Analysing normal Distribution

Mariza Spyropoulou

2023-09-20

R Markdown

This is an R Markdown document describing how we can create posterior distributions having likelihood and prior follow a normal distribution. We will explain as well how we can integrate out parameters that are not of interest from the final distribution.

When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

we present an very simple example of $x \sim N(\mu, Q^{-1})$ in order to showcase how we can easily identify the precision matrix, Q and the mean value μ.

$$f(x) \propto \exp\left(-\frac{1}{2}(x-\mu)^T Q(x-\mu)\right)$$
$$f(x) \propto \exp\left(-\frac{1}{2}(x^T Qx - \mu^T Qx - x^T Q\mu + \mu^T Q\mu)\right)$$
$$f(x) \propto \exp\left(-\frac{1}{2}(\color{red}{x^T Qx} - \color{blue}{2x^T Q\mu} + \mu^T Q\mu)\right)$$

From the function above knowing which is the value x we can easily identify the value of Q form the red part. The μ value though sometimes is not straightforward to extract it. We showcase how we can go from the blue value to get the parameter.

$$\color{blue}{2x^T Q\mu} \rightarrow x \quad \text{is known}$$
$$\mu = Q^{-1} Q\mu$$

Hence, having the blue term we can get the $Q\mu$ value where simply multiplying with the Q^{-1} we end up with the value of interest, μ.

Next, we will show an example where we can use the above information to get the posterior distribution.

$$\text{Likelihood: } Y = a + \beta X + \epsilon, \epsilon \sim N(0, \Sigma_1)$$
$$\text{Prior: } \beta \sim N(\mu, \Sigma_2)$$

$$\text{Posterior for } \beta\text{: } p(\beta|.) \propto p(y|\beta)p(\beta)$$

Analysing the likelihood distribution in terms of the beta parameter

$$p(y|\beta) \propto \exp(Y - a - X\beta)^T \Sigma_1^{-1} (Y - a - X\beta)$$
$$\propto (Y-a)^T \Sigma_1^{-1}(Y-a) - (Y-a)^T \Sigma_1^{-1} X\beta - (X\beta)^T \Sigma_1^{-1}(Y-a) + (X\beta)^T \Sigma_1^{-1}(X\beta)^T$$
$$\propto (Y-a)^T \Sigma_1^{-1}(Y-a) - 2\beta^T \color{blue}{X^T \Sigma_1^{-1}(Y-a)} + \beta^T \color{red}{X^T \Sigma_1^{-1} X}\beta$$

The term in the red colour is the new precision matrix $Q = X^T \Sigma_1^{-1} X$ or Var = $\left(X^T \Sigma_1^{-1} X\right)^{-1}$.
The term in the blue colour is the $Q\mu$ where we need to multiply with Q to get the μ term
$\mu = Q^{-1}\color{blue}{Q\mu} = Q^{-1} X^T \Sigma_1^{-1}(Y-a)$.

To proceed with the posterior distribution we need to get in the same format the normal prior. Prior for β:

$$p(\beta) \propto \exp(\beta - \mu_2)^T \Sigma_2^{-1}(\beta - \mu_2)$$
$$\propto \exp\left(\beta^T \color{red}{\Sigma_2^{-1}}\beta - 2\beta^T \color{blue}{\Sigma_2^{-1}\mu_2} + \mu_2^T \Sigma_2^{-1}\mu_2\right)$$

The new precision matrix is $Q_2 = \Sigma_2^{-1}$ and $\mu = Q_2^{-1} Q_2 \mu = Q_2^{-1}\color{blue}{\Sigma_2^{-1}\mu_2}$ for the prior. Combining likelihood and prior distributions we can spot different precision matrix, Q and μ vector.

$$\beta^T \color{red}{\left(X^T \Sigma_1^{-1} X + \Sigma_2^{-1}\right)} \beta$$
$$2\beta^T \color{blue}{\left(X^T \Sigma_1^{-1}(Y-a) + \Sigma_2^{-1}\mu_2\right)}$$

