Chapitre 1

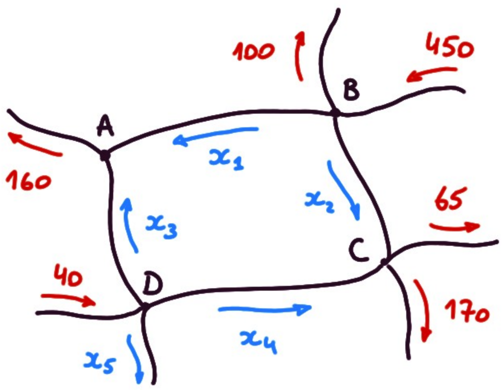
Systèmes linéaires

1.1 Motivation

Ce premier chapitre présente les systèmes d'équations qui seront étudiés dans ce cours. Voyons comment ce type de système peut apparaître, dans des situations très pratiques.

Traffic routier

Dans une petite ville ne possédant que 4 croisements, on a mesuré les flux de voitures sur quelques axes routiers entrants et sortants de la ville, dans le but de prévoir les flux résultant sur le réseau interne, et de préparer les aménagements nécessaires :



Ces mesures indiquent, par exemple, que le flux de voitures entrant au croisement B, venant de l'est de la ville, est de 450 voitures par heure.

Étant données ces contraintes, se pose la question de savoir s'il est possible de calculer les flux résultants sur les autres axes, indiqués par les lettres x_1 à x_5 sur la figure.

Le principe régissant les flux à un croisement est le même que celui utilisé dans les réseaux électriques (Loi de Kirchoff) : en chaque noeud du réseau, la somme des flux entrants doit être égal à la somme des flux sortants, ce qui donne, en chacun des points du réseau,

$$
\begin{aligned}
A : &\quad +x_1 + x_3 = 160 \\
B : &\quad 450 = 100 + x_1 + x_2 \\
C : &\quad x_2 + x_4 = 65 + 170 \\
D : &\quad 40 = x_3 + x_4 + x_5
\end{aligned}
$$

On peut récrire ces relations comme suit :

$$\begin{cases} x_1 & & & + & x_3 & & & & & = & 160 \\ x_1 & + & x_2 & & & & & & & = & 350 \\ & & x_2 & & & + & x_4 & & & = & 235 \\ & & & & x_3 & + & x_4 & + & x_5 & = & 40 \end{cases}$$

Plusieurs questions se posent :

- ⋆ Existe-t-il des nombres x_1, x_2, x_3, x_4, x_5 satisfaisant simultanément à ces 4 conditions ?
- ⋆ Si oui, ces nombres sont-ils tous positifs, pour respecter les sens imposés sur la figure, ou alors certains de ces nombres sont-ils négatifs (auquel cas on devra inverser le sens du traffic sur les axes concernés) ?
- ⋆ Si oui toujours, est-ce que ces nombres sont uniques ? Existe-t-il plusieurs solutions ? Et si ils ne sont pas uniques, quelles contraintes y a-t-il sur les choix que l'on peut faire ?
- ⋆ Et si la ville contenait des milliers de croisements, avec des centaines de flux entrants/sortants ?

Le système de 4 équations à 5 inconnues ci-dessus est un exemple de ce qu'on appelle un *système linéaire*. Ce type de système forme une part importante de ce cours, et on commencera leur étude dans la section suivante.

Informel 1.1. La dernière question ("et si la ville était beaucoup plus grande ?") montre qu'il est important d'aborder l'étude de ces systèmes de façon rigoureuse, en acceptant qu'ils peuvent être de taille arbitrairement grande.

Discrétisation d'équations différentielles

(faire)

1.2 Définition et exemples

Définition 1.2. Un **système linéaire en les variables** $x_1, x_2, \dots, x_n$ est une famille d'équations du type suivant :

$$(*) \begin{cases} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & b_2 \\ \vdots & & & & & & & & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & b_m \end{cases}$$

Un tel système est dit "$m \times n$" : il contient m équations, et n variables. Les nombres a_{kj} ($1 \leqslant k \leqslant m$, $1 \leqslant j \leqslant n$) sont appelés les **coefficients** du système, les b_k ($1 \leqslant k \leqslant m$) forment le **second membre**.

Remarque 1.3. Insistons sur le fait que les coefficients a_{ij}, ainsi que le second membre, sont des *nombres fixés* qui ne dépendent pas des x_i ; en général ils sont donnés par une situation pratique. ⋄

On peut voir un système $(*)$ comme une famille de m **contraintes** que les variables $x_1, \ldots, x_n$ doivent satisfaire, où la $k$ème contrainte est

$$a_{k1}x_1 \quad + \quad a_{k2}x_2 \quad + \quad \cdots \quad + \quad a_{kn}x_n \quad = \quad b_k$$

(On appelle cette contrainte une **équation linéaire**.)

Résoudre un système linéaire

Considérons un système $m \times n$ donné, comme $(*)$.

Définition 1.4. Une famille de nombres $(\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n)$ est **solution** de $(*)$ si elle satisfait simultanément aux m contraintes spécifiées par $(*)$. L'ensemble des solutions de $(*)$ est noté $S_{(*)}$.

- ⋆ S'il existe au moins une solution, $S_{(*)} \neq \varnothing$, on dit que $(*)$ est **compatible**.
- ⋆ S'il n'existe aucune solution, $S_{(*)} = \varnothing$, on dit que $(*)$ est **incompatible** (ou **singulier**).

Lorsqu'un système est compatible, le **résoudre** signifiera trouver *toutes* ses solutions. Dans ce cas, on devra aussi savoir décrire précisément $S_{(*)}$. Voyons deux exemples simples.

