 4y + 2z) + \lambda_2(-1) = 0 \\
&2y + 8z + \lambda_1(2y + 4z) + \lambda_2(-4) = 0 \\
&\lambda_1(x^2 - 2xy + 2y^2 + 2yz + 2z^2 - 9) = 0 \\
&\lambda_2(-x - y - 4z + 8) = 0 \\
&\lambda_1, \lambda_2 \geq 0
\end{aligned}
$$

además el punto debe ser factible, es decir, debe cumplir las ecuaciones

$$
\begin{aligned}
&x^2 - 2xy + 2y^2 + 2yz + 2z^2 \leq 9 \\
&x + y + 4z \geq 8
\end{aligned}
$$

Para dicho punto las restricciones son activas luego los multiplicadores de Lagrange pueden ser mayores que 0 de acuerdo con las ecuaciones de la complementariedad de holguras. Sustituyendo el punto $(1,-1,2)$ en las condiciones KKT nos queda

$$
\begin{aligned}
&1 + 4\lambda_1 - \lambda_2 = 0 \\
&7 - 2\lambda_1 - \lambda_2 = 0 \\
&14 + 6\lambda_1 - 4\lambda_2 = 0
\end{aligned}
$$

que da lugar a estos valores de los multiplicadores $\lambda_1 = 1$ y $\lambda_2 = 5$ que verifican las ecuaciones. Luego el punto es un punto estacionario.

Calculamos el hessiano del lagrangiano

$$
\nabla^2 L(x,y,z) = \begin{pmatrix} \dfrac{\partial^2 L}{\partial x^2} & \dfrac{\partial^2 L}{\partial x \partial y} & \dfrac{\partial^2 L}{\partial x \partial z} \\ \dfrac{\partial^2 L}{\partial y \partial x} & \dfrac{\partial^2 L}{\partial y^2} & \dfrac{\partial^2 L}{\partial y \partial z} \\ \dfrac{\partial^2 L}{\partial z \partial x} & \dfrac{\partial^2 L}{\partial z \partial y} & \dfrac{\partial^2 L}{\partial z^2} \end{pmatrix}
$$

siendo $\dfrac{\partial^2 L}{\partial x^2} = 2 + 2\lambda_1$, $\dfrac{\partial^2 L}{\partial x \partial y} = -2\lambda_1$, $\dfrac{\partial^2 L}{\partial x \partial z} = 0$, $\dfrac{\partial^2 L}{\partial y^2} = 2 + 4\lambda_1$, $\dfrac{\partial^2 L}{\partial y \partial z} = 2 + 2\lambda_1$, $\dfrac{\partial^2 L}{\partial z^2} = 8 + 4\lambda_1$. Luego el hessiano es

$$\nabla^2 L(x,y,z) = \begin{pmatrix} 2+2\lambda_1 & -2\lambda_1 & 0 \\ -2\lambda_1 & 2+4\lambda_1 & 2+2\lambda_1 \\ 0 & 2+2\lambda_1 & 8+4\lambda_1 \end{pmatrix}_{\lambda_1=1} = \begin{pmatrix} 4 & -2 & 0 \\ -2 & 6 & 4 \\ 0 & 4 & 12 \end{pmatrix}$$

que es una matriz definida positiva ya que los determinantes en orden creciente son todos estrictamente positivos $|4| = 4$, $\begin{vmatrix} 4 & -2 \\ -2 & 6 \end{vmatrix} = 20$ y $\begin{vmatrix} 4 & -2 & 0 \\ -2 & 6 & 4 \\ 0 & 4 & 12 \end{vmatrix} = 176$. Los autovalores son 2.1770, 5.7411 y 14.0819.

Alternativamente, se puede comprobar que la función objetivo y las restricciones son convexas en dicho punto, es decir, el hessiano es matriz semidenifida positiva.

$$\nabla^2 f(x,y,z) = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 2 \\ 0 & 2 & 8 \end{pmatrix}$$

que es una matriz definida positiva ya que los determinantes en orden creciente son todos estrictamente positivos $|2| = 2$, $\begin{vmatrix} 2 & 0 \\ 0 & 2 \end{vmatrix} = 4$ y $\begin{vmatrix} 2 & 0 & 0 \\ 0 & 2 & 2 \\ 0 & 2 & 8 \end{vmatrix} = 24$. Los autovalores son 1.3944, 2.0000 y 8.6056.

$$\nabla^2 g_1(x,y,z) = \begin{pmatrix} 2 & -2 & 0 \\ -2 & 4 & 2 \\ 0 & 2 & 4 \end{pmatrix}$$

que es una matriz definida positiva ya que los determinantes en orden creciente son todos estrictamente positivos $|2| = 2$, $\begin{vmatrix} 2 & -2 \\ -2 & 4 \end{vmatrix} = 4$ y $\begin{vmatrix} 2 & -2 & 0 \\ -2 & 4 & 2 \\ 0 & 2 & 4 \end{vmatrix} = 8$. Los autovalores son 0.3961, 3.1099 y 6.4940.

La segunda restricción es una recta. Por tanto, es cóncava y convexa a la vez

$$\nabla^2 g_2(x,y,z) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

que es una matriz semidefinida positiva y semidefinida negativa ya que los determinantes son todos 0.

IV. Optimización estocástica

IV.1. Estocasticidad

La estocasticidad o incertidumbre aparece en todos los sistemas pero hasta ahora no era posible la solución de problemas de optimización de grandes sistemas considerando explícitamente ésta. La incertidumbre puede deberse a carencia de datos fiables, errores de medida o tratarse de parámetros que representan información sobre el futuro. Por ejemplo, en el caso de planificación de sistemas de energía eléctrica la incertidumbre surge principalmente en la demanda y precios futuros, las aportaciones hidráulicas o la disponibilidad de los elementos de generación y red.

En optimización determinista se supone que los parámetros del problema son conocidos con certeza. En optimización estocástica se relaja esta condición. No se conocen sus valores, sólo sus distribuciones y habitualmente se supone que éstas son discretas con un número finito de estados posibles.

Los tipos de modelos que aparecen en programación lineal[11] estocástica son motivados principalmente por problemas con decisiones de tipo *aquí y ahora* (*here and now*) (Dantzig, 1955), decisiones previas bajo futuro incierto. Esto es, decisiones que deben tomarse basándose en información a priori, existente o supuesta, sobre situaciones futuras sin realizar observaciones adicionales.

[11] La restricción a problemas lineales es por razones prácticas: primero, son los más frecuentemente utilizados, y segundo, aquéllos con métodos de solución más eficientes y avanzados. Las técnicas que se presentan admiten también funciones no lineales si región factible y función objetivo son convexas

Recurso[12] es la capacidad de tomar una acción correctora después de que un suceso aleatorio haya ocurrido. Por ejemplo, se toman hoy un conjunto de decisiones x_1 con valores de los parámetros conocidos (es decir, deterministas), durante la noche se producen unos sucesos aleatorios y mañana se toman un conjunto de acciones correctoras x_2^ω que mitigan (corrigen) los efectos de los sucesos aleatorios ω sobre las decisiones de hoy. Las decisiones de la segunda etapa x_2^ω son los recursos. La función objetivo asociada a estas decisiones es la *función de recursos*, depende de las decisiones de la primera etapa x_1 y de los valores ω de los parámetros aleatorios. Esta circunstancia incrementa exponencialmente el tamaño del problema. Este problema, denominado *problema lineal estocástico bietapa*[13] PLE-2, determina el valor óptimo de x_1 y x_2^ω. En el ámbito de planificación, conceptualmente *asignación óptima de recursos*, de sistemas de energía eléctrica se pueden citar como ejemplo modelos de:

3. Coordinación hidrotérmica (Morton, 1993), (Jacobs, 1995), (Morton, 1996, (Pereira, 1991)

   Ésta consiste en la planificación de la explotación de un sistema hidrotérmico en varios periodos considerando las aportaciones hidráulicas como parámetro aleatorio. El objetivo es determinar la explotación del subsistema hidráulico que minimiza los costes de explotación esperados derivados de la utilización del subsistema térmico.

4. Planificación a largo plazo de la expansión de la red (Perez, 1991), (Stanford, 1989)

   La expansión de la red se formula como un problema estocástico bietapa. La primera etapa recoge las decisiones de inversión en corredores y la segunda las

---

[12] El recurso se denomina *completo* cuando para cualquier decisión de la primera etapa (independientemente de su factibilidad) existen decisiones factibles para la segunda etapa. Se denomina *relativamente completo* si se cumple sólo para decisiones factibles de la primera etapa y *parcial* o *simple* cuando no siempre las decisiones de la segunda etapa son factibles para las decisiones de la primera etapa.

[13] Se denomina bietapa porque tiene dos etapas, una asociada a las decisiones anteriores a la realización de la incertidumbre y otra a las decisiones correctoras una vez realizada la incertidumbre.

decisiones de explotación. Los distintos escenarios se deben a variaciones en la demanda, la hidraulicidad y la disponibilidad de elementos de generación y red. En el modelo PERLA la expansión de la red se formula como un problema de planificación estática que optimiza el coste total anual, suma del coste anualizado de inversión, explotación e indisponibilidad. Este ha sido utilizado por REE como herramienta de ayuda a la planificación táctica (con horizontes de 15 a 30 años). El modelo ha sido desarrollado en FORTRAN y utiliza MINOS (Murtagh, 1987) y ZOOM para la solución de los problemas de optimización.

5. Determinación del precio de la electricidad para diferentes tipos de consumidores en sistemas hidráulicos (Mathiesen, 1992)

   Este modelo analiza la formación de precios en el nuevo mercado de electricidad de Noruega. Explícitamente considera la gran variación interanual entre aportaciones hidráulicas (estocásticas) y el reparto anual del agua teniendo en cuenta los diferentes estados. Considera varias regiones y periodos en el año. El modelo propuesto se formula como problema no lineal (cuadrático) estocástico multietapa. La principal hipótesis es la estrecha relación entre el precio de la electricidad y otros precios de la energía. Se supone que todos los agentes son neutrales frente al riesgo y compran al precio estipulado. Se distingue entre dos tipos de consumidores: aquéllos que no tienen fuentes de energía alternativas (demanda contractual cuyo precio está estipulado y garantizado durante el año) y aquéllos que sí (demanda puntual cuyo precio es incierto hasta el comienzo del periodo).

6. Otros ámbitos de aplicación pueden ser:

   - planificación a largo plazo de la expansión de la generación (MIT, 1982), (Stanford, 1989) o de la red de distribución
   - planificación a largo plazo de la expansión de inversiones en elementos de potencia reactiva
   - planificación de capacidad de una línea de producción (Ramos, 1992)
   - gestión de una cartera de valores (Dantzig, 1991c)

Veamos un ejemplo sencillo de planificación lineal estocástica bietapa. Se trata de un problema de planificación óptima de la expansión de la generación, es decir, determinar la inversión óptima en generación. Se dispone de varios tipos de generadores candidatos para satisfacer la demanda durante varios periodos. El primero de los periodos tiene una demanda aleatoria. La decisión de inversión debe ser única para todos los periodos. Además existe una restricción presupuestaria a la inversión total y una potencia mínima a instalar.

```
$TITLE Planificación óptima de la expansión de la generación

SETS
   I generadores           / gen-1 * gen-4 /
   J periodos              / per-1 * per-3 /
   S escenarios de demanda / s-1 * s-3 /

PARAMETERS
   F(i) coste fijo de inversión [€]
   / gen-1 10
     gen-2  7
     gen-3 16
     gen-4  6 /

   PROB(s) probabilidad de cada escenario [p.u.]
   / s-1  0.2
     s-2  0.5
     s-3  0.3 /

   DEM(j) demanda para un escenario [MW]

TABLE V(i,j) coste variable de operación [€/MW]

         per-1  per-2  per-3
gen-1     40     24     4
gen-2     45     27     4.5
gen-3     32     19.2   3.2
gen-4     55     33     5.5

TABLE DEMS(s,j) demanda estocástica [MW]

     per-1 per-2 per-3
s-1    3     3     2
s-2    5     3     2
s-3    7     3     2

SCALARS
   POTMIN potencia mínima a instalar [MW] / 12 /
   PRSPTO límite presupuestario     [€] / 120 /

VARIABLES
   X(i)      potencia a instalar               [MW]
   Y(j,i)    potencia en operación             [MW]
   YS(s,j,i) potencia en operación estocástica [MW]
   COSTE     coste total

POSITIVE VARIABLES X, Y, YS

EQUATIONS
   COST    coste total               [€]
   COSTS   coste total estocástico   [€]
   PRESUP  limitación presupuestaria [€]
   INSMIN  potencia mínima instalada                          [MW]
   BALPOT  potencia en operación menor que instalada          [MW]
```

```
   BALPOTS potencia en operación menor que instalada estocástica [MW]
   BALDEM  balance de demanda                                    [MW]
   BALDEMS balance de demanda estocástico                        [MW] ;

COST  .. COSTE =E= SUM(i, F(i) * X(i))
                 + SUM((j,i),             V(i,j) * Y(j,i))    ;
COSTS .. COSTE =E= SUM(i, F(i) * X(i))
                 + SUM((s,j,i), PROB(s) * V(i,j) * YS(s,j,i)) ;

PRESUP .. SUM(i, F(i) * X(i)) =L= PRSPTO ;

INSMIN .. SUM(i,        X(i)) =G= POTMIN ;

BALPOT(j,i)    .. Y(j,i)    =L= X(i) ;
BALPOTS(s,j,i) .. YS(s,j,i) =L= X(i) ;

BALDEM(j)    .. SUM(i, Y(j,i))    =G= DEM(j)    ;
BALDEMS(s,j) .. SUM(i, YS(s,j,i)) =G= DEMS(s,j) ;

MODEL DETERM / COST,  INSMIN, PRESUP, BALPOT,  BALDEM  / ;
MODEL ESTOCA / COSTS, INSMIN, PRESUP, BALPOTS, BALDEMS / ;

* cada escenario determinista por separado

LOOP (s,
   DEM(j) = DEMS(s,j) ;
   SOLVE DETERM MINIMIZING COSTE USING LP ;
) ;

* escenario de demanda media

DEM(j) = SUM(s, PROB(s) * DEMS(s,j)) ;
SOLVE DETERM MINIMIZING COSTE USING LP ;

* problema estocástico

SOLVE ESTOCA MINIMIZING COSTE USING LP ;
```

El problema se resuelve en primer lugar de forma determinista para cada escenario de demanda por separado. A continuación se resuelve el problema para el valor medio de los parámetros estocásticos y finalmente el problema estocástico y sus resultados se presentan a continuación.

	Determ 1	**Determ 2**	**Determ 3**	**Medio**	**Estocást**
Generador 1	.	0.33	3.67	0.67	0.67
Generador 2	.	.	.	.	2
Generador 3	3	4.67	3.33	4.53	4.33
Generador 4	9	7	5	6.8	5

Coste total	262	346.67	437.33	355.73	362.47

Para la solución de problemas estocásticos se ha recurrido históricamente a diferentes aproximaciones:

- Una es solucionar el problema para cada posible escenario. Ésta se denomina *espera y observa* (*wait and see* o *scenario analysis* o *what-if analysis*), corresponde a tomar las decisiones una vez que se ha resuelto la incertidumbre. A partir de las soluciones se observan sistemáticamente semejanzas y diferencias entre ellas y mediante criterios subjetivos tomar decisiones *flexibles*, es decir, aquéllas cuyos componentes forman parte de las decisiones óptimas bajo numerosos escenarios. Evidentemente este procedimiento no garantiza la obtención de la solución óptima. De hecho la solución óptima estocástica no tiene por qué ser óptima para ningún escenario. Con este criterio el generador 2 en el problema de expansión previo no aparecería puesto que no aparece en la solución para ninguno de los escenarios ni para el de demanda media. Sin embargo, sí aparece en la solución óptima del problema estocástico.
- Otra es la solución del problema estocástico para el valor medio de los parámetros aleatorios convirtiéndolo, por tanto, en un problema determinista.

El paso de estas aproximaciones a la optimización estocástica supone un salto cualitativo en cuanto a complejidad de solución y tamaño del problema. La optimización estocástica se debe utilizar cuando las decisiones pueden no ser utilizables o realizables si no se tiene en cuenta expresamente la incertidumbre. En el ejemplo anterior se aprecia que las decisiones de inversión en las que se está interesado cambian según se haga la hipótesis sobre la ocurrencia de cada escenario futuro o del escenario de demanda media. De hecho, las soluciones deterministas bajo un escenario pueden ser infactibles bajo otro.

Las decisiones en un momento dado dependen exclusivamente de la información disponible en ese momento, no pueden utilizar información futura. Las decisiones de la primera etapa (de inversión) son independientes del escenario que ocurra, se toman con antelación a resolver la incertidumbre. Esta característica se denomina propiedad de *implantabilidad* o *no anticipatividad* de las decisiones.

Los diferentes estados que pueden tomar los parámetros aleatorios a lo largo del tiempo se representan mediante un *árbol de probabilidad* o *de escenarios* (figura 4.1.). Se define un *escenario* como cualquier camino (trayectoria) que va desde la raíz del árbol hasta las hojas de manera que los escenarios que comparten una misma información hasta una cierta etapa también comparten esa parte del árbol. El árbol de probabilidad es la forma natural y explícita de representar la no anticipatividad de las decisiones. En el caso ejemplo previo, el árbol es simplemente una raíz y tres hojas. Las decisiones de la primera etapa se comparten (son las mismas para cualquiera de los escenarios) y por eso la raíz es única. Las decisiones de la segunda etapa son múltiples (dependen de cada escenario) y por eso tiene tres hojas. La determinación del árbol de probabilidad debe considerar las dependencias temporales y/o espaciales que pudieran existir entre los parámetros aleatorios. Si el proceso aleatorio es markoviano la dependencia temporal sólo alcanzará un periodo. En un caso general se puede extender más allá.

En la formulación que se utiliza en este ejemplo y en algunas de las técnicas de descomposición (e.g., descomposición de Benders), se utiliza expresamente la estructura del árbol de probabilidad para definir las variables y los parámetros aleatorios. Así en el ejemplo, las variables de la primera etapa no dependen de los escenarios mientras que los de la segunda sí. Existe otra técnica de resolución alternativa (*scenario aggregation with splitting variable formulation* (Rockafellar)) que considera variables independientes para cada escenario (como se muestra en la figura 4.2) y establece restricciones de no anticipatividad que hacen que ciertas variables sean iguales entre sí en ciertos periodos. En esta técnica se utiliza el método de relajación lagrangiana, especialmente indicado cuando existe un número reducido de restricciones de complicación. Es decir, el árbol de la figura 4.1 se convierte en el de la figura 4.2 mediante restricciones. En el ejemplo previo sería equivalente a haber definido una variable de inversión para cada escenario y haber impuesto restricciones que las obligan a ser iguales.

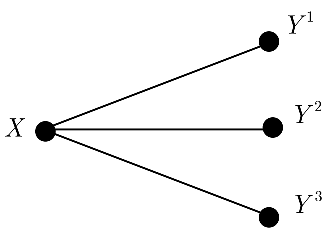


Figura 4.1 Árbol de probabilidad.

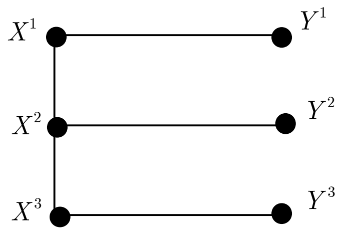


Figura 4.2 Árbol de probabilidad para agregación de escenarios.

La estructura de la matriz de restricciones del problema estocástico se observa en la figura 4.3. En ella se muestran las restricciones que involucran variables de cada etapa por separado y las que unen variables de ambas etapas. El problema estocástico es de tamaño mucho mayor que el de un problema determinista y tiene la estructura no anticipativa del árbol de probabilidad. Las técnicas de descomposición aprovechan esta estructura repetitiva de la matriz de restricciones.

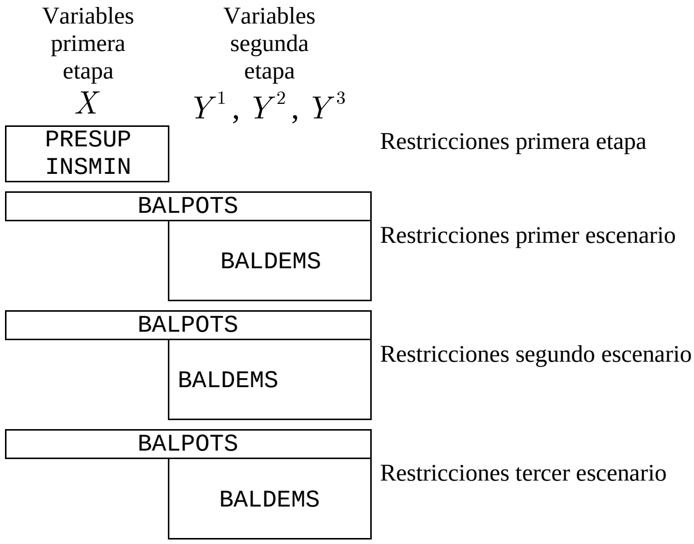


Figura 4.3 Estructura de la matriz de coeficientes de las restricciones.

La función objetivo del problema estocástico es el valor de la función objetivo para las decisiones de la primera etapa más el valor esperado de la función objetivo para las decisiones de la segunda etapa (362.47 en el ejemplo). El *valor de la solución estocástica* (VSS) es la diferencia entre la función objetivo del problema estocástico y la del problema determinista para el valor medio de los parámetros (362.47–355.73=6.73).

Se denomina *valor esperado con información perfecta* a la suma ponderada para cada escenario de la función objetivo total sabiendo que dicho escenario va a ocurrir con certeza (356.93 para el ejemplo). Es decir, se revela la incertidumbre antes de tomar las decisiones de la primera etapa. Este valor siempre será menor o igual, si se está minimizando, que la función objetivo del problema estocástico.

Para cada escenario, la solución del problema estocástico es siempre peor o igual que la solución con información perfecta (280, 349.33 y 439.33 respectivamente). Se denomina *arrepentimiento* a la diferencia entre ambas (280–262=18, 349.33–346.67=2.66, 439.33–437.33=2) y *valor esperado de la información perfecta* (EVPI) a su esperanza.

Cuando esta diferencia en algún escenario pueda ser elevada (de consecuencias catastróficas) o existe gran no linealidad en la función objetivo para los diferentes escenarios la optimización estocástica se puede formular como minimización del máximo arrepentimiento (o criterio minimax). Éste es un criterio conservador de protección frente el riesgo que determina planes *robustos* (es decir, estables frente a perturbaciones en los datos).

Veamos otro ejemplo de un problema de optimización de una compañía de gas con demanda y precio estocásticos. Cuando la compañía compra gas suministra parte a sus consumidores y el resto lo almacena. Cuando vende gas lo toma de sus compras o de sus depósitos. Luego sus decisiones son cuánto gas comprar y vender directamente, cuánto comprar y almacenar, y cuánto tomar del depósito y vender. Las decisiones óptimas dependerán del precio del gas ahora y en el futuro, del coste de almacenamiento, del tamaño del depósito y de la demanda en cada periodo. La demanda anual y el precio del gas para cada escenario se presentan en la siguiente tabla. El precio de almacenamiento de una unidad de gas es de 1 €/año. Se quieren tomar las decisiones óptimas para los dos primeros años sabiendo que el año actual está siendo normal y el año próximo puede ser normal, frío o muy frío con las probabilidades que aparecen en la tabla.

Escenario	**Coste del gas (€)**	**Demanda**	**Probabilidad**
Normal	5.0	100	1/3
Frío	6.0	150	1/3
Muy frío	7.5	180	1/3

```
$TITLE Planificación óptima de la compra de gas

SETS
   S escenarios de temperatura / nr, fr, mf /
   O opciones de gestión    / cyv, cya, ayv /

PARAMETERS
   CVS(s) coste del gas en cada escenario [€]
   / nr 5.0
     fr 6.0
     mf 7.5 /

   PROBS(s) probabilidad de cada escenario [p.u.]
   / nr 0.333333
     fr 0.333333
     mf 0.333334 /

   DEMS(s) demanda en un escenario
   / nr 100
     fr 150
     mf 180 /

   CV   coste del gas
   PROB probabilidad
   DEM  demanda en un escenario

SCALARS
   CALM coste de almacenamiento [€ por año] / 1 /

VARIABLES
   X(o)    cantidad que gestiona el primer  año
   Y(o)    cantidad que gestiona el segundo año
   YS(s,o) cantidad que gestiona el segundo año estocástico
   COSTE   coste total

POSITIVE VARIABLES X, Y, YS

EQUATIONS
   COST     coste total
   COSTS    coste total estocástico
   BALDEM1  balance de demanda del primer  año
   BALDEM2  balance de demanda del segundo año
   BALDEMS2 balance de demanda del segundo año
   GSTDEP   gestión del depósito ;

COST  .. COSTE =E= CVS('nr') * (X('cyv')+X('cya'))
                 + CALM * X('cya')
                 + CV * (Y('cyv')+Y('cya')) ;

COSTS .. COSTE =E= CVS('nr') * (X('cyv')+X('cya'))
                 + CALM * X('cya')
                 + SUM(s, PROBS(s) * CVS(s) * (YS(s,'cyv')+YS(s,'cya'))) ;

BALDEM1 .. X('cyv') =E= DEMS('nr') ;

BALDEM2     .. X('ayv') + Y('cyv')    =E= DEM     ;
BALDEMS2(s) .. X('ayv') + YS(s,'cyv') =E= DEMS(s) ;

GSTDEP .. X('cya') =E= X('ayv') ;

MODEL DETERM / COST , BALDEM1, BALDEM2 , GSTDEP / ;
MODEL PROBAB / COSTS, BALDEM1, BALDEMS2, GSTDEP / ;

* cada escenario determinista por separado

LOOP (s,
   CV   = CVS(s) ;
   DEM  = DEMS(s) ;
```

```
   SOLVE DETERM MINIMIZING COSTE USING LP ;
) ;

* escenario de demanda media

CV  = SUM(s, PROBS(s) * CVS(s))  ;
DEM = SUM(s, PROBS(s) * DEMS(s)) ;
SOLVE DETERM MINIMIZING COSTE USING LP ;

* problema estocástico

SOLVE PROBAB MINIMIZING COSTE USING LP ;
```

Los resultados para cada escenario determinista, el escenario medio y el problema estocástico se presentan a continuación.

	Compra y Vende	Compra y almacena	Depósito a venta	Compra y vende	Compra y almacena	Coste
	Año 1			Año 2		
Normal	100	.	.	100	.	1000
Frío	100	.	.	150	.	1400
Muy frío	100	180	180	.	.	1580
Medio	100	143.33	143.33	.	.	1360
				.	.	
Estocástico	100	100	100	50	.	1400
				80	.	

Para este ejemplo también se observa que las soluciones en cada escenario son claramente diferentes a las del escenario medio y a las decisiones del problema estocástico.

Aunque las primeras investigaciones aparecen en 1955 con trabajos de Dantzig (Dantzig, 1955) y Beale (Véale, 1955), sólo recientemente el avance en la tecnología de los ordenadores ha permitido la solución de problemas de muy gran tamaño y ha devuelto interés al tema de la *optimización estocástica* produciendo además un avance en la teoría matemática que lo sustenta. Los libros de Ermoliev y Wets (eds.), Kall y Walace, Birge y Louveaux y Prékopa (Ermoliev, 1988), (Birge, 1997), (Kall, 1995), (Prekopa, 1995) pueden servir de compendio a la investigación reciente en este campo. La resolución de este tipo de problemas abre nuevas posibilidades en numerosos ámbitos.

La programación estocástica apareció como una extensión de la programación lineal con énfasis en el gran número de variables y parámetros (Véale, 1955), (Dantzig, 1955), (Dantzig, 1963), (Dantzig, 1961). Por otra parte, como otra extensión de la programación lineal para grandes sistemas con estructuras especiales en la matriz de coeficientes de las restricciones aparecieron las técnicas de descomposición (Benders, 1962), (Dantzig, 1963), (Dantzig, 1960), (Van Slyke, 1969), denominadas de *optimización matemática a gran escala*.

Algunas técnicas que permiten la solución de problemas de programación lineal estocástica de gran tamaño son las siguientes:

- *descomposición*

  Las técnicas de descomposición resuelven problemas de muy gran tamaño con una estructura especial, que se aprovecha desde un punto de vista teórico y computacional con el uso de algoritmos especializados. Constituye una forma flexible y elegante de resolver problemas.

