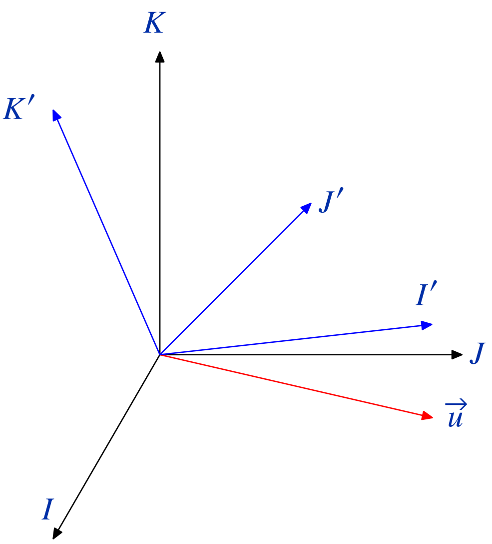
Mathématiques

...

T. Horsin

Algèbre linéaire. Version $\sum_{i=1}^{10} \frac{1}{10^{i!}}$.

Voici quelques notes sur l'algèbre linéaire et plus spécifiquement l'algèbre linéaire en dimension finie. Je ne prétends pas à une exposition originale et on peut trouver tout ce qui suit dans une littérature déjà existante et très abondante. Il n'est pas impossible que je me trompe de temps en temps, je corrige quand je m'en rends compte et me signaler des erreurs me semblerait appréciable. En vert les parties nécessaires, il me semble, à la compréhension du cours mva107 et aussi du mva101. Lire le reste n'est pas nécessairement inintéressant.



Angles d'Euler

Figure 1

1 Rappel: les espaces vectoriels

1.1 Généralités

Ce paragraphe est indiqué pour rendre ce cours à peu près auto-suffisant. On peut tout à fait ne pas le lire et tout de même comprendre la suite.

Un espace vectoriel (notion introduite de manière axiomatique par Peano)[1] se définit comme un ensemble de vecteurs qui, mutatis mutandis, se comportent comme les vecteurs usuels du plan ou

[1] Giuseppe, mathématicien italien 1858-1932

Case 2D5000, 292 rue Saint-Martin 75141 Paris Cedex 03
Bureau 17.0.14
Tel: 0158808765, Mel: thierry.horsin@cnam.fr

Conservatoire National des Arts et Métiers
Département d'Ingénierie Mathématique

de l'espace. Pour définir un espace vectoriel, on a donc besoin d'un ensemble de scalaires qui vont jouer le rôle de multiplicateurs des vecteurs.
La définition générale est la suivante:
Soit $\mathbb{K}$ un corps, c'est à dire, un anneau pour lequel la multiplication définit une loi de groupe sur l'ensemble des éléments non nuls. En particulier il y a un élément neutre pour la multiplication. Dans toute la suite $\mathbb{K}$ sera soit $\mathbb{R}$ soit $\mathbb{C}$, éventuellement $\mathbb{Q}$. Noter que, par exemple, $\mathbb{Z}$ n'est pas un corps (l'inverse d'un élément de $\mathbb{Z}$ n'étant pas dans $\mathbb{Z}$).
Un $\mathbb{K}$ espace vectoriel E est d'abord (mais pas seulement) un ensemble E muni d'une addition qui en fait un groupe commutatif.
Ces éléments qui sont donc des vecteurs (attention: un ensemble de vecteurs n'est pas nécessairement représenté ni représentable, du moins aisément, avec des flèches unilatérales.), peuvent être ajoutés entre eux (c'est l'addition) pour donner un vecteur (elle est donc une loi interne), on la note – et c'est probablement une des rares notations quasi-obligatoires en maths– $+$ car on demande que cette addition soit commutative, c'est à dire que

$$x + y = y + x$$

si x et y sont n'importe quels vecteurs de E.
Cette addition possède un élément neutre (le vecteur nul) noté 0. Tout vecteur x possédant un opposé $-x$ c'est à dire que

$$x + (-x) = (-x) + x = 0.$$

Noter que si y est un opposé de x alors

$$y + (x + (-x)) = y.$$

Pour pouvoir enlever la parenthèse on demande l'associativité de l'addition c'est à dire que pour tous vecteurs x_1 x_2 x_3, on doit avoir

$$(x_1 + x_2) + x_3 = x_1 + (x_2 + x_3).$$

L'égalité $y + (x + (-x)) = y$ devient donc $(y + x) + (-x) = y$ et donc si y est un opposé de x on a

$$-x = y$$

qui prouve l'unicité de l'opposé.
On demande ensuite à E d'être muni d'une loi externe (multiplication à gauche par les scalaires):

$$\mathbb{K} \times E \to E$$
$$(k, x) \mapsto k.x$$

et cette loi doit vérifier

$$\forall (\lambda, \mu, x, y) \in K \times K \times E \times E,\ (\lambda \mu).x = \lambda.(\mu.x)$$
$$(\lambda + \mu).x = \lambda.x + \mu.x$$
$$\lambda.(x + y) = \lambda.x + \lambda.y$$
$$1.x = x.$$

Définition 1.1 *Quand E répond à ces définitions avec $\mathbb{K}$, on dit que E est un $\mathbb{K}$-espace vectoriel, ou un espace vectoriel sur $\mathbb{K}$.*

Case 2D5000, 292 rue Saint-Martin 75141 Paris Cedex 03
Bureau 17.0.14
Tel: 0158808765, Mel: thierry.horsin@cnam.fr
Conservatoire National des Arts et Métiers
Département d'Ingénierie Mathématique

Remarquons alors $(0+\lambda).x = \lambda.x = 0.x + \lambda.x$ donc

$$0.x = 0,$$

et

$$\lambda.(x+0) = \lambda.x + \lambda.0 = \lambda.x$$

donc

$$\lambda.0 = 0,$$

puis

$$(1+(-1))x = 0 = x + (-1).x,$$

donc

$$-x = (-1).x \quad .$$

1.2 Les bases

Soit E un espace vectoriel
Soit $\mathcal{F} := (u_i)_{i\in I}$ une famille d'élément de E (finie ou infinie –mais on pourra se contenter de considérer le cas finie–).

Définition 1.2 *On dit que la famille $\mathcal{F}$ est une famille libre, si toute combinaison linéaire finie nulle de ses éléments a ses coefficients nuls.*

Autrement dit, si pour tout $n \in \mathbb{N}$, pour tout $i_1, ..., i_N \in I^N$

$$\forall(\lambda_1, ..., \lambda_N) \in \mathbb{K}^N, \sum_{k=1}^{N} \lambda_k u_{i_k} = 0 \Rightarrow \forall k = 1, ..., N,\ \lambda_k = 0.$$

Ainsi aucun des u_i ne peut être exprimé en fonction des autres d'où l'expression libre.

Définition 1.3 *On dit que la famille $\mathcal{F}$ est une famille génératrice si $\forall u \in E, \exists N \in \mathbb{N},\ i_1, .., i_N \in I^N,\ \lambda_1, ..., \lambda_N \in \mathbb{K}^N$ tels que $u = \sum_{k=1}^{N} \lambda_k u_{i_k}$.*

Ainsi grâce à la famille $\mathcal{F}$, on peut exprimer chaque vecteur de E. Noter que pour chaque vecteur u, on a besoin d'un nombre fini N de vecteurs de E.

Définition 1.4 *On dit que $\mathcal{F}$ est une base de E si c'est une famille à la fois libre et génératrice.*

Intuitivement une famille libre doit contenir peu de vecteurs pour éviter des combinaisons linéaires non nulles entre ces vecteurs, tandis qu'une famille génératrice doit en contenir beaucoup afin par les combinaisons linéaires de ceux-ci d'engendrer tous les vecteurs de l'espace.
Une base en contient donc juste assez.

Case 2D5000, 292 rue Saint-Martin 75141 Paris Cedex 03
Bureau 17.0.14
Tel: 0158808765, Mel: thierry.horsin@cnam.fr

Conservatoire National des Arts et Métiers
Département d'Ingénierie Mathématique

De fait on a

Proposition 1.5 *Toute sous-famille d'une famille libre est libre, toute sur-famille d'une famille génératrice est génératrice.*

La démonstration est suffisament simple pour que je l'omette.
Tout d'abord, on peut démontrer, mais cela dépend de l'axiome de Zorn[2] (lui-même indécidable)

Théorème 1.6 *Tout $\mathbb{K}$-espace vectoriel possède une base.*

Définition 1.7 *On dit que E est dimension finie sur $\mathbb{K}$ si il existe une base de E constituée d'un nombre fini de vecteurs de E.*

Il est facile de construire des espaces vectoriels de dimension finie (et infinie aussi d'ailleurs).
On dispose alors du

Théorème 1.8 *Si E possède une base avec un nombre fini d'éléments, alors toutes les bases ont le même nombre d'éléments.*

En effet, cela provient du résultat suivant.

Proposition 1.9 *Soit $u_1, ..., u_k$, k vecteurs formant une famille libre, $k+1$ combinaisons linéaires de ceux-ci forment une famille liée.*

En effet, si l'on note ces combinaisons linéaires,

$$\begin{aligned}
c_1 &:= a_1^1 u_1 + ... + a_1^k u_k \\
c_2 &:= a_2^1 u_1 + ... + a_2^k u_k \\
&\vdots \\
c_{k+1} &:= a_{k+1}^1 u_1 + ... + a_{k+1}^k u_k,
\end{aligned}$$

si tous les coefficiens a_i^{k+1} sont nuls, on est ramené à démontrer le résultat pour $k-1$ vecteurs formant une famille libre. Sinon on peut toujours supposer que a_1^k est non nul.
Les k vecteurs

$$c_2 - \frac{a_2^k}{a_1^k} c_1 ... c_{k+1} - \frac{a_{k+1}^k}{a_1^k} c_1$$

sont combinaisons linéaire des $k-1$ vecteurs $u_1 ... u_{k-1}$, il suffit alors de raisonner par récurrence sur k. Je laisse la suite en exercice.

---

[2] Max, mathématicien américain, 1906–1993

Case 2D5000, 292 rue Saint-Martin 75141 Paris Cedex 03
Bureau 17.0.14
Tel: 0158808765, Mel: thierry.horsin@cnam.fr

Conservatoire National des Arts et Métiers
Département d'Ingénierie Mathématique

Cette proposition étant démontrée. Une base constituant une famille génératrice, si l'une d'entre elle a N éléments, tout base contenant plus de $N+1$ élément serait une famille liée ce qui est faux. D'où le théorème.

1.3 sous-espace vectoriel

Définition 1.10 *Un sous-espace vectoriel d'un $\mathbb{K}$-espace vectoriel E est un sous-ensemble qui muni des lois faisant de E un $\mathbb{K}$-espace vectoriel est lui-même un $\mathbb{K}$-espace vectoriel .*

En particulier si F est un sous-espace vectoriel de E alors $\forall x, y \in F^2\ x+y \in F$. Plus précisément on a

Théorème 1.11 *Soit F un sous-ensemble d'un $\mathbb{K}$-espace vectoriel , F est un sous-espace vectoriel de E ssi*

$$\forall(\lambda, \mu) \in \mathbb{K}^2,\ \forall(x, y) \in F^2,\ \lambda x + \mu y \in F.$$

Il suffit d'écrire les définitions pour vérifier ce théorème.

Définition 1.12 *Soit F et G deux sous-espaces vectoriels de E, on dira que F et G sont en somme directe et on écrira $F \oplus G$ si*

$$F \cap G = \{0\}.$$

Supposons que F et G soient en somme directe, alors tout élément qui appartient à $F+G$ s'écrit de manière unique $f+g$ où $f \in F$ et $g \in G$.
En effet si $f+g = f'+g'$ avec $f, f' \in F$ et $g, g' \in G$ alors $f - f' = g' - g$ est un élément de $G \cap F$, c'est donc l'élément nul.

