Kurt Schmidheiny/Klaus Neusser
University of Basel

Elements of Matrix Algebra

Contents

Version: 20-9-2022, 22:16

Foreword

These lecture notes are supposed to summarize the main results concerning matrix algebra as they are used in econometrics and economics. For a deeper discussion of the material, the interested reader should consult the references listed at the end.

1 Definitions

A matrix is a rectangular array of numbers. Here we consider only real numbers. If the matrix has n rows and m columns, we say that the matrix is of dimension $(n \times m)$. We denote matrices by capital bold letters:

$$\mathbf{A} = (\mathbf{A})_{ij} = (a_{ij}) = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{pmatrix}$$

The numbers a_{ij} are called the elements of the matrix.

An $(n \times 1)$ matrix is a column vector with n elements. Similarly, a $(1 \times m)$ matrix is a row vector with m elements. We denote vectors by bold letters.

$$\mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \qquad \mathbf{b} = \begin{pmatrix} b_1 & b_2 & \dots & b_m \end{pmatrix}.$$

A (1×1) matrix is a scalar which is denoted by an italic letter.

The *null matrix* $(\mathbf{O})$ is a matrix whose elements are all equal to zero, i.e. $a_{ij} = 0$ for all $i = 1, \dots, n$ and $j = 1, \dots, m$.

A *square matrix* is a matrix with the same number of columns and rows, i.e. $n = m$.

A *symmetric matrix* is a square matrix such that $a_{ij} = a_{ji}$ for all $i = 1, \ldots, n$ and $j = 1, \ldots, m$.

A *diagonal matrix* is a square matrix such that the off-diagonal elements are all equal to zero, i.e. $a_{ij} = 0$ for $i \neq j$.

The *identity matrix* is a diagonal matrix with all diagonal elements equal to one. The identity matrix is denoted by $\mathbf{I}$ or $\mathbf{I}_n$.

A square matrix is said to be *upper triangular* whenever $a_{ij} = 0$ for $i > j$ and *lower triangular* whenever $a_{ij} = 0$ for $i < j$.

Two vectors $\mathbf{a}$ and $\mathbf{b}$ are said to be *linearly dependent* if there exist scalars α and β both not equal to zero such that $\alpha\mathbf{a} + \beta\mathbf{b} = \mathbf{0}$. Otherwise they are said to be *linearly independent*.

2 Matrix Operations

2.1 Equality

Two matrices or two vectors are equal if they have the same dimension and if their respective elements are all equal:

$$\mathbf{A} = \mathbf{B} \qquad \Longleftrightarrow \qquad a_{ij} = b_{ij} \quad \text{for all } i \text{ and } j$$

2.2 Transpose

Definition 1. The matrix $\mathbf{B}$ is called the transpose of matrix $\mathbf{A}$ if and only if

$$b_{ij} = a_{ji} \qquad \text{for all } i \text{ and } j.$$

The matrix $\mathbf{B}$ is denoted by $\mathbf{A}'$ or $\mathbf{A}^T$.

Taking the transpose of a matrix is equivalent to interchanging rows and columns. If $\mathbf{A}$ has dimension $(n \times m)$ then $\mathbf{A}'$ has dimension $(m \times n)$. The transpose of a column vector is a row vector and vice versa. Note:

- $(\mathbf{A}')' = \mathbf{A}$ for any matrix $\mathbf{A}$ (2.1)
- $\mathbf{A}' = \mathbf{A}$ for a symmetric matrix $\mathbf{A}$ (2.2)

2.3 Addition and Subtraction

The addition and subtraction of matrices is only defined for matrices with the same dimension.

Definition 2. The sum of two matrices $\mathbf{A}$ and $\mathbf{B}$ of the same dimensions is given by the sum of their elements, i.e.

$$\mathbf{C} = \mathbf{A} + \mathbf{B} \qquad \Longleftrightarrow \qquad c_{ij} = a_{ij} + b_{ij} \quad \text{for all } i \text{ and } j$$

The sum of a matrix $\mathbf{A}$ and a scalar b is a matrix $\mathbf{C} = \mathbf{A} + b$ with $c_{ij} = a_{ij} + b$. Note that $\mathbf{A} + b = b + \mathbf{A}$.

We have the following calculation rules if matrix dimensions agree:

- $\mathbf{A} + \mathbf{O} = \mathbf{A}$ (2.3)
- $\mathbf{A} - \mathbf{B} = \mathbf{A} + (-\mathbf{B})$ (2.4)
- $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$ (2.5)
- $(\mathbf{A} + \mathbf{B}) + \mathbf{C} = \mathbf{A} + (\mathbf{B} + \mathbf{C})$ (2.6)
- $(\mathbf{A} + \mathbf{B})' = \mathbf{A}' + \mathbf{B}'$ (2.7)

2.4 Product

Definition 3. The *inner product* (dot product, scalar product) of two vectors $\mathbf{a}$ and $\mathbf{b}$ of the same dimension $(n \times 1)$ is a scalar (real number) defined as:

$$\mathbf{a}'\mathbf{b} = \mathbf{b}'\mathbf{a} = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n = \sum_{i=1}^{n} a_i b_i.$$

The product of a scalar c and a matrix $\mathbf{A}$ is a matrix $\mathbf{B} = c\mathbf{A}$ with $b_{ij} = c a_{ij}$. Note that $c\mathbf{A} = \mathbf{A}c$ when c is a scalar.

