

# Interpolation et Approximation par des B-splines

Pierre Pansu

February 9, 2004

## 1 Interpolation

*Interpoler* une famille de  $N + 1$  points  $Q_j$  de  $\mathbf{R}^n$ , c'est trouver une courbe qui passe par tous ces points. En plus de prescrire des positions, on peut aussi vouloir prescrire des tangentes.

Le prototype du problème d'interpolation consiste à se donner  $a = t_0 < t_1 < \dots < t_N = b$  et  $Q_0, \dots, Q_N \in \mathbf{R}$ , et chercher une fonction  $x : [a, b] \rightarrow \mathbf{R}$  telle que  $f(t_i) = Q_i$  pour tout  $i = 0, \dots, N$ , voir figure 1.

### 1.1 Polynôme d'interpolation de Lagrange

Une réponse classique à la question consiste à chercher  $f$  parmi les polynômes de degré au plus  $N$ .

**Théorème 1** *Etant donnés  $a = t_0 < t_1 < \dots < t_N = b$  et  $Q_0, \dots, Q_N \in \mathbf{R}$ , il existe un unique polynôme  $x : [a, b] \rightarrow \mathbf{R}$  de degré  $\leq N$  tel que  $x(t_i) = Q_i$  pour tout  $i = 0, \dots, N$ . On l'appelle le polynôme d'interpolation de Lagrange.*

**Preuve.** Soit  $E$  l'espace vectoriel des polynômes à coefficients réels, de degré  $\leq N$ . Alors  $\dim E = N + 1$ . On considère l'application linéaire

$$E \rightarrow \mathbf{R}^{N+1}, \quad x \mapsto (x(t_0), \dots, x(t_N)).$$

Pour montrer qu'elle est bijective, il suffit de vérifier qu'elle est injective. Or si  $x(t_0) = \dots = x(t_N) = 0$ , le polynôme  $x$  a  $N + 1 > \deg(f)$  racines distinctes, donc il est nul. ■

### 1.2 Erreur d'interpolation

La qualité d'un procédé d'interpolation se mesure par une *estimation d'erreur*. Il s'agit de vérifier si l'interpolante approche bien une fonction donnée suffisamment régulière. Autrement dit, étant

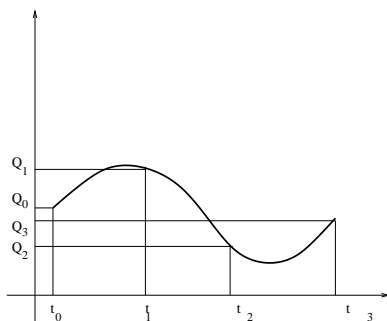


Figure 1: Interpolation en dimension 1

donnée une fonction  $f$  suffisamment différentiable, de majorer la différence entre  $f$  et la fonction  $x$  qui *interpole*  $f$  aux points  $t_i$ , i.e. telle que  $x(t_i) = f(t_i)$  pour  $i = 0, \dots, N$ .

Pour le polynôme d'interpolation de Lagrange, l'estimation d'erreur classique est la suivante.

**Théorème 2** Soit  $f$  une fonction de classe  $C^{N+1}$  sur  $[a, b]$  et  $x_{Lag, N+1}$  le polynôme d'interpolation de Lagrange de  $f$  aux points  $a = t_0 < \dots < t_N = b$ . Alors

$$\|f - x_{Lag, N+1}\|_\infty \leq \frac{1}{(N+1)!} \|q_{N+1}\|_\infty \|f^{(N+1)}\|_\infty$$

où  $q_{N+1}(t) = \prod_{i=0}^N (t - t_i)$ .