- *simulación y reducción de varianza*

  Las técnicas de simulación y reducción de varianza son necesarias cuando la resolución explícita del número de problemas que surgen debido a la aleatoriedad en los parámetros resulta inviable.

La conjunción de varias de estas técnicas fue propuesta por Dantzig (Dantzig, 1987), (Dantzig, 1990) y ha dado resultados espectaculares por los tamaños de los problemas de planificación resueltos (Dantzig, 1991a), (Dantzig, 1991c), (Entriken, 1990), (Infanger, 1991), (Infanger, 1994), (Stanford, 1989).

Las técnicas de simulación y descomposición utilizadas en la solución de estos problemas de optimización de muy gran tamaño hacen que el uso de procesamiento paralelo o distribuido[14] sea muy conveniente (Dantzig, 1987), (Dantzig, 1990). Con ellas

[14] Las técnicas de descomposición son de uso matemático general y tienen sentido independientemente de la tecnología del ordenador/es utilizados en su resolución. Sin embargo, dada su naturaleza resulta muy conveniente la utilización de cálculo en paralelo (un ordenador con múltiples CPUs débil o fuertemente acopladas) o cálculo distribuido (múltiples ordenadores trabajando en colaboración) para reducir los tiempos de cálculo de manera sustancial.

la solución repetida de los subproblemas o maestros se puede efectuar en paralelo entre los diferentes procesadores disponibles (Entriken, 1988), (Entriken, 1989), (Ruszczynski, 1993).

Actualmente existe una herramienta para la solución de problemas lineales estocásticos bietapa de gran tamaño denominada DECIS que se puede ejecutar en un ordenador paralelo (Dantzig, 1991b) o convencional (Infanger, 1991). Las reducciones de tiempo con respecto a su solución en ordenador convencional logradas para este tipo de problema son del orden de la mitad del número de procesadores utilizados. Por ejemplo, con un problema concreto resuelto con 64 procesadores en paralelo tomando 600 muestras en la simulación el tiempo de cálculo se redujo en 37.5 veces. Este decremento depende del número de muestras y de procesadores.

Otra aplicación denominada MSLiP permite la solución de problemas lineales estocásticos multietapa de tamaño medio para su uso en un ordenador convencional (Gassmann, 1990). Existe también una versión estocástica OSL-SE del optimizador de programación lineal entera mixta OSL. En (Ermoliev, 1988) se citan además otros códigos de programación lineal estocástica.

La formulación de problemas estocásticos se está empezando a realizar también con lenguajes algebraicos de modelado como GAMS (Ramos, 1998) o AMPL (Gassmann, 1995), (Gassmann, 1996). Tanto DECIS como OSL-SE están disponibles con GAMS.

IV.2. Optimización lineal bietapa y multietapa

El *problema lineal bietapa* PL-2 se representa matemáticamente de la forma siguiente, considerando el vector x_1 las variables de la primera etapa y x_2 las variables de la segunda:

$$\begin{aligned} &\min_{x_1,x_2}(c_1^T x_1 + c_2^T x_2) \\ &A_1 x_1 \qquad\qquad = b_1 \\ &B_1 x_1 \quad + A_2 x_2 \quad = b_2 \\ &x_1, \qquad x_2 \qquad \geq 0 \end{aligned} \tag{1.161}$$

donde la matriz A_1 es $m_1 \times n_1$ y A_2 es $m_2 \times n_2$. Las dimensiones de los demás vectores y matrices se derivan de éstas. Por ser optimización determinista se suponen conocidas las matrices y vectores. El tamaño del problema completo es $(m_1 + m_2) \times (n_1 + n_2)$. La estructura de la matriz de restricciones se presenta en la figura 4.4.

$$
\begin{array}{|c|c|}
\hline
A_1 & \multicolumn{1}{c}{} \\
\hline
B_1 & A_2 \\
\hline
\end{array}
$$

Figura 4.4 Estructura de la matriz de coeficientes de las restricciones en PL-2.

Los dos ejemplos del capítulo anterior corresponden a problemas lineales bietapa, deben proporcionar decisiones antes de resolver la incertidumbre, siendo la segunda etapa estocástica. Un problema de planificación estática (para un momento fijo en el tiempo) se formula frecuentemente como minimización de una función objetivo suma de costes totales de inversión y explotación sujeta a restricciones propias de inversión y de explotación. Las decisiones de inversión corresponden a la primera etapa y las de explotación a la segunda.

Los problemas de planificación dinámica deben proporcionar decisiones óptimas para momentos discretos del tiempo. Éstos se conocen genéricamente como *problemas lineales multietapa* PL-P. Aparecen en el estudio de procesos dependientes del tiempo donde las actividades de un periodo están conectadas con las de los periodos anterior pero no con otros (Dantzig, 1963). Para P etapas el problema se puede formular como:

$$
\begin{aligned}
&\min_{x_p} \sum_{p=1}^{P} c_p^T x_p \\
&B_{p-1} x_{p-1} + A_p x_p = b_p \quad p = 1, \ldots, P \\
&x_p \geq 0 \\
&B_0 \equiv 0
\end{aligned}
\tag{1.162}
$$

donde las matrices A_p y B_p y los vectores b_p y c_p son conocidos por ser optimización determinista y x_p son las variables. La matriz A_p es $m_p \times n_p$. El problema completo tiene $m = \sum_{p=1}^{P} m_p$ filas y $n = \sum_{p=1}^{P} n_p$ columnas.

Característica de estos problemas es una estructura de la matriz de coeficientes de las restricciones en escalera, tal como se presenta en la figura 4.5.

$$
\begin{array}{|c|c|c|c|c|}
\hline
A_1 & & & & \\
\hline
B_1 & A_2 & & & \\
\hline
 & B_2 & A_3 & & \\
\hline
 & & \ddots & & \\
 & & & B_{P-1} & A_P \\
\hline
\end{array}
$$

Figura 4.5 Estructura de la matriz de coeficientes de las restricciones en problemas lineales multietapa.

El tamaño del problema crece linealmente con el número de etapas. Aunque para problemas pequeños o medianos se pueden utilizar técnicas convencionales de programación lineal, esto resulta inviable para problemas de gran tamaño. Por ello, para su resolución se emplean técnicas de descomposición que aprovechan su especial estructura.

IV.3. Optimización lineal estocástica bietapa

El *problema lineal estocástico bietapa* PLE-2, donde se considera aleatoriedad en la segunda etapa, se formula matemáticamente como[15]:

$$
\begin{array}{lll}
\min_{x_1, x_2^\omega} c_1^T x_1 + \sum_{\omega \in \Omega} p^\omega c_2^{\omega T} x_2^\omega & & \\
A_1 x_1 & & = b_1 \\
B_1^\omega x_1 & + A_2^\omega x_2^\omega & = b_2^\omega \\
x_1, & x_2^\omega & \geq 0
\end{array}
\tag{1.163}
$$

donde función objetivo corresponde a las decisiones de la primera etapa más el valor esperado correspondiente a las decisiones de la segunda etapa y donde $\omega \in \Omega$ toma un

[15] El superíndice ω indica la dependencia de las matrices y vectores con respecto al valor de ω.

número finito, $|\Omega|$ (cardinal de Ω), de valores cada uno con probabilidad p^ω. Si A_2^ω es independiente de ω este problema se denomina con *recurso fijo*. La estructura de la matriz de restricciones de un problema PLE-2 para tres escenarios en la segunda etapa se representa en la siguiente figura.

$$
\begin{array}{|c|c|c|c|}
\hline
A_1 & & & \\
\hline
B_1^1 & A_2^1 & & \\
\hline
B_1^2 & & A_2^2 & \\
\hline
B_1^3 & & & A_2^3 \\
\hline
\end{array}
$$

Figura 4.14 Estructura de la matriz de coeficientes de las restricciones en PLE-2.

El problema de minimización de máximo arrepentimiento tiene una estructura similar al anterior y se formula como:

$$
\begin{aligned}
&\min_{\alpha, x_1, x_2^\omega} \alpha \\
&\alpha \ - c_1^T x_1 - p^\omega c_2^{\omega T} x_2^\omega \geq -f^\omega \quad \forall \omega \in \Omega \\
&A_1 x_1 \qquad\qquad\quad = b_1 \\
&B_1^\omega x_1 \quad + A_2^\omega x_2^\omega \quad = b_2^\omega \\
&\ x_1, \qquad\ x_2^\omega \qquad \geq 0
\end{aligned}
\tag{1.164}
$$

siendo f^ω el valor óptimo de la función objetivo para cada escenario ω (i.e., la función objetivo para la decisión óptima sabiendo que va a ocurrir un determinado escenario ω).

El uso de técnicas de descomposición se justifica por la estructura especial que presenta la matriz de restricciones del problema completo. Los subproblemas son separables para cada escenario y tienen la misma estructura en las restricciones.

IV.4. Optimización lineal estocástica multietapa

El *problema lineal estocástico multietapa* PLE-P se puede formular como:

$$
\begin{aligned}
&\min_{x_p^{\omega_p}} \sum_{p=1}^{P} \sum_{\omega_p \in \Omega_p} p_p^{\omega_p} c_p^{\omega_p T} x_p^{\omega_p} \\
&B_{p-1}^{\omega_p} x_{p-1}^{\omega_{p-1}} + A_p^{\omega_p} x_p^{\omega_p} = b_p^{\omega_p} \qquad p = 1,\ldots,P \\
&x_p^{\omega_p} \geq 0 \\
&B_0^{\omega_1} \equiv 0
\end{aligned}
\tag{1.165}
$$

donde Ω_p representa el conjunto de los posibles valores de ω_p en la etapa p cada uno con probabilidad de ocurrencia $p_p^{\omega_p}$. Estas probabilidades son, en general, dependientes de los valores tomados en la etapa anterior. Cuando se van a tomar las decisiones al comienzo de la etapa p los valores de $c_p^{\omega_p}$, $B_{p-1}^{\omega_p}$, $A_p^{\omega_p}$, $b_p^{\omega_p}$ son conocidos y se conocen las distribuciones condicionadas de los vectores para las etapas futuras $p+1,\ldots,P$. Los diferentes valores de los parámetros en las sucesivas etapas forman una estructura en árbol.

Si $\left|\Omega_p\right|$ representa el cardinal de Ω_p, el problema lineal completo tiene $\sum_{p=1}^{P} m_p \left|\Omega_p\right|$ restricciones y $\sum_{p=1}^{P} n_p \left|\Omega_p\right|$ variables, siendo $A_p^{\omega_p}$ una matriz $m_p \times n_p$. La matriz de restricciones de un problema PLE-3 con dos escenarios en la segunda etapa y cuatro en la tercera se representa en la siguiente figura.

$$
\begin{array}{|c|c|c|c|c|c|c|}
A_1 & & & & & & \\
B_1^1 & A_2^1 & & & & & \\
B_1^2 & & A_2^2 & & & & \\
 & B_2^1 & & A_3^1 & & & \\
 & B_2^2 & & & A_3^2 & & \\
 & & B_2^3 & & & A_3^3 & \\
 & & B_2^4 & & & & A_3^4
\end{array}
$$

Figura 4.15 Estructura de la matriz de coeficientes de las restricciones en problemas lineales multietapa.

El método de resolución es descomposición anidada estocástica de Benders (también denominado en la literatura *programación dinámica dual estocástica* (Pereira, 1991)) teniendo presente que los subproblemas para cada etapa son separables.

IV.5. Referencias

Beale, E.M.L. "On Minimizing a Convex Function Subject to Linear Inequalities" *Journal of the Royal Statistical Society*. Vol 17b, pp 173-184. 1955.

Benders, J.F. "Partitioning Procedures for Solving Mixed-Variable Programming Problems" *Numerische Mathematik*. Vol 4, pp 238-252. 1962.

Birge, J.R. and Louveaux, F. (1997) *Introduction to Stochastic Programming*. Springer-Verlag.

Dantzig, G.B. "Linear Programming Under Uncertainty" *Management Science*. Vol 1, No 3-4, pp 197-206. April-July 1955.

Dantzig, G.B. *Linear Programming and Extensions*. Princeton University Press. Princeton, USA. 1963.

Dantzig, G.B. "Planning Under Uncertainty Using Parallel Computing" *Systems Optimization Laboratory. Department of Operations Research. Stanford University*. SOL 87-1. January 1987.

Dantzig, G.B. and Glynn, P.W. "Parallel Processors for Planning Under Uncertainty" *Annals of Operations Research*. Vol 22, pp 1-21. 1990.

Dantzig, G.B. and Infanger, G. "Large-Scale Stochastic Linear Programs: Importance Sampling and Benders Decomposition" *Systems Optimization Laboratory. Department of Operations Research. Stanford University*. SOL 91-4. March 1991.

Dantzig, G.B., Ho, J.K. and Infanger, G. "Solving Stochastic Linear Programs on a Hypercube Multicomputer" *Systems Optimization Laboratory. Department of Operations Research. Stanford University*. SOL 91-10. August 1991.

Dantzig, G.B. and Infanger, G. “Multi-Stage Stochastic Linear Programs for Portfolio Optimization” *Systems Optimization Laboratory. Department of Operations Research. Stanford University*. SOL 91-11. September 1991.

Dantzig, G.B. and Madansky, A. “On the Solution of Two-Stage Linear Programs Under Uncertainty” *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*. University of California Press. Berkeley. Vol I, pp 165-176. 1961.

Dantzig, G.B. and Wolfe, P. “Decomposition Principle for Linear Programs” *Operations Research*. Vol 8, No 1, pp 101-111. January-February 1960.

Entriken, R. “A Parallel Decomposition Algorithm for Staircase Linear Programs” *Systems Optimization Laboratory. Department of Operations Research. Stanford University*. SOL 88-21. December 1988.

Entriken, R. “The Parallel Decomposition of Linear Programs” *Systems Optimization Laboratory. Department of Operations Research. Stanford University*. SOL 89-17. November 1989.

Entriken, R. and Infanger, G. “Decomposition and Importance Sampling for Stochastic Linear Models” *IBM Research Report*. RC 15545 (#69107). March 1990.

Ermoliev, Y. and Wets, R.J-B. (eds.) *Numerical Techniques for Stochastic Optimization*. Springer-Verlag. Berlin, Germany. 1988.

Gassmann, H.I. “MSLiP: A Computer Code for the Multistage Stochastic Linear Programming Problem” *Mathematical Programming*. Vol 47, pp 407-423. August 1990.

Gassmann, H.I. and Ireland, A.M. “Scenario Formulation in an Algebraic Modelling Language” *Annals of Operations Research*. Vol 59, pp 45-75. 1995.

Gassmann, H.I. and Ireland, A.M. “On the Formulation of Stochastic Linear Programs using Algebraic Modelling Languages” *Annals of Operations Research*. Vol 64, pp 83-112. 1996.

Infanger, G. “Monte Carlo (Importance) Sampling within a Benders Decomposition Algorithm for Stochastic Linear Programs. Extended Version: Including Results of Large-Scale Problems” *Systems Optimization Laboratory. Department of Operations Research. Stanford University*. SOL 91-6. March 1991.

Infanger, G. *Planning Under Uncertainty: Solving Large-Scale Stochastic Linear Programs* Boyd Fraser. 1994.

Jacobs, J, Freeman, G., Grygier, J., Morton, D. P., Schultz, G., Staschus, K. and Stedinger, J. “SOCRATES: A System for Scheduling Hydroelectric Generation under Uncertainty” *Annals of Operations Research*. Vol 59, pp 99-133. 1995.

Kall, P. and Wallace, S.W. (1995) *Stochastic Programming*. John Wiley and Sons.

Massachusetts Institute of Technology (MIT) “Electric Generation Expansion Analysis System. Volume 1: Solution Techniques, Computing Methods, and Results. Volume 2: Details of Solution Techniques, Data of Test Systems, and Glossary of Terms” *Electric Power Research Institute*. EPRI EL-2561. August 1982.

Mathiesen, L. “Pricing of Electricity in the Presence of Water Uncertainty” February 1992.

Mayer, J. (1998) *Stochastic Linear Programming Algorithms. A Comparison Based on a Model Management System.* Gordon and Breach Science Publishers.

Morton, D. P. “Algorithmic Advances in Stochastic Programming” Systems Optimization Laboratory. Department of Operations Research. Stanford University. SOL 93-6. July 1993.

Morton, D. P. “An Enhanced Decomposition Algorithm for Multistage Stochastic Hydroelectric Scheduling” *Annals of Operations Research*. Vol 64, pp 211-235. 1996.

Murtagh, B.A. and Saunders, M.A. “MINOS 5.1 User's Guide” *Systems Optimization Laboratory. Department of Operations Research. Stanford University*. SOL 83-20R. December 1983. Revised January 1987.

Pereira, M.V.F. and Pinto, L.M.V.G. "Multi-Stage Stochastic Optimization Applied to Energy Planning" *Mathematical Programming*. Vol 52, pp 359-375. August 1991.

Pérez Arriaga, J.I., Ramos, A. y Latorre, G. "Volumen 1: Guía de usuario del programa de planificación estática de la red de transporte a largo plazo. PERLA. Volumen 2: Casos ejemplo" *Instituto de Investigación Tecnológica*. Julio 1991.

Prékopa, A. *Stochastic Programming*. Kluwer Academic Publishers, 1995.

Ramos, A. "Integrated Model for Capacity Planning for Manufacturing Systems" March 1992.

Rockafellar, R.T. and Wets, R.J.B. "Scenarios and Policy Aggregation in Optimization Under Uncertainty" Mathematics in Operations Research. Vol 16, pp 119-147.

Ruszczynski, A. "Parallel Decomposition of Multistage Stochastic Programming Problems" *Mathematical Programming*. Vol 58, pp 201-228. 1993.

Stanford University "Decomposition Techniques for Multi-Area Generation and Transmission Planning Under Uncertainty" *Electric Power Research Institute*. EPRI EL-6484. August 1989.

Van Slyke, R.M. and Wets, R.J.-B. "L-Shaped Linear Programs with Application to Optimal Control and Stochastic Programming" *SIAM Journal on Applied Mathematics*. Vol 17, No 4, pp 638-663. July 1969.

IV.6. Biblioteca de problemas

PROBLEMA 1

Considérese una explotación agrícola en la que se puede cosechar trigo, maíz y azúcar de remolacha. La superficie dedicada al cultivo es de 500 Ha, y hay que decidir cuánta superficie se va a dedicar a cada uno de los tres tipos de cultivo; esta decisión debe efectuarse en invierno para proceder a la correspondiente siembra. La explotación agrícola tiene, además, una ganadería, y para su cría y cuidado se requieren 200 Tm de trigo y 240 Tm de maíz, que pueden ser producidas en la propia explotación o compradas en el exterior. Una vez atendidas las necesidades de la explotación, la producción en exceso de estos dos cultivos se destina a la venta, pudiéndose vender a precio de mercado, 170 y 150 euros/Tm, respectivamente para trigo y maíz. El otro tipo de cosecha es azúcar de remolacha, que se puede vender a 36 euros/Tm. Sin embargo, la Comisión Europea impone una cuota máxima de producción y la cuota asignada a la explotación es de 6000 Tm; cualquier producción en exceso sólo puede ser vendida a 10 euros/Tm.

El coste de cultivo es de 150, 230 y 260 euros/Ha para cada uno los cultivos, maíz, trigo y azúcar, respectivamente. Si es necesario comprar trigo o maíz para satisfacer la demanda mínima, los precios de compra tienen un 40% de incremento sobre los precios de mercado debido a costes de transporte y otros.

Basado en la experiencia de años anteriores, el responsable de la explotación agrícola considera que el rendimiento esperado por Ha es de 2.5 Tm, 3 Tm y 20 Tm para trigo, maíz y remolacha de azúcar, respectivamente. La tabla 1 recoge todos estos datos.

	Trigo	Maíz	Remolacha
Cosecha (Tm/Ha)	2.5	3	20
Costo de plantación por Ha	150	230	260
Precio de venta por Tm	170	150	36(menos de 6000 Tm)
			10 (más de 6000 Tm)
Precio de compra por Tm	283	210	--
Requerimiento mínimo	200	240	--

1. Plantear y resolver un modelo de Programación Lineal que sirva para determinar la superficie que debe dedicarse a cada tipo de cultivo de modo que se satisfagan las condiciones del problema y se maximice el beneficio total.

2. Por otra parte, la experiencia indica que en el pasado se han obtenido rendimientos muy distintos al considerado como rendimiento esperado o promedio, debido a los

cambios climáticos. La posible variación del rendimiento puede representarse bajo diversos tipos de escenarios. Supóngase, para simplificar el problema, cierta correlación entre los rendimientos de los tipos de cultivo y que la temporada puede ser buena, normal o mala para todas las cosechas. La temporada buena proporciona unos rendimientos un 20% superiores al promedio y la temporada mala un 20% inferiores al promedio. Por último, se supone que el rendimiento agrícola no tiene influencia sobre los precios. Analizar la sensibilidad de la solución óptima a la variación del rendimiento.

3. Plantear y resolver el problema estocástico asociado a la situación descrita en el apartado b).

PROBLEMA 2

Una compañía aérea tiene que decidir cómo particionar un nuevo avión que cubrirá el vuelo Madrid-Argentina. El avión tiene capacidad para 200 pasajeros de clase turista. Se puede reservar una parte del avión para asientos de primera clase, pero cada uno de estos asientos ocupa el espacio de 2 asientos de clase turista. También se puede incluir una sección de clase business, en la que cada asiento ocupa el lugar de 1,5 de clase turista. El beneficio neto por un billete de primera es triple del de uno de clase turista, y el beneficio de un billete business es doble del de uno de clase turista.

Una vez que el avión ha sido particionado, esta partición no puede ser modificada entre vuelos. La compañía sabe, sin embargo, que el avión no siempre se llena en todas sus clases. Han decidido considerar tres escenarios equiprobables: (i) Vuelos matinales entre semana. Piensan que pueden vender 20 billetes de primera, 50 business y 200 clase turista. (ii) Vuelos fin de semana. 10 de primera, 20 business y 75 clase turista. (iii) Vuelos vespertinos entre semana. 5 primera, 10 business y 150 clase turista. Considerando que no se permite overbooking (vender más billetes de una clase que asientos haya),

a) Modelizar el problema lineal determinístico correspondiente al primer escenario (vuelos matinales entre semana).

b) Modelizar el problema estocástico que se le presenta a la compañía,

- Utilizando representación compacta.
- -Utilizando la representación con variables divididas.

c) Resolver el problema estocástico.

d) Calcular el valor de la información perfecta y el valor de la solución estocástica

PROBLEMA 3

Tomates S.A. produce pasta de tomate, ketchup y salsa a partir de cuatro recursos: mano de obra, tomates, azúcar y especias. Cada caja de pasta de tomate necesita 0.5 horas de trabajo, 1 caja de tomates, nada de azúcar y 0.25 latas de especias. Cada bote de ketchup necesita 0.8 horas, 0,5 cajas de tomates, 0,5 paquetes de azúcar y 1 lata de especias. Cada sobre de salsa de tomate necesita 1 hora de trabajo, 0.5 cajas de tomates, 1 paquete de azúcar y 3 latas de especias.

La compañía debe decidir la producción para el próximo periodo. Sólo tiene disponibles 200 horas de trabajo, 250 cajas de tomates, 300 paquetes de azúcar y 100 latas de especias en cada periodo. Sin embargo, puede conseguir más recursos pagándolos a 2 euros/hora, 0.5 euros/caja tomate, 1 euro/paquete azúcar y 1 euro/lata especias. Los costes unitarios de producción son de 1 euro /caja de pasta de tomate, 1,5 euro/bote de ketchup y 2.5 euros/sobre de salsa.

La demanda no se conoce hasta que los productos están elaborados. Tomates S.A. espera dos situaciones equiprobables posibles, correspondientes a un buen o mal momento económico. Espera vender 200 cajas de pasta, 40 botes de ketchup y 20 sobres de salsa en el primer caso, y 100, 30, y 5 en el segundo, respectivamente. La producción sobrante puede almacenarse hasta el periodo siguiente con un coste de 0.5, 0.25 y 0.2 euros por unidad de pasta, ketchup y salsa, respectivamente. Existen también costes estimados de 2, 3 y 6 euros por unidad de demanda no satisfecha de pasta, ketchup y salsa, respectivamente.

a) Modelizar el problema lineal determinístico correspondiente al primer escenario (buen momento económico).

b) Modelizar el problema estocástico de recurso completo utilizando representación compacta y representación vía variables divididas.

c) Resolver el problema estocástico.

d) Calcular el valor de la información perfecta y el valor de la solución estocástica

IV.7. Resultados de la biblioteca de problemas

Problema 1

a)

$$\max \quad \sum_j \left(pv_j \left(Q_j + S_j - \min_j \right) - c_j H_j - pc_j S_j \right) - (pv_{RML} - pexc) * REM_2$$

$$\begin{aligned}
& Q_j = r_j H_j \quad \forall j \\
& Q_j + S_j \geq \min_j \quad \forall j \\
& \sum_j H_j \leq 500 \\
& Q_{RML} = REM_1 + REM_2 \\
& REM_1 \leq 6000
\end{aligned}$$

Solución: Beneficio: 118600

Q de trigo 300, maíz 240, remolacha 6000

H de trigo 120, maíz 80, remolacha 300, y no compra nada fuera.

b) Para cada escenario:

s1 (alto) : Beneficio: 167666.667

Q de trigo 550, maíz 240, remolacha 6000

H de trigo 183.333, maíz 66.667, remolacha 250, y no compra nada fuera.

S3 (bajo) : Beneficio: 59950

Q de trigo 200, maíz 60, remolacha 6000

H de trigo 100, maíz 25, remolacha 375, y compra 180 Tm de maíz.

c) Estocástico:

$$\max \quad \sum_s pr_s \sum_j \left(pv_j \left(QS_{sj} + SS_{sj} - \min_j \right) - c_j H_j - pc_j SS_{sj} \right) - (pv_{RML} - pexc) * REMS_{s2}$$
$$QS_{sj} = rs_{sj} H_j \quad \forall j, s$$
$$QS_{sj} + SS_{sj} \geq \min_j \quad \forall j, s$$
$$\sum_j H_j \leq 500$$
$$QS_{s,RML} = REMS_{s1} + REMS_{s2} \quad \forall s$$
$$REMS_{s1} \leq 6000 \quad \forall s$$

Solución: Beneficio esperado: 108281.61

	Trigo	Maíz	Remolacha
Hectáreas	170	80	250
Q s1	510	288	6000
Q s2	425	240	5000
Q s3	340	192	4000
SS s1	0	0	0
SS s2	0	0	0
SS s3	0	48	0

Problema 2

$$\max \quad \sum_j b_j O_j$$

a)
$$O_j \leq AS_j \quad \forall j$$
$$O_j \leq d_j \quad \forall j$$
$$\sum_j e_j AS_j \leq 200$$

b)
$$\max \quad \sum_s pr_s \sum_j b_j O_{sj}$$
$$O_{sj} \leq AS_j \quad \forall j, s$$
$$O_{sj} \leq d_{sj} \quad \forall j, s$$
$$\sum_j e_j AS_j \leq 200$$

$$\max \quad \sum_s pr_s \sum_j b_j O_{sj}$$
$$O_{sj} \leq AS_{sj} \quad \forall j, s$$
$$O_{sj} \leq d_{sj} \quad \forall j, s$$
$$\sum_j e_j AS_{sj} \leq 200 \quad \forall s$$
$$AS_{s_1 j} = AS_{s_2 j} = AS_{s_3 j}$$

c) Estocástico: Beneficio esperado: 183.333

	Primera	Business	Turista

Asientos	10	20	150
Ocup mat	10	20	150
Ocup finde	10	20	75
Ocup vesp	5	10	150

d) Valor esperado con Información perfecta: 193.333

Matinal Benef: 250 P: 20 B: 50 T: 85
Finde Benef: 145 P: 10 B: 20 T: 75
Vesp Benef: 185 P: 5 B: 10 T: 150
Valor de la solución estocástica (media – estoc) = 224.975 – 183.333 = 41.642
Valor esperado de la información perfecta (arrepentimiento medio):
1/3 (250- 220) + 1/3 (145-145) + 1/3 (185-185) = 193.333-183.333=10
Problema 3

$$\min \sum_j \left(cprod_j Q_j + calm_j E_j + cdns_j F_j \right) + \sum_i cexc_i R_i$$

a)
$$Q_j = d_j - F_j + E_j \quad \forall j$$
$$\sum_j req_{ji} Q_j \le \lim_i + R_i \quad \forall i$$

$$\min \sum_j cprod_j Q_j + \sum_i cexc_i R_i + \sum_s pr_s \sum_j \left(calm_j E_{sj} + cdns_j F_{sj} \right)$$

b)
$$Q_j = d_{sj} - F_{sj} + E_{sj} \quad \forall j,s$$
$$\sum_j req_{ji} Q_j \le \lim_i + R_i \quad \forall i$$

$$\min \sum_s pr_s \left(\sum_j \left(cprod_j Q_{sj} + calm_j E_{sj} + cdns_j F_{sj} \right) + \sum_i cexc_i R_{si} \right)$$
$$Q_{sj} = d_{sj} - F_{sj} + E_{sj} \quad \forall j,s$$
$$\sum_j req_{ji} Q_{sj} \le \lim_i + R_{si} \quad \forall i,s$$
$$Q_{s_1 j} = Q_{s_1 j} \quad \forall j \qquad R_{s_1 i} = R_{s_2 i} \quad \forall i$$

c) Estocástico: Coste esperado: 313.5

	Pasta	Ketchup	Salsa
Producción	100	30	15
Bueno (almac, falta)	(0,100)	(0,10)	(0,5)
Bueno (almac, falta)	(0,0)	(0,0)	(10,0)

No compra nada de ningún recurso

d) Valor esperado con Información perfecta: 258.75

Bueno Coste: 360 P: 200 K: 40 S: 20 (compra 50 especias, justo dem)

Malo   Coste: 157.5   P: 100 B: 30   T: 5   (no compra nada, justo dem)

Valor de la solución estocástica -(media – estoc) = 313.5 – 243.75 = 71.75
Valor esperado de la información perfecta (arrepent. medio) = 313.5 - 258.75 = 54.75

V. Programación dinámica

V.1. Introducción

Hay una gran variedad de situaciones de muy diversa naturaleza que se pueden modelar como procesos de decisión secuenciales de tipo polietápico, para los cuales el método conocido como programación dinámica (*dynamic programming* DP) se muestra como una herramienta eficaz en la mayoría de los casos.

