计量经济学(Econometrics)

胡华平

西北农林科技大学

经济管理学院数量经济教研室

huhuaping01@hotmail.com

2021-09-07

第6章：多元回归：矩阵部分

6.1k变量模型的矩阵表达

k变量线性回归模型

k变量总体回归模型(PRM)

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \cdots + \beta_k X_{ki} + u_i$$

其n个联立方程组为：

$$\begin{aligned}
Y_1 &= \beta_1 + \beta_2 X_{21} + \beta_3 X_{31} + \cdots + \beta_k X_{k1} + u_1 \\
Y_2 &= \beta_1 + \beta_2 X_{22} + \beta_3 X_{32} + \cdots + \beta_k X_{k2} + u_2 \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
Y_n &= \beta_1 + \beta_2 X_{2n} + \beta_3 X_{3n} + \cdots + \beta_k X_{kn} + u_n
\end{aligned}$$

k变量线性回归模型

如果样本数为n，则可以将上述PRM模型表达为矩阵形式：

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \cdots \\ Y_n \end{bmatrix} = \begin{bmatrix} 1 & X_{21} & X_{31} & \cdots & X_{k1} \\ 1 & X_{22} & X_{32} & \cdots & X_{k2} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & X_{2n} & X_{3n} & \cdots & X_{kn} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{bmatrix} + \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}$$

进一步地，可以得到精简化的PRM矩阵形式：

$$\begin{array}{cccc} \boldsymbol{y} = & \boldsymbol{X\beta} & & +\boldsymbol{u} \\ (n \times 1) & (n \times k)(k \times 1) & + & (n \times 1) \end{array}$$

k变量线性回归模型

或者进一步紧凑表达为：

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{u}$$

其中：

- **向量**（默认为列向量）用加粗体的小写字母表达
- **矩阵**用大写粗体字母表达
- 矩阵或向量的**维度**需要注意标明

6.2 模型假设的矩阵表达

经典线性回归模型假设(N-CLRM)的矩阵表达

进一步地，我们可以用矩阵方法表达正态经典线性回归模型假设（N-CLRM）：

N-CLRM假设1-1：模型是正确设置的。（这里大有学问，也是一切计量分析问题的根本来源）

N-CLRM假设1-2：模型应该是参数线性的。也即模型中**参数**必须线性，变量可以不是线性。

CLRM假设2-1：X是固定的（给定的）或独立于误差项。也即自变量X**不是**随机变量。或者表达为矩阵 $\boldsymbol{X}_{n*k}$是非随机的，即它由固定数的一个集合构成。

N-CLRM假设2-2：多元回归情形下，自变量 X间无**完全共线性**。可记为 $\rho(\boldsymbol{X}) = k$，也即矩阵 $\boldsymbol{X}$为**列满秩**

- 矩阵 $\boldsymbol{X}$是列满秩(full column rank) 的， 也即其秩等于矩阵的列数。
- 矩阵 $\boldsymbol{X}$的列是线性独立的，即在 X_{ki} 变量之间无完全的线性关系即无**完全共线性**

经典线性回归模型假设(N-CLRM)的矩阵表达

N-CLRM假设3-1：随机干扰项期望为0。可记为 $E(\boldsymbol{u})=\mathbf{0}$

具体地：

$$E(\boldsymbol{u})=E\begin{bmatrix}u_1\\u_2\\\vdots\\u_n\end{bmatrix}=\begin{bmatrix}E(u_1)\\E(u_2)\\\vdots\\E(u_n)\end{bmatrix}=\begin{bmatrix}0\\0\\\vdots\\0\end{bmatrix}=\mathbf{0}$$

N-CLRM假设3-2/3：随机干扰项同方差且无自相关。可记为 $E(\boldsymbol{u}\boldsymbol{u}')=\sigma^2\boldsymbol{I}$

在正态性假设下，关于随机干扰项的全部假设可以记为 $\boldsymbol{u}\sim N(\mathbf{0},\sigma^2\boldsymbol{I})$

随机干扰项的方差协方差矩阵

随机干扰项的方差协方差矩阵为：

$$
\begin{aligned}
var-cov(\boldsymbol{u}) &= E(\boldsymbol{u}\boldsymbol{u}') \\
&= E\begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}\begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix} = E\begin{bmatrix} u_1^2 & u_1u_2 & \cdots & u_1u_n \\ u_2u_1 & u_2^2 & \cdots & u_2u_n \\ \vdots & \vdots & \vdots & \vdots \\ u_nu_1 & u_nu_2 & \cdots & u_n^2 \end{bmatrix} \\
&= \begin{bmatrix} E(u_1^2) & E(u_1u_2) & \cdots & E(u_1u_n) \\ E(u_2u_1) & E(u_2^2) & \cdots & E(u_2u_n) \\ \vdots & \vdots & \vdots & \vdots \\ E(u_nu_1) & E(u_nu_2) & \cdots & E(u_n^2) \end{bmatrix}
\end{aligned}
$$

随机干扰项的方差协方差矩阵

如果满足N-CLRM假设，则随机干扰项的方差协方差矩阵进一步可以写成：

$$
var-cov(\boldsymbol{u}) = E(\boldsymbol{uu}') = \begin{bmatrix} \sigma_1^2 & \sigma_{12}^2 & \cdots & \sigma_{1n}^2 \\ \sigma_{21}^2 & \sigma_2^2 & \cdots & \sigma_{2n}^2 \\ \vdots & \vdots & \vdots & \vdots \\ \sigma_{n1}^2 & \sigma_{n2}^2 & \cdots & \sigma_n^2 \end{bmatrix} \qquad \leftarrow (E(u_i)=0)
$$

$$
= \begin{bmatrix} \sigma^2 & \sigma_{12}^2 & \cdots & \sigma_{1n}^2 \\ \sigma_{21}^2 & \sigma^2 & \cdots & \sigma_{2n}^2 \\ \vdots & \vdots & \vdots & \vdots \\ \sigma_{n1}^2 & \sigma_{n2}^2 & \cdots & \sigma^2 \end{bmatrix} \qquad \leftarrow \left[var(u_i)=\sigma^2; cov(u_i,u_j)=0, i \neq j\right]
$$

$$
= \sigma^2 \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} = \sigma^2 \boldsymbol{I}
$$

6.3 OLS估计的矩阵表达

OLS估计的矩阵表达：代数过程

给定如下的样本回归模型（SRM）：

$$Y_i = \hat{\beta}_1 + \hat{\beta}_2 X_{2i} + \hat{\beta}_3 X_{3i} + \cdots + \hat{\beta}_k X_{ki} + e_i \qquad \text{(SRM)}$$

可以表达为：

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \cdots \\ Y_n \end{bmatrix} = \begin{bmatrix} 1 & X_{21} & X_{31} & \cdots & X_{k1} \\ 1 & X_{22} & X_{32} & \cdots & X_{k2} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & X_{2n} & X_{3n} & \cdots & X_{kn} \end{bmatrix} \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \vdots \\ \hat{\beta}_k \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix}$$

$$\begin{array}{ccc} \boldsymbol{y} = & \boldsymbol{X}\hat{\boldsymbol{\beta}} & +\boldsymbol{e} \\ (n \times 1) & (n \times k)(k \times 1) & +(n \times 1) \end{array}$$

OLS估计的矩阵表达：代数过程

$$
\begin{aligned}
\sum e_i^2 &= \sum\left(Y_i-\hat{\beta}_1-\hat{\beta}_2 X_{2i}-\cdots-\hat{\beta}_k X_{ki}\right)^2 \\
&= e_1^2+e_2^2+\cdots+e_n^2=\left[e_1 e_2 \cdots e_n\right]\begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix}
\end{aligned}
$$

$$
\left\{\begin{array}{ll}
\dfrac{\partial \sum e_i^2}{\partial \hat{\beta}_1}=2 \sum\left(Y_i-\hat{\beta}_1-\hat{\beta}_2 X_{2i}-\cdots-\hat{\beta}_k X_{ki}\right)(-1) & =0 \\
\dfrac{\partial \sum e_i^2}{\partial \hat{\beta}_2}=2 \sum\left(Y_i-\hat{\beta}_1-\hat{\beta}_2 X_{2i}-\cdots-\hat{\beta}_k X_{ki}\right)\left(-X_{2i}\right) & =0 \\
\qquad\cdots & \cdots \\
\dfrac{\partial \sum e_i^2}{\partial \hat{\beta}_k}=2 \sum\left(Y_i-\hat{\beta}_1-\hat{\beta}_2 X_{2i}-\cdots-\hat{\beta}_k X_{ki}\right)\left(-X_{ki}\right) & =0
\end{array}\right.
$$

