# Linear Algebra 2

Zambelli Lorenzo
BSc Applied Mathematics

February 2021-March 2021

## Inner Product spaces

**Goal:** move from $\mathbb{R}^n$ to the higher level of any arbitrary real vector space

**Definition 1** ***Inner product*** *$V$ a real vector space. An inner product on $V$ is a function $\langle\, , \rangle : V \times V \to \mathbb{R}$[1] that satisfies the three properties:*

1. $\langle x, x\rangle \geq o\ \forall x \in V$ *and* $\langle x, x\rangle = 0 \Leftrightarrow x = 0$
2. $\langle x, y\rangle = \langle y, x\rangle\ \forall x, y \in V$
3. $\langle \alpha x + \beta y, z\rangle = \alpha \langle x, z\rangle + \beta \langle y, z\rangle\ \forall \alpha\beta \in \mathbb{R}$ , $x, y, z \in V$

**Definition 2** ***Inner product space*** *Real vector space $V$ together with an inner product $\langle\, , \rangle$*

**Note:**

- **Norm of** $v \in V$: $\|v\| = \langle v, v\rangle^{1/2}$ and
- If $v, u$ are orthogonal than $\langle v, u\rangle = 0$

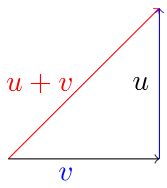
**Exemples:**

1. $V = \mathbb{R}^n$, $\langle x, y\rangle := x^T y$ (scalar product)
2. $\mathbb{R}^{m\times n}$ $\langle A, B\rangle = \sum_{i=1}^{n}\sum_{j=1}^{m} A_{ij}B_{ij}$
3. $V = C[a, b]$ $\langle f, g\rangle := \int_a^b f(x)g(x)dx$
4. $P_n = \left\{p(x)|p(x) = p_0 + p_1 x + \cdots + p_{n-1}x^{n-1}\right\}$ real polynomials of degree at most $n-1$.
   Take $x_1, x_2, ..., x_n \in \mathbb{R}$ $n$ distinct reals.
   Inner Product on $P_n$:
   $$\langle p, q\rangle := \sum_{i=1}^{n} p(x_i)q(x_i)$$

**Theorem 1** ***Pythagoras Theorem*** *Let $u, v \in V$ be orthogonal. Then*

$$\|u + v\|^2 = \|u\|^2 + \|v\|^2$$

[1] $V \times V$ means cartesian product

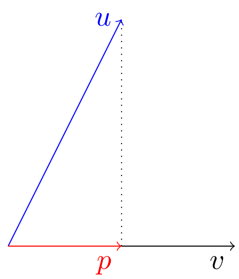



**Note:** For $A \in \mathbb{R}^{m \times n}$ $\|A\| = (trA^T A)^{1/2}$

**Definition 3** ***Projection*** *For arbitrary $(V, \langle\ ,\ \rangle)$ we define for given $u, v$*

$$\begin{aligned} p &:= \frac{\langle u, v\rangle}{\|v\|} \cdot \frac{v}{\|v\|} \\ &= \alpha \cdot \textit{unit vector} \end{aligned}$$

*Projection of $u$ onto $v$*




The concept of vector projection indeed gives te following two geometric properties:

1. $u - p$ and $p$ are orthogonal
2. $u = p \Leftrightarrow \exists \beta \in \mathbb{R}$ such that $u = \beta v$

**Theorem 2** ***Cauchy-Schwarz*** *Let $(V, \langle\ ,\ \rangle)$ be an inner product space and let $u, v \in V$ then*

$$|\langle u, v\rangle| \leq \|u\| \cdot \|v\|$$

*There, equality holds if and only if $u$ and $v$ are linearly dependent*

Now we can also define the angle between two vectors $u, v \in V$ with $V$ be arbitrary inner product space. Since $|\langle u, v\rangle| \leq \|u\|\|v\|$, we have

$$-1 \leq \frac{\langle u, v\rangle}{\|u\|\|v\|} \leq 1$$

whose angle: $\theta \in [-\pi, \pi]$ such that

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

## General Normed spaces $(V, \|.\|)$

**Theorem 3** *Let $\langle \, , \, \rangle$ be an inner product on $V$. For $v \in V$, define*

$$\|v\| = \sqrt{\langle v, v \rangle}$$

*Then $\|.\|$ is a norm on $V$*

In mathematics, we also define the concept of norm on a real vector space

**Definition 4** *Let $V$ be a real vector space. A **norm on V** is a function $\|.\| : V \to \mathbb{R}$ that satisfies the following properties:*

1. $\|v\| \geq 0 \ \forall v \in V$ *and* $\|v\| = 0 \Leftrightarrow v = 0$
2. $\|\alpha v\| = |\alpha| \|v\| \ \forall \alpha \in \mathbb{R}, \ v \in V$
3. $\|v + w\| \leq \|v\| + \|w\| \ \forall v, w \in V$ *(triangle inequality)*

*If $\|.\|$ is a norm of $V$, then the pair $(V, \|.\|)$ is called **a normed linear space***

