STAT 210
Applied Statistics and Data Analysis
Multiple Linear Regression

Joaquin Ortega

Fall 2020

Introduction

Introduction

We aim to build models that explain the linear dependence of a response variable y on a set of explanatory variables $x_1, \ldots, x_p$.

This type of model appeared initially in Astronomy and Physics, where Gauss and Laplace used them.

The response variable y is always a random variable, while the explanatory variables, or regressors, are frequently variables whose value the experimenter controls.

However, this does not have to be so, i.e., the x_i can also be r.v. whose measurement is taken simultaneously as the y-variable value

Multiple Linear Regression

The basic hypothesis of the model is that

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_p x_p + \epsilon$$

where

- y and ϵ are random variables,
- $x_1, \ldots, x_p$ are the regressors or explanatory or independent variables,
- ϵ includes all random factors, and
- $\beta_0, \ldots, \beta_p$ are the regression coefficients we want to estimate.

β_j represents the increase in the response y when x_j increases one unit.

Multiple Linear Regression

We can write the model in matrix notation as

$$\mathbf{Y} = \mathbf{X}\beta + \boldsymbol{\epsilon}$$

where

$$\mathbf{Y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}; \quad \mathbf{X} = \begin{pmatrix} 1 & x_{11} & \cdots & x_{1p} \\ 1 & x_{21} & \cdots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & \cdots & x_{np} \end{pmatrix}; \quad \boldsymbol{\beta} = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{pmatrix}; \quad \boldsymbol{\epsilon} = \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{pmatrix}$$

and the dimensions are

$$\mathbf{Y} : (n \times 1) \quad \mathbf{X} : (n \times (p+1)) \quad \boldsymbol{\beta} : ((p+1) \times 1) \quad \boldsymbol{\epsilon} : (n \times 1)$$

Multiple Linear Regression

We make the following hypothesis on ϵ.

i) $E(\epsilon_i) = 0, i = 1, \ldots, n$.

ii) $Var(\epsilon_i) = \sigma^2, i = 1, \ldots, n$.

iii) $Cov(\epsilon_i \epsilon_j) = E(\epsilon_i \epsilon_j) = 0$ for all $i \neq j$.

iv) $\epsilon_i \sim N(0, \sigma^2)$.

This means that $\boldsymbol{\epsilon} \sim N_n(0, \sigma^2 I_n)$. We also have the following hypothesis on the regressors

v) The sample size n is bigger than or equal to $p + 1$, i.e., we have enough data to estimate the $p + 1$ parameters.

vi) Regressors are linearly independent, i.e., none of them is entirely determined by the rest.

Multiple Linear Regression

These hypotheses imply the following for the dependent variable:

i') $E(Y|X_1, \ldots, X_p) = \beta_0 + \beta_1 X_1 + \cdots + \beta_p X_p$.

ii') $Var(Y|X_1, \ldots, X_p) = \sigma^2$.

iii') The Y_i's are not correlated.

iv') $\mathbf{Y}$ has a normal distribution, and its components are independent.

Estimation

Estimation

The sample is

$$\begin{matrix} (y_1, x_{11}, \ldots, x_{1p}) \\ (y_2, x_{21}, \ldots, x_{2p}) \\ \vdots \quad\quad \vdots \\ (y_n, x_{n1}, \ldots, x_{np}) \end{matrix}$$

and we are looking for a hyperplane that best fits all of these points.

We fit it by the method of least squares that is equivalent to maximum likelihood under the hypothesis of normality.

As in the case of simple regression, the normal equation is $\mathbf{X}'\mathbf{X}\hat{\beta} = \mathbf{X}'\mathbf{Y}$ with solution

$$\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}.$$

Estimation

The unbiased estimator for the variance σ^2 of the errors is $\sum_i \hat{\epsilon}_i^2$ divided by the corresponding number of degrees of freedom.

From the normal equation, we have that

$$\mathbf{X}'(\mathbf{X}\hat{\beta} - \mathbf{Y}) = 0$$

and since $\mathbf{X}\hat{\beta}$ is the vector of fitted values, this is equivalent to

$$\mathbf{X}'\hat{\epsilon} = 0$$

so we have $p + 1$ restrictions.

