STAC67: Regression Analysis

Lecture 8

Sohee Kang

Feb. 4, 2021

Chapter 5: Matrix Approach to Simple Linear Regression Analysis

Matrix

- Definition: A **matrix** is a rectangular array of numbers or symbolic elements
- The dimension of a matrix is its number of rows and columns, often denoted as $r \times c$ (r rows by c columns)
- Can be represented in full form or abbreviated form:

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1c} \\ a_{21} & a_{22} & \dots & a_{2c} \\ \vdots & \vdots & \vdots & \vdots \\ a_{r1} & a_{r2} & \dots & a_{rc} \end{bmatrix} = [a_{ij}], \quad i = 1, \dots, r; j = 1, \dots, c$$

- Special Types of Matrics

Squre Matrix: number of rows = number of columns ($r = c$)
Vector: matrix with one column (column vector) or one row (row vector)

$$\mathbf{D} = \begin{bmatrix} d_1 \\ d_2 \\ d_3 \end{bmatrix}, \qquad \mathbf{E}' = [17 \quad 31], \qquad \mathbf{F}' = [f_1 \quad f_2 \quad f_3]$$

Special Types of Matrices

Transponse: matrix formed by interchanging rows and columns of matrix (use "prime" to denote transpose)

$$\mathbf{G}_{2\times 3} = \begin{bmatrix} 6 & 15 & 22 \\ 8 & 13 & 25 \end{bmatrix}, \quad \mathbf{G}'_{3\times 2} = \begin{bmatrix} 6 & 8 \\ 15 & 13 \\ 22 & 25 \end{bmatrix}$$

- Regression Example

Response Vector: $\mathbf{Y}_{n\times 1} = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}, \quad \mathbf{Y}'_{1\times n} = [Y_1 \quad Y_2 \quad \dots \quad Y_n]$

or $\vec{Y}$

Design Matrix: $\mathbf{X}_{n\times 2} = \begin{bmatrix} 1 & X_1 \\ 1 & X_2 \\ \vdots & \\ 1 & X_n \end{bmatrix}, \quad \mathbf{X'}_{2\times n} = \begin{bmatrix} 1 & 1 & \dots & 1 \\ X_1 & X_2 & \dots & X_n \end{bmatrix}$

or $\underset{\sim}{X}$

Matrix Addition and Subtraction

- Addition/subtraction: add (or subtract) the corresponding elements of the two matrices, defined if and only if the matrices are of the same order.
- Addition is commutative, i.e. $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$.
- Regression Example

$$\mathbf{Y}_{n\times 1} = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}, \quad \boldsymbol{E}(\boldsymbol{Y})_{n\times 1} = \begin{bmatrix} E(Y_1) \\ E(Y_2) \\ \vdots \\ E(Y_n) \end{bmatrix}, \quad \boldsymbol{\epsilon}_{n\times 1} = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{bmatrix}$$

$$\underset{n\times 1}{\boldsymbol{Y}} = \underset{n\times 1}{\boldsymbol{E}(\boldsymbol{Y})} + \underset{n\times 1}{\boldsymbol{\epsilon}} \quad since \quad \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} E(Y_1) \\ E(Y_2) \\ \vdots \\ E(Y_n) \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{bmatrix} = \begin{bmatrix} E(Y_1) + \epsilon_1 \\ E(Y_2) + \epsilon_2 \\ \vdots \\ E(Y_n) + \epsilon_n \end{bmatrix}$$

Matrix Multiplication

- **Multiplication** of a Matrix by a Scalar (single number)

$$k = 3, \quad \boldsymbol{A} = \begin{bmatrix} 2 & 1 \\ -2 & 7 \end{bmatrix} \Rightarrow k\boldsymbol{A} = \begin{bmatrix} 3(2) & 3(1) \\ 3(-2) & 3(7) \end{bmatrix} = \begin{bmatrix} 6 & 3 \\ -6 & 21 \end{bmatrix}$$

- **Multiplication** of a Matrix by a Matrix (# cols(**A**) = #rows(**B**))

