

2. CONCEPTOS BÁSICOS

2.1 ACERCA DE ESTE CAPÍTULO

En el presente capítulo se presentan los elementos teóricos básicos que nos permitirán abordar la solución de un programa lineal.

Las secciones 2.2 a 2.5 contienen sólo un breve repaso de algunos resultados de álgebra lineal que serán utilizados posteriormente. A partir de la sección 2.6 se introducen, en forma detallada, los conceptos específicos inherentes a la teoría de programación lineal.

2.2 VECTORES

Si consideramos el espacio vectorial P^v sobre P con las operaciones suma de vectores y producto por un escalar, un vector de P^v es un elemento del espacio expresado como una fila o como una columna de n elementos, siendo n la **dimensión** del vector.

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix} \text{ o } x = (x_1; x_2; \dots; x_n)$$

OPERACIONES CON VECTORES

Suma. Los vectores de la misma dimensión pueden sumarse (o restarse) y la suma (o diferencia) se efectúa componente a componente. La suma (o diferencia) de los vectores x e $y \in P^v$ es otro vector $z \in P^v$, cuyas componentes son: $z_i = x_i \pm y_i, \forall i = 1, 2, \dots, n$.

La suma de vectores goza de las propiedades asociativa, conmutativa, elemento neutro y elemento simétrico.

Producto de un escalar por un vector. Dado un escalar $\alpha \in P$ y un vector $x \in P^v$ se define el producto del escalar α por el vector x , como un nuevo vector $y \in P^v$ cuyas componentes son $y_i = \alpha x_i, \forall i = 1, 2, \dots, n$.

Producto escalar de dos vectores. Dados dos vectores x e $y \in P^v$ el producto escalar de ambos es un número real que se obtiene como $xy = x_1y_1 + x_2y_2 + \dots + x_ny_n$.

DEPENDENCIA E INDEPENDENCIA LINEAL

Definición 2-1. Combinación lineal. Se dice que el vector $y \in P^v$ es una combinación lineal de los vectores $x_1, x_2, \dots, x_k \in P^v$ si $y = c_1x_1 + c_2x_2 + \dots + c_kx_k$, con $c_1, c_2, \dots, c_k \in P$.

Los escalares c_1, c_2, \dots, c_k se denominan **coeficientes de la combinación lineal**.

Definición 2-2. Dependencia lineal. Un conjunto de vectores $x_1, x_2, \dots, x_k \in P^n$ es linealmente dependiente si existe una colección de escalares c_1, c_2, \dots, c_k , no todos nulos, tales que $c_1x_1 + c_2x_2 + \dots + c_kx_k = 0$ (con 0 =vector nulo).

Definición 2-3. Independencia lineal. Un conjunto de vectores $x_1, x_2, \dots, x_k \in P^V$ es linealmente independiente si la relación $c_1x_1 + c_2x_2 + \dots + c_kx_k = 0$ (con $0 = \text{vector nulo}$), se verifica sólo cuando los escalares c_1, c_2, \dots, c_k son todos nulos.

Definición 2-4. Conjunto generador. Una colección de vectores $x_1, x_2, \dots, x_k \in P^V$ se denomina conjunto generador de P^V , si cualquier vector $y \in P^V$ puede expresarse como combinación lineal de x_1, x_2, \dots, x_k .

Observación. La representación puede no ser única. El valor de k puede ser mayor o igual que n .

Definición 2-5. Base de un espacio vectorial. Un conjunto de vectores $x_1, x_2, \dots, x_m \in P^n$ constituyen una base del espacio vectorial P^n si:

- x_1, x_2, \dots, x_m son *linealmente independientes* y
- constituyen un *conjunto generador* de P^n .

Teorema 2-1. Si un espacio vectorial tiene una base constituida por m vectores, entonces toda base de dicho espacio tiene m vectores.

Definición 2-6. Espacio dimensionalmente finito. Si un espacio vectorial tiene una base constituida por m vectores, se dice que dicho espacio es dimensionalmente finito, siendo m su dimensión.

Teorema 2-2. Cualquier vector de un espacio vectorial puede escribirse de manera *única* como combinación lineal de los vectores de una base. Los escalares de tal combinación se denominan *coordenadas* del vector en la base dada.

Teorema 2-3. Reemplazo de un vector de la base por otro vector.² Sea $B = \{u_1, u_2, \dots, u_n\}$ una base del espacio vectorial P^V . Todo $u_j \neq 0$ que no integre dicha base y que por ende puede expresarse como combinación lineal de la misma en la forma $y_1u_1 + y_2u_2 + \dots + y_nu_n$, con $y_1, y_2, \dots, y_n \in P$, puede reemplazar a cualquier vector u_r de B para el cual y_r sea distinto de cero, y el nuevo conjunto de vectores seguirá siendo linealmente independiente, es decir, constituirá una nueva base.

Demostración. Reemplacemos u_r de B por u_j y razonemos por el absurdo admitiendo que $u_1, u_2, \dots, u_j, \dots, u_n$ son linealmente dependientes; por consiguiente debemos hallar n coeficientes $\mu_1, \mu_2, \dots, \mu_j, \dots, \mu_n$, no todos nulos, tales que

$$\mu_1u_1 + \mu_2u_2 + \dots + \mu_ju_j + \dots + \mu_nu_n = 0 \quad (2-1)$$

Sabemos que $u_1, u_2, \dots, u_r, \dots, u_n$ constituyen una base, y que cualquier subconjunto de un conjunto linealmente independiente también es linealmente independiente. Luego, para que se

verifique (2-1) debe darse que $\sum_{\substack{i=1 \\ i \neq j}}^n \mu_i u_i \neq 0$ y $\mu_j \neq 0$.

Podemos entonces dividir (2-1) por μ_j

$$\lambda_1u_1 + \lambda_2u_2 + \dots + u_j + \dots + \lambda_nu_n = 0, \text{ con } \lambda_i = \mu_i/\mu_j$$

reemplazando u_j por $y_1u_1 + y_2u_2 + \dots + y_ru_r + \dots + y_nu_n$ (combinación lineal de B)

$$(\lambda_1 + y_1)u_1 + (\lambda_2 + y_2)u_2 + \dots + y_ru_r + \dots + (\lambda_n + y_n)u_n = 0$$

² Este resultado será utilizado más adelante en el desarrollo y ejecución del método Simplex.

Puesto que $y_r \neq 0$, los vectores $u_1, u_2, \dots, u_r, \dots, u_n$ serían linealmente dependientes y no constituirían una base, lo cual es un absurdo que contradice la hipótesis y provino de suponer la dependencia lineal de $u_1, u_2, \dots, u_j, \dots, u_n$. Por lo tanto, los vectores $u_1, u_2, \dots, u_j, \dots, u_n$ constituyen una nueva base.

2.3 MATRICES

Una matriz A de orden $m \times n$ es un arreglo rectangular de $m \times n$ números en forma de m filas y n columnas.

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}. \text{ También se escribe, } A = (a_{ij})_{m \times n}$$

TIPOS DE MATRICES

Matriz cuadrada. Es aquella matriz en donde $m = n$, es decir, tiene igual número de filas que de columnas.

Matriz rectangular. Es aquella matriz en donde $m \neq n$.

Matriz diagonal. Es aquella matriz cuadrada en donde $a_{ij} = 0 \forall i \neq j$.

Matriz escalar. Es aquella matriz diagonal en donde todos los elementos de la diagonal son iguales.

Matriz identidad: Es aquella matriz escalar con todos los elementos de la diagonal iguales a 1. La matriz identidad se nota con I .

Matriz traspuesta. Se llama *matriz traspuesta*, A^T , de una matriz A a aquella matriz que tiene por columnas las filas de A , mientras que sus filas son las columnas de A .

Matriz simétrica. Una matriz A se denomina *simétrica* si $A^T = A$.

Matriz triangular. Es aquella matriz cuadrada cuyos elementos arriba o debajo de la diagonal principal son nulos.

Matriz nula. Es aquella que posee todos sus elementos iguales a cero. Se nota con 0 .

OPERACIONES CON MATRICES

Suma. Dadas dos matrices $A = (a_{ij})_{m \times n}$ y $B = (b_{ij})_{m \times n}$ se define como suma (o diferencia) de ambas, $A \pm B$, a una nueva matriz $C = (c_{ij})_{m \times n}$ tal que $c_{ij} = a_{ij} \pm b_{ij} \forall i, j$.

La suma de matrices goza de las propiedades asociativa y conmutativa.

