

# Problem Set 6–STAT 7200

1. We discussed in class how  $\hat{Y} = P\mathbb{Y} = \mathbb{X}\hat{B}$  is the projection of  $\mathbb{Y}$  on the column space  $\mathcal{V}$  of  $\mathbb{X}$ :

$$\mathcal{V} = \{\mathbb{X}A \mid A \text{ is a } (q+1) \times p \text{ matrix}\}.$$

Use this fact to show that  $\mathbb{X}\hat{B}$  is also the solution to the following least squares problem:

$$\min_{V \in \mathcal{V}} \text{tr}(\Omega(\mathbb{Y} - V)^T(\mathbb{Y} - V)),$$

where  $\Omega$  is an arbitrary  $p \times p$  positive definite matrix.

2. For this question, we will use the `sheishu` dataset in the package `ACSWR`:

```
> library(ACSWR)
> data(sheishu)
```

Answer the following questions:

- Find the least squares estimate  $\hat{B}$  for the regression of (Taste, Odor) on the other eight covariates, and test for overall significance.
- Test the significance of (Alcohol, Formyl\_nitrogen) adjusted for the other covariates.
- Test the significance of (Sake\_meter, Direct\_reducing\_sugar, Total\_sugar) adjusted for the other covariates.
- Test the significance of (pH, Acidity\_1, Acidity\_2) adjusted for the other covariates.
- Using all subset selection, find the model with the lowest AIC.