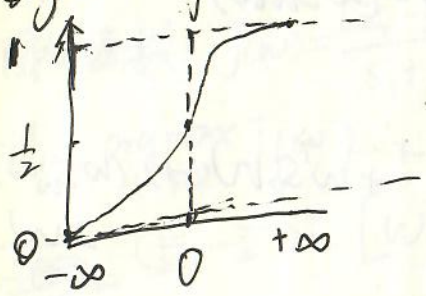
Logistic Regression.(广义线性模型)

sigmoid function: $\sigma(z) = \frac{1}{1+e^{-z}}$



$P_1 = P(y=1|x) = \sigma(w^Tx) = \frac{1}{1+e^{-w^Tx}}, \quad y=1$

$P_0 = P(y=0|x) = 1-P(y=1|x) = \frac{e^{-w^Tx}}{1+e^{w^Tx}}, \quad y=0$

$P(y|x) = P_1^{y} P_0^{1-y}$

MLE: $\hat{w} = \underset{w}{\arg\max} \log P(Y|X)$

Maximum Likelihood Estimate. 最大似然估计.

$= \underset{w}{\arg\max} \log \prod_{i=1}^{N} P(y_i|x_i)$

$= \underset{w}{\arg\max} \sum_{i=1}^{N} \log P(y_i|x_i)$

$= \underset{w}{\arg\max} \sum_{i=1}^{N} y_i \log P_1 + (1-y_i) \log P_0$

$= \underset{w}{\arg\max} \sum_{i=1}^{N} y_i \log P_1 + (1-y_i) \log P_0$

MLE $\Rightarrow$ loss function

max min.

概率生成模型.

Gaussian Discriminant Analysis.

$P(y|x) \propto P(x|y)P(y)$

$\hat{y} = \underset{y\in\{0,1\}}{\arg\max} P(y|x) = \underset{y}{\arg\max} P(y)P(x|y)$

$x|y=1 \sim N(\mu_1, \Sigma)$

$x|y=0 \sim N(\mu_2, \Sigma)$

$y \sim \text{Bernoulli}(\phi) \Rightarrow \begin{array}{c|cc} y & 1 & 0 \\ \hline p & \phi & 1-\phi \end{array} \quad \begin{cases} \phi^{y} & y=1 \\ (1-\phi)^{1-y}, & y=0 \end{cases}$

log-likelihood:

$L(\theta) = \log \prod_{i=1}^{N} P(x_i, y_i) = \sum_{i=1}^{N} \log(P(x_i|y_i)P(y_i))$

$= \sum_{i=1}^{N} [\log P(x_i|y_i) + \log P(y_i)]$

$= \sum_{i=1}^{N} [\log N(\mu_1, \Sigma)^{y_i} + \log N(\mu_2, \Sigma)^{1-y_i} + \log \phi^{y_i}(1-\phi)^{1-y_i}]$