**Preuve.** Posons  $g = f - x_{Lag, N+1}$ . Fixons  $t \in [a, b]$  distinct des  $t_i$ . Posons

$$k(u) = g(u) - q_{N+1}(u)g(t)/q_{N+1}(t).$$

Alors  $k(t_i) = 0$ ,  $k(t) = 0$  et  $k^{(N+1)} = f^{(N+1)} - (N+1)!g(t)/q_{N+1}(t)$ . Le théorème de Rolle donne un zéro de  $k'$  dans chacun des  $N+1$  intervalles délimités par les  $t_i$  et  $t$ . Une nouvelle application du théorème de Rolle donne  $N$  zéros pour  $k''$ . On continue jusqu'à trouver un zéro  $\xi$  pour  $k^{(N+1)}$  dans l'intervalle  $[a, b]$ . Il vient

$$\begin{aligned} |g(t)| &= \frac{1}{(N+1)!} q_{N+1}(t) f^{(N+1)}(\xi) \\ &\leq \frac{1}{(N+1)!} \|q_{N+1}\|_\infty \|f^{(N+1)}\|_\infty . \blacksquare \end{aligned}$$

### 1.3 Des polynômes aux splines

La faiblesse de l'interpolation de Lagrange, c'est que l'erreur d'interpolation croît avec  $N$ . Cela se traduit expérimentalement par de grandes oscillations du polynôme d'interpolation, même si  $f$  est très simple. Par exemple (Runge 1901), lorsqu'on interpole la fonction  $x \mapsto 1/(1+25x^2)$  en des points uniformément répartis sur l'intervalle  $[-1, 1]$ , les polynômes de Lagrange  $x_{Lag, N}$  ne convergent pas vers  $f$ . D'où l'idée d'interpoler par des fonctions polynomiales par morceaux, dont le degré n'augmente pas avec le nombre de points d'interpolation.

Par exemple, la fonction  $X_1$  qui est affine sur chaque intervalle  $[t_i, t_{i+1}]$  et interpole une fonction  $f$  de classe  $C^2$  satisfait, d'après le théorème 2 appliqué sur chaque intervalle,

$$\|f - X_1\|_\infty \leq 2h^2 \|f^{(2)}\|_\infty .$$

Comme on souhaite interpoler par des fonctions plus différentiables que  $X_1$ , on cherche l'interpolante dans l'espace des fonctions de classe  $C^{k-1}$  polynomiales de degré  $k$  sur chaque intervalle  $[t_i, t_{i+1}]$ , appelées *fonctions splines* (la terminologie a été introduite par Schoenberg en 1946 : en anglais, **spline** désigne une bande de métal souple utilisée par les dessinateurs pour tracer une jolie courbe entre deux points). C'est ainsi que sont nées les fonctions et les courbes B-splines.

### 1.4 Interpolation par des courbes B-splines

On se limite aux courbes de degré 3, pour simplifier. On cherche à faire passer une courbe B-spline de degré  $k$  de positions et vitesses aux extrémités prescrites par  $N-1$  points  $Q_i$ . Le problème se divise en deux phases.

Première phase : On se fixe un vecteur de noeuds  $\mathbf{t}$  et on cherche un polygone de contrôle  $\mathbf{P}$  tel que la courbe B-spline  $X_k$  correspondante passe par les  $Q_i$  aux noeuds. L'interpolation se traduit alors par la résolution d'un système linéaire.

Deuxième phase : on cherche à optimiser le choix du vecteur de noeuds. C'est typiquement non linéaire.

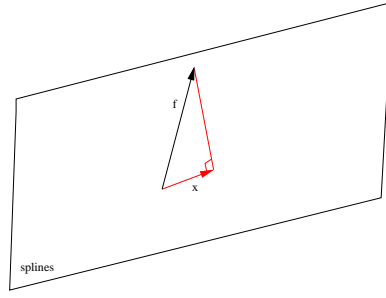


Figure 2: L'interpolante est une projection orthogonale

## 1.5 Le problème linéaire

A nouveau, on se contente d'énoncer le théorème pour des B-splines de degré 3. Une généralisation à tout degré impair se trouve dans [K], Theorem 8.30 page 172.

