PCSI - mathématiques

Systèmes linéaires

1 Généralités

n et p sont deux entiers naturels non nuls. $\mathbb{K}$ désigne le corps des réels ou des complexes.

Définition 1 *Un* système linéaire *de n équations à p inconnues à coefficients dans $\mathbb{K}$ est un ensemble de conditions de la forme*

$$(S)\left\{\begin{array}{ccccccc} a_{1,1}x_1 & + & a_{1,2}x_2 & +\ldots+ & a_{1,p}x_p & = & b_1 \\ a_{2,1}x_1 & + & a_{2,2}x_2 & +\ldots+ & a_{2,p}x_p & = & b_2 \\ \vdots & & & & \vdots & & \vdots \\ a_{n,1}x_1 & + & a_{n,2}x_2 & +\ldots+ & a_{n,p}x_p & = & b_n \end{array}\right.$$

où $A=(a_{i,j})_{\substack{1\leq i\leq n\\1\leq j\leq p}}\in\mathcal{M}_{n,p}(\mathbb{K})$ et $b_1,\ \ldots,\ b_n\in\mathbb{K}$.

- A est la *matrice* de (S).
- Le *rang* de (S) est $r=\operatorname{rg}(S)=\operatorname{rg}(A)$.
- On associe à (S) le *système homogène* (S_0) obtenu en remplaçant les seconds membres b_1, ..., b_n par 0.
- Une *solution* de (S) est un p-uplet de scalaires $(x_1,\ \ldots,\ x_p)$ vérifiant les n équations.
- (S) est *compatible* s'il admet au moins une solution. Il est *incompatible* dans le cas contraire.
  Notons qu'un système homogène est toujours compatible, car il admet au moins la solution triviale (nulle).
- *Résoudre* (S), c'est trouver toutes ses solutions.

Quelques remarques générales permettent de dégager une méthode d'étude.

1.1 Interprétations de (S)

1.1.1 Matricielle

Notons pour x_1, ..., $x_p\in\mathbb{K}$: $X=\begin{pmatrix}x_1\\\vdots\\x_p\end{pmatrix}$ et $B=\begin{pmatrix}b_1\\\vdots\\b_n\end{pmatrix}$. Alors

$$(S)\Longleftrightarrow AX=B$$

(forme matricielle de (S)).
Dans cette notation abrégée, $(S_0)\Longleftrightarrow AX=0$.

1.1.2 Vectorielle

Notons pour $j=1$, ..., p, $C_j=\begin{pmatrix}a_{1,j}\\\vdots\\a_{n,j}\end{pmatrix}$ ($j^{\text{ième}}$ colonne de la matrice A). Alors

$$(S)\Longleftrightarrow x_1C_1+\ldots+x_pC_p=B$$

ce qui montre que :

- (S) est compatible *ssi* B est c.l. de $(C_1,\ \ldots,\ C_p)$ (les solutions de (S) correspondant aux telles c.l.) ;
- Les solutions éventuelles de (S_0) correspondent aux relations de dépendance linéaire entre C_1, ..., C_p.

1.1.3 En termes d'application linéaire

Notons $\mathcal{E}_n$ et $\mathcal{E}_p$ les bases canoniques de $\mathbb{K}^n$ et $\mathbb{K}^p$ respectivement. Soient les vecteurs $x\in\mathbb{K}^p$ et $b\in\mathbb{K}^n$ représentés par X et B dans $\mathcal{E}_p$ et $\mathcal{E}_n$ respectivement.

