2010 年度 学部 エコノメトリックス 第 12 回講義メモ (訂正版)

2010 年 11 月 16 日

3.6 定式化の誤りとその影響

a) overspecification(説明変数過剰) の場合

いま、

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{Z}\boldsymbol{\gamma} + \boldsymbol{\epsilon} \quad \text{where } \boldsymbol{\epsilon} \sim IID(\boldsymbol{0}, \sigma^2 \boldsymbol{I}) \tag{3.6.1}$$

というモデルを考えよう。ここで、$\boldsymbol{X}$ は $n \times k$、$\boldsymbol{Z}$ は $n \times \ell$ の説明変数行列であるとする。(3.6.1) 式は

$$\boldsymbol{y} = \begin{pmatrix} \boldsymbol{X} & \boldsymbol{Z} \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta} \\ \boldsymbol{\gamma} \end{pmatrix} + \boldsymbol{u} \tag{3.6.2}$$

と変形できるから、(3.6.2) 式における $\boldsymbol{\beta}, \boldsymbol{\gamma}$ の OLS 推定量 $\tilde{\boldsymbol{b}}, \tilde{\boldsymbol{c}}$ は、

$$\begin{pmatrix} \tilde{\boldsymbol{b}} \\ \tilde{\boldsymbol{c}} \end{pmatrix} = \left[\begin{pmatrix} \boldsymbol{X} & \boldsymbol{Z} \end{pmatrix}' \begin{pmatrix} \boldsymbol{X} & \boldsymbol{Z} \end{pmatrix} \right]^{-1} \begin{pmatrix} \boldsymbol{X} & \boldsymbol{Z} \end{pmatrix}' \boldsymbol{y}$$

$$= \begin{pmatrix} \boldsymbol{X}'\boldsymbol{X} & \boldsymbol{X}'\boldsymbol{Z} \\ \boldsymbol{Z}'\boldsymbol{X} & \boldsymbol{Z}'\boldsymbol{Z} \end{pmatrix}^{-1} \begin{pmatrix} \boldsymbol{X}'\boldsymbol{y} \\ \boldsymbol{Z}'\boldsymbol{y} \end{pmatrix}$$

で表される。

分割行列の逆行列については、次の補題が知られている。

補題 3.6.1 matrix inversion lemma

$\boldsymbol{B}$ を $m \times m$ の行列とし、次のような小行列に分割できるとする。

$$\boldsymbol{B} = \begin{pmatrix} \boldsymbol{B}_{11} & \boldsymbol{B}_{12} \\ \boldsymbol{B}_{21} & \boldsymbol{B}_{22} \end{pmatrix}$$

ここで、$\boldsymbol{B}_{11}$ は $m_1 \times m_1$ 行列、$\boldsymbol{B}_{12}$ は $m_1 \times m_2$ 行列、$\boldsymbol{B}_{21}$ は $m_2 \times m_1$ 行列、$\boldsymbol{B}_{22}$ は $m_2 \times m_2$ 行列であり、$m_1 + m_2 = m$ である。

このとき、$\boldsymbol{B}$ に逆行列が存在すれば、小行列 $\boldsymbol{B}_{ij}$, $(i, j = 1, 2)$ を用い

て以下のように表現できる。

$$\boldsymbol{B}^{-1} = \begin{pmatrix} \boldsymbol{C}^{-1} & -\boldsymbol{C}^{-1}\boldsymbol{B}_{12}\boldsymbol{B}_{22}^{-1} \\ -\boldsymbol{B}_{22}^{2}\boldsymbol{B}_{21}\boldsymbol{C}^{-1} & \boldsymbol{B}_{22}^{-1} + \boldsymbol{B}_{22}^{-1}\boldsymbol{B}_{21}\boldsymbol{C}^{-1}\boldsymbol{B}_{12}\boldsymbol{B}_{22}^{-1} \end{pmatrix} \tag{3.6.3}$$