If $c_A = r_B = c$:

$$\underset{r_A \times c_A}{\boldsymbol{A}} \underset{r_B \times c_B}{\boldsymbol{B}} = [ab_{ij}] = \left[\sum_{k=1}^{c} a_{ik} b_{kj} \right] \quad i = 1, \dots, r_A; j = 1, \dots, c_B$$

- ab_{ij} = sum of the products of the c elements of i^{th} row of **A** and j^{th} column of **B**

$$\underset{(3\times 2)}{\boldsymbol{A}} \underset{(2\times 2)}{\boldsymbol{B}} = \begin{bmatrix} 2 & 5 \\ 3 & -1 \\ 0 & 7 \end{bmatrix} \begin{bmatrix} 3 & -1 \\ 2 & 4 \end{bmatrix} = \begin{bmatrix} 16 & 18 \\ 7 & -7 \\ 14 & 28 \end{bmatrix}$$

Matrix Multiplication Example

- Simultaneous Equations: $a_{11}x_1 + a_{12}x_2 = y_1 \quad a_{21}x_1 + a_{22}x_2 = y_2$

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

- Sum of Squares: $4^2 + (-2)^2 + 3^2 = \begin{bmatrix} 4 & -2 & 3 \end{bmatrix} \begin{bmatrix} 4 \\ -2 \\ 3 \end{bmatrix} = [29]$

- Regression Equation (Expected Values):

Regression Coef. Vector

Design Matrix

$$\begin{bmatrix} 1 & X_1 \\ 1 & X_2 \\ \vdots & \vdots \\ 1 & X_n \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} \beta_0 + \beta_1 X_1 \\ \beta_0 + \beta_1 X_2 \\ \vdots \\ \beta_0 + \beta_1 X_n \end{bmatrix}$$

Matrix Multiplication Examples

- Matrices used in simple linear regression (that generalize to multiple regression):

$$\boldsymbol{Y}'\boldsymbol{Y} = \begin{bmatrix} Y_1 & Y_2 & \dots & Y_n \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \sum Y_i^2$$

$$\boldsymbol{X}'\boldsymbol{X} = \begin{bmatrix} 1 & 1 & \dots & 1 \\ X_1 & X_2 & \dots & X_n \end{bmatrix} \begin{bmatrix} 1 & X_1 \\ 1 & X_2 \\ \vdots & \vdots \\ 1 & X_n \end{bmatrix} = \begin{bmatrix} n & \sum x_i \\ \sum x_i & \sum x_i^2 \end{bmatrix}$$

$$\boldsymbol{X}'\boldsymbol{Y} = \begin{bmatrix} 1 & 1 & \dots & 1 \\ X_1 & X_2 & \dots & X_n \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} \sum y_i \\ \sum x_i y_i \end{bmatrix}$$

Special Matrix Types

- **Symmetric Matrix**: Squre matrix with a transpose equal to itself: $\boldsymbol{A} = \boldsymbol{A}'$
- **Diagonal Matrix**: Squre matrix with all off-diagonal elements equal to 0:
- **Identity Matrix**: Diagonal matrix with all diagonal elements equal to 1 (acts like multiplying a scalar by 1)

$$\underset{3\times 3}{\boldsymbol{I}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \underset{3\times 3}{\boldsymbol{A}} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \quad \Rightarrow \boldsymbol{IA} = \boldsymbol{AI} = \boldsymbol{A}$$

- **1-Vector and Matrix and Zero Vector**:

$$\underset{r\times 1}{\mathbf{1}} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \quad \underset{r\times r}{\boldsymbol{J}} = \begin{bmatrix} 1 & \dots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \dots & 1 \end{bmatrix} \quad \underset{r\times 1}{\mathbf{0}} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Linear Dependence and Rank of a Matrix

- Linear Dependence: When a linear function of the columns (rows) of a matrix produces a zero vector (one or more columns (rows) can be written as linear function of the other columns (rows))
- Rank of a matrix: Number of linearly independent columns (rows) of the matrix. Rank cannot exceed the minimum of the number of rows or columns of the matrix. $rank(\boldsymbol{A}) \leq min(r_A, c_A)$
- A matrix is full rank if $rank(A) = min(r_A, c_A)$
- **Matrix Inversion**

