Hierarchical Prior Models and Krylov-Bayes Iterative Methods: Part 1

Daniela Calvetti

Case Western Reserve University

- D Calvetti, E Somersalo: A Gaussian hypermodel to recover blocky objects. Inverse Problems 23 (2007) 733
- D Calvetti, E Somersalo: Hypermodel in Bayesian imaging framework. Inverse Problems 24 (2008) 034013
- D Calvetti E Somersalo: Statistical Methods in Imaging. in Otmar Scherzer (ed.), Handbook of Mathematical Methods in Imaging. Springer Science+Business Media LLC (2011)
- J Bardsley, D Calvetti E Somersalo: Hierarchical regularization for edge-preserving reconstruction of PET images. Inverse Problems 26 (2010) 035010 (16pp)

Linear discrete Inverse Problems

We consider the linear discrete inverse problem of estimating $x \in \mathbb{R}^n$ from

$$b = \mathrm{A}x + \epsilon$$

where

- A is an $m \times n$ matrix, typically badly conditioned and of ill-determined rank
- ϵ is additive noise, which we assume zero-mean white Gaussian
- x is a discretized signal

$$x_j = f(t_j), \quad t_j = \frac{j}{n}, \quad 0 \le j \le n,$$

Priors as beliefs

Consider two prior models expressing our belief about the signal before taking the data into account.

1. **Prior model 1**:
   - We know that $x_0 = 0$,
   - We believe that the absolute value of the slope of f is bounded by some $m_1 > 0$.
2. **Prior model 2**:
   - We know that $x_0 = x_n = 0$
   - We believe that the curvature of f is bounded by some $m_2 > 0$.

Formalization of prior beliefs

1. **Prior model 1**:
   - Slope:
   $$f'(t_j) \approx \frac{x_j - x_{j-1}}{h}, \quad h = \frac{1}{n},$$
   - Prior information: We believe that
   $$|x_j - x_{j-1}| \leq h\, m_1 \text{ with some uncertainty.}$$
2. **Prior model 2**:
   - Curvature:
   $$f''(t_j) \approx \frac{x_{j-1} - 2x_j + x_{j+1}}{h^2}.$$
   - Prior information: We believe that
   $$|x_{j-1} - 2x_j + x_{j+1}| \leq h^2 m_2 \text{ with some uncertainty.}$$

Assigning boundary conditions

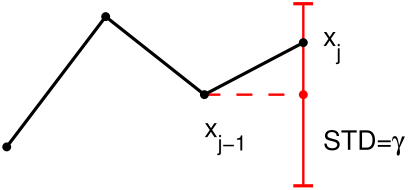
Following the Bayesian paradigm, in both cases, we assume that x_j is a realization of a random variable X_j.
Boundary conditions:

1. **Prior model 1**: $X_0 = 0$ with certainty. Probabilistic model for X_j, $1 \leq j \leq n$.
2. **Prior model 2** $X_0 = X_n = 0$ with certainty. Probabilistic model for X_j, $1 \leq j \leq n-1$.

Priors as autoregressive models

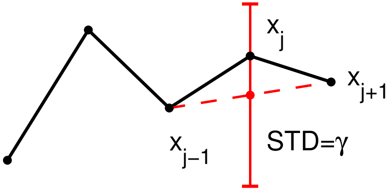
1. First order prior:

$$X_j = X_{j-1} + \gamma W_j, \quad W_j \sim \mathcal{N}(0,1), \quad \gamma = h\, m_1.$$



2. Second order prior:

$$X_j = \frac{1}{2}(X_{j-1} + X_{j+1}) + \gamma W_j, \quad W_j \sim \mathcal{N}(0,1), \quad \gamma = \frac{1}{2} h^2\, m_2.$$



Matrix form: first order model

System of equations:

$$
\begin{aligned}
X_1 = X_1 - X_0 &= \gamma W_1 \\
X_2 - X_1 &= \gamma W_2 \\
\vdots \quad &\quad \vdots \\
X_n - X_{n-1} &= \gamma W_n
\end{aligned}
$$

$$
\mathsf{L}_1 = \begin{bmatrix} 1 & & & \\ -1 & 1 & & \\ & \ddots & \ddots & \\ & & -1 & 1 \end{bmatrix} \in \mathbb{R}^{n\times n}, \quad X = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix}, \quad W = \begin{bmatrix} W_1 \\ W_2 \\ \vdots \\ W_n \end{bmatrix}.
$$

$$
\mathsf{L}_1 X = \gamma W, \quad W \sim \mathcal{N}(0, \mathsf{I}_n),
$$

Matrix form: second order model

System of equations:

