Resumen IN3701 - Modelamiento y Optimización

Basado en el libro: *Introduction to Linear Optimization*

Por: Cristian Aguayo Quintana
cristian.aguayo@ug.uchile.cl

1. Geometría de la Programación Lineal

1.1. Poliedros y conjuntos convexos

Definición 1.1. $\mathcal{P} \subseteq \mathbb{R}^n$ es un **poliedro** si puede ser representado como un conjunto finito de inecuaciones lineales. Esto es:

$$\mathcal{P} := \{\mathbf{x} = (x_1, ..., x_n) \in \mathbb{R}^n \mid \mathbf{A}\mathbf{x} \geq \mathbf{b}\}$$

donde $\mathbf{A} \in \mathcal{M}_{m\times n}$, $\mathbf{b} \in \mathbb{R}^m$ y m es el número de inecuaciones. La región factible de un *problema de programación lineal* es un poliedro.

Definición 1.2. Un **hiperplano** es un poliedro determinado por una única restricción de igualdad.

$$\mathcal{H} := \left\{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{a}'\mathbf{x} = \mathbf{b}\right\}$$

Definición 1.3. Un **semiespacio** es un poliedro determinado por una única restricción de desigualdad.

$$\mathcal{S} := \left\{\mathbf{x} \in \mathbb{R} \mid \mathbf{a}'\mathbf{x} \geq \mathbf{b}\right\}$$

Todo semiespacio está delimitado por el hiperplano correspondiente.

Definición 1.4. Un conjunto $S \subseteq \mathbb{R}^n$ es **convexo** si:

$$\forall \mathbf{x}, \mathbf{y} \in S,\ \forall \lambda \in [0, 1] \qquad \lambda\mathbf{x} + (1 - \lambda)\mathbf{y} \in S$$

En particular, se tiene que los poliedros son conjuntos convexos.

Definición 1.5. Sean $\mathbf{x}^1, ..., \mathbf{x}^k$ vectores $\in \mathbb{R}^n$ y sean $\lambda_1, ..., \lambda_k \geq 0$ escalares tales que $\sum_{i=1}^k \lambda_i = 1$:

(a) El vector $\sum_{i=1}^k \lambda_i \mathbf{x}^i$ es una **combinación convexa** de los vectores $\mathbf{x}^1, ..., \mathbf{x}^k$.

(b) La **envoltura convexa** de los vectores $\mathbf{x}^1, ..., \mathbf{x}^k$ es el conjunto de todas las combinaciones convexas de estos vectores

Teorema 1.1.

(a) La intersección de conjuntos convexos es un conjunto convexo.

(b) Todos los poliedros son conjuntos convexos.

(c) Una combinación convexa de un número finito de elementos de un conjunto convexo también pertenece a ese conjunto.

(d) La envoltura convexa de un número finito de vectores es un conjunto convexo.

1.2. Puntos extremos, vértices y soluciones básicas factibles

Definición 1.6. Un **punto extremo** es aquel que no puede ser expresado como una combinación convexa de otros dos puntos de un poliedro. Es decir:

$$\nexists \mathbf{y}, \mathbf{z} \in \mathcal{P} \setminus \{\mathbf{x}\}, \lambda \in [0,1] : \qquad \mathbf{x} = \lambda \mathbf{y} + (1-\lambda)\mathbf{z}$$

Definición 1.7. Un punto $\mathbf{x}$ es **vértice** de un poliedro $\mathcal{P}$ si existe una recta que intersecta a $\mathcal{P}$ sólo en el punto $\mathbf{x}$. Esto es:

$$\exists \mathbf{c} \mid \mathbf{c}'\mathbf{x} < \mathbf{c}'\mathbf{y} \;\; \forall \mathbf{y} \neq \mathbf{x} \in \mathcal{P}$$

Definición 1.8. Si un vector $\mathbf{x}^*$ satisface $\mathbf{a}_i^{\mathbf{t}}\mathbf{x}^* = b_i$ para algún i, se dice que la i-ésima restricción es **activa** en $\mathbf{x}^*$.

Teorema 1.2. Sea $\mathbf{x}^* \in \mathbb{R}^{\ltimes}$ y sea $I = \left\{ i : \mathbf{a}_i^{\mathbf{t}}\mathbf{x}^* = \mathbf{b}_i \right\}$ el conjunto de todas las restricciones activas en $\mathbf{x}^*$. Entonces, las siguientes son equivalentes:

(a) Existen n vectores en el conjunto $\{\mathbf{a}_i \mid i \in I\}$ que son linealmente independientes.

(b) La colección de vectores $\mathbf{a}$ con $i \in I$ son una base de $\mathbb{R}^n$

(c) El sistema de ecuaciones de todas las restricciones activas tiene solución única.

Definición 1.9. Sea un poliedro $\mathcal{P}$ definido por restricciones lineales de igualdad y desigualdad, y sea $\mathbf{x}^* \in \mathbb{R}^n$:

(a) $\mathbf{x}^*$ es una **solución básica** si:

   (i) Todas las restricciones de igualdad son activas.

   (ii) De las restricciones que son activas en $\mathbf{x}^*$, hay n de ellas que son linealmente independientes.

(b) Si $\mathbf{x}^*$ es una solución básica que satisface todas las restricciones, se dice que es una **solución básica factible**.

Teorema 1.3. Sea $\mathcal{P}$ un poliedro no vacío, y sea $\mathbf{x}^* \in \mathcal{P}$. Entonces, las siguientes son equivalentes:

(a) $\mathbf{x}^*$ es un vértice;

(b) $\mathbf{x}^*$ es un punto extremo;

(c) $\mathbf{x}^*$ es solución básica factible.

Corolario 1.1. Dado un número finito de restricciones lineales de desigualdad, solo puede haber un número finito de soluciones básicas o soluciones básicas factibles.

1.3. Poliedros en forma estándar

Un poliedro en **forma estándar** es aquel que puede ser representado solo con restricciones de igualdad y solo con variables no negativas. Es decir:

$$\mathcal{P} := \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{A}\mathbf{x} = \mathbf{b},\ \mathbf{x} \geq \mathbf{0}\}$$

Sin pérdida de generalidad, se asume que las m restricciones son linealmente independientes, lo que implica que $m \leq n$.
Para representar la región factible de un PPL como un poliedro en forma estándar, puede ser útil considerar:

- Sumar o restar *variables de holgura* $x_{h_k} \geq 0$ en desigualdades para transformarlas en igualdades.
- En el caso que $x_j \leq 0$, cambiar x_j por $-x_j$ y así poder definir $x_j \geq 0$.
- Escribir como $x_i^+ - x_i^-$ con $x_i^+, x_i^- \geq 0$ las variables x_i que no tengan restricción de signo ($x_i \in \mathbb{R}$).

Teorema 1.4. Considere las restricciones $\mathbf{A}\mathbf{x} = \mathbf{b}$ y $\mathbf{x} \geq 0$, y asuma que la matriz $\mathbf{A}$ tiene filas linealmente independientes. Un vector $\mathbf{x} \in \mathbb{R}^n$ es una solución básica factible si y sólo si $\mathbf{A}\mathbf{x} = \mathbf{b}$, y si existen índices $B(1), ..., B(m)$ tales que:

(a) Las columnas $\mathbf{A}_{B(1)}, ..., \mathbf{A}_{B(m)}$ son linealmente independientes;

(b) Si $i \neq B(1), ..., B(m)$, entonces $x_i = 0$.

El procedimiento para construir soluciones básicas es:

1. Elegir m columnas linealmente independientes $A_{\cdot B(1)}, ..., A_{\cdot B(m)}$.
2. Dejar $x_i = 0$ para todos los $i \neq B(1), ..., B(m)$.
3. Resolver el sistema de m ecuaciones $\mathbf{A}\mathbf{x} = \mathbf{b}$ para las incógnitas $x_{B(1)}, ..., x_{B(m)}$. Estas incógnitas recibirán el nombre de **variables básicas**.

1.4. Degenerancia

Definición 1.10. Una solución básica $\mathbf{x}$ se dice **degenerada** si hay más de n restricciones activas en $\mathbf{x}$,

Definición 1.11. Considere el poliedro en forma estándar $\mathcal{P} = \{\mathbf{x} \in \mathbb{R}^n | \mathbf{A}\mathbf{x} = \mathbf{b}, \mathbf{x} \geq 0\}$ y sea $\mathbf{x}$ una solución básica. Sea m el número de filas de $\mathbf{A}$. El vector $\mathbf{x}$ es solución básica **degenerada** si hay más de $n - m$ componentes de $\mathbf{x}$ que son cero.