ここで、$\boldsymbol{C} = \boldsymbol{B}_{11} - \boldsymbol{B}_{12}\boldsymbol{B}_{22}^{-1}\boldsymbol{B}_{21}$ である。

また、$\boldsymbol{B}_{12} = \boldsymbol{B}_{21} = \boldsymbol{O}$ のときには、

$$\boldsymbol{B}^{-1} = \begin{pmatrix} \boldsymbol{B}_{11}^{-1} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{B}_{22}^{-1} \end{pmatrix} \tag{3.6.4}$$

になる。

補題 3.6.1 の結果を用いると、

$$\begin{aligned} \boldsymbol{W} &= \boldsymbol{X}'\boldsymbol{X} - \boldsymbol{X}'\boldsymbol{Z}(\boldsymbol{Z}'\boldsymbol{Z})^{-1}\boldsymbol{Z}'\boldsymbol{X} \\ &= \boldsymbol{X}'(\boldsymbol{I} - \boldsymbol{Z}(\boldsymbol{Z}'\boldsymbol{Z})^{-1}\boldsymbol{Z}')\boldsymbol{X} = \boldsymbol{X}'\boldsymbol{M}_{\boldsymbol{Z}}\boldsymbol{X} \end{aligned}$$

とおくと、$\tilde{\boldsymbol{\beta}}$ は

$$\begin{aligned} \tilde{\boldsymbol{b}} &= \boldsymbol{W}^{-1}\boldsymbol{X}'\boldsymbol{y} - \boldsymbol{W}^{-1}\boldsymbol{X}'\boldsymbol{Z}(\boldsymbol{Z}'\boldsymbol{Z})^{-1}\boldsymbol{Z}'\boldsymbol{y} \\ &= \boldsymbol{W}^{-1}\boldsymbol{X}'(\boldsymbol{I} - \boldsymbol{Z}(\boldsymbol{Z}'\boldsymbol{Z})^{-1}\boldsymbol{Z}')\boldsymbol{y} \\ &= (\boldsymbol{X}'\boldsymbol{M}_{\boldsymbol{Z}}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{M}_{\boldsymbol{Z}}\boldsymbol{y} \end{aligned}$$

で表すことができる。

もし、データ生成プロセス（DGP: Data Generating Process）では $\boldsymbol{\gamma} = \boldsymbol{0}$、すなわち $\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$ であるならば、

$$\begin{aligned} \tilde{\boldsymbol{b}} &= (\boldsymbol{X}'\boldsymbol{M}_{\boldsymbol{Z}}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{M}_{\boldsymbol{Z}}(\boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}) \\ &= \boldsymbol{\beta} + (\boldsymbol{X}'\boldsymbol{M}_{\boldsymbol{Z}}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{M}_{\boldsymbol{Z}}\boldsymbol{\epsilon} \end{aligned}$$

となる。これより、

$$\mathrm{E}(\tilde{\boldsymbol{b}} \mid \boldsymbol{X}, \boldsymbol{Z}) = \boldsymbol{\beta}$$

がえられるので、$\tilde{\boldsymbol{b}}$ は不偏性をみたしている。ところが、$\tilde{\boldsymbol{b}}$ の分散をみると、

$$\mathrm{Var}(\tilde{\boldsymbol{b}} \mid \boldsymbol{X}, \boldsymbol{Z}) = \sigma^2(\boldsymbol{X}'\boldsymbol{M}_{\boldsymbol{Z}}\boldsymbol{X})^{-1} \tag{3.6.2}$$

になっている。$\boldsymbol{M}_{\boldsymbol{Z}} + \boldsymbol{P}_{\boldsymbol{Z}} = \boldsymbol{I}$ であることを用いると、

$$^\forall \boldsymbol{w} \neq \boldsymbol{0}, \quad \boldsymbol{w}'\boldsymbol{M}_{\boldsymbol{Z}}\boldsymbol{w} \leq \boldsymbol{w}'\boldsymbol{w} \quad \Longrightarrow \quad (\boldsymbol{w}'\boldsymbol{M}_{\boldsymbol{Z}}\boldsymbol{w})^{-1} \geq (\boldsymbol{w}'\boldsymbol{w})^{-1}$$

