Inverse Problems Exercise 7
Sommersemester 2022
Return your written solutions either in person or by email
to vesa.kaarnioja@fu-berlin.de by Monday 20 June, 2022, 12:15

1. Suppose that our quantity of interest is a function $f : [0,1] \to \mathbb{R}$, $f(0) = 0$, and we have *a priori* knowledge that the function f may have a large jump at few places in the interval, but the locations of these places are unknown to us. One way to construct a prior in this case is to consider the finite difference approximation of the derivative of f and assume that it follows an impulse noise probability distribution.

   Let $t_j = j/d$, $j \in \{0, \ldots, d\}$, be a discretization of the computational domain and let $x_j = f(t_j)$. Let us consider the problem of drawing a sample from the d-dimensional prior density

   $$\pi(x) = \left(\frac{\alpha}{\pi}\right)^d \prod_{j=1}^{d} \frac{1}{1 + \alpha^2 (x_j - x_{j-1})^2}, \quad \alpha > 0.$$

   Define $u = (u_1, \ldots, u_d)^{\mathrm{T}}$ by

   $$u_j = x_j - x_{j-1}, \quad j = 1, \ldots, d. \tag{1}$$

   The new random variable u is distributed according to the d-dimensional Cauchy distribution (*without positivity constraint!*)

   $$\pi(u) = \left(\frac{\alpha}{\pi}\right)^d \prod_{j=1}^{d} \frac{1}{1 + \alpha^2 u_j^2}. \tag{2}$$

   The components u_j are independent of each other and can be drawn from a univariate Cauchy distribution using inverse transform sampling. Note that the inverse cumulative distribution function corresponding to a univariate Cauchy density (*without positivity constraint!*) is given by

   $$\Phi^{-1}(t) = \frac{1}{\alpha} \tan(\pi(t - \tfrac{1}{2})), \quad t \in (0,1).$$

   Implement the following algorithm in MATLAB:

   - Fix `d = 1200`, set α `= 1`, and define `t = (1:d)'/d`;
   - Use componentwise inverse transform sampling to draw a realization from the d-dimensional Cauchy distribution (2).
   - Note that (1) implies that we can transform the realization back into the original coordinates x via

     $$x_1 = u_1,\ x_2 = u_1 + u_2,\ x_3 = u_1 + u_2 + u_3, \ldots,\ x_d = u_1 + u_2 + u_3 + \cdots + u_d.$$

     In MATLAB, this can be achieved with the command `x = cumsum(u);`
   - Finally, visualize the sample you obtained using the `plot` command. Plot the realization as a function over the original spatial mesh `t`.

2. Let $y \in \mathbb{R}^2$ and $x \in \mathbb{R}$ and

$$y = \begin{pmatrix} 2 \\ 1 \end{pmatrix} x + \eta, \quad \eta \sim \mathcal{N}(0, \gamma^2 I_2),$$

where $I_2 \in \mathbb{R}^{2\times 2}$ is an identity matrix. Suppose the prior distribution is given by $x \sim \mathcal{N}(0, 2)$. What is the posterior distribution if we observe $\hat{y} = (1\ 2)^\top$? What is the posterior variance? What happens to posterior distribution and variance under decreasing noise ($\gamma \downarrow 0$)?

3. Let $A \in \mathbb{R}^{k\times d}$, $x \in \mathbb{R}^d$, $y, \eta \in \mathbb{R}^k$, and consider the linear measurement model with additive noise:

$$y = Ax + \eta.$$

During the lecture, we proved that if x is endowed with a Gaussian prior distribution $\mathcal{N}(x_0, \Gamma_{\text{pr}})$, the noise η is assumed to have the Gaussian distribution $\mathcal{N}(\eta_0, \Gamma_{\text{n}})$, and x and η are mutually independent, then the posterior distribution is Gaussian with posterior covariance

$$\Gamma_{\text{post}} = (\Gamma_{\text{pr}}^{-1} + A^{\text{T}}\Gamma_{\text{n}}^{-1}A)^{-1} \tag{3}$$

and posterior mean

$$\mu_{\text{post}} = \Gamma_{\text{post}}(A^{\text{T}}\Gamma_{\text{n}}^{-1}(y - \eta_0) + \Gamma_{\text{pr}}^{-1}x_0). \tag{4}$$

Prove that

$$\Gamma_{\text{post}} = \Gamma_{\text{pr}} - \Gamma_{\text{pr}}A^{\text{T}}(A\Gamma_{\text{pr}}A^{\text{T}} + \Gamma_{\text{n}})^{-1}A\Gamma_{\text{pr}} \tag{5}$$

and

$$\mu_{\text{post}} = x_0 + \Gamma_{\text{pr}}A^{\text{T}}(A\Gamma_{\text{pr}}A^{\text{T}} + \Gamma_{\text{n}})^{-1}(y - Ax_0 - \eta_0). \tag{6}$$

Hint: Use the Sherman–Morrison–Woodbury formula: for *any* conformable matrices A, B, C, and D such that A and C are invertible (square) matrices, it holds that

$$(A + BCD)^{-1} = A^{-1} - A^{-1}B(C^{-1} + DA^{-1}B)^{-1}DA^{-1},$$

if $A + BCD$ is invertible (or, equivalently, if $C^{-1} + DA^{-1}B$ is invertible).

Begin by applying the Sherman–Morrison–Woodbury formula on (3); this should yield the formula (5). The formula (6) can then be proved by plugging the formula (5) into (4) and simplifying the resulting expression.