A fin de plantear el problema en líneas generales, podemos suponer un sistema físico S cuyo estado en una *etapa, fase* o *instante k* cualquiera viene especificado por un vector x_k, a cuyas componentes daremos el nombre de *variables de estado*, las cuales pueden evolucionar en el tiempo de forma discreta o continua, dependiendo en cada instante dicha evolución bien de la opción adoptada por un decisor, bien de una intervención humana seguida de una intervención del azar o viceversa. Dichos cambios se hacen mediante un vector u_k a cuyas componentes daremos el nombre de *variables de control* o *de decisión*.

La programación dinámica es una metodología matemática orientada a la solución de problemas con decisiones divisibles o separables en k etapas sucesivas[16] donde se debe minimizar el coste total (o cualquier otra función objetivo) de dichas decisiones. En cada etapa se valora no sólo el *coste actual* o *inmediato* de tomar una decisión sino el *coste futuro* que se origina a partir de ella. Cada etapa se caracteriza por x_k estados (situaciones en que puede encontrarse el sistema en cada etapa). El número de estados puede ser finito o infinito. Cada estado guarda toda la información necesaria (i.e., variables de estado) para tomar las decisiones futuras sin necesidad de conocer cómo se

[16] Estas etapas muchas veces se hallan asociadas a tiempo, de ahí el nombre de programación dinámica, pero esto no es necesario.

ha alcanzado dicho estado. Mediante una decisión u_k en la etapa k se va de un estado x_k al comienzo de la etapa k a otro estado x_{k+1} al comienzo de la siguiente. En cada etapa se evalúa la decisión óptima para cada uno de sus estados.

Variables de estado de etapa k		Decisiones		Variables de estado de etapa $k+1$
$\boxed{x_k}$	$\rightarrow$	u_k	$\rightarrow$	$\boxed{x_{k+1}}$

La programación dinámica es una técnica de divide y vencerás. Resuelve el problema de optimización del conjunto de todas las etapas mediante un *procedimiento recursivo* que se resuelve de manera iterativa, incorporando cada vez una etapa, es decir, partes cada vez mayores del problema original. El procedimiento puede hacerse *hacia delante* (*forward* DP) o *hacia atrás* (*backward* DP). La recursión hacia atrás define la política óptima en la etapa k conocida la política óptima en cualquier estado de la etapa $k+1$, produciendo un modelo del tipo, por ejemplo,

$$f_k^*(x_k) = \min_{u_k} \left\{ c_{x_k u_k} + f_{k+1}^*(x_{k+1}) \right\} \tag{1.166}$$

siendo x_k el estado actual en la etapa k, x_{k+1} el estado al que se llega en la etapa $k+1$ dependiente del estado inicial x_k y de la decisión u_k en la etapa k, $f_k(x_k)$ el valor acumulado de la función objetivo para el estado x_k desde la etapa k hasta la N y $c_{x_k u_k}$ el valor inmediato de la función objetivo al tomar la decisión u_k desde el estado x_k.

El conjunto de estados en cada etapa debe estar dentro de un conjunto de valores admisibles y las decisiones que permiten el paso de un estado en una etapa a otro en la siguiente también puede estar sometido a un conjunto de restricciones. El problema en cada etapa se supone determinista, es decir, se sabe con certeza cuál es la transición entre los estados y su coste. La función objetivo no está restringida a ser lineal, también puede ser no lineal. La programación dinámica tiene la “maldición” de la dimensionalidad debido a la explosión de estados que se necesita analizar en cada etapa. Esto hace que computacionalmente resulte muy costosa de resolver. Como contrapartida la programación dinámica proporciona mucha información. Por ejemplo, la función de coste futuro en cada etapa para cada uno de los estados posibles de dicha etapa y, además, no sólo la decisión óptima sino otras soluciones cuasióptimas con los valores correspondientes de la función objetivo.

De forma completa el problema de programación dinámica de naturaleza aditiva se puede formular como

$$\begin{aligned}
&\min_{u_k} J = f_0(x_0) = \sum_{k=0}^{N} c_{x_k u_k} \\
&x_{k+1} = \phi_k(x_k, u_k) \\
&g_{k,i_k}(x_k, u_k) = 0 \\
&x_k \in X_k,\ u_k \in U_k
\end{aligned} \tag{1.167}$$

siendo x_0 el estado en la etapa inicial conocido, X_k el conjunto de estados posibles en la etapa k, U_k el conjunto de decisiones factibles en la etapa k, $g_{k,i_k}(x_k, u_k) = 0$ las restricciones o ligaduras del sistema en esa etapa y $\phi_k(x_k, u_k)$ la ecuación que gobierna la transición o paso del sistema de la etapa *k* a la siguiente. Obviamente, no todos los modelos susceptibles de ser tratados mediante programación dinámica son de naturaleza aditiva.

Al conjunto de decisiones tomadas a lo largo del tiempo le daremos el nombre de *política* en el caso de situaciones deterministas y el de *estrategia* cuando se hace intervenir el azar. El objetivo en todo proceso de decisión polietápico del tipo que vamos a analizar es el de determinar la política o estrategia que optimiza una función criterio.

En lo que sigue supondremos que los posibles cambios sólo pueden tener lugar en instantes discretos del tiempo, no forzosamente equidistantes. Al intervalo de tiempo transcurrido entre la adopción de dos decisiones consecutivas es lo que hemos llamado fase o etapa. En la práctica la dificultad radica en los aspectos siguientes:

- la identificación de las fases o etapas, junto con las variables de estado y las de decisión
- la formulación del criterio que se pretende optimizar y de las restricciones en términos de dichas variables

La metodología de la programación dinámica es difícil de explicar en un contexto puramente teórico. A efectos de una introducción, que permita una posterior profundización en el tema, es mucho más eficaz presentar dicha metodología a partir de ejemplos como los que se exponen en las secciones siguientes.

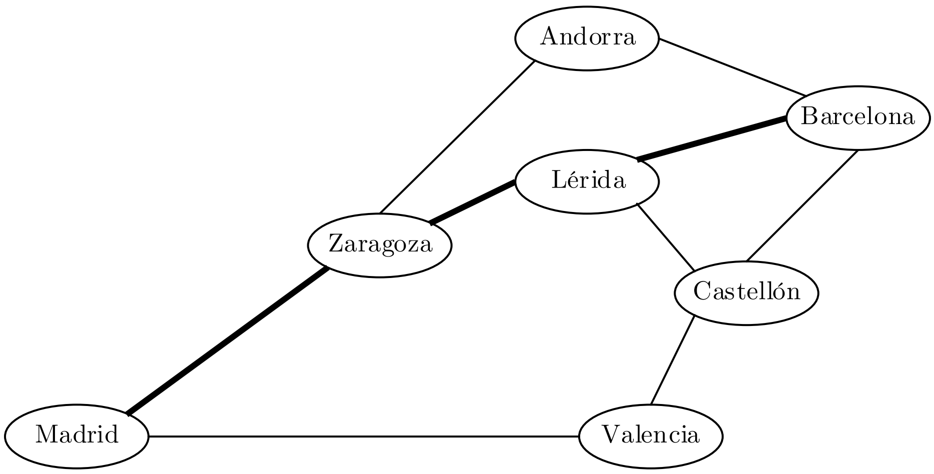
V.2. Principio de optimalidad de la programación dinámica o de Bellman

El principio en el que se basa la programación dinámica fue establecido por Richard Bellman en los años 50. De forma resumida, el principio de optimalidad de Bellman dice que “toda política óptima está formada por subpolíticas óptimas”. A continuación se explica e interpreta con más detalle este principio.

Dado un estado, la política óptima para las siguientes etapas no depende de la política tomada en las etapas anteriores. La decisión óptima inmediata sólo depende del estado en el que se está, no de cómo se llegó hasta él. Toda la información sobre el pasado se resume en el estado en que se encuentra.

Una vez conocida la solución óptima global, cualquier solución parcial que involucre sólo una parte de las etapas es también una solución óptima. Todo subconjunto de una solución óptima es a su vez una solución óptima para un problema parcial.

Aplicando este principio a la búsqueda del camino más corto entre Madrid y Barcelona averiguamos que la solución del problema pasa por Zaragoza. Si nos preguntamos por el camino más corto entre Zaragoza y Barcelona, es obvio que será el mismo que el utilizado en la solución del problema global (Madrid-Barcelona). Si existiera un camino más corto entre Zaragoza y Barcelona (problema parcial), lo habríamos tomado como parte de la solución del problema global.



A continuación y a efectos de ilustrar el método de trabajo de la programación dinámica vamos a analizar algunos modelos de aplicación muy general.

V.2.1. Un ejemplo introductorio

Vamos a considerar un problema clásico soportado por una red valorada (es decir, una red a cada uno de cuyos arcos le ha sido asignado uno o más valores, que pueden ser una distancia, un tiempo o un coste) y sin circuitos.

Supongamos que se quiere construir una carretera entre las ciudades de Aburgo y Teburgo, que constará de seis tramos y pasará por una serie de ciudades. Los costes de las distintas variantes que el trazado de la carretera puede tener vienen reflejados en la figura 1 y en ellos se han tenido en cuenta los trabajos y construcciones de carácter técnico que la obra lleva implícitos, así como los gastos de expropiaciones, indemnizaciones, etc.

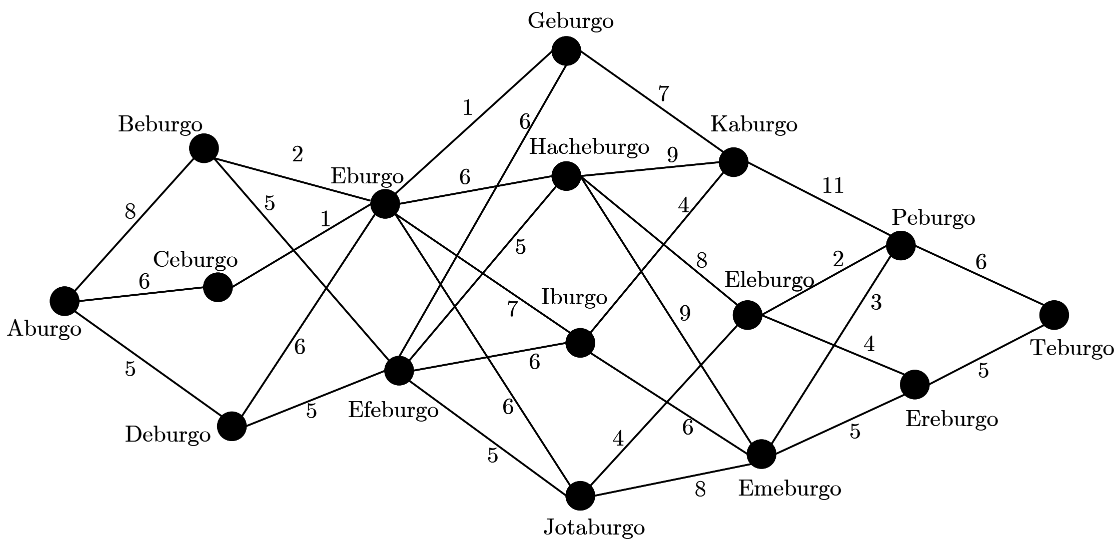


Figura 1

Aunque la carretera puede ser construida en cualquiera de los dos sentidos supondremos que lo será de Aburgo a Teburgo y que el criterio para establecer el trazado de la misma es solamente el objetivo económico de minimizar los costes.

Comenzaremos, de acuerdo con el convenio anterior, por considerar la red orientada de la figura 2. La clasificación en niveles (en este caso, cada nivel será uno de los tramos en la construcción de la carretera) de los vértices de la red es la dada en la Tabla 1. En el nivel 0 se encuentra la ciudad inicial Aburgo y en el nivel *i-ésimo* se encuentran aquellos vértices (ciudades) que sólo tienen antecesores inmediatos en los niveles precedentes. Hay algoritmos, muy sencillos de programar en el ordenador, para proceder a esta clasificación en niveles.

Llamaremos x_i, $i = 0,1,\ldots,6$, a las variables de decisión relativas a cada tramo. La variable x_i puede tomar como valores las ciudades correspondientes al *i-ésimo* tramo.

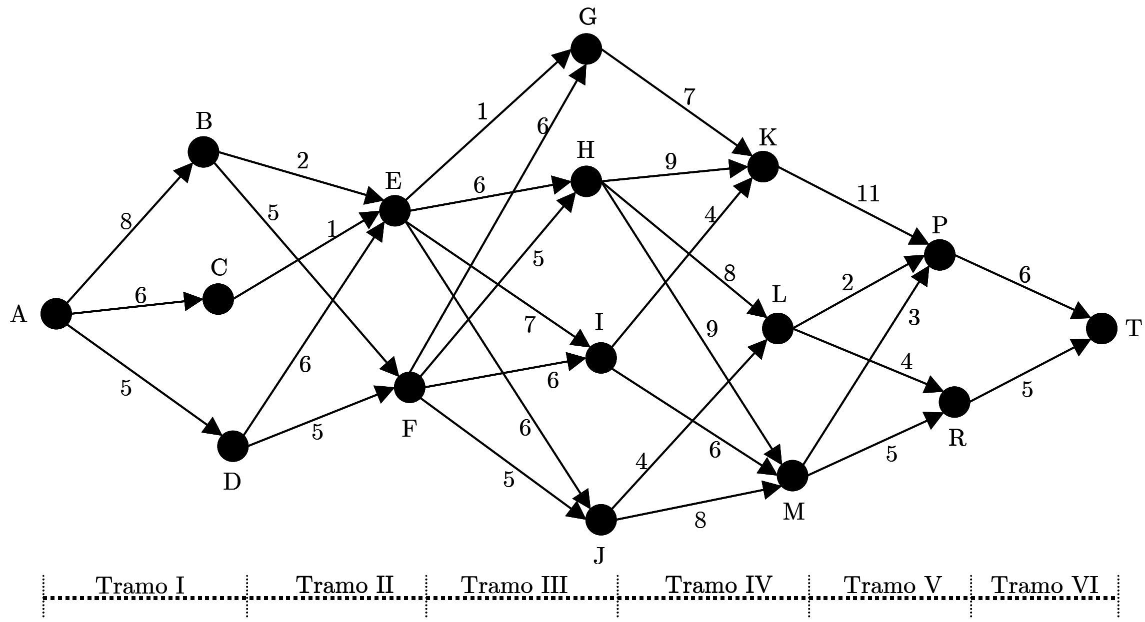


Figura 2

Nivel	Ciudades
0	A
1	B, C, D
2	E, F
3	G, H, I, J
4	K, L, M
5	P, R
6	T

Tabla 1

Designaremos por $v_i\left(x_{i-1}, x_i\right)$ al coste del tramo *i-ésimo*, el cual dependerá de los valores que demos a x_{i-1} y a x_i. La expresión que nos proporciona el coste total es

$$G\left(x_0, x_1, \ldots, x_6\right) = \sum_{i=1}^{6} v_i\left(x_{i-1}, x_i\right) \tag{1.168}$$

1) El coste mínimo del tramo 1 para cada una de las ciudades *B*, *C* y *D* que lo configuran es

$$g_1(B) = 8, \quad g_1(C) = 6, \quad g_1(D) = 5$$

2) Calculamos ahora el coste mínimo conjunto de construcción de tramos 1 y 2. Las ciudades en las que puede acabar el conjunto de los dos primeros tramos son E y F. Tendremos

$$\begin{aligned} g_{1,2}(E) &= \min_{x_1}\left(g_1(x_1)+v_2(x_1,E)\right) \\ &= \min\left(g_1(B)+v_2(B,E), g_1(C)+v_2(C,E), g_1(D)+v_2(D,E)\right) = \\ &= \min(8+2, 6+1, 5+6) = 7 \end{aligned}$$

Es decir, el camino más barato hasta *E* es el que, partiendo de *A*, pasa por *C*.

$$\begin{aligned} g_{1,2}(F) &= \min_{x_1}\left(g_1(x_1)+v_2(x_1,F)\right) \\ &= \min\left(g_1(B)+v_2(B,E), g_1(C)+v_2(C,E), g_1(D)+v_2(D,E)\right) = \\ &= \min(8+5, 6+\infty, 5+5) = 10 \end{aligned}$$

Hemos asignado un coste infinito a aquellos posibles trazados cuya construcción se ha estimado como no realizable y que, en consecuencia, no aparecen en el diseño de la figura 1. El camino más barato hasta *F* es el que, partiendo de *A*, pasa por *D*.

Pasaríamos luego a determinar el coste mínimo para los tramos I, II y III considerados conjuntamente y para los vértices *G*, *H*, *I* y *J*. Este coste es, para *G*,

$$\begin{aligned} g_{1,2,3}(G) &= \min_{x_2}\left[g_{1,2}(x_2)+v_3(x_2,G)\right] = \\ &= \min\left[g_{1,2}(E)+v_3(E,G), g_{1,2}(F)+v_3(F,G)\right] = \\ &= \min[7+1, 10+6] = 8 \end{aligned}$$

y de la misma forma obtendríamos

$$g_{1,2,3}(H) = 13; \quad g_{1,2,3}(I) = 14; \quad g_{1,2,3}(J) = 13$$

que corresponden a los caminos *ACEG*, *ACEH*, *ACEI* y *ACEJ*.

Prosiguiendo en esta línea, obtendríamos para el conjunto de las cuatro primeras fases

$g_{1,2,3,4}(K) = 15$, correspondiente al camino *ACEGK*

$g_{1,2,3,4}(L) = 17$, correspondiente al camino *ACEJL*

$g_{1,2,3,4}(M) = 20$, correspondiente al camino *ACEIM*

y para las cinco primeras fases se tendría

$g_{1,2,3,4,5}(P) = 19$, a lo largo del camino *ACEJPL*

$g_{1,2,3,4,5}(R) = 21$, a lo largo del camino *ACEJLR*

Finalmente, el coste mínimo del proyecto completo sería

$$G^*(T) = \min G(x_0, x_1, x_2, x_3, x_4, x_5, x_6) = \min_{x_5}\left[g_{1,2,3,4,5}(x_5) + v_6(x_5, T)\right] = 25$$

correspondiente a la secuencia *ACEGKPT*.

Este método es característico de los procedimientos de la programación dinámica. Su peculiaridad es la de ir investigando subpolíticas óptimas que abarquen de paso en paso más fases unitivas hasta encontrar la o las políticas óptimas. Es, asimismo, posible, en aplicación del principio de optimalidad de Bellman, empezar por el otro extremo o desdoblar el problema en módulos.

Así, si la construcción de la carretera se abordara simultáneamente desde los dos extremos *A* y *T*, a fin de coincidir en una de las ciudades del tramo III, para cada uno de los valores de x_3 tendríamos, por una parte, según hemos calculado, que los caminos más económicos desde *A* son

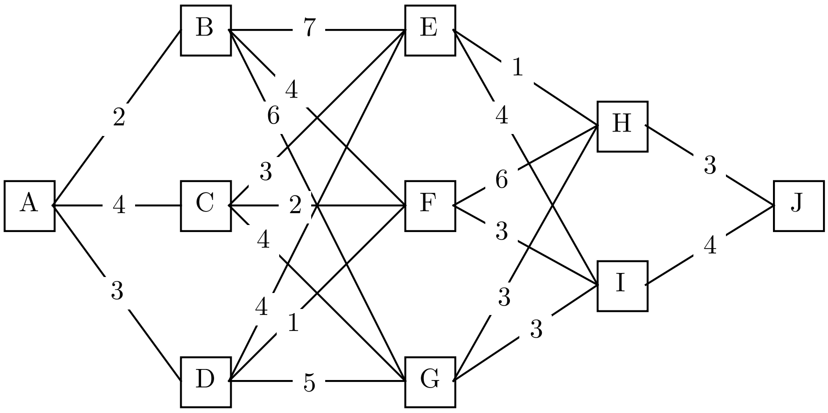
ACEG: coste 8, *ADFH*: coste 13, *ACEI*: coste 14, *ACEJ*: coste 13

mientras que los caminos óptimos desde *T* serían

TPKG: coste 8, *TPLH*: coste 16, *TPMI*: coste 15, *TPLJ*: coste 13

de donde el camino de *A* a *T* de coste total mínimo sería el que resulta de "empalmar" los dos que terminan en *J*.

A continuación vamos a presentar este mismo problema desde otro ángulo. Se trata de un viajante de comercio que desea ir de la ciudad *A* a la *J* por el camino más corto. Vamos a resolver el problema mediante programación dinámica hacia atrás. Notar que el avance a través de la red está dividido en etapas, necesarias para que se pueda utilizar la programación dinámica. Por consiguiente, no es un problema general de distancia mínima en una red (para ese caso también es posible aplicar el principio de optimalidad y ecuaciones para implantar la solución, pero la definición de las etapas y estados no es tan evidente).



Empezamos por la etapa $k = 4$

Estados x_4	Distancia acumulada f_4^*	Decisión óptima u_4^*
H	3	*J*
I	4	*J*

Para la etapa $k = 3$

	Estados x_4			
Estados x_3	*H*	*I*	Distancia acumulada f_3^*	Decisión óptima u_3^*
E	4	8	4	*H*
F	9	7	7	*I*
G	6	7	6	*H*

Para la etapa $k = 2$

	Estados x_3				
Estados x_2	*E*	*F*	G	Distancia acumulada f_2^*	Decisión óptima u_2^*
B	11	11	12	11	*E*, *F*
C	7	9	10	7	*E*
D	8	8	11	8	*E*, *F*

Finalmente en la etapa $k = 1$

	Estados x_2				
Estado x_1	*B*	*C*	*D*	Distancia acumulada f_1^*	Decisión óptima u_1^*
A	13	11	11	11	*C*, *D*

A partir de la tabla final se deduce que las rutas óptimas son: *ACEHJ*, *ADEHJ* y *ADFIJ* con distancia total de 11. Obsérvese que si en cada ciudad se hubiera tomado la decisión miope, es decir, ir a la ciudad más próxima, la distancia total recorrida sería de 13, que no coincide con la óptima.

Veamos ahora la resolución mediante programación dinámica hacia adelante. Para la etapa $k=2$

	Estado x_1		
Estados x_2	*A*	Distancia acumulada f_2^*	Decisión óptima u_2^*
B	2	2	*A*
C	4	4	*A*
D	3	3	*A*

Para $k=3$

	Estados x_2				
Estados x_3	*B*	*C*	*D*	Distancia acumulada f_3^*	Decisión óptima u_3^*
E	9	7	7	7	*C*, *D*
F	6	6	4	4	*D*
G	8	8	8	8	*B*, *C*, *D*

Para $k=4$

	Estados x_3				
Estados x_4	*E*	*F*	*G*	Distancia acumulada f_4^*	Decisión óptima u_4^*
H	8	10	11	8	*E*
I	11	7	11	7	*F*

Finalmente para la etapa $k=5$

	Estados x_4			
Estados x_5	*H*	*I*	Distancia acumulada f_5^*	Decisión óptima u_5^*
J	11	11	11	*H*, *I*

Las rutas óptimas son: *JHECA*, *JHEDA* y *JIFDA* con distancia óptima de 11.

V.2.2. Otro ejemplo: Planificación de la expansión de la generación

Se trata de minimizar los costes totales (fijos y variables) de expansión del equipo generador para un alcance de varios años. Las *decisiones* son la potencia a instalar de cada tipo de generación en cada año del alcance del modelo. Habitualmente se tienen en cuenta estas *restricciones en la expansión*: potencia instalada inicial conocida, máxima (mínima) potencia instalable, inversión máxima (mínima), número máximo (mínimo) de generadores instalables en cada año. Además también se consideran estas *restricciones de operación*: el balance generación demanda en cada año. Los *estados* se definen como el número total de generadores instalados al comienzo de cada año.

Supongamos un sistema eléctrico con los siguientes datos

Año	Demanda (MW)	Coste de inversión por generador de 1 GW [€/GW año]
1999	1000	50
2000	2000	55
2001	4000	60
2002	6000	65
2003	7000	45
2004	8000	40

Existe un coste adicional de 15 €/año por año si se construye al menos un generador. No se pueden instalar más de 3000 MW de generación en ningún año y se parte de un sistema eléctrico sin ningún generador instalado.

Resolvemos el problema mediante programación dinámica hacia atrás. En la etapa $n = 2004$ se tiene

	Est 2004		
Est 2003	8000	Cst fut	Inst ópt
7000	15+40=55	**55**	1000
8000	0	**0**	0

Para la etapa $n = 2003$

	Est 2003			
Est 2002	7000	8000	Cst fut	Inst ópt
6000	15+45+**55**=115	15+90=105	**105**	2000
7000	**55**	15+45=60	**55**	0
8000		0	**0**	0

Para la etapa $n = 2002$

	Est 2002				
Est 2001	6000	7000	8000	Cst fut	Inst ópt
4000	15+130+**105**=250	15+195+**55**=265		**250**	2000
5000	15+65+**105**=185	15+130+**55**=200	15+195=210	**185**	1000
6000	**105**	15+65+**55**=135	15+130=145	**105**	0
7000		**55**	15+65=80	**55**	0
8000			0	**0**	0

Para la etapa $n = 2001$

	Est 2001						
Est 2000	4000	5000	6000	7000	8000	Cst fut	Inst ópt
2000	15+120+**250**=385	15+180+**185**=380				**380**	3000
3000	15+60+**250**=325	15+120+**185**=320	15+180+**105**=300			**300**	3000
4000	**250**	15+60+**185**=260	15+120+**105**=240	15+180+**55**=250		**240**	2000
5000		**185**	15+60+**105**=180	15+120+**55**=190	15+180=195	**180**	1000
6000			**105**	15+60+**55**=130	15+120=135	**105**	0

Para la etapa $n = 2000$

	Est 2000						
Est 1999	2000	3000	4000	5000	6000	Cst fut	Inst ópt
1000	15+55+**380**=445	15+110+**300**=425	15+165+**240**=420			**420**	3000
2000		15+55+**300**=365	15+110+**240**=365	15+165+**180**=360		**360**	3000
3000			15+55+**240**=305	15+110+**180**=305	15+165+**105**=285	**285**	3000

Para la etapa $n = 1999$

	Est 1999				
Est 1998	1000	2000	3000	Cst fut	Inst ópt
0	15+50+**420**=485	15+100+**360**=475	15+150+**285**=450	**450**	3000

La potencia óptima a instalar en cada año a partir de 1999 es: 3000, 3000, 0, 0, 2000 y 0 MW con un coste total de 450 €.

