Oral I - Agrégation interne Algèbre
Leçon 150
Diverses factorisations de matrices. Applications.

J. Mellac

Décembre 2015-Janvier 2016

Table des matières

1 Décomposition A=LU

On considère un corps $\mathbb{K}$ et $A \in GL_n(\mathbb{K})$.

1.1 Définitions. Rappel

Définition. On appelle opération élémentaire sur les lignes de la matrice A toute opération de l'un des types suivants : En notant L_i la i-ème ligne de A :

— $L_i \leftarrow L_i + \lambda L_j$, où $\lambda \in \mathbb{K}^*$: on ajoute à L_i la j-ème ligne de A multipliée par λ.

— $L_i \leftarrow \lambda L_i$: on multiplie L_i par un scalaire $\lambda \in \mathbb{K}^*$.

— $L_i \longleftrightarrow L_j$: permutation de 2 lignes de A.

On peut définir de la même manière des opérations élémentaires sur les colonnes d'une matrice.

1.2 Traduction en terme de produits matriciels

Proposition. Chaque opération sur les lignes de $A \in GL_n(\mathbb{K})$ consiste à multiplier A à gauche par une matrice inversible.

$L_i \leftarrow L_i + \lambda L_j$ par $T_{i,j}(\lambda) = I_n + \lambda E_{i,j} \in GL_n(\mathbb{K})$. Matrice de transvection

$L_i \leftarrow \lambda L_i$ par $D_i(\lambda) = I_n + (\lambda - 1)E_{i,i} \in GL_n(\mathbb{K})$. Matrice de dilatation

$L_i \longleftrightarrow L_j$ par $P_{i,j}(\lambda) = I_n - (E_{i,i} + E_{j,j}) + (E_{i,j} + E_{j,i}) \in GL_n(\mathbb{K})$. Matrice de permutation

Démonstration. Montrons le premier résultat, les autres se traitant de la même manière. On considère la base canonique $(E_{kl})_{(k,l)\in[\![1,n]\!]\times[\![1,n]\!]}$ de $\mathcal{M}_n(\mathbb{K})$, $A = \sum_{k=1}^{n}\sum_{l=1}^{n} a_{kl}E_{kl}$. On utilise le résultat

$E_{i,j}\, E_{kl} = \delta_{jk}\, E_{il}$ où δ_{jk} est le symbole de kronecker

$$(I_n + \lambda E_{ij})A = A + \sum_{k=1}^{n}\sum_{l=1}^{n} a_{kl}E_{ij}\, E_{kl} = A + \lambda \sum_{l=1}^{n} a_{jl}E_{il}$$

ce qui montre que la ligne λL_j est ajoutée à la ligne L_i.

De même chaque opération sur les colonnes de A consiste à multiplier A à droite par une matrice inversible.

Définition. Soit A_j, $j \in [\![1, n]\!]$ les matrices extraites de A obtenues en gardant les j premières lignes et les j premières colonnes. Les déterminants de ces n matrices sont appelés déterminants principaux de A.

Proposition. La matrice A admet une factorisation sous la forme $A = LU$, où L et U sont des matrices respectivement triangulaires inférieures et supérieures L ne contenant que des 1 sur la diagonale si et seulement si tous les déterminants principaux de A sont non nuls. Une telle décomposition est alors unique.

Démonstration. Faisons une récurrence sur n.

Pour $n = 1$ le résultat est vérifié.

supposons le résultat vérifié pour $n-1$ avec $n \geq 2$. Soit $A \in GL_n(\mathbb{K})$ ayant tous les déterminants principaux non nuls.

Ecrivons A sous forme de blocs :

$A=\begin{pmatrix} A_1 & B_1 \\ C_1 & a_{nn}\end{pmatrix}$ avec $A_1 \in \mathcal{M}_{n-1}(\mathbb{K}),\ B_1 \in \mathcal{M}_{n-1,1}(\mathbb{K}),\ C_1 \in \mathcal{M}_{1,n-1}(\mathbb{K}),\ a_{nn} \in \mathbb{K}$.
On a $A_1 \in GL_{n-1}(\mathbb{K}),\ a_{nn} \in \mathbb{K}^*$ car les déterminants principaux de A sont non nuls.
D'après l'hypothèse de récurrence il existe L_1 triangulaire inférieure à diagonale unité et U_1 triangulaire supérieure, toutes deux dans $GL_{n-1}(\mathbb{K})$ telles que $A_1 = L_1U_1$.