となるので、これを用いると

$$(\boldsymbol{X}'\boldsymbol{X} - \boldsymbol{X}'\boldsymbol{M}_{\boldsymbol{Z}}\boldsymbol{X})_{jj} = (\boldsymbol{X}'\boldsymbol{P}_{\boldsymbol{Z}}\boldsymbol{X})_{jj} = \boldsymbol{x}_j'\boldsymbol{P}_{\boldsymbol{Z}}\boldsymbol{x}_j \geq 0 \quad (j = 1, \ldots, k)$$

が得られる。（もし、$\boldsymbol{X}'\boldsymbol{Z} = \boldsymbol{O}$ であれば $\boldsymbol{X}'\boldsymbol{P}_{\boldsymbol{Z}}\boldsymbol{X} = \boldsymbol{O}$）

このことより、$\tilde{\boldsymbol{b}}$ は効率性を満たしていないことが分かる。

b) underspecification(説明変数過少) の場合

いま、DGP が $\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{Z\gamma} + \boldsymbol{\epsilon}$ であるとしよう。このとき、推定モデルが $\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{u}$ であるならば、OLS 推定量にどのような影響があるだろうか。

$$\boldsymbol{b} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y} = \boldsymbol{\beta} + (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{Z\gamma} + (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\epsilon}$$

となるので、

$$\begin{aligned} \mathrm{E}(\boldsymbol{b} \mid \boldsymbol{X}, \boldsymbol{Z}) &= \boldsymbol{\beta} + (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{Z\gamma} \\ \mathrm{MSE}\,(\boldsymbol{b}) &= \mathrm{E}\left[(\boldsymbol{b}-\boldsymbol{\beta})(\boldsymbol{b}-\boldsymbol{\beta})'\right] \\ &= \sigma^2(\boldsymbol{X}'\boldsymbol{X})^{-1} + (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{Z\gamma\gamma}'\boldsymbol{Z}'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1} \end{aligned}$$

が得られる。したがって、$\boldsymbol{X}'\boldsymbol{Z} = \boldsymbol{O}$ でない限り、$\boldsymbol{b}$ は偏りのある推定量であることが分かる。

3.7 多重共線性とその対策

いままで、$\boldsymbol{X}'\boldsymbol{X}$ に逆行列が存在すること、すなわち $rank(\boldsymbol{X}'\boldsymbol{X}) = k$ を仮定してきた。しかしながら、$\boldsymbol{X}$ を構成する列ベクトルが一次従属である場合には、$rank(\boldsymbol{X}'\boldsymbol{X}) < k$ となるため、逆行列が存在しない。このようなケースを、説明変数間に**多重共線性** (multi-collinearlity) が存在する、と呼ぶ。

説明変数間に完全な多重共線性が存在するケースもあるが、通常は説明変数間の相関がかなり高い場合に多重共線性に近い現象が起きる。

幾何的に見てゆくと、二つの説明変数ベクトル $\boldsymbol{x}_1, \boldsymbol{x}_2$ の相関が高いときには、

$$\mathrm{Cor}(x_1, x_2) = \frac{< \boldsymbol{x}_1, \boldsymbol{x}_2 >}{\|\boldsymbol{x}_1\|\|\boldsymbol{x}_2\|} = \cos\theta$$

この二つのベクトルのなす角 θ が 0 か π に近いことを意味する。すると、この二つのベクトルを含む線形空間が不安定であることが分かる。

したがって、多重共線性に近い現象が起きるとき、$(\boldsymbol{X}'\boldsymbol{X})^{-1}$ は大きくなるので、

(1) OLS 推定量の分散が大きくなる。
(2) (1) より回帰係数の検定において、帰無仮説を棄却しにくくなる。

ということが起きる。

対策

数学的な対策としては、

(a) リッジ回帰 (ridge regression)

(b) 主成分回帰

があるが、経済分析の分野では余り用いられない。