Exemple 1.5. Le système 2×2

$$(*) \begin{cases} x_1 & + & x_2 & = & 1 \\ x_1 & + & x_2 & = & 0 \end{cases}$$

est incompatible. En effet, quelle que soit la valeur de x_1 et x_2, la somme $x_1 + x_2$ ne peut pas être à la fois égale à 1 et à 0. Donc $S_{(*)} = \varnothing$. ◇

Exemple 1.6. Considérons le système 1×2 suivant :

$$\left\{ \begin{array}{ccccc} x_1 & + & x_2 & = & 1 \end{array} \right.$$

On trouve facilement des solutions : $(1, 0)$, $(2, -1)$, $(3, -2)$, etc. Donc ce système est compatible, et semble même posséder une infinité de solutions. Pour décrire son ensemble de solutions précisément (pour n'en oublier aucune!), il suffit de remarquer que l'on peut toujours *choisir* une des variables, et prendre l'autre en fonction de façon à ce que la relation soit satisfaite. Par exemple, en choisissant x_1, on garantit que la contrainte est satisfaite en prenant

$$x_2 = 1 - x_1 \,.$$

Lorsqu'on peut ainsi choisir une variable, appelée **variable libre**, on a avantage à y penser comme à un *paramètre*, et à utiliser une autre lettre pour la décrire. Si on utilise la lettre t pour ce paramètre, on a

$$\begin{aligned} x_1 &= t \,, \\ x_2 &= 1 - t \,. \end{aligned}$$

Les variables x_1 et x_2 étant exprimées en fonction des variables libres, on les appelle **variables liées** (ou **variables de base**). On peut finalement exprimer l'ensemble des solutions comme suit :

$$S = \{(t, 1-t) \,:\, t \in \mathbb{R}\} \,.$$

◇

1.3 Sur le nombre de solutions d'un système linéaire

Un de nos objectifs dans ce qui suit sera de trouver des conditions suffisantes pour déterminer si un système est compatible ou incompatible.

Mais avant cela, nous allons énoncer une propriété générale, satisfaite par n'importe quel système linéaire, concernant le *nombre* de solutions qu'il peut posséder.

Théorème 1.7. *Si un système est compatible, alors soit il possède exactement une solution, soit il en possède une infinité.*

Preuve: **Remarque :** Un peu plus loin dans le cours, nous redonnerons la preuve de ce théorème, mais en utilisant le langage vectoriel, ce qui la rendra plus transparente.

Considérons un système $m \times n$ de la forme $(*)$, que l'on suppose être compatible. Si sa solution n'est pas unique, c'est qu'il existe au moins deux familles distinctes, que l'on notera $(\bar{x}_1, \dots, \bar{x}_n)$ et $(\bar{y}_1, \dots, \bar{y}_n)$, qui sont toutes deux solutions de $(*)$; cela signifie qu'elles satisfont toutes les contraintes : pour tout $k = 1, 2, \dots, m$, on a d'une part que

$$a_{k1}\bar{x}_1 \quad + \quad \cdots \quad + \quad a_{kn}\bar{x_n} \quad = \quad b_k$$

et d'autre part que

$$a_{k1}\bar{y}_1 \quad + \quad \cdots \quad + \quad a_{kn}\bar{y_n} \quad = \quad b_k$$

Prenons alors un réel λ quelconque, différent de 0 et de 1, et définissons la famille $(\bar{z}_1, \dots, \bar{z}_n)$, où

$$z_j := \lambda\bar{x}_j + (1-\lambda)\bar{y}_j\,, \quad j = 1, 2, \dots, n\,.$$

Montrons alors que $(\bar{z}_1, \dots, \bar{z}_n)$ est aussi solution, en montrant qu'elle satisfait chacune des m contraintes du système. En effet,

$$\begin{aligned}
&a_{k1}\bar{z}_1 + \cdots + a_{kn}\bar{z_n} \\
&\quad = a_{k1}(\lambda\bar{x}_1 + (1-\lambda)\bar{y}_1) + \cdots + a_{kn}(\lambda\bar{x}_n + (1-\lambda)\bar{y}_n) \\
&\quad = \lambda\big(\underbrace{a_{k1}\bar{x}_1 + \cdots + a_{kn}\bar{x_n}}_{=b_k}\big) + (1-\lambda)\big(\underbrace{a_{k1}\bar{y}_1 + \cdots + a_{kn}\bar{y_n}}_{=b_k}\big) \\
&\quad = \lambda b_k + (1-\lambda)b_k \\
&\quad = b_k\,,
\end{aligned}$$

ce qui signifie que la $k$ème contrainte est satisfaite.
Puisque $(\bar{x}_1, \dots, \bar{x}_n)$ et $(\bar{y}_1, \dots, \bar{y}_n)$ sont distinctes, il existe au moins un k tel que $\bar{x}_k \neq \bar{y}_k$. Cela signifie que si λ n'est égal ni à 0 ni à 1, le nombre $z_k = \lambda\bar{x}_k + (1-\lambda)\bar{y}_k$ est différent à la fois de $\bar{x}_k$ et de $\bar{y}_k$. On peut donc, en choisissant λ, construire autant de nouvelles solutions. Ceci signifie que le système possède une infinité de solutions. □

Informel 1.8. En d'autres termes, le nombre de solutions de n'importe quel système linéaire ne peut être que 0 (s'il est incompatible), 1 ou ∞. Plus tard on se référera à ce résultat comme **le Théorème "**$0, 1, \infty$**"**.

Pour des petites valeurs de n, l'affirmation du Théorème "$0, 1, \infty$" peut s'interpréter géométriquement.

Interprétation géométrique dans le cas $n = 2$

Fixons $n = 2$, et considérons un système $m \times 2$:

$$(*) \begin{cases} a_{11}x_1 & + & a_{12}x_2 & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & = & b_2 \\ & & & \vdots & \\ a_{m1}x_1 & + & a_{m2}x_2 & = & b_m \end{cases}$$

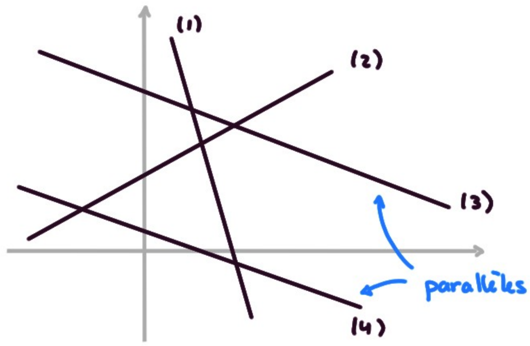
Ici, une paire (x_1, x_2) peut s'interpréter comme les coordonnées d'un point dans le plan, relativement à un repère orthonormé fixé. Aussi, on sait (voir cours de géométrie analytique) qu'une contrainte de la forme

$$a_{k1}x_1 \quad + \quad a_{k2}x_2 \quad = \quad b_k$$

signifie que le point de coordonnées (x_1, x_2) est sur une **droite**.

