

الجمهورية الجزائرية الديمقراطية الشعبية

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE

وزارة التعليم العالي والبحث العلمي

MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE  
SCIENTIFIQUE

جامعة عمارة تليدجي بالأغواط

UNIVERSITE AMAR TELIDJI LAGHOUAT



كلية العلوم

FACULTE DES SCIENCES

DEPARTEMENT DE MATHEMATIQUE ET INFORMATIQUE

Mémoire de MASTER

Domaine : Mathématique informatique (MI)

Filière : Mathématique

Option : Analyse mathématique

Présenté par :

TIMMAOUI Ahmed

**THEME**

---

**Méthodes itératives pour les problèmes inverses linéaires**

---

Soutenu publiquement devant le jury composé de :

Encadreur : NOURI Brahim

Co-Encadreur: BENABDERRAHMANE Benyattou

Président : ALLAOUI Salaheddine

Examineur : BOUHKATEM Yamna

Examineur : RAHMOUNE Abita

Année universitaire 2012-2013



---

# Remerciement

Ce travail a été réalisé au sein du laboratoire d'Informatique et de Mathématiques (LIM) à l'université de Laghouat sous la direction de Monsieur NOUIRI Brahim, Maître de conférences à l'université de Laghouat. Je tien à le remercier pour leur disponibilité leurs conseils et pour avoir guidé ce travail avec beaucoup d'intérêt.

Mes remerciements s'adressent a Monsieur le Professeur BENABDERRAHMANE Benyattou ; Co-encadreur et chef d'équipe de Mathématiques Appliquées d'avoir disponibilité tous les moyens de laboratoire durant la période de la préparation de ce mémoire.

Je remercie les membres du Jury pour leurs acceptations d'examiner ce mémoire.

Mes derniers et profonds remerciements vont à mes chers parents à qui je dédie ce travail ainsi qu'à toute ma famille et mes amis pour leur grand soutien.

Aussi, je remercie tout mes collègues et aux qui m'ont aidé de près ou de loin en vue de réaliser ce mémoire.

---

# Table des matières

<b>Introduction générale</b>	<b>1</b>
<b>1 Espaces de Hilbert et opérateurs linéaires continus</b>	<b>4</b>
1.1 Définitions et exemples . . . . .	4
1.2 Opérateurs linéaires dans un espaces de Hilbert . . . . .	12
1.3 Opérateurs compacts . . . . .	15
1.4 Décomposition spectrale des opérateurs auto-adjoints compacts . . . . .	17
<b>2 Opérateurs intégraux, équations intégrales et Méthode de Moindres Carrés linéaire</b>	<b>19</b>
2.1 Opérateurs intégraux . . . . .	19
2.2 Equations intégrales . . . . .	21
2.2.1 Discrétisation des équations intégrales . . . . .	22
2.2.1.1 Discrétisation par quadrature . . . . .	22
2.2.1.2 Discrétisation par la méthode de Galerkin . . . . .	23
2.3 Propriétés mathématiques des problèmes de moindres carrés . . . . .	25
2.3.0.3 Cas de la dimension finie . . . . .	28
2.4 Décomposition en valeurs singulières de matrices . . . . .	29
2.4.0.4 Applications de SVD aux problèmes de moindres carrés . . . . .	32
2.4.1 Développement en valeur singulières des opérateurs compacts . . . . .	33
2.4.2 Applications de SVD aux problèmes de moindres carrés . . . . .	35

<b>3 Méthodes itératives</b>	<b>37</b>
3.1 Méthode de Landweber . . . . .	37
3.2 Méthode du gradient conjugué . . . . .	41
3.2.1 Méthode des directions conjuguées . . . . .	41
3.2.2 Propriétés de base de la méthode du gradient conjugué . . . . .	43
3.2.3 Vitesse de convergence . . . . .	47
3.2.4 Méthode du gradient conjugué pour l'équation normale . . . . .	50
<b>Conclusion générale</b>	<b>51</b>
<b>Bibliographie</b>	<b>52</b>

---

# Introduction générale

Un problème inverse est une situation dans laquelle on tente de déterminer les paramètres d'un modèle à partir de mesures expérimentales.

En mathématiques, un problème inverse linéaire a la forme d'une équation

$$Ax = y \tag{1}$$

Où  $y$  représente les mesures effectuées,  $x$  représente les valeurs des paramètres du phénomène et  $A$  est un opérateur linéaire, d'un espace de Hilbert  $H$  dans  $H$  ; qui représente la relation entre les mesures et les paramètres du modèle.

Les problèmes inverses linéaires généralement sont des problèmes mal posés car si l'on cherche à résoudre l'équation (1) ; cela nécessite l'inversion de l'opérateur  $A$ . Cette opération n'est pas forcément évidente d'un point de vue numérique. Et d'après Hadamard [9] un problème est bien posé s'il vérifie les trois conditions suivantes :

- La solution existe ;
- Elle est unique ;
- Elle dépend continument des données.

Donc, si l'une des trois conditions n'est pas satisfaite, on dit que le problème est mal posé.

La résolution du problème inverse passe en général par une étape initiale de modélisation du phénomène, dite problème direct qui décrit comment les paramètres du modèle se traduisent en effets observables expérimentalement. Ensuite, à partir des mesures obtenues sur le phénomène réel, la démarche va consister à approximer au mieux les paramètres qui permettent de rendre compte de ces mesures. Cette résolution peut se faire par *simulation numérique* ou de façon analytique.

Dans ce mémoire, nous avons examiné deux méthodes itératives pour les problèmes inverses linéaires :

La première est la méthode de Landweber [12], qu'est construite une suite de solutions approchées converge vers la solution désirée (dans le cas non bruité). Dans le contexte des problèmes inverses linéaires en présence de bruit, la suite construite ne converge pas, en général, vers une solution du problème de départ. Il est, encore une fois, nécessaire de régulariser le processus itératif, et c'est l'indice d'itération lui-même qui joue le rôle de paramètre de régularisation. En d'autres termes, il convient d'arrêter les itérations plus tôt qu'on ne le ferait dans un cas non bruité. Malheureusement, la méthode de Landweber converge trop lentement.

La deuxième est la méthode du gradient conjugué et ses variantes qui est la plus employée en pratique [13].

Ce mémoire se décompose en trois chapitres de la manière suivante :

Dans le premier chapitre, nous allons commencer par donner un rappel sur les espaces de Hilbert et leurs propriétés, ainsi que quelques résultats indispensables sur les opérateurs linéaires continus et les propriétés les plus importantes dans les espaces de Hilbert.

Dans le second chapitre, nous introduisons une source importante de problèmes inverses linéaires : les équation de première espèce. Après avoir les principales propriétés des opérateurs intégraux, nous expliquerons en quoi ils sont mal posés. Enfin, nous introduirons des méthodes de discrétisation, conduisons à des problèmes de moindres carrés, nous étudierons leur propriétés mathématiques, dans un cadre Hilbertien : l'aspect géométrique, et le lien avec les équations normales, ainsi que les questions d'existence et d'unicité des solutions. Nous introduirons également l'outil fondamental, tant pour l'analyse théorique que pour l'approximation numérique, qu'est la décomposition en valeurs singulières pour les opérateurs entre espaces de Hilbert.

Dans le dernier chapitre, nous introduisons deux méthodes itératives pour la simulation numérique des problèmes inverses linéaires : la méthode de Landweber et la méthode du gradient conjugué.

Ce mémoire se termine par une conclusion et quelques perspectives.

## NOTATIONS

- $(x, y)$  : Produit scalaire de  $x$  et  $y$ .
- $\|x\|$  : Norme de  $x$ .
- $X_1 \oplus X_2$  : Somme directe de  $X_1$  et  $X_2$ .
- $(E, \|\cdot\|)$  : Espace normé.
- $x \perp y$  :  $x$  et  $y$  sont orthogonaux.
- $A_1 \perp A_2$  :  $A_1$  est orthogonal à  $A_2$ .
- $A^\perp$  : Complémentaire orthogonale de  $A$ .
- $L^2(\Omega)$  : L'espace vectoriel des fonctions de carré intégrable sur  $\Omega$ ,

$$L^2(\Omega) = \left\{ f : \Omega \longrightarrow \mathbb{R}, \int_{\Omega} |f(x)|^2 dx < \infty \right\}$$

- $\text{Ker}T$  : Noyau de  $T$ ,  $\text{Ker}T = \{x \in H_1, Tx = 0\}$ .
- $\text{Im}T$  : Image de  $T$ ,  $\text{Im}T = \{y \in H_2, \exists x \in H_1, Tx = y\}$ .
- $\dim T$  : Dimension de  $T$ ,
- $T^*$  : L'adjoint de  $T$ .
- $T^t$  : Le transposé de  $T$ .
- $\bar{A}$  : La fermeture de  $A$ .
- $\bar{B}_E$  : La boule unité fermée d'un espace normé  $E$ .
- $\mathcal{L}(E, F)$  : L'ensemble des applications linéaires continus de  $E$  dans  $F$ .
- $\mathcal{L}(E)$  : L'ensemble des applications linéaires continus de  $E$  dans  $E$ .
- $K(E, F)$  : L'ensemble des opérateurs compacts de  $E$  dans  $F$ .
- $K(E)$  : L'ensemble des opérateurs compacts de  $E$  dans  $E$ .
- $\sigma(T)$  : Le spectre de  $T$ ,

$$\sigma(T) = \{\lambda \in \mathbb{C}, T - \lambda I \text{ n'est pas inversible dans } \mathcal{L}(H)\}.$$

- $\sigma_p(T)$  : Le spectre ponctuel i.e. l'ensemble des valeurs propres de  $T$ .
- $\text{vect}(e_1, e_2, \dots, e_n)$  : L'espace vectoriel engendré par les vecteurs  $e_1, e_2, \dots, e_n$ .
- $\text{span}(e_1, e_2, \dots, e_n)$  : L'espace vectoriel engendré par les vecteurs  $e_1, e_2, \dots, e_n$ .

---

---

# Chapitre 1

---

## Espaces de Hilbert et opérateurs linéaires continues

Nous donnerons dans ce chapitre les principaux résultats d'analyse fonctionnelle dont nous aurons besoin, ainsi que des compléments concernant les opérateurs dans les espaces de Hilbert.

### 1.1 Définitions et exemples

**Définition 1.1 (Semi-norme)** Soit  $E$  un espace vectoriel sur  $\mathbb{R}$ . On appelle semi-norme pour  $E$  une application

$P : E \longrightarrow \mathbb{R}$ , telle que  $\forall x, y \in E$  et  $\lambda \in \mathbb{R}$  on a

- $P(x) \geq 0$ , propriété de non-négativité.
- $P(\lambda x) = |\lambda| P(x)$ , propriété d'homogénéité absolue.
- $P(x + y) \leq P(x) + P(y)$ , propriété de sous-additivité.

L'espace vectoriel  $E$  muni d'une semi-norme s'appelle espace semi-normé, noté par  $(E, P)$

**Remarque 1.1** Pour une semi-norme, on a  $P(0) = 0$ .

**En effet.**

$$P(0) = P(0 \cdot x) = 0 \cdot P(x) = 0$$

A noter qu'il est possible que  $P(x) = 0$  et  $x \neq 0$ .

**Définition 1.2 (Convergence)** Si  $(E, P)$  est un espace vectoriel semi-normé, alors on dit qu'une suite  $(x_n)_{n \in \mathbb{N}}$  de  $E$  converge vers  $x \in E$  (en abrégé  $x_n \longrightarrow x$ ) si

$$\lim_{n \rightarrow \infty} P(x_n - x) = 0.$$

**Définition 1.3 (Ensemble fermé)** Soit  $S$  un sous-ensemble de  $E$ , alors la fermeture de  $S$  dans  $E$  pour la semi-norme  $P$  est

$$\bar{S} = \{x \in E : \exists x_n \in S, \forall n \in \mathbb{N} \text{ tels que } x_n \longrightarrow x \text{ dans } E\}.$$

$S$  est dit fermé si  $S = \bar{S}$

A noter que  $\bar{S}$  est toujours fermé.

**Définition 1.4 (Suite de Cauchy)** Soit  $(E, P)$  un espace vectoriel semi-normé. Une suite  $(x_n)_{n \in \mathbb{N}}$  de  $E$  est de Cauchy si et seulement si

$$\forall \epsilon > 0, \exists N > 0 \text{ tel que } P(x_n - x_m) < \epsilon, \forall n, m \geq N.$$

**Définition 1.5 (Espace complet)** Soit  $(E, P)$  un espace vectoriel semi-normé. On dit que  $(E, P)$  est complet si toute suite de Cauchy de  $E$  est convergente.

**Définition 1.6 (Norme)** Soit  $E$  un espace vectoriel sur  $\mathbb{R}$ . On appelle norme sur  $E$  une application  $\|\cdot\| : E \longrightarrow \mathbb{R}$ , telle que  $\|\cdot\|$  est une semi-norme pour  $E$  et de plus

$$\|x\| = 0 \Leftrightarrow x = 0$$

Un espace vectoriel  $E$  muni d'une norme s'appelle espace vectoriel normé, noté par  $(E, \|\cdot\|)$ .

**Exemple 1.1** Dans l'espace vectoriel réel  $\mathbb{R}$ , l'application  $\|\cdot\| : \mathbb{R} \longrightarrow \mathbb{R}$  donnée par  $\|x\| = |x|, \forall x \in \mathbb{R}$  où  $|x|$  est la valeur absolue de  $x$ , est une norme pour  $\mathbb{R}$ .

**Exemple 1.2** Soit  $\Omega$  un ouvert de  $\mathbb{R}^n$ . L'espace vectoriel des fonctions de carré intégrable sur  $\Omega$  est :

$$L^2(\Omega) = \left\{ f : \Omega \longrightarrow \mathbb{R}, \int_{\Omega} |f(x)|^2 dx < \infty \right\}$$

l'application

$$\|\cdot\| : L^2(\Omega) \longrightarrow \mathbb{R}$$

donné par

$$\|f\| = \left( \int_{\Omega} |f(x)|^2 dx \right)^{\frac{1}{2}}$$

est une norme pour  $L^2(\Omega)$ .

**Définition 1.7 (Espace de Banach)** *Un espace vectoriel normé complet est appelé espace de **Banach**.*

**Définition 1.8 (Produit scalaire)** *Soit  $E$  un espace vectoriel sur le corps  $IR$ . Un produit scalaire sur  $E$  est une application  $\varphi : E \times E \rightarrow IR$  telle que,  $\forall x, x_1, x_2, y \in E$ , et  $\lambda \in IR$ , on a :*

1.  $\varphi(x, x) \geq 0$  et  $\varphi(x, x) = 0 \Leftrightarrow x = 0$ .
2.  $\varphi(x, y) = \varphi(y, x)$ .
3.  $\varphi(x_1 + x_2, y) = \varphi(x_1, y) + \varphi(x_2, y)$ .
4.  $\varphi(\lambda x, y) = \lambda\varphi(x, y)$ .

