高级计量经济学

Lecture 7: Hypothesis Testing

黄嘉平

工学博士 经济学博士
深圳大学中国经济特区研究中心 讲师

办公室 粤海校区汇文楼1510
E-mail huangjp@szu.edu.cn
Website https://huangjp.com

R 的典型回归结果

Kleiber & Zeileis, *Applied Econometrics with R*, Springer.: Section 3.2

```
> data("CPS1988")
> cps_lm <- lm(log(wage) ~ experience + I(experience^2) + education + ethnicity, data = CPS1988)
> summary(cps_lm)

Call:
lm(formula = log(wage) ~ experience + I(experience^2) + education + ethnicity, data = CPS1988)

Residuals:
    Min      1Q  Median      3Q     Max 
-2.9428 -0.3162  0.0580  0.3756  4.3830 

Coefficients:
                  Estimate Std. Error t value Pr(>|t|)    
(Intercept)      4.321e+00  1.917e-02  225.38   <2e-16 ***
experience       7.747e-02  8.800e-04   88.03   <2e-16 ***
I(experience^2) -1.316e-03  1.899e-05  -69.31   <2e-16 ***
education        8.567e-02  1.272e-03   67.34   <2e-16 ***
ethnicityafam   -2.434e-01  1.292e-02  -18.84   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.5839 on 28150 degrees of freedom
Multiple R-squared:  0.3347,Adjusted R-squared:  0.3346 
F-statistic:  3541 on 4 and 28150 DF,  p-value: < 2.2e-16
```

MATLAB 的典型回归结果

```
>> load hospital.mat;
>> mdl = fitlm(hospital,'interactions','ResponseVar','Weight',...
    'PredictorVars',{'Sex','Age','Smoker'},...
    'CategoricalVar',{'Sex','Smoker'});

mdl =


Linear regression model:
    Weight ~ 1 + Sex*Age + Sex*Smoker + Age*Smoker

Estimated Coefficients:
                         Estimate      SE        tStat       pValue  
                         ________    _______    ________    __________

    (Intercept)             118.7     7.0718      16.785     6.821e-30
    Sex_Male               68.336     9.7153      7.0339    3.3386e-10
    Age                   0.31068    0.18531      1.6765      0.096991
    Smoker_1               3.0425     10.446     0.29127       0.77149
    Sex_Male:Age         -0.49094    0.24764     -1.9825      0.050377
    Sex_Male:Smoker_1      0.9509     3.8031     0.25003       0.80312
    Age:Smoker_1         -0.07288    0.26275    -0.27737       0.78211


Number of observations: 100, Error degrees of freedom: 93
Root Mean Squared Error: 8.75
R-squared: 0.898,  Adjusted R-Squared: 0.892
F-statistic vs. constant model: 137, p-value = 6.91e-44
```

应用研究中的典型回归结果

Busse et al, (2015). The Psychological Effect of Weather on Car Purchases. *QJE*, 130(1):371-414.
https://doi.org/10.1093/qje/qju033

TABLE III

EFFECT OF WEATHER ON FOUR-WHEEL-DRIVE PURCHASES

	Dependent variable: indicator equal to 1 if purchase was a four-wheel-drive				
	Full year	Quarter 1	Quarter 2	Quarter 3	Quarter 4
Temperature	−.032**	−.038**	−.018**	−.029**	−.038**
	(.001)	(.002)	(.003)	(.004)	(.003)
Rainfall	.084**	.119**	.081**	.054**	.132**
	(.014)	(.036)	(.026)	(.023)	(.032)
Snowfall	1.81**	1.67**	.72	125*	2.11**
	(.26)	(.33)	(.82)	(53)	(.48)
Slushfall	.504**	.540**	.27	−.029	.769**
	(.077)	(.110)	(.22)	(.167)	(.166)
Cloud cover	.461**	.337**	.512**	.383**	.598**
	(.039)	(.072)	(.077)	(.087)	(.076)
Year fixed effects	X	X	X	X	X
DMA*week-of-the-year fixed effects	X	X	X	X	X
R-squared	0.086	0.086	0.074	0.084	0.097
Observations	39,984,509	9,150,047	10,205,673	10,608,527	10,020,262

Notes. Coefficient values and clustered standard errors are presented from OLS regressions of an indicator for whether a car sold was a four-wheel-drive on weather variables: temperature (°F), rain (inches), snow (liquidized inches), slush (liquidized inches), and cloud cover (fraction of sky covered). Fixed effects are included for each year and for DMA*week-of-the-year (week 1–week 52). The first column uses all the data and the next four columns present results separately for the four quarters of the year. All coefficients and standard errors have been multiplied by 100 for ease of presentation. Thus each coefficient represents the percentage point change in probability of purchasing a covertible. Standard errors are clustered at the DMA*day. * significant at 5%; ** significant at 1%.

