3 Basic Concepts from Linear algebra)

Linear algebra is an important prerequisite in order to understand the model formulation and calculations within Mixed Model. The following slides served as a brush-up on the theory, with presentation of the most important concepts and results.

Link to the full screen presentation[1]

[1] http://www.jbs.agrsci.dk/biometri/Courses/HSVmixed2001/LinAlg.f.pdf

Why Linear Algebra??

- Many statistical models used in practice are assumed to have some kind of a linear structure. (Linear regression and analysis of variance are classical examples.)
- Linear algebra is the branch of mathematics that deals with linear structures.
- Linear algebra is a convenient tool for handling models with linear structures.
- Moreover, many concepts from linear algebra can be given geometrical interpretation.

- Hence geometry can be a way to understand statistical models with linear structures

Vectors

Vectors: A column vector is a list of numbers stacked on top of each other, e.g.

$$a = \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix}$$

A row vector is a list of numbers written one after the other, e.g.

$$b = (2, 1, 3)$$

In both cases, the list is ordered, i.e.

$$(2, 1, 3) \neq (1, 2, 3).$$

- Note In what follows all vectors are column vectors unless otherwise stated.

In general an n–vector has the form

$$a = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$$

where the a_is are numbers.

Transpose of vectors: This means that a column vector is turned into a row vector and that a row vector is turned into a column vector. The transpose is denoted by "$^\top$". For example,

$$a^\top = (a_1, a_2, \dots, a_n)$$

Hence transposing twice takes us back to where we started:

$$a = (a^\top)^\top$$

- Example:

$$\begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}^\top = [1, 3, 2] \quad \text{og} \quad [1, 3, 2]^\top = \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}$$

Multiplying a vector by a number: If a is a vector and α is a number then αa is the vector

$$\alpha a = \begin{bmatrix} \alpha a_1 \\ \alpha a_2 \\ \vdots \\ \alpha a_n \end{bmatrix}$$

- Example:

$$7 \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix} = \begin{bmatrix} 7 \\ 21 \\ 14 \end{bmatrix}$$

Sum of vectors: Let a and b be n–vectors. The sum $a + b$ is the n–vector

$$a + b = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = \begin{bmatrix} a_1 + b_1 \\ a_2 + b_2 \\ \vdots \\ a_n + b_n \end{bmatrix} = b + a$$

- Note Only vectors of the same dimension can be added !
- Example:

$$\begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix} + \begin{bmatrix} 2 \\ 8 \\ 9 \end{bmatrix} = \begin{bmatrix} 1+2 \\ 3+8 \\ 2+9 \end{bmatrix} = \begin{bmatrix} 3 \\ 11 \\ 11 \end{bmatrix}$$

Inner product of vectors: Let a and b be n–vectors. The inner product $a \cdot b$ is the *number*

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

- Note The product is a number – not a vector
- Note Only vectors of the same dimension can be multiplied!
- Example:

$$\begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ 8 \\ 9 \end{bmatrix} = 1 \cdot 2 + 3 \cdot 8 + 2 \cdot 9 = 44$$

The length (norm) of a vector: The length (or norm) of a vector a is

$$||a|| = \sqrt{a \cdot a} = \sqrt{\sum_{i=1}^{n} a_i^2}$$

The 0–vector and the 1–vector: The 0-vector (1–vector) is a vector with 0 (1) on all entries. The 0–vector (1–vector) is frequently written simply as 0 (1) or as 0_n (1_n) to emphasize that it is of length n.

Orthogonal (perpendicular) vectors: Two vectors a and b with $a \neq 0$ and $b \neq 0$ are orthogonal if their inner product is zero, written

$$a \perp b \Leftrightarrow a \cdot b = 0$$

Matrices

Matrix: A matrix A with r rows og c columns is an $r \times c$ table of the form

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1c} \\ a_{21} & a_{22} & \dots & a_{2c} \\ \vdots & \vdots & \ddots & \vdots \\ a_{r1} & a_{r2} & \dots & a_{rc} \end{bmatrix}$$

It is said that A has the dimension $r \times c$.

