Inverse Problems: From deterministic methods to probabilistic 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, Stockholm, Sweden
21/04/2008
Content

- Invers Problems : Examples and general formulation
- Inversion methods :
  - analytical, parametric and non parametric
- Deterministic regularization
- Probabilistic methods
- Bayesian inference approach
- Prior models 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 length 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 radiograph $g_\phi(r, z)$

▶ Example 1 : Deconvolution  
▶ Example 2 : Image restoration  
▶ Example 3 : Image reconstruction
Measuring variation of temperature with a thermometer

- $f(t)$ variation of temperature over time
- $g(t)$ variation of length of the liquid in thermometer

Forward model: Convolution

$$g(t) = \int f(t') h(t - t') \, dt' + \epsilon(t)$$

$h(t)$: 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 thermometer

Forward model: Convolution

\[ g(t) = \int f(t') h(t - t') \, dt' + \epsilon(t) \]

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
  \[
  g(x, y) = \int \int f(x', y') h(x - x', y - y') \, dx' \, dy' + \epsilon(x, y)
  \]
  \( h(x, y) \) : 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 \( g(x_i, y_i), i = 1, \ldots, M \), find \( f(x, y) \)
Making an image with an unfocused camera

Forward model: 2D Convolution

\[
g(x, y) = \int \int f(x', y') h(x - x', y - y') \, dx' \, dy' + \epsilon(x, y)
\]

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 radiograph $g_\phi(r, z)$

**Forward model:**
Line integrals or Radon Transform

$$g_\phi(r) = \int_{L_{r,\phi}} f(x, y) \, dl + \epsilon_\phi(r)$$

$$= \iint f(x, y) \delta(r - x \cos \phi - y \sin \phi) \, dx \, dy + \epsilon_\phi(r)$$

**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

\[ g_\phi(r_1, r_2) = \int_{L_{r_1, r_2, \phi}} f(x, y, z) \, dl \quad g_\phi(r) = \int_{L_{r, \phi}} f(x, y) \, dl \]

Forward problem: \( f(x, y) \) or \( f(x, y, z) \) \( \rightarrow \) \( g_\phi(r) \) or \( g_\phi(r_1, r_2) \)

Inverse problem: \( g_\phi(r) \) or \( g_\phi(r_1, r_2) \) \( \rightarrow \) \( f(x, y) \) or \( f(x, y, z) \)
X ray Tomography and Radon Transform

\[ g(r, \phi) = -\ln \left( \frac{l}{l_0} \right) = \int_{L_{r,\phi}} f(x, y) \, dl \]

\[ g(r, \phi) = \int_{D} f(x, y) \delta(r - x \cos \phi - y \sin \phi) \, dx \, dy \]

\[ f(x, y) \xrightarrow{\text{RT}} g(r, \phi) \]

IRT

\[ p(r, \phi) \]

\[ \phi \]

\[ r \]
General formulation of inverse problems

- General non linear inverse problems:
  \[ g(s) = [\mathcal{H}f(r)](s) + \epsilon(s), \quad r \in \mathcal{R}, \quad s \in \mathcal{S} \]

- Linear models:
  \[ g(s) = \int f(r) h(r, s) \, dr + \epsilon(s) \]
  If \( h(r, s) = h(r - s) \) \( \rightarrow \) Convolution.

- Discrete data:
  \[ g(s_i) = \int h(s_i, r) f(r) \, dr + \epsilon(s_i), \quad i = 1, \ldots, m \]

- Inversion: Given the forward model \( \mathcal{H} \) and the data \( g = \{g(s_i), i = 1, \ldots, m\} \) estimate \( f(r) \)

- Well-posed and **ill-posed** problems (Hadamard):
  existence, uniqueness and stability

- Need for **prior information**
Analytical methods (mathematical physics)

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

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

\[ \hat{f}(r) = \int w(s, r) g(s) \, ds \]

