

Domaine : Sciences, Technologie et Santé

Master Mention MABS

**MICROBIOLOGIE-AGROBIOSCIENCES-BBIOINFORMATIQUE ET
BIOLOGIE DES SYSTEMES**

Spécialité : Bioinformatique et Biologie des Systèmes

Mémento

Cours Mathématiques pour la Biologie

Septembre 2011

Enseignants (2011-2012) :

Régine André-Obrecht (Professeur)

Roland Barriot (Maître de Conférences)



Table des matières

I.	Eléments d’algèbre linéaire	3
I.1	Matrices	3
I.1.a	Rappels.....	3
I.1.b	Matrices orthogonales.....	4
I.1.c	Vecteurs et valeurs propres	4
I.1.d	Diagonalisation.....	6
	APPLICATION :.....	6
1-	Calcul de la puissance d’une matrice.	6
2-	Résolution des relations de récurrence linéaires homogènes d’ordre k.....	6
3-	Analyse en Composantes Principales (TP).....	7
I.2	Systèmes linéaires	11
I.2.a	Quelques résultats théoriques.....	11
I.2.b	Le cas simple des matrices triangulaires	12
I.2.c	Des méthodes exactes	13
I.2.d	Des Méthodes itératives	17
II.	Notions d’analyse.....	21
II.1	Calcul différentiel.....	21
II.1.a	Fonction dérivable et différentiable	21
II.1.b	Recherche du minimum d’une fonction ; algorithme de descente du gradient.....	21
II.2	Recherche du zéro d’une fonction.....	21
II.3	Equations différentiables linéaires	21
III.	Notions de probabilités	21
III.1	Conditionnement et vraisemblance	21
III.2	Régression linéaire – Modèles linéaires	21

I. Éléments d'algèbre linéaire

I.1 Matrices

Soit A une matrice d'ordre $m \times n$, à coefficients réels ou complexes $\{a_{i,j}, i=1, \dots, m, j=1, \dots, n\}$, à m lignes et n colonnes:

$$A = \begin{pmatrix} a_{1,1} & \dots & a_{1,j} & \dots & a_{1,n} \\ \dots & & & & \dots \\ a_{i,1} & & a_{i,j} & \dots & a_{i,n} \\ \dots & & & & \dots \\ a_{m,1} & \dots & a_{m,j} & \dots & a_{m,n} \end{pmatrix}$$

I.1.a Rappels

Opérations sur les matrices : A et B matrices $m \times n$, D matrice $n \times p$

- Somme $S = A + B$: $s_{i,j} = a_{i,j} + b_{i,j}$ S est une matrice $m \times n$.
- Produit par un scalaire $C = \alpha \cdot A$: $c_{i,j} = \alpha a_{i,j}$ C est une matrice $m \times n$.
- Produit de deux matrices $P = A \cdot D$: $p_{i,j} = \sum_{k=1}^n a_{i,k} d_{k,j}$ P est une matrice $m \times p$.
- Transposée d'une matrice A^T : $(A^T)_{i,j} = A_{j,i}$

$$\text{Propriétés : } (A^T)^T = A, \quad (A + B)^T = A^T + B^T, \quad (\alpha A)^T = \alpha A^T, \quad (A \cdot B)^T = B^T A^T,$$

- A est dite triangulaire supérieure (resp. inférieure) si tous les termes au dessus de la diagonale (resp. en dessous) sont nuls. A est dite diagonales si seuls les termes de la diagonale sont non nuls.

Cas des matrices carrées $m = n$:

- Trace d'une matrice

$$\text{tr}(A) = \sum_{k=1}^n a_{k,k}, \quad \text{tr}(\alpha A) = \alpha \text{tr}(A), \quad \text{tr}(A + B) = \text{tr}(A) + \text{tr}(B) \quad \text{tr}(A \cdot B) = \text{tr}(B \cdot A)$$

- Déterminant noté $|A|$ ou $\det(A)$

$$\det(A) = \sum_{i=1}^n (-1)^{i+j} a_{i,j} \det(A_{i,j}) \quad \text{où } A_{i,j} \text{ est la matrice } A \text{ à laquelle on a ôté la } i^{\text{ème}} \text{ ligne et la } j^{\text{ème}} \text{ colonne.}$$

Exercice 1 : calculer le déterminant de la matrice $A = \frac{1}{6} \begin{pmatrix} 1 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 1 \end{pmatrix}$.

Si A est triangulaire ou diagonale, $|A| = \prod_{i=1}^n a_{i,i}$

- Inverse d'une matrice : si A est inversible, l'inverse A^{-1} vérifie $A^{-1}A = AA^{-1} = I$

Théorème : A est inversible si et seulement si $|A|$ est non nul

Propriétés : $(A^{-1})^T = (A^T)^{-1}$, $(A.B)^{-1} = B^{-1}.A^{-1}$, $|A^{-1}| = \frac{1}{|A|}$

Définition et propriété : On appelle matrice extraite de A, toute matrice \tilde{A} obtenue en supprimant k lignes et k colonnes de A ; \tilde{A} est d'ordre $(n-k) \times (n-k)$. On appelle rang d'une matrice, noté $rg(A)$, la plus grande valeur de $(n-k)$ telle que \tilde{A} soit inversible.

Si A est inversible, $rg(A) = n$.

I.1.b Matrices orthogonales

Définitions : Une matrice carrée A est dite :

- Symétrique si $A^T = A$
- Singulière si $|A| = 0$, régulière sinon.
- Idempotente si $A.A = A$
- Orthogonale si $A^T A = AA^T = I$ ou $A^{-1} = A$

Exercice 2 : quelles sont les propriétés de la matrice $A = \frac{1}{6} \begin{pmatrix} 1 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 1 \end{pmatrix}$?

