Mineurs de la co-matrice

Jean-François Burnol, 10 octobre 2018; v2 12 octobre

Soit $\mathbb{K}$ un corps. Soit A une matrice de taille $M \times N$ à coefficients dans $\mathbb{K}$. Nous numérotons les lignes et colonnes à partir de 1.

Un multi-indice (pour les lignes) est une partie I de $\{1, \ldots, M\}$. On appelle *longueur* de I sa cardinalité et on la note $\ell(I)$ ou $\#I$. De même pour un multi-indice $J \subset \{1, \ldots, N\}$ pour les colonnes. On appelle *poids* de I la quantité $\sum_{i \in I} i$ et on la note $|I|$. [1] [2] Le complémentaire de I est noté I^c.

Pour I et J non vides, la *sous-matrice* (ou matrice extraite) de multi-indices I (pour les lignes) et J (pour les colonnes) est la matrice A_{IJ} de taille $\ell(I) \times \ell(J)$ dont les coefficients c_{pq} sont donnés par la formule $c_{pq} = a_{i_p j_q}$ avec $I = \{i_1 < i_2 < \cdots < i_k\}$, $k = \ell(I)$, $J = \{j_1 < j_2 < \cdots < j_l\}$, $l = \ell(J)$.

Pour $\#I = \#J$, le *mineur* $|A|_{IJ}$ est le déterminant de A_{IJ}. Attention qu'on note souvent A_{IJ} le mineur, mais j'ai besoin d'une notation aussi pour les sous-matrices. Et à propos comme il est usuel avec les indices de matrices on ne met pas de virgule entre les indices, sauf en cas de confusion possible.

Par convention $|A|_{\emptyset\emptyset} = 1$.

1 Mineurs et rang

L'ordre maximal possible pour un mineur est $\min(M, N)$ si A possède M lignes et N colonnes. L'ordre maximal pour un mineur non nul est un entier important associé à la matrice A.

Théorème 1 (et définition)**.** *Soit A une matrice de taille $M \times N$ à coefficients dans $\mathbb{K}$. Les trois entiers :*
— *la dimension n de l'espace engendré par les colonnes,*
— *la dimension n' de l'espace engendré par les lignes,*
— *l'entier $0 \leqslant k \leqslant \min(M, N)$ tel qu'il existe un mineur non nul d'ordre k et aucun mineur non nul d'ordre strictement supérieur,*
sont identiques. Cet entier commun s'appelle, par définition, le rang *de la matrice A.*

Pour la définition de k on rappelle qu'au pire, c'est-à-dire si A est nulle, elle possède en tout cas le mineur $|A|_{\emptyset\emptyset} = 1$ non nul.

Preuve. Si A est la matrice nulle, il n'y a pas grand chose à faire. Sinon, soit n le nombre maximal de colonnes linéairement indépendantes. Sans perte de généralité, on supposera que ce sont les n premières C_1, …, $C_n \in \mathbb{K}^M$ (c'est juste pour noter plus facilement les indices ; bien sûr les C_i sont des colonnes *distinctes* de A). Au passage on note que $n \leqslant \dim \mathbb{K}^M = M$. Par le théorème de la base incomplète on peut choisir parmi la base canonique des vecteurs $E_{i_k} \in \mathbb{K}^M$, $1 \leqslant k \leqslant M - n$ pour obtenir une base de $\mathbb{K}^M$ formée de C_1, …, C_n, et des E_{i_k}. Considérons la matrice $M \times M$ avec ces M colonnes. Par permutation de lignes on la transforme de manière à avoir en bas à droite la matrice identité de taille $M - n$. Le déterminant est non nul et il est au signe près un mineur de taille $n \times n$ de la matrice d'origine de taille $M \times N$.

Donc il existe un mineur $n \times n$ non-nul.

Toute sous-matrice carrée de A de taille $> n$ aura ses colonnes linéairement dépendantes (si $n = N$, il n'existe pas $n + 1$ colonnes distinctes de A et donc aucun mineur d'ordre $> n$) et donc tout mineur d'ordre $> n$ (s'il en existe…) sera nul. Ainsi $k = n$.