**Théorème 3** Soient  $Q_0, \dots, Q_N$  des points de  $\mathbf{R}^n$ . Soient  $v_a, v_b$  deux vecteurs de  $\mathbf{R}^n$ . Soit  $\mathbf{t}$  un vecteur de noeuds vissé aux extrémités, de la forme

$$t_0 = t_1 = t_2 = t_3 = a < t_4 < \dots < t_{N+2} < b = t_{N+3} = t_{N+4} = t_{N+5} = t_{N+6}.$$

Il existe un unique polygone de contrôle  $\mathbf{P} = (P_0, \dots, P_{N+2})$  tel que la courbe B-spline de degré 3 associée satisfasse

$$\forall j = 0, \dots, N, X_3(t_{j+3}) = Q_j, \quad X_3'(a) = v_a, \quad \text{et } X_3'(b) = v_b.$$

**Preuve.** Comme chaque coordonnée se traite indépendamment, on peut supposer que  $n = 1$ . Dans ce cas, on considère l'application linéaire

$$\mathbf{R}^{N+3} \rightarrow \mathbf{R}^{N+3}, \quad (P_0, \dots, P_{N+2}) \mapsto (X_3'(a), X_3(t_3), \dots, X_3(t_{N+3}), X_3'(b)).$$

Pour prouver qu'elle est bijective, il suffit de prouver qu'elle est injective. Cela résulte immédiatement du lemme suivant appliqué à la fonction  $f = 0$ . En effet, si les points et vecteurs interpolés sont tous nuls, le lemme donne  $X_3'' \equiv 0$ . Avec la condition initiale  $X_3(a) = X_3'(a) = 0$ , cela entraîne que  $X_3 \equiv 0$ , et donc que les  $P_i$  sont tous nuls. ■

**Lemme 1.1** Soient  $f, x : [a, b] \rightarrow \mathbf{R}$  deux fonctions de classe  $C^2$ . On suppose que

- $x$  est polynômiale de degré 3 sur chaque intervalle  $[t_i, t_{i+1}]$ ,  $i = 3, \dots, N + 2$  ;
- $f(t_i) = x(t_i)$  pour  $i = 3, \dots, N + 3$  et  $f'(a) = x'(a)$ ,  $f'(b) = x'(b)$ .

Alors

$$\int_a^b (f''(t) - x''(t))^2 dt = \int_a^b f''(t)^2 dt - \int_a^b x''(t)^2 dt.$$

**Remarque.** Ce lemme signifie que la spline  $x$  qui interpole une fonction  $f$  est la projection orthogonale de  $f$  sur le sous-espace des splines, pour le produit scalaire  $f \cdot g = \int_a^b f''(t)g''(t) dt$ , voir figure 2.

**Preuve.** On vérifie immédiatement que

$$\int_a^b (f''(t) - x''(t))^2 dt - \int_a^b f''(t)^2 dt + \int_a^b x''(t)^2 dt = -2R$$

où

$$R = \int_a^b (f''(t) - x''(t))x''(t)dt.$$

On intègre par parties sur chaque intervalle

$$\begin{aligned} \int_{t_i}^{t_{i+1}} (f''(t) - x''(t))x''(t) dt &= (f'(t_{i+1}) - x'(t_{i+1}))x''(t_{i+1}) - (f'(t_i) - x'(t_i))x''(t_i) \\ &\quad - \int_{t_i}^{t_{i+1}} (f'(t) - x'(t))x'''(t) dt \\ &= (f'(t_{i+1}) - x'(t_{i+1}))x''(t_{i+1}) - (f'(t_i) - x'(t_i))x''(t_i) \\ &\quad - (f(t_{i+1}) - x(t_{i+1}))x'''(t_{i+1-}) - (f(t_i) - x(t_i))x'''(t_i+) \\ &\quad + \int_{t_i}^{t_{i+1}} (f(t) - x(t))x^{(4)}(t) dt \\ &= (f'(t_{i+1}) - x'(t_{i+1}))x''(t_{i+1}) - (f'(t_i) - x'(t_i))x''(t_i), \end{aligned}$$

car  $x^{(4)} \equiv 0$  et  $f(t_{i+1}) - x(t_{i+1}) = f(t_i) - x(t_i) = 0$ . En additionnant, il vient

$$\begin{aligned} R &= (f'(b) - x'(b))x''(b) - (f'(a) - x'(a))x''(a) \\ &= 0, \end{aligned}$$

