# Linear Algebra

Janis Hutz
https://janishutz.com

January 9, 2026

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

*"Lineare Algebra ist Kulturgut"*
- Robert Weismantel, 2024

HS2024, ETHZ
Exam-Summary of the lecture notes
https://ti.inf.ethz.ch/ew/courses/LA24/notes_part_I.pdf

and

https://ti.inf.ethz.ch/ew/courses/LA24/notes_part_II.pdf

# Contents

# 0    Introduction

This summary is designed to provide the optimal balance between depth and broadness for the exam. To print this summary, print two pages per A4 paper (so A5 size), leaving out the first three pages (Title page, Table of Contents, Introduction)

The introductory sections were kept as short as possible to allow for more space for the more difficult parts

# 1 Vectors

*SIMPLIFY RESULTS AS MUCH AS POSSIBLE* $(21 \div 12 = 7 \div 4)$

**D 1.5:** **Linear Combination**: $\sum_{i=1}^{n} \lambda_i v_i$, scaled combination of $n$ vectors $v_i$.

**D 1.8:** It is called an **Affine Combination** if $\sum_{i=1}^{n} \lambda_i = 1$, a **Conic Combination** if $\lambda_i \geq 0$ for $i = 1, 2, \ldots, n$ and a **Convex Combination** if it is affine and conic. Linear combination as a matrix: $Ax = b$, where the matrix's columns are the vectors $v_i$ and $x$'s components are $\lambda_i$   **D 1.10:** **Scalar-Product**: Multiply component wise, add all components together. Results in number, alternatively $v^\top w$ (S2.2.3), or $\sum_{i=1}^{n} v_i w_i$;



linear          (i) affine          (ii) conic          (iii) convex
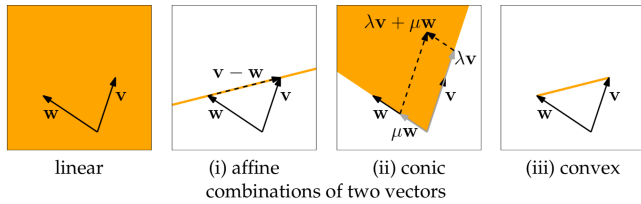        combinations of two vectors

Figure 1.1: Different kinds of combinations of two vectors

**Dot-Free Notation**: Similar to summation notation, we can use it to define e.g. vector addition: $[v_i]_{i=1}^m + [w_i]_{i=1}^m := [v_i + w_i]_{i=1}^m$;   **D 1.12:** **Euclidean Norm**: $||v|| = \sqrt{v \cdot v} = \sqrt{v^\top v}$; **Squared norm**: $v^\top v = ||v||^2$ **Unit-Vector**: $||v|| = 1$, obtaining: $\frac{v}{||v||} = \frac{1}{||v||} \cdot v$;

**L 1.13:** **Cauchy-Schwarz inequality**: $|v \cdot w| \leq ||v|| \cdot ||w||$ for any vectors $v, w \in \mathbb{R}^m$. Equality if $v\lambda = w$ or $w\lambda = v$;   **D 1.15:** **Angles**: $\cos(\alpha) = \frac{v \cdot w}{||v|| \cdot ||w||}$. If $v \perp w \in \mathbb{R}^m$, then $v \cdot w = 0$;   **L 1.17:** **Triangle inequality**: $||v + w|| \leq ||v|| + ||w||$; **Hyperplane**: $\{v \in \mathbb{R}^n : v \cdot d = 0\}$, $d \in \mathbb{R}^n$, $d \neq 0$;   **L 1.20:** **Linear independence**: Vectors are linearly independent if (a) no vector is a linear combination of the others or (b) there are no $\lambda_i$, such that $\sum_{i=1}^{n} \lambda_i v_i = 0$. For matrices: $\neg \exists x \neq 0 : Ax = 0$ (no column vector is a linear combination of another).   **D 1.23:** **Span of vectors**: Set of all linear combinations of a vector; **Standard Basis vector**: Vector with just one component being 1, all others 0;

# 2 Matrices

**Size**: $m \times n$: $m$ rows, $n$ cols (*Zeilen zuerst, Spalten später*);   **D 2.3:** **Addition**: $A + B := [a_{ij} + b_{ij}]_{i=1, j=1}^{m, \ n}$; **Scalar-multiple**: $\lambda A := [\lambda a_{ij}]_{i=1, j=1}^{m, \ n}$;   **D 2.4:** **Square matrices**: **Identity matrix**: quadratic matrix, diagonals 1, $A = AI = IA$; **Diagonal matrix**: $a_{ij} = 0$ if $i \neq j$; **Triangle matrix**: *lower* if $a_{ij} = 0$ for $i < j$, *upper* else; **Symmetric matrix**: $a_{ij} = a_{ji} \forall i, j$, $A^\top = A$;

**D 2.5:** **Matrix-Vector-Product**: Rows of matrix ($m \times n$) with vector ($n$ elements), i.e. $u_1 = \sum_{i=1}^{m} a_{1,i} \cdot v_i$, $Ix = x$; **Trace**: Sum of the diagonal entries;

**D 2.9:** **Column Space**: $\{Ax : x \in \mathbb{R}^n\}$, i.e. the span of the column vectors;   **D 2.10:** **Rank**: $\text{rank}(A) :=$ the number of linearly independent column vectors;   **D 2.12:** **Transpose**: Mirror the matrix along its diagonal. **L 2.20:** $(AB)^\top = B^\top A^\top$, $(A^\top)^\top = A$ (O2.12);   **D 2.14:** **Row Space**: $R(A) = C(A^\top)$;

**D 2.17:** **Matrix Multiplication**: $A \times B = C$, $c_{ij} = \sum_{k=1}^{n} a_{i,k} b_{k,j}$. Dimension restrictions: $A$ is $m \times n$, $B$ is $n \times p$, result is $m \times p$. *For each entry, multiply the i-th row of A with the j-th column of B.* **Not** commutative, but associative & distributive;   **L 2.22:** **Outer product**: $\text{rank}(A) = 1 \Leftrightarrow \exists$ non-zero vectors $v \in \mathbb{R}^m$, $w \in \mathbb{R}^n$, s.t. $A$ is outer product, i.e. $A = vw^\top$, thus $\text{rank}(vw^\top) = 1$

**Rotation matrix**: $R(\xi) = \begin{bmatrix} \cos(\xi) & -\sin(\xi) \\ \sin(\xi) & \cos(\xi) \end{bmatrix}$;

**T 2.24:** **CR-Decomposition**: $A = CR$. Get $R$ from (reduced) row echelon form, $C$ is the columns from $A$ where there is a pivot in $R$. $C \in \mathbb{R}^{m \times r}, R \in \mathbb{R}^{r \times n}$ (in RREF), $r = \text{rank}(A)$; Row Echelon Form: To find REF, try to create pivots: $R_0 = \begin{bmatrix} 1 & 0 & 2 & 3 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$, use Gauss-Jordan-Elimination to find it (row-transformations); Reduced REF: RREF is simply REF without any zero rows (i.e. in $R_0$, $R$ (in RREF) would be $R_0$ without the last row).

**D 2.26:** **Transformations**: A matrix can be understood as a re-mapping of the unit vectors, scaling and re-orienting them. Each column vector can then be understood as the new unit vector $e_i$, hence essentially adding another coordinate system to the original one, which is moved and rotated a certain way. The rotation matrix under 2 is such an example. To prove that $T$ is linear transformation, use: $T(x + y) = T(x) + T(y)$ and $T(\lambda x) = \lambda T(x)$.

