ALGÈBRE LINEAIRE Module 2
Structure Euclidienne
PAD - Notes de cours

S. Rigal, D. Ruiz, et J. C. Satgé

December 5, 2008

Table des Matières

1 Espaces euclidiens – Orthogonalité **1**

1-1 Espaces euclidiens . . . 3
1-1.1 Espaces vectoriels normés – Généralités . . . 3
1-1.2 Produit scalaire canonique dans $\mathbb{R}^n$ – Norme euclidienne . . . 4
1-1.3 Produit scalaire canonique dans $\mathbb{C}^n$ . . . 5
1-1.4 Produit scalaire sur un espace vectoriel – Espaces euclidiens . . . 6
1-1.5 Exemples . . . 6
1-2 Bases orthonormées – Matrices orthogonales . . . 9
1-2.1 Orthogonalité . . . 9
1-2.2 Bases orthonormées . . . 10
1-2.3 Matrices orthogonales dans $\mathcal{M}_n(\mathbb{R})$ . . . 11
1-2.4 Matrices unitaires dans $\mathcal{M}_n(\mathbb{C})$ . . . 11
1-3 Procédé d'orthogonalisation de Schmidt . . . 13
1-3.1 Introduction par un exemple . . . 13
1-3.2 Généralisation . . . 14
1-4 Factorisation **QR** . . . 15
1-4.1 Définition – Propriétés . . . 15
1-4.2 Application . . . 16

2 Formes bilinéaires et quadratiques **19**

2-1 Forme bilinéaire – Matrice d'une forme bilinéaire . . . 21
2-1.1 Formes bilinéaires . . . 21
2-1.2 Représentation matricielle d'une forme bilinéaire . . . 21
2-1.3 Exemple dans $\mathbb{R}^3$ . . . 22
2-2 Formes quadratiques . . . 23
2-2.1 Propriétés . . . 24
2-3 Formes quadratiques définies positives . . . 25
2-3.1 Produit scalaire . . . 25
2-3.2 Exemples . . . 25
2-4 Réduction en somme de carrés d'un polynôme homogène de degré 2 . . . 27
2-4.1 Exemple . . . 27
2-4.2 Méthode générale . . . 28
2-5 Diagonalisation des endomorphismes symétriques . . . 29
2-5.1 Introduction . . . 29

2-5.2 Généralisation 30
2-6 Diagonalisation d'une forme quadratique 31

3 Projections et symétries – Premiers problèmes d'optimisation 33
3-1 Projecteurs et symétries 35
3-1.1 Exemples dans $\mathbb{R}^3$ 35
3-1.2 Définitions – Propriétés 37
3-1.3 Projection orthogonale – Projection sur un convexe – Caractérisation 38
3-2 Résolution de systèmes linéaires surdéterminés 41
3-3 Approximation d'une fonction au sens des moindres carrés 43
3-3.1 Approximation en moyenne quadratique 43
3-3.2 Approximation au sens des moindres carrés discrets 44
3-4 Minimisation de fonctionnelles quadratiques généralisées 47

Chapitre 3

Projections et symétries – Premiers problèmes d'optimisation

Dans ce deuxième chapitre :

- Nous étudierons (à la fois de manière géométrique et algébrique) deux endomorphismes particuliers d'un espace vectoriel E : les projecteurs et les symétries, dans le cas général et dans le cas où l'espace E est un espace euclidien.
- Nous introduirons ensuite la notion de système linéaire surdéterminé, et la résolution de tels systèmes au sens des moindres carrés. Nous verrons que cela se ramène à un problème particulier de projections orthogonale.
- En guise d'application, nous construirons des problèmes d'approximation que nous résoudrons à l'aide des résultats et techniques développés auparavant.
- Pour finir, nous nous intéresserons à une classe de problèmes d'optimisation très importante, à savoir la minimisation de fonctionnelles quadratiques généralisées. Nous verrons le lien direct qui existe entre ces problèmes d'optimisation et la résolution de systèmes linéaires particuliers, et nous verrons aussi comment interpréter tout cela de manière géométrique à l'aide de la décomposition en vecteurs propres et valeurs propres.