I.1.c Vecteurs et valeurs propres

Définitions : On appelle vecteur propre d'une matrice carrée $n \times n$, A, tout vecteur x de R^n , non nul, pour lequel il existe un réel λ vérifiant :

$$Ax = \lambda x$$

Le réel λ est alors appelé valeur propre de A, associée à x .

Propriétés : Soit A une matrice carrée réelle $n \times n$,

- Les valeurs propres sont solutions de l'équation $\det(A - \lambda I) = 0$. Cette expression est un polynôme de degré n , on dit que λ est une valeur propre de multiplicité algébrique k , si λ est une racine d'ordre k de ce polynôme, appelé polynôme caractéristique associé à A.

- Le nombre de valeurs propres est inférieur ou égal à n (degré du polynôme) et des valeurs propres peuvent être nulles.
- La somme des valeurs propres est égale à la somme des éléments diagonaux de la matrice ($tr(A)$),
- Le produit des valeurs propres est égal au déterminant de la matrice $|A|$,
- Le rang de A est égal au nombre de valeurs propres non nulles.
- Les valeurs propres de A sont les carrés des valeurs propres de A et ses deux matrices ont les mêmes vecteurs propres.
- Si A est **inversible**, toutes les valeurs propres sont non nulles et les valeurs propres de A^{-1} sont les inverses des valeurs propres de A et les vecteurs propres sont les mêmes.
- Si A est une matrice **idempotente**, ses valeurs propres sont égales à 0 ou 1, et sa trace est égale à son rang.
- Si A est une matrice symétrique,
 - o Les vecteurs propres associés à des valeurs propres différentes sont orthogonaux,
 - o Si une valeur propre est de multiplicité k, il existe k vecteurs orthogonaux associés à cette même valeur,
 - o La matrice A possède n valeurs propres $\lambda_1, \lambda_2, \dots, \lambda_n$, pas forcément différentes, et il existe un ensemble de n vecteurs propres orthogonaux x_1, x_2, \dots, x_n tels que :

$$x_i^T x_j = 0, \quad i \neq j; \quad i, j = 1, \dots, n$$

Exercice 3 : Rechercher les vecteurs propres et valeurs propres pour les matrices suivantes, après avoir étudié l'inversibilité de la matrice :

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

Exercice 4 : Cas particulier des matrices dites de covariance : Soit $X = \begin{pmatrix} -1 & 2 & 2 \\ -1 & 4 & 4 \\ 1 & 4 & 2 \\ 1 & 2 & 4 \end{pmatrix}$, une matrice dite

des individus dans R^3 , chaque ligne représente un individu d'une population de 4 individus (X_1, X_2, X_3, X_4) .

Calculer G le centre de gravité de cette population. Donner la matrice centrée \underline{X} associée à X et calculer $\frac{1}{4} \underline{X}^T \underline{X}$; cette matrice est la matrice de covariance associée à la population initiale. Quelles sont les propriétés de cette matrice ? (dimension, rang, symétrie...). Généraliser pour une matrice X de dimension p x n, de p individus de R^n .

I.1.d Diagonalisation

Définition : Deux matrices d'ordre $n \times n$ A et B sont dites semblables s'il existe une matrice P carrée d'ordre n telle que

$$B = P^{-1}AP.$$

On dit que A est diagonalisable si A est semblable à une matrice diagonale, c'est-à-dire s'il existe P et une suite de n valeurs réelles (d_1, d_2, \dots, d_n) telle que :

$$P^{-1}AP = \text{diag} (d_1, d_2, \dots, d_n)$$

Théorème : Si A est symétrique, soit X la matrice formée par un ensemble de n vecteurs propres unitaires (de norme 1) orthogonaux x_1, x_2, \dots, x_n associés aux valeurs propres $\lambda_1, \lambda_2, \dots, \lambda_n$, la matrice X est orthogonale et A est diagonalisable avec :

$$X^T AX = \text{diag} (\lambda_1, \lambda_2, \dots, \lambda_n)$$

APPLICATION :

1- Calcul de la puissance d'une matrice.

Propriété (exercice 5) : Si A est diagonalisable, A^m est diagonalisable et si P est la matrice de passage, on a :

$$P^{-1}AP = \text{diag} (d_1, d_2, \dots, d_n)$$

$$P^{-1}A^mP = \text{diag} (d_1^m, d_2^m, \dots, d_n^m)$$

Application à la résolution des systèmes de récurrence de la forme $X^{m+1} = AX^m$.

Exercice 6 : Un modèle naïf de la reproduction de l'aphès des peupliers.

(l'unité de temps est la durée d'une génération \rightarrow indice n dans ce qui suit)

Les femelles adultes déposent des galles sur les feuilles de l'arbre. Tous les descendants d'un aphès sont contenues dans une galle et une fraction d'entre eux survivent jusqu'à l'âge adulte. Il est fait l'hypothèse très naïve que les taux de reproduction et de survie sont constants et indépendants de l'environnement. Soit r le taux de reproduction, f le nombre de descendants par femelle et m le taux de mortalité. A chaque génération n , a_n désigne le nombre d'insectes femelles adultes, et ρ_n le nombre de descendants à la génération n .

Donner l'expression de a_n en fonction des données initiales et de n .

On suppose que le taux de mortalité est de 80%, que la proportion de femelles est 50%. Quel doit être le taux de fécondité minimum pour éviter l'extinction de cette population ?

2- Résolution des relations de récurrence linéaires homogènes d'ordre k

(Polynôme caractéristique associée, matrice associée)

3- Analyse en Composantes Principales (TP)

Dans quel espace d'observations doit on travailler ?