Since there are n data points, the sum of squared residuals has $n - p - 1$ degrees of freedom.

Estimation

The unbiased estimator for σ^2 is then

$$MSE = \frac{SSE}{n-p-1} = \frac{1}{n-p-1}\sum_{i=1}^{n}\hat{\epsilon}_i^2.$$

We have that

$$\frac{\sum \hat{\epsilon}_i^2}{\sigma^2} \sim \chi^2_{n-p-1}.$$

Geometric Interpretation

Consider vectors $\mathbf{1}, \mathbf{X}_1, \ldots, \mathbf{X}_p$, columns of matrix $\mathbf{X}$. The objective of the estimation is to determine a linear combination of these vectors

$$\hat{\mathbf{Y}} = \beta_0 \mathbf{1} + \beta_1 \mathbf{X}_1 + \cdots + \beta_p \mathbf{X}_p$$

so that the norm of $\hat{\epsilon} = \mathbf{Y} - \hat{\mathbf{Y}}$ is minimal.

$\hat{\mathbf{Y}}$ is the projection of vector $\mathbf{Y}$ on the subspace generated by $(\mathbf{1}, \mathbf{X}_1, \ldots, \mathbf{X}_p)$ and $\hat{\epsilon} = \mathbf{Y} - \hat{\mathbf{Y}}$ must be orthogonal to any vector of this subspace, in particular to the generators, that is,

$$\mathbf{1}'\hat{\epsilon} = \mathbf{X}_1'\hat{\epsilon} = \cdots = \mathbf{X}_p'\hat{\epsilon} = 0$$

or in matrix notation

$$\mathbf{X}'\hat{\epsilon} = 0.$$

Geometric Interpretation

This is

$$\mathbf{X}'(\mathbf{Y} - \mathbf{X}\hat{\beta}) = 0 \quad \Rightarrow \quad \mathbf{X}'\mathbf{Y} = \mathbf{X}'\mathbf{X}\hat{\beta}$$

which is the normal equation.

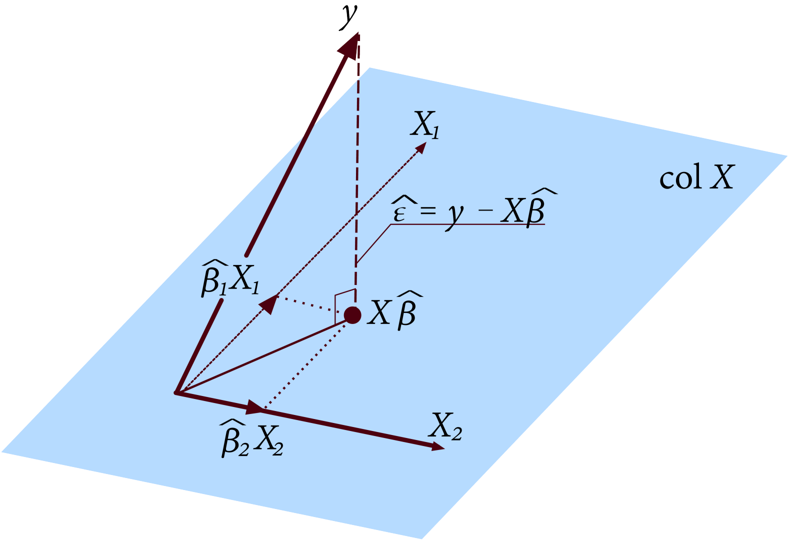
By orthogonality

$$||\mathbf{Y}||^2 = ||\hat{\mathbf{Y}}||^2 + ||\hat{\boldsymbol{\epsilon}}||^2$$

which is equivalent to

$$\sum y_i^2 = \sum \hat{y}_i^2 + \sum \hat{\epsilon}_i^2.$$

Geometric Interpretation



Maximum Likelihood

Another important estimation method is Maximum Likelihood, which in this case gives the same parameter estimates for the regression model.

