

D

Dipartimento

S

Scienze

E

Economiche

# Quaderni di Didattica

Università  
Ca' Foscari  
Venezia

Dipartimento  
di Scienze  
Economiche

Giacomo Pasini  
Yuri Pettinicchi

Matrix Algebra and  
Vector Spaces for  
Econometrics



## Matrix Algebra and Vector Spaces for Econometrics

**Giacomo Pasini, Yuri Pettinicchi**

*Università Ca' Foscari Venezia*

### **Abstract**

This document is the result of a reorganization of lecture notes used by the authors while TAing the Econometrics course at the PhD program at the School for advanced Studies in Economics at the University of Venice. It collects a series of results in Matrix Algebra and Vector Spaces analysis useful as a background for a course in Econometric Theory at the P.h.D. level. Most of the material is taken from Appendix A of Mardia et al. (1979) and chapters 4, 5 of Noble and Daniel (1988)

### **Parole Chiave**

Matrix Algebra, Vector Spaces

:

Giacomo Pasini  
Dipartimento di Scienze Economiche  
Università Ca' Foscari Venezia  
Cannaregio 873, Fondamenta S.Giobbe  
30121 Venezia - Italia  
Telefono: (+39)041 2349228  
Fax: (+39)041 2349176  
giacomo.pasini@unive.it

*I Quaderni di Didattica sono pubblicati a cura del Dipartimento di Scienze Economiche dell'Università di Venezia. I lavori riflettono esclusivamente le opinioni degli autori e non impegnano la responsabilità del Dipartimento. I Quaderni di Didattica vogliono promuovere la circolazione di studi ancora preliminari e incompleti, per suscitare commenti critici e suggerimenti. Si richiede di tener conto della natura provvisoria dei lavori per eventuali citazioni o per ogni altro uso*

# 1 Introduction

This document is the result of a reorganization of lecture notes used by the authors while TAing the Econometrics course at the PhD program at the School for advanced Studies in Economics at the University of Venice. It collects a series of results in Matrix Algebra and Vector Spaces analysis useful as a background for a course in Econometric Theory at the P.h.D. level. Most of the material is taken from Appendix A of Mardia et al. (1979) and chapters 4, 5 of Noble and Daniel (1988)

## 2 Matrix Algebra

### 2.1 basic definitions

**Definition 1 (Matrix)** A matrix  $\mathbf{A}$  is a rectangular array of numbers. If  $\mathbf{A}$  has  $n$  rows and  $p$  columns, we say it is of order  $n \times p$

$$\mathbf{A}_{n \times p} = \begin{bmatrix} a_{11} & \dots & a_{1p} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{np} \end{bmatrix} = (a_{ij})$$

**Definition 2 (Column vector)** A matrix with column-order 1 is called column vector

$$\mathbf{a} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$$

Row-vectors are column vector transposed:  $\mathbf{a}' = (a_1 \dots a_n)$

A matrix can be written in terms of its column (row) vectors:

$$\mathbf{A} = (\mathbf{a}_{(1)}, \dots, \mathbf{a}_{(p)}) = \begin{pmatrix} \mathbf{a}'_1 \\ \vdots \\ \mathbf{a}'_n \end{pmatrix}$$

Notation: we write the  $j$ -th column of matrix  $\mathbf{A}$  as  $\mathbf{a}_{(j)}$ , the  $i$ -th row as  $\mathbf{a}_i$  (if written as a column vector), the  $ij$ -th element as  $(a_{ij})$ .

**Definition 3 (Partitioned Matrix)** A matrix written in terms of its sub-matrices is called a partitioned matrix:

$$\text{Example: } \mathbf{A}(n \times p) = \begin{bmatrix} \mathbf{A}_{11}(r \times s) & \mathbf{A}_{12}(r \times (p-s)) \\ \mathbf{A}_{21}((n-r) \times s) & \mathbf{A}_{22}((n-r) \times (p-s)) \end{bmatrix}$$

**Definition 4 (Diagonal Matrix)** Given a  $n$ -dimensional vector  $\mathbf{a}$ , a  $n \times n$  diagonal matrix  $\mathbf{B} = \text{diag}(\mathbf{a})$  is a matrix with  $b_{ii} = a_i; b_{ij} = 0 \forall i \neq j$ :

$$\mathbf{B} = \text{diag}(\mathbf{a}) = \begin{bmatrix} a_1 & \dots & 0 \\ \vdots & a_i & \vdots \\ 0 & \dots & a_n \end{bmatrix}$$

Given a matrix  $\mathbf{A}_{n \times p}$ ,  $\mathbf{B} = \text{diag}(\mathbf{A})$  is a matrix with  $b_{ii} = a_{ii}; b_{ij} = 0 \forall i \neq j$

Table 1: common types of matrices

Name	Definition	Notation
Scalar	$p = n = 1$	$a$
Column vector	$p = 1$	$\mathbf{a}$
Unit vector	$p = 1; a_i = 1$	$\mathbf{1}$ or $\mathbf{1}_n$
Square	$p = q$	
Symmetric	$a_{ij} = a_{ji}$	
Unit matrix	$a_{ij} = 1; p = n$	$\mathbf{J}_p = \mathbf{1}\mathbf{1}'$
(Upper) Triangular	$a_{ij} = 0 \quad \forall j > i$	
Null matrix	$a_{ij} = 0 \quad \forall i, j$	$\mathbf{0}$

Table 2: Basic matrix operations

Operation	Restrictions	Definitions
Addition	$\mathbf{A}, \mathbf{B}$ same order	$\mathbf{A} + \mathbf{B} = (a_{ij} + b_{ij})$
Subtraction	$\mathbf{A}, \mathbf{B}$ same order	$\mathbf{A} - \mathbf{B} = (a_{ij} - b_{ij})$
Scalar multiplication		$c\mathbf{A} = (ca_{ij})$
Inner product	$\mathbf{a}, \mathbf{b}$ same order	$\mathbf{a}'\mathbf{b} = \sum_i a_i b_i$