- Recherche de l'espace discriminant lors d'une phase d'apprentissage à partir d'un ensemble de réalisations ou d'individus de \mathbb{R}^d , toutes classes confondues.
- Analyse des données

Soit Y l'ensemble des individus, $Y = \{y_1, y_2, \dots, y_N\}$, $y_i \in \mathbb{R}^d$. (nuage de réalisations, nuage d'observations)

Matrice des individus \mathfrak{R} , la matrice $N \times d$ suivante :

$$\begin{pmatrix} y_{1,1} & y_{1,2} & \dots & y_{1,d} \\ y_{2,1} & y_{2,2} & \dots & y_{2,d} \\ \dots & \dots & \dots & \dots \\ y_{N,1} & y_{N,2} & \dots & y_{N,d} \end{pmatrix}$$

Objectif : réduire la dimension de l'espace des observations ou projeter l'espace des observations sur un sous espace de dimension $p < d$, en perdant un minimum d'information.

Soit W_p , un sous espace de dimension p de \mathbb{R}^d . W_p est caractérisé par un point A de \mathbb{R}^d , et p vecteurs directeurs $\langle u_1, u_2, \dots, u_p \rangle$ qui représentent une base du sous espace vectoriel associé.

Soit \bar{y}_n la projection de chaque individu y_n sur W_p .

Définition : On appelle moyenne des carrés des distances des individus à leur projection sur W_p le critère suivant :

$$I_{W_p} = \frac{1}{N} \sum_{n=1}^N \|y_n - \bar{y}_n\|^2$$

Reformulation : trouver le sous espace W_p qui minimise le critère I_{W_p}

Proposition : Pour p fixé, le sous espace W_p qui minimise le critère I_{W_p} contient nécessairement le centre de gravité du nuage Y .

Désormais on recherche W_p tel que A soit le centre de gravité du nuage.

Par commodité, il est fait un changement d'origine tel que A soit l'origine de l'espace \mathbb{R}^d et du nuage de points Y . Cela revient à « centrer les données » et remplacer chaque y_n par sa valeur diminuée de la moyenne de Y , ce qui est désormais supposé.

Sous l'hypothèse $O = G$:

Proposition : Minimiser le critère J_{W_p} revient à maximiser le critère d'inertie des projections, à savoir :

$$J_{W_p} = \frac{1}{N} \sum_{n=1}^N \|\bar{y}_n\|^2$$

démonstration : théorème de Pythagore

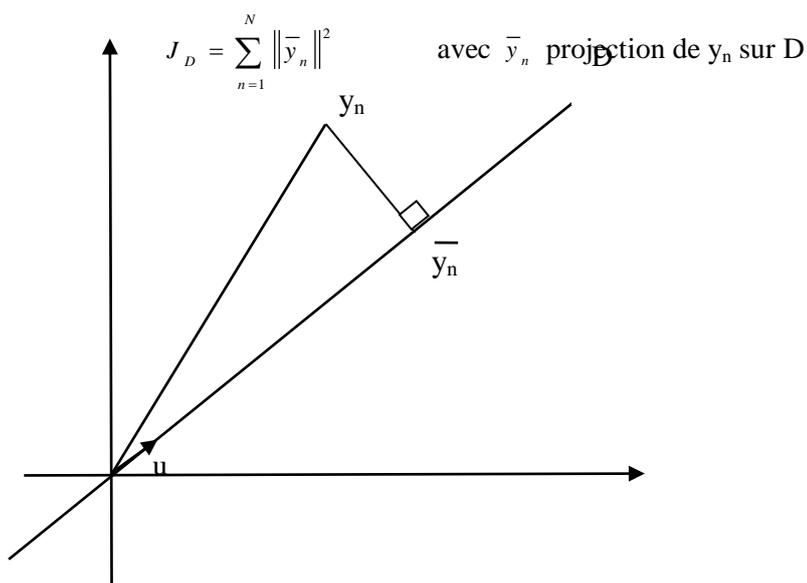
$$\max_W \sum_{n=1}^N \|\bar{y}_n\|^2 = \min_W \sum_{n=1}^N \|y_n - \bar{y}_n\|^2$$

Les solutions W_p sont emboîtées, on recherche donc W_1 , et par récurrence on en déduira ensuite W_p , pour tout p , à partir de W_{p-1} d'où les deux paragraphes suivants

I Projection sur une droite passant par l'origine

Objectif : réduire la dimension de l'espace des observations ou projeter l'espace des observations sur un sous espace de dimension $1 < d$, en perdant un minimum d'information.

Reformulation : trouver la droite $D(O,u)$ passant par l'origine de direction u , qui maximise la dispersion du nuage des points projetés, c'est-à-dire



$$\max_{u_1} \sum_{n=1}^N \|\bar{y}_n\|^2 = \min_{u_1} \sum_{n=1}^N \|y_n - \bar{y}_n\|^2$$

Théorème : le vecteur U recherché est le vecteur propre associé à la plus grande valeur propre de la matrice $\mathfrak{R}'\mathfrak{R}$.

Définition : la matrice $\frac{1}{N}\mathfrak{R}'\mathfrak{R}$ est la matrice de covariance des individus.

II Généralisation : l'Analyse en Composantes principales

Reprise de la formulation : Minimiser le critère J_{W_p} revient à maximiser le critère d'inertie des projections, à savoir :

$$J_{W_p} = \frac{1}{N} \sum_{n=1}^N \|\bar{y}_n\|^2$$

Ce critère est celui que l'on a cherché à maximiser pour trouver la meilleure droite de projection passant par l'origine. On sait donc trouver W_1 . On trouve les autres sous espaces par itération en utilisant le résultat suivant :

Théorème (des solutions emboîtées) : Soit W_k le sous espace de dimension k maximisant l'inertie des projections pour cette dimension, alors W_{k+1} , le sous espace de dimension (k+1) maximisant l'inertie des projections est la somme directe (orthogonale) de W_k et du sous espace de dimension 1 orthogonale à W_k et maximisant l'inertie des projections sur une droite dans le sous espace orthogonal à W_k .