Soit enfin $f\in L(\mathbb{K}^p,\mathbb{K}^n)$ l'unique application linéaire telle que $\operatorname{mat}(f\ ;\ \mathcal{E}_n,\mathcal{E}_p)=A$. Alors

$$(S)\Longleftrightarrow f(x)=b$$

donc la recherche des solutions de (S) est celle des antécédents de b par f. (S_0) a pour ensemble de solutions $\ker(f)$: c'est donc un sev de $\mathbb{K}^n$. En outre, si $X_0=\operatorname{mat}(x_0\ ;\ \mathcal{E}_p)$ est une solution particulière de (S) alors

$$\begin{array}{rcl} X & = & \operatorname{mat}(x\ ;\ \mathcal{E}_p)\ \text{solution de}\ (S) \\ & \Leftrightarrow & f(x)=b=f(x_0) \\ & \Leftrightarrow & f(x)-f(x_0)=0=f(x-x_0) \\ & \Leftrightarrow & Y=X-X_0\ \text{solution de}\ (S_0) \end{array}$$

de sorte que les solutions de (S) s'obtiennent en ajoutant à l'une d'entre elles toutes les solutions de (S_0). Elles forment donc un sous-espace *affine* de $\mathbb{K}^p$.

D'autre part, $\operatorname{rg}(f)=\operatorname{rg}(A)=\operatorname{rg}(S)=r$ donc le th. du rang pour f s'écrit $p=r+\dim(\ker f)$ donc :

Proposition 1 *La dimension de l'espace affine (*resp. *vectoriel) des solutions de (S) (*resp.*(S_0)) est $\boldsymbol{p-r}$.*

1.1.4 En termes de formes linéaires

Notons pour $i=1$, ..., n, φ_i la forme linéaire[1] sur $\mathbb{K}^p$ définie par $\varphi_i(X)=a_{i,1}x_1+\ldots+a_{i,p}x_p$ si $X=\begin{pmatrix}x_1\\\vdots\\x_p\end{pmatrix}$. Alors

$$(S)\Longleftrightarrow\left\{\begin{array}{ccc}\varphi_1(X) & = & b_1\\ \vdots & & \vdots\\ \varphi_n(X) & = & b_n\end{array}\right.$$

et pour $i=1$, ..., n, $(\mathcal{H}_i)\mid\varphi_i(X)=b_i$ définit un hyperplan affine de $\mathbb{K}^p$ (dont la direction est l'hyperplan vectoriel $\ker(\varphi_i)$). La résolution de (S) consiste à étudier l'intersection de ces n hyperplans.

[1] Une *forme linéaire* sur $\mathbb{K}^p$ est une application linéaire de $\mathbb{K}^p$ dans $\mathbb{K}$. Si ce n'est pas l'application nulle, le th. du rang montre que son noyau est de dimension $p-1$: on l'appelle un *hyperplan* de $\mathbb{K}^p$.

2 Systèmes de Cramer

On se limite dans ce paragraphe au cas de systèmes ayant autant d'équation que d'inconnues : $p = n$. La matrice A est donc carrée de taille n.

2.1 Condition de compatibilité

Théorème 1 *Il y a équivalence entre les conditions :*

(1) *A est inversible ;*

(2) $\det(A) \neq 0$ *;*

(3) $\mathrm{rg}(S) = n$ *;*

(4) *Pour tout second membre B, (S) admet* au moins *une solution ;*

(5) *Pour tout second membre B, (S) admet* au plus *une solution ;*

(6) *Pour tout second membre B, (S) admet* exactement *une solution ;*

(7) *(S_0) admet comme unique solution la solution triviale.*

Définition 2 *Le système (S) est* de Cramer *si l'une des conditions équivalentes précédentes est remplie.*

La résolution d'un système de Cramer revient donc précisément à inverser sa matrice. C'est même la méthode la plus générale de calcul d'un inverse, *cf.* 2.3.

2.2 Formules de Cramer

Supposons le système (S) de Cramer. On note toujours C_1, ..., C_n les colonnes de la matrice A. Si l'on désigne par $(x_1^*,\ ...,\ x_n^*)$ l'unique n-uplet de $\mathbb{K}^n$ solution de (S) alors on peut obtenir cette solution par le :

Théorème 2 (formules de Cramer) *Pour* $i = 1$, ..., n,

$$x_i^* = \frac{\det(C_1,\ ...,\ C_{i-1},\ B,\ C_{i+1},\ ...,\ C_n)}{\det(A)}.$$

L'intérêt de ces formules est surtout théorique (p. ex. pour montrer que si les coefficients du système sont des fonctions continues d'une certaine variable t, les solutions sont également fonctions continues de t).