3-1 Projecteurs et symétries

3-1.1 Exemples dans $\mathbb{R}^3$

Soit $\mathcal{P}$ le plan vectoriel

$$\mathcal{P} = \left\{ \mathbf{u} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \text{ t.q. } 2x + 3y + z = 0 \right\},$$

de vecteurs directeurs (base de $\mathcal{P}$)

$$\mathcal{P} = \mathsf{Vect} \left\{ \mathbf{e}_1 \begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix}, \mathbf{e}_2 \begin{pmatrix} 0 \\ 1 \\ -3 \end{pmatrix} \right\}.$$

et $\mathcal{D}$ la droite vectorielle

$$\mathcal{D} = \mathsf{Vect} \left\{ \mathbf{e}_3 \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \right\}.$$

1. Soit p le projecteur sur $\mathcal{P}$ de direction $\mathcal{D}$. Dans la base $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$, la matrice de p est

$$\mathbf{P} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

   sachant que les vecteurs $\mathbf{e}_1$ et $\mathbf{e}_2$ sont invariants par p, puisque appartenant à $\mathcal{P}$, et que le vecteur $\mathbf{e}_3$, qui dirige $\mathcal{D}$, appartient aussi à $\mathsf{Ker}\, p$ puisque p est une projection selon $\mathcal{D}$.

   On vérifie naturellement que $\mathbf{P}^2 = \mathbf{P}$, ce qui correspond bien à $p \circ p = p$.

   Remarques :

   - $\mathcal{P}$ correspond à l'image de p, et $\mathcal{D}$ correspond au noyau de p.
   - $\mathcal{P}$, qui est aussi l'ensemble des vecteurs invariants par p, est en fait l'espace propre associé à la valeur propre 1.
   - On note aussi que $E = \mathsf{Im}\, p \oplus \mathsf{Ker}\, p$.

2. Soit s la symétrie d'axe $\mathcal{P}$ et de direction $\mathcal{D}$. Dans la base $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$, la matrice de s est

$$\mathbf{S} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix},$$

   sachant que, cette fois, $\mathcal{P}$ est l'espace propre associé à la valeur propre 1, et que $\mathcal{D}$ est l'espace propre associé à la valeur propre -1.

   On vérifie naturellement que $\mathbf{S}^2 = \mathbf{I}_3$, ce qui correspond bien à $s \circ s = \mathbf{I}_E$.

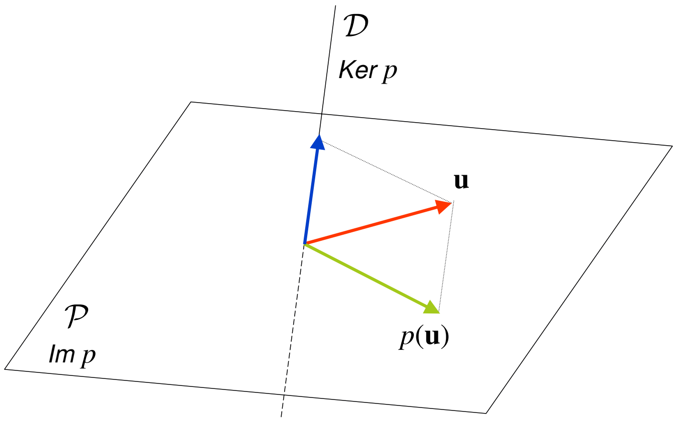

Figure 3-1.1: p est un projecteur sur $\mathcal{P}$ selon $\mathcal{D}$ de $\mathbb{R}^3$

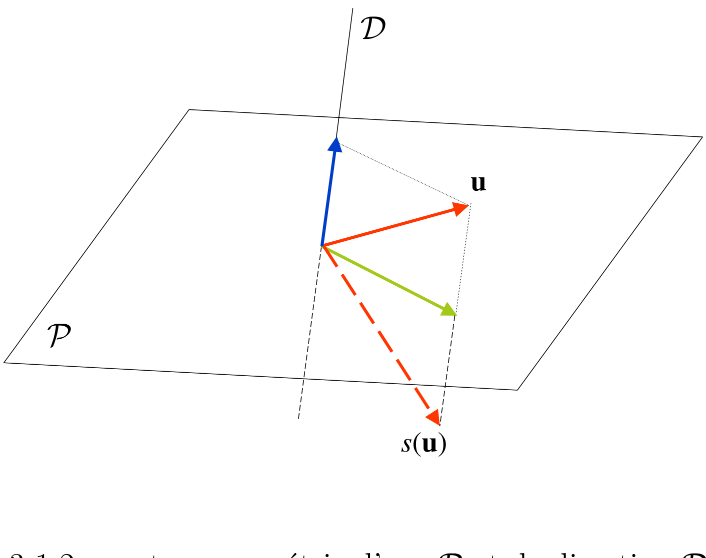

Figure 3-1.2: s est une symétrie d'axe $\mathcal{P}$ et de direction $\mathcal{D}$ de $\mathbb{R}^3$