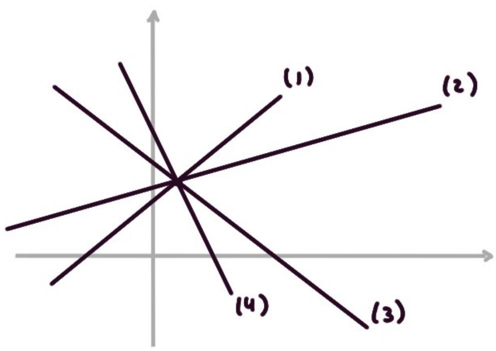
Donc une paire (x_1, x_2) sera solution de $(*)$, $(x_1, x_2) \in S_{(*)}$, si et seulement si le point (x_1, x_2) appartient en même temps à chacune des m droites spécifiées dans $(*)$. Or appartenir à m droites en même temps est une contrainte en général difficile à satisfaire, surtout si m est grand.

- $\star$ $S_{(*)}$ est en général vide, surtout si on parle de plus de deux droites, ou dès que deux de ces droites sont parallèles et distinctes :

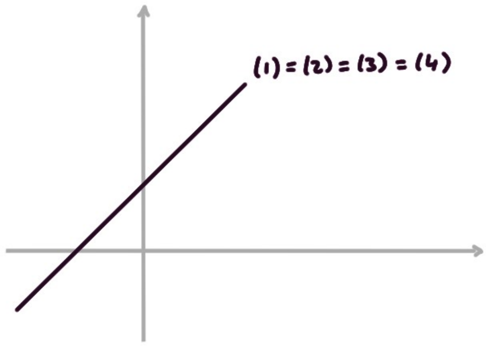


  Sur ce dessin, on voit qu'il n'existe aucun point (x_1, x_2) qui appartient aux quatre droites à la fois.

- $\star$ $S_{(*)}$ contient seulement un élément si les droites s'intersectent en exactement un point :



- $\star$ $S_{(*)}$ contient une infinité d'éléments si les droites sont confondues :



On comprend que géométriquement, il est impossible de créer m droites dans le plan qui s'intersectent, par exemple, en exactement 4 points.

Interprétation géométrique dans le cas $n = 3$

Fixons $n = 3$, et considérons un système $m \times 3$:

$$(*) \left\{ \begin{array}{ccccccc} a_{11}x_1 & + & a_{12}x_2 & + & a_{13}x_3 & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & a_{23}x_3 & = & b_2 \\ & & & & & \vdots & \\ a_{m1}x_1 & + & a_{m2}x_2 & + & a_{m3}x_3 & = & b_m \end{array} \right.$$

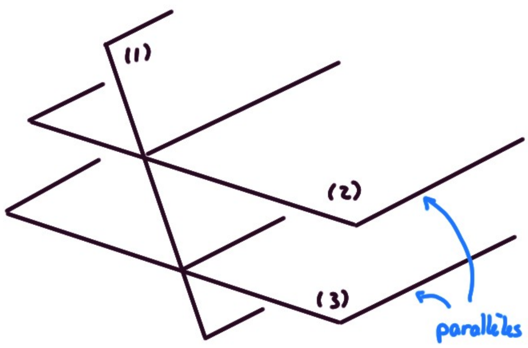
Ici, un triplet (x_1, x_2, x_3) peut s'interpréter comme les coordonnées d'un point dans l'espace, relativement à un repère orthonormé fixé. Aussi, on sait (voir cours de géométrie analytique) qu'une contrainte de la forme

$$a_{k1}x_1 \quad + \quad a_{k2}x_2 \quad + \quad a_{k3}x_3 \quad = \quad b_k$$

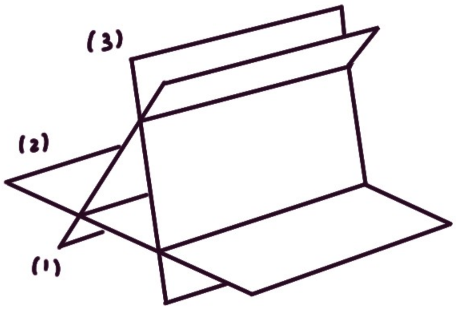
signifie que le point de coordonnées (x_1, x_2, x_3) est sur un **plan**.

Donc un triplet (x_1, x_2, x_3) sera solution de $(*)$ si et seulement si le point (x_1, x_2, x_3) appartient en même temps à chacun des m plans spécifiés. Or ici aussi il est géométriquement clair que $S_{(*)}$ ne peut contenir que 0, 1 ou une infinité de points.

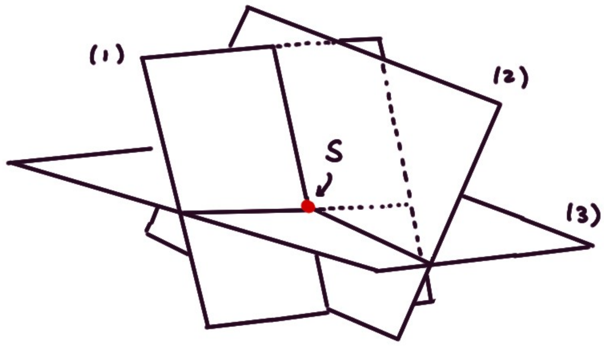
- $\star$ $S_{(*)}$ est vide dès que 2 de ces plans sont parallèles, distincts :



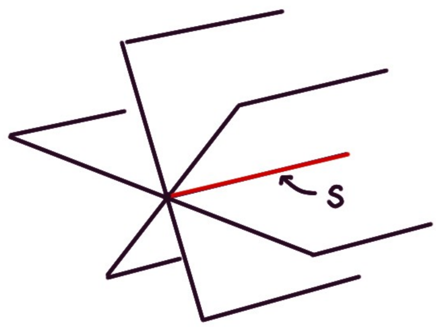
Ou alors ils peuvent aussi n'avoir aucun point en commun mais s'intersecter 2 à 2, sans que certains soient parallèles :



- ⋆ $S_{(*)}$ contient exactement un élément si les plans s'intersectent en seulement un point :



- ⋆ $S_{(*)}$ contient une infinité d'éléments si les plans sont confondus, ou s'intersectent selon une droite :



1.4 Transformer un système en un autre

Notre but pour la suite du chapitre est de présenter une méthode qui permet de savoir si un système est compatible ou incompatilbe et qui, lorsqu'il est compatible, permet en plus de décrire précisément l'ensemble de toutes ses solutions.

