Probabilistic Modelling and Reasoning, Tutorial Question Sheet 4 (for Week 6)

School of Informatics, University of Edinburgh

Instructor: Amos Storkey
Prob. Sheet. ID: 18204.1
Thanks: Chris Williams for some of the questions.

1. **Maximum likelihood estimation**. The Poisson distribution is a discrete distribution on the non-negative integers, with

$$P(x) = \frac{e^{-\lambda}\lambda^x}{x!} \qquad x = 0,\ 1,\ 2, \ldots$$

You are given a sample of n observations $x_1, \ldots, x_n$ drawn from this distribution. Determine the maximum likelihood estimator of the Poisson parameter λ.

2. If $\mathbf{a}$ and $\mathbf{b}$ are $d \times 1$ column vectors and M is a $d \times d$ symmetric matrix, show that $\mathbf{a}^T M \mathbf{b} = \mathbf{b}^T M \mathbf{a}$.

3. Consider the Gaussian distribution

$$p(\mathbf{x}) \propto \exp\left\{-\frac{1}{2}(\mathbf{x}^T A \mathbf{x} - 2\mathbf{x}^T \mathbf{b})\right\}$$

where A is a symmetric matrix. Show that the mean and covariance are given by

$$\text{mean}(\mathbf{x}) = A^{-1}\mathbf{b}. \qquad \text{cov}(\mathbf{x}) = A^{-1}.$$

4. **Try this to help your understanding of Gaussian distributions and samples from Gaussian distributions.** The bivariate Gaussian has covariance matrix

$$\Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix}.$$

σ_{11} defines the variance on axis 1, and similarly σ_{22} for the variance on axis 2. Of course because of the symmetry of the covariance matrix $\sigma_{12} = \sigma_{21}$. The correlation coefficient ρ is defined as $\rho = \sigma_{12}/\sqrt{\sigma_{11}\sigma_{22}}$. It can be shown that ρ lies between -1 and 1.

A bivariate Gaussian with zero mean is defined by

$$p(\mathbf{x}) = \frac{1}{(2\pi)^{d/2}|\Sigma|^{1/2}} \exp\left\{-\frac{1}{2}\mathbf{x}^T \Sigma^{-1} \mathbf{x}\right\}.$$

Copy the matlab function plotquad: `http://www.inf.ed.ac.uk/teaching/courses/pmr/code/plotquad.m` from the PMR webpage to your filespace, and use it to make contour and surface plots of the bivariate Gaussian. The function takes three arguments $\sigma_{11}, \sigma_{22}, \rho$. For example `plotquad(1,1,0)` plots the isotropic Gaussian with variance 1 on each axis and zero correlation. Experiment with different values of these three parameters and observe the effects.

If you are unfamiliar with matlab, there is an "Introduction to MATLAB" document available from the course webpage that should get you started.

5. **This questions helps you see how all multivariate Gaussians are related**. Generate a sample of 100 points from the bivariate Gaussian defined in question 4. You should use the matlab functions randn and chol (the Cholesky decomposition) to help you. Hint: the Cholesky decomposition U of a matrix Σ is such that $U^T U = \Sigma$. If $\mathbf{x}$ is a random vector drawn from $N(0, I)$, then $\mathbf{y} = U^T \mathbf{x}$ is a random vector that has mean 0 and covariance $E[\mathbf{y}\mathbf{y}^T] = E[U^T \mathbf{x}\mathbf{x}^T U] = U^T U = \Sigma$.