OLS估计的矩阵表达：代数过程

正规方程组如下：

$$
\begin{cases}
n\hat{\beta}_1 & + \hat{\beta}_2 \sum X_{2i} & + \hat{\beta}_3 \sum X_{3i} & + \cdots + \hat{\beta}_k \sum X_{ki} & = \sum Y_i \\
\hat{\beta}_1 \sum X_{2i} & + \hat{\beta}_2 \sum X_{2i}^2 & + \hat{\beta}_3 \sum X_{2i}X_{3i} & + \cdots + \hat{\beta}_k \sum X_{2i}X_{ki} & = \sum X_{2i}Y_i \\
\hat{\beta}_1 \sum X_{3i} & + \hat{\beta}_2 \sum X_{3i}X_{2i} & + \hat{\beta}_3 \sum X_{3i}^2 & + \cdots + \hat{\beta}_k \sum X_{3i}X_{ki} & = \sum X_{3i}Y_i \\
\dots & \dots & \dots & \dots \quad \dots & \dots \\
\hat{\beta}_1 \sum X_{ki} & + \hat{\beta}_2 \sum X_{ki}X_{2i} & + \hat{\beta}_3 \sum X_{ki}X_{3i} & + \cdots + \hat{\beta}_k \sum X_{ki}^2 & = \sum X_{ki}Y_i
\end{cases}
$$

OLS估计的矩阵表达：代数过程

写成矩阵形式：

$$
\begin{bmatrix}
n & \sum X_{2i} & \sum X_{3i} & \cdots & \sum X_{ki} \\
\sum X_{2i} & \sum X_{2i}^2 & \sum X_{2i}X_{3i} & \cdots & X_{2i}X_{ki} \\
\sum X_{3i} & \sum X_{3i}X_{2i} & \sum X_{3i}^2 & \cdots & \sum X_{3i}X_{ki} \\
\cdots & \cdots & \cdots & \cdots & \cdots \\
\sum X_{ki} & \sum X_{ki}X_{2i} & \sum X_{ki}X_{3i} & \cdots & \sum X_{ki}^2
\end{bmatrix}
\begin{bmatrix}
\hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_3 \\ \vdots \\ \hat{\beta}_k
\end{bmatrix}
=
\begin{bmatrix}
1 & 1 & \cdots & 1 \\
X_{21} & X_{22} & \cdots & X_{2n} \\
X_{31} & X_{32} & \cdots & X_{3n} \\
\cdots & \cdots & \cdots & \cdots \\
X_{k1} & X_{k2} & \cdots & X_{kn}
\end{bmatrix}
\begin{bmatrix}
Y_1 \\ Y_2 \\ Y_3 \\ \vdots \\ Y_n
\end{bmatrix}
$$

从而有：

$$\boldsymbol{X}'\boldsymbol{X}\hat{\boldsymbol{\beta}} = \boldsymbol{X}'\boldsymbol{y}$$

如果矩阵 $\boldsymbol{X}'\boldsymbol{X}$的逆矩阵存在，则两边同时左乘 $(\boldsymbol{X}'\boldsymbol{X})^{-1}$，得到OLS估计量：

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

OLS估计的矩阵表达：纯矩阵过程

最小二乘方法将求解最小化过程：

$$
\begin{aligned}
Q &= \sum e_i^2 \\
&= \boldsymbol{e}'\boldsymbol{e} && \leftarrow \left[y = X\hat{\beta} + e; \quad e = y - X\hat{\beta} \right] \\
&= (\boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\beta}})'(\boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\beta}}) && \leftarrow \begin{cases} \left(\hat{\boldsymbol{\beta}}'\boldsymbol{X}'\boldsymbol{y}\right)' = \boldsymbol{y}'\boldsymbol{X}\hat{\boldsymbol{\beta}} = \hat{\boldsymbol{\beta}}'\boldsymbol{X}'\boldsymbol{y} \\ (1^*\mathrm{k}) \cdot (\mathrm{k}^*\mathrm{n}) \cdot (\mathrm{n}^*1) = (1 * 1) \end{cases} \\
&= \boldsymbol{y}'\boldsymbol{y} - \boldsymbol{2}\hat{\boldsymbol{\beta}}'\boldsymbol{X}'\boldsymbol{y} + \hat{\boldsymbol{\beta}}'\boldsymbol{X}'\boldsymbol{X}\hat{\boldsymbol{\beta}}
\end{aligned}
$$

> 提示：标量的转置还是自身！

OLS估计的矩阵表达：纯矩阵过程

进一步可以得到：

$$
\begin{aligned}
\frac{\partial Q}{\partial \hat{\boldsymbol{\beta}}} &= 0 \\
\frac{\partial(\boldsymbol{y}'\boldsymbol{y} - \boldsymbol{2}\hat{\boldsymbol{\beta}}'\boldsymbol{X}'\boldsymbol{y} + \hat{\boldsymbol{\beta}}'\boldsymbol{X}'\boldsymbol{X}\hat{\boldsymbol{\beta}})}{\partial \hat{\boldsymbol{\beta}}} &= 0 \\
-2\boldsymbol{X}'\boldsymbol{y} + 2\boldsymbol{X}'\boldsymbol{X}\hat{\boldsymbol{\beta}} &= 0 \\
-\boldsymbol{X}'\boldsymbol{y} + \boldsymbol{X}'\boldsymbol{X}\hat{\boldsymbol{\beta}} &= 0 \\
\boldsymbol{X}'\boldsymbol{X}\hat{\boldsymbol{\beta}} &= \boldsymbol{X}'\boldsymbol{y}
\end{aligned}
$$

如果矩阵 $\boldsymbol{X}'\boldsymbol{X}$的逆矩阵存在，则两边同时左乘 $(\boldsymbol{X}'\boldsymbol{X})^{-1}$，得到OLS估计量：

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

提示：

$$
\begin{aligned}
F &= AZ, & \frac{\partial F}{\partial Z} &= A' \\
F &= ZA, & \frac{\partial F}{\partial Z} &= A \\
F &= Z'A, & \frac{\partial F}{\partial Z} &= A \\
F &= A'ZB, & \frac{\partial F}{\partial Z} &= AB' \\
F &= A'Z'B, & \frac{\partial F}{\partial Z} &= AB' \\
F &= A'Z'B, & \frac{\partial F}{\partial Z} &= BA' \\
F &= Z'AZ, & \frac{\partial F}{\partial Z} &= AZ + A'Z
\end{aligned}
$$

OLS估计的矩阵表达：纯矩阵过程(u的方差)

