

1 Vectors

1.1 Linear Combination (Def. 1.4)

Let $v_1, v_2, \dots, v_n \in \mathbb{R}^m$ and $\lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{R}$, then

$$\lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_n v_n$$

is a linear combination of v_1, v_2, \dots, v_n .

1.1.1 Special Linear Combinations (Def. 1.7)

A linear combination $\lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_n v_n$ is

- (i) *affine* if $\lambda_1 + \lambda_2 + \dots + \lambda_n = 1$,
- (ii) *conic* if $\lambda_i \geq 1$ for $i = 1, 2, \dots, n$,
- (iii) *convex* if it is both affine and conic.

1.2 Euclidean Norm (Def. 1.11)

$$\|v\| := \sqrt{v \cdot v} = \sqrt{v_1^2 + v_2^2 + \dots + v_n^2}.$$

To *normalize* a vector, i.e. make it of *length* one, we can divide by the Euclidean norm: $v' = v/\|v\|$.

1.3 Dot Product and Angles (Def. 1.14)

The angle α between two non-zero vectors v and w is

$$\cos(\alpha) = \frac{v \cdot w}{\|v\| \|w\|},$$

where $-1 \leq \alpha \leq 1$.

1.3.1 Cauchy-Schwarz (Lemma 1.12)

For any two vectors $v, w \in \mathbb{R}^m$

$$|v \cdot w| \leq \|v\| \cdot \|w\|.$$

Equality holds when one is a scalar multiple of the other.

1.3.2 Triangle Inequality (Lemma 1.16)

Let $v, w \in \mathbb{R}^m$. Then

$$\|v + w\| \leq \|v\| + \|w\|.$$

1.3.3 Orthogonal Vectors (Def. 1.15)

Let $v, w \in \mathbb{R}^m$. v and w are orthogonal if and only if

$$v \cdot w = 0,$$

i.e., the angle between them is $\alpha = 90^\circ$.

1.4 Linear Independence (Cor. 1.20)

Let $v_1, v_2, \dots, v_n \in \mathbb{R}^m$. They are linearly dependent if

- (i) none of the vectors is a linear combination of the other ones,
- (ii) there are no scalars $\lambda_1, \lambda_2, \dots, \lambda_n$ besides 0, 0, ..., 0 such that $\sum_{j=1}^n \lambda_j v_j = 0$,
- (iii) none of the vectors is a linear combination of the previous ones.

Following, the columns of a matrix A are linearly independent if there is no x besides 0 such that $Ax = 0$.

1.5 The Span (Def. 1.22)

Let $v_1, v_2, \dots, v_n \in \mathbb{R}^m$, their Span is defined as the set of all linear combination, i.e.,

$$\text{Span}(v_1, v_2, \dots, v_n) := \left\{ \sum_{j=1}^n \lambda_j v_j : \lambda_j \in \mathbb{R} \forall j \in [n] \right\}.$$

1.6 Example

Let $v = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$ and $w = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} \in \mathbb{R}^2$ be arbitrary and assume $v \neq 0$ and $w \neq \lambda v$, $\forall \lambda \in \mathbb{R}$. Show that $\text{Span}(v, w) = \mathbb{R}^2$.

To show this, prove that there are λ and μ for each $u \in \mathbb{R}^2$ such that $u = \lambda v + \mu w$, i.e.

$$\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \lambda \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} + \mu \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} \Rightarrow \begin{cases} (1) u_1 = \lambda v_1 + \mu w_1 \\ (2) u_2 = \lambda v_2 + \mu w_2 \end{cases}$$

Assume $v_1 \neq 0$, thus from (1) follows

$$\lambda = \frac{u_1 - \mu w_1}{v_1}.$$

Using this in (2) (assuming $w_2 - \frac{v_2 w_1}{v_1} \neq 0$), we get

$$u_2 = \left(\frac{u_1 - \mu w_1}{v_1} \right) v_2 + \mu w_2 \Rightarrow \mu = \frac{u_2 - \frac{u_1 v_2}{v_1}}{w_2 - \frac{v_2 w_1}{v_1}}.$$

Thus, choosing λ and μ is possible for all $u \in \mathbb{R}^2$.

2 Matrices (Def. 2.1)

An $m \times n$ matrix is a rectangular array of numbers with m columns and n rows, i.e.,

$$A \in \mathbb{R}^{m \times n} := \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}.$$

2.1 The Rank (Def. 2.9)

The rank of a matrix $A \in \mathbb{R}^{m \times n}$ is its number of *linearly independent* columns.

2.2 The Transpose (Def. 2.11)

Let $A \in \mathbb{R}^{m \times n}$, the transpose of A , A^T , is equal to A with interchanged rows and columns.

Moreover, $(AB)^T = B^T A^T$ for any $A \in \mathbb{R}^{n \times k}$ and $B \in \mathbb{R}^{k \times m}$ (Lemma 2.19).

2.3 Square Matrix Classes (Def. 2.3)

Let $A \in \mathbb{R}^{m \times m}$, A is

- (i) *upper triangular* if all entries below the diag. are 0,
- (ii) *lower triangular* if all entries above the diag. are 0,
- (iii) *diagonal* if it is both *upper* and *lower* triangular, i.e., the matrix only has non-zero values on its diagonal,
- (iv) *symmetric* if the values above and below the diagonal are equal.

2.4 Symmetric Matrices (Obs. 2.12)

Let $A \in \mathbb{R}^{m \times m}$, A is symmetric $\Leftrightarrow A = A^T$.

2.5 Matrix-Vector Multiplication (Def 2.4)

The product Ax can be understood in three different ways:

2.5.1 Matrix Picture

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a_{11} \cdot x_1 + \dots + a_{1n} \cdot x_n \\ a_{21} \cdot x_1 + \dots + a_{2n} \cdot x_n \\ \vdots \\ a_{m1} \cdot x_1 + \dots + a_{mn} \cdot x_n \end{bmatrix}.$$

2.5.2 Row Picture

$$\begin{bmatrix} -u_1 - \\ -u_2 - \\ \vdots \\ -u_m - \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} u_1 \cdot x \\ u_2 \cdot x \\ \vdots \\ u_m \cdot x \end{bmatrix}.$$

2.5.3 Column Picture

$$\begin{bmatrix} | & | & | \\ w_1 & \dots & w_n \\ | & | & | \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1 \cdot w_1 + x_2 \cdot w_2 + \dots + x_n \cdot w_n.$$

2.6 Matrix-Matrix Multiplication (Def. 2.16)

For any two matrices $A \in \mathbb{R}^{m \times k}$ and $B \in \mathbb{R}^{k \times n}$, the product $AB \in \mathbb{R}^{m \times n}$ can also be understood in three ways:

2.6.1 The Matrix Picture

$$\begin{bmatrix} -u_1 - \\ -u_2 - \\ \vdots \\ -u_m - \end{bmatrix} \begin{bmatrix} | & | & | \\ v_1 & v_2 & \dots & v_n \\ | & | & | \end{bmatrix} = \begin{bmatrix} u_1 \cdot v_1 & u_1 \cdot v_2 & \dots & u_1 \cdot v_n \\ u_2 \cdot v_1 & u_2 \cdot v_2 & \dots & u_2 \cdot v_n \\ \vdots & \vdots & \ddots & \vdots \\ u_m \cdot v_1 & u_m \cdot v_2 & \dots & u_m \cdot v_n \end{bmatrix}.$$

2.6.2 The Row Picture

$$\begin{bmatrix} -u_1 - \\ -u_2 - \\ \vdots \\ -u_m - \end{bmatrix} B = \begin{bmatrix} -u_1 B - \\ -u_2 B - \\ \vdots \\ -u_m B - \end{bmatrix}.$$

2.6.3 The Column Picture

$$A \begin{bmatrix} | & | & | \\ v_1 & v_2 & \dots & v_n \\ | & | & | \end{bmatrix} = \begin{bmatrix} | & | & | \\ Av_1 & Av_2 & \dots & Av_n \\ | & | & | \end{bmatrix}.$$

2.6.4 Distributivity and Associativity (Lemma 2.22)

Let A, B, C and D be four matrices, then

- (i) $A(B + C) = AB + AC$ and $(B + C)D = BD + CD$ (distributivity)
- (ii) $(AB)C = A(BC)$ (associativity)

Keep in mind, that MATRIX MULTIPLICATION IS NOT COMMUTATIVE.

2.7 $A = CR$

Let $A \in \mathbb{R}^{m \times n}$ with rank r . Let $C \in \mathbb{R}^{m \times r}$ be the submatrix of A containing all its independent columns. Then there exists a unique $R \in \mathbb{R}^{r \times n}$ such that

$$A = CR.$$

In other words, C describes the linearly independent columns, while R shows how to combine them to create A . For instance,

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

2.8 Linear Transformation (Def. 2.27)

Let $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ be a function. T is called a linear transformation if $\forall v, w \in \mathbb{R}^n$ and $\forall \lambda, \mu \in \mathbb{R}$,

$$T(\lambda v + \mu w) = \lambda T(v) + \mu T(w).$$

2.9 Example

Let $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$, is T a linear transformation?

