Inverse Problems: From deterministic regularization theory to Bayesian inference

Ali Mohammad-Djafari

Groupe Problèmes Inverses
Laboratoire des Signaux et Systèmes
(UMR 8506 CNRS - SUPELEC - Univ Paris Sud 11)
Supélec, Plateau de Moulon, 91192 Gif-sur-Yvette, FRANCE.

djafari@lss.supelec.fr
http://djafari.free.fr
http://www.lss.supelec.fr

KTH, Dept. of Mathematics, Stockhol, Sweden
21/04/2008

Content

- Invers Problems : Examples and general formulation
- Inversion methods :
  analytical, parametric and non parametric
- Determinitic regularization
- Probabilistic methods
- Bayesian inference approach
- Prior moedels for images
- Bayesian computation
- Application in Computed Tomography
- Conclusions
- Questions and Discussion

Inverse problems : 3 examples

- Example 1 :
  Measuring variation of temperature with a therometer
  - $f(t)$ variation of temperature over time
  - $g(t)$ variation of legth of the liquid in thermometer
- Example 2 :
  Making an image with a camera, a microscope or a telescope
  - $f(x,y)$ real scene
  - $g(x,y)$ observed image
- Example 3 : Making an image of the interior of a body
  - $f(x,y)$ a section of a real 3D body $f(x,y,z)$
  - $g_\phi(r)$ a line of observed radiographe $g_\phi(r,z)$

- Example 1 : Deconvolution
- Example 2 : Image restoration
- Example 3 : Image reconstruction

Measuring variation of temperature with a therometer

- $f(t)$ variation of temperature over time
- $g(t)$ variation of legth of the liquid in thermometer
- Forward model : Convolution

$$g(t) = \int f(t')\, h(t-t')\, \mathrm{d}t'$$

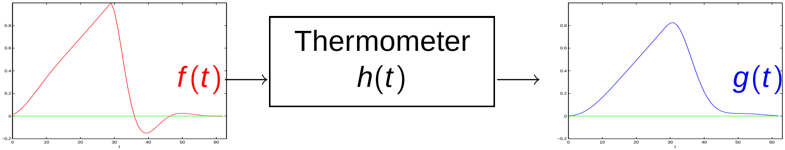
  $h(t)$: impulse response of the measurement system
- Inverse problem : Deconvolution

  Given the forward model $\mathcal{H}$ (impulse response $h(t)$))
  and a set of data $g(t_i), i = 1, \cdots, M$
  find $f(t)$

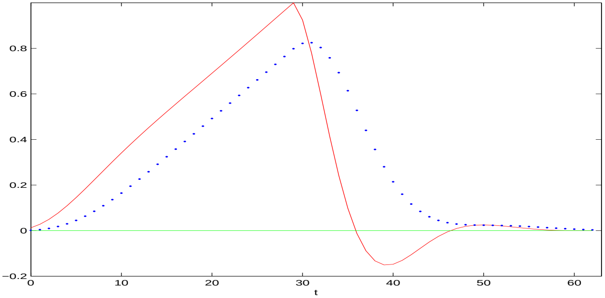
Measuring variation of temperature with a therometer

Forward model : Convolution

$$g(t) = \int f(t')\, h(t-t')\, \mathrm{d}t'$$

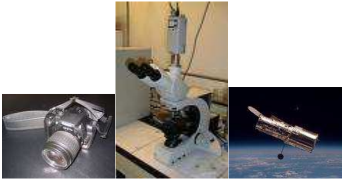


Inversion : Deconvolution



Making an image with a camera, a microscope or a telescope

- $f(x,y)$ real scene
- $g(x,y)$ observed image



- Forward model : Convolution

$$g(x,y) = \iint f(x',y')\, h(x-x',y-y')\, \mathrm{d}x'\, \mathrm{d}y'$$

$h(x,y)$: Point Spread Function (PSF) of the imaging system

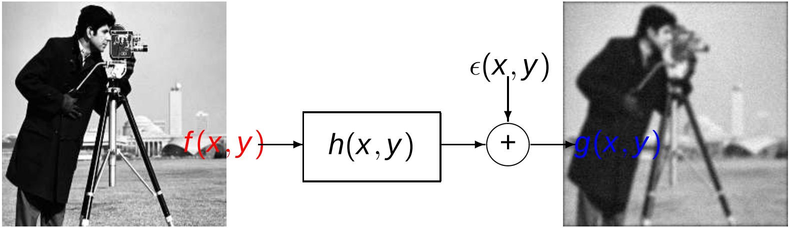
- Inverse problem : Image restoration

Given the forward model $\mathcal{H}$ (PSF $h(x,y)$))
and a set of data $g(x_i,y_i), i = 1, \cdots, M$
find $f(x,y)$

