

ANIMAL BREEDING NOTES

CHAPTER 5

DIFFERENTIAL CALCULUS IN MATRIX NOTATION

Vectors of partial derivatives

Let x be an $m \times 1$ vector and y be a **scalar function** of the elements of x , i.e.,

$$y = f(x_1, \dots, x_m).$$

Then, $\partial y / \partial x$ is defined to be the **vector of partial derivatives** $\partial y / \partial x_i$, i.e.,

$$\frac{\partial y}{\partial x} = \begin{bmatrix} \frac{\partial y}{\partial x_1} \\ \frac{\partial y}{\partial x_2} \\ \vdots \\ \frac{\partial y}{\partial x_m} \end{bmatrix}$$

Example: $y = 5x_1 + x_2 - 3x_3$

Notice that

$$y = [5 \ 1 \ -3] \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} 5 \\ 1 \\ -3 \end{bmatrix}$$

Thus,

$$\frac{\partial y}{\partial x} = \begin{bmatrix} \frac{\partial y}{\partial x_1} \\ \frac{\partial y}{\partial x_2} \\ \frac{\partial y}{\partial x_3} \end{bmatrix} = \begin{bmatrix} \frac{\partial}{\partial x_1} \\ \frac{\partial}{\partial x_2} \\ \frac{\partial}{\partial x_3} \end{bmatrix} y = \begin{bmatrix} 5 \\ 1 \\ -3 \end{bmatrix}.$$

In general, for $y = a'x$

$$\frac{\partial y}{\partial x} = \frac{\partial}{\partial x} a'x = \frac{\partial}{\partial x} x'a = a$$

Matrices of partial derivatives

(1) Let x be an $m \times 1$ vector and y be a **vector a scalar functions** of the elements of x , i.e., $\{y_i = f_i(x_1, \dots, x_n)\}$. Then, $\partial y / \partial x$ is defined to be the **matrix of partial derivatives** $\partial y / \partial x_i$, i.e.,

$$\frac{\partial y}{\partial x} = \left[\frac{\partial y_1}{\partial x} \frac{\partial y_2}{\partial x} \dots \frac{\partial y_n}{\partial x} \right]$$

Let $y_i = a_i'x$. Thus, y can be written as

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} a_1'x \\ a_2'x \\ \vdots \\ a_n'x \end{bmatrix} = Ax$$

where the $\{a_i'\}$ are the **rows** of A .

Thus,

$$\frac{\partial y}{\partial x} = \partial Ax = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} = A'$$

On the other hand, if $y = x'A$, i.e., y is a function of the columns of A ,

$$\frac{\partial y}{\partial x} = \frac{\partial x'A}{\partial x} = [c_1 \ c_2 \ \dots \ c_n] = A$$

where the $\{c_i\}$ are the **columns** of A .

Example:

$$(a) \quad y = Ax = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} x_1 + 2x_2 + 3x_3 \\ 4x_1 + 5x_2 + 6x_3 \\ 7x_1 + 8x_2 + 9x_3 \end{bmatrix}$$

$$\frac{\partial y}{\partial x} = \left[\frac{\partial y_1}{\partial x} \frac{\partial y_2}{\partial x} \frac{\partial y_3}{\partial x} \right] = \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_2}{\partial x_1} & \frac{\partial y_3}{\partial x_1} \\ \frac{\partial y_1}{\partial x_2} & \frac{\partial y_2}{\partial x_2} & \frac{\partial y_3}{\partial x_2} \\ \frac{\partial y_1}{\partial x_3} & \frac{\partial y_2}{\partial x_3} & \frac{\partial y_3}{\partial x_3} \end{bmatrix} = \begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{bmatrix} = A'$$

$$(b) \quad y = x'A = \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} = \begin{bmatrix} x_1 + 4x_2 + 7x_3 \\ 2x_1 + 5x_2 + 8x_3 \\ 3x_1 + 6x_2 + 9x_3 \end{bmatrix}$$

$$\frac{\partial y}{\partial x} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} = A$$

(2) Let $y = x'Ax$. Then, $\frac{\partial y}{\partial x} = Ax + A'x$, and if A is **symmetric**, $\frac{\partial y}{\partial x} = 2Ax$.

