Université Joseph Fourier, Grenoble I
MAP350

Fonctions, Espaces vectoriels : un bref résumé [1]

1 Espaces vectoriels

Nous ne referons pas le cours d'algèbre linéaire : le contenu d'un cours de première année est une base de travail. Au fur et à mesure des notions de ce chapitre, on révisera donc les concepts d'espaces vectoriels, de sous-espace vectoriel, de système générateur, de rang, de bases, de dimension, d'application linéaire. On révisera également les notions fondamentales sur les matrices et le calcul matriciel.

2 Fonctions

2.1 Fonction d'une variable réelle

Définition 1 *On appelle fonction réelle de la variable réelle une application définie sur une partie $D \in \mathbb{R}$ à valeurs réelles. On note $\mathbb{R}^D$ l'ensemble de ces fonctions.*

En général D sera un intervalle.

2.2 Opérations sur les fonctions de $\mathbb{R}^D$

On sait de façon naturelle définir l'addition, la multiplication, la multiplication par un scalaire réel.

On se remémorera aussi l'opération de composition des applications.

On a quelquefois besoin de définir des opérations "ensemblistes" :

– restriction : Si f est définie sur D et $A \subset D$ on peut considérer la restriction de f à A :

$$f_{|A} : \begin{array}{l} A \to \mathbb{R} \\ x \to f(x) \end{array}$$

– prolongement : inversement si $f \in \mathbb{R}^D$ est la restriction d'une fonction $g \in \mathbb{R}^{D'}$ avec $D \subset D'$ alors g est un prolongement de f.

2.3 Propriétés des fonctions

On révisera : image, image réciproque, fonction croissante, strictement croissante, décroissante, strictement décroissante, monotone, majorée, minorée, bornée, paire, impaire, périodique,...

Nous allons manipuler de façon intensive des fonctions définies "par morceaux" et non pas par une formule unique. Ceci justifie un passage en revue de notions classiques de limite, continuité, dérivabilité,... en insistant sur les dites notions "à gauche" et "àdroite".

2.4 Limite

La notion de limite demande la notion de voisinage d'un point a pour une fonction $f \in \mathbb{R}^D$. Dans un premier temps nous pouvons considérer un voisinage de a comme un intervalle ouvert contenant a.

1. Pour une révision des notions de base, on pourra consulter les cours en ligne de l'année L1 à l'UJF :
http ://ljk.imag.fr/membres/Bernard.Ycart/mel/
http ://dlst.ujf-grenoble.fr/index.php ?module=cours

Définition 2 *Soit a un réel et f une fonction définie sur un voisinage de a sauf peut-être en a, et à valeurs dans $\mathbb{R}$. Soit ℓ un réel. On dit que f tend vers ℓ quand x tend vers a, ou que f a pour limite ℓ en a si :*

$$\forall \epsilon > 0, \exists \eta > 0, (0 <\mid x - a \mid \leq \eta) \Longrightarrow (\mid f(x) - \ell \mid \leq \epsilon).$$

Notation On note : $\lim\limits_{x \to a} f(x) = \ell$ ou $f(x) \underset{x \to a}{\longrightarrow} \ell$.

On peut compléter naturellement cette définition par les suivantes.

Définition 3 *Soit a un réel et f une fonction définie sur un voisinage à gauche de a sauf peut-être en a, et à valeurs dans $\mathbb{R}$. Soit ℓ un réel. On dit que f tend à gauche vers ℓ quand x tend vers a, ou que f a pour limite à gauche ℓ en a si :*

$$\forall \epsilon > 0, \exists \eta > 0, (a - \eta \leq x < a) \longrightarrow (\mid f(x) - \ell \mid \leq \epsilon).$$

Notation On note : $\lim\limits_{x \to a^-} f(x) = \ell$ ou $f(x) \underset{x \to a^-}{\longrightarrow} \ell$. On note également $\ell = f(a^-)$.

Définition 4 *Soit a un réel et f une fonction définie sur un voisinage à droite de a sauf peut-être en a, et à valeurs dans $\mathbb{R}$. Soit ℓ un réel. On dit que f tend à droite vers ℓ quand x tend vers a, ou que f a pour limite à droite ℓ en a si :*

$$\forall \epsilon > 0, \exists \eta > 0, (a < x \leq a + \eta) \longrightarrow (\mid f(x) - \ell \mid \leq \epsilon).$$

Notation On note : $\lim\limits_{x \to a^+} f(x) = \ell$ ou $f(x) \underset{x \to a^+}{\longrightarrow} \ell$, et $\ell = f(a^+)$.

Proposition 1 *La fonction f tend vers ℓ lorsque x tend a si et seulement si f tend à droite vers ℓ quand x tend vers a et f tend à gauche vers ℓ quand x tend vers a,*

Cette proposition nous sera particulièrement utile lors de la manipulation de fonctions définies par morceaux.

2.5 Continuité

Définition 5 *Une fonction f est continue (resp. à gauche, à droite) en a quand elle admet $f(a)$ comme limite (resp. à gauche, à droite) en a.*

Proposition 2 *Une fonction est continue en a si et seulement si elle est continue à gauche et à droite en a*

Il convient évidemment d'adapter les définitions lorsque le point a constitue l'extrémité de l'intervalle de définition.

2.6 Dérivabilité

Pour simplifier la présentation, on considérera des fonctions définies sur un intervalle ouvert $I =]a, b[$, $I =]-\infty, b[$, $I =]a, +\infty[$, $I = \mathbb{R}$. Les extensions se feront de façon naturelle.

La notion de dérivée d'une fonction $f \in \mathbb{R}^I$ est bien connue et on peut la résumer ainsi :

Définition 6 *On dit que f est dérivable en a si le taux d'accroissement $\frac{f(x)-f(a)}{x-a}$ tend vers une limite lorsque x tend vers a. Cette limite est notée $f'(a)$.*

On a évidemment des notions identiques pour les dérivées à gauche est à droite :

Définition 7 *On dit que f est dérivable à gauche (resp. à droite) en a si le taux d'accroissement $\frac{f(x)-f(a)}{x-a}$ tend vers une limite lorsque x tend vers a à gauche (resp. à droite).*

Proposition 3 *La dérivabilité est équivalente à l'existence et l'égalité des dérivées à droite et à gauche.*

Si une fonction f est dérivable en chaque point d'un intervalle I ceci nous permet de définir la fonction dérivée f' sur I.

2.7 Des ensembles de fonctions liées à la notion de continuité et de dérivée

Définition 8 *On note* $\mathcal{C}^0([a,b])$ *ou* $\mathcal{C}([a,b])$ *l'ensemble des fonctions continues sur* $[a,b]$. *On note* $\mathcal{C}^1([a,b])$ *l'ensemble des fonctions continues sur* $[a,b]$ *possédant une dérivée continue sur* $[a,b]$.

On peut de proche en proche définir les dérivées successives et par exemple :

Définition 9 *On note* $\mathcal{C}^2([a,b])$ *l'ensemble des fonctions continues sur* $[a,b]$ *possédant une dérivée et une dérivée seconde continues sur* $[a,b]$.

Par récurrence on définira $\mathcal{C}^n([a,b])$, l'ensemble des fonctions n fois continûment dérivable sur $[a,b]$. On dit également que f est de classe $\mathcal{C}^n$.

2.8 Les fonctions définies par morceaux

On s'intéressera souvent à des fonctions définies par morceaux. Par exemple, on se donne des réels $x_0 < x_1 < x_2$ et une fonction f définie de la façon suivante :

$$\begin{cases} f(x) = g(x) & x \in \ [x_0, x_1[\\ f(x) = h(x) & x \in \ ]x_1, x_2] \end{cases}$$

où g et h sont des fonctions continues respectivement sur les intervalles fermés $[x_0, x_1]$ et $[x_1, x_2]$, par exemple des polynômes.

2.8.1 Continuité

Pour obtenir la continuité de f on fera les remarques suivantes :
- les restrictions de f à chacun des intervalles $[x_0, x_1[$ et $]x_1 x_2]$ sont continues,
- la seule condition à vérifier pour que l'on puisse définir une fonction f continue est donc que :

$$\lim_{x \to {x_1}^-} f(x) = \lim_{x \to {x_1}^-} g(x) = \lim_{x \to {x_1}^+} h(x) = \lim_{x \to {x_1}^+} f(x),$$

- puisque les fonctions g et h sont fonctions continues respectivement sur les intervalles fermés $[x_0, x_1]$ et $[x_1, x_2]$ cette condition est équivalente à dire que $g(x_1) = h(x_1)$.

En fait on peut alors définir f de façon cohérente par :

$$\begin{cases} f(x) = g(x) & x \in \ [x_0, x_1] \\ f(x) = h(x) & x \in \ [x_1, x_2] \end{cases}$$

avec des intervalles fermés.

