Autour de la diagonalisation

Cédric Gérot, Pierre Granjon, Nicolas Le Bihan
Laboratoire des Images et des Signaux
Grenoble

January 22, 2003

Contents

1 Eléments d'algèbre **2**
1.1 Espace vectoriel 2
1.1.1 Définition 2
1.1.2 Sous-espace vectoriel 2
1.1.3 Base d'un espace vectoriel 2
1.2 Application linéaire 3
1.2.1 Définition 3
1.2.2 Sous espaces vectoriels $Im\varphi$ et $Ker\varphi$ 3
1.2.3 Déterminant 3
1.2.4 Spectre d'un endomorphisme 4
1.3 Matrice 4
1.3.1 Définition 4
1.3.2 Matrice d'une application linéaire 5
1.3.3 Matrice d'un endomorphisme 5

2 Diagonalisation et SVD **5**
2.1 Définitions 5
2.1.1 Qu'est-ce qu'être diagonalisable? 5
2.1.2 Exemples 5
2.1.3 Diagonalisation 6
2.1.4 Singular Values Decomposition 6
2.2 Interprétations algébriques 7
2.2.1 Changement de bases 7
2.2.2 Diagonalisation: un même changement de bases pour le départ et l'arrivée 8
2.2.3 SVD: des changements de bases pour que les bases de départ et d'arrivée soient toutes deux orthogonales 8
2.2.4 Diagonalisation orthogonale: réservée aux matrices symétriques . . 9

3 Utilisations **9**
3.1 SVD: décomposer dans une meilleure base 9
3.2 Diagonalisation: de l'itération au développement de Taylor 9

Introduction

Ne cherchez pas: il n'y a rien de nouveau dans ce petit rapport. Celui-ci est né de questions que nous nous posions chacun pour soi et que nous avons voulu clarifier ensemble. L'implusion initiale pourrait se résumer à “mais que signifie réellement ce que nous utilisons quotidiennement sans y penser”. Ou encore “posons clairement une fois pour toute la signification stricte de nos outils afin de chasser les légers doutes qui peuvent nous envahir quand nous les utilisons”. Ce rapport sera donc assez sobre; l'introduction en contient sans aucun doute toutes les phrases creuses. Alors, ne tardons plus: bonne lecture!

1 Eléments d'algèbre

Dans cette première section, nous rappelons la définition des divers animaux algébriques que nous manipulons.

1.1 Espace vectoriel

1.1.1 Définition

Définition 1 (Espace vectoriel) *On appelle* espace vectoriel *sur K (les scalaires) un ensemble non vide E muni*

- *d'une loi de composition interne $+$ (c'est la somme de vecteurs), telle que $(E, +)$ soit un groupe commutatif,*
- *d'une loi de composition externe, application de $K \times E$ dans E (c'est le produit par un scalaire, ou encore produit externe) notée $(\lambda, x) \mapsto \lambda.x$ et telle que*

  $1.x = x$ *(1 est l'élément neutre multiplicatif de K)*

  $(\lambda + \mu).x = \lambda.x + \mu.x$

  $\lambda.(\mu.x) = (\lambda\mu).x$

  $\lambda.(x + y) = \lambda.x + \mu.y$

1.1.2 Sous-espace vectoriel

Définition 2 (Sous-espace vectoriel) *Un* sous-espace vectoriel *de E est une partie non vide F de E,*

- *stable pour l'addition de E ($(x, y) \in F^2 \Rightarrow x + y \in F$)*
- *stable pour le produit externe ($(\lambda, x) \in K \times F \Rightarrow \lambda.x \in F$)*

