

MAXIMUM D'ENTROPIE ET PROBLÈMES INVERSES

Ali MOHAMMAD-DJAFARI

Laboratoire des signaux et systèmes

Unité mixte de recherche 008506

Supélec, Plateau de Moulon

91192 Gif-sur-Yvette Cedex, FRANCE.

djafari@lss.supelec.fr

<http://djafari.free.fr>

<http://www.lss.supelec.fr/perso/djafari>

PROBLÈMES INVERSES : FORMALISATION

$$\mathcal{H} (\quad \mathbf{g} \quad , \quad \mathbf{f} \quad , \quad \mathbf{z} \quad , \quad \boldsymbol{\epsilon} \quad) = 0$$

modèle
mesures
grandeurs
inconnues
grandeurs inconnues
intermédiaires
erreurs
et bruit

Cas particuliers :

- Modèle implicite avec relation entre \mathbf{f} et \mathbf{z} :

$$\begin{cases} \mathbf{g} = \mathcal{H}_1(\mathbf{f}, \mathbf{z}) + \boldsymbol{\epsilon} \\ \mathcal{H}_2(\mathbf{f}, \mathbf{z}) = 0 \end{cases}$$

- Modèle non linéaire simple :

$$\mathbf{g} = \mathcal{H}(\mathbf{f}) + \boldsymbol{\epsilon}$$

- Modèle linéaire + bruit additif :

$$\mathbf{g} = \mathcal{H}\mathbf{f} + \boldsymbol{\epsilon}$$

SUIVIE DE FRONT DE FUSION LIQUIDE-SOLIDE

Solide :
$$\frac{\partial T_s}{\partial t} = \alpha_s \left(\frac{\partial^2 T_s}{\partial x^2} + \frac{\partial^2 T_s}{\partial y^2} \right)$$

Liquide :
$$\frac{\partial T_l}{\partial t} = \alpha_l \left(\frac{\partial^2 T_l}{\partial x^2} + \frac{\partial^2 T_l}{\partial y^2} \right)$$

Bilan d'énergie à l'interface :
$$k_s \frac{\partial T_s}{\partial n} - k_l \frac{\partial T_l}{\partial n} = \rho L_f \vec{v} \cdot \vec{n}$$

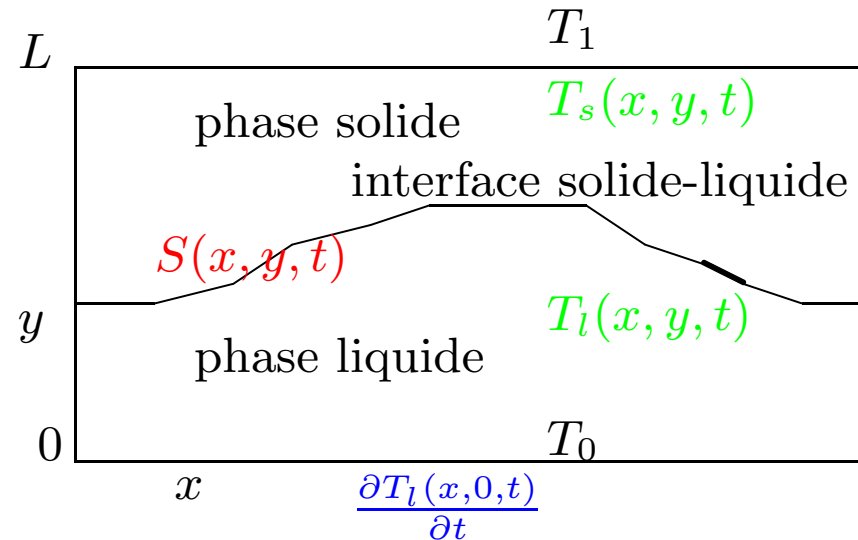
\vec{v} : vitesse de l'interface solide-liquide

\vec{n} : vecteur normal à l'interface

Mesures : flux de chaleur sur la paroi chaude $\frac{\partial T_s(x,0,t)}{\partial t}$

