Compléments sur les matrices

Ces compléments sont **hors programme**. Ils constituent des développements pour approfondir ce qui a été vu.
Plus que les énoncés, ce sont les preuves qui sont intéressantes.

1 Calcul des puissances

Le calcul des puissances d'une matrice est souvent très utile.
Nous en verrons un exemple avec les suites récurrentes linéaires d'ordre p, ou avec les calculs de probabilités.

Méthode *(Calcul des puissances de A)*

Quelques méthodes usuelles :

1. **Par calcul direct avec une récurrence.**
   On calcule les premières puissances à la main et on en déduit une conjecture sur la forme des puissances de A.
   On démontre ensuite cette conjecture à l'aide d'une récurrence.

2. **Par le binôme de Newton.**
   On écrit $A = B+C$ avec B et C deux matrices **qui commutent** et dont on sait calculer facilement les puissances. En particulier, il faut penser aux matrices :

   - les matrices scalaires : λI_n dont on sait qu'elles commutent avec toutes les matrices. $(\lambda I_n)^p = \lambda^p I_n$.
   - la matrice J composée que de 1 : pour $p \geq 1$, $J^p = n^{p-1} J$.
     Attention à bien vérifier la commutativité et à traiter à part le cas $J^0 = I_n$ pour lequel la formule précédente est fausse.
   - une matrice nilpotente N telle que $N^r = 0$ pour une certaine puissance r.
     On sait que toutes les matrices triangulaires supérieures strictes (ou inférieures strictes) sont nilpotentes d'indice $r \leq n$
     (*Ce résultat est hors programme est doit être redémontré dans chaque cas particulier*).
     Attention à bien vérifier la **commutativité.**

et pour mémoire...

3. **Par diagonalisation** (deuxième année) : voir plus loin dans ce chapitre.

4. **Avec un polynôme annulateur :** vous serez guidé.

5. **Avec les applications linéaires :** on en reparlera...

Exemple

On pose $A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$

Calculer les puissances de A.

Solution :

On propose ici les deux premières méthodes. Celle avec le binôme de Newton est la plus naturelle (et la plus rapide) en raison de la forme particulière de la matrice.

1. **Conjecture puis récurrence** :

   On calcule à la main les premières puissances de A :

$$A^2 = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix} \qquad \text{et} \qquad A^3 = \begin{pmatrix} 1 & 3 & 6 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{pmatrix}$$

   On conjecture que la matrice A^p a des 1 sur la diagonale, p sur la sur-diagonale, et un autre coefficient non nul dans le coin.
   (*par stabilité des matrices triangulaires, on est sûr que les coefficients sous la diagonale resteront nuls*)
   Sauf si l'on a beaucoup d'intuition, il est difficile d'imaginer la valeur du coefficient dans le coin et on le pose comme une inconnue dans notre récurrence :
   On cherche donc à montrer qu'il existe une suite (α_p) telle que $\forall p \in \mathbf{N}$, $A^p = \begin{pmatrix} 1 & p & \alpha_p \\ 0 & 1 & p \\ 0 & 0 & 1 \end{pmatrix}$
   Initialisation : pour $p = 0$ le résultat est vrai et $\alpha_0 = 0$.
   Hérédité : on suppose le résultat vrai au rang p, et on cherche à le montrer au rang $p + 1$.

$$A^{p+1} = A^p\, A = \begin{pmatrix} 1 & p+1 & \alpha_p + p + 1 \\ 0 & 1 & p+1 \\ 0 & 0 & 1 \end{pmatrix}$$

   Ainsi, par principe de récurrence, le résultat est vrai pour tout $p \in \mathbf{N}$, avec la suite $(\alpha_p)_{p \in \mathbf{N}}$ définie par $\alpha_0 = 0$ et $\forall p \in \mathbf{N}$, $\alpha_{p+1} = \alpha_p + (p + 1)$.
   Il reste à trouver la forme de α_p en fonction de p.
   Assez simplement, on calcule successivement

$$\begin{array}{ll} p = 0 & \alpha_0 = 0 \\ p = 1 & \alpha_1 = 0 + 1 \\ p = 2 & \alpha_2 = 0 + 1 + 2 \\ p = 3 & \alpha_3 = 0 + 1 + 2 + 3 \\ \vdots & \\ p & \alpha_p = 0 + 1 + 2 + 3 + \cdots + p = \frac{p(p+1)}{2} \end{array}$$