- (i) Show $T(v + w) = T(v) + T(w)$ for all $v, w \in \mathbb{R}^n$.
- (ii) Show $T(\lambda v) = \lambda T(v)$ for all $v \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}$.

2.9.1 $T = T_A$ (Theo. 2.29)

Let $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ be a linear transformation, there exists a unique $A \in \mathbb{R}^{m \times n}$ such that $T = T_A$.

2.9.2 Kernel and Image (Def. 2.31)

Let $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ be a linear transformation,

- (i) the *kernel* of T is $\text{Ker}(T) = N(T_A)$,
- (ii) the *image* of T is $\text{Im}(T) = C(T_A)$.

2.10 Rotation Matrix

In $\mathbb{R}^{2 \times 2}$, to rotate the xy plane counterclockwise around an angle θ we use

$$R = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}.$$

In $\mathbb{R}^{3 \times 3}$, there are three different rotation matrices, one for each axis:

$$R_x(\theta) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{bmatrix}$$

$$R_y(\theta) = \begin{bmatrix} \cos \theta & 0 & \sin \theta \\ 0 & 1 & 0 \\ -\sin \theta & 0 & \cos \theta \end{bmatrix}$$

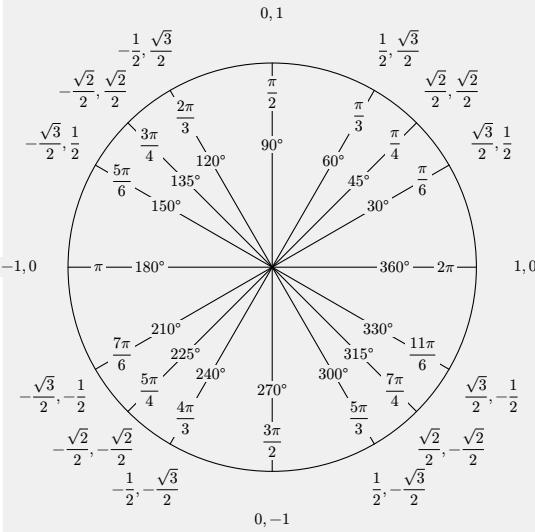
$$R_z(\theta) = \begin{bmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

2.10.1 Composition of Rotations

Let $R(\theta_1), R(\theta_2) \in \mathbb{R}^{2 \times 2}$ be two rotation matrices, then

$$R(\theta_1)R(\theta_2) = R(\theta_1 + \theta_2).$$

Moreover, for every $R(\theta) \in \mathbb{R}^{2 \times 2}$, there is some $R(-\theta) \in \mathbb{R}^{2 \times 2}$ such that $R(\theta)R(-\theta) = R(0) = I$.



3 Linear Equations

A system of m linear equations in n variables,

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n &= b_m, \end{aligned}$$

can be expressed using matrices in the shape $Ax = b$:

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}.$$

3.1 Gauss Elimination

3.1.1 Back Substitution

We can directly solve $Ax = b$, when A is an *upper triangular matrix* by solving the equations in reverse order.

$$\left[\begin{array}{ccc|c} 2 & 3 & 4 & 19 \\ 0 & 5 & 6 & 17 \\ 0 & 0 & 7 & 14 \end{array} \right] \quad \text{Solve for } x_3 \text{ first and then work your way up (backwards substitution).}$$

3.1.2 Elimination

Transform $Ax = b$ to $Ux = c$, where U is upper triangular, by performing

- (i) row exchanges,
- (ii) scalar multiplication,
- (iii) row subtraction.

u_{11}	\dots	c_1
0	u_{22}	\dots
0	0	\ddots
0	0	u_{rr}
\vdots	\vdots	\vdots
0	0	0
		c_m

Crucially, $Ax = b$ and $Ux = c$ have the same solutions (Lemma 3.3).

3.1.3 Elimination and Permutation Matrices

Elimination and row exchanges requires matrices as *linear combinations of the rows* (multiply from the left).

$$\text{row 2 minus } c \cdot \text{row 1} \quad \text{exchange row 2 and row 3}$$

$$E_{21} = \begin{bmatrix} 1 & 0 & 0 \\ -c & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad P_{23} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}.$$

3.2 Inverse Matrices (Def. 3.7)

Let $M \in \mathbb{R}^{m \times m}$, M is invertible, if there exists an $M^{-1} \in \mathbb{R}^{m \times m}$ such that

$$MM^{-1} = M^{-1}M = I.$$

For invertible matrices, the following hold

- (i) inverses are unique, i.e. if $AM = MA = I$ and $BM = MB = I$, $A = B$ (Lemma 3.8),
- (ii) if A and B are invertible, then $(AB)^{-1} = B^{-1}A^{-1}$ (Lemma 3.9),
- (iii) if A is invertible, then $(A^T)^{-1} = (A^{-1})^T$ (Lemma 3.10).

3.2.1 Inverse Theo. (Theorem 3.11)

Let $A \in \mathbb{R}^{m \times m}$, the following is equivalent

- (i) A is invertible,
- (ii) For every $b \in \mathbb{R}^m$, $Ax = b$ has a unique solution x ,
- (iii) The columns of A are linearly independent.

3.2.2 Example

Given $A, B \in \mathbb{R}^{m \times m}$ such that $AB = I$, prove $BA = I$.

- (i) Show B has linearly independent columns: Let $x \in \mathbb{R}^m$ such that $Bx = 0$, then

$$x = Ix = ABx = A0 = 0.$$

Hence by Lemma 1.19, they are linearly independent.

- (ii) Show A also has linearly independent columns: Let $y \in \mathbb{R}^m$ such that $Ay = 0$, then by Theorem 3.11 there is some $x \in \mathbb{R}^m$ such that $Bx = y$, then

$$y = Bx = B(Ix) = B(ABx) = B(Ay) = BO = 0.$$

- (iii) Show $BA - I = 0$:

$$A(BA - I) = ABA - A = IA - A = 0.$$

Thus, $BA = I$.

3.2.3 Inverse of a 2×2 matrix

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

$$A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \quad (\text{if } ad - bc \neq 0).$$

3.2.4 Inverse of a $n \times n$ matrix

TODO: DID WE HAVE THAT??

3.3 Gauss-Jordan Elimination

3.3.1 Reduced Row Echelon Form

Let $M \in \mathbb{R}^{m \times n}$ with rank r , M is in RREF, if

- (i) it contains the unit vectors e_0, \dots, e_r as its columns in ascending order,
- (ii) the columns in between contain values only up to the i -th row, where i refers to the index of the most recent unit vector.

3.3.2 Solving $Ax = b$

We add the following step to Gauss Elimination:

- (i) multiply each row with a pivot p with $1/p \Rightarrow p = 1$
- (ii) eliminate any value above the pivot
- (iii) (at the end remove any zero-rows, $R_0 \rightarrow R$)

Then $Ax = b$ has solution if $b_i = 0 \forall i > r$. Then x is equal to the *direct solution* where

$$x_i = \begin{cases} b_i & \text{if } j = j_i \\ 0 & \text{otherwise.} \end{cases}$$

4.5 Bases and Dimensions

4.5.1 Basis (Def. 4.16)

Let V be a vector space. A subset $B \subseteq V$ is called a *basis* of V if B is linearly independent and $\text{Span}(B) = V$.

4.5.2 Dimensions (Def. 4.23)

Let V be a finitely generated vector space. Then $\dim(V) = |B|$, for any basis $B \subseteq V$. Moreover, the size of all bases of a vector space is equal (Theo. 4.20).

4.6 Example

Let $V \subseteq \mathbb{R}^m$, what is the dimension of V ?

- (i) Find a basis of V ,
- (ii) show it is a basis (linearly independent & span is V),
- (iii) dimension is the cardinality of the base.

4.6.1 Steinitz Exchange Lemma (Lemma 4.19)

Let V be a vector space, $F \subseteq V$ a finite set of linearly independent vectors, and $G \subseteq V$ a finite set of vectors with $\text{Span}(G) = V$. Then

- (i) $|F| \leq |G|$,
- (ii) there is some subset $E \subseteq G$ with $|E| = |G| - |F|$ such that $\text{Span}(F \cup E) = V$.

4.7 The Column Space of A (Def. 2.8)

Let $A \in \mathbb{R}^{m \times n}$. The *column space* of A is the span of its columns, $C(A) := \{Ax : x \in \mathbb{R}^n\} \subseteq \mathbb{R}^m$.

