Economic Forecasting and Analysis

林朕陞 博士
Dr. Chen-Sheng Lin

科技部
人文社會科學研究中心
2022.3.8

- Forecasting vs. Economic forecasting

  - **Forecasting** is the process of making predictions based on past and present data.

  - **Economic forecasting** is the process of making predictions about the economy based on economic theory.

■ **The theory of Inflation Prediction**

- Monetary model (Hallman, Poter and Small, 1991)

$$\pi_t = \beta_0 + \sum_{j=1}^{p} \beta_j \pi_{t-j} + \theta Mgap_{t-i}$$

- Phillips curve (Phillips, 1958)

$$\pi_t = \pi_t^e + \alpha(u_t^n - u_t)$$

- Markup model (Brouwer and Erisson, 1998)

$$lnP_t = ln\mu_0 + \gamma lnUCL_t + \delta LnIP_t + \kappa lnPET_t$$

where UCL: nominal cost of labor per unit of output, IP: import price, and PET: petrol price

■ How to produce a high-quality economic forecasts?

- Understand the ins and outs of the data

- Have some economic knowledge

- Be familiar with econometric methods

- To widely read financial newspapers and magazines
  (e.g., Business Week, Economist, Financial Times, Bloomberg,
  CNBC... etc.)
  ⇒ A good forecasting always needs timeliness and effectiveness.

■ Note:

- Forecasting is very time-consuming and costly.

- To make a forecast would be worthwhile only for significant effect on the policy making process of the authority.

  ⇒ Government officials use economic forecasts to determine fiscal and monetary policies.

■ Textbook:

Diebold, F. X. (2007), *Elements of Forecasting*, 4th ed., Cincinnati: South-Western College Publishing. [free download]

Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani, (2017), *An Introduction to Statistical Learning with Applications in R*, Springer, New York. [free download]

■ Today

- 1st period: Construct a linear regression model

- 2nd period: Estimating the coefficients

- 3rd period: Model Selection

- 4th period: Example for Taiwan Inflation Forecast

The **Phillips curve** is an economic concept developed by A. W. Phillips stating that inflation and unemployment have an inverse relationship.

Fig.1 Taiwan Inflation Rate and Unemployment Rate (1982Q1~2016Q4)
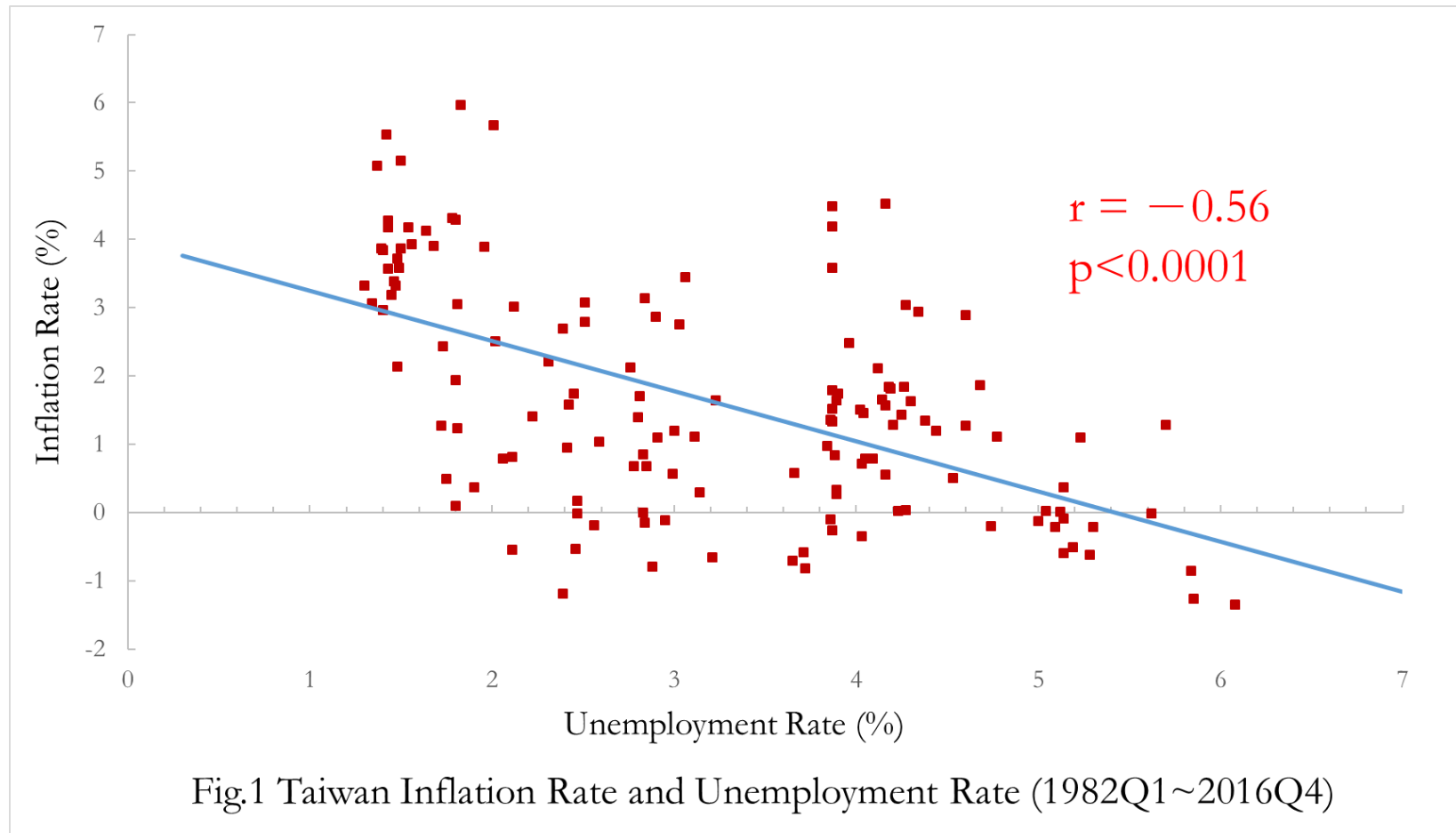
Linear Regression Model