1.5. Existencia de puntos extremos y optimalidad

Definición 1.12. Un poliedro $\mathcal{P} \subset \mathbb{R}^n$ **contiene una línea** si existe un vector $\mathbf{x} \in \mathcal{P}$ y un vector no nulo $\mathbf{d} \in \mathbb{R}^n$ tal que $\mathbf{x} + \lambda\mathbf{d} \in \mathcal{P}$ para cualquier $\lambda \in \mathbb{R}$.

Teorema 1.5. Dado un poliedro $\mathcal{P} := \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{a}_i^{\mathbf{t}}\mathbf{x} \geq b_i,\ i = 1, ..., m\}$ no vacío, entonces, las siguientes son equivalentes:

(a) El poliedro $\mathcal{P}$ tiene al menos un punto extremo.

(b) El poliedro $\mathcal{P}$ no contiene una recta.

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

Corolario 1.2. Todo poliedro no vacío acotado y todo poliedro no vacío en forma estándar tiene al menos una solución básica factible.

1.6. Optimalidad de puntos extremos

Considere el problema de programación lineal que consiste en minimizar $\mathbf{c}'\mathbf{x}$ con $\mathbf{x} \in \mathcal{P}$. Suponga además que $\mathcal{P}$ tiene al menos un punto extremo. Entonces, o el costo óptimo es $-\infty$, o existe un punto extremo de $\mathcal{P}$ que es óptimo.

2. Geometría de la Programación Lineal

2.1. Poliedros y conjuntos convexos

Definición 2.1. $P \subseteq \mathbb{R}^n$ es un **poliedro** si puede ser representado como un conjunto finito de inecuaciones lineales. Esto es:

$$P := \{\mathbf{x} = (x_1, ..., x_n) \in \mathbb{R}^n \mid \mathbf{A}\mathbf{x} \geq \mathbf{b}\}$$

donde $\mathbf{A} \in \mathcal{M}_{m\times n}$, $\mathbf{b} \in \mathbb{R}^m$ y m es el número de inecuaciones. La región factible de un *problema de programación lineal* es un poliedro.

Definición 2.2. Un conjunto $S \subset \mathbb{R}^n$ es **acotado** si existe una constante K tal que el valor absoluto de cada componente de cada elemento de S es menor o igual que K.

Definición 2.3. Sea $\mathbf{a} \neq \mathbf{0} \in \mathbb{R}^n$ y sea b un escalar:

(a) El conjunto $\{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{a}'\mathbf{x} = b\}$ corresponde a un **hiperplano**

(b) El conjunto $\{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{a}'\mathbf{x} \geq b\}$ corresponde a un **semiespacio**

Todo semiespacio está delimitado por el hiperplano correspondiente.

Definición 2.4. Un conjunto $S \subseteq \mathbb{R}^n$ es **convexo** si:

$$\forall \mathbf{x}, \mathbf{y} \in S,\ \forall \lambda \in [0, 1] \qquad \lambda\mathbf{x} + (1-\lambda)\mathbf{y} \in S$$

En particular, se tiene que los poliedros son conjuntos convexos.

Definición 2.5. Sean $\mathbf{x}^1, ..., \mathbf{x}^k$ vectores $\in \mathbb{R}^n$ y sean $\lambda_1, ..., \lambda_k \geq 0$ escalares tales que $\sum_{i=1}^k \lambda_i = 1$:

(a) El vector $\sum_{i=1}^k \lambda_i \mathbf{x}^i$ es una **combinación convexa** de los vectores $\mathbf{x}^1, ..., \mathbf{x}^k$.

(b) La **envoltura convexa** de los vectores $\mathbf{x}^1, ..., \mathbf{x}^k$ es el conjunto de todas las combinaciones convexas de estos vectores

Teorema 2.1.

(a) La intersección de conjuntos convexos es un conjunto convexo.

(b) Todos los poliedros son conjuntos convexos.

(c) Una combinación convexa de un número finito de elementos de un conjunto convexo también pertenece a ese conjunto.

(d) La envoltura convexa de un número finito de vectores es un conjunto convexo.

2.2. Puntos extremos, vértices y soluciones básicas factibles

Definición 2.6. Un **punto extremo** es aquel que no puede ser expresado como una combinación convexa de otros dos puntos de un poliedro. Es decir:

$$\nexists \mathbf{y}, \mathbf{z} \in P \setminus \{\mathbf{x}\}, \lambda \in [0,1] : \qquad \mathbf{x} = \lambda \mathbf{y} + (1-\lambda)\mathbf{z}$$

Definición 2.7. Un punto $\mathbf{x}$ es **vértice** de un poliedro P si existe una recta que intersecta a P sólo en el punto $\mathbf{x}$. Esto es:

$$\exists \mathbf{c} \mid \mathbf{c}'\mathbf{x} < \mathbf{c}'\mathbf{y} \;\; \forall \mathbf{y} \neq \mathbf{x} \in P$$

Definición 2.8. Si un vector $\mathbf{x}^*$ satisface $\mathbf{a}_i'\mathbf{x}^* = b_i$ para algún i, se dice que la i-ésima restricción es **activa** en $\mathbf{x}^*$.

Teorema 2.2. Sea $\mathbf{x}^* \in \mathbb{R}^{\ltimes}$ y sea $I = \{i : \mathbf{a}_i'\mathbf{x}^* = \mathbf{b}_i\}$ el conjunto de todas las restricciones activas en $\mathbf{x}^*$. Entonces, las siguientes son equivalentes:

(a) Existen n vectores en el conjunto $\{\mathbf{a}_i \mid i \in I\}$ que son linealmente independientes.

(b) La colección de vectores $\mathbf{a}$ con $i \in I$ son una base de $\mathbb{R}^n$

(c) El sistema de ecuaciones de todas las restricciones activas tiene solución única.

Definición 2.9. Sea un poliedro P definido por restricciones lineales de igualdad y desigualdad, y sea $\mathbf{x}^* \in \mathbb{R}^n$:

(a) $\mathbf{x}^*$ es una **solución básica** si:

  (i) Todas las restricciones de igualdad son activas.
  (ii) De las restricciones que son activas en $\mathbf{x}^*$, hay n de ellas que son linealmente independientes.

(b) Si $\mathbf{x}^*$ es una solución básica que satisface todas las restricciones, se dice que es una **solución básica factible**.

Teorema 2.3. Sea P un poliedro no vacío, y sea $\mathbf{x}^* \in P$. Entonces, las siguientes son equivalentes:

(a) $\mathbf{x}^*$ es un vértice;

(b) $\mathbf{x}^*$ es un punto extremo;

(c) $\mathbf{x}^*$ es solución básica factible.

Corolario 2.1. Dado un número finito de restricciones lineales de desigualdad, solo puede haber un número finito de soluciones básicas o soluciones básicas factibles.

2.3. Poliedros en forma estándar

Un poliedro en **forma estándar** es aquel que puede ser representado solo con restricciones de igualdad y solo con variables no negativas. Es decir:

$$P := \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{Ax} = \mathbf{b},\ \mathbf{x} \geq \mathbf{0}\}$$

Sin pérdida de generalidad, se asume que las m restricciones son linealmente independientes, lo que implica que $m \leq n$.
Para representar la región factible de un PPL como un poliedro en forma estándar, puede ser útil considerar:

- Sumar o restar *variables de holgura* $x_{h_k} \geq 0$ en desigualdades para transformarlas en igualdades.
- En el caso que $x_j \leq 0$, cambiar x_j por $-x_j$ y así poder definir $x_j \geq 0$.
- Escribir como $x_i^+ - x_i^-$ con $x_i^+, x_i^- \geq 0$ las variables x_i que no tengan restricción de signo ($x_i \in \mathbb{R}$).

Teorema 2.4. Considere las restricciones $\mathbf{Ax} = \mathbf{b}$ y $\mathbf{x} \geq 0$, y asuma que la matriz $\mathbf{A}$ tiene filas linealmente independientes. Un vector $\mathbf{x} \in \mathbb{R}^n$ es una solución básica si y sólo si $\mathbf{Ax} = \mathbf{b}$, y si existen índices $B(1), ..., B(m)$ tales que:

(a) Las columnas $\mathbf{A}_{B(1)}, ..., \mathbf{A}_{B(m)}$ son linealmente independientes;

(b) Si $i \neq B(1), ..., B(m)$, entonces $x_i = 0$.

