STATS 762 Assignment 4 Extra Question

1. (a) Consider each of the following sets of vectors in R^4 where a, b and c represent any real numbers.

   All vectors of the form:

$$\text{(i)} \begin{bmatrix} a \\ b \\ a+b \\ 1 \end{bmatrix} \qquad \text{(ii)} \begin{bmatrix} a \\ b \\ a+b \\ 0 \end{bmatrix} \qquad \text{(iii)} \begin{bmatrix} a \\ b \\ c \\ a+b-2c \end{bmatrix}$$

   In each case, determine whether the set of vectors represents a subspace of R^4 (you need to justify your answer).

   (b) Consider the subspace of R^4 defined as all vectors of the form

$$\begin{bmatrix} a \\ a+b \\ a-b \\ -2b \end{bmatrix}$$

   What is the dimension of this subspace? Give a set of vectors that form a basis for this subspace.

   (c) Consider a random vector $X \sim \text{N}_3(\mu_X,\ \Sigma_X)$:

$$X = \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix}, \quad \mu_X = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix}, \quad \Sigma_X = \begin{bmatrix} 2 & 1 & 2 \\ 1 & 2 & -1 \\ 2 & -1 & 3 \end{bmatrix}$$

   i. Find a matrix A such that if we let $W = AX$ then:

$$\mu_W = \begin{bmatrix} \mu_1 - \mu_2 \\ \mu_2 + 2\mu_3 \\ 3\mu_1 - 2\mu_2 - \mu_3 \end{bmatrix}$$

   ii. Find the covariance matrix for W, i.e. Σ_W.

   (d) The Best Linear Unbiased Estimate (BLUE) of a parameter is defined as the estimator that has the smallest variance among the set of all unbiased linear estimators. Note a linear estimator is simply any estimator which is a linear combination of the data, i.e. can be written as $\mathbf{a}^t\mathbf{Y}$.

   Given the linear model,

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon} \quad \text{where} \quad \boldsymbol{\beta} \sim \text{N}(\mathbf{0}, \sigma^2\mathbf{I}_n)$$

   and the least squares estimates of $\boldsymbol{\beta}$

$$\hat{\boldsymbol{\beta}} = \left(\mathbf{X}^t\mathbf{X}\right)^{-1}\mathbf{X}^t\mathbf{Y},$$

   consider estimating a linear combination of the parameters $\mathbf{a}^t\boldsymbol{\beta}$. We can show that among all possible linear unbiased estimators of $\mathbf{a}^t\boldsymbol{\beta}$, $\mathbf{a}^t\hat{\boldsymbol{\beta}}$ has the smallest variance using the following steps.

i. First, consider any linear estimator $\mathbf{b}^t\mathbf{Y}$ and show:

A. $\mathbf{b}^t\mathbf{Y}$ is unbiased for $\mathbf{a}^t\boldsymbol{\beta}$ if $\mathbf{b}^t\mathbf{X} = \mathbf{a}^t$.

B. $\text{Var}(\mathbf{b}^t\mathbf{Y}) = \sigma^2\mathbf{b}^t\mathbf{b}$.

ii. Now consider the estimator based on the least squares estimates $\mathbf{a}^t\boldsymbol{\beta}$. Show the following:

A. $\mathbf{a}^t\hat{\boldsymbol{\beta}}$ is unbiased for $\mathbf{a}^t\boldsymbol{\beta}$,

B. $\text{Var}(\mathbf{a}^t\hat{\boldsymbol{\beta}}) = \sigma^2\,\mathbf{a}^t(\mathbf{X}^t\mathbf{X})^{-1}\mathbf{a}$

iii. Now set $\mathbf{b}^* = \mathbf{X}(\mathbf{X}^t\mathbf{X})^{-1}\mathbf{a}$. For any $\mathbf{b}$

$$(\mathbf{b} - \mathbf{b}^*)^t(\mathbf{b} - \mathbf{b}^*) \geq 0$$

and the equality holds only if $\mathbf{b} = \mathbf{b}^*$. Show that

$$\sigma^2(\mathbf{b} - \mathbf{b}^*)^t(\mathbf{b} - \mathbf{b}^*) = \text{Var}(\mathbf{b}^t\mathbf{Y}) - \text{Var}\left(\mathbf{b}^{*t}\mathbf{Y}\right)$$

which means $\text{Var}\,(\mathbf{b}^t\mathbf{Y}) \geq \text{Var}\,(\mathbf{b}^{*t}\mathbf{Y})$ and the equality holds only if $\mathbf{b} = \mathbf{b}^*$.

Hint: expand the left hand side and simplify using the results in (a) and (b).