Systèmes d'équations linéaires

Lycée Berthollet, MPSI1 2023-24

Exemple introductif (fil rouge)

Exemple 1 On considère le système suivant :

$$(S)\quad\left\{\begin{array}{rrrrcr} x & +2y & -2z & +3t & = & 2\\ 2x & +4y & -3z & +4t & = & 5\\ 5x & +10y & -8z & +11t & = & 12\end{array}\right.$$

qu'on résout par la méthode du pivot en expliquant chaque étape, et notamment en décrivant les opérations élémentaires effectuées sur les lignes. On fait ensuite remarquer que seuls les coefficients comptent, pour peu que le système soit bien présenté ("en colonnes"). On peut donc abstraire le calcul précédent en ne gardant que des tableaux de nombres, appelés matrices. On fait enfin la même résolution directement sur les matrices "augmentées".

Deux autres exemples à traiter par les élèves :

Exemples 2

1.
$$\left\{\begin{array}{rrrcr} x & +2y & -z & = & 1\\ 2x & +y & -z & = & 5\\ x & & -z & = & 5\end{array}\right.$$

2.
$$\left\{\begin{array}{rrrrcr} 2x & +y & -2z & +3t & = & 1\\ 3x & +2y & -z & +2t & = & 4\\ 3x & +3y & +3z & -3t & = & 5\end{array}\right.$$

I Généralités

1 Systèmes linéaires

Comme précédemment $\mathbb{K}$ désignera soit le corps $\mathbb{R}$, soit le corps $\mathbb{C}$.

Définition 3 Un *équation linéaire à p inconnues x_j*, $j \in [\![1,p]\!]$, est une équation du type

$$\sum_{j=1}^{p} a_j x_j = b,$$

où les *coefficients a_j*, ainsi que le *second membre b*, sont des éléments de $\mathbb{K}$. Les *solutions* de cette équation sont les $(x_1,\dots,x_p) \in \mathbb{K}^p$ tels que l'équation soit satisfaite.

Exemples 4 Équations de droites affines en dimension 2 (les équations de la forme $y = ax + b$ ne permettent pas de représenter toutes les droites) et de plans affines en dimension 3. Vecteur normal obtenu avec une équation et interprétation en termes de produit scalaire “usuel”.
Remarque : plus généralement, on appelle hyperplan affine une partie de $\mathbb{R}^p$ définie par une équation linéaire à coefficients non tous nuls.

Exercice 1 Dans le plan $\mathcal{P}$ euclidien usuel muni d'un RON, trouver une équation de la droite passant par les points $A(2,1)$ et $B(-1,-2)$ en déterminant d'abord un vecteur normal à cette droite.

Définition 5 Un *système* (S) *de* n *équations linéaires à* p *inconnues* x_j, $j \in [\![1,p]\!]$ est la conjonction de n équations linéaires :

$$\forall i \in [\![1,n]\!], \quad \sum_{j=1}^{p} a_{i,j} x_j = b_i,$$

où les *coefficients* $a_{i,j}$, $(i,j) \in [\![1,n]\!] \times [\![1,p]\!]$, ainsi que les *seconds membres* b_i, $i \in [\![1,n]\!]$, sont des éléments de $\mathbb{K}$. Les *solutions* de ce système sont les $(x_1, \ldots, x_p) \in \mathbb{K}^p$ tels que les n équations soient satisfaites.

Exemples 6 Intersections de droites affines en dimension 2, de plans affines en dimension 3. (Vecteur directeur de l'intersection de deux plans affines de $\mathbb{R}^3$ à l'aide d'un “produit vectoriel”.)