In matrix form if $\boldsymbol{A}$ is a **square matrix** and **full rank** (all rows and columns are linearly independent), then $\boldsymbol{A}$ has an inverse: $\boldsymbol{A}^{-1}$ such that:
$\boldsymbol{A}^{-1}\boldsymbol{A} = \boldsymbol{A}\boldsymbol{A}^{-1} = \boldsymbol{I}$

Computing an Inverse of 2 X 2 Matrix

$$A_{2\times 2} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \text{full rank (columns/rows are linear indepenent)}$$

Determinant of $\boldsymbol{A} = |\boldsymbol{A}| = a_{11}a_{22} - a_{12}a_{21}$

Note: If $\boldsymbol{A}$ is not full rank then $|\boldsymbol{A}| = 0$

- **Regression Example**: find the expression for $(\boldsymbol{X}'\boldsymbol{X})^{-1}$

$$|X'X| = \left| \begin{bmatrix} n & \sum x_i \\ \sum x_i & \sum x_i^2 \end{bmatrix} \right| = n\sum x_i^2 - \left(\sum x_i\right)^2$$

$$= n\left(\sum x_i^2 - \frac{(\sum x_i)^2}{n}\right)$$

$$= n\sum (x_i - \bar{x})^2$$

$$= n\,S_{xx}$$

Thus $$(X'X)^{-1} = \frac{1}{nS_{xx}} \cdot \begin{bmatrix} \sum x_i^2 & -\sum x_i \\ -\sum x_i & n \end{bmatrix} = \begin{bmatrix} \frac{1}{n} + \frac{\bar{x}^2}{S_{xx}} & -\frac{\bar{x}}{S_{xx}} \\ -\frac{\bar{x}}{S_{xx}} & \frac{1}{S_{xx}} \end{bmatrix}$$

Useful Matrix Results

- Addition Rules:

$$\boldsymbol{A} + \boldsymbol{B} = \boldsymbol{B} + \boldsymbol{A}, \quad (\boldsymbol{A} + \boldsymbol{B}) + \boldsymbol{C} = \boldsymbol{A} + (\boldsymbol{B} + \boldsymbol{C})$$

- Multiplcation Rules:

$$(\boldsymbol{AB})\boldsymbol{C} = \boldsymbol{A}(\boldsymbol{BC}), \quad \boldsymbol{C}(\boldsymbol{A} + \boldsymbol{C}) = \boldsymbol{CA} + \boldsymbol{CB}, \quad k(\boldsymbol{A} + \boldsymbol{B}) = k\boldsymbol{A} + k\boldsymbol{B}, k = scalar$$

- Transpose Rules:

$$(\boldsymbol{A}')' = \boldsymbol{A}, \quad (\boldsymbol{A} + \boldsymbol{B})' = \boldsymbol{A}' + \boldsymbol{B}', \quad (\boldsymbol{AB})' = \boldsymbol{B}'\boldsymbol{A}', \quad (\boldsymbol{ABC})' = \boldsymbol{C}'\boldsymbol{B}'\boldsymbol{A}'$$

- Inverse Rules (Full Rank, Square Matrices)

$$(\boldsymbol{AB})^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1}, (\boldsymbol{ABC})^{-1} = \boldsymbol{C}^{-1}\boldsymbol{B}^{-1}\boldsymbol{A}^{-1}, \ (\boldsymbol{A}^{-1})^{-1} = \boldsymbol{A}, (\boldsymbol{A}')^{-1} = (\boldsymbol{A}^{-1})'$$

Useful Matrix Resuls

- Properties of the determinant:

$\det(I) = 1$

$\det(cA_{n\times n}) = c^n \det(A)$

$\det(A^{-1}) = \frac{1}{\det(A)}$

$\det(A') = \det(A)$

$\det(AB) = \det(A)\det(B)$

- Properties of the trace:

$Tr(A+B) = Tr(A) + Tr(B)$

$Tr(cA) = c\,Tr(A)$

Trace is invariant under cyclic permutations

$Tr(ABC) = Tr(CAB)$

Random Vectors and Matrices

- Shown for case of $n = 3$, generazlies to any n:
- Random variables: $Y_1,\ Y_2,\ Y_3 \Rightarrow \underset{\sim}{\boldsymbol{Y}} = \begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \end{bmatrix}$ with expectation: $\boldsymbol{E}(\underset{\sim}{\boldsymbol{Y}}) = \begin{bmatrix} E(Y_1) \\ E(Y_2) \\ E(Y_3) \end{bmatrix}$
- Variance-covariance Matrix:

$$\sigma^2 \underset{\sim}{\boldsymbol{Y}} = E\left[(\underset{\sim}{Y} - E(\underset{\sim}{Y}))(\underset{\sim}{Y} - E(\underset{\sim}{Y}))'\right] =$$

$$E\left[\begin{bmatrix} Y_1 - E(Y_1) \\ Y_2 - E(Y_2) \\ Y_3 - E(Y_3) \end{bmatrix} \begin{bmatrix} Y_1 - E(Y_1) & Y_2 - E(Y_2) & Y_3 - E(Y_3) \end{bmatrix}\right] =$$

$$= \begin{bmatrix} Cov(Y_1, Y_1) & \cdots & Cov(Y_1, Y_3) \\ \vdots & & \\ (Y_3, Y_1) & \cdots & Cov(Y_3, Y_3 \end{bmatrix}$$

$$= \begin{bmatrix} 6^2 & 0 & 0 \\ 0 & 6^2 & 0 \\ 0 & 0 & 6^2 \end{bmatrix} = 6^2 I$$

assumption
ε_i i.i.d. $N(0, 6^2)$

Simple Linear Regression in Matrix form

- Simple Linear regression Model: $Y_i = \beta_0 + \beta_1 X_i + \epsilon_i, \quad i = 1, \ldots, n$
- Define: $\vec{Y} = \begin{bmatrix} Y_1 \\ \vdots \\ Y_n \end{bmatrix}$ response vector $\quad \vec{B} = \begin{bmatrix} B_0 \\ B_1 \end{bmatrix}$ regression coef vector

  $X = \begin{bmatrix} 1 & X_1 \\ \vdots & \vdots \\ 1 & X_n \end{bmatrix}$ design matrix $\quad \vec{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{bmatrix}$ error vector
- Assuming constant variance and independence of error term, ϵ_i:

$$\sigma^2(\underset{\sim}{Y}) = \sigma^2(\underset{\sim}{\epsilon}) = \sigma^2 I_{n \times n} \text{ as shown above}$$

$$E(Y) = E(X\vec{B} + \vec{\varepsilon}) = X\vec{B}$$

- Further, assuming normal distribution for error tersm, ϵ_i:

$$\vec{\varepsilon} \sim N(\vec{0}, \sigma^2 I) \qquad Y \sim N(X\vec{B}, \sigma^2 I)$$

Estimating Parameters by Least Squares

- Mimimizing the sum of squares of errors:

$$Q(\underset{\sim}{\beta}) = \sum_{i=1}^{n} \epsilon_i^2 = \underset{\sim}{\epsilon}'\underset{\sim}{\epsilon} = (\underset{\sim}{Y} - \boldsymbol{X}\underset{\sim}{\beta})'(\underset{\sim}{Y} - \boldsymbol{X}\underset{\sim}{\beta})$$

- Normal equations:

$$\begin{bmatrix} n & \sum X_i \\ \sum X_i & \sum X_i^2 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} \sum Y_i \\ \sum X_i Y_i \end{bmatrix}, \quad or \quad (\boldsymbol{X}'\boldsymbol{X})\underset{\sim}{\beta} = \boldsymbol{X}'\underset{\sim}{y};$$

The soltion of $\underset{\sim}{\beta}$ is given by:

$$\hat{\vec{\beta}} = (X'X)^{-1}X'\vec{y}$$

Note: $(\boldsymbol{X}'\boldsymbol{X})$ should be invertable, full rank.

- Find $(\boldsymbol{X}'\boldsymbol{X})^{-1}$.

