STA9792 - Homework 1

Christopher Lang

11/12/2017

Question 1.1

Let $\hat{y} = \beta_0 + \beta_1 x_1 + \beta_2 x_2^2$

Our error function (residual sum of squares) is:

$$L = (y - \hat{y})^2$$

Taking the partial derivative of L with respect to β and setting to zero, we get the following:

$$\frac{\partial L}{\partial \beta_j} = 2(y - \hat{y})(-\frac{\partial}{\partial \beta_j}\hat{y}) = 0$$

$$\frac{\partial L}{\partial \beta_j} = 2(y - (\beta_0 + \beta_1 x_1 + \beta_2 x_2^2))(-\frac{\partial}{\partial \beta_j}(\beta_0 + \beta_1 x_1 + \beta_2 x_2^2)) = 0$$

$$\frac{\partial L}{\partial \beta_j} = 2(y - (\beta_0 + \beta_1 x_1 + \beta_2 x_2^2))(-\frac{\partial}{\partial \beta_j}(\beta_0 + \beta_1 x_1 + \beta_2 x_2^2)) = 0$$

Then our first order conditions are:

$$\frac{\partial L}{\partial \beta_0} = 2(y - (\beta_0 + \beta_1 x_1 + \beta_2 x_2^2))(-1) = 0$$

$$\frac{\partial L}{\partial \beta_1} = 2(y - (\beta_0 + \beta_1 x_1 + \beta_2 x_2^2))(-x_1) = 0$$

$$\frac{\partial L}{\partial \beta_2} = 2(y - (\beta_0 + \beta_1 x_1 + \beta_2 x_2^2))(-x_2^2) = 0$$

Rewritting so that constants are on the right hand side:

$$\beta_0 n + \beta_1 x_1 + \beta_2 x_2^2 = y$$

$$\beta_0 x_1 + \beta_1 x_1 x_1 + \beta_2 x_2^2 x_1 = y x_1$$

$$\beta_0 x_2^2 + \beta_1 x_1 x_2^2 + \beta_2 x_2^2 x_2^2 = y x_2^2$$

And finally in matrix form:

$$\begin{bmatrix} n & \sum x_{i,1} & \sum x_{i,2}^2 \\ \sum x_{i,1} & \sum x_{i,1}^2 & \sum x_{i,2}^2 x_{i,1} \\ \sum x_{i,2}^2 & \sum x_{i,1} x_{i,2}^2 & \sum x_{i,2}^4 \end{bmatrix} \cdot \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix} = \begin{bmatrix} \sum y_i \\ \sum y_i x_{i,1} \\ \sum y_i x_{i,2}^2 \end{bmatrix}$$

Where $\sum \iff \sum_{i=1}^{n}$

Question 1.2

Let $\hat{y} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3$

Our error function (residual sum of squares) is:

$$L = \sum_{i=1}^{n} (y - \hat{y})^2$$

subject to the constraint of $\sum_{i=0}^{2} \beta_i = 1$. Let:

$$g(\beta_j) = \sum_{i=0}^{2} \beta_j - 1 = 0$$

Therefore our Lagrange objective function is:

$$E = \sum_{i=1}^{n} (y - \hat{y})^2 + \lambda \cdot g(\beta_j)$$

$$E = \sum_{i=1}^{n} (y - \hat{y})^2 + \lambda(\sum_{i=0}^{2} \beta_j - 1)$$

Taking the partial derivative of E with respect to β and setting to zero, we get the following:

$$\frac{\partial}{\partial \beta_j} E = \frac{\partial}{\partial \beta_j} [\sum_{i=1}^{n} (y - \hat{y})^2 + \lambda(\sum_{i=0}^{2} \beta_j - 1) = 0$$

$$\frac{\partial}{\partial \beta_j} E = 2 \sum_{i=1}^{n} (y - \hat{y})(0 - \frac{\partial}{\partial \beta_j} \hat{y}) + \lambda(\frac{\partial}{\partial \beta_j} \sum_{i=0}^{2} \beta_j - 0) = 0$$

$$\frac{\partial}{\partial \beta_j} E = 2 \sum_{i=1}^{n} (y - \hat{y})(-\frac{\partial}{\partial \beta_j} \hat{y}) + \lambda(\frac{\partial}{\partial \beta_j} \sum_{i=0}^{2} \beta_j) = 0$$

Hence, the first order conditions will be the following:

$$\frac{\partial E}{\partial \beta_0} = 2 \sum_{i=1}^{n} (y - \beta_0 - \beta_1 x_{i,1} - \beta_2 x_{i,2} - \beta_3 x_{i,3})(-1) + \lambda = 0$$

$$\frac{\partial E}{\partial \beta_1} = 2 \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_{i,1} - \beta_2 x_{i,2} - \beta_3 x_{i,3})(-x_{i,1}) + \lambda = 0$$

$$\frac{\partial E}{\partial \beta_2} = 2 \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_{i,1} - \beta_2 x_{i,2} - \beta_3 x_{i,3})(-x_{i,2}) + \lambda = 0$$

$$\frac{\partial E}{\partial \beta_3} = 2 \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_{i,1} - \beta_2 x_{i,2} - \beta_3 x_{i,3})(-x_{i,3}) + \lambda = 0$$