- Note One can regard A as consisting of c columns vectors put after each other:

$$A = [a_1 : a_2 : \cdots : a_c]$$

Transpose of matrices: A matrix is transposed by interchanging rows and columns and is denoted by "$^\top$". That is,

$$A^\top = \begin{bmatrix} a_{11} & a_{21} & \dots & a_{r1} \\ a_{12} & a_{22} & \dots & a_{r2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1c} & a_{2c} & \dots & a_{rc} \end{bmatrix}$$

Example:

$$\begin{bmatrix} 1 & 2 \\ 3 & 8 \\ 2 & 9 \end{bmatrix}^\top = \begin{bmatrix} 1 & 3 & 2 \\ 2 & 8 & 9 \end{bmatrix}$$

- Note If A is an $r \times c$ matrix then $A^\top$ is a $c \times r$ matrix.
- Note One can regard a column vector of length r as an $r \times 1$ matrix and a row vector of length c as a $1 \times c$ matrix.

Multiplying a matrix with a number: For a number α and a matrix A, the product αA is the matrix

$$\alpha A = \begin{bmatrix} \alpha a_{11} & \alpha a_{12} & \dots & \alpha a_{1c} \\ \alpha a_{21} & \alpha a_{22} & \dots & \alpha a_{2c} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha a_{r1} & \alpha a_{r2} & \dots & \alpha a_{rc} \end{bmatrix}$$

Example:

$$7 \begin{bmatrix} 1 & 2 \\ 3 & 8 \\ 2 & 9 \end{bmatrix} = \begin{bmatrix} 7 & 14 \\ 21 & 56 \\ 14 & 63 \end{bmatrix}$$

Sum of matrices: Let $A = [a_1 : a_2 : \cdots : a_c]$ and $B = [b_1 : b_2 : \cdots : b_c]$ be $r \times c$ matrices.

The sum $A + B$ is the $r \times c$ matrix given by

$$\begin{aligned} A + B &= [a_1 + b_1 : a_2 + b_2 : \cdots : a_s + b_s] \\ &= \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1c} \\ a_{21} & a_{22} & \dots & a_{2c} \\ \vdots & \vdots & \ddots & \vdots \\ a_{r1} & a_{r2} & \dots & a_{rc} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1c} \\ b_{21} & b_{22} & \dots & b_{2c} \\ \vdots & \vdots & \ddots & \vdots \\ b_{r1} & b_{r2} & \dots & b_{rc} \end{bmatrix} \\ &= \begin{bmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \dots & a_{1c} + b_{1c} \\ a_{21} + b_{21} & a_{22} + b_{22} & \dots & a_{2c} + b_{2c} \\ \vdots & \vdots & \ddots & \vdots \\ a_{r1} + b_{r1} & a_{r2} + b_{r2} & \dots & a_{rc} + b_{rc} \end{bmatrix} = B + A \end{aligned}$$

- Note Only matrices with the same dimensions can be added.

Example:

$$\begin{bmatrix} 1 & 2 \\ 3 & 8 \\ 2 & 9 \end{bmatrix} + \begin{bmatrix} 5 & 4 \\ 8 & 2 \\ 3 & 7 \end{bmatrix} = \begin{bmatrix} 6 & 6 \\ 11 & 10 \\ 5 & 16 \end{bmatrix}$$

Multiplication of a matrix and a vector: Let A be an $r \times c$ matrix and let b be a c-dimensional column vector. The product Ab is the $r \times 1$ matrix

$$Ab = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1c} \\ a_{21} & a_{22} & \dots & a_{2c} \\ \vdots & \vdots & \ddots & \vdots \\ a_{r1} & a_{r2} & \dots & a_{rc} \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_c \end{bmatrix} = \begin{bmatrix} a_{11}b_1 + a_{12}b_2 + \cdots + a_{1c}b_c \\ a_{21}b_1 + a_{22}b_2 + \cdots + a_{2c}b_c \\ \vdots \\ a_{r1}b_1 + a_{r2}b_2 + \cdots + a_{rc}b_c \end{bmatrix}$$

