# Elements of Matrix Algebra 

Klaus Neusser Kurt Schmidheiny

September 30, 2015

## Contents

1 Definitions ..... 2
2 Matrix operations ..... 3
3 Rank of a Matrix ..... 5
4 Special Functions of Quadratic Matrices ..... 6
4.1 Trace of a Matrix ..... 6
4.2 Determinant ..... 6
4.3 Inverse of a Matrix ..... 7
5 Systems of Equations ..... 8
6 Eigenvalue, Eigenvector and Spectral Decomposition ..... 9
7 Quadratic Forms ..... 11
8 Partitioned Matrices ..... 12
9 Derivatives with Matrix Algebra ..... 13
10 Kronecker Product ..... 14

## 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:

$$
\mathbf{A}=(\mathbf{A})_{i j}=\left(a_{i j}\right)=\left(\begin{array}{cccc}
a_{11} & a_{12} & \ldots & a_{1 m} \\
a_{21} & a_{22} & \ldots & a_{2 m} \\
\vdots & \vdots & \ddots & \vdots \\
a_{n 1} & a_{n 2} & \ldots & a_{n m}
\end{array}\right)
$$

The numbers $a_{i j}$ are called the elements of the matrix.
A $(n \times 1)$ 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.

$$
\mathbf{a}=\left(\begin{array}{c}
a_{1} \\
a_{2} \\
\vdots \\
a_{n}
\end{array}\right) \quad \mathbf{b}=\left(b_{1}, b_{2}, \ldots, b_{m}\right) .
$$

A $(1 \times 1)$ matrix is a scalar which is denoted by an italic letter.
The null matrix $(\mathbf{O})$ is a matrix all elements equal to zero, i.e. $a_{i j}=0$ for all $i=1, \ldots, n$ and $j=1, \ldots, m$.
A quadratic matrix is a matrix with the same number of columns and rows, i.e. $n=m$.

A symmetric matrix is a quadratic matrix such that $a_{i j}=a_{j i}$ for all $i=$ $1, \ldots, n$ and $j=1, \ldots, m$.
A diagonal matrix is a quadratic matrix such that the off-diagonal elements
are all equal to zero, i.e. $a_{i j}=0$ 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 quadratic matrix is said to be upper triangular whenever $a_{i j}=0$ for $i>j$ and lower triangular whenever $a_{i j}=0$ for $i<j$.
Two vectors $\mathbf{a}$ and $\mathbf{b}$ are said to be linearly dependent if there exists scalars $\alpha$ and $\beta$ both not equal to zero such $\alpha \mathbf{a}+\beta \mathbf{b}=\mathbf{0}$. Otherwise they are said to be linearly independent.

## 2 Matrix operations

## Equality

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

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

## Transpose

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

$$
b_{i j}=a_{j i} \quad \text { for all } i \text { and } j .
$$

The matrix $\mathbf{B}$ is denoted by $\mathbf{A}^{\prime}$ or $\mathbf{A}^{T}$.
Taking the transpose of some matrix is equivalent to interchanging rows and columns. If $\mathbf{A}$ has dimension $(n \times m)$ then $\mathbf{B}$ has dimension $(m \times n)$.
Taking the transpose of a column vector gives a row vector and vice versa. In general we mean vectors to be column vectors.

Remark 1. For any matrix $\mathbf{A},\left(\mathbf{A}^{\prime}\right)^{\prime}=\mathbf{A}$. For symmetric matrices $\mathbf{A}^{\prime}=\mathbf{A}$.

## Addition and Subtraction

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

Definition 2. The sum (difference) between two matrices $\mathbf{A}$ and $\mathbf{B}$ of the same dimension is given by the sum (difference) of its elements, i.e.

