Tutorial Week 6

COMP9418 – Advanced Topics in Statistical Machine Learning, 17s2, UNSW Sydney

Instructor: Edwin V. Bonilla

Last Update: Friday 25th August, 2017 at 10:20

1. (PRML, Bishop, 2006) Let $\mathbf{z}$ be a D-dimensional random variable having a Gaussian distribution with zero mean and unit covariance matrix, and suppose that the positive definite symmetric matrix $\mathbf{\Sigma}$ has the Cholesky decomposition $\mathbf{\Sigma} = \mathbf{L}\mathbf{L}^T$, where $\mathbf{L}$ is a lower-triangular matrix (i.e., one with zeros above the leading diagonal). Show that the variable $\mathbf{y} = \boldsymbol{\mu} + \mathbf{L}\mathbf{z}$ has a Gaussian distribution with mean $\boldsymbol{\mu}$ and covariance $\mathbf{\Sigma}$. Explain how this provides a technique for generating samples from a general multivariate Gaussian using samples from a univariate Gaussian having zero mean and unit variance.

2. (PRML, Bishop, 2006) In this exercise, we show more carefully that rejection sampling does indeed draw samples from the desired distribution $p(\mathbf{z})$. Suppose the proposal distribution is $q(\mathbf{z})$ and show that the probability of a sample value $\mathbf{z}$ being accepted is given by $\tilde{p}(\mathbf{z})/kq(\mathbf{z})$ where $\tilde{p}$ is any unnormalized distribution that is proportional to $p(\mathbf{z})$, and the constant k is set to the smallest value that ensures $kq(\mathbf{z}) \geq \tilde{p}(\mathbf{z})$ for all values of $\mathbf{z}$. Note that the probability of drawing a value $\mathbf{z}$ is given by the probability of drawing that value from $q(\mathbf{z})$ times the probability of accepting that value given that it has been drawn. Make use of this, along with the sum and product rules of probability, to write down the normalized form for the distribution over $\mathbf{z}$, and show that it equals $p(\mathbf{z})$.

3. (PRML, Bishop, 2006) Show that the Gibbs sampling algorithm, as described in the lecture, satisfies detailed balance.

4. (PRML, Bishop, 2006) Consider the simple graph shown in Figure 1 in which the observed node x is given by a Gaussian distribution $\mathcal{N}(x|\mu, \tau^{-1})$, with mean μ and precision τ. Suppose that the prior distributions over the mean and the precision are given by $\mathcal{N}(\mu|\mu_0, s_0)$ and $\mathrm{Gamma}(\tau|a, b)$, where $\mathrm{Gamma}(\cdot|\cdot, \cdot)$ denotes a gamma distribution. Write down expressions for the conditional distributions $p(\mu|x, \tau)$ and $p(\tau|x, \mu)$ that would be required in order to apply Gibbs sampling to the posterior distribution $p(\mu, \tau|x)$.

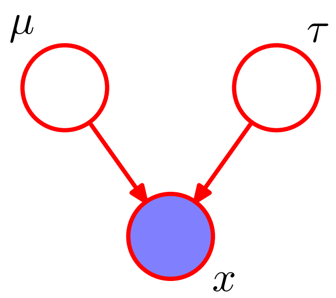

Figure 1: A graph involving an observed Gaussian variable x with prior distributions over its mean μ and precision τ.

5. (BRML, Barber, 2017) Consider the symmetric Gaussian proposal distribution:

$$q(\mathbf{z}'|\mathbf{z}) = \mathcal{N}(\mathbf{z}'|\mathbf{z}, \sigma_q^2\mathbf{I}) \tag{1}$$

and the target distribution

$$p(\mathbf{z}) = \mathcal{N}(\mathbf{z}|\mathbf{0}, \sigma_p^2\mathbf{I}), \tag{2}$$

where $\mathbf{z} \in \mathbb{R}^D$. Show that:

$$\mathbb{E}_{q(\mathbf{z}'|\mathbf{z})}\left[\log \frac{p(\mathbf{z}')}{p(\mathbf{z})}\right] = -\frac{D\sigma_q^2}{2\sigma_p^2}. \tag{3}$$

Discuss how this results relates to the probability of accepting a Metropolis-Hastings update under a Gaussian proposal distribution in high dimensions.