$$L=\begin{pmatrix} L_1 & 0 \\ E_1 & 1\end{pmatrix} U=\begin{pmatrix} U_1 & F_1 \\ 0 & a\end{pmatrix}$$

avec $E_1 \in \mathcal{M}_{1,n-1}(\mathbb{K}),\ F_1 \in \mathcal{M}_{n-1,1}(\mathbb{K}),\ a \in \mathbb{K}^*$. L est triangulaire inférieure à diagonale unité et U triangulaire supérieure, $U \in GL_n(\mathbb{K})$ car $\det(U) = \det(A)$.
L'égalité $A = LU$ est équivalente à $A_1 = L_1U_1,\ B_1 = L_1F_1,\ C_1 = E_1U_1,\ a_{nn} = a + E_1F_1$ en utilisant les produits par blocs. L_1 est inversible, le système $L_1F_1 = B_1$ a une solution unique $F_1 \in \mathcal{M}_{n-1,1}(\mathbb{K})$, $E_1U_1 = C_1 \Leftrightarrow {}^tU_1{}^tE_1 = {}^tC_1$ a une solution unique $E_1 \in \mathcal{M}_{1,n-1}(\mathbb{K})$ et enfin $a = a_{nn} - E_1F_1$. ce qui donne une décomposition LU.

Réciproquement, si $A \in GL_n(\mathbb{K})$ est égale à LU sous les hypothèses précédentes faites sur ces matrices, U est inversible et la décomposition par bloc précédente montre que tous les déterminants principaux de A sont non nuls car $A_1 = L_1U_1$ et $\det(A_1) = \dfrac{\det(U)}{a} \neq 0$, tous les déterminants principaux de A_1 sont non nuls d'après l'hypothése de récurrence. Montrons l'unicité de cette décomposition :
supposons $A = LU = L'U'$ sous les hypothèses précédentes. On a alors $L'^{-1}L = U'U^{-1}$, matrice à la fois triangulaire inférieure et supérieure, donc diagonale, cette diagonale ne contenant que des 1 avec l'hypothése faite sur U et sur U', d'où $L'^{-1}L = U'U^{-1} = I_n,\ L' = L,\ U' = U$.

Remarque. La méthode de Pivot de Gauss permet d'obtenir L et U de la manière suivante :
Sous l' hypothèse que les déterminants principaux sont non nuls on a :

> **Proposition.** Toute matrice inversible, pour laquelle les déterminants principaux sont non nuls, peut être transformée en une matrice triangulaire supérieure par une suite d'opérations élémentaires ne nécessitant pas de permutation de ligne.

Démonstration. Récurrence sur n taille de la matrice. Pour $n = 1$ c'est immédiat.
Supposons le résultat vrai pour $n - 1 \geq 1$ et soit $A \in GL_n(\mathbb{K})$. On a $a_{11} \neq 0$ d'aprés l'hypothèse.
On effectue les opérations élémentaires suivantes : $L_i \longleftarrow L_i - \dfrac{a_{i1}}{a_{11}}L_1, i \in [\![2, n]\!]$. La matrice obtenue est de la forme : $A_1 = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a'_{22} & \cdots & a'_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ 0 & a'_{n2} & \cdots & a'_{nn}\end{pmatrix}$ on applique la récurrence à la matrice inversible
$A' = \begin{pmatrix} a'_{22} & \cdots & a'_{2n} \\ \vdots & \vdots & \vdots \\ a'_{n2} & \cdots & a'_{nn}\end{pmatrix}$
ce qui est possible car $a'_{22} = \dfrac{a_{11}a_{22} - a_{12}a_{21}}{a_{11}} \neq 0$.

Chacune de ces opérations élémentaires correspond à multiplier A à gauche par une matrice de transvection qui est triangulaire inférieure avec des 1 sur la diagonale. Ce qui donne $P_rP_{r-1}\cdots P_1A = U$ où les P_i sont des matrices de transvection et U la matrice triangulaire supérieure obtenue. Il suffit de prendre $L = (P_rP_{r-1}\cdots P_1)^{-1}$. On peut remarquer que la matrice $P_rP_{r-1}\cdots P_1 = P_rP_{r-1}\cdots P_1\, I_n$ peut être obtenue en appliquant la même suite d'opérations élémentaires, dans le même ordre à I_n.

1.3 Algorithme de factorisation LU : méthode de Crout

Pour déterminer L et U on utilise la méthode des coefficients indéterminés, c'est à dire qu'on pose $A = LU$, on effectue le produit et on identifie, ce qui donne :

$$A = LU = \begin{pmatrix} 1 & 0 & \cdots & \cdots & 0 \\ l_{21} & 1 & 0 & \cdots & 0 \\ l_{31} & l_{32} & 1 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ l_{n1} & l_{n2} & \cdots & l_{nn-1} & 1 \end{pmatrix} \begin{pmatrix} u_{11} & u_{12} & \cdots & \cdots & u_{1n} \\ 0 & u_{22} & \cdots & \cdots & u_{2n} \\ 0 & 0 & u_{33} & \cdots & u_{3n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ u_{n1} & u_{n2} & u_{n3} & \cdots & u_{nn} \end{pmatrix}$$

$u_{11} = a_{11}$
pour $j = 2$ à n :
$u_{1j} = a_{1j}$
$l_{j1} = \dfrac{a_{j1}}{a_{11}}$