   Une méthode un peu plus élégante pour obtenir ce résultat est d'écrire que

$$\forall p \in \mathbf{N},\ \alpha_{p+1} - \alpha_p = p + 1.$$

   On reconnait alors le terme général d'une somme télescopique :

$$\sum_{k=0}^{p-1} (\alpha_{k+1} - \alpha_k) = \sum_{k=0}^{p-1} (k+1) \iff \alpha_p - \alpha_0 = \sum_{k=1}^{p} k \iff \alpha_p = \frac{p(p+1)}{2}$$

2. **Méthode du binôme de Newton** :
   On écrit $A = I + N$ avec $N = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$.
   Les deux matrices commutent et N est nilpotente d'indice 3. En effet,
   $$N = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \qquad \text{et} \qquad N^2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \qquad \text{et} \qquad N^3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$
   Donc pour tout $k \geq 3$, $N^k = 0$, ainsi, d'après le binôme de Newton,
   $$A^p = \sum_{k=0}^{p} \binom{p}{k} N^k I^{p-k} = I + pN + \frac{p(p-1)}{2} N^2$$
   Donc
   $$A^p = \begin{pmatrix} 1 & p & \frac{p(p+1)}{2} \\ 0 & 1 & p \\ 0 & 0 & 1 \end{pmatrix}$$
3. **Méthode par diagonalisation.**
   vous verez l'an prochain que cette matrice n'est pas diagonalisable et qu'on est ramené à la méthode précédente).

2 Diagonalisation d'une matrice

Cette partie est hors programme, mais vous l'aborderez en profondeur en seconde année.
Aucun des résultats qui suivent ne peut être réutilisé sans démonstration.

Définition 2.1 *(Matrice diagonalisable)*

Soit $A \in \mathcal{M}_n(\mathbf{R})$,

On dit que A est diagonalisable sur $\mathbf{R}$ s'il existe une matrice inversible $P \in \mathrm{GL}_n(\mathbf{R})$, et une matrice diagonale $D \in \mathcal{M}_n(\mathbf{R})$ telles que
$$A = P\,D\,P^{-1}.$$

Outre certains intérêts intrinsèques qu'il est trop tôt pour présenter, la diagonalisation d'une matrice permet de calculer facilement ses puissances :

Théorème 2.2 *(Puissance d'une matrice diagonalisée)*

Si $A \in \mathcal{M}_n(\mathbf{R})$, $P \in \mathrm{GL}_n(\mathbf{R})$ et $(\lambda_1,\ \lambda_2, \cdots,\ \lambda_n)$ tels que

$$A = PDP^{-1} \qquad \text{avec} \quad D = \begin{pmatrix} \lambda_1 & & & (0) \\ & \lambda_2 & & \\ & & \ddots & \\ (0) & & & \lambda_n \end{pmatrix}$$

$$\text{alors } \forall p \in \mathbf{N}, \qquad A^p = P\,D^p P^{-1} \qquad \text{avec} \quad D^p = \begin{pmatrix} \lambda_1^p & & & (0) \\ & \lambda_2^p & & \\ & & \ddots & \\ (0) & & & \lambda_n^p \end{pmatrix}$$

Preuve

Par récurrence immédiate. ■

Très souvent, en première année, les questions des devoirs se suivront sur ce modèle :

1. On donne la matrice P et on demande de montrer qu'elle est inversible et de calculer P^{-1} avec l'algorithme de Gauss-Jordan.

2. On demande de montrer que $P^{-1}AP$ est une matrice diagonale.

3. On calcule les puissances de A à l'aide du théorème précédent que l'on démontre par récurrence.

Parfois, le sujet ira un peu plus loin et vous fera refaire une partie de ce qui suit :

Théorème 2.3 *(Condition suffisante de diagonalisabilité)*

Si $A - \lambda I$ n'est pas inversible pour n valeurs de λ **distinctes**,
alors A est diagonalisable.
Avec les notations précédentes, on peut alors écrire la matrice diagonale D sous la forme

$$D = \begin{pmatrix} \lambda_1 & & & (0) \\ & \lambda_2 & & \\ & & \ddots & \\ (0) & & & \lambda_n \end{pmatrix}$$

avec $(\lambda_1, \lambda_2, \cdots, \lambda_n)$ les n valeurs distinctes trouvées pour λ.

La $i^{\text{ème}}$ colonne de la matrice P est solution non nulle de l'équation $AX = \lambda_i X$.