The set of all vectors you can get by combining the columns of the matrix. For a matrix transformation $Ax = b$, the column space is the set of all possible output vectors b for which a solution x exists.

4.7.1 Basis of $C(A)$ (Theo. 4.25)

The basis of $C(A)$ is the set of its *linearly independent columns*, and hence $\dim(C(A)) = \text{rank}(A)$. These can be calculated using Gauss-Jordan (C matrix of $A = CR$).

4.8 The Row Space of A (Def. 2.13)

Let $A \in \mathbb{R}^{m \times n}$. The *row space* of A is the span of its rows,

$$R(A) := \{A^T x : x \in \mathbb{R}^m\} = C(A^T) \subseteq \mathbb{R}^n.$$

It tells about the fundamental, independent equations in a system. If a new equation is a combination of existing rows, it's in the row space. I.e. if a vector satisfies this new equation, it's in the row space.

For example, let $A = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \end{bmatrix}$, with row vectors $r_1 = [1, 0, 2]$ and $r_2 = [0, 1, 0]$. A linear combination of these is:

$$c_1[1, 0, 2] + c_2[0, 1, 0] = [c_1, c_2, 2c_1]$$

The row space is the set of vectors $(x, y, z) \in \mathbb{R}^3$ satisfying $z = 2x$ (one less free variable \Rightarrow a plane through the origin).

4.8.1 Basis of $R(A)$ (Theo. 4.28)

The basis of $R(A)$ is the set of its *linearly independent rows*, and hence $\dim(R(A)) = \dim(C(A)) = \text{rank}(A)$. These correspond to the first r rows of A in REF, where r is the rank of A .

4.9 Bases in $A = CR$ (Theo. 4.30)

Let $A = CR$, then the columns of C form a basis of $C(A)$ and the columns of R form a basis of $R(A)$.

4.10 The Nullspace of A (Def. 4.31)

Let $A \in \mathbb{R}^{m \times n}$. The *nullspace* of A is the set of all solutions to $Ax = 0$,

$$N(A) := \{x \in \mathbb{R}^n : Ax = 0\} \subseteq \mathbb{R}^n.$$

4.10.1 Basis of $N(A)$ (Lemma 4.34)

We can calculate a basis of $N(A)$, by converting A to R in RREF through Gauss-Jordan and then solving $Rx = 0$ with the *special cases* where $x_i = 1$ for every $i = 1, \dots, n$ for every linearly **dependent** columns in R .

For example, $A \in \mathbb{R}^{2 \times 4}$,

- (i) convert $A \rightarrow R$ through Gauss-Jordan,
- (ii) solve $Rx = 0$ with special cases (per special case: set one free variable (one variable of the **dependent** rows) e.g. $x_2 = 1$, the other free variables to 0), solve for x

$$\begin{array}{c} \left[\begin{array}{cccc} 1 & 2 & 0 & 3 \\ 0 & 0 & 1 & -2 \end{array} \right] \xrightarrow{\text{A in RREF}} \left[\begin{array}{cccc} x_1 \\ x_2 \\ x_3 \\ x_4 \end{array} \right] = 0 \text{ with } x = \underbrace{\left[\begin{array}{c} x_1 \\ 1 \\ x_3 \\ x_4 \end{array} \right]}_{\text{special cases}} \text{ and } \left[\begin{array}{c} x_1 \\ x_2 \\ x_3 \\ 1 \end{array} \right]. \end{array}$$

(iii) The solutions x_1, \dots, x_n form a Basis of $N(R)$,

$$x_1 = \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix}, x_2 = \begin{bmatrix} -3 \\ 0 \\ 2 \\ 1 \end{bmatrix} \Rightarrow \{x_1, x_2\} \text{ is a basis of } N(R).$$

(iv) Any basis of $N(R)$ is also a basis of $N(A)$ (Lemma 4.33).

Following, $\dim(N(A)) = n - \text{rank}(A)$ (Theo. 4.35).

4.11 Left Nullspace of A (Def. 4.36)

Let $A \in \mathbb{R}^{m \times n}$. The *left nullspace* of A is the set of all solutions to $A^T y = 0$,

$$LN(A) := N(A^T) \subseteq \mathbb{R}^m.$$

4.11.1 Basis of $LN(A)$ (Theo. 3.48)

We can calculate a basis of $LN(A)$ by,

- (i) convert $A \rightarrow R_0$ through Gauss-Jordan,
- (ii) the last $m - r$ rows of the row operation matrix M used in Gauss-Jordan form a basis of $LN(A)$.

Following, $\dim(LN(A)) = m - \text{rank}(A)$.

subspace	$C(A)$	$R(A)$	$N(A)$	$LN(A)$
dimensions	r	r	$n - r$	$m - r$

4.12 The Solution Space of $Ax = b$ (Def. 4.39)

Let $A \in \mathbb{R}^{m \times n}$, and $b \in \mathbb{R}^m$, then *solution space* of $Ax = b$ is the set

$$\text{Sol}(A, b) := \{x \in \mathbb{R}^n : Ax = b\} \subseteq \mathbb{R}^n.$$

For any $Ax = b$ we have three options, 1. no solutions, 2. one solution and 3. infinite solutions.

- (i) If A is not invertible and $b \notin C(A)$ then no solution exists.
- (ii) If A is invertible $\Rightarrow N(A) = \{0\}$ then exactly one solution exists.

- (iii) If A is not invertible, but $b \in C(A)$, then $\exists s$ and we can shift the non-trivial nullspace using s to get infinite solutions:

$$\underbrace{A(s + n)}_x = As + 0 = b \text{ with } n \in N(A).$$

	$r = n$ (full rank)	$r < n$ (dependent cols)
$r = m$ (full rank)	invertible \Rightarrow one solution	underdetermined $\Rightarrow \infty$ solutions
$r < m$ (zero rows)	overdetermined $\Rightarrow 0$ or 1 solution	0 or ∞ solutions

4.12.1 Sol. Space is shifted Nullspace (Theo. 4.40)

Let $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$. Let s some solution for x to $Ax = b$, then

$$\text{Sol}(A, b) = \{s + n : n \in N(A)\}.$$

5 Orthogonality

5.1 Orthogonal Subspaces (Def. 5.1.1)

Two subspaces V and W are orthogonal if for all $v \in V$ and $w \in W$, $v \cdot w = 0$.

More specifically, this also holds for the bases, i.e.,

$hhjkjhkjlkjhjhjkhjkjh v_1, \dots, v_k$ be a basis of V and w_1, \dots, w_l be a basis of W . V and W are orthogonal if and only if $v_i \cdot w_j = 0 \forall i \in [k]$ and $j \in [l]$ (Lemma 5.1.2).

5.2 Orthogonal Complement (Def. 5.1.5)

Let V be a subspace of \mathbb{R}^n , the orthogonal complement to V ,

$$V^\perp := \{w \in \mathbb{R}^n : w \cdot v = 0 \text{ for all } v \in V\}.$$

5.3 Orthogonal Decomps. (Theo. 5.1.7)

Let V, W be subspaces of \mathbb{R}^n , then these are equivalent

- (i) $W = V^\perp$,
- (ii) $\dim(V) + \dim(W) = n$,
- (iii) Every $u \in \mathbb{R}^n$ can be written as $u = v + w$ with $v \in V$ and $w \in W$.

5.4 Orthogonal Matrix Subspaces (Cor. 5.1.9)

Let $A \in \mathbb{R}^{m \times n}$, then

- (i) $N(A) = C(A^T)^\perp$, also $N(A) = N(A^T A)$,
- (ii) $C(A^T) = N(A)^\perp$, also $C(A^T) = C(A^T A)$.

5.5 Decomp. of the Sol. Space (Theo. 5.1.10)

Let $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$. $\text{Sol}(A, b) = x_1 + N(A)$ with $x_1 \in R(A)$ such that $Ax_1 = b$.

6 Projections

6.1 Projection Matrix (Theo. 5.2.6)

Let S be a subspace of \mathbb{R}^m and A be a matrix with columns that are a basis of S . The projection of $b \in \mathbb{R}^m$ on S is given by

$$\text{proj}_S(b) = Pb \text{ with } P = A(A^T A)^{-1} A^T.$$

Moreover,

- (i) $A^T A$ is invertible $\Leftrightarrow A$ has linearly independent columns (Lemma 5.2.4),
- (ii) If A has linearly independent columns, then $A^T A$ is square, invertible and symmetric (Cor. 5.2.5).

6.1.1 Remark 5.2.7

For any projection matrix P and corresponding subspace S ,

- (i) $P^2 = (A(A^T A)^{-1} A^T)^2 = A(A^T A)^{-1} \underbrace{A^T A}_{I} (A^T A)^{-1} A^T = A(A^T A)^{-1} A^T = P$
- (ii) $\text{proj}_{S^\perp}(b) = b - Pb = (I - P)b$.