Puis pour $i = 2..n-1$

$u_{ii} = a_{ii} - \displaystyle\sum_{k=1}^{i-1} l_{ik}u_{ki}$

Pour $j = i+1..n$

$u_{ij} = a_{ij} - \displaystyle\sum_{k=1}^{i-1} l_{ik}u_{kj}$

$l_{ji} = \dfrac{1}{u_{ii}}\left(a_{ji} - \displaystyle\sum_{k=1}^{i-1} l_{jk}u_{ki}\right)$

$u_{nn} = a_{nn} - \displaystyle\sum_{k=1}^{n-1} l_{nk}u_{kn}.$

Complexité de l'algorithme

La ligne i nécessite $N_i = \displaystyle\sum_{j=1}^{i-1} j + \sum_{j=1}^{i-1}(i-1) = \frac{i(i-1)}{2} + (n-i+1)(i-1)$ multiplications soit $N_1 = \displaystyle\sum_{i=1}^{n} = \frac{n(n^2-1)}{3}$ multiplications et un calcul analogue donne $N_2 = \dfrac{n(n-1)(2n-1)}{6}$ additions, on obtient donc une complexité en $O(n^3)$.

Exemple

$A = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix}$, en utilisat l'algorithme ci-dessus on obtient :

$$L = \begin{pmatrix} 1 & 0 & 0 \\ -\frac{1}{2} & 1 & 0 \\ 0 & -\frac{2}{3} & 1 \end{pmatrix} \quad U = \begin{pmatrix} 2 & -1 & 0 \\ 0 & \frac{3}{2} & -1 \\ 0 & 0 & \frac{4}{3} \end{pmatrix}.$$

1.4 Factorisation A=PLU

Lorsque A est une matrice inversible pour laqulle la méthode de Gauss nécessite des permutations de lignes la factorisation précédente n'est plus possible. Ceci peut être dû à des déterminants principaux nuls ou à des pivots trop petits en valeur absolue nécéssitant d'utiliser la méthode du pivot partiel consistant à choisir à l'étape i le coefficient le plus grand en valeur absolue sur la colonne i,c'est à dire à permuter deux lignes. Au point de vue matriciel ceci revient à mutiplier la matrice par la matrice de permutation $I_n - (E_{i,i} + E_{j,j}) + (E_{i,j} + E_{j,i})$, matrice comportant un 1 sur chaque ligne et chaque colonne et des zéros ailleurs. Il est aisé de constater que la produit de matrices de permutations est une matrice de permutation.

Proposition. La matrice inversible A admet une factorisation sous la forme $PA = LU$, où L et U sont des matrices respectivement triangulaires inférieures et supérieures L ne contenant que des 1 sur la diagonale et P une matrice de permutation.

Démonstration. En tenant compte des permutations éventuelles de lignes, on peut écrire en utilisant le pivot de Gauss, $T_rP_r \cdots T_1P_1A = U$ où les T_i sont des matrices de transvection et P_i de permutation, éventuellement $P_i = I_n$. On peut remarquer que, avec les notations du début $P_{ij}\,T\,P_{ij}$ est la matrice obtenue en permutant les colonnes i et j de T, puis en réalisant la meme opération sur les lignes de la matrice transformée, d'ù $P_{ij}\,T = T'P_{ij}$ où T' est une matrice de transvection et $T_rP_r \cdots T_1P_1A = T'_rT'_{r-1} \cdots T'_1PA = U$ où P est une matrice de permutation et T'_i sont des matrices de transvection. En posant $L = \left(T'_rT'_{r-1} \cdots T'_1\right)^{-1}$ on obtient le résultat cherché.

1.5 Applications

On veut résoudre le système linéaire de Cramer :

$$(S)\begin{cases} a_{11}x_1 + a_{12}x_2 + ... + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + ... + a_{2n}x_n = b_2 \\ \vdots \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + ... + a_{nn}x_n = b_n \end{cases}$$

appelé système de n équations à n inconnues où $(b_1, b_2, \cdots, b_n) \in \mathbb{K}^n$ est appelé second membre du système et $A = (a_{ij})_{(i,j)\in[\![1,n]\!]^2} \in GL_n(K)$ est la matrice du système qui peut être écrit $AX = B$.
Si l'on connait une décomposition de $A = LU$ résoudre ce système $A\,X = B$ revient à résoudre les deux systèmes, en posant $Z = UX$:
$LZ = B$, $UX = Z$ où X, B et Z sont des vecteurs colonnes.

Résolution de $LZ = B$.

$z_1 = b_1$.
Pour $i = 2..n$,

$$z_i = b_i - \sum_{j=1}^{i-1} l_{ij}z_j$$

fin

Résolution de $UX = Z$.