**Examples**

1. $V \in \mathbb{R}^n$ for $v = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$ define $\|v\| = \sum_{i=1}^{n} |v_i|$ This is a norm on $\mathbb{R}^n$. We often denote it by $\| \, \|_1$, "1_norm"
2. $V = \mathbb{R}^n$, $\|v\| = \max_{i=1,\ldots,n} |v_i|$ it is also a norm, denoted by $\| \, \|_\infty$, or infinity norm
3. $V = C[a, b]$, $\|f\|_\infty = \max_{x \in [a,b]} |f(x)|$
4. $V = C[a, b]$, $\|f\|_1 = \int_a^b |f(x)| dx$
5. $V = \mathbb{R}^n$, $\|v\| = \left( \sum_{i=1}^{n} |v_i|^P \right)^{1/P}$, P-norm

**Theorem 4** *Let $V$ be a real linear space and let $\|.\|$ be a norm on $V$. Then there exists an inner product $\langle \, , \, \rangle$ on $V$ such that*

$$\langle v, v \rangle^{1/2} = \|v\| \quad \forall v \in V$$

*if and only if*

$$2\|u\|^2 + 2\|v\|^2 = \|u + v\|^2 + \|u - v\|^2 \quad \forall u, v \in V$$

**Definition 5** *Let $x, y$ be vectors in a normed linear space. The distance between $x, y$ is defined to be the number $\|y - x\|$*

## Orthonormal sets

Recall: the **standard basis** in $R^n$:

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \cdots, e_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix},$$

**Definition 6** *Let $V$ be an inner product space with inner product $\langle u, v\rangle$. Let $v_1, ..., v_n \in V$ be all nonzero. The set $\{v_1, ..., v_n\}$ is called orthogonal if $\langle v_i, v_j\rangle = 0\ \forall i \neq j$.*
*It is called **orthonormal** if $\|v_i\| = 1$ for $i = 1, 2, ..., n$*

$$\delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases} \quad \textit{Kronecker delta symbol}$$

**Note:** $\{v_1, ..., v_n\}$ is an orthonormal set iff[2] $\langle v_i, v_j\rangle = \delta_{ij}$

**Definition 7** *An orthonormal set of vectors is an orthogonal set of unit vectors*

**Theorem 5** ***orthogonal set and linearly independent*** *If $\{v_1, ..., v_n\}$ is an orthogonal set, then $v_1, ..., v_n$ are linearly independent*

**Note:** Suppose $(V, \langle\ ,\ \rangle)$ inner product space, let $\{u_1, ..., u_n\}$ be orthonormal set in $V$. Since $u_1, ...u_n$ are linearly independent and span the subspace $S := span(u_1, ..., u_n)$, the set $\{u_1, ..., u_n\}$ is a basis of $S$

**Definition 8** *We call $\{u_1, ..., u_n\}$ an orthonormal basis of $S$*

**Theorem 6** *Let $\{u_1, ..., u_n\}$ be an orthonormal basis of the inner product space $V$. Let $v \in V$ then*

$$v = \sum_{i=1}^{n} \langle v, u_i\rangle \cdot u_i$$

*$\langle v, u_i\rangle$ is the coordinates of $V$ with respect to the basis $\{u_1, ..., u_n\}$*

**Theorem 7** *Let $\{u_1, ..., u_n\}$ be an orthonormal basis of $V$. Let $u = \sum_{i=1}^{n} a_i u_i$ and $v = \sum_{i=1}^{n} b_i u_i$ be two vectors in $V$. Then the inner product of $u$ and $v$ is equal to*

$$\langle u, v\rangle = \sum_{i=1}^{n} a_i b_i{}^3$$

**Corollary 1** ***Formula of Parseval*** *Let $\{u_1, ..., u_n\}$ be an orthonormal basis of $V$ and $v = \sum_{i=1}^{n} c_i u_i$ any vector in $V$. Then it norm is equal to*

$$\|v\| = \left(\sum_{i=1}^{n} c_i^2\right)^{1/2}$$

*Note that the RHS was the old norm of the coordinate vector* $c = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} \in \mathbb{R}^n$

[2] if and only if

[3] scalar product $a^T b$

### Least squares problem

Recall the previous least square problem:
Given $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$, find $\hat{x} \in \mathbb{R}^n$ such that $\|A\hat{x} - b\|^2$ is minimal.

**More mathematical**: minimize the function $r : \mathbb{R}^n \to [0, \infty)$ defined by $r(x) = \|Ax - b\|^2$.

**Geometric interpretation:** find the orthogonal projection $p$ of $b$ onto the surface $R(A)$, by using the normal equation: $A^T A\hat{x} = A^T b$.
Under the additioanl assumption $rank(A) = n$ (i.e. $A$ is injective) we know that $A^T A$ non singular, in that case $\hat{x}$ is unique and given by $\hat{x} = (A^T A)^{-1} A^T b$.

if the columns of $A$ are an orthonormal set, then $A^T A = I$[4] so $\hat{x} = A^T b$ and the projection $p$ equal

$$p = AA^T b$$

this is called the **projector** onto $R(A)$

**Theorem 9** *Let $S$ be a finite dimensional subspace of $V$ and $\{u_1, ..., u_n\}$ an orthonormal basis of $S$. For $x \in V$ define*

$$p = \sum_{i=1}^{n} \langle x, u_i\rangle\, u_i$$

*then $p \in S$ and $x - p \in S^\perp$*

**Theorem 10** $p = \sum_{i=1}^{n} \langle x, u_i\rangle\, u_i$ *is the vector in $S$ that is closet to $x$. That is, for any $y \in S$ with $y \neq p$ we have*

$$\|x - y\| > \|x - p\|$$

**Definition 10** ***orthogonal projection*** *The vector $p$ is called the orthogonal projection of $x$ onto $S$*

**Corollary 2** *Let $S$ be a nonzero subspace of $\mathbb{R}^m$ and let $b \in \mathbb{R}^m$. If $\{u_1, ..., u_k\}$ is an orthonormal basis for $S$ and $U = (u_1, ..., u_k)$, then the projection $p$ of $b$ onto $S$ is given by*

$$p = UU^T b$$

**Note:** the last few theorems are use in many application, but most commonly are used for approximate a function. The process consist into project a function onto a subspace.

