

## Chapitre 5

# Moindres carrés et inversion généralisée

Nous avons vu qu'un problème est dit mal-posé s'il ne satisfait pas à une des trois conditions : existence, unicité et stabilité. Lorsque la difficulté vient de la non unicité de la solution l'inversion généralisée permet d'y remédier. En effet, la résolution de l'équation  $\mathbf{g} = \mathbf{H}\mathbf{f}$  présente trois possibilités :

- Il y a une solution au problème ;
- Il y a une infinité de solutions au problème ;
- Il n'y pas de solution au problème.

Dans le premier cas, il ne reste qu'à calculer l'inverse. Dans le second cas, le problème devient celui d'un choix parmi l'ensemble des solutions suivant un critère (par exemple, la norme minimale). Dans le troisième cas, on souhaiterait aussi pouvoir définir une solution, cela est possible en définissant un critère du type distance dans l'espace  $G$  entre  $\mathbf{g}$  et  $\mathbf{H}\mathbf{f}$ , (par exemple la distance euclidienne), et en choisissant comme solution celle qui la minimise. Dans le cas où la distance choisie est une distance euclidienne la solution est dite *au sens des moindres carrés* et nous verrons qu'il y a un lien entre la solution au sens des MC et celle de l'inverse généralisée.

## 5.1 Définition en dimension infinie

Soit  $\mathbf{H} : F \mapsto G$ ,

$\mathbf{P}$  le projecteur orthogonal de  $F$  sur  $\mathcal{Ker}(\mathbf{H})^\perp = \overline{\mathcal{Im}(\mathbf{H}^*)}$  et

$\mathbf{Q}$  le projecteur orthogonal de  $G$  sur  $\mathcal{Ker}(\mathbf{H}^*)^\perp = \overline{\mathcal{Im}(\mathbf{H})}$ . On a alors

$$\mathbf{H}\mathbf{f} = \mathbf{H}\mathbf{P}\mathbf{f} \quad \mathbf{H}^*\mathbf{g} = \mathbf{H}^*\mathbf{Q}\mathbf{g}$$

**Définition 1** *L'inverse généralisée  $\mathbf{H}^+$  de  $\mathbf{H}$  est l'opérateur dont le domaine est  $\mathcal{D}(\mathbf{H}^+) = \mathcal{Im}(\mathbf{H}) \oplus \mathcal{Im}(\mathbf{H})^\perp$ , tel que  $\mathbf{H}^+\mathbf{g} = \mathbf{f}_0 \in F$ , où  $\mathbf{f}_0 \in S = \{\mathbf{f} \in F : \mathbf{H}\mathbf{f} = \mathbf{Q}\mathbf{g}\}$  et  $\|\mathbf{f}_0\| \leq \|\mathbf{f}\|$ ,  $\forall \mathbf{f} \in S, \mathbf{f} \neq \mathbf{f}_0$ .*

Notons que

$$\|\mathbf{g} - \mathbf{H}\mathbf{f}\|^2 = \|\mathbf{Q}\mathbf{g} - \mathbf{H}\mathbf{f}\|^2 + \|\mathbf{g} - \mathbf{Q}\mathbf{g}\|^2$$

Si  $\mathbf{g} \in \mathcal{Im}(\mathbf{H}) \oplus \mathcal{Im}(\mathbf{H})^\perp$ , l'ensemble des solutions des MC est non-vide, fermé et convexe. Il existe donc un élément unique de norme minimale, qui est  $\mathbf{H}^+\mathbf{g}$ .

**Limitations :**

- $\forall \mathbf{g} \in G$ , il existe une solution inverse généralisée ssi  $\mathcal{Im}(\mathbf{H})$  est fermée.
- En général  $\mathbf{H}$  est compact,  $\mathcal{Im}(\mathbf{H})$  n'est pas fermé, sauf si  $\mathbf{H}$  est de dimension finie.
- Existence : Non  $\mathcal{Im}(\mathbf{H}) \oplus \mathcal{Im}(\mathbf{H})^\perp \neq G$
- Unicité : Oui
- Stabilité : Non  $\mathcal{Im}(\mathbf{H}) \neq \overline{\mathcal{Im}(\mathbf{H})}$

## 5.2 Définition en dimension finie

Considérons l'équation

$$\mathbf{g} = \mathbf{H} \mathbf{f}$$

où  $\mathbf{H}$  est une matrice de dimension  $[m, n]$ .