$$\begin{aligned}
X_2 - 2X_1 = X_2 - 2X_1 + X_0 &= \gamma W_1 \\
X_3 - 2X_2 + X_1 &= \gamma W_2 \\
\vdots \quad & \quad \vdots \\
-2X_{n-1} - X_{n-2} = X_n - 2X_{n-1} + X_{n-2} &= \gamma W_{n-1}
\end{aligned}$$

$$\mathsf{L}_2 = \begin{bmatrix} -2 & 1 & & \\ 1 & -2 & 1 & \\ & \ddots & \ddots & \\ & & 1 & -2 \end{bmatrix} \in \mathbb{R}^{(n-1)\times(n-1)}, \quad X = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_{n-1} \end{bmatrix},$$

$$\mathsf{L}_2 X = \gamma W, \quad W \sim \mathcal{N}(0, \mathsf{I}_{n-1}),$$

Testing a Prior

But

$$\Gamma^{-1} = \mathsf{L}^{\mathsf{T}}\mathsf{L} \Rightarrow \Gamma = \left(\mathsf{L}^{\mathsf{T}}\mathsf{L}\right)^{-1} = \mathsf{L}^{-1}\mathsf{L}^{-\mathsf{T}}$$

Therefore,

$$\mathsf{L}\Gamma\mathsf{L}^{\mathsf{T}} = \mathsf{L}\left(\mathsf{L}^{-1}\mathsf{L}^{-\mathsf{T}}\right)\mathsf{L}^{\mathsf{T}} = \mathsf{I}_n.$$

Conclusion: Given $X \sim \mathcal{N}(x_0, \Gamma)$,

$$\boxed{W = \mathsf{L}(X - x_0) \sim \mathcal{N}(0, \mathsf{I}_n).}$$

This transform is called *whitening* of X, or *Mahalanobis transformation* of X.
(Recall: a noise vector $E \sim \mathcal{N}(0, \mathsf{I}_n)$ is called *white noise*.)

Testing a Prior

Conversely:

$$W = \mathsf{L}(X - x_0) \Rightarrow X = \mathsf{L}^{-1}W + x_0.$$

Therefore, if W is white noise, then $X \sim \mathcal{N}(x_0, \Gamma)$.
This observation allows us to generate random draws from the distribution $\mathcal{N}(x_0, \Gamma)$:

Sampling from Gaussian densities

Repeat N times:

1. Draw a realization $w \sim \mathcal{N}(0, \mathsf{I}_n)$
2. Set $x = x_0 + \mathsf{L}^{-1}w$.

Random draws from priors

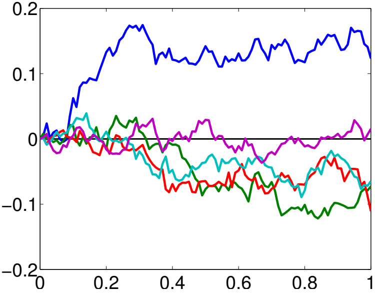
Generate m draws from the prior using the Matlab command `randn`.

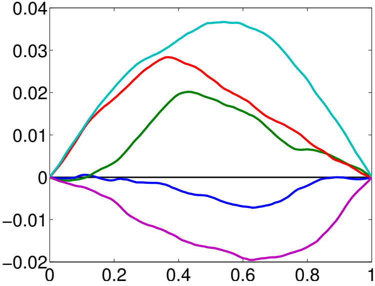
```
n = 100;         % number of discretization intervals
t = (0:1/n:1);
m = 5;          % number of draws

% First order model. Boundary condition X_0 = 0

L1    = diag(ones(1,n),0) - diag(ones(1,n-1),-1);
gamma = 1/n;  % m_1 = 1
W     = gamma*randn(n,m);
X     = L1\W;
```

Plots of the random draws





Adding structure

Assume that there is a reason to believe that the slope (or the curvature) may be 10 times higher at isolated points.
If t_k is such point, we replace the condition

$$X_k - X_{k-1} = \gamma W_k$$

by the modified condition

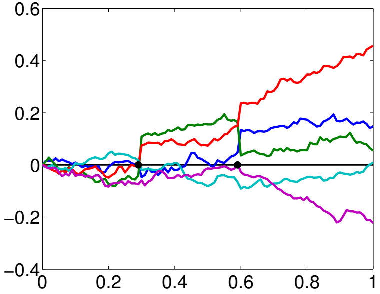
$$X_k - X_{k-1} = 10\gamma W_k.$$

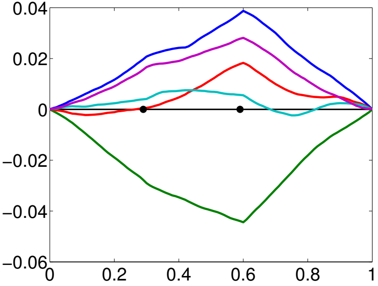
In Matlab

```
n = 100;        % number of discretization intervals
t = (0:1/n:1);
m = 5;         % number of draws
k = 30;        % position of the jump

L1     = diag(ones(1,n),0) - diag(ones(1,n-1),-1);
gamma = (1/n)*ones(n,1);
gamma(k) = 10*gamma(k);
W     = diag(gamma)*randn(n,m);
X     = L1\W;
```