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

\[ Q(w(s, r)) = \left\| g(s) - [\mathcal{H} \hat{f}(r)](s) \right\|_2^2 = \int \left\| g(s) - [\mathcal{H} \hat{f}(r)](s) \right\|_2^2 \, ds \]

\[ = \int \left\| g(s) - \int h(s, r) \hat{f}(r) \, dr \right\|^2 \, ds \]

\[ = \int \left\| g(s) - \int h(s, r) \left[ \int w(s, r) g(s) \, ds \right] \, dr \right\|^2 \, ds \]

\[ = \int \left\| g(s) - \int \int h(s, r) w(s, r) g(s) \, ds \, dr \right\|^2 \, ds \]
Analytical methods

- Trivial solution:

\[ w(s, r) = h^{-1}(s, r) \]

Example: Fourier Transform:

\[ g(s) = \int f(r) \exp\{-js.r\} \, dr \]

\[ h(s, r) = \exp\{-js.r\} \implies w(s, r) = \exp\{+js.r\} \]

\[ \hat{f}(r) = \int g(s) \exp\{+js.r\} \, ds \]

- Known classical solutions for specific expressions of \( h(s, r) \):
  - 1D cases: 1D Fourier, Hilbert, Weil, Melin, ...
  - 2D cases: 2D Fourier, Radon, ...
Analytical Inversion methods

Radon:

\[ g(r, \phi) = \int_L f(x, y) \, dl \]

\[ f(x, y) = \left(-\frac{1}{2\pi^2}\right) \int_0^\pi \int_{-\infty}^{+\infty} \frac{\partial}{\partial r} g(r, \phi) \left(\frac{r - x \cos \phi - y \sin \phi}{r - x \cos \phi - y \sin \phi}\right) \, dr \, d\phi \]
Filtered Backprojection method

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

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

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

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

- Backprojection of filtered projections:

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

\[\begin{array}{cccccc}
g(r, \phi) & \xrightarrow{\text{FT}} & \mathcal{F}_1 & \xrightarrow{\text{Filter}} & |\Omega| & \xrightarrow{\text{IFT}} g_1(r, \phi) & \xrightarrow{\text{Backprojection}} & f(x, y) \end{array}\]
Limitations: Limited angle or noisy data

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

- $f(r)$ is described in a parametric form with a very few number of parameters $\theta$ and one searches $\hat{\theta}$ which minimizes a criterion such as:

- **Least Squares (LS):**
  
  $Q(\theta) = \sum_i |g_i - [\mathcal{H} f(\theta)]_i|^2$

- **Robust criteria:**
  
  $Q(\theta) = \sum_i \phi(|g_i - [\mathcal{H} f(\theta)]_i|)$

  with different functions $\phi$ ($L_1$, Hubert, ...).

- **Likelihood:**
  
  $\mathcal{L}(\theta) = -\ln p(g|\theta)$

- **Penalized likelihood:**
  
  $\mathcal{L}(\theta) = -\ln p(g|\theta) + \lambda \Omega(\theta)$

**Examples:**

- **Spectrometry:** $f(t)$ modelled as a sum of gaussians

  $f(t) = \sum_{k=1}^{K} a_k \mathcal{N}(t|\mu_k, \nu_k) \quad \theta = \{a_k, \mu_k, \nu_k\}$

- **Tomography in CND:** $f(x, y)$ is modelled as a superposition of circular or elliptical discs

  $\theta = \{a_k, \mu_k, r_k\}$
Non parametric methods

\[ g(s_i) = \int h(s_i, r) f(r) \, dr + \epsilon(s_i), \quad i = 1, \cdots, M \]

- \( f(r) \) is assumed to be well approximated by
  \[ f(r) \simeq \sum_{j=1}^{N} f_j b_j(r) \]

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

\[ g(s_i) = g_i \simeq \sum_{j=1}^{N} f_j \int h(s_i, r) b_j(r) \, dr, \quad i = 1, \cdots, M \]

\[ g = H f + \epsilon \quad \text{with} \quad H_{ij} = \int h(s_i, r) b_j(r) \, dr \]

- \( H \) is huge dimensional

- LS solution : \( \hat{f} = \arg \min_f \{Q(f)\} \) with
  \[ Q(f) = \sum_i |g_i - [H f]_i|^2 = \|g - H f\|^2 \]

does not give satisfactory result.
CT as a linear inverse problem

\[ g(s_i) = \int_{L_i} f(r) \, dl_i + \epsilon(s_i) \rightarrow \text{Discretization} \rightarrow g = Hf + \epsilon \]
Classical methods in CT

\[ g(s_i) = \int_{L_i} f(r) \, dl_i + \epsilon(s_i) \rightarrow \text{Discretization} \rightarrow g = Hf + \epsilon \]

