

# Linear algebra

Kristoffer Klokke

2022

# Contents

<b>1</b>	<b>Systems of linear Systems of Linear Equations</b>	<b>3</b>
1.1	Gauss-Jordan Elimination . . . . .	3
1.2	Matrices and Matrix operations . . . . .	5
1.2.1	Matrix operations . . . . .	5
1.3	Inverses . . . . .	6
1.4	Determinants by cofactor expansion . . . . .	7
1.5	Properties of determinants . . . . .	8
1.6	Matrix transformation . . . . .	9
1.7	Projections . . . . .	10
<b>2</b>	<b>Vector spaces</b>	<b>10</b>
2.1	Subspaces . . . . .	11
2.2	Spanning sets . . . . .	11
2.3	Linearly independent . . . . .	12
2.4	Coordinates and basis . . . . .	12
2.5	Dimensions . . . . .	13
2.6	Change of basis . . . . .	13
<b>3</b>	<b>General Linear Transformation</b>	<b>13</b>
3.1	Matrices for general linear transformations . . . . .	16
3.2	Composition . . . . .	16
3.3	Isomorphism . . . . .	16
3.4	Similarity . . . . .	17
<b>4</b>	<b>Eigenvalues and eigenvectors</b>	<b>17</b>
4.1	Finding eigenvalues . . . . .	17
4.2	Finding eigenvectors . . . . .	18
4.3	Diagonalization . . . . .	18
<b>5</b>	<b>Angles</b>	<b>18</b>
5.1	Orthonormal basis . . . . .	19
5.2	Gram-Schmidt process for finding orthonormal basis . . . . .	20
5.3	Orthogonal matrices . . . . .	20
5.4	Orthogonal diagonalization . . . . .	20

# 1 Systems of linear Systems of Linear Equations

A system of linear equations are multiple equations containing unknowns which are shared. Ex.

$$2x + y + 3z = 10$$

$$x + y + z = 6$$

$$x + 3y + 2z = 13$$

The same system may be written in form of a matrix

$$\begin{bmatrix} 2 & 1 & 3 & 10 \\ 1 & 1 & 1 & 6 \\ 1 & 3 & 2 & 13 \end{bmatrix}$$

When solving a system there may be

- No solutions - No possible value can be assigned to the variable such that all equations are true (inconsistent)
- One solution - A combination of values can be assigned to make every equation true (consistent)
- Infinite solutions - One or more unknowns may have an infinite amount of possible assignable values (consistent)

By transforming a system of equations to a matrix, the following row operations can be performed:

- Multiply a row through by a nonzero constant.
- Interchange two rows.
- Add a constant times one row to another.

## 1.1 Gauss-Jordan Elimination

Solving a system of equations in a matrix, can be done such that the matrix has the following requirements

- If the row contains nothing but zeroes the first number should be a 1, called the leading 1.

- If a row is made of nothing but zeroes it should be grouped at the bottom
- In two rows the top row should contain a leading 1 further to the left than the bottom
- Each column which contains a leading 1, every number in the same column below should be 0

This form is called row echelon form.

The solution may also be written as:

$$\{(x_1 = 4, x_2 = 6, x_3 = t, x_4 = v, x_5 = 1) | t \in \mathbb{R}\}$$

Where variables means they can be any possible assignment or a function with given restrictions, called free variables.

In case of the leading 1 column is zero both above and underneath the matrix is in reduced echelon form.

To make a matrix into echelon form the following algorithm can be used:

1. Locate the leftmost column that does not consist entirely of zeros and exchange it to the top
2. Multiply the top row by  $\frac{1}{a}$  where  $a$  is the leading number in the row
3. subtract top row from every other below row, such the top row is the only non zero value in the column
4. Repeat 2 and 3 but ignore the top row and let the second top row be the top

To reduce the echelon form, from the bottom the bottom row is added to the above rows until the leading 1 is the only in the column. This is repeated until the top is reached.

A homogeneous linear system are systems which all constant (right part of equal) are 0 and the trivial solution (all variables are assigned 0) are a possible solution.

A free variable is the term for a variable which can be assigned multiple values. The number of free variables will be equal to the number of variables minus zero rows.

In a homogenous linear system if the number of unknowns exceed the number of equation there will be an infinite amount of solutions.

Back substitution is a method taking the echelon form and starting from the bottom isolating the leading variable and substituting it upwards.

