Linear Algebra 2 Notes (2023/2024)

Griffin Reimerink

Contents

1 Vector spaces

Definition *Vector space*

Consider a **field** $\mathbb{F}$ of scalars and a set V with elements called **vectors**.
We define two operations on V:

- $\oplus : V \times V \to V$ Vector addition
- $\odot : \mathbb{F} \times V \to V$ Scalar multiplication

V is a **vector space** over $\mathbb{F}$ if and only if these axioms hold:

1. There exists an element $\mathbf{0}$ (the **zero vector**) in V such that $\mathbf{0} \oplus u = u$ for all $u \in V$.
2. Vector addition is commutative: $u \oplus v = v \oplus u$
3. Vector addition is associative: $u \oplus (v \oplus w) = (u \oplus v) \oplus w$
4. For each $u \in V$, there exists $z \in V$ (the **additive inverse**) such that $u \oplus z = \mathbf{0}$
5. For all $u \in V$, $1 \odot u = u$
6. For all $a, b \in \mathbb{F}$ and all $u \in V$, $a \odot (b \odot u) = (ab) \odot u$
7. For all $c \in \mathbb{F}$ and all $u, v \in V$, $c \odot (u \oplus v) = (c \odot u) \oplus (c \odot v)$
8. For all $a, b \in \mathbb{F}$ and all $u \in V$, $(a + b) \odot u = (a \odot u) \oplus (b \odot u)$

Theorem

Let V be a vector space and let $u \in V$. Then,

1. The zero vector $\mathbf{0}$ is unique.
2. $0 \odot u = \mathbf{0}$
3. The additive inverse z of u is unique.
4. $-1 \odot u = z$

Alternative notation:
additive inverse $-u$ zero vector 0 vector addition $u + v$ scalar multiplication cu

Definition *Subspace*

Let V be a vector space over F. $U \subseteq V$ is a **subspace** of V if
U is a vector space over F with the same vector addition and scalar multiplication operations as V.

Only axioms 1 and 4 have to be checked to see if U is a subspace.

Theorem

Let V be a vector space over $\mathbb{F}$ and $U \subseteq V$. Then U is a subspace of V if and only if:

1. U is nonempty
2. $u + v \in U$ for all $u, v \in U$
3. $cu \in U$ for all $c \in \mathbb{F}$ and $u \in U$

1.1 Linear combinations

Definition *Linear combination*

Let V be a vector space over $\mathbb{F}$.
For given $c_1, c_2, \ldots, c_2 \in \mathbb{F}$ and $v_1, v_2 \ldots, v_r \in V$, we say that an expression of the form $c_1v_1 + c_2v_2 + \ldots + c_rv_r$ is a **linear combination** of $v_1, v_2 \ldots, v_r$.

Definition *Span*

The set of all linear combinations of $v_1, v_2, \ldots, v_r \in V$ is called the **span** of $v_1, v_2, \ldots, v_r$.
We denote it by $\text{span}(v_1, v_2 \ldots, v_r) = \{c_1v_1 + c_2v_2 + \ldots + c_rv_r \mid c_1, c_2 \ldots, c_r \in \mathbb{F}\}$

Theorem

Let V be a vector space and $v_1, v_2, \ldots, v_r \in V$. Then $\text{span}(v_1, v_2, \ldots, v_r)$ is a subspace of V.

Definition *Spanning set*

We say that $v_1, v_2, \ldots, v_r$ is a **spanning set** for V (or $v_1, v_2, \ldots, v_r$ **span** V)
if every vector in V can be written as a linear combination of $v_1, v_2, \ldots, v_r$.

Theorem

1. If $v_1, v_2, \ldots, v_r$ span V and one of these vectors is a linear combination of the others, then these $r-1$ vectors span V.
2. One of the vectors $v_1, v_2, \ldots, v_r$ is a linear combination of the others if and only if there exist scalars $c_1, c_2, \ldots, c_r \in \mathbb{F}$, not all zero, such that $c_1 v_1 + c_2 v_2 + \ldots + c_r v_r = 0$

Definition *Linear dependence and independence*

The vectors $v_1, v_2, \ldots, v_r$ are **linearly dependent** if there exist scalars $c_1, c_2, \ldots, c_r \in \mathbb{F}$, not all zero, such that $c_1 v_1 + c_2 v_2 + \ldots + c_r v_r = 0$. Otherwise, they are **linearly independent**.

