1 Plantejament del problema

Notació. Al llarg de l'exposició, un element a_i^j denotarà l'element de la columna i i de la fila j d'una matriu. Un element u_i denotarà un vector columna, mentre que b^j representarà la coordenada j d'un vector b. L'únic cas en què un superíndex representarà un exponent és en el cas A^{-1} (matriu inversa de A) o $(\det A)^{-1}$ (invers de $\det A$).

Volem resoldre el següent problema: suposem que tenim un sistema d'equacions lineals

$$\left\{ \begin{array}{rcl} a_1^1 x^1 + \ldots + a_n^1 x^n & = & b^1 \\ \cdots\cdots\cdots\cdots\cdots & = & \cdots \\ a_1^m x^1 + \ldots + a_n^m x^n & = & b^m \end{array} \right.$$

on a_i^j, b^j són elements coneguts d'un cos K. Es tracta de trobar les n-tuples $(x^1, \ldots, x^n) \in K^n$ que satisfan simultàniament totes aquestes equacions. Si prenem

$$A = \begin{pmatrix} a_1^1 & \cdots & a_n^1 \\ \vdots & \ddots & \vdots \\ a_1^m & \cdots & a_n^m \end{pmatrix}, \quad b = \begin{pmatrix} b^1 \\ \vdots \\ b^m \end{pmatrix}, \quad x = \begin{pmatrix} x^1 \\ \vdots \\ x^n \end{pmatrix},$$

llavors podem escriure el sistema anterior en forma matricial: $Ax = b$. Aquest mateix problema es pot plantejar d'una forma equivalent en el llenguatge de les aplicacions lineals: siguin E, F dos K-espais vectorials amb bases $u_1, \ldots, u_n$ i $v_1, \ldots, v_m$ respectivament. Sabem que existeix una única aplicació lineal $f : E \longrightarrow F$ que té com a matriu associada en aquestes bases la matriu A. Designem per b el vector $b = b^1 v_1 + \ldots + b^m v_m \in F$ i per x el vector $x = x^1 u_1 + \ldots + x^n u_n \in E$. Llavors, la condició $Ax = b$ equival a $f(x) = b$, i el problema es converteix en trobar les antiimatges del vector b per l'aplicació f.

En qualsevol de les dues interpretacions, el nostre objectiu és:

1. Saber quan el problema té solució i quan no.
2. Saber quantes solucions té.
3. Donar un mètode per trobar totes les solucions.

Per resoldre aquestes qüestions farem servir la interpretació que ens sigui més còmoda en cada cas. En general, els raonaments teòrics són més simples en el llenguatge de les aplicacions lineals, mentre que la resolució de casos concrets és més còmoda amb el llenguatge de matrius.

2 Existència de solucions

En el problema plantejat en l'apartat anterior tenim que

$$\begin{aligned}\exists\, x \in E \text{ tal que } f(x) = b \Leftrightarrow b \in \mathrm{Im} f \Leftrightarrow \langle f(u_1), \dots, f(u_n)\rangle = \langle f(u_1), \dots, f(u_n), b\rangle \\ \Leftrightarrow \dim \mathrm{Im} f = \dim\langle f(u_1), \dots, f(u_n), b\rangle,\end{aligned}$$

ara bé, $\dim \mathrm{Im} f$ és el que anomenem $\mathrm{rang} f$, que coincideix amb $\mathrm{rang} A$ (perquè les columnes de A no són altra cosa que les coordenades de $f(u_1), \dots, f(u_n)$ posades en columna). De la mateixa forma, $\dim\langle f(u_1), \dots, f(u_n), b\rangle = \mathrm{rang}(A, b)$, on (A, b) representa la matriu que s'obté afegint a A una columna formada per les coordenades de b. Així, en el llenguatge de matrius, el sistema $Ax = b$ té solució si i només si $\mathrm{rang} A = \mathrm{rang}(A, b)$.

Ara, suposem que x_0 és una solució particular del sistema, això és, $f(x_0) = b$. Sigui x una altra solució qualsevol; aleshores, per linealitat de f podem escriure:

$$f(x) = f(x_0) \Rightarrow f(x - x_0) = \vec{0} \Rightarrow x - x_0 \in \mathrm{Ker} f \Rightarrow x = x_0 + u, \text{ amb } u \in \mathrm{Ker} f,$$

de manera que, fent un abús de notació, podem escriure que la solució general és $x = x_0 + \mathrm{Ker} f$, és a dir, tota solució s'obté sumant un vector del nucli a la solució particular, i per tant, hi haurà tantes solucions com vectors de $\mathrm{Ker} f$ (en termes de dimensions, el conjunt de solucions forma un subespai afí que té la mateixa dimensió que $\mathrm{Ker} f$). En particular, la solució serà única si $\mathrm{Ker} f = \{\vec{0}\}$, però aleshores

$$\dim E = \dim \mathrm{Ker} f + \dim \mathrm{Im} f \Rightarrow n = \mathrm{rang} A,$$

de manera que la solució és única si i només si el rang de A coincideix amb el nombre de columnes de A, això és, amb el nombre d'incògnites del sistema. Aquests resultats s'agrupen sota el nom de teorema de Rouché-Frobenius (tot i que en altres països se'l coneix com a teorema de Rouché-Capelli, Rouché-Fontené o Kronecker-Capelli):