A echelon form is not unique to a system but a reduced echelon form is unique and the number of zero rows will be unique.

## 1.2 Matrices and Matrix operations

A matrix is a rectangular array of numbers called entries.

The size of a matrix is written as rows x columns

A single row matrix is called a row vector and single column is called column vector.

Standard basis is the set of column vectors in a identity matrix.

Variables in matrices are called scalars and unless stated is in the realm of real numbers.

When referring to a number in the matrix  $A$  it will be  $a_{row\ column}$  and the value can be found by  $(A)_{row\ column}$

A matrix scalar matrix can be written as  $A = [a_{ij}]_{m \times n}$  the  $m \times n$  is optional if the size matters.

### 1.2.1 Matrix operations

Addition/subtraction - every number in corresponding entries are added/-subtracted.

Addition and subtraction can not be done on two different sized matrices.

Multiplying a matrix with a scalar, is done by multiplying the scalar on every entry.

For multiplying two matrices an entry is found by  $a_{ij} = \sum_{t=0}^n b_{i(j+t)} \cdot c_{(i+t)j}$

where  $n$  is the size of row size of matrix  $b$  and column size of matrix  $c$ , if this value does not match the two matrices can not be multiplied.

The resulting product matrix of  $a = b \times c$  will have the size rows = c columns and columns = b rows

Partitioning a matrix is the act of splitting it into a smaller valid matrices.

This can be used if only certain entries are wanted for a matrix multiplication

Linear combinations are an array of matrices (length  $n$ ) of the same size matrix  $A$  and an array of scalars (length  $n$ )  $c$  computed as  $\sum_{i=0}^n A_i \cdot c_i$

Multiplying a vector row and a matrix (with appropriate lengths) can therefore be expressed as an linear combination.

Column row expansion is the act of splitting two matrices being multiplied into rows and columns, and then taking the sum of each product of row and column, and thus getting the multiplied matrix.

Matrix form of linear system is the act of expressing a linear system in the form of matrix  $A$  containing all coefficients, matrix  $X$  containing a column vector of all unknown variables, and matrix  $B$  which is a column vector of all results of the linear system. in this way the linear system can be expressed as  $AX = B$

Transposing a matrix is flipping column and rows such in matrix  $A$  the entry  $a_{ij}$  is now entry  $a_{ji}$  in the matrix  $A^T$ .

A trace operation of matrix  $A$  is taking the addition of the diagonal line called  $tr(A)$ , can only be performed on square matrices.

The trace of a square matrix is defined as  $tr(A) = a_{11} + a_{22} + \dots + a_{nn}$  or  $tr(A) = \sum_{i=1}^n a_{ii}$

For the trace the following is true

- $tr(A^T) = tr(A)$
- $tr(A \pm B) = tr(A) \pm tr(B)$
- $tr(cA) = c \cdot tr(A)$
- $tr(AB) = tr(BA)$

The matrix polynomials are a way of describing multiple equations in matrix form by

If  $A$  contains all the coefficients,  $X$  contains all variables in a column vector, and  $B$  contains all constant the equations are equal to in a column vector.

Then the equations can be described as  $A \cdot X = B$

### 1.3 Inverses

For matrix operations are the following true, if the size allows for it

- Commutative law for matrix addition -  $A + B = B + A$
- Associative law for matrix addition -  $A + (B + C) = (A + B) + C$
- Associative law for matrix multiplication -  $A(BC) = (AB)C$
- Left and Right distributive law -  $A(B + C) = AB + AC$

Zero matrices are matrices consisting of only zeros and are denoted by  $0_{m \times n}$ . Identity matrix  $I$  is the matrix containing 1's on the diagonal, and has the property that if multiplied by  $A$  the product is  $A$ .

The inverse matrix  $A^{-1}$  is the matrix  $B$  to the square  $A$  if  $AB = I$ , and  $A$  is said to be invertible. If  $B$  does not exist  $A$  is singular.

Only one inverse exists to a matrix.

For two inverse matrices is the following true  $B^{-1}A^{-1} = (AB)^{-1}$ .

Powers of a matrix are defined as expected, and for  $A^0 = I$  and if  $A$  is invertible then  $A^{-n} = (A^{-1})^n$ .

$A^n$  is invertible and  $(A^n)^{-1} = A^{-n} = (A^{-1})^n$ .