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

We have the following calculation rules:

$$
\begin{aligned}
\mathbf{A}+\mathbf{O} & =\mathbf{A} & & \text { addition of null matrix } \\
\mathbf{A}-\mathbf{B} & =\mathbf{A}+(-\mathbf{B}) & & \\
\mathbf{A}+\mathbf{B} & =\mathbf{B}+\mathbf{A} & & \text { kommutativity } \\
(\mathbf{A}+\mathbf{B})+\mathbf{C} & =\mathbf{A}+(\mathbf{B}+\mathbf{C}) & & \text { associativity } \\
(\mathbf{A}+\mathbf{B})^{\prime} & =\mathbf{A}^{\prime}+\mathbf{B}^{\prime} & &
\end{aligned}
$$

## 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:

$$
\mathbf{a}^{\prime} \mathbf{b}=\mathbf{b}^{\prime} \mathbf{a}=a_{1} b_{1}+a_{2} b_{2}+\cdots+a_{n} b_{n}=\sum_{i=1}^{n} a_{i} b_{i} .
$$

The product of a scalar $c$ and a matrix $\mathbf{A}$ is a matrix $\mathbf{B}=c \mathbf{A}$ with $b_{i j}=c a_{i j}$. 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

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

Remark 2. 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 A B may be defined, $\mathbf{B} \mathbf{A}$ is only defined if $n=m$. Thus for quadratic matrices both $\mathbf{A B}$ and $\mathbf{B A}$ are defined.

Remark 3. The product of two matrices is in general not commutative, i.e. AB $\neq \mathbf{B}$ A.

Remark 4. The product AB may also be defined as

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

where $\mathbf{a}_{\bullet}^{\prime}$ 。 denotes the $i$-th row of $\mathbf{A}$ and $\mathbf{b}_{\bullet j}$ the $j$-th column of $\mathbf{B}$.
Under the assumption that dimensions agree, we have the following calculation rules:

$$
\begin{aligned}
\mathbf{A I} & =\mathbf{A}, \quad \mathbf{I A}=\mathbf{A} & & \\
\mathbf{A O} & =\mathbf{O}, \quad \mathbf{O A}=\mathbf{O} & & \\
(\mathbf{A B}) \mathbf{C} & =\mathbf{A}(\mathbf{B C})=\mathbf{A B C} & & \text { associativity } \\
\mathbf{A}(\mathbf{B}+\mathbf{C}) & =\mathbf{A B}+\mathbf{A C} & & \text { distributivity } \\
(\mathbf{B}+\mathbf{C}) \mathbf{A} & =\mathbf{B A}+\mathbf{C A} & & \text { distributivity } \\
c(\mathbf{A}+\mathbf{B}) & =c \mathbf{A}+c \mathbf{B} & & \text { distributivity with scalar } c \\
(\mathbf{A B})^{\prime} & =\mathbf{B}^{\prime} \mathbf{A}^{\prime} & & \text { reverse ordering after transpose } \\
(\mathbf{A B C})^{\prime} & =\mathbf{C}^{\prime} \mathbf{B}^{\prime} \mathbf{A}^{\prime} & &
\end{aligned}
$$

## 3 Rank of a Matrix

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). For quadratic matrices the column rank is always equal to the row rank. In this case we just speak of the rank of a matrix. The rank of a quadratic matrix is denoted by $\operatorname{rank}(\mathbf{A})$.
For quadratic matrices we have:

$$
\begin{aligned}
\operatorname{rank}(\mathbf{A B}) & \leq \min \{\operatorname{rank}(\mathbf{A}), \operatorname{rank}(\mathbf{B})\} \\
\operatorname{rank}\left(\mathbf{A}^{\prime}\right) & =\operatorname{rank}(\mathbf{A}) \\
\operatorname{rank}(\mathbf{A}) & =\operatorname{rank}\left(\mathbf{A A}^{\prime}\right)=\operatorname{rank}\left(\mathbf{A}^{\prime} \mathbf{A}\right)
\end{aligned}
$$

## 4 Special Functions of Quadratic Matrices

In this section only quadratic $(n \times n)$ matrices with dimensions are considered.