Définition 7 Avec les notations précédentes, on appelle *système linéaire homogène associé à* (S) le système (H) :

$$\forall i \in [\![1,n]\!], \quad \sum_{j=1}^{p} a_{i,j} x_j = 0.$$

2 Représentation par des matrices

On représente le système linéaire précédent par le tableau à n lignes et p colonnes des coefficients $a_{i,j}$ ($(i,j) \in [\![1,n]\!] \times [\![1,p]\!]$), qu'on appelle *matrice du système* (H) :

$$A = (a_{i,j})_{i,j} = (a_{i,j})_{\substack{1 \le i \le n \\ 1 \le j \le p}} = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,p} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,p} \end{pmatrix}$$

et par le tableau à n lignes et une colonne des b_i ($i \in [\![1,n]\!]$), qu'on appelle *vecteur colonne (ou matrice colonne) des seconds membres* :

$$B = (b_i)_{i=1}^{n} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}.$$

On représente souvent la donnée totale du système dans un tableau à n lignes et $p+1$ colonnes formé en juxtaposant A et B, séparés par un trait vertical, appelé *matrice augmentée du système* (S) :

$$(A|B) = \left(\begin{array}{cccc|c} a_{1,1} & a_{1,2} & \cdots & a_{1,p} & b_1 \\ a_{2,1} & a_{2,2} & \cdots & a_{2,p} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,p} & b_n \end{array}\right).$$

Remarque 8 Il arrive qu'on remplace les parenthèses par des crochets dans la notation matricielle, ce qui est plus pratique pous les matrices ayant beaucoup de lignes. La matrice augmentée se note aussi $[A|B]$.

Exemple 9 Pour l'exemple introductif, on a comme matrice augmentée :

$$\left(\begin{array}{rrrr|r} 1 & 2 & -2 & 3 & 2 \\ 2 & 4 & -3 & 4 & 5 \\ 5 & 10 & -8 & 11 & 12 \end{array}\right) = \left[\begin{array}{rrrr|r} 1 & 2 & -2 & 3 & 2 \\ 2 & 4 & -3 & 4 & 5 \\ 5 & 10 & -8 & 11 & 12 \end{array}\right].$$

3 Opérations élémentaires sur les lignes

On traite simultanément le cas des systèmes et des matrices. On a **trois sortes d'opérations élémentaires sur les lignes** :

Échange de deux lignes $L_i \leftrightarrow L_k$

Pour tout $j \in [\![1,p]\!]$, on échange les coefficients $a_{i,j}$ et $a_{k,j}$.

Ajout à une ligne d'un multiple d'une autre $L_i \leftarrow L_i + \lambda L_k$ $(i \neq k, \lambda \in \mathbb{K})$

Pour tout $j \in [\![1,p]\!]$, on remplace $a_{i,j}$ par $a_{i,j} + \lambda a_{k,j}$.

Produit d'une ligne par un scalaire non nul $L_i \leftarrow \lambda L_i$ $(\lambda \in \mathbb{K}^\star)$

Pour tout $j \in [\![1,p]\!]$, on remplace $a_{i,j}$ par $\lambda a_{i,j}$.

Remarque 10 Pour l'échange de lignes, il peut être pratique d'accepter le cas $i = k$, qui ne change rien, lorsqu'on ne sait pas s'il y a égalité ou non. Par exemple, dans l'algorithme du pivot pour les systèmes réels, d'un point de vue numérique, il est préférable de choisir le pivot de valeur absolue maximale, parmi ceux disponibles. On échange donc la ligne courante avec la ligne dont le coefficient approprié est de valeur absolue maximale, qui peut être la ligne courante elle-même. On peut ainsi décrire l'algorithme de manière épurée, même si en pratique on ne fait aucune opération pour cet échange trivial. De la même manière, il peut être utile d'accepter le cas $\lambda = 0$ pour l'ajout à une ligne d'un multiple d'une autre.

En revanche, il est important dans le cas d'ajout de supposer que $i \neq k$, sous peine de transformer un système linéaire en un système qui ne lui serait pas logiquement équivalent. De manière analogue, il est crucial de supposer que $\lambda \neq 0$ lorsqu'on multiplie une ligne par λ.

