Homework Assignment #1

1. (5 points) For example 2 in lecture notes 1, what is the rank of the design matrix X? Assume $n > p$. You may have to describe cases. For cases where the rank is $< p$, describe what this means in "practical" terms – what does the scatterplot of the data look like?

Solution: Recall the model is

$$Y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \epsilon_i$$

where Y_i and x_i are the blood pressure and weight of subject i, respectively. The design matrix is

$$\mathbf{X} = \begin{bmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ \vdots & \vdots & \vdots \\ 1 & x_n & x_n^2 \end{bmatrix} = \begin{bmatrix} \mathbf{1}_n \ (x_i)_n \ (x_i^2)_n \end{bmatrix}$$

With no constraints on the x's, then $\mathbf{X}$ could have rank 1, 2 or 3. If x_i are all equal, then all columns are the proportional to the first column $\mathbf{1}_n$ so that rank($\mathbf{X}$) = 1. If there are only two different values of x_i, a row reduced echelon form has two pivots so that rank($\mathbf{X}$) = 2. Otherwise, $\mathbf{X}$ has rank($\mathbf{X}$) = 3 (check this by row reduction). In "practical" terms, rank=1 means we would be trying to fit a quadratic term to data with one possible value of a covariate and there will be an infinite number of solution; rank=2 means we would be trying to fit a quadratic term to data with two possible values of a covariate and there will be an infinite number of solutions.

2. (5 points) For example 4 in lecture notes 1, what is the rank of the design matrix $\mathbf{X}$? Assume $n > p$. You may have to describe cases. For cases where the rank is $< p$, describe what this means in "practical" terms – what does the scatterplot of the data look like?

Solution: The Inverse Square Law states that the force of gravity F between two bodies a distance D apart is given by

$$F = \frac{c}{D^\beta}$$

Taking logarithms on both sides yields the linear model

$$\log F = \log c - \beta \log D \Leftrightarrow Y_i = \beta_0 + \beta_1 x_i$$

with $Y_i = \log F$, $\beta_0 = \log c$ and $x_i = -\log D$. Hence if we have a sample $(Y_1, x_i), \ldots, (Y_n, x_n)$ where $n > 2$ then the design matrix is $\mathbf{X} =$

$$\begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix}$$

The rank of $\mathbf{X}$ could be 1 or 2. Using similar reasoning as problem 1, $\text{rank}(\mathbf{X}) = 1$ implies that $x_1 = \ldots = x_n$, that is, all the bodies in the sample are the same distance apart. The scatterplot in this case would look like the left side of figure 1.

3. (3 points) For example 5 in lecture notes 1, what is the rank of the design matrix $\mathbf{X}$? Assume $J \geq 2$.

Solution: Recall the design matrix $\mathbf{X}$ is

$$\begin{bmatrix} 1 & 1 & 0 \\ \vdots & \vdots & \vdots \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ \vdots & \vdots & \vdots \\ 1 & 0 & 1 \end{bmatrix}$$

Note that the first column of $\mathbf{X}$ is the sum of the second and third columns. The column space of $\mathbf{X}$ is

$$\mathcal{R}(\mathbf{X}) = \text{span}\{ \begin{pmatrix} 1 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \}$$

hence $\text{rank}\mathbf{X} = 2$.

4. (3 points) For example 6 in lecture notes 1, what is the rank of the design matrix $\mathbf{X}$? Assume $K \geq 2$.

Solution: Recall the design matrix $\mathbf{X}$ is

$$\begin{bmatrix} 1 & 1 & 0 & 1 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \mathbf{1}_K & \mathbf{1}_K & \mathbf{0}_K & \mathbf{1}_K & \mathbf{0}_K \\ \mathbf{1}_K & \mathbf{1}_K & \mathbf{0}_K & \mathbf{0}_K & \mathbf{1}_K \\ \mathbf{1}_K & \mathbf{0}_K & \mathbf{1}_K & \mathbf{1}_K & \mathbf{0}_K \\ \mathbf{1}_K & \mathbf{0}_K & \mathbf{1}_K & \mathbf{0}_K & \mathbf{1}_K \end{bmatrix}$$

Performing elementary row operations, we can reduce the matrix to the echelon form

$$\begin{bmatrix} 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ \vdots & & & & \vdots \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \text{ thus } \operatorname{rank}(X) = 3.$$

5. (5 points) Repeat the exercise on page 10 of lecture notes 2 for the dataset $\{(x, y) = (1, 0), (1, 2), (0, 0)\}$.

