C2, Systèmes linéaires, méthodes directes

$A \in \mathcal{M}_{n,p}(\mathbb{R})$, $b \in \mathbb{R}^n$. On veut calculer $x \in \mathbb{R}^p$ t.q. $Ax = b$

1. $p = n$, A inversible. Méthodes directes (Gauss, LU, Choleski)
   Une méthode directe donne la solution après un nombre fini d'opérations
2. $p \neq n$ (mais aussi $p = n$).
   Echelonnement (généralisation de Gauss)

Méthode de Gauss, Etape 1

$A \in \mathcal{M}_n(\mathbb{R})$, $b \in \mathbb{R}^n$. On veut calculer $x \in \mathbb{R}^n$ t.q. $Ax = b$

$$\begin{bmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & \vdots & \vdots \\ a_{n,1} & \cdots & a_{n,n} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$$

Si $a_{1,1} \neq 0$, on pose $u_{1,j} = a_{1,j}$ pour tout j et $y_1 = b_1$

$u_{1,1}$ s'appelle "le premier pivot"

Pour $i > 1$, on remplace L_i par $L_i - \ell_{i,1} L_1$, $(L_1 = L_1(U))$

Pour $i > 1$, on remplace b_i par $b_i - \ell_{i,1} y_1$,

$\ell_{i,1} = a_{i,1}/u_{1,1}$

Le nouveau système (équivalent au premier) s'écrit

$$\begin{bmatrix} u_{1,1} & \cdot & \cdots & u_{1,n} \\ 0 & \tilde{a}_{2,2} & \cdots & \tilde{a}_{2,n} \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \cdot & \cdots & \cdot \end{bmatrix} \begin{bmatrix} x_1 \\ \cdot \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} y_1 \\ \tilde{b}_2 \\ \vdots \\ \cdot \end{bmatrix}$$

Méthode de Gauss, Etape 2

$$\begin{bmatrix} u_{1,1} & \cdot & \cdots & u_{1,n} \\ 0 & \tilde{a}_{2,2} & \cdots & \tilde{a}_{2,n} \\ . & \vdots & \vdots & \vdots \\ 0 & \cdot & \cdots & \cdot \end{bmatrix} \begin{bmatrix} x_1 \\ \cdot \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} y_1 \\ \tilde{b}_2 \\ \vdots \\ \cdot \end{bmatrix}$$

Si $\tilde{a}_{2,2} \neq 0$, on pose $u_{2,j} = \tilde{a}_{2,j}$ pour tout j et $y_2 = \tilde{b}_2$
$u_{2,2}$ s'appelle "le deuxième pivot"
Pour $i > 2$, on remplace $\tilde{L}_i$ par $\tilde{L}_i - \ell_{i,2}\tilde{L}_2$, $(\tilde{L}_2 = L_2(U))$
Pour $i > 2$, on remplace $\tilde{b}_i$ par $\tilde{b}_i - \ell_{i,2} y_2$,
$\ell_{i,2} = \tilde{a}_{i,2}/u_{2,2}$
Le nouveau système (équivalent au premier) s'écrit

$$\begin{bmatrix} u_{1,1} & \cdot & \cdot & \cdots & u_{1,n} \\ 0 & u_{2,2} & \cdot & \cdots & u_{2,n} \\ 0 & 0 & \tilde{a}_{3,3} & \cdots & \tilde{a}_{3,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdot & \cdots & \cdot \end{bmatrix} \begin{bmatrix} x_1 \\ \cdot \\ \cdot \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \tilde{b}_3 \\ \vdots \\ \cdot \end{bmatrix}$$

Méthode de Gauss, Etape k ($k \leq n$)

Si $\tilde{a}_{k,k} \neq 0$, on pose $u_{k,j} = \tilde{a}_{k,j}$ pour tout j et $y_k = \tilde{b}_k$
$u_{k,k}$ s'appelle "le k-ieme pivot"
Pour $i > k$, on remplace L_i par $\tilde{L}_i - \ell_{i,k}\tilde{L}_k$, $(\tilde{L}_k = L_k(U))$
Pour $i > k$, on remplace $\tilde{b}_i$ par $\tilde{b}_i - \ell_{i,k} y_k$,
$\ell_{i,k} = \tilde{a}_{i,k}/u_{k,k}$

Méthode de Gauss, système final

Après les n étapes le système devient $Ux = y$.
La matrice U est triangulaire supérieure avec $u_{1,1}, \ldots, u_{n,n}$ comme termes diagonaux.
La résolution du système est très facile :
$x_n = \frac{y_n}{u_{n,n}}$, $x_{n-1} = \frac{y_{n-1} - u_{n-1,n} x_n}{u_{n-1,n-1}}$ ...
Remarque importante : À chaque étape, le déterminant de la matrice et les déterminants des sous matrices "principales" ne changent pas (les sous matrices principales sont les matrices de $M_k(\mathbb{R})$ obtenues aves les k premières lignes et colonnes de A)
On en déduit

