Mathematics 2
Homework for Part 1

The solutions are to be submitted as one .zip file to the appropriate mailbox on ucilnica. The solutions should contain a .pdf file containing a clear and well described procedure and a code. The homework is worth 3 points. The written exam is worth additional 7 points.

1 Importance sampling

We would like to compute the integral

$$I = \int_1^\infty x^{-7/4} \cdot e^{-1/x} \, dx.$$

1. Plot the integrand $f(x) = x^{-7/4} \cdot e^{-1/x}$ on some large interval $[1, N]$, $N \in \mathbb{N}$.
2. Estimate the integral I by sampling from the uniform distribution on $[1, N]$ for some large N using the Monte Carlo integration with samples of size $n = 10^7$. Compute the average and the standard deviation for 10 samples and comment what is a good choice of N.
3. Estimate I as $E_q(\frac{f}{q})$ for some appropriately chosen $q(x)$ by sampling form q using the inversion sampling and repeating step 2 above for this choice of q.
4. Compare and comment the results obtained by both methods.

2 Markov chains: countable state space

Let $X_0, X_1, X_2, \ldots$ be a Markov chain with a countable state space $S = \{w_i : i \in \mathbb{N}\}$. Let $p_{ij} = P(X_1 = w_j \mid X_0 = w_i)$. and

$$f_{ij}(n) = P(X_n = w_j, X_n - 1 \neq w_j, \ldots, X_1 \neq w_j \mid X_0 = w_i)$$

A state w_i is called **recurrent** if $\sum_{n=1}^{\infty} f_{ii}(n) = 1$ and **transcient** otherwise. Time $T_{ij} = \min\{n : X_n = w_j \mid X_0 = w_i\}$ is called a first-passage time from w_i to w_j.

1. For which states is T_{ii} a random variable and what is its pmf it this case?

States w_i and w_j **communicate**, i.e., $w_i \sim w_j$, if there exists $n \in \mathbb{N}_0$ and $m \in \mathbb{N}_0$ such that $P(X_n = w_j : X_0 = w_i) > 0$ and $P(X_m = w_i : X_0 = w_j) > 0$. A Markov chain is **irreducible**, if all pairs of states communicate.

2. Show that the relation $\sim$ is an equivalence relation and that all in each class all states are recurrent or all are transcient.

3. If w_i and w_j are recurrent states in the same class, show that $\sum_{n=1}^{\infty} f_{ij}(n) = 1$ and T_{ij} is a random variable.

A state w_i is called **positive-recurrent** if it is recurrent and $E[T_{ii}] < \infty$. A recurrent state, which is not positive-recurrent is called **null-recurrent**.

A **stationary distribution** of a Markov chain is a sequence $\{\pi_i : i \in \mathbb{N}\}$, such that

$$\pi_j = \sum_i \pi_i p_{ij} \quad \text{for } j \in \mathbb{N}, \quad \pi_j \geq 0 \quad \text{for } j \in \mathbb{N} \quad \text{and} \quad \sum_{j=1}^{\infty} \pi_j = 1.$$

Let $X_0 = w_i$ and let N_j be the expected number of visits of the state w_j before returning to the state w_i.

4. Show that if all states in an irreducible Markov chain are positive-recurrent, then a stationary distribution π exists and is equal to $\pi_j = \frac{N_{ij}}{E[T_{ii}]}$. In the computation use a simple observation that

$$N_{ij} = \sum_{t=0}^{\infty} P(X_t = w_j, T_{ii} > t \mid X_0 = w_i). \tag{1}$$

5. Argue that π_i from 4. is $1/E[T_{ii}]$. By showing that a stationary distribution of an irreducible, positive recurrent Markov chain is unique, argue that its i-th component is equal to $1/E[T_{ii}]$. To prove uniqueness follow the following procedure. Without loss of generality we can assume $i = 1$. Let $v = \{v_i : i \in \mathbb{N}\}$ be a stationary distrubution with $v_1 = \frac{1}{E[T_{11}]}$. We have to prove that $v_j = \frac{N_{1j}}{E[T_{11}]}$.