### 4.1 Trace of a Matrix

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

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

The following calculation rules hold:

$$
\begin{aligned}
\operatorname{tr}(c \mathbf{A}) & =c \operatorname{tr}(\mathbf{A}) \\
\operatorname{tr}\left(\mathbf{A}^{\prime}\right) & =\operatorname{tr}(\mathbf{A}) \\
\operatorname{tr}(\mathbf{A}+\mathbf{B}) & =\operatorname{tr}(\mathbf{A})+\operatorname{tr}(\mathbf{B}) \\
\operatorname{tr}(\mathbf{A B}) & =\operatorname{tr}(\mathbf{B A}) \\
\operatorname{tr}(\mathbf{A B C}) & =\operatorname{tr}(\mathbf{B C A})=\operatorname{tr}(\mathbf{C A B})
\end{aligned}
$$

### 4.2 Determinant

The determinant can be computed according to the following formula:

$$
|\mathbf{A}|=\sum_{i=1}^{n} a_{i j}(-1)^{i+j}\left|A_{i j}\right| \quad \text { for some arbitrary } j
$$

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

$$
\mathbf{A}_{i j}=\left(\begin{array}{ccccc}
a_{11} & \cdots & a_{1 j} & \cdots & a_{1 n} \\
\vdots & & & & \vdots \\
\boldsymbol{a}_{i 1} & & a_{i j} & \cdots & a_{i n} \\
\vdots & & & & \vdots \\
a_{n 1} & \cdots & a_{n j} & \cdots & a_{n n}
\end{array}\right)
$$

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.
For $n=2$ and $n=3$, the determinant is given by a tractable formula:

$$
\begin{array}{rlrl}
n=2: & & |\mathbf{A}|= & a_{11} a_{22}-a_{12} a_{21} \\
n=3: & & |\mathbf{A}|= & a_{11}\left|\begin{array}{ll}
a_{22} & a_{23} \\
a_{32} & a_{33}
\end{array}\right|-a_{21}\left|\begin{array}{cc}
a_{12} & a_{13} \\
a_{32} & a_{33}
\end{array}\right|+a_{31}\left|\begin{array}{cc}
a_{12} & a_{13} \\
a_{22} & a_{23}
\end{array}\right| \\
& =a_{11} a_{22} a_{33}-a_{11} a_{23} a_{32}-a_{21} a_{12} a_{33}
\end{array}
$$

Calculation rules for the determinant are:

$$
\begin{aligned}
\left|\mathbf{A}^{\prime}\right| & =|\mathbf{A}| \\
|\mathbf{A B}| & =|\mathbf{A}| \cdot|\mathbf{B}| \\
|c \mathbf{A}| & =c^{n}|\mathbf{A}|
\end{aligned}
$$

### 4.3 Inverse of a Matrix

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

$$
\mathbf{A}^{-1}=\frac{1}{|\mathbf{A}|}\left(\begin{array}{cccc}
(-1)^{1+1}\left|\mathbf{A}_{11}\right| & (-1)^{2+1}\left|\mathbf{A}_{21}\right| & \ldots & (-1)^{n+1}\left|\mathbf{A}_{n 1}\right| \\
(-1)^{1+2}\left|\mathbf{A}_{12}\right| & (-1)^{2+2}\left|\mathbf{A}_{22}\right| & \ldots & (-1)^{n+2}\left|\mathbf{A}_{n 2}\right| \\
\vdots & \vdots & \ddots & \vdots \\
(-1)^{1+n}\left|\mathbf{A}_{1 n}\right| & (-1)^{2+n}\left|\mathbf{A}_{2 n}\right| & \ldots & (-1)^{n+n}\left|\mathbf{A}_{n n}\right|
\end{array}\right)
$$

where $\mathbf{A}_{i j}$ 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.

$$
\mathbf{A}_{i j}=\left(\begin{array}{ccccc}
a_{11} & \cdots & a_{1 j} & \cdots & a_{1 n} \\
\vdots & & & & \vdots \\
a_{i 1} & \cdots & & a_{i j} & \cdots \\
\vdots & & & a_{i n} \\
a_{n 1} & \cdots & & & \\
l_{n j} & \cdots & & a_{n n}
\end{array}\right)
$$

The term $(-1)^{i+j}\left|\mathbf{A}_{i j}\right|$ is called the cofactor of $a_{i j}$.
For $n=2$, the inverse is given by

$$
\mathbf{A}^{-1}=\frac{1}{a_{11} a_{22}-a_{12} a_{21}}\left(\begin{array}{rr}
a_{22} & -a_{12} \\
-a_{21} & a_{11}
\end{array}\right) .
$$