1. Si la méthode aboutit, $\det(A) = \prod_{i=1}^{n} u_{i,i} \neq 0$
2. La méthode aboutit si et seulement si les mineurs de A sont non nuls (les mineurs sont les déterminants des sous matrices "principales")

En pratique L et U sont stockés dans A, y et x dans b

Méthode LU

On note L la matrice triangulaire inférieure obtenue lors de la méthode de Gauss, en ajoutant $\ell_{i,i} = 1$ pour tout i.
Un calcul simple montre alors que $A = LU$ et $Ly = b$. Le système $Ax = b$ est équivalent à $Ux = y$, $Ly = b$.

Nombres d'opérations, conservation du profil

$A \in \mathcal{M}_n(\mathbb{R})$, $b \in \mathbb{R}^n$. On veut calculer $x \in \mathbb{R}^n$ t.q. $Ax = b$

Nombre d'opérations pour Gauss (et donc LU) :

$\frac{2}{3}n^3 + C_n$ (avec $C_n > 0$ borné par Cn^2)

Nombre d'opérations pour calculer y et x : $\tilde{C}_n > 0$ borné par Cn^2

Ceci explique l'intérêt de LU par rapport à Gauss si il y a beaucoup de systèmes à résoudre avec la même matrice

Question importante en pratique :

Comment minimiser les besoins en stockage ?

Conservation du profil :

Ligne i :

$a_{i,j} = 0$ pour $j < p(i) \leq i$ implique $\ell_{i,j} = 0$ pour $j < p(i) \leq i$

Colonne j :

$a_{i,j} = 0$ pour $i < p(j) \leq j$ implique $u_{i,j} = 0$ pour $i < p(j) \leq j$

Il peut être intéressant de choisir l'ordre des équations et des inconnues de manière à optimiser le stockage

Unicité de LU

$A = L_1 U_1 = L_2 U_2$
Les matrices L_i et U_i sont inversibles; et donc
$U_1 (U_2)^{-1} = (L_1)^{-1} L_2$
La matrice $U_1 (U_2)^{-1}$ est triangulaire supérieure, La matrice $(L_1)^{-1} L_2$ est triangulaire inférieure avec termes diagonaux égaux à 1, donc
$U_1 (U_2)^{-1} = (L_1)^{-1} L_2 = I$
et donc $U_1 = U_2$, $L_1 = L_2$

1. U triangulaire supérieure inversible. Donc $E_i = ev\{e_1, \ldots, e_i\}$ stable par U. Ceci prouve que E_i est aussi stable par U^{-1} et donc U^{-1} triangulaire supérieure.
2. L est inversible, $(L^{-1})^t = (L^t)^{-1}$ (car $(L^{-1})^t L^t x \cdot y = x \cdot L L^{-1} y = x \cdot y$). L triangulaire inférieure implique L^{-1} triangulaire inférieure.

Gauss avec “recherche de pivot”

Question :
Que faire si $a_{1,1} = 0$ (à l’étape 1) ou si $\tilde{a}_{k,k} = 0$ à l’étape k ?
On suppose $\det(A) \neq 0$
A l’étape 1, si $a_{1,1} = 0$, on échange la ligne 1 avec une ligne m pour laquelle $a_{m,1} \neq 0$ (possible car $\det(A) \neq 0$). On échange aussi b_1 et b_m
A l’étape k, si $\tilde{a}_{k,k} = 0$, on échange la ligne k avec une ligne $m > k$ pour laquelle $a_{m,k} \neq 0$ (possible car $\det(A) \neq 0$). On échange aussi $\tilde{b}_i$ et $\tilde{b}_m$
On trouve ainsi une matrice de permutation P et des matrices L et U tels que $PA = LU$ (et le système est devenu $LUx = Pb$, car $Ly = Pb$, $Ux = y$)

En fait pour toute matrice A, il existe une matrice de permutation, P, et des matrices L et U avec L triangulaire inférieure, avec des termes diagonaux égaux à 1, et U triangulaire supérieure tels que $PA = LU$ (mais il n’y a pas toujours unicité). Ceci faisait partie de l’examen de mai 2020

$\det(A) = 0$ si et seulement si $\ker(A) \neq \{0\}$

Si $\ker(A) \neq \{0\}$, $\mathrm{rang(A) < n}$. Les vecteurs colonnes de A sont liés et donc $\det(A) = 0$ (cette implication a été vue au 1er cours).
On montre maintenant que $\det(A) = 0$ implique $\ker(A) \neq \{0\}$
On suppose que $\det(A) = 0$. Cela signifie que nécessairement, lors de l'une des étapes de Gauss avec pivot il est impossible trouver un pivot.

