Mathématiques Numériques
Chapitre 2 : systèmes linéaires
Telecom Nancy

Bruno Pinçon (I.E.C.L.) bruno.pincon@univ-lorraine.fr

2018-2019

Introduction I

Résoudre un système linéaire est une opération très courante :

- Un grand nombre de problèmes d'ingénierie conduisent, après modélisation puis éventuellement une procédure de discrétisation, à résoudre des systèmes linéaires : calculs sur des réseaux électriques ou hydrauliques en régime stationnaire, calculs de structures, etc.
- Cette opération intervient aussi comme "sous-méthode" dans de nombreux algorithmes (problèmes d'interpolation, problèmes de moindres carrés, résolution de systèmes non-linéaires, optimisation, automatique, traitement du signal, etc.).

Notations I

Le problème est le suivant : “étant donnés une matrice A et un vecteur b (les données du problème), trouver le vecteur x solution de” :

$$Ax = b \quad \text{avec} \quad \begin{cases} A \in \mathcal{M}_{nn}(\mathbb{R}) = \mathbb{R}^{n\times n} \\ x \in \mathbb{R}^n \\ b \in \mathbb{R}^n \end{cases} \tag{1}$$

Notations :

- Si A est une matrice (n, m), A_i désigne la matrice $(1, m)$ (appelée aussi vecteur ligne) formée par la i ème ligne de A et A^j désigne la matrice $(n, 1)$ (appelée aussi vecteur colonne) formée par la j ème colonne de A.

Notations II

- On remarque que pour tout vecteur $x \in \mathbb{R}^n$ (on considèrera toujours chaque vecteur comme une matrice unicolonne) :

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1 \underbrace{\begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}}_{e^1} + x_2 \underbrace{\begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \end{bmatrix}}_{e^2} + \cdots + x_n \underbrace{\begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}}_{e^n} = \sum_{k=1}^{n} x_k e^k$$

qui fait apparaître la base $(e^1, e^2, \ldots, e^n)$ appelée base canonique de $\mathbb{R}^n$ (cette famille est génératrice par définition mais on montre très facilement qu'elle est libre). On a $(e^j)_i = \delta_{i,j}$ (le symbole de Kronecker).

Notations III

- Si A est une matrice (n, m) et B une matrice (m, p) alors le produit matriciel AB est bien défini et donne une matrice de taille (n, p) avec :

$$(AB)_{i,j} = \sum_{k=1}^{m} a_{i,k} b_{k,j}, \ \forall (i, j) \in [\![1, n]\!] \times [\![1, p]\!]$$

- Si A est une matrice (n, m) alors $A^\top$, la transposée de A est une matrice (m, n) avec $(A^\top)_{i,j} = a_{j,i}$ (les lignes de A forment les colonnes de $A^\top$).

Notations IV

- Si x et y sont deux vecteurs de $\mathbb{R}^n$ alors :

$$(x|y) := x_1 y_1 + \cdots + x_n y_n = \sum_{k=1}^{n} x_k y_k$$

est le produit scalaire canonique de $\mathbb{R}^n$. Il permet de généraliser la notion d'orthogonalité usuelle. On remarque que $(x|y) = x^\top y = y^\top x$.

Interprétation géométrique d'un système linéaire I

On connaît une condition **algébrique** qui assure l'existence (quelque soit le second membre b) et l'unicité d'une solution de (1), c'est à dire :

$$\forall b \in \mathbb{R}^n, \exists ! x \in \mathbb{R}^n \ : \ Ax = b$$

qui est $\det(A) \neq 0$. Cependant il est utile d'avoir en tête la vue géométrique comme **intersection d'hyperplans affines** :
En dimension 2 (1) s'écrit :

$$\begin{array}{rcl} a_{1,1}x_1 + a_{1,2}x_2 & = & b_1 \\ a_{2,1}x_1 + a_{2,2}x_2 & = & b_2 \end{array}$$

il s'agit de deux équations de droite (on suppose que les coefficients de chaque droite ne sont pas tous nuls : $\forall i, \exists j : a_{i,j} \neq 0$) et on cherche donc l'intersection de ces 2 droites (on cherche x qui appartient à la droite 1 **et** à la droite 2).

Interprétation géométrique d'un système linéaire II

On sait que :

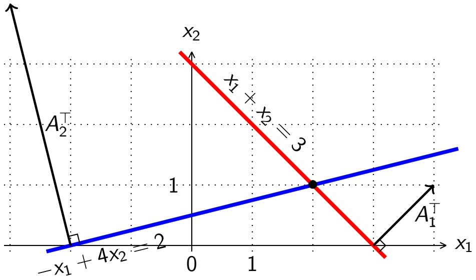
- si les deux droites sont confondues, il y a une infinité de solutions ;
- si les droites sont non confondues mais parallèles il n'y a pas de solution ;
- sinon (droites non parallèles) il y a un et un seul point d'intersection, c'est à dire une et une seule solution.

On peut remarquer que ce dernier cas est "générique" : si les coefficients de chaque droite étaient choisis au hasard (selon une certaine loi à expliciter...), il y aurait vraiment très peu de chance que les droites soient parallèles. Rmq : le fait que les droites sont ou ne sont pas parallèles ne dépend pas du second membre b.

Interprétation géométrique d'un système linéaire III

Exemple le système linéaire $Ax = b$ suivant :

$$\underbrace{\begin{bmatrix} 1 & 1 \\ -1 & 4 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}}_{x} = \underbrace{\begin{bmatrix} 3 \\ 2 \end{bmatrix}}_{b} \iff \left\{ \begin{array}{rcl} x_1 + x_2 & = & 3 \\ -x_1 + 4x_2 & = & 2 \end{array} \right.$$



Interprétation géométrique d'un système linéaire IV

Remarque : dans cette figure apparaît le fait géométrique suivant bien connu : une droite d'équation $\alpha_1 x_1 + \alpha_2 x_2 = \beta$ est perpendiculaire au vecteur $\alpha := [\alpha_1, \alpha_2]^\top$ constitué par ses coefficients. En effet pour obtenir un vecteur v parallèle à cette droite il suffit de faire la différence entre deux points quelconques qui sont sur la droite, par exemple x et x' dont les coordonnées vérifient l'équation de la droite :

$$\begin{array}{rcl} \alpha_1 x_1 + \alpha_2 x_2 & = & \beta \\ \alpha_1 x_1' + \alpha_2 x_2' & = & \beta \end{array}$$

En faisant la différence de ces deux équations, il vient :

$$\alpha_1 (\underbrace{x_1 - x_1'}_{=v_1}) + \alpha_2 (\underbrace{x_2 - x_2'}_{=v_2}) = 0$$

