Part 2： Simultaneous-Equation Models (SEM)

Chapter 18. Why Should We Concern SEM ?

18.1 The Nature of Simultaneous-Equation Models

Definition and basic format of SEM

- **Simultaneous Equations Models (SEM)**: A system of equations consisting of several equations with interrelated or jointly influence.
- The basic and simple SEM is

$$\begin{cases} Y_{1i} = \beta_{10} + \gamma_{12} Y_{2i} + \beta_{11} X_{1i} + u_{i1} \\ Y_{2i} = \beta_{20} + \gamma_{21} Y_{1i} + \beta_{21} X_{1i} + u_{i2} \end{cases}$$

Example 1: Demand-and-Supply System

Demand-and-Supply System:

$$
\begin{cases}
\text{Demand function: } Q_t^d = \alpha_0 + \alpha_1 P_t + u_{1t}, \quad \alpha_1 < 0 \\
\text{Supply function: } Q_t^s = \beta_0 + \beta_1 P_t + u_{2t}, \quad \beta_1 > 0 \\
\text{Equilibrium condition: } Q_t^d = Q_t^s
\end{cases}
$$

Example 2: Keynesian Model of Income Determination

Keynesian Model of Income Determination:

$$\begin{cases} C_t = \beta_0 + \beta_1 Y_t + \varepsilon_t & \text{(consumption function)} \\ Y_t = C_t + I_t & \text{(income identity)} \end{cases}$$

Example 3: The IS Model

Macroeconomics goods market equilibrium model, also known as **IS Model**:

$$
\begin{cases}
\text{Consumption function: } C_t = \beta_0 + \beta_1 Y_{dt} + u_{1t} & < \beta_1 < 1 \\
\text{Tax function: } T_t = \alpha_0 + \alpha_1 Y_t + u_{2t} & 0 < \alpha_1 < 1 \\
\text{Investment function: } I_t = \gamma_0 + \gamma_1 r_t + u_{3t} & \\
\text{Definition: } \gamma_{dt} = Y_t - T_t & \\
\text{Government expenditure: } G_t = \bar{G} & \\
\text{National income identity: } Y_t = C_t + I_t + G_t &
\end{cases}
$$

> where:
>
> Y =national income; Y_d =disposable income; r =interest rate; $\bar{G}$ =given level of government expenditure

Example 4: The LM Model

Macroeconomics money market equilibrium system, also known as **LM Model**:

$$\begin{cases} \text{Money demand function: } M_t^d = a + bY_t - cr_t + u_t \\ \text{Money supply function: } M_t^s = \bar{M} \\ \text{Equilibrium condition: } M_t^d = M_t^s \end{cases}$$

> Where:
>
> Y =income; r =interest rate; $\bar{M}$ =assumed level of money supply.

Klein's model I:

$$\begin{cases} \text{Consumption function: } C_t = \beta_0 + \beta_1 P_t + \beta_2 \left(W + W'\right)_t + \beta_3 P_{t-1} + u_{t1} \\ \text{Investment function: } I_t = \beta_4 + \beta_5 P_t + \beta_6 P_{t-1} + \beta_7 K_{t-1} + u_{t2} \\ \text{Demand for labor: } w_t = \beta_8 + \beta_9 \left(Y + T - W'\right)_t + \beta_{10} \left(Y + T - W'\right)_{t-1} + \beta_{11} t + u_{t3} \\ \text{Identity: } Y_t = C_t + I_t + C_t \\ \text{Identity: } Y_t = W'_t + W_t + P_t \\ \text{Identity: } K_t = K_{t-1} + I_t \end{cases}$$

> Where:
>
> C =consumption expenditure; Y =income after tax; P =profits; W =private wage bill;
>
> W' =government wage bill; K =capital stock; T =taxes.

Example 6: Murder Rates and Size of the police Force

Cities often want to determine how much additional **law enforcement** will decrease their **murder rates**.

$$\begin{cases} \text{murdpc} = \alpha_1 \,\text{polpc} + \beta_{10} + \beta_{11}\text{incpc} + u_1 \\ \quad \text{polpc} \;\; = \alpha_2 \,\text{murdpc} + \beta_{20} + \text{ other factors.} \end{cases}$$

> Where :
>
> $murdpc$ =murders per capita; $polpc$ =number of police officers per capita; $incpc$ =income per capita.

Example 7: Housing Expenditures and Saving

For a random household in the population, we assume that annual **housing expenditures** and **saving** are jointly determined by:

$$\begin{cases} \text{housing} = \alpha_1 \text{saving} + \beta_{10} + \beta_{11}\text{inc} + \beta_{12}educ + \beta_{13}\text{age} + u_1 \\ \text{saving} = \alpha_2 \text{housing} + \beta_{20} + \beta_{21}\text{inc} + \beta_{22}educ + \beta_{23}\text{age} + u_2 \end{cases}$$

> Where:
>
> inc =annual income; $saving$ =household saving; $educ$ =education measured in years;
>
> age =age measured in years.

The Nature of SEM

The essence of simultaneous equation model is **endogenous variable** problem:

- Each of these equations has its economic **causality effect**.
- Some of these equations contain **endogenous variables**.
- Sample data is only the end result of various variables, which lies complex causal interaction behind them.
- Estimation all of the **parameters** directly by OLS method may induce problems.

Truffles example: the story

Truffles are delicious food materials. They are edible fungi that grow below the ground. Consider a supply and demand model for truffles:

$$\begin{cases} \text{Demand: } Q_{di} = \alpha_1 + \alpha_2 P_i + \alpha_3 PS_i + \alpha_4 DI_i + e_{di} \\ \text{Supply: } Q_{si} = \beta_1 + \beta_2 P_i + \beta_3 PF_i + e_{si} \\ \text{Equity: } Q_{di} = Q_{si} \end{cases}$$

> where:
>
> - Q_i =the quantity of truffles traded in a particular marketplace;
> - P_i =the market price of truffles;
> - PS_i =the market price of a substitute for real truffles;
> - DI_i =per capita monthly disposable income of local residents;
> - PF_i =the price of a factor of production, which in this case is the hourly rental price of truffle-pigs used in the search process.

Truffles example: model variables

All variables

vars	label	measure
P	market price of truffles	dollar/ounce
Q	market quantity of truffles	ounce
PS	market price of substitute	dollar/ounce
DI	disposable income	dollar/capita, monthly
PF	rental price of truffles-pigs	dollar/hour

Truffles example: the data set

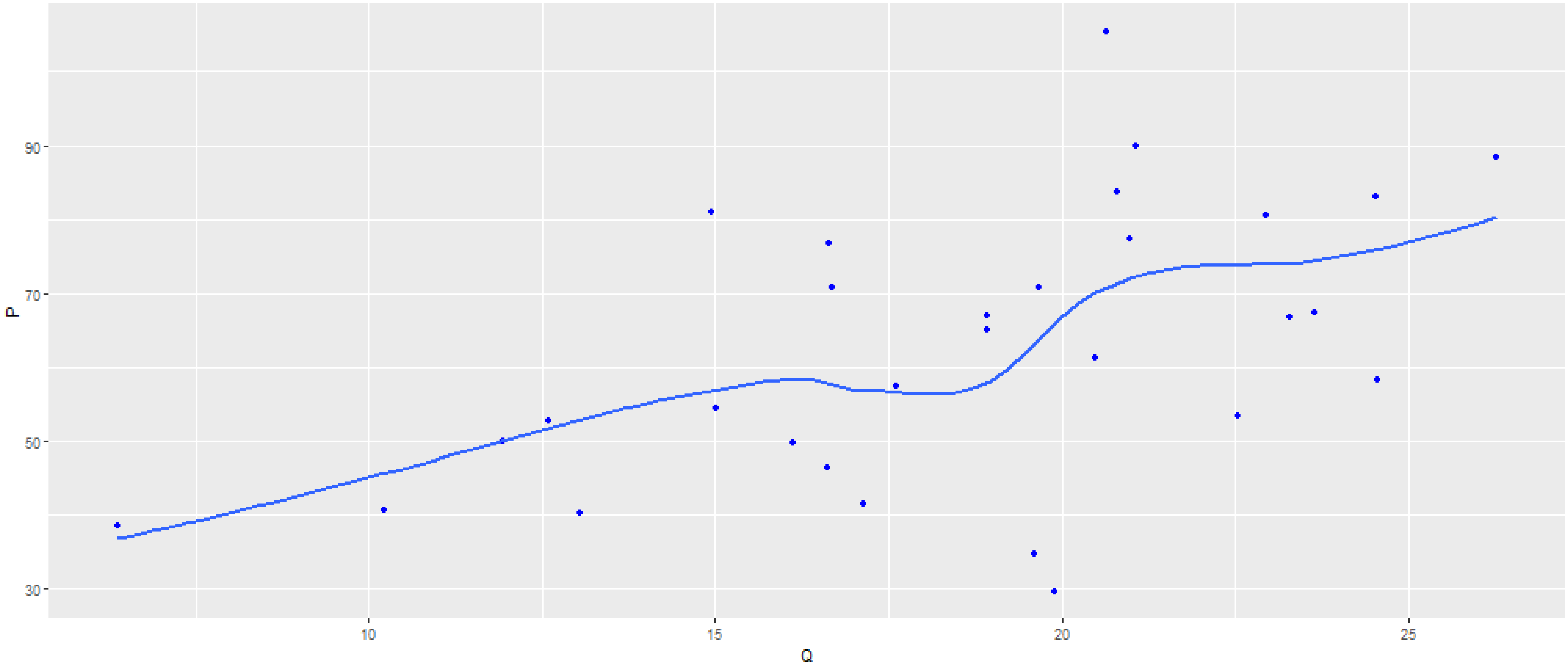
Truffles data set (n = 30)

