Statistical Sciences Honours
Matrix Methods
Lecture 1 – Nomenclature and Revision of Basics

Stefan S. Britz
stefan.britz@uct.ac.za

Department of Statistical Sciences
University of Cape Town

1.1 Nomenclature

Different mathematical objects we will be working with:

Concept	Symbol	Dimension	Handwritten
Scalar	a	1×1	
Matrix	$\boldsymbol{A}$	$n \times p$	
Column Vector	$\boldsymbol{a}$	$n \times 1$	
Row Vector	$\boldsymbol{a}'$	$1 \times p$	

1.1 Nomenclature

From a given matrix we can also extract specific rows, columns, or scalars:

1.2 Special Matrices

- Square matrix

$$\boldsymbol{A}\colon n \times n = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

1.2 Special Matrices

- Square matrix

$$\boldsymbol{A}\colon n \times n = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

- Diagonal matrix – $\boldsymbol{A}\colon n \times n$ where $a_{ij} = 0 \ \forall \ i \neq j$

$$\boldsymbol{A} = diag(d_i) = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix}$$

1.2 Special Matrices

- Symmetric matrix – $\boldsymbol{A}$: $n \times n$ where $a_{ij} = a_{ji}$

1.2 Special Matrices

- Symmetric matrix – $\boldsymbol{A}$: $n \times n$ where $a_{ij} = a_{ji}$
- Upper triangular matrix

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix}$$

1.2 Special Matrices

- Symmetric matrix – $\boldsymbol{A}: n \times n$ where $a_{ij} = a_{ji}$
- Upper triangular matrix

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix}$$

- Identity matrix

$$\boldsymbol{I}: n \times n = diag(\mathbf{1}) = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

1.3 Operations

- Transpose

  If $\boldsymbol{A} = \{a_{ij}\}$, then its transpose is $\boldsymbol{A}' = \boldsymbol{A}^T = \{a_{ji}\}$

  Note that for symmetric matrices $\boldsymbol{A} = \boldsymbol{A}'$

1.3 Operations

- Transpose

  If $\boldsymbol{A} = \{a_{ij}\}$, then its transpose is $\boldsymbol{A}' = \boldsymbol{A}^T = \{a_{ji}\}$

  Note that for symmetric matrices $\boldsymbol{A} = \boldsymbol{A}'$

- Addition

  $\boldsymbol{C} = \boldsymbol{A} + \boldsymbol{B}$ is the matrix $\{a_{ij} + b_{ij}\}$

1.3 Operations

- Transpose

  If $\boldsymbol{A} = \{a_{ij}\}$, then its transpose is $\boldsymbol{A}' = \boldsymbol{A}^T = \{a_{ji}\}$

  Note that for symmetric matrices $\boldsymbol{A} = \boldsymbol{A}'$

- Addition

  $\boldsymbol{C} = \boldsymbol{A} + \boldsymbol{B}$ is the matrix $\{a_{ij} + b_{ij}\}$

- Scalar multiplication

  $\boldsymbol{B} = c\boldsymbol{A}$ is the matrix $\{ca_{ij}\}$

1.3 Operations

- Transpose

  If $\boldsymbol{A} = \{a_{ij}\}$, then its transpose is $\boldsymbol{A}' = \boldsymbol{A}^T = \{a_{ji}\}$

  Note that for symmetric matrices $\boldsymbol{A} = \boldsymbol{A}'$

- Addition

  $\boldsymbol{C} = \boldsymbol{A} + \boldsymbol{B}$ is the matrix $\{a_{ij} + b_{ij}\}$

- Scalar multiplication

  $\boldsymbol{B} = c\boldsymbol{A}$ is the matrix $\{ca_{ij}\}$

- Trace

  For a square matrix, the trace is the sum of the diagonal

$$tr(\boldsymbol{A}) = \sum_{i=1}^{n} a_{ii}$$

1.3 Operations

- Partitioning

  It is often convenient to look at rectangular subsets of the matrix elements, which are themselves matrices (or vectors/scalars) of lower dimensions. For example, $\boldsymbol{A}\colon r \times c$ could be split up, or "partitioned" as follows:

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{A}_{11}\colon p \times q & \boldsymbol{A}_{12}\colon p \times (c-q) \\ \boldsymbol{A}_{21}\colon (r-p) \times q & \boldsymbol{A}_{22}\colon (r-p) \times (c-q) \end{bmatrix}$$

- Partitioning

  It is often convenient to look at rectangular subsets of the matrix elements, which are themselves matrices (or vectors/scalars) of lower dimensions. For example, $\boldsymbol{A}\colon r \times c$ could be split up, or "partitioned" as follows:

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{A}_{11}\colon p \times q & \boldsymbol{A}_{12}\colon p \times (c-q) \\ \boldsymbol{A}_{21}\colon (r-p) \times q & \boldsymbol{A}_{22}\colon (r-p) \times (c-q) \end{bmatrix}$$

- Other examples:

1.4 Linear Algebra

The concept of a matrix is intimately linked with the linear algebraic concept of the linear transformation of one vector to another.