此外，很用以证明OLS方法下，利用样本回归模型得到的估计量 $\hat{\sigma}^2$，是对总体回归模型参数 σ^2的无偏估计，也即：

$$\hat{\sigma}^2 = \frac{\sum e_i^2}{n-k} = \frac{\boldsymbol{e}'\boldsymbol{e}}{n-k} = \frac{= \mathbf{y}'\mathbf{y} - \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y}}{n-k} \qquad \leftarrow E(\hat{\sigma}^2) = \sigma^2$$

$$\begin{aligned}
\mathbf{e}'\mathbf{e} &= (\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})'(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}) \\
&= \left(\mathbf{y}' - \hat{\boldsymbol{\beta}}'\mathbf{X}'\right)(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}) \\
&= \mathbf{y}'\mathbf{y} - \mathbf{y}'\mathbf{X}'\mathbf{y} - \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y} + \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}} \\
&= \mathbf{y}'\mathbf{y} - \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y} - \left(\mathbf{y}' - \hat{\boldsymbol{\beta}}'\mathbf{X}'\right)\mathbf{X}\hat{\boldsymbol{\beta}} \\
&= \mathbf{y}'\mathbf{y} - \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y} - \mathbf{e}'\mathbf{X}\hat{\boldsymbol{\beta}} \\
&= \mathbf{y}'\mathbf{y} - \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y} - (\mathbf{X}'\mathbf{e})'\hat{\boldsymbol{\beta}} \\
&= \mathbf{y}'\mathbf{y} - \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y}
\end{aligned}$$

$$\begin{aligned}
\mathbf{y} &= \mathbf{X}\hat{\boldsymbol{\beta}} + \mathbf{e} \\
\mathbf{y}' &= \hat{\boldsymbol{\beta}}'\mathbf{X}' + \mathbf{e}' \\
\mathbf{y}' - \hat{\boldsymbol{\beta}}'\mathbf{X}' &= \mathbf{e}'
\end{aligned}$$

$$\begin{aligned}
\boldsymbol{y} &= \boldsymbol{X}\hat{\boldsymbol{\beta}} + \mathbf{e} \\
\mathbf{X}'\mathbf{y} &= \mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}} + \mathbf{X}'\mathbf{e} \\
\mathbf{X}'\mathbf{y} &= \mathbf{X}'\mathbf{X}\left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{y} + \mathbf{X}'\mathbf{e} \\
&= \mathbf{X}'\mathbf{y} + \mathbf{X}'\mathbf{e} \\
\mathbf{X}'\mathbf{e} &= \mathbf{0}
\end{aligned}$$

回归系数的方差协方差矩阵

对于回归系数的OLS估计量 $\hat{\boldsymbol{\beta}}$，进一步讨论其方差和协方差矩阵 $var-cov(\hat{\boldsymbol{\beta}})$，一般记为：

$$
\begin{aligned}
var-cov(\hat{\boldsymbol{\beta}}) &= E\left(\left(\hat{\beta}-E(\hat{\beta})\right)\left(\hat{\beta}-E(\hat{\beta})\right)'\right) \\
&= \begin{bmatrix}
var(\hat{\beta}_1) & cov(\hat{\beta}_1,\hat{\beta}_2) & \cdots & cov(\hat{\beta}_1,\hat{\beta}_k) \\
cov(\hat{\beta}_2,\hat{\beta}_1) & var(\hat{\beta}_2) & \cdots & cov(\hat{\beta}_2,\hat{\beta}_k) \\
\vdots & \vdots & \vdots & \vdots \\
cov(\hat{\beta}_k,\hat{\beta}_1) & cov(\hat{\beta}_k,\hat{\beta}_2) & \cdots & var(\hat{\beta}_k)
\end{bmatrix}
\end{aligned}
$$

如果满足N-CLRM假设，则回归系数的OLS估计量 $\hat{\boldsymbol{\beta}}$的方差和协方差矩阵 $var-cov(\hat{\boldsymbol{\beta}})$可以进一步写成：

$$
\begin{aligned}
var-cov(\hat{\boldsymbol{\beta}}) &\equiv \sigma^2_{\hat{\boldsymbol{\beta}}} \\
&= \boldsymbol{E}\left(\left(\hat{\boldsymbol{\beta}}-\boldsymbol{E}(\hat{\boldsymbol{\beta}})\right)\left(\hat{\boldsymbol{\beta}}-\boldsymbol{E}(\hat{\boldsymbol{\beta}})\right)'\right) \\
&= \boldsymbol{E}\left(\left(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right)\left(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right)'\right) \\
&= \boldsymbol{E}\left(((\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{u})\,((\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{u})'\right) \\
&= \boldsymbol{E}\left((\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{u}\boldsymbol{u}'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\right) \\
&= (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{E}(\boldsymbol{u}\boldsymbol{u}')\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1} \\
&= (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\sigma^2\boldsymbol{I}\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1} \\
&= \sigma^2(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1} \\
&= \sigma^2(\boldsymbol{X}'\boldsymbol{X})^{-1}
\end{aligned}
$$

$$
\begin{aligned}
\hat{\boldsymbol{\beta}} &= \left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\boldsymbol{y} \\
&= \left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}\boldsymbol{X}'(\boldsymbol{X}\boldsymbol{\beta}+\boldsymbol{u}) \\
&= \left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\beta}+\left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\boldsymbol{u} \\
&= \boldsymbol{\beta}+\left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\boldsymbol{u}
\end{aligned}
$$

$$
\hat{\boldsymbol{\beta}}-\boldsymbol{\beta} = \left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\boldsymbol{u}
$$

$$
(\mathrm{AB})' = \mathrm{B}'\mathrm{A}'
$$

$$
\begin{aligned}
\left[\left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}\right]' &= \left[\left(\boldsymbol{X}'\boldsymbol{X}\right)'\right]^{-1} \\
&= \left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}
\end{aligned}
$$

回归系数的方差协方差矩阵

那么，可以很快得到回归系数的OLS估计量 $\hat{\boldsymbol{\beta}}$的样本方差和协方差矩阵 $S_{ij}^2(\hat{\boldsymbol{\beta}})$

$$
\begin{aligned}
S_{ij}^2(\hat{\boldsymbol{\beta}}) &= \hat{\sigma}^2(\boldsymbol{X}'\boldsymbol{X})^{-1} \\
&= \frac{\boldsymbol{e}'\boldsymbol{e}}{n-k}(\boldsymbol{X}'\boldsymbol{X})^{-1} \\
&= \frac{\mathbf{y}'\mathbf{y} - \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y}}{n-k} \cdot (\boldsymbol{X}'\boldsymbol{X})^{-1}
\end{aligned}
$$

OLS估计的性质：BLUE

下面我们将证明高斯-马尔可夫定理(Gauss-Markov Theorem)：在正态经典线性回归模型假设（N-CLRM）下，采用普通最小二乘法（OLS），得到的估计量 $\hat{\beta}$，是真实参数 β最优的、线性的、无偏估计量（BLUE）。记为：

$$\xrightarrow[\text{N-CLRM}]{\text{OLS}} \hat{\boldsymbol{\beta}} \xrightarrow{\text{BLUE}} \boldsymbol{\beta}$$

因为模型参数的OLS估计为：

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

又因为矩阵 $\boldsymbol{X}$为列满秩，也即 $\rho(\boldsymbol{X}) = k$，所以 $\hat{\boldsymbol{\beta}}$关于 $\boldsymbol{y}$是线性的。

OLS估计的性质：BLUE

根据模型参数OLS估计，容易得到如下过程：

$$
\begin{aligned}
\hat{\boldsymbol{\beta}} &= (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y} \\
&= (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'(\boldsymbol{X}\boldsymbol{\beta}+\boldsymbol{u}) \\
&= (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\beta}+(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{u} \\
&= \boldsymbol{\beta}+(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{u}
\end{aligned}
$$

进一步可证明

$$
\begin{aligned}
E(\hat{\boldsymbol{\beta}}) &= E(\boldsymbol{\beta}+(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{u}) \\
&= \boldsymbol{E}(\boldsymbol{\beta})+(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{E}(\boldsymbol{u}) \\
&= \boldsymbol{\beta}
\end{aligned}
$$

因此，$\hat{\boldsymbol{\beta}}$是参数$\boldsymbol{\beta}$的无偏估计量得证。

OLS估计的性质：BLUE

假设存在用其他方法估计的线性无偏估计量 $\boldsymbol{\beta}^*$，则要求 $\boldsymbol{C}$满足如下条件：

$$\boldsymbol{C}\boldsymbol{X} = 0$$

从而保证如下式子成立：

$$\begin{aligned}
\boldsymbol{\beta}^* &= \left((\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}' + \boldsymbol{C}\right)\boldsymbol{y} \\
&= \left((\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}' + \boldsymbol{C}\right)(\boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{u}) \\
&= \boldsymbol{\beta} + \boldsymbol{C}\boldsymbol{X}\boldsymbol{\beta} + (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{u} + \boldsymbol{C}\boldsymbol{u} \\
&= \boldsymbol{\beta} + (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{u} + \boldsymbol{C}\boldsymbol{u}
\end{aligned}$$