---

[4]

**Definition 9** *An $n \times n$ matrix is said to be orthogonal matrix if the column vectors of the matrix form an orthonormal set in $\mathbb{R}^n$*

**Theorem 8** *An $n \times n$ matrix $Q$ is orthogonal iff $Q^t Q = I$*

### Fourier approximation

**Definition 11** ***The nth order Fourier approximation of*** $f(x)$ *Let* $n$ *be a positive integer. A trygonometric function of degree* $n$ *is any function of the form*

$$t_n^*(x) = \frac{a_0}{2} + \sum_{k=1}^{n} a_k \cos(kx) + b_k \sin(kx)$$

*where* $a_i, b_i \in \mathbb{R}$. $a_0 = \frac{1}{\pi}\int_a^b f(x)dx$ *and* $a_k = \langle f, \cos(kx)\rangle = \frac{1}{\pi}\int_a^b f(x)cos(kx)dx$, $b_k = \langle f, \sin(kx)\rangle$

**Aim:** given $n$, find $a_i, b_i$ so that $\|f - t_n\|$ is minimal
The real numbers $a_i, b_i$ are called **the Fourier coefficients of** $f$

## Gram-Schmidt Orthogonalization

**General Problem:** Given an inner product space $(V, \langle\, , \,\rangle)$. Let $S$ be a finite dimensional subspace $V$. Find an orthonormal basis for it.

**Theorem 11** ***The Gram-Schidt Process*** *Let* $\{x_1, ..., x_n\}$ *be a basis for the inner product space* $V$. *Let*

$$u_1 = \left(\frac{1}{\|x_1\|}\right) x_1$$

*and define* $u_2, ..., u_n$ *recursively by*

$$u_{k+1} = \frac{1}{\|x_{k+1} - p_k\|}(x_{k+1} - p_k) \quad \text{for } k = 1, ..., n-1$$

*where*

$$\begin{aligned} p_k &= \langle x_{k+1}, u_1\rangle u_1 + \langle x_{k+1}, u_2\rangle u_2 + \cdots + \langle x_{k+1}, u_k\rangle u_k \\ &= \sum_{i=1}^{k} \langle x_{k+1}, u_i\rangle u_i \end{aligned}$$

*is the projection of* $x_{k+1}$ *onto* $Span(u_1, ..., u_k)$. *Then the set*

$$\{u_1, ..., u_n\}$$

*is an orthonormal basis for* $V$

**Theorem 12** *Let* $A \in \mathbb{R}^{n\times n}$ *be nonsingular. Then there exists an orthogonal matrix* $Q \in \mathbb{R}^{n\times n}$ *and an upper triangular matrix* $R \in \mathbb{R}^{n\times n}$ *with positive diagonal elements such that*

$$A = QR$$

**Theorem 13** ***QR factorization*** *Let* $A \in \mathbb{R}^{m\times n}$, *have rank(n), columns of* $A$ *is linearly independent. Then, there exists* $Q \in \mathbb{R}^{m\times n}$ *whose columns form an orthonormal set in* $\mathbb{R}^m$, *and* $R \in \mathbb{R}^{n\times n}$ *upper triangular with positive diagonal elements such that*

$$A = QR$$

**Remark:** $Q$ is called column-orthogonal. Here, follow QR factorization of a $\mathbb{R}^{3\times3}$ matrix in term of column vectors:

$$\begin{bmatrix} a_1 & a_2 & a_3 \end{bmatrix} = \begin{bmatrix} q_1 & q_2 & q_3 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ 0 & r_{22} & r_{23} \\ 0 & 0 & r_{33} \end{bmatrix}$$

**Theorem 14** *If $A$ is an $m \times n$ matrix of rank $n$, then the least squares solution of $Ax = b$ is given by $\hat{x} = R^{-1}Q^Tb$, where $Q$ and $R$ are the matrices obtained from the QR factorization. The solution $\hat{x}$ may be obtained by using back sobstitution to solve $Rx = Q^Tb$*

# Eigenvalues and eigenvectors

Remember the definition from LA1, here we will write only the new theorems, definition,facts or we rewrite previous theorem in a better way

For $A \in \mathbb{R}^{n\times n}$ the characteristic polynomial is $p(s) = det(A - sI)$. This is always a $n$th degree polynomial with real coefficients. In fact:

$$p(s) = (-1)^n s^n + p_{n-1}s^{n-1} + \cdots + p_1 s + p_0$$

with $p_i \in \mathbb{R}$.
**Fact:** $\lambda$ eigenvalue $\Leftrightarrow$ $\bar{\lambda}$ eigenvalue. Two explanations for this (only true if $A \in \mathbb{R}^{n\times n}$):

1. $p_A(s)$ is a real polynomial. If $\lambda$ root of $p_A(s) \Leftrightarrow \bar{\lambda}$ is a root of $p_A(s)$
2. $\lambda$ eigenvalue of $A$, then $Ax = \lambda x \Leftrightarrow \bar{A}x = \bar{\lambda x} \Leftrightarrow \bar{A}\bar{x} = \bar{\lambda}\bar{x} \Leftrightarrow A\bar{x} = \bar{\lambda}\bar{x}$ where $\bar{\lambda}$ eigenvalue, $\bar{x}$ eigenvector.