Producto de un escalar por una matriz. Dado un escalar $\alpha \in P$ y una matriz $A = (a_{ij})_{m \times n}$ se define el producto del escalar por la matriz, αA , a una nueva matriz $B = (b_{ij})_{m \times n}$ tal que $b_{ij} = \alpha a_{ij} \forall i, j$.

Producto de matrices. El producto de una matriz $A = (a_{ij})_{m \times k}$ por una matriz $B = (b_{ij})_{k \times n}$ sólo está definido en el caso en que el número de columnas de A sea igual al número de filas de

B , y es igual a una nueva matriz $C = (c_{ij})_{m \times n}$ tal que, $c_{ij} = \sum_{p=1}^k a_{ip} b_{pj}$, $i = 1, 2, \dots, m$; $j = 1, 2, \dots, n$.

DETERMINANTE DE UNA MATRIZ

El **determinante** de una matriz cuadrada $A = (a_{ij})_{m \times m}$ es una función que le asigna a la matriz un único número real denominado determinante (o *módulo de la matriz*) y denotado como $\det(A)$ o $|A|$.

Definición 2-7. $|A| = a_{11}$, si $m = 1$.

Definición 2-8. Menor de una matriz. Dada una matriz A de orden $m \geq 2$, se denomina menor de la matriz asociado al elemento a_{ij} de A , y se nota con M_{ij} , al determinante que se obtiene de eliminar la fila i y la columna j de A .

Definición 2-9. Cofactor. Dada una matriz A de orden $m \geq 2$, se denomina cofactor de A asociado al elemento a_{ij} , y se nota con c_{ij} , al valor obtenido por $c_{ij} = -1^{i+j} M_{ij}$.

Definición 2-10. $|A| = \sum_{j=1}^m a_{1j} c_{1j}$, si $m \geq 2$. Es decir, el determinante de la matriz A es la

suma de los elementos de la primera fila de A multiplicados por sus respectivos cofactores.

Cuando el determinante de una matriz resulta nulo se dice que la matriz es **singular**, en caso contrario se denomina **no singular** o **regular**.

Propiedades de los determinantes

- Si A es una matriz de orden m , resulta

$$|A| = \sum_{j=1}^m a_{ij} c_{ij} = a_{i1} c_{i1} + a_{i2} c_{i2} + \dots + a_{im} c_{im}, \text{ para } i = 1, 2, \dots, m$$

$$\text{o bien, } |A| = \sum_{i=1}^m a_{ij} c_{ij} = a_{1j} c_{1j} + a_{2j} c_{2j} + \dots + a_{mj} c_{mj}, \text{ para } j = 1, 2, \dots, m.$$

- Si todos los elementos de una fila o de una columna de una matriz A son nulos, $|A| = 0$.
- $|A| = |A^T|$
- Si B se obtiene de A , intercambiando dos filas o dos columnas, entonces $|B| = -|A|$.
- Si A tiene dos filas o dos columnas idénticas o proporcionales, entonces $|A| = 0$.
- $|AB| = |A||B|$
- Si $B = \alpha A$, $\alpha \in P$, entonces $|B| = \alpha |A|$.

Regla de Sarrus. Proporciona un esquema sencillo para calcular determinantes de orden tres. *No es aplicable a matrices de orden mayor.*

Como se muestra a continuación, se escribe la matriz A repitiendo sus dos primeras columnas. Luego se efectúan los productos indicados por las flechas y se suman teniendo en cuenta que, los productos correspondientes a las flechas que se dirigen hacia abajo tendrán signo positivo y los correspondientes a las flechas que se dirigen hacia arriba, negativo.

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{11} & a_{12} \\ a_{21} & a_{22} & a_{23} & a_{21} & a_{22} \\ a_{31} & a_{32} & a_{33} & a_{31} & a_{32} \end{pmatrix}$$

$$|A| = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}$$

MATRIZ INVERSA

Si A y B son dos matrices no singulares de orden $m \times m$ tales que $AB = BA = I$, entonces B es la **inversa de A** y se nota con A^{-1} , o bien A es la **inversa de B** y se nota con B^{-1} . Si existe, la inversa de una matriz es única. Una matriz es invertible si, y sólo si, su determinante es distinto de cero.

Algunos resultados relativos a la inversión de matrices

- Si A es una matriz no singular (es decir, posee inversa), A^T también posee inversa y se cumple $(A^T)^{-1} = (A^{-1})^T$.
- Si A y B son no singulares $(AB)^{-1} = (B)^{-1}(A)^{-1}$.
- Toda matriz triangular posee inversa.
- Si A es invertible, el cálculo de su inversa viene dado por $A^{-1} = \frac{\text{adj}(A)}{|A|}$ donde **adj(A)**

es la **matriz adjunta** de A y se obtiene trasponiendo la matriz que resulta de reemplazar cada elemento de A por su correspondiente cofactor c_{ij} .³

RANGO DE UNA MATRIZ

Se llama **submatriz** de una matriz A a la matriz que resulta de suprimir filas o columnas de A .

Dada una matriz A se dice que el **rango** de A es r , si existe una submatriz cuadrada de orden r con determinante distinto de cero, y toda submatriz cuadrada de orden $r + 1$ tiene determinante nulo, conviniendo que el rango de la matriz nula es cero.

Propiedades de los rangos

- El rango r de una matriz cuadrada A de orden m , es menor que m si, y sólo si, $|A| = 0$.
- Dada una matriz cuadrada A , si $|A| = 0$, alguna fila o columna de A es combinación lineal de sus paralelas.

OPERACIONES ELEMENTALES CON MATRICES

Se dice que se efectúa una **transformación** u **operación elemental** en una matriz cuando:

- Se intercambian dos filas o dos columnas.
- Se multiplica una fila (columna) por un real distinto de cero.
- Se suma a una fila (columna) el producto de otra fila (columna) por un real.

Dos matrices son **equivalentes** si una se obtiene de la otra al realizar un número finito de operaciones elementales.

Teorema 2-4. Si en una matriz se efectúan operaciones elementales su rango no se modifica.

2.4 SISTEMAS DE ECUACIONES LINEALES

Un sistema de m ecuaciones lineales con n incógnitas es una expresión de la forma

$$\sum_{j=1}^n a_{ij}x_j = b_i, \quad i = 1, 2, \dots, m, \quad m (< ; = ; >) n$$

³ En el Ejemplo 2-2 de este capítulo se muestra un método alternativo para el cálculo de la inversa.

O bien,

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n &= b_2 \\ \dots + \dots + \dots + \dots &= \dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n &= b_m \end{aligned}$$

en forma matricial, $Ax = b$, siendo $A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$; $x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \dots \\ x_n \end{pmatrix}$; $b = \begin{pmatrix} b_1 \\ b_2 \\ \dots \\ b_m \end{pmatrix}$

Se llama **solución** del sistema a toda n -upla real $(x_1, x_2, \dots, x_n)^T$ que satisface cada una de las ecuaciones del sistema.⁴

Si $b \neq 0$ el sistema se denomina **no homogéneo**, en caso contrario es **homogéneo**.

Se denomina **matriz ampliada** o **matriz aumentada** a la matriz de m filas y $n + 1$ columnas formada por:

$$(A|b) = \left(\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{array} \right)$$

PROPIEDADES DE LOS SISTEMAS DE ECUACIONES LINEALES

- (Teorema de *Rouche Frobenius*) Un sistema de m ecuaciones con n incógnitas tiene solución si, y sólo si, el rango de la matriz ampliada y el rango de la matriz A de los coeficientes son iguales. En este caso se dice que **el sistema es compatible**.
En el caso en que el rango de A sea distinto al de la matriz ampliada el sistema no tiene solución y se denomina **sistema incompatible**.
- Si el rango r de la matriz ampliada es igual al rango de la matriz A y,
 - i. r coincide con el número n de incógnitas, la **solución es única**. Se denomina **sistema compatible determinado**.
 - ii. $r < n$, hay $n - r$ incógnitas que pueden tomar valores arbitrarios, existiendo por lo tanto **infinitas soluciones**. Se denomina **sistema compatible indeterminado**.
- Un sistema de igual número de ecuaciones e incógnitas tiene solución única si, y sólo si, la matriz A tiene determinante no nulo.