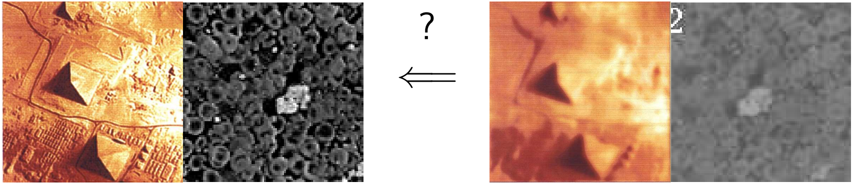
Making an image with an unfocused camera

Forward model : 2D Convolution

$$g(x,y) = \iint f(x',y')\, h(x-x', y-y')\, \mathrm{d}x'\, \mathrm{d}y'$$

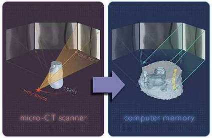


Inversion : Deconvolution



Making an image of the interior of a body

- $f(x,y)$ a section of a real 3D body $f(x,y,z)$
- $g_\phi(r)$ a line of observed radiographe $g_\phi(r,z)$



- Forward model :
  Line integrals or Radon Transform

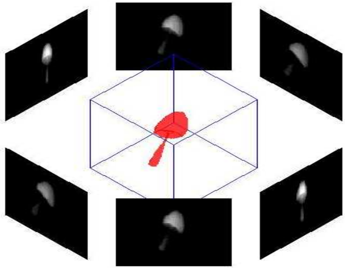
$$g_\phi(r) = \int_{L_{r,\phi}} f(x,y) \, \mathrm{d}l = \iint f(x,y) \, \delta(r - x\cos\phi - y\sin\phi) \, \mathrm{d}x \, \mathrm{d}y$$

- Inverse problem : Image reconstruction

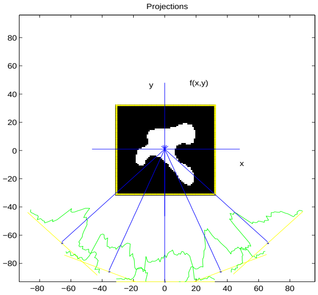
  Given the forward model $\mathcal{H}$ (Radon Transform) and
  a set of data $g_{\phi_i}(r), i = 1, \cdots, M$
  find $f(x,y)$

2D and 3D Computed Tomography

3D



2D

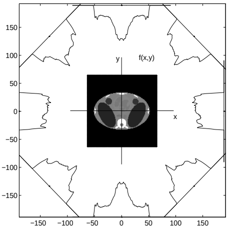


$$g_\phi(r_1,r_2) = \int_{\mathcal{L}_{r_1,r_2,\phi}} f(x,y,z)\,\mathrm{d}l \qquad g_\phi(r) = \int_{\mathcal{L}_{r,\phi}} f(x,y)\,\mathrm{d}l$$

Forward probelm : $f(x,y)$ or $f(x,y,z) \longrightarrow g_\phi(r)$ or $g_\phi(r_1,r_2)$

Inverse problem : $g_\phi(r)$ or $g_\phi(r_1,r_2) \longrightarrow f(x,y)$ or $f(x,y,z)$

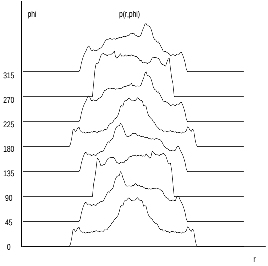
X ray Tomography and Radon Transform



$$g(r,\phi) = -\ln\left(\frac{I}{I_0}\right) = \int_{L_{r,\phi}} f(x,y)\, dl$$

$$g(r,\phi) = \iint_D f(x,y)\,\delta(r - x\cos\phi - y\sin\phi)\, dx\, dy$$

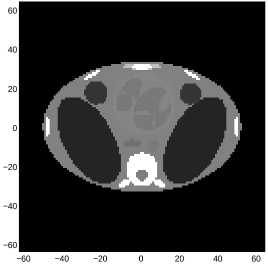
$f(x,y) \longrightarrow$ RT $\longrightarrow g(r,\phi)$



IRT

?

