

# Resumen Optimización

ICS1113 - Prof. José Tomás Marquinez V. - 2016

## Capítulo 0.- 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 *Construcción del problema*, la *Solución del modelo*, la *Verificación del modelo*, y la *Implementación y Control del modelo*.

## Capítulo 1.- 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.
- *Continuos*.
- *Determinísticos*.
- *Lineales*.
- *Estáticos*.
- *Discretos*.
- Estocásticos.
- *No Lineales*.

**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  $h_j(\vec{x}) = 0, g_i(\vec{x}) \leq 0$ .
- *Restricciones de las variables*: Limitaciones en los valores que pueden tomar las variables por sí solas, de la forma  $\vec{x} \in \Omega : \Omega \subset \mathbb{R}^n$ . Acá caen las restricciones de *Naturaleza de las Variables*.
- *Función Objetivo*: Medida para comparar las alternativas posibles. Se busca maximizar o minimizar este objetivo.

## Capítulo 2.- Conceptos básicos.

Se define el *espacio de soluciones factibles*  $\Omega$ , o simplemente espacio factible, como la intersección de todas las restricciones. Si esta 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  $\vec{x}^1$ .

Sea el siguiente problema de optimización:

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

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.

### Existencia de Soluciones Óptimas:

**Teorema:** El *Teorema 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 (al menos una) solución óptima (pues  $f$  alcanza sus puntos extremos).

**Teorema:** El *Teorema Práctico de Existencia de Soluciones Óptimas de la Optimización Lineal* dice que, en un problema lineal<sup>1</sup>, 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.

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

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

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

---

<sup>1</sup>Ver Capítulo 3

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

■ **Equivalencia I:**

$$P_1) \min_{\text{s.a.}} f(\vec{x}) \sim P_2) \max_{\vec{x} \in \Omega} -f(\vec{x})$$

■ **Equivalencia II:**

$$P_1) \min_{\text{s.a.}} f(\vec{x}) \sim P_2) \min_{\substack{\text{s.a.} \\ \vec{x} \in \Omega}} \mu \\ f(\vec{x}) \leq \mu$$

$$\vec{x} \in \Omega, \mu \in \mathbb{R}$$

■ **Equivalencia III:**

$$P_1) \min_{\text{s.a.}} \max_{\vec{x} \in \Omega} \{f_1(\vec{x}), \dots, f_n(\vec{x})\} \sim P_2) \min_{\substack{\text{s.a.} \\ \vec{x} \in \Omega}} \mu \\ f_i(\vec{x}) \leq \mu \quad \forall i = 1, \dots, n$$

$$\vec{x} \in \Omega, \mu \in \mathbb{R}$$

■ **Equivalencia IV:**

$$P_1) \min_{\text{s.a.}} \sum_{i=1}^r f_i(\vec{x}) \sim P_2) \min_{\substack{\text{s.a.} \\ \vec{x} \in \Omega}} \sum_{i=1}^r \mu_i \\ f_i(\vec{x}) \leq \mu_i \quad \forall i = 1, \dots, r$$

$$\vec{x} \in \Omega, \mu_i \in \mathbb{R} \quad \forall i = 1, \dots, r$$

■ **Equivalencia V:**

$$P_1) \min_{\text{s.a.}} f(\vec{x}) \sim P_2) \min_{\text{s.a.}} g(f(\vec{x}))$$

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:**

$$P_1) \min_{\text{s.a.}} \frac{1}{f(\vec{x})} \sim P_2) \max_{\vec{x} \in \Omega} f(\vec{x})$$

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

### Nociones Básicas de Convexidad:

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 (necesariamente) es convexa, la desigualdad lineal es convexa. Luego, todos los poliedros definidos por desigualdades lineales son convexas.

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 cónica 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 cónica.

La función  $f$  es una *función estrictamente cónica 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:** El *Teorema de Convexidad* indica que si, para el problema  $P$ ) convexo,  $\vec{x}$  es mínimo local de  $f$  en  $D$ , entonces  $\vec{x}$  es mínimo global de  $f$  en  $D$ . Sin embargo, no necesariamente este óptimo será único.