1.1.3 Base d'un espace vectoriel

Définition 3 (Base) *Une famille $(x_i)_{i \in I}$ de vecteurs de E est une* base *de E ssi elle est libre (ou encore composée de vecteurs linéairement indépendants: on ne peut en écrire un comme une combinaison d'autres de la famille)*

$$(\sum_{i \in I} \lambda_i x_i = 0) \Rightarrow (\forall i \in I, \lambda_i = 0)$$

et génératrice (tout vecteur de E peut s'écrire comme une combinaison de vecteurs de la famille)

$$\forall x \in E,\ \exists \lambda_i : x = \sum_{i \in I} \lambda_i x_i$$

1.2 Application linéaire

1.2.1 Définition

Définition 4 (Application linéaire) *Soient E et F deux espaces vectoriels sur le même corps K. Une application φ de E dans F est* linéaire *ssi* $\varphi(\lambda.x + \mu.y) = \lambda.\varphi(x) + \mu.\varphi(y)$.

Définition 5 *Une application linéaire est également appelée* morphisme *d'espaces vectoriels.*
Lorsque φ est bijective, c'est un isomorphisme.
Lorsque φ est linéaire de E dans E, c'est un endomorphisme.
Un endomorphisme bijectif est appelé automorphisme.
Enfin, une application linéaire de E dans K s'appelle une forme linéaire *sur E.*

1.2.2 Sous espaces vectoriels $Im\varphi$ et $Ker\varphi$

Définition 6 (Noyau) *Le* noyau *d'une application linéaire φ de E dans F est défini par*

$$Ker\varphi = \varphi^{-1}(0)$$

(c'est l'ensemble des vecteurs de l'espace de départ E qui sont envoyés par φ dans le 0 *de F).*

Définition 7 (Image) *L'*image *d'une application linéaire φ de E dans F est définie par*

$$Im\varphi = \varphi(E)$$

(c'est l'ensemble des vecteurs de l'espace d'arrivée F qui sont atteints par φ).

Comme toute image d'un sous-espace vectoriel par une application linéaire, le *noyau* d'une application linéaire est un sous-espace vectoriel (de l'espace de départ E). Comme toute image réciproque d'un sous-espace vectoriel par une application linéaire, l'*image* d'une application linéaire est un sous-espace vectoriel (de l'espace d'arrivée F).

1.2.3 Déterminant

Définition 8 (Déterminant suivant une base) *Le* déterminant *suivant une base $B = (e_1, \ldots, e_n)$ de E notée* $\det_B$ *est la seule forme de E^n dans K qui soit*

- *n-linéaire (ses n applications partielles sont linéaires): $\forall (a_1, \ldots, a_n) \in E^n$, $\forall k \in 1 \ldots n$, $v \in E \mapsto \det_B(\ldots, a_{k-1}, v, a_{k+1}, \ldots) \in K$ est linéaire,*
- *alternée (si on échange deux vecteurs, les signe est changé),*
- *et qui prenne la valeur 1 en $(e_1, \ldots, e_n)$.*

Notons que toutes les formes n-linéaires alternées ψ sur E de dimension n sont proportionnelles:

$$\psi(v_1, \dots, v_n) = \psi(e_1, \dots, e_n)\mathrm{det}_B(v_1, \dots, v_n)$$

Définition 9 (Déterminant d'un endomorphisme) *On appelle* déterminant d'un endomorphisme φ *de* E *le scalaire noté* $\mathrm{det}\varphi$ *défini par*

$$\mathrm{det}\varphi = \mathrm{det}_B(\varphi(e_1), \dots, \varphi(e_n))$$

où $B = (e_1, \dots, e_n)$ *est une base de* E*.*

Notons que ce scalaire est indépendant de la base B.

