Analyse numérique, Matmeca 1ere année
Corrigé de la feuille 4[1]

1. Interpolation de Lagrange

1.a Notons ϕ l'application définie par

$$\phi \,:\, \mathrm{P} \in \mathcal{P}_n \longmapsto \Big(\mathrm{P}(x_0), \mathrm{P}(x_1), \ldots, \mathrm{P}(x_n)\Big) \in \mathbb{R}^{n+1}$$

ϕ est clairement un homomorphisme d'espaces vectoriels. Montrons qu'il est injectif. Pour cela, on étudie son noyau. Soit $\mathrm{P} \in \mathrm{Ker}\ \phi$. On a alors $\deg \mathrm{P} = n$, et pour tout $i \in [\![\, 0\,;\, n\,]\!]$, $\mathrm{P}(x_i) = 0$. On a donc au moins $n+1$ racines distinctes pour P, qui est de degré n. On en déduit que P est le polynôme nul, donc que ϕ est injective.

Comme de plus, on a $\dim \mathbb{R}^{n+1} = \dim \mathcal{P}_n$, on en déduit que ϕ est un isomorphisme d'espaces vectoriels. On sait alors que l'image d'une base par ϕ ou par ϕ^{-1} est une base. Comme $\{\mathrm{L}_i\}_{i\in[\![\,0\,;\,n\,]\!]}$ apparaît comme l'image de la base canonique de $\mathbb{R}^{n+1}$ par ϕ^{-1}, il s'agit d'une base de $\mathcal{P}_n$.

Afin de déterminer les L_i, on remarque que l'on en connaît déjà $n-1$ racines : les x_j, pour $j \neq i$. Le polynôme L_i peut donc s'écrire sous la forme

$$\mathrm{L}_i(\mathrm{X}) = \mathrm{A}_i(\mathrm{X}) \prod_{j=0, j\neq i}^{n} (\mathrm{X} - x_j)$$

Et on a alors

$$\deg(\mathrm{L}_i) = \deg(\mathrm{A}_i) + n$$

or on sait que $\deg \mathrm{L}_i \leqslant n$, et on en déduit que $\deg \mathrm{A}_i = 0$, c'est à dire que A_i est constant. Pour déterminer cette constante, on utilise la valeur de L_i en x_i. On a

$$\mathrm{L}_i(x_i) = 1 = \mathrm{A}_i \prod_{j=0, j\neq i}^{n} (x_i - x_j)$$

d'où

$$\mathrm{A}_i = \frac{1}{\displaystyle\prod_{j=0, j\neq i}^{n} (x_i - x_j)}$$

Finalement, on trouve

$$\mathrm{L}_i(\mathrm{X}) = \prod_{j=0, j\neq i}^{n} \frac{\mathrm{X} - x_j}{x_i - x_j}$$

1.b On note p_n le polynôme d'interpolation de f en les points x_i, c'est à dire le polynôme de $\mathcal{P}_n$ tel que pour tout $i \in [\![\, 0\,;\, n\,]\!]$ on ait $p_n(x_i) = f(x_i)$. Comme $p_n \in \mathcal{P}_n$ et que les $\{\mathrm{L}_i\}_{i\in[\![\,0\,;\,n\,]\!]}$ forment une base de $\mathcal{P}_n$, il existe $\lambda_0, \ldots, \lambda_n$ tels que

$$p_n = \sum_{i=0}^{n} \lambda_i \mathrm{L}_i$$

[1] généré avec LaTeX 2_ε. Tous les commentaires, compléments, insultes et remarques désobligeantes sont les bienvenus à `perrier@math.u-bordeaux1.fr`

En évaluant cette expression en x_i, on trouve

$$p_n(x_i) = \lambda_i \mathrm{L}_i(x_i) = \lambda_i$$

on a donc

$$p_n = \sum_{i=0}^{n} f(x_i)\,\mathrm{L}_i$$

Montrons maintenant l'unicité d'un tel polynôme. Supposons qu'il existe p_n et $\bar{p}_n$ tels que

$$\forall i \in [\![\, 0\,;\, n \,]\!] \qquad p_n(x_i) = \bar{p}_n(x_i) = f(x_i)$$

On en déduit que

$$\forall i \in [\![\, 0\,;\, n \,]\!] \qquad (p_n - \bar{p}_n)(x_i) = 0$$

Le polynôme $(p_n - \bar{p}_n)$ admet $n+1$ racines, et comme ce polynôme est de degré au plus n, il est nul. On a donc $p_n = \bar{p}_n$, d'où l'unicité de p_n.

1.c Pour cette question, on note

$$(x_0, x_1, x_2, x_3, x_4) = \left(-1, -\frac{1}{2}, 0, \frac{1}{2}, 1\right)$$

Le polynôme que l'on cherche est alors égal à

$$\begin{aligned}\mathrm{P} &= f(x_0)\mathrm{L}_0 + f(x_1)\mathrm{L}_1 + f(x_2)\mathrm{L}_2 + f(x_3)\mathrm{L}_3 + f(x_4)\mathrm{L}_4 \\ &= -\frac{3}{2}\,\mathrm{L}_0 + \frac{1}{4}\,\mathrm{L}_2\end{aligned}$$

Il suffit donc de calculer L_0 et L_2.

$$\mathrm{L}_0(\mathrm{X}) = \frac{\mathrm{X}\left(\mathrm{X} - \frac{1}{2}\right)\left(\mathrm{X} + \frac{1}{2}\right)\left(\mathrm{X} - 1\right)}{\left(-1\right)\left(-\frac{3}{2}\right)\left(-\frac{1}{2}\right)\left(-2\right)} = \frac{2\mathrm{X}}{3}\left(\mathrm{X} - \frac{1}{2}\right)\left(\mathrm{X} + \frac{1}{2}\right)\left(\mathrm{X} - 1\right)$$

$$\mathrm{L}_2(\mathrm{X}) = \frac{\left(\mathrm{X} + 1\right)\left(\mathrm{X} + \frac{1}{2}\right)\left(\mathrm{X} - \frac{1}{2}\right)\left(\mathrm{X} - 1\right)}{\left(-1\right)\left(-\frac{1}{2}\right)\left(1\right)\left(\frac{1}{2}\right)} = (\mathrm{X}+1)(2\mathrm{X}+1)(2\mathrm{X}-1)(\mathrm{X}-1)$$

On trouve finalement

$$\begin{aligned}\mathrm{P}(\mathrm{X}) &= -\frac{3}{2}\,\mathrm{L}_0(\mathrm{X}) + \frac{1}{4}\,\mathrm{L}_2(\mathrm{X}) \\ &= -\mathrm{X}\left(\mathrm{X} - \frac{1}{2}\right)\left(\mathrm{X} + \frac{1}{2}\right)\left(\mathrm{X} - 1\right) + \left(\mathrm{X} + 1\right)\left(\mathrm{X} - \frac{1}{2}\right)\left(\mathrm{X} + \frac{1}{2}\right)\left(\mathrm{X} - 1\right) \\ &= \left(\mathrm{X} - 1\right)\left(\mathrm{X} - \frac{1}{2}\right)\left(\mathrm{X} + \frac{1}{2}\right)\end{aligned}$$

1.d On cherche à montrer que

$$\forall x \in [\, a\,;b\,] \quad \exists \xi \in \,]\, a\,;b\,[\qquad f(x) - p_n(x) = \frac{\prod_{i=0}^{n}(x - x_j)}{(n+1)!} f^{(n+1)}(\xi)$$

Notons que cette égalité est évidente si $f \in \mathcal{P}_n$, ou alors si x est égal à l'un x_i.

Dans les autres cas, considérons la fonction

$$\varphi \,:\, t \in \mathbb{R} \longmapsto f(t) - p_n(t) - \mathrm{K}_x \prod_{i=0}^{n} (x - x_j)$$

où K_x est choisi tel que $\varphi(x) = 0$ (c'est à dire tel que l'égalité que l'on cherche à montrer soit vérifiée). On va chercher à exprimer K_x en fonction d'un $f^{(n+1)}(\xi)$.