**Product:** Since $p(s) = (\lambda_1 - s)(\lambda_2 - s)\cdots(\lambda_n - s)$ and $p(s) = det(A - SI)$ we have

$$det(A) = p(0) = \lambda_1\lambda_2\cdots\lambda_n$$

**Sum:** for given $A$ let $tr(A) = \sum_{i=1}^n A_{ii}$ then the trace of $A$ it is also equal to the sum of eigenvalues: $tr(A) = \lambda_1 + \cdots + \lambda_n$

## Similarity of matrices

**Definition 12** $A, B \in \mathbb{R}^{n\times n}$, *or in $\mathbb{C}^{n\times n}$ if there exists a nonsingular $S \in \mathbb{C}^{n\times n}$ such that*

$$B = S^{-1}AS$$

**Recall:** If $\alpha : \mathbb{C}^n \to \mathbb{C}^n$ is a linear map and $\{v_1, v_2, ..., v_n\}$ a basis of $\mathbb{C}^n$ then the matrix of $\alpha$ with respect to $\{v_1, ..., v_n\}$ is defined as $A \in \mathbb{C}^{n\times n}$, $M = (a_{ij})$, where the $a_{ij}$ satisfy

$$\begin{aligned} \alpha(v_1) &= a_{11}v_1 + a_{21}v_2 + \cdots + a_{n1}v_n \\ \alpha(v_2) &= a_{12}v_1 + a_{22}v_2 + \cdots + a_{n2}v_n \\ &\vdots \qquad\qquad \vdots \\ \alpha(v_n) &= a_{1n}v_1 + a_{2n}v_2 + \cdots + a_{nn}v_n \end{aligned}$$

Now defined by $\alpha(x) = Ax$. So, $A$ and $B$ are similar means: they define one and the same linear map, or: they are matrices of one and the same linear map.

**Theorem 15** *Let $A, B \in \mathbb{R}^{n \times n}$. If $A$ and $B$ are similar, they have the same characteristic polynomial and the same eigenvalues.*

**Theorem 16** ***Cayley-Hamilton*** *Let $A$ be an $n \times n$ matrix, and let $p_A(t) = det(A - tI)$ be the corresponding characteristic polynomial. Then, $p_A(A) = 0$.*

**Note:**

1. $I, A, A^2, ..., A^n$ are linearly dependent, thus

$$p_A(A) = A^n + p_{n-1}A^{n-1} + \cdots + p_1 A + p_0 I = 0$$

## Hermitian matrices

**Definition 13** ***Field*** *set $k$ with two binary operations, called addition and multiplication, satisfying some axioms.*

**Complex scalar product:**

$$z = \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix}, \quad w = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} \to \sum_{i=1}^{n} \bar{z}_i w_i$$

**Note:** $\overline{z}^T$ is equal to $z^H$ and it is called Hermitian transponse

**Definition 14** ***Norm on*** *$\mathbb{C}^n$ induce by $z^H w$:*

$$\|z\|^2 = z^H z = \sum_{i=1}^{n} \bar{z}_i z_i = \sum_{i=1}^{n} |z_i|^2$$

**Definition 15** ***Complex inner product space*** *Let $V$ be a complex vector space. An inner product on $V$ is a map $\langle\, , \rangle : V \times V \to \mathbb{C}$ and satisfying the following conditions:*

1. $\langle z, z \rangle \geq 0 \; \forall z \in V$ *and* $\langle z, z \rangle = 0 \Leftrightarrow z = 0$
2. $\langle z, w \rangle = \overline{\langle w, z \rangle} \; \forall z, w \in V$
3. $\langle \alpha z + \beta w, u \rangle = \alpha \langle z, u \rangle + \beta \langle w, u \rangle \; \forall \alpha, \beta \in \mathbb{C} \; \forall z, w, u \in V$

All the results on real inner product spaces carry over to complex ones

- $z, w \in V$ orthogonal if $\langle z, w \rangle = 0$
- norm induced by inner product: $\|z\|^2 = \langle z, z \rangle$
- orthonormal set in $V$: $\{z_1, ..., z_n\}$ such that $\|z_i\| = 1 \; \forall i$ and $\langle z_i, z_j \rangle = 0 \; i \neq j$
- $\{z_1, ..., z_n\}$ are linearly independent

- Parseval theorem: if $\{w_1, ..., w_n\}$ orthonormal basis of $V$, and $z \in V$, then

$$z = \sum_{i=1}^{n} \langle z, w_i \rangle \, w_i$$

and

$$\|z\|^2 = \sum_{i=1}^{n} | \langle z, w_i \rangle |^2$$

**Definition 16** ***Norm*** *$V$ a complex vector space. A norm on $V$ is a map $\|.\| : V \to \mathbb{R}$ such that*

1. $\forall z \in \mathbb{C}$ $\|z\| \geq 0$ *and* $\|z\| = 0 \Leftrightarrow z = 0$
2. $\forall \alpha \in \mathbb{C}$ $\forall z \in V$ $\|\alpha z\| = |\alpha| \|z\|$
3. $\forall z, w \in V$ $\|z + w\| \leq \|z\| + \|w\|$

## Complex matrices

$\mathbb{C}^{m \times n}$: the complex vector space of all $m \times n$ matrices with entries in $\mathbb{C}$.