**Notation :** *On note  $\varphi(x, y)$  par  $(x, y)$ .*

**Définition 1.9 (Espace préhilbertien)** *Un espace vectoriel muni d'un produit scalaire est appelé un espace **préhilbertien**.*

**Exemple 1.3** *Sur  $IR^n$ , le produit scalaire euclidien usuel est*

$$(x, y) = \sum_{i=1}^n x_i y_i.$$

**Exemple 1.4**  *$L^2(\Omega)$  est un espace préhilbertien si on le munit du produit scalaire*

$$(f, g) = \int_{\Omega} f(x)g(x)dx$$

**Remarque 1.2** *Un produit scalaire sur  $E$  définit une norme sur  $E$  par la formule suivante*

$$\|x\| = (x, x)^{\frac{1}{2}}$$

**Lemme 1.1 (Inégalité de Cauchy-Schwarz)** *Soit  $E$  un espace préhilbertien.  $\forall x, y \in E$ ,*

$$|(x, y)| \leq (x, x)^{1/2}(y, y)^{1/2} \tag{1.1}$$

**Preuve.** L'inégalité (1.1) est trivialement satisfaite si  $(x, y) = 0$ . Nous supposons donc que  $(x, y) \neq 0$ . Alors nous obtenons

$$\begin{aligned}
 0 &\leq \left( \left[ \frac{x}{(x, x)^{\frac{1}{2}}} - \frac{(x, y)}{|(x, y)|} \frac{y}{(y, y)^{\frac{1}{2}}} \right], \left[ \frac{x}{(x, x)^{\frac{1}{2}}} - \frac{(x, y)}{|(x, y)|} \frac{y}{(y, y)^{\frac{1}{2}}} \right] \right) \\
 &= 1 - \frac{2|(x, y)|}{(x, x)^{\frac{1}{2}}(y, y)^{\frac{1}{2}}} + 1 \\
 &= 2 - \frac{2|(x, y)|}{(x, x)^{\frac{1}{2}}(y, y)^{\frac{1}{2}}} \\
 &\Leftrightarrow \frac{|(x, y)|}{(x, x)^{\frac{1}{2}}(y, y)^{\frac{1}{2}}} \leq 1 \\
 &\Leftrightarrow |(x, y)| \leq (x, x)^{\frac{1}{2}}(y, y)^{\frac{1}{2}}
 \end{aligned}$$

**Lemme 1.2 (Loi du parallélogramme)** *La norme induite par un produit scalaire satisfait l'égalité*

$$\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2) \quad (1.2)$$

**Preuve.** Nous avons

$$\begin{aligned}
 \|x + y\|^2 + \|x - y\|^2 &= (x + y, x + y) + (x - y, x - y) \\
 &= (x, x) + (x, y) + (y, x) + (y, y) + (x, x) - (x, y) - (y, x) + (y, y) \\
 &= 2(\|x\|^2 + \|y\|^2)
 \end{aligned}$$

**Définition 1.10 (Espace de Hilbert)** *Un espace de Hilbert est un espace complet par rapport à la norme induite par un produit scalaire.*

**Exemple 1.5**  $L^2(\Omega)$  est un espace de Hilbert pour la produit scalaire défini en l'exemple 1.4

**Définition 1.11 (Ensemble convexe)** *Un ensemble  $U \in E$  (espace vectoriel), est convexe si*

$$\forall x, y \in U, \forall \lambda \in [0, 1] \text{ on a } \lambda x + (1 - \lambda)y \in U.$$

**Théorème 1.1 (Projection)** Soient  $H$  un espace de Hilbert,  $A \in H$  un ensemble non vide convexe et fermé et  $x_0 \in H - A$ . Alors il existe un et seulement un  $y_0 \in A$ , tel que

$$\|x_0 - y_0\| = \inf_{y \in A} \|x_0 - y\| \quad (1.3)$$

De plus  $y_0$  est caractérisé par la propriété :  $y_0 \in A$

$$(x_0 - y_0, y - y_0) \leq 0 \quad \forall y \in A \quad (1.4)$$

**Preuve.** a) Montrons d'abord l'existence de  $y_0$ . Soit  $d = \inf_{y \in A} \|x_0 - y\| > 0$ .

Choisissons une suite  $\{y_n; \|x_0 - y_n\| \leq d + 1/n; n \in \mathbb{N}\} \subset A$ .

D'après la loi du parallélogramme, nous obtenons

$$\begin{aligned} \|y_m - y_n\|^2 &= \|(y_m - x_0) - (y_n - x_0)\|^2 = 2\|y_m - x_0\|^2 + 2\|y_n - x_0\|^2 - 4\left\|\frac{y_m + y_n}{2} - x_0\right\|^2 \\ &\leq 2(d + 1/m)^2 + 2(d + 1/n)^2 - 4d^2 \\ &= \frac{4}{m}d + \frac{2}{m^2} + \frac{4}{n}d + \frac{2}{n^2} < \epsilon \text{ si } m, n > n_0(\epsilon). \end{aligned}$$

Puisque  $A$  est un ensemble convexe,  $\frac{y_m + y_n}{2} \in A$  et  $\left\|\frac{y_m + y_n}{2} - x_0\right\|^2 \geq d^2$ . Donc  $\{y_n; n \in \mathbb{N}\} \subset A$  est une suite de Cauchy et il existe un  $y_0 \in H$  tel que  $y_n \rightarrow y_0$ . Parce que  $A$  est fermé,  $y_0 \in A$ .

b) Montrons que  $\|x_0 - y_0\| = d$ . Nous avons :

$$d \leq \|x_0 - y_0\| \leq \|x_0 - y_n\| + \|y_n - y_0\| < d + \frac{1}{n} + \epsilon \text{ si } n > n_0(\epsilon), \text{ d'où } \|x_0 - y_0\| = d.$$

c) Il nous reste à montrer l'unicité de  $y_0$ . Soit  $z_0 \in A$ , tel que  $\|z_0 - x_0\| = d$ .

En utilisant la loi du parallélogramme, on à

$$\begin{aligned} \|z_0 - y_0\|^2 &= \|(z_0 - x_0) - (y_0 - x_0)\|^2 \\ &= 2\|z_0 - x_0\|^2 + 2\|y_0 - x_0\|^2 - \|z_0 + y_0 - 2x_0\|^2 \\ &= 4d^2 - 4\left\|\frac{z_0 + y_0}{2} - x_0\right\|^2 \\ &\leq 4d^2 - 4d^2 \\ &= 0. \end{aligned}$$

d'où  $z_0 = y_0$ .

**On montre l'équivalence de (1.3) et (1.4).** Soit  $y_0 \in A$  vérifiant (1.3) et soit  $z \in A$ .

On a  $y = (1 - t)y_0 + tz \in A$  pour  $t \in ]0, 1]$

et donc

$$\begin{aligned} \|x_0 - y_0\| &\leq \|x_0 - [(1 - t)y_0 + tz]\| \\ &= \|(x_0 - y_0) - t(z - y_0)\| \end{aligned}$$

Donc

$$\begin{aligned} \|x_0 - y_0\|^2 &\leq \|(x_0 - y_0) - t(z - y_0)\|^2 \\ &= ((x_0 - y_0) - t(z - y_0), (x_0 - y_0) - t(z - y_0)) \\ &= (x_0 - y_0, x_0 - y_0) - 2t(x_0 - y_0, z - y_0) + (t(z - y_0), t(z - y_0)) \\ &= \|x_0 - y_0\|^2 - 2t(x_0 - y_0, z - y_0) + t^2 \|z - y_0\|^2 \end{aligned}$$

Alors  $2(x_0 - y_0, z - y_0) \leq t \|z - y_0\|^2$

Quand  $t \rightarrow 0$  on obtient (1.4)

**Inversement**, soit  $y_0$  vérifiant (1.4), Alors  $\forall y \in A$  Nous avons :

$\forall y \in A$

$$\begin{aligned} \|x_0 - y_0\|^2 - \|x_0 - y\|^2 &= (x_0 - y_0, x_0 - y_0) - (x_0 - y, x_0 - y) \\ &= (x_0, x_0) + (y_0, y_0) - 2(x_0, y_0) - (x_0, x_0) - (y, y) + 2(x_0, y) \\ &= 2(x_0, y - y_0) + (y_0, y_0) - (y, y) + 2(y_0, y - y_0) - 2(y_0, y - y_0) \\ &= 2(x_0 - y_0, y - y_0) + (y_0, y_0) - (y, y) + 2(y_0, y) - 2(y_0, y_0) \\ &= 2(x_0 - y_0, y - y_0) - (y_0, y_0) - (y, y) + 2(y_0, y) \\ &= 2(x_0 - y_0, y - y_0) - (y_0 - y, y_0 - y) \\ &= 2(x_0 - y_0, y - y_0) - \|y_0 - y\|^2 \leq 0. \end{aligned}$$

d'où (1.3)

**Corollaire 1.1** Soit  $A \subset H$  un sous espace vectoriel fermé, et soit  $x_0 \in H$  Alors, la projection de  $x_0$  sur  $A$  est caractérisée par :

$$y_0 \in A \text{ et } (x_0 - y_0, y) = 0, \forall y \in A \quad (1.5)$$

**Preuve.** D'après (1.4) on a

$$(x_0 - y_0, y - y_0) \leq 0 \quad \forall y \in A$$

et donc

$$(x_0 - y_0, ty - y_0) \leq 0 \quad \forall y \in A, \forall t \in \mathbb{R}$$

Il en résulte que

$$(x_0 - y_0, y) = 0 \quad \forall y \in A$$

**Inversement** si  $y_0$  vérifie (1.5) on a :

$$(x_0 - y_0, y - y_0) = 0 \quad \forall y \in A$$

**Définition 1.12** Dans un espace de Hilbert, on dit que deux vecteurs  $x, y \in H$  sont orthogonaux si  $(x, y) = 0$ . On note  $x \perp y$ .

L'orthogonal d'un sous espace vectoriel  $A$  et

$$A^\perp = \{x \in H, (x, y) = 0, \forall y \in A\}$$

$A^\perp$  s'appelle le complément orthogonal de  $A$ .

**Remarque 1.3**  $A^\perp$  est un sous espace vectoriel fermé de  $H$ .

**Théorème 1.2 (Projection orthogonale)** Soit  $H_1$  un sous espace fermé d'un espace de Hilbert  $H$ , et soit  $P$  la projection orthogonale de  $H$  sur  $H_1$ . Alors

$$\text{Ker}P = \{x \in H; Px = 0\} = H_1^\perp$$

De plus  $(I - P)$  est l'opérateur de la projection orthogonale de  $H$  sur  $H_1^\perp$ . L'application  $I$  signifie l'identité, i.e.,  $Ix = x \quad \forall x \in H$ .

**Preuve.**

1. Soit  $x \in H_1^\perp$ , c-à-d.  $(x, y) = 0 \quad \forall y \in H_1$ .  
D'autre part  $(x - Px, y) = 0$ , et donc  $(Px, y) = 0, \forall y \in H_1$ . Pour  $y = Px, \|Px\|^2 = 0$ , c-à-d.  $x \in \text{Ker}P$ .
2. Si  $x \in \text{Ker}P$ , alors  $\forall y \in H_1, 0 = (x - Px, y) = (x, y)$ . Donc  $x \perp H_1$ .

3. Parceque  $H_1^\perp$  est un sous espace fermé de  $H$ , la projection orthogonale  $P_1$  de  $H$  sur  $H_1^\perp$  existe. Soit  $y \in H_1^\perp$ . Alors  $(x - (I - P)x, y) = (Px, y) = 0, \forall x \in H$ , ce qui montre que  $P_1 = I - P$ .

**Théorème 1.3 (Décomposition orthogonale)** *Si  $H_1$  est un sous espace fermé d'un espace de Hilbert  $H$ , alors tout  $x \in H$  se décompose d'une manière unique ;*

$$x = y + z, \quad y \in H_1, \quad z \in H_1^\perp.$$

**Preuve.** L'existence d'une telle décomposition vient du fait que

$$x = Px + (I - P)x,$$

où  $P$  est la projection orthogonale de  $H$  sur  $H_1$ . Supposons maintenant que  $x = y + z, y \in H_1$ . Alors  $Px = Py - Pz = y$  et  $(I - P)x = (I - P)y + (1 - P)z = z$  cette décomposition est donc unique.

L'écriture précédent est correcte car  $P$  est une application linéaire de  $H$  sur  $H_1$ ,  
**en effet.**

Pour tout  $x_1, x_2 \in H, \alpha, \beta \in IR, y \in H_1 (x - Px, y) = 0$ .

Nous avons d'une part

$$(\alpha x_1 - \alpha Px_1 + \beta x_2 - \beta Px_2, y) = (\alpha x_1 + \beta x_2 - (\alpha Px_1 + \beta Px_2), y) = 0,$$

et D'autre part  $(\alpha x_1 + \beta x_2 - P(\alpha x_1 + \beta x_2), y) = 0$

L'unicité de la projection orthogonale de  $\alpha x_1 + \beta x_2$  entraîne que  $P(\alpha x_1 + \beta x_2) = \alpha Px_1 + \beta Px_2$ .

**Corollaire 1.2** *Soit  $H_1$  un sous espace fermé d'un espace de Hilbert  $H$ .*

*Alors  $H = H_1 \oplus H_1^\perp$  c-à-d.  $H$  admet une décomposition orthogonale.*

**Preuve.** Si  $x = Px + (I - P)x \in H_1 \cap H_1^\perp$ , alors  $(x, Px) = (x, (I - P)x)$ , ce qui entraîne que  $(x, Px + (I - P)x) = (x, x) = \|x\|^2 = 0$ , donc  $x = 0$ ,

utilisant le théorème 1.3, on a  $H = H_1 \oplus H_1^\perp$ .

**Définition 1.13 (Ensemble dense)** *Une partie  $G$  de  $H$  est dite dense dans  $H$*

*si  $\forall h \in H, \forall \epsilon > 0, \exists g \in G; \|g - h\| < \epsilon$ .*

*ou de manière équivalente si tout  $h$  de  $H$  est limite d'une suite d'élément  $g_n$  de  $G$  :*

$$\|g_n - h\| \longrightarrow 0.$$

## 1.1.2 Opérateurs linéaires dans un espaces de Hilbert

Crétoire de densité :

$$F \subset H \text{ est dense dans } H \iff F^\perp = 0.$$

**Définition 1.14 (Base Hilbertienne)** On appelle base Hilbertienne de  $H$ , une suite  $(e_n)_{IN}$  telle que

- $\forall n \in IN, \|e_n\| = 1.$
- $\forall n, m \in IN, (e_n, e_m) = \delta_{nm} = \begin{cases} 1 & \text{si } n = m \\ 0 & \text{si } n \neq m \end{cases}$
- L'espace vectoriel engendré par les  $e_n$  est dense dans  $H$ .

**Théorème 1.4** (Voir H.Brezis)

Soit  $(e_n)_{IN}$  une base Hilbertienne de  $H$  et  $x \in H$ . Alors

- $x = \sum_{n \in IN} x_n e_n$  avec  $x_n = (x, e_n).$
- $\|x\|^2 = \sum_{n \in IN} \|x_n\|^2.$

**Définition 1.15 (Espace de Hilbert séparable)** Un espace de Hilbert est séparable s'il possède un sous ensemble dénombrable dense.

**Théorème 1.5** (Voir H.Brezis)

Toute espace de Hilbert séparable admet une base Hilbertienne.

**Théorème 1.6 (Inégalité de Bessel)** (voir W. Hengartner)

Soit  $E = \{e_i; i \in I\}$  un système orthonormal dans un espace de Hilbert  $H$  et soit, pour  $x \in H$ ,  $c_i(x) = (x, e_i), i \in I$ . Alors

$$\sum_{i \in I} |c_i(x)|^2 \leq \|x\|^2. \quad (1.6)$$

## 1.2 Opérateurs linéaires dans un espaces de Hilbert

**Définition 1.16 (Opérateur linéaire continu)** Soit  $H_1, H_2$  deux espace de Hilbert. On appelle opérateur linéaire continu de  $H_1$  dans  $H_2$  une application linéaire continue  $T$  de  $H_1$  dans  $H_2$ , c-à-d. elle vérifie :

- $\forall (x, y) \in H_1 \times H_2, \forall (\alpha, \beta) \in IR^2, T(\alpha x + \beta y) = \alpha T x + \beta T y.$
- $\exists M > 0, \forall x \in H_1, \|T x\|_{H_2} \leq M \|x\|_{H_1}.$

## 1.1.2 Opérateurs linéaires dans un espaces de Hilbert

---

On défini la norme de l'opérateur  $T$  par

$$\|T\| = \sup_{x \in H_1} \frac{\|Tx\|_{H_2}}{\|x\|_{H_1}}.$$

**Définition 1.17 (Noyau et image de  $T$ )** – Le noyau de  $T$  est le sous espace de  $H_1$  défini par :  $\text{Ker}T = \{x \in H_1, Tx = 0\}$ .

– L'image de  $T$  est le sous espace de  $H_2$  défini par :  $\text{Im}T = \{y \in H_2, \exists x \in H_1, Tx = y\}$ .

**Théorème 1.7 (De l'application ouvert)** (Voir H.Brezis)

Soit  $E, F$  deux espace de Banach et  $T \in \mathcal{L}(E, F)$  (i.e  $T$  linéaire et continu)

Si  $T$  est surjective, alors l'image par  $T$  d'un ouvert de  $E$  est un ouvert de  $F$ .

On dit que  $T$  est ouverte.

**Corollaire 1.3** Soit  $E, F$  deux espace de Hilbert  $T \in \mathcal{L}(E, F)$  bijective, Alors  $T^{-1}$  est continu.

**Définition 1.18 (Dual de  $H$ )** Soit  $H$  un espace de Hilbert. On note  $H'$  l'espace de formes linéaires continues sur  $H$ , c'est-à-dire, l'espace  $\mathcal{L}(H, \mathbb{R})$ . On appelle  $H'$  l'espace dual de  $H$ .