Proof:

$$y = x'Ax = \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j$$

$$\frac{\partial y}{\partial x} = \begin{bmatrix} \frac{\partial}{\partial x_1} \\ \vdots \\ \frac{\partial}{\partial x_n} \end{bmatrix} \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j = \begin{bmatrix} \sum_{j=1}^n a_{1j} x_j + \sum_{i=1}^n a_{i1} x_i \\ \vdots \\ \sum_{j=1}^n a_{nj} x_j + \sum_{i=1}^n a_{in} x_i \end{bmatrix}$$

$$\frac{\partial y}{\partial x} = \begin{bmatrix} a_1' x \\ a_2' x \\ \vdots \\ a_n' x \end{bmatrix} + \begin{bmatrix} c_1' x \\ c_2' x \\ \vdots \\ c_n' x \end{bmatrix} = Ax + A'x$$

where $\{a_j\}$ are the rows of A and $\{c_j\}$ are the columns of A .

Thus, if A is symmetric,

$$\frac{\partial y}{\partial x} = Ax + A'x = 2Ax$$

Example:

$$y = x'Ax = [x_1 \ x_2 \ x_3] \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$$y = [x_1 + 2x_2 + 3x_3 : 2x_1 + 4x_2 + 5x_3 : 3x_1 + 5x_2 + 6x_3] \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$$y = x_1^2 + 4x_1x_2 + 6x_1x_3 + 4x_2^2 + 10x_2x_3 + 6x_3^2$$

$$\frac{\partial \mathbf{y}}{\partial \mathbf{x}} = \begin{bmatrix} \frac{\partial \mathbf{y}}{\partial x_1} \\ \frac{\partial \mathbf{y}}{\partial x_2} \\ \frac{\partial \mathbf{y}}{\partial x_3} \end{bmatrix} = \begin{bmatrix} 2x_1 + 4x_2 + 6x_3 \\ 4x_1 + 8x_2 + 10x_3 \\ 6x_1 + 10x_2 + 12x_3 \end{bmatrix} = 2A\mathbf{x}$$

(3) Let t be a **scalar** and \mathbf{Y} a **matrix**, whose elements are functions of the scalar t . Then,

$$\frac{\partial \mathbf{y}}{\partial t} = \begin{bmatrix} \frac{\partial y_{11}}{\partial t} & \dots & \frac{\partial y_{1m}}{\partial t} \\ \vdots & \vdots & \vdots \\ \frac{\partial y_{n1}}{\partial t} & \dots & \frac{\partial y_{nm}}{\partial t} \end{bmatrix}$$

Example:

$$\mathbf{y} = \begin{bmatrix} 2t + t^2 & 3t \\ 3t & 4t + 2t^2 \end{bmatrix}$$

$$\frac{\partial \mathbf{y}}{\partial t} = \begin{bmatrix} 2+2t & 3 \\ 3 & 4+4t \end{bmatrix}$$

(4) If A and B are functions of the **scalar** t , and if $C = AB$, then

$$c_{ij} = \sum_{k=1}^n a_{ik} b_{kj}$$

and

$$\frac{\partial c_{ij}}{\partial t} = \sum_{k=1}^n \left[a_{ik} \frac{\partial b_{ki}}{\partial t} + \frac{\partial a_{ik}}{\partial t} b_{kj} \right]$$

Thus, by applying (3) above

$$\frac{\partial AB}{\partial t} = A \frac{\partial B}{\partial t} + \frac{\partial A}{\partial t} B$$

Example:

Let $A = \begin{bmatrix} 3t & 2t^2 \\ 2t^2 & 4t \end{bmatrix}$ and $B = \begin{bmatrix} 5t^2 & t \\ t & 6t \end{bmatrix}$