Plots of the random draws





The jumps (or kinks) are allowed, but not forced.

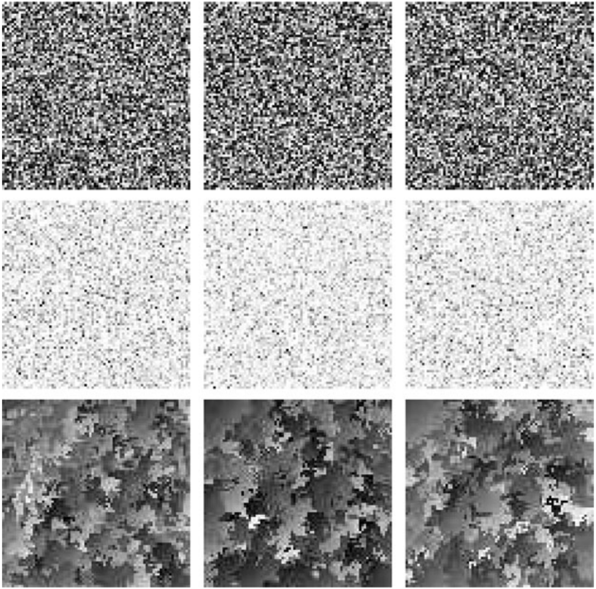


Fig. 21-2

Random draws from various MRF priors. *Top row*: white noise prior. *Middle row*: sparsity prior or ℓ^1-prior with positivity constraint. *Bottom row*: total variation prior

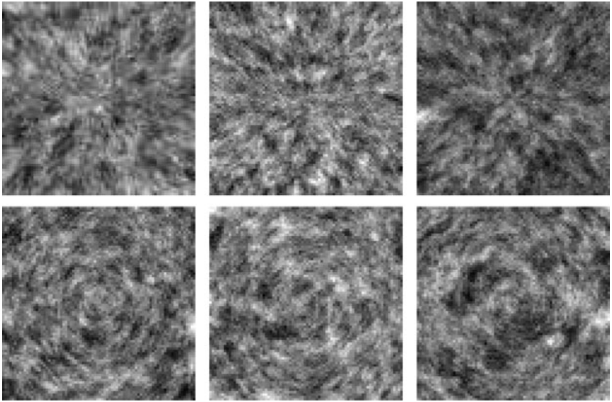


Fig. 21-1

Random draws from anisotropic Markov models. In the *top row*, the Markov model assumes stronger dependency between neighboring pixels in the radial than in angular direction, while in the *bottom row* the roles of the directions are reversed. See text for a more detailed discussion

Hypermodels

If the bound on the slope (curvature) is not known

- we can model it also as a random variable θ.
- The prior takes on the form

$$X_k - X_{k-1} = \theta_k^{1/2} W_k$$

- The parameter θ_k is the variance of the Gaussian innovation
- θ_k quantifies the uncertainty in going from X_{k-1} to X_k

Matrix form of hypermodels

In matrix-vector terms

$$LX = D^{1/2}W$$

where

$$D = diag\{\theta_1, \theta_2, \dots, \theta_n\}.$$

Since W is an n-variate standard normal, we can write the probability density of X as

$$\pi_{\mathrm{prior}}(x) \propto \exp\left(-\frac{1}{2}\|D^{-1/2}LX\|^2\right).$$

Quantitative prior

If we have information about the

- location
- number
- expected amplitude

of the jumps, it should be encoded in the first order Markov model by setting the corresponding θs.