## 2.2 Matrix operations

**Definition 5 (Matrix multiplication)** If  $\mathbf{A}, \mathbf{B}$  are conformable, i.e. the number of columns of  $\mathbf{A}$  equals the number of rows of  $\mathbf{B}$ , then  $\mathbf{AB}$  is the matrix with at each entry  $ab_{ij}$  has the inner product of the  $i$ -th row vector of  $\mathbf{A}$  and the  $j$ -th column vector of  $\mathbf{B}$ :

$$\mathbf{AB} = (\mathbf{a}'_i \mathbf{b}_{(j)})$$

Note that in general  $\mathbf{AB} \neq \mathbf{BA}$

**Definition 6 (Transpose)**  $\mathbf{A}'$ , the transpose of  $\mathbf{A}$  is the  $(p \times n)$  matrix where  $(a'_{ij}) = (a_{ji})$ , i.e. the matrix where the  $j$ -th column corresponds to the  $j$ -th row of  $\mathbf{A}$ :  $\mathbf{A}' = (\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n)$ .

The transpose satisfies

$$(\mathbf{A}')' = \mathbf{A}; (\mathbf{A} + \mathbf{B})' = \mathbf{A}' + \mathbf{B}'; (\mathbf{AB})' = \mathbf{B}'\mathbf{A}'.$$

For a partitioned  $\mathbf{A}$

$$\mathbf{A}' = \begin{bmatrix} \mathbf{A}'_{11} & \mathbf{A}'_{21} \\ \mathbf{A}'_{12} & \mathbf{A}'_{22} \end{bmatrix}$$

If  $\mathbf{A}$  is symmetric then  $a_{ij} = a_{ji} \Leftrightarrow \mathbf{A} = \mathbf{A}'$

**Definition 7 (Trace)** If  $\mathbf{A}$  is a square matrix, then the trace function is

$$\text{tr}\mathbf{A} = \sum_i a_{ii}$$

It satisfies the following properties for  $\mathbf{A}, \mathbf{B}$  square matrices,  $\mathbf{C}, \mathbf{D}$  and  $\mathbf{D}, \mathbf{C}$  conformable, i.e.  $\mathbf{C}_{n \times p}, \mathbf{D}_{n \times p}$ , a scalar  $\alpha$  and a set of vector  $\mathbf{x}_i, i = 1, \dots, t$ :

$$\text{tr } \alpha = \alpha; \text{tr } \mathbf{A} \pm \mathbf{B} = \text{tr } \mathbf{A} \pm \text{tr } \mathbf{B}; \text{tr } \alpha \mathbf{A} = \alpha \text{tr } \mathbf{A}$$

$$\text{tr } \mathbf{CD} = \text{tr } \mathbf{DC} = \sum_i \sum_j c_{ij} d_{ji} \text{ and } \text{tr } \mathbf{CC}' = \text{tr } \mathbf{C}'\mathbf{C} = \sum_i \sum_j c_{ij}^2$$

$$\sum_i \mathbf{x}'_i \mathbf{A} \mathbf{x}_i = \text{tr } \mathbf{AT} \text{ where } \mathbf{T} = \sum_i \mathbf{x}_i \mathbf{x}'_i$$

To prove the last property note that  $\mathbf{x}'_i \mathbf{A} \mathbf{x}_i$  is a scalar and so it is  $\sum_i \mathbf{x}'_i \mathbf{A} \mathbf{x}_i$ . Therefore,

$$\text{tr } \sum_i \mathbf{x}'_i \mathbf{A} \mathbf{x}_i = \sum_i \text{tr } \mathbf{x}'_i \mathbf{A} \mathbf{x}_i = \sum_i \text{tr } \mathbf{A} \mathbf{x}_i \mathbf{x}'_i = \text{tr } \mathbf{A} \sum_i \mathbf{x}_i \mathbf{x}'_i = \text{tr } \mathbf{AT}$$

### 2.3 Determinants and their properties

**Definition 8 (Minors and Cofactors)** Given  $\mathbf{A}_{p \times p}$ ,

1. The  $ij$ -minor of  $\mathbf{A}$ ,  $M_{ij}$  is the determinant of the  $(p-1) \times (p-1)$  matrix formed by deleting the  $i$ -th row and the  $j$ -th column of  $\mathbf{A}$ .
2. The  $ij$ -cofactor of  $\mathbf{A}$ ,  $A_{ij}$ , is  $(-1)^{i+j} M_{ij}$ .

Note that  $\text{sign}(-1)^{i+j}$  forms an easy-to-remember pattern on a matrix:

$$\begin{bmatrix} + & - & + & \dots \\ - & + & - & \dots \\ + & - & + & \dots \\ \dots & \dots & \dots & \dots \end{bmatrix}$$

**Definition 9 (Determinant)**

1. The determinant of a  $1 \times 1$  matrix  $\alpha$  is  $|\alpha| = \alpha$ .
2. The determinant of a  $p \times p$  matrix  $\mathbf{A}$  is  $|\mathbf{A}| = \sum_{j=1}^p a_{1j} A_{1j}$ , i.e. the sum of the products between the entries of the first row and the corresponding cofactors.

Notation:  $|\mathbf{A}| = \det \mathbf{A}$ .

Note that the definition of determinant is given with respect to the first row for simplicity.  $|\mathbf{A}| = \sum_{j=1}^p a_{ij} A_{ij}$  is the same for any  $i$ , and it can be computed with respect to a column as well:  $|\mathbf{A}| = \sum_{i=1}^p a_{ij} A_{ij}$ .