Définition 1 : On appelle **premier axe principal des individus Y**, le vecteur $u_1 \in R^d$, de norme 1 (au signe près !), vecteur directeur de la droite maximisant le critère d'inertie des projections du nuage.

Pour chaque individu y de R^d , on appelle première composante principale de y, la projection de y sur cette droite.

Détermination du premier axe principal : le vecteur u_1 est « un » des vecteurs propres associés à la plus grande valeur propre λ_1 de la matrice de covariance des individus Cov,

$$Cov = \frac{1}{N} \sum_{n=1}^N y_n y_n'$$

Définition 2 : On appelle **deuxième (ième) axe principal des individus Y**, le vecteur $u_2 \in R^d$ ($u_p \in R^d$), de norme 1, vecteur orthogonal à u_1 (W_{p-1}) et vecteur directeur de la droite maximisant le critère d'inertie des projections du nuage.

Pour chaque individu y de R^d , on appelle deuxième (ième) composante principale de y, la projection de y sur cette droite.

Détermination du deuxième axe principal : le vecteur u_2 est un des vecteurs propres associés à la plus grande valeur propre $\lambda_2 \leq \lambda_1$ de la matrice de covariance des individus Cov, orthogonal à u_1

Détermination des autres axes principaux : les axes principaux successifs sont les vecteurs propres de la matrice Cov rangés suivant l'ordre des valeurs propres décroissantes.

Remarques :

- La matrice Cov est en général symétrique définie positive, elle est diagonalisable sur une base orthonormée de vecteurs propres avec des valeurs propres positives. La décomposition en axes principaux équivaut à chercher une base de vecteurs propres et à ordonner cette base selon la valeur décroissante des valeurs propres.
- La valeur propre rapportée à la somme des valeurs propres exprime la partie de l'inertie portée par l'axe principal correspondant.

- En pratique, le rapport $\frac{\sum_{i=1}^k \lambda_i}{\sum_{i=1}^d \lambda_i}$ sert à choisir la dimension de l'espace de projections retenue ; on

choisit k de telle sorte que le rapport soit suffisamment proche de 1 (0,95) pour estimer que le nuage de points Y est suffisamment bien représenté par l'ensemble de ses projections sur le sous espace correspondant.

Exercice 7 Trouver les axes principaux correspondant à une Analyse en Composantes Principales du nuage de points suivants :

$$y_1 = \begin{pmatrix} 2 \\ 2 \end{pmatrix}, y_2 = \begin{pmatrix} 4 \\ 2 \end{pmatrix}, y_3 = \begin{pmatrix} 6 \\ 4 \end{pmatrix}, y_4 = \begin{pmatrix} 8 \\ 4 \end{pmatrix},$$

Donner la droite de régression associée à ces points dans le repère canonique.

I.2 Systèmes linéaires

On considère le système linéaire de n inconnues à m équations :

$$\left\{ \begin{array}{l} a_{1,1}x_1 + a_{1,2}x_2 + \dots + a_{1,n}x_n = b_1 \quad (L_1) \\ \dots \\ a_{i,1}x_1 + a_{i,2}x_2 + \dots + a_{i,n}x_n = b_i \quad (L_i) \\ \dots \\ a_{m,1}x_1 + a_{m,2}x_2 + \dots + a_{m,n}x_n = b_m \quad (L_m) \end{array} \right. \quad (S)$$

Sous forme compacte $\sum_{j=1}^n a_{i,j}x_j = b_i, i = 1, \dots, m$

Sous forme matricielle $Ax = b$, avec $A = \begin{pmatrix} a_{1,1} & \dots & a_{1,j} & \dots & a_{1,n} \\ \dots & & & & \dots \\ a_{i,1} & & a_{i,j} & \dots & a_{i,n} \\ \dots & & & & \dots \\ a_{m,1} & \dots & a_{m,j} & \dots & a_{m,n} \end{pmatrix}, x = \begin{pmatrix} x_1 \\ \dots \\ x_i \\ \dots \\ x_n \end{pmatrix}, b = \begin{pmatrix} b_1 \\ \dots \\ b_i \\ \dots \\ b_m \end{pmatrix}$. Le

vecteur x est le vecteur inconnu à n coordonnées.

I.2.a Quelques résultats théoriques

Théorème : On ne modifie pas les solutions d'un système linéaire si on ajoute à une ligne une combinaison linéaire des autres lignes (second membre compris !!)

$$L_i \leftarrow L_i + \sum_{j \neq i} \alpha_j L_j$$

Théorème : Le système (S) admet une solution unique si et seulement si $n=m$ et $\det(A) \neq 0$. Dans ce cas, la matrice A est inversible et la solution est donnée par $x = A^{-1}b$.

Théorème : (Les formules de Cramer) Lorsque $\det(A) \neq 0$, la solution est donnée pour tout $j=1, \dots, n$ par

$$x_j = \frac{1}{\det(A)} \det \begin{pmatrix} a_{1,1} & \dots & a_{1,j-1} & b_1 & a_{1,j+1} & \dots & a_{1,n} \\ a_{2,1} & & a_{2,j-1} & b_2 & a_{2,j+1} & & a_{2,n} \\ \dots & & \dots & \dots & \dots & & \dots \\ a_{n,1} & & a_{n,j-1} & b_n & a_{n,j+1} & & a_{n,n} \end{pmatrix}$$

Remarque : Ces formules sont inutilisables dès lors que $n > 10$!!

Intérêt des systèmes n=m et des systèmes de Cramer

Nombre de systèmes sont sur contraints : trop d'équations pour le nombre d'inconnues « m > n », comme dans le cas des problèmes d'optimisation (régression : recherche d'un sous espace de faible dimension approchant au mieux un grand nombre de points)

Mais

$$Ax = b \text{ (S)} \Rightarrow A^T Ax = A^T b \text{ (S1)}$$

La matrice $A^T A$ est carrée de dimension $n \times n$, une solution de (S) est solution de (S1) et de plus :

Théorème : $\det(A^T A) \neq 0$ si et seulement si $\text{rang}(A) = \min(m, n) = n$.