$\Longrightarrow$



General formulation of inverse problems

- General non linear inverse problems :

$$g(\boldsymbol{s}) = [\mathcal{H} f(\boldsymbol{r})](\boldsymbol{s}), \quad \boldsymbol{r} \in \mathcal{R}, \quad \boldsymbol{s} \in \mathcal{S}$$

- Linear models :

$$g(\boldsymbol{s}) = \int f(\boldsymbol{r})\, h(\boldsymbol{r}, \boldsymbol{s})\, \mathrm{d}\boldsymbol{r}$$

  If $h(\boldsymbol{r}, \boldsymbol{s}) = h(\boldsymbol{r} - \boldsymbol{s}) \longrightarrow$ Convolution.

- Discrete data :

$$g(\boldsymbol{s}_i) = \int h(\boldsymbol{s}_i, \boldsymbol{r})\, f(\boldsymbol{r})\, \mathrm{d}\boldsymbol{r} + \epsilon(\boldsymbol{s}_i), \quad i = 1, \cdots, m$$

- Inversion : Given the forward model $\mathcal{H}$ and the data $\boldsymbol{g} = \{g(\boldsymbol{s}_i), i = 1, \cdots, m)\}$ estimate $f(\boldsymbol{r})$
- Well-posed and Ill-posed problems (Hadamard) : existance, uniqueness and stability
- Need for prior information

Analytical methods (mathematical physics)

$$g(s_i) = \int h(s_i, r)\, f(r)\, dr + \epsilon(s_i), \quad i = 1, \cdots, m$$

$$g(s) = \int h(s, r)\, f(r)\, dr$$

$$\widehat{f}(r) = \int w(s, r)\, g(s)\, ds$$

$w(s, r)$ minimizing a criterion :

$$\begin{aligned} Q(w(s,r)) &= \left\| g(s) - [\mathcal{H}\widehat{f}(r)](s) \right\|_2^2 = \int \left| g(s) - [\mathcal{H}\widehat{f}(r)](s) \right|^2 \, ds \\ &= \int \left| g(s) - \int h(s,r)\, \widehat{f}(r)\, dr \right|^2 \, ds \\ &= \int \left| g(s) - \int h(s,r) \left[\int w(s,r)\, g(s)\, ds \right] \, dr \right|^2 \, ds \\ &= \int \left| g(s) - \int \int h(s,r) w(s,r)\, g(s)\, ds\, dr \right|^2 \, ds \end{aligned}$$

Analytical methods

- Trivial solution :

$$w(\mathbf{s}, \mathbf{r}) = h^{-1}(\mathbf{s}, \mathbf{r})$$

Example : Fourier Transform :

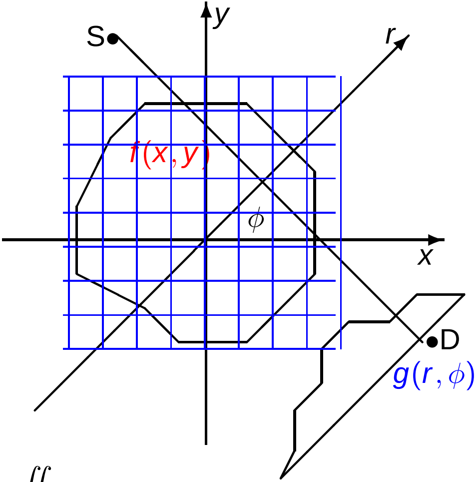
$$g(\mathbf{s}) = \int f(\mathbf{r}) \exp\{-j\mathbf{s}.\mathbf{r}\} \, \mathrm{d}\mathbf{r}$$

$$h(\mathbf{s}, \mathbf{r}) = \exp\{-j\mathbf{s}.\mathbf{r}\} \longrightarrow w_i(\mathbf{r}) = \exp\{+j\mathbf{s}.\mathbf{r}\}$$

$$f(\mathbf{r}) = \int g(\mathbf{s}) \exp\{+j\mathbf{s}.\mathbf{r}\} \, \mathrm{d}\mathbf{s}$$

- Known classical solutions for specific expressions of $h(\mathbf{s}, \mathbf{r})$:
  - 1D cases : 1D Fourier, Hilbert, Weil, Melin, ...
  - 2D cases : 2D Fourier, Radon, ...

