Université Bordeaux 1 **Algèbre et calcul formel – Agrégation**
Mathématiques Année 2021–2022

FEUILLE D'EXERCICES n° 7

Algèbre linéaire : décompositions de matrices définies sur un corps

1. Réduction

Nous avons déjà travaillé sur la diagonalisation des matrices. Sur un corps algébriquement clos, les matrices non diagonalisables peuvent-être triangulées, et même mises sous forme de Jordan. Pour cela, on dispose de la fonction `Jordan_form`.

On va construire un exemple de matrice non diagonalisable sur $\mathbb{C}$, dont la réduction est formée de plusieurs blocs de Jordan.

```
J1 = jordan_block(5,3)
J2 = jordan_block(1,4)
J3 = jordan_block(2,5)
J = block_diagonal_matrix(J1,J2,J3)
```

Cela doit donner une matrice sous forme de Jordan, dont on peut prendre un conjugué, pour construire un exemple non-trivial :

```
N1 = jordan_block(0,12)
N2 = transpose(N1)
D = diagonal_matrix([1]+[2 for i in range(11)])
P = N1+N2+D
A = P^(-1)*J*P
```

On s'est arrangé pour que $\det P = 1$. On peut aussi prendre une matrice au hasard, qui aura de grandes chances d'être inversible.

```
M12Q = MatrixSpace(QQ,12)
P = M12Q.random_element()
A = P^(-1)*J*P
```

Quelle est la décomposition de Dunford de la matrice A précédente ? Soit K un corps. On rappelle que la décomposition de Dunford d'une matrice A de $\mathcal{M}_n(K)$ est l'écriture de A sous la forme $A = D + N$ où $DN = ND$, où N est nilpotente et où D est diagonalisable dans une clôture algébrique de K. Je ne connais pas de commande directe sur `sage` pour cette décomposition. Nous verrons plus tard un algorithme qui la calcule. Par contre, il existe une commande qui calcule la forme de Jordan : `jordan_form`. Si l'on utilise l'option `transformation=true`, cette fonction `jordan_form` donne également une matrice de passage.

Exemple piège à faire sans calculs : quelle est la décomposition de Dunford de $\left(\begin{smallmatrix} 1 & 2 \\ 0 & 3 \end{smallmatrix}\right)$?

Remarque. On a construit ci-dessus des matrices diagonales par blocs en utilisant `block_diagonal_matrix`. Remarquons au passage l'existence de la fonction `block_matrix`, qui permet la construction de matrices par blocs.

2. Décomposition LU

Théorème 1. *Soient K un corps et $A \in \mathrm{GL}_n(K)$. Alors il existe une matrice de permutation P, une matrice triangulaire supérieure U et une matrice triangulaire inférieure L dont les éléments diagonaux sont égaux à 1 telles que $A = PLU$. De plus, pour P fixée, cette décomposition est unique si elle existe.*

Cette décomposition s'obtient grâce à l'algorithme du pivot de Gauss. Une décomposition $A = LU$ (c'est-à-dire où $P = I_n$) existe si et seulement si on ne rencontre pas de « pivot nul » dans cet algorithme. En fait, on montre que la décomposition $A = LU$ existe si et seulement si pour tout $k \in [[1,n]]$, $\det A_k \neq 0$, où A_k est la sous-matrice supérieure gauche de taille k de A. Par exemple, les matrices de $\mathrm{GL}_n(\mathbb{C})$ à *diagonale strictement dominante* admettent une décomposition $A = LU$, ainsi que les matrices réelles *symétriques définies positives* et les matrices complexes *hermitiennes définies positives*.

Sur Sage, la décomposition $A = PLU$ de A est donnée par `A.LU()`.

Soit $A = \begin{pmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 10 \end{pmatrix}$. Donner la décomposition LU de A (on pourra essayer la fonction `LU` sans option, et aussi avec l'option `pivot='nonzero'`). Soit $B = \begin{pmatrix} 1 \\ 3 \\ 5 \end{pmatrix}$. Pour entrer B, on peut utiliser `vector`. En utilisant cette décomposition LU, et en résolvant deux systèmes triangulaires sur papier, résoudre $AX = B$.

Définir la concaténée C de A et B, appliquer la fonction `LU` à C et retrouver la solution de l'équation par résolution d'une équation triangulaire.

Vérifier le résultat obtenu en utilisant `A.solve_right(B)`.

Même exercice avec $A = \begin{pmatrix} 1 & 4 & 7 \\ 2 & 8 & 8 \\ 3 & 6 & 10 \end{pmatrix}$ et $B = \begin{pmatrix} 1 \\ 3 \\ 5 \end{pmatrix}$.