2.8.2 Dérivabilité

Si les fonctions g et h sont dérivables respectivement sur les intervalles fermés $[x_0, x_1]$ et $[x_1, x_2]$, alors aux conditions suivantes il faut rajouter :

$$\lim_{x \to {x_1}^-} \frac{f(x) - f(a)}{x - a} = \lim_{x \to {x_1}^-} \frac{g(x) - g(a)}{x - a} = \lim_{x \to {x_1}^+} \frac{h(x) - h(a)}{x - a} = \lim_{x \to {x_1}^+} \frac{f(x) - f(a)}{x - a}.$$

Cette condition est équivalente à : $g'(x_1) = h'(x_1)$.

De plus si les fonctions g et h sont dérivables et à dérivées continues respectivement sur les intervalles fermés $[x_0, x_1]$ et $[x_1, x_2]$, alors il en est de même de la fonction f sur l'intervalle $[x_0, x_2]$.

2.8.3 Dérivée seconde

On peut énoncer de la même manière que si les fonctions g et h sont deux fois dérivables respectivement sur les intervalles fermés $[x_0, x_1]$ et $[x_1, x_2]$, alors il en est de même de la fonction f si on ajoute la condition : $g''(x_1) = h''(x_1)$.

De plus si les fonctions g et h sont deux fois continuement dérivables respectivement sur les intervalles fermés $[x_0, x_1]$ et $[x_1, x_2]$, alors il en est de même de la fonction f sur l'intervalle $[x_0, x_2]$.

2.8.4 Généralisation : dérivée $n^{ième}$

On étendra ces propositions à des fonctions g et h dérivables n fois sur $[x_0, x_1]$ et $[x_1, x_2]$, et à des fonctions g et h, n fois continuement dérivables sur $[x_0, x_1]$ et $[x_1, x_2]$

2.8.5 A titre d'exercice : une autre généralisation

On considère une fonction définie par morceaux sur chacun des intervalles :

$$x_0 < x_1 < x_2 \cdots < x_n$$

chaque restriction aux intervalles $[x_i, x_{i+1}]$ étant un polynôme.

Énoncer une propriété assurant que cette fonction appartienne à $\mathcal{C}^m([x_0, x_n])$.

S'agit-il d'une condition nécessaire et suffisante ?

3 Espaces vectoriels de fonctions

3.1 L'espace vectoriel $\mathbb{R}^D$

L'ensemble $\mathbb{R}^D$ muni de l'addition de fonctions et de la multiplication par un scalaire réel en fait un espace vectoriel sur $\mathbb{R}$. Les espaces vectoriels E dont nous parlerons par la suite, seront en fait des sous-espaces vectoriels de $\mathbb{R}^D$ et en fait pour chacun de ces ensembles F il suffira de vérifier les propriétés d'un sous-espace vectoriel :
- $\forall f, g \in F, f + g \in F$
- $\forall f \in F, \forall \lambda \in \mathbb{R}, \lambda f \in F$.

3.2 Des sous-espaces vectoriels bien utiles

Dans l'espace vectoriels de $\mathbb{R}^I$ où I est un intervalle de $\mathbb{R}$ on vérifie sans difficulté que les ensembles suivant sont des sous-espaces vectoriels : $\mathcal{C}(I)$, $\mathcal{C}^1(I)$, $\mathcal{C}^2(I)$, ..., $\mathcal{C}^n(I)$,... On rencontrera particulièrement $\mathcal{C}([a, b])$, $\mathcal{C}^1([a, b])$, $\mathcal{C}^2([a, b])$, ..., $\mathcal{C}^n([a, b])$.

Les **polynômes** constituent un sous-espace vectoriel de $\mathbb{R}^{\mathbb{R}}$.

Les **polynômes de degré inférieur ou égal à** n forment un sous espace vectoriel de $\mathbb{R}^{\mathbb{R}}$ de dimension $n + 1$ dont on sait que les polynômes $1, x, x^2, \cdots, x^n$ forment une base naturelle. Ce n'est souvent pas la base la plus adaptée à la résolution de problèmes et on verra que le choix d'autres bases est souvent plus pertinent.

Les **fonctions définies par morceaux sur des subdivisions** : d'autres exemples auront beaucoup d'importance en CAO et dessin par ordinateur par exemple ceux de fonctions définies des morceaux de polynômes sur une subdivision :

$$a = x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n = b.$$

On y reviendra en détail.

Université de Grenoble
MAP350 – L3 MI

Interpolation polynomiale

1 Interpolation polynomiale de Lagrange

On suppose donnés $n+1$ points distincts dans $\mathbb{R}$: $x_0,\ x_1,\ \cdots,\ x_n$ ainsi que $n+1$ nombres réels : $y_0,\ y_1,\ \cdots,\ y_n$ et on cherche un polynôme vérifiant :

$$p(x_i) = y_i, \quad i = 0, 1,\ \ldots,\ n.$$

Le nombre de conditions (égal à $n+1$), conduit naturellement à chercher un polynôme de degré inférieur ou égal à n c'est-à-dire un polynôme dans $\mathbb{R}_n[x]$ qui est un espace vectoriel de dimension $n+1$.

1.1 Rappels

Soit $p(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n$ un polynôme de degré n (donc $a_n \neq 0$).

- $\bar{x} \in \mathbb{R}$ (ou $\mathbb{C}$) est une racine de multiplicité k de $p(x)$ si et seulement si

$$p(\bar{x}) = p'(\bar{x}) = p''(\bar{x}) = \cdots = p^{(k-1)}(\bar{x}) = 0 \quad \text{et} \quad p^{(k)}(\bar{x}) \neq 0.$$

- $\bar{x} \in \mathbb{R}$ (ou $\mathbb{C}$) est une racine de multiplicité k de $p(x)$ si et seulement si $p(x)$ se factorise par $(x-\bar{x})^k$ et pas par $(x-\bar{x})^{k+1}$, autrement dit s'il existe un polynôme $q(x)$ de degré $n-k$ tel que

$$p(x) = (x-\bar{x})^k\, q(x) \quad \text{et} \quad q(\bar{x}) \neq 0.$$

- Tout polynôme de degré n, à coefficients réels ou complexes, admet exactement n racines complexes, comptées avec leur multiplicité :

$$p(x) = a_n\,(x-z_1)^{k_1}\,(x-z_2)^{k_2}\,\cdots(x-z_r)^{k_r} \qquad \text{avec } \ k_1 + k_2 + \cdots + k_r = n,$$

  où $z_1, z_2, ..., z_r$ sont les racines distinctes de $p(x)$.
- Ainsi, si $p(x)$ admet $n+1$ racines comptées avec leur multiplicité, alors $p(x)$ est le polynôme identiquement nul.

1.2 Existence de la solution

Considérons l'application ϕ :

$$\begin{array}{lcl} \mathbb{R}_n[x] & \xrightarrow{\phi} & \mathbb{R}^{n+1} \\ p & \longrightarrow & \phi(p) = \big(p(x_0), p(x_1), \cdots, p(x_n)\big). \end{array}$$

On montre sans difficulté que ϕ est une application linéaire.

Théorème 1 *Si les points $x_0,\ x_1,\ \cdots,\ x_n$ sont distincts l'application linéaire ϕ est bijective.*

Démonstration : On sait que :

$$\dim(\mathbb{R}_n[x]) = \dim \text{ Ker } (\phi) + \dim \text{ Im } (\phi).$$

Le noyau de ϕ est constitué des polynômes de degré inférieur ou égal à n s'annulant en les $n+1$ points x_i et ce noyau est donc réduit au polynôme nul. Le noyau est donc de dimension zéro ce qui signifie que ϕ est injective. Il en résulte que Im (ϕ) est par conséquent de dimension $n+1$ et donc égal à $\mathbb{R}^{n+1}$ ce qui signifie que ϕ est surjective.
Fin de la démonstration

Il existe donc un unique polynôme $p \in \mathbb{R}_n[x]$ vérifiant les conditions :

$$p(x_i) = y_i,\ i = 0, 1,\ \cdots,\ n,$$

on l'appelle le polynôme d'interpolation.

Quand les valeurs y_i sont les valeurs d'une fonction f on parle du polynôme d'interpolation de f aux points x_i et on le note souvent $P(., f)$.

On peut en fait montrer l'équivalence suivante :

Proposition 1 *L'application linéaire ϕ est bijective, si et seulement si les $n+1$ points x_i, sont tous distincts.*

2 Détermination du polynôme d'interpolation

La méthode de détermination du polynôme d'interpolation de Lagrange dépend de la base choisie pour exprimer ce polynôme. Nous considérons dans un premier temps la détermination de ce polynôme dans la base canonique (ou base des monômes). Nous verrons ensuite la détermination du polynôme d'interpolation dans 2 bases spécialement adaptées au problème d'interpolation de Lagrange : la base de Lagrange, puis la base de Newton.