Cette méthode est utile non-seulement parce qu'elle mène à un algorithme (l'*algorithme de Gauss*) que l'on peut facilement implémenter sur une machine à l'aide d'un programme de quelques lignes, mais aussi parce qu'elle fournit un résultat théorique qui sera utilisé souvent dans la suite du cours.

Informel 1.9. Attention : Ce que nous présentons ci-dessous est une *méthode de calcul*. Elle sera utilisée souvent dans la suite du cours, mais ne constitue pas, en soi, "l'essence de l'algèbre linéaire" !

Un idéal : les systèmes triangulaires

Pour comprendre l'idée derrière la méthode générale qui va suivre, commençons par considérer un type de système dont la structure suggère elle-même une méthode de résolution.

Exemple 1.10. Considérons le système 3×3 suivant :

$$(*) \left\{ \begin{array}{ccccccc} x_1 & - & x_2 & + & 2x_3 & = & 4 \\ & & x_2 & - & 3x_3 & = & -5 \\ & & & & 5x_3 & = & 10 \end{array} \right. ,$$

que l'on peut comprendre comme étant en fait

$$(*) \left\{ \begin{array}{ccccccc} x_1 & - & x_2 & + & 2x_3 & = & 4 \\ \color{red}{0x_1} & + & x_2 & - & 3x_3 & = & -5 \\ \color{red}{0x_1} & + & \color{red}{0x_2} & + & 5x_3 & = & 10 \end{array} \right.$$

La présence des zéros en bas à gauche donne à ce système une structure triangulaire, qui suggère une résolution simple, "du bas vers le haut" :

a) Dans la troisième équation, on calcule $x_3 = \frac{10}{5} = 2$.

b) On injecte x_3 dans la deuxième équation, pour trouver

$$x_2 = -5 + 3x_3 = -5 + 6 = 1 \,.$$

c) On injecte x_3 et x_2 dans la première équation, pour trouver

$$x_1 = 4 + x_2 - 2x_3 = 4 + 1 - 4 = 1 \,.$$

Donc la solution est unique, donnée par $(x_1, x_2, x_3) = (1, 1, 2)$. ◇

Le système de ce dernier exemple était très particulier, puisque les coefficients $a_{21} = a_{31} = a_{32} = 0$, lui conférant une structure simple à traiter. Mais même si le système était très grand, toujours avec la même structure triangulaire,

$$(*) \left\{ \begin{array}{ccccccccccccc} x_1 & + & 2x_2 & - & 4x_3 & - & 8x_4 & + & x_5 & + & 9x_6 & = & -3 \\ & & x_2 & + & 4x_3 & - & 5x_4 & + & x_5 & - & x_6 & = & 6 \\ & & & & x_3 & + & 6x_4 & + & x_5 & - & 6x_6 & = & -2 \\ & & & & & & -x_4 & + & x_5 & + & x_6 & = & 3 \\ & & & & & & & & 4x_5 & + & x_6 & = & 0 \\ & & & & & & & & & & 7x_6 & = & 11 \end{array} \right.$$

on traiterait le problème de la même façon, "du bas vers le haut"...

Voyons un autre exemple dans lequel on profite de la présence de coefficients nuls dans la partie inférieure gauche, mais où le nombre de variables est supérieur au nombre d'équations :

Exemple 1.11. Considérons

$$\left\{ \begin{array}{ccccccc} 2x_1 & + & x_2 & - & x_3 & = & 0 \\ \color{red}{0x_1} & + & x_2 & + & 2x_3 & = & 3 \end{array} \right.$$

Ce système 2×3 n'est pas "aussi triangulaire" que l'on voudrait. Pourtant, une opération naturelle est de déplacer les termes contenant x_3 du côté droit, pour récrire ce système comme

$$\left\{ \begin{array}{ccccc} 2x_1 & + & x_2 & = & x_3 \\ \color{red}{0x_1} & + & x_2 & = & 3 - 2x_3 \end{array} \right.$$

Écrit sous cette forme, on voit qu'on peut choisir la valeur de x_3, ce qui fixe le côté droit, et qu'on se retrouve ensuite avec un système 2×2 triangulaire dont second membre dépend de x_3. Puisque la valeur de x_3 peut être choisie, on dit que c'est une variable **libre**, elle joue le rôle de paramètre ; on la notera plutôt t :

$$\left\{\begin{array}{rcrcl} 2x_1 & + & x_2 & = & t \\ & & x_2 & = & 3-2t \end{array}\right.$$

Procédant "du bas vers le haut", on obtient $x_2 = 3-2t$, puis

$$x_1 = \tfrac{1}{2}(t-x_2) = \tfrac{1}{2}(t-(3-2t)) = \tfrac{3}{2}(t-1)\,.$$

On a donc, pour tout choix de t, une solution (x_1, x_2, x_3) donnée par

$$\begin{aligned} x_1 &= \tfrac{3}{2}(t-1) \\ x_2 &= 3-2t \\ x_3 &= t\,. \end{aligned}$$

On peut donc écrire l'ensemble de toutes les solutions de la façon suivante :

$$S = \{(\tfrac{3}{2}(t-1), 3-2t, t) \,:\, t \in \mathbb{R}\}\,.$$

◇

Ainsi, la stratégie générale, pour résoudre un système quelconque, sera d'arriver à le *transformer en un système aussi triangulaire que possible*. Pour que ce nouveau système soit utile, il faudra être sûr que son ensemble de solutions soit *exactement le même* que le système de départ.

Opérations élémentaires

Considérons un système $m \times n$, noté $(*)$. Dans la suite, on utilisera le symbole L_i pour représenter la $i$ème ligne de $(*)$:

$$L_i : \qquad a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n = b_i\,.$$

Définissons plusieurs façons d'agir sur les lignes d'un système :

Définition 1.12. a) L'opération consistant à échanger les lignes i et j est appelée **opération élémentaire de Type I**, et sera notée

$$\boxed{L_i \leftrightarrow L_j}$$

b) L'opération consistant à multiplier la $i$ème ligne par un scalaire non-nul $\lambda \in \mathbb{R}^*$ est appelée **opération élémentaire de Type II**, et sera notée

$$\boxed{L_i \leftarrow \lambda L_i}$$

c) L'opération consistant à rajouter à la $i$ème ligne un multiple (par un scalaire λ) de la $j$ème est appelée **opération élémentaire de Type III**, et sera notée

$$\boxed{L_i \leftarrow L_i + \lambda L_j}$$

Une opération élémentaire a pour effet de transformer un système linéaire $(*)$ en un autre système linéaire $(*)'$, de même dimension; ces deux systèmes sont alors dits **équivalents selon les lignes** (ou **ligne-équivalents**).