Interprétation géométrique d'un système linéaire V

ce qui s'écrit encore $(\alpha|v) = 0$, tout vecteur v parallèle à la droite est orthogonal avec le vecteur α.
Cette interprétation permet de comprendre à quoi correspond une équation linéaire de la forme suivante dans $\mathbb{R}^n$ (où au moins un des coefficients appelé α_{i^*} est non nul) :

$$\alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n = \beta \tag{2}$$

Il s'agit d'un hyperplan affine (dans un espace vectoriel à n dimensions c'est un sous espace affine de dimension $n - 1$) passant par :

$$x^* = \frac{\beta}{\alpha_{i^*}} e^{i^*}$$

et orthogonal au vecteur $\alpha = [\alpha_1, \ldots, \alpha_n]^\top$. En effet d'après cette interprétation géométrique, un point x appartiendrait à l''hyperplan si

Interprétation géométrique d'un système linéaire VI

et seulement si : $(\alpha|x - x^*) = 0$. En développant ce produit scalaire il vient :

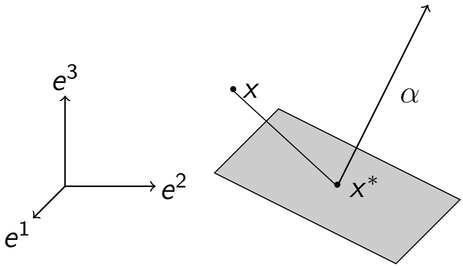
$$(\alpha|x) - \underbrace{(\alpha|x^*)}_{=\beta} = 0 \iff \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n = \beta$$

et on retrouve bien l'équation (2). Une notion qui sera utilisée plus tard (en particulier pour le cours de GRO) est celle de demi-espace : un tel hyperplan permet de séparer l'espace en 2 parties. Par exemple les points qui sont "au-dessus" de l'hyperplan ("au-dessus" au sens de la direction du vecteur α) et comprenant l'hyperplan lui même seront caractérisés par :

$$(\alpha|x - x^*) \geq 0 \iff \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n \geq \beta$$

Interprétation géométrique d'un système linéaire VII

Exemple en 3D, avec l'hyperplan $0x_1 + x_2 + 2x_3 = 3$:



Interprétation géométrique d'un système linéaire VIII

Le point x dessiné vérifie $(\alpha | x - x^*) > 0$ (c'est à dire $x_2 + 2x_3 > 3$) il est bien strictement "au-dessus" du plan vis à vis de l'orientation donnée par le vecteur $\alpha = [0, 1, 2]^\top$ (le point du plan x^* choisi a pour coordonnée $x^* = [0, 3, 0]^\top$).

Ainsi pour un système linéaire $Ax = b$ a n équations et n inconnues, l'interprétation géométrique se généralise de la façon suivante :

L'ensemble des solutions de $Ax = b$ est l'intersection des n hyperplans d'équation
$A_i x = b_i \iff a_{i,1}x_1 + a_{i,2}x_2 + \cdots + a_{i,n}x_n = b_i, \; i = 1, \ldots, n.$

Sur le déterminant d'une matrice I

définition

On peut définir le déterminant de $A \in \mathbb{R}^{n \times n}$ comme l'application multi-linéaire alternée des colonnes de A dans $\mathbb{R}$ qui vaut 1 sur la matrice identité. C'est à dire que $\det(A) := \det(A^1, A^2, \dots, A^n)$ où :

- multilinéaire veut dire linéaire en chacun des arguments :

$$\det(\dots, \alpha u + \beta v, \dots) = \alpha \det(\dots, u, \dots) + \beta \det(\dots, v, \dots)$$

- alternée veut dire que le déterminant est nul si deux arguments sont identiques :

$$\det(\dots, u, \dots, u, \dots) = 0$$

- et enfin $\det(e^1, \dots, e^n) = 1$.

Sur le déterminant d'une matrice II

Exercice

Montrer que si on échange deux colonnes d'un déterminant, il change de signe :

$$\det(\ldots, u, \ldots, v, \ldots) = -\det(\ldots, v, \ldots, u, \ldots)$$

Aide : partir de $\det(\ldots, u+v, \ldots, u+v, \ldots) = 0$.

Sur le déterminant d'une matrice III

Correction

$$
\begin{aligned}
\det(\dots, u+v, \dots, u+v, \dots) &= 0 \\
\det(\dots, u, \dots, u+v, \dots) + \det(\dots, v, \dots, u+v, \dots) &= 0 \\
\underbrace{\det(\dots, u, \dots, u, \dots)}_{=0} + \det(\dots, u, \dots, v, \dots) & \\
+ \det(\dots, v, \dots, u, \dots) + \underbrace{\det(\dots, v, \dots, v, \dots)}_{=0} &= 0 \\
\det(\dots, u, \dots, v, \dots) + \det(\dots, v, \dots, u, \dots) &= 0
\end{aligned}
$$

Sur le déterminant d'une matrice IV

Quelques propriétés :

- Un déterminant peut se calculer récursivement par la formule de Laplace, en notant $[A]_{(-i,-j)}$ la matrice obtenue en supprimant la ligne i et la colonne j de A, on a :
  - développement par rapport à la colonne j de la matrice :

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j} a_{i,j} \det([A]_{(-i,-j)})$$

  - développement par rapport à la ligne i de la matrice :

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{i,j} \det([A]_{(-i,-j)})$$

Sur le déterminant d'une matrice V

On se ramène alors à des calculs de déterminants de matrices $(n-1, n-1)$, et, comme on sait calculer le déterminant d'une matrice $(1,1)$: $\det(a) = a$; on peut en déduire une méthode de calcul (très inefficace) d'un déterminant quelconque.

Exercice

1. En utilisant la définition donnée précédemment du déterminant d'une matrice, montrer que si A est une matrice $(1,1), A = [a]$, alors $det(A) = a$.

2. En utilisant la formule de Laplace, retrouver la formule du déterminant d'une matrice $(2,2)$ (Si $A = \begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix}$ alors $det(A) = a_{1,1}a_{2,2} - a_{1,2}a_{2,1}$).

Sur le déterminant d'une matrice VI

Correction

1. $\det(a) = a\det(1)$ par linéarité par rapport au premier (et seul) argument, puis $\det(1) = 1$ car le déterminant vaut 1 sur la "matrice" identité $(1, 1)$.

2. En utilisant un développement par rapport à la première colonne :

$$
\begin{aligned}
det(A) = \begin{vmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{vmatrix} &= \sum_{i=1}^{2} (-1)^{i+1} a_{i,1} det([A]_{(-i,-1)}) \\
&= a_{1,1} det(a_{2,2}) - a_{2,1} det(a_{1,2}) \\
&= a_{1,1} a_{2,2} - a_{1,2} a_{2,1}
\end{aligned}
$$

Sur le déterminant d'une matrice VII

- $\det(A^1, A^2, \dots, A^n)$ est aussi égal à la mesure signée du parallélépipède généralisé défini par les n vecteurs dans $\mathbb{R}^n$: pour $n = 2$ c'est l'aire (orienté) du parallélogramme A^1, A^2, pour $n = 3$ c'est le volume (orienté) du parallélépipède A^1, A^2,A^3, etc. Ainsi pour $n = 2$ si A^1 et A^2 sont colinéaires, le parallélogramme correspondant est dégénéré et son aire est nulle (donc $\det(A) = 0$). De même pour $n = 3$, si les 3 vecteurs sont liés, le parallélépipède est aussi dégénéré et son volume est nul, etc.
- On a $\det(A^\top) = \det(A)$.
- Si A et B sont 2 matrices (n, n), $\det(AB) = \det(A)\det(B)$.