**Définition 2 (Quasi-solution)** On appelle  $\hat{\mathbf{f}}$  une quasi-solution de l'équation  $\mathbf{g} = \mathbf{H} \mathbf{f}$  une solution qui minimise une distance  $\Delta(\mathbf{g}, \mathbf{H} \mathbf{f})$  entre  $\mathbf{g}$  et  $\mathbf{H} \mathbf{f}$  :

$$\hat{\mathbf{f}} = \arg \min_{\mathbf{f}} \{ \Delta(\mathbf{g}, \mathbf{H} \mathbf{f}) \}$$

Lorsque la distance choisie est une distance euclidienne  $\Delta(\mathbf{g}, \mathbf{H} \mathbf{f}) = \|\mathbf{g} - \mathbf{H} \mathbf{f}\|^2$ , la solution ainsi définie est dite "au sens des moindres carrés".

**Définition 3 (Moindre carré)** On appelle  $\hat{\mathbf{f}}$  une solution au sens des moindres carrés de l'équation  $\mathbf{g} = \mathbf{H} \mathbf{f}$  une solution qui minimise la norme euclidienne  $\|\mathbf{g} - \mathbf{H} \mathbf{f}\|^2$  :

$$\hat{\mathbf{f}} = \arg \min_{\mathbf{f}} \{ \|\mathbf{g} - \mathbf{H} \mathbf{f}\|^2 \}.$$

Cet ensemble coïncide avec l'ensemble des solutions de l'équation normale

$$[\mathbf{H}^t \mathbf{H}] \hat{\mathbf{f}} = \mathbf{H}^t \mathbf{g}.$$

- Si  $M = N$  et  $\text{rang}\{\mathbf{H}\} = N$ , alors  $\hat{\mathbf{f}} = \mathbf{f} = \mathbf{H}^{-1} \mathbf{g}$ .
- Si  $N \leq M$  et  $\text{rang}\{\mathbf{H}\} = N$ , alors  $[\mathbf{H}^t \mathbf{H}]$  est de rang plein et  $\hat{\mathbf{f}} = [\mathbf{H}^t \mathbf{H}]^{-1} \mathbf{H}^t \mathbf{g}$ .
- Si  $M < N$ , alors  $[\mathbf{H}^t \mathbf{H}]$  est forcément singulière. Elle a un rang au maximum égal à  $M$ , et l'équation (1) n'a pas de solution unique (Elle a une infinité de solutions).

**Définition 4 (Inverse généralisée)** On appelle  $\mathbf{f}^+ = \mathbf{H}^+ \mathbf{g}$  la solution inverse généralisée de l'équation  $\mathbf{g} = \mathbf{H} \mathbf{f}$  la solution MC de norme minimale de l'équation normale  $[\mathbf{H}^t \mathbf{H}] \hat{\mathbf{f}} = \mathbf{H}^t \mathbf{g}$ . Si  $\mathbf{g} \in \text{Im}(\mathbf{H})$  alors cette solution existe et elle est unique. Elle est donnée par :

$$\mathbf{f}^+ = \mathbf{H}^+ \mathbf{g}.$$

où  $\mathbf{H}^+$  est l'inverse généralisée de  $\mathbf{H}$ .

Si  $\text{rang}\{\mathbf{H}\} = N$  alors  $\mathbf{H}^+ = [\mathbf{H}^t \mathbf{H}]^{-1} \mathbf{H}^t$  et si  $\text{rang}\{\mathbf{H}\} = M$  alors  $\mathbf{H}^+ = \mathbf{H}^t [\mathbf{H} \mathbf{H}^t]^{-1}$ .

Pour être complet nous mentionnons deux autres définitions pour l'inverse généralisée d'une matrice.

**Définition 5 (MOORE)** Une matrice  $\mathbf{G}$  de dimension  $[N, M]$  est dite inverse généralisée de la matrice  $\mathbf{H}$  si

$$\mathbf{H} \mathbf{G} = \mathbf{P}_A \quad \text{et} \quad \mathbf{G} \mathbf{H} = \mathbf{P}_G$$

où  $\mathbf{P}_A$  est l'opérateur (matrice) de projection des vecteurs de l'espace  $\mathbb{R}^m$  sur  $\mu(\mathbf{H})$  et  $\mathbf{P}_G$  est l'opérateur (matrice) de projection des vecteurs de l'espace  $\mathbb{R}^n$  sur  $\mu(\mathbf{G})$ .

**Définition 6** (PENROSE) Une matrice  $\mathbf{G}$  de dimension  $[N, M]$  est dite inverse généralisée de la matrice  $\mathbf{H}$  si

$$\mathbf{HGH} = \mathbf{H}, \quad (\mathbf{HG})^t = \mathbf{HG}, \quad \mathbf{GHG} = \mathbf{G} \quad \text{et} \quad (\mathbf{GH})^t = \mathbf{GH}$$

Ces définitions sont identiques si les normes associées aux espaces  $\mathbf{R}^m$  et  $\mathbf{R}^n$  sont du type euclidien  $\|\mathbf{f}\| = \langle \mathbf{f}, \mathbf{f} \rangle^{1/2}$ .

