FYS-STK4155 Sept 23

TP = True positive, equ with a hit

TN = True negative equ with correct rejection

Accuracy score:

$$\frac{\sum TP + \sum TN}{N}$$

$N \leftarrow$ number of outputs

FP = False positive equ with false alarm

FN = False negative

Confusion matrix

TP	FP
FN	TN

Precision $= \dfrac{\Sigma TP}{\Sigma TP + \Sigma FP}$

True negative rate $= TNR$

$= \dfrac{TN}{N}$

Recall $= \dfrac{\Sigma TP}{\Sigma TP + \Sigma FN}$

ROC : Receiver - operating characteristic

plot true positive rate against False positive rate

Gains Curv

$$= \frac{\Sigma TP + \Sigma FP}{N}$$

Optimization problem

$C(\hat{\beta}) =$

Taylor expand around

$\hat{\beta} - \beta^{(n)}$ iteration n

$C(\beta^{(n)}) + (g^{(n)})^T(\hat{\beta} - \beta^{(n)})$

$$\boxed{\begin{aligned} \left.\frac{\partial C}{\partial \beta}\right|_{\beta = \beta^{(n)}} &= g^{(n)} \\ \left.\frac{\partial^2 C}{\partial \beta \partial \beta^T}\right|_{\beta = \beta^{(n)}} &= H^{(n)} \end{aligned}}$$

$+ \frac{1}{2}(\hat{\beta} - \beta^{(n)})^T \times H^{(n)} \times (\hat{\beta} - \beta^{(n)}) + \ldots$

$b = \hat{\beta} - \beta^{(n)} = b^{(n)}$

$$C(\hat{\beta}) = C(\beta^{(n)}) + (g^{(n)})^T b + \frac{1}{2} b^T H^{(n)} b + \ldots$$

$$\left(f(x) = \text{const} + \gamma^T x + \tfrac{1}{2} x^T A x \right)$$

$$\frac{\partial c}{\partial b} = \left(g^{(n)}\right) + H^{(n)} b = 0$$

$$\Rightarrow \quad b = \hat{\beta} - \beta^{(n)} =$$

$$- \left(H^{(n)}\right)^{-1} \cdot g^{(n)} \Rightarrow$$

$$\hat{\beta} = \beta^{(n+1)} = \beta^{(n)} - \left(H^{(n)}\right)^{-1} g^{(n)}$$

$$H^{(n)} \to \gamma$$

$$\beta^{(n+1)} = \beta^{(n)} - \gamma g^{(n)}$$

learning rate

Expand c around

$$\beta^{(n)} - \gamma g^{(n)}$$

$$c\Big(\underbrace{\beta^{(n)} - \gamma g^{(n)}}_{\beta^{(n+1)}}\Big) = c\left(\beta^{(n)}\right)$$

$$- \gamma \left(g^{(n)}\right)^T g^{(n)} + \frac{1}{2}\gamma^2 \left(g^{(n)}\right)^T H^{(n)}$$

$$\times g^{(n)} + \ldots$$

Take derivative wrt $\gamma \Rightarrow$

$(g^{(n)} \to g \quad H^{(n)} \to H)$

$$\gamma = \frac{g^T g}{g^T H g}$$

assume $Hg = \lambda g$

$$\gamma = \frac{g^T g}{\lambda g^T g} = \frac{1}{\lambda}$$

$$\gamma_{min} = \frac{1}{\lambda_{max}}$$

$$\gamma_{max} = \frac{1}{\lambda_{min}}$$

The convergence of the scheme is given

$$\gamma < \frac{2}{\lambda_{max}}$$

λ_{max} is the largest eigenvalue of H

General optimization problem