Neighborhood: Vector

Definition 1

Let $\mathbf{c}$ represent an arbitrary *m*–dimensional column vector. Taking the norm for $\mathbb{R}^{m\times 1}$ to be the usual norm, a *neighborhood* of $\mathbf{c}$ is a set of the general form

$$\{\mathbf{x} \in \mathbb{R}^{m\times 1} : ||\mathbf{x} - \mathbf{c}|| < r\}$$

where r is a positive number called the *radius* of the neighborhood.

Neighborhood: Matrix

The definition of neighborhood can be extended to row vectors. In fact, it can be extended to matrices of any dimensions.

Definition 2

Let $\mathrm{C} \in \mathbb{R}^{m \times n}$ represent an arbitrary matrix of order $m \times n$. Taking the norm for $\mathbb{R}^{m \times n}$ to be the usual norm, a *neighborhood* of C is a set of the general form

$$\{\mathrm{X} \in \mathbb{R}^{m \times n} : ||\mathrm{X} - \mathrm{C}|| < r\}.$$

Interior point

Definition 3

Let S represent an arbitrary set of m–dimensional column vectors (or more generally of $m \times n$ matrices), that is, let S represent a subset of $\mathbb{R}^{m\times 1}$ (or more generally of $\mathbb{R}^{m\times n}$). Then, a vector $\mathbf{x}$ (or matrix X) in S is said to be an *interior point* of S if there exists a neighborhood of $\mathbf{x}$ (or X), all of whose points belong to S.

Open set

Definition 4

The set S is said to be *open* if all of its points are interior points.

First-order partial derivative

Definition 5

Let f represent a function, defined on a set S, of a vector $\mathbf{x} = (x_1, \ldots, x_m)'$ of m variables. Suppose that S contains at least some interior points, and let $\mathbf{c} = (c_1, \ldots, c_m)'$ represent an arbitrary one of those points. Further, let $\mathbf{u}_j$ represent the jth column of I_m. Consider the limit

$$\lim_{t \to 0} \frac{f(\mathbf{c} + t\mathbf{u}_j) - f(\mathbf{c})}{t}.$$

When this limit exists, it is called the jth (first–order) *partial derivative* of f at $\mathbf{c}$ and is denoted by $D_j f(\mathbf{c})$.

Derivative

- The scalar $D_j f(\mathbf{c})$ can be regarded as the value assigned to the point $\mathbf{c}$ by a function.
- This function is denoted by the symbol $D_j f$ (and is referred to as the jth partial derivative of f).
- Its domain consists of those interior points (of S) at which the jth partial derivative (of f) is defined.

Derivative

- The symbol $\mathrm{D}f$ represents the row vector $(D_1 f, \ldots, D_m f)$ whose elements are the partial derivatives of f.
- Accordingly the symbol $\mathrm{D}f(\mathbf{c})$ represents the row vector $[D_1 f(\mathbf{c}), \ldots, D_m f(\mathbf{c})]$ whose elements are the values of the functions $D_1 f, \ldots, D_m f$ at $\mathbf{c}$.
- Note that $\mathrm{D}f(\mathbf{c})$ is defined only if $\mathbf{c}$ is such that all m of the partial derivatives of f at $\mathbf{c}$ exist.
- The column vector $(\mathrm{D}f)'$ is referred to as the *gradient* (or gradient vector) of f.

Derivative

An alternative notation is obtained by writing

- $\partial f(\mathbf{x})/\partial x_j$ for the jth partial derivative of f at $\mathbf{x}$,
- $\partial f(\mathbf{x})/\partial \mathbf{x}'$ for the row vector $[\partial f(\mathbf{x})/\partial x_1, \ldots, \partial f(\mathbf{x})/\partial x_m]$ of partial derivatives of f at $\mathbf{x}$,
- $\partial f(\mathbf{x})/\partial \mathbf{x}$ for the column vector $[\partial f(\mathbf{x})/\partial x_1, \ldots, \partial f(\mathbf{x})/\partial x_m]'$ of partial derivatives of f at $\mathbf{x}$,
- The symbols $\partial f(\mathbf{x})/\partial x_j$, $\partial f(\mathbf{x})/\partial \mathbf{x}'$ and $\partial f(\mathbf{x})/\partial \mathbf{x}$ have the same interpretations as $D_j f(\mathbf{x})$, $\mathrm{D}f(\mathbf{x})$ and $[\mathrm{D}f(\mathbf{x})]'$, respectively.