- Eksempel:

$$\begin{bmatrix} 1 & 2 \\ 3 & 8 \\ 2 & 9 \end{bmatrix} \begin{bmatrix} 5 \\ 8 \end{bmatrix} = \begin{bmatrix} 1 \cdot 5 + 2 \cdot 8 \\ 3 \cdot 5 + 8 \cdot 8 \\ 2 \cdot 5 + 9 \cdot 8 \end{bmatrix} = \begin{bmatrix} 21 \\ 79 \\ 82 \end{bmatrix}$$

Multiplication of matrices: Let A be an $r \times c$ matrix and B a $c \times t$ matrix, i.e. $B = [b_1 : b_2 : \cdots : b_t]$. The product AB is the $r \times t$ matrix given by:

$$AB = A[b_1 : b_2 : \cdots : b_t] = [Ab_1 : Ab_2 : \cdots : Ab_t]$$

Example:

$$\begin{bmatrix} 1 & 2 \\ 3 & 8 \\ 2 & 9 \end{bmatrix} \begin{bmatrix} 5 & 4 \\ 8 & 2 \end{bmatrix} = \left[\begin{bmatrix} 1 & 2 \\ 3 & 8 \\ 2 & 9 \end{bmatrix} \begin{bmatrix} 5 \\ 8 \end{bmatrix} : \begin{bmatrix} 1 & 2 \\ 3 & 8 \\ 2 & 9 \end{bmatrix} \begin{bmatrix} 4 \\ 2 \end{bmatrix} \right]$$

$$= \begin{bmatrix} 1 \cdot 5 + 2 \cdot 8 & 1 \cdot 4 + 2 \cdot 2 \\ 3 \cdot 5 + 8 \cdot 8 & 3 \cdot 4 + 8 \cdot 2 \\ 2 \cdot 5 + 9 \cdot 8 & 2 \cdot 4 + 9 \cdot 2 \end{bmatrix} = \begin{bmatrix} 21 & 8 \\ 79 & 28 \\ 82 & 26 \end{bmatrix}$$

- Note The product AB can only be formed if the number of rows in B and the number of columns in A are the same. On that case, A and B are said to be conforme.
- Note In general AB and BA are not identical.

A mnemonic for matrix multiplication is :

$$\begin{bmatrix} 1 & 2 \\ 3 & 8 \\ 2 & 9 \end{bmatrix} \begin{bmatrix} 5 & 4 \\ 8 & 2 \end{bmatrix} = \begin{array}{cc|cc} & & 5 & 4 \\ & & 8 & 2 \\ \hline 1 & 2 & 1 \cdot 5 + 2 \cdot 8 & 1 \cdot 4 + 2 \cdot 2 \\ 3 & 8 & 3 \cdot 5 + 8 \cdot 8 & 3 \cdot 4 + 8 \cdot 2 \\ 2 & 9 & 2 \cdot 5 + 9 \cdot 8 & 2 \cdot 4 + 9 \cdot 2 \end{array} = \begin{bmatrix} 21 & 8 \\ 79 & 28 \\ 82 & 26 \end{bmatrix}$$

Special matrices:

- An $n \times n$ matrix is said to be a *square matrix*
- A matrix with 0 on all entries is the 0*–matrix* and is often written simply as 0 (or as $0_{r \times c}$ to emphasize the dimension).
- A matrix consisting of 1s in all entries is of written J (or as $J_{r \times c}$ to emphasize the dimension).
- A square matrix with 0 on all off–diagonal entries and elements $d_1, d_2, \ldots, d_n$ on the diagonal is said to be a *diagonal matrix* and is iften written $diag\{d_1, d_2, \ldots, d_n\}$
- A diagonal matrix 1s on the diagonal is called the *unity matrix* and is denoted I (or $I_{n \times n}$ to emphasize the dimension).
- A matrix A is a *symmetric matrix* $A = A^\top$.