**Théorème 1.8 (Riesz)** Pour tout élément  $f \in H'$  il existe un unique vecteur  $x_0 \in H$  tel que :

$$f(x) = (x, x_0) \quad \forall x \in H.$$

**Preuve.** Soit  $G = \text{Ker}f$ , Alors  $G$  est un sous espace fermé de  $H$ .

– Si  $G = H$ , alors  $f \equiv 0$  et on pose  $x_0 = 0$ .

– Si  $G \neq H$ , Soit  $y_0 \in H - G$ ,  $y_0 \neq 0$ .

Pour  $x \in H$ , on a  $f(x)y_0 - f(y_0)x \in G$ , et donc

$$(f(x)y_0 - f(y_0)x, y_0) = 0, \text{ d'où } (f(x)y_0, y_0) = (f(y_0)x, y_0) = (x, \overline{f(y_0)y_0}).$$

Posons  $x_0 = \frac{f(y_0)y_0}{\|y_0\|^2}$ , Alors  $f(x) = (x, x_0)$ .

$x_0$  est unique, car si  $x_1$  et  $x_2$  vérifiant  $f(x) = (x, x_1) = (x, x_2)$ , alors  $\forall x \in H \quad (x, x_1 - x_2) = 0$  et donc  $x_1 = x_2$ .

### Adjoint d'un opérateur

**Proposition 1.1** Soient  $H_1$  et  $H_2$  des espace de Hilbert et  $T \in \mathcal{L}(H_1, H_2)$ . Alors il existe un unique  $T^* \in \mathcal{L}(H_2, H_1)$  tel que,

$$\forall x \in H_1, \text{ et } y \in H_2, \text{ on ait : } (Tx, y) = (x, T^*y).$$

## 1.1.2 Opérateurs linéaires dans un espaces de Hilbert

---

On a de plus  $\|T^*\| = \|T\|$ .

**Preuve.**  $\forall y \in H_2$  L'application  $x \rightarrow (Tx, y)$  est linéaire et continue ( de norme inférieure à  $\|T\| \|y\|$  d'après l'**inégalité de Cauchy-Schwarz**).

D'après le théorème de représentation de Riesz, il existe donc un unique élément noté  $T^*y$  tel que  $(Tx, y) = (x, T^*y)$ . On vérifie facilement que  $\forall y, z \in H_2, \text{ et } \lambda \in \mathbb{R}, T^*y + \lambda T^*z$  vérifie la propriété qui définit  $T^*(y + \lambda z)$ . Par unicité  $T^*y + \lambda T^*z = T^*(y + \lambda z)$ . Ce qui prouve que  $T^*$  est linéaire.

Enfin, la continuité résulte de l'inégalité de Cauchy-Schwarz :

$$\|T^*y\|_{H_1}^2 = (T^*y, T^*y) = (TT^*y, y) \leq \|T\| \|T^*y\|_{H_1} \|y\|_{H_2}$$

Ce qui montre de plus que  $\|T^*\| \leq \|T\|$ . Comme  $T$  et  $T^*$  jouent le même rôle on a :  $\|T^*\| = \|T\|$ . et  $(T^*)^* = T$  car la définition est symétrique.

**Définition 1.19 (Adjoint d'un opérateur)** Soient  $H_1, H_2$  deux espaces de Hilbert et  $T \in \mathcal{L}(H_1, H_2)$ .

L'unique application linéaire  $T^* \in \mathcal{L}(H_2, H_1)$  telle que :  $\forall x \in H_1, y \in H_2$  on ait  $(Tx, y) = (x, T^*y)$  est appelée **l'adjoint de T**.

**Proposition 1.2** Soient  $T$  et  $S$  deux opérateurs linéaires, et  $\alpha, \beta \in \mathbb{R}$ . Alors :

- $(\alpha T + \beta S)^* = \alpha T^* + \beta S^*$ .
- $(TS)^* = T^*S^*$ .

**Proposition 1.3** Soient  $H_1$  et  $H_2$  deux espaces de Hilbert et soit  $T \in \mathcal{L}(H_1, H_2)$ ; alors

1.  $\text{Ker}T^* = (\text{Im}T)^\perp$ ;
2.  $(\text{Ker}T)^\perp = \overline{\text{Im}T^*}$ ;

**Preuve.**

1. Si  $y \in H_2$ , on voit que  $y \in \text{Ker}T^*$  si et seulement si  $\forall x \in H_1, 0 = (T^*y, x) = (y, Tx)$ ; clairement, ceci équivaut à dire que  $y \in (\text{Im}T)^\perp$ , d'où la première assertion.
2. d'après 1., nous avons  $\overline{\text{Im}T^*} = ((\text{Im}T^*)^\perp)^\perp = (\text{Ker}T)^\perp$ .

**Définition 1.20** Un opérateur  $T \in \mathcal{L}(H)$  est dit

- **hermitien** ou **auto-adjoint** si  $T^* = T$ .
- **positif** s'il est hermitien et de plus  $(Th, h) \geq 0 \forall h \in H$ .

- **unitaire** si  $T$  est inversible et  $T^* = T^{-1}$ .
- **normal** si  $TT^* = T^*T$ .

**Remarque 1.4** En dimension fini, les opérateurs auto-adjoint sont ceux qui ont une matrice symétrique.

## 1.3 Opérateurs compacts

**Définition 1.21** Soit  $E$  et  $F$  deux espace vectoriels normés.

Une application  $T \in \mathcal{L}(E, F)$  est dit **compacte** (ou un opérateur compact) si :  $\overline{T(\bar{B}_E)}$  est un compact de  $F$  où  $\bar{B}_E$  est la boule unité fermé de  $E$ .

**Notation :** On note par  $K(E, F)$  l'ensemble des opérateurs compacts de  $E$  dans  $F$ . Dans le cas  $E = F$  on note simplement cet espace par  $K(E)$ .

**Proposition 1.4** Soit  $E, F$  deux espaces vectoriels normés et  $T \in K(E, F)$ . les conditions suivantes sont équivalents :

1.  $T$  est compact.
2.  $\forall B \subset E, B$  borné  $\overline{T(B)}$  est compact.
3. Toute suite bornée  $(x_n)_{n \in \mathbb{N}}$  de  $E$ ,  $(Tx_n)_{n \in \mathbb{N}}$  admet une valeur d'adhérence.

**Preuve.**

1. (1.  $\Rightarrow$  2.) Soit  $B \in E$  borné, alors  $\exists r > 0$  tel que  $B \subset r\bar{B}_E$  d'où  $\overline{T(B)} \subset r\overline{T(\bar{B}_E)}$ . Ainsi,  $\overline{T(B)}$  est compact, comme fermé du compact  $r\overline{T(\bar{B}_E)}$ .

2. (2.  $\Rightarrow$  3.) Il suffit de poser  $B = \{x_n, n \in \mathbb{N}, x_n \text{ convergente}\}$ .

3. (3.  $\Rightarrow$  1.) Soit  $(y_n)_{n \in \mathbb{N}}$  une suite de  $\overline{T(\bar{B}_E)}$ .

$\forall n \in \mathbb{N}, \exists z_n \in \overline{T(\bar{B}_E)}$  tel que  $\|y_n - z_n\| \leq 2^{-n}$ .

Comme par hypothèse  $(z_n)_{n \in \mathbb{N}}$  admet une valeur d'adhérence, il en est même pour  $(y_n)_{n \in \mathbb{N}}$ .

**Proposition 1.5** Soit  $E$  et  $F$  deux espace de Banach

- L'ensemble des opérateurs compacts  $K(E, F)$  est un sous espace vectoriel fermé de  $\mathcal{L}(E, F)$ .
- Soient  $T_1 \in \mathcal{L}(E, F)$  et  $T_2 \in \mathcal{L}(F, G)$ , alors si  $T_1$  ou  $T_2$  est compact,  $T_2 \circ T_1 \in \mathcal{L}(E, G)$ .

En particulier,  $K(E)$  est un idéal bilatère de  $\mathcal{L}(E)$ .

**Corollaire 1.4 (les isomorphismes ne sont pas compacts)** Soit  $E$  un espace vectoriel normé de dimension infini. Alors, l'opérateur identité de  $E$  n'est pas compact. De plus généralement, tout isomorphisme  $T : E \rightarrow E$  n'est pas compact.

**Preuve.** Pour l'opérateur identité  $I$  sur  $E$ , on a  $\overline{I(\bar{B}_E)} = \bar{B}_E$ , qui ne peut être compacte car la dimension de  $E$  est infinie. Quand à l'assumption générale, si un isomorphisme  $T : E \rightarrow E$  est compact alors l'opérateur identité  $I = T^{-1} \circ T$  serait compact, ce qui serait une contradiction.

**Corollaire 1.5** Soient  $E$  et  $F$  deux espaces vectoriels normés de dimension infinie, et  $T$  un opérateur compact de  $E$  dans  $F$ . Alors  $T$  n'est jamais inversible dans  $\mathcal{L}(E, F)$ .

**Preuve.** Si  $T$  est inversible, son inverse  $T^{-1}$  vérifie

$$T \circ T^{-1} = I$$

comme  $T$  est compact, et que  $I$  ne peut l'être d'après le corollaire 1.4, nous avons une contradiction.

**Définition 1.22 (opérateur de rang fini)** Soient  $E$  et  $F$  deux espaces vectoriels normés. Une application  $T \in \mathcal{L}(E, F)$  est dite **de rang fini** si la dimension de l'image de  $T$  est fini.

**Remarque 1.5** Tout opérateur de rang fini est compact.

**En effet.**,  $\overline{T(\bar{B}_E)}$  est un fermé borné de l'espace de dimension fini  $T(E)$ , est donc compact. Comme  $K(E, F)$  est fermé, il s'ensuit que tout opérateur qui peut être approché par des opérateurs de rang fini est également compact.

**Théorème 1.9** Soient  $H_1, H_2$  deux espace de Hilbert. Alors  $T \in K(H_1, H_2)$  si et seulement s'il existe une suite  $T_n \in \mathcal{L}(E, F)$ ,  $T_n$  de rang fini, telle que

$$\lim_{n \rightarrow \infty} \|T_n - T\| = 0.$$

**Preuve.** Soit  $T \in K(H_1, H_2)$  et  $\epsilon > 0$ . Comme  $T$  est compact,  $T(\bar{B}_E)$  est précompact, il existe donc  $N_\epsilon \in \mathbb{N}$  et  $\{y_1, \dots, y_{N_\epsilon}\}$  tels que

$$T(\bar{B}_E) \subset \cup_{i=1}^{N_\epsilon} B(y_i, \epsilon) \tag{1.7}$$

On pose  $F_\epsilon = \text{Vect} \{y_1, \dots, y_{N_\epsilon}\}$ , et  $P_\epsilon : H_2 \rightarrow H_2$  la projection orthogonale sur  $F_\epsilon$ .

Soit  $T_\epsilon = P_\epsilon \circ T$ . comme  $T_\epsilon(H_1) \subset F_\epsilon$ ,  $T_\epsilon$  est de rang fini.

D'autre part d'après (1.7),

$$\forall x \in \bar{B}_E, \quad \|Tx - T_\epsilon x\| = \|Tx - P_\epsilon(Tx)\| = \inf_{y \in F_\epsilon} \|Tx - y\| \leq \epsilon.$$

D'où  $\|T - T_\epsilon\| \leq \epsilon$ .

### 1.1.4 Décomposition spectrale des opérateurs auto-adjoints compacts

**Théorème 1.10 (Riesz)** (Voir H.Brezis)

Soit  $E$  un espace vectoriel normé tel que  $B_E$  soit compact. Alors  $E$  est de dimension finie.

**Théorème 1.11 (Ascoli-Arzelà)** (voir A.KOLMOGOROV)

Soit  $(E, \|\cdot\|)$  un espace normé compact.  $A \subseteq C(E, \mathbb{R})$ , Alors  $A$  est relativement compact (i.e.  $\bar{A}$  est compact) si et seulement si

1.  $A$  est bornée i.e.  $\exists c > 0 ; \forall u \in A, \|u\| \leq c$ .
2.  $A$  équicontinue i.e.  $\forall \epsilon > 0, \exists \delta > 0; \forall f \in A; \forall x, t \in E : \|x - t\| \leq \delta \Rightarrow |f(x) - f(t)| \leq \epsilon$ .

**Théorème 1.12 (Alternative de Fredholm)** (Voir H.Brezis)

Soit  $T \in K(E)$ . Alors

1.  $\text{Ker}(I-T)$  est de dimension finie,
2.  $\text{Im}(I-T)$  est fermé, et plus précisément  $\text{Im}(I-T) = \text{Ker}(I-T^*)^\perp$
3.  $\text{Ker}(I-T) = 0 \Leftrightarrow \text{Im}(I-T) = E$
4.  $\dim(I-T) = \dim(I-T^*)$

**Remarque 1.6** L'Alternative de Fredholm concerne la résolution de l'équation  $\mathbf{u-Tu=f}$ . Elle exprime que :

**ou bien** pour tout  $f \in E$  l'équation  $u - Tu = f$  admet une solution unique,

**ou bien** l'équation homogène  $u - Tu = 0$  admet  $n$  solutions linéairement indépendantes et, dans ce cas, l'équation non homogène  $u - Tu = f$  est résoluble si et seulement si  $f$  vérifie  $n$  **conditions d'orthogonalité** (i.e.  $f \in \text{Ker}(I-T^*)^\perp$ )

## 1.4 Décomposition spectrale des opérateurs auto-adjoints compacts

Dans toute cette section,  $T$  désigne un opérateur auto-adjoint compact dans un espace de Hilbert  $H$ .

**Définition 1.23 (Spectre d'un opérateur auto-adjoint compact)** Le spectre de  $T$  est l'ensemble

$$\sigma(T) = \{\lambda \in \mathbb{C}, T - \lambda I \text{ n'est pas inversible dans } \mathcal{L}(H)\}.$$

**Définition 1.24 (Valeur propre)** Un nombre  $\lambda \in \mathbb{C}$  est une valeur propre si et seulement si  $T - \lambda I$  n'est pas injectif.

### 1.1.4 Décomposition spectrale des opérateurs auto-adjoints compacts

---

**Remarque 1.7** L'ensemble des valeurs propres de  $T$  s'appelle le spectre ponctuel, **notée**  $\sigma_p(T)$ .  
C'est-à-dire :

$$\sigma_p(T) = \{\lambda \in \mathbb{C}, \text{Ker}(T - \lambda I) \neq \{0\}\}$$

On appelle multiplicité de la valeur propre  $\lambda$ , la dimension du sous espace propre  $\text{Ker}(T - \lambda I)$ .

**Proposition 1.6** 1.  $\sigma(T) = \{0\} \cup \sigma_p(T)$

2. Toute valeur propre non-nulle est de multiplicité finie ;

3. Les valeurs propres de  $T$  sont réelles et des vecteurs propres correspondant à des valeurs propres distinctes sont orthogonaux ;

4. L'un des nombres  $\pm \|T\|$  est une valeur propre de  $T$ .

**Théorème 1.13** (Voir H.Brezis)

$T$  admet une suite (finie ou infinie) de valeurs propres  $\lambda_n$  tel que :

$$|\lambda_1| \geq |\lambda_2| \geq \dots \geq |\lambda_n| \geq \dots$$

Et  $\lim_{n \rightarrow \infty} \lambda_n = 0$ . Dans ce cas la suite des vecteurs propres  $(e_n)_{n \in \mathbb{N}}$  associés aux valeurs propres  $\lambda_n$  constitue une base hilbertienne de  $(\text{Ker}T)^\perp$  tel que :

$$\forall x \in H, x = x_0 + \sum_{n \in \mathbb{N}} (x, e_n) e_n, \quad Tx = \sum_{n \in \mathbb{N}} \lambda_n (x, e_n) e_n, \text{ où } x_0 \in \text{Ker}T.$$

---

---

## Chapitre 2

---

# Opérateurs intégraux, équations intégrales et Méthode de Moindres Carrés linéaire

Dans ce chapitre, nous introduisons une source importante de problèmes inverses linéaires : les équations intégrales de première espèce. Nous étudierons ensuite l'approximation numérique des équations de première espèce par la méthode de quadratique-collocation et la méthode de Galerkin. Puis, on donne quelques propriétés mathématiques de la méthode de moindres carrés. Ensuite, on donne l'outil fondamental que constitue la décomposition en valeurs singulières. Enfin, nous montrerons comment la décomposition en valeurs singulières permet d'analyser les problèmes de moindres carrés.

