# Probabilistic Modelling and Reasoning
# Inference with Gaussian Random Variables

Division of Informatics, University of Edinburgh

Instructor: Dr Chris Williams

October 2001

We have two variables $X$ and $Y$. Imagine the $X$ models the position of an object in the world (in one dimension) and $Y$ is an observation, say in a camera, of the position of the object in the camera. A camera calibration procedure tells us the relationship between $X$ and $Y$; in our case we assume

$$Y = 2X + 8 + N_y$$

where $N_y$ is some Gaussian measurement noise with zero mean and variance 1. Thus our model for $P(y|x)$ is

$$P(y|x) = \frac{1}{\sqrt{2\pi}} \exp\{-\frac{1}{2}(y - 2x - 8)^2\}.$$

Also we assume that $x \sim N(0, 1/\alpha)$ so that

$$P(x) = \sqrt{\frac{\alpha}{2\pi}} \exp -\frac{\alpha x^2}{2}.$$

Given this, we want to infer the distribution of $X$ given that $Y = y$. To do this we compute the mean and covariance of $(X, Y)^T$, and then condition on $Y = y$. The mean vector is easily calculated as

$$\boldsymbol{\mu} = \begin{pmatrix} \mu_x \\ \mu_y \end{pmatrix} = \begin{pmatrix} 0 \\ 8 \end{pmatrix}.$$

For the covariance matrix, we have that $\text{var}(X) = 1/\alpha$. For $\text{var}(Y)$ we find

$$\text{var}(Y) = E[(Y - \mu_y)^2] = E[(2X + N_y)^2] = \frac{4}{\alpha} + 1,$$

and for $\text{covar}(XY)$ we find

$$\text{covar}XY = E[(X - \mu_x)(Y - \mu_y)] = E[X(2X + N_y)] = \frac{2}{\alpha}$$

and thus

$$\Sigma = \begin{pmatrix} 1/\alpha & 2/\alpha \\ 2/\alpha & 4/\alpha + 1 \end{pmatrix}.$$

Given a vector of random variables split into two parts $\mathbf{X}_1$ and $\mathbf{X}_2$ with

$$\boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{pmatrix}$$

and

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

the general expression for obtaining the conditional distribution of $\mathbf{X}_1$ given $\mathbf{X}_2$, is

$$\boldsymbol{\mu}^c_{1|2} = \boldsymbol{\mu}_1 + \Sigma_{12}\Sigma_{22}^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2),$$

$$\Sigma^c_{1|2} = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}.$$

Applying this to the case above, we obtain

$$\mu_{x|y} = 0 + \frac{2}{\alpha}.\frac{\alpha}{4+\alpha}(y-8) = \frac{2}{4+\alpha}(y-8)$$

and

$$\mathrm{var}(x|y) = \frac{1}{\alpha} - \frac{2}{\alpha}.\frac{\alpha}{4+\alpha}.\frac{2}{\alpha} = \frac{1}{4+\alpha}.$$

The obvious estimator of $X$ is $(y-8)/2$, which is obtained from inverting the dependence between $y$ and $x$ on the assumption that the noise is zero. We see that this is obtained in the limit $\alpha \to 0$, which corresponds to an improper prior on $X$ with infinite variance. For non-zero $\alpha$, the effect is to "shrink" $\mu_{x|y}$ towards zero, which corresponds to the information in the prior on $X$ that zero is its most likely value. Note that if $\alpha \to \infty$, which corresponds to being certain at the outset that $X = 0$, then this information overwhelms the information coming from the observation, and in this limit $\mu_{x|y} = 0$. Notice also that the posterior variance $1/(4+\alpha)$ is smaller than the prior variance $1/\alpha$.