7 Linear Regression

7.1 Least Squares Approximation

A linear regression through the data points $(t_1, b_1), (t_2, b_2), \dots, (t_m, b_m)$ can be expressed in algebraic terms as minimizing the sum of the squared errors (where α_0 is the intercept, and α_1 the slope),

$$\min_{\alpha_0, \alpha_1} \left\| b - A \underbrace{\begin{bmatrix} \alpha_0 \\ \alpha_1 \end{bmatrix}}_{\alpha} \right\|^2, \text{ where } b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} \text{ and } A = \begin{bmatrix} 1 & t_1 \\ 1 & t_2 \\ \vdots & \vdots \\ 1 & t_m \end{bmatrix}.$$

To minimize error, the error vector must be orthogonal to columns of A , so

$$\begin{aligned} A^T(b - A\alpha) &= 0 \\ A^T A \alpha &= A^T b \quad (\text{"normal equation"}) \\ \alpha &= (A^T A)^{-1} A^T b \\ \begin{bmatrix} \alpha_0 \\ \alpha_1 \end{bmatrix} &= \begin{bmatrix} \frac{1}{m} & \sum_{k=1}^m t_k \\ 0 & \sum_{k=1}^m t_k^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum_{k=1}^m b_k \\ \sum_{k=1}^m t_k b_k \end{bmatrix}. \end{aligned}$$

7.2 Remark 5.3.3

If the columns of A are pairwise orthogonal, then $A^T A$ is a diagonal matrix and we can further simplify the expression,

$$\begin{bmatrix} \alpha_0 \\ \alpha_1 \end{bmatrix} = \begin{bmatrix} \frac{1}{m} & 0 \\ 0 & \frac{1}{\sum_{k=1}^m t_k^2} \end{bmatrix} \begin{bmatrix} \sum_{k=1}^m b_k \\ \sum_{k=1}^m t_k b_k \end{bmatrix} = \begin{bmatrix} \frac{1}{m} \sum_{k=1}^m b_k \\ \frac{\sum_{k=1}^m t_k b_k}{\sum_{k=1}^m t_k^2} \end{bmatrix}.$$

8 Orthonormal Bases

8.1 Orthonormal Vectors (Def. 5.4.1)

Vectors $q_1, \dots, q_n \in \mathbb{R}^m$ are orthonormal if they are orthogonal and have norm 1, i.e., if for all $i, j \in [n]$

$$q_i \cdot q_j = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise.} \end{cases}$$

8.2 Orthogonal Matrix (Def. 5.4.3)

A square matrix $Q \in \mathbb{R}^{n \times n}$ is orthogonal when $Q^T Q = I$. Then $Q Q^T = I$ and the columns of Q form an orthonormal basis of \mathbb{R}^n .

Additionally, orthogonal matrices *preserve norm and dot product*, i.e.,

$$\|Qx\| = \|x\| \text{ and } (Qx)^T (Qy) = x^T y,$$

for all $x, y \in \mathbb{R}^n$ (Prop. 5.4.6).

8.3 Gram-Schmidt

Given n linearly independent vectors a_1, \dots, a_n that span S , q_1, \dots, q_n can be constructed as follows:

- (i) $q_1 = \frac{a_1}{\|a_1\|}$
 - (ii) For $k = 2, \dots, n$,
- $$q_k = a_k - \sum_{i=1}^{k-1} (a_k^T q_i) q_i$$
- $$q_k = \frac{a_k}{\|a_k - \sum_{i=1}^{k-1} (a_k^T q_i) q_i\|}.$$

The vectors q_1, \dots, q_n form an orthonormal basis of S (Theorem 5.4.10).

8.4 Example

An upper triangular $n \times n$ matrix with non-zero diagonals that does **not** yield the canonical basis after the Gram-Schmidt process is $-I$.

8.5 QR-Decomposition

Let $A \in \mathbb{R}^{m \times n}$ with linearly independent columns.

$$A = QR$$

where Q is an $m \times n$ matrix with orthonormal columns and R is an upper triangular matrix given by $R = Q^T A$.

8.6 Projections with QR (Fact 5.4.13)

- (i) Any projection on $C(A)$ can also be done by Q , following $\text{proj}_{C(A)}(b) = Q(Q^T Q)^{-1} Q^T b = Q Q^T b$.
- (ii) The normal equation can be re-written as

$$A^T A \hat{x} = A^T b$$

$$R^T Q \hat{x} = R^T Q^T b \quad (R^T \text{ is invertible})$$

$$R \hat{x} = Q^T b. \quad (\text{useful for least squares})$$

9 Pseudoinverses

9.1 Full Column Rank (Def. 5.5.1)

For any $A \in \mathbb{R}^{m \times n}$ with $\text{rank}(A) = n$, $A_{\text{left}}^+ = (A^T A)^{-1} A^T$ is a **left inverse**, $\Rightarrow A^+ A = I$.

9.2 Full Row Rank (Def. 5.5.3)

For any $A \in \mathbb{R}^{m \times n}$ with $\text{rank}(A) = m$, $A_{\text{right}}^+ = A^T (A A^T)^{-1}$ is a **right inverse**, $\Rightarrow A A^+ = I$.

9.3 All Matrices (Def. 5.5.7)

For any $R \in \mathbb{R}^{m \times n}$ with $A = CR$,

$$A^+ = (CR)^+ = R^+ C^+$$

$$= R^T (RR^T)^{-1} (C^T C)^{-1} C^T$$

$$= R^T (C^T A R^T)^{-1} C^T.$$

9.4 Pseudoinverses with rank r (Prop. 5.5.9)

For $A \in \mathbb{R}^{m \times n}$ with rank r , let $S \in \mathbb{R}^{m \times r}$ and $T \in \mathbb{R}^{r \times n}$, such that $A = ST$, then $A^+ = T^+ S^+$.

9.5 Properties of A^+ (Theo. 5.5.11)

- (i) $A A^+ A = A$
- (ii) $A^+ A A^+ = A^+$

- (iii) A^+ is symmetric & the projection matrix on $C(A)$.
(iv) A^+A is symmetric & the projection matrix on $C(A^T)$.
(v) $(A^T)^+ = (A^+)^T$.

9.6 Properties of Pseudoinverses

- (i) Prove that if $\text{rank}(A) = \text{rank}(B) = n$, we have $(AB)^+ = B^+A^+$.

$C(AB) = C(A)$ (as $\text{rank}(B) = n$ implies that $C(B) = n$). Then using Prop. 5.5.9 we get $(AB)^+ = B^+A^+$.

- (ii) Prove that $A^+AA^+ = A^+$.

$$A^+AA^+ = (CR)^+CR(CR)^+ = R^+(C^+C)(RR^+)C^+ = R^+C^+ = A^+.$$

- (iii) Prove that $(A^T)^+ = (A^+)^T$.

prove for full row & column rank separately, then use Prop. 5.5.9 to get $(A^T)^+ = (C^T)^+(R^T)^+ = (C^+)^T(R^+)^T = (R^+C^+)^T = (A^+)^T$.

- (iv) Prove that A^+A is symmetric and that it is the projection matrix for the subspace $C(A^T)$.

projection matrix: $A^+A = (CR)^+CR = R^+C^+CR = R^T(RR^T)^{-1}R$. ($C(A^T) = C(R^T)$)

$$\text{symmetric: } (A^+A)^T = (R^T(RR^T)^{-1}R)^T = R^T((RR^T)^{-1})^T = R^T(RR^T)^{-1}R = A^+A.$$

10 Certificate of Unsolvability (Theorem 6.2.4)

Let $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$. The system of linear equations has no solution iff there exists vector z s.t.:

$$\begin{aligned} \{x \in \mathbb{R}^n \mid Ax = b\} &= \emptyset \\ \Leftrightarrow \{z \in \mathbb{R}^m \mid A^Tz = 0, b^Tz = 1\} &\neq \emptyset \end{aligned}$$

10.1 The Algebraic Certificate

The vector z serves as a “certificate” that $Ax = b$ is impossible. If both x and z existed, we would reach the following contradiction:

$$0 = 0^T x = (z^T A)x = z^T(Ax) = z^T b = 1$$

10.2 Characterizing Solvability

Row Independence: If the rows of A are linearly independent, then $A^Tz = 0$ only has the trivial solution $z = 0$. Since $b^T 0 = 0 \neq 1$, a certificate z can never exist. Therefore, $Ax = b$ always has a solution for every b .

Linear Independence: A vector b is linearly independent from the columns of A if and only if $Ax = b$ has no solution, which can be verified by finding the certificate z .