**Definition 17** ***Hermetian transponse*** *For given $M = (m_{ij})$ , $\overline{M} = (\overline{m_{ij}})$ Hermetian transponse of $M$:*

$$M^H = \overline{M}^T$$

**Inner product on $\mathbb{C}^{m \times n}$:**

$$\langle M, N \rangle = tr(M^H N) = \sum_{i=1}^{m} \sum_{j=1}^{n} \overline{m_{ij}} n_{ij}$$

**Basic Rules:**

- $(A^H)^H = A$
- $(\alpha A + \beta B)^H = \overline{\alpha} A^h + \overline{\beta} B^H$
- $(AB)^H = B^H A^H$

**Definition 18** ***Hermitian Matrix*** *A matrix $M$ is said to be* ***Hermitian*** *if $M = M^H$*

**Note:** every symmetric matrices is Hermitian. If $M$ is a matrix with real entries, then $M^H = M^T$

**Theorem 17** *Let $A \in \mathbb{C}^{n \times n}$ be Hermitian. Then all its eigenvalues are real. Eigenvectors belonging to* ***distinct*** *eigenvalues are* ***orthogonal***

**Definition 19** *$U \in \mathbb{C}^{n \times n}$ is called* ***unitary*** *if its columns form an* ***orthonormal set*** *in $\mathbb{C}^n$*

**Theorem 18** ***Schur's Theorem*** *For each $n \times n$ matrix $A$, there exists a unitary matrix $U$ such that $\Lambda$ is upper triangular*

**Theorem 19** ***Schur decomposition*** *Let $A \in \mathbb{C}^{n\times n}$. There exists a unitary $U \in \mathbb{C}^{n\times n}$ and an upper triangular $T \in \mathbb{C}^{n\times n}$ such that*

$$U^H AU = T$$

*or equivalently*

$$A = UTU^H$$

*called* ***Schur decomposition*** *or factorization*

**Theorem 20** *If $A$ is Hermitian then there exists a unitary $U$ that diagonalizes $A$ and a real diagonal matrix $\Lambda$ such that*

$$U^H AU = \Lambda$$

*or equivalently*

$$A = U\Lambda U^H$$

So, for a Hermitian matrix $A \in \mathbb{C}^{n\times n}$ there exists an othonormal basis $\mathbb{C}^n$ of eigenvectors of $A$

**Theorem 21** ***Real Schur Decomposition*** *Let $A \in \mathbb{R}^{n\times n}$. There exist an orthogonal $Q \in \mathbb{R}^{n\times n}$ and a quasi upper triangular matrix $T \in \mathbb{R}^{n\times n}$ such that*

$$A = QTQ^T$$

*where*

$$T = \begin{pmatrix} B_1 & \cdots & \cdots & \cdots \\ 0 & B_2 & \cdots & \cdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & B_k \end{pmatrix}$$

*and where $B_i$'s are either $1\times 1$ or $2\times 2$ matrices.*

We determined as follows:

- Compute all eigenvalues $\lambda_1, ..., \lambda_n$ of $A$
- Suppose $\lambda_1, ..., \lambda_r$ are not real and $\lambda_{r+1}, ..., \lambda_n$ are real
- $\lambda_1, ..., \lambda_r$ appears in complex conjugate pairs, says $\lambda_1, \overline{\lambda_1}, \lambda_2, \overline{\lambda_2}, ..., \lambda_{r/2}, \overline{\lambda_{r/2}}$
- Suppose $\lambda_j = a_j + ib_j$ and $\overline{\lambda_j} = a_j - ib_j$.
  This gives $r/2$ $2\times 2$ matrices $B_j = \begin{pmatrix} a_j & b_j \\ -b_j & a_j \end{pmatrix}$
- The remaining real $\lambda_{r+1}, ..., \lambda_n$ gives $n-r$ $1\times 1$ matrices $B_j = \lambda_j$

**Theorem 22** *Let $A \in \mathbb{R}^{n\times n}$ be symmetruc. There exists an orthogonal matrrix $Q \in \mathbb{R}^{n\times n}$ and a real diagonal matrix $\Lambda \in \mathbb{R}^{n\times n}$ such that*

$$A = Q\Lambda Q^T$$

**Definition 20** *Let $A \in \mathbb{C}^{n\times n}$. $A$ is called* ***normal*** *if*

$$A^H A = AA^H$$

**Theorem 23** *Let $A \in \mathbb{C}^{n\times n}$. There exixts unitary $U \in \mathbb{C}^{n\times n}$ and a diagonal matrix $\Lambda \in \mathbb{C}^{n\times n}$ such that $A = U\Lambda U^H$ if and only if $A$ is normal*

## Singular value decomposition

Let $A \in \mathbb{R}^{m\times m}$. For the moment let assume $n \leq m$ ($A$ is tall). We call $A$ **rank deficient** if $rank(A) < n$.
Often we want to know: how close is $A$ from being rank deficient.