El procedimiento para construir soluciones básicas es:

1. Elegir m columnas linealmente independientes $A_{\cdot B(1)}, ..., A_{\cdot B(m)}$.
2. Dejar $x_i = 0$ para todos los $i \neq B(1), ..., B(m)$.
3. Resolver el sistema de m ecuaciones $\mathbf{Ax} = \mathbf{b}$ para las incógnitas $x_{B(1)}, ..., x_{B(m)}$. Estas incógnitas recibirán el nombre de **variables básicas**.

Teorema 2.5. Sea $P = \{\mathbf{x} \mid \mathbf{Ax} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$ un poliedro no vacío, donde $\mathbf{A} \in \mathcal{M}_{m\times n}$, $\mathbf{a}_1', ..., \mathbf{a}_m'$ son las filas de $\mathbf{A}$, rango$(\mathbf{A}) = k < m$, y las filas $\mathbf{a}_{i_1}', ..., \mathbf{a}_{i_k}'$ son linealmente independientes. Sea

$$Q = \left\{\mathbf{x} \mid \mathbf{a}_{i_1}'\mathbf{x} = b_{i_1}, ..., \mathbf{a}_{i_k}'\mathbf{x} = b_{i_k}, \mathbf{x} \geq \mathbf{0}\right\}$$

Entonces, $Q = P$.

2.4. Degenerancia

Definición 2.10. Una solución básica $\mathbf{x}$ se dice **degenerada** si hay más de n restricciones activas en $\mathbf{x}$,

Definición 2.11. Considere el poliedro en forma estándar $P = \{\mathbf{x} \in \mathbb{R}^n | \mathbf{Ax} = \mathbf{b}, \mathbf{x} \geq 0\}$ y sea $\mathbf{x}$ una solución básica. Sea m el número de filas de $\mathbf{A}$. El vector $\mathbf{x}$ es solución básica **degenerada** si hay más de $n - m$ componentes de $\mathbf{x}$ que son cero.

2.5. Existencia de puntos extremos y optimalidad

Definición 2.12. Un poliedro $P \subset \mathbb{R}^n$ **contiene una línea** si existe un vector $\mathbf{x} \in P$ y un vector no nulo $\mathbf{d} \in \mathbb{R}^n$ tal que $\mathbf{x} + \lambda \mathbf{d} \in P$ para cualquier $\lambda \in \mathbb{R}$.

Teorema 2.6. Dado un poliedro $P := \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{a}_i'\mathbf{x} \geq b_i,\ i = 1, ..., m\}$ no vacío, entonces, las siguientes son equivalentes:

(a) El poliedro P tiene al menos un punto extremo.

(b) El poliedro P no contiene una recta.

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

Corolario 2.2. Todo poliedro no vacío acotado y todo poliedro no vacío en forma estándar tiene al menos una solución básica factible.

2.6. Optimalidad de puntos extremos

Teorema 2.7. Considere el problema de programación lineal que consiste en minimizar $\mathbf{c}'\mathbf{x}$ con $\mathbf{x} \in P$. Suponga además que P tiene al menos un punto extremo. Entonces, o el costo óptimo es $-\infty$, o existe un punto extremo de P que es óptimo.

Corolario 2.3. Considere el problema de programación linear de minimizar $\mathbf{c}'\mathbf{x}$ sobre un poliedro no vacío. Entonces, o el costo óptimo es $-\infty$, o existe una solución óptima.

3. Método Simplex

3.1. Condiciones de optimalidad

3.1.1. Propiedades de la forma estándar

Dado un problema en forma estándar:

$$\begin{array}{lll} (\mathrm{P}) & \text{mín} & \displaystyle\sum_{i=1}^{n} c_i x_i \\ & \text{s.a:} & \begin{array}{rcll} \mathbf{A}\mathbf{x} & = & \mathbf{b} & \\ x_i & \geq & 0 & i = 1, ..., n \end{array} \end{array}$$

y considerando $B(1), ..., B(m)$ los índices de las variables básicas de una solución básica factible, se construye la *matriz básica* $\mathbf{B}$ con las columnas $A_{\cdot B(1)}, ..., A_{\cdot B(m)}$ y la *matriz no básica* $\mathbf{N}$ con las demás columnas de $\mathbf{A}$. Reescribiendo $\mathbf{x} = (\mathbf{x}_B, \mathbf{x}_N)$, entonces el problema se puede escribir separando las variables básicas de las no básicas.

$$\begin{array}{llccccl} (\mathrm{P}) & \text{mín} & \displaystyle\sum_{i \in B} c_{B(i)} x_{B(i)} & + & \displaystyle\sum_{j \in N} c_{N(j)} x_{N(j)} & & \\ & \text{s.a:} & \displaystyle\sum_{i \in \mathcal{B}} B_{\cdot i} x_{B(i)} & + & \displaystyle\sum_{j \in \mathcal{N}} N_{\cdot j} x_{N(j)} & = & \mathbf{b} \\ & & & & x_i & \geq & 0 \quad i = 1, ..., n \end{array}$$

o bien, en forma matricial:

$$\begin{array}{llccccc} (\mathrm{P}) & \text{mín} & \mathbf{c}_B^{\mathbf{t}} \mathbf{x}_B & + & \mathbf{c}_N^{\mathbf{t}} \mathbf{x}_N & & \\ & \text{s.a:} & \mathbf{B}\mathbf{x}_B & + & \mathbf{N}\mathbf{x}_N & = & \mathbf{b} \\ & & & & \mathbf{x}_B, \mathbf{x}_N & \geq & 0 \end{array}$$

Si $\mathbf{x}$ es solución básica factible de (P), entonces $\mathbf{x}_N = \mathbf{0}$ y por lo tanto:

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

3.1.2. Dirección factible

Sea $\mathbf{x} \in \mathcal{P}$. Un vector $\mathbf{d} \in \mathbb{R}^n$ se dice dirección factible en $\mathbf{x}$ si existe un escalar θ para el cual $\mathbf{x}+\theta\mathbf{d} \in \mathcal{P}$. Dada $\mathbf{x}$ una solución básica factible de un problema en forma estándar (P), se considera la posibilidad de moverse de $\mathbf{x}$ a un nuevo vector $\mathbf{x}+\theta\mathbf{d}$ seleccionando una variable no básica x_j inicialmente nula, e incrementándola a un valor positivo θ manteniendo las otras variables no básicas en cero. Al mismo tiempo, $\mathbf{x}_B$ cambia a $\mathbf{x}+\theta\mathbf{d}_B$. Estamos interesados solo en soluciones factibles. Entonces:

$$0 = \mathbf{Ad} = \mathbf{Bd}_B + A_{\cdot j}$$

Por lo tanto, $\mathbf{d}_B = -\mathbf{B}^{-1}A_{\cdot j}$ recibe el nombre de *j-ésima dirección factible*. Como solo se está incrementando la variable no básica x_j, bastará con analizar el comportamiento de las variables básicas. A partir de esto, se distinguen dos casos:

- Si $\mathbf{x}$ es no-degenerado, entonces $\mathbf{x}_B > 0$ con lo que $\mathbf{x}_B + \theta\mathbf{d}_B \geq 0$ y la factibilidad se mantiene para θ lo suficientemente pequeño. En este caso, $\mathbf{d}$ es una dirección factible.

- Si $\mathbf{x}$ es degenerado, entonces es posible que una de las variables básicas $x_{B(i)}$ sea nula y su componente correspondiente de $\mathbf{d}$ sea negativa, por lo que violaría inmediatamente la restricción de no negatividad.

3.1.3. Costos reducidos

Sea $\mathbf{x}$ una solución básica de (P), $\mathbf{B}$ su matriz básica asociada y sean $\mathbf{c}_B, \mathbf{c}_N$ los vectores de costos de las variables básicas y no básicas respectivamente. Entonces, se define el *vector de costos reducidos* $\overline{\mathbf{c}_N^{\mathbf{t}}}$ como:

$$\overline{\mathbf{c}_N^{\mathbf{t}}} = \mathbf{c}_N - \mathbf{c}_B^{\mathbf{t}}\mathbf{B}^{-1}\mathbf{N}$$

Como $\mathbf{x}$ es solución básica, satisface $\mathbf{Bx}_B + \mathbf{Nx}_N = \mathbf{b}$. Despejando $\mathbf{x}_B$ y reemplazándolo en la función objetivo:

$$\mathbf{c}_B^{\mathbf{t}}\mathbf{x}_B + \mathbf{c}_N^{\mathbf{t}}\mathbf{x}_N = \mathbf{c}_B^{\mathbf{t}}\mathbf{B}^{-1}\mathbf{b} + \underbrace{\left(\mathbf{c}_N^{\mathbf{t}} - \mathbf{c}_B^{\mathbf{t}}\mathbf{B}^{-1}\mathbf{N}\right)}_{\overline{\mathbf{c}_N^{\mathbf{t}}}}\mathbf{x}_N$$

Dado que $\mathbf{c}_B^{\mathbf{t}}\mathbf{B}^{-1}\mathbf{b}$ es una constante, el problema de minimizar $\mathbf{c}_B^{\mathbf{t}}\mathbf{x}_B + \mathbf{c}_N^{\mathbf{t}}\mathbf{x}_N$ se transforma en analizar el vector de costos reducidos.