Dado un sistema lineal homogéneo de cualquier número de ecuaciones con n incógnitas, como la columna de los términos independientes es una columna de ceros, el rango de la matriz A resulta siempre igual al rango de la matriz ampliada, por lo tanto existe siempre solución. Además,

- Un sistema homogéneo tendrá solución única si, y sólo si, el rango de A es igual al número de incógnitas n . Dicha solución estará dada por $x_j = 0 \forall j = 1, 2, \dots, n$ y se denomina **solución trivial** del sistema.
- Un sistema homogéneo de igual número de ecuaciones e incógnitas admite infinitas soluciones además de la trivial si, y sólo si, el determinante de la matriz A es nulo.

⁴ El superíndice T significa trasposición.

SOLUCIÓN DE UN SISTEMA DE ECUACIONES LINEALES

Los métodos de resolución de sistemas de ecuaciones lineales se dividen en dos grandes grupos:

Los **métodos exactos** que requieren un determinado número de pasos para llegar a la solución del sistema.

Los **métodos aproximados** que utilizan algoritmos iterativos que calculan la solución del sistema por aproximaciones sucesivas.

Contrariamente a lo que pueda parecer, en numerosas ocasiones los métodos aproximados permiten obtener un grado de exactitud superior al que se obtiene mediante el empleo de métodos exactos, debido fundamentalmente a los errores de truncamiento que se producen en el proceso.

Entre los métodos exactos se encuentran el *método de Gauss*, una modificación de éste denominado *método de Gauss-Jordan* y la conocida *regla de Cramer*. Entre los métodos aproximados podemos citar los métodos de *Richardson*, *Jacobi* y *Gauss-Seidel*.

A los fines de nuestro estudio nos centraremos, a través de ejemplos, en el método de Gauss-Jordan y en Cramer.

Ejemplo 2-1 Método de Gauss-Jordan

Empleando este método, un sistema de m ecuaciones lineales con m incógnitas se transforma, mediante una combinación de ecuaciones que permite eliminar incógnitas, en un sistema equivalente de resolución inmediata.

Consideremos el siguiente sistema,

$$\begin{aligned} 2x_1 &+ 3x_3 + x_4 = 2 \\ x_1 - x_2 + 2x_3 + x_4 &= 4 \\ x_1 + x_2 + x_3 + x_4 &= -1 \\ 4x_1 + 2x_2 - x_3 + 2x_4 &= -3 \end{aligned}$$

Primero expresamos los coeficientes y el término independiente como una matriz amplia da.

$$\begin{array}{cccc|c} 2 & 0 & 3 & 1 & 2 \\ 1 & -1 & 2 & 1 & 4 \\ 1 & 1 & 1 & 1 & -1 \\ 4 & 2 & -1 & 2 & -3 \end{array}$$

El paso siguiente consiste en normalizar la primera fila dividiéndola por 2 (el valor 2 recibe el nombre de **elemento pivote**).

$$\begin{array}{cccc|c} 1 & 0 & 3/2 & 1/2 & 1 \\ 1 & -1 & 2 & 1 & 4 \\ 1 & 1 & 1 & 1 & -1 \\ 4 & 2 & -1 & 2 & -3 \end{array}$$

Luego, mediante operaciones elementales, se hacen cero los restantes elementos de la primera columna. Para ello, a la segunda fila le restamos la primera, a la tercera le restamos la primera y, a la última, le restamos la primera multiplicada por 4. Observemos que, a través de este procedimiento, hemos eliminado x_1 en todas las filas excepto en la primera.

$$\begin{array}{cccc|c} 1 & 0 & 3/2 & 1/2 & 1 \\ 0 & -1 & 1/2 & 1/2 & 3 \\ 0 & 1 & -1/2 & 1/2 & -2 \\ 0 & 2 & -7 & 0 & -7 \end{array}$$

A continuación, normalizamos la segunda fila multiplicándola por -1 (ahora el elemento pivote es el valor -1) y eliminamos x_2 en las filas tercera y cuarta (a la tercera fila le restamos la segunda y, a la cuarta, le restamos la segunda multiplicada por 2).

$$\begin{array}{cccc|c} 1 & 0 & 3/2 & 1/2 & 1 \\ 0 & 1 & -1/2 & -1/2 & -3 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & -6 & 1 & -1 \end{array}$$

Intercambiamos las filas tercera y cuarta.

$$\begin{array}{cccc|c} 1 & 0 & 3/2 & 1/2 & 1 \\ 0 & 1 & -1/2 & -1/2 & -3 \\ 0 & 0 & -6 & 1 & -1 \\ 0 & 0 & 0 & 1 & 1 \end{array}$$

Normalizamos la tercera fila dividiendo por el elemento pivote -6 y eliminamos x_3 en las filas primera y segunda (a la primera fila le restamos la tercera multiplicada por $3/2$ y, a la segunda, le sumamos la tercera multiplicada por $1/2$).

$$\begin{array}{cccc|c} 1 & 0 & 0 & 3/4 & 3/4 \\ 0 & 1 & 0 & -7/12 & -35/12 \\ 0 & 0 & 1 & -1/6 & 1/6 \\ 0 & 0 & 0 & 1 & 1 \end{array}$$

Finalmente eliminamos x_4 en las filas primera, segunda y tercera (a la primera fila le restamos la cuarta multiplicada por $3/4$, a la segunda le sumamos la cuarta multiplicada por $7/12$ y, a la tercera, le sumamos la cuarta multiplicada por $1/6$).

$$\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & -7/3 \\ 0 & 0 & 1 & 0 & 1/3 \\ 0 & 0 & 0 & 1 & 1 \end{array}$$

Este último sistema, equivalente al inicial, nos provee la solución dada por $x_1 = 0$, $x_2 = -7/3$, $x_3 = 1/3$ y $x_4 = 1$.

Ejemplo 2-2 Inversión de matrices

Sabemos que, si A es una matriz no singular de orden m , resulta $AA^{-1} = I$. Haciendo $X = A^{-1}$ y sustituyendo en la ecuación anterior, obtenemos $AX = I$.

Podemos considerar entonces que esta última ecuación matricial representa un sistema de ecuaciones simultáneas, en donde no hay un solo vector de términos independientes sino m , los m vectores que forman la matriz identidad I .

Además, no existe un solo vector de incógnitas, sino m , los que corresponden a cada columna de la matriz X . Es decir, $AX = I$.

O bien,

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mm} \end{pmatrix} \begin{pmatrix} x_{11} & x_{12} & \dots & x_{1m} \\ x_{21} & x_{22} & \dots & x_{2m} \\ \dots & \dots & \dots & \dots \\ x_{m1} & x_{m2} & \dots & x_{mm} \end{pmatrix} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{pmatrix}$$

De esta forma, es posible determinar la inversa mediante el método de Gauss–Jordan aplicando operaciones elementales sobre las filas de la matriz ampliada (A/I) , de manera de transformar A en I . Cuando lo hayamos logrado, obtendremos la inversa buscada A^{-1} .

A modo de ejemplo, supongamos que queremos hallar la inversa de la siguiente matriz.

$$A = \begin{pmatrix} 1 & -6 & 2 \\ 2 & -2 & -1 \\ 1 & -3 & -5 \end{pmatrix}$$

Aumentamos la matriz de coeficientes con la matriz identidad.

$$\begin{array}{ccc|ccc} 1 & -6 & 2 & 1 & 0 & 0 \\ 2 & -2 & -1 & 0 & 1 & 0 \\ 1 & -3 & -5 & 0 & 0 & 1 \end{array}$$

Normalizamos la primera fila (el valor del elemento pivote es 1) y hacemos cero los restantes elementos de la primera columna.

$$\begin{array}{ccc|ccc} 1 & -6 & 2 & 1 & 0 & 0 \\ 0 & 10 & -5 & -2 & 1 & 0 \\ 0 & 3 & -7 & -1 & 0 & 1 \end{array}$$

Observemos que para hallar la primera columna de A^{-1} debemos resolver el siguiente sistema.

$$\begin{pmatrix} 1 & -6 & 2 \\ 2 & -2 & -1 \\ 1 & -3 & -5 \end{pmatrix} \begin{pmatrix} x_{11} \\ x_{22} \\ x_{33} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

Para las columnas segunda y tercera el procedimiento será análogo, variando el vector de las incógnitas y el de los términos independientes. Es decir, para hallar A^{-1} resolvemos en forma simultánea tres sistemas de ecuaciones con tres incógnitas, cada uno de los cuales tiene la misma matriz de coeficientes A .