**Theorem 24** *Let $A \in \mathbb{R}^{m\times m}$, $n \leq m$. There exist **orthogonal matrices** $U \in \mathbb{R}^{m\times m}$ and $V \in \mathbb{R}^{m\times m}$ and real numbers*

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0$$

*Such that $A = U\Sigma V^T$ and $\sigma_i = \sqrt{\lambda_i}$ for $i = 1, .., n$, where*

$$\Sigma = \begin{pmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \sigma_n \\ -- & -- & -- & -- \\ & 0_{(m-n)\times n} & & \end{pmatrix}$$

*Moreover, if $rank(A) = r$ then*

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$$
$$\sigma_{r+1} = \sigma_{r+2} = \cdots = \sigma_n = 0$$

*The numbers of $\sigma_i$ are **unique**. They are called the **singular values of** $A$.*

**Idea:** The smallest singular values quantify how close $A$ is to lose a rank.

**Some observations on the SVD**

1. Given $A \in \mathbb{R}^{m\times n}$ ($m \geq n$) its singular values $\sigma_1, ..., \sigma_n$ are **unique** since $\sigma_1^2, ..., \sigma_n^2$ are the eigenvalues of $A^TA$
2. $U$ and $V$ are **not unique**
3. Since $V^TA^TAV = \Lambda$ with $\sigma_1^2, ..., \sigma_n^2$ eigenvalues, the columns of $V$ are always eigenvectors of $A^TA$
4. Since $AA^T = U \begin{pmatrix} \Sigma_1 \\ 0 \end{pmatrix} V^TV \begin{pmatrix} \Sigma_1 & 0 \end{pmatrix} U^T = U \begin{pmatrix} \Sigma_1^2 & 0 \\ 0 & 0 \end{pmatrix} U^T$ the columns of $U$ are always eigenvectors of $AA^T$
5. if $A$ has rank $r$, then
   (a) $v_1, ..., v_r$ form an orthonormal basis for $R(A^T)$
   (b) $v_{r+1}, ..., v_n$ form an orthonormal basis for $N(A)$
   (c) $u_1, ..., u_r$ form an orthonormal basis for $R(A)$
   (d) $U_{r+1}, ..., u_n$ forma an orthonormal basis for $N(A^T)$
6. let $A \in \mathbb{R}^{m\times n}$ and $AV = U\Sigma$, then $A^Tu_j = \sigma_j v_j$ for $j = 1, ..., n$ and $A^Tu_j = 0$ for $j = n+1, ..., m$

**Theorem 25** ***General*** *Let $A \in \mathbb{R}^{m\times n}$ $rank(A) = r$ ($\leq min(n,m)$). There exist $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$ and orthogonal $U \in \mathbb{R}^{m\times m}$, $V \in \mathbb{R}^{n\times n}$ such that*

$$A = U \begin{pmatrix} \Sigma_1 & 0_{r\times n-r} \\ 0_{m-r\times r} & 0_{m-r\times n-r} \end{pmatrix} V^T$$

**Remark:** If $rank(A) = n$ (injective) then the 0 matrices on the right are absent. If $rank(A) = m$ surjective then the 0 matrices on the bottom are absent. If $A$ is square nonsingular, then all 0 matrices are absent

**Application**: Given $A \in \mathbb{R}^{m\times n}$, $rank(A) = r$, and $0 \leq k \leq r$, compute the distance of $A$ to $M_k$ define by

$$d(A, M_k) = inf\{\, \|A - S\|_F \,|S \in M_k\}$$

$\|.\|_F$ is the Frobenious norm on $\mathbb{R}^{m\times n}$ which is define by

$$\|M\|_F = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} m_{ij}^2} = \sqrt{tr(M^T M)}$$

Also find $X \in M_k$ such that $d(A, M_k) = \|A - X\|_F$

**Theorem 26** *Let $A \in \mathbb{R}^{m\times n}$ and $k < r = rank(A)$. Let $M_k = \{S \in \mathbb{R}^{m\times n} \,|\, rank(S) \leq k\}$. Let $\sigma_1, ..., \sigma_r$ be the non-zero eigenvalues of $A$, then*

$$d(A, M_k) = \sqrt{\sigma_{k+1}^2 + \cdots + \sigma_r^2}$$

*Let $A = U\Sigma V^T$ be the SVD of $A$. Then $X = U\Sigma V^T$ where $\Sigma$ have now the $\sigma_{k+1}, ...., \sigma_r = 0$. Then $X \in M_k$ and*

$$\|A - X\|_F = d(A, M_k)$$

*$X$ is called the best approximation in $M_k$ of $A$*

## Quadratic Form

**Definition 21** ***A quadratic equation*** *in two variables $x, y$ is an equation of the form*

$$ax^2 + 2bxy + cy^2 + dx + ey + f = 0 \tag{1}$$

*Equation (1) may be rewritten in the form*

$$\begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} a & b \\ b & c \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} d & e \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + f = 0$$

*Let*

$$x = \begin{pmatrix} x \\ y \end{pmatrix} \quad A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$$

*The term*

$$x^T A x = ax^2 + 2bxy + cy^2$$

*is called the **quadratic form** associated with (1)*

**Definition 22** *Let $A \in \mathbb{R}^{n\times n}$ be symmetric then the expression $x^T Ax$ with $x \in \mathbb{R}^n$ is called* ***the quadratic form associated with*** *$A$.*
*It is also define a* ***quadratic function***

$$F : \mathbb{R}^n \to \mathbb{R}^n \quad F(x) = x^T Ax$$

**Definition 23** *Given $F : \mathbb{R}^n \to \mathbb{R}$, a point $x_0 \in \mathbb{R}^n$ is called* ***stationary*** *if $\frac{\partial F}{\partial x_i}(x_0) = 0$ for $i = 1, 2, ..., n$*

