

L 4.11 Column space and Nullspace is a subspace: Let $A \in \mathbb{R}^{m \times n}$, then $\mathbf{C}(A) = \{A\mathbf{x} : \mathbf{x} \in \mathbb{R}^n\}$ is a subspace of $\mathbb{R}^m$. Moreover, $\mathbf{R}(A) = \mathbf{C}(A^\top)$ is a subspace of $\mathbb{R}^n$. The nullspace $\mathbf{N}(A) = \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} = \mathbf{0}\}$ is a subspace of $\mathbb{R}^n$.

L 4.14 Subspaces are vector spaces: V is a vector space and U is its subspace. Then U is also a vector space with the same $\oplus$ and $\odot$ as V .

4.2 Bases and dimension

D 4.18 Basis: 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** $V : \text{Span}(B) = V$.

A basis is not always a set of *vectors*, it is a set of elements of U . Thus for the subspace of diagonal matrices D_m , the basis is made up of *diagonal matrices*.

The basis of $\{0\}$ is $\emptyset$. Since there is no vector, $\emptyset$ is vacuously independent and **Span** $(\emptyset) = \{0\}$ since an empty sum yields 0.

O 4.20 Non-uniqueness of basis: Every set $B = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\} \subseteq \mathbb{R}^m$ of m linearly independent vectors is a basis of $\mathbb{R}^m$. $\mathbb{R}[x]$ is not finitely generated for example.

D 4.21 Finitely generated vector space: A vector space V is called finitely generated if there exists a finite subset $G \subseteq V$ with $\text{Span}(G) = V$ Then V has a basis $B \subseteq G$.

T 4.22 Finitely generated VS has a basis: If V is finitely generated, then V has a basis $B \subseteq V$.

L 4.23 Steinitz exchange lemma: Let $F \subseteq V$ be a finite set of linearly independent vectors, $G \subseteq V$ a finite set of vectors with $\text{Span}(G) = V$. Then:

- $|F| \leq |G|$
- $\exists$ subset $E \subseteq G$ of size $|G| - |F|$ such that $\text{Span}(F \cup E) = V$.

T 4.24 All bases have the same size: All bases have the same size: $B, B' \in V \Rightarrow |B| = |B'|$.

D 4.25 Dimension: If V is finitely generated, then $d = \dim(V)$ is the size of any basis B of V .

D 4.26 Linear transformation between vector spaces: Let V, W be vector spaces. A function $T : V \rightarrow W$ is linear if, for all $\mathbf{x}_1, \mathbf{x}_2 \in V$ and $\lambda_1, \lambda_2 \in \mathbb{R}$, $T(\lambda_1\mathbf{x}_1 + \lambda_2\mathbf{x}_2) = \lambda_1T(\mathbf{x}_1) + \lambda_2T(\mathbf{x}_2)$.

L 4.27 Bijective linear transformations preserve basis: If $T : V \rightarrow W$ is a bijective linear map, then $B \subseteq V$ is a basis of $V \Leftrightarrow T(B)$ is a basis of W , and hence $\dim(V) = \dim(W)$.

D 4.28 Isomorphic vector spaces: $V \cong W \Leftrightarrow \exists T : V \rightarrow W$ linear and bijective.

T 4.29 Basis writes vectors as unique linear combinations: Let V be a finitely generated vector space with basis $B = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m\}$. then every $\mathbf{v} \in V$ can be written **uniquely** as $\mathbf{v} = \sum_{j=1}^m \lambda_j \mathbf{v}_j$, for unique scalars $\lambda_1, \lambda_2, \dots, \lambda_m \in \mathbb{R}$.

L 4.30 Less than $\dim(V)$ vectors do not span V : If $|G| < \dim(V)$, then $\text{Span}(G) \neq V$.