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

## Capítulo 3.- Introducción a la 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 da que en los Problemas de Programación Lineal (PPL) continuos existen tres posibilidades:

- No existe solución factible.
- El problema es no acotado.
- Existe (al menos) un punto extremo tal que es solución óptima (global) del problema.

**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, el (o los) punto que esté asociado a la curva de nivel de menor valor, será la solución óptima.

**Definición:** Un *poliedro* es un conjunto definido por restricciones lineales (afines).

$$P = \{x \in \mathbb{R}^n : Ax \leq b\} \text{ o bien } P = \{x \in \mathbb{R}^n : Ax = b, x \geq 0\}$$

**Definición:** Si el poliedro es acotado, se le dice *polítopo*.

**Teorema:** Sea  $P = \{x \in \mathbb{R}^n : Ax \leq b\}$  no vacío. El *Teorema de No Acotamiento* indica que  $P$  es un conjunto no acotado si y sólo si  $\exists h \in \mathbb{R}^n, h \neq 0$  tal que  $Ah \leq 0$ .

**Definición:** Un vector  $\vec{x}$  se dice *punto extremo* (o vértice) de  $\Omega$  si no existen  $\vec{u}, \vec{v} \in \Omega, \vec{u} \neq \vec{v}$  tales que  $\vec{x} = \lambda \vec{u} + (1 - \lambda) \vec{v}, 0 < \lambda < 1$ .

## 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:

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

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

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

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

con  $A \in \mathbb{R}^{n \times m}$ ;  $b \in \mathbb{R}^m$ ;  $c, x \in \mathbb{R}^n$ , tales que  $n \geq m$ ;  $b \geq 0$  y  $A$  tiene rango completo.

**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,  $x_i = \tilde{x}_i - \tilde{\tilde{x}}_i$ .
- 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:

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

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** o matriz básica.
- $R$  es una submatriz de  $A$  formada por  $n - m$  columnas.

**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)  $x^p$ :

$$\begin{array}{ll} \min & z = (c_B^p)^T \cdot \bar{b}^p + (\bar{c}_R^p)^T \cdot x_R^p \\ \text{s.a.} & Ix_B^p + \bar{R}x_R^p = \bar{b} \\ & x_B^p \geq 0 \\ & x_R^p = 0 \end{array}$$

### 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 sobreeterminado).

- III) Se verifica el **criterio de optimidad**, 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:

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

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  $e$ , se escoge arbitrariamente uno para continuar. Sea  $I_R$  el conjunto de índices de las variables no básicas, entonces:

$$\min_{\bar{c}_j < 0} \{\bar{c}_j : j \in I_R\} = \bar{c}_e \rightarrow x_e \text{ entra a la base.}$$

- 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. Sea  $I_B$  el conjunto de índices de las variables básicas, entonces:

$$\min_{\bar{a}_{ie} > 0} \left\{ \frac{\bar{b}_i}{\bar{a}_{ie}} : i \in I_B \right\} = \frac{\bar{b}_s}{\bar{a}_{se}} \rightarrow x_s \text{ sale de la base.}$$

- 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*:

Cabe mencionar que esta técnica dejó de verse en el curso ICS1113, sin embargo, este resumen lo incluye dado que antes se revisaba.

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

- 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 cuociente 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 cuociente 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.

## Capítulo 5.- Método de Simplex en su Fase I.

**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 (como si fuese 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 6.- 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  $B^*$ .

**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 base 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 7.- Teoría de Dualidad.

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

$$P) \quad z^* = \min_{\text{s.a.}} z = \vec{c}^T \vec{x}$$

$$\begin{aligned} & A\vec{x} \geq \vec{b} \\ & \vec{x} \geq 0 \end{aligned}$$

Se le llama *problema dual a P* al problema

$$D) \quad w^* = \min_{\text{s.a.}} w = \vec{y}^T \vec{b}$$

$$\begin{aligned} & A^T \vec{y} \geq \vec{c} \\ & \vec{y} \geq 0 \end{aligned}$$

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
$\geq 0$	$\leq$
$\leq 0$	$\geq$
Irrestrictas	$=$
Restricciones	Variables
$\leq$	$\leq 0$
$\geq$	$\geq 0$
$=$	Irrestrictas

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

**Corolario:** Si alguno de los problemas es factible y 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 *Lema de Farkas* nos dice lo siguiente: Considerando un problema en su forma estándar con una función objetivo cualquiera  $\vec{c}$ , y su dual, entonces

$$Ax = b, x \geq 0 \text{ infactible} \Leftrightarrow \exists y \neq 0 : A^T y \leq 0, b^T y > 0$$

**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  $D$ .

**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

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

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 8.- 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.