进一步得到：

$$\boldsymbol{\beta}^* - \boldsymbol{\beta} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{u} + \boldsymbol{C}\boldsymbol{u}$$

根据方差定义，有：

$$
\begin{aligned}
var-cov(\boldsymbol{\beta}^*) &= \boldsymbol{E}\left((\boldsymbol{\beta}^*-\boldsymbol{\beta})(\boldsymbol{\beta}^*-\boldsymbol{\beta})'\right) \\
&= \boldsymbol{E}\left(\left((\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{u}+\boldsymbol{C}\boldsymbol{u}\right)\left((\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{u}+\boldsymbol{C}\boldsymbol{u}\right)'\right) \\
&= \boldsymbol{\sigma}^2(\boldsymbol{X}'\boldsymbol{X})^{-1}+\boldsymbol{\sigma}^2\boldsymbol{C}\boldsymbol{C}' \\
&= var-cov(\hat{\boldsymbol{\beta}})+\boldsymbol{\sigma}^2\boldsymbol{C}\boldsymbol{C}'
\end{aligned}
$$

其中，我们可以证明$\boldsymbol{\sigma}^2\boldsymbol{C}\boldsymbol{C}'$是半正定矩阵，矩阵对角线元素$\geq 0$，因此有：

$$
var-cov(\boldsymbol{\beta}^*) \geq var-cov(\hat{\boldsymbol{\beta}})
$$

从而表明N-CLRM假设下，OLS方法估计得到的$\hat{\boldsymbol{\beta}}$，方差最小。

6.4 假设检验的矩阵表达

平方和分解的矩阵表达

对于多元回归模型：

$$
\begin{aligned}
Y_i &= \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \cdots + \beta_k X_{ki} + u_i && \text{(PRM)} \\
Y_i &= \hat{\beta}_1 + \hat{\beta}_2 X_{2i} + \hat{\beta}_3 X_{3i} + \cdots + \hat{\beta}_k X_{ki} + e_i && \text{(SRM)} \\
\hat{Y}_i &= \hat{\beta}_1 + \hat{\beta}_2 X_{2i} + \hat{\beta}_3 X_{3i} + \cdots + \hat{\beta}_k X_{ki} && \text{(SRF)}
\end{aligned}
$$

平方和分解的矩阵表达

通过对 Y_i 的变异及其来源的分解，可以得到：

$$
\begin{aligned}
(Y_i-\bar{Y}_i) &= (\hat{Y}_i-\bar{Y}_i)+(Y_i-\bar{Y}_i) \\
y_i &= \hat{y}_i+e_i \\
\sum y_i^2 &= \sum \hat{y_i}^2+\sum e_i^2 \\
TSS &= ESS+RSS
\end{aligned}
$$

其中TSS表示总离差平方和，ESS表示回归平法和，RSS表示残差平方和。它们分别可以用矩阵表达为：

$$
\begin{aligned}
TSS &= \boldsymbol{y}'\boldsymbol{y}-n\bar{Y}^2 \\
RSS &= \boldsymbol{e}\boldsymbol{e}' = \boldsymbol{y}\boldsymbol{y}' - \hat{\boldsymbol{\beta}}'\boldsymbol{X}'\boldsymbol{y} \\
ESS &= \hat{\boldsymbol{\beta}}'\boldsymbol{X}'\boldsymbol{y}-n\bar{Y}^2
\end{aligned}
$$

平方和分解的矩阵表达（ANOVA）

进一步地，可以得到方差分析表（ANOVA）：

变异来源	平方和符号 SS	平方和计算公式	自由度 df	均方和符号 MSS	均方和计算公式
回归平方和	ESS	$\sum \hat{y}_i^2$	k-1	MSS_{ESS}	$=(\hat{\boldsymbol{\beta}}'\boldsymbol{X}'\boldsymbol{y}-n\bar{Y}^2)/(k-1)$
残差平方和	RSS	$\sum e_i^2$	n-k	MSS_{RSS}	$=(\boldsymbol{y}\boldsymbol{y}'-\hat{\boldsymbol{\beta}}'\boldsymbol{X}'\boldsymbol{y})/(n-k)$
总平方和	TSS	$\sum y_i^2$	n-1	MSS_{TSS}	$=(\boldsymbol{y}'\boldsymbol{y}-n\bar{Y}^2)/(n-1)$

拟合优度的矩阵表达

根据拟合优度的定义，判定系数 R^2和调整判定系数 $\bar{R}^2$的矩阵计算公式为：

$$R^2 = \frac{ESS}{TSS} = \frac{\hat{\boldsymbol{\beta}}'\boldsymbol{X}'\boldsymbol{y} - n\bar{Y}^2}{\boldsymbol{y}'\boldsymbol{y} - n\bar{Y}^2}$$

$$\bar{R}^2 = 1 - \frac{RSS/f_{RSS}}{TSS/f_{TSS}} = 1 - \frac{MSS_{RSS}}{MSS_{TSS}} = 1 - \frac{\left(\boldsymbol{y}\boldsymbol{y}' - \hat{\boldsymbol{\beta}}'\boldsymbol{X}'\boldsymbol{y}\right)/(k-1)}{\left(\boldsymbol{y}'\boldsymbol{y} - n\bar{Y}^2\right)/(n-1)}$$

回归系数显著性检验（t检验）的矩阵方法实现

根据回归系数显著性检验的定义，利用矩阵方法实现t检验的过程如下：

对于多元回归模型

$$
\begin{aligned}
Y_i &= \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \cdots + \beta_k X_{ki} + u_i \\
\boldsymbol{y} &= \boldsymbol{X\beta} + \boldsymbol{u} && \text{(PRM)} \\
\boldsymbol{y} &= \boldsymbol{X\hat{\beta}} + \boldsymbol{e} && \text{(SRM)}
\end{aligned}
$$

回归系数显著性检验（t检验）的矩阵方法实现

在N-CLRM假设下，采用OLS估计方法，可以证明：

$$\boldsymbol{u} \sim N(\mathbf{0}, \sigma^2 \boldsymbol{I})$$
$$\hat{\boldsymbol{\beta}} \sim N\left(\boldsymbol{\beta}, \sigma^2 \boldsymbol{X}'\boldsymbol{X}^{-1}\right)$$

从而可以构造t统计量

$$\boldsymbol{T}_{\hat{\beta}} = \frac{\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}}{\boldsymbol{S}_{\hat{\beta}}} \sim \boldsymbol{t(n-k)}$$

$$T_{\hat{\beta}} = \frac{\hat{\beta}_j - \beta_j}{\sqrt{c_{ij} \frac{\mathbf{y}'\mathbf{y} - \hat{\beta}'\mathbf{X}'\mathbf{y}}{n-k}}}$$

提示：

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

$$\begin{aligned} var - cov(\hat{\boldsymbol{\beta}}) &\equiv \sigma^2_{\hat{\beta}} \\ &= \sigma^2(\boldsymbol{X}'\boldsymbol{X})^{-1} \end{aligned}$$

$$\begin{aligned} S^2_{ij}(\hat{\boldsymbol{\beta}}) &= \hat{\sigma}^2(\boldsymbol{X}'\boldsymbol{X})^{-1} \\ &= \frac{\boldsymbol{e}'\boldsymbol{e}}{n-k}(\boldsymbol{X}'\boldsymbol{X})^{-1} \\ &= \frac{\mathbf{y}'\mathbf{y} - \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y}}{n-k} \cdot (\boldsymbol{X}'\boldsymbol{X})^{-1} \end{aligned}$$

回归系数显著性检验（t检验）的矩阵方法实现

对于总体回归模型的任一参数 $\boldsymbol{\beta_j}, j \in (1,2,\cdots,k)$提出假设：

$$
\begin{aligned}
H_0 &: \boldsymbol{\beta_j} = 0 \\
H_1 &: \boldsymbol{\beta_j} \neq 0
\end{aligned}
$$

