Le déterminant expliqué à mon petit-fils

PAR PATRICK TELLER

Peut-on définir le déterminant et établir ses propriétés nécessaires en évitant le recours aux formes multilinéaires antisymétriques ?

L'objectif de ce travail est de définir le déterminant d'une matrice carrée au moyen d'une factorisation adaptée de la décomposition de Bruhat M=UPV.

On décrira un algorithme qui détermine la matrice P en un maximum de $\frac{n^3-n}{3}$ multiplications.

1 Le cahier des charges

Nous cherchons une fonction $d\colon \mathcal{M}_n(K) \longrightarrow K$ telle que

i) $\rho(\mathrm{I})=1$

ii) $\forall M, j, k, d((M_1, ..., M_{j-1}, \mathrm{kM}_j, M_{j+1}, ..., M_n)) = \mathrm{kd}(M)$

iii) $\forall M, k, i, j,\ d(M(I + \mathrm{kE_{ij}})) = d(M)$

iv) $\forall M, S$(matrice de permutation), $d(\mathrm{MS}) = \mathrm{sgn}(S) d(M)$

2 Le déterminant d'une matrice carrée

Définition 1. *Matrices principales*

Une matrice $P \in \mathcal{GL}_n(K)$ sera dite matrice principale lorsqu'elle possède un terme non nul et un seul par ligne et par colonne; leur ensemble sera noté $\mathcal{P}_n(K)$.

Exemple 2. $\begin{pmatrix} 0 & 0 & 3 & 0 \\ 4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 \\ 0 & -9 & 0 & 0 \end{pmatrix}$

Définition 3. *Le poids d'une matrice principale*

Soit une matrice principale P et la permutation φ telle que $\forall j, p_{\mathrm{ij}} \neq 0 \Longleftrightarrow i = \varphi(j)$; on appellera poids de P le réel sgn(φ)$\times \prod_{j=1}^n p_{\varphi(j)j}$.

Définition 4. *On désignera par $\mathcal{U}_n$, $(\mathrm{resp}\mathcal{L}_n)$, les ensembles de matrices triangulaires supérieures (resp triangulaires inférieures) dont la diagonale ne comporte que des « 1 » (on dira « unipotentes »).*

Théorème 5. *D'après la Décomposition de Bruhat*

Quelle que soit la matrice inversible M il existe un triplet $(P, U, V) \in \mathcal{P}_n \times \mathcal{U}_n^2$ telles que $M = UPV$ et P est unique

Démonstration.

Existence:

On désignera par $(E_1, \ldots, E_n)$ *la base canonique de* K^n

Soit $M_1 = \sum \lambda_{k1} E_k$ *on pose* $\varphi(1) = \max\{k, \lambda_{k1} \neq 0\}$, $E'_1 = \sum \lambda_{k1} E_k$ *et* $a_{\varphi(1)1} = \lambda_{\varphi(1)1}$

$M_2 = \sum_{k \neq \varphi(1)} \lambda_{k2} E_k + \mathrm{Vect}(E'_1)$ *on pose* $\varphi(2) = \max\{k, \lambda_{k2} \neq 0\}$, $E'_2 = \sum \lambda_{k2} E_k$ *et* $a_{\varphi(2)2} = \lambda_{\varphi(2)2}$

.... de manière générale pour chaque j $M_j = \sum_{k \notin \varphi([\![1, \ldots, j-1]\!])} \lambda_{\mathrm{kj}} E_k + \mathrm{Vect}(E'_1, \ldots, E'_{j-1})$ *et* $\varphi(j) = \max\{k, \lambda_{\mathrm{kj}} \neq 0\}$, $E'_j = \sum_{k \notin \varphi([\![1, \ldots, j-1]\!])} \lambda_{\mathrm{kj}} E_k$ *et* $a_{\varphi(j)j} = \lambda_{\varphi(j)j}$

On résumera cela ainsi : pour chaque j *on pose* $\varphi(j) = \min\{i, M_j = a_{\mathrm{ij}} E_i \bmod (M_1, \ldots, M_{j-1}, E_1, \ldots, E_{i-1})\}$.

Ceci définit une matrice principale $P = (p_{\mathrm{ij}})$, *où* $p_{\mathrm{ij}} = \delta_{i\varphi(j)} a_{\mathrm{ij}}$ *et deux matrices triangulaires supérieures U et V unipotentes telles que M=UPV.*

Unicité de P:

$M = UPV \iff \mathrm{MV}^{-1} = \mathrm{UP}$

Supposons UPV=U'P'V', φ *et* φ' *associees a P et P', d'où* $U^{-1}U'P' = P\mathbb{V}V'^{-1}$ *, posons* $W = U^{-1}U'$ *et* $Z = \mathbb{V}V'^{-1}$, *W et Z appartiennent à* $\mathcal{U}_n$.