Therefore the posterior distribution is defined as:

$$Q = X^T \Sigma_1^{-1} X + \Sigma_2^{-1}$$
$$Q\mu = \left(X^T \Sigma_1^{-1}(Y-a) + \Sigma_2^{-1}\mu_2\right)$$
$$\beta|y \sim N\left(Q^{-1}\left(X^T \Sigma_1^{-1}(Y-a) + \Sigma_2^{-1}\mu_2\right), Q^{-1}\right)$$

Next we will talk about integrating out parameters from a normal distribution which can also be generalised to other distributions. Let's define likelihood and priors first

$$Y = a + X\zeta + D\beta + B\delta + \epsilon, \quad \epsilon \sim N\left(0, \sigma^2\right)$$

Parameters of interest are the $\begin{pmatrix} \zeta \\ \delta \end{pmatrix}$

$$X\zeta + B\delta = \begin{pmatrix} X & B \end{pmatrix} \begin{pmatrix} \zeta \\ \delta \end{pmatrix}$$
$$U = \begin{pmatrix} X & B \end{pmatrix}$$
$$\psi = \begin{pmatrix} \zeta \\ \delta \end{pmatrix}$$

$$p(\zeta, \delta|.) \propto \exp\left(-\frac{1}{2}(Y - a - D\beta - (X\zeta + B\delta))^T \left(\sigma^2 I\right)^{-1} (Y - a - D\beta - (X\zeta + B\delta))\right)$$
$$\propto \exp\left(-\frac{1}{2}(Y - a - D\beta - U\psi)^T \left(\sigma^2 I\right)^{-1} (Y - a - D\beta - U\psi)\right)$$
$$\propto \exp\left(-\frac{1}{2}\left((Y-a-D\beta)^T \left(\sigma^2 I\right)^{-1}(Y-a-D\beta) - 2\psi^T \color{blue}{U^T(\sigma^2 I)^{-1}(Y-a-D\beta)} + \psi^T \color{red}{U^T(\sigma^2 I)^{-1} U}\psi\right)\right)$$

$$Q = U^T(\sigma^2 I)^{-1} U$$
$$P = U^T(\sigma^2 I)^{-1}(Y - a - D\beta)$$
$$\psi = \begin{pmatrix} \zeta \\ \delta \end{pmatrix} |. \sim N\left(Q^{-1}P, Q^{-1}\right)$$

From the above function,posterior distribution we want to integrate out the ψ parameter. We know that any probability density function needs to be

$$f(x) > 0, \quad \int_x f(x)dx = 1$$

In our case it will be

$$\int_x f(x)dx = 1 \rightarrow \int_x (2\pi)^{-1/2} \det(Q^{-1})^{-1/2} \exp\left(-\frac{1}{2}(\psi - Q^{-1}P)^T Q(\psi - Q^{-1}P)\right)dx = 1$$
$$\det(Q^{-1})^{-1/2} \int_x \exp\left(-\frac{1}{2}(\psi^T Q\psi - 2\psi^T QQ^{-1}P + P^T Q^{-1} QQ^{-1}P)\right)dx = (2\pi)^{1/2}$$
$$\int_x \exp\left(-\frac{1}{2}(\psi^T Q\psi - 2\psi^T QQ^{-1}P\right) \exp\left(-\frac{1}{2}P^T Q^{-1}P\right)dx = (2\pi)^{1/2}\det(Q)^{-1/2}$$
$$\int_x \exp\left(-\frac{1}{2}(\psi^T Q\psi - 2\psi^T QQ^{-1}P\right)dx = (2\pi)^{1/2}\det(Q)^{-1/2}\exp\left(\frac{1}{2}P^T Q^{-1}P\right)$$

Therefore, a full conditional marginalizing over ζ and δ will be proportional to

$$\det(Q)^{-1/2}\exp\left(\frac{1}{2}P^T Q^{-1}P\right)$$