Normal Model

Zhenhua Wang

12/7/2019

Normal distribution

TODO: introduction to normal

Univariate Normal Model

Suppose our sampling model is univariate normal distribution.

$$
\begin{aligned}
y_1, ..., y_n|\theta, \sigma^2 &\sim Normal(\theta, \sigma^2) \\
P(y_i|\theta, \sigma^2) &= \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y_i-\theta)^2}{2\sigma^2}}
\end{aligned}
$$

Notice that normal distribution have two parameters, namely θ, σ^2. So we need to find priors for both parameters.

Conjugate Prior

To find conjugate priors for θ, we rewrite the sampling model proportional to θ

$$
\begin{aligned}
P(y_1, ..., y_n|\theta, \sigma^2) &\propto_\theta exp(-\frac{1}{2\sigma^2}\sum(y_i-\theta)^2) \\
&\propto_\theta exp(-\frac{1}{2\sigma^2}\sum(\theta^2 - 2y_i\theta) \\
&\propto_\theta exp(-\frac{1}{2\sigma^2}(n\theta^2 - 2\sum y_i\theta)
\end{aligned}
$$

This sugguests that our prior should contains 2nd order terms in exponent. The most simple prior that contains similar terms is Normal distribution. So $\theta \sim Normal(\mu_0, \tau_0^2)$, $P(\theta) \propto exp(-\frac{1}{2\tau_0^2}(\theta^2 - 2\mu_0\theta))$. We could easily compute the full conditional distribution for θ.

$$
\begin{aligned}
P(\theta|y_1, ..., y_n, \sigma^2) &\propto P(y_1, ..., y_n|\theta, \sigma^2)P(\theta) \\
&\propto exp(-\frac{1}{2\sigma^2}(n\theta^2 - 2\sum y_i\theta) \times exp(-\frac{1}{2\tau_0^2}(\theta^2 - 2\mu_0\theta)) \\
&\propto exp(-\frac{1}{2}([\frac{1}{\tau_0^2} + \frac{n}{\sigma^2}]\theta^2 - 2[\frac{\mu_0}{\tau_0^2} + \frac{n\bar{y}}{\sigma^2}]\theta)) \\
\theta|y_1, ..., y_n, \sigma^2 &\sim Normal([\frac{1}{\tau_0^2} + \frac{n}{\sigma^2}]^{-1}[\frac{\mu_0}{\tau_0^2} + \frac{n\bar{y}}{\sigma^2}], [\frac{1}{\tau_0^2} + \frac{n}{\sigma^2}]^{-1})
\end{aligned}
$$

As can be seen from the pdf of normal distribution, the conjugate prior for σ^2 is inverse-gamma distribution.

$$
\begin{aligned}
\sigma^2 &\sim InverseGamma(\nu_0/2, \nu_0\sigma_0^2/2) \\
P(\sigma^2) &\propto (\frac{1}{\sigma^2})^{\frac{\alpha_0}{2}+1} exp(-\frac{\nu_0\sigma_0^2}{2}\frac{1}{\sigma^2})
\end{aligned}
$$

We could easily compute the full conditional distribution for σ^2.

$$
\begin{aligned}
P(\sigma^2|y_1, ..., y_n, \theta) &\propto P(y_1, ..., y_n|\theta, \sigma^2)P(\sigma^2) \\
&\propto (\frac{1}{\sigma^2})^{\frac{n}{2}} exp(\frac{\sum(y_i - \theta)^2}{2}\frac{1}{\sigma^2}) \times (\frac{1}{\sigma^2})^{\frac{\nu_0}{2}+1} exp(-\frac{\nu_0\sigma_0^2}{2}\frac{1}{\sigma^2}) \\
\sigma^2|y_1, ..., y_n, \theta &\sim InverseGamma(\frac{\nu_0 + n}{2}, \frac{\nu_0\sigma_0^2 + \sum(y_i - \theta)^2}{2})
\end{aligned}
$$

