Applications linéaires et matrices

J.Ch. Gilbert, INRIA

Applications linéaires

- Soient E et F deux espaces vectoriels *sur le même corps commutatif* $\mathbb{K}$ et f une application de E dans F, On dit que f est **linéaire** si

  $(L.1)$ $f(x+y) = f(x) + f(y)$, $\forall x, y \in E$,
  $(L.2)$ $f(\alpha x) = \alpha f(x)$, $\forall \alpha \in \mathbb{K}$, $\forall x \in E$.

  On note $L(E,F)$ l'ensemble des applications linéaires de E dans F.

- Propriétés immédiates :
  1) On a toujours $f(0) = 0$.
  2) $L(E,F)$ est un sous-espace vectoriel de $\mathcal{F}(E,F)$, l'ensemble des applications de E dans F.

- Applications linéaires particulières :
  1) On appelle **endomorphisme** de E une application linéaire de E dans E. On note $\operatorname{End}(E)$ ou $L(E)$ l'ensemble des endomorphismes de E.
  2) On appelle **isomorphisme** de E sur F une application linéaire *bijective* de E dans F. On note $\operatorname{Isom}(E,F)$ l'ensemble des isomorphismes de E dans F. S'il existe un isomorphisme de E dans F, on dit que E et F son **isomorphes**.
  3) On appelle **forme linéaire** sur E une application linéaire de E dans $\mathbb{K}$. On note $E^* = L(E,\mathbb{K})$ l'ensemble des formes linéaires sur E.

- Exemples.
  1) L'**application identité** ou **identité** de E est l'application linéaire qui envoie tout élément $x \in E$ sur lui-même :

  $$\operatorname{Id}_E : E \to E : x \mapsto \operatorname{Id}_E(x) = x.$$

  2) L'**application nulle** de E dans F est l'application linéaire qui à $x \in E$ associe $0 \in F$.

Image et noyau

- Soit $f \in L(E,F)$.
- L'**image** de f, notée $R(f)$ ou $\operatorname{Im}(F)$, est l'ensemble des éléments de F qui sont l'image par f d'un vecteur de E :

  $$R(f) := \{f(x) : x \in E\}.$$

  C'est un sous-espace vectoriel de F. En fait, l'image par une application linéaire d'un sous-espace vectoriel de E est un sous-espace vectoriel de F. On a

  $$f \text{ est surjective} \quad \Longleftrightarrow \quad R(f) = F.$$

- On appelle **rang** de f, que l'on note $\mathrm{rg}(f)$, la dimension de son image : $\mathrm{rg}(f) = \dim R(f)$.
- Le **noyau** de f, noté $N(f)$ ou $\mathrm{Ker}(f)$, est l'ensemble des éléments de E pour lesquels f s'annule :

$$N(f) := \{x \in E : f(x) = 0\}.$$

  C'est un sous-espace vectoriel de E. On a

$$f \text{ est injective} \quad \Longleftrightarrow \quad N(f) = \{0\}.$$

Théorèmes sur la dimension

- Commençons par un résultat décrivant comment se transforment les familles libre ou génératrice par des applications linéaires injective ou surjective.

> Soient $f \in L(E, F)$ et $\mathcal{F} := \{v_1, \ldots, v_p\} \subset E$.
> (i) Si f est injective et $\mathcal{F}$ est libre dans E, alors $f(\mathcal{F})$ est libre dans F.
> (ii) Si f est surjective et $\mathcal{F}$ est génératrice dans E, alors $f(\mathcal{F})$ est génératrice dans F.
> (iii) Si f est bijective et $\mathcal{F}$ est une base de E, alors $f(\mathcal{F})$ est une base de F.

- Espaces vectoriels isomorphes.

> Deux espaces vectoriels (*de dimension finie*) sont isomorphes si et seulement si ils ont la même dimension.

- **Théorème de la dimension.**

  Si $f \in L(E, F)$ alors

$$\boxed{\dim E = \dim N(f) + \dim R(f).}$$

  <u>Corollaire</u> :

> Si $f \in L(E, F)$ et $\dim E = \dim F$.
> alors les propriétés suivantes sont équivalentes
> (i) f est injective,
> (ii) f est surjective,
> (iii) f est bijective.

