AN INTRODUCTION

TO MATRIX ALGEBRA

WITH MATLAB

Ab Mooijaart

Matthijs Warrens

&

Eeke van der Burg

Section of Psychometrics and Research Methodology

Department of Psychology

Leiden University

The Netherlands

August, 2006

Introduction

In the bachelors / masters program for Psychology in Leiden we have an extensive program for Multivariate Analysis in general. This program consists of courses in: Multivariate Analysis, Test Theory, Multidimensional Scaling, Cluster Analysis, Computational Statistics, Structural Equations Models.

The practice in teaching these courses is that students learn how to use some statistical packages for carrying out the analyses. These packages are very welcome. However, a disadvantage is that students are not familiar with the basic principles of these techniques. This is in particular important for students who are willing to learn more than “pushing buttons”. We think of students who want to know the basic principles behind these methods and want to extend their knowledge to other new techniques. For these students Matrix Algebra is absolutely important. Therefore in this monograph we give an introduction to Matrix Algebra.

This is not a course for mathematicians and we will not give rigid proofs. Mostly we only give simple examples. However, we think that the material we present is sufficient for students who want to know the basic ideas.

In each chapter the main theory is given. Because in the courses following to this course, we will give in this course sometimes the MATLAB code of the problems. Also, in cases where it is easy to compare the solutions with SPSS, the SPSS output is given too. However, the main purpose of this monograph is to understand some basic principles of matrix algebra and not the use of computer programs.

The structure of the chapters is as follows: each chapter consists of the theory and ends with some exercises. The solutions of the exercises are given in the last part of the monograph.

Used Notation

Matrices: capitals, bold. Examples: **A**, **B**, **X**,…
Vectors: lower case, bold. Examples: **a**, **b**, **x**,…
Scalars: lower case. Examples: a, b, x,…
Random variables: Examples: A, B, X,…

If matrix operations are written, then it is assumed that these operations exist. For instance, the product of two matrices **A** and **B**, **AB**, means that the number of columns of **A** is equal to the number of rows of **B**.

Table of contents

Chapter 1

Some basics

Data can be presented in a special kind of table, a so-called matrix. Matrices are useful with linear algebra.

1.1 What a matrix is

A matrix will be denoted by capital Latin letters, like **A** or **B**, and is notated as

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & . & . & a_{1p} \\ a_{21} & a_{22} & . & . & a_{2p} \\ . & . & . & . & . \\ . & . & . & a_{ij} & . \\ . & . & . & . & . \\ a_{n1} & a_{n2} & . & . & a_{np} \end{pmatrix},$$

where a_{ij} is element in row *i* and column *j*. Mind the order of the indices. The order of the matrix is denoted as $(n \times p)$. Here, we will consider two *kinds of matrices*, namely *square* matrices, for example, (2×2), (3×3) or $(i \times i)$ in general, and *rectangular* matrices, for example, (2×4) or (3×5). Examples of square matrices are

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 3 \\ 5 & 4 & 1 \\ 3 & 6 & 2 \end{pmatrix} \text{ asymmetric matrix,} \qquad \mathbf{A} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 2 \end{pmatrix} \text{ symmetric matrix, } a_{ij} = a_{ji},$$

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 6 \end{pmatrix} \text{ diagonal matrix, } a_{ij} = 0 \text{ for } i \neq j,$$

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \text{ identity matrix,} \qquad \mathbf{A} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \text{ null matrix.}$$

The *transpose* of a matrix $\mathbf{A}$ is denoted by $\mathbf{A}'$. An example:

$$\mathbf{A} = \begin{pmatrix} 1 & 3 \\ 2 & 4 \\ 5 & 6 \end{pmatrix} \text{ and } \mathbf{A}' = \begin{pmatrix} 1 & 2 & 5 \\ 3 & 4 & 6 \end{pmatrix}.$$

Sometimes the transpose of a matrix is denoted as $\mathbf{A}^{\mathrm{T}}$. If the order of a matrix is $(n \times p)$, then the transpose has order $(p \times n)$.

Besides matrices there are *vectors* and *scalars*. Example vector:

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ . \\ x_n \end{pmatrix} \text{ column vector, and } \mathbf{x}' = \begin{pmatrix} x_1 & x_2 & . & x_n \end{pmatrix} \text{ row vector.}$$

Example scalar:

$$x = (x) \text{ scalar, number.}$$

1.2 Simple operations

Let the following matrices and vectors be defined

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix},\ \mathbf{B} = \begin{pmatrix} 4 & 3 \\ 3 & 4 \\ 5 & 6 \end{pmatrix},\ \mathbf{a} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \text{ and } \mathbf{b} = \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix}.$$

Addition of two matrices is defined if both matrices have the same number of rows and the same number of columns, hence the order of the matrices is equal:

$$\mathbf{A} + \mathbf{B}' = \begin{pmatrix} 1+4 & 2+3 & 3+5 \\ 3+3 & 2+4 & 1+6 \end{pmatrix} = \begin{pmatrix} 5 & 5 & 8 \\ 6 & 6 & 7 \end{pmatrix},$$

$\mathbf{A} + \mathbf{B} =$ not possible, the order of the matrices are not equal to each other.

For *subtraction* the order of the matrices must also be equal:

$$\mathbf{A} - \mathbf{B}' = \begin{pmatrix} 1-4 & 2-3 & 3-5 \\ 3-3 & 2-4 & 1-6 \end{pmatrix} = \begin{pmatrix} -3 & -1 & -2 \\ 0 & -2 & -5 \end{pmatrix}.$$

The *product* of a scalar and a matrix is defined as:

$$c\mathbf{A} = c\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} = \begin{pmatrix} ca_{11} & ca_{12} & ca_{13} \\ ca_{21} & ca_{22} & ca_{23} \end{pmatrix}.$$

An example; if $c = 3$ and **A** is as above, then

$$c\mathbf{A} = 3\begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 6 & 9 \\ 9 & 6 & 3 \end{pmatrix}.$$

The *product* of a row vector and a column vector, with an equal number of elements, is defined as:

$$\mathbf{a}'\mathbf{b} = \begin{pmatrix} a_1 & a_2 & . & a_p \end{pmatrix}\begin{pmatrix} b_1 \\ b_2 \\ . \\ b_p \end{pmatrix} = a_1b_1 + a_2b_2 + ... + a_pb_p$$

An example; if **a** and **b** are defined as above, then

$$\mathbf{a}'\mathbf{b} = \begin{pmatrix} 1 & 2 & 3 \end{pmatrix}\begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix} = 1\times4 + 2\times5 + 3\times6 = 4 + 10 + 18 = 32$$

The *pre multiplication* of a vector **b** by a matrix **A** is defined if the number of columns of **A** is equal to the number of rows of **b**:

$$\mathbf{Ab} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix}\begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} a_{11}b_1 + a_{12}b_2 + a_{13}b_3 \\ a_{21}b_1 + a_{22}b_2 + a_{23}b_3 \end{pmatrix}$$

There is a similar definition for **bA**, the *post multiplication* of **b** by **A** (if the number of elements in **b** equals the number of rows in **A**). An example; if **A** and **b** are defined as above, then

$$\mathbf{Ab} = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}\begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix} = \begin{pmatrix} (1\times4) + (2\times5) + (3\times6) \\ (3\times4) + (2\times5) + (1\times6) \end{pmatrix} = \begin{pmatrix} 32 \\ 28 \end{pmatrix}$$

The *pre multiplication* of a matrix **B** by a matrix **A** is defined as **AB** if the number of columns of **A** are equal to the number of rows of **B**:

$$\mathbf{AB} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{pmatrix} = \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31} & a_{11}b_{12} + a_{12}b_{22} + a_{13}b_{32} \\ a_{21}b_{11} + a_{22}b_{21} + a_{23}b_{31} & a_{21}b_{12} + a_{22}b_{22} + a_{23}b_{32} \end{pmatrix}.$$

There is a similar definition for **BA**, the *post multiplication* of **B** by **A**.
An example; if **A** and **B** are defined as above, then

$$\mathbf{AB} = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} \begin{pmatrix} 4 & 3 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} = \begin{pmatrix} (1\times4)+(2\times3)+(3\times5) & (1\times3)+(2\times4)+(3\times6) \\ (3\times4)+(2\times3)+(1\times5) & (3\times3)+(2\times4)+(1\times6) \end{pmatrix} = \begin{pmatrix} 25 & 29 \\ 23 & 23 \end{pmatrix}.$$

Let $\mathbf{C} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, an identity matrix. *Pre* (or *post*) *multiplying* a matrix with an identity matrix results in the same matrix:

$$\mathbf{AC} = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1+0+0 & 0+2+0 & 0+0+3 \\ 3+0+0 & 0+2+0 & 0+0+1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}.$$

Let $\mathbf{D} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$, a diagonal matrix. *Post multiplying* a matrix with a diagonal matrix results in a matrix in which the columns of the original matrix are multiplied with the diagonal elements of the diagonal matrix:

$$\mathbf{AD} = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} = \begin{pmatrix} 1+0+0 & 0+4+0 & 0+0+9 \\ 3+0+0 & 0+4+0 & 0+0+3 \end{pmatrix} = \begin{pmatrix} 1 & 4 & 9 \\ 3 & 4 & 3 \end{pmatrix}.$$

Pre multiplying a matrix with a diagonal matrix results in a matrix in which the rows of the original matrix are multiplied with the diagonal elements of the diagonal matrix:

$$\mathbf{DB} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} 4 & 3 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} = \begin{pmatrix} 4+0+0 & 3+0+0 \\ 0+6+0 & 0+8+0 \\ 0+0+15 & 0+0+18 \end{pmatrix} = \begin{pmatrix} 4 & 3 \\ 6 & 8 \\ 15 & 18 \end{pmatrix}.$$

In general, the following properties hold:

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

$$\mathbf{A} - \mathbf{B} = -1(\mathbf{B} - \mathbf{A})$$

$$\mathbf{AB} \neq \mathbf{BA}\ .$$

1.3 Examples and MATLAB code

To define a matrix in MATLAB one must specify all the elements per row between square brackets. The different rows need to be separated by a ';'. Of course, all rows must be of equal length. The MATLAB code to define a matrix **B**, is:

```
% Define matrix B
B = [1 2 3;5 4 1;3 6 2]
```

```
B =
   1   2   3
   5   4   1
   3   6   2
```

In a similar way vectors can be defined by:

```
% Define vectors x and y
x = [1; 2; 3; 4; 5]
y = [1 2 3 4 5]
```

```
x =
   1
   2
   3
   4
   5

y =
   1   2   3   4   5
```

The transpose of the matrix **B** can be obtained with:

```
% Transpose of B
B'
```

```
ans =
   1   5   3
   2   4   6
   3   1   2
```

Note: when the result of a statement is not assigned to a specific source, MATLAB simply refers to the obtained result as 'ans'.