根据原假设 H_0，可以得到：

$$
\boldsymbol{t^*_{\hat{\beta}}} = \frac{\hat{\boldsymbol{\beta}}}{\sqrt{\boldsymbol{S^2_{ij}(\hat{\beta}_{kk})}}} = \frac{\hat{\beta}_j}{\sqrt{c_{ij}\frac{\mathbf{y}'\mathbf{y}-\hat{\beta}'\mathbf{X}'\mathbf{y}}{n-k}}}
$$

其中 $\boldsymbol{S^2_{ij}(\hat{\beta_{kk}})}$表示，由 $\hat{\boldsymbol{\beta}}$的样本方差和协方差矩阵 $S^2_{ij}(\hat{\boldsymbol{\beta}})$的对角线元素组成的列向量，即

$$
S^2_{ij}(\hat{\beta}_{kk}) = [\boldsymbol{s}^2_{\hat{\beta}_1}, \boldsymbol{s}^2_{\hat{\beta}_2}, \cdots, \boldsymbol{s}^2_{\hat{\beta}_k}]^T
$$

回归系数显著性检验（t检验）的矩阵方法实现

若给定显著性水平 α和自由度 $(n-k)$，很快可以得到t分布的查表t值，也即 $t_{(1-\alpha/2)}(n-k)$。

然后比较样本t统计量 $\boldsymbol{t}^*_{\hat{\boldsymbol{\beta}}}$与理论t分布查的表t值 $(t_{(1-\alpha/2)}(n-2))$的关系。根据如下法则做出参数 β_2的显著性检验结论：

- 如果列向量 $\boldsymbol{t}^*_{\hat{\boldsymbol{\beta}}}$的第 k个元素 $t^*_{\hat{\beta}_k} > t_{(1-\alpha/2)}(n-2)$，则表明参数 β_k的t检验在 α水平下是**显著**的，也即显著地拒绝 $H_0: \beta_k = 0$，从而接受 $H_1: \beta_k \neq 0$。
- 如果列向量 $\boldsymbol{t}^*_{\hat{\boldsymbol{\beta}}}$的第 k个元素 $t^*_{\hat{\beta}_k} \leq t_{(1-\alpha/2)}(n-2)$，则表明参数 β_k的t检验在 α水平下是**不显著**的，也即不能显著地拒绝 $H_0: \beta_k = 0$，从而只能暂时接受 $H_0: \beta_k = 0$。

模型整体显著性检验（F检验）的矩阵方法实现

对于多元回归模型

$$
\begin{aligned}
Y_i &= \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \cdots + \beta_k X_{ki} + u_i && \text{(U-PRM)}\\
Y_i &= \hat{\beta}_1 + \hat{\beta}_2 X_{2i} + \hat{\beta}_3 X_{3i} + \cdots + \hat{\beta}_k X_{ki} + e_i && \text{(U-SRM)}\\
\boldsymbol{y} &= \boldsymbol{X\beta} + \boldsymbol{u} && \text{(PRM)}\\
\boldsymbol{y} &= \boldsymbol{X\hat{\beta}} + \boldsymbol{e} && \text{(SRM)}
\end{aligned}
$$

我们称总体回归模型和对应的样本回归模型为**无约束模型**（unrestricted model）。

对于总体回归模型的斜率参数 $\boldsymbol{\beta_j}, j \in (2, \cdots, k)$提出如下**联合假设**（joint hypothesis）：

$$
\begin{aligned}
H_0 :& \beta_2 = \beta_3 = \cdots = \beta_k = 0\\
H_1 :& \beta_j \text{ not all } 0, \text{ for } j \in (2, \cdots, k)
\end{aligned}
$$

模型整体显著性检验（F检验）的矩阵方法实现

在原假设 $H_0 : \beta_2 = \beta_3 = \cdots = \beta_k = 0$下，我们可以得到如下模型：

$$
\begin{aligned}
Y_i &= \beta_1 + u_i && \text{(R-PRM)} \\
Y_i &= \hat{\beta}_1 + e_i && \text{(R-SRM)}
\end{aligned}
$$

此时，我们称总体回归模型和对应的样本回归模型为**受约束模型**（restricted model）。在备择假设 $H_1 : \beta_j$不全为0，$j \in (2, \cdots, k)$下，我们可以得到该假设下的一种特殊回归模型（如 $\beta_j \neq 0, j \in (2, \cdots, k)$），也即无约束总体回归模型和无约束样本回归模型。

受约束模型：一般也称为参数约束回归模型（restricted model），是指总体参数满足某种约束条件的一类回归模型。

无约束模型：一般也称为参数无约束回归模型（unrestricted model），是指总体参数没有被指定满足某种约束条件的一类回归模型。

模型整体显著性检验（F检验）的矩阵方法实现

根据回归系数显著性检验的定义，利用矩阵方法实现F检验的过程如下：

在N-CLRM假设下，采用OLS估计方法，容易证明：

对于无约束总体回归模型有

$$
\begin{aligned}
u_i &\sim i.i.d\ N(0,\sigma^2) \\
Y_i &\sim i.i.d\ N(\beta_1+\beta_2X_i+\cdots+\beta_kX_i,\sigma^2) \\
RSS_U &= \sum(Y_i-\hat{Y}_i)^2 \sim \chi^2(n-k)
\end{aligned}
$$

对于受约束总体回归模型有

$$
\begin{aligned}
u_i &\sim i.i.d\ N(0,\sigma^2) \\
Y_i &\sim i.i.d\ N(\beta_1,\sigma^2) \\
RSS_R &= \sum(Y_i-\hat{Y}_i)^2 \sim \chi^2(n-1)
\end{aligned}
$$

进一步地可以构造得到随机变量 $\tilde{F}$，它将服从如下的F分布：

$$\tilde{F}=\frac{(RSS_R-RSS_U)/(k-1)}{RSS_U/(n-k)}=\frac{ESS_U/df_{ESS_U}}{RSS_U/df_{RSS_U}}\sim F(df_{ESS},df_{RSS})$$

模型整体显著性检验（F检验）的矩阵方法实现

变异来源	平方和符号 SS	平方和计算公式	自由度 df	均方和符号 MSS	均方和计算公式
回归平方和	ESS	$\sum \hat{y}_i^2$	k-1	MSS_{ESS}	$=(\hat{\boldsymbol{\beta}}'\boldsymbol{X}'\boldsymbol{y}-n\bar{Y}^2)/(k-1)$
残差平方和	RSS	$\sum e_i^2$	n-k	MSS_{RSS}	$=(\boldsymbol{y}\boldsymbol{y}'-\hat{\boldsymbol{\beta}}'\boldsymbol{X}'\boldsymbol{y})/(n-k)$
总平方和	TSS	$\sum y_i^2$	n-1	MSS_{TSS}	$=(\boldsymbol{y}'\boldsymbol{y}-n\bar{Y}^2)/(n-1)$

模型整体显著性检验（F检验）的矩阵方法实现

基于原假设 $H_0: \beta_2 = \cdots = \beta_k = 0$（也即斜率系数全部等于0，或者说约束模型的 $RSS_R = 0$），然后根据ANOVA分析表，我们可以计算得到一个样本F统计量（F^*）

$$F^* = \frac{ESS_U/df_{ESS_U}}{RSS_U/df_{RSS_U}} = \frac{\left(\hat{\boldsymbol{\beta}}'\boldsymbol{X}'\boldsymbol{y} - n\bar{Y}^2\right)/(k-1)}{\left(\boldsymbol{y}\boldsymbol{y}' - \hat{\boldsymbol{\beta}}'\boldsymbol{X}'\boldsymbol{y}\right)/(n-k)}$$

此外，我们还可以通过拟合优度 R^2，计算得到 F^*

$$F^* = \frac{ESS_U/df_{ESS_U}}{RSS_U/df_{RSS_U}} = \frac{R_U^2/(k-1)}{\left(1-R_U^2\right)/(n-k)}$$

模型整体显著性检验（F检验）的矩阵方法实现

若给定显著性水平 α和样本数 (n)，很快可以得到F分布的查表F值，也即 $F_{(1-\alpha)}(k-1, n-k)$,然后比较其与样本F统计量（ F^*）的关系。