古典正态模型下的精确检验

精确检验与古典正态模型

Exact Tests and the Classical Normal Linear Model

当检验统计量在 H_0 下的分布为已知时，该检验被称为精确检验（exact test）。

假设古典正态线性模型：

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{u}\ , \quad \boldsymbol{u} \sim N(\mathbf{0}, \sigma^2 \boldsymbol{I})$$

其中 $\boldsymbol{u}$ 独立于 $\boldsymbol{X}$（$\boldsymbol{X}$ 固定或在 $\boldsymbol{y}$ 之前生成）。

"古典" 一般指 $\boldsymbol{X}$ 为固定变量或 $\boldsymbol{X}$ 与 $\boldsymbol{u}$ 独立，在此假设下我们只需要考虑普通的期望值。"正态" 指 $\boldsymbol{u}$ 服从正态分布。只有在针对 $\boldsymbol{\beta}$ 进行假设检验时才需要关于分布的完整信息。

单一约束的检验

Tests for a Single Restriction

首先我们考虑针对单个系数的假设检验。我们可以把系数向量写成 $\boldsymbol{\beta} = [\boldsymbol{\beta}_1^\top, \beta_2]^\top$，并针对 $H_0 : \beta_2 = 0$ 进行检验。

这里的 $\beta_2 = 0$ 可以看作系数的线性约束条件，而零假设和备择假设分别是在有约束和没有约束下的回归。

原回归模型可以写成

$$\boldsymbol{y} = \boldsymbol{X}_1\boldsymbol{\beta}_1 + \boldsymbol{x}_2\beta_2 + \boldsymbol{u} \ , \quad \boldsymbol{u} \sim N(\boldsymbol{0}, \sigma^2\boldsymbol{I})$$

$\boldsymbol{X}_1$ 是 $n \times (k-1)$ 矩阵，$\boldsymbol{x}_2$ 是 $n \times 1$ 向量，$\boldsymbol{X} = [\boldsymbol{X}_1, \boldsymbol{x}_2]$。

根据 FWL 定理，$\hat{\beta}_2$ 和 $\hat{\boldsymbol{u}}$ 可以从下面的短模型获得

$$\boldsymbol{M}_1\boldsymbol{y} = \boldsymbol{M}_1\boldsymbol{x}_2\beta_2 + \boldsymbol{v} \ , \quad \boldsymbol{M}_1 = \boldsymbol{I} - \boldsymbol{X}_1(\boldsymbol{X}_1^\top\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1^\top$$

此时可得 $\hat{\beta}_2 = \dfrac{\boldsymbol{x}_2^\top\boldsymbol{M}_1\boldsymbol{y}}{\boldsymbol{x}_2^\top\boldsymbol{M}_1\boldsymbol{x}_2}, \qquad \mathrm{Var}[\hat{\beta}_2] = \dfrac{\sigma^2}{\boldsymbol{x}_2^\top\boldsymbol{M}_1\boldsymbol{x}_2}$

单一约束的检验：σ^2 已知

当 σ^2 已知时，我们可以用下面的统计量检验 $H_0 : \beta_2 = 0$

$$z_{\beta_2} \equiv \frac{\hat{\beta}_2 - 0}{\sqrt{\mathrm{Var}[\hat{\beta}_2]}} = \frac{\boldsymbol{x}_2^\top \boldsymbol{M}_1 \boldsymbol{y}}{\boldsymbol{x}_2^\top \boldsymbol{M}_1 \boldsymbol{x}_2} \cdot \frac{\sqrt{\boldsymbol{x}_2^\top \boldsymbol{M}_1 \boldsymbol{x}_2}}{\sigma} = \frac{\boldsymbol{x}_2^\top \boldsymbol{M}_1 \boldsymbol{y}}{\sigma\sqrt{\boldsymbol{x}_2^\top \boldsymbol{M}_1 \boldsymbol{x}_2}}$$

