1 Vectors and Matrices

1.1 Vectors

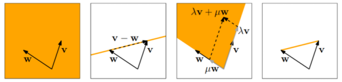
1.1.1 Linear Combination $\lambda\mathbf{v} + \mu\mathbf{w}$

$\mathbf{v}, \mathbf{w} \in \mathbb{R}^m$, $\lambda, \mu \in \mathbb{R}$

$\sum_{i=1}^{n} \lambda_i \mathbf{v}_i$, e.g. $5\begin{pmatrix}2\\3\end{pmatrix} - 3\begin{pmatrix}3\\-1\end{pmatrix} = \begin{pmatrix}10\\15\end{pmatrix} - \begin{pmatrix}9\\-3\end{pmatrix} = \begin{pmatrix}1\\18\end{pmatrix}$

Every vector in $\mathbb{R}^m$ can be written as $\sum_{i=1}^{m} u_i \mathbf{e}_i$.

(i) **affine**: if $\lambda_1 + \lambda_2 + ... + \lambda_n = 1$
(ii) **conic**: if $\lambda_j \geq 0$ for $j = 1, 2, ..., n$
(iii) **convex**: if both *affine* and *conic* combination



(i) linear (ii) affine (iii) conic (iv) convex

1.1.2 Lengths and Angles from Dot Products

Scalar (or dot, inner) product : $\mathbf{v} \cdot \mathbf{w} = v_1 w_1 + \cdots + v_n w_n$.

Outer product $rank(A) = 1 \iff \exists$ non-zero vectors $\mathbf{v} \in \mathbb{R}^m, \mathbf{w} \in \mathbb{R}^n$, s.t. A is an outer product, i.e. $A = \mathbf{v}\mathbf{w}^\top$, thus $rank(\mathbf{v}\mathbf{w}^\top) = 1$.

Length (Euclidian Norm): $||\mathbf{v}|| = \sqrt{\mathbf{v} \cdot \mathbf{v}} = \sqrt{\mathbf{v}^\top \mathbf{v}} = \sqrt{v_1^2 + \cdots + v_n^2}$.

$||\mathbf{v}||^2 := \mathbf{v}^\top \mathbf{v}$

Unit vector : $||\mathbf{u}|| = 1 = \frac{\mathbf{v}}{||\mathbf{v}||}$, for $\mathbf{v} \neq \mathbf{0}$.

Orthogonal (perpendicular) vectors: $\mathbf{v} \cdot \mathbf{w} = 0$.

Angle between vectors (Cosine Formula): $\cos\alpha = \frac{\mathbf{v} \cdot \mathbf{w}}{||\mathbf{v}||\,||\mathbf{w}||}$, for $\mathbf{v} \neq \mathbf{0}$.

Because $\cos\alpha \leq 1$, **Cauchy-Schwarz** inequality:

$$\underbrace{|\mathbf{v} \cdot \mathbf{w}|}_{|\cos\alpha|\,||\mathbf{v}||\,||\mathbf{w}||} \leq ||\mathbf{v}||\,||\mathbf{w}||.$$

Triangle inequality : $||\mathbf{v} + \mathbf{w}|| \leq ||\mathbf{v}|| + ||\mathbf{w}||$.

1.1.3 Linear (In)dependence of Vectors

$\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_k$ are linearly **independent** if:
(i) no vector is a linear combination of the previous ones. Or
(ii) no vector is a linear combination of the other ones. Or
(iii) there are no $c_1, c_2, \ldots, c_k$ besides $0, 0, \ldots, 0$ such that $c_1\mathbf{w}_1 + c_2\mathbf{w}_2 + \cdots + c_k\mathbf{w}_k = \mathbf{0}$.

All are equivalent: $(i) \Leftrightarrow (ii) \Leftrightarrow (iii)$.
Replace no by <u>**some**</u> for linear **dependence**.

A has **independent columns** if (iii) there is no $\mathbf{x}$ besides $\mathbf{0}$ such that $A\mathbf{x} = \mathbf{0}$.

1.1.4 Span

Span of vectors is the set of all linear combinations of those vectors.

1.1.5 Hyperplane through origin

Let $\mathbf{d} \in \mathbb{R}^m$, $\mathbf{d} \neq 0$, $H_\mathbf{d} = \{\mathbf{v} \in \mathbb{R}^m : \mathbf{v} \cdot \mathbf{d} = 0\}$

1.2 Matrices Basics

$A \in \mathbb{R}^{m \times n}$: a **matrix** A with m rows, n columns. *Zeilen zuerst, Spalten später*

$A = [a_{ij}]_{i=1,\ j=1}^{m\ \ \ n}$

Trace : Sum of the diagonal entries.

rank(A) = # independent columns (rows).

1.2.1 Matrix addition, scalar multiplication $A + B, \lambda A$

$$\begin{bmatrix}1 & 2\\3 & 4\end{bmatrix} + \begin{bmatrix}5 & 6\\7 & 8\end{bmatrix} = \begin{bmatrix}6 & 8\\10 & 12\end{bmatrix}, \qquad 2\begin{bmatrix}1 & 2\\3 & 4\end{bmatrix} = \begin{bmatrix}2 & 4\\6 & 8\end{bmatrix}$$

1.2.2 Matrix-Vector Multiplication

$$\underbrace{\begin{bmatrix}a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ \vdots & \vdots & \ddots & \vdots\\ a_{m1} & a_{m2} & \cdots & a_{mn}\end{bmatrix}}_{\begin{bmatrix}— & \mathbf{u}_1 & —\\ & \vdots & \\ — & \mathbf{u}_m & —\end{bmatrix}} \begin{pmatrix}x_1\\x_2\\ \vdots\\x_n\end{pmatrix} = \underbrace{\begin{pmatrix}a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n\\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n\\ \vdots\\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n\end{pmatrix}}_{\begin{pmatrix}\mathbf{u}_1 \cdot \mathbf{x}\\ \vdots\\ \mathbf{u}_m \cdot \mathbf{x}\end{pmatrix}}$$

$$\begin{bmatrix}| & & |\\ \mathbf{v}_1 & \cdots & \mathbf{v}_n\\ | & & |\end{bmatrix}\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix} = \underbrace{x_1\mathbf{v}_1 + \cdots + x_n\mathbf{v}_n}_{combination} = \sum_{i=1}^{n} x_i \mathbf{v}_i$$

1.3 Matrix Multiplication AB

If $A \in \mathbb{R}^{m \times k}$ and $B \in \mathbb{R}^{k \times n}$, then $AB \in \mathbb{R}^{m \times n}$.
BA: Square matrices: usually, $BA \neq AB$ (matrix multiplication is **not commutative**). General matrices: BA can be undefined (if $m \neq n$), or of different size than AB.

$$\begin{bmatrix}— & \mathbf{u}_1 & —\\ & \vdots & \\ — & \mathbf{u}_m & —\end{bmatrix}\begin{bmatrix}| & & |\\ \mathbf{v}_1 & \cdots & \mathbf{v}_n\\ | & & |\end{bmatrix} = \begin{bmatrix}\mathbf{u}_1\mathbf{v}_1 & \mathbf{u}_1\mathbf{v}_2 & \cdots & \mathbf{u}_1\mathbf{v}_n\\ \mathbf{u}_2\mathbf{v}_1 & \mathbf{u}_2\mathbf{v}_2 & \cdots & \mathbf{u}_2\mathbf{v}_n\\ \vdots & \vdots & \ddots & \vdots\\ \mathbf{u}_m\mathbf{v}_1 & \mathbf{u}_m\mathbf{v}_2 & \cdots & \mathbf{u}_m\mathbf{v}_n\end{bmatrix}$$

$$\begin{bmatrix}— & \mathbf{u}_1 B & —\\ & \vdots & \\ — & \mathbf{u}_m B & —\end{bmatrix} = \begin{bmatrix}| & & |\\ A\mathbf{v}_1 & \cdots & A\mathbf{v}_n\\ | & & |\end{bmatrix}$$

1.3.1 Distributivity and associativity

$A(B + C) = AB + AC$, $(B + C)D = BD + CD$. And $(AB)C = A(BC) = ABC$, $(AB)(CD) = A((BC)D) = \cdots = ABCD$. Matrix multiplication isn't commutative.

1.4 The Transpose of A

$$\begin{bmatrix}a & b\\c & d\\e & f\end{bmatrix}^\top = \begin{bmatrix}a & c & e\\b & d & f\end{bmatrix}$$

Row i of A = column i of $A^\top$, column j of A = row j of $A^\top$.

Product: $(AB)^\top = B^\top A^\top$.
More matrices: $(ABC)^\top = C^\top B^\top A^\top$.

Inverse: $(A^{-1})^\top = (A^\top)^{-1}$.
Proof: $(A^{-1})^\top A^\top = (AA^{-1})^\top = I^\top = I$.

Permutation: $P^\top = P^{-1}$.
Proof: $\mathbf{p}_i \cdot \mathbf{p}_j = I_{ij} \Leftrightarrow PP^\top = I$.

1.5 Inverse Matrices

$M \in \mathbb{R}^{n \times n}$ (square!) is invertible if there is a matrix $M^{-1} \in \mathbb{R}^{n \times n}$ (the inverse of M) such that $MM^{-1} = M^{-1}M = I$.
There can only be **one inverse**: If $MX = YM = I$, then $X = IX = (YM)X = Y(MX) = YI = Y$.

$\mathbb{R}^{1 \times 1}$: $[x]^{-1} = [\frac{1}{x}]$ (if $x \neq 0$)

$\mathbb{R}^{2 \times 2}$: $\begin{bmatrix}a & b\\c & d\end{bmatrix}^{-1} = \dfrac{1}{ad - bc}\begin{bmatrix}d & -b\\-c & a\end{bmatrix}$ if $(ad - bc \neq 0)$

$\mathbb{R}^{n \times n}$: $[M|I] \to [I|M^{-1}]$ using Gauss-Jordan elimination (3.2.1)

1.5.1 The Inverse Theorem

$A \in \mathbb{R}^{n \times n}$ is invertible
$\Leftrightarrow \forall b \in \mathbb{R}^n$, $A\mathbf{x} = \mathbf{0}$ has a unique solution $\mathbf{x}$
$\Leftrightarrow$ The columns of A are independent.

For any $A \in \mathbb{R}^{n \times n}, B \in \mathbb{R}^{n \times n}$, if $AB = I$, then $BA = I$.
Proof: $AB = I \Leftrightarrow B = A^{-1} \Leftrightarrow BA = A^{-1}A = I$.

