IN77O - Modelos y Algoritmos de Optimización

Resumen Teoría de Dualidad

Basado en el Capítulo 4 del libro *Introduction to linear optimization*

Profesor: Fernando Ordóñez
Auxiliares: Benjamín Barrientos, José Miguel González

3. El método *simplex* (resumen útil para entender dualidad)

El método *simplex* se basa en ir iterando sobre distintas soluciones básicas factibles del poliedro definido por las restricciones de un problema en forma estándar. Cada iteración tiene asociada una base B, matriz formada por las columnas de la matriz del problema correspondientes a las m variables distintas de 0 en cada solución básica (completando con otras variables iguales a 0 en el caso de soluciones básicas degeneradas).

Definición 3.1. Sea $\mathbf{x}$ una solución básica, $\mathbf{B}$ la matriz básica asociada y $\mathbf{c}_B$ el vector de costos de las variables básicas. Para cada j, definimos el **costo reducido** $\mathbf{c}_j$ de la variable $\mathbf{x}_j$ como:

$$\bar{\mathbf{c}}_j = \mathbf{c}_j - \mathbf{c}_B'\mathbf{B}^{-1}\mathbf{A}_j$$

Observación: $\bar{\mathbf{c}}_j$ se interpreta como el cambio en el valor de la función objetivo que supondría el aumento de una unidad en el valor de la j-ésima variable.

Teorema 3.1. Considere una solución básica factible $\mathbf{x}$ asociada con una matriz básica $\mathbf{B}$, y sea $\bar{\mathbf{c}}$ el correspondiente vector de costos reducidos.

(a) Si $\bar{\mathbf{c}} \geq \mathbf{0}$, entonces $\mathbf{x}$ es óptima.

(b) Si $\mathbf{x}$ es óptima y no degenerada, entonces $\bar{\mathbf{c}} \geq \mathbf{0}$.

Definición 3.2. Una matriz básica $\mathbf{B}$ se dice **óptima** si:

(a) $\mathbf{B}^{-1}\mathbf{b} \geq \mathbf{0}$, y

(b) $\bar{\mathbf{c}}' = \mathbf{c}' - \mathbf{c}_B'\mathbf{B}^{-1}\mathbf{A} \geq \mathbf{0}$

Teorema 3.2. Asumamos que el conjunto factible de un problema de optimización lineal es no vacío y que toda solución básica factible es no degenerada, o bien, que hay soluciones degeneradas pero se utiliza una regla de pivoteo para evitar ciclaje. Entonces, el método *simplex* finaliza luego de un número finito de iteraciones. Al terminar, existen dos posibilidades:

(a) Se tiene una base óptima $\mathbf{B}$ y una solución básica factible que es óptima.

(b) Se tiene un vector $\mathbf{d}$ que satisface $\mathbf{Ad} = \mathbf{0}$, $\mathbf{d} \geq \mathbf{0}$ y $\mathbf{c}'\mathbf{d} < \mathbf{0}$, y el costo óptimo es $-\infty$.

4. Teoría de Dualidad

4.1. Motivación

Dado un problema de programación lineal, se le asocia una variable de "precio" a cada restricción, y se comienzan a buscar los valores de estos precios, de tal forma que la presencia o ausencia de restricciones no afecte el costo óptimo. Los valores de estos precios se pueden encontrar resolviendo un nuevo problema de programación lineal.

Dado el siguiente *problema primal* en forma estándar

$$\begin{array}{lrcl} \text{mín} & \mathbf{c}'\mathbf{x} & & \\ \text{s.a:} & \mathbf{A}\mathbf{x} & = & \mathbf{b} \\ & \mathbf{x} & \geq & \mathbf{0} \end{array}$$

Asumiendo que este problema tiene solución óptima $\mathbf{x}^*$, se presenta el *problema relajado*, en el cual la restricción $\mathbf{A}\mathbf{x} = \mathbf{b}$ es remplazada por una penalización $\mathbf{y}'(\mathbf{b} - \mathbf{A}\mathbf{x})$ donde $\mathbf{y}$ es un vector que tiene la misma dimensión que $\mathbf{b}$:

$$\begin{array}{ll} \text{mín} & \mathbf{c}'\mathbf{x} + \mathbf{y}'(\mathbf{b} - \mathbf{A}\mathbf{x}) \\ \text{s.a:} & \mathbf{x} \geq \mathbf{0} \end{array}$$