Then insert the linear transformation given by the task and replace $x$ (or whatever variable there is) with $x + y$ or $\lambda x$. $Ax = \sum_{i=1}^n (x_i v_i)$, where $v_i$ is the $i$-th column of $A$; **D 2.26:** **Kernel**: $\text{Ker}(T) := \{x \in \mathbb{R}^n : T(x) = 0\} \subseteq \mathbb{R}^n$; **Image**: $\text{Im}(t) := \{T(x) : x \in \mathbb{R}^n\} \subseteq \mathbb{R}^m$;

# 3   Solving Linear Equations

Put the system of linear equation's factors (i.e. for a linear equation $ax + by + cz = u$, we would put $a, b, c$ into the matrix and $u$ into the vector) into a matrix, where each row is an equation and the result into a vector $b$. Then, we solve $Ax = b$ by Gauss elimination.

**Gauss elimination**: Transform matrix $A$ into upper triangle matrix by performing row transformations (adding, adding scalar multiples, multiplying by scalar) on it. All operations performed on $A$ have to also be performed on $b$. Typically, write down both as a matrix with a dividing line between $A$ and $b$. Then solve by back-substitution. Gauss elimination succeeds iff $A \in \mathbb{R}^{m \times m}$ and $A$'s columns are linearly independent. (Runtime: $\mathcal{O}\left(m^3\right)$)

## 3.1   Inverse

**D 3.8:** Inverse: Perform Gauss-elimination on a matrix of form $\begin{bmatrix} A \mid I \end{bmatrix}$ until we get $\begin{bmatrix} I \mid A^{-1} \end{bmatrix}$. $A$ is invertible, iff $\det(A) \neq 0$. Alternative: $MM^{-1} = M^{-1}M = I$. $M$ has to be square. 0 matrix has no inverse

**Inverse for specific sizes**: $1 \times 1$ $M = [a]$, $M^{-1} = [\frac{1}{a}]$ (if $a \neq 0$); $2 \times 2$ $M = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ $M^{-1} = \frac{1}{\det(M)} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$;

**L 3.10:** **Inverse product**: $(AB)^{-1} = B^{-1}A^{-1}$; **L 3.11:** $(A^{-1})^\top = (A^\top)^{-1}$; **T 3.12:** **Inverse theorem**: $A$ is invertible $\Leftrightarrow Ax = b$ has a unique solution $\forall b \in \mathbb{R}^n \Leftrightarrow$ the columns of $A$ are independent. **Diagonal matrix**: each element reciprocal. (Might require proof)

## 3.2   LU-Decomposition

**T 3.14:** **LU-Decomposition**: $A = LU$. $U$ upper triangle, result of Gauss elimination, $L$ lower triangle, $(E_1 \times E_2 \times \ldots \times E_n)^{-1}$. Transformation matrices $E$ $(E \cdot A = A_1)$: transformation is a single entry in lower triangle, where the $i$ and $j$ are the two rows involved and the value of $e_{ij}$ is the operation performed on the two. $L^{-1}$ for size 3: Diagonal, $L_{1,2}^{-1} = -L_{1,2}$, $L_{2,3}^{-1} = -L_{2,3}$, $L_{1,3}^{-1} = -L_{1,3}$. When multiplying up to three different ones, copies values to multiplied matrix. If it is impossible to decompose $A$ into $LU$ without row exchanges, we get $PA = LU$, where $P$ is a permutation matrix (indicating which rows have been swapped). Time complexity is improved significantly with this $\mathcal{O}\left(m^2\right)$.

**D 3.15:** **Permutations**: bijective function $\pi$ on matrix; Reorders the input structure (i.e. vector or matrix);

**D 3.16:** **Permutation matrix**: $p_{ij} = \begin{cases} 1 & \text{if } j = \pi(i) \\ 0 & \text{else} \end{cases}$ **L 3.17:** $P^{-1} = P^\top$.

**T 3.19:** **LUP-Decomposition**: $PA = LU$, where $U = P \times LA$. $P_j = I$ if no row swaps are performed at each step and $P = P_m \times P_{m-1} \times \ldots \times P_1$. Rewriting this as $A = P^\top LU$, we can simply solve an SLE using LUP-Decomposition

## 3.3   Gauss-Jordan-Elimination

**Gauss-Jordan-Elimination**: Generalization of Gauss elimination to $m \times n$ matrices, still works similarly to Gauss elimination. We aim to find REF or RREF (see 2). To describe, we say $\text{REF}(j_1, j_2, \ldots, j_r)$ or equivalently with RREF, where $j_r$ is the $r$-th pivot.

The solution is then in a vector, whose components are either 0 or the $r$-th component of $b$. Example:

$$\begin{bmatrix} 0 & 1 & 0 & 0 & 2 & 0 \\ 0 & 0 & 1 & 0 & 3 & 0 \\ 0 & 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ b_1 \\ b_2 \\ b_3 \\ 0 \\ b_4 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \\ 0 \end{bmatrix}$$

If the green marked entry in $b$ were not to be 0, then the SLE would not have a solution.

**CR-Decomposition**: see 2 for exaplanation. **T 3.25:** is described there

# 4 The four fundamental subspaces

## 4.1 Vector space

**D 4.1:** **Vector space**: Abstract mathematical concept, elements are vectors; Two operations ($\oplus$/+ and $\odot$/· (depending on notation)); Triple $(V, +, \cdot)$, $V$ is a set of vectors, satisfying the vector space axioms *commutativity, associativity, existance of zero and negative vectors and identity element (1), compatibility of $\oplus$ with · (in $\mathbb{R}$), distributivity over $\oplus$ ($\lambda(v + w) = \lambda v + \lambda w$) and distributivity over + (in $\mathbb{R}$) ($(\lambda + \mu)v = \lambda v + \mu v$).*

**Defining a vector space**: We need to define addition and scalar multiplication for the elements in a ***canonical*** way (= *according to accepted standard*)

**D 4.9:** **Subspaces**: Two subset axioms, $v + w \in U$, $\lambda v \in U$, $\lambda \in \mathbb{R}$ (closed under addition & scalar multiplcation), where $U \subseteq V$, **L 4.10:** at least ***always*** contains the zero vector ($0 \in U$); $C(A) = \{Ax : x \in \mathbb{R}^n\}$ is a subspace of $\mathbb{R}^m$, if $A \in \mathbb{R}^{m \times n}$; **L 4.13:** $U$ is also a vector space with the same + and · as $V$; $U \cap W$ is subspace, $U \cup W$ not;

**Prove of vector space / subspace** : Use the axioms and prove each and every one of them, then, if that is successful, conclude that is subspace / vector space. **D 4.16:** **Span**$(G)$: Set of all linear combinations of $G$, $G \subseteq V$. Set is **linearly independent** if no vector $v \in G$ is a linear combination of $G \backslash \{v\}$. Vectors are not linearly independent if non-trivial linear combination of them is 0-vector

**D 4.17:** **Basis**: Set of vectors that are linearly independent and span $B$ (number equals dimension of space), subspace of $V$. For $\mathbb{R}^m$, the set of unit vectors is a basis. For a matrix, all linearly independent columns form a basis of the column space $C(A)$. Every set of $m$ linearly independent vectors is a basis of $\mathbb{R}^m$. ***Calculating:*** If we have a matrix with full column / row rank, then the basis are all column / row vectors.

