Tema 5: Sistemas de ecuaciones lineales.

1 Definiciones generales

Definición. 1.1 Una ecuación lineal con n incognitas es una expresión del tipo

$$a_1x_1 + \cdots + a_nx_n = b$$

donde $a_1, \ldots, a_n, b \in \mathbb{R}$ y $x_1, \ldots, x_n$ son variables.

Llamaremos sistema de ecuaciones lineales a un conjunto de ecs. lineales:

$$(S) : \left\{ \begin{array}{ccc} a_{11}x_1 + & \cdots & +a_{1n}x_n = b_1 \\ \vdots & \ddots & \vdots \\ a_{m1}x_1 + & \cdots & +a_{mn}x_n = b_m \end{array} \right.$$

La matriz $A = (a_{ij}) \in \mathcal{M}_{m\times n}$ se denomina matriz de coeficientes del sistema (S). La matriz

$$(A|\mathbf{b}) = \begin{bmatrix} a_{11} & \ldots & a_{1n} & b_1 \\ \vdots & \ldots & \vdots & \vdots \\ a_{m1} & \ldots & a_{mn} & b_m \end{bmatrix} \in \mathcal{M}_{m\times(n+1)}$$

se denomina matriz ampliada del sistema.

El vector $\mathbf{b} = (b_1, \ldots, b_m) \in \mathbb{R}^m$ se denomina término independiente. Si escribimos $\mathbf{x}= (x_1, \ldots, x_n)$ el sistema (S) puede expresarse como:

$$(S) : \; A\mathbf{x} = \mathbf{b}.$$

Diremos que (S) es homogéneo si $\mathbf{b} = \mathbf{0}$.

Definición. 1.2 Dado un S. E. L. $A\mathbf{x} = \mathbf{b}$ con $A \in \mathcal{M}_{m\times n}$, diremos que $\mathbf{a} \in \mathbb{R}^n$ es una solución del sistema si

$$A\mathbf{a} = \mathbf{b}.$$

Diremos que (S) es compatible si posee al menos una solución e incompatible si no poseee ninguna.

Un sistema se llamará compatible determinado si poseee una única solución y compatible indeterminado si posee más de una.

Teorema. 1.3 Sean L y L_0 la soluciones de $A\mathbf{x} = \mathbf{b}$ y $A\mathbf{x} = \mathbf{0}$, $A \in \mathcal{M}_{m\times n}$. Se verifica:

1. $L_0 \neq \emptyset$.
2. L_0 es una variedad lineal de $\mathbb{R}^n$
3. Si $\mathbf{a} \in L$, $L = \mathbf{a} + L_0$

Definición. 1.4 Dos sistemas se ecuaciones $(S) : A\mathbf{x} = \mathbf{b}$ y $(S') : A'\mathbf{x} = \mathbf{b}'$ se denominan equivalentes si poseen las mismas soluciones, es decir:

$$\forall\, \mathbf{a} \in \mathbb{R}^n, \quad A\mathbf{a} = \mathbf{b} \quad \Longleftrightarrow \quad A'\mathbf{a} = \mathbf{b}'.$$

Definición. 1.5 Un sistema $(S) : A\mathbf{x} = \mathbf{b}$ se dice triangular superior (resp. inferior) si su matriz de coeficientes, A, es triangular superior (resp. inferior).

Proposición. 1.6 Sea $A \in \mathcal{M}_{m\times n}$ y $C \in \mathcal{M}_n$ regular, entonces el sistema $(S): \ A\mathbf{x} = \mathbf{b}$ es equivalente a $(S'): \ (CA)\mathbf{x} = C\mathbf{b}$.

Corolario. 1.7 Si en un sistema

- Intercambiamos 2 ecuaciones.
- Multiplicamos una ecuación por un número $\alpha \neq 0$.
- Sumamos a una ecuación otra multiplicada por α.

obtenemos otro equivalente al primero.

Corolario. 1.8 Sea $(S): A\mathbf{x} = \mathbf{b}$ un S. E. L. con $A \in \mathcal{M}_{m\times n}$, entonces (S) es equivalente a un sistema en el cual la matriz ampliada es escalonada por filas.

Este último corolario es la base del método de Gauss para resolver sistemas de ecuaciones lineales, que describiremos un poco más adelante.

Corolario. 1.9 Sea $(S): A\mathbf{x} = \mathbf{b}$ un S. E. L. con $A \in \mathcal{M}_{m\times n}$, entonces (S) es equivalente a un sistema cuya matriz ampliada es escalonada canónica por filas.