Sea $g(\mathbf{y})$ el costo óptimo del problema relajado en función del vector de precios $\mathbf{y}$. Se tiene que:

$$g(\mathbf{y}) = \min_{\mathbf{x} \geq \mathbf{0}} \left[\mathbf{c}'\mathbf{x} + \mathbf{y}'(\mathbf{b} - \mathbf{A}\mathbf{x})\right] \leq \mathbf{c}'\mathbf{x}^* + \mathbf{y}'(\mathbf{b} - \mathbf{A}\mathbf{x}^*) = \mathbf{c}'\mathbf{x}^*$$

Es decir, cada $\mathbf{y}$ entrega una cota inferior $g(\mathbf{y})$ para el costo óptimo $\mathbf{c}'\mathbf{x}^*$. Entonces, maximizar $g(\mathbf{y})$ - que en este caso es un problema de maximización sin restricciones - se interpreta como la búsqueda de la mejor cota inferior para $\mathbf{c}'\mathbf{x}^*$. Reordenando la expresión de $g(\mathbf{y})$:

$$g(\mathbf{y}) = \mathbf{y}'\mathbf{b} + \min_{\mathbf{x} \geq \mathbf{0}} \left(\mathbf{c}' - \mathbf{y}'\mathbf{A}\right)\mathbf{x}$$

y notando que:

$$\min_{\mathbf{x} \geq \mathbf{0}} \left(\mathbf{c}' - \mathbf{y}'\mathbf{A}\right)\mathbf{x} = \begin{cases} 0 & \mathbf{c}' - \mathbf{y}'\mathbf{A} \geq \mathbf{0} \\ -\infty & \sim \end{cases}$$

se obtiene que solo se deben considerar los valores de $g(\mathbf{y})$ para los cuales $g(\mathbf{y}) \neq -\infty$. Entonces, se desea

$$\begin{array}{lrcl} \text{máx} & \mathbf{y}'\mathbf{b} & & \\ \text{s.a:} & \mathbf{y}'\mathbf{A} & \leq & \mathbf{c}' \end{array}$$

Este problema se conoce como el *dual* del problema original.

4.2. El problema dual

Sea $\mathbf{A}$ una matriz de filas $\mathbf{a}_i'$ y columnas $\mathbf{A}_j$. De forma general, un *problema primal* de minimización (P), posee un *problema dual* (D), donde las componentes de ambos problemas se relacionan de la siguiente forma:

$$\begin{array}{llll} \text{(P)} & \text{mín} & \mathbf{c}'\mathbf{x} & \\ & \text{s.a:} & \mathbf{a}_i'\mathbf{x} \geq b_i & i \in M_1 \\ & & \mathbf{a}_i'\mathbf{x} \leq b_i & i \in M_2 \\ & & \mathbf{a}_i'\mathbf{x} = b_i & i \in M_3 \\ & & x_j \geq 0 & j \in N_1 \\ & & x_j \leq 0 & j \in N_2 \\ & & x_j \in \mathbb{R} & j \in N_3 \end{array} \qquad \begin{array}{llll} \text{(D)} & \text{máx} & \mathbf{y}'\mathbf{b} & \\ & \text{s.a:} & y_i \geq 0 & i \in M_1 \\ & & y_i \leq 0 & i \in M_2 \\ & & y_i \in \mathbb{R} & i \in M_3 \\ & & \mathbf{y}'\mathbf{A}_j \leq c_j & j \in N_1 \\ & & \mathbf{y}'\mathbf{A}_j \geq c_j & j \in N_2 \\ & & \mathbf{y}'\mathbf{A}_j = c_j & j \in N_3 \end{array}$$

Notar que por cada restricción en el primal (sin contar las de signo), se introduce una variable en el problema dual; por cada variable en el primal, se introduce una restricción en el dual. A continuación, se resumen las relaciones de signo entre variables y restricciones de los problemas primal y dual:

PRIMAL	minimizar	maximizar	DUAL
restricciones	$\geq b_i$ $\leq b_i$ $= b_i$	≥ 0 ≤ 0 $\mathbb{R}$	**variables**
variables	≥ 0 ≤ 0 $\mathbb{R}$	$\leq c_j$ $\geq c_j$ $= c_j$	**restricciones**

Teorema 4.1. Si se transforma el dual en un problema equivalente de minimización y luego se toma su dual, se obtiene un problema equivalente al problema original.