1. Si c'est à l'étape 1, $Ae_1 = 0$,
2. Si c'est à l'étape 2, il existe $\alpha \in \mathbb{R}$ tel que $A(e_2 - \alpha e_1) = 0$
3. Si c'est à l'étape i, $1 \leq 1 \leq n$, il existe $\alpha_1, \ldots, \alpha_{i-1}$ tel que $A(e_i - \sum_{j=1}^{i-1} \alpha_j e_j) = 0$

Ceci donne bien $\ker(A) \neq \{0\}$.

Méthode de Choleski, existence par LU

$A \in \mathcal{M}_n(\mathbb{R})$, A s.d.p.. Alors, il existe une unique matrice L triangulaire inférieure avec $\ell_{i,i} > 0$ pour tout i telle que $A = LL^t$

Existence par LU:

Les matrices principales de A sont s.d.p., donc les mineurs sont strictement positifs et A admet une décomposition LU. On note D la partie diagonale de U

$A = LU = LD\tilde{U}$, L et $\tilde{U}$ ont des 1 en termes diagonaux

$LD\tilde{U} = (\tilde{U})^t DL^t$, $D\tilde{U}(L^t)^{-1} = L^{-1}(\tilde{U})^t D$

$D\tilde{U}(L^t)^{-1}$ est triang. sup., $L^{-1}(\tilde{U})^t D$ est triang. inf.

Donc $\tilde{U}(L^t)^{-1} = I$ $\tilde{U} = L^t$, $A = L\sqrt{D}\sqrt{D}L^t$

$\sqrt{D}$ matrice diagonale a termes diagonaux strictement positifs

Méthode de Choleski, existence par récurrence sur n

$A \in \mathcal{M}_n(\mathbb{R})$, A s.d.p.. On veut montrer qu'il existe une matrice L triangulaire inférieure avec $\ell_{i,i} > 0$ pour tout i t.q. $A = LL^t$

Initialisation : $n = 1$. $\ell_{1,1} = \sqrt{a_{1,1}}$

Itérations : On suppose le résultat vrai si $A \in \mathcal{M}_k(\mathbb{R})$, $k < n$

$A \in \mathcal{M}_n(\mathbb{R})$, A s.d.p.. $A = \begin{bmatrix} B & a \\ a^t & \alpha \end{bmatrix}$, $B \in \mathcal{M}_{n-1}(\mathbb{R})$,

$a \in \mathcal{M}_{n-1,1}(\mathbb{R}) = \mathbb{R}^{n-1}$, $\alpha \in \mathbb{R}$.

B est s.d.p. donc $B = MM^t$. On cherche $L = \begin{bmatrix} M & 0 \\ b^t & \beta \end{bmatrix}$

$$LL^t = \begin{bmatrix} M & 0 \\ b^t & \beta \end{bmatrix} \begin{bmatrix} M^t & b \\ 0 & \beta \end{bmatrix} = \begin{bmatrix} MM^t & Mb \\ b^t M^t & b^t b + \beta^2 \end{bmatrix} = \begin{bmatrix} B & Mb \\ b^t M^t & b^t b + \beta^2 \end{bmatrix}$$

$b = M^{-1} a$, $\beta = \sqrt{\alpha - b^t b}$

possible si $\alpha - b^t b = \alpha - a^t (M^t)^{-1} M^{-1} a = \alpha - a^t B^{-1} a > 0$

Méthode de Choleski, existence, fin

$A \in \mathcal{M}_n(\mathbb{R})$, A s.d.p.. $A = \begin{bmatrix} B & a \\ a^t & \alpha \end{bmatrix}$, $B \in \mathcal{M}_{n-1}(\mathbb{R})$,
$a \in \mathcal{M}_{n-1,1}(\mathbb{R})$, $\alpha \in \mathbb{R}$
Il reste à montrer que $\alpha - a^t B^{-1} a > 0$

$z \in \mathbb{R}^{n-1}$,

$$\begin{bmatrix} B & a \\ a^t & \alpha \end{bmatrix} \begin{bmatrix} z \\ 1 \end{bmatrix} \cdot \begin{bmatrix} z \\ 1 \end{bmatrix} = \begin{bmatrix} Bz + a \\ a^t z + \alpha \end{bmatrix} \cdot \begin{bmatrix} z \\ 1 \end{bmatrix} = Bz \cdot z + 2a^t z + \alpha > 0$$

On choisit $z = -B^{-1}a$, ceci donne $\alpha - a^t B^{-1} a > 0$

Finalement $L = \begin{bmatrix} M & 0 \\ b^t & \beta \end{bmatrix}$, $b = M^{-1}a$, $\beta = \sqrt{\alpha - b^t b}$
$A = LL^t$, L triangulaire inférieure avec $\ell_{i,i} > 0$ pour tout i