当 $H_0 : \beta_2 = 0$ 成立时，$\boldsymbol{M}_1 \boldsymbol{y} = \boldsymbol{M}_1 \boldsymbol{X}_1 \boldsymbol{\beta}_1 + \boldsymbol{M}_1 \boldsymbol{u} = \boldsymbol{M}_1 \boldsymbol{u}$，因此

$$z_{\beta_2} = \frac{\boldsymbol{x}_2^\top \boldsymbol{M}_1 \boldsymbol{u}}{\sigma\sqrt{\boldsymbol{x}_2^\top \boldsymbol{M}_1 \boldsymbol{x}_2}}$$

z_{β_2} 是 $u_1, \ldots, u_n$ 的线性结合，因此服从正态分布。$E[z_{\beta_2}] = 0$,

$$\mathrm{Var}[z_{\beta_2}] = \frac{E[\boldsymbol{x}_2^\top \boldsymbol{M}_1 \boldsymbol{u}\boldsymbol{u}^\top \boldsymbol{M}_1 \boldsymbol{x}_2]}{\sigma^2 \boldsymbol{x}_2^\top \boldsymbol{M}_1 \boldsymbol{x}_2} = \frac{\boldsymbol{x}_2^\top \boldsymbol{M}_1 E[\boldsymbol{u}\boldsymbol{u}^\top] \boldsymbol{M}_1 \boldsymbol{x}_2}{\sigma^2 \boldsymbol{x}_2^\top \boldsymbol{M}_1 \boldsymbol{x}_2} = \frac{\sigma^2 \boldsymbol{x}_2^\top \boldsymbol{M}_1 \boldsymbol{x}_2}{\sigma^2 \boldsymbol{x}_2^\top \boldsymbol{M}_1 \boldsymbol{x}_2} = 1$$

所以 $z_{\beta_2} \sim N(0,1)$。

单一约束的检验：σ^2 未知

当 σ^2 未知时，我们可以用回归标准误 s 替代 σ。因为 $s^2 = \frac{1}{n-k}\sum_{t=1}^{n}\hat{u}_t^2 = \frac{\boldsymbol{y}^\top \boldsymbol{M}_X \boldsymbol{y}}{n-k}$，我们可以定义下面的统计量

$$t_{\beta_2} \equiv \frac{\hat{\beta}_2 - 0}{\sqrt{\widehat{\mathrm{Var}}[\hat{\beta}_2]}} = \frac{\boldsymbol{x}_2^\top \boldsymbol{M}_1 \boldsymbol{y}}{s\sqrt{\boldsymbol{x}_2^\top \boldsymbol{M}_1 \boldsymbol{x}_2}} = \sqrt{\frac{n-k}{\boldsymbol{y}^\top \boldsymbol{M}_X \boldsymbol{y}}} \cdot \frac{\boldsymbol{x}_2^\top \boldsymbol{M}_1 \boldsymbol{y}}{\sqrt{\boldsymbol{x}_2^\top \boldsymbol{M}_1 \boldsymbol{x}_2}} = \frac{z_{\beta_2}}{\sqrt{\frac{\boldsymbol{y}^\top \boldsymbol{M}_X \boldsymbol{y}}{(n-k)\sigma^2}}}$$

已知 $z_{\beta_2} \sim N(0,1)$，如果我们能证明 $\boldsymbol{y}^\top \boldsymbol{M}_X \boldsymbol{y} \big/ \sigma^2 \sim \chi^2(n-k)$，且 z_{β_2} 和 $\boldsymbol{y}^\top \boldsymbol{M}_X \boldsymbol{y} \big/ \sigma^2$ 独立，即可得到 $t_{\beta_2} \sim t(n-k)$。

$\boldsymbol{y}^\top \boldsymbol{M}_X \boldsymbol{y} \big/ \sigma^2 \sim \chi^2(n-k)$：因为 $\boldsymbol{M}_X$ 是投影到 $\mathcal{S}^\perp(\boldsymbol{X})$ 的投影矩阵，所以

$$\frac{\boldsymbol{y}^\top \boldsymbol{M}_X \boldsymbol{y}}{\sigma^2} = \frac{\boldsymbol{u}^\top \boldsymbol{M}_X \boldsymbol{u}}{\sigma^2} = \boldsymbol{\varepsilon}^\top \boldsymbol{M}_X \boldsymbol{\varepsilon}, \qquad \boldsymbol{\varepsilon} = \boldsymbol{u}/\sigma \sim N(\boldsymbol{0}, \boldsymbol{I})$$