Teorema de Rouché-Frobenius. *Sigui $Ax = b$ l'expressió matricial d'un sistema lineal de m equacions amb n incògnites. Llavors, el sistema té solució si i només si* $\mathrm{rang} A = \mathrm{rang}(A, b)$, *i la solució és única si aquest rang coincideix amb el nombre n d'incògnites.* □

Quan un sistema té una única solució es diu que és *compatible determinat*; si no és única, se'n diu *compatible indeterminat*; quan no té solució, s'anomena *incompatible.*

Observació. Tal com hem vist, la solució general d'un sistema d'equacions és $x = x_0 + \mathrm{Ker} f$, on x_0 n'és una solució particular. Els vectors de $\mathrm{Ker} f$ tenen per coordenades les solucions del sistema homogeni associat, $Ax = \vec{0}$, per tant, també podem dir que les solucions de $Ax = b$ s'obtenen sumant a una solució particular les solucions del sistema homogeni associat.

3 Regla de Cramer

Donarem ara un mètode per trobar les solucions d'un sistema d'equacions lineals en un cas molt particular: suposem que $Ax = b$ on A és una matriu quadrada i singular, això és, A és $n \times n$ i $\det A \neq 0$. Dit de forma equivalent, en el llenguatge de les aplicacions lineals, $f(x) = b$, on f és un isomorfisme. En aquest cas, la solució existeix i és única, i fent ús de la matriu inversa A^{-1} (que també sabem que existeix en ser A regular), tenim que $x = A^{-1}b$. Si escrivim $A = (a_i^j)$, $A^T = (a_j^i)$ la seva transposada, i (X_j^i) la matriu dels adjunts d'aquesta última, sabem que $A^{-1} = (X_j^i)(\det A)^{-1}$ com a conseqüència de la regla de Laplace. Llavors podem escriure

$$x = A^{-1}b = (\det A)^{-1} \begin{pmatrix} X_1^1 & \cdots & X_1^n \\ \vdots & \ddots & \vdots \\ X_n^1 & \cdots & X_n^n \end{pmatrix} \begin{pmatrix} b^1 \\ \vdots \\ b^n \end{pmatrix} \Rightarrow x^i = (\det A)^{-1} \sum_{j=1}^{n} X_i^j b^j,$$

però l'expressió $\sum_{j=1}^{n} X_i^j b^j$ coincideix amb el desenvolupament del determinant de A per la columna i, havent-la substituïda pel vector b, això és, $\det(a_1, \ldots, a_{i-1}, b, a_{i+1}, \ldots, a_n)$:

$$x^i = \frac{\det(a_1, \ldots, a_{i-1}, b, a_{i+1}, \ldots, a_n)}{\det A}, \quad i = 1, \ldots, n.$$

Aquesta fórmula es coneix amb el nom de *regla de Cramer*.

4 Resolució d'un sistema general d'equacions lineals

Donat el sistema general $Ax = b$ del principi,

$$\left\{ \begin{array}{rcl} a_1^1 x^1 + \ldots + a_n^1 x^n & = & b^1 \\ \cdots\cdots\cdots\cdots\cdots & = & \cdots \\ a_1^m x^1 + \ldots + a_n^m x^n & = & b^m \end{array} \right.$$

suposem que té solució, això és, $\mathrm{rang}A = \mathrm{rang}(A, b) = r$. Sabem que el rang d'una matriu coincideix amb el més gran dels ordres dels menors de A amb determinant no nul. Llavors, reordenant les equacions i les incògnites convenientment podem suposar que

$$M = \begin{vmatrix} a_1^1 & \cdots & a_r^1 \\ \vdots & \ddots & \vdots \\ a_1^r & \cdots & a_r^r \end{vmatrix} \neq 0.$$