Inconnue : position de l'interface $S(x, y, t)$

Grandeurs intermédiaire : champs de température $T_s(x, y, t)$ et $T_l(x, y, t)$



IMAGERIE À ONDES DIFFRACTÉES

Mesures : champ diffracté par objet $\phi_d(\mathbf{r}_i)$

Grandeur inconnue : $f(\mathbf{r}) = k_0^2(n^2(\mathbf{r}) - 1)$

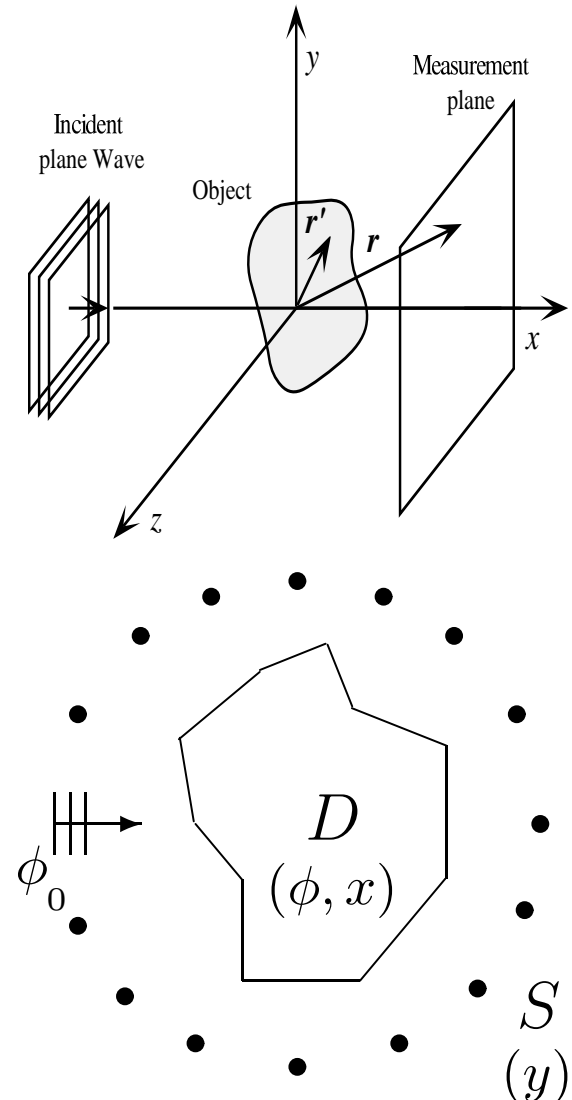
Grandeurs intermédiaires : $\phi(\mathbf{r})$

$$\phi_d(\mathbf{r}_i) = \iint_D G_m(\mathbf{r}_i, \mathbf{r}') \phi(\mathbf{r}') f(\mathbf{r}') d\mathbf{r}', \quad \mathbf{r}_i \in S$$

$$\phi(\mathbf{r}) = \phi_0(\mathbf{r}) + \iint_D G_o(\mathbf{r}, \mathbf{r}') \phi(\mathbf{r}') f(\mathbf{r}') d\mathbf{r}', \quad \mathbf{r} \in D$$

Approximation de Born ($\phi(\mathbf{r}') \simeq \phi_0(\mathbf{r}')$) :

$$\phi_d(\mathbf{r}_i) = \iint_D G_m(\mathbf{r}_i, \mathbf{r}') \phi_0(\mathbf{r}') f(\mathbf{r}') d\mathbf{r}'$$



PROBLÈMES INVERSES LINÉAIRES

$$g(\mathbf{s}_i) = \int f(\mathbf{r}) h(\mathbf{r}, \mathbf{s}_i) d\mathbf{r} + \epsilon(\mathbf{s}_i), \quad i = 1, \dots, M$$

- Débruitage & Interpolation: $g(\mathbf{r}_i) = f(\mathbf{r}) + \epsilon(\mathbf{r}_i)$

- Déconvolution: $g(t_i) = \int f(t') h(t_i - t') dt' + \epsilon(t_i)$

- Restauration d'image :

$$g(x_i, y_j) = \iint f(x', y') h(x_i - x', y_j - y') dx' dy' + \epsilon(x_i, y_j)$$

- Reconstruction d'image :

$$g(r_i, \phi_j) = \iint f(x, y) \delta(r_i - x \cos \phi_j - y \sin \phi_j) dx dy + \epsilon(r_i, \phi_j)$$

- Synthèse de Fourier :

$$g(\Omega_i, \phi_j) = \iint f(x, y) \exp [j(x\Omega_i \cos \phi_j + y\Omega_i \sin \phi_j)] dx dy + \epsilon(\Omega_i, \phi_j)$$