Normalizamos ahora la segunda fila (el valor del elemento pivote es 10) y anulamos el resto de los elementos de la segunda columna.

$$\begin{array}{ccc|ccc} 1 & 0 & -1 & -1/5 & 3/5 & 0 \\ 0 & 1 & -1/2 & -1/5 & 1/10 & 0 \\ 0 & 0 & -11/2 & -2/5 & -3/10 & 1 \end{array}$$

Por último, normalizamos la tercera fila y hacemos cero los restantes elementos de la tercera columna.

$$\begin{array}{ccc|ccc} 1 & 0 & 0 & -7/55 & 36/55 & -2/11 \\ 0 & 1 & 0 & -9/55 & 7/55 & -1/11 \\ 0 & 0 & 1 & 4/55 & 3/55 & -2/11 \end{array}$$

La inversa buscada será:

$$A^{-1} = \begin{pmatrix} -7/55 & 36/55 & -2/11 \\ -9/55 & 7/55 & -1/11 \\ 4/55 & 3/55 & -2/11 \end{pmatrix}$$

Ejemplo 2-3 Regla de Cramer

La regla de *Cramer* nos permite resolver un sistema de ecuaciones lineales en términos de determinantes, cuando la matriz A de los coeficientes es no singular.

Dado el sistema:

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mm} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_m \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \dots \\ b_m \end{pmatrix} \text{ o bien, } Ax = b$$

puesto que $|A| \neq 0$, podemos calcular A^{-1} mediante la expresión $A^{-1} = \text{adj}(A)/|A|$, por lo tanto, premultiplicando por A^{-1} en la ecuación $Ax = b$ resulta $x = A^{-1}b = \frac{\text{adj}(A)}{|A|}b$

Operando se tiene

$$\frac{1}{|A|} \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1m} \\ A_{21} & A_{22} & \dots & A_{2m} \\ \dots & \dots & \dots & \dots \\ A_{m1} & A_{m2} & \dots & A_{mm} \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ \dots \\ b_m \end{pmatrix} = x$$

en particular,

$$x_i = \frac{1}{|A|} (A_{i1}b_1 + A_{i2}b_2 + \dots + A_{im}b_m)$$

Para recordar la expresión anterior, si escribimos $|A|$ y reemplazamos la i -ésima columna por b , se obtiene el determinante $|B_i|$.

$$|B_i| = \begin{vmatrix} a_{11} & a_{12} & \dots & \overset{\text{\tiny \textit{i-ésima columna}}}{b_1} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & b_2 & \dots & a_{2m} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & b_m & \dots & a_{mm} \end{vmatrix}$$

El cofactor de b_k en $|B_i|$ es el cofactor de a_{ki} en $|A|$, si se desarrolla $|B_i|$ por la columna i , resulta $|B_i| = A_{1i}b_1 + A_{2i}b_2 + \dots + A_{mi}b_m$, de esta forma

$$x_i = \frac{1}{|A|} (A_{1i}b_1 + A_{2i}b_2 + \dots + A_{mi}b_m) = \frac{|B_i|}{|A|} \text{ para } i = 1, 2, \dots, m$$

siendo $x_1 = \frac{|B_1|}{|A|}$; $x_2 = \frac{|B_2|}{|A|}$; ... y $x_m = \frac{|B_m|}{|A|}$ la solución única del sistema propuesto.

Veamos como resolver por Cramer el siguiente sistema.

$$\begin{aligned} 2x_1 - x_2 + 3x_3 &= 4 \\ 4x_1 + x_2 - 5x_3 &= 6 \\ 3x_1 - 2x_2 + 2x_3 &= 2 \end{aligned}$$

$$|A| = \begin{vmatrix} 2 & -1 & 3 \\ 4 & 1 & -5 \\ 3 & -2 & 2 \end{vmatrix} = -26; x_1 = \frac{\begin{vmatrix} 4 & -1 & 3 \\ 6 & 1 & -5 \\ 2 & -2 & 2 \end{vmatrix}}{-26} = \frac{-52}{-26} = 2; x_2 = \frac{\begin{vmatrix} 2 & 4 & 3 \\ 4 & 6 & -5 \\ 3 & 2 & 2 \end{vmatrix}}{-26} = \frac{-78}{-26} = 3; x_3 = \frac{\begin{vmatrix} 2 & -1 & 4 \\ 4 & 1 & 6 \\ 3 & -2 & 2 \end{vmatrix}}{-26} = \frac{-26}{-26} = 1$$

2.5 CONVEXIDAD

A la hora de resolver un problema de optimización, el estudio de la convexidad de conjuntos y funciones adquiere fundamental relevancia, tanto en la búsqueda de óptimos como en el desarrollo de algoritmos eficientes de solución.

Definición 2-11. Recta en P^v . Dados dos puntos x_1 y $x_2 \in P^v$, la recta definida por dichos puntos puede expresarse mediante el conjunto:

$$R = \{x \in \mathbb{R}^n / x = \lambda x_1 + (1 - \lambda)x_2; x_1, x_2 \in \mathbb{R}^n; x_1 \neq x_2; \lambda \in \mathbb{R}\}$$

o bien,

$$R = \{x \in \mathbb{R}^n / x = x_2 + \lambda(x_1 - x_2); x_1, x_2 \in \mathbb{R}^n; x_1 \neq x_2; \lambda \in \mathbb{R}\}$$

Definición 2-12. Segmento en P^v . El segmento delimitado por los puntos $x_1, x_2 \in P^v$ puede expresarse mediante el conjunto:

$$[x_1, x_2] = \{x \in \mathbb{R}^n / x = \lambda x_1 + (1 - \lambda)x_2; x_1, x_2 \in \mathbb{R}^n; x_1 \neq x_2; \lambda \in [0; 1]\}$$

o bien,

$$[x_1, x_2] = \{x \in \mathbb{R}^n / x = x_2 + \lambda(x_1 - x_2); x_1, x_2 \in \mathbb{R}^n; x_1 \neq x_2; \lambda \in [0; 1]\}$$

Se trata de un **segmento lineal cerrado** ya que incluye los puntos extremos x_1 y x_2 . Cuando el segmento no incluye los dos puntos extremos se denomina segmento abierto y se nota,

$$]x_1, x_2[= \{x \in \mathbb{R}^n / x = \lambda x_1 + (1 - \lambda)x_2; x_1, x_2 \in \mathbb{R}^n; x_1 \neq x_2; \lambda \in]0; 1[\}$$

Definición 2-13. Hiperplano. Un hiperplano en P^v generaliza el concepto de recta en P^2 y el concepto de plano en P^3 . Un hiperplano en P^v es un conjunto dado por:

$$H = \{x \in \mathbb{R}^n / c^T x = z; z \in \mathbb{R}; c \in \mathbb{R}^n; c \neq 0\}$$

es decir, se trata de una expresión de la forma $c_1x_1 + c_2x_2 + \dots + c_nx_n = z$. El vector c se denomina **vector normal al hiperplano**.

Definición 2-14. Semiespacio. A partir del concepto de hiperplano se define el concepto de *semiespacio*. Un hiperplano $c^T x = z$ puede dividir P^n en tres semiespacios:

$$\{x \in \mathbb{R}^n / c^T x < z\}, \{x \in \mathbb{R}^n / c^T x = z\} \text{ y } \{x \in \mathbb{R}^n / c^T x > z\}$$

El primero se denomina *semiespacio inferior*, el segundo es el propio hiperplano y, el último, el denominado *semiespacio superior*. En los casos primero y tercero se trata de *semiespacios abiertos* puesto que las desigualdades son estrictas, es decir, menor (mayor) estrictamente. El propio hiperplano es un *semiespacio cerrado*.

La unión de un semiespacio abierto con el hiperplano da lugar a un semiespacio cerrado.

$$\begin{aligned} \{x \in \mathbb{R}^n / c^T x < z\} \cup \{x \in \mathbb{R}^n / c^T x = z\} &= \{x \in \mathbb{R}^n / c^T x \leq z\} \text{ semiespacio cerrado} \\ \{x \in \mathbb{R}^n / c^T x > z\} \cup \{x \in \mathbb{R}^n / c^T x = z\} &= \{x \in \mathbb{R}^n / c^T x \geq z\} \text{ semiespacio cerrado} \end{aligned}$$

Observemos que ambos semiespacios cerrados incluyen los puntos “sobre” el hiperplano H . Un *punto de frontera* de un conjunto es aquel que está ubicado en los límites de dicho conjunto, es decir, existe un entorno de él que contiene puntos pertenecientes y no pertenecientes al conjunto.

