Trigonalisation

Il est parfois utile (par exemple si un endomorphisme n'est pas diagonalisable) de chercher une base dans laquelle sa matrice est triangulaire.

.

2.1 Lemme. Toute matrice triangulaire supérieure est semblable à une matrice triangulaire inférieure.

• • *Démonstration.* La permutation des éléments de la base canonique $p(e_i) = e_{n-i+1}$ convertit les matrices triangulaires supérieures en matrices triangulaires inférieures. • •

On a vu que les valeurs propres d'une matrice triangulaire sont ses éléments diagonaux $\lambda_1, ..., \lambda_n$ et son polynôme caractéristique est

$p_A(x) = (-1)^n \Pi_{i=1}^n (x - \lambda_i)$.

.

Définition. Un endomorphisme est **trigonalisable** si il existe une base dans laquelle sa matrice est triangulaire (supérieure ou inférieure).

2.2 Théorème. Un endomorphisme est trigonalisable dans K si et seulement si son polynôme caractéristique est scindé dans K.

Plus précisement:

- $K = \mathbf{C}$: tout endomorphisme est trigonalisable dans $\mathbf{C}$.

- $K = \mathbf{R}$: un endomorphisme est trigonalisable dans $\mathbf{R}$ si et seulement si toutes les racines complexes de son polynôme caractéristique sont réelles.

.

• • *Démonstration.* On procède par récurrence sur la dimension de E. Supposons le théorème vrai pour la dimension inférieure à n.

Soit dim $E = n$. Soit λ une valeur propre de f et v un vecteur propre, $f(v) = \lambda v$. Prenons $v_1 = v$ et complétons v_1 en une base $(v_1, ..., v_n)$ de E. La matrice A de f dans cette base est triangulaire supérieure par blocs: le premier bloc diagonal est λ et le deuxième bloc (de dimension $n-1$) sera noté B. On a $p_f(x) = (\lambda - x)p_B(x)$, donc $p_B(x)$ est aussi scindé dans K. On peut appliquer l'hypothèse de récurrence à la matrice B; il exsite une matrice de passage P telle que $P^{-1}BP$ soit triangulaire supérieure. Soit $\tilde{P}$ la matrice diagonale par bloc avec le premier bloc de taille 1 egal à 1 et le deuxième bloc de taille $n-1$ égal à P. Alors $\tilde{P}^{-1}A\tilde{P}$ est triangulaire supérieure.

• •

La trigonalisation dans $\mathbf{C}$ permet de calculer rapidement les valeurs propres des puissances de A:

2.3. Proposition. Soit $p_A(x) = (-1)^n \Pi_{i=1}^n (x - \lambda_i)$ la factorisation de p_A dans $\mathbf{C}$. Alors $p_{A^k}(x) = (-1)^n \Pi_{i=1}^n (x - \lambda_i^k)$.

• • *Démonstration.* Si A est semblable à une matrice triangulaire T avec les éléments diagonaux $\lambda_1, ..., \lambda_n$, alors $q(A)$ est semblable à la matrice triangulaire $q(T)$ avec les éléments diagonaux $q(\lambda_1), ..., q(\lambda_n)$. • •

Remarque: De manière plus générale, soit $q(t) = a_0 + a_1 t + ... + a_k t^k$, et $q(A) = a_0 + a_1 A + ... + a_k A^k$. Alors $p_{q(A)}(x) = (-1)^n \Pi_{i=1}^n (x - q(\lambda_i))$.

.

2.4. Corollaire. Les valeurs propres réelles de la matrice $q(A)$ sont les nombres réels dans la liste $q(\lambda_1), ..., q(\lambda_n)$.

.

Rappel: la **trace** d'une matrice est la somme de ses éléments diagonaux: $tr(A) = \sum_i a_{ii}$.

Lemme. Pour toute deux matrice A et B on a $tr(AB) = tr(BA)$.

Corollaire. Les matrices semblables ont la même trace.

2.5. Corollaire. Soit $K = C$. La trace d'une matrice est égal à la somme de ses valeurs propres comptées avec leurs multiplicités: $tr(A) = \lambda_1 + ... + \lambda_n$.

On a aussi $tr(A^k) = \lambda_1^k + ... + \lambda_n^k$ et de manière plus générale, pour un polynôme $q(t) = a_0 + a_1 t + ... + a_k t^k$, on a $tr(q(A)) = q(\lambda_1) + ... + q(\lambda_n)$

.

Remarque 1. Les coefficients du polynôme caractéristique

$p_A(x) = (-1)^n x^n + c_1 x^{n-1} + ... + c_{n-1} x + c_n = (-1)^n \Pi_{i=1}^n (x - \lambda_i)$ sont liés avec ses racines par les formules de Viète (ou Vieta):

$c_k = (-1)^{n-k} \sum_{1 \leq i_1 < ... < i_k \leq n} \lambda_{i_1} ... \lambda_{i_k}$.

