Evidence Function

Estimates for α and β can alternatively be obtained by first integrating the product of likelihood and prior over parameters w.

$$p(t|\alpha,\beta) = \int p(t|w,\beta)\, p(w|\alpha)\, dw$$

and then maximizing the resulting marginal likelihood or evidence function w.r.t α and β. This approach is known as empirical Baye's

Log Likelihood is given by :-

$$\log p(t|\alpha,\beta) = \frac{M}{2}\log\alpha + \frac{N}{2}\log\beta - E(m_N) - \frac{1}{2}\log|S_N^{-1}| - \frac{N}{2}\log(2\pi)$$

where

$$E(m_N) = \frac{\beta}{2}\|t - \Phi m_N\|^2 + \frac{\alpha}{2} m_N^T m_N$$

For completeness, the relationship between evidence, likelihood, prior, posterior is of course given by baye's theorem

$$p(w|t,\alpha,\beta) = \frac{p(t|w,\beta)\, p(w|\alpha)}{p(t|\alpha,\beta)}$$

Date
Page

Maximization of log marginal likelihoods gives its solution :-

$$\alpha = \frac{\gamma}{m_N^T m_N}$$

$$\frac{1}{\beta} = \frac{1}{N-\gamma} \sum_{i=1}^{N} \left(t_i - m_N^T \phi(x_i) \right)^2$$

where

$$\gamma = \sum_{i=0}^{M-1} \frac{\lambda_i}{\alpha + \lambda_i}$$

$$E(m_N) = \frac{N-\gamma}{2 \sum_{i=1}^{N} (t_i - m_N^T \phi(x_i))^2} \|t - \phi m_N\|^2 + \frac{\gamma}{2 m_N^T m_N} m_N^T m_N$$

$$E(m_N) = \frac{N-\gamma}{2 \sum_{i=1}^{N} (t_i - m_N^T \phi(x_i))^2} \|t - \phi_{m_N}\|^2 + \frac{\gamma}{2}$$

$$E(m_N) = \frac{N}{2} - \frac{\gamma}{2} + \frac{\gamma}{2}$$

$$E(m_N) = \frac{N}{2}$$

$$\boxed{2E(m_N) = N}$$

Scanned with CamScanner