Analytical Inversion methods



$g(r,\phi) = \int_L f(x,y)\, \mathrm{d}l$

Radon :

$$g(r,\phi) = \iint_D f(x,y)\, \delta(r - x\cos\phi - y\sin\phi)\, \mathrm{d}x\, \mathrm{d}y$$

$$f(x,y) = \left(-\frac{1}{2\pi^2}\right) \int_0^{\pi} \int_{-\infty}^{+\infty} \frac{\frac{\partial}{\partial r} g(r,\phi)}{(r - x\cos\phi - y\sin\phi)}\, \mathrm{d}r\, \mathrm{d}\phi$$

Filtered Backprojection method

$$f(x,y) = \left(-\frac{1}{2\pi^2}\right) \int_0^{\pi} \int_{-\infty}^{+\infty} \frac{\frac{\partial}{\partial r} g(r,\phi)}{(r - x\cos\phi - y\sin\phi)} \, \mathrm{d}r \, \mathrm{d}\phi$$

Derivation $\mathcal{D}$: $$\bar{g}(r,\phi) = \frac{\partial g(r,\phi)}{\partial r}$$

Hilbert Transform $\mathcal{H}$: $$g_1(r',\phi) = \frac{1}{\pi} \int_0^{\infty} \frac{\bar{g}(r,\phi)}{(r - r')} \, \mathrm{d}r$$

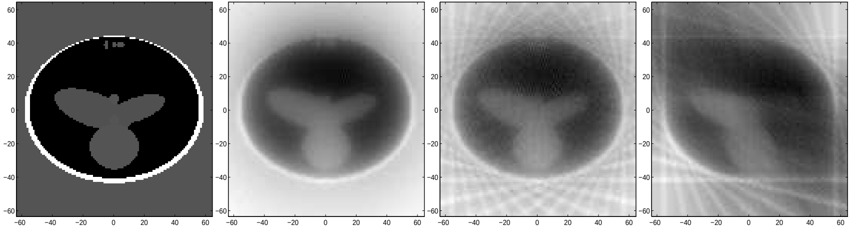
Backprojection $\mathcal{B}$: $$f(x,y) = \frac{1}{2\pi} \int_0^{\pi} g_1(r' = x\cos\phi + y\sin\phi, \phi)$$

$$f(x,y) = \mathcal{B}\,\mathcal{H}\,\mathcal{D}\, g(r,\phi) = \mathcal{B}\,\mathcal{F}_1^{-1}\,|\Omega|\,\mathcal{F}_1\, g(r,\phi)$$

- Backprojection of filtered projections :

$$\xrightarrow{g(r,\phi)} \boxed{\text{FT } \mathcal{F}_1} \longrightarrow \boxed{\text{Filter } |\Omega|} \longrightarrow \boxed{\text{IFT } \mathcal{F}_1^{-1}} \xrightarrow{g_1(r,\phi)} \boxed{\text{Backprojection } \mathcal{B}} \xrightarrow{f(x,y)}$$

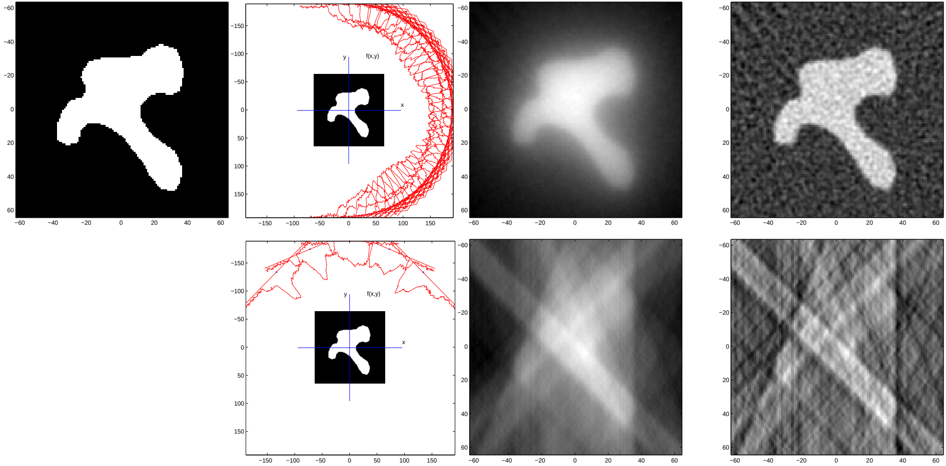
Limitations : Limited angle or noisy data



Original | 64 proj. | 16 proj. | 8 proj. $[0, \pi/2]$

- Limited angle or noisy data
- Accounting for detector size
- Other measurement geometries : fan beam, ...