Définition 1.13 *On dira que F et G sont supplémentaires dans E si $F \oplus G$ et si $F+G = E$.*

Ainsi d'après la définition si F et G sont supplémentaires tout élément x de E s'écrit de manière unique $x = f+g$ avec $f \in F$ et $g \in G$. Et réciproquement l'unicité d'une telle écriture pour tout $x \in E$ implique que F et G sont supplémentaires.
On peut toujours construire un supplémentaire d'un sous-espace vectoriel . Démontrons le en dimension finie:

Théorème 1.14 [de la base incomplète]. *Si $e_1, ..., e_k$ constitue une famille libre d'un espace vectoriel de dimension N, il existe $N-k$ vecteurs $e_{k+1}, ..., e_N$ tels que $e_1, ..., e_N$ soit une base de E.*

Case 2D5000, 292 rue Saint-Martin 75141 Paris Cedex 03
Bureau 17.0.14
Tel: 0158808765, Mel: thierry.horsin@cnam.fr

Conservatoire National des Arts et Métiers
Département d'Ingénierie Mathématique

En effet: si pour tout vecteur $x \in E$, x est combinaison linéaire de $e_1, ..., e_k$ cela signifie que cette famille de k vecteurs est génératrice de E, mais E est alors de dimension k ce qu'il serait absurde de supposer.
Donc il existe x tel que la famille $e_1, ..., e_k, x$ soit une famille libre. On pose $e_{k+1} = x$. On recommence....
Si $e_1...e_k$ est une base de F alors on peut prendre comme supplémentaire le sous-espace vectoriel engendré par $e_{k+1}, ..., e_N$.
On dispose également du résultat suivant :

Théorème 1.15 *Les propriétés suivantes sont équivalentes:*

i. F *et* G *sont supplémentaires.*

ii. $F + G = E$ *et* $\dim(E) = \dim(F) + \dim(G)$.

iii. $\dim(E) = \dim(F) + \dim(G)$ *et* $F \cap G = \{0\}$.

En effet si F et G sont supplémentaires on a $i \Rightarrow iii$.
Voyons que $iii \Rightarrow ii$. Prenons une base B_1 de F et une base B_2 de G. La réunion de B_1 et de B_2 est une famille libre. Si ce n'est pas une base de E on la peut compléter en une base de E, mais alors $\dim(E) > \dim(F) + \dim(G)$.

2 Rappel: les applications linéaires

Soit E et F deux espaces vectoriels sur $\mathbb{R}$.
Soit $f : E \to F$ une application linéaire, c'est-à-dire une application f telle que

$$\forall (x, y) \in E \times E,\ \forall (\lambda, \mu) \in \mathbb{R} \times \mathbb{R},\ f(\lambda x + \mu y) = \lambda f(x) + \mu f(y).$$

Soit $(u_1, ..., u_N)$ une base de E et $(v_1, ..., v_M)$ une base de F.
Tout vecteur $x \in E$ s'écrit de manière unique sous la forme

$$x = \sum_{i=1}^{N} x_i u_i$$

et donc

$$f(x) = \sum_{i=1}^{N} x_i f(u_i).$$

Ainsi pour tout vecteur x, pour connaître $f(x)$, il suffit de connaitre les x_i qui sont les coordonnées de x dans la base $(u_1, ..., u_N)$ et chacun des vecteurs $f(u_1), ..., f(u_N)$. Ces mêmes vecteurs sont, quant à eux, déterminés par leurs coordonnées dans la base $(v_1, ..., v_M)$.

Case 2D5000, 292 rue Saint-Martin 75141 Paris Cedex 03
Bureau 17.0.14
Tel: 0158808765, Mel: thierry.horsin@cnam.fr

Conservatoire National des Arts et Métiers
Département d'Ingénierie Mathématique

On écrit donc $f(u_j) = \sum_{i=1}^{M} a_{ij} v_i$. Il s'ensuit que pour connaitre $f(x)$, il suffit de connaitre ces coefficients a_{ij} $i = 1...N, j = 1...M$ et les coordonnées de x dans la base $(u_1, .., u_N)$.
La coordonnée de $f(x)$ sur le vecteur v_i est $\sum_{j=1}^{N} a_{ij} x_j$.
Ces calculs montrent la propriété essentielle suivante qui permettra la construction des matrices

Théorème 2.1 *Une application linéaire est parfaitement déterminé par la connaissance de l'image de tous les vecteurs d'une base.*

2.1 Noyau d'une application linéaire

Soit $f : E \to F$ une application linéaire, on définit
$$\mathfrak{K}erf := \{u \in E, f(u) = 0\}.$$
C'est un sous-espace vectoriel de E car si u, v en sont des éléments

$$f(\lambda u + \mu v) = \lambda f(u) + \mu f(v) = 0.$$

Remarquons que 0 est toujours dans le noyau qui ne peut donc jamais être vide. En effet

$$f(\lambda 0) = f(0) = \lambda f(0).$$

Donc ceci implique que $f(0) = 0$.
Que dire si $\ker(f) = \{0\}$. C'est dire que $f(x) = 0$ implique $x = 0$. Donc si $f(x) = f(y)$ alors $f(x - y) = 0$ donc $x - y = 0$ donc $x = y$.
Finalement on voit que

Théorème 2.2 *Une application linéaire est injective ssi son noyau est réduit à $\{0\}$.*

2.2 Image d'une application linéaire

Soit $f : E \to F$ une application linéaire, on définit
$$\mathfrak{I}mf = \{f(x), x \in E\}$$
l'image de f. C'est un sous-espace vectoriel de F.
On dispose également du résultat

Théorème 2.3 *f est surjective ssi $\mathfrak{I}mf = F$.*

En effet si f est surjective, tout $y \in F$ s'écrit $f(x)$ pour un certain x.

Case 2D5000, 292 rue Saint-Martin 75141 Paris Cedex 03
Bureau 17.0.14
Tel: 0158808765, Mel: thierry.horsin@cnam.fr

Conservatoire National des Arts et Métiers
Département d'Ingénierie Mathématique

2.3 Théorème du rang

Soit E un $\mathbb{K}$-espace vectoriel de dimension finie. On a

Théorème 2.4 *Si f est une application linéaire de E dans F, alors*
$$\dim(E) = \dim(\ker(f)) + \dim(\Im(f)).$$

En effet, prenons H un supplémentaire du noyau de f dans E.
L'application linéaire $f : E \to F$ est surjective quand on la considère de $E \to \Im(f)$. Restreinte à H, elle est aussi injective, donc f définit une bijection de H sur $\Im(f)$.
Or une application linéaire bijective préserve les dimensions car l'image d'une base est alors une base.
Plus précisément:

Proposition 2.5 *Si $f : E \to F$ est injective, l'image par f d'une famille libre est libre, et si f est surjective, l'image par f d'une famille génératrice est génératrice. Si f est bijective, l'image d'une base est donc une base. Réciproquement si E et F ont même dimension, si l'image d'une base est une base alors f est bijective.*

Il suffit d'écrire les choses pour démontrer cette proposition.

3 Les matrices[3] [4]

Etant donné $f : E \to F$ une application linéaire, B_1 une base de E, B_2 une base de F, on appelle matrice de f relativement aux bases B_1 de E et B_2 de F le tableau à M lignes et N colonnes

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & \dots & a_{1N} \\ a_{21} & a_{22} & a_{23} & \dots & \dots & a_{2N} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{M1} & a_{M2} & a_{M3} & \dots & \dots & a_{MN} \end{pmatrix}$$

Si x est un vecteur de E, pour déterminer les coordonnées de $(y_1, ..., y_M)$ $f(x)$ dans la base F connaissant la matrice de f relativement aux bases choisies, on écrit A la matrice de f,

[3] Attribuable à Cayley
[4] Cayley Arthur, mathématicien anglais, 1821–1895

Case 2D5000, 292 rue Saint-Martin 75141 Paris Cedex 03
Bureau 17.0.14
Tel: 0158808765, Mel: thierry.horsin@cnam.fr

Conservatoire National des Arts et Métiers
Département d'Ingénierie Mathématique

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ \vdots \\ y_M \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & \dots & a_{1N} \\ a_{21} & a_{22} & a_{23} & \dots & \dots & a_{2N} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{M1} & a_{M2} & a_{M3} & \dots & \dots & a_{MN} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ \vdots \\ \vdots \\ x_N \end{pmatrix},$$

on se déplace horizontalement dans la matrice A à la profondeur j, chacun des coefficients étant multipliés par la coordonnée respective de x.

On voit donc qu'une fois les bases fixées dans les e.v de départ et d'arrivée, se donner une application linéaire ou se donner une matrice est identique. *Deux matrices identiques ne représentent donc pas nécessairement la même application linéaire sauf si les matrices ont été construites dans les mêmes bases.*

3.1 Remarques

-Une fois les bases fixées dans E et F on voit grâce à la représentation matricielle d'une application linéaire que l'on peut toujours supposer que $E = \mathbb{R}^N$ et $F = \mathbb{R}^M$.
-Il est indispensable de préciser les bases dans lesquelles est exprimée la matrice.
Lorsque $E = F$, on utilise fréquemment la même base pour E comme espace de départ et d'arrivée.
On notera $\mathscr{M}^{M,N}(\mathbb{R})$ l'ensemble des matrices réelles ayant M lignes et N colonnes. On peut définir la somme de deux telles matrices, le produit d'une telle matrice par un réel qui font de cet ensemble un espace vectoriel de dimension $M \times N$.
On dit qu'une matrice est carrée si le nombre de ses lignes égale le nombre de ses colonnes.
Une matrice carrée représente un endomorphisme.

3.2 Transposée

Soit $f : E \to F$ une application linéaire. On appelle ${}^t f$ l'application transposée de f définie par

$${}^t f : F^* \to E^*,$$
$${}^t f(v^*)(x) = v^*(f(x)).$$

Si on exprime la matrice de ${}^t f$ dans la base duale $v_1^*, ..., v_M^*$ de F^* et la base $u_1^*, ..., u_N^*$ de E^* cette matrice est la transposée de A donnée par

$${}^t A = \begin{pmatrix} a_{11} & a_{21} & \dots & \dots & a_{M1} \\ a_{12} & a_{22} & \dots & \dots & a_{M2} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{1N} & a_{2N} & \dots & \dots & a_{MN} \end{pmatrix}$$

Case 2D5000, 292 rue Saint-Martin 75141 Paris Cedex 03
Bureau 17.0.14
Tel: 0158808765, Mel: thierry.horsin@cnam.fr

Conservatoire National des Arts et Métiers
Département d'Ingénierie Mathématique

Théorème 3.1 *Une matrice et sa transposée ont le même rang.*

Soit $y^* \in \ker({}^t f)$. Alors $\forall x \in E$ on a $y^*(f(x)) = 0$.
Donc $y^* \in (\mathrm{Im}()f)^\perp$ et réciproquement. On conclut donc par le théorème du rang.
-Comment interprète-t-on le rang en terme de matrices ?
Rappelons que le rang est la dimension de l'image. C'est donc le nombre maximum de vecteurs linéairement indépendants dans l'image. Il s'ensuit que le rang est le nombre de colonnes linéairement indépendantes de la matrice représentant une application linéaire.
Comme le rang d'une application linéaire et de sa transposée sont les mêmes, on en déduit que le rang d'une matrice est aussi le nombre maximum de lignes linéairement indépendantes.

4 Produit de matrices

Soit $f : E \to F$ et $g : F \to G$ deux applications linéaires sur des espaces vectoriels de dimension finie respectives N, M, P. L'application $g \circ f : E \to G$ est également une application linéaire. Si l'on a choisi une base $\mathscr{B}$, $\mathscr{B}_1$, $\mathscr{B}_2$ de E, F, et G respectivement, et que l'on considère les matrices de f et g, A et B. Comment trouve-t-on la matrice de $g \circ f$?
En se souvenant que les colonnes de la matrice de $g \circ f$ sont les coordonnées des images des éléments de $\mathscr{B}$ exprimés dans la base $\mathscr{B}_2$.Il faut donc calculer $g(f(u_1)), .., g(f(u_N))$. On doit donc faire N produit de B par les N colonnes successives de A.
Pour ce faire on juxtapose donc la matrice A à la droite de B.

$$BA = \begin{pmatrix} b_{11} & b_{12} & \dots & \dots & b_{1M} \\ b_{21} & b_{22} & \dots & \dots & b_{2M} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ b_{P1} & b_{P2} & \dots & \dots & b_{PM} \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & \dots & a_{1N} \\ a_{21} & a_{22} & a_{23} & \dots & \dots & a_{2N} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{M1} & a_{M2} & a_{M3} & \dots & \dots & a_{MN} \end{pmatrix}$$

4.1 Remarque importante

Le produit de matrices n'est pas commutatif.
En effet, il faut que le nombre de colonnes de la matrice à gauche égale le nombre de lignes de la matrice à droite.

