アクチュアリー ゼミ 第4回 補足

教科書 1.5 では. $V(\hat{\alpha})$, $V(\hat{\beta})$, $Cov(\hat{\alpha}, \hat{\beta})$ だけ求めてから. $(\hat{\alpha}, \hat{\beta})$ が $N\left(\begin{pmatrix}\alpha\\ \beta\end{pmatrix}, \frac{\sigma^2}{n}\begin{pmatrix}1 & \bar{x}\\ \bar{x} & \overline{x^2}\end{pmatrix}\right)$ に従うと結論づけているが.

疑問① $\hat{\alpha}$, $\hat{\beta}$ がそれぞれ正規分布に従うのか?

疑問② $\hat{\alpha}$. $\hat{\beta}$ が正規分布に従うからといって. $\hat{\alpha}$. $\hat{\beta}$ は明らかに独立でない (どちらも (x, y) のデータから推定していることから) のに. $(\hat{\alpha}, \hat{\beta})$ が多次元正規分布に従うのか?

等の疑問が生じるだろう. なので補足として §1. $(\hat{\alpha}, \hat{\beta})$ の平均と分散共分散行列の導出

§2. 線型変換に対する多次元の再生性 について論じようと思う.

§1.

$\begin{pmatrix}\hat{\alpha}\\ \hat{\beta}\end{pmatrix}$ の推定

そもそも (α, β) の推定には正規方程式が用いられており,

$$\begin{pmatrix}\hat{\alpha}\\ \hat{\beta}\end{pmatrix} = \begin{pmatrix}1 & \bar{x}\\ \bar{x} & \overline{x^2}\end{pmatrix}^{-1}\begin{pmatrix}\bar{y}\\ \overline{xy}\end{pmatrix}$$

と表されている. ここで説明変数は与えられた確定値という仮定から. $\begin{pmatrix}1 & \bar{x}\\ \bar{x} & \overline{x^2}\end{pmatrix}^{-1}$ は定行列.

しかし $\begin{pmatrix}\bar{y}\\ \overline{xy}\end{pmatrix}$ の値にばらつきが生じるため. $\begin{pmatrix}\hat{\alpha}\\ \hat{\beta}\end{pmatrix}$ がばらつくのであり. $\begin{pmatrix}\bar{y}\\ \overline{xy}\end{pmatrix}$ のばらつきについて詳しくみていこう.

$$\bar{y} = \frac{1}{n}\sum y_i = \frac{1}{n}\sum \beta x_i + \alpha + \varepsilon_i = \beta\bar{x} + \alpha + \frac{1}{n}\sum \varepsilon_i$$

$$\overline{xy} = \frac{1}{n}\sum x_i y_i = \frac{1}{n}\sum \beta x_i^2 + \alpha x_i + x_i\varepsilon_i = \beta\overline{x^2} + \alpha\bar{x} + \frac{1}{n}\sum x_i\varepsilon_i$$

なので $\begin{pmatrix}\bar{y}\\ \overline{xy}\end{pmatrix} = \begin{pmatrix}1 & \bar{x}\\ \bar{x} & \overline{x^2}\end{pmatrix}\begin{pmatrix}\alpha\\ \beta\end{pmatrix} + \frac{1}{n}\begin{pmatrix}\sum\varepsilon_i\\ \sum x_i\varepsilon_i\end{pmatrix}$

これを正規方程式に代入すると $\begin{pmatrix}\hat{\alpha}\\ \hat{\beta}\end{pmatrix} = \begin{pmatrix}\alpha\\ \beta\end{pmatrix} + \frac{1}{n}\begin{pmatrix}1 & \bar{x}\\ \bar{x} & \overline{x^2}\end{pmatrix}^{-1}\begin{pmatrix}\sum\varepsilon_i\\ \sum x_i\varepsilon_i\end{pmatrix}$ となっている. …①

$E\left[\begin{pmatrix}\hat{\alpha}\\ \hat{\beta}\end{pmatrix}\right]$ について

定行列なので外に出た

$$E\left[\begin{pmatrix}\hat{\alpha}\\ \hat{\beta}\end{pmatrix}\right] = E\left[\begin{pmatrix}\alpha\\ \beta\end{pmatrix}\right] + \frac{1}{n}\begin{pmatrix}1 & \bar{x}\\ \bar{x} & \overline{x^2}\end{pmatrix}^{-1} E\left[\begin{pmatrix}\sum\varepsilon_i\\ \sum x_i\varepsilon_i\end{pmatrix}\right] = E\left[\begin{pmatrix}\alpha\\ \beta\end{pmatrix}\right] = \begin{pmatrix}\alpha\\ \beta\end{pmatrix}$$ となっている

$E[\varepsilon_i] = 0$ を利用すると 0

つまり期待値が真の値になっている.

$\begin{pmatrix}\hat{\alpha}\\ \hat{\beta}\end{pmatrix}$ の分散 について

多変数の解析を行うとき. 分散共分散行列 (Σ) を求めることがよくあり.