Exemple 1.13. Considérons le système

$$(*)\left\{\begin{array}{rcrcrcrcr} x_1 & + & 3x_2 & + & x_3 & + & 4x_4 & = & 3 \\ 2x_1 & + & 5x_2 & + & x_3 & - & x_4 & = & -1 \\ -x_1 & - & 2x_2 & - & 3x_3 & - & 9x_4 & = 0 & \end{array}\right.$$

⋆ En appliquant à $(*)$ l'opération de Type I donnée par $L_1 \leftrightarrow L_2$, on obtient

$$(*)'\left\{\begin{array}{rcrcrcrcr} 2x_1 & + & 5x_2 & + & x_3 & - & x_4 & = & -1 \\ x_1 & + & 3x_2 & + & x_3 & + & 4x_4 & = & 3 \\ -x_1 & - & 2x_2 & - & 3x_3 & - & 9x_4 & = & 0 \end{array}\right.$$

⋆ En appliquant à $(*)$ l'opération de Type II donnée par $L_2 \leftarrow \frac{1}{2}L_2$, on obtient

$$(*)'\left\{\begin{array}{rcrcrcrcr} x_1 & + & 3x_2 & + & x_3 & + & 4x_4 & = & 3 \\ x_1 & + & \frac{5}{2}x_2 & + & \frac{1}{2}x_3 & - & \frac{1}{2}x_4 & = & -\frac{1}{2} \\ -x_1 & - & 2x_2 & - & 3x_3 & - & 9x_4 & = & 0 \end{array}\right.$$

⋆ En appliquant à $(*)$ l'opération de Type III donnée par $L_3 \leftarrow L_3 + L_1$, on obtient

$$(*)'\left\{\begin{array}{rcrcrcrcr} x_1 & + & 3x_2 & + & x_3 & + & 4x_4 & = & 3 \\ 2x_1 & + & 5x_2 & + & x_3 & - & x_4 & = & -1 \\ & & x_2 & - & 2x_3 & - & 5x_4 & = & 3 \end{array}\right.$$

◇

Un propriété remarquable des opérations définies ci-dessus est qu'elles sont toutes *inversibles*, dans le sens suivant : si $(*)'$ s'obtient en appliquant une opération élémentaire (de Type I, II ou III) à $(*)$, alors il existe une opération élémentaire **réciproque** qui permet, si on l'applique à $(*)'$, de revenir au système $(*)$ de départ.

a) La réciproque de $L_i \leftrightarrow L_j$ est $L_i \leftrightarrow L_j$ (évident).
b) La réciproque de $L_i \leftarrow \lambda L_i$ ($\lambda \in \mathbb{R}^*$) est $L_i \leftarrow \frac{1}{\lambda}L_i$ (en effet, on peut "défaire" la multiplication de la ligne i par λ en... divisant la ligne i par λ).
c) La réciproque de $L_i \leftarrow L_i + \lambda L_j$ est $L_i \leftarrow L_i - \lambda L_j$ (évident).

Comme conséquence de l'inversibilité :

Théorème 1.14. *Soient deux systèmes de mêmes dimensions, $(*)$ et $(*)'$. Si $(*)'$ est obtenu à partir de $(*)$ par une d'opération élémentaire, alors $(*)$ et $(*)'$ ont le même ensemble de solutions :*

$$S_{(*)} = S_{(*)'}\,.$$

L'usage du théorème ci-dessus se fera comme suit :

Corollaire 1. *Soient deux systèmes de mêmes dimensions, $(*)_1$ et $(*)_2$. Si $(*)_2$ peut être obtenu à partir de $(*)_1$ par une suite finie d'opérations élémentaires,*

$$(*)_1 \rightsquigarrow \cdots \rightsquigarrow (*)_2\,,$$

alors $(*)_1$ et $(*)_2$ ont le même ensemble de solutions :

$$S_{(*)_1} = S_{(*)_2}\,.$$

Preuve: Par le théorème, on sait que lors de chaque étape, l'ensemble de solutions est préservé. Puisqu'il y un nombre fini d'étapes, ceci implique $S_{(*)_1} = S_{(*)_2}$. □

Ce dernier corollaire suggère une méthode de résolution d'un système $m \times n$ quelconque : puisque les opérations élémentaires ne changent pas l'ensemble des solutions, on pourra appliquer au système de départ des opérations qui font peu à peu apparaître des zéros dans la partie inférieure gauche. Une fois le système "triangularisé", on procédera "du bas vers le haut", comme vu précédemment.

Plutôt que d'écrire explicitement un algorithme très général, considérons un premier exemple, qui contient déjà l'idée de la méthode, et montre dans quel ordre les opérations élémentaires sont choisies :

Exemple 1.15. Considérons le système

$$(*)_1 \left\{ \begin{array}{rcrcrcl} 3x_1 & + & 5x_2 & + & 4x_3 & = & 1 \\ 6x_1 & + & 12x_2 & + & 6x_3 & = & 0 \\ -2x_1 & - & 2x_2 & - & 7x_3 & = & 5 \end{array} \right.$$

D'abord, simplifions la deuxième ligne en la divisant par 6, ce qui revient à faire $L_2 \leftarrow \frac{1}{6}L_2$:

$$(*)_1' \left\{ \begin{array}{rcrcrcl} 3x_1 & + & 5x_2 & + & 4x_3 & = & 1 \\ x_1 & + & 2x_2 & + & x_3 & = & 0 \\ -2x_1 & - & 2x_2 & - & 7x_3 & = & 5 \end{array} \right.$$

Regardons maintenant la première colonne, et les coefficients présents devant x_1. Pour ce qui va suivre, on a avantage à faire $L_1 \leftrightarrow L_2$:

$$(*)_1'' \left\{ \begin{array}{rcrcrcl} x_1 & + & 2x_2 & + & x_3 & = & 0 \\ 3x_1 & + & 5x_2 & + & 4x_3 & = & 1 \\ -2x_1 & - & 2x_2 & - & 7x_3 & = & 5 \end{array} \right.$$

