

# Resumen Optimización

José Tomás Marquinez - 2012

## Capítulo 1.- Investigación de Operaciones.

**Definición:** La *Investigación de Operaciones* es una disciplina científica que aplica métodos analíticos avanzados para ayudar a tomar mejor decisiones.

La metodología incluye la *Definición del problema*, la *Contrucción del problema*, la *Solución del modelo*, la *Verificación del modelo*, y la *Implementación y Control del modelo*.

## Capítulo 2.- Introducción al Modelamiento.

**Definición:** Un *modelo* es un esquema teórico, generalmente en forma matemática, de un sistema o de una realidad compleja. Es una herramienta que ayuda a la toma de decisiones. Debe ser una simplificación de la realidad.

**Definición:** En particular, los *modelos matemáticos* pretenden optimizar. Existen:

- |              |                        |
|--------------|------------------------|
| ■ Dinámicos: | ■ Estáticos:           |
| ● Continuos. | ● Estocásticos.        |
| ● Discretos. | ● Determinísticos.     |
|              | ○ <i>Lineales</i>      |
|              | ○ <i>No Lineales</i> . |

**Terminología:** Un modelo matemático cuenta con:

- *Variables de Decisión*: Cantidades que se buscan determinar y que inciden en el objetivo.
- *Variables de Estado*: Cantidades que buscan definir el estado del sistema o modelo.
- *Variables Auxiliares*: Cantidades que buscan relacionar variables entre sí, principalmente.
- *Restricciones del problema*: Relaciones entre las variables que limitan sus valores, de la forma  $h_i(\vec{x}) = 0, g_j(\vec{x}) \geq 0$ .
- *Restricciones de las variables*: Limitaciones en los valores que pueden tomar las variables por sí solas, de la forma  $\vec{x} \in \Omega : \Omega \subset \mathbb{R}^n$ . Aquí caen las restricciones de *Naturaleza de las Variables*.
- *Función Objetivo*: Medida para comparar las alternativas posibles. Se busca maximizar o minimizar este objetivo.

## Capítulo 3.- Programación Lineal.

**Definición:** Un *modelo de Programación Lineal* es aquel cuyas variables son continuas, y tanto sus restricciones como su función objetivo son lineales.

Se define el *espacio de soluciones factibles*  $\Omega$ , o simplemente espacio factible, como la intersección de todas las restricciones. Si ésta no incluye ningún punto (si el espacio es vacío), se le dice *espacio infactible*.

Se define la *solución o punto óptimo* como el punto  $\vec{x} \in \Omega$  que entrega el mejor valor de la función objetivo.

Se define un *punto vecino* como aquél punto  $\vec{x}^2$  que comparte todas las restricciones, excepto una, con un punto  $\vec{x}^1$ .

Se da que en los Problemas de Programación Lineal (PPL) continuos existen tres posibilidades:

- No existe solución factible.
- Existe una única solución.
- Existen infinitas soluciones.

**Solución gráfica:** Para determinar la solución gráficamente, es necesario graficar todas las restricciones, determinar el espacio común de la intersección de todas las restricciones, y graficar las curvas de nivel de la función objetivo. En un problema de minimización, la curva de nivel de menor valor dentro del dominio corresponde a la solución óptima.

**Existencia de Soluciones Óptimas:** Sea el siguiente problema de optimización:

$$P) \quad \begin{array}{ll} \text{mín} & f(\vec{x}) \\ \text{s.a.} & \vec{x} \in \Omega \end{array}$$

Se define un *punto extremo* como un punto que es mínimo o máximo, local o global, de una función  $f$  sobre un dominio  $\Omega$ .

Un punto  $\vec{x}$  se define *mínimo global* de  $f$  en  $\Omega$  (y solución óptima del problema  $P$ )), si

$$f(\vec{x}) \leq f(\vec{y}), \forall \vec{y} \in \Omega.$$

Un punto  $\vec{x}$  se define *mínimo local* de  $f$  en  $\Omega$  si existe un  $\epsilon > 0$  tal que

$$f(\vec{x}) \leq f(\vec{y}), \quad \forall \vec{y} : \|\vec{x} - \vec{y}\| \leq \epsilon, \vec{y} \in \Omega$$

Se definen los *mínimos estrictos* si las desigualdades son estrictas.

**Teorema:** El *Teorema de Existencia de Soluciones Óptimas* dice que si  $f$  es continua sobre  $\Omega$ , con  $\Omega$  cerrado y no vacío sobre  $\mathbb{R}^n$ , entonces si  $f(\vec{x}) \rightarrow \infty$  cuando  $\|\vec{x}\| \rightarrow \infty, \vec{x} \in \Omega$ , entonces  $P$ ) admite al menos una solución óptima.

**Teorema:** El *Teorema de de Bolzano-Weierstrass* indica que si  $f$  es continua sobre un dominio  $\Omega$  no vacío y compacto (cerrado y acotado), entonces el problema necesariamente tendrá solución óptima.

**Teorema:** El *Teorema Práctico de Existencia de Soluciones Óptimas de la Optimización Lineal* dice que si el dominio  $\Omega$  es no vacío y cerrado, y la función objetivo está acotada inferiormente (en caso de minimización), entonces  $P$  tiene solución óptima.

**Problemas Equivalentes:** Sean los siguientes problemas de optimización  $P_1$  y  $P_2$ :

$$\begin{array}{ll} P_1) & \begin{array}{l} \text{mín} \quad f(\vec{x}) \\ \text{s.a.} \quad g_i(\vec{x}) \leq 0 \quad i = 1, \dots, m \\ \vec{x} \in \Omega_1 \end{array} \end{array} \quad \begin{array}{ll} P_2) & \begin{array}{l} \text{mín} \quad h(\vec{y}) \\ \text{s.a.} \quad l_j(\vec{y}) \leq 0 \quad j = 1, \dots, r \\ \vec{y} \in \Omega_2 \end{array} \end{array}$$

Se **dicen** problemas equivalentes si la solución óptima de  $P_1$  provee la solución óptima de  $P_2$ , y viceversa.

■ **Equivalencia I:**

$$\begin{array}{ll} P_1) & \begin{array}{l} \text{mín} \quad f(\vec{x}) \\ \text{s.a.} \quad \vec{x} \in \Omega \end{array} \end{array} \sim \begin{array}{ll} P_2) & \begin{array}{l} \text{máx} \quad -f(\vec{x}) \\ \text{s.a.} \quad \vec{x} \in \Omega \end{array} \end{array}$$

■ **Equivalencia II:**

$$\begin{array}{ll} P_1) & \begin{array}{l} \text{mín} \quad f(\vec{x}) \\ \text{s.a.} \quad \vec{x} \in \Omega \end{array} \end{array} \sim \begin{array}{ll} P_2) & \begin{array}{l} \text{mín} \quad \mu \\ \text{s.a.} \quad f(\vec{x}) \leq \mu \\ \vec{x} \in \Omega, \mu \in \mathbb{R} \end{array} \end{array}$$

■ **Equivalencia III:**

$$\begin{array}{ll} P_1) & \begin{array}{l} \text{mín} \quad \text{máx}\{f_1(\vec{x}), \dots, f_n(\vec{x})\} \\ \text{s.a.} \quad \vec{x} \in \Omega \end{array} \end{array} \sim \begin{array}{ll} P_2) & \begin{array}{l} \text{mín} \quad \mu \\ \text{s.a.} \quad f_i(\vec{x}) \leq \mu \quad \forall i = 1, \dots, n \\ \vec{x} \in \Omega, \mu \in \mathbb{R} \end{array} \end{array}$$

■ **Equivalencia IV:**

$$\begin{array}{ll} P_1) & \begin{array}{l} \text{mín} \quad \sum_{i=1}^r f_i(\vec{x}) \\ \text{s.a.} \quad \vec{x} \in \Omega \end{array} \end{array} \sim \begin{array}{ll} P_2) & \begin{array}{l} \text{mín} \quad \sum_{i=1}^r \mu_i \\ \text{s.a.} \quad f_i(\vec{x}) \leq \mu_i \quad \forall i = 1, \dots, r \\ \vec{x} \in \Omega, \mu_i \in \mathbb{R} \quad \forall i = 1, \dots, r \end{array} \end{array}$$

■ **Equivalencia V:**

$$\begin{array}{ll} P_1) & \begin{array}{l} \text{mín} \quad f(\vec{x}) \\ \text{s.a.} \quad \vec{x} \in \Omega \end{array} \end{array} \sim \begin{array}{ll} P_2) & \begin{array}{l} \text{mín} \quad g(f(\vec{x})) \\ \text{s.a.} \quad \vec{x} \in \Omega \end{array} \end{array}$$

donde la función  $g : f(\Omega) \subset \mathbb{R} \rightarrow \mathbb{R}$  es estrictamente creciente sobre  $f(\Omega)$ . Notemos que los valores óptimos no serán necesariamente los mismos.