Ces opérations permettent de définir la notion suivante :

Équivalence par lignes

Définition 11 Deux systèmes linéaires (resp. matrices, resp matrices augmentées) sont *équivalents par lignes* ssi l'on peut passer de l'un à l'autre par une suite finie d'opérations élémentaires sur les lignes.

Dans ce cas, on notera $(S) \underset{L}{\sim} (S')$ (resp. $A \underset{L}{\sim} A'$, resp. $(A|B) \underset{L}{\sim} (A'|B')$).

Proposition 12 *L'équivalence par lignes est une relation d'équivalence (i.e. réflexive, symétrique et transitive) sur l'ensemble des systèmes linéaires sur $\mathbb{K}$ (resp. l'ensemble des matrices à n lignes et p colonnes à coefficients dans $\mathbb{K}$, resp. sur l'ensemble des matrices augmentées à n lignes et $p+1$ colonnes à coefficients dans $\mathbb{K}$).*

Démonstration: À faire en exercice, le point clé étant la symétrie. On montre cela en remarquant que toute opération élémentaire sur les lignes est réversible : on trouve une deuxième opération élémentaire sur les lignes (qui peut éventuellement être la première opération elle-même) qui "défait" ce que la première a fait. $\square$

Le point crucial pour la résolution des systèmes linéaires est le suivant :

Proposition 13 *Si deux systèmes sont équivalents par lignes, alors ils sont équivalents d'un point de vue logique, c'est-à-dire qu'ils ont le même ensemble de solutions.*

Démonstration: Soient (S), (S') deux systèmes linéaires sur $\mathbb{K}$ tels que $(S) \underset{L}{\sim} (S')$. Comme l'équivalence par lignes est une relation d'équivalence, donc symétrique, il suffit de montrer que $(S) \Longrightarrow (S')$ et pour cela qu'on a une implication pour chaque opération élémentaire sur les lignes. C'est évident pour l'échange. Pour la multiplication par un scalaire cela provient du fait qu'en multipliant chaque membre d'une égalité par le même scalaire, on obtient encore une égalité. Pour l'ajout d'un multiple d'une autre ligne à une ligne, cela provient de la remarque précédente et du fait qu'on peut additionner membre à membre les égalités. $\square$

La justification de la présentation matricielle est le résultat suivant :

Proposition 14 *Deux systèmes linéaires sont équivalents par lignes si et seulement si leur matrices augmentées le sont.*

Démonstration: C'est trivial. $\square$

Exemple 15 Notre exemple fil rouge.

II Échelonnement et algorithme du pivot

Définition 16 Une matrice est dite *échelonnée par lignes* si elle vérifie les deux propriétés suivantes :

1. Si une ligne est nulle, toutes les lignes suivantes le sont aussi ;
2. À partir de la deuxième ligne, dans chaque ligne non nulle, le premier coefficient non nul à partir de la gauche est situé plus à droite que le premier coefficient non nul de la ligne précédente.

Remarque 17 L'échelonnement se traduit par le fait que la matrice vérifie un schéma "en escalier" :

DESSIN

Définition 18 Pour une matrice échelonnée, on appelle *pivot* le premier coefficient non nul de chaque ligne non nulle.
Une matrice échelonnée par lignes est dite *échelonnée réduite par lignes* (si elle est nulle ou) si tous ses pivots sont égaux à 1 et sont les seuls éléments non nuls de leur colonne.