$\dim(\mathcal{S}^\perp(\boldsymbol{X})) = n-k \;\; \Rightarrow \;\; \mathrm{rank}(\boldsymbol{M}_X) = n-k$，因此 $\boldsymbol{y}^\top \boldsymbol{M}_X \boldsymbol{y} \big/ \sigma^2 \sim \chi^2(n-k)$。

单一约束的检验：σ^2 未知

z_{β_2} 和 $\boldsymbol{y}^\top \boldsymbol{M}_X \boldsymbol{y} / \sigma^2$ 独立：两个变量中随机要素全都来自 $\boldsymbol{u}$。$\boldsymbol{y}^\top \boldsymbol{M}_X \boldsymbol{y} / \sigma^2$ 只通过 $\boldsymbol{M}_X \boldsymbol{u}$ 受 $\boldsymbol{u}$ 的影响，而由

$$\boldsymbol{x}_2^\top \boldsymbol{M}_1 \boldsymbol{u} = \boldsymbol{x}_2^\top \boldsymbol{P}_X \boldsymbol{M}_1 \boldsymbol{u} = \boldsymbol{x}_2^\top (\boldsymbol{P}_X - \boldsymbol{P}_X \boldsymbol{P}_1) \boldsymbol{u} \overset{\boldsymbol{P}_X \boldsymbol{P}_1 = \boldsymbol{P}_1 \boldsymbol{P}_X}{=} \boldsymbol{x}_2^\top \boldsymbol{M}_1 \boldsymbol{P}_X \boldsymbol{u}$$

可知 z_{β_2} 只通过 $\boldsymbol{P}_X \boldsymbol{u}$ 受 $\boldsymbol{u}$ 的影响。

$\boldsymbol{M}_X \boldsymbol{u}$ 与 $\boldsymbol{P}_X \boldsymbol{u}$ 的期望值都为 $\mathbf{0}$，因此它们的协方差是

$$E[\boldsymbol{M}_X \boldsymbol{u} \boldsymbol{u}^\top \boldsymbol{P}_X] = \boldsymbol{M}_X E[\boldsymbol{u} \boldsymbol{u}^\top] \boldsymbol{P}_X = \sigma^2 \boldsymbol{M}_X \boldsymbol{P}_X \overset{\boldsymbol{P}_X \boldsymbol{M}_X = \boldsymbol{O}}{=} \boldsymbol{O}$$

因此 $\boldsymbol{M}_X \boldsymbol{u}$ 与 $\boldsymbol{P}_X \boldsymbol{u}$ 不相关。因为 $\boldsymbol{M}_X \boldsymbol{u}$ 与 $\boldsymbol{P}_X \boldsymbol{u}$ 都服从正态分布，所以 $\boldsymbol{M}_X \boldsymbol{u}$ 独立于 $\boldsymbol{P}_X \boldsymbol{u}$，所以。$\boldsymbol{y}^\top \boldsymbol{M}_X \boldsymbol{y} / \sigma^2$ 独立于 z_{β_2}。

综上，我们可得 $t_{\beta_2} \sim t(n-k)$。

多重约束的联合检验

Joint Tests for Multiple Restrictions

当存在 r 个线性约束条件时（$r \leq k$），我们可以将其写成 $H_0 : \boldsymbol{\beta}_2 = \mathbf{0}$。此时，在不同的假设下我们可以得到下面的回归模型：

$$H_0 : \boldsymbol{y} = \boldsymbol{X}_1\boldsymbol{\beta}_1 + \boldsymbol{u} \;, \quad \boldsymbol{u} \sim N(\mathbf{0}, \sigma^2\boldsymbol{I})$$
$$H_1 : \boldsymbol{y} = \boldsymbol{X}_1\boldsymbol{\beta}_1 + \boldsymbol{X}_2\boldsymbol{\beta}_2 + \boldsymbol{u} \;, \quad \boldsymbol{u} \sim N(\mathbf{0}, \sigma^2\boldsymbol{I})$$

因为需要检验的参数大于 1，我们不能继续使用 z 检验或 t 检验，而需要找到另一个合适的检验统计量。这里我们用有约束（restricted）回归和无约束（unrestricted）回归的拟合（fit）结果来建立检验统计量，并用两个回归的残差平方和作为拟合的指标。

若将有约束回归的残差平方和写为 RSSR，将无约束回归的写为 USSR，我们可以定义下面的统计量

$$F_{\boldsymbol{\beta}_2} \equiv \frac{(\text{RSSR} - \text{USSR})\big/ r}{\text{USSR}\big/ (n-k)}$$

