Variance of Coefficients in Linear Models: Exact vs. Simplified Calculation

September 4, 2025

1 General Formula

Consider the linear model:

$$y = X\beta + \varepsilon, \quad \text{with} \quad \mathrm{Var}(\varepsilon) = \Sigma.$$

The least squares estimator is:

$$\hat{\beta} = (X^\top X)^{-1} X^\top y.$$

The exact variance of $\hat{\beta}$ is:

$$\mathrm{Var}(\hat{\beta}) = (X^\top X)^{-1} X^\top \Sigma X (X^\top X)^{-1}.$$

If $\Sigma = \sigma^2 I$ (i.i.d. errors), this simplifies to:

$$\mathrm{Var}(\hat{\beta}) = \sigma^2 (X^\top X)^{-1}.$$

2 Design Matrix (Hadamard, 4 Runs)

With 4 runs and 3 factors plus intercept, the design matrix is:

$$X = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}.$$

We have $X^\top X = 4I_4$, so:

$$\mathrm{Var}(\hat{\beta}) = \frac{\sigma^2}{4} I_4 \quad \text{(i.i.d. errors)}.$$

3 Example: Heteroscedastic Errors

Let $\Sigma = \mathrm{diag}(1, 4, 1, 4)$. Then:

$$\mathrm{Var}(\hat{\beta}) = \begin{bmatrix} 0.625 & 0 & -0.375 & 0 \\ 0 & 0.625 & 0 & -0.375 \\ -0.375 & 0 & 0.625 & 0 \\ 0 & -0.375 & 0 & 0.625 \end{bmatrix}.$$

Observations:

- Variances increase to 0.625.
- Covariances between parameters are nonzero.
- Using $\sigma^2(X^\top X)^{-1}$ would ignore these covariances.

4 Example: Correlated Errors

Let Σ have correlation $\rho = 0.5$ between all errors:

$$\Sigma_{ij} = \begin{cases} 1 & i = j, \\ 0.5 & i \neq j. \end{cases}$$

Then:

$$\mathrm{Var}(\hat{\beta}) = \begin{bmatrix} 0.625 & 0 & 0 & 0 \\ 0 & 0.125 & 0 & 0 \\ 0 & 0 & 0.125 & 0 \\ 0 & 0 & 0 & 0.125 \end{bmatrix}.$$

Intercept variance increases, factor variances decrease. The simplification $\sigma^2(X^\top X)^{-1}$ would incorrectly give 0.25 for all.

5 Key Takeaways

1. Exact formula: $\mathrm{Var}(\hat{\beta}) = (X^\top X)^{-1} X^\top \Sigma X (X^\top X)^{-1}$.
2. Simplified formula holds only if $\Sigma = \sigma^2 I$ (i.i.d. errors).
3. Heteroscedasticity introduces parameter correlations and changes variances.
4. Correlated errors can drastically change variance estimates.
5. Always prefer the exact formula if Σ is not diagonal or constant.

This demonstrates that assuming $\Sigma = \sigma^2 I$ is risky: it hides dependencies and can under- or overestimate parameter uncertainty.