1.2.4 Spectre d'un endomorphisme

Définition 10 (Valeur propre, Vecteur propre) $\lambda \in K$ *est valeur propre de l'endomorphisme* $\varphi : E \to E$ *s'il existe un vecteur* v *de* E *tel que*

$$\varphi(x) = \lambda.x$$

Et x *est un* vecteur propre *associé à* λ*.*

Définition 11 (Spectre) *On appelle* spectre *de l'endomorphisme* φ *l'ensemble de ses valeurs propre.*

Définition 12 (Sous-espace propre) *Le sous-espace vectoriel défini par* $Ker(\varphi - \lambda\mathrm{id}_E)$ *où* λ *est une valeur propre de* φ*, est le* sous-espace propre *associé à* λ*. Privé de* 0*, c'est l'ensemble des vecteurs propres associés à* λ*.*

Définition 13 (Polynôme caractéristique) *On appelle* polynôme caractéristique *de l'endomorphisme* φ *le polynôme de* K *dans* K *défini par*

$$P_\varphi(\lambda) = \mathrm{det}(\varphi - \lambda\mathrm{id}_E)$$

où l'on note $\lambda\mathrm{id}_E$ *l'endomorphisme qui à tout vecteur* v *associe* $\lambda.v$*.*

Notons que les valeurs propres de φ sont les racines de son polynôme caractéristique.

1.3 Matrice

1.3.1 Définition

Définition 14 (Matrice) *On appelle* matrice *de type* (n, p) *sur un corps* K *toute application*

$$\begin{array}{rcl} M : I\!N_n \times I\!N_p & \to & K \\ (i, j) & \mapsto & m_{ij} \end{array}$$

1.3.2 Matrice d'une application linéaire

Définition 15 (Matrice d'une application linéaire dans des bases) *Soit une application linéaire φ de E dans F. Soient $B_E = (e_1, \ldots, e_p)$ une base de E et $B_F = (f_1, \ldots, f_p)$ une base de F. Il existe alors (a_{ij}) tels que $\forall j \in 1 \ldots p$,*

$$\varphi(e_j) = \sum_{i=1}^{n} a_{ij} f_j$$

La matrice *de φ* relativement à *B_E et B_F est l'application* $\mathrm{Mat}\varphi_{B_E,B_F} : (i,j) \mapsto a_{ij}$.

Notons que φ est déterminée de façon unique par $(a_{ij})_{i=1\ldots n,\ j=1\ldots p}$.
Notons également que les colonnes d'une matrice sont les coordonnées dans B_F de l'image par φ des vecteurs de la base B_E.

1.3.3 Matrice d'un endomorphisme

Les matrices associées aux endomorphismes sont des matrices carrées. La réciproque n'est pas vraie: deux espaces vectoriels différents peuvent avoir la même dimension. Cependant, toute matrice carrée est la matrice d'un endomorphisme. On peut donc malgré tout définir pour les matrices carrées les notions de déterminant et de spectre.

2 Diagonalisation et SVD

Dans cette section nous tentons d'éclaircir l'action de diagonaliser ou décomposer une application linéaire ou une matrice. Notez à ce propos que l'on parlera indifféremment d'application linéaire ou de matrice.

2.1 Définitions

2.1.1 Qu'est-ce qu'être diagonalisable?

Théorème 1 (CNS pour être diagonalisable) *Un endomorphisme φ (matrice carrée) est diagonalisable ssi*

- *P_φ est scindé, et*
- *pour toute valeur propre λ_i de φ,* $\dim Ker(\varphi - \lambda_i \mathrm{id}_E) = \mathrm{m}_\varphi(\lambda_i)$.

Notez que P_φ est scindé s'il peut s'écrire $P_\varphi = \Pi_{i=1}^{\kappa}(\lambda_i - \lambda)^{\alpha_i}$, $\sum_{i=1}^{\kappa} \alpha_i = n$. Et alors, $\alpha_i = \mathrm{m}_\varphi(\lambda_i)$.

En d'autres termes, la multiplicité algébrique ($\alpha_i = \mathrm{m}_\varphi(\lambda_i)$) d'une valeur propre doit être égale à sa multiplicité géométrique (la dimension du sous-espace propre associé).