MÉTHODES ALGÈBRIQUES DÉTERMINISTES

$$g(\mathbf{s}_i) = \int \mathbf{f}(\mathbf{r}) h(\mathbf{r}, \mathbf{s}_i) d\mathbf{r} \quad \longrightarrow \text{discrétisation} \longrightarrow \quad \mathbf{g} = \mathbf{H} \mathbf{f} + \epsilon$$

- Rétrojection ou filtre adapté : $\hat{\mathbf{f}} = \mathbf{H}^t \mathbf{g}$, mais $\mathbf{H}^t \mathbf{H} \neq \mathbf{I}$
- Inversion directe lorsque possible: $\hat{\mathbf{f}} = \mathbf{H}^{-1} \mathbf{g}$
- Inverse généralisée:
 - $M > N$ et $\text{rang}\{A\} = N \longrightarrow \hat{\mathbf{f}} = (\mathbf{H}^t \mathbf{H})^{-1} \mathbf{H}^t \mathbf{g}$
 - $M < N$ et $\text{rang}\{A\} = M \longrightarrow \hat{\mathbf{f}} = \mathbf{H}^t (\mathbf{H} \mathbf{H}^t)^{-1} \mathbf{g}$
 - Décomposition en valeurs singulières : $\hat{\mathbf{f}} = \sum_{k=1}^r \frac{\langle \mathbf{u}_k, \mathbf{g} \rangle}{\lambda_k} \mathbf{v}_k$
- Restriction de l'espace des solutions :

$\min \Omega(\mathbf{f})$ s.c. $\mathbf{H} \mathbf{f} = \mathbf{g}$, avec, par exemple

$$\Omega(\mathbf{f}) = \sum_{j=1}^n f_j^2, \quad \Omega(\mathbf{f}) = - \sum_{j=1}^n f_j \log f_j, \quad \Omega(\mathbf{f}) = \sum_{j=1}^n \log f_j, \quad \dots$$

MÉTHODES PROBABILISTES

On se place dans le cadre de l'estimation :

- Maximum de vraisemblance (MV)
- Maximum d'entropie (ME)
- Maximum d'entropie sur la moyenne (MEM)
- Estimation bayésienne (EB)

Avantages :

- Prise en compte de la nature du bruit et l'incertitude du modèle
- Prise en compte d'information *a priori* sur la solution
- Cadre cohérent permettant aussi la détermination des hyperparamètres, marginalisation des paramètres de nuisance, ...

Limitations : Mise en œuvre parfois plus difficile en pratique

ENTROPIE (SHANNON 1945)

- Source X : $\begin{cases} x_1, & x_2, & \dots, & x_n \\ p_1, & p_2, & \dots, & p_n \end{cases} \longrightarrow \begin{cases} \frac{1}{p_1}, & \frac{1}{p_2}, & \dots, & \frac{1}{p_n} \\ \ln \frac{1}{p_1}, & \ln \frac{1}{p_2}, & \dots, & \ln \frac{1}{p_n} \end{cases}$
- Information moyenne de la source :

$$S = \sum_{i=1}^n p_i \ln \left(\frac{1}{p_i} \right) = - \sum_{i=1}^n p_i \ln p_i$$

- Extension au cas continu :

$$S = - \int p(x) \ln p(x) dx$$

- Entropie relative (divergence de Kullback 1959) :

$$K(p : q) = \int p(x) \ln \left(\frac{p(x)}{q(x)} \right) dx$$

- Information moyenne apportée par $p(x)$ par rapport à $q(x)$
- Si $q(x)$ uniforme $\min K(p, q) = \max S(p)$

PRINCIPE DU MAXIMUM D'ENTROPIE (JAYNES 1954)

$$E \{ \phi_k(X) \} = \int \phi_k(x) p(x) dx = d_k \quad k = 1, \dots, K \longrightarrow p(x)?$$

- Une infinité de solutions possibles
- **PME** : parmi ces solutions possibles choisir celle qui a l'entropie maximale

$$\max S = - \int p(x) \ln p(x) dx \quad \text{s.c.} \quad \int \phi_k(x) p(x) dx = d_k \quad k = 1, \dots, K$$

$$\text{Lagrangien :} \quad \mathcal{L} = - \int p(x) \ln p(x) dx - \sum_{k=1}^K \lambda_k \left[\int \phi_k(x) p(x) dx - d_k \right]$$

$$\frac{\partial \mathcal{L}}{\partial p} = 0 \longrightarrow p(x) = \frac{1}{Z(\lambda_1, \dots, \lambda_k)} \exp \left[- \sum_{k=1}^K \lambda_k \phi_k(x) \right]$$