La fonction φ admet au moins $n+2$ zéros distincts (les x_i, et x). En appliquant le théorème de Rolle sur les intervalles successifs définis par les zéros de φ, on montre que φ' admet au moins $n+1$ zéros distincts. En appliquant à nouveau le théorème de Rolle sur les intervalles définis par les zéros de φ', on montre que φ'' admet au moins n zéros. En continuant de cette manière, on montre que $\varphi^{(n+1)}$ admet au moins un zéro :

$$\exists \xi \qquad \varphi^{(n+1)}(\xi) = 0$$

Par ailleurs, lorsqu'on dérive $n+1$ fois φ :

- La dérivée $(n+1)$ème de p_n est nulle, car p_n est de degré inférieur à n.
- La dérivée $(n+1)$ème de $\mathrm{K}_x \prod_{i=0}^{n} (x - x_j)$ est égale à $(n+1)!\mathrm{K}_x$.

on a donc

$$\varphi^{(n+1)}(t) = f^{(n+1)}(t) - \mathrm{K}_x(n+1)!$$

d'où

$$\varphi^{(n+1)}(\xi) = 0 = f^{(n+1)}(\xi) - \mathrm{K}_x(n+1)!$$

soit

$$\mathrm{K}_x = \frac{f^{(n+1)}(\xi)}{(n+1)!}$$

Enfin, en évaluant φ en x, on trouve

$$\varphi(x) = 0 = f(x) - p_n(x) - \frac{\prod_{j=0}^{n} (x - x_j)}{(n+1)!} f^{(n+1)}(\xi)$$

ce qui donne

$$\forall x \in [\,a\,;b\,] \quad \exists \xi \in \,]\,a\,;b\,[\qquad f(x) - p_n(x) = \frac{\prod_{j=0}^{n} (x - x_j)}{(n+1)!} f^{(n+1)}(\xi)$$

$\boxed{\textbf{1.e}}$ D'après l'égalité trouvée à la question précédente, on a

$$\forall x \in [\,a\,;b\,] \qquad |f(x) - p_n(x)| = \frac{\prod_{j=0}^{n} |x - x_j|}{(n+1)!} \left| f^{(n+1)}(\xi) \right|$$

d'où

$$\forall x \in [\,a\,;b\,] \qquad |f(x) - p_n(x)| \leqslant \frac{\|f^{(n+1)}\|_\infty}{(n+1)!} \prod_{i=0}^{n} |b - a|$$

soit

$$\forall x \in [\,a\,;b\,] \qquad |f(x) - p_n(x)| \leqslant \frac{\|f^{(n+1)}\|_\infty (b-a)^{n+1}}{(n+1)!}$$

Finalement

$$\|f - p_n\|_\infty \leqslant \frac{\|f^{(n+1)}\|_\infty (b-a)^{n+1}}{(n+1)!}$$

- **Premier exemple** Si $f : x \mapsto \cos(x)$, alors on peut majorer $\|f^{(n+1)}\|_\infty$ par 1. On a donc
$$\|f - p_n\|_\infty \leqslant \frac{1}{(n+1)!}$$
Pour avoir $\|f - p_n\|_\infty \leqslant \mathrm{E}$, il suffit donc d'avoir $1/(n+1)! \leqslant \mathrm{E}$.
  - si on se donne $\mathrm{E} = 0,1$, alors $n = 3$ convient (c'est à dire que 4 points suffisent).
  - si on se donne $\mathrm{E} = 0,01$, alors $n = 4$ convient (c'est à dire que 5 points suffisent).
  - si on se donne $\mathrm{E} = 0,001$, alors $n = 6$ convient (c'est à dire que 7 points suffisent).
- **Second exemple** On s'interesse à présent à $g : x \mapsto \mathrm{e}^{3x}$. On a alors
$$\forall x \in \mathbb{R} \qquad g^{(n+1)}(x) = 3^{n+1}\mathrm{e}^{3x}$$
On en déduit que si on se place sur $[\,0\,;1\,]$, $\|g^{(n+1)}\|_\infty = 3^{n+1}\mathrm{e}$. Il suffit donc d'avoir $\dfrac{3^{n+1}\mathrm{e}}{(n+1)!} \leqslant \mathrm{E}$ si on veut être sûr que $\|g - p_n\|_\infty \leqslant \mathrm{E}$.
  - si on se donne $\mathrm{E} = 0,1$ alors $n = 9$ convient (c'est à dire que 10 points suffisent).
  - si on se donne $\mathrm{E} = 0,01$ alors $n = 11$ convient (c'est à dire que 12 points suffisent).
  - si on se donne $\mathrm{E} = 0,001$ alors $n = 12$ convient (c'est à dire que 13 points suffisent).

2. Interpolation d'Hermite

2.a Soit ϕ le morphisme d'évaluation :

$$\phi \; : \; p \in \mathcal{P}_3 \longmapsto (p(x_1), p'(x_1), p(x_2), p'(x_2)) \in \mathbb{R}^4$$

L'application ϕ est linéaire. On étudie son noyau :

$$\text{Ker } \phi = \Big\{ p \in \mathcal{P}_3 \qquad p(x_1) = p'(x_1) = p(x_2) = p'(x_2) = 0 \Big\}$$

Soit $p \in \text{Ker } \phi$. Alors x_1 et x_2 sont racines doubles de p. Donc il existe K tel que

$$p(\mathrm{X}) = \mathrm{K}(\mathrm{X} - x_1)^2 \, (\mathrm{X} - x_2)^2$$

comme p est de degré inférieur à 3, on en déduit que p est nul. Le morphisme ϕ est donc injectif. De plus, $\dim \mathcal{P}_3 = \dim \mathbb{R}^4$, donc d'après le théorème du rang, ϕ est un isomorphisme.

Si on se donne une fonction $f \in \mathscr{C}^1([\, a\, ;b\,])$, alors le polynôme

$$\phi^{-1} \left(f(x_1), f'(x_1), f(x_2), f'(x_2) \right)$$

est tel que

$$\left\{ \begin{aligned} p(x_1) &= f(x_1) \\ p'(x_1) &= f'(x_1) \\ p(x_2) &= f(x_2) \\ p'(x_2) &= f'(x_2) \end{aligned} \right.$$

d'où l'existence du polynôme d'interpolation d'Hermite.

Montrons à présent l'unicité d'un tel polynôme. Soient p et $\bar{p}$ deux polynômes tels que

$$\left\{ \begin{aligned} p(x_1) &= f(x_1) \\ p'(x_1) &= f'(x_1) \\ p(x_2) &= f(x_2) \\ p'(x_2) &= f'(x_2) \end{aligned} \right. \qquad \text{et} \qquad \left\{ \begin{aligned} \bar{p}(x_1) &= f(x_1) \\ \bar{p}'(x_1) &= f'(x_1) \\ \bar{p}(x_2) &= f(x_2) \\ \bar{p}'(x_2) &= f'(x_2) \end{aligned} \right.$$

On en déduit que $\phi(p) = \phi(\bar{p})$, donc comme ϕ est injective, $p = \bar{p}$, d'où l'unicité d'un tel polynôme.

2.b Soit $x \in [\, a\, ; b\,]$. Considérons la fonction

$$\varphi \; : \; t \in [\, a\, ; b\,] \longmapsto f(t) - \mathrm{K}_x (t - x_1)^2 (t - x_2)^2 \in \mathbb{R}$$

où K_x est choisi de telle sorte que $\varphi(x) = 0$. La fonction φ admet trois zéros distincts : x, x_1, x_2. Sur les intervalles définis par ces zéros de φ, le théorème de Rolle assure l'existence de deux zéros à φ'. De plus, on avait déjà deux zéros pour φ' : x_1, x_2. On en déduit que φ' admet au moins quatre zéros. En appliquant à nouveau le théorème de Rolle sur les intervalles définis par les zéros de φ', on voit que φ'' admet au moins trois zéros. En procédant de même, on montre que $\varphi^{(3)}$ admet au moins deux zéros, puis que $\varphi^{(4)}$ admet au moins un zéro, que l'on note ξ.

Par ailleurs :

$$\varphi^{(4)}(t) = f^{(4)}(t) - 4!\mathrm{K}_x$$

Comme $\varphi^{(4)}(\xi) = 0$,

$$\mathrm{K}_x = \frac{f^{(4)}(\xi)}{4!}$$

On a finalement montré que

$$\forall x \in [\,a\,;b\,] \quad \exists \xi \in [\,a\,;b\,] \qquad f(x) - p_n(x) = \frac{f^{(4)}(\xi)}{4!}\,(x-x_1)^2(x-x_2)^2$$

2.c On voit immédiatement que les polynômes suivants conviennent :

$$\begin{cases} \mathrm{A}_1 = \phi^{-1}\Big((1,0,0,0)\Big) \\ \mathrm{B}_1 = \phi^{-1}\Big((0,1,0,0)\Big) \\ \mathrm{A}_2 = \phi^{-1}\Big((0,0,1,0)\Big) \\ \mathrm{B}_2 = \phi^{-1}\Big((0,0,0,1)\Big) \end{cases}$$

où ϕ est le morphisme défini à la question (a).

On remarque que A_1 et A_2 sont les mêmes en changeant les rôles de x_1 et x_2. De même pour B_1 et B_2. On ne va donc calculer que A_1 et B_1.