4.3 Computing the three fundamental subspaces

T 4.31 Basis of $\mathbf{C}(A)$: Pivots columns of RREF: If R is RREF of A , then all columns at pivots of R form a basis of $\mathbf{C}(A) : \dim(\mathbf{C}(A)) = \text{rank}(A) = r$

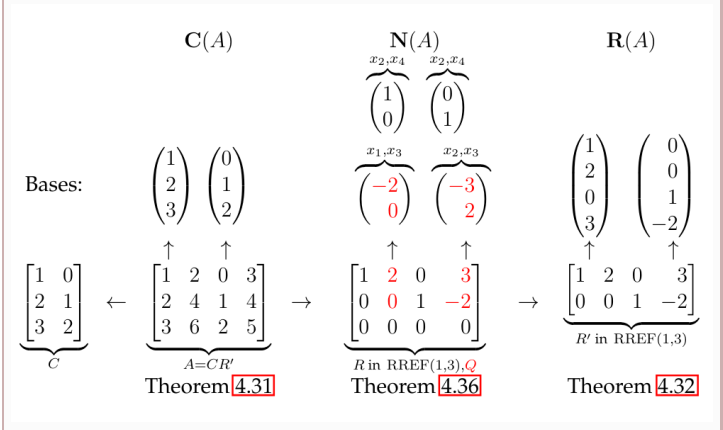
T 4.32 Basis of $\mathbf{R}(A)$: Non-zero rows of RREF of A : Non-zero rows of RREF of A form a basis of $\mathbf{R}(A)$, so $\dim(\mathbf{R}(A)) = r$ (multiplying by and invertible matrix M does not change the row-space of MA on the left.)

T 4.33 Row rank equals column rank: $\text{rank}(A) = \text{rank}(A^\top)$

C 4.34 Rank is at most min of the matrix dimensions: A is a $m \times n$ matrix with $\text{rank } r \Rightarrow r \leq \min(n, m)$.

T 4.36 Basis of $N(A)$: Non-pivot columns of RREF(A): If $\text{rank}(A) = r$, then $\dim(N(A)) = n - r = n - \text{rank}(A)$.

T 4.36 Basis of $N(A^\top)$: If $\text{rank}(A) = r$, then $\dim(N(A^\top)) = m - r = m - \text{rank}(A)$.



4.4 All solutions of $A\mathbf{x} = \mathbf{b}$

D 4.37 Solution space: Set of all solutions of $A\mathbf{x} = \mathbf{b}$, thus $\text{Sol}(A, \mathbf{b}) := \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} = \mathbf{b}\} \subseteq \mathbb{R}^n$.

T 4.38 Solution space from shifting the nullspace: Let $\mathbf{s}$ be some solution of $A\mathbf{x} = \mathbf{b}$, then $\text{Sol}(A, \mathbf{b}) := \{\mathbf{s} + \mathbf{x} \in \mathbb{R}^n : \mathbf{x} \in N(A)\}$. We can also compute $\text{Sol}(A, \mathbf{b})$, although it is not a subspace. To describe all solutions, we need *some* solutions.

T 4.39 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: $\dim(\text{Sol}(A, \mathbf{b})) = n - r$, and $\dim(\text{Sol}(A, \mathbf{b})) := \dim(N(A))$

T 4.40 Systems of rank m are solvable: Let $A \in \mathbb{R}^{m \times n}$ with $\text{rank}(A) = m$, $A\mathbf{x} = \mathbf{b}$ is solvable for all $\mathbf{b} \in \mathbb{R}^m$.

T 4.41 Systems of rank less than m are typically unsolvable: Systems of rank $r < m$ are typically unsolvable.

D 4.42 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) — **typically solvable**.
- $m < n \Rightarrow$ underdetermined (A is a wide matrix) — **typically solvable**.
- $m > n \Rightarrow$ overdetermined (A is a tall matrix) — **typically unsolvable**.

5 Orthogonality and Projections

5.1 Definition

D 5.1.1 Orthogonal subspaces: Two vectors are orthogonal if their scalar product is 0 : $\mathbf{v}^\top \mathbf{w} = \sum_{i=1}^n \mathbf{v}_i \mathbf{w}_i = 0$. Two subspaces are orthogonal if all $\mathbf{v}$ and $\mathbf{w}$ are orthogonal.

L 5.1.2 Orthogonality of bases: Let $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ and $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_l$ be bases of subspaces V and W . V and W are orthogonal iff. $\mathbf{v}_i$ and $\mathbf{w}_j$ are orthogonal for all $i \in \{1, \dots, k\}$ and $j \in \{1, \dots, l\}$.