- What is it?

- When to use it?

- How to use it?

- How to explain the results?

- Why? (suitability and assumptions)

■ Linear Regression Model

- Input vector: $X^T = (X_1, X_2, \ldots X_P)$

- A real-valued output: Y

$$Y = f(X) \textcolor{red}{+ \varepsilon} = \beta_0 + \sum_{j=1}^{p} \beta_j X_j \textcolor{red}{+ \varepsilon}$$

$$E(Y|X) = f(X) = \beta_0 + \sum_{j=1}^{p} \beta_j X_j$$

Note: "linear" means?

- **Model is linear in the parameters β**

$$Model\ 1: Y = \color{red}{\beta_0} + \color{red}{\beta_1} X_1^2 + \color{red}{\beta_2} \frac{1}{X_1} + \boldsymbol{\varepsilon} \qquad \textbf{Linear}$$

$$Model\ 2: Y = \color{red}{\beta_0^2} + \sqrt{\color{red}{\beta_1}} X_1 + \color{red}{\beta_2} X_2 X_3 + \boldsymbol{\varepsilon} \quad \textbf{Non Linear}$$

$$Model\ 3: Y = \color{red}{\beta_0} + \color{red}{\beta_1} X_1 + \color{red}{\beta_2}\color{red}{\beta_3} X_2 + \boldsymbol{\varepsilon} \qquad \textbf{Non Linear}$$

■ **Variables** X_j

- Quantitative inputs (continuous or discrete)

- Transformations of quantitative inputs, such as log, square-root or square

- Basis expansions, such as $X_2 = X_1^2$, $X_3 = X_1^3$ leading to a polynomial representation