Let $\boldsymbol{x}: n \times 1$ and $\boldsymbol{y}: m \times 1$ be n- and m-dimensional vectors respectively, such that each element of $\boldsymbol{y}$ is related to $\boldsymbol{x}$ through the linear relationship:

$$y_i = \sum_{j=1}^{n} a_{ij} x_j, \quad i = 1, \ldots, m$$

1.4 Linear Algebra

The concept of a matrix is intimately linked with the linear algebraic concept of the linear transformation of one vector to another.

Let $\boldsymbol{x}: n \times 1$ and $\boldsymbol{y}: m \times 1$ be n- and m-dimensional vectors respectively, such that each element of $\boldsymbol{y}$ is related to $\boldsymbol{x}$ through the linear relationship:

$$y_i = \sum_{j=1}^{n} a_{ij} x_j, \quad i = 1, \ldots, m$$

This can be expressed as:

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}$$

where $\boldsymbol{A}$ is the matrix $\{a_{ij}\}$

1.4 Linear Algebra

We can now define the multiplication of two matrices by considering a matrix as a collection of columns (or rows):

$$\begin{aligned}
\boldsymbol{C} = \boldsymbol{AB} &= \boldsymbol{A}[\boldsymbol{b}_1 \ \boldsymbol{b}_2 \ \dots \ \boldsymbol{b}_p] \\
&= [\boldsymbol{A}\boldsymbol{b}_1 \ \boldsymbol{A}\boldsymbol{b}_2 \ \dots \ \boldsymbol{A}\boldsymbol{b}_p] \\
&= \begin{bmatrix} \sum_{k=1}^n a_{1k}b_{k1} & \dots & \sum_{k=1}^n a_{1k}b_{kp} \\ \vdots & \ddots & \vdots \\ \sum_{k=1}^n a_{mk}b_{k1} & \dots & \sum_{k=1}^n a_{mk}b_{kp} \end{bmatrix}
\end{aligned}$$

such that

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

1.4 Linear Algebra

Note that the multiplication rules for matrices defined in terms of their elements apply also to manipulation on the component matrices of partitioned matrices.

For example,

$$\begin{bmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} \\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} \end{bmatrix} \begin{bmatrix} \boldsymbol{B}_{11} & \boldsymbol{B}_{12} \\ \boldsymbol{B}_{21} & \boldsymbol{B}_{22} \end{bmatrix} = \begin{bmatrix} \boldsymbol{A}_{11}\boldsymbol{B}_{11} + \boldsymbol{A}_{12}\boldsymbol{B}_{21} & \boldsymbol{A}_{11}\boldsymbol{B}_{12} + \boldsymbol{A}_{12}\boldsymbol{B}_{22} \\ \boldsymbol{A}_{21}\boldsymbol{B}_{11} + \boldsymbol{A}_{22}\boldsymbol{B}_{21} & \boldsymbol{A}_{21}\boldsymbol{B}_{12} + \boldsymbol{A}_{22}\boldsymbol{B}_{22} \end{bmatrix},$$

conditional on all products of partitions being conformable.

1.4 Linear Algebra

Matrix multiplication is crucial to countless statistical applications:

Exercises

Exercise 1.1

Show that if $\boldsymbol{D} = diag(d_i)$, then $\boldsymbol{DA} = \{d_i a_{ij}\}$; note that this implies that $\boldsymbol{IA} = \boldsymbol{A}$, where $\boldsymbol{I}$ is the identity matrix.