4.2 Associativité

On dispose du résultat suivant

Case 2D5000, 292 rue Saint-Martin 75141 Paris Cedex 03
Bureau 17.0.14
Tel: 0158808765, Mel: thierry.horsin@cnam.fr

Conservatoire National des Arts et Métiers
Département d'Ingénierie Mathématique

Théorème 4.1 *Si chacun des produits matriciels ci-dessous est licite, alors*
$$(CB)A = C(BA).$$

Il suffit de faire le calcul pour vérifier cette propriété.

5 Déterminant[5] [6]

5.1 Définition

Dans $\mathbb{R}^N$, on pose
$$\langle u, v\rangle = {}^t u.v,$$
et
$$\|u\| = \sqrt{\langle u, v\rangle}.$$
Ces objets permettent de définir un produit scalaire[7]. Cependant la théorie des déterminants peut se définir sans ces notions. Nous l'allons présenter cependant de manière euclidienne[8] c'est à dire en utilisant la mesure des vecteurs.
Soit $f : E \to E$ une application linéaire et A sa matrice exprimée dans une base de E.
Supposons que $\dim(E) = 2$.
On appelle
$$\det(A) = a_{11}a_{22} - a_{12}a_{21}.$$
Comme les colonnes de A sont les coordonnées des vecteurs $f(u_1)$,$f(u_2)$ exprimés dans la base (u_1, u_2), le déterminant mesure le volume du parallélélogramme formé par les vecteurs de $\mathbb{R}^2$ dont les coordonnées sont $\begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix}$ et $\begin{pmatrix} a_{21} \\ a_{22} \end{pmatrix}$. Ce volume est compté positivement lorsque on passe du premier au second par un angle inférieur à l'angle plat positif, sinon il est compté négativement.
On peut vérifier que $\det(AB) = \det(BA) = \det(A)\det(B)$ si bien sûr le produit est bien défini pour AB et BA, donc si B est carrée de taille 2×2.
Comment interprêter cette égalité ?
Le déterminant d'une matrice mesure le volume des images de la base usuelle de $\mathbb{R}^2$.
L'image de cette base par A définit un parallélogramme. Ce parallélogramme Notons w_1 et w_2 les vecteurs colonnes de A. Si ces deux vecteurs sont colinéaires, alors l'égalité est vérifiée.

[5] Notion attribuée à Cardan ou Cardano
[6] Cardano, Gerolamo, mathématicien italien, 1501–1576
[7] La définition moderne du produit scalaire est difficile à attribuer
[8] Euclide, mathématicien grec, dates de naissance et décès imprécises.

Case 2D5000, 292 rue Saint-Martin 75141 Paris Cedex 03
Bureau 17.0.14
Tel: 0158808765, Mel: thierry.horsin@cnam.fr

Conservatoire National des Arts et Métiers
Département d'Ingénierie Mathématique

On écrit $w_1 = ae_1 + be_2$ et $w_2 = ce_1 + de_2$ (en fait a, b, c, d sont les coefficients de A). $Bw_1 = aBe_1 + bBe_2$ et $Bw_2 = cBe_1 + dBe_2$. Le volume déterminé par Bw_1 et Bw_2 a pour valeur $(ad - bc)\det(Be_1, Be_2) = \det(A)\det(B)$. Donc $\det(AB) = \det(BA) = \det(B)\det(A)$.
Supposons avoir défini le déterminant de toutes les matrices de tailles $p \times p$ avec $p < N$.
Posons A_1 la matrice carrée obtenue en rayant la première ligne et la première colonne de $(C_1, ...-..., C_N)$, A_2 la matrice carrée obtenue en rayant la deuxième ligne et la première colonne,.....
Dans la matrice A_1 on raye également les lignes ce qui donne des matrices B_{1j}.
On pose

$$\det(C_1, ..., C_N) = a_{11}\det(A_1) - a_{21}\det(A_2) + a_{31}\det(A_3).... + (-1)^N a_{N1}\det(A_N).$$

Remarquons que

$$\begin{aligned}
\det(A_1) &= \sum_{j\neq 1}(-1)^j a_{j2}\det(B_{1j}),\\
\det(A_2) &= \sum_{j<2}(-1)^{j-1} a_{j2}\det(B_{2j}) + \sum_{j>2}(-1)^j a_{j2}\det(B_{2j}),\\
\det(A_k) &= \sum_{j<k}(-1)^{j-1} a_{j2}\det(B_{2j}) + \sum_{j>k}(-1)^j a_{j2}\det(B_{2j}).
\end{aligned}$$

Donc

$$\begin{aligned}
\det(C_1, ..., C_N) = \sum_{k=1}^{N}\sum_{j<k}(-1)^k(-1)^{j-1}a_{k1}a_{j2}\det(B_{2j})\\
+ \sum_{k=1}^{N}\sum_{j>k}(-1)^k(-1)^j a_{k1}a_{j2}\det(B_{2j})
\end{aligned}.$$

On voit donc qu'on aurait pu prendre comme définition du déterminant le développement par rapport à la deuxième colonne, en renversant les signes.
Il s'ensuit que

$$\det(C_1, C_2, ..., C_N) = -\det(C_2, C_1, ..., C_N).$$

Donc $\det(C_1, C_1, C_3, ..., C_N) = 0$.
De plus si on suppose par récurrence que toute permutation de deux colonnes du déterminant change son signe en son opposé pour tout déterminant de p vecteurs à p coordonnées avec $p < N$, il en est de même pour le déterminant de N vecteurs à N coordoonnées.
On vérifie sans peine grâce à la définition de $\det(C_1, ., \lambda C_i, .., C_N) = \lambda\det(C_1, ..., C_N)$.
On identifiera toujours par la suite, le déterminant de N vecteurs de $\mathbb{R}^N$ avec le déterminant de la matrice formée par ces vecteurs rangés dans le même ordre.
Il s'ensuit de cette propriété que l'on peut calculer un déterminant à partir de n'importe quelle colonne en faisant glisser cette colonne à la première place et laissant ordonner les $N - 1$ autres. On observera que cela induit pour la $k$ième colonne $k - 1$ changement de signe. Il faut donc multiplier le résultat du déterminant calculé par $(-1)^k$.

Case 2D5000, 292 rue Saint-Martin 75141 Paris Cedex 03
Bureau 17.0.14
Tel: 0158808765, Mel: thierry.horsin@cnam.fr

Conservatoire National des Arts et Métiers
Département d'Ingénierie Mathématique

5.2 Interprétation du déterminant

5.2.1 Dans $\mathbb{R}^2$

Dans $\mathbb{R}^2$ le déterminant de deux vecteurs est l'aire algébrique du parallélogramme formé par ces deux vecteurs.

En effet soit $\vec{u}\begin{pmatrix} a \\ b \end{pmatrix}$ et $\vec{v}\begin{pmatrix} c \\ d \end{pmatrix}$. On écrit $\vec{v} = \langle \vec{u}, \vec{v} \rangle \frac{\vec{u}}{\| \vec{u} \|^2} + \vec{w}$.

Donc la surface du parallélogramme est $\| \vec{u} \| \| \vec{w} \|$.

On a $\vec{w} = \begin{pmatrix} c - a(ac+bd)\frac{1}{a^2+b^2} \\ d - b(ac+bd)\frac{1}{a^2+b^2} \end{pmatrix}$

Donc

$$\vec{w} = \begin{pmatrix} \frac{cb^2 - abd}{a^2+b^2} \\ \frac{da^2 - bac}{a^2+b^2} \end{pmatrix},$$

$$\| \vec{w} \|^2 = \frac{(c^2b^4 + a^2b^2d^2 - 2adcbb^2 + d^2a^4 + b^2a^2c^2 - 2abcda^2)}{(a^2+b^2)^2} = \frac{(ad-bc)^2}{a^2+b^2}.$$

Ainsi la quantité trouvée est bien l'aire (absolue) du parallélogramme en question. Reste le signe à déterminer.

La convention de la positivité est que si l'on va de $\vec{u}$ à $\vec{v}$ en faisant moins d'un angle plat dans le sens trigo, l'aire est comptée positivement.

Remarquons que $\vec{w} = \frac{ad-bc}{a^2+b^2}\begin{pmatrix} -b \\ a \end{pmatrix}$.

Or $\begin{pmatrix} -b \\ a \end{pmatrix}$ est un vecteur directement orthogonal à $\vec{u}$ ce qui implique la convention de signe.

5.2.2 Dans $\mathbb{R}^3$

Soit $\vec{u}, \vec{v}, \vec{w}$ trois vecteurs de $\mathbb{R}^3$.

On observe que $\det(\vec{u}, \vec{v}, \vec{w}) = \vec{u}.(\vec{v} \wedge \vec{w})$ et donc là encore le déterminant est la mesure du volume algébrique déterminé par les trois vecteurs $\vec{u}, \vec{v}, \vec{w}$.

5.2.3 Dans $\mathbb{R}^N$

On dit que le déterminant de N vecteurs de $\mathbb{R}^N$ est la mesure du volume algébrique déterminé par ces N vecteurs.

Case 2D5000, 292 rue Saint-Martin 75141 Paris Cedex 03
Bureau 17.0.14
Tel: 0158808765, Mel: thierry.horsin@cnam.fr

Conservatoire National des Arts et Métiers
Département d'Ingénierie Mathématique

5.3 Déterminant d'un produit

On se propose de démontrer que si A et B sont deux matrices carrées de même tailles, alors

$$\det(AB) = \det(BA) = \det(B)\det(A).$$

Pour ce faire, on écrit

$$AB(e_1) = A(B(e_1)) = \sum_{i=1}^{N} A(b_{i1}e_i),$$

donc

$$AB(e_1) = \sum_{i=1}^{N} b_{i1}Ae_i, ..., AB(e_k) = \sum_{i=1}^{N} b_{ik}Ae_i, ..., AB(e_N) = \sum_{i=1}^{N} b_{iN}Ae_i.$$

On a donc

$$\det(AB) = \sum_{i_1=1}^{N}\sum_{i_2=1}^{N}...\sum_{i_N=1}^{N} b_{i_1 1}b_{i_2 2}...b_{i_N N}\det(A(e_{i_1}), ..., A(e_{i_N})).$$

Dans le second membre il est nécessaire que tous les i_k soient deux à deux distincts. A une permutation près on a donc $\det(A)$

Donc

$$\det(AB) = \det(A)\sum_{(i_1,i_2,...,i_N)\in\{1,...,N\}, i_k\neq i_m \text{ si } m\neq n} \pm b_{i_1 1}b_{i_2 2}...b_{i_N N}.$$

On a un signe $+$ lorsque on est passé de $\det(A(e_{i_1}), ..., A(e_{i_N}))$ à $\det(A)$ avec un nombre pair de permutations de colonnes et $-$ sinon. On voit alors facilement par récurrence que la quantité trouvée est $\det(B)$. Notons au passage que ceci démontre que $\det(B) = \det({}^tB)$.

Finalement

$$\det(AB) = \det(A)\det(B).$$

6 Matrices inversibles.

Dans tout le paragraphe E désigne un ev de dim finie N, et $f : E \to E$ une application linéaire (ici un endomorphisme) dont on considère la matrice A dans une base $e_1, ..., e_N$ fixée une fois pour toute.

6.1 Généralités

On dira qu'une matrice carrée A est inversible, si l'application linéaire qu'elle représente est bijective.

Rappelons que pour E espace vectoriel de dimension finie, on dispose de

Théorème 6.1 *Soit $f : E \to E$ une application linéaire, alors f est bijective ssi f est injective ssi f est surjective.*

Case 2D5000, 292 rue Saint-Martin 75141 Paris Cedex 03
Bureau 17.0.14
Tel: 0158808765, Mel: thierry.horsin@cnam.fr

Conservatoire National des Arts et Métiers
Département d'Ingénierie Mathématique

Ainsi pour qu'une matrice soit inversible il faut et il suffit que les colonnes de cette matrice soient linéairement indépendantes donc ssi $\det(A) \neq 0$.
Cette dernière équivalence n'est pas prouvée formellement pour l'instant mais repose sur l'interprétation du déterminant en tant que volume du parallélépipède formé par les vecteurs.
Démontrons cette affirmation plus formellement.