If we define

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 4 & 3 & 5 \\ 3 & 4 & 6 \end{pmatrix}, \ \mathbf{y} = \begin{pmatrix} -1 \\ 0 \\ 2 \end{pmatrix} \ \text{and} \ \ c = 3$$

then, the *addition* of the matrices **A** and **B** is given by:

```
% Add A and B
C = A+B
```

```
C =
   5   5   8
   6   6   7
```

Subtraction of both matrices can be done by:

```
% Subtract A and B
D = A-B
```

```
D =
  -3  -1  -2
   0  -2  -5
```

Multiplication of the two matrices:

```
% Multiply A and B
E = A*B'
F = B'*A
```

```
E =
  25  29
  23  23
```

```
F =
  13  14  15
  15  14  13
  23  22  21
```

Multiplication of matrix **A** and vector **y** / scalar *c*:

```
% Multiply A and y or c
A*y
c*A
```

```
ans =
     5
    -1
```

```
ans =
   3  6  9
   9  6  3
```

Let *n* denote the number of rows and *p* the number of columns, then the order of matrix **A** can be determined by using the '`size(A)`' function:

```
% Order of A
[n,p] = size(A)
```

```
n =
   2

p =
   3
```

1.4 Exercises

Let the following matrices and vectors be defined:

$$\mathbf{A}=\begin{pmatrix}3&4&0\\6&4&2\end{pmatrix}\qquad \mathbf{B}=\begin{pmatrix}2&6\\1&0\\4&2\end{pmatrix}\qquad \mathbf{C}=\begin{pmatrix}3&2&1\\3&4&1\end{pmatrix}\qquad \mathbf{D}=\begin{pmatrix}3&6\\6&1\end{pmatrix}$$

$$\mathbf{X}=\begin{pmatrix}1&3\\-1&0\\6&-2\\1&-1\end{pmatrix}\qquad \mathbf{u}'=\begin{pmatrix}2&3\end{pmatrix}\qquad \mathbf{v}=\begin{pmatrix}3\\4\end{pmatrix}$$

1. Compute by hand:

a. **A+C** and **A-C**
b. $\mathbf{A}+\mathbf{B}$ and **A+B’**
c. $\mathbf{AB}$
d. $\mathbf{AC}$ and **AC’**
e. $\mathbf{u}'\mathbf{Du}$
f. $\mathbf{u}'\mathbf{v}$
g. $(\mathbf{A}+\mathbf{C})'$
h. $3\mathbf{C}$
i. $\mathbf{BA}$
j. $\mathbf{X}'\mathbf{X}$
k**. u’X** and **Xu’ ; vX** and **Xv**

2. Compute the above exercises with MATLAB and compare the results.

Chapter 2

The determinant of a matrix

The determinant is a function defined on square matrices. It is uniquely defined.

2.1 What a determinant is

Let **A** be a square matrix, then the determinant of **A** is denoted as $|\mathbf{A}|$ or det(**A**). For a (2 × 2) matrix **A**, the determinant is defined as

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}.$$

Example: let $\mathbf{A} = \begin{pmatrix} 4 & 1 \\ 1 & 2 \end{pmatrix}$, then $|\mathbf{A}| = (4\times 2) - (1\times 1) = 7.$

<u>Definition</u>: the minor of element a_{ij} is the determinant of the matrix formed by removing row *i* and column *j* of the matrix **A**.

Example: let **A** be

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix},$$

then the minor of a_{11} is $\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}$ and the minor of a_{12} is $\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix}$.
This are the determinants of two 2×2 matrices.

Example: Let **A** be

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 2 & 1 \\ 3 & 1 & 4 \end{pmatrix}$$

The minor of a_{22} is $\begin{vmatrix} 1 & 3 \\ 3 & 4 \end{vmatrix} = (1 \times 4) - (3 \times 3) = -5.$

<u>Definition</u>: The cofactor c_{ij} of element a_{ij} is $c_{ij} = (-1)^{i+j} \times$ minor(a_{ij}).

Example (see example above): $c_{22} = (-1)^{2+2} \times -5 = 1 \times -5 = -5.$

<u>Definition</u>: the determinant of **A** is:

$$|A| = a_{11}c_{11} + a_{12}c_{12} + a_{13}c_{13} \quad \textit{in terms of row } 1$$

$$= a_{i1}c_{i1} + a_{i2}c_{i2} + a_{i3}c_{i3} \quad \textit{in terms of row } i$$

$$= a_{1j}c_{1j} + a_{2j}cj + a_{3j}c_{3j} \quad \textit{in terms of column } j$$

Example (see example above):

Element	minor	cofactor	Element $\times$ cofactor
$a_{11} = 1$	$\begin{vmatrix} 2 & 1 \\ 1 & 4 \end{vmatrix} = 7$	7	7
$a_{12} = 2$	$\begin{vmatrix} 2 & 1 \\ 3 & 4 \end{vmatrix} = 5$	–5	–10
$a_{13} = 3$	$\begin{vmatrix} 2 & 2 \\ 3 & 1 \end{vmatrix} = -4$	– 4	–12
$\lvert\mathbf{A}\rvert =$			–15

Exercize: Compute **|A|** by means of the elements of column 2.

2.2 Example and MATLAB code

To obtain the determinant of a square matrix:

```
% Determinant of A
A = [1 2 3;2 2 1; 3 1 4];
det(A)
```

```
ans =
  -15
```

2.3 Exercise

Let matrix **A**, **B** and **C** be defined as

$$\mathbf{A} = \begin{pmatrix} 2 & 1 & 3 \\ 3 & 2 & 4 \\ 1 & 1 & 2 \end{pmatrix},\ \mathbf{B} = \begin{pmatrix} 2 & 1 & 3 \\ 0 & 0 & 0 \\ 1 & 1 & 2 \end{pmatrix} \text{ and } \mathbf{C} = \begin{pmatrix} 2 & 1 & 3 \\ 4 & 2 & 6 \\ 1 & 1 & 2 \end{pmatrix}.$$

1: Compute the determinant of **A, B,** and **C**, i.e. $|\mathbf{A}|$, $|\mathbf{B}|$, and $|\mathbf{C}|$. Use different rows or columns and compare the results.

2. Compute the determinant of the following 4×4 matrix

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 0 & 2 & 1 \\ 2 & 3 & 2 & 1 \\ 1 & 1 & 2 & 1 \end{pmatrix},$$

3: Check the results with MATLAB.

2.4 Properties of determinants

1: The determinant of a matrix is equal to the determinant of the transpose of the matrix.

Example: $\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$, then $\mathbf{A}' = \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix}$. It is easy to verify that $|\mathbf{A}| = |\mathbf{A}'|$.

2: If one row or column of **A** is the null vector, then $|\mathbf{A}| = 0$.

Let row/column i be consist of zero elements, then the determinant can be computed as the sum of products (see definition), where the products are the elements of row/column i times its corresponding cofactor. So all the products are zero and so the determinant is zero too.

3: $|\mathbf{AB}| = |\mathbf{A}||\mathbf{B}|$

Example: $\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ $\mathbf{B} = \begin{pmatrix} 2 & -3 \\ 4 & 5 \end{pmatrix}$

$$|\mathbf{A}| = 4 - 6 = -2$$
$$|\mathbf{B}| = 10 - (-12) = 22$$
$$\mathbf{AB} = \begin{pmatrix} 10 & 7 \\ 22 & 11 \end{pmatrix}$$

$|\mathbf{AB}| = 110 - 154 = -44.$ This is indeed $|\mathbf{A}||\mathbf{B}|$.

4: The determinant of an under- or upper triangular matrix is equal to the product of the diagonal elements.

Example: $\mathbf{A} = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{pmatrix}$.

Take row 3 and find
$$\begin{aligned} |\mathbf{A}| &= 0 \times \begin{vmatrix} 2 & 3 \\ 4 & 5 \end{vmatrix} - 0 \times \begin{vmatrix} 1 & 3 \\ 0 & 5 \end{vmatrix} + 6 \times \begin{vmatrix} 1 & 2 \\ 0 & 4 \end{vmatrix} \\ &= 0 \qquad 0 \qquad + 6(1 \times 4 - 0 \times 2) \\ &= 6 \times 4 \times 1, \end{aligned}$$

which is the product of the diagonal elements.

Notice that the diagonal matrix $\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 6 \end{pmatrix}$ has the same determinant.

5: If **B** is formed by interchanging two rows or columns of matrix **A**, then $|\mathbf{B}| = -|\mathbf{A}|$.

Example: Let $\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ then $|\mathbf{A}| = -2.$ Interchanging row 1 and 2 of **A** gives $\mathbf{B} = \begin{pmatrix} 3 & 4 \\ 1 & 2 \end{pmatrix}$ and $|\mathbf{B}| = 2,$ which is indeed $-|\mathbf{A}|$.

A more elegant proof is:
$\mathbf{B} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\mathbf{A},$ and so according to property 1 it holds $|\mathbf{B}| = \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix} |\mathbf{A}| = -|\mathbf{A}|.$

6: If **B** is formed by multiplying one row or column of **A** by a constant *k*, then $|\mathbf{B}| = k|\mathbf{A}|$.

Example: $\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ then $|\mathbf{A}| = -2.$ Multiplying row one with $k = 3,$ then

$|\mathbf{B}| = \begin{vmatrix} 3 & 6 \\ 3 & 4 \end{vmatrix} = 12 - 18 = -6$, which is indeed $3 \times |\mathbf{A}|$

Another way is

$\mathbf{B} = \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix}\mathbf{A},$ and so according to property 1 it holds $|\mathbf{B}| = \begin{vmatrix} 3 & 0 \\ 0 & 1 \end{vmatrix} |\mathbf{A}| = 3 \times |\mathbf{A}|.$

7: If **B** is formed by multiplying one row (or column) of **A** with a constant *k* and adding this to another row (or column) of **A**, then $|\mathbf{A}| = |\mathbf{B}|$.

Example:

$\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$, then $|\mathbf{A}| = -2.$ Multiplying row two with $k = 2,$ and adding this to row 1, then $|\mathbf{B}| = \begin{vmatrix} 7 & 10 \\ 3 & 4 \end{vmatrix} = 28 - 30 = -2.$

Another way is

$$\mathbf{B} = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \mathbf{A}, \text{ and so according to property 1 it holds } |\mathbf{B}| = \begin{vmatrix} 1 & 2 \\ 0 & 1 \end{vmatrix} |\mathbf{A}| = |\mathbf{A}|.$$

8: If two rows or columns of **A** are equal to each other, then the determinant of **A** is 0.

If row/column i is equal to row/column j of matrix **A**, then by applying property 7 row/column i of matrix **A** can be made a zero vector without changing the determinant of **A**. So by property 2 the determinant of **A** is zero.