1.5.2 Properties of the inverse

If $A \in \mathbb{R}^{n \times n}$ and $B \in \mathbb{R}^{n \times n}$ are invertible, then $(AB)^{-1} = B^{-1}A^{-1}$.
Proof: $(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AIA^{-1} = I$.
More matrices: $(ABC)^{-1} = C^{-1}B^{-1}A^{-1}$.
$(A^{-1})^{-1} = A$
$(A^\top)^{-1} = (A^{-1})^\top$

1.6 Special Matrices

- **Identity Matrix** ($a_{ii} = 1$ for all i): I
- **Diagonal Matrix** ($a_{ij} = 0$ for all $i \neq j$)
- **Upper triangular matrix** ($a_{ij} = 0$ for all $i > j$) : U
- **Lower triangular matrix** ($a_{ij} = 0$ for all $i < j$) : L
- **Symmetric matrix** ($a_{ij} = a_{ji}$ for all i, j) : $A = A^\top$
- **Skew-symmetric matrix** ($a_{ij} = -a_{ji}$ for all i, j) : $A = -A^\top$
- **Nilpotent matrix** $A^k = 0$ for some $k \in \mathbb{N}$

1.6.1 Symmetric Matrices

A matrix $S_{n \times n}$ satisfying $S = S^\top$ is symmetric.

If S is symmetric, then S^{-1} is also symmetric (if it exists).
Proof: $S^{-1} = (S^\top)^{-1} = (S^{-1})^\top$.

$AA^\top$ (and $A^\top A$) is symmetric.
Proof: $AA^\top = (A^\top)^\top A^\top = (AA^\top)^\top$.

1.7 *CR*-Decomposition $A = CR$

$C \in \mathbb{R}^{m \times r}$: the independent columns.
$R \in \mathbb{R}^{r \times n}$: how to combine them to get all columns.
$r = rank(A)$

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

Interpretation: $\mathbf{v}_1 = 1\mathbf{c}_1 + 0\mathbf{c}_2, \ldots, \mathbf{v}_4 = 3\mathbf{c}_1 + -2\mathbf{c}_2$.
Computation: Gauss-Jordan elimination (3.2.1). Get R from RREF. C is the columns from A where there is a pivot in R.

2 Linear Transformations

2.1 Injective, surjective, bijective

Let X, Y be sets and $f : X \to Y$ a function.

1. f is **injective** if for every $y \in Y$, there is at most one $x \in X$ with $f(x) = y$.
   "For every possible output, at most one input leads to it."
2. f is **surjective** if for every $y \in Y$, there is at least one $x \in X$ with $f(x) = y$.
   "For every possible output, at least one input leads to it."
3. f is **bijective** (undoable) if f is both injective and surjective.
   "For every possible output, exactly one input leads to it."
4. The **inverse** of a bijective function f is $f^{-1} : Y \to X$, $y \mapsto$ the unique $x \in X$ s.t. $f(x) = y$.

$f^{-1} \circ f = id$, $(f^{-1})^{-1} = f$

2.2 Linear Transformations

$$A \in \mathbb{R}^{m \times n} : \mathbb{R}^n \longrightarrow \mathbb{R}^m$$
$$\mathbf{x} \longmapsto A\mathbf{x}.$$

2.2.1 Definition of Linear Transformations

Given two vector spaces U and V, a Linear Transformation is a function $T : U \to V$ such that, for all $\mathbf{v}, \mathbf{w} \in U$ and $\lambda \in \mathbb{R}$ we have

$$(i)\ T(\mathbf{v} + \mathbf{w}) = T(\mathbf{v}) + T(\mathbf{w}), \qquad (ii)\ T(\lambda\mathbf{v}) = \lambda T(\mathbf{v}).$$

2.2.2 Facts about Linear Transformations

Fact 1. Let $T : U \to V$ be a linear transformation and $k \in \mathbb{N}$. For all $\mathbf{u}_1, \dots, \mathbf{u}_k \in U$ and $\alpha_1, \dots, \alpha_k \in \mathbb{R}$ we have

$$T(\alpha_1 \mathbf{u}_1 + \cdots + \alpha_k \mathbf{u}_k) = \alpha T(\mathbf{u}_1) + \cdots + \alpha T(\mathbf{u}_k).$$

Fact 2. The value of T in a basis of U fully determines T. Let $T : U \to V$ and $L : U \to V$ be two linear transformations that take the same value in a basis $\mathbf{u}_1, \dots, \mathbf{u}_n$ of U. Then $T = L$. *Proof:*

$$\begin{aligned} T(\mathbf{u}) &= T(\alpha_1 \mathbf{u}_1 + \cdots + \alpha_n \mathbf{u}_n) \\ &= \alpha_1 T(\mathbf{u}_1) + \cdots + \alpha_n T(\mathbf{u}_n) \\ &= \alpha_1 L(\mathbf{u}_1) + \cdots + \alpha_n L(\mathbf{u}_n) \\ &= L(\alpha_1 \mathbf{u}_1 + \cdots + \alpha_n \mathbf{u}_n) = L(\mathbf{u}) \end{aligned}$$

Fact 3. Given a basis $\mathbf{u}_1, \dots, \mathbf{u}_n$ of U and any $\mathbf{v}_1, \dots, \mathbf{v}_n \in V$, there is a Linear Transformation $T : U \to V$ such that, for all $1 \le i \le n$, $T(\mathbf{u}_i) = \mathbf{v}_i$.

Fact 4. For any Linear Transformation $T : \mathbb{R}^n \to \mathbb{R}^m$, there exists $A \in \mathbb{R}^{m \times n}$ such that $T(\mathbf{x}) = A\mathbf{x}$ for all $\mathbf{x} \in \mathbb{R}^n$.
Constructing A: If $\mathbf{e}_1, \dots, \mathbf{e}_n$ is the canonical basis of $\mathbb{R}^n$,

$$A = \begin{bmatrix} | & & | \\ T(\mathbf{e}_1) & \cdots & T(\mathbf{e}_n) \\ | & & | \end{bmatrix}.$$

Fact 5. Given $T : \mathbb{R}^n \to \mathbb{R}^m$ and $L : \mathbb{R}^m \to \mathbb{R}^p$, with corresponding matrices $A \in \mathbb{R}^{m \times n}$ and $B \in \mathbb{R}^{p \times m}$, the linear transformation $L \circ T$ given by $L \circ T(\mathbf{x}) = L(T(\mathbf{x}))$ corresponds to multiplying by BA. In other words $L \circ T(\mathbf{x}) = BA\mathbf{x}$.

2.2.3 Prove, that T is linear transformation

Use $T(x+y) = T(x) + T(y)$ and $T(\lambda x) = \lambda T(x)$. Insert the linear transformatione given by the task and replace x with $x+y$ and with λx.

2.2.4 (Counter)examples of Linear Transformations

Functions that are linear transformations:

(1) The identity map $T : \mathbb{R}^n \to \mathbb{R}^n$ given by $T(x) = x$,
(2) For any matrix A, the map $x \to Ax$,
(3) For a vector $v \in \mathbb{R}^n$ the map $T : \mathbb{R}^n \to \mathbb{R}$ given by $T(x) = v^\top x$,
(4) The map $T : \mathbb{R}^n \to \mathbb{R}^n$ given by $T(x) = 0$.

Functions that are not linear transformations:

(1) For a vector $v \in \mathbb{R}^n$ (such that $v \neq 0$) the map $T : \mathbb{R}^n \to \mathbb{R}$ given by $T(x) = v + x$,
(2) The map $T : \mathbb{R}^n \to \mathbb{R}$ given by $T(x) = \|x\|$,
(3) The map $T : \mathbb{R}^n \to \mathbb{R}$ given by $T(x) = \|x\|^2$,
(4) The map $T : \mathbb{R}^n \to \mathbb{R}^n$ given by $T(x) = \frac{1}{\|x\|}x$.

3 The 4 Fundamental Subspaces

Span, Basis	Example
V: vector space	$\mathbf{C}(A)$
S: sequence of vectors in V	the columns of A
S **spans** V: V = all combinations of S	the columns span $\mathbf{C}(A)$
S **basis** of V: S independent, S spans V	the independent columns: basis of $\mathbf{C}(A)$

3.1 Vector Spaces and Subspaces

Abstract concept of combinations that we can do with vectors: $\mathbf{v} + \mathbf{w}$, $\lambda \mathbf{v}$.
Examples: $\mathbb{R}^n$, $\mathbb{C}^n$, $\mathbb{R}^{m \times n}$, $\mathbb{R}^{\mathbb{R}}$ (functions $f : \mathbb{R} \to \mathbb{R}$), $\{0,1\}^n$ (bit vectors).
Vector space is a triple $(V, \oplus, \odot)$ where V is a set with two operations. They base on fields and satisfy axioms: *commutativity, associativity, zero vector, negative element, identity element, compatibility of multiplications of vectors and scalars ($\in \mathbb{R}$), distributivity over $\oplus$ both for vectors and scalars ($\in \mathbb{R}$).*

3.1.1 Subspaces of Vector Spaces

Let V be a vector space. **Subspace** : nonempty $U \subseteq V$ satisfying: if $\mathbf{v}, \mathbf{w} \in U$ and c is a scalar, then

$$(i)\ \mathbf{v} + \mathbf{w} \in U, \qquad (ii)\ c\mathbf{v} \in U.$$

A subspace is a vector space.

U always contains $\mathbf{0}$: Let $\mathbf{u} \in U$, then $0\mathbf{u} \in U$ by (ii).
Smallest subspace: $U = \{\mathbf{0}\}$, largest subspace: $U = V$.

$\mathbf{C}(A)$, $\mathbf{C}(A^\top)$, $\mathbf{N}(A)$ are subspaces.

3.2 Fundamental subspaces of a matrix

3.2.1 The **Column Space** of $A \in \mathbb{R}^{m \times n}$

$\mathbf{C}(A) = \{A\mathbf{x} \in \mathbb{R}^m : \mathbf{x} \in \mathbb{R}^n\}$ (subspace of $\mathbb{R}^m$)
Combinations (span) of the columns.

The columns of A span the vector space $\mathbf{C}(A)$.