Some rules for matrix operations: For (conformable) matrices A, B and C the following rules apply

$$(A + B)^\top = A^\top + B^\top$$

$$(AB)^\top = B^\top A^\top$$
$$A(B + C) = AB + AC$$
$$AB = AC \not\Rightarrow B = C$$

Inverse of a matrix: The inverse of an $n \times n$ matrix A is the matrix B (which is also $n \times n$) which multiplied with A gives the identity matrix I. That is,

$$AB = BA = I.$$

One says that B is A's inverse and writes $B = A^{-1}$.

- Note Only square matrices can have an inverse.
- Note Not all square matrices have an inverse.
- Note When the inverse exists, it is unique.
- Note Finding the inverse of a large matrix A is numerically complicated.

Example 1. It is easy find the inverse for a 2×2 matrix. When

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

then the inverse is

$$A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

under the assumption that $ab - bc \neq 0$. The number $ab - bc$ is called the *determinant* of A, sometimes written $\det(A)$.

If the determinant $\det(A) = 0$, then A has no inverse. fin

Example 2. Finding the inverse of a diagonal matrix is easy: Let

$$A = \begin{bmatrix} a_1 & 0 & \dots & 0 \\ 0 & a_2 & & 0 \\ \vdots & & \ddots & 0 \\ 0 & 0 & \dots & a_n \end{bmatrix}$$

where all $a_i \neq 0$. Then the inverse is

$$A^{-1} = \begin{bmatrix} \frac{1}{a_1} & 0 & \dots & 0 \\ 0 & \frac{1}{a_2} & & 0 \\ \vdots & & \ddots & 0 \\ 0 & 0 & \dots & \frac{1}{a_n} \end{bmatrix}$$

If one $a_i = 0$ then A^{-1} does not exist. fin

Generalized inverse: Not all square matrices have an inverse. However all square matrices have a *generalized inverse*.

A generalized inverse of a square matrix A is a matrix A^- satisfying that

$$AA^-A = A$$

Any square matrix has an infinite number of generalized inverses.

Linear Combinations

Let $a_1, a_2, \ldots, a_c$ be r–vectors and let $A = [a_1 : a_2 : \cdots : a_c]$ be the corresponding $r \times c$ matrix.

Let $vv = (v_1, v_2, \ldots, v_c)^\top$ be a c-vector and let

$$x \quad = \quad Av = a_1v_1 + a_2v_2 + \cdots + a_cv_c = \sum_j a_jv_j$$

Then the r–vector x is said to be a *linear combination* of $a_1, a_2, \ldots, a_c$.

Let $w = (w_1, w_2, \ldots, w_c)^\top$ be another c vector and let correspondingly $y = Aw = \sum_j a_jw_j$.

Then the following can be noted:

- For a number α the vector $\alpha x = \alpha(Av) = A(\alpha v)$ is also a linear combination of $a_1, a_2, \ldots, a_c$.
- The sum $x + y = Av + Aw = A(v + w)$ is also a linear combination of $a_1, a_2, \ldots, a_c$.
- Hence if x and y are both linear combination $a_1, a_2, \ldots, a_c$ then so is the sum $\alpha x + \beta y$ where α and β are numbers.

n–dimensional Spaces

A 2–vector $x = (x_1, x_2)$ can be regarded as the point with coordinates (x_1, x_2) in a 2–dimensional coordinate system, i.e. in the plane.

Likewise a 3–vector $x = (x_1, x_2, x_3)$ can be regarded as the point with coordinates (x_1, x_2, x_3) in a 3–dimensional coordinate system, i.e. in space.

In general an n–vector $x = (x_1, x_2, \ldots, x_n)$ can be regarded as the point with coordinates $(x_1, x_2, \ldots, x_n)$ in an n–dimensional coordinate system, i.e. in an n–dimensional space. Such as space shall here be referred to as R^n. Its hard to draw!