Chapter 3

The inverse of a matrix

The inverse of a square matrix **A** is notated as $\mathbf{A}^{-1}$ and is defined as

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

The inverse of a matrix may be used for solving a set of linear equations.

3.1 Linear equations

Suppose we have two linear equations with two unknowns. As an example we have

$$2x_1+3x_2=5 \tag{3.1}$$
$$3x_1-6x_2=-3 \tag{3.2}$$

This set of equation can be written as $\mathbf{Ax}=\mathbf{b}$, where

$$\mathbf{A}=\begin{pmatrix}2 & 3\\ 3 & -6\end{pmatrix},\ \mathbf{b}=\begin{pmatrix}5\\ -3\end{pmatrix} \text{ and } \mathbf{x}=\begin{pmatrix}x_1\\ x_2\end{pmatrix}.$$

For a solution of this set of equations we follow the following steps:

<u>Step</u> 1: multiply (3.1) with 2; this gives

$$4x_1+6x_2=10 \tag{3.3}$$
$$3x_1-6x_2=-3 \tag{3.4}$$

<u>Step</u> 2: add (3.3) to (3.4); this gives

$$7x_1+0x_2=7 \tag{3.5}$$
$$3x_1-6x_2=-3 \tag{3.6}$$

<u>Step</u> 3: divide (3.5) by 7; this gives

$$x_1+0x_2=1 \tag{3.7}$$
$$3x_1-6x_2=-3 \tag{3.8}$$

Step 4: divide (3.8) by 3; this gives

$$x_1 = 1 \quad (3.9)$$
$$x_1 - 2x_2 = -1 \quad (3.10)$$

Step 5: subtract (3.9)-(3.10); this gives

$$x_1 = 1 \quad (3.11)$$
$$2x_2 = 2 \quad (3.12)$$

Step 6: divide (3.12) by 2; this gives

$$x_1 = 1 \quad (3.13)$$
$$x_2 = 1 \quad (3.14)$$

So the solution is $\mathbf{x} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

We do now all these steps by matrices. We first define a super matrix

$$\left(\mathbf{A} \mid \mathbf{b}\right) = \begin{pmatrix} 2 & 3 & 5 \\ 3 & -6 & -3 \end{pmatrix}.$$

We will show that all the steps above can be done by pre-multiplying this matrix by an elementary matrix (i.e. a matrix with elementary row operations: addition and multiplication of rows) (Namboodiri,1984, p.53).

Step 1:

$$\begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 2 & 3 & 5 \\ 3 & -6 & -3 \end{pmatrix} = \begin{pmatrix} 4 & 6 & 10 \\ 3 & -6 & -3 \end{pmatrix}$$

Step 2:

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 4 & 6 & 10 \\ 3 & -6 & -3 \end{pmatrix} = \begin{pmatrix} 7 & 0 & 7 \\ 3 & -6 & -6 \end{pmatrix}$$

Step 3:

$$\begin{pmatrix} 1/7 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 7 & 0 & 7 \\ 3 & -6 & -3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 \\ 3 & -6 & -3 \end{pmatrix}$$

Step 4:

$$\begin{pmatrix} 1 & 0 \\ 0 & 1/3 \end{pmatrix}\begin{pmatrix} 1 & 0 & 1 \\ 3 & -6 & -3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 \\ 1 & -2 & -1 \end{pmatrix}$$

Step 5:

$$\begin{pmatrix} 1 & 0 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 1 \\ 1 & -2 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 2 \end{pmatrix}$$

Step 6:

$$\begin{pmatrix} 1 & 0 \\ 0 & 1/2 \end{pmatrix}\begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 2 \end{pmatrix} = \left(\begin{array}{cc|c} 1 & 0 & 1 \\ 0 & 1 & 1 \end{array}\right)$$

So the third column gives the solution of **x**.

Remember: the problem was to solve $\mathbf{Ax} = \mathbf{b}$. This was done each step by pre-multiplying a matrix with an elementary matrix. In the end we used 6 elementary matrices, so we can write the procedure as:

$$E_6E_5E_4E_3E_2E_1\mathbf{Ax} = E_6E_5E_4E_3E_2E_1\mathbf{b}$$

$$\mathbf{CAx} = \mathbf{Cb}$$

$$\mathbf{I}_2\mathbf{x} = \mathbf{Cb}.$$

So $\mathbf{CA} = \mathbf{I_2}$, the identity matrix of order (2×2). The matrix **C** is called the inverse of **A**. This inverse of **A** will be noted as $\mathbf{A}^{-1}$. So the solution of *x* can be written as

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

where it holds:

$$\mathbf{A}^{-1} = E_6E_5E_4E_3E_2E_1.$$

For the example it holds:

$$\mathbf{A}^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & 1/2 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 1/3 \end{pmatrix}\begin{pmatrix} 1/7 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 2/7 & 1/7 \\ 1/7 & -2/21 \end{pmatrix}.$$

Indeed, if we pre multiply **b** with $\mathbf{A}^{-1}$, then we get:

$$\mathbf{x} = \begin{pmatrix} 2/7 & 1/7 \\ 1/7 & -2/21 \end{pmatrix}\begin{pmatrix} 5 \\ -3 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix},$$ which is indeed the solution.

3.2 Alternative

Generally, we can write

$$\mathbf{A}^{-1} = \frac{1}{|\mathbf{A}|}\left(\mathbf{A}^{c}\right)',$$

where $\mathbf{A}^{c}$ is the matrix with co-factors (see Chapter 2.1). Furthermore, the determinant of **A** must be unequal to 0.

Example: See matrix **A** above:

$$|\mathbf{A}| = \begin{vmatrix} 2 & 3 \\ 3 & -6 \end{vmatrix} = -21$$

$\mathbf{A}^{c} = \begin{pmatrix} -6 & -3 \\ -3 & 2 \end{pmatrix} = \left(\mathbf{A}^{c}\right)',$ because $\mathbf{A}^{c}$ is symmetric. So the inverse of **A** is

$$\mathbf{A}^{-1} = \frac{1}{-21}\begin{pmatrix} -6 & -3 \\ -3 & 2 \end{pmatrix} = \begin{pmatrix} 6/21 & 3/21 \\ 3/21 & -2/21 \end{pmatrix} = \begin{pmatrix} 2/7 & 1/7 \\ 1/7 & -2/21 \end{pmatrix},$$

which is indeed the inverse of **A** as shown above.

3.3 Examples and MATLAB code

To obtain the inverse of a square matrix:

```
A = [2 3;3 -6]
inv(A)
```

```
A =
    2    3
    3   -6

ans =
   0.2857    0.1429
   0.1429   -0.0952
```

Solving $\mathbf{A}\mathbf{x} = \mathbf{b}$ for **x** can be obtained with:

```
b = [5;-3]
x = inv(A)*b
```

```
b =
    5
   -3

x =
   1.0000
   1.0000
```

3.4 Exercise

1a: Let **D** be defined as in the Exercise of Chapter 1, i.e. $\mathbf{D}=\begin{pmatrix}3 & 6\\ 6 & 1\end{pmatrix}$. Compute $\mathbf{D}^{-1}$, and check that $\mathbf{D}^{-1}\mathbf{D}=\mathbf{D}\mathbf{D}^{-1}=\mathbf{I}_2$.

1b: Solve the system $\mathbf{Ax}=\mathbf{b}$, where $\mathbf{A}=\begin{pmatrix}2 & 3\\ 4 & 9\end{pmatrix}$ and $\mathbf{b}=\begin{pmatrix}1\\ 2\end{pmatrix}$.

1c: replace in 1b matrix **A** by $\mathbf{A}=\begin{pmatrix}2 & 3\\ 4 & 6\end{pmatrix}$.

2: Given matrix **A** and vector **b**:

$$\mathbf{A} = \begin{pmatrix}2 & 3 & 4\\ 0 & 2 & 1\\ 1 & 2 & 1\end{pmatrix} \text{ and } \mathbf{b}=\begin{pmatrix}1\\ 2\\ 1\end{pmatrix}.$$

Solve the vector **x** from the equations **Ax=b** by means of elementary row operations. Compute $\mathbf{A}^{-1}$ by the elementary matrices. Compute $\mathbf{A}^{-1}$ also in the alternative way and compare the results.

3: Compare the results with the MATLAB output.

3.5 Properties of the inverse

1: The inverse of a nonsingular matrix, i.e. a matrix with determinant unequal to 0, is unique.

The definition of the inverse is $\mathbf{A}^{-1}=\frac{1}{|\mathbf{A}|}\left(\mathbf{A}^{c}\right)'$. For a nonsingular matrix it holds that the determinant is not 0, so the inverse is uniquely defined.

2: The inverse of the inverse of matrix **A** is equal to matrix **A**.

Proof: $(\mathbf{A}^{-1})^{-1}\mathbf{A}^{-1}=\mathbf{I}\rightarrow(\mathbf{A}^{-1})^{-1}\mathbf{A}^{-1}\mathbf{A}=\mathbf{A}\rightarrow(\mathbf{A}^{-1})^{-1}=\mathbf{A}$

3: $\left(\mathbf{AB}\right)^{-1}=\mathbf{B}^{-1}\mathbf{A}^{-1}$.

Proof: If $\left(\mathbf{AB}\right)^{-1}=\mathbf{B}^{-1}\mathbf{A}^{-1}$. then also $\mathbf{B}^{-1}\mathbf{A}^{-1}\left(\mathbf{AB}\right)=\mathbf{I}$ and $\left(\mathbf{AB}\right)\mathbf{B}^{-1}\mathbf{A}^{-1}=\mathbf{I}$.

--- $\mathbf{B}^{-1}\mathbf{A}^{-1}\left(\mathbf{AB}\right)=\mathbf{B}^{-1}(\mathbf{A}^{-1}\mathbf{A})\mathbf{B}=\mathbf{B}^{-1}(\mathbf{I})\mathbf{B}=\mathbf{B}^{-1}\mathbf{B}=\mathbf{I}$

--- $\left(\mathbf{AB}\right)\mathbf{B}^{-1}\mathbf{A}^{-1}=\mathbf{A}(\mathbf{B}\mathbf{B}^{-1})\mathbf{A}^{-1}=\mathbf{A}(\mathbf{I})\mathbf{A}^{-1}=\mathbf{A}\mathbf{A}^{-1}=\mathbf{I}$.

4: If **A** is nonsingular, i.e. |**A**|≠0, then $\mathbf{A}^{-1}$ is also nonsingular.

$|\mathbf{A}\ \mathbf{A}^{-1}| = |\mathbf{A}||\mathbf{A}^{-1}| = |\mathbf{I}| = 1$, thus $|\mathbf{A}^{-1}| = \dfrac{1}{|\mathrm{A}|} \neq 0$

5: The inverse of the transpose of matrix **A** is equal to the transpose of the inverse of **A**.