Qualitative prior

If we only know that *jumps may occur* but no information about how many, where and how big is available, then

- The variance of the innovation is unknown.
- The variance is modeled as a random variable
- The estimation of the variance of the Markov process is part of the inverse problem
- The prior for the problem is the joint prior for X and Θ

$$\pi_{\mathrm{prior}}(x, \theta) = \pi_{\mathrm{prior}}(x \mid \theta)\pi_{\mathrm{hyper}}(\theta)$$

Conditional smoothness prior

If we had the variance information, the original smoothness prior for X would be determined. Since the variance vector is unknown, we cannot ignore the normalizing factor

$$\pi_{\mathrm{prior}}(x \mid \theta) = \left(\frac{det(L^T D_\theta^{-1} L)}{(2\pi)^n} \right)^{1/2} \exp\left(-\frac{1}{2} \| D_\theta^{-1/2} L x \|^2 \right)$$

If L is invertible, there is an analytic expression for the determinant

- If L is not invertible, introduce

$$Z_j = X_j - X_{j-1}, \quad LX = Z$$

Then

$$\begin{aligned}
\pi_{\mathrm{prior}}(z \mid \theta) &= \left(\frac{det(D_\theta^{-1})}{(2\pi)^n}\right)^{1/2} \exp\left(-\frac{1}{2}\|D_\theta^{-1/2} z\|^2\right) \\
&= \frac{1}{(2\pi)^{n/2}} \exp\left(-\frac{1}{2}\|D_\theta^{-1/2} z\|^2 - \frac{1}{2}\sum_{j=1}^{n} \log \theta_j\right)
\end{aligned}$$

Sparsity promoting priors

Assume that we expect a priori a smooth signal with few discontinuities. Then:

- The jumps should be sudden thus the variances should be independent
- There is no preference about the locations of the jumps, thus the components should be identically distributed
- Only a few variances can be large, most should be close to zero Therefore the hyperprior $\pi_{\mathrm{hyper}}(\theta)$ should be such that
  - a few variances can be significantly large
  - most of the variances are zero

Hyperpriors

The candidate probability densities for hyperpriors are

1. The gamma distribution

$$\Theta_j \sim \mathrm{Gamma}(\alpha, \theta_0), \quad \pi_{\mathrm{hyper}}(\theta) \propto \prod_{j=1}^{n} \theta_j^{\alpha-1} \exp\left(-\frac{\theta_j}{\theta_0}\right)$$

   with mean and variance $\alpha\theta_0$ and $\alpha\theta_0^2$

2. The inverse gamma distribution

$$\Theta_j \sim \mathrm{InvGamma}(\alpha, \theta_0), \quad \pi_{\mathrm{hyper}}(\theta) \propto \prod_{j=1}^{n} \theta_j^{-\alpha-1} \exp\left(-\frac{\theta_0}{\theta_j}\right)$$

   with mean and variance $\theta_0/(\alpha-1)$ and $\theta_0^2/(\alpha-1)^2(\alpha-2)$ when $\alpha > 2$.

Gamma vs Inverse Gamma

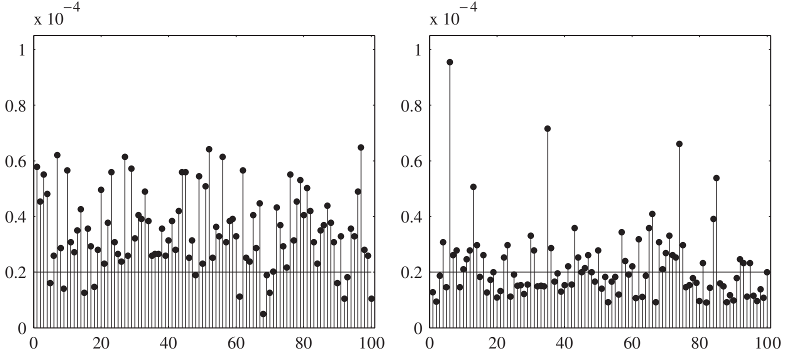


Figure 2. One hundred random draws from the gamma distribution (left) and inverse gamma distribution (right). The parameters are set so that the means and variance match. The mean is indicated in the figures by a horizontal line.

Assume the data are contaminated by additive zero-mean Gaussian noise with covariance $\sigma^2 I$

$$B = AX + E, \qquad E \sim N(0, \sigma^2 I)$$

leading to

$$\pi(b \mid z) \approx \exp\left(-\frac{1}{2\sigma^2}\|b - AL^{-1}z\|^2\right)$$

Note that

$$\pi(b \mid z, \theta) = \pi(b \mid z),$$

i.e., the likelihood does not depend on the hyperparameters.
It follows from Bayes formula that the posterior of (Z, Θ) given $B = b$ is

$$\pi(z, \theta \mid b) \propto \pi_{\rm hyper}(\theta)\pi_{prior}(z \mid \theta)\pi(b \mid z)$$

MAP estimates computation

To calculate the MAP estimate

$$(z_{\mathrm{MAP}}, \theta_{\mathrm{MAP}}) = \mathrm{argmax} \pi(z, \theta \mid b)$$

we apply the following Iterative Alternating Scheme (IAS):