Théorème 6.2 *Le déterminant d'une matrice est nulle ssi elle représente une application linéaire non bijective.*

On procède par récurrence sur la dimension, sachant qu'on a vérifié ces propriétés pour $N = 1$ où rien n'est à faire, $N = 2$ car le déterminant mesure le degré de proportionnalité entre deux vecteurs.
On prend $N \geq 3$ et on suppose avoir démontré la propriété pour toutes les applications linéaires de n'importe quel espace vectoriel de dimension $< N$ à valeurs dans lui-même.
Supposons $a_{11} \neq 0$. Si tel n'est pas le cas, et si tous les éléments de la première colonne de A sont nuls alors $\det(A) = 0$ et l'image de e_1 est par l'application linéaire représentée est nulle donc f est non injective.
Donc, si $\det(A) \neq 0$ l'un au moins des éléments de la première ligne est $\neq 0$. Supposons par exemple que ce soit a_{12}
Soit P_{12} la matrice dont tous les coefficients sont nuls sauf $p_{21} = 1$ $p_{12} = 1$ et $a_{jj} = 1$ pour tout $j = 3, ..., N$. Un calcul immédiat montre que $\det(P_{12}) = \pm 1$.
Par la propriété du produit $\det(AP_{12}) = \pm \det(A)$ et on observe que $P_{21}A$ a son premier coefficient valant a_{12}.
Supposons donc que $a_{11} \neq 0$.
Considérons la matrice M_{12} dont tous les coefficients sont nuls sauf les éléments diagonaux qui valent 1 et $m_{12} = -\dfrac{a_{21}}{a_{11}}$.
Observons que $\det(AM_{12}) = \det(A)$ car M_{12} a clairement un déterminant égal à 1. Dans $M_{21}A$ le coefficient de position $(1, 2)$ est nul.
Procédant ainsi sans changer le déterminant de A, à la suite d'au plus N transformations, on est ramené à une matrice où la première ligne est

$$\begin{pmatrix} 1 & 0 & \dots & \dots & 0 \end{pmatrix}$$

Donc le déterminant de A est égal au déterminant d'une matrice carrée de taille $N - 1 \times N - 1$ donc les colonnes sont les coordonnées de

$$Ae_2 - \frac{a_{12}}{a_{11}}Ae_1, ..., Ae_N - \frac{a_{1N}}{a_{11}}Ae_1,$$

le long des vecteurs $e_2, ..., e_N$.
Si $\det(A) = 0$ alors notre hypothèse de récurrence affirme que ces $N - 1$ colonnes sont liées.
Il existe donc une combinaison linéaire nulle de ces vecteurs

$$\alpha_2(Ae_2 - \frac{a_{12}}{a_{11}}Ae_1) + ... + \alpha_N(Ae_N - \frac{a_{1N}}{a_{11}}Ae_1) = 0,$$

Case 2D5000, 292 rue Saint-Martin 75141 Paris Cedex 03
Bureau 17.0.14
Tel: 0158808765, Mel: thierry.horsin@cnam.fr
Conservatoire National des Arts et Métiers
Département d'Ingénierie Mathématique

où au moins l'un des $\alpha_2, ..., \alpha_N$ est non nul. Cette combinaison linéaire est une combinaire linéaire nulle des vecteurs

$$Ae_1, ..., Ae_N$$

où l'un des coefficients au moins est non nul.
Il s'ensuit que la matrice A représente une application linéaire non inversible puisque son image est de dimension $\leq N-1$.
Réciproquement s'il existe une combinaise linéaire nulle

$$\alpha_1 Ae_1 + \alpha_2 Ae_2 + ... + \alpha_N Ae_N$$

où au moins un des coefficients est non nul, c'est à dire si l'image de l'application linéaire représentée par A est de dimension $\leq N-1$ alors on peut toujours supposer que

$$Ae_1 = \sum_{i=2}^{N} \mu_i Ae_i.$$

Compte tenu des propriétés du déterminant (linéarité par rapport à une colonne – ou une ligne – et nullité si deux colonnes –ou deux lignes– sont égales), il s'ensuit que $\det(A) = 0$. Le théorème est donc démontré.
Notons qu'il est *intuitivement* plus fréquent d'avoir un nombre non nul que nul, il s'ensuit qu'il est *intuitivement* rare qu'une matrice ne soit pas inversible. Mais la rareté ne signifie pas non-existence.

6.2 Calcul de l'inverse

Si f est bijective alors f^{-1} existe et on notera A^{-1} la matrice de f^{-1} relative à la base choisie. Dans tout ce paragraphe, f sera supposé bijective.
Puisque $f \circ f^{-1} = f^{-1} \circ f = Id_E$, on en déduit que

$$AA^{-1} = A^{-1}A = I.$$

Comment calculer A^{-1} ?
Plusieurs méthodes sont possibles au moins théoriquement mais le calcul effectif lorsque N est grand s'avère toujours (ou presque) très difficile sans propriétés supplémentaires de A.
Je renvoie à des cours plus spécialisés pour la mise en oeuvre de ces méthodes (et leurs explications), mais il est important d'avoir en tête ce délicat problème dans l'utilisation des matrices.
Pour déterminer A^{-1}, l'on se doit de trouver $g_1, ..., g_N$ tels que $Ag_1 = e_1, ..., Ag_N = e_N$.

6.2.1 Formules de Cramer[9]

Elles sont essentiellement inutilisables d'un point de vue pratique mais présentent l'avantage de ne pas l'être dans certaines démonstrations.

[9] Gabriel, Mathématicien suisse, 1704–1752

Case 2D5000, 292 rue Saint-Martin 75141 Paris Cedex 03
Bureau 17.0.14
Tel: 0158808765, Mel: thierry.horsin@cnam.fr
Conservatoire National des Arts et Métiers
Département d'Ingénierie Mathématique

Observons que $\det(A) = \langle Ae_1, y_1\rangle$ où y_1 est un vecteur dont les coordonnées sont constitués de $\pm$ des déterminants de tailles $N-1 \times N-1$. Chacun de ces déterminants affecté du signe est appelé cofacteur de la coordonnée correspondantes de Ae_1.
Noter que si l'on calcule $\langle Ae_k, y_1\rangle$ pour $k = 2, .., N$ cela revient à calculer le déterminant d'une matrice dont la première colonne est l'une des $N-1$ dernières, c'est donc une quantité nulle.
Procédons de même pour chacune des colonnes de la matrice et déterminons en développant le déterminant par rapport à la k-ième colonne un vecteur colonne y_k tel que $\det(A) = \langle Ae_k, y_k\rangle$.
De même que pour y_1, si l'on remplace Ae_k par y_k dans la matrice A le déterminant de la matrice obtenue est nul.
Il s'ensuit que la matrice $\tilde{A} = {}^t\begin{pmatrix} y_1 & y_2 & \dots & \dots & y_N \end{pmatrix}$ satisfait à

$$\tilde{A}A = \det(A) = A\tilde{A}.$$

Finalement

$$A^{-1} = \frac{1}{\det(A)}\tilde{A},$$

formule souvent résumée en disant que l'inverse d'une matrice est ``un sur déterminant de A fois transposée de la comatrice''.

6.2.2 Autres méthodes

Toute méthode faisant intervenir le déterminant d'une matrice dont beaucoup de coefficients non diagonaux ne sont pas nuls est vouée à l'échec d'un point de vue calculatoire.
On prend donc le partie de résoudre les N systèmes $Ag_1 = e_1, ..., Ag_N = e_N$.
La méthode du pivot de Gauss[10] est assez efficace pour les résoudre et a l'avantage de se programmer assez simplement.

6.3 Déterminant de l'inverse

Compte tenu que $A.A^{-1} = I$ alors $\det(A^{-1}) = \dfrac{1}{\det(A)}$.
Hors ``programme'': Application: si A est une matrice, considérons la matrice $A - XI$ à coefficients dans le corps $K(X)$. Son déterminant est non nul.
On a $(XI - A) * \mathrm{com}(A - XI_N) = \det(XI - A)I_N$, où la matrice $\mathrm{com}(A - X_N I)$ désigne la comatrice de $A - XI$.
Toutes les quantités qui apparaissent sont des polynômes en X à coefficients.
Si on écrit $\mathrm{com}(A - XI) = X^{N-1}P_{N-1}(A) + ... + P_0(A)$, et que l'on développe

$$(XI - A)(X^{N-1}P_{N-1}(A) + ... + P_0(A))$$

[10] Carl Friedrich, mathématicien allemand, 1777–1855

Case 2D5000, 292 rue Saint-Martin 75141 Paris Cedex 03
Bureau 17.0.14
Tel: 0158808765, Mel: thierry.horsin@cnam.fr
Conservatoire National des Arts et Métiers
Département d'Ingénierie Mathématique

et qu'on identifie à $\det(XI_N - A) = (X^N + a_{N-1}X^{N-1} ... + a_0)I$ on obtient

$$\begin{aligned} &P_{N-1}(A) = I_N \\ &-AP_{N-1}(A) + P_{N-2}(A) = a_{N-1}I_N \\ &-AP_{N-2}(A) + P_{N-3}(A) = a_{N-2}I_N \\ &\vdots \\ &-AP_1(X) + P_0(A) = a_1 I_N \\ &-AP_0(A) = a_0 I_N \end{aligned}$$

En multipliant par A l'avant-dernière, A^2 l'antepénultième,...,A^N la première et en ajoutant, le membre de gauche fait 0 et celui de droite est la valeur du polynôme caractéristique de f en A, ce qui démontre le théorème de Cayley-Hamilton[11] énoncé et démontré autrement plus loin.

7 Changement de bases

Soit $f : E \to F$ une application linéaire entre deux espaces de dimension finie N et M.
Soit A la matrice de f exprimée dans deux bases $\mathscr{B}_1$ de E et $\mathscr{B}_2$ de F. Soit ${\mathscr{B}_1}'$ et ${\mathscr{B}_2}'$ deux autres bases de E et F respectivement.
Notons P la matrice constituée des vecteurs de ${\mathscr{B}_1}'$ exprimées dans la base $\mathscr{B}_1$ et Q la matrice correspondante dans F.
Soit u un vecteur de E et X les coordonnées de u dans la base $\mathscr{B}_1$ et X' dans la base ${\mathscr{B}_1}'$. On a $u = \sum u_i e_i = \sum u'_i e_{\prime} i$
Comme ${e_i}' = Pe_i$ on obtient $X = PX'$, ou encore $X' = P^{-1}X$.
De même $Y' = Q^{-1}Y$. Donc $A' = Q^{-1}AP$.

8 Réduction des endomorphismes

8.1 Polynôme d'endomorphisme

Soit P un polynôme à coefficients réels. Et soit $f : E \to E$ un endomorphisme.
On écrit $P(X) = a_0 + a_1X + a_2X^2 + ... + a_pX^p$ avec $a_p \neq 0$, p étant donc le degré de P.
On définit l'endomorphisme $P(f)$

8.2 Diagonalisation

Soit $f : E \to E$ une application linéaire.

[11] William, mathématicien irlandais, 1805–1865

Case 2D5000, 292 rue Saint-Martin 75141 Paris Cedex 03
Bureau 17.0.14
Tel: 0158808765, Mel: thierry.horsin@cnam.fr
Conservatoire National des Arts et Métiers
Département d'Ingénierie Mathématique

Définition 8.1 *Soit $u \neq 0$ un vecteur de E. On dira que u est un vecteur propre de E ssi, $\exists \lambda \in \mathbb{R}$ tel que $f(u) = \lambda u$.*

On dit alors que λ est la valeur propre associée à u.
Il revient alors au même de dire que $f - \lambda Id_E$ est une application linéaire non injective.
Ainsi on dispose du

Théorème 8.2 *λ est une valeur propre de f ssi $f - \lambda Id_E$ est non injective, ssi $f - \lambda Id_E$ est non surjective.*

Il s'ensuit que les valeurs propres sont calculables en résolvant

$$\det(A - \lambda I) = 0.$$

Définition 8.3 *On dira que f est diagonalisable si il existe une base $\mathscr{B}$ de E formée de vecteurs propres de f.*

Il existe des applications linéaires qui ne sont pas diagonalisables.
Si f est diagonalisable, alors la matrice de f dans la base de vecteurs propres est diagonale et les éléments diagonaux sont les valeurs propres de f.