2 Regla de Cramer. El teorema de Rouché-Frobenius

Definición. 2.1 Un S. E. L. $(S): \ A\mathbf{x} = \mathbf{b}$ se denomina cuadrado si $A \in \mathcal{M}_n$.

Un sistema $(S): \ A\mathbf{x} = \mathbf{b}$ se denomina de Cramer si es cuadrado y, además,

$$det(A) \neq 0.$$

Teorema. 2.2 (Regla de Cramer)
Si $(S): \ A\mathbf{x} = \mathbf{b}$, $A \in \mathcal{M}_n$, es un sistema de Cramer, entonces (S) tiene una única solución $\mathbf{a} = (\alpha_1, \ldots, \alpha_n)$, dada por

$$\alpha_i = \frac{1}{det(A)} \begin{vmatrix} a_{11} & \cdots & \overset{(i}{b_1} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & b_n & \cdots & a_{nn} \end{vmatrix} \quad \forall \ i = 1, \ldots, n.$$

Este teorema proporciona un método, la llamada regla de Cramer, para resolver un sistema de ecuaciones cuadrado compatible. Sin embargo, este método no resulta conveniente en la práctica, debido al excesivo número de operaciones que requiere. Debido a ello, suele utilizarse el Método de Gauss que consiste en lo siguiente:

Dada $A \in \mathcal{M}_n$ regular y $\mathbf{b} \in \mathbb{R}^n$, podemos encontar una matriz $P \in \mathcal{M}_n$ regular, producto de transformaciones elementales por filas, tal que

$$T = PA$$

es triangular superior. En consecuencia, por la proposición 1.6, el sistema de ecuaciones $(S): A\mathbf{x} = \mathbf{b}$ es equivalente al sistema

$$(S'): T\mathbf{x} = \mathbf{b}'$$

siendo $\mathbf{b}' = P\mathbf{b}$. Por tanto, las soluciones de (S) son las soluciones de (S'), que resulta mucho mas fácil resolver.

Número de operaciones

Utilizaremos las conocidas fórmulas (que pueden demostrarse por inducción):

$$\sum_{i=1}^{n} i = \frac{n(n+1)}{2}, \quad \sum_{i=1}^{n} i^2 = \frac{n(n+1)(2n+1)}{6}$$

- **Método de Gauss**

1. Triangularizar:
   Si suponemos que hemos triangularizado hasta la fila i, podemos suponer que $a_{ii} \neq 0$, pues si no intercambiamos dos filas que no supone ninguna operación. Para anular a_{i+1i} tenemos que realizar $\mathbf{f}_{i+1} - \frac{a_{i+1i}}{a_{ii}}\mathbf{f}_i$, llamando $\mathbf{f}_i$ a la fila i-ésima, lo que supone 1 división y $n-i+1$ multiplicaciones (quedan $n-(i-1)$ elementos en la fila i de la matriz del sistema, pero hay que descontar el de a_{i+1i} que sabemos que sale 0 y sumar el del término independiente, o sea en total $n-i+1$) y $n-i+1$ restas; y eso hay que hacerlo en $n-i$ filas, luego quedan en total $2(n-i+1)(n-i)+(n-i)$ operaciones (el primer sumando corresponde a las multiplicaciones y restas y el segundo a las divisiones). Por tanto, sumando las correspondientes a las n filas:

$$\sum_{i=1}^{n}[2(n-i+1)(n-i)+(n-i)] = \sum_{i=1}^{n}(n-i)\,[2(n-i)+3] = 2\sum_{i=1}^{n}(n-i)^2+\sum_{i=1}^{n}(n-i) =$$

$$2\frac{(n-1)n(2(n-1)+1)}{6}+\frac{3(n-1)n}{2} = \frac{4n^3+3n^2-7n}{6} \approx \frac{2}{3}n^3$$

2. Método de subida (sustitución regresiva)
   Para despejar x_i en la ecuación $a_{ii}x_i + a_{ii+1}x_{i+1} + \ldots + a_{in}x_n = b_i$ hace falta realizar $n-i$ multiplicaciones, $n-i$ sumas o restas y 1 división, luego en total tenemos:

$$\sum_{i=1}^{n}[2(n-i)+1] = 2\frac{(n-1)n}{2}+n = n^2$$

Luego para resolver el sistema por el método de Gauss el número de operaciones a realizar es:

$$\frac{4n^3+3n^2-7n}{6}+n^2 = \frac{4n^3+9n^2-7n}{6}$$

- **Regla de Cramer**
  Tenemos que calcular $n+1$ determinantes.