Exercise 1.2

Show that

(a) $(\boldsymbol{AB})' = \boldsymbol{B}'\boldsymbol{A}'$

(b) $tr(\boldsymbol{AB}) = tr(\boldsymbol{BA})$

1.5 Commutative, associative and distributive laws

The following laws apply to matrix operations:

- Associativity

$$(\boldsymbol{A} + \boldsymbol{B}) + \boldsymbol{C} = \boldsymbol{A} + (\boldsymbol{B} + \boldsymbol{C})$$

1.5 Commutative, associative and distributive laws

The following laws apply to matrix operations:

- Associativity

$$(\boldsymbol{A} + \boldsymbol{B}) + \boldsymbol{C} = \boldsymbol{A} + (\boldsymbol{B} + \boldsymbol{C})$$

- Distributivity

$$\boldsymbol{A}(\boldsymbol{B} + \boldsymbol{C}) = \boldsymbol{A}\boldsymbol{B} + \boldsymbol{A}\boldsymbol{C}$$

1.5 Commutative, associative and distributive laws

The following laws apply to matrix operations:

- Associativity

$$(\boldsymbol{A}+\boldsymbol{B})+\boldsymbol{C}=\boldsymbol{A}+(\boldsymbol{B}+\boldsymbol{C})$$

- Distributivity

$$\boldsymbol{A}(\boldsymbol{B}+\boldsymbol{C})=\boldsymbol{A}\boldsymbol{B}+\boldsymbol{A}\boldsymbol{C}$$

- Commutativity in addition...

$$\boldsymbol{A}+\boldsymbol{B}=\boldsymbol{B}+\boldsymbol{A}$$

1.5 Commutative, associative and distributive laws

The following laws apply to matrix operations:

- Associativity
$$(\boldsymbol{A}+\boldsymbol{B})+\boldsymbol{C}=\boldsymbol{A}+(\boldsymbol{B}+\boldsymbol{C})$$
- Distributivity
$$\boldsymbol{A}(\boldsymbol{B}+\boldsymbol{C})=\boldsymbol{A}\boldsymbol{B}+\boldsymbol{A}\boldsymbol{C}$$
- Commutativity in addition...
$$\boldsymbol{A}+\boldsymbol{B}=\boldsymbol{B}+\boldsymbol{A}$$
- ...but **NOT** in multiplication!
$$\boldsymbol{A}\boldsymbol{B}\neq\boldsymbol{B}\boldsymbol{A}$$

1.6 Other types of matrix multiplication

Two other "products" derive from different considerations:

- Hadamard product (Schur product) – $\boldsymbol{A} \circ \boldsymbol{B}$

  The element-wise or Hadamard product of two matrices of the same size is the matrix with elements $\{a_{ij}b_{ij}\}$.

1.6 Other types of matrix multiplication

Two other "products" derive from different considerations:

- Hadamard product (Schur product) – $\boldsymbol{A} \circ \boldsymbol{B}$

  The element-wise or Hadamard product of two matrices of the same size is the matrix with elements $\{a_{ij}b_{ij}\}$.

- Kronecker product – $\boldsymbol{A} \otimes \boldsymbol{B}$ The direct or Kronecker product of any two matrices (there are no size matching restrictions), say $\boldsymbol{A}\colon p \times q$ and $\boldsymbol{B}\colon m \times n$, is defined as the $pm \times qn$ matrix

$$\boldsymbol{A} \otimes \boldsymbol{B} = \begin{bmatrix} a_{11}\boldsymbol{B} & a_{12}\boldsymbol{B} & \dots & a_{1q}\boldsymbol{B} \\ a_{12}\boldsymbol{B} & a_{22}\boldsymbol{B} & \dots & a_{2q}\boldsymbol{B} \\ \vdots & \vdots & \ddots & \vdots \\ a_{p1}\boldsymbol{B} & a_{p2}\boldsymbol{B} & \dots & a_{pq}\boldsymbol{B} \end{bmatrix}$$

  Note that $(\boldsymbol{A} \otimes \boldsymbol{B})' = \boldsymbol{A}' \otimes \boldsymbol{B}'$

Class Exercise

Suppose you are given the following two matrices:

$$\boldsymbol{S} = \begin{bmatrix} 6 & -2 \\ 3 & 9 \end{bmatrix} \quad \text{and} \quad \boldsymbol{T} = \begin{bmatrix} 5 & 5 \\ 0 & -4 \end{bmatrix}$$

Calculate:

- $\boldsymbol{ST} =$
- $\boldsymbol{TS} =$
- $\boldsymbol{S} \circ \boldsymbol{T} =$
- $\boldsymbol{S} \otimes \boldsymbol{T} =$