Theorem

Let $v_1, v_2, \ldots, v_r \in V$ be vectors. Every vector in $\operatorname{span}(v_1, v_2, \ldots, v_r)$ can be written uniquely as a linear combination of $v_1, v_2, \ldots, v_r$ if and only if $v_1, v_2, \ldots, v_r$ are linearly independent.

1.2 Subspace operations

Theorem

If U, W are subspaces of a vector space V, then $U \cap W$ is a subspace of V.

This is not necessarily true for $U \cup W$.

Definition *Sum of subspaces*

The **sum** of two subspaces U and W of V is defined as $U + W = \{u + w \mid u \in U \text{ and } w \in W\}$

Theorem

If U, W are subspaces of a vector space V, then $U + W$ is a subspace of V.

Definition *Direct sum of subspaces*

The sum of two subspaces U, W is called a **direct sum** if $U \cap W = 0$

Every vector in a direct sum $U + W$ can be uniquely expressed as a sum of $u \in U$ and $w \in W$.

1.3 Bases

Definition *Basis*

Let V be a vector space over F. The vectors $v_1, v_2, \ldots, v_n$ form a **basis** of V if

1. $v_1, v_2, \ldots, v_n$ are linearly independent
2. $\operatorname{span}(v_1, v_2, \ldots, v_n) = V$

Bases are not unique.

Lemma *Replacement lemma*

Let V be a nonzero vector space over F and let r be a positive integer.

Suppose that $u_1, u_2, \ldots, u_r$ span V. Let $v = \sum_{i=1}^{r} c_i u_i$ be a nonzero vector in V. Then,

1. $c_j \neq 0$ for some $j \in \{1, 2, \ldots, r\}$.
2. If $c_j \neq 0$ then $v, u_1, \ldots, u_{j-1}, u_{j+1} \ldots, u_r$ span V.
3. If $\{u_1, u_2, \ldots, u_r\}$ is a basis for V then $\{v, u_1, \ldots, u_{j-1}, u_{j+1} \ldots, u_r\}$ is also a basis for V.

Theorem

Let n and r be positive integers and let V be a vector space over F.

Let $\{v_1, v_2, \ldots, v_n\}$ be a basis of V and let $u_1, u_2, \ldots, u_r$ be linearly independent. Then,

1. $r \leq n$
2. If $r = n$, then $\{u_1, u_2, \ldots, u_r\}$ is a basis of V.

Theorem

If $\{v_1, v_2, \ldots, v_n\}$ and $\{u_1, u_2, \ldots, u_r\}$ are bases of V, then $n = r$.

Definition *Dimension*

Let V be a vector space and let n be a positive integer.

- If $\{v_1, v_2, \ldots, v_n\}$ is a basis of v, then n is the **dimension** of V.
- If $V = \{0\}$, then V has dimension zero.
- If V has dimension n for some $n \in \mathbb{N}$, then V is finite dimensional (notation: $\dim V = n$)
- Otherwise, V is infinite dimensional.

If $\operatorname{span}(v_1, v_2, \ldots, v_r) = V$, then $\dim V \leq r$

If V is finite dimensional, every set of linearly independent vectors in V can be extended to a basis.

If V is finite dimensional and U is a subspace of V, then $\dim U \leq \dim V$.

The set of sequences in $\mathbb{R}$ is an example of an infinite-dimensional vector space.

Theorem *Grassman's formula*

$$\dim(U \cap W) + \dim(U + W) = \dim U + \dim W$$

1.4 Linear transformations

Definition *Linear transformation*

Let V and W be vector spaces over the same field $\mathbb{F}$.
A function $T : V \to W$ is called a **linear transformation** (a **linear operator** if $V = W$) if

$$T(au + bv) = aT(u) + bT(v) \quad \forall a, b \in \mathbb{F} \quad \forall u, v \in V$$

Let $A \in \mathbb{F}^{n \times m}$ and $T(x) = Ax$. Then T is a linear transformation from $\mathbb{F}^m$ to $\mathbb{F}^n$.