L 5.1.3 Linear independence of bases of orthogonal subspaces: Let V and W be two orthogonal subspaces of $\mathbb{R}^n$ with the bases $\mathbf{v}_1, \dots, \mathbf{v}_k$, $\mathbf{w}_1, \dots, \mathbf{w}_l$. The set of vectors $\{\mathbf{v}_1, \dots, \mathbf{v}_k, \mathbf{w}_1, \dots, \mathbf{w}_l\}$ is linearly independent.

C 5.1.4 Combinations of subspaces: Let V and W be orthogonal subspaces. Then $V + W$ is a subspace of $\mathbb{R}^n$, and $V \cap W = \{0\}$. If $\dim(V) = k$ and $\dim(W) = l$, then $\dim(V + W) = k + l \leq n$, for $V, W \subseteq \mathbb{R}^n$.

D 5.1.5 Orthogonal Complement: Let V be a subspace of $\mathbb{R}^n$, its *orthogonal complement*: $V^\perp = \{\mathbf{w} \in \mathbb{R}^n | \mathbf{w}^\top \mathbf{v} = 0, \forall \mathbf{v} \in V\}$

T 5.1.6 Relations between subspaces: $N(A) = C(A^\top)^\perp = R(A)^\perp$ and $C(A^\top) = N(A)^\perp$

T 5.1.7 Vector decomposition by orthogonal complements: $W = V^\perp \Leftrightarrow \dim(V) + \dim(W) = n \Leftrightarrow \mathbf{u} = \mathbf{v} + \mathbf{w}, \forall \mathbf{u} \in \mathbb{R}^n$ with unique vectors $\mathbf{v} \in V, \mathbf{w} \in W$.
Example We can write: $\mathbb{R}^n = V + V^\perp = \{v + w \mid v \in V, w \in V^\perp\}$

T 5.1.8 Orthogonal: Let V be a subspace of $\mathbb{R}^n$. Then $V = (V^\perp)^\perp$

L 5.1.10 Justification of existence solutions for normal equations: Let $A \in \mathbb{R}^{m \times n}$. Then $N(A) = N(A^\top A)$ and $C(A^\top) = C(A^\top A)$.

$N(A) = N(A^\top A)$: if $N(A) = \{0\}$ i.e. A lin. indep. columns then $N(A^\top A)$ invertible as $N(A^\top A) = \{0\}$ and square.

$C(A^\top) = C(A^\top A)$: needed so $A^\top A \hat{x}$ is in the same subspace as $A^\top b$.

Proof of $N(A) = N(A^\top A)$:
If $x \in N(A)$ then $Ax = 0 \Rightarrow A^\top Ax = A^\top 0 \Rightarrow A^\top Ax = 0$.
If $x \in N(A^\top A)$ then $A^\top Ax = 0$, which means $0 = x^\top 0 = x^\top A^\top Ax = (Ax)^\top (Ax) = ||Ax||^2 \Rightarrow Ax = 0$

Proof of $C(A^\top) = C(A^\top A)$:
If $x \in C(A^\top A)$ then let $\exists y A^\top Ay = x$ and if we set $z = Ay$ then $A^\top z = x$ thus $x \in C(A^\top)$.
We have $C(A^\top) = N(A)^\perp = N(A^\top A)^\perp$ as $(N(A^\top A) = N(A))$. Then $N(A^\top A)^\perp = C((A^\top A)^\top)$ (by $N(B)^\perp = C(B^\top)$) and $(A^\top A)^\top = A^\top A$ thus $C(A^\top) = C(A^\top A)$.

5.2 Projections

D 5.2.1 Projections: Projecting a vector onto a subspace is done with $\text{proj}_S(\mathbf{b}) = \underset{\mathbf{p} \in S}{\text{argmin}} \|\mathbf{b} - \mathbf{p}\|$. and yields the closest point in the new subspace S .
 $b = p + e \Rightarrow b - p = e$, with e the error.

L 5.2.2 One-dimensional projection formula: Projection of $\mathbf{b}$ on $S = \{\lambda \mathbf{a} | \lambda \in \mathbb{R}\} = C(\mathbf{a})$:
 $\text{proj}_S(\mathbf{b}) = \frac{\mathbf{a} \mathbf{a}^\top}{\mathbf{a}^\top \mathbf{a}} \mathbf{b}$, where $\mathbf{a} \in \mathbb{R}^m \setminus \{0\}$.