Example:

$$\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

$$\begin{aligned} |\mathbf{A}| &= aA_{11} + bA_{12} \\ &= a|d| - b|c| \\ &= ad - bc \end{aligned}$$

**Definition 10 (Non-singular matrix)** A square matrix is non-singular if  $|\mathbf{A}| \neq 0$ ; otherwise it is singular

It can be proved that:

- If  $\mathbf{A}$  is triangular (or diagonal),  $|\mathbf{A}| = \prod_i a_{ii}$
- $|\alpha\mathbf{A}| = \alpha|\mathbf{A}|$
- $|\mathbf{AB}| = |\mathbf{A}||\mathbf{B}|$

## 2.4 Inverse and other useful matrices

**Definition 11 (Inverse)** The inverse of  $\mathbf{A}$  is the unique matrix  $\mathbf{A}^{-1}$  satisfying

$$\mathbf{AA}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$$

The inverse exists if and only if  $\mathbf{A}$  is non singular, i.e. if  $|\mathbf{A}| \neq 0$

The following properties holds:

1.  $\mathbf{A}^{-1} = \frac{1}{|\mathbf{A}|}(\mathbf{A}_{ij})'$ , where  $(\mathbf{A}_{ij})$  is the *adjoint* matrix, the matrix whose  $i, j$ -th entry is the  $i, j$ -th cofactor
2.  $(c\mathbf{A}^{-1}) = c^{-1}\mathbf{A}^{-1}$
3.  $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$
4.  $(\mathbf{A}')^{-1} = (\mathbf{A}^{-1})'$

The first property follows from the definition of determinant, the others from the definition of inverse applied to  $\mathbf{AB}$ .

- For a partitioned matrix  $\mathbf{P} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}$ ,  
 $|\mathbf{P}| = |\mathbf{A}||\mathbf{D} - \mathbf{CA}^{-1}\mathbf{B}| = |\mathbf{D}||\mathbf{A} - \mathbf{BA}^{-1}\mathbf{C}|$
- For  $\mathbf{B}(p \times n)$  and  $\mathbf{C}(n \times p)$ ,  
 $|\mathbf{A} + \mathbf{BC}| = |\mathbf{A}||\mathbf{I}_p + \mathbf{A}^{-1}\mathbf{BC}| = |\mathbf{A}||\mathbf{I}_n + \mathbf{CA}^{-1}\mathbf{B}|$

**Definition 12 (Kroneker product)** Let  $\mathbf{A}$  be a  $(n \times p)$  matrix and  $\mathbf{B}$  a  $(m \times q)$  one. Then, the Kroneker product of  $\mathbf{A}$  and  $\mathbf{B}$  is the  $(nm \times pq)$  matrix

$$\mathbf{A} \otimes \mathbf{B} = \begin{bmatrix} a_{11}\mathbf{B} & a_{12}\mathbf{B} & \dots & a_{1p}\mathbf{B} \\ a_{21}\mathbf{B} & a_{22}\mathbf{B} & \dots & a_{2p}\mathbf{B} \\ \vdots & \vdots & & \vdots \\ a_{n1}\mathbf{B} & a_{n2}\mathbf{B} & \dots & a_{np}\mathbf{B} \end{bmatrix}$$

**Definition 13 (Orthogonal matrices)** A square matrix  $\mathbf{A}$  is orthogonal if  $\mathbf{AA}' = \mathbf{I}$ .

The following properties hold:

1.  $\mathbf{A}^{-1} = \mathbf{A}'$

2.  $\mathbf{A}'\mathbf{A} = \mathbf{I}$
3.  $|\mathbf{A}| = \pm 1$
4. The sum of squares in each columns (rows) is unity whereas the sum of cross-products of the elements of any two columns (rows) is zero:  
 $\mathbf{a}'_i \mathbf{a}_i = 1; \mathbf{a}'_i \mathbf{a}_j = 0$  if  $i \neq j$
5.  $\mathbf{C} = \mathbf{A}\mathbf{B}$  is orthogonal if  $\mathbf{A}$  and  $\mathbf{B}$  are orthogonal

**Definition 14 (Quadratic form)** A quadratic form in the vector  $\mathbf{x}$  is a function of the form

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

with  $\mathbf{A}$  symmetric.

**Definition 15 (Definiteness)** A symmetric matrix  $\mathbf{A}$  is called positive definite (p.d) /positive semi-definite (p.s.d) and we write  $\mathbf{A} > 0$  /  $\mathbf{A} \geq 0$  respectively if

$$Q(\mathbf{x}) > 0 \quad \forall \mathbf{x} \neq \mathbf{0}; \quad Q(\mathbf{x}) \geq 0 \quad \forall \mathbf{x} \neq \mathbf{0}$$

Negative definite and semi-definite are similarly defined.

**Theorem 1** If  $\mathbf{A} \geq 0$  is a  $(p \times p)$  matrix, then for any  $(n \times p)$  matrix  $\mathbf{C}$ ,  $\mathbf{C}'\mathbf{A}\mathbf{C} \geq 0$ . If  $\mathbf{A} > 0$  and  $\mathbf{C}$  is non-singular (thus  $p = n$ ), then  $\mathbf{C}'\mathbf{A}\mathbf{C} > 0$ .

*Proof:*  $\mathbf{A} \geq 0$  implies that  $\forall \mathbf{x} \neq \mathbf{0}$

$$\mathbf{x}'\mathbf{C}'\mathbf{A}\mathbf{C}\mathbf{x} = (\mathbf{C}\mathbf{x})'\mathbf{A}(\mathbf{C}\mathbf{x}) \geq 0 \implies \mathbf{C}'\mathbf{A}\mathbf{C} \geq 0$$

If  $\mathbf{A} > 0$  and  $\mathbf{C}$  is non-singular then  $\mathbf{C}\mathbf{x} \neq \mathbf{0}$  and the result follows from the previous statement.