- Symmetric matrices
  (a) If $\boldsymbol{A}$ and $\boldsymbol{B}$ are both symmetric, then $(\boldsymbol{AB})' = \boldsymbol{B}'\boldsymbol{A}' = \boldsymbol{BA} \neq \boldsymbol{AB}$ in general. Therefore, $\boldsymbol{AB}$ is not (necessarily) symmetric.
  (b) For any matrix $\boldsymbol{A}$, both $\boldsymbol{A}'\boldsymbol{A}$ and $\boldsymbol{AA}'$ are symmetric. This also applies to the $n \times n$ matrix $\boldsymbol{xx}'$ where $\boldsymbol{x}$ is an n-vector.
  (c) $tr(\boldsymbol{A}'\boldsymbol{A})$ is the sum of the squares of the elements of $\boldsymbol{A}$.

1.7 Properties of some Special Matrices

- Symmetric matrices
  (a) If $\boldsymbol{A}$ and $\boldsymbol{B}$ are both symmetric, then $(\boldsymbol{AB})' = \boldsymbol{B}'\boldsymbol{A}' = \boldsymbol{BA} \neq \boldsymbol{AB}$ in general. Therefore, $\boldsymbol{AB}$ is not (necessarily) symmetric.
  (b) For any matrix $\boldsymbol{A}$, both $\boldsymbol{A}'\boldsymbol{A}$ and $\boldsymbol{AA}'$ are symmetric. This also applies to the $n \times n$ matrix $\boldsymbol{xx}'$ where $\boldsymbol{x}$ is an n-vector.
  (c) $tr(\boldsymbol{A}'\boldsymbol{A})$ is the sum of the squares of the elements of $\boldsymbol{A}$.

Exercise 1.3

Show that for a square matrix $\boldsymbol{A}$: $n \times n$, $tr(\boldsymbol{A}'\boldsymbol{A}) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}^2$

1.7 Properties of some Special Matrices

- Matrices of ones

  Define $\mathbf{1}$ as the vector consisting only of 1's and $\boldsymbol{J} = \mathbf{1}\mathbf{1}'$ as the square matrix consisting only of 1's.

1.7 Properties of some Special Matrices

- Matrices of ones

  Define $\mathbf{1}$ as the vector consisting only of 1's and $\boldsymbol{J} = \mathbf{1}\mathbf{1}'$ as the square matrix consisting only of 1's.

Exercise 1.4

Show that if $\boldsymbol{x}$ and $\mathbf{1}$ have length n, then

(a) $\mathbf{1}'\mathbf{1} = n$

(b) $\mathbf{1}'\boldsymbol{x} = \sum_{i=1}^{n} x_i$

(c) $\left(\boldsymbol{I} - \frac{1}{n}\boldsymbol{J}\right)\boldsymbol{x}$ is the vector with elements $x_i - \bar{x}$ where $\bar{x}$ is the mean of the elements of $\boldsymbol{x}$. The matrix $\boldsymbol{I} - \frac{1}{n}\boldsymbol{J}$ is termed the centring matrix.

1.7 Properties of some Special Matrices

- Idempotency

  A matrix $\boldsymbol{K}$ such that $\boldsymbol{K}^2 = \boldsymbol{K}$ is said to be idempotent.

1.7 Properties of some Special Matrices

- Idempotency

  A matrix $\boldsymbol{K}$ such that $\boldsymbol{K}^2 = \boldsymbol{K}$ is said to be idempotent.

Exercise 1.5

Show that the centring matrix is idempotent. (Applying centring to an already centred vector does nothing further to it!)

1.7 Properties of some Special Matrices

- Idempotency

  A matrix $\boldsymbol{K}$ such that $\boldsymbol{K}^2 = \boldsymbol{K}$ is said to be idempotent.

Exercise 1.5

Show that the centring matrix is idempotent. (Applying centring to an already centred vector does nothing further to it!)

- Orthogonal vectors

  Vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ are orthogonal if $\boldsymbol{x'y} = \boldsymbol{y'x} = 0$.

1.7 Properties of some Special Matrices

- Idempotency

  A matrix $\boldsymbol{K}$ such that $\boldsymbol{K}^2 = \boldsymbol{K}$ is said to be idempotent.

Exercise 1.5

Show that the centring matrix is idempotent. (Applying centring to an already centred vector does nothing further to it!)

- Orthogonal vectors

  Vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ are orthogonal if $\boldsymbol{x}'\boldsymbol{y} = \boldsymbol{y}'\boldsymbol{x} = 0$.