$x_n = \dfrac{z_n}{u_{nn}}$
Pour $i = n - 1..1$,

$$x_i = \frac{z_i - \sum_{j=i+1}^{n} u_{ij}x_j}{u_{ii}}$$

fin

Le premier algorithme contient $\sum_{i=1}^{n}(i-1) = \frac{n(n-1)}{2}$ opérations, le deuxième algorithme contient $\sum_{i=1}^{n-1}(n-i+1) = \frac{n(n+1)}{2}$ opérations, soit n^2 opérations au total, la résolution de ces systèmes est donc de complexité $O(n^2)$, à comparer avec la complexité en $O(n^3)$ de la méthode de Gauss, si la factorisation est connue cette méthode est donc économe en temps de calcul. Elle est intéressante lorsqu'il y a plusieurs valeurs possibles de b avec la même matrice A.

1.6 Décomposition LD^tL d'une matrice symétrique réelle

On considère une matrice A symétrique réelle que l'on suppose factorisable sous forme LU. En divisant chaque ligne de de U par son terme diagonal (non nul d'après l'hypothèse de factorisation) on peut écrire $U = DU'$ où la matrice diagonale D contient les termes diagonaux de U, ce qui donne $A = LDU'$ et ${}^tA = A \Rightarrow U' = {}^tL$ par l'unicité de la décomposition $A = LU$.
Si la matrice A est la matrice d'une forme quadratique Q dans une base donnée, en notant p le nombre de termes positifs sur la diagonale de D, la signatures de Q est $(p, n-p)$ car les matrices A et D sont congruentes.

2 Décomposition de Cholesky

Soit A une matrice réelle symétrique d'ordre n.

Définition. La matrice symétrique réelle A est définie positive si : $\forall X \in \mathbb{R}^n,\ X \neq 0,\ {}^tXAX > 0$. Ceci revient à dire que A est la matrice d'un produit scalaire sur $\mathbb{R}^n$ défini par $< X|Y >= {}^tXAY$.

Proposition. Soit $A \in \mathcal{S}_n(\mathbb{R})$ matrice symétrique réelle d'ordre n. Les conditions suivantes sont équivalentes :
1) Les valeurs propres de A sont strictement positives.
2) Pour tout vecteur X non nul de $\mathbb{R}^n$, on a ${}^tXAX > 0$, c'est à dire A est définie positive..
3) Il existe une matrice triangulaire inférieure inversible B telle que $A = B^tB$. Cette factorisation est la factorisation de Cholesky de A. La décomposition est unique si on impose la stricte positivité des coefficients de B.

Démonstration. Montrons l'équivalence entre 1 et 2.
On considère une base orthonormale de vecteurs propres de l'automorphisme canoniquement associé à A, $\mathcal{B} = (f_1, \cdots f_n)$ soit $X = \sum_{i=1}^{n} x_i f_i \in \mathbb{R}^n$, non nul :
${}^tXAX = \sum_{i=1}^{n} \lambda_i x_i^2 > 0$. Réciproquement en prenant $X = f_i$ on a ${}^tXAX = \lambda_i > 0$. Montrons l'équivalence entre 1 et 3 :
Une telle matrice admet une décomposition LU comme le montre la proposition ci-dessous.
Si A est symétrique définie positive $A = LD^tL$ et les coefficients diagonaux de D sont strictements positifs, on peut écrire $D = D'^2$ et en posant $B = LD'$ on obtient la décomposition cherchée.
Réciproquement, A est symétrique et
$\forall x \in \mathbb{R}^n,\ x \neq 0,\ ,\ (Ax|x) = (B^tBx|x) = ({}^tBx|{}^tBx) > 0$ et A est définie positive.

Remarques. 1) Le procédé d'orthonormalisation de Gram Schmidt permet de démontrer ce résultat et de déterminer B : On munit $\mathbb{R}^n$ du poduit scalaire $(X|Y) = {}^tXAY$ car A est symétrique définie positive. en partant de la base canonique $\mathcal{B}' = (e_1, \cdots e_n)$ on construit une base orthonormale $\mathcal{B}$ de la manière suivante :

$g_1 = e_1,\ f_1 = \dfrac{g_1}{||g_1||}$

$g_i = e_i - \displaystyle\sum_{k=1}^{i-1}(e_k|f_k)f_k,\ f_i = \dfrac{g_i}{||g_i||},\ i \in [\![2, n]\!]$.

On a $\mathrm{Vect}(e_1, \cdots, e_k) = \mathrm{Vect}(f_1, \cdots, f_k),\ k \in [\![1, n]\!]$ et la base $\mathcal{B} = (f_1, f_2, \cdots, f_n)$ est orthonormale La matrice de passage de $\mathcal{B}'$ à $\mathcal{B}$ est triangulaire supérieure de terme diagonaux $\dfrac{1}{||g_i||} > 0$. la matrice du produit scalaire dans ctte base $\mathcal{B}$ est $I_n = {}^tPAP$ ce qui donne $A = B{}^tB$ avec $B = {}^tP^{-1}$.