- **Calcul de** A_1

  On sait que $\mathrm{A}_1(x_2) = \mathrm{A}_1'(x_2) = 0$. Cela signifie que A_1 est divisible par $(\mathrm{X}-x_2)^2$. Soient L_1 et L_2 les polynômes d'interpolation de Lagrange associés à $\{x_1, x_2\}$. Comme on l'a vu à l'exercice précédent, on a

$$\mathrm{L}_1(\mathrm{X}) = \frac{\mathrm{X}-x_2}{x_1-x_2} \qquad \text{et} \qquad \mathrm{L}_2(\mathrm{X}) = \frac{\mathrm{X}-x_1}{x_2-x_1}$$

  Comme A_1 est divisible par $(\mathrm{X}-x_2)^2$, et que L_1^2 est proportionnel à ce polynôme, A_1 est également divisible par L_1^2 :

$$\exists \mathrm{Q}_1 \in \mathcal{P}_1 \qquad \mathrm{A}_1 = \mathrm{Q}_1\mathrm{L}_1^2$$

  On va utiliser les autres conditions que l'on a sur A_1 pour déterminer Q_1 :

$$\begin{cases} \mathrm{A}_1(x_1) = 1 = \mathrm{Q}_1(x_1)\mathrm{L}_1^2(x_1) = \mathrm{Q}_1(x_1) \\ \mathrm{A}_1'(x_1) = 0 = \mathrm{Q}_1'(x_1)\mathrm{L}_1^2(x_1) + 2\mathrm{Q}_1(x_1)\mathrm{L}_1(x_1)\mathrm{L}_1'(x_1) \\ \qquad\qquad\;\; = \mathrm{Q}_1'(x_1) + 2\mathrm{Q}_1(x_1)\mathrm{L}_1'(x_1) \end{cases}$$

  On en déduit que

$$\begin{cases} \mathrm{Q}_1'(x_1) = -2\mathrm{L}_1'(x_1) \\ \mathrm{Q}_1(x_1) = 1 \end{cases}$$

  d'où

$$\mathrm{A}_1(\mathrm{X}) = (1 - 2\mathrm{L}_1'(x_1)(\mathrm{X}-x_1))\mathrm{L}_1^2(\mathrm{X})$$

- **Calcul de** B_1

  De même que pour le calcul de A_1, on a

$$\exists \mathrm{R}_1 \in \mathcal{P}_1 \qquad \mathrm{B}_1 = \mathrm{R}_1\mathrm{L}_1^2$$

  Les autres conditions sur B_1 s'expriment sous la forme

$$\begin{cases} \mathrm{B}_1(x_1) = 0 = \mathrm{R}_1(x_1) \\ \mathrm{B}_1'(x_1) = 1 = \mathrm{R}_1'(x_1) + 2\mathrm{R}_1(x_1)\mathrm{L}_1'(x_1) \end{cases}$$

On en déduit que

$$\begin{cases} R_1'(x_1) = 1 \\ R_1(x_1) = 0 \end{cases}$$

d'où

$$B_1(X) = (X - x_1)L_1^2(X)$$

2.d D'une manière générale, on peut tenir le même raisonnement qu'à la question précédente pour trouver

$$\begin{cases} A_k(X) = (1 - 2L_k'(x_k)(X - x_k))L_k^2(X) \\ B_k(X) = (X - x_k)L_k^2(X) \end{cases}$$

où L_k est le $k$ème polynôme d'interpolation de Lagrange associé aux points où l'on interpole.

Remarque
Étant donné une fonction f suffisamment régulière, et n points, on peut interpoler à n'importe quel ordre en chacun de ces points, l'ordre n'étant pas forcément le même pour chaque point.

On se donne n points $x_1, x_2, \ldots, x_n$, et n entiers $\alpha_1, \alpha_2, \ldots \alpha_n$. On note $M = \sum_{i=1}^{n} (\alpha_i + 1)$. Alors pour tout $f \in \mathscr{C}^{\max \alpha_i}$, il existe un unique polynôme P de $\mathcal{P}_{M-1}$ tel que

$$\forall k \in [\![1\,;\, n]\!] \quad \forall 0 \leqslant j \leqslant \alpha_k \qquad f^{(j)}(x_k) = P^{(j)}(x_k)$$

Si f est de classe $\mathscr{C}^M$, alors

$$\forall x \in [\,a\,;b\,] \quad \exists \xi \in [\,a\,;b\,] \qquad f(x) - P(x) = \frac{\prod_{k=1}^{n} (X - x_k)^{\alpha_k}}{M!} f^{(M)}(\xi)$$

Attention cependant, si on désire interpoler à l'ordre N en un point x_0, c'est à dire imposer

$$f^{(N)}(x_0) = P^{(N)}(x_0)$$

on doit <u>obligatoirement</u> prescrire les valeurs en ce point à <u>tout</u> ordre inférieur à N, c'est à dire que l'on doit également avoir

$$\forall i \in [\![0\,;\, N]\!] \qquad f^{(i)}(x_0) = P^{(i)}(x_0)$$

afin

1. de garantir l'unicité du polynôme d'interpolation.
2. d'utiliser la méthode vue dans les cas des interpolations d'Hermite et de Lagrange pour déterminer une formule d'erreur.

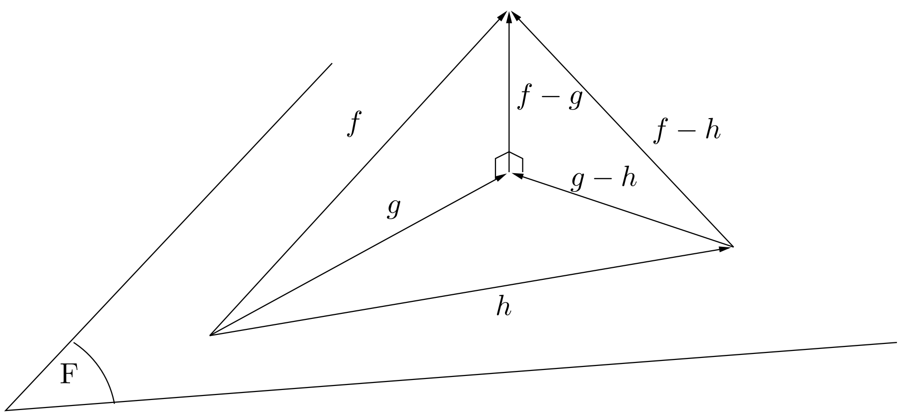

FIG. 1 – Pour la première inégalité, on écrit le théorème de Pythagore dans le triangle défini par les vecteurs $f-g$, $g-h$, $f-h$.

3. PROJECTION HILBERTIENNE SUR UN ESPACE DE DIMENSION FINIE / MÉTHODE DE GALERKIN

3.a On suppose que g est un vecteur de F tel que

$$\forall h \in \mathrm{F} \qquad (f-g, h) = 0$$

Alors on peut écrire, pour tout $h \in \mathrm{F}$ (voir Figure 1) :

$$\begin{aligned}
\|f-h\|_2^2 &= \|f-g+g-h\|_2^2 \\
&= \|f-g\|_2^2 + 2(f-g, g-h) + \|g-h\|_2^2 \\
&= \|f-g\|_2^2 + \|g-h\|_2^2 \\
\|f-h\|_2^2 &\geqslant \|f-g\|_2^2
\end{aligned}$$

On a en effet $(f-g, g-h) = 0$, car $g-h \in \mathrm{F}$, donc $g-h$ est orthogonal à $f-g$, par hypothèse.

Réciproquement, supposons qu'il existe g tel que pour tout $h \in \mathrm{F}$, on ait $\|f-g\|_2 \leqslant \|f-h\|_2$.

Alors pour tout v, et pout tout $t \in \mathbb{R}$, $g+tv$ est « plus éloigné » de f que g ne l'est (voir Figure 2). On a ainsi

$$\|f-(g+tv)\|_2 \geqslant \|f-g\|_2$$

on commence par développer le premier membre :

$$\begin{aligned}
\|f-(g+tv)\|_2^2 &= \|f-g-tv\|_2^2 \\
&= \|f-g\|_2^2 - 2t(f-g, v) + t^2\|v\|_2^2
\end{aligned}$$

on en déduit que pour tout $t \in \mathbb{R}$ et pour tout $v \in \mathrm{F}$

$$-2t(f-g, v) + t^2\|v\|_2^2 \geqslant 0$$

Si $t > 0$, on peut simplifier par t l'expression trouvée, puis faire tendre t vers 0, pour trouver $(f-g, v) \leqslant 0$. Si t est négatif, l'inégalité trouvée change de sens lorsqu'on

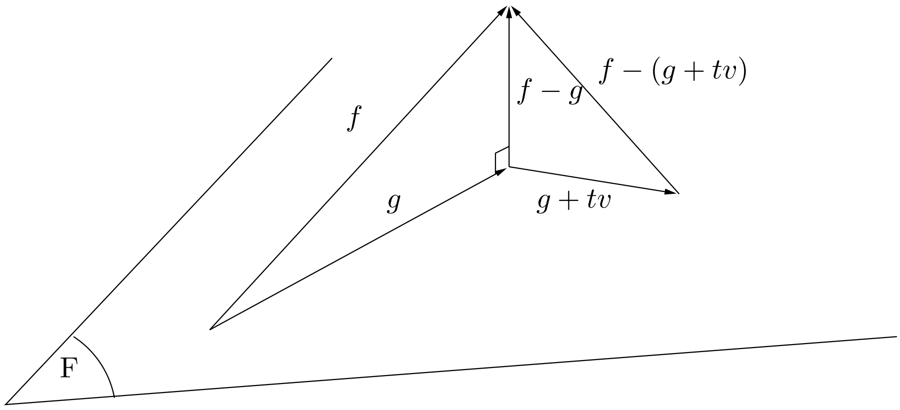

Fig. 2 – Pour la seconde inégalité, on écrit le théorème de Pythagore dans le triangle défini par les vecteurs $f - (g + tv)$, $g + tv$, $f - g$.

la simplifie par t. On trouve alors, en faisant tendre t vers 0 : $(f - g, v) \geqslant 0$. On en déduit finalement que

$$(f - g, v) = 0$$

3.b Supposons qu'il existe deux vecteurs g_1 et g_2 tels que g_1 et g_2 minimisent $\|f - h\|_2$ pour $h \in \mathrm{F}$. Alors, en appliquant le théorème de Pythagore, on a

$$\|f - g_1\|_2^2 = \|f - g_2\|_2^2 + \|g_1 - g_2\|_2^2$$

et de même

$$\|f - g_2\|_2^2 = \|f - g_1\|_2^2 + \|g_2 - g_1\|_2^2$$

en injectant la seconde égalité dans la première, on trouve

$$\|f - g_1\|_2^2 = \|f - g_1\|_2^2 + 2\|g_1 - g_2\|_2^2$$

d'où $\|g_1 - g_2\|_2^2 = 0$, soit $g_1 = g_2$.