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

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

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

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

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

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

$$\begin{aligned} PR) \quad z^0 = & \min z = \vec{c}^T \vec{x} \\ \text{s.a.} \quad & A\vec{x} \geq \vec{b} \\ & \vec{x} \geq 0 \\ & 0 \leq x_j \leq 1, \quad \forall j = s + 1, \dots, p \end{aligned}$$

**Definición:** Una matriz  $A$  es *totalmente unimodular* si el determinante de cada una de sus submatrices cuadradas es 0, 1 ó  $-1$ .

**Teorema:** Si la matriz  $A$  es totalmente unimodular y  $\vec{b}$  es un vector con todos sus valores enteros, entonces el poliedro  $P = \{\vec{x} : A\vec{x} = \vec{b}\}$  tiene todos sus puntos extremos enteros.

**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:

$$\text{Para } PR_1 \rightarrow R_1 = R_0 \cap \{x : x_k \geq \lceil x_k \rceil\}$$

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

- 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:

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

Sin embargo, es más práctico y claro ir definiendo las mejores cotas superiores y cotas inferiores para el problema entero que se tiene hasta el momento.

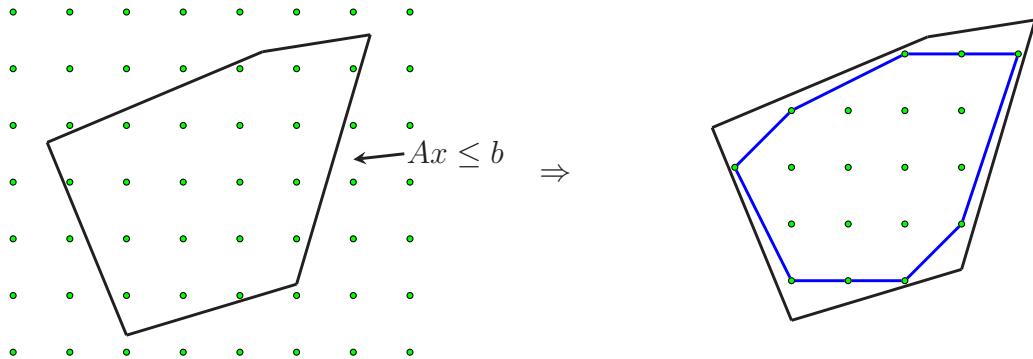
### Planos Cortantes:

**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, y corresponde a la envoltura convexa.

**Definición:** La **envoltura convexa**  $C(X)$  de un conjunto de puntos  $\Omega$  es la intersección de todos los conjuntos convexos que contienen a  $\Omega$ .

$$C(X) = \left\{ \sum_{i=1}^{|\Omega|} \alpha_i x_i \mid (\forall \alpha_i : \alpha_i \geq 0) \wedge \sum_{i=1}^{|\Omega|} \alpha_i = 1 \right\}$$



**Técnica de Covers:**

Consideremos restricciones del tipo

$$\sum_{i=1}^n a_i x_i \leq b, \text{ con } a_i \geq 0 \quad \forall i = 1, \dots, n \quad \text{y} \quad x_i \in \{0, 1\} \quad \forall i = 1, \dots, n$$

**Definición:** Un *cover*  $C$  es un subconjunto de  $\{1, \dots, n\}$  tal que  $\sum_{i \in C} a_i > b$ .

**Teorema:** De un *cover*  $C$ , se puede definir el siguiente plano cortante válido:

$$\sum_{i \in C} x_i \leq |C| - 1$$

**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:

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

en donde  $I_R$  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).

La restricción definida en (II) se define como *corte de Gomory*.