### 2.1 Opérateurs intégraux

**Définition 2.1 (Opérateur intégral)** Soit  $k$  une fonction de l'espace  $L^2([a, b] \times [c, d])$ . L'opérateur :

$$Ay(x) = \int_c^d k(x, t)y(t)dt, \quad x \in [a, b]. \quad (2.1)$$

est bien défini en tant qu'opérateur de  $L^2[a, b]$  dans  $L^2[c, d]$ . On dira que l'opérateur  $A$  est l'opérateur intégral de noyau  $k$ .

**Exemple 2.1 (Opérateurs de Volterra)** Il s'agit d'opérateurs de la forme :

$$Ay(x) = \int_0^x k(x, t)y(t)dt, \quad x \in [0, 1].$$

avec  $k \in L^2([0, 1] \times [0, 1])$ .

**Théorème 2.1** Soit  $A$  l'opérateur intégral de noyau  $k$ . Alors

1. L'adjoint  $A^*$  est l'opérateur intégral de noyau  $k^*$  avec :

$$k^*(x, t) = k(t, x). \quad (2.2)$$

2. L'opérateur intégral  $A$  de noyau  $k$  est auto-adjoint si, et seulement si, le noyau  $k$  est symétrique c'est à dire :

$$k(x, t) = k(t, x), \forall (x, t) \in [a, b] \times [c, d]. \quad (2.3)$$

3. Soient  $A_1$ , et  $A_2$  les opérateurs intégraux de noyaux, respectivement  $k_1 \in L^2([a, b] \times [c, d])$  et  $k_2 \in L^2([c, d] \times [e, f])$ . Alors  $A_1 A_2 \in L^2((a, b) \times (e, f))$  est un opérateur intégral de noyau

$$k(x, r) = \int_c^d k_1(x, t) k_2(t, r) dt \quad (2.4)$$

**Preuve.**

1. Il suffit de partir de la définition. Soit  $u \in L^2(a, b)$ ,  $v \in L^2(c, d)$ .

$$(Au, v) = \int_c^d \left( \int_a^b k(x, t) u(t) dt \right) v(x) dx = \int_{[a, b] \times [c, d]} k(x, t) u(t) v(x) dt dx$$

par le théorème de Fubini. En échangeant encore l'ordre d'intégration, il vient :

$$(Au, v) = \int_a^b \left( \int_c^d k(x, t) v(x) dt \right) u(t) dx = (u, A^*v)$$

d'après la définition de l'adjoint. En permuttant le nom des variables, on obtient la définition 2.2.

2. découle de 1.

3. Ici encore, il suffit de suivre les définitions :

$$\begin{aligned} (A_1 A_2)(u)(t) &= \int_c^d k_1(x, t) \left( \int_e^f k_2(t, r) u(r) dr \right) dt \\ &= \int_e^f \left( \int_c^d k_1(x, t) k_2(t, r) dt \right) u(r) dr \end{aligned}$$

**Définition 2.2 (Opérateurs intégraux à noyaux dégénérés)** Un opérateur à noyau dégé-

nére est un opérateur de noyau se la forme :

$$k(x, t) = \sum_{j=1}^n a_j(x)b_j(t).$$

Les opérateurs correspondants sont de rang fini.

## 2.2 Equations intégrales

On classe les équations intégrales que l'on peut associer à l'opérateur intégral A en deux catégories :

1. Équations de première espèce : Il s'agit de l'équation de la forme suivante :

$$Au = f, \text{ où } f \in L^2[c, d] \text{ est donnée;} \quad (2.5)$$

2. Équations de seconde espèce : Il s'agit de l'équation de la forme suivante :

$$u - Au = f, \text{ où } f \in L^2[c, d] \text{ est donnée;} \quad (2.6)$$

**Théorème 2.2** Soit  $k \in L^2([a, b] \times [c, d])$ . L'opérateur A de noyau k est compact de  $L^2[a, b]$  dans  $L^2[c, d]$ .

**Preuve.** On utilisons le thérème de Ascoli-Arzela.

Soit B un borné de  $C[a, b] \Leftrightarrow \exists c > 0, \forall u \in B : \|u\| \leq c$

1.  $A(B)$  est bornée.
2. Montrons que  $A(B)$  est équicontinue.  $(\forall \epsilon > 0, \exists \delta > 0, \forall v \in A(B), \forall x, x' \in [a, b]; |x - x'| < \delta \Rightarrow |v(x) - v(x')| < \epsilon)$

$$\begin{aligned} \text{On a } |v(x) - v(x')| &= |Au(x) - Au(x')| \\ &= \left| \int_c^d (k(x, t) - k(x', t))u(t)dt \right| \\ &\leq \int_c^d |k(x, t) - k(x', t)| |u(t)| dt \\ &\leq K \int_c^d |k(x, t) - k(x', t)| dt \end{aligned}$$

$k(x, t)$  est continu sur  $[a, b] \times [c, d] \Rightarrow k(x, t)$  est uniformément continu sur  $[a, b] \times [c, d]$

$\forall \epsilon' > 0, \exists \delta' > 0, \forall x, x' \in [a, b]; |x - x'| < \delta' \Rightarrow |k(x, t) - k(x', t)| < \epsilon'$ . Donc  $|v(x) - v(x')| <$

$K\epsilon'(b-a)$ ,  $\epsilon' = \frac{\epsilon}{K(b-a)}$ ,  $|v(x) - v(x')| < \epsilon$ .

Donc  $A(B)$  est équicontinue, de 1 et 2 l'opérateur  $A$  est compact.

### 2.2.1 Discrétisation des équations intégrales

Nous allons nous borner à deux méthodes pour discrétiser une équation intégrale : la méthode de quadrature-collocation et la méthode de Galerkin.

#### 2.2.1.1 Discrétisation par quadrature

Cette méthode consiste à appliquer les méthodes numériques de calcul intégral pour aboutir à un système linéaire.

En générale, une formule de quadrature s'écrit se la forme :

$$I^Q = \sum_{j=1}^n w_j f(x_j), \quad j = 1, \dots, n. \quad (2.7)$$

Où  $x_j$  sont les noeuds et  $w_j$  sont les poids.

**Application d'une formule de quadrature à une équation intégrale de première espèce** On exprime que l'équation intégrale est vérifiée en un nombre fini de points  $t_i$ ,  $i = 1, \dots, m$ ,

$$\int_a^b k(x_i, t)u(t)dt = f(x_i).$$

On remplace cette intégrale par une forme quadrature choisie dans l'équation (2.7). On obtient le système :

$$\sum_{j=1}^n w_j k(x_i, t_j)u_j = f(x_i). \quad (2.8)$$

On remarque que c'est un système linéaire

$$Ay = b, \quad A = [A_{i,j}], \quad b = [b_i], \quad y = [y_j]$$

avec :

$$A_{i,j} = w_j k(x_i, t_j), \quad b_i = f(x_i), \quad y_j = u_j$$

pour  $j = 1, \dots, n$ ,  $i = 1, \dots, m$ .

**Remarque 2.1** Si  $m > n$ , on obtient un système sur-déterminé. Le système (2.8) devra en général être résolu au sens des moindres carrés.

**Exemple 2.2** Si l'opérateur intégral

$$Au(x) = \int_0^1 u(t)dt, \quad t \in [0, 1]$$

alors la matrice  $A_{i,j}$  est triangulaire, avec :

$$A_{i,j} = h, \quad j \geq 1$$

et le système d'équation (2.8) est :

$$h \sum_{j=1}^m u_j = f_i, \quad i = 1, \dots, m.$$

Par récurrence, on obtient que la solution s'écrit :

$$u_j = \frac{f_j - f_{j-1}}{h}, \quad j = 2, \dots, n.$$

avec

$$u_1 = f_1/h.$$

### 2.2.1.2 Discrétisation par la méthode de Galerkin

La méthode de Galerkin est une méthode qui permet de transformer un problème continu en un problème discret.

On approche les espaces  $L^2(a, b)$  et  $L^2(c, d)$  par une suite de sous-espaces  $E_n$  et  $F_m$  respectivement, on suppose que  $\dim E_n = n$  et  $\dim F_m = m$  tel que  $E_n \subset L^2(a, b)$  et  $F_m \subset L^2(c, d)$ .

On projette l'équation (2.5) sur  $F_m$ , c'est-à-dire que l'on cherche  $u_n \in E_n$  solution de l'équation :

$$(Au_n, v_m) = (f, v_m), \quad \forall v_m \in F_m \tag{2.9}$$

Cette équation est l'équation de Galerkin pour  $u_n$ . Pour expliciter cette équation, nous introduisons une base  $\{e_1, \dots, e_n\}$  dans l'espace  $E_n$  et une base  $\{f_1, \dots, f_m\}$  dans l'espace  $F_m$ .

Développons  $u_n$  dans cette base sous la forme :

$$u_n = \sum_{j=1}^n x_j e_j.$$

et prenons  $v_m = f_i$  dans l'équation de Galerkin (2.9). Il vient :

$$\sum_{j=1}^n (Ae_j, f_i) x_j = (f, f_i), \quad i = 1, \dots, m. \quad (2.10)$$

C'est un système d'équations qu'il admet un traitement numérique.

Les éléments de la matrice de la méthode de Galerkin sont des intégrales doubles et simples tel que :

$$A_{ij} = \int \int_{[a,b] \times [c,d]} k(x,t) f_i(x) e_j(t) ds dt, \quad b_i = \int_c^d f(x) f_i(x) dx \quad (2.11)$$

**Exemple 2.3** Nous choisissons la fonction constante par morceaux pour les deux sous-espaces d'approximation.

Plus précisément, posons  $h_s = (b-a)/n$ ,  $h_t = (d-c)/m$ , et subdivisons les intervalles  $]a, b[$  et  $]c, d[$  en  $n$  et  $m$  intervalles de taille  $h_s$  et  $h_t$  respectivement. Notons  $I_j^s = ]a + (j-1)h_s, a + jh_s[$  et  $I_i^t = ]c + (i-1)h_t, c + ih_t[$  ces intervalles, et définissons les fonction de base par :

$$e_j(t) = \begin{cases} h_s^{-1/2}, & \text{si } t \in I_j^s \\ 0 & \text{sinon,} \end{cases} \quad j = 1, \dots, n, \quad (2.12)$$

et

$$f_i(x) = \begin{cases} h_t^{-1/2}, & \text{si } x \in I_i^t \\ 0 & \text{sinon,} \end{cases} \quad i = 1, \dots, m, \quad (2.13)$$

Les éléments de la matrice et du second membre se calculent alors par

$$A_{ij} = \frac{1}{\sqrt{h_s h_t}} \int_{I_i^t} \int_{I_j^s} k(x,t) dx dt, \quad b_i = \int_{I_i^t} f(x) dx \quad (2.14)$$

## 2.3 Propriétés mathématiques des problèmes de moindres carrés

Dans tout ce paragraphe,  $A$  est un opérateur intégral compact de  $E$  dans  $F$  tel que  $E = L^2[a, b]$  et  $F = L^2[c, d]$  deux espaces de Hilbert.

Étant donné  $\hat{z} \in F$ , nous cherchons  $\hat{x} \in E$  solution de

$$A\hat{x} = \hat{z}. \quad (2.15)$$

Le problème (2.15) est dit mal posé si l'un des trois conditions suivantes est satisfaite :

1. L'opérateur  $A$  n'est pas surjectif.
2. L'opérateur  $A$  n'est pas injectif.
3. Si l'inverse existe et n'est pas continu.

La première difficulté n'est pas sérieuse : il suffit de se restreindre à  $ImA$ . La seconde est plus gênante car il faut sélectionner parmi plusieurs solutions une seule solution. La troisième est fondamentale par ce qu'elle est liée à la fermeture ou non de  $ImA$ .

**Théorème 2.3** *Soit  $A \in \mathcal{L}(E, F)$ ,  $E$  et  $F$  deux espaces de Hilbert. Supposons que  $A$  soit injectif, et notons  $A^{-1} : ImA \rightarrow E$  l'inverse de  $A$ . On a :*

$$ImA \text{ fermé} \iff A^{-1} \text{ est continu.}$$

**Preuve.**  $\Rightarrow$  Dans ce cas, soit  $W = ImA$ , donc  $W$  est un espace de Hilbert. L'opérateur  $\tilde{A} : E \rightarrow W$ ,  $\tilde{A}u = Au$  pour  $u \in E$  est un opérateur linéaire continu et bijectif. Une conséquence classique du théorème de l'application ouverte (théorème 1.7) est que  $\tilde{A}^{-1}$  est continu. Il en est donc de même pour  $A^{-1}$ .

$\Leftarrow$  Puisque  $A^{-1}$  est continu, et que  $E = ImA^{-1}$  est fermé,  $ImA = (A^{-1})^{-1}(E)$  est fermé dans  $F$ .

**Remarque 2.2** *L'opérateur  $A$  peut être injectif ou non donc  $ImA$  peut être n'est pas fermée. Le corollaire 1.5 montre que si  $A$  est compact,  $A$  n'a pas d'inverse continu, et dans ce cas  $ImA$  ne sera pas fermé.*

**Remarque 2.3** *Le problème (2.15) n'a de solution que pour les seconds membres dans l'image de  $A$ . Nous cherchons donc une autre formulation du problème original, qui permette d'étendre la notion de solution à un sous-espace plus grand.*

### 2.2.3 Propriétés mathématiques des problèmes de moindres carrés

Nous proposons une formulation comme un problème de *moindres carrés* :

$$\min_{x \in E} \frac{1}{2} \|Ax - \hat{z}\|_F^2 \quad (2.16)$$

Nous allons voir que ce problème est équivalent à une équation linéaire, mais pour un opérateur différent de  $A$ .

**Théorème 2.4** *Soit  $A \in \mathcal{L}(E, F)$ ,  $E, F$  deux espaces de Hilbert, et soit  $\hat{z} \in F$ . Un élément  $\hat{x} \in E$  est une solution de (2.16) si et seulement si*

$$A^*A\hat{x} = A^*\hat{z} \quad (2.17)$$

**Preuve.** Soit  $x$  vérifiant (2.17). On a pour tout  $y \in E$  :

$$\hat{z} - Ay = \hat{z} - Ax + A(x - y).$$

L'équation normale (2.17) implique que les deux termes de la somme sont orthogonaux ( le résidu  $\hat{z} - Ax$  est orthogonal à l'image de  $A$ ). Le théorème de Pythagore implique :

$$\|\hat{z} - Ay\|_F^2 = \|\hat{z} - Ax\|_F^2 + \|A(x - y)\|_F^2 \geq \|\hat{z} - Ax\|_F^2.$$

$x$  est donc bien solution de (2.16).

Réciproquement, soit  $x$  tel que  $A^*(\hat{z} - Ax) = w \neq 0$ . Choisissons  $y = x + \epsilon w$ , avec  $\epsilon > 0$ . On a alors :

$$\|\hat{z} - Ay\|_F^2 = (\hat{z} - Ay, \hat{z} - Ay) = \|\hat{z} - Ax\|_F^2 - 2\epsilon(\hat{z} - Ax, w) + \|Aw\|_F^2 < \|\hat{z} - Ax\|_F^2$$

si  $\epsilon$  est suffisamment petit.  $x$  n'est pas solution de (2.16).

**Remarque 2.4** *L'équation normale (2.17) se réécrit :*

$$A^*(A\hat{x} - \hat{z}) = 0,$$

*ce qui exprime simplement que le résidu  $\hat{z} - A\hat{x}$  est dans le noyau de  $A^*$ , c'est-à-dire orthogonal à (la fermeture de )l'image de  $A$  (voir la proposition 1.3).*

**Remarque 2.5** *La solution du problème de moindres carrés est telle que  $Ax$  est la projection de  $\hat{z}$  sur l'image de  $A$ .*

### 2.2.3 Propriétés mathématiques des problèmes de moindres carrés

---

**Lemme 2.1** *La solution du problème (2.16) est unique si, et seulement si, l'opérateur  $A$  est injectif.*

**Preuve.** Notons tout d'abord que  $\ker A^*A = \ker A$ .

$$A^*Ax = 0 \Leftrightarrow (A^*Ax, x) = 0 \Leftrightarrow \|Ax\|_F = 0 \Leftrightarrow Ax = 0$$

Par conséquent,  $A$  et  $A^*A$  sont injectifs en même temps. Supposons que le problème (2.16) admet deux solutions  $x_1$  et  $x_2$  alors :

$$\begin{aligned} A^*Ax_1 = \hat{z} \text{ et } A^*Ax_2 = \hat{z} &\Rightarrow A^*A(x_1 - x_2) = 0 \\ &\Rightarrow (x_1 - x_2) = 0 \text{ (} A^*A \text{ est injectif)} \\ &\Rightarrow x_1 = x_2. \end{aligned}$$

d'où le résultat.