3.c Montrons dans un premier temps que la famille des fonctions φ^i est libre. Soient $\lambda_0 \dots \lambda_{n+1}$ tels que

$$\sum_{i=0}^{n} \lambda_i \varphi^i = 0$$

En évaluant cette expression en un x_j, on trouve $\lambda_j = 0$. On en déduit que la combinaison linéaire est triviale. La famille des φ^i est donc libre. Voir la Figure 3 pour un tracé de la fonction φ^i.

Montrons à présent qu'elle est génératrice. Pour cela, il suffit de montrer que l'espace V_h est de dimension $n + 1$. Notons Φ l'homomorphisme de $\mathcal{P}_1^n$ dans $\mathbb{R}^{n-1}$ défini par

$$\Phi \,:\, (\mathrm{P}_1, \mathrm{P}_2, \dots, \mathrm{P}_n) \longmapsto \Big((\mathrm{P}_2 - \mathrm{P}_1)(x_1), (\mathrm{P}_3 - \mathrm{P}_2)(x_2), \dots, (\mathrm{P}_n - \mathrm{P}_{n-1})(x_{n-1}) \Big)$$

Montrons que ce morphisme est surjectif. Notons $(y_1, \dots, y_{n-1}) \in \mathbb{R}^{n-1}$. On choisit

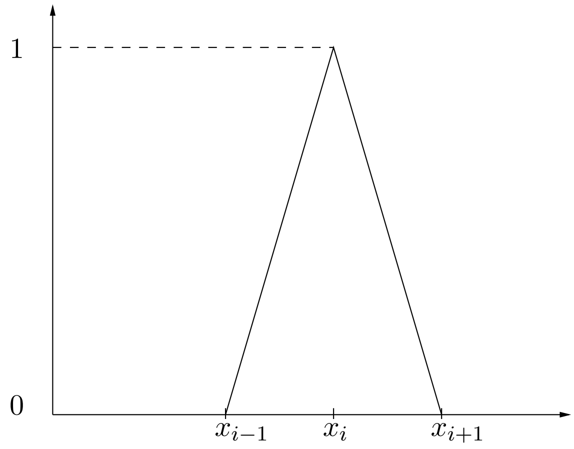

FIG. 3 – La fonction de base φ^i. On sait qu'elle vaut 0 en x_j pour $j \neq i$, et 1 en x_i. Il n'existe qu'une seule manière de joindre les points $(x_{i-1}, 0)$ et $(x_i, 1)$ par une droite, ainsi que pour les points $(x_i, 1)$ et $(x_{i+1}, 0)$.

alors $\mathrm{P}_1 \ldots \mathrm{P}_n$ tels que

$$\begin{cases} \mathrm{P}_1(x_1) = 0 \\ \mathrm{P}_2(x_1) = y_1 \quad \text{et} \quad \mathrm{P}_2(x_2) = 0 \\ \ldots \\ \mathrm{P}_i(x_{i-1}) = y_i \quad \text{et} \quad \mathrm{P}_i(x_i) = 0 \\ \ldots \\ \mathrm{P}_{n-1}(x_{n-2}) = y_{n-2} \quad \text{et} \quad \mathrm{P}_{n-1}(x_{n-1}) = 0 \\ \mathrm{P}_n(x_{n-1}) = y_{n-1} \end{cases}$$

des tels P_i existent : il suffit de choisir $\mathrm{P}_i(\mathrm{X}) = y_i \dfrac{\mathrm{X} - x_i}{x_{i-1} - x_i}$, puis $\mathrm{P}_1 = 0$ et $\mathrm{P}_n = y_{n-1}$. Ainsi, Φ est surjective. On a alors, d'après le théorème du rang

$$\dim\Big((\mathcal{P}_1)^n\Big) = \dim \mathrm{V}_h + \mathrm{rg}(\Phi)$$

et comme Φ est surjective, on a $\mathrm{rg}(\Phi) = n - 1$. Finalement, on trouve que

$$\dim \mathrm{V}_h = 2n - (n-1) = n + 1$$

Les φ^i forment donc une base de V_h.

On décompose la projection de f sur V_h dans la base des φ^i :

$$\mathrm{P}_{\mathrm{V}_h}(f) = \sum_{i=0}^{n} f_i \varphi^i$$

$\mathrm{P}_{\mathrm{V}_h}(f)$ est caractérisée par le fait que $f - \mathrm{P}_{\mathrm{V}_h}(f)$ est orthogonale à V_h :

$$\forall v \in \mathrm{V}_h \qquad (f - \mathrm{P}_{\mathrm{V}_h}(f), v) = 0$$

Comme φ^i est une base de V_h, cette dernière assertion est équivalente à

$$\forall j \in [\![\, 0\, ;\, n \,]\!] \qquad (f - \mathrm{P}_{\mathrm{V}_h}(f), \varphi^j) = 0$$

d'où

$$\forall j \in [\![0\,;\,n]\!] \qquad (f,\varphi^j) - \sum_{i=0}^{n} f_i(\varphi^i,\varphi^j) = 0$$

ceci peut encore s'écrire

$$\mathrm{A}f = b$$

où le coefficient général de A est (φ^i,φ^j), f est le vecteur des f_i, et b est le vecteur dont les coefficients sont les (f,φ^i). La matrice A est symétrique, par symétrie du produit scalaire. (La matrice A est appelée *matrice de Gram* de la famille $\left\{\varphi^i\right\}_{0\leqslant i\leqslant n}$).

La matrice A est définie positive, car pour tout v, on a

$$\begin{aligned}
(\mathrm{A}v,v) &= \sum_{j=0}^{n} v_j(\mathrm{A}v)_j \\
&= \sum_{j=0}^{n} v_j \sum_{i=0}^{n} v_i(\varphi^j,\varphi^i) \\
&= \sum_{j=0}^{n}\sum_{i=0}^{n} (\varphi^j,\varphi^i)v_iv_j \\
&= (\sum_{j=0}^{n} v_j\varphi^j, \sum_{i=0}^{n} v_i\varphi^i) \\
&= \|\widetilde{v}\|_2^2 \geqslant 0
\end{aligned}$$

où la fonction $\widetilde{v}$ est définie par

$$\widetilde{v} = \sum_{i=0}^{n} v_i\varphi^i$$

Bien que ça ne soit pas nécessaire pour la suite de l'exercice, on peu calculer les coefficients de la matrice A. Rappelons que l'on a

$$a_{i,j} = \int_a^b \varphi^i(t)\,\varphi^j(t)\,\mathrm{d}t$$

On remarque que les supports de φ^i et φ^j sont disjoints sauf si $j = i$ ou $j = i-1$ ou $j = i+1$. Donc à part dans ces cas, $a_{i,j} = 0$. Il reste donc à calculer $a_{i,i}$ et $a_{i,i+1}$.

Comme φ^i s'identifie sur $[\,x_{i-1}\,;x_i\,]$ au polynôme interpolant les valeurs $(0,1)$ aux points (x_{i-1},x_i), et au polynôme interpolant les valeurs $(1,0)$ aux points (x_i,x_{i+1}), on a

$$\varphi^i(t) = \begin{cases} \dfrac{t-x_{i-1}}{x_i-x_{i-1}} & \text{si } t \in [\,x_{i-1}\,;x_i\,] \\ \dfrac{t-x_{i+1}}{x_i-x_{i+1}} & \text{si } t \in [\,x_i\,;x_{i+1}\,] \\ 0 & \text{sinon} \end{cases}$$

$$\begin{aligned}
a_{i,i} &= \int_{x_{i-1}}^{x_{i+1}} \left(\varphi^2(t)\right)^2 \mathrm{d}t \\
&= \int_{x_{i-1}}^{x_i} \left(\varphi^2(t)\right)^2 \mathrm{d}t + \int_{x_i}^{x_{i+1}} \left(\varphi^2(t)\right)^2 \mathrm{d}t \\
&= \int_{x_{i-1}}^{x_i} \frac{(t-x_{i-1})^2}{(x_i-x_{i-1})^2}\,\mathrm{d}t + \int_{x_i}^{x_{i+1}} \frac{(t-x_{i+1})^2}{(x_i-x_{i+1})^2}\,\mathrm{d}t \\
&= \left[\frac{(t-x_{i-1})^3}{3(x_i-x_{i-1})^2}\right]_{x_{i-1}}^{x_i} + \left[\frac{(t-x_{i+1})^3}{3(x_{i+1}-x_i)^2}\right]_{x_i}^{x_{i+1}} \\
a_{i,i} &= \frac{x_{i+1}-x_{i-1}}{3}
\end{aligned}$$

