# SDS 383D Peer Review for Jesse Miller

February 20, 2017

## Ex1_ozone_se.r

For coding style, I suggest using more spaces, such as around operators and after commas. For example, instead of

```
sigmahat1 <- (1/(nrow(x)-ncol(x)))*sum((y-yhat)^2)
```

consider:

```
sigmahat1 <- (1 / (nrow(x) - ncol(x))) * sum((y - yhat)^2)
```

## Ex1_bootstrap_A.r

Good comments on the function to compute the bootstrap. From line 38 to 40, the variables betahat, yhat and residuals seem redundant for this function.

## Ex1_bootstrap_B.r

You generate $10,000$ samples for the bi-variate Normal distribution. I think that is too much data for two dimensions. You can also try a smaller number of samples (e.g 1000) to see how well the bootstrap recovers the true covariance.

In line 34, instead of

```
sigma_hat_boot <- sigma_hat_boot + (t(mvn_samp_temp - mu_hat_temp)
%*% (mvn_samp_temp - mu_hat_temp)/N)
```

I suggest:

```
sigma_hat_boot <- sigma_hat_boot + (crossprod(mvn_samp_temp - mu_hat_temp)/N)
```

This is shorter and surprisingly faster. The former takes 8.7 second while the latter takes 5.8 second. In general, optimizing what is inside a big loop will make a big difference.

## 03_CrossValidationB.r

You use sin(100x) as the wiggly function. I think it is too wiggly that after adding Normal error, it looks a bit like white noise.