Limitations : Limited angle or noisy data



Original | Data | Backprojection | Filtered Backprojectio

Parametric methods

- $f(\boldsymbol{r})$ is described in a parametric form with a very few number of parameters $\boldsymbol{\theta}$ and one searches $\widehat{\boldsymbol{\theta}}$ which minimizes a criterion such as :
- Least Squares (LS) : $Q(\boldsymbol{\theta}) = \sum_i |g_i - [\mathcal{H} f(\boldsymbol{\theta})]_i|^2$
- Robust criteria : $Q(\boldsymbol{\theta}) = \sum_i \phi\left(|g_i - [\mathcal{H} f(\boldsymbol{\theta})]_i|\right)$
  with different functions ϕ (L_1, Hubert, ...).
- Likelihood : $\mathcal{L}(\boldsymbol{\theta}) = -\ln p(\boldsymbol{g}|\boldsymbol{\theta})$
- Penalized likelihood : $\mathcal{L}(\boldsymbol{\theta}) = -\ln p(\boldsymbol{g}|\boldsymbol{\theta}) + \lambda\phi(\boldsymbol{\theta})$

Examples :

- Spectrometry : $f(t)$ modelled as a sum og gaussians
  $f(t) = \sum_{k=1}^{K} a_k \mathcal{N}(t|\mu_k, v_k)$ $\quad \boldsymbol{\theta} = \{a_k, \mu_k, v_k\}$
- Tomography in CND : $f(x, y)$ is modelled as a superposition of circular or elleiptical discs
  $\boldsymbol{\theta} = \{a_k, \mu_k, r_k\}$

Non parametric methods

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

- $f(\boldsymbol{r})$ is assumed to be well approximated by

$$f(\boldsymbol{r}) \simeq \sum_{j=1}^{N} f_j\, b_j(\boldsymbol{r})$$

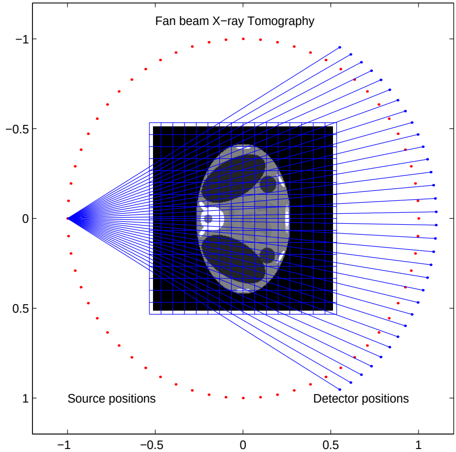
with $\{b_j(\boldsymbol{r})\}$ a basis or any other set of known functions

$$g(\boldsymbol{s}_i) = g_i \simeq \sum_{j=1}^{N} f_j \int h(\boldsymbol{s}_i, \boldsymbol{r})\, b_j(\boldsymbol{r})\, \mathrm{d}\boldsymbol{r}, \quad i = 1, \cdots, M$$

$$\boldsymbol{g} = \boldsymbol{H}\boldsymbol{f} + \boldsymbol{\epsilon} \text{ with } H_{ij} = \int h(\boldsymbol{s}_i, \boldsymbol{r})\, b_j(\boldsymbol{r})\, \mathrm{d}\boldsymbol{r}$$

- $\boldsymbol{H}$ is huge dimensional
- LS solution : $\widehat{\boldsymbol{f}} = \arg\min_{\boldsymbol{f}} \{Q(\boldsymbol{f})\}$ with
  $Q(\boldsymbol{f}) = \sum_i |g_i - [\boldsymbol{H}\boldsymbol{f}]_i|^2 = \|\boldsymbol{g} - \boldsymbol{H}\boldsymbol{f}\|^2$
  does not give satisfactory result.