On calcule à présent $a_{i,i+1}$.

$$\begin{aligned}
a_{i,i+1} &= \langle \varphi^i, \varphi^{i+1} \rangle \\
&= \int_{x_i}^{x_{i+1}} \frac{(t-x_i)(t-x_{i+1})}{(x_i-x_{i+1})(x_{i+1}-x_i)}\,\mathrm{d}t
\end{aligned}$$

On va donc calculer $\displaystyle\int_{x_i}^{x_{i+1}} (t-x_i)(t-x_{i+1})\,\mathrm{d}t$

$$\begin{aligned}
\int_{x_i}^{x_{i+1}} (t-x_i)(t-x_{i+1})\,\mathrm{d}t &= \int_{x_i}^{x_{i+1}} (t-x_{i+1})^2\,\mathrm{d}t + (x_{i+1}-x_i)\int_{x_i}^{x_{i+1}} (t-x_{i+1})\,\mathrm{d}t \\
&= \left[\frac{(t-x_{i+1})^3}{3}\right]_{x_i}^{x_{i+1}} + (x_{i+1}-x_i)\left[\frac{(t-x_{i+1})^2}{2}\right]_{x_i}^{x_{i+1}} \\
&= (x_{i+1}-x_i)^3\left(\frac{1}{3}-\frac{1}{2}\right) \\
\int_{x_i}^{x_{i+1}} (t-x_i)(t-x_{i+1})\,\mathrm{d}t &= -\frac{(x_{i+1}-x_i)^3}{6}
\end{aligned}$$

On trouve finalement

$$a_{i,i+1} = \frac{x_{i+1}-x_i}{6}$$

3.d Par comparaison de $\mathrm{P}_{\mathrm{V}_h}(f)$, qui est la meilleure approximation de f, et de $\Pi_h(f)$, qui est un élément de V_h, on a

$$\|f - \mathrm{P}_{\mathrm{V}_h}(f)\|_2 \leqslant \|f - \Pi_h(f)\|_2$$

3.e En utilisant la question précédente, on a

$$\|f - \mathrm{P}_{\mathrm{V}_h}(f)\|_2 \leqslant \|f - \Pi_h(f)\|_2$$

Afin d'utiliser les formules de majoration des erreurs d'interpolation sur chacun des intervalles $[\,x_i\,;x_{i+1}\,]$, on va commencer par découper l'intégrale sur $[\,a\,;b\,]$ en somme d'intégrales sur $[\,x_i\,;x_{i+1}\,]$:

$$\begin{aligned}
\|f - \mathrm{P}_{\mathrm{V}_h}(f)\|_2 &\leqslant \|f - \Pi_h(f)\|_2 \\
&= \sqrt{\sum_{i=0}^{n-1}\int_{x_i}^{x_{i+1}} (f-\Pi_h(f))^2(t)\,\mathrm{d}t}
\end{aligned}$$

Sur $[\,x_i\,;x_{i+1}\,]$, $\Pi_h(f)$ est une interpolation aux points x_i et x_{i+1} de f. On a donc, d'après le cours sur l'interpolation

$$\forall x \in [\,x_i\,;x_{i+1}\,] \quad \exists \xi \in [\,x_i\,;x_{i+1}\,] \qquad f(x) - \Pi_h(f)(x) = \frac{(x-x_{i+1})(x-x_i)}{2} f^{(2)}(\xi)$$

On a alors

$$\forall x \qquad (f-\Pi_h(f))^2(x) \leqslant \frac{\|f^{(2)}\|_\infty^2}{4}(x_{i+1}-x)^2(x_i-x)^2$$

d'où

$$\int_{x_i}^{x_{i+1}} (f-\Pi_h(f))^2(x)\,\mathrm{d}x \leqslant \frac{\|f^{(2)}\|_\infty^2}{4}\int_{x_i}^{x_{i+1}} (x_{i+1}-x)^2(x_i-x)^2\,\mathrm{d}x$$

Afin de calculer cette dernière intégrale, on écrit l'intégrande sous la forme

$$\begin{aligned}(x-x_{i+1})^2(x-x_i)^2 &= (x-x_i+x_i-x_{i+1})^2(x-x_i)^2\\ &= \Big((x-x_i)^2+2(x_i-x_{i+1})(x-x_i)\\ &\quad +(x_i-x_{i+1})^2\Big)\Big(x-x_i\Big)^2\\ &= (x-x_i)^4+2(x_i-x_{i+1})(x-x_i)^3+(x_i-x_{i+1})^2(x-x_i)^2\end{aligned}$$

On en déduit que

$$\begin{aligned}\int_{x_i}^{x_{i+1}}(x-x_{i+1})^2(x-x_i)^2\,\mathrm{d}x &= \int_{x_i}^{x_{i+1}}(x-x_i)^4\,\mathrm{d}x\\ &\quad +2(x_i-x_{i+1})\int_{x_i}^{x_{i+1}}(x-x_i)^3\,\mathrm{d}x\\ &\quad +(x_i-x_{i+1})^2\int_{x_i}^{x_{i+1}}(x-x_i)^2\,\mathrm{d}x\\ &= \frac{(x_{i+1}-x_i)^5}{5}-2\,\frac{(x_{i+1}-x_i)^5}{4}+\frac{(x_{i+1}-x_i)^5}{3}\\ &= \frac{(x_{i+1}-x_i)^5}{30}\end{aligned}$$

On trouve ainsi, en ramassant les petits bouts :

$$\begin{aligned}\|f-\mathrm{P}_{\mathrm{V}_h}(f)\|_2 &\leqslant \sqrt{\sum_{i=0}^{n-1}\frac{\|f^{(2)}\|_\infty^2(x_{i+1}-x_i)^5}{120}}\\ &\leqslant \sqrt{\frac{\|f^{(2)}\|_\infty^2\max\limits_{i\in[\![0\,;n-1]\!]}(x_{i+1}-x_i)^4}{120}\sum_{i=0}^{n-1}(x_{i+1}-x_i)}\\ &\leqslant \|f^{(2)}\|_\infty\max_{i\in[\![0\,;n-1]\!]}(x_{i+1}-x_i)^2\sqrt{\frac{b-a}{120}}\end{aligned}$$

4. Dérivation numérique

Au cours des précédents TDs, on a déjà vu plusieurs manières d'approcher numériquement des dérivées. Ces approximations étaient basées sur des combinaisons de plusieurs développements de Taylor.

$$f(x_0+h)=f(x_0)+hf'(x_0)+\frac{h^2}{2}f''(x_0)+\mathrm{O}\left(h^3\right)$$

On obtient alors

$$\frac{f(x_0+h)-f(x_0)}{h}=f'(x_0)+\frac{h}{2}f''(x_0)+\mathrm{O}\left(h^2\right)$$

qui est une approximation de f' à l'ordre 1 (car l'erreur est de l'ordre de h). On l'avait utilisé dans la l'exercice 3 de la feuille 4.

À plusieurs reprises, on a également utilisé une discretisation du Laplacien centrée, qui combine un développement de Taylor à droite et à gauche du point que l'on considère : on additionne

$$f(x_0+h)=f(x_0)+hf'(x_0)+\frac{h^2}{2}f''(x_0)+\mathrm{O}\left(h^3\right)$$

et

$$f(x_0 - h) = f(x_0) - hf'(x_0) + \frac{h^2}{2} f''(x_0) + \mathrm{O}\left(h^3\right)$$

pour trouver

$$f''(x_0) = \frac{f(x_0+h) + f(x_0-h) - 2f(x_0)}{h^2} + \mathrm{O}\left(h\right)$$

On peut trouver des formules d'ordre plus grand, mais la règle est que plus les formules sont d'ordre élevé, plus on utilise de points. Par exemple, en combinant différemment les développements aux points $x_0 + h$ et $x_0 - h$, on a

$$f'(x_0) = \frac{f(x_0+h) - f(x_0-h)}{2h} + \mathrm{O}\left(h^2\right)$$

Pour obtenir une approximation à l'ordre 2 de la dérivée, on peut également utiliser les points x_0, $x_0 + h$ et $x_0 + 2h$:

$$f(x_0 + 2h) = f(x_0) + 2hf'(x_0) + 2h^2 f''(x_0) + \frac{4h^3}{3} f^{(3)}(x_0) + \mathrm{O}\left(h^4\right)$$

$$f(x_0 + h) = f(x_0) + hf'(x_0) + \frac{h^2}{2} f''(x_0) + \frac{h^3}{6} f^{(3)}(x_0) + \mathrm{O}\left(h^4\right)$$

que l'on combine de manière :

- à avoir un coefficient non nul devant f'
- à avoir un coefficient nul devant f'' (pour faire monter l'ordre)

On trouve alors

$$f'(x_0) = \frac{\frac{1}{2} f(x_0+2h) - 2f(x_0+h) + \frac{3}{2} f(x_0)}{h} + \frac{h^2}{6} f^{(3)}(x_0) + \mathrm{O}\left(h^3\right)$$

qui est d'ordre 2.