Solution: Here the design matrix is $X = (1\ 1\ 0)^T$, so $\hat{\beta} = (X^T X)^{-1} X^T Y = 1$. The fitted values are $\hat{Y} = (1, 1, 0) = X$. See figure 3 for the plots. In this case, there are three data points, so we must visualize the vectors in three dimensions. The dimension of Ω is 1, while the dimension of $\Omega^\perp$ is 2. The set $\Omega^\perp$ is the set of all vectors perpendicular to a vector $X = (1\ 1\ 0)^T$, that is, $\{c(1, -1, k) : c, k \in \mathbb{R}\}$. See figure 4 for a scatterplot. If we fit the model $Y_i = \beta_0 + \beta_1 x_i + \epsilon_i$ then the design matrix is

$$X^* = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 0 \end{bmatrix}$$

which has two linearly independent columns, hence in this case, the dimension of Ω is 2, while the dimension of $\Omega^\perp$ is 1. The set $\Omega^\perp$ is the set of all vectors perpendicular to the plane $\{c(1, 1, k) : c, k \in \mathbb{R}\}$ spanned by the two column vectors of the design matrix, that is, $\{c(1, -1, 0) : c, \in \mathbb{R}\}$. Furthermore, $\hat{\beta} =$

$$\begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

So the intercept is zero and hence this model reduces to $Y_i = \beta x_i + \epsilon_i$, so the other previous answers do not change.