Otras aplicaciones muy características de la programación dinámica en sistemas de energía eléctrica son la programación semanal de los generadores térmicos y la coordinación hidrotérmica. En el primer caso, las etapas son las horas (o días) de la semana, los estados corresponden a las combinaciones posibles de acoplamiento o parada de los grupos térmicos (por ejemplo, 2^n siendo n el número de grupos térmicos cada uno con dos posibles estados, acoplamiento o parada) y las decisiones son los arranques y paradas de los grupos entre horas consecutivas. En el segundo caso, las etapas son las semanas (o meses) del año, los estados son las combinaciones de los niveles de reserva de los embalses (por ejemplo, 5^n siendo n el número de embalses cada uno con cinco posibles niveles) y las decisiones son las producciones de las centrales hidráulicas. En ambos casos, para sistemas eléctricos reales se puede apreciar el elevado número de estados que hay que evaluar en cada etapa. En el problema de coordinación hidrotérmica, además es necesaria la discretización artificial de los niveles de los embalses.

V.2.3. Modelos de asignación

Los ejemplos utilizados muestran, también, una característica fundamental de los modelos de programación dinámica, que es el carácter recurrente del proceso: el óptimo en la fase n no tiene en cuenta todas las fases previas, sino sólo las k precedentes siendo, por lo general $k = 1$, de forma que en este último caso el valor óptimo para las diferentes valores de las variables de estado $\overline{x}_n$ de la fase n se obtiene como una función de los óptimos parciales, previamente calculados, de los valores de las variables $\overline{x}_{n-1}$ y de la relación que ligue a $\overline{x}_n$ y $\overline{x}_{n-1}$.

Es precisamente en la determinación de esta recurrencia donde radica la dificultad de elaborar un modelo de programación dinámica pero, una vez conseguido el modelo, su aplicación supone una reducción drástica de las opciones a considerar a efectos de la determinación del óptimo.

Los ejemplos anteriores son también ilustración de un caso general, que podemos denominar de *funciones separables en fases,* y el proceso de optimización empleado se dice que es de *tipo secuencial.* Estas denominaciones son válidas no sólo para aquellas funciones para las que

$$G(x_0,x_1,\ldots,x_n) = v_1(x_0,x_1) + v_2(x_1,x_2) + \ldots + v_n(x_{n-1},x_n) \tag{1.169}$$

sino también para aquellos casos en los que

$$G(x_0,x_1,\ldots,x_n) = v_1(x_0,x_1)v_2(x_1,x_2)\ldots v_n(x_{n-1},x_n) \tag{1.170}$$

Vamos a continuación a analizar algunos modelos dinámicos que, por su generalidad y frecuente aplicación, merecen mención aparte. Son los denominados modelos de asignación. Bajo esta denominación se engloban una serie de problemas cuyo objetivo es determinar el máximo de

$$F(x_1,x_2,\ldots,x_n) = f_1(x_1) + f_2(x_2) + \ldots + f_n(x_n) \tag{1.171}$$

sujetas las variables a las restricciones

$$x_1 + x_2 + \ldots + x_n = K\,;\; x_i \geq 0 \tag{1.172}$$

Una interpretación económica de este problema consiste en suponer que tenemos una cierta cantidad de un recurso *económico* que puede ser de muy diversa naturaleza: hombres, agua, energía, dinero, máquinas, etc. , y que puede ser utilizado de diferentes formas. Cada una de estas diferentes maneras de utilizar el recurso recibe el nombre de ***actividad*** y como resultado de utilizar todo o parte del recurso en una determinada actividad obtenemos un ***ingreso*** o ***retorno*** que, a su vez, puede venir expresado de muy diferentes maneras. La magnitud del ingreso depende de la cantidad de recurso asignado y de la actividad sobre la que se hace la asignación. A estos efectos supondremos que:

- Los ingresos procedentes de distintas actividades pueden ser medidos con una misma unidad.
- El ingreso o utilidad total es la suma de los ingresos o utilidades de cada actividad.
- El ingreso procedente de una determinada actividad es independiente de las asignaciones hechas en las otras actividades.

Para ilustrar el problema de la asignación y las condiciones matemáticas que deben cumplirse para garantizar la existencia, y eventualmente, la unicidad de las soluciones vamos a analizar la situación siguiente, en la que suponemos que disponemos de una cantidad C_0 que tenemos que dividir en dos partes x_0 y $C_0 - x_0$, obteniendo de la primera cantidad x_0 un retorno $g(x_0)$ y de la segunda un retorno $h(C_0 - x_0)$. Si deseamos hacer el reparto de forma que maximicemos el retorno total estamos enfrentados al problema analítico de obtener el máximo de la función

$$R_1(C_0, x_0) = g(x_0) + h(C_0 - x_0), \quad x_0 \in [0, C_0] \tag{1.173}$$

Supondremos que g y h son funciones continuas $\forall x \in [0, C_0]$, de forma que su máximo siempre existe en ese intervalo.

Consideremos ahora un proceso bietápico en el que, como precio de obtener un retorno $g(x_0)$, la cantidad x_0 se ve reducida a cx_0, $0 < c < 1$, y, análogamente, $C_0 - x_0$ se ve reducida a $b(C_0 - x_0)$, $0 < b < 1$. Con la cantidad restante total,

$$C_1 = cx_0 + b(C_0 - x_0)$$

volvemos a repetir el proceso, haciendo

$$C_1 = x_1 + (C_1 - x_1), \quad 0 < x_1 < C_1$$

y, como resultado de la nueva asignación, obtenemos el retorno parcial

$$g(x_1) + h(C_1 - x_1)$$

y un retorno global para el conjunto del proceso bietápico

$$\begin{gathered} R_2(C_0, x_0, x_1) = g(x_0) + h(C_0 - x_0) + g(x_1) + h(C_1 - x_1) \\ 0 \le x_0 \le C_0; \quad 0 \le x_1 \le C_1 \end{gathered}$$

Si damos un salto a un proceso en N etapas, en cada una de las cuales repetimos el proceso de asignación anterior, el retorno total vendrá dado por

$$\begin{gathered} R_N(C_0, x_0, x_1, \ldots, x_{N-1}) = g(x_0) + h(C_0 - x_0) + g(x_1) + h(C_1 - x_1) + \ldots \\ \ldots + g(x_{N-1}) + h(C_{N-1} - x_{N-1}) \end{gathered}$$

$$\begin{aligned}
&x_1 = cx_0 + b(C_0 - x_0), \quad 0 \le x_0 \le C_0 \\
&x_2 = cx_1 + b(C_1 - x_1), \quad 0 \le x_2 \le C_1 \\
&\dots \\
&x_{N-1} = cx_{N-2} + b(C_{N-2} - x_{N-2}), \quad 0 \le x_{N-2} \le C_{N-2}, \quad 0 \le x_{N-1} \le C_{N-1}
\end{aligned}$$

y, en principio, parece que estamos enfrentados a un problema de maximizar una función en un dominio N-dimensional. Pero observamos que, en cada etapa lo que hacemos es una elección unidimensional, por lo que si conseguimos formular el problema de manera que se preserva esta peculiaridad unidimensional, nos habremos librado del engorroso tratamiento del análisis multidimensional.

Con este fin definimos la función

$$f_N(C_0) = \text{máximo retorno obtenido en un proceso de } N \text{ etapas}$$

$$(N = 1,2,\dots) \text{ comenzando con una cantidad inicial}$$

$$C_0 = \max_{\{x_0,x_1,\dots,x_{N-1}\}} R_N(C_0, x_0, \dots, x_{N-1}), \quad N = 2,3,\dots$$

con

$$f_1(C_0) = \max_{0 \le x_0 \le C_0} R_1(C_0, x_0) = \max_{0 \le x_0 \le C_0} \left[g(x_0) + h(C_0 - x_0) \right]$$

Nuestro siguiente objetivo será obtener una ecuación para $f_2(C_0)$ en términos de $f_1(C_0)$. Considerando el proceso bietápico, vemos que el retorno total será la suma del retorno de la primera etapa más el de la segunda, cuando en esta última se ha asignado una cantidad $C_1 = cx_0 + b(C_0 - x_0)$. Es claro que cualquiera que sea el valor de x_0 elegido en la etapa inicial, la cantidad restante C_1 debe ser usada de la mejor manera posible en la segunda etapa, si deseamos obtener una asignación óptima. Si lo hacemos así, en la segunda etapa conseguiremos un retorno óptimo

$$f_1\left(cx_0 + b(C_0 - x_0)\right)$$

si x_1 es elegido de forma óptima. Como resultado el retorno total para el conjunto de las dos primeras etapas, resultante de la asignación inicial x_0, viene dado por

$$R_2(C_0, x_0, x_1) = g(x_0) + h(C_0 - x_0) + f_1\left(ax_0 + b(C_0 - x_0)\right)$$

Como x_0 es elegido de manera que haga máxima la expresión anterior, deducimos la relación de recurrencia

$$f_2(C_0) = \max_{0 \le x_0 \le C_0} \left\{ g(x_0) + h(C_0 - x_0) + f_1\left(ax_0 + b(C_0 - x_0)\right) \right\}$$

conectando a las funciones f_1 y f_2. Usando el mismo argumento para el proceso en *N* etapas, obtenemos la ecuación funcional

$$f_N(C_0) = \max_{0 \le x_0 \le C_0} \left\{ g(x_0) + h(C_0 - x_0) + f_{N-1}\left(ax_0 + b(C_0 - x_0)\right) \right\}, \;\; N \ge 2$$

de forma que, comenzando con f_1, continuamos evaluando sucesivamente $f_2, f_3, \ldots$ En cada etapa del proceso, obtenemos, no sólo $f_k(C_0)$, sino también $x_k(C_0)$, la asignación óptima que debemos hacer al comienzo de la *k-ésima* etapa cuando inicialmente se parte con una cantidad C_0. Así, la solución de un problema específico, dados C_0 y N, es de la forma

$$\begin{aligned}
&x_0^* = x_N(C_0) \\
&x_1^* = x_{N-1}\left(ax_0^* + b\left(C_0 - x_0^*\right)\right) \\
&x_2^* = x_{N-2}\left(ax_1^* + b\left(C_1 - x_1^*\right)\right) \\
&\ldots \\
&x_{N-1}^* = x_1\left(ax_{N-2}^* + b\left(C_{N-2} - x_{N-2}^*\right)\right)
\end{aligned}$$

siendo $\left(x_0^*, x_1^*, \ldots, x_{N-1}^*\right)$ un conjunto de asignaciones, que puede no ser único, que maximiza el retorno total tras N etapas.

Las ventajas del método son evidentes: hemos reducido un simple problema N-dimensional a una secuencia de N problemas unidimensionales.

¿Qué ocurre si el proceso continua indefinidamente, es decir, si $N \to \infty$? Aunque un proceso no acotado es siempre una ficción física, como proceso matemático es atractivo, ya que el paso al límite de la ecuación [7] nos permite reducir el problema al de resolver la ecuación funcional

$$f(C_0) = \max_{0 \le x \le C_0} \left[g(x) + h(C_0 - x) + f\left(ax + b(C_0 - x)\right) \right]$$

cuya solución $f(C_0)$ es el retorno total del proceso, con una función de asignación $x(C_0)$ proporcionada por la ecuación anterior.

Este ejemplo es un caso particular del modelo general consistente en obtener el máximo de la función

$$F(x_1, x_2, \ldots, x_N) = \sum_{i=1}^{N} f_i(x_i) \text{ (todas las } f_i \text{ continuas } \forall x \ge 0\text{)}$$

sobre la región definida por

$$\begin{aligned} &x_1 + x_2 + \ldots + x_N = C \\ &x_i \ge 0 \end{aligned}$$

Como el máximo depende de *N* y *c*, definimos la secuencia de funciones

$$g_N(C) = \max_{\{x_i\}} F(x_1, x_2, \ldots, x_N), \quad N = 1, 2, \ldots$$

y, argumentando como ejemplos anteriores, tenemos la relación de recurrencia

$$f_N(C) = \max_{0 \le x \le C} [\ ]$$

$$g_N(C) = \max_{0 \le x \le C} \left(f_N(x) + g_{N-1}(C - x) \right), \ N = 2, 3, \ldots$$

con $g_1(C) = f_1(C)$.

La ecuación [8] nos plantea la necesidad de garantizar la *existencia y unicidad de sus soluciones*. Los siguientes teoremas, cuya importancia no es meramente teórica, sino también computacional, establecen las condiciones requeridas.

Teorema 1:

Supongamos que

1) $g(x)$ $h(x)$ son funciones continuas de x para $x \ge 0$, con $g(0) = h(0) = 0$.

2) $0 \le a < 1, \quad 0 \le b < 1$ $\sum_{n=0}^{\infty} M(c^n x) < \infty, \ \forall x \ge 0, \ \text{siendo}$

$$M(x) = \max_{0 \le y \le x} \max\left(\left|g(y)\right|, \left|h(y)\right|\right) \text{ y } c = \max(a,b)$$

Bajo estas hipótesis hay una solución única de [8] que es continua y que se anula para $x = 0$.

Definición 1:

Una función $f(x)$ continua es **convexa** en el intervalo $[a,b]$ si $\forall u,v \in [a,b]$ y $\forall \lambda \in [0,1]$

$$f\left(\lambda u + (1-\lambda)v\right) \le \lambda f(u) + (1-\lambda) f(v)$$

Se dice que $f(x)$ es **cóncava** en el intervalo $[a,b]$ si $\forall u,v \in [a,b]$ y $\forall \lambda \in [0,1]$

$$f\left(\lambda u + (1-\lambda)v\right) \ge \lambda f(u) + (1-\lambda) f(v)$$

Una función convexa (cóncava) en $[a,b]$ tiene un valor mínimo (respectivamente, máximo) en dicho intervalo.

Teorema 2:

Si, además de las condiciones del teorema 1, g y h son funciones convexas de C_0 entonces $f(C_0)$ es también una función convexa de C_0 y, para cada valor de C_0, la elección *x* será igual a 0 o a C_0.

Teorema 3:

Si, además de las condiciones del teorema 1, g y h son funciones estrictamente cóncavas de C_0, entonces $f(C_0)$ es también una función estrictamente cóncava de C_0 y la política óptima es única.

Teorema 4:

Supongamos que $g(x)$ y $h(x)$ son estrictamente cóncavas para $x \ge 0$, monótonas crecientes con derivadas continuas y que $g(0) = h(0) = 0$.

$$\frac{g'(0)}{1-a} > \frac{h'(0)}{1-b}, \quad h'(0) > g'(0), \quad a < b$$

Entonces la política óptima tiene la siguiente forma:

a) $x = C_0$ para $0 \le C_0 \le C^*$, donde C^* es la raíz de

$$h'(0) = g'(u) + (b-a)\,g'(au) + a(b-a)\,g'\left(a^2u\right) + \ldots$$

b) $x = x(C_0)$ para $C_0 > C^*$, donde $x(C_0)$ es una función que satisface las desigualdades $0 < x(C_0) < C_0$ y es la solución de

$$g'(x) - h'(C_0 - x) + (a-b)\,f'\left(ax + b(C_0 - x)\right) = 0$$

Veamos un ejemplo de problema de asignación. Supongamos, a título de ejemplo, que una empresa quiere vender sus productos en cuatro zonas que denominaremos I, II, III y IV y dispone de una cantidad de 9 millones de euros para invertir en publicidad. Supuesto que se conocen en cada zona las ganancias correspondientes a cada una de las posibles inversiones se desea determinar cual debe ser la distribución óptima de los 9 millones. Suponiendo, para simplificar, que las asignaciones se hacen por millones enteros de euros, podemos considerar que los beneficios, también en millones, para cada posible asignación vienen reflejados en la tabla 2.1

Inversión	0	1	2	3	4	5	6	7	8	9
Zona 1	0	0.5	0.6	0.8	1.1	1.3	1.4	1.46	1.50	1.52
Zona II	0	0.4	0.48	0.6	0.78	0.91	1	1.05	1.08	1.1
Zona III	0	0.6	0.71	0.9	1.08	1.26	1	1.48	1.55	1.6
Zona IV	0	0.3	0.45	0.7	0.85	0.95	1.02	I 1.07	1.1 1	1.13

Tabla 2.1

Llamando $f_i(x_i)$ a la ganancia obtenida en la zona i con una inversión x_i en publicidad, se trataría de hacer máximo

$$F(x_1, x_2, x_3, x_4) = f_1(x_1) + f_2(x_2) + f_3(x_3) + f_4(x_4)$$

con las restricciones

$$x_1 + x_2 + x_3 + x_4 = 9 \ ; \ x_i \ge 0, \ x_i \in \mathbb{Z}$$

En estos modelos de asignación es muy común sustituir la variable x_i por otra variable y_i que representa la cantidad total invertida en el conjunto de las i primeras zonas. Es claro que $x_i = y_i - y_{i-1}$ y la función [9] puede ser escrita como

$$F(y_1, y_2, y_3, y_4) = f_1(y_1) + f_2(y_2 - y_1) + f_3(y_3 - y_2) + f_4(9 - y_3)$$

Para construir un algoritmo recurrente, que es una característica de la programación dinámica, comenzamos definiendo

$$g_0(y_0) = 0 \ , \ y_0 = 0$$

Si el algoritmo fuera multiplicativo esta fase inicial sería definida por

$$g_0(y_0) = 1 \ , \ y_0 = 1$$

A partir de aquí definimos para la zona i la función de optimización

$$g_i(y_i) = \max_{0 \le y_{i-1} \le y_i} \left[g_{i-1}(y_{i-1}) + f_i(y_i - y_{i-1}) \right]$$

y para $i = 4$ obtendremos el óptimo deseado con $y_4 = 9$

$$g_4(9) = \max_{0 \le y_3 \le 9} \left[g_3(y_3) + f_4(9 - y_3) \right]$$

Comenzaremos el desarrollo del algoritmo con ayuda de la tabla 2.2

y_1, y_2	0	1	2	3	4	5	6	7	8	9
	0	0.5	0.6	0.8	1.1	1.3	1.4	1.46	1.50	1.52
$f_2(y_2 - y_1) = f_2(x_2)$	0	0.4	0.48	0.6	0.78	0.91	1	1.05	1.08	1.1
$g_2(y_2)$	0	0.5	0.9	1	1.2	1.5	1.7	1.8	1.9	2.08
Subpolíticas óptimas para las zonas I y II	0,0	1,0	1,1	2,1	3,1	4,1	5,1	6,1	5,3	5,4

Tabla 2.2

para obtener las subpolíticas óptimas para el conjunto de las dos primeras zonas, en la que, por ejemplo, $g_2(3)$ se ha obtenido de la forma siguiente

$$g_2(3) = \max\left[f_1(0) + f_2(3), f_1(1) + f_2(2), f_1(2) + f_2(1), f_1(3) + f_2(0)\right] =$$
$$= \max[0 + 0.6,\ 0.5 + 0.48,\ 0.6 + 0.4,\ 0.8 + 0] = 1$$

y los demás en forma similar.

Para determinar $g_3(y_3)$ obtenemos la tabla 2.3 con ayuda de la tabla anterior y de la fórmula [12]. Esta tabla y la fórmula [13] nos permitirán obtener la tabla 2.4 y determinar la política óptima.

x_3, y_2	0	1	2	3	4	5	6	7	8	9
$g_2(y_2)$	0	0.5	0.9	1	1.2	1.5	1.7	1.8	1.9	2.08
$f_3(x_3) = f_3(y_3 - y_2)$	0	0.6	0.71	0.9	1.08	1.26	1.4	1.48	1.55	1.6
$g_3(y_3)$	0	0.06	1.1	1.5	1.61	1.8	2.1	2.8	2.41	2.6
Subpolítica óptima para zonas I y II	0,0	1,0	1,1	2,1	3,1	4,1	5,1	6,1	5,3	5,4
Subpolíticas óptimas para zonas I a III	0,0,0	0,0,1	1,0,1	1,1,1	1,1,2	3,1,1 1,1,3	4,1,1	5,1,1	5,1,2	5,1,3

Tabla 2.3

x_4, y_3	0	1	2	3	4	5	6	7	8	9
$g_3(y_3)$	0	0.6	1.1	1.5	1.61	1.8	2.1	2.3	2.41	2.6
$f_4(x_4) = f_4(9 - y_3)$	0	0.3	0.45	0.7	0.85	0.95	1.02	1.07	1.1 1	1.13
$g_4(9)$	0	0.6	1.1	1.5	1.8	1.95	2.2	2.4	2.6	2.8
Subpolíticas Óptimas para Zonas I, II y III	0,0,0	0,0,1	1,0,1	1,1,1	1,2,2	3,1,1 1,1,3	4,1,1	5,1,1	5,1,2	5,1,3
Políticas óptimas	0,0,0,0	0,0,1,0	1,0,1,0	1,1,1,0	1,1,1,1	1,1,1,2	1,1,1,3	4,1,1,1	5,1,1,1	4,1,1,3

Tabla 2.4

La política óptima es, por tanto, la (4,1,1,3), pudiendo verificarse que todas las subpolíticas son óptimas. Así, por ejemplo, la tabla 2.2 nos muestra que la mejor manera de distribuir 5 millones entre las zonas I y II es 4 y 1 respectivamente, mientras que el análisis combinado de las tablas 2.2 y 2.4 nos permitiría establecer que la manera óptima de distribuir 4 millones entre las zonas III y IV es asignar 1 millón a III y 3 a IV con un beneficio de 1.3 millones.

Debe hacerse resaltar la gran eficiencia del método. Una descripción de todas las políticas hubiera precisado el análisis de 220 posibilidades, mientras que si el número de millones a invertir hubiera sufrido un incremento sólo de 3 millones, el número de posibilidades hubiera llegado a 455.

En general, el número de posibilidades de repartir n unidades monetarias entre k zonas viene dado por las combinaciones de $(n+k-1)$ elementos seleccionados de $k-1$ en $k-1$.

V.2.4. Caminos óptimos sobre redes valoradas sin circuitos

Al comienzo vimos un ejemplo en el que sobre una red valorada, en la que cada arco tenía asociado el coste de la construcción del correspondiente tramo de carretera, tratábamos de obtener el trazado de una carretera entre dos ciudades con coste mínimo.

En otras ocasiones el objetivo, por el contrario, será determinar el camino de máximo valor entre la entrada y la salida de una red valorada. Una tal situación se da en los modelos de *gestión de proyectos* en los que se utiliza la técnica PERT-CPM.

En el caso siguiente se analiza una red del tipo conocido como *red Manhattan,* que adopta una estructura rectangular. En el ejemplo considerado se trata de encontrar el camino de duración mínima entre los puntos O y *B* de una ciudad, separados por siete manzanas de casas, conociéndose el tiempo, en minutos, que se tarda en atravesar cada una de las manzanas que constituyen las diversas rutas que unen *O* y *B.* Se trata de obtener el camino que lleva de O hasta *B* con la condición de que las direcciones de avance son de izquierda a derecha o de abajo hacia arriba. A tal efecto, consideraremos un origen de coordenadas en el punto O y asociaremos a cada encrucijada posible un par de coordenadas y en concreto, al punto *B* le asociaremos el par (3,4), tal como se muestra en la figura 3, que es una red Manhattan 3 x 4.

En lo que sigue designaremos por $t(i,j)$ al mínimo tiempo que se tarda en alcanzar el vértice (i,j) desde O. Además llamamos $i=0$ al tiempo que se tarda en llegar al vértice (i,j) desde el vértice "inferior" $(i,j-1)$ *y* $D(i,j)$ al tiempo que se invierte en el trayecto entre los vértices $(i-1,j)$ e (i,j). El proceso de búsqueda del camino al que corresponde el tiempo mínimo de viaje entre el punto O $(0,0)$ y el punto *B* (m,n) lo dividiremos en $m+n$ etapas o fases, para una red Manhattan general de dimensión $m \times n$. La variable de estado x_k asociada a la fase k tomará como argumentos todos los vértices (i,j) tales que $i+j=k$.

El algoritmo recurrente que nos conduce al óptimo se define para la etapa k, en la que se supone que ya ha sido aplicado en las $k-1$ etapas anteriores, como

$$\underset{i+j=k}{t(i,j)} = \begin{cases} \min\left[t(i-1,j)+D(i,j), t(i,j-1)+A(i,j)\right], \ j \neq 0, i \neq 0 \\ t(i-1,j)+D(i,j), \text{ si } j=0 \\ t(i,j-1)+A(i,j), \text{ si } i=0 \end{cases}$$

con la condición inicial evidente $t(0,0)=0$

Para cada fase, y para cada vértice de la misma tal que $i=0$, $j=0$, la decisión a tomar es si debe ser alcanzado desde el vértice $(i-1,j)$, que representaremos por una Derecha, o desde el vértice $(i,j-1)$, que representaremos por una Arriba. Para los vértices en los que $i=0$ o $j=0$ la decisión viene forzada: A o *D*, respectivamente.

(0,4)	5	(1,4)	4	(2,4)	3	(3,4)
2		3		4		6
(0,3)	2	(1,3)	3	(2,3)	2	(3,3)
3		5		4		1
(0,2)	1	(1,2)	3	(2,2)	6	(3,2)
4		3		6		5
(0,1)	2	(1,1)	4	(2,1)	4	(3,1)
3		1		3		5
(0,0)	2	(1,0)	3	(2,0)	4	(3,0)

Figura 3

Nótese que el algoritmo debe ser evaluado en los $(m+1)\times(n+1)-1$ vértices de la red (exceptuando el vértice 0). En cada uno de los $m\times n$ vértices para los que $i\neq 0$, $j\neq 0$ se realizan tres operaciones: dos sumas y una selección de mínimo; en cada uno de los $m+n$ vértices para los que $i=0$ o $j=0$ se realiza una sola operación. En total, y para alcanzar el óptimo, se precisan $4(m+n)$ operaciones; en nuestro caso 28.

Considerando la red de la figura *3* la aplicación del algoritmo a las sucesivas fases nos conduce a los siguientes resultados, entre los cuales los correspondientes a la fase quinta aparecen detallados.

Fase 1	Vértice	$t(i,j)$	Decisión
	(0,1)	3	A
	(1,0)	2	D

Fase 2	Vértice	$t(i,j)$	Decisión
	(0,2)	7	A
	(1,1)	3	A
	(2,0)	5	D

Fase 3	Vértice	$t(i,j)$	Decisión
	(0,3)	10	A
	(1,2)	6	A
	(2,1)	7	D
	(3,0)	9	D

Fase 4	Vértice	$t(i,j)$	Decisión
	(0,4)	12	A
	(1,3)	11	A
	(2,2)	9	D
	(3,1)	11	D

Fase 5	Vértice	$t(i,j)$	Decisión
	(1,4)	$\min[t(0,4)+D(1,4),t(1,3)+A(1,4)]=$ $=\min(12+5,11+3)=14$	A
	(2,3)	$\min[t(1,3)+D(2,3),t(2,2)+A(2,3)]=$ $=\min(11+3,9+4)=13$	A
	(3,2)	$\min[t(2,2)+D(3,2),t(3,1)+A(3,2)]=$ $=\min(9+6,11+5)=15$	D

Fase 6	Vértice	$t(i,j)$	Decisión
	(2,4)	17	A
	(3,3)	15	D

Fase 7	Vértice	$t(i,j)$	Decisión
	B(3,4)	20	D

La política óptima se obtiene ahora analizando las decisiones óptimas desde el final al principio. La decisión óptima final era alcanzar el vértice B desde la derecha; ello nos sitúa en el vértice (2,4). Para este vértice la decisión óptima era A, esto es, ser alcanzado desde el vértice (2,3). Para el vértice (2,3) la decisión óptima era A, ser alcanzado desde el vértice (2,2). Este último vértice era alcanzado de forma óptima desde (1,2), según la decisión D, y el vértice (1,2) lo era desde (1,1), por la decisión óptima A. A su vez (1,1) debe ser alcanzado desde (1,0) y este último desde el origen O. Así pues, la secuencia óptima de vértices es

$$0(0,0) \rightarrow (1,0) \rightarrow (1,1) \rightarrow (1,2) \rightarrow (2,2) \rightarrow (2,3) \rightarrow (2,4) \rightarrow B(3,4)$$

En este punto son interesantes algunas consideraciones. Observemos que si hubiéramos analizado todos los caminos que conducen de O a B hubiéramos tenido que evaluar

V.2.5. Procesos polietápicos aleatorios de decisión

Vamos a aclarar lo anterior estudiando una situación en la que un buscador de oro tiene dos pequeños minas, Áurea y Benéfica, cuyas reservas se estiman respectivamente en las cantidades x e y. El afortunado tiene un pequeño extractor mecánico para sacar el oro de las minas. Cuando el buscador lo emplea en Áurea tiene una probabilidad p_A de que extraiga una fracción r_A del oro que haya en ese momento en la mina y siga funcionando y una probabilidad $(1-p_A)$ de que se averíe y no extraiga nada. La avería, de producirse, es catastrófica, sin posibilidad de reparación. En forma similar, Benéfica tiene asociadas las probabilidades p_B y $(1-p_B)$ y la fracción r_B.

