CSE 515T (Spring 2019) Assignment 3

Due Wednesday, 10 April 2019

1. (Sampling from a multivariate Gaussian distribution). Assume you can sample from a univariate standard normal distribution $\mathcal{N}(0, 1^2)$. In the last assignment, this allowed you to sample trivially from a multivariate standard normal distribution $\mathcal{N}(\mathbf{0}, \mathbf{I})$. Use the closure of the Gaussian distribution to affine transformations to derive a procedure to sample from an arbitrary multivariate Gaussian distribution

   $$p(\mathbf{x}) = \mathcal{N}(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma})$$

   using an oracle that can draw a sample from the univariate standard normal distribution.

   Hint: the *Cholesky decomposition* may be helpful.

Now we will consider a series of questions relating to an application of Bayesian inference to numerical analysis, specifically quadrature.

We are going to consider the function

$$f(x) = \exp(-x^2)$$

and its definite integral

$$Z = \int_{-\infty}^{\infty} f(x)\,\mathrm{d}x.$$

The function f has no elementary antiderivative, so the calculation of Z is not straightforward. There is a famous method for computing Z with the trick of considering Z^2 instead, rewriting the resulting $2d$ integral in polar coordinates, and making a convenient substitution.[1] If you haven't seen this, it's beautiful and worth checking out. The result is

$$Z = \sqrt{\pi}.$$

We will consider modeling f with a Gaussian process prior distribution:

$$p(f) = \mathcal{GP}(f; \mu, K),$$

and conditioning on the following set of data $\mathcal{D} = (\mathbf{x}, \mathbf{y})$:

$$\begin{aligned}
\mathbf{x} &= [-2.5, -1.5, -0.5, 0.5, 1.5, 2.5]^\top; \\
\mathbf{y} &= \exp(-\mathbf{x}^2) \\
&= [0.0019305, 0.1054, 0.7788, 0.7788, 0.1054, 0.0019305]^\top.
\end{aligned}$$

We will fix the prior mean function μ to be identically zero; $\mu(x) = 0$.

2. First, let us consider the question of model, specifically kernel, selection.

   Consider the following three choices for the covariance function K:

$$\begin{aligned}
K_1(x, x') &= \exp\bigl(-|x - x'|^2\bigr) \\
K_2(x, x') &= \exp\bigl(-|x - x'|\bigr) \\
K_3(x, x') &= \bigl(1 + \sqrt{3}|x - x'|\bigr) \exp\bigl(-\sqrt{3}|x - x'|\bigr).
\end{aligned}$$

   Note that I am not parameterizing any of these kernels; please consider them to be fixed.

   Each kernel defines a Gaussian process model for the data in a natural way:

$$p(f \mid \mathcal{M}_i) = \mathcal{GP}(f; \mu, K_i).$$

   Consider a uniform prior distribution over these models:

$$\Pr(\mathcal{M}_i) = 1/3 \qquad i = 1, 2, 3.$$

   (a) Compute the log marginal likelihood for each model given the data $\mathcal{D}$ above.

[1] This is commonly credited to Gauss, but the idea goes back at least to Poisson.

(b) Compute the model posterior $\Pr(\mathcal{M} \mid \mathcal{D})$. Is there strong evidence for one model being the best?

3. Now we will consider integration.

   Perform Bayesian quadrature to estimate the definite integral $\int_{-5}^{5} f(x)\,\mathrm{d}x$, using the model $\mathcal{M}_1$ from question 1. What is the predictive mean and standard deviation, $p(Z \mid \mathcal{D}, \mathcal{M}_1)$? (You may give a numeric answer.) How does this compare with the true answer?