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

Data matching

- Observation model
  \[ g_i = h_i(f) + \epsilon_i, \quad i = 1, \ldots, M \rightarrow g = H(f) + \epsilon \]

- Mismatch between data and output of the model \( \Delta(g, H(f)) \)

  \[ \hat{f} = \arg \min_f \{ \Delta(g, H(f)) \} \]

- Examples:
  - LS
    \[ \Delta(g, H(f)) = \| g - H(f) \|^2 = \sum_i |g_i - h_i(f)|^2 \]
  - \( L_p \)
    \[ \Delta(g, H(f)) = \| g - H(f) \|^p = \sum_i |g_i - h_i(f)|^p, \quad 1 < p < 2 \]
  - KL
    \[ \Delta(g, H(f)) = \sum_i g_i \ln \frac{g_i}{h_i(f)} \]

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

Inverse problems = Ill posed problems

\[ \rightarrow \quad \text{Need for prior information} \]

Functional space : Tikhonov :

\[
g = \mathcal{H}(f) + \epsilon \quad \rightarrow \quad J(f) = \| g - \mathcal{H}(f) \|^2 + \lambda \| Df \|^2
\]

Finite dimensional space : \( g = H(f) + \epsilon \) Philips & Tawmey :

- Minimum norme LS (MNLS) :
  \[ J(f) = \| g - H(f) \|^2 + \lambda \| f \|^2 \]

- Classical regularization :
  \[ J(f) = \| g - H(f) \|^2 + \lambda \| Df \|^2 \]

- More general regularization :
  \[ J(f) = Q(g - H(f)) + \lambda \Omega(Df) \]
  or
  \[ J(f) = \Delta_1(g, H(f)) + \lambda \Delta_2(f, f_\infty) \]

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 → 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

\[ \mathcal{M} : \quad g = Hf + \epsilon \]

- Observation model \( \mathcal{M} \) + Hypothesis on the noise \( \epsilon \)
  \begin{equation}
  p(g|f; \mathcal{M}) = p_\epsilon(g - Hf)
  \end{equation}

- A priori information
  \[ p(f|\mathcal{M}) \]

- Bayes:
  \begin{equation}
  p(f|g; \mathcal{M}) = \frac{p(g|f; \mathcal{M}) p(f|\mathcal{M})}{p(g|\mathcal{M})}
  \end{equation}

Link with regularization:

Maximum A Posteriori (MAP):

\[ \hat{f} = \arg \max_f \{ p(f|g) \} = \arg \max_f \{ p(g|f) p(f) \} \]
\[ = \arg \min_f \{ -\ln p(g|f) - \ln p(f) \} \]

with \( Q(g, Hf) = -\ln p(g|f) \) and \( \lambda \Omega(f) = -\ln p(f) \)
Case of linear models and Gaussian priors

\[ g = Hf + \epsilon \]

- Hypothesis on the noise: \( \epsilon \sim \mathcal{N}(0, \sigma^2_\epsilon I) \)
  \[ p(g | f) \propto \exp \left\{ -\frac{1}{2\sigma^2_\epsilon} \| g - Hf \|^2 \right\} \]

- Hypothesis on \( f \): \( f \sim \mathcal{N}(0, \sigma^2_f (D^t D)^{-1}) \)
  \[ p(f) \propto \exp \left\{ -\frac{1}{2\sigma^2_f} \| Df \|^2 \right\} \]

- A posteriori:
  \[ p(f | g) \propto \exp \left\{ -\frac{1}{2\sigma^2_\epsilon} \| g - Hf \|^2 \frac{1}{2\sigma^2_f} \| Df \|^2 \right\} \]

- MAP:
  \[ \hat{f} = \arg \max_f \{ p(f | g) \} = \arg \min_f \{ J(f) \} \]
  with \[ J(f) = \| g - Hf \|^2 + \lambda \| Df \|^2, \quad \lambda = \frac{\sigma^2_\epsilon}{\sigma^2_f} \]