Si les espaces $\mathcal{P}$ et $\mathcal{D}$ sont orthogonaux, on parle alors de **projecteurs et symétries orthogonaux**. Ce qui caractérise les projecteurs orthogonaux (respectivement symétries orthogonales) est qu'il existe une base orthonormée de E, qui décompose l'espace E en $\mathcal{P} \overset{\perp}{\oplus} \mathcal{D}$ (somme directe orthogonale), dans laquelle la matrice du projecteur orthogonal est formée d'une diagonale de 1 suivis d'une diagonale de 0 (respectivement formée d'une diagonale de 1 suivis d'une diagonale de -1 pour ce qui est de la symétrie orthogonale).

3-1.2 Définitions – Propriétés

Définition 3-1.1 *Soit E un espace vectoriel, et p et s deux endomorphismes de E :*

- *Si $p \circ p = p$, on dit que p est un **projecteur** de E.*
- *Si $s \circ s = \mathbf{I}_E$, on dit que s est une symétrie de E.*

Propriétés : Soit E un espace vectoriel de dimension n. On a alors :

- Si $\mathbf{P}$ est la matrice d'un projecteur exprimé dans une base quelconque de E, alors $\mathbf{P}^2 = \mathbf{P}$.
- Si $\mathbf{S}$ est la matrice d'une symétrie exprimée dans une base quelconque de E, alors $\mathbf{S}^2 = \mathbf{I}_n$.

Dans le cas particulier des espaces vectoriels sur $\mathbb{R}$, on a aussi les résultats suivants :

- Soit E un $\mathbb{R}$-espace vectoriel euclidien de dimension n. Si $\mathbf{P}$ est la matrice d'un projecteur orthogonal exprimé dans la base canonique de E, alors il existe une matrice $\mathbf{Q}$ orthonormale (matrice de changement de base) telle que la matrice $\mathbf{P}$ vérifie la relation de similitude suivante (compte tenu que $\mathbf{Q}^{-1} = \mathbf{Q}^T$) :

$$\mathbf{P} = \mathbf{Q}^T \begin{pmatrix} \mathbf{I}_m & \\ & \mathbf{0}_{n-m} \end{pmatrix} \mathbf{Q}\,,$$

  où m est la dimension de l'espace invariant par $\mathbf{P}$ ($m = \dim(\mathit{Im}\,\mathbf{P})$).

- La relation précédente indique en outre que la matrice d'un projecteur orthogonal, exprimée dans n'importe quelle **base orthonormée**, est symétrique. C'est aussi une caractérisation des projecteurs orthogonaux, c'est à dire qu'ils vérifient $p \circ p = p$ et leur matrice dans toute base orthonormée est symétrique.

- Soit E un $\mathbb{R}$ espace vectoriel de dimension n. Si $\mathbf{S}$ est la matrice d'une symétrie orthogonale exprimé dans la base canonique de E, alors il existe une matrice $\mathbf{Q}$ orthonormale (matrice de changement de base) telle que la matrice $\mathbf{S}$ vérifie la relation de similitude suivante :

$$\mathbf{S} = \mathbf{Q}^T \begin{pmatrix} \mathbf{I}_m & \\ & -\mathbf{I}_{n-m} \end{pmatrix} \mathbf{Q}\,,$$

  où m est la dimension de l'espace propre associé à la valeur propre 1.

- La relation précédente indique, là encore, que la matrice d'une symétrie orthogonale, exprimée dans n'importe quelle **base orthonormée**, est symétrique. C'est aussi une caractérisation, à savoir que les symétries orthogonales vérifient $s \circ s = \mathbf{I}_E$ et leur matrice dans une base orthonormée quelconque est symétrique.

Pour les espaces euclidiens sur le corps $\mathbb{C}$, les propriétés sont très similaires, à ceci près que la notion de symétrie se transforme en "*symétrie hermitienne*".

3-1.3 Projection orthogonale – Projection sur un convexe – Caractérisation

Etant donné un espace euclidien E, V un sous ensemble de E, et $\mathbf{x}$ un élément de E, on est souvent amené à se poser les questions suivantes :

1. Existe-t'il $\mathbf{v}^* \in V$ tel que $\|\mathbf{x} - \mathbf{v}^*\| = \min_{\mathbf{v} \in V} \|\mathbf{x} - \mathbf{v}\|$?