2) Le calcul effectif des coefficients de B se fait par identification :

$$b_{11}^2 = a_{11}$$

$$b_{ii}^2 = a_{ii} - \sum_{k=1}^{i-1} b_{ik}^2,\ i \in [\![2, n]\!]$$

$$b_{ij} = \frac{a_{ij} - \sum_{k=1}^{j-1} b_{ik}b_{jk}}{b_{jj}},\ j \in [\![1, i-1]\!].$$

Ce qui montre l'unicité de la décomposition si les coefficients diagonaux sont positifs.

3) La complexité de cet algorithme est en $O(n^3)$:

Il y a $\displaystyle\sum_{i=1}^{n}(i-1) = \frac{n(n-1)}{2}$ divisions.

Il y a $\displaystyle\sum_{i=1}^{n}\left((i-1) + \sum_{j=1}^{i-1}(j-1)\right)$ additions et le même nombre de multiplications soit au total $\dfrac{2n^3 + 3n^2 + n}{6}$ opérations.

Proposition. Une matrice symétrique réelle A est définie positive si et seulement si tous ses déterminants principaux sont strictement positifs.

Démonstration. Si la matrice symétrique A est définie positive, en considérant $E_i = \mathrm{Vect}(f_1, \cdots, f_i),\ i \in [\![1, i]\!]$ où les vecteurs f_i sont des vecteurs propres associés respectivement aux valeurs propres λ_i, la restriction de l'automorphisme canoniquement associé à A, au sous espace stable E_i admet des valeurs propres strictement positives son déterminant est strictement positif.
Réciproquement supposons tous les mineurs principaux strictement positifs. Faisons une récurrence sur n.
Si $n = 1$ A est définie positive.
Si la propriété est vraie pour $n - 1 \geq 1$, notons ϕ la forme bilinéaire symétrique de matrice A dans la base canonique, q la forme quadrique associée et $E = \mathbb{R}^{n-1} \times \{0\}$. $q|_E$ est par hypothèse définie positive et q non dégénérée car $\det(A) \neq 0$. On a $\mathbb{R}^n = E \oplus E^\perp$ avec $\dim E^\perp = 1,\ E^\perp = \mathbb{R}e$. On note $a = q(e)$.
en faisant un raisonnement analogue à celui de la partie directe on déduit que $a, \det(A_{n-1})$ et $\det(A)$ sont de même signe, ce qui montre que q et donc A est définie positive.

Exemple. $A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 5 & 5 & 5 \\ 1 & 5 & 14 & 14 \\ 1 & 5 & 14 & 15 \end{pmatrix}$ est symétrique définie positive, les calculs précédents permettent d'obtenir :

$$B = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 1 & 2 & 3 & 0 \\ 1 & 2 & 3 & 1 \end{pmatrix}.$$

3 Factorisation QR

3.1 Généralités

Proposition. Soit $A \in GL_n(\mathbb{R})$. Il existe une matrice orthogonale Q et une matrice R triangulaire supérieure à coefficients diagonaux strictement positifs tels que $A = QR$. Cette décomposition est unique.

Démonstration. On munit $\mathbb{R}^n$ du produit scalaire canonique pour lequel la base canonique $\mathcal{B} = (e_1, \cdots, e_n)$ est orthonormale. Soit $\mathcal{B}' = (C_1, \cdots, C_n)$ la base de $\mathbb{R}^n$ formée des vecteurs colonnes de A. On peut, à l'aide du procédé d'orthonormalisation de Gram Schmidt obtenir une base orthonormale à partir de $\mathcal{B}'$, appelons $\mathcal{B}''$ cette base. La matrice de passage de $\mathcal{B}$ à $\mathcal{B}''$, $P_{\mathcal{B},\mathcal{B}''}$ orthogonale car ces deux bases sont orthonormales et la relation entre matrice de passage, $A = P_{\mathcal{B},\mathcal{B}'} = P_{\mathcal{B},\mathcal{B}''} P_{\mathcal{B}''\mathcal{B}'}$, $Q = P_{\mathcal{B},\mathcal{B}''}$ est orthognale et comme on l'a vu précédemment $R = P_{\mathcal{B}'',\mathcal{B}'} = (P_{\mathcal{B}',\mathcal{B}''})^{-1}$ est triangulaire supérieure avec des termes diagonaux $\langle g_k | f_k \rangle > 0$ les notations étant celles utilisées précédemment.
Supposons $A = Q_1 R_1 = Q_2 R_2$ avec $Q_i,\ i = 1, 2$ orthogonales et $R_i,\ i = 1, 2$ triangulaires supérieures à coefficients diagonaux strictement positifs, on a alors $U = R_1 (R_2)^{-1} = (Q_1)^{-1} Q_2 = {}^t Q_1 Q_2$ est triangulaire supérieure orthogonale à coefficients diagonaux strictement positifs et ${}^t U = U^{-1}$ est à la fois triangulaire supérieure et inférieure donc diagonale et orthogonale, d'où $U = I_n$ car les coefficients diagonaux sont strictement positifs. On a donc bien $Q_1 = Q_2$ et $R_1 = R_2$.