- Advantage: characterization of the solution
  \[ f | g \sim \mathcal{N} (\hat{f}, \hat{P}) \text{ with } \hat{f} = \hat{P} H^t g, \quad \hat{P} = (H^t H + \lambda D^t D)^{-1} \]
MAP estimation with other priors:

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

Separable priors:

- Gaussian: \( p(f_j) \propto \exp \left\{ -\alpha |f_j|^2 \right\} \rightarrow \Omega(\mathbf{f}) = \alpha \sum_j |f_j|^2 \)
- Gamma: \( p(f_j) \propto f_j^\alpha \exp \left\{ -\beta f_j \right\} \rightarrow \Omega(\mathbf{f}) = \alpha \sum_j \ln f_j + \beta f_j \)
- Beta: \( p(f_j) \propto f_j^\alpha (1 - f_j)^\beta \rightarrow \Omega(\mathbf{f}) = \alpha \sum_j \ln f_j + \beta \sum_j \ln(1 - f_j) \)
- Generalized Gaussian: \( p(f_j) \propto \exp \left\{ -\alpha |f_j|^p \right\}, \quad 1 < p < 2 \rightarrow \Omega(\mathbf{f}) = \alpha \sum_j |f_j|^p \)

Markovian models:

\[ p(f_j | \mathbf{f}) \propto \exp \left\{ -\alpha \sum_{i \in N_j} \phi(f_j, f_i) \right\} \rightarrow \Omega(\mathbf{f}) = \alpha \sum_j \sum_{i \in N_j} \phi(f_j, f_i) \]
MAP estimation with markovien priors:

\[ \hat{f} = \arg \min_f \{J(f)\} \quad \text{with} \quad J(f) = \|g - Hf\|^2 + \lambda \Omega(f) \]

\[ \Omega(f) = \sum_j \phi(f_j - f_{j-1}) \]

with \( \phi(t) : \)

Convex functions:

\[ |t|^\alpha, \sqrt{1 + t^2} - 1, \log(\cosh(t)), \left\{ \begin{array}{ll} t^2 & |t| \leq T \\ 2T|t| - T^2 & |t| > T \end{array} \right. \]

or Non convex functions:

\[ \log(1 + t^2), \quad \frac{t^2}{1 + t^2}, \quad \arctan(t^2), \left\{ \begin{array}{ll} t^2 & |t| \leq T \\ T^2 & |t| > T \end{array} \right. \]
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
  - ...

Full Bayesian approach

\[ \mathcal{M} : \quad g = H f + \epsilon \]

- **Forward & errors model**: \[ p(g|f, \theta_1; \mathcal{M}) \]
- **Prior models**: \[ p(f|\theta_2; \mathcal{M}) \]
- **Hyperparameters\( \theta = (\theta_1, \theta_2) \)**: \[ p(\theta|\mathcal{M}) \]
- **Bayes**: \[ p(f, \theta|g; \mathcal{M}) = \frac{p(g|f, \theta; \mathcal{M}) p(f|\theta; \mathcal{M}) p(\theta|\mathcal{M})}{p(g|\mathcal{M})} \]
- **Joint MAP**: \[ (\hat{f}, \hat{\theta}) = \text{arg max}_{(f, \theta)} \{ p(f, \theta|g; \mathcal{M}) \} \]
- **Marginalization**:
  \[
  \begin{align*}
  p(f|g; \mathcal{M}) &= \int p(f, \theta|g; \mathcal{M}) \, df \\
  p(\theta|g; \mathcal{M}) &= \int p(f, \theta|g; \mathcal{M}) \, d\theta
  \end{align*}
  \]
- **Posterior means**:
  \[
  \begin{align*}
  \hat{f} &= \int f \, p(f, \theta|g; \mathcal{M}) \, df \, d\theta \\
  \hat{\theta} &= \int \theta \, p(f, \theta|g; \mathcal{M}) \, df \, d\theta
  \end{align*}
  \]
- **Evidence of the model**:
  \[
  p(g|\mathcal{M}) = \int \int p(g|f, \theta; \mathcal{M}) p(f|\theta; \mathcal{M}) p(\theta|\mathcal{M}) \, df \, d\theta
  \]
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)
Which images I am looking for?
Which image I am looking for?