D'une manière générale, on peut utiliser l'interpolation afin de déterminer des formules d'ordre élevé. Rappelons en effet que l'on a la formule suivante sur l'erreur d'interpolation :

$$\forall x \quad \exists \xi \qquad f(x) - \mathrm{P}(x) = \frac{\prod_{i=0}^{n} (x - x_i)}{(n+1)!} f^{(n+1)}(\xi)$$

Formellement, on dérive, pour trouver :

$$f'(x) - \mathrm{P}'(x) = \sum_{j=0}^{n} \frac{\prod_{i=0, i\neq j}^{n} (x - x_i)}{(n+1)!} f^{(n+1)}(\xi) + \frac{\prod_{i=0}^{n} (x - x_i)}{(n+1)!} \frac{\mathrm{d}\xi}{\mathrm{d}x} f^{(n+2)}(\xi)$$

cette dérivation est purement formelle, car on ne connait rien *a priori* sur la régularité de la fonction ξ. On évalue l'expression trouvée en l'un des points d'interpolation x_j, et on a alors

$$f'(x_j) - \mathrm{P}'(x_j) = \frac{\prod_{i=0, i\neq j}^{n} (x_j - x_i)}{(n+1)!} f^{(n+1)}(\xi)$$

On suppose alors que $x_j = x_0$, et que les autres points d'interpolation sont de la forme $x_0 + ih$ avec $i \in \mathbb{Z}$. Dans ce cas, l'erreur $f'(x_j) - \mathrm{P}'(x_j)$ est de l'ordre de h^n, et on a ainsi construit une formule de dérivation numérique d'ordre n, faisant intervenir n points.

5. Approximation au sens des moindres carrés discrets

5.a On a

$$\begin{aligned}\mathrm{J}(a) &= \sum_{i=0}^{k}\left|y_i - \sum_{j=0}^{n} a_j x_i^j\right|^2\\ &= \langle y - \mathrm{U}a | y - \mathrm{U}a\rangle\end{aligned}$$

où la matrice U est la matrice $(k+1)\times(n+1)$ dont le coefficient général est $u_{i,j} = x_i^j$

2.b Le problème auquel on s'intéresse revient à chercher le minimum, au sens de la norme 2 dans $\mathbb{R}^{k+1}$ de la distance entre le vecteur $(y_0,\dots,y_k)$ et les éléments de $\mathrm{Vect}(\mathrm{U}_1,\mathrm{U}_2,\dots,\mathrm{U}_n)$ où les U_i sont les colonnes de la matrice U. D'après ce qu'on a vu à l'exercice 3, ce minimum est atteint pour la projection orthogonale de y sur $\mathrm{Vect}(\mathrm{U}_1,\mathrm{U}_2,\dots,\mathrm{U}_n)$. Le vecteur a s'écrit $a = \sum_{i=0}^{k} a_i\mathrm{U}_i$. Le fait que a est la projection de y sur $\mathrm{Vect}(\mathrm{U}_1,\mathrm{U}_2,\dots,\mathrm{U}_n)$ se traduit par

$$\forall j\in [\![\,0\,;\,n\,]\!] \qquad \langle y - a, \mathrm{U}_j\rangle = 0$$

ou encore

$$\forall j\in [\![\,0\,;\,n\,]\!] \qquad \langle y - \sum_{i=0}^{n} a_i\mathrm{U}_i, \mathrm{U}_j\rangle = 0$$

soit

$$\forall j\in [\![\,0\,;\,n\,]\!] \qquad \sum_{i=0}^{n} a_i\langle \mathrm{U}_i, \mathrm{U}_j\rangle = \langle y, \mathrm{U}_j\rangle$$

ce qui peut encore s'écrire

$${}^t\mathrm{U}\,\mathrm{U}a = {}^t\mathrm{U}\,y$$

> Attention, il est hors de question ici de simplifier l'expression par ${}^t\mathrm{U}$ à gauche. En effet, ${}^t\mathrm{U}$ a peu de chances d'être inversible, puisque U n'est *a priori* même pas carrée.

5.c Dans le cas où $n = k$, la matrice U est carrée. De plus, son déterminant est un déterminant de Van–Der–Monde, qui est non nul car les points x_i sont supposés distincts. Les matrices U et ${}^t\mathrm{U}$ sont donc inversibles, et l'équation ${}^t\mathrm{U}\,\mathrm{U}a = {}^t\mathrm{U}\,y$ devient

$$a = \mathrm{U}^{-1}y$$

On retrouve le polynôme interpolant les points (x_i, y_i).

5.d La matrice U s'écrit, dans ce premier exemple

$$\mathrm{U} = \begin{pmatrix} 1 & -1 & 1\\ 1 & -\dfrac{1}{2} & \dfrac{1}{4}\\ 1 & 0 & 0\\ 1 & \dfrac{1}{2} & \dfrac{1}{4}\\ 1 & 1 & 1\end{pmatrix}$$

puis

$$
{}^t\mathrm{U}\,\mathrm{U} = \begin{pmatrix} 5 & 0 & \frac{5}{2} \\ 0 & \frac{5}{2} & 0 \\ \frac{5}{2} & 0 & \frac{17}{8} \end{pmatrix}
$$

Enfin, si $y = {}^t\left(-\frac{3}{2}, 0, \frac{1}{4}, 0, 0\right)$, alors

$$
{}^t\mathrm{U}\,y = \begin{pmatrix} -\frac{5}{4} \\ \frac{3}{2} \\ -\frac{3}{2} \end{pmatrix}
$$

On est finalement amené à résoudre le système

$$
\begin{pmatrix} 5 & 0 & \frac{5}{2} \\ 0 & \frac{5}{2} & 0 \\ \frac{5}{2} & 0 & \frac{17}{8} \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} -\frac{5}{4} \\ \frac{3}{2} \\ -\frac{3}{2} \end{pmatrix}
$$

La résolution de ce système donne finalement

$$
a_0 = \frac{1}{4} \qquad a_1 = \frac{3}{5} \qquad a_2 = -1
$$

soit

$$
\mathrm{P}(x) = -x^2 \frac{3}{5} x + \frac{1}{4}
$$

5.e Les matrices U et ^{t}U sont les mêmes qu'à la question précédente. La matrice ${}^t\mathrm{U}\,\mathrm{U}$ est donc inchangée. Par ailleurs, dans ce cas, on a $y = {}^t\left(1, \frac{1}{2}, 0, \frac{1}{2}, 1\right)$, donc

$$
{}^t\mathrm{U}\,y = \begin{pmatrix} 3 \\ 0 \\ \frac{9}{4} \end{pmatrix}
$$

on est amené à résoudre

$$
\begin{pmatrix} 5 & 0 & \frac{5}{2} \\ 0 & \frac{5}{2} & 0 \\ \frac{5}{2} & 0 & \frac{17}{8} \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 3 \\ 0 \\ \frac{9}{4} \end{pmatrix}
$$

On trouve alors

$$
a_0 = \frac{6}{35} \qquad a_1 = 0 \qquad a_2 = \frac{6}{7}
$$

6. Minimisation

6.a Le problème auquel on s'intéresse est le même qu'à l'exercice précédent, dans le cas où $n = 1$. En reprenant les notations de l'exercice précédent, on a

$$\mathrm{U} = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_{n-1} \\ 1 & x_n \end{pmatrix}$$

Comme on l'a vu auparavant, le minimum que l'on recherche est atteint, et le vecteur ${}^t(a, b)$ qui minimise cette quantité vérifie

$${}^t\mathrm{U}\,\mathrm{U}\begin{pmatrix} a \\ b \end{pmatrix} = {}^t\mathrm{U}\,y$$

où y est le vecteur ${}^t(y_1, \ldots, y_n)$. En explicitant les différentes matrices, on trouve

$$\begin{pmatrix} n & \sum_{i=1}^{n} x_i \\ \sum_{i=1}^{n} x_i & \sum_{i=1}^{n} x_i^2 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{n} y_i \\ \sum_{i=1}^{n} x_i y_i \end{pmatrix}$$

Afin de simplifier les expressions (et de faire comme si on savait faire des statistiques, pour crâner un peu), on va noter respectivement les esperances, variances et covariances :

$$\begin{aligned}
\mathrm{E(X)} &= \frac{1}{n}\sum_{i=1}^{n} x_i \\
\mathrm{Var(X)} &= \frac{1}{n}\sum_{i=1}^{n} (x_i - \mathrm{E(X)})^2 \\
\mathrm{Cov(X, Y)} &= \frac{1}{n}\sum_{i=1}^{n} (x_i - \mathrm{E(X)})(y_i - \mathrm{E(Y)})
\end{aligned}$$

