Sunday, June 16, 2019 23:53

$$J(\mu, \Sigma) = \left(\max\left(0,\ 1 - \mu^T \phi(x) y \right) \right)^2 + \phi(x)^T \Sigma \phi(x)$$

$$+ \gamma \left\{ \log \frac{\det(\tilde{\Sigma})}{\det(\Sigma)} + tr(\tilde{\Sigma}^{-1} \Sigma) + (\mu - \tilde{\mu})^T \tilde{\Sigma}^{-1} (\mu - \tilde{\mu}) - d \right\}$$

$$\frac{\partial J}{\partial \mu} = \frac{\partial \max(0, 1 - \mu^T \phi(x) y)^2}{\partial \mu} + 2\gamma \tilde{\Sigma}^{-1} (\hat{\mu} - \tilde{\mu}) = 0$$

$$\frac{\partial \max(0, 1 - \mu^T \phi(x) y)}{\partial \mu} = \begin{cases} -\phi(x) y & \text{for } 1 - \mu^T \phi(x) y > 0 \\ 0 & \text{otherwise} \end{cases}$$

$$\frac{\partial J}{\partial \mu} = -2\phi(x) y + 2\hat{\mu} \phi(x) y + 2\gamma \tilde{\Sigma}^{-1} (\mu - \tilde{\mu}) = 0$$

$$\left[\tilde{\Sigma}^{-1} + \frac{\phi(x) y \phi(x) y}{\gamma} \right] \hat{\mu} = \tilde{\Sigma}^{-1} \tilde{\mu} + \frac{\phi(x) y}{\gamma}$$

$$\hat{\mu} = \left(\tilde{\Sigma}^{-1} + \frac{\phi(x) y \phi(x) y}{\gamma} \right)^{-1} \left\{ \tilde{\Sigma}^{-1} \tilde{\mu} + \frac{\phi(x) y}{\gamma} \right\}$$

$$= \left(y\tilde{\Sigma} - \frac{\tilde{\Sigma} \phi(x) y \phi(x) y \tilde{\Sigma}}{\phi(x)^T \tilde{\Sigma} \phi(x) + \gamma} \right) \left(\tilde{\Sigma}^{-1} \tilde{\mu} + \frac{\phi(x) y}{\gamma} \right)$$

$$= \tilde{\mu} - \frac{\tilde{\Sigma} \phi(x) y \phi(x)^T y \tilde{\mu}}{\phi^T(x) \tilde{\Sigma} \phi(x) + \gamma} + \frac{\phi(x) y}{\gamma} \left(\frac{\gamma \tilde{\Sigma}}{\phi^T(x) \tilde{\Sigma} \phi(x) + \gamma} \right)$$

$$= \tilde{\mu} + y \frac{1 - \tilde{\mu}\phi(x)^T y}{\phi(x)\tilde{\Sigma}\phi(x)^T + \gamma} \tilde{\Sigma}\phi(x)$$

$$= \tilde{\mu} + \frac{\max(0, 1 - \tilde{\mu}\phi(x)^T y)}{\phi(x)\tilde{\Sigma}\phi(x)^T + \gamma} \tilde{\Sigma}\phi(x)$$