Matrix polynomials are inserting a matrix into a polynomial function.

Transpose is the interchanging of rows in a matrix, and are expressed as  $(A^T)^{-1} = (A^{-1})^T$ .

The transposed matrix has the following properties:

- $(A^T)^T = A$
- $(A \pm B)^T = A^T \pm B^T$
- $(cA)^T = cA^T$
- $(AB)^T = B^T A^T$

## 1.4 Determinants by cofactor expansion

Minor of entry  $a_{ij}$  is denoted  $M_{ij}$  and is defined as the determinant of the submatrix that remains when  $i$  row and  $j$  column is deleted.

The cofactor  $C_{ij}$  of  $M_{ij}$  is defined as, if  $i + j$  is even then  $1^{i+j}M_{ij}$  and  $(-1)^{i+j}M_{ij}$  if odd.

The determinant of an  $n \times n$  matrix can be found either by a row or column in the following way:

$$\det(A) = a_{1j}C_{1j} + a_{2j}C_{2j} + \dots + a_{nj}C_{nj}$$

This is via the column and by switching  $i$  and  $j$  it is by the row.

Therefore making the definition recursive.

If we make the function  $f(i, j) = 1 - ((i + j) \% 2) \cdot 2$  such it will be  $-1$  when the sum of  $i$  and  $j$  is odd and  $1$  when the sum is even.

$$\det(A) = a_{1j}f(1, j)^{1+j}M_{1j} + a_{2j}f(2, j)^{2+j}M_{2j} + \dots + a_{nj}f(n, j)^{n+j}M_{nj}$$

if we also denote the function  $g(A, i, j)$  to be a function which removes row  $i$  and  $j$

Then we can make a true recursive definition.

$$\det(A) = a_{1j}f(1, j)^{1+j}\det(g(A, 1, j)) + a_{2j}f(2, j)^{2+j}\det(g(A, 2, j)) + \dots + a_{nj}f(n, j)^{n+j}\det(g(A, n, j))$$

Where for the 1x1 matrix  $\det([a_{11}]) = a_{11}$  If  $A$  is triangular then  $\det(A)$  is equal to the product of the main diagonal.

It can also be noted that determination of the zero matrix is zero

And for the transpose  $\det(A) = \det(A^T)$

To make life easier row operation can be done such less operations are needed to find the determinant.

When performing row operation the following is the for the  $A$  matrix with size  $n \times n$

- If  $B$  is the result of a row or column in  $A$  being multiplied by a constant  $k$  then  $\det(B) = k \cdot \det(A)$
- if  $B$  is the result of two rows or two columns being interchanged then  $\det(B) = -\det(A)$
- if  $B$  is the result of one row being multiplied and added to another row then  $\det(B) = \det(A)$

For the special case where  $A = I_n$  and  $E$  is the result of an row operation on  $A$  the following applies

- If  $E$  is the result of multiplying a row with the constant  $k$  which is not 0 then  $\det(E) = k$
- If  $E$  is the result of interchanging two rows then  $\det(E) = -1$
- If  $E$  is the result of adding a constant to a row and adding to another then  $\det(E) = 1$

If  $A$  has a row/column which is proportional to another then  $\det(A) = 0$

Reducing a matrix the echelon, in doing so instead of reducing the a row by  $1/\text{the leading value}$ , the leading value is taken to the side, then the product of every value taken to the side will be equal to the determinant.

## 1.5 Properties of determinants

If  $A$  is a  $n \times n$  matrix and  $C_{ij}$  is the cofactor of  $A_{ij}$ , then the matrix is called matrix of cofactors from  $A$

The transpose of the matrix of cofactors from  $A$  is called adjoint of  $A$



- $\det(kA) = k^n \det(A)$
- $\det(C) = \det(A) + \det(B)$  If A and B only differ in one row/column and C is equal to A and the the differ row/column sum
- $\det(AB) = \det(A)\det(B)$  if both A and B are  $n \times n$  matrices
- A is only invertible if  $\det(A) \neq 0$
- If A is invertible then  $\det(A^{-1}) = \frac{1}{\det(A)}$
- If A is invertible then  $A^{-1} = \frac{1}{\det(A)} \text{adj}(A)$

In the system of linear equation  $Ax = b$  then the solutions can be found by

$$x_n = \frac{\det(A_n)}{\det(A)}$$

## 1.6 Matrix transformation

The matrix transformation is function  $T_A(x) = A \cdot x$ , which transforms the axis linearly, such origo stays in the same place.