■ Equivalencia VI:

$$P_1) \quad \min_{\vec{x} \in \Omega} \frac{1}{f(\vec{x})} \quad \sim \quad P_2) \quad \max_{\vec{x} \in \Omega} f(\vec{x})$$

con  $f(\vec{x}) > 0, \forall \vec{x} \in \Omega$ .

Nociones Básicas de Convexidad:

**Teorema:** El *Teorema de Convexidad* indica que si, para el problema  $P$ ),  $f$  es una función convexa y  $\Omega$  es un conjunto convexo, entonces si  $\vec{x}$  es un mínimo local de  $f$  en  $\Omega$ ,  $\vec{x}$  es un mínimo global de  $f$  en  $\Omega$ .

Se **define** que un conjunto  $\Omega$  es un *conjunto convexo* si

$$\forall \vec{x}^1, \vec{x}^2 \in \Omega \text{ se tiene } \vec{x} = \lambda \vec{x}^1 + (1 - \lambda) \vec{x}^2 \in \Omega, \forall \lambda \in [0, 1]$$

Notemos que la intersección de conjuntos convexos es convexa, la unión de conjuntos convexos no es convexa, la desigualdad lineal es convexa. Luego, todos los poliedros definidos por desigualdades lineales son convexos.

Se **define** que  $f(\vec{x}) : \Omega \rightarrow \mathbb{R}$ ,  $\Omega$  convexo, es una *función convexa sobre  $\Omega$*  si

$$f(\lambda \vec{x}^1 + (1 - \lambda) \vec{x}^2) \leq \lambda f(\vec{x}^1) + (1 - \lambda) f(\vec{x}^2), \quad \forall \vec{x}^1, \vec{x}^2 \in \Omega, \lambda \in [0, 1].$$

Una función definida sobre un dominio no convexo no puede ser convexa. La suma de funciones convexas es convexa.

La función  $f$  es una *función estrictamente convexa sobre  $\Omega$*  si la desigualdad anterior es estricta.

Se **define** que  $f(\vec{x}) : \Omega \rightarrow \mathbb{R}$ ,  $\Omega$  convexo, es una *función cóncava sobre  $\Omega$*  si

$$f(\lambda \vec{x}^1 + (1 - \lambda) \vec{x}^2) \geq \lambda f(\vec{x}^1) + (1 - \lambda) f(\vec{x}^2), \quad \forall \vec{x}^1, \vec{x}^2 \in \Omega, \lambda \in [0, 1].$$

Luego, si  $f(\vec{x})$  es una función convexa, entonces  $-f(\vec{x})$  es una función cóncava.

La función  $f$  es una *función estrictamente cóncava sobre  $\Omega$*  si la desigualdad anterior es estricta.

Se **define** que el problema de optimización  $P$  es un *problema convexo* si  $\Omega$  es una parte convexa de  $\mathbb{R}^n$  y  $f(\vec{x})$  es convexa sobre  $\Omega$ .

**Teorema:** Se especifica el *Teorema de Convexidad* al indicar que si  $P$  es un problema convexo, entonces si  $\vec{x}$  es un mínimo local del problema, es también su mínimo global. Sin embargo, no necesariamente este óptimo será único.

**Corolario:** Si  $f(\vec{x})$  es estrictamente convexo sobre  $\Omega$  (cerrado), entonces todo punto mínimo local de  $f(\vec{x})$  es también su único mínimo global.

## Capítulo 4.- Método Simplex en Programación Lineal.

En este capítulo y el siguiente se omitirá la notación vectorial con flecha a propósito.

**Notación:** Para iniciar, se considerará un PPL inicial escrito de la siguiente manera matricial:

$$\begin{array}{ll} \text{mín} & c^T x \\ \text{s.a.} & Ax \leq b \end{array}$$

con  $A \in \mathbb{R}^{n \times m}$ ;  $\vec{c}, \vec{x} \in \mathbb{R}^n$ ;  $\vec{b} \in \mathbb{R}^m$ .

**Definición:** Se dice que un PPL está escrito en su *formato estándar* si se escribe de la siguiente manera matricial:

$$\begin{array}{ll} \text{mín} & c^T x \\ \text{s.a.} & Ax = b \\ & x \geq 0 \end{array}$$

con  $A \in \mathbb{R}^{n \times m}$ ;  $\vec{c}, \vec{x} \in \mathbb{R}^n$ ;  $\vec{b} \geq \vec{0} \in \mathbb{R}^m$ .

**Definición:** La *fase II del método Simplex* es un algoritmo que itera viajando de un *punto factible* hacia su mejor vecino, hasta obtener el resultado deseado. Se basa en que si existe exactamente una solución óptima, entonces esta debe ser un vértice, y si existen múltiples, al menos dos de ellas deben ser vértices adyacentes (en problemas acotados). Y que si una solución en un vértice es igual o mejor que todas las soluciones factibles en los vértices adyacentes a ella, entonces es igual o mejor que todas las demás soluciones en los vértices. Luego, es óptima.

El método requiere llevar el problema en su formato estándar. Para ello:

- Si existen variables libres, se debe sustituir por la diferencia de dos variables auxiliares no negativas. Es decir,  $x_1 = \tilde{x}_1 - \tilde{\tilde{x}}_1$ .
- Si hay restricciones de desigualdades, se debe agregar *variables de holgura* para los casos de  $\leq$  o agregar *variables de exceso* (o restar variables de holgura) para los casos de  $\geq$ .
- Si el problema requiere maximizar una función, se debe cambiar al problema equivalente utilizando la Equivalencia I.

Por notación, el número de restricciones es  $m$  y el número de variables totales en este formato es  $n$ .

**Definición:** Una *base* es una submatriz de  $A$  formada por  $m$  columnas linealmente independientes. Permite despejar unas variables en términos de otras para fijar algunas con el fin de resolver un sistema de  $n \times n$ .

**Definición:** Una *solución básica* del sistema  $Ax = b, x \geq 0$  es una solución de la forma  $x = [x_B, x_R]$ , donde  $x_R = 0$ .

Una vez escogida la base, se debe reordenar el problema para que quede escrito de la siguiente manera:

$$\begin{array}{ll} \text{mín} & z = c_B^T \cdot x_B + c_R^T \cdot x_R \\ \text{s.a.} & Bx_B + Rx_R = b \\ & x_R, x_B \geq 0 \end{array}$$

donde

- $x_B \in \mathbb{R}^m$  tiene las coordenadas de  $x$  que están consideradas en la base actual. Es decir, son las *variables básicas*. Por construcción, en la iteración actual  $x_B \geq 0$ .
- $x_R \in \mathbb{R}^{n-m}$  tiene las coordenadas de  $x$  que no están consideradas en la base actual. Es decir, son las *variables no básicas*. Por construcción, en la iteración actual  $x_R = 0$ .
- $B$  es la *base*.
- $R$  es una submatriz de  $A$  formada por  $n - m$  columnas linealmente independientes.

**Definición:** Una solución básica del sistema  $Ax = b, x \geq 0$  se llama *solución básica factible* si  $x_B \geq 0$ . Corresponde a un vértice del dominio.

**Definición:** Se dice que un PPL está escrito en su *formato canónico* si se escribe de la siguiente manera matricial, asociada a una base particular (no necesariamente óptima)  $x^p$ :

$$\begin{array}{ll} \text{mín} & z = (c_B^p)^T \cdot x_B^p + (c_R^p)^T \cdot x_R^p \\ \text{s.a.} & Bx_B^p + Rx_R^p = b \\ & x_R^p, x_B^p \geq 0 \end{array}$$

### Una iteración de Simplex:

- I) Se calculan los siguientes componentes:

$$\bar{R} = B^{-1}R \quad \bar{b} = B^{-1}b$$

- II)  $\bar{b}$  indica el valor actual de las variables básicas, las cuales deben ser positivas por construcción. Por lo tanto, se verifica que se cumpla el *criterio de factibilidad*. En caso de no cumplirse, la base actual no es una base factible:

$$\bar{b} = x_B = B^{-1}b \geq 0$$

Si el valor de una variable básica es igual a cero, existe solución degenerada (punto sobre-determinado).