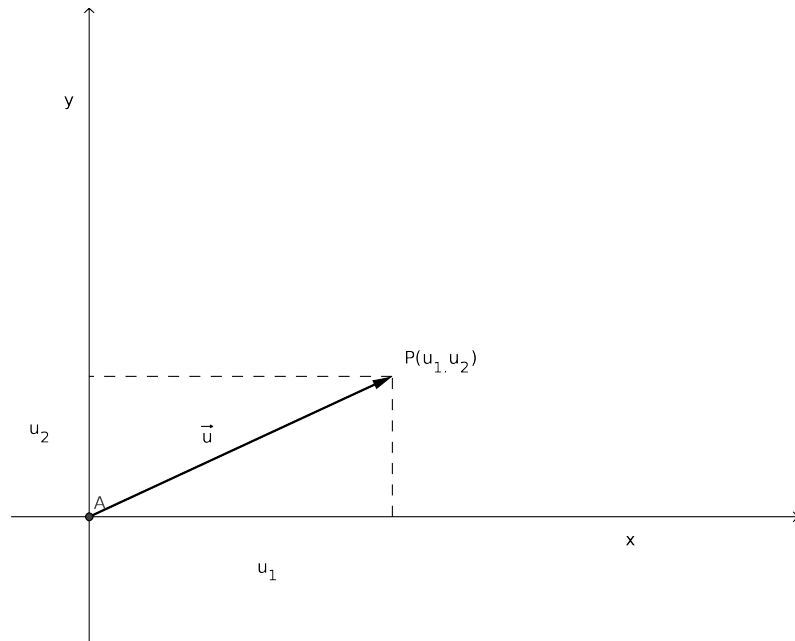
## 3 Vector Spaces

### 3.1 Geometry of $2 \times 1$ vectors

- $2 \times 1$  geometrical vector

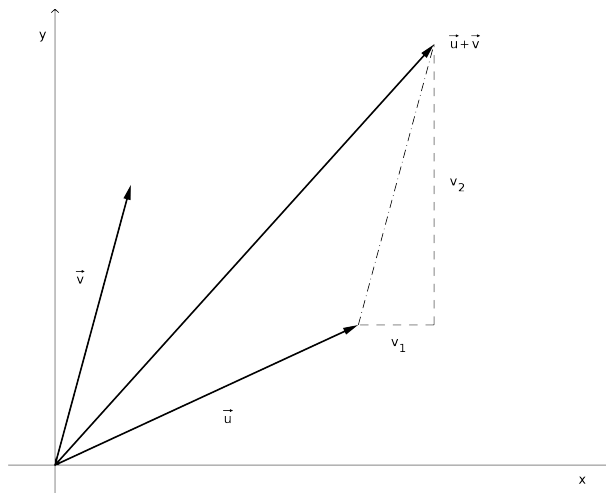
A  $2 \times 1$  vector has a natural representation on a standard x-y coordinate system as a segment from the origin to a given point  $P$ :

$$P(u_1, u_2) \longrightarrow \mathbf{u} = [u_1 \quad u_2]'$$



The sum of two 2-dimensional geometrical vectors has a natural geometrical meaning:

$$\mathbf{u} + \mathbf{v} = [u_1 \quad u_2]' + [v_1 \quad v_2]' = [u_1 + v_1 \quad u_2 + v_2]'$$



The parallelogram rule: place the tail of a vector parallel to  $\mathbf{v}$  and with  $\mathbf{v}$ 's length on the head of vector  $\mathbf{u}$ . Call the "sum" of the two vectors the new vector  $\mathbf{u} + \mathbf{v}$  connecting the tail of  $\mathbf{u}$  with the head of  $\mathbf{v}$ .

The length of a geometric vector, the shortest distance between its tail and its head, can be derived by Pythagorean theorem:

$$d = \sqrt{u_1^2 + u_2^2} = (\mathbf{u}^T \mathbf{u})^{1/2}$$

Therefore, the matrix product  $\mathbf{u}'\mathbf{v} = u_1v_1 + u_2v_2 = 0$  has a geometric meaning:  $\mathbf{u}'\mathbf{v} = 0$  implies that the two vectors are orthogonal. Since the slope of  $\mathbf{u}$  is  $u_2/u_1$ , and the slope of  $\mathbf{v}$  is  $v_2/v_1$ , it is easily to check that the two vectors are perpendicular. The product of their slopes  $(u_2/u_1)(v_2/v_1)$  equals  $-1$ . Rewriting it,  $u_1v_1 + u_2v_2 = 0$ .

### 3.2 Algebraic Structures

We start this part introducing a definition of binary operation. We see when it is internal or external with respect to a set. Then we describe the characteristics of an algebraic structure.

**Definition 16 (Binary Operation)** *Let  $A, B, C$  be sets. We define a binary operation any application  $\varphi : A \times B \rightarrow C$ .*

*If  $A = B = C$  we can say that  $\varphi : A \times A \rightarrow A$  is an internal binary operation on  $A$ .*

For algebraic structure we mean a  $n$ -tuple formed by sets and operation on themselves. The simplest one is a couple  $(X, \varphi)$  where  $X$  is a set and  $\varphi : X \times X \rightarrow X$  is an internal binary operation on  $X$

**Definition 17** *Given  $(X, \varphi)$*

1. *The operation  $\varphi$  is called associative if  $\forall x, y, z \in X, (x\varphi y)\varphi z = x\varphi(y\varphi z)$ .*
2. *The operation  $\varphi$  is called commutative if  $\forall x, y \in X, x\varphi y = y\varphi x$ .*
3. *An element  $u \in X$  is neutral with respect to the operation  $\varphi$  if  $\forall x \in X, x\varphi u = u\varphi x = x$ .*
4. *If  $(X, \varphi)$  admits the neutral element  $u$ , an element  $x \in X$  is called invertible if  $\exists x' \in X, x\varphi x' = x'\varphi x = u$ . In this case,  $x$  and  $x'$  are called opposites.*

Now, we focus on a special algebraic structure  $(X, +, *)$  formed by a set  $X$  and by two internal operations on  $X$ .