Derivatives and integrals are linear transformations.

Lemma

Let V and W be vector spaces over $\mathbb{F}$ and let $T : V \to W$ be a linear transformation. Then,

1. $T(cv) = cT(v)$ for all $c \in \mathbb{F}$ and $v \in V$
2. $T(0) = 0$ (the first 0 is in V, the second 0 is in W)
3. $T(a_1v_1 + a_2v_2 + \ldots + a_nv_n) = a_1T(v_1) + a_2T(v_2) + \ldots + a_nT(v_n) \quad \forall v_k \in V \quad \forall a_k \in \mathbb{F}$

Definition *Kernel and range*

Let $T : V \to W$ be a linear transformation.

- The **kernel** of T is defined as $\ker T = \{v \in V \mid T(v) = 0\}$
- The **range** of T is defined as $\operatorname{ran} T = \{w \in W \mid \exists v \in V \quad \text{s.t.} \quad T(v) = w\}$

Theorem

The kernel of T is a subspace of V and the range of T is a subspace of W.

Any element of an n-dimensional vector space can be represented by the
***E*-coordinate vector** $\langle c_1, c_2, \ldots, c_n \rangle \in \mathbb{F}^n$ by fixing an ordered basis E.

Definition *E-basis representation*

Let $E = (v_1, v_2, \ldots, v_n)$ be an ordered basis of V.
For any $u \in V$, write $u = c_1v_1 + c_2v_2 + \ldots + c_nv_n$ where $c_k \in \mathbb{F}$.
The function $[\,\cdot\,]_E : V \to \mathbb{F}^n$, defined by $[u]_E = \langle c_1, c_2, \ldots, c_n \rangle$ is called the ***E*-basis representation**.

Theorem

$[\,\cdot\,]_E : V \to \mathbb{F}^n$ is a linear transformation.

Theorem *Isomorphism V and $\mathbb{F}^n$*

$[\cdot]_E$ is a bijection between V and $\mathbb{F}^n$.
We say that $[\cdot]_E : V \to \mathbb{F}^n$ is a (linear) **isomorphism** and V and $\mathbb{F}^n$ are **isomorphic**.

Theorem *Dimension theorem*

Let $T : V \to W$ be a linear transformation. Then $\dim \ker(T) + \dim \operatorname{ran}(T) = \dim V$

The range of a matrix is equal to the span of its columns.

Lemma

Let $T : V \to W$ be a linear map. Then $\ker(T) = \{0\}$ if and only if T is injective.

Lemma

Let $T : V \to W$ be a linear map with $\dim V = \dim W < \infty$. Then T is injective $\iff$ T is surjective.

1.5 Matrix representations

Definition *Matrix representation of a linear transformation*

Let V be a n-dimensional vector space over F with basis β and linear transformation β.
Then the **matrix representation** with respect to β is the matrix ${}_\beta[T]_\beta$
where each column is a vector of the basis with T applied to it.

Every n-dimensional vector space (mapped to $\mathbb{F}^n$) has the **standard basis** $\left\{ \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \dots, \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} \right\}$

Definition *Change of basis matrix*

The **change of basis matrix** from β to γ is the $n \times n$ matrix ${}_\beta[I]_\gamma$
whose columns are the elements of β expressed in γ.

Theorem

Let V be a finite dimensional vector space with bases β, γ and linear transformation T.
Let $S =_\gamma [I]_\beta$. Then ${}_\beta[I]_\gamma = S^{-1}$ and ${}_\gamma[T]_\gamma = S\, {}_\beta[T]_\beta\, S^{-1}$

Definition *Similarity*

Two matrices A and B are **similar** if and only if $A = Q^{-1}BQ$ for some Q.