- III) Se verifica el *criterio de optimalidad*, que indica si la base actual es la base óptima. En caso de que no, se debe continuar con la iteración. El criterio indica que los costos reducidos de las variables no básicas no deben ser negativos (éstos indican un aporte marginal a la función objetivo). Si alguno de los costos reducidos es 0, entonces el problema tiene soluciones múltiples:

$$\bar{c}_R = c_R - c_B B^{-1}R = c_R - c_B \bar{R} \geq 0$$

Si la base actual es la óptima, el valor de las variables es  $x_R = 0, x_B = \bar{b}$ .

- IV) Se verifica el *criterio de entrada*. Entrará a la base aquella variable no básica que tenga costos reducidos menores (entre las que poseen costos reducidos negativos). Si el criterio entrega más de un  $p$ , se escoge arbitrariamente uno para continuar.

$$\min\{\bar{c}_{Ri}\} = \bar{c}_{Rp} \rightarrow x_p \text{ entra a la base.}$$

- V) Se verifica el *criterio de salida*. Saldrá de la base aquella variable que deba hacerse cero para llegar al vértice vecino más cercano. Si el criterio de salida entrega más de un  $s$ , se escoge arbitrariamente. Si el criterio de salida no entrega ningún  $s$ , el problema es no acotado:

$$\min_{\bar{a}_{ip} > 0} \left\{ \frac{\bar{b}}{\bar{a}_{ip}} \right\} = \frac{\bar{b}}{\bar{a}_{sp}} \rightarrow x_s \text{ sale de la base.}$$

- VI) Se actualizan los componentes  $x_B, x_R, B$  y  $R$  y se comienza nuevamente.

Una iteración de Simplex usando la técnica de *tableau*:

- I) Se escribe la siguiente matriz (la primera fila es usada como una guía para el calculista de qué va en qué columna):

$x_B$	$x_R$	$z$
$x_B$	$x_R$	$z$
$B$	$R$	$b$

- II) *Pivotear* la tabla completa de tal manera que en el cuadrante de  $B$  quede la matriz identidad, y los números sobre esta matriz sean ceros. De esa manera, la matriz *pivoteada* quedará de la siguiente manera:

$x_B$	$x_R$	$z$
0	$\bar{c}_R$	$z_{\text{iter}} = -c_B B^{-1}b$
$I$	$R$	$b$

- III) Se verifica el *criterio de optimalidad*. Se verifica que todos los valores en la casilla de  $\bar{c}_R$  sean mayor o igual a cero.
- IV) Se verifica el *criterio de entrada*. Entrará a la base aquella variable no básica que tenga costos reducidos menores (entre las que poseen costos reducidos negativos). Por lo tanto, entra a la base la variable que se encuentra inmediatamente arriba del menor valor de los  $\bar{c}_R$  en la guía para el calculista.
- V) Se verifica el *criterio de salida*. Para ello, se destaca la columna de la variable que entra, y se calcula el cociente de los valores entre la columna de  $z$  y la columna destacada (siempre y cuando estos últimos sean  $> 0$ ). Se debe destacar la fila que contenga el menor cociente entre los calculados. Luego, saldrá aquella variable que se encuentre inmediatamente arriba (en la guía del calculista) del *pivote* (el 1 de la matriz identidad) de la fila seleccionada.

- VI) Se cambia toda la columna de la variable entrante por la de la variable saliente y se comienza de nuevo.

### Fase I de método Simplex:

**Definición:** La *fase I del método Simplex* utiliza el algoritmo de Simplex para encontrar una solución factible del problema original. Para ello, agrega variables auxiliares que hacen factible una solución que no lo es, al agregar dimensiones al problema.

- I) Se agregan *variables auxiliares* en aquellas restricciones que no presentan pivotes para la identidad que genera una solución factible. Generalmente se agrega un variable  $t_j$  por cada restricción  $j$  del problema.
- II) Se escoge la matriz  $B$  del problema como aquella que forme la matriz *identidad*. En el caso de haber agregado una variable por restricción, la base estará conformada por las variables auxiliares.
- III) Se intenta que las variables auxiliares no sean necesarias para estar en una solución factible del problema original, por lo que se modifica la función objetivo a la *suma simple de todas las variables auxiliares* que se hayan agregado.
- IV) Se comienza a iterar Simplex Fase II de este nuevo problema auxiliar hasta que el criterio de optimalidad lo indique.
- V) Si la base actual es tal que el valor óptimo de Fase I es igual a 0, entonces esa es la base que se utiliza para comenzar a iterar Fase II del problema original. Si la base actual es tal que el valor óptimo de Fase I es mayor a 0, entonces no existe solución factible para comenzar a iterar Simplex Fase II del problema original.



## Capítulo 5.- Análisis de Sensibilidad.

**Definición:** Se le dice *Análisis de Sensibilidad* al arte de modificar parámetros del problema para ver cómo se comporta la solución del mismo. Asume una solución óptima  $x^*$  obtenida de una base óptima  $B^*$ .

**Definición:** Los *precios sombra* indican el cambio global en la función objetivo al variar en una unidad el lado derecho de una restricción (una componente del vector  $b$ ).

Variación en los recursos  $b_j$ :

**Definición:** Los *precios sombra* indican el cambio global en la función objetivo al variar en una unidad el lado derecho de una restricción (una componente del vector  $b$ ).

$$\pi^T = \frac{\Delta z}{\Delta b} = c_B^T \cdot [B^*]^{-1}$$

Notemos que para el cálculo de  $\pi^T$  se requiere que no cambie  $B^*$ . Por lo tanto, se requiere estudiar el criterio de factibilidad.

Variación en los costos  $c_i$ :

Esto varía la pendiente de las curvas de nivel, por lo que podría cambiar la solución óptima: Requiere que el criterio de optimalidad de Simplex siga cumpliéndose.

Variación en los factores tecnológicos  $a_{ij}$ : Generalmente requiere un estudio desde cero, porque ocurren ambas cosas simultáneamente.

Agregación de una variable  $x_{n+1}$ : Proceder como si hubiera cambiado un costo no básico, ya que no cambia el tamaño de la base.

Agregación de una restricción: Verificar si la solución actual sigue siendo factible. Notar que en este caso la base sí cambia su tamaño.

## Capítulo 6.- Teoría de Dualidad.

**Definición:** Sea el problema lineal siguiente:

$$P) \quad z^* = \begin{array}{ll} \text{mín} & z = \vec{c}^T \vec{x} \\ \text{s.a.} & A\vec{x} \geq \vec{b} \\ & \vec{x} \geq 0 \end{array}$$

Se le llama *problema dual a P* al problema

$$D) \quad w^* = \begin{array}{ll} \text{mín} & w = \vec{y}^T \vec{b} \\ \text{s.a.} & A^T \vec{y} \geq \vec{c} \\ & \vec{y} \geq 0 \end{array}$$

Para pasar de primal a dual, se requiere hacer lo siguiente:

- I) Por cada restricción  $j$  del problema primal se define una variable dual  $y_j$ .
- II) Los costos del problema dual serán los recursos del problema primal  $\vec{b}$ .
- III) Los recursos del problema dual serán los costos del problema primal  $\vec{c}$ .
- IV) La matriz A será transpuesta, y las variables que la acompañan serán las variables  $\vec{y}$ .
- V) Los signos de desigualdad de las restricciones y de la naturaleza variable del problema dual va directamente relacionada con los del primal, según la siguiente tabla:

Minimización	Maximización
Variables	Restricciones
$\geq 0$	$\leq$
$\leq 0$	$\geq$
Irrestringidas	$=$
Restricciones	Variables
$\leq$	$\leq 0$
$\geq$	$\geq 0$
$=$	Irrestringidas

**Teorema:** El *Teorema Débil de Dualidad* nos dice lo siguiente: Sea  $\vec{x}$  una solución factible del problema primal y sea  $\vec{y}$  una solución factible del problema dual. Entonces, si el problema primal es de minimización en  $\vec{x}$  y el dual de maximización en  $\vec{y}$ , entonces  $\forall \vec{x}, \vec{y}$  factibles se cumple que  $z(\vec{x}) \geq w(\vec{y})$ .

**Corolario:** Si alguno de los problemas es factible pero no acotado, entonces el otro problema es infactible.

**Corolario:** Si alguno de los problemas es factible y el otro es infactible, entonces el problema que admite solución factible es no acotado.