Definition 4. The product of two matrices $\mathbf{A}$ and $\mathbf{B}$ with dimensions $(n \times k)$ and $(k \times m)$, respectively, is given by the matrix $\mathbf{C}$ with dimension

$(n \times m)$ such that

$$\mathbf{C} = \mathbf{A}\,\mathbf{B} \quad \Longleftrightarrow \quad c_{ij} = \sum_{s=1}^{k} a_{is} b_{sj} \quad \text{for all } i \text{ and } j$$

Remark 1. The matrix product is only defined if the number of columns of the first matrix is equal to the number of rows of the second matrix. Thus, although $\mathbf{A}\,\mathbf{B}$ may be defined, $\mathbf{B}\,\mathbf{A}$ is only defined if $n = m$. Thus for square matrices both $\mathbf{A}\,\mathbf{B}$ and $\mathbf{B}\,\mathbf{A}$ are defined.

Remark 2. The product of two matrices is in general **not** commutative, i.e. $\mathbf{A}\,\mathbf{B} \neq \mathbf{B}\,\mathbf{A}$.

Remark 3. The product $\mathbf{A}\,\mathbf{B}$ may also be defined as

$$c_{ij} = (\mathbf{C})_{ij} = \mathbf{a}'_{i\bullet}\mathbf{b}_{\bullet j}$$

where $\mathbf{a}'_{i\bullet}$ denotes the i-th row of $\mathbf{A}$ and $\mathbf{b}_{\bullet j}$ the j-th column of $\mathbf{B}$.

We have the following calculation rules if matrix dimensions agree:

- $\mathbf{AI} = \mathbf{A}, \quad \mathbf{IA} = \mathbf{A}$ (2.8)
- $\mathbf{AO} = \mathbf{O}, \quad \mathbf{OA} = \mathbf{O}$ (2.9)
- $(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC}) = \mathbf{ABC}$ (2.10)
- $\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}$ (2.11)
- $(\mathbf{B} + \mathbf{C})\mathbf{A} = \mathbf{BA} + \mathbf{CA}$ (2.12)
- $c(\mathbf{A} + \mathbf{B}) = c\mathbf{A} + c\mathbf{B}$ (2.13)
- $(\mathbf{AB})' = \mathbf{B}'\mathbf{A}'$ (order!) (2.14)
- $(\mathbf{ABC})' = \mathbf{C}'\mathbf{B}'\mathbf{A}'$ (order!) (2.15)

3 Rank of a Matrix

A set of vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ is *linearly independent* if $\sum_{i=1}^{n} c_i \mathbf{x}_i = \mathbf{0}$ implies $c_i = 0$ for all $i = 1, \ldots, n$.

The *column rank* of a matrix is the maximal number of linearly independent columns. The *row rank* of a matrix is the maximal number of linearly independent rows. A matrix is said to have full column (row) rank if the column rank (row rank) equals the number of columns (rows).

The column rank of an $n \times k$ matrix $\mathbf{A}$ is equal to its row rank. We can therefore just speak of the rank of a matrix denoted by $\text{rank}(\mathbf{A})$.

For an $(n \times k)$ matrix $\mathbf{A}$, a $(k \times m)$ matrix $\mathbf{B}$ and an $(n \times n)$ square matrix $\mathbf{C}$, we have

- $\text{rank}(\mathbf{A}) \leq \min(n, k)$ (3.1)
- $\text{rank}(\mathbf{A}') = \text{rank}(\mathbf{A})$ (3.2)
- $\text{rank}(\mathbf{A}'\mathbf{A}) = \text{rank}(\mathbf{A}\mathbf{A}') = \text{rank}(\mathbf{A})$ (3.3)
- $\text{rank}(\mathbf{AB}) \leq \min(\text{rank}(\mathbf{A}), \text{rank}(\mathbf{B}))$ (3.4)
- $\text{rank}(\mathbf{AB}) = \text{rank}(\mathbf{B})$ if $\mathbf{A}$ has full column rank (3.5)
- $\text{rank}(\mathbf{AB}) = \text{rank}(\mathbf{A})$ if $\mathbf{B}$ has full row rank (3.6)
- $\text{rank}(\mathbf{A}'\mathbf{CA}) = \text{rank}(\mathbf{CA})$ if $\mathbf{C}$ is nonnegative definite (3.7)
- $\text{rank}(\mathbf{A}'\mathbf{CA}) = \text{rank}(\mathbf{A})$ if $\mathbf{C}$ is positive definite (3.8)

4 Special Functions of Square Matrices

In this section only square $(n \times n)$ matrices are considered.