Cet entier k reste invariant par passage à la transposée et est ainsi égal aussi au nombre maximal de colonnes indépendantes dans la transposée, qui est le nombre maximal n' de lignes indépendantes dans la matrice A. □

1. C'est un peu ennuyeux car souvent $|X|$ est la notation pour la cardinalité d'un ensemble X.
2. La quantité $\sum_p (i_p - p)$ avec $I = \{i_1 < i_2 < \cdots < i_k\}$ est parfois plus naturelle, mais $|I| + |J|$ pour $\ell(I) = \ell(J) = k$ conserverait la même parité. De même si l'on décidait de numéroter lignes et colonnes à partir de 0, cela ne changerait pas la parité $\sum_{i \in I} i + \sum_{j \in J} j$ lorsque I et J ont la même cardinalité.

2 Mineurs de la co-matrice

Le *co-facteur* c_{ij} est $(-1)^{i+j}$ fois le mineur $|A|_{\{i\}^c\{j\}^c}$. La *co-matrice* est la matrice des co-facteurs. Enfin, le plus dur, la *transposée de la co-matrice* est la transposée de la co-matrice.
L'objet principal de cette fiche est de donner une formule pour les mineurs de la co-matrice.

Théorème 2. *Soit A une matrice de taille $N \times N$. Soit C la co-matrice de A. Pour deux multi-indices I et J de même longueur $1 \leqslant k \leqslant N$ on a :*

$$|C|_{IJ} = (\det A)^{k-1}(-1)^{|I|+|J|}|A|_{I^cJ^c} \tag{1}$$

Remarques

— La formule vaut aussi pour $k = 0$ dans le sens où à gauche on aura le déterminant de la matrice de taille 0×0, donc 1 et à droite l'expression $(\det A)^{-1} \cdot \det A$, donc une formule valable en tout cas si $\det A \neq 0$.

— Pour $k = N$, la formule dit $\det C = (\det A)^{N-1}$. Normal... dois-je rappeler que $A \cdot {}^tC = (\det A) I_N$? [3]

— Pour $k = 1$, on retrouve la définition des co-facteurs.

— Pour A non-inversible, on constate que tous les mineurs de la co-matrice d'ordres $k \geqslant 2$ sont nuls. Donc la co-matrice est de rang au plus 1. Dire qu'elle est de rang 0, c'est dire qu'elle est nulle, c'est dire que tous les mineurs de taille $(N-1)\times(N-1)$ de A sont nuls, c'est dire que A est de rang au plus $N-2$. Ainsi C est de rang 1 exactement si A est de rang $N-1$. Ça surprend la première fois ! Retrouvez-le directement !

— Avec $k = N-1$, on en conclut que $\boxed{\textit{la co-matrice de la co-matrice est } (\det A)^{N-2} \textit{ fois la matrice } A}$! (on a utilisé que $|I^c| + |J^c|$ a la même parité que $|I| + |J|$). C'est même vrai pour $N = 1$ (dans le même sens généralisé discuté plus haut pour $k = 0$), puisque la co-matrice de (a) est (1) dont la co-matrice est (1)...

— Pour les matrices inversibles de déterminant 1, passer deux fois à la co-matrice reconstruit la matrice d'origine, mais pas de surprise puisque la co-matrice est alors la transposée de l'inverse.

Preuve dans le cas des mineurs principaux. On va montrer le théorème si $I = J = \{1, \ldots, k\}$. Notons L_1, ..., L_N les lignes de la matrice C et considérons M_k la matrice de colonnes ${}^tL_1, \ldots, {}^tL_k$ puis les colonnes extraites de la matrice identité. Donc $\det M_k = |C|_{II}$. Si l'on multiplie M_k à gauche par A on obtient d'abord $\det A$ fois les k premières colonnes de la matrice identité, puis les colonnes $k+1$, ..., n de A. Donc le déterminant est $(\det A)^k |A|_{I^cJ^c}$. Ainsi