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

$$
\begin{aligned}
\left(\mathbf{A}^{-1}\right)^{-1} & =\mathbf{A} \\
(\mathbf{A B})^{-1} & =\mathbf{B}^{-1} \mathbf{A}^{-1} \quad \text { order reversed } \\
\left(\mathbf{A}^{\prime}\right)^{-1} & =\left(\mathbf{A}^{-1}\right)^{\prime} \\
\left|\mathbf{A}^{-1}\right| & =|\mathbf{A}|^{-1}
\end{aligned}
$$

## 5 Systems of Equations

Consider the following system of $n$ equations in $m$ unknowns $x_{1}, \ldots, x_{m}$ :

$$
\begin{aligned}
& a_{11} x_{1}+a_{12} x_{2}+\cdots+a_{1 m} x_{m}=b_{1} \\
& a_{21} x_{1}+a_{22} x_{2}+\cdots+a_{2 m} x_{m}=b_{2} \\
& \cdots \\
& \cdots \\
& a_{n 1} x_{1}+a_{n 2} x_{2}+\cdots+a_{n m} x_{m}=b_{n}
\end{aligned}
$$

If we collect the unknowns into a vector $\mathbf{x}=\left(x_{1}, \ldots, x_{m}\right)^{\prime}$, the coefficients $b_{1}, \ldots, b_{m}$ in to a vector $\mathbf{b}$, and the coefficients $\left(a_{i j}\right)$ into a matrix $\mathbf{A}$, we can
rewrite the equation system compactly in matrix form as follows:

$$
\underbrace{\left(\begin{array}{cccc}
a_{11} & a_{12} & \ldots & a_{1 m} \\
a_{21} & a_{22} & \ldots & a_{2 m} \\
\vdots & \vdots & \ddots & \vdots \\
a_{n 1} & a_{n 2} & \ldots & a_{n m}
\end{array}\right)}_{\mathbf{A}} \underbrace{\left(\begin{array}{c}
x_{1} \\
x_{2} \\
\vdots \\
x_{m}
\end{array}\right)}_{\mathbf{x}}=\underbrace{\left(\begin{array}{c}
b_{1} \\
b_{2} \\
\vdots \\
b_{n}
\end{array}\right)}_{\mathbf{b}}
$$

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

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

Remark 5. 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, Eigenvector and Spectral Decomposition

In this section we only consider quadratic matrices of dimension $n \times n$.

## Eigenvalue and Eigenvector

A scalar $\lambda$ is said to be an eigenvalue for the matrix $\mathbf{A}$ if there exists a vector $\mathrm{x} \neq 0$ such

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

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

## Characteristic equation

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

$$
\mathbf{A} \mathbf{x}=\lambda \mathbf{x}=\lambda \mathbf{I} \mathbf{x} \quad \Longleftrightarrow \quad(\mathbf{A}-\lambda \mathbf{I}) \mathbf{x}=0
$$

This equation system has a nontrivial solution, $\mathbf{x} \neq 0$, if and only if the $\operatorname{matrix}(\mathbf{A}-\lambda \mathbf{I})$ 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 $\lambda$ :

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

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 $\lambda$ can be obtained from the equation $(\mathbf{A}-\lambda \mathbf{I}) \mathbf{x}=0$.
We have the following important relations:

$$
\operatorname{tr}(\mathbf{A})=\sum_{i=1}^{n} \lambda_{i} \quad \text { and } \quad|\mathbf{A}|=\prod_{i=1}^{n} \lambda_{i}
$$

## Eigenvalues and eigenvectors of symmetric matrices

If $\mathbf{A}$ is a symmetric matrix, all eigenvalues are real and there exist $n$ linearly independent eigenvectors $\mathbf{x}_{1}, \ldots, \mathbf{x}_{n}$ with the properties $\mathbf{x}_{i}^{\prime} \mathbf{x}_{j}=0$ for $i \neq j$ and $\mathbf{x}_{i}^{\prime} \mathbf{x}_{i}=1$, i.e the eigenvectors are orthogonal to each other and of length one. If we collect of the eigenvectors into an $(n \times n)$ matrix $\mathbf{X}=\left(\mathbf{x}_{1}, \ldots, \mathbf{x}_{n}\right)$, we can write