Représentation matricielle

- On suppose que $\dim E = n$ et $\dim F = p$ et on se donne des bases $\mathcal{U} := \{u_1, \ldots, u_n\}$ de E et $\mathcal{V} := \{v_1, \ldots, v_p\}$ de F. On note

$$\boxed{a_{ij} := \text{composante } i \text{ de } f(u_j) \text{ dans la base } \mathcal{V}.} \tag{1}$$

  Ces pn scalaires sont déterminés par f et déterminent complètement f.

- L'ensemble des scalaires a_{ij} peuvent être rangés dans un tableau

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pn} \end{pmatrix},$$

que l'on appelle **matrice** de type (p, n), car elle a p lignes et n colonnes. On note que le premier indice de a_{ij} est l'indice de ligne et que le second est l'indice de colonne. L'ensemble des matrices de type (p, n) avec éléments dans $\mathbb{K}$ est noté

$$\mathcal{M}_{p,n}(\mathbb{K}) \quad \text{ou} \quad \mathbb{K}^{p\times n}.$$

Si $p = n$ on dit que la matrice est **carrée** d'**ordre** n. Les éléments a_{ii}, $1 \leq i \leq n$, forment ce que l'on appelle la **diagonale** d'une matrice carrée.

- Lorsque les éléments a_{ij} d'une matrice A sont déterminés par une application linéaire en utilisant la formule (1), on dit que A est la **représentation matricielle** de f dans les bases $\mathcal{U}$ et $\mathcal{V}$ et on note

$$A = M_{\mathcal{U},\mathcal{V}}(f).$$

- <u>Exemples</u>.
  1) Tous les éléments de la matrice associée à l'application nulle sont nuls. On obtient ainsi la **matrice nulle** de type (p, n) :

$$0 := \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{pmatrix}.$$

  2) Si $f = \mathrm{Id}_E$ et si on prend la même base au départ et à l'arrivée dans E pour représenter f, on obtient la **matrice identité**

$$I := \begin{pmatrix} 1 & 0 & 0 & & \\ 0 & 1 & 0 & \ddots & \\ 0 & 0 & 1 & \ddots & 0 \\ & \ddots & \ddots & \ddots & 0 \\ & & 0 & 0 & 1 \end{pmatrix},$$

  qui a des 1 sur sa diagonale et des 0 ailleurs.

Structure vectorielle de $\mathcal{M}_{p,n}(\mathbb{K})$

- L'application $M_{\mathcal{U},\mathcal{V}} : L(E, F) \to \mathcal{M}_{p,n}(\mathbb{K})$ est bijective et permet de mettre sur $\mathcal{M}_{p,n}(\mathbb{K})$ une structure d'espace vectoriel au moyen des formules suivantes : $\forall f, g \in L(E, F)$ et $\forall \alpha \in \mathbb{K}$,

$$M_{\mathcal{U},\mathcal{V}}(f) + M_{\mathcal{U},\mathcal{V}}(g) = M_{\mathcal{U},\mathcal{V}}(f + g),$$
$$\alpha M_{\mathcal{U},\mathcal{V}}(f) = M_{\mathcal{U},\mathcal{V}}(\alpha f).$$

On en déduit

$$\boxed{\dim L(E, F) = (\dim E)(\dim F)}$$

et des formules pour la somme de deux matrices et le produit d'une matrice par un scalaire.

- Somme de deux matrices $A, B \in \mathcal{M}_{p,n}(\mathbb{K})$:

$$\boxed{(A+B)_{ij} = A_{ij} + B_{ij}.}$$

- Produit d'une matrice $A \in \mathcal{M}_{p,n}(\mathbb{K})$ par un scalaire $\alpha \in \mathbb{K}$:

$$\boxed{(\alpha A)_{ij} = \alpha(A_{ij}).}$$

Produit d'une matrice par un vecteur-colonne

- Soient $x_1, \ldots, x_n$, les composantes d'un vecteur $x \in E$ dans la base $\mathcal{U}$: $x = \sum_{j=1}^{n} x_j u_j$. On peut ranger ces composantes sous la forme d'un **vecteur-colonne**, qui est une matrice de type $(n, 1)$, que l'on note encore x :

$$x := \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

- Soient $f \in L(E, F)$ et $A = M_{\mathcal{U},\mathcal{V}}(f)$ la représentation matricielle de f dans les bases $\mathcal{U}$ et $\mathcal{V}$. Si les composantes de $y = f(x)$ dans la base $\mathcal{V}$ sont $y_1, \ldots, y_p$, on a les formules

$$y_i = \sum_{j=1}^{i} a_{ij} x_j, \quad i = 1, \ldots, p.$$

Ces formules se retiennent en constatant que y_i est obtenu en multipliant les éléments de la i-ième ligne de la matrice A par les éléments du vecteur-colonne x :

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pn} \end{pmatrix} \begin{pmatrix} \boxed{\begin{matrix} x_1 \\ \vdots \\ x_n \end{matrix}} \end{pmatrix}. \tag{2}$$

Observons que ce produit a bien un sens, car les lignes de A ont n éléments, tout comme le vecteur-colonne x.

