Due Friday, October 19th, 2007

1. **Singular value decomposition and eigenvalues:**

(a) Generate a 100×50 matrix X with i.i.d. $\mathcal{N}(0, 1)$ entries. Compute the singular value decomposition of X using `svd` in R. Report the diagonal entries in the D matrix.

(b) Write the SVD of X as $X = UDV'$. In this case, the sample covariance matrix

$$A = X'X/100\,,$$

show that $A = VD^2V'/100$.

(c) From our data generation procedure, we know the covariance matrix for X is identity I_{50}, whose eigenvalue is 1 with multiplicity 50. What are the eigenvalues for the sample covariance matrix A?

(d) Generate a larger 1000×500 matrix Y with i.i.d. $\mathcal{N}(0, 1)$ entries. Make a QQ-plot for the eigenvalues for the sample covariance matrix of Y (use `svd` to find them) against those of A in part (c). What phenomenon do you see? What does it mean?

2. **Robust regression method:**

(a) Read the help file for `rlm` in R and answer the following questions:

i. What are the default initial values for the coefficients?

ii. What is the default weight for the observations?

iii. What is the default fitting method for the function?

(b) Generate a data set in R as the following:

```
> x <- 1:50
> y <- 4 + 2 * x + rnorm(50, mean = 0, sd = 0.8)
> y[35:40] <- 7 + x[35:40]^2
```

Fit a linear regression of y on x using least square estimation.

(c) Fit a robust linear regression of y on x using the default method in `rlm`. Make a plot of your fit and compare with your results in part (b).