Gràcies al rang, sabem que a la matriu (A, b), les $m - r$ últimes files són combinació lineal de les anteriors, és a dir, en el sistema donat, les $m - r$ últimes equacions són combinació lineal de les r primeres. Per tant, el sistema original té exactament les mateixes solucions que

$$\left\{ \begin{array}{rcl} a_1^1 x^1 + \ldots + a_n^1 x^n & = & b^1 \\ \cdots\cdots\cdots\cdots\cdots & = & \cdots \\ a_1^r x^1 + \ldots + a_n^r x^n & = & b^r \end{array} \right.$$

i llavors n'hi ha prou de trobar les solucions d'aquest sistema més petit. Si l'escrivim d'aquesta forma:

$$\left\{ \begin{array}{rcl} a_1^1 x^1 + \ldots + a_r^1 x^r & = & b^1 - a_{r+1}^1 x^{r+1} + \ldots + a_n^1 x^n \\ \cdots\cdots\cdots\cdots\cdots & = & \cdots \\ a_1^r x^1 + \ldots + a_r^r x^r & = & b^r - a_{r+1}^r x^{r+1} + \ldots + a_n^r x^n \end{array} \right.$$

llavors per a cada valor arbitrari que donem a $x^{r+1}, \ldots, x^n$ obtenim un sistema de r equacions i r incògnites, amb determinant de la matriu de coeficients igual a $M \neq 0$. Per tant, es pot aplicar la regla de Cramer i obtenir uns valors únics per a $x^1, \ldots, x^r$. Com tenim $n - r$ incògnites a què podem donar valors arbitraris, diem que el sistema té $n - r$ *graus de llibertat*.

Tanmateix, la regla de Cramer no és eficient per resoldre un sistema genèric donat l'elevat nombre d'operacions que cal fer per calcular els $r+1$ determinants (un per a cada incògnita, i $\det A$), sobretot si r és gran. Seguidament veiem un mètode molt més eficient per portar a la pràctica.

5 Mètode de Gauss-Jordan

La justificació teòrica d'aquest mètode rau en uns raonaments molt simples. Sigui $f : E \longrightarrow F$ una aplicació lineal i siguin $u_1, \ldots, u_n$ i $v_1, \ldots, v_m$ bases de E i F respectivament. Denotem per A la matriu associada a f en aquestes bases i $b \in F$ amb coordenades $(b^1, \ldots, b^m)$ (vector columna). Com sempre, volem trobar els $x \in E$ tals que $f(x) = b$. Un canvi de base en F dóna lloc a un canvi en la matriu associada A i en les coordenades de b, mentre que les coordenades de x resten invariables. Per tant, si A' i b' són la nova matriu associada i les noves coordenades de b, les solucions del sistema $A'x = b'$ són exactament les mateixes que les del sistema original $Ax = b$. Així doncs, ens proposem fer canvis en la base de F de manera que la matriu A' sigui prou simple perquè el fet de trobar la solució del sistema no impliqui cap càlcul. Els canvis que efectuarem sobre la base de F seran de tres tipus:

1. Permutació de l'ordre dels vectors de la base de F. Naturalment, aleshores, les coordenades de

$$\begin{aligned} f(u_i) &= a_i^1 v_1 + \ldots + a_i^m v_m, \\ b &= b^1 v_1 + \ldots + b^m v_m \end{aligned}$$

   queden permutades, cosa que equival al fet que en (A, b) les files quedin permutades.

2. Substitució d'un vector v_j de la base per kv_j amb $k \neq 0$. Llavors,

$$\begin{aligned} f(u_i) &= a_i^1 v_1 + \ldots + (a_i^j k^{-1}) k v_j + \ldots + a_i^m v_m, \\ b &= b^1 v_1 + \ldots + (b^j k^{-1}) k v_j + \ldots + b^m v_m, \end{aligned}$$

   és a dir, a la matriu (A, b) la fila j queda multiplicada per k^{-1}.

3. Substitució d'un vector v_j de la base per $v_j + kv_h$ amb $h \neq j$. Llavors,

$$\begin{aligned} f(u_i) &= a_i^1 v_1 + \ldots + a_i^j (v_j + k v_h) + \ldots + (a_i^h - a_i^j k) v_h + \ldots + a_i^m v_m, \\ b &= b^1 v_1 + \ldots + b^j (v_j + k v_h) + \ldots + (b^h - b^j k) v_h + \ldots + b^m v_m, \end{aligned}$$

   és a dir, a la matriu (A, b), a la fila h se li resta la fila j multiplicada per k.

