Lecture 10 Code: Least Squares Using Matrices

Using the menu pricing data set, the code below demonstrates how to manually compute the least squares estimates for the parameters of the regression model using the equation $\hat{\boldsymbol{\beta}} = (\boldsymbol{X'X})^{-1}\boldsymbol{X'Y}$. The results are then compared to the output from `lm()`. Note that it is always better to use `lm()` than to do the computations manually. I am just showing this to verify the formulas.

```
nyc <- read.csv("https://ericwfox.github.io/data/nyc.csv")

# response vector
Y <- matrix(nyc$Price, ncol=1)
head(Y, n=5)
```

```
##      [,1]
## [1,]   43
## [2,]   32
## [3,]   34
## [4,]   41
## [5,]   54
```

```
# design matrix
X <- cbind(Intercept = 1, nyc[, c("Food", "Decor", "East")])
X <- as.matrix(X)
rownames(X) <- nyc$Restaurant
X[1:5,]
```

```
##                     Intercept Food Decor East
## Daniella Ristorante         1   22    18    0
## Tello's Ristorante          1   20    19    0
## Biricchino                  1   21    13    0
## Bottino                     1   20    20    0
## Da Umberto                  1   24    19    0
```

```
# manually compute least squares estimates
betaHat <- solve(t(X) %*% X) %*% t(X) %*% Y
betaHat
```

```
##                 [,1]
## Intercept -24.026880
## Food        1.536346
## Decor       1.909373
## East        2.067013
```

```
# compare with lm()
lm1 <- lm(Price ~ Food + Decor + East, data=nyc)
coef(lm1)
```

```
## (Intercept)        Food       Decor        East
##  -24.026880    1.536346    1.909373    2.067013
```