Corollaire. Soit $A \in \mathcal{M}_{m,n}(\mathbb{R})$ avec $m > n$ et $rg(A) = n$. Il existe une matrice orthogonale $Q \in O_m(\mathbb{R})$ et une matrice $R \in \mathcal{M}_{m,n}(\mathbb{R})$ triangulaire supérieure à coefficients diagonaux strictement positifs tels que $A = QR$, autrement dit les lignes $L_i,\ i > n$ de R sont nulles

Démonstration. On compléte A par B telle que $\tilde{A} = (A, B) \in GL_m(\mathbb{R})$, on en déduit des matrices $\tilde{Q},\ \tilde{R}$ telles que $\tilde{A} = \tilde{Q}\tilde{R}$, $\tilde{R} = (R_1, R_2)$ avec $R_1 \in \mathcal{M}_{m,n}(\mathbb{R})$, on a alors $A = QR_1$, mais dans ce cas il n'y a pas d'unicité dans le choix de B donc dans la décomposition de A.

On déduit de ce résultat la décomposition d'Iwasawa

Proposition. soit $A \in GL_n(\mathbb{R})$, Il existe une matrice orthogonale Q, une matrice diagonale à coefficients diagonaux strictement positifs et une matrice R triangulaire supérieure à coefficients diagonaux égaux à 1 tels que $A = QDR$. Cette décomposition est unique.

Démonstration. En utilisant la décomposition précédente, $A = QR$, on peut écrire $R = DR'$, D diagonale contenant les termes diagonaux de R, et la diagonale de R' formée de 1, ce qui donne la décomposition recherchée, l'unicité de la décomposition QR entraine l'unicité de cette dernière décomposition.

Remarque. Algorithme de Gram Schmidt

En notant $Q = (q_1, \cdots q_i, \cdots q_n)$ où $q_i,\ i \in [\![1, n]\!]$ est la i-ème colonne de Q, on a d'après l'étude précédente $q_i = \dfrac{a_i - \sum_{j=1}^{i-1}({}^t q_j a_i) q_j}{||a_i - \sum_{j=1}^{i-1}({}^t q_j a_i) q_j||}$ où $a_i,\ i \in [\![1, n]\!]$ est la i-ème colonne de A.

On obtient par identification :

$a_1 = r_{11} q_1$

$a_i = \displaystyle\sum_{j=1}^{1} r_{ji} q_j,\ j \in [\![1, ni]\!]$, ce qui permet d'obtenir les matrices Q et R :

Pour $i \in [\![1, n]\!]$ # Détermination de la i-ème colonne de Q et de R

Pour $j \in [\![1, i-1]\!]$

$r_{ji} = {}^t q_j a_i$

$b_i = a_i - \displaystyle\sum_{j=i}^{i-1} r_{ji} q_j$

$r_{ii} = b_i^2$

$q_i = \dfrac{b_i}{r_{ii}}$

fin

Cet algorithme est numériquement instable et produit par erreur d'arrondi des vecteurs qui ne sont pas toujours othogonaux.

3.2 Méthode de Householder

Cette méthode est beaucoup plus stable numériquement pour obtenir la décomposition QR de la matrice. Elle repose sur l'utilisation de réflexions, c'est à dire de symétries orthogonales par rapport à des hyperplans de $\mathbb{R}^n$, sur lequel on considère le produit scalaire canonique.

Proposition. Etant donné deux vecteurs x et y de $\mathbb{R}^n$, distincts, de même norme, il existe une réflexion unique s_H d'hyperplan H qui échange x et y

Démonstration. Notons $H = (\mathbb{R}(x-y))^\perp$, on a $< x+y | x-y >= ||x||^2 - ||y||^2 = 0$ ce qui montre que $x + y \in H$, on a alors :
$s_H(x-y) = -x + y,\ s_H(x+y) = x + y \Longrightarrow s_H(x) = y.$

Soit $A \in GL_n(\mathbb{R}), A = (c_1, c_2, \cdots, c_n),\ c_i,\ i \in [\![1, n]\!]$ étant la i-ème colonne de A. La base canonique de $\mathbb{R}^n$ est notée $(e_1, e_2, \cdots e_n)$. En utilisant la proposition précédente, si c_1 et e_1 ne sont pas colinéaires, il existe une matrice H_1 de réflexion, telle que $H_1(c_1) = ||c_1|| e_1$, les termes de la première colonne de la matrice $H_1 A$ sont $||c_1|| > 0$ pour le premier, les autres étant nuls. On peut réitérer cette méthode pour la matrice A' constituée des lignes et des colonnes 2 à n de $H_1 A$, en réitérant cette méthode on obtient $H_n H_{n-1} \cdots H_1 A = R$ triangulaire supérieure, à coefficients diagonaux strictement positifs, certaines de ces matrices pouvant être I_n si les zéros recherchés existent déjà.