Asimismo, un *punto interior* de un conjunto es aquel que pertenece al conjunto y no se considera punto de frontera.

En base a lo expuesto, un conjunto que sólo contiene sus puntos interiores se denomina *abierto*. Si un conjunto contiene sus puntos interiores y sus puntos de frontera, es *cerrado*.

Definición 2-15. Polítopo. Se denomina polítopo a todo conjunto de P^n que puede expresarse como intersección de un número finito de semiespacios cerrados. La generalización a cualquier dimensión de un polígono bidimensional y un poliedro tridimensional es un polítopo.

Cuando el polítopo está acotado se llama *poliedro*.

Definición 2-16. Combinación lineal convexa. Se dice que un punto x es combinación lineal de n puntos si x puede expresarse como $x = \sum_{i=1}^n \lambda_i x_i$, con $\lambda_i \in \mathbb{R}$. Si además

$\lambda_i \geq 0 \forall i$ y $\sum_{i=1}^n \lambda_i = 1$, se trata de una combinación lineal convexa.

Como caso particular, el segmento delimitado por los puntos $x_1, x_2 \in P^n$ es el conjunto de combinaciones lineales convexas de dichos puntos.

Definición 2-17. Conjunto convexo. Un conjunto K perteneciente a P^n es convexo si dados dos puntos cualesquiera $x_1, x_2 \in K$ con $x_1 \neq x_2$, el segmento lineal cerrado que los une está totalmente contenido en K (es decir, cualquier combinación lineal convexa de tales puntos pertenece al conjunto).

Los hiperplanos son conjuntos convexas. Los semiespacios, tanto abiertos como cerrados, son conjuntos convexas.

Al definir un semiespacio se estableció que un conjunto es cerrado si incluye sus puntos de frontera. Naturalmente este concepto se aplica también a los conjuntos convexas. De esta forma, un conjunto convexo puede ser cerrado o abierto. A su vez, como se vio en el Ejemplo 1-15, un conjunto convexo puede ser *acotado* o *no acotado*, dependiendo de si el mismo se halla “limitado” o no. Por definición:

- El conjunto vacío es un conjunto convexo.
- Los conjuntos formados por un único punto son convexas.

Definición 2-18. Punto extremo. Un punto x de un conjunto convexo K es extremo o *vértice* de K si no puede expresarse como combinación lineal convexa de otros dos puntos de K , es decir, si no existe $x_1, x_2 \in K$, $x_1 \neq x_2$, $x = \lambda x_1 + (1 - \lambda)x_2$, con $0 \leq \lambda \leq 1$.

Un conjunto convexo puede tener tanto un número finito como infinito de puntos extremos.

Propiedades de los conjuntos convexos

- La intersección finita o infinita de conjuntos convexos es un conjunto convexo (en general, la unión de conjuntos convexos no es un conjunto convexo).
- La suma de conjuntos convexos es un conjunto convexo.
- La combinación lineal de conjuntos convexos es un conjunto convexo.

Todo polígono es un conjunto convexo por ser la intersección de un número finito de semiespacios cerrados.

2.6 TEOREMAS BÁSICOS DE LA PROGRAMACIÓN LINEAL

Dado un programa lineal en su forma estándar:⁵

$$\text{Max } z = cx$$

s. a

$$Ax = b$$

$$x \geq 0$$

donde

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}, x = \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix}, b = \begin{pmatrix} b_1 \\ b_2 \\ \dots \\ b_m \end{pmatrix} \text{ y } c = (c_1, c_2, \dots, c_n)$$

Definición 2-19. Solución. Es toda n -upla $(x_1, x_2, \dots, x_n)^T$ que satisface el sistema de restricciones $Ax = b$.

Definición 2-20. Solución factible. Es toda solución que además satisface la condición de no negatividad de las variables.

Definición 2-21. Solución básica. Es toda solución que posee, a lo sumo, m componentes distintas de cero (es decir al menos $n - m$ componentes nulas).

Sabemos que si $m < n$ y el rango de A es igual al rango de la matriz ampliada (A/b) e igual a m , el sistema $Ax = b$ tiene infinitas soluciones. En tal caso podemos extraer m columnas de A para formar una submatriz cuadrada B de rango m (estamos extrayendo una base B de vectores de la matriz A). Si llamamos con x_B al vector formado por las componentes de x asociadas a la matriz B (esto es, si la columna j forma parte de la matriz B entonces se toma la componente x_j del vector x) resulta que el sistema $Bx_B = b$ es compatible y determinado con solución $x_B = B^{-1}b$.

Luego, igualando a cero las $n - m$ componentes del vector x que no intervinieron en x_B , tendremos una solución básica compuesta por las componentes de x_B y las $n - m$ componentes nulas.

⁵ A partir aquí, en la forma estándar, designaremos con n a la cantidad de variables del vector x incluyendo, si las hubiera, holguras y/o excesos.

Así, el sistema $Ax = b$ puede expresarse siempre en términos de una base cualquiera como

$$Bx_B + Nx_N = b$$

siendo

B : una base de vectores de la matriz A

x_B : vector de **variables básicas** (cada componente de x_B recibe el nombre de variable básica)

N : matriz correspondiente a la “no base” (integrada por los $n - m$ vectores restantes de A no incluidos en B)

x_N : vector de **variables no básicas** (cada una de las $n - m$ componente nulas de x_N recibe el nombre de variable no básica)

Como mencionáramos anteriormente, para que el sistema $Ax = b$ tenga solución y en particular una solución básica, vamos a suponer que la matriz A es de orden $m \times n$, con $m < n$, y que su rango es máximo (igual a m). Tal supuesto recibe el nombre de **hipótesis de rango completo**.⁶

Si se cumple la hipótesis de rango completo el sistema $Ax = b$ tiene infinitas soluciones, pero no necesariamente todas ellas serán soluciones factibles.

Definición 2-22. Solución básica factible. Es toda solución básica en la cual todas sus variables básicas son no negativas. En general en desarrollos posteriores utilizaremos frecuentemente la expresión *solución básica* para referirnos a una *solución básica factible*.

Definición 2-23. Solución básica factible no degenerada. Es toda solución básica factible con exactamente m componentes estrictamente positivas. Si la solución posee menos de m componentes estrictamente positivas (es decir, más de $n - m$ componentes nulas) se denomina *solución básica factible degenerada*.

Definición 2-24. Región de factibilidad. Es el conjunto de todas las soluciones factibles.

Teorema 2-5. El conjunto de todas las soluciones factibles de un programa lineal es un conjunto convexo cerrado.

Demostración. Si x es una solución factible de un programa lineal satisface las condiciones

$Ax = b$ y $x \geq 0$. Cada una de las igualdades $\sum_{j=1}^n a_{ij}x_j = b_i$ con $i = 1, 2, \dots, m$, puede entenderse

como la intersección de dos semiespacios cerrados en P^v : $\sum_{j=1}^n a_{ij}x_j \leq b_i$ con $i = 1, 2, \dots, m$ y

$\sum_{j=1}^n a_{ij}x_j \geq b_i$ con $i = 1, 2, \dots, m$. Además las desigualdades $x_j \geq 0$ con $j = 1, 2, \dots, n$, también son

semiespacios cerrados en P^v . La región de factibilidad es entonces la intersección de un número finito de semiespacios cerrados, concretamente un polítopo, siendo por tanto un conjunto convexo cerrado.

La región de factibilidad puede ser:

- El conjunto vacío, en cuyo caso el programa lineal es **no factible** (no tiene solución).

⁶ La hipótesis de rango completo equivale a suponer que las m ecuaciones que componen el sistema $Ax = b$ son linealmente independientes.

- Un conjunto formado por un único punto, en cuyo caso dicho punto constituye la solución trivial.
- Un conjunto con más de un punto, en cuyo caso tiene infinitos puntos. En efecto, si x_1 y x_2 son soluciones factibles, también lo son los puntos del segmento cerrado delimitado por x_1 y x_2 .

Este último caso es el que nos interesa puesto que en los dos precedentes la solución es trivial. Pretendemos encontrar entre los infinitos puntos de la región de factibilidad aquel a aquellos que optimicen la función objetivo del programa lineal.

Teorema 2-6. Teorema fundamental de la programación lineal. Dado un programa lineal en su forma estándar que verifica la hipótesis de rango completo:

- si hay una solución factible también hay una solución básica factible;
- si hay una solución factible óptima también hay una solución básica factible óptima.

Demostración.