Recall that Y has a normal distribution with mean

$$\beta_0 + \beta_1 X_1 + \cdots + \beta_p X_p$$

and variance σ^2 and they are independent. In matrix notation:

$$\mathbf{Y} \sim N(\boldsymbol{\beta}_0 + \boldsymbol{\beta}_1 \mathbf{X}_1 + \cdots + \boldsymbol{\beta}_p \mathbf{X}_p, \sigma^2 \mathbf{I}_n) = N(\mathbf{X}\beta, \sigma^2 \mathbf{I}_n)$$

The density function for $\mathbf{Y}$ is

$$\prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{ -\frac{(Y_i - (\beta_0 + \beta_1 X_{1i} + \cdots + \beta_p X_{pi}))^2}{2\sigma^2} \right\} \tag{1}$$

Maximum Likelihood

After we observe the sample, and we replace these values in (1), this becomes a function of the unknown parameters $\boldsymbol{\beta}$ and σ^2. This is known as the **likelihood function**.

$$\mathcal{L}(\boldsymbol{\beta}, \sigma^2|\mathbf{X}) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(Y_i - (\beta_0 + \beta_1 X_{1i} + \cdots + \beta_p X_{pi}))^2}{2\sigma^2}\right\} \tag{2}$$

In the maximum likelihood method, we obtain the estimators by maximizing the likelihood function.

Equivalently, we can obtain the parameters maximizing the natural log of the likelihood, known as the log-likelihood function

$$\ell(\boldsymbol{\beta}, \sigma^2|\mathbf{X}) = \ln \mathcal{L}(\boldsymbol{\beta}, \sigma^2|\mathbf{X})$$

Maximum Likelihood

To get the estimators, we differentiate either of these expressions with respect to the parameters and set the derivatives equal to zero.

The estimator for β is the same we obtained with least squares

$$\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}.$$

The estimator for σ^2 is

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}\hat{\epsilon}_i^2$$

which is biased.

We prefer to use MSE, since it is unbiased.

Example

Example

For this example, we use the `mtcars` data set, available in the base package.

```r
str(mtcars, vec.len = 2)
```

```
## 'data.frame':	32 obs. of  11 variables:
##  $ mpg : num  21 21 22.8 21.4 18.7 ...
##  $ cyl : num  6 6 4 6 8 ...
##  $ disp: num  160 160 108 258 360 ...
##  $ hp  : num  110 110 93 110 175 ...
##  $ drat: num  3.9 3.9 3.85 3.08 3.15 ...
##  $ wt  : num  2.62 2.88 ...
##  $ qsec: num  16.5 17 ...
##  $ vs  : num  0 0 1 1 0 ...
##  $ am  : num  1 1 1 0 0 ...
##  $ gear: num  4 4 4 3 3 ...
##  $ carb: num  4 4 1 1 2 ...
```

```r
head(mtcars, 3)
```

```
##                mpg cyl disp  hp drat    wt  qsec vs am gear carb
## Mazda RX4     21.0   6  160 110 3.90 2.620 16.46  0  1    4    4
## Mazda RX4 Wag 21.0   6  160 110 3.90 2.875 17.02  0  1    4    4
## Datsun 710    22.8   4  108  93 3.85 2.320 18.61  1  1    4    1
```

Example

Although all the variables appear as numerical, some of them, such as `am` and `vs` are, in fact, categorical. They represent whether the transmission is automatic (`am=1` for manual) and whether the engine is V-shaped (`vs=0` for V-shaped).

For our example, we will only keep

- ‘mpg‘, which measures fuel consumption and will be the response variable in the model,
- ‘disp‘, the displacement in cubic inches,
- ‘hp‘, the gross horsepower,
- ‘drat‘, the rear axle ratio and
- ‘wt‘, the weight.

