Algoritmo de mejoramiento de una solución básica factible

Consideramos el problema

$$\begin{array}{ll} \text{Minimizar} & z = \mathbf{cx} \\ \text{sujeto a} & \mathbf{Ax} = \mathbf{b} \\ & \mathbf{x} \geq \mathbf{0} \end{array}$$

siendo $\mathbf{A}$ una matriz $m \times n$ de rango m.

Construiremos un algoritmo que nos permitirá partir de una solución básica factible para encontrar otra que ofrezca una mejor imagen en la función objetivo.

Sea $\begin{pmatrix} \mathbf{B}^{-1}\mathbf{b} \\ \mathbf{0} \end{pmatrix}$ una solución básica factible. Su imagen por la función objetivo será $\mathbf{c}\begin{pmatrix} \mathbf{B}^{-1}\mathbf{b} \\ \mathbf{0} \end{pmatrix} = (\mathbf{c_B}, \mathbf{c_N})\begin{pmatrix} \mathbf{B}^{-1}\mathbf{b} \\ \mathbf{0} \end{pmatrix} = \mathbf{c_B}\mathbf{B}^{-1}\mathbf{b} + \mathbf{c_N}\mathbf{0} = \boxed{\mathbf{c_B}\mathbf{B}^{-1}\mathbf{b} = \mathbf{z}_0}$

Sea además $\begin{pmatrix} \mathbf{x_B} \\ \mathbf{x_N} \end{pmatrix}$ una solución factible cualquiera (no necesariamente básica). Es decir, se cumplirá que $\mathbf{x_B} \geq \mathbf{0}$, $\mathbf{x_N} \geq \mathbf{0}$, $\mathbf{A}\begin{pmatrix} \mathbf{x_B} \\ \mathbf{x_N} \end{pmatrix} = \mathbf{b} \Leftrightarrow (\mathbf{B}, \mathbf{N})\begin{pmatrix} \mathbf{x_B} \\ \mathbf{x_N} \end{pmatrix} = \mathbf{b} \Leftrightarrow \mathbf{x_B} = \mathbf{B}^{-1}\mathbf{b} - \mathbf{B}^{-1}\mathbf{N}\mathbf{x_N}$.

Expresaremos de otro modo esta última igualdad. Sea R el conjunto de los índices de las variables no básicas. Esto implica que la matriz $\mathbf{N}$ está formada por las columnas de $\mathbf{A}$ cuyos índices pertenecen a R. Por eso: $\boxed{\mathbf{x_B} = \mathbf{B}^{-1}\mathbf{b} - \sum_{j \in R} \mathbf{B}^{-1}\mathbf{a}_j x_j}$. Nótese que con esta expresión podemos determinar el valor de las variables básicas una vez fijado el valor de las no básicas.

Reescribimos ahora la imagen de la función objetivo:

$$z = \mathbf{cx} = \mathbf{c_B}\mathbf{x_B} + \mathbf{c_N}\mathbf{x_N} = \mathbf{c_B}\left(\mathbf{B}^{-1}\mathbf{b} - \sum_{j \in R} \mathbf{B}^{-1}\mathbf{a}_j x_j\right) + \sum_{j \in R} c_j x_j = \underbrace{\mathbf{c_B}\mathbf{B}^{-1}\mathbf{b}}_{=z_0} - \mathbf{c_B}\sum_{j \in R} \mathbf{B}^{-1}\mathbf{a}_j x_j + \sum_{j \in R} c_j x_j =$$

$$z_0 - \sum_{j \in R}\left(\underbrace{\mathbf{c_B}\mathbf{B}^{-1}\mathbf{a}_j}_{=z_j} - c_j\right)x_j = \boxed{z_0 - \sum_{j \in R}\left(z_j - c_j\right)x_j}$$

Como nuestro objetivo es minimizar esta imagen, procuraremos hacer que el sustraendo $\sum_{j \in R}\left(z_j - c_j\right)x_j$ sea el máximo posible, ya que esto nos dará un valor de z mejor que z_0. Así que partiendo de la solución básica factible considerada, cambiaremos el valor de las variables no básicas (que hasta ahora todas valían 0) para que se mejore la función objetivo. Para ser ordenados, dejaremos que todas sigan valiendo 0, salvo una, que elegiremos convenientemente.

Si fuera x_k la elegida para ser positiva (dejando las restantes en 0), resultará que $z = z_0 - (z_k - c_k)x_k$.

Si el factor $(z_k - c_k)$ es negativo, cualquier incremento en el valor de x_k provocará un valor de z mayor que z_0, es decir, que empeora la imagen de la función. Por lo tanto, esa variable deberá seguir valiendo cero. Pero si todos los posibles factores $(z_j - c_j)$ son negativos, no nos conviene hacer positiva a ninguna de las variables x_k, porque esto empeoraría la función objetivo. En ese caso, estamos frente a una solución básica factible óptima, no es posible mejorarla.