- Si $(x_1, x_2, \dots, x_n)^T$ es una solución factible se verifica que

$$x_1 A_1 + x_2 A_2 + \dots + x_n A_n = b$$

con $x_j \geq 0 \forall j = 1, 2, \dots, n$.

Supongamos que las p primeras de tales variables son no nulas, es decir

$$x_1 A_1 + x_2 A_2 + \dots + x_p A_p + 0A_{p+1} + \dots + 0A_n = b$$

podemos entonces analizar dos casos:

- Los vectores A_i con $i = 1, 2, \dots, p$ son linealmente independientes.
Como m es el rango de A , debe ser $p \leq m$:
si $p = m$ la solución es básica,
si $p < m$ basta con tomar $m - p$ columnas de A tales que agregadas a las p dadas formen una base de m vectores. Asignando el valor cero a las variables asociada a las $m - p$ columnas agregadas se tiene una solución básica degenerada.
- Los vectores A_i con $i = 1, 2, \dots, p$ son linealmente dependientes.
En este caso existen p escalares y_1, y_2, \dots, y_p , no todos nulos, tales que

$$y_1 A_1 + y_2 A_2 + \dots + y_p A_p + 0A_{p+1} + \dots + 0A_n = 0$$

Supongamos que al menos uno de los y_p es mayor que cero. Multiplicamos esta última ecuación por $\varepsilon \in \mathbb{P}$ y realizamos la siguiente operación

$$\begin{array}{r} x_1 A_1 + x_2 A_2 + \dots + x_p A_p + 0A_{p+1} + \dots + 0A_n = b \\ - \varepsilon y_1 A_1 - \varepsilon y_2 A_2 - \dots - \varepsilon y_p A_p - \varepsilon 0A_{p+1} - \dots - \varepsilon 0A_n = \varepsilon 0 \end{array}$$

resultando

$$(x_1 - \varepsilon y_1) A_1 + (x_2 - \varepsilon y_2) A_2 + \dots + (x_p - \varepsilon y_p) A_p + 0 + \dots + 0 = b$$

Luego, el vector $[(x_1 - \varepsilon y_1), (x_2 - \varepsilon y_2), \dots, (x_p - \varepsilon y_p), 0, \dots, 0]^T$ es solución del sistema $Ax = b$, sin embargo no podemos asegurar que todas sus componentes sean no negativas (evidentemente, si $\varepsilon = 0$ obtenemos la solución factible inicial). Eligiendo

$$\varepsilon = \min \left\{ \frac{x_i}{y_i}, \text{ para los } y_i > 0 \text{ con } i = 1, 2, \dots, p \right\}$$

obtenemos una solución factible que tiene como máximo $p - 1$ variables positivas (puesto que la o las componentes a las que corresponda el mínimo de las x_i/y_i se anulan).

Si los correspondientes $p - 1$ vectores son linealmente independientes se aplica el caso 1 y la solución obtenida es básica. En caso contrario, repitiendo el proceso anterior podemos eliminar variables positivas en cada nueva solución obtenida, hasta conseguir una solución factible a la que correspondan vectores linealmente independientes y aplicar el caso 1.

Si en la expresión

$$(x_1 - \varepsilon y_1)A_1 + (x_2 - \varepsilon y_2)A_2 + \dots + (x_p - \varepsilon y_p)A_p + 0 + \dots + 0 = b$$

no hubiera ningún $y_p > 0$, ε podría tomar valores positivos muy grandes y la solución no estaría acotada (el programa lineal no tiene solución óptima finita).

Queda demostrado entonces el apartado i.

ii. Sea $x = (x_1, x_2, \dots, x_n)^T$, una solución factible óptima. Supongamos que las p primeras de tales variables son no nulas, es decir,

$$x_1 A_1 + x_2 A_2 + \dots + x_p A_p + 0 A_{p+1} + \dots + 0 A_n = b$$

Si los vectores A_i con $i = 1, 2, \dots, p$ son linealmente independientes, la demostración es análoga al caso 1 de i. Ahora $x = (x_1, x_2, \dots, x_p, 0, \dots, 0)^T$ es una solución básica ($p \leq m$) y por hipótesis es óptima, por tanto x es una solución básica factible óptima.

Si los vectores A_i con $i = 1, 2, \dots, p$ son linealmente dependientes existen p escalares y_1, y_2, \dots, y_p , no todos nulos, tales que

$$y_1 A_1 + y_2 A_2 + \dots + y_p A_p + 0 A_{p+1} + \dots + 0 A_n = 0$$

En forma similar al caso 2 de i., si a $x_1 A_1 + x_2 A_2 + \dots + x_p A_p = b$ le sumamos $y_1 A_1 + y_2 A_2 + \dots + y_p A_p = 0$ multiplicada por un real θ , tenemos que

$$(x_1 + \theta y_1)A_1 + (x_2 + \theta y_2)A_2 + \dots + (x_p + \theta y_p)A_p = b$$

por tanto, $x + \theta y$ es una solución que satisface el sistema de restricciones.

Como las variables x_1, x_2, \dots, x_p son estrictamente positivas y los escalares y_1, y_2, \dots, y_p son finitos, es posible hallar $\varepsilon > 0$ tal que para $-\varepsilon \leq \theta \leq \varepsilon$ se tenga la factibilidad de la solución anterior.

Luego,

$$z = c(x + \theta y) = cx + \theta cy$$

Si $cy \neq 0$, puede ser $cy > 0$ o $cy < 0$. Para el caso en que $cy > 0$ tomamos $\theta = \varepsilon$ entonces

$$cx < cx + \theta cy$$

por lo tanto, x no es óptimo. Asimismo, cuando $cy < 0$ tomamos $\theta = -\varepsilon$ resultando nuevamente

$$cx < cx + \theta cy$$

es decir, x no es óptimo lo cual es un absurdo. Así pues, ha de ser $cy = 0$. En conclusión $x + \theta y$ también es una solución óptima para cualquier valor de $\theta \in [-\varepsilon, \varepsilon]$, por lo que siguiendo los mismos pasos del caso 2 de i., concluimos que es posible encontrar una solución básica factible óptima.

Teorema 2-7. Teorema de equivalencia. Dado un programa lineal en su forma estándar que verifica la hipótesis de rango completo. Sea K el conjunto convexo de sus soluciones factibles (polítopo). Un punto x es punto extremo de K si, y sólo si, es una solución básica factible.

Demostración. Sea $x = (x_1, x_2, \dots, x_m, 0, \dots, 0)^T$ una solución básica factible, entonces

$$x_1 A_1 + x_2 A_2 + \dots + x_m A_m + 0 A_{m+1} + \dots + 0 A_n = b$$

siendo los vectores A_i con $i = 1, 2, \dots, m$, linealmente independientes.

Si suponemos que x es combinación lineal convexa de otros dos puntos y, q pertenecientes a K , resultará

$$x = \lambda y + (1 - \lambda)q, \text{ con } 0 \leq \lambda \leq 1$$

Dado que x, y y q tienen componentes no negativas y $0 \leq \lambda \leq 1$, las $n - m$ últimas componentes de y y q tienen que ser nulas, luego

$$\begin{aligned} y_1 A_1 + y_2 A_2 + \dots + y_m A_m &= b \\ q_1 A_1 + q_2 A_2 + \dots + q_m A_m &= b \end{aligned}$$

Como los vectores A_i con $i = 1, 2, \dots, m$ son linealmente independientes debe ser $y = q$ y consecuentemente, $x = y = q$. Hemos probado que la solución básica factible es un punto extremo de K .

Recíprocamente, supongamos que $x = (x_1, x_2, \dots, x_k, 0, \dots, 0)^T$ es un punto extremo de K que posee las k primeras componentes no nulas (y por ende positivas). Por tanto

$$x_1 A_1 + x_2 A_2 + \dots + x_k A_k = b$$

Si comprobamos que los vectores A_1, A_2, \dots, A_k son linealmente independientes, el punto extremo será una solución básica factible.