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

## Capítulo 9.- 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 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  $G = (N, A)$ , en el que a cada *nodo*  $i \in N$  ( $|N| = n$ ) 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  $(i, j) \in A$ , por

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

Para los alcances del curso ICS1113, considerar  $l_{ij} = 0$  y  $u_{ij} = \infty, \forall (i, j) \in A$ .

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).

**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:

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

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.

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

- 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:

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

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 vuelve a III.

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

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

**Denotación:** Para estos problemas se tiene un grafo dirigido  $G = (N, A)$ , con  $|N| = n$ , 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**  $s \in N$  y un **nodo de destino**  $t \in N$ .

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

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

**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 partitiona en dos subconjuntos  $V$  (también llamado  $S$ ) y  $NV$  (también llamado  $(V - S)$ ). Un nodo pertenece a  $V$  si se conoce la ruta más corta para llegar a él desde el nodo origen  $s$ . Por lo tanto,  $NV$  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  $p_i = j$  que indica el nodo  $j$  desde el cual es conveniente entrar al nodo  $i$ .

II) Sea  $V = s$ ,  $d_s = 0$ ,  $n_s = s$ .

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

$$d_j \leftarrow \min\{d_j, d_i + c_{ij}\}, \forall (i, j) \in A : j \in NV$$

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

V) Sea  $k \in (NV)$  tal que

$$d_k = \min\{d_i : i \in NV\}$$

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

VII) El camino óptimo para llegar del nodo  $s$  al nodo  $t$  queda determinado por los valores  $p_v$ , 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  $V = N$ .

**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, dada las capacidades de los arcos.

**Denotación:** Para estos problemas se tiene un grafo dirigido  $G = (N, A)$ , en el que cada **arco**  $(i, j) \in A$  tiene una se le asocia 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**  $s \in N$  y un **nodo de destino**  $t \in N$ .

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

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

**Definición:** Un **corte de la red que separa  $s$  y  $t$** , denotado como  $\text{corte-}(s, t)$ , es una partición de  $N$  en dos subconjuntos,  $S$  y  $T$ , tales que  $s \in S$  y  $t \in T$ . **Definición:** La **capacidad** o valor de corte está dada por

$$\text{cap}(S, T) = \sum_{(i,j) \in A: i \in S, j \in T} u_{ij}$$

El corte con la capacidad más pequeña se denomina **corte mínimo** y limita el flujo de la red. Es decir,  $F \leq \text{cap}(S, T), \forall (S, T) : S \subseteq V - \{t\}, s \in S, t \in T$ .

**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. Es decir **Teorema:** Además, se tiene que si  $F^*$  es el flujo máximo entre  $s$  y  $t$ , entonces  $F^* = \min_{(S,T)} \{\text{cap}(S, T)\}$ .

**Definición:** Se llama **grafo de capacidades residuales** al grafo  $G' = (N, A')$ , en que  $A'$  es:

- Para  $a = (i, j) \in A$ , si  $f_a < u_a$ , se pone un arco  $(i, j)$  con capacidad igual a  $u_a - f_a$ .
- Para  $a = (i, j) \in A$ , si  $f_a > l_a = 0$ , se pone un arco  $(j, i)$  con capacidad igual a  $f_a$ .

**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, a partir de la red residual.

- I) Se construye el grafo residual (o de capacidades residuales) a partir de algún flujo factible  $\vec{f}$  (o  $\vec{x}$ ), que tiene asociado un flujo saliente de  $s$  (o entrante a  $t$ ) igual a  $F$ .
- II) Mientras exista un camino  $p$  desde  $s$  hacia  $t$  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  $s$  a  $t$  con  $f_p > 0$ , se ha terminado el algoritmo.  
El flujo máximo es

$$F^* = F - \sum_{p \text{ encontrados}} f_p$$

## Capítulo 10.- Optimización No Lineal Sin Restricciones.

Esta sección considerará un problema estándar

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

**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  $f$  continuamente diferenciable. Si  $x^*$  es 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  $C^2$ . 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>2</sup>.

**Teorema:** (*Condición suficiente*) Sea  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  una función de clase  $C^2$ . Sea el punto  $x^*$  un punto crítico de  $f$ . 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  $\Omega$ .