2 Inner product spaces

Definition *Inner product space over $\mathbb{R}$*

An **inner product space** over $\mathbb{R}$ is an $\mathbb{R}$-vectorspace V together with a map

$$V \times V \to \mathbb{R} : (v_1, v_2) \mapsto \langle v_1, v_2 \rangle$$

This map satisfies the axioms of an inner product:
1. **Linearity** For any fixed v_2, the map $V \to \mathbb{R} : v \in V \mapsto \langle v, v_2 \rangle$ is linear.
2. **Symmetry** $\langle v_1, v_2 \rangle = \langle v_2, v_1 \rangle$ for all $v_1, v_2 \in V$
3. **Positivity** $\langle v, v \rangle \geq 0$ and if $v \neq 0, \langle v, v \rangle > 0$

Definition *Hermitian inner product space (Inner product space over $\mathbb{C}$)*

An **inner product space** over $\mathbb{C}$ is a $\mathbb{C}$-vectorspace V together with a map

$$V \times V \to \mathbb{C} : (v_1, v_2) \mapsto \langle v_1, v_2 \rangle$$

This map satisfies the following axioms:

1. **Linearity** For any fixed v_2, the map $V \to \mathbb{C} : v \in V \mapsto \langle v, v_2 \rangle$ is linear. The map is not necessarily linear if v_1 is fixed instead of v_2.
2. **Symmetry** $\langle v_1, v_2 \rangle = \overline{\langle v_2, v_1 \rangle}$ for all $v_1, v_2 \in V$
3. **Positivity** $\langle v, v \rangle \in \mathbb{R}$ and if $v \neq 0, \langle v, v \rangle > 0$

Notation: $v_1 \perp v_2$ means that v_1 and v_2 have inner product 0.

The standard inner product for $\mathbb{C}^n$ is $\begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} \cdot \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} = \sum_{j=1}^{n} a_j \bar{b}_j$

2.1 Norms

Definition *Norm*

Let V be an $\mathbb{F}$-vectorspace.
A vector-space **norm** on V is a map $\| \cdot \| : V \to \mathbb{R}$ where $\|v\|$ is the norm of $v \in V$.
This map has to satisfy the following conditions for all $v \in V$:

1. **Non-negativity** $\|v\| \geq 0$
2. **Positivity** $\|v\| = 0 \iff v = 0$
3. **Homogeneity** $\|\lambda v\| = |\lambda| \cdot \|v\| \quad \forall \lambda \in \mathbb{F}$
4. **Triangle inequality** $\|v + w\| \leq \|v\| + \|w\|$

Theorem

Let V be an inner product space. Then we can define a norm as $\|v\| = \sqrt{\langle v, v \rangle}$

Theorem *Cauchy-Schwarz-Bunyakovsky inequality*

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

2.2 Orthonormal vectors

Definition *Orthogonal system*

An orthogonal system in a vector space V is a set of n vectors $v_1, \ldots, v_n \in V$ such that $\langle v_i, v_j \rangle = 0$ if and only if $i \neq j$

Definition *Orthonormal system*

An orthonormal system in a vector space V is a set of n vectors $v_1, \ldots, v_n \in V$ such that $\langle v_i, v_j \rangle = \begin{cases} 1 \text{ if and only if } i = j \\ 0 \text{ if and only if } i \neq j \end{cases}$

Lemma

The vectors in any orthogonal system form an independent set.

Definition *Orthonormal basis*

We call $\{v_i\}_{i \in I} \subseteq V$ an **orthonormal basis** iff $\{v_i\}_{i \in I}$ is an orthonormal system that spans V.

Theorem

Let $\{v_i\}_{i \in I}$ be a basis of V. Take any vector $v \in V$, which can be written uniquely as $a_{i_1} v_{i_1} + \ldots + a_{i_m} v_{i_m}$
Then $\langle v, v_j \rangle = \begin{cases} a_j \text{ if } j \in \{i_1, \ldots, i_m\} \\ 0 \text{ otherwise.} \end{cases}$

Algorithm *Gram-Schmidt process*

Let $\dim V < \infty$. Let $v_1, v_2, \ldots, v_n$ be a basis of V.
Define the vectors $u_1, u_2, \ldots, u_n$ recursively by $u_1 = \frac{v_1}{||v_1||}$ and $u_k = \frac{x_k}{||x_k||}$
where $x_k = v_k - \sum_{i=1}^{k-1} \langle u_i v_k \rangle u_i$ for all $k \leq n$.
Then, $\{u_1, u_2, \ldots, u_n\}$ is an orthonormal basis of V.