2.1 Dans la base canonique

Si on écrit le polynôme p :

$$p(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n,$$

les conditions d'interpolation se traduisent par le système linéaire de $n+1$ équations à $n+1$ inconnues :

$$\begin{array}{ccccccccc}
a_0 & + & a_1 x_0 & + & a_2 x_0^2 & + \cdots + & a_n x_0^n & = & y_0 \\
a_0 & + & a_1 x_1 & + & a_2 x_1^2 & + \cdots + & a_n x_1^n & = & y_1 \\
a_0 & + & a_1 x_2 & + & a_2 x_2^2 & + \cdots + & a_n x_2^n & = & y_2 \\
\vdots & & \vdots & & \vdots & & \vdots & & \vdots \\
a_0 & + & a_1 x_n & + & a_2 x_n^2 & + \cdots + & a_n x_n^n & = & y_n
\end{array}$$

ou

$$\begin{pmatrix} 1 & x_0 & x_0^2 & \cdots & x_0^n \\ 1 & x_1 & x_1^2 & \cdots & x_1^n \\ 1 & x_2 & x_2^2 & \cdots & x_2^n \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^n \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} y_0 \\ y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

La matrice du système est ce que l'on appelle la matrice de Van der Monde de déterminant :

$$\prod_{0 \leq i < j \leq n} (x_j - x_i),$$

qui n'est évidemment pas nul si les points sont distincts.

Cette méthode, n'est numériquement pas très bonne : elle n'est pas très stable ce qui entraîne que les erreurs d'arrondis dans les calculs perturbent fortement le résultat.

2.2 Base de Lagrange

Les polynômes de Lagrange associés aux $n+1$ points distincts :

$$x_0,\ x_1,\ x_2\ , \cdots,\ x_n,$$

sont définis par :

$$L_i(x) = \prod_{\substack{0 \leq j \leq n \\ j \neq i}} \frac{x - x_j}{x_i - x_j},\ 0 \leq i \leq n.$$

Les propriétés suivantes sont faciles à démontrer :

Propriété 1 *Pour* $0 \leq i \leq n$ *et* $0 \leq j \leq n$ *:*

$$L_i(x_j) = \delta_{ij}.$$

Propriété 2 *Les polynômes de Lagrange :*

$$L_i(x),\ 0 \leq i \leq n,$$

forment une base de $\mathbb{R}_n[x]$.

Propriété 3 *L'unique polynôme* $p \in \mathbb{R}_n[x]$ *vérifiant* $p(x_i) = y_i,\ i = 0, 1, \ldots, n$ *est donné par :*

$$p = \sum_{i=0}^{i=n} y_i\, L_i,$$

ou :

$$p(x) = \sum_{i=0}^{i=n} y_i\, L_i(x).$$

En dépit de son expression simple cette expression du polynôme d'interpolation n'est pas toujours la mieux adaptée pour les calculs.

2.3 Base de Newton et différences divisées

On considère les polynômes suivants de $\mathbb{R}_n[x]$:

$$
\begin{aligned}
&N_0(x) = 1, \\
&N_1(x) = x - x_0, \\
&N_2(x) = (x - x_0)(x - x_1), \\
&\cdots\cdots \\
&N_n(x) = (x - x_0)(x - x_1)\cdots(x - x_{n-1}).
\end{aligned}
$$

Ces polynômes sont donc déterminés par le polynôme $N_0(x)$ et la relation de récurrence :

$$N_{i+1}(x) = (x - x_i)\, N_i(x), \quad i = 0, 1, \ldots, n-1,$$

de sorte que l'on peut écrire :

$$N_i(x) = \prod_{j=0}^{j=i-1} (x - x_j),\ i = 1, 2, \ldots, n.$$

Remarque 1 *Chaque polynôme $N_k(x)$ est exactement de degré k et son terme de plus haut degré est exactement x^k :*

$$N_k(x) = x^k + q(x) \quad \textit{avec} \quad \deg\big(q(x)\big) \le k - 1\,.$$

Proposition 2 *Les polynômes $N_i(x)$, $i = 0, 1, \ldots, n$, forment une base de $\mathbb{R}_n(x)$, appellée base de Newton (associée aux points x_i).*

Démonstration : La dimension de $\mathbb{R}_n(x)$ étant égale à $n + 1$, il suffit de montrer que les $n + 1$ polynômes $N_i(x)$ sont linéairement indépendants. On montre facilement que si l'on prend une combinaison linéaire :

$$\sum_{i=0}^{n} \alpha_i N_i(x),$$

nulle pour toute valeur de x alors tous les α_i sont nuls. On commence par exemple en considérant la valeur $x = x_0$ ce qui donne :

$$\sum_{i=0}^{n} \alpha_i N_i(x_0) = \alpha_0 = 0.$$

On choisit ensuite $x = x_1$ et ainsi de suite.
Fin de la démonstration

Lemme 1 *Soit $p(x)$ un polynôme quelconque de degré k que l'on exprime dans la base canonique ainsi que dans la base de Newton (associée aux points x_i) :*

$$
\begin{aligned}
p(x) &= a_0 + a_1\, x + a_2\, x^2 + \cdots + a_k\, x^k, \quad (a_k \neq 0), \\
&= d_0\, N_0(x) + d_1\, N_1(x) + d_2\, N_2(x) + \cdots + d_k\, N_k(x).
\end{aligned}
$$

Alors : $d_k = a_k$.

• Soit f une fonction définie sur un intervalle I contenant l'ensemble des points d'interpolation x_j (ces points étant 2 à 2 distincts et en nombre quelconque). Pour tout entier k, on note $P_k(.,f)$ le polynôme d'interpolation de f aux points $x_0, x_1, \ldots, x_k$, que l'on exprime dans la base de Newton (associée à ces points) :

$$P_k(x,f) = \delta_0\, N_0(x) + \delta_1\, N_1(x) + \delta_2\, N_2(x) + \cdots + \delta_k\, N_k(x).$$

On a alors les résultats suivants.

Proposition 3 *Pour $k = 1, 2, ...,$ on a :* $\quad P_k(x,f) = P_{k-1}(x,f) + \delta_k\, N_k(x).$

Démonstration : Le polynôme $P_k(x,f) - P_{k-1}(x,f)$ s'annule aux points $x_0, \ldots, x_{k-1}$ et il est de degré inférieur ou égal k. Il est donc de la forme $\alpha N_k(x)$ et cette constante α est égale à δ_k car le terme de plus haut degré ne figure que dans $N_k(x)$.
Fin de la démonstration

Corollaire 1 *On déduit de cette proposition que si le polynôme d'interpolation $P_{k-1}(.,f)$ de f aux points $x_0, x_1, \ldots, x_{k-1}$ est connu, alors la détermination du polynôme d'interpolation $P_k(.,f)$ ne nécessite que le calcul d'un seul coefficient supplémentaire, à savoir δ_k.*

Les coefficients δ_k sont appelés *différences divisées* d'ordre k de f aux points $x_0,\ x_1, \ldots, x_k$ et sont notés plus précisément $\delta_k = \delta[x_0, \ldots, x_k; f]$. Un calcul simple montre que

$$\delta_0 = \delta[x_0; f] = f(x_0) \quad \text{ et } \quad \delta_1 = \delta[x_0, x_1; f] = \frac{f(x_1) - f(x_0)}{x_1 - x_0}.$$

Plus généralement, on a le résultat suivant qui permet le calcul des différences divisées selon un schéma dit *triangulaire.*

Proposition 4 *Pour $k = 1, 2, ...,$ on a :*

$$\delta[x_0, x_1, \ldots, x_{k-1}, x_k; f] = \frac{\delta[x_1, \ldots, x_{k-1}, x_k; f] - \delta[x_0, x_1, \ldots, x_{k-1}; f]}{x_k - x_0}.$$

Démonstration : On considère $q_{k-1}(x)$ le polynôme d'interpolation de f aux points $x_1, \cdots, x_k$, puis on construit :

$$p(x) = \frac{x - x_0}{x_k - x_0} \times q_{k-1}(x) - \frac{x - x_k}{x_k - x_0} \times P_{k-1}(x,f).$$

Des calculs simples montrent que polynôme $p(x)$ vérifie $p(x_i) = f(x_i),\ i = 0, \cdots, k$, que son degré est inférieur ou égal à k et que donc il s'agit de $P_k(.,f)$. Son terme de degré k est obtenu en multipliant par $\frac{x}{x_k - x_0}$ le terme de degré $k-1$ de q_{k-1} que l'on peut appeler $\delta[x_1, \cdots, x_{k-1}, x_k; f]$ et en lui retranchant le produit par $\frac{x}{x_k - x_0}$ du terme de degré $k-1$ du polynôme $P_{k-1}(x,f)$ ce qui nous donne la formule.
Fin de la démonstration

Cette proposition fournit un algorithme de calcul des différences divisées, et donc du polynôme d'interpolation sur la base de Newton.

2.4 Evaluation d'un polynôme : schéma de Horner

Soit $p(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$. Le tracé du graphe de cette fonction polynomiale sur un intervalle $[a, b]$ conduit à évaluer $p(t_i)$ pour de nombreuses valeurs $t_i = a + i * (b - a)/N,\ i = 0, 1, ..., N$; D'où, l'intérêt de développer un algorithme performant pour l'évaluation d'un polynôme.