ここでも $(\hat{\alpha}, \hat{\beta})$ に対して分散共分散行列を求めることを考えよう

$V[\varepsilon_i] = \sigma^2$, $Cov(\varepsilon_i, \varepsilon_j) = 0$ が使えればうれしいのだが. $(\hat{\alpha}, \hat{\beta})$ では.

$$\frac{1}{n}\begin{pmatrix}1 & \bar{x}\\ \bar{x} & \overline{x^2}\end{pmatrix}^{-1}\begin{pmatrix}\sum\varepsilon_i\\ \sum x_i\varepsilon_i\end{pmatrix} = \underbrace{\frac{1}{n}\begin{pmatrix}1 & \bar{x}\\ \bar{x} & \overline{x^2}\end{pmatrix}^{-1}\begin{pmatrix}1 & 1 & 1 & \cdots & 1\\ x_1 & x_2 & x_3 & \cdots & x_n\end{pmatrix}}_{\text{定行列による変換}}\begin{pmatrix}\varepsilon_1\\ \varepsilon_2\\ \vdots\\ \varepsilon_n\end{pmatrix}$$

というように線型変換されているため.

少々複雑な結果となりそうであり. 一旦線型変換時に分散共分散行列がどうなるかを考えてみよう.

線型変換時のΣの変化

$x_1 = a_{11}\varepsilon_1 + a_{12}\varepsilon_2 + \cdots + a_{1n}\varepsilon_n$

$x_2 = a_{21}\varepsilon_1 + a_{22}\varepsilon_2 + \cdots + a_{2n}\varepsilon_n$ (ただし $V(\varepsilon_i)=\sigma^2$, $Cov(\varepsilon_i, \varepsilon_j)=0$) のように線型変換したときに

(x_1, x_2) の分散共分散行列Σは

$$\Sigma = \begin{pmatrix} V(x_1) & Cov(x_1 x_2) \\ Cov(x_1 x_2) & V(x_2) \end{pmatrix}$$ で与えられ、それぞれの成分について考えていく。

$$V(x_1) = Cov(x_1, x_1) = Cov(a_{11}\varepsilon_1 + a_{12}\varepsilon_2 + \cdots + a_{1n}\varepsilon_n,\ a_{11}\varepsilon_1 + a_{12}\varepsilon_2 + \cdots + a_{1n}\varepsilon_n)$$

ここで $V(\varepsilon_i)=\sigma^2$, $Cov(\varepsilon_i, \varepsilon_j)=0$ を利用し $a_{11}^2\sigma^2 + a_{12}^2\sigma^2 + \cdots + a_{1n}^2\sigma^2$ と計算される。

同様にして $Cov(x_1, x_2) = a_{11}a_{21}\sigma^2 + a_{12}a_{22}\sigma^2 + \cdots + a_{1n}a_{2n}\sigma^2$

$V(x_2) = a_{21}^2\sigma^2 + a_{22}^2\sigma^2 + \cdots + a_{2n}^2\sigma^2$ であるから

$$\Sigma = \sigma^2 \begin{pmatrix} a_{11}^2 + a_{12}^2 + \cdots + a_{1n}^2 & a_{11}a_{21} + a_{12}a_{22} + \cdots + a_{1n}a_{2n} \\ a_{11}a_{21} + a_{12}a_{22} + \cdots + a_{1n}a_{2n} & a_{21}^2 + a_{22}^2 + \cdots + a_{2n}^2 \end{pmatrix} = \sigma^2 \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \end{pmatrix} \begin{pmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \\ \vdots & \vdots \\ a_{1n} & a_{2n} \end{pmatrix}$$ となる。

← 線型変換時の分散分母散行列!

実際に (α, β) のΣを求める。

これを用いると、求める分散母分散行列は

$(A^{-1})^T = (A^T)^{-1}$ を利用し、A^T が A に一致することから

$$\Sigma = \sigma^2 \frac{1}{n^2} \begin{pmatrix} 1 & \bar{x} \\ \bar{x} & \overline{x^2} \end{pmatrix}^{-1} \begin{pmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_n \end{pmatrix} \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix} \left\{ \begin{pmatrix} 1 & \bar{x} \\ \bar{x} & \overline{x^2} \end{pmatrix}^{-1} \right\}^T = \frac{\sigma^2}{n^2} \begin{pmatrix} 1 & \bar{x} \\ \bar{x} & \overline{x^2} \end{pmatrix}^{-1} \begin{pmatrix} n & x_1 + x_2 + \cdots + x_n \\ x_1 + x_2 + \cdots + x_n & x_1^2 + x_2^2 + \cdots + x_n^2 \end{pmatrix} \begin{pmatrix} 1 & \bar{x} \\ \bar{x} & \overline{x^2} \end{pmatrix}^{-1}$$