- Find $(\boldsymbol{X}'\boldsymbol{X})^{-1}(\boldsymbol{X}'\underset{\sim}{y})$. Check whether these estiamtes agrees with $\hat{\beta}_0, \hat{\beta}_1$

Fitted Values and Residuals

$$\underset{\sim}{e} = \begin{bmatrix} e_1 \\ \vdots \\ e_n \end{bmatrix} = \begin{bmatrix} Y_1 - \hat{Y}_1 \\ \vdots \\ Y_n - \hat{Y}_n \end{bmatrix} = \underset{\sim}{Y} - \underset{\sim}{\hat{Y}} = \underset{\sim}{Y} - \boldsymbol{X}\underset{\sim}{\hat{\beta}}$$

$$\underset{\sim}{\hat{Y}} = \boldsymbol{X}\underset{\sim}{\hat{\beta}} = \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\underset{\sim}{Y} = \boldsymbol{H}\underset{\sim}{Y}, \quad \boldsymbol{H} = \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'$$

- $\boldsymbol{H}$ is called the **hat** or **projection** matrix. Note that $\boldsymbol{H}$ is idempotent matrix ($\boldsymbol{HH} = \boldsymbol{H}$)

$Var(e) = (I-H)\,Var(y)\,(I-H)'$

$= (I-H)\,Var(y)$

$$\underset{\sim}{e} = \underset{\sim}{Y} - \boldsymbol{H}\underset{\sim}{Y} = (\boldsymbol{I} - \boldsymbol{H})\underset{\sim}{Y}$$

$$\hat{\sigma}^2 = \frac{1}{n-2}\sum_{i=1}^{n} e_i^2 = \frac{1}{n-2}\underset{\sim}{e}'\underset{\sim}{e}$$

$$(\boldsymbol{X}'\boldsymbol{X})\underset{\sim}{\hat{\beta}} = \boldsymbol{X}'\underset{\sim}{y}$$

$x'(\vec{Y} - x\hat{\beta}) = \vec{0}$

$x'\vec{e} = \vec{0}$

$$\boxed{\begin{array}{l} \sum e_i = 0 \\ \sum x_i e_i = 0 \end{array}}$$

Inferences in Linear Regression

$$\underset{\sim}{\hat{\beta}} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\underset{\sim}{Y}$$

$$E(\underset{\sim}{\hat{\beta}}) = E((X'X)^{-1}X'Y) = (X'X)^{-1}X'X\vec{B} = \vec{B}$$

$$\sigma^2(\underset{\sim}{\hat{\beta}}) = (X'X)^{-1}X'\sigma^2 I \times (X'X)^{-1} = \sigma^2 (X'X)^{-1}$$

$$\hat{\sigma}^2(\underset{\sim}{\hat{\beta}}) = MSE\,(X'X)^{-1}$$

Reall:

$$(\boldsymbol{X}'\boldsymbol{X})^{-1}\sigma^2 = \begin{bmatrix} Var(\hat{B}_0) & Cov(B_0, B_1) \\ Cov(B_1, B_0) & Var(\hat{B}_1) \end{bmatrix} = \begin{bmatrix} \frac{1}{n} + \frac{\bar{X}^2}{S_{xx}} & \frac{-\bar{X}}{S_{xx}} \\ \frac{-\bar{X}}{S_{xx}} & \frac{1}{S_{xx}} \end{bmatrix} \sigma^2$$

- The diagonal elements of this matrix gives $Var(\hat{\beta}_0)$ and $Var(\hat{\beta}_1)$.
- Off diagonal elements gives $Cov(\hat{\beta}_0, \hat{\beta}_1)$

Chapter 6: Mutiple Regression I

Models with Multiple Predictors

- Most Practical Problems have **more than one** potential predictor variable
- Goal is to determine effects (if any) of each predictor, controlling for others
- Can include polynomial terms to allow for nonlinear relations
- Can include product terms to allow for interactions when effect of one variable depends on level of another variable
- Can include "dummy" variables for categorical predictors
- First-Order Model with 2 Numeric Predictors

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \epsilon \Rightarrow E(Y) = \beta_0 + \beta_1 X_1 + \beta_2 X_2$$

Example - Multiple Regression with Two predictor Variable

- This dataset, which can be found at UCI Machine Learning Repository contains a response variable **mpg** which stores the city fuel efficiency of cars, as well as several predictor variables for the attributes of the vehicles. we would like to model the fuel efficiency (mpg) of a car as a function of its weight (**wt**) and model year (**year**).