.

Remarque 2. Les traces des puissances $s_k = tr(A^k)$ sont liées avec les coefficients du polynôme caractéristique $p_A(x) = (-1)^n x^n + c_1 x^{n-1} + ... + c_{n-1} x + c_n$ par les *formules de Newton:*

$$(-1)^n s_k + c_1 s_{k-1} + ... + c_{k-1} s_1 + k c_k = 0$$

$k = 1, 2, ..., n$. Cela permet d'exprimer les c_k en termes des s_k (ou réciproquement) par récurrence.

Endomorphismes nilpotents.

Définition. Un endomorphisme est **nilpotent** s'il existe k tel que $f^k = 0$. La valeur minimal de k s'appelle l'**indice de nilpotence** de f.

2.6. Proposition. Un endomorphisme est nilpotent si et seulement si il a le même polynôme caractéristique que l'endomorphisme nul: $p_f(x) = (-1)^n x^n$.

• • *Démonstration.* 1. Soit f nilpotent. Si λ est une valeur propre de f, alors λ^k est une valeur propre de $f^k = 0$, donc $\lambda = 0$. Donc le polynôme $p_f(x)$ n'a pas de racine non-nulle. Si $K = C$, cela entraine que $p_f(x) = (-1)^n x^n$. Si $K = R$, on se ramène au cas $K = C$ en considérant la matrice de f dans une base comme une matrice à coefficients complèxes.

2. Soit $p_f(x) = (-1)^n x^n$ - un polynôme scindé avec une seule racine 0 (de multiplicité n). f est donc trigonalisable par une matrice triangulaire T avec la diagonale principale nulle. Pour une telle matrice on a $T^n = 0$ (évident). • •

2.7. Corollaire. Un endomorphisme nilpotent non-nul n'est pas diagonalisable.

En effet, les valeurs propres de f étant nulles, si f est diagonalisable, la matrice diagonale semblable à f serait nulle.

.

On a déjà remarqué que $p_{f+\mu Id}(x) = p_f(x-\mu)$. En notant $g = f + \mu Id$ on déduit de la proposition précédente:

2.8. Proposition. Soit g un endomorphisme. On a
$p_g(x) = (-1)^n(x-\mu)^n$ si et seulement si $g - \mu Id$ est nilpotent.

.

Sommes directes.

Définition. Soit E un espace vectoriel sur K, ($K = \mathbf{R}$ ou $\mathbf{C}$). Soit $E_1, ..., E_k$ ses sous-espaces vectoriels. La **somme** $E_1 + ... + E_k$ est le sous-espace formé de tous les vecteurs $v = v_1 + ... + v_k$ où $v_i \in E_i$. La somme est **directe** (noté $E_1 \bigoplus ... \bigoplus E_k$) si une telle décomposition est **unique:** si $v = v_1 + ... + v_k = u_1 + ... + u_k$ avec $v_i \in E_i$ et $u_i \in E_i$ alors $v_i = u_i$, $(i = 1, ..., k)$.

.

2.9. Lemme. Les propriétés 1.-5. suivantes sont équivalentes:

1. La somme $E_1 + ... + E_k$ est directe.

2. La relation $v_1 + ... + v_k = 0$ où $v_i \in E_i$ entraine $v_1 = 0$, ... , $v_k = 0$.

3. Soit B_1, ... , B_k des bases des sous-espaces $E_1, ..., E_k$. Alors leur réunion $B = B_1 \cup ... \cup B_k$ est libre et donc est une base de la somme $E_1 + ... + E_k$ (base adaptée).

4. Soit dim $(E) < \infty$, alors dim $(E_1 + ... + E_k) = \dim E_1 + ...+ \dim E_k$.

5. Pour tout i on a $E_i \cap (E_1 + ... + E_{i-1} + E_{i+1} + ... + E_k) = \{0\}$.

.

A noter: la somme de deux sous-espaces E_1 et E_2 est directe si et seulement si $E_1 \cap E_2 = \{0\}$.

Exemple: soit $e_1, ..., e_k$ des vecteurs non-nuls de E. La somme $Ke_1 + ... + Ke_k$ est directe si et seulement si les vecteurs $e_1, ..., e_k$ sont linéairement indépendants. On a $E = Ke_1 \bigoplus ... \bigoplus Ke_k$ si et seulement si $(e_1, ..., e_k)$ est une base de E.

.

Sous-espaces stables. Décomposition en blocs.