1. Initialize $\theta = \theta_0, \quad k = 1$
2. Update the estimate of the increments z:
   $z^k = \mathrm{argmax} \pi(z, \theta^{k-1} \mid b)$
3. Update the variances θ: $\theta^k = \mathrm{argmax} \pi(z^k, \theta \mid b)$.
4. Repeat from 2. until convergence

Gamma and inverse gamma hyperprior

If we denote the negative log-posterior by

$$F(z, \theta \mid b) = -\log(\pi(z, \theta \mid b))$$

and use a gamma hyperprior then

$$\begin{aligned} F(z, \theta \mid b) \quad \simeq \quad & \frac{1}{2\sigma^2}\|AL^{-1}z - b\|^2 + \frac{1}{2}\|D^{-1/2}z\|^2 + \\ & \frac{1}{\theta_0}\sum_{j=1}^{n}\theta_j - \left(\alpha - \frac{3}{2}\right)\sum_{j=1}^{n}\log\theta_j \end{aligned}$$

while with the inverse gamma hyperprior

$$\begin{aligned} F(z, \theta \mid b) \quad \simeq \quad & \frac{1}{2\sigma^2}\|AL^{-1}z - b\|^2 + \frac{1}{2}\|D^{-1/2}z\|^2 + \\ & \theta_0\sum_{j=1}^{n}\frac{1}{\theta_j} - \left(\alpha + \frac{3}{2}\right)\sum_{j=1}^{n}\log\theta_j \end{aligned}$$

The IAS algorithm

The efficiency of the alternating scheme comes form the fact that

- The functional to be minimized is quadratic in z
- The parameters θ_j are mutually independent
- There is an explicit formula for the minimizer with respect to θ

Next we will look at the two steps separately.

Updating z and θ

1. Updating z is tantamount to solving the linear system

$$\begin{bmatrix} \sigma^{-1}AL^{-1} \\ D^{-1/2} \end{bmatrix} z = \begin{bmatrix} \sigma^{-1}b \\ 0 \end{bmatrix}, \qquad x = L^{-1}z.$$

2. Updating θ_j requires solving

$$\frac{\partial}{\partial \theta_j} F(z, \theta \mid b) = 0, \qquad 1 \leq j \leq n.$$

When the hyperprior is the gamma distribution the solution is

$$\theta_j = \theta_0 \left(\eta + \sqrt{\frac{z_j^2}{2\theta_0} + \eta^2} \right), \quad \eta = \frac{1}{2}\left(\alpha - \frac{3}{2}\right)$$

and when the hyperprior is the inverse gamma

$$\theta_j = \frac{1}{\alpha + 3/2}\left(\theta_0 + \frac{1}{2} z_j^2\right)$$

Extension to 2D

Consider a gray scale image supported on unit square and represent it as an $n \times n$ matrix. Recalling that the Kronecker product of two matrices (A, B) is defined as

$$A \otimes B = [a_{ij} B],$$

representing the columns of the image matrix by x_j,

$$L_1 x = \begin{bmatrix} L & & \\ & \ddots & \\ & & L \end{bmatrix} \begin{bmatrix} x^{(1)} \\ \vdots \\ x^{(n)} \end{bmatrix} = \begin{bmatrix} Lx^{(1)} \\ \vdots \\ Lx^{(n)} \end{bmatrix},$$

that is,

$$L_1 = I_n \otimes L.$$

Extension to 2D

Similarly

$$L_2 x = \begin{bmatrix} I_n & & & & \\ -I_n & I_n & & & \\ & \ddots & \ddots & & \\ & & & -I_n & I_n \end{bmatrix} \begin{bmatrix} x^{(1)} \\ \vdots \\ x^{(n)} \end{bmatrix} = \begin{bmatrix} x^{(1)} \\ x^{(2)} - x^{(1)} \\ \vdots \\ x^{(n)} - x^{(n-1)} \end{bmatrix},$$

that is,

$$L_2 = L \otimes I_n.$$

Denote the vertical and horizontal jump vectors by

$$v = L_1 x, \qquad h = L_2 x.$$

and by V and H their stochastic extensions.

$$\pi_{\mathrm{prior}}(v, h \mid \theta) = \frac{1}{(2\pi)^N} \exp\left(-\frac{1}{2} \sum_{j=1}^{N} \frac{v_j^2 + h_j^2}{\theta_j} - \sum_{j=1}^{N} \log \theta_j \right)$$

If the variance of the jumps is unknown, we model it as a random variable ad use either a gamma or an inverse gamma hyperprior.