Les (λ_i) s'appellent les **valeurs propres** de A et le $i^{\text{ème}}$ vecteurs colonne de P est un **vecteur propre** associé à la valeur propre λ_i.

⚠ La réciproque est fausse : cette condition est suffisante mais pas nécessaire. Vous verrez cela plus en détail en deuxième année, le but n'est pas de faire ici tout le programme de l'an prochain.

Remarque : Si on change l'ordre des λ_i dans l'écriture de D, alors cela revient simplement à échanger l'ordre des colonnes dans la matrice P. La matrice P n'est pas unique : on peut multiplier ses colonnes par des coefficients non nuls sans changer la relation.

Preuve

Supposons qu'il existe n valeurs distinctes $\lambda_1, \lambda_2, \cdots, \lambda_n$ telles que

$$\forall i \in [\![1, n]\!], \quad A - \lambda_i I \notin \mathrm{GL}_n(\mathbf{R}).$$

On sait que : $A - \lambda_i I \notin \mathrm{GL}_n(\mathbf{R}) \iff (A - \lambda_i I)X = 0$ n'est pas un système de Cramer
$\iff (A - \lambda_i I)X = 0$ admet une solution X non nulle

Ceci nous assure de l'existence de n vecteurs colonne non nuls $K_1, K_2, \cdots, K_n$ tels que

$$\forall i \in [\![1, n]\!], \quad (A - \lambda_i I)\, K_I = 0$$

C'est-à-dire

$$AK_i = \lambda_i K_i$$

On définit la matrice P en plaçant côte à côte les n vecteurs colonne K_i :

$$P = (K_1 | K_2 | \cdots | K_n).$$

Et on pose

$$D = \begin{pmatrix} \lambda_1 & & & (0) \\ & \lambda_2 & & \\ & & \ddots & \\ (0) & & & \lambda_n \end{pmatrix}$$

Alors, un calcul immédiat donne :

$$PD = (K_1 | K_2 | \cdots | K_n) \times \begin{pmatrix} \lambda_1 & & & (0) \\ & \lambda_2 & & \\ & & \ddots & \\ (0) & & & \lambda_n \end{pmatrix} = (\lambda_1 K_1 | \lambda_2 K_2 | \cdots | \lambda_n K_n) = AP$$

(la multiplication à droite de P par la matrice D revient à faire des dilatations sur les colonnes de P)

On admet[1] ici que la matrice P est inversible. On a donc

$$A = PDP^{-1}.$$

■

Dans les devoirs, on vous guidera donc de la façon suivante :

1. Trouver les valeurs de λ telles que $A - \lambda I$ n'est pas inversible (en général, en s'aidant du pivot de Gauss-Jordan).
2. Résoudre le système $AX = \lambda X$ et en déduire une matrice P qui convient.

1. Ce résultat se démontre très facilement quand on a étudié les applications linéaires en dimension finie.

3 Suites récurrentes linéaires d'ordre p

Définition 3.1 *(Suite récurrente)*

Une suite récurrente linéaire d'ordre p (à coefficients constants) est définie par

$$\forall n \in \mathbf{N}, \qquad u_{n+p} = a_1 u_{n+p-1} + a_2 u_{n+p-2} + \cdots + a_p u_n$$

avec $(a_1, a_2, \cdots, a_p) \in \mathbf{R}^p$.

Remarque : Comme dans l'ensemble du cours, il est sous-entendu que les suites considéres sont dans $\mathbf{R}^{\mathbf{N}}$.

Propriété 3.2 *(Mise sous forme matricielle)*

Toute suite récurrence linéaire d'ordre p peut être écrite sous forme matricielle comme une suite linéaire matricielle d'ordre 1 :

$$U_{n+1} = AU_n$$

avec $(U_n) \in (\mathcal{M}_{p,1}(\mathbf{R}))^{\mathbf{N}}$, et $A \in \mathcal{M}_p(\mathbf{R})$.
Avec les notations précédentes, on définit U_n et A par

$$U_n = \begin{pmatrix} u_n \\ u_{n+1} \\ \vdots \\ u_{n+p-1} \end{pmatrix} \qquad A = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & 0 \\ 0 & \cdots & \cdots & 0 & 1 \\ a_p & \cdots & \cdots & a_2 & a_1 \end{pmatrix}$$

Preuve

Il suffit de vérifier par le calcul que $U_{n+1} = AU_n$ ■

Propriété 3.3

Soit la suite la suite (U_n) définie pour tout $n \in \mathbf{N}$ par $U_{n+1} = AU_n$
Alors pour tout $m \geq 0$,

$$U_{n+m} = A^m U_n$$

En particulier $U_m = A^m U_0$

⚠ Les matrices A^m et U_0 ne commutent pas (les dimensions seraient incompatibles).