CT as a linear inverse problem



$$\color{blue}{g(s_i)} = \int_{L_i} \color{red}{f(\boldsymbol{r})} \, \mathrm{d}l_i \longrightarrow \text{Discretization} \longrightarrow \color{blue}{\boldsymbol{g}} = \boldsymbol{H} \color{red}{\boldsymbol{f}} + \boldsymbol{\epsilon}$$

Classical methods in CT

$$g(s_i) = \int_{L_i} f(\boldsymbol{r}) \; \mathrm{d}l_i \longrightarrow \text{Discretization} \longrightarrow \boldsymbol{g} = \boldsymbol{H}\boldsymbol{f} + \boldsymbol{\epsilon}$$

- $\boldsymbol{H}$ is a huge dimensional matrix of line integrals
- $\boldsymbol{H}\boldsymbol{f}$ is the forward or projection operation
- $\boldsymbol{H}^t\boldsymbol{g}$ is the backward or backprojection operation
- $(\boldsymbol{H}^t\boldsymbol{H})^{-1}\boldsymbol{H}^t\boldsymbol{g}$ is the filtered backprojection minimizing the LS criterion $Q(\boldsymbol{f}) = \|\boldsymbol{g} - \boldsymbol{H}\boldsymbol{f}\|^2$
- Iterative methods :
  $$\widehat{\boldsymbol{f}}^{(k+1)} = \widehat{\boldsymbol{f}}^{(k)} + \alpha^{(k)}\boldsymbol{H}^t\left(\boldsymbol{g} - \boldsymbol{H}\widehat{\boldsymbol{f}}^{(k)}\right)$$
  try to minimize the Least squares criterion
- Other criteria :
  - Robust criteria : $Q(\boldsymbol{f}) = \sum_i \phi(|g_i - [\boldsymbol{H}\boldsymbol{f}]_i|)$
  - Likelihood : $\mathcal{L}(\boldsymbol{f}) = p(\boldsymbol{g}|\boldsymbol{f})$
  - Regularization : $J(\boldsymbol{f}) = \|\boldsymbol{g} - \boldsymbol{H}\boldsymbol{f}\|^2 + \lambda\|\boldsymbol{D}\boldsymbol{f}\|^2$.

Inversion : Deterministic methods

Data matching

- Observation model
  $g_i = h_i(\boldsymbol{f}) + \epsilon_i, \quad i = 1, \ldots, M \longrightarrow \boldsymbol{g} = \boldsymbol{H}(\boldsymbol{f}) + \boldsymbol{\epsilon}$
- Misatch between data and output of the model $\Delta(\boldsymbol{g}, \boldsymbol{H}(\boldsymbol{f}))$

$$\widehat{\boldsymbol{f}} = \arg\min_{\boldsymbol{f}} \{\Delta(\boldsymbol{g}, \boldsymbol{H}(\boldsymbol{f}))\}$$

- Examples :

– LS $\quad \Delta(\boldsymbol{g}, \boldsymbol{H}(\boldsymbol{f})) = \|\boldsymbol{g} - \boldsymbol{H}(\boldsymbol{f})\|^2 = \sum_i |g_i - h_i(\boldsymbol{f})|^2$

– L_p $\quad \Delta(\boldsymbol{g}, \boldsymbol{H}(\boldsymbol{f})) = \|\boldsymbol{g} - \boldsymbol{H}(\boldsymbol{f})\|^p = \sum_i |g_i - h_i(\boldsymbol{f})|^p, \quad 1 < p < 2$

– KL $\quad \Delta(\boldsymbol{g}, \boldsymbol{H}(\boldsymbol{f})) = \sum_i g_i \ln \frac{g_i}{h_i(\boldsymbol{f})}$

- In general, does not give satisfactory results for inverse problems.

Regularization theory

Inverse problems = Ill posed problems

$\longrightarrow$ Need for prior information

Functional space : Tikhonov :

$$g = \mathcal{H}f \longrightarrow J(f) = ||g - \mathcal{H}(f)||_2^2 + \lambda||\mathcal{D}f||_2^2$$

Finite dimensional space : Philips & Towmey :

- Minimum norme LS (MNLS) : $J(\boldsymbol{f}) = ||\boldsymbol{g} - \boldsymbol{H}(\boldsymbol{f})||^2 + \lambda||\boldsymbol{f}||^2$
- Classical regularization : $J(\boldsymbol{f}) = ||\boldsymbol{g} - \boldsymbol{H}(\boldsymbol{f})||^2 + \lambda||\boldsymbol{D}\boldsymbol{f}||^2$
- More general regularization :

$$J(\boldsymbol{f}) = \mathcal{Q}(\boldsymbol{g} - \boldsymbol{H}(\boldsymbol{f})) + \lambda\Phi(\boldsymbol{D}\boldsymbol{f})$$

or

$$J(\boldsymbol{f}) = \Delta_1(\boldsymbol{g}, \boldsymbol{H}(\boldsymbol{f})) + \lambda\Delta_2(\boldsymbol{f}, \boldsymbol{f}_\infty)$$

