Matt Dickenson `mcd31`
STA561/CS571 — Fall 2013 **Homework 2** Due: 16 September, 2013

Homework Notes: I did not work with anyone else on this homework or refer to resources other than the course notes, textbook, and course Piazza page.

Problem 1

A Using the normal equation $\hat{\beta} = (X^T X)^{-1}(X^T Y)$, we can calculate $\hat{\beta}$ in $\mathcal{R}$ with the following program (after loading the data as `lindata`):

Listing 1: R Code for 1A

```r
X = as.matrix(lindata[1:1000 , 1:2])
Y = as.matrix(lindata[1:1000 , 3])
X.prime.X = t(X) %*% X
X.prime.X.inverse = solve(X.prime.X)
X.prime.Y = t(X) %*% Y
beta.hat = X.prime.X.inverse %*% X.prime.Y
beta.hat #=> 3.0117879, 0.7832938
```

Letting β_i correspond to X_i, $\hat{\beta}_1 = 3.01, \hat{\beta}_2 = 0.78$.

B We can also estimate β using online stochastic gradient descent using $\mathcal{R}$'s `optim` function for minimization:

Listing 2: R Code for 1B

```r
linreg = function(X, Y){
  RSS = function(b, x, y){
    residuals = y - (x %*% b )
    sq.resid = residuals^2
    rss = sum(sq.resid)
    return(rss)
  }
  results = optim(rep(0, ncol(X)), RSS,
    hessian=TRUE, method="BFGS", x=X, y=Y)
  list(beta=results$par,
    vcov=solve(results$hessian),
    converged=results$convergence==0)
}
model = linreg(X, Y)
model$beta #=> 3.0117879, 0.7832938
```

Using this method we reach the same results as above: $\hat{\beta}_1 = 3.01, \hat{\beta}_2 = 0.78$.

C

D

Problem 2