如果 H_0 不成立，RSSR 将大于 USSR。
因此当 $F_{\boldsymbol{\beta}_2}$ 取值很大时，我们倾向于拒绝 H_0。

多重约束的联合检验

Joint Tests for Multiple Restrictions

我们将要证明 $F_{\boldsymbol{\beta}_2} \equiv \dfrac{(\mathrm{RSSR}-\mathrm{USSR})/r}{\mathrm{USSR}/(n-k)}$ 在 H_0 下服从 $F(r,n-k)$。

$\mathrm{RSSR}=\boldsymbol{y}^\top\boldsymbol{M}_1\boldsymbol{y}$，$\mathrm{USSR}=\boldsymbol{y}^\top\boldsymbol{M}_X\boldsymbol{y}$。根据 FWL 定理，USSR 等于短模型 $\boldsymbol{M}_1\boldsymbol{y}=\boldsymbol{M}_1\boldsymbol{X}_2\boldsymbol{\beta}_2+\boldsymbol{v}$ 的残差平方和

$$\hat{\boldsymbol{v}}^\top\hat{\boldsymbol{v}}=\boldsymbol{y}^\top\boldsymbol{M}_1\boldsymbol{y}-\boldsymbol{y}^\top\boldsymbol{M}_1\boldsymbol{X}_2(\boldsymbol{X}_2^\top\boldsymbol{M}_1\boldsymbol{X}_2)^{-1}\boldsymbol{X}_2^\top\boldsymbol{M}_1\boldsymbol{y}$$

勾股定理

因此，$\mathrm{RSSR}-\mathrm{USSR}=\boldsymbol{y}^\top\boldsymbol{M}_1\boldsymbol{X}_2(\boldsymbol{X}_2^\top\boldsymbol{M}_1\boldsymbol{X}_2)^{-1}\boldsymbol{X}_2^\top\boldsymbol{M}_1\boldsymbol{y}$。一般情况下 $\boldsymbol{M}_X\boldsymbol{y}=\boldsymbol{M}_X\boldsymbol{u}$，在 H_0 成立时 $\boldsymbol{M}_\mathbf{1}\boldsymbol{y}=\boldsymbol{M}_\mathbf{1}\boldsymbol{u}$，则

$$F_{\boldsymbol{\beta}_2}=\frac{\boldsymbol{\varepsilon}^\top\boldsymbol{M}_1\boldsymbol{X}_2(\boldsymbol{X}_2^\top\boldsymbol{M}_1\boldsymbol{X}_2)^{-1}\boldsymbol{X}_2^\top\boldsymbol{M}_1\boldsymbol{\varepsilon}\big/r}{\boldsymbol{\varepsilon}^\top\boldsymbol{M}_X\boldsymbol{\varepsilon}\big/(n-k)},\qquad \boldsymbol{\varepsilon}=\boldsymbol{u}/\sigma\sim N(\mathbf{0},\boldsymbol{I})$$

$\mathrm{rank}(\boldsymbol{X}_2)=r$

已知 $\boldsymbol{\varepsilon}^\top\boldsymbol{M}_X\boldsymbol{\varepsilon}\sim\chi^2(n-k)$，而分子可以写成 $\boldsymbol{\varepsilon}^\top\boldsymbol{P}_{\boldsymbol{M}_1\boldsymbol{X}_2}\boldsymbol{\varepsilon}$ 并服从 $\chi^2(r)$。又因为 $\boldsymbol{M}_X$ 和 $\boldsymbol{P}_{\boldsymbol{M}_1\boldsymbol{X}_2}$ 正交，加上 $\boldsymbol{u}$ 服从正态分布可得分子与分母相互独立。因此 $F_{\boldsymbol{\beta}_2}\sim F(r,n-k)$。

大样本检验

精确检验的条件

我们把精确检验（t 检验和 F 检验）所需的条件总结如下：

- $\boldsymbol{X}$ 与 $\boldsymbol{u}$ 独立
- $\boldsymbol{u} \sim N(\mathbf{0}, \sigma^2 \boldsymbol{I})$