Supposons donc que WP'=PZ, dans WP' les termes qui vérifient $i > \varphi'(j)$ *sont nuls et dans PZ les termes qui vérifient* $i = \varphi(j)$ *ne le sont pas, donc* $\forall j, \varphi(j) \leqslant \varphi'(j)$ *et comme il s'agit de permutations elles sont égales et, en utilisant le caractère unipotent de W et Z, on en déduit P=P'* □

Exemple 6. $M = \begin{pmatrix} 1 & 3 & -1 \\ -2 & -2 & 1 \\ 1 & 1 & 1 \end{pmatrix}$

$M_1 = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}, M_2 = \begin{pmatrix} 3 \\ -2 \\ 1 \end{pmatrix}, M_3 = \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}$

a. $M_1 = E_3 - 2E_2 + E_1$

b. $M_2 = 2E_1 + M_1$

c. $M_3 = 2M_1 - M_2 + 3E_2$; d'où $(\, M_1 \;\; M_2 \;\; M_3 \,) \begin{pmatrix} 1 & -1 & -2 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 0 \\ -2 & 0 & 3 \\ 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 2 & 0 \\ 0 & 0 & 3 \\ 1 & 0 & 0 \end{pmatrix}$, c'est à dire $M = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 2 & 0 \\ 0 & 0 & 3 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & -1 & -2 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 2 & 0 \\ 0 & 0 & 3 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}$.

Ici $U = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix}, V = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}, P = \begin{pmatrix} 0 & 2 & 0 \\ 0 & 0 & 3 \\ 1 & 0 & 0 \end{pmatrix}$

Remarque 7. le couple (U,V) n'est pas forcément unique

le c. pourrait tout aussi bien s'écrire $M_3 = M_1 + 3E_2 - 2E_1$; d'où $(\, M_1 \;\; M_2 \;\; M_3 \,) \begin{pmatrix} 1 & -1 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & -2 \\ -2 & 0 & 3 \\ 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & -2/3 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 2 & 0 \\ 0 & 0 & 3 \\ 1 & 0 & 0 \end{pmatrix}$, c'est à dire $M = \begin{pmatrix} 1 & -2/3 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 2 & 0 \\ 0 & 0 & 3 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & -1 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & -2/3 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 2 & 0 \\ 0 & 0 & 3 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

Ici $U = \begin{pmatrix} 1 & -2/3 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix}$, $V = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, $P = \begin{pmatrix} 0 & 2 & 0 \\ 0 & 0 & 3 \\ 1 & 0 & 0 \end{pmatrix}$

Définition 8. *Le poids d'une matrice*

Soient $M \in \mathcal{GL}_n(K)$ *et U,P,V comme au-dessus alors on posera* $\rho(M) = \rho(P)$

et si rang(M)<n on posera $\rho(M)$=0

Lemme 9. *L'effet de la transposition de deux colonnes consécutives de M*

Soient les colonnes M_j et M_{j+1} et M' la matrice obtenue à partir de M en échangeant les colonnes M_j et M_{j+1}; on désignera par φ' et P' la permutation et la matrice principale associées à M'.

1er cas $\varphi(j) < \varphi(j+1)$ les colonnes M_j et M_{j+1} sont comme suit $\begin{pmatrix} & & & \ldots\ldots & \ldots. \\ 0 & \ldots & 0 & m_{\varphi(j)j} & ? \\ & & & 0 & ? \\ 0 & \ldots & 0 & \ldots. & m_{\varphi(j+1)j+1} \\ & & & \ldots. & 0 \\ & & & \ldots & \ldots \\ & & & 0 & 0 \end{pmatrix}$

d'où $\varphi'(j) = j+1$ *et* $\varphi'(j+1) = j$ *et* $p'_{\varphi'(j)j} p'_{\varphi'(j+1)j+1} = p_{\varphi(j)j} p_{\varphi(j+1)j+1}$

2nd cas $\varphi(j) > (j+1)$ *les colonnes M_j et M_{j+1} sont comme suit* $\begin{pmatrix} 0 & \ldots & 0 & 0 & m_{\varphi(j+1)j+1} \\ & & & ? & * \\ & & & ? & * \\ 0 & \ldots & 0 & m_{\varphi(j)j} & m_{\varphi(j)j+1} \\ & & & \ldots. & 0 \\ & & & \ldots & \ldots \\ & & & 0 & 0 \end{pmatrix}$