- Numeric or "dummy" coding of the levels of qualitative inputs

- Interactions between variables, for example $X_4 = X_1 X_2$

■ **Least squares estimate (OLS)**

$$\min_{\beta} \text{RSS}(\beta) = \sum_{i=1}^{N} \left(y_i - f(x_i) \right)^2$$

$$= \sum_{j=1}^{p} \left(y_i - \left(\beta_0 + \sum_{j=1}^{p} \beta_j X_{ij} \right) \right)^2$$
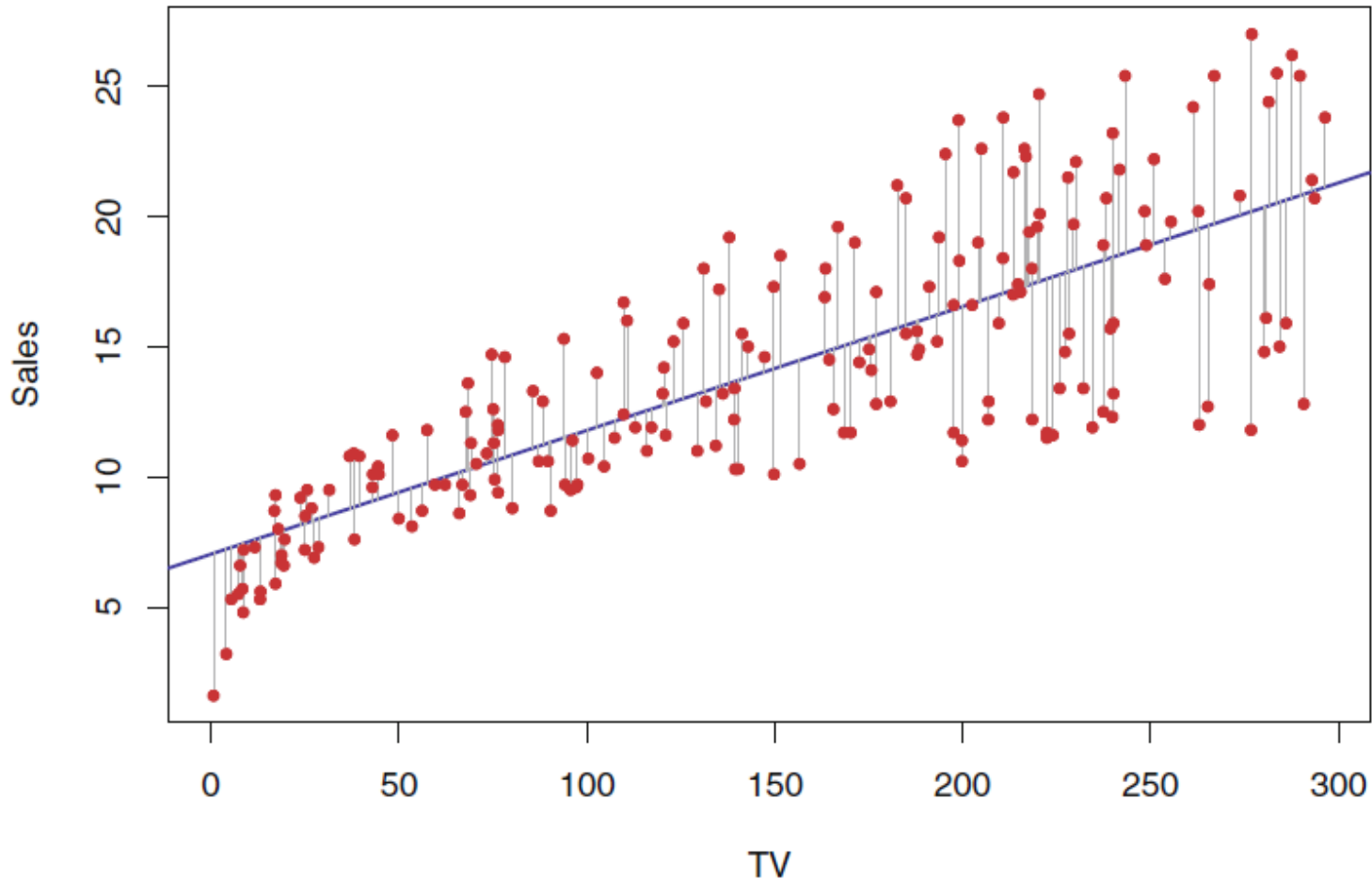