Définition 8.4 *On appelle polynôme caractéristique de f, la fonction de X donnée par*

$$\det(XI - A).$$

Comme pour une matrice inversible P on a $\det(P^{-1}) = \dfrac{1}{\det(P)}$, le polynôme caractéristique de f ne dépend que de f.
Nous allons démontrer un théorème fondamental pour la suite:

Théorème 8.5 [de Cayley-Hamilton]. *Soit f un endomorphisme et soit P son polynôme caractéristique, alors $P(f) = 0$.*

Pour ce faire, prenons $x \in E$. Nous allons montrer que $P(f)(x) = 0$.
Pour cela, on considère la famille $x, f(x), f^2(x), \ldots$. Il existe un entier $p \leq N$ tel que la famille $x, f(x), \ldots, f^{p-1}(x)$ soit libre mais pas $x, \ldots, f^p(x)$. Complétons la famille $x, f(x), \ldots, f^{p-1}(x)$ en une base de E et exprimons la matrice A de f dans cette base.
Elle a la forme

Case 2D5000, 292 rue Saint-Martin 75141 Paris Cedex 03
Bureau 17.0.14
Tel: 0158808765, Mel: thierry.horsin@cnam.fr

Conservatoire National des Arts et Métiers
Département d'Ingénierie Mathématique

$$A=\begin{pmatrix} 0 & 0 & \dots & \dots & \dots & \dots & \dots & 0 & a_0 & L_0 \\ 1 & 0 & 0 & \dots & \dots & \dots & \dots & 0 & a_1 & L_1 \\ 0 & 1 & 0 & 0 & \dots & \dots & \dots & 0 & a_2 & L_2 \\ 0 & 0 & 1 & 0 & 0 & \dots & \dots & 0 & a_3 & L_3 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \dots & \vdots & \vdots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & 0 & 0 & a_{p-4} & L_{p-4} \\ \vdots & \ddots & \ddots & \ddots & \ddots & 1 & 0 & 0 & a_{p-3} & L_{p-3} \\ 0 & 0 & \dots & \dots & \dots & \dots & 1 & 0 & a_{p-2} & L_{p-2} \\ 0 & 0 & 0 & \dots & \dots & \dots & 0 & 1 & a_{p-1} & L_{p-1} \\ 0 & 0 & 0 & 0 & \dots & \dots & \dots & 0 & 0 & L_p \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \dots & \dots & \dots & 0 & L_N \end{pmatrix}$$

où $L_0,\dots,L_N$ sont des lignes à $N-p$ éléments.
Si on calcule le polynôme caractéristique (qui ne dépend pas de la base)
on trouve $\det(C_x)\det(B-XI_{N-p})$ où B est la matrice formée par les $N-p$ lignes $L_p,\dots,L_N$.
En effet quand on développe le déterminant par rapport à la première linge, tout se passe comme si on développer d'abord le déterminant de la matrice

$$\begin{pmatrix} -X & 0 & \dots & \dots & \dots & \dots & \dots & 0 & a_0 \\ 1 & -X & 0 & \dots & \dots & \dots & \dots & 0 & a_1 \\ 0 & 1 & -X & 0 & \dots & \dots & \dots & 0 & a_2 \\ 0 & 0 & 1 & -X & 0 & \dots & \dots & 0 & a_3 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \dots & \vdots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & 0 & 0 & a_{p-4} \\ \vdots & \ddots & \ddots & \ddots & \ddots & 1 & -X & 0 & a_{p-3} \\ 0 & 0 & \dots & \dots & \dots & \dots & 1 & -X & a_{p-2} \\ 0 & 0 & 0 & \dots & \dots & \dots & 0 & 1 & a_{p-1}-X \end{pmatrix}$$

et ce développement fait

$$-X\begin{vmatrix} -X & 0 & \dots & \dots & \dots & \dots & 0 & a_1 \\ 1 & -X & 0 & \dots & \dots & \dots & 0 & a_2 \\ 0 & 1 & -X & 0 & \dots & \dots & 0 & a_3 \\ \ddots & \ddots & \ddots & \ddots & \ddots & \dots & \vdots & \vdots \\ \ddots & \ddots & \ddots & \ddots & \ddots & 0 & 0 & a_{p-4} \\ \ddots & \ddots & \ddots & \ddots & 1 & -X & 0 & a_{p-3} \\ 0 & \dots & \dots & \dots & \dots & 1 & -X & a_{p-2} \\ 0 & 0 & \dots & \dots & \dots & 0 & 1 & a_{p-1}-X \end{vmatrix}+(-1)^{p-1}a_0.$$

On voit alors facilement par récurrence que l'on obtient

$$(-1)^p(X^p-a_{p-1}X^{p-1}-\dots-a_0).$$

Or précisément on a choisi $a_0,\dots,a_{p-1}$ pour que

$$f^p(x)=a_{p-1}f^{p-1}(x)+\dots+a_1f(x)+a_0x,$$

Case 2D5000, 292 rue Saint-Martin 75141 Paris Cedex 03
Bureau 17.0.14
Tel: 0158808765, Mel: thierry.horsin@cnam.fr
Conservatoire National des Arts et Métiers
Département d'Ingénierie Mathématique

ce qui finit de démontrer le théorème, puisque $P(f)(x) = det(B - XI_{N-p})_{X=f} * (f^p - a_{p-1}f^{p-1} - ... - a_0 I)(x)$.
Le polynôme $X^p - a_{p-1}X^{p-1} - ... - a_0$ s'appelle annulateur minimal de x pour f.
On vient de plus de prouver que l'annulateur minimal de x pour f divise le polynôme caractéristique.

Définition 8.6 *Soit λ une valeur propre de f, on appelle sous-espace propre associé à la valeur propre λ, $\mathfrak{K}er(f - \lambda I)$.*

Définition 8.7 *On dira que f est scindé sur $\mathbb{R}$ ssi $\det(XI - A)$ a toutes ses racines réelles.*

Supposons que f soit scindé.
Cela signifie que (notant $P(X) = \det(XI - A)$) on peut écrire $P(X) = (X - \lambda_1)^{\nu_1}....(X - \lambda_k)^{\nu_k}$.

Définition 8.8 *On appelle sous-espace caractéristique associé à λ_1*

$$\ker(f - \lambda_1 I)^{\nu_1}.$$

Pour chacune des valeurs propres, on peut définir le sous-espace caractéristique associé: si λ est cette valeur propre, et si ν est l'exposant de $(X - \lambda)$ dans le polynôme caractéristique, c'est $\ker(f - \lambda I)^{\nu}$.

Remarquons que $\ker(f - \lambda I) \subset \ker(f - \lambda I)^{\nu}$.
Afin de poursuivre, il est nécessaire de faire un peu d'arithmétique des polynômes.

8.2.1 L'anneau $\mathbb{K}[X]$

C'est l'ensemble des polynômes à coefficients dans $\mathbb{K}$. Je ne développerai pas la structure d'anneau, mais dis juste que dans un anneau les éléments non nuls ne sont pas tous inversibles.
Dans $\mathbb{K}[X]$ on dispose d'une division euclidienne et plus précisément on a

Théorème 8.9 *Soit P_1 et P_2 deux polynômes, il existe alors Q et R uniques tels que $P_1 = P_2 Q + R$ avec $\deg(R) < \deg(P_2)$*

La démonstration est immédiate par récurrence sur $\deg(P_1)$.
Etant donné $P_1, ..., P_k$, k polynômes. Regardons le sous-ensemble de $\mathbb{K}[X]$ formé par

$$I = \{R_1 P_1 + ... + R_k P_k,\ R_i \in \mathbb{K}[X]\}.$$

Considérons d le plus petit degré possible d'un élément non nul de cet ensemble, avec la convention que $\deg(0) = -\infty$.
Si $d = 0$ alors on peut trouver k polynômes $R_1, ..., R_k$ tels que

$$R_1 P_1 + ... + R_k P_k = 1.$$

Case 2D5000, 292 rue Saint-Martin 75141 Paris Cedex 03
Bureau 17.0.14
Tel: 0158808765, Mel: thierry.horsin@cnam.fr

Conservatoire National des Arts et Métiers
Département d'Ingénierie Mathématique

Notons que ceci implique que les seuls polynômes qui divisent chacun des P_i sont les polynômes constants.
Sinon, il existe un polynôme m de degré d tel que
$$R_1P_1 + ... + R_kP_k = m$$
pour un choix de $R_1, ..., R_k$.
Ainsi tout polynôme qui divise chacun des P_i divise m. Le polynôme m est donc un multiple du $PGCD$ de $P_1, ..., P_l$.
Prenons maintenant un polynôme P de la forme $R'_1P_1 + ... + R'_2P_2$.
On effectue la division de P par m. $P = mQ + R$ avec $\deg(R) < \deg(m)$.
On a donc $R = P - mQ = R'_1P_1 + ... + R'_kP_k - QR_1P_1 - ... - QR_kP_k$ donc $R \in I$. La seule possibilité puisque m est de degré minimal est que $R = 0$, donc m divise P. En particulier m divise P_1, P_2,...,P_k, c'est donc un diviseur commun de ces polynômes. Donc le polynôme m est un diviseur de chacun de ces polynômes et tout diviseur commun à chacun de ces polynômes divise m, qui est donc le plus grand diviseur commun des fonctions $P_1, ..., P_k$.
Supposons que les polynômes $P_1, ..., P_k$ soient premiers entre eux deux à deux.
Soit E un espace vectoriel de dimension finie, alors on a

Théorème 8.10 [des noyaux].
$$\ker(P_1(f) \circ .... \circ P_k(f)) = \oplus_{i=1}^k \ker(P_i(f)).$$

On démontre ce theorème par récurrence sur le nombre de polynômes.
Supposons $k = 2$. D'après les résultats exposés sur les polynômes.
Il existe deux polynômes R_1 et R_2 tels que
$$1 = P_1R_1 + P_2R_2.$$
On a donc $P_1(f)P_2(f)(x) = P_2(f)P_1(f)R_1(f)(x) + P_2(f)R_2(f)(x)$.
Donc si $x \in \ker(P_1(f)P_2(f))$ alors $R_1(f)P_1(f)(x) \in \ker(R_2(f))$ et $R_2(f)P_2(f) \in \ker(R_1(f))$.
Finalement
$$\ker(P_1(f)P_2(f)) \subset \ker(P_1(f)) + \ker(P_2(f)).$$
De plus on a
$$\ker(P_1(f)) \subset \ker(P_1(f)P_2(f))$$
et
$$\ker(P_2(f)) \subset \ker(P_1(f) \circ P_2(f)).$$
Donc
$$\ker(P_1(f)P_2(f)) = \ker(P_1(f)) + \ker(P_2(f)).$$
Soit $x \in \ker(P_1(f)) \cap \ker(P_2(f))$.
Comme $x = R_1(f)P_1(f)(x) + R_2(f)P_2(f)(x)$ alors $x = 0$.
Finalement

Case 2D5000, 292 rue Saint-Martin 75141 Paris Cedex 03
Bureau 17.0.14
Tel: 0158808765, Mel: thierry.horsin@cnam.fr
Conservatoire National des Arts et Métiers
Département d'Ingénierie Mathématique

$$\ker(P_1(f)P_2(f)) = \ker(P_1(f)) \oplus \ker(P_2(f)).$$

La récurrence est alors immédiate.
Ainsi, grâce au théorème de Cayley-Hamilton, si P désigne le polynôme caractéristique de f, on a dans le cas où f est scindé on a

$$E = \oplus_{i=1}^{k} \ker(f - \lambda_i I)^{\nu_i}$$

Donc $N = \sum_{i=1}^{k} \dim(\ker(f - \lambda_i I)^{\nu_i}$.
Comme $\ker(f - \lambda_i I) \subset \ker(f - \lambda_i I)^{\nu_i}$ on en déduit que
f est diagonalisable ssi $\dim(\ker(f - \lambda I)) = \dim(\ker(f - \lambda_i I)^{\nu_i})$.
Finalement on a