1. Calculando los determinantes por $\sum_{\sigma}(-1)^{\sigma}a_{1\sigma(1)}\ldots a_{n\sigma(n)}$ tenemos que realizar $n-1$ multiplicaciones y eso $n!$ veces y sumar los $n!$ sumandos, luego para el cálculo de cada determinante necesitamos $n!(n-1)+(n!-1)$ operaciones y eso para $n+1$ determinantes, y luego realizar las n divisiones, por tanto tenemos:

$$(n+1)\,[n!(n-1)+(n!-1)]+n = (n+1)!(n-1)+(n+1)!-n-1+n =$$

$$= (n+1)!(n-1+1)-1 = (n+1)!n-1$$

2. Usando eliminación gaussiana para calcular cada determinante, es decir, triangularizando cada determinante:
   Suponiendo que se ha triangularizado hasta la fila i, tenemos que calcular $\mathbf{f}_{i+1} - \frac{a_{i+1i}}{a_{ii}}\mathbf{f}_i$, luego tenemos que hacer 1 división, $n-i$ multiplicaciones y $n-i$ restas y eso en $n-i$ filas, luego nos queda:

$$\sum_{i=1}^{n}\left[2(n-i)^2+(n-i)\right] = \frac{4n^3-3n^2-n}{6}$$

Y ahora tenemos que hacer $n-1$ multiplicaciones para el cálculo de un determinante; como son $n+1$ determinantes y n divisiones tenemos:

$$(n+1)\left[\frac{4n^3-3n^2-n}{6}+n-1\right]+n$$

Para $n=25$, realizando 10^6 operaciones por segundo se tardaría:

- Por Gauss: $\frac{4\times 25^3+9\times 25^2-7\times 25}{6}=11325$ operaciones. Por tanto: $\frac{11325}{10^6}=0.011325$ segundos.
- Por Cramer: 1.00822×10^{28} operaciones. Luego sería en años: $\frac{1.00822\times 10^{22}}{36\times 24\times 60\times 60}=3.197\times 10^{14}$ años.
- Por Cramer, triangularizando los determinantes: $26\left[\frac{4\times 25^3-3\times 25^2-25}{6}+24\right]+25 = 263249$ operaciones, lo que supondría 0.263249 segundos.

Nota. 2.3 *Por tanto: En términos del número de operaciones que cada método requiere para resolver un sistema cuadrado de orden n, el método de Gauss resulta ser mucho más eficiente que la regla de Cramer. De manera aproximada, el método de Gauss requiere, aproximadamente,* $\frac{2}{3}n^3$ *operaciones mientras que la regla de Cramer exige, también de manera aproximada y suponiendo que los determinantes se calculan mediante eliminación gaussiana,* $\frac{2}{3}n^4$.

Teorema. 2.4 Dada $A\in\mathcal{M}_n$ el sistema $(S):\ A\mathbf{x}=\mathbf{b}$ es compatible determinado si y sólo si $det(A)\neq 0$. En otras palabras, un sistema cuadrado tiene una única solución si y sólo si es de Cramer.

Definición. 2.5 Dada $A\in\mathcal{M}_{m\times n}$ definimos:

- El espacio fila de A, $F(A)$, es la variedad lineal de $\mathbb{R}^n$ generada por las filas de A. $F(A)=\{\mathbf{x}^tA:\mathbf{x}\in\mathbb{R}^m\}$ o bien escritos por columnas como $F(A)=\{A^t\mathbf{x}:\mathbf{x}\in\mathbb{R}^m\}$.
- El espacio columna de A es la v. l. de $\mathbb{R}^m$, generada por las columnas de A. $C(A)=\{A\mathbf{x}:\mathbf{x}\in\mathbb{R}^n\}$.
- El espacio nulo de A se define como $N(A)=\{\mathbf{u}\in\mathbb{R}^n:\ A\mathbf{u}=\mathbf{0}\}$.

Proposición. 2.6 Dada $A\in\mathcal{M}_{m\times n}$, $F(A)$, $C(A)$ y $N(A)$ son variedades lineales.