$\lambda_1, \dots, \lambda_k$ multiplicateurs de Lagrange

$$\frac{\partial \ln Z(\lambda_1, \dots, \lambda_k)}{\partial \lambda_k} = d_k \quad k = 1, \dots, K$$

EXEMPLES

$\phi_1(x) = 1$	$x \in [a, b]$	$p(x) = \frac{1}{b-a}$	Uniforme
$\phi_1(x) = x$	$x \in \mathbb{R}_+$	$p(x) = \mu \exp[-\mu x]$	Exponentielle
$\phi_1(x) = x^2$	$x \in \mathbb{R}$	$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{x^2}{2\sigma^2}\right]$	Gaussienne
$\left\{ \begin{array}{l} \phi_1(x) = x \\ \phi_2(x) = x^2 \end{array} \right.$	$x \in \mathbb{R}$	$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x-m)^2}{2\sigma^2}\right]$	Gaussienne
$\left\{ \begin{array}{l} \phi_1(x) = x \\ \phi_2(x) = \ln x \end{array} \right.$	$x \in \mathbb{R}_+$	$p(x) \propto x^{-\lambda} \exp[-\mu x]$	Gamma
$\left\{ \begin{array}{l} \phi_1(x) = \ln x \\ \phi_2(x) = \ln(1-x) \end{array} \right.$	$x \in [0, 1]$	$p(x) \propto x^{-\lambda}(1-x)^\mu$	Béta

Maximum d'entropie classique (ME): $\mathbf{g} = \mathbf{H} \mathbf{f}$

maximiser $S(\mathbf{f}) = - \sum_j f_j \ln f_j$ s.c. $g_i = \sum_j H_{ij} f_j, i = 1, \dots, M$

minimiser $K(\mathbf{f} : \mathbf{m}) = \sum_j \left[f_j \ln \frac{f_j}{m_j} + (f_j - m_j) \right]$ s.c. $\mathbf{g} = \mathbf{H} \mathbf{f}$

Solution :

Lagrangien : $\mathcal{L}(\mathbf{f}, \boldsymbol{\lambda}) = K(\mathbf{f} : \mathbf{m}) - \sum_i \lambda_i (g_i - [\mathbf{H} \mathbf{f}]_i)$

$$\frac{\partial \mathcal{L}}{\partial f_j} = 0 \longrightarrow f_j = m_j \exp [-\lambda_0 - [\mathbf{H}^t \boldsymbol{\lambda}]_j]$$

$$\frac{\partial \mathcal{L}}{\partial \lambda_i} = 0 \longrightarrow \sum_j H_{ij} m_j \exp [-\lambda_0 - [\mathbf{H}^t \boldsymbol{\lambda}]_j] = g_i$$

Prise en compte du bruit : $\mathbf{g} = \mathbf{H} \mathbf{f} + \boldsymbol{\epsilon}$

$\min K(\mathbf{f} : \mathbf{m})$ s.c. $\|\mathbf{g} - \mathbf{H} \mathbf{f}\|^2 \leq c \longrightarrow \hat{\mathbf{f}} = \arg \min_{\mathbf{f}} \{ \|\mathbf{g} - \mathbf{H} \mathbf{f}\|^2 + \lambda K(\mathbf{f} : \mathbf{m}) \}$

Maximum d'entropie sur la moyenne (MEM) : $\mathbf{g} = \mathbf{H} \langle \mathbf{f} \rangle$

- Mesure de référence : $\mathbf{f} \in \mathcal{C}$, $\mu(\mathbf{f})$: $\mathbf{m} = \int \mathbf{f} \, d\mu(\mathbf{f})$

- Modèle de mesure : $\mathbf{g} = \mathbf{H} \langle \mathbf{f} \rangle$

- Objective : chercher $p(\mathbf{f})$ ou $\langle \mathbf{f} \rangle = \int_{\mathcal{C}} \mathbf{f} p(\mathbf{f}) \, d\mu(\mathbf{f})$