- Si $\overline{\mathbf{c}_N^{\mathbf{t}}} \geq \mathbf{0}$, entonces el óptimo se alcanza para $\mathbf{x}_N = \mathbf{0}$ y la base $\mathbf{B}$ es óptima.

- Si al menos una componente j de $\overline{\mathbf{c}_N^{\mathbf{t}}}$ es negativa, la j-ésima dirección básica $\mathbf{d}$ es una dirección factible de reducción de costo, y por lo tanto la base actual no es óptima.

3.2. Desarrollo del método Simplex

3.2.1. Entrada y salida a la base

Sea $\mathbf{x}$ solución básica factible no degenerada. Si $\mathbf{x}$ tiene el j-ésimo costo reducido negativo, es conveniente moverse en la j-ésima dirección. En particular, conviene moverse lo máximo posible manteniendo la factibilidad. Se escoge:

$$\theta^* = \max\{\theta \geq 0 \mid \mathbf{x} + \theta\mathbf{d} \in \mathcal{P}\}$$

La única forma en que se viola la factibilidad del problema al moverse en una dirección factible $\mathbf{d}$ es haciendo negativa alguna de las componentes del nuevo punto.

- Si $\mathbf{d} \geq \mathbf{0}$, entonces $\mathbf{x} + \theta \mathbf{d} \geq \mathbf{0}$ nunca se vuelve infactible. En este caso, $\theta^* = \infty$ (problema no acotado)

- Si $d_i < 0$ para algún i, entonces la restricción $x_i + \theta d_i \geq 0$ se transforma en $\theta \leq -\dfrac{x_i}{d_i}$ Esta restricción debe satisfacerse para cada i tal que $d_i < 0$, por lo que se escoge:

$$\theta^* = \min_{i|d_i<0} \left\{ -\frac{x_i}{d_i} \right\}$$

  En este caso, la variable j-ésima deja de ser nula y vale θ^*, mientras que la variable i-ésima ahora es nula. De esta forma, x_j reemplaza a x_i en la base. Es decir, x_j *entra a la base* y x_i *sale de la base*. Se construye entonces una nueva matriz básica $\overline{\mathbf{B}}$ en la que se reemplaza la columna asociada a x_i por la columna asociada a x_j. El punto $\mathbf{y} = \mathbf{x} + \theta^* \mathbf{d}$ es solución básica factible asociada a la matriz $\overline{\mathbf{B}}$.

3.2.2. Iteración del método Simplex

1. Se elige una base $A_{\cdot B(1)}, ..., A_{\cdot B(m)}$ y una solución básica factible asociada no degenerada $\mathbf{x}$.
2. Calcular los costos reducidos para las variables no-básicas ($j \in \mathbf{N}$):

$$\bar{c}_j = c_j - \mathbf{c}_B^{\mathbf{t}} \mathbf{B}^{-1} A_{\cdot j}$$

   - Si todos son no negativos, entonces la base actual es óptima y el algoritmo termina.
   - Si no, escoger algún j tal que $\bar{c}_j = 0$.

3. Calcular la j-ésima dirección factible $\mathbf{d} = -\mathbf{B}^{-1} A_{\cdot j}$. Si ninguna de sus componentes es negativa, entonces el problema es no acotado (el óptimo es $-\infty$) y el algoritmo termina.
4. Si alguna componente de $\mathbf{d}$ es negativa, encontrar:

$$\theta^* = \min_{i|d_i<0} \left\{ -\frac{x_i}{d_i} \right\}$$

5. Sea k tal que $\theta^* = -\dfrac{x_k}{d_k}$. Se procede a formar una nueva matriz base reemplazando la columna $A_{\cdot k}$ por la columna $A_{\cdot j}$. Las componentes de la nueva solución básica factible $\mathbf{y}$ están dadas por:

$$\begin{aligned} y_j &= \theta^* \\ y_{B(i)} &= x_{B(i)} + \theta^* d_{B(i)} \end{aligned}$$

   Asumiendo que el conjunto factible es no vacío y que todas las soluciones básicas factibles son no degeneradas, entonces el método Simplex termina en un número finito de iteraciones y entrega una de las siguientes posibilidades:

   - Una base $\mathbf{B}$ óptima y su correspondiente solución básica factible $\mathbf{x}^*$ óptima.
   - Una dirección $\mathbf{d}$ de crecimiento factible e infinito. Luego, el problema es no acotado y el costo óptimo es $-\infty$.

3.2.3. Simplex para problemas degenerados

Hasta ahora, se ha asumido que las soluciones básicas factibles son no degeneradas. Para el caso degenerado, hay dos casos a estudiar:

a) Si la actual solución básica factible $\mathbf{x}$ es degenerada, θ^* puede ser igual a cero, y en ese caso, la nueva solución básica factible $\mathbf{y}$ es igual a $\mathbf{x}$. Esto ocurre si alguna variable básica $x_{B(l)}$ es nula y la correspondiente componente $d_{B(l)}$ del vector dirección es negativa. Sin embargo, se puede definir una nueva matriz básica $\overline{\mathbf{B}}$, reemplazando $A_{\cdot l}$ por $A_{\cdot j}$.

b) Incluso si $\theta^* > 0$, puede pasar que más de una de las variables básicas originales se vuelva cero en el nuevo punto $\mathbf{x} + \theta^*$. Dado que una de ellas entra a la base, las otras se mantienen en la base con valor cero, y la nueva solución básica factible es cero. Para evitar caer en un loop, se deben seguir las *reglas de pivoteo*:

   - Escoger la columna con el costo reducido más negativo.
   - Escoger la columna con costo reducido negativo para la cual el correspondiente decrecimiento de costo $\theta^* |\bar{c}_j|$ es mayor.
   - Escoger la columna con costo reducido negativo que tenga el menor índice.

3.3. Implementaciones del método Simplex

3.3.1. Simplex revisado

1. Se elige una base $A_{\cdot B(1)}, ..., A_{\cdot B(m)}$ y una solución básica factible asociada $\mathbf{x}$.

2. Calcular el vector $\mathbf{p}' = \mathbf{c}_B^{\mathbf{t}}\mathbf{B}^{-1}$ y calcular los costos reducidos $\bar{c}_j = c_j - \mathbf{p}'\mathbf{A}_{\cdot j}$. Si son todos no negativos, la solución básica factible actual es óptima y el algoritmo termina. Si no, elegir un j tal que $\bar{c}_j < 0$.

3. Calcular el vector $\mathbf{u} = -\mathbf{d} = \mathbf{B}^{-1}A_{\cdot j}$. Si ninguna componente de $\mathbf{u}$ es positiva, entonces el problema es no acotado y el algoritmo termina.

4. Si hay componentes de $\mathbf{u}$ positivas, sea

$$\theta^* = \min_{i=1,...,m|u_i>0} \left\{ \frac{x_{B(i)}}{u_i} \right\}$$

5. Sea l tal que $\theta^* = x_{B(l)}/u_l$. Se forma una nueva base reemplazando la columna $A_{\cdot l}$ por la columna $A_{\cdot j}$. Si $\mathbf{y}$, los valores de las nuevas variables básicas son $y_j = \theta^*$ e $y_{B(i)} = x_{B(i)} - \theta^* u_i$ para $i \neq l$.

6. Formar la matriz de $m \times (m+1)$ $[\mathbf{B}^{-1} \mid \mathbf{u}]$. Pivotear hasta que la última columna se convierta en el l-ésimo vector canónico $\mathbf{e}_l$. Las primeras m columnas del resultado, forman la matriz $\overline{\mathbf{B}}^{-1}$.

3.3.2. Método full tableau

Consiste en construir una tabla (*tableau*) a partir de coeficientes conocidos entregados por una solución básica factible y ejecutar implícitamente el método Simplex. La estructura del tableau es:

$-\mathbf{c}_B^{\mathbf{t}}\mathbf{B}^{-1}\mathbf{b}$	$\mathbf{c}' - \mathbf{c}_B\mathbf{B}^{-1}\mathbf{A}$
$\mathbf{B}^{-1}\mathbf{b}$	$\mathbf{B}^{-1}\mathbf{A}$

Una iteración típica de full tableau consiste en:

1. Comenzar con el tableau asociado con una matriz base $\mathbf{B}$ y su correspondiente solución básica factible $\mathbf{x}$.