FIGURE 3.1. *For the* Advertising *data, the least squares fit for the regression of* sales *onto* TV *is shown. The fit is found by minimizing the sum of squared errors. Each grey line segment represents an error, and the fit makes a compromise by averaging their squares. In this case a linear fit captures the essence of the relationship, although it is somewhat deficient in the left of the plot.*
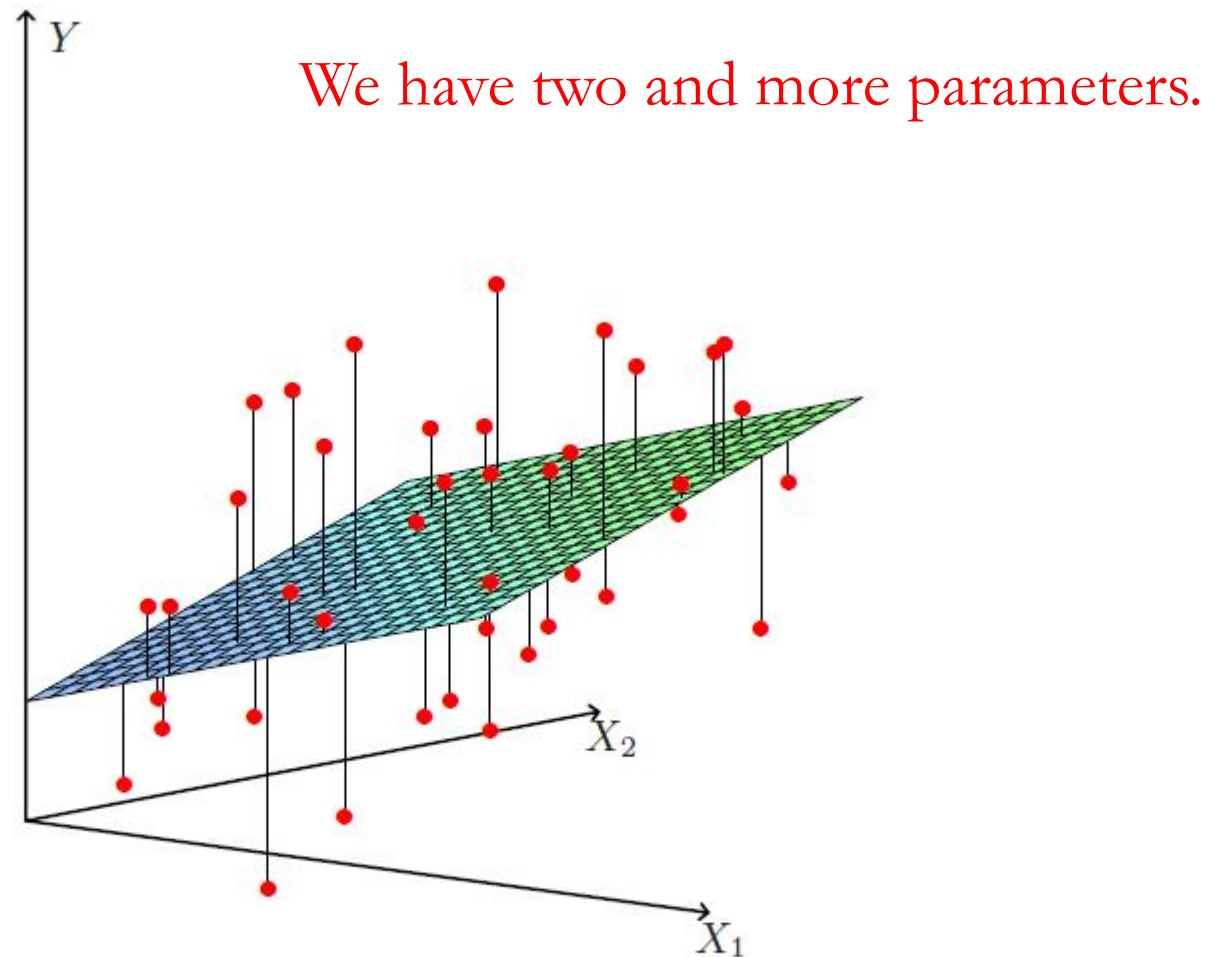
We have two and more parameters.