Challenges in 2D

$$\begin{bmatrix} V \\ H \end{bmatrix} = MX, \qquad M = \begin{bmatrix} L_1 \\ L_2 \end{bmatrix} \in B^{2N \times N}$$

$$M = QR = [Q_1 \quad Q_2] \begin{bmatrix} R_1 \\ 0 \end{bmatrix}$$

$$R_1 X = Q_1^T \begin{bmatrix} V \\ H \end{bmatrix}, \qquad Q_2^T \begin{bmatrix} V \\ H \end{bmatrix} = 0$$

The second expression is a compatibility condition:

- if the vectors V and H come from an image, the circulation around a vertex where 4 pixels meet must vanish

Likelihood in 2D

We write the likelihood as a singular Gaussian density in terms of V and H

$$\begin{aligned}\pi(b \mid v, h) \quad &\propto \quad \delta\left(Q_2^T \left[\begin{array}{c} v \\ h \end{array}\right]\right) \cdot \\ &\qquad \exp\left(-\frac{1}{2\sigma^2}\|AR_1^{-1}Q_1^T \left[\begin{array}{c} v \\ h \end{array}\right] - b\|^2\right)\end{aligned}$$

where the Dirac delta ensures that the support of the density is in the subspace orthogonal to the range of Q_2.

MAP in 2D

Due to the singularity of the posterior, the MAP is the maximizer of the non-singular part restricted to the constraint subspace. In the case of the gamma hyperprior we minimize

$$
\begin{aligned}
F(v, h, \theta \mid b) \quad \simeq \quad & \frac{1}{2\sigma^2}\|AR_1^{-1}Q_1^T \begin{bmatrix} v \\ h \end{bmatrix} - b\|^2 \\
+ \quad & \frac{1}{2}\sum_{j=1}^{N} \frac{v_j^2 + h_j^2}{\theta_j} + \sum_{j=1}^{N} \frac{\theta_j}{\theta_0} - (\alpha - 2)\sum_{j=1}^{N} \log \theta_j
\end{aligned}
$$

subject to

$$
Q_2^T \begin{bmatrix} v \\ h \end{bmatrix} = 0
$$

This is equivalent to minimizing

$$
\begin{aligned}
F(x,\theta \mid b) &\simeq \frac{1}{2\sigma^2}\|Ax - b\|^2 \\
&+ \frac{1}{2}\sum_{j=1}^{N}\frac{(L_1x)_j^2 + (L_2x)_j^2}{\theta_j} + \sum_{j=1}^{N}\frac{\theta_j}{\theta_0} - (\alpha - 2)\sum_{j=1}^{N}\log\theta_j
\end{aligned}
$$

where

$$
L_1x = v, \qquad L_2x = h.
$$

In particular, updating (v,h) requires solving in the least squares sense

$$
\left[\begin{array}{c} \sigma^{-1}A \\ D^{-1/2}L_1 \\ D^{-1/2}L_2 \end{array}\right] x = \left[\begin{array}{c} \sigma^{-1}b \\ 0 \\ 0 \end{array}\right].
$$

Updating θ in 2D

The formula for updating θ_j for the gamma hyperprior

$$\theta_j = \theta_0 \left(\eta + \sqrt{\frac{v_j^2 + h_j^2}{2\theta_0} + \eta^2} \right) \qquad \eta = \frac{1}{2}(\alpha - 2)$$

and for the inverse gamma hyperprior

$$\theta_j = \frac{1}{\alpha + 2} \left(\theta_0 + \frac{1}{2}(v_j^2 + h_j^2) \right)$$

TV and Perona Malik

TV penalty

$$TV(f) = \int_{\Omega} | \nabla f|$$

The discrete counterpart of the total variation is

$$TV(x) = \sum_{j=1}^{N} \sqrt{v_j^2 + h_j^2}, \qquad v = L_1 x,\ h = L_2 x.$$

The TV-penalized solution of $Ax = b$ is

$$\widehat{x}_{\mathrm{TV}} = \mathrm{argmin}(\|Ax - b\|^2 + \delta\, TV(x)), \qquad \delta > 0.$$

while the Tikhonov regularized solution with a quadratic penalty is

$$\widehat{x}_{\mathrm{Tikh}} = \mathrm{argmin} \left(\|Ax - b\|^2 + \delta \sum_{j=1}^{N} (v_j^2 + h_j^2) \right),$$

Write

$$TV(x) = \sum_{j=1}^{N} \sqrt{v_j^2 + h_j^2} = \sum_{j=1}^{N} \frac{v_j^2 + h_j^2}{\sqrt{v_j^2 + h_j^2}} = \|W_x^{1/2} L x\|^2$$

where

$$L = \begin{bmatrix} L_1 \\ L_2 \end{bmatrix}, \qquad W_x = I_2 \otimes diag\left(\frac{1}{\sqrt{(L_1 x)_j^2 + (L_2 x)_j^2)}}\right)$$

The nonlinear normal equations corresponding to TV regularization are

$$A^T A x + \delta L^T W_x L x = A^T b$$

that can be seen as an asymptotic state of a nonlinear diffusion equation

$$\frac{\partial x}{\partial t}(t) = A^T (A x(t) - b) + \delta L^T W_{x(t)} L x(t)$$

and use a time marching scheme to drive the initial state to a steady state.