To find z , use Gaussian elimination (RREF) on the augmented matrix:

$$\begin{bmatrix} A^T \\ b^T \end{bmatrix} z = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

10.3 The Set of All Solutions (Theorem 6.2.2)

If the system is solvable, the set of all solutions is a shifted copy of the nullspace:

$$\{x \in \mathbb{R}^n \mid Ax = b\} = x_r + N(A)$$

where x_r is the **unique** vector in the row space $R(A) = C(A^T)$ such that $Ax_r = b$. This is because every solution x is a mix of two perpendicular parts:

$$x = \underbrace{x_r}_{\text{Row Space (unique)}} + \underbrace{x_n}_{\text{Nullspace (any)}}$$

(i) **The “Pure” Solution (x_r):** There is exactly one solution living entirely in the Row Space (the “Active Zone” of A). This is the solution with the *shortest length* because it contains no “waste.”

(ii) **The “Invisible” Noise (x_n):** You can add *any* vector x_n from the Nullspace to x_r without changing the result. **Why?** Because the matrix is blind to it:

$$A(x_r + x_n) = Ax_r + Ax_n = b + 0 = b$$

11 Determinant (Def. 6.0.6)

Let $A \in \mathbb{R}^{n \times n}$, the determinant $\det(A)$ is defined as

$$\det(A) = \sum_{\sigma \in \Pi_n} \text{sgn}(\sigma) \prod_{i=1}^n A_{i,\sigma(i)},$$

where Π_n is the set of all permutations of n elements. Moreover,

$$\det(I) = 1 \quad (\text{Prop. 6.0.7})$$

$$\text{If } A \in \mathbb{R}^{1 \times 1}, \det(A) = A \quad (\text{Def. 7.2.3})$$

$$\det(A^T) = \det(A) \quad (\text{Theo. 6.0.9})$$

$$\text{If } A \text{ is triangular, } \det(A) = \prod_{k=1}^n A_{kk} \quad (\text{Prop. 6.0.8})$$

$$\text{If } A \text{ is orthogonal, } \det(A) = 1 \text{ or } -1 \quad (\text{Prop. 6.0.10})$$

$$A \text{ is invertible if and only if } \det(A) \neq 0 \quad (\text{Prop. 6.0.11})$$

$$\text{If } \det(A) \neq 0, \det(A^{-1}) = \frac{1}{\det(A)} \quad (\text{Prop. 6.0.13})$$

$$\text{Given some } B \in \mathbb{R}^{n \times n}, \det(AB) = \det(A) \cdot \det(B) \quad (\text{Prop. 6.0.12})$$

- If any two rows are equal, then $\det(A) = 0$.
- If A has a row of zeros, then $\det(A) = 0$.
- If any Eigenvalue of A is 0, then $\det(A) = 0$.
- If we swap the rows of $A \rightarrow B$ once, then $\det(B) = -\det(A)$.
- The determinant is a linear function of each row separately.
 - If a single row of A is multiplied by some scalar t , then $\det(A') = t \cdot \det(A)$.
 - If the whole matrix is multiplied by t (i.e. all n rows are multiplied by t), then $\det(t \cdot A) = t^n \cdot \det(A)$.
 - If a row of A is replaced by the sum of itself and a multiple of another row, the determinant stays unchanged.

11.1 Determinant of a 2×2 Matrix

Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, then $\det(A) = ad - bc$.

11.2 Det. through Co-Factors

11.2.1 Co-Factors (Def. 6.0.15)

Let $A \in \mathbb{R}^{n \times n}$, for each $1 \leq i, j \leq n$, let \mathcal{A}_{ij} denote the matrix obtained by removing the i -th row and j -th column from A . The co-factors of A are

$$C_{ij} = (-1)^{i+j} \det(\mathcal{A}_{ij}).$$

11.2.2 Determinant (Prop. 6.0.16)

We can then rewrite the determinant of A as

$$\det(A) = \sum_{j=1}^n A_{ij} C_{ij},$$

for some $1 \leq i \leq n$. As in make a $++\dots$ grid, pick a row or column and calculate $\pm A_{i,j} \det(\dots)$ for that whole row or column recursively.

11.3 Cramer’s Rule (Prop. 6.0.19)

Let $A \in \mathbb{R}^{n \times n}$, such that $\det(A) \neq 0$ and $b \in \mathbb{R}^n$. Then the solution $x \in \mathbb{R}^n$ for $Ax = b$ is given by

$$x_j = \frac{\det(\mathcal{B}_j)}{\det(A)},$$

where \mathcal{B}_j is the matrix obtained by replacing the j -th column of A with b .

11.4 Linearity of the Determinant (Prop. 6.0.22)

The determinant is linear in each row or each column, i.e.,

$$\left| \begin{array}{c} -a_0a_1^T + a_1a_0^T \\ -a_2^T \\ \vdots \\ -a_n^T \end{array} \right| = \alpha_0 \left| \begin{array}{c} -a_1^T \\ -a_2^T \\ \vdots \\ -a_n^T \end{array} \right| + \alpha_1 \left| \begin{array}{c} -a_2^T \\ -a_3^T \\ \vdots \\ -a_n^T \end{array} \right|.$$

11.5 Example

Let $v_1, v_2, u_1, u_2 \in \mathbb{R}^n$ and $M \in \mathbb{R}^{(n-2) \times n}$ be arbitrary and consider the four $n \times n$ matrices

$$A = \begin{bmatrix} -v_1^T \\ -u_1^T \\ M \end{bmatrix}, B = \begin{bmatrix} -v_1^T \\ -u_2^T \\ M \end{bmatrix},$$

$$C = \begin{bmatrix} -v_2^T \\ -u_1^T \\ M \end{bmatrix}, D = \begin{bmatrix} -v_2^T \\ -u_2^T \\ M \end{bmatrix}$$

as well as the following $n \times n$ matrix

$$E = \begin{bmatrix} -(v_1 - v_2)^T \\ -(u_1 - u_2)^T \\ M \end{bmatrix}.$$

Find a formula for $\det(E)$ in terms of $\det(A)$, $\det(B)$, $\det(C)$ and $\det(D)$.

Use Prop. 7.3.7 to get

$$\begin{aligned} \det(A) - \det(B) &= \det \begin{bmatrix} -v_1^T \\ -u_1^T \\ M \end{bmatrix} - \det \begin{bmatrix} -v_2^T \\ -u_2^T \\ M \end{bmatrix} \\ &= \det \begin{bmatrix} -v_1^T \\ -(u_1 - u_2)^T \\ M \end{bmatrix} \end{aligned}$$

and

$$\begin{aligned} \det(C) - \det(D) &= \det \begin{bmatrix} -v_2^T \\ -u_1^T \\ M \end{bmatrix} - \det \begin{bmatrix} -v_2^T \\ -u_2^T \\ M \end{bmatrix} \\ &= \det \begin{bmatrix} -v_2^T \\ -(u_1 - u_2)^T \\ M \end{bmatrix} \end{aligned}$$

together this gives us

$$\begin{aligned} \det \begin{bmatrix} -v_1^T \\ -(u_1 - u_2)^T \\ M \end{bmatrix} - \det \begin{bmatrix} -v_2^T \\ -(u_1 - u_2)^T \\ M \end{bmatrix} \\ = \det \begin{bmatrix} -(v_1 - v_2)^T \\ -(u_1 - u_2)^T \\ M \end{bmatrix}. \end{aligned}$$

12 Eigenvalues/vectors (Def. 7.1.1)

Let $A \in \mathbb{R}^{n \times n}$, $\lambda \in \mathbb{C}$ is an eigenvalue of A and $v \in \mathbb{C}^n \setminus \{0\}$ is the associated eigenvector of A , if

$$Av = \lambda v.$$

To calculate eigenvalue/vector pairs, we use Prop. 7.1.2:

- (i) $\det(A - \lambda I) = 0 \iff \lambda$ is an eigenvalue of A ,
- (ii) v is an eigenvector of A (associated with λ) $\iff v \in N(A - \lambda I)$ and $v \neq 0$.

Every matrix $A \in \mathbb{R}^{n \times n}$ has an eigenvalue (Theo 7.1.4).

12.1 Example

Consider the sequence of numbers given by $a_0 = 1$, $a_1 = 1$ and $a_n = -a_{n-1} + 6a_{n-2}$ for $n \geq 2$. Find $\alpha, \beta \in \mathbb{R}$ such that $a_n = \frac{4}{5}\alpha^n + \frac{1}{5}\beta^n$ for all $n \in \mathbb{N}_0$. Prove your answer.