FIGURE 3.1. *Linear least squares fitting with $X \in \mathbb{R}^2$. We seek the linear function of X that minimizes the sum of squared residuals from Y.*

The red dot represents the pair of least squares estimates $(\hat{\beta}_0, \hat{\beta}_1)$. These values clearly minimize the RSS.

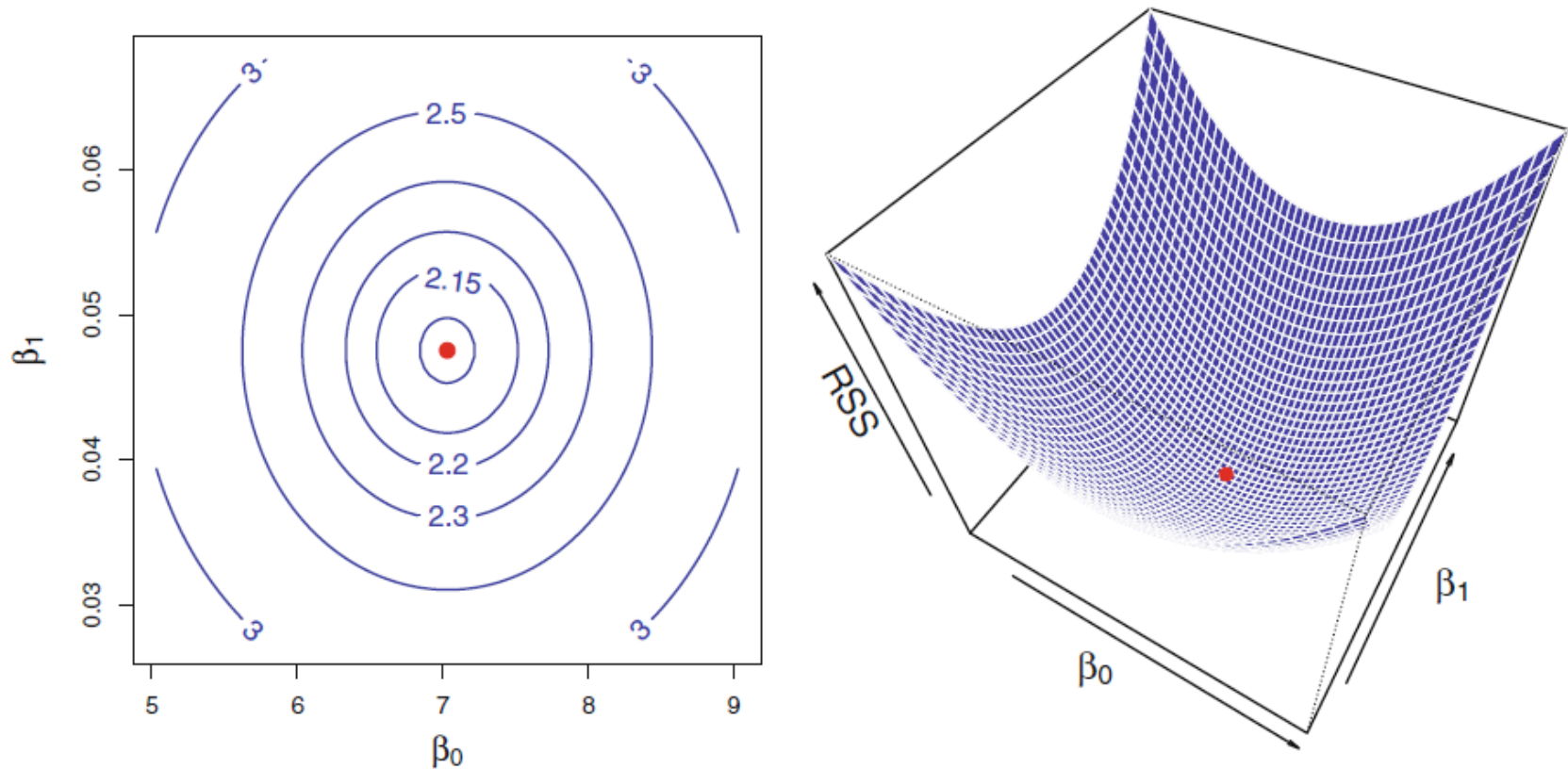**FIGURE 3.2.** *Contour and three-dimensional plots of the RSS on the* **Advertising** *data, using* **sales** *as the response and* **TV** *as the predictor. The red dots correspond to the least squares estimates* $\hat{\beta}_0$ *and* $\hat{\beta}_1$, *given by (3.4).*