**Proposition 2.1** *1. L'équation (2.17) admet une solution si et seulement si  $\hat{z} \in \text{Im}A \oplus \text{Im}A^\perp$ .*  
*2. si  $\hat{z} \in \text{Im}A \oplus \text{Im}A^\perp$ , l'ensemble  $S$  des solutions de (2.17) est un convexe fermé non vide de  $E$ .*

**Preuve.**

1. Soit  $x \in E$  une solution de (2.17). On a donc  $Ax - \hat{z} \in \ker A^* = \overline{\text{Im}A}^\perp = (\text{Im}A)^\perp$ . Donc  $\hat{z} = Ax + (\hat{z} - Ax) \in \text{Im}A \oplus (\text{Im}A)^\perp$ .  
 Inversement, soit  $\hat{z} = z^1 + z^2$ , avec  $z^1 \in \text{Im}A$ ,  $z^2 \in \text{Im}A^\perp$ . Il existe donc  $x \in E$ , tel que  $Ax = z^1$ .  
 Évidemment  $A^*Ax = A^*z^1$ . Mais, toujours parce que  $(\text{Im}A)^\perp = \ker A^*$ ,  $A^*z^2 = 0$ , c'est-à-dire que  $A^*\hat{z} = A^*z^1 = A^*Ax$ , et  $\hat{z}$  est une solution de (2.17).
2. L'ensemble des solutions est non-vidé d'après le point 1. C'est un espace affine, c'est donc en particulier un convexe, et il est fermé puisque c'est l'image réciproque de  $\{A^*\hat{z}\}$  par l'opérateur continu  $A^*A$ . Ceci prouve que  $S = x_0 + \ker A$ , où  $x_0$  est une solution quelconque de (2.17).

**Corollaire 2.1** *Si  $\hat{z} \in \text{Im}A \oplus \text{Im}A^\perp$ , le problème (2.16) admet une unique solution de norme minimale.*

### 2.2.3 Propriétés mathématiques des problèmes de moindres carrés

---

**Preuve.** Notons  $S$  l'ensemble de solutions de (2.16). Chercher une solution de norme minimale de ce problème de moindres carrés revient à résoudre le problème suivant :

$$\min_{x \in S} \|x\|_E, \quad S = \{x \in E, \|Ax - b\|_F \text{ minimal}/b \text{ donné}\}$$

c'est-à-dire à projeter l'origine sur l'ensemble  $S$ . D'après la proposition 2.1,  $S$  est un convexe fermé non-vidé de  $E$ . Le théorème de projection 1.1 implique que  $S$  possède un élément de norme minimal, qui est la solution cherchée.

Nous noterons  $\hat{x}$  cette solution particulière.

#### 2.3.0.3 Cas de la dimension finie

Si  $E$  et  $F$  sont deux espaces de dimension finie, on peut supposer que  $E = \mathbb{R}^n$  et  $F = \mathbb{R}^m$ . Dans ce cas,  $A$  s'identifie à une matrice  $A \in \mathbb{R}^{m \times n}$ . Nous supposons en général que  $m > n$ , c'est-à-dire que le problème est sur-déterminé.

On peut préciser, dans ce cas, les résultats du paragraphe précédent.

**Proposition 2.2** *Quand  $E$  et  $F$  sont deux espaces de dimension finie, le problème de moindres carrés admet toujours au moins une solution.*

*Cette solution est unique si, et seulement si,  $A$  est de rang maximal (de rang  $n$  si le problème est sur-déterminé).*

**Preuve.** Nous montrons que les équations normales ont toujours au moins une solution, c'est-à-dire que  $A^t \hat{z} \in \text{Im}(A^t)$ . Pour cela, commençons par remarquer que  $\text{Im} A^t = \text{Im}(A^t A)$ . En effet,  $\text{Im}(A^t A) = \ker(A^t A)^\perp = \ker(A)^\perp = \text{Im}(A^t)$ . Il suffit donc de vérifier que  $A^t \hat{z} \in \ker(A)^\perp$ , or c'est immédiat, puisque, si  $x \in \ker(A)^\perp$ ,  $(A^t \hat{z}, x) = (\hat{z}, Ax) = 0$ .

**Proposition 2.3** *Sous l'hypothèse que  $A$  est de rang  $n$ , la matrice des équations normales  $A^t A$  est définie positive.*

**Preuve.** Nous avons

$$(A^t A x, x) = (A x, A x) = \|x\|^2 \geq 0$$

donc  $A^t A$  est semi-définie positive. De plus, quand  $A$  est de rang  $n$ , nous avons vu (au lemme 2.1) que la matrice (carrée)  $A^t A$  est injective, donc inversible. Elle est donc définie positive.

## 2.4 Décomposition en valeurs singulières de matrices

La décomposition en valeurs singulières souvent abrégée par **SVD**, est un outil fondamental pour étudier un nombre croissant de problèmes linéaires.

**Théorème 2.5** Soit  $A \in IR^{m \times n}$  une matrice de rang  $r$ . Il existe deux matrices orthogonales  $U \in IR^{m \times m}$ , ( $U^t U = U U^t = I_m$ ) et  $V \in IR^{n \times n}$ , ( $V^t V = V V^t = I_n$ ) telles que

$$A = U \Sigma V^t, \quad \Sigma = \begin{pmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{pmatrix} \quad (2.18)$$

où  $\Sigma \in IR^{m \times n}$ ,  $\Sigma_1 = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_r)$ , et

$$\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0.$$

Composante par composante, l'identité matricielle (2.18) devient :

$$A v_j = \sigma_j u_j, \quad A^t u_j = \sigma_j v_j, \quad \text{pour } j = 1, \dots, n, \quad (2.19)$$

$$A^t u_j = 0 \quad \text{pour } j = n + 1, \dots, m. \quad (2.20)$$

**Remarque 2.6** Si l'on note  $U = (u_1, \dots, u_m)$ ,  $V = (v_1, \dots, v_n)$  les colonnes des matrices  $U$  et  $V$ , les vecteurs  $u_j$  et  $v_j$  sont, respectivement, les vecteurs singuliers droits et gauche associés à la valeur singulière  $\sigma_j$ .

**Preuve.** La preuve se fait par récurrence sur  $n$ .

Par définition de ce qu'est une norme matricielle subordonnée, il existe un vecteur  $v_1 \in IR^n$  tel que

$$\|v_1\|_2 = 1, \quad \|A v_1\|_2 = \|A\|_2 \stackrel{\text{def}}{=} \sigma,$$

où  $\sigma$  est strictement positif (si  $\sigma = 0$ , alors  $A = 0$ , et il n'y a rien à démontrer). Posons  $u_1 = 1/\sigma A v_1 \in IR^m$ .

Complétons le vecteur  $v_1$  en une base orthogonale de  $IR^n$ , et notons  $V = (v_1, V_1) \in IR^{n \times n}$  la matrice formée par les vecteurs de base. Faisons de même pour  $u_1$  et  $IR^m$ , notant  $U = (u_1, U_1) \in IR^{m \times m}$ . Remarquons que les matrices  $U$  et  $V$  sont orthogonales par construction.

D'après notre choix de  $U_1$ ,  $U_1^t A v_1 = \sigma U_1^t u_1 = 0$ , et donc le produit  $U^t A V$  a la structure par blocs suivante :

$$A_1 \stackrel{\text{def}}{=} U^t A V = \begin{pmatrix} \sigma & w^t \\ 0 & B \end{pmatrix}$$

## 2.2.4 Décomposition en valeurs singulières de matrices

---

avec  $w^t = u_1^t A V_1$  et  $B = U_1^t A V_1 \in IR^{(m-1) \times (n-1)}$ .

Comme  $U$  et  $V$  sont orthogonales,  $\|A_1\|_2 = \|A\|_2 = \sigma$ . Mais la double inégalité

$$\|A_1\|_2 (\sigma^2 + w^t w)^{1/2} \geq \left\| A_1 \begin{pmatrix} \sigma \\ w \end{pmatrix} \right\|_2 = \left\| \begin{pmatrix} \sigma^2 + w^t w \\ Bw \end{pmatrix} \right\|_2 \geq \sigma^2 + w^t w,$$

montre que  $\|A_1\|_2 \geq (\sigma^2 + w^t w)^{1/2}$ . On doit donc avoir  $w = 0$ . On peut alors terminer la démonstration en appliquant l'hypothèse de récurrence à  $B$ .

**Proposition 2.4** *Soit  $A \in IR^{m \times n}$  une matrice. Notons  $A = U \Sigma V^t$  sa décomposition en valeurs singulières.*

1. *Les valeurs propres de la matrice  $A^t A$  sont les nombres  $\sigma_j^2$ ,  $j = 1, \dots, n$ , et ses vecteurs propres sont les vecteurs singuliers à gauche de  $A$ ,  $v_j$ ,  $j = 1, \dots, n$ ;*

2. *Les valeurs propres de la matrice  $\begin{bmatrix} 0 & A^t \\ A & 0 \end{bmatrix}$  sont les nombres  $\pm \sigma_j$ ,  $j = 1, \dots, n$ , et ses*

*vecteurs propres sont  $\frac{1}{\sqrt{2}} \begin{bmatrix} v_j \\ \pm u_j \end{bmatrix}$*

Cette proposition nous permet de préciser en quel sens la SDV d'une matrice  $A$  est unique. Puisque les valeurs singulières sont les valeurs propres de  $A^t A$ , elle sont déterminées par  $A$ , donc unique. Les vecteurs singuliers à droite appartenant à une valeur singulière simple sont également uniques (à un facteur  $\pm 1$  près), par contre pour une valeur singulière multiple, seul le sous-espace est unique. Enfin, les vecteurs singuliers à gauche correspondant à une valeur singulière simple sont également uniques. Par contre, les  $m - n$  derniers vecteurs singuliers à droite ne sont pas déterminés uniquement par  $A$ , seul le sous-espace engendré par  $u_{n+1}, \dots, u_m$  est unique.

Il est important d'écrire la SDV de  $A$  sous une forme simplement différente de (2.18) :

– Tout d'abord, en posant  $U_n = (u_1, \dots, u_n)$ , on a

$$A = U_n \Sigma V^t \tag{2.21}$$

– De même si  $A$  n'est pas de rang  $n$ , notons  $U_r = (u_1, \dots, u_r)$ ,  $V_r = (v_1, \dots, v_r)$  les matrices formées par  $r$  premiers vecteurs singuliers, et  $\Sigma_r = \text{diag}(\sigma_1, \dots, \sigma_r)$ . Alors :

$$A = U_r \Sigma_r V_r^t = \sum_{i=1}^r \sigma_i u_i^t v_i. \tag{2.22}$$

## 2.2.4 Décomposition en valeurs singulières de matrices

---

**Remarque 2.7** Les matrices  $U_r$  et  $V_r$  sont orthogonales, puisque c'est le cas de vecteurs  $u_1, \dots, u_r$  et  $v_1, \dots, v_r$ . La décomposition (2.22) correspond à une somme de matrices de rang 1.

Cette dernière forme exprime que les colonnes de  $V_r$  forment une base de  $(\ker A)^\perp$ , que celles de  $U_r$  forment une base de  $ImA$ , et qu'en développant les vecteurs dans cette base, on a

$$x = x^0 + \sum_{i=1}^r (x, v_i) v_i \implies Ax = \sum_{i=1}^r \sigma_i (x, v_i) u_i, \quad \text{où } x^0 \in \ker A. \quad (2.23)$$

**Proposition 2.5** On a les relations suivantes :

1. Le rang de  $A$  est égal au nombre de valeurs singulières non-nulles ;
2.  $\ker A = \text{vect}(v_{r+1}, \dots, v_n)$ ,  $ImA = \text{vect}(u_1, \dots, u_r)$ ;
3.  $\ker A^t = \text{vect}(v_1, \dots, v_r)$ ,  $ImA^t = \text{vect}(u_{r+1}, \dots, u_m)$ ;
4.  $\|A\| = \sigma_1$ .

**Preuve.** Les matrices  $U$  et  $V$  sont des isométries, donc bijection. Les différentes propriétés sont donc des conséquences de ce qu'elles sont vraies pour  $\Sigma$ .

1. puisque le rang d'une matrice diagonale est égal au nombre d'éléments diagonaux non-nuls. Le résultat général est une conséquence de la remarque ci-dessus.
2. Soit  $v \in \ker A$ , alors  $V^t v \in \ker U^t A V = \ker \Sigma$ , et réciproquement. Le noyau de  $\Sigma$  est engendré par les colonnes  $r + 1$  à  $n$  de l'identité, et donc le noyau de  $A$  est engendré par le produit de  $V$  par ces colonnes, c'est-à-dire les vecteurs  $v_{r+1}, \dots, v_n$ .

Un raisonnement analogue donne le résultat pour l'image de  $A$ .

3. Le raisonnement est identique à celui du point précédent.
4. Les deux matrices  $U$  et  $V$  sont des isométries, donc pour tout  $x \in IR^n$  on a :

$$\frac{\|Ax\|_2}{\|x\|_2} = \frac{\|U\Sigma V^t x\|_2}{\|x\|_2} = \frac{\|\Sigma y\|_2}{\|y\|_2}$$

avec  $y = V^t x$  (et donc  $\|y\|_2 = \|x\|_2$ ). Donc :

$$\frac{\|Ax\|_2}{\|x\|_2} = \frac{(\sum_{j=1}^r \sigma_j^2 y_j^2)^{1/2}}{(\sum_{j=1}^r y_j^2)^{1/2}} \leq \sigma_1,$$

avec égalité pour  $y = (1, 0, \dots, 0)^t$ .

## 2.2.4 Décomposition en valeurs singulières de matrices

---

**Proposition 2.6** *Pour  $k = 1, \dots, n - 1$ , notons  $A_k = U\Sigma_k V^t$ . On a  $\|A - A_k\|_2 = \sigma_{k+1}$ , et la matrice  $A_k$  est la matrice de rang  $k$  qui minimise l'écart  $\|A - B\|_2$*

### 2.4.0.4 Applications de SVD aux problèmes de moindres carrés

La SVD fournit la solution la plus claire du problème de moindre carré.

On suppose que  $A$  est de rang différent de  $n$ , parceque il y a des points forts de SVD est justement de permettre de traiter ce cas. Quand  $A$  est de rang  $r < n$ , nous chercherons la solution de norme minimale.

Introduisons la SVD de  $A$  (équation (2.18)) dans le problème de moindre carré (2.16), on obtient :

$$\|Ax - z\|_2^2 = \|U\Sigma V^t x - z\|_2^2 = \|\Sigma V^t x - U^t z\|_2^2 \quad (2.24)$$

On pose  $w = U^t z$  et  $y = V^t x$ . Nous avons :

$$\|Ax - z\|_2^2 = \|\Sigma y - w\|_2^2 = \sum_{i=1}^r |\sigma_i y_i - w_i|^2. \quad (2.25)$$

On obtient donc toute les solutions du problème (2.16) en posant :

$$y_i = \begin{cases} w_i/\sigma_i, & \text{pour } i = 1, \dots, r \\ \text{quelconque} & \text{pour } i = r + 1, \dots, n. \end{cases} \quad (2.26)$$

Si  $A$  est de rang plein, il y a bien une solution unique, et dans le cas contraire, la solution est définie à l'addition d'un élément du noyau de  $A$  près, la norme de cette solution est  $\sum_{i=1}^r |w_i/\sigma_i|^2 + \|\text{élément du noyau}\|_2^2$ . La solution de norme minimale est donc celle qui n'a pas de composante dans  $\ker A$ .