- (i) Define sequence algebraically: Let $v_n = \begin{bmatrix} a_{n+1} \\ a_n \end{bmatrix}$. We translate the recurrence $a_n = -a_{n-1} + 6a_{n-2}$ into matrix form.

$$\begin{bmatrix} a_n \\ a_{n-1} \end{bmatrix} = \underbrace{\begin{bmatrix} -1 & 6 \\ 1 & 0 \end{bmatrix}}_A \underbrace{\begin{bmatrix} a_{n-2} \\ a_{n-1} \end{bmatrix}}_{v_{n-2}}$$

Note: Since $v_0 = \begin{bmatrix} a_1 \\ a_0 \end{bmatrix}$ has a_0 at the bottom, our target a_n (which we'll need later) will always correspond to the **bottom component** of v_n . This is more clear if we shift the above stated matrix equation:

$$\begin{bmatrix} a_{n+1} \\ a_n \end{bmatrix} = \underbrace{\begin{bmatrix} -1 & 6 \\ 1 & 0 \end{bmatrix}}_A \underbrace{\begin{bmatrix} a_{n-1} \\ a_{n-2} \end{bmatrix}}_{v_{n-1}}$$

- (iii) Find the corresponding eigenvectors $v_i \in N(A - \lambda_i I)$:

$$(A - \lambda_i I)v_i = 0$$

$$\begin{bmatrix} -3 & 6 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} (v_1)_1 \\ (v_1)_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow v_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix},$$

$$\begin{bmatrix} 2 & 6 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} (v_2)_1 \\ (v_2)_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow v_2 = \begin{bmatrix} -3 \\ 1 \end{bmatrix},$$

- (iv) Express v_0 as a linear combination $c_1 v_1 + c_2 v_2$: We solve the system for the initial vector $v_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$:

$$c_1 \begin{bmatrix} 2 \\ 1 \end{bmatrix} + c_2 \begin{bmatrix} -3 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \Rightarrow \begin{cases} 2c_1 - 3c_2 = 1 \\ c_1 + c_2 = 1 \end{cases}$$

From the second equation, $c_1 = 1 - c_2$. Substituting into the first:

$$2(1 - c_2) - 3c_2 = 1$$

$$2 - 5c_2 = 1 \Rightarrow 5c_2 = 1 \Rightarrow c_2 = \frac{1}{5}$$

Then, $c_1 = 1 - \frac{1}{5} = \frac{4}{5}$. Thus:

$$\begin{bmatrix} a_1 \\ a_0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \frac{4}{5}\mathbf{v}_1 + \frac{1}{5}\mathbf{v}_2$$

(v) Rewrite $\mathbf{v}_n = \begin{bmatrix} a_{n+1} \\ a_n \end{bmatrix} = A^n \mathbf{v}_0$ as

$$\begin{aligned} \mathbf{v}_n &= A^n \left(\frac{4}{5}\mathbf{v}_1 + \frac{1}{5}\mathbf{v}_2 \right) \\ &= \frac{4}{5}A^n\mathbf{v}_1 + \frac{1}{5}A^n\mathbf{v}_2 \\ &= \frac{4}{5}\lambda_1^n\mathbf{v}_1 + \frac{1}{5}\lambda_2^n\mathbf{v}_2 \\ &= \frac{4}{5}2^n \begin{bmatrix} 2 \\ 1 \end{bmatrix} + \frac{1}{5}(-3)^n \begin{bmatrix} -3 \\ 1 \end{bmatrix}. \end{aligned}$$

(vi) Consider second component: $a_n = \frac{4}{5} \cdot 2^n \cdot 1 + \frac{1}{5} \cdot (-3)^n \cdot 1$.

12.2 Reverse Diagonalization Example

Construct a square matrix B with eigenvalues 0, 1, 2, such that B is not a diagonal matrix.

Let A be the diagonal matrix with 0, 1, 2 on its diagonals and let

$$V = \begin{bmatrix} 1 & 1 & 1 \\ e_1 + e_2 & e_1 - e_2 & e_3 \\ 1 & 1 & 1 \end{bmatrix}, \text{ then } B = VAV^{-1}. \text{ Ensure } \det(V) \neq 0.$$

12.3 Quadratic Formula (for finding λ)

To find zeros of $ax^2 + bx + c = 0$ are given by

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}.$$

12.4 Eigenvalues, Trace & Determinant (Def.)

8.3.4, Lemma 8.3.6

The characteristic polynomial of $A \in \mathbb{R}^{n \times n}$ is given by:

$$\begin{aligned} P(z) &= (-1)^n \det(A - zI) \\ &= \det(zI - A) = (z - \lambda_1)(z - \lambda_2)\dots(z - \lambda_n) \end{aligned}$$

The roots λ_i are the eigenvalues, and their **algebraic multiplicity** is the number of times they appear as roots.

Key Properties: The coefficients of the polynomial reveal the sum and product of eigenvalues:

(i) **Determinant:** The product of eigenvalues.

$$\det(A) = \prod_{i=1}^n \lambda_i$$

(ii) **Trace:** The sum of diagonal elements equals the sum of eigenvalues.

$$\text{Tr}(A) = \sum_{i=1}^n A_{ii} = \sum_{i=1}^n \lambda_i$$

12.4.1 Complex Eigenvalues & Conjugates

If A is a real matrix, complex eigenvalues come in **conjugate pairs**. If $Av = \lambda v$, then taking the conjugate of both sides gives:

$$A |(\bar{v})| = |(Av)| = |(\lambda)v| = |(\lambda)| |(v)|$$

Conclusion: If λ is an eigenvalue with eigenvector v , then $|(\lambda)|$ is an eigenvalue with eigenvector $|(v)|$.

12.5 Special Eigenvalues

(i) If λ and v are an eigenvalue-eigenvector pair of A , then λ^k and v are one for A^k . Induction Proof: $A^k v = A(A^{k-1}v) = A(\lambda^{k-1}v) = \lambda^{k-1}(Av) = \lambda^k v$ (Prop. 8.3.1).

- (ii) Let A be invertible, if λ and v are an eigenvalue-eigenvector pair of A , then $\frac{1}{\lambda}$ and v are an eigenvalue-eigenvector pair of A^{-1} . Proof: $Av = \lambda v \iff v = A^{-1}(\lambda v) \iff \lambda A^{-1}v = v \iff A^{-1}v = \frac{1}{\lambda}v$ (works since $\lambda \neq 0$) (Prop. 8.3.1).
- (iii) Let $A \in \mathbb{R}^{n \times n}$, the eigenvalues of A are the same ones as of A^T . Proof: $\det(A - zI) = \det((A - zI)^T) = \det(A^T - zI)$ (Prop. 8.3.5).
- (iv) Let $Q \in \mathbb{R}^{n \times n}$ be an orthogonal matrix, if λ is an eigenvalue of Q , then $|\lambda| = 1$. Proof: $\|v\|^2 = \|Qv\|^2 = \|\lambda v\|^2 = |\lambda| \cdot \|v\|^2$ (Prop. 8.2.7).
- (vi) Let $A \in \mathbb{R}^{n \times n}$, if (λ, v) is an eigenvalue-eigenvector pair of A , then $(\bar{\lambda}, \bar{v})$ is an eigenvalue-eigenvector pair of A too. Thus, if $\lambda \in \mathbb{C}$ is an eigenvalue of A , then $\bar{\lambda}$ is also an eigenvalue of A (Prop. 8.2.8).
- (vii) Let $P \in \mathbb{R}^{n \times n}$ be a projection matrix, then P has two eigenvalues, 0 and 1 and a *complete set* of eigenvectors (Prop. 7.1.21).
- (viii) Let $D \in \mathbb{R}^{n \times n}$ be a diagonal matrix, then its eigenvalues are its diagonal entries and the canonical basis is a set of eigenvectors of D (Ex. 7.1.23).
- (ix) Let $T \in \mathbb{R}^{n \times n}$ be a triangular matrix, then its eigenvalues are its diagonal entries, however, T might not have a complete set of eigenvectors (Fact 45).

12.6 Important Words Of Caution

- (i) Even though the **eigenvalues** of A and A^T are the same, **the eigenvectors are not!**
- (ii) The eigenvalues of $A + B$ **are not the sum** of the eigenvalues of A and the eigenvalues of B !
- (iii) The eigenvalues of AB **are not the product** of the eigenvalues of A and the eigenvalues of B !
- (iv) **Gaussian Elimination does not preserve eigenvalues and eigenvectors!**

12.7 Distinct Eigenvalues (Theo. 7.1.9)

Let $A \in \mathbb{R}^{n \times n}$ with n distinct, real eigenvalues, then there is a basis of \mathbb{R}^n made up of eigenvectors of A . We also say that A has a *complete set* of eigenvectors (Def. 7.1.20).