2.1.2 Exemples

Exemple 1

$$A = \begin{bmatrix} 2 & -1 & 1 \\ 1 & 0 & 1 \\ 0 & 0 & 2 \end{bmatrix}$$

n'est pas diagonalisable.
En effet, si $P_A(\lambda) = (\lambda - 1)^2(\lambda - 2)$ est bien scindé, en revanche

$$\begin{aligned} \dim Ker(A - I) &= \dim Ker \begin{bmatrix} 1 & -1 & 1 \\ 1 & -1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \\ &= 3 - \text{rang} \begin{bmatrix} 1 & -1 & 1 \\ 1 & -1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \\ &= 1 \\ &< 2 = \mathrm{m}_A(1) \end{aligned}$$

<u>Exemple 2</u>

$$A = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

est diagonalisable dans $\mathbb{C}$, mais pas dans $\mathbb{R}$. En effet,

$$P_A(\lambda) = -(\lambda - 1)(\lambda - j)(\lambda - j^2)$$

où $j = \exp(\sqrt{-1}\frac{2\pi}{3})$.

2.1.3 Diagonalisation

Diagonaliser une matrice carrée (n, n) diagonalisable A revient à écrire (on dit aussi décomposer en valeurs propres, EVD)

$$A = V \Lambda V^{-1}$$

où

- V est une matrice inversible, dont les colonnes sont des vecteurs propres de A,
- Λ est une matrice diagonale dont les éléments diagonaux sont les valeurs propres de A

Les colonnes v_i de V sont également appelées "vecteurs propres droits" de A:

$$A v_i = \lambda_i v_i$$

afin de les différencier des colonnes ν_i de V^{-1^T} qui sont des "vecteurs propres gauches" de A:

$$\nu_i^T A = \lambda_i \nu_i^T$$

2.1.4 Singular Values Decomposition

Décomposer une matrice (n, p) A en valeurs singulières revient à écrire

$$A = U \Sigma W^T$$

où

- U est une matrice (n, n) unitaire : $U^{-1} = U^T$,
- Σ est une matrice (p, n) dont seuls les éléments "diagonaux" Σ_{ii} peuvent être non nuls (elles sont appelées valeurs singulières),
- W est une matrice (p, p) unitaire : $W^{-1} = W^T$,

2.2 Interprétations algébriques

2.2.1 Changement de bases

Définition 16 (Matrice de passage) *On appelle* matrice de passage *d'une base $B_E = (e_1, \ldots, e_n)$ à une base $\overline{B_E} = (\overline{e_1}, \ldots, \overline{e_n})$ du même espace vectoriel E la matrice P dont les colonnes sont les coordonnées de $(\overline{e_1}, \ldots, \overline{e_n})$ dans la base $(e_1, \ldots, e_n)$.*

Il faut comprendre l'expression "passage de B_E à $\overline{B_E}$" comme "cette matrice permet d'exprimer un vecteur donné par ses coordonnées dans B_E (on le note V_{B_E}) en fonction de ses coordonnées dans $\overline{B_E}$ (on le note alors $V_{\overline{B_E}}$) ":

$$V_{B_E} = PV_{\overline{B_E}}$$

Théorème 2 *Soit φ une application linéaire d'un espace vectoriel E dans un espace vectoriel F.*
Soient B_E et $\overline{B_E}$ deux bases de E et P la matrice de passage de B_E à $\overline{B_E}$.
Soient B_F et $\overline{B_F}$ deux bases de F et Q la matrice de passage de B_F à $\overline{B_F}$.
L'application φ ramenée dans les bases B_E et B_F est notée φ_{B_E,B_F}.
L'application φ ramenée dans les bases $\overline{B_E}$ et $\overline{B_F}$ est notée $\varphi_{\overline{B_E},\overline{B_F}}$.
Alors

$$\mathrm{Mat}\varphi_{\overline{B_E},\overline{B_F}} = Q^{-1}\mathrm{Mat}\varphi_{B_E,B_F}P$$