Gauss-Markov

Generalized GM

Piecewise Gaussian

Mixture of GM
Markovian prior models for images

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

- Gauss-Markov: $$\phi(t) = |t|^2$$
- Generalized Gauss-Markov: $$\phi(t) = |t|^\alpha$$
- Piecewise Gauss-Markov or GGM: $$\phi(t) = \begin{cases} t^2 & |t| \leq T \\ T^2 & |t| > T \end{cases}$$

or equivalently:

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

$q$ line process (contours)

- Mixture of Gaussians:

$$\Omega(f|z) = \sum_k \sum_{\{j:z_j=k\}} \left( \frac{f_j - m_k}{v_k} \right)^2$$

$z$ region labels process.
Gauss-Markov-Potts prior models for images

\[ f(r) \quad z(r) \quad c(r) = 1 - \delta(z(r) - z(r')) \]

\[ p(f(r)|z(r) = k, m_k, v_k) = \mathcal{N}(m_k, v_k) \]

\[ p(f(r)) = \sum_k P(z(r) = k) \mathcal{N}(m_k, v_k) \] Mixture of Gaussians

- Separable iid hidden variables : \( p(z) = \prod_r p(z(r)) \)
- Markovian hidden variables : \( p(z) \) Potts-Markov :

\[ p(z(r)|z(r'), r' \in V(r)) \propto \exp \left\{ \gamma \sum_{r' \in V(r)} \delta(z(r) - z(r')) \right\} \]

\[ p(z) \propto \exp \left\{ \gamma \sum_{r \in R} \sum_{r' \in V(r)} \delta(z(r) - z(r')) \right\} \]
Four different cases

To each pixel of the image is associated 2 variables $f(r)$ and $z(r)$

- $f \mid z$ Gaussian iid, $z$ iid : Mixture of Gaussians
- $f \mid z$ Gauss-Markov, $z$ iid : Mixture of Gauss-Markov
- $f \mid z$ Gaussian iid, $z$ Potts-Markov : Mixture of Independent Gaussians (MIG with Hidden Potts)
- $f \mid z$ Markov, $z$ Potts-Markov : Mixture of Gauss-Markov (MGM with hidden Potts)
Case 1: \( f \mid z \) Gaussian iid, \( z \) iid

Independent Mixture of Independent Gaussians (IMIG):

\[
p(f(r)|z(r) = k) = \mathcal{N}(m_k, v_k), \quad \forall r \in \mathcal{R}
\]
\[
p(f(r)) = \sum_{k=1}^{K} \alpha_k \mathcal{N}(m_k, v_k), \text{ with } \sum_k \alpha_k = 1.
\]
\[
p(z) = \prod_r p(z(r) = k) = \prod_r \alpha_{k} = \prod_k \alpha_k^{n_k}
\]

Noting

\[
m_z(r) = m_k, v_z(r) = v_k, \alpha_z(r) = \alpha_k, \forall r \in \mathcal{R}_k
\]

we have:

\[
p(f|z) = \prod_{r \in \mathcal{R}} \mathcal{N}(m_z(r), v_z(r))
\]
\[
p(z) = \prod_r \alpha_z(r) = \prod_k \alpha_k^{\sum_{r \in \mathcal{R}} \delta(z(r)-k)} = \prod_k \alpha_k^{n_k}
\]
Case 2: \( f \sim \text{Gauss-Markov}, \quad z \sim \text{iid} \)

Independent Mixture of Gauss-Markov (IMGM):

\[
p(f(r) | z(r), z(r'), f(r'), r' \in \mathcal{V}(r)) = \mathcal{N}(\mu_z(r), \nu_z(r)), \forall r \in \mathcal{R}
\]

\[
\mu_z(r) = \frac{1}{|\mathcal{V}(r)|} \sum_{r' \in \mathcal{V}(r)} \mu_{z}^{*}(r')
\]

\[
\mu_{z}^{*}(r') = \delta(z(r') - z(r)) f(r') + (1 - \delta(z(r') - z(r))) m_z(r')
\]

\[
= (1 - c(r')) f(r') + c(r') m_z(r')
\]

\[
p(f | z) \propto \prod_r \mathcal{N}(\mu_z(r), \nu_z(r)) \propto \prod_k \alpha_k \mathcal{N}(m_k 1, \Sigma_k)
\]

\[
p(z) = \prod_r \nu_z(r) = \prod_k \alpha_k^{n_k}
\]

with \( 1_k = 1, \forall r \in \mathcal{R}_k \) and \( \Sigma_k \) a covariance matrix \((n_k \times n_k)\).
Case 3: \( f|z \) Gauss iid, \( z \) Potts