$\mathbf{C}(A) = \mathbf{Im}(T) := \{T(\mathbf{x}) : \mathbf{x} \in \mathbb{R}^m\} \subseteq \mathbb{R}^m$ ($A \in \mathbb{R}^{m \times n}$, s.t. $T = T_A$)
The set of all outputs that T can produce.

3.2.2 **Row Space** $\mathbf{R}(A) := \mathbf{C}(A^\top)$

$\mathbf{R}(A) = \mathbf{R}(R_0)$
The rows of R form a basis for $\mathbf{R}(A)$.
independent rows = # independent columns!

3.2.3 **Null Space** of $A \in \mathbb{R}^{m \times n}$

$\mathbf{N}(A) = \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} = \mathbf{0}\} \subseteq \mathbb{R}^n$
$\mathbf{N}(A) = \mathbf{Ker}(T) := \{\mathbf{x} \in \mathbb{R}^n : T(\mathbf{x}) = 0\} \subseteq \mathbb{R}^n$ ($A \in \mathbb{R}^{m \times n}$, s.t. $T = T_A$)

3.2.4 Left Nullspace $\mathbf{N}(A^\top)$

Why "left"? All solutions of $A^\top \mathbf{y} = \mathbf{0}$ = all solutions of $\mathbf{y}^\top A = \mathbf{0}^\top$.

3.2.5 Dimensions of the Four Subspaces

subspace	of	definition	dimension
$\mathbf{C}(A)$	$\mathbb{R}^m$	combinations of the columns of A	$r = \mathbf{rank}(A)$
$\mathbf{R}(A) = \mathbf{C}(A^\top)$	$\mathbb{R}^n$	combinations of the rows of A = columns of $A^\top$	r
$\mathbf{N}(A)$	$\mathbb{R}^n$	solutions of $A\mathbf{x} = \mathbf{0}$	$n - r$
$\mathbf{N}(A^\top)$	$\mathbb{R}^m$	solutions of $A^\top \mathbf{y} = \mathbf{0}$	$m - r$

3.3 Bases and dimension

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

Independant columns form basis of $\mathbf{C}(A)$

Non-uniqueness of basis: Every set $B \subseteq \mathbb{R}^m$ of m linearly independent vectors is a basis of $\mathbb{R}^m$.

Finitely generated vector space : $\exists G \subseteq V$ with $Span(G) = V$. Then V has a basis $B \subseteq G$.

Finitely generated Vector Space has a basis

Steinitz exchange lemma : "exchanging elements between G and F": V is finitely generated vector space, $F \subseteq V$ a finite set of lin. independent vectors, and $G \subseteq V$ a finite set of vectors with $Span(G) = V$, then:

- $|V| \le |G|$
- $\exists E \subseteq G$ of size $|G| - |F|$, s.t. $Span(F \cup E) = V$.

All bases have the same size : $B, B' \implies |B| = |B'|$

Dimension of V ($dim(V)$) is the size of an arbitrary basis B of V

Linear transformation between vector spaces Let V, W be vector spaces. A function $T : V \to W$ is linear if, for all $x_1, x_2 \in V$ and $\lambda_1, \lambda_2 \in \mathbb{R}$, $T(\lambda_1 x_1 + \lambda_2 x_2) = \lambda_1 T(x_1) + \lambda_2 T(x_2)$.

Bijective lin. transformations preserve basis If $T : V \to W$ is a bijective linear map, then $B \subseteq V$ is a basis of $V \Leftrightarrow T(B)$ is a basis of W, and hence $\dim(V) = \dim(W)$.

Isomorphic vector spaces $V \cong W \iff \exists T : V \to W$ linear and bijective.

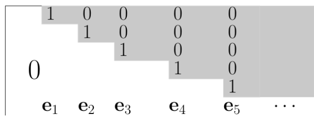
Basis writes vectors as a unique lin. combination Let V be a finite-dimensional vector space with basis $B = \{v_1, \dots, v_m\}$. Then every $v \in V$ can be written uniquely as $v = \sum_{j=1}^m \lambda_j v_j$, for unique scalars $\lambda_1, \dots, \lambda_m$.

Less than $\dim(V)$ vectors do not span V If $|G| < \dim V$, then $\mathrm{span}(G) \neq V$.

3.4 Computing the three fundamental subspaces

$\mathbf{N}(A) = \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} = \mathbf{0}\}$ is a subspace of $\mathbb{R}^n$.
"Computing" a subspace: find a basis of it! If all columns are independent: $\mathbf{N}(A) = \{\mathbf{0}\}$. Row operations don't change solutions: $A\mathbf{x} = 0 \Leftrightarrow R\mathbf{x} = 0$, $\mathbf{N}(A) = \mathbf{N}(R)$.

R is in **reduced row echelon form**:



(standard unit vectors)

How-to:

Transform A to R using Gauss-Jordan elimination.
Read a basis of $\mathbf{N}(R)$ off R.

3.4.1 Gauss-Jordan Elimination

We add those steps to Gauss elimination:

For each column

For a row with a pivot p, we multiply the row by $1/p$ to get $p = 1$. We eliminate above the pivot.

At the end ($R_0 \to R$)

Remove the bottom zero-rows

3.4.2 Bases of subspaces

The basis of $\mathbf{C}(A)$: Columns of A where Pivot columns in $RREF(A)$

The basis of $\mathbf{R}(A) = \mathbf{C}(A^\top)$: Nonzero rows of $RREF(A)$

Row rank equals columns rank : $rank(A) = rank(A^\top)$

Rank is $\le$ min of the matrix dimensions: $r \le min(n,m)$, $A \in \mathbb{R}^{m \times n}$

Nullspace isomorphism : $R = \mathrm{RREF}(A)$, then $T : N(R) \to \mathbb{R}^{n-r}$ is an isomorphism between $N(R)$ and $\mathbb{R}^{n-r} \Rightarrow \dim(N(R)) = n - r$.

Basis of $N(A)$: Non-pivot columns of $RREF(A)$

$$\begin{bmatrix} 1 & 2 & 0 & 3 \\ 0 & 0 & 1 & -2 \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \mathbf{0} \Leftrightarrow \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}}_{I} \begin{pmatrix} x_1 \\ x_3 \end{pmatrix} + \underbrace{\begin{bmatrix} 2 & 3 \\ 0 & -2 \end{bmatrix}}_{F} \begin{pmatrix} x_2 \\ x_4 \end{pmatrix} = \mathbf{0}$$

$$\Leftrightarrow \underbrace{\begin{pmatrix} x_1 \\ x_3 \end{pmatrix}}_{\mathbf{x}_I} = -F \underbrace{\begin{pmatrix} x_2 \\ x_4 \end{pmatrix}}_{\mathbf{x}_F} \qquad (1)$$

For each free variable f_i in $\mathbf{x}_F$, solve (1) with $f_i = 1$ and all other free variables $= 0$. You get the $n - r$ basis vectors:

$$\underbrace{\begin{pmatrix} x_1 \\ x_3 \end{pmatrix} = -F \begin{pmatrix} 1 \\ 0 \end{pmatrix}}_{[-2 \quad 1 \quad 0 \quad 0]^\top}, \qquad \underbrace{\begin{pmatrix} x_1 \\ x_3 \end{pmatrix} = -F \begin{pmatrix} 0 \\ 1 \end{pmatrix}}_{[-3 \quad 0 \quad 2 \quad 1]^\top}$$

3.5 The Complete Solution to $Ax = b$

Solution space of $A\mathbf{x} = \mathbf{b}$:
$\mathbf{Sol}(A, \mathbf{b}) := \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} = \mathbf{b}\} \subseteq \mathbb{R}^n$

Solution space from shifting the nullspace : Let $\mathbf{s}$ be some solution of $A\mathbf{x} = \mathbf{b}$, then:
$\mathbf{Sol}(A, \mathbf{b}) := \{\mathbf{s} + \mathbf{x} \in \mathbb{R}^n : \mathbf{x} \in N(A)\}$.
We can also compute $\mathbf{Sol}(A, \mathbf{b})$, although it is not a subspace.

Dimension of a solution space Let $A \in \mathbb{R}^{m \times n}$ with rank r. If $A\mathbf{x} = \mathbf{b}$ is solvable, then:
$\dim(\mathbf{Sol}(A, \mathbf{b})) = n - r$, and $\dim(\mathbf{Sol}(A, \mathbf{b})) := \dim(N(A))$.

Systems of rank m are solvable Let $A \in \mathbb{R}^{m \times n}$ with $\mathrm{rank}(A) = m$, $A\mathbf{x} = \mathbf{b}$ is solvable for all $\mathbf{b} \in \mathbb{R}^m$.

Systems of rank less than m are typ. unsolvable Systems of rank $r < m$ are typically unsolvable.

Types of systems Let $A \in \mathbb{R}^{m \times n}$ and $\mathbf{b} \in \mathbb{R}^m$. The system $A \in \mathbb{R}^{m \times n}$

is called:

- $m = n \Rightarrow$ square (A is a square matrix) $\star$ **typ. solvable**
- $m < n \Rightarrow$ underdetermined (A is a wide matrix) $\star$ **typ. solvable**
- $m > n \Rightarrow$ overdetermined (A is a tall matrix) $\star$ **typ. unsolvable**

"Typical" matrices are with $m \leq n$ and have rank $r = m$.
Apply row operations also to $\mathbf{b}$ ($A \to R_0, \mathbf{b} \to \mathbf{c}$).

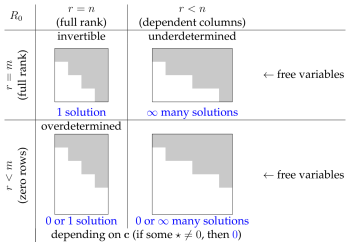
3.5.1 The set of all solutions to a SLE

Injectivity of A on $C(A^\top)$, uniqueness of sol. :
$A \in \mathbb{R}^{m \times n}, \ x, y \in C(A^\top): \ Ax = Ay \Leftrightarrow x = y$
This leads to: $C(A^T) \cap N(A) = \{0\}$

Set of all solution of linear equations
Set of all sol. : $\{x \in \mathbb{R}^n | Ax = b\} \neq \emptyset$, then:
$\{x \in \mathbb{R}^n \mid Ax = b\} = x_1 + N(A), \quad x_1 \in R(A)$ is unique s.t. $Ax_1 = b$.

Linear equations with no solution
Linear equations has no solution:
$\{x \in \mathbb{R}^n \mid Ax = b\} = \emptyset \iff \{z \in \mathbb{R}^m \mid A^T z = 0, \ b^T z = 1\} \neq \emptyset$.