$$(\det A)^k |A|_{I^cI^c} = \det(AM_k) = \det A \det M_k = (\det A)|C|_{II}$$

d'où le résultat lorsque $\det A \neq 0$. Le cas général peut se voir en disant qu'on fait d'abord le raisonnement avec $\mathbb{K}$ égal au corps des fractions de l'anneau commutatif intègre $\mathbb{Z}[a_{ij}]$ où les a_{ij} sont des indéterminées, alors $\det A$ est un polynôme non nul et $A = (a_{ij})$ est inversible, d'où (1) dans l'algèbre $\mathbb{Z}[a_{ij}]$ d'abord puis pour les matrices A à coefficients dans $\mathbb{K}$ par spécialisation des indéterminées a_{ij}. (vaudououou...) □

3 Permutations et co-matrices

Soit A une matrice de taille $M \times N$. On dit qu'on permute les lignes par une permutation $\sigma \in S_M$ si l'on remplace A par la matrice $\sigma \cdot A$ obtenue en mettant la ligne L_i de A en $\sigma(i)^{\text{e}}$ ligne de $\sigma \cdot A$.

De telles opérations de lignes sont représentées par la multiplication à gauche par une matrice carrée $P(\sigma)$. Le nombre de colonnes N n'a pas d'influence sur le procédé, et $P(\sigma) = \sigma \cdot I_M$. Autrement dit en

3. Et la convention $0^0 = 1$ est importante ici !

j^{e} colonne le 1 est trouvé sur la $\sigma(j)^{\mathrm{e}}$ ligne, c'est-à-dire $P(\sigma)$ est la matrice de l'endomorphisme ϕ_σ de $\mathbb{K}^M$ qui envoie le vecteur e_j sur $e_{\sigma(j)}$. Comme $\phi_\tau \circ \phi_\sigma(e_j) = e_{\tau(\sigma(j))}$, on a $P(\tau\sigma) = P(\tau)P(\sigma)$, et on a défini une action à gauche de S_M sur les matrices avec M lignes (et un nombre arbitraire de colonnes). Les coefficients de $P(\sigma)$ sont donnés par la formule $\delta_{i,\sigma(j)} = \delta_{\sigma^{-1}(i),j}$ donc la transposée de $P(\sigma)$ est aussi son inverse.

On dit qu'on permute les colonnes par $\sigma \in S_N$ si l'on déplace la colonne A_j en position $A_{\sigma(j)}$ pour chaque j. Cela signifie une multiplication à droite par une matrice $R(\sigma)$. On obtient $R(\sigma)$ en appliquant la transformation à la matrice identité. Cela veut dire que dans la ligne i, le 1 est déplacé en $\sigma(i)^{\mathrm{e}}$ position. Ou encore dans la colonne j, on trouve un 1 pour l'indice $i = \sigma^{-1}(j)$. Donc $R(\sigma) = P(\sigma^{-1}) = P(\sigma)^{-1} = {}^tP(\sigma)$. Remarquons que $R(\sigma)R(\tau) = {}^t(P(\tau)P(\sigma)) = {}^tP(\tau\sigma) = R(\tau\sigma)$, autrement dit (attention !) on a une action *à gauche* de S_N, ici sur les matrices avec N colonnes et un nombre quelconque de lignes.

Pour les matrices carrées (qui nous intéressent principalement) de taille $N \times N$, on a donc deux actions à gauche de S_N : par permutation de lignes ou par permutation de colonnes. La première est donnée par multiplication à gauche par $P(\sigma)$, la seconde par multiplication à droite par $R(\sigma) = P(\sigma^{-1}) = P(\sigma)^{-1} = {}^tP(\sigma)$. Ces deux actions commutent.

Théorème 3. *Soit A une matrice carrée de taille $N \times N$ et C sa co-matrice. Soit $\sigma \in S_N$. La co-matrice de $\sigma \cdot A$ est $\varepsilon(\sigma)$ fois $\sigma \cdot C$, pour chacune des deux actions à gauche : permutations des lignes ou permutations des colonnes.*

On a noté $\varepsilon(\sigma)$ la signature de la permutation σ. C'est le déterminant de $P(\sigma)$. On notera $\sigma \cdot A$ l'action de permutation de lignes, et $\sigma \diamond A$ l'action de permutation de colonnes.