$(\mathbf{A}')^{-1}$ is the inverse of $\mathbf{A}'$. So it must hold $\mathbf{A}'\left(\mathbf{A}^{-1}\right)' = \mathbf{I}$ and $\left(\mathbf{A}^{-1}\right)'\mathbf{A}' = \mathbf{I}$.

--- $\mathbf{A}'\left(\mathbf{A}^{-1}\right)' = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$

--- $\left(\mathbf{A}^{-1}\right)'\mathbf{A}' = \mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$.

Chapter 4

Linear (in)dependence of vectors

4.1 Linear (in)dependence

Let **a** and **b** two vectors. A linear combination of these vectors can be written as:

$$k_1\mathbf{a} + k_2\mathbf{b},$$

where k_1 and k_2 are scalars, which are not both equal to 0.

<u>Definition</u>: A set of vectors is linearly dependent if one of the vectors can be written as a linear combination of (some of) the other vectors.

Example: Suppose the following vectors are defined:

$$\mathbf{a} = \begin{pmatrix} 1 \\ 3 \\ 5 \end{pmatrix} \quad \mathbf{b} = \begin{pmatrix} 2 \\ 6 \\ 10 \end{pmatrix} \quad \mathbf{c} = \begin{pmatrix} 8 \\ 4 \\ 2 \end{pmatrix} \quad \mathbf{d} = \begin{pmatrix} 11 \\ 13 \\ 17 \end{pmatrix}$$

This set is linearly dependent because it holds

$$\mathbf{b} = 2\mathbf{a}.$$

<u>General</u>: A set of vectors is linearly dependent if a combination of vectors is equal to the null vector and not all k's are equal to 0. So dependence if

$$\mathbf{Ak} = k_1\mathbf{a}_1 + k_2\mathbf{a}_2 + \ldots + k_n\mathbf{a}_n = \mathbf{0}, \text{ and } \textbf{not all } k_i = 0.$$

A set of vectors is linearly <u>in</u>dependent if a combination of vectors is equal to the null vector and all k's are equal to 0. So independence if

$$\mathbf{Ak} = k_1\mathbf{a}_1 + k_2\mathbf{a}_2 + \ldots + k_n\mathbf{a}_n = \mathbf{0}, \text{ and } \textbf{all } k_i = 0.$$

Example 1: Are the set of vectors $\begin{pmatrix} 2 \\ 4 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ 3 \end{pmatrix}$ linearly dependent?

Investigate the set

$$k_1\begin{pmatrix} 2 \\ 4 \end{pmatrix} + k_2\begin{pmatrix} 1 \\ 3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

This set can be written as

$$\begin{pmatrix} 2k_1 \\ 4k_1 \end{pmatrix} + \begin{pmatrix} 1k_2 \\ 3k_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

or

$$2k_1 + k_2 = 0$$
$$4k_1 + 3k_2 = 0$$

So it follows

$$k_2 = -2k_1 \longrightarrow 4k_1 - 6k_1 = 0 \longrightarrow k_1 = 0 \text{ and } k_2 = 0.$$

So there is no solution of the set of equations in which one of the scalars k is unequal to 0. Conclusion: the set of vectors is independent.

Example 2: Are the set of vectors $\begin{pmatrix} 2 \\ 4 \end{pmatrix}$ and $\begin{pmatrix} 6 \\ 12 \end{pmatrix}$ linearly dependent?

Investigate the set

$$k_1 \begin{pmatrix} 2 \\ 4 \end{pmatrix} + k_2 \begin{pmatrix} 6 \\ 12 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

This set can be written as

$$\begin{pmatrix} 2k_1 \\ 4k_1 \end{pmatrix} + \begin{pmatrix} 6k_2 \\ 12k_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

or

$$2k_1 + 6k_2 = 0$$
$$4k_1 + 12k_2 = 0$$

So it follows

$$k_1 = -3k_2 \longrightarrow -12k_2 + 12k_2 = 0 \longrightarrow 0k_2 = 0.$$

So each value of k_2 suffices(e.g. k_2 =1 (and k_1. =-3). And the same holds for k_1. Conclusion: the set of vectors is dependent.

Example 3: If a_i is the null vector, then the set $(a_1 a_2 ... a_n)$ is dependent.

If a_i is the null vector, then for each scalar k_i (also non zero) it holds $k_i a_i = 0$, and so the set of vectors is dependent.

Example 4: If $a_i = a_j$, then the set $(a_1 a_2 ... a_n)$ is dependent.

If $k_i = -k_j$, then $k_i a_i + (-k_j) a_j = 0$, and so there are scalars k_i and k_j unequal to zero which holds. And so the set of vectors is dependent.

4.2 Geometry and interpretation in two dimensions

Definition: The rank of a matrix **A** is the size of the largest sub-matrix of **A**, which is not singular. Notation: r(**A**).

- Rank of a matrix can never be larger than the minimum of the number of rows or columns of a matrix. Example: if the order of a matrix is (4×6), then the rank cannot be larger than 4.

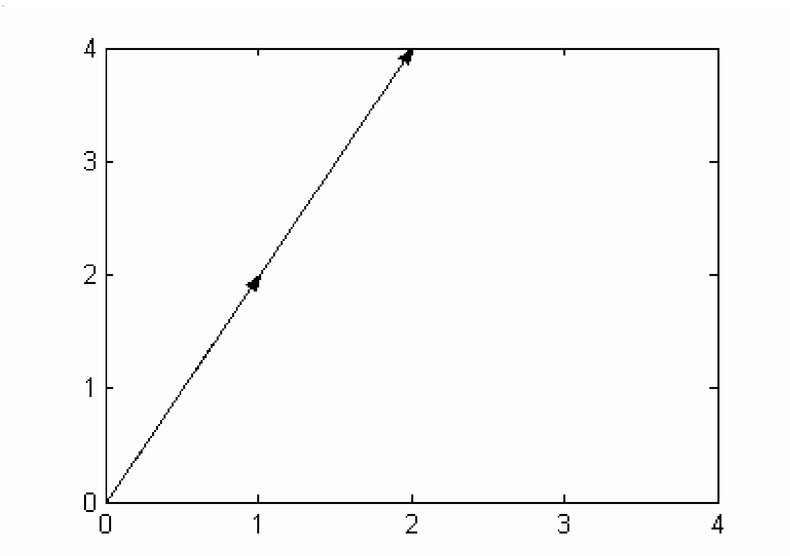
Example 5: Let the following vectors be defined

$$\mathbf{a} = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \quad \mathbf{b} = \begin{pmatrix} 2 \\ 4 \end{pmatrix}.$$

This set of vectors is dependent, because $\mathbf{b} = 2\mathbf{a}$. The matrix **A**, formed by the two vectors is

$$\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix},$$

which is rank 1 because $|\mathbf{A}| = 0$. The two vectors are presented in the figure below:



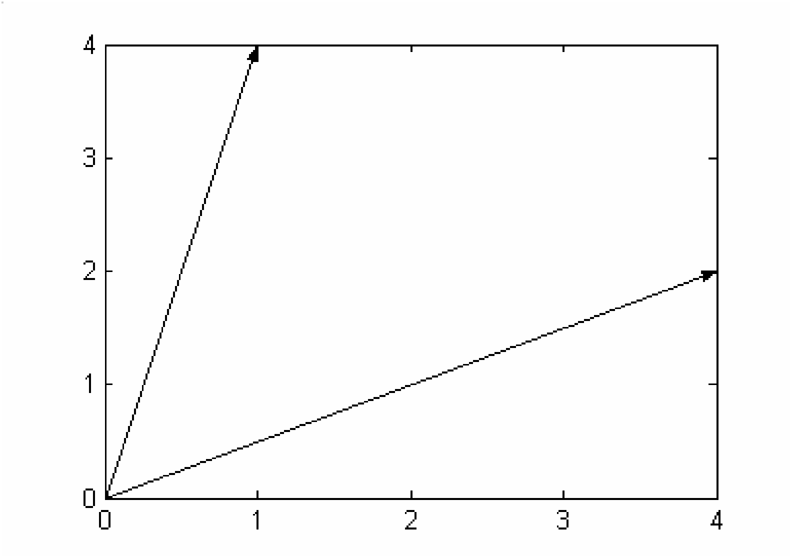
We have a different situation if we define the following vectors:

$$\mathbf{c} = \begin{pmatrix} 1 \\ 4 \end{pmatrix} \quad \mathbf{d} = \begin{pmatrix} 4 \\ 2 \end{pmatrix}.$$

The matrix **B**, formed by the two vectors is

$$\mathbf{B} = \begin{pmatrix} 1 & 4 \\ 4 & 2 \end{pmatrix},$$

which is rank 2 because |**B**|=-14. The two vectors are presented in the figure below:



Example 6: Let the following vectors be defined

$$\mathbf{a} = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} \quad \mathbf{b} = \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix} \quad \mathbf{c} = \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}$$

The matrix **A**, formed by the three vectors is

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 2 \\ 1 & 2 & 2 \\ 2 & 2 & 1 \end{pmatrix}$$

The determinant is $|\mathbf{A}| = 1(2-4) - 2(1-4) + 2(2-4) = -2 + 6 - 4 = 0.$

So

- matrix **A** is singular
- r(**A**) is smaller than 3
- the (2×2) sub matrix formed by rows 2 and 3 and columns 1 and 2 is $\begin{pmatrix} 1 & 2 \\ 2 & 2 \end{pmatrix}$, which is not singular. So r(**A**) = 2.

4.3 Consistent set of linear equations: $\mathbf{Ax} = \mathbf{b}$

Let the following set of equations be given

$$\begin{aligned} 2x_1 + x_2 \quad\quad &= 5 \quad (1) \\ 3x_1 + 6x_2 + x_3 &= 1 \quad (2) \\ 5x_1 + 7x_2 + x_3 &= 8 \quad (3) \end{aligned}$$

Then it follows

$$\begin{aligned} 2x_1 + x_2 &= 5 \quad (1) \\ 2x_1 + x_2 &= 7 \quad (3) - (2) \end{aligned}$$

Obviously there is no solution for this set of equations, i.e. the set of equations is <u>inconsistent</u>.

<u>In general</u>: A set $\mathbf{Ax} = \mathbf{b}$ is consistent if and only if $r(\mathbf{A}) = r(\mathbf{A} \mid \mathbf{b})$.