**Teorema:** La función  $f$  es convexa en  $S \subseteq \mathbb{R}^n$ , conjunto convexo y no vacío, si y sólo si  $f(\vec{y}) \geq f(\vec{x}) + \nabla f(\vec{x})^T(\vec{y} - \vec{x}), \forall \vec{x}, \vec{y} \in S$ .

**Definición:**  $d \neq 0$  se dice *dirección de descenso* de  $f$  en  $x$  si existe  $r$  tal que  $f(x + \lambda d) < f(x), \forall \lambda : 0 \leq \lambda \leq r$ . Si  $d$  es tal que  $d^T \nabla f(x) < 0$ , entonces  $d$  es una dirección de descenso.

**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

---

<sup>2</sup>

Una matriz cuadrada, invertible y simétrica  $H_f$  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  $H_f$  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  $H_f$  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. De manera similar, lo mismo ocurre si todos los valores propios de  $H$  son negativos o cero.

Una matriz cuadrada, invertible y simétrica  $H_f$  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. De manera similar, lo mismo ocurre si todos los valores propios de  $H$  son estrictamente negativos.

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  $\vec{d}_k = -\nabla f(\vec{x}_k)$ .
- III) Si  $\vec{d}_k = \vec{0}$  terminar de iterar, pues tenemos un punto estacionario de  $f$ . Si no, resolver el problema:
 
$$\begin{array}{ll} \min & f(\vec{x}_k + t_k \cdot \vec{d}_k) = f(\vec{x}_k - t_k \cdot \nabla f(\vec{x}_k)) \\ \text{s.a.} & t > 0 \end{array}$$

$t_k$  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.

Este algoritmo cumple la propiedad de que  $(\vec{d}_k)^T (\vec{d}_{k+1}) = 0$ , es decir, el método avanza en direcciones mutuamente ortogonales.

**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 satisface 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  $\vec{d}_k = -[\nabla^2 f(\vec{x}_k)]^{-1} \cdot \nabla f(\vec{x}_k)$  la dirección de descenso a utilizar. Resolver el problema:

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

$t_k$  se denomina *tamaño del paso*.

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

Se pueden determinar otros criterios de parada.

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

## Capítulo 11.- Optimización No Lineal Restringida.

Esta sección considerará un problema estándar

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

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:** Un *conjunto de índices activos*  $I(\vec{x})$  (o, simplemente, conjunto activo) 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:** Dado los vectores  $\vec{v}_1, \dots, \vec{v}_r$ , el *cono generado por*  $\vec{v}_1, \dots, \vec{v}_r$  es el conjunto

$$\{x : x = \sum_{i=1}^r \lambda_i \vec{v}_i, \lambda_i \geq 0, \forall i = 1, \dots, r\}$$

**Definición:** Sea  $\vec{x} \in \Omega$ . Un vector  $\vec{d} \in \mathbb{R}^n, \vec{d} \neq 0$  es una *dirección factible* para  $\Omega$  con respecto a  $\vec{x}$  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 *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 Necesaria de Segundo Orden*) Sea  $f : \mathbb{R} \rightarrow \mathbb{R}$ , una función dos veces diferenciable. Si  $x^* \in [a, b]$  es un punto mínimo local de  $P$ , entonces

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

**Teorema:** (*Condición Suficiente de Segundo Orden*) Sea  $f : \mathbb{R} \rightarrow \mathbb{R}$  perteneciente a la clase  $C^2$ . Sea  $x^* \in [a, b]$ .

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

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:**

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

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

$$\nabla f(\vec{x}^*) + \sum_{j=1}^r \lambda_j \nabla h_j(\vec{x}^*) = 0$$

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

$$\mathcal{L}(\vec{x}, \vec{\lambda}) = f(\vec{x}) + \sum_{j=1}^r \lambda_j (h_j(\vec{x}) - a_j)$$

Es posible reescribir el problema como

$$\begin{aligned} \min_{(\vec{x}, \vec{\lambda})} \quad & \mathcal{L}(\vec{x}, \vec{\lambda}) \\ \text{s.a.} \quad & \vec{x} \in \mathbb{R}^n \\ & \vec{\lambda} \in \mathbb{R}^r \end{aligned}$$

Luego,  $(\vec{x}^*, \vec{\lambda}^*)$  es un punto estacionario del Lagrangeano. Así,  $\nabla_x \mathcal{L} = \vec{0}$  y  $\nabla_{\lambda} \mathcal{L} = h(\vec{x}) - a = \vec{0}$ .