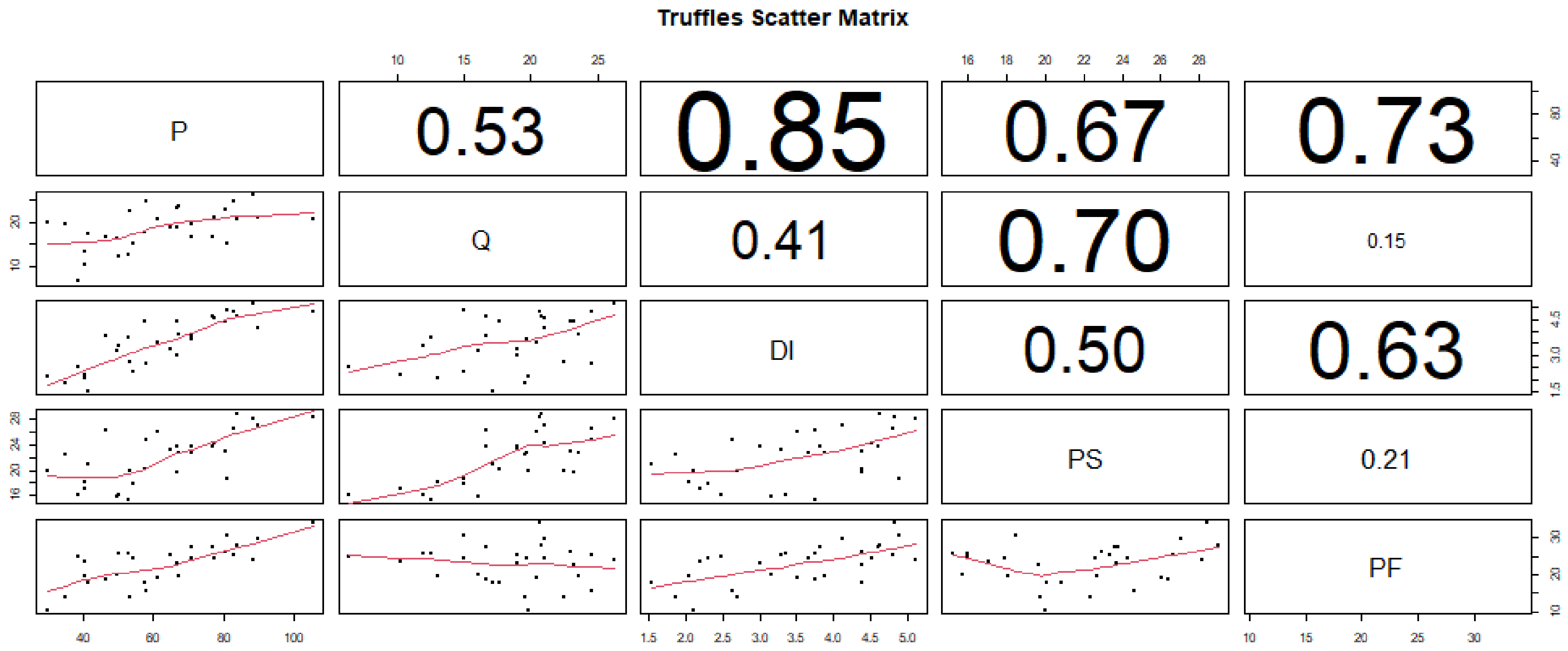
id	P	Q	PS	DI	PF
1	29.64	19.89	19.97	2.103	10.52
2	40.23	13.04	18.04	2.043	19.67
3	34.71	19.61	22.36	1.87	13.74
4	41.43	17.13	20.87	1.525	17.95
5	53.37	22.55	19.79	2.709	13.71
6	38.52	6.37	15.98	2.489	24.95
7	54.33	15.02	17.94	2.294	24.17
8	40.56	10.22	17.09	2.196	23.61

Showing 1 to 8 of 30 entries

Previous 1 2 3 4 Next

Truffles example: the Scatter plot (P VS Q)




Truffles Scatter Matrix
P
Q
DI
PS
PF
0.53
0.85
0.67
0.73
0.41
0.70
0.15
0.50
0.63
0.21

Let's start with the simplest linear regression model.

Generally, we use price (P) and output (Q) data to directly conduct simple linear regression modeling:

$$\begin{cases} P = \hat{\beta}_1 + \hat{\beta}_2 Q + e_1 & \text{(simple P model)} \\ Q = \hat{\beta}_1 + \hat{\beta}_2 P + e_2 & \text{(simple Q model)} \end{cases}$$

Truffles example: the simple linear regression

As we all know, the linear regression of two variables is asymmetrical, so there is:

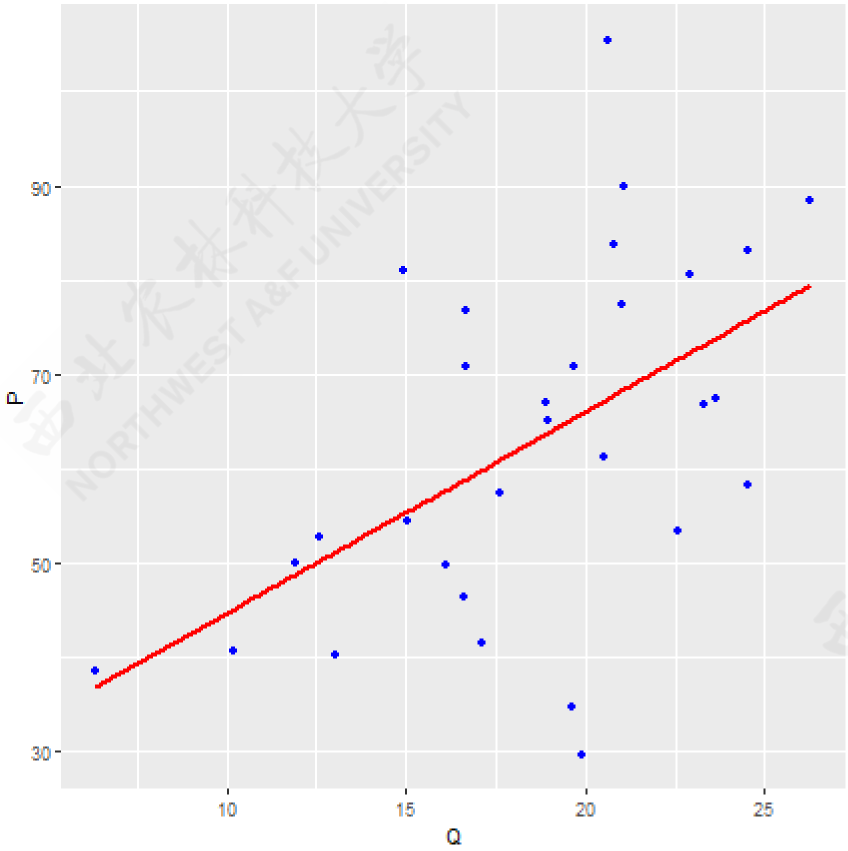
- the simple **Price** regression is:

$$
\begin{aligned}
\widehat{P} = & \quad +23.23 & \quad +2.14Q \\
(\text{t}) & \quad (1.8748) & \quad (3.2831) \\
(\text{se}) & \quad (12.3885) & \quad (0.6518) \\
(\text{fitness}) & R^2 = 0.2780; \bar{R}^2 = 0.2522 \\
& F^* = 10.78;\ p = 0.0028
\end{aligned}
$$