**Teorema:** El *Teorema Fuerte de Dualidad* nos dice lo siguiente: Dado un par de problemas primal-dual, si uno de ellos admite solución óptima, entonces el otro también la admite y los respectivos valores óptimos son iguales:

$$z(\vec{x}^*) = \vec{c}^T \vec{x}^* = \vec{b}^T \vec{y}^* = w(\vec{y}^*)$$

Cada base óptima en el primal *mapea* una base óptima en el dual.

Se tiene, también, que  $\vec{y}^* = \vec{\pi}^*$ , es decir, los valores de las variables duales en su óptimo representan los precios sombras de las restricciones respectivas en el primal.

Las holguras del dual son los costos reducidos del primal (y en el óptimo son  $\geq 0$ ). Además, si  $\vec{x}$  no es óptimo para el problema  $P$ , entonces  $\vec{y}$  no es factible para el problema  $D$ .

Sea  $\vec{x}$  una solución factible del problema primal y sea  $\vec{y}$  una solución factible del problema dual. Entonces, si el problema primal es de minimización en  $\vec{x}$  y el dual de maximización en  $\vec{y}$ , entonces  $\forall \vec{x}, \vec{y}$  factibles se cumple que  $z(\vec{x}) \geq w(\vec{y})$ .

**Teorema:** El *Teorema de Holgura Complementaria* nos dice que para el problema  $P$  y el problema  $D$ , ambos factibles, se tiene que si  $\vec{x}^*$  es óptima del primal y si  $\vec{y}^*$  es óptima del dual, entonces

$$\begin{aligned} y_i^* \left( b_i - \sum_{j=1}^n a_{ij} x_j^* \right) &= 0, & i = 1, \dots, m \\ x_j^* \left( c_j - \sum_{i=1}^m a_{ij} y_i^* \right) &= 0, & j = 1, \dots, n \end{aligned}$$

Es decir, si una restricción no es activa, entonces su correspondiente variable dual es nula.

Con respecto a las variables, se obtienen las siguientes relaciones (son bidireccionales):

En el Primal	En el Dual
Variables	Costos Reducidos asociados a Holguras
Holguras	Costos Reducidos asociados a Variables
Variables Básicas	Costos Reducidos de Variables No Básicas
Variables No Básicas	Costos Reducidos de Variables Básicas.
Múltiples Soluciones	Solución Degenerada

## Capítulo 7.- Programación Lineal Entera.

**Definición:** Un *modelo de Programación Lineal Entera* (PLE) es aquél cuyas variables son enteras, y tanto sus restricciones como su función objetivo son lineales. Existen dos tipos que se ven:

- Problema *entero puro*: todas las variables son enteras.

$$P) \quad \begin{aligned} z^* = \quad & \text{mín} \quad z = \vec{c}^T \vec{x} \\ \text{s.a.} \quad & A\vec{x} \geq \vec{b} \\ & \vec{x} \in \mathbb{Z}^+ \end{aligned}$$

- Problema *entero binario*: todas las variables toman valores 0 ó 1.

$$PB) \quad \begin{aligned} z^* = \quad & \text{mín} \quad z = \vec{c}^T \vec{x} \\ \text{s.a.} \quad & A\vec{x} \geq \vec{b} \\ & \vec{x} \in [0, 1] \end{aligned}$$

**Definición:** Un *modelo de Programación Lineal Mixto* (PLM) es aquél con algunas variables enteras o binarias y con otras continuas, y tanto sus restricciones como su función objetivo son lineales ( $1 \leq s \leq p < n$ ).

$$PB) \quad \begin{aligned} z^* = \quad & \text{mín} \quad z = \vec{c}^T \vec{x} \\ \text{s.a.} \quad & A\vec{x} \geq \vec{b} \\ & x_j \geq 0, \quad \forall j = 1, \dots, s \\ & x_j \in [0, 1], \quad \forall j = s + 1, \dots, p \\ & x_j \in \mathbb{Z}^+, \quad \forall j = p + 1, \dots, n \end{aligned}$$

**Definición:** Sea el problema binario  $PB$ . El siguiente problema se llama *relajación lineal* del problema entero mixto binario original.

$$PRB) \quad \begin{aligned} z^0 = \quad & \text{mín} \quad z = \vec{c}^T \vec{x} \\ \text{s.a.} \quad & A\vec{x} \geq \vec{b} \\ & \vec{x} \geq 0 \\ & 0 \leq x_j \leq 1, \quad \forall j \in J \subset \{1, \dots, n\} \end{aligned}$$

**Definición:** Sea el problema  $P$ . El siguiente problema se llama *relajación lineal* del problema entero original.

$$PR) \quad \begin{aligned} z^0 = \quad & \text{mín} \quad z = \vec{c}^T \vec{x} \\ \text{s.a.} \quad & A\vec{x} \geq \vec{b} \\ & \vec{x} \geq \vec{0} \end{aligned}$$

*Branch & Bound*: (o Ramificación y Acotamiento): Algoritmo que ramifica el espacio de soluciones cada vez que obtiene una solución fraccionaria en alguna componente.

**Definición:** En este contexto, se le llama *incumbente* a la mejor solución entera obtenida hasta el instante.

Supongamos que se desea resolver el problema  $P$ , cuyo espacio factible sin considerar la naturaleza variable denotaremos  $R_0$ :

- I) Se resuelve el problema relajado lineal del problema,  $PR$ . Si la solución es entera, no se acota más la rama. Si el valor de la función objetivo es peor que el valor del incumbente, no se acota más la rama. Si el problema actual es infactible, no se resuelve ni se ramifica. Si la solución posee alguna componente fraccionaria, digamos la componente  $k$ , entonces se debe ramificar.

- II) Si se decidió ramificar, entonces se crean dos subproblemas  $PR_1$  y  $PR_2$ , cuyos dominios serán:

$$\text{Para } PR_1 \rightarrow R_1 = R_0 \cap \{x : x_k \geq \lfloor x_k \rfloor + 1\}$$

$$\text{Para } PR_2 \rightarrow R_2 = R_0 \cap \{x : x_k \leq \lfloor x_k \rfloor\}$$

- III) Se continúa con cada rama reiniciando el algoritmo recursivamente.

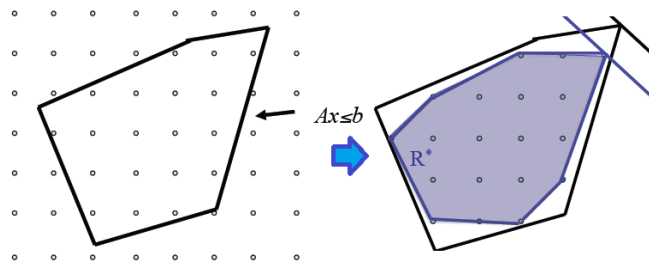
Se tiene que el valor óptimo de las ramas será siempre peor al valor óptimo del nodo desde donde fue ramificado.

A su vez, se puede **definir** el *GAP* o Brecha de Optimalidad, que indica qué tan lejos estamos del óptimo, con  $z_{RL}$  el mejor valor no entero actual y  $z_L$  el mejor valor entero actual:

$$\beta = \frac{z_{RL} - z_L}{z_{RL}}$$

### Planos Cortantes:

**Definición:** Un *plano cortante* es una restricción redundante para la formulación del problema pero que corta soluciones fraccionarias de la relajación lineal. De esta manera, se intenta hacer que el poliedro formado por las restricciones tengan como vértices soluciones enteras, como se muestra en la siguiente figura. Esto se puede hacer con la ayuda del gráfico del problema.



### Método de Gomory:

**Definición:** El *Método de Gomory* es una técnica antigua de corte del espacio, utilizando un algoritmo claro de resolución que impide que la solución óptima del problema relajado sea solución del próximo problema a resolver, sin excluir soluciones enteras del espacio. Este es el siguiente:

- I) Resolver el problema de relajación lineal  $PR$ . Si la solución resultante tiene todas las componentes enteras, no seguir. Si alguna es entera, realizar un corte.
- II) Obligar a que se cumpla lo siguiente:

$$\sum_{j \in \text{VNB}} (\bar{a}_{kj} - \lfloor \bar{a}_{kj} \rfloor)(x_R)_j \geq (\bar{b}_k - \lfloor \bar{b}_k \rfloor)$$

en donde VNB es el conjunto de índices de las variables no básicas,  $\bar{a}_{kj}$  son coeficientes de  $B^{-1}R$  y  $\bar{b} = B^{-1}\vec{b}$ . De esta manera, se obtienen cortes válidos para los valores de las variables no básicas. De las restricciones del problema estándar se pueden obtener igualdades que permitan escribir estas nuevas restricciones en términos de lo que convenga (en particular, de las variables principales para poder graficarlo).