En pratique, après s'être assuré que (S) est de Cramer, il vaut souvent mieux résoudre (S) en éliminant progressivement les inconnues et en rendant le système triangulaire, ce qui permet un calcul de proche en proche (méthode du pivot) :

$$(S) \Leftrightarrow \left\{\begin{array}{ccccccc} a_{1,1}x_1 & + & a_{1,2}x_2 & +\dots+ & a_{1,n}x_n & = & b_1 \\ & & a'_{2,2}x_2 & +\dots+ & a'_{2,n}x_n & = & b'_2 \\ & & & \ddots & \vdots & & \vdots \\ & & & & a'_{n,n}x_n & = & b'_n \end{array}\right.$$

La résolution d'un tel système de Cramer triangulaire est alors triviale : on calcule $x_n = \frac{b'_n}{a'_{n,n}}$, puis x_{n-1} grâce à l'équation précédente et ainsi de suite jusqu'à x_1.

Exemple 1 *systèmes de* Cramer

1. $\left\{\begin{array}{l} 2x + y + \ \ z = -1 \\ \ \ x - y \qquad = -1 \\ 3x + y + 2z = -1 \end{array}\right.$

2. $\left\{\begin{array}{l} x + y + z \qquad = 0 \\ x + y \qquad + t = 1 \\ x \qquad + z + t = 2 \\ \qquad y + z + t = 3 \end{array}\right.$

2.3 Inversion d'une matrice carrée

Soit $A \in \mathcal{M}_n(\mathbb{K})$ inversible. Pour calculer A^{-1} on peut former le système (de Cramer) $(S) : AX = Y$ soit

$$(S) \left\{\begin{array}{ccccccc} a_{1,1}x_1 & + & a_{1,2}x_2 & + ... + & a_{1,n}x_n & = & y_1 \\ a_{2,1}x_1 & + & a_{2,2}x_2 & + ... + & a_{2,n}x_n & = & y_2 \\ \vdots & & & & & & \vdots \\ a_{n,1}x_1 & + & a_{n,2}x_2 & + ... + & a_{n,n}x_n & = & y_n \end{array}\right.$$

La résolution de (S) donne l'expression des inconnues x_1, ..., x_n en fonction des seconds membres y_1, ..., y_n soit les relations (linéaires)

$$\left\{\begin{array}{ccccccc} b_{1,1}y_1 & + & b_{1,2}y_2 & + ... + & b_{1,n}y_n & = & x_1 \\ b_{2,1}y_1 & + & b_{2,2}y_2 & + ... + & b_{2,n}y_n & = & x_2 \\ \vdots & & & & & & \vdots \\ b_{n,1}y_1 & + & b_{n,2}y_2 & + ... + & b_{n,n}y_n & = & x_n \end{array}\right.$$

ou encore, matriciellement, $BY = X$. Alors la matrice $B \in \mathcal{M}_n(\mathbb{K})$ obtenue est l'inverse de A puisque

$$AX = Y \Leftrightarrow A^{-1}AX = A^{-1}Y \Leftrightarrow X = A^{-1}Y$$

Exemple 2 *Soit $\mathcal{U} = (u_1, u_2, u_3)$ une base d'un $\mathbb{K}$-ev E et soit $\mathcal{V} = (v_1, v_2, v_3)$ la famille de vecteurs définie par les relations*

$$\left\{\begin{array}{l} v_1 = u_1 + u_2 \\ v_2 = u_1 - u_2 \\ v_3 = u_1 + u_2 + u_3 \end{array}\right.$$

Pour montrer que $\mathcal{V}$ est une base de E et obtenir la matrice de passage inverse (de $\mathcal{V}$ à $\mathcal{U}$) on peut résoudre le système $Y = AX$ après avoir constaté qu'il est de Cramer*, où A est la matrice* $\left(\begin{smallmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 0 & 0 & 1 \end{smallmatrix}\right)$.