Proposition. Le vecteur $u \in \mathbb{R}^n$ étant normé, la matrice $H = I_n - 2u^t u$ est la matrice de la réflexion par rapport à $H = (\mathbb{R}u)^\perp$.
Cette matrice est symétrique et orthogonale.

Démonstration. On vérifie sans difficulté, que $Hu = -u$ et si $x \perp u$, $Hx = x$. Par ailleurs H est la matrice d'une symétrie orthogonale dans la base canonique qui est orthonormale, elle est donc symétrique et orthogonale.

Ainsi pour avoir $Ha = b$ où a et b sont deux vecteurs distincts de même norme, on choisit d'après l'étude précédente le vecteur $u = \dfrac{b-a}{||b-a||}$, pour avoir $Hc_1 = ||c_1||e_1$ on prendra $u = \dfrac{c_1 - ||c_1||e_1}{||c_1 - ||c_1||e_1||}$.

La matrice $Q = H_n \cdots H_1$ est orthogonale et $R = H_n \cdots H_1 A$ est triangulaire supérieure à coefficients diagonaux strictement positifs, on a $A = QR$ avec $Q = {}^t(H_n \cdots H_1) = H_1 \cdots H_n$ et $R = {}^tQ$.

Exemple.

$$A = \begin{pmatrix} -4 & 20 & 35 & 5 \\ -4 & -30 & -15 & 55 \\ -8 & 40 & -80 & -65 \\ 23 & -15 & 30 & 15 \end{pmatrix}$$

Détaillons les deux premières étapes :
Première étape :
$||c_1|| = 25,\ v_1 = c_1 - 25e_1 = {}^t(-29, -4, -8, 23),\ ||v_1||^2 = 1450 = 2.725$ d'où

$$H_1 = I_4 - \frac{1}{725}\begin{pmatrix} -29 \\ -4 \\ -8 \\ 23 \end{pmatrix}\begin{pmatrix} -29 & -4 & -8 & 23 \end{pmatrix} = \frac{1}{725}\begin{pmatrix} -116 & -116 & -232 & 667 \\ -116 & 709 & -32 & 92 \\ -232 & -32 & 661 & 184 \\ 667 & 92 & 184 & 196 \end{pmatrix}$$

On a

$$A_1 = H_1 A = \frac{25}{29}\begin{pmatrix} 29 & -29 & 58 & 29 \\ 0 & -42 & -15 & 167 \\ 0 & 32 & -88 & -69 \\ 0 & 24 & 21 & -1 \end{pmatrix}$$

Deuxième étape : $c_2 = \dfrac{25}{29}{}^t(-29, -42, 32, 24),\ w_2 = \dfrac{25}{29}{}^t(0, -42, 32, 24),\ ||w_2|| = 50, v_2 = w_2 - 50e_2 = \dfrac{100}{29}{}^t(0, -25, 8, 6)$ on simplifie en choisissant
$a_2 = {}^t(0, -25, 8, 6),\ ||a_2||^2 = 725$, d'où

$$H_2 = I_4 - \frac{2}{725}\begin{pmatrix} 0 \\ -25 \\ 8 \\ 6 \end{pmatrix}\begin{pmatrix} 0 & -25 & 8 & 6 \end{pmatrix} = \frac{1}{725}\begin{pmatrix} 725 & 0 & 0 & 0 \\ 0 & -525 & 400 & 300 \\ 0 & 400 & 597 & -96 \\ 0 & 300 & -96 & 653 \end{pmatrix}$$

Ona alors

$$A_2 = H_2 A_1 = \begin{pmatrix} 25 & -25 & 50 & 25 \\ 0 & 50 & -25 & -75 \\ 0 & 0 & -72 & -17 \\ 0 & 0 & 21 & 31 \end{pmatrix}$$

La troisième étape donne

$$H_3 = \frac{1}{25}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -24 & 7 \\ 0 & 0 & 7 & 24 \end{pmatrix},\ R = H_3 A_2 = \begin{pmatrix} 25 & -25 & 50 & 25 \\ 0 & 50 & -25 & -75 \\ 0 & 0 & 75 & 25 \\ 0 & 0 & 0 & 25 \end{pmatrix}$$

$$Q = H_1 H_2 H_3 = \frac{1}{25}\begin{pmatrix} -4 & 8 & 17 & 16 \\ -4 & -17 & -8 & 16 \\ -8 & 16 & -16 & 7 \\ 23 & 4 & -4 & 8 \end{pmatrix}$$

3.3 Applications

3.4 Systèmes linéaires

A l'aide de la factorisation $A = QR \in \mathcal{M}_{m,n}(\mathbb{R})$ il est possible de résoudre des systèmes linéaires lorsque $m > n$ en utilisant des relations vues ci-dessus : $AX = B \Longleftrightarrow QZ = B,\;$ et $RX = Z$, l'algorithme de résolution étant de même nature que dans le cas de la factorisation LU.