PD:Matrix

- Suppose that the domain of the function f to be differentiated is the set $\mathbb{R}^{m \times n}$ of all $m \times n$ matrices or, more generally, is a set S in $\mathbb{R}^{m \times n}$ that contains at least some interior points.
- Then f can be regarded as a function of an $m \times n$ matrix $\mathrm{X} = \{x_{ij}\}$ of mn "independent" variables.
- For purposes of differentiating f, the elements of X can be rearranged in the form of an mn–dimensional column vector $\mathbf{x}$, and f can be reinterpreted as a function of $\mathbf{x}$, in which case the domain of f is the set, say S^*, obtained by rearranging the elements of each $m \times n$ matrix in S in the form of a column vector.

PD:Matrix

- By definition, the elements $\partial f/\partial x_{ij}$ $(i = 1, \ldots, m; j = 1, \ldots, n)$ of the mn–dimensional column vector $\partial f/\partial \mathbf{x}$ are the first-order partial derivatives of f at $\mathbf{x}$.
- Instead of representing $\partial f/\partial \mathbf{x}$ as a column vector, it is represented as $m \times n$ matrix with elements $\partial f/\partial x_{ij}$ $(i = 1, \ldots, m; j = 1, \ldots, n)$.
- This matrix is to be denoted by the symbol $\partial f(\mathrm{X})/\partial \mathrm{X}$ and is to be called the derivative of $f(\mathrm{X})$ with respect to X.

PD:Matrix

- let us write $\partial f(\mathrm{X})/\partial \mathrm{X}'$ for the $n \times m$ matrix $[\partial f(\mathrm{X})/\partial \mathrm{X}]'$.
- Refer to this matrix as the derivative of $f(\mathrm{X})$ with respect to X' .

PD:Matrix

Lemma 6

Let f represent a function, defined on a set S, of a vector $\mathbf{x} = (x_1, \ldots, x_m)'$ of m variables, and suppose that $f(\mathbf{x})$ is constant or does not vary with x_j. Then, for any interior point $\mathbf{c}$ of S, $D_j f(\mathbf{c}) = 0$.

PD:Matrix

Lemma 7

Let f and g represent functions, defined on a set S and let a and b represent constants. Define,

$$l = af + bg, \; h = fg, \; \text{and} \;\; r = f/g,$$

so that l and h are functions, each of whose domain is S, and r is a function whose domain is $S^* = \{x \in S : g(x) \neq 0\}$.

PD:Matrix

Lemma 7

If f and g are continuously differentiable at an interior point $\mathbf{c}$ of S, then l and h are also continuously differentiable at $\mathbf{c}$, and

$$D_j l(\mathbf{c}) = a(\mathbf{c}) D_j f(\mathbf{c}) + b(\mathbf{c}) D_j g(\mathbf{c}) \tag{1}$$

$$D_j h(c) = f(\mathbf{c}) D_j g(\mathbf{c}) + g(\mathbf{c}) D_j f(cv). \tag{2}$$

PD:Matrix

Lemma 7

And, if f and g are continuously differentiable at an interior point $\mathbf{c}$ of S^*, then r is also continuously differentiable at $\mathbf{c}$, and

$$D_j r(\mathbf{c}) = [g(\mathbf{c}) D_j f(\mathbf{c}) - f(\mathbf{c}) D_j g(\mathbf{c})]/[g(\mathbf{c})]^2.$$

Differentiation of Linear Forms

Consider the function $f(\mathbf{x}) = \mathbf{a}'\mathbf{x} = \sum_{i=1} a_i x_i$.

$$\frac{\partial x_i}{\partial x_j} = \begin{cases} 1, & \text{if } i = j \\ 0, & \text{if } i = j. \end{cases} \tag{1}$$