El proceso de extracción comienza usando el extractor en una de las dos minas. Si la máquina no se ha averiado tras esta operación inicial, se vuelve a elegir en cual de las dos minas se va a realizar la segunda extracción, y así sucesivamente hasta que la máquina se avería o hasta un determinado número de extracciones.

En los casos de procesos aleatorios, a la secuencia de decisiones se le suele dar el nombre de *estrategia*, en vez del de política. En este caso, una estrategia será representada por una sucesión de A's y B's, según se elija Áurea o Benéfica. Así, la sucesión BAABABBBA… representa una estrategia que comienza eligiendo Benéfica para, después, elegir dos veces seguidas Áurea, suponiendo siempre que la máquina no se avería, y así sucesivamente.

Dada la intervención del azar, el valor de una estrategia se suele medir utilizando algún tipo de promedio sobre los valores aleatorios que puede tomar la magnitud que se quiere maximizar, en este caso la cantidad total de oro extraída. Suele ser habitual tomar como medida de valoración el *valor esperado*.

Volveremos a emplear la aproximación mediante ecuaciones funcionales. Definimos

$f_N(x,y) =$ cantidad esperada recogida de oro antes de que el extractor sea averíe cuando A tiene x; B tiene y y se ha empleado una estrategia óptima que dura como mucho N etapas o extracciones.

Si consideramos el proceso unietápico, vemos que elegir A conduce a una cantidad esperada de oro dada por $p_A r_A x$, mientras que optar por B produce $p_B r_B y$. Por tanto

$$f_1(x,y) = \max\left(p_A r_A x, p_B r_B y\right)$$

Si nos situamos en el caso general de un proceso en N etapas, cualquiera que sea la elección inicial, la continuación sobre las $N-1$ etapas siguientes debe ser óptima si deseamos obtener una estrategia global óptima. Como en el caso de haber elegido A el retorno total esperado es

$$f(A,x,y) = p_A\left(r_A x + f_{N-1}\left((1-r_A)x, y\right)\right)$$

mientras que el retorno global esperado habiendo elegido B es

$$f(B,x,y) = p_B\left(r_B y + f_{N-1}\left(x,(1-r_B)y\right)\right)$$

y obtenemos para el retorno total la relación de recurrencia

$$\begin{aligned} f_N(x,y) &= \max\left[f(A,x,y), f(B,x,y)\right] \\ &= \max\left[p_A\left(r_A x + f_{N-1}\left((1-r_A)x, y\right)\right), p_B\left(r_B y + f_{N-1}\left(x,(1-r_B)y\right)\right)\right] \end{aligned}$$

Al igual que nos ocurrió en los modelos de asignación podemos plantearnos que sucedería si el número de etapas crece indefinidamente. En este caso, el retorno total $f(x,y)$, suponiendo que exista, satisface la ecuación funcional

$$f(x,y) = \max\left[p_A\left(r_A x + f\left((1-r_A)x, y\right)\right), p_B\left(r_B y + f\left(x,(1-r_B)y\right)\right)\right]$$

Como antes también se hace necesario establecer las condiciones para la existencia y unicidad de las soluciones de esta ecuación.

V.2.6. Control óptimo de sistemas lineales

La programación dinámica es un método que también se puede utilizar para el diseño de controles óptimos. En este caso, la función objetivo puede ser: el tiempo necesario para moverse de una condición operativa a otra, la suma de errores en estado estacionario, la suma de errores transitorios, el coste de operación, la amplitud o la energía de las señales de control, etc. En los controles habituales la función objetivo suele ser el sumatorio en el tiempo, cuando se trata de modelos de tiempo discreto, de los cuadrados de los errores y de las variables de entrada.

Supongamos que tenemos un sistema dinámico lineal expresado en la forma

$$x_{k+1} = A_k x_k + B_k u_k \tag{1.174}$$

siendo el valor inicial x_0 y las matrices A_k y B_k conocidas, x_k y u_k los n estados y m controles en el intervalo de muestreo k, $k = 0,\dots,N$. La función objetivo cuadrática se define como

$$J = \sum_{k=0}^{N-1}\left(x_k^T Q_k x_k + u_k^T R_k u_k\right) + \frac{1}{2}x_N^T Q_N x_N \tag{1.175}$$

siendo Q_k y R_k matrices cuadradas simétricas de la dimensión apropiada. Q_k es una matriz semidefinida positiva y R_k es definida positiva. El sumatorio total se ha dividido en dos partes: 0 a $N-1$ y N. No tienen sentido los controles u_N en el intervalo N. Las matrices Q_k y R_k puede ser iguales o diferentes en cada intervalo. Esta función objetivo cuadrática trata de mantener el sistema próximo al origen puesto que induce penalizaciones elevadas para grandes desviaciones y penalizaciones pequeñas para desviaciones menores.

El problema de programación dinámica consiste en encontrar la secuencia de controles óptimos $u_0, u_1, u_2, \ldots, u_{N-1}$ que hacen que la función objetivo sea mínima sujeto a las restricciones propias del sistema dinámico lineal.

Veamos cómo se resuelve este problema mediante programación dinámica hacia atrás. Empezando en la etapa N el valor de la función objetivo será

$$J_N^* = x_N^T Q_N x_N \tag{1.176}$$

La función objetivo en el intervalo $N-1$ se puede expresar como

$$J_{N-1} = \left(x_{N-1}^T Q_{N-1} x_{N-1} + u_{N-1}^T R_{N-1} u_{N-1}\right) + J_N^* \tag{1.177}$$

Expandimos ahora el último término

$$J_{N-1} = \left(x_{N-1}^T Q_{N-1} x_{N-1} + u_{N-1}^T R_{N-1} u_{N-1}\right) + x_N^T Q_N x_N \tag{1.178}$$

$$\begin{aligned} J_{N-1} &= \left(x_{N-1}^T Q_{N-1} x_{N-1} + u_{N-1}^T R_{N-1} u_{N-1}\right) + \\ &+ \left(A_{N-1} x_{N-1} + B_{N-1} u_{N-1}\right)^T Q_N \left(A_{N-1} x_{N-1} + B_{N-1} u_{N-1}\right) \end{aligned} \tag{1.179}$$

$$\begin{aligned} J_{N-1} &= \left(x_{N-1}^T Q_{N-1} x_{N-1} + u_{N-1}^T R_{N-1} u_{N-1}\right) + \\ &+ x_{N-1}^T A_{N-1}^T Q_N A_{N-1} x_{N-1} + u_{N-1}^T B_{N-1}^T Q_N B_{N-1} u_{N-1} + 2 x_{N-1}^T A_{N-1}^T Q_N B_{N-1} u_{N-1} \end{aligned} \tag{1.180}$$

Para determinar el óptimo se deriva J_{N-1} con respecto a u_{N-1} y se iguala a cero dicha derivada

$$\frac{\partial J_{N-1}}{\partial u_{N-1}} = \left(R_{N-1} + B_{N-1}^T Q_N B_{N-1}\right) u_{N-1} + B_{N-1}^T Q_N A_{N-1} x_{N-1} = 0$$

$$u_{N-1}^{*} = -\left(R_{N-1} + B_{N-1}^{T} Q_N B_{N-1}\right)^{-1} B_{N-1}^{T} Q_N A_{N-1} x_{N-1}$$

Si se sustituye el valor del control óptimo en el intervalo $N-1$, u_{N-1}^{*}, en la ecuación de la función objetivo de dicho intervalo J_{N-1} resulta ser

$$\begin{aligned}
&J_{N-1}^{*} = \left[x_{N-1}^{T}\left(Q_{N-1} + A_{N-1}^{T} Q_N A_{N-1}\right) x_{N-1} + u_{N-1}^{*T}\left(R_{N-1} + B_{N-1}^{T} Q_N B_{N-1}\right) u_{N-1}^{*}\right] - \\
&-2 x_{N-1}^{T} A_{N-1}^{T} Q_N B_{N-1}\left(R_{N-1} + B_{N-1}^{T} Q_N B_{N-1}\right)^{-1} B_{N-1}^{T} Q_N A_{N-1} x_{N-1} = \\
&x_{N-1}^{T}\left(Q_{N-1} + A_{N-1}^{T} Q_N A_{N-1}\right) x_{N-1} \\
&-x_{N-1}^{T} A_{N-1}^{T} Q_N B_{N-1}\left(R_{N-1} + B_{N-1}^{T} Q_N B_{N-1}\right)^{-1} B_{N-1}^{T} Q_N A_{N-1} x_{N-1}
\end{aligned}$$

es decir, $J_{N-1}^{*} = x_{N-1}^{T} K_{N-1} x_{N-1}$, con

$$K_{N-1} = A_{N-1}^{T}\left[Q_N - Q_N B_{N-1}\left(R_{N-1} + B_{N-1}^{T} Q_N B_{N-1}\right)^{-1} B_{N-1}^{T} Q_N\right] A_{N-1} + Q_{N-1}$$

Ahora se pasa al intervalo $N-2$ y se obtiene la regla general del control óptimo para cualquier etapa k.

$$u_k^{*} = L_k x_k$$

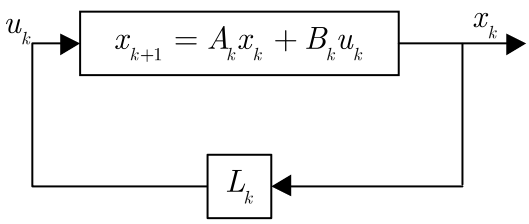
donde la ganancia L_k viene dada por la ecuación

$$L_k = -\left(R_k + B_k^{T} K_{k+1} B_k\right)^{-1} B_k^{T} K_{k+1} A_k$$

estando las matrices K_k definidas recursivamente por estas expresiones

$$\begin{aligned}
K_N &= Q_N \\
K_k &= A_k^{T}\left[K_{k+1} - K_{k+1} B_k\left(R_k + B_k^{T} K_{k+1} B_k\right)^{-1} B_k^{T} K_{k+1}\right] A_k + Q_k
\end{aligned}$$

Esta solución del problema resulta especialmente atractiva porque cada estado x_k es realimentado como control u_k mediante una ganancia lineal L_k.



La última ecuación se conoce con el nombre de ecuación de Riccati de tiempo discreto.

Cuando las matrices A_k, B_k, Q_k y R_k son constantes, entonces para tiempo infinito las matrices K_k convergen al estado estacionario K satisfaciendo la denominada ecuación de Riccati algebraica

$$K = A^T \left[K - KB\left(R + B^T KB\right)^{-1} B^T K \right] A + Q$$

La ley de control del sistema se convierte ahora en

$$u^* = Lx \tag{1.181}$$

y la ganancia en estado estacionario es

$$L = -\left(R + B^T KB\right)^{-1} B^T KA \tag{1.182}$$

V.3. Biblioteca de problemas

PROBLEMA 1

Una empresa de aparatos electrónicos tiene un contrato para entregar el siguiente número de radios durante los tres meses siguientes: mes 1 200 radios, mes 2 300 y mes 3 otras 300. Por cada radio que se produce los meses 1 y 2, se incurre en un coste variable de 10 € y si es en el mes 3 de 12 €. El coste de almacenamiento es de 1.5 € por cada radio en inventario al final de un mes. El coste de preparar la producción durante un mes es de 250 €. Las radios que se fabrican durante un mes pueden servir para abastecer la demanda de ese mes o de alguno futuro. Suponga que la producción durante cada mes ha de ser múltiplo de 100. Dado que el nivel inicial de inventario es 0, utilizar la programación dinámica para dar un calendario óptimo de producción.

PROBLEMA 2

Un dispositivo electrónico está compuesto por tres componentes. Las tres componentes están en serie con lo que el fallo de una de ellas provoca el fallo del dispositivo. La fiabilidad (probabilidad de no fallar) del dispositivo puede ser mejorada instalando una o dos componentes de reserva en cada tipo de componente. La siguiente tabla muestra la fiabilidad r y el coste en miles de euros. El presupuesto disponible para la construcción del dispositivo es de un millón de euros. ¿Cómo se debe construir el dispositivo? (Nota: el objetivo es maximizar la fiabilidad $r_1 r_2 r_3$ del dispositivo, lo que implica que la descomposición del objetivo es más multiplicativa que aditiva.)

Número de unidades	Componente 1		Componente 2		Componente 3	
en paralelo	r_1	c_1	r_2	c_2	r_3	c_3
1	0.6	100	0.7	300	0.5	200
2	0.8	200	0.8	500	0.7	400
3	0.9	300	0.9	600	0.9	500

PROBLEMA 3

El ingeniero responsable de una central eléctrica tiene que decidir cada día, a la vista de la demanda prevista para ese día, cuáles de los dos generadores de 600 y 1000 MW de que dispone deben ser puestos en funcionamiento y en qué momento. La demanda requerida varía a lo largo de cuatro periodos en los que el día se considera dividido a estos efectos, de acuerdo con la siguiente tabla:

Periodo	Demanda [MW]
0 a.m. – 6 a.m.	400
6 a.m. – 12 m.	900
12 m. – 18 p.m.	1200
18 p.m. – 24 p.m.	500

Al comienzo de cada periodo los generadores pueden ser conectados o desconectados. Los costes de arranque de los generadores son, respectivamente, de 100 y 75 miles de euros respectivamente, en tanto que los costes operativos (costes por usar el generador, no proporcionales a la demanda que satisfagan) por cada periodo de utilización son de 175 y 210 miles de euros respectivamente. Si los dos generadores son siempre desconectados a media noche y pueden volver a arrancar de forma inmediata, mediante programación dinámica obtener la programación de funcionamiento de los generadores que minimiza el coste diario.

PROBLEMA 4

Con la llegada del verano, la empresa artesanal HeladoSA desea determinar la producción de helados de turrón que debe realizar para la temporada estival. La producción mensual máxima de la empresa es de cuatro mil helados.

Tras realizar un estudio de mercado, la empresa estima que habrá una demanda de 3000 helados en julio, 2000 helados en agosto y 1000 helados en septiembre. Los pedidos de helados deben hacerse en cajas de 1000 unidades. Con el fin de simplificar el cálculo de costes, se considerará que cada helado almacenado al final de mes supondrá un coste de 10 €. Los gastos de producción dependen del número de cajas de 1000 unidades producidas mensualmente según se muestra en la tabla.

Miles de unidades producidas	1	2	3	4
Coste unitario por helado (euros)	10	10	5	4

Se supone que los helados tienen una caducidad de 6 meses, por lo que los helados que no se vendan un mes servirán para el mes siguiente. La empresa parte de un inventario de cero unidades y no desea ningún inventario al final del verano. Se pide:

1) Determine cuál debe ser la producción de la empresa durante los meses de junio, julio y agosto.

2) Razone el resultado del apartado anterior, ¿por qué sale esa solución? ¿qué modificaciones variarían la solución?

PROBLEMA 5

Una empresa de televisión desea gestionar su programación anual en las horas de prime time nocturnas (20 a 23 h). Para llenar esas horas dispone de tres programas de éxito reconocido: Informativos 2, el magazine Operación Musical y la serie Hospital Comarcal. La programación anual se divide en las temporadas de invierno, primavera, verano y otoño. Los índices de audiencia (estimados por una empresa especialista en investigación de mercados audiovisuales y expresados en número previsto de millones de espectadores) de cada programa cambian en cada temporada dependiendo de la hora de emisión de prime time y se muestran en la siguiente tabla:

	Invierno	Invierno	Invierno	Primavera	Primavera	Primavera
	20–21	21–22	22–23	20–21	21–22	22–23
Informativos 2	2.5	2.25	2	2.7	2.9	2.95
Hospital Comarcal	2.7	3.6	4.2	4	3.6	3.4
Operación Musical	3.5	3.6	3.8	3	3.1	2.8

	Verano	Verano	Verano	Otoño	Otoño	Otoño
	20–21	21–22	22–23	20–21	21–22	22–23
Informativos 2	2.6	2.35	2.2	2.7	2.9	2.95
Hospital Comarcal	2.7	3.7	4.3	4	3.6	3.4
Operación Musical	3.3	3.5	3.9	3	3.5	2.8

Al final del año pasado la empresa de televisión emitía en primer lugar Informativos 2, a continuación Hospital Comarcal y por último Operación Musical. La empresa quiere conocer la programación óptima para el presente año que maximice el número de espectadores previstos, sabiendo que cada cambio en la parrilla horaria de los programas entre una temporada y la siguiente supone una pérdida de 250000 espectadores. Adicionalmente, la Dirección General ha impuesto la restricción de que entre temporadas consecutivas no se produzcan tres cambios de horario de emisión entre la parrilla de una temporada y la anterior. Por el éxito tenido en los años anteriores se ha decidido que el orden de emisión, no necesariamente consecutivo, de Hospital Comarcal primero y Operación Musical después se mantenga tanto en invierno como en primavera, aunque es revisable para las otras dos temporadas, y que en verano se emita Operación Musical el último mientras que en otoño se emitirá Informativos 2 en primer lugar. Proponer un modelo de optimización para este problema, y resolverlo.

PROBLEMA 6

Una compañía aérea tiene una flota de 15 aeronaves: 5 de cada uno de tres tipos A, B y C, cuyas respectivas capacidades para el transporte de viajeros son de 80, 70 y 55 personas. Una agencia de viajes le solicita presupuesto para trasladar a 372 personas. La compañía analiza sus costes, que dependen del número de aviones de cada tipo que quiera utilizar para transportar a esas personas, datos que se dan en la siguiente tabla en miles de euros

Tipo	1	2	3	4	5
A	11	20	30	40	50
B	9	17	24	34	45
C	8	15	21	26	31

Además la compañía aérea incurre en un coste fijo adicional de 6 k€ por cada tipo de aviones que utilice.

Proponer un modelo de programación dinámica (etapas, decisiones, estados, función en cada etapa, transiciones, ...) cuyo objetivo sea determinar la composición óptima de la flotilla de aviones que va a realizar el transporte para minimizar los costes de la operación y mostrar la resolución de la primera etapa.

PROBLEMA 7

RentCar está desarrollando un plan de reemplazamiento para su flota de 1 coche para los próximos 5 años (2005-2009). Se supone la flota adquirida a 1 de enero de 2005 con un precio por coche de 20000 € y necesariamente a 1 de enero 2009 empieza con otra flota totalmente nueva (liquidando todos los coches). Al principio de cada año, se decide si un coche debe ser mantenido o reemplazado. Un coche debe estar en servicio al menos un año y a lo sumo tres. En la siguiente tabla se dan los costes de reemplazamiento (en €) por coche en función del año en que es adquirido y de los años que lleva en funcionamiento.

Año adquisición	Años funcionando		
	1	2	3
1-ene-2005	4000	5400	9800
1-ene-2006	4300	6200	8700
1-ene-2007	4800	7100	–
1-ene-2008	4900	–	–

Además se considera un coste de mantenimiento anual de 200 € el primer año de uso, 300 en segundo y 500 el tercero.

Resolver el problema de encontrar la política óptima que se debe seguir para minimizar el coste mediante programación dinámica.

PROBLEMA 8

Un alumno dispone de 5 días para preparar tres asignaturas. Dependiendo del número de días que dedique a cada asignatura obtendrá diferentes puntuaciones según se muestra en la tabla.

	Asignaturas		
Días dedicados	CS	ITF	MM
0	0	0	0
1	4	5	5
2	5	6	7
3	6	7	9
4	8	8	9
5	10	9	10

Desarrollar un modelo de programación dinámica que optimice la distribución de tiempo en orden a obtener una puntuación global media máxima, dando la solución óptima.

V.4. Resultados de la biblioteca de problemas

RESULTADO DEL PROBLEMA 1

Etapas: principio de mes. Estados: nivel de inventario. Decisiones: cantidad a producir. La solución es producir 200 unidades en el primer mes, 600 en el segundo y no producir en el tercero. El coste total es 8950 €.

RESULTADO DEL PROBLEMA 2

Etapas: las piezas. Estados: capital disponible. Decisiones: 1, 2 ó 3 componentes en paralelo. La solución es poner dos piezas de tipo 1, una de tipo 2 y tres de tipo 3. La fiabilidad es 0.504.

RESULTADO DEL PROBLEMA 3

Etapas: periodos. Estados: grupos acoplados o parados. Decisiones: uso o no del generador en el periodo. La solución es arrancar inicialmente el 2 (grande) en la etapa 0, arrancar el 1 (pequeño) en la etapa 3 y parar el 2 (grande) en la etapa 4. Coste total: 1155 k€

RESULTADO DEL PROBLEMA 5

Las etapas son las temporadas. Los estados en cada etapa son las seis combinaciones posibles de orden de emisión de los tres programas. Los denominamos $(1,2,3)_k$, $(1,3,2)_k$, $(2,1,3)_k$, $(2,3,1)_k$, $(3,1,2)_k$, $(3,2,1)_k$, para una etapa k cualesquiera.

La función de beneficio (número de espectadores) a maximizar en cada etapa estará formada por el número de telespectadores de la configuración de esa etapa, más el que corresponda por cambio de horario. A esto se añade el coste pasado para llegar al estado (hacia delante), o el coste futuro si es hacia atrás.

Resolviendo hacia delante: Partimos del estado inicial $(1,2,3)_0$.

	$(1,2,3)_0$	
$(1,2,3)_1$	2.5+3.6+3.8=9.9	$(1,2,3)_0$
$(2,1,3)_1$	2.25+2.7+3.8–0.25–0.25=8.25	$(1,2,3)_0$
$(2,3,1)_1$	2.25+4.2+3.5–0.25–0.25–0.25=9.2	$(1,2,3)_0$

	$(1,2,3)_1$	$(2,1,3)_1$	$(2,3,1)_1$		
$(1,2,3)_2$	9.9+9.1=19	8.25+9.1–0.5=16.85	–	19	$(1,2,3)_1$
$(2,1,3)_2$	9.9+9.7–0.5=19.1	8.25+9.7=17.95	9.2+9.7–0.5=18.4	19.1	$(1,2,3)_1$
$(2,3,1)_2$	–	8.25+9.3–0.5=17.05	9.2+9.3=18.5	18.5	$(2,3,1)_1$

	$(1,2,3)_2$	$(2,1,3)_2$	$(2,3,1)_2$		
$(1,2,3)_3$	19+10.2=29.2	19.1+10.2–0.5=28.8	–	29.2	$(1,2,3)_2$
$(2,1,3)_3$	19+8.95–0.5=27.45	19.1+8.95=28.05	18.5+8.95–0.5=26.95	28.05	$(2,1,3)_2$

	$(1,2,3)_3$	$(2,1,3)_3$		
$(1,2,3)_4$	29.2+9.1=38.3	28.05+9.1–0.5=36.65	38.3	$(1,2,3)_3$
$(1,3,2)_4$	29.2+9.6–0.5=38.3	–	38.3	$(1,2,3)_3$

Luego la solución óptima es emitir los programas en el siguiente orden $(1,2,3)_1$, $(1,2,3)_2$, $(1,2,3)_3$, $(1,2,3)_4$ ó $(1,3,2)_4$ con un total de 38.3 millones de espectadores.

Resolviendo hacia atrás (estados en las filas y decisiones en columnas):

	$(1,2,3)_4$	$(1,3,2)_4$		
$(1,2,3)_3$	9.1	9.6–0.5=9.1	9.1	$(1,2,3)_4$ o $(1,3,2)_4$
$(2,1,3)_3$	9.1–0.5=8.6	–	8.6	$(1,2,3)_4$

	$(1,2,3)_3$	$(2,1,3)_3$		
$(1,2,3)_2$	10.2+9.1=19.3	8.95+8.6–0.5=17.05	19.3	$(1,2,3)_3$
$(2,1,3)_2$	10.2–0.5+9.1=18.8	8.95+8.6=17.55	18.8	$(1,2,3)_3$
$(2,3,1)_2$	–	8.95+8.6–0.5=17.05	17.05	$(2,1,3)_3$

	$(1,2,3)_2$	$(2,1,3)_2$	$(2,3,1)_2$		
$(1,2,3)_1$	9.1+19.3=28.4	9.7–0.5+18.8=28	–	28.4	$(1,2,3)_2$
$(2,1,3)_1$	9.1–0.5+19.3=27.9	9.7+18.8=28.5	9.3+17.05–0.5=25.85	28.5	$(2,1,3)_2$
$(2,3,1)_1$	–	9.7–0.5+18.8=28	9.3+17.05=26.35	28	$(2,1,3)_2$

	$(1,2,3)_1$	$(2,1,3)_1$	$(2,3,1)_1$ (no posible)		
$(1,2,3)_0$	9.9+28.4=38.3	8.25+28.5–0.5=36.25	9.2+28–0.5=36.7	38.3	$(1,2,3)_1$

RESULTADO DEL PROBLEMA 6

Etapas: tipo de aviones (A, B, C)

Decisiones: número de aviones de ese tipo

Estados: capacidad acumulada

Función objetivo: coste de los aviones decididos de ese tipo + 6 si es distinto de 0 + coste acumulado

Transiciones: de un estado a otro se pasa sumando la capacidad acumulada

A	0	1	2	3	4	5
(372)	0 (0)	17 (80)	26 (160)	36 (240)	46 (320)	56 (400)

B	0	1	2	3	4	5
0	0 (0)	17 (70)	23 (140)	30 (210)	40 (280)	51 (350)
80	17 (80)	34 (150)	40 (220)	47 (290)	57 (360)	68 (430)
160	26 (160)	43 (230)	49 (300)	56 (370)	66 (440)	Sobra
240	36 (240)	53 (310)	59 (380)	Sobra	Sobra	Sobra
320	46 (320)	63 (390)	Sobra	Sobra	Sobra	Sobra
400	56 (400)	Sobra	Sobra	Sobra	Sobra	Sobra

C	0	1	2	3	4	5	
0	Falta	Falta	Falta	Falta	Falta	Falta	Falta
70	Falta	Falta	Falta	Falta	Falta	Falta	Falta
80	Falta	Falta	Falta	Falta	Falta	Falta	Falta
140	Falta	Falta	Falta	Falta	Falta	31+6+23	60
150	Falta	Falta	Falta	Falta	Falta	31+6+34	71
160	Falta	Falta	Falta	Falta	26+6+26	Sobra	58
210	Falta	Falta	Falta	21+6+30	Sobra	Sobra	57
220	Falta	Falta	Falta	21+6+40	Sobra	Sobra	67
230	Falta	Falta	Falta	21+6+43	Sobra	Sobra	70
240	Falta	Falta	Falta	21+6+36	Sobra	Sobra	63
280	Falta	Falta	15+6+40	Sobra	Sobra	Sobra	61
290	Falta	Falta	15+6+47				68
300	Falta	Falta	15+6+49				70
310	Falta	Falta	15+6+53				74
320	Falta	8+6+46					60
350	Falta	8+6+51					65
360	Falta	8+6+57					71
370	Falta	8+6+56					70
380	59						59
390	63						63
400	56						56
430	68						68
440	66						66
Mínimo:							56

Mínimo: 56 k€ con 0 de tipo C, 0 de tipo B y 5 de tipo A.

RESULTADO DEL PROBLEMA 7

Las etapas son los años, de 2006 a 2009. Los estados de cada etapa son el año de antigüedad del coche. Las decisiones al principio de cada etapa son comprar o no coche.

	A^{2005}		
A^{2006}	4000+200=4200	A^{2005}	4200
$C^{2006–A2005}$	200	A^{2005}	200

	A^{2006}	$C^{2006-A2005}$		
A^{2007}	4300+200+4200=8700	5400+300+200=5900	5900	$C^{2006-A2005}$
$C^{2007-A2006}$	200+4200=4400	–	4400	A^{2006}
$C^{2007-A2005}$	–	300+200=500	500	$C^{2006-A2005}$

	A^{2007}	$C^{2007-A2006}$	$C^{2007-A2005}$		
A^{2008}	4800+200+5900=10900	9800+300+4400=14500	9800+500+500=10800	10800	$C^{2007-A2005}$
$C^{2008-A2007}$	200+5900=6100	–	–	6100	A^{2007}
$C^{2008-A2006}$	–	300+4400=4700	–	4700	$C^{2007-A2006}$

	A^{2008}	$C^{2008-A2007}$	$C^{2008-A2006}$		
A^{2009}	4900+200+10800=15900	7100+300+6100=13500	8700+500+4700=13900	13500	$C^{2008-A2007}$

La solución óptima es adquirir coche en 2007 y el resto de los años no comprarlo.