4.1 Trace of a Matrix

Definition 5. The *trace* of a matrix $\mathbf{A}$, denoted by $\mathrm{tr}(\mathbf{A})$, is the sum of its diagonal elements:

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

The following calculation rules hold if matrix dimensions agree:

- $\mathrm{tr}(c\mathbf{A}) = c\,\mathrm{tr}(\mathbf{A})$ (4.1)
- $\mathrm{tr}(\mathbf{A}') = \mathrm{tr}(\mathbf{A})$ (4.2)
- $\mathrm{tr}(\mathbf{A}+\mathbf{B}) = \mathrm{tr}(\mathbf{A}) + \mathrm{tr}(\mathbf{B})$ (4.3)
- $\mathrm{tr}(\mathbf{AB}) = \mathrm{tr}(\mathbf{BA})$ (4.4)
- $\mathrm{tr}(\mathbf{ABC}) = \mathrm{tr}(\mathbf{BCA}) = \mathrm{tr}(\mathbf{CAB})$ (4.5)

4.2 Determinant

The *determinant* of a$(n \times n)$ matrix $\mathbf{A}$ with $n > 1$ can be computed according to the following formula:

$$|\mathbf{A}| = \sum_{i=1}^{n} a_{ij}(-1)^{i+j}|\mathbf{A}_{ij}| \qquad \text{for some arbitrary } j$$

The determinant, computed as above, is said to be developed according to the j-th column. The term $(-1)^{i+j}|\mathbf{A}_{ij}|$ is called the cofactor of the element a_{ij}. Thereby $\mathbf{A}_{ij}$ is a matrix of dimension $((n-1) \times (n-1))$ which is obtained by deleting the i-th row and the j-th column.

$$\mathbf{A}_{ij} = \begin{pmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{i1} & \cdots & \boxed{a_{ij}} & \cdots & a_{in} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & a_{nj} & \cdots & a_{nn} \end{pmatrix}$$

For $n = 1$, i.e. if $\mathbf{A}$ is a scalar, the determinant $|\mathbf{A}|$ is defined as the absolute value. For $n = 2$, the determinant is given by:

$$|\mathbf{A}| = a_{11}a_{22} - a_{12}a_{21}.$$

If at least two columns (rows) are linearly dependent, the determinant is equal to zero and the inverse of $\mathbf{A}$ does not exist. The matrix is called *singular* in this case. If the matrix is *nonsingular* then all columns (rows) are linearly independent. If a column or a row has just zeros as its elements, the determinant is equal to zero. If two columns (rows) are interchanged, the determinant changes its sign.
Calculation rules for the determinant are:

- $|\mathbf{A}| = |\mathbf{A}'|$ (4.6)
- $|\mathbf{AB}| = |\mathbf{A}| \cdot |\mathbf{B}|$ (4.7)
- $|c\mathbf{A}| = c^n|\mathbf{A}|$ (4.8)

4.3 Inverse of a Matrix

If $\mathbf{A}$ is a square matrix, there may exist a matrix $\mathbf{B}$ with property $\mathbf{AB} = \mathbf{BA} = \mathbf{I}$. If such a matrix exists, it is called the *inverse* of A and is denoted by $\mathbf{A}^{-1}$, hence $\mathbf{AA}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$. The inverse of a matrix can be computed as follows

$$\mathbf{A}^{-1} = \frac{1}{|\mathbf{A}|} \begin{pmatrix} (-1)^{1+1}|\mathbf{A}_{11}| & (-1)^{2+1}|\mathbf{A}_{21}| & \ldots & (-1)^{n+1}|\mathbf{A}_{n1}| \\ (-1)^{1+2}|\mathbf{A}_{12}| & (-1)^{2+2}|\mathbf{A}_{22}| & \ldots & (-1)^{n+2}|\mathbf{A}_{n2}| \\ \vdots & \ddots & & \vdots \\ (-1)^{1+n}|\mathbf{A}_{1n}| & (-1)^{2+n}|\mathbf{A}_{2n}| & \ldots & (-1)^{n+n}|\mathbf{A}_{nn}| \end{pmatrix}$$

where $\mathbf{A}_{ij}$ is the matrix of dimension $(n-1)\times(n-1)$ obtained from $\mathbf{A}$ by deleting the i-th row and the j-th column.

$$\mathbf{A}_{ij} = \begin{pmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{i1} & \cdots & \boxed{a_{ij}} & \cdots & a_{in} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & a_{nj} & \cdots & a_{nn} \end{pmatrix}$$

The term $(-1)^{i+j}|\mathbf{A}_{ij}|$ is called the cofactor of a_{ij}.