Limitations :

- Errors are considered implicitly white and Gaussian
- Limited prior information on the solution
- Lack of tools for the determination of the hyperparameters

Inversion : Probabilistic methods

Taking account of errors and uncertainties $\longrightarrow$ Probability theory

- Maximum Likelihood (ML)
- Minimum Inaccuracy (MI)
- Probability Distribution Matching (PDM)
- Maximum Entropy (ME) and Information Theory (IT)
- Bayesian Inference (BAYES)

Advantages :

- Explicit account of the errors and noise
- A large class of priors via explicit or implicit modeling
- A coherent approach to combine information content of the data and priors

Limitations :

- Practical implementation and cost of calculation

Bayesian estimation approach

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

- Observation model $\mathcal{M}$ + Hypothesis on the noise $\boldsymbol{\epsilon} \longrightarrow$
  $p(\boldsymbol{g}|\boldsymbol{f};\mathcal{M}) = p_{\boldsymbol{\epsilon}}(\boldsymbol{g} - \boldsymbol{H}\boldsymbol{f})$
- A priori information $\quad p(\boldsymbol{f}|\mathcal{M})$
- Bayes : $\quad p(\boldsymbol{f}|\boldsymbol{g};\mathcal{M}) = \dfrac{p(\boldsymbol{g}|\boldsymbol{f};\mathcal{M})\,p(\boldsymbol{f}|\mathcal{M})}{p(\boldsymbol{g}|\mathcal{M})}$

Link with regularisation :

Maximum A Posteriori (MAP) :

$$\begin{aligned} \widehat{\boldsymbol{f}} &= \arg\max_{\boldsymbol{f}} \{p(\boldsymbol{f}|\boldsymbol{g})\} = \arg\max_{\boldsymbol{f}} \{p(\boldsymbol{g}|\boldsymbol{f})\, p(\boldsymbol{f})\} \\ &= \arg\min_{\boldsymbol{f}} \{-\ln p(\boldsymbol{g}|\boldsymbol{f}) - \ln p(\boldsymbol{f})\} \end{aligned}$$

with $\quad Q(\boldsymbol{g}, \boldsymbol{H}\boldsymbol{f}) = -\ln p(\boldsymbol{g}|\boldsymbol{f}) \quad$ and $\quad \lambda\Phi(\boldsymbol{f}) = -\ln p(\boldsymbol{f})$

Case of linear models and Gaussian priors

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

- Hypothesis on the noise : $\boldsymbol{\epsilon} \sim \mathcal{N}(0, \sigma_\epsilon^2 \boldsymbol{I}) \longrightarrow$
  $p(\boldsymbol{g}|\boldsymbol{f}) \propto \exp\left\{-\frac{1}{2\sigma_\epsilon^2}\|\boldsymbol{g} - \boldsymbol{H}\boldsymbol{f}\|^2\right\}$
- Hypothesis on $\boldsymbol{f}$: $\boldsymbol{f} \sim \mathcal{N}(0, \sigma_f^2 (\boldsymbol{D}^t\boldsymbol{D})^{-1}) \longrightarrow$
  $p(\boldsymbol{f}) \propto \exp\left\{-\frac{1}{2\sigma_f^2}\|\boldsymbol{D}\boldsymbol{f}\|^2\right\}$
- A posteriori :
  $$p(\boldsymbol{f}|\boldsymbol{g}) \propto \exp\left\{-\frac{1}{2\sigma_\epsilon^2}\|\boldsymbol{g} - \boldsymbol{H}\boldsymbol{f}\|^2 \frac{1}{2\sigma_f^2}\|\boldsymbol{D}\boldsymbol{f}\|^2\right\}$$
- MAP : $\widehat{\boldsymbol{f}} = \arg\max_{\boldsymbol{f}} \{p(\boldsymbol{f}|\boldsymbol{g})\} = \arg\min_{\boldsymbol{f}} \{J(\boldsymbol{f})\}$
  with $J(\boldsymbol{f}) = \|\boldsymbol{g} - \boldsymbol{H}\boldsymbol{f}\|^2 + \lambda\|\boldsymbol{D}\boldsymbol{f}\|^2, \qquad \lambda = \frac{\sigma_\epsilon^2}{\sigma_f^2}$
- Advantage : characterization of the solution