3.5 Détermination de valeurs propres

La détermination de valeurs propres en utilisant le polynôme caractéristique est numériquement instable, car une faible variation des coefficients de ce polynôme peut entrainer une erreur importante sur les racines. L'utilisation de la factorisation QR permet d'obtenir des valeurs précises de ces valeurs propres, avec néanmoins des hypothèses restrictives pour assurer la convergence de la méthode.

Principe de la méthode

On considère $A \in GL_n(\mathbb{R})$:
On définit $A_1 = A = Q_1R_1$ en utilisant la factorisation QR.
$A_2 = R_1Q_1 = Q_2R_2$ on a $A_2 = {}^tQ_1(Q_1R_1)Q_1$ les matrices A_2 et A_1 sont semblables. On définit par récurrence
$A_k = Q_kR_k,\; A_{k+1} = R_kQ_k = Q_{k+1}R_{k+1}$ on a de même A_k et A_{k+1} qui sont semblables. Ces différentes matrices ont donc les mêmes valeurs propres.

Théorème. Soit $A \in GL_n(\mathbb{R})$ diagonalisables et telle que ses valeurs propres vérifient :
$|\lambda_1| > |\lambda_2| > \cdots > |\lambda_n|$.
On suppose de plus que $A = PDP^{-1}$ avec D diagonale et P admettant une décomposition LU. On note $D_{ii} = \lambda_i$.
En utilisant les matrices définies ci-dessus on a $\lim\limits_{k\to\infty} (A_k)_{ii} = \lambda_i$ et $\lim\limits_{k\to\infty} (A_k)_{ij} = 0$ pour $i > j$.

Démonstration. En remarquant que $A_{k+1} = {}^tQ_kA_nQ_k$ on obtient $A_{k+1} = {}^t\Omega_kA_k\Omega_k$ avec $\Omega_k = Q_1Q_2\cdots Q_k$.
On a $A^k = (Q_1R_1)^k = Q_1\left(R_1Q_1\right)^{k-1}R_1 = Q_1A_2^{k-1}R_1$ ce qui entraine $A^k = (Q_1Q_2\cdots Q_k)\left(R_kR_{k-1}\cdots R_1\right)$ qui est la décomposition QR de A^k avec $Q = \Omega_k$.

Ecrivons $A^k = PD^kP^{-1}$ on peut écrire $P = Q'R',\; P^{-1} = LU$.
On a $A^k = Q'R'D^kLU$. Notons $R'D^kLD^{-k}R'^{-1} = \tilde{Q}_k\tilde{R}_k$ décomposition QR de cette matrice.

$A^k = Q'R'D^kLU = Q'(R'D^kLD^{-k}R'^{-1})R'D^kU = Q'\tilde{Q}_k\tilde{R}_kR'D^kU$.La matrice $Q'\tilde{Q}_k$ est orthogonale et la matrice $\tilde{R}_kR'D^kU$ est triangulaire supérieure, c'est presque la décomposition QR de A^k à ceci près que les coefficients diagonaux de la deuxième matrice peuvent ne pas être strctement positifs. Il existe une matrice diagonale Λ_k dont les termes diagonaux sont de valeurs absolues 1 telle que $A^k = \left[Q'\tilde{Q}_k\Lambda_k\right]\left[(\Lambda_k)^{-1}\tilde{R}_kR'D^kU\right]$, décomposition QR de A^k.

Par identification on a $\Omega_k = Q'\tilde{Q}_k\Lambda_k$.
On remarque que :

$$D^kLD^{-k} = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i < j \\ \dfrac{\lambda_i^k}{\lambda_j^k} & \text{si } i > j \end{cases}$$

Ce qui entraine que la matrice $R'D^kLD^{-k}R'^{-1}$ tend vers la matrice identité quand $k \to \infty$.

La norme euclidienne de $\tilde{Q_k}$ est constante, la suite $(\tilde{Q_k})$ admet donc une suite extraite convergente $\tilde{Q_{k_i}}$ ce qui montre que la suite

$(\tilde{R_{k_i}})$ converge également car $(\tilde{Q_{k_i}}\tilde{R_{k_i}}$ converge vers l'identité. $\tilde{Q}\tilde{R}$ est donc la décomposition QR de I_n, on a alors $\tilde{Q} = I_n$, la suite compacte $(\tilde{Q_k})$ n'a qu'une valeur d'adhérence, elle converge donc vers ctte valeur I_n.
On en déduit que la suite $(\Lambda_k\Omega_k)$ converge vers la matrice Q', facteur orthogonal de P.
$\Lambda_k A^{k+1}\Lambda_k \longrightarrow {}^tQ'AQ' = {}^tQ'(PDP^{-1})Q' = {}^tQ'(Q'R'DR'^{-1t}Q')Q' = R'DR'^{-1}$ cette dernière matrice étant triangulaire supérieure dont la diagonale contient les valeurs propres de A.