The Standard Bayesian Linear Model

Please cite as

Movellan J. R. (2011) *The Standard Bayesian Linear Model. Revision 1.* MPLab Tutorials, University of California San Diego

Let

$$Y = bX + \sqrt{K}W \tag{1}$$

where $Y \in \mathcal{R}^n, X \in \mathcal{R}^p, W \in \mathcal{R}^n, K \in \mathcal{R}$ are random variables and $b, \in \mathcal{R}^{n\times p}$ are fixed matrices. We let W be zero mean Gaussian with covariance matrix σ_w. For fixed parameters $\alpha, \beta \in \mathcal{R}$ we let

$$p(k) = \frac{(\beta/2)^{\alpha/2}}{\Gamma(\alpha/2)} k^{-(\alpha/2+1)} e^{-\beta/(2k)} \tag{2}$$

Thus K is an Inverse Gamma variable with parameters $\alpha/2, \beta/2$ (see Appendix). For each k we let X be a multivariate Gaussian random vector with mean μ and covariance matrix $k\sigma$, i.e.,

$$p(x \mid k) = \frac{1}{(2\pi k)^{p/2}|\sigma|^{1/2}} e^{-\frac{1}{2k}(x-\mu)'\sigma^{-1}(x-\mu)} \tag{3}$$

0.1 Marginal Distributions

We already know the marginal distribution of k. For the marginal distribution of x note

$$\begin{aligned} p(x) &= \int p(k)p(x|k)dk \\ &= \int \frac{(\beta/2)^{\alpha/2}}{\Gamma(\alpha/2)} k^{-(\alpha/2+1)} e^{-\beta/(2k)} \frac{1}{(2\pi k)^{p/2}|\sigma|^{1/2}} e^{-\frac{1}{2k}(x-\mu)'\sigma^{-1}(x-\mu)}\, dk \end{aligned} \tag{4}$$

Taking out all the terms constant with respect to k we get

$$p(x) = \frac{(\beta/2)^{\alpha/2}}{\Gamma(\alpha/2)(2\pi)^{p/2}|\sigma|^{1/2}} \int k^{-(\alpha/2+p/2+1)} e^{-\frac{1}{2k}(\beta+(x-\mu)'\sigma^{-1}(x-\mu))}\, dk \tag{5}$$

We note that the integrand looks like an Inverse Gamma distribution with parameters

$$\tilde{\alpha}/2 = (\alpha + p)/2 \tag{6}$$

$$\tilde{\beta}/2 = (\beta + (x-\mu)'\sigma^{-1}(x-\mu))/2 \tag{7}$$

Thus

$$p(x) \propto \int k^{-(\alpha/2+p/2+1)} e^{-\frac{1}{2k}(\beta+(x-\mu)'\sigma^{-1}(x-\mu))}\, dk = \left(\frac{(\tilde{\beta}/2)^{\tilde{\alpha}/2}}{\Gamma(\tilde{\alpha}/2)}\right)^{-1} \tag{8}$$

and

$$p(x) \propto (\beta + (x-\mu)'\sigma^{-1}(x-\mu))^{-(\alpha+p)/2} \tag{9}$$

$$\propto \Big(1 + (x-\mu)'(\beta\sigma)^{-1}(x-\mu)\Big)^{-(\alpha+p)/2} \tag{10}$$

This is a p-dimensional t distribution with α degrees of freedom.

0.2 Statistics of the Prior Distribution

$$E[X \mid k] = \mu \tag{11}$$
$$\mathrm{Var}[X \mid k] = k\sigma \tag{12}$$
$$E[K] = \frac{\beta}{\alpha - 2} \tag{13}$$
$$\mathrm{Mode}[K] = \frac{\beta}{\alpha + 2} \tag{14}$$
$$\mathrm{Var}[K] = \frac{2\beta}{(\alpha - 2)^2(\alpha - 4)} \tag{15}$$
$$E[X] = \int p(k)E[X|k]dk = \mu \tag{16}$$
$$\begin{aligned}\mathrm{Var}[X] &= \int p(k)(X - E[X|k] + E[X|k] - E[X])^2 dk \\ &= \int p(k)\mathrm{Var}[X|k]dk + \int p(k)(E[X|k] - E[X])^2 dl \\ &= \int p(k)k\sigma dk = E[K]\sigma = \frac{\beta}{\alpha - 2}\sigma \end{aligned}\tag{17}$$