Es útil mezclar los métodos, y ejecutar *Branch & Bound* con algunos cortes incluidos en el problema, para fortalecer la formulación.

## Capítulo 8.- Flujo en Redes.

Se asumen conocimientos básicos de grafos, árbol, camino y circuito.

**Problema de Flujo a Costo Mínimo:** Se busca determinar los flujos en los arcos de la red de modo que los elementos ofrecidos (o demandados) sean transportados hasta los puntos de demanda al menor costo posible, satisfaciendo las restricciones de capacidad y flujo mínimo por arco.

**Notación:** Para estos problemas se tiene un grafo dirigido  $G = (V, A)$ , en el que a cada *nodo*  $i \in V$  se le asocia una *oferta*  $b_i$  (demanda si es  $< 0$ ), con  $\sum_{i=1}^n b_i = 0$ . A cada *arco*  $(i, j) \in A$  se le asocia un *costo unitario* de transporte  $c_{ij}$ , una *cota mínima* de transporte  $l_{ij} \geq 0$  una *capacidad máxima* de transporte  $u_{ij} \geq l_{ij}$ .

El modelo quedará dado, con  $x_{ij}$  la cantidad de productos que irán del nodo  $i$  al nodo  $j$  por el arco  $(i, j)$ , por

$$\begin{aligned} \text{mín} \quad & \sum_{(i,j) \in A} c_{ij} x_{ij} \\ \text{s.a.} \quad & \sum_{(i,j) \in A} x_{ij} + \sum_{(k,i) \in A} x_{ki} = b_i \quad \forall i \in V \quad \text{Conservación de flujo} \\ & x_{ij} \leq u_{ij} \quad \forall (i, j) \in A \quad \text{Capacidad máxima} \\ & x_{ij} \geq l_{ij} \quad \forall (i, j) \in A \quad \text{Capacidad mínima (no negatividad)} \end{aligned}$$

La estructura de estos problemas tienen una bondad adicional debido a que la matriz de restricciones del problema es *totalmente unimodular* (también se dice que es de capacidad máxima) (determinante de cada una de sus submatrices cuadradas es 0, 1 ó -1): hace que los vértices del poliedro sean puntos enteros del espacio vectorial.

**Simplex Especializado en Redes:** Permite resolver los problemas de flujo a costo mínimo y análogos. Se mueve de un árbol generador (o base) a otro, hasta encontrar el óptimo.

- I) Se debe encontrar un *criterio de factibilidad* de la red, es decir, un grafo conexo y sin circuitos que abarque todos los nodos del problema sin generar circuitos. Además, este grafo debe permitir un flujo inicial factible.
- II) Con la ayuda del árbol, determinar las variables básicas y las no básicas. Las *variables básicas* son aquellos flujos  $x_{ij}$  tales que  $l_{ij} \leq x_{ij} \leq u_{ij}$ . Las *variables no básicas* son aquellas que  $x_{ij} = l_{ij}$  ó  $x_{ij} = u_{ij}$ . Luego, es fácil ver que todo flujo que cumple  $l_{ij} < x_{ij} < u_{ij}$  debe ser un flujo básico.
- III) Dado que los *costos reducidos de las variables básicas* deben ser 0, se calculan los valores de las variables duales  $\pi_i$ , teniendo en cuenta que los grados de libertad del problema permiten fijar una de esas variables arbitrariamente.

$$\bar{c}_{ij} = c_{ij} - \pi_i + \pi_j = 0$$

- IV) Se verifica el **criterio de optimalidad**, que indica si la base actual es óptima. Para ello, se verifican que se cumplan las siguientes condiciones para las variables *no* básicas:

$$\begin{aligned}\bar{c}_{ij} &= c_{ij} - \pi_i + \pi_j \geq 0 & \text{si } x_{ij} &= l_{ij} \\ \bar{c}_{ij} &= c_{ij} - \pi_i + \pi_j \leq 0 & \text{si } x_{ij} &= u_{ij}\end{aligned}$$

Si el criterio no se cumple, entonces se debe continuar al siguiente punto. Si el criterio se cumple, entonces se termina de iterar, y la base actual es la óptima. Si el valor de alguno de estos costos reducidos es cero, el problema presenta múltiples soluciones.

- V) Se verifica el **criterio de entrada**, que indica qué variable entrará a la base, es decir, cuál arco se incorpora al árbol actual. Si el criterio entrega más de un par  $(p, q)$ , se escoge arbitrariamente uno para continuar. La variable que entra será aquella con máxima violación de optimalidad.

$$\min\left\{\min_{\bar{c}_{ij}<0} \min_{x_{ij}=l_{ij}} \{\bar{c}_{ij}\}, -\max_{\bar{c}_{ij}>0} \max_{x_{ij}=u_{ij}} \{\bar{c}_{ij}\}\right\} = \bar{c}_{pq} \rightarrow x_{pq} \text{ entra a la base.}$$

- VI) Al ingresar el arco  $(p, q)$  a la base, se generará un circuito en el árbol generador, que indicará los arcos cuyo flujo se verá modificado por la reasignación. El **criterio de salida** nos indica cuál es el arco que acotará dicha reasignación de flujo  $\epsilon$ . Ese arco, una vez reasignado el flujo, será el que salga de la base. Si el criterio de salida entrega más de un par  $(r, s)$ , se escoge arbitrariamente uno para continuar:

$$\min\{u_{pq} - l_{pq}, \min_{(i,j) \in C_1} \{x_{ij} - l_{ij}\}, \min_{(i,j) \in C_2} \{u_{ij} - x_{ij}\}\} = \epsilon_{rs} \rightarrow x_{rs} \text{ sale de la base.}$$

donde  $C_1$  son los arcos del circuito formado que siguen el sentido del flujo, y  $C_2$  son los arcos del circuito formado que siguen el sentido contrario del flujo. La variable entrante domina el sentido del flujo.

Si el criterio de salida no entrega ningún par  $(r, s)$ , entonces el problema es un problema no acotado.

Notemos que puede ocurrir que la base no cambie (en el caso de que el mínimo sea  $u_{pq} - l_{pq}$ ), ya que para encontrar el punto de la iteración siguiente sólo bastaba reasignar el flujo.

- VII) Se **actualizan** los flujos, haciendo:

$$\begin{aligned}x_{ij}^{\text{nuevo}} &= x_{ij} + \epsilon & \text{si } x_{ij} &\in C_1 \\ x_{ij}^{\text{nuevo}} &= x_{ij} - \epsilon & \text{si } x_{ij} &\in C_2\end{aligned}$$

y se comienza nuevamente.

**Análisis de Sensibilidad para Flujo a costo mínimo:** Sigue el mismo criterio que Simplex lineal, analogizándolo de manera correcta.

**Problema de Flujo Máximo:** Consiste en determinar la máxima cantidad de flujo que se puede transportar a través de una red desde un nodo de origen a otro de destino.



**Denotación:** Para estos problemas se tiene un grafo dirigido  $G = (V, A)$ , en el que cada *arco*  $(i, j) \in A$  tiene una *cota mínima* de transporte  $l_{ij}$  y una *capacidad máxima* de transporte  $u_{ij}$ , y no tiene costo asociado. Existen dos nodos especiales: un *nodo de origen*  $r \in V$  y un *nodo de destino*  $s \in V$ .

El modelo quedará dado, con  $x_{ij}$  el flujo en el arco  $(i, j)$  y  $F$  el valor del flujo que entra en el origen, por

$$\begin{aligned} & \text{máx} && F \\ \text{s.a.} &&& \sum_{(i,j) \in A} x_{ij} - \sum_{(k,i) \in A} x_{ki} = \begin{cases} F & \text{si } i = r \\ 0 & \text{si } i \neq r, s \\ -F & \text{si } i = s \end{cases} \quad \forall i \in N \\ &&& l_{ij} \leq x_{ij} \leq u_{ij} \quad (i, j) \in A \end{aligned}$$

**Definición:** Un *corte de la red que separa  $r$  y  $s$*  define un subconjunto de nodos  $S \subseteq V$  tal que  $r \in S$  y  $s \in (V - S)$ , cuyo valor de corte o *capacidad* está dada por

$$C(S) = \sum_{(i,j): i \in S, j \in (V-S)} u_{ij}$$

El corte con la capacidad más pequeña se **denomina *corte mínimo*** y limita el flujo de la red. Es decir,  $F \leq C(S), \forall S \subseteq V$ .

