Spring 16 – AMS256 Homework 2 Solutions

1. (a) Let $\underset{(a+b+1)\times 1}{\boldsymbol{\beta}} = [\mu, \alpha_1, \ldots, \alpha_a, \beta_1, \ldots, \beta_b]^T$. Then if we arrange the data by cycling through index j for each i,

$$\underset{(ab)\times(a+b+1)}{\boldsymbol{X}} = \begin{bmatrix} 1 & 1 & 0 & 0 & \cdots & 0 & 0 & 1 & 0 & \cdots & 0 & 0 \\ 1 & 1 & 0 & 0 & \cdots & 0 & 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & 1 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 1 & 0 \\ 1 & 1 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 1 \\ 1 & 0 & 1 & 0 & \cdots & 0 & 0 & 1 & 0 & \cdots & 0 & 0 \\ 1 & 0 & 1 & 0 & \cdots & 0 & 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & 0 & 1 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 1 & 0 \\ 1 & 0 & 1 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 1 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & 0 & 0 & 0 & \cdots & 0 & 1 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix}.$$

The first column is the sum of columns 2 through $a{+}1$ and column 2 is the sum of the last b columns minus columns 3 through $a{+}1$. The remaining columns are linearly independent, so $dim(C(\boldsymbol{X})) = rank(\boldsymbol{X}) = a + b - 1$ and $dim(\mathcal{N}(\boldsymbol{X})) = a + b + 1 - (a + b - 1) = 2$.

(b) $\boldsymbol{X}^T\boldsymbol{X}$ is made up of dot products of every column with every other column (which in this case counts how many 1s are in common between the two columns). It is given by

$$\boldsymbol{X}^T\boldsymbol{X} = \begin{bmatrix} ab & b\underset{1\times a}{\mathbf{1}} & a\underset{1\times b}{\mathbf{1}} \\ b\underset{a\times 1}{\mathbf{1}} & b\boldsymbol{I}_a & \underset{a\times b}{\mathbf{1}} \\ a\underset{b\times 1}{\mathbf{1}} & \underset{b\times a}{\mathbf{1}} & a\boldsymbol{I}_b \end{bmatrix}.$$

We know $\boldsymbol{G}$ is a generalized inverse of $\boldsymbol{A}$ if $\boldsymbol{AGA} = \boldsymbol{A}$. So