   (a) Let $j > 1$. From the equality

$$v_j = \sum_{i_1=1}^{\infty} v_{i_1} p_{i_1 j} = v_1 p_{1j} + \sum_{i_1=2}^{\infty} v_{i_1} p_{i_1 j} = \frac{p_{1j}}{E[T_{11}]} + \sum_{i_1=2}^{\infty} v_{i_1} p_{i_1 j} \tag{2}$$

argue that

$$v_j \geq \frac{1}{E[T_{11}]} \cdot P(X_1 = w_j, T_{11} > 1 \mid X_0 = w_1).$$

(b) Write down the formula (2) for $j = i_1$, i.e., $v_{i_1} = \cdots$. Plug this expression into (2) to obtain

$$v_j = \frac{p_{1j}}{E[T_{11}]} + \frac{\sum_{i_1=2}^{\infty} p_{1i_1} \cdot p_{i_1 j}}{E[T_{11}]} + \sum_{i_2=2}^{\infty} \sum_{i_1=2}^{\infty} v_{i_2} p_{i_2 i_1} p_{i_1 j}. \tag{3}$$

Argue that

$$v_j \geq \frac{1}{E[T_{11}]} \cdot \sum_{\ell=1}^{2} P(X_\ell = w_j, T_{11} > \ell \mid X_0 = w_1).$$

(c) Using induction on k show that

$$\begin{aligned} v_j = {} & \frac{p_{1j}}{E[T_{11}]} + \frac{\sum_{i_1=2}^{\infty} p_{1i_1} \cdot p_{i_1 j}}{E[T_{11}]} + \dots \\ & + \frac{\sum_{i_k=2}^{\infty} \sum_{i_{k-1}=2}^{\infty} \cdots \sum_{i_1=2}^{\infty} p_{1i_k} \cdot p_{i_k i_{k-1}} \cdots p_{i_2 i_1} \cdot p_{i_1 j}}{E[T_{11}]} + \\ & + \sum_{i_{k+1}=2}^{\infty} \sum_{i_k=2}^{\infty} \cdots \sum_{i_1=2}^{\infty} v_{i_{k+1}} p_{i_{k+1} i_k} \cdots p_{i_2 i_1} p_{i_1 j} \end{aligned}$$

and hence

$$v_j \geq \frac{1}{E[T_{11}]} \cdot \sum_{\ell=1}^{k+1} P(X_\ell = w_j, T_{11} > \ell \mid X_0 = w_1). \tag{4}$$

(d) Argue from (4) that $v_j \geq \frac{N_{1j}}{E[T_{11}]}$ (observe (1)) and hence $v_j \geq \pi_j$.

(e) Observe that $\{u_i \colon i \in \mathbb{N}\}$, $u_i = v_i - \pi_i$ for $i \in \mathbb{N}$, satisfies

$$u_j = \sum_i u_i p_{ij} \quad \text{for } j \in \mathbb{N}, \quad u_i \geq 0 \ \text{ for } i \in \mathbb{N} \quad \text{and} \quad u_1 = 0. \tag{5}$$

Using (5) and the irreducibility of the Markov chain argue that $u_j = 0$ for all $j \in \mathbb{N}$ and hence $v_j = \pi_j$ for all $j \in \mathbb{N}$.

3 Metropolis-Hastings algorithm

Let X be a random variable following the (α, η)-Weibull distribution with the parameters $\alpha, \eta \in (0, \infty)$, support $x \in (0, \infty)$, and the pdf

$$p(x \mid \alpha, \eta) = \alpha\eta x^{\alpha-1} e^{-x^{\alpha}\eta}.$$

Let the prior distribution $\pi(\alpha, \eta)$ be proportional to $e^{-\alpha-2\eta} \cdot \eta$. We observe the data $x = 0.5$, $x = 0.4$, $x = 0.6$, $x = 0.8$, $x = 0.3$. Approximate the posterior distribution for α and η and compute the mean and the variance of $\alpha - s$ and $\eta - s$ using Metropolis-Hastings algorithm with:

1. a multivariate normal proposal. Use mean 0, while tune the covariance matrix by hand via trial.

2. a proposal distribution

$$q(\alpha', \eta' \mid \alpha, \eta) = \frac{1}{\alpha\eta} \exp(-\frac{\alpha'}{\alpha} - \frac{\eta'}{\eta}).$$

For each of the above scenarios:

1. Generate 5 independent chains of 1000 samples.

2. Apply standard MCMC diagnostics for each algorithm/run (traceplot for each parameter and all chains at the same time), autocovariance and the ESS.

3. Compare both algorithms in sampling from the target distribution. Include a discussion of how difficult/easy it was to tune MCMC parameters.

4. Estimate the probability $(\alpha, \eta) \in [2, \infty) \times [2, \infty)$.