CS542 Machine Learning

Homework 3

Chih Wei Tung (U17550453)

5.3

$p(T|x,\ w,\ \Sigma) = \prod_{n=1}^{N} N(t_n|y(x_n, w), \Sigma)$

$\ln p(T|x,\ w,\ \Sigma) = \frac{-N}{2}(\ln|\Sigma| + K\ln(2\pi)) - \frac{1}{2}\sum_{n=1}^{N}(t_n - y_n)^T \Sigma^{-1}(t_n - y_n)$ (Likelihood Function)

Where $y_n = y(x_n, w)$

$E(w) = \frac{1}{2}\sum_{n=1}^{N}(t_n - y_n)^T \Sigma^{-1}(t_n - y_n)$ (Error Function)

Maximize likelihood function w.r.t Σ:

$\frac{-N}{2}\ln|\Sigma| - \frac{1}{2}\sum_{n=1}^{N}(t_n - y_n)^T \Sigma^{-1}(t_n - y_n) = \frac{-N}{2}\ln|\Sigma| - \frac{1}{2}Tr[\Sigma^{-1}\sum_{n=1}^{N}(t_n - y_n)(t_n - y_n)^T]$

$\therefore \Sigma = \frac{1}{N}\sum_{n=1}^{N}(t_n - y_n)(t_n - y_n)^T$

5.4

$p(t = 1|x) = \sum_{k=0}^{1} p(t = 1|k)p(k|n)$

$= (1 - \epsilon)y(x, w) + \epsilon(1 - y(x, w))$

$p(t|x) = p(t = 1|x)^t(1 - p(t = 1|x))^{1-t}$

$\therefore E(w) = -\sum_{n=1}^{N}\{t_n \ln[(1 - \epsilon)y(x_n, w) + \epsilon(1 - y(x_n, w))] + (1 - t_n)\ln[1 - (1 - \epsilon)y(x_n, w) - \epsilon(1 - y(x_n, w))]\}$

5.26

$\Omega_n = \frac{1}{2}\sum_k \left(\sum_i \tau_{ni}\frac{\partial y_{nk}}{\partial x_{ni}}\right)^2$

$= \frac{1}{2}\sum_k(\sum_i \tau_{ni}J_{nki})^2$ (Jacobian J)

$\beta_{nj} = \sum_i w_{ji}\alpha_{ni}$

$= \sum_i w_{ji}\mathcal{G}x_{ni}$

$= \sum_i w_{ji}\sum_{i'}\tau_{ni'}\frac{\partial x_{ni}}{\partial x_{ni'}}$

$= \sum_i w_{ji}\tau_{ni}$

$$\Omega_n = \frac{1}{2}\sum_k (\mathcal{G}y_{nk})^2 = \frac{1}{2}\sum_k \alpha_{nk}^2$$

$$\frac{\partial \Omega_n}{\partial w_{rs}} = \sum_k (\mathcal{G}y_{nk})\mathcal{G}(\delta_{nkr}Z_{ns})$$

$$= \sum_k \alpha_{nk}(\phi_{nkr}Z_{ns} + \delta_{nkr}\alpha_{ns})$$

$$\delta_{nkr} = h'(a_{nr})\sum_l w_{lr}\,\delta_{nkl}$$

$$\phi_{nkr} = \mathcal{G}\delta_{nkr}$$

$$= \mathcal{G}(h'(a_{nr})\sum_l w_{lr}\delta_{nkl})$$

$$= h''(a_{nr})\beta_{nr}\sum_l w_{lr}\delta_{nkl} + h'(a_{nr})\sum_l w_{lr}\phi_{nkl}$$