car  $f'(a) - x'(a) = f'(b) - x'(b) = 0$ . ■

## 1.6 Estimation de l'erreur d'interpolation

**Théorème 4** Soit  $f : [a, b] \rightarrow \mathbf{R}$  une fonction de classe  $C^2$ . Soit  $X_3$  la fonction B-spline de degré 3 qui l'interpole en  $N + 1$  point plus les dérivées aux bornes, selon le théorème 3. Alors

$$\|f - X_3\|_\infty \leq \frac{h^{3/2}}{2} \|f''\|_2 \quad \text{et} \quad \|f' - X_3'\|_\infty \leq h^{1/2} \|f''\|_2$$

où  $h = \max\{t_{i+1} - t_i\}$ .

**Preuve.** Posons  $g = f - X_3$ . Le lemme 1.1 donne

$$\begin{aligned} \|g''\|_2^2 &= \int_a^b g''(t)^2 dt \\ &= \int_a^b f''(t)^2 dt - \int_a^b X_3''(t)^2 dt \\ &\leq \|f''\|_2^2. \end{aligned}$$

Par construction,  $g$  s'annule aux  $t_i$ , donc, d'après le théorème de Rolle,  $g'$  s'annule au moins une fois dans chaque intervalle  $[t_i, t_{i+1}]$ . Tout point  $t \in [a, b]$  est donc à distance au plus  $h$  d'un point  $t'$  tel que  $g'(t') = 0$ . On écrit

$$\begin{aligned} |g'(t)| &= |g'(t) - g'(t')| \\ &= \left| \int_{t'}^t g''(s) ds \right| \\ &\leq \left( \int_{t'}^t ds \right)^{1/2} \left( \int_{t'}^t g''(s)^2 ds \right)^{1/2} \\ &\leq h^{1/2} \|g''\|_2 \\ &\leq h^{1/2} \|f''\|_2, \end{aligned}$$

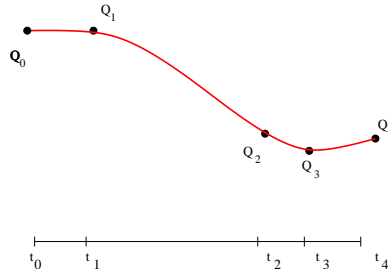


Figure 3: Paramétrisation chordale

d'après l'inégalité de Cauchy-Schwarz. Ceci montre que

$$\|g'\|_{\infty} \leq h^{1/2} \|f''\|_2$$

Tout point  $t \in [a, b]$  est à distance au plus  $h/2$  d'un point  $t''$  tel que  $g(t'') = 0$ . Par conséquent

$$\begin{aligned} |g(t)| &= |g(t) - g(t'')| \\ &= \left| \int_{t''}^t g'(s) ds \right| \\ &\leq |t - t''| \|g'\|_{\infty} \\ &\leq \frac{h^{1/2}}{2} \|f''\|_2 \cdot \blacksquare \end{aligned}$$

**Remarque.** Dans [K], Theorem 8.34 page 178, on trouve l'amélioration suivante. *Sous les hypothèses du Théorème 4, et si de plus  $f$  est de classe  $C^4$ , alors*

$$\|f - X_3\|_{\infty} \leq \frac{h^4}{16} \|f^{(4)}\|_{\infty}.$$

## 1.7 Choix du vecteur de noeuds

Etant donnés les points  $Q_i$  à interpoler, quel est le meilleur choix des  $t_i$  ? Pour éviter que la dérivée de la courbe interpolante soit grande, il vaut mieux que des points  $Q_i$  et  $Q_{i+1}$  éloignés soient interpolés en des valeurs  $t_i$  et  $t_{i+1}$  éloignées. Autrement dit, il faut corrélérer les espacements  $t_{i+1} - t_i$  avec les distances  $\|Q_{i+1} - Q_i\|$ .

Un choix simple consiste à poser

$$t_{i+1} - t_i = \|Q_{i+1} - Q_i\|.$$

Ce choix s'appelle la *paramétrisation chordale* (**chordal parametrization**). Elle a pour effet de produire une courbe interpolante paramétrée à vitesse à peu près constante, voir figure 3.

