HW 0:

Problem 1:

1. $\|Xw - y\|^2 = \|y\|^2 + \|Xw\|^2 - 2y^TXw$

$= \|y\|^2 + w^TX^TXw - 2y^TXw$

gradient $= 0 + 2X^TXw - 2X^T y = 0$

$\Longrightarrow \quad X^TXw = X^Ty$

$\because n \geq m$, if X^TX is inveritble

$w = (X^TX)^{-1}X^Ty$.

2. $\|Xw - y\|^2 + \lambda\|w\|^2$

$= \|y\|^2 + w^TX^TXw - 2y^TXw + \lambda w^Tw$.

gradient: $= 0 + 2X^TXw - 2X^Ty + 2\lambda w = 0$

$(X^TX + \lambda I)w = X^Ty$

if $(X^TX + \lambda I)$ is invertible, then

$w = (X^TX + \lambda I)^{-1}X^Ty$

Problem 2:

1. $Pr(\langle H,H,T,T,H\rangle) = p^3(1-p)^2$ // $\log Pr(\langle H,H,T,T,H\rangle) = 3\log p + 2\log(1-p)$

Given a particular sequence, i stands for the number of 'H' in seq

$Pr(seq) = p^i(1-p)^{5-i}$

2. a. $Pr(fair \cap \langle H,H,T,T,H\rangle) = \frac{1}{2} \times (\frac{1}{2})^3 \times (1-\frac{1}{2})^2 = (\frac{1}{2})^6 = \frac{1}{64}$

b. $Pr(biased \cap \langle H,H,T,T,H\rangle) = \frac{1}{2} \times (\frac{2}{3})^3 \times (1-\frac{2}{3})^2 = \frac{1}{2} \times \frac{8}{27} \times \frac{1}{9} = \frac{4}{243}$

扫描全能王 创建