For $n=2$, the inverse is given by

$$\mathbf{A}^{-1} = \frac{1}{a_{11}a_{22}-a_{12}a_{21}} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}.$$

We have the following calculation rules if both $\mathbf{A}^{-1}$ and $\mathbf{B}^{-1}$ exist and matrix dimensions agree:

- $\left(\mathbf{A}^{-1}\right)^{-1} = \mathbf{A}$ (4.9)
- $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$ (order!) (4.10)
- $(\mathbf{A}')^{-1} = \left(\mathbf{A}^{-1}\right)'$ (4.11)
- $|\mathbf{A}^{-1}| = |\mathbf{A}|^{-1} = \dfrac{1}{|\mathbf{A}|}$ (4.12)

4.4 Nonsingular Square Matrices

The following statements about a square $(n\times n)$ matrix $\mathbf{A}$ are equivalent:

- $\mathbf{A}$ is nonsingular (4.13)
- $|\mathbf{A}| \neq 0$ (4.14)
- $\mathbf{A}^{-1}$ exists (4.15)
- $rank(\mathbf{A}) = n$ (full rank) (4.16)
- $\lambda_i \neq 0$ for all $i = 1, ..., n$ (4.17)

5 Systems of Equations

Consider the following system of m equations in n unknowns $x_1,\ldots,x_n$:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

If we collect the unknowns into a vector $\mathbf{x} = (x_1,\ldots,x_n)'$, the coefficients $b_1,\ldots,b_n$ in to a vector $\mathbf{b}$, and the coefficients (a_{ij}) into a matrix $\mathbf{A}$, we can rewrite the equation system compactly in matrix form as follows:

$$\underbrace{\begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix}}_{\mathbf{A}} \underbrace{\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}}_{\mathbf{x}} = \underbrace{\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}}_{\mathbf{b}}$$

$$\mathbf{A}\,\mathbf{x} = \mathbf{b}$$

This equation system has a unique solution if $m = n$, i.e. if $\mathbf{A}$ is a square matrix, and $\mathbf{A}$ is nonsingular, i.e $\mathbf{A}^{-1}$ exits. The solution is then given by

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$$

Remark 4. To achieve numerical accuracy it is preferable not to compute the inverse explicitly. There are efficient numerical algorithms which can solve the equation system without computing the inverse.

6 Eigenvalue, -vector and Decomposition

6.1 Eigenvalue and Eigenvector

A scalar λ is said to be an *eigenvalue* of the square matrix $\mathbf{A}$ if there exists a vector $\mathbf{x} \neq 0$ such that

$$\mathbf{A}\,\mathbf{x} = \lambda \mathbf{x}$$

The vector $\mathbf{x}$ is called an *eigenvector* corresponding to λ. If $\mathbf{x}$ is an eigenvector then $\alpha\,\mathbf{x}$, $\alpha \neq 0$, is also an eigenvector. Eigenvectors are therefore not unique. It is sometimes useful to normalize the length of the eigenvectors to one, i.e. to choose the eigenvector such that $\mathbf{x}'\mathbf{x} = 1$.

6.2 Characteristic Equation

In order to find the eigenvalues and eigenvectors of a square matrix, one has to solve the equation system

$$\mathbf{A}\,\mathbf{x} = \lambda \mathbf{x} = \lambda \mathbf{I}\,\mathbf{x} \qquad \Longleftrightarrow \qquad (\mathbf{A} - \lambda\,\mathbf{I})\mathbf{x} = 0.$$

This equation system has a nontrivial solution, $\mathbf{x} \neq 0$, if and only if the matrix $(\mathbf{A} - \lambda\,\mathbf{I})$ is singular, or equivalently if and only if the determinant of $(\mathbf{A} - \lambda\,\mathbf{I})$ is equal to zero. This leads to an equation in the unknown parameter λ:

$$|\mathbf{A} - \lambda\,\mathbf{I}| = 0.$$

This equation is called the *characteristic equation* of the matrix $\mathbf{A}$ and corresponds to a polynomial equation of order n. The n solutions of this equation (roots) are the eigenvalues of the matrix. The solutions may be complex numbers. Some solutions may appear several times. Eigenvectors corresponding to some eigenvalue λ can be obtained from the equation $(\mathbf{A} - \lambda\,\mathbf{I})\mathbf{x} = 0$.

We have the following relations for an $(n \times n)$ matrix $\mathbf{A}$:

- $\mathrm{tr}(\mathbf{A}) = \sum_{i=1}^{n} \lambda_i$ (6.1)
- $|\mathbf{A}| = \prod_{i=1}^{n} \lambda_i$ (6.2)

6.3 Decomposition of Symmetric Matrices

If $\mathbf{A}$ is a symmetric $(n \times n)$ matrix, all n eigenvalues $\lambda_1, ..., \lambda_n$ are real and there exist n linearly independent eigenvectors $\mathbf{x}_1, \ldots, \mathbf{x}_n$ with the properties $\mathbf{x}_i'\mathbf{x}_j = 0$ for $i \neq j$ and $\mathbf{x}_i'\mathbf{x}_i = 1$, i.e the eigenvectors are orthogonal to each other and of length one. The eigenvector $\mathbf{x}_i$ corresponds to the eigenvalue λ_i.