Théorème 8.11 *Un endomorphisme scindé est diagonalisable ssi les sous-espaces propres et les sous-espaces caractéristiques ont même dimension.*

Remarquons que $\dim \ker(f - \lambda_i I)^{\nu_i} = \nu_i$.
En effet il existe un polynôme de degré le plus petit possible m tel que $m(f) = 0$.
En effet, $\{P/P(f) = 0\}$ est non vide. Soit m un élément de cet ensemble de degré minimum. Faisons la division de P par m, il existe Q et R tels que $P = mQ + R$ avec $\deg(R) < \deg(m)$.
Comme $P(f) = m(f) = 0$ alors $R(f) = 0$, donc R est nul car m est de degré minimal.
Ainsi il existe un polynôme non nul de degré le plus petit possible, tel que
$m(f) = 0$ et tout polynôme qui annule f est divisible par m, donc également le polynôme caractéristique.
Si on prend le coefficient de plus haut degré de f égal à 1 alors

Définition 8.12 *m s'appelle le polynôme minimal de f.*

Lorsque le polynôme est scindé, le polynôme minimal est donc de la forme

$$m = (X - \lambda_1)^{\alpha_1} \ldots\ldots (X - \lambda_k)^{\alpha_k},$$

où, a priori, $0 \leq \alpha_i \leq \nu_i$.
Remarquons que nécessairement $\alpha_i \geq 1$. Puisque λ_i est valeur propre donc $1 \leq \alpha_i \leq \nu_i$.
Mais d'après le théorème des noyaux, on a

$$N = \sum_{i=1}^{k} \dim(\ker(f - \lambda_i I)^{\alpha_i}) = \sum_{i=1}^{k} \dim(\ker(f - \lambda_i I)^{\nu_i}),$$

Remarquons que m étant le polynôme minimal, on a

$$\dim(\ker(f - \lambda_i I)^{\alpha_i}) \leq \nu_i.$$

En effet, si pour un i, on a $\dim(\ker(f - \lambda_i I)^{\alpha_i} > \nu_i$, pour tout $x \in \ker(f - \lambda_i I)^{\alpha_i}$non nul, les α_i vecteurs

$$x, (f - \lambda_i I)(x), \ldots, (f - \lambda_i I)^{\alpha_i - 1}(x)$$

Case 2D5000, 292 rue Saint-Martin 75141 Paris Cedex 03
Bureau 17.0.14
Tel: 0158808765, Mel: thierry.horsin@cnam.fr

Conservatoire National des Arts et Métiers
Département d'Ingénierie Mathématique

ne peuvent pas être liés (sinon le polynôme m serait de degré plus petit), donc il existe x tel que ces $\alpha_i > \nu_i$ vecteurs soient libres et $x \in \ker(f - \lambda_i I)^{\alpha_i}$. Pour cet x là, si P_f est le polynôme caractéristique de f alors $P_f(f)(x) \neq 0$, ce qui est impossible.
Finalement $N = \sum_{i=1}^{k} \dim \alpha_i \leq \sum_{i=1}^{K} \nu_i$. Donc $\forall i = 1, ..., k$ on a $N = \sum_{i=1}^{k} \dim(\ker(f - \lambda_i)_i^{\alpha} \leq \sum_{i=1}^{k} \nu_i = N$.
Il s'ensuit que $\forall i = 1, ..., k$ on a $\dim(\ker(f - \lambda_i)_i^{\alpha}) = \nu_i$.
Finalement pour un endomorphisme scindé on a

Théorème 8.13 *Si f est scindé, f est diagonalisable ssi pour toute valeur propres λ de f, la dimension des sous-espaces propre est égal à la multiplicité de la valeur propre.*

8.3 Trigonalisation

Soit $f : E \to E$ un endomorphisme. Soit P_f son polynôme caractéristique. On dira que f est trigonalisable (ou triangulisable) si il existe une famille de sous-espaces vectoriels de E, E_1,....,E_N de dimension respective $1...N$ et tels que $f(E_i) \subset E_i$ avec $E_i \subset E_{i+1}$ pour tout $i = 1, ..., N - 1$.
La matrice d'une telle application linéaire exprimée dans une base $e_1, ..., e_N$ telle que $e_1 \in E_1$ (e_1, e_2) base de E_2,..., $(e_1, ..., e_N)$ base de $E_N = E$ est triangulaire supérieure.
On voit que les éléments diagonaux sont les valeurs propres de f.
Il suffit en effet de calculer $\det(A - \lambda I)$.
On dispose alors du théorème

Théorème 8.14 *f est triangulisable ssi son polynôme caractéristique est scindé.*

En effet si $P_f = \pi_{i=1}^{k}(X - \lambda_i)^{\nu_i}$. On prend une base de $\ker(f - \lambda_1 I)$, $e_1^1, .., e_{k_1}^1$, que l'on complète en une base de $\ker(f - \lambda_1 I)^2$ par $e_{k_1+1}^1 ... e_{k_2}^1$.....,jusqu'à avoir constitué une base de $\ker(f - \lambda_1 I)^{\nu_1}$ et on continue.
Une fois ce théorème établi, pour que f soit triangulisable, il faut et il suffit que P_f ait toutes ses racines dans $\mathbb{K}$.
En particulier toute matrice à coefficients dans $\mathbb{R}$ ou $\mathbb{C}$ est trigonalisable sur $\mathbb{C}$.

9 Exemples d'endomorphismes diagonalisables

On suppose ici que $\mathbb{K} = \mathbb{R}$.

Théorème 9.1 *Soit f un endomorphisme de E dont le matrice dans une base donnée est symétrique, alors f est diagonalisable.*

Case 2D5000, 292 rue Saint-Martin 75141 Paris Cedex 03
Bureau 17.0.14
Tel: 0158808765, Mel: thierry.horsin@cnam.fr
Conservatoire National des Arts et Métiers
Département d'Ingénierie Mathématique

Remarque 9.2 *Ainsi la matrice d'inertie d'un solide indéformable est diagonalisable dans une BON. On rappelle que si S est un solide indéformable, la matrice d'inertie est en notant $\rho(M)$ la densité de masse en un point M et $dm = \rho(M)dM$*

$$\begin{pmatrix} A & -F & -E \\ -F & B & -D \\ -E & -D & C \end{pmatrix},$$

où

$$A = \int_S (y^2 + z^2)dm, \; B = \int_S (x^2 + z^2)dm, \; \int_S (y^2 + z^2)dm,$$

et

$$F = \int_S xydm, \; E = \int_S xzdm, \; D = \int_S yzdm.$$

Cette matrice intervient quand on étudie le moment cinétique propre d'un solide

$$\vec{L} = \int_S \overrightarrow{GM} \wedge (\vec{\omega} \wedge \overrightarrow{GM})dm,$$

$\vec{\Omega}$ désignant le vecteur rotation de S. Les éléments de la matrice d'inertie apparaissent en écrivant les coordonnées de $\vec{L}$ dans la base usuelle de $\mathbb{R}^3$ (attachée à G). Le moment cinétique propre intervient dans le calcul du moment dynamique et donc dans le PFD.

Remarque 9.3 *La matrice de covariance d'une famille de v.a. $X_1, ..., X_n$ est symétrique:*

$$cov(X_i, X_j) := E(X_i X_j) - E(X_i)E(X_j).$$

Pour deux v.a. indépendantes on a $cov(X_i, X_j) = 0$. Il s'ensuit que cette matrice est un indicateur d'indépendance (cependant l'indépendance n'est pas caractérisée par la covariance).

Preuve du théorème:
Soit λ et x une valeur propre (sur $\mathbb{C}$) et son vecteur propre respectif.
Définissons un produit scalaire $\langle | \rangle$ par la formule usuelle dans la base de E choisie. Cette base est alors orthonormale pour ce produit scalaire.
$\langle A(x)|\bar{x}\rangle = \lambda\langle x|\bar{x}\rangle$ donc λ est réel. Et finalement $\Re(x)$ est vecteur propre associé à λ.
Donc toute matrice symétrique a une valeur propre réelle λ et un vecteur réel x qui lui est associé.
Soit $x^\perp = \{y \in E, \langle x|y\rangle = 0\}$. Si $x' \in x^\perp$ alors
$\langle x|Ax'\rangle = \lambda\langle x|x'\rangle$ donc $x^\perp$ est stable par A. La restriction de A à $x^\perp$ est également symétrique, on peut donc conclure par récurrence.
Remarquons que la procédure montre que les vecteurs propres seront orthogonaux entre eux.
Soit P la matrice de passage de la base initiale à la base de vecteurs propres.
Les colonnes de P sont les coordonnées des vecteurs propres dans la base initiale. Elles sont donc orthogonales entre elles, et on peut choisir les vecteurs propres v de fapropcon à avoir ${}^t vv = 1$.
On a donc ${}^t PP = Id$, donc $P^{-1} = P$. Une telle matrice est dite orthogonale
Ainsi

Proposition 9.4 *On peut diagonaliser une matrice symétrique dans une base orthonormale de vecteurs propres via une matrice de passage orthogonale.*

Théorème 9.5 *Si A commute avec ${}^tA$ alors A est diagonalisable (éventuellement sur $\mathbb{C}$).*

Définition 9.6 *Une telle matrice est dite normale.*

Il faut prendre garde à la diagonalisibilité dans ce cas. En effet elle est à entendre sur $\mathbb{C}$.
En effet la matrice $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ est normale, mais n'est pas diagonalisable sur $\mathbb{R}$, ses valeurs propres étant i et $-i$.
Soit λ une valeur propre (complexe) de A.
On a

$$A(x) = \lambda x$$

donc

$${}^tAA(x) = \lambda\, {}^tA(x).$$

Ainsi ${}^tA(x)$ est donc un vecteur propre associé à λ.
Soit y tel que

$$\langle x|\bar{y}\rangle = 0,\ \forall x \text{ vecteur propre associé à } \lambda,$$

alors

$$\langle A(y)|\bar{x}\rangle = \langle y|{}^tA(\bar{x})\rangle$$

Or ${}^tA(x)$ est aussi un vecteur propre associé à λ, donc $A(y)$ est orthogonal à tout vecteur propre associé à λ.
Maintenant

$$\langle {}^tA(y)|\bar{x}\rangle = \langle y|A(\bar{x})\rangle = \langle y|\bar{\lambda}\bar{x}\rangle = 0.$$

En raisonnant sur

$$F := \{y \in E,\ \langle y|\bar{x}\rangle,\ x \text{ vecteur propre associé à } \lambda\}$$

stable par A et ${}^tA$ qui est de dimension $< \dim(E)$, on voit qu'on peut maintenant faire un raisonnement par récurrence puisque la restriction de A à F est stable par A et ${}^tA$, et sa restriction à F est alors normale sur F.
En fait les raisonnements précédents, s'étendent aux matrices complexes qui commutent avec leur transconjuguée (et en particulier qui lui sont égales on parle alors de matrice hermitienne).