**Teorema:** El *Teorema de Ford & Fulkerson* indica que el máximo valor de flujo (si existe) que se puede enviar desde el origen al destino es igual a la capacidad de un corte mínimo.

**Algoritmo de Ford-Fulkerson:** Construye iterativamente el flujo máximo a partir de un flujo factible, en que en cada iteración se identifica un camino de aumento de flujo con respecto al flujo existente.

- I) Se comienza considerando que todos los arcos  $(i, j)$  de un grafo  $G(V, A)$  poseen flujo 0 o algún flujo factible.
- II) Mientras existe un camino  $p$  desde  $r$  hacia  $s$  en la red residual, se calcula la capacidad máxima de dicho camino:

$$f_p = \min_{(i,j) \in p} \{u_{ij}\}$$

- III) A todos los arcos  $(i, j) \in p$  se le resta la cantidad  $f_p$  en la red. Es decir, se asume que se pasará esa cantidad por camino. Se recomienda anotar en cada arco la cantidad  $f_p$  de esa iteración, para después saber cuál es el flujo por cada arco. Se vuelve al paso II).

- IV) Una vez que no queden caminos posibles de  $r$  a  $s$  con  $f_p > 0$ , se ha terminado. El flujo máximo es

$$F = \sum_{p \text{ encontrados}} f_p$$

**Problema de la Ruta más Corta:** Consiste en determinar un camino dirigido en el grafo que conecte a un nodo con otro nodo a costo mínimo.

**Denotación:** Para estos problemas se tiene un grafo dirigido  $G = (V, A)$ , en el que cada *arco*  $(i, j) \in A$  tiene una se le asocia un *costo* de transporte  $c_{ij}$ . Se denotan dos nodos como especiales: un *nodo de origen*  $r \in V$  y un *nodo de destino*  $s \in V$ .

El modelo quedará dado, con  $x_{ij}$  variable binaria relajada que indica si se usará el arco  $(i, j)$  para llegar de  $r$  a  $s$ , por

$$\begin{aligned} \text{mín} \quad & \sum_{i=1}^n \sum_{j=1}^n c_{ij} x_{ij} \\ \text{s.a.} \quad & \sum_{j=1}^n x_{ij} - \sum_{k=1}^n x_{ki} = \begin{cases} 1 & \text{si } i = r \\ -1 & \text{si } i = s \\ 0 & \text{si } i \neq r, s \end{cases} \quad \forall i \in N \\ & x_{ij} \in \{0, 1\} \quad (i, j) \in A \end{aligned}$$

**Algoritmo de Dijkstra:** Iterativamente determina el camino más corto para llegar a un nodo nuevo hasta determinar el del nodo origen.

- I) El conjunto  $V$  de nodos se particiona en dos subconjuntos  $S$  y  $(V - S)$ . Un nodo pertenece a  $S$  si se conoce la ruta más corta para llegar a él desde el nodo origen  $r$ . Por lo tanto,  $(V - S)$  contiene el resto de los nodos.

Para cada nodo  $i$  se define un valor  $d_i = \infty$  que indica la distancia o costo total necesario para llegar al nodo  $i$  utilizando el mejor camino, y un valor  $n_i = j$  que indica el nodo  $j$  desde el cual es conveniente entrar al nodo  $i$ .

- II) Sea  $S = r$ ,  $d_r = 0$ ,  $n_r = r$ .

- III) Para el último nodo  $i$  en ingresar a  $S$ , se evalúan todos los nodos  $j \notin S$  que tienen en común un arco con  $i$ . Para todos esos nodos  $j$ , se actualiza el valor

$$d_j \leftarrow \min\{d_j, d_i + c_{ij}\}, (i, j) \in A, j \notin S$$

- IV) Si  $d_j^{\text{antiguo}} < d_j^{\text{nuevo}}$ , entonces actualizar  $n_j \leftarrow i$ .

- V) Sea  $k \in (V - S)$  tal que

$$d_k = \min\{d_i : i \in (V - S)\}$$

- VI) Si  $k = s$ , entonces seguir al paso VII). Si  $k \neq s$ , hace  $S = S \cup \{k\}$  y retomar en el paso III).

- VII) El camino óptimo para llegar del nodo  $r$  al nodo  $s$  queda determinado por los valores  $n_v$ , reconstruyendo el camino desde el destino hacia el origen.

Notemos que el algoritmo permite encontrar el camino óptimo desde un nodo origen hacia cualquier nodo  $i$  si se decide que el algoritmo termine una vez que  $S = V$ .

## Capítulo 9.- Optimización No Lineal Sin Restricciones.

Esta sección considerará un problema estándar

$$\begin{array}{ll} \text{mín} & f(\vec{x}) \\ \text{s.a.} & \vec{x} \in \mathbb{R}^n \end{array}$$

**Definición:** Un *modelo de Programación No Lineal* es aquél cuyas variables son continuas, y alguna de sus restricciones o función objetivo son no lineales.

**Teorema:** (*Condición necesaria*) Sea  $f$  diferenciable y sea  $x^*$  mínimo local de  $f$ . Entonces  $\nabla f(x^*) = \vec{0}$ .

**Definición:** Un *punto estacionario* o punto crítico es aquél punto  $x^*$  que satisface el teorema anterior. Estos puntos se pueden identificar como candidatos a mínimo.

**Teorema:** (*Condición necesaria de segundo orden*) Sea  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  una función de clase  $\mathcal{C}^2$ . Si  $x^* \in \mathbb{R}^n$  un mínimo local de  $f(\vec{x})$ , entonces se tiene que  $\nabla f(\vec{x}) = \vec{0}$  y que  $H(f(x^*)) = \nabla^2 f(x^*)$  es semidefinida positiva<sup>1</sup>.

**Teorema:** (*Condición suficiente*) Sea  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  una función clase  $\mathcal{C}^2$ . Sea el punto  $x^*$  tal que  $\nabla f(x^*) = \vec{0}$ . Si  $H(f(x^*)) = \nabla^2 f(x^*)$  es definida positiva,  $x^*$  es mínimo local estricto de  $f$ .

**Teorema:** Una función  $f$  es *convexa* en  $\Omega$  si y solo si  $H(f(\vec{x}))$  es una matriz semidefinida positiva, para todo  $\vec{x} \in \Omega$ . (Y es cóncava si y solo si es una matriz semidefinida negativa).

**Teorema:** Si  $H(f(\vec{x}))$  es una matriz definida positiva para todo  $\vec{x} \in \Omega$ , entonces la función  $f$  es *estrictamente convexa* en  $\Omega$ .

**Método del Gradiente o Descenso más pronunciado de Cauchy:** Método iterativo que comienza en un punto factible y se aproxima a un punto que satisfaga condiciones locales de optimalidad, siguiendo una dirección de movimiento por iteración.

- I) Comenzar el algoritmo con un punto factible inicial  $\vec{x}_k = \vec{x}_0 \in \mathbb{R}^n$ .
- II) Determinar la *dirección de descenso* o dirección de máximo decrecimiento. Esta queda determinada por  $\vec{h}_k = -\nabla f(\vec{x}_k)$ .

---

1

Una matriz cuadrada, invertible y simétrica  $H$  es *semidefinida positiva* si todos los determinantes de sus submatrices son positivos o cero. De manera similar, lo mismo ocurre si todos los valores propios de  $H$  son positivos o cero.

Una matriz cuadrada, invertible y simétrica  $H$  es *definida positiva* si todos los determinantes de sus submatrices son estrictamente positivos. De manera similar, lo mismo ocurre si todos los valores propios de  $H$  son estrictamente positivos.

Una matriz cuadrada, invertible y simétrica  $H$  es *semidefinida negativa* si los determinantes de sus submatrices  $A_{n \times n}$  con  $n$  impar son negativos o cero, y todos los determinantes de sus submatrices  $A_{n \times n}$  con  $n$  par son positivos o cero.

Una matriz cuadrada, invertible y simétrica  $H$  es *definida negativa* si los determinantes de sus submatrices  $A_{n \times n}$  con  $n$  impar son estrictamente negativos, y todos los determinantes de sus submatrices  $A_{n \times n}$  con  $n$  par son estrictamente positivos.