$$
\mathbf{C}^{\prime} \mathbf{C}=\mathbf{C C}^{\prime}=\mathbf{I}
$$

If we collect all the eigenvalues into a diagonal matrix $\boldsymbol{\Lambda}$,

$$
\boldsymbol{\Lambda}=\left(\begin{array}{cccc}
\lambda_{1} & 0 & \ldots & 0 \\
0 & \lambda_{2} & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & \lambda_{n}
\end{array}\right)
$$

we can diagonalize the matrix $\mathbf{A}$ as follows:

$$
\mathrm{C}^{\prime} \mathrm{AC}=\mathrm{C}^{\prime} \mathrm{C} \boldsymbol{\Lambda}=\mathbf{I} \boldsymbol{\Lambda}=\boldsymbol{\Lambda} .
$$

This implies that we can decompose $\mathbf{A}$ into the sum of $n$ matrices as follows:

$$
\mathbf{A}=\mathbf{C} \boldsymbol{\Lambda} \mathbf{C}^{\prime}=\sum_{i=1}^{n} \lambda_{i} \mathbf{x}_{i} \mathbf{x}_{i}^{\prime}
$$

where the matrices $\mathbf{x}_{i} \mathbf{x}_{i}^{\prime}$ have all rank one. The above decomposition is called the spectral decomposition or eigenvalue decomposition of $\mathbf{A}$. The inverse of the matrix $\mathbf{A}$ is now easily calculated by

$$
\mathbf{A}^{-1}=\mathbf{C} \boldsymbol{\Lambda}^{-1} \mathbf{C}^{\prime}=\sum_{i=1}^{n} \frac{1}{\lambda_{i}} \mathbf{x}_{i} \mathbf{x}_{i}^{\prime}
$$

as $\mathbf{C}^{-1}=\mathbf{C}^{\prime}$.
Remark 6. Note that 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 quadratic matrix $\mathbf{A}$ of dimension $(n \times n)$ the scalar function

$$
f(\mathbf{x})=\mathbf{x}^{\prime} \mathbf{A} \mathbf{x}=\sum_{j=1}^{n} \sum_{i=1}^{n} x_{i} x_{j} a_{i j}=\mathbf{x}^{\prime} \mathbf{A} \mathbf{x}
$$

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

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

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

$$
\begin{array}{rlrl}
\lambda_{i} & >0 \text { for all } i & & \text { all eigenvalues are positive } \\
|\mathbf{A}| & >0 & \text { the determinant is positive } \\
\mathbf{A}^{-1} \text { exists } & \\
\operatorname{tr}(\mathbf{A}) & >0 &
\end{array}
$$

The first property can serve as an alternative definition for a positive definite matrix.
The quadratic form $\mathbf{x}^{\prime} \mathbf{A x}$ and therefore the matrix $\mathbf{A}$ is called nonnegative definite or positive semi-definite, if and only if

$$
\mathbf{x}^{\prime} \mathbf{A} \mathbf{x} \geq 0 \quad \text { for all } x
$$

For nonnegative definite matrices we have:

$$
\begin{aligned}
\lambda_{i} & \geq 0 \quad \text { for all } i \\
|\mathbf{A}| & \geq 0 \\
\operatorname{tr}(\mathbf{A}) & \geq 0
\end{aligned} \quad \text { the determinant is nonnegative }
$$

The first property can serve as an alternative definition for a nonnegative definite matrix.

Theorem 1. If the matrix $\mathbf{A}$ of dimension $(n \times m), n>m$, has full rank then $\mathbf{A}^{\prime} \mathbf{A}$ is positive definite and $\mathbf{A} \mathbf{A}^{\prime}$ is nonnegative definite.

## 8 Partitioned Matrices

Consider a quadratic 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 $\mathbf{P}_{22}$ :

$$
\mathbf{P}=\left(\begin{array}{ll}
\mathbf{P}_{11} & \mathbf{P}_{12} \\
\mathbf{P}_{21} & \mathbf{P}_{22}
\end{array}\right)
$$