**L 4.20:** **Steiniz exchange lemma**: Let $F \subseteq V$ be a finite set of linearly independent vectors, $G \subseteq V$ a finite set of vectors with Span$(G) = V$. Then: $|F| \leq |G|$, $\exists$ subset $E \subseteq G$ of size $|G| - |F|$ s.t. Span$(F \cup E) = V$. **T 4.21:** $B, B' \subseteq V$, finite bases of $V$, then $|B| = |B'|$; **D 4.22:** **Finitely generated vector space**: $\exists G \subseteq V$ with Span$(G) = V$. **T 4.23:** If $V$ finitely generated, then $V$ has a basis $B \subseteq G$.

**D 4.24:** **Dimension**: $V$ finitely generated, then $d = \dim(V)$ is the size of any basis $B$ of $V$. **L 4.25:** Let $F \subseteq V$ be a set of $d$ linearly independent vectors, then $F$ is basis of $V$. Let $G$ be a set of $d$ vectors with Span$(G) = V$, then $G$ is a basis of $V$

## 4.2 Computing the fundamental subspaces

Let $M$ be an invertible $m \times m$ matrix.

**Column Space**: Set of all linearly indep. columns ($C(A) = \{Ax : x \in \mathbb{R}^n\} \subseteq \mathbb{R}^m$. **T 4.26:** The number of pivots in REF form of matrix $A$ is the rank$(A) = \dim(C(A))$. The column vectors at the pivots form the basis of $A$;

**Row Space**: $R(A) = C(A^\top) \subseteq \mathbb{R}^n$. **L 4.28:** $R(A) = R(MA)$. **T 4.29:** $\dim(R(A)) = $ rank$(A)$ **T 4.30:** rank$(A) = $ rank$(A^\top)$; The rows in RREF form a basis of $A$;

**D 4.32:** **Null Space**: $N(A) = \text{Ker}(A) = \{x \in \mathbb{R}^n : Ax = 0\}$. **L 4.33:** $N(A) \subseteq \mathbb{R}^n$. **L 4.34:** $N(A) = N(MA)$. Can be easily obtained from RREF by setting $b = 0$ in $Rx = b$. If there are any free variables, choose any real (or complex) number satisfying the condition. **To find basis**, we can rewrite and apply below Lemma

$Rx = 0 \Leftrightarrow I \cdot x(I) + Q \cdot x(Q) = 0$, e.g. for $R = \begin{bmatrix} 1 & 2 & 0 & 3 \\ 0 & 0 & 1 & -2 \end{bmatrix}$ to $I \begin{bmatrix} x_1 \\ x_3 \end{bmatrix} + \begin{bmatrix} 2 & 3 \\ 0 & -2 \end{bmatrix} \begin{bmatrix} x_2 \\ x_4 \end{bmatrix} = 0 \Leftrightarrow x(I) = -Q \cdot x(Q)$

We may then freely choose the *free variables* $x(Q)$, then find *basic variables* $x(I)$ using the above, typically choose $e_1, \ldots, e_k$ for $x(Q)$ to get the $k$-the vec of basis ($k$ = num cols of $Q$). Finally combine the vectors into one.

**L 4.35:** $N(R) = N(A)$. The vectors obtained from the above procedure form a basis of $N(R)$

**T 4.36:** $\dim(N(A)) = n - $ rank$(A)$;

**D 4.37:** **Left Nullspace**: $LN(A) := N(A^\top)$. **L 4.38:** $LN(A) \subseteq \mathbb{R}^m$. **T 4.39:** $\dim(LN(A)) = m - $ rank$(A)$

**D 4.40:** **Solution space**: Set of all solutions of $Ax = b$, thus $\text{Sol}(A, b) := \{x \in \mathbb{R}^n : Ax = b\} \subseteq \mathbb{R}^n$. **T 4.41:** Let $s$ be some solution of $Ax = b$, then $\text{Sol}(A, b) = \{s + x : x \in N(A)\}$. This means, we can also compute $\text{Sol}(A, b)$ despite it not being a subspace. To describe all solutions, we simply need *some* solutions.

**L 4.42:** Let $A \in \mathbb{R}^{m \times n}$. If rank$(A) = m$, $Ax = b$ has a solution for every $b \in \mathbb{R}^m$

# 5   Orthogonality

## 5.1   Definition

**D 5.1:** **Orthogonality**: Two vectors are orthogonal, if the their scalar product is 0, i.e. $v^\top w = \sum_{i=1}^{n} v_i w_i = 0$.
**L 5.2:** Two subspaces are orthogonal to each other if for each $v \in V$ and $w \in W$, $v$ and $w$ are orthogonal.
**L 5.3:** As a consequence, if that is true, all these vectors are linearly independent.   **C 5.4:** $V \cap W = \{0\}$ and their union is $V \cup W = \{\lambda v + \mu w : \lambda, \mu \in \mathbb{R}, v \in V, w \in W\}$. If $\dim(V) = k$ and $\dim(W) = l$, then $\dim(V \cup W) = k + l \leq n$, for $V, W \subseteq \mathbb{R}^n$;

**D 5.5:** **Orthogonal complement**: $V^\perp := \{w \in \mathbb{R}^n : w^\top v = 0, \forall v \in V\}$.   **T 5.6:** $N(A) = C(A^\top)^\perp = R(A)^\perp$ and $C(A^\top) = N(A)^\perp$.   **T 5.7:** The following is equivalent for orthogonal subspaces of $\mathbb{R}^n$: $W = V^\perp \Leftrightarrow \dim(V) + \dim(W) = n \Leftrightarrow u = v + w \forall u \in \mathbb{R}^n$ with unique vectors $v \in V, w \in W$.   **L 5.8:** $V = (V^\perp)^\perp$.   **C 5.9:** $N(A) = C(A^\top)^\perp$ and $C(A^\top) = N(A)^\perp$   **T 5.10:** $\{x \in \mathbb{R}^n : Ax = b\} = x_1 + N(A)$, where $x_1 \in \mathbb{R}(A)$ such that $Ax_1 = b$
**C 5.11:** $N(A) = N(A^\top A)$ and $C(A^\top) = C(A^\top A)$

## 5.2   Projections

**D 5.1:** **Projection**: Projecting a vector onto a subspace is done with $\text{proj}_S(b) = \text{argmin}_{p \in S} ||b - p||$ and yields the closest point (or vector) in the new subspace $S$   **L 5.2:** **1-Dimensional**: $\text{proj}_S(b) = \dfrac{aa^\top}{a^\top a} b$, where we project $b \in \mathbb{R}^m$ on $S = \{\lambda a : \lambda \in \mathbb{R}\} = C(a)$ where $a \in \mathbb{R}^m \setminus \{0\}$. We note that $(b - \text{proj}_S(b)) \perp \text{proj}_S(b)$, i.e. the "error-vector" is perpendicular to $a$.

**L 5.3:** **General case**: PREFER 5.2.6! $S$ is a subspace in $\mathbb{R}^m$ with $\dim(S) = n$. Let $a_1, a_2, \ldots, a_n$ be a basis of $S$, i.e. $S = \text{Span}(a_1, \ldots, a_n) = C(A) = \{A\lambda : \lambda \in \mathbb{R}^n\}$ where $A$ is a matrix with column vectors $a_1, \ldots, a_n$. We project $b \in \mathbb{R}^m$ onto the subspace $S$, then $\text{proj}_S(b) = A\widehat{x}$, where $\widehat{x}$ satisfies $A^\top A \widehat{x} = A^\top b$.   **L 5.4:** $A^\top A$ is invertible $\Leftrightarrow A$ has linearly independent columns $\Rightarrow$   **C 5.5:** $A^\top A$ is square, invertible and symmetric.