En effet, dans B_E,

$$e_i = \begin{bmatrix} 0 & & \\ \vdots & & \\ 0 & & \\ 1 & \leftarrow & i \\ 0 & & \\ \vdots & & \\ 0 & & \end{bmatrix}$$

est transformé par P en la $i^{\text{ème}}$ colonne de P qui est égale à $\overline{e_i}$ (exprimé dans B_E).
Soit

$$\overline{x} = \begin{bmatrix} \vdots \\ \overline{x_i} \\ \vdots \end{bmatrix}$$

dans $\overline{B_E}$. Dans B_E, il s'écrit

$$\sum_i \overline{x_i e_i} \quad = \quad \sum_i \overline{x_i} P e_i$$

$$
\begin{aligned}
&= P\sum_i \overline{x_i} e_i \\
&= P\begin{bmatrix} \vdots \\ \overline{x_i} \\ \vdots \end{bmatrix} \\
&= P\overline{x}
\end{aligned}
$$

Donc, en notant $A = \mathrm{Mat}\varphi_{B_E,B_F}$, l'image de $\overline{x}$ par φ s'écrit dans B_F

$$
\begin{aligned}
y &= AP\overline{x} \\
&= \begin{bmatrix} \vdots \\ \overline{y_i} \\ \vdots \end{bmatrix} \\
&= \sum_i y_i f_i
\end{aligned}
$$

Aussi, comme on a écrit $\overline{e_i} = Pe_i$, on peut écrire $f_i = Q^{-1}\overline{f_i}$.

$$
y = \sum_i y_i Q^{-1}\overline{f_i}
$$

Donc, dans $\overline{B_F}$, ce vecteur s'écrit

$$
\begin{aligned}
\overline{y} &= Q^{-1}\begin{bmatrix} \vdots \\ \overline{y_i} \\ \vdots \end{bmatrix} \\
&= Q^{-1}AP\overline{x}
\end{aligned}
$$

C.Q.F.D.

2.2.2 Diagonalisation: un même changement de bases pour le départ et l'arrivée

La diagonalisation d'une matrice carrée diagonalisable (d'un endomorphisme diagonalisable φ) peut alors s'interpréter comme la recherche d'un changement de base V appliqué à E, tant comme espace de départ de φ que comme son espace d'arrivée, pour que φ s'exprime alors simplement comme une multiplication des coordonnées d'un vecteur par un scalaire (*a priori* différent pour chaque coordonnée).

La nouvelle base est une base de vecteurs propres de φ et les scalaires multipliant les coordonnées des vecteurs sont les valeurs propres de φ.

2.2.3 SVD: des changements de bases pour que les bases de départ et d'arrivée soient toutes deux orthogonales

La décomposition en valeurs singulières d'une matrice (d'une application linéaire φ) peut alors s'interpréter comme la recherche du changement de base W de l'espace de départ E et du changement de base U de l'espace d'arrivée F qui permettent la situation suivante:

- les nouvelles bases de départ et d'arrivée sont orthonormées ($U^T U = I$ et $W^T W = I$),
- φ s'exprime alors simplement comme une multiplication des coordonnées d'un vecteur par un scalaire (*a priori* différent pour chaque coordonnée).

Les scalaires multipliant les coordonnées des vecteurs sont les éléments "diagonaux" de Σ (les valeurs singulières).

2.2.4 Diagonalisation orthogonale: réservée aux matrices symétriques

Bien sûr, lorsque l'on souhaite diagonaliser un endomorphisme diagonalisable φ de telle sorte que la nouvelle base de E soit une base orthonormée, on fait alors une *diagonalisation orthogonale* qui est à la fois une diagonalisation (EVD) et une SVD.

On peut montrer que cela n'est réalisable (et réalisé) que pour les endomorphismes (matrices carrées) symétriques.

3 Utilisations

3.1 SVD: décomposer dans une meilleure base

3.2 Diagonalisation: de l'itération au développement de Taylor