Lorsque les distances  $\|Q_{i+1} - Q_i\|$  varient brutalement, il y a de meilleurs choix, voir [HL], figures 4.36a et 4.36b pages 205 et 206.

On gagne de la souplesse à fixer des valeurs  $s_j$  où interpoler indépendantes des noeuds et à fixer ensuite les noeuds de façon à optimiser le conditionnement de la matrice du système linéaire, voir [HL].

## 2 Approximation

### 2.1 Position du problème

Le problème d'*approximation* d'une famille de  $N + 1$  points  $Q_j$  par une courbe consiste, au lieu d'exiger que la courbe passe exactement par les  $Q_j$ , i.e.  $x(s_j) = Q_j$  (ce qui nécessite autant de

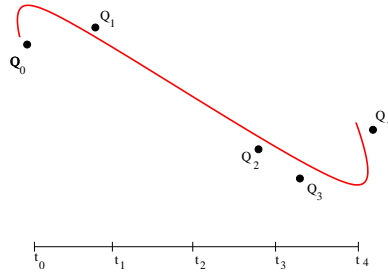


Figure 4: Approximation d'un nuage de points par une courbe

degré de liberté que d'équations, à savoir  $n(N + 1)$ ), à chercher à minimiser l'écart quadratique

$$\sum_j \|x(s_j) - Q_j\|^2$$

où  $x$  varie dans un espace possédant typiquement bien moins de degrés de liberté, voir figure 4.

**Exemple.** Le problème d'approximation de  $N + 1$  points  $Q_j$  par une application constante, est bien posé.

En effet, l'intervalle  $[0, 1]$  est subdivisé par exemple par  $s_j = j/N$ ,  $j = 0, \dots, N$ . L'inconnue est une application constante de  $[0, 1]$  dans  $\mathbf{R}^n$ , i.e. un point  $x \in \mathbf{R}^n$ . Il doit minimiser

$$E(P, v) = \sum_{j=0}^N \|x - Q_j\|^2 .$$

La solution est le centre de gravité des points  $Q_j$ . Dans ce problème, l'inconnue  $x$  vit dans un espace de dimension  $n < n(N + 1)$ .

## 2.2 Approximation par une courbe B-spline

Montrons que le problème se ramène à nouveau à la résolution d'un système linéaire.

Soit

$$a = t_0 = \dots = t_k < t_{k+1} < \dots < t_{m-k} = \dots = t_m = b$$

un vecteur de noeuds vissé aux extrémités. Soit  $\mathbf{P} = (P_0, \dots, P_{m-k-2}) \in \mathbf{R}^{n(m-k-1)}$  un polygone de contrôle. La courbe B-spline de degré  $k$  correspondante est donnée sur  $[a, b]$  par  $X_k(t) = \sum_{i=0}^{m-k-2} B_{i,k}(t)P_i$ . Soient  $a \leq s_0 < \dots < s_N \leq b$  des valeurs distinctes dans  $[a, b]$  et  $Q_0, \dots, Q_N$  des points dans  $\mathbf{R}^n$ .

Le problème d'approximation consiste à trouver  $\mathbf{P}$  qui minimise

$$E(\mathbf{P}) = \sum_{j=0}^N \|X_k(s_j) - Q_j\|^2 .$$

On remarque que  $E$  est un polynôme du second degré sur  $\mathbf{R}^{n(m-k-1)}$ . Si on représente les coordonnées de tous les  $P_i$  simultanément par une colonne  $V$  à  $n(m - k - 1)$  coefficients, alors  $E$  s'écrit

$$E(\mathbf{P}) = {}^T V A V + 2 {}^T B V + c$$

où  $A$  est une matrice symétrique positive,  $B$  est la colonne des coordonnées des vecteurs

$$\sum_j B_{i,k}(s_j) Q_j,$$

et

$$c = \sum_j \|Q_j\|^2.$$

La fonction  $E$  atteint donc son minimum aux vecteurs  $V$  solutions de l'équation vectorielle

$$AV + B = 0.$$

Il s'agit d'un système linéaire à  $n(m - k - 1)$  équations et  $n(m - k - 1)$  inconnues. Noter que la taille ne dépend pas du nombre de points à approximer.