$$\begin{aligned}
\boldsymbol{X}^T\boldsymbol{X}\boldsymbol{G}\boldsymbol{X}^T\boldsymbol{X} &= \boldsymbol{X}^T\boldsymbol{X}\begin{bmatrix} 1/(ab) & \underset{1\times a}{\mathbf{0}} & \underset{1\times b}{\mathbf{0}} \\ -1/(ab)\underset{a\times 1}{\mathbf{1}} & 1/b\boldsymbol{I}_a & \underset{a\times b}{\mathbf{0}} \\ -1/(ab)\underset{b\times 1}{\mathbf{1}} & \underset{b\times a}{\mathbf{0}} & 1/a\boldsymbol{I}_b \end{bmatrix}\begin{bmatrix} ab & b\underset{1\times a}{\mathbf{1}} & a\underset{1\times b}{\mathbf{1}} \\ b\underset{a\times 1}{\mathbf{1}} & b\boldsymbol{I}_a & \underset{a\times b}{\mathbf{1}} \\ a\underset{b\times 1}{\mathbf{1}} & \underset{b\times a}{\mathbf{1}} & a\boldsymbol{I}_b \end{bmatrix} \\
&= \boldsymbol{X}^T\boldsymbol{X}\begin{bmatrix} \frac{ab}{ab}+0+0 & \frac{b}{ab}\underset{1\times a}{\mathbf{1}} + \underset{1\times a}{\mathbf{0}} + \underset{1\times a}{\mathbf{0}} & \frac{a}{ab}\underset{1\times b}{\mathbf{1}} + \underset{1\times b}{\mathbf{0}} + \underset{1\times b}{\mathbf{0}} \\ \frac{-ab}{ab}\underset{a\times 1}{\mathbf{1}} + \frac{b}{b}\underset{a\times 1}{\mathbf{1}} + \underset{a\times 1}{\mathbf{0}} & \frac{-b}{ab}\underset{a\times a}{\mathbf{1}} + \frac{b}{b}\boldsymbol{I}_a + \underset{a\times a}{\mathbf{0}} & \frac{-a}{ab}\underset{a\times b}{\mathbf{1}} + \frac{1}{b}\underset{a\times b}{\mathbf{1}} + \underset{a\times b}{\mathbf{0}} \\ \frac{-ab}{ab}\underset{b\times 1}{\mathbf{1}} + \underset{b\times 1}{\mathbf{0}} + \frac{a}{a}\underset{b\times 1}{\mathbf{1}} & \frac{-b}{ab}\underset{b\times a}{\mathbf{1}} + \underset{b\times a}{\mathbf{0}} + \frac{1}{a}\underset{b\times a}{\mathbf{1}} & \frac{-a}{ab}\underset{b\times b}{\mathbf{1}} + \underset{b\times b}{\mathbf{0}} + \frac{a}{a}\boldsymbol{I}_a \end{bmatrix} \\
&= \begin{bmatrix} ab & b\underset{1\times a}{\mathbf{1}} & a\underset{1\times b}{\mathbf{1}} \\ b\underset{a\times 1}{\mathbf{1}} & b\boldsymbol{I}_a & \underset{a\times b}{\mathbf{1}} \\ a\underset{b\times 1}{\mathbf{1}} & \underset{b\times a}{\mathbf{1}} & a\boldsymbol{I}_b \end{bmatrix}\begin{bmatrix} 1 & 1/a\underset{1\times a}{\mathbf{1}} & 1/b\underset{1\times b}{\mathbf{1}} \\ \underset{a\times 1}{\mathbf{0}} & -1/a\underset{a\times a}{\mathbf{1}} + \boldsymbol{I}_a & \underset{a\times b}{\mathbf{0}} \\ \underset{b\times 1}{\mathbf{0}} & \underset{b\times a}{\mathbf{0}} & -1/b\underset{b\times b}{\mathbf{1}} + \boldsymbol{I}_b \end{bmatrix} \\
&= \begin{bmatrix} ab+0+0 & \frac{ab}{a}\underset{1\times a}{\mathbf{1}} + \frac{-ab}{a}\underset{1\times a}{\mathbf{1}} + b\underset{1\times a}{\mathbf{1}} + \underset{1\times a}{\mathbf{0}} & \frac{ab}{b}\underset{1\times b}{\mathbf{1}} + \underset{1\times b}{\mathbf{0}} + \frac{-ab}{b}\underset{1\times b}{\mathbf{1}} + a\underset{1\times b}{\mathbf{1}} \\ b\underset{a\times 1}{\mathbf{1}} + \underset{a\times 1}{\mathbf{0}} + \underset{a\times 1}{\mathbf{0}} & \frac{b}{a}\underset{a\times a}{\mathbf{1}} + \frac{-b}{a}\underset{a\times a}{\mathbf{1}} + b\boldsymbol{I}_a + \underset{a\times a}{\mathbf{0}} & \frac{b}{b}\underset{a\times b}{\mathbf{1}} + \underset{a\times b}{\mathbf{0}} + \frac{-b}{b}\underset{a\times b}{\mathbf{1}} + \underset{a\times b}{\mathbf{1}} \\ a\underset{b\times 1}{\mathbf{1}} + \underset{b\times 1}{\mathbf{0}} + \underset{b\times 1}{\mathbf{0}} & \frac{a}{a}\underset{b\times a}{\mathbf{1}} + \frac{-a}{a}\underset{b\times a}{\mathbf{1}} + \underset{b\times a}{\mathbf{1}} + \underset{b\times a}{\mathbf{0}} & \frac{a}{b}\underset{b\times b}{\mathbf{1}} + \underset{b\times b}{\mathbf{0}} + \frac{-a}{b}\underset{b\times b}{\mathbf{1}} + a\boldsymbol{I}_b \end{bmatrix} \\
&= \begin{bmatrix} ab & b\underset{1\times a}{\mathbf{1}} & a\underset{1\times b}{\mathbf{1}} \\ b\underset{a\times 1}{\mathbf{1}} & b\boldsymbol{I}_a & \underset{a\times b}{\mathbf{1}} \\ a\underset{b\times 1}{\mathbf{1}} & \underset{b\times a}{\mathbf{1}} & a\boldsymbol{I}_b \end{bmatrix} = \boldsymbol{X}^T\boldsymbol{X}\,.
\end{aligned}$$

(c) We need to show that $\boldsymbol{u}_j \in \mathcal{N}(\boldsymbol{X})$ and u_j $j = 1, 2$ are linearly independent. First check that $\boldsymbol{X}\boldsymbol{u}_1 = \mathbf{0}$ and $\boldsymbol{X}\boldsymbol{u}_2 = \mathbf{0}$, so $\boldsymbol{u}_1 \in \mathcal{N}(\boldsymbol{X})$ and $\boldsymbol{u}_2 \in \mathcal{N}(\boldsymbol{X})$. Clearly, $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$ are linearly independent since obtaining a 0 in the first element of $c_1\boldsymbol{u}_1 + c_2\boldsymbol{u}_2$ requires $c_1 = -c_2$. However, this will not produce a 0 in any other element of the sum unless $c_1 = -c_2 = 0$.

Because the dimension of $\mathcal{N}(\boldsymbol{X})$ is 2 and we have 2 linearly independent vectors in $\mathcal{N}(\boldsymbol{X})$, $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$ together form a basis for $\mathcal{N}(\boldsymbol{X})$.

2. (a) Let $\underset{(8)\times 1}{\boldsymbol{\beta}} = (\mu, \alpha_1, \alpha_2, \beta_{11}, \beta_{12}, \beta_{21}, \beta_{22}, \beta_{23})^T$. And let $\underset{(10)\times 1}{\boldsymbol{y}} = (100, 80, 80, 80, 110, 90, 100, 140, 110, 15$

The data is as follows:

$$\underset{(10)\times(8)}{\boldsymbol{X}} = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

This matrix has rank 5.

