CONSEQUENCE OF THE GAUSS-MARKOV THEOREM

(sometimes it is called G-M thm)

An easy consequence of the Gauss-Markov theorem is the following.

Proposition. *If $r = p$, then for any $\mathbf{b} \in \mathbb{R}^p$, the statistic $\mathbf{b}^T \hat{\mathbf{a}}$ is a linear unbiased estimate of the univariate parameter function $\mathbf{b}^T \mathbf{a}$, and it has minimum variance among all of such (linear, unbiased) estimates of $\mathbf{b}^T \mathbf{a}$ (BLUE).*

Proof. Unbiasedness is obvious; for any $\mathbf{b} \in \mathbb{R}^p$:

$$\mathrm{Var}\,(\mathbf{b}^T \hat{\mathbf{a}}) = \mathbf{b}^T \mathrm{Var}\,(\hat{\mathbf{a}}) \mathbf{b} \quad \text{and} \quad \mathrm{Var}\,(\mathbf{b}^T \tilde{\mathbf{a}}) = \mathbf{b}^T \mathrm{Var}\,(\tilde{\mathbf{a}}) \mathbf{b}.$$

In view of the Gauss–Markov theorem, $\mathrm{Var}\,(\tilde{\mathbf{a}}) - \mathrm{Var}\,(\hat{\mathbf{a}})$ is a positive semidefinite matrix, which means that for any vector $\mathbf{b} \in \mathbb{R}^p$:

$$\mathrm{Var}\,(\mathbf{b}^T \tilde{\mathbf{a}}) - \mathrm{Var}\,(\mathbf{b}^T \hat{\mathbf{a}}) \geq 0.$$

□

Definition. The parameter function $\mathbf{b}^T \mathbf{a}$ (with $\mathbf{b} \in \mathbb{R}^p$) is said to be *estimable* (in a linear and unbiased way) if there exists a vector $\mathbf{c} \in \mathbb{R}^n$ such that $\mathbf{E}(\mathbf{c}^T \mathbf{Y}) = \mathbf{b}^T \mathbf{a}$.

Proposition. *The parameter function $\mathbf{b}^T \mathbf{a}$ is estimable if and only if $\mathbf{b}$ is within the linear subspace of $\mathbb{R}^p$ spanned by the row vectors of $\mathbf{X}$.*

Proof. The following are equivalent steps:

$$\begin{aligned}
\mathbf{c}^T \mathbf{E}(\mathbf{Y}) &= \mathbf{b}^T \mathbf{a} \qquad (\forall\, \mathbf{a} \in \mathbb{R}^p)\,, \\
\mathbf{c}^T \mathbf{X} \mathbf{a} &= \mathbf{b}^T \mathbf{a} \qquad (\forall\, \mathbf{a} \in \mathbb{R}^p)\,, \\
\mathbf{c}^T \mathbf{X} &= \mathbf{b}^T\,, \\
\mathbf{b} &= \mathbf{X}^T \mathbf{c},
\end{aligned}$$

which means that the vector $\mathbf{b}$ is within the subspace spanned by the column vectors of $\mathbf{X}^T$, which is the same as the subspace spanned by the row vectors of $\mathbf{X}$. □

If $r = p$, this holds for any $\mathbf{b} \in \mathbb{R}^p$; therefore, any parameter function $\mathbf{b}^T \mathbf{a}$ is estimable. However, if $r < p$, then the BLUE estimate of the first proposition can be extended only for an estimable $\mathbf{b}^T \mathbf{a}$.