Typiquement, l'approximation est utilisée pour alléger des données : remplacer une courbe composite ou comportant un grand nombre de points de contrôle par une seule B-spline ayant peu de points de contrôle.

### 3 Résolution numérique d'un problème d'interpolation

On se donne le vecteur de noeuds

$$t_0 = t_1 = t_2 = t_3 = 0 < t_4 = 1 < \dots < t_{N+2} = N - 1 < N = t_{N+3} = t_{N+4} = t_{N+5} = t_{N+6}$$

dans l'intervalle  $[0, N]$ . Il s'agit de trouver le polygone de contrôle (à  $N + 3$  sommets) de la B-spline qui passe par le point  $Q_i$  en  $t_{i+3}$  et a pour dérivées  $v_0$  (resp.  $v_N$ ) aux extrémités. On peut supposer que  $n = 1$ .

On utilise les valeurs des fonctions B-splines relatives au vecteur de noeuds ci-dessus (en exercice du cours de géométrie différentielle).

$$X_3(0) = P_0, \tag{1}$$

$$X_3'(0) = 3(P_1 - P_0), \tag{2}$$

$$X_3(t_4) = \frac{1}{4}P_1 + \frac{7}{12}P_2 + \frac{1}{6}P_3, \tag{3}$$

$$X_3(t_{i+3}) = \frac{1}{6}P_i + \frac{2}{3}P_{i+1} + \frac{1}{6}P_{i+2} \tag{4}$$

pour  $i \geq 2$ .

Il s'agit de résoudre le système  $A\mathbf{P} = \mathbf{Q}$  pour  $\mathbf{Q} = (Q_0, v_0, Q_1, \dots, Q_{N-1}, v_N, Q_N)$  et  $\mathbf{P} \in \mathbf{R}^{N+3}$  et

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & \dots \\ -3 & 3 & 0 & 0 & 0 & \dots \\ 0 & \frac{1}{4} & \frac{7}{12} & \frac{1}{6} & 0 & \dots \\ 0 & 0 & \frac{1}{6} & \frac{2}{3} & \frac{1}{6} & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

La matrice  $A$  n'est pas tout à fait symétrique. Comme on préfère les matrices symétriques, on résoud le système équivalent

$${}^T A A \mathbf{P} = {}^T A \mathbf{Q}.$$

La matrice  ${}^T A A$  est *coercive*,

$$A\mathbf{p} \cdot A\mathbf{p} \geq \frac{1}{3} \|\mathbf{P}\|^2.$$

Cela indique que la résolution devrait être efficace. On applique la méthode du *gradient conjugué*.

### 3.1 Rappels sur la méthode du gradient conjugué

Soit  $E$  un espace de Hilbert et  $q$  une forme quadratique définie positive sur  $E$ ,  $q(x) = Ax \cdot x$ . Alors  $x$  est solution de l'équation  $Ax = y$  si et seulement si  $x$  est le minimum de la fonction  $z \mapsto f(z) = \frac{1}{2}Az \cdot z - y \cdot z$ . La méthode consiste à chercher les minima  $x_r$  de la fonction  $f$  sur les sous-espaces vectoriels croissants

$$E_r = \text{vect}\{y, Ay, \dots, A^r y\}.$$

Alors  $x_r - x_{r-1}$  est donné par une formule simple au moyen de produits scalaires. La suite  $x_r$  converge vers la solution véritable  $x_\infty$ . En particulier, si  $E$  est de dimension  $N$ , alors  $x_N = x_\infty$ . L'expérience prouve que  $x_{N/5}$  est souvent déjà très proche de  $x_\infty$ .