Systèmes linéaires généraux I

On peut aussi s'intéresser à des systèmes linéaires n'ayant pas forcément le même nombre d'équations (m) que d'inconnues (n) : étant donné $A \in \mathbb{R}^{m \times n}$ et $b \in \mathbb{R}^m$ on cherche $x \in \mathbb{R}^n$ vérifiant $Ax = b$. Ici A prend un vecteur de $\mathbb{R}^n$ (l'espace de départ) et lui associe un vecteur de $\mathbb{R}^m$ (l'espace but) : $\left\{ \begin{array}{ccc} \mathbb{R}^n & \to & \mathbb{R}^m \\ x & \mapsto & Ax \end{array} \right.$

Le point de vue géométrique précédent est toujours valable :

L'ensemble des solutions est l'intersection (peut être vide) des m hyperplans affines d'équation $A_i x = b_i$ de $\mathbb{R}^n$.

Rmq : une autre interprétation est la suivante :

> *l'ensemble des solutions est l'ensemble (peut être vide) des $x \in \mathbb{R}^n$ tels qu'on obtient b par combinaison linéaire des colonnes de A ($\sum_{j=1}^{n} x_j A^j = b$).*

Systèmes linéaires généraux II

Existence d'une solution : l'espace image. L'ensemble des $y \in \mathbb{R}^m$ tels qu'il existe $x \in \mathbb{R}^n$ avec $y = Ax$ est appelée ImA :

$$ImA := \{y \in \mathbb{R}^m : \exists x \in \mathbb{R}^n, y = Ax\} = \{Ax : x \in \mathbb{R}^n\}$$

Exercice

1. Montrer que ImA est un sous-espace vectoriel de l'espace but $\mathbb{R}^m$.
2. Montrer que $(A^1, \ldots, A^n)$ est une famille génératrice de ImA, en d'autres termes : $ImA = Vect(A^1, \ldots, A^n)$.
   Aide : montrer que $Ax = \sum_j x_j A^j$.

Systèmes linéaires généraux III

Correction

1. Il est clair que $ImA \neq \emptyset$ (par exemple $O_m = A0_n \in ImA$). Soient $y, y' \in ImA$ et $\alpha, \alpha' \in \mathbb{R}$, est-ce que $y'' := \alpha y + \alpha' y' \in ImA$? La réponse est oui si on trouve un vecteur $x'' \in \mathbb{R}^n$ tel que $y'' = Ax''$. Comme $y \in ImA$, $\exists x \in \mathbb{R}^n$ tel que $y = Ax$. De même $\exists x' \in \mathbb{R}^n$ tel que $y' = Ax'$. Alors $x'' := \alpha x + \alpha' x'$ vérifie :

$$Ax'' = A(\alpha x + \alpha' x') = \alpha Ax + \alpha' Ax' = \alpha y + \alpha' y' \quad \square$$

2.

$$\begin{bmatrix} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,n}x_n \\ a_{2,1}x_1 + a_{2,2}x_2 + \cdots + a_{2,n}x_n \\ \\ a_{m,1}x_1 + a_{m,2}x_2 + \cdots + a_{m,n}x_n \end{bmatrix} = x_1 \begin{bmatrix} a_{1,1} \\ a_{2,1} \\ \\ a_{m,1} \end{bmatrix} + x_2 \begin{bmatrix} a_{1,2} \\ a_{2,2} \\ \\ a_{m,2} \end{bmatrix} + \cdots + x_n \begin{bmatrix} a_{1,n} \\ a_{2,n} \\ \\ a_{m,n} \end{bmatrix}$$

Systèmes linéaires généraux IV

L'existence d'au moins une solution pour b donné est donc équivalente à :

$$b \in ImA$$

Rmq :

- Si $ImA = \mathbb{R}^m$ (l'espace but entier est atteint par A) on a existence pour tout second membre b.
- Par contre lorsque $rangA := \dim(ImA) < m$, on aura pas forcément existence d'une solution. Si b est choisi au hasard, il y a très peu de chance qu'il y en ait une : supposons par exemple $rangA = 2$ et $m = 3$, ImA est donc un plan vectoriel et b choisi au hasard dans $\mathbb{R}^3$ a peu de chance de se trouver sur ce plan.

Systèmes linéaires généraux V

Unicité d'une (éventuelle) solution : l'espace noyau de A **(**$KerA$**)**. Le noyau de A est l'ensemble suivant :

$$KerA := \{x \in \mathbb{R}^n : Ax = 0\}$$

On montre facilement que $KerA$ est un sous-espace vectoriel de l'espace de départ.

Proposition : on a unicité si et seulement si $KerA = \{0\}$.

Preuve : Soient 2 solutions x et x', vérifiant donc $Ax = b$ et $Ax' = b$. Par différence $Ax - Ax' = A(x - x') = b - b = 0$ et donc $x - x' \in KerA$. Ainsi si $KerA$ ne contient que le vecteur nul, $x = x'$.

Systèmes linéaires généraux VI

Réciproquement si $dim(KerA) \geq 1$ (Si $KerA$ n'est pas restreint au vecteur nul, il contient au moins une droite vectorielle), et si x vérifie $Ax = b$ alors tout vecteur $x + u$ avec $u \in KerA$ est aussi solution du système linéaire :

$$A(x + u) = \underbrace{Ax}_{=b} + \underbrace{Au}_{=0} = b$$

dans un tel cas ($b \in ImA$ et $KerA \neq \{0\}$) on a donc une infinité de solutions. $\square$

Systèmes linéaires généraux VII

Théorème du rang : on a la relation suivante :

$$dim(KerA) + \underbrace{dim(ImA)}_{=rangA} = dim(\mathbb{R}^n) = n$$

Si on revient maintenant au cas carré ($n = m$) on voit que si $KerA = \{0\}$ alors $rangA = n$, l'unicité implique l'existence (pour tout second membre). Inversement si $rangA = n$ alors $KerA = \{0\}$, l'existence (pour tout second membre) implique l'unicité. Pour vérifier qu'une matrice est inversible il suffit donc de vérifier l'une de ces deux conditions.

En pratique... I

On veut donc résoudre un système linéaire carré. Soit :

- $\det(A) \neq 0$ (ou $KerA = \{0\}$ ou $rangA = n$) et on a existence et unicité pour tout second membre,
- soit $\det(A) = 0$ (ce qui est exceptionnel si les coefficients de A sont choisis au hasard) et alors on a existence que si $b \in ImA$ (ce qui est exceptionnel si b est choisi au hasard) et dans ce cas on a une infinité de solutions.