3 Étude du cas général

On revient à la situation générale : (S) est un système de n équations, à p inconnues, de rang r.

$r = \mathrm{rg}(A) = \mathrm{rg}(C_1,\ ...,\ C_p)$ donc il existe des indices j_1, ..., $j_r \in [\![1, p]\!]$ deux à deux distincts tels que $(C_{j_1},\ ...,\ C_{j_r})$ soit une base de $\langle C_1,\ ...,\ C_p\rangle$, sev de $\mathbb{K}^n$ engendré par les vecteurs-colonne de A.

Notons également L_1, ..., L_n $(\in \mathbb{K}^p)$ les vecteurs-ligne de la matrice A. On a aussi $r = \mathrm{rg}(A) = \mathrm{rg}(L_1,\ ...,\ L_n)$.

Soit C la matrice (n, r) extraite de A en ne retenant que les colonnes C_{j_1}, ..., C_{j_r}. Les r colonnes de C sont libres (et bien sûr $r \leq n$) de sorte que $\mathrm{rg}(C) = r$.

C'est aussi le rang les lignes L'_1, ..., L'_n de C donc il existe des indices i_1, ..., $i_r \in [\![1, n]\!]$ deux à deux distincts tels que $\left(L'_{i_1},\ ...,\ L'_{i_r}\right)$ soit une base de $\langle L'_1,\ ...,\ L'_n\rangle \subset \mathbb{K}^r$.

Soit D la sous-matrice de C obtenue en ne gardant que les lignes i_1, ..., i_r. D est donc la sous-matrice extraite de A en retenant les lignes i_1, ..., i_r et les colonnes j_1, ..., j_r :

$$\begin{array}{ccc} & & j_1 \quad\; j_r \\ & & \vdots \quad\;\; \vdots \\ A = & \left(\begin{array}{c} \\ \\ i_1 \cdots \\ i_r \cdots \end{array} \right. & \left(\begin{array}{c} C \\ \\ (D) \end{array} \right) \left. \vphantom{\begin{array}{c} C \\ \\ (D) \\ x \end{array}} \right) \end{array}$$

D est de rang r et appartient à $\mathcal{M}_r(\mathbb{K})$ donc : D est une matrice carrée de taille r *inversible*.

Remarque 1 *Il est clair que toute sous-matrice de A est de rang $\leq r$. On a ainsi mis en évidence que* le rang r d'une matrice A est la taille de la plus grande sous-matrice carrée inversible extraite de A.

Les équations n° i_1, ..., i_r sont appelées des *équations principales*, les autres étant *équations auxiliaires*.

Les inconnues x_{j_1}, ..., x_{j_r} s'appellent des *inconnues principales* de (S) (les autres étant des *inconnues auxiliaires*).

Remarque 2 *La détermination des équations / inconnues principales est* globale *et* non unique *en général. Cela n'a donc pas de sens de dire "la $i^{ème}$ équation (*resp. *inconnue) est principale". On peut seulement dire qu'elle fait partie d'un ensemble d'équations (*resp. *d'inconnues) principales. Mais une autre détermination pourrait tout aussi bien conduire à la considérer comme auxiliaire.*

3.1 Raisonnement sur les colonnes

$(C_{j_1}, \ ..., \ C_{j_r})$ est une base de $\langle C_1, \ ..., \ C_p \rangle$ donc

$$\begin{aligned} (S) \text{ est compatible} \quad &\Longleftrightarrow \quad B \in \langle C_1, \ ..., \ C_p \rangle \\ &\Longleftrightarrow \quad B \in \langle C_{j_1}, \ ..., \ C_{j_r} \rangle \\ &\Longleftrightarrow \quad (C_{j_1}, \ ..., \ C_{j_r}, \ B) \text{ liée,} \end{aligned}$$

puisque la famille $(C_{j_1}, \ ..., \ C_{j_r})$ est libre par définition.