- **Least squares estimate (OLS)**

$$\min_{\beta} \text{RSS}(\beta) = (\boldsymbol{y} - \boldsymbol{X}\beta)^T(\boldsymbol{y} - \boldsymbol{X}\beta)$$

$$\frac{\partial \text{RSS}(\beta)}{\partial \beta} = -2\boldsymbol{X}^T(\boldsymbol{y} - \boldsymbol{X}\beta)$$

$$\frac{\partial^2 \text{RSS}(\beta)}{\partial \beta \partial \beta^T} = 2\boldsymbol{X}^T\boldsymbol{X} \quad \text{positive definite}$$

- Assuming (for the moment) that $\boldsymbol{X}$ has full column rank, and hence $\boldsymbol{X}^T\boldsymbol{X}$ is positive definite, we set the first derivative to zero

$$\boldsymbol{X}^T(\boldsymbol{y} - \boldsymbol{X}\beta) = 0$$

$$\widehat{\beta} = (\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{y}$$

Notes:

• The columns of $\boldsymbol{X}$ are not linearly independent $\Rightarrow \boldsymbol{X}$ is not full rank.

• If two of the inputs were perfectly correlated, (e.g., $X_2 = 3X_1$)

• Then $\boldsymbol{X}^T\boldsymbol{X}$ is singular and the least squares coefficients $\widehat{\beta}$ are not uniquely defined.
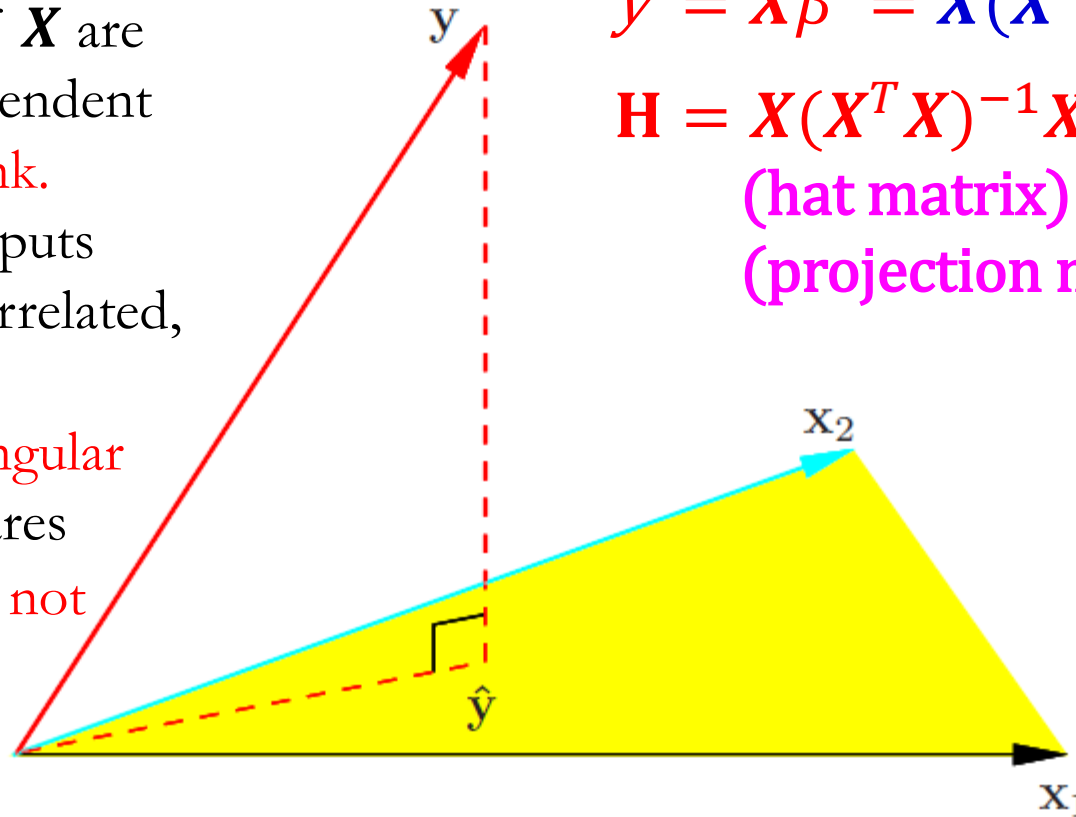
$$\widehat{y} = \boldsymbol{X}\widehat{\beta} = \boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{y}$$

$$\mathbf{H} = \boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T$$

(hat matrix)
(projection matrix)

FIGURE 3.2. *The N-dimensional geometry of least squares regression with two predictors. The outcome vector* **y** *is orthogonally projected onto the hyperplane spanned by the input vectors* **x₁** *and* **x₂***. The projection* **ŷ** *represents the vector of the least squares predictions*

- With a large number of predictors ($\boldsymbol{X}$), we often would like to determine a smaller subset that exhibit the strongest effects. In order to get the "big picture," we are willing to sacrifice some of the small details.