En pratique... II

Quelques questions et remarques :

1. Vous avez appris en 1er cycle comment résoudre un système linéaire par une méthode de type Gauss. Est-ce la bonne méthode ?

   Réponse : oui ! Voilà une bonne nouvelle. Attention en premier cycle on apprend plutôt la méthode de Gauss-Jordan qui permet d'obtenir directement la solution ; dans ce cours on va utiliser la méthode de Gauss qui conduit, partant de $Ax = b$ *à obtenir un système linéaire équivalent* $Ux = b'$ *où* U *est une matrice triangulaire supérieure (il y a donc une étape supplémentaire) ; néanmoins on peut montrer (malgré l'étape supplémentaire de la résolution du système triangulaire) que c'est la méthode la plus économique et c'est donc elle qui est dans les bonnes bibliothèques d'algèbre linéaire. Il existe cependant d'autres méthodes (dite itératives comme le gradient conjugué, GMRES,*

En pratique... III

etc.) qui sont utilisées pour résoudre certains grands systèmes linéaires creux (creux veut dire que la matrice comporte une très grande majorité de coefficients nuls).

2. Doit-on d'abord vérifier que $\det(A) \neq 0$ (ou par exemple que $KerA = \{0\}$) avant de tenter un calcul ?

   Réponse : non on découvre l'inversibilité par la méthode de Gauss (le déterminant de A s'obtient comme un sous-produit de cette méthode, dans la plupart des cas c'est la méthode la plus rapide pour le calculer !).

En pratique... IV

3. On a souvent plusieurs systèmes linéaires à résoudre avec la même matrice : $Ax^{(k)} = b^{(k)}, k = 1, \dots, K$. Les différents seconds membres $b^{(k)}$ n'étant pas forcément connus au même moment (par exemple $b^{(2)}$ peut dépendre de $x^{(1)}$, etc. Comment optimiser les calculs ?

   Réponse : en réinterprétant la méthode de Gauss comme une factorisation de la matrice A (cf troisième partie de ce chapitre).

4. En arithmétique flottante on sait que chaque calcul élémentaire peut être entaché d'une (généralement très petite) erreur ; si A est inversible (auquel cas on devrait pouvoir calculer la solution) quels critères pourraient nous permettre d'avoir une indication sur la précision de la solution obtenue ?

   Réponse : on pourrait penser que la magnitude de $|det(A)|$ pourrait donner une indication : si $|det(A)| >> 1$ est grand alors

En pratique... V

(avec une bonne méthode de Gauss) les petites erreurs ne s'amplifie pas trop... Il n'en est rien, la bonne notion est celle de conditionnement de la matrice $A = \|A\| \|A^{-1}\|$ *qui sera abordée dans la dernière partie de ce chapitre.*

Résoudre un système triangulaire I

Résoudre $Tx = b$ lorsque la matrice T est triangulaire supérieure (c'est à dire lorsque $t_{i,j} = 0$ pour $j < i$) est assez simple ; en tenant compte des coefficients nuls de la matrice, $Tx = b$ s'écrit :

$$
\begin{array}{cccccccccc}
t_{1,1}x_1 + & t_{1,2}x_2 & + \dots & \dots & \dots & \dots & + \, t_{1,n}x_n & = & b_1 \\
 & t_{2,2}x_2 & + \dots & \dots & \dots & \dots & + \, t_{2,n}x_n & = & b_2 \\
 & & \ddots & \dots & \vdots & \vdots & \vdots & \vdots & \vdots \\
 & & & t_{i,i}x_i & + \dots & \dots & + \, t_{i,n}x_n & = & b_i \\
 & & & & \ddots & \vdots & \vdots & \vdots & \vdots \\
 & & & & & t_{n-1,n-1}x_{n-1} & + \, t_{n-1,n}x_n & = & b_{n-1} \\
 & & & & & & t_{n,n}x_n & = & b_n
\end{array}
$$

On peut montrer que $\det(T) = \prod_{k=1}^{n} t_{k,k}$ (cf exercice 2 feuille 2) et par conséquent T est inversible si et seulement si tous ses coefficients diagonaux sont non nuls.

Résoudre un système triangulaire II

Pour résoudre un tel système il suffit de commencer (**étape 1**) par résoudre la dernière équation :

$$t_{n,n}x_n = b_n \iff x_n = \frac{b_n}{t_{n,n}}$$

ce qui nous donne x_n. Puis (**étape 2**) connaissant x_n l'avant dernière équation nous donne x_{n-1} :

$$t_{n-1,n-1}x_{n-1} + t_{n-1,n}x_n = b_{n-1} \iff x_{n-1} = (b_{n-1} - t_{n-1,n}x_n)/t_{n-1,n-1}$$

L'étape $n - i + 1$ utilise la i ème équation et à ce moment de l'algorithme on connait $x_{i+1}, \ldots, x_n$, on en déduit donc x_i (sachant que $t_{i,i} \neq 0$) :

$$t_{i,i}x_i + t_{i,i+1}x_{i+1} + \cdots + t_{i,n}x_n = b_i \iff x_i = \left(b_i - \sum_{j=i+1}^{n} t_{i,j}x_j\right)/t_{i,i}$$

Résoudre un système triangulaire III

D'où l'algorithme :

$$
\begin{array}{l}
\textbf{pour } i = n, n-1, \ldots, 1 \\
\qquad x_i \leftarrow \left(b_i - \sum_{j=i+1}^{n} t_{i,j} x_j \right) / t_{i,i}
\end{array}
$$

ou encore en écrivant la somme en pseudo-code :

$$
\begin{array}{l}
\textbf{pour } i = n, n-1, \ldots, 1 \\
\qquad temp \leftarrow b_i \\
\qquad \textbf{pour } j = i+1, \ldots, n \\
\qquad \qquad temp \leftarrow temp - t_{i,j} x_j \\
\qquad x_i \leftarrow temp / t_{i,i}
\end{array}
$$

Résoudre un système triangulaire IV

Pour cet algorithme (parfois appelé “remontée”) il est assez simple de calculer le nombre d’opérations arithmétiques (cf exercice 5 feuille 2), on obtient :

$$C_{remontee} : \frac{n(n-1)}{2} \text{ multiplications}, \quad \frac{n(n-1)}{2} \text{ additions}, \quad n \text{ divisions}$$

De même il est tout aussi facile de résoudre un système triangulaire inférieur (exercice).