We may write the example above as

$$\mathbf{A} = \begin{pmatrix} 2 & 1 & 0 \\ 3 & 6 & 1 \\ 5 & 7 & 1 \end{pmatrix} \text{ and } \mathbf{b} = \begin{pmatrix} 5 \\ 1 \\ 8 \end{pmatrix}$$

- $|\mathbf{A}| = 2(6-7) - 1(3-5) = 0$
- $r(\mathbf{A}) \le 2$
- the sub matrix formed by rows 2 and 3 and columns 1 and 2 is $\begin{pmatrix} 3 & 6 \\ 5 & 7 \end{pmatrix}$, which is not singular. So $r(\mathbf{A}) = 2$.
- $(\mathbf{A} \mid \mathbf{b}) = \begin{pmatrix} 2 & 1 & 0 & 5 \\ 3 & 6 & 1 & 1 \\ 5 & 7 & 1 & 8 \end{pmatrix}$
- The matrix formed by columns 1,2 and 4 is

$$\begin{pmatrix} 2 & 1 & 5 \\ 3 & 6 & 1 \\ 5 & 7 & 8 \end{pmatrix}.$$

- The determinant of this matrix is $2(41) - 1(19) + 5(-9) = 82 - 19 - 45 = 18$.
- So $r(\mathbf{A} \mid \mathbf{b}) = 3$.
- So $r(\mathbf{A}) \ne r(\mathbf{A} \mid \mathbf{b})$ and the set of equations is inconsistent.

Example: let **A** be as above and let $\mathbf{b}=\begin{pmatrix}5\\1\\6\end{pmatrix}$. Investigate whether the set of equations is consistent or not.

The set of equation can be written as

$$\begin{aligned}2x_1+x_2 \quad\quad &=5 \quad (1)\\ 3x_1+6x_2+x_3 &=1 \quad (2)\\ 5x_1+7x_2+x_3 &=6 \quad (3)\end{aligned}$$

Then it follows

$$\begin{aligned}2x_1+x_2 &=5 \quad (1)\\ 2x_1+x_2 &=5 \quad (3)\text{ - }(2)\end{aligned}$$

Obviously, there is just one equation with two unknowns. So there are many solutions. For instance,

$$\begin{aligned}x_1=0 &\longrightarrow x_2=5\\ x_1=1 &\longrightarrow x_2=3\end{aligned}$$

etc.. The set of equations is now <u>consistent</u>. This can be verified by checking that $r(\mathbf{A})=2$ (as we saw before), and $r\left(\mathbf{A}\mid\mathbf{b}\right)=2,$ also.

<u>Possibilities of the solution of</u> $\mathbf{Ax}=\mathbf{b}$ (where **A** is a <u>square</u> matrix)

1: no solution: inconsistent set of equations. Here it holds: $r(\mathbf{A})\neq r\left(\mathbf{A}\mid\mathbf{b}\right)$

2: one solution: consistent set of equations, $r(\mathbf{A})=r\left(\mathbf{A}\mid\mathbf{b}\right)$ and $|\mathbf{A}|\neq 0$

3: many solutions: consistent set of equations, $r(\mathbf{A})=r\left(\mathbf{A}\mid\mathbf{b}\right)$ and |**A**|=0 (thus A is singular).

4.4 Examples and MATLAB code

Investigate the rank of the matrix **A**:

```
A =[1 2 3;1 3 2;1 1 4]
inv(A)
rank(A)
```

```
A =
    1    2    3
    1    3    2
    1    1    4

Warning: Matrix is singular to
working precision.

ans =
  Inf  Inf  Inf
  Inf  Inf  Inf
  Inf  Inf  Inf

ans =
     2
```

Next, form a set of linear equations and determine rank:

```
b1 = [1;2;0]
C = [A b1]
rank(C)
```

```
b1 =
     1
     2
     0

C =
    1    2    3    1
    1    3    2    2
    1    1    4    0

ans =
     2
```

Again, form a set of linear equations:

```
b2 = [1;2;3]
C = [A b2]
rank(C)
```

```
b2 =
     1
     2
     3

C =
    1    2    3    1
    1    3    2    2
    1    1    4    3

ans =
    3
```

4.5 Exercises

1: Are the following set of vectors dependent or independent?

1a: $\begin{pmatrix}1\\2\end{pmatrix}\ \begin{pmatrix}2\\3\end{pmatrix}\ \begin{pmatrix}3\\4\end{pmatrix}$ 1b: $\begin{pmatrix}1\\2\end{pmatrix}\ \begin{pmatrix}2\\3\end{pmatrix}\ \begin{pmatrix}0\\0\end{pmatrix}$

1c: $\begin{pmatrix}1\\2\\-1\\6\end{pmatrix}\ \begin{pmatrix}3\\8\\9\\10\end{pmatrix}\ \begin{pmatrix}2\\-1\\2\\-2\end{pmatrix}$ 1d: $\begin{pmatrix}3\\2\\1\\-4\\1\end{pmatrix}\ \begin{pmatrix}2\\3\\0\\-1\\-1\end{pmatrix}\ \begin{pmatrix}1\\-6\\3\\-8\\7\end{pmatrix}$

2: Is the vector $\begin{pmatrix}6\\10\\-2\end{pmatrix}$ a linear combination of the following vectors

$$\begin{pmatrix}1\\3\\2\end{pmatrix}\ \begin{pmatrix}2\\8\\-1\end{pmatrix}\ \begin{pmatrix}-1\\9\\2\end{pmatrix}\ ?$$

3: Is the vector $\begin{pmatrix}5\\1\\8\end{pmatrix}$ a linear combination of the following vectors

$$\begin{pmatrix}2\\3\\5\end{pmatrix}\ \begin{pmatrix}1\\6\\7\end{pmatrix}\ \begin{pmatrix}0\\1\\1\end{pmatrix}\ ?$$

4: Investigate the linear system $\mathbf{Ax} = \mathbf{b}$, that means are there zero, one or many solutions of the system?

4a: with $\mathbf{A} = \begin{pmatrix}1 & 2 & 3\\1 & 3 & 2\\1 & 1 & 4\end{pmatrix}$ and $\mathbf{b} = \begin{pmatrix}1\\2\\0\end{pmatrix}$.

4b: the same as in 4a, but now with $\mathbf{b} = \begin{pmatrix}1\\2\\3\end{pmatrix}$.

5: Are the following set of vectors dependent or independent? Determine the rank of the corresponding matrix with columns the given vectors.

5a: $\begin{pmatrix}1\\2\end{pmatrix}\ \begin{pmatrix}2\\4\end{pmatrix}$

5b: $\begin{pmatrix}1\\1\\2\end{pmatrix}\ \begin{pmatrix}2\\2\\2\end{pmatrix}\ \begin{pmatrix}0\\0\\0\end{pmatrix}$

5c: $\begin{pmatrix}1\\0\\1\end{pmatrix}\ \begin{pmatrix}1\\2\\0\end{pmatrix}\ \begin{pmatrix}0\\1\\2\end{pmatrix}$

5d: $\begin{pmatrix}1\\0\\2\\0\end{pmatrix}\ \begin{pmatrix}2\\2\\0\\1\end{pmatrix}\ \begin{pmatrix}1\\-2\\6\\-1\end{pmatrix}$

5e: $\begin{pmatrix}2\\0\\1\\0\end{pmatrix}\ \begin{pmatrix}0\\1\\2\\-1\end{pmatrix}\ \begin{pmatrix}1\\-1\\-1\\1\end{pmatrix}\ \begin{pmatrix}0\\0\\1\\2\end{pmatrix}$

5f: $\begin{pmatrix}1\\2\\1\end{pmatrix}\ \begin{pmatrix}1\\1\\0\end{pmatrix}\ \begin{pmatrix}0\\1\\-1\end{pmatrix}\ \begin{pmatrix}2\\1\\3\end{pmatrix}$

6: Is the vector $\begin{pmatrix}1\\3\end{pmatrix}$ a linear combination of the vectors $\begin{pmatrix}1\\2\end{pmatrix}$ and $\begin{pmatrix}2\\4\end{pmatrix}$?

7: The null-space of a matrix **A** is the set of all vectors **x**, for which it holds $\mathbf{Ax} = 0$.

7a: What is the null-space of $\mathbf{A} = \begin{pmatrix}1 & 2\\3 & 4\end{pmatrix}$?

7b: What is the null-space of $\mathbf{A} = \begin{pmatrix}1 & 2\\3 & 6\end{pmatrix}$?

Chapter 5

Eigenvalues and eigenvectors

5.1 Theory

A very important set of equations is

$$\mathbf{Ax} = \lambda \mathbf{x}, \tag{5.1}$$

where **A** is a square matrix. The vector **x** is called an eigenvector of **A**, and λ is called an eigenvalue of **A**. To solve **x** and λ in (5.1) we write

$$\begin{aligned} &\mathbf{Ax} = \lambda \mathbf{x} \rightarrow \mathbf{Ax} - \lambda \mathbf{x} = \mathbf{0} \rightarrow \\ &(\mathbf{A} - \lambda \mathbf{I})\mathbf{x} = \mathbf{0}. \end{aligned} \tag{5.2}$$

Obviously, one solution for **x** is the null-vector. This solution is called a trivial solution. To find non trivial solutions **x** must be in the null-space of $\mathbf{A} - \lambda \mathbf{I}$, and so

$$\left| \mathbf{A} - \lambda \mathbf{I} \right| = 0, \tag{5.3}$$

see exercise 7 of chapter 4. From (5.2) we can solve the eigenvalue(s) of matrix **A**.

Example: Let $\mathbf{A} = \begin{pmatrix} 3 & 5 \\ -2 & -4 \end{pmatrix}$,

Then $\mathbf{A} - \lambda \mathbf{I} = \begin{pmatrix} 3 & 5 \\ -2 & -4 \end{pmatrix} - \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} = \begin{pmatrix} 3-\lambda & 5 \\ -2 & -4-\lambda \end{pmatrix}$. Now λ must be solved from the equation

$$\begin{vmatrix} 3-\lambda & 5 \\ -2 & -4-\lambda \end{vmatrix} = 0.$$

So

$$(3-\lambda)(-4-\lambda)-5(-2)=0 \rightarrow$$
$$-12-3\lambda+4\lambda+\lambda^2+10=0 \rightarrow$$
$$\lambda^2+\lambda-2=0.$$

This equation is called the <u>characteristic function</u>. We see in this example that there are two possible solutions of λ.

In general, if matrix **A** is of order $(p \times p)$, then there are in principle *p* eigenvalues. These eigenvalues may all be different, or some may be equal to each other, or some may be imaginary. In the example above we find two different solutions for λ, namely λ=1 and λ=-2.