2. Examinar los costos reducidos (ubicados en la llamada *fila cero*). Si todos son no negativos, la actual solución básica factible es óptima y el algoritmo termina. Si no, elegir algún j para el cual $\bar{c}_j < 0$

3. Considere el vector $\mathbf{u} = \mathbf{B}^{-1}\mathbf{A}$ que corresponde a la j-ésima columna (la *columna pivote*) del tableau. Si ninguna componente de $\mathbf{u}$ es positiva, entonces el costo óptimo es $-\infty$ y el algoritmo termina.

4. Para cada i tal que $u_i > 0$, calcular $x_{B(i)}/u_i$. Sea l el índice de la fila que corresponde al mínimo del cuociente. La columna $A_{\cdot B(l)}$ sale de la base y la columna $A_{\cdot j}$ entra a la base.

5. Pivotear sobre la columna de pivoteo, de tal forma de convertirla en un vector canónico normalizando la componente u_l.

3.3.3. Simplex fase I

En determinados problemas, puede ser difícil encontrar una solución básica factible para iniciar el método Simplex. Para encontrar una solución básica factible, se considera el problema auxiliar

$$\begin{array}{llrcl}
\text{(P Aux)} & \text{mín} & \displaystyle\sum_{i=1}^{m} y_i & & \\
& \text{s.a:} & \mathbf{Ax}+\mathbf{y} & = & \mathbf{b} \\
& & \mathbf{x},\mathbf{y},\mathbf{b} & \geq & 0
\end{array}$$

En este nuevo problema, $\mathbf{x} = \mathbf{0}$ e $\mathbf{y} = \mathbf{b}$ es solución básica factible y la matriz de base es la identidad de $m \times m$. Si la solución óptima de este problema es $\mathbf{y} = \mathbf{0}$, entonces la base óptima del problema auxiliar corresponde a una base factible del problema original. En caso contrario, el problema original es infactible.

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

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

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

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

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

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

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:

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

4.4. Variables duales como costos marginales

Dado el problema en forma estándar

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

suponiendo que existe una solución básica factible óptima no degenerada de $\mathbf{x}$. Luego:

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

Si se añade una perturbación pequeña $\mathbf{d}$ en $\mathbf{b}$:

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

para algún valor de $\mathbf{d}$. Como el vector $\mathbf{b}$ no tiene relación con el vector de costos reducidos, significa que la base actual del problema es también una base factible del problema perturbado. Por dualidad fuerte, se sabe que:

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

La solución óptima del dual equivale a $\mathbf{y}^* = \mathbf{c}_B\mathbf{B}^{-1}$. Entonces, el costo óptimo después de la perturbación cambia a:

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

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 $\mathbf{b} \in \mathbb{R}^n$. Entonces, exactamente una de las siguientes alternativas se cumple:

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

(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 $\mathbf{A}_1, ..., \mathbf{A}_n$ y $\mathbf{b}$ vectores dados, y suponga que cualquier vector $\mathbf{y}$ que satisfaga $y'\mathbf{A}_i \geq 0$, $i = 1, ..., n$, también debe satisfacer $\mathbf{y}'\mathbf{b} \geq 0$. Entonces, $\mathbf{b}$ puede ser expresado como una combinación lineal no negativa de los vectores $\mathbf{A}_1, ..., \mathbf{A}_n$.

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 $\mathbf{x} \in C$.

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) El vector cero es un punto extremo de C.

(b) El cono 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:

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

(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 $\mathcal{P}$.

4.9. Representación de poliedros

Teorema 4.9. 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 $C = \{\mathbf{x} | \mathbf{A}\mathbf{x} \geq \mathbf{0}\}$ contiene al $\mathbf{0}$. Entonces, cada elemento de C puede ser expresado como una combinación lineal no negativa de los rayos extremos de C.

5. Análisis de sensibilidad

5.1. Análisis de sensibilidad local

Dado un problema de programación lineal con su base óptima $\mathbf{B}$ y su respectiva solución óptima $\mathbf{x}^*$. Se analizarán casos en que se modifica algún valor de $\mathbf{A}, \mathbf{b}$ o $\mathbf{c}$, se agrega una nueva restricción o se agrega una nueva variable. En primera instancia se buscan condiciones bajo las cuales la base actual siga siendo óptima. Si estas condiciones son violadas, se buscará un algoritmo que encuentre una nueva solución óptima sin tener que resolver el problema desde cero.
Como $\mathbf{B}$ es una base óptima para el problema original, cumple la *condición de factibilidad* $\mathbf{B}^{-1}\mathbf{b} \geq \mathbf{0}$ y la de optimalidad $\mathbf{c}' - \mathbf{c}_B^{\mathbf{t}}\mathbf{B}^{-1}\mathbf{A} \geq \mathbf{0}$. Cuando el problema cambia, se busca estudiar cómo estas condiciones son afectadas. Imponiendo que las condiciones de factibilidad y optimalidad se mantengan para el problema modificado, se obtienen condiciones bajo las cuales $\mathbf{B}$ sigue siendo óptima.

5.1.1. Adición de una nueva variable

Suponga que se añade una nueva variable x_{n+1} con su respectiva columna $A_{\cdot n+1}$. Se obtiene el nuevo problema:

$$\begin{array}{lcclcl} \text{mín} & \mathbf{c}'\mathbf{x} & + & c_{n+1}x_{n+1} & & \\ \text{s.a:} & \mathbf{A}\mathbf{x} & + & A_{\cdot n+1}x_{n+1} & = & \mathbf{b} \\ & & & \mathbf{x} & \geq & \mathbf{0} \end{array}$$

Es claro que $(\mathbf{x}, x_{n+1}) = (\mathbf{x}^*, 0)$ es solución básica factible asociada a $\mathbf{B}$. Para mantener la optimalidad de $\mathbf{B}$, es necesario y suficiente que el costo reducido de x_{n+1} sea no negativo

$$\overline{c}_{n+1} = c_{n+1} - \mathbf{c}_B^{\mathbf{t}}\mathbf{B}^{-1}A_{\cdot n+1} \geq \mathbf{0}$$

Si se satisface esta condición, $(\mathbf{x}^*, 0)$ es solución óptima del nuevo problema. Si no, se puede añadir una columna en el tableau, asociada a la nueva variable, y aplicar el algoritmo Simplex primal comenzando desde la base actual $\mathbf{B}$.

5.1.2. Adición de una restricción de desigualdad

Se introduce una nueva restricción $\mathbf{a}_{m+1}^{\mathbf{t}}\mathbf{x} \geq b_{m+1}$ donde $\mathbf{a}_{m+1}, b_{m+1}$ están dados. Si la solución óptima $\mathbf{x}^*$ del problema original cumple con la restricción, entonces $\mathbf{x}^*$ es solución óptima de este problema también. Si la nueva restricción es violada, se agrega una variable de holgura no negativa x_{n+1} con lo que la nueva restricción se reescribe como $\mathbf{a}_{m+1}^{\mathbf{t}}\mathbf{x} - x_{n+1} = b_{m+1}$. Se obtiene un problema en forma estándar en el cual la matriz $\mathbf{A}$ es reemplazada por:

$$\overline{\mathbf{A}} = \begin{bmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{a}_{m+1}^{\mathbf{t}} & -1 \end{bmatrix}$$

A partir de la base óptima del problema original $\mathbf{B}$, se forma una base para el nuevo problema considerando las variables básicas originales junto a x_{n+1}. La nueva matriz base $\overline{\mathbf{B}}$ es de la forma:

$$\overline{\mathbf{B}} = \begin{bmatrix} \mathbf{B} & \mathbf{0} \\ \mathbf{a}' & -1 \end{bmatrix}$$

donde el vector $\mathbf{a}'$ contiene a aquellas componentes de $\mathbf{a}_{m+1}^{\mathbf{t}}$ asociadas con las columnas básicas originales. La solución básica asociada a esta base es $(\mathbf{x}^*, \mathbf{a}_{m+1}^{\mathbf{t}}\mathbf{x}^* - b_{m+1})$ y es infactible porque se asumió que $\mathbf{x}^*$ violaba la nueva restricción. La inversa de $\overline{\mathbf{B}}$ está dada por:

$$\overline{\mathbf{B}}^{-1} = \begin{bmatrix} \mathbf{B}^{-1} & \mathbf{0} \\ \mathbf{a}' & -1 \end{bmatrix}$$