Supposons que  $\mathbf{H}\mathbf{f} = \mathbf{g}$  et calculons l'erreur quadratique moyenne entre une estimation  $\hat{\mathbf{f}}$  linéaire quelconque  $\hat{\mathbf{f}} = \mathbf{G}^+\mathbf{g}$  de la solution et  $\mathbf{f}$ :

$$EQM = [\mathbf{f} - \hat{\mathbf{f}}]^t[\mathbf{f} - \hat{\mathbf{f}}] = \text{Tr} \left\{ [\mathbf{f} - \hat{\mathbf{f}}][\mathbf{f} - \hat{\mathbf{f}}]^t \right\} = \text{Tr} \left\{ \mathbf{f}\mathbf{f}^t[\mathbf{I} - \mathbf{H}^+\mathbf{H}]^t \right\}$$

Il est alors facile de vérifier que cette erreur est minimale si  $\hat{\mathbf{f}} = \mathbf{H}^+\mathbf{g}$  et qu'elle est donnée par

$$EQM = [\mathbf{f} - \hat{\mathbf{f}}]^t[\mathbf{f} - \hat{\mathbf{f}}] = \text{Tr} \left\{ \mathbf{f}\mathbf{f}^t[\mathbf{I} - \mathbf{H}^+\mathbf{H}]^t \right\}$$

l'estimation est parfaite si  $\mathbf{H}^+\mathbf{H} = \mathbf{I}$ .

En résumé, lorsque nous voulons calculer une solution au sens inverse généralisée de l'équation  $\mathbf{g} = \mathbf{H}\mathbf{f}$  avec  $\mathbf{H}$  une matrice de dimensions  $[M, N]$ , quatre situations se présentent :

- si  $M = N$  et  $\text{rang} \{ \mathbf{H} \} = N$  alors  $\mathbf{H}^+ = \mathbf{H}^{-1}$
- si  $M > N$  et  $\text{rang} \{ \mathbf{H} \} = N$  alors  $\mathbf{H}^+ = (\mathbf{H}^t\mathbf{H})^{-1}\mathbf{H}^t$
- si  $M < N$  et  $\text{rang} \{ \mathbf{H} \} = M$  alors  $\mathbf{H}^+ = \mathbf{H}^t(\mathbf{H}\mathbf{H}^t)^{-1}$
- si  $\text{rang} \{ \mathbf{H} \} = K < \inf(M, N)$  alors  $\left\{ \begin{array}{l} \text{Décomp. en valeurs singulières} \\ \text{Méthodes itératives} \end{array} \right.$

### 5.3 Algorithmes d'inversion généralisée

### 5.4 Décomposition en valeurs singulières

Considérons l'équation

$$\mathbf{g} = \mathbf{H} \mathbf{f}$$

et notons  $\{\mathbf{v}_j, j = 1, \dots, K\}$  les vecteurs propres de  $\mathbf{H}^t \mathbf{H}$ ,  $\{\mathbf{u}_i, i = 1, \dots, K\}$  les vecteurs propres de  $\mathbf{H} \mathbf{H}^t$  et  $\lambda_j^2$  les valeurs propres correspondantes. Nous avons alors :

$$\mathbf{H}^t \mathbf{H} \mathbf{v}_j = \lambda_j^2 \mathbf{v}_j, \quad j = 1, \dots, K$$

$$\mathbf{H} \mathbf{H}^t \mathbf{u}_i = \lambda_i^2 \mathbf{u}_i, \quad i = 1, \dots, K$$

$$\mathbf{H} \mathbf{v}_j = \lambda_j \mathbf{u}_j, \quad j = 1, \dots, K$$

$$\mathbf{H}^t \mathbf{u}_i = \lambda_i \mathbf{v}_i, \quad i = 1, \dots, K$$

$$\mathbf{H} = \mathbf{U} \mathbf{\Lambda} \mathbf{V}^t$$

$$\mathbf{U} = (\mathbf{u}_1 : \mathbf{u}_2 : \dots : \mathbf{u}_K), \quad \mathbf{V} = (\mathbf{v}_1 : \mathbf{v}_2 : \dots : \mathbf{v}_K), \quad \mathbf{\Lambda} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_K)$$

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_K, \quad K \leq \inf(M, N).$$