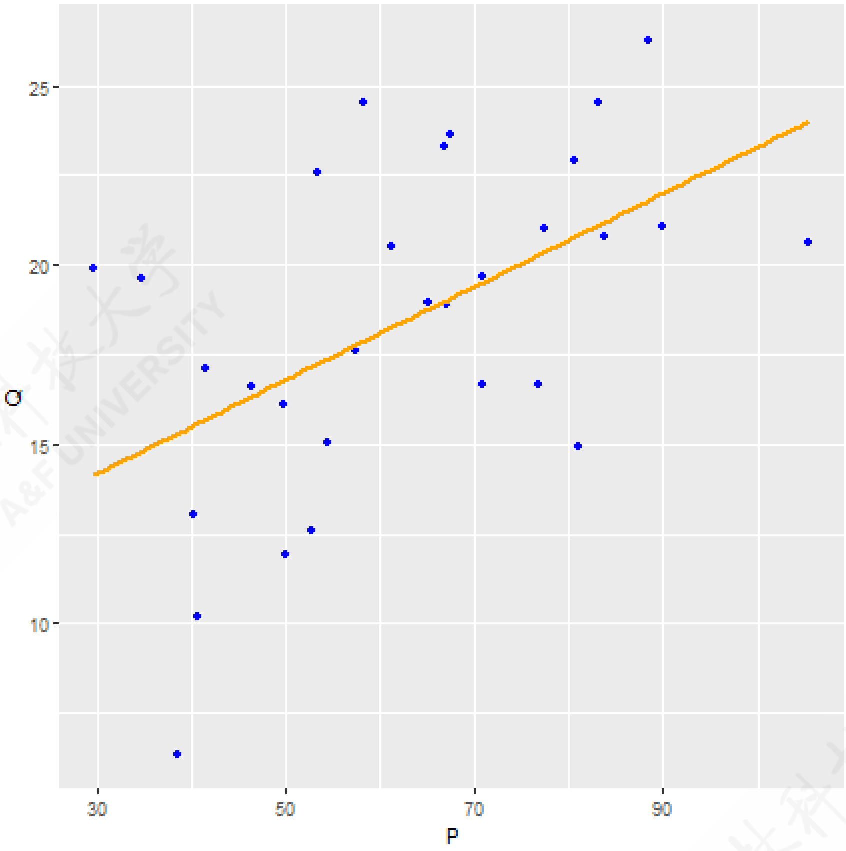
- the simple **Quantity** regression is:

$$
\begin{aligned}
\widehat{Q} = & \quad +10.31 & \quad +0.13P \\
(\text{t}) & \quad (3.9866) & \quad (3.2831) \\
(\text{se}) & \quad (2.5863) & \quad (0.0396) \\
(\text{fitness}) & R^2 = 0.2780; \bar{R}^2 = 0.2522 \\
& F^* = 10.78;\ p = 0.0028
\end{aligned}
$$

Truffles example: the sample regression line (SRL)



SRL of the Price model



SRL of the Quantity model

Truffles example: the multi-variables regression model

Of course, we can also use more independent variables X to build the regression models:

$$\begin{cases} P = \hat{\beta}_1 + \hat{\beta}_2 Q + \hat{\beta}_3 DI + \hat{\beta}_2 PS + e_1 & \text{(added P model)} \\ Q = \hat{\beta}_1 + \hat{\beta}_2 P + \hat{\beta}_2 PF + e_2 & \text{(added Q model)} \end{cases}$$

Truffles example: the multi-variables regression model

- the estimation result of multi-vars **Price** regression model is:

$$
\begin{aligned}
\widehat{P} = \quad & -13.62 & +0.15Q & +12.36DI + 1.36PS \\
\text{(t)} \quad & (-1.4987) & (0.3032) & (6.7701) \quad (2.2909) \\
\text{(se)} \quad & (9.0872) & (0.4988) & (1.8254) \quad (0.5940) \\
\text{(fitness)} & R^2 = 0.8013; \bar{R}^2 = 0.7784 \\
& F^* = 34.95;\ p = 0.0000
\end{aligned}
$$

- the estimation result of multi-vars **Quantity** regression model is:

$$
\begin{aligned}
\widehat{Q} = \quad & +20.03 & +0.34P & -1.00PF \\
\text{(t)} \quad & (16.3938) & (15.5436) & (-13.1028) \\
\text{(se)} \quad & (1.2220) & (0.0217) & (0.0764) \\
\text{(fitness)} & R^2 = 0.9019; \bar{R}^2 = 0.8946 \\
& F^* = 124.08; p = 0.0000
\end{aligned}
$$

18.2 Notations and Definitions

Structural SEM (1): algebraic expression A

Structural equations: System of equations that directly characterize economic structure or behavior.

The **algebraic expression** of structural SEM is:

$$
\begin{cases}
Y_{t1} = \quad + \gamma_{21}Y_{t2} + \cdots + \gamma_{m1}Y_{tm} + \beta_{11}X_{1t} + \beta_{21}X_{t2} + \cdots + \beta_{k1}X_{tk} + \varepsilon_{t1} \\
Y_{t2} = \gamma_{12}Y_{t1} + \quad \cdots + \gamma_{m2}Y_{tm} + \beta_{12}X_{t1} + \beta_{22}X_{t2} + \cdots + \beta_{k2}X_{tk} + \varepsilon_{t2} \\
\vdots \qquad \vdots \qquad \vdots \qquad \vdots \\
Y_{tm} = \gamma_{1m}Y_{t1} + \gamma_{2m}Y_{t2} + \cdots \quad + \beta_{1m}X_{t1} + \beta_{2m}X_{t2} + \cdots + \beta_{km}X_{tk} + \varepsilon_{tm}
\end{cases}
$$

Structural SEM (1): Structural coefficients

Structural coefficients: Parameters in structural equation that represents an economic outcome or behavioral relationship, including:

$$
\begin{cases}
Y_{t1} = \quad + \gamma_{21} Y_{t2} + \cdots + \gamma_{m1} Y_{tm} + \beta_{11} X_{1t} + \beta_{21} X_{t2} + \cdots + \beta_{k1} X_{tk} + \varepsilon_{t1} \\
Y_{t2} = \gamma_{12} Y_{t1} + \quad \cdots + \gamma_{m2} Y_{tm} + \beta_{12} X_{t1} + \beta_{22} X_{t2} + \cdots + \beta_{k2} X_{tk} + \varepsilon_{t2} \\
\vdots \quad \vdots \quad \vdots \quad \vdots \\
Y_{tm} = \gamma_{1m} Y_{t1} + \gamma_{2m} Y_{t2} + \cdots \quad + \beta_{1m} X_{t1} + \beta_{2m} X_{t2} + \cdots + \beta_{km} X_{tk} + \varepsilon_{tm}
\end{cases}
$$

- **Endogenous structural coefficients**: $\gamma_{11}, \gamma_{21}, \cdots, \gamma_{m1}; \cdots; \gamma_{1m}, \gamma_{2m}, \cdots, \gamma_{mm}$
- **Exogenous structural coefficients**: $\beta_{11}, \beta_{21}, \cdots, \beta_{m1}; \cdots; \beta_{1m}, \beta_{2m}, \cdots, \beta_{mm}$;

Structural SEM (1): Structural variables

- **Endogenous variables**: Variables determined by the model.
- **Predetermined variables**: Variables which values are not determined by the model in the **current** time period.

$$
\begin{cases}
Y_{t1} = \quad + \gamma_{21}Y_{t2} + \cdots + \gamma_{m1}Y_{tm} + \beta_{11}X_{1t} + \beta_{21}X_{t2} + \cdots + \beta_{k1}X_{tk} + \varepsilon_{t1} \\
Y_{t2} = \gamma_{12}Y_{t1} + \quad \cdots + \gamma_{m2}Y_{tm} + \beta_{12}X_{t1} + \beta_{22}X_{t2} + \cdots + \beta_{k2}X_{tk} + \varepsilon_{t2} \\
\vdots \quad \vdots \quad \vdots \quad \vdots \\
Y_{tm} = \gamma_{1m}Y_{t1} + \gamma_{2m}Y_{t2} + \cdots \quad + \beta_{1m}X_{t1} + \beta_{2m}X_{t2} + \cdots + \beta_{km}X_{tk} + \varepsilon_{tm}
\end{cases}
$$

Endogenous variables:

- Such as: $Y_{t1}; Y_{t2}; \cdots; Y_{tm}$

Predetermined variables:

- Such as: $X_{..}$

Predetermined variables: Variables which values are not determined by the model in the **current** time period, including:

- the **exogenous variables**
- the **lagged endogenous variables**.

$$\begin{cases} Y_{t1} = \quad\quad + \gamma_{21}Y_{t2} + \cdots + \gamma_{m1}Y_{tm} & +\beta_{11}X_{1t} + \beta_{21}X_{t2} + \cdots + \beta_{k1}X_{tk} & +\varepsilon_{t1} \\ Y_{t2} = \gamma_{12}Y_{t1} + \quad\quad \cdots + \gamma_{m2}Y_{tm} & +\beta_{12}X_{t1} + \beta_{22}X_{t2} + \cdots + \beta_{k2}X_{tk} & +\varepsilon_{t2} \\ \vdots \quad\quad \vdots \quad\quad \vdots & \vdots & \\ Y_{tm} = \gamma_{1m}Y_{t1} + \gamma_{2m}Y_{t2} + \cdots & +\beta_{1m}X_{t1} + \beta_{2m}X_{t2} + \cdots + \beta_{km}X_{tk} & +\varepsilon_{tm} \end{cases}$$