Preuve

par récurrence immédiate ■

Exemple

La suite de Fibonacci définie par $u_{n+2} = u_{n+1} + u_n$ et $u_0 = 0$ et $u_1 = 1$

Exemple

La suite définie par $u_{n+3} = -u_{n+2} + 5u_{n+1} - 3u_n$

Théorème 3.4 *(Solutions de la récurrence d'ordre 2)*

Soit $(b, c) \in \mathbf{R}^2$. On définit la suite récurrente linéaire d'ordre 2 à coefficients constants par

$$\forall n \in \mathbf{N},\ u_{n+2} + bu_{n+1} + cu_n = 0$$

On rappelle que **équation caractéristique** est $x^2 + bx + c = 0$. On note Δ son discriminant.

- Si $\Delta > 0$, on note (r_1, r_2) les deux solutions de l'équation caractéristique, Alors $\exists(\lambda, \mu) \in \mathbf{R}^2,$ tel que

$$\forall n \in \mathbf{N},\ u_n = \lambda r_1^n + \mu r_2^n$$

- Si $\Delta = 0$, on note r la racine double de l'équation caractéristique, Alors $\exists(\lambda, \mu) \in \mathbf{R}^2,$ tel que

$$\forall n \in \mathbf{N},\ u_n = \lambda r^n + \mu \boldsymbol{n} r^n$$

- Si $\Delta < 0$, on note $r_\pm = \rho\, \mathrm{e}^{\pm i\theta}$ les racines complexes conjuguées de l'équation caractéristique, Alors $\exists(\lambda, \mu) \in \mathbf{R}^2,$ tel que

$$\forall n \in \mathbf{N},\ u_n = \rho^n \left(\lambda \cos(n\theta) + \mu \sin(n\theta)\right)$$

λ et μ sont déterminés de façon unique par u_0 et u_1.

Preuve

Nous avions donné une preuve par récurrence dans le chapitre sur les suites usuelles. Nous proposons ici, une preuve utilisant les matrices. Elle est un peu plus difficile, mais elle permet de mieux comprendre d'où viennent les disjonctions cas $\Delta = 0$, $\Delta > 0$...
Si on pose pour tout $n \in \mathbf{N}$,

$$U_n = \begin{pmatrix} u_0 \\ u_1 \end{pmatrix} \qquad \text{et} \qquad A = \begin{pmatrix} 0 & 1 \\ -c & -b \end{pmatrix}$$

Alors $\forall n \in \mathbf{N},\ U_{n+1} = A^n U_n$ et par récurrence immédiate, $U_n = A^n U_0$.
Pour calculer les puissances de A, on cherche à la diagonaliser lorsque c'est possible. On étudie donc pour quelles valeurs de λ la matrice $A - \lambda I$ n'est pas inversible.

$$A - \lambda I = \begin{pmatrix} -\lambda & 1 \\ -c & -b-\lambda \end{pmatrix}$$

Or cette matrice n'est pas inversible si et seulement si son déterminant est nul, c'est-à-dire si et seulement si

$$-\lambda\,(-b-\lambda) - 1 \times (-c) = 0$$

On retrouve l'équation caractéristique : $\lambda^2 + b\lambda + c = 0$. C'est magique ![2]
$\star$ Si $\Delta > 0$, alors il y a deux valeurs propres distinctes r_1 et r_2, et A est diagonalisable. Il existe donc une matrice inversible $P \in \mathrm{GL}_n(\mathbf{R})$ telle que

$$A = P \begin{pmatrix} r_1 & 0 \\ 0 & r_2 \end{pmatrix} P^{-1}$$

Si on note $V_n = P^{-1}U_n$, alors $\forall n \in \mathbf{N},\ V_n = P^{-1}A^nU_0 = P^{-1}PD^nP^{-1}U_0 = D^nV_0$.
On peut écrire pour tout entier naturel n, $V_n = \begin{pmatrix} v_n \\ w_n \end{pmatrix}$
Et dans ce cas, $\quad \forall n \in \mathbf{N}, \quad v_n = r_1^n v_0 \quad$ et $\quad w_n = r_2^n w_0$.
Or $\forall n \in \mathbf{N},\ U_n = PV_n$, donc si on écrit P sous la forme