Définition. Soit $f : E \to E$ un endomorphisme. Un sous-espace F de E est **stable** ou **invariant** par f si $f(F) \subset F$, (donc, si pour tout $v \in F$ on a $f(v) \in F$).

A noter: $\mathrm{Ker}(f)$ et $\mathrm{Im}(f)$ sont des sous-espaces stables par f.

2.10. Lemme. Si g commmute avec f, $fg = gf$, alors $\mathrm{Ker}(g)$ et $\mathrm{Im}(g)$ sont des sous-espaces stables par f.

[En particulier, on peut prendre $g = a_0 Id + a_1 f + ... + a_k f^k$.]

• • *Démonstration.* Si $v \in Ker(g)$ on a $g(f(v)) = f(g(v)) = 0$ donc $f(v) \in Ker(g)$.

Si $v \in Im(g)$, alors $v = g(u)$ et $f(v) = f(g(u)) = g(f(u))$, donc
$f(v) \in Im(g)$. • •

Si F est stable par f, on définit **l'endomorphisme unduit** $f_F : F \to F$ par $f_F(v) = f(v)$ si $v \in F$. Dans une base de E où les premiers vecteurs forment une base de F la matrice A de f est triangulaire par blocs.

Rappel: Le déterminant d'une matrice triangulaire par blocs est le produit des déterminants des blocs diagonaux.

2.11. Corollaire. Le polynôme caractéristique de l'endomorphisme induit f_F divise le polynôme caractéristique de f.

• • *Démonstration.* Le premier bloc diagonal de A est la matrice de f_F. • •

Soit E la somme directe des sous-espaces stables par f, $E = E_1 \bigoplus ... \bigoplus E_k$; soit f_i l'endomorphisme induit dans E_i. Alors l'étude de f se réduit à l'étude de chaque f_i séparement: f est une sorte la "somme directe" des f_i. La matrice de f dans une base adaptée est diagonale par blocs, le i-ème bloc diagonal étant la matrice de f_i. Un des objectifs de la réduction est de décomposer f en blocs de taille minimum (blocs"indécomposables").

Remarque. Si E est la somme directe des sous-espaces stables par f, $E = E_1 \bigoplus ... \bigoplus E_k$, le polynôme caractéristique de f est le produit des polynômes caractéristiques des endomorphismes f_i induits dans E_i $(i = 1, ..., k)$.

Donc la décomposition de f en blocs est liée à la factorisation du polynôme caractéristique de f.

.

Espaces propres

Définition. Soit $\lambda \in K$. Le sous-espace

$E_\lambda = \mathrm{Ker}(f - \lambda Id) = \{v \in E : f(v) = \lambda v\}$ s'appelle **l'espace propre associé à** λ. (Noter que $E_0 = \mathrm{Ker}\ (f)$).

Si λ n'est pas une valeur propre, $E_\lambda = \{0\}$.

Evidemment, E_λ est **stable** par f.

On a vu qu'un endomorphisme est diagonalisable si et seulement si ses vecteurs propres engendrent tout l'espace E. Autrement dit, un endomorphisme est diagonalisable si et seulement si E est la somme de ses espaces propres.

Le théorème 1.3 (des vecteurs propres associés à des valeurs propres deux à deux distinctes sont linéairement indépendants) peut être reformulé de façon suivante:

2.12. Théorème. Les espaces propres associés à des valeurs propres deux à deux distinctes sont en somme directe.

2.13. Corollaire. Un endomorphisme est **diagonalisable** si et seulement si la somme des dimensions de ses espaces propres est égale à $\dim(E)$. (Rappelons que $\dim(E) < \infty$.)

2.14. Proposition. La dimension de l'espace propre E_λ ne dépasse pas la multiplicité de λ dans le polynôme caractéristique.

• • *Démonstration.* L'endomorphisme $\widetilde{f}$ induit par f dans E_λ est proportionnel à l'identité: $\widetilde{f} = \lambda Id$. Donc $p_{\widetilde{f}}(x) = (-1)^k(x-\lambda)^k$, où $k = dim(E_\lambda)$. Pour conclure, il suffit de se rappeler que $p_{\widetilde{f}}(x)$ divise $p_f(x)$. • •

2.15. Corollaire. Un endomorphisme est diagonalisable si et seulement si p_f est scindé et la dimension de chaque espace propre E_λ est égale à la multiplicité de λ.

• • *Démonstration.* $E = E_{\lambda_1} \bigoplus ... \bigoplus E_{\lambda_s}$ si et seulement si

$dim(E) = dim(E_{\lambda_1}) + ... + dim(E_{\lambda_s})$. Mais

$dim(E_{\lambda_i}) \leq mult(\lambda_i)$ et $mult(\lambda_1) + ... + mult(\lambda_s) = dim(E)$, d'où le résultat. • •