Le résultat fondamental de ce chapitre est le suivant :

Théorème 19 *Toute matrice est équivalente par lignes à une unique matrice échelonnée réduite par lignes.*

Démonstration: On admet l'unicité, qui peut se montrer à l'aide des outils d'algèbre linéaire qu'on verra au second semestre. L'existence repose sur l'algorithme du pivot de Gauss-Jordan, qu'on décrit ci-dessous. □

Algorithme du pivot

Soit A une matrice à n lignes et p colonnes. Cet algorithme mène par une suite d'opérations élémentaires sur les lignes à l'unique matrice échelonnée réduite par lignes qui est équivalente par lignes à la matrice A. Il se compose de trois étapes :

1. la **descente**, qui mène à une matrice échelonnée par lignes ;
2. la **remontée**, qui mène à une matrice échelonnée par lignes telle que les pivots sont les seuls éléments non nuls de leur colonne ;
3. la **normalisation**, qui mène à une matrice échelonnée réduite par lignes.

Il est à noter que les étapes de remontée et de normalisation peuvent éventuellement être interverties.
N.B. Pour la commodité de la description des algorithmes, on prend la convention d'appeler toujours $a_{i,j}$ les coefficients des différentes matrices obtenues après chaque opération élémentaire. Cette convention est cohérente avec la vision informatique des choses, où la *variable* représentant la matrice garde le même nom tout en étant sans cesse modifée.
Descente

Elle consiste à échelonner par lignes la matrice, mais cela en effectuant un traitement colonne par colonne de gauche à droite.

Pour bien faire comprendre l'algorithme, plutôt que de faire directement le cas général, on décrit d'abord le cas de la première colonne (mais cela n'est pas logiquement nécessaire). Si la première colonne est nulle, elle vérifie déjà les critères de l'échelonnement. Si elle n'est pas entièrement nulle, **quitte à échanger deux lignes**, on s'arrange pour que le coefficient $a_{1,1}$ (le *pivot*) soit non nul. On fait ensuite les **opérations** $L_i \leftarrow L_i - \frac{a_{i,1}}{a_{1,1}} L_1$ pour $i \in [\![2,n]\!]$ ce qui fait apparaître des 0 sous le pivot. Remarquons qu'on peut faire ces opérations élémentaires dans l'ordre qu'on veut car aucune ne modifie la ligne L_1.

Le cas général est analogue : supposons qu'on ait déja traité les premières colonnes de la matrice jusqu'à la colonne $j_0 - 1 \leq p - 1$ comprise. Le dernier pivot exhibé (*i.e.* le plus à droite) est en colonne $j < j_0$ et en ligne $i_0 - 1 \leq n$, ou il n'existe pas et dans ce cas, on pose $i_0 = 1$. On traite alors la colonne j_0 ainsi :

— Si les coefficients a_{i,j_0} pour $i \geq i_0$ sont nuls, alors la colonne j_0 convient. C'est en particulier le cas lorsqu'il n'y a pas de tels indices i, c'est-à-dire si $i_0 = n+1$.

— Sinon, quitte à faire un **échange de la ligne d'indice i_0 avec une ligne au-dessous** (ce qui ne change pas les colonnes d'indice 1 à $j_0 - 1$, puisque elles ont des coefficients nuls dans les lignes d'indice $i \geq i_0$), on peut s'arranger pour que $a_{i_0,j_0} \neq 0$: c'est alors notre pivot, dont on se sert pour faire apparaître des 0 en dessous de lui, par les opérations élémentaires

$$L_i \leftarrow L_i - \frac{a_{i,j_0}}{a_{i_0,j_0}} L_{i_0}, \text{ pour } i \in [\![i_0+1, n]\!]$$

puis la colonne j_0 convient. Remarquons qu'on peut faire ces opérations élémentaires dans l'ordre qu'on veut car aucune ne modifie la ligne L_{i_0}.

On continue ainsi jusqu'à la dernière colonne et on obtient une matrice échelonnée par lignes.

Remontée

On parcourt cette fois les colonnes **par ordre décroissant des indices**, et dans chaque colonne **contenant un pivot**, on effectue des opérations élémentaires pour faire apparaître des 0 au dessus du pivot : si la colonne a comme indice j_0 et le pivot est en ligne i_0, on effectue :

$$L_i \leftarrow L_i - \frac{a_{i,j_0}}{a_{i_0,j_0}} L_{i_0}, \text{ pour } i \in [\![1, i_0-1]\!].$$

Remarquer qu'on peut faire ces opérations en "remontant" cette colonne ou en la descendant car ces opérations sont indépendantes les unes des autres.