Therefore,

$$\frac{\partial \mathbf{a}'\mathbf{x}}{\partial x_j} = \frac{\partial \sum_{i=1} a_i x_i}{\partial x_j} = a_j.$$

Thus,

$$\frac{\partial(\mathbf{a}'\mathbf{x})}{\partial \mathbf{x}} = \mathbf{a}$$

or alternatively as

$$\frac{\partial(\mathbf{a}'\mathbf{x})}{\partial \mathbf{x}'} = \mathbf{a}'.$$

Differentiation of Linear Forms

Consider the function $f(\mathbf{x}) = \mathrm{B}\mathbf{x}$ where $\mathrm{B} = \{b_{ij}\}$ represent a $p \times m$ matrix of constants. The ith element of $\mathrm{B}\mathbf{x}$ is the linear form $\mathbf{b}_i'\mathbf{x}$, whose coefficient vector is $\mathbf{b}_i' = (b_{i1}, \ldots, b_{im})$.

The jth partial derivative of this linear form is b_{ij}. Thus, the partial derivative of $\mathrm{B}\mathbf{x}$ with respect to $\mathbf{x}'$ is

$$\frac{\partial(\mathrm{B}\mathbf{x})}{\partial\mathbf{x}'} = \mathrm{B},$$

and the partial derivative of $(\mathrm{B}\mathbf{x})'$ with respect to $\mathbf{x}$ is

$$\frac{\partial(\mathrm{B}\mathbf{x})'}{\partial\mathbf{x}} = \mathrm{B}',$$

VERIFY!

Differentiation of Quadratic Forms

Consider the function $f(\mathbf{x}) = \mathbf{x}'\mathrm{A}\mathbf{x} = \sum_{i,k} a_{ik} x_i x_k$.

$$\frac{\partial(x_i x_k)}{\partial x_j} = \begin{cases} 2x_j, & \text{if } i = k = j, \\ x_i, & \text{if } k = j \text{ but } i \neq j, \\ x_k, & \text{if } i = j \text{ but } k \neq j, \\ 0, & \text{otherwise.} \end{cases} \tag{2}$$

Differentiation of Quadratic Forms

Then

$$\begin{aligned}
\frac{\partial(\mathbf{x}'\mathrm{A}\mathbf{x})}{\partial x_j} &= \frac{\partial(\sum_{i,k} a_{ik}x_i x_k)}{\partial x_j} \\
&= \frac{\partial(a_{jj}x_j^2 + \sum_{i\neq j} a_{ij}x_i x_j) + \sum_{k\neq j} a_{jk}x_j x_k) + \sum_{i\neq j,k\neq j} a_{ik}x_i x_k)}{\partial x_j} \\
&= 2a_{jj}x_j + \sum_{i\neq j} a_{ij}x_i + \sum_{k\neq j} a_{jk}x_k + 0 \\
&= \sum_i a_{ij}x_i + \sum_k a_{jk}x_k .
\end{aligned}$$

Differentiation of Quadratic Forms

Therefore

$$\frac{\partial(\mathbf{x}'\mathrm{A}\mathbf{x})}{\partial\mathbf{x}} = (\mathrm{A}+\mathrm{A}')\mathbf{x},$$

as is evident upon observing that $\sum_k a_{jk}x_k$ is the jth element of the column vector $\mathrm{A}\mathbf{x}$ and $\sum_i a_{ij}x_i$ is the jth element of $\mathrm{A}'\mathbf{x}$.

Matrices of Constants

Lemma 8

Let $F = \{f_{is}\}$ represent a $p \times q$ matrix of functions, defined on a set S, of a vector $\mathbf{x} = (x_1, \ldots, x_m)'$ of m variables; and suppose that (for $\mathbf{x} \in S$) $F(\mathbf{x})$ is constant or (more generally) does not vary with x_j. Then, at any interior point of S, $\partial F / \partial x_j = 0$.

