8. Exercise sheet

Due by: Friday, 14 June 2024, 11:59 pm (CEST)
Please refer to **Assignment Submission Guideline** on Moodle
Implementational details about the Metropolis–Hastings algorithm and Gibbs sampler can be found in the note "Markov chain Monte Carlo" on the course Moodle page

Problem 1. Let $\pi(x_1, x_2)$ be the probability density function corresponding to the bivariate Gaussian random variable

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \sim \mathrm{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & p \\ p & 1 \end{pmatrix} \right),$$

where $p \in [0, 1)$ is a parameter.

(a) Implement a Gibbs sampler to draw a sample of size $N = 10\,000$ from the distribution $\pi(x_1, x_2)$. Plot the sample you obtained and compare visually against a contour plot of the probability density π. Using the generated sample, estimate the expected value of π and the marginal standard deviations of x_1 and x_2.

(b) Implement part (a) for parameter values $p = 0.5, 0.9, 0.99$, and 0.999. How does the degree of correlation between x_1 and x_2 affect the performance of the Gibbs sampler?

Problem 2. Let us consider the following measurement model:

$$y = (x_1^2 + x_2^2)^{1/2} + \eta, \quad \eta \sim \mathrm{N}(0, 0.1^2),$$

where $y \in \mathbb{R}$ is the measurement, $x = (x_1, x_2)^{\mathrm{T}} \in \mathbb{R}^2$ is the unknown parameter, and $\eta \in \mathbb{R}$ is observational noise. Let us endow the unknown parameter x with the prior

$$\pi(x_1, x_2) = C \exp\left(-\frac{1}{2}\big(|x_1 - 1| + (x_2 - 1)^2\big) \right), \quad (x_1, x_2) \in \mathbb{R}^2,$$

where the normalization constant C is chosen to satisfy $\int_{\mathbb{R}^2} \pi(x_1, x_2)\,\mathrm{d}x_1\,\mathrm{d}x_2 = 1$. Here, x and η are assumed to be independent.

(a) Write down the posterior density of $x|y$. You may ignore the constant of normalization.

(b) Suppose that we observe $y = 1$. Implement the random walk Metropolis–Hastings algorithm and draw a sample of size $N = 5000$ from the posterior distribution $\pi(x|y = 1)$ using this algorithm. Use the origin as the initial point. Try using step sizes $\gamma = 0.05$, 0.30, and 1.50. Plot the sample you obtained and compare visually against a contour plot of the (unnormalized) posterior density for each value of γ. Plot also the acceptance ratio of these samples. Which step size appears to work the best?

(c) Using the best step size you obtained in part (b), use the corresponding sample average to find an approximation for the conditional mean (CM) estimate

$$x_{\mathrm{CM}} = \mathbb{E}[x|y] = \int_{\mathbb{R}^2} x\,\pi(x|y=1)\,\mathrm{d}x.$$

Problem 3. Consider a discrete Markov chain with three possible states $\mathcal{X} = \{1,2,3\}$ and the transition matrix $\mathrm{P} = [P_{i,j}]_{i,j=1}^3$ defined by

$$\mathrm{P} = \begin{pmatrix} 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \end{pmatrix}.$$

(a) What is the invariant probability distribution for this Markov chain?

(b) Consider the probability mass function (PMF) $p^*(1) = 4/12$, $p^*(2) = 3/12$, and $p^*(3) = 5/12$ for the states. Your task is to implement an MCMC sampler to draw a sample from the distribution with PMF p^*. Following Definition 5.15 in the textbook, you can proceed as follows:

Choose an (arbitrary) initial state $x^{(0)} \in \mathcal{X}$.

For $j = 0, 1, 2, \ldots$, do

Draw a new proposal $y \in \mathcal{X}$ from the distribution with the PMF $p(1) = P_{1,x^{(j)}}$, $p(2) = P_{2,x^{(j)}}$, and $p(3) = P_{3,x^{(j)}}$.

Compute the acceptance probability $\alpha = \min\{1, p^*(y)/p^*(x^{(j)})\}$.

Draw t from the uniform distribution $\mathrm{U}[0,1]$.

If $\alpha > t$, do

Set $x^{(j+1)} = y$.

Else

Set $x^{(j+1)} = x^{(j)}$.

End if

End for

Use this algorithm to draw a sample of size $N = 10\,000$. Verify the correctness of your implementation by computing the relative frequencies of the occurrence of each state $i \in \{1,2,3\}$, in the sample that you obtained.