2. Peut-on obtenir une caractérisation de $\mathbf{v}^*$ quand il existe ?

La réponse à ces deux questions peut être en particulier donnée dans le cas où V est un sous-espace vectoriel de l'espace euclidien E, ou bien encore lorsque le sous-ensemble V est un convexe fermé dans E. Ces résultats sont résumés dans les deux théorèmes qui suivent :

Théorème 3-1.1 ***Projection orthogonale***
Soit E un espace euclidien, et soit $V \subset E$ un sous-espace vectoriel de E. Pour tout $\mathbf{x} \in E$ il existe un unique $\mathbf{v}^* \in V$ (appelé projection orthogonale de $\mathbf{x}$ sur V*) tel que*

$$\|\mathbf{x} - \mathbf{v}^*\| \leq \|\mathbf{x} - \mathbf{v}\| \ \text{ pour tout } \mathbf{v} \in V \ .$$

Ce nombre : $\|\mathbf{x} - \mathbf{v}^*\|$ est appelé distance de $\mathbf{x}$ à V

De plus, la projection orthogonale de $\mathbf{x}$ sur V est l'unique élément $\mathbf{v}^* \in V$ tel que

$$(\mathbf{x} - \mathbf{v}^*) \in V^{\perp} \ .$$

Théorème 3-1.2 ***Projection sur un convexe fermé***
Soit E un espace euclidien, et soit $V \subset E$ un sous-ensemble convexe fermé dans E. Pour tout $\mathbf{x} \in E$ il existe un unique $\mathbf{v}^* \in V$ (appelé meilleure approximation de $\mathbf{x}$ sur V*) tel que*

$$\|\mathbf{x} - \mathbf{v}^*\| \leq \|\mathbf{x} - \mathbf{v}\| \ \text{ pour tout } \mathbf{v} \in V \ .$$

De plus, la meilleure approximation de $\mathbf{x}$ sur V est l'unique élément $\mathbf{v}^* \in V$ tel que

$$\mathcal{R}e \, \langle \mathbf{x} - \mathbf{v}^* \, , \, \mathbf{v} - \mathbf{v}^* \rangle \leq 0 \ \text{ pour tout } \mathbf{v} \in V \ .$$

La caractérisation de la projection orthogonale donnée dans le théorème 3-1.1 permet de calculer $\mathbf{v}^*$ lorsque $V \subset E$ est un sous-espace de dimension finie :

Corollaire 3-1.1 *Soit E un espace euclidien, et soit $V \subset E$ un sous-espace vectoriel de dimension finie dans E. Soit $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ une base de V, et soit $\mathbf{v}^* = \sum_{j=1}^{n} \alpha_j \mathbf{v}_j$ le développement de la projection orthogonale de $\mathbf{x} \in E$ sur V dans cette base de V.*

Les coefficients α_j, $j = 1, \ldots, n$, sont alors entièrement déterminés par la solution du système linéaire :

$$\begin{aligned}
&\mathbf{G}\boldsymbol{\alpha} = \mathbf{b} \ \text{ avec} \\
&\mathbf{G} = (g_{ij})_{1 \leq i,j \leq n} \ \text{ et } \ g_{ij} = \langle \mathbf{v}_j \,,\, \mathbf{v}_i \rangle \\
&\mathbf{b} = (b_i)_{1 \leq i \leq n} \ \text{ et } \ b_i = \langle \mathbf{x} \,,\, \mathbf{v}_i \rangle \ .
\end{aligned}$$

En outre, la matrice **G** *définie ci-dessus, appelée aussi* matrice de Gram associée à la base $\mathbf{v}_1, \ldots, \mathbf{v}_n$*, est hermitienne définie positive.*

On remarque aussi, au vu du résultat énoncé dans le corollaire pécédent, l'intérêt de disposer d'une base orthogonale de l'espace V puisque dans ce cas précis, la matrice de Gram précédente est diagonale.