**T 5.6:** Projection in terms of projection matrix $P = A(A^\top A)^{-1} A^\top$: $\text{proj}_S(b) = Pb$, $A$ is the matrix of task

## 5.3   Least squares, Linear regression

**Least squares**: Approximate a solution to System of equations. Concept $\min\limits_{\widehat{x} \in \mathbb{R}^n} ||A\widehat{x} - b||^2$. Using the normal equations, we get $A^\top A \widehat{x} = A^\top b$. Using the definition of $\widehat{x} = (A^\top A)^{-1} A^\top b$ to solve the least squares problem borders insanity, so use $A^\top A \widehat{x} = A^\top b$ to solve.

> ### Least squares                                                                       Usage
>
> (i) Calculate $M = A^\top A$ (matrix)
> (ii) Calculate $b' = A^\top b$ (vector)
> (iii) Solve resulting System of Equations $M\widehat{x} = b'$ normally

**Linear regression**: Application of least squares problem, problem is to find $A$ and $b$ such that we can solve the system. We define a matrix $A = \begin{bmatrix} 1 & t_1 \\ \vdots & \vdots \\ 1 & t_n \end{bmatrix}$ and a result vector $b = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$ where $n$ is the total number of data points and $t_i$ is the slope of the $i$-th function, where $b_i$ is its output. The first column is all 1s because the constant element has no scalar, for obvious reasons. This comes from the following concept: $f(t) = \alpha_0 + \alpha_1 t$, so if the first data point is $(1, 2)$, we get $\alpha + \alpha_1 \cdot 1 = 2$, which we will then transform into a SLE with other equations.

**L 5.3:** The columns in $A$ are linearly dependent $\Leftrightarrow t_i = t_j \, \forall i \neq j$

## 5.4   Gram Schmidt

**D 5.1:** **Orthonormal vector**: Orthogonal and norm is 1. Alternatively: $q_i^\top q_j = \delta_{ij}$, ***Kronecker delta*** $\delta_{ij} = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases}$;   **D 5.4:** **Orthogonal matrix**: If $Q^\top Q = I$, $QQ^\top = I$ (if $Q$ square), so $Q^{-1} = Q^\top$, columns of $Q$

form orthonormal basis for $\mathbb{R}^n$. **Ex 5.5:** Rotation & permutation matrices. **P 5.7:** Orthogonal matrices preserve norm and inner product of vectors. If $Q \in \mathbb{R}^{n \times n}$, then $\forall x, y \in \mathbb{R}^n$, $||Qx|| = ||x||$ and $(Qx)^\top (Qy) = x^\top y$; Product of any two orthogonal matrices is orthogonal; For $Q$ orthogonal, we want $a \cdot b + c \cdot d = 0$

**Projections with orthonormal bases**: Much simpler, because $A^\top A = I$ if $A$ has orthonormal columns. **P 5.8:** The least squares solution to $Qx = b$, where $Q$ is the matrix whose columns are the vectors forming the orthonormal base of $S \subseteq \mathbb{R}^m$, is given by $\hat{x} = Q^\top b$ and the projection matrix is given by $QQ^\top$;

**D 5.10:** **Gram-Schmidt**: Used to construct orthonormal bases. We have linearly independent vectors $a_1, \ldots, a_n$ that span a subspace $S$, then Gram-Schmidt constructs $q_1, \ldots, q_n$ by setting $q_1 = \frac{a_1}{||a_1||}$ and for $k = 2, \ldots, n$, $q_k' = a_k - \sum_{i=1}^{k-1} (a_k^\top q_i) q_i$ then setting $q_k = \frac{q_k'}{||q_k'||}$;

**D 5.12:** **QR-Decomposition**: $A = QR$, where $R = Q^\top A$ and $Q$ is obtained from the Gram-Schmidt process, it is made up of the vectors $q_i$ as columns. **L 5.13:** $R$ is upper triangle and invertible. $QQ^\top A = A$, meaning $A = QR$ is well-defined. **F 5.14:** This greatly simplifies calculations involving projections and least squares, since $C(A) = C(Q)$, so $\text{proj}_{C(A)}(b) = QQ^\top b$ and for least squares, we have $R\hat{x} = Q^\top b$. This can efficiently be solved because $R$ is triangular using back-substitution.

## 5.5   Pseudoinverse

**Pseudoinverse**: $A^+ = (A^\top A)^{-1} A^\top$; $\text{rank}(A) = \text{rank}(A^+)$

Let $A \in \mathbb{R}^{m \times n}$; **D 5.1:** **Full column rank**: $\text{rank}(A) = n$. $A^+ = (A^\top A)^{-1} A^\top$. **P 5.2:** $A$ full column rank, $A^+ A = I_n$ (left inverse); **D 5.3:** **Full row rank**: $\text{rank}(A) = m$. $A^+ = AA^\top (AA^\top)^{-1}$. **L 5.4:** $A$ full row rank, $AA^+ = I_m$; **L 5.5:** For any matrix $A$ and vector $b \in C(A)$, the unique solution for the least squares problem is given by vector $\hat{x} \in C(A^\top)$ and satisfies $A\hat{x} = b$. **P 5.6:** For a full row rank matrix $A$, the solution is given by $\hat{x} = A^+ b$ with $A\hat{x} = b$; **D 5.7:** **General case**: $A^+ = R^+ C^+ = R^\top (C^\top A R^\top)^{-1} C^\top$. We can use any full-rank factorization, not just $CR$, i.e. **P 5.10:** let $S \in \mathbb{R}^{m \times r}$ and $T \in \mathbb{R}^{r \times n}$ s.t. $A = ST$, then $A^+ = T^+ S^+$.

**T 5.12:** **Properties of Pseudoinverse**: $AA^+ A = A$, $A^+ AA^+ = A^+$, $AA^+$ is symmetric, it is the projection matrix for projections on $C(A)$. $A^+ A$ is symmetric, and is the projection matrix on $C(A^\top)$. $(A^\top)^+ = (A^+)^\top$;

## 5.6   Farka's Lemma & Projections of sets

**T 5.8:** **Farka's Lemma** Let $A \in Q^{m \times m}$, $b \in Q^m$. There either:

- exists a vector $x \in \mathbb{R}^n$ such that $Ax \leq b$ or
- there exists a vector $y \in \mathbb{R}^m$ such that $y \geq 0, y^\top A = 0$ and $y^\top b < 0$

# 6  Determinant

> **Determinant**          **Definition**
>
> The determinant can be understood as a number that corresponds to how much the associated linear transformation inflates space, it corresponds to the area of the unit cube, which can be scaled by a matrix operation.

**2 × 2 matrices** Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, then $\det(A) = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$.