To justify such n–dimensional spaces, suppose x consists of a location of an object (that takes 3 coordinates), the temperature of the object (that occupies one coordinate) and the time (that also occupies one coordinate). Hence the total information about the object can be regarded as a point in a 5–dimensional space.

Note that If x and y are both vectors in R^n then so is the sum $\alpha x + \beta y$.

Linear Subspaces

Consider a set $a_1, a_2, \ldots, a_c$ of r–vectors.

We can regard these vectors as "building blocks" for creating new vectors as linear combinations of the building blocks. Any such vector is an r–vector

The set of vectors which can be created as linear combinations of the "building blocks" is called a *linear subspace* of R^r.

Such a space, let us call it L, is said to be *spanned* by $a_1, a_2, \ldots, a_c$ and we write $L = \text{span}(a_1, a_2, \ldots, a_c)$.

Example 3. Consider the vectors

$$a_1 = \begin{bmatrix} 2 \\ 6 \\ 4 \end{bmatrix} \quad , \quad a_2 = \begin{bmatrix} 1 \\ 5 \\ 7 \end{bmatrix}$$

Hence $\text{span}(a_1, a_2)$ is the set of vectors which can be written as

$$y = \begin{bmatrix} 2 \\ 6 \\ 4 \end{bmatrix} v_1 + \begin{bmatrix} 1 \\ 5 \\ 7 \end{bmatrix} v_2$$

for alle possible choices of $v = (v_1, v_2)$. fin

More precisely, L consists of all vectors of the form

$$a_1v_1 + a_2v_2 + \cdots + a_cv_c$$

for all possible choices of c–vectors $v = (v_2, \ldots, v_c)$.

It is common to organize the building blocks as a matrix $A = [a_1 : \cdots : a_c]$. Then another way of describing L is as the set of vectors that can be written as Av, or more precisely

$$L = \{y|y = Av \text{ for all possible vectors } v\}$$

Frequenly one uses the name $\mathrm{span}(A)$ for L.

There are some additional aspects of subspaces of which a few will be illustrated:

Example 4. Consider again the subspace $L = \mathrm{span}(a_1, a_2)$ where

$$a_1 = (2, 6, 4)^\top \quad a_2 = (1, 5, 7)^\top$$

- A question is whether all vectors $y = (y_1, y_2, y_3)^\top$ can be written as $y = a_1v_1 + a_2v_2$?

  The answer is "no", for example $y = (1, 5, 3)$ can not be written in that form.

- Another question is whether there are other ways of representing L?

  The answer is "yes" – there are infinitely many. To pick one, let $b_1 = a_1 + a_2$ and $b_2 = a_1 - a_2$. Then $L = \mathrm{span}(b_1, b_2)$.

fin

- Note The 0-vector belongs to all linear subspaces. In the previous example one gets $y = 0$ when choosing $\alpha = (0, 0, 0)$.)

Linear dependence and independence

Linearly dependent vectors: A set of vectors $a_1, ..., a_c$ are *linearly dependent* if one of them can be written as a linear combination of the others, for example if

$$a_c = \sum_{j=1}^{c-1} a_j q_j$$

where the v_js are numbers.

Linearly independent vectors: If none of the vectors $a_1, ..., a_c$ can be written as a linear combination of the others, the set is said to be *linearly independent*.

Throw–out–technique: If one vector, say a_c, can be written as a linear combination of the other vectors, then it can be thrown away with changing the structure of the space, i.e.

$$\mathrm{span}(a_1, \ldots, a_c) = \mathrm{span}(a_1, \ldots, a_{c-1})$$

This process can go on until one ends up with a set of linearly independent vectors.

This allow us to find a representation of the which is as simple (economical) as possible.

