SDS 385: Exercises 1 - Preliminaries

August 23, 2016

 $Professor\ James\ Scott$

Spencer Woody

Problem 1

(A)

$$\hat{\beta} = \operatorname*{arg\,min}_{\beta \in \mathbb{R}^p} \sum_{i=1}^{N} \frac{w_i}{2} \left(y_i - x_i^T \beta \right)^2 \tag{1}$$

$$= \underset{\beta \in \mathbb{R}^p}{\arg \min} \frac{1}{2} (Y - X\beta)^T W (Y - X\beta)$$
 (2)

$$\frac{1}{2}(Y - X\beta)^T W(Y - X\beta) = \frac{1}{2}(Y^T - \beta^T X^T)W(Y - X\beta)$$
(3)

$$= \frac{1}{2} (Y^T W - \beta^T X^T W)(Y - X\beta) \tag{4}$$

$$= \frac{1}{2} (Y^T W Y - \beta^T X^T W Y - Y^T W X \beta + \beta^T X^T W X \beta)$$
 (5)

$$= \frac{1}{2} (Y^T W Y - 2(X\beta)^T W Y + \beta^T X^T W X \beta)$$
 (6)

$$= \frac{1}{2}Y^TWY - (X\beta)^TWY + \frac{1}{2}\beta^TX^TWX\beta, \tag{7}$$

because

$$\beta^T X^T W Y = (X\beta)^T W Y, \tag{8}$$

and

$$Y^T W X \beta = (Y^T W X \beta)^T :: Y^T W X \beta \in \mathbb{R}^1$$
 (9)

$$(Y^T W X \beta)^T = (W X \beta)^T Y = (X \beta)^T W^T Y = (X \beta)^T W Y. \tag{10}$$

We want to minimize the objective function from Eqn. (7), so we take the gradient with respect to β and set it equal to zero. For each of the three terms, their are respective gradients with respect to β are

(i)

$$\frac{\partial}{\partial \beta} \frac{1}{2} Y^T W Y = 0 \tag{11}$$

(ii)

$$\frac{\partial}{\partial \beta} - (X\beta)^T WY = -X^T WY \tag{12}$$

(iii)

$$\frac{\partial}{\partial \beta} \frac{1}{2} \beta^T X^T W X \beta = \frac{1}{2} \beta^T (X^T W X + (X^T W X)^T)$$
 (13)

$$= X^T W X \beta. \tag{14}$$

Summing these terms and equaling them to zero yields

$$X^T W X \beta - X^T W Y = 0 \tag{15}$$

$$\therefore (X^T W X)\hat{\beta} = X^T W Y \tag{16}$$

(B)

(C)

(D)

Problem 2

- (A)
- (B)
- (C)