La solution inverse généralisée au sens du Moore Penrose est définie par

$$\mathbf{f}^+ = \mathbf{H}^+ \mathbf{g}, \quad \text{avec} \quad \mathbf{H}^+ = \mathbf{V} \mathbf{\Lambda}^+ \mathbf{U}^t$$

et où

$$\mathbf{\Lambda}^+ = \text{diag} \{ \alpha_i \}, \quad \begin{cases} \alpha_i = \frac{1}{\lambda_i} & \text{si } \lambda_i \neq 0 \\ \alpha_i = 0 & \text{si } \lambda_i \simeq 0 \end{cases}$$

On a aussi les relations suivantes :

$$\text{Ker}(\mathbf{H}) = \mathbf{k}(z) = \mathbf{V}(\mathbf{I} - \Lambda^+ \Lambda)z, \quad \forall z \in \mathbb{R}^n$$

Les composantes d'indices 0 à  $K$  des vecteurs  $\mathbf{k}(z)$  sont nulles. On peut alors écrire :

$$\mathbf{k}(z) = \sum_{k=K+1}^N z_k \mathbf{v}_k \quad \forall z \in \mathbb{R}^n$$

et définir ainsi l'ensemble  $S$  des solutions de l'équation  $\mathbf{g} = \mathbf{H}\mathbf{f}$  par

$$S = \{ \mathbf{f} : \mathbf{f} = \mathbf{f}^+ + \mathbf{k}(z) \quad \forall z \in \mathbb{R}^n \}$$

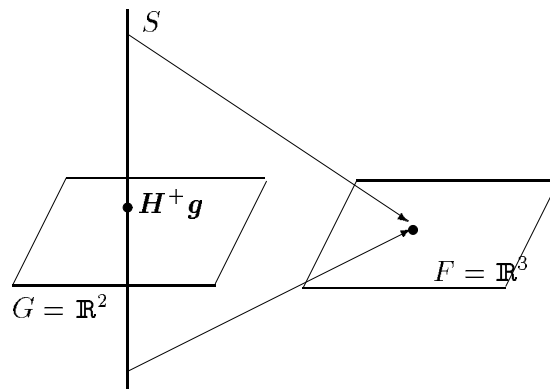


FIG. 5.1 - Inversion généralisée et espace des solutions.

### 5.4.1 Méthode de BIALY

Cette méthode est basée sur le théorème suivant :

**Théorème 1** Soit  $\mathbf{H} : H \mapsto H$  un opérateur borné et non négatif. Soit  $\mathbf{P}$  l'opérateur de projection orthogonale sur  $\text{Ker}(\mathbf{H})$ . Alors, s'il existe une solution pour  $\mathbf{H}\mathbf{f} = \mathbf{g}$ , pour tout  $\mathbf{g} \in H$ ,  $\mathbf{f}_0 \in H$  la suite des itérations

$$\mathbf{f}_{n+1} = \mathbf{f}_n + \alpha(\mathbf{g} - \mathbf{H}\mathbf{f}_n), \quad \text{avec } 0 < \alpha < 2/\|\mathbf{H}\|$$

converge vers  $\mathbf{P}\mathbf{f}_0 + \mathbf{f}_{IG}$ , où  $\mathbf{f}_{IG}$  est la solution IG de  $\mathbf{H}\mathbf{f} = \mathbf{g}$ .

- si  $\mathbf{f}_0 = \mathbf{0}$ , alors  $\lim_{n \rightarrow \infty} \mathbf{f}_n = \mathbf{f}_{IG}$
- $\mathbf{f}_{IG}$  est la solution de norme minimale de  $\mathbf{H}^t \mathbf{H}\mathbf{f} = \mathbf{H}^t \mathbf{g}$ .

### 5.4.2 Méthode de LANDWEBER

Cette méthode, qui est une extension de la méthode précédente, est basée sur le théorème suivant :

**Théorème 2** Soient  $\mathbf{H} : H_1 \mapsto H_2$  un opérateur borné,  $\mathbf{g} \in \text{Im}(\mathbf{H}) + \text{Im}(\mathbf{H})^\perp$ . Alors la suite des itérations

$$\begin{aligned} \mathbf{f}_0 &= \mathbf{0} \\ \mathbf{f}_{n+1} &= \mathbf{f}_n + \alpha \mathbf{H}^t(\mathbf{g} - \mathbf{H}\mathbf{f}_n), \quad \text{avec } 0 < \alpha < 2/\|\mathbf{H}^t \mathbf{H}\| \end{aligned}$$

converge vers la solution IG de  $\mathbf{H}\mathbf{f} = \mathbf{g}$ .