The function then takes a vector of length  $n$  and multiply it / transforms it with a matrix of size  $n \times n$

The transformation which has the following properties:

- $T_A(0) = 0$  Origin stays in place
- $T_A(k \cdot x) = k \cdot T_A(x)$
- $T_A(x + y) = T_A(x) + T_A(y)$
- $T_A(x - y) = T_A(x) - T_A(y)$

To think of the matrix A which the transformation use, it can be though of in the coordination system x,y,z A would then be a 3x3 matrix.

Each column then represent the standard basis for the axis.

Such for the Cartesian (standard) coordination system, the matrix would be a 3x3 identity matrix.

Generally speaking the identity matrix is also called to standard matrix

The transformation may also be composited together like  $(T_B \cdot T_A)(x) = T_B(T_A(X))$

Likewise multiplication the transformation is not commutative  $T_A \cdot T_B \neq T_B \cdot T_A$

The exception to this is rotation and reflection transformations.

If  $A$  is invertible then  $T_A : \mathbb{R}^n \rightarrow \mathbb{R}^n$  is also invertible.

This is written as  $T_A^{-1} = T_{A^{-1}}$

A linear transformation is when  $T : \mathbb{R}^N \rightarrow \mathbb{R}^M$  satisfies the following linear conditions

- $T(k \cdot x) = k \cdot T(x)$
- $T(x + y) = T(x) + T(y)$

## 1.7 Projections

A projection is a transformation to a lower dimension system.

An orthogonal projection is a projection on a line, plane or the appropriate amount of dimensions.

## 2 Vector spaces

If  $V$  defines all objects which are non empty then

- $u, v \in V \rightarrow v + u \in V$
- $u + v = v + u$
- $u + (v + w) = (u + v) + w$
- The zero vector exist called 0 with the property  $v + 0 = v$
- For each  $u$  in  $V$  exists a  $-u$  such  $u + (-u) = 0$
- If  $k$  is a scalar and  $u$  is in  $V$  then  $uk$  is in  $V$
- $k(u + v) = ku + kv$
- $k(mu) = (km)u$
- $1u = u$

If an object satisfies all these axioms then we can classify a set of objects as being in the vector space.

Therefore  $\mathbb{R}$  is in the vector space,  $\mathbb{R}^n$  being the vector of real number length  $n$  is in the vector space.

Likewise  $\mathbb{R}^{n \cdot m}$  matrix is in the vector space and even functions like the polynomial functions are in the vector space.

## 2.1 Subspaces

A subspace  $W$  of  $V$  is a subspace if  $u$  and  $v$  are in  $W$  and  $u + v$  is in  $W$ .  
And if  $k$  is a scalar and  $u$  is in  $W$  then  $ku$  is in  $W$

## 2.2 Spanning sets

$w$  is said to be a linear combination if

$$w = k_1v_1 + k_2v_2 + \dots + k_rv_r$$

Where  $k$  is coefficients and  $v$  is vectors in the vector space  $V$ .

To say that the linear combination is spanning means by the given vectors a linear combination can be made such every vector is made in the space.

So if  $V$  is the vector space and  $W \leq V$  a subspace such that  $W = \text{span}\{w_1, \dots, w_n\}$  then the set is the spanning set of  $W$

To find if a vector can be expressed as a combination of two vectors, the two vectors is added with an unknown scalar at each vector in the addition.

The expression can then be split up such for the vector of size  $n$  there will be  $n$  equation with  $q$  unknowns equal to the amount of addition.

$$\begin{aligned}(9, 2, 7) &= k_1(1, 2, -1) + k_2(6, 4, 2) \\ &= (k_1 + 6k_2, 2k_1 + 4k_2, -k_1 + 2k_2)\end{aligned}$$

$$9 = k_1 + 6k_2$$

$$2 = 2k_1 + 4k_2$$

$$77 = -k_1 + 2k_2$$

$$k_1 = -3$$

$$k_2 = 2$$

This method can be more generalised by setting it up in a matrix form and instead of finding a single vector it can just be a general vector of variables, from which Gauss-Jordan elimination can be used to find if the solution is consistent. To test if a set of vectors span an entire space, they can be setup as a matrix and if the determinant is non zero then it spans the entire space.