On a alors

$$\begin{aligned}
\mathrm{Var(X)} &= \frac{1}{n}\sum_{i=1}^{n} (x_i - \mathrm{E(X)})^2 \\
&= \frac{1}{n}\sum_{i=1}^{n} \Big(x_i^2 - 2\mathrm{E(X)}x_i + (\mathrm{E(X)})^2\Big) \\
&= \frac{1}{n}\sum_{i=1}^{n} x_i^2 - 2\mathrm{E(X)}\,\frac{1}{n}\sum_{i=1}^{n} x_i + (\mathrm{E(X)})^2 \\
\mathrm{Var(X)} &= \frac{1}{n}\sum_{i=1}^{n} x_i^2 - (\mathrm{E(X)})^2
\end{aligned}$$

donc

$$\frac{1}{n}\sum_{i=1}^{n} x_i^2 = \mathrm{Var(X)} + (\mathrm{E(X)})^2$$

Par ailleurs

$$\begin{aligned}
\mathrm{Cov(X, Y)} &= \frac{1}{n}\sum_{i=1}^{n} (x_i - \mathrm{E(X)})(y_i - \mathrm{E(Y)}) \\
&= \frac{1}{n}\sum_{i=1}^{n} \Big(x_i y_i - \mathrm{E(Y)}x_i - \mathrm{E(X)}x_i + \mathrm{E(X)}\mathrm{E(Y)}\Big) \\
&= \frac{1}{n}\sum_{i=1}^{n} x_i y_i - \mathrm{E(Y)}\,\frac{1}{n}\sum_{i=1}^{n} x_i - \mathrm{E(X)}\,\frac{1}{n}\sum_{i=1}^{n} y_i + (\mathrm{E(X)})^2 \\
\mathrm{Cov(X, Y)} &= \frac{1}{n}\sum_{i=1}^{n} x_i y_i - \mathrm{E(X)}\mathrm{E(Y)}
\end{aligned}$$

soit

$$\frac{1}{n}\sum_{i=1}^{n} x_i y_i = \mathrm{Cov(X,Y)} + \mathrm{E(X)E(Y)}$$

Finalement, le système que l'on cherche à résoudre s'écrit :

$$\begin{pmatrix} 1 & \mathrm{E(X)} \\ \mathrm{E(X)} & \mathrm{Var(X)} + (\mathrm{E(X)})^2 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \mathrm{E(Y)} \\ \mathrm{Cov(X,Y)} + \mathrm{E(X)E(Y)} \end{pmatrix}$$

le déterminant de ce système est égal à

$$\begin{vmatrix} 1 & \mathrm{E(X)} \\ \mathrm{E(X)} & \mathrm{Var(X)} + (\mathrm{E(X)})^2 \end{vmatrix} = \mathrm{Var(X)}$$

On trouve finalement, en utilisant les formules de Cramer

$$\begin{aligned}
a &= \frac{1}{\mathrm{Var(X)}}\begin{vmatrix} \mathrm{E(Y)} & \mathrm{E(X)} \\ \mathrm{Cov(X,Y)} + \mathrm{E(X)E(Y)} & \mathrm{Var(X)} + (\mathrm{E(X)})^2 \end{vmatrix} \\
&= \frac{1}{\mathrm{Var(X)}}\left(\mathrm{E(Y)Var(X)} + \mathrm{E(Y)}(\mathrm{E(X)})^2 - \mathrm{E(X)Cov(X,Y)} - \mathrm{E(Y)}(\mathrm{E(X)})^2\right) \\
&= \frac{\mathrm{E(Y)Var(X)} - \mathrm{E(X)Cov(X,Y)}}{\mathrm{Var(X)}} \\
a &= \mathrm{E(Y)} - \frac{\mathrm{Cov(X,Y)}}{\mathrm{Var(X)}}\mathrm{E(X)}
\end{aligned}$$

et

$$\begin{aligned}
b &= \frac{1}{\mathrm{Var(X)}}\begin{vmatrix} 1 & \mathrm{E(Y)} \\ \mathrm{E(X)} & \mathrm{Cov(X,Y)} + \mathrm{E(X)E(Y)} \end{vmatrix} \\
&= \frac{\mathrm{Cov(X,Y)} + \mathrm{E(X)E(Y)} - \mathrm{E(X)E(Y)}}{\mathrm{Var(X)}} \\
b &= \frac{\mathrm{Cov(X,Y)}}{\mathrm{Var(X)}}
\end{aligned}$$

On obtient ainsi une droite approchant le nuage de points $(x_i, y_i)_{1 \leqslant i \leqslant n}$. Cette droite est appelée *droite des moindres carrés* ou *droite de régression linéaire*. Elle a pour équation

$$\mathrm{Var(X)}\Big(y - \mathrm{E(Y)}\Big) - \mathrm{Cov(X,Y)}\Big(x - \mathrm{E(X)}\Big) = 0$$

6.b En échangeant les rôles de X et Y, on trouve

$$\begin{cases} c = \mathrm{E(X)} - \dfrac{\mathrm{Cov(X,Y)}}{\mathrm{Var(Y)}}\mathrm{E(Y)} \\ d = \dfrac{\mathrm{Cov(X,Y)}}{\mathrm{Var(Y)}} \end{cases}$$

La droite que l'on trouve alors a pour équation

$$\mathrm{Var(Y)}\Big(x - \mathrm{E(X)}\Big) - \mathrm{Cov(X,Y)}\Big(y - \mathrm{E(Y)}\Big) = 0$$

On a ainsi construit deux droites pouvant approcher les points (x_i, y_i). Si ces points sont alignés, on trouve la même droite. Sinon, un bon test pour voir si les points son effectivement « presque » alignés est de dire que les coefficients directeurs de ces deux droites sont égaux. Cela donne

$$\frac{\mathrm{Cov}(\mathrm{X},\mathrm{Y})}{\mathrm{Var}(\mathrm{X})} = \frac{\mathrm{Var}(\mathrm{Y})}{\mathrm{Cov}(\mathrm{X},\mathrm{Y})}$$

c'est à dire que

$$\frac{\Big(\mathrm{Cov}(\mathrm{X},\mathrm{Y})\Big)^2}{\mathrm{Var}(\mathrm{X})\mathrm{Var}(\mathrm{Y})} = 1$$

cette dernière quantité est appelée *coefficient de correlation.* Ce coefficient est toujours inférieur à 1 (d'après l'inégalité de Cauchy-Schwarz), et s'il est « proche » de 1, on considère que le regression linéaire approche bien le nuage de points.

7. Meilleure approximation au sens de Tchebychev

7.a Les polynômes $\mathrm{T}_n(x)$ sont définis, pour $x \in [-1\,;1]$ par

$$\mathrm{T}_n(x) = \cos\left(n \operatorname{Arccos}(x)\right)$$

Si $n = 0$, alors

$$\mathrm{T}_0(x) = \cos(0) = 1$$

Si $n = 1$,

$$\mathrm{T}_n(x) = \cos\left(\operatorname{Arccos}(x)\right) = x$$

En additionnant les relations

$$\begin{aligned} \cos((n+1)u) &= \cos(nu)\cos(u) - \sin(nu)\sin(u) \\ \cos((n-1)u) &= \cos(nu)\cos(u) + \sin(nu)\sin(u) \end{aligned}$$

on trouve

$$\cos((n+1)u) + \cos((n-1)u) = 2\cos(nu)\cos(u)$$

en évaluant cette expression pour $u = \operatorname{Arccos}(x)$, on trouve

$$\mathrm{T}_{n+1}(x) + \mathrm{T}_{n-1}(x) = 2x\mathrm{T}_n(x)$$

d'où la relation de récurrence

$$\mathrm{T}_{n+1}(x) = 2x\mathrm{T}_n(x) - \mathrm{T}_{n-1}(x) \tag{1}$$

Montrons que la suite de polynômes définie d'une part par la relation de récurrence (1), et d'autre part par la donnée de

$$\left\{ \begin{aligned} \mathrm{T}_0(\mathrm{X}) &= 1 \\ \mathrm{T}_1(\mathrm{X}) &= \mathrm{X} \end{aligned} \right.$$

définit une suite de polynômes de degré n, et de coefficient dominant 2^{n-1}. On note

$$\mathscr{P}(n)\ :\ \text{« } \forall k \leqslant n\ \mathrm{T}_k \text{ est de degré } k \text{ et de coefficient dominant } 2^{k-1} \text{ »}$$

- $\mathscr{P}(0)$, $\mathscr{P}(1)$: le résultat est clair.
- $\mathscr{P}(n) \Longrightarrow \mathscr{P}(n+1)$ On sait que T_{n-1} et T_n son respectivement de degré $n-1$ et n, et que leurs coefficients dominants sont 2^{n-2} et 2^{n-1}. Alors $2\mathrm{X}\mathrm{T}_n(\mathrm{X})$ est de degré $n+1$, et T_{n-1} est de degré $n-1$, donc T_{n+1} est de degré $n+1$. De plus, son coefficient dominant est le double de celui de T_n, donc est égal à 2^n.
- Conclusion pour tout n entier naturel, T_n est de degré n, et son coefficient dominant est 2^{n-1}.

