

Lecture 13: Aitken's theorem: the direct proof; GLS versus OLS

Econometrics ■ Unit I: Introduction and generalized least squares

Kapil Kumar

Department of Statistics
Chaudhary Charan Singh University, Meerut

B.Sc. (Hons) Statistics, Semester V
Academic Session 2026–2027

Previous lecture

Lecture 12: Aitken's theorem: statement and short proof.

Today

- Gauss–Markov theorem for GLSE (Aitken's theorem)
 - Proof II (long): a direct proof

Proof.

Let $\tilde{\beta} = Ay + c$ be any linear estimator, where A is a $k \times n$ matrix and c is a $k \times 1$ vector, both free of y (they may depend on X and Ω).

Step 1 (condition for unbiasedness). Since $E(y) = X\beta$,

$$E(\tilde{\beta}) = AE(y) + c = AX\beta + c.$$

This must equal β for every β .

- Putting $\beta = 0$ gives $c = 0$.
- Then $AX\beta = \beta$ for every β . Taking $\beta = e_j$, the j th unit vector, gives $AXe_j = e_j$, that is, the j th column of AX is the j th column of I_k . Hence $AX = I_k$.

So $\tilde{\beta}$ is unbiased if and only if $c = 0$ and $AX = I_k$.

Proof (cont.)

Step 2 (sampling error). With $c = 0$ and $AX = I_k$,

$$\tilde{\beta} = A(X\beta + \epsilon) = AX\beta + A\epsilon = \beta + A\epsilon, \quad \text{so} \quad \tilde{\beta} - \beta = A\epsilon.$$

Step 3 (the GLS matrix). Write $\hat{\beta}_G = A_G y$ with $A_G = (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}$. It satisfies

$$A_G X = (X'\Omega^{-1}X)^{-1}(X'\Omega^{-1}X) = I_k,$$

so GLS is itself a linear unbiased estimator with $c = 0$.

Proof (cont.)

Step 4 (difference from GLS). Define $D = A - A_G$, so that $A = A_G + D$. Then

$$DX = AX - A_GX = I_k - I_k = 0, \quad \text{and so} \quad X'D' = (DX)' = 0.$$

Step 5 (variance of the competitor). From Step 2 and Rule P4,

$$\text{Var}(\tilde{\beta}) = A \text{Var}(\varepsilon) A' = \sigma^2 A \Omega A' = \sigma^2 (A_G + D) \Omega (A_G + D)'.$$

Multiplying out,

$$\text{Var}(\tilde{\beta}) = \sigma^2 \{A_G \Omega A'_G + A_G \Omega D' + D \Omega A'_G + D \Omega D'\}.$$

Proof (cont.)

Step 6 (first term).

$$\begin{aligned}A_G \Omega A'_G &= (X' \Omega^{-1} X)^{-1} X' \Omega^{-1} \Omega \Omega^{-1} X (X' \Omega^{-1} X)^{-1} \\&= (X' \Omega^{-1} X)^{-1} X' \Omega^{-1} X (X' \Omega^{-1} X)^{-1} \\&= (X' \Omega^{-1} X)^{-1}.\end{aligned}$$

So $\sigma^2 A_G \Omega A'_G = \text{Var}(\hat{\beta}_G)$. (Here we used $(X' \Omega^{-1} X)^{-1}$ symmetric, so $A'_G = \Omega^{-1} X (X' \Omega^{-1} X)^{-1}$.)

Proof (cont.)

Step 7 (cross terms vanish).

$$\begin{aligned}A_G \Omega D' &= (X' \Omega^{-1} X)^{-1} X' \Omega^{-1} \Omega D' \\&= (X' \Omega^{-1} X)^{-1} X' D' \\&= (X' \Omega^{-1} X)^{-1} \cdot 0 = 0 \quad (\text{Step 4}).\end{aligned}$$

The other cross term is its transpose: $D \Omega A'_G = (A_G \Omega D')' = 0$.

Step 8 (the key identity). Substituting Steps 6 and 7 into Step 5,

$$\boxed{\text{Var}(\tilde{\beta}) = \text{Var}(\hat{\beta}_G) + \sigma^2 D \Omega D'}.$$

Proof (cont.)

Step 9 (the extra term is positive semidefinite). Since $T\Omega T' = I$ (Section 2.3 of the notes),

$$\Omega = T^{-1}(T')^{-1} = KK' \quad \text{with } K = T^{-1} \text{ nonsingular.}$$

For any $k \times 1$ vector ℓ ,

$$\begin{aligned} \ell'(D\Omega D')\ell &= \ell'DKK'D'\ell \\ &= (K'D'\ell)'(K'D'\ell) = \sum_i (K'D'\ell)_i^2 \geq 0. \end{aligned}$$

So $D\Omega D'$ is positive semidefinite.

Proof (cont.)

Step 10 (conclusion for every linear combination). For any ℓ , by Step 8 and Rule P4,

$$\begin{aligned}\text{Var}(\ell' \tilde{\beta}) - \text{Var}(\ell' \hat{\beta}_G) &= \ell' \{ \text{Var}(\tilde{\beta}) - \text{Var}(\hat{\beta}_G) \} \ell \\ &= \sigma^2 \ell' D \Omega D' \ell \geq 0.\end{aligned}$$

Taking $\ell = e_j$ gives $\text{Var}(\tilde{\beta}_j) \geq \text{Var}(\hat{\beta}_{G,j})$ for each coefficient j .

Step 11 (when equality holds). Equality for every ℓ means $K' D' \ell = 0$ for every ℓ , that is, $K' D' = 0$. Since K is nonsingular, $D' = 0$, so $D = 0$ and $A = A_G$. Hence GLS is the *unique* best linear unbiased estimator. $\square$

Result 2.3

OLS is linear and unbiased in model (2.1), so by Aitken's theorem $\text{Var}(b) - \text{Var}(\hat{\beta}_G)$ is positive semidefinite. GLS is at least as efficient as OLS. When $\Omega = I$, $A_G = (X'X)^{-1}X'$ and Proof II becomes a proof of the ordinary Gauss–Markov theorem for OLS.

Exam tip

For “State and prove the Gauss–Markov theorem for GLSE”, write Proof II in full. Always state the unbiasedness condition $AX = I$ and show clearly why the cross terms vanish. Proof I is enough for a short note.

Practice

Re-work the derivations and examples of this lecture without looking at the slides.

Next lecture

Lecture 14: A worked example of GLS.