Preuve. Admettons le résultat pour les lignes.

Clairement $\sigma \diamond A = {}^t(\sigma \cdot {}^tA)$, et par ailleurs la co-matrice de la transposée est la transposée de la co-matrice. Donc la co-matrice de $\sigma \diamond A$ est la transposée de la co-matrice de $\sigma \cdot {}^tA$, soit encore la transposée de $\varepsilon(\sigma)$ fois $\sigma \cdot {}^tC$, c'est-à-dire $\varepsilon(\sigma)\sigma \diamond C$.

Pour les lignes, si le résultat est vrai pour σ_1 et pour σ_2 il est vrai pour $\sigma_2\sigma_1$, car la co-matrice de $\sigma_2 \cdot \sigma_1 \cdot A$ sera $\varepsilon(\sigma_2)$ fois $\sigma_2 \cdot \mathrm{comat}(\sigma_1 \cdot A)$ donc $\varepsilon(\sigma_2)\varepsilon(\sigma_1)\sigma_2 \cdot \sigma_1 \cdot C = \varepsilon(\sigma_2\sigma_1)\sigma_2\sigma_1 \cdot C$.

Il suffit de le montrer pour les transpositions. Soit donc τ qui échange la ligne i_1 avec la ligne i_2, $i_1 < i_2$. Si $i \notin \{i_1, i_2\}$ et j est quelconque la sous-matrice obtenue en supprimant dans $\tau \cdot A$ la i^{e} ligne et la j^{e} colonne ne diffère que par une transposition de lignes de la sous-matrice obtenue pareillement à partir de A. Donc le co-facteur de la position (i, j) dans $\tau \cdot A$ est l'opposé de celui de la position (i, j) dans A.

Soit maintenant $c_{i_1 j}$ un co-facteur de $\tau \cdot A$ pris dans la $i_1{}^{\mathrm{e}}$ ligne. Il vaut $(-1)^{i_1+j}$ fois un déterminant, où l'on a supprimé la j^{e} colonne de $\tau \cdot A$, et pris les lignes $1, \ldots, i_1 - 1, i_1 + 1, \ldots, i_2, \ldots, N$ de ce $\tau \cdot A$, qui sont les lignes originelles de A de numéros $1, \ldots, i_1 - 1, i_1, i_1 + 1, \ldots, i_2 - 1, i_2 + 1, \ldots, N$, [4] sauf que la $i_1{}^{\mathrm{e}}$ a été déplacée après la $(i_2 - 1)^{\mathrm{e}}$. Ce déterminant est donc $(-1)^{i_2+j}$ fois le co-facteur en (i_2, j) de A fois le signe de la permutation qui déplace i_1 par exactement $i_2 - i_1 - 1$ unités : si $i_2 = i_1 + 1$ pas de déplacement, si $i_2 = i_1 + 2$ une transposition, si $i_2 = i_1 + 3$ deux transpositions, etc... Cela donne donc $(-1)^{i_2+j}(-1)^{i_2-i_1-1}$ fois le co-facteur en (i_2, j), et comme on avait aussi un terme $(-1)^{i_1+j}$, on voit au final que $c_{i_1 j}$ est l'opposé du co-facteur de la position (i_2, j) en A.