12.8 Complex Numbers

Complex numbers are of the form $z = (a + ib) \in \mathbb{C}$ with $a, b \in \mathbb{R}$ and $i^2 = -1$. The following operations are defined:

- (i) $(a + ib) + (x + iy) = (a + x) + i(b + y)$.
- (ii) $(a + ib) \cdot (x + iy) = (ax - by) + i(ay + bx)$.
- (iii) $(a + ib) \cdot (a - ib) = a^2 + b^2$.
- (iv) $\frac{a+ib}{x+iy} = \frac{(a+ib)(x-iy)}{(x+iy)(x-iy)} = \frac{(ax+by)+i(bx-ay)}{(x^2+y^2)} = \left(\frac{ax+by}{x^2+y^2} \right) + i \left(\frac{bx-ay}{x^2+y^2} \right)$ (modulus).
- (v) $|z| = \sqrt{a^2 + b^2}$ (conjugate).
- (vi) $a + \bar{b} = a - ib$
- (vii) $|z|^2 = z\bar{z}$.
- (viii) $\overline{z_1 + z_2} = \overline{z_1} + \overline{z_2}$.
- (ix) $\frac{1}{z} = \frac{\bar{z}}{|z|^2}$.

A complex number $z = (a + ib) \in \mathbb{C}$ can be written as $z = re^{i\theta}$ where $r = |z|$ and $\theta = \tan^{-1}(\frac{b}{a})$.

12.8.1 Complex Matrices/Vectors

Let $A \in \mathbb{C}^{m \times n}$, the *conjugate transpose* $A^* = \bar{A}^T$.

Given $\mathbf{v} \in \mathbb{C}^n$, we have

$$\|\mathbf{v}\|^2 = \mathbf{v}^* \mathbf{v} = \bar{\mathbf{v}}^T \mathbf{v} = \sum_{i=1}^n \bar{v}_i v_i = \sum_{i=1}^n |v_i|^2.$$

The dot-product in \mathbb{C}^n is given by $\langle \mathbf{v}, \mathbf{w} \rangle = \mathbf{w}^* \mathbf{v}$.

12.9 Fund. Theorem of Algebra (Cor. 8.1.3)

Any degree $n \geq 1$ polynomial $P(z) = \alpha_n z^n + \dots + a_1 z + a_0$ with $a_n \neq 0$ has n zeros: $\lambda_1, \dots, \lambda_n \in \mathbb{C}$ such that

$$P(z) = \alpha_n(z - \lambda_1)\dots(z - \lambda_n).$$

The number of times $\lambda \in \mathbb{C}$ appears in this expression is called the *algebraic multiplicity* of the zero (i.e. of the $P(\lambda) = 0$). This guarantees that an $n \times n$ matrix always has exactly n eigenvalues (if you count repeats and complex eigenvalues).

12.10 Geometric Multiplicity

Let $A \in \mathbb{R}^{n \times n}$ with eigenvalue λ , we call the dimension of $N(A - \lambda I)$ the *geometric multiplicity* of λ .

12.11 Characteristic Polynomial

Let $A \in \mathbb{R}^{m \times n}$, the *characteristic polynomial* of A is

$$\begin{aligned} P(z) &= (-1)^n \det(A - zI) = (z - \lambda_1)(z - \lambda_2)\dots(z - \lambda_n) \\ &= z^n + \underbrace{\left(- \sum_{i=1}^n \lambda_i \right) z^{n-1}}_{-\text{Tr}(A)} + \underbrace{\sum_{k=1}^{n-2} b_k z^k}_{\text{messy middle terms}} + (-1)^n \underbrace{\prod_{i=1}^n \lambda_i}_{(-1)^n \det(A)}. \end{aligned}$$

12.12 Trace and Determinant (Prop. 7.1.12)

Let $A \in \mathbb{R}^{n \times n}$ and $\lambda_1, \dots, \lambda_n$ its n eigenvalues, then

$$\text{Tr}(A) = \sum_{i=1}^n \lambda_i \text{ and } \det(A) = \prod_{i=1}^n \lambda_i.$$

Following, for matrices A, B and $C \in \mathbb{R}^{n \times n}$,

- (i) $\text{Tr}(AB) = \text{Tr}(BA)$,
- (ii) $\text{Tr}(ABC) = \text{Tr}(BCA) = \text{Tr}(CAB)$.

12.13 Diagonalization (Theo. 7.2.1)

Let $A \in \mathbb{R}^{n \times n}$ be a matrix with a complete set of eigenvectors.

Let $V = \begin{bmatrix} | & | & | \\ v_1 & \dots & v_n \end{bmatrix} \in \mathbb{R}^{n \times n}$ be the matrix whose columns are the eigenvectors, and $\Lambda \in \mathbb{R}^{n \times n}$ the matrix whose diagonal entries are the eigenvalues ($\Lambda_{ii} = \lambda_i$ for all $i \in [n]$), then

$$A = V\Lambda V^{-1}.$$

12.14 Diagonalizable Matrix (Def. 7.2.2)

A matrix $A \in \mathbb{R}^{n \times n}$ is said to be *diagonalizable* if there exists an invertible matrix V , such that $VAV^{-1} = \Lambda$, where Λ is a diagonal matrix.

12.15 Similar Matrices (Def. 7.2.3)

Two matrices A and $B \in \mathbb{R}^{n \times n}$ are *similar*, if exists an invertible matrix S , such that

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

Similar matrices are clones of each other. They represent the exact same linear transformation, just viewed from a different coordinate system. Similar matrices have the same eigenvalues (Prop. 7.2.4).

12.16 Example

Assume that $A, B \in \mathbb{R}^{n \times n}$ are similar, prove that their characteristic polynomials are equal.

As A and B are similar, there exists a matrix S such that $B = S^{-1}AS$. Recall that $\det(S)\det(S^{-1}) = 1$. Thus, $\det(A - zI) = \det(S^{-1})\det(A - zI)\det(S) = \det(S^{-1}(A - zI)S) = \det(B - zS^{-1}IS) = \det(B - zI)$.

12.17 Spectral Theorem (Theo. 7.3.1)

Any symmetric matrix $A \in \mathbb{R}^{n \times n}$ has n real eigenvalues and an orthonormal basis made of eigenvectors of A .

12.17.1 Diag. for Sym. Matrices (Cor. 7.3.2)

For any symmetric matrix $A \in \mathbb{R}^{n \times n}$, there exists an orthogonal matrix $V \in \mathbb{R}^{n \times n}$ (whose columns are the eigenvectors of A) and a diagonal matrix Λ whose entries are the eigenvalues of A , such that

$$A = V\Lambda V^T \text{ and } V^TV = I.$$

(This is also called the *eigendecomposition*).

12.18 Example

Find a matrix A with orthonormal eigenvectors

$$\mathbf{v}_1 = \frac{1}{9} \begin{bmatrix} 1 \\ 8 \\ -4 \end{bmatrix}, \mathbf{v}_2 = \frac{1}{9} \begin{bmatrix} -4 \\ 4 \\ 7 \end{bmatrix}, \mathbf{v}_3 = \frac{1}{9} \begin{bmatrix} 8 \\ 1 \\ 4 \end{bmatrix}$$

and corresponding eigenvalues $\lambda_1 = 1, \lambda_2 = -1, \lambda_3 = 0$.

Let V be the 3×3 matrix with $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ as its columns and D the diagonal matrix with $\lambda_1, \lambda_2, \lambda_3$ on its diagonals, then $A = VDV^T$.

12.19 Eigenvalues of Sym. Matrices

The rank of a real, symmetric matrix A is the number of non-zero eigenvalues (Cor. 7.3.4).

Every symmetric matrix has a real eigenvalue (Cor. 7.3.8).

12.20 Rayleigh Quotient (Prop 7.3.10)

Let $A \in \mathbb{R}^{n \times n}$ be a symmetric matrix, the *Rayleigh Quotient*, defined for $x \in \mathbb{R}^n \setminus \{0\}$, as

$$R(x) = \frac{x^T Ax}{x^T x},$$

attains its maximum at $R(\mathbf{v}_{\max}) = \lambda_{\max}$ and its minimum at $R(\mathbf{v}_{\min}) = \lambda_{\min}$, where λ_{\max} and λ_{\min} are the maximum and minimum eigenvalues of A and \mathbf{v}_{\max} and \mathbf{v}_{\min} their associated eigenvectors.

12.21 Positive (Semi)-Definite (Def. 7.3.11)

A symmetric matrix $A \in \mathbb{R}^{n \times n}$ is said to be *positive semidefinite* (PSD) if all its eigenvalues are ≥ 0 and *positive definite* they are > 0 . Moreover, (as per Prop. 7.3.12) A is

- (i) PSD $\iff x^T Ax \geq 0$ for all $x \in \mathbb{R}^n \setminus \{0\}$,
- (ii) PD $\iff x^T Ax > 0$ for all $x \in \mathbb{R}^n \setminus \{0\}$.

Given two matrices A and B that are PSD (PD), their sum is also PSD (PD) (Fact 7.3.13).