Corresponding to each eigenvalue there is an eigenvector. Such an eigenvector can be solved from (5.2). For the example above we find for the <u>eigenvalue λ=1</u>:

$$(\mathbf{A}-\lambda\mathbf{I})\mathbf{x}=\left(\begin{pmatrix}3 & 5\\ -2 & -4\end{pmatrix}-\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}\right)\begin{pmatrix}x_1\\ x_2\end{pmatrix}=\begin{pmatrix}3-1 & 5\\ -2 & -4-1\end{pmatrix}\begin{pmatrix}x_1\\ x_2\end{pmatrix}=\begin{pmatrix}0\\ 0\end{pmatrix}.$$

This gives the equations

$$2x_1+5x_2=0 \qquad (1)$$
$$-2x_1-5x_2=0 \qquad (2)$$

From (1) we find $x_1=-5/2\,x_2$. (Notice that (2) gives the same solution. Why?) From this solution we see that we do not find a unique solution for x_1 and x_2 and so we do not find an unique eigenvector **x**. This is obvious, because from (5.1) we see directly that each eigenvector can be multiplied with a constant for which the set of equations still holds. For instance, let *c* be a constant, then we can write

$$\mathbf{A}(c\mathbf{x})=\lambda(c\mathbf{x}), \qquad (5.4)$$

where now the new eigenvector is $c\mathbf{x}$. However, the equation in (5.4) is the same as in (5.1). So each eigenvector may be multiplied/divided by any constant.
It is a standard convention to choose *c* such that $\mathbf{x}'\mathbf{x}=1$.
For instance, if in the example above we choose $x_2=1 \rightarrow x_1=-5/2$. So a solution for the eigenvector is $\begin{pmatrix}-5/2\\ 1\end{pmatrix}$. Now by dividing this vector by the square root of the sum of squares of the elements we find

$$\mathbf{x}=\begin{pmatrix}-5/2\\ 1\end{pmatrix}/\sqrt{25/4+1}=\begin{pmatrix}-5/\sqrt{29}\\ 2/\sqrt{29}\end{pmatrix}.$$

It is easy to verify that now it holds $\mathbf{x}'\mathbf{x} = 1$.

For the example above we find for the <u>eigenvalue λ = -2</u> :

$$(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \left(\begin{pmatrix} 3 & 5 \\ -2 & -4 \end{pmatrix} - \begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix} \right) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 3-(-2) & 5 \\ -2 & -4-(-2) \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0.$$

This gives the equations

$$5x_1 + 5x_2 = 0 \qquad (1)$$
$$-2x_1 - 2x_2 = 0 \qquad (2)$$

From (1) we find $x_1 = -x_2$. So if $x_2 = 1 \rightarrow x_1 = -1$, and so the eigenvector is $\mathbf{x} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$.

Now if we divide this by $\sqrt{2}$, we find $\mathbf{x} = \begin{pmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}$, for which it holds $\mathbf{x}'\mathbf{x} = 1$.

5.2 Example and MATLAB code

Obtain the eigenvectors and eigenvalues of matrix A:

```
A=[3 5;-2 -4]

[X,D]=eig(A)
```

```
A =
    3    5
   -2   -4

X =                        D =
   0.9285   -0.7071            1    0
  -0.3714    0.7071            0   -2
```

5.3 Exercise

1: Compute the eigenvalues and eigenvectors of the matrix $\begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}$.

2: Do the same for the matrix $\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$.

3: Compare the solutions of 1 and 2 with the MATLAB output.

5.4 Properties of eigenvectors / eigenvalues

Because in most situations we deal with symmetric matrices, like covariance or correlation matrices, we give here some properties, which hold for symmetric matrices.

1: The sum of the eigenvalues of a matrix is equal to the sum of the diagonal elements of that matrix. (This sum of diagonal elements of a matrix is also called the trace of a matrix, often denoted as tr(**A**)).

Example: the eigenvalues of matrix $\mathbf{A}=\begin{pmatrix}1 & 2\\ 2 & 3\end{pmatrix}$ are $(2+\sqrt{5})$ and $(2-\sqrt{5})$, so the sum of the eigenvalues is 4. The trace of matrix **A** is indeed 4.

2: The product of the eigenvalues of a matrix **A** is equal to the determinant of that matrix.

Example: for the same matrix **A** it holds that the product of the eigenvalues is $(2+\sqrt{5})(2-\sqrt{5})=4-5=-1.$ The determinant of matrix **A** is also $-1.$
Consequence: if one (or more) of a matrix is (are) zero, then the determinant is zero and the matrix is singular.

3: The inner product of the eigenvectors of a matrix is equal to 0.

Example: The eigenvectors of matrix **A** are $\begin{pmatrix}1\\ \frac{1}{2}(1+\sqrt{5})\end{pmatrix}$ and $\begin{pmatrix}1\\ \frac{1}{2}(1-\sqrt{5})\end{pmatrix}$. So the inner product is $1+\frac{1}{4}(1-5)=0.$

4: If we collect the eigenvectors of matrix **A** in a matrix **X** and the eigenvalues in a diagonal matrix $\mathbf{\Lambda},$ then we can write $\mathbf{AX}=\mathbf{X\Lambda}$, where $\mathbf{X}'\mathbf{X}=\mathbf{I}.$ (Here it is assumed that an eigenvalue and the corresponding eigenvector are in the same column of matrix $\mathbf{\Lambda}$ and **X**, respectively).

Example: Let $\mathbf{A}=\begin{pmatrix}1 & 2\\ 2 & 4\end{pmatrix}$, then **X** can be written as

$\begin{pmatrix}-2/\sqrt{5} & 1/\sqrt{5}\\ 1/\sqrt{5} & 2/\sqrt{5}\end{pmatrix}$ and $\mathbf{\Lambda}=\begin{pmatrix}0 & 0\\ 0 & 5\end{pmatrix}$.

Then $\mathbf{AX} = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} \begin{pmatrix} -2/\sqrt{5} & 1/\sqrt{5} \\ 1/\sqrt{5} & 2/\sqrt{5} \end{pmatrix} = \begin{pmatrix} 0 & 5/\sqrt{5} \\ 0 & 10/\sqrt{5} \end{pmatrix}$. Indeed this is equal to

$$\mathbf{X\Lambda} = \begin{pmatrix} -2/\sqrt{5} & 1/\sqrt{5} \\ 1/\sqrt{5} & 2/\sqrt{5} \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 5 \end{pmatrix} = \begin{pmatrix} 0 & 5/\sqrt{5} \\ 0 & 10/\sqrt{5} \end{pmatrix}.$$

5: If we define for a matrix **A** the matrix of eigenvectors and eigenvalues **X** and $\mathbf{\Lambda}$ as in property 4, then we can write in general $\mathbf{A} = \mathbf{X\Lambda X}'$.

This is easy to prove from $\mathbf{AX} = \mathbf{X\Lambda}$ and the fact that $\mathbf{X}'\mathbf{X} = \mathbf{I}$. because multiplying both sides of the equation sign gives the result.

6: In general it holds for a symmetric matrix $\mathbf{A}^{p} = \mathbf{X\Lambda}^{p}\mathbf{X}'$. Remark, this equation also holds for negative values of *p*, or fractional values of *p*.

Example: Let $p = 2,$ then $\mathbf{A}^2 = (\mathbf{X\Lambda X}')(\mathbf{X\Lambda X}') = \mathbf{X\Lambda X}'\mathbf{X\Lambda X}' = \mathbf{X\Lambda}^2\mathbf{X}'$.

5.5 Singular value decomposition

Every matrix $\mathbf{B}(n \times p)$ with $n \geq p$, can be written as a product of matrices:

$$\mathbf{B}=\mathbf{K\Lambda L}' \quad \text{with } \mathbf{K}'\mathbf{K}=\mathbf{I} \text{ and } \mathbf{L}'\mathbf{L}=\mathbf{I}.$$

$\mathbf{K}(n \times p)$ are called the left eigenvectors of **B**, $\mathbf{\Lambda}(p \times p)$ (diagonal) the singular values, and $\mathbf{L}(p \times p)$ the right eigenvectors**.** The columns of **K** are also the first p eigenvectors of $\mathbf{BB}'$ and the columns of **L** are the *p* eigenvectors of $\mathbf{B}'\mathbf{B}$, and $\mathbf{\Lambda}^2$ are the (first p) eigenvalues of $\mathbf{B}'\mathbf{B}$ and $\mathbf{BB}'$, thus

$$\mathbf{B}'\mathbf{B} = \mathbf{L\Lambda}^2\mathbf{L}' \text{ and } \mathbf{BB}'=\mathbf{K\Lambda}^2\mathbf{K}'.$$

$\mathbf{B}=\mathbf{K\Lambda L}'$ is called the singular value decomposition of **B**.

Chapter 6

Application in statistics: Multiple Regression and Principal Component Analysis

In this chapter we discuss the basic principles of two important methods in statistics: Multiple Regression (MR) and Principal Component Analysis (PCA). These methods are the basic of a lot of other methods.

6.1 Multiple Regression

Suppose there are scores on *p* variables and we want to "predict" the scores on one variable. Then we can define the following vector / matrix

y : vector of order $(n \times 1)$; often named: criterion variable, dependent variable
X : matrix of order $(n \times m)$; the first column of this matrix contains only 1's, the next columns of this matrix are the scores on the independent variables, or predictors
b : vector of regression weights
e : vector of residuals

The regression equation can be written as

$$\mathbf{y} = b_1+\mathbf{x}_2 b_2+\mathbf{x}_3 b_3+\ldots+\mathbf{x}_n b_n+\mathbf{e} = \mathbf{Xb}+\mathbf{e},$$

where the vector **y** and the matrix **X** are known, and the vector **b** has to be estimated. A standard way of solving the vector **b** is by the so-called least squares method. In this method the sum of squares of the residuals is minimized. This means that we minimize

$$\sum_{i=1}^{n} e_i^2 = \mathbf{e}'\mathbf{e}.$$

So, the problem is to minimize

$$f = \mathbf{e}'\mathbf{e} = (\mathbf{y}-\mathbf{Xb})'(\mathbf{y}-\mathbf{Xb}).$$

This function can be written as

$$\begin{aligned} f &= \mathbf{y}'\mathbf{y} - \mathbf{y}'\mathbf{X}\mathbf{b} - \mathbf{b}'\mathbf{X}'\mathbf{y} + \mathbf{b}'\mathbf{X}'\mathbf{X}\mathbf{b} \\ &= \mathbf{y}'\mathbf{y} - 2(\mathbf{y}'\mathbf{X})\mathbf{b} + \mathbf{b}'(\mathbf{X}'\mathbf{X})\mathbf{b}. \end{aligned}$$

Now we write $\mathbf{a}' \equiv \mathbf{y}'\mathbf{X}$ and $\mathbf{A} \equiv \mathbf{X}'\mathbf{X}$, then we have the function to minimize

$$f = \mathbf{y}'\mathbf{y} - 2\mathbf{a}'\mathbf{b} + \mathbf{b}'\mathbf{A}\mathbf{b}.$$

One way of minimizing this function is by taking the derivatives of this function with respect to the unknown vector **b**, and equalize these derivatives to zero.