Example 5. Consider the vectors

$$a_1 = \begin{bmatrix} 2 \\ 6 \\ 4 \end{bmatrix} , \quad a_2 = \begin{bmatrix} 1 \\ 5 \\ 7 \end{bmatrix} , \quad a_3 = \begin{bmatrix} 0 \\ 2 \\ 5 \end{bmatrix} \text{ og } x = \begin{bmatrix} 3 \\ 13 \\ 16 \end{bmatrix}$$

1. The vector x is a linear combination of a_1, a_2 and a_3, since $x = a_1 + a_2 + a_3$.

2. Since $a_3 = a_2 - \frac{1}{2}a_1$, the a_i–vectors are linearly dependent. Consequently x can be written as a linear combination of only a_1 og a_2, because $x = \frac{1}{2}a_1 + 2a_2$.

3. The vectors a_1, a_2 are linearly independent and so are the sets a_1, a_3 and a_2, a_3.

fin

Basis of a subspace: If the vectors $a_1, ..., a_c$ span a given subspace L and are linearly independent, the are said to be a basis for L.

Any linear subspace has infinitely many different bases.

Dimension of a linear subspace: Yet all bases of a linear subspace shares have a common feature: They have the same number of elements. The number of elements of a basis is the *dimension* of the subspace.

Throw–away: Having a linearly dependent set of vectors $a_1, ..., a_c$ on can always apply the throw–away–technique to obtain a linearly independent set of vectors. This set is then a basis

for $\text{span}(a_1, \ldots, a_c)$.

Example 6. Consider the vectors

$$a_1 = \begin{bmatrix} 2 \\ 6 \\ 4 \end{bmatrix}, \quad a_2 = \begin{bmatrix} 1 \\ 5 \\ 7 \end{bmatrix}, \quad a_3 = \begin{bmatrix} 0 \\ 2 \\ 5 \end{bmatrix}$$

$$b_1 = \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix} \quad \text{and} \quad b_2 = \begin{bmatrix} 2 \\ 8 \\ 9 \end{bmatrix}$$

and the corresponding matrices $A = [a_1 : a_2 : a_3]$, $\tilde{A} = [a_1 : a_2]$ og $B = [b_1 : b_2]$.

1. Since $a_3 = a_2 - \frac{1}{2}a_1$, the a_i vectors are linearly dependent.

fin

- Note Since $L = span(A) = span(B)$ one can think of the matrices A and B as two *different* ways of representing the *same* linear subspace.

Projections onto Linear Subspaces

Example 7. Consider the vector $a = (2, 2)$ and $y = (1, 2)$.

Clear y is not in $\text{span}(a)$. In statistics the following question is extremely important: Can we find a vector $\hat{y}$ in $\text{span}(a)$ which is as "close to" y as possible?

The answer is "yes": Find the *(orthogonal) projection* of the point y onto the line going through a. There is a simple mathematical expression for obtaining $\hat{y}$, namely

$$\hat{y} = a(a^\top a)^{-1}a^\top y = \begin{bmatrix} 2 \\ 2 \end{bmatrix} \frac{1}{8} [2, 2] \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} \frac{3}{2} \\ \frac{3}{2} \end{bmatrix}$$

The property of $\hat{y}$ is that the length of $y - \hat{y}$ is as small as possible.

Moreover, $y - \hat{y}$ and $\hat{y}$ are orthogonal. fin

In general let y be an r–vector and let $A = [a_1 : \cdots : a_c]$ be an $r \times c$ matrix.

Then there always exist a vector $\hat{y}$ in $\mathrm{span}(A)$ which is as close to y as possible.

If y is in $\mathrm{span}(A)$, then $\hat{y} = y$ because in this case the lenght of $y - \hat{y}$ is zero.

If y is not in $\mathrm{span}(A)$ then the expression is as follows: Assume that all columns of A are linearly independent. (Recall that if that is not

the case we can throw away redundant columns without changing the space spanned by those remaining.)

Then $\hat{y} = Py$ where

$$P = A(A^\top A)^{-1}A^\top$$

is the projection matrix onto $\mathrm{span}(A)$.