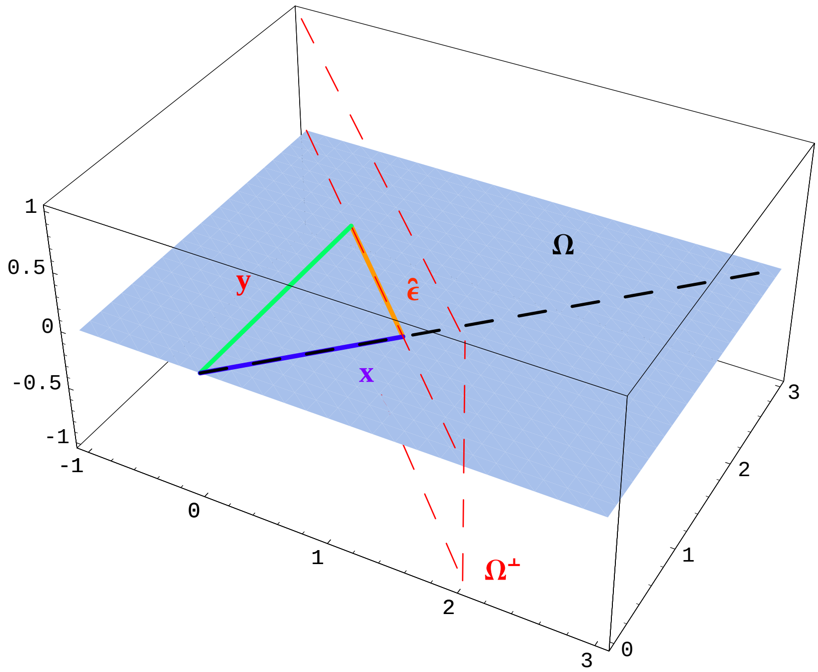

Figure 1: Visualization of problem 6

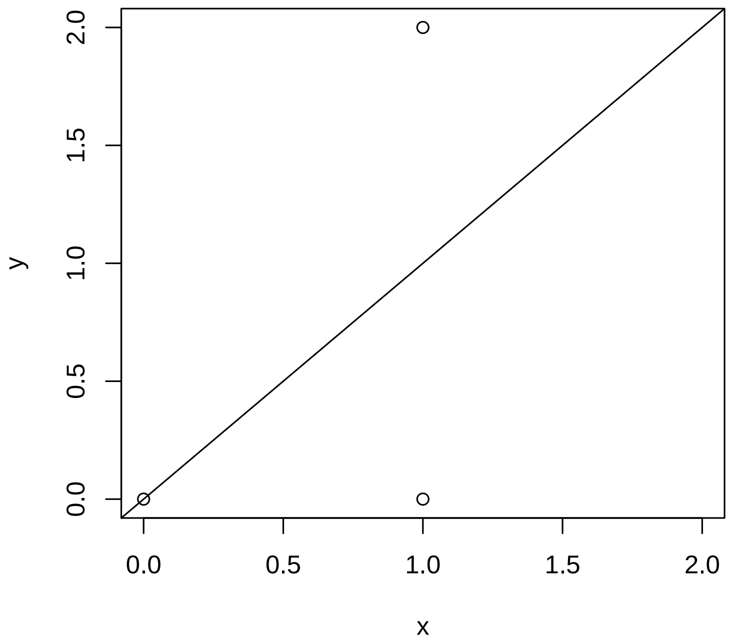

Figure 2: Scatterplot for problem 6

6. (5 points) Let $\mathbf{P} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$ and $\mathbf{Q} = \mathbf{I} - \mathbf{P}$ and assume that $\mathbf{X}'\mathbf{X}$ is non-singular. Prove: (a) $\mathbf{P}$ and $\mathbf{Q}$ are symmetric and idempotent; (b) $\mathbf{QX} = 0$ and $\mathbf{X}'\mathbf{Q} = \mathbf{0}$.

Solution: (a) Using the properties of matrix transpose,

$$[\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']' = \mathbf{X}[(\mathbf{X}'\mathbf{X})^{-1}]'\mathbf{X}' = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$$

Moreover,

$$[\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'][\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'] =$$
$$\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}[\mathbf{X}'\mathbf{X}](\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$$

proving that $\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$ is symmetric and idempotent. As for $\mathbf{Q}$, note that matrix transpose distributes over sums,

$$(\mathbf{I} - \mathbf{P})' = \mathbf{I}' - \mathbf{P}' = \mathbf{I} - \mathbf{P}$$

since $\mathbf{I}$ and $\mathbf{P}$ are symmetric. Moreover,

$$(\mathbf{I} - \mathbf{P})(\mathbf{I} - \mathbf{P}) = \mathbf{I} - \mathbf{P} - \mathbf{P} + \mathbf{P}^2 = \mathbf{I} - \mathbf{P} - \mathbf{P} + \mathbf{P} = \mathbf{Q}$$

since $\mathbf{P}$ is idempotent.

(b) By straightforward calculation,

$$\mathbf{QX} = (\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')\mathbf{X} = \mathbf{X} - \mathbf{X} = \mathbf{0}$$

Since $\mathbf{Q}$ is symmetric, take transpose on both sides of the equation above to get

$$\mathbf{X}'\mathbf{Q} = \mathbf{0}$$

7. (3 points) Let $\mathbf{J}_n$ be the $n \times n$ matrix of 1's. Show $\frac{1}{n}\mathbf{J}_n$ is idempotent.

Solution: Again, by straightfoward calculation,

$$\frac{1}{n}\mathbf{J}_n \frac{1}{n}\mathbf{J}_n = \frac{1}{n^2}\begin{bmatrix} \mathbf{1}_n'\mathbf{1}_n & \dots & \mathbf{1}_n'\mathbf{1}_n \\ \vdots & \ddots & \vdots \\ \mathbf{1}_n'\mathbf{1}_n & \dots & \mathbf{1}_n'\mathbf{1}_n \end{bmatrix} = \frac{1}{n^2}\begin{bmatrix} n & \dots & n \\ \vdots & \ddots & \vdots \\ n & \dots & n \end{bmatrix} = \frac{1}{n}\mathbf{J}_n$$

8. (3 points) Let A be a symmetric $n \times n$ matrix, let λ_i be its eigenvalues and let s be a nonzero integer. Prove $\text{trace}(A^s) = \sum_{i=1}^n \lambda_i^s$

Solution: Since A is symmetric, applying the Spectral Theorem gives

$$A = T\Lambda T'$$

where T is an orthonormal matrix and $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$. Thus, we have

$$\begin{aligned} A^s &= (T\Lambda T')(T\Lambda T')\cdots(T\Lambda T') \\ &= T\Lambda(T'T)\Lambda T'\cdots T\Lambda T' \\ &= T\Lambda I\Lambda T'\cdots T\Lambda T' \\ &= T\Lambda^2 T'\cdots T\Lambda T' \\ &\vdots \\ &= T\Lambda^s T' \end{aligned}$$

Hence

$$\begin{aligned} \text{trace}(A^s) &= \text{trace}(T\Lambda^s T') \\ &= \text{trace}(T'T\Lambda^s) \\ &= \text{trace}(\Lambda^s) \\ &= \text{trace}(\text{diag}(\lambda_1^s, \dots, \lambda_n^s)) \\ &= \sum_{i=1}^n \lambda_i^s. \end{aligned}$$