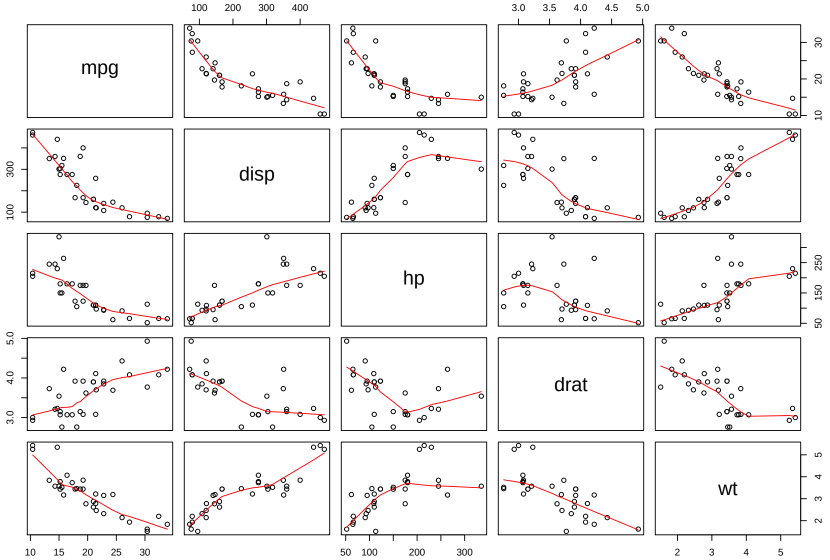
Example

```r
data.cars <- mtcars[,c('mpg','disp','hp','drat','wt')]
head(data.cars)
```

```
##                    mpg disp  hp drat    wt
## Mazda RX4         21.0  160 110 3.90 2.620
## Mazda RX4 Wag     21.0  160 110 3.90 2.875
## Datsun 710        22.8  108  93 3.85 2.320
## Hornet 4 Drive    21.4  258 110 3.08 3.215
## Hornet Sportabout 18.7  360 175 3.15 3.440
## Valiant           18.1  225 105 2.76 3.460
```

Example

```
pairs(data.cars, panel=panel.smooth)
```



Example

The top graphs have `mpg` on the y axis and show that this variable decreases with `disp`, `hp`, and `wt` but increases with `drat`.

We fit a model with the four variables using the function `lm()`.

```
model1 <- lm(mpg ~ hp + wt + disp + drat, data = data.cars)
```

Example

```r
summary(model1)
```

```
##
## Call:
## lm(formula = mpg ~ hp + wt + disp + drat, data = data.cars)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.5077 -1.9052 -0.5057  0.9821  5.6883
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 29.148738   6.293588   4.631  8.2e-05 ***
## hp          -0.034784   0.011597  -2.999  0.00576 **
## wt          -3.479668   1.078371  -3.227  0.00327 **
## disp         0.003815   0.010805   0.353  0.72675
## drat         1.768049   1.319779   1.340  0.19153
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.602 on 27 degrees of freedom
## Multiple R-squared:  0.8376, Adjusted R-squared:  0.8136
## F-statistic: 34.82 on 4 and 27 DF,  p-value: 2.704e-10
```

Example

Results show that the coefficients for `hp` and `wt` are negative. For `drat` it is positive, and for `disp`, it is positive but small.

`hp` and `wt` have a significant *p*-values, but the other two variables do not.

The R^2 is nearly .838 while the adjusted R^2 is .814.

Going back to the graphs, we see that `hp` and `disp` have a strong linear relation, and they are probably adding the same information to the model. In fact, their correlation is

```
with(data.cars, cor(hp,disp))
```

```
## [1] 0.7909486
```

Example

Part of the design matrix:

```r
head(model.matrix(model1),10)
```

```
##                   (Intercept)  hp    wt  disp drat
## Mazda RX4                   1 110 2.620 160.0 3.90
## Mazda RX4 Wag               1 110 2.875 160.0 3.90
## Datsun 710                  1  93 2.320 108.0 3.85
## Hornet 4 Drive              1 110 3.215 258.0 3.08
## Hornet Sportabout           1 175 3.440 360.0 3.15
## Valiant                     1 105 3.460 225.0 2.76
## Duster 360                  1 245 3.570 360.0 3.21
## Merc 240D                   1  62 3.190 146.7 3.69
## Merc 230                    1  95 3.150 140.8 3.92
## Merc 280                    1 123 3.440 167.6 3.92
```

Example

A fundamental principle in statistical modeling is parsimony, also known as Occam's razor.

As Einstein said, 'Everything should be made as simple as possible, but not simpler'.