L 5.2.3 General Projection Formula: Let S be a subspace in $\mathbb{R}^m$ with basis $\mathbf{a}_1, \dots, \mathbf{a}_n$ that span S . The projection to $S = C(A)$ is well-defined. The general formula: $\text{proj}_S(\mathbf{b}) = A \hat{\mathbf{x}}$, where $\hat{\mathbf{x}}$ s.t. $A^\top A \hat{\mathbf{x}} = A^\top \mathbf{b}$.
Note that there might be multiple if A doesn't have linearly indep. columns.

L 5.2.4 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, and invertible.

T 5.2.5 Projection in terms of projection matrix: Let A be the matrix with column vectors $\mathbf{a}_1, \dots, \mathbf{a}_n$ (lin. indep. columns). The projection of $b \in \mathbb{R}^m$ to S is given by $\text{proj}_S(\mathbf{b}) = P \mathbf{b}$ with projection matrix $P = A(A^\top A)^{-1} A^\top$. A needs to have **linearly independent** columns.

T 5.2.6 Idempotent: $PPb = Pb \Leftrightarrow P^2 = P$.

Projection Matrix Symmetric: $P^\top = P$ (has to be reproven in exam)
 $P^\top = (A(A^\top A)^{-1} A^\top)^\top = (A^\top)^\top (A^\top A)^{-1 \top} A^\top = A(A^\top A)^{-1} A^\top = P$. We use the fact that for invertible matrices $M^{-1 \top} = M^{\top -1}$.

T 5.2.6 Projection on $S^\perp$: $(I - P)$ is the projection matrix that maps $b \in \mathbb{R}^m$ to $\text{proj}_{S^\perp}(b)$.

6 Applications of Orthogonality and Projections

6.1 Least Squares Approximation

Fact 6.1.1: A minimiser of $\min_{x' \in \mathbb{R}^n} ||Ax - b||^2$ is also a solution of $A^\top A \hat{x} = A^\top b$.
When A has linearly independent columns, the minimiser is given by $\hat{x} = (A^\top A)^{-1} A^\top b$.

For $A = QR$: The minimiser is a solution of $Rx = Q^\top b$.

Least Squares: Assume $Ax = b$ does not always have a solution, however we want to get the "best" solution according to $\min_{x' \in \mathbb{R}^n} ||Ax' - b||^2$. We can solve this using projections as follows.

- First write down the equation in form of e.g. $b = \alpha_3 t^3 + \alpha_2 t^2 + \alpha_1 t + \alpha_0$

- Now write using matrices
$$\begin{bmatrix} 1 & \dots & 1 \\ t^3 & \dots & 1 \\ \vdots & \dots & \vdots \\ 1 & \dots & 1 \end{bmatrix} \begin{pmatrix} \alpha_3 \\ \vdots \\ \alpha_0 \end{pmatrix} = \begin{pmatrix} b_3 \\ \vdots \\ b_0 \end{pmatrix}$$

- Normal Equations: $(A^\top A)x' = (A^\top b) \Rightarrow Nx' = y \Rightarrow x' = N^{-1}y$

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 define a matrix $A = \begin{bmatrix} 1 & \mathbf{t}_1 \\ \vdots & \vdots \\ 1 & \mathbf{t}_n \end{bmatrix}$ and a result vector $\mathbf{b} = \begin{bmatrix} \mathbf{b}_1 \\ \vdots \\ \mathbf{b}_n \end{bmatrix}$

We solve $\min_{\alpha_0, \alpha_1} \sum_{k=1}^m (b_k - [\alpha_0 + \alpha_1 t_k])^2$ here. Note that we could also fit a parabola, etc...

L 6.1.2: The columns of the $m \times 2$ matrix defined before are linearly dependent if and only if $t_i = t_j$ for all $i \neq j$. (All datapoints are unique in time)

6.2 The set of all solutions to a system of linear equations

L 6.2.1 Injectivity of A on $C(A^\top)$, uniqueness of solutions: $A \in \mathbb{R}^{m \times n}$, $\mathbf{x}, \mathbf{y} \in C(A^\top) : A\mathbf{x} = A\mathbf{y} \Leftrightarrow \mathbf{x} = \mathbf{y}$, which leads to: $C(A^\top) \cap N(A) = \{0\}$

Proof of $x, y \in C(A^\top)$ then $Ax = Ay \iff x = y$: $Ax = Ay \iff Ax - Ay = 0 \iff A(x - y) = 0 \iff x - y \in N(A)$ and $x - y \in C(A^\top) \iff x - y = 0$ thus $x = y$.