如果以上条件不能被满足，就无法获得 t 统计量或 F 统计量的精确分布。

在大样本（$n \to \infty$）下，我们可以获得检验统计量的渐进分布。

大样本下的单一约束检验

我们假设回归模型满足 $\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{u}$，$\boldsymbol{u} \sim \mathrm{IID}\,(\boldsymbol{0}, \sigma_0^2 \boldsymbol{I})$，其中误差项满足 $E[u_t \mid \boldsymbol{X}_t] = 0$，$E[u_t^2 \mid \boldsymbol{X}_t] = \sigma_0^2$。同时假设 $\operatorname*{plim}_{n\to\infty} \frac{1}{n}\boldsymbol{X}^\top\boldsymbol{X} = \boldsymbol{S}_{\boldsymbol{X}^\top\boldsymbol{X}}$，$\boldsymbol{S}_{\boldsymbol{X}^\top\boldsymbol{X}}$ 是有限非随机正定矩阵。在这个假设下，OLS 估计量 $\hat{\boldsymbol{\beta}}$ 满足一致性。

针对 $H_0 : \beta_2 = 0$，已知 t 统计量可以写成

$$t_{\beta_2} = \sqrt{\frac{n-k}{\boldsymbol{y}^\top \boldsymbol{M}_X \boldsymbol{y}}} \cdot \frac{\boldsymbol{x}_2^\top \boldsymbol{M}_1 \boldsymbol{y} / \sqrt{n}}{\sqrt{\boldsymbol{x}_2^\top \boldsymbol{M}_1 \boldsymbol{x}_2 / n}}$$

根据 LLN，第一项依概率收敛于 $\frac{1}{\sigma_0}$，同时在 H_0 成立时 $\boldsymbol{M}_1\boldsymbol{y} = \boldsymbol{M}_1\boldsymbol{u}$，因此

$$t_{\beta_2} \xrightarrow{p} \frac{\boldsymbol{x}_2^\top \boldsymbol{M}_1 \boldsymbol{u} / \sqrt{n}}{\sigma_0 \sqrt{\boldsymbol{x}_2^\top \boldsymbol{M}_1 \boldsymbol{x}_2 / n}}$$

右侧概率极限的分子符合 CLT 的形式，且其期望值为零，方差等于分母，因此可得 $t_{\beta_2} \overset{a}{\sim} N(0,1)$（渐进分布是标准正态分布）。

（详见 pp. 152-154）

大样本下的多重约束检验

已知多重约束检验的 F 统计量是

$$F_{\beta_2} = \frac{\boldsymbol{\varepsilon}^\top \boldsymbol{M}_1 \boldsymbol{X}_2 (\boldsymbol{X}_2^\top \boldsymbol{M}_1 \boldsymbol{X}_2)^{-1} \boldsymbol{X}_2^\top \boldsymbol{M}_1 \boldsymbol{\varepsilon} \big/ r}{\boldsymbol{\varepsilon}^\top \boldsymbol{M}_X \boldsymbol{\varepsilon} \big/ (n-k)}, \qquad \boldsymbol{\varepsilon} = \boldsymbol{u}/\sigma_0$$

可以将其改写成

$$F_{\beta_2} = \frac{n^{-1/2} \boldsymbol{\varepsilon}^\top \boldsymbol{M}_1 \boldsymbol{X}_2 (n^{-1} \boldsymbol{X}_2^\top \boldsymbol{M}_1 \boldsymbol{X}_2)^{-1} n^{-1/2} \boldsymbol{X}_2^\top \boldsymbol{M}_1 \boldsymbol{\varepsilon} \big/ r}{\boldsymbol{\varepsilon}^\top \boldsymbol{M}_X \boldsymbol{\varepsilon} \big/ (n-k)}$$

当 $n \to \infty$ 时，可得 $rF_{\beta_2} \sim \chi^2(r)$，或 $F_{\beta_2} \sim F(r, \infty)$。