TV penalty and gamma

Alternatively, given initial guess x^0 , let the kth iterate satisfy

$$A^T Ax + \delta L^T W^{k-1} Lx = A^T b, \qquad W^{k-1} = W_{x^{k-1}}.$$

- This is equivalent to the fixed point iteration for TV penalty.
- This is also the IAS updating step with gamma hyperprior and $\alpha = 2$, $\delta = \sqrt{2/\sigma_0}\sigma^2$.
- The choice $\alpha = 2 + \epsilon$ prevents prior variance from vanishing, stabilizing IAS algorithm and TV penalty as well.

Perona-Malik and inverse gamma

If we let

$$W_x = I_2 \otimes \mathrm{diag}\left(\frac{1}{1+\beta\left((L_1x)_j^2+(L_2x)_j^2\right)}\right)$$

- the solution of the nonlinear normal equations is the Perona-Malik regularized solution.
- The fixed point scheme is equivalent to the IAS for inverse gamma hyperprior with $\beta = 1/2\theta_0$, and $\delta = (\alpha+2)\theta_0\sigma^2$.

Quantification of uncertainty

- Sampling the posterior is a means of quantifying uncertainly in the solution
- In the case of images the number of unknowns is (twice) the number of pixels
- Using MCMC techniques may be unfeasible for such high dimensions
- Sampling the posterior in a ROI reduces the complexity of the problem

Sampling the ROI

- Let I_{ROI} be the set of indices of the pixels X' in the ROI.
- Partition $X = [X'; X"]$, $\Theta = [\Theta'; \Theta"]$
- Given an estimate of $(X_{\mathrm{MAP}}, \Theta_{\mathrm{MAP}})$ consider the conditional posterior

$$\pi(x', \theta' \mid b, \theta"_{\mathrm{MAP}}, \theta"_{\mathrm{MAP}}) \propto \pi(x, \theta \mid b)_{x"=x"_{\mathrm{MAP}}, \theta"=\theta"_{\mathrm{MAP}}}$$

MCMC for ROI

- Initialize: $\theta^0 = \theta_{\mathrm{MAP}}, \quad x^0 = x_{\mathrm{MAP}}; \quad k = 0$
- Update x: Draw x'_+ from conditional distribution

$$\pi(x' \mid b, \theta^k, x"_{\mathrm{MAP}})$$

  and set $x^{k+1} = [x'_+; x"_{\mathrm{MAP}}]$.
- Update θ: draw θ'_+ from the conditional distribution

$$\pi(\theta' \mid b, x^{k+1}, \theta"_{\mathrm{MAP}})$$

  and set $\theta^{k+1} = [\theta'_+; \theta"_{\mathrm{MAP}}]$
- Continue until desired sample size is reached

Efficiency of scheme

- The conditional density is Gaussian, therefore each random draw requires the solution of one linear system of the size of the ROI
- The θ_j are independent, thus the draw of θ can be done component by component
- More on sampling this afternoon and tomorrow

Computed examples

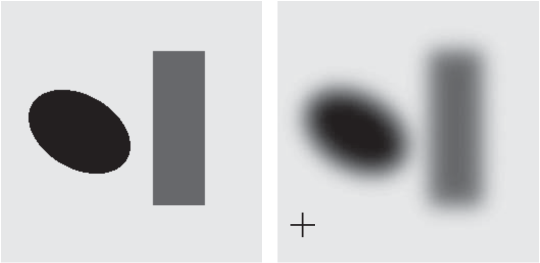


Figure 3. Original and blurred image. The cross in the bottom left corner of the blurred image indicates the half-width of the Gaussian blur.