Linear sum of Matrices

Lemma 9

Let $\mathrm{F} = \{f_{is}\}$ and $\mathrm{G} = \{g_{is}\}$ represent $p \times q$ matrices of functions, defined on a set S, of a vector $\mathbf{x} = (x_1, \ldots, xm)'$ of m variables. And, let a and b represent constants or (more generally) functions (defined on S) that are continuous at every interior point of S and are such that $a(\mathbf{x})$ and $b(\mathbf{x})$ do not vary with x_j. Then, at any interior point $\mathbf{c}$ (of S) at which F and G are continuously differentiable, $a\mathrm{F} + b\mathrm{G}$ is continuously differentiable and

$$\frac{\partial(a\mathrm{F} + b\mathrm{G})}{\partial x_j} = a\frac{\partial \mathrm{F}}{\partial x_j} + b\frac{\partial \mathrm{G}}{\partial x_j}. \tag{3}$$

Linear sum of Matrices

Proof: Let $\mathrm{L} = a\mathrm{F} + b\mathrm{G}$.The *is*th element of L is

$$l_{is} = af_{is} + bg_{is}.$$

The functions f_{is} and g_{is} are continuously differentiable at $\mathbf{c}$ implying that l_{is} is continuously differentiable at $\mathbf{c}$ and

$$\frac{\partial l_{is}}{\partial x_j} = a\frac{\partial f_{is}}{\partial x_j} + b\frac{\partial g_{is}}{\partial x_j}.$$

Linear sum of Matrices

Proof: It follows that L is continuously differentiable at $\mathbf{c}$. Since $\partial l_{is}/\partial x_j$, $\partial f_{is}/\partial x_j$, and $\partial g_{is}/\partial x_j$ are the *is*th elements of $\partial \mathrm{L}/\partial x_j$, $\partial \mathrm{F}/\partial x_j$, and $\partial \mathrm{G}/\partial x_j$ respectively. Thus,

$$\frac{\partial \mathrm{L}}{\partial x_j} = a\frac{\partial \mathrm{F}}{\partial x_j} + b\frac{\partial \mathrm{G}}{\partial x_j}.$$

Product of Matrices

Lemma 10

Let $\mathrm{F} = \{f_{is}\}$ *and* $\mathrm{G} = \{g_{is}\}$ *represent* $p \times q$ *and* $q \times r$ *matrices of functions, defined on a set* S*, of a vector* $\mathbf{x} = (x_1, \ldots, x_m)'$ *of* m *variables. Then, at any interior point* $\mathbf{c}$ *(of* S*) at which* F *and* G *are continuously differentiable,* FG *is continuously differentiable and*

$$\frac{\partial \mathrm{FG}}{\partial x_j} = \mathrm{F}\frac{\partial \mathrm{G}}{\partial x_j} + \frac{\partial \mathrm{F}}{\partial x_j}\mathrm{G}.$$

Product of Matrices

Proof: Let $\mathrm{H} = \mathrm{FG}$. The *it*th element of H is

$$h_{it} = \sum_{s=1}^{q} f_{is} s_{st}.$$

The functions f_{is} and g_{st} are are continuously differentiable at $\mathbf{c}$, implying that $f_{is}g_{st}$ is continuously differentiable at $\mathbf{c}$ and that

$$\frac{\partial f_{is} g_{st}}{\partial x_j} = f_{is} \frac{\partial g_{st}}{\partial x_j} + \frac{\partial f_{is}}{\partial x_j} g_{st}.$$

Product of Matrices

Proof: Thus, h_{it} is continuously differentiable at $\mathbf{c}$, and

$$\frac{\partial h_{it}}{\partial x_j} = \sum_{s=1}^{q} \frac{\partial (f_{is} g_{st})}{\partial x_j} = \sum_{s=1}^{q} f_{is} \frac{\partial g_{st}}{\partial x_j} + \sum_{s=1}^{q} \frac{\partial f_{is}}{\partial x_j} g_{st}.$$