割る

$$= \frac{\sigma^2}{n} \begin{pmatrix} 1 & \bar{x} \\ \bar{x} & \overline{x^2} \end{pmatrix}^{-1} \begin{pmatrix} 1 & \bar{x} \\ \bar{x} & \overline{x^2} \end{pmatrix} \begin{pmatrix} 1 & \bar{x} \\ \bar{x} & \overline{x^2} \end{pmatrix}^{-1} = \frac{\sigma^2}{n} \begin{pmatrix} 1 & \bar{x} \\ \bar{x} & \overline{x^2} \end{pmatrix}^{-1}$$ となる

ゆえに $\begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \end{pmatrix}$ の平均は $\begin{pmatrix} \alpha \\ \beta \end{pmatrix}$ で分散共分散行列は $\frac{\sigma^2}{n} \begin{pmatrix} 1 & \bar{x} \\ \bar{x} & \overline{x^2} \end{pmatrix}^{-1}$ である。

§2.

$$
\begin{aligned}
x_1 &= a_{11}\varepsilon_1 + a_{12}\varepsilon_2 + \cdots + a_{1n}\varepsilon_n \\
x_2 &= a_{21}\varepsilon_1 + a_{22}\varepsilon_2 + \cdots + a_{2n}\varepsilon_n \\
&\ \vdots \\
x_n &= a_{n1}\varepsilon_1 + a_{n2}\varepsilon_2 + \cdots + a_{nn}\varepsilon_n
\end{aligned}
\quad \text{つまり} \quad
\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}
= \overset{A\ \text{とする}}{\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & & \vdots \\ \vdots & & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}}
\begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}
$$

という変数変換を考える.
ただし $|A| \neq 0$ とする.

このとき $f_{x_1 x_2 \cdots x_n}(x_1\, x_2 \cdots x_n) = f_{E_1 E_2 \cdots E_n}(\varepsilon_1\ \varepsilon_2 \cdots \varepsilon_n)\left|\dfrac{\partial(\varepsilon_1\,\varepsilon_2 \cdots \varepsilon_n)}{\partial(x_1\, x_2 \cdots x_n)}\right| = f_{E_1}(\varepsilon_1) f_{E_2}(\varepsilon_2) \cdots f_{E_n}(\varepsilon_n) \dfrac{1}{\|A\|}$

$$\therefore f(x_1 x_2 \cdots x_n) = \frac{1}{\|A\|}\left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n \exp\left(-\frac{\varepsilon_1^2 + \varepsilon_2^2 + \cdots + \varepsilon_n^2}{2\sigma^2}\right) = \frac{1}{\|A\|}\left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n \exp\left(-\frac{1}{2}(\varepsilon_1\,\varepsilon_2 \cdots \varepsilon_n)\begin{pmatrix}\sigma^2 & & \\ & \sigma^2 & \\ & & \ddots \ \sigma^2\end{pmatrix}^{-1}\begin{pmatrix}\varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_n\end{pmatrix}\right)$$

ここで $(\varepsilon_1 \cdots \varepsilon_n)\begin{pmatrix}\sigma^2 & & \\ & \ddots & \\ & & \sigma^2\end{pmatrix}^{-1}\begin{pmatrix}\varepsilon_1 \\ \vdots \\ \varepsilon_n\end{pmatrix} = (x_1 \cdots x_n)(A^T)^{-1}\begin{pmatrix}\sigma^2 & & \\ & \ddots & \\ & & \sigma^2\end{pmatrix}^{-1} A^{-1}\begin{pmatrix}x_1 \\ \vdots \\ x_n\end{pmatrix} = (x_1 \cdots x_n)\Bigg\{\underbrace{A\begin{pmatrix}\sigma^2 & & \\ & \sigma^2 & \\ & & \ddots \ \sigma^2\end{pmatrix}A^T}_{=\Sigma\ \text{とおく}}\Bigg\}^{-1}\begin{pmatrix}x_1 \\ \vdots \\ x_n\end{pmatrix}$

$$\therefore f(x_1 \cdots x_n) = \frac{1}{\|A\|}\left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n \exp\left(-\frac{1}{2}(x_1 \cdots x_n)\Sigma^{-1}\begin{pmatrix}x_1 \\ \vdots \\ x_n\end{pmatrix}\right)$$

ゆえに $\begin{pmatrix}x_1 \\ \vdots \\ x_n\end{pmatrix} \sim N\left(0\ ,\ A\begin{pmatrix}\sigma^2 & & \\ & \sigma^2 & \\ & & \ddots \ \sigma^2\end{pmatrix}A^T\right)$ となる.

ここから,(x_1, x_2) が多次元正規分布に従うことが納得できるだろう.

追記

§2 の議論を使えば 線型変換A によって変換された値の Σ は 元のΣ を $A\Sigma A^T$ としたものであることが言える気がする. 弱点は正規分布を仮定しているところだけど…