According to the derivative rules as given in the Appendix we have:

$$\partial \mathbf{a}'\mathbf{b} / \partial \mathbf{b} = \mathbf{a} \equiv \mathbf{X}'\mathbf{y}$$

and

$$\partial(\mathbf{b}'\mathbf{A}\mathbf{b}) / \partial \mathbf{b} = 2\mathbf{A}\mathbf{b} \equiv 2\mathbf{X}'\mathbf{X}\mathbf{b}.$$

So we have the equation

$$\partial f / \partial \mathbf{b} = -2\mathbf{X}'\mathbf{y} + 2\mathbf{X}'\mathbf{X}\mathbf{b} = 0.$$

From this equation it follows

$$\mathbf{X}'\mathbf{X}\mathbf{b} = \mathbf{X}'\mathbf{y}.$$

Now if we assume that the vectors of columns are a set of independent vectors, the determinant of $\mathbf{X}'\mathbf{X}$ is unequal to zero and so we can write for an estimate of **b**

$$\hat{\mathbf{b}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}.$$

6.2 Example and MATLAB code

Let matrix be $\mathbf{X} = \begin{pmatrix} 1 & .2 \\ 1 & .3 \\ 1 & .1 \\ 1 & .5 \\ 1 & .4 \end{pmatrix}$ and $\mathbf{y} = \begin{pmatrix} .3 \\ .2 \\ .1 \\ .6 \\ .4 \end{pmatrix}$.

The regression weights in the regression equation $\mathbf{y} = \mathbf{Xb} + \mathbf{e}$ can then be estimated as $\hat{\mathbf{b}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$. Note that the first regression weight in **b** is constant in all equations. This weight is called the intercept. We can verify now

$$\mathbf{X}'\mathbf{X} = \begin{pmatrix} 5 & 1.5 \\ 1.5 & .55 \end{pmatrix}.$$

Furthermore, $|\mathbf{X}'\mathbf{X}| = 2.75 - 2.25 = .5.$ So

$$(\mathbf{X}'\mathbf{X})^{-1} = \frac{1}{.5}\begin{pmatrix} .55 & -1.5 \\ -1.5 & 5 \end{pmatrix}' = 2\begin{pmatrix} .55 & -1.5 \\ -1.5 & 5 \end{pmatrix} = \begin{pmatrix} 1.1 & -3 \\ -3 & 10 \end{pmatrix}.$$

And

$$\mathbf{X}'\mathbf{y} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ .2 & .3 & .1 & .5 & .4 \end{pmatrix}\begin{pmatrix} .3 \\ .2 \\ .1 \\ .6 \\ .4 \end{pmatrix} = \begin{pmatrix} 1.6 \\ .59 \end{pmatrix}. \text{ Thus } \bar{y} = 1.6.$$

It follows

$$\hat{\mathbf{b}} = \begin{pmatrix} 1.1 & -3 \\ -3 & 10 \end{pmatrix}\begin{pmatrix} 1.6 \\ .59 \end{pmatrix} = \begin{pmatrix} -.01 \\ 1.1 \end{pmatrix}.$$

We write the estimated scores on the dependent variable **y** as

$$\hat{\mathbf{y}} = \mathbf{X}\hat{\mathbf{b}} = \begin{pmatrix} 1 & .2 \\ 1 & .3 \\ 1 & .1 \\ 1 & .5 \\ 1 & .4 \end{pmatrix}\begin{pmatrix} -.01 \\ 1.1 \end{pmatrix} = \begin{pmatrix} .21 \\ .32 \\ .10 \\ .54 \\ .43 \end{pmatrix}. \text{ Now } Xv = \begin{pmatrix} 1 & 3 \\ -1 & 0 \\ 6 & -2 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 3 \\ 4 \end{pmatrix} = \begin{pmatrix} 15 \\ -3 \\ 10 \\ -1 \end{pmatrix}$$

For each case in this system we can write

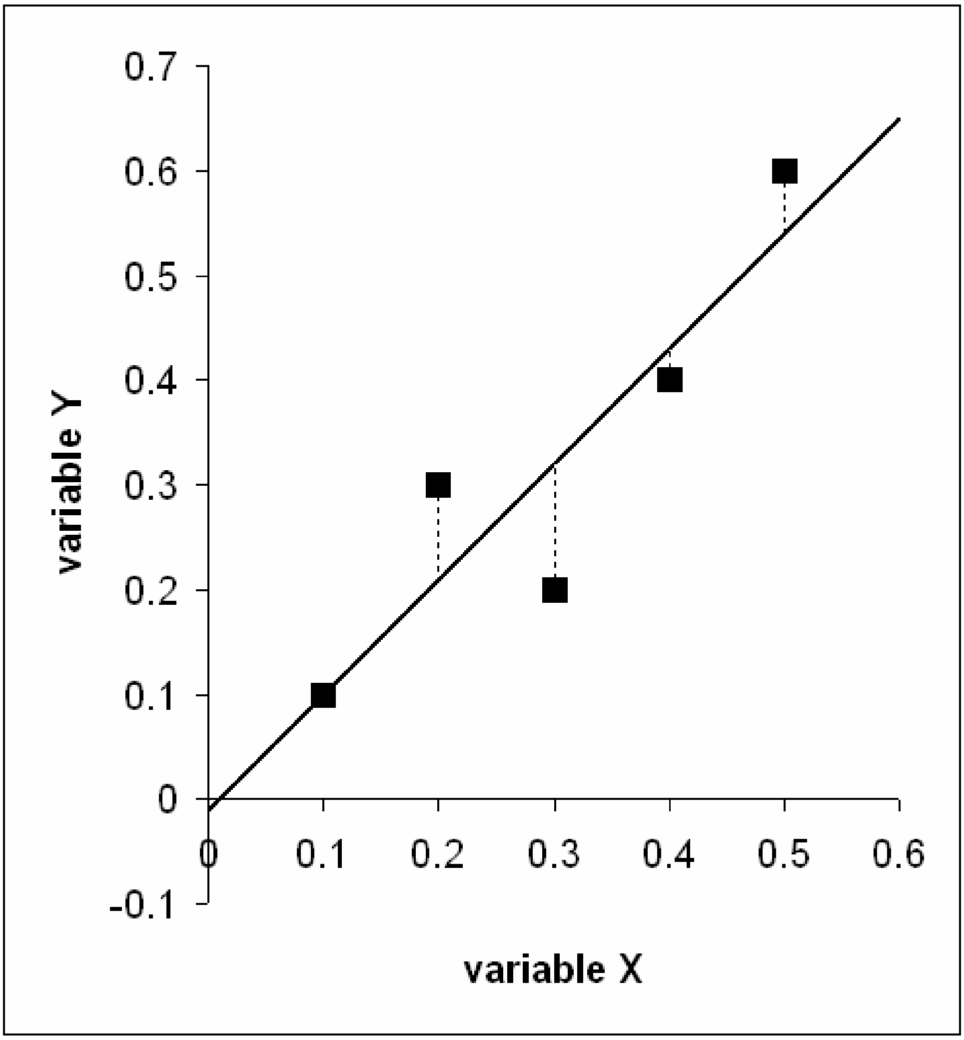
$$\hat{y}_i = \hat{b}_1 + \hat{b}_2 x_{i2},$$

where $\hat{b}_1$ is the estimate of the intercept.

For each of the cases we can write

$$\hat{y}_1 = x_{11}\hat{b}_1 + x_{12}\hat{b}_2 = \hat{b}_1 + x_{12}(1.1) = .21$$
$$\hat{y}_2 = x_{21}\hat{b}_1 + x_{22}\hat{b}_2 = \hat{b}_1 + x_{22}(1.1) = .32$$
$$\hat{y}_3 = x_{31}\hat{b}_1 + x_{32}\hat{b}_2 = \hat{b}_1 + x_{32}(1.1) = .10$$
$$\hat{y}_4 = x_{41}\hat{b}_1 + x_{42}\hat{b}_2 = \hat{b}_1 + x_{42}(1.1) = .54$$
$$\hat{y}_5 = x_{51}\hat{b}_1 + x_{52}\hat{b}_2 = \hat{b}_1 + x_{52}(1.1) = .43$$

A picture of this regression system can be written as



Remarks:

1: The difference between y_i and $\hat{y}_i$ are the errors (residuals) e_i. These are depicted in the picture as the vertical line segments. In matrix notation we have: $e = y - \hat{y}$

2: The regression is chosen in such a way that the sum of squares of the errors is minimal. So here we have for the minimum value of $f = e'e = \sum_i e_i^2$, and

$$f = \sum_i e_i^2 = (.3-.21)^2 + (.2-.32)^2 + (.1-.10)^2 + (.6-.54)^2 + (.4-.43)^2 = .027.$$

3: To compare this with the output of SPSS we find

Regression

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.904[a]	.818	.757	9.487E-02

a. Predictors: (Constant), X

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	.121	1	.121	13.444	.035[a]
	Residual	2.700E-02	3	9.000E-03		
	Total	.148	4			

a. Predictors: (Constant), X

b. Dependent Variable: Y

Remark:
SSregression $= (\hat{y}-\bar{y})'(\hat{y}-\bar{y})$; SSresidual $= e'e = (y-\hat{y})'(y-\hat{y})$; SStotal $= (y-\bar{y})'(y-\bar{y})$

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-1.00E-02	.099		-.101	.926
	X	1.100	.300	.904	3.667	.035

a. Dependent Variable: Y

4: In MATLAB, the regression weights may be obtained with:

```
X = [1 .2;1 .3;1 .1;1 .5;1 .4];
y = [.3;.2;.1;.6;.4];
b = inv(X'*X)*X'*y
```

```
b =
  -0.0100
   1.1000
```

6.3 Principal Component Analysis

Let **X** be a matrix of scores of *n* subjects on *m* variables, so **X** has the order $(n \times m)$. We want to reduce this matrix to, say, one vector **y**. So we try to find a linear combination of the columns of **X** which represents the data in some optimal way. This can be written as

$$\mathbf{Xb} = \mathbf{y}.$$

Interpretation of this model:
--- **y** is an weighted sum of variables **X** (means 0)
--- the vector **y** is unknown, in opposite to the regression model.

We have chosen here that the means of the columns of **X** are zero. So $\mathbf{1}'\mathbf{X} = 0'$. Then the sum of squares of **y** is $\mathbf{y}'\mathbf{y}$ and the variance is $\mathbf{y}'\mathbf{y}/n$.

Now we look for the weights **b** such that the variance of **y** is maximal. This means that we look for a **y** such that **y** discriminates optimally between subjects.

The function to be optimized is

$$f = \mathbf{y}'\mathbf{y} = \mathbf{b}'\mathbf{X}'\mathbf{Xb}.$$

Because this function is unbounded, we have to put a restriction on **b**. The restriction is $\mathbf{b}'\mathbf{b} = 1.$ By defining $\mathbf{A} \equiv \mathbf{X}'\mathbf{X},$ we find by the result given in the Appendix, that **b** is the eigenvector of **A** corresponding to the largest eigenvalue of **A**.

6.4 Example and MATLAB code

Let $\mathbf{X} = \begin{pmatrix} 2 & -1 \\ 0 & 2 \\ -1 & 0 \\ -1 & -1 \end{pmatrix}$. We will reduce this matrix to one column vector as discussed above.