- Solution :

$$\text{maximiser } - \int_{\mathcal{C}} p(\mathbf{f}) \ln p(\mathbf{f}) \, d\mu(\mathbf{f}) \quad \text{s.c. } \mathbf{g} = \mathbf{H} \langle \mathbf{f} \rangle$$

$$p(\mathbf{f}|\boldsymbol{\lambda}) = \frac{1}{Z(\boldsymbol{\lambda})} \mu(\mathbf{f}) \exp [\boldsymbol{\lambda}^t [\mathbf{H} \mathbf{f}]]$$

$$Z(\boldsymbol{\lambda}) = \int_{\mathcal{C}} \mu(\mathbf{f}) \exp [\boldsymbol{\lambda}^t [\mathbf{H} \mathbf{f}]] \, d\mathbf{f}$$

$$\frac{\partial \ln Z(\boldsymbol{\lambda})}{\partial \lambda_i} = 0, \quad i = 1, \dots, M \quad \longrightarrow \boldsymbol{\lambda} \longrightarrow p(\mathbf{f}|\boldsymbol{\lambda}) \longrightarrow \hat{\mathbf{f}}(\boldsymbol{\lambda}) = \int \mathbf{f} p(\mathbf{f}|\boldsymbol{\lambda}) \, d\mathbf{f}$$

Critère primal / Critère dual :

$$\mathbf{s} = \mathbf{H}^t \boldsymbol{\lambda}, \quad \Omega^*(\mathbf{s}) = \ln Z(\mathbf{s}) = \ln \int_{\mathcal{C}} \exp [\mathbf{s}^t \mathbf{f}] \, d\mu(\mathbf{f}),$$

$$\Omega(\mathbf{f}) = \mathbf{s}^t \mathbf{f} - \Omega^*(\mathbf{s}), \quad D(\boldsymbol{\lambda}) = \boldsymbol{\lambda}^t \mathbf{g} - \Omega^*(\mathbf{H}^t \boldsymbol{\lambda})$$

$$\widehat{\boldsymbol{\lambda}} = \arg \min_{\boldsymbol{\lambda}} \{D(\boldsymbol{\lambda})\}$$

Critère dual

$$\widehat{\mathbf{f}} = \arg \min_{\mathbf{f} \in \mathcal{C}} \{\Omega(\mathbf{f})\} \quad \text{s.c.} \quad \mathbf{g} = \mathbf{H} \mathbf{f}$$

Critère primal

$$\widehat{\mathbf{f}}(\mathbf{s}) = \frac{d\Omega^*(\mathbf{s})}{d\mathbf{s}}$$

relation explicite

- Ω et D dépendent de la mesure de référence $d\mu(\mathbf{f})$.

- Si on note $\mathbf{m} = \int \mathbf{f} \, d\mu(\mathbf{f})$ alors on montre que

$\Omega(\mathbf{f}) = \Omega(\mathbf{f}, \mathbf{m})$ est une mesure de distance entre \mathbf{f} et \mathbf{m} :

- $\Omega(\mathbf{f}, \mathbf{m}) \geq 0$, et $\Omega(\mathbf{f}, \mathbf{m}) = 0$ ssi $\mathbf{f} = \mathbf{m}$
- $\Omega(\mathbf{f}, \mathbf{m})$ est dérivable et strictement convexe sur \mathcal{C}
- $\Omega(\mathbf{f}, \mathbf{m}) = \infty$ si $\mathbf{f} \notin \mathcal{C}$.

- Si $\mu(\mathbf{f}) = \prod_{j=1}^N \mu_j(f_j)$ alors $p(\mathbf{f}, \boldsymbol{\lambda}) = \prod_{j=1}^N p_j(f_j, \boldsymbol{\lambda})$

et

$$\Omega^*(\mathbf{s}, \mathbf{m}) = \sum_j \Omega_j^*(s_j, m_j), \quad \Omega(\mathbf{f}, \mathbf{m}) = \sum_j \Omega_j(f_j, m_j), \quad \hat{f}_j = \frac{\partial \Omega_j^*(s_j, m_j)}{\partial s_j}$$

$\Omega_j^*(s, m)$ et $\Omega_j(f, m)$ dépendent des mesures de référence $\mu_j(f)$.