Case 2D5000, 292 rue Saint-Martin 75141 Paris Cedex 03
Bureau 17.0.14
Tel: 0158808765, Mel: thierry.horsin@cnam.fr

Conservatoire National des Arts et Métiers
Département d'Ingénierie Mathématique

10 Formes normales de Jordan

Soit f un endomorphisme trigonalisable.
Il existe une base de E telle que la matrice de f dans cette base soit triangulaire supérieure.
Notons D la matrice diagonale construite et N telle que $A = D + N$.
Si $e_1, ..., e_N$ est la base de trigonalisation (supérieure) alors

$$N(e_i) \subset \text{vect}(e_1, .., e_{i-1}) \text{ et } N(e_1) = 0.$$

Ainsi il existe un plus petit k tel que $N^k = 0$.
Remarquons qu'on peut choisir cette base de trigonalisation à l'aide d'une base de chacun des $\ker(f - \lambda_i I)^{\alpha_i}$.
En particulier si l'on fait cela, D et N commutent.
On a ainsi prouvé:

Théorème 10.1 *Tout endomorphisme f s'écrit $f = d+n$ où d et n commutent d est diagonale et n est nilpotente, d et n commutent, pourvu que $\mathbb{K}$ soit algébriquement clos (par exemple $\mathbb{C}$).*

Soit n un endomorphisme nilpotent.
Soit k l'indice de nilpotence, c'est à dire le plus petit entier tel que $n^k = 0$.
Choisissons une base de $\ker(n^{k-1})$ formée de la manière suivante:
$\ker(n^{k-1}) = F_{k-1} \oplus \ker(n^{k-2})$.
On choisit une base de F_{k-1} $e_1^{k-1}, ...e_{f_{k-1}}^{k-1}$, l'image de cette base par n est une famille libre de $\ker(n^{k-2})$
On a donc $\ker(n^{k-2}) = n(F_{k-1}) \oplus G_{k-2} \oplus \ker(n^{k-3})$.
On note $H_{k-2} = n(F_{k-1}) \oplus G_{k-2}$ dont on construit une base en complétant la famille $n(e_1^{k-1}), ..., n(e_{f_{k-1}}^{k-1})$ par une base de H_{k-1} $e_1^{k-2}, ..., e_{f_{k-2}}^{k-2}$, on poursuit cette opération.
$\ker(n^{k-3}) = n^2(F_{k-1}) \oplus n(G_{k-2}) \oplus G_{k-3} \oplus \ker(n^{k-4})$.
Finalement on construit une base de E en partant de bas en haut de gauche à droite où la matrice de n est de la forme

$$\begin{pmatrix}
0 & 1 & 0 & \dots & \dots & \dots & \dots & \dots & \dots & \dots & 0 \\
0 & 0 & 1 & 0 & \dots & \dots & \dots & \dots & \dots & \dots & 0 \\
0 & 0 & 0 & 1 & 0 & \dots & \dots & \dots & \dots & \dots & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & \dots & \dots & 0 \\
0 & \dots & \dots & \dots & 0 & 1 & 0 & \dots & \dots & \dots & 0 \\
\multicolumn{11}{c}{\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots} \\
0 & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & 0 & 1 \\
0 & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & 0
\end{pmatrix}$$

Case 2D5000, 292 rue Saint-Martin 75141 Paris Cedex 03
Bureau 17.0.14

Tel: 0158808765, Mel: thierry.horsin@cnam.fr

Conservatoire National des Arts et Métiers

Département d'Ingénierie Mathématique

11 Matrices compagnons: invariants de similitude (théorème de Frobenius)

Soit $f : E \to E$ un endomorphisme et soit m son polynôme minimal de degré d.
Soit $x \in E$ $\{P / P(f)(x)\}$ est non vide puisqu'il contient m. Supposons $x \neq 0$.
Soit P_x le polynôme de degré le plus petit possible et non nul dans cet ensemble de polynôme.
En faisant la division de m par P_x on voit que P_x divise m.
Montrons qu'il existe x tel que $P_x = m$. Notons $m_1..m_k$ les facteurs premiers de m.
D'après le théorème des noyaux, on a $E = \ker(m_1(f)) \oplus ... \oplus \ker(m_k(f))$.
Tous les éléments x de $\ker(m_1(f))$ ont un P_x qui divise m_1. Il en existe donc un tel que $P_x = m_1$.
Soit $x_1 ... x_k$ tels que $m_i = P_{x_i}$. Posons $x = x_1 + ... + x_k$. Il est clair que $P_x = m$.
Soit $E_x = \mathrm{vect}(x, f(x), ...,)$. E_x est de dimension d.
Montrons qu'il existe E_x possède un sev supplémentaire stable par f.
Soit $l \in E^*$ tel que $l(x), ..., l(f^{d-2}(x))$ sont nuls et tel que $l(f^{d-1}(x)) = 1$.
Soit $F = \{l, ..., l \circ f^{d-1}\}^{\perp}$. Montrons F est stable par f.
Si $y \in F$, $l(f(y)) = 0$. De plus si ces d éléments vérifient on a $a_0 l + ... + a_{d-1} l \circ f^{d-1} = 0$. On a alors $a_{d-1} = 0$, etc, puis $a_0 = 0$.
F est donc de dimension $N - d$. De plus comme F est stable par f, ainsi que E_x, alors $F \cap E_x = \{0\}$.
Finalement $E = E_x \oplus F$. Poursuivant cette méthode, on voit que

Théorème 11.1 *$E = E_{x_1} \oplus E_{x_2} \oplus ... \oplus E_{x_k}$, où $P_{x_1} = m$ P_{x_2} / P_{x_1} et $P_{x_2} = m_{f/E_{x_2} \oplus ... \oplus E_{x_k}}$,...$P_{x_k} / P_{x_{k-}}$ et $P_{x_k} = m_{f/E_{x_k}}$.*

Le choix des x_i n'est pas unique, mais le nombre et les polynômes le sont.
En effet, supposons qu'on ait deux telles décompositions, $E = E_{y_1} \oplus ... \oplus E_{y_q}$
$P_{x_1} = P_{y_1}$ par construction.
$P_{x_2}(f)(E) = P_{x_2}(f)(E_{x_1}) = P_{x_2}(f)(E_{y_1} \oplus ... \oplus E_{y_q})$.
$(P_{x_1} / P_{x_2}) = P_{P_{x_2}(f)(x_1)} = P_{P_{x_2}(f)(y_1)}$.
Finalement $E_{P_{x_2}(f)(x_1)}$ et $E_{P_{x_2}(f)(y_1)}$ ayant même annulateur minimal ont même dimension. Finalement P_{y_2} / P_{x_2}. En échangeant les rôles, on a $P_{x_2} = P_{y_2}$.....

Théorème 11.2 *Les polynômes (en prenant le terme de plus haut degré de facteur 1) intervenant dans le théorème précédent et leur nombre sont uniques. On les appelle invariants de similitude de f*

Remarquons que dans la base de E_{x_1} formée par $x_1, ..., f^{d-1}(x_1)$ la matrice de f a une forme très particulière déjà rencontrée lors de la deuxième preuve du théorème de Cayley-Hamilton. Une telle matrice est appelée matrice compagnon du polynôme qui sert à l'écrire.

Case 2D5000, 292 rue Saint-Martin 75141 Paris Cedex 03
Bureau 17.0.14
Tel: 0158808765, Mel: thierry.horsin@cnam.fr

Conservatoire National des Arts et Métiers
Département d'Ingénierie Mathématique

12 Forme normale de Jordan[12]

On a vu que toute matrice est trigonalisable (éventuellelement sur $\mathbb{C}$)
On peut être plus précis: Soit $f : \mathbb{R}^n \to \mathbb{R}^n$ une application linéaire.

Théorème 12.1 *Il existe une base dans laquelle la matrice de f est de la forme $D+N$ où D est une matrice diagonale et N une matrice nilpotente telles que $DN = ND$ et telle que N possède des 1 ou des 0 au dessus de la diagonale, tous les autres coefficients de N étant nuls.*

Evidemment si la trigonalisation est possible sur $\mathbb{R}$ cette opération se fait sur $\mathbb{R}$, sinon elle est toujours possible sur $\mathbb{C}$.

Exemple: $f : \mathbb{R}^3 \to \mathbb{R}^3$ dont la matrice dans la base usuelle est donnée par $\begin{pmatrix} 1 & 2 & -2 \\ 2 & 1 & -2 \\ 1 & 1 & -1 \end{pmatrix}$.

Soit $\vec{u_1} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ vecteur propre associé à 1, $\vec{u_2} = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$ vecteur propre associé à -1 et $\vec{u_3} = \vec{u_1} \wedge \vec{u_2} = \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}$.

On a vu que $f(\vec{u_3}) = \vec{u_3} + 6\,\vec{u_1}$.
Dans la base $\vec{u_1}, \vec{u_3}/6, \vec{u_2}$ la matrice de l'endomorphisme est

$$\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

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

Application de la réduction des endomorphismes:
De multiples modèles où le temps intervient comme paramètre d'évolution s'écrivent sous la forme $X'(t) = MX(t)$ où X est un vecteur colonne dépendant du temps et M une matrice carrée de taille $n \times n$.
On démontre que la solution d'un tel système est $X(t) = \exp(tM)X(0)$ où

$$\exp(tM) = \sum_{k=0}^{\infty} \frac{t^k M^k}{k!}.$$

Si on réduit M (diagonalisation ou trigonalisation), c'est à dire $A = P^{-1}MP$ avec A diagonale ou $A = D + N$.
On a $M = PAP^{-1}$ et on voit que $\exp(tM) = P\exp(tA)P^{-1}$. On est donc réduit à calculer $\exp(tA)$.

[12] Camille, mathématicien franpropcais, 1838-1922

Case 2D5000, 292 rue Saint-Martin 75141 Paris Cedex 03
Bureau 17.0.14
Tel: 0158808765, Mel: thierry.horsin@cnam.fr
Conservatoire National des Arts et Métiers
Département d'Ingénierie Mathématique

Si A est diagonale $\exp(A)$ est une matrice diagonale dont la diagonale est constituée des exp des éléments diagonaux de A.
Si $A = D + N$ où D et N commutent alors

$$\exp(t(D+N)) = \exp(tD)\exp(tN).$$

Par exemple si $N = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$,

$$N^2 = [\,0\,].$$

Donc

$$\exp(N) = I + N,$$

finalement

$$\exp(D+N) = \exp(D)(I+N).$$

13 Espaces euclidiens

13.1 Définition

Définition 13.1 *On appelle produit scalaire sur $\mathbb{R}^n$ une application*

$$p : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$$

telle que pour tout vecteur et scalaire apparaissant, on ait

i. $p(\vec{u}, \vec{v}) = p(\vec{v}, \vec{u})$

ii. $p(\vec{u}, \vec{u}) \geq 0$ *et* $p(\vec{u}, \vec{u}) = 0 \iff u = 0.$

iii. $p(\alpha\,\vec{u} + \beta\,\vec{v}, \vec{w}) = \alpha p(\vec{u}, \vec{w}) + \beta p(\vec{v}, \vec{w}).$

On dit que p est une forme bilinéaire symétrique définie positive.

Définition 13.2 *Un espace euclidien est un espace vectoriel de dim finie muni d'un produit scalaire.*

Exemple: $p(\vec{u}, \vec{v}) := \sum_{i=1}^n u_i v_i$ où les (u_i) et les (v_i) sont les composantes des vecteurs $\vec{u}$ et $\vec{v}$ dans la base usuelle de $\mathbb{R}^n$.
On dispose d'une inégalité fondamentale

Case 2D5000, 292 rue Saint-Martin 75141 Paris Cedex 03
Bureau 17.0.14
Tel: 0158808765, Mel: thierry.horsin@cnam.fr

Conservatoire National des Arts et Métiers
Département d'Ingénierie Mathématique

Théorème 13.3 [Inégalité de Cauchy-Schwartz]. *On a*
$$|p(\vec{u},\vec{v})| \leq \sqrt{p(\vec{u},\vec{u})}\sqrt{p(\vec{v},\vec{v})},$$

qui se démontre en utilisant le fait que $p(\vec{u}+\alpha\vec{v},\vec{u}+\alpha\vec{v}) \geq 0$ pour tout α.
En particulier l'égalité dans l'inégalité de Cauchy-Schwarz est vrai ssi $\vec{u},\vec{v}$ est une famille liée.

Définition 13.4 *Si p est une produit scalaire sur $\mathbb{R}^n$ alors on définit la norme associée à p par*
$$\| \vec{u} \|_p = \sqrt{p(\vec{u},\vec{u})}.$$

On dispose de

Proposition 13.5 *Si $\|\cdot\|_p$ est la norme associée à p alors*

i. $\| \vec{u} + \vec{v} \|_p \leq \| \vec{u} \|_p + \| \vec{v} \|_p$.

ii. $\|\lambda \vec{u} \|_p = |\lambda| \| \vec{u} \|_p$.

13.2 Définir un espace euclidien

Proposition 13.6 *Etant donné une base de $\mathbb{R}^n$ constituée de vecteurs $\vec{u_1},\dots,\vec{u_n}$, il existe un produit scalaire p pour lesquels ces n vecteurs sont 2 à 2 orthogonaux entre eux et chacun de norme 1.*

Il nous faut définir $p(\vec{v},\vec{w})$ pour deux vecteurs quelconques.
On écrit $\vec{v} = \sum_1^n v_i \vec{u_i}$ et $\vec{w} = \sum_1^n w_i \vec{w_i}$. On doit alors avoir
$$p(\vec{v},\vec{w}) = \sum_i v_i w_i.$$
On vérifie que cette définition donne bien un produit scalaire.