**Proposition 3.1** *On définit trois suites  $x_r$ ,  $v_r$  et  $e_r$  par  $x_0 = 0$ ,  $v_0 = y$ , les relations de récurrence*

$$\begin{aligned} x_{r+1} &= x_r - \frac{e_r \cdot v_r}{Av_r \cdot v_r} v_r, \\ v_{r+1} &= e_{r+1} - \frac{Av_r \cdot e_{r+1}}{Av_r \cdot v_r} v_r \end{aligned}$$

et

$$e_r = Ax_r - y.$$

Alors la fonction  $f : x \mapsto \frac{1}{2}Ax \cdot x - y \cdot x$  atteint son minimum sur  $E_r$  en  $x_r$ . La suite  $x_r$  converge vers l'unique solution  $x_\infty$  de l'équation  $Ax = y$ . De plus, on a l'estimation a posteriori

$$A(x_\infty - x_r) \cdot (x_\infty - x_r) = e_r \cdot (x_{r+1} - x_r).$$

Si  $A$  est coercive, i.e.  $\|x\|^2 \leq C Ax \cdot x$ , on a l'estimation a posteriori

$$\|x_\infty - x_r\| \leq C \|Ax_r - y\|.$$

**Preuve.** Définissons  $x_r$  comme le point de  $E_r$  où la restriction de  $f$  à  $E_r$  atteint son minimum. Il existe et est unique par stricte convexité. Pour la même raison, si  $E_\infty$  est l'adhérence de la réunion des  $E_r$ , la fonction uniformément convexe  $f$  atteint son minimum sur  $E_\infty$  en un unique point  $x_\infty$ .

On utilise le produit scalaire associé à la forme quadratique  $q$ ,

$$a \cdot_A b = Aa \cdot b.$$

Relativement à cette structure euclidienne, les ensembles de niveau de  $f$  sont des sphères centrées en  $x_\infty$ . Les ensembles de niveau de la restriction de  $f$  à  $E_r$  sont des sphères centrées en  $x_r$ . Soit  $s \leq r - 1$ . Au point  $x_s$  où  $f$  atteint son minimum sur  $E_s$  cette sphère est tangente à  $E_s$ , donc  $x_r - x_s \perp_A E_s$ . En particulier  $x_\infty - x_s \perp_A E_s$ .

Posons  $e_r = Ax_r - y$ . Alors  $e_r \in E_{r+1}$ . Or  $e_r = A(x_r - x_\infty) \perp E_r$ , donc en particulier  $e_r \cdot e_{r-1} = 0$ .

Considérons la suite  $v_r$  définie par  $v_0 = y$  et par la relation de récurrence

$$v_{r+1} = e_{r+1} - \frac{v_r \cdot_A e_{r+1}}{v_r \cdot_A v_r} v_r.$$

Montrons par récurrence sur  $r$  que  $v_r \in E_{r+1}$ ,  $v_r \perp_A E_r$  et

$$x_{r+1} = x_r - \frac{e_r \cdot v_r}{Av_r \cdot v_r} v_r.$$

Lorsque  $r = 0$ ,  $v_0 = y \in E_1$  par définition et  $E_0 = \{0\}$ .



Supposons connu que  $v_{r-1} \in E_r$ ,  $v_{r-1} \perp_A E_{r-1}$  et

$$x_r = x_{r-1} - \frac{e_{r-1} \cdot v_{r-1}}{A v_{r-1} \cdot v_{r-1}} v_{r-1}.$$

On remarque que  $e_r \in A E_r \subset E_{r+1}$  donc  $v_r \in E_{r+1}$ .

D'autre part,  $x_\infty - x_r \perp_A E_r$ , donc  $e_r = A(x_\infty - x_r) \perp_A A^{-1} E_r$ . Or  $A E_{r-1} \subset E_r$  donc  $e_r \perp_A E_{r-1}$ . Comme  $v_r$  est un vecteur du plan engendré par  $e_r$  et  $v_{r-1}$ , il est orthogonal à  $E_{r-1}$ . Par construction,  $v_r \perp_A v_{r-1}$ .

Si  $v_{r-1} = 0$ , alors d'après l'hypothèse de récurrence,  $x_r = x_{r-1}$ , donc  $e_r = e_{r-1}$ . Comme  $e_r \cdot e_{r-1} = 0$ , nécessairement  $e_{r-1} = 0$ , donc  $x_\infty = x_{r-1}$ . On conclut que  $x_{r+1} = x_r$ ,  $v_r = 0$  et la relation de récurrence pour  $x_{r+1}$  est satisfaite.