III) Si  $\vec{h}_k = \vec{0}$  terminar de iterar, pues tenemos un punto estacionario de  $f$ . Si no, resolver el problema:

$$\begin{aligned} \text{mín} \quad & f(\vec{x}_k + t_k \cdot \vec{h}_k) = f(\vec{x}_k - t_k \cdot \nabla f(\vec{x}_k)) \\ \text{s.a.} \quad & t > 0 \end{aligned}$$

$t_k$  se denomina **tamaño del paso**

IV) Actualizar el punto  $\vec{x}_{k+1} \leftarrow \vec{x}_k + t_k \cdot \vec{h}_k$ , y la iteración  $k \leftarrow k + 1$ . Volver al paso II.

Es posible determinar otros criterios de parada, como que  $\|\nabla f(\vec{x}_k)\| \leq \epsilon$ ,  $\|\vec{x}_{k+1} - \vec{x}_k\| \leq \epsilon$ ,  $\|f(\vec{x}_{k+1}) - f(\vec{x}_k)\| \leq \epsilon$ , o  $k > r$ .

Este algoritmo es de convergencia lineal.

**Teorema:** Sea  $\vec{x}^*$  mínimo local de  $f$ . Entonces, existe  $r > 0$  tal que si el método del gradiente es iniciado desde un punto  $\vec{x}_0$  tal que  $\|\vec{x}^* - \vec{x}_0\| < r$ , este converge a  $\vec{x}^*$ .

Luego, si  $f$  es estrictamente convexa y tiene mínimo, entonces el método converge desde cualquier punto de partida.

**Método de Newton:** Método iterativo que comienza en un punto factible y se aproxima a un punto que satisfaga condiciones locales de optimalidad, siguiendo una dirección de movimiento por iteración. Utiliza información de segundo orden.

**Algoritmo de Newton: (primera versión)**

- I) Comenzar el algoritmo con un punto factible inicial  $\vec{x}_k = \vec{x}_0 \in \mathbb{R}^n$ .
- II) Determinar el  $\nabla f(\vec{x}_k) = \vec{0}$ . Si resulta ser igual a  $\vec{0}$ , entonces terminar de iterar, pues tenemos un punto estacionario de  $f$ . Si no, continuar.
- III) Actualizar el punto  $\vec{x}_{k+1} \leftarrow \vec{x}_k - [\nabla^2 f(\vec{x}_k)]^{-1} \cdot \nabla f(\vec{x}_k)$ , y la iteración  $k \leftarrow k + 1$ . Volver al paso II).

**Algoritmo de Newton: (segunda versión)**

- I) Comenzar el algoritmo con un punto factible inicial  $\vec{x}_k = \vec{x}_0 \in \mathbb{R}^n$ .
- II) Determinar el  $\nabla f(\vec{x}_k) = \vec{0}$ . Si resulta ser igual a  $\vec{0}$ , entonces terminar de iterar, pues tenemos un punto estacionario de  $f$ . Si no, continuar.
- III) Sea  $\vec{h}_k = -[\nabla^2 f(\vec{x}_k)]^{-1} \cdot \nabla f(\vec{x}_k)$ . Resolver el problema:

$$\begin{aligned} \text{mín} \quad & f(\vec{x}_k + t_k \cdot \vec{h}_k) = f(\vec{x}_k - t_k \cdot [\nabla^2 f(\vec{x}_k)]^{-1} \cdot \nabla f(\vec{x}_k)) \\ \text{s.a.} \quad & t \geq 0 \end{aligned}$$

$t_k$  se denomina **tamaño del paso**

IV) Actualizar el punto  $\vec{x}_{k+1} \leftarrow \vec{x}_k + t_k \cdot \vec{h}_k$ , y la iteración  $k \leftarrow k + 1$ . Volver al paso II).

Se pueden determinar otro criterios de parada.

El algoritmo converge cuadráticamente. Si la función es cuadrática, el método converge en una iteración.

## Capítulo 10.- Optimización No Lineal Restringida.

Esta sección considerará un problema estándar

$$\begin{array}{ll} \min & f(\vec{x}) \\ \text{s.a.} & g_i(\vec{x}) \leq 0 \quad \forall i = 1, \dots, m \\ & \vec{x} \in \mathbb{R}^n \end{array}$$

Denotaremos el conjunto de soluciones factibles de  $P$ ) como  $\Omega = \{\vec{x} \in \mathbb{R}^n : g_i(\vec{x}) \leq b_i, \forall i = 1, \dots, m\}$ .

**Definición:** Sea  $\vec{x} \in \Omega$ . Un vector  $\vec{d} \in \mathbb{R}^n$  es una *dirección factible* en  $\Omega$  con respecto  $\vec{x}$  si existe  $\epsilon > 0$  tal que  $(\vec{x} + \lambda \vec{d}) \in \Omega, \forall \lambda \in (0, \epsilon]$ .

**Definición:** El *conjunto de direcciones factibles en  $\vec{x}$*  es el conjunto  $D(\vec{x}) = \{\vec{d} \in \mathbb{R}^n : \vec{d} \text{ es dirección factible de } \Omega \text{ con respecto a } \vec{x}\}$ .

**Definición:** Un *conjunto de índices activos  $I(\vec{x})$*  es el conjunto de los índices de las restricciones que están activas en un punto:  $I(\vec{x}) = \{i \in \{1, \dots, m\} : g_i(\vec{x}) = b_i\}$

**Definición:** Un *cono tangente  $T(\vec{x})$*  es el cono generado por los gradientes de las restricciones activas en ese punto, definido como  $T(\vec{x}) = \{\vec{h} \in \mathbb{R}^n : \nabla g_i(\vec{x})^T \cdot \vec{h} \leq 0, \forall i \in I(\vec{x})\}$ . Las direcciones factibles están en el cono tangente  $T(\vec{x})$ .

Caso Unidimensional:

$$\begin{array}{ll} P) \min & f(x) \\ \text{s.a.} & a \leq x \leq b \\ & x \in \mathbb{R} \end{array}$$

**Teorema: (Condición Necesaria de Primer Orden)** Sea  $f : \mathbb{R} \rightarrow \mathbb{R}$ . Si  $x^* \in [a, b]$  es un punto mínimo local de  $P$ ), entonces

- Si  $x^* = a \Rightarrow f'(x^*) \geq 0$ .
- Si  $x^* = b \Rightarrow f'(x^*) \leq 0$ .
- Si  $a < x^* < b \Rightarrow f'(x^*) = 0$ .

**Teorema: (Condición Suficiente de Primer Orden)** Sea  $f : \mathbb{R} \rightarrow \mathbb{R}$ . Si  $x^* \in [a, b]$  es un punto mínimo local de  $P$ ), entonces

- Si  $x^* = a$  y  $f'(x^*) > 0 \Rightarrow x^*$  es mínimo local estricto de  $P$ ).
- Si  $x^* = b$  y  $f'(x^*) < 0 \Rightarrow x^*$  es un mínimo local estricto de  $P$ ).

**Teorema: (Condición Suficiente de Segundo Orden)** Sea  $f : \mathbb{R} \rightarrow \mathbb{R}$  perteneciente a la clase  $\mathcal{C}^2$ . Si  $x^* \in [a, b]$  es un punto mínimo local de  $P$ ), entonces

- Si  $x^* = a \Rightarrow (f'(a) \geq 0) \wedge (f''(a) \geq 0 \text{ si } f'(a) = 0)$ .

- Si  $x^* = b \Rightarrow (f'(b) \leq 0) \wedge (f''(b) \geq 0 \text{ si } f'(b) = 0)$ .
- Si  $a < x^* < b \Rightarrow (f'(x^*) = 0) \vee (f''(x^*) \geq 0)$ .

Si se tiene más de un mínimo local, se deben comparar todos para encontrar el mínimo global del problema  $P$ ).