D'où la recherche de la solution de (S1) !

I.2.b Le cas simple des matrices triangulaires

$$\begin{cases} l_{1,1}x_1 & = b_1 \\ l_{2,1}x_1 + l_{2,2}x_2 & = b_2 \\ \dots & \\ l_{i,1}x_1 + \dots + l_{i,i}x_i & = b_i \\ \dots & \\ l_{n,1}x_1 + l_{n,2}x_2 + \dots + l_{n,n}x_n & = b_n \end{cases}$$

La matrice du système est triangulaire (ici inférieure, les termes supérieurs sont nuls). Il existe une unique solution si et seulement si $l_{i,i} \neq 0, \forall i = 1, n$. Par substitution successive, on montre que la

solution est donnée par : $x_1 = \frac{b_1}{l_{1,1}}, x_i = \frac{b_i - \sum_{j=1}^{i-1} l_{i,j}x_j}{l_{i,i}}, \forall i = 2, \dots, n,$

Propriété : Le nombre d'opérations est de l'ordre de n^2 .

Trouver un compromis entre ce nombre d'opérations « n^2 » et celui trop élevé impliqué par les formules de Cramer ($(n+1)!$), tant que n reste inférieur à 100 !

- Méthode du pivot de Gauss
- Méthode de décomposition de la matrice A en un produit de matrices triangulaires

Exercice 8 : Résolution dans le cas d'une matrice triangulaire

Résoudre le système $Ax = b$, avec $A = \begin{pmatrix} 2 & 1 & 0 & 4 \\ 0 & -1 & -2 & 0 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$ et $b = \begin{pmatrix} 2 \\ -2 \\ -5 \\ -2 \end{pmatrix}$

I.2.c Des méthodes exactes

a) Méthode de Gauss (dite du pivot de Gauss)

Exemple : Les systèmes suivants sont équivalents, obtenus par substitution et élimination à chaque étape d'une inconnue, appelée le pivot.

$$\begin{cases} 2x_1 + x_2 + 4x_4 = 2 \\ -4x_1 - 2x_2 + 3x_3 - 7x_4 = -9 \\ 4x_1 + x_2 - 2x_3 + 8x_4 = 2 \\ -3x_2 - 12x_3 - x_4 = 2 \end{cases}$$

Élimination de x_1 dans les trois dernières lignes :

$$\begin{cases} 2x_1 + x_2 + 4x_4 = 2 \\ 3x_3 + x_4 = -5 \\ -x_2 - 2x_3 = -2 \\ -3x_2 - 12x_3 - x_4 = 2 \end{cases}$$

Comme on ne peut utiliser l'inconnue x_2 dans la 2^e équation, on inverse les lignes 2 et 3

$$\begin{cases} 2x_1 + x_2 + 4x_4 = 2 \\ -x_2 - 2x_3 = -2 \\ 3x_3 + x_4 = -5 \\ -3x_2 - 12x_3 - x_4 = 2 \end{cases}$$

Élimination de x_2 dans les deux dernières lignes :

$$\begin{cases} 2x_1 + x_2 + 4x_4 = 2 \\ -x_2 - 2x_3 = -2 \\ 3x_3 + x_4 = -5 \\ -6x_3 - x_4 = 8 \end{cases}$$

Élimination de x_3 dans la dernière ligne :

$$\begin{cases} 2x_1 + x_2 + 4x_4 = 2 \\ -x_2 - 2x_3 = -2 \\ 3x_3 + x_4 = -5 \\ x_4 = -2 \end{cases}$$

Le système est un système triangulaire que l'on résout comme précédemment : $x = \begin{pmatrix} 3 \\ 4 \\ -1 \\ -2 \end{pmatrix}$

Généralisation : La méthode consiste à trouver un système triangulaire équivalent par élimination successive des inconnues et combinaison linéaire des lignes du système. A chaque étape, une inconnue est éliminée, elle s'appelle le pivot (après inversion de lignes si nécessaire). Pour éviter des erreurs numériques, il faut chercher à éliminer l'inconnue de plus fort coefficient en valeur absolue.

Exercice 9 Donner par la méthode du pivot les solutions des deux systèmes suivants

$$\begin{cases} 2x_1 - x_2 - 2x_3 = 1 \\ 4x_1 - x_2 - x_3 = 4 \\ 6x_1 - 5x_2 - 13x_3 = -2 \\ 8x_1 - 7x_2 - 15x_3 = 0 \end{cases}$$

Exercice 10 : Considérons le système linéaire $Ax = b$ avec : $A = \begin{pmatrix} 0.001 & 1 \\ 1 & 1 \end{pmatrix}, b = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$

- a) Résoudre de manière exacte le système linéaire
- b) On suppose que les nombres sont représentés en virgule flottante dans une base décimale avec 3 chiffres significatifs et que le résultat est arrondi à 3 chiffres significatifs.
 - a. Appliquer l'algorithme de Gauss avec 0.001 comme pivot.
 - b. Appliquer l'algorithme de Gauss avec 1 comme pivot.
 - c. Conclure.

Propriété Le nombre d'opérations est de l'ordre de $\frac{2}{3}n^3$.

b) la méthode de Crout

Utilisation de la décomposition de A en un produit de deux matrices triangulaires de même ordre n, $A = L \times U$ avec L matrice triangulaire inférieure (« lower ») et U triangulaire supérieure (« upper »).

Proposition : Si A est inversible et vérifie pour tout $k=1, \dots, n$ $\det \begin{pmatrix} a_{1,1} & \dots & a_{1,j} & \dots & a_{1,k} \\ \dots & & & & \dots \\ a_{i,1} & & a_{i,j} & \dots & a_{i,n} \\ \dots & & & & \dots \\ a_{k,1} & \dots & a_{k,j} & \dots & a_{k,k} \end{pmatrix} \neq 0$,

la décomposition existe, L et U sont inversibles et la décomposition est unique.