À la fin de cette étape, on a une matrice échelonnée, dont chaque pivot est le seul élément non nul de sa colonne.

Normalisation

Il suffit alors de multiplier chaque ligne contenant un pivot par l'inverse de ce pivot pour obtenir une matrice échelonnée réduite par ligne.

Exemple 20 L'exemple fil rouge et d'autres faisant en particulier faire des échanges de lignes.

Exemple 21 Utilisation de Python pour l'algorithme du pivot, formel ou numérique.

Dans le cas numérique, exemples d'instabilité.

III Ensemble des solutions d'un système linéaire

On considère ici un système linéaire à n équations et p inconnues, dont la matrice augmentée est $(A|B)$. On applique l'algorithme du pivot à la matrice A, mais en faisant les opérations élémentaires nécessaires directement sur la matrice augmentée $(A|B)$.

À la fin de la descente on peut définir les notions suivantes :

Définition 22 Les *inconnues principales* sont celles correspondant aux colonnes contenant un pivot, les *inconnues secondaires* (ou *paramètres*) sont les inconnues non principales.

Exemple 23 Dans notre fil rouge, les inconnues principales sont x et z et les inconnues secondaires sont y et t.

On a alors trivialement :

Proposition 24 *Le nombre de paramètres est égal au nombre d'inconnues moins le nombre de pivots.*

Exemple 25 Dans notre fil rouge, le nombre de pivots est 2 et le nombre de paramètres est 2 aussi. Dans le premier des exemples 2, le nombre de pivots est 3 et le nombre de paramètres est 0.

Définition 26 On dit que le système est *incompatible* lorsqu'à la fin de la descente, il existe des lignes avec un premier membre nul et un second membre non nul. Si le système n'est pas incompatible, il est dit *compatible*.

Exemple 27 Notre fil rouge est un système compatible. Les exemples 2 sont, le premier, compatible et le second, incompatible.

L'intérêt de la notion de compatibilité est le suivant :

Proposition 28 *Le système admet des solutions si et seulement s'il est compatible.*

Démonstration: S'il y a incompatibilité, il est clair qu'il n'y a pas de solutions, puisque l'une des équations a la forme $0 = b$, avec $b \neq 0$. S'il y a compatibilité, c'est la suite de l'algorithme qui prouve qu'il y a des solutions, comme on va le voir bientôt. □

En poursuivant l'algorithme du pivot (remontée et normalisation), puis en revenant à l'écriture du système et passant les paramètres au second membre, on obtient l'expression des solutions du système linéaire : on donne différentes écritures de cette forme paramétrique.

Plus précisément, si on note $a'_{i,j}$ les coefficients du système final obtenu par l'algorithme du pivot, b'_i ses seconds membres, r le nombre de pivots et $J = \{j_i;\ i \in [\![1,r]\!]\}$ l'ensemble des indices des colonnes des pivots (en particulier, pour $i \in [\![1,r]\!]$, $a'_{i,j_i} = 1$), alors (S) est équivalent à

$$\forall i \in [\![1,r]\!], \quad x_{j_i} = b_i - \sum_{j \in [\![1,p]\!] \setminus J} a'_{i,j} x_j.$$

En intercalant dans ce système les lignes $x_j = x_j$, pour $j \in [\![1,p]\!] \setminus J$, on obtient une représentation paramétrique de l'ensemble des solutions.

Interprétation : description des solutions comme somme d'une solution particulière du système avec second membre et de la “solution générale” du système homogène associé.

Exemple 29 Application aux problèmes d'intersection en géométrie du plan et de l'espace. Résoudre les systèmes exhibés précédemment.

Exemple 30 Exemples provenant de l'électricité.