**Théorème 2.6** *La solution de norme minimale du problème de moindre carré (2.16) est donnée par :*

$$x = \sum_{i=1}^r (z, u_i)/\sigma_i v_i \quad (2.27)$$

**Preuve.** Par définition,  $y = V^t x$  s'inverse en  $x = Vy$ , donc les composantes sont  $(x, v_i) =$

## 2.2.4 Décomposition en valeurs singulières de matrices

---

$(Vy, v_i) = (y, V^t v_i) = y_i$  En utilisant (2.26), nous avons

$$x = \sum_{i=1}^r (x, v_i) v_i = \sum_{i=1}^r y_i v_i = \sum_{i=1}^r (z, u_i) / \sigma_i v_i$$

### 2.4.1 Développement en valeur singulières des opérateurs compacts

Soit  $A$  un opérateur compact de  $E$  dans  $F$ . Nous allons généraliser la décomposition en valeurs singulières à cette situation. La principale différence sera l'existence d'une infinité dénombrable de valeurs singulières.

**Théorème 2.7** *Il existe une suite  $(\sigma_j)_{j \in \mathbb{N}} \in \mathbb{R}_+$ , et deux familles orthonormales  $(e_j)_{j \in \mathbb{N}} \in E$ ,  $(f_j)_{j \in \mathbb{N}} \in F$  telles que :*

1.  $(\sigma_j)_{j \in \mathbb{N}}$  décroissante,  $\lim_{j \rightarrow +\infty} \sigma_j = 0$
2.  $Ae_j = \sigma_j f_j$ ;  $A^* f_j = \sigma_j e_j$ ,  $j \in \mathbb{N}$
3. Pour tout  $x \in E$ , on a le développement

$$x = x_0 + \sum_{j=1}^{+\infty} (x, e_j) e_j, \quad \text{où } x_0 \in \ker A \tag{2.28}$$

4. Pour tous  $x \in E$  et  $y \in F$  on a :

$$Ax = \sum_{j=1}^{+\infty} \sigma_j (x, e_j) f_j, \quad A^* y = \sum_{j=1}^{+\infty} \sigma_j (y, f_j) e_j \tag{2.29}$$

La suite  $(e_j)_{j \in \mathbb{N}}$  est une base hilbertienne de  $\ker A^\perp$ , la suite  $(f_j)_{j \in \mathbb{N}}$  est une base hilbertienne de  $\overline{\text{Im} A}$ .

**Preuve.** Si on pose  $T = A^* A$ ,  $T$  est un opérateur auto adjoint compact comme composé de deux opérateurs compacts. la théorème 1.13 implique que  $T$  a une suite de valeurs propres non nulles  $(\lambda_j)_{j \in \mathbb{N}} \in \mathbb{R}^*$ , et de vecteurs propres  $(e_j)_{j \in \mathbb{N}}$ , tels que

$$Te_j = \lambda_j e_j$$

et

$$\forall x \in E, \exists x_0 \in \ker T, \quad x = x_0 + \sum_{j=1}^{+\infty} (x, e_j) e_j.$$

## 2.2.4 Décomposition en valeurs singulières de matrices

---

Si on range les  $(\lambda_j)_{j \in IN}$  en une suite décroissante, et chaque valeurs étant comptée avec son ordre de multiplicité, alors  $\forall j, \lambda_j > 0$  puisque :

$$\lambda_j = (Te_j, e_j) = (A^*Ae_j, e_j) = \|Ae_j\|^2 \geq 0$$

et  $\lambda_j \neq 0$  par hypothèse. On définit  $\sigma_j = \lambda_j^{1/2}$ ,  $\sigma_j > 0$  et on pose  $f_j = \frac{1}{\sigma_j}Ae_j \in ImA \subset F$ .

Nous avons  $A^*f_j = \frac{1}{\sigma_j}A^*Ae_j = \frac{1}{\sigma_j}e_j = \sigma_j e_j$ . et on a

$$(f_j, f_k) = \frac{1}{\sigma_j \sigma_k} (Ae_j, Ae_k) = \frac{\lambda_k}{\sigma_j \sigma_k} (e_j, e_k) = \sqrt{\frac{\sigma_k}{\sigma_j}} \delta_{jk} = \begin{cases} 1 & \text{si } j = k \\ 0 & \text{si } j \neq k, \end{cases} .$$

donc  $(f_j)_{j \in IN}$  est une suite orthonormale dans  $F$ .

Les identités 2 sont immédiates.

Nous avons  $\ker T = \ker(A^*A) = \ker A$  (voir le lemme 2.1). et on a donc le développement (2.28). et pour le premier développement de (2.29) on a pour  $x \in E$ , posons

$$X = \sum_{j=1}^{+\infty} \sigma_j(x, e_j) f_j.$$

Cette série est convergente dans  $E$  avec :

$$\begin{aligned} \|X\|^2 &= \sum_{j=1}^{+\infty} |\sigma_j(x, e_j)|^2 = \sum_{j=1}^{+\infty} \sigma_j^2 |\sigma_j(x, e_j)|^2 \\ &\leq \sum_{j=1}^{+\infty} \sigma_1^2 |\sigma_j(x, e_j)|^2 \quad (\text{car } \sigma_1 > \sigma_j) \\ &\leq \sigma_1^2 \sum_{j=1}^{+\infty} |\sigma_j(x, e_j)|^2 = \sigma_1^2 \|x\|^2 \\ &\Rightarrow \|X\| \leq \sigma_1 \|x\| \end{aligned}$$

et  $X \in \overline{ImA}$  car pour

$$x = x_0 + \sum_{j=1}^{+\infty} (x, e_j) e_j$$

## 2.2.4 Décomposition en valeurs singulières de matrices

---

on a

$$\begin{aligned}
 Ax &= A(x_0 + \sum_{j=1}^{+\infty} (x, e_j) e_j) \\
 &= \sum_{j=1}^{+\infty} A((x, e_j) e_j) \\
 &= \sum_{j=1}^{+\infty} (x, e_j) \sigma_j f_j \\
 &= \sum_{j=1}^{+\infty} \sigma_j (x, e_j) f_j
 \end{aligned}$$

et la même idée pour

$$A^*y = \sum_{j=1}^{+\infty} \sigma_j (y, f_j) e_j$$

seulement  $\forall y \in F$

$$y = \sum_{j=1}^{+\infty} (y, f_j) f_j$$

ce qui prouve que  $X \in \overline{\text{Im}A}$ .

**Définition 2.3** Les nombre  $\sigma_j$  sont appelés les valeurs singulières de  $A$ . Les vecteurs  $e_j$  et  $f_j$  sont les vecteurs singuliers. Le développement obtenu en (2.28) s'appelle le développement en valeurs singulières de  $A$

### 2.4.2 Applications de SVD aux problèmes de moindres carrés

**Théorème 2.8** Soit  $z \in F$ . L'équation (2.15) possède une solution dans  $E$  si et seulement si  $z \in \overline{\text{Im}A}$  et de plus :

$$\sum_{j=1}^{+\infty} \frac{|(z, f_j)|^2}{\sigma_j^2} < \infty \quad (2.30)$$

Dans ce cas l'ensemble des solutions de (2.15) est donné par

$$x = \sum_{j=1}^{+\infty} \frac{(z, f_j)}{\sigma_j} e_j + \ker A \quad (2.31)$$

**Preuve.** Le développement de  $z$  sur la base hilbertienne des  $(f_j)_{j \in \mathbb{N}} \subset F$  est :  $z = \sum_{j=1}^{+\infty} (z, f_j) f_j$ , et le développement de  $x$  sur la base hilbertienne des  $(e_j)_{j \in \mathbb{N}} \subset E$  est :  $x = \sum_{j=1}^{+\infty} (x, e_j) e_j$ ,

## 2.2.4 Décomposition en valeurs singulières de matrices

---

Nous avons :

$$\begin{aligned}
 Ax &= A \left( \sum_{j=1}^{+\infty} (x, e_j) e_j \right) \\
 &= \sum_{j=1}^{+\infty} (x, e_j) A e_j \quad (\text{car } A \text{ est compact} \Rightarrow A \text{ est continue}) \\
 &= \sum_{j=1}^{+\infty} (x, e_j) \sigma_j f_j = \sum_{j=1}^{+\infty} x_j \sigma_j f_j \quad \text{avec } x_j = (x, e_j)
 \end{aligned}$$

$x$  est la solution de l'équation (2.15) équivalente à

$$\begin{aligned}
 Ax &= z \\
 \Leftrightarrow \sum_{j=1}^{+\infty} (x, e_j) \sigma_j f_j &= \sum_{j=1}^{+\infty} (z, f_j) f_j \\
 \Leftrightarrow (x, e_j) \sigma_j &= (z, f_j) \\
 \Leftrightarrow (x, e_j) &= \frac{(z, f_j)}{\sigma_j} \\
 \Leftrightarrow x &= \sum_{j=1}^{+\infty} (x, e_j) e_j = \sum_{j=1}^{+\infty} \frac{(z, f_j)}{\sigma_j} e_j
 \end{aligned}$$

Alors l'ensemble de solutions de l'équation (2.15) est  $x$  tel que :

$$x = \sum_{j=1}^{+\infty} \frac{(z, f_j)}{\sigma_j} e_j + \ker A.$$

**Remarque 2.8** La condition (2.30) s'obtient simplement en exprimant que la série des coefficients du développement de  $x$  doit être de carré intégrable est lui même la norme minimale de  $x$ , et cette condition s'appelle la condition de Picard.

**Remarque 2.9** La solution de norme minimale s'obtient quand  $\ker A = 0$ , en effet la série est orthogonale à  $\ker A$ . Puisque  $f_j \in \text{Im} A$ ;  $\forall j \in \mathbb{N}$  et donc la somme est dans  $\overline{\text{Im} A}$  (car si  $z_n \in \text{Im} A \Rightarrow \lim z_n \in \text{Im} A$ ).

---

---

# Chapitre 3

---

## Méthodes itératives

Dans ce chapitre, nous examinerons deux méthodes itératives pour les problèmes inverses linéaires : la méthode de Landweber et la méthode du gradient conjugué. La méthode de Landweber a pour principal avantage de se prêter à une analyse simple. Malheureusement, elle converge trop lentement pour être utilisable en pratique. La méthode du gradient conjugué est la plus employée.

### 3.1 Méthode de Landweber

Landweber [12] est proposé de réécrire l'équation  $Ax = \hat{z}$  sous la forme suivante :

$$x = (I - \omega A^* A)x + \omega A^* \hat{z} \quad (3.1)$$

où  $\omega > 0$ . Le schéma itérative de cette équation est le suivant :

$$\begin{cases} x_0 = 0; \\ x_{n+1} = x_n + \omega (A^* \hat{z} - A^* A x_n), \quad n \in \mathbb{N} \end{cases} \quad (3.2)$$

**Remarque 3.1** 1. Nous avons :  $\forall n \in \mathbb{N} : x_n \in \text{Im}A \subset \ker A^\perp$  ;

2. Si  $(x_n)$  converge vers un élément  $x \in E$ , on doit avoir  $x_n \in \ker A^\perp$  (car  $\ker A^\perp$  est un sous-espace fermé) ;

3. Par continuité,  $x$  satisfait l'équation (2.17), et donc  $x = \hat{x}$  est la solution de norme minimale de (2.15).

Il reste donc à trouver une condition sur  $\omega$  qui assure la convergence de  $(x_n)$  vers  $x = \hat{x}$

Soient  $A$  un opérateur compact,  $(\sigma_j)_{j \in \mathbb{N}}$  ses valeurs singulières et  $(u_j)_{j \in \mathbb{N}}$ ,  $(v_j)_{j \in \mathbb{N}}$  ses vecteurs singuliers. Nous avons le résultat de convergence suivant :

**Proposition 3.1** 1. On a

$$\forall n \in \mathbb{N}^* : x_n = \omega \sum_{j=0}^{n-1} (I - \omega A^* A)^j A^* \hat{z}; \quad (3.3)$$

2. Si on choisit  $0 < \omega < 2/\sigma_1^2$ , la suite itérés  $(x_n)$  de la méthode de Landweber converge vers  $\hat{x}$  pour  $\hat{z} \in \text{Im}A$ ;
3. Sous l'hypothèse (de régularité) supplémentaire  $\hat{x} \in \text{Im}A^*$ , on a l'estimation suivant

$$\|x_n - \hat{x}\| = O(1/n) \quad (3.4)$$

**Preuve.**

1. L'équation (3.3) se démontre par récurrence.
2. Si on pose  $e_n = x_n - \hat{x}$ . En soustrayant (3.3) de (2.17), nous obtenons pour  $n > 0$

$$e_n = (I - \omega A^* A)e_{n-1} = (I - \omega A^* A)^n e_0. \quad (3.5)$$

Si nous introduisons le développement en valeurs singulières de  $e_0 = \sum_{j=1}^{\infty} (e_0, u_j) u_j$ , l'équation (3.5) se réécrit

$$e_n = \sum_{j=1}^{\infty} (e_0, u_j) (1 - \omega \sigma_j^2)^n u_j$$

Comme les vecteurs singuliers sont orthogonaux, on déduit

$$\|e_n\|^2 = \sum_{j=1}^{\infty} |(e_0, u_j)|^2 |1 - \omega \sigma_j^2|^{2n}, \quad (3.6)$$

et si  $\omega < 2/\|A\|^2 = 2/\sigma_1^2$ , chaque terme  $|1 - \omega \sigma_j^2|$  est en valeur absolue strictement inférieur à 1. Pour passe à la limite dans (3.6) nous utilisons le théorème de convergence dominé de Lebesgue. Tout d'abord, chaque terme de la somme tend vers 0, et est majoré par  $|(e_0, u_j)|$ . Or d'après l'inégalité de Bessel

$$\sum_{j=1}^{\infty} |(e_0, u_j)|^2 \leq \|e_0\|^2.$$

Le théorème de Lebesgue permet donc de conclure que  $e_n \rightarrow 0$ , et donc que la suite  $x_n$  converge vers  $\hat{x}$ .

3. Par hypothèse, il existe  $w \in E$ , tel que  $\hat{x} = A^*w$  et (3.6) devient

$$\|e_n\|^2 = \sum_{j=1}^{\infty} \sigma_j^2 |(w, v_j)|^2 |1 - \omega \sigma_j^2|^{2n}$$

En utilisant alors l'inégalité (démontrée plus bas)

$$|1 - \omega \sigma_j^2|^n \leq \frac{1}{n\omega} \frac{1}{\sigma_j^2} \quad (3.7)$$

nous obtenons

$$\|e_n\|^2 \leq \frac{1}{n^2\omega^2} \sum_{j=1}^{\infty} |(w, v_j)|^2 = \frac{1}{n^2\omega^2} \|w\|_F^2 \quad (3.8)$$

c'est l'estimation du théorème.

Démontrons maintenant l'inégalité (3.7), ainsi qu'une autre qui nous sera utile pour la démonstration du théorème suivant.

**Lemme 3.1** *On a les inégalités, valable pour tout  $x \in ]0, 1[$*

$$(1 - x)^n \leq \frac{1}{nx} \quad (3.9)$$

$$1 - (1 - x)^n \leq nx \quad (3.10)$$

**Preuve.** Pour la première inégalité, il suffit de montrer que

$$nx(1 - x)^n \leq 1, \quad \text{pour } x \in [0, 1]$$

Nous avons la fonction  $f(x) = nx(1 - x)^n$  est positive en  $[0, 1]$ , et les valeurs 0 et 1 sont des zéros pour  $f$ , et la dérivée de  $f$  est  $f'(x) = n(1 - x)^n(1 - \frac{nx}{1-x})$  qui s'annule au point  $\frac{1}{n+1}$ , où la fonction vaut  $\left(\frac{n}{n+1}\right)^{n+1} \leq 1$ .

La second inégalité est une conséquence de la concavité de la fonction  $x \mapsto (1 - x)^n$  :

$$1 - (1 - x)^n \leq 1 - (1 - nx) = nx$$

Nous voyons donc que la méthode itérative converge pour des données non bruitées. Considérons maintenant ce qui se passe si nous remplaçons  $z$  par  $z^\delta$  avec  $\|z - z^\delta\| = \delta$ . Nous noterons

$x^{n,\delta}$  la suite des itérés correspondants, et nous introduirons la suite  $x^n$  construite à partir de la donnée non bruitée  $z$ . Nous voulons estimer l'erreur  $x^{n,\delta} - \hat{x}$ , et montrer que nous pouvons choisir  $n$  en fonction de  $\delta$  pour obtenir la convergence.