A symmetric $(n \times n)$ matrix $\mathbf{A}$ can be diagonalized as

$$\mathbf{H}'\mathbf{A}\mathbf{H} = \mathbf{\Lambda}, \tag{6.3}$$

where the diagonal matrix $\mathbf{\Lambda}$ collects the eigenvalues of $\mathbf{A}$

$$\mathbf{\Lambda} = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix},$$

and the $(n \times n)$ matrix $\mathbf{H} = (\mathbf{x}_1, \ldots, \mathbf{x}_n)$ collecting the corresponding eigenvectors $\mathbf{x}_1, \ldots, \mathbf{x}_n$ is orthogonal

$$\mathbf{H}'\mathbf{H} = \mathbf{I},$$

hence $\mathbf{H}^{-1} = \mathbf{H}'$ and $\mathbf{H}\mathbf{H}' = \mathbf{I}$. We can therefore decompose $\mathbf{A}$ into the sum of n matrices:

$$\mathbf{A} = \mathbf{H}\mathbf{\Lambda}\mathbf{H}' = \sum_{i=1}^{n} \lambda_i \mathbf{x}_i \mathbf{x}_i'$$

where the matrices $\mathbf{h}_i\mathbf{h}_i'$ have all rank one. This decomposition is called the *spectral decomposition* or *eigendecomposition* of $\mathbf{A}$.

The inverse of a nonsingular symmetric matrix $\mathbf{A}$ can be calculated as

$$\mathbf{A}^{-1} = \mathbf{H}\mathbf{\Lambda}^{-1}\mathbf{H}' = \sum_{i=1}^{n} \frac{1}{\lambda_i} \mathbf{x}_i \mathbf{x}_i'.$$

Remark 5. Beside symmetric matrices, many other matrices, but not all matrices, are also diagonalizable.

7 Quadratic Forms

For a vector $\mathbf{x} \in \mathbb{R}^n$ and a symmetric matrix $\mathbf{A}$ of dimension $(n \times n)$ the scalar function

$$f(\mathbf{x}) = \mathbf{x}'\mathbf{A}\mathbf{x} = \sum_{j=1}^{n} \sum_{i=1}^{n} x_i x_j a_{ij}$$

is called a *quadratic form*.

The quadratic form $\mathbf{x}'\mathbf{A}\mathbf{x}$ and therefore the matrix $\mathbf{A}$ is called *positive (negative) definite*, if and only if

$$\mathbf{x}'\mathbf{A}\mathbf{x} > 0 (< 0) \quad \text{for all } x \neq 0.$$

The property that $\mathbf{A}$ is positive definite implies that

- $\lambda_i > 0$ for all $i = 1, ..., n$ (7.1)
- $|\mathbf{A}| > 0$ (7.2)
- $\mathbf{A}^{-1}$ exists and is positive definite (7.3)
- $\mathrm{tr}(\mathbf{A}) > 0$ (7.4)

The first property is an alternative definition for a positive definite matrix.

The quadratic form $\mathbf{x}'\mathbf{A}\mathbf{x}$ and therefore the matrix $\mathbf{A}$ is called *non-negative definite* or *positive semi-definite*, if and only if

$$\mathbf{x}'\mathbf{A}\mathbf{x} \geq 0 \quad \text{for all } x.$$

For nonnegative definite matrices we have:

- $\lambda_i \geq 0$ for all $i = 1, ..., n$ (7.5)
- $|\mathbf{A}| \geq 0$ (7.6)
- $\mathrm{tr}(\mathbf{A}) \geq 0$ (7.7)

The first property is an alternative definition for nonnegative definiteness.

For an $(n \times m)$ matrix $\mathbf{B}$,

- $\mathbf{B}'\mathbf{B}$ is nonnegative definite (7.8)
- $\mathbf{B}'\mathbf{B}$ is positive definite if $\mathbf{B}$ has full column rank (7.9)
- $\mathbf{B}\mathbf{B}'$ is nonnegative definite (7.10)

If the $(n \times m)$ matrix $\mathbf{B}$ has rank m (full column rank) and the $(n \times n)$ matrix $\mathbf{A}$ is positive definite then

- $\mathbf{B}'\mathbf{A}\mathbf{B}$ is positive definite (7.11)

The inverse of a symmetric positive definite $(n \times n)$ matrix $\mathbf{A}$ can be decomposed into

$$\mathbf{A}^{-1} = \mathbf{C}'\mathbf{C} \quad \text{where} \quad \mathbf{C}\mathbf{A}\mathbf{C}' = \mathbf{I}.$$

where C is a $(n \times n)$ matrix.