- Lorsque $\mu(\mathbf{f})$ n'est pas séparable il devient très difficile de mener les calcul (excepté le cas gaussien).
- Cas gaussien : $\mu(\mathbf{f}) = \mathcal{N}(\mathbf{m}, \mathbf{R}_f)$ où on a

$$\Omega(\mathbf{f}, \mathbf{m}) = -\frac{1}{2}(\mathbf{f} - \mathbf{m})^t \boldsymbol{\Sigma}^{-1}(\mathbf{f} - \mathbf{m}) \quad \text{et} \quad \Omega^*(\boldsymbol{\lambda}) = -\frac{1}{2}\boldsymbol{\lambda}^t \boldsymbol{\lambda}$$

Exemples :

Cas	$\mu_j(f)$	$\Omega_j^*(s, m)$	$\Omega_j(f, m)$
Gaussienne	$\exp \left[-\frac{1}{2}(f - m)^2 \right]$	$\frac{1}{2}(s - m)^2$	$\frac{1}{2}(f - m)^2$
Poissonnienne	$\frac{m^f}{f!} \exp[-m]$	$\exp[m - s]$	$-\frac{f}{m} \ln \frac{f}{m} + m - f$
Gamma	$f^{\alpha-1} \exp[-f/m]$	$\ln(s - m)$	$-\ln \frac{f}{m} + \frac{f}{m} - 1$
Gaussienne	$\mathcal{N}(m, \Sigma)$	$\ s - m\ ^2$	$\ f - m\ _{\Sigma}^2$

Prise en compte du bruit :

(Thèses de G. Le Besnerais (93) et J.F. Bercher(95))

Deux approches possibles :

1- Remplacer $\mathbf{g} = \mathbf{H}\mathbf{f}$ par

$$\left\{ \begin{array}{l} |g_i - [\mathbf{H}\mathbf{f}]_i| < \epsilon \\ \text{ou} \\ \|\mathbf{g} - \mathbf{H}\mathbf{f}\|^2 < \sigma^2 \end{array} \right.$$

$$\left\{ \begin{array}{l} \hat{\mathbf{f}} = \arg \min_{\mathbf{f}} \{\Omega(\mathbf{f})\} \text{ s.c. } |g_i - [\mathbf{H}\mathbf{f}]_i| < \epsilon, \text{ avec } \Omega(\mathbf{f}) = \sum_j \Omega_j(f_j) \\ \text{ou} \\ \hat{\boldsymbol{\lambda}} = \arg \min_{\boldsymbol{\lambda}} \{\tilde{D}(\boldsymbol{\lambda})\} \text{ avec } \tilde{D}(\boldsymbol{\lambda}) = D(\boldsymbol{\lambda}) + \alpha \|\boldsymbol{\lambda}\|^2 \end{array} \right.$$

α dépend de σ^2 .

2- Remplacer $\mathbf{g} = \mathbf{H}\mathbf{f}$ par $\mathbf{g} = \mathbf{H}\mathbf{f} + \boldsymbol{\epsilon}$ et la réécrire

$$\mathbf{g} = [\mathbf{H}|\mathbf{I}] \begin{bmatrix} \mathbf{f} \\ \boldsymbol{\epsilon} \end{bmatrix} \longrightarrow \mathbf{g} = \tilde{\mathbf{H}} \tilde{\mathbf{f}}$$

Si $\mu(\tilde{\mathbf{f}}) = \mu_f(\mathbf{f}) \mu_\epsilon(\boldsymbol{\epsilon})$ alors

$$\hat{\tilde{\mathbf{f}}} = \arg \min_{\mathbf{f} \in \mathcal{C}} \{ \mathcal{Q}(\mathbf{g} - \mathbf{H}\mathbf{f}) + \alpha \Omega(\mathbf{f}) \}$$

avec

$$\Omega(\mathbf{f}) = \sum_{j=1}^N \Omega_j(f_j), \quad \text{et} \quad \mathcal{Q}(\mathbf{z}) = \sum_{i=1}^M q_i(z_j),$$

$\Omega_j(f_j)$ et $q_i(z_i)$ dépendent des mesures de référence $\mu_f(\mathbf{f})$ et $\mu_b(\mathbf{f})$.

Cependant la détermination du α reste empirique.