Iterative MAP estimation: GMRES

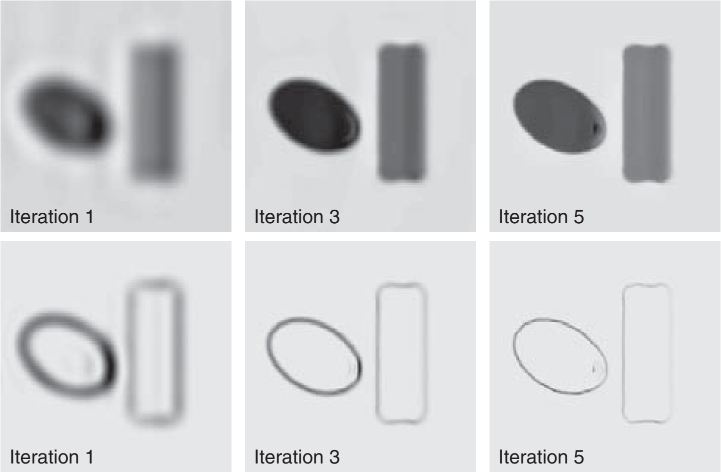


Figure 4. Approximation of the MAP Estimate of the image (top row) and of the variance (bottom row) after 1, 3 and 5 iteration of the cyclic algorithm when using the GMRES method to compute the updated of the image at each iteration step.

Iterative MAP estimation: CGLS

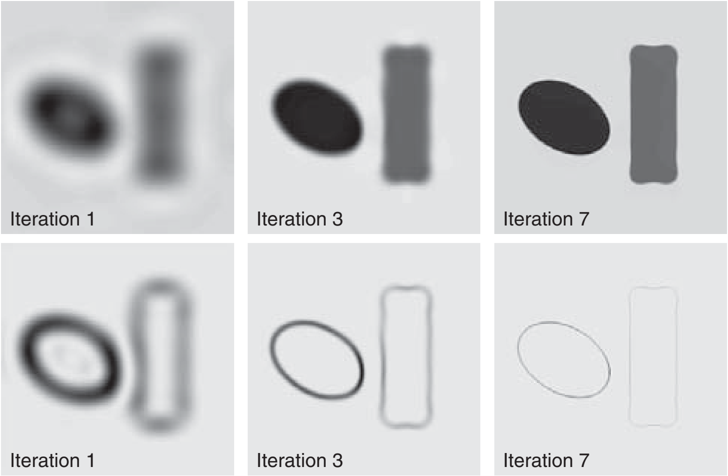


Figure 5. Approximation of the MAP estimate of the image (top row) and of the variance (bottom row) after 1, 3 and 5 iteration of the cyclic algorithm when using the CGLS method to compute the updated of the image at each iteration step

Comparing hyperpriors

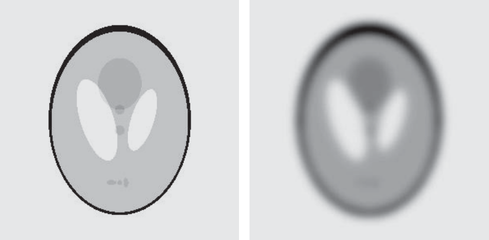


Figure 7. Original, blurred and noisy Shepp Logan phantom.

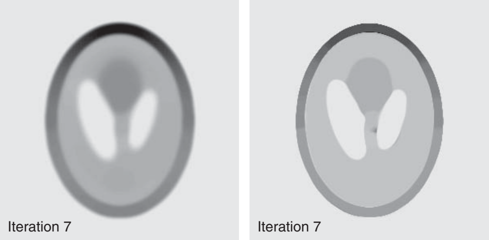


Figure 8. Approximation of the MAP estimate of the Shepp Logan phantom using a gamma (left) and an inverse gamma (right) hyperprior.

Uncertain boundaries

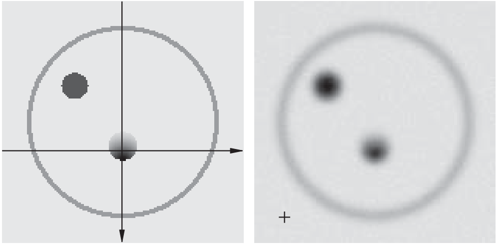


Figure 9. The true image of example 3 (left) and the blurred noisy version of it (right). The image profiles are computed along the horizontal and vertical line segments marked on the true image.

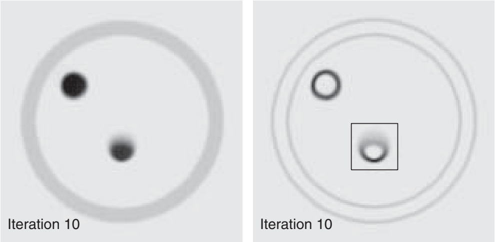


Figure 10. MAP estimates of the image and the variance using gamma hyperprior and CGLS iterative solver. The ROI is marked in the variance image.