On fait le même raisonnement pour $c_{i_2 j}$ et on conclut que la co-matrice de $\tau \cdot A$ est l'opposé de $\tau \cdot C$. □

Deuxième preuve (pour les fainéants). Soit $A_\sigma = P(\sigma)A$ la matrice A avec les lignes permutées par σ. Soit C_σ sa co-matrice. Supposons A inversible, donc aussi A_σ. Alors $C_\sigma = (\det A_\sigma){}^tA_\sigma^{-1} = \varepsilon(\sigma)(\det A)$ ${}^tP(\sigma)^{-1}{}^tA^{-1} = \varepsilon(\sigma)P(\sigma)C$, où l'on a noté $C = (\det A){}^tA^{-1}$ la co-matrice de A et utilisé que la tranposée de $P(\sigma)$ est aussi son inverse. Même preuve pour les permutations de colonnes. Finalement pour les A non inversibles, on peut invoquer l'argument avec l'algèbre polynomiale $\mathbb{Z}[a_{ij}]$. □

4. Je rédige comme si $i_2 > i_1 + 2$, pour les cas $i_2 = i_1 + 1$ ou $i_1 + 2$ il faut interpréter convenablement les notations.

Troisième preuve (pour les grands fainéants). On a une formule générale $\operatorname{comat}(PQ) = \operatorname{comat}(P)\operatorname{comat}(Q)$ et il suffit de l'appliquer avec $P = P(\sigma)$ et $Q = A$, sachant que $\operatorname{comat}(P(\sigma)) = \det(P(\sigma))\,{}^tP(\sigma)^{-1} = \varepsilon(\sigma)P(\sigma)$. □

Pour la formule $\operatorname{comat}(AB) = \operatorname{comat}(A)\operatorname{comat}(B)$: par exemple on le fait d'abord pour A et B inversibles par le lien avec la transposée de l'inverse et la multiplicativité du déterminant. Puis en utilisant le corps $\mathbb{K}(X)$ des fractions rationnelles en une indéterminée et les matrices $A + XI_N$, $B + XI_N$ (inversibles puisque de déterminants non nuls) on obtient $\operatorname{comat}((A + XI_N)(B + XI_N)) = \operatorname{comat}(A + XI_N)\operatorname{comat}(B + XI_N)$ qu'on spécialise en $X = 0$, puisque les entrées sont polynomiales. (« petit vaudou »)

Fin de la preuve de théorème 2. Soit donc $I = \{i_1 < i_2 < \cdots < i_k\}$ et $J = \{j_1 < j_2 < \cdots < j_k\}$ deux multi-indices de même longueur $k \geqslant 1$.

Soit σ' la permutation avec $\sigma'(p) = i_p$ pour $p \leqslant k$ et qui est croissante à partir de $k+1$. Soit $\sigma = \sigma'^{-1}$. Ainsi $\sigma(I^c) = \{k+1, \ldots, N\}$ en préservant l'ordre.

De même soit τ' la permutation avec $\tau'(q) = j_q$ pour $q \leqslant l$ et qui est croissante à partir de $k+1$. Soit $\tau = \tau'^{-1}$. Ainsi $\tau(J^c) = \{k+1, \ldots, N\}$ en préservant l'ordre.

Notons $B = \tau \diamond \sigma \cdot A$. Par construction $|A|_{I^cJ^c} = |B|_{\{k+1,\ldots,N\}\{k+1,\ldots,N\}}$. On sait déjà que

$$|D|_{\{\leqslant k\}\{\leqslant k\}} = (\det B)^{k-1}(-1)^{\sum_{i\leqslant k} i + \sum_{j\leqslant k} j}|B|_{\{>k\}\{>k\}} = (\varepsilon(\sigma)\varepsilon(\tau))^{k-1}(\det A)^{k-1}|A|_{I^cJ^c}$$

avec D la co-matrice de B. Par le théorème 3 :

$$D = \varepsilon(\tau)\varepsilon(\sigma)\tau \diamond \sigma \cdot C$$

et compte tenu de $\sigma(\{i_1, \ldots, i_k\}) = \{1, \ldots, k\} = \tau(\{j_1, \ldots, j_k\})$ en préservant l'ordre, il vient

$$|D|_{\{\leqslant k\}\{\leqslant k\}} = (\varepsilon(\sigma)\varepsilon(\tau))^k|C|_{IJ}$$

soit finalement

$$|C|_{IJ} = \varepsilon(\sigma)\varepsilon(\tau)(\det A)^{k-1}|A|_{I^cJ^c}$$

Il reste donc à calculer $\varepsilon(\sigma)$ et $\varepsilon(\tau)$. On obtient σ en faisant $i_1 - 1$ transpositions, puis $i_2 - 2$, puis $i_3 - 3$, ... donc $\varepsilon(\sigma) = (-1)^{\sum_{p\leqslant k}(i_p - p)} = (-1)^{|I|}(-1)^{k(k+1)/2}$. De même $\varepsilon(\tau) = (-1)^{|J|}(-1)^{k(k+1)/2}$. Ce qui achève la preuve du théorème 2. □