**Definition 24** *Let $A \in \mathbb{R}^{n\times n}$ be symmetric*

1. *$A$ is* ***positive definite*** *if $x^T Ax > 0$ $x \neq 0$ (denoted by $A > 0$)*
2. *$A$* ***negative definite*** *if $x^T Ax < 0$ $x \neq 0$ (denoted by $A < 0$*
3. *$A \geq 0$, or $A$ is* ***positive semi-definite*** *if $x^T Ax \geq 0$ $\forall x \in \mathbb{R}^n$*
4. *$A \leq 0$, or $A$ is* ***negative semi-definite*** *if $x^T Ax \leq 0$ $\forall x \in \mathbb{R}^n$*
5. *$A$* ***indefinite*** *if $x^T Ax$ taken both positive as well as negative real values*

**Theorem 27** *Let $A \in \mathbb{R}^{n\times n}$ be symmetric and $\lambda_1, \lambda_2, ..., \lambda_n$ be its real eigenvalues. We have:*

1. *$A > 0 \Leftrightarrow \lambda_i > 0$ $i = 1, 2, ..., n$*
2. *$A < 0 \Leftrightarrow \lambda_i < 0$ $i = 1, 2, ..., n$*
3. *$A \geq 0 \Leftrightarrow \lambda_i \geq 0$ $i = 1, 2, ..., n$*
4. *$A \leq 0 \Leftrightarrow \lambda_i \leq 0$ $i = 1, 2, ..., n$*
5. *$A$ is indefinite $\Leftrightarrow$ $A$ has eigenvalues of different sign*

**Definition 25** *Let $F : \mathbb{R}^n \to \mathbb{R}$ be $C^2$. Let $x_0$ be stationary point, the* ***Hessian matrix*** *at $x_0$ is*

$$H(x_0) = \begin{pmatrix} F_{x_1x_1}(x_0) & F_{x_1x_2}(x_0) & \cdots & F_{x_1x_n}(x_0) \\ F_{x_2x_1}(x_0) & F_{x_2x_2}(x_0) & & \vdots \\ \vdots & & \ddots & \vdots \\ F_{x_nx_1} & \cdots & \cdots & F_{x_nx_n}(x_0) \end{pmatrix}$$

**Theorem 28** *Let $F : \mathbb{R}^2 \to \mathbb{R}$ be twice continuosly differentiable and $\begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$ be stationary point. Then*

1. *$H(x_0, y_0) < 0 \Rightarrow \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$* ***local maximum***
2. *$H(x_0, y_0) > 0 \Rightarrow \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$* ***local minimum***
3. *$H(x_0, y_0)$ indefinite $\Rightarrow \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$* ***saddle point***

### Positive definite matrices

Some necessary conditions are:

1. $A$ is positive definite $\Rightarrow$ $A$ is nonsingular
2. $A$ is positive definite $\Rightarrow$ $det(A) > 0$

**Theorem 29** *Let $A \in \mathbb{R}^{n\times n}$ be symmetric. Then the following are equivalent*

1. *$A$ is **positive definite***
2. *$det(A_r) > 0$ for $r = 1, 2, ..., n$*
3. *$A$ can be reduced to upper triangular form using only type III row operations, and the pivot elements will all be positive*
4. *There exist a **lower triangular matrix** $L$ **with positive diagonal elements** such that*
$$A = LL^T$$
*called it **Cholesky decomposition***
5. *There exist **nonsingular** matrix $B$ such that*
$$A = B^T B$$

## Jordan canonical/Normal form

The aim is to find a nonsingular $T \in \mathbb{C}^{n\times n}$ such that $T^{-1}AT = J$,where $A \in \mathbb{C}^{n\times n}$ and $J$ has a simple form.
The simple matrix $Y$ is the matrix pf the linear map $A : \mathbb{C}^n \to \mathbb{C}^n$ with respect to the basis $\{t_1, t_2, ..., t_n\}$.
A psecific case is the diagonal matrix, but not all matrix are diagonalizable.

**Theorem 30** *Let $A \in \mathbb{C}^{n\times n}$. There exists a nonsingular $T \in \mathbb{C}^{n\times n}$ such that $T^{-1}AT = J$, with*

$$J = \begin{pmatrix} j_1 & 0 & \cdots & 0 \\ 0 & j_2 & 0 & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & j_n \end{pmatrix}$$

*a blockdiagonal matrix with diagonal blocks $j_i$ of the form*

$$j_i = \lambda I + N$$

*where $I$ is an identity matrix and $N$ has the form*

$$N = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

*and $\lambda$ is an eigenvalue of $A$. The matrix $J$ is called the Jordan Normal form of $A$.*

**Note:**

- $J$ is just the matrix of $A : \mathbb{C}^n \to \mathbb{C}^n$ with respect to a suitable basis of $\mathbb{C}^n$
- It is possible that a block $j_i$ is $1 \times 1$. In that case it has the form $j_i = \lambda$, where $\lambda$ is an eigenvalue of $A$.