Teorema 4.2. Supongamos ahora que se transformó un problema de programación lineal Π_1 en otro problema de programación lineal Π_2 mediante una secuencia de transformaciones del tipo:

(a) Reemplazar una variable libre como diferencia de dos variables no negativas.

(b) Agregar variables de holgura para convertir desigualdades en igualdades.

(c) Si alguna fila de la matriz $\mathbf{A}$ en un problema en forma estándar factible es linealmente dependiente, eliminar la respectiva restricción de igualdad.

Entonces, los duales de Π_1 y Π_2 son equivalentes, es decir, o los dos son infactibles, o los dos tienen el mismo costo óptimo.

4.3. Teoremas de dualidad

Teorema 4.3. (Dualidad débil) Si $\mathbf{x}$ es solución factible del problema primal (minimización) e $\mathbf{y}$ es solución factible del problema dual (maximización), entonces:

$$\mathbf{y}'\mathbf{b} \leq \mathbf{c}'\mathbf{x}$$

Corolario 4.1.

(a) Si el costo óptimo del primal es $-\infty$, el dual debe ser infactible.

(b) Si el costo óptimo del dual es ∞, el primal debe ser infactible.

Corolario 4.2. Sean $\mathbf{x}$ e $\mathbf{y}$ soluciones factibles al problema primal y dual respectivamente, y supongamos que $\mathbf{y}'\mathbf{b} = \mathbf{c}'\mathbf{x}$. Entonces, $\mathbf{x}$ e $\mathbf{y}$ son soluciones óptimas del primal y el dual respectivamente.

Teorema 4.4. (Dualidad fuerte) Si un problema de programación lineal tiene una solución óptima, entonces su dual también la tiene y sus respectivos costos óptimos son iguales.

De los teoremas de dualidad débil y fuerte, se puede concluir sobre si los problemas tienen óptimo finito, si son no acotados o si son infactibles de acuerdo a la siguiente tabla de posibilidades para el primal y el dual:

	Óptimo finito	**No acotado**	**Infactible**
Óptimo finito	Posible	Imposible	Imposible
No acotado	Imposible	Imposible	Posible
Infactible	Imposible	Posible	Posible

Teorema 4.5. (Holgura complementaria) Sean $\mathbf{x}, \mathbf{y}$ soluciones factibles del problema primal y del dual respectivamente. Los vectores $\mathbf{x}$ e $\mathbf{y}$ son soluciones óptimas para los respectivos problemas si y sólo si:

$$\begin{array}{rcll} y_i(\mathbf{a}_i'\mathbf{x} - b_i) & = & 0 & \forall i \\ (c_j - \mathbf{y}'\mathbf{A}_j)x_j & = & 0 & \forall j \end{array}$$

4.4. Variables duales como costos marginales

Dado el problema en forma estándar

$$\begin{array}{llrcl} \text{(P)} & \text{mín} & \mathbf{c}'\mathbf{x} & & \\ & \text{s.a:} & \mathbf{A}\mathbf{x} & = & \mathbf{b} \\ & & \mathbf{x} & \geq & \mathbf{0} \end{array}$$

y suponiendo que existe una solución básica factible óptima no degenerada $\mathbf{x}^*$, se tiene que:

$$\mathbf{x} = \mathbf{B}^{-1}\mathbf{b} > 0$$

$$\mathbf{c}'\mathbf{x}^* = \mathbf{y}^{*\prime}\mathbf{b}$$

Ahora bien, si se añade una perturbación pequeña $\mathbf{d}$ en en el lado derecho $\mathbf{b}$, se tendrá que:

$$\bar{\mathbf{x}}^* = \mathbf{B}^{-1}(\mathbf{b} + \mathbf{d}) > 0$$

para algún valor de $\mathbf{d}$. Es decir, la base actual B es también una base factible del problema perturbado. Más aún, como el vector del lado derecho no tiene relación con el vector de costos reducidos, la base seguirá siendo óptima y el óptimo dual seguirá siendo $\mathbf{y}^* = \mathbf{c}_B\mathbf{B}^{-1}$. Por dualidad fuerte, el costo óptimo del problema perturbado verificará:

$$\mathbf{c}'\bar{\mathbf{x}}^* = \mathbf{c}_B\mathbf{B}^{-1}(\mathbf{b} + \mathbf{d}) = \mathbf{y}^{*\prime}(\mathbf{b} + \mathbf{d}) = \mathbf{y}^{*\prime}\mathbf{b} + \sum_{i=1}^{n} y_i^* d_i = \mathbf{c}'\mathbf{x}^* + \sum_{i=1}^{n} y_i^* d_i$$

Es decir, cada componente y_i^* del óptimo dual $\mathbf{y}^*$ puede interpretarse como el *costo marginal* o *precio sombra* de aumentar en una unidad la i-ésima componente de $\mathbf{b}$.

4.6. Lema de Farkas y desigualdades lineales

Teorema 4.6. (Lema de Farkas) Sea $\mathbf{A} \in \mathcal{M}_{m\times n}$ y sea $\mathbf{b} \in \mathbb{R}^m$. Entonces, exactamente una de las siguientes alternativas se cumple:

(a) Existe algún $\mathbf{x} \geq \mathbf{0}$ tal que $\mathbf{A}\mathbf{x} = \mathbf{b}$.

(b) Existe algún vector $\mathbf{y}$ tal que $\mathbf{y}'\mathbf{A} \geq \mathbf{0}'$ e $\mathbf{y}'\mathbf{b} < 0$.

Corolario 4.3. Sean $\mathbf{A}_{\cdot 1}, ..., \mathbf{A}_{\cdot n}$ y $\mathbf{b}$ vectores dados, y suponga que cualquier vector $\mathbf{y}$ que satisfaga $y'\mathbf{A}_{\cdot i} \geq 0$, $i = 1, ..., n$, también debe satisfacer $\mathbf{y}'\mathbf{b} \geq 0$. Entonces, $\mathbf{b}$ puede ser expresado como una combinación lineal no negativa de los vectores $\mathbf{A}_{\cdot 1}, ..., \mathbf{A}_{\cdot n}$.

Teorema 4.7. Suponga que el sistema de inecuaciones lineales $\mathbf{Ax} \leq \mathbf{b}$ tiene al menos una solución, y sea d algún escalar. Entonces, las siguientes son equivalentes:

(a) Cada solución factible del sistema $\mathbf{Ax} \leq \mathbf{b}$ satisface $\mathbf{c}'\mathbf{x} \leq d$.

(b) Existe algún $\mathbf{y} \geq 0$ tal que $\mathbf{y}'\mathbf{A} = \mathbf{c}'$ y $\mathbf{y}'\mathbf{b} \leq d$.

4.8. Conos y rayos extremos

Definición 4.1. Un conjunto $C \subset \mathbb{R}^n$ es un **cono** si $\lambda\mathbf{x} \in C$ para todo $\lambda \geq 0$ y todo $\mathbf{x} \in C$.

Observaciones:

- El único posible punto extremo de un cono es el vector $\mathbf{0}$. Si un cono tiene al $\mathbf{0}$ como punto extremo, se dice que el cono es **puntiagudo**.
- Un poliedro de la forma $\mathcal{P} = \{\mathbf{x} \in \mathbb{R}^n | \mathbf{Ax} \geq \mathbf{0}\}$ es llamado **cono poliedral**.

Teorema 4.8. Sea $C \subset \mathbb{R}^n$ el cono poliedral definido por las restricciones $\mathbf{a}_i'\mathbf{x} \geq 0$, $i = 1, ..., m$. Entonces, las siguientes son equivalentes:

(a) C es puntiagudo.