## 2.3 Linearly independent

A set of vectors is set to be linearly independent set if no vector in the set can be expressed as linear combination of other vectors in the set.

A set is linearly independent if the only solution is  $k_1 = 0, k_2 = 0, \dots, k_r = 0$

$$k_1 v_1 + k_2 v_2 + \dots k_r v_r = 0$$

Where  $v$  is the vectors in the vector set.

This is again tested in a matrix form using elimination.

A set of non infinite vectors that contains 0 is linearly dependent. - This can be seen by the zero coefficient test will result in 0 using any coefficient.

A set with exactly two vectors which is not a scalar multiple of each other are linearly independent.

If a set contains more vectors than the size of the vectors then the set is linearly dependent.

To test if a set of functions  $f_1, f_2, \dots, f_n$  are linearly dependent they must be continuous derivable to  $n - 1$ .

They can then be setup in a matrix of size  $n \times n$  where the first row is the functions, the next is the function derived once, this continuous til  $n - 1$ .

If the determinant is not 0 on  $(-\infty, \infty)$  then the functions form a linearly independent set of vectors.

## 2.4 Coordinates and basis

Coordinate system are defined by the basic vectors, in a rectangular 2D system they are (0,1) and (1,0).

They both describe the positive direction and the unit spacing between integers.

A vector space  $V$  is set to be finite-dimensional if there exist a finite set of vectors in  $V$  which spans  $V$ .

An infinite-dimensional set has no such set which spans  $V$ .

If  $S$  is a set of vectors in a finite-dimensional vector space  $V$  then  $S$  is a basis for  $V$  if:

- $S$  spans  $V$
- $S$  is linearly independent

The basis for a vector space  $V$  there will be a unique combination for every vector in  $V$ .

In the ordered basis  $S$  the coefficients in the linear combination is called the

coordinates of  $v$  relative to the basis  $S$

The vector made from the combination is called the coordinate vector of  $v$  relative to  $S$  and is denoted by  $(v)_S = (c_1, c_2, \dots, c_n)$  called comma-delimited form and may be expressed in a column matrix called matrix form

## 2.5 Dimensions

For the vector space with  $n$  dimensions:

If a set in  $V$  has fewer than  $n$  vector then it does not span  $V$  - due to some dimension not being represented

If a set in  $V$  has more than  $n$  vectors then it not linearly dependent

$\dim(V)$  is a function which gives the number of dimensions of  $V$ .

## 2.6 Change of basis

When changing coordinate system a vector  $V$  will be changed in to  $(v)_s$  is then is displayed in matrix form.

Therefore the before coordinate is changed in the coefficients used to create the vector.

Changing the basis can be described as a transformation matrix, which transform the old basis vectors into the new basis vectors.

It is denoted by  $P_{B \rightarrow B'} = [[u'_1]_B | [u'_2]_B | \dots | [u'_n]_B]$  where  $B$  is the old basis and  $B'$  is the new basis

It can here be observed that  $P_{B \rightarrow B'} \cdot P_{B' \rightarrow B} = I$

To find the transition matrix, a matrix can be made of [new basis—old basis] where the basis's are column vectors of the basis vectors.

Then the matrix is reduced to echelon form which results in [I—transition matrix from old to new]

## 3 General Linear Transformation

If  $T : V \rightarrow W$  is a mapping between vector spaces then  $T$  is a linear transformation if the two properties hold

- $T(ku) = kT(u)$
- $T(u + v) = T(u) + T(v)$

It can here be observed that the following is true

- $T(0) = 0$
- $T(u - v) = T(u) - T(v)$
- $T(-v) = -T(v)$

The mapping  $I : v \rightarrow$  defined by  $I(v) = v$  is the identity operator.

Let  $T : V \rightarrow W$  a linear transformation. If  $S = \{v_1, v_2, \dots, v_n\}$  is the basis for  $V$  then any vector in  $V$  can be described as

$$T(v) = c_1T(v_1) + c_2T(v_2) + \dots + c_nT(v_n)$$

Where  $c_1, c_2, \dots, c_n$  are the coefficients required to express  $v$  as a linear combination of vector basis in  $S$

For the transformation  $T : V \rightarrow W$  then the kernel denoted by  $\ker(T)$  is all vectors in  $V$  which maps into 0

Given a set of vectors the range denoted  $R(T)$  is the possible outcome given the set of vector inputs.