根据如下法则做出总体回归模型整体显著性检验结论：

- 如果 $F^* > F_{(1-\alpha)}(k-1, n-k)$，则表明总体回归模型的F检验在 α水平下是**显著**的，也即显著地拒绝 $H_0: \beta_2 = \beta_3 = \cdots = \beta_k = 0$，从而接受 $H_1: \beta_j$不全为0，$j \in (2, \cdots, k)$，认为模型整体统计上是有意义的！
- 如果 $F^* \leq F_{(1-\alpha)}(k-1, n-k)$，则表明总体回归模型的F检验在 α水平下是**不显著**的，也即不能显著地拒绝 $H_0: \beta_2 = \beta_3 = \cdots = \beta_k = 0$，从而只能暂时接受 $H_0: \beta_2 = 0$，认为模型整体在统计上是无意义的！

6.5 模型预测的矩阵表达

样本外预测的矩阵方法实现

根据一元线性回归样本外预测的知识内容，下面将用矩阵方法实现：

- 样本外均值预测 $\boldsymbol{E(Y_0|X_0)}$
- 样本外个值预测 $\boldsymbol{(Y_0|X_0)}$。

其中，给定样本外数据 $\boldsymbol{X_0} = [1, X_{20}, X_{30}, \cdots, X_{k0}]^t$（列向量）。

对于多元回归模型：

$$
\begin{aligned}
Y_i &= \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \cdots + \beta_k X_{ki} + u_i \\
Y_i &= \hat{\beta}_1 + \hat{\beta}_2 X_{2i} + \hat{\beta}_3 X_{3i} + \cdots + \hat{\beta}_k X_{ki} + e_i \\
\boldsymbol{y} &= \boldsymbol{X\beta} + \boldsymbol{u} \\
\boldsymbol{y} &= \boldsymbol{X\hat{\beta}} + \boldsymbol{e} \\
\boldsymbol{\hat{y}} &= \boldsymbol{X\hat{\beta}} \\
\hat{Y}_0 &= \boldsymbol{X_0\hat{\beta}}
\end{aligned}
$$

样本外预测的矩阵方法实现：均值预测

对于样本外均值预测 $\boldsymbol{E}(\boldsymbol{Y_0}|\boldsymbol{X_0})$，矩阵实现步骤如下：

在N-CLRM假设下，已知 $\hat{Y}_0$ 的期望和**真实方差**为：

$$
\begin{aligned}
E(\hat{Y}_0) &= E(\boldsymbol{X_0}\hat{\boldsymbol{\beta}}) = \boldsymbol{X_0}\boldsymbol{\beta} = E(\boldsymbol{Y_0}) \\
var(\hat{Y}_0) &= E(\boldsymbol{X_0}\hat{\boldsymbol{\beta}} - \boldsymbol{X_0}\boldsymbol{\beta})^2 && \leftarrow (var.def) \\
&= E\left(\boldsymbol{X_0}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})'\boldsymbol{X_0'}\right) && \leftarrow (\text{trans.}) \\
&= \boldsymbol{X_0}\boldsymbol{E}\left((\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})'\right)\boldsymbol{X_0'} && \leftarrow def. \quad var(\hat{\boldsymbol{\beta}}) \\
&= \sigma^2\boldsymbol{X_0}(\boldsymbol{X'X})^{-1}\boldsymbol{X_0'} && \leftarrow var(\hat{\boldsymbol{\beta}}) = \sigma^2(\boldsymbol{X'X})^{-1}
\end{aligned}
$$

同时，$\hat{Y}_0$的样本方差为：

$$
S^2_{\hat{Y}_0} = \hat{\sigma}^2\boldsymbol{X_0}(\boldsymbol{X'X})^{-1}\boldsymbol{X_0'}
$$

样本外预测的矩阵方法实现：均值预测

因此 $\hat{\boldsymbol{Y}}_{\boldsymbol{0}}$服从如下正态分布：

$$\hat{Y}_0 \sim N(\mu_{\hat{Y}_0}, \sigma^2_{\hat{Y}_0})$$
$$\hat{Y}_0 \sim N\left(E(Y_0|X_0), \sigma^2 \boldsymbol{X_0}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'_{\boldsymbol{0}}\right)$$

因此可以构造t统计量：

$$t_{\hat{Y}_0} = \frac{\hat{Y}_0 - E(Y|X_0)}{S_{\hat{Y}_0}} \qquad \sim t(n-k)$$

其中：

$$\boldsymbol{S}_{\hat{\boldsymbol{Y}}_{\boldsymbol{0}}} = \sqrt{\hat{\sigma}^2 X_0(X'X)^{-1}X'_0}$$
$$\hat{\sigma}^2 = \frac{\boldsymbol{ee}'}{(n-k)}$$

样本外预测的矩阵方法实现：均值预测

给定显著性水平 α的情况下，可以查表得到理论t值 $t_{1-\alpha/2}(n-k)$，从而可以计算得到均值预测的置信区间：

$$\hat{Y}_0 - t_{1-\alpha/2}(n-k) \cdot S_{\hat{Y}_0} \leq E(Y|X_0) \leq \hat{Y}_0 + t_{1-\alpha/2}(n-k) \cdot S_{\hat{Y}_0}$$

其中：

$$\boldsymbol{S}_{\hat{\mathbf{Y}}_0} = \sqrt{\hat{\sigma}^2 X_0 (X'X)^{-1} X_0'} \quad \leftarrow \hat{\sigma}^2 = \frac{\boldsymbol{ee}'}{(n-k)}$$

$$\hat{Y}_0 = \boldsymbol{X_0}\hat{\boldsymbol{\beta}}$$

样本外预测的矩阵方法实现：个值预测

对于多元线性回归模型，样本外个值预测 $(\mathbf{Y_0}|\boldsymbol{X}_\mathbf{0})$ 的矩阵实现步骤如下：

因为有： $e_0 = Y_0 - \hat{Y}_0$

所以 e_0 的期望为：

$$
\begin{aligned}
E(e_0) &= E(Y_0 - \hat{Y}_0) \\
&= E(\boldsymbol{X_0\beta} + u_0 - \boldsymbol{X_0\hat{\beta}}) \\
&= E\left(u_0 - \boldsymbol{X_0}(\boldsymbol{\hat{\beta}} - \boldsymbol{\beta})\right) \\
&= E\left(u_0 - \boldsymbol{X_0}(\boldsymbol{X'X})^{-1}\boldsymbol{X'u}\right) \\
&= 0
\end{aligned}
$$

提示：

$$
\begin{aligned}
\boldsymbol{\hat{\beta}} &= \left(\boldsymbol{X'X}\right)^{-1}\boldsymbol{X'y} \\
&= \left(\boldsymbol{X'X}\right)^{-1}\boldsymbol{X'}(\boldsymbol{X\beta} + \boldsymbol{u}) \\
&= \left(\boldsymbol{X'X}\right)^{-1}\boldsymbol{X'X\beta} + \left(\boldsymbol{X'X}\right)^{-1}\boldsymbol{X'u} \\
&= \boldsymbol{\beta} + \left(\boldsymbol{X'X}\right)^{-1}\boldsymbol{X'u}
\end{aligned}
$$

$$
\boldsymbol{\hat{\beta}} - \boldsymbol{\beta} = \left(\boldsymbol{X'X}\right)^{-1}\boldsymbol{X'u}
$$

样本外预测的矩阵方法实现：个值预测

同时，e_0的真实方差为：

$$
\begin{aligned}
var(e_0) &= E(Y_0 - \hat{Y}_0)^2 \\
&= E(e_0^2) \\
&= E\big(u_0 - \boldsymbol{X_0}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{u}\big)^2 \\
&= \sigma^2 \left(1 + \boldsymbol{X_0}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}_0'\right)
\end{aligned}
$$

进一步地，e_0服从如下正态分布：

$$
\begin{aligned}
e_0 &\sim N(\mu_{e_0}, \sigma^2_{e_0}) \\
e_0 &\sim N\left(0, \sigma^2 \left(1 + \boldsymbol{X_0}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}_0'\right)\right)
\end{aligned}
$$