Soit maintenant une matrice A qui n'est pas supposée inversible, ni même carrée. Soit $A \in \mathcal{M}_{m,n}(K)$, le résultat précédent s'étend à ce cas : il existe une décomposition $A = PLU$, où P est une matrice de permutation de taille m, L est une matrice triangulaire inférieure de $\mathrm{GL}_m(K)$ dont les coefficients diagonaux sont égaux à 1, et U est une matrice *échelonnée par lignes* de $\mathcal{M}_{m,n}(K)$. Il n'y a pas unicité en général, même si P est fixée.

Rappelons ce qu'est une matrice échelonnée par lignes.

Définition 2. Soit $E = (e_{ij})_{i,j} \in \mathcal{M}_{m,n}(K)$. On dit que E est *échelonnée par lignes* s'il existe un entier $r \leqslant \min(m,n)$ et une fonction $f\colon [[1,r]] \longrightarrow [[1,n]]$ strictement croissante telle que les propriétés suivantes sont vérifiées.

(1) Pour tout $i \in [[1,r]]$, $e_{if(i)} \neq 0$.

(2) Si $j < f(i)$, alors $e_{ij} = 0$.

(3) Si $i > r$, alors $e_{ij} = 0$.

Le rang de E est alors égal à r.

La matrice E est dite *échelonnée réduite par lignes* si de plus pour tout $i \in [[1, r]]$:

(1) $e_{if(i)} = 1$

(2) pour tout $k < i$, $e_{kf(i)} = 0$.

Dans sage, la matrice U donnée par `A.LU()` n'est pas échelonnée par lignes, mais seulement triangulaire supérieure. Par contre, la commande `A.echelon_form()` rend une matrice $E \in \mathcal{M}_{m,n}(K)$ échelonnée réduite par lignes telle qu'il existe $G \in \mathrm{GL}_m(K)$ vérifiant $GA = E$. On peut souhaiter récupérer la matrice G. La fonction `extended_echelon_form` donne E et G, sous la forme concaténée $(E|G)$.

Essayer la commande `LU` sur `A=matrix(QQ,[[1,2,3],[3,2,1],[1,1,1]])`. On voit sur sa décomposition LU que A n'est pas inversible. En utilisant cette décomposition, calculer $\operatorname{Im} A$, $\operatorname{Ker} A$, puis résoudre $AX = B$, où $B = {}^t(1, 2, 3)$. Essayer aussi avec $B = {}^t(2, 2, 1)$. Retrouver ces résultats en utilisant les commandes `column_space`, `kernel`, et `solve_right`.

Soit $A = \begin{pmatrix} 0 & 1 & 2 & 1 & 1 \\ 0 & 1 & 2 & 0 & 2 \\ 0 & 2 & 4 & 3 & 1 \\ 0 & -1 & -2 & 3 & 0 \end{pmatrix}$. En utilisant les fonctions citées ci-dessus, Calculer $\operatorname{Ker} A$ et $\operatorname{Im} A$, puis résoudre $AX = B$, où $B = {}^t(1, 1, 1, 1)$, puis où $B = {}^t(1, 1, 2, 0)$.

3. Décomposition QR

Théorème 1. *Soit $A \in \mathrm{GL}_n(\mathbb{R})$. Alors il existe une matrice triangulaire supérieure R dont les coefficients diagonaux sont strictement positifs et une matrice orthogonale Q telles que $A = QR$. De plus, cette décomposition est unique.*

Cette décomposition est utile pour la résolution du problème des "moindres carrés". Nous verrons cela plus tard.

Elle permet aussi de ramener la résolution d'une équation $AX = B$ à la résolution d'une équation triangulaire. En effet, si $A = QR$ est la décomposition QR de A (supposée inversible), alors l'équation $AX = B$ est équivalente à l'équation $RX = {}^tQB$. Mais surtout, cette décomposition est utile pour la résolution du problème des "moindres carrés".

Par exemple, on définit la matrice

```
A=matrix(ZZ,[[1,2,3],[1,-2,0],[-2,2,1 ]])
```

Si on essaie `A.QR`, ça ne marche pas. Dans l'aide à la décomposition QR, le corps de base utilisé est une clôture algébrique de $\mathbb{Q}$. Essayons.

```
M3Qbar = MatrixSpace(QQbar,3,3); A = M3Qbar(A); A.QR()
```

On doit trouver la décomposition QR de A (vérifier que QR est bien une approximation numérique de A).