- The latter operation  $+$  is called addition, it is associative and commutative,  $(X, +)$  admits neutral element, denoted by  $0$ .
- The former operation  $*$  is called multiplication, it is associative,  $(X, *)$  admits neutral element, denoted by  $1$ .

If  $\forall x, y, z \in X, x*(y+z) = (x*y) + (x*z)$  and  $(y+z)*x = (y*x) + (z*x)$ , we can say that the multiplication is distributive with respect to the addition.

**Definition 18 (External Binary Operation)**  *$V$  is a set and  $\mathbb{K}$  is an algebraic structure  $(\mathbb{K}, +, *)$ . (i.e., the set  $\mathbb{R}$  of real numbers or the set  $\mathbb{C}$  of complex numbers). Any application  $\varphi$  such that  $\mathbb{K} \times V \rightarrow V$  is said an external operation on  $V$  with coefficients in  $\mathbb{K}$ .*

### 3.3 Vector Spaces

**Definition 19 (Vector Space)** We define a vector space on  $\mathbb{K}$  an algebraic structure  $\mathcal{V} = (\mathbb{K}, \mathbf{V}, \oplus, \odot)$ , formed by an algebraic structure  $\mathbb{K} = (\mathbb{K}, +, *)$  with 0 and 1 as respective neutrals, a set  $V$ , an internal operation  $\oplus$  on  $V$ :

$$\oplus : \mathbf{V} \times \mathbf{V} \rightarrow \mathbf{V}$$

and an external operation  $\odot$  on  $V$  with coefficients in  $\mathbb{K}$ :

$$\odot : \mathbb{K} \times V \rightarrow V$$

satisfying the following axioms:

*SV1* The algebraic structure  $(\mathbf{V}, \oplus)$  is commutative, associative, admits the neutral element and the opposite for each own element.

*SV2*  $\forall \alpha$  and  $\forall \beta \in \mathbb{K}$  and  $\mathbf{v}, \mathbf{w} \in \mathbf{V}$

- (i)  $(\alpha + \beta) \odot \mathbf{v} = (\alpha \odot \mathbf{v}) \oplus (\beta \odot \mathbf{v})$
- (ii)  $\alpha \odot (\mathbf{v} \oplus \mathbf{w}) = (\alpha \odot \mathbf{v}) \oplus (\alpha \odot \mathbf{w})$
- (iii)  $\alpha \odot (\beta \odot \mathbf{v}) = (\alpha * \beta) \odot \mathbf{v}$
- (iv)  $1 \odot \mathbf{v} = \mathbf{v}$

If  $\mathbb{K}$  is  $\mathbb{R}$ , then we have a real vector space. If  $\mathbb{K}$  is  $\mathbb{C}$ , then we have a complex vector space.

The elements of set  $\mathbf{V}$  are called *vectors* and the elements of  $\mathbb{K}$  are called *scalars*. Operation  $\oplus$  is said vector addition. Hereafter we denote it only with  $+$ .

Operation  $\odot$  is said multiplication by a scalar. Hereafter we denote it only with  $*$ . We often omit it. It is up to the reader to understand when they mean vector operation or when they mean internal operation in  $\mathbb{K}$ .

The unique neutral element of  $(\mathbf{V}, +)$  is called null vector and denoted by  $\mathbf{0}$ . The unique opposite vector of a  $\mathbf{v} \in \mathbf{V}$  is denoted by  $-\mathbf{v}$ .

**Proposition 1** Let  $\mathcal{V}$  be a vector space on  $\mathbb{K}$ . For any  $\alpha$  and  $\beta \in \mathbb{K}$  and for any  $\mathbf{v} \in \mathbf{V}$ , we have:

- (i)  $\alpha \mathbf{v} = \mathbf{0}$  if and only if  $\alpha = 0$  or  $\mathbf{v} = \mathbf{0}$ ;
- (ii)  $(-\alpha)\mathbf{v} = \alpha(-\mathbf{v}) = -(\alpha\mathbf{v})$ ;
- (iii) if  $\alpha\mathbf{v} = \beta\mathbf{v}$  and  $\mathbf{v} \neq \mathbf{0}$ , then  $\alpha = \beta$
- (iv)  $1 \odot \mathbf{v} = \mathbf{v}$

- $\mathcal{G}^3$ , the real vector space of all geometrical vectors in three-dimensional physical space.
- $\mathbb{R}^p$  ( $\mathbb{C}^p$ ), the real (complex) vector space of all real (complex)  $p \times 1$  column matrices.

- Suppose that  $\mathcal{V}$  and  $\mathcal{W}$  are both real or both complex vector spaces. The product space  $\mathcal{V} \times \mathcal{W}$  is the vector space of ordered pairs  $(\mathbf{v}, \mathbf{w})$  with  $\mathbf{v}$  in  $\mathcal{V}$  and with  $\mathbf{w}$  in  $\mathcal{W}$ , where

$$(\mathbf{v}, \mathbf{w}) + (\mathbf{v}', \mathbf{w}') = (\mathbf{v} + \mathbf{v}', \mathbf{w} + \mathbf{w}')$$

and

$$\alpha(\mathbf{v}, \mathbf{w}) = (\alpha\mathbf{v}, \alpha\mathbf{w})$$

using the same scalars as for  $\mathcal{V}$ .