“Triangulariser” un système linéaire I

La méthode de Gauss consiste à effectuer des transformations successives ($n-1$ en tout) sur (1) qui conduisent à un système linéaire équivalent dont la matrice est triangulaire supérieure (on a vu précédement qu'il était très simple de le résoudre !) :

$$Ax = b \xrightarrow{tr.1} A^{(1)}x = b^{(1)} \xrightarrow{tr.2} A^{(2)}x = b^{(2)} \dots \xrightarrow{tr.n-1} A^{(n-1)}x = b^{(n-1)}$$

Écrivons l'équation matricielle (1) comme système d'équations linéaires :

$$\left\{ \begin{array}{ccc} a_{1,1}x_1 + a_{1,2}x_2 + \dots \cdots + a_{1,n}x_n & = & b_1 \\ a_{2,1}x_1 + a_{2,2}x_2 + \dots \cdots + a_{2,n}x_n & = & b_2 \\ \vdots & & \vdots \\ a_{i,1}x_1 + a_{i,2}x_2 + \dots \cdots + a_{i,n}x_n & = & b_i \\ \vdots & & \vdots \\ a_{n,1}x_1 + a_{n,2}x_2 + \dots \cdots + a_{n,n}x_n & = & b_n \end{array} \right. \quad (3)$$

“Triangulariser” un système linéaire II

Gauss première étape. Idée : éliminer la variable x_1 des équations $2, 3, \ldots, n$.

(i) Recherche d'un pivot non nul : si $a_{1,1} = 0$, on cherche alors $a_{i^*,1}$ tel que $a_{i^*,1} \neq 0$ et on échange les équations 1 et i^*. Rmq : si tous les coefficients de la première colonne de A sont nuls alors $\det(A) = 0$.

(ii) Supposons donc $a_{1,1} \neq 0$ (ce qui est donc toujours possible modulo un échange de lignes si A est inversible), l'idée de la méthode est de soustraire la première équation multipliée par le “bon” coefficient à la deuxième équation, de façon à éliminer la variable x_1 dans la nouvelle (deuxième) équation ainsi obtenue. On recommence ensuite cette manip sur l'équation 3, puis sur la 4, etc, et enfin sur la dernière.

“Triangulariser” un système linéaire III

exercice

À vous de trouver le bon coefficient pour éliminer la variable x_1 dans les équations $2, 3, \dots, n$. On suppose donc (via un éventuel échange de lignes et rénumérotation) que $a_{1,1} \neq 0$. La manip proposée est :

$$equ_i^{(1)} := equ_i - coef_i^{(1)} \times equ_1$$

1. Écrire la nouvelle équation i ($equ_i^{(1)}$) ainsi obtenue.
2. En déduire la valeur coefficient $coef_i^{(1)}$ (celui qui permet à la nouvelle équation i ne ne plus dépendre de x_1).
3. On obtient donc un nouveau système linéaire $A^{(1)}x = b^{(1)}$: donner l'expression des coefficients de $A^{(1)}$ et $b^{(1)}$.

"Triangulariser" un système linéaire IV

Solution :

1. La nouvelle équation i s'écrit ($i \in [\![2, n]\!]$) :

$$\underbrace{(a_{i,1} - coef_i^{(1)} a_{1,1})}_{=a_{i,1}^{(1)}} x_1 + \underbrace{(a_{i,2} - coef_i^{(1)} a_{1,2})}_{=a_{i,2}^{(1)}} x_2 + \cdots + \underbrace{(a_{i,j} - coef_i^{(1)} a_{1,j})}_{=a_{i,j}^{(1)}} x_j + \cdots + \underbrace{(a_{i,n} - coef_i^{(1)} a_{1,n})}_{=a_{i,n}^{(1)}} x_n = \underbrace{b_i - coef_i^{(1)} b_1}_{=b_i^{(1)}}$$

2. On a $a_{i,1}^{(1)} = 0 \iff coef_i^{(1)} = \frac{a_{i,1}}{a_{1,1}}$.

3.
   - Pour $j \in [\![1, n]\!]$: $a_{1,j}^{(1)} = a_{1,j}$ et $b_1^{(1)} = b_1$ (première ligne inchangée).
   - Pour $i \in [\![2, n]\!]$: $a_{i,1}^{(1)} = 0$ (plus de dépendance en x_1)
   - Pour $i, j \in [\![2, n]\!]$: $a_{i,j}^{(1)} = a_{i,j} - coef_i^{(1)} a_{1,j}$, $b_i^{(1)} = b_i - coef_i^{(1)} b_1$

"Triangulariser" un système linéaire V

On obtient donc le (nouveau mais équivalent) système linéaire :

$$A^{(1)}x = b^{(1)} : \left\{ \begin{array}{cccc} a_{1,1}x_1 & +a_{1,2}x_2 + \ldots \cdots + a_{1,n}x_n & = & b_1 \\ 0 & +a^{(1)}_{2,2}x_2 + \ldots \cdots + a^{(1)}_{2,n}x_n & = & b^{(1)}_2 \\ \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ 0 & +a^{(1)}_{n,2}x_2 + \ldots \cdots + a^{(1)}_{n,n}x_n & = & b^{(1)}_n \end{array} \right. \tag{4}$$

Remarques sur la transformation $Ax = b \leftrightarrow A^{(1)}x = b^{(1)}$:

1. la transformation est **inversible** : si on utilise $equ_i^{(1)} + coef_i^{(1)} \times equ_1^{(1)}$ on retrouve les équations initiales ! Ainsi toute solution de $Ax = b$ sera une solution de $A^{(1)}x = b^{(1)}$ et inversement : les deux systèmes sont équivalents (en ce sens qu'ils ont les mêmes solutions).

“Triangulariser” un système linéaire VI

2. les manipulations sur les équations du système sont équivalentes à des manipulations sur les lignes de la matrice A et du vecteur b : ainsi éliminer l'inconnue x_1 des équations $2, 3, \ldots, n$ correspond à faire apparaître des zéros sous la diagonale dans la première colonne de la matrice $(A^{(1)})$ avec les opérations :

$$ligne_i^{(1)} = ligne_i - coef_i^{(1)} \times ligne_1 \;\text{ pour } 2 \leq i \leq n.$$

   Il faut bien sûr appliquer la même transformation sur b.

3. En choisissant $coef_1^{(1)} = 0$ la manipulation sur la première équation (cad l'identité !) rentre dans ce formalisme.

4. la transformation est **linéaire**, on verra qu'elle correspond à la multiplication par une matrice inversible $M^{(1)}$, on a $A^{(1)} = M^{(1)}A$ et $b^{(1)} = M^{(1)}b$;

"Triangulariser" un système linéaire VII

Et la suite ? La deuxième étape revient à faire exactement la même chose sur le sous-système (en rouge dans un transparent précédent) constitué par les équations $2, 3, \dots, n$. Comme ce sous-système ne dépend pas de la variable x_1, c'est un système de $n-1$ équations à $n-1$ inconnues, sur lequel on applique la même technique :

- si le coefficient $a_{2,2}^{(1)} \neq 0$ on procède tout de suite à la phase d'élimination de l'inconnue x_2 dans les équations $3, 4, \dots, n$;
- sinon on cherche un coefficient non nul :

$$a_{i,2}^{(1)} \neq 0, \qquad 2 < i \leq n$$

  puis on échange les équations 2 et i et l'on procède alors à la phase d'élimination.