Comme pour la décomposition LU, le théorème précédent s'étend au cas où A n'est pas inversible, et au cas où A n'est pas carrée. Soit donc $A \in \mathcal{M}_{m,n}(\mathbb{R})$.

Il existe une décomposition $A = QR$, où $Q \in \mathcal{O}_m(\mathbb{R})$ et où R est une matrice *échelonnée par lignes* de $\mathcal{M}_{m,n}(\mathbb{R})$. Il n'y a pas unicité en général.

En utilisant la décomposition QR, retrouver les résultats déjà obtenus ci-dessus (grâce à la décomposition LU de A), pour $A = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \\ 1 & 1 & 1 \end{pmatrix}$. Il ne s'agit pas de faire les calculs à la main : on peut utiliser les commandes `right_kernel` et `solve_right` sur des matrices triangulaires.

4. Décomposition QR et moindres carrés

On cherche à trouver un vecteur X de taille n minimisant l'expression $||AX - B||_2$, où A est une matrice de type $m \times n$ et B un vecteur de taille m, les coefficients de ces matrices étant réels. On suppose que $m > n$ et que le rang de A est égal à n.

Un exemple de problème : l'approximation d'une fonction. Soit y une fonction de $\mathbb{R}$ dans $\mathbb{R}$ qui dépend linéairement de n paramètres x_i :

$$y = \sum_{j=1}^{n} x_j f_j$$

où les fonctions f_j sont linéairement indépendantes. Les fonctions f_j sont fixées. On cherche y, d'après la connaissance de valeurs mesurées $y_i = y(t_i)$, pour $i \in \{1, \ldots, m\}$. On veut donc trouver les valeurs des x_i. En général, par exemple à cause des incertitudes de mesure, il n'existe pas n valeurs x_j qui vérifient simultanément les m équations

$$y_i = \sum_{j=1}^{n} x_j f_j(t_i).$$

On cherche donc des x_j en cherchant à satisfaire au mieux les équations. Par exemple, on peut chercher des x_j minimisant la somme des carrés des écarts entre y_i et $\sum_{j=1}^{n} x_j f_j(t_i)$, c'est-à-dire

$$E(x_1, \ldots, x_2) = \sum_{i=1}^{n} \left(\sum_{j=1}^{n} x_j f_j(t_i) - y_i \right)^2.$$

On appelle solution au sens des moindres carrés un vecteur qui minimise E. C'est donc le problème annoncé en introduction, avec

$$A = \begin{pmatrix} f_1(t_1) & \ldots & f_n(t_1) \\ \vdots & & \vdots \\ f_1(t_m) & \ldots & f_n(t_m) \end{pmatrix} \qquad \text{et} \qquad B = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}.$$

Existence et unicité. Un vecteur colonne X réalise le minimum de $E(X) = ||AX - B||_2^2$ si et seulement si ${}^tAAX = {}^tAB$. Pour le voir, on peut utiliser le fait que ce minimum est réalisé si et seulement si AX est égal à la projection orthogonale de B sur l'espace vectoriel engendré par les colonnes de A.

Il suffit donc de résoudre le système ${}^tAAX = {}^tAB$. Comme on a supposé A de rang n, ce qui arrive très souvent, la matrice ${}^tAA$ appartient à $\mathrm{GL}_n(\mathbb{R})$, et ce système admet une solution unique. Remarquons qu'on ne peut simplifier cette équation par ${}^tA$, puisque cette matrice n'est pas inversible.

Mise en œuvre. Pour résoudre le problème, on peut donc calculer ${}^tAA$, puis résoudre ${}^tAAX = {}^tAB$ à l'aide de la méthode du pivot de Gauss. Cette méthode peut poser des problèmes d'instabilité numérique. Par exemple, considérons la matrice

$$A = \begin{pmatrix} 1 & 1 & 1 \\ \varepsilon & 0 & 0 \\ 0 & \varepsilon & 0 \\ 0 & 0 & \varepsilon \end{pmatrix}.$$

Alors si ε est trop petit, ${}^tAA$ sera arrondie en une matrice de rang 1. C'est le cas si $\varepsilon = 10^{-7}$, avec la précision par défaut de sage.

L'utilisation de la décomposition QR de A conduit à une méthode plus stable. On peut écrire $A = QR$, où Q est une matrice de $\mathcal{O}_m(\mathbb{R})$ et où R est de la forme

$$R = \begin{pmatrix} R' \\ 0 \end{pmatrix}$$

R' étant une matrice triangulaire supérieure inversible de taille n et 0 la matrice nulle de type $(m-n) \times n$. Cette décomposition est donnée par la commande `QR`.