3-2 Résolution de systèmes linéaires surdéterminés

Soient m et n deux entiers naturels tels que

$$m > n \ ,$$

$\mathbf{A}$ une matrice rectangulaire à m lignes et n colonnes, et $\mathbf{b}$ un vecteur de $\mathbb{R}^m$.
On se propose de résoudre le système rectangulaire suivant :

$$\mathcal{P} \left\{ \begin{array}{l} \text{Déterminer } \ \mathbf{x} \in \mathbb{R}^n \ \text{ solution de} \\ \qquad \mathbf{A}\mathbf{x} = \mathbf{b} \ . \end{array} \right.$$

$$\begin{bmatrix} \\ \quad \mathbf{A} \quad \\ \\ \end{bmatrix} \begin{bmatrix} \mathbf{x} \end{bmatrix} = \begin{bmatrix} \\ \mathbf{b} \\ \\ \end{bmatrix}$$

Sauf cas très particulier, ce problème n'a généralement pas de solution, car il contient plus d'équations que d'inconnues. On dira à ce propos que le système linéaire considéré est surdéterminé.

Rappelons un certains nombre de définitions et résultats.

Définition 3-2.1 Rang d'une matrice
On appelle **rang** *d'une matrice le plus grand nombre de vecteurs lignes ou de vecteurs colonnes linéairement indépendants que l'on peut extraire de cette matrice.*

Proposition 3-2.1 *Soit* $\mathbf{A}$ *une matrice carrée de taille* $m \times m$*. Le système linéaire* $\mathbf{A}\mathbf{x} = \mathbf{b}$
a une solution et une seule pour tout vecteur $\mathbf{b} \in \mathbb{R}^m$ *si et seulement si le rang de la matrice* $\mathbf{A}$ *est égal à* m*.*

Puisqu'on ne peut pas espérer, en général, trouver de solution au problème $\mathcal{P}$, on va alors considérer le problème suivant :

$$\widetilde{\mathcal{P}} \left\{ \begin{array}{l} \text{Déterminer } \ \tilde{\mathbf{x}} \in \mathbb{R}^n \ \text{ tel que} \\ \qquad \|\mathbf{A}\tilde{\mathbf{x}} - \mathbf{b}\|_2^2 \leq \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2^2 \ , \quad \forall \mathbf{x} \in \mathbb{R}^n \ . \end{array} \right.$$

Définition 3-2.2 *Soit* $\mathbf{R}(\mathbf{x}) = \mathbf{A}\mathbf{x} - \mathbf{b}$ *le* **résidu** *associé au système linéaire* $\mathbf{A}\mathbf{x} = \mathbf{b}$*.*
On dira que $\tilde{\mathbf{x}}$ *est* **pseudo-solution** *du problème* $\mathcal{P}$*, ou encore solution au sens des moindres carrés du problème* $\mathcal{P}$*, si* $\tilde{\mathbf{x}}$ *réalise le minimum de* $\|\mathbf{R}(\mathbf{x})\|_2^2 = \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2^2$*, c'est à dire si* $\tilde{\mathbf{x}}$ *est solution du problème* $\widetilde{\mathcal{P}}$*.*

Proposition 3-2.2 *Notons J la fonctionnelle de $\mathbb{R}^n$ dans $\mathbb{R}$ définie par :*

$$\forall \mathbf{x} \in \mathbb{R}^n \ , \quad J(\mathbf{x}) = \|\mathbf{R}(\mathbf{x})\|_2^2 = \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2^2 \ .$$

La fonctionnelle J est une fonctionnelle quadratique généralisée de terme quadratique semi-défini positif sur $\mathbb{R}^n$ et, si la matrice $\mathbf{A}$ est de rang maximal égal à n, alors le terme quadratique de la fonctionnelle J est défini positif sur $\mathbb{R}^n$.

Remarque : On peut vérifier facilement que le terme quadratique de la fonctionnelle $J(\mathbf{x})$ est associé à la matrice $\mathbf{A}^T\mathbf{A}$. Si la matrice $\mathbf{A}$ est de rang maximal, la matrice $\mathbf{A}^T\mathbf{A}$ est alors symétrique définie positive.

Proposition 3-2.3 *Soit $\mathbf{A}$ une matrice à m lignes et n colonnes, avec $m > n$, et soit $\mathbf{b}$ un vecteur de $\mathbb{R}^m$. Une condition nécessaire et suffisante pour que $\mathbf{x} \in \mathbb{R}^n$ réalise le minimum de $\|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2^2$ est que :*

$$\mathbf{A}^T\mathbf{A}\mathbf{x} = \mathbf{A}^T\mathbf{b} \ .$$

Ce système admet toujours au moins une solution. Si de plus la matrice $\mathbf{A}^T\mathbf{A}$ est régulière, c'est à dire que $\mathrm{rang}(\mathbf{A}) = n$, *alors la solution est unique.*

Remarque : Le théorème de projection orthogonale permet d'établir les résultats ci-dessus. En effet, minimiser $\|\mathbf{A}\mathbf{x}-\mathbf{b}\|_2^2$ revient à trouver le vecteur $\mathbf{y} = \mathbf{A}\mathbf{x}$ le plus proche du vecteur $\mathbf{b}$, c'est à dire à projeter orthogonalement $\mathbf{b}$ sur $\mathit{Im}\,(\mathbf{A})$ ...