8 Partitioned Matrices

Consider a square matrix $\mathbf{P}$ of dimensions $((p+q)\times(r+s))$ which is partitioned into the $(p\times r)$ matrix $\mathbf{P}_{11}$, the $(p\times s)$ matrix $\mathbf{P}_{12}$, the $(q\times r)$ matrix $\mathbf{P}_{21}$ and the $(q\times s)$ matrix $\mathbf{P}_{22}$:

$$\mathbf{P} = \begin{pmatrix} \mathbf{P}_{11} & \mathbf{P}_{12} \\ \mathbf{P}_{21} & \mathbf{P}_{22} \end{pmatrix}$$

Assuming that dimensions in the involved multiplications agree, two partitioned matrices are multiplied as

$$\begin{pmatrix} \mathbf{P}_{11} & \mathbf{P}_{12} \\ \mathbf{P}_{21} & \mathbf{P}_{22} \end{pmatrix}\begin{pmatrix} \mathbf{Q}_{11} & \mathbf{Q}_{12} \\ \mathbf{Q}_{21} & \mathbf{Q}_{22} \end{pmatrix} = \begin{pmatrix} \mathbf{P}_{11}\mathbf{Q}_{11}+\mathbf{P}_{12}\mathbf{Q}_{21} & \mathbf{P}_{11}\mathbf{Q}_{12}+\mathbf{P}_{12}\mathbf{Q}_{22} \\ \mathbf{P}_{21}\mathbf{Q}_{11}+\mathbf{P}_{22}\mathbf{Q}_{21} & \mathbf{P}_{21}\mathbf{Q}_{12}+\mathbf{P}_{22}\mathbf{Q}_{22} \end{pmatrix}$$

Assuming that $\mathbf{P}_{11}^{-1}$ exists, the determinant of a partitioned matrix is

$$\begin{vmatrix} \mathbf{P}_{11} & \mathbf{P}_{12} \\ \mathbf{P}_{21} & \mathbf{P}_{22} \end{vmatrix} = |\mathbf{P}_{11}|\cdot|\mathbf{P}_{22}-\mathbf{P}_{21}\mathbf{P}_{11}^{-1}\mathbf{P}_{12}| \tag{8.1}$$

and the inverse is

$$\begin{pmatrix} \mathbf{P}_{11} & \mathbf{P}_{12} \\ \mathbf{P}_{21} & \mathbf{P}_{22} \end{pmatrix}^{-1} = \begin{pmatrix} \mathbf{P}_{11}^{-1}+\mathbf{P}_{11}^{-1}\mathbf{P}_{12}\mathbf{F}^{-1}\mathbf{P}_{21}\mathbf{P}_{11}^{-1} & -\mathbf{P}_{11}^{-1}\mathbf{P}_{12}\mathbf{F}^{-1} \\ -\mathbf{F}^{-1}\mathbf{P}_{21}\mathbf{P}_{11}^{-1} & \mathbf{F}^{-1} \end{pmatrix} \tag{8.2}$$

where $\mathbf{F}=\mathbf{P}_{22}-\mathbf{P}_{21}\mathbf{P}_{11}^{-1}\mathbf{P}_{12}$ is assumed nonsingular.

The determinant of a block diagonal matrix is

$$\begin{vmatrix} \mathbf{P}_{11} & \mathbf{O} \\ \mathbf{O} & \mathbf{P}_{22} \end{vmatrix} = |\mathbf{P}_{11}|\cdot|\mathbf{P}_{22}|$$

and its inverse is, assuming that $\mathbf{P}_{11}^{-1}$ and $\mathbf{P}_{22}^{-1}$ exist,

$$\begin{pmatrix} \mathbf{P}_{11} & \mathbf{O} \\ \mathbf{O} & \mathbf{P}_{22} \end{pmatrix}^{-1} = \begin{pmatrix} \mathbf{P}_{11}^{-1} & \mathbf{O} \\ \mathbf{O} & \mathbf{P}_{22}^{-1} \end{pmatrix}.$$

9 Derivatives with Matrix Algebra

A *linear function* f from the n-dimensional vector space of real numbers, $\mathbb{R}^n$, to the real numbers, $\mathbb{R}$, $f:\mathbb{R}^n \longrightarrow \mathbb{R}$ is determined by the coefficient vector $\mathbf{a}=(a_1,\ldots,a_n)'$:

$$y = f(\mathbf{x}) = \mathbf{a}'\mathbf{x} = \sum_{i=1}^{n} a_i x_i = a_1x_1+a_2x_2+\cdots+a_nx_n$$

where $\mathbf{x}$ is a column vector of dimension n and y a scalar.

The derivative of $y=f(\mathbf{x})$ with respect to the column vector $\mathbf{x}$ is defined as follows:

$$\frac{\partial y}{\partial \mathbf{x}} = \frac{\partial \mathbf{a}'\mathbf{x}}{\partial \mathbf{x}} = \frac{\partial \mathbf{x}'\mathbf{a}}{\partial \mathbf{x}} = \begin{pmatrix} \partial y/\partial x_1 \\ \partial y/\partial x_2 \\ \vdots \\ \partial y/\partial x_n \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \mathbf{a}$$

and with respect to the row vector $\mathbf{x}'$ as follows:

$$\frac{\partial y}{\partial \mathbf{x}'} = \frac{\partial \mathbf{a}'\mathbf{x}}{\partial \mathbf{x}'} = \frac{\partial \mathbf{x}'\mathbf{a}}{\partial \mathbf{x}'} = \begin{bmatrix} \frac{\partial y}{\partial x_1} & \frac{\partial y}{\partial x_2} & \ldots & \frac{\partial y}{\partial x_n} \end{bmatrix} = \begin{bmatrix} a_1 & a_2 & \ldots & a_n \end{bmatrix} = \mathbf{a}'$$

