Stats Modelling Example Sheet 1

In all the questions that follow, X is an n by p design matrix with full column rank and H is the orthogonal projection onto the column space of X. Also, let X_0 be the matrix formed from the first $p_0 < p$ columns of X and let H_0 be the orthogonal projection on to the column space of X_0. The vector $Y \in \mathbb{R}^n$ will be a vector of responses and we will define $\hat{\boldsymbol{\beta}} := (X^TX)^{-1}X^TY$, $\hat{\boldsymbol{\beta}}_0 := (X_0^TX_0)^{-1}X_0^TY$ and $\tilde{\sigma}^2 := ||(I-H)Y||^2/(n-p)$. By normal linear model, we mean the model $Y = X\beta + \epsilon$, $\epsilon|X \sim N_n(0, \sigma^2 I)$.

1 Question 1

Show that

$$||(H-H_0)Y||^2 = ||(I-H_0)Y||^2 - ||(I-H)Y||^2 = ||HY||^2 - ||H_oY||^2$$

We know that since H and H_0 are projection matrices, they have the properties: $H_0H = HH_0 = H_0$ and $H^T = H$, $H_0^T = H_0$

$$\begin{aligned}
||(H-H_0)Y||^2 &= Y^T(H-H_0)^T(H-H_0)Y && (1)\\
&= Y^T[H^TH - H_0^TH - H^TH_0 + H_0^TH_0]Y && (2)\\
&= Y^T[H - H_0 - H_0 + H_0]Y && (3)\\
&= Y^T[H - H_0]Y && (4)
\end{aligned}$$

$$\begin{aligned}
||(I-H_0)Y||^2 - ||(I-H)Y||^2 &= ((I-H_0)Y - (I-H)Y)^T((I-H_0)Y + (I-H)Y) && (5)\\
&= Y^T(H-H_0)^T(2I - H - H_0)Y && (6)\\
&= Y^T(2H^T - 2H_0^T - H^TH + H_0^TH - H^TH_0 + H_0^TH_0)Y && (7)\\
&= Y^T[H - H_0]Y && (8)
\end{aligned}$$

$$\begin{aligned}
||HY||^2 - ||H_0Y||^2 &= ((H-H_0)Y)^T((H+H_0)Y) && (9)\\
&= Y^T(H^TH - H_0^TH + H^TH_0 - H_0^TH_0)Y && (10)\\
&= Y^T(H - H_0)Y && (11)
\end{aligned}$$