Homework 4

! For each problem, please clearly show your reasoning and write all the steps.
G As data scientists, you should feel free to google it whenever you see something unfamiliar.
☺ Group discussion for the homework is encouraged, but you have to write your answer by yourself. Also, you are always welcome to discuss the problems with me.

Task 0.

Read the relevant chapters in the textbooks listed on Sakai.

- [Al] stands for *Introduction to Machine Learning* by Alpaydin;
- [Bi] stands for *Patern Recognition and Machine Learning* by Bishop;
- [HaTF] stands for *The Elements of Statistical Learning* by Hastie, Tibshirani and Friedman.

Problem 1. Programming: MCMC (10pt)

Suppose we want to evaluate

$$s = \mathbb{E}_p f = \int f(x_1, x_2) p(x_1, x_2) dx_1 dx_2 \ ,$$

where

$$f(x_1, x_2) = \frac{1}{1 + \exp(-(x_1 + x_2))}$$

and p is the Rosenbrock density

$$p(x_1, x_2) \propto \exp\left(-\frac{(1 - x_1)^2 + 100(x_2 - x_1^2)^2}{20}\right) \ .$$

1. Use Metropolis-Hastings methods, sample 10000 points following $p(x_1, x_2)$ and plot them. You can use your favourite proposal kernel, or just employ the Gaussian proposal used in the lecture.
2. Approximate s accordingly.

Problem 2. Rejection sampling (10pt) [Bi] Ex.11.6

In this exercise, we show that rejection sampling indeed draws samples from the desired distribution. Suppose the proposed distribution is $q(z)$. Also suppose $p = \tilde{p}(z)/Z$ for some fixed but unknown Z, and k is the smallest value for which $kq(z) \geq \tilde{p}(z)$ for all z.

1. Show that the probability of a sample value z being accepted is given by $\tilde{p}(z)/kq(z)$.

2. Explain in plain language why we are sampling from an equivalent density

$$\frac{q(z)p(\text{acceptance}|z)}{p(\text{acceptance})}\ .$$

3. Prove that $p(\text{acceptance}) = Z/k$.

4. According to Parts 1-3, prove that by doing rejection sampling, we are sampling from $p(z)$.

Problem 3. Entropy calculation (10pt) [Bi] Ex.1.39

Consider two binary variables x and y taking values in $\{0, 1\}$, such that $\mathbb{P}\Big((x, y) = (0, 0)\Big) = \mathbb{P}\Big((x, y) = (0, 1)\Big) = \mathbb{P}\Big((x, y) = (1, 1)\Big) = \frac{1}{3}$ and $\mathbb{P}\Big((x, y) = (1, 0)\Big) = 0$. Evalutate the following:

1. $H[x]$
2. $H[y]$
3. $H[y|x]$
4. $H[x|y]$
5. $H[x, y]$
6. $I[x, y]$

Draw a diagram to show the relationship between these various quantities.

Problem 4. Entropy (10pt + 5pt Bonus)

1. (a) (Bonus) Let $\mathbf{x}$ be a D-dimensional Gaussian random variable with $p(\mathbf{x}) = \mathcal{N}(\mathbf{x}|0, \mathbf{\Sigma})$, where $\mathbf{\Sigma}$ is a $D \times D$ covariance matrix. Prove:

$$H[\mathbf{x}] = \frac{1}{2}\ln\left[(2\pi e)^D\,|\mathbf{\Sigma}|\right]\ .$$

   (b) In light of Part (a), using the fact that $H[\mathrm{x}_1, \cdots, \mathrm{x}_D] \le \sum_{i=1}^{D} H[\mathrm{x}_i]$, prove:

$$|\mathbf{\Sigma}| \le \prod_{i=1}^{D} \Sigma_{ii}\ .$$

2. By a change of variable, prove:

$$H[a\mathbf{x}] = H[\mathbf{x}] + \ln|a|\ .$$