Méthode de Crout : La méthode consiste à substituer au système de n équations à n inconnues (x), un système de 2n équations à 2n inconnues (x,y) :

$$Ax = b \Leftrightarrow Ly = b \quad \text{et} \quad Ux = y$$

Mais il faut pouvoir résoudre auparavant le système suivant de n^2 équations à n^2 inconnues (l,u) :

$$A = \begin{pmatrix} a_{1,1} & \dots & a_{1,j} & \dots & a_{1,n} \\ \dots & & & & \\ a_{i,1} & & a_{i,j} & \dots & a_{i,n} \\ \dots & & & & \\ a_{n,1} & \dots & a_{n,j} & \dots & a_{n,n} \end{pmatrix} = \begin{pmatrix} 1 & \dots & 0 & \dots & 0 \\ \dots & & & & \\ l_{i,1} & & 1 & \dots & 0 \\ \dots & & & & \\ l_{n,1} & \dots & l_{n,j} & \dots & l_{n,n} \end{pmatrix} \begin{pmatrix} u_{1,1} & \dots & u_{1,j} & \dots & u_{1,n} \\ \dots & & & & \\ 0 & & u_{i,i} & \dots & u_{i,n} \\ \dots & & & & \\ 0 & \dots & 0 & \dots & u_{n,n} \end{pmatrix} = LU$$

Ce système est non linéaire et il est a priori plus compliqué que le système initial. L'algorithme de Crout consiste à ordonner les équations pour trouver la solution unique !!

L'ordre est donné comme suit :

$$\begin{pmatrix} a_{1,1} & a_{1,j} & \dots & a_{1,n} \\ \dots & & & \\ a_{i,1} & a_{i,j} & \dots & a_{i,n} \\ \dots & & & \\ a_{n,1} & a_{n,j} & \dots & a_{n,n} \end{pmatrix} = \begin{pmatrix} 1 & \dots & 0 & \dots & 0 \\ \dots & & & & \\ l_{i,1} & & 1 & \dots & 0 \\ \dots & & & & \\ l_{n,1} & \dots & l_{n,j} & \dots & l_{n,n} \end{pmatrix} \begin{pmatrix} u_{1,1} & \dots & u_{1,j} & \dots & u_{1,n} \\ \dots & & & & \\ 0 & & u_{i,i} & \dots & u_{i,n} \\ \dots & & & & \\ 0 & \dots & 0 & \dots & u_{n,n} \end{pmatrix}$$

Exemple : cas n=3

$$\begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ l_{2,1} & 1 & 0 \\ l_{3,1} & l_{3,2} & 1 \end{pmatrix} \begin{pmatrix} u_{1,1} & u_{1,2} & u_{1,3} \\ 0 & u_{2,2} & u_{2,3} \\ 0 & 0 & u_{3,3} \end{pmatrix}$$

$$\Leftrightarrow \begin{cases} a_{1,1} = u_{1,1} \\ a_{2,1} = l_{2,1} u_{1,1} \\ a_{3,1} = l_{3,1} u_{1,1} \\ a_{1,2} = u_{1,2} \\ a_{2,2} = l_{2,1} u_{1,2} + u_{2,2} \\ a_{3,2} = l_{3,1} u_{1,2} + l_{3,2} u_{2,2} \\ a_{1,3} = u_{1,3} \\ a_{2,3} = l_{2,1} u_{1,3} + u_{2,3} \\ a_{3,3} = l_{3,1} u_{1,3} + l_{3,2} u_{2,3} + u_{3,3} \end{cases}$$

(l' inconnue calculée)

Algorithme de Crout

Pour i de 1 à j , faire $u_{i,j} = a_{i,j} - \sum_{k=1}^{i-1} l_{i,k} u_{k,j}$

Pour i de $j+1$ à n , faire $l_{i,j} = (a_{i,j} - \sum_{k=1}^{j-1} l_{i,k} u_{k,j}) / u_{j,j}$

Application : Utilisation de l'algorithme de Crout pour trouver l'inverse d'une matrice M

$$M = \begin{pmatrix} m_{1,1} & & m_{1,n} \\ m_{i,1} & m_{i,i} & m_{i,n} \\ m_{n,1} & & m_{n,n} \end{pmatrix}, M^{-1} = \begin{pmatrix} x_{1,1} & & x_{1,n} \\ x_{i,1} & x_{i,i} & x_{i,n} \\ x_{n,1} & & x_{n,n} \end{pmatrix}$$

$$MM^{-1} = I \Leftrightarrow M \begin{pmatrix} x_{j,1} \\ x_{j,j} \\ x_{j,n} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad j = 1, \dots, n$$

Exercice 11 : Inverser la matrice $\begin{pmatrix} 1 & -1 & -1 \\ 2 & 1 & 0 \\ 3 & 0 & 1 \end{pmatrix}$

Remarque : De manière générale, la méthode de Crout est utilisée dès lors que la matrice A intervient dans plusieurs systèmes linéaires.

Exercice 12 : Proposer deux stratégies basées sur l'algorithme de Crout pour résoudre le système $A^2x=b$. Quelle stratégie est la meilleure en termes de nombre d'opérations ?

c) la méthode de Cholesky

Au lieu d'utiliser la décomposition de Crout, A est décomposée en un produit LL^T où L est une matrice triangulaire de même ordre n, triangulaire inférieure. Cette décomposition n'existe pas toujours ! En particulier A doit être symétrique ($A^T=A$) et A inversible implique que $l_{k,k}$ est non nul, pour tout k.

