Homework 2

Psychology 312

Instructions. Feel free to email me for hints if you get stumped.

1. Suppose you have a vector $\boldsymbol{x}$, and that $\boldsymbol{x}$ has at least one nonzero element. Define $\boldsymbol{P_x} = \boldsymbol{x}(\boldsymbol{x'x})^{-1}\boldsymbol{x'}$. Note that $\boldsymbol{x'x}$ is a scalar, so $\boldsymbol{P_x}$ may also be written as

$$\boldsymbol{P_x} = \frac{\boldsymbol{xx'}}{\boldsymbol{x'x}} = \left(\frac{1}{\boldsymbol{x'x}}\right)\boldsymbol{xx'}$$

$\boldsymbol{P_x}$ is known as the "orthogonal projector for $\boldsymbol{x}$." Define $\boldsymbol{Q_x} = \boldsymbol{I} - \boldsymbol{P_x}$. Given the above definitions, answer the following:

(a) Prove that $\boldsymbol{P_x}$ is idempotent.

(b) Prove that $\boldsymbol{P_x}$ is symmetric.

(c) Now consider any other non-null vector $\boldsymbol{y}$. Prove that $\boldsymbol{a} = \boldsymbol{P_x y}$ is collinear with $\boldsymbol{x}$ i.e., that $\boldsymbol{P_x y}$ can be written in the form $c\boldsymbol{x}$ for some scalar c.

(d) Prove that $\boldsymbol{Q_x} = \boldsymbol{I} - \boldsymbol{P_x}$ is also symmetric and idempotent.

(e) Consider the vector $\boldsymbol{b} = \boldsymbol{Q_x y}$. Prove that $\boldsymbol{b}$ is orthogonal to $\boldsymbol{a} = \boldsymbol{P_x y}$, i.e., $\boldsymbol{a'b} = \boldsymbol{b'a} = 0$, and that $\boldsymbol{b}$ is also orthogonal to $\boldsymbol{x}$.

(f) Prove that $\boldsymbol{a} + \boldsymbol{b} = \boldsymbol{y}$, thus showing that $\boldsymbol{P_x}$ and $\boldsymbol{Q_x}$ can be used to break a vector $\boldsymbol{y}$ into two component vectors, one orthogonal to $\boldsymbol{x}$, one collinear with $\boldsymbol{x}$.

(g) Consider a data vector $\boldsymbol{x}$, and $\boldsymbol{1}$, a conformable "summing vector" of ones . Explain why $\boldsymbol{x}$ is in deviation score form if and only if $\boldsymbol{1'x} = 0$.

(h) Consider a data matrix $\boldsymbol{X}$. Prove that if the columns of $\boldsymbol{X}$ are in deviation score form, any linear combination of the variables in $\boldsymbol{X}$, i.e., $\boldsymbol{y} = \boldsymbol{Xb}$, will also be in deviation score form.

(i) Consider $\boldsymbol{Q_1}$. Explain succinctly (in terms of what you discovered from problems a–h) why $\boldsymbol{Q_1}$ carries any vector of scores into deviation score form.

(j) Using R, create a vector $\boldsymbol{x}$ such that $\boldsymbol{x'} = [2, -3, 2, 2, 0]$. Using this $\boldsymbol{x}$, create $\boldsymbol{P_x}$ and $\boldsymbol{Q_x}$, create another vector $\boldsymbol{y}$, such that $\boldsymbol{y'} = [1, 2, 3, 4, 5]$, and then demonstrate numerically all the properties proven in problems a–h.