3.5.2 Number of solutions of $A\mathbf{x} = \mathbf{b}$



4 Linear Equations $A\mathbf{x} = \mathbf{b}$

$$\left.\begin{array}{rcl} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &=& b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &=& b_2 \\ &\vdots& \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &=& b_m \end{array}\right\} \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

4.1 Elimination and Back Substitution

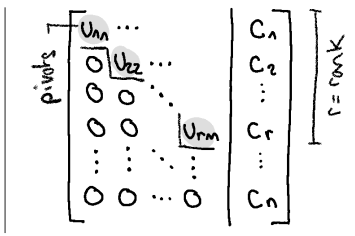
4.1.1 Back Substitution

If A upper triangular, work your way up!

$$\begin{bmatrix} a_1 & a_2 \\ & a_3 \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \Rightarrow \begin{cases} x_1 &= (b_1 - a_2 x_2)/a_1 \\ x_2 &= b_2/a_3 \end{cases} \uparrow$$

4.1.2 Gauss Elimination

Transform $Ax = b$ to $Ux = c$ with U upper triangular. Permited row operations: **exchange**, scalar **multiplication** (not 0), and **substraction**. Solving $Ux = c$ also solves $Ax = b$.



4.1.3 Elimination and Permutation Matrices

Subtract $d \cdot$(Row 1) from (Row 2):

$$E_{21} = \begin{bmatrix} 1 & 0 & 0 \\ -d & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Exchange Rows 2 and 3:

$$P_{23} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

$E_{**} \cdots E_{**} P_{**} \cdots P_{**} A = EPA = U$, $EPb = c$. The elimination matrices are in decreasing order of indices. E is lower triangular, P has unit vectors as columns.

4.2 Permutations

Permutation matrix P: reordering (*permutation*) of the rows of I. PA: permutation of the rows of A.
If P, P' are permutation matrices, then also PP': reordering twice is another reordering.
There are $n!$ permutation matrices, since n things can be ordered in $n!$ ways.

5 Orthogonality, Projections, and Least Squares

used to decompose space into subspaces

5.1 Vectors and Subspaces Orthogonality

Orthogonal subspaces : Two subspaces V and W of $\mathbb{R}^n$ are orthogonal if $\forall_{\mathbf{v} \in V, \ \mathbf{w} \in W} \ \mathbf{v} \cdot \mathbf{w} = 0$ (all vectors are orthogonal).

Orthogonality of bases : Let $\mathbf{v_1}, ..., \mathbf{v_2}$ and $\mathbf{w_1}, ..., \mathbf{w_2}$ be bases of subspaces W and V. W and V are orthogonal $\Leftrightarrow$ all $\mathbf{v_i}$ orthogonal to all $\mathbf{w_j}$

Combinations and interaction of subspaces

- The set of vectors $\{v_1, ..., v_2, w_1, ..., w_2\}$ are linearly independent.
- The union of bases of two subspaces gives a basis for the new subspace: $V \cup W = V + W = \{\lambda v + \mu w \mid \lambda, \mu \in \mathbb{R}, \ v \in V, \ w \in W\}$.
- If V and W are subspaces of $\mathbb{R}^n$, then $V + W$ is a subspace of $\mathbb{R}^n$.
- $V \cap W = \{0\}$ if subspaces are orthogonal.
- If V and W orthogonal subspaces, then $\dim(V + W) = \dim(V) + \dim(W) \leq n$.

Proof. Let $k = \dim(V), \ell = \dim(W)$, $\mathbf{v}_1, \ldots, \mathbf{v}_k$ a basis of V, $\mathbf{w}_1, \ldots, \mathbf{w}_\ell$ a basis of W. Want to show: these $k + \ell$ vectors are independent.

Suppose $\underbrace{c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k}_{\mathbf{v} \in V} + \underbrace{d_1\mathbf{w}_1 + \cdots d_\ell\mathbf{w}_\ell}_{\mathbf{w} \in W \ (\Rightarrow -\mathbf{w} \in W)} = \mathbf{0}$. Then $\mathbf{v} = -\mathbf{w} \in V \cap W$, so $\mathbf{v} = \mathbf{w} = \mathbf{0}$. $\mathbf{v}_1, \ldots, \mathbf{v}_k$ and $\mathbf{w}_1, \ldots, \mathbf{w}_\ell$ are independent $\Rightarrow c_1, \ldots, c_k = 0$ and $d_1, \ldots, d_\ell = 0 \Rightarrow$ $\mathbf{v}_1, \ldots, \mathbf{v}_k, \mathbf{w}_1, \ldots, \mathbf{w}_\ell$ are independent (1.3.3) $\Rightarrow k + \ell \leq n$.

5.1.1 Orthogonal complement $V^\perp$

If V is a subspace of $\mathbb{R}^n$, then its **orthogonal complement**: $V^\perp = \{w \in \mathbb{R}^n | w^\top v = 0 \text{ for all } v \in V\}$.

$V = (V^\perp)^\perp$

Vector decomposition by orth. complements : If V, W are orthogonal subspaces of $\mathbb{R}^n$, then $W = V^\perp \Leftrightarrow \dim(V) + \dim(W) = n \Leftrightarrow$ $\mathbf{u} = \mathbf{v} + \mathbf{w}$ for every $\mathbf{u} \in \mathbb{R}^n$ with *unique* vectors $\mathbf{v} \in V$, $\mathbf{w} \in W$.

Relations between subspaces : $N(A) = C(A^\top)^\perp = R(A)^\perp$ and $R(A) = C(A^\top) = N(A)^\perp$
Proof: $A\mathbf{v} = 0$ and $\mathbf{w} = A^\top \mathbf{x} \Rightarrow \mathbf{v}^\top \mathbf{w} = \mathbf{v}^\top(A^\top \mathbf{x}) = (\mathbf{v}^\top A^\top)\mathbf{x} = (A\mathbf{v})^\top \mathbf{x} = 0$.

$N(A) = N(A^\top A)$ and $C(A^\top) = C(A^\top A)$. This is justification to normal equations. Proof: $x \in N(A^\top A)$, then $0 = \mathbf{x}^\top 0 = \mathbf{x}^\top A^\top A\mathbf{x} = (A\mathbf{x})^\top(A\mathbf{x}) = ||A\mathbf{x}||^2$. This implies $A\mathbf{x} = 0 \implies \mathbf{x} \in N(A)$.

5.2 Projections

Projection of $\mathbf{b} \in \mathbb{R}^m$ on a subspace S (of $\mathbb{R}^m$) is the point in S that is closest to $\mathbf{b}$: $\text{proj}_S(b) = \arg\min_{\mathbf{p} \in S} ||\mathbf{b} - \mathbf{p}||$.

5.2.1 One-dimensional case

Let $\mathbf{a} \in \mathbb{R}^m \backslash \{0\}$. Projection of $\mathbf{b} \in \mathbb{R}^m$ on subspace $S = \{\lambda \mathbf{a} | \lambda \in \mathbb{R}\} = C(a)$: $\text{proj}_S(\mathbf{b}) = \frac{aa^\top}{a^\top a}\mathbf{b}$.
"Error vector" is perpendicular to projection ($(b - \text{proj}_S(b)) \perp \text{proj}_S(b)$).

5.2.2 General case

Let S be a subspace in $\mathbb{R}^m$ with a basis $a_1, ..., a_n$ that span S. Let A be the matrix with column vectors $a_1, ..., a_n$.
The general formula: $\text{proj}_S(b) = A\hat{x}$, where $\hat{x}$ is $A^T A\hat{x} = A^T b$.

Properties of $A^\top A$: $A^\top A$ is invertible $\Leftrightarrow$ A has linearly independent columns. $\Rightarrow A^\top A$ is a square matrix, symmetric, invertible.

Projection Matrix : $\text{proj}_S(b) = Pb$ with $P = A(A^\top A)^{-1} A^\top$.

As $\text{proj}_S(\text{proj}_S(b)) = \text{proj}_S(b)$: $P^2 = P$

$I - P$ *is projection matrix to* $\text{proj}_{S^\perp}(b)$, *as* $b = e + \text{proj}_S(b) = e + Pb$, *where* $e \in S^\perp$.

$(I - P)^2 = I - 2P + P^2 = I - P$

5.3 Least Squares Approximation

Goal: Approximate a solution to System of equations (if it has no solution $\mathbf{x}$: find $\mathbf{x}$ for which $A\mathbf{x}$ is as close as possible to $\mathbf{b}$: $\min_{\hat{\mathbf{x}} \in \mathbb{R}^n} \|A\hat{\mathbf{x}} - \mathbf{b}\|^2$

$$\underbrace{A^\top A\hat{\mathbf{x}} = A^\top \mathbf{b}}_{\text{Normal Equation}}.$$

Usage: find $A^\top A$ and $A^\top \mathbf{b}$ and solve system. When A has independent columns: $\hat{\mathbf{x}} = (A^\top A)^{-1} A^\top \mathbf{b}$ (unique).

For any A, $\mathbf{C}(A^\top) = \mathbf{C}(A^\top A)$.

5.3.1 Linear Regression

Application of least squares problem, in which it is to find A and $\mathbf{b}$ such that we can solve the system. We have data-points (values b_k over time t_k). We define a matrix $A = \begin{bmatrix} 1 & \mathbf{t}_1 \\ \vdots & \vdots \\ 1 & \mathbf{t}_n \end{bmatrix}$ and a result vector $\mathbf{b} = \begin{bmatrix} \mathbf{b}_1 \\ \vdots \\ \mathbf{b}_n \end{bmatrix}$ where n is the total number of data points and $\mathbf{t}_i$ is the slope of the ith function, where $\mathbf{b}_i$ is its output. The first column is all 1s because the constant element has no scalar. This comes from the following concept: $f(t) = \alpha_0 + \alpha_1$, so if the first data point is $(1, 2)$, we get $\alpha + \alpha_1 \cdot 1 = 2$, which will then transform into a SLE with other equations.

Fitting a parabola:
$(t_k, b_k) = \{(0, 1), (1, 2), (2, 5)\}$, $b_k \approx \alpha_0 + \alpha_1 t_k + \alpha_2 t_k^2$
$A = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \end{pmatrix}, b = \begin{pmatrix} 1 \\ 2 \\ 5 \end{pmatrix}, \hat{\alpha} = (A^T A)^{-1} A^T b = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \hat{b}(t) = 1 + t^2$.

