Mathematical Notes MSc Thesis

Lucas Paul Unterweger

April 3, 2024

1 Important Prior Distributions

Normal Distribution

$$f(x|\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

Gamma Distribution

$$f(x|\alpha,\beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}, \quad x > 0$$

where α is the shape parameter, β is the rate parameter, and $\Gamma(\cdot)$ is the gamma function.

Beta Distribution

$$f(x|a,b) = \frac{x^{a-1}(1-x)^{b-1}}{B(a,b)}$$

where $x \in [0,1]$, $a > 0$, $b > 0$, and $B(a,b)$ is the Beta function defined as:

$$B(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$$

where Γ is the gamma function.

2 Other Functions

Confluent Hyper-Geometrics Function of the second kind (Tricomi's (confluent hypergeometric) function)

$$U(a,b,z) = \frac{1}{\Gamma(a)} \int_0^\infty e^{-zt} t^{a-1} (1+t)^{b-a-1} dt$$

Marginal Prior with $\phi^\xi = \frac{2c^\xi}{\kappa_B^2 a^\xi}$

$$p(\sqrt{\theta_j}|\phi^\xi, a^\xi, c^\xi) = \frac{\Gamma(c^\xi + \frac{1}{2})}{\sqrt{2\pi\phi^\xi} \cdot B(a^\xi, c^\xi)} \cdot U\left(c^\xi + \frac{1}{2}, \frac{3}{2} - a^\xi, \frac{\theta_j}{2\phi^\xi}\right)$$
$$\propto U\left(c^\xi + \frac{1}{2}, \frac{3}{2} - a^\xi, \frac{\theta_j}{2\phi^\xi}\right)$$

3 Bayesian Regression

Consider a Bayesian regression model where the relationship between the independent variable x and the dependent variable y is modeled as:

$$y = \beta_0 + \beta_1 x + \epsilon$$

where ϵ is the error term assumed to follow a normal distribution with mean 0 and variance σ^2.

We assume the following conjugate prior distributions for the parameters:

$$\begin{aligned}\beta_0 &\sim \mathcal{N}(\mu_0, \tau_0^2)\\ \beta_1 &\sim \mathcal{N}(\mu_1, \tau_1^2)\\ \sigma^2 &\sim \text{Inv-Gamma}(\alpha, \beta)\end{aligned}$$

The likelihood function for the observed data (x_i, y_i), $i = 1, \ldots, n$, is given by:

$$L(\beta_0, \beta_1, \sigma^2 | x, y) = \prod_{i=1}^{n} f(y_i | \beta_0, \beta_1, x_i, \sigma^2)$$

where $f(y_i|\beta_0, \beta_1, x_i, \sigma^2)$ is the probability density function (PDF) of the normal distribution.

The posterior distribution of the parameters, denoted as $\pi(\beta_0, \beta_1, \sigma^2|x, y)$, is proportional to the product of the likelihood function and the prior distributions:

$$\pi(\beta_0, \beta_1, \sigma^2|x, y) \propto L(\beta_0, \beta_1, \sigma^2|x, y) \times \pi(\beta_0) \times \pi(\beta_1) \times \pi(\sigma^2)$$

The maximum a posteriori (MAP) estimation involves maximizing the log-posterior function:

$$\ell(\beta_0, \beta_1, \sigma^2|x, y) = \sum_{i=1}^{n} \log f(y_i|\beta_0, \beta_1, x_i, \sigma^2) + \log \pi(\beta_0) + \log \pi(\beta_1) + \log \pi(\sigma^2)$$

with respect to the parameters β_0, β_1, and σ^2.

The estimates of the parameters, denoted as $\hat{\beta}_0$, $\hat{\beta}_1$, and $\hat{\sigma}^2$, can be obtained by maximizing the log-posterior function.

4 Derviation Idea

$$post(\beta, \sigma^2|x, y) \propto L(\beta, \sigma^2|x, y) \times prior(\beta, \sigma^2)$$