Exercises for Probabilistic Graphical Models
Sheet No.6

Bernt Schiele, Bjoern Andres, Eldar Insafutdinov, Evgeny Levinkov

Due Date: Monday, 30th January

Submit during the lecture or to Eldar via eldar@mpi-inf.mpg.de or in the office (629).

Begin the title of email with [PGM]. Only send the code that you have written, no data files.

1 Rejection Sampling

Points: 5

When a sampling by direct transformation is not easy to find, a powerful indirect method known as *rejection sampling* (or the Accept/Reject Algorithm) can often provide a solution.

Let $\mathcal{N}_+(\mu, \mu^-, \sigma^2)$ denote the (one-sided) truncated normal distribution with left truncation point μ^-. The corresponding pdf is proportional to

$$p^*(x) = \begin{cases} \exp(-(x-\mu)^2/2\sigma^2) & x \geq \mu^-, \\ 0 & \text{otherwise.} \end{cases}$$

The task is to sample $X \sim \mathcal{N}_+(\mu, \mu^-, \sigma^2)$.

- A straightforward approach would be to sample $X \sim \mathcal{N}(\mu, \sigma^2)$ from the standard normal distribution and only keep samples $x_i \geq \mu^-$. Discuss when this method could work and when it would be too inefficient.

- Now consider rejection sampling. For the proposal distribution, use translated exponential distribution with the following pdf

$$q(x) = \begin{cases} \lambda \exp(-\lambda(x-\mu^-)) & x \geq \mu^-, \\ 0 & \text{otherwise.} \end{cases}$$

  Find M as a function of $(\lambda, \mu, \mu^-, \sigma)$ such that $Mq(x) \geq p^*(x)$ for all x.

- Let $U \sim \text{Uniform}(0, 1)$. Find g such that $X = g(U)$ has pdf $q(x)$.

- Let $\lambda = 1$, $\mu = 0$, $\mu^- = -1$, $\sigma = 1$. Perform computer simulation using both methods discussed above (see Algorithm 27.2 in Barber for Rejection sampling). Generate $n = 10000$ candidates and report the acceptance ratio of both methods. Plot histograms of samples from both methods.
- Same as above, but use $\mu^- = 1$. Discuss the difference (if any).

2 Gibbs Sampling

Points: 7
Consider a (unnormalized) two-dimensional Gaussian distribution

$$p^*(x) = \exp\left(-\frac{1}{2}(x-\mu)^\top \Sigma^{-1}(x-\mu)\right),$$

where $x = (x_1, x_2)^\top$, $\mu = (\mu_1, \mu_2)^\top$, and the covariance matrix is given by

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}.$$

- Derive the conditional distribution $p(x_1 \mid x_2)$. How are $X_1 \mid X_2$ and $X_2 \mid X_1$ distributed?
- Implement a Gibbs sampler.
- Let $\sigma_1 = \sigma_2 = 1$ and let $\mu_1 = \mu_2 = 0$. Plot the trajectory of 10,000 samples generated by the sampler for each $\rho = 0$, 0.5, 0.8. Briefly explain in which case the sampler is more effective and why.
- Let $\rho = 0.5$. Estimate the value of

  $$\mathrm{P}(1 \le X_2 \le 2 \mid 1 \le X_1 \le 2).$$

  Perform 100 experiments sampling 10,000 points each time and use the obtained samples to approximate the above probability via the Monte-Carlo estimate:

  $$\frac{1}{m}\sum_{i=1}^{m} f(x_i) \approx \mathbb{E}\, f(X)$$

  (you have to choose the right f and x_i). Evaluate the estimator performance in terms of (empirical) bias and variance.

3 Metropolis-Hastings

3.1 Implementation

Points: 4
Consider a mixture of two one-dimensional Gaussians

$$p(x) = w_1\,\mathcal{N}(x \mid \mu_1, \sigma_1^2) + w_2\,\mathcal{N}(x \mid \mu_2, \sigma_2^2)$$

and let the proposal distribution be a Gaussian as well

$$\tilde{q}(x') = \mathcal{N}(x' \,|\, x, \sigma_p^2).$$

- Implement a Metropolis-Hastings sampler.
- Let the parameters of the target distribution be given as follows:

  $$w_1 = 0.3, \quad w_2 = 0.7, \quad \mu_1 = 0, \quad \mu_2 = 10, \quad \sigma_1 = 2, \quad \sigma_2 = 2.$$

  Sample $N = 100,\ 1000,\ 10000$ points for each $\sigma_p = 0.1,\ 1,\ 10,\ 100$. Plot histograms of samples in each experiment and discuss the differences.

3.2 Metropolis-Hastings in High Dimensions

Points: 4

- Let $X = (X_1, \ldots, X_n)^\top$, where X_i is a random variable with mean μ_i and variance σ_i^2. Compute the expected value of $X^\top X$.
- Exercise 27.5. in Barber.