Sea $\mathbf{c}_B$ el vector de costos de las m variables del problema original. Entonces, el vector de costos reducidos asociado con la base $\overline{\mathbf{B}}$ del nuevo problema está dado por:

$$\begin{bmatrix} \mathbf{c}' & 0 \end{bmatrix} - \begin{bmatrix} \mathbf{c}_B^{\mathbf{t}} & 0 \end{bmatrix} \overline{\mathbf{B}}^{-1}\overline{\mathbf{A}} = \begin{bmatrix} \mathbf{c}' - \mathbf{c}_B^{\mathbf{t}}\mathbf{B}^{-1}\mathbf{A} & 0 \end{bmatrix}$$

que, dada la optimalidad de $\mathbf{B}$ en el problema original, es no negativo. Se tiene entonces que $\overline{\mathbf{B}}$ es una base dual factible y por lo tanto, se puede aplicar el método Simplex dual. Nótese además que:

$$\overline{\mathbf{B}}^{-1}\overline{\mathbf{A}} = \begin{bmatrix} \mathbf{B}^{-1}\mathbf{A} & \mathbf{0} \\ \mathbf{a}'\mathbf{B}^{-1}\mathbf{A} - \mathbf{a}_{m+1}^{\mathbf{t}} & 1 \end{bmatrix}$$

donde $\mathbf{B}^{-1}\mathbf{A}$ se tiene del tableau final Simplex del problema original.

5.1.3. Cambios en el vector b

Suponga que una componente b_i de $\mathbf{b}$ cambia a $b_i + \delta$. Se desea determinar el rango de valores de δ para el cual, la base actual siga siendo óptima. Las condiciones de optimalidad no varían al variar $\mathbf{b}$. Entonces, basta con revisar la condición de factibilidad:

$$\mathbf{B}^{-1}(\mathbf{b} + \delta\mathbf{e}_i)$$

Sea $\mathbf{g} = (\beta_{1i}, \beta_{2i}, ..., \beta_{mi})$ la i-ésima columna de $\mathbf{B}^{-1}$. Entonces:

$$x_{B(j)} + \delta\beta_{ji} \geq 0, \quad j = 1, ..., m$$

Por lo tanto:

$$\max_{\{j|\beta_{ji}>0\}} \left(-\frac{x_{B(j)}}{\beta_{ji}}\right) \leq \delta \leq \min_{\{j|\beta_{ji}<0\}} \left(-\frac{x_{B(j)}}{\beta_{ji}}\right)$$

Para δ en este rango, el costo óptimo como función de δ está dado por $\mathbf{c}_B^{\mathbf{t}}\mathbf{B}^{-1}(\mathbf{b} + \delta\mathbf{e}_i) = \mathbf{y}'\mathbf{b} + \delta y_i$ donde $\mathbf{y}' = \mathbf{c}_B^{\mathbf{t}}\mathbf{B}^{-1}$ es la solución dual (óptima) asociada a la base actual $\mathbf{B}$.

5.1.4. Cambios en el vector c

Suponga que algún coeficiente de costo c_j se convierte en $c_j + \delta$. La condición primal de factibilidad no se ve afectada. Por lo tanto, bastará enfocarse en la condición de optimalidad:

$$\mathbf{c}_B^{\mathbf{t}}\mathbf{B}^{-1}\mathbf{A} \leq \mathbf{c}'$$

Si c_j es el coeficiente de costo de una variable no básica x_j, se necesita que:

$$\mathbf{c}_B^{\mathbf{t}}\mathbf{B}^{-1}A_{\cdot j} \leq c_j + \delta$$

o equivalentemente, $\delta \geq -\bar{c}_j$. Si esta condición se mantiene, la base actual sigue siendo óptima; de otra forma, se puede aplicar el método Simplex primal comenzando por la actual solución básica factible.
Si c_j es el coeficiente de costo de la l-ésima variable básica ($j = B(l)$), entonces $\mathbf{c}_B$ se convierte en $\mathbf{c}_B + \delta\mathbf{e}_l$ y todas las condiciones de optimalidad se ven afectadas. Las condiciones de optimalidad para el nuevo problema son:

$$(\mathbf{c}_B + \delta\mathbf{e}_l)^{\mathbf{t}}\mathbf{B}^{-1}A_{\cdot i} \leq \bar{c}_i \quad \forall i \neq j$$

dado que x_j es una variable básica y por lo tanto, su costo reducido se mantiene en cero y no necesita ser examinada. Equivalentemente, $\delta q_{li} \leq \bar{c}_j$, donde q_{li} es la l-ésima entrada de $\mathbf{B}^{-1}A_{\cdot i}$, que puede ser obtenida del tableau.

6. Programación entera

6.1. Métodos de plano cortante

Considere el problema de programación entera:

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

y su relajación lineal:

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

A continuación, se enuncian algunos *métodos de plano cortante* para resolver problemas de programación entera.

6.1.1. Algoritmo genérico de plano cortante

Una iteración clásica del algoritmo genérico de plano cortante procede como sigue:

1. Resolver la relajación lineal. Sea $\mathbf{x}^*$ una solución óptima.

2. Si $\mathbf{x}$ es entero, el algoritmo termina.

3. Si no, añadir una restricción de desigualdad lineal al problema relajado que todas las soluciones enteras al problema de programación entera satisfagan, pero que $\mathbf{x}^*$ no, y repetir el paso 1.

6.1.2. Algoritmo de plano cortante Gomory

Sea $\mathbf{x}^* = (\mathbf{x}_B, \mathbf{x}_N)$ solución básica factible óptima de la relajación lineal con su base asociada $\mathbf{B}$. Sea $\mathcal{N}$ el conjunto de índices no básicos y sea $\mathbf{N}$ la sub matriz de $\mathbf{A}$ con columnas $A_{\cdot i}$ con $i \in \mathcal{N}$. Del tableau óptimo, se obtienen los coeficientes de las restricciones:

$$\mathbf{x}_B + \mathbf{B}^{-1}\mathbf{N}\mathbf{x}_N = \mathbf{B}^{-1}\mathbf{b}$$

Sea $\overline{a}_{ij} = (\mathbf{B}^{-1}A_{\cdot j})_i$ y $\overline{a}_{i0} = (\mathbf{B}^{-1}\mathbf{b})_i$. Considere una igualdad del tableau óptimo en que $\overline{a}_{i0}$ sea fraccionario:

$$x_i + \sum_{j \in \mathcal{N}} \overline{a}_{ij} x_j = \overline{a}_{i0}$$

Dado que $x_j \geq 0$ para todo j, se tiene que:

$$x_i + \sum_{j \in \mathcal{N}} \lfloor \overline{a}_{ij} \rfloor x_j \leq x_i + \sum_{j \in \mathcal{N}} \overline{a}_{ij} x_j = \overline{a}_{i0}$$

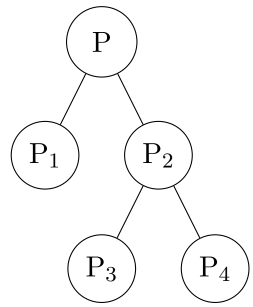
Dado que x_j debe ser entero, se obtiene que:

$$x_i + \sum_{j \in \mathcal{N}} \lfloor \overline{a}_{ij} \rfloor x_j \leq \lfloor \overline{a}_{i0} \rfloor$$

Esta desigualdad es válida para todas las soluciones enteras, pero no es cumplida por $\mathbf{x}$. La razón es que $x_i^* = \overline{a}_{i0}$, $x_j^* = 0$ para todo índice $j \in \mathcal{N}$ y $\lfloor \overline{a}_{i0} \rfloor < \overline{a}_{i0}$ pues $\overline{a}_{i0}$ se asumió fraccionario.

6.2. Branch & Bound

El método *Branch* & *Bound* resuelve problemas de programación lineal entera resolviendo una secuencia ordenada de problemas de programación lineal, que se obtienen relajando las restricciones de integridad y añadiendo restricciones adicionales. Estas restricciones permiten separar la región factible en subregiones complementarias



El algoritmo entonces opera como sigue:

1. Se establece una cota superior (∞) y una inferior ($-\infty$) de la solución óptima. Se resuelve el problema original relajando las restricciones de integridad.

   - Si el problema relajado es infactible, el original también lo es.
   - Si la solución satisface la condición de integridad, es óptima.
   - En cualquier otro caso, se actualiza el valor de la cota inferior con el valor de la función objetivo del problema relajado.