“Triangulariser” un système linéaire VIII

Remarque : si A est inversible, on peut toujours trouver $i \in [\![2, n]\!]$ tel que $a_{i,2}^{(1)} \neq 0$; en effet comme la matrice $A^{(1)}$ a le découpage par blocs suivant :

$$A^{(1)} = \left(\begin{array}{c|c} a_{1,1} & a_{1,2} \dots a_{1,n} \\ \hline 0 & \\ \vdots & [A^{(1)}]_{[-1,-1)} \\ 0 & \end{array}\right)$$

le calcul de son déterminant (en développant par rapport à la première colonne) donne :

$$\det(A^{(1)}) = a_{1,1} \det([A^{(1)}]_{(-1,-1)}).$$

D'autre part comme $Ax = b$ et $A^{(1)}x = b^{(1)}$ sont équivalents, A est inversible **ssi** $A^{(1)}$ l'est ! Donc $det(A) \neq 0 \iff \det(A^{(1)}) \neq 0$.

"Triangulariser" un système linéaire IX

Ainsi, comme $a_{1,1} \neq 0$, la première colonne de la matrice $[A^{(1)}]_{(-1,-1)}$ ne peut être nulle (sinon son déterminant serait nul et par conséquent celui de $A^{(1)}$ aussi), c-a-d qu'il existe bien $i \in [\![2, n]\!]$ tel que $a_{i,2}^{(1)} \neq 0$. De la même façon, on montre que ce résultat est aussi vrai pour la suite : si A est inversible on peut trouver, à chaque étape k de Gauss, un coefficient non nul :

$$a_{i,k}^{(k-1)} \neq 0, \quad \text{avec } i \in [\![k, n]\!]$$

Ainsi la méthode de Gauss marche à tous les coups sur une matrice inversible (en arithmétique exacte !).

“Triangulariser” un système linéaire X

Et la fin ? Au bout de $n-1$ étapes (qui consistent donc toutes à faire la même chose sur des systèmes linéaires de plus en plus petits), on obtient le système linéaire équivalent $A^{(n-1)}x = b^{(n-1)}$, où la matrice $A^{(n-1)}$ est triangulaire supérieure :

$$A^{(n-1)} = \begin{pmatrix} a_{1,1} & \times & \dots & \times \\ 0 & a_{2,2}^{(1)} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \times \\ 0 & \dots & 0 & a_{n,n}^{(n-1)} \end{pmatrix}$$

avec ses éléments diagonaux (les pivots) $a_{i,i}^{(i-1)}$ tous non nuls. Il suffit alors d'utiliser l'algorithme vu dans la section précédente pour résoudre ce système triangulaire.

“Triangulariser” un système linéaire XI

exercice : assimiler ce qui précéde et l’écrire algorithmiquement

1. Écrire l’étape k de Gauss (on suppose que l’on a pu mener les étapes $1, \dots, k-1$).
2. Écrire l’algorithme des $n-1$ étapes, en travaillant “en place” (on suppose que l’on a sauvegardé une copie de A et b) en utilisant de plus la stratégie du pivot partiel (recherche du pivot max en valeur absolue dans la colonne). Ne pas écrire l’algorithme de résolution du système linéaire.

L'écriture matricielle de la méthode de Gauss

Plan

Écrire chaque étape de Gauss "matriciellement" va nous conduire à la notion très fructueuse de "factorisation" de la matrice initiale sous la forme :

$$A = LU \text{ ou } PA = LU$$

où :

- L est une matrice triangulaire inférieure (elle contient "l'historique" de la méthode de Gauss qui ne dépend que de la matrice A (et pas du second membre b).
- U est simplement la matrice triangulaire supérieure obtenue à la fin du procédé.
- P est une matrice de permutation qui "enregistre" les éventuelles permutations de lignes pour trouver un pivot non nul (ou le plus grand pivot possible si on utilise la stratégie du pivot partiel).

Action d'une matrice du type $M = I - z(e^k)^\top$ I

Soit B une matrice (n, m), on regarde la transformation opérée par la multiplication MB où $M := I - z(e^k)^\top$ avec $z \in \mathbb{R}^n$, e^k k-ème vecteur de la base canonique de $\mathbb{R}^n$ et I la matrice identité.

exercice

1. Montrer que la matrice M correspond à une matrice identité dans laquelle on a ajouté le vecteur $-z$ dans la k ème colonne.
2. Montrer que $(MB)_i = B_i - z_i B_k, \forall i \in [\![1, n]\!]$, autrement dit :

$$MB = \left(\begin{array}{c} B_1 - z_1 B_k \\ \hline B_2 - z_2 B_k \\ \hline \vdots \\ \hline B_n - z_n B_k \end{array} \right)$$

Aide : (cf ex 1 feuille 2) $(e^i)^\top B = B_i$ et $Be^j = B^j$.

Action d'une matrice du type $M = I - z(e^k)^\top$ II

Solution

1. cf tableau.
2. $$MB = (I - z(e^k)^\top)B = B - z(e^k)^\top B = B - zB_k$$

   d'où :

   $$\forall i \in [\![1, n]\!] : (MB)_i = (e^i)^\top(B - zB_k) = B_i - z_i B_k$$

   Ainsi la transformation est comprise dans le sens suivant : **à chaque ligne i de la matrice B initiale, on a soustrait sa ligne k multipliée par le coefficient z_i.**

Action d'une matrice du type $M = I - z(e^k)^\top$ III

exercice

Montrer que si $z_k = 0$ alors la matrice M est nécessairement inversible d'inverse :

$$L = (M^{-1}) = I + z(e^k)^\top$$

Aide : effectuer le produit LM et utiliser l'associativité du produit matriciel.

Action d'une matrice du type $M = I - z(e^k)^\top$ IV

correction

$$\begin{aligned}
LM &= (I + z(e^k)^\top)(I - z(e^k)^\top) \\
&= I + \underbrace{z(e^k)^\top - z(e^k)^\top}_{0} - z(e^k)^\top z(e^k)^\top \\
&= I - z\left(\left(e^k\right)^\top z\right)\left(e^k\right)^\top \\
&= I - z\,\underbrace{(z|e^k)}_{z_k=0}(e^k)^\top \\
&= I
\end{aligned}$$

Avec cet outil, on obtient facilement les matrices qui permettent de passer d'une étape à l'autre dans la méthode de Gauss (l'échange des lignes pour trouver un pivot non nul, est lui obtenu par une matrice de permutation élémentaire ; nous en parlerons plus loin.) :

Action d'une matrice du type $M = I - z(e^k)^\top$ V

- pour l'étape 1 : $A^{(1)} = M^{(1)}A$ et $b^{(1)} = M^{(1)}b$ avec $M^{(1)} = I - z^{(1)}(e^1)^\top$ le vecteur $z^{(1)}$ étant :

$$z^{(1)} = \begin{pmatrix} 0 \\ a_{2,1}/a_{1,1} \\ \vdots \\ a_{n,1}/a_{1,1} \end{pmatrix}$$

En effet $M^{(1)}$ effectue bien les opérations attendues sur les lignes :

$$\begin{cases} A_1^{(1)} = A_1 - 0 \times A_1 \\ A_i^{(1)} = A_i - z_i^{(1)} \times A_1 \ \text{ pour } 2 \leq i \leq n \end{cases}$$

(et on a le même effet sur b bien sûr).