而且 e_0的样本方差为：：

$$
S^2_{Y_0 - \hat{Y}_0} \equiv S^2_{e_0} = \hat{\sigma}^2 \left(1 + X_0(X'X)^{-1}X_0'\right) \qquad \leftarrow \hat{\sigma}^2 = \frac{\boldsymbol{ee}'}{(n-k)}
$$

样本外预测的矩阵方法实现：个值预测

因此可以构造t统计量：

$$T = \frac{\hat{Y}_0 - Y_0}{S_{e_0}} \sim t(n-k)$$

给定显著性水平 α的情况下，可以查表得到理论t值 $t_{1-\alpha/2}(n-k)$，从而可以计算得到均值预测的置信区间：

$$\hat{Y}_0 - t_{1-\alpha/2}(n-2) \cdot S_{Y_0-\hat{Y}_0} \leq (Y_0|X_0) \leq \hat{Y}_0 + t_{1-\alpha/2}(n-2) \cdot S_{Y_0-\hat{Y}_0}$$

其中：

$$S_{Y_0-\hat{Y}_0} = \sqrt{\hat{\sigma}^2\left(1 + (X_0(X'X)^{-1}X_0')\right)} \qquad \leftarrow \hat{\sigma}^2 = \frac{\boldsymbol{e}\boldsymbol{e}'}{(n-k)}$$

$$\hat{Y}_0 = \boldsymbol{X_0}\hat{\boldsymbol{\beta}}$$

6.6 矩阵方法总结（示例）

矩阵方法总结：消费支出案例（数据）

Year	Y	one	X2	X3
1956	1673	1	1839	1
1957	1688	1	1844	2
1958	1666	1	1831	3
1959	1735	1	1881	4
1960	1749	1	1883	5

Showing 1 to 5 of 15 entries

Previous 1 2 3 Next

$$Y_i = \hat{\beta}_1 + \hat{\beta}_2 X_{2i} + \hat{\beta}_3 X_{3i} + e_i$$

- Y_i表示人均私人消费支出
- X_{2i}表示人均可支配收入
- X_{3i}表示时间 $t \in 1, 2, \cdots n$

矩阵方法总结：消费支出案例（软件报告）

我们可以构建如下的回归模型：

$$Y = +\hat{\beta}_1 + \hat{\beta}_2 X2 + \hat{\beta}_3 X3 + e_i$$

统计软件自动计算结果整理如下（便于后续手动计算的比较）：

$$\begin{array}{lllll}
\widehat{Y} = & +300.29 & +0.74X2 & +8.04X3 \\
\text{(t)} & (3.8342) & (15.6096) & (2.6960) \\
\text{(se)} & (78.3176) & (0.0475) & (2.9835) \\
\text{(fitness)} & R^2 = 0.9976; & \bar{R}^2 = 0.9972 \\
 & F^* = 2513.52; p = 0.0000
\end{array}$$

矩阵方法总结：消费支出案例（软件报告）

利用R软件给出更为详细的分析报告如下（便于后续手动计算的比较）：

```
Call:
lm(formula = mod_mat, data = data_PCE)

Residuals:
    Min      1Q  Median      3Q     Max
-22.380  -6.141   3.414   6.686  22.183

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 300.28626   78.31763   3.834  0.00238 **
X2            0.74198    0.04753  15.610 2.46e-09 ***
X3            8.04356    2.98355   2.696  0.01945 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 12.84 on 12 degrees of freedom
Multiple R-squared:  0.9976,    Adjusted R-squared:  0.9972
F-statistic:  2514 on 2 and 12 DF,  p-value: < 2.2e-16
```

矩阵方法总结：消费支出案例（回归系数）

$$\boldsymbol{X}=\begin{bmatrix}1 & 1839 & 1\\ 1 & 1844 & 2\\ 1 & 1831 & 3\\ \cdots & \cdots & \cdots\\ 1 & 2595 & 15\end{bmatrix}$$

$$\boldsymbol{y}=\begin{bmatrix}1673\\ 1688\\ 1666\\ \cdots\\ 2324\end{bmatrix}$$

$$\boldsymbol{X}'\boldsymbol{X}=\begin{bmatrix}15 & 31895 & 120\\ 31895 & 68922513 & 272144\\ 120 & 272144 & 1240\end{bmatrix}$$

$$(\boldsymbol{X}'\boldsymbol{X})^{-1}=\begin{bmatrix}37.2328 & -0.0225 & 1.3367\\ -0.0225 & 0 & -8e-04\\ 1.3367 & -8e-04 & 0.054\end{bmatrix}$$

$$\boldsymbol{X}'\boldsymbol{y}=\begin{bmatrix}29135\\ 62905821\\ 247934\end{bmatrix}$$

$$\hat{\boldsymbol{\beta}}=(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y}=\begin{bmatrix}300.2863\\ 0.742\\ 8.0436\end{bmatrix}$$

求出回归标准误差方差

可以根据如下公式计算回归误差方差 $\hat{\sigma}^2$：

$$\hat{\sigma}^2 = \frac{\sum e_i^2}{n-k} = \frac{\boldsymbol{e}'\boldsymbol{e}}{n-k} = \frac{= \mathbf{y}'\mathbf{y} - \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y}}{n-k}$$

$$\boldsymbol{y}'\boldsymbol{y} = \begin{bmatrix} 1673 & 1688 & 1666 & \cdots & 2324 \end{bmatrix} \begin{bmatrix} 1673 \\ 1688 \\ 1666 \\ \cdots \\ 2324 \end{bmatrix}$$

$$\hat{\boldsymbol{\beta}}'\boldsymbol{X}'\boldsymbol{y} = \begin{bmatrix} 300.2863 & 0.742 & 8.0436 \end{bmatrix} \begin{bmatrix} 29135 \\ 62905821 \\ 247934 \end{bmatrix}$$

求出回归标准误差方差

因此有：

$$
\begin{aligned}
RSS = \sum e_i^2 = \boldsymbol{e}'\boldsymbol{e} = \boldsymbol{y}\boldsymbol{y}' - \hat{\boldsymbol{\beta}}'\boldsymbol{X}'\boldsymbol{y} \\
= 57420003 - 57418026.1446 = 1976.8554
\end{aligned}
$$

$$
\begin{aligned}
\hat{\sigma}^2 &= \frac{\boldsymbol{e}'\boldsymbol{e}}{\boldsymbol{n}-\boldsymbol{k}} \\
&= \frac{1976.8554}{12} = 164.7379
\end{aligned}
$$

求出回归系数的样本方差-协方差矩阵

根据如下公式：

$$
\begin{aligned}
S_{ij}^2(\hat{\boldsymbol{\beta}}) &= \hat{\sigma}^2(\boldsymbol{X}'\boldsymbol{X})^{-1} \\
&= \frac{\boldsymbol{e}'\boldsymbol{e}}{n-k}(\boldsymbol{X}'\boldsymbol{X})^{-1} \\
&= \frac{\mathbf{y}'\mathbf{y} - \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y}}{n-k} \cdot (\boldsymbol{X}'\boldsymbol{X})^{-1}
\end{aligned}
$$

可以计算得到回归系数的方差协方差矩阵为：

$$
S_{ij}^2(\hat{\boldsymbol{\beta}}) = \hat{\sigma}^2(\boldsymbol{X}'\boldsymbol{X})^{-1} = \begin{bmatrix} 6133.6505 & -3.7079 & 220.2063 \\ -3.7079 & 0.0023 & -0.1371 \\ 220.2063 & -0.1371 & 8.9015 \end{bmatrix}
$$

回归系数的显著性检验（t检验）

```
SS_b <- matrix(diag(mat_cov_b), byrow = F)
S_b <- sqrt(SS_b)
```

回归系数的方差协方差矩阵中：

对角线元素即为回归系数的样本方差 $S^2_{\hat{\beta}}$:

$$S^2_{\hat{\beta}} = \begin{bmatrix} 6133.6505 \\ 0.0023 \\ 8.9015 \end{bmatrix}$$