APPROCHE BAYÉSIENNE DE L'ESTIMATION

- Modèle d'observation et caractéristiques du bruit : $\longrightarrow p(\mathbf{g}|\mathbf{f}, \boldsymbol{\theta}_1)$
- Information *a priori* sur \mathbf{f} : $\longrightarrow p(\mathbf{f}|\boldsymbol{\theta}_2)$
- Bayes : $\longrightarrow p(\mathbf{f}|\mathbf{g}, \boldsymbol{\theta}) = \frac{p(\mathbf{g}|\mathbf{f}, \boldsymbol{\theta}_1) p(\mathbf{f}|\boldsymbol{\theta}_2)}{p(\mathbf{g}|\boldsymbol{\theta})}$ avec $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)$
- Règle d'estimation : $\longrightarrow \hat{\mathbf{f}}$

Exemple : *Maximum A Posteriori* (MAP) :

$$\hat{\mathbf{f}} = \arg \min_{\mathbf{f}} \{-\ln p(\mathbf{f}|\mathbf{g}, \boldsymbol{\theta})\} = \arg \min_{\mathbf{f}} \{-\ln p(\mathbf{g}|\mathbf{f}, \boldsymbol{\theta}_1) - \ln p(\mathbf{f}|\boldsymbol{\theta}_2)\}$$

Cadre général: fonction coût $c(\mathbf{f}, \hat{\mathbf{f}})$

$$\hat{\mathbf{f}}(\mathbf{g}, \boldsymbol{\theta}_1, \boldsymbol{\theta}_2) = \arg \min_{\mathbf{z}} \left\{ \int c(\mathbf{f}, \mathbf{z}) p(\mathbf{f}|\mathbf{g}, \boldsymbol{\theta}) d\mathbf{f} \right\}$$

CAS LINÉAIRE GAUSSIEN

$$\mathbf{g} = \mathbf{H}\mathbf{f} + \boldsymbol{\epsilon}$$

- Hypothèse sur le bruit : $\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \sigma_\epsilon^2 \mathbf{I}) \longrightarrow \mathbf{g}|\mathbf{f} \sim \mathcal{N}(\mathbf{H}\mathbf{f}, \sigma_\epsilon^2 \mathbf{I})$

$$p(\mathbf{g}|\mathbf{f}) \propto \exp \left[-\frac{1}{2\sigma_\epsilon^2} \|\mathbf{g} - \mathbf{H}\mathbf{f}\|^2 \right]$$

- Hypothèse gaussienne sur \mathbf{f} : $\mathbf{f} \sim \mathcal{N}(\mathbf{f}_0, \sigma_f^2 \mathbf{P}_0)$:

$$p(\mathbf{f}) \propto \exp \left[-\frac{1}{2\sigma_f^2} [\mathbf{f} - \mathbf{f}_0]^t \mathbf{P}_0^{-1} [\mathbf{f} - \mathbf{f}_0] \right]$$

- loi a posteriori

$$p(\mathbf{f}|\mathbf{g}) \propto \exp \left[-\frac{1}{2\sigma_\epsilon^2} \|\mathbf{g} - \mathbf{H}\mathbf{f}\|^2 - \frac{1}{2\sigma_f^2} [\mathbf{f} - \mathbf{f}_0]^t \mathbf{P}_0^{-1} [\mathbf{f} - \mathbf{f}_0] \right]$$

Cadre général de l'estimé MAP :

$$\hat{\mathbf{f}} = \arg \min_{\mathbf{f}} \{J(\mathbf{f})\} \quad \text{avec} \quad J(\mathbf{f}) = \|\mathbf{g} - \mathbf{H}\mathbf{f}\|^2 + \lambda\Omega(\mathbf{f})$$

- Hypothèse Gaussienne sur \mathbf{f} :

$$p(f_j) \propto \exp[-\alpha(f_j - m_j)^2] \longrightarrow \Omega(\mathbf{f}) = \alpha \sum_j (f_j - m_j)^2$$

- Hypothèse Gamma sur \mathbf{f} :

$$p(f_j) \propto (f_j/m_j)^\alpha \exp[-f_j/m_j] \longrightarrow \Omega(\mathbf{f}) = \alpha \sum_j \ln \frac{f_j}{m_j} + \frac{f_j}{m_j},$$

- Hypothèse Béta sur \mathbf{f} :

$$p(f_j) \propto f_j^\alpha (1 - f_j)^\beta \longrightarrow \Omega(\mathbf{f}) = \alpha \sum_j \ln f_j + \beta \sum_j \ln(1 - f_j),$$

- Hypothèse gaussienne généralisée sur \mathbf{f} :

$$p(f_j) \propto \exp[-\alpha|f_j - m_j|^p], \quad 1 < p < 2 \longrightarrow \Phi(\mathbf{f}) = \alpha \sum_j |f_j - m_j|^p,$$