Gauss iid as in Case 1:

\[
p(f|z) = \prod_{r \in \mathcal{R}} \mathcal{N}(m_z(r), v_z(r)) = \prod_k \prod_{r \in \mathcal{R}_k} \mathcal{N}(m_k, v_k)
\]

Potts-Markov

\[
p(z(r)|z(r'), r' \in \mathcal{V}(r)) \propto \exp \left\{ \gamma \sum_{r' \in \mathcal{V}(r)} \delta(z(r) - z(r')) \right\}
\]

\[
p(z) \propto \exp \left\{ \gamma \sum_{r \in \mathcal{R}} \sum_{r' \in \mathcal{V}(r)} \delta(z(r) - z(r')) \right\}
\]
Case 4: $f|z$ Gauss-Markov, $z$ Potts

Gauss-Markov as in Case 2:

$$p(f(r)|z(r), z(r'), f(r'), r' \in \mathcal{V}(r)) = \mathcal{N}(\mu_z(r), \nu_z(r)), \forall r \in \mathcal{R}$$

$$\mu_z(r) = \frac{1}{|\mathcal{V}(r)|} \sum_{r' \in \mathcal{V}(r)} \mu_z^*(r')$$

$$\mu_z^*(r') = \delta(z(r') - z(r)) f(r') + (1 - \delta(z(r') - z(r)) m_z(r')$$

$$p(f|z) \propto \prod_r \mathcal{N}(\mu_z(r), \nu_z(r)) \propto \prod_k \alpha_k \mathcal{N}(m_k1, \Sigma_k)$$

Potts-Markov as in Case 3:

$$p(z) \propto \exp \left\{ \gamma \sum_{r \in \mathcal{R}} \sum_{r' \in \mathcal{V}(r)} \delta(z(r) - z(r')) \right\}$$
Summary of the two proposed models

\( f | \sim \text{Gaussian iid} \)
\( z \sim \text{Potts-Markov} \)

(MIG with Hidden Potts)

\( f | \sim \text{Markov} \)
\( z \sim \text{Potts-Markov} \)

(MGM with hidden Potts)
Bayesian Computation

\[
p(f, z, \theta | g) \propto p(g | f, z, \nu) p(f | z, m, v) p(z | \gamma, \alpha) p(\theta)
\]

\[
\theta = \{ \nu, (\alpha_k, m_k, v_k), k = 1, \ldots, K \} \quad p(\theta) \quad \text{Conjugate priors}
\]

- Direct computation and use of \( p(f, z, \theta | g; M) \) is too complex

- Possible approximations:
  - Gauss-Laplace (Gaussian approximation)
  - Exploration (Sampling) using MCMC methods
  - Separable approximation (Variational techniques)

- Main idea in Variational Bayesian methods:
  Approximate
  \[
p(f, z, \theta | g; M) \quad \text{by} \quad q(f, z, \theta) = q_1(f) q_2(z) q_3(\theta)
\]

  - Choice of approximation criterion: \( KL(q : p) \)
  - Choice of appropriate families of probability laws for \( q_1(f) \), \( q_2(z) \) and \( q_3(\theta) \)
MCMC based algorithm

\[ p(f, z, \theta | g) \propto p(g | f, z, \theta) p(f | z, \theta) p(z) p(\theta) \]

General scheme:

\[ \hat{f} \sim p(f | \hat{z}, \hat{\theta}, g) \rightarrow \hat{z} \sim p(z | \hat{f}, \hat{\theta}, g) \rightarrow \hat{\theta} \sim (\theta | \hat{f}, \hat{z}, g) \]

- Estimate \( f \) using \( p(f | \hat{z}, \hat{\theta}, g) \propto p(g | f, \theta) p(f | \hat{z}, \hat{\theta}) \)
  Needs optimisation of a quadratic criterion.