- Norm

  The norm or length of $\boldsymbol{x}$ is defined by $||x|| = \sqrt{\boldsymbol{x}'\boldsymbol{x}}$. The vector $\boldsymbol{u} = \frac{\boldsymbol{x}}{||\boldsymbol{x}||}$ has a norm of 1 and is said to be normalised.

1.7 Properties of some Special Matrices

- Orthonormal vectors

  If $\boldsymbol{x}$ and $\boldsymbol{y}$ are orthogonal and normalised, they are an orthonormal pair of vectors. Clearly, if the columns of the matrix $\boldsymbol{P}$ are an orthonormal set of vectors, then $\boldsymbol{P}'\boldsymbol{P} = \boldsymbol{I}$. Similarly, if the rows form an orthonormal set then $\boldsymbol{P}\boldsymbol{P}' = \boldsymbol{I}$.

1.7 Properties of some Special Matrices

- Orthogonal matrix

  If the matrix $\boldsymbol{P}$ is square and the rows and columns are mutually orthonormal, $\boldsymbol{P}$ is an orthogonal matrix.

1.7 Properties of some Special Matrices

- Orthogonal matrix

  If the matrix $\boldsymbol{P}$ is square and the rows and columns are mutually orthonormal, $\boldsymbol{P}$ is an orthogonal matrix.

- Bilinear and quadratic forms

  Observe that

$$\boldsymbol{x}'\boldsymbol{A}\boldsymbol{y} = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_i y_j$$

  where $\boldsymbol{x}$ and $\boldsymbol{y}$ are $n \times 1$ vectors and $\boldsymbol{A}$ an $n \times n$ matrix. This is termed a bilinear form and, if $\boldsymbol{x} = \boldsymbol{y}$, a quadratic form.

Class Exercise

- Write out the quadratic form of a vector x and the embedded matrix

$$\begin{bmatrix} 2 & -1 & -1 \\ -1 & 1 & 1 \\ -1 & 1 & 4 \end{bmatrix}$$

Class Exercise

- Write out the quadratic form of a vector $\boldsymbol{x}$ and the embedded matrix

$$\begin{bmatrix} 2 & -1 & -1 \\ -1 & 1 & 1 \\ -1 & 1 & 4 \end{bmatrix}$$

- Determine an embedded matrix that would yield the quadratic form

$$g = x_1^2 - 2x_1x_2 - x_1x_3 + 3x_2^2 + 5x_3^2$$

1.7 Properties of some Special Matrices

- Positive definiteness

  A symmetric matrix $\boldsymbol{A}$ is:

  - Positive definite if $\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x} > 0$ for any $\boldsymbol{x} \neq 0$.
  - Positive semi-definite if $\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x} \geq 0$ and $\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x} = 0$ for at least one $\boldsymbol{x} \neq 0$.
  - Negative definite if $\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x} < 0$ for any $\boldsymbol{x} \neq 0$.

1.7 Properties of some Special Matrices

- Positive definiteness

  A symmetric matrix $\boldsymbol{A}$ is:

  - Positive definite if $\boldsymbol{x'Ax} > 0$ for any $\boldsymbol{x} \neq 0$.
  - Positive semi-definite if $\boldsymbol{x'Ax} \geq 0$ and $\boldsymbol{x'Ax} = 0$ for at least one $\boldsymbol{x} \neq 0$.
  - Negative definite if $\boldsymbol{x'Ax} < 0$ for any $\boldsymbol{x} \neq 0$.

Examples:

(a) $\boldsymbol{J}: 2 \times 2$ yields $\boldsymbol{x'Jx} = (x_1 + x_2)^2$ which is non-negative, but zero whenever $x_1 = -x_2$. Thus $\boldsymbol{J}$ is positive semi-definite.

(b) The matrix $\boldsymbol{A} = \begin{bmatrix} 2 & 1 \\ -2 & -1 \end{bmatrix}$ yields $\boldsymbol{x'Ax} = (x_1 - x_2)(2x_1 + x_2)$ and hence $\boldsymbol{A}$ is indefinite.

Exercises

Exercise 1.6

Show that

$$\boldsymbol{x}' \left(\boldsymbol{I} - \frac{1}{n}\boldsymbol{J} \right) \left(\boldsymbol{I} - \frac{1}{n}\boldsymbol{J} \right) \boldsymbol{x} = \boldsymbol{x}' \left(\boldsymbol{I} - \frac{1}{n}\boldsymbol{J} \right) \boldsymbol{x} = \sum_{i=1}^{n} (x_i - \bar{x})^2$$