L'algorithme ci-dessous (dit naïf) calcule $p(x)$ à l'aide de $2n$ multiplications.

```
% data : un réel x et le vecteur a des coefficients a(i)
xn = 1
p = a(0)
pour i = 1 jusqu'à n faire
    xn = xn * x
    p = p + a(i) * xn
fin pour
```

L'algorithme suivant, appelé **schéma de Horner** calcule le polynôme $p(x)$ à l'aide de seulement n multiplications, de la façon suivante.

```
% data : un réel x et le vecteur a des coefficients a(i)
p = a(n)
pour i = n-1 jusqu'à 0 avec un pas de -1 faire
    p = p * x + a(i)
fin pour
```

Pour un polynôme écrit dans la base de Newton, le schéma de Horner s'applique sous la forme suivante.

```
% data : un réel x et le vecteur a des coefficients c(i)
% et des points d'interpolation x(i)
p = c(n)
pour i = n-1 jusqu'à 0 avec un pas de -1 faire
    p = p * (x - x(i)) + c(i)
fin pour
```

3 Erreur dans l'interpolation de Lagrange

On se donne $n+1$ points distincts dans un intervalle $I = [a, b]$: $x_0,\ x_1,\ \cdots,\ x_n$ et une fonction $f \in \mathcal{C}^{n+1}(I)$.

Désignons par p_n, le polynôme d'interpolation de f aux points x_i.

Pour x fixé dans I et différent des points x_i, considérons la fonction :

$$\phi(u) = f(u) - p_n(u) - (f(x) - p_n(x)) \times \frac{\prod_{i=0}^{i=n}(u - x_i)}{\prod_{i=0}^{i=n}(x - x_i)}.$$

La fonction ϕ s'annule aux points x_i et également en x donc en $n+2$ points de I. D'après le théorème de Rolle on en déduit que sa dérivée s'annule en $n+1$ points distincts de I. Etant donnée que la fonction ϕ appartient également à $\mathcal{C}^{n+1}(I)$ on en déduit par récurrence que $\phi^{(n+1)}$ s'annule en un point ξ de I. La dérivée $(n+1)^{\text{ème}}$ du polynôme $\prod_{i=0}^{i=n}(u-x_i)$ est constante et égale à $(n+1)!$ et la dérivée $(n+1)^{\text{ème}}$ de p_n est identiquement nulle. Il vient donc :

$$\phi^{(n+1)}(\xi) = 0 = f^{(n+1)}(\xi) - (f(x) - p_n(x)) \times \frac{(n+1)!}{\prod_{i=0}^{i=n}(x-x_i)},$$

et la formule :

$$f(x) - p_n(x) = \frac{\prod_{i=0}^{i=n}(x-x_i)}{(n+1)!} f^{(n+1)}(\xi).$$

Cette formule est également valable pour les valeurs x_i. Elle permet donc d'énoncer le théorème suivant :

Théorème 2 *Soit $I = [a,b]$, $f \in \mathcal{C}^{n+1}(I)$, $x_0,\ x_1,\ \cdots,\ x_n$ $n+1$ points distincts de I et p_n le polynôme d'interpolation de f en ces points.*

Pour tout $x \in I$ il existe $\xi \in I$ tel que :

$$f(x) - p_n(x) = \frac{\prod_{i=0}^{i=n}(x-x_i)}{(n+1)!} f^{(n+1)}(\xi).$$

On en déduit par exemple que :

Corollaire 2

$$\mid f(x) - p_n(x) \mid \leq \frac{\mid \prod_{i=0}^{i=n}(x-x_i) \mid}{(n+1)!} \underset{a \leq \xi \leq b}{\text{Max}} \mid f^{(n+1)}(\xi) \mid, \forall\, tx \in\ [a,b].$$

4 Interpolation de Hermite

On considère $n+1$ points distincts $x_0, x_1, \ldots, x_n$ d'un intervalle $[a,b]$ ainsi que deux familles de $n+1$ nombres réels :

$$y_0, y_1, \ldots, y_n \qquad \text{et} \qquad y'_0, y'_1, \ldots, y'_n\,,$$

et on cherche un polynôme $p(x)$ vérifiant :

$$\begin{cases} p(x_i) = y_i, & i = 0, 1, ..., n, \\ p'(x_i) = y'_i, & i = 0, 1, ..., n. \end{cases}$$

Les valeurs y_i et y'_i pourront être les valeurs d'une fonction $f \in \mathcal{C}^1([a,b])$ et de sa dérivée, à savoir $y_i = f(x_i)$ et $y'_i = f'(x_i)$. C'est dans ce cadre que nous présenterons ce que l'on appelle l'**interpolation d'Hermite**.

L'observation du nombre de degré de liberté nous indique que nous avons $2n+2$ conditions de sorte qu'il est naturel de chercher notre polynôme $p(x) = p_{2n+1}(x)$ dans $\mathbb{R}_{2n+1}[x]$.

Par analogie avec l'interpolation de Lagrange, nous proposons de construire une base de polynômes $\{h_i(x), \bar{h}_i(x),\ i = 0, 1, ..., n\}$ de $\mathbb{R}_{2n+1}[x]$ permettant d'écrire le polynôme d'interpolation de Hermite sous la forme :

$$p_{2n+1}(x) = \sum_{i=0}^{i=n} f(x_i)\, h_i(x) + \sum_{i=0}^{i=n} f'(x_i)\, \bar{h}_i(x).$$

Précisément, ces polynômes $h_i(x)$ et $\bar{h}_i(x)$ doivent satisfaire les conditions résumées dans le tableau suivant.

	h_0	h_1	$\cdots$	h_i	$\cdots$	h_n	$\bar{h}_0$	$\bar{h}_1$	$\cdots$	$\bar{h}_i$	$\cdots$	$\bar{h}_n$
$f(x_0)$	1	0	$\cdots$	0	$\cdots$	0	0	0	$\cdots$	0	$\cdots$	0
$f(x_1)$	0	1	$\cdots$	0	$\cdots$	0	0	0	$\cdots$	0	$\cdots$	0
$\vdots$	$\vdots$	$\vdots$	$\ddots$			$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$
$f(x_i)$	0	0		1		0	0	0	$\cdots$	0	$\cdots$	0
$\vdots$	$\vdots$	$\vdots$			$\ddots$	$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$
$f(x_n)$	0	0	$\cdots$	0	$\cdots$	1	0	0	$\cdots$	0	$\cdots$	0
$f'(x_0)$	0	0	$\cdots$	0	$\cdots$	0	1	0	$\cdots$	0	$\cdots$	0
$f'(x_1)$	0	0	$\cdots$	0	$\cdots$	0	0	1	$\cdots$	0	$\cdots$	0
$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$	$\vdots$	$\ddots$			$\vdots$
$f'(x_i)$	0	0	$\cdots$	0	$\cdots$	0	0	0		1		0
$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$	$\vdots$			$\ddots$	$\vdots$
$f'(x_n)$	0	0	$\cdots$	0	$\cdots$	0	0	0	$\cdots$	0	$\cdots$	1

Construction des polynômes $h_i(x)$

Le tableau ci-dessus donne les conditions à imposer à chaque polynôme $h_i(x)$:

$$\begin{cases} h_i(x_i) = 1 \quad \text{et} \quad h_i(x_j) = 0 \quad \text{si } j \neq i\,, \\ h_i'(x_j) = 0 \quad \forall\, j\,. \end{cases}$$

Nous allons rechercher des polynômes vérifiant ces conditions en utilisant les polynômes de Lagrange. Précisément, le polynôme $h_i(x)$ devant s'annuler ainsi que sa dérivée en tout point $x_j \neq x_i$, il s'écrit sous la forme :

$$h_i(x) = L_i^2(x)\, r_i(x),$$

où $r_i(x)$ est un polynôme de degré inférieur ou égal à 1.

Les conditions supplémentaires à réaliser pour $h_i(x)$ se traduisent donc par :

$$\begin{cases} h_i(x_i) = L_i^2(x_i) r_i(x_i) = r_i(x_i) = 1, \\ h_i'(x_i) = L_i^2(x_i) r_i'(x_i) + 2L_i(x_i)L_i'(x_i) r_i(x_i) = r_i'(x_i) + 2L_i'(x_i) = 0, \end{cases}$$

et nous en déduisons après un calcul simple :

$$r_i(x) = 1 - 2(x - x_i)L_i'(x_i).$$

Construction des polynômes $\bar{h}_i(x)$

De la même manière, on déduit du tableau ci-dessus les conditions à imposer à chaque polynôme $\bar{h}_i(x)$:

$$\begin{cases} \bar{h}_i(x_j) = 0 \quad \forall\, j\,, \\ \bar{h}'_i(x_i) = 1 \quad \text{et} \quad \bar{h}'_i(x_j) = 0 \quad \text{si } j \neq i\,. \end{cases}$$

La encore, nous allons rechercher le polynôme $\bar{h}_i(x)$ sous la forme :

$$\overline{h}_i(x) = L_i^2(x)s_i(x),$$

où $s_i(x)$ est un polynôme de degré inférieur ou égal à 1.