$$P = \begin{pmatrix} \alpha & \beta \\ \delta & \gamma \end{pmatrix}$$

alors $\quad \forall n \in \mathbf{N},\ u_n = \alpha v_n + \beta w_n = r_1^n(\alpha v_0) + r_2^n(\beta w_0)$. Dans le théorème, on a noté $\alpha v_0 = \lambda$ et $\beta w_0 = \mu$ (que l'on trouve grâce aux premiers termes de la suite u).
On remarque en particulier que l'on n'a pas besoin ici de connaître les coefficients de la matrice P.
$\star$ Si $\Delta < 0$, alors on diagonalise la matrice sur $\mathbf{C}$ (deux racines complexes distinctes) et on se ramène ensuite sur $\mathbf{R}$ comme cela est détaillé dans la preuve du chapitre sur les suites usuelles.
$\star$ Si $\Delta = 0$, alors la condition suffisante de diagonalisabilité n'est pas vérifiée (dans le cas présent, on pourrait montrer très simplement que la matrice ne peut pas être diagonalisable : vous pouvez chercher à le démontrer en exercice).
On note $r = \frac{-b}{2}$ la racine double et on sait donc que $\Delta = b^2 - 4c = 0$, ainsi $c = \frac{b^2}{4} = r^2$.
Dans la suite, on suppose $r \neq 0$ (sinon, la suite est nulle pour $n \geq 2$)
On peut écrire A sous la forme

$$A = \begin{pmatrix} 0 & 1 \\ -r^2 & 2r \end{pmatrix}$$

On s'inspire du cas diagonalisable et on étudie la matrice $N = A - rI$:

$$N = A - rI = \begin{pmatrix} 0 & 1 \\ -r^2 & 2r \end{pmatrix} - \begin{pmatrix} r & 0 \\ 0 & r \end{pmatrix} = \begin{pmatrix} -r & 1 \\ -r^2 & r \end{pmatrix}$$

Et on trouve que $N^2 = 0$.
On peut donc appliquer avec profit le binôme de Newton (les matrices commutent) :

$$\begin{aligned} \forall n \geq 1,\ A^n = (N + rI)^n &= \sum_{k=0}^{n} \binom{n}{k} N^k r^{n-k} = r^n I + nr^{n-1}N = r^{n-1}(r + nN) \\ &= r^{n-1} \begin{pmatrix} r(1-n) & p \\ -nr^2 & r(1+n) \end{pmatrix} \end{aligned}$$

Or $U_n = A^nU_0$ et donc avec la première ligne du produit :

$$u_n = r^n(1-n)u_0 + nr^{n-1}u_1 = r^n u_0 + nr^n(\frac{u_1}{r} - u_0).$$

Dans le théorème, on a pose $\lambda = u_0$ et $\mu = \frac{u_1}{r} - u_0$. ■

2. En fait, c'est mathématique, mais selon une acception très courante face au tableau : c'est magique. Et comme pour tout bon tour de magie, ce n'est pas drôle d'aller chercher à dévoiler les ficelles qui le permettent, et on se garde donc bien d'essayer de comprendre ce que dit l'enseignant pour garder intact l'émerveillement.

Utilisation similaire :
Cette méthode de diagonalisation peut aussi être utilisée pour des suites ayant des relations linéaire entre elles.
Par exemple, si on définit deux suite (u_n) et (v_n) telles que

$$\forall n \in \mathbf{N}, \begin{cases} u_{n+1} = au_n + bv_n \\ v_{n+1} = cu_n + dv_n \end{cases}$$

alors on pourra poser

$$U_n = \begin{pmatrix} u_n \\ v_n \end{pmatrix} \quad \text{avec} \quad U_{n+1} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} U_n = A\,U_n$$

Ainsi,

$$\forall n \in \mathbf{N}, \quad U_n = A^n\,U_0$$

Et le problème revient à calculer les puissances de A par une méthode au choix.

Remarque : Dans le cas présent avec uniquement deux suites, il est souvent plus simple de faire apparaître une relation de récurrence d'ordre 2 pour la suite (u_n) et de résoudre cette récurrence. Si vous avez oublié cette méthode, un petit passage par le chapitre sur les suites usuelles pourrait vous éviter un passage par la case prison.

Nous retrouverons abondamment ce type de relations linéaires **en probabilités**.