0.3 Posterior Distributions

Given a prior distribution $p(x, k)$ and an observation vector y, our goal is to find the posterior distribution $p(x, k \mid y)$. Note

$$\begin{aligned} p(x,k \mid y) &\propto p(k)p(x \mid k)p(y \mid k, x) \\ &\propto k^{-(\alpha/2+1)} e^{-\beta/(2k)} \frac{1}{k^{p/2}} e^{-\frac{1}{2k}(x-\mu)'\sigma^{-1}(x-\mu)} \frac{1}{k^{n/2}} e^{-\frac{1}{2k}(y-bx)'\sigma_w^{-1}(y-bx)} \end{aligned}\tag{18}$$

Rearranging terms we get

$$p(x, k \mid y) \propto k^{-(\alpha+p+n+2)/2} e^{-q/(2k)} \tag{19}$$

where

$$q = \beta + (x - \mu)'\sigma^{-1}(x - \mu) + (y - bx)'\sigma_w^{-1}(y - bx) \tag{20}$$

We will now aim to write q in the following form

$$q = \tilde{\beta} + (x - \tilde{\mu})'\tilde{\sigma}^{-1}(x - \tilde{\mu}) \tag{21}$$

To do so we first gather all the terms quadratic on x

$$x'\sigma^{-1}x + x'b'\sigma_w^{-1}bx = x'\tilde{\sigma}^{-1}x \tag{22}$$

Thus

$$\tilde{\sigma}^{-1} = \sigma^{-1} + b'\sigma_w^{-1}b \tag{23}$$

Next we gather the terms linear on x

$$-2x'\sigma^{-1}\mu - 2x'b'\sigma_w^{-1}y = -2x'\tilde{\sigma}^{-1}\tilde{\mu} \tag{24}$$

Thus

$$\tilde{\mu} = \tilde{\sigma}\Big(\sigma^{-1}\mu + b'\sigma_w^{-1}y\Big) \tag{25}$$

We note that $\tilde{\mu}$ and $\tilde{\sigma}$ can also be expressed as follows

$$\tilde{\mu} = \mu + g(y - b\mu) \tag{26}$$

$$\tilde{\sigma} = (I - gb)\sigma \tag{27}$$

where g can be thought of as a gain term that modulates the effect of the error term $y - b\mu$

$$g = \sigma b'(\sigma_w + b'\sigma b)^{-1} \tag{28}$$

$$\tag{29}$$

Next we gather the terms constant with respect to x

$$\beta + \mu'\sigma^{-1}\mu + y'\sigma_w^{-1}y = \tilde{\beta} + \tilde{\mu}'\tilde{\sigma}^{-1}\tilde{\mu} \tag{30}$$

Thus

$$\tilde{\beta} = \beta + \mu'\sigma^{-1}\mu + y'\sigma_w^{-1}y - \tilde{\mu}'\tilde{\sigma}^{-1}\tilde{\mu} \tag{31}$$

Finally we let

$$\tilde{\alpha} = \alpha + n \tag{32}$$

and note

$$p(x, k \mid y) \propto k^{-(\tilde{\alpha}/2+1)} e^{-\tilde{\beta}/(2k)} \frac{1}{k^{p/2}} e^{-\frac{1}{2k}(x-\tilde{\mu})'\tilde{\sigma}^{-1}(x-\tilde{\mu})} \tag{33}$$

which we recognize as the product of an Inverse Gamma and a Gaussian distribution.

$$p(k \mid y) \propto k^{-(\tilde{\alpha}/2+1)} e^{-\tilde{\beta}/(2k)} \tag{34}$$

$$p(x \mid k, y) \propto \frac{1}{k^{p/2}} e^{-\frac{1}{2k}(x-\tilde{\mu})'\tilde{\sigma}^{-1}(x-\tilde{\mu})} \tag{35}$$