Les conditions supplémentaires à réaliser pour $h_i(x)$ sont :

$$\begin{cases} \overline{h}_i(x_i) = L_i^2(x_i)s_i(x_i) = s_i(x_i) = 0, \\ \overline{h}_i{}'(x_i) = s'_i(x_i)L_i^2(x_i) + 2L_i(x_i)L'_i(x_i)s_i(x_i) = s'_i(x_i) + 2L'_i(x_i)s_i(x_i) = s'_i(x_i) = 1, \end{cases}$$

et nous en déduisons :

$$s_i(x) = (x - x_i).$$

Polynôme d'interpolation de Hermite

Le polynôme recherché existe bien et se met donc sous la forme :

$$p_{2n+1}(x) = \sum_{i=0}^{i=n} h_i(x)f(x_i) + \sum_{i=0}^{i=n} \overline{h}_i(x)f'(x_i),$$

et plus précisément :

$$p_{2n+1}(x) = \sum_{i=0}^{i=n} [1 - 2(x - x_i)L'_i(x_i)]L_i^2(x)f(x_i) + \sum_{i=0}^{i=n} (x - x_i)L_i^2(x)f'(x_i).$$

Les polynômes $h_i(x)$ et $\overline{h}_i(x)$ sont appelés polynômes d'interpolation de Hermite.

Proposition 5 *Le polynôme d'interpolation d'Hermite est unique.*

Démonstration : Supposons qu'il existe deux polynômes $p_{2n+1}(x), q_{2n+1}(x) \in \mathbb{R}_{2n+1}[x]$ vérifiant les conditions d'interpolation. Le polynôme $r_{2n+1}(x) = p_{2n+1}(x) - q_{2n+1}(x)$ admet $n+1$ racines doubles distinctes. Comme il est de degré inférieur ou égal à $2n+1$ ce ne peut être que le polynôme nul.
Fin de la démonstration

Université de Grenoble
MAP350 – L3 MI

Fonction spline

1 Fonction spline : premières définitions

Il existe de très nombreuses variantes de la définition d'une fonction spline. De façon générale on peut dire qu'une fonction spline est une fonction polynomiale par morceaux, vérifiant des conditions de dérivabilité à différents ordres.

Dans ce chapitre, on considèrera :

$$a = x_0 < x_1 < x_2 < \cdots < x_n < x_{n+1} = b.$$

Attention, les notations sont légèrement différentes des chapitres précédents

On cherchera des ensembles Π_k^m de fonctions f tels que :
- f est sur chaque intervalle $[x_i, x_{i+1}]$ un polynome de degré $\leq m$,
- f appartient à $\mathcal{C}^k([a, b])$.

Propriété 1 *L'ensemble Π_k^m est un espace vectoriel de dimension finie.*

Démonstration : Il est facile de voir que Π_k^m est un espace vectoriel. Pour montrer qu'il est de dimension finie on le considérera comme sous-espace vectoriel d'un espace vectoriel de polynômes de degré $\leq m$ par morceaux que l'on formalisera.
Fin de la démonstration

Il n'est pas évident de parler toute de suite rigoureusement de dimension mais on peut faire quelques remarques sur le nombre de degrés de libertés :
- il y a $n+1$ morceaux de polynômes de degré $\leq m$ ce qui fournit $(n+1)\times(m+1)$ paramètres,
- les conditions de régularités aux points intérieurs $x_1 < x_2 < \cdots < x_n$ imposent pour l'ensemble des dérivées (y compris la continuité) : $n \times (k+1)$ conditions.

Ceci nous suggérerait des espaces vectoriels de dimension $(n+1)\times(m+1) - n \times (k+1)$ sans pour autant que ce soit une démonstration. En particulier, l'ensemble n'est jamais vide, même pour k grand, car un polynôme de degré inférieur ou égal à m appartient toujours à Π_k^m.

2 Fonctions splines cubiques

Nous nous intéresserons donc maintenant au cas :

$$a = x_0 < x_1 < x_2 < \cdots < x_n < x_{n+1} = b,$$

et à l'espace vectoriel Π_2^3.

Propriété 2 *L'ensemble* Π_2^3 *est un espace vectoriel de dimension* $n+4$.

Avant de procéder à la démonstration introduisons pour $i = 1, 2, \cdots n$ les fonctions $(x - x_i)_+$ définies sur $[a, b]$ par :

$$\begin{cases} (x - x_i)_+ = 0, \ x \leq x_i \\ (x - x_i)_+ = (x - x_i), \ x \geq x_i \end{cases}$$

On montre sans difficulté que pour $x \in [a, b]$:

$$\int_a^x (u - x_i)_+ du = \frac{1}{2}(x - x_i)_+^2,$$

et que si on intégre encore une fois de a à x on obtient :

$$\frac{1}{6}(x - x_i)_+^3.$$

Démonstration : Pour cette démonstration on remarquera qu'une fonction $f \in \Pi_2^3$ admet pour dérivée seconde une fonction linéaire par morceaux sur la subdivision

$$a = x_0 < x_1 < x_2 < \cdots < x_n < x_{n+1} = b$$

dont nous savons que c'est espace vectoriel. Nous utliserons comme base de cet espace vectoriel les fonctions :

$$\begin{cases} 1, x, \\ (x - x_i)_+, \ i = 1, 2, \cdots, n. \end{cases}$$

En intégrant deux fois la fonction f'' nous montrons que pour $x \in [a, b]$:

$$f(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \sum_{i=1}^{n} \delta_i \frac{1}{6}(x - x_i)_+^3.$$

On montre sans difficulté que les fonctions $1, x, x^2, x^3$ et $(x - x_i)_+^3, \ i = 1, 2, \cdots, n$ sont indépendantes et donc ceci montre le résultat.
Fin de la démonstration

3 Fonctions splines naturelles

Les notations sont similaires à celles du paragraphe précédent.

Nous disposons maintenant de $n + 4$ degrés de liberté : à priori si nous souhaitons imposer des conditions d'interpolation aux points $x_1, x_2, \cdots, x_n$ il nous reste 4 paramètres.

Définition 1 *L'espace* S *des* **splines naturelles** *consiste en l'ensemble des focntions splines cubiques de* Π_2^3 *vérifiant les conditions :*

$$f''(x) = 0, \ x \in [a, x_1] \cup [x_n, b].$$

Nous allons énoncer un certain nombre de propriétés que nous démontrerons ensuite.

Propriété 3 *L'ensemble S des splines naturelles est un espace vectoriel de dimension n.*

Propriété 4 *Etant données n valeurs $y_1, y_2, \cdots, y_n$ il existe une et une seule spline naturelle s vérifiant :*

$$s(x_i) = y_i,\ \ i = 1, 2, \cdots, n.$$

Et la dernière propriété et non des moindres :

Propriété 5 *Parmi toutes les fonctions de $\mathcal{C}^2([a,b]]$ vérifiant les conditions d'interpolation, la fonction spline naturelle minimise la quantité :*

$$\int_a^b (s^{"}(x))^2 dx$$

Démontrons la propriété 3

<u>*Démonstration :*</u> La structure d'espace vectoriel est facile à montrer. Nous avons vu que nous pouvions décrire tout élément de Π_2^3 sous la forme :

$$f(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \sum_{i=1}^{n} \delta_i \frac{1}{6}(x - x_i)_+^3.$$

La condition $f''(x) = 0,\ x \in [a, x_1]$ est équivalente à :

$$a_2 = a_3 = 0.$$

Prenant en compte ces conditions, l'autre condition $f''(x) = 0,\ x \in [x_n, b]$ est équivalente à ce que le polynôme :

$$f(x) = a_0 + a_1 x + \sum_{i=1}^{n} \delta_i \frac{1}{6}(x - x_i)^3,$$

soit de degré inférieur ou égal à 1 ce qui se traduit par :

$$\sum_{i=1}^{n} \delta_i = 0,\ \text{ et } \sum_{i=1}^{n} \delta_i x_i = 0.$$

On peut donc identifier l'ensemble des splines naturelles au sous espace vectoriel des $(n+4)$-uplets :

$$(a_0, a_1, a_2, a_3, \delta_1, \delta_2, \cdots, \delta_n),$$

vérifiant :

$$a_2 = 0, a_3 = 0, \sum_{i=1}^{n} \delta_i = 0,\ \text{ et } \sum_{i=1}^{n} \delta_i x_i = 0.$$

Pour montrer qu'il s'agit d'un sous-espace vectoriel de $\mathbb{R}^{n+4}$ de dimension n, nous considérons cet ensemble comme le noyau de l'application linéaire de $\mathbb{R}^{n+4}$ dans $\mathbb{R}^4$:

$$(a_0, a_1, a_2, a_3, \delta_1, \delta_2, \cdots, \delta_n) \to (a_2, a_3, \sum_{i=1}^{n} \delta_i, \sum_{i=1}^{n} \delta_i x_i).$$

Il est facile de voir que l'application linéaire est surjective et le théorème du rang nous donne le résultat final.
Fin de la démonstration

4 Une méthode de construction de la fonction spline naturelle d'interpolation

4.1 Les grandes lignes de la méthode

Nous allons en fait démontrer la propriété 4 en proposant une méthode de construction. Pour simplifier les calculs nous allons examiner uniquement le cas de points d'interpolations équidistants :

$$a = x_0 < x_1 < x_2 < \cdots < x_n < x_{n+1} = b,$$

avec $x_{i+1} - x_i = h,\ i = 0, 1, \cdots, n$.