则回归系数的样本标准差 $S_{\hat{\beta}}$ 为：

$$S_{\hat{\beta}} = \begin{bmatrix} 78.3176 \\ 0.0475 \\ 2.9835 \end{bmatrix}$$

回归系数的显著性检验（t检验）

根据原假设 H_0，可以得到：

$$t^*_{\hat{\beta}} = \frac{\hat{\beta}}{\sqrt{S^2_{ij}(\hat{\beta}_{kk})}} = \frac{\hat{\beta}}{S_{\hat{\beta}}}$$

计算结果为：

$$t^*_{\hat{\beta}} = \begin{bmatrix} 3.8342 \\ 15.6096 \\ 2.696 \end{bmatrix}$$

给定 $\alpha = 0.05, k = 3, n = 15$，我们可以查表得 $t_{1-\alpha/2}(n-k) = t_{0.975}(13) = 2.1788$。

因此表明全部回归系数的t检验的都是显著的。

平方和分解和ANOVA分析表

$$ESS == \boldsymbol{y}\boldsymbol{y}' - \hat{\boldsymbol{\beta}}'\boldsymbol{X}'\boldsymbol{y} = 1976.8554$$

$$RSS = \hat{\boldsymbol{\beta}}'\boldsymbol{X}'\boldsymbol{y} - n\bar{Y}^2 = 1976.8554$$

$$TSS = \boldsymbol{y}'\boldsymbol{y} - n\bar{Y}^2 = 830121.3333$$

变异来源	平方和符号SS	平方和计算公式	自由度df	均方和符号MSS	均方和计算公式
回归平方和	ESS	$\sum \hat{y}_i^2$	2	MSS_{ESS}	$\hat{\boldsymbol{\beta}}'\boldsymbol{X}'\boldsymbol{y} - n\bar{Y}^2$ =828144.4779
残差平方和	RSS	$\sum e_i^2$	12	MSS_{RSS}	$\boldsymbol{y}\boldsymbol{y}' - \hat{\boldsymbol{\beta}}'\boldsymbol{X}'\boldsymbol{y}$=1976.8554
总平方和	TSS	$\sum y_i^2$	14	MSS_{TSS}	$\boldsymbol{y}'\boldsymbol{y} - n\bar{Y}^2$=830121.3333

判定系数和调整判定系数

根据判定系数公式可以计算得到：

$$R^2 = \frac{ESS}{TSS} = \frac{\hat{\boldsymbol{\beta}}'\boldsymbol{X}'\boldsymbol{y} - n\bar{Y}^2}{\boldsymbol{y}'\boldsymbol{y} - n\bar{Y}^2} = \frac{828144.4779}{830121.3333} = 0.9976$$

根据调整判定系数公式可以计算得到：

$$\bar{R}^2 = \frac{ESS}{TSS} = 1 - \frac{(\boldsymbol{y}\boldsymbol{y}' - \hat{\boldsymbol{\beta}}'\boldsymbol{X}'\boldsymbol{y})/(k-1)}{(\boldsymbol{y}'\boldsymbol{y} - n\bar{Y}^2)/(n-1)} = 1 - \frac{1976.8554/12}{830121.3333/14} = 0.9972$$

进行F检验

首先计算得到方差分析表（ANOVA）：

变异来源	平方和符号SS	平方和计算公式	自由度df	均方和符号MSS	均方和计算公式
回归平方和	ESS	$\sum \hat{y}_i^2$	2	MSS_{ESS}	$\hat{\boldsymbol{\beta}}'\boldsymbol{X}'\boldsymbol{y}-n\bar{Y}^2$ =828144.4779
残差平方和	RSS	$\sum e_i^2$	12	MSS_{RSS}	$\boldsymbol{y}\boldsymbol{y}'-\hat{\boldsymbol{\beta}}'\boldsymbol{X}'\boldsymbol{y}$=1976.8554
总平方和	TSS	$\sum y_i^2$	14	MSS_{TSS}	$\boldsymbol{y}'\boldsymbol{y}-n\bar{Y}^2$=830121.3333

进行F检验

根据方差分析表ANOVA和样本F统计量计算公式，可以得到：

$$F^* = \frac{ESS_U/df_{ESS_U}}{RSS_U/df_{RSS_U}} = \frac{\left(\hat{\boldsymbol{\beta}}'\boldsymbol{X}'\boldsymbol{y} - n\bar{Y}^2\right)/(k-1)}{\left(\boldsymbol{y}\boldsymbol{y}' - \hat{\boldsymbol{\beta}}'\boldsymbol{X}'\boldsymbol{y}\right)/(n-k)} = \frac{828144.4779/2}{1976.8554/12} = 2513.5207$$

得到显著性检验的判断结论。因为 $F^* = 2513.5207$ 大于 $F_{1-\alpha}(k-1, n-k) = F_{0.95}$(2,12)=3.8853，所以模型整体显著性的F检验结果显著。

进行样本外预测

给定样本外 $\boldsymbol{X_0}$ 矩阵为：

$$\boldsymbol{X_0} = \begin{bmatrix} 1 & 2610 & 16 \end{bmatrix}$$

已经求得斜率向量为：

$$\hat{\boldsymbol{\beta}} = \begin{bmatrix} 300.2863 \\ 0.742 \\ 8.0436 \end{bmatrix}$$

则线性拟合值为：

$$\hat{Y}_0 = \boldsymbol{X_0}\hat{\boldsymbol{\beta}} = 2365.553$$

进行样本外预测（均值预测）

已知：

$$(\boldsymbol{X}'\boldsymbol{X})^{-1} = \begin{bmatrix} 37.2328 & -0.0225 & 1.3367 \\ -0.0225 & 0 & -8e-04 \\ 1.3367 & -8e-04 & 0.054 \end{bmatrix}$$

所以有：

$$\boldsymbol{X}_0(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}_0' = 0.2953$$

因此进一步得到：

$$\boldsymbol{S}_{\hat{\boldsymbol{Y}}_0} = \sqrt{\hat{\sigma}^2 X_0(X'X)^{-1}X_0'} = \sqrt{164.7379 * 0.2953} = 6.9744$$

进行样本外预测（均值预测）

又因为给定 $\alpha = 0.05$时可以查到理论t值 $t_{1-\alpha/2}(n-k) = t_{0.975}(12)$=2.1788

因此可以计算得到均值预测的 $1-\alpha$置信区间为：

$$\hat{Y}_0 - t_{1-\alpha/2}(n-k) \cdot S_{\hat{Y}_0} \leq E(Y|X_0) \leq \hat{Y}_0 + t_{1-\alpha/2}(n-k) \cdot S_{\hat{Y}_0}$$
$$2365.5532 - 2.1788 * 6.9744 \leq E(Y|X_0) \leq 2365.5532 + 2.1788 * 6.9744$$
$$2350.3573 \leq E(Y|X_0) \leq 2380.7492$$

进行样本外预测（个值预测）

已知：

$$(\boldsymbol{X}'\boldsymbol{X})^{-1} = \begin{bmatrix} 37.2328 & -0.0225 & 1.3367 \\ -0.0225 & 0 & -8e-04 \\ 1.3367 & -8e-04 & 0.054 \end{bmatrix}$$

所以有：

$$\boldsymbol{X}_0(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}_0' = 0.2953$$

因此进一步得到：

$$\boldsymbol{S}_{e0} = \sqrt{\hat{\sigma}^2(1 + X_0(X'X)^{-1}X_0')} = \sqrt{164.7379 * (1 + 0.2953)} = 7.0458$$

进行样本外预测（个值预测）

又因为给定 $\alpha=0.05$时可以查到理论t值 $t_{1-\alpha/2}(n-k)=t_{0.975}(12)$=2.1788

因此可以计算得到均值预测的 $1-\alpha$置信区间为：

$$\hat{Y}_0-t_{1-\alpha/2}(n-k)\cdot S_{e0}\leq(Y_0|X_0)\leq\hat{Y}_0+t_{1-\alpha/2}(n-k)\cdot S_{e0}$$
$$2365.5532-2.1788*7.0458\leq(Y_0|X_0)\leq 2365.5532+2.1788*7.0458$$
$$2350.2019\leq(Y_0|X_0)\leq 2380.9046$$

本章结束