Structural SEM (1): Predetermined variables

- **Exogenous variables**: The variables not determined by the model, neither in the **current period** nor in the **lagged period**.
- **Lagged endogenous variables**: The lag variable of the endogenous variable in the current period。

current period exogenous:

- $X_{t1}, X_{t2}, \cdots, X_{tk}$.

lagged period exdogenous:

- lagged from X_{t1}:
  $X_{t-1,1}; X_{t-2,1}; \cdots; X_{t-(T-1),1}$
- and lagged from X_{tk}:
  $X_{t-1,k}; X_{t-2,k}; \cdots; X_{t-(T-1),k}$
- $\cdots$

lagged endogenous:

- lagged from Y_{t1}:
  $Y_{t-1,1}; Y_{t-2,1}; \cdots, Y_{t-(T-1),1}$
- and lagged from Y_{tm}:
  $Y_{t-1,m}; Y_{t-2,m}; \cdots; Y_{t-(T-1),m}$
- $\cdots$

Structural SEM (1): Predetermined coefficients

Predetermined coefficients: coefficients before predetermined variables.

$$
\begin{cases}
Y_{t1} = \quad + \gamma_{21}Y_{t2} + \cdots + \gamma_{m1}Y_{tm} \; + \beta_{11}X_{1t} + \beta_{21}X_{t2} + \cdots + \beta_{k1}X_{tk} \; + \varepsilon_{t1} \\
Y_{t2} = \gamma_{12}Y_{t1} + \quad \cdots + \gamma_{m2}Y_{tm} \; + \beta_{12}X_{t1} + \beta_{22}X_{t2} + \cdots + \beta_{k2}X_{tk} \; + \varepsilon_{t2} \\
\quad \vdots \qquad \vdots \qquad \vdots \qquad \vdots \\
Y_{tm} = \gamma_{1m}Y_{t1} + \gamma_{2m}Y_{t2} + \cdots \quad + \beta_{1m}X_{t1} + \beta_{2m}X_{t2} + \cdots + \beta_{km}X_{tk} + \varepsilon_{tm}
\end{cases}
$$

> Such as:
>
> - all $\beta_{..}$

By simple transformation, the **algebraic expression** of SEM can also show as:

$$A:\begin{cases} Y_{t1} = \quad + \gamma_{21}Y_{t2} + \cdots + \gamma_{m1}Y_{tm} + \beta_{11}X_{1t} + \beta_{21}X_{t2} + \cdots + \beta_{k1}X_{tk} + \varepsilon_{t1} \\ Y_{t2} = \gamma_{12}Y_{t1} + \quad \cdots + \gamma_{m2}Y_{tm} + \beta_{12}X_{t1} + \beta_{22}X_{t2} + \cdots + \beta_{k2}X_{tk} + \varepsilon_{t2} \\ \vdots \quad \vdots \quad \vdots \quad \vdots \\ Y_{tm} = \gamma_{1m}Y_{t1} + \gamma_{2m}Y_{t2} + \cdots \quad + \beta_{1m}X_{t1} + \beta_{2m}X_{t2} + \cdots + \beta_{km}X_{tk} + \varepsilon_{tm} \end{cases}$$

$$\Rightarrow B:\begin{cases} \gamma_{11}Y_{t1} + \gamma_{21}Y_{t2} + \cdots + \gamma_{m-1,1}Y_{t,m-1} + \gamma_{m1}Y_{tm} + \beta_{11}X_{t1} + \beta_{21}X_{t2} + \cdots + \beta_{k1}X_{tk} = \varepsilon_{t1} \\ \gamma_{12}Y_{t1} + \gamma_{22}Y_{t2} + \cdots + \gamma_{m-1,2}Y_{t,m-1} + \gamma_{m2}Y_{tm} + \beta_{12}X_{t1} + \beta_{22}X_{t2} + \cdots + \beta_{k2}X_{tk} = \varepsilon_{t2} \\ \vdots \quad \vdots \quad \vdots \quad \vdots \\ \gamma_{1m}Y_{t1} + \gamma_{2m}Y_{t2} + \cdots + \gamma_{m-1,m}Y_{t,m-1} + \gamma_{mm}Y_{tm} + \beta_{1m}X_{t1} + \beta_{2m}X_{t2} + \cdots + \beta_{km}X_{tk} = \varepsilon_{tm} \end{cases}$$

Structural SEM (2): matrix expression

With the Matrix language, the **matrix expression** of SEM was noted as:

$$
\begin{bmatrix} Y_1 & Y_2 & \cdots & Y_m \end{bmatrix}_t
\begin{bmatrix}
\gamma_{11} & \gamma_{12} & \cdots & \gamma_{1m} \\
\gamma_{21} & \gamma_{22} & \cdots & \gamma_{2m} \\
\cdots & \cdots & \cdots & \cdots \\
\gamma_{m1} & \gamma_{m2} & \cdots & \gamma_{mm}
\end{bmatrix} +
$$

$$
\begin{bmatrix} X_1 & X_2 & \cdots & X_m \end{bmatrix}_t
\begin{bmatrix}
\beta_{11} & \beta_{12} & \cdots & \beta_{1m} \\
\beta_{21} & \beta_{22} & \cdots & \beta_{2m} \\
\cdots & \cdots & \cdots & \cdots \\
\beta_{k1} & \beta_{k2} & \cdots & \beta_{km}
\end{bmatrix}
$$

$$
= \begin{bmatrix} \varepsilon_1 & \varepsilon_2 & \cdots & \varepsilon_m \end{bmatrix}_t
$$

Structural SEM (2): matrix expression

For Simplicity, we can generized the **matrix expression** of SEM :

$$
\begin{array}{ccc}
\boldsymbol{y}_t'\boldsymbol{\Gamma} & +\boldsymbol{x}_t'\boldsymbol{B} & =\boldsymbol{\varepsilon}_t' \\
(1*m)(m*m) & (1*k)(k*m) & (1*m)
\end{array}
$$

> **where:**
>
> - Bold upper letter and greek means a **matrix**
> - Bold lower letter and greek means a **column vector**

Structural SEM (2): Endogenous coefficients matrix

For the **Endogenous parameter matrix $\boldsymbol{\Gamma}$**:

- To ensure that each equation has a **dependent variable**, then the matrix $\boldsymbol{\Gamma}$ each column has at least one element of 1
- If matrix $\boldsymbol{\Gamma}$ is upper triangular matrix, then the SEM is a **recursive** model system.
- For the SEM solution to exist, $\boldsymbol{\Gamma}$ must be **nonsingular**.

$$\boldsymbol{\Gamma}=\begin{bmatrix}\gamma_{11} & \gamma_{12} & \cdots & \gamma_{1m}\\ \gamma_{21} & \gamma_{22} & \cdots & \gamma_{2m}\\ \cdots & \cdots & \cdots & \cdots\\ \gamma_{m1} & \gamma_{m2} & \cdots & \gamma_{mm}\end{bmatrix}$$

$$\text{if } \Rightarrow \begin{bmatrix}\gamma_{11} & \gamma_{12} & \cdots & \gamma_{1m}\\ 0 & \gamma_{22} & \cdots & \gamma_{2m}\\ \cdots & \cdots & \cdots & \cdots\\ 0 & 0 & \cdots & \gamma_{mm}\end{bmatrix}$$

$$\begin{cases} y_{1t} = & f_1\left(\mathbf{x}_t\right)+\varepsilon_{t1}\\ y_{2t} = & f_2\left(y_{t1},\mathbf{x}_t\right)+\varepsilon_{t2}\\ \vdots & \vdots\\ y_{mt} = & f_m\left(y_{t1},y_{t2},\ldots,\mathbf{x}_t\right)+\varepsilon_{mt}\end{cases}$$

Structural SEM (2): Exdogenous coefficients matrix

The **Exogenous coefficients matrix** $\boldsymbol{B}$:

$$
\boldsymbol{B}=\begin{bmatrix}
\beta_{11} & \beta_{12} & \cdots & \beta_{1m} \\
\beta_{21} & \beta_{22} & \cdots & \beta_{2m} \\
\cdots & \cdots & \cdots & \cdots \\
\beta_{k1} & \beta_{k2} & \cdots & \beta_{km}
\end{bmatrix}
$$

Reduced SEM (1): algebraic expression

Reduced equations: The equation expresses an **endogenous variable** with all the **predetermined variables** and the **stochastic disturbances**.

$$
\begin{cases}
Y_{t1} = +\pi_{11}X_{t1} + \pi_{21}X_{t2} + \cdots + \pi_{k1}X_{tk} + v_{t1} \\
Y_{t2} = +\pi_{12}X_{t1} + \pi_{22}X_{t2} + \cdots + \pi_{k2}X_{tk} + v_{t2} \\
\quad \vdots \qquad\qquad\qquad \vdots \qquad\qquad\qquad \vdots \\
Y_{tm} = +\pi_{1m}X_{t1} + \pi_{2m}X_{t2} + \cdots + \pi_{km}X_{tk} + v_{tm}
\end{cases}
$$

Reduced SEM (1): Reduced coefficients and disturbance

- **Reduced coefficients**: parameters in the reduced SEM.
- **Reduced disturbance**: stochastic disturbance terms in the reduced SEM.

$$
\begin{cases}
Y_{t1} = +\pi_{11}X_{t1} + \pi_{21}X_{t2} + \cdots + \pi_{k1}X_{tk} + v_{t1} \\
Y_{t2} = +\pi_{12}X_{t1} + \pi_{22}X_{t2} + \cdots + \pi_{k2}X_{tk} + v_{t2} \\
\vdots \qquad\qquad \vdots \qquad\qquad \vdots \\
Y_{tm} = +\pi_{1m}X_{t1} + \pi_{2m}X_{t2} + \cdots + \pi_{km}X_{tk} + v_{tm}
\end{cases}
$$

Reduced coefficients:

- $\pi_{11}, \pi_{21}, \cdots, \pi_{k1}$
- $\pi_{1m}, \pi_{2m}, \cdots, \pi_{km}$.