The rank of a transformation is denoted by  $\text{rank}(T)$  and is equal to the number of dimension in the transformation

The Nullity of the transformation is denoted by  $\text{nullity}(T)$  and is equal to the number of dimensions which is in the kernel.

It can here be seen in the transformation  $T : V \rightarrow W$  that  $\text{rank}(T) + \text{nullity}(T) = \dim(V)$

A transformation is set to be one-to-one if it maps every distinct vector to a new distinct vector. -  $\ker(T) = \{0\}$

A transformation  $T : V \rightarrow W$  is set to be onto  $W$  if every vector in  $W$  can be obtained through a mapping of  $V$

$T_A$  is one-to-one only if the columns of  $A$  are linearly independent- due to if not linearly independent the nullity would be more than 0

If  $A$  is an  $n \times n$  matrix in which there are no duplicate rows and no duplicate columns, then the following statements are equivalent.

- $A$  is invertible
- $Ax = 0$  has only the trivial solution
- The reduced row echelon form of  $A$  is  $I_n$ .
- $A$  is expressible as a product of elementary matrices.
- $Ax = b$  is consistent for every  $n \times 1$  matrix  $b$ .

- $Ax = b$  has exactly one solution for every  $n \times 1$  matrix  $b$ .
- $\det(A) \neq 0$ .
- The column vectors of  $A$  are linearly independent.
- The row vectors of  $A$  are linearly independent.
- The column vectors of  $A$  span  $R^n$ .
- The row vectors of  $A$  span  $R^n$ .
- The column vectors of  $A$  form a basis for  $R^n$ .
- The row vectors of  $A$  form a basis for  $R^n$ .
- $A$  has rank  $n$ .
- $A$  has nullity 0.
- The orthogonal complement of the null space of  $A$  is  $R^n$ .
- The orthogonal complement of the row space of  $A$  is  $\{0\}$ .
- $\lambda = 0$  is not an eigen value of  $A$ .
- $A^T A$  is invertible.
- The kernel of  $T_A$  is  $\{0\}$ .
- The range of  $T_A$  is  $R^n$ .
- $T_A$  is one-to-one.

The inverse transformation is defined by if  $T_A : V \rightarrow W$  then  $T_{A^{-1}} = W \rightarrow V = T^{-1}$

The row space is the space defined by the rows. By reducing the transformation the number of non zero rows are the row space rank.

The same goes for columns in the column space.

It can here be observed that the row space and column space will have the same rank.

### 3.1 Matrices for general linear transformations

For the two vector spaces  $V$  with basis  $B$  and  $W$  with basis  $B'$ . Here vector space  $V$  is  $n$  dimensions and  $W$  is  $m$  dimensions.

Our goal is to find a matrix  $A$  such a vector  $x$  can be transformed from  $V$  basis  $B$  to  $W$  basis  $B'$  -  $A[x]_B = [T(x)]_{B'}$

If we have an  $A \in \mathbb{M}_{m,n}(\mathbb{R})$  then  $T(x)$  can be computed in the following steps

1. First compute  $[x]_B \in \mathbb{R}^n$
2. Then compute  $A \cdot [x]_B \in \mathbb{R}^m$
3. Reconstruct  $T(x)$  from its  $B'$  coordinates  $[T(x)]_{B'} = A \cdot [x]_B$

To find  $A$  we can do the following

1. Write  $B = \{u_1, \dots, u_n\}$  and  $B' = \{v_1, \dots, v_m\}$
2. Then  $A = [[T(u_1)]_{B'} | \dots | [T(u_n)]_{B'}] \in \mathbb{M}_{m,n}(\mathbb{R})$

The matrix  $A$  is denoted  $[T]_{B',B}$  and the order is chosen such  $[T(x)]_{B'} = [T]_{B',B} \cdot [x]_B$

This is called matrix of  $T$  relative to  $B$  and  $B'$

### 3.2 Composition

The transformation of  $T_1 : U \rightarrow V$  and  $T_2 : V \rightarrow W$  can be combined to the linear transformation  $(T_1 \cdot T_2) : U \rightarrow W$

This is called the composition of the two transformations.