RESULTADO DEL PROBLEMA 8

Etapas: asignaturas. Decisiones: días a dedicar. Estados: días ya utilizados.

Resolución hacia atrás:

MM

Decisiones / Estados	0	1	2	3	4	5	Ópt	f.ópt.
0	0	5	7	9	9	10	5	10
1	0	5	7	9	9	–	3 ó 4	9
2	0	5	7	9	–	–	3	9
3	0	5	7	–	–	–	2	7
4	0	5	–	–	–	–	4	5
5	0	–	–	–	–	–	0	0

ITF

Decisiones / Estados	0	1	2	3	4	5	Ópt	f.ópt.
0	10	14	15	14	13	9	2	15
1	9	14	13	12	8	–	1	14
2	9	12	11	7	–	–	1	12
3	7	10	6	–	–	–	1	10
4	5	5	–	–	–	–	0.1	5
5	0	–	–	–	–	–	0	0

CS

Decisiones / Estados	0	1	2	3	4	5	Ópt	f.ópt.
0	15	18	17	16	13	10	1	18

Solución: dedicar a CS 1 día, a ITF 1 día y a MM 3 días. Notas: 4, 5 y 9. Nota media 6.

	0	Nota total acumulada	Decisión óptima
0_{MM}	0	0	0_{MM}
1_{MM}	5	5	1_{MM}
2_{MM}	7	7	2_{MM}
3_{MM}	9	9	3_{MM}
4_{MM}	9	9	4_{MM}
5_{MM}	10	10	5_{MM}

	0_{MM}	1_{MM}	2_{MM}	3_{MM}	4_{MM}	5_{MM}	Nota total acumulada	Decisión óptima
0_{ITF}	0+0	–	–	–	–	–	0	0_{ITF}
1_{ITF}	0+5	5+0	–	–	–	–	5	1_{ITF}, 0_{ITF}
2_{ITF}	0+6	5+5	7+0	–	–	–	10	1_{ITF}
3_{ITF}	0+7	5+6	7+5	9+0	–	–	12	1_{ITF}
4_{ITF}	0+8	5+7	7+6	9+5	9+0	–	14	1_{ITF}
5_{ITF}	0+9	5+8	7+7	9+6	9+5	10+0	15	2_{ITF}

	0_{ITF}	1_{ITF}	2_{ITF}	3_{ITF}	4_{ITF}	5_{ITF}	Nota total acumulada	Decisión óptima
0_{CS}	0+0	–	–	–	–	–	0	0_{CS}
1_{CS}	0+4	5+0	–	–	–	–	5	0_{CS}
2_{CS}	0+5	5+4	10+0	–	–	–	10	1_{CS}
3_{CS}	0+6	5+5	10+4	12+0	–	–	14	1_{CS}
4_{CS}	0+8	5+6	10+5	12+4	14+0	–	16	1_{CS}
5_{CS}	0+10	5+8	10+6	12+5	14+4	15+0	18	1_{CS}

VI. Programación matemática multicriterio

VI.1. Introducción: Conceptos básicos

En la vida real, y tanto en el ámbito profesional como el personal, nos vemos enfrentados a multitud de situaciones en las que tenemos que decidir entre varias alternativas. La propia optimización no es más que una forma de tomar una decisión entre unas alternativas factibles.

Así, en su dimensión más básica, un proceso de toma de decisión puede entenderse como la elección de lo "mejor" entre lo "posible". Ahora bien, según se defina qué es lo mejor y qué es lo posible nos enfrentaremos a distintas situaciones de decisión.

La optimización clásica tiene como característica general que lo mejor, el objetivo, es único y está claramente determinado y que lo posible, las soluciones factibles, no vienen expresadas explícitamente sino en forma de restricciones y sin incertidumbre (excepto en optimización estocástica, que no es precisamente clásica). Pero además de estos contextos de decisión de optimización clásica, existen otros que configuran lo que se suele denominar en términos amplios la teoría de la decisión.

Dentro de este contexto surge la *decisión multicriterio*, en la que si bien dada una decisión sus consecuencias están perfectamente determinadas, lo que no está definido tan claramente es qué es lo mejor, existiendo varios objetivos en conflicto.

Respecto a lo posible, se trata de establecer las alternativas o puntos factibles existentes. El conjunto puede ser discreto o continuo. En general, se considera discreto y se aplica la metodología apropiada cuando es factible enumerar y tratar explícitamente cada una de las alternativas posibles. En el caso continuo o caso discreto donde no viene explícitamente definido el conjunto de alternativas es cuando se habla de conjunto o región factible. Este conjunto o región factible, a su vez, puede venir definido de forma rígida mediante restricciones o de forma más flexible mediante lo que se conoce como *metas* y *niveles de aspiración.*

Respecto a lo mejor, se puede definir según un único criterio o según varios criterios. Los problemas de decisión con un único criterio y conjunto factible continuo (entendiendo por éste la extensión a conjuntos discretos no definidos explícitamente) son básicamente problemas de optimización "clásica": lineal, entera o no lineal. Si además incluyen aleatoriedad, estaríamos ante un problema de optimización estocástica. Si el conjunto factible es discreto, sólo tiene sentido plantearse el problema en el caso de que haya aleatoriedad, siendo entonces un problema de los conocidos como problemas clásicos de decisión.

En el caso en que haya varios criterios, si la región factible es continua, se puede resolver el problema mediante métodos denominados de *optimización multiobjetivo, métodos de compromiso* o mediante *métodos satisfacientes* (*programación por metas*). Si lo posible viene definido por un conjunto discreto de alternativas (pudiendo incluso no ser numérico el valor de los criterios), existen métodos multicriterio discretos para resolver el problema.

VI.1.1. Conceptos básicos

De forma general un problema de decisión multicriterio vendría formulado de la siguiente forma:

$$\begin{aligned}&\text{opt } z(x) = (z_1(x), ..., z_p(x))\\&x \in F\end{aligned}$$

donde F es el espacio de decisiones o soluciones (si es continuo, se denomina región factible, $F \subseteq \mathbb{R}^n$).

Al conjunto $z(F)$ se le denomina *espacio de objetivos* o *resultados* (en el caso de que sean criterios numéricos $z(F) \subseteq \mathbb{R}^p$).

A continuación se exponen algunos conceptos básicos de la decisión multicriterio:

Atributo: "valor" observado (medido) de una decisión independientemente del decisor. Los atributos suelen ser competetidores o contradictorios entre sí.

Objetivo: dirección de mejora de un atributo. Esta dirección será de maximización o minimización en el caso de atributos numéricos y en el caso de atributos no numéricos vendrá dado por un sistema de preferencias (por ejemplo, si el problema fuera la selección de un automóvil, el color sería un atributo no numérico y se establecería un sistema de preferencias sobre este atributo).

Nivel de aspiración: es un nivel aceptable de logro para un atributo.

Meta: es la combinación de un atributo con su nivel de aspiración.

Criterio: son los atributos, objetivos o metas relevantes en un problema de decisión.

Dadas estas definiciones previas, lo primero que se necesita en un problema de decisión multicriterio es dar un concepto de solución. El concepto de solución utilizado habitualmente es el concepto de óptimo de Pareto. Este concepto está basado en el *criterio de optimalidad paretiana,* enunciado por Pareto en 1896:

"Una alternativa es *eficiente* (o *Pareto óptima*) si toda alternativa que proporcione una mejora en un atributo produce un empeoramiento en al menos otro de los atributos."

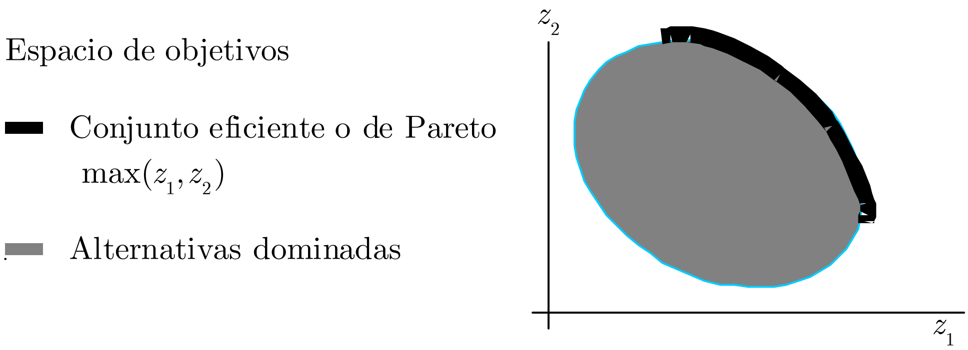
De esta definición, se deriva la de *alternativa dominada* o *no eficiente,* como una alternativa para la que existe otra alternativa con todos los atributos iguales o mejores.

Así se define el *conjunto eficiente* o *de Pareto* o también llamada *frontera de Pareto* (por su representación en el caso continuo como parte de la frontera del espacio de objetivos o resultados) como el conjunto de alternativas eficientes o no dominadas. Es decir, el conjunto de alternativas tales que no existe otra con todos los atributos preferidos o iguales.

Para atributos numéricos con objetivos de maximización la frontera de Pareto se expresa como

$$\varepsilon = \left\{ x \in F : \nexists x' \in F / z_k(x') \geq z_k(x) \ \forall k \text{ y } z_t(x') > z_t(x) \text{ para al menos un } t \in \{1, \ldots, p\} \right\}$$

Un ejemplo gráfico para dos atributos que se desea que tomen el mayor valor posible podría ser el siguiente:



Sin embargo, en la mayoría de las situaciones, el fin último es dar una única solución, no un conjunto de posibles soluciones. Se denomina *solución de mejor compromiso* a la solución del conjunto eficiente que es seleccionada por el decisor.

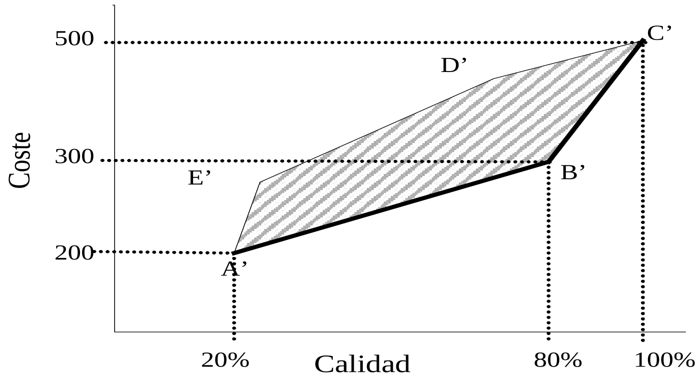
Otro concepto fundamental en el análisis de decisiones multicriterio es la matriz de pagos, que respresenta el grado de conflicto entre los objetivos. La *matriz de pagos* (pay-off matrix) es una matriz donde se representan por filas el valor óptimo de un atributo sin considerar el resto de atributos (resolviendo el problema independientemente), y los valores que resultarían para los demás objetivos con esa solución. Esta matriz representa el grado de conflicto que hay entre los objetivos.

Por ejemplo, supóngase un problema en que hay que determinar la composición de una mezcla a partir de unas ciertas componentes básicas. Supóngase que se desea que el coste de la mezcla sea mínimo pero que la calidad sea máxima, medida ésta como una función de las componentes. Se tendrían entonces dos objetivos contrapuestos: minimizar el coste y maximizar la calidad. La matriz de pagos sería de la forma

	Coste	Calidad
Coste	m_{11}	m_{12}
Calidad	m_{21}	m_{22}

donde m_{11} es el mínimo coste que se puede lograr con las condiciones impuestas en el problema y m_{12} es la calidad para la composición que da el mínimo coste; en la siguiente fila sería al revés, es decir, m_{22} sería la máxima calidad que se puede lograr de la mezcla, y m_{21} sería el coste de esa mezcla. A la diagonal de esta matriz se le conoce como *punto ideal* ya que tiene los mejores valores posibles para cada uno de los atributos, y si existe realmente conflicto entre criterios, será un punto inalcanzable.

Así, si el espacio de objetivos fuera el siguiente, donde el conjunto eficiente está marcado en negrita, y se considera un problema continuo,



la matriz de pagos sería

	Coste	Calidad
Coste	**200**	20
Calidad	500	**100**

y el punto ideal sería (200,100), inalcanzable a la vista del espacio de objetivos.

Por otra parte, las *tasas de intercambio* (trade-offs o costes de oportunidad) entre los atributos representan lo que se está dispuesto a empeorar un atributo por mejorar en una unidad otro objetivo. Serían las pendientes de los segmentos que forman el conjunto eficiente. Así, en el segmento A'B' la tasa de intercambio entre el coste y la calidad será $T_{A'B'} = \frac{300-200}{80-20} = 1.66$, es decir, en ese segmento cada unidad más de calidad "cuesta" 1.66 unidades monetarias. Análogamente, para el segmento B'C' la tasa será $T_{B'C'} = \frac{500-300}{100-80} = 10$. Es decir, cada unidad más de calidad "cuesta" 10 unidades monetarias.

Problemas sencillos como éste pueden ser resueltos con Excel, utilizando el complemento que viene por defecto, Solver.

VI.1.2. Un ejemplo básico

Un empresario ha de producir un compuesto basado en un determinado componente básico. Este componente puede ir en una cantidad variable entre el 50 y el 100% del total de la composición. Por otra parte del componente básico hay dos calidades, una calidad superior que cuesta 200€/unidad y otra que cuesta 100€. El mercado le obliga a que al menos un 20% ha de ser del componente de calidad superior. El empresario se plantea maximizar la calidad y minimizar el coste. Plantear un modelo para este problema, mostrar la región factible o espacio de soluciones, mostrar el espacio de objetivos y la frontera de Pareto, y obtener las tasas de intercambio.

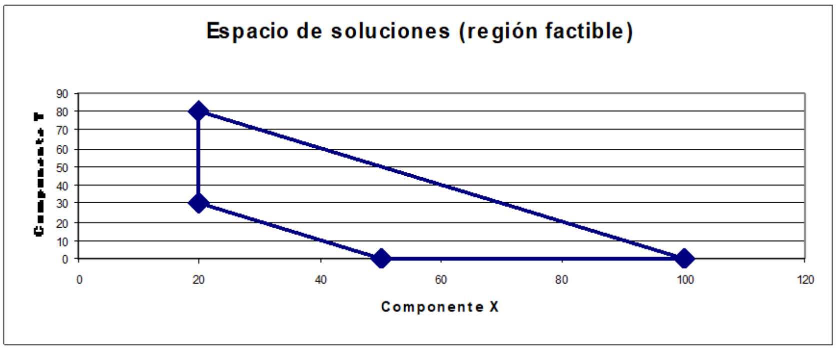
El siguiente modelo de optimización sirve para determinar el espacio de soluciones, donde *X* representa la calidad superior del componente e *Y* la calidad inferior:

$$
\begin{aligned}
&X \geq 20 \\
&X + Y \leq 100 \\
&X + Y \geq 50 \\
&X, Y \geq 0
\end{aligned}
$$

Por otra parte, los dos atributos planteados y su dirección de optimización, serían:

$$
\begin{array}{lll}
\text{Calidad:} & \max & X \\
\text{Coste:} & \min & 200X + 100Y
\end{array}
$$

La región factible o espacio de soluciones del modelo se presenta a continuación. Los vértices de esa región son puntos especialmente relevantes del problema.

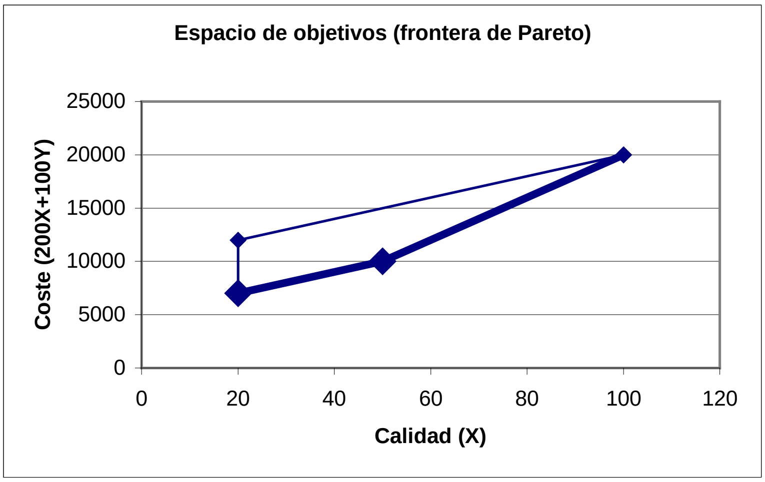


Las soluciones obtenidas optimizando cada criterio por separado forman la matriz de pagos del problema:

	Calidad	Coste
Calidad	**100**	20000
Coste	20	**7000**

en la que la diagonal forma el ideal (inalcanzable) que sería (100,7000).

Para ir más allá, y poder obtener las tasas de intercambio, es necesario obtener la frontera de Pareto. En este caso, se representan en el plano las imágenes según los atributos de Calidad y Coste de los vértices de la región factible o espacio de soluciones, obteniendo el espacio de objetivos del problema:



Y la frontera de Pareto es la formada por los puntos A (20, 7000), B (50,10000), C (100, 20000). Pueden observarse las siguientes tasas de intercambio:

$$T_{AB} = \frac{10000 - 7000}{50 - 20} = 100 \qquad T_{BC} = \frac{20000 - 10000}{100 - 50} = 200$$

Obsérvese que para obtener las tasas de intercambio ha sido necesario tener la Frontera de Pareto. Y por otra parte, para obtener la frontera de Pareto, hemos obtenido todo el espacio de objetivos, a partir de las imágenes de los vértices del espacio de soluciones. Sin embargo, este procedimiento es altamente costoso, y nada trivial. Por lo tanto, lo primero que se va a estudiar es cómo obtener la frontera de Pareto directamente a través de modelos de programación matemática, que son los métodos multiobjetivo de la siguiente sección.

VI.2. Métodos de optimización multiobjetivo

Estos métodos se utilizan para obtener la frontera de Pareto en un problema. En general, se plantean para conjuntos de alternativas continuos, ya que en el caso discreto la eliminación de alternativas dominadas se reduce a un simple algoritmo de comparación entre pares.

El planteamiento inicial del problema sería

$$\begin{aligned} &\max z = (z_1(x), ..., z_p(x)) \\ &x \in F \end{aligned}$$

donde $z_i(x)$ es la función matemática que describe el atributo i-ésimo, $x \subseteq \mathbb{R}^n$ es el vector de variables de decisión y F es el conjunto de restricciones que definen las posibles soluciones.

Dado que de forma natural surge que la decisión que debería tomarse en un problema con criterios múltiples debería encontrarse en la frontera de Pareto, es lógico querer conocer esta frontera, que incluso permite determinar las tasas de intercambio. Así dentro de los métodos de análisis existe la denominada optimización multiobjetivo, que son un conjunto de métodos para generar el conjunto eficiente en su totalidad.

Las técnicas generadoras del conjunto eficiente, son técnicas de carácter mecánico, en las que no se incluyen las preferencias del decisor. El fin para el que están diseñadas es la obtención de todo el conjunto eficiente, en general, tras aplicar métodos de programación paramétrica, aunque no llegaremos a ese nivel de profundidad.

VI.2.1. Método de las ponderaciones [Zadeh, 1963]

Este método consiste en multiplicar cada objetivo por un peso o factor no negativo y agregarlos en una única función. Variando los pesos se puede obtener todo el conjunto eficiente, obteniendo las distintas soluciones mediante programación paramétrica. Sin embargo, en este curso nos conformaremos con resolver los distintos problemas planteados para los valores de los parámetros mediante Solver.

$$\left.\begin{array}{l} \max \sum_{i=1}^{p} \lambda_i z_i(x) \\ x \in F \\ \lambda \geq 0 \end{array}\right\} P(\lambda)$$

Uno de los resultados en que se fundamenta este método dice que si $\lambda_i > 0 \quad \forall i$, entonces cualquier solución óptima del problema $P(\lambda)$ es eficiente. El recíproco es cierto sólo bajo ciertas condiciones (por ejemplo, si todas las funciones objetivo y las restricciones son lineales).

En cualquier caso, hay que tener en cuenta que para aplicar este método (y no sólo éste) es conveniente haber normalizado previamente los criterios (para que no influya la diferencia de unidades de los criterios).

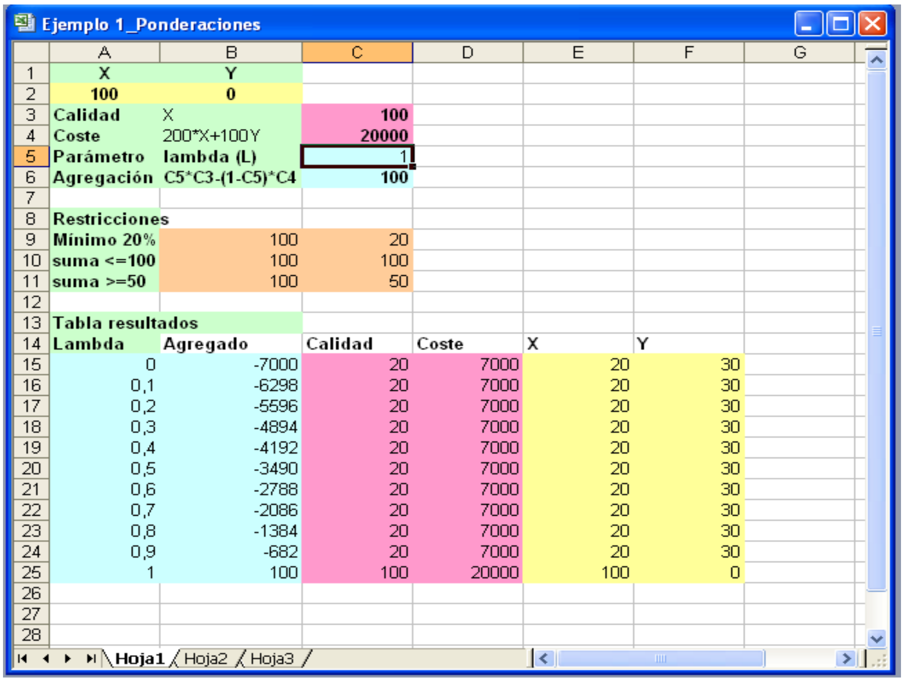
Sobre el ejemplo 1 de la sección anterior, obtener la frontera de Pareto mediante el método de las ponderaciones.

En primer lugar, hay que tener en cuenta que para agregar los objetivos ambos tienen que ir en el mismo sentido, es decir, no podemos tener uno de maximización (Calidad) y otro de minimización (Coste). Para ello basta con cambiar de signo la expresión de uno para que ya estén ambos en el mismo sentido. Por ejemplo, poniendo un menos en el Coste, ya será un criterio de maximización.

Por lo tanto, el modelo que hay que resolver para distintos valores de lambda es:

$$\begin{array}{l} \max \quad \lambda X + (1-\lambda)\left[-(200X + 100Y)\right] \\ \qquad X \geq 20 \\ \qquad X + Y \leq 100 \\ \qquad X + Y \geq 50 \\ X, Y \geq 0 \end{array}$$

Para ello utilizamos un archivo de Excel, con una celda para el valor de lambda que iremos introduciendo, y una celda para el objetivo agregado. Este modelo lo resolveremos cambiando los valores de lambda entre 0 y 1, e iremos guardando en una tabla los resultados obtenidos.



	A	B	C	D	E	F	G
1	X	Y					
2	100	0					
3	Calidad	X	100				
4	Coste	200*X+100Y	20000				
5	Parámetro	lambda (L)	1				
6	Agregación	C5*C3-(1-C5)*C4	100				
7							
8	Restricciones						
9	Mínimo 20%	100	20				
10	suma <=100	100	100				
11	suma >=50	100	50				
12							
13	Tabla resultados						
14	Lambda	Agregado	Calidad	Coste	X	Y	
15	0	-7000	20	7000	20	30	
16	0,1	-6298	20	7000	20	30	
17	0,2	-5596	20	7000	20	30	
18	0,3	-4894	20	7000	20	30	
19	0,4	-4192	20	7000	20	30	
20	0,5	-3490	20	7000	20	30	
21	0,6	-2788	20	7000	20	30	
22	0,7	-2086	20	7000	20	30	
23	0,8	-1384	20	7000	20	30	
24	0,9	-682	20	7000	20	30	
25	1	100	100	20000	100	0	
26							
27							
28							

Obsérvese que la solución es siempre la misma excepto para la última fila de la tabla. El problema es que la programación lineal siempre da un vértice de la región factible como solución, por lo que ésta se va a ir repitiendo hasta el coste pierda fuerza en el objetivo agregado. Por otra parte, dado que las unidades son muy distintas entre uno y otro criterio resulta muy difícil llegar a compensar el coste con la calidad. Para evitar este problema, puesto que con estos pocos valores de lambda y tal diferencia de unidades es difícil creer que ésa sea la frontera de Pareto (las combinaciones lineales de ambas soluciones), se puede normalizar cada uno de los criterios en la función agregada. Hay muchas formas de normalizar pero en decisión multicriterio, y dado que ya se tiene la matriz de pagos, una forma es restar al ideal el criterio y dividir por la diferencia entre el ideal y el anti-ideal, entendido como el peor valor en la matriz de pagos para ese criterio. En nuestro ejemplo, la matriz de pagos es:

	Calidad	Coste
Calidad	**100**	20000
Coste	20	**7000**

Entonces para nuestros criterios se plantearía minimizar la agregación de los siguientes criterios normalizados:

$$\frac{100 - Calidad}{100 - 20} \qquad \frac{7000 - Coste}{7000 - 20000}$$

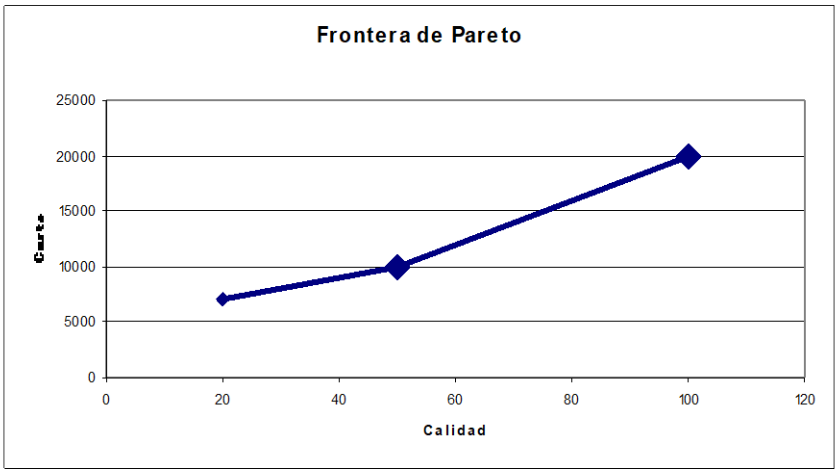
Obsérvese además que ahora ambos criterios son de minimización puesto que queremos aproximarnos en ambos casos al ideal de cada criterio.

Ejemplo 1_Ponderaciones

	A	B	C	D	E	F	G
1	X	Y					
2	100	0					
3	Calidad	X	100				
4	Coste	200*X+100Y	20000				
5	Parámetro	lambda (L)	1				
6	Agregación objetivos		0	C5*(100-C3)/(100-20)+(1-C5)*(7000-C4)/(7000-20000)			
7							
8	Restricciones						
9	Mínimo 20%	100	20				
10	suma <=100	100	100				
11	suma >=50	100	50				
12							
13	Tabla resultados						
14	Lambda	Agregado	Calidad	Coste	X	Y	
15	0	0	20	7000	20	30	
16	0,1	0,1	20	7000	20	30	
17	0,2	0,2	20	7000	20	30	
18	0,3	0,3	20	7000	20	30	
19	0,4	0,3884	50	10000	50	0	
20	0,5	0,4279	50	10000	50	0	
21	0,6	0,4	100	20000	100	0	
22	0,7	0,3	100	20000	100	0	
23	0,8	0,2	100	20000	100	0	
24	0,9	0,1	100	20000	100	0	
25	1	1	100	20000	100	0	
26							
27							
28							

Hoja1 Hoja2 Hoja3

Puede observarse que se han obtenido tres puntos distintos, cuyos valores de los criterios son: (20,7000), (50,10000) y (100,20000), cuya unión da lugar a toda la frontera de Pareto según el siguiente gráfico:



Esta forma de aplicar el procedimiento no es la mejor, ya que con variaciones de lambda de 0,1 no se puede asegurar que se obtengan todos los puntos extremos o vértices de la frontera. La programación paramétrica es la herramienta idónea para ello.

