Design Matrix Example (For Illustration)

We look at the design matrix in the context of simple linear regression. We start with the standard analysis in R and get the estimated coefficients.

```
Corn <- read.csv("C:/hess/STAT512/RNotes/Intro and R/Corn.csv")
Corn
```

```
##    Yield X
## 1     12 2
## 2     13 2
## 3     13 3
## 4     14 3
## 5     15 4
## 6     15 4
## 7     14 5
## 8     16 5
## 9     17 6
## 10    18 6
```

```
Fit <- lm(Yield ~ X, data = Corn)
summary(Fit)
```

```
##
## Call:
## lm(formula = Yield ~ X, data = Corn)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.8500 -0.3000  0.2250  0.4125  1.0000
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  10.1000     0.7973   12.67 1.42e-06 ***
## X             1.1500     0.1879    6.12 0.000283 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.8404 on 8 degrees of freedom
## Multiple R-squared:  0.824,  Adjusted R-squared:  0.802
## F-statistic: 37.45 on 1 and 8 DF,  p-value: 0.0002832
```

Now we look at the design matrix and estimate the coefficients "by hand". solve() takes the inverse of a matrix. t() transposes a matrix. %*% is used for matrix multiplication.

```
model.matrix(Fit)
```

```
##    (Intercept) X
## 1            1 2
## 2            1 2
```

```
## 3            1 3
## 4            1 3
## 5            1 4
## 6            1 4
## 7            1 5
## 8            1 5
## 9            1 6
## 10           1 6
## attr(,"assign")
## [1] 0 1
```

```
X <- model.matrix(Fit)
is.matrix(X)
```

```
## [1] TRUE
```

```
Y <- Corn$Yield
BetaHat <- solve(t(X)%*%(X))%*%t(X)%*%Y
BetaHat
```

```
##              [,1]
## (Intercept) 10.10
## X            1.15
```