Notons que cet algorithme peut être considéré comme un algorithme du type gradient pour la minimisation du critère des MC  $J(\mathbf{f}) = \|\mathbf{g} - \mathbf{H}\mathbf{f}\|^2$ . En effet, on peut remarquer que le gradient de ce critère est  $\nabla J(\mathbf{f}) = -2\mathbf{H}^t(\mathbf{g} - \mathbf{H}\mathbf{f})$ .

### 5.4.3 Méthode de VAN-CITTERT

Cette méthode est une extension de la méthode précédente et permet d'apporter des contraintes supplémentaires sur la solution. En effet, si  $\mathbf{T}$  est un opérateur traduisant ces contraintes on peut envisager l'algorithme suivant :

$$\begin{cases} \mathbf{f}_0 = \mathbf{0} \\ \mathbf{f}_{n+1} = \mathbf{f}_n + \alpha \mathbf{H}^t(\mathbf{g} - \mathbf{H}\mathbf{T}\mathbf{f}_n), \quad \text{avec } 0 < \alpha < 2/\|\mathbf{T}^t \mathbf{H}^t \mathbf{H}\mathbf{T}\| \end{cases}$$

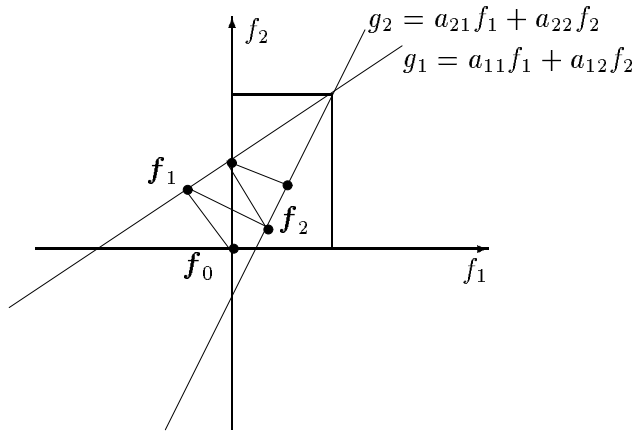
#### 5.4.4 Méthode de KACZMARZ

L'idée de base dans cette méthode est d'utiliser les données  $g_i$  une par une. Ainsi, une mise en œuvre en ligne serait possible. Dans une relation linéaire  $\mathbf{g} = \mathbf{H}\mathbf{f}$  chaque équation  $g_i = \sum_{j=1}^N h_{ij} f_j = \langle \mathbf{h}_i, \mathbf{f} \rangle$  définit un sous-espace dans l'espace des solutions. Cet algorithme recherche une solution à l'itération  $k+1$  en projetant la solution à l'itération  $k$  sur le sous-espace créé par l'équation  $g_k = \sum_{j=1}^N h_{kj} f_j$ .

$$\begin{cases} \mathbf{f}_0 = \mathbf{0} \\ \mathbf{f}_{n+1} = \mathbf{f}_n + \frac{g_i - \langle \mathbf{f}_n, \mathbf{h}_i \rangle}{\langle \mathbf{h}_i, \mathbf{h}_i \rangle} \mathbf{h}_i \quad i = 1, 2, \dots, M, 1, 2, \dots, M, 1, 2 \end{cases}$$

où  $\mathbf{h}_i$  est la  $i^{\text{ème}}$  ligne de la matrice  $\mathbf{H}$ .

Pour bien comprendre l'essentiel de cet algorithme, considérons le cas d'une matrice  $\mathbf{H}$  de dimensions  $(2 \times 2)$ . Dans ce cas, chaque équation décrit une ligne droite dans l'espace des solutions  $(f_1, f_2)$  et la figure ci-dessous montre alors le principe de cet algorithme.



Évidemment, lorsqu'on veut imposer d'autres contraintes, comme par exemple la positivité, on peut introduire au cours de ces itérations. Ainsi, si on note par  $p_i(\mathbf{f})$ ,  $i = 1, \dots, m$  les  $m$  contraintes supplémentaires que l'on veut imposer à la solution, l'algorithme s'écrit :

$$\begin{cases} \mathbf{f}_0 = \mathbf{0} \\ \mathbf{f}_{n+1} = P(\mathbf{f}_n) = p_m(\dots p_2(p_1(\mathbf{f})) \dots) \\ p_i(\mathbf{f}) = \mathbf{f}_n + \frac{g_i - \langle \mathbf{f}_n, \mathbf{h}_i \rangle}{\langle \mathbf{h}_i, \mathbf{h}_i \rangle} \mathbf{h}_i \quad i = 1, 2, \dots, M, 1, 2, \dots, M, 1, 2 \end{cases}$$