**Theorem 31** ***Dimension theorem*** *For a linear map $A : V \to V$ we have*

$$dim\, V = dim\, N(A) + dim\, R(A)$$

**Note:** The $dim\, N(A - \lambda I$ is equal to the number of linearly independent eigenvectors associeted with $\lambda$

**Definition 26** *Let $\lambda_1, \lambda_2, ..., \lambda_k$ be the distict eigenvalues of $A$. We will denote*

$$g_i = dim\, N(A - \lambda_i I)$$

*The integers $g_1, g_2, ...,_k$ are called **the geometric multiplicities** of the eigenvalues $\lambda_1, \lambda_2, ..., \lambda_k$*

**Definition 27** *Suppose $\lambda_1, \lambda_2, ..., \lambda_k$ are the distinct eigenvalues of $A$, then the characteristic polynomial is given by*

$$p_A(z) = (-1)^n (z - \lambda_1)^{a_1} (z - \lambda_2)^{a_2} \cdots (z - \lambda_k)^{a_k}$$

*for certain positive integers $a_i$ such that $a_1 + a_2 + \cdots + a_k = n$*
*The integer $a_i$ is called **agebraic multiplicity** of the eigenvalue $\lambda_i$ for $i = 1, ..., k$*

**Note:** $g_i \leq a_i$ for $i = 1, ..., k$

**Definition 28** *Let $V$ be a (complex) vector space, and let $V_1, ..., V_k$ be linear subspaces of $V$. We define their **sum** as*

$$V_1 + V_2 + \cdots + V_k = \{x_1 + x_2 + \cdots + x_k \in V \mid x_i \in V_i\}$$

**Definition 29** *we will write*

$$V = V_1 \oplus V_2 \oplus \cdots \oplus V_k$$

*if*

1. $V =_1 + V_2 + \cdots + V_k$
2. *every $x \in V$ can be written as $x = x_1 + \cdots + x_k$ with $x_i \in V_i$ in exactly one way,*

*i.e: $x = x_1 + x_2 + \cdots + x_k$ with $x_i \in V_i$ and $x = \overline{x_1} + \cdots + \overline{x_k}$ with $\overline{x_i} \in V_i \Rightarrow x_i = \overline{x_i}$ for $i = 1, ...., k$.*
*We then say that $V$ is the **direct sum** of $V_1, ..., V_k$*

**Definition 30** *Let $W \subset V$ is a subspace. We call $W$ **A-invariant** if all $x \in W$ we have $Ax \in W$. We write*

$$AW \subset W$$

Now the jordan form under the condition $a_i = g_i$ $(i = 1, ..., k)$, that is $A$ is diagonalizable since it have lineraly independent eigenvectors, is given by:

$$J = \begin{pmatrix} \lambda_1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & \lambda_1 & & & & & & \\ & & & \lambda_2 & & & & & \\ & & & & \ddots & & & & \\ & & & & & \lambda_2 & & & \\ & & & & & & \ddots & & \\ & & & & & & & \lambda_k & \\ & & & & & & & & \ddots & \\ & & & & & & & & & \lambda_k \end{pmatrix}$$

that is all Jordan blocks $j_i$ are $1 \times 1$, and these blocks are $\lambda_1(a_1$ times $)$, $\lambda_2(a_2$ times $)$ up to $\lambda_k(a_k$ times $)$.

**Definition 31** *Let $q(z)$ be a polynomial. We say that **annihilates** $A$ if $q(A) = 0$*

**Note:** By the CH theorem, $p_A(z)$ annihilates $A$. For a given $A$ there are many polynomials with this property

**Theorem 32** *Let $A \in \mathbb{C}^{n \times n}$. There exist exactly one **monic polynomial** $p_{min}(z)$ of **minimum degree** that annihilates $A$*

**Note:**

- $p_{min}(z)$ is called minimum polynomial of $A$.
- Monic means that the coeficcient of the higher power of $z$ is 1
- the minimum polynomial is unique

**Corollary 3** *Let $p(z)$ be any polynomial tha annihilates $A$. Then $p_{min}(z)$ is a divisor of $p(z)$, i.e. there eixts $q(z)$ such that*

$$p(z) = q(z)p_{min}(z)$$

**Theorem 33** *Let $A \in \mathbb{C}^{n \times n}$. Every eigenvalue of $A$ is a root of $p_{min}(z)$. Conversely, every root of $p_{min}(z)$ is an eigenvalue of $A$*

Let $A \in \mathbb{C}^{n \times n}$ and $\lambda_1, ..., \lambda_k$ be distinct eigenvalues, then $p_{min}(z)$ can be written as

$$p_{min}(z) \; = \; (z - \lambda_1)^{m_1}(z - \lambda_2)^{m_2} \cdots (z - \lambda_k)^{m_k}$$

for certain positive integers $m_1, ..., m_k$

**Theorem 34** *Let $A \in \mathbb{C}^{n \times n}$ and let $\lambda_1, ..., \lambda_k$ be distinct eigenvalues. Let $m_1, ..., m_k$ be the integers for $i = 1, ..., k$ define subspace*

$$V_i = N\left((A - \lambda_i I)^{m_i}\right)$$

*Then we have*

*1.* $V_i$ *is* $A-$*invariant*

*2.* $\mathbb{C}^{n\times n} = V_1 \oplus V_2 \oplus \cdots \oplus V_k$

**Summary**:

1. The total size of the blocks corresponding to an eigenvalue $\lambda$ is equal to $a$, the algebraic multiplicity

2. The number of Jordan blocks corresponding to an eigenvalue $\lambda$ is equal to $g$, its geometric multiplicity

3. the maximum size of blocks corresponding to $\lambda$ is qual to $m$, its multiplicity as a root of the minimal polynomial