

CSE 515T (Spring 2018) Assignment 3

Due Monday, 19 March 2018

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.

2. (Sampling from a Gaussian process). Consider the squared exponential covariance function

$$K(x, x'; \lambda, \ell) = \lambda^2 \exp\left(-\frac{\|x - x'\|^2}{2\ell^2}\right). \quad (1)$$

Take a dense grid (somewhere around 1 000 points) in the domain $x \in [-4, 4]$. For $(\lambda, \ell) = (1, 1)$, draw five samples from a Gaussian process prior with mean function $\mu(x) = 0$ and this covariance. Plot your samples. Repeat with $(\lambda, \ell) = (2, 1)$ and $(\lambda, \ell) = (1, 2)$. What changed?

Note: you may encounter numerical problems here. The most common issues are that the sample covariance \mathbf{K} is not quite symmetric or that it is not quite positive definite. To ensure positive definiteness, you may need to add a small constant (on the order of 10^{-6} perhaps) to the diagonal to the covariance matrix \mathbf{K} in case of numerical instability. If the matrix is not exactly symmetric, a simple trick to symmetrize it is to take $\frac{1}{2}(\mathbf{K} + \mathbf{K}^\top)$.

Repeat this sampling procedure with the following covariance functions:

$$\begin{aligned} K(x, x') &= \exp(|x - x'|) \\ K(x, x') &= (1 + |x - x'|^2). \end{aligned}$$

3. (Gaussian process regression). Consider the following data:

$$\begin{aligned} \mathbf{x} &= [-2.26, -1.31, -0.43, 0.32, 0.34, 0.54, 0.86, 1.83, 2.77, 3.58]^\top; \\ \mathbf{y} &= [1.03, 0.70, -0.68, -1.36, -1.74, -1.01, 0.24, 1.55, 1.68, 1.53]^\top. \end{aligned}$$

Fix the observation noise variance at $\sigma^2 = 0.5^2$.

- Examining a scatter plot of the data, guess which values of (λ, ℓ) in the above covariance (1) (if any) might explain this data well.
- Perform Gaussian process regression for these data on the interval $x_* \in [-4, 4]$ using the squared exponential covariance (1) for the same set of hyperparameters (λ, ℓ) above. Plot the posterior mean and the pointwise 95% credible interval for each. Which predictions look the best?
- Visualize the model evidence $p(\mathbf{y} \mid \mathbf{x}, \lambda, \ell, \sigma^2)$ as a function of (λ, ℓ) . You can choose to make, for example, a heatmap or a contour plot of this function. Probably it will be a good idea to compute and plot the logarithm of the model evidence rather than the model evidence directly.