On peut donc supposer que  $v_{r-1} \neq 0$ . Alors  $E_r$  est engendré par  $E_{r-1}$  et  $v_{r-1}$ , donc  $v_r \perp_A E_r$ .

Si  $v_r = 0$ , alors  $e_r$  est colinéaire à  $v_{r-1}$ . Or  $e_r \perp E_{r-1}$ . Nécessairement  $e_r = 0$ , donc  $x_{r+1} = x_r$  et la relation de récurrence pour  $x_{r+1}$  est à nouveau satisfaite.

On peut donc supposer que  $v_r \neq 0$ . Alors  $v_r$  est une base de l'orthogonal de  $E_{r-1}$  dans  $E_r$ , donc il existe un réel  $\lambda$  tel que  $x_{r+1} - x_r = \lambda v_r$ . Comme cette valeur de  $\lambda$  minimise  $f(x_r + \lambda v_r)$ , nécessairement

$$\lambda = -\frac{x_r \cdot_A v_r - y \cdot v_r}{v_r \cdot_A v_r},$$

comme annoncé.

Par construction, la distance (mesurée au moyen de  $x \mapsto AX \cdot x$ ) de  $x_\infty$  au sous-espace  $E_r$  est atteinte en  $x_r$ . Comme la réunion des  $E_r$  est dense dans  $E_\infty$ , cette distance, égale à  $A(x_\infty - x_r) \cdot (x_\infty - x_r)$ , tend vers 0. Remarquer que les vecteurs  $x_{r+1} - x_r$  sont deux à deux orthogonaux (ils sont colinéaires aux vecteurs obtenus en orthonormalisant la base  $y, Ay, \dots, A^r y \dots$ ). On a donc

$$\sum_{r=0}^{\infty} A(x_{r+1} - x_r) \cdot (x_{r+1} - x_r) = A x_\infty \cdot x_\infty.$$

Enfin,

$$\begin{aligned} \|x_r - x_\infty\|^2 &\leq C A(x_r - x_\infty) \cdot (x_r - x_\infty) \\ &= (A x_r - y) \cdot (x_r - x_\infty) \\ &\leq \|A x_r - y\| \|x_r - x_\infty\| \end{aligned}$$

d'où

$$\|x_\infty - x_r\| \leq C \|A x_r - y\|. \blacksquare$$

### 3.2 Méthode de Jacobi

On peut aussi utiliser la méthode de résolution de Jacobi (voir [K], Theorem 4.2 page 55). Soit  $D$  est la diagonale de  $A$ . Alors

$$Ax = y \Leftrightarrow x = -D^{-1}(A - D)x + D^{-1}y$$

et on itère la transformation affine  $x \mapsto -D^{-1}(A - D)x + D^{-1}y$ . Celle-ci est contractante lorsque  $\mathbf{R}^{n+1}$  est muni de la norme  $\ell^1$ . En effet,

$$q = \|D^{-1}(A - D)\|_{\ell^1 \rightarrow \ell^1} = \max\{|A_{jj}|^{-1} \sum_{i \neq j} |A_{ji}|; j = 0, \dots, m\}$$

qui dans le cas présent vaut  $q = 1/2$ . Posant  $x_0 = O$  et

$$x_{\nu+1} = -D^{-1}(A - D)x_\nu + D^{-1}y,$$

on peut affirmer que la solution  $x$  satisfait

$$\|x - x_\nu\|_{\ell^1} \leq \frac{q^\nu}{1 - q} \|D^{-1}y\|_{\ell^1}.$$

et dans le cas présent  $\frac{q^\nu}{1 - q} = 2^{1-\nu}$ .

## References

- [HL] J. HOSCHEK, D. LASSER, *Grundlagen der geometrischen Datenverarbeitung*. Teubner, Stuttgart (1989). **English translation** : *Fundamentals of computer aided geometric design*. A. K. Peters, Wellesley, Mass. (1993).
- [K] R. KRESS, *Numerical Analysis*. Graduate Texts in Math. **181**. Springer, Heidelberg (1998).
- [R] J.-J. RISLER, *Méthodes mathématiques pour la CAO*. Masson, Paris (1991).