Soit $C = {}^tQB = {}^t(\gamma_1, \ldots, \gamma_m)$. Soient $C_1 = (\gamma_1, \ldots, \gamma_n)$ et $C_2 = (\gamma_{n+1}, \ldots, \gamma_m)$. Alors

$$||AX - B||_2^2 = ||R'X - C_1||_2^2 + ||C_2||_2^2.$$

Donc X réalise le minimum de $||AX - B||_2$ si et seulement si $R'X = C_1$.

Remarquons qu'il est inutile de calculer Q. Le vecteur C peut en effet se calculer pas à pas en même temps que l'on calcule R. En fait, il suffit de concaténer A et B avant d'appliquer la commande `QR`.

Exemple 1. Essayer ces deux méthodes avec les données

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 10^{-7} & 0 & 0 \\ 0 & 10^{-7} & 0 \\ 0 & 0 & 10^{-7} \end{pmatrix} \qquad \text{et} \qquad B = \begin{pmatrix} 0.01 \\ 0.51 \\ 0.49 \\ -0.98 \end{pmatrix}.$$

Pour la première méthode, si l'on écrit la matrice sur le corps `RR`, la matrice ${}^tAA$ est de rang 1. Il faut utiliser une précision plus grande pour pouvoir espérer trouver la solution.

Exemple 2. Trouver par les deux méthodes la droite qui approche le plus les points $(1, 11)$, $(2, 11)$, $(3, 16)$, $(4, 19)$ et $(5, 21)$ au sens des moindres carrés, puis illustrer la solution sur un graphe.

On peut utiliser les commandes suivantes (où $ax + b$ est la droite trouvée).

```
pts=points([[1,11],[2,11],[3,16],[4,19],[5,21]],color='red')
dte=plot(a*x+b,(1,5))
pts+dte
```

Exemple 3. Trouver le polynôme de degré inférieur ou égal à 3 qui approche le plus les points $(1,4)$, $(2,21)$, $(3,56)$, $(4,120)$ et $(5,211)$ au sens des moindres carrés. Fraire un graphe.

5. Décomposition en valeurs singulières

Soit $A \in M_{m,n}(\mathbb{R})$. Soit $S = {}^tAA \in M_n(\mathbb{R})$. La matrice Q est symétrique donc il existe $V \in \mathcal{O}_n(\mathbb{R})$ telle que ${}^tVSV = \mathrm{Diag}(\lambda_1, \dots, \lambda_n)$. Comme A est définie positive, $\lambda_i \geqslant 0$ pour tout i. L'égalité ci-dessus s'écrit aussi

$$^t(AV)(AV) = \mathrm{Diag}(\lambda_1, \dots, \lambda_n).$$

Si l'on pose $AV = (Y_1|\dots|Y_n)$ alors ${}^tY_iY_j = 0$ si $i \neq j$ et ${}^tY_iY_i = \lambda_i \geqslant 0$.

Définition 1. On appelle valeurs singulières de A les racines carrées des valeurs propres de ${}^tAA$.

On ordonne les valeurs propres λ_i (et donc aussi les colonnes de V) de telle sorte que $\lambda_i > 0$ si $i \leqslant r$ et λ_i si $i > r$. On écrit $V = (V_1|\dots|V_n)$ (V_i est vecteur propre associé à la valeur propre λ_i). On note $\sqrt{\mu_i}$ pour tout $i \in [[1,n]]$. Pour tout j, le vecteur $Y_j = AV_j$ est la j-ème colonne de AV. On pose $U_j = Y_j/\mu_j$ pour $i \in [[1,r]]$. Alors $(U_1, \dots, U_r)$ est une famille orthonormée de $\mathbb{R}^m$, que l'on complète en une base orthonormée $(U_1, \dots, U_m)$. La matrice $U = (U_1|\dots|U_m)$ appartient à $\mathcal{O}_m(\mathbb{R})$ et pour tout j, $Y_j = \mu_j U_j$.

Soit $D = {}^tUAV$. Alors on peut écrire la matrice D par blocs $D = \begin{pmatrix} D_0 & 0 \\ 0 & 0 \end{pmatrix} \in M_{m,n}(\mathbb{R})$ où $D_0 = \mathrm{Diag}(\mu_1, \dots, \mu_r)$.