This may again also be chained with other transformations. The composition will be linear and if both transformation is one-to-one then the composition is also one-to-one.

If both transformations are isomorphisms then the composition and the inverse of the composition are isomorphisms

### 3.3 Isomorphism

A transformation  $T : V \rightarrow W$  which is both one-to-one and onto is isomorphism, and  $W$  is isomorphic to  $V$  written as  $W \simeq V$

Therefore there will exist an inverse transformation such  $T^{-1} \cdot T(v) = v$

Two vector spaces are isomorphic iff they have the same dimension



### 3.4 Similarity

To find a transformation in basis  $B$  in another basis  $B'$  then  $[T]_{B'} = P_{B' \rightarrow B}^{-1} \cdot [T]_B \cdot P_{B' \rightarrow B}$ . It can then be said that  $[T]_{B'}$  and  $T_B$  is similar.

## 4 Eigenvalues and eigenvectors

An eigenvector is vector which only get scaled in a transformation, which is not a zero vector.

This scaled value is the eigenvalue.

$$Ax = \lambda x$$

Where  $x$  is the eigen vector,  $A$  is a transformation and  $\lambda$  is the eigenvalue. For the matrix  $A$  it will be invertible only if  $\lambda = 0$  is not an eigenvalue of  $A$ .

### 4.1 Finding eigenvalues

To find an eigenvalue of a transformation, first it can be rewritten as

$$Ax = \lambda \cdot Ix$$

Such it is matrix multiplication on both sides and it can therefore be reduced

$$(\lambda \cdot I - A)x = 0$$

Then an value for  $\lambda$  can be found such the matrix in the parentheses will be linearly dependent.

When the matrix is linear dependent, this makes it possible for there to exist an eigenvector of which the first equation holds true.

When expanding  $(\lambda \cdot I - A)$  it is known as the characteristic polynomial denoted  $p(\lambda)$

This polynomial is then solved to find the eigenvalues

An easy way to find eigenvalues for a 2x2 is by the formula

$$\lambda = m \pm \sqrt{m^2 - p}$$

Where  $m$  is the mean of the diagonal and  $p$  is the determinant.

In case of an upper triangle or lower triangle the diagonal is the eigen values.

## 4.2 Finding eigenvectors

To find the eigenvector first the eigen values have to be known.  
Then the equation

$$(\lambda \cdot I - A) \times x = 0$$

Where  $x$  is the matrix of same size of  $A$ 's row and 0 is representing the zero vector.

Then a system of equation can be made and solved for.

The solutions can then be inserted in  $x$  to find the eigenvector.

## 4.3 Diagonalization

Diagonalization describes the relation between eigenvalues, eigenvectors and the original vector

$$A = XDX^{-1}$$

Where  $A$  is an matrix square with eigen values  $\lambda_1, \lambda_2, \dots$

$X$  is the matrix consisting of the eigenvectors  $[x_1, x_2, \dots]$  which all are linear independent

$D$  is the matrix consisting of 0's and the eigenvalues on the diagonal in the same order as the eigenvectors for  $X$ .

In general the diagonalization where  $B = P^{-1}AP$  it is said that  $B$  and  $A$  are similar and will share many similarities and properties.

Diagonalization is useful since the eigenvalue to the power of the integer  $c$  is the eigenvalue to  $A$  to the power of  $c$

So the transformation using  $X$  can be used to transform  $D$  to the wanted power

## 5 Angles

To find angles for objects the formula

$$\cos \theta = \frac{\langle u, v \rangle}{\|u\| \|v\|}$$

Where  $\langle u, v \rangle$  is the inner product, which is a extension of dot product.

Found by taking  $a_1b_1 + a_2b_2 + \dots a_nb_n$ .

$\|u\|$  is the length of  $u$  found by taking  $\sqrt{a_1^2 + a_2^2 + \dots + a_n^2}$ .

The innerproduct and length has to following properties

$$\bullet \langle u, v \rangle^2 \leq \langle u, u \rangle \langle v, v \rangle$$

- $\langle u, v \rangle^2 \leq \|u\|^2 \|v\|^2$
- $\|u + v\| \leq \|u\| + \|v\|$
- $d(u, v) \leq d(u, w) + d(w, v)$  where  $d$  is the distance function
- $\langle u, v \rangle = 0$  the angle will be  $90^\circ$  there the objects are orthogonal
- $\|u + v\|^2 = \|u\|^2 + \|v\|^2$

Orthogonal complements denoted  $W^\perp$  are orthogonal objects for a subspace  $W$  of  $V$ .