La démarche est exactement la même dans le cas général mais techniquement plus complexe.

L'esprit de la méthode est la suivante : sur chacun des intervalles $[x_i, x_{i+1}]$ la fonction spline $s(x)$ est un polynôme $p_i(x)$ de degré inférieur ou égal à 3. D'après ce que nous avons vu à propos de l'interpolation d'Hermite, nous pouvons l'écrire de façon unique en fonction de ses valeurs et des valeurs de ses dérivées aux points x_i et x_{i+1}. Les valeurs de ces dérivées aux points x_i seront notées $y_i^{'}$: elles sont inconnues mais en écrivant les conditions de raccord des dérivées secondes :

$$p_{i-1}^{''}(x_i) = p_i^{''}(x_i),\ i = 2, 3, \cdots, n-1,$$

ainsi que :

$$p_1^{''}(x_1) = p_{n-1}^{''}(x_n) = 0,$$

nous allons pouvoir déterminer de façon unique les valeurs de ces dérivées $y_i^{'}$.

Avant de procéder au calcul il nous faut établir le lemme technique suivant :

Lemme 1 *Soit $\alpha \neq \beta$, $h = \beta - \alpha$ et p le polynôme de degré inférieur ou égal à 3 vérifiant :*

$$p(\alpha) = y_\alpha,\ p(\beta) = y_\beta,\ p^{'}(\alpha) = y_\alpha^{'},\ p^{'}(\beta) = y_\beta^{'}.$$

On a :

$$p^{''}(\alpha) = \frac{2}{h^2}(3y_\beta - 3y_\alpha - 2hy_\alpha^{'} - hy'_\beta), \quad (1)$$

et

$$p^{''}(\beta) = \frac{2}{h^2}(3y_\alpha - 3y_\beta + 2hy_\beta^{'} + hy'_\alpha). \quad (2)$$

Démonstration : Remarquons tout d'abord que les deux formules sont identiques à la permutation de α et β près sans oublier que h devient alors $-h$! Nous pourrions passer par une forme d'Hermite mais ce n'est pas le plus facile pour ce calcul. Ecrivons donc le polynôme que nous cherchons sous la forme d'un dévelopement de Taylor en α :

$$p(x) = y_\alpha + (x-\alpha)y_\alpha^{'} + \frac{(x-\alpha)^2}{2}p^{''}(\alpha) + \frac{(x-\alpha)^3}{6}p^{'''}(\alpha).$$

Nous avons tout d'abord :

$$p(\beta) = y_\alpha + (\beta-\alpha)y_\alpha^{'} + \frac{(\beta-\alpha)^2}{2}p^{''}(\alpha) + \frac{(\beta-\alpha)^3}{6}p^{'''}(\alpha) = y_\alpha + hy_\alpha^{'} + \frac{h^2}{2}p^{''}(\alpha) + \frac{h^3}{6}p^{'''}(\alpha),$$

puis ;

$$p^{'}(x) = y_\alpha^{'} + (x-\alpha)p^{''}(\alpha) + \frac{(x-\alpha)^2}{2}p^{'''}(\alpha),$$

et donc :

$$p^{'}(\beta) = y_\alpha^{'} + (\beta-\alpha)p^{''}(\alpha) + \frac{(\beta-\alpha)^2}{2}p^{'''}(\alpha) = y_\alpha' + hp^{''}(\alpha) + \frac{h^2}{2}p^{'''}(\alpha).$$

Le système d'équations :

$$\begin{cases} p(\beta) = y_\beta = y_\alpha + hy_\alpha^{'} + \frac{h^2}{2}p^{''}(\alpha) + \frac{h^3}{6}p^{'''}(\alpha) \\ p^{'}(\beta) = y_\beta^{'} = y_\alpha^{'} + hp^{''}(\alpha) + \frac{h^2}{2}p^{'''}(\alpha) \end{cases}$$

nous donne après élimination de $p_\alpha^{'''}$:

$$p^{''}(\alpha) = \frac{2}{h^2}(3y_\beta - 3y_\alpha - 2hy_\alpha^{'} - hy_\beta^{'}).$$

Fin de la démonstration

4.2 La méthode de construction

Nous allons écrire maintenant les conditions de raccord des dérivées secondes au point $x_i,\ 1 < i < n$. La dérivée seconde en x_i du polynôme définit sur l'intervalle $[x_{i-1}, x_i]$ est donné par la formule 2 en prenant

$$\alpha = x_{i-1}, \beta = x_i,$$

et avec des notations naturelles sa valeur est :

$$\frac{2}{h^2}(3y_{i-1} - 3y_i + 2hy_i^{'} + hy_{i-1}^{'}).$$

La dérivée seconde en x_i du polynôme définit sur l'intervalle $[x_i, x_{i+1}]$ est donné par la formule 1 en prenant

$$\alpha = x_i, \beta = x_{i+1},$$

et avec des notations naturelles sa valeur est :

$$\frac{2}{h^2}(3y_{i+1} - 3y_i - 2hy_i^{'} - hy_{i+1}^{'}).$$

En écrivant l'égalité on a :

$$y'_{i-1} + 4y'_i + y'_{i+1} = \frac{3}{h}(y_{i+1} - y_{i-1}),$$

et ceci pour $i = 2, 3, \cdots n-1$.

La condition de nullité de la dérivée seconde en x_1 s'écrit en utilisant la formule 1 avec :

$$\alpha = x_1, \beta = x_2,$$

et donne :

$$0 = \frac{2}{h^2}(3y_2 - 3y_1 - 2hy'_1 - hy'_2),$$

ou :

$$2y'_1 + y'_2 = \frac{3}{h}(y_2 - y_1).$$

De même la condition de nullité de la dérivée seconde en x_n s'écrit en utilisant la formule 2 avec :

$$\alpha = x_{n-1}, \beta = x_n,$$

et donne :

$$0 = \frac{2}{h^2}(3y_{n-1} - 3y_n + 2hy'_n + hy'_{n-1}).$$

ou :

$$y'_{n-1} + 2y'_n = \frac{3}{h}(y_n - y_{n-1}).$$

Globalement les conditions néessaires et suffisantes que doivent vérifier les dérivées aux points $x_1,\ x_2,\ \cdots,\ x_n$ sont données par le système linéaire :

$$\begin{array}{ccccccccccccccc}
2y'_1 & + & y'_2 & + & 0 & + & 0 & +\cdots+ & 0 & + & 0 & = & \frac{3}{h}(y_2 - y_1) \\
y'_1 & + & 4y'_2 & + & y'_3 & + & 0 & +\cdots+ & 0 & + & 0 & = & \frac{3}{h}(y_3 - y_1) \\
0 & + & y'_2 & + & 4y'_3 & + & y'_4 & +\cdots+ & 0 & + & 0 & = & \frac{3}{h}(y_4 - y_2) \\
\vdots & & \vdots & & \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\
0 & + & 0 & + & 0 & + & 0 & +\cdots+ & y'_{n-1} & + & 2y'_n & = & \frac{3}{h}(y_n - y_{n-1})
\end{array} \qquad (3)$$

Ceci peut s'écrire sous forme matricielle :

$$\begin{pmatrix}
2 & 1 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\
1 & 4 & 1 & 0 & 0 & \cdots & 0 & 0 & 0 \\
0 & 1 & 4 & 1 & 0 & \cdots & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & 0 & 0 & \cdots & 1 & 4 & 1 \\
0 & 0 & 0 & 0 & 0 & \cdots & 0 & 1 & 2
\end{pmatrix}
\begin{pmatrix}
y'_1 \\ y'_2 \\ y'_3 \\ y'_4 \\ y'_5 \\ \vdots \\ y'_{n-2} \\ y'_{n-1} \\ y'_n
\end{pmatrix}
= \frac{3}{h}
\begin{pmatrix}
y_2 - y_1 \\ y_3 - y_1 \\ y_4 - y_2 \\ y_5 - y_3 \\ y_6 - y_4 \\ \vdots \\ y_{n-1} - y_{n-3} \\ y_n - y_{n-2} \\ y_n - y_{n-1}
\end{pmatrix}$$

Propriété 6 *Le système linéaire 3 admet une solution unique.*

Cette propriété peut se démontrer de différentes manières.

5 Propriété de minimisation de la fonctions spline

Maintenant que nous savons que la spline naturelle d'interpolation existe, examinons maintenant la propriété 5.