4 Algèbre extérieure et co-matrice

L'algèbre extérieure n'est pas au programme de l'agrégation interne, néanmoins pour les curieux, j'ajoute quelques éléments. Soit donc à nouveau A de taille $N \times N$ et ϕ l'endomorphisme associé de $\mathbb{K}^N$. Soit $e_1, \ldots, e_N$ les vecteurs de la base canonique $\mathcal{B}$. Soit $f_j = e_1 \wedge \cdots \wedge e_{j-1} \wedge e_{j+1} \wedge \cdots \wedge e_N$, ils forment une base $\mathcal{B}'$ de $\bigwedge^{N-1}\mathbb{K}^N$. Déterminons la matrice de $\wedge^{N-1}\phi$ dans cette base. La i^e coordonnée c_{ij} de $(\wedge^{N-1}\phi)(f_j)$ vérifie $e_i \wedge ((\wedge^{N-1}\phi)(f_j)) = (-1)^{i-1}c_{ij}e_1 \wedge \cdots \wedge e_N$. Il s'agit donc d'évaluer $e_i \wedge \phi(e_1) \wedge \cdots \wedge \phi(e_{j-1}) \wedge \phi(e_{j+1}) \wedge \cdots \wedge \phi(e_N)$, qui est aussi

$$\det_{\mathcal{B}}(e_i, \phi(e_1), \ldots, \phi(e_{j-1}), \phi(e_{j+1}), \ldots, \phi(e_N)) \cdot e_1 \wedge \cdots \wedge e_N$$

Par conséquent

$$c_{ij} = (-1)^{i-1}\det_{\mathcal{B}}(e_i, \phi(e_1), \ldots, \phi(e_{j-1}), \phi(e_{j+1}), \ldots, \phi(e_N)) = \det_{\mathcal{B}}(\phi(e_1), \ldots, \phi(e_{j-1}), e_i, \phi(e_{j+1}), \ldots, \phi(e_N))$$

est le déterminant de taille $N \times N$ formée en remplaçant dans la matrice A la j^e colonne par une colonne avec seulement un 1 en i^e ligne. En annulant toute la i^e ligne sauf le 1 puis en déplaçant cette ligne en première position ainsi que la colonne on voit que ce déterminant est exactement $(-1)^{i+j}$ fois le mineur $|A|_{\{i\}^c\{j\}^c}$, autrement dit c_{ij} est le co-facteur de la position (i, j) dans A.

Donc la co-matrice C de A est exactement la matrice de $\wedge^{N-1}\phi$ dans la base $\mathcal{B}'$.

Ce qui permet à propos de revoir que la co-matrice d'un produit AB est le produit des co-matrices.

Au passage on note que l'interprétation du co-facteur comme un déterminant de taille $N \times N$ marche pour tous les mineurs : la quantité $(-1)^{|I|+|J|}|A|_{IJ}$ est le déterminant de taille $N \times N$ obtenu en remplaçant dans A les lignes et les colonnes d'indices dans I^c, respectivement J^c de la manière suivante : pour j le plus petit indice de J^c on utilise la colonne n'ayant qu'un 1 en i^e position avec i le plus petit indice de I^c et des zéros ailleurs. Puis on itère. Il reste encore des entrées avec $j \in J$ et $i \in I^c$, on les annule. Ainsi dans la i^e ligne avec $i \in I^c$ la seule entrée non nulle est un 1 en la j^e colonne telle que l'ordinal de i dans I^c soit identique à l'ordinal de j dans J^c.

Pas clair ?