Matrix A $(m \times 2)$ has linearly dependent columns $\Leftrightarrow t_i = t_j \forall i \neq j$
.

5.4 Orthonormal Bases and Gram Schmidt

5.4.1 Orthonormal vectors

Vectors are orthonormal if they are orthogonal and have norm 1.
($\mathbf{q}_i^\top \mathbf{q}_i = \delta_{ij} = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}$, with δ_{ij} being the *Kronecker delta*.)
For example the canonical basis $\mathbf{e}_1, ..., \mathbf{e}_n \in \mathbb{R}^{n \times n}$.

5.4.2 Orthogonal Matrix

A square matrix $Q \in \mathbb{R}^{n \times n}$ is an Orthogonal Matrix when $Q^\top Q = I$. In this case $QQ^\top = I$, $Q^{-1} = Q^\top$, and the columns of Q form an orthonormal basis of $\mathbb{R}^n$.

Examples: permutation matrices, the 2×2 matrix R that corresponds to rotating counterclockwise the plane by θ

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

Orthogonal matrices preserve norm and inner product of vectors. I.e., if $Q \in \mathbb{R}^{n \times n}$ is orthogonal then, for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$.

$$||Q\mathbf{x}|| = ||\mathbf{x}|| \text{ and } (Q\mathbf{x})^\top(Q\mathbf{y}) = \mathbf{x}^\top \mathbf{y}.$$

5.4.3 Projections/Least squares with Orthonormal Basis

Let S be a subspace of $\mathbb{R}^n$ and $\mathbf{q}_1, \ldots, \mathbf{q}_n$ be an orthonormal basis for S. Let $Q \in \mathbb{R}^{m \times n}$ be the matrix whose columns are the $\mathbf{q}_i$'s. The Projection Matrix that projects to S is given by $QQ^\top$ and the Least Squares solution to $Q\mathbf{x} = \mathbf{b}$ is given by $\hat{\mathbf{x}} = Q^\top \mathbf{b}$.

5.4.4 Gram-Schmidt algorithm

Given n linearly independent vectors $\mathbf{a}_1, \dots, \mathbf{a}_n$ that span a subspace S, the Gram-Schmidt process constructs an orthonormal basis $\mathbf{q}_1, \dots, \mathbf{q}_n$ the following way:

- $\mathbf{q}_1 = \frac{\mathbf{a}_1}{||\mathbf{a}_1||}$
- For $k = 2, \dots, n$ do
  - $\mathbf{q'}_k = \mathbf{a}_k - \sum_{i=1}^{k-1} (\mathbf{a}_k^\top \mathbf{q}_i)\mathbf{q}_i$
  - $\mathbf{q}_k = \frac{\mathbf{q'}_k}{||\mathbf{q'}_k||}$

5.4.5 QR decomposition

The QR decomposition of $A \in \mathbb{R}^{m\times n}$ is $A = QR$ where $Q \in \mathbb{R}^{m \times n}$ is the output of Gram-Schmidt on A and $R \in \mathbb{R}^{n\times n} = Q^\top A$ **is an upper triangular, invertible matrix** and $QQ^\top A = A$, hence the QR-Decomposition is well-defined.

Since $\mathbf{N}(A) = \mathbf{N}(R) = \{\mathbf{0}\}$, both R and $R^\top$ are invertible.

Projections

Since $\mathbf{C}(A) = \mathbf{C}(Q)$, the projections on $\mathbf{C}(A)$ can be done with Q which means they are given by $\text{proj}_{\mathbf{C}(A)}(b) = QQ^\top \mathbf{b}$.

Least Squares

$$\begin{aligned} A^\top A\hat{\mathbf{x}} = A^\top \mathbf{b} &\Rightarrow (QR)^\top (QR)\hat{\mathbf{x}} = (QR)^\top \mathbf{b} \\ &\Rightarrow R^\top Q^\top QR\hat{\mathbf{x}} = R^\top Q^\top \mathbf{b} \\ &\Rightarrow R^\top R\hat{\mathbf{x}} = R^\top Q^\top \mathbf{b} \Rightarrow R\hat{\mathbf{x}} = Q^\top \mathbf{b} \end{aligned}$$

We can use back-substitution, as R is triangular.

5.5 The Pseudoinverse $A^\dagger$

5.5.1 Full Column Rank Matrices

For $A \in \mathbb{R}^{m\times n}$ with $\text{rank}(A) = n$, $A^\dagger = (A^\top A)^{-1} A^\top$.
$A^\dagger$ is a **left inverse** of A, meaning that $A^\dagger A = I$.

5.5.2 Full Row Rank Matrices

For $A_{m\times n}$ with $\text{rank}(A) = m$, $A^\dagger = A^\top (AA^\top)^{-1}$.
$A^\dagger$ is a **right inverse** of A, meaning that $AA^\dagger = I$.

The unique solution $\hat{\mathbf{x}} \in \mathbf{C}(A^\top)$ to $\min_{\mathbf{x}\in\mathbb{R}^n} ||\mathbf{x}||^2$ such that $A\mathbf{x} = \mathbf{b}$ is given by $\hat{\mathbf{x}} = A^\dagger \mathbf{b}$.

5.5.3 All Matrices

For $A_{m\times n}$ with $\text{rank}(A) = r$ and decomposition $A = CR$,

$$\begin{aligned} A^\dagger = R^\dagger C^\dagger &= R^\top (RR^\top)^{-1} (C^\top C)^{-1} C^\top \\ &= R^\top (C^\top CRR^\top)^{-1} C^\top \\ &= R^\top (C^\top AR^\top)^{-1} C^\top. \end{aligned}$$

The unique solution $\hat{\mathbf{x}} \in \mathbf{C}(A^\top)$ to $\min_{\mathbf{x}\in\mathbb{R}^n} ||\mathbf{x}||^2$ such that $A^\top A\mathbf{x} = A^\top \mathbf{b}$ is given by $\hat{\mathbf{x}} = A^\dagger \mathbf{b}.$

We can compute the pseudoinverse from any full rank factorization, not just specifically the CR decomposition.

5.5.4 Some Pseudoinverse Facts

- $AA^\dagger A = A$ and $A^\dagger AA^\dagger = A^\dagger$ and $(A^\dagger)^\top = (A^\top)^\dagger$
- $AA^\dagger$ is symmetric and is the projection matrix on $\mathbf{C}(A)$
- $A^\dagger A$ is symmetric and is the projection matrix on $\mathbf{C}(A^\top)$
- $((AB)^\dagger = B^\dagger A^\dagger$, as long as $\text{rank}(A) = \text{rank}(B) = n$.; needs proof)

6 The Determinant $\det(A)$

6.1 2×2-matrices

$$\det(\begin{vmatrix} a & b \\ c & d \end{vmatrix}) := \det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc$$

$$\begin{vmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{vmatrix} = A_{11} \begin{vmatrix} A_{22} & A_{23} \\ A_{32} & A_{33} \end{vmatrix} - A_{12} \begin{vmatrix} A_{21} & A_{23} \\ A_{31} & A_{33} \end{vmatrix} + A_{13} \begin{vmatrix} A_{21} & A_{22} \\ A_{31} & A_{32} \end{vmatrix}$$

6.2 General case

6.2.1 Definition

Given a permutation $\sigma : \{1, \dots, n\} \to \{1, \dots, n\}$ (there exist $n!$), its sign is defined as

$$\text{sgn}(\sigma) = \begin{cases} 1 & \text{if} \quad |(i,j) \in \{1,\dots,n\} \times \{1,\dots,n\} \text{ such that } i < j \text{ and } \sigma(i) > \sigma(j)| \text{ is even,} \\ -1 & \text{if} \quad |(i,j) \in \{1,\dots,n\} \times \{1,\dots,n\} \text{ such that } i < j \text{ and } \sigma(i) > \sigma(j)| \text{ is odd.} \end{cases}$$

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

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

From this definition one can verify the following propositions.

A **is invertible** $\Leftrightarrow \det(A) \neq 0$,

$\det(A^{-1}) = \frac{1}{\det(A)}$ (if A^{-1} exists), $\det(A^\top) = \det(A)$,

$\det(AB) = \det(A)\det(B)$, $\det(A^n) = \det(A)^n$

A permutation matrix P corresponding to a permutation σ has $\det(P) = \text{sgn}(\sigma)$,

A triangular matrix T has $\det(T) = \prod_{k=1}^n T_{kk}$, $\det(I) = 1$,

an orthogonal matrix Q has $\det(Q) = 1$ or -1
Proof: $1 = \det(I) = \det(Q^\top Q) = \det(Q^\top)\det(Q) = \det(Q)^2$.
if $\det(Q) = 1$, Q is a rotation matrix. If $\det(Q) = -1$, Q is a reflection matrix.
$\det(\lambda A) = \lambda^n \det(A)$

6.2.2 Co-Factors

Given $A_{n\times n}$, for each $1 \leq i, j \leq n$ let $\mathcal{A}_{ij}$ denote the $(n-1) \times (n-1)$ matrix obtained by removing row i and column j from A. Then we define the co-factors of A as $C_{ij} = (-1)^{i+j} \det(\mathcal{A}_{ij})$.

The determinant can be written in terms of co-factors:
For any row $1 \leq i \leq n$, $\det(A) = \sum_{j=1}^n A_{ij}\mathbf{C}_{ij}$.

We can (inefficiently) compute the inverse using co-factors: $A^{-1} = \frac{1}{\det(A)} C^\top$, where $C \in \mathbb{R}^{n\times n}$ is the matrix with the co-factors of A as entries. In other words, $AC^\top = \det(A)I$.

6.2.3 Cramer's Rule

The solution $\mathbf{x}$ of $A\mathbf{x} = \mathbf{b}$, if $\det(A) \neq 0$, is given by $\mathbf{x}_j = \frac{\det(\mathcal{B}_j)}{\det(A)}$, where $\mathcal{B}_j$ is the matrix obtained by replacing the j-th column of A with $\mathbf{b}$. This is inefficient to compute.

6.2.4 Elimination and the Determinant

If P is a permutation that swaps two rows of A then $\det(PA) = -\det(A)$.