On va maintenant profiter du "x_1" en haut à gauche, appelé **pivot**, pour faire apparaître des zéros dans la partie inférieure de la première colonne. Par exemple, pour faire disparaître le "$3x_1$" de la deuxième ligne, on a besoin de lui soustraire 3 fois la première ligne. Donc en faisant $L_2 \leftarrow L_2 - 3L_1$, on obtient

$$(*)_1''' \left\{ \begin{array}{rcrcrcl} {\color{blue}x_1} & + & 2x_2 & + & x_3 & = & 0 \\ {\color{red}0x_1} & - & x_2 & + & x_3 & = & 1 \\ -2x_1 & - & 2x_2 & - & 7x_3 & = & 5 \end{array} \right. .$$

Ensuite, en faisant $L_3 \leftarrow L_3 + 2L_1$,

$$(*)_1'''' \left\{ \begin{array}{rcrcrcl} {\color{blue}x_1} & + & 2x_2 & + & x_3 & = & 0 \\ {\color{red}0x_1} & - & x_2 & + & x_3 & = & 1 \\ {\color{red}0x_1} & + & 2x_2 & - & 5x_3 & = & 5 \end{array} \right. .$$

Ensuite, on passe à la deuxième colonne. C'est maintenant le "$-x_2$", dans L_2, qui joue le rôle de pivot et dicte le choix de l'opération suivante, qui fait apparaître encore un zéro au bas de la deuxième colonne : $L_3 \leftarrow L_3 + 2L_2$, qui donne

$$(*)_2 = (*)_1''''' \left\{ \begin{array}{rcrcrcl} x_1 & + & 2x_2 & + & x_3 & = & 0 \\ {\color{red}0x_1} & {\color{blue}-} & {\color{blue}x_2} & + & x_3 & = & 1 \\ {\color{red}0x_1} & + & {\color{red}0x_2} & - & 3x_3 & = & 7 \end{array} \right.$$

Remarquons que dans cette dernière étape, l'opération élémentaire n'a pas affecté les zéros de la première colonne !

En transformant $(*)_1$ en $(*)_2$, nous dirons plus loin que nous avons *échelonné* le système.

En procédant "du bas vers le haut" dans $(*)_2$, on obtient

$$S_{(*)_2} = \{(\tfrac{27}{3}, -\tfrac{10}{3}, -\tfrac{7}{3})\}\,,$$

et le corollaire permet de conclure que $S_{(*)_1} = S_{(*)_2}$. ◇

1.5 Matrices et algorithme de Gauss

On comprend que la méthode illustrée sur l'exemple précédent, appelée **algorithme de Gauss**, s'applique à des systèmes de tailles quelconques. Il consiste à utiliser des opérations élémentaires qui font progressivement apparaître des zéros dans la partie inférieure gauche de la matrice.

Nous verrons d'autres exemples plus loin, mais arrêtons-nous un instant pour introduire une certaine simplification d'écriture.

Matrices associées à un système

Les opérations élémentaires que l'on effectue sur un système général du type

$$(*)\left\{\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & b_2 \\ \vdots & & & & & & & \vdots & \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & b_m \end{array}\right. ,$$

agissent sur les coefficients a_{ij}, et sur les termes du second membre, les b_i. Dans ces opérations, le nom que l'on donne aux variables (jusqu'à présent : $x_1, \dots, x_n$) n'a pas d'importance. Il est donc utile de simplifier la manipulation des systèmes en ne gardant que la structure numérique des coefficients et du second membre.

Définition 1.16. Soit $(*)$ le système ci-dessus.

a) La **matrice associée à** $(*)$ est le tableau de $m \times n$ nombres défini par

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

b) La **matrice augmentée associée à** $(*)$ est le tableau de $m \times (n+1)$ nombres défini par

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

(La ligne verticale, qui sépare les a_{kn} des b_k, est là pour rappeler que la dernière colonne doit être interprétée comme le second membre dans $(*)$.)

Exemple 1.17. La matrice augmentée du système

$$\left\{\begin{array}{rcrcrcrcl} -x_1 & + & x_2 & & & + & x_4 & = & \alpha \\ x_1 & & & + & x_3 & + & 7x_4 & = & \beta \\ & & x_2 & - & x_3 & + & x_4 & = & \gamma \end{array}\right.$$

est donnée par

$$\left(\begin{array}{rrrr|c} -1 & 1 & 0 & 1 & \alpha \\ 1 & 0 & 1 & 7 & \beta \\ 0 & 1 & -1 & 1 & \gamma \end{array}\right)$$

◇

Opérations élémentaires sur les matrices

Il est clair que les opérations élémentaires (de Type I,II,III) que l'on effectue sur un système se traduisent naturellement en des opérations sur les lignes de la matrice et de la matrice augmentée associée.

De manière générale, on peut effectuer des opérations élémentaires sur une matrice sans se référer au système dont elle est issue. On peut donc définir deux matrices comme étant **équivalentes selon les lignes** (ou **ligne-équivalentes**) si l'une peut s'obtenir à partir de l'autre par un nombre fini d'opérations élémentaires.

Matrices échelonnées

L'algorithme de Gauss mène, en général, à une matrice qui est ce que nous appelions "aussi triangulaire que possible" ; définissons ce terme précisément.

Définition 1.18. Pour une matrice $m \times n$ quelconque,

a) Le **coefficient principal de la i-ème ligne** est, s'il y en a un, le premier coefficient non-nul trouvé, en partant de la gauche.

b) Une matrice est **échelonnée** si les deux conditions suivantes sont satisfaites :

- ⋆ Si elle possède des lignes nulles (c'est-à-dire ne contenant que des coefficients nuls), alors celles-ci sont toutes dans la partie inférieure de la matrice.
- ⋆ Si elle possède des lignes non-nulles, alors le coefficient principal de chacune de ces lignes se trouve strictement à droite du coefficient principal de la ligne du dessus.