L'existence d'une solution (indépendamment de toute hypothèse) est directement assurée par l'existence de cette projection orthogonale, et la caractérisation donnée dans la proposition est une conséquence (que vous pouvez vérifier en exercice).

3-3 Approximation d'une fonction au sens des moindres carrés

Le problème de l'approximation d'une fonction f sur un intervalle $I \subset \mathbb{R}$ est fondamentalement différent de celui de l'interpolation. Il consiste à remplacer la fonction f considérée par une autre fonction $\mathcal{P}(x)$ (en général plus régulière, et facile à manipuler) de sorte que la distance entre f et $\mathcal{P}$ soit aussi petite que possible.

On peut chercher par exemple un polynôme de bas degré, qui approche la fonction f en un sens à préciser sur l'intervalle I, ce qui diffère du problème d'interpolation qui consiste a trouver un polynôme de degré en général élevé qui coïncide au maximum avec la fonction f.

La notion de distance entre les fonctions f et $\mathcal{P}$ est bien évidemment fondamentale dans la définition du procédé d'approximation. On pourra par exemple distinguer :

1. **l'approximation au sens de la convergence uniforme**, où il s'agit de minimiser
$$\max_{x\in I} |f(x) - \mathcal{P}(x)| = \|f - \mathcal{P}\|_\infty ,$$
2. **l'approximation en moyenne quadratique** où il s'agit de minimiser la quantité
$$\int_I (f(x) - \mathcal{P}(x))^2 \, dx \;=\; \|f - \mathcal{P}\|^2_{\mathrm{L}^2(I)} ,$$
3. **l'approximation au sens des moindres carrés discrets**, utile lorsque f n'est connue que de manière discrète (c'est à dire sur un ensemble fini de points $x_i \in I$, $1 \le i \le m$); cette approximation consiste alors à minimiser la quantité
$$\sum_{i=1}^{m} (f(x_i) - \mathcal{P}(x_i))^2 .$$

3-3.1 Approximation en moyenne quadratique

On se propose d'approcher f sur l'intervalle $I = [a, b]$ par une fonction $\mathcal{P}(x)$, où $\mathcal{P}$ est une combinaison linéaire d'un ensemble de n fonctions données $u_j \in \mathrm{L}^2(I)$, $j = 1, \ldots, n$,
$$\mathcal{P}(x) = \sum_{j=1}^{n} \alpha_j u_j(x) ,$$
les coefficients α_j étant donc les inconnues à déterminer de façon à minimiser la quantité
$$J(\alpha_1, \ldots, \alpha_n) \;=\; \|f - \mathcal{P}\|^2_{\mathrm{L}^2(I)} \;=\; \int_a^b (f(x) - \sum_{j=1}^{n} \alpha_j u_j(x))^2 \, dx .$$

Remarque : Si on souhaite réaliser une approximation polynômiale de f, il suffira de choisir
$$u_j(x) = x^{j-1} , \quad j = 1, 2, \ldots, n .$$

Proposition 3-3.1 *Soit $\alpha \in \mathbb{R}^n$ le vecteur des coefficients α_j, $j = 1, \ldots, n$. Une condition nécessaire et suffisante pour que α réalise le minimum de la fonctionnelle $J(\alpha)$, est que α soit solution du système linéaire suivant*

$$\sum_{j=1}^{n} < u_j, u_i > \ \alpha_j \ = < f, u_i > \ , \quad i = 1, \ldots, n \ ,$$

avec

$$\begin{aligned} < u_j, u_i > &= \int_a^b u_j(x)\, u_i(x)\, dx \\ < f, u_i > &= \int_a^b f(x)\, u_i(x)\, dx \end{aligned}$$

3-3.2 Approximation au sens des moindres carrés discrets

Soit f une fonction dont on connait la valeur sur un sous ensemble fini de points $x_1, \ldots, x_m$. On se propose alors d'approcher la fonction f par une fonction $\mathcal{P}$ de telle façon que la quantité

$$\sum_{i=1}^{m} (f(x_i) - \mathcal{P}(x_i))^2 \ ,$$

soit minimale.