(b) C no contiene una línea.

(c) Existen n vectores de la familia $\mathbf{a}_1, ..., \mathbf{a}_m$ que son linealmente independientes.

Definición 4.2.

(a) Dado un poliedro $\mathcal{P} = \{\mathbf{x} \in \mathbb{R}^n | \mathbf{Ax} \geq \mathbf{b}\}$ y un vector $\mathbf{y} \in \mathcal{P}$, se define el **cono de recesión en y** como el conjunto de todas las direcciones $\mathbf{d}$ en las cuales nos podemos mover indefinidamente lejos de $\mathbf{y}$, sin dejar el conjunto $\mathcal{P}$. Esto es:

$$\{\mathbf{d} \in \mathbb{R}^n | \mathbf{A}(\mathbf{y} + \lambda\mathbf{d}) \geq \mathbf{b}, \forall\lambda \geq 0\} = \{\mathbf{d} \in \mathbb{R}^n | \mathbf{Ad} \geq \mathbf{0}\}$$

(b) Los elementos distintos de cero del cono de recesión se llaman **rayos** del poliedro $\mathcal{P}$.

(c) Un elemento $\mathbf{x}$ distinto de cero de un cono poliedral $C \subset \mathbb{R}^n$ es llamado un **rayo extremo** si hay $(n-1)$ restricciones linealmente independientes que son activas en $\mathbf{x}$.

(d) Un rayo extremo del cono de recesión asociado con un poliedro no vacío $\mathcal{P}$ también se denomina **rayo extremo** de $\mathcal{P}$.

Teorema 4.9. Consideremos el problema de minimizar $\mathbf{c}'\mathbf{x}$ sobre un cono poliedral puntiagudo $C = \{\mathbf{x} \in \mathbb{R}^n | \mathbf{a_i}'\mathbf{x} \geq \mathbf{0},\ i = 1, \dots, m\}$. Entonces, el costo óptimo es igual a $-\infty$ si y sólo si algún rayo extremo $\mathbf{d}$ de C satisface $\mathbf{c}'\mathbf{d} < 0$.

Teorema 4.10. Consideremos el problema de minimizar $\mathbf{c}'\mathbf{x}$ sujeto a $\mathbf{Ax} \geq \mathbf{b}$, y asumamos que el conjunto factible tiene al menos un punto extremo. Entonces, el costo óptimo es igual a $-\infty$ si y sólo si algún rayo extremo $\mathbf{d}$ del conjunto factible satisface $\mathbf{c}'\mathbf{d} < 0$.

4.9. Representación de poliedros

Teorema 4.11. Sea

$$\mathcal{P} = \{\mathbf{x} \in \mathbb{R}^n | \mathbf{A}\mathbf{x} \geq \mathbf{b}\}$$

un poliedro no vacío con al menos un punto extremo. Sean $\mathbf{x}^1, ..., \mathbf{x}^k$ los puntos extremos, y sea $\mathbf{w}^1, ..., \mathbf{w}^r$ el conjunto completo de los rayos extremos de $\mathcal{P}$. Sea además:

$$\mathcal{Q} = \left\{ \sum_{i=1}^{k} \lambda_i \mathbf{x}^i + \sum_{j=1}^{r} \theta_j \mathbf{w}^j \middle| \lambda_i \geq 0, \theta_j \geq 0, \sum_{i=1}^{k} \lambda_i = 1 \right\}$$

Entonces, $\mathcal{Q} = \mathcal{P}$.

Corolario 4.4. Un poliedro no vacío y acotado es la envoltura convexa de sus puntos extremos.

Corolario 4.5. Asuma que el cono $C = \{\mathbf{x} | \mathbf{A}\mathbf{x} \geq \mathbf{0}\}$ es puntiagudo. Entonces, cada elemento de C puede ser expresado como una combinación lineal no negativa de los rayos extremos de C.

Referencias

[1] D. Bertsimas y J.N. Tsitsiklis. *Introduction to linear optimization.* Athena Scientific, 1997.