Then we reformulate the set of equations so that all constants are on the right hand side:

$$\beta_0 n + \beta_1 \sum_{i=1}^{n} x_{i,1} + \beta_2 \sum_{i=1}^{n} x_{i,2} + \beta_3 \sum_{i=1}^{n} x_{i,3} = -\frac{\lambda}{2} + \sum_{i=1}^{n} y_i$$

$$\beta_0 \sum_{i=1}^{n} x_{i,1} + \beta_1 \sum_{i=1}^{n} x_{i,1}x_{i,1} + \beta_2 \sum_{i=1}^{n} x_{i,1}x_{i,2} + \beta_3 \sum_{i=1}^{n} x_{i,1}x_{i,3} = -\frac{\lambda}{2} + \sum_{i=1}^{n} y_i x_{i,1}$$

$$\beta_0 \sum_{i=1}^{n} x_{i,2} + \beta_1 \sum_{i=1}^{n} x_{i,2}x_{i,1} + \beta_2 \sum_{i=1}^{n} x_{i,2}x_{i,2} + \beta_3 \sum_{i=1}^{n} x_{i,2}x_{i,3} = -\frac{\lambda}{2} + \sum_{i=1}^{n} y_i x_{i,2}$$

$$\beta_0 \sum_{i=1}^{n} x_{i,3} + \beta_1 \sum_{i=1}^{n} x_{i,3}x_{i,1} + \beta_2 \sum_{i=1}^{n} x_{i,3}x_{i,2} + \beta_3 \sum_{i=1}^{n} x_{i,3}x_{i,3} = -\frac{\lambda}{2} + \sum_{i=1}^{n} y_i x_{i,3}$$

Rewritten in matrix form:

$$\begin{bmatrix} n & \sum x_{i,1} & \sum x_{i,2} & \sum x_{i,3} \\ \sum x_{i,1} & \sum x_{i,1}^2 & \sum x_{i,2}x_{i,1} & \sum x_{i,3}x_{i,1} \\ \sum x_{i,2} & \sum x_{i,1}x_{i,2} & \sum x_{i,2}^2 & \sum x_{i,3}x_{i,2} \\ \sum x_{i,3} & \sum x_{i,1}x_{i,3} & \sum x_{i,3}x_{i,2} & \sum x_{i,3}^2 \end{bmatrix} \cdot \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} = \begin{bmatrix} \sum y_i \\ \sum y_i x_{i,1} \\ \sum y_i x_{i,2} \\ \sum y_i x_{i,3} \end{bmatrix} - \begin{bmatrix} \frac{\lambda}{2} \\ \frac{\lambda}{2} \\ \frac{\lambda}{2} \\ \frac{\lambda}{2} \end{bmatrix}$$

Where $\sum \iff \sum_{i=1}^{n}$

Question 1.3

Let $\hat{y} = \beta_0 + \beta_1 x_i + \beta_2 \sqrt{x_2}$

Subject to:

- $\sum \beta_1 = 1$
- $\beta_1 = 0.5 \cdot \beta_2 + 0.05$

Let:

- $g(\beta_i) = \sum \beta_i - 1 = 0$
- $h(\beta_1, \beta_2) = \beta_1 - 0.5\beta_2 - 0.05 = 0$

Then we minimize the least squares problem with the following objective function:

$$E = (y - \hat{y})^2 + \lambda_1 g(\beta_i) + \lambda_2 h(\beta_1, \beta_2)$$
$$E = (y - \hat{y})^2 + \lambda_1 (\sum \beta_i - 1) + \lambda_2 (\beta_1 - 0.5\beta_2 - 0.05)$$

$$2(y - \hat{y})(-\frac{\partial}{\partial \beta}\hat{y}) + \lambda_1 (\frac{\partial}{\partial \beta} \sum \beta_i) + \lambda_2 \frac{\partial}{\partial \beta}(\beta_1 - 0.5\beta_2) = 0$$

The first order conditions are:

$$\frac{\partial E}{\partial \beta_0} = 2(y - (\beta_0 + \beta_1 x_1 + \beta_2 \sqrt{x_2}))(-1) + \lambda_1 = 0$$

$$\frac{\partial E}{\partial \beta_1} = 2(y - (\beta_0 + \beta_1 x_1 + \beta_2 \sqrt{x_2}))(-x_1) + \lambda_1 + \lambda_2 = 0$$

$$\frac{\partial E}{\partial \beta_2} = 2(y - (\beta_0 + \beta_1 x_1 + \beta_2 \sqrt{x_2}))(-\sqrt{x_2}) + \lambda_1 + \lambda_2(-0.5) = 0$$

Rewritting the set of equations so that constants are on the right hand side:

$$\beta_0 + \beta_1 x_1 + \beta_2 \sqrt{x_2} = -\frac{\lambda_1}{2} + y$$

$$\beta_0 x_1 + \beta_1 x_1 x_1 + \beta_2 x_1 \sqrt{x_2} = \frac{-\lambda_1 - \lambda_2}{2} + y x_1$$

$$\beta_0 \sqrt{x_2} + \beta_1 \sqrt{x_2} x_1 + \beta_2 x_2 = \frac{-\lambda_1 + 0.5\lambda_2}{2} + y\sqrt{x_2}$$

Rewritten in matrix form:

$$\begin{bmatrix} n & \sum x_{i,1} & \sum \sqrt{x_{i,2}} \\ \sum x_{i,1} & \sum x_{i,1}^2 & \sum x_{i,1}\sqrt{x_{i,2}} \\ \sum \sqrt{x_{i,2}} & \sum x_{i,1}\sqrt{x_{i,2}} & \sum x_{i,2} \end{bmatrix} \cdot \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix} = \begin{bmatrix} \sum y_i \\ \sum y_i x_{i,1} \\ \sum y_i \sqrt{x_{i,2}} \end{bmatrix} + \begin{bmatrix} -\frac{\lambda_1}{2} \\ \frac{-\lambda_1 - \lambda_2}{2} \\ \frac{-\lambda_1 + 0.5\lambda_2}{2} \end{bmatrix}$$