**Théorème 3.1**    1. Si  $n$  est choisi tel que  $n\delta \rightarrow 0$  quand  $n \rightarrow \infty$ , on a  $\lim_{\delta \rightarrow 0} x^{n(\delta),\delta} = \hat{x}$ .  
 2. Sous l'hypothèse supplémentaire  $\hat{x} \in \text{Im}A^*$ , on peut choisir  $n(\delta) = E(1/\sqrt{\delta})$  et alors  $\|x^{n(\delta),\delta} - \hat{x}\| = O(\sqrt{\delta})$  ( $E$  est la fonction partie entière).

**Preuve.** Nous avons

$$\|x^{n,\delta} - \hat{x}\| \leq \|x^n - \hat{x}\| + \|x^{n,\delta} - x^n\| \quad (\text{inégalité triangulaire})$$

Dans le premier cas, nous avons  $\|x^n - \hat{x}\| \rightarrow 0$ , dans le second terme nous savons de plus qu'il se comporte (au moins) comme  $O(1/n)$ .

Pour estimer le second terme, soustrayons l'équation définissant  $x^{n,\delta}$  de celle définissant  $x^n$ . En notant  $d_n = x^{n,\delta} - x^n$ , il vient :

$$d^{n+1} = (I - \omega A^* A)d^n + \omega A^*(z^\delta - z), \quad d_0 = 0.$$

Cette résultat se démontre par la même récurrence qui définit la suite  $x_n$  originale. Nous pouvons donc utiliser (3.3). Un raisonnement semblable à celui de la proposition 3.1 conduit à la représentation :

$$\|d^{n+1}\|^2 = \sum_{j=1}^{\infty} \frac{1 - (1 - \omega \sigma_j^2)^{2n}}{\sigma_j^2} |(z^\delta - z, v_j)|^2. \quad (3.11)$$

soit, d'après le lemme 3.1

$$\|d^{n+1}\|^2 \leq \omega n \sum_{j=1}^{\infty} |(z^\delta - z, v_j)|^2 = n\omega \|z^\delta - z\|^2 \leq n\omega\delta. \quad (3.12)$$

1. Dans le cas général, on a seulement

$$\|x^{n,\delta} - \hat{x}\| \leq \|x^n - \hat{x}\| + n\omega\delta$$

et le premier terme tend vers 0. La condition  $n\delta \rightarrow 0$  suffit à assurer la convergence.

2. Dans ce cas on peut préciser la convergence, en utilisant 2. du théorème précédent.

$$\|x^{n,\delta} - \hat{x}\| \leq \frac{1}{n\omega} \|w\|_F^2 + n\omega\delta,$$

La somme au second membre est minimale quand  $n\omega = \|w\|_F \delta^{-1/2}$ , et le minimum vaut  $\|w\|_F \delta^{1/2}$ .

On remarque que la méthode de Landweber converge trop lentement pour être utilisable en pratique. Il existe d'autres méthodes itératives, comme la méthode du gradient conjugué, qui converge bien plus rapidement pour un coût essentiellement équivalent.

## 3.2 Méthode du gradient conjugué

La méthode du gradient conjugué ne s'agit non seulement d'une des techniques les plus utiles pour résoudre de grands systèmes linéaires.

La méthode analysée ci-dessous est celle trouvée dans les années 50 par Hestens et Stiefel, à savoir la méthode du gradient conjugué linéaire. Cette dernière se base sur la recherche de directions successives permettant d'atteindre la solution exacte d'un système linéaire de matrice symétrique et définie positive.

La méthode du gradient conjugué peut être utilisée pour résoudre les problèmes (2.17) notons que la matrice  $A^T A$  est symétrique et définie positive (pour tout  $x$ ,  $x^T A^T A x = \|Ax\|^2 \geq 0$ ).

La méthode du gradient conjugué linéaire est une méthode qui résout deux problèmes équivalents possédant la même solution unique. Ces problèmes sont le système d'équation linéaires

$$Ax = b \tag{3.13}$$

et le problème de minimisation suivant :

$$\Phi(x) = \frac{1}{2} x^T A x - b^T x.$$

Dans ces deux problèmes,  $A$  est une matrice  $n \times n$ , symétrique et définie positive. notons encore que le gradient de  $\Phi$  est égal au résidu du système linéaire :

$$\nabla\Phi(x) = Ax - b := r(x).$$

### 3.2.1 Méthode des directions conjuguées

Une des multiples propriétés intéressantes de la méthode du gradient conjugué est la création économique d'une suite de vecteurs  $(p_k)_{k \in \mathbb{N}}$  de  $n$  vecteurs  $A$ -conjugués, avec  $A$  une matrice symétrique et définie positive. L'ensemble de ces vecteurs non-nuls de  $\mathbb{R}^n$   $\{p_0, p_1, \dots, p_{n-1}\}$  est

### 3.3.2 Méthode du gradient conjugué

dit A-conjugué si

$$p_i^T A p_j = 0, \quad \text{pour tout } i \neq j$$

et forme une base de  $\mathbb{R}^n$ . Grâce à l'A-conjugaison des vecteurs de direction  $\{p_0, p_1, \dots, p_{n-1}\}$ , il est possible de minimiser  $\Phi$  en seulement n pas, en minimisant  $\Phi$  successivement selon les n direction. Soient  $x_0 \in \mathbb{R}^n$  un point de départ quelconque et  $x^* \in \mathbb{R}^n$  la solution exacte de la problème (3.13). Définitions la suite

$$x_{k+1} = x_k + \alpha_k p_k \tag{3.14}$$

où  $\alpha_k$  est le scalaire minimisant  $\Phi$  selon la direction  $x_k + \alpha_k p_k$  et qui défini par

$$\alpha_k = -\frac{r_k^T p_k}{p_k^T A p_k} \tag{3.15}$$

**Théorème 3.2**  $\forall x_0 \in \mathbb{R}^n$ , la suite  $\{x_k\}$ , définie par l'algorithme des directions conjuguées (défini par (3.14) et (3.15)), converge en au maximum n pas vers  $x^*$ .

**Preuve.** Sachant que  $p_0, \dots, p_{n-1}$  engendrent  $\mathbb{R}^n$  parce que  $\{p_0, p_1, \dots, p_{n-1}\}$  forme une base de  $\mathbb{R}^n$ , on définit  $\sigma_k = \frac{p_k^T A(x^* - x_0)}{p_k^T A p_k}$  tel que :  $x^* - x_0 = \sigma_0 p_0 + \sigma_1 p_1 + \dots + \sigma_{n-1} p_{n-1}$ .

Selon l'algorithme (3.13), nous avons  $x_k = x_0 + \alpha_0 p_0 + \alpha_1 p_1 + \dots + \alpha_{k-1} p_{k-1}$ .

En multipliant cette expression par  $p_k^T A$ , on voit que  $p_k^T A(x_k - x_0) = 0$  et donc que

$$p_k^T A(x^* - x_0) = p_k^T A(x^* - x_k) = p_k^T (b - A x_k) = -p_k^T r_k.$$

Ainsi, finalement  $\sigma_k = \alpha_k$ , ce qui démontre le théorème

**Remarque 3.2** Nous avons  $r_{k+1} = r_k + \alpha_k A p_k$ .

**Théorème 3.3** Soit  $x_0 \in \mathbb{R}^n$  un point de départ quelconque et soit la suite  $\{x_k\}$  générée par l'algorithme des directions conjuguées.

Alors

$$r_k^T p_i = 0 \quad \text{pour } i = 0, \dots, k-1 \tag{3.16}$$

et  $x_k$  est le minimiseur de  $\Phi(x)$  sur l'ensemble

$$\{x | x = x_0 + \text{span} \{p_0, \dots, p_{k-1}\}\} \tag{3.17}$$

**Preuve.** Nous montrons d'abord qu'un point  $\tilde{x}$  minimise  $\Phi$  sur l'ensemble (3.17) si et seulement si  $r(\tilde{x})p_i = 0$ , pour tout  $i = 0, 1, \dots, k-1$ . Définissons  $h(\sigma) = \Phi(x_0 + \sigma_0 p_0 + \dots + \sigma_{k-1} p_{k-1})$ , où  $\sigma = (\sigma_0, \sigma_1, \dots, \sigma_{k-1})^T$ . Comme  $h(\sigma)$  est une fonction quadratique strictement convexe, elle a un unique minimiseur, qui satisfait

$$\frac{\partial h(\sigma^*)}{\partial \sigma_i} = 0 \quad i = 0, 1, \dots, k-1. \quad (3.18)$$

Par la règle de la chaîne, ceci implique que

$$\nabla \Phi(x_0 + \sigma_0^* p_0 + \dots + \sigma_{k-1}^* p_{k-1})^T p_i = 0 = r(x_k)^T p_i \quad i = 0, 1, \dots, k-1.$$

Ce qui nous donne le résultat.

Ensuite, nous montrons par récurrence que  $x_k$  satisfait (3.16).

Ancrage : pour  $k = 1$ , on a  $r_1^T p_0 = 0$ , pour le choix de  $\alpha_0$  dans (3.16).

Hypothèse de récurrence :  $r_{k-1}^T p_i = 0$  pour  $i = 0, 1, \dots, k-2$ .

Par

$$r_k = r_{k-1} + \alpha_{k-1} A p_{k-1}$$

nous obtenons

$$p_{k-1}^T r_k = p_{k-1}^T r_{k-1} + \alpha_{k-1} p_{k-1}^T A p_{k-1} = 0$$

Pour les autres vecteurs  $p_i$ ,  $i = 0, 1, \dots, k-2$ , nous avons

$$p_i^T r_k = p_i^T r_{k-1} + \alpha_{k-1} p_i^T A p_{k-1} = 0$$

par l'hypothèse de récurrence et la conjugaison de  $p_i$ . Nous concluons que  $r_k^T p_i = 0$ , pour  $i = 0, 1, \dots, k-1$ .

Le théorème 3.3 nous dit donc que le résidu  $r_k$  est orthogonal à toutes les directions de recherche précédentes  $p_0, p_1, \dots, p_{k-1}$ .

**Remarque 3.3** *Il y a plusieurs méthodes pour choisir les directions de recherche. On pourrait par exemple utiliser les valeurs propres de la matrice  $A$ .*

### 3.2.2 Propriétés de base de la méthode du gradient conjugué

La méthode du gradient conjugué est une méthode de directions conjuguées qui possède la propriété particulière de n'utiliser que le vecteur  $p_{k-1}$  et le résidu  $r_k$  pour générer le vecteur  $p_k$

### 3.3.2 Méthode du gradient conjugué

de l'ensemble des vecteurs A-conjugués.

Dans la méthode du gradient conjugué, les directions  $p_k$  sont choisies telles qu'elles sont des combinaisons linéaires entre la direction de la plus grande pente  $-\nabla\Phi(x_k) = -r_k$  et la direction de  $p_{k-1}$  :

$$p_k = -r_k + \beta_k p_{k-1}$$

où  $\beta_k$  est choisi que l'A-conjugaison entre  $p_k$  et  $p_{k-1}$  soit vérifiée. Après la multiplication par  $p_{k-1}^T A$ , on obtient

$$\beta_k = \frac{r_k^T A p_{k-1}}{p_{k-1}^T A p_{k-1}}.$$

On choisit la direction de la plus grande pente au point initial  $x_0$  comme la première direction de recherche,  $p_0$ . Voici une première version d'un algorithme pour la méthode du gradient conjugué :

**Algorithme 3.1** (*Algorithme du gradient conjugué*)

**Initialisation**  $k = 0$  : choix de  $x_0 \in \mathbb{R}^n$  et posons  $r_0 = Ax_0 - b$  et  $p_0 = -r_0$  ;

**Itération**  $k$  :

(a) Si  $r_k = 0$  Stop ;

(b) Sinon :

– Nouvelle solution :

$$\begin{cases} \alpha_k = -\frac{r_k^T p_k}{p_k^T A p_k}; \\ x_{k+1} = x_k + \alpha_k p_k; \\ r_{k+1} = Ax_{k+1} - b; \end{cases}$$

– Nouvelle direction :

$$\begin{cases} \beta_{k+1} = \frac{r_{k+1}^T A p_k}{p_k^T A p_k}; \\ p_{k+1} = -r_{k+1} + \beta_{k+1} p_k; \end{cases}$$

(c)  $k := k + 1$ .

**Théorème 3.4** *Supposons que le  $k$ -ième itéré, généré par la méthode du gradient conjugué,*

### 3.3.2 Méthode du gradient conjugué

n'est pas le point  $x^*$ . Alors les quatre propriétés suivantes sont vraies :

$$\begin{aligned} r_k^T &= 0, \text{ pour } i = 0, \dots, k-1 \\ \text{span} \{r_0, r_1, \dots, r_k\} &= \text{span} \{r_0, Ar_0, \dots, A^k r_0\}, \\ \text{span} \{p_0, p_1, \dots, p_k\} &= \text{span} \{r_0, Ar_0, \dots, A^k r_0\}, \\ p_k^T Ap_i &= 0, \text{ pour } i = 0, \dots, k-1. \end{aligned}$$

Voilà pourquoi la séquence  $\{x_k\}$  converge vers  $x^*$  en au plus  $n$  pas.

**Preuve.** Nous allons démontrer les 3 dernières propriétés par récurrence. Ce qui concerne l'ancrage, les propriétés 2 et 3 sont trivialement vraies pour  $k = 0$  et par constriction, la propriété 4 restent vraies pour  $k + 1$ . Par ces hypothèses, on a que

$$r_k \in \text{span} \{r_0, Ar_0, \dots, A^k r_0\} \text{ et } p_k \in \text{span} \{r_0, Ar_0, \dots, A^k r_0\}$$

En multipliant cette deuxième expression par  $A$  et en se rappelant que  $r_{k+1} = r_k + \alpha_k Ap_k$ , on voit facilement que

$$r_{k+1} \in \text{span} \{r_0, Ar_0, \dots, A^{k+1} r_0\}.$$

Grâce à cette constatation et au moyen de l'hypothèse de récurrence de la deuxième propriété, nous avons

$$\text{span} \{r_0, r_1, \dots, r_{k+1}\} \subset \text{span} \{r_0, Ar_0, \dots, A^{k+1} r_0\}$$

Par l'hypothèse de récurrence pour la troisième propriété, on peut déduire que

$$A^{k+1} r_0 = A(A^k r_0) \in \text{span} \{Ap_0, Ap_1, \dots, Ap_k\}$$

Comme  $r_{k+1} = r_k + \alpha_k Ap_k$ , il en suite que  $Ap_k = \frac{(r_{k+1} - r_k)}{\alpha_k}$  et donc

$$A^{k+1} r_0 \in \text{span} \{r_0, r_1, \dots, r_{k+1}\}$$

Ensemble avec l'hypothèse de récurrence de la deuxième propriété, on trouve que

$$\text{span} \{r_0, Ar_0, \dots, A^{k+1} r_0\} \subset \text{span} \{r_0, r_1, \dots, r_{k+1}\}$$

### 3.3.2 Méthode du gradient conjugué

et donc la deuxième propriété est démontrée. Montrons encore que la troisième propriété reste vraie pour  $k + 1$  :

$$\begin{aligned}
 \text{span} \{p_0, p_1, \dots, p_{k+1}\} &= \text{span} \{p_0, p_1, \dots, p_k, r_{k+1}\}, \quad \text{par l'avant - dernière ligne de l'algorithme} \\
 &= \text{span} \{r_0, Ar_0, \dots, A^k r_0, r_{k+1}\}, \quad \text{par l'hypothèse de récurrence de la troisième propriété} \\
 &= \text{span} \{r_0, r_1, \dots, r_{k+1}\} \quad \text{par la deuxième propriété qu'on vient de démontrer} \\
 &= \text{span} \{r_0, Ar_0, \dots, A^{k+1} r_0\} \quad \text{par la deuxième propriété pour } k + 1
 \end{aligned}$$

Passons à la quatrième propriété à démontrer.

Nous multiplions l'avant dernière ligne de l'algorithme avec  $Ap_i$ ,  $i = 0, 1, \dots, k$  pour obtenir

$$p_{k+1}^T Ap_i = -r_{k+1}^T Ap_i + \beta_{k+1} p_k^T Ap_i \quad (3.19)$$

Par la quatrième ligne de l'algorithme, nous avons que si  $i = k$ , alors  $p_{k+1}^T Ap_i = 0$ . L'hypothèse de récurrence de la quatrième propriété à démontrer nous dit que  $p_0, \dots, p_k$  sont A-conjugués et elle permet ainsi d'utiliser le théorème 3.3. Il en suite que

$$r_{k+1}^T p_i = 0, \quad \text{pour } i = 0, 1, \dots, k$$

Ensuite, après plusieurs applications de la troisième propriété, on trouve pour  $k = 0, \dots, k - 1$  que

$$Ap_i \in A \text{span} \{r_0, Ar_0, \dots, A^i r_0\} = \text{span} \{Ar_0, A^2 r_0, \dots, A^{i+1} r_0\} \subset \text{span} \{p_0, \dots, p_{i+1}\}$$

Des deux dernière expressions trouvées, nous concluons que  $r_{k+1}^T Ap_i = 0$ , pour  $i = 0, 1, \dots, k - 1$ .