The proportionality constants follow

$$p(k \mid y) = \frac{(\tilde{\beta}/2)^{\tilde{\alpha}/2}}{\Gamma(\tilde{\alpha}/2)} k^{-(\tilde{\alpha}/2+1)} e^{-\tilde{\beta}/(2k)} \tag{36}$$

$$p(x \mid x, y) = \frac{1}{(2\pi k)^{p/2} |\tilde{\sigma}|^{1/2}} e^{-\frac{1}{2k}(x-\tilde{\mu})'\tilde{\sigma}^{-1}(x-\tilde{\mu})} \tag{37}$$

0.4 Statistics of the Posterior Distribution

$$g = \sigma b'(\sigma_w + b'\sigma b)^{-1} \tag{38}$$

$$\tilde{\mu} = \mu + g(y - b\mu) \tag{39}$$

$$\tilde{\sigma} = (I - gb)\sigma \tag{40}$$

$$\tilde{\alpha} = \alpha + n \tag{41}$$

$$\tilde{\beta} = \beta + \mu'\sigma^{-1}\mu + y'\sigma_w^{-1}y - \tilde{\mu}\tilde{\sigma}^{-1}\tilde{\mu} \tag{42}$$

$$E[X \mid y, k] = \tilde{\mu} \tag{43}$$

$$\mathrm{Var}[X \mid k] = k\tilde{\sigma} \tag{44}$$

$$E[K] = \frac{\tilde{\beta}}{\tilde{\alpha} - 2} \tag{45}$$

$$\mathrm{Mode}[K] = \frac{\tilde{\beta}}{\tilde{\alpha} + 2} \tag{46}$$

$$\mathrm{Var}[K] = \frac{2\tilde{\beta}}{(\tilde{\alpha} - 2)^2(\tilde{\alpha} - 4)} \tag{47}$$

$$E[X] = \tilde{\mu} \tag{48}$$

$$\mathrm{Var}[X] = \frac{\tilde{\beta}}{\tilde{\alpha} - 2}\tilde{\sigma} \tag{49}$$

1 Appendix

1.1 Gamma Probability Density

$$p(x \mid \alpha, \beta) = \frac{x^{\alpha-1} e^{-x/\beta}}{\beta^\alpha \Gamma(\alpha)} \tag{50}$$

where Γ is the gamma function.

$$\Gamma(u) = \int_0^\infty f^{u-1} e^{-t} dt \tag{51}$$

For the case in which α is an integer it can be shown that if

$$Y = \sum_{i=1}^{\alpha} Z_i \tag{52}$$

where Z_i are iid exponential random variables with expected value β then Y is a Gamma distribution with parameters α, β. It can be shown

$$E[X \mid \alpha, \beta] = \alpha\beta \tag{53}$$

$$\text{Var}[X \mid \alpha, \beta] = \alpha\beta^2 \tag{54}$$

$$\text{Mode}[X \mid \alpha, \beta] = (\alpha - 1)\beta \tag{55}$$

1.2 Inverse Gamma Probability Density

If X has a Gamma $\alpha, 1/\beta$ distribution then $Y = 1/X$ has an inverse Gamma density function

$$p(y \mid \alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} y^{-(\alpha+1)} e^{-\frac{\beta}{y}} \tag{56}$$

This follows from the fact that if $y = f(x)$ is a monotonic function then

$$p_Y(y) = \frac{p_X(x)}{|f'(y)|} \tag{57}$$

where f' is the derivative of f.It can be shown that

$$E[Y \mid \alpha, \beta] = \frac{\beta}{\alpha - 1} \tag{58}$$

$$\text{Mode}[Y \mid \alpha, \beta] = \frac{\beta}{\alpha + 1} \tag{59}$$

$$\text{Var}[Y \mid \alpha, \beta] = \frac{\beta^2}{(\alpha - 1)^2 (\alpha - 2)} \tag{60}$$

$$E[Y^{-1} \mid \alpha, \beta] = \frac{\alpha}{\beta} \tag{61}$$