Then,

$$\begin{aligned} \frac{\partial AB}{\partial t} &= \begin{bmatrix} 3t & 2t^2 \\ 2t^2 & 4t \end{bmatrix} \begin{bmatrix} 10t & 1 \\ 1 & 6 \end{bmatrix} + \begin{bmatrix} 3 & 4t \\ 4t & 4 \end{bmatrix} \begin{bmatrix} 5t^2 & t \\ t & 6t \end{bmatrix} \\ &= \begin{bmatrix} 32t^2 & 3t + 12t^2 \\ 4t + 20t^3 & 24t + 2t^2 \end{bmatrix} + \begin{bmatrix} 19t^2 & 3t + 24t^2 \\ 4t + 20t^3 & 24t + 4t^2 \end{bmatrix} \\ &= \begin{bmatrix} 51t^2 & 6t + 36t^2 \\ 8t + 40t^3 & 48t + 6t^2 \end{bmatrix} \end{aligned}$$

(5) Let A be a nonsingular matrix whose elements are functions of the **scalar** t . Then, the elements of A^{-1} will also be functions of t . Thus,

$$\frac{\partial A^{-1}}{\partial t} = -A^{-1} \frac{\partial A}{\partial t} A^{-1}$$

Proof: Consider differentiating $AA^{-1} = I$, i.e.,

$$\frac{\partial AA^{-1}}{\partial t} - \frac{\partial I}{\partial t} = 0$$

By (4) above

$$\begin{aligned} \frac{\partial AA^{-1}}{\partial t} &= A \frac{\partial A^{-1}}{\partial t} + \frac{\partial A}{\partial t} A^{-1} = 0 \\ \Rightarrow A^{-1} A \frac{\partial A^{-1}}{\partial t} + A^{-1} \frac{\partial A}{\partial t} A^{-1} &= 0 \end{aligned}$$

$$\Rightarrow \frac{\partial A^{-1}}{\partial t} = -A^{-1} \frac{\partial A}{\partial t} A^{-1}$$

Special cases of (5)

$$(5.1) \quad \frac{\partial A^{-1}}{\partial a_{ij}} = -a^j (a')^i$$

where

a_{ij} = ijth element of A

a^j = jth column of A^{-1}

$(a')^i$ = ith row of A^{-1}

Proof:

$$\frac{\partial A}{\partial a_{ij}} = E_{ji} = e_j e_i^{'},$$

where

E_{ji} = n × n matrix having a 1 in the ijth location and 0's

elsewhere

e_j = n × 1 column vector with a 1 in the jth location and 0's

elsewhere

$e_i^{'}$ = n × 1 row vector with a 1 in the ith location and 0's

elsewhere

Thus, by (5),

$$\frac{\partial A^{-1}}{\partial a_{ij}} = -A^{-1} E_{ji} A^{-1} = -(A^{-1} e_j)(e_i^{'}) A^{-1} = -a^j (a')^i$$

$$(5.2) \quad \frac{\partial a^{km}}{\partial a_{ij}} = -a^{kj} a^{im}, \quad a^{pq} = pq^{\text{th}} \text{ element of } A^{-1}$$

(6) Let each element of a matrix $A_{n \times n}$ be a function of the scalar t and let $B_{n \times n}$ be a matrix

whose elements are independent of t .

Then,

$$\text{tr}(AB) = \sum_{i=1}^n \sum_{j=1}^n a_{ij} b_{ji}$$

Thus,

$$\frac{\partial \text{tr}(AB)}{\partial t} = \sum_{i=1}^n \sum_{j=1}^n \frac{\partial a_{ij}}{\partial t} b_{ji}$$

So, by (3) we have that

$$\frac{\partial \text{tr}(AB)}{\partial t} = \text{tr} \left[\frac{\partial A}{\partial t} B \right]$$

(6.1) If $t = a_{ij}$, i.e., the ij^{th} element of A , then:

$$\frac{\partial \text{tr}(AB)}{\partial a_{ij}} = b_{ji}, \quad b_{ji} = ji^{\text{th}} \text{ element of } B$$

Proof:

$$\frac{\partial \text{tr}(AB)}{\partial a_{ij}} = \sum_{k=1}^n \sum_{m=1}^n \frac{\partial a_{km}}{\partial a_{ij}} b_{mk} = b_{ji}$$