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \epsilon_i, \quad i = 1, 2, \ldots, n, \quad where\ \epsilon_i \sim N(0, \sigma^2)$$

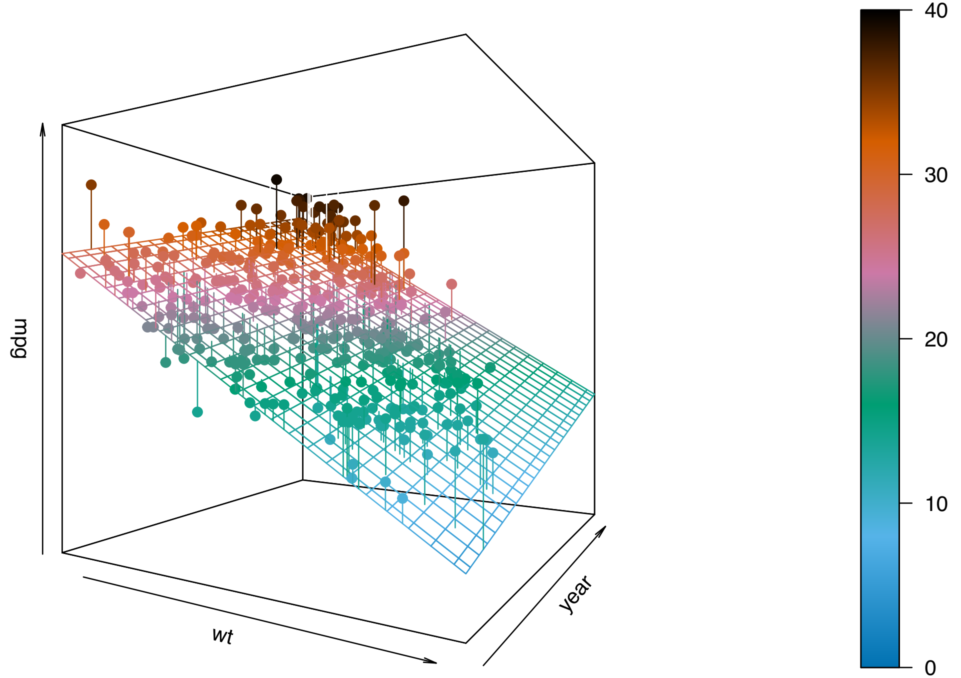
```
##
## Call:
## lm(formula = mpg ~ wt + year, data = autompg)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -8.852 -2.292 -0.100  2.039 14.325
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept) -1.464e+01  4.023e+00  -3.638 0.000312 ***
## wt          -6.635e-03  2.149e-04 -30.881  < 2e-16 ***
## year         7.614e-01  4.973e-02  15.312  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.431 on 387 degrees of freedom
```

3-dimensional image

- what happens if we fit two simple regressions?

```
##  (Intercept)           wt
## 46.198395049 -0.007643021
```

```
## (Intercept)        year
##  -70.155395    1.231997
```



Interpretation of regression parameters

1. Change of $E(Y)$ per unit corresponding to a unit change in $X_j, \quad j = 1, \ldots, p$ when other predictor variables are held constant.
2. β_2: effect of X_2 after **adjusting** for X_1
3. β_1: effect of X_1 after **adjusting** for the effect of X_2

- If X_1 and X_2 are **uncorrelated** then the estimates will be the same as the estimates in the simple models but in general, this is not true.
- Consider the following sequence of model:

1. Fit a simple linear model between Y and X_1, and find the residual, e_1.
2. Fit a regression with X_2 as a response and X_1 as a covariate and find the residual, e_2.
3. Fit a regression with e_1 as a response and e_2 as a covariate.

```
# Fit a simple regression between Y and X1 and find e1
```