2. Sea x_k variable que debiese ser entera pero que no lo es en la solución óptima del problema relajado. Es claro que:

   $$\lfloor x_k \rfloor \leq x_k \leq \lceil x_k \rceil$$

   por lo que los problemas fruto de la ramificación serían:

   a) El problema original con relajación de integridad y la nueva restricción $x_k \leq \lfloor x_k \rfloor$.

   b) El problema original con relajación de integridad y la nueva restricción $x_k \geq \lceil x_k \rceil$.

   Estos problemas se colocan ordenadamente en una lista de problemas a procesar y se resuelven secuencialmente o en paralelo.

3. Se resuelve el problema siguiente en la lista de problemas a procesar.

4. Se procede a actualizar las cotas:

   - Si la solución del problema actual satisface las condiciones de integridad y el valor de la función objetivo es menor que la cota superior actual, la cota superior se actualiza al valor objetivo de la función objetivo del problema resuelto.
   - Si la solución obtenida no satisface las condiciones de integridad y el valor de la correspondiente función objetivo se encuentra entre la cota inferior y superior, se actualiza en valor de la cota inferior al valor de la función objetivo del problema resuelto y se procede a ramificar. Los problemas generados se añaden a la lista de problemas a procesar.

5. Se verifica si existen ramas que deben *podarse*:

   - Si la solución del problema actual cumple con la condición de integridad, no da lugar a ramificaciones adicionales con esa solución, por lo tanto, se dice que esta rama se *poda por razones de integridad.*
   - Si la solución no satisface la condición de integridad y el valor de la función objetivo del problema resuelto es mayor que la cota superior, no es posible obtener soluciones mediante ramificaciones adicionales, y por lo tanto, se *poda por cotas.*
   - Si el problema es infactible, se *poda por infactibilidad.*

6. Si la lista de problemas a procesar no está vacía, se continúa con el paso 3. Si está vacía, se concluye. En caso de concluir, si existe un candidato a minimizador, este es efectivamente el minimizador; si no existe, el problema es infactible.

El proceso de ramificación termina por alguna de las razones posibles.

- El problema es considerado infactible.
- La solución obtenida satisface las condiciones de integridad.
- La cota inferior obtenida es superior a la cota superior disponible.

Por lo tanto, la rama disponible se poda por infactibilidad, optimalidad o por cotas.

6.3. Programación dinámica

La técnica de programación dinámica resuelve problemas de programación entera secuencialmente. Para construir un algoritmo de programación dinámica, se procede como sigue:

1. Ver la elección de una solución factible como una secuencia de decisiones que se producen en etapas, y de modo que el costo total es la suma de los costos de las decisiones individuales.
2. Definir el estado como una suma de todas las decisiones anteriores pertinentes.
3. Determinar cuáles transiciones de estado son posible. Dejar que el costo de cada transición sea el costo de la decisión correspondiente.
4. Escribir una recursión en el costo óptimo del estado de origen a un estado de destino.

7. Optimalidad y dualidad en programación no lineal

7.1. Definiciones básicas

7.1.1. Problema de programación no lineal

Un problema de programación no lineal tiene la forma:

$$\begin{array}{lrcl} \text{mín} & f(x_1,...,x_n) & & \\ \text{s.a:} & h_1(x_1,...,x_n) & = & 0 \\ & \vdots & \vdots & \vdots \\ & h_l(x_1,...,x_n) & = & 0 \\ & g_1(x_1,...,x_n) & \leq & 0 \\ & \vdots & \vdots & \vdots \\ & g_m(x_1,...,x_n) & \leq & 0 \end{array}$$

Cualquier vector $\mathbf{x} = (x_1,...,x_n)$ que satisface las restricciones, se denomina, al igual que en el caso lineal, *solución factible*, y el conjunto de todas las soluciones factibles se denomina *región factible*.

7.1.2. Teorema de Weirestrass

Sea $S \neq \varnothing \subseteq \mathbb{R}^n$ compacto y $f : S \to \mathbb{R}$ una función continua. El problema:

$$\begin{array}{ll} \text{mín} & f(\mathbf{x}) \\ \text{s.a:} & \mathbf{x} \in S \end{array}$$

posee al menos una solución óptima.

7.1.3. Mínimo global

Una función $f(\mathbf{x})$ tiene un *mínimo global* $\mathbf{x}^*$ en S si y sólo si $f(\mathbf{x}^*) \leq f(\mathbf{x})\ \forall \mathbf{x} \in S$.

7.1.4. Mínimo local

Una función $f(\mathbf{x})$ tiene un *mínimo local* $\mathbf{x}^*$ en S si y sólo si existe $\varepsilon > 0$ tal que $f(\mathbf{x}^*) \leq f(\mathbf{x})$ si $0 < \|\mathbf{x} - \mathbf{x}^*\| \leq \varepsilon$.

7.1.5. Diferenciabilidad

Se dice que $f : \mathbb{R}^n \to \mathbb{R}$ es *diferenciable* en $\mathbf{x}$ si las derivadas parciales $\dfrac{\partial f(\mathbf{x})}{\partial x_i}$ para $i = 1,...,n$ existen y:

$$\lim_{\mathbf{y} \to \mathbf{x}} \frac{f(\mathbf{y}) - f(\mathbf{x}) - \nabla^{\mathbf{t}} f(\mathbf{x})(\mathbf{y} - \mathbf{x})}{\|\mathbf{y} - \mathbf{x}\|} = 0$$

donde:

$$\nabla f(\mathbf{x}) = \begin{bmatrix} \dfrac{\partial f(\mathbf{x})}{\partial x_1} \\ \vdots \\ \dfrac{\partial f(\mathbf{x})}{\partial x_n} \end{bmatrix}$$

7.1.6. Función continuamente diferenciable

Una función f se dice continuamente diferenciable en $\overline{\mathbf{x}}$ si todas sus derivadas parciales son continuas en $\overline{\mathbf{x}}$.

7.1.7. Funciones dos veces diferenciables

Se dice que $f : \mathbb{R}^n \to \mathbb{R}$ es dos veces diferenciable en $\mathbf{x}$ si existe un vector columna $\nabla f(\mathbf{x})$ y una matriz de $n \times n$ $\nabla^2 f(\mathbf{x})$ que cumple:

$$\lim_{\mathbf{y}\to\mathbf{x}} \frac{f(\mathbf{y}) - f(\mathbf{x}) - \nabla^{\mathbf{t}} f(\mathbf{x})(\mathbf{y}-\mathbf{x}) - \frac{1}{2}(\mathbf{y}-\mathbf{x})\nabla^2 f(\mathbf{x})(\mathbf{y}-\mathbf{x})}{\|\mathbf{y}-\mathbf{x}\|} = 0$$

La matriz $\nabla^2 f(\mathbf{x})$ recibe el nombre de *matriz hessiana* de f en el punto $\mathbf{x}$. El elemento ij de $\nabla^2 f(\mathbf{x})$ es la segunda derivada parcial $\dfrac{\partial^2 f(\mathbf{x})}{\partial x_i \partial x_j}$. Además, si estas segundas derivadas parciales son continuas, entonces f se denomina *dos veces continuamente diferenciable* y en este caso, $\nabla^2 f(\mathbf{x})$ es una *matriz simétrica*.

7.1.8. Matriz semidefinida positiva

Sea $\mathbf{A} \in \mathcal{M}_{n\times n}$ una matriz simétrica. Se dice que $\mathbf{A}$ es *semidefinida positiva* si $\forall \mathbf{x} \in \mathbb{R}^n$:

$$\mathbf{x}'\mathbf{A}\mathbf{x} \geq 0$$

Además, si se tiene la igualdad única y exclusivamente con el vector $\mathbf{x} = \mathbf{0}$, entonces $\mathbf{A}$ se dice *definida positiva*.
Otra forma de mostrar que $\mathbf{A}$ es semidefinida positiva, es verificando que sus valores propios son mayores o iguales que cero. En el caso de que sus valores propios sean mayores estrictos que cero, se dice que $\mathbf{A}$ es definida positiva.

7.2. Convexidad y optimización no lineal

7.2.1. Función convexa

Sea $f : S \neq \varnothing \subseteq \mathbb{R}^n \to \mathbb{R}$. Se dice que f es convexa en S si $\forall \mathbf{x}, \mathbf{y} \in S$, y $\forall \lambda \in [0, 1]$ se tiene que:

$$f(\lambda\mathbf{x} + (1-\lambda)\mathbf{y}) \leq \lambda f(\mathbf{x}) + (1-\lambda) f(\mathbf{y})$$

7.2.2. Mínimos locales y globales

Sea $f : S \neq \varnothing \subseteq \mathbb{R}^n \to \mathbb{R}$. Todo *mínimo local* de f en S es también un *mínimo global* de f en S. Además, si la función es *estrictamente convexa*, es decir, $f(\lambda\mathbf{x}+(1-\lambda)\mathbf{y}) < \lambda f(\mathbf{x})+(1-\lambda)f(\mathbf{y})$ para todo $\mathbf{x}, \mathbf{y} \in S$, $\lambda \in [0, 1]$, entonces ese mínimo es único.