**Definición:** Se dice que el punto  $\vec{x}$  es *regular* (o que cumple las condiciones de regularidad) si el Jacobiano<sup>3</sup>  $\mathcal{J}(\vec{x})$  de las restricciones evaluado en el punto es de rango máximo  $r$ , es decir, las  $r$  columnas son linealmente independientes. De lo contrario, se dice que el punto es *singular*.

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).

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

**Teorema:** Sea  $\vec{x} \in \Omega$ , un mínimo local del problema  $P$ ). Entonces,  $\nabla^T f(\vec{x})d \geq 0, \forall d \in D(\vec{x})$ .

**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*.

$${}^3 \mathcal{J}(\vec{x}) = \begin{bmatrix} \nabla h_1(\vec{x}) \\ \nabla h_2(\vec{x}) \\ \vdots \\ \nabla h_m(\vec{x}) \end{bmatrix}$$

**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. Es decir, si el Jacobiano de las restricciones activas tiene rango completo.

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  $\min_{(\vec{x}, \vec{\mu}), \vec{\mu} \geq \vec{0}} \mathcal{L}(\vec{x}, \vec{\mu})$ .

Luego, las condiciones de KKT necesarias para el siguiente problema de minimización/**maximización**<sup>4</sup>

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

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

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

**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:

<sup>4</sup>Puede ser más recomendable convertir el problema de  $\max f(\vec{x})$  a  $-\min -f(\vec{x})$ , y continuar con las condiciones de mínimo.

- 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.

## Capítulo 12.- Programación Dinámica.

Programación Dinámica es un capítulo que suele estar en el programa del curso ICS1113. Sin embargo, el semestre 2016'2 se excluyó del programa. Sin embargo, esta sección del resumen permanece para futuras referencias.

Programación dinámica se basa en el Principio de Optimalidad de Bellman, y consiste en solucionar el presente suponiendo que en cada etapa futura del problema siempre se tomarán las decisiones correctas.

**Teorema:** *Principio de Optimalidad de Bellman:* Sea  $x_1, x_2, \dots, x_n$  las decisiones óptimas de cada período, para un problema de  $n$  períodos. Entonces,  $x_k, x_{k+1}, \dots, x_n$  son decisiones óptimas para un problema de  $n - k + 1$  períodos.

Para que un problema pueda ser resuelto con esta técnica, debe cumplir con ciertas *características*:

- Naturaleza secuencial de las decisiones: El problema puede ser dividido en etapas.
- Cada etapa tiene un número de estados asociados a ella.
- La decisión óptima de cada etapa depende sólo del estado actual y no de las decisiones anteriores.
- La decisión tomada en una etapa determina cuál será el estado de la etapa siguiente.

Así, la política óptima desde un estado  $s$  de la etapa  $k$  a la etapa final está constituida por una decisión que transforma  $s$  en un estado  $s'$  de la etapa  $k + 1$  y por la política óptima desde el estado  $s'$  hasta la etapa final.

Para resolver el problema, se siguen los siguientes pasos:

- I) *Identificación de etapas, estados y variables de decisión:*
  - Cada etapa debe tener asociado una o mas decisiones (problema de optimización), cuya dependencia de las decisiones anteriores esta dada exclusivamente por las variables de estado.
  - Cada estado debe contener toda la información relevante para la toma de decisión asociada al período.
  - Las variables de decisión son aquellas sobre las cuales debemos definir su valor de modo de optimizar el beneficio acumulado y modificar el estado de la próxima etapa.
- II) *Descripción de ecuaciones de recurrencia:* Deben indicar cómo se acumula la función objetivo y como varían las funciones de estado de una etapa a otra.
- III) *Resolución* Se debe optimizar cada subproblema por etapas en función de los resultados de la resolución del subproblema siguiente. Notar que las para que las recurrencias estén bien definidas, se requiere de condiciones borde.

**Aclaración 1:** Claramente, este RESUMEN no abarca toda la materia en exhaustividad.

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