We conclude that H is continuously differentiable at $\mathbf{c}$ and [since $\sum_{s=1}^{q} f_{is}(\partial g_{st}/\partial x_j)$ and $\sum_{s=1}^{q}(\partial f_{is}/\partial x_j) g_{st}$ are the *it*th elements of $\mathrm{F}(\partial \mathrm{G}/\partial x_j)$ and $(\partial \mathrm{F}/\partial x_j)\mathrm{G}$, respectively] that

$$\frac{\partial \mathrm{FG}}{\partial x_j} = \mathrm{F}\frac{\partial \mathrm{G}}{\partial x_j} + \frac{\partial \mathrm{F}}{\partial x_j}\mathrm{G}.$$

Differentiation of a vector wrt its elements

Let $\mathbf{x} = \{x_s\}$ represent an m–dimensional column vector, and let $\mathbf{u}_j$ represent the jth column of an identity matrix (of unspecified dimensions). Then

$$\frac{\partial \mathbf{x}}{\partial x_j} = \mathbf{u}_j. \tag{4}$$

Differentiation of a matrix wrt its elements

Result (4) cab be generalized to an $m \times n$ matrix $\mathrm{X} = \{x_{st}\}$.

$$\frac{\partial x_{st}}{\partial x_{ij}} = \begin{cases} 1, & \text{if } s = i \text{ and } t = j \\ 0, & \text{otherwise.} \end{cases}$$

Or, in matrix notation,

$$\frac{\partial \mathrm{X}}{\partial x_{ij}} = \mathbf{u}_i \mathbf{u}_j'.$$

Differentiation of a symmetric matrix wrt its elements

Suppose now that $\mathrm{X} = \{x_{st}\}$ is a symmetric matrix of order $m \times m$. Then

$$\frac{\partial x_{st}}{\partial x_{ii}} = \begin{cases} 1, & \text{if } s = t = i \\ 0, & \text{otherwise.} \end{cases}$$

and, for $j < i$ (or alternatively for $j > i$),

$$\frac{\partial x_{st}}{\partial x_{ij}} = \begin{cases} 1, & \text{if } s = i \text{ and } t = j \text{ or } s = j \text{ and } t = i, \\ 0, & \text{otherwise.} \end{cases}$$

Differentiation of a symmetric matrix wrt its elements

Or, in matrix notation,

$$\frac{\partial \mathrm{X}}{\partial x_{ii}} = \mathbf{u}_i \mathbf{u}_i',$$

and, for $j < i$ (or alternatively for $j > i$),

$$\frac{\partial \mathrm{X}}{\partial x_{ij}} = \mathbf{u}_i \mathbf{u}_j' + \mathbf{u}_j \mathbf{u}_i'.$$

Differentiation of a Trace of a Matrix

Lemma 11

Let $\mathrm{F} = \{f_{is}\}$ *represent a* $p \times p$ *matrix of functions, defined on a set* S*, of a vector* $\mathbf{x} = (x_1, \ldots, x_m)'$ *of* m *variables. Then, at any interior point* $\mathbf{c}$ *(of* S*) at which* F *is continuously differentiable,* $tr(\mathrm{F})$ *is continuously differentiable and*

$$\frac{\partial tr(\mathrm{F})}{\partial x_j} = tr\left(\frac{\partial \mathrm{F}}{\partial x_j}\right).$$

Differentiation of a Trace of a Matrix

Proof: As is evident upon observing that $tr(\mathrm{F}) = f_{11} + f_{22} + \cdots + f_{pp}$ [which establishes that $tr(\mathrm{F})$ is continuously differentiable at $\mathbf{c}$] and that (at $\mathbf{x} = \mathbf{c}$)

$$\frac{\partial tr(\mathrm{F})}{\partial x_j} = \frac{\partial f_{11}}{\partial x_j} + \cdots + \frac{\partial f_{pp}}{\partial x_j} = tr\left(\frac{\partial \mathrm{F}}{\partial x_j}\right).$$

Differentiation of a Trace of Product of two Matrices

Lemma 12

Let F *and* G *represent* $p \times q$ *and* $q \times p$ *matrices of functions, defined on a set* S*, of a vector* $\mathbf{x} = (x_1, \ldots, x_m)'$ *of* m *variables. Then*

$$\frac{\partial(tr(\mathrm{FG}))}{\partial x_j} = \frac{\partial(tr(\mathrm{GF}))}{\partial x_j} = tr\left(\mathrm{F}\frac{\partial \mathrm{G}}{\partial x_j}\right) + tr\left(\frac{\partial \mathrm{F}}{\partial x_j}\mathrm{G}\right).$$