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

$$
\left(\begin{array}{ll}
\mathbf{P}_{11} & \mathbf{P}_{12} \\
\mathbf{P}_{21} & \mathbf{P}_{22}
\end{array}\right)\left(\begin{array}{ll}
\mathbf{Q}_{11} & \mathbf{Q}_{12} \\
\mathbf{Q}_{21} & \mathbf{Q}_{22}
\end{array}\right)=\left(\begin{array}{ll}
\mathbf{P}_{11} \mathbf{Q}_{11}+\mathbf{P}_{12} \mathbf{Q}_{21} & \mathbf{P}_{11} \mathbf{Q}_{12}+\mathbf{P}_{12} \mathbf{Q}_{22} \\
\mathbf{P}_{21} \mathbf{Q}_{11}+\mathbf{P}_{22} \mathbf{Q}_{21} & \mathbf{P}_{21} \mathbf{Q}_{12}+\mathbf{P}_{22} \mathbf{Q}_{22}
\end{array}\right)
$$

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

$$
\left|\begin{array}{ll}
\mathbf{P}_{11} & \mathbf{P}_{12} \\
\mathbf{P}_{21} & \mathbf{P}_{22}
\end{array}\right|=\left|\mathbf{P}_{11}\right| \cdot\left|\mathbf{P}_{22}-\mathbf{P}_{21} \mathbf{P}_{11}^{-1} \mathbf{P}_{12}\right|
$$

and the inverse is

$$
\left(\begin{array}{ll}
\mathbf{P}_{11} & \mathbf{P}_{12} \\
\mathbf{P}_{21} & \mathbf{P}_{22}
\end{array}\right)^{-1}=\left(\begin{array}{cc}
\mathbf{P}_{11}^{-1}+\mathbf{P}_{11}^{-1} \mathbf{P}_{12} \mathbf{F}^{-1} \mathbf{P}_{21} \mathbf{P}_{11}^{-1} & -\mathbf{P}_{11}^{-1} \mathbf{P}_{12} \mathbf{F}^{-1} \\
-\mathbf{F}^{-1} \mathbf{P}_{21} \mathbf{P}_{11}^{-1} & \mathbf{F}^{-1}
\end{array}\right)
$$

where $\mathbf{F}=\mathbf{P}_{22}-\mathbf{P}_{21} \mathbf{P}_{11}^{-1} \mathbf{P}_{12}$ non-singular.
The determinant of a block diagonal matrix is

$$
\left|\begin{array}{cc}
\mathbf{P}_{11} & \mathbf{O} \\
\mathbf{O} & \mathbf{P}_{22}
\end{array}\right|=\left|\mathbf{P}_{11}\right| \cdot\left|\mathbf{P}_{22}\right|
$$

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

$$
\left(\begin{array}{cc}
\mathbf{P}_{11} & \mathbf{O} \\
\mathbf{O} & \mathbf{P}_{22}
\end{array}\right)^{-1}=\left(\begin{array}{cc}
\mathbf{P}_{11}^{-1} & \mathbf{O} \\
\mathbf{O} & \mathbf{P}_{22}^{-1}
\end{array}\right) .
$$

## 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 $\mathbf{a}=\left(a_{1}, \ldots, a_{n}\right)^{\prime}:$

$$
y=f(\mathbf{x})=\mathbf{a}^{\prime} \mathbf{x}=\sum_{i=1}^{n} a_{i} x_{i}=a_{1} x_{1}+a_{2} x_{2}+\cdots+a_{n} x_{n}
$$

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:

$$
\frac{\partial y}{\partial \mathbf{x}}=\frac{\partial \mathbf{a}^{\prime} \mathbf{x}}{\partial \mathbf{x}}=\left(\begin{array}{c}
\partial y / \partial x_{1} \\
\partial y / \partial x_{2} \\
\vdots \\
\partial y / \partial x_{n}
\end{array}\right)=\left(\begin{array}{c}
a_{1} \\
a_{2} \\
\vdots \\
a_{n}
\end{array}\right)=\mathbf{a}
$$

The simultaneous equation system $\mathbf{y}=\mathbf{A x}$ can be viewed as $m$ linear functions $y_{i}=\mathbf{a}_{i}^{\prime} \mathbf{x}$ where $\mathbf{a}_{i}^{\prime}$ 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

$$
\frac{\partial y_{i}}{\partial \mathbf{x}}=\frac{\partial \mathbf{a}_{\mathbf{\prime}}^{\prime} \mathbf{x}}{\partial \mathbf{x}}=\mathbf{a}_{i}
$$