**$n \times n$ matrices** Can be solved using the Triangle rule, Cramer's Rule, Co-Factors or LU-Decomposition ($\det(A) = \det(LU) = \det(L) \cdot \det(U)$, whereas $L$ and $U$ are triangle matrices, so $\det(U)$ is just the product of all entries on the main diagonal, see properties below)

> **Determinant**          **Properties**
>
> (I)   **P 6.9:** Matrix $T \in \mathbb{R}^{n \times n}$ is triangular, $\det(T) = \prod_{k=1}^{n} T_{kk}$, in particular, $\det(I) = 1$
>
> (II)   **T 6.10:** Matrix $A \in \mathbb{R}^{n \times n}$ then $\det(A) = \det(A^\top)$
>
> (III)   **P 6.11:** Matrix $Q \in \mathbb{R}^{n \times n}$ orthogonal $\Leftrightarrow \det(Q) = 1$ or $\det(Q) = -1$
>
> (IV)   **P 6.12:** Matrix $A \in \mathbb{R}^{n \times n}$ invertible $\Leftrightarrow \det(A) \neq 0$
>
> (V)   **P 6.13:** Matrices $A, B \in \mathbb{R}^{n \times n}$, then $\det(AB) = \det(A) \cdot \det(B)$, in particular, $\det(A^n) = \det(A)^n$
>
> (VI)   **P 6.14:** Matrix $A \in \mathbb{R}^{n \times n}$, then $\det(A^{-1}) = \dfrac{1}{\det(A)}$
>
> (VII)   $\det(2A) = 2^n \det(A)$

**Ex 6.15:** **Triangle rule**: Number of permutations: $n!$. Use diagonals to calculate:

$$\det(A) = \begin{bmatrix} A_{11} & & \\ & A_{22} & \\ & & A_{33} \end{bmatrix} + \begin{bmatrix} & & A_{13} \\ A_{21} & & \\ & A_{23} & \end{bmatrix} + \begin{bmatrix} & A_{12} & \\ & & A_{23} \\ A_{31} & & \end{bmatrix} \text{ and } A = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix}$$

and the same the oposite direction (instead of top left to bottom right, top right to bottom left) and subtract that. **Don't forget brackets**. Written out, we have

$$\det(A) = A_{11}A_{22}A_{33} + A_{13}A_{21}A_{23} + A_{12}A_{23}A_{31} - (A_{13}A_{22}A_{31} + A_{12}A_{21}A_{33} + A_{11}A_{23}A_{32})$$

**D 6.16:** **Co-Factors**: $C_{ij} = (-1)^{i+j} \det(\mathscr{A}_{ij})$, where $\mathscr{A}$ is the $(n-1) \times (n-1)$ matrix that is obtained when removing the $i$-th row and $j$-th column from $A$. Or more simple: In a $3 \times 3$ matrix, the co-factor are the entries on its first row, $A_{ij}$ the matrix obtained from $A$ without the first row and column $j$ (i.e. the one on which the current co-factor is) **P 6.17:** Using this, $\det(A) = \sum_{j=1}^{n} A_{ij} C_{ij}$.

**P 6.18:** What we had found as the inverse for $A \in \mathbb{R}^{2 \times 2}$ (see 3.1) has a form in $\mathbb{R}^{n \times n}$, where $A^{-1} = \frac{1}{\det(A)} C^\top$, where $C$ is an $n \times n$ matrix containing the co-factors of $A$ as entries. This can be rewritten as $AC^\top = \det(A)I$;

**P 6.20:** **Cramer's Rule**: The idea here is that we solve a linear system of type $Ax = b$, then using that the detminant is multiplicative, we get $\det(A)x_1 = \det(\mathscr{B}_1)$, where $\mathscr{B}$ is the matrix obtained from $A$ by replacing the first column with $b$. So, the solution $x \in \mathbb{R}^n$ for $Ax = b$ is given by $x_j = \dfrac{\det(\mathscr{B}_j)}{\det(A)}$

# 7 Eigenvalues and Eigenvectors

## 7.0 Complex numbers

**Operations**: $i^2 = -1$ (NOT $i = \sqrt{-1}$ bc. otherwise $1 = -1$). Complex number $z_j = a_j + b_j i$. *Addition, Subtraction* $(a_1 \pm a_2) + (b_1 \pm b_2)i$. *Multiplication* $(a_1 a_2 - b_1 b_2) + (a_1 b_2 + a_2 b_1)i$. *Division* $\dfrac{a_1 b_1 + a_2 b_2}{b_1^2 + b_2^2} + \dfrac{a_2 b_1 - a_1 b_2}{b_1^2 b_2^2} i$;

**Parts**: $\Re(a + bi) := a$ (Real part), $\Im(a + bi) := b$ (imaginary part), $|z| := \sqrt{a^2 + b^2}$ (modulus), $\overline{a + bi} := a - bi$ (complex conjugate); **F 7.3: Polar coordinates**: $a + bi$ (normal form), $r \cdot e^{i\phi}$ (polar form). Transformation polar $\to$ normal: $r \cdot \cos(\phi) + r \cdot \sin(\phi)i$. Transformation normal $\to$ polar: $|z| \cdot e^{i \cdot \arcsin(\frac{b}{|z|})}$; **Conjugate Transpose**: $A^* = \overline{A}^\top$ (also knows as *hermitian transpose*);

**T 7.4: Fundamental Theorem of Algebra**: Any polynomial of degree $n \geq 1$ has a zero $\lambda \in \mathbb{C}$ s.t. $P(\lambda) = 0$, precisely, $n$ zeros (=roots). The number of times a certain $\lambda$ appears is called the *algebraic multiplicity.*

**Complex valued matrices and vectors**: Works very similar to with real numbers, but the transpose has to be replaced by the conjugate transpose.

## 7.1 Eigenvalues & Eigenvectors

> **Eigenvalues & Eigenvectors**                                            **Definition 7.1**
>
> We call $\lambda \in \mathbb{C}$ an eigenvalue of $A \in \mathbb{R}^{n \times n}$ and $v \in \mathbb{C}^n \backslash \{0\}$ an eigenvector of $A$ if $Av = \lambda v$. The two are then called an eigenvalue-eigenvector pair and if $\lambda \in \mathbb{R}$, then it is a real eigenvalue and we have a real eigenvalue-eigenvector pair. $\lambda_1^2, \dots, \lambda_n^2$ eigenvalues of $A^2$ for $\lambda_1, \dots, \lambda_n$ eigenvalues of $A$

To find an Eigenvalue and Eigenvector of a matrix $M \in \mathbb{R}^{n \times n}$, simply calculate the eigenvalues first, using the zeros of the polynomial obtained from calculating $\det(M - \lambda I)$, which is obtained from **P 7.2:** $\det(M - \lambda I) = 0$. This means, we simply need to calculate the determinant of $M - \lambda I$, which is fairly straight forward. We can then try to find eigenvectors $v$ such that $Mv = \lambda v$, or in other words a non-zero element of $N(M - \lambda I) \backslash \{0\}$, i.e. the null space of $M - \lambda I$. This means we try to find a solution such that $0 = (M - \lambda I)v$, where $v$ is not the zero vector.

**P 7.3:** Characteristic polynomial: $(-1)^n \det(A - zI) = \det(zI - A) = (z - \lambda_1)(z - \lambda_2) \dots (z - \lambda_n)$. The coefficient of the $\lambda^n$ term is $(-1)^n$. Usually determined from $\det(M - \lambda I)$. **T 7.4:** From this, the fundamental theorem of algebra implies that every matrix has an eigenvalue $\lambda \in \mathbb{C}$.

