Theorem 5.1. *Under the Gauss-Markov model, if $\boldsymbol{\lambda}^\intercal \mathbf{b}$ is (linearly) estimable, then the least squares estimator $\boldsymbol{\lambda}^\intercal \hat{\mathbf{b}}$ is the best (i.e., minimum variance) linear unbiased estimator (BLUE) of $\boldsymbol{\lambda}^\intercal \mathbf{b}$.*

LSE is BLUE

在G-M模型中，若$\lambda^T b$可估，则其

LSE 和 BLUE 等价且唯一。

最小误差所得的估计量具有最小方差。

$y = Xb + e \Rightarrow$ 对$\lambda^T b$估计，若$a^T y$无偏，则可对 a 作正交分解 $a = \underbrace{P_x a}_{a_0 \in C(X)} + \underbrace{(I-P_x)a}_{a_1 \in [C(X)]^\perp}$

$(\lambda = X^T a)$

$Var(a^T y) = \sigma^2 (a_0^T a_0 + a_1^T a_1)$, $a_1^T a_1 \to 0$

距离和方差等价

$a_0 = \underbrace{X(X^TX)^g}_{(X^T)^g} \lambda = \underbrace{X(X^TX)^g X^T}_{P_x} a$

7. (1) Note 5 Section 2 Example 2中，广义逆的不同选择$(\mathbf{X}\mathbf{X})^-$是否会影响最终的方差计算结果？ ~~(2) 如何理解Theorem 5.1? (3) 为何Theorem 5.3 中，$\hat{\sigma}^2 = \mathrm{SSE}/(N-r)$的分母是$N-r$? (曹家豪)~~

Example 2

Consider a Gauss-Markov model. Let $\boldsymbol{\lambda}^\intercal \mathbf{b}$ be an estimable function. By Theorem 4.7, the least squares estimator $\boldsymbol{\lambda}^\intercal \hat{\mathbf{b}}$ of $\boldsymbol{\lambda}^\intercal \mathbf{b}$ is unique, and invariant of the choice of $\hat{\mathbf{b}} = (\mathbf{X}^\intercal\mathbf{X})^-\mathbf{X}^\intercal\mathbf{y} + [\mathbf{I} - (\mathbf{X}^\intercal\mathbf{X})^-(\mathbf{X}^\intercal\mathbf{X})]\mathbf{z}$ (and the choice of generalized inverse $(\mathbf{X}^\intercal\mathbf{X})^-$). Therefore, we can represent $\boldsymbol{\lambda}^\intercal \hat{\mathbf{b}}$ as $\boldsymbol{\lambda}^\intercal(\mathbf{X}^\intercal\mathbf{X})^-\mathbf{X}^\intercal\mathbf{y}$.

$$\begin{aligned}
\mathsf{Var}(\boldsymbol{\lambda}^\intercal \hat{\mathbf{b}}) = \mathsf{Var}(\boldsymbol{\lambda}^\intercal(\mathbf{X}^\intercal\mathbf{X})^-\mathbf{X}^\intercal\mathbf{y}) &= \sigma^2 \boldsymbol{\lambda}^\intercal(\mathbf{X}^\intercal\mathbf{X})^-\mathbf{X}^\intercal\mathbf{X}[(\mathbf{X}^\intercal\mathbf{X})^-]^\intercal \boldsymbol{\lambda} \\
&= \sigma^2 \boldsymbol{\lambda}^\intercal(\mathbf{X}^\intercal\mathbf{X})^-[\boldsymbol{\lambda}^\intercal(\mathbf{X}^\intercal\mathbf{X})^-\mathbf{X}^\intercal\mathbf{X}]^\intercal \\
&= \sigma^2 \boldsymbol{\lambda}^\intercal(\mathbf{X}^\intercal\mathbf{X})^- \boldsymbol{\lambda}
\end{aligned}$$

due to that $(\mathbf{X}^\intercal\mathbf{X})^-\mathbf{X}^\intercal$ is a generalized inverse of $\mathbf{X}$ (Lemma 3.4) and Theorem 4.5(iii).

(1) 广义逆的选择不会影响方差的计算。

例2中：$Var(\lambda^T \hat{b}) = Var(\lambda^T (X^TX)^- X^T y) = \sigma^2 \lambda^T (X^TX)^- \lambda$

两种解释方法

- i). 广义逆的选取不影响 $\hat{b}$ 的值，因此不影响 $Var(\lambda^T \hat{b})$
- ii) 由于 $\lambda^T b$ 可估，则可令 $\lambda = X^T a$，提 $\lambda^T (X^TX)^- \lambda = a^T X (X^TX)^- X^T a = a^T P_x a$
  P_x 与广义逆选取无关。