Méthode de Choleski, unicité et calcul

$A \in \mathcal{M}_n(\mathbb{R})$, A s.d.p.. On cherche L triangulaire inférieure avec $\ell_{i,i} > 0$ pour tout i t.q. $A = LL^t$

$a_{i,j} = \sum_{k=1}^{n} \ell_{i,k}\ell_{j,k} = \sum_{k=1}^{\min\{i,j\}} \ell_{i,k}\ell_{j,k}$

Calcul $C_1(L)$: $\ell_{1,1} = \sqrt{a_{1,1}}$, $\ell_{i,1} = \frac{a_{i,1}}{\ell_{1,1}}$ $(i > 1)$ n opérations

$C_1(L) \ldots C_{q-1}$ connues

Calcul $C_q(L)$: $\ell_{q,q} = \sqrt{a_{q,q} - \sum_{k=1}^{q-1} \ell_{q,k}^2}$ $2q-1$ opérations

$\ell_{i,q} = \frac{a_{i,q} - \sum_{k=1}^{q-1} \ell_{i,k}\ell_{q,k}}{\ell_{q,q}}$ $(i > q)$ $(n-q)(2q-1)$ opérations

Nombre total d'opérations :
$\sum_{q=1}^{n}(n-q+1)(2q-1) = \frac{n^3}{3} + \frac{n^2}{2} + \frac{n}{6}$

Rappel pour LU : $\frac{2}{3}n^3 + \ldots$

Méthode de Choleski, résumé

$A \in \mathcal{M}_n(\mathbb{R})$, A s.d.p., $b \in \mathbb{R}^n$. On cherche x t.q. $Ax = b$

1. Calcul de L triangulaire inférieure avec $\ell_{i,i} > 0$ et $A = LL^t$
2. Calcul de $y \in \mathbb{R}^n$ t.q. $Ly = b$ Descente
3. Calcul de $x \in \mathbb{R}^n$ t.q. $L^t x = y$ Remontée

Conservation du profil
(pour minimiser le nombre de calcul et le stockage) :
Ligne i :
$a_{i,j} = 0$ pour $j < p(i) \leq i$ implique $\ell_{i,j} = 0$ pour $j < p(i) \leq i$

Echelonnement (généralisation de Gauss)

$A \in \mathcal{M}_{n,p}(\mathbb{R})$, $A \neq 0$, $b \in \mathbb{R}^n$. On cherche $x \in \mathbb{R}^p$ t.q. $Ax = b$

Principe de la méthode :

On transforme le système $Ax = b$ en un système équivalent $Ex = y$ (avec $E \in \mathcal{M}_{n,p}(\mathbb{R})$, $y \in \mathbb{R}^n$)

La matrice E possède les propriétés suivantes :

1. Si la ligne i de E a des termes non nuls, on note $p(i)$ l'indice de colonne du premier terme non nul (de sorte que $e_{i,j} = 0$ si $j < i$). Sinon on pose $p(i) = 0$
2. Pour tout $1 \leq i \leq n$, si $p(i) > 0$, $e_{i,p(i)} = 1$
3. Pour tout $1 < i \leq n$, si $p(i) \neq 0$, alors $1 \leq p(i-1) < p(i)$
   En reprenant la méthode de Gauss, on voit qu'il est possible de construire E et y

Non existence de solution

$A \in \mathcal{M}_{n,p}(\mathbb{R})$, $A \neq 0$, $b \in \mathbb{R}^n$

Après échelonnement le système $Ax = b$ est sous la forme $Ex = y$

On note $j_m = \max\{p(i), 1 \leq i \leq n\}$ et i_m tel que $p(i_m) = j_m$

Non existence de solution :
Le système $Ax = b$ n'a pas de solution si il existe $i > i_m$ tel que $y_i \neq 0$

On va montrer ensuite que $Ax = b$ a au moins une solution si $y_i = 0$ pour tout $i > i_m$

Calcul des solutions

$A \in \mathcal{M}_{n,p}(\mathbb{R})$, $A \neq 0$, $b \in \mathbb{R}^n$

Après échelonnement le système $Ax = b$ est sous la forme $Ex = y$

On suppose que $y_i = 0$ si $i > i_m$

Le système $Ax = b$ a alors au moins une solution

1. La valeur de x_j est arbitraire si $j \notin \mathrm{Im(p)}$
2. Le calcul de $x_{p(i)}$ se fait avec la ligne i du système $Ex = y$ (en prenant les valeurs de i dans l'ordre décroissant)

Calcul de la dimension du noyau de A

$A \in \mathcal{M}_{n,p}(\mathbb{R})$, $A \neq 0$

Exercice :

1. Comment obtenir la dimension du noyau de A ?
2. Comment obtenir une base du noyau de A ?