Here is the sampler for normal model

```
normal_sampler = function(S, y, mu0, tau20, nu0, sigma20) {
  # init
  sigma2 = tau20
  n = length(y)
  ybar = mean(y)

  theta.mcmc = c()
  sigma2.mcmc = c()
  for (s in 1:S) {
    # update theta
    tau2n = 1 / (1/tau20 + n/sigma2)
    mun = tau2n * (mu0/tau20 + n*ybar / sigma2)
    theta = rnorm(1, mean = mun, sd = sqrt(tau2n))

    # update sigma2
    nun = nu0 + 1
    sigma2n = nu0*sigma20 + sum((y - theta)^2)
    sigma2 = 1/rgamma(1, nun, sigma2n)

    # save
    theta.mcmc[s] = theta
    sigma2.mcmc[s] = sigma2
  }
  return(list(theta = theta.mcmc, sigma2 = sigma2.mcmc))
}
```

Multivariate Normal Model

Now our data sampling model becomes multivariate normal distribution

$$
\begin{aligned}
y_1, ..., y_n|\theta, \Sigma &\sim MVNormal(\theta, \Sigma) \\
P(y_i|\theta, \Sigma) &= (2\pi)^{k/2}|\Sigma|^{1/2}e^{-\frac{1}{2}(x-\theta)^T\Sigma^{-1}(x-\theta)}
\end{aligned}
$$

Let's first derive the full conditional for θ. Similar to the univariate case, we choose our conjugate prior for θ to be $MVN(\mu_0, \Lambda_0)$.

$$
\begin{aligned}
P(\theta|y_1,...,y_n,\Sigma) &\propto P(y_1,...,y_n|\theta,\Sigma)P(\theta) \\
&\propto exp\{-\frac{1}{2}\sum(y_i-\theta)^T\Sigma^{-1}(y_i-\theta)\} \times exp\{-\frac{1}{2}\sum(\theta-\mu_0)^T\Lambda_0^{-1}(\theta-\mu_0)\} \\
&\propto exp\{-\frac{1}{2}[\sum(\theta^T\Sigma^{-1}\theta - 2\theta^T\Sigma^{-1}y_i) + (\theta^T\Lambda_0^{-1}\theta) - 2\theta^T\Lambda_0^{-1}\mu_0)]\} \\
&\propto exp\{-\frac{1}{2}[\theta^T(n\Sigma^{-1}+\Lambda_0^{-1})\theta - 2\theta^T(n\Sigma^{-1}\bar{y}+\Lambda_0^{-1}\mu_0)]\} \\
\theta|y_1,...,y_n,\Sigma &\sim MVN([n\Sigma^{-1}+\Lambda_0^{-1}]^{-1}(n\Sigma^{-1}\bar{y}+\Lambda_0^{-1}\mu_0), [n\Sigma^{-1}+\Lambda_0^{-1}]^{-1})
\end{aligned}
$$

Similarly, we choose the conjugate prior for Σ to be $InverseWishart(\nu_0, S_0^{-1})$

$$
\begin{aligned}
\Sigma &\sim InverseWishart(\nu_0, S_0^{-1}) \\
P(\Sigma) &\propto |\Sigma|^{\frac{-(\nu_0+p+1)}{2}} exp\{\frac{-tr(S_0\Sigma^{-1})}{2}\}
\end{aligned}
$$

We could easily compute the full conditional distribution for Σ. Let $S_\theta = \sum(y_i-\theta)(y_i-\theta)^T$

$$
\begin{aligned}
P(\Sigma|y_1,...,y_n,\theta) &\propto P(y_1,...,y_n|\theta,\Sigma)P(\Sigma) \\
&\propto |\Sigma|^{n/2} exp\{-\frac{1}{2}\sum(y_i-\theta)^T\Sigma^{-1}(y_i-\theta)\} \times |\Sigma|^{\frac{-(\nu_0+p+1)}{2}} exp\{\frac{-tr(S_0\Sigma^{-1})}{2}\} \\
&\propto |\Sigma|^{n/2} exp\{-\frac{1}{2}S_\theta\Sigma^{-1}\} \times |\Sigma|^{\frac{-(\nu_0+p+1)}{2}} exp\{\frac{-tr(S_0\Sigma^{-1})}{2}\} \\
&\propto |\Sigma|^{\frac{-((\nu_0+n)+p+1)}{2}} exp\{-\frac{1}{2}tr((S_0+S_\theta)\Sigma^{-1})\} \\
\Sigma|y_1,...,y_n,\theta &\sim InverseWishart(\nu_0+n, [S_0+S_\theta]^{-1})
\end{aligned}
$$