Reduced disturbance:

- $v_1, v_2, \cdots, v_m$。

Reduced SEM (2): matrix expression

$$
\begin{cases}
Y_{t1} = +\pi_{11}X_{t1} + \pi_{21}X_{t2} + \cdots + \pi_{k1}X_{tk} + v_{t1} \\
Y_{t2} = +\pi_{12}X_{t1} + \pi_{22}X_{t2} + \cdots + \pi_{k2}X_{tk} + v_{t2} \\
\vdots \qquad\qquad \vdots \qquad\qquad \vdots \\
Y_{tm} = +\pi_{1m}X_{t1} + \pi_{2m}X_{t2} + \cdots + \pi_{km}X_{tk} + v_{tm}
\end{cases}
$$

For this algebraic reduced SEM, we can note its matrix form as:

$$
\begin{bmatrix} Y_1 & Y_2 & \cdots & Y_m \end{bmatrix}_t =
$$

$$
\begin{bmatrix} X_1 & X_2 & \cdots & X_m \end{bmatrix}_t
\begin{bmatrix}
\pi_{11} & \pi_{12} & \cdots & \pi_{1m} \\
\pi_{21} & \pi_{22} & \cdots & \pi_{2m} \\
\cdots & \cdots & \cdots & \cdots \\
\pi_{m1} & \pi_{m2} & \cdots & \pi_{mm}
\end{bmatrix}
+ \begin{bmatrix} v_1 & v_2 & \cdots & v_m \end{bmatrix}_t
$$

Reduced SEM (2): matrix expression

For simplicity, the matrix expression of reduced SEM can be noted further.

$$\begin{array}{ccc} \boldsymbol{y}_t' & =\boldsymbol{x}_t'\boldsymbol{\Pi} & +\boldsymbol{v}_t' \\ (1*m) & (1*k)(k*m) & (1*m) \end{array}$$

- the reduced coefficients matrix is :

$$\boldsymbol{\Pi} = \begin{bmatrix} \pi_{11} & \pi_{12} & \cdots & \pi_{1m} \\ \pi_{21} & \pi_{22} & \cdots & \pi_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ \pi_{m1} & \pi_{m2} & \cdots & \pi_{mm} \end{bmatrix}$$

- the reduced disturbances vector is:

$$\boldsymbol{v}_t' = \begin{bmatrix} v_1 & v_2 & \cdots & v_m \end{bmatrix}_t$$

Structural VS Reduced SEM: the two systems

We can induce Reduced Equations from Structural Equations:

$$
\begin{cases}
Y_{t1} = & + \gamma_{21} Y_{t2} + \cdots + \gamma_{m1} Y_{tm} & + \beta_{11} X_{1t} + \beta_{21} X_{t2} + \cdots + \beta_{k1} X_{tk} & + \varepsilon_{t1} \\
Y_{t2} = \gamma_{12} Y_{t1} + & \cdots + \gamma_{m2} Y_{tm} & + \beta_{12} X_{t1} + \beta_{22} X_{t2} + \cdots + \beta_{k2} X_{tk} & + \varepsilon_{t2} \\
\vdots \qquad \vdots & \vdots & \vdots & \\
Y_{tm} = \gamma_{1m} Y_{t1} + \gamma_{2m} Y_{t2} + \cdots & & + \beta_{1m} X_{t1} + \beta_{2m} X_{t2} + \cdots + \beta_{km} X_{tk} & + \varepsilon_{tm}
\end{cases}
$$

$$
\Rightarrow
\begin{cases}
Y_{t1} = & + \pi_{11} X_{t1} + \pi_{21} X_{t2} + \cdots + \pi_{k1} X_{tk} & + v_{t1} \\
Y_{t2} = & + \pi_{12} X_{t1} + \pi_{22} X_{t2} + \cdots + \pi_{k2} X_{tk} & + v_{t2} \\
\vdots & \vdots & \vdots \\
Y_{tm} = & + \pi_{1m} X_{t1} + \pi_{2m} X_{t2} + \cdots + \pi_{km} X_{tk} & + v_{tm}
\end{cases}
$$

Structural VS Reduced SEM: coefficients

The Structural SEM :

$$\boldsymbol{y}_t'\boldsymbol{\Gamma} + \boldsymbol{x}_t'\boldsymbol{B} = \boldsymbol{\varepsilon}_t'$$

The Reduced SEM:

$$\boldsymbol{y}_t' = \boldsymbol{x}_t'\boldsymbol{\Pi} + \boldsymbol{v}_t'$$

- where:

$$\begin{aligned} \boldsymbol{\Pi} &= -\boldsymbol{B}\boldsymbol{\Gamma}^{-1} \\ \boldsymbol{v}_t' &= \boldsymbol{\varepsilon}_t'\boldsymbol{\Gamma}^{-1} \end{aligned}$$

- and:

$$\boldsymbol{\Gamma} = \begin{bmatrix} \gamma_{11} & \gamma_{12} & \cdots & \gamma_{1m} \\ \gamma_{21} & \gamma_{22} & \cdots & \gamma_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ \gamma_{m1} & \gamma_{m2} & \cdots & \gamma_{mm} \end{bmatrix}$$

Structural VS Reduced SEM: Moments

Now we concern the first and second moments of the disturbance:

- first , let us assumed the moments of **structural disturbances** satisfy:

$$
\begin{aligned}
\mathbf{E}[\varepsilon_{\mathbf{t}}|\mathbf{x}_{\mathbf{t}}] &= \mathbf{0} \\
\mathbf{E}[\varepsilon_{\mathbf{t}}\varepsilon_{\mathbf{t}}'|\mathbf{x}_{\mathbf{t}}] &= \boldsymbol{\Sigma} \\
E\left[\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_s'|\mathbf{x}_t, \mathbf{x}_s\right] &= \mathbf{0}, \quad \forall t, s
\end{aligned}
$$

- then, we can prove that the **reduced disturbances** satify:

$$
\begin{aligned}
E\left[\mathbf{v}_t|\mathbf{x}_t\right] &= \left(\boldsymbol{\Gamma}^{-1}\right)'\mathbf{0} = \mathbf{0} \\
E\left[\mathbf{v}_t\mathbf{v}_t'|\mathbf{x}_t\right] &= \left(\boldsymbol{\Gamma}^{-1}\right)'\boldsymbol{\Sigma}\boldsymbol{\Gamma}^{-1} = \boldsymbol{\Omega} \\
\text{where: } \boldsymbol{\Sigma} &= \boldsymbol{\Gamma}'\boldsymbol{\Omega}\boldsymbol{\Gamma}
\end{aligned}
$$

Structural VS Reduced SEM: useful expression*

In a sample of data, each joint observation will be one row in a data matrix (with T observations):

$$[\mathbf{Y} \quad \mathbf{X} \quad \mathbf{E}] = \begin{bmatrix} \mathbf{y}_1' & \mathbf{x}_1' & \boldsymbol{\varepsilon}_1' \\ \mathbf{y}_2' & \mathbf{x}_2' & \boldsymbol{\varepsilon}_2' \\ \vdots & & \\ \mathbf{y}_T' & \mathbf{x}_T' & \boldsymbol{\varepsilon}_T' \end{bmatrix}$$

then the structural SEM is:

$$\mathbf{Y\Gamma} + \mathbf{XB} = \mathbf{E}$$

the first and second moment of structural disturbances is:

$$\begin{aligned} E[\mathbf{E}|\mathbf{X}] &= \mathbf{0} \\ E\left[(1/T)\mathbf{E}'\mathbf{E}|\mathbf{X}\right] &= \mathbf{\Sigma} \end{aligned}$$

