

Lecture 1: Preliminaries: notation; differentiation with respect to a vector; expectations and variances

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

Unit I: Introduction and generalized least squares

Lecture 1 begins Unit I: Introduction and generalized least squares.

Today

- Notation
- Rules for differentiating with respect to a vector
- Rules for expectations and variances

- These results are *not* new syllabus topics.
- They are tools from earlier courses that every derivation in this course uses.
- Read this part once and come back to it whenever a step says “by Rule P1” and so on.

In the linear model

$$y = X\beta + \varepsilon,$$

- y is the $n \times 1$ vector of observations on the study (dependent) variable;
- X is the $n \times k$ matrix of observations on the explanatory variables;
- β is the $k \times 1$ vector of unknown coefficients;
- ε is the $n \times 1$ vector of disturbances.

If the model has an intercept, the first column of X is a column of ones and k counts the intercept. A prime ($'$) means transpose.

Derivative with respect to a vector

Let $b = (b_1, \dots, b_k)'$. The derivative of a scalar function $f(b)$ with respect to b is the column of partial derivatives

$$(\partial f / \partial b_1, \dots, \partial f / \partial b_k)'.$$

Rule P1

If a is a constant vector, then $\frac{\partial(a'b)}{\partial b} = a$.

Reason. $a'b = a_1b_1 + \dots + a_kb_k$, so $\partial(a'b)/\partial b_j = a_j$ for every j .

Rule P2

If A is a constant symmetric matrix, then $\frac{\partial(b'Ab)}{\partial b} = 2Ab$.

Check for $k = 2$. With $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix}$,

$$b'Ab = a_{11}b_1^2 + 2a_{12}b_1b_2 + a_{22}b_2^2.$$

Differentiating,

$$\frac{\partial(b'Ab)}{\partial b_1} = 2a_{11}b_1 + 2a_{12}b_2, \quad \frac{\partial(b'Ab)}{\partial b_2} = 2a_{12}b_1 + 2a_{22}b_2.$$

Stacking these two entries gives exactly $2Ab$. The same calculation works for any k .

Rule P3 (minimum)

If the first derivative is zero at b_0 and the matrix of second derivatives is positive definite, then b_0 gives the minimum. For $f(b) = c - 2b'd + b'Ab$ the matrix of second derivatives is $2A$.

For a random vector v with mean $E(v)$, its variance (covariance) matrix is

$$\text{Var}(v) = E\{(v - E v)(v - E v)'\}.$$

Its i th diagonal entry is $\text{Var}(v_i)$, and its (i, j) entry is $\text{Cov}(v_i, v_j)$.

Rule P4

If A is a constant matrix and c a constant vector, then

$$E(Av + c) = AE(v) + c, \quad \text{Var}(Av + c) = A \text{Var}(v) A'.$$

Rule P5 (trace)

The trace $\text{tr}(A)$ is the sum of the diagonal entries. For matrices of suitable orders, $\text{tr}(AB) = \text{tr}(BA)$. For a scalar s , $s = \text{tr}(s)$.

Rule P6

If $E(\epsilon) = 0$ and $\text{Var}(\epsilon) = \Sigma$, then for any constant matrix A ,

$$E(\epsilon' A \epsilon) = \text{tr}(A \Sigma).$$

Reason. $\epsilon' A \epsilon$ is a scalar, so $\epsilon' A \epsilon = \text{tr}(\epsilon' A \epsilon) = \text{tr}(A \epsilon \epsilon')$ by Rule P5. Taking expectations,

$$E\{\text{tr}(A \epsilon \epsilon')\} = \text{tr}\{A E(\epsilon \epsilon')\} = \text{tr}(A \Sigma).$$

Practice

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

Next lecture

Lecture 2: Positive definite matrices and eigenvalues; the OLS criterion and normal equations.