(b) The normal equations are $\boldsymbol{X}^T\boldsymbol{X}\boldsymbol{\beta} = \boldsymbol{X}^T\boldsymbol{y}$. The solutions to the normal equations are given by $\hat{\boldsymbol{\beta}} = (\boldsymbol{X}^T\boldsymbol{X})^-\boldsymbol{X}^T\boldsymbol{y} + (\boldsymbol{I}_8 - (\boldsymbol{X}^T\boldsymbol{X})^-(\boldsymbol{X}^T\boldsymbol{X}))\boldsymbol{z}$ for any $\boldsymbol{z} \in \mathbb{R}^8$, where $(\boldsymbol{X}^T\boldsymbol{X})^-\boldsymbol{X}^T$ is any generalized inverse of $\boldsymbol{X}$. Letting $\boldsymbol{G}$ the default output from the ginv function in the MASS package from R and fixing z to be zero, we end up with a particular solution of

$$\hat{\boldsymbol{\beta}} = \begin{bmatrix} 59.66 \\ 16.90 \\ 42.76 \\ 13.45 \\ 3.45 \\ -2.41 \\ 17.59 \\ 27.59 \end{bmatrix},$$

but an infinite number of solutions can be found using the equations above.

(c) The vectors below form a basis for $\mathcal{N}(\boldsymbol{X})$ $(dim(\mathcal{N}(\boldsymbol{X})) = 8 - 5 = 3)$;

$$\begin{aligned} \boldsymbol{u}_1 &= [1, -1, -1, 0, 0, 0, 0, 0]^T \\ \boldsymbol{u}_2 &= [-1, 0, 0, 1, 1, 1, 1, 1]^T \\ \boldsymbol{u}_3 &= [0, 1, -1, -1, -1, 1, 1, 1]^T \end{aligned}$$

(d) Since the matrix $\boldsymbol{X}$ has rank 5, any 5 linearly independent rows, and every linear combination of those rows, form estimable functions. We'll take the odd numbered rows, and we end up with these 5 linearly independent functions:

$$\begin{aligned} \boldsymbol{\lambda}_1\boldsymbol{\beta} &= \mu + \alpha_1 + \beta_{11} \\ \boldsymbol{\lambda}_2\boldsymbol{\beta} &= \mu + \alpha_1 + \beta_{12} \\ \boldsymbol{\lambda}_3\boldsymbol{\beta} &= \mu + \alpha_2 + \beta_{21} \\ \boldsymbol{\lambda}_4\boldsymbol{\beta} &= \mu + \alpha_2 + \beta_{22} \\ \boldsymbol{\lambda}_5\boldsymbol{\beta} &= \mu + \alpha_2 + \beta_{23} \end{aligned}$$

(e) The function $\alpha_1 - \alpha_2$ is not estimable because it cannot be formed as a linear combination of the above equations. Or we can show that the corresponding $\boldsymbol{\lambda} = [0, 1, -1, 0, 0, 0, 0, 0]^T$ is not orthogonal to $\boldsymbol{u}_3$ in (c).

(f) $\boldsymbol{\beta}_2$ and $\boldsymbol{\beta}_5$ have the same vector of predicted values.

(g) For both of these vectors, $\boldsymbol{X}\boldsymbol{\beta} = [100, 100, 90, 90, 110, 110, 120, 120, 120, 120]^T$

3. Let Model1 be the model with β_0, β_1 and β_2 and Model2 the model with β_0 and β_1 assuming that $\beta_2 = 0$. Under Model1,

$$\underset{(3)\times(3)}{\boldsymbol{X}} = \begin{bmatrix} 1 & -1 & 1 \\ 1 & 0 & -2 \\ 1 & 1 & 1 \end{bmatrix}.$$

$\boldsymbol{X}$ is nonsingular, so the solution for the NEs is;

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{y} = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 5 \end{bmatrix}^{-1} \boldsymbol{X}^T\boldsymbol{y} = \begin{bmatrix} 1/3 & 1/3 & 1/3 \\ -1/2 & 0 & 1/2 \\ 1/5 & -2/5 & 1/5 \end{bmatrix} \boldsymbol{y}$$

Note that the columns of $\boldsymbol{X}$ are orthogonal (that is, $\boldsymbol{X}$ is an orthogonal column matrix). Use $\boldsymbol{V}$ to denote the design matrix under Model2.

$$\underset{(3)\times(2)}{\boldsymbol{V}} = \begin{bmatrix} 1 & -1 \\ 1 & 0 \\ 1 & 1 \end{bmatrix}.$$

The LSE under Model2 is

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{V}^T\boldsymbol{V})^{-1}\boldsymbol{V}^T\boldsymbol{y} = \begin{bmatrix} 1/3 & 1/3 & 1/3 \\ -1/2 & 0 & 1/2 \end{bmatrix} \boldsymbol{y}$$

So the least squared estimates of β_0 and β_1 in Model2 are the same as the exact solutions (and LSEs) for β_0 and β_1 in Model1 (this is a consequence of orthogonal columns in $\boldsymbol{X}$).