Le terme de droite de (3.19) disparaît donc pour  $i = 0, 1, \dots, k - 1$ . Finalement, par le théorème 3.3 et par réarrangement de l'avant dernière ligne de l'algorithme, nous trouons que

$$p_i = -r_i + \beta_i p_{i-1},$$

tel que  $r_i \in \text{span} \{p_i, p_{i-1}\}$  pour tout  $i = 0, 1, \dots, k - 1$ . On conclut que la première affirmation du théorème 3.4 est vérifiée.

Le théorème 3.4, nous dit que les directions  $p_0, \dots, p_{n-1}$  sont effectivement A-conjuguées. D'autant plus, il nous dit que les résidus  $r_i$  sont deux à deux orthogonaux et chaque  $p_k$  et

chaque résidu  $r_k$  sont contenus dans le sous-espace de Krylov de degré  $k$  pour  $r_0$ , défini par

$$K(r_0; k) := \text{span} \{r_0, Ar_0, \dots, A^k r_0\}.$$

### 3.2.3 Vitesse de convergence

Grâce à la troisième propriété du théorème 3.4 et deuxième pas de l'algorithme du gradient conjugué, nous avons que

$$x_{k+1} = x_0 + \alpha_0 p_0 + \dots + \alpha_k p_k = x_0 + \gamma_0 r_0 + \gamma_1 A r_0 + \dots + \gamma_k A^k r_0.$$

On définit maintenant  $P_k^*$  comme étant un polynôme de degré  $k$  avec les coefficients  $\gamma_0, \gamma_1, \dots, \gamma_k$ . Remplaçons notre polynôme dans l'expression ci-dessus :

$$x_{k+1} = x_0 + P_k^*(A)r_0$$

Nous allons voir maintenant que parmi toutes les méthode possibles dont les  $k$  premier pas sont restreints sur le sous-espace de Krylov  $K(r_0; k)$ , l'algorithme du gradient conjugué donne le meilleur résultat en termes de minimisation de la distance vers la solution, si cette distance est mesuré par  $\|\cdot\|_A$ , défini comme

$$\|z\|_A^2 = z^T A z$$

D'après cette définition, on constate que

$$\frac{1}{2} \|x - x^*\|_A^2 = \frac{1}{2} (x - x^*)^T A (x - x^*) = \Phi(x) - \Phi(x^*)$$

Par conséquent le théorème 3.3 nous dit que  $x_{k+1}$  minimise  $\|x - x^*\|_A^2$  sur l'ensemble  $x_0 + \text{span} \{p_0, \dots, p_k\}$ . Comme  $x_{k+1} = x_0 + P_k^*(A)r_0$ , il en suit que  $P_k^*$  est la solution de

$$\min_{P_k} \|x_0 + P_k(A)r_0 - x^*\|_A$$

De plus, nous avons que

$$x_{k+1} - x^* = x_0 + P_k^*(A)r_0 - x^* = [I + P_k^*(A)A](x_0 - x^*).$$

### 3.3.2 Méthode du gradient conjugué

Soient  $0 < \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$  les valeurs propres de  $A$  et  $v_1, \dots, v_n$  les vecteurs propres orthonormés correspondants. Nous avons bien que ces vecteurs propres engendrent  $\mathbb{R}^n$ , donc

$$x_0 - x^* = \sum_{i=1}^n \psi_i v_i$$

et il est aussi facile à voir que les vecteurs propres de  $A$  sont aussi des vecteurs propres de  $P_k(A)$  pour tout polynôme. Nous pouvons donc écrire que  $P_k(A)v_i = P_k(\lambda_i)v_i$ ,  $i = 1, 2, \dots, n$  et donc

$$x_{k+1} - x^* = \sum_{i=1}^n [1 + \lambda_i P_k^*(\lambda_i)] \psi_i v_i$$

La norme définie ci-dessus nous permet de dire que

$$\begin{aligned} \|x_{k+1} - x^*\|_A^2 &= \min_{p_k} \sum_{i=1}^n \lambda_i [1 + \lambda_i P_k(\lambda_i)]^2 \psi_i^2 \\ &\leq \min_{p_k} \max_{1 \leq i \leq n} [1 + \lambda_i P_k(\lambda_i)]^2 \left( \sum_{j=1}^n \lambda_j \psi_j^2 \right) \\ &= \min_{p_k} \max_{1 \leq i \leq n} [1 + \lambda_i P_k(\lambda_i)]^2 \|x_0 - x^*\|_A^2, \end{aligned}$$

sachant que  $\|x_0 - x^*\|_A^2 = \sum_{j=1}^n \lambda_j \psi_j^2$ . Cette dernière expression nous permet de quantifier le taux de convergence de la méthode du gradient conjugué.

**Théorème 3.5** *Si  $A$  a seulement  $r$  (avec  $r < n$ ) valeurs propres distinctes, alors l'algorithme du gradient conjugué arrive à la solution en au plus  $r$  itérations.*

**Preuve.** Soient  $\tau_1 < \dots < \tau_r$  les valeurs distinctes des valeurs propres  $\lambda_1, \dots, \lambda_n$ .

Définissons maintenant  $Q_r(\lambda)$  par

$$Q_r(\lambda) = \frac{(-1)^r}{\tau_1 \tau_2 \dots \tau_r} (\lambda - \tau_1) \dots (\lambda - \tau_r)$$

Il est facile à voir que  $Q_r(\lambda) - 1$  est un polynôme de degré  $r$  avec une racine  $\lambda = 0$ . Définissons ensuite  $\bar{P}_{r-1}$ , un polynôme de degré  $r-1$  par

$$\bar{P}_{r-1}(\lambda) = (Q_r(\lambda) - 1)/\lambda$$

En posant  $k = r - 1$  dans  $\min_{p_k} \max_{1 \leq i \leq n} [1 + \lambda_i P_k(\lambda_i)]^2$ , nous obtenons

$$0 \leq \min_{p_{r-1}} \max_{1 \leq i \leq n} [1 + \lambda_i P_{r-1}(\lambda_i)]^2 \leq \max_{1 \leq i \leq n} [1 + \lambda_i P_{r-1}(\lambda_i)]^2 = \max_{1 \leq i \leq n} Q_r(\lambda_i) = 0$$

Donc, pour  $k = r - 1$  on trouve que  $\|x_r - x^*\|_A = 0$  et alors  $x_r = x^*$ .

**Théorème 3.6** *Si  $A$  a des valeurs propres  $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$ , nous avons que*

$$\|x_{k+1} - x^*\|_A^2 \leq \left( \frac{\lambda_{n-k} - \lambda_1}{\lambda_{n-k} + \lambda_1} \right)^2 \|x_0 - x^*\|_A^2$$

**Preuve.** On donne juste une idée comment ce résultat a été trouvé. On choisit un polynôme  $\bar{P}_k$  de degré  $k$  tel que le polynôme  $Q_{k+1}(\lambda) = 1 + \lambda \bar{P}_k(\lambda)$  a comme racines les plus grandes valeurs propres  $\lambda_n, \lambda_{n-1}, \dots, \lambda_{n-k+1}$  et le point milieu entre  $\lambda_1$  et  $\lambda_{n-k}$ . On peut montrer que la valeur maximale atteinte par  $Q_{k+1}$  sur  $\lambda_1, \dots, \lambda_{n-k+1}$  est  $(\lambda_{n-k} - \lambda_1)/(\lambda_{n-k} + \lambda_1)$ .

Le théorème 3.6 nous permet de prédire le comportement de la méthode du gradient conjugué. Supposons par exemple que les valeurs propres de  $A$  consistent en  $m$  grandes valeurs propres et les  $n - m$  valeurs restantes sont situées autour de 1. Si nous définissons  $\epsilon = \lambda_{n-k} - \lambda_1$ , le théorème 3.6 nous dit qu'après  $m + 1$  pas de la méthode du gradient conjugué, nous avons que

$$\|x_{m+1} - x^*\|_A \approx \epsilon \|x_0 - x^*\|_A.$$

Pour une petite valeur de  $\epsilon$  après  $m + 1$  pas, la méthode va nous donner une bonne approximation de la solution. Mais attention, le théorème 3.6 nous donne une borne supérieure. Il se peut que la méthode nous donne déjà des bons résultats après les premières itérations. Il est généralement vrai que si les valeurs propres apparaissent en  $r$  groupe, alors la méthode du gradient conjugué résoud approximativement le problème après  $r$  pas.

Une autre expression de convergence pour la méthode du gradient conjugué est celle basée sur le nombre de condition spectral. Celui-ci est défini par

$$K = \|A\|_2 \cdot \|A^{-1}\|_2 = \frac{\lambda_n}{\lambda_1}.$$

On arrive ensuite à montrer que

$$\|x_k - x^*\|_A \leq \left( \frac{\sqrt{K} - 1}{\sqrt{K} + 1} \right)^{2k} \|x_0 - x^*\|_A.$$

Cette expression est plus approximative que celle du théorème 3.6 et donne souvent d'importantes surestimation de l'erreur. Mais en réalité, souvent, on n'a pas beaucoup d'information sur  $A$  et ses valeurs propres. Cette dernière méthode ne demande que les valeurs propres extrêmes ou des approximations de celles-ci.

### 3.2.4 Méthode du gradient conjugué pour l'équation normale

La méthode du gradient conjugué peut être utilisée pour résoudre les problèmes (2.17). On prend à la place de  $A$  dans (3.13) la matrice  $A^T A$ , et à la place de  $b$  le vecteur  $A^T \hat{z}$ . donc la méthode du gradient conjugué pour résoudre le problème de moindres carrés (2.17) est donnée par l'algorithme suivant :

**Algorithme 3.2** (*Algorithme du gradient conjugué pour l'équation normale*)

**Initialisation**  $k = 0$  : choix de  $x_0 \in \mathbb{R}^n$  et posons  $r_0 = A^T A x_0 - A^T \hat{z}$  et  $p_0 = -r_0$ ;

**Itération**  $k$  :

(a) Si  $r_k = 0$  Stop;

(b) Sinon :

– Nouvelle solution :

$$\begin{cases} \alpha_k = -\frac{r_k^T p_k}{p_k^T A^T A p_k}; \\ x_{k+1} = x_k + \alpha_k p_k; \\ r_{k+1} = A^T A x_{k+1} - A^T \hat{z}; \end{cases}$$

– Nouvelle direction :

$$\begin{cases} \beta_{k+1} = \frac{r_{k+1}^T A^T A p_k}{p_k^T A^T A p_k}; \\ p_{k+1} = -r_{k+1} + \beta_{k+1} p_k; \end{cases}$$

(c)  $k := k + 1$ .

---

## Conclusion générale

Dans ce mémoire, nous avons examiné deux méthodes itératives pour les problèmes inverses linéaires :

- 1) La première est la méthode de Landweber, qu'est construite une suite de solutions approchées converge vers la solution désirée (dans le cas non bruité). Dans le contexte des problèmes inverses linéaires en présence de bruit, la suite construite ne converge pas, en général, vers une solution du problème de départ. Il est, encore une fois, nécessaire de régulariser le processus itératif, et c'est l'indice d'itération lui-même qui joue le rôle de paramètre de régularisation. En d'autres termes, il convient d'arrêter les itérations plus tôt qu'on ne le ferait dans un cas non bruité. Malheureusement, la méthode de Landweber converge trop lentement.
- 2) La deuxième est la méthode du gradient conjugué et ses variantes qui est la plus employée en pratique.

---

## Bibliographie

- [1] A. KOLMOGOROV, S. FOMINE, Élément de la théorie des fonctions et de l'analyse fonctionnelle, MIR-MOSCOU, 1973.
- [2] A. Quarteroni, R. Sacco, F. Saleri. Méthodes Numériques-Algorithmes, analyse et applications, Springer-Verlag Italia, Milano 2004.
- [3] Delft. Nederland, Conjugate Gradients and Conjugate Residuals type methods for solving Least Squares problems from Tomography, Tamara Kloek 2012.
- [4] D. KAUTH, Optimisation numérique-Méthodes du gradient conjugué linéaire, 2009.
- [5] E. fricain, Analyse fonctionnelle et théorie des opérateurs-cours et exercices-2009-2010.
- [6] F. Jedrzejewski, Introduction aux méthodes numériques, Springer-Verlag France, Paris 2005.
- [7] F. Arif, Problèmes inverses linéaires et application à la Tomographie à rayons X, Mémoire de Master à l'Université ABDELHAMID IBN BADIS de Mostaganem, 2012 ; Algérie.
- [8] H. Brezis, Analyse fonctionnelle-Théorie et applications, Masson, Paris 1987.
- [9] H. Brezis, Functional Analysis, Sobolev Spaces and Partial Differential Equations, Springer Science+Business Media, LLC 2011.
- [10] J. Erhel, N. Nassif, B. Philippe, Calcul matriciel et systèmes linéaires, Beyrouth, Liban 2004.

- [11] J. Hadamard. Lectures on Cauchy's problem in linear partial Differential Equation. Yale University press. 1923.
- [12] J. B. Conway, A Course in Functional Analysis, Springer, 1990.
- [13] J. Richard Shewchuk, An Introduction to the Conjugate Gradient Method Without the Agonizing Pain, 1994.
- [14] L. Landweber. An iteration formula for Fredholm integral equations of the first kind. Amer. J. Math., 73 :615–624, 1951.
- [15] M. Hanke. Conjugate Gradient Type Methods for Ill-Posed Problems, volume 327 of Pitman Research Notes in Mathematics. Longman, Harlow, 1995.
- [16] M. KERN, Cours de DEA, Problèmes inverses, Ecole supérieure d'ingénieurs, LEONARD DE VINCI, INRIA, 2012.
- [17] M.L KERN, Cours de DEA, Problèmes inverses, Ecole supérieure d'ingénieurs, LEONARD DE VINCI, INRIA, 2003.
- [18] T. Hohage, Lecture notes on Inverse Problems, University of *Göttingen*, Summer 2002.
- [19] S. Maingot, D. Manceau, Théorie spectrale.
- [20] W. HENGARTNER, M. LAMBERT, C. REKISCHER. Introduction à L'analyse Fonctionnelle. Sillery, Québec, 1981.
- [21] [http ://www.math-uvsq.fr](http://www.math-uvsq.fr).





**Résumé.** Dans ce mémoire, nous avons examiné deux méthodes itératives pour les problèmes inverses linéaires : la première est la méthode de Landweber, qu'est construite une suite de solutions approchées converge vers la solution désirée (dans le cas non bruité). Dans le contexte des problèmes inverses linéaires en présence de bruit, la suite construite ne converge pas, en général, vers une solution du problème de départ. Il est, encore une fois, nécessaire de régulariser le processus itératif, et c'est l'indice d'itération lui-même qui joue le rôle de paramètre de régularisation. En d'autres termes, il convient d'arrêter les itérations plus tôt qu'on ne le ferait dans un cas non bruité. Malheureusement, la méthode de Landweber converge trop lentement. La deuxième est la méthode du gradient conjugué et ses variantes qui est la plus employée en pratique.

Mots clés : Problème inverse linéaire, Problème mal posé, Méthode itérative, Méthode de Landweeber, Méthode du Gradient conjugué.

**Abstract.** In this memoir, we examined two iterative methods for linear inverse problems : the first is the Landweber method, what build a sequence of approximate solutions converges to the desired solution (in the noiseless case). In the context of linear inverse problems with noise, subsequently built does not converge, in general, to a solution of the original problem. It is again necessary to regularize the iterative process, and it is the iteration itself plays the role of regularization parameter index. In other words, it is necessary to stop the iterations sooner than they would in a non-noisy case. Unfortunately, the Landweber method converges too slowly. The second is the conjugate gradient method and its variants is the most used in practice.

Keywords : Linear inverse problem, ill-posed problem, iterative method, Landweeber method, the conjugate gradient method.