Algorithme de Cholesky (Exercice 11): Les colonnes sont déterminées par récurrence.

$$1) \quad A = LL^T \Rightarrow \forall j \leq i, \quad a_{i,j} = \sum_{s=1}^j l_{i,s} l_{j,s}$$

$$2) \quad (\text{1ère colonne}) \quad l_{1,1} = \sqrt{a_{1,1}}, \quad \forall i = 2, \dots, n, \quad l_{i,1} = \frac{a_{i,1}}{l_{1,1}}$$

$$3) \quad \text{pour } k = 2, n$$

$$l_{k,k} = \sqrt{a_{k,k} - \sum_{s=1}^{k-1} l_{k,s}^2},$$

$$\text{pour } i = (k+1), n$$

$$l_{i,k} = \frac{a_{i,k} - \sum_{s=1}^{k-1} l_{i,s} l_{k,s}}{l_{k,k}}$$

Exercice 13 : Appliquer la méthode de Cholesky pour trouver la matrice L vérifiant

$$LL^T = A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 2 \end{pmatrix}$$

Propriété : le nombre d'opérations pour trouver la matrice L est de l'ordre de $\frac{1}{3}n^3$. Résoudre un système linéaire avec cette méthode est donc d'un coût inférieur à la méthode de Gauss.

I.2.d Des Méthodes itératives

Dès lors que le nombre d'opérations devient trop important, on a recours à des résolutions approchées. Les méthodes exploitent la réécriture de l'équation « $Ax=b$ », avec A inversible, sous la forme « $x=Bx+c$ » avec la matrice (I-B) inversible. Cette nouvelle écriture suggère la définition d'une méthode itérative de résolution du système linéaire « $Ax=b$ » : soit $x^{(0)}$ un vecteur initial,

$$x^{(k+1)} = Bx^{(k)} + c, k \geq 0$$

La méthode itérative est convergente si $\lim_{k \rightarrow \infty} x^{(k)} = x, \forall x^{(0)}$

(inspirée de la méthode du point fixe)

a) Le principe général

Trouver une décomposition de A sous la forme « M-N » avec M une matrice **inversible** et **facile à inverser**.

$$Ax = b \Leftrightarrow Mx = Nx + b \Leftrightarrow x = M^{-1}Nx + M^{-1}b$$

Cette dernière forme est sous la forme $x=Bx+c$, avec $B=M^{-1}N$ et $c=M^{-1}b$. Si A et M sont inversibles,

$M^{-1}A=M^{-1}(M-N)=I-B$ est aussi inversible. Les méthodes de Jacobi et de Gauss-Seidel sont issues de la décomposition très simple de A en trois matrices, une diagonale D, une triangulaire inférieure E et une triangulaire supérieure F:

$$A = D - E - F$$

$$D_{i,j} = a_{i,j} \delta_{i,j}$$

$$(-E)_{i,j} = \begin{cases} a_{i,j}, & \text{si } i > j \\ 0, & \text{sin on} \end{cases}$$

$$(-F)_{i,j} = \begin{cases} a_{i,j}, & \text{si } i < j \\ 0, & \text{sin on} \end{cases}$$

Hypothèse : $a_{i,i} \neq 0, \forall i = 1, \dots, n$

b) La méthode de Jacobi

On choisit ici $M = D$ et $N = E + F$. Étant donné l'hypothèse faite sur les éléments diagonaux de A , M est inversible et facile à inverser (diagonale).

On a alors

$$Ax = b \Leftrightarrow Dx = (E + F)x + b \Leftrightarrow x = D^{-1}(E + F)x + D^{-1}b,$$

ce qui conduit à la **méthode itérative de Jacobi** :

$$Dx^{(k+1)} = (E + F)x^{(k)} + b, k \geq 0.$$

La matrice de cette méthode itérative est donc

$$J = D^{-1}(E + F) = D^{-1}(D - A) = I - D^{-1}A$$

et s'appelle la **matrice de Jacobi**.

$$\begin{aligned} a_{11}x_1^{(k+1)} &= -a_{12}x_2^{(k)} - a_{13}x_3^{(k)} \dots - a_{1,n-1}x_{n-1}^{(k)} - a_{1n}x_n^{(k)} + b_1 \\ a_{22}x_2^{(k+1)} &= -a_{21}x_1^{(k)} - a_{23}x_3^{(k)} \dots - a_{2,n-1}x_{n-1}^{(k)} - a_{2n}x_n^{(k)} + b_2 \\ &\vdots \\ a_{n-1,n-1}x_{n-1}^{(k+1)} &= -a_{n-1,1}x_1^{(k)} \dots - a_{n-1,n-2}x_{n-2}^{(k)} - a_{n-1,n}x_n^{(k)} + b_{n-1} \\ a_{nn}x_n^{(k+1)} &= -a_{n,1}x_1^{(k)} \dots - a_{n,n-2}x_{n-2}^{(k)} - a_{n,n-1}x_{n-1}^{(k)} + b_n. \end{aligned}$$

c) La méthode de Gauss-Seidel

Il semble qu'on puisse améliorer la méthode en utilisant "mieux" les quantités déjà calculées. En effet, pour calculer la seconde composante $x_2^{(k+1)}$, pourquoi ne pas utiliser la nouvelle valeur $x_1^{(k+1)}$ au lieu de l'ancienne $x_1^{(k)}$? Cela conduit au système

$$\begin{aligned} a_{11}x_1^{(k+1)} &= -a_{12}x_2^{(k)} - a_{13}x_3^{(k)} \dots - a_{1,n-1}x_{n-1}^{(k)} - a_{1n}x_n^{(k)} + b_1 \\ a_{22}x_2^{(k+1)} &= -a_{21}x_1^{(k+1)} - a_{23}x_3^{(k)} \dots - a_{2,n-1}x_{n-1}^{(k)} - a_{2n}x_n^{(k)} + b_2 \\ &\vdots \\ a_{n-1,n-1}x_{n-1}^{(k+1)} &= -a_{n-1,1}x_1^{(k+1)} \dots - a_{n-1,n-2}x_{n-2}^{(k+1)} - a_{n-1,n}x_n^{(k)} + b_{n-1} \\ a_{nn}x_n^{(k+1)} &= -a_{n,1}x_1^{(k+1)} \dots - a_{n,n-2}x_{n-2}^{(k+1)} - a_{n,n-1}x_{n-1}^{(k+1)} + b_n. \end{aligned}$$