Procederemos por el absurdo, suponiendo que A_1, A_2, \dots, A_k son linealmente dependientes. Existen entonces k escalares y_1, y_2, \dots, y_k , no todos nulos, tales que

$$y_1 A_1 + y_2 A_2 + \dots + y_k A_k = 0$$

Para el vector $y = (y_1, y_2, \dots, y_k, 0, \dots, 0)^T$ se comprueba que, tomando un ε adecuado, los vectores $x + \varepsilon y$ y $x - \varepsilon y$ pertenecen al conjunto K de soluciones factibles. En efecto, si a la ecuación $x_1 A_1 + x_2 A_2 + \dots + x_k A_k = b$ le restamos $y_1 A_1 + y_2 A_2 + \dots + y_k A_k = 0$ multiplicada por un real ε mayor que cero, obtenemos

$$(x_1 - \varepsilon y_1) A_1 + (x_2 - \varepsilon y_2) A_2 + \dots + (x_k - \varepsilon y_k) A_k = b$$

Si ahora a $x_1 A_1 + x_2 A_2 + \dots + x_k A_k = b$ le sumamos $y_1 A_1 + y_2 A_2 + \dots + y_k A_k = 0$ multiplicada por un real ε mayor que cero, tenemos que

$$(x_1 + \varepsilon y_1) A_1 + (x_2 + \varepsilon y_2) A_2 + \dots + (x_k + \varepsilon y_k) A_k = b$$

Luego, como los x_i son todos positivos, escogiendo $\varepsilon \leq \min \left\{ \frac{x_i}{|y_i|}, y_i \neq 0, |y_i| = \text{valor absoluto de } y_i \right\}$, los vectores $x + \varepsilon y$ y $x - \varepsilon y$ son soluciones factibles del programa lineal. Por otra parte, $x = 0,5(x - \varepsilon y) + 0,5(x + \varepsilon y)$ lo cual nos dice que x es combinación convexa de dos puntos distintos de K , contradiciendo la hipótesis original que indica que x es un punto extremo de K . El absurdo provino de suponer que los vectores A_1, A_2, \dots, A_k eran linealmente dependientes.

En consecuencia, tales vectores son linealmente independientes y el punto extremo es una solución básica factible (que puede ser degenerada si $k < m$).



Teorema 2-8. La función objetivo de un programa lineal alcanza su óptimo en al menos un punto extremo del conjunto de soluciones factibles. Si el óptimo se produce en más de un punto extremo, entonces también se produce en todo punto que es combinación convexa de ellos.

Demostración. Sea K la región factible y x_1, x_2, \dots, x_q los puntos extremos de K . Sea x un punto cualquiera perteneciente a K , entonces x puede expresarse como combinación lineal convexa de los puntos extremos. Esto es

$$x = \sum_{i=1}^q \lambda_i x_i, \text{ con } \lambda_i \in \mathbb{R}, \lambda_i \geq 0 \forall i \text{ y } \sum_{i=1}^q \lambda_i = 1$$

en consecuencia

$$z = cx = c \sum_{i=1}^q \lambda_i x_i = \sum_{i=1}^q \lambda_i (cx_i) = \sum_{i=1}^q \lambda_i z_i$$

donde z_i es el valor que toma la función objetivo en $x_i, i = 1, 2, \dots, q$.

Sea $z_0 = \max\{z_i\}, i = 1, 2, \dots, q$, entonces $z_0 \geq cx_i, i = 1, 2, \dots, q$. Luego

$$z_0 = 1z_0 = \left(\sum_{i=1}^q \lambda_i\right)z_0 = \sum_{i=1}^q \lambda_i z_0 \geq \sum_{i=1}^q \lambda_i (cx_i) = cx = z$$

es decir $z_0 \geq z$ para todo x , por lo tanto z_0 es el óptimo y se produce en al menos un punto extremo.

Si el óptimo se produce en p puntos extremos x_1, x_2, \dots, x_p tenemos que

$$z^* = cx_1 = cx_2 = \dots = cx_i = \dots = cx_p$$

si x es una combinación convexa de ellos

$$x = \sum_{i=1}^p \lambda_i x_i, \text{ con } \lambda_i \in \mathbb{R}, \lambda_i \geq 0 \forall i \text{ y } \sum_{i=1}^p \lambda_i = 1$$

luego la función objetivo resulta

$$z = cx = c \sum_{i=1}^p \lambda_i x_i = \sum_{i=1}^p \lambda_i (cx_i) = \sum_{i=1}^p \lambda_i z^* = z^* \sum_{i=1}^p \lambda_i = z^*$$

con lo cual x también es óptimo.

CONCLUSIONES

Los teoremas anteriores nos permiten deducir las siguientes consecuencias inmediatas.

- Si el conjunto de soluciones factibles de un programa lineal que verifica la hipótesis de rango completo es no vacío, entonces tiene al menos un punto extremo.
- Cada solución básica factible de un programa lineal corresponde a un punto extremo del conjunto de soluciones factibles (y viceversa).

- **El conjunto de soluciones factibles de un programa lineal que verifica la hipótesis de rango completo, tiene un número finito de puntos extremos.**

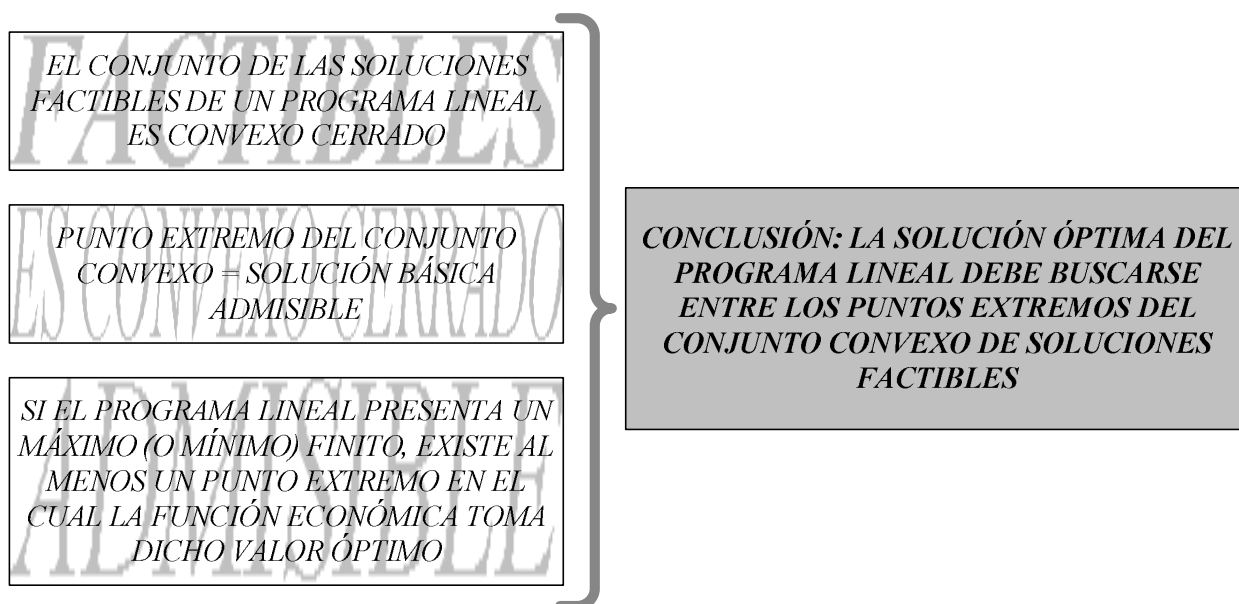
Efectivamente, el *número máximo* de puntos extremos de una región factible proveniente de un programa lineal con n variables y m restricciones está dado por todas las combinaciones de los n vectores de la matriz A tomados de a m .

$$N^{\circ} \text{ máximo de puntos extremos} = C_{n,m} = \binom{n}{m} = \frac{n!}{(n-m)!m!}$$

- **Si un programa lineal tiene una solución óptima finita la misma se produce en al menos un punto extremo del conjunto de soluciones factibles.**

En la Figura 2-1 se resumen los conceptos precedentes.

Figura 2-1. ELEMENTOS FUNDAMENTALES DE LA TEORÍA DE PROGRAMACIÓN LINEAL



2.7 SOLUCIÓN CONCEPTUAL

En base a lo expuesto se deduce que para obtener la solución óptima habrá que calcular el valor que toma la función económica en cada punto extremo y seleccionar el mejor:

Para obtener un punto extremo podemos proceder de la siguiente manera:

- Expresamos el programa lineal en su forma estándar.
- De la matriz A seleccionamos una base (m vectores columna linealmente independientes).
- Resolvemos el sistema normal de Cramer que se obtiene haciendo cero las $n - m$ incógnitas que no acompañan a la base.
- Si todas las incógnitas que resultan de resolver el sistema de Cramer son no negativas, esta solución, con el agregado de los $n - m$ ceros anteriores, nos proporciona un punto extremo.