2.3 Dual vector spaces

Definition *Dual vector space*

$L(V, W)$ is the set of linear maps from V to W.
Let V be an $\mathbb{F}$-vectorspace. The **dual vector space** of V is denoted V^* or $L(V, \mathbb{F})$.
The elements of V^* are called **functionals**.

Let V, W be an $\mathbb{F}$-vectorspace. Then $L(V, W)$ is also an $\mathbb{F}$-vectorspace.

If $V = \mathbb{F}^n$, then $V^* = \mathbb{F}^{1 \times n}$.

Inner products of $\mathbb{R}$-vectorspaces are functionals.

Theorem

If V has $\dim V = n < \infty$, then $V^* \cong V$ ($\cong$ means isomorphic)
Any basis for V yields such an isomorphism.

Theorem *Riesz representation theorem*

If $\dim V < \infty$, then the map $V \to V^*$ is bijective.

Definition *Adjoint*

Let $T \in L(W, V)$. Then T^* (or T^{adj}) is a function such that $\langle \langle T^*(\cdot), v \in V \rangle \rangle = \langle \cdot, T^*(v) \rangle$
T^* is a linear map from V to W.

T^* is the conjugate transpose of T.

Definition *Self-adjoint*

A linear transformation is **self-adjoint** or **Hermitian** if $T^* = T$

2.4 Orthogonal complements

Definition *Orthogonal complement*

Let V be an inner product space and $W \subseteq V$. Then $W^\perp := \{v \in V : \langle v, w \rangle = 0 \ \ \forall w \in W\}$

If W is a subspace of V, then $W^\perp$ is a subspace of V.

$V^\perp \cap V = \{0\}$

Theorem

If V is finite-dimensional, then $(V^\perp)^\perp = V$. Otherwise, $V \subseteq (V^\perp)^\perp$.

Theorem

If V is finite-dimensional and $W \subseteq V$, then $\dim W + \dim W^\perp = \dim V$.

Theorem

If U is a finite-dimensional subspace of V, then V is the direct sum of U and $U^\perp$.

Definition *Orthogonal projection*

Let $V \subseteq W$. Every $w \in W$ can be written uniquely as $v \in V + v' \in V^\perp$ where v is the **orthogonal projection** of w. (notation: $P_v w$)

Theorem

$$\ker T = (\operatorname{image} T^*)^\perp$$

Theorem *Minimum norm solution*

Let V_1, V_2 be inner product spaces and $T : V_1 \to V_2$ $\mathbb{F}$-linear.
Let $b \in V_2$. The set of solutions to $T(x) = b$ is $\{x \in V_1 \mid T(x) = b\} = \{a + v \mid a \in V_1, v \in \ker(T)\}$
If we solve $T \circ T^*(x) = b$, then $T^*(x)$ is the **minimum norm solution** to $T(x) = b$.

Theorem

If $v_1, \ldots, v_n$ is an orthonormal basis for U, then $P_u(v) = \langle v, v_1\rangle v_1 + \langle v, v_2\rangle v_2 + \ldots + \langle v, v_n\rangle v_n$

Theorem

Let V be a finite-dimensional vectorspace. Let T be a linear operator. Then

$$T \text{ is an orthogonal projection} \iff T \text{ is idempotent and self-adjoint}$$

Theorem

Let $U = \operatorname{span}\{u_1, u_2, \ldots, u_n\}$ be a finite-dimensional subspace of an inner product space V. Let $v \in V$.
Then $P_u v = \sum_{j=1}^{n} c_j u_j$ where $[c_1\, c_2\, \ldots\, c^n]^T \in \mathbb{F}^n$ is a solution to the **normal equation**:

$$\begin{bmatrix} \langle u_1, u_1\rangle & \langle u_1, u_2\rangle & \cdots & \langle u_1, u_n\rangle \\ \langle u_2, u_1\rangle & \langle u_2, u_2\rangle & \cdots & \langle u_2, u_n\rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle u_n, u_1\rangle & \langle u_n, u_2\rangle & \cdots & \langle u_n, u_n\rangle \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} = \begin{bmatrix} \langle v, u_1\rangle \\ \langle v, u_2\rangle \\ \vdots \\ \langle v, u_n\rangle \end{bmatrix}$$