Si el factor $(z_k - c_k)$ es positivo, implica que al aumentar el valor de x_k mejoraremos la función objetivo. Si hay varias opciones de factores $(z_j - c_j)$ positivos, elegiremos aquél que haga disminuir más *velozmente* la función objetivo. Para ello, tomamos k tal que $z_k - c_k = \max\{(z_j - c_j),\ j \in R\}$. Pero hay que ser cuidadosos, porque recordemos que al cambiar los valores de las variables no básicas deberán actualizarse los valores de las variables básicas (de acuerdo a la expresión de $\mathbf{x_B}$ recuadrada anteriormente), y si esto provocara un valor negativo en una variable básica, dejaría de ser factible, y no podemos permitirlo. Tendremos que ver cuándo detener el incremento de la variable x_k.

Retomemos: $\mathbf{x_B} = \mathbf{B}^{-1}\mathbf{b} - \mathbf{B}^{-1}\mathbf{a}_k x_k$. Si notamos $\overline{\mathbf{b}} = \mathbf{B}^{-1}\mathbf{b}$, $\mathbf{y}_k = \mathbf{B}^{-1}\mathbf{a}_k$ tenemos que $\mathbf{x_B} = \overline{\mathbf{b}} - \mathbf{y}_k x_k$. Expresado vectorialmente: $\begin{bmatrix} x_{\mathbf{B}1} \\ x_{\mathbf{B}2} \\ \vdots \\ x_{\mathbf{B}m} \end{bmatrix} = \begin{bmatrix} \overline{b}_1 \\ \overline{b}_2 \\ \vdots \\ \overline{b}_m \end{bmatrix} - \begin{bmatrix} y_{1k} \\ y_{2k} \\ \vdots \\ y_{mk} \end{bmatrix} x_k$.

Si todos los y_{ik} cumplen que $y_{ik} \le 0$, entonces x_k puede crecer tranquilamente porque los $x_{\mathbf{B}i}$ nunca serán negativos, y la imagen de la función objetivo será cada vez mejor. Estamos frente a un caso de solución óptima no acotada.

Pero si para algún i sucede que $y_{ik} > 0$, x_k podrá crecer hasta que alguna de las variables básicas llegue a valer 0. Sea entonces $x_k = \min\limits_{1 \le i \le m} \left\{ \dfrac{\overline{b}_i}{y_{ik}},\ y_{ik} > 0 \right\} = \dfrac{\overline{b}_r}{y_{rk}}$.

Ahora los valores de las $x_{\mathbf{B}i}$ se actualizan y resultan $x_{\mathbf{B}i} = \overline{b_i} - \dfrac{y_{ik}}{y_{rk}} \overline{b_r}$, con $i = 1, \ldots, m$.

En particular, $x_{\mathbf{B}r} = \overline{b_r} - \dfrac{y_{rk}}{y_{rk}} \overline{b_r} = 0$. Decimos que la variable x_k se vuelve positiva y "entra" a la base, mientras que $x_{\mathbf{B}r}$ se anula y "sale" de la base. O sea que estamos cambiando el conjunto de variables básicas, y por eso estamos ahora frente a una solución básica factible que es diferente a la inicial, pero que nos ofrece una mejor imagen por la función objetivo.

Es importante aclarar que, efectivamente, las columnas de la matriz $\mathbf{A}$ correspondientes a las nuevas variables básicas, también forman una base del espacio (es decir, que son linealmente independientes). Esto es cierto porque hemos sustituido la columna $\mathbf{a}_{\mathbf{B}r}$ por la $\mathbf{a}_k$, y como hemos definido $\mathbf{y}_k = \mathbf{B}^{-1}\mathbf{a}_k$, luego $\mathbf{a}_k = \mathbf{B}\mathbf{y}_k = \sum_{i=1}^{m} \mathbf{a}_{\mathbf{B}i} y_{ik}$ con $y_{rk} \neq 0$; es decir que $\mathbf{a}_k$ es combinación lineal de las columnas de $\mathbf{B}$ con un coeficiente no nulo para $\mathbf{a}_{\mathbf{B}r}$ y por tanto puede sustituirla en la base y el conjunto nuevo también es LI.

Hemos visto aquí un algoritmo que nos permite partir de una solución básica factible y encontrar otra que sea mejor que la primera. La aplicación reiterada de este algoritmo será el corazón del método simplex.

Los casos particulares (degenerados) no estudiados aquí serán analizados más adelante.