**Definition 20 (Subspaces)** *Suppose that  $\mathcal{V}_0$  and  $\mathcal{V}$  are both real or both complex vector spaces, that  $\mathbf{V}_0$  is a subset of  $\mathbf{V}$ , and that the operations on elements of  $\mathbf{V}_0$  as  $\mathcal{V}_0$ -vectors are the same as the operations on them as  $\mathcal{V}$ -vectors. Then  $\mathcal{V}_0$  is said to be a subspace of  $\mathcal{V}$ .*

**Theorem 2 (Subspace Theorem)** *Suppose that  $\mathcal{V}$  is a vector space and that  $\mathbf{V}_0$  is a subset of  $\mathbf{V}$ ; define vector addition and multiplication by scalars for elements of  $\mathbf{V}_0$  exactly as in  $\mathbf{V}$ . Then  $\mathcal{V}_0$  is a subspace of  $\mathcal{V}$  if and only if the following three conditions hold:*

1.  $\mathcal{V}_0$  is nonempty.
2.  $\mathcal{V}_0$  is closed under multiplication in the sense that  $\alpha\mathbf{v}_0$  is in  $\mathcal{V}_0$ ,  $\forall \mathbf{v}_0$  in  $\mathcal{V}_0$  and all scalars  $\alpha$ .
3.  $\mathcal{V}_0$  is closed under vector addition in the sense that  $\mathbf{v}_0 + \mathbf{v}'_0$  is in  $\mathcal{V}_0$  for all vectors  $\mathbf{v}_0$  in  $\mathcal{V}_0$  and  $\mathbf{v}'_0$  in  $\mathcal{V}_0$ .

**Example 1** *Suppose that  $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_r\}$  is some nonempty set of vectors from  $\mathbf{V}$ . Define  $\mathbf{V}_0$  as the set of all linear combinations*

$$\mathbf{v}_0 = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \dots + \alpha_r\mathbf{v}_r$$

where the scalars  $\alpha_i$  are allowed to range over all arbitrary values.

Then  $\mathcal{V}_0$  is a subspace of  $\mathcal{V}$ . (Prove it using the Subspace Theorem)

### 3.4 Linear dependence and linear independence

**Definition 21** *A linear combination of the vectors  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$  is an expression of the form  $\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \dots + \alpha_n\mathbf{v}_n$  where the  $\alpha_i$  are scalars.*

**Definition 22** *A vector  $\mathbf{v}$  is said to be linearly dependent on the set of vectors  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$  if and only if can be written as some linear combination of  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ ; otherwise  $\mathbf{v}$  is said to be linearly independent of the set of vectors.*

**Definition 23** *A set  $S$  of vectors  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$  in  $\mathbf{V}$  is said to span (or generate) some subspace  $\mathcal{V}_0$  of  $\mathcal{V}$  if and only if  $S$  is a subset of  $\mathbf{V}_0$  and every vector  $\mathbf{v}_0$  in  $\mathbf{V}_0$  is linearly dependent on  $S$ ;  $S$  is said to be a spanning set or generating set for  $\mathcal{V}_0$ .*

A natural spanning set for  $\mathbb{R}^3$  is the set of three unit vectors.

A four vectors set, where at least three are linearly independent, is still a spanning set of  $\mathbb{R}^3$ .

Therefore, given a spanning set  $S$  we can always delete all the linearly dependent vectors and still obtain a spanning set of *linearly independent* vectors.

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad \mathbf{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

**Definition 24** Let  $\mathbf{L} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$  be a nonempty set of vectors.

- Suppose that

$$\alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \dots + \alpha_k \mathbf{v}_k = \mathbf{0}$$

implies that  $\alpha_1 = \alpha_2 = \dots = \alpha_k = 0$ . Then,  $\mathbf{L}$  is said to be linearly independent.

- A set that is not linearly independent is said to be linearly dependent. Equivalently,  $\mathbf{L}$  is linearly dependent if and only if there are scalars  $\alpha_1, \alpha_2, \dots, \alpha_k$  not all zero, with

$$\alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \dots + \alpha_k \mathbf{v}_k = \mathbf{0}$$

**Example 2** The set  $\{1, 2 - 3t, 4 + t\}$  is linearly dependent in the vector space  $\mathcal{P}^3$  of polynomials of degree strictly less than three. To determine this, suppose that

$$\alpha_1(1) + \alpha_2(2 - 3t) + \alpha_3(4 + t) = 0$$

that means, for all  $t$

$$(\alpha_1 + 2\alpha_2 + 4\alpha_3) + (-3\alpha_2 + \alpha_3)t = 0.$$

Then, we get

$$\begin{cases} \alpha_1 + 2\alpha_2 + 4\alpha_3 = 0 \\ -3\alpha_2 + \alpha_3 = 0 \end{cases}$$

a system of two equations for the three  $\alpha_i$ . The general solution is  $\alpha_1 = \frac{-10k}{3}, \alpha_2 = \frac{k}{3}$  where  $k = \alpha_3$ .

**Theorem 3 (Linear Independence)**

- Suppose that  $\mathbf{L} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$  with  $k \geq 2$  and with all the vectors  $\mathbf{v}_i \neq \mathbf{0}$ . Then  $\mathbf{L}$  is linearly dependent if and only if at least one of the  $\mathbf{v}_j$  is linearly dependent on the remaining vectors  $\mathbf{v}_i$  where  $i \neq j$ .
- Any set containing  $\mathbf{0}$  is linearly dependent.
- $\{\mathbf{v}\}$  is linearly independent if and only if  $\mathbf{v} \neq \mathbf{0}$ .
- Suppose that  $\mathbf{v}$  is linearly dependent on a set  $\mathbf{L} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$  and that  $\mathbf{v}_j$  is linearly dependent on the others in  $\mathbf{L}$ , namely  $\mathbf{L}'_j = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{j-1}, \mathbf{v}_{j+1}, \dots, \mathbf{v}_k\}$ . Then  $\mathbf{v}$  is linearly dependent on  $\mathbf{L}'_j$ .