The simultaneous equation system $\mathbf{y}=\mathbf{A}\mathbf{x}$ can be viewed as m linear functions $y_i=\mathbf{a}_i'\mathbf{x}$ where $\mathbf{a}_i'$ denotes the i-th row of the $(m\times n)$ dimensional matrix $\mathbf{A}$. Thus the derivative of y_i with respect to $\mathbf{x}$ is given by

$$\frac{\partial y_i}{\partial \mathbf{x}} = \frac{\partial \mathbf{a}_i'\mathbf{x}}{\partial \mathbf{x}} = \mathbf{a}_i$$

Consequently the derivative of $\mathbf{y}=\mathbf{A}\mathbf{x}$ with respect to row vector $\mathbf{x}'$ can be defined as

$$\frac{\partial \mathbf{y}}{\partial \mathbf{x}'} = \frac{\partial \mathbf{A}\mathbf{x}}{\partial \mathbf{x}'} = \begin{pmatrix} \partial y_1/\partial \mathbf{x}' \\ \partial y_2/\partial \mathbf{x}' \\ \vdots \\ \partial y_m/\partial \mathbf{x}' \end{pmatrix} = \begin{pmatrix} \mathbf{a}_1' \\ \mathbf{a}_2' \\ \vdots \\ \mathbf{a}_m' \end{pmatrix} = \mathbf{A}.$$

The derivative of $\mathbf{y} = \mathbf{A}\mathbf{x}$ with respect to column vector $\mathbf{x}$ is therefore

$$\frac{\partial \mathbf{y}}{\partial \mathbf{x}} = \frac{\partial \mathbf{A}\mathbf{x}}{\partial \mathbf{x}} = \mathbf{A}'.$$

For a square matrix $\mathbf{A}$ of dimension $(n \times n)$ and the quadratic form $\mathbf{x}'\mathbf{A}\mathbf{x} = \sum_{j=1}^{n}\sum_{i=1}^{n} x_i x_j a_{ij}$ the derivative with respect to the column vector $\mathbf{x}$ is defined as

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

If $\mathbf{A}$ is a symmetric matrix this reduces to

$$\frac{\partial \mathbf{x}'\mathbf{A}\mathbf{x}}{\partial \mathbf{x}} = 2\mathbf{A}\mathbf{x}.$$

The derivative of the quadratic form $\mathbf{x}'\mathbf{A}\mathbf{x}$ with respect to the matrix elements a_{ij} is given by

$$\frac{\partial \mathbf{x}'\mathbf{A}\mathbf{x}}{\partial a_{ij}} = x_i x_j.$$

Therefore the derivative with respect to the matrix $\mathbf{A}$ is given by

$$\frac{\partial \mathbf{x}'\mathbf{A}\mathbf{x}}{\partial \mathbf{A}} = \mathbf{x}\mathbf{x}'.$$

10 Kronecker Product

The Kronecker Product of a $m \times n$ Matrix $\mathbf{A}$ with a $p \times q$ Matrix $\mathbf{B}$ is a $mp \times nq$ Matrix $\mathbf{A} \otimes \mathbf{B}$ defined as follows:

$$\mathbf{A} \otimes \mathbf{B} = \begin{pmatrix} a_{11}\mathbf{B} & a_{12}\mathbf{B} & \dots & a_{1n}\mathbf{B} \\ a_{21}\mathbf{B} & a_{22}\mathbf{B} & \dots & a_{21}\mathbf{B} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}\mathbf{B} & a_{m2}\mathbf{B} & \dots & a_{mn}\mathbf{B} \end{pmatrix}$$

The following calculation rules hold if matrix dimensions agree:

- $(\mathbf{A} \otimes \mathbf{B}) + (\mathbf{C} \otimes \mathbf{B}) = (\mathbf{A} + \mathbf{C}) \otimes \mathbf{B}$ (10.1)
- $(\mathbf{A} \otimes \mathbf{B}) + (\mathbf{A} \otimes \mathbf{C}) = \mathbf{A} \otimes (\mathbf{B} + \mathbf{C})$ (10.2)
- $(\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D}) = (\mathbf{A}\mathbf{C}) \otimes (\mathbf{B}\mathbf{D})$ (10.3)
- $(\mathbf{A} \otimes \mathbf{B})^{-1} = \mathbf{A}^{-1} \otimes \mathbf{B}^{-1}$ (10.4)
- $\text{tr}(\mathbf{A} \otimes \mathbf{B}) = \text{tr}(\mathbf{A})\text{tr}(\mathbf{B})$ (10.5)

References

[1] Abadir, K.M. and J.R. Magnus, *Matrix Algebra*, Cambridge: Cambridge University Press, 2005.

