Posterior(s) for Normal(s)

Elad Zippory

September, 14, 2015

1 unknown mean, known variance

$Y \backsim N(\mu, \sigma^2); \mu \backsim N(\mu_0, \sigma_0^2)$
WTS: $P(\mu|y)$ and the weights of $P(y|\mu)$ and $P(\mu)$
$P(\mu) \propto exp[-\frac{(\mu-\mu_0)^2}{2\sigma_0^2}]$
$P(y|\mu) \propto exp[-\frac{\sum(y-\mu)^2}{2\sigma^2}]$
$P(\mu|y) \propto exp[-\frac{(\mu-\mu_0)^2}{2\sigma_0^2} - \frac{\sum(y-\mu)^2}{2\sigma^2}]$ which tells us nothing about the weights or even a nice Gaussian kernel.

Luckily, we know that

$$\begin{aligned} f(x) &= a(y-x)^2 + b(x-z)^2 \\ &= (a+b)\left[x - \frac{ay+bz}{a+b}\right]^2 + \frac{ab}{a+b}(y-z)^2 \end{aligned}$$

$$\begin{aligned} f(x) &= exp\left[a(y-x)^2 + b(x-z)^2\right] \\ &\backsim exp\left[(a+b)\left[x - \frac{ay+bz}{a+b}\right]^2\right] \end{aligned}$$

$$exp[-\frac{(\mu-\mu_0)^2}{2\sigma_0^2} - \frac{\sum(y-\mu)^2}{2\sigma^2}] \propto exp\left[-\frac{1}{2}\left[\frac{1}{\sigma_0^2} + \frac{n}{\sigma^2}\right]\right]\left[\mu - \frac{\frac{\mu_0}{\sigma_0^2} + \frac{n\bar{y}}{\sigma^2}}{\frac{1}{\sigma_0^2} + \frac{n}{\sigma^2}}\right]$$

So we got a nice kernel. Let's interpret the dynamics of the weights?

2 known mean, unknown variance

$Y \backsim N(\mu, \sigma^2); \sigma^2 \backsim \Gamma^{-1}(a, b)$
We use the inverse gamma because:

1. it's positive

2. it is a conjugate of the normal (the posterior will be inverse gamma)

3. it is a general case of χ^{-1}

$P(\sigma^2) \propto (\sigma^2)^{-\frac{a+2}{2}} exp\left[\frac{-ab}{2\sigma^2}\right]$
$P(y|\mu, \sigma^2) \propto (\sigma^2)^{\frac{-n}{2}} exp\left[\frac{-\sum(y_i-\mu)^2}{2\sigma^2}\right]$
$P(\sigma^2|y) \propto (\sigma^2)^{-\frac{a+2+n}{2}} exp\left[\frac{-(ab+\sum(y_i-\mu)^2)}{2\sigma^2}\right]$, which is the kernel of an inverse gamma. Note that it reduces to $\frac{1}{\sigma^2}$ as $a, b \to 0$ because it reduces to the sample estimate.

3 A look at regression with known variance

$y \backsim N(X\beta, \sigma^2 I_n); P(\beta) \propto exp\left[(\beta-\beta_0)'\Sigma_0^{-1}(\beta-\beta_0)\right]$
$P(\beta|y) \propto exp\left[(y-x\beta)'(\sigma^2 I_n)^{-1}(y-x\beta) + (\beta-\beta_0)'\Sigma_0^{-1}(\beta-\beta_0)\right]$
for simplicity, let $\beta_0 = \mathbf{0}$ as we are now more interested in Σ_0.

$$\Sigma_0 = \begin{bmatrix} \tau_1^2 & 0 \\ 0 & \tau_K^2 \end{bmatrix}$$

$$\begin{aligned} P(\beta|y) &\propto exp\left[-\frac{1}{2}(y-x\beta)'(\sigma^2 I_n)^{-1}(y-x\beta) + \beta'\Sigma_0^{-1}\beta\right] \\ &\propto exp\left[-\frac{1}{2}\beta'(x'x(\sigma^2 I_n)^{-1} + \Sigma_0^{-1})\beta - 2\beta' x' y(\sigma^2 I_n)^{-1}\right] \end{aligned}$$

Let $V = x'x(\sigma^2 I_n)^{-1} + \Sigma_0^{-1}$, look at $\beta V\beta - 2\beta' x' y\sigma^2 I_n^{-1}$. Can we see the $(X-?)??(X-?)$ format in it? (hint: add a constant)
$P(\beta|y) \propto exp\left[-\frac{1}{2}(\beta - V^{-1}x'y\sigma^2 I_n^{-1})'V(\beta - V^{-1}x'y\sigma^2 I_n^{-1})\right]$, which is our beloved Gaussian kernel.
$\beta|y \backsim N(V^{-1}x'y\sigma^2 I_n^{-1}, V^{-1})$
So, as $\tau_k \to \infty$, $V \to \sigma^2(X'X)^{-1}$
and $E[\beta|y, x] = (x'x)^{-1}x'y$