### 3.5 Basis and dimension

**Definition 25 (Basis)** A basis for a vector space  $V$  is a linearly independent spanning set for  $V$ .

**Example 3** Recalling the definition of a spanning set, the vectors forming the basis should belong to the vector space.

We seek a basis for the subspace  $\mathcal{V}_0$  of  $\mathbb{R}^3$  consisting of all solutions to  $x_1 + x_2 + x_3 = 0$ . We can try with  $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ . These vectors are linearly independent and any vector in  $\mathbb{R}^3$  is a linear combination of these three. Any vector in  $\mathcal{V}_0$  is a linear combination of these three, as well. So we can say that  $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$  is a basis for  $\mathbb{R}^3$  but not for  $\mathcal{V}_0$  because they do not belong to  $\mathcal{V}_0$ .

A solution for  $x_1 + x_2 + x_3 = 0$  could be  $x_1 = \alpha$  and  $x_2 = \beta$ , so that  $x_3 = -\alpha - \beta$ . The general vector

$$\mathbf{v}_0 = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} \alpha \\ \beta \\ -\alpha - \beta \end{bmatrix} = \alpha \mathbf{v}_1 + \beta \mathbf{v}_2$$

where

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$$

form a basis for  $\mathcal{V}_0$ .

**Theorem 4 (Unique basis representation theorem)** Let  $\mathbf{B} = \{\mathbf{v}_1 \dots \mathbf{v}_r\}$  be a basis. Then the representation of each  $\mathbf{v}$  with respect to  $\mathbf{B}$  is unique:

if  $\mathbf{v} = \alpha_1 \mathbf{v}_1 + \dots + \alpha_r \mathbf{v}_r$  and  $\mathbf{v} = \alpha'_1 \mathbf{v}_1 + \dots + \alpha'_r \mathbf{v}_r$   
then  $\forall i, \alpha_i = \alpha'_i$ .

**Definition 26 (Dimension)** The number of vectors in a basis for a vector space is called dimension of the vector space.

### 3.6 Matrix rank

**Definition 27** Let  $\mathbf{A}$  be a  $p \times q$  matrix

1. The real (or complex) column space of  $\mathbf{A}$  is the subspace of  $\mathbb{R}^p$  (or  $\mathbb{C}^p$ ) that it is spanned by the set of columns of  $\mathbf{A}$ .
2. The real (or complex) row space of  $\mathbf{A}$  is the subspace of real (or complex) vector space of all real (or complex)  $1 \times q$  matrices that it is spanned by the set of the rows of  $\mathbf{A}$ .

The rank of a matrix  $\mathbf{A}$  is equal to the dimension of the row space or of the column space.

- The row space dimension is equal to the column space dimension.  
Think to a  $p \times p$  full rank matrix. This is invertible, so it is its transpose. Therefore  $\mathbf{A}'$  rank is the same as  $\mathbf{A}$  rank
- In a  $p \times q$  matrix the rank is at most  $\min p, q$   
Think to a  $3 \times 4$  matrix: If three columns are linearly independent, we have three  $3 \times 1$  vectors spanning  $\mathbb{R}^3$ , therefore the fourth is linearly dependent

### 3.7 Norms

**Definition 28** A norm (or vector norm) on  $V$  is a function that assigns to each vector  $\mathbf{v}$  in  $V$  a nonnegative real number, called the norm of  $\mathbf{v}$  and denoted by  $\|\mathbf{v}\|$  satisfying:

1.  $\|\mathbf{v}\| > 0$  for  $\mathbf{v} \neq 0$ , and  $\|\mathbf{0}\| = 0$ .
2.  $\|\alpha\mathbf{v}\| = |\alpha| \|\mathbf{v}\|$  for all scalars  $\alpha$  and vectors  $\mathbf{v}$ .
3. (the triangle inequality)  $\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$  for all vectors  $\mathbf{u}$  and  $\mathbf{v}$ .

**Definition 29** For vectors  $\mathbf{x} = [x_1 x_2 \dots x_p]^T$  in  $\mathbb{R}^p$  or  $\mathbb{C}^p$ , the norms  $\|\cdot\|_1, \|\cdot\|_2, \|\cdot\|_\infty$  (called the 1-norm, 2-norm,  $\infty$ -norm) are defined as

$$\begin{aligned} \|\mathbf{x}\|_1 &= |x_1| + |x_2| + \dots + |x_p| \\ \|\mathbf{x}\|_2 &= (|x_1|^2 + |x_2|^2 + \dots + |x_p|^2)^{1/2} \\ \|\mathbf{x}\|_\infty &= \max\{|x_1|, |x_2|, \dots, |x_p|\}. \end{aligned} \tag{1}$$

**Theorem 5 (Schwarz inequality)** Let  $\mathbf{x}$  and  $\mathbf{y}$  be  $p \times 1$  column matrices. Then

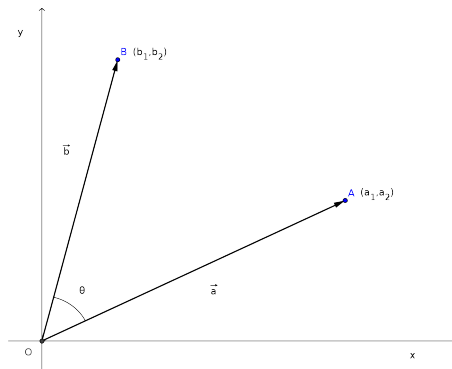
$$|\mathbf{x}^H \mathbf{y}| \leq \|\mathbf{x}\|_2 \|\mathbf{y}\|_2$$

where  $\mathbf{x}^H$  is the hermitian transpose, namely a matrix formed by the complex conjugates of the entries of the transpose matrix.