Proposición. 2.7

1. Si A y B son equivalentes por filas, entonces $F(A)=F(B)$ y $N(A)=N(B)$.
2. Si A y B son equivalentes por columnas $C(A)=C(B)$

Teorema. 2.8 Si $A\in\mathcal{M}_{m\times n}$ entonces

$$\begin{aligned} n &= dim(N(A))+dim(C(A))\\ &= dim(N(A))+r(A)\end{aligned}$$

Teorema. 2.9 (de Rouché-Frobenius)
Sea $A\in\mathcal{M}_{m\times n}$, $\mathbf{b}\in\mathbb{R}^m$ y $(S):\ A\mathbf{x}=\mathbf{b}$. Sea $(A|\mathbf{b})$ la matriz ampliada de (S). Se verifica:

1) (S) es compatible si sólo si $r(A)=r((A|\mathbf{b}))$.

2) a) Si $\mathbf{b}=\mathbf{0}$, el conjunto de soluciones de (S) es una variedad lineal de dimensión $n-r$, siendo $r=r(A)$.

b) Si $\mathbf{b} \neq \mathbf{0}$ y (S) es compatible, toda solución de (S) es de la forma $\mathbf{x}_1 + \mathbf{x}_0$, donde $A\mathbf{x}_1 = \mathbf{b}$ y $A\mathbf{x}_0 = \mathbf{0}$.

Nota. 2.10 *Como consecuencia del teorema de Rouché-Frobenius, dado el sistema* $(S): \ A\mathbf{x} = \mathbf{b}$, *con* $A \in \mathcal{M}_{m\times n}$ *y* $\mathbf{b} \in \mathbb{R}^m$, *podemos asegurar que:*

$$(S) \text{ es compatible} \quad \Longleftrightarrow \quad r(A) = r((A|\mathbf{b})),$$

y además, en tal caso:

1) Si $r(A) = n$ *entonces* (S) *es compatible determinado.*

2) Si $r(A) < n$ *entonces* (S) *es compatible indeterminado.*

3 Ecuaciones de variedades lineales

1) Ecuaciones paramétricas:

Sea $W \subseteq \mathbb{R}^n$ una variedad lineal y $\mathcal{B} = \{\mathbf{a}_1, \ldots, \mathbf{a}_r\}$ una base de W. Dado $\mathbf{x} \in W$, existen $\lambda_1, \ldots, \lambda_r \in \mathbb{R}$ tales que

$$\mathbf{x} = \sum_{i=1}^{r} \lambda_i \mathbf{a}_i.$$

Si $\mathbf{x} = (x_1, \ldots, x_n)$ y, para cada $i = 1, \ldots, r$,

$$\mathbf{a}_i = (a_{1i}, \ldots, a_{ni}),$$

entonces,

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \lambda_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{bmatrix} + \cdots + \lambda_r \begin{bmatrix} a_{1r} \\ a_{2r} \\ \vdots \\ a_{nr} \end{bmatrix} = \begin{bmatrix} \lambda_1 a_{11}+ & \cdots & \lambda_r a_{1r} \\ \lambda_1 a_{21}+ & \cdots & \lambda_r a_{2r} \\ \vdots & & \vdots \\ \lambda_1 a_{n1}+ & \cdots & \lambda_r a_{nr} \end{bmatrix}.$$

Por tanto, $\mathbf{x} \in W$ si y sólo si existen $\lambda_1, \ldots, \lambda_r \in \mathbb{R}$ verificando:

$$(1) \left\{ \begin{array}{lclcl} x_1 & = & \lambda_1 a_{11}+ & \cdots & \lambda_r a_{1r} \\ x_2 & = & \lambda_1 a_{21}+ & \cdots & \lambda_r a_{2r} \\ \vdots & & & & \\ x_n & = & \lambda_1 a_{n1}+ & \cdots & \lambda_r a_{nr} \end{array} \right.$$

Estas son unas ecuaciones paramétricas de W.