$$\boldsymbol{f}|\boldsymbol{g} \sim \mathcal{N}(\widehat{\boldsymbol{f}}, \widehat{\boldsymbol{P}}) \text{ with } \widehat{\boldsymbol{f}} = \widehat{\boldsymbol{P}}\boldsymbol{H}^t\boldsymbol{g}, \quad \widehat{\boldsymbol{P}} = \left(\boldsymbol{H}^t\boldsymbol{H} + \lambda\boldsymbol{D}^t\boldsymbol{D}\right)^{-1}$$

MAP estimation with other priors :

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

Separable priors :

- Gaussian : $p(f_j) \propto \exp\{-\alpha|f_j|^2\} \longrightarrow \Omega(\boldsymbol{f}) = \alpha \sum_j |f_j|^2$
- Gamma : $p(f_j) \propto f_j^\alpha \exp\{-\beta f_j\} \longrightarrow \Omega(\boldsymbol{f}) = \alpha \sum_j \ln f_j + \beta f_j$
- Beta :
  $p(f_j) \propto f_j^\alpha (1-f_j)^\beta \longrightarrow \Omega(\boldsymbol{f}) = \alpha \sum_j \ln f_j + \beta \sum_j \ln(1-f_j)$
- Generalized Gaussian :
  $p(f_j) \propto \exp\{-\alpha|f_j|^p\}, \quad 1 < p < 2 \longrightarrow \quad \Phi(\boldsymbol{f}) = \alpha \sum_j |f_j|^p,$

Markovian models :

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

MAP estimation with markovien priors :

$$\widehat{\boldsymbol{f}} = \arg\min_{\boldsymbol{f}} \{J(\boldsymbol{f})\} \quad \text{with} \quad J(\boldsymbol{f}) = \|\boldsymbol{g} - \boldsymbol{H}\boldsymbol{f}\|^2 + \lambda\Omega(\boldsymbol{f})$$

$$\Omega(\boldsymbol{f}) = \sum_j \phi(f_j - f_{j-1})$$

with $\phi(t)$:

Convex functions :

$$|t|^\alpha, \; \sqrt{1+t^2} - 1, \; \log(\cosh(t)), \; \begin{cases} t^2 & |t| \leq T \\ 2T|t| - T^2 & |t| > T \end{cases}$$

or Non convex functions :

$$\log(1+t^2), \quad \frac{t^2}{1+t^2}, \quad \arctan(t^2), \quad \begin{cases} t^2 & |t| \leq T \\ T^2 & |t| > T \end{cases}$$

Two main steps in the Bayesian approach

- Prior modeling
  - Separable :
    Gaussian, Generalized Gaussian, Gamma,
    mixture of Gaussians, mixture of Gammas, ...
  - Markovian : Gauss-Markov, GGM, ...
  - Separable or Markovian with hidden variables
    (contours, region labels)
- Choice of the estimator and computational aspects
  - MAP, Posterior mean, Marginal MAP
  - MAP needs optimization algorithms
  - Posterior mean needs integration methods
  - Marginal MAP needs integration and optimization
  - Approximations :
    - Gaussian approximation (Laplace)
    - Numerical exploration MCMC
    - Variational Bayes (Separable approximation)

Main advantages of the Bayesian approach

- MAP = Regularization
- Posterior mean ? Marginal MAP ?
- More information in the posterior law than only its mode or its mean
- Meaning and tools for estimating hyper parameters
- Meaning and tools for model selection
- More specific and specialized priors, particularly through the hidden variables
- More computational tools :
  - Expectation-Maximization for computing the maximum likelihood parameters
  - MCMC for posterior exploration
  - Variational Bayes for analytical computation of the posterior marginals
  - ...

End of the first talk

- Thanks for your attention
- In the second talk, we go more in details of some of these points :
  - A class of Gauss-Markov-Potts priors for images
  - Computational aspects : MCMC and Variational methods
  - Application in Computed Tomographie with very limited number of projections
  - Other applications : Image fusion, Image separation, Hyperspectral image segmentation, Superresolution, ...

Questions and Discussions

- Thanks for your attentions
- ...
- ...
- Questions ?
- ...
- Discussions ?
- ...
- ...