COMP9417 - Week 3 Tutorial notes

Linear Regression II

$$N(\mu, 1) \;,\; p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-(x-\mu)^2}{2\sigma^2}\right)$$

$$\log L(m) = \log\left(\prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-(x_i-m)^2}{2\sigma^2}\right)\right)$$

$$= \sum_{i=1}^{n} \log\left(\frac{1}{2\sqrt{\pi\cancel{\sigma^2}}} \exp\left(\frac{-(x_i-m)^2}{2\cancel{\sigma^2}}\right)\right)$$

$$= \sum_{i=1}^{n} \left[\log\left(\frac{1}{\sqrt{2\pi}}\right) - \frac{1}{2}(x_i-m)^2\right]$$

$$= \hat{\mu}_{MLE} = m = \arg\max_{m} \;\ldots$$

$$\frac{\partial l}{\partial m} = 2\left(-\frac{1}{2}\right) \sum_{i=1}^{m} (x_i - m)$$

$$0 = \left(\sum_{i=1}^{n} x_i\right) - nm \rightarrow m = \frac{1}{n} \sum_{i=1}^{n} x_i = \hat{\mu}_{MLE}$$

Bias-Variance Decomposition

Proof:

$$MSE = E[(\hat{\theta} - \theta)^2] \qquad \begin{cases} Var(\hat{\theta}) = E[(\hat{\theta} - E(\hat{\theta}))^2] \\ Bias(\hat{\theta}) = E(\hat{\theta}) - \theta \end{cases}$$

$$= E[((\hat{\theta} - E(\hat{\theta})) + (E(\hat{\theta}) - \theta))^2]$$

$$= E[(\hat{\theta} - E(\hat{\theta}))^2 + 2E\underbrace{(\hat{\theta}) E(E(\hat{\theta}))(E(\hat{\theta}) - }_{}\theta) + E(\hat{\theta} - \theta)^2]$$

$$= E[(\hat{\theta} - E(\hat{\theta}))^2] + 2E[\;---\;] + E[(E(\hat{\theta}) - \theta)^2]$$

$$\downarrow \qquad\qquad\qquad\qquad\qquad\qquad \downarrow$$

$$Var(\hat{\theta}) \; + 0 \qquad\qquad + \quad Bias(\hat{\theta})^2$$

$y_i | x_i = x_i^T \beta^* + \varepsilon$, $\varepsilon \sim N(0, \sigma^2)$

$y_i | x_i \sim N(x^T \beta^*, \sigma^2 I)$

$$\log L(\beta) = \log(P(y|X,\beta))$$

$$= \log\left(\prod_{i=1}^{n} P(y|X,\beta)\right)$$

$$= \sum_{i=1}^{n} \log(P(\cdots))$$

$$= \sum_{i=1}^{n} \log\left[\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-(y_i - x_i^T\beta)^2}{2\sigma^2}\right)\right]$$

$$= \sum_{i=1}^{n} \left(\log\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right) - \frac{1}{2\sigma^2}(y_i - x_i^T\beta)^2\right)$$

$$= -\frac{n}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - x_i^T\beta)^2$$

$$= -\frac{n}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}\|y - X\beta\|_2^2$$

$$\hat{\mu}_{MLE} = \arg\min_{\beta} \|y - X\beta\|_2^2$$

① triangular inequality

② $g(cx) = |c| g(x)$

③ $g(0) = 0$

Cauchy-Swartz inequality:

$\|x\|_p \geq \|x\|_q$, $p < q$