> **Eigenvalues & Eigenvectors**                                                    **Properties**
>
> **(I)** **P 7.7:** $\lambda$, $v$ eigenvalue-eigenvector pair of matrix $A$, then $\lambda^k$ and $v$ are eigenvalue-eigenvector pair of matrix $A^k$ for $k \geq 1$
>
> **(II)** **P 7.8:** $\lambda$, $v$ eigenvalue-eigenvector pair of matrix $A$, then $\frac{1}{\lambda}$ and $v$ are eigenvalue-eigenvector pair of matrix $A^{-1}$
>
> **(III)** **P 7.9:** If $\lambda_1, \dots, \lambda_k$ are all distinct, the corresponding eigenvectors $v_1, \dots, v_k$ are all linearly independent.
>
> **(IV)** **T 7.10:** If matrix $A \in \mathbb{R}^{n \times n}$ has $n$ distinct eigenvalues, then there is a basis of $\mathbb{R}^n$, $v_1, \dots, v_n$ made up of eigenvectors of $A$.
>
> **(V)** **P 7.11:** The eigenvalues of $A$ are the same as for $A^\top$ (but not the eigenvectors)
>
> **(VI)** **D 7.12:** The trace (see 2) **P 7.13:** $\mathrm{Tr}(A) = \sum_{i=1}^n \lambda_i$ and $\det(A) = \prod_{i=1}^n \lambda_i$, when the eigenvalues $\lambda_1, \dots, \lambda_n$ are the $n$ eigenvalues of $A \in \mathbb{R}^{n \times n}$ as they show up in the characteristic polynomial.
>
> **(VII)** **L 7.15:** We have for matrices $A, B, C \in \mathbb{R}^{n \times n}$:
>   (i) $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$
>   (ii) $\mathrm{Tr}(ABC) = \mathrm{Tr}(BCA) = \mathrm{Tr}(CAB)$
>
> **(VIII)** **P 7.18:** If $\lambda \in \mathbb{C}$ is an eigenvalue of $Q \in \mathbb{R}^{n \times n}$ ($Q$ orthogonal), then $|\lambda| = 1$
>
> **(IX)** **D 7.21:** A matrix has a *complete set of real eigenvectors*, if we can build a basis of $\mathbb{R}^n$ from them.
>
> **(X)** **P 7.22:** A projection matrix on the subspace $U \subseteq \mathbb{R}^n$ has eigenvalues 0 and 1 and a complete set of real eigenvectors.

**Properties**

**(XI)** **D 7.23:** *geometric multiplicity* is the dimension of $N(A - \lambda I)$

**(XII)** **Ex 7.24:** For a diagonal matrix $D$, the eigenvalues of $D$ are the diagonal entries. The canonical basis $e_1, \ldots, e_n$ is a set of eigenvectors of $D$.

**(XIII)** $AB$ and $BA$ have the same eigenvalues (follows from **P7.1.12** and **L7.1.14**)

## 7.2   Change of basis

What impress uses to project the original coordinate system onto the camera's coordinate system. We have a linear transformation $L : \mathbb{R}^n \to \mathbb{R}^m$ (given by e.g. $x \in \mathbb{R}^n \to Ax \in \mathbb{R}^m$), for which we want to find a matrix $B$ that maps it from a basis $u_1, \ldots, u_n$ to another one $v_1, \ldots, v_n$. Now that $B$ helps us map a vector $\alpha$ to a vector $\beta$, which has the different basis. We now define $U$ as the matrix whose columns are the first basis and $V$ as the matrix whose columns are the second basis. Now, if $L(x) = V\beta$ and $x = U\alpha$, so $\beta = V^{-1}AU\alpha$, now $B = V^{-1}AU$.

## 7.3   Diagonalization

**T 7.1:** $A = V\Lambda V^{-1}$, where $V$'s columns are its eigenvectors and $\Lambda$ is a diagonal matrix with $\Lambda_{ii} = \lambda_i$ and all other entries 0. $A \in \mathbb{R}^{n \times n}$ and has to have a complet set of real eigenvectors.

**D 7.2:** **Diagonalizable matrix**: A matrix is diagonalizable if there exists an invertible matrix $V$ such that $V^{-1}AV = \Lambda$ where $\Lambda$ is a diagonal matrix.

**D 7.3:** **Similar matrix**: Two matrices are similar if there exists and invertible matrix $S$ such that $B = S^{-1}AS$.

**P 7.4:** Similar matrices have the same eigenvalues.

## 7.4   Spectral Theorem and symmetric matrices

**T 7.1:** **Spectral theorem**: Any symmetric matrix $A \in \mathbb{R}^{n \times n}$ has $n$ real eigenvalues and an orthonormal basis made of eigenvectors of $A$. **C 7.2:** There also exists an orthogonal matrix $V \in \mathbb{R}^{n \times n}$ (whose columns are the eigenvectors of $A$) such that $A = V\Lambda V^\top$, where $\Lambda \in \mathbb{R}^{n \times n}$ is diagonal with the diagonal entries being the eigenvalues of $A$ and $V^\top V = I$. This decomposition is called the ***eigendecomposition***. **C 7.5:** $\mathrm{rank}(A)$ is the number of non-zero eigenvalues (including/counting repetitions repetitions).

For general $n \times n$ matrices, the rank is $n - \dim(N(M))$, i.e. the geometric multiplicity of $\lambda = 0$.

**P 7.7:** $A \in \mathbb{R}^{n \times n}$, $v_1, \ldots, v_n$ orthonormal basis of of eigenvectors of $A$ and $\lambda_1, \ldots, \lambda_n$ the associated eigenvectors, then $A = \sum_{k=1}^{n} \lambda_i v_i v_i^\top$ $\lambda \in \mathbb{R}$ if $A$ is real symmetric and $\lambda$ is an eigenvalue of $A$. **C 7.9:** Every real symmetric matrix has a real eigenvalue $\lambda$;

**P 7.11:** **Rayleigh Quotient**: Let $A \in \mathbb{R}^{n \times n}$ be symmetric, then $R(x) = \dfrac{x^\top Ax}{x^\top x}$. The minimum of it is at $R(v_{\max}) = \lambda_{\max}$ and the minimum correspondingly at the smallest eigenvalue;

**D 7.12:** **Positive Semidefinite (PSD)**: all eigenvalues $\lambda_i \geq 0$   for all eigenvalues of $A \in \mathbb{R}^{n \times n}$ (symmetric), **Positive Definite (PD)**: all eigenvalues $\lambda_i > 0$   for all eigenvalues of $A \in \mathbb{R}^{n \times n}$ (symmetric);

**F 7.14:** If $A$ and $B$ PSD or PD, then also $A + B$; $A$ is PSD $\Leftrightarrow x^\top Ax \geq 0$ for all $x \in \mathbb{R}^n$;

**D 7.15:** **Gram Matrix**: $G_{ij} = v_i^\top v_j$, where $v_1, \ldots, v_n \in \mathbb{R}^m$. We have $i, j \leq n$, because $G \in \mathbb{R}^{n \times n}$. If $V \in \mathbb{R}^{m \times n}$'s columns are the $n$ vectors, then $G = V^\top V$. Abuse of notation: $AA^\top$ is also sometimes called a gram matrix of $A$. If $a_1, \ldots, a_n \in \mathbb{R}^m$ are the columns of $A$, then $AA^\top$ is an $m \times m$ matrix and $AA^\top = \sum_{i=1}^{n} a_i a_i^\top$.

**P 7.17:** The non-zero eigenvalues of $A^\top A \in \mathbb{R}^{n \times n}$ and $AA^\top \in \mathbb{R}^{m \times m}$ are the same for a matrix $A \in \mathbb{R}^{m \times n}$.

**P 7.18:** **Cholesky decomposition**: Every symmetric positive semidefinite matrix $M$ is a gram matrix of an upper triangular matrix $C$, so $M = C^\top C$ (Cholesky decomposition).