Dado que el número de puntos extremos (soluciones básicas factibles) es finito, repetimos el proceso hasta agotar todos los casos posibles es decir, todas las combinaciones de n vectores de la matriz A tomados de a m .

Ejemplo 2-4 Solución conceptual

Veamos como resolver conceptualmente el problema planteado en el Ejemplo 1-1.

$$\text{Max } z = 20x_1 + 45x_2 \text{ (maximizar contribución marginal)}$$

s. a

$$x_1 + 2x_2 \leq 40 \text{ (restricción debida a la mano de obra)}$$

$$3x_1 + 1,5x_2 \leq 75 \text{ (restricción debida a la materia prima)}$$

$$x_2 \leq 15 \text{ (restricción debida a la demanda)}$$

$$x_1; x_2 \geq 0 \text{ (condición de no negatividad de las variables)}$$

Llevamos el problema a su forma estándar

$$\text{Max } z = 20x_1 + 45x_2 + 0x_3 + 0x_4 + 0x_5$$

s. a

$$x_1 + 2x_2 + x_3 = 40$$

$$3x_1 + 1,5x_2 + x_4 = 75$$

$$x_2 + x_5 = 15$$

$$x_j \geq 0 \quad j = 1; 2; \dots; 5$$

Indiferentemente de cuál sea la función económica, para hallar un punto extremo debemos encontrar una base, en este caso 3 vectores columna de la matriz A que resulten linealmente independientes ($m = 3$). Probemos con (A_1, A_2, A_3) .

$$\det(A_1, A_2, A_3) = \begin{vmatrix} 1 & 2 & 1 \\ 3 & 3/2 & 0 \\ 0 & 1 & 0 \end{vmatrix} = 3 \Rightarrow (A_1, A_2, A_3) \text{ linealmente independientes}$$

Luego, resolvemos por Cramer el sistema resultante de anular x_4 y x_5 ($n - m = 5 - 3 = 2$ variables que no acompañan a la base).

$$x_1 = \frac{\begin{vmatrix} 40 & 2 & 1 \\ 75 & 3/2 & 0 \\ 15 & 1 & 0 \end{vmatrix}}{3} = 35/2 \quad x_2 = \frac{\begin{vmatrix} 1 & 40 & 1 \\ 3 & 75 & 0 \\ 0 & 15 & 0 \end{vmatrix}}{3} = 15 \quad x_3 = \frac{\begin{vmatrix} 1 & 2 & 40 \\ 3 & 3/2 & 75 \\ 0 & 1 & 15 \end{vmatrix}}{3} = -15/2$$

Al ser $x_3 < 0$, el punto $(17,5; 15; -7,5; 0; 0)^T$ no es solución factible.

Intentemos ahora con (A_1, A_2, A_4) .

$$\det(A_1, A_2, A_4) = \begin{vmatrix} 1 & 2 & 0 \\ 3 & 3/2 & 1 \\ 0 & 1 & 0 \end{vmatrix} = -1 \Rightarrow (A_1, A_2, A_4) \text{ linealmente independientes}$$

$$x_1 = \frac{\begin{vmatrix} 40 & 2 & 0 \\ 75 & 3/2 & 1 \\ 15 & 1 & 0 \end{vmatrix}}{-1} = 10 \quad x_2 = \frac{\begin{vmatrix} 1 & 40 & 0 \\ 3 & 75 & 1 \\ 0 & 15 & 0 \end{vmatrix}}{-1} = 15 \quad x_4 = \frac{\begin{vmatrix} 1 & 2 & 40 \\ 3 & 3/2 & 75 \\ 0 & 1 & 15 \end{vmatrix}}{-1} = 45/2$$

El punto $(10; 15; 0; 22,5; 0)^T$ es un punto extremo. El valor de la función objetivo en dicho punto es $z = 20 \cdot 10 + 45 \cdot 15 = 875$.

En forma similar, debemos repetir este procedimiento hasta analizar todas las posibilidades que en este caso son $C_{5,3} = 10$.

Observemos que tomando como base a los vectores A_3 , A_4 y A_5 que acompañan a las variables de holgura, obtenemos una ***solución básica factible trivial***. En efecto, haciendo $x_1 = 0$ y $x_2 = 0$, resulta inmediato $x_3 = 40$, $x_4 = 75$ y $x_5 = 15$. Lógicamente para este punto $z = 0$.

La Tabla 2-1 resume los resultados obtenidos.

Tabla 2-1

Base	x_1	x_2	x_3	x_4	x_5	Observaciones	z
A_1, A_2, A_3	17,5	15	-7,5	0	0	No factible	
A_1, A_2, A_4	10	15	0	22,5	0	Factible Óptima	875
A_1, A_2, A_5	20	10	0	0	5	Factible	850
A_1, A_3, A_4	A_1, A_3, A_4 no son base						
A_1, A_3, A_5	25	0	15	0	15	Factible	500
A_1, A_4, A_5	40	0	0	-45	15	No factible	
A_2, A_3, A_4	0	15	10	52,5	0	Factible	675
A_2, A_3, A_5	0	50	-60	0	-35	No factible	
A_2, A_4, A_5	0	20	0	45	-5	No factible	
A_3, A_4, A_5	0	0	40	75	15	Factible	0

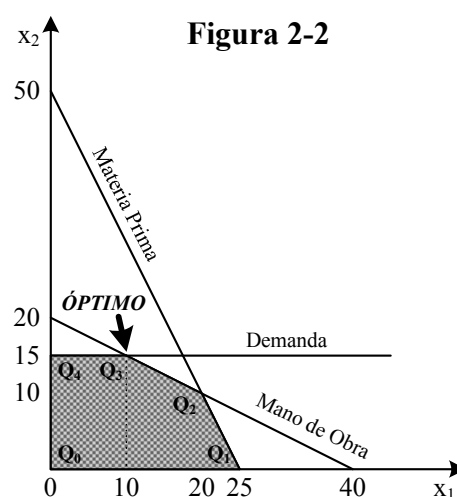
Veamos ahora cómo podemos representar en forma estándar el espacio de soluciones que obtuvimos al aplicar el método gráfico (Figura 2-2).

Puesto que x_3 , x_4 y x_5 son las holguras, cualquier punto perteneciente al polígono convexo de soluciones factibles puede ser representado en términos de las variables x_1 , x_2 , x_3 , x_4 y x_5 .

Si $x_3 = 0$ la restricción de *mano de obra* se reduce a $x_1 + 2x_2 = 40$, luego cualquier punto sobre la recta que incluye al segmento $[Q_2, Q_3]$ debe tener $x_3 = 0$. En forma similar, $x_4 = 0$ y $x_5 = 0$ definen los segmentos $[Q_1, Q_2]$ y $[Q_3, Q_4]$ respectivamente.

Si $x_3 = 15$ la restricción de *mano de obra* se transforma en $x_1 + 2x_2 = 25$ lo cual equivale a bajar con una paralela a $[Q_2, Q_3]$ hasta satisfacer esta última ecuación. Consecuentemente, $x_3 \geq 0$ junto con las restricciones de no negatividad $x_1 \geq 0$ y $x_2 \geq 0$ definen el área del triángulo comprendido entre los puntos $(0; 0)$, $(40; 0)$ y $(0; 20)$. Podemos hacer una interpretación similar para $x_4 \geq 0$ y $x_5 \geq 0$.

Evidentemente el punto $x_1 = 10$; $x_2 = 10$ puede representarse en términos de x_3 ; x_4 y x_5 , bastará con reemplazar en las respectivas ecuaciones.



NECESIDAD DE UN PROCEDIMIENTO ALTERNATIVO DE SOLUCIÓN

Desde el punto de vista conceptual el problema de la programación lineal ha quedado totalmente resuelto.

Sin embargo, el procedimiento antes descrito puede convertirse en una ardua tarea si tenemos en cuenta por ejemplo, que una aplicación pequeña con 20 variables y 10 restricciones podría tener $C_{20, 10} = 184.756$ puntos extremos. Aún empleando computadoras de última generación, el número de sistemas de ecuaciones lineales a resolver puede demandar tiempos de cálculo inadmisibles.

Afortunadamente, en 1947 *George B. Dantzig (1914 – 2005)* desarrolló un método denominado **Simplex**. El método o **algoritmo** Simplex, que estudiaremos en el capítulo siguiente, permite resolver problemas lineales sin necesidad de analizar el valor de la función económica en *todos* los puntos extremos. En la actualidad son numerosos los paquetes de software que hacen uso del Simplex y de sus variantes.