Fent canvis d'aquests tres tipus i permutant, si convé, l'ordre de les incògnites, cosa que equival a permutar l'ordre de les n primeres columnes de (A, b), podem obtenir una matriu de la forma

$$\left(\begin{array}{cccccccc|c} 1 & 0 & \cdots & 0 & c_{r+1}^1 & \cdots & c_n^1 & d^1 \\ 0 & 1 & \cdots & 0 & c_{r+1}^2 & \cdots & c_n^2 & d^2 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & c_{r+1}^r & \cdots & c_n^r & d^r \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & d^{r+1} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & d^m \end{array}\right),$$

de manera que el sistema és compatible si i només si $d^{r+1} = \ldots = d^m = 0$. En tal cas, la solució general és $x^i = d^i - c_{r+1}^i x^{r+1} - \ldots - c_n^i x^n$ per a cada $i = 1, \ldots, r$, i els x^i per a $i = r + 1, \ldots, n$ prenen valors arbitraris.

A la pràctica, per aplicar el mètode de Gauss, els canvis en la matriu (A, b) es fan de la següent manera: si la primera columna és tota 0 es passa al lloc n. Si hi ha un element no nul (idealment, es cerca un 1), es permuten les files de manera que aquesta quedi la primera. Amb un canvi del tipus 2 s'aconsegueix que aquest element sigui un 1 (en cas que no ho fos ja), i amb canvis del tipus 3 es pot aconseguir que la resta de la columna sigui 0. Suposem que ja tenim h columnes en la forma desitjada. Si a la columna $h+1$ els elements de les files $h+1, \ldots, n$ són tots 0, la posem al lloc n. En cas contrari, col·loquem un element no nul (idealment un 1) en la fila $h+1$ permutant les dues files. Amb canvis del tipus 2 i 3 podem aconseguir que aquest element sigui un 1 (si no ho era ja) i la resta de la columna sigui 0. Observem que procedint d'aquesta manera, les columnes anteriors no varien, això és, no destruïm els zeros que ja havíem fet. El procés pot continuar fins a obtenir una matriu com la que hem descrit més amunt.

6 Càlcul de la matriu inversa

Com a exemple d'aplicació del mètode de Gauss, anem a veure com es pot calcular la inversa d'una matriu. Ja sabem que $A^{-1} = (X_j^i)(\det A)^{-1}$, però a la pràctica, si l'ordre de A és gran, es necessiten moltes operacions per calcular la matriu dels adjunts.

Donada una matriu A quadrada ($n \times n$) es tracta de trobar, si existeix, una matriu $B = (b_i^j)$ que faci $AB = I$, on I és la matriu identitat d'ordre n. Això equival a buscar les n columnes $b_i = (b_i^1, \ldots, b_i^n)$ de B tals que $Ab_i = e_i$, on $e_i = (0, \ldots, 1, \ldots, 0)$ és la columna i de I. En altres paraules, hem de resoldre els n sistemes d'equacions $Ax = e_i$, $i = 1, \ldots, n$, tots amb la mateixa matriu de coeficients A. Si considerem la matriu ampliada $(A, e_1, \ldots, e_n)$ i apliquem el mètode de Gauss, només cal fer les transformacions una sola vegada. Observem que $(e_1, \ldots, e_n)$ és precisament la matriu identitat. Partim doncs del sistema (A, I) i un cop fetes les transformacions haurem arribat a una matriu del tipus (I, B):

$$\left(\begin{array}{ccc|ccc} 1 & \cdots & 0 & b_1^1 & \cdots & b_n^1 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 1 & b_1^n & \cdots & b_n^n \end{array}\right).$$

Per tant, la columna i de la matriu B és la solució del sistema $Ax = e_i$, de manera que B és la matriu A^{-1} buscada. Naturalment, si A no té inversa no podrem transformar A en I, i algun dels sistemes $Ax = e_i$ serà incompatible.

Exemple. Calculem la inversa de

$$A = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

Partim de (A, I) i apliquem el mètode de Gauss-Jordan per acabar obtenint (I, B):

$$\left(\begin{array}{ccc|ccc} 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \end{array}\right) \Rightarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \end{array}\right) \Rightarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & -1 & 0 & -1 & 1 \end{array}\right) \Rightarrow$$

$$\Rightarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & -2 & -1 & -1 & 1 \end{array}\right) \Rightarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & \frac{1}{2} & \frac{1}{2} & \frac{-1}{2} \end{array}\right) \Rightarrow$$

$$\Rightarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & \frac{1}{2} & \frac{-1}{2} & \frac{1}{2} \\ 0 & 0 & 1 & \frac{1}{2} & \frac{1}{2} & \frac{-1}{2} \end{array}\right) \Rightarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{-1}{2} & \frac{1}{2} & \frac{1}{2} \\ 0 & 1 & 0 & \frac{1}{2} & \frac{-1}{2} & \frac{1}{2} \\ 0 & 0 & 1 & \frac{1}{2} & \frac{1}{2} & \frac{-1}{2} \end{array}\right).$$

Per tant,

$$A^{-1} = \frac{1}{2} \begin{pmatrix} -1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{pmatrix}.$$