De plus, on cherche $\mathcal{P}$ sous la forme

$$\mathcal{P}(x) \ = \ \sum_{j=1}^{n} \alpha_j\, u_j(x); \, ,$$

où les u_j sont des fonctions connues et les $\{\alpha_j\}_{j=1}^n$ sont les inconnues à déterminer. On suppose de plus que $m > n$, pour ne pas être ramené à un problème d'interpolation. Ce problème correspond au *lissage d'une fonction f donnée par une combinaison linéaire de fonctions quelconques.*

Introduisons alors la matrice $\mathbf{A}$ à m lignes et n colonnes, ainsi que les vecteurs $\alpha \in \mathbb{R}^n$ et $\mathbf{F} \in \mathbb{R}^m$ définis par

$$\mathbf{A} = \begin{pmatrix} u_1(x_1) & \cdots & u_n(x_1) \\ u_1(x_2) & \cdots & u_n(x_2) \\ \vdots & & \vdots \\ u_1(x_m) & \cdots & u_n(x_m) \end{pmatrix} \ , \quad \alpha = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} \ , \quad \text{et} \ \ \mathbf{F} = \begin{pmatrix} f(x_1) \\ f(x_2) \\ \vdots \\ f(x_m) \end{pmatrix} \ .$$

En utilisant ces notations, on a alors le résultat suivant :

Proposition 3-3.2 *Il est facile de voir que le problème d'approximation décrit ci-dessus se ramène en fait à minimiser par rapport à $\alpha \in \mathbb{R}^n$ la quantité*

$$J(\alpha) = \|\mathbf{A}\alpha - \mathbf{F}\|_2^2 \ ,$$

et donc que le vecteur de coefficients $\alpha \in \mathbb{R}^n$ recherché correspond à la solution au sens des moindres carrés du système linéaire surdéterminé

$$\mathbf{A}\alpha = \mathbf{F} \ .$$

De plus, si les points x_i, $i = 1, \ldots, m$, et les fonctions u_j, $j = 1, \ldots, n$, sont choisis de telle façon que la matrice $\mathbf{A}$ soit de rang maximal, alors le problème d'approximation au sens des moindres carrés discrets précédent admet une solution unique $\alpha \in \mathbb{R}^n$, solution du système linéaire

$$\mathbf{A}^T\mathbf{A}\alpha = \mathbf{A}^T\mathbf{F} \ .$$

3-4 Minimisation de fonctionnelles quadratiques généralisées

Définition 3-4.1 Fonctionnelle quadratique généralisée
On appelle **Fonctionnelle quadratique généralisée** *toute application f de $\mathbb{R}^n$ dans $\mathbb{R}$ sous la forme :*

$$\forall \mathbf{x} \in \mathbb{R}^n \quad , \; f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T\mathbf{A}\mathbf{x} - \mathbf{b}^T\mathbf{x} + c \, ,$$

où $\mathbf{A}$ est une matrice de $\mathcal{M}_n(\mathbb{R})$, $\mathbf{b}$ un vecteur de $\mathbb{R}^n$, et c une constante réelle.

On appelle **terme quadratique** *associé à la fonctionnelle f le terme $\frac{1}{2}\mathbf{x}^T\mathbf{H}\mathbf{x}$.*

Remarque : On peut toujours se ramener au cas où la matrice $\mathbf{A}$ est symétrique, car on a :

$$\forall \mathbf{u} \in \mathbb{R}^n \; , \quad \mathbf{u}^T\mathbf{A}\mathbf{u} \; = \; \mathbf{u}^T\left(\frac{\mathbf{A}+\mathbf{A}^T}{2}\right)\mathbf{u} \; .$$

Définition 3-4.2 *Si la matrice $\mathbf{A}$ introduite dans la définition précédente est* **symétrique définie positive**, *alors la la fontionnelle quadratique généralisée est dite* **associée à une forme quadratique définie positive sur** $\mathbb{R}^n$.