7.b Cherchons les racines de T_n dans $[-1\,;1]$. Dans ce cas, les racines x_i peuvent s'écrire sous la forme $\cos(\theta_i)$, où $\theta_i \in [0\,;\pi]$. On a alors

$$\begin{aligned} \mathrm{T}_n(\cos\theta_i) = 0 \quad &\Longleftrightarrow \quad \cos(n\theta_i) = 0 \\ &\Longleftrightarrow \quad \exists k \in \mathbb{N} \quad n\theta_i = \frac{\pi}{2} + k\pi \\ &\Longleftrightarrow \quad \exists k \in \mathbb{N} \quad \theta_i = \frac{(2k+1)\pi}{2n} \end{aligned}$$

De cette manière, on a trouvé n racines distinctes : les $\cos\left(\frac{(2k+1)\pi}{2n}\right)$, pour $k \in [\![0\,;\, n-1]\!]$. Comme le polynôme T_n est de degré n, on a ainsi déterminé toutes les racines de T_n.

Afin de déterminer les extrema de T_n, on va également se placer sur $[-1\,;1]$. On a alors $\mathrm{T}_n(x) = \cos(n \operatorname{Arccos}(x))$.

$$\begin{aligned}\frac{\mathrm{dT}_n}{\mathrm{d}x}(x) &= \frac{\mathrm{d}}{\mathrm{d}x}\left(\cos(n \operatorname{Arccos}(x))\right)\\ &= \frac{n}{\sqrt{1-x^2}} \sin(n \operatorname{Arccos}(x))\end{aligned}$$

Si, comme pour la recherche des racines de T_n, on pose $x_k = \theta_k$ les abscisses des extrema de T_n, alors on trouve

$$\forall k \in [\![1\,;\, n-1]\!] \qquad x_k = \cos\left(\frac{k\pi}{n}\right)$$

On a ainsi trouvé $n-1$ extrema pour T_n, qui est de degré n. On les a donc tous trouvés (puisque les abscisses des extrema de T_n annulent le polynôme de degré $n-1$ T'_n). En ces extrema, on a $|\mathrm{T}_n(x_k)| = 1$ (les valeurs ± 1 sont prises alternativement). De plus, ces valeurs extremales sont également atteintes en -1 et en 1.

7.c On note $q(x)$ un polynôme de degré n, et de même coefficient dominant que T_n. Alors le polynôme $\mathrm{T}_n - q$ est de degré au plus $n-1$. Supposons que q vérifie

$$\max_{x\in[-1\,;1]} |q(x)| < \max_{x\in[-1\,;1]} |\mathrm{T}_n(x)|$$

Alors, lorsqu'on évalue $\mathrm{T}_n - q$ en les abscisses des extrema de T_n, la quantité évaluée est du même signe que T_n. En effet, en ces points, T_n est extrêmal, et on sait qu'en tous ces points, q est strictement inférieur à l'extremum atteint. Le polynôme $\mathrm{T}_n - q$ change donc au moins n fois de signe sur $[-1\,;1]$. Sur ces intervalles, $\mathrm{T}_n - q$ s'annule donc à chaque fois au moins une fois. On en déduit que $\mathrm{T}_n - q$ est de degré $n-1$, et admet au moins n racines, donc il s'agit du polynôme nul. Ceci contredit le fait que $\mathrm{T}_n \neq q$, et on en déduit que

$$\max_{x\in[-1\,;1]} |\mathrm{T}_n(x)| \leqslant \max_{x\in[-1\,;1]} |q(x)|$$

pour tout polynôme q de degré n et de coefficient dominant 2^{n-1}.

7.d Afin de généraliser au cas d'un intervalle $[a\,;b]$ quelconque, on cherche une transformation affine envoyant $[a\,;b]$ sur $[-1\,;1]$: on cherche α et β tels que

$$\begin{cases}\alpha a + \beta = -1\\ \alpha b + \beta = 1\end{cases}$$

on trouve alors

$$\begin{cases}\alpha = \dfrac{2}{b-a}\\ \beta = -\dfrac{a+b}{b-a}\end{cases}$$

Soit P un polynôme de degré n. On a alors :

$$\max_{x\in[a\,;b]} |\mathrm{P}(x)| \geqslant \max_{x\in[-1\,;1]} \left|\mathrm{P}\left(\frac{-2x+a+b}{a-b}\right)\right| \geqslant \max_{x\in[-1\,;1]} \left|\mathrm{T}_n\left(\frac{-2x+a+b}{a-b}\right)\right|$$

Les points de Tchebychev sur $[\,a\,;b\,]$ sont donc les points de la forme

$$x_k = \frac{1}{2}\left((a+b) - (a-b)\cos\left(\frac{k\pi}{n}\right)\right)$$

Application

Lorsqu'on choisit $[\,-8\,;8\,]$, et $x \mapsto \dfrac{1}{1+x^2}$, les points équidistants sont de la forme

$$-8 + \frac{16k}{n} \quad k \in [\![\,0\,;\,n\,]\!]$$

alors les points de Tchebychev sont de la forme

$$8\cos\left(\frac{k\pi}{n}\right) \quad k \in [\![\,0\,;\,n\,]\!]$$

Les tracés de la fonction exacte, de l'interpolée avec des points equidistants, et de l'interpolée avec les points de Tchebychev ont été réalisés pour 11, puis 21 points. Les résultats sont sur les figures 4 et 5. Pour le choix des points uniformément répartis sur $[\,-8\,;8\,]$, on observe de fortes oscillations sur les bords de l'intervalle. Ces oscillations sont appelées *Phénomène de Runge.*

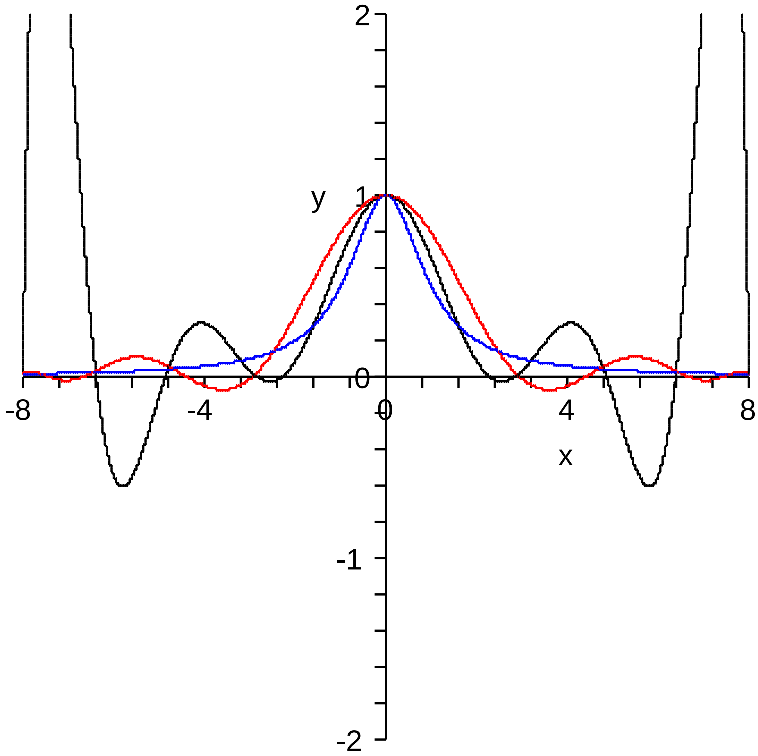

FIG. 4 – Le tracé de la fonction (courbe bleue), de son interpolée avec des points equidistants (courbe noire), et de son interpolée avec les points de Tchebychev (courbe rouge) pour $n = 10$ (11 points)

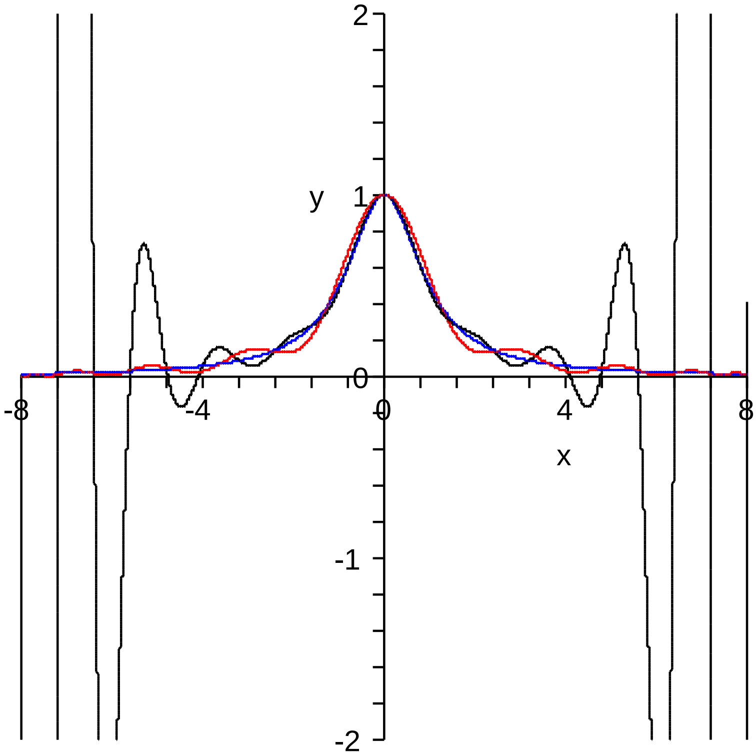

Fig. 5 – Le tracé de la fonction (courbe bleue), de son interpolée avec des points equidistants (courbe noire), et de son interpolée avec les points de Tchebychev (courbe rouge) pour $n = 20$ (21 points)