3.2 Raisonnement sur les lignes

Soit $k \in [\![1, n]\!] - \{i_1, \ ..., \ i_r\}$. L_k est c.l. de $(L_{i_1}, \ ..., \ L_{i_r})$:

$$L_k = \lambda_k^1 L_{i_1} + ... + \lambda_k^r L_{i_r} \tag{1}$$

ce qui exprime une relation entre les *premiers membres* de l'équation k et des équations principales.

Cas 1 *Pour tout $k \in [\![1, n]\!] - \{i_1, \ ..., \ i_r\}$,*

$$b_k = \lambda_k^1 b_{i_1} + ... + \lambda_k^r b_{i_r}.$$

Alors les seconds membres b_k vérifient *tous* les mêmes relations (1) que les premiers membres. Les $(p-r)$ équations auxiliaires sont conséquences des r équations principales (elles sont redondantes). Pour chaque valeur des inconnues auxiliaires, on obtiendra les valeurs correspondantes des inconnues principales en résolvant le système (S') obtenu...

- en ne retenant que les équations i_1, ..., i_r ;
- en faisant passer les inconnues autres que x_{j_1}, ..., x_{j_r} dans les seconds membres :

$$(S') \left\{ \begin{array}{lcl} a_{i_1, j_1} x_{j_1} \quad + ... + \quad a_{i_1, j_r} x_{j_r} & = & b'_{i_1} \\ \vdots & & \vdots \\ a_{i_r, j_1} x_{j_1} \quad + ... + \quad a_{i_r, j_r} x_{j_t} & = & b'_{i_r} \end{array} \right.$$

(où $b'_{i_k} = b_{i_k} - \sum_{j \notin \{j_i, \ ..., \ j_r\}} a_{i_k, j} x_j$).

(S') a pour matrice D donc c'est un système de Cramer, dont l'unique solution exprime les inconnues principales en fonction des inconnues auxiliaires qui jouent le rôle de paramètres.

Cas 2 *Il existe $k \in [\![1, n]\!] - \{i_1, \ ..., \ i_r\}$ tel que*

$$b_k \neq \lambda_k^1 b_{i_1} + ... + \lambda_k^r b_{i_r}.$$

Alors l'équation k est incompatible (contradictoire) avec les équations principales et : (S) n'admet pas de solution.

3.3 Méthode de résolution

Convertissons les remarques précédentes en une méthode d'étude d'un système (S).

1. On détermine le rang r de (S).
2. On détermine
   - r équations principales $(i_1, \ ..., \ i_r)$,
   - r inconnues principales $(x_{j_1}, \ ..., \ x_{j_r})$ de (S).

   Cela se fait simultanément, en mettant en évidence dans la matrice A une sous-matrice carrée $r \times r$ inversible.
3. On écarte provisoirement les équations auxiliaires et on fait passer les inconnues auxiliaires (s'il en existe) dans les seconds membres (système (S')). On résout le système de Cramer obtenu.
4. On teste la compatibilité des solutions obtenues pour (S') avec les $(n-r)$ équations auxiliaires (s'il en existe) pour vérifier que les solutions calculées en 3. satisfont effectivement toutes les équations de (S).

Remarque 3 *Dans le cas d'un système de* Cramer, *il n'y a bien sûr que l'étape 3. D'autre part, il se peut que dans les cas simples on puisse voir dès le début la (les) condition(s) de compatibilité. Notamment pour un système homogène, toujours compatible, l'étape 4. est superflue.*

Exemple 3 *Cas général*

1. $\left\{ \begin{array}{l} x + \ \ y + \ \ z + \ \ t = 2 \\ 2x + \ \ y + \ \ z + \ \ t = 1 \\ x + 2y + 2z \qquad = 2 \end{array} \right.$

2. $\left\{ \begin{array}{l} x \qquad\; + \ \ z = 1 \\ \qquad y + \ \ z = 0 \\ x + \ \ y \qquad = 1 \\ 2x + 3y \qquad = 0 \end{array} \right.$

3. $\left\{ \begin{array}{l} \qquad ay + bz = \ u \\ bx \qquad + az = \ v \\ ax + by \qquad = w \end{array} \right.$