16831 Statistical Techniques, Fall 2014: Problem Set 3

Name:

Due: Tuesday, November 11, beginning of class

1 Bayesian Linear Regression Prediction

In the Bayesian linear regression problem, we assume that our data points $(x_1, y_1), \ldots, (x_n, y_n)$ are generated as:

$$y_t = \theta^T x_t + \epsilon_t, \quad \epsilon_t \sim \mathcal{N}(0, \sigma^2),$$

with a prior on parameter θ given by $\mathcal{N}(\mu_\theta, \Sigma_\theta)$.
Given an x_t, we would like to predict y_t. In class, we showed, using linearity of expectation, that the mean of the *predictive distribution* over y_t was given by:

$$\begin{aligned}
\mathbb{E}[y_t] &= \mathbb{E}[\theta^T x_t + \epsilon] \\
\mathbb{E}[y_t] &= \mathbb{E}[\theta^T x_t] + \mathbb{E}[\epsilon] \\
\mathbb{E}[y_t] &= \mathbb{E}[\theta]^T x_t + 0 \\
\mathbb{E}[y_t] &= \mu_\theta^T x_t
\end{aligned}$$

What is the variance of y_t?
Hint: For uncorrelated random variables X_i

$$Var(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} Var(X_i)$$

2 Conditional Independence in a Gaussian

In the Natural / Canonical Parameterization of a Gaussian, P's sparsity pattern is essentially encoding the graphical model structure of the Gaussian random variables.

2.1

Show that if $P_{ij} = 0$ then conditioned on everything else, x_i and x_j are independent.

2.2

Argue by example why $P_{ij} = 0$ does not imply independence even if it implies conditional independence.