（尝试推导此结论）

置信域

线性回归系数的置信域

Confidence Region of Linear Regression Coefficients

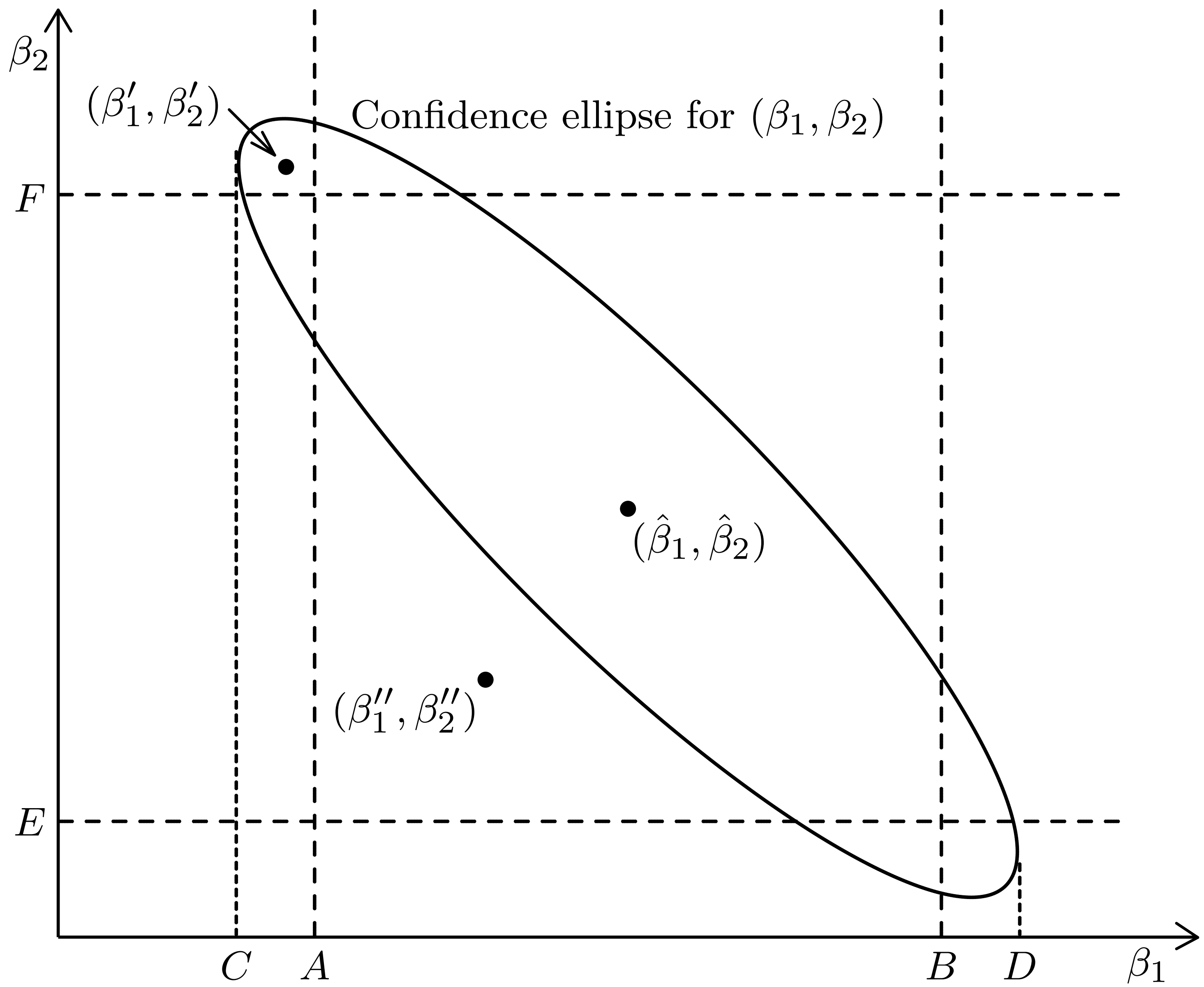

Figure 5.3 Confidence ellipses and confidence intervals

多变量估计量的置信域通常可以写成

$$\text{变量的二次函数} \leq C$$

的形式，其图形含义为椭圆或椭圆体。

注意图中置信域和一维置信区间的区别！

异方差稳健统计量

异方差性及其影响

Heteroskedasticity and its Consequences

异方差性：$\mathrm{Var}[\boldsymbol{u} \mid \boldsymbol{X}] = \boldsymbol{\Omega}$，$\boldsymbol{\Omega}$ 的非对角要素为零，对角要素 ω_t^2 不相同。

在外生性成立时，$\hat{\boldsymbol{\beta}}$ 的协方差矩阵可以写成

$$\begin{aligned}\mathrm{Var}[\hat{\boldsymbol{\beta}}] &= E\left[(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0)(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0)^\top\right] \\ &= (\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top\boldsymbol{\Omega}\,\boldsymbol{X}(\boldsymbol{X}^\top\boldsymbol{X})^{-1}\end{aligned}$$

最后一行的表达式被称为 sandwich covariance matrix。 详见 Lecture 6

异方差性的影响：

- OLS 估计量 $\hat{\boldsymbol{\beta}}$ 不再是最有效的，但还是一致的。
- $s^2(\boldsymbol{X}^\top\boldsymbol{X})^{-1}$ 不再是协方差矩阵的非偏估计量，因此影响假设检验的准确性。