It turned out that we have to optimize $f = \mathbf{b}'\mathbf{X}'\mathbf{Xb},$ under the restriction that $\mathbf{b}'\mathbf{b} = 1.$ Therefore we have to find the eigenvector of $\mathbf{X}'\mathbf{X}$ corresponding to the largest eigenvalue of $\mathbf{X}'\mathbf{X}$. We find

$$\mathbf{X}'\mathbf{X} = \begin{pmatrix} 2 & 0 & -1 & -1 \\ -1 & 2 & 0 & -1 \end{pmatrix} \begin{pmatrix} 2 & -1 \\ 0 & 2 \\ -1 & 0 \\ -1 & -1 \end{pmatrix} = \begin{pmatrix} 6 & -1 \\ -1 & 6 \end{pmatrix}.$$

For the eigenvalues we have to solve

$$\begin{vmatrix} 6-\lambda & -1 \\ -1 & 6-\lambda \end{vmatrix} = 36 - 12\lambda + \lambda^2 - 1 = 0.$$

This gives the solution $\lambda = 7$ and $\lambda = 5$.

For the eigenvector corresponding to the largest eigenvalue we have to solve

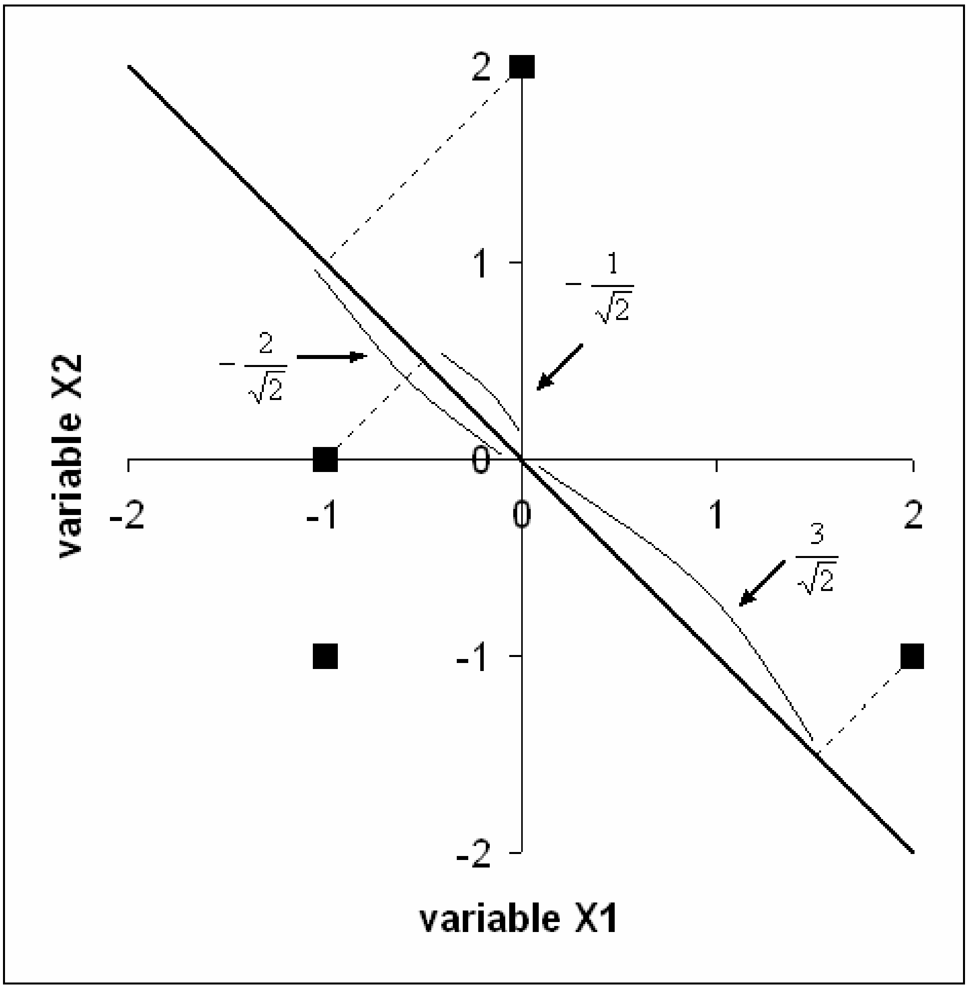
$$(6-7)b_1 - b_2 = 0 \rightarrow b_2 = -b_1.$$

So the solution of **b** is $\hat{\mathbf{b}} = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} \end{pmatrix}$. This means that the optimal **y** is

$$\mathbf{X}\hat{\mathbf{b}} = \hat{\mathbf{y}} \rightarrow \begin{pmatrix} 2 & -1 \\ 0 & 2 \\ -1 & 0 \\ -1 & -1 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} \frac{3}{\sqrt{2}} \\ \frac{-2}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} \\ 0 \end{pmatrix} . \approx \begin{pmatrix} 2.1 \\ -1.4 \\ 0.7 \\ 0.0 \end{pmatrix} \text{(see bends in picture next page)}$$

Remark 1: $\hat{\mathbf{y}}'\hat{\mathbf{y}} = \begin{pmatrix} \frac{3}{\sqrt{2}} & \frac{-2}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0 \end{pmatrix} \begin{pmatrix} \frac{3}{\sqrt{2}} \\ \frac{-2}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} \\ 0 \end{pmatrix} = 7,$ which is indeed the largest eigenvalue of $\mathbf{X}'\mathbf{X}$.

Remark 2: A picture of this PCA is.



3: In MATLAB, we may obtain the eigenvalues with:

```
X = [2 -1;0 2;-1 0;-1 -1];
eig(X'*X)
```

```
ans =
   5.0000
   7.0000
```

6.5 Exercises

1. Let the scores of 5 subjects on two independent variables be given in matrix

$$\mathbf{X} = \begin{pmatrix} 1 & 1 \\ 1 & 3 \\ 1 & 0 \\ 1 & 2 \\ 1 & 4 \end{pmatrix},\ \text{and the scores on the dependent variable be given in vector }\ \mathbf{y} = \begin{pmatrix} 2 \\ 1 \\ 4 \\ 3 \\ 5 \end{pmatrix}.$$

   1a: Carry out a multiple regression by estimating the regression weights. It holds $\hat{\mathbf{b}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$.
   1b: Make a plot of the second column of **X** and the observed and predicted **y** scores. Interpret the results; show what the residuals are and compute the sum of squares of the residuals.
   1c: Compare the results in 1a and 1b with the SPSS output.

2. The scores of 5 subjects on two variables are given in matrix $\mathbf{X} = \begin{pmatrix} 1 & -1 \\ 2 & 1 \\ 0 & 0 \\ -2 & 2 \\ -1 & -2 \end{pmatrix}$.

   2a: Carry out a Principal Component Analysis with one principal component.
   2b: Make a plot of the first principal component and the two observed variables. Interpret the results.

Literature

Namboodiri, K (1984). *Matrix Algebra: An Introduction*. London: Sage.
Healy, M.J.R. (1986). *Matrices for Statistics*. Oxford Science Publications.
Graybill, F.A. (1969). *Matrices with Applications in Statistics*. Belmont (CA), Wadsworth Company, Inc..

Appendix

1: Let $f = \mathbf{a}'\mathbf{x},$ then we can write $f = \begin{pmatrix} a_1 & a_2 & a_3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = a_1x_1 + a_2x_2 + a_3x_3.$ Now it holds

$$\begin{aligned} \partial f / \partial x_1 &= a_1 \\ \partial f / \partial x_2 &= a_2 \\ \partial f / \partial x_3 &= a_3 \end{aligned}$$

In vector/matrix notation we can write this as

$$\partial f / \partial x = \begin{pmatrix} \partial f / \partial x_1 \\ \partial f / \partial x_2 \\ \partial f / \partial x_3 \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = \mathbf{a}.$$

2: Let $f = \mathbf{x}'\mathbf{A}\mathbf{x},$ where **A** is a symmetric matrix, then we can write

$$f = \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = a_{11}x_1^2 + a_{21}x_1x_2 + a_{12}x_1x_2 + a_{22}x_2^2 = a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2,$$

because **A** is a symmetric matrix. Now it holds

$$\begin{aligned} \partial f / \partial x_1 &= 2a_{11}x_1 + 2a_{12}x_2 \\ \partial f / \partial x_2 &= 2a_{12}x_1 + 2a_{22}x_2 \end{aligned}$$

In vector/matrix notation we can write this as

$$\partial f / \partial \mathbf{x} = 2 \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} . = 2\mathbf{A}\mathbf{x}$$

3: Optimizing the function $f = \mathbf{b}'\mathbf{Ab}$ under the restriction $\mathbf{b}'\mathbf{b} = 1$ can be done by optimizing the following function (this function is called the Lagrange function):

$$f^* = \mathbf{b}'\mathbf{Ab} - \lambda(\mathbf{b}'\mathbf{b} - 1),$$

where λ is called the Lagrange multiplier ($\lambda \neq 0$).
Taking derivatives of this function with respect to **b** and λ, and equalizing to zero gives

$$\partial f^* / \partial \mathbf{b} = 2\mathbf{Ab} - 2\lambda\mathbf{b} = 0 \qquad \text{(A.1)}$$

$$\partial f^* / \partial \lambda = \mathbf{b}'\mathbf{b} - 1 = 0 \qquad \text{(A.2)}$$

From (A.1) it follows $\mathbf{Ab} = \lambda\mathbf{b}$. In addition with (A.2) it follows that **b** is an eigenvector of **A**, such that the sum of squares of **b** is equal to 1.

Because we optimize $f = \mathbf{b}'\mathbf{Ab} = \mathbf{b}'\lambda\mathbf{b} = \lambda\mathbf{b}'\mathbf{b} = \lambda,$ the optimum of *f* is equal to the largest eigenvalue of **A** and **b** is the corresponding eigenvector.

4: If we have two vectors **x** and **y**, then geometrically speaking there is an angle θ between the two vectors. It can be proven that the cosine of this angle is

$$\cos\theta = \frac{\mathbf{x}'\mathbf{y}}{\sqrt{\mathbf{x}'\mathbf{x}}\sqrt{\mathbf{y}'\mathbf{y}}}.$$

Note that $\sqrt{\mathbf{x}'\mathbf{x}}$ and $\sqrt{\mathbf{y}'\mathbf{y}}$ are the lengths of the vectors **x** and **y**, respectively.
In addition, if **x** and **y** have means zero (or $\mathbf{u}'\mathbf{x}$=0 and $\mathbf{u}'\mathbf{y}$=0), then $\cos\theta$=cor(**x,y**).
Furthermore if **x** and **y** are standardized, i.e. vectors with mean zero and variance 1 (or $\mathbf{x}'\mathbf{x}$/n=1 and $\mathbf{y}'\mathbf{y}$/n=1), then $\cos\theta$=cor(**x,y**)=$\mathbf{x}'\mathbf{y}$/n.