Structural VS Reduced SEM: useful expression*

Assume that:

$$
\begin{aligned}
&(1/T)\mathbf{X}'\mathbf{X} \rightarrow \mathbf{Q} \text{ (a finite positive definite matrix)} \\
&(1/T)\mathbf{X}'\mathbf{E} \rightarrow \mathbf{0}
\end{aligned}
$$

then the reduced SEM can be noted as:

$$
\mathbf{Y} = \mathbf{X\Pi} + \mathbf{V} \qquad \leftarrow \mathbf{V} = \mathbf{E\Gamma}^{-1}
$$

And we may have following useful results:

$$
\frac{1}{T}\begin{bmatrix} \mathbf{Y}' \\ \mathbf{X}' \\ \mathbf{V}' \end{bmatrix} \begin{bmatrix} \mathbf{Y} & \mathbf{X} & \mathbf{V} \end{bmatrix} \rightarrow \begin{bmatrix} \mathbf{I}'\mathbf{QI} + \mathbf{\Omega} & \mathbf{II}'\mathbf{Q} & \mathbf{\Omega} \\ \mathbf{QI} & \mathbf{Q} & \mathbf{0}' \\ \mathbf{\Omega} & \mathbf{0} & \mathbf{\Omega} \end{bmatrix}
$$

Case 1: Keynesian income model (structural SEM)

The Keynesian model of income determination (structural SEM):

$$\begin{cases} C_t = \beta_0 + \beta_1 Y_t + \varepsilon_t & \text{(consumption function)} \\ Y_t = C_t + I_t & \text{(income equity)} \end{cases}$$

So the structural SEM contains:

2 **endogenous variables**:

- $c_t; Y_t$

1 **predetermined variables**:

- 1 **exogenous variables**: I_t
- 0 **lagged endogenous variable**.

> **Exercise**: can you get the reduced SEM from this structural SEM ?

Case 1: Keynesian income model (reduced SEM)

We can get the reduced SEM from the former structural SEM and denoted (the right):

$$\begin{cases} Y_t = \dfrac{\beta_0}{1-\beta_1} + \dfrac{1}{1-\beta_1} I_t + \dfrac{\varepsilon_t}{1-\beta_1} \\ C_t = \dfrac{\beta_0}{1-\beta_1} + \dfrac{\beta_1}{1-\beta_1} I_t + \dfrac{\varepsilon_t}{1-\beta_1} \end{cases}$$

$$\begin{cases} Y_t = \pi_{11} + \pi_{21} I_t + v_{t1} \\ C_t = \pi_{12} + \pi_{22} I_t + v_{t2} \end{cases}$$

where:

$$\begin{cases} \pi_{11} = \dfrac{\beta_0}{1-\beta_1}; \quad \pi_{21} = \dfrac{\beta_0}{1-\beta_1}; \quad v_{t1} = \dfrac{\varepsilon_t}{1-\beta_1}; \\ \pi_{12} = \dfrac{1}{1-\beta_1}; \quad \pi_{22} = \dfrac{\beta_1}{1-\beta_1}; \quad v_{t2} = \dfrac{\varepsilon_t}{1-\beta_1}; \end{cases}$$

> there are 2 **structural coefficients** $\beta_0; \beta_1$ totally ; and 4 **reduced coefficients** $\pi_{11}, \pi_{21}; \pi_{12}, \pi_{22}$ (There are actually three only !)

Case 2: Macroeconomic Model (structural SEM)

Consider the **Small Macroeconomic Model** (Structural SEM):

$$\begin{cases} \text{consumption: } c_t = \alpha_0 + \alpha_1 y_t + \alpha_2 c_{t-1} + \varepsilon_{t,c} \\ \text{investment: } i_t = \beta_0 + \beta_1 r_t + \beta_2 (y_t - y_{t-1}) + \varepsilon_{t,j} \\ \text{demand: } y_t = c_t + i_t + g_t \end{cases}$$

where: c_t = consumption; y_t = output; i_t = investment; r_t = rate; g_t = government expenditure.

3 **endogenous variables**: $c_t; i_t; Y_t$

4 **predetermined variables**:

- 2 **exogenous variables**: $r_t; g_t$.
- 2 **lagged endogenous variables**: $y_{t-1}; c_{t-1}$

totally 6 **strutural coefficients**: $\alpha_0, \alpha_1, \alpha_2; \beta_0, \beta_1, \beta_2;$

Case 2: Macroeconomic Model（reduced SEM）

We can get the reduced SEM from the former structural SEM: (HOW TO??)

$$
\begin{cases}
c_t = & [\alpha_0(1-\beta_2)+\beta_0\alpha_1+\alpha_1\beta_1 r_t+\alpha_1 g_t+\alpha_2\left(1-\beta_2\right)c_{t-1}-\alpha_1\beta_2 y_{t-1} \\
 & +(1-\beta_2)\,\varepsilon_{t,c}+\alpha_1\varepsilon_{t,j}]/\Lambda \\
i_t = & [\alpha_0\beta_2+\beta_0\left(1-\alpha_1\right)+\beta_1\left(1-\alpha_1\right)r_t+\beta_2 g_t+\alpha_2\beta_2 c_{t-1}-\beta_2\left(1-\alpha_1\right)y_{t-1} \\
 & +\beta_2\varepsilon_{t,c}+\left(1-\alpha_1\right)\varepsilon_{t,j}]/\Lambda \\
y_t = & [\alpha_0+\beta_0+\beta_1 r_t+g_t+\alpha_2 c_{t-1}-\beta_2 y_{t-1}+\varepsilon_{t,c}+\varepsilon_{t,j}]/\Lambda
\end{cases}
$$

where: $\Lambda = 1-\alpha_1-\beta_2$。 For simplicity, denote the **reduced SEM** as:

$$
\begin{cases}
c_t = \pi_{11}+\pi_{21}r_t+\pi_{31}g_t+\pi_{41}c_{t-1}+\pi_{51}y_{t-1}+v_{t1} \\
i_t = \pi_{12}+\pi_{22}r_t+\pi_{32}g_t+\pi_{42}c_{t-1}+\pi_{52}y_{t-1}+v_{t2} \\
i_t = \pi_{13}+\pi_{23}r_t+\pi_{33}g_t+\pi_{43}c_{t-1}+\pi_{53}y_{t-1}+v_{t3}
\end{cases}
$$

> So we have 15 **reduced coefficients** totally!

Thinking:

- What are the purposes of structural SEM and reduced SEM respectively?
- Note the consumption function (in structural SEM): the rate i_t does not impact the consumption c_t!
  - It will be obvious from the reduced SEM that $\frac{\Delta c_t}{\Delta r_t} = \frac{\alpha_1 \beta_1}{\Lambda}$
- Note the consumption function (in structural SEM): What are the reasons for the impact of income y_t on consumption c_t?
  - It's also easy to get the answer by transformation:
    $$\frac{\Delta c_t}{\Delta y_t} = \frac{\Delta c_t / \Delta r_t}{\Delta y_t / \Delta r_t} = \frac{\alpha_1 \beta_1 / \Lambda}{\beta_1 / \Lambda} = \alpha_1$$

According to the relationship between Structural SEM and Reduced SEM:

$$\boldsymbol{y}_t' = -\,\boldsymbol{x}_t'\boldsymbol{\Pi} + \boldsymbol{v}_t' = -\boldsymbol{x}_t'\boldsymbol{B}\boldsymbol{\Gamma}^{-1} + \boldsymbol{\varepsilon}_t'\boldsymbol{\Gamma}^{-1}$$

Then, the following matrixes can be easily obtained:

$$\begin{aligned}\mathbf{y}' &= \begin{bmatrix} c & i & y \end{bmatrix}\\ \boldsymbol{x}' &= \begin{bmatrix} 1 & r & g & c_{-1} & y_{-1} \end{bmatrix}\end{aligned}$$

$$\mathbf{B} = \begin{bmatrix} -\alpha_0 & -\beta_0 & 0 \\ 0 & -\beta_1 & 0 \\ 0 & 0 & -1 \\ -\alpha_2 & 0 & 0 \\ 0 & \beta_2 & 0 \end{bmatrix}$$

$$\Gamma = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ -\alpha_1 & -\beta_2 & 1 \end{bmatrix}$$

$$\boldsymbol{\Gamma}^{-1} = \frac{1}{\Lambda}\begin{bmatrix} 1-\beta_2 & \beta_2 & 1 \\ \alpha_1 & 1-\alpha_1 & 1 \\ \alpha_1 & \beta_2 & 1 \end{bmatrix}$$

We can get the same answers: (It's so easy!)