Exemple 1.19. La matrice

$$\left(\begin{array}{cccc|c} 3 & 1 & 2 & 0 & 7 \\ 0 & 0 & 4 & 3 & -1 \\ 0 & 0 & 0 & -\frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

est échelonnée. En effet, la seule ligne ne contenant que des zéros est tout en bas, et sur chacune des autres lignes, le coefficient principal est strictement à droite du coefficient principal de la ligne du dessus : $-\frac{1}{2}$ (coefficient principal de la troisième ligne) est à droite de 4 (coefficient principal de la deuxième ligne), qui est lui-même à droite de 3 (coefficient principal de la première ligne). ◇

Exemple 1.20. La matrice

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & -1 & 0 \end{pmatrix}$$

n'est pas échelonnée, car le coefficient principal de la troisième ligne est juste sous le (et non *strictement à droite du*) coefficient principal de la deuxième ligne. ◇

Théorème 1.21. *Toute matrice peut être transformée, à l'aide d'un nombre fini de transformations élémentaires, en une matrice échelonnée. (En d'autres termes : toute matrice est équivalente selon les lignes à une matrice échelonnée.)*

Preuve: La méthode utilisée dans les exemples traités plus haut se généralise sans difficulté à n'importe quelle matrice. □

Exemple 1.22. Considérons la matrice augmentée du système vu plus haut :

$$\left(\begin{array}{ccc|c} 3 & 5 & 4 & 1 \\ 6 & 12 & 6 & 0 \\ -2 & -2 & -7 & 5 \end{array}\right).$$

En appliquant successivement les opérations élémentaires

$$L_2 \leftarrow \tfrac{1}{6}L_2\,, \quad L_1 \leftrightarrow L_2\,, \quad L_2 \leftarrow L_2 - 3L_1\,, L_3 \leftarrow L_3 + 2L_1\,, \quad L_3 \leftarrow L_3 + 2L_2\,,$$

on obtient comme on a vu la matrice

$$\left(\begin{array}{ccc|c} 1 & 2 & 1 & 0 \\ 0 & -1 & 1 & 1 \\ 0 & 0 & -3 & 7 \end{array}\right),$$

qui est échelonnée. ◇

Remarque 1.23. La matrice échelonnée obtenue dépend du choix des opérations élémentaires faites en chemin ! Par exemple, en partant de

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 3 \end{pmatrix},$$

alors en faisant $L_2 \leftarrow L_2 - L_1$ on obtient une première échelonnée

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 2 \end{pmatrix},$$

mais en faisant $L_1 \leftrightarrow L_2$ suivie de $L_2 \leftarrow L_2 - L_1$ on obtient

$$\begin{pmatrix} 1 & 1 & 3 \\ 0 & 0 & -2 \end{pmatrix},$$

qui est une autre forme échelonnée. Donc une matrice peut posséder plusieurs formes échelonnées. ◇

Concrètement, dans la transformation d'un système à l'aide d'opérations élémentaires, c'est une fois que la matrice obtenue est échelonnée que l'on peut déjà savoir si le système est compatible, si oui identifier les variables libres, et exprimer l'ensemble des solutions. Voyons un exemple assez complet :

Exemple 1.24. Supposons que l'on parte du système

$$(*)\left\{\begin{array}{rcrcrcrcr} 3x_1 & + & x_2 & + & 2x_3 & & & = & 7\\ 6x_1 & + & 2x_2 & + & 12x_3 & + & 7x_4 & = & 12\\ 6x_1 & + & 2x_2 & + & 8x_3 & + & 4x_4 & = & 13\\ 3x_1 & + & x_2 & + & 6x_3 & + & 4x_4 & = & 6 \end{array}\right. ,$$

dont la matrice augmentée est

$$\left(\begin{array}{cccc|c} 3 & 1 & 2 & 0 & 7\\ 6 & 2 & 12 & 7 & 12\\ 6 & 2 & 8 & 4 & 13\\ 3 & 1 & 6 & 4 & 6 \end{array}\right)$$

Faisons d'abord apparaître des zéros dans la première colonne, en appliquant successivement $L_2 \leftarrow L_2 - 2L_1$, $L_3 \leftarrow L_3 - 2L_1$, $L_4 \leftarrow L_4 - L_1$:

$$\left(\begin{array}{cccc|c} 3 & 1 & 2 & 0 & 7\\ 0 & 0 & 8 & 7 & -2\\ 0 & 0 & 4 & 4 & -1\\ 0 & 0 & 4 & 4 & -1 \end{array}\right)$$

Comme les deux dernières lignes sont égales, on peut faire $L_4 \leftarrow L_4 - L_3$:

$$\left(\begin{array}{cccc|c} 3 & 1 & 2 & 0 & 7\\ 0 & 0 & 8 & 7 & -2\\ 0 & 0 & 4 & 4 & -1\\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

Ensuite, $L_2 \leftarrow L_2 - 2L_3$, suivie de $L_2 \leftrightarrow L_3$, nous donne une version échelonnée :

$$\left(\begin{array}{cccc|c} 3 & 1 & 2 & 0 & 7\\ 0 & 0 & 4 & 4 & -1\\ 0 & 0 & 0 & -1 & 0\\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

Comment poursuivre, pour obtenir $S_{(*)}$? Pour y voir plus clair, revenons au système correspondant à cette dernière matrice augmentée :

$$(*)'\left\{\begin{array}{rcrcrcrcr} 3x_1 & + & x_2 & + & 2x_3 & + & 0x_4 & = & 7\\ 0x_1 & + & 0x_2 & + & 4x_3 & + & 4x_4 & = & -1\\ 0x_1 & + & 0x_2 & + & 0x_3 & - & x_4 & = & 0\\ 0x_1 & + & 0x_2 & + & 0x_3 & + & 0x_4 & = & 0 \end{array}\right. ,$$

que l'on peut écrire plus simplement comme

$$(*)'\left\{\begin{array}{rcrcrcrcr} 3x_1 & + & x_2 & + & 2x_3 & & & = & 7\\ & & & & 4x_3 & + & 4x_4 & = & -1\\ & & & & & - & x_4 & = & 0 \end{array}\right. .$$

On a supprimé la dernière ligne "$0x_1 + 0x_2 + 0x_3 + 0x_4 = 0$". En effet, cette contrainte n'en est pas une, puisqu'elle est satisfaite par n'importe quel quadruplet (x_1, x_2, x_3, x_4).