异方差时的一致估计

Consistent Estimation under Heteroskedasticity

当 ω_t^2 未知时，我们通常需要对其进行估计。但是我们只有 n 个观测值，却需要估计 n 个 ω_t^2，因此无法直接得到 $\boldsymbol{\Omega}$ 的一致估计量。

但是我们可以估计 OLS 估计量的协方差矩阵。这里我们用 $\sqrt{n}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0)$ 替代 $\hat{\boldsymbol{\beta}}$，则有

$$\begin{aligned}\operatorname{Var}\left[\sqrt{n}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0)\right] &= E\left[n(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0)(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0)^\top\right] \\ &= (\tfrac{1}{n}\boldsymbol{X}^\top\boldsymbol{X})^{-1}(\tfrac{1}{n}\boldsymbol{X}^\top\boldsymbol{\Omega}\,\boldsymbol{X})(\tfrac{1}{n}\boldsymbol{X}^\top\boldsymbol{X})^{-1}\end{aligned}$$

已知 $\operatorname*{plim}_{n\to\infty} \frac{1}{n}(\boldsymbol{X}^\top\boldsymbol{X})^{-1} = (\boldsymbol{S}_{\boldsymbol{X}^\top\boldsymbol{X}})^{-1}$，我们可以用 $\frac{1}{n}(\boldsymbol{X}^\top\boldsymbol{X})^{-1}$ 作为该极限的一致估计量。

中间项的极限 $\operatorname*{plim}_{n\to\infty} \frac{1}{n}\boldsymbol{X}^\top\boldsymbol{\Omega}\,\boldsymbol{X}$ 是 $k \times k$ 的对称矩阵，因此只有 $\frac{1}{2}k(k+1)$ 个参数需要估计。在一定条件下，我们可以通过 $\boldsymbol{\Omega}$ 的某些非一致估计量 $\hat{\boldsymbol{\Omega}}$ 对该项进行一致估计，即 $\frac{1}{n}\boldsymbol{X}^\top\hat{\boldsymbol{\Omega}}\,\boldsymbol{X}$。（White, 1980）

在实际应用中，我们可以忽略 $1/n$ 而直接估计 $\hat{\boldsymbol{\beta}}$ 的协方差矩阵：

$$\widehat{\operatorname{Var}}_{\mathrm{h}}[\hat{\boldsymbol{\beta}}] = (\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top\hat{\boldsymbol{\Omega}}\,\boldsymbol{X}(\boldsymbol{X}^\top\boldsymbol{X})^{-1}$$

这种估计量被称为 heteroskedasticity-consistent covariance matrix estimator (HCCME)，或 heteroskedasticity-robust estimator。

HCCMEs

HCCME 的关键是如何找到合适的估计量 $\hat{\boldsymbol{\Omega}}$。因为 $\boldsymbol{\Omega}$ 是对角矩阵，$\hat{\boldsymbol{\Omega}}$ 也是对角矩阵。下面通过定义 $\hat{\boldsymbol{\Omega}}$ 的第 t 对角要素介绍几种常用的 HCCME。

- $HC_0 : \hat{u}_t^2$
- $HC_1 : \frac{n}{n-k}\hat{u}_t^2$
- $HC_2 : \hat{u}_t^2 \big/ (1-h_t)$，$h_t$ 是 $\boldsymbol{P_X}$ 的第 t 对角要素
- $HC_3 : \hat{u}_t^2 \big/ (1-h_t)^2$

这四个 HCCME 都满足一致性，但在有限样本下表现都不够好。四个当中 HC_0 表现最差，HC_2 或 HC_3 表现最好。

需要注意的是，有些软件里的默认设定是使用 HC_0，在实践操作中需要人为指定。

课外阅读

- White, H. (1980).
  A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity.
  Econometrica, 48:4, 817-838.
  http://www.jstor.org/stable/1912934

- MacKinnon, J. G. and White, H. (1985).
  Some Heteroskedasticity-Consistent Covariance Matrix Estimators with Improved Finite Sample Properties.
  Journal of Econometrics, 29:3, 305-325.
  https://doi.org/10.1016/0304-4076(85)90158-7

- MacKinnon, J. G. (2005).
  Thirty Years of Heteroskedasticity-Robust Inference.
  In: Chen, X. and Swanson, N. R. (eds.), *Recent Advances and Future Directions 437 in Causality, Prediction, and Specification Analysis*, 437-461, Springer.
  https://doi.org/10.1007/978-1-4614-1653-1_17