Composition d'applications linéaires et produit matriciel

- Voici quelques propriétés immédiates, liées à la composition d'applications linéaires.

> Soient E, F et G trois espaces vectoriels sur le même corps $\mathbb{K}$. Si $f \in L(E, F)$ et $g \in L(F, G)$, alors $g \circ f \in L(E, G)$. De plus, on a les propriétés suivantes :
> (i) $g \circ (f_1 + f_2) = (g \circ f_1) + (g \circ f_2), \quad \forall f_1, f_2 \in L(E, F), \forall g \in L(F, G)$,
> (ii) $(g_1 + g_2) \circ f = (g_1 \circ f) + (g_2 \circ f), \quad \forall f \in L(E, F), \forall g_1, g_2 \in L(F, G)$,
> (iii) $g \circ (\alpha f) = (\alpha g) \circ f = \alpha(g \circ f), \quad \forall \alpha \in \mathbb{K}, \forall f \in L(E, F), \forall g \in L(F, G)$.

- Reprenons les notations du résultat ci-dessus : E, F et G trois espaces vectoriels sur le même corps $\mathbb{K}$, $f \in L(E,F)$ et $g \in L(F,G)$. On note

$$n = \dim E, \quad p = \dim F, \quad q = \dim G.$$

On se donne des bases $\mathcal{U}$ de E, $\mathcal{V}$ de F et $\mathcal{W}$ de G, et on note $A = M_{\mathcal{U},\mathcal{V}}(f)$ de type (p,n) et $B = M_{\mathcal{V},\mathcal{W}}(g)$ de type (q,p). Comme $g \circ f \in L(E,G)$, la matrice $C = M_{\mathcal{U},\mathcal{W}}(g \circ f)$ est de type (q,n). Son élément (i,j) peut se calculer à partir des éléments de A et B par la formule :

$$\boxed{c_{ij} = \sum_{k=1}^{p} b_{ik}\, a_{kj}.} \tag{3}$$

On appelle **produit** des matrices B et A, la matrice C dont l'élément (i,j) est donné par la formule ci-dessus. On note

$$C = BA.$$

On a donc

$$\boxed{M_{\mathcal{U},\mathcal{W}}(g \circ f) = M_{\mathcal{V},\mathcal{W}}(g) M_{\mathcal{U},\mathcal{V}}(f).}$$

- <u>Remarques</u> :
  1) Pour obtenir c_{ij}, il faut additionner les produits des éléments de la ligne i de B par les éléments de la colonne j de A :

$$\begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1p} \\ \vdots & \vdots & & \vdots \\ \boxed{b_{i1} \quad b_{i2} \quad \cdots \quad b_{ip}} \\ \vdots & \vdots & & \vdots \\ b_{q1} & b_{q2} & \cdots & b_{qp} \end{pmatrix} \begin{pmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2j} & \cdots & a_{2n} \\ \vdots & & \vdots & & \vdots \\ a_{p1} & \cdots & a_{pj} & \cdots & a_{pn} \end{pmatrix}.$$

  2) Pour que le produit matriciel BA ait un sens (en particulier la somme dans (3)), il faut que le nombre de lignes de A soit égal au nombre de colonnes de B. Ceci découle d'ailleurs du fait que l'espace vectoriel d'arrivée de f est l'espace vectoriel de départ de g.
  3) Le produit matriciel *n'*est *pas* commutatif en général. Il peut d'abord y avoir un problème de dimension. Mais même si les matrices A et B sont carrées, en général

$$BA \neq AB.$$

  Par exemple

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \neq \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