Problema con Restricciones de Igualdad:

$$\begin{aligned} P) \quad & \text{mín} \quad f(\vec{x}) \\ \text{s.a.} \quad & h_i(\vec{x}) = a_i \quad \forall i = 1, \dots, m \\ & x \in \mathbb{R}^n \end{aligned}$$

**Teorema de Lagrange:** (*Condición Necesaria de Primer Orden*) El *Teorema de Lagrange* dice que si  $\vec{x}^*$  (punto factible regular) es óptimo local del problema, entonces existen *multiplicadores de Lagrange*  $\lambda_1, \dots, \lambda_m$  tales que

$$\nabla f(\vec{x}^*) + \sum_{i=1}^m \lambda_i \nabla h_i(\vec{x}^*) = 0$$

**Definición:** La función  $\mathcal{L}$  a continuación se llama función Lagrangeana o *Lagrangeano*:

$$\mathcal{L}(\vec{x}, \vec{\lambda}) = f(\vec{x}) + \sum_{i=1}^m \lambda_i (h_i(\vec{x}) - a_i)$$

Es posible reescribir el problema como

$$\begin{aligned} \text{mín}_{(\vec{x}, \vec{\lambda})} \quad & \mathcal{L}(\vec{x}, \vec{\lambda}) \\ \text{s.a.} \quad & \frac{\partial \mathcal{L}}{\partial x} \mathcal{L}(\vec{x}^*, \vec{\lambda}^*) = 0 \\ & \frac{\partial \mathcal{L}}{\partial \lambda} \mathcal{L}(\vec{x}^*, \vec{\lambda}^*) = 0 \end{aligned}$$

Luego,  $(\vec{x}^*, \vec{\lambda}^*)$  es un punto estacionario del Lagrangeano.

**Definición:** Se dice que el punto  $\vec{x}$  es *regular* (o que cumple las condiciones de regularidad) si el Jacobiano<sup>2</sup>  $\mathcal{J}(\vec{x})$  de las restricciones es de rango máximo, es decir, son linealmente independientes. De lo contrario, se dice que el punto es *singular*.

$$^2 \mathcal{J}(\vec{x}) = \begin{bmatrix} \nabla h_1(\vec{x}) \\ \nabla h_2(\vec{x}) \\ \vdots \\ \nabla h_m(\vec{x}) \end{bmatrix}$$

Para que los puntos estacionarios  $\vec{x}^*$  sean óptimos locales, debe cumplirse que

$$\Delta x \frac{\partial^2 \mathcal{L}(\vec{x}^*, \vec{\lambda}^*)}{\partial x^2} \Delta x^T \geq 0$$

Así, la **condición necesaria de segundo orden** será que el Hessiano del Lagrangeano sea semidefinido positivo en el subespacio definido por las restricciones. Y la **condición suficiente de segundo orden** será que el Hessiano del Lagrangeano debe ser definido positivo en dicho subespacio. En este caso, el punto será mínimo local estricto.

Luego, basta evaluar los puntos en el Hessiano para determinar la naturaleza de los puntos.

**Interpretación de los Multiplicadores de Lagrange:** El valor  $\lambda_i$  es la sensibilidad del valor óptimo o **precio sombra** frente a una variación unitaria en el parámetro  $a_i$  de la restricción  $i$ . Si  $\vec{x}^*$  es un punto mínimo local:

$$\frac{\partial \mathcal{L}(\vec{x}^*)}{\partial a_i} = \frac{\partial f(\vec{x}^*)}{\partial a_i} = -\lambda_i$$

**Problema con Restricciones de Desigualdad:** El problema difiere en que las condiciones de optimalidad son distintas si se trata de un punto en el borde del dominio (restricciones activas) o de un punto interior (todas inactivas).

$$\begin{aligned} P) \quad & \text{mín} \quad f(\vec{x}) \\ \text{s.a.} \quad & g_i(\vec{x}) \leq b_i \quad \forall i = 1, \dots, m \\ & \vec{x} \in \mathbb{R}^n \end{aligned}$$

**Teorema:** La **condición necesaria de Karush-Kuhn-Tucker (KKT)** dice que para un  $\vec{x}^*$  mínimo local del problema  $P$ ), si se cumple la condición de regularidad en el punto  $\vec{x}^*$ , entonces existen multiplicadores  $\mu_1 \geq 0, \mu_2 \geq 0, \dots, \mu_m \geq 0$  tales que:

$$\nabla f(\vec{x}^*) + \sum_{i=1}^m \mu_i \nabla g_i(\vec{x}^*) = 0$$

$$\mu_i (g_i(\vec{x}^*) - b_i) = 0, \quad \forall i = 1, \dots, m \quad \text{Condiciones de holgura complementaria}$$

Estas son las denominadas **condiciones de KKT**.

**Definición:** Se dice que el punto  $\vec{x}$  cumple con la **condición de regularidad** (o es regular) si el conjunto  $T(\vec{x})$  define al conjunto de las direcciones factibles.

**Definición:** Se dice que el punto  $\vec{x}$  es **regular para las restricciones del problema** si los gradientes  $\nabla g_i(\vec{x}), i \in I(\vec{x})$  son linealmente independientes.

El problema se puede redefinir con el uso del **Lagrangeano** del problema:

$$\mathcal{L}(\vec{x}, \vec{\mu}) = f(\vec{x}) + \sum_{i=1}^m \mu_i (g_i(\vec{x}) - b_i)$$

en el que  $(\vec{x}^*, \vec{\mu}^*)$  es punto estacionario del problema  $\min_{(\vec{x}, \vec{\mu}), \vec{\mu} \geq 0} \mathcal{L}(\vec{x}, \vec{\mu})$ .

Luego, las condiciones de KKT necesarias para el siguiente problema de minimización/**maximización**<sup>3</sup>

$$P) \quad \begin{array}{ll} \text{mín} & f(\vec{x}) \\ \text{máx} & \\ \text{s.a.} & g_i(\vec{x}) \leq b_i \quad \forall i = 1, \dots, m \\ & x_j \geq 0 \quad j \in J, J \subseteq \{1, \dots, n\} \end{array}$$

y usando el Lagrangeano  $\mathcal{L}(\vec{x}, \vec{\mu}) = f(\vec{x}) + \sum_{i=1}^m \mu_i (g_i(\vec{x}) - b_i)$ , son

$$\begin{array}{llll} 1 \dots n) & \frac{\partial \mathcal{L}}{\partial x_j} & \begin{array}{l} \geq 0 \\ \leq 0 \end{array} & \forall j = 1, \dots, n \\ n+1 \dots n+m) & \frac{\partial \mathcal{L}}{\partial \mu_i} & \leq 0 & \forall i = 1, \dots, m \\ n+m+1 \dots 2n+m) & x_j \frac{\partial \mathcal{L}}{\partial x_j} & = 0 & \forall j = 1 \dots, m \\ 2n+m+1 \dots 2n+2m) & \mu_i \frac{\partial \mathcal{L}}{\partial \mu_i} & = 0 & \forall i = 1, \dots, m \\ 2n+2m+\text{card}(J) & x_j & \geq 0 & \forall j \in J \\ 2n+3m+\text{card}(J) & \mu_i & \begin{array}{l} \geq 0 \\ \leq 0 \end{array} & \forall i = 1, \dots, m \end{array}$$

**Teorema:** (*Suficiencia de las condiciones de KKT*) Si  $\vec{x}^*$ , solución factible de  $P$ ), es un punto que satisface las condiciones de KKT, y  $P$  es un problema convexo, entonces ese punto es mínimo global del problema (no requiere regularidad del punto óptimo).

**Teorema:** (*Condición de segundo orden de KKT*) Sea  $\vec{x}^*$  punto regular tal que se cumple KKT con multiplicadores  $\vec{\lambda}^*$ . Se tiene que:

- Si  $\vec{x}^*$  es mínimo local, entonces

$$\vec{d}^T \frac{\partial^2 \mathcal{L}(\vec{x}^*, \vec{\lambda}^*)}{\partial x^2} \vec{d} \geq 0, \quad \forall \vec{d} \in K(\vec{x}^*)$$

- Por otra parte, si

$$\vec{d}^T \frac{\partial^2 \mathcal{L}(\vec{x}^*, \vec{\lambda}^*)}{\partial x^2} \vec{d} > 0, \quad \forall \vec{d} \in K(\vec{x}^*), \vec{d} \neq \vec{0}$$

entonces  $\vec{x}^*$  es mínimo local estricto del problema.

<sup>3</sup>Puede ser más recomendable convertir el problema de máx  $f(\vec{x})$  a  $-\text{mín} -f(\vec{x})$ , y continuar con las condiciones de mínimo.



## Capítulo 11.- Programación Dinámica.

Pendiente

**Aclaración 1:** *Esto es un RESUMEN, por lo que puede haber parte de la materia no incluida intencionalmente.*

**Aclaración 2:** *Es posible que algo esté mal tipeado.*