Théorème 2. *Soit $A \in M_{m,n}(\mathbb{R})$ et soient $\mu_1, \dots, \mu_r$ les valeurs singulières non nulles de A. Il existe $U \in \mathcal{O}_m(\mathbb{R})$ et $V \in \mathcal{O}_n(\mathbb{R})$ telles que ${}^tUAV = D$ où $D = \begin{pmatrix} D_0 & 0 \\ 0 & 0 \end{pmatrix}$ et où $D_0 = Diag(\mu_1, \dots, \mu_r)$. On appelle la décomposition $A = UD^tV$ une décomposition en valeurs singulières (ou SVD) de A.*

Remarque. L'ensemble des valeurs singulières est uniquement déterminé, mais il n'y a pas unicité de la décomposition en valeurs singulières.

Théorème 3. *Soient $A \in M_{m,n}(\mathbb{R})$ et $B \in M_{n,1}(\mathbb{R})$. Soit $A = UD^tV$ une SVD de A. On note*

$$\delta = \min\left\{||AX - B||_2^2 \ : \ X \in M_{n,1}(\mathbb{R})\right\}$$

et

$$\mathcal{S} = \left\{X \in M_{n,1}(\mathbb{R}) \ : \ ||AX - B||_2^2 = \delta\right\}.$$

Alors
$$X_0 = \sum_{i=1}^{r} \frac{{}^tU_iB}{\mu_i} V_i$$
est l'élément de $\mathcal{S}$ de plus petite norme et $\mathcal{S}$ est l'espace affine
$$\mathcal{S} = X_0 + \sum_{i=r+1}^{n} \mathbb{R}V_i.$$
Enfin,
$$\delta = \sum_{i=r+1}^{m} ({}^tU_ib)^2 \quad et \quad ||X_0||_2^2 = \sum_{i=1}^{r} \left(\frac{{}^tU_iB}{\mu_i}\right)^2.$$

Exemple. Soient
$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 5 & -6 & -1 \\ 7 & -1 & 6 \\ 4 & -2 & 2 \end{pmatrix} \quad \text{et} \quad B = \begin{pmatrix} 5 \\ 11 \\ -2 \\ 11 \\ 4 \end{pmatrix}$$
En utilisant la fonction `SVD`, trouver le X_0 du théorème 3 correspondant.

6. Pseudo-inverse

Dans la suite, on considère une SVD de A fixée $A = UD{}^tV$

Proposition 1. *(1)* $A = \sum_{k=1}^{r} \mu_k U_k V_k$
(2) $\operatorname{Ker} A = Vect\{V_j \ : \ r+1 \leqslant j \leqslant n\}$
(3) $\operatorname{Im} A = Vect\{U_j \ : \ 1 \leqslant j \leqslant r\}$
(4) $\operatorname{rg} A = r$

Corollaire 2. *(1)* $S = {}^tAA = \sum_{k=1}^{r} \mu_k V_k {}^tV_k$
(2) $\operatorname{Ker} S = \operatorname{Ker} A = Vect\{V_j \ : \ r+1 \leqslant j \leqslant n\}$
(3) $\operatorname{Im} S = \operatorname{Im} {}^tA = Vect\{V_j \ : \ 1 \leqslant j \leqslant r\}$

Définition 3. Soit $D^\dagger$ la matrice de $M_{n,m}$ définie par blocs par $D^\dagger = \begin{pmatrix} D_0^{-1} & 0 \\ 0 & 0 \end{pmatrix}$ (remarquer que la taille est (n, m) alors que celle de D est (m, n)). On appelle pseudo-inverse de A la matrice
$$A^\dagger = VD^{\dagger t}U.$$

Exercice. Si $A \in \mathrm{GL}_n(\mathbb{R})$, alors $A^\dagger = A^{-1}$.

La pseudo-inverse semble dépendre de la SVD de A choisie. Il n'en est rien. La proposition suivante prouve l'unicité de $A^\dagger$.

Proposition 4. *Soit $B \in M_{m,n}(\mathbb{R})$. Alors $B = A^\dagger$ si et seulement si les propriétés suivantes sont vérifiées.*

(1) BA *et* AB *sont symétriques.*

(2) $ABA = A$ *et* $BAB = B$

Proposition 5. *(1)* $A^\dagger = \sum_{k=1}^{r} \dfrac{1}{\mu_k} V_k {}^t U_k$

(2) $AA^\dagger = \sum_{k=1}^{r} U_k {}^t U_k$. *C'est la projection orthogonale sur* $\operatorname{Im} A$.

(3) $A^\dagger A = \sum_{k=1}^{r} V_k {}^t V_k$. *C'est la projection orthogonale sur* $(\operatorname{Ker} A)^\perp$.

Théorème 6. *Soit $A \in M_{m,n}(\mathbb{R})$. Soient $B \in M_{m,1}(\mathbb{R})$. La solution X_0 de norme minimale du problème des moindres carrés associé à A et B est $X_0 = A^\dagger B$.*