<u>*Démonstration :*</u> Soit $s(x)$ la spline cubique naturelle et $y(x)$ une autre fonction interpolante de $\mathcal{C}^2[a,b]$. On écrit :

$$y(x) = s(x) + [y(x) - s(x)].$$

On peut écrire :

$$\begin{aligned}\int_{x_0}^{x_{n+1}} (y^{''}(x))^2 dx &= \int_{x_0}^{x_{n+1}} (s^{''}(x))^2 dx + 2\int_{x_0}^{x_{n+1}} (s^{''}(x))(y^{''}(x) - s^{''}(x))dx \\ &+ \int_{x_0}^{x_{n+1}} (y^{''}(x) - s^{''}(x))^2 dx.\end{aligned}$$

En intégrant par parties l'intégrale du milieu s'écrit :

$$\int_{x_0}^{x_{n+1}} (s^{''}(x))(y^{''}(x) - s^{''}(x))dx = (s^{''}(x)(y^{'}(x) - s^{'}(x)) \mid_{x_0}^{x_{n+1}} - \int_{x_0}^{x_{n+1}} (s^{'''}(x))(y^{'}(x) - s^{'}(x))dx.$$

La première expression est nulle à cause de conditions aux extrémités :

$$s^{''}(x_0) = s^{''}(x_{n+1}) = 0.$$

Dans la seconde expression, qui est, au signe près, tout d'abord égale à :

$$\int_{x_1}^{x_n} (s^{'''}(x))(y^{'}(x) - s^{'}(x)dx)$$

la dérivée troisième est constante par morceaux, ce qui implique :

$$\int_{x_1}^{x_n} (s^{'''}(x))(y^{'}(x) - s^{'}(x))dx = \sum_{i=1}^{i=n-1} s_j^{'''}(y(x) - s(x)) \mid_{x_i}^{x_{i+1}},$$

où $s_j^{'''}$ désigne une constante valeur de la dérivée troisième de s dans l'intervalle adéquat.

Etant données les conditions d'interpolation, cette quantité est donc nulle et on obtient finalement :

$$\int_{x_0}^{x_{n+1}} (y^{''}(x))^2 dx = \int_{x_0}^{x_{n+1}} (s^{''}(x))^2 dx + \int_{x_0}^{x_{n+1}} (y^{''}(x) - s^{''}(x))^2 dx.$$

d'où l'on déduit :

$$\int_{x_0}^{x_n} (y^{''}(x))^2 dx \geq \int_{x_0}^{x_n} (s^{''}(x))^2 dx.$$

.

<u>*Fin de la démonstration*</u>

Université Joseph Fourier, Grenoble I
MAP350

Résumé de cours sur les moindres carrés : partie théorique

Notations : Pour une matrice $A \in \mathcal{M}_{n,m}(\mathbb{R})$ on note sa transposée A' (elle est notée ${}^tA$ dans d'autres cours). Pour un vecteur $x \in \mathbb{R}^n$, en fait $x \in \mathcal{M}_{n,1}$, on identifie le scalaire $\| x \|_2^2$ (le carré de la norme euclidienne usuelle) avec la matrice 1×1 : $x'x$.

1 Systèmes linéaires aux moindres carrés

On considère une matrice $A \in \mathcal{M}_{n,m}(\mathbb{R})$ et un vecteur $b \in \mathbb{R}^n$. En général il n'existe pas de vecteur $x \in \mathbb{R}^m$ qui vérifie $Ax = b$ et on remplace le problème par le problème d'optimisation suivant :

$$\underset{x \in \mathbb{R}^m}{\text{Min}} \| Ax - b \|_2 \tag{1}$$

où $\| \ \|_2$ est la norme euclidienne usuelle de $\mathbb{R}^n$.

Nous montrerons dans la suite pourquoi ce problème admet toujours une solution (pas forcèment unique) mais nous l'admettrons pour le moment.

Propriété 1 *Soit S l'ensemble des solutions du problème de minimisation (1), c'est à dire :*

$$S = \{\hat{x} \in \mathbb{R}^m : \| A\hat{x} - b \|_2 = \underset{x \in \mathbb{R}^m}{\text{Min}} \| Ax - b \|_2\}.$$

On peut affirmer que $\hat{x} \in S$ si et seulement si :

$$A'(b - A\hat{x}) = 0.$$

Démonstration : Soit $\hat{x}$ et $\hat{r} = b - A\hat{x}$ vérifiant $A'\hat{r} = 0$. En posant pour un $x \in \mathbb{R}^m$ quelconque : $r = b - Ax$, on a :

$$r = b - Ax = b - A\hat{x} + A\hat{x} - Ax = \hat{r} + A(\hat{x} - x).$$

En notant : $e = \hat{x} - x$ on obtient :

$$r'r = (\hat{r}' + e'A')(\hat{r} + Ae) = \hat{r}'\hat{r} + \hat{r}'Ae + e'A'\hat{r} + e'A'Ae.$$

Comme il s'agit de matrices 1×1, les matrices $\hat{r}'Ae$ et $e'A'\hat{r}$ sont égales et la quantité précédente est égale à :

$$\hat{r}'\hat{r} + e'A'Ae.$$

Il vient donc que :

$$\| r \|_2^2 = \| \hat{r} \|_2^2 + \| Ae \|_2^2 \geq \| \hat{r} \|_2^2,$$

d'où l'on déduit que $\hat{x} \in S$.

Réciproquement, nous allons prouver que :

$$\hat{x} \in S \Longrightarrow A'(b - A\hat{x}) = 0,$$

que nous allons montrer en considérant la contraposée :

$$A'(b - A\hat{x}) \neq 0 \Longrightarrow \hat{x} \notin S.$$

Pour un tel $\hat{x}$ on pose pour $\epsilon \in \mathbb{R}$: $x = \hat{x} + \epsilon A'\hat{r}$ et il vient donc :

$$r = b - Ax = b - A\hat{x} + A(\hat{x} - x) = \hat{r} - \epsilon AA'\hat{r}.$$

Un calcul simple montre que :

$$r'r = \hat{r}\,'\hat{r} - 2\epsilon\hat{r}\,'AA'\hat{r} + \epsilon^2 \parallel AA'\hat{r} \parallel = \parallel \hat{r} \parallel_2^2 - 2\epsilon \parallel A'\hat{r} \parallel_2^2 + \epsilon^2 \parallel AA'\hat{r} \parallel_2^2 .$$

Puisque $A'\hat{r} \neq 0$, pour des valeurs de $\epsilon > 0$ assez petites, cette quantité est inférieure à $\hat{r}\,'\hat{r}$ d'où le résultat.
<u>*Fin de la démonstration*</u>

Propriété 2 *La matrice $A'A$ est une matrice (symétrique) définie positive de $\mathcal{M}_{m,m}(\mathbb{R})$ si et seulement si* $\mathrm{rang}(A) = m$ *et alors le problème (1) admet une solution unique donnée par :*

$$A'Ax = A'b.$$

<u>*Démonstration :*</u> Supposons que la $A \in \mathcal{M}_{n,m}(\mathbb{R})$ est de rang m, alors les vecteurs colonnes de A sont indépendantes. Dans ce cas :

$$x \neq 0 \Longrightarrow Ax \neq 0,$$

et donc :

$$x \neq 0 \Longrightarrow x'A'Ax > 0,$$

et la matrice est définie positive. Réciproquement, supposons la matrice définie postive et montrons que A est de rang m, où plutôt démontrons la contraposée, c'est à dire que si A n'est pas de rang m alors A n'est pas définie positive. On suppose donc qu'il existe x_0 tel que $Ax_0 = 0$ et il vient immédiatement que la matrice $A'A$ n'est pas définie positive car $x'A'Ax = 0$. On trouvera dans les cours des années précédentes qu'une matrice sysmétrique définie positive est inversible.
<u>*Fin de la démonstration*</u>

2 Interprétations géométriques

2.1 Moindres carrés et projections orthogonales

On considère une matrice $A \in \mathcal{M}_{n,m}(\mathbb{R})$ et un vecteur $b \in \mathbb{R}^n$. On considère le sous espace vectoriel de $\mathbb{R}^n$:

$$V = \{Ax,\ x \in \mathbb{R}^m\}.$$

On cherche en fait $\hat{y} \in V$ qui minimise la norme euclidienne $\| y - b \|$. Ce vecteur est la projection orthogonale de b sur V et il est donc caractérisé par :

$$< \hat{y} - b, y >= 0, \forall y \in V,$$

condition que l'on peut écrire en posant $\hat{y} = A\hat{x}$ et $y = Ax$:

$$< A\hat{x} - b, Ax >= 0, \forall x \in \mathbb{R}^m,$$

c'est à dire :

$$< x', A'A\hat{x} - A'b >= 0, \forall x \in \mathbb{R}^m,$$

ce qui est équivalent à : $A'A\hat{x} = A'b$.

La projection $\hat{y}$ de b est unique mais il peu exister un seul ou une infinité de vecteurs $\hat{x}$ tels que $\hat{y} = A\hat{x}$ suivant que le rang de A est m ou non.

2.2 Une propriété plus fine

Avec toujours $A \in \mathcal{M}_{n,m}(\mathbb{R})$ et $b \in \mathbb{R}^n$ on note :

$$V = \text{Im } (A) = \{z = Ax, x \in \mathbb{R}^m\},$$

et

$$V^{\perp} = \{y \in \mathbb{R}^m :< y, Ax >= 0, \forall x \in \mathbb{R}^m\}.$$

le sous espace complémentaire orthogonal de Im (a) dans $\mathbb{R}^n$. Toute solution aux moindres carrées $\hat{x}$ est telle que $b - A\hat{x} \in V^{\perp}$.

