# Written Homework 1

Tim Stolp

06-11-2018

## 1 Introduction

## 2 Regression

a $(r - Xw)^T(r - Xw)$
$Xw = \hat{r}$
$(r - \hat{r})^T(r - \hat{r})$
$\hat{r}^i = w^T\hat{x}^i$
$(r - w^T\hat{x}^i)^T(r - w^T\hat{x}^i)$
$\sum_{i=1}^{N}(r^i - w^T\hat{x}^i)^2$

b $(r - Xw)^T(r - Xw)$
$r^T - w^TX^T(r - Xw)$
$r^Tr - w^TX^Tr - r^TXw + w^TX^TXw$

c $r =$Nx1
$r - Xw$ so $Xw$ must also be Nx1
$X =$Nx(d+1)
$w =$(d+1)x1

d $r - Xw =$ Nx1
$(r - Xw)^T =$ 1xN
$(r - Xw)^T(r - Xw) =$ 1x1
$E(w|X, r) =$ 1x1

e $r^Tr - w^TX^Tr - r^TXw + w^TX^TXw$
$-X^Tr - X^Tr + (X^TX + X^TX)w$
$wX^TXw - 2X^Tr$

f $wX^TXw - 2X^Tr = 0$
$X^TXw = X^Tr$
$Xw = r$
$w = X^{-1}r$

# 3 Bayesian Decision Theory

P = Positive result
I = infected
$p(I|P) = \frac{p(P|I)p(I)}{p(P|I)p(I)+p(P|I^C)p(I^C)}$
p(I) = 0,0052
p($I^C$) = 1 - 0,0052 = 0,9948
p(P|I) = 1
p(P|$I^C$) = 0,03
$\frac{1 \cdot 0{,}0052}{1 \cdot 0{,}0052+0{,}03 \cdot 0{,}9948} \approx 0,148$
-

# 4 Association Rules

$Chair \rightarrow Table$
$Support : \frac{1}{3}$
$Confidence : 1$
$Lift : 2$

$Table \rightarrow Chair$
$Support : \frac{1}{3}$
$Confidence : \frac{2}{3}$
$Lift : 2$

$Flowerpot \rightarrow Table$
$Support : \frac{1}{6}$
$Confidence : \frac{1}{2}$
$Lift : 1$

$Table \rightarrow Flowerpot$
$Support : \frac{1}{6}$
$Confidence : \frac{1}{3}$
$Lift : 1$

## 4.1 Discrimant Functions

a $C_1$ = spam, $C_2$ = not spam
$g(x,y) = p(C_1|x,y) - p(C_2|x,y)$
choose$\begin{cases} C_1 & \text{if } g(x,y) > 0 \\ C_2 & \text{otherwise} \end{cases}$

b without normalizing term.
$g(x,y) = p(x,y|C_1)p(C_1) - p(x,y|C_2)p(C_2)$

$p(C_1) = 0,17$ and $p(C_2) = 0,83$
Take $x = 4$ and $y = 5$

$p(x,y|C_1) = 0,85$
$p(x,y|C_2) = 0,06$

$g(x,y) = 0,85 \cdot 0,17 - 0,06 \cdot 0,83 = 0,0947 \approx 0,095$
$0,095 > 0$ so the email is considered spam

c $g(x,y) = \frac{p(C_1|x,y)}{p(C_2|x,y)}$
choose$\begin{cases} C_1 & \text{if } g(x,y) > 1 \\ C_2 & \text{otherwise} \end{cases}$
Bayes and ignoring the normalization terms:
$g(x,y) = \frac{p(x,y|C_1)}{p(x,y|C_2)} \cdot \frac{p(C_1)}{p(C_2)}$

d $g(x,y) = log\frac{p(C_1|x,y)}{p(C_2|x,y)}$
choose$\begin{cases} C_1 & \text{if } g(x,y) > 0 \\ C_2 & \text{otherwise} \end{cases}$
Bayes and ignoring the normalization terms:
$g(x,y) = log\frac{p(x,y|C_1)}{p(x,y|C_2)} + log\frac{p(C_1)}{p(C_2)}$