Action d'une matrice du type $M = I - z(e^k)^\top$ VI

- pour une étape k quelconque ($1 \leq k \leq n-1$), la matrice $A^{(k-1)}$ étant de la forme :

$$
\begin{pmatrix}
a_{1,1} & \times & \times & \times & \times & \times \\
0 & a_{2,2}^{(1)} & \times & \times & \times & \times \\
\vdots & \ddots & \ddots & \times & \times & \times \\
\vdots & \dots & 0 & a_{k,k}^{(k-1)} & \times & \times \\
\vdots & \dots & \vdots & \vdots & \times & \times \\
0 & \dots & 0 & a_{n,k}^{(k-1)} & \times & \times
\end{pmatrix}
$$

si $a_{kk}^{(k-1)} \neq 0$ (dans le cas contraire il faut au préalable, échanger la ligne k avec une ligne $i > k$ de façon à obtenir un pivot non nul) on peut procéder à l'élimination de la variable x_k

Action d'une matrice du type $M = I - z(e^k)^\top$ VII

dans les équations $k+1, \dots, n$, en utilisant la matrice $M^{(k)} = I - z^{(k)}(e^{(k)})^\top$ avec :

$$z^{(k)} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ a^{(k-1)}_{k+1,k}/a^{(k-1)}_{kk} \\ \vdots \\ a^{(k-1)}_{n,k}/a^{(k-1)}_{kk} \end{pmatrix}$$

Remarquons qu'avec ces vecteurs $z^{(k)}$ particuliers, les matrices $M^{(k)}$ sont inversibles d'inverse $I + z^{(k)}(e^{(k)})^\top$, et qu'elles sont aussi triangulaires inférieures puisque les k premières composantes des $z^{(k)}$ sont nulles (cf exercice 2 feuille 2).

Action d'une matrice du type $M = I - z(e^k)^\top$ VIII

De plus comme les coefficients diagonaux de ces matrices triangulaires sont tous égaux à 1, on a aussi $\det(M^{(k)}) = 1$ (de même pour leurs inverses, cf exercice 2 feuille 2).

La factorisation $A = LU$ I

Théorème : la factorisation $A = LU$

Soit une matrice A (n, n) inversible et telle qu'à chaque étape de la méthode de Gauss on ait (sans procéder à des échanges de lignes) :

$$a_{k,k}^{(k-1)} \neq 0 \qquad \forall k \in [\![1, n-1]\!].$$

Alors il existe une unique factorisation de la matrice A de la forme :

$$A = LU$$

où L est une matrice triangulaire inférieure à diagonale unité (c-a-d que $l_{i,i} = 1, \ \forall i$) et U est une matrice triangulaire supérieure inversible U.

La factorisation $A = LU$ II

Preuve abrégée : On pose $U = A^{(n-1)}$ la dernière matrice obtenue (qui possède donc les propriétés attendues, sauf qu' il reste à montrer que $a_{n,n}^{(n-1)}$ est non nul), on a :

$$U = A^{(n-1)} = M^{(n-1)}A^{(n-2)} = M^{(n-1)}M^{(n-2)}A^{(n-3)} = M^{(n-1)}M^{(n-2)}\dots M^{(1)}A \tag{5}$$

où les matrices $M^{(k)}$ sont de la forme $M^{(k)} = I - z^{(k)}(e^k)^\top$ avec $z_i^{(k)} = 0$ pour $i \in [\![1, k]\!]$. Si on prend les déterminants de l'équation ci-dessus :

$$\det(U) = \prod_{i=1}^{n} a_{ii}^{(i-1)} = \left(\prod_{i=1}^{n-1} \det(M^{(i)})\right)\det(A) = \det(A)$$

car $\det(M^{(i)}) = 1$, par conséquent on a nécessairement $a_{n,n}^{(n-1)} \neq 0$. Les inverses de ces matrices sont appelées $L^{(k)}$ et donc

La factorisation $A = LU$ III

$L^{(k)} = I + z^{(k)}(e^k)^\top$ d'après un résultat précédent. En multipliant (5) à gauche par successivement $L^{(n-1)}$, $L^{(n-2)}$, ..., $L^{(1)}$, on obtient :

$$A = L^{(1)} L^{(2)} \dots L^{(n-1)} U \tag{6}$$

On pose alors $L = L^{(1)} L^{(2)} \dots L^{(n-1)}$. Comme produit de matrices triangulaires inférieures à diagonale unité, L est aussi triangulaire inférieure à diagonale unité (cf exercice 2 feuille 2) L'unicité fait l'objet de l'exercice 4 de la feuille 2. $\square$

Pour le calcul effectif de la matrice L on a un résultat assez fort : elle s'obtient sans aucun calcul supplémentaire puisque :

$$L = I + \sum_{k=1}^{n-1} z^{(k)}(e^k)^\top = I + \left(\begin{array}{c|c|c|c|c} z^{(1)} & z^{(2)} & \dots & z^{(n-1)} & 0 \end{array} \right)$$

La factorisation $A = LU$ IV

ce résultat se montrant en généralisant le calcul suivant : le produit de deux matrices $L^{(k)}L^{(k')}$ avec $k < k'$ est égal à :

$$L^{(k)}L^{(k')} = I + z^{(k)}(e^k)^\top + z^{(k')}(e^{k'})^\top$$

(faire le calcul).

Remarques :

1. Cette première factorisation repose sur l'hypothèse suivante : à chaque étape de la méthode de Gauss on suppose que $a_{kk}^{(k-1)} \neq 0$ (sans avoir à effectuer des échanges de lignes). Par exemple la matrice :
$$\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$$
ne convient pas pour cette méthode. Cependant pour certaines matrices qui interviennent dans la pratique (matrices à diagonale

La factorisation $A = LU$ V

dominante, matrices définies positives), cette condition est vérifiée, la factorisation étant de plus stable vis à vis des erreurs d'arrondi numérique dues aux calculs avec les flottants.

2. Pour une matrice générale (ne possédant pas l'une de ces propriétés) cet algorithme peut être très mauvais car un petit pivot peut amplifier les erreurs d'arrondi. La méthode standard consiste, à chaque étape k, à rechercher d'abord le pivot maximum en valeur absolue (ou module dans le cas complexe) dans la colonne k (à partir de la ligne k) :

$$\max_{i \in [\![k,n]\!]} \left| a_{i,k}^{(k-1)} \right|$$

(que l'on suppose atteind en i_0 par exemple) et à échanger les lignes k et i_0. Matriciellement cela revient à multiplier en premier par une matrice de permutation $P^{(k)}$ avant de procéder à la

La factorisation $A = LU$ VI

phase d'élimination (correspondant à la multiplication par la matrice $M^{(k)}$) :

$$A^{(k)} = M^{(k)} P^{(k)} A^{(k-1)}.$$

Cette stratégie appelée "méthode de Gauss à pivot partiel" conduit à une factorisation du type $PA = LU$ où P est une matrice de permutation.