### 3.8 Inner product

**The angle between geometrical vectors**

We want to size the angle between the geometrical vectors  $\mathbf{a} = [a_1 \ a_2]'$  and  $\mathbf{b} = [b_1 \ b_2]'$



From trigonometry we have that

$$|AB|^2 = |OA|^2 + |OB|^2 - 2|OA||OB|\cos\theta$$

Using the Pythagorean Theorem we can rewrite it as

$$(a_1 - b_1)^2 + (a_2 - b_2)^2 = (a_1^2 + a_2^2) + (b_1^2 + b_2^2) - 2(a_1^2 + a_2^2)^{1/2}(b_1^2 + b_2^2)^{1/2}\cos\theta$$

and, after rearranging the terms, we get

$$(a_1b_1 + a_2b_2) = (a_1^2 + a_2^2)^{1/2}(b_1^2 + b_2^2)^{1/2}\cos\theta.$$

Defining  $(a_1b_1 + a_2b_2)$  as  $\mathbf{a}'\mathbf{b}$  and using 2-norm concept, we can compute the angle between two nonzero geometrical vectors with the following formula:

$$\cos\theta = \frac{\mathbf{a}'\mathbf{b}}{\|\mathbf{a}\|_2\|\mathbf{b}\|_2}.$$

The special product defined above ( $\mathbf{a}'\mathbf{b}$ ) could be easily extend to  $p$ -dimensional vectors.

**Definition 30 (Inner product)** Let  $\mathcal{V}$  be a real vector space. An inner product on  $\mathcal{V}$  is a function that assigns to each ordered pair of vectors  $\mathbf{u}$  and  $\mathbf{v}$  in  $\mathcal{V}$  a real number, denoted by  $\langle \mathbf{u}, \mathbf{v} \rangle$ , satisfying:

1.  $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$  for all  $\mathbf{u}$  and  $\mathbf{v}$  in  $\mathcal{V}$ .
2. for all  $\mathbf{u}, \mathbf{v}, \mathbf{w}$  in  $V$  and all real numbers  $\alpha, \beta$   $\langle \alpha\mathbf{u} + \beta\mathbf{v}, \mathbf{w} \rangle = \alpha\langle \mathbf{u}, \mathbf{w} \rangle + \beta\langle \mathbf{v}, \mathbf{w} \rangle$  and  $\langle \mathbf{w}, \alpha\mathbf{u} + \beta\mathbf{v} \rangle = \alpha\langle \mathbf{w}, \mathbf{u} \rangle + \beta\langle \mathbf{w}, \mathbf{v} \rangle$
3.  $\langle \mathbf{u}, \mathbf{u} \rangle \geq 0$  if  $\mathbf{u} \neq 0$ ,  
and  $\langle \mathbf{u}, \mathbf{u} \rangle = 0$  if and only if  $\mathbf{u} = 0$ .

The angle between two nonzero vectors  $\mathbf{u}$  and  $\mathbf{v}$  is defined by its cosine:

$$\cos\theta = \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\langle \mathbf{u}, \mathbf{u} \rangle^{1/2}\langle \mathbf{v}, \mathbf{v} \rangle^{1/2}}.$$

**Theorem 6 (Schwarz inequality)** Let  $\langle \mathbf{u}, \mathbf{v} \rangle$  be an inner product on the real vector space  $\mathcal{V}$ . Then  $|\langle \mathbf{u}, \mathbf{v} \rangle| \leq \langle \mathbf{u}, \mathbf{u} \rangle^{1/2}\langle \mathbf{v}, \mathbf{v} \rangle^{1/2}$  for all  $\mathbf{u}, \mathbf{v} \in \mathcal{V}$ .

**Theorem 7 (Inner product norms)** Let  $\langle \mathbf{u}, \mathbf{v} \rangle$  be an inner product on the real vector space  $\mathcal{V}$ , and define  $\|\mathbf{v}\| = \langle \mathbf{v}, \mathbf{v} \rangle^{1/2}$ . Then  $\|\cdot\|$  is a norm on  $\mathcal{V}$  induced by the inner product.

**Definition 31 (Orthogonality)** Let  $\langle \cdot, \cdot \rangle$  be an inner product on  $V$  and let  $\|\cdot\|$  be its induced norm.

1. Two vectors  $\mathbf{u}$  and  $\mathbf{v}$  are said to be orthogonal if and only if  $\langle \mathbf{u}, \mathbf{v} \rangle = 0$  in that set.
2. A set of vectors is said to be orthogonal if and only if every two vectors from the set are orthogonal:  $\langle \mathbf{u}, \mathbf{v} \rangle = 0 \forall \mathbf{u} \neq \mathbf{v}$ .
3. If a nonzero vector  $\mathbf{u}$  is used to produce  $\mathbf{v} = \frac{\mathbf{u}}{\|\mathbf{u}\|}$  so that  $\|\mathbf{v}\| = 1$ , then  $\mathbf{u}$  is said to have been normalized to produce the normalized vector  $\mathbf{v}$ .
4. A set of vectors is said to be orthonormal if and only if the set is orthogonal and  $\|\mathbf{v}\| = 1$  for all  $\mathbf{v}$  in the set.

In any vector space with an inner product,  $\mathbf{0}$  is orthogonal to every vector.

## References

- Mardia, K. V., J. T. Kent, and J. M. Bibby (1979). *Multivariate analysis*. Probability and Mathematical Statistics, London: Academic Press.
- Noble, B. and J. W. Daniel (1988). *Applied Linear Algebra* (3rd ed.). Prentice Hall.