- From any set of p predictors: $\boldsymbol{2^p}$ alternative models can be constructed

- **All (Possible) Subsets**

  - Coefficient of Determination (R^2)

  - Adjusted Coefficient of Determination (R_a^2)

  - Akaike Information Criterion (AIC)

  - Schwarz Bayesian Criterion (SBC)

  - Mallows' Cp

■ Coefficient of Determination (R^2)

$$R^2 = 1 - \frac{RSS}{TSS} = 1 - \frac{\sum_{i=1}^{N}(y_i - f(x_i))^2}{\sum_{i=1}^{N}(y_i - \bar{y})^2}$$

- the proportion of variability in Y that can be explained using $\boldsymbol{X}$

■ Adjusted Coefficient of Determination (R_a^2)

$$R^2 = 1 - \frac{\dfrac{RSS}{N - p - 1}}{\dfrac{TSS}{N - 1}} = 1 - \frac{N - 1}{N - p - 1}(1 - R^2)$$

■ Akaike Information Criterion (AIC)

- $AIC = 2(p + 1) + N \times \ln(RSS/N)$

- $AIC = 2(p + 1) - 2lnL$

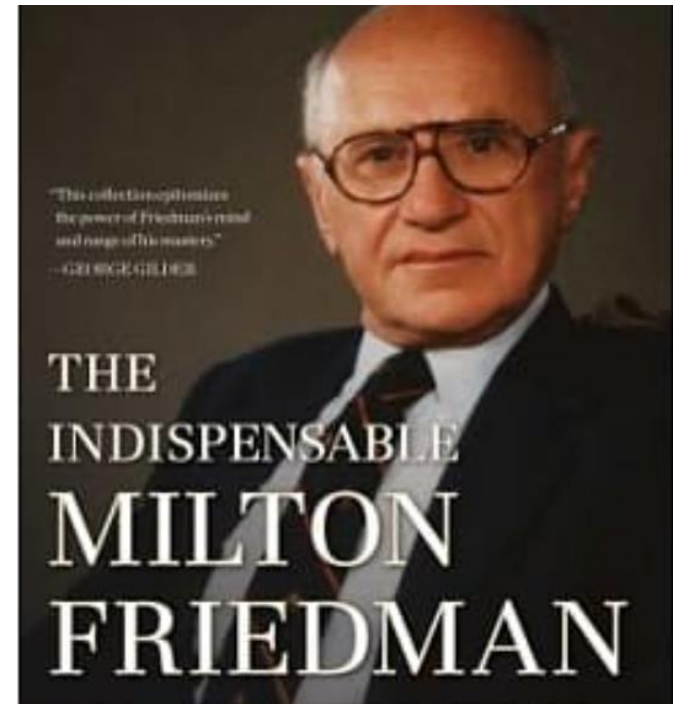■ Schwarz Bayesian Criterion (SBC) or Bayesian Information Criterion (BIC)

- $BIC = (p + 1)lnN + N \times \ln(RSS/N)$

- $BIC = (p + 1)lnN - 2lnL$

■ Mallows' Cp

- $C_P = \frac{1}{N}[RSS + 2(p + 1)\,\hat{\sigma}^2]$

- $C_P = \frac{RSS}{\hat{\sigma}^2} + 2(p + 1) - N$

- Y_t : Inflation rate

- X_t :