Ex. in  $R^3$  and the subspace describing a plane the orthogonal complement is vector of which the whole plane the orthogonal to.

Therefore  $W \cap W^\perp = \{0\}$  and  $(W^\perp)^\perp = W$

## 5.1 Orthonormal basis

Orthonormal is the set of vectors in norm 1 which are orthogonal to an object. By norm 1, it means that every value in the vector is divided by the length of the vector.

The orthonormal form can be seen to be linear independent and therefore a basis for  $V$ .

To then find a vector  $u$  in the basis, it can be found as

$$u = \frac{\langle u, v_1 \rangle}{\|v_1\|^2} v_1 + \frac{\langle u, v_2 \rangle}{\|v_2\|^2} v_2 + \dots + \frac{\langle u, v_n \rangle}{\|v_n\|^2} v_n$$

Or just by

$$u = \langle u, v_1 \rangle v_1 + \langle u, v_2 \rangle v_2 + \dots + \langle u, v_n \rangle v_n$$

If  $W$  is a finite dimensional subspace of inner product space  $V$ , then every vector  $u$  in  $V$  can be described by

$$u = w_1 + w^\perp$$

Where  $w_1$  is in  $W$

The projected vector  $u$  is found by a linear combination of the basis vectors.

## 5.2 Gram-Schmidt process for finding orthonormal basis

The gram-schmidt process works by taking an original basis and turning it into orthogonal basis.

For the basis  $B = \{b_1, b_2, \dots, b_n\}$  to the orthogonal basis  $U = \{u_1, u_2, \dots, u_n\}$  it can be done by

$$u_k = b_k - \sum_{i=1}^{k-1} \frac{\langle b_k, u_i \rangle}{|u_i|^2} u_i$$

Therefore making the first three terms

$$u_1 = b_1 \tag{1}$$

$$u_2 = b_2 - \frac{\langle b_2, u_1 \rangle}{|u_1|^2} u_1 \tag{2}$$

$$u_3 = b_3 - \frac{\langle b_3, u_1 \rangle}{|u_1|^2} u_1 - \frac{\langle b_3, u_2 \rangle}{|u_2|^2} u_2 \tag{3}$$

## 5.3 Orthogonal matrices

A matrix which inverse is equal to the transpose is said to be orthogonal.

$$AA^T = A^A = I$$

$$A^{-1} = A^T$$

It can be seen that if  $A$  is orthogonal the row and column vectors of  $A$  form an orthonormal set in  $R^n$  with the euclidean inner product

A product of orthogonal matrices is orthogonal.

if  $A$  is orthogonal then  $\det(A) = 1 \vee \det(A) = -1$

$\|Ax\| = \|x\|$  for all  $x$  in  $R^n$  and  $Ax \cdot Ay = x \cdot y$  for all  $x$  and  $y$  in  $R^n$

## 5.4 Orthogonal diagonalization

If  $A$  and  $B$  are square matrices, then they are orthogonally similar matrices, if there exists an orthogonal matrix  $P$  where  $B = P^T A P$  and therefore  $A = P B P^T$

$P^T A P = D$   $P$  orthogonally diagonalizes and  $A$  is orthogonally diagonalizable

It can be seen that  $A$  must be symmetric if it is orthogonally diagonalizable, and  $A$  has an orthonormal set of  $n$  eigenvectors.

To find the orthogonal diagonalization, first the each eigenspace is found and the gram schmidt process is applied to find basis for each eigenvector.

Then a matrix  $P$  is constructed from the eigenvectors in the columns. The diagonalization will then have the eigenvalues in the diagonal in the same order as the column vectors.

The spectral decomposition comes from  $A = PDP^T$  which can be written as  $A = \lambda_1 u_1 u_1^T + \lambda_2 u_2 u_2^T + \lambda_n u_n u_n^T$

It can be seen that multiplying a vector  $x$  to  $A$ , the result can be found by multiplying  $x$  into each  $\lambda u u^T$ .

There can for  $D$  be found matrices in the form of an upper right triangle denoted  $S$  and in the form of upper right triangle plus an extra diagonal denoted  $H$ .

These forms can help reduce the number of operations in algorithms.