- Modèles markoviens pour \mathbf{f} :

$$p(f_j|\mathbf{f}) \propto \exp \left[-\alpha \sum_{i \in N_j} \phi(f_j, f_i) \right] \longrightarrow \Phi(\mathbf{f}) = \alpha \sum_j \sum_{i \in N_j} \phi(f_j, f_i),$$

Estimation au sens du MAP ou régularisation:

$$\hat{\mathbf{f}} = \arg \min_{\mathbf{f}} \{J(\mathbf{f})\} \quad \text{avec} \quad J(\mathbf{f}) = \|\mathbf{g} - \mathbf{h}(\mathbf{f})\|^2 + \lambda \Omega(\mathbf{f})$$

- Lois gaussiennes: $\Omega(\mathbf{f})$ quadratique
 $\longrightarrow J(\mathbf{f})$ quadratique $\longrightarrow \hat{\mathbf{f}}$ fonction linéaire de \mathbf{g} \longrightarrow Algorithmes rapides

- Lois non gaussiens mais $\Omega(\mathbf{f})$ convexe:

Exemples: $\Omega(\mathbf{f}) = \sum_j \phi(f_j)$ ou $\Omega(\mathbf{f}) = \sum_j \phi(f_j - f_{j-1})$

avec $\phi(t) = \{|t|^p, (|t| \ln |t| - |t|), (\ln(1 + |t|))\}$

$\longrightarrow J(\mathbf{f})$ convexe \longrightarrow Estimation nonlinéaire mais facile à calculer

- Lois non gaussiens et $\Omega(\mathbf{f}) = \sum_j \Omega(f_j - f_{j-1})$ non strictement convexe:

Exemples:

$$\Omega(t) = \begin{cases} |t|^2 & \text{if } |t| < \alpha, \\ \alpha^2 & \text{else,} \end{cases}, \begin{cases} t^2 & \text{if } |t| < \alpha, \\ 2\alpha t - \alpha^2 & \text{else,} \end{cases}, \frac{\alpha^2 t^2}{1+t^2}, \log \cosh(t/\alpha)$$

$\longrightarrow J(\mathbf{f})$ non convexe \longrightarrow Minima locaux \longrightarrow Optimisation globale

BAYES OU MEM ?

MEM :

- Choix de \mathcal{C} et de $\mu(\mathbf{f})$:
 - Différents choix pour \mathcal{C} : \mathbb{R}^N , \mathbb{R}_+^N , $[a, b]^N$
 - Seule $\mu(\mathbf{f}) = \prod_{j=1}^N \mu_j(f_j)$ est actuellement utilisée
 - μ_j qui permettent de mener les calculs jusqu'au bout sont limitées
 - Choix de $\mu(\mathbf{f})$ non séparable se heurte au calcul du $Z(\boldsymbol{\lambda})$.
- Prise en compte du bruit :
 - Dans la première approche seule le support du bruit est utilisé
 - Dans la deuxième approche le choix du α est ad hoc.
- Calcul effective de la solution :
 - Pas de difficulté spécifique car par construction les critères obtenues sont convexes.

BAYES :

- Choix des lois a priori :
 - $p(\mathbf{f})$ peut être séparable ou non.
 - Prise en compte des corrélations peut se faire par des modèles markoviens
- Choix d'un estimateur MAP, MP, MAPM
ou choix de la fonction coût : Plutôt considérations calculatoires :
 - MAP est le plus facile car ne nécessite pas d'intégration.
 - MP ou MAPM nécessite les calculs intégrales mais il y a des outils du type échantillonneur de Gibbs.
- Calcul effective de la solution
 - MAP = Régularisation
 - Si la loi *a posteriori* n'est pas unimodale on peut utiliser
 - soit les algorithmes d'optimisation du type recuit simulé,
 - soit les algorithmes du type relaxation déterministe (GNC)
- Détermination des hyperparamètres β et θ :

PROBLÈMES OUVERTS

- Modélisation du problème direct :
 - MEM limitée aux problèmes inverses linéaires ?
 - BAYES n'est pas limitée d'un point de vue formel.
- Prise en compte de l'incertitude du modèle
- Détermination des hyperparamètres et sélection du modèle
 - Dans BAYES il y a des outils nécessaires.
- Mesure de la qualité de la solution:
 - Dans BAYES la loi *a posteriori* résume tout sur la solution. On peut alors calculer les variances, les covariances, etc.
 - Dans MEM on n'utilise pas la loi *a posteriori* pour ce fin. Cependant on peut calculer la sensibilité de la solution vis-à-vis des erreurs.