7.2.3. Diferenciabilidad de primer orden y convexidad

Sea $S \neq \varnothing \subseteq \mathbb{R}^n$ un conjunto convexo y sea $f : \mathbb{R}^n \to \mathbb{R}$ función diferenciable en S. Entonces, f es una *función convexa* en S si y sólo si:

$$f(\mathbf{x}) - f(\mathbf{y}) \leq \nabla^{\mathbf{t}} f(\mathbf{x})(\mathbf{y} - \mathbf{x})$$

para cualquier $\mathbf{x}, \mathbf{y} \in S$. Si la desigualdad se cumple de manera estricta para cualquier $\mathbf{x}, \mathbf{y} \in S$ tales que $\mathbf{x} \neq \mathbf{y}$, entonces se dice que f es *estrictamente convexa* en S.

7.2.4. Diferenciabilidad de segundo orden y convexidad

Sea $S \neq \varnothing \subseteq \mathbb{R}^n$ convexo abierto, y sea $f : \mathbb{R}^n \to \mathbb{R}$ función dos veces diferenciable en S. Entonces, f es *convexa* en S si y sólo si $\nabla^2 f(\tilde{\mathbf{x}})$ es una matriz *semidefinida positiva* $\forall \tilde{\mathbf{x}} \in S$:

$$\mathbf{y}'\nabla^2 f(\tilde{\mathbf{x}})\mathbf{y} \geq 0 \qquad \forall \mathbf{y} \in \mathbb{R}^n$$

Si la desigualdad es estricta $\forall \mathbf{y} \neq \mathbf{0} \in \mathbb{R}^n$, se dice que es *estrictamente convexa*.

7.2.5. Propiedades adicionales de funciones convexas

Sea $\{f_i(\mathbf{x})\}_{i=1}^k$ familia de funciones convexas, $\{\lambda_i\}_{i=1}^k$ escalares positivos. Se tiene que:

- $h(\mathbf{x}) = \sum_{i=1}^k \lambda_i f_i(\mathbf{x})$ es convexa.
- $h(\mathbf{x}) = \sup_{i=1,\ldots,k}\{f_i(\mathbf{x})\}$ es convexa.
- Si h es convexa y $\mathbf{T}$ es una función lineal (o afín), entonces $(h \circ \mathbf{T})(\mathbf{x})$ es convexa.
- Si h es convexa y g es una función de una variable convexa no decreciente, entonces $(g \circ h)(\mathbf{x})$ es convexa.

7.3. Condiciones de Karush-Kuhn-Tucker

Sea el problema de minimización no lineal:

$$\begin{array}{rrcl}
\text{mín} & f(x_1,...,x_n) & & \\
\text{s.a:} & h_1(x_1,...,x_n) & = & 0 \\
 & \vdots & \vdots & \vdots \\
 & h_l(x_1,...,x_n) & = & 0 \\
 & g_1(x_1,...,x_n) & \leq & 0 \\
 & \vdots & \vdots & \vdots \\
 & g_m(x_1,...,x_n) & \leq & 0
\end{array}$$

Las *condiciones de Karush-Kuhn-Tucker (KKT)* para el problema anterior, están dadas por:

$$\begin{array}{rcll}
\nabla f(\overline{\mathbf{x}}) + \sum_{k=1}^l \lambda_k \nabla h_k(\overline{\mathbf{x}}) + \sum_{j=1}^m \mu_j \nabla g_j(\overline{\mathbf{x}}) & = & \mathbf{0} & \\
h_k(\overline{\mathbf{x}}) & = & 0 & k = 1,...,l \\
g_j(\overline{\mathbf{x}}) & \leq & 0 & j = 1,...,m \\
\mu_j g_j(\overline{\mathbf{x}}) & = & 0 & j = 1,...,m \\
\mu_j & \geq & 0 & j = 1,...,m
\end{array}$$

Los vectores $\boldsymbol{\lambda}$, $\boldsymbol{\mu}$ se conocen como *multiplicadores de Kuhn-Tucker* Las condiciones $h_k(\overline{\mathbf{x}}) = 0$ y $g_j(\overline{\mathbf{x}}) \leq 0$ se denominan *condiciones de factibilidad primal*. La condición $\mu_j g_j(\overline{\mathbf{x}}) = 0$ es conocida como *condición de holgura complementaria*, y la condición $\mu_j \geq 0$ recibe el nombre de *condición de factibilidad dual*.

7.3.1. Condiciones necesarias de Karush-Kuhn-Tucker

Si $\overline{\mathbf{x}}$ es un mínimo local de $f(\mathbf{x})$, entonces satisface las condiciones de KKT. Las condiciones de KKT son *condiciones necesarias* para la mayoría de los problemas de programación no lineal.

7.3.2. Condiciones suficientes de Karush-Kuhn-Tucker

Si $f(\mathbf{x}), \{g_j(\mathbf{x})\}_{j=1}^m$ son funciones convexas, $\{h_i(\mathbf{x})\}_{k=1}^l$ son lineales y $(\overline{\mathbf{x}}, \boldsymbol{\lambda}, \boldsymbol{\mu})$ satisfacen las condiciones de KKT, entonces $\overline{\mathbf{x}}$ es mínimo local de $f(\mathbf{x})$.

7.4. Condiciones de regularidad

Se pueden dar algunos ejemplos en los que la solución óptima de un problema de optimización no tiene asociado un punto de KKT. Se han estudiado ciertas condiciones suficientes que evitan estas "patologías", las cuales reciben el nombre de *condiciones de regularidad.*

7.5. Condición de Slater

Debe existir $\tilde{\mathbf{x}} \in \mathbb{R}^n$ tal que $g_j(\tilde{\mathbf{x}}) < 0 \quad \forall j = 1, ..., m$.

7.5.1. Condición de independencia lineal

Los gradientes de las restricciones activas deben ser linealmente independientes.

7.6. Dualidad en problemas no lineales

7.6.1. Función dual

Sea el problema primal no lineal:

$$\begin{array}{llrcl} \text{(P)} & \text{mín} & f(x_1, ..., x_n) & & \\ & \text{s.a:} & \mathbf{h}(\mathbf{x}) & = & 0 \\ & & \mathbf{g}(\mathbf{x}) & \leq & 0 \end{array}$$

El problema dual requiere la introducción de la *función dual* definida por:

$$\theta(\boldsymbol{\lambda}, \boldsymbol{\mu}) = \min_{\mathbf{x}} \left\{ f(\mathbf{x}) + \boldsymbol{\lambda}^{\mathbf{t}} \mathbf{h}(\mathbf{x}) + \boldsymbol{\mu}^{\mathbf{t}} \mathbf{g}(\mathbf{x}) \right\}$$

Definiendo el *lagrangeano* como:

$$\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = f(\mathbf{x}) + \boldsymbol{\lambda}^{\mathbf{t}} \mathbf{h}(\mathbf{x}) + \boldsymbol{\mu}^{\mathbf{t}} \mathbf{g}(\mathbf{x})$$

la función dual se escribe como $\theta(\boldsymbol{\lambda}, \boldsymbol{\mu}) = \min_{\mathbf{x}} \{\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu})\}$

7.6.2. Problema dual no lineal

Dado un problema no lineal primal (P), se define el *dual no lineal* (D) como:

$$\begin{array}{llrcl} \text{(D)} & \text{máx} & \theta(\boldsymbol{\lambda}, \boldsymbol{\mu}) & & \\ & \text{s.a:} & \boldsymbol{\mu} & \geq & 0 \end{array}$$

7.6.3. Teorema de dualidad débil

Cualquier solución factible $\mathbf{x}$ del problema primal y cualquier solución del problema dual cumplen:

$$f(\mathbf{x}) \geq \theta(\boldsymbol{\lambda}, \boldsymbol{\mu})$$

7.6.4. Teorema de dualidad fuerte

Dado un problemas convexo, $\mathbf{x}^*$ resuelve el problema primal si y sólo si $(\boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)$ resuelve el problema dual y $\boldsymbol{\mu}^* \mathbf{g}(\mathbf{x}^*) = 0$.