Important formulas for EE292D, "Statistical learning and pattern classification"

Spring, 2003, Stanford University

David G. Stork

$$P(\omega_i|\mathbf{x}) = \frac{p(\mathbf{x}|\omega_i)P(\omega_i)}{\sum_{k=1}^{c} P(\omega_k)p(\mathbf{x}|\omega_k)}$$

$$p(\mathbf{x}) = \frac{1}{(2\pi)^{d/2}|\mathbf{\Sigma}|^{1/2}} \exp\left[-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^t \mathbf{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right]$$

$$\begin{aligned}
\mathbf{\Sigma} &= \mathcal{E}[(\mathbf{x}-\boldsymbol{\mu})(\mathbf{x}-\boldsymbol{\mu})^t] \\
\mathbf{\Sigma} &= \frac{1}{n}\sum_{i=1}^{n}(\mathbf{x}-\boldsymbol{\mu})(\mathbf{x}-\boldsymbol{\mu})^t \\
\mathbf{\Sigma} &= \int (\mathbf{x}-\boldsymbol{\mu})(\mathbf{x}-\boldsymbol{\mu})^t p(\mathbf{x}) d\mathbf{x}
\end{aligned}$$

$$l(\boldsymbol{\theta}) = \ln p(\mathcal{D}|\boldsymbol{\theta})$$

$$\hat{\boldsymbol{\theta}} = \arg\max_{\boldsymbol{\theta}} l(\boldsymbol{\theta})$$

$$p(\mathbf{x}|\mathcal{D}) = \int p(\mathbf{x}|\boldsymbol{\theta}) p(\boldsymbol{\theta}|\mathcal{D})\ d\boldsymbol{\theta})$$

$$Q(\boldsymbol{\theta};\boldsymbol{\theta}^i) = \mathcal{E}_{\mathcal{D}_b}\left[\ln p(\mathcal{D}_g, \mathcal{D}_b;\boldsymbol{\theta})|\mathcal{D}_g;\boldsymbol{\theta}^i\right]$$

$$P^* \leq P \leq P^*\left(2 - \frac{c}{c-1}P^*\right)$$

$$\begin{aligned}
J_p(\mathbf{a}) &= \sum_{\mathbf{y}\in\mathcal{Y}}(-\mathbf{a}^t\mathbf{y}) \\
\mathbf{a}(k+1) &= \mathbf{a}(k) + \eta(k)\sum_{\mathbf{y}\in\mathcal{Y}_k}\mathbf{y}
\end{aligned}$$

$$\begin{aligned}
\Delta w_{kj} &= \eta\delta_i y_j = \eta(t_k - z_k)f'(net_k)y_j \\
\Delta w_{ji} &= \eta\delta_j x_i = \eta\left[\sum_{k=1}^{c} w_{kj}\delta_k\right] f'(net_j)x_i
\end{aligned}$$

$$f(net) = a\ \tanh[b\ net] = \frac{2a}{1 + e^{-b\ net}} - a$$

$$i(N) = -\sum_j P(\omega_j) \log_2 P(\omega_j)$$

$$\mathcal{E}_{\mathcal{D}}\left[(g(\mathbf{x};\mathcal{D}) - F(\mathbf{x}))^2\right] = (\mathcal{E}_{\mathcal{D}}\left[g(\mathbf{x};\mathcal{D}) - F(\mathbf{x})\right])^2 + \mathcal{E}_{\mathcal{D}}\left[g(\mathbf{x};\mathcal{D}) - \mathcal{E}_{\mathcal{D}}[g(\mathbf{x};\mathcal{D})]\right)^2]$$

$$\begin{aligned} \hat{\theta}_{(i)} &= \hat{\theta}(x_1, x_2, \cdots, x_{i-1}, x_{i+1}, \cdots, x_n) \\ \hat{\theta}_{(\cdot)} &= \frac{1}{n}\sum_{i=1}^{n} \hat{\theta}_{(i)} \end{aligned}$$