It then holds that

1. Py is in $\mathrm{span}()$.
2. Py is the vector in $\mathrm{span}(A)$ which is closest to y (in the sense that the lenght of $y - \hat{y}$ is minmized.
3. $Py = y$ if and only if y is already in $\mathrm{span}(A)$.

Example 8. Consider the 3×2 matrix $A = [a_1 : a_2]$, where

$$a_1 = \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix} \quad \text{og} \quad a_2 = \begin{bmatrix} 2 \\ 8 \\ 9 \end{bmatrix}$$

Then the projection matrix onto $\mathrm{span}(A)$ is $P = A(A^\top A)^{-1} A^\top$. To find P we first calculate

$$A^\top A = \begin{bmatrix} 1 & 3 & 2 \\ 2 & 8 & 9 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 8 \\ 2 & 9 \end{bmatrix} = \begin{bmatrix} 14 & 44 \\ 44 & 149 \end{bmatrix}$$

Hence

$$(X^\top X)^{-1} = \frac{1}{150} \begin{bmatrix} 149 & -44 \\ -44 & 14 \end{bmatrix}$$

From this we find

$$\begin{aligned} (X^\top X)^{-1} X^\top &= \frac{1}{150} \begin{bmatrix} 149 & -44 \\ -44 & 14 \end{bmatrix} \begin{bmatrix} 1 & 3 & 2 \\ 2 & 8 & 9 \end{bmatrix} \\ &= \frac{1}{150} \begin{bmatrix} 61 & 95 & 98 \\ -16 & -20 & 38 \end{bmatrix} \end{aligned}$$

Finally we find

$$P = A(A^\top A)^{-1}A^\top = \frac{1}{150}\begin{bmatrix} 1 & 2 \\ 3 & 8 \\ 2 & 9 \end{bmatrix}\begin{bmatrix} 61 & 95 & 98 \\ -16 & -20 & 38 \end{bmatrix}$$

$$= \frac{1}{150}\begin{bmatrix} 29 & 55 & -22 \\ 55 & 125 & 10 \\ -22 & 10 & 146 \end{bmatrix}$$

fin

Exercises in linear algebra

Exercise 1. *1. Are the vectors $(1,1)$ and $(1,2)$ orthogonal?*

2. Are $(1,1)$ and $(2,-2)$?

3. Are $(1,1)$ and $(-1,-1)$?

4. Make a drawing which illustrates these vectors

Exercise 2. *Let*

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}.$$

1. *Is A symmetrical?*

2. *Is $A^\top A$ symmetrical?*

3. *Is $AA^\top$ symmetrical?*

4. *What is the result from adding A and $A^\top$?*

Exercise 3. *Let*

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \text{ and } B = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}.$$

Calculate AB and BA. What can be concluded from this?

Exercise 4. *Let $a = (1, 1, 1, 0, 0, 0)^\top$ be a 6×1 matrix. Find $aa^\top$ and $a^\top a$.*

Exercise 5. *Let*

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

and

$$B = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

Calculate AB. What can be concluded from this?

Exercise 6. *What is the inverse to the 3×3 matrix $diag(1, 4, 9)$?*

Exercise 7. *Two equations with two unknowns. COnvince yourself* *that the system of equations*

$$\begin{aligned} x_1 + 2x_2 &= 3 \\ 2x1 + 3x_2 &= 4 \end{aligned}$$

can be written as

$$\begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 4 \end{bmatrix},$$

i.e. as $Ax = b$. Find A^{-1} and use this for solving the system of equations as follows:

$$x = Ix = A^{-1}Ax = A^{-1}b.$$

Exercise 8. *Let*

$$A = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}.$$

1. *How do vectors of the form Av look when $v = (v_1, v_2)^\top$?*
2. *Find the projection matrix $P = A(A^\top A)^{-1}A^\top$.*
3. *Let $y = (1, 3, 5, 7)^\top$. Find Py.*