# 8  Singular Value Decomposition

## 8.1  Calculating

**D 8.1:** **SVD** *Normal form*: $A = U\Sigma V^\top$; *Compact Form*: If $A$ has rank $r$, we can write SVD as $A = U_r\Sigma_r V_r^\top$

**SVD** — **Usage**

We can use remark 8.1.3 from the lecture notes to calculate the SVD. One of the two will be beneficial if $m \neq n$ for $A$ (smaller dimension = easier determinant):

$$AA^\top = U(\Sigma\Sigma^\top)U^\top \qquad \text{and} \qquad A^\top A = V(\Sigma^\top\Sigma)V^\top$$

to more easily calculate the SVD, because the left signular vectors of $A$ (cols of $U$) are the eigenvectors of $AA^\top$ and analogously for the right hand side, the eigenvectors of $A^\top A$. The singular values are the square-root of the eigenvalues of $AA^\top$ and $A^\top A$ respectively, which then are put (sorted in ascending order) into $\Sigma$ (diagonal matrix) and if $m > n$(for $A \in \mathbb{R}^{m\times n}$, (or $n > m$ if using $A^\top A$) the not determined eigenvalues are to filled with 0
$\implies$ **Important:** Take $\sqrt{\ldots}$!

## 8.2  Show that SVD is valid

To show that a given $V$, $U$ and $\Sigma$ form an SVD, we need to show that $V$ and $U$ are orthogonal (calculate $V^\top V$ and $U^\top U$ and check if it is the identitiy), then check if $A = U\Sigma T^\top$.   **Quick & dirty:** $\Sigma$ should only contain entries $\geq 0$, due to square root, check that $U, V$ orthogonal & "normalized" (= common factors extracted) and verify dimensions of the matrices (s.t. matrix multiplication works)

**SVD of** $A^{-1}$**:** $V\Sigma^{-1}U^\top$. Requires proof: $AA^{-1} = I = A^\top A \Leftrightarrow AV\Sigma^{-1}U^\top = U\Sigma V^\top V\Sigma^{-1}U^\top = I$

**SVD** — **Properties**

1. **Theorem 8.6:** Every matrix $A \in \mathbb{R}^{m\times n}$ has an SVD

2. **D 8.2:** *Frobenius norm*: $||A||_F = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} A_{ij}^2}$

3. **D 8.2:** *Spectral / Operator norm*: $||A||_{op} = \max_{x\in\mathbb{R}^n} ||Ax||$ s.t. $||x|| = 1$

4. *Special cases*: $\sigma_1 \geq \ldots \geq \sigma_{\min\{m,n\}}$ are eigenvalues of $A \in \mathbb{R}^{m\times n}$

   (a) $||A||_F^2 = \text{Tr}(A^\top A)$

   (b) $||A||_F^2 = \sum_{i=1}^{\min\{m,n\}} \sigma_i^2$

   (c) $||A||_{op} = \sigma_1$

   (d) $||A||_{op} \leq ||A||_F \leq \sqrt{\min\{m,n\}}||A||_{op}$



When the system of linear equations has the same number of variables as the number of equations

PERFECTLY BALANCED
THIS MEME WAS MADE BY
CRAMER'S RULE GANG
AS ALL THINGS SHOULD BE
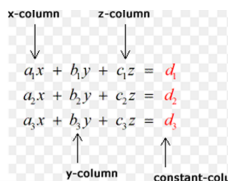
Good Luck!

# 9    Exercise tricks

$\sin, \cos, \tan$ values ($+\pi \Leftrightarrow x \cdot (-1)$) (to the right)

| °    | rad               | $\sin(\xi)$             | $\cos(\xi)$            | $\tan(\xi)$             |
|------|-------------------|------------------------|-----------------------|------------------------|
| 0°   | 0                 | 0                      | 1                     | 1                      |
| 30°  | $\frac{\pi}{6}$   | $\frac{1}{2}$          | $\frac{\sqrt{3}}{2}$  | $\frac{\sqrt{3}}{2}$   |
| 45°  | $\frac{\pi}{4}$   | $\frac{\sqrt{2}}{2}$   | $\frac{\sqrt{2}}{2}$  | 1                      |
| 60°  | $\frac{\pi}{3}$   | $\frac{\sqrt{3}}{3}$   | $\frac{1}{2}$         | $\sqrt{3}$             |
| 90°  | $\frac{\pi}{2}$   | 1                      | 0                     | $\varnothing$          |
| 120° | $\frac{2\pi}{3}$  | $\frac{\sqrt{3}}{2}$   | $-\frac{1}{2}$        | $-\sqrt{3}$            |
| 135° | $\frac{3\pi}{4}$  | $\frac{\sqrt{2}}{2}$   | $-\frac{\sqrt{2}}{2}$ | $-1$                   |
| 150° | $\frac{5\pi}{6}$  | $\frac{1}{2}$          | $-\frac{\sqrt{3}}{2}$ | $-\frac{\sqrt{3}}{2}$  |
| 180° | $\pi$             | 0                      | $-1$                  | 0                      |

**IMPORTANT:** For the multiple-choice exercises, try to come up with a short proof for each option

Non-Trivial null-space $\Rightarrow \operatorname{rank}(A) < n$ if $A \in \mathbb{R}^{n \times n}$

$A \in \mathbb{R}^{2 \times 2}$ for which $A^\top = -A \Leftrightarrow \begin{bmatrix} 0 & a \\ -a & 0 \end{bmatrix}$

## 9.1    Vector product

$a \times b = \begin{pmatrix} a_x \\ a_y \\ a_z \end{pmatrix} \times \begin{pmatrix} b_x \\ b_y \\ b_z \end{pmatrix} = \begin{pmatrix} a_y b_z - a_z b_y \\ a_z b_x - a_x b_z \\ a_x b_y - a_y b_x \end{pmatrix} = c$ gives us a vector $c$ for which $c \perp a$ and $c \perp b$

## 9.2    Basics

**Squares of numbers:** $14 : 196, 15 : 225, 16 : 256, 17 : 289, 18 : 324, 19 : 361, 21 : 441, 22 : 484, 23 : 529, 24 : 576$

**Long multiplication:** $a \cdot b$, for $n = \operatorname{len}(a), m = \operatorname{len}(b)$ we have $\sum_{i=0}^{n-1} \left( \sum_{j=0}^{m-1} a[i] \cdot b[j] * 10^{m-1-j} \right) \cdot 10^{n-1-i}$

## 9.3    Proof patterns

**(I)   Composition of Implication:** If $S \Rightarrow T$ and $T \Rightarrow U$ are both true, then $S \Rightarrow U$ is true.

**(II)   Direct proof of Implication:** Prove $S \Rightarrow T$ assuming $S$ and then proving $S$ under that assumption.

**(III)   Indirect proof of Implication:** Prove $S \Rightarrow T$ by assuming $\neg T$ and proving $\neg S$ under that assumption.

**(IV)   Case distinction:** Prove $S$ by finding a list of $R_1, \ldots, R_n$ (cases) proving at least one $R_i$, then showing that $R_i \Rightarrow S$ for all $R_i$.

**(V)   Proof by contradiction** : Prove $S$ by assuming it to be false, deriving statements from it until reaching a contradiction.

**(VI)   Existence proof:** Prove $S$ is true for at least one value

**(VII)   Proof by Induction** : Prove $P(0)$ (base case), then prove for any $k$ that $P(k) \rightarrow P(k+1)$ is true (Induction step). Using an induction hypothesis can be helpful

## 9.4    Proving bijection

We need to prove surjectivity and injectivity separately.