Theorem *Best approximation theorem*

Let U be a finite-dimensional subspace of an inner product space V.
Then $\|v - P_u v\| \le \|v - w\| \quad \forall w \in U$ with equality if and only if $P_u v = w$

2.5 Definite matrices

Definition *Definite matrix*

Define the inner product as the standard inner product in $\mathbb{F}^n$. Let $A \in \mathbb{F}^n \times n$ be a Hermitian matrix.
We say that A is:

1. **positive semidefinite** if $\langle Ax, x\rangle \ge 0 \quad \forall x \in \mathbb{F}^n$
2. **positive definite** if $\langle Ax, x\rangle > 0$ for all nonzero $x \in \mathbb{F}^n$
3. **negative semidefinite** if $\langle Ax, x\rangle \le 0 \quad \forall x \in \mathbb{F}^n$
4. **negative definite** if $\langle Ax, x\rangle < 0$ for all nonzero $x \in \mathbb{F}^n$
5. **indefinite** if $\langle Ax, x\rangle$ takes both positive and negative values.

Notation: $A \ge 0$, $A > 0$, $A \le 0$, $A < 0$

Theorem

Let $A \in \mathbb{F}^n$ be a Hermitian matrix with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{R}$ We have that:

1. $A \ge 0$ if and only if $\lambda_i \ge 0 \quad \forall i \in \{1, 2, \ldots, n\}$
2. $A > 0$ if and only if $\lambda_i > 0 \quad \forall i \in \{1, 2, \ldots, n\}$
3. $A \le 0$ if and only if $\lambda_i \le 0 \quad \forall i \in \{1, 2, \ldots, n\}$
4. $A < 0$ if and only if $\lambda_i < 0 \quad \forall i \in \{1, 2, \ldots, n\}$
5. A is indefinite if and only if $\lambda_i > 0$ and $\lambda_j < 0$ for some $i, j \in \{1, 2, \ldots, n\}$

Lemma

- A is positive semidefinite if and only if there exists a $B \in \mathbb{F}^{n\times n}$ such that $A = B^*B$
- A is positive definite if and only if there exists a nonsingular B such that $A = B^*B$
- $A \leq 0 \iff -A \geq 0$
- If A is positive semidefinite, then $\langle Ax, x\rangle = 0 \implies Ax = 0$
- $A > 0 \implies \det A > 0$

Definition *Square root of a matrix*

If $A = B^2$, then B is the square root of A.

Some matrices have many square roots. Not every matrix has a square root.

Theorem

Let $A \in \mathbb{F}^{n\times n}$ be positive semidefinite. Then A has a unique positive semidefinite square root.

Theorem *Cholesky factorization*

Let $F \in \mathbb{F}^{n\times n}$ be a positive semidefinite matrix.
There exists a lower triangular matrix $L \in \mathbb{F}^{n\times n}$ with real nonnegative diagonal entries such that $P = LL^*$

Lemma *QR factorization*

Let $A \in \mathbb{F}^{n\times n}$.
There exists a unitary Q and upper triangular R with real nonnegative diagonal entries such that $A = QR$

2.6 Singular value decomposition

Definition *Rank deficient matrix*

Let $a \in \mathbb{R}^{m\times n}$ be a real matrix with $m \geq n$. A is called **rank deficient** if $\operatorname{rank} A < n$.

Theorem *Singular value decomposition (tall matrices)*

Let $A \in \mathbb{R}^{m\times n}$ with $m \geq n$
There exist two orthogonal matrices $U \in R^{m\times m}$ and $V \in \mathbb{R}^{n\times n}$ and real numbers $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n \geq 0$ such that $A = U\Sigma V^T$, where

$$\Sigma = \left[\begin{array}{cccc} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_n \\ \hline O & & & \end{array}\right]$$

and O is a $(m-n)\times n$ zero matrix (with the same m, n as A)
Moreover if $\operatorname{rank} A = r$ then

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

The numbers σ_i are unique. They are called the **singular values** of A.

Σ is unique in general, but U and V are not.

Lemma

Let $r = \operatorname{rank} A$.
Let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be the eigenvalues of A^TA, and order them such that $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$
Partition U into U_1 (first r columns) and U_2 (last $n-r$ columns). Similarly, partition V into V_1 and V_2.

