Estimable Functions and Contrasts

Peter von Rohr

2022-03-21

Models Not Of Full Rank

- Model

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{e}$$

- Least squares normal equations

$$\mathbf{X}^T\mathbf{X}\mathbf{b}^{(0)} = \mathbf{X}^T\mathbf{y}$$

Solutions

- matrix $\mathbf{X}$ not of full rank
- $\mathbf{X}^T\mathbf{X}$ cannot be inverted
- solution

$$\mathbf{b}^{(0)} = (\mathbf{X}^T\mathbf{X})^{-}\mathbf{X}^T\mathbf{y}$$

where $(\mathbf{X}^T\mathbf{X})^{-}$ stands for a **generalized inverse**

Generalized Inverse

- matrix $\mathbf{G}$ is a generalized inverse of matrix $\mathbf{A}$, if

$$\mathbf{AGA} = \mathbf{A}$$

$$(\mathbf{AGA})^T = \mathbf{A}^T$$

- Use `MASS::ginv()` in R

Systems of Equations

- For a consistent system of equations

$$Ax = y$$

- Solution

$$x = Gy$$

if G is a generalized inverse of A.

$$x = Gy$$
$$Ax = AGy$$
$$Ax = AGAx$$

Non Uniqueness

- Solution $x = Gy$ is not unique

$$\tilde{\mathbf{x}} = \mathbf{G}\mathbf{y} + (\mathbf{G}\mathbf{A} - \mathbf{I})\mathbf{z}$$

yields a different solution for an arbitrary vector $\mathbf{z}$

$$\mathbf{A}\tilde{\mathbf{x}} = \mathbf{A}\mathbf{G}\mathbf{y} + (\mathbf{A}\mathbf{G}\mathbf{A} - \mathbf{A})\mathbf{z}$$

Least Squares Normal Equations

- Instead of $Ax = y$, we have

$$\mathbf{X}^T\mathbf{X}\mathbf{b}^{(0)} = \mathbf{X}^T\mathbf{y}$$

- With generalized inverse $\mathbf{G}$ of $\mathbf{X}^T\mathbf{X}$

$$\mathbf{b}^{(0)} = \mathbf{G}\mathbf{X}^T\mathbf{y}$$

is a solution to the least squares normal equations

Parameter Estimator

But $\mathbf{b}^{(0)}$ is not an estimator for the parameter $\mathbf{b}$, because

- it is not unique
- Expectation $E(\mathbf{b}^{(0)}) = E(\mathbf{G}\mathbf{X}^T\mathbf{y}) = \mathbf{G}\mathbf{X}^T\mathbf{X}\mathbf{b} \neq \mathbf{b}$

Estimable Functions

Animal	Breed	Observation
1	Angus	16
2	Angus	10
3	Angus	19
4	Simmental	11
5	Simmental	13
6	Limousin	27

Model

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{e}$$

$$\mathbf{y} = \begin{bmatrix} 16 \\ 10 \\ 19 \\ 11 \\ 13 \\ 27 \end{bmatrix}, \mathbf{X} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix} \text{ and } \mathbf{b} = \begin{bmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix}$$

Normal Equations

$$X^T X b^0 = X^T y$$

$$\begin{bmatrix} 6 & 3 & 2 & 1 \\ 3 & 3 & 0 & 0 \\ 2 & 0 & 2 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mu^0 \\ \alpha_1^0 \\ \alpha_2^0 \\ \alpha_3^0 \end{bmatrix} = \begin{bmatrix} 96 \\ 45 \\ 24 \\ 27 \end{bmatrix}$$

Solutions to Normal Equations

Elements of Solution	b_1^0	b_2^0	b_3^0	b_4^0
μ^0	16	14	27	-2982
α_1^0	-1	1	-12	2997
α_2^0	-4	-2	-15	2994
α_3^0	11	13	0	3009

Functions of Solutions

Linear Function	b_1^0	b_2^0	b_3^0	b_4^0
$\alpha_1^0 - \alpha_2^0$	3.0	3.0	3.0	3.0
$\mu^0 + \alpha_1^0$	15.0	15.0	15.0	15.0
$\mu^0 + 1/2(\alpha_2^0 + \alpha_3^0)$	19.5	19.5	19.5	19.5

- $\alpha_1^0 - \alpha_2^0$: estimate of the difference between breed effects for Angus and Simmental
- $\mu^0 + \alpha_1^0$: estimate of the general mean plus the breed effect of Angus
- $\mu^0 + 1/2(\alpha_2^0 + \alpha_3^0)$: estimate of the general mean plus mean effect of breeds Simmental and Limousin

Definition of Estimable Functions

$$\mathbf{q}^T \mathbf{b} = \mathbf{t}^T E(\mathbf{y})$$

- Example

$$E(y_{1j}) = \mu + \alpha_1$$

with $\mathbf{t}^T = \left[\begin{array}{cccccc} 1 & 1 & 1 & 0 & 0 & 0 \end{array} \right]$ and $\mathbf{q}^T = \left[\begin{array}{cccc} 1 & 1 & 0 & 0 \end{array} \right]$

- Properties

$$\mathbf{q}^t = \mathbf{t}^T \mathbf{X}$$

- Test

$$\mathbf{q}^T \mathbf{H} = \mathbf{q}^T$$

with $\mathbf{H} = \mathbf{G}\mathbf{X}^T\mathbf{X}$