**Surjectivity** Given a function $f : X \rightarrow Y$, it is surjective, iff $\forall y \in Y, \exists x \in X : f(x) = y$ (continuous function)

**Injectivity** $x_1 \neq x_2 \Rightarrow f(x_1) \neq f(x_2)$

## 9.5    Subspace of matrix vector space

**Example 9.1:** $U = \{ A \in \mathbb{R}^{2 \times 2} : \operatorname{Tr}(A) = 0 \}$. Prove $\dim(U) = 3$. $U \subseteq \mathbb{R}^{2 \times 2}$:

Claim that the standard basis of $\mathbb{R}^{2 \times 2}$ form a basis of $U$, thus implying $\dim(U) = 3$. These are $B_1 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, B_2 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$ and $B_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$. Prove that they are a basis by proving that they span $U$.

## 9.6    Proof of symmetry

**Example 9.1:** $A \in \mathbb{R}^{n \times n}$ satisfying $AA^\top = I$ and $A^2 = I$. Prove that $A$ is symmetric.

$A = AI = A(AA^\top) = (AA)A^\top = (A^2)A^\top = IA^\top = A^\top \Rightarrow A$ is symmetric

## 9.7    SVD

**Example 9.1:** $A \in \mathbb{R}^{n \times n}$, $A$ invertible. $\sigma_1$ largest SV of $A$, $\sigma_1'$ largest SV of $A^{-1}$. Prove that $\sigma_1 \sigma_1' \geq 1$

Use SVD of $A^{-1}$. From this we know that SV of $A^{-1}$ are given by $\frac{1}{\sigma_1} \leq \ldots \leq \frac{1}{\sigma_n}$ (bc. $\Sigma$ is diagonal). Thus, $\sigma_1' = \frac{1}{\sigma_n}$ (largest SV) $\Rightarrow \sigma_1 \cdot \sigma_1' = \frac{\sigma_1}{\sigma_n} \geq 1$ since $\sigma_1 \geq \sigma_n > 0$

## 9.8    Least squares

Least squares is much more versatile than it might seem. It can even be used for optimization problems in multiple variables or for finding coefficients of quadratic, etc equations. Let's take $ax^2 + b$ as an example. We want to find $a$ and $b$. The $A$ matrix is then simply 1s on the first column and $x_1^2, \ldots$ on the second column. Insert the values for $x_1, \ldots$ and compute least squares

## 9.9    Finding values for which a matrix is an inverse for another

We can interpret a matrix $A^{-1} = \begin{bmatrix} | & | & | \\ x_1 & x_2 & x_3 \\ | & | & | \end{bmatrix}$, then solve SLEs, using $AA^{-1} = I$, whereas $I = \begin{bmatrix} | & | & | \\ e_1 & e_2 & e_3 \\ | & | & | \end{bmatrix}$, where $e_1$ is a standard basis vector. Thus, we get $Ax_1 = e_1$, $Ax_2 = e_2$, $\ldots$

Solving all these SLE gives us solutions for all the variables in the original matrix $A$.

## 9.10    Proving commutativity of matrix operation

For some matrices, the matrix product is commutative. To prove that, prove that both matrices have linearly independent columns, using the statements from the task and the proof of the first matrix' linear independence. Then finally, show commutativity, e.g. if $AB = I$ and $BA = I$, by showing that $A(BA - I) = 0$

## 9.11    Dimensions of subspaces

Simply argue with the size of the basis. Thus: Find basis, then argue that the basis specified is actually a basis (by showing that all its elements are linearly independent), then count the number of elements, which is the dimension. $U_1 \backslash U_2$ can never be a subspace, because 0 is missing!

## 9.12    Vector combination independence

*(Other vectors that also form basis)* Given a basis of a vector space, we have $n$ new vectors, formed form a basis. To decide if the new set forms a basis, try to construct original vectors from the new ones, or to disprove, show that 0 vector is a linear combination of the vectors with non-zero coefficients.

## 9.13    CR-Decomposition

Perform row sign-inversion only at the very end, as it can lead to nastiness along the way. $R$ is in RREF, $C$ the columns with pivot in $R$

## 9.14    Eigenvalues

For triangle and diagonal matrices, the eigenvalues are on the diagonals. Matrices of all 0's are positive semi-definite.

For exercises with a complete set of distinct eigenvalues, we can use the $A = V\Lambda V^{-1}$ decomposition, which for $A^n$ simplifies to $V \cdot \Lambda^n V^{-1}$ and then compute $\Lambda^n$, which is simple because $\Lambda$ is a diagonal matrix (so all entries on diagonal $^n$) **Alternate approach**: $\det(A) = \prod_{i=1}^n \lambda_i$ and $\det(A^n) = \det(A)^n$, then check all determinants

## 9.15    SVD

The pseudo-inverse using the SVD uses the concepts of the SVD with CR-Decomposition. $A = U_r \Sigma_r V_r^\top$, where $U_r = C$ and $\Sigma_r V_r^\top = R$

## 9.16    Quadratic solution formula

$\dfrac{\text{-}b \pm \sqrt{b^2 - 4ac}}{2a}$, for $f(x) = ax^2 + bx + c$, result $\mathbb{R}$ iff $b^2 - 4ac \geq 0$. Not that it is strictly needed, but there was still some space. There's an excellent song to remember it: `https://www.youtube.com/watch?v=E2eVZFy9yzk`

## 9.17    Multiplying out with transposes

For vectors, we have $(v - u)^\top (v - u) = v^\top v - 2v^\top u + u^\top u$

## 9.18    Number of Solutions of SLE

System $Ax = b$ has two characteristic numbers, $m$ (number of equations), $n$ (number of variables). For solution space, we have a third one, $r = \text{rank}(A)$. $A$ in REF doesn't guarantee solution. $A$ in RREF does (in $\mathbb{R}$)!

**Number of solutions** :

| $R_0$ | $r = n$ | $r < n$ |
|---|---|---|
| $r = m$ | 1 | $\infty$ |
| $r < m$ | 0 / 1 | 0 / $\infty$ |

## 9.19    Farka's Lemma

This task is from Assignment 10. Let $P_1 \subseteq \mathbb{R}^2$, $P^2 \subseteq \mathbb{R}^2$ in the plane. Assume $P_1 \cap P_2 = \emptyset$. We want to prove that there exists a vector $v \in \mathbb{R}^2$ and a scalar $w$ such that $P_1 \subseteq \{x \in \mathbb{R}^2 : x \cdot v \leq w\}$ and $P_2 \subseteq \{x \in \mathbb{R}^2 : x \cdot v > w\}$. Intuition: We can draw a line to separate the two polyhedra.

The approach to solve this now, is to assume $P_1 = \{x \in \mathbb{R}^2 : A_1 x \leq b_1\}$ and $P_2 = \{x \in \mathbb{R}^2 : A_2 x \leq b_2\}$ for some $A_1, A_2 \in \mathbb{Q}^{m \times 2}$ and $b_1, b_2 \in \mathbb{Q}^m$ and $m \in \mathbb{N}$. We observe that the system $\begin{bmatrix} A_1 \\ A_2 \end{bmatrix} x \leq \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$ has no solution since $P_1 \cap P_2 = \emptyset$. Farka's lemma implies the existence of a vector $y \in \mathbb{R}^{2m}$ with $y \geq 0$, $y^\top \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} = 0$ and $y^\top \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} < 0$.

## CHECK FOR SMALL ERRORS!

These errors could be something like a missing minus. Verify all solutions using computations.

**Example 1:** For Eigenvalues check that $\det(A - \lambda I) = 0$ for all eigenvalues.

**Example 2:** For Eigenvectors check that $Av = \lambda v$ for all eigenvalue-eigenvector pairs.