*Le * dans une ligne d'indice $k \in]\!]\varphi(j+1), \varphi(j)[\![$ signifiant soit 0, soit un scalaire quelconque s'il existe un indice p<j tel que $\varphi(p) = k$.*

- si $m_{\varphi(j)j+1}$=0 les choses sont simples: $\varphi'(j) = j+1$ et $\varphi'(j+1) = j$ et $p'_{\varphi'(j)j} p'_{\varphi'(j+1)j+1} = p_{\varphi(j)j} p_{\varphi(j+1)j+1}$

- si $m_{\varphi(j)j+1} \neq 0$ $\varphi'(j) = j$ et $M'_{j+1} = \begin{pmatrix} 0 \\ ? \\ ? \\ m_{\varphi(j)j} \\ \ldots. \\ \ldots \\ 0 \end{pmatrix} = \frac{m_{\varphi(j)j}}{m_{\varphi(j)j+1}} M'_j - \frac{m_{\varphi(j+1)j+1}}{m_{\varphi(j)j+1}} m_{\varphi(j)j} E_{\varphi(j+1)} [M_1, \ldots, M_{j-1}, E_1, \ldots., E_{\varphi(j+1)-1}]$ *et donc* $\varphi'(j) = j$ *et* $\varphi'(j+1) = j+1$, $p'_{\varphi'(j)j} = m_{\varphi(j)j+1}$, $p'_{\varphi'(j+1)j+1} = -\frac{m_{\varphi(j+1)j+1}}{m_{\varphi(j)j+1}} m_{\varphi(j)j}$ *et donc* $p'_{\varphi'(j)j} p'_{\varphi'(j+1)j+1} = p_{\varphi(j)j} p_{\varphi(j+1)j+1}$.

Dans tous les cas le poids de P a été multiplié par -1.

Théorème 10. *Propriétés du poids*

i) $\forall M, \rho(M) = 0 \Longleftrightarrow \operatorname{rang}(M) < n$

ii) $\rho(I)$=1

iii) $\forall M, j, k, \rho((M_1, \ldots, M_{j-1}, \mathrm{kM}_j, M_{j+1}, \ldots, M_n)) = k\rho(M)$

iv) $\forall M, k, i, j,\ \rho(M(I + \mathrm{kE_{ij}})) = \rho(M)$

v) $\forall M, S$*(matrice de permutation),* $\rho(\mathrm{MS}) = \operatorname{sgn}(S)\rho(M)$

Démonstration. □

i), ii), iii) immédiats.

iv) Soient M et M'=M(I+kE_{ij})

si i<j la matrice I+kE_{ij} est triangulaire supérieure, donc M(I+kE_{ij})=UPV(I+kE_{ij}) => ρ(M(I+kE_{ij}))=$\rho(P) = \rho(M)$

si i>j M'=(M_1,.....,M_j,...,M'$_i$=M_i+kM_j,M_{i+1},..,M_n)

$\forall k < i, \varphi_{M'}(k) = \varphi_M(k)$

*M'$_i$=M_i+kM_j=M_i*mod (Vect(M'_1, ..., M'_{i-1}, E_1, ..., $E_{\varphi(i)-1}$))=$M_i$mod (Vect(M_1, ..., M_{i-1}, E_1, ..., $E_{\varphi(i)-1}$) *donc* $\varphi_{M'}(i) = \varphi_M(i)$

et alors $\forall k > i$ $\varphi_{M'}(k) = \varphi_M(k)$

v) si S est une matrice de permutation on la décompose en un produit de transpositions (j,j+1) et on applique le résultat du lemme.

Conclusion: ρ c'est le déterminant

3 Les trois grandes propriétés

Théorème 11. $\forall A$, *ρ(A)=0*$\Longleftrightarrow$rang$(A) < n$

découle de la définition de ρ.

Théorème 12. $\forall (A, B)$, $\rho(\mathrm{AB}) = \rho(A)\rho(B)$

Démonstration.

découle du Théorème 7 (iii,iv,v) appliqué à M$\longmapsto \frac{\rho(\mathrm{AM})}{\rho(A)}$

$\square$

Remarque 13.

Si nous avions défini la matrice Q par les conditions

Pour chaque j on pose $\psi(j) = \max \{i, M_j = b_{\mathrm{ij}} E_i \mathrm{mod} \mathrm{Vect}(M_{j+1}, ..., M_n, E_{i+1}, ..., E_n)\}$, ce qui définit une matrice principale Q=(q_{ij}), où q$_{\mathrm{ij}}$=$\delta_{i\varphi(j)} b_{\mathrm{ij}}$ et deux matrices triangulaires supérieures L et L' unipotentes telles que M=LQL'.