Chain Rule

- Let $\mathrm{H} = \{h_{is}\}$ is an $n \times r$ matrix of functions, defined on S, of $\mathbf{x}$,
- g is a function, defined on a set T, of an $n \times r$ matrix $\mathrm{Y} = \{y_{is}\}$ of nr variables,
- $\mathrm{H}(\mathbf{x}) \in T$ for every $\mathbf{x}$ in S,
- f is the composite function defined (on S) by $f(\mathbf{x}) = g[\mathrm{H}(\mathbf{x})]$.

Chain Rule

- Suppose that the elements of H and Y are rearranged in the form of column vectors $\mathbf{h}$ and $\mathbf{y}$, respectively,
- for purposes of differentiation, g is reinterpreted as a function of $\mathbf{y}$,
- If h or, equivalently, H is continuously differentiable at an interior point $\mathbf{c}$ of S
- g is continuously differentiable at $\mathbf{h}(\mathbf{c})$ or, equivalently, $\mathrm{H}(\mathbf{c})$,

Chain Rule

- Then f is continuously differentiable at $\mathbf{c}$ and

$$\frac{\partial f}{\partial x_j} = \sum_{i=1}^{n} \sum_{s=1}^{r} \frac{\partial g}{\partial y_{is}} \frac{\partial h_{is}}{\partial x_j}$$

- Equivalently

$$\frac{\partial f}{\partial x_j} = tr\left[\left(\frac{\partial g}{\partial \mathrm{Y}}\right)' \left(\frac{\partial \mathrm{H}}{\partial x_j}\right)\right].$$

Partial Derivatives of Determinants

- $$\frac{\partial \det(\mathrm{X})}{\partial \mathrm{X}} = [\mathrm{adj}(\mathrm{X})]'.$$
- $$\frac{\partial \det(\mathrm{F})}{\partial x_j} = tr\left[\mathrm{adj}(\mathrm{F})\frac{\partial \mathrm{F}}{\partial x_j}\right]$$
- Moreover, if F is non–singular as well as continuously differentiable at $\mathbf{c}$, then
$$\frac{\partial \det(\mathrm{F})}{\partial x_j} = |\mathrm{F}|\, tr\left[\mathrm{F}^{-1}\frac{\partial \mathrm{F}}{\partial x_j}\right].$$

Partial Derivatives of Determinants

Result:

$$\frac{\partial \text{logdet}(\mathrm{X})}{\partial \mathrm{X}} = (\mathrm{X}^{-1})'$$

Proof:

$$\begin{aligned}\frac{\partial \text{logdet}(\mathrm{X})}{\partial x_{ij}} &= tr\left[\mathrm{X}^{-1}\frac{\partial \mathrm{X}}{\partial x_{ij}}\right] = tr(\mathrm{X}'\mathbf{u}_i\mathbf{u}_j') \\ &= \mathbf{u}_j'(\mathrm{X}^{-1})\mathbf{u}_i = y_{ij},\end{aligned}$$

where y_{ji} is the jith element of X^{-1} or, equivalently, the ijth element of $(\mathrm{X}^{-1})'$.

Partial Derivatives of Inverse of a Matrix

Result: Let F, A, and B are $p \times p$, $k \times p$, and $p \times r$ matrices, all are function of $\mathbf{x}$. F, A, and B all are continuously differentiable at an interior point $\mathbf{c}$ of S. Then

$$\frac{\partial(\mathrm{AF}^{-1}\mathrm{B})}{\partial x_j} = \mathrm{AF}^{-1}\frac{\partial \mathrm{B}}{\partial x_j} - \mathrm{AF}^{-1}\frac{\partial \mathrm{F}}{\partial x_j}\mathrm{F}^{-1}\mathrm{B} + \frac{\partial \mathrm{A}}{\partial x_j}\mathrm{F}^{-1}\mathrm{B}.$$