Nota. 3.1 *Las ecuaciones paramétricas* **NO** *son únicas.*

2) Ecuaciones implícitas:

Continuemos con la variedad W y su base $\mathcal{B}$ consideradas anteriormente. Las ecs. paramétricas de W dadas por (1), pueden ser escritas como un S. E. L.

$$(E): \ A\lambda = \mathbf{x},$$

siendo $\lambda = (\lambda_1, \ldots, \lambda_r)$ y $A = [\mathbf{a}_1 | \ldots | \mathbf{a}_r] \in \mathcal{M}_{n\times r}$. Es obvio que

$$\begin{array}{rcl} \mathbf{x} \in W & \Longleftrightarrow & (E) \text{ es compatible} \\ & \Longleftrightarrow & r(A) = r((A|\mathbf{x})). \end{array}$$

Sea $P \in \mathcal{M}_n$ regular tal que PA es triangular superior. Puesto que $r(A) = r$, tenemos

$$PA = \left[\begin{array}{c} T_r \\ \hline \Theta \end{array}\right]$$

siendo $T_r \in \mathcal{M}_r$ regular. Si descomponemos P en dos bloques,

$$P\mathbf{x} = \left[\begin{array}{c} P_r \\ \hline P_{n-r} \end{array}\right] \mathbf{x} = \left[\begin{array}{c} P_r\mathbf{x} \\ \hline P_{n-r}\mathbf{x} \end{array}\right].$$

Por tanto, multiplicando por bloques:

$$P \cdot (A|\mathbf{x}) = (PA|P\mathbf{x}) = \left[\begin{array}{c|c} T_r & P_r\mathbf{x} \\ \hline \Theta & P_{n-r}\mathbf{x} \end{array}\right]$$

En consecuencia,

$$r((A|\mathbf{x})) = r(A) = r \quad \Longleftrightarrow \quad P_{n-r}\mathbf{x} = \mathbf{0}.$$

Lo que nos da como ecuaciones implícitas de W el S. E. L.:

$$P_{n-r}\mathbf{x} = \mathbf{0}.$$

Nota. 3.2 *Al igual que antes las ecuaciones implícitas de W* **NO** *son únicas. Además, si $dim(W) = r$ entonces $W \subseteq \mathbb{R}^n$ tiene $n - r$ ecs. implícitas linealmente independientes.*

4 Operaciones con variedades

Definición. 4.1 Dadas $L_1,\ L_2 \subseteq \mathbb{R}^n$ variedades lineales, definimos:

$$\begin{aligned} L_1 \cap L_2 &= \{\mathbf{u} \in \mathbb{R}^n : \ \mathbf{u} \in L_1 \text{ y } \mathbf{u} \in L_2\} \\ L_1 + L_2 &= \{\mathbf{u} \in \mathbb{R}^n : \ \exists\ \mathbf{v}_1 \in L_1,\ \exists\ \mathbf{v}_2 \in L_2,\ \mathbf{u} = \mathbf{v}_1 + \mathbf{v}_2\}. \end{aligned}$$

$L_1 \cap L_2$ se denomina variedad intersección de L_1 y L_2. $L_1 + L_2$ se denomina variedad suma.

Proposición. 4.2 Si $L_1,\ L_2 \subseteq \mathbb{R}^n$ son v. l., entonces $L_1 \cap L_2$ y $L_1 + L_2$ también son variedades lineales. Además, $L_1 + L_2$ es la menor variedad lineal que contiene a L_1 y a L_2.

Definición. 4.3 Dadas $L,\ L_1,\ L_2 \subseteq \mathbb{R}^n$, diremos que L es suma directa de L_1 y L_2, y lo denotamos por $L = L_1 \oplus L_2$, si

$$L_1 + L_2 = L \quad \text{y} \quad L_1 \cap L_2 = \{\mathbf{0}\}.$$

Proposición. 4.4 Dadas $L,\ L_1,\ L_2 \subseteq \mathbb{R}^n$, entonces $L = L_1 \oplus L_2$ si y sólo si

$$\forall\ \mathbf{x} \in L,\ \exists!\ \mathbf{x}_1 \in L_1,\ \exists!\ \mathbf{x}_2 \in L_2 : \quad \mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2.$$

Proposición. 4.5 Dada $L \subseteq \mathbb{R}^n$, una variedad lineal, existe una variedad $L' \subseteq \mathbb{R}^n$ tal que:

$$\mathbb{R}^n = L \oplus L'.$$

Dicha variedad L' se denomina v. l. complementaria de L.

Nota. 4.6 *La variedad complementaria no es única.*

Teorema. 4.7 (Fórmula de la dimensión)
Dadas las variedades lineales $L_1,\ L_2 \in \mathbb{R}^n$, se verifica:

$$dim(L_1 + L_2) = dim(L_1) + dim(L_2) - dim(L_1 \cap L_2).$$

En particular, si $L = L_1 \oplus L_2$, entonces

$$dim(L) = dim(L_1) + dim(L_2).$$