- Estimate \( z \) using \( p(z | \hat{f}, \hat{\theta}, g) \propto p(g | \hat{f}, \hat{z}, \hat{\theta}) p(z) \)
  Needs sampling of a Potts Markov field.

- Estimate \( \theta \) using
  \[ p(\theta | \hat{f}, \hat{z}, g) \propto p(g | \hat{f}, \sigma_c^2 I) p(\hat{f} | \hat{z}, (m_k, v_k)) p(\theta) \]
  Conjugate priors \( \rightarrow \) analytical expressions.
Application of CT in NDT

Reconstruction from only 2 projections

\[ g_1(x) = \int f(x, y) \, dy, \quad g_2(y) = \int f(x, y) \, dx \]

- Given the marginals \( g_1(x) \) and \( g_2(y) \) find the joint distribution \( f(x, y) \).
- Infinite number of solutions: \( f(x, y) = g_1(x) \, g_2(y) \, \Omega(x, y) \)

\( \Omega(x, y) \) is a Copula:

\[ \int \Omega(x, y) \, dx = 1 \quad \text{and} \quad \int \Omega(x, y) \, dy = 1 \]
Application in CT

\[
g | f = Hf + \epsilon
\]

\[
g | f \sim \mathcal{N}(Hf, \sigma^2 I)
\]

\[f | z \text{ iid Gaussian}
\]

\[f | z \text{ or Gauss-Markov}
\]

\[z \text{ iid or Potts}
\]

\[c(r) \in \{0, 1\}
\]

\[1 - \delta(z(r) - z(r'))
\]

binary
Proposed algorithm

\[ p(f, z, \theta | g) \propto p(g | f, z, \theta) p(f | z, \theta) p(\theta) \]

General scheme:

\[ \hat{f} \sim p(f | \hat{z}, \hat{\theta}, g) \rightarrow \hat{z} \sim p(z | \hat{f}, \hat{\theta}, g) \rightarrow \hat{\theta} \sim (\theta | \hat{f}, \hat{z}, g) \]

- Estimate \( f \) using \( p(f | \hat{z}, \hat{\theta}, g) \propto p(g | f, \theta) p(f | \hat{z}, \hat{\theta}) \)
  Needs optimisation of a quadratic criterion.

- Estimate \( z \) using \( p(z | \hat{f}, \hat{\theta}, g) \propto p(g | \hat{f}, \hat{z}, \hat{\theta}) p(z) \)
  Needs sampling of a Potts Markov field.

- Estimate \( \theta \) using
  \[ p(\theta | \hat{f}, \hat{z}, g) \propto p(g | \hat{f}, \sigma^2_\epsilon I) p(\hat{f} | \hat{z}, (m_k, v_k)) p(\theta) \]
  Conjugate priors \( \rightarrow \) analytical expressions.
Results

Original

Backprojection

Filtered BP

LS

Gauss-Markov+pos

GM+Line process

GM+Label process
Application in Microwave imaging

\[ g(\omega) = \int f(r) \exp \{-j(\omega \cdot r)\} \, dr + \epsilon(\omega) \]

\[ g(u, v) = \int f(x, y) \exp \{-j(ux + vy)\} \, dx \, dy + \epsilon(u, v) \]

\[ g = H \hat{f} + \epsilon \]

\[ f(x, y) \quad g(u, v) \quad \hat{f} \quad \text{IFT} \quad \hat{f} \quad \text{Proposed method} \]
Conclusions

- Bayesian Inference for inverse problems
- Approximations (Laplace, MCMC, Variational)
- Gauss-Markov-Potts are useful prior models for images incorporating regions and contours
- Separable approximations for Joint posterior with Gauss-Markov-Potts priors
- Application in different CT (X ray, US, Microwaves, PET, SPECT)

Perspectives:

- Efficient implementation in 2D and 3D cases
- Evaluation of performances and comparison with MCMC methods
- Application to other linear and non linear inverse problems: (PET, SPECT or ultrasound and microwave imaging)
Some references


Questions and Discussions

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