The determinant is linear in each row (and column): For any $a_0, a_1, a_2, \dots, a_n \in \mathbb{R}^n$ and $\alpha_0, \alpha_1 \in \mathbb{R}$ we have

$$\begin{vmatrix} | & | & & | \\ \alpha_0 a_0 + \alpha_1 a_1 & a_2 & \cdots & a_n \\ | & | & & | \end{vmatrix} = \alpha_0 \begin{vmatrix} | & | & & | \\ a_0 & a_2 & \cdots & a_n \\ | & | & & | \end{vmatrix} + \alpha_1 \begin{vmatrix} | & | & & | \\ a_1 & a_2 & \cdots & a_n \\ | & | & & | \end{vmatrix}.$$

For example, $\det \begin{bmatrix} \alpha_0 a_0^T + \alpha_1 a_1^T \\ a_2^T \end{bmatrix} = \alpha_0 \det \begin{bmatrix} a_0^T \\ a_2^T \end{bmatrix} + \alpha_1 \det \begin{bmatrix} a_1^T \\ a_2^T \end{bmatrix}$.

7 Eigenvalues and Eigenvectors

7.1 Complex Numbers $\mathbb{C}$

$$\mathbb{C} = \{z = a + ib : a, b \in \mathbb{R},\ i^2 = -1\}$$

We can do operations with complex numbers:

- $(a + ib) + (x + iy) = (a + x) + i(b + y)$
- $(a + ib)(x + iy) = (ax - by) + i(ay + bx)$
- $(a + ib)(a - ib) = a^2 + b^2$
- $\frac{a+ib}{x+iy} = \frac{ax+by}{x^2+y^2} + i\frac{bx-ay}{x^2+y^2}$

Given $z \in \mathbb{C}$ with $z = a + ib$ we have the following notation

$$\Re(a+ib) := a \quad \text{called the real part of } z = a+ib, \tag{14}$$

$$\Im(a+ib) := b \quad \text{called the imaginary part of } z = a+ib, \tag{15}$$

$$|z| := \sqrt{a^2+b^2} \quad \text{called the modulus of } z = a+ib, \tag{16}$$

$$\overline{a+ib} := a - ib \quad \text{called the complex conjugate of } z = a+ib. \tag{17}$$

Note that for $z_1, z_2 \in \mathbb{C}$, we have $|z|^2 = z\bar{z}$, $z_1 z_2 = z_2 z_1$, $\overline{z_1 + z_2} = \overline{z_1} + \overline{z_2}$, and $\frac{1}{z} = \frac{\bar{z}}{|z|^2}$.
Euler's Formula. Given $\theta \in \mathbb{R}$, $e^{i\theta} = \cos\theta + i\sin\theta$.
This means, that $e^{i\pi} + 1 = 0$.

Polar Coordinates. $z = re^{i\theta}$, where $r = |z|$ and θ is called the argument of z.

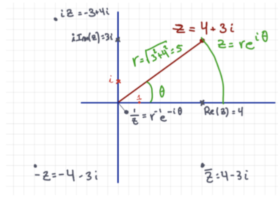


FIGURE 4. A complex number $z = 4 + 3i$ in the complex plane.

Fundamental Theorem of Algebra . Any degree $n \geq 1$ polynomial $P(z) = \alpha_n z^n + \alpha_{n-1} z^{n-1} + \cdots + \alpha_1 z + \alpha_0$ (with $\alpha_n \neq 0$) has a zero $\lambda \in \mathbb{C}$ such that $P(\lambda) = 0$.

Any such polynomial $P(z)$ has n zeros $\lambda_1, \dots, \lambda_n \in \mathbb{C}$, perhaps with repetitions, such that $P(z) = \alpha_n(z - \lambda_1)\cdots(z - \lambda_n)$. The number of times λ_i appears in this expansion is called the **algebraic multiplicity of λ_i** .

7.1.1 Complex-valued Matrices and Vectors

The natural operation of "transposing" for complex vectors and matrices is that of "**conjugate transpose**" or "hermitian transpose": $A^* = \overline{A}^\top$.

Length: $||\mathbf{v}|| = \sqrt{|v_1|^2 + \cdots + |v_n|^2}$.
Scalar (inner) Product: $\langle \mathbf{v}, \mathbf{w} \rangle = \mathbf{w}^* \mathbf{v}$.

7.2 Introduction to Eigenvalues and Eigenvectors

Given $A \in \mathbb{R}^{n\times n}$, we say $\lambda \in \mathbb{C}$ and $\mathbf{v} \in \mathbb{C}^n \setminus \{0\}$ are an **eigenvalue-eigenvector pair** of A when $A\mathbf{v} = \lambda\mathbf{v}$ holds. If $\lambda \in \mathbb{R}$, then we have a real eigenvalue-eigenvector pair.

7.2.1 Finding Eigenvalues and Eigenvectors

An eigenvalue $\lambda \in \mathbb{C}$ is a solution of $\det(A - \lambda I) = 0$, which is a polynomial equation.
An associated eigenvector $\mathbf{v} \in \mathbf{N}(A - \lambda I)$ is a non-zero solution to $(A - \lambda I)\mathbf{v} = \mathbf{0}$.

7.2.2 Theorems and Propositions

Complex EW exist in conjugate pairs for real A : Let $A \in \mathbb{R}^{n\times n}$. If $(\lambda, \mathbf{v})$ is an eigenvalue–eigenvector pair, then $(\overline{\lambda}, \overline{\mathbf{v}})$ is also an eigenvalue–eigenvector pair.

Every $A \in \mathbb{R}^{n\times n}$ has an eigenvalue (perhaps complex-valued).

If λ and $\mathbf{v}$ are an eigenvalue-eigenvector pair of A, then, for $k \geq 1$, λ^k and $\mathbf{v}$ are an eigenvalue-eigenvector pair of A^k.
Proof: Induction Step: $A^k\mathbf{v} = A(A^{k-1}\mathbf{v}) = A(\lambda^{k-1}\mathbf{v}) = \lambda^k\mathbf{v}$.

If λ and $\mathbf{v}$ are an eigenvalue-eigenvector pair of an invertible A (and $\lambda \neq 0$), then, $\frac{1}{\lambda}$ and $\mathbf{v}$ are an eigenvalue-eigenvector pair of A^{-1}.
Proof: $A\mathbf{v} = \lambda\mathbf{v} \Rightarrow A^{-1}(\lambda\mathbf{v}) = \mathbf{v} \Rightarrow A^{-1}\mathbf{v} = \frac{1}{\lambda}\mathbf{v}$.

Let $A \in \mathbb{R}^{n\times n}$ and let $\mathbf{v}_1, \dots, \mathbf{v}_k \in \mathbb{R}^n$ be eigenvectors corresponding to eigenvalues $\lambda_1, \dots, \lambda_k \in \mathbb{R}$. If $\lambda_1, \dots, \lambda_k$ are all distinct, the eigenvectors $\mathbf{v}_1, \dots, \mathbf{v}_k$ are **linearly independent** .

If $A_{n\times n}$ has n distinct real eigenvalues, then the eigenvectors **form a basis for $\mathbb{R}^n$** (Eigenbasis).

The eigenvalues of $A \in \mathbb{R}^{n\times n}$ **are the same as those of A^T** . *Proof:* $\det(A - zI) = \det((A - zI)^\top) = det(A^\top - zI)$.

7.2.3 Characteristic Polynomial

$$(-1)^n \det(A - zI) = \det(zI - A) = (z - \lambda_1) \cdots (z - \lambda_n)$$

Let $A \in \mathbb{R}^{n\times n}$ and $\lambda_1, \dots, \lambda_n$ its n eigenvalues as they show up in the characteristic polynomial (meaning that a value λ may be repeated, the number of times it shows up is the **algebraic multiplicity** of λ), then
$\mathbf{Tr}(A) = \sum_{i=1}^n A_{ii} = \sum_{i=1}^n \lambda_i$, $\det(A) = \prod_{i=1}^n \lambda_i$.
*This is useful to **verify** computations.*
For $A, B, C \in \mathbb{R}^{n\times n}$:
$\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$, **and** $\mathrm{Tr}(ABC) = \mathrm{Tr}(BCA) = \mathrm{Tr}(CAB)$.

7.2.4 Words of Caution

Even thought the eigenvalues of A and $A^\top$ are the same, the eigenvectors are not!

The eigenvalues of $A + B$ are not easily computed from the eigenvalues of A and B, they are not their sum!

The eigenvalues of AB are not easily computed from the eigenvalues of A and B, they are not their product!

Gauss elimination doesn't preserve eigenvalues and eigenvectors.

7.2.5 Eigenvalues of Orthogonal Matrices

Let $Q \in \mathbb{R}^{n\times n}$ be an orthogonal matrix. If $\lambda \in \mathbb{C}$ is an eigenvalue of Q, then $|\lambda| = 1$.
Proof: $||\mathbf{v}||^2 = ||Q\mathbf{v}||^2 = ||\lambda\mathbf{v}||^2 = |\lambda|^2||\mathbf{v}||^2$. Since $\mathbf{v} \neq \mathbf{0}$ we have $|\lambda| = 1$.

7.2.6 Repeated Eigenvalues

Repeated eigenvalues can (but doesn't have to) pose a real obstacle to building a basis.

If, given a matrix $A \in \mathbb{R}^{n\times n}$, we can build a basis of $\mathbb{R}^n$ with eigenvectors of A we say that A has a **complete set of real eigenvectors** .

Given A and an eigenvalue λ of A we call the dimension of $\mathbf{N}(A - \lambda I)$ the **geometric multiplicity** of λ.

A matrix has a complete set of real eigenvectors if and only if all its eigenvalues are real and the geometric multiplicities equal the algebraic multiplicities for all eigenvalues.

Let P be the **Projection Matrix** on the subspace $U \subseteq \mathbb{R}^n$. Then P has two eigenvalues, 0 and 1, and a complete set of real eigenvectors.

For a **Diagonal Matrix** $D_{n\times n}$, the eigenvalues of D are the diagonal entries of D. The canonical basis $\mathbf{e}_1, \dots, \mathbf{e}_n$ is a set of eigenvectors of D.

8 Diagonalization and Singular Value Decomposition

8.1 Diagonalization

8.1.1 Change of Basis of a Linear Transformation

Let $U_{n\times n}$ have the basis elements $\mathbf{u}_1, \dots, \mathbf{u}_n$ for columns and $V_{m\times m}$ the basis elements $\mathbf{v}_1, \dots, \mathbf{v}_m$.
Let $A_{m\times n}$ correspond to a linear transformation $L : \mathbb{R}^n \to \mathbb{R}^m$ in the basis $\mathbf{u}_1, \dots, \mathbf{u}_n$.
The matrix $B_{m\times n}$, corresponding to L written in the new basis $\mathbf{v}_1, \dots, \mathbf{v}_m$, is $B = V^{-1}AU$.

8.1.2 Diagonalization of a Matrix

Let $A \in \mathbb{R}^{n\times n}$ have a complete set of real eigenvectors (eigenbasis). Let $\mathbf{v}_1, \dots, \mathbf{v}_n \in \mathbb{R}^{n\times n}$ be a basis formed with eigenvectors of A and let $\lambda_1, \dots, \lambda_n$ be the associated eigenvalues (λ_i associated to $\mathbf{v}_i$). Let V be the matrix whose columns are the eigenvectors $\mathbf{v}_i$. Then, $A = V\Lambda V^{-1}$, where Λ is a diagonal matrix with $\Lambda_{ii} = \lambda_i$.
Equivalently, $\Lambda = V^{-1}AV$, since V is invertible.
A matrix $A \in \mathbb{R}^{n\times n}$ is called **diagonalizable** if there exists an invertible matrix V such that $V^{-1}AV = \Lambda$, where Λ is a diagonal matrix.

We say that $A \in \mathbb{R}^{n\times n}$ and $B \in \mathbb{R}^{n\times n}$ are **similar matrices** if there exists an invertible matrix S such that $B = S^{-1}AS$.
Similar matrices have the same eigenvalues.

8.2 Symmetric Matrices and the Spectral Theorem

8.2.1 Spectral Theorem

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

Eigendecomposition: 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 eigenvectors of A) such that $A = V\Lambda V^\top$, where $\Lambda \in \mathbb{R}^{n\times n}$ is a diagonal matrix with the eigenvalues of A in its diagonal and $V^\top V = I$.

Rank-One Spectral Decomposition: Let $A \in \mathbb{R}^{n\times n}$ be symmetric, and let $v_1, \dots, v_n$ be an orthonormal basis of eigenvectors of A (the columns of V), with associated eigenvalues $\lambda_1, \dots, \lambda_n$. Then $A = \sum_{k=1}^n \lambda_k\, v_k v_k^T$.
A real symmetric matrix is a weighted sum of orthogonal projections onto its eigenvector directions, with weights given by the eigenvalues.

Rank of real symmetric matrix: For a real symmetric matrix A, $\mathrm{rank}(A) = \#$ non-zero eigenvalues of A (counting repetitions).

Orthogonality of EV : Let $A \in \mathbb{R}^{n\times n}$ be symmetric and let $\lambda_1 \neq \lambda_2 \in \mathbb{R}$ be two distinct eigenvalues of A with corresponding eigenvectors v_1, v_2. Then v_1 and v_2 are orthogonal.

Symmetric matrix has real EW : A symmetric matrix $A \in \mathbb{R}^{n\times n}$ has only real eigenvalues: $\lambda \in \mathbb{C} \Rightarrow \lambda \in \mathbb{R}$. Indeed, if $Av = \lambda v$: $\lambda\|v\|^2 = \overline{\lambda}\, v^*v = (\lambda v)^*v = (Av)^*v = v^*A^*v = v^*Av = v^*\lambda v = \lambda\|v\|^2$. $\Rightarrow$ every symmetric matrix $A \in \mathbb{R}^{n\times n}$ has a real eigenvalue. (All eigenvalues of symmetric matrices are real. And every symmetric matrix has a real eigenvalue.)

8.2.2 Rayleigh Quotient

Given a symmetric matrix A the Rayleigh Quotient, is defined for $\mathbf{x} \neq \mathbf{0}$, as $R(\mathbf{x}) = \frac{\mathbf{x}^\top A\mathbf{x}}{\mathbf{x}^\top\mathbf{x}}$. The minimum of R is $R(v_{\min}) = \lambda_{\min}$, and the maximum $R(v_{\max}) = \lambda_{\max}$. Here $\lambda_{\max}/\lambda_{\min}$ are the largest/smallest eigenvalues of A, and $v_{\max}/v_{\min}$ their associated eigenvectors.

8.2.3 Positive (Semi)definite Matrix

A symmetric matrix A is Positive Semidefinite (PSD) if all its eigenvalues are non-negative. If all its eigenvalues are strictly positive, then A is Positive Definite (PD).
A matrix A is PSD if and only if $\mathbf{x}^\top A\mathbf{x} \geq 0$ for all $\mathbf{x} \in \mathbb{R}^n$.
A matrix A is PD if and only if $\mathbf{x}^\top A\mathbf{x} > 0$ for all $\mathbf{x} \neq 0$.

(If A and B are PSD (or PD) then their sum is PSD (or PD).)

Given a real matrix $A \in \mathbb{R}^{m\times n}$, the non-zero eigenvalues of $A^\top A \in \mathbb{R}^{n\times n}$ are the same as the ones of $AA^\top \in \mathbb{R}^{m\times m}$. Both are symmetric and positive semidefinite.

8.2.4 Gram Matrix

Given vectors $v_1, \dots, v_n \in \mathbb{R}^m$, their *Gram matrix* is $G \in \mathbb{R}^{n\times n}$ defined by $G_{ij} = v_i^\top v_j$. If $V = [v_1 \cdots v_n] \in \mathbb{R}^{m\times n}$, then $G = V^\top V$.
If $A = [a_1 \cdots a_n] \in \mathbb{R}^{m\times n}$, one also calls $AA^\top$ a Gram matrix; note that $AA^\top = \sum_{i=1}^n a_i a_i^\top$. It is a $m \times m$ matrix.

Every positive semidefinite matrix M is a gram matrix of an upper triangular matrix C. $M = C^\top C$ is known as the **Cholesky Decomposition**.

8.3 Singular Value Decomposition

Every matrix $A \in \mathbb{R}^{m\times n}$ has a SVD (Singular Value Decomposition) of the form
$$A = U\Sigma V^\top$$
where $\Sigma \in \mathbb{R}^{m\times n}$ is a rectangular diagonal matrix with entries $\sigma_i = \Sigma_{ii}$ called **Singular Values** of A and ordered $\sigma_1 \geq \cdots \geq \sigma_{\min(m,n)}$, $U \in \mathbb{R}^{m\times m}$ ($V \in \mathbb{R}^{n\times n}$) is orthogonal with columns called Left (Right) **Singular Vectors** of A.
This also means: Every linear transformation is diagonal in orthonormal bases of singular vectors.

8.3.1 Compact SVD

If $\mathrm{rank}(A) = r$, then $A = U_r\Sigma_r V_r^\top$ where U_r, Σ_r, V_r contain the first r singular vectors or values.

SVD as a sum of rank-one matrices: We can write any rank-r matrix A as a sum of r rank-1 matrices: $A = \sum_{k=1}^r \sigma_k \mathbf{u}_k\mathbf{v}_k^\top$ where $\mathbf{u}_k$ ($\mathbf{v}_k$) is the k-th column of U (V) (the corresponding singular vectors).

8.3.2 Constructing the matrices

Since $AA^\top = U(\Sigma\Sigma^\top)U^\top$, the columns of U are the eigenvectors of $AA^\top$ and the singular values of A are the square-root of the eigenvalues of $AA^\top$.
If $m > n$, A has n singular values and $AA^\top$ has m eigenvalues, but the "missing" ones are 0.

Since $A^\top A = V(\Sigma^\top\Sigma)V^\top$, the columns of V are the eigenvectors of $A^\top A$ and the singular values of A are the square-root of the eigenvalues of $A^\top A$.
If $n > m$, A has m singular values and $A^\top A$ has n eigenvalues, but the "missing" ones are 0.

8.4 Show that SVD is valid

To show that a given V, U and Σ 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: Σ should only contain entries ≥ 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: $AA^{-1} = I = A^\top A \Leftrightarrow AV\Sigma^{-1}U^\top = U\Sigma V^\top V\Sigma^{-1}U^\top = I$

9 Exercise Tricks

Non-Trivial null-space $\Rightarrow$ 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 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 $n = \text{len}(a), m = \text{len}(b)$ we have

$$\sum_{i=0}^{n-1}\left(\sum_{j=0}^{m-1} a[i]\cdot b[j] * 10^{m-1-j}\right)\cdot 10^{n-1-i}$$

9.2 Subspace of matrix vector space

$U = \{A \in \mathbb{R}^{2\times 2} : \text{Tr}(A) = 0\}$. 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.

9.3 Calculating the Determinant

- **Method A** (2×2) : $\det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc$.
- **Method B (Triangular):** Product of diagonal entries.
- **Method C (General** $n \times n$**:**
  1. Use row operations to convert A to an upper triangular matrix U.
  2. **Track changes:** Row swap: multiply det by -1; Row subtraction $(R_i - kR_j)$: det does **not** change; Scalar multiplication (kR_i): multiply det by k.
- $\det(A)$ = (corresponding factors) $\times \prod \mathbf{u}_{ii}$.

9.4 Linear Independence

To check if vectors $\mathbf{v}_1, \dots, \mathbf{v}_n$ are linearly independent:

1. Form matrix $A = [\mathbf{v}_1 \dots \mathbf{v}_n]$.
2. Perform **Gaussian Elimination** to get REF.
3. If every column has a pivot (no free variables), they are *independent*; otherwise, they are *dependent*.

9.5 Invertible matrix properties

A is invertible
$\Leftrightarrow$ $A\mathbf{x} = \mathbf{b}$ has a unique solution for every $\mathbf{b}$
$\Leftrightarrow$ $N(A) = \mathbf{0}$
$\Leftrightarrow$ $rank(A) = n$
$\Leftrightarrow$ columns of A are lin. independent

9.6 Homogenous system solutions

A homogenous system $A\mathbf{x} = 0$ has either one solution $x = 0$ or infinitely many solutions: all scaled $\alpha\mathbf{x}$.

9.7 SVD

$A \in \mathbb{R}^{n\times n}$, A invertible. σ_1 largest SV of A, σ_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 $\frac{1}{\sigma_1} \leq \dots \leq \frac{1}{\sigma_n}$ (bc. Σ 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 $\sigma_1 \geq \sigma_n > 0$

9.7 Computing SVD Step-by-Step

Target: $A = U\Sigma V^\top$. (Rank r).

1. **Right Singular Vectors** (V): Compute $M = A^\top A$. Find eigenvalues λ_i and orthonormal eigenvectors $\mathbf{v}_i$ of M. Sort $\lambda_1 \geq \lambda_2 \geq \dots \geq 0$. $V = [\mathbf{v}_1 \dots \mathbf{v}_n]$.
2. **Singular Values** (Σ)**:** $\sigma_i = \sqrt{\lambda_i}$. Matrix Σ has σ_i on diagonal.
3. **Left Singular Vectors** (U)**:** For $i = 1 \dots r$ $(\sigma_i \neq 0)$: $\mathbf{u}_i = \frac{1}{\sigma_i} A\mathbf{v}_i$. For $i > r$: Extend to orthonormal basis of $\mathbb{R}^m$ (Gram-Schmidt on Nullspace of $A^\top$).

9.7 Singular Values inversion

Since $\sigma_1 \geq \dots \geq \sigma_n > 0$, it follows that $\frac{1}{\sigma_1} \leq \dots \leq \frac{1}{\sigma_n}$, and therefore, when ordered decreasingly, the singular values of A^{-1} are $\frac{1}{\sigma_n}, \dots, \frac{1}{\sigma_1}$.

9.8 Fitting a Polynomial (Least Squares)

Task: Fit $p(t) = \alpha_0 + \alpha_1 t + \dots + \alpha_k t^k$ to points $(t_1, y_1), \dots, (t_m, y_m)$.

1. Setup $A\mathbf{x} = \mathbf{b}$ where unknowns $\mathbf{x} = (\alpha_0, \dots, \alpha_k)^\top$.
2. Matrix A (Vandermonde structure):

$$A = \begin{bmatrix} 1 & t_1 & t_1^2 & \dots \\ \vdots & \vdots & \vdots & \\ 1 & t_m & t_m^2 & \dots \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix}$$

3. Solve Normal Equations: $A^\top A\hat{\mathbf{x}} = A^\top\mathbf{b}$.

9.9 Tricks for 2×2 Eigenvalues

$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. Char Poly: $\lambda^2 - \text{Tr}(A)\lambda + \det(A) = 0$.

- $\lambda_{1,2} = \frac{\text{Tr} \pm \sqrt{\text{Tr}^2 - 4\det}}{2}$.
- **Real Eigenvalues:** iff Discriminant $D \geq 0$.
- **One Real Eigenvalue:** iff $D = 0$.
- **Complex Eigenvalues:** iff $D < 0$ (conjugate pair $a \pm bi$).

9.10 Calculate inverse

1. Form the augmented matrix $[A|I]$.
2. Perform **Gauss-Jordan Elimination**.
3. Once $[I|B]$ is reached, then $B = A^{-1}$. If you get a row of zeros on the left side, A is not invertible (singular).

9.10 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 $Ax_1 = e_1$, $Ax_2 = e_2$, ... Solving all these SLE gives us solutions for all the variables in the original matrix A.

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 Λ^n, which is simple because Λ 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

$\frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$, for $f(x) = ax^2 + bx + c$, result $\mathbb{R}$ iff $b^2 - 4ac \geq 0$.

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, $r = \text{rank}(A)$.

Number of solutions:

R_0	$r = n$	$r < n$
$r = m$	1	∞
$r < m$	0 / 1	0 / ∞

9.19 Quick Sanity Checks

- **Trace:** $\text{Tr}(A) = \sum a_{ii} = \sum \lambda_i$ (Sum of diagonal = sum of eigenvalues).
- **Determinant:** $\det(A) = \prod \lambda_i$ (Product of eigenvalues).
- **Rank:** Rank = Dimension of $C(A)$ = Dimension of $R(A)$ = Number of non-zero singular values.
- **Symmetry:** If A is symmetric, eigenvalues are real, eigenvectors are orthogonal.
- **Orthogonal matrix** Q**:** $Q^\top Q = I$. Determinant is ± 1. Preserves lengths $(\|Q\mathbf{x}\| = \|\mathbf{x}\|)$

10 Proof Patterns

- **Composition of Implication:** If $S \Rightarrow T$ and $T \Rightarrow U$ are both true, then $S \Rightarrow U$ is true.
- **Direct proof of Implication:** Prove $S \Rightarrow T$ assuming S and then proving S under that assumption.
- **Indirect proof of Implication:** Prove $S \Rightarrow T$ by assuming $\neg T$ and proving $\neg S$ under that assumption.
- **Case distinction:** Prove S by finding a list of $R_1, \dots, R_n$ (cases) proving at least one R_i, then showing that $R_i \Rightarrow S$ for all R_i.
- **Proof by contradiction**: Prove S by assuming it to be false, deriving statements from it until reaching a contradiction.
- **Existence proof:** Prove S is true for at least one value
- **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

10.1 How to prove U is a Subspace

To prove $U \subseteq V$ is a subspace:

1. **Check 1 (Zero):** Show $\mathbf{0} \in U$. (Usually easy, if fails $\rightarrow$ not a subspace).
2. **Check 2 (Closure):** Let $\mathbf{u}, \mathbf{v} \in U$ and $\lambda \in \mathbb{R}$. Show $\lambda\mathbf{u} + \mathbf{v} \in U$.

Counter-example: To disprove, find specific vectors where closure fails or show $\mathbf{0} \notin U$.

10.2 How to prove Linear Independence

To prove $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ are L.I.:

1. Set up equation: $\sum_{i=1}^k \lambda_i \mathbf{v}_i = \mathbf{0}$.
2. Show that this implies $\lambda_1 = \dots = \lambda_k = 0$.
3. *Matrix way:* Form $A = [\mathbf{v}_1 \dots \mathbf{v}_k]$. Show $N(A) = \{\mathbf{0}\}$ (e.g. rank k).

10.3 How to prove Surjectivity / Injectivity

Let $T : V \rightarrow W$ be linear (matrix A).

- **Injective (1-to-1):** Show $\text{Ker}(T) = \{\mathbf{0}\}$. (Solve $A\mathbf{x} = \mathbf{0} \implies \mathbf{x} = \mathbf{0}$).
- **Surjective (Onto):** Show $\text{Im}(T) = W$. (Rank = dim(W)).
- **Bijective:** Show both (or if $\dim(V) = \dim(W)$, just one is enough).

10.4 Proving Matrix Properties

- **Symmetric:** Show $A^\top = A$. (Use $(AB)^\top = B^\top A^\top$).
- **Orthogonal:** Show $Q^\top Q = I$. (Cols are orthonormal).
- **Positive Definite:** Show $\mathbf{x}^\top A\mathbf{x} > 0$ for all $\mathbf{x} \neq \mathbf{0}$.

10.4 Proof of symmetry

$A \in \mathbb{R}^{n\times n}$ 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

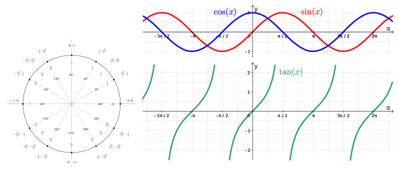
10.5 Proving commutativity of matrix operation

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 $A(BA - I) = 0$

10.6 Prove the dimension of a subspace

Write the general matrix, apply the constraints, deconstruct into linear combination of independent variables, prove that they span the space, show that they are linearly independent.

11 Trigonometry



12 Example Matrices

Symmetric $(A = A^\top)$: $\lambda_i \in \mathbb{R}$. Distinct eigenvectors orthogonal.

- 2×2: $A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \implies \det = -3, \text{Tr} = 2, \lambda = \{3, -1\}$
- 3×3: $A = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} \implies \det = 4, \text{Tr} = 6, \lambda = \{2, 2 \pm \sqrt{2}\}$ (PSD)

Skew-Symmetric $(A = -A^\top)$: $\text{Tr} = 0, \lambda = 0$ or imaginary. det = 0 if n is odd.

- 2×2: $A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \implies \det = 1, \lambda = \{i, -i\}$
- 3×3: $A = \begin{bmatrix} 0 & 1 & 2 \\ -1 & 0 & 3 \\ -2 & -3 & 0 \end{bmatrix} \implies \det = 0, \lambda = \{0, \pm i\sqrt{14}\}$

Orthogonal $(Q^\top Q = I)$: $\det = \pm 1, |\lambda_i| = 1$. Preserves $\|x\|$ and $\angle(x, y)$.

- 2×2 (Rot): $Q = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \implies \det = 1, \lambda = \{e^{i\pi/4}, e^{-i\pi/4}\}$
- 3×3 (Perm): $P = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \implies \det = 1, \lambda = \{1, e^{i2\pi/3}, e^{i4\pi/3}\}$

Nilpotent $(A^k = 0)$: All $\lambda = 0 \implies \det = 0, \text{Tr} = 0$.

- 2×2: $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \implies A^2 = 0$, only one eigenvector $v = (1, 0)^\top$
- 3×3: $A = \begin{bmatrix} 0 & 1 & 2 \\ 0 & 0 & 3 \\ 0 & 0 & 0 \end{bmatrix} \implies A^3 = 0, \lambda = \{0, 0, 0\}$

Diagonally Dominant $(|a_{ii}| > \sum_{j\neq i}|a_{ij}|)$: Always invertible $(\det \neq 0)$.

- 3×3: $A = \begin{bmatrix} 3 & 1 & 1 \\ 1 & 4 & 2 \\ 0 & 1 & 2 \end{bmatrix} \implies$ Row 1: $3 > 1 + 1$, Row 2: $4 > 1 + 2$

Rank-1 $(A = uv^\top)$: rank $= 1, \lambda = \{u^\top v, 0, \dots, 0\}$.

- 3×3: $A = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \\ 3 & 3 & 3 \end{bmatrix} \implies \text{Tr} = 6, \lambda = \{6, 0, 0\}, \det = 0$

Projection $(P^2 = P)$: $\lambda \in \{0, 1\}$. $\text{rank}(P) = \text{Tr}(P)$.

- 3×3: $P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \implies \lambda = \{1, 1, 0\}, \text{rank} = 2, \text{Tr} = 2$

12.1 Transformations of $\mathbb{R}^2$

Examples for linear transformations from $\mathbb{R}^2$ to itself, corresponding to square matrices $A \in \mathbb{R}^{2\times 2}$:

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$$

corresponds to **stretching** by a factor of 2 in the vertical axis.

$$A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$$

corresponds to a **shearing** transformation.

$$A = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

corresponds to a counter-clockwise **rotation** by $\frac{\pi}{4}$.

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

corresponds to a **reflection** by the diagonal line $x_2 = x_1$.