VI.2.2. Método de las ε-restricciones [Marglin, 1967]

Consiste en optimizar uno de los objetivos e incorporar el resto como restricciones paramétricas, resolviendo el problema resultante para los distintos valores del parámetro.

$$\left.\begin{array}{l} \max z_l(x) \\ x \in F \\ z_k(x) \geq \varepsilon_k \quad k = 1,\ldots,l-1,l+1,\ldots,p \end{array}\right\} P_l(\varepsilon)$$

El fundamento de este procedimiento se recoge en los dos resultados siguientes:

Teorema: Si la solución del problema $P_l(\varepsilon)$ es única, entonces es una solución eficiente.

Teorema: Si x^* es eficiente, $\forall l \; \exists \varepsilon_k$ tales que x^* es solución óptima de $P_l(\varepsilon)$.

Por lo tanto, de estos dos resultados se infiere que variando el lado derecho siempre obtenemos soluciones eficientes, y que haciéndolo variar adecuadamente debemos obtener todas las soluciones eficientes. Sin embargo, su aplicación práctica suele ser relativamente compleja.

Apliquemos este procedimiento al ejemplo 1 que venimos tratando. Para ello seleccionamos un criterio que quedará como objetivo, y en nuestro caso, otro que pasará a las restricciones. Por ejemplo, mantengamos como objetivo el Coste, y pasemos a las restricciones la Calidad.

Para manejar coherentemente el parámetro del lado derecho de la restricción, normalizaremos este lado, de modo que variando E entre 0 y 1 obtengamos todos los valores de la Calidad dentro de la Frontera de Pareto. Esto podemos hacerlo multiplicando por E la diferencia entre lo mejor y lo peor de este criterio, y sumando lo peor. Es decir, si hacemos (100-20)*E + 20, haremos corresponder E=0 con la peor calidad, y E=1 con la mejor posible.

El objetivo del modelo será minimizar el coste, y añadiremos en las restricciones la restricción paramétrica de la calidad:

$$
\begin{aligned}
\min \quad & Coste = 200X + 100Y \\
& X \geq 20 \\
& X + Y \leq 100 \\
& X + Y \geq 50 \\
& Calidad : X \geq (100-20)\varepsilon + 20 \\
& X, Y \geq 0
\end{aligned}
$$

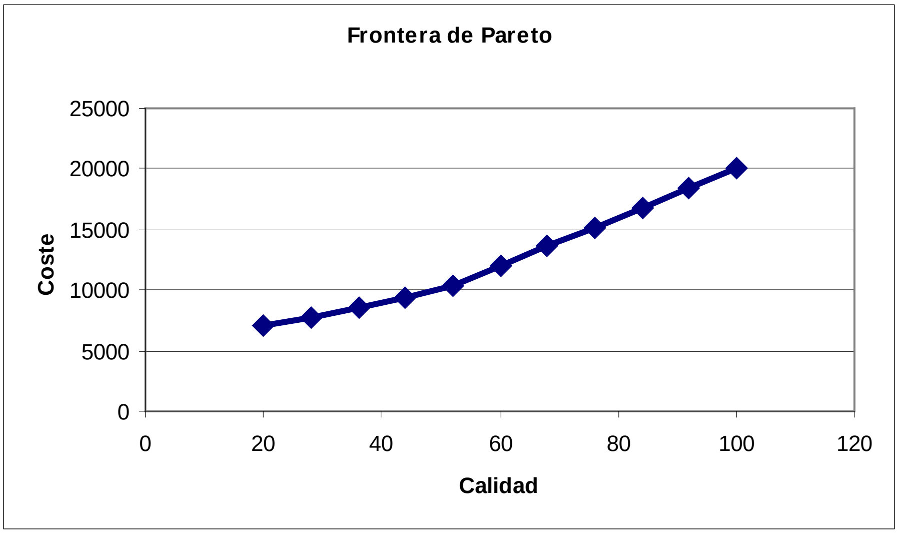
A continuación, se ejecuta secuencialmente el modelo haciendo variar E de 0 a 1, y tomando nota de las soluciones obtenidas en cada paso, obteniendo la siguiente tabla:

Ejemplo 1_eRestricciones

	A	B	C	D	E	F	G
1	X	Y					
2	100	0					
3	Calidad	X	100				
4	Coste	200*X+100Y	20000				
5	E (parámetro restricción)		1				
6							
7	Restricciones						
8	Mínimo 20%	100	20				
9	suma <=100	100	100				
10	suma >=50	100	50				
11	Calidad >=E	100	100	(100-20)*C5+20			
12							
13	Tabla resultados						
14	E	Calidad	Coste	X	Y		
15	0	20	7000	20	30		
16	0,1	28	7800	28	22		
17	0,2	36	8600	36	14		
18	0,3	44	9400	44	6		
19	0,4	52	10400	52	0		
20	0,5	60	12000	60	0		
21	0,6	68	13600	68	0		
22	0,7	76	15200	76	0		
23	0,8	84	16800	84	0		
24	0,9	92	18400	92	0		
25	1	100	20000	100	0		

Hoja1 / Hoja2 / Hoja3

cuyo gráfico de frontera de Pareto es:



Obsérvese que este procedimiento no plantea problema de agregación, problemas de unidades, ni de sentido de la optimización. Además, la frontera sale por puntos distintos, no por unión de otros, ya que estamos modificando la región factible y sus vértices. Sin embargo, con varios criterios en las restricciones es algo más difícil de manejar que las ponderaciones. A pesar de eso, este método resulta mejor, en general, que el de ponderaciones.

VI.3. Métodos de decisión multicriterio continuos

VI.3.1. Programación compromiso

En la programación compromiso el fin último es seleccionar un punto del conjunto eficiente y, para ello, naturalmente, ha de incluir las preferencias del decisor. Las técnicas que se presentan fueron desarrolladas inicialmente en los trabajos de Yu (1973) y Zeleny (1973, 1974).

Se define el *punto* o *alternativa ideal*, como un punto del espacio de objetivos que recoge los valores óptimos de los objetivos individualmente tratados y se denota por

$z^* = (z_1^*, ..., z_i^*, ..., z_p^*)$, siendo $z_i^* = \max_{x \in F} z_i(x)$ (en caso de maximizar el objetivo, en otro caso será el mínimo), denominándose este valor individual *punto ancla*.

Se define la *solución óptima* o *mejor solución compromiso*, como la solución eficiente más próxima al punto ideal (axioma de [Zeleny, 1973]).

La cuestión ahora es definir una distancia que ha de ser minimizada sobre el conjunto eficiente. Para ello se define el grado de proximidad del objetivo i-ésimo normalizado (ver sección VI.2.1) como $d_i(x) = \dfrac{z_i^* - z_i(x)}{z_i^* - z_{*i}}$, siendo z_{*i} el anti-ideal del objetivo, es decir, el peor valor posible para el objetivo sobre el conjunto eficiente.

Estos grados de proximidad son agregados para definir una métrica:

$$\min_{x \in F} L_\pi = \left[\sum_{i=1}^{p} w_i^\pi \left(\frac{z_i^* - z_i(x)}{z_i^* - z_{*i}} \right)^\pi \right]^{1/\pi}$$

donde el conjunto de los pesos w_i suponen una ponderación preferencial de los criterios (ponderación subjetiva dada por el decisor, este sistema de ponderación se puede intentar obtener del decisor por varios métodos descritos en la literatura, como son el método de ordenación, el método de Saaty que puede verse en la sección VI.3.3).

Para los distintos valores de π se obtienen diferentes métricas. Así $\pi = \infty$, denota la distancia de Tchebychev o lo que es lo mismo minimizar la máxima distancia (así se obtiene una solución equilibrada, en que todos los grados de proximidad individuales están acotados, siendo a su vez un problema lineal).

Para el caso $\pi = 1$, se trata de una agregación lineal de distancias ponderadas, que también mantiene el carácter lineal del problema.

En general, para diferentes valores de π se obtiene diferente solución, de ahí que se haya definido el *conjunto compromiso*, como el conjunto de soluciones que se obtienen al variar π. Para el caso de dos objetivos, es usual dar el conjunto definido por el segmento $[L_1, L_\infty]$ (es decir, el segmento que une las soluciones obtenidas con las dos métricas comentadas anteriormente), ya que bajo ciertas condiciones es cierto que el conjunto compromiso coincide con este segmento. En nuestro caso, trabajaremos con el caso lineal y Tchebychev.

Apliquemos el método al ejemplo que venimos tratando.

En primer lugar, de la matriz de pagos obtenemos los puntos ancla que nos dan el siguiente ideal $z^*= (100,7000)$, y el anti-ideal $z_* = (20,20000)$.

A continuación normalizamos los criterios:

$$\begin{aligned}
&Calidad: \; d_1(X,Y) = \frac{100-X}{100-20} = \frac{100-X}{80}\\
&Coste: \; d_2(X,Y) = \frac{7000-(200X+100Y)}{7000-20000} = \frac{200X+100Y-7000}{13000}
\end{aligned}$$

Ahora tenemos que dar la ponderación subjetiva de los criterios. Ésta es la parte más complicada, ya que a veces es difícil decir cuánto más importante es un criterio respecto al otro. Por otra parte, se suelen dar pesos cuya suma sea 1, por normalizar. De momento, supongamos que consideramos que la calidad es 3 veces más importante que el coste, en cuyo caso el peso de la calidad sería 0,75, y el del coste 0,25.

El problema a resolver según la métrica L_1 sería:

$$\begin{aligned}
\min \quad & 0,75\frac{100-X}{80} + 0,25\frac{200X+100Y-7000}{13000}\\
s.a. \quad & X \geq 20\\
& X+Y \leq 100\\
& X+Y \geq 50\\
& X,Y \geq 0
\end{aligned}$$

El problema para la métrica L_∞ sería:

$$\begin{aligned}
\min & \left[\max\left(0,75\frac{100-X}{80}; \quad 0,25\frac{200X+100Y-7000}{13000}\right)\right]\\
s.a. \quad & X \geq 20\\
& X+Y \leq 100\\
& X+Y \geq 50\\
& X,Y \geq 0
\end{aligned}$$

que linealmente se formula de la siguiente forma:

$$
\begin{aligned}
\min \quad & D \\
s.a. \quad & X \geq 20 \\
& X + Y \leq 100 \\
& X + Y \geq 50 \\
& 0,75\frac{100 - X}{80} \leq D \\
& 0,25\frac{200X + 100Y - 7000}{13000} \leq D \\
& X, Y \geq 0
\end{aligned}
$$

Al resolver este modelo puede observarse que la solución X=100 e Y=0, es claramente sesgada hacia el criterio de Calidad ya que es al que hemos dado más peso. En general, cuando hay dos criterios esto es lo que ocurre, que la solución es la mejor para el criterio de mayor importancia. Habiendo más, puede estar más equilibrado pero por lo general alguno de los criterios aparece muy desviado, y por eso se utiliza mucho la métrica infinito o Tchebychev, que al minimizar la máxima desviación, da soluciones más equilibradas.

Puede observarse que la solución es X=81,09 e Y=0, que no da la máxima calidad ni el peor coste, de hecho se puede ver que el desvío ponderado al ideal es el mismo para ambos objetivos. Es una solución más equlibrada.

Se pueden obtener diferentes soluciones si se modifican los pesos o ponderaciones subjetivas de los criterios. En general este método no es para hacer variar todos los pesos posibles, pero sí se debe si es que se puede, interactuar con el decisor hasta que considere que el modelo refleja sus preferencias. Esta interacción es habitual ya que no es fácil captar numéricamente las preferencias subjetivas, con lo que el ajuste de los parámetros del modelo es necesario, e incluso en ocasiones ayuda a saber qué grado de preferencia se tiene realmente por los distintos criterios.

VI.3.2. Métodos satisfacientes: programación por metas

Los métodos satisfacientes se basan en la denominada lógica satisfaciente enunciada por Simon en 1955:

"en los contextos actuales de decisión (información incompleta, recursos limitados, conflictos de intereses,...) el decisor más que optimizar unas funciones objetivo intenta que una serie de metas se aproximen lo más posible a unos niveles de aspiración prefijados."

De ahí que se desarrollara, basándose en esta filosofía, la denominada *programación por metas* ([Charnes y Cooper, 1961], [Lee, 1972] e [Ignizio, 1976]).

La idea de este método es que si un atributo tiene por expresión matemática $z_i(x)$ y para este atributo se tiene un *nivel de aspiración*, es decir, un nivel aceptable de logro para un atributo, $\hat{z}_i$, entonces una *restricción meta* sería una restricción de la forma $z_i(x) \geq \hat{z}_i$ (para el caso de desear "al menos" ese valor).

Sin embargo, dado que son niveles de aspiración, una restricción meta, en general, no puede ser incluida como una restricción rígida en el problema, ya que soluciones que no logren ese nivel de aspiración también son soluciones admisibles. Así, una restricción meta se formula utilizando *variables de desviación*: $z_i(x) + n_i - p_i = \hat{z}_i$. De estas variables, una de ellas es una variable no deseada: si la meta es "al menos" un valor, la variable no deseada será n_i, y si es del tipo "a lo sumo", la variable de desviación no deseada será p_i. En ocasiones, hay metas de ser igual a cierto nivel de aspiración, en cuyo caso ambas variables son no deseadas, pero no es habitual.

Así el problema a resolver, suponiendo que todos los objetivos fueran de tipo maximizar sería:

$$\min_{x \in F \cap \text{restricciones meta}} \sum_{i=1}^{p} n_i$$

Este procedimiento se puede aplicar también a las restricciones del problema para relajarlas, admitiendo soluciones "cercanas" a la región factible (el mismo concepto se ha aplicado en lógica difusa para resolver problemas de programación lineal).

La programación por metas aplicada en la realidad muestra que se están obteniendo valiosos resultados, siendo una de las técnicas de decisión multicriterio que está proporcionando mejores resultados. Es especialmente útil cuando hay muchos criterios.

Sin embargo, casi nunca se utiliza en su versión básica. Algunas variantes de este método se han desarrollado con gran éxito y se exponen a continuación.

VI.3.2.1. Variantes de la programación por metas

VI.3.2.1.1. Programación por metas ponderadas

$$\begin{aligned}
&\min \sum_{i=1}^{p} (\alpha_i n_i + \beta_i p_i) \\
&z_i(x) + n_i - p_i = \hat{z}_i \quad i = 1, \ldots, p \\
&x \in F \\
&n_i \geq 0,\ p_i \geq 0 \quad i = 1, \ldots, p
\end{aligned}$$

La idea básica es ponderar las variables de desviación, ya que pueden tener diferente relevancia las metas o diferentes unidades si no han sido previamente normalizados los criterios.

Los distintos criterios, en general, vendrán dados en distintas unidades, así que la función objetivo estaría agregando valores de distintas unidades. Una primera medida, muy deseable y ampliamente utilizada, es ponderar las desviaciones dividiendo por el nivel de aspiración, con lo que serían desviaciones porcentuales que no tienen unidades y que corrigen el efecto de las distintas magnitudes de éstas.

Por otra parte, si sólo se ponderan dividiendo por el nivel de aspiración, implícitamente lo que se está haciendo es dar una misma importancia a todos los criterios. Si no es ése el caso, se deben multiplicar por pesos que muestren la relevancia dada por el decisor a cada meta.

A continuación aplicamos esta técnica al ejemplo con el que trabajamos.

En primer lugar, hemos de dar un nivel de aspiración para la calidad y uno para el coste. Por ejemplo, supongamos que queremos una calidad del 75% y un coste de 10000€. Y supongamos que no damos preferencias subjetivas más allá de las marcadas por los propios niveles de aspiración. En ese caso el modelo sería:

$$
\begin{aligned}
\min \quad & \frac{n_1}{75} + \frac{p_2}{10000} \\
s.a. \quad & X \geq 20 \\
& X + Y \leq 100 \\
& X + Y \geq 50 \\
& X + n_1 - p_1 = 75 \\
& 200X + 100Y + n_2 - p_2 = 10000 \\
& X, Y, n_1, n_2, p_1, p_2 \geq 0
\end{aligned}
$$

La solución es 50 de Calidad y 10000 de Coste, es decir, se cumple la meta del coste a costa de incumplir la de la calidad. Esto muestra que no es posible cumplir las dos metas a la vez, pero además, este método en muchas ocasiones da una solución de este tipo, en la que unas metas se cumplen por completo y otras se incumplen totalmente (normalmente por un tema de unidades). Por ello, se presenta el siguiente método, que al igual que en el caso de programación compromiso, da soluciones más equlibradas.

VI.3.2.1.2. Programación por metas MINIMAX o Tchebychev

En este caso se busca una solución "equilibrada", de modo que ninguna de las metas se desvíe en exceso de su nivel de aspiración. Para ello se minimiza la máxima distancia a este nivel, es decir, el problema se plantearía en estos términos

$$
\begin{aligned}
& \min D \\
& \alpha_i n_i + \beta_i p_i \leq D \quad i = 1, \ldots, p \\
& z_i(x) + n_i - p_i = \hat{z}_i \quad i = 1, \ldots, p \\
& x \in F \\
& n_i \geq 0, \; p_i \geq 0 \quad i = 1, \ldots, p
\end{aligned}
$$

Para el ejemplo que estamos viendo y con los mismos datos de la sección anterior, el modelo sería:

$$\begin{aligned}
\min \quad & D \\
s.a. \quad & X \geq 20 \\
& X + Y \leq 100 \\
& X + Y \geq 50 \\
& X + n_1 - p_1 = 75 \\
& 200X + 100Y + n_2 - p_2 = 10000 \\
& n_1 / 75 \leq D \\
& p_2 / 10000 \leq D \\
& X, Y, n_1, n_2, p_1, p_2 \geq 0
\end{aligned}$$

y la solución obtenida es 60 de Calidad y 12000 de Coste, una solución equilibrada, que se desvía un 20% de los niveles de aspiración de ambos criterios.

VI.3.2.1.3. Programación por metas lexicográficas

Este procedimiento establece niveles de prioridad en las metas y resuelve secuencialmente el problema para cada nivel de prioridad, manteniendo los valores previamente obtenidos para metas con mayor nivel de prioridad. Así, si en un problema se formulan k metas es posible que se consideren agrupadas en $h\,(\leq k)$ niveles de prioridad. Por ejemplo, si el problema incluye 6 metas, que denominaremos $g_i(n, p), i = 1,\ldots,6$, agrupadas en 3 niveles de prioridad de la forma las más importantes las dos primeras, después la tercera, y por último las 3 últimas, el problema se representa como

$$Lex \min a = [g_1(n, p) + g_2(n, p), g_3(n, p), g_4(n, p) + g_5(n, p) + g_6(n, p)]$$

Para resolver el problema, se resolvería primero con las dos primeras metas. Obtenidos los valores de las variables de desviación no deseadas para éstas, se fijan y se resuelve el problema para la tercera meta (segundo nivel de prioridad). Una vez resuelto, se fija el valor de la variable no deseada también para esta meta, y se resuelve el tercer nivel incluyendo las tres últimas metas.

Es un procedimiento que requiere de varias ejecuciones secuenciales del modelo de metas modificando el objetivo y fijando las variables no deseadas de los niveles anteriores. Es útil cuando hay muchas metas y claramente diferenciadas (por ejemplo metas económicas, metas de clientes, etc.) teniendo una prioridad clara unas sobre otras.

En nuestro ejemplo no tiene mucho sentido presentarlo. Por ejemplo, si damos prioridad a la calidad, y un nivel de aspiración de 75%, el primer nivel será aproximarse a este nivel. Dado que es posible, la solución del primer nivel fijará la calida en 75% o más, y el segundo nivel, el del coste, nos dará la solución de menos coste para esa calidad, es decir, 75% de Calidad y 15000 de Coste.

VI.3.2.2. Temas críticos en programación por metas

A pesar de que la programación por metas resulta un potente instrumento de gran aplicabilidad e indudable éxito en la realidad, diferentes autores han apuntado algunas aparentes debilidades de los modelos de metas. Estas consideraciones han llevado a denominar *temas críticos en programación por metas* a anomalías aparentemente causadas por debilidades lógicas de la programación por metas, aunque en realidad se deben al uso insatisfactorio del enfoque. La mayoría se presentan en las metas lexicográficas. Algunas de ellas se exponen a continuación.

Posible equivalencia de soluciones entre modelos de programación por metas y modelos de optimización de un solo criterio.

Es cierto, que, por ejemplo, en las metas lexicográficas si se fijan unos niveles muy pesimistas en los primeros niveles (fáciles de alcanzar) y muy optimistas en los últimos (difíciles de alcanzar), al resolverlo para los primeros niveles se obtendrá un valor 0 de las desviaciones, y puede llegar un momento en que para una meta no haya óptimos alternativos y, por lo tanto, fijar el valor de su desviación es ya dar una solución haciendo redundantes las metas de prioridades más bajas; de este modo, sería equivalente a fijar las metas de los niveles anteriores en su nivel de aspiración y optimizar la meta para la que no hay óptimos alternativos. Sin embargo, esta equivalencia se da por un mal planteamiento de los niveles de aspiración. La equivalencia puede darse también en metas ponderadas, por ejemplo, si se plantean unos niveles de aspiración muy pesimistas para las metas, excepto para una en que resulte muy optimista, prácticamente inalcanzable. En tal caso, es cierto que el problema puede ser equivalente a optimizar ese atributo incluyendo las demás metas como restricciones. En cualquier caso, estos problemas se derivan de la formulación del modelo y no a debilidades lógicas de la metodología.

Conclusiones equivocadas por agregar metas lexicográficas en una función objetivo

Cuando se plantea un modelo de programación por metas lexicográficas, suele haber una tendencia (casi natural) a intentar plantear un modelo equivalente que se resuelva en una sola iteración, sin tener que resolver un problema para cada nivel de prioridad. Así es común ver que se plantea una función objetivo como en el procedimiento de metas ponderadas donde se multiplica cada variable de desviación por un peso asignado a su nivel de prioridad (así para el primer nivel el peso es mucho mayor que para el segundo, etc.). En general, este planteamiento, además de ser lógicamente incorrecto, puede llevar a obtener soluciones que difieran mucho de las que se obtendrían con el modelo lexicográfico, y por lo tanto, si se intenta razonar sobre ellas como si fueran soluciones del modelo lexicográfico, a conclusiones claramente incorrectas.

Problemas derivados de la omisión de una variable de desviación

Otros de los problemas que han surgido en la aplicación, o mejor dicho, en la incorrecta aplicación de la programación por metas, han sido los derivados por no incluir una variable de desviación. En ocasiones, se ha prescindido de la variable de desviación que no es considerada, por ejemplo, si se desea que un atributo sea superior a un nivel de aspiración, la variable no deseada es la de holgura al nivel de aspiración, n, sin embargo de la otra variable, la de exceso p, no se considera en el objetivo. Obsérvese que prescindir de ella implica que el atributo no pueda tomar valores superiores al nivel de aspiración, sólo inferiores (no deseados) o igual, lo que restringe el espacio de soluciones, eliminando del espacio de objetivos toda la región en que el atributo sea superior. Por lo tanto, la variable de desviación que no es la no deseada, también debe aparecer en el modelo o formular la meta como una desigualdad y no una igualdad. Ante el riesgo que puede presentar esta última formulación, se aconseja mantener ambas variables con la formulación lógica clásica de las metas, además de que, en general, no incluirlas no conlleva una mejora computacional (los métodos de resolución suelen añadir una variable en las desigualdades como primera medida).

Problemas derivados de la inclusión de metas con dos lados innecesariamente

Las metas con dos lados no deseados son menos habituales que las de un solo lado, ya que es raro que un decisor desee el logro exacto de una meta más que superar, o no hacerlo, un determinado nivel. Incluir entre las no deseadas una desviación que no lo es, puede llevar a obtener soluciones subóptimas. Es claro, que hay que tener un especial cuidado en identificar cuáles son las variables de desviación no deseadas.

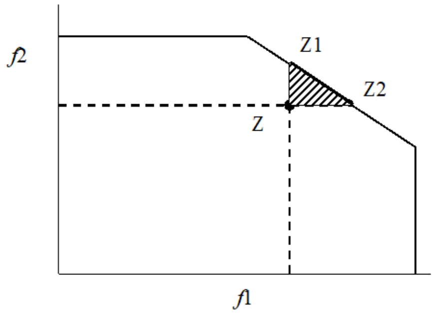
Incompatibilidad entre ordenaciones lexicográficas y la existencia de una función de utilidad

Una de las críticas más fuertes al modelo lexicográfico ha sido su incompatibilidad con una función de utilidad del decisor. Esta incompatibilidad demostrada, que aquí no se niega, se deriva de la incompatibilidad con uno de los axiomas que se atribuye a una función de utilidad. Los axiomas que se supone que deben ser cumplidos por una función de utilidad son comparabilidad, reflexividad, transitividad y continuidad. De ellos, el que resulta incompatible con las ordenaciones lexicográficas es el de la continuidad. Según este axioma, los conjuntos de nivel de la función de utilidad son superficies continuas, es decir, dada una combinación de dos elecciones posibles de un decisor, siempre se podrá reducir la cantidad de una elección encontrando un incremento de la otra elección que compense exactamente la reducción. Es fácil comprobar que en las metas lexicográficas los conjuntos de nivel o superficies de indiferencia están formadas por un único punto (ver estos conjuntos como intersección del conjunto mejor y el peor respecto a un punto), por lo tanto, no es posible obtener el mismo valor con dos puntos distintos, lo que supone que no se cumple el axioma. Sin embargo, la continuidad de las preferencias dista mucho de ser un hecho, ni tan siquiera una hipótesis corroborada de forma empírica. Es tan sólo una hipótesis necesaria para axiomatizar la teoría neoclásica del consumo y asegurar la existencia de un equilibrio competitivo. Es más, hay situaciones en las que suponer válido este axioma puede ser un disparate; por ejemplo, supóngase un problema de planificación forestal donde se tienen en cuenta varios criterios, por ejemplo, uno económico que fuera la producción de madera y otro que fuera un índice que mida el riesgo de colapso ecológico del bosque. En este caso el supuesto de continuidad diría que siempre existe un incremento en el volumen de madera producida que compense un incremento del riesgo de colapso del bosque, por alto que éste sea, lo que puede ser un auténtico disparate. Es obvio, que en este caso y en otros muchos, las preferencias no son continuas y por lo tanto no existe una función de utilidad para representarlas. Es decir, no es absolutamente necesario aceptar la continuidad de las preferencias para poder representar el problema de decisión.

Posible ineficiencia paretiana de la solución obtenida por metas, y forma de resolverlo

La programación por metas no está pensada para obtener soluciones eficientes sino para obtener soluciones satisfacientes. En este sentido, es posible que la solución de un modelo de programación por metas no sea eficiente. Para ello, lo que tiene que pasar es que haya óptimos alternativos del problema de metas ponderadas o del último nivel de metas lexicográficos (esta condición es necesaria aunque no es suficiente). Para algunos autores es un gran inconveniente, pero no ha de olvidarse que el objetivo de la programación por metas, como ya se ha dicho, no es obtener soluciones eficientes sino satisfacientes. Para comprobar si una solución obtenida mediante programación por metas es eficiente, se fijan los valores de las variables de desviación no deseadas obtenidos, y se maximizan las variables opuestas; si la solución no cambia, es que la solución es eficiente, mientras que si varía, la nueva solución será una solución eficiente con los mismos valores de las variables no deseadas. Existe un test propuesto por Hannan (1980) que permite a partir de una solución obtenida por metas, obtener todas las soluciones eficientes que tienen el mismo valor de las variables de desviación no deseadas. Este test consiste en aplicar la técnicas multiobjetivo a un problema que sea el original pero incluyendo los valores de las variables de desviación (tanto las deseadas como las no deseadas) obtenidas previamente mediante programación por metas. Así si una meta es que x_1

sea al menos 5, y al resolverlo por metas se ha obtenido $n_1 = 0, p_1 = 1$, se formularía una restricción para el test de Hannan que fuera: $x_1 \geq 5 - n_1^* + p_1^* = 5 - 0 + 1 = 6$.. Gráficamente, si Z es el punto obtenido por programación por metas, toda la zona rayada son puntos mejores, y Z1 y Z2 serían puntos eficientes mejores que Z y todas las combinaciones lineales convexas de ellos también:



De aquí se deriva un procedimiento para resolver modelos de programación por metas si se desea que además la solución sea eficiente:

1) Resolver el modelo de metas inicial. Si no hay óptimos alternativos la solución es eficiente, en otro caso ir al paso 2.