In this spirit, we could try a simpler model excluding the variable `disp`.

Example

```r
model2 <- lm(mpg ~ hp + wt + drat, data = data.cars)
summary(model2)
```

```
##
## Call:
## lm(formula = mpg ~ hp + wt + drat, data = data.cars)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.3598 -1.8374 -0.5099  0.9681  5.7078
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 29.394934   6.156303   4.775 5.13e-05 ***
## hp          -0.032230   0.008925  -3.611 0.001178 **
## wt          -3.227954   0.796398  -4.053 0.000364 ***
## drat         1.615049   1.226983   1.316 0.198755
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.561 on 28 degrees of freedom
## Multiple R-squared:  0.8369, Adjusted R-squared:  0.8194
## F-statistic: 47.88 on 3 and 28 DF,  p-value: 3.768e-11
```

Example

We see that the R^2 has decreased marginally while the adjusted R^2, which considers the number of variables in the model, has increased.

Let us compare the coefficients

```
round(coef(model1),3); round(coef(model2),3)
```

```
## (Intercept)          hp          wt        disp        drat
##      29.149      -0.035      -3.480       0.004       1.768

## (Intercept)          hp          wt        drat
##      29.395      -0.032      -3.228       1.615
```

We see that the change is small.

Example

We can also try removing the variable `drat`.

```
model3 <- lm(mpg ~ hp + wt , data = data.cars)
summary(model3)
```

```
##
## Call:
## lm(formula = mpg ~ hp + wt, data = data.cars)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -3.941 -1.600 -0.182  1.050  5.854
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 37.22727    1.59879  23.285  < 2e-16 ***
## hp          -0.03177    0.00903  -3.519  0.00145 **
## wt          -3.87783    0.63273  -6.129 1.12e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.593 on 29 degrees of freedom
## Multiple R-squared:  0.8268, Adjusted R-squared:  0.8148
## F-statistic: 69.21 on 2 and 29 DF,  p-value: 9.109e-12
```

The R^2's have diminished slightly, and the estimated coefficients are similar to those we had before.

Example

Since `disp` and `hp` had a large correlation, we can also try a model with `disp` instead of `hp`.

```
model4 <- lm(mpg ~ disp + wt, data = data.cars)
summary(model4)
```

```
##
## Call:
## lm(formula = mpg ~ disp + wt, data = data.cars)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.4087 -2.3243 -0.7683  1.7721  6.3484
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 34.96055    2.16454  16.151 4.91e-16 ***
## disp        -0.01773    0.00919  -1.929  0.06362 .
## wt          -3.35082    1.16413  -2.878  0.00743 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.917 on 29 degrees of freedom
## Multiple R-squared:  0.7809, Adjusted R-squared:  0.7658
## F-statistic: 51.69 on 2 and 29 DF,  p-value: 2.744e-10
```

but this does not seem to be an improvement.

Example

If we change the order in which the regressors appear in the model, we get:

```
model4a <- lm(mpg ~ wt + disp, data = data.cars)
summary(model4a)
```

```
##
## Call:
## lm(formula = mpg ~ wt + disp, data = data.cars)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.4087 -2.3243 -0.7683  1.7721  6.3484
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 34.96055    2.16454  16.151 4.91e-16 ***
## wt          -3.35082    1.16413  -2.878  0.00743 **
## disp        -0.01773    0.00919  -1.929  0.06362 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.917 on 29 degrees of freedom
## Multiple R-squared:  0.7809, Adjusted R-squared:  0.7658
## F-statistic: 51.69 on 2 and 29 DF,  p-value: 2.744e-10
```

Sampling Distribution

Sampling Distribution

We have $\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$.

As in the simple regression case, the expected value is

$$E(\hat{\beta}) = E[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}] = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'E(\mathbf{Y}) = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X}\beta = \beta$$

so $\hat{\beta}$ is an unbiassed estimator for β. As for the variance,

$$\begin{aligned} Var(\hat{\beta}) &= Var((\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}) \\ &= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'Var(\mathbf{Y})\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} \\ &= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\sigma^2\mathbf{I}_n\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} \\ &= \sigma^2(\mathbf{X}'\mathbf{X})^{-1} \qquad (3) \end{aligned}$$

Sampling Distribution

For each estimated parameter $\hat{\beta}_j$, we have

$$\hat{\beta}_j \sim N(\beta_j, \sigma^2 v_{jj}),$$

where the v_{jj}'s are the diagonal elements of matrix $V = (\mathbf{X}'\mathbf{X})^{-1}$.