Université Joseph Fourier, Grenoble I
MAP350

Résumé de cours sur les moindres carrés : applications

Notations : Les notations utilisées dans le chapitre précédent concernant les systèmes linéaires aux moindres carrés $Ax = b$ vont nous amener à choisir d'autres notations moins classiques pour d'autres problèmes. Attention à la confusion qui peut résulter des habitudes de notations !

1 Régression linéaire

1.1 Enoncé du problème

Une des applications les plus souvent rencontrées est celle de la droite de régression que l'on trouve également dans les cours de statistique.

Nous nous plaçons dans le plan euclidien où un point M sera repéré pas ses coordonnées (u, v). On donne n points $M_i = (u_i, v_i),\ i = 1, 2, \cdots, n$ et on cherche une doite passant "approximativement" par ces points. Désignons cette droite par :

$$v = \alpha u + \beta,$$

et nous cherchons à "satisfaire au mieux" le système d'équations :

$$\left\{ \begin{array}{l} v_1 = \alpha u_1 + \beta \\ v_2 = \alpha u_2 + \beta \\ \vdots \\ v_n = \alpha u_n + \beta \end{array} \right. \tag{1}$$

Une solution raisaonnable est de considérer un système aux moindres carrés, consistant à minimiser la quantité :

$$\mid v_1 - \alpha u_1 - \beta \mid^2 + \mid v_2 - \alpha u_2 - \beta \mid^2 + \cdots + \mid v_n - \alpha u_n - \beta \mid^2 = \sum_{i=1}^{n} \mid v_i - \alpha u_i - \beta \mid^2 .$$

Les inconnues étant α et β nous écrirons donc le système :

$$\left\{ \begin{array}{l} \alpha u_1 + \beta = v_1 \\ \alpha u_2 + \beta = v_2 \\ \vdots \\ \alpha u_n + \beta = v_n \end{array} \right. \tag{2}$$

Nous écrirons ce système d'équations :

$$Ax = \begin{pmatrix} u_1 & 1 \\ u_2 & 1 \\ \vdots & \vdots \\ u_2 & 1 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = b$$

et nous pouvons alors appliquer les résultats du chapitre précédent.

1.2 Discussion et solution du problème

Tout d'abord que signifie que la matrice A soit de rang 2 ? tout simplement que les deux vecteurs colonnes soient indépendants et donc que les u_i ne sont pas tous égaux. L'interprétation géométrique en est très naturelle : il faut qu'il y ait aux moins deux abscisses de mesures différentes.

Ensuite, la solution théorique de ce système est donnée par :

$$A'Ax = A'b,$$

et le calcul montre qu'il s'agit du système 2×2 :

$$\begin{pmatrix} u_1 & u_2 & \dots & u_n \\ 1 & 1 & \dots & 1 \end{pmatrix} \begin{pmatrix} u_1 & 1 \\ u_2 & 1 \\ \vdots & \vdots \\ u_2 & 1 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} u_1 & u_2 & \dots & u_n \\ 1 & 1 & \dots & 1 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$$

ou :

$$\begin{pmatrix} \sum_{i=1}^n u_i^2 & \sum_{i=1}^n u_i \\ \sum_{i=1}^n u_i & n \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^n u_i v_i \\ \sum_{i=1}^n v_i \end{pmatrix}$$

La solution de ce système est donnée par :

$$\alpha = \frac{n \sum_{i=1}^n u_i v_i - \sum_{i=1}^n u_i \sum_{i=1}^n v_i}{n \sum_{i=1}^n u_i^2 - (\sum_{i=1}^n u_i)^2}$$

$$\beta = \frac{\sum_{i=1}^n u_i^2 \sum_{i=1}^n v_i - \sum_{i=1}^n u_i v_i \sum_{i=1}^n u_i}{n \sum_{i=1}^n u_i^2 - (\sum_{i=1}^n u_i)^2}$$

On trouvera d'autres formes du résultat, en particulier en faisant intervenir des quantités reliées aux statistiques. Par exemple pour le dénominateur on vérifiera :

$$n \sum_{i=1}^n u_i^2 - (\sum_{i=1}^n u_i)^2 = n \sum_{i=1}^n (u_i - \overline{u})^2,$$

avec :

$$\overline{u} = \frac{1}{n} \sum_{i=1}^n u_i.$$

2 Approximation aux moindres carrés dans une famille linéaire

2.1 Le cas général

On se donne toujours des points $M_i = (u_i, v_i),\ i = 1, 2, \cdots, n$ du plan euclidien.

On considère maintenant des fonctions $f_1, f_2, \cdots, f_m \in \mathbb{R}^{\mathbb{R}}$, linéairement indépendantes.

Nous souhaitons trouver parmi les fonctions de l'espace vectoriel V engendré par $f_1, f_2, \cdots, f_m \in \mathbb{R}^{\mathbb{R}}$, celle qui passe le mieux par les points M_i. Une des modélisations de ce problème, consiste à trouver une fonction f de l'espace vectoriel V minimisant la quantité :

$$\sum_{i=1}^{n} \mid f(u_i) - v_i \mid^2,$$

ce qui revient, en exprimant la fonction f sur la base des f_i à résoudre le système aux moindres carrés :

$$\left\{ \begin{array}{l} \sum_{k=1}^{m} \alpha_k f_k(u_1) = v_1 \\ \sum_{k=1}^{m} \alpha_k f_k(u_2) = v_2 \\ \cdots \\ \sum_{k=1}^{m} \alpha_k f_k(u_n) = v_n \end{array} \right. \tag{3}$$

La formulation matricielle de ce système est :

$$\begin{pmatrix} f_1(u_1) & f_2(u_1) & \cdots & f_m(u_1) \\ f_1(u_2) & f_2(u_2) & \cdots & f_m(u_2) \\ \vdots & \vdots & \vdots & \vdots \\ f_1(u_n) & f_2(u_n) & \cdots & f_m(u_n) \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_m \end{pmatrix} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}.$$

Nous pouvons donc lui appliquer la formulation d'un système aux moindres carrés $Ax = b$ avec :

$$A = \begin{pmatrix} f_1(u_1) & f_2(u_1) & \cdots & f_m(u_1) \\ f_1(u_2) & f_2(u_2) & \cdots & f_m(u_2) \\ \vdots & \vdots & \vdots & \vdots \\ f_1(u_n) & f_2(u_n) & \cdots & f_m(u_n) \end{pmatrix} \quad , \quad b = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$$

et les inconnues :

$$x = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_m \end{pmatrix}.$$

Nous ne ferons pas de discussion générale sur le problème. Nous examinerons simplement la méthode sur des cas particuliers.

2.2 Comment ajuster une parabole ?

Où comment faire passer une parabole par les points $M_i = (u_i, v_i),\ i = 1, 2, \cdots, n$?

Il s'agit du problème précédent en prenant les fonctions :

$$f_1(u) = 1,\ f_2(u) = u,\ f_3(u) = u^2,$$

que nous pouvons renuméroter de façon naturelle :

$$f_0(u) = 1,\ f_1(u) = u,\ f_2(u) = u^2,$$

De la même façon nous renumérotons les inconnues $\alpha_0,\ \alpha_1,\ \alpha_2$ et voulons minimiser la quantité :

$$\sum_{i=1}^{n} \mid \alpha_0 + \alpha_1 u_i + \alpha_2 u_i^2 - v_i \mid^2 .$$

La matrice A du système est :

$$A = \begin{pmatrix} 1 & u_1 & u_1^2 \\ 1 & u_2 & u_2^2 \\ \vdots & \vdots & \vdots \\ 1 & u_n & u_n^2 \end{pmatrix}$$

et le système $A'Ax = A'b$ est en fait :

$$\begin{pmatrix} 1 & 1 & \cdots & 1 \\ u_1 & u_2 & \cdots & u_n \\ u_1^2 & u_2^2 & \cdots & u_n^2 \end{pmatrix} \begin{pmatrix} 1 & u_1 & u_1^2 \\ 1 & u_2 & u_2^2 \\ \vdots & \vdots & \vdots \\ 1 & u_n & u_n^2 \end{pmatrix} = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ u_1 & u_2 & \cdots & u_n \\ u_1^2 & u_2^2 & \cdots & u_n^2 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$$

donc le système 3×3 :

$$\begin{pmatrix} n & \sum_{i=1}^n u_i & \sum_{i=1}^n u_i^2 \\ \sum_{i=1}^n u_i & \sum_{i=1}^n u_i^2 & \sum_{i=1}^n u_i^3 \\ \sum_{i=1}^n u_i^2 & \sum_{i=1}^n u_i^3 & \sum_{i=1}^n u_i^4 \end{pmatrix} \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^n v_i \\ \sum_{i=1}^n u_i v_i \\ \sum_{i=1}^n u_i^2 v_i \end{pmatrix}.$$