1. $\lambda_i \in \mathbb{R}$ and $\lambda_i \geq 0$ for all $i \in 1, 2, \ldots, n$
2. $\lambda_{r+1} = \lambda_{r+2} = \ldots = \lambda_n = 0$
3. $AV_2 = 0$
4. $AV_1V_1^T = A$
5. Let $u_i = \frac{1}{\sigma_i}Av_i$ where v_i is the i-th column of V_1.
   Then $\{u_1, u_2, \ldots, u_n\}$ is an orthonormal set in $\mathbb{R}^m$.

Theorem

The squares of the singular values of A are always equal to the eigenvalues of $A^T A$.

Definition *Left and right singular vectors*

- The columns of U are eigenvectors of AA^T. We call them **left singular vectors** of A.
- The columns of V are eigenvectors of $A^T A$. We call them **right singular vectors** of A.

Theorem

Let $r = \operatorname{rank} A$ and

$$\Sigma_1 = \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_r \end{bmatrix}$$

Then $A = U_1 \Sigma_1 V_1^T$ and $A^T = V_1 \Sigma_1 U_1^T$. We call this the **reduced** or **compact** SVD.

Theorem

1. $\operatorname{ran} A = \operatorname{ran} U_1$
2. $\operatorname{ran} A^T = \operatorname{ran} V_1$
3. $\operatorname{null} A = \operatorname{null} V_1^T = \operatorname{ran} V^2$
4. $\operatorname{null} A^T = \operatorname{null} U_1^T = \operatorname{ran} U^2$

Theorem *Generalized singular value decomposition*

Let $A \in \mathbb{R}^{m\times n}$ be a matrix of rank r.
There exist $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n > 0$ and orthogonal $U \in \mathbb{R}^{m\times m}$ and $V \in \mathbb{R}^{n\times n}$ such that

$$A = U \begin{bmatrix} \Sigma & 0_{(n-r)\times r} \\ 0_{r\times(m-r)} & 0_{(n-r)\times(m-r)} \end{bmatrix} V^T$$

where Σ is a diagonal matrix containing $\sigma_1, \sigma_2, \ldots, \sigma_r$ and $0_{m\times n}$ is the $m \times n$ zero matrix.

- If $\operatorname{rank} A = n$, then the zero blocks on the right are absent.
- If $\operatorname{rank} A = m$, then the zero blocks on the bottom are absent.
- If A is square and nonsingular, then all three zero blocks are absent.

Definition

$$\|M\|_2 = \sup_{x\in\mathbb{R}^n\setminus 0} \frac{\|Mx\|}{\|x\|} \qquad M_k := \{S \in \mathbb{R}^{m\times n} \mid \operatorname{rank} S \leq k\} \qquad \alpha(A, M_k) := \inf\{\|A - S\|_2 \mid S \in M_k\}$$

Theorem *Best rank n approximation*

Let $A \in \mathbb{R}^{m\times n}$ and $k < r = \operatorname{rank} A$. Let $M_k := \{S \in \mathbb{R}^{m\times n} \mid \operatorname{rank} S \leq k\}$
Let $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_k \geq \sigma_{k+1} \geq \ldots \geq \sigma_r > 0$ be the nonzero singular values of A
Then $\alpha(A, M_k) = \sigma_{k+1}$.
Let

$$A = U \begin{bmatrix} \Sigma & 0_{(n-r)\times r} \\ 0_{r\times(m-r)} & 0_{(n-r)\times(m-r)} \end{bmatrix} V^T$$

where Σ is a diagonal matrix containing $\sigma_1, \ldots, \sigma_k, \sigma_{k+1}, \ldots, \sigma r$ and $0_{m\times n}$ is the $m \times n$ zero matrix.
Then

$$X = U \begin{bmatrix} \widetilde{\Sigma} & 0_{(n-r)\times r} \\ 0_{r\times(m-r)} & 0_{(n-r)\times(m-r)} \end{bmatrix} V^T$$

where $\widetilde{\Sigma}$ is Σ with $\sigma_{k+1}, \ldots, \sigma r$ replaced by zeroes.
This matrix X is the **best approximation** of A in M_k.