Dans le cas des fonctionnelles quadratiques généralisées associées à une forme quadratique définie positive sur $\mathbb{R}^n$, on a le résultat suivant :

Proposition 3-4.1 *Soit $f : \mathbb{R}^n \to \mathbb{R}$ une fonctionnelle quadratique généralisée associée à une forme quadratique définie positive sur $\mathbb{R}^n$. Alors :*

1. *f admet un minimum global unique sur $\mathbb{R}^n$, noté $\hat{\mathbf{x}}$.*
2. *$\hat{\mathbf{x}}$ est l'unique solution du système linéaire $\mathbf{A}\mathbf{x} = \mathbf{b}$.*
3. *$\hat{\mathbf{x}}$ est le centre de symétrie d'un réseau d'ellipsoïdes homothétiques (E_α) definis par :*
$$E_\alpha = \{\mathbf{x} \in \mathbb{R}^n \; / \; f(\mathbf{x}) \leq \alpha\} \;\; , \;\; \forall \alpha \geq f(\hat{\mathbf{x}}) \; .$$
4. *Le minimum de f sur $\mathbb{R}^n$ vaut*
$$f(\hat{\mathbf{x}}) \;\; = \;\; -\frac{1}{2}\mathbf{b}^T\mathbf{A}^{-1}\mathbf{b} \; + \; c \; .$$
5. *Pour tout vecteur $\mathbf{x} \in \mathbb{R}^n$ on a :*
$$f(\mathbf{x}) \;\; = \;\; f(\hat{\mathbf{x}}) + \frac{1}{2}(\mathbf{x}-\hat{\mathbf{x}})^T\mathbf{A}(\mathbf{x}-\hat{\mathbf{x}}) \; .$$

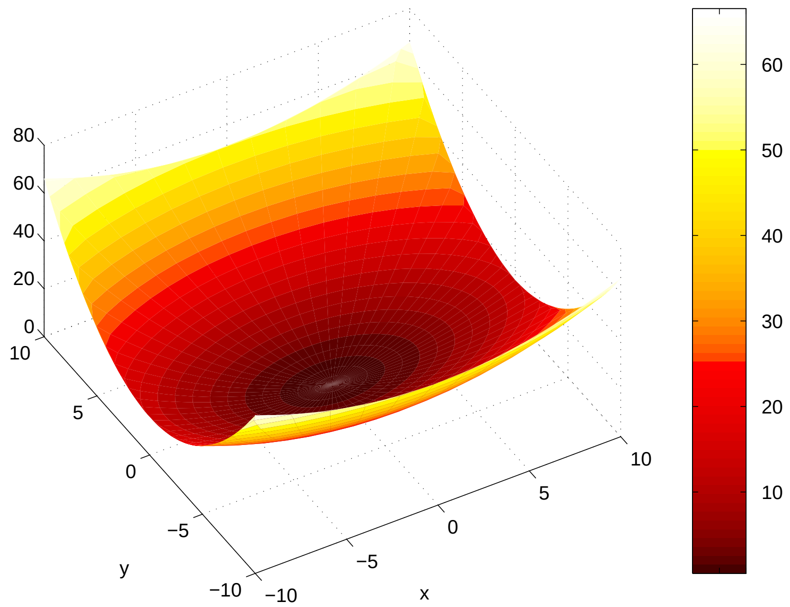

Figure 3-4.3: Un exemple de forme quadratique en dimension 2

La figure ci dessous illustre la forme géométrique d'une nappe quadratique, à savoir le dessin dans $\mathbb{R}^3$ d'une forme quadratique de $\mathbb{R}^2$ dans $\mathbb{R}$, où (x, y) jouent le rôle de $\mathbf{x} \in \mathbb{R}^2$ et $z = \frac{1}{2}\mathbf{x}^T\mathbf{A}\mathbf{x}$ (avec $\mathbf{A}$ matrice 2×2 symétrique définie positive).

Remarque : On peut noter qu'il y a une certaine dualité entre la recherche du minimum d'une fonctionnelle quadratique associée à une forme quadratique définie positive et la résolution d'un système linéaire associé à une matrice symétrique définie positive.

Remarque : La matrice $\mathbf{A}$ étant symétrique définie positive, elle définit une norme sur $\mathbb{R}^n$ par l'égalité

$$\|\mathbf{x}\|_{\mathbf{A}} \;=\; \sqrt{\mathbf{x}^T\mathbf{A}\mathbf{x}}\;.$$

Cette norme est dite ellipsoïdale, car les iso-contours $\{\|\mathbf{x}\|_{\mathbf{A}} = C^{ste}\}$ forment des ellipsoïdes dans $\mathbb{R}^n$ (c.f. figure 3-4.3). De plus, cette norme est associée au produit scalaire suivant :

$$(\mathbf{x}, \mathbf{y})_{\mathbf{A}} \;=\; \mathbf{x}^T\mathbf{A}\mathbf{y}\;.$$