In general, the matrix $\mathbf{X}'\mathbf{X}$ is not diagonal. Therefore, its inverse will not be diagonal either, and the estimated coefficients $\hat{\beta}$ will not be independent as they do not have null covariance.

The Hat Matrix

The Hat Matrix

Since $\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$ and $\hat{\mathbf{Y}} = \mathbf{X}\hat{\beta}$ we have that

$$\hat{\mathbf{Y}} = \mathbf{X}\hat{\beta} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y} = \mathbf{H}\mathbf{Y}$$

where $\mathbf{H}$ is known as the 'hat' matrix.

$\mathbf{H}$ is idempotent ($\mathbf{H}^2 = \mathbf{H}$) and symmetric.

It is a projection matrix that sends points onto the column space of the design matrix.

We have already interpreted the elements of this matrix in terms of 'leverage', particularly the diagonal elements, in the simple regression case.

The Hat Matrix

Properties:

1. $h_{ii} = \mathbf{X}_i'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}_i, \quad h_{ij} = \mathbf{X}_i'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}_j$ only depend on $\mathbf{X}$, $\mathbf{X}_i'$ is the i-th row of $\mathbf{X}$.
2. We have
$$rank(\mathbf{H}) = rank(\mathbf{X}) = p + 1, \quad tr(\mathbf{H}) = \sum h_{ii} = p + 1$$
3. $\mathbf{HX} = \mathbf{X} \Rightarrow \mathbf{H1} = \mathbf{1} \Rightarrow \sum_{j=1}^{n} h_{ij} = 1$.
4. $\mathbf{H}^2 = \mathbf{H} \Rightarrow h_{ii} = \sum_j h_{ij}^2 \Rightarrow h_{ii} \geq h_{ii}^2 \Rightarrow |h_{ii}| \leq 1$.
5. It is also true that $h_{ij}^2 \leq h_{ii}$. Thus $nh_{ii} \geq 1 \Rightarrow h_{ii} \geq 1/n$ and we get
$$\frac{1}{n} \leq h_{ii} \leq 1$$

The Hat Matrix

6.

$$
\begin{aligned}
Cov(\hat{\mathbf{Y}}) = Cov(\mathbf{X}\hat{\beta}) &= Cov(\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}) \\
&= Cov(\mathbf{H}\mathbf{Y}) = \sigma^2\mathbf{H}
\end{aligned}
$$

and $Var(\hat{\mathbf{Y}}(\mathbf{X}_i)) = \sigma^2 h_{ii}$. Therefore

$$
\frac{1}{n} \leq \frac{Var(\hat{\mathbf{Y}}(\mathbf{X}_i))}{\sigma^2} \leq 1 \quad \text{and} \quad \frac{\sigma^2}{n} \leq Var(\hat{\mathbf{Y}}(\mathbf{X}_i)) \leq \sigma^2.
$$

7. If there are many observations, the h_{ij} are small, and the residuals are practically uncorrelated. On the other hand, if h_{ii} is big, $Var(\hat{\epsilon}_i)$ is small, and the i-th observation will attract the regression line or hyperplane.

The Hat Matrix

As an example we are going to obtain the hat matrix for the model with regressors `disp` and `wt` in the previous example. Remember that this is `model4`.

```
X <-model.matrix(model4)
n <-nrow(X)
p <-ncol(X)
H <- X%*%solve(t(X)%*%X)%*%t(X)
hii <-diag(H)
```

Let us verify that the sum of the diagonal elements of $\mathbf{H}$ (the trace of $\mathbf{H}$) is p:

```
sum(hii)
```

```
## [1] 3
```

The Hat Matrix

```
plot(hii,type = 'h', ylim = c(0,.2))
```