  4) Attention aussi au fait que

$$AB = 0 \quad \not\Rightarrow \quad A \text{ ou } B \text{ est nulle.}$$

  De même

$$A \text{ non nulle et } AB = AC \quad \not\Rightarrow \quad B = C.$$

5) Si les matrices A, B et C sont de type convenable pour que les produits ci-dessous aient un sens, on a

$$A(B+C) = AB + AC,$$
$$(A+B)C = AC + BC,$$
$$A(\alpha B) = (\alpha A)B = \alpha(AB),$$
$$(AB)C = A(BC).$$

6) Si on note $M_{\mathcal{U}}(x)$ le vecteur-colonne formé des composantes de x dans la base $\mathcal{U}$, la formule (2) peut se récrire sous forme de produit matriciel :

$$M_{\mathcal{V}}(f(x)) = M_{\mathcal{U},\mathcal{V}}(f) M_{\mathcal{U}}(x).$$

Isomorphisme et inverse de matrice

- Si $f \in L(E,F)$ est bijective (c'est un isomorphisme) alors $f^{-1} : F \to E$ est linéaire.
- On dit qu'une matrice *carrée* A d'ordre n est **inversible**, s'il existe une matrice B d'ordre n telle que

$$BA = AB = I.$$

  La matrice B est appelée l'**inverse** de A et notée A^{-1}.
- Ces deux notions sont liées :

  Soit $f \in L(E,F)$. Alors f est bijective si et seulement si $M_{\mathcal{U},\mathcal{V}}(f)$ est inversible. Dans ce cas

$$(M_{\mathcal{U},\mathcal{V}}(f))^{-1} = M_{\mathcal{V},\mathcal{U}}(f^{-1}).$$

Effet d'un changement de base

- Soient $\mathcal{U}$ et $\mathcal{U}'$ deux bases d'un espace vectoriel E. On appelle **matrice de passage** de la base $\mathcal{U}$ à la base $\mathcal{U}'$, la matrice

$$P_{\mathcal{U}\to\mathcal{U}'} := M_{\mathcal{V},\mathcal{U}}(\mathrm{Id}_E).$$

  Cette matrice est inversible et on a

$$P_{\mathcal{U}'\to\mathcal{U}} = (P_{\mathcal{U}\to\mathcal{U}'})^{-1}.$$

- Si $x \in \mathbb{K}^n$ et $x' \in \mathbb{K}^n$ représentent un même vecteur de E dans des bases $\mathcal{U}$ et $\mathcal{U}'$, on a

$$x' = (P_{\mathcal{U}\to\mathcal{U}'})^{-1} x.$$

- Soient E et F deux espaces vectoriels, $f \in L(E,F)$, $\mathcal{U}$ et $\mathcal{U}'$ deux bases de E, $\mathcal{V}$ et $\mathcal{V}'$ deux bases de F, $A = M_{\mathcal{U},\mathcal{V}}(f)$ et $A' = M_{\mathcal{U}',\mathcal{V}'}(f)$. Alors

$$A' = (P_{\mathcal{V}\to\mathcal{V}'})^{-1} A (P_{\mathcal{U}\to\mathcal{U}'}).$$

Notions de rang

- Voici quelques notions de rang :
  - Soit $\mathcal{F} = \{v_1, \dots, v_p\}$ une famille de vecteurs. On appelle **rang** de $\mathcal{F}$ la dimension de vect$\mathcal{F}$, l'espace vectoriel engendré par la famille $\mathcal{F}$. C'est le nombre maximal de vecteurs linéairement indépendants que l'on peut trouver dans $\mathcal{F}$.
  - Soit $A \in \mathcal{M}_{p,n}(\mathbb{K})$. On appelle **rang** de A, noté $\operatorname{rg} A$, le rang dans $\mathbb{K}^p$ de la famille des vecteurs formés des colonnes de A.
  - On appelle **rang** d'une application linéaire $f \in L(E,F)$ la dimension de son image : $\operatorname{rg} f = \dim R(f)$.

  Les deux dernières notions sont liées par le fait que si $A = M_{\mathcal{U},\mathcal{V}}(f)$, on a $\operatorname{rg} f = \operatorname{rg} A$.

- Soient E_0, E, F, F_0 des espaces vectoriels et $f \in L(E,F)$. Alors
  1) si $g \in L(E_0, E)$ est *surjective*, alors $\operatorname{rg}(f \circ g) = \operatorname{rg} f$,
  2) si $h \in L(F, F_0)$ est *injective*, alors $\operatorname{rg}(h \circ f) = \operatorname{rg} f$.