13.2.1 Exemples

Produits scalaires usuels:
$$p(\vec{v},\vec{w}) = \sum_1^n i v_i w_i,$$
$$p(\vec{v},\vec{w}) = \sum_1^n (1+i^2)^{1/2} v_i w_i.$$

14 Procédé d'orthonormalisation

Etant donné n vecteurs de $\mathbb{R}^n$ $\vec{u_1}, \ldots, \vec{u_n}$ et p un produit scalaire. On dispose du résultat suivant:

Théorème 14.1 *Ces n vecteurs forment une base de $\mathbb{R}^n$ ssi la matrice suivante (dite matrice de Gramm-Schmidt) est inversible.*

$$(p(\vec{u_i}, \vec{u_j}))_{i=1,\ldots,n, j=1,\ldots,n}.$$

En effet si il existe une combinaison linéaire de ces vecteurs qui est nulle, alors la même combinaison linéaire des lignes (ou des colonnes) de cette matrice est nulle.
Réciproquement: si la matrice est non inversible, il existe une combinaison linéaire des lignes (ou des colonnes) dont les coefficients sont non tous nuls qui donne une ligne nulle. Les vecteurs concernés par cette combinaison linéaire ont la même combinaison linéaire nulle. Sinon....on peut toujours supposer (quitte à diminuer n) que ces combinaisons linéaires ne concerent que p vecteurs linéairement indépendants. Mais en faisant la même combinaison des lignes on obtient la norme de cette combinaison nulle, donc cette combinaison est nulle contredisant que ces vecteurs sont linéairement indépendants.
Etant donné p vecteurs linéairement indépendants dans un espace de dimension n euclidien. Il existe p vecteurs engendrant le même sous-espace vectoriel et formant une famille orthonormale. En effet soit $\vec{v_1}, \ldots, \vec{v_p}$ cette famille. On pose

$$\vec{w_1} = \frac{\vec{v_1}}{\| \vec{v_1} \|}; \overrightarrow{W_2} = \alpha_2 \vec{w_1} + \vec{v_2}.$$

On résoud

$$\alpha + \beta\langle \vec{v_2}, \vec{w_1} \rangle = 0$$

ce qui donne α. On définit ensuite

$$\vec{w_2} = \frac{\overrightarrow{W_2}}{\| \overrightarrow{W_2} \|}.$$

Puis

$$\overrightarrow{W_3} = \alpha_3 \vec{w_1} + \beta_3 \vec{w_2} + \vec{v_3}.$$

Ce qui donne

$$\alpha_3 + \langle \vec{v_3}, \vec{w_1} \rangle = 0, \text{ et } \beta_3 + \langle \vec{v_3}, \vec{w_1} \rangle = 0,$$

Case 2D5000, 292 rue Saint-Martin 75141 Paris Cedex 03
Bureau 17.0.14
Tel: 0158808765, Mel: thierry.horsin@cnam.fr

Conservatoire National des Arts et Métiers
Département d'Ingénierie Mathématique

puis on pose

$$\vec{w_3} = \frac{\overrightarrow{W_3}}{\| \overrightarrow{W_3} \|}.$$

On poursuit ce procédé. dit procédé d'orthonormalisation de Gramm-Schmidt.

14.1 Diagonalisation

Soit $(\vec{u_i})$ une base de E et $(\vec{e_i})$ une autre base de E. Soit P la matrice de passage (contenant les coordonnées des $\vec{e_i}$ dans la base $\vec{u_j}$. On notera les coefficients de P avec la lettre π. Soit $\vec{x}$ et $\vec{y}$ deux vecteurs et x_i et X_i les coordonnées de $\vec{x}$ dans la 1ère et 2ème bases respectivement. Notons $u_{ij} = \left\langle \vec{u_i} \mid \vec{u_j} \right\rangle$ et $e_{pq} = \left\langle \vec{e_p} \mid \vec{e_q} \right\rangle$.
On a

$$\left\langle \vec{x} \mid \vec{y} \right\rangle = \sum_{i=1}^{N} \sum_{j=1}^{N} x_i y_j u_{ij} = \sum_{p=1}^{N} \sum_{q=1}^{N} X_p Y_q e_{pq} = {}^t x U y = {}^t X E Y$$

Or $x = PX$ et $y = PY$.
donc

$$E = {}^t P U P$$

En particulier si les deux bases sont orthonormales alors ${}^t PP = I$.

Définition 14.2 *Si ${}^t PP = P\,{}^t P = Id$, on dit que P est orthogonale.*

On considère $f : E \to E$ un endomorphisme de $(E, \langle \cdot | \cdot \rangle)$ espace euclidien.
Soit A la matrice de f dans une base quelconque.
On dira que f est symétrique si

$$\forall x, y, \langle f(x) \,|\, y \rangle = \langle x \,|\, f(y) \rangle .$$

Si M est la matrice de f dans une base $\langle \cdot | \cdot \rangle$ orthonormale de E, il revient au même de dire que M est symétrique. Car $\left\langle M(e_i) \,|\, e_j \right\rangle = \left\langle e_i \,|\, {}^t M(\vec{e_j}) \right\rangle = \left\langle e_i \,|\, M e_j \right\rangle = m_{ij} = m_{ji}$ si (e_i) désigne cette base orthonormale.
Il s'ensuit que

Théorème 14.3 *Si f est un endomorphisme symétrique de $(E, \langle \cdot | \cdot \rangle)$, f est diagonalisable dans une base orthonormale de vecteurs propres*

Remarque 14.4 *Dans le cas d'une matrice symétrique dans la base canonique de $\mathbb{R}^N$, la matrice de passage à une base orthonormale de vecteurs propres est alors orthogonale.*

Case 2D5000, 292 rue Saint-Martin 75141 Paris Cedex 03
Bureau 17.0.14
Tel: 0158808765, Mel: thierry.horsin@cnam.fr
Conservatoire National des Arts et Métiers
Département d'Ingénierie Mathématique

Exemple: Soit A une matrice symétrique et B une matrice symétrique à valeurs propres strictement > 0.
B définit un produit scalaire sur $\mathbb{R}^N$: $\langle \cdot | \cdot \rangle_B := \langle \cdot | B \cdot \rangle$.
Soit v l'endomorphisme symétrique défini par B et u celui définit par A.
On a

$$\left\langle v^{-1}u(x) \,|\, y \right\rangle_B = \left\langle v^{-1}u(x) \,|\, v(y) \right\rangle = \left\langle x \,|\, vv^{-1}u(y) \right\rangle = \left\langle x \,|\, v^{-1}u(y) \right\rangle.$$

Il s'ensuit que $v^{-1}u$ est un endomorphisme symétrique de $(E, \langle \cdot | \cdot \rangle_B)$. Il existe donc une base $\langle \cdot | \cdot \rangle_B$ orthonormale de vecteurs propres de $v^{-1}u$. On peut donc trouver une matrice P telle que

$$\forall x, y, \ \langle P(x) \,|\, P(y) \rangle_B = \langle x \,|\, y \rangle_B$$

et telle que

$$P^{-1}B^{-1}AP = D$$

où D est diagonale. Donc

$${}^tPBP = I, \ B^{-1} = P\ ({}^tP) \text{ donc } \ {}^tPAP = D.$$

Soit X un vecteur colonne et $Y = ({}^tP)^{-1}X$.
On a

$${}^tXAX = \ {}^tY\,{}^tPAPY = \ {}^tYDY$$

et

$${}^tXBX = \ {}^tYY.$$

Exemple en mécanique: En présence d'un potentiel quadratique A constant et d'un tenseur d'énergie cinétique B constant , le lagrangien[13]associé est de la forme

$$L(q, \dot{q}) = \frac{1}{2}\left({}^t\dot{q}B\dot{q} + \ {}^tqAq \right),$$

où $(q, \dot{q}) \in \mathbb{R}^N \times \mathbb{R}^N$ sont voués à désigner le vecteur position et le vecteur vitesse.
Si l'on effectue le changement donnée par la réduction précédente.
On voit que L prend la forme

$$\tilde{L}(Q, \dot{Q}) = \frac{1}{2}\sum_{i=1}^{N}(\dot{Q}_k^2 + \lambda_k Q_k^2)$$

---

[13] Giuseppe Lodovico Lagrange, mathématicien italien, 1736–1813

Case 2D5000, 292 rue Saint-Martin 75141 Paris Cedex 03
Bureau 17.0.14
Tel: 0158808765, Mel: thierry.horsin@cnam.fr

Conservatoire National des Arts et Métiers
Département d'Ingénierie Mathématique

Les équations de Lagrange dans ce système de coordonnées deviennent

$$\frac{d}{dt}\frac{\partial \tilde{L}}{\partial \dot{Q}_k} - \frac{\partial \tilde{L}}{\partial Q_k} = 0$$

ce qui donne

$$\ddot{Q}_k - \lambda_k Q_k = 0,\ i = 1...N.$$

Noter qu'en général $\lambda_k < 0$. Par exemple dans le cas d'un potentiel élastique c'est le cas:
Par exemple si un point de masse M glisse sans frottement sur une tige horizontale et est attaché à un ressort de raideur K et longueur au repos l_0, l'énergie cinétique est $\frac{1}{2}Mv^2$ ($v = \dot{q}$) et l'énergie potentielle est $K(l - l_0)^2/2$ ($l = q$)), le lagrangien est

$$\frac{1}{2}Mv^2 - K\frac{(l - l_0)^2}{2}$$

Dans le cas général, on a donc "dévissé" le système en la superposition de N oscillateurs harmoniques (si les $\lambda_k < 0$).
Remarque: le formalisme lagrangien est une manière de traiter les problèmes mécaniques.
Dans le cas de la masselotte accrochée au ressort, alors on sait que

$$M\ddot{l} = -K(l - l_0) \tag{1}$$

Donc en multipliant par $\dot{l}$ on en déduit que

$$M\frac{\dot{l}^2}{2} + K\frac{(l - l_0)^2}{2}$$

est constante.
Le lagrangien est une notion associée au minimum de l'action. Soit L une fonction de $\mathbb{R}^N \times \mathbb{R}^N \to \mathbb{R}$ dit le lagrangien et supposée au moins de classe C^2 (on peut améliorer cette hypothèse).
On considère pour cela

$$A(q) := \int_0^1 L(q, \dot{q})ds$$

où q est une fonction de classe C^1 sur $[0, 1]$ telle que $q(0) = q_0$ et $\dot{q}(0) = \dot{q}_0$ et de même en $t = 1$ sont fixés.
On peut calculer la dérivée de A par rapport à q en considérant des perturbations $q + t\delta q$ avec $\delta q(0) = \delta q(1) = 0$ (d'autres choix sont possibles)
On a

Case 2D5000, 292 rue Saint-Martin 75141 Paris Cedex 03
Bureau 17.0.14
Tel: 0158808765, Mel: thierry.horsin@cnam.fr

Conservatoire National des Arts et Métiers
Département d'Ingénierie Mathématique

$$(A(q+t\delta q))'(t)_{t=0} = (\int_0^1 L(q+t\delta q, \dot{q}+t\dot{\delta q})ds)'(t)_{t=0}$$
$$= \int_0^1 (\partial_q L \delta_q + \dot{\delta q}\partial_{\dot{q}} L)ds$$

et en intégrant par partie le deuxième terme, puis en utilisant le fait que δq s'annulle en 0 et 1, on obtient

$$(A(q+t\delta q))'(t)_{t=0} = \int_0^1 \left(\frac{\partial L}{\partial q} - \frac{d}{dt}(\frac{\partial L}{\partial \dot{q}})\right)\delta q ds = 0.$$

Ceci étant valable quelque soit δq, on montre que cela implique que

$$\frac{\partial L}{\partial q} - \frac{d}{dt}(\frac{\partial L}{\partial \dot{q}}) = 0 \text{ sur } [0,1].$$

Case 2D5000, 292 rue Saint-Martin 75141 Paris Cedex 03
Bureau 17.0.14

Tel: 0158808765, Mel: thierry.horsin@cnam.fr

Conservatoire National des Arts et Métiers

Département d'Ingénierie Mathématique