2) Si se desea una única solución eficiente ir a 3, si se desea explorar un conjunto de soluciones eficientes ir a 4.

3) Maximizar la suma de las variables desviacionales opuestas sin empeorar los valores alcanzados. La solución será eficiente

4) Transformar el modelo de metas en uno multiobjetivo con el procedimiento de Hannan para obtener todos los puntos eficientes del correspondiente modelo de metas.

El concepto de meta redundante y repercusión en las metas lexicográficas

El concepto de metas redundante puede surgir tanto en metas ponderadas como en lexicográficas, aunque con distinto sentido. Así en metas ponderadas puede surgir porque se plantee una meta inalcanzable y las demás de modo pesimista, con lo que el problema puede desviarse hacia la meta inalcanzable olvidando las demás. Sin embargo, donde es más relevante y claro este concepto es en las metas lexicográficas, ya que si en un determinado nivel de prioridad se llega a un único óptimo, todas las metas incluidas en niveles inferiores son un mero ornamento, ya que no pueden variar la solución. Se define como meta redundante aquella cuya omisión no influye en el resultado. La existencia de metas redundantes puede darse principalmente por tres causas:

- Una excesiva priorización de las metas, es decir, una agrupación en un número excesivamente elevado de niveles de prioridad
- Ser excesivamente optimista, fijando los niveles de aspiración muy próximos al ideal del atributo.
- Incluir muchas metas con dos lados, obligando a la minimización de ambas variables de desviación.

Cuando se da la redundancia es interesante saber en qué nivel de prioridad se ha dado, y si parece oportuno, replantear el problema. El replanteamiento puede hacerse redefiniendo los niveles de aspiración de las metas anteriores o viendo si se han incluido metas de dos lados innecesariamente, o agrupando en distintos niveles (por ejemplo, se puede plantear un último nivel con la última meta no redundante ponderada para darle más importancia y las restantes metas, las que resultaron redundantes, con ponderación menor).

VI.3.3. Obtención de preferencias mediante programación por metas

Dentro de los métodos de decisión multicriterio discretos, Saaty en 1977 introdujo el método AHP (procesos analíticos jerarquizados) que tuvo un gran impacto teórico y aplicado. En el caso que nos ocupa, nos centramos no en el método sino en la metodología para obtener preferencias entre criterios (ponderaciones) o preferencias entre los atributos que incluye Saaty en su método, aunque ya se había establecido antes: la comparación por pares.

Hay algunas formas más sencillas de obtener los pesos, como es pedirle al decisor que clasifique los criterios por orden de importancia, de modo que si hay n criterios, al más importante le asigne el valor 1 y al menos importante el valor n; a continuación, con el fin de que la suma de los pesos sea 1 se le asigna al criterio en posición j-ésima el peso $W_j = \frac{1/r_j}{\sum_{i=1}^{n} r_i}$ o el peso $W_j = \frac{n - r_j + 1}{\sum_{i=1}^{n} (n - r_i + 1)}$. El problema de este método es que no tiene en cuenta la diferencia de importancia entre criterios, es decir, cuánto más importante es uno que el siguientes.

Para evitarlo Saaty propone otro método que consiste en comparar los criterios por parejas, de modo que se asigne un 1 si ambos son de la misma importancia, un 3 si hay una moderada importancia de un criterio respecto a otro, un 7 si hay una demostrada importancia y un 9 si hay extrema importancia. Con ello se forma una matriz cuadrada a_{ij} que valora la importancia del criterio *i* respecto al *j*. Si la importancia de un criterio respecto a otro es a_{ij}, a la inversa será $1/a_{ij}$. Obtenida esta matriz, se busca un vector de pesos para los criterios que sea solución del sistema $\frac{W_i}{W_j} = a_{ij}$, o lo que es lo mismo $W_i = a_{ij} W_j$. Lamentablemente, ese sistema no suele tener solución dadas las normales inconsistencias del decisor y hay que buscar los que más se aproximen, por ejemplo, con la programación por metas vista en la sección anterior (planteando las restricciones meta $W_i - a_{ij} W_j + n_{ij} - p_{ij} = 0$ y penalizando ambas variables de desviación).

Aunque no puede hablarse de cuál es el mejor método para estimar pesos preferenciales, siempre que la situación lo permita parecen tener más solidez los métodos propuestos por Saaty que los anteriores.

Supóngase que se quiere elegir el trazado de una autopista, supóngase que el centro decisor ha emitido sus juicios de valor obteniendo la siguiente matriz de comparaciones.

	Coste	Impacto Medioambiental	Tiempo ejecución
Coste	1	2	5
Impacto Medioambiental	1/2	1	3
Tiempo ejecución	1/5	1/3	1

A partir de esta matriz se buscan los pesos preferenciales de los criterios resolviendo el siguiente problema donde se han introducido las variables de desviación para resolverlo (ver la sección VI.3.2):

$$\begin{array}{ll}
\min \sum_{i=1}^{3}(n_i + p_i) & \\
w_1 - 2w_2 + n_{12} - p_{12} = 0 & w_2 - 1/2w_1 + n_{21} - p_{21} = 0 \\
w_1 - 5w_3 + n_{13} - p_{13} = 0 & w_3 - 1/5w_1 + n_{31} - p_{31} = 0 \\
w_2 - 3w_3 + n_{23} - p_{23} = 0 & w_3 - 1/3w_2 + n_{32} - p_{32} = 0 \\
w_1 + w_2 + w_3 = 1 &
\end{array}$$

La solución obtenida es $(w_1, w_2, w_3) = (.588,.294,.118)$, siendo las variables de desviaciones $n_{12} = n_{13} = n_{21} = n_{31} = n_{32} = p_{12} = p_{13} = p_{21} = p_{23} = 0, n_{23} = 0.06, p_{32} = 0.02$.

Esta misma estrategia puede utilizarse para comparar alternativas (numéricas o no) y asignarles un valor numérico de preferencias.

VI.4. Referencias

Benayoun, R., Roy, B., Sussman, B. Electre: Une Méthode pour Guider le Choix en Présence de Vue Multiples. *Sema* (*Metra International*), Direction Scientifique. Note de travail, 49.

Charnes, A., Cooper, W.W (1961) *Management Models and Industrial Applications of Linear Programming*. John Wiley and Sons.

DeGroot, M.H. (1970) *Optimal Statistical Decisions*, McGraw Hill, New York

French, S. (1986) *Decision Theory: An introduction to the mathematics of rationality*, Ellis Horwood, Chichester.

Hannan, E.L. (1980) Nondominance in Goal Programming. *INFOR, Canadian Journal of Operational Research and Information Processing* **18**

Ignizio, J. P. (1976) *Goal Programming and Extensions.* Lexington Books.

Lee, S. M. (1972) *Goal Programming for Decision Analysis*. Auerbach Publishers.

Keeney, R.L. and Raiffa, H. (1976) *Decisions with Multiple Objectives: Preferences and Value Trade-Offs.* John Wiley and Sons.

Ríos, S., M.J. Ríos-Insúa, S. Ríos-Insúa. (1989) *Procesos de Decisión Multicriterio*. Eudema.

Romero, C. (1991) *Handbook of Critical Issues in Goal Programming.* Pergamon Press.

Romero, C. (1993) *Teoría de la Decisión Multicriterio: Conceptos, Técnicas y Aplicaciones.* Alianza Universidad.

Saaty, T. L. (1977) A Scaling Method for Priorities in Hierarchical Structures. *Journal of Mathematical Psychology,* **15,** 234-281.

Saaty, T.L. (1980) *The Analytic Hierarchy Process.* McGraw Hill

Simon, H. A. (1955) A Behavioral Model of Rational Choice. *Quarterly Journal of Economics,* **69,** 99-118.

Yu, P.L. (1973) A Class of Solutions for Group Decision Problems. *Management Science,* **19,** 936-946.

Yu, P.L. (1985) *Multiple criteria decision making: Concepts, techniques and extensions*. Plenum, New York.

Zadeh, L.A. (1963) Optimality and Non-Scalar-Valued Performance Criteria. *IEEE Transactions on Automatic Control,* **8**, 59-60.

Zeleny, M. (1973) Compromise Programming, in *Multiple Criteria Decision Making* (Cochrane, J.L. and Zeleny, M. ed.). University of South Carolina Press, 262-301

Zeleny, M. (1974) A Concept of Compromise Solutions and the Method of the Displaced Ideal. *Computers and Operations Research*, **1,** 479-496.

Zeleny, M. (1982) *Multiple criteria decision making*. McGraw-Hill, New York.

VI.5. Ejemplos y Problemas

PROBLEMA 1

Se presenta una planificación eléctrica muy simplificada de un país. Se supone una demanda eléctrica anual de 150000 GWh, que puede ser satisfecha mediante el empleo de las siguientes fuentes de energía: carbón, energía nuclear, gas natural, energía eólica, biomasa procedente de cultivos energéticos, y energía hidráulica. Se desea que tanto el coste de generación, así como las emisiones de CO2 producidas sean mínimas.

Las posibilidades de utilización de cada una de las opciones están determinadas por el potencial existente, o por condicionantes de otro tipo, que se detallan a continuación:

- la energía nuclear está sujeta a una moratoria que impide la expansión de la potencia instalada, con una producción anual de 53000 GWh.
- respecto al gas natural, la producción potencial sería de 12.845 GWh.
- el potencial de la energía eólica ha sido evaluado en unos 2.000 MW, que, para una utilización media de 2.500 h, produciría un máximo de 5.000 GWh.
- los recursos de biomasa para producción energética han sido estimados en 22 Mtep, que equivaldrían a unos 40000 GWh de producción eléctrica.
- por último, se supone un potencial hidráulico de 31.755 GWh.

Toda la demanda debe ser satisfecha con estas opciones, sin tener en cuenta otras como fuel-oil, energía solar… La tabla 1 muestra los costes de generación y emisiones de CO2 de las opciones consideradas, referidos al kWh generado 1GWh= 10^6kWh

Tabla 1. Costes y emisiones de CO_2 de la generación eléctrica

Fuentes de energía	**Coste (c€/kWh)**	**Emisiones de CO_2 (g/kWh)**
Carbón	5,85	1015
Nuclear	8,24	-
Gas natural	5,00	401
Eólica	9,00	-
Biomasa	12,00	-
Hidráulica	6,00	-

Analizar el problema, obteniendo la matriz de pagos, la frontera de Pareto, la mejor solución compromiso suponiendo igualdad de preferencias, y la solución por metas, planteando como niveles de aspiración un 125% de los respectivos ideales.

PROBLEMA 2

Un particular desea construir un chalet. A efectos de planificación ha considerado las siguientes actividades, con sus duraciones, y mano de obra por actividad:

	Compra de terrenos	12 días	2 trabajadores
	Planos del edificio y parcela	6 días	4 trabajadores
	Autorización municipal para construir	5 días	1 trabajador
	Financiación de la construcción	7 días	2 trabajadores
	Cimientos del edificio	9 días	10 trabajadores
	Cavar zanjas alcantarillado de la parcela	8 días	5 trabajadores
	Colocar tuberías de alcantarillado	12 días	5 trabajadores
	Rellenar zanjas	4 días	5 trabajadores
	Construcción de la estructura de la casa	20 días	10 trabajadores
	Tabicado	5 días	5 trabajadores
	Instalación eléctrica	6 días	5 trabajadores
	Conexión desagües del edificio de alcantarillado	2 días	5 trabajadores
	Instalación fontanería	5 días	5 trabajadores
	Solado de la casa	4 días	10 trabajadores
	Pintura	3 días	5 trabajadores
	Pavimentación exteriores y ajardinamiento	27 días	5 trabajadores
	Cédula de habitabilidad	2 días	1 trabajador

Entre estas actividades existen las siguientes relaciones de precedencia:

A y B preceden a C y D	I precede a J
D y C preceden a E y F	J precede a K y a L
E precede a I	K y M preceden a N y O
F precede a G	L precede a M
G precede a H	N y O preceden a Q

H precede a L y a P	

Supóngase ahora que los tiempos señalados en la tabla son fijos, pero que podrían acelerarse las ejecuciones de las actividades hasta un máximo de dos días, excepto L y Q que sólo admiten un máximo de un día. Los costes de acelerar un día el tiempo de ejecución de una actividad son los siguientes:

A, C, B y G	6000€	K, O, H	12000€
D, F, E	8000€		
I, J, L, Q	10000€	M, N, P	5000€

Analizar el problema, obteniendo la matriz de pagos, la frontera de Pareto, la mejor solución compromiso suponiendo igualdad de preferencias, y la solución por metas, planteando como niveles de aspiración 60 días de tiempo, y un coste de 50000€.

Nota: El siguiente modelo sirve para representar las precedencias y las reducciones de tiempo en la duración de las tareas, y obtener el tiempo del proyecto y el coste.

DI_i : duración inicial tarea i
R_i : máxima reducción de duración tarea i
C_i : coste unitario reducción tarea i
t_i : variable instante inicio tarea i
r_i : variable reducción aplicada a tarea i
tt : tiempo total de ejecución del proyecto

$$\begin{cases} t_i + DI_i - r_i \le t_j & \forall (i,j) \in P \\ t_i + DI_i - r_i \le tt & \forall i \in I \\ r_i \le R_i & \forall i \in I \\ t_i, r_i \ge 0 & \forall i \in I \\ Tiempo: \; tt & \\ Coste: \; \sum_i C_i r_i & \end{cases}$$

VI.6. Resultados de los ejemplos y problemas

RESULTADO DEL PROBLEMA 1

El modelo que representa la región factible es el siguiente:

$$
\begin{aligned}
&X_C + X_H + X_B + X_E + X_G + X_N = 150000 \\
&X_H \leq 31755 \\
&X_B \leq 40000 \\
&X_E \leq 5000 \\
&X_G \leq 12845 \\
&X_N \leq 53000 \\
&X_i \geq 0 \quad \forall i
\end{aligned}
$$

Y las expresiones de los dos criterios (Coste y Emisiones), expresados en c€/kWh y g/kWh (para ello basta dividir por la producción total), son las siguientes:

$$
\begin{aligned}
&Coste: \quad \frac{5,85X_C + 8,24X_N + 5X_G + 9X_E + 12X_B + 6X_H}{150000} \\
&Emisiones: \quad \frac{1015X_C + 401X_G}{150000}
\end{aligned}
$$

siendo la matriz de pagos:

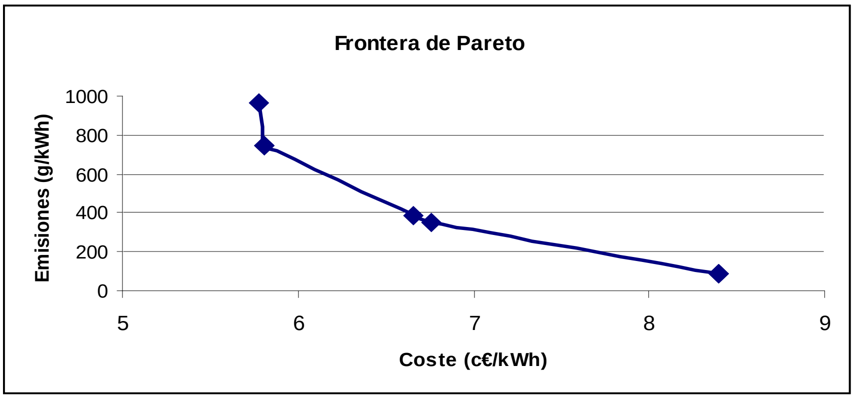
Matriz de pagos

	Coste	Emisiones
Coste	**5,777**	**962,421**
Emisiones	**8,398**	**84,413**

Por el procedimiento de ponderaciones la tabla para obtener la frontera de Pareto es:

Lambda	**Agregado**	**Coste**	**Emisiones**	**C**	**N**	**G**	**E**	**B**	**H**
0	0	**8,3984**	**84,4123**	**7400**	**53000**	**12845**	**5000**	**40000**	**31755**
0,1	0,1	**8,3984**	**84,4123**	**7400**	**53000**	**12845**	**5000**	**40000**	**31755**
0,2	0,2	**8,3984**	**84,4123**	**7400**	**53000**	**12845**	**5000**	**40000**	**31755**
0,3	0,3	**8,3984**	**84,4123**	**7400**	**53000**	**12845**	**5000**	**40000**	**31755**
0,4	0,3347	**6,7584**	**355,0789**	**47400**	**53000**	**12845**	**5000**	**0**	**31755**
0,5	0,3406	**6,6534**	**388,9123**	**52400**	**53000**	**12845**	**0**	**0**	**31755**
0,6	0,3094	**5,8089**	**747,5456**	**105400**	**0**	**12845**	**0**	**0**	**31755**
0,7	0,2351	**5,8089**	**747,5456**	**105400**	**0**	**12845**	**0**	**0**	**31755**
0,8	0,1608	**5,8089**	**747,5456**	**105400**	**0**	**12845**	**0**	**0**	**31755**
0,9	0,0865	**5,8089**	**747,5456**	**105400**	**0**	**12845**	**0**	**0**	**31755**
1	0	**5,7772**	**962,4211**	**137155**	**0**	**12845**	**0**	**0**	**0**

y el gráfico obtenido:

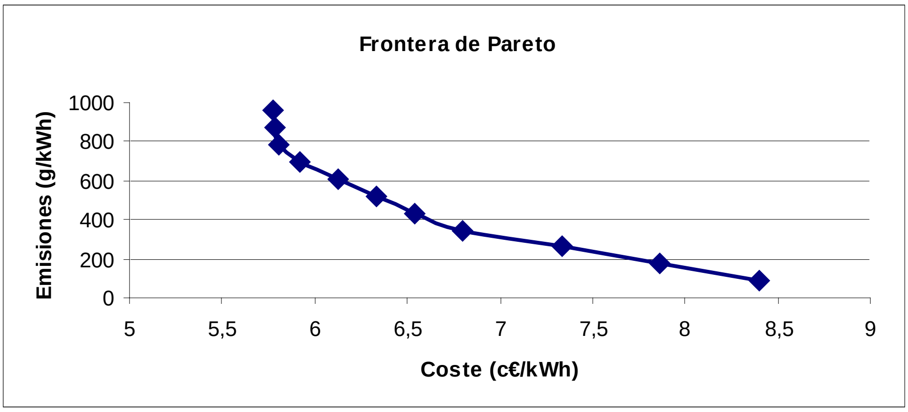


La tabla mediante las e-restricciones es:

Tabla resultados

e	Coste	Emisiones	C	N	G	E	B	H
0	8,3984	84,4123	7400	53000	12845	5000	40000	31755
0,1	7,8664	172,2132	20375,5	53000	12845	5000	27024,5	31755
0,2	7,3344	260,0141	333501	53000	12845	5000	14049	31755
0,3	6,8024	347,8149	46326,5	53000	12845	5000	1073,5	31755
0,4	6,5434	435,6158	59302	46098	12845	0	0	31755
0,5	6,3367	523,4167	72277,5	33122,5	12845	0	0	31755
0,6	6,1299	611,2176	85253	20147	12845	0	0	31755
0,7	5,9232	699,0185	98228,5	7171,5	12845	0	0	31755
0,8	5,8031	786,8193	111204	0	12845	0	0	25951
0,9	5,7902	874,6202	124179,5	0	12845	0	0	12975,5
1	5,7772	962,4211	137155	0	12845	0	0	0

y el gráfico obtenido:



Para buscar la mejor solución compromiso, primero lo hacemos con la métrica L1, normalizando la distancia al ideal de los criterios, y agregado de forma lineal en un objetivo en que multiplicamos por 0,5 cada una de esas distancias (dado que dice que se considere igualdad de preferencias). Con ello, obtenemos la solución:

C	N	G	E	B	H
52400	**53000**	**12845**	**0**	**0**	**31755**

con los siguientes valores de ambos criterios:

Coste	**6,65343333**
Emisiones	**388,9123**

Respecto a utilizar la distancia Tchebychev, en la siguiente hoja del archivo se encuentra el modelo y la solución, observándose que la solución es la misma.

Por último, aplicamos la programación por metas con un 125% respecto al ideal para ambos criterios, lo que supone un nivel de aspiración para el coste de 7,22 c€/kWh y para las emisiones de 106 g/kWh. Suponemos preferencias iguales como en el caso compromiso. Por otra parte, podemos hacerlo mediante metas ponderadas o Tchebychev. Se tiene la solución por metas ponderadas:

C	N	G	E	B	H
10590,3005	**53000**	**12845**	**5000**	**36809,6995**	**31755**

cuyo coste y emisiones son:

Coste	**8,26763101**
Emisiones	**106**

siendo las desviaciones

	n	p
Desviaciones coste	**0**	1,04763102
Desviaciones emisiones	**0**	0

Como puede observarse, se cumple la meta de emisiones, y para esa meta se obtiene el coste correspondiente.

Sin embargo, si aplicamos las metas Tchebychev la solución es:

C	N	G	E	B	H
12677,6333	**53000**	**12845**	**5000**	**34722,3667**	**31755**

con coste y emisiones

Coste	**8,18205037**
Emisiones	**120,124285**

y desviaciones

	n	p
Desviaciones coste	**0**	0,96205037
Desviaciones emisiones	**0**	14,1242852

que suponen para ambos criterios un incumplimiento respecto al nivel de aspiración de un 13%.

RESULTADO DEL PROBLEMA 2

El modelo que se ha utilizado es el descrito en la nota del enunciado:

DI_i : duración inicial tarea i
R_i : máxima reducción de duración tarea i
C_i : coste unitario reducción tarea i
t_i : variable instante inicio tarea i
r_i : variable reducción aplicada a tarea i
tt : tiempo total de ejecución del proyecto

$$\begin{cases} t_i + DI_i - r_i \le t_j & \forall (i,j) \in P \\ t_i + DI_i - r_i \le tt & \forall i \in I \\ r_i \le R_i & \forall i \in I \\ t_i, r_i \ge 0 & \forall i \in I \\ Tiempo: \ tt & \\ Coste: \ \sum_i C_i r_i & \end{cases}$$

En primer lugar se obtiene la matriz de pagos (el coste se introduce en miles de euros para que sean unidades semejantes al tiempo). Hay que tener cuidado en los resultados del anti-ideal, puesto que pueden ser peores de los realmente alcanzablees:

Matriz de pagos		
	Tiempo	Coste (k€)
Tiempo	**58**	**116**
Coste (k€)	**70**	**0**

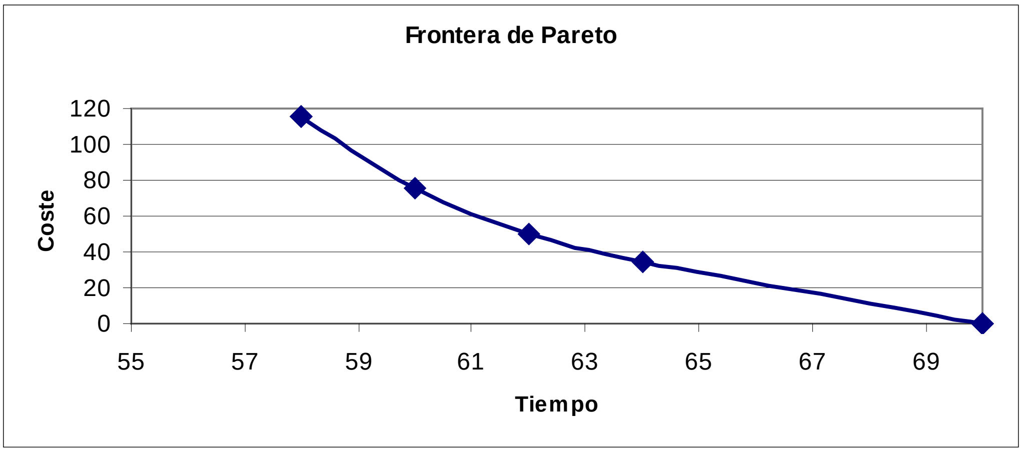
El coste 0 se obtiene sin ninguna reducción, dando una duración de 70 días. Mientras que para una duración de 58 días, coste 116.000 euros, las reducciones son:

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
r	**2**	**0**	**0**	**2**	**2**	**2**	**2**	**2**	**0**	**0**	**0**	**0**	**1**	**1**	**0**	**2**	**0**

La frontera de Pareto la obtenemos a continuación mediante ponderaciones obteniendo la siguiente tabla y el siguiente gráfico:

Tabla	

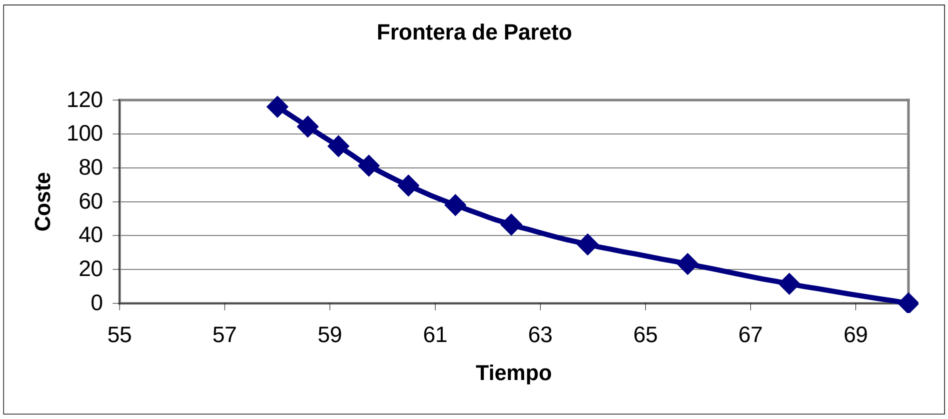
resultados				
Lambda	Agregado	Tiempo	Coste	Reducciones
0	0	70	0	
0,1	0,1	70	0	
0,2	0,2	70	0	
0,3	0,3	70	0	
0,4	0,37586	64	34	A, G, P 2
0,5	0,38218	62	50	A, D, G, P 2
0,6	0,36207	60	76	A, D, F, M, G, P 2 N 1
0,7	0,3	58	116	A, D, E, F, G, H, M, P 2 N 1
0,8	0,2	58	116	
0,9	0,1	58	116	
1	0	58	116	



Si lo hacemos mediante las e-restricciones la tabla y el gráfico resultan:

Tabla resultados		
Lambda	Tiempo	Coste
0	70	0
0,1	67,7333	11,6
0,2	65,8	23,2
0,3	63,9	34,8
0,4	62,45	46,4

0,5	61,3846	58
0,6	60,4923	69,6
0,7	59,74	81,2
0,8	59,16	92,8
0,9	58,58	104,4
1	58	116



A la hora de elegir un punto de esta frontera como solución, aplicamos la programación compromiso. Si se aplica con la métrica L1 y suponiendo igualdad de preferencias sobre los criterios (aunque normalizados con el ideal y anti-ideal), se tiene:

Tiempo	**62**	0,33333
Coste (miles de euros)	**50**	0,43103

aplicando las siguientes reducciones: A, D, G, P 2 días

Si utilizamos la norma infinito o Tchebychev el resultado es:

Tiempo	**62,6415**	0,38679
Coste (miles de euros)	**44,8679**	0,38679

aplicando las siguientes reducciones: A, G, P 2 días y D 1,3585 días.

Si cambiamos el enfoque y olvidamos la frontera de Pareto para plantear niveles de aspiración de los criterios y sus respectivas metas, el modelo de programación por metas ponderadas con niveles de 60 días de duración y 50.000 euros, e igualdad de preferencias da el siguiente resultado:

Tiempo	**62**
Coste (miles de euros)	**50**

que cumple la meta del coste pero no la del tiempo, con reducciones: A, D, G, P 2 días.

Si utilizamos las metas Tchebychev se tiene:

Tiempo	**61,8795**
Coste (miles de euros)	**51,5663**

que incumple ambas metas, en un 3,1% respecto a su nivel de aspiración, con las reducciones: A, D, G, P 2 días y F y M 0,1205 días.