  - X_{1t} : unemployment rate

  - X_{2t} : M2 growth rate

  - X_{3t} : nonminal interest rate

- Sample period: 1982 Q1~2016 Q4

- Frequency: Quarterly

- Databank: AREMOS

Nobel Prize in Economics, 1976

"Inflation is always and everywhere a monetary phenomenon."
通貨膨脹無論何時何地皆是一種貨幣現象。

```
#### Using for loop
library(leaps)
ex=read.csv("inf.csv", header=TRUE)
colnames(ex)<-c("date","Y","X1","X2","X3")
attach(ex)
nn<-length(ex[,1])
vars <- c("Y","X1","X2","X3")
N <- list(0,1,2,3)#
COMB <- sapply(N, function(m) combn(x=vars[2:4], m))
COMB2 <- list()
k=0
exlmf<-lm(Y~X1+X2+X3)
SSEF<-deviance(aov(exlmf))
MSEF<-deviance(aov(exlmf))/(nn-4)
res.table<-NULL
res.table0<-NULL
for(i in seq(COMB))
{
tmp <- COMB[[i]]
for(j in seq(ncol(tmp)))
```

```
{k <- k + 1
if(length(tmp)==0) COMB2[[k]] <- formula(paste("Y~1"))
if(length(tmp)>0) COMB2[[k]] <- formula(paste("Y", "~",
paste(tmp[,j], collapse=" + ")))
exlm0<-lm(COMB2[[k]],data=ex)
exlm<-summary(exlm0)
np<-length(tmp[,1])+1
R2p<-exlm$r.squared
R2ap<-exlm$adj.r.squared
SSEp<-deviance(aov(exlm0))
Cp<-SSEp/MSEF-(nn-2*(np))
AICp<-nn*log(SSEp)-nn*log(nn)+2*np
SBCp<-nn*log(SSEp)-nn*log(nn)+log(nn)*np
#PRESSp<-sum((resid(exlm0)/(1 - lm.influence(exlm0)$hat))^2)
res.table<-rbind(res.table,c(noquote(paste(tmp[,j], collapse=" + ")),
format(round(c(np,R2p,R2ap,Cp,AICp,BICp),3), nsmall = 3)))
res.table0<-rbind(res.table0,c(np,R2p,R2ap,Cp,AICp,SBCp))}
}
library(xtable)
res.table
```

■ Output (X_{1t}: unemployment rate X_{2t}: M2 growth rate X_{3t}: interest rate)

```
            [,1]                [,2]        [,3]        [,4]       [,5]       [,6]       [,7]
[1,] ""                 "  1.000" "  0.000" "  0.000" " 89.597" "146.111" "149.052"
[2,] "X1"               "  2.000" "  0.313" "  0.308" " 20.314" " 95.511" "101.395"
[3,] "X2"               "  2.000" "  0.037" "  0.030" " 83.201" "142.848" "148.731"
[4,] "X3"               "  2.000" "  0.156" "  0.150" " 56.026" "124.319" "130.202"
[5,] "X1 + X2"          "  3.000" "  0.401" "  0.393" "  2.268" "78.297"  "87.122"
[6,] "X1 + X3"          "  3.000" "  0.321" "  0.311" " 20.486" " 95.865" "104.690"
[7,] "X2 + X3"          "  3.000" "  0.179" "  0.167" " 52.762" "122.427" "131.252"
[8,] "X1 + X2 + X3"     "  4.000" "  0.402" "  0.389" "  4.000" "80.022"  "91.788"
```

No.	Variables	P	R^2	R_a^2	Cp	AIC	SBC
1	None	1	0	0	89.597	146.111	149.052
2	$X1$	2	0.313	0.308	20.314	95.511	101.395
3	$X2$	2	0.037	0.030	83.201	142.848	148.731
4	$X3$	2	0.156	0.150	56.026	124.319	130.202
5	$X1+X2$	3	0.401	0.393	2.268	78.297	87.122
6	$X1+X3$	3	0.321	0.311	20.486	95.865	104.690
7	$X2+X3$	3	0.179	0.167	52.762	122.427	131.252
8	$X1+X2+X3$	4	0.402	0.389	4.000	80.022	91.788

■ **Final regressions model**

$$Inflation_t = 6.339 - 1.131 \times \text{Unemployment}_t - 0.095 \times \text{M2 growth}_t$$

(p=0.000) (p=0.000)   (p=0.000)

$$Inflation_t = 6.091 - 1.086 \times \text{Unemployment}_t - 0.100 \times \text{M2 growth}_t$$

(p=0.000) (p=0.000)   (p=0.000)

$$+0.041 \times \text{Nominal interest rate}_t$$

(p=0.606)

Thank You for Your Time and Attention!