Show that

$$\sum (X_i - \bar{X})^2 = X'AX \text{ where } A = I_n - \frac{1}{n}J_n$$

Solution. Starting on the right hand side, we get

$$\begin{aligned}
X'AX &= [X_1, X_2, \dots, X_n] \begin{bmatrix} 1-\frac{1}{n} & -\frac{1}{n} & \cdots & -\frac{1}{n} \\ -\frac{1}{n} & 1-\frac{1}{n} & \cdots & -\frac{1}{n} \\ \vdots & \vdots & \ddots & \vdots \\ -\frac{1}{n} & -\frac{1}{n} & \cdots & 1-\frac{1}{n} \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix} \\
&= \begin{bmatrix} X_1 - \frac{1}{n}\sum X_i & X_2 - \frac{1}{n}\sum X_i & \dots & X_n - \frac{1}{n}\sum X_i \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix} \\
&= \begin{bmatrix} X_1 - \bar{X} & X_2 - \bar{X} & \dots & X_n - \bar{X} \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix} \\
&= \left[X_1^2 - X_1\bar{X} + X_2^2 - X_2\bar{X} + \cdots + X_n^2 - X_n\bar{X}\right] \\
\text{(adding zero this step)} &= \left[X_1^2 - X_1\bar{X} - X_1\bar{X} + X_1\bar{X} + X_2^2 - X_2\bar{X} - X_2\bar{X} + X_2\bar{X} + \cdots + X_n^2 - X_n\bar{X} - X_n\bar{X} + X_n\bar{X}\right] \\
&= \left[X_1^2 - 2X_1\bar{X} + X_1\bar{X} + X_2^2 - 2X_2\bar{X} + X_2\bar{X} + \cdots + X_n^2 - 2X_n\bar{X} + X_n\bar{X}\right] \\
&= \sum X_i^2 - 2\bar{X}\sum X_i + \bar{X}\sum X_i \\
&= \sum X_i^2 - 2\bar{X}\sum X_i + n\bar{X}\bar{X} \\
&= \sum (X_i^2 - 2\bar{X}X_i + \bar{X}^2) \\
&= \sum (X_i - \bar{X})^2
\end{aligned}$$

□

Table 1: Significant and Important Variable Prediction

Model	X1	X2	X3
Random Survival Forest	.990	.316	.710
Cox Proportional Hazards	.964	.062	.346

- Used $\alpha = .05$ for significance level for Cox model
- For each variable, we took the mean number of times that the said variable had a higher importance score than all three noise variables.