(6.2) If $t = \{a_{ij}\}$, i.e., t is a **matrix**, then we can define:

$$\frac{\partial}{\partial A} \text{tr}(AB) \equiv \left\{ \frac{\partial}{\partial a_{ij}} \text{tr}(AB) \right\} = \{b_{ji}\} = B'$$

Proof:

$$\frac{\partial}{\partial A} \text{tr}(AB) = \left\{ \sum_{k=1}^n \sum_{m=1}^n \frac{\partial a_{km}}{\partial a_{ij}} b_{mk} \right\} = \{b_{ji}\} = B'$$

(6.3) Generalizing, if the elements of the matrix A are functions of the elements of a matrix $T_{n \times n}$, and B is a constant matrix,

$$\frac{\partial}{\partial T} \text{tr}(AB) \equiv \left\{ \frac{\partial}{\partial t_{ij}} \text{tr}(AB) \right\} = \left\{ \text{tr} \left[\frac{\partial A}{\partial t_{ij}} B \right] \right\}$$

by (3) and (6.2) above.

$$(6.4) \quad \frac{\partial}{\partial A} \text{tr}(AB) = \frac{\partial}{\partial A} \text{tr}(BA) = B'$$

Proof: This is true because $\text{tr}(AB) = \text{tr}(BA)$.

$$(6.5) \quad \frac{\partial}{\partial A} \text{tr}(A' B) = \frac{\partial}{\partial A} \text{tr}(B' A) = B$$

Proof:

$$\frac{\partial}{\partial A} \text{tr}(B' A) = \left\{ \sum_{k=1}^n \sum_{m=1}^n b_{mk} \frac{\partial a_{km}}{\partial a_{ij}} \right\} = \{b_{ij}\} = B$$

$$\frac{\partial}{\partial A} \text{tr}(B' A) = \frac{\partial}{\partial A} \text{tr}(A' B) \quad \text{because } \text{tr}(A'B) = \text{tr}(B'A).$$

$$(6.6) \quad \frac{\partial}{\partial A} \text{tr}(CA'BAD) = BADC + B'AC'D'$$

Proof:

$$\frac{\partial}{\partial A} \text{tr}(CA'BAD) = \text{tr}(A' BADC)$$

$$\frac{\partial}{\partial A} \text{tr}(CA'BAD) = \text{tr}(DCA' BA)$$

By the derivative of a quadratic form (2), we have that:

$$\frac{\partial}{\partial A} \text{tr}(CA'BAD) = \frac{\partial}{\partial A} \text{tr}(A' BADC) + \frac{\partial}{\partial A} \text{tr}(DCA' BA)$$

$$\frac{\partial}{\partial A} \text{tr}(CA'BAD) = BADC \quad \text{by (6.5)}$$

$$+ \quad B'AC'D' \quad \text{by (6.2)}$$

$$(6.7) \quad \frac{\partial}{\partial A} \text{tr}(BCADE) = C'B'E'D'$$

Proof:

$$\frac{\partial}{\partial A} \text{tr}(BCADE) = \frac{\partial}{\partial A} \text{tr}(ADEBC)$$

$$\frac{\partial}{\partial A} \text{tr}(BCADE) = C'B'E'D' \quad \text{by (6.2)}$$

$$(6.8) \quad \frac{\partial}{\partial A} u'BACv = B'uv'C'$$

Proof: $u'BACv$ is a scalar. Because a scalar equals its own trace,

$$\frac{\partial}{\partial A} u'BACv = \frac{\partial}{\partial A} \text{tr}(u'BACv)$$

$$\frac{\partial}{\partial A} u'BACv = \frac{\partial}{\partial A} \text{tr}(ACvu'B)$$

$$\frac{\partial}{\partial A} u'BACv = B'uv'C' \quad \text{by (6.2)}$$

References

Goldberger, A. S. 1964. Econometrics Theory. John Wiley and Sons, Inc., NY.

Searle, S. R. 1982. Matrix Algebra Useful for Statistics. John Wiley and Sons, Inc., NY.