12.22 Gram Matrix (Def. 7.3.14)

Let $V \in \mathbb{R}^{m \times n}$, the Gram matrix of V is the inner product of the columns of V , i.e.,

$$G = V^T V.$$

Sometimes VV^T is also called a Gram matrix of V (the inner product of the rows) (Remark 7.3.15).

12.23 Gram and Eigenvalues (Prop. 7.3.16)

Let $A \in \mathbb{R}^{m \times n}$, the non-zero eigenvalues of $A^T A$ are the same as the ones of AA^T . Both matrices are symmetric and PSD.

12.24 Cholesky Decomposition (Prop. 7.3.17)

Every symmetric, PSD matrix M is a gram matrix of an upper triangular matrix C , i.e.,

$$M = C^T C.$$

12.24.1 Calculating the Cholesky Decomposition

- (i) Let M be symmetric and PSD, the eigendecomposition (Cor. 7.3.2) gives us $M = V\Lambda V^T$
- (ii) We build $\Lambda^{\frac{1}{2}}$ by taking the square root of each entry of Λ , following, $M = (V\Lambda^{\frac{1}{2}})(V\Lambda^{\frac{1}{2}})^T$.
- (iii) We then take the QR decomposition $(V\Lambda^{\frac{1}{2}})^T = QR$, following, $M = (QR)^T(QR) = R^T Q^T QR = R^T R$.

13 Singular Value Decomposition (Def. 8.1.1)

Let $A \in \mathbb{R}^{m \times n}$. There exists orthogonal matrices $U \in \mathbb{R}^{m \times m}$ and $V \in \mathbb{R}^{n \times n}$ such that

$$A = U\Sigma V^T,$$

where $\Sigma \in \mathbb{R}^{m \times n}$ is a diagonal matrix whose entries are non-zero and ordered in descending order. Moreover,

- (i) $U^T U = V^T V = I$.
- (ii) $\sigma_1 \geq \dots \geq \sigma_{\min\{m,n\}}$.
- (iii) The columns of U are called the *left singular vectors*.
- (iv) The columns of V are called the *right singular vectors*.
- (v) The diagonal entries of Σ are called the *singular values*.

13.1 Existence of SVD (Theo. 8.1.5)

Every matrix $A \in \mathbb{R}^{m \times n}$ has a SVD.

13.2 Compact Form (Remark 8.1.2)

If A has rank r , the SVD can be written in a compact form:

$$A = U_r \Sigma_r V_r^T,$$

where U_r and V_r contain the first r left/right singular vectors respectively and Σ_r contains the first r singular values.

13.3 Calculating the SVD (Remark 8.1.3)

Let $A \in \mathbb{R}^{m \times n}$ and $A = U\Sigma V^T$ be its SVD, then

$$AA^T = U(\Sigma\Sigma^T)U^T \text{ and } A^T A = V(\Sigma^T\Sigma)V^T.$$

In other words, the SVD of A can be calculated by

- (i) Taking the eigendecomposition of AA^T or $A^T A$.
- (ii) Then U corresponds to the eigenvectors of AA^T ,
- (iii) V to the eigenvectors of $A^T A$,
- (iv) Σ to the square roots of the shared eigenvalues.

13.4 Rank r Matrices (Prop. 8.1.4)

Let $A \in \mathbb{R}^{m \times n}$ with rank r and $\sigma_1, \dots, \sigma_r$ be the non-zero singular values of A , u_1, \dots, u_r the corresponding left singular vectors and v_1, \dots, v_r the corresponding right singular vectors. Then $A = \sum_{k=1}^r \sigma_k u_k v_k^T$.

14 Vector and Matrix Norms

14.1 Vector Norms

For $1 \leq p \leq \infty$, the l_p norm is given by

$$\|x\|_p = \left(\sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}},$$

for $p < \infty$, and $\|x\|_\infty = \max_i |x_i|$.

In particular let $\|x\|_p$ for $p = 1$ denote the *Manhattan distance*.

14.2 Matrix Norms (Def. 8.2.1)

Let $A \in \mathbb{R}^{m \times n}$ then consider

$$\begin{aligned} \text{Frobenius norm} \quad \|A\|_F &:= \sqrt{\sum_{i=1}^m \sum_{j=1}^n A_{ij}^2} \\ \text{operator/spectral norm} \quad \|A\|_{\text{op}} &:= \max_{x \in \mathbb{R}^n \text{ s.t. } \|x\|=1} \|Ax\| \end{aligned}$$

15 Multiple Choice Solutions

15.1 HS23

1. Let $A \in \mathbb{R}^{5 \times 6}$ with $\text{rank}(A) = 5$ and $b \in \mathbb{R}^5$ be arbitrary. Which of the following statements must be true?
 - The linear system $Ax = b$ has no solution, i.e. there is no $x \in \mathbb{R}^6$ such that $Ax = b$.
 - The linear system $Ax = b$ has exactly one solution, i.e. there is exactly one $x \in \mathbb{R}^6$ such that $Ax = b$.
 - **The linear system $Ax = b$ has infinitely many solutions, i.e. there exists infinitely many $x \in \mathbb{R}^6$ such that $Ax = b$.**
 - The linear system $Ax = b$ has exactly two solutions, i.e. there are exactly two $x \in \mathbb{R}^6$ such that $Ax = b$.
2. Let $n \in \mathbb{N}^+$. Consider arbitrary non-zero vectors $v, w \in \mathbb{R}^n$. Let A be the $n \times n$ matrix $A = vw^T$. Which of the following statements must be true?
 - A has a real eigenvalue λ with $\lambda \neq 0$.
 - **The largest singular value of A is $\sigma = \|v\|\|w\|$.**
 - If v is orthogonal to w , then $A = 0$.
 - $\det(A - \lambda I) = (-1)^n \lambda^n$ for all $\lambda \in \mathbb{R}$.
3. Let $n \in \mathbb{N}^+$. We call a matrix $A \in \mathbb{R}^{n \times n}$ nilpotent if there exists a $k \in \mathbb{N}^+$ such that $A^k = 0$. Which of the following statements must be true?
 - Every $A \in \mathbb{R}^{n \times n}$ with nullspace $N(A) \neq \{0\}$ is nilpotent.
 - If $A \in \mathbb{R}^{n \times n}$ is nilpotent and $B \in \mathbb{R}^{n \times n}$ is nilpotent, then $A + B$ is nilpotent.
 - **If $A \in \mathbb{R}^{n \times n}$ is nilpotent, then $\text{rank}(A) < n$.**
 - If $A \in \mathbb{R}^{n \times n}$ is nilpotent and $B \in \mathbb{R}^{n \times n}$ is nilpotent, then AB is nilpotent.
4. Let $n \in \mathbb{N}^+$ and let $A \in \mathbb{R}^{n \times n}$ be a symmetric positive semidefinite matrix. Which of the following statements must be true?
 - $AB = BA$ for all matrices $B \in \mathbb{R}^{n \times n}$.
 - All diagonal entries of A are non-zero
 - The trace of A is strictly positive.
 - **All diagonal entries of A are non-negative.**
6. Consider the 2×2 matrix $A = \begin{bmatrix} 3 & 2 \\ -4 & -3 \end{bmatrix}$ with eigenvalues 1 and -1 . Which of the following statements must be true?
 - $A^{1024} = A$.
 - $A^{1024} = 2A$.
 - **$A^{1024} = I$.**
 - $A^{1024} = 2I$.

15.2 FS24

1. Let $A \in \mathbb{R}(5 \times 5)$ with $\sum_{j=1}^5 A_{ij} = 0$ for all $i \in \{1, 2, \dots, 5\}$.

Which of the following statements must be true?

- A is the zero matrix, i.e. $A = 0$.
- A has an inverse A^{-1} .
- **The rank of A is less than 5, i.e. $\text{rank}(A) < 5$.**
- There are only finitely many vectors $x \in \mathbb{R}^5$ with $Ax = 0$.

2. Let $n \in \mathbb{N}^+$ and let $A \in \mathbb{R}^{n \times n}$ be an arbitrary symmetric and positive semidefinite matrix. Which of the following statements must be true?

- $\text{Tr}(A^2) = \text{Tr}(A)^2$.
- $\text{Tr}(A^2) \leq \text{Tr}(A)^2$.
- $\text{Tr}(A^2) \geq \text{Tr}(A)^2$.
- $\text{Tr}(A^2) \neq \text{Tr}(A)^2$.

6. Let $m, n \in \mathbb{N}^+$ with $m < n$. Let $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$ and consider the system of equations $Ax = b$. Which of the following statements must be true?

- The system $Ax = b$ has infinitely many solutions.
- There exist choices for A and b such that the system has a unique solution.
- **The system $Ax = b$ either has no solution or infinitely many.**
- If the system $Ax = b$ has a solution for some specific b , then it must have a solution for all possible choices of b .