[2] Amemiya, T., *Introduction to Statistics and Econometrics*, Cambridge, Massachusetts: Harvard University Press, 1994.

[3] Dhrymes, P.J., *Introductory Econometrics*, New York : Springer-Verlag, 1978.

[4] Meyer, C.D., *Matrix Analysis and Applied Linear Algebra*, Philadelphia: SIAM, 2000.

[5] Strang, G., *Linear Algebra and its Applications*, 3rd Edition, San Diego: Harcourt Brace Jovanovich, 1986.

[6] Magnus, J.R., and H. Neudecker, *Matrix Differential Calculus with Applications in Statistics and Econometrics*, Chichester: John Wiley, 1988.

Formula Sources and Proofs

(2.8) Abadir and Magnus (2005), p. 28, ex. 2.18 (b).

(2.10) Abadir and Magnus (2005), p. 25, ex. 2.14 (a).

(2.11) Abadir and Magnus (2005), p. 25, ex. 2.14 (b).

(2.14) Abadir and Magnus (2005), p. 26, ex. 2.15 (a).

(2.15) Abadir and Magnus (2005), p. 26, ex. 2.15 (b).

(3.1) Abadir and Magnus (2005), p. 78 - 79, ex. 4.7 (a).

(3.2) Abadir and Magnus (2005), p. 77 - 78, ex. 4.5.

(3.3) Abadir and Magnus (2005), p. 81, ex. 4.13 (d).

(3.4) Abadir and Magnus (2005), p. 81, ex. 4.15 (b).

(3.5) Abadir and Magnus (2005), p. 85, ex. 4.25 (c).

(3.6) Abadir and Magnus (2005), p. 85, ex. 4.25 (d).

(3.7) Abadir and Magnus (2005), p. 221, ex. 8.27 (a).

(3.8) Abadir and Magnus (2005), p. 221, ex. 8.26 (a).

(4.1) Abadir and Magnus (2005), p. 30, ex. 2.24 (b).

(4.2) Abadir and Magnus (2005), p. 30, ex. 2.24 (c).

(4.3) Abadir and Magnus (2005), p. 30, ex. 2.24 (a).

(4.4) Abadir and Magnus (2005), p. 30, ex. 2.26 (a).

(4.5) Abadir and Magnus (2005), p. 31, ex. 2.26 (c).

(4.6) Abadir and Magnus (2005), p. 88, ex. 4.30.

(4.7) Abadir and Magnus (2005), p. 94, ex. 4.42.

(4.8) Abadir and Magnus (2005), p. 90, ex. 4.35 (a).

(4.9) Abadir and Magnus (2005), p. 84, ex. 4.22 (b).

(4.10) Abadir and Magnus (2005), p. 84, ex. 4.22 (d).

(4.11) Abadir and Magnus (2005), p. 84, ex. 4.22 (c).

(4.12) Abadir and Magnus (2005), p. 95, ex. 4.44 (a).

(4.13) Abadir and Magnus (2005), p. 83-84, ex. 4.21.

(4.14) Abadir and Magnus (2005), p. 94, ex. 4.43.

(4.15) Abadir and Magnus (2005), p. 83-84, ex. 4.21.

(4.16) Abadir and Magnus (2005), p. 83-84, ex. 4.21

(4.17) Abadir and Magnus (2005), p. 164, ex. 7.16.

(6.1) Abadir and Magnus (2005), p. 168, ex. 7.27.

(6.2) Abadir and Magnus (2005), p. 167, ex. 7.26.

(6.3) Abadir and Magnus (2005), p. 177, ex. 7.46.

(7.1) Abadir and Magnus (2005), p. 215, ex. 8.11 (a).

(7.2) Abadir and Magnus (2005), p. 215, ex. 8.12 (a).

(7.3) Abadir and Magnus (2005), p. 216, ex. 8.14 (c).

(7.4) Abadir and Magnus (2005), p. 215, ex. 8.12 (b).

(7.5) Abadir and Magnus (2005), p. 215, ex. 8.11 (b).

(7.6) Abadir and Magnus (2005), p. 216, ex. 8.13 (a)

(7.7) Abadir and Magnus (2005), p. 216, ex. 8.13 (b).

(7.8) Abadir and Magnus (2005), p. 214, ex. 8.9 (a).

(7.10) Abadir and Magnus (2005), p. 214, ex. 8.9 (a).

(7.11) Abadir and Magnus (2005), p. 221, ex. 8.26 (b).

(8.1) Abadir and Magnus (2005), p. 114, ex. 5.30 (a).

(8.2) Abadir and Magnus (2005), p. 106, ex. 5.16 (a).

(10.1) Abadir and Magnus (2005), p. 275, ex. 10.3 (a).

(10.2) Abadir and Magnus (2005), p. 275, ex. 10.3 (b).

(10.3) Abadir and Magnus (2005), p. 275, ex. 10.3 (d).

(10.4) Abadir and Magnus (2005), p. 278, ex. 10.8.

(10.5) Abadir and Magnus (2005), p. 277, ex. 10.7 (b).