Ceci définit une nouvelle méthode itérative qu'on appelle **méthode itérative de Gauss-Seidel** et qui s'écrit, en notation matricielle,

$$Dx^{(k+1)} = Ex^{(k+1)} + Fx^{(k)} + b \Leftrightarrow (D - E)x^{(k+1)} = Fx^{(k)} + b.$$

Étant donné l'hypothèse faite sur les éléments diagonaux de A , la matrice $(D - E)$ est inversible et la **matrice de Gauss-Seidel** est

$$\mathcal{L}_1 = (D - E)^{-1}F.$$

d) *Convergence des méthodes itératives*

Définition

Soit $A \in M_n(\mathbb{R})$. On dit que A est à diagonale strictement dominante si

$$\forall i = 1, 2, \dots, n, |a_{ii}| > \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}|.$$

Proposition:

Si $A \in M_n(\mathbb{R})$ est à diagonale strictement dominante, alors A est inversible.

Théorèmes :

Si $A \in M_n(\mathbb{R})$ est à diagonale strictement dominante, alors quel que soit le vecteur de départ, la méthode de Jacobi converge vers l'unique solution de $Ax = b$.

Si $A \in M_n(\mathbb{R})$ est à diagonale strictement dominante, alors quel que soit le vecteur de départ, la méthode de Gauss-Seidel converge vers l'unique solution de $Ax = b$.

Exercice 14

On considère le système $Ax = b$ avec

$$A = \begin{pmatrix} 1 & 1/2 \\ 1/2 & 1 \end{pmatrix} \text{ et } b = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \text{ dont la solution est } x = \begin{pmatrix} 0 \\ 2 \end{pmatrix}$$

Montrer que les méthodes de Jacobi, puis Gauss-Seidel conduisent aux relations et solutions suivantes :

$$x^{(k+1)} = \begin{pmatrix} 0 & -1/2 \\ -1/2 & 0 \end{pmatrix} x^{(k)} + \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

$$x^{(k+1)} = \begin{pmatrix} 0 & -1/2 \\ -1/2 & 0 \end{pmatrix} \begin{pmatrix} 1/(2^k) \\ (2^{k+1}-1)/(2^k) \end{pmatrix} + \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 1/(2^{k+1}) \\ (2^{k+2}-1)/(2^{k+1}) \end{pmatrix} \rightarrow \begin{pmatrix} 0 \\ 2 \end{pmatrix}$$

$$x^{(k+1)} = \begin{pmatrix} 0 & -1/2 \\ 0 & 1/4 \end{pmatrix} x^{(k)} + \begin{pmatrix} 1 \\ 3/2 \end{pmatrix}$$

$$\begin{aligned}
x^{(k+1)} &= \begin{pmatrix} 0 & -1/2 \\ 0 & 1/4 \end{pmatrix} \begin{pmatrix} 1/(2^{2k-1}) \\ (2^{2k+1} - 1)/(2^{2k}) \end{pmatrix} + \begin{pmatrix} 1 \\ 3/2 \end{pmatrix} = \begin{pmatrix} 1/(2^{2k+1}) \\ (2^{2k+3} - 1)/(2^{2k+2}) \end{pmatrix} \\
&= \begin{pmatrix} 1/(2^{2(k+1)-1}) \\ (2^{2(k+1)+1} - 1)/(2^{2(k+1)}) \end{pmatrix} \\
&\rightarrow \begin{pmatrix} 0 \\ 2 \end{pmatrix}.
\end{aligned}$$

Définition : Si M est une matrice diagonalisable, $M = P^{-1} \text{diag} (\lambda_1, \dots, \lambda_n) P$, On appelle rayon spectral de M, le réel $\rho(M) = \max_i |\lambda_i|$

Théorème : Si $\rho(M) < 1$ (resp. > 1), alors la suite $x^{(k+1)} = Mx^{(k)} + C$ converge (resp. diverge) pour tout vecteur C et tout vecteur $x^{(0)}$.

Définition : Une application V de l'ensemble des matrices réelles n x n dans \mathbb{R}^+ est une norme matricielle si elle vérifie les 4 conditions suivantes :

1. $V(A) = 0 \Leftrightarrow A = 0$
2. $\forall \lambda \in \mathbb{R}, V(\lambda A) = |\lambda| V(A)$
3. $V(A + B) \leq V(A) + V(B)$
4. $V(AB) \leq V(A)V(B)$

Théorème

Soit N une norme vectorielle sur \mathbb{R}^n , l'application V définie sur l'ensemble des matrices n x n, par

$$V(A) = \sup_{x \neq 0} \frac{N(Ax)}{N(x)}$$

est une norme matricielle

Exemples : $\|\cdot\|_1, \|\cdot\|_2, \|\cdot\|_F, \|\cdot\|_\infty$

Théorème : s'il existe une norme matricielle V telle que $V(M) < 1$, alors la suite converge $x^{(k+1)} = Mx^{(k)} + C$, pour tout vecteur C et tout vecteur $x^{(0)}$.

II. Notions d'analyse

II.1 Calcul différentiel

II.1.a Fonction dérivable et différentiable

II.1.b Recherche du minimum d'une fonction ; algorithme de descente du gradient

II.2 Recherche du zéro d'une fonction

II.3 Equations différentiables linéaires

III. Notions de probabilités

III.1 Conditionnement et vraisemblance

III.2 Régression linéaire – Modèles linéaires