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

$$
\frac{\partial \mathbf{y}}{\partial \mathbf{x}^{\prime}}=\frac{\partial \mathbf{A} \mathbf{x}}{\partial \mathbf{x}^{\prime}}=\left(\begin{array}{c}
\partial y_{1} / \partial \mathbf{x}^{\prime} \\
\partial y_{2} / \partial \mathbf{x}^{\prime} \\
\vdots \\
\partial y_{n} / \partial \mathbf{x}^{\prime}
\end{array}\right)=\left(\begin{array}{c}
\mathbf{a}_{1}^{\prime} \\
\mathbf{a}_{2}^{\prime} \\
\vdots \\
\mathbf{a}_{n}^{\prime}
\end{array}\right)=\mathbf{A}
$$

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

$$
\frac{\partial \mathbf{y}}{\partial \mathbf{x}}=\frac{\partial \mathbf{A} \mathbf{x}}{\partial \mathbf{x}}=\mathbf{A}^{\prime}
$$

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

$$
\frac{\partial \mathbf{x}^{\prime} \mathbf{A} \mathbf{x}}{\partial \mathbf{x}}=\left(\mathbf{A}+\mathbf{A}^{\prime}\right) \mathbf{x}
$$

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

$$
\frac{\partial \mathbf{x}^{\prime} \mathbf{A} \mathbf{x}}{\partial \mathbf{x}}=2 \mathbf{A} \mathbf{x}
$$

The derivative of the quadratic form $\mathbf{x}^{\prime} \mathbf{A} \mathbf{x}$ with respect to the matrix elements $a_{i j}$ is given by

$$
\frac{\partial \mathbf{x}^{\prime} \mathbf{A} \mathbf{x}}{\partial a_{i j}}=x_{i} x_{j} .
$$

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

$$
\frac{\partial \mathbf{x}^{\prime} \mathbf{A} \mathbf{x}}{\partial \mathbf{A}}=\mathbf{x x}^{\prime}
$$

## 10 Kronecker Product

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

$$
\mathbf{A} \otimes \mathbf{B}=\left(\begin{array}{cccc}
a_{11} \mathbf{B} & a_{12} \mathbf{B} & \ldots & a_{1 n} \mathbf{B} \\
a_{21} \mathbf{B} & a_{22} \mathbf{B} & \ldots & a_{21} \mathbf{B} \\
\vdots & \vdots & \ddots & \vdots \\
a_{m 1} \mathbf{B} & a_{m 1} \mathbf{B} & \ldots & a_{m n} \mathbf{B}
\end{array}\right)
$$

The following calculation rules hold:

$$
\begin{aligned}
(\mathbf{A} \otimes \mathbf{B})+(\mathbf{C} \otimes \mathbf{B}) & =(\mathbf{A}+\mathbf{C}) \otimes \mathbf{B} \\
(\mathbf{A} \otimes \mathbf{B})+(\mathbf{A} \otimes \mathbf{C}) & =\mathbf{A} \otimes(\mathbf{B}+\mathbf{C}) \\
(\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D}) & =(\mathbf{A C}) \otimes(\mathbf{B D}) \\
(\mathbf{A} \otimes \mathbf{B})^{-1} & =\mathbf{A}^{-1} \otimes \mathbf{B}^{-1} \\
\operatorname{tr}(\mathbf{A} \otimes \mathbf{B}) & =\operatorname{tr}(\mathbf{A}) \operatorname{tr}(\mathbf{B})
\end{aligned}
$$

## References

[1] Amemiya, T., Introduction to Statistics and Econometrics, Cambridge, Massachusetts: Harvard University Press, 1994.
[2] Dhrymes, P.J., Introductory Econometrics, New York : Springer-Verlag, 1978.
[3] Greene, W.H., Econometrics, New York: Macmillan, 1997.
[4] Meyer, C.D., Matrix Analysis and Applied Linear Algebra, Philadelphia: SIAM, 2000.
[5] Strang, G., Linear Algeba and its Applications, 3rd Edition, San Diego: Harcourt Brace Jovanovich, 1986.
[6] Magnus, J.R., and H. Neudecker, Matrix Differentiation Calculus with Applications in Statistics and Econometrics, Chichester: John Wiley, 1988.