Utilisation d'une factorisation pour résoudre un système linéaire

Lorsque l'on a obtenu une factorisation $A = LU$, résoudre un système linéaire consiste à résoudre deux systèmes triangulaires :
$Ax = b \Leftrightarrow LUx = b$, on pose $Ux = y$ et l'on résoud :

$$\begin{array}{lll} (i) & Ly = b & \text{on obtient } y \text{ (résolution d'un système tri inf)} \\ (ii) & Ux = y & \text{on obtient } x \text{ (résolution d'un système tri sup)} \end{array}$$

Le coût est donc d'environ n^2 multiplications et n^2 additions.
Dans le cas d'une factorisation $PA = LU$, la méthode ci-dessus est précédée de l'application de la permutation sur le second membre :
$Ax = b \Leftrightarrow PAx = Pb \Leftrightarrow LUx = Pb$.

Algorithme et coût de la méthode de Gauss/LU I

Question : différences entre “méthode de Gauss” et “factorisation $A = LU$ ou $PA = LU$?

Réponse : presque rien ! Pour obtenir une factorisation ($A = LU$ ou $PA = LU$) on procède avec une méthode de Gauss qui agit **uniquement sur la matrice** (pas d'opérations sur un second membre) et qui **enregistre les coefficients d'élimination successifs** pour former la matrice L (qui s'obtient directement avec ces coefficients). Dans le cas d'une méthode avec échange de lignes (pivot partiel qui conduit à une factorisation $PA = LU$), on enregistre aussi les permutations successives (en fait un seul tableau d'entiers de dimension n suffit). Dans la méthode de Gauss classique on procède aussi à la fin à la résolution du système triangulaire supérieur.

Algorithme et coût de la méthode de Gauss/LU II

Algorithme basique factorisation $A = LU$. Il utilise un tableau `LU` initialisé par copie du tableau `A`. La factorisation est calculée "en place" dans le tableau `LU`, la place laissée par les 0 de l'étape k est utilisée pour stocker les coefficients d'élimination (stockage "compact" de la factorisation).

```
LU ← A                    initialisation du tableau LU par copie
pour k de 1 à n−1         boucle des n−1 étapes
    si LU_{k,k} = 0
        arrêt de l'algorithme (gérer l'exception...)
    pour i de k+1 à n     les opérations sur les lignes
        LU_{i,k} ← LU_{i,k}/LU_{k,k}     calcul et stockage du coef d'élimination z_i^(k)
        pour j de k+1 à n    mise à jour de la ligne i
            LU_{i,j} ← LU_{i,j} − LU_{i,k} × LU_{k,j}
```

Algorithme et coût de la méthode de Gauss/LU III

Nombre d'opérations arithmétiques de cet algorithme (cf exercice 5 feuille 2) :

$$C_{LU} \ : \left\{ \begin{array}{ll} \frac{n(n-1)(2n-1)}{6} = \frac{1}{3}n^3 + O(n^2) & \text{additions et multiplications} \\ \frac{n(n-1)}{2} = 0(n^2) & \text{divisions} \end{array} \right.$$

Soit $C_{LU} \simeq \frac{1}{3}n^3$ additions et multiplications.

Le coût de la factorisation $PA = LU$ avec la stratégie du pivot partiel n'est que légèrement supérieur car la recherche du coefficient maximum intervient avec un coût en $O(n^2)$ (de même que les opérations d'échanges de lignes, et le coût total de la mise à jour de la permutation peut se faire avec seulement $O(n)$ opérations). De même le coût de la méthode de Gauss classique, est juste un peu plus élevé puisque les opérations sur le second membre et la résolution finale du système triangulaire supérieur rajoutent simplement environ n^2 multiplications et additions.

Pourquoi formaliser Gauss en factorisation ? I

Dans la pratique, on a souvent à résoudre des systèmes linéaires qui comportent tous la même matrice :

$$Ax^{(i)} = b^{(i)} \qquad 1 \leq i \leq m \tag{7}$$

et où les vecteurs $b^{(i)}$ ne sont pas tous connus au même moment [1], par exemple $b^{(i+1)}$ s'obtient avec un calcul qui utilise $x^{(i)}$. Cette situation est très courante dans les problèmes d'ingénierie : souvent la matrice A correspond à la modélisation d'un système, le second membre est l'entrée imposée à ce système et la solution $x = A^{-1}b$ est la réponse du système pour l'entrée b. On a bien sûr envie de calculer les réponses $x^{(i)}$ qui correspondent à diverses entrées $b^{(i)}$.

Pourquoi formaliser Gauss en factorisation ? II

Si on utilise bêtement une méthode de Gauss classique à chaque fois, on refait toujours les mêmes opérations sur la matrice A alors que c'est inutile ! On obtient donc un coût calcul d'environ :

$$C_{meth1} = m \times \left(C_{fact} + n^2 \right)$$

où $C_{fact} \simeq \frac{1}{3} n^3$ est le nombre d'opérations (multiplications et additions/soustraction) correspondant au travail sur A (identique à celui de la factorisation donc), la partie n^2 correspond à la somme du travail d'élimination sur le second membre et du travail pour résoudre le système triangulaire supérieur final.
Si on procède d'abord par une factorisation, puis par m "descentes-remontées" on obtient :

$$C_{meth2} = C_{fact} + m \times n^2$$

Pourquoi formaliser Gauss en factorisation ? III

Et :

$$\frac{C_{meth1}}{C_{meth2}} = \frac{m(C_{fract} + n^2)}{C_{fact} + mn^2} = \frac{m(C_{fract} + n^2)}{C_{fact} + n^2 + (m-1)n^2} = \frac{m}{1 + \frac{(m-1)n^2}{C_{fact}+n^2}}$$

et en utilisant l'approximation $C_{fact} \simeq \frac{1}{3}n^3$, on obtient :

$$\frac{C_{meth1}}{C_{meth2}} \simeq \frac{m}{1 + \frac{3(m-1)}{n}}$$

Ainsi pour $n >> m$ la méthode 2 tend à être presque m fois plus rapide que la méthode 1 et pour $m >> n$ la méthode 2 est environ $n/3$ plus rapide que la 1.

Le fait de pouvoir résoudre des systèmes linéaires assez rapidement (en n^2) une fois la factorisation calculée à d'autres avantages comme celui de pouvoir estimer $\|A^{-1}\|$ sans avoir à calculer explicitement A^{-1}.

1. Car il est simple de généraliser la méthode de Gauss classique pour qu'elle s'adapte à plusieurs seconds membres.