On observe maintenant que la variable $x_2 = t$ est **libre**, puisqu'on peut la passer du côté droit pour obtenir un système triangulaire en (x_1, x_3, x_4), qui sont les variables **de base** :

$$(*)'\left\{\begin{array}{rcrcrcr} 3x_1 & + & 2x_3 & & & = & 7-t\\ & & 4x_3 & + & 4x_4 & = & -1\\ & & & - & x_4 & = & 0 \end{array}\right. .$$

Maintenant, en procédant "de bas en haut", on obtient

$$x_4 = 0\,, \qquad x_3 = -\tfrac{1}{4}\,, \qquad x_1 = \frac{15-2t}{6}\,,$$

et donc :

$$S_{(*)} = S_{(*)'} = \left\{\left(\tfrac{15-2t}{6}, t, -\tfrac{1}{4}, 0\right) \,:\, t \in \mathbb{R}\right\}$$

Le système est donc compatible, et possède une infinité de solutions. ◇

Exemple 1.25. Considérons le système

$$(*)\left\{\begin{array}{ccccccc} x_1 & + & x_2 & - & x_3 & = & 5 \\ 3x_1 & + & 4x_2 & & & = & 4 \\ 4x_1 & + & 5x_2 & - & x_3 & = & 7 \end{array}\right.$$

Après échelonnage, sa matrice devient

$$\left(\begin{array}{ccc|c} 1 & 1 & -1 & 5 \\ 0 & 1 & 3 & -11 \\ 0 & 0 & 0 & -3 \end{array}\right).$$

Le système correspondant est

$$(*)'\left\{\begin{array}{ccccccc} x_1 & + & x_2 & - & x_3 & = & 5 \\ 0x_1 & + & x_2 & + & 3x_3 & = & -11 \\ 0x_1 & + & 0x_2 & + & 0x_3 & = & -3 \end{array}\right.,$$

qui est incompatible, puisque la dernière contrainte ne peut jamais être satisfaite, quel que soit (x_1, x_2, x_3). Donc $(*)$ est aussi incompatible. ◇

La réduite de Gauss

Définition 1.26. Une matrice est **échelonnée réduite** si elle est échelonnée et si, de plus,

- ⋆ tous ses coefficients principaux sont égaux à 1, et si
- ⋆ chaque coefficient principal est l'unique élément non-nul de sa colonne.

Les coefficients principaux d'une matrice échelonnée réduite sont appelés **pivots**.

Exemple 1.27. La matrice

$$\begin{pmatrix} 0 & 1 & 0 & 3 & 8 & 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 5 & 9 & 0 & 2 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 3 & 0 & 6 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 5 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

est échelonnée réduite (les pivots sont indiqués en bleu). ◇

Voyons comment obtenir une échelonnée réduite d'une matrice. Une fois encore, la méthode présentée dans cet exemple particulier montre comment procéder en général :

Exemple 1.28. Considérons encore une fois la matrice du premier exemple de cette section :

$$\left(\begin{array}{ccc|c} 3 & 5 & 4 & 1 \\ 6 & 12 & 6 & 0 \\ -2 & -2 & -7 & 5 \end{array}\right).$$

Pour obtenir sa réduite, on commence par l'échelonner, comme vu plus haut :

$$\left(\begin{array}{ccc|c} 1 & 2 & 1 & 0 \\ 0 & -1 & 1 & 1 \\ 0 & 0 & -3 & 7 \end{array}\right).$$

On poursuit en rendant tous les coefficients principaux égaux à 1, à l'aide de $L_2 \leftarrow (-1)L_2$ et $L_3 \leftarrow \frac{1}{-3}L_3$:

$$\left(\begin{array}{ccc|c} 1 & 2 & 1 & 0 \\ 0 & 1 & -1 & -1 \\ 0 & 0 & 1 & -7/3 \end{array}\right),$$

Remarquons que maintenant, lorsqu'on additionne un multiple d'une ligne L_i à une ligne L_j située *au-dessus*, la présence de zéros fait *qu'on ne modifie pas les coefficients de L_j situés à droite du coefficient principal de L_i*.

On procède donc "du bas vers le haut" pour faire apparaître des zéros au-dessus des "1". D'abord, on utilise $L_2 \leftarrow L_2 + L_3$ pour faire partir le "-1" de la troisième colonne, ce qui donne

$$\left(\begin{array}{ccc|c} 1 & 2 & 1 & 0 \\ 0 & 1 & 0 & -10/3 \\ 0 & 0 & 1 & -7/3 \end{array}\right).$$

Ensuite, on peut faire disparaître le "2" de la deuxième colonne en faisant $L_1 \leftarrow L_1 - 2L_2$,

$$\left(\begin{array}{ccc|c} 1 & 0 & 1 & 20/3 \\ 0 & 1 & 0 & -10/3 \\ 0 & 0 & 1 & -7/3 \end{array}\right),$$

puis le "1" de la troisième colonne en faisant $L_1 \leftarrow L_1 - L_3$:

$$\left(\begin{array}{ccc|c} 1 & 0 & 0 & 27/3 \\ 0 & 1 & 0 & -10/3 \\ 0 & 0 & 1 & -7/3 \end{array}\right),$$

qui est *la* forme échelonnée réduite de la matrice de départ. On remarque que cette matrice correspond au système

$$\left\{\begin{array}{rrrcr} x_1 & & & = & 27/3 \\ & x_2 & & = & -10/3 \\ & & x_3 & = & -7/3 \end{array}\right.,$$

pour lequel on a la solution "sous les yeux" : $(x_1, x_2, x_3) = (\frac{27}{3}, \frac{-10}{3}, \frac{-7}{3})$. ⋄

On l'a mentionné plus haut, une matrice peut être échelonnée d'une infinité de façons différentes. Pour la réduite, c'est différent :

Théorème 1.29. *Toute matrice A peut être transformée, à l'aide d'un nombre fini de transformations élémentaires, en une matrice échelonnée réduite. De plus, cette échelonnée réduite est unique et ne dépend pas des opérations élémentaires avec lesquelles elle a été obtenue ; on l'appelle* ***la réduite de Gauss de A***.

Preuve: (Omise pour l'instant.) □

Dans la résolution d'un système, on n'a pas besoin d'aller jusqu'à la réduite pour obtenir la solution ; n'importe quelle échelonnée suffit. Mais puisque la réduite est *unique*, comme l'affirme le théorème ci-dessus, elle fournit des informations importantes sur la matrice de départ, que nous exploiterons dans les chapitres suivants, en particulier dans l'étude des applications linéaires. D'un point de vue pratique, elle nous permet toujours d'identifier les variables de base et les variables libres, et donc de résoudre complètement le système.