$$\mathbf{\Pi} = -\boldsymbol{B}\mathbf{\Gamma}^{-1} = \frac{1}{\Lambda}\begin{bmatrix} \alpha_0(1-\beta_2)+\beta_0\alpha_1 & \alpha_0\beta_2+\beta_0(1-\alpha_1) & \alpha_0+\beta_0 \\ \alpha_1\beta_1 & \beta_1(1-\alpha_1) & \beta_1 \\ \alpha_1 & \beta_2 & 1 \\ \alpha_2(1-\beta_2) & \alpha_2\beta_2 & \alpha_2 \\ -\beta_2\alpha_1 & -\beta_2(1-\alpha_1) & -\beta_2 \end{bmatrix}$$

$$\mathbf{\Pi}' = \frac{1}{\Lambda}\begin{bmatrix} \alpha_0(1-\beta_2)+\beta_0\alpha_1 & \alpha_1\beta_1 & \alpha_1 & \alpha_2(1-\beta_2) & -\beta_2\alpha_1 \\ \alpha_0\beta_2+\beta_0(1-\alpha_1) & \beta_1(1-\alpha_1) & \beta_2 & \alpha_2\beta_2 & -\beta_2(1-\alpha_1) \\ \alpha_0+\beta_0 & \beta_1 & 1 & \alpha_2 & -\beta_2 \end{bmatrix}$$

- Where:

$$\Lambda = 1-\alpha_1-\beta_2$$

- Remeber that:

$$\mathbf{x}' = \begin{bmatrix} 1 & r & g & c_{-1} & y_{-1} \end{bmatrix}$$

Supplement: inverse matrix solution and procedure*

Use the elementary row operation (Gauss-Jordan) to find the inverse matrix:

1. Construct **augmented matrix**
2. Transform the augmented matrix for many times until the goal is achieved.

Use cofactor, algebraic cofactor and adjoint matrix to get the inverse matrix:

1. Calculate **cofactor matrix** and **algebraic cofactor matrix**;
2. Calculate **adjoint matrix**: it is the **transpose** of the cofactor matrix;
3. Calculate the **determinant** of original matrix : each element of **top row** in the original matrix is multiplied by its corresponding **top row** element in the "cofactor matrix";
4. Calculated the inverse matrix: **1/ determinant** × **adjoint matrix**

18.3 Is the OLS Method Still applicable ?

Endogenous variable problem

Consider Keynes's model of income determination, We will be able to show that Y_t and u_t will be correlated, thus violating the CLRM **A2** assumption.

$$\begin{cases} C_t = \beta_0 + \beta_1 Y_t + u_t \qquad (0 < \beta_1 < 1) \qquad \text{(consumption function)} \\ Y_t = C_t + I_t \qquad \text{(Income Identity)} \end{cases}$$

By transforming the above structural equation, we obtained:

$$Y_t = \beta_0 + \beta_1 Y_t + I_t + u_t$$

$$Y_t = \frac{\beta_0}{1-\beta_1} + \frac{1}{1-\beta_1} I_t + \frac{1}{1-\beta_1} u_t \qquad \text{(eq1: Reduced equation)}$$

$$E(Y_t) = \frac{\beta_0}{1-\beta_1} + \frac{1}{1-\beta_1} I_t \qquad \text{(eq2: Take the expectation for both sides)}$$

Endogenous variable problem

Further more:

$$\begin{aligned} Y_t - E(Y_t) &= \frac{u_t}{1-\beta_1} && \text{(eq 1 - eq 2)} \\ u_t - E(u_t) &= u_t && \text{(eq 3: Expectation is equal to 0)} \\ cov(Y_t, u_t) &= E([Y_t - E(Y_t)][u_t - E(u_t)]) && \text{(eq 4: Covariance definition)} \\ &= \frac{E(u_t^2)}{1-\beta_1} && \text{(eq 5: Variance definition)} \\ &= \frac{\sigma^2}{1-\beta_1} \neq 0 && \text{(eq 6: The variance is not 0)} \end{aligned}$$

- Therefore, the consumption equation of the Keynesian model will not satisfy the CLRM A2 assumption.
- Thus, OLS method cannot be used to obtain **Best linear unbiased estimator** (BLUE) for consumption equation.

The OLS estimator of the coefficient is biased

Furthermore, the OLS estimator is biased to its true β_1, which means $E(\hat{\beta}_1) \neq \beta_1$. The proof show as below.

$$\begin{cases} C_t = \beta_0 + \beta_1 Y_t + u_t & (0 < \beta_1 < 1) \quad \text{(consumption function)} \\ Y_t = C_t + I_t & \text{(Income Identity)} \end{cases}$$

$$\hat{\beta}_1 = \frac{\sum c_t y_t}{\sum y_t^2} = \frac{\sum C_t y_t}{\sum y_t^2} = \frac{\sum [(\beta_0 + \beta_1 Y_t + u_t) y_t]}{\sum y_t^2} = \beta_1 + \frac{\sum u_t y_t}{\sum y_t^2} \qquad \text{(eq 1)}$$

Take the expectation of both sides in eq 1, so:

$$E(\hat{\beta}_1) = \beta_1 + E\left(\frac{\sum u_t y_t}{\sum y_t^2}\right)$$

> Question: is the expactation $E\left(\frac{\sum u_t y_t}{\sum y_t^2}\right)$ equal to zero?

Supplement: Proof 1/2

$$\frac{\sum c_t y_t}{\sum y_t^2} = \frac{\sum (C_t - \bar{C})(Y_t - \bar{Y})}{\sum y_t^2} = \frac{\sum (C_t - \bar{C}) y_t}{\sum y_t^2}$$

$$= \frac{\sum C_t y_t - \sum \bar{C} y_t}{\sum y_t^2} = \frac{\sum C_t y_t - \sum \bar{C}(Y_t - \bar{Y})}{\sum y_t^2}$$

$$= \frac{\sum C_t y_t - \bar{C}\sum Y_t - \sum \bar{C}\bar{Y}}{\sum y_t^2} = \frac{\sum C_t y_t - \bar{C}\sum Y_t - n\bar{C}\bar{Y}}{\sum y_t^2} = \frac{\sum C_t y_t}{\sum y_t^2}$$

$$\hat{\beta}_1 = \frac{\sum (\beta_0 + \beta_1 Y_t + u_t)\, y_t}{\sum y_t^2} = \frac{\sum \beta_0 y_t + \sum \beta_1 Y_t y_t + \sum u_t y_t}{\sum y_t^2}$$

$$= \frac{\beta_1 \sum (y_t + \bar{Y}) y_t + \sum u_t y_t}{\sum y_t^2} = \beta_1 + \frac{\sum y_t u_t}{\sum y_t^2}$$

$$\Leftarrow \sum y_t = 0; \qquad \frac{\sum Y_t y_t}{y_t^2} = 1$$

Conduct the limit to probability:

$$\begin{aligned}\text{plim}\left(\hat{\beta}_1\right) &= \text{plim}(\beta_1) + \text{plim}\left(\frac{\sum y_t u_t}{\sum y_t^2}\right) \\ &= \text{plim}(\beta_1) + \text{plim}\left(\frac{\sum y_t u_t / n}{\sum y_t^2 / n}\right) = \beta_1 + \frac{\text{plim}(\sum y_t u_t / n)}{\text{plim}\left(\sum y_t^2 / n\right)}\end{aligned}$$

And we've shown that:

$$cov(Y_t, u_t) = E([Y_t - E(Y_t)][u_t - E(u_t)]) = \frac{E(u_t^2)}{1 - \beta_1} = \frac{\sigma^2}{1 - \beta_1} \neq 0$$

Therefore we finaly prove: $E\left(\frac{\sum u_t y_t}{\sum y_t^2}\right) \neq 0$

Simulation: artificially population

Here, we construct an artificially controlled population for our Keynes's SEM model.

$$\begin{cases} C_t = \beta_0 + \beta_1 Y_t + u_t & (0 < \beta_1 < 1) & \text{(consumption function)} \\ Y_t = C_t + I_t & & \text{(Income Identity)} \end{cases}$$

$$\begin{cases} C_t = 2 + 0.8 Y_t + u_t & (0 < \beta_1 < 1) & \text{(consumption function)} \\ Y_t = C_t + I_t & & \text{(Income Identity)} \end{cases}$$

The artificially controlled population is set to:

- $\beta_0 = 2, \beta_1 = 0.8, I_t \leftarrow$ given values
- $E(u_t) = 0, var(u_t) = \sigma^2 = 0.04$
- $E(u_t u_{t+j}) = 0, j \neq 0$
- $cov(u_t, I_t) = 0$

Simulation: the data sets

The simulation data under given conditions are:

Y	C	I	u
18.1570	16.1570	2	-0.3686
19.5998	17.5998	2	-0.0800
21.9347	19.7347	2.2	0.1869
21.5514	19.3514	2.2	0.1103
21.8843	19.4843	2.4	-0.0231
22.4265	20.0265	2.4	0.0853
25.4094	22.8094	2.6	0.4819
22.6952	20.0952	2.6	-0.0610

Showing 1 to 8 of 20 entries

Previous 1 2 3 Next

Simulation: manual calculation

According to the above formula, the regression coefficient can be calculated as follows:

Easy to calculate: $\sum u_t y_t$ =3.8000

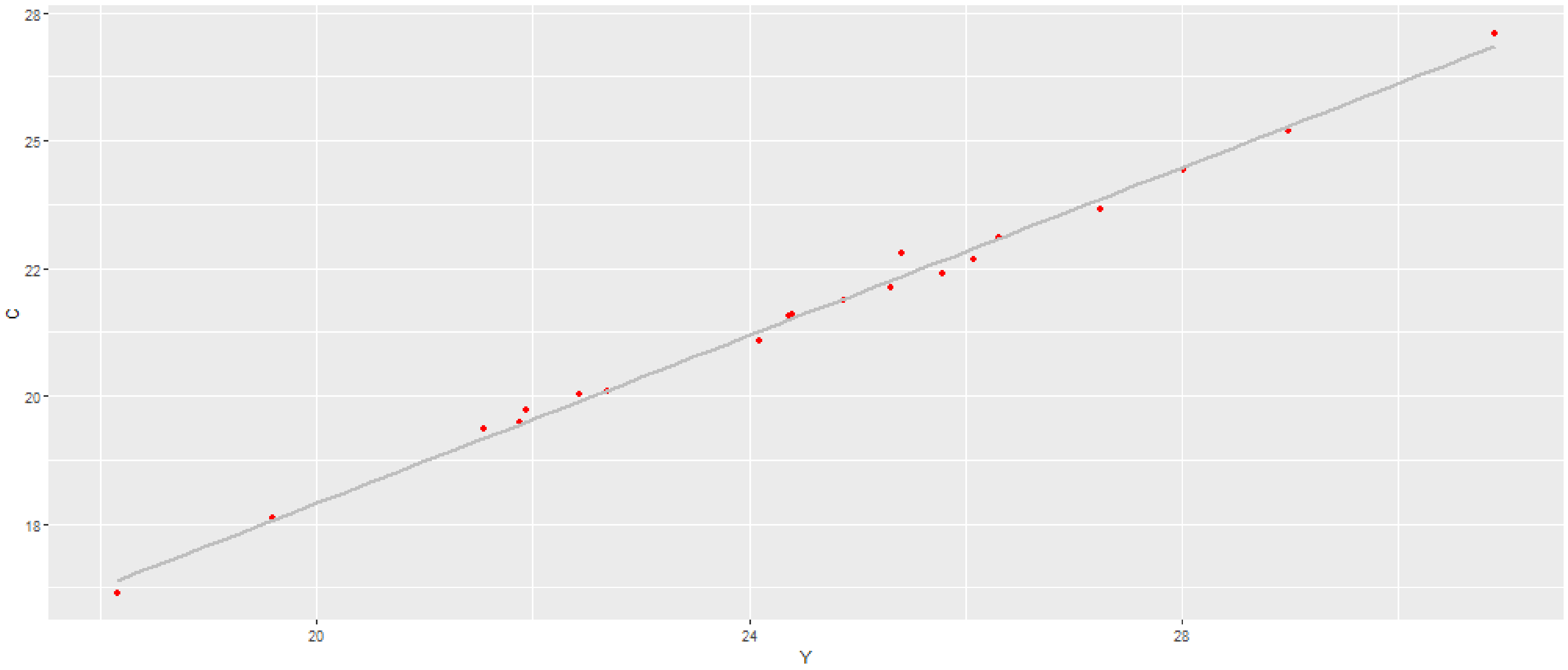
And: $\sum y_t^2$ =184.0000

And: $\frac{\sum u_t y_t}{\sum y_t^2}$ =0.0207

Hence： $\hat{\beta}_1 = \beta_1 + \frac{\sum u_t y_t}{\sum y_t^2}$ =0.8+0.0207= 0.8207

This also means that $\hat{\beta}_1$ is different from $\beta_1 = 0.8$, and the differnce is 0.0207.

Simulation: scatter plots



Simulation: regression report 1

Next, we used the simulated data for R analysis to obtain the original OLS report.

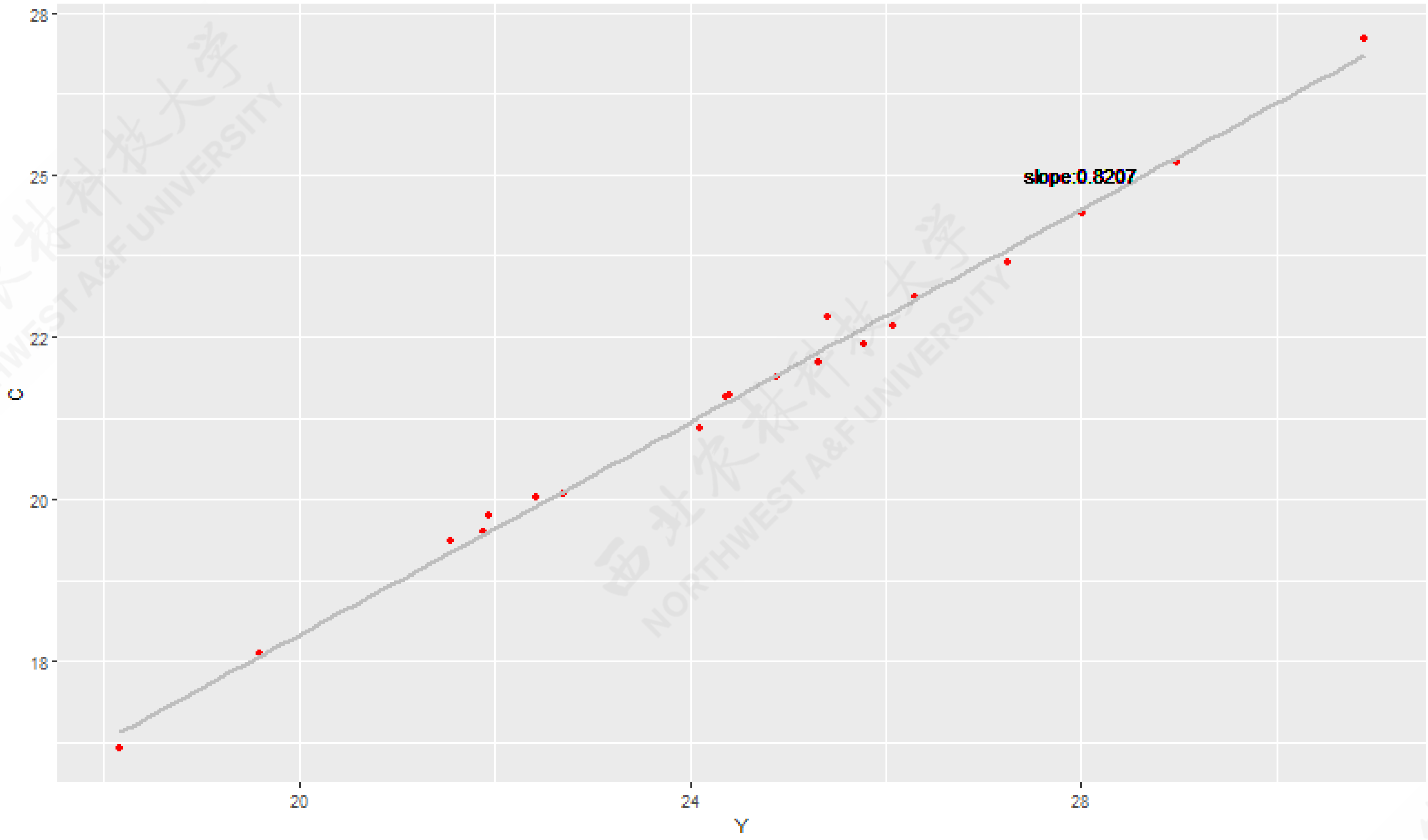
```
Call:
lm(formula = mod_monte$mod.C, data = monte)

Residuals:
    Min      1Q  Median      3Q     Max
-0.2700 -0.1586 -0.0013  0.0927  0.4631

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   1.4940     0.3541    4.22  0.00052 ***
Y             0.8207     0.0143   57.21  < 2e-16 ***
---
Signif. codes:
0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.2 on 18 degrees of freedom
Multiple R-squared:  0.995,    Adjusted R-squared:  0.994
F-statistic: 3.27e+03 on 1 and 18 DF,  p-value: <2e-16
```

Simulation: regression report 2

The tidy report of OLS estimation shows below.

$$
\begin{array}{lll}
\widehat{C} = & +1.49 & +0.82Y \\
\text{(t)} & (4.2188) & (57.2090) \\
\text{(se)} & (0.3541) & (0.0143) \\
\text{(fitness)} n = 20; & & R^2 = 0.9945; \bar{R}^2 = 0.9942 \\
 & F^* = 3272.87; p = 0.0000 &
\end{array}
$$

Simulation: sample regression line (SRL)



The SRL

Conclusions and points

So let's summarize this chapter.

- Compared with the single-equation model, the SEM involves more than one dependent or endogenous variable. So there must be as many equations as endogenous variables.
- SEM always show that the endogenous variables are correlated with stochastic terms in other equations.
- Classical OLS may not be appropriate because the estimators are inconsistent.

End Of This Chapter