On pourrait démontrer que la fonction M$\longmapsto \rho(Q)$ possède les propriétés du Théorème 7.

Théorème 14.

Avec les notations au-dessus $\rho(P) = \rho(Q)$

Démonstration.

Soit M=LQL', alors JMJ=JLJ.JQJ.JL'J, où JLJ et JL'J appartiennent à $\mathcal{U}_n$ *et JQJ est principale, donc* $\rho(\mathrm{JMJ}) = \rho(\mathrm{JQJ}) = \rho(Q)$ *par définition mais aussi* $\rho(\mathrm{JMJ}) = \rho(J)\rho(M)\rho(J)$, *d'après le Théorème précédent, ce qui donne* $\rho(M)$*; en fin de compte* $\rho(P) = \rho(Q)$ $\square$

Théorème 15. *Le déterminant de la transposée*

Soit A=UPV, alors ${}^tA = {}^tV{}^tP{}^tU$ *d'où, d'après le Théorème précédent,* $\rho({}^tA)$=$\rho({}^tP)$=$\rho(P) = \rho(A)$

4 Un algorithme de factorisation

Une fois établie l'unicité de P, tout algorithme qui fournit P dans une factorisation M=UPV conviendra.

Pour cela il nous faut décrire trois procédures: une qui repère les positions des termes non nuls de P, deux qui « nettoyent » les colonnes pas à pas.

Définition 16. *Procédure Pivot(j) appliquée à la matrice A*

si j=1 $\varphi(j){:=}\max\{i, a_{\mathrm{ij}}{\neq}0\}$

si j>1 $\varphi(j){:=}\max\{i{\notin}\varphi([\![1,....,j-1]\!]), a_{\mathrm{ij}}{\neq}0\}$

Définition 17. *Procédure Nettoyage Vertical(j) appliquée à la matrice A*

$A{:=}\prod_{i<\varphi(j), a_{\mathrm{ij}}\neq 0, i\notin\varphi([\![1,...,j-1]\!])}\left(I-\frac{a_{\mathrm{ij}}}{a_{\varphi(j)j}}E_{i\varphi(j)}\right)A$

Définition 18. *Procédure Nettoyage Horizont(j) appliquée à la matrice A*

$A{:=}A\prod_{k>j, a_{\varphi(j)k}\neq 0}\left(I-\frac{a_{\varphi(j)k}}{a_{\varphi(j)j}}E_{\mathrm{jk}}\right)$

L'algorithme:

Pivot(1)

j:2

tant que j<n

Pivot(j)

Nettoyage Vertical(j)

Nettoyage Horizontal(j)

j:j+1

fin.

Complexité de la recherche de P:

Les opérations évoquées au-dessus reviennent en principe à ajouter à une ligne ou une colonne un multiple d'une autre; en fait comme Nettoyage Horizontal(j) n'est appelée que lorsque la colonne C_j ne comporte plus qu'un terme nul il vaut mieux le considérer comme l'opération suivante

pour tout k>j, $a_{\varphi(j)k}$:=0

Quant à Nettoyage Vertical(j) il fait appel au plus à $\varphi(j-1)$ multiplications par des matrices de transvection, comme dans les colonnes C_1,...,C_{j-1} les termes de la ligne $L_{\varphi(j)}$ sont nuls, Nettoyage Vertical(j) se traduit simplement par les affectations suivantes: $\forall i<\varphi(j)$, $\forall k>j$, $a_{\mathrm{ik}}:=a_{\mathrm{ik}}-\frac{a_{\mathrm{ij}}}{a_{\varphi(j)j}}a_{\varphi(j)k}$ et $\forall i<\varphi(j), a_{\mathrm{ij}}:=0$.

D'où Nettoyage Vertical(j) coûte 1 division et 2*(n-j)*($\varphi(j)-1\Big)$multiplications de scalaires , par suite la détermination de P exige n divisions et (au plus) $2\sum_{j=1}^{n-1}(n-j)^2=\frac{(n-1)n(n+1)}{3}$ donc l'algorithme réclame n divisions et (au plus) $\frac{n^3-n}{3}$ multiplications.

Conclusion : On peut définir le déterminant de manière plus économique aussi bien en termes d'outils mathématiques qu'en complexité: rien de ce qui a été employé ne déborde du programme de L1.

Paris, octobre 2017