T 6.2.2 Set of all solutions of linear equations: Suppose the set of all solutions, $\{\mathbf{x} \in \mathbb{R}^n | A\mathbf{x} = \mathbf{b}\} \neq \emptyset$, then $\{\mathbf{x} \in \mathbb{R}^n | \mathbf{x} = \mathbf{b}\} = \mathbf{x}_1 + N(A)$, $\mathbf{x}_1 \in R(A)$ is unique such that $A\mathbf{x}_1 = \mathbf{b}$.

C 6.2.3: Suppose that $\{\mathbf{x} \in \mathbb{R}^n | A\mathbf{x} = \mathbf{b}\} \neq \emptyset$. Then there exists a unique vector $\mathbf{x}_1 \in C(A^\top A)$ such that $A\mathbf{x}_1 = \mathbf{b}$.

T 6.2.4 Linear equations with no solution: For linear equations that have no solutions, these statements are equivalent:

$$\{\mathbf{x} \in \mathbb{R}^n | A\mathbf{x} = \mathbf{b}\} = \emptyset \iff \{\mathbf{z} \in \mathbb{R}^m | A^\top \mathbf{z} = 0, \mathbf{b}^\top \mathbf{z} = 1\} \neq \emptyset$$

Show that a vector $\mathbf{b}$ is linearly independent from a set of vectors $a_1, \dots, a_n$. We just put them into the matrix equation $A\mathbf{x} = \mathbf{b}$. If there is no solution, $\mathbf{b}$ is independent.

6.3 Orthogonal Bases and Gram Schmidt

D 6.3.1 Orthogonal vectors: $\mathbf{q}_i^\top \mathbf{q}_j = \delta_{ij} = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}$, with δ_{ij} being the *Kronecker delta*. // All vectors are pairwise orthogonal $\mathbf{q}_i^\top \mathbf{q}_j = 0$ and all of them have norm 1: $\mathbf{q}^\top \mathbf{q} = ||\mathbf{q}||^2 = ||\mathbf{q}|| = 1$.

D 6.3.2 Orthogonal Matrix: Q orthogonal $\iff Q^\top Q = I$

D 6.3.2 Orthogonal matrix: If it is **square**, then $QQ^\top = I$, $Q^{-1} = Q^\top$, and the columns of Q form an orthogonal basis for $\mathbb{R}^n$.

P 6.3.6 Preserving qualities of orthogonal matrices: Orthogonal matrices preserve norm and inner product of vectors: $||Q\mathbf{x}|| = ||\mathbf{x}||$ and $(Q\mathbf{x})^\top (Q\mathbf{y}) = \mathbf{x}^\top \mathbf{y}$

P 6.3.7 Least square solution to $Q\mathbf{x} = \mathbf{b}$: The Least Squares solution to $Q\mathbf{x} = \mathbf{b}$, where Q is the matrix whose columns are the vectors forming the orthogonal basis of $S \subseteq \mathbb{R}^m$, is given by $\hat{\mathbf{x}} = Q^\top \mathbf{b}$ and the projection matrix is given by $QQ^\top$.

D 6.3.8 Gram-Schmidt algorithm: This algorithm is used to construct orthogonal bases. We have linearly independent vectors $\mathbf{a}_1, \dots, \mathbf{a}_n$, that span a subspace S , then Gram-Schmidt constructs $\mathbf{q}_1, \dots, \mathbf{q}_n$ by:

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

D 6.3.10 QR Decomposition: $A = QR$, where $R = Q^\top A$ and Q is obtained from the Gram-Schmidt process, is made up of the vectors $\mathbf{q}_i$ as columns.

L 6.3.11 Well-Defined QR Decomposition: R is an upper triangular matrix and invertible. Moreover, $QQ^\top A = A$ (Projection of A onto A), and hence $A = QR$ is well-defined.

P 6.3.12: This greatly simplifies calculations involving projections and least squares, since $C(A) = C(Q)$, so $\text{proj}_{C(A)}(\mathbf{b}) = QQ^\top \mathbf{b}$ and for least squares, we have $R\hat{\mathbf{x}} = Q^\top \mathbf{b}$. This can efficiently be solved using back-substitution because R is upper triangular. $A^\top A \hat{\mathbf{x}} = A^\top \mathbf{b} \iff R^\top Q^\top QR \hat{\mathbf{x}} = R^\top Q^\top \mathbf{b} \iff R^\top R \hat{\mathbf{x}} = R^\top Q^\top \mathbf{b} \iff R \hat{\mathbf{x}} = Q^\top \mathbf{b}$ as (R invertible, $(R^\top)^{-1}$ exists.)

6.4 Pseudoinverse

D 6.4.1 Left Pseudoinverse (Full column rank): For $A \in \mathbb{R}^{m \times n}$ with full-column rank(A) = n , we get pseudoinverse $A^\dagger \in \mathbb{R}^{n \times m}$ as $A^\dagger = (A^\top A)^{-1} A^\top$ (Basically Least-Squares). $A^\dagger$ is a left inverse: $A^\dagger A = I$.

D 6.4.3 Right Pseudoinverse (Full row rank): For $A \in \mathbb{R}^{m \times n}$ with full row rank, rank(A) = m we get $A^\dagger \in \mathbb{R}^{n \times m}$ as $A^\dagger = A^\top (AA^\top)^{-1}$. $A^\dagger$ is a right inverse: $AA^\dagger = I$. Same as the "transpose of the pseudoinverse of the transpose". It picks the vector with smallest norm (from column space).

D 6.4.7 CR Decomposition with pseudoinverse: For $A \in \mathbb{R}^{m \times n}$ with rank(A) = r and a CR -decomposition $A = CR$, we define $A^\dagger = R^\dagger C^\dagger$. In general, $A^\dagger = R^\top (RR^\top)^{-1} (C^\top C)^{-1} C^\top = R^\top (C^\top CRR^\top)^{-1} C^\top = R^\top (C^\top AR^\top)^{-1} C^\top$.

L 6.4.8 Unique solution of least squares with pseudoinverse: For any matrix A and vector $\mathbf{x} \in C(A)$, then unique solution of the least squares problem is given by a vector $\hat{\mathbf{x}} \in C(A)$ satisfying $A\hat{\mathbf{x}} = \mathbf{b}$.

Note $\hat{\mathbf{x}} \in C(A^\top)$ as this minimises norm since any nullspace component is zeroed out. $\hat{\mathbf{x}}$ indeed in $C(A^\top)$ as $A^\dagger = R^\top \dots$ or $A^\dagger = A^\top \dots$

P 6.4.9 TS Decomposition: For $A \in \mathbb{R}^{m \times n}$ with rank(A) = r , let $S \in \mathbb{R}^{m \times r}$, $T \in \mathbb{R}^{r \times n}$ such that $A = ST$. Then, $A^\dagger = T^\dagger S^\dagger$.

T 6.4.10 Properties of Pseudoinverse: Let $A \in \mathbb{R}^{m \times n}$.

- $AA^\dagger A = A$ and $A^\dagger AA^\dagger = A^\dagger$ and $(A^\top)^\dagger = (A^\dagger)^\top$.
- $AA^\dagger$ is symmetric $((AA^\dagger)^\top = AA^\dagger)$, and the projection matrix for the projection on $C(A)$.
- $A^\dagger A$ is symmetric $((A^\dagger A)^\top = A^\dagger A)$, and the projection matrix for the projection on $C(A^\top)$.

$N(A^\top) = N(A^\dagger)$: Intuitively this holds as $A^\dagger \hat{\mathbf{x}} \in C(A^\top)$ (we pick the minimal $\hat{\mathbf{x}}$). Thus anything orthogonal to the subspace $C(A)^\perp = N(A^\top)$ is projected to 0: $\forall \mathbf{x} \in C(A)^\perp$ we have $A^\dagger \mathbf{x} = 0$. We conclude that $N(A^\top) = C(A)^\perp = N(A^\dagger)$.

Rank	Pseudo-inverse	Inverse property	Projection
Full column rank	$A^\dagger = (A^\top A)^{-1} A^\top$	Left inverse	$A^\dagger \mathbf{b} = \min_{\mathbf{x} \in \mathbb{R}^n} A\mathbf{x} - \mathbf{b} $
Full row rank	$A^\dagger = A^\top (AA^\top)^{-1}$	Right inverse	$\hat{\mathbf{x}} = A^\dagger \mathbf{b}$ is the solution to $A\mathbf{x} = \mathbf{b}$ with the smallest norm
General	$A^\dagger = R^\dagger C^\dagger = R^\top (C^\top AR^\top)^{-1} C^\top$	$A^\dagger AA^\dagger = A^\dagger$	$\hat{\mathbf{x}} = A^\dagger \mathbf{b}$ is the solution to $\min_{\mathbf{x} \in \mathbb{R}^n} \mathbf{x} $ s.t. $A^\top A \mathbf{x} = A^\top \mathbf{b}$

7 The Determinant

D 7.1.1 2×2 Determinant: For $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, $\det(A) = ad - bc$.

D 7.2.1 Permutation sign: The sign of a permutation is defined as the number of swaps of rows or columns. $\det(\text{permutation matrix}) = (-1)^k \det(\text{original matrix})$, where k is the number of swaps. Even number of swaps $\Rightarrow +1$, odd number $\Rightarrow -1$.

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

Parity of the *number of row swaps* necessary to get back to the identity.

(i) The sign of a permutation is multiplicative: $\text{sgn}(\sigma \circ \lambda) = \text{sgn}(\sigma) \cdot \text{sgn}(\lambda)$.
(ii) For all $n \geq 2$, exactly half of the permutations have sign 1 and the rest have sign -1 .

D 7.2.3 Determinant big formula: For a square matrix $A \in \mathbb{R}^{n \times n}$, $\det(A) = \sum_{\sigma \in \Pi_n} \text{sgn}(\sigma) \prod_{i=1}^n A_{i, \sigma(i)}$, (*Number of permutations: $n!$*)

Determinant Properties:

- Matrix $T \in \mathbb{R}^{n \times n}$ is triangular, then $\det(T) = \prod_{k=1}^n T_{kk}$, in particular $\det(I) = 1$.
- Matrix $A \in \mathbb{R}^{n \times n}$, $\det(A) = \det(A^\top)$
- Matrix $Q \in \mathbb{R}^{n \times n}$ is orthogonal $\iff \det(Q) = 1$ or $\det(Q) = -1$.
- Matrix $A \in \mathbb{R}^{n \times n}$ is invertible $\iff \det(A) \neq 0$
- Matrices $A, B \in \mathbb{R}^{n \times n}$, $\det(AB) = \det(A) \det(B)$, in particular $\det(A^n) = \det(A)^n$
- Matrix $A \in \mathbb{R}^{n \times n}$, $\det(A^{-1}) = \frac{1}{\det(A)}$
- $\det(\lambda A) = \lambda^n \det(A)$

P 7.2.4 Determinant of orthogonal matrices:

- Given a permutation matrix $P \in \mathbb{R}^{n \times n}$ corresponding to a permutation σ , then $\det(P) = \text{sgn}(\sigma)$. We sometimes also write $\text{sgn}(P)$.
- Given a triangular (either upper- or lower) matrix $T \in \mathbb{R}^{n \times n}$ we have $\det(T) = \prod_{k=1}^n T_{kk}$, in particular, $\det(I) = 1$.
- If $Q \in \mathbb{R}^{n \times n}$ is an orthogonal matrix then $\det(Q) = 1$ or $\det(Q) = -1$.

$1 = \det(I) = \det(Q^\top Q) = \det(Q^\top) \det(Q) = \det(Q)^2$, so $\det(Q) = \pm 1$. If the determinant is 1, then Q is a rotation matrix. If the determinant is -1 it's a reflection matrix.

P 7.3.2 Cofactor determinant calculation: Let $A \in \mathbb{R}^{n \times n}$, for any $1 \leq i \leq n$, $\det(A) = \sum_{j=1}^n A_{ij} C_{ij}$, where the co-factors are $C_{ij} = (-1)^{i+j} \det(\mathcal{A}_{ij})$, with $\mathcal{A}_{ij}$ being the $(n-1) \times (n-1)$ submatrix obtained by removing row i and column j from A .

The signs are $\begin{bmatrix} + & - & + \dots \\ - & + & - \dots \\ + & - & + \dots \end{bmatrix}$ Use sign of the current matrix not original matrix.

P 7.3.3: $A^{-1} = \frac{1}{\det(A)} C^\top$, C is the $n \times n$ matrix with the cofactors of A as entries.

P 7.3.5 Cramer's Rule: The idea here is that we solve a linear system of type $A\mathbf{x} = \mathbf{b}$, then, due to the determinant being multiplicative, we get $\det(A)\mathbf{x}_1 = \det(\mathcal{B})$, where $\mathcal{B}$ is the matrix obtained from A by replacing the first column with $\mathbf{b}$. So, the solution $\mathbf{x} \in \mathbb{R}^n$ for $A\mathbf{x} = \mathbf{b}$ is given by $\mathbf{x}_j = \frac{\det(\mathcal{B}_j)}{\det(A)}$

P 7.3.6 Swapping rows permutation matrix: If $A \in \mathbb{R}^{n \times n}$ and P is a permutation matrix that swaps two elements, meaning that PA corresponds to swapping two rows of A , then $\det(PA) = -\det(A)$.

P 7.3.7 Linearity of the determinant: The determinant is linear in each row (and column). For example:

$$\det \begin{bmatrix} \alpha_0 \mathbf{a}_0^\top + \alpha_1 \mathbf{a}_1^\top \\ \mathbf{a}_2^\top \end{bmatrix} = \alpha_0 \det \begin{bmatrix} \mathbf{a}_0^\top \\ \mathbf{a}_2^\top \end{bmatrix} + \alpha_1 \det \begin{bmatrix} \mathbf{a}_1^\top \\ \mathbf{a}_2^\top \end{bmatrix}$$

We can often use G-J-Elimination to make computation easier (adding / subtracting rows doesn't change the determinant). *Note that* we cannot exchange rows or multiply them without keeping track of it.

During the determinant calculation we have to look at all possible permutations and sum those. If $\sigma(i) \neq i$, i.e. we venture off the diagonal (identity permutation), we're going to have to venture off the diagonal for another j as well: $\sigma(j) \neq j$.

8 Eigenvalues and Eigenvectors

8.1 Complex Numbers

Operations: $i^2 = -1$:

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

Parts: $\Re(a + bi) := a$ (Real part), $\Im(a + bi) := b$ (imaginary part), $|z| := \sqrt{a^2 + b^2}$ (modulus), $a + bi := a - bi$ (complex conjugate)

Shortcuts:

- $|z|^2 = z\bar{z}$
- $z_1 \cdot z_2 = z_2 \cdot z_1$
- $z_1 + z_2 = \bar{z}_1 + \bar{z}_2$
- $\frac{1}{z} = \frac{\bar{z}}{|z|^2}$
- if $x = \bar{x}$ then $x \in \mathbb{R}$
- if $x_1 + ix_2 = y_1 + iy_2$ then we know $x_1 = y_1$ and $x_2 = y_2$

R 8.1.1 Euler's formula: For $\theta \in \mathbb{R}$, $e^{i\theta} = \cos \theta + i \sin \theta \Rightarrow e^{i\pi} = -1$
Polar form of a complex number: $z = r e^{i\theta}$, $z \in \mathbb{C}$, $r > 0$ is the modulus of z , $\theta \in [0, 2\pi]$.
T 8.1.2 Fundamental Theorem of Algebra: Any degree n non-constant ($n \geq 1$) polynomial $P(z) = \alpha_n z^n + a_{n-1}^{n-1} + \dots + \alpha_1 z + \alpha_0$, ($\alpha_n \neq 0$) has a zero: there exist $\lambda \in \mathbb{C}$ such that $P(\lambda) = 0$. A degree n polynomial has at most n distinct zeros (roots).
C 8.1.3 Algebraic multiplicity, number of 0: in polynomial Any degree n non-constant ($n \geq 1$) polynomial has n zeros $\lambda_1, \dots, \lambda_n \in \mathbb{C}$, and $P(z) = \alpha_n (z - \lambda_1)(z - \lambda_2) \dots (z - \lambda_n)$. The number of times $\lambda \in \mathbb{C}$ appears in the expression is called the <i>algebraic multiplicity</i> of zero.
Inner product on $\mathbb{C}^n$ and Conjugate Transpose: The inner product on $\mathbb{C}^n$ is given by $\langle v, w \rangle = w^* v$. $A^* = \bar{A}^\top = A^\top$
$\lambda