Department of Applied Mathematics and Theoretical Physics
Hanne Kekkonen & Yury Korolev

Inverse Problems

Example sheet **4** Presentation **20 Jan. 2020**, **2-3:30pm**, **MR15**.

Exercise 1 (Gaussian model)
Assume that we observe measurement $F = AU + N$, where $A : \mathbb{R}^d \to \mathbb{R}^k$ is a known matrix, $N \sim \mathcal{N}(0, \Sigma_n)$ and $U \sim \pi = \mathcal{N}(\theta_u, \Sigma_u)$, where Σ_n and Σ_u are both invertible.
i) Show that the posterior covariance Σ and mean $\overline{u}$ can be written as

$$\Sigma = (A^\top \Sigma_n^{-1} A + \Sigma_u^{-1})^{-1}$$

and

$$\overline{u} = \Sigma(A^\top \Sigma_n^{-1} f + \Sigma_u^{-1} \theta_u).$$

ii) What happens on the small noise limit $\delta \to 0$, $\Sigma_\eta = \delta^2 \Sigma_0$, if we assume that $k = n$ and A invertible?
iii) What happens if we only assume that $\mathrm{null}(A) = \{0\}$?

Exercise 2 (Underdetermined Gaussian model*)
Assume the same model as in 1. but this time $F \in \mathbb{R}^k$ and $U \in \mathbb{R}^d$ with $k < d$, and $\mathrm{rank}(A) = k$. We can then write

$$A = (A_0 \ \ 0) Q^\top,$$

with $Q \in \mathbb{R}^{d\times d}$ being an orthonormal matrix, $Q^\top Q = I$, and $A_0 \in \mathbb{R}^{k\times k}$ an invertible matrix. We denote $L_u = \Sigma_u^{-1}$ and write

$$Q^\top L_u Q = \begin{bmatrix} L_{11} & L_{12} \\ L_{12}^\top & L_{22} \end{bmatrix}, \quad L_{11} \in \mathbb{R}^{k\times k},\ L_{22} \in \mathbb{R}^{(d-k)\times(d-k)}.$$

We also write $Q = (Q_1 \ Q_2)$ with $Q_1 \in \mathbb{R}^{d\times k}$ and $Q_1 \in \mathbb{R}^{d\times(d-k)}$. Define $z \in \mathbb{R}^k$ to be the unique solution of $A_0 z = f$. Let $w \in \mathbb{R}^k$ and $w' \in \mathbb{R}^{d-k}$ be defined via $\Sigma_u^{-1}\theta_u = Q(w \ w')^\top$. Show that on the small noise limit $\delta \to 0$, $\Sigma_n = \delta^2 \Sigma_0$,

$$\Pi^f \rightharpoonup \mathcal{N}(\overline{\theta}_f, \overline{\Sigma}_f)$$

where

$$\overline{\theta}_f = Q(z \ z')^\top \quad \text{and} \quad \overline{\Sigma}_f = Q_2 L_{22}^{-1} Q_2^\top.$$

Above $z' = -L_{22}^{-1} L_{12}^\top z + L_{22}^{-1} w' \in \mathbb{R}^{d-k}$.

Exercise 3 (Estimators)
Let $U \in \mathbb{R}$ and assume that the posterior distribution is given by

$$\pi^f(u) = \frac{c}{\sigma_1}\varphi\Big(\frac{u}{\sigma_1}\Big) + \frac{1-c}{\sigma_2}\varphi\Big(\frac{u-1}{\sigma_2}\Big),$$

where $0 < c < 1$, $\sigma_1, \sigma_2 > 0$ and φ is density function of standard normal distribution $\varphi(u) = (2\pi)^{-1/2}\exp(-u^2/2)$. Calculate the conditional mean (CM) and maximum a posterior (MAP) estimates, and the posterior variance. Does MAP or CM always give a better estimator for u?

Exercise 4 (Sampling)
Let V be a real valued random variable with probability density $\pi(v)$, such that $\pi(v) = 0$ only at isolated points. We define the cumulative distribution function

$$\Phi(t) = \int_{-\infty}^{t} \pi(v)dv.$$

Define a new random variable $T = \Phi(V)$. Show that $T \sim \mathcal{U}([0,1])$.

Exercise 5 (Hyperpriors)
Assume that we observe a measurement $F{+}AU{+}N$ and the null space of A is zero. We model $U \sim \mathcal{U}$, where $\mathcal{U}$ is uninformative and improper prior with constant density on $\mathbb{R}^d$, that is, $\pi(u) = c > 0$. Furthermore, assume that $N \,|\, \delta \sim \mathcal{N}(0, \delta^2 I)$, where $\delta > 0$ is unknown. The noise amplitude is modelled by assuming $1/\delta^2 = \gamma \sim \Gamma(\alpha, \beta)$, where $\alpha, \beta > 0$, and $\Gamma(\alpha, \beta)$ is the Gamma distribution, with the density

$$\pi_h(\gamma) \propto \gamma^{\alpha-1}\exp(-\beta\gamma).$$

Write down the posterior distribution $\pi^f(u, \gamma)$ and the densities of $u \,|\, \gamma, f$ and $\gamma \,|\, u, f$. Give the MAP estimators for u and γ.

Exercise 6 (Towards continuous models 1)
The Lebesgue measure ν_n on Euclidean space $\mathbb{R}^n$ is countably additive and translation invariant. Show that there is no analogue of Lebesgue measure on infinite-dimensional Banach space X.

Exercise 7 (Towards continuous models 2)
Let $U \sim \mathcal{N}(0, I)$, where $0 \in \mathbb{R}^d$ and $I \in \mathbb{R}^{d\times d}$ is identity matrix. Where does most of the probability mass lie when d is large? Hint: What happens to the volume of d-dimensional unit ball $B_d(0,1)$ when $d \to \infty$?

Exercise 8 (Hellinger distance 1)
Let μ and μ' be two probability measures on a separable Banach space X. Let $(Y, \|\cdot\|)$ be a separable Banach space and assume that $g : X \to Y$ is measurable and has second moments with respect to both μ and μ'. Show that

$$\|\mathbb{E}^{\mu}(g) - \mathbb{E}^{\mu'}(g)\| \le 2\Big(\mathbb{E}^{\mu}\|g\|^2 + \mathbb{E}^{\mu'}\|g\|^2\Big)^{\frac{1}{2}} d_{Hell}(\mu, \mu'),$$

Exercise 9 (Hellinger distance 2)
Assume that the measures μ' and μ are equivalent, that is, $\mu' \ll \mu$ and $\mu \ll \mu'$. The Kullback–Leibler divergence between μ' and μ is defined as

$$D_{KL}(\mu'||\mu) = \int \log\Big(\frac{d\mu'}{d\mu}\Big) d\mu'.$$

Is D_{KL} a metric? Assume that the measures μ' and μ are equivalent. Show that

$$d_{Hell}(\mu, \mu')^2 \leq \frac{1}{2} D_{KL}(\mu||\mu').$$

Exercise 10 (Matlab exercises)
1. Sample from ℓ^1, Cauchy and Gaussian priors using Matlab. Plot the samples as a 2D image.

2. Assume that we have a posterior distribution with density

$$\pi(x, y) = \exp\big(- 10(x^2 - y)^2 - (y - 1/4)^4\big).$$

Write a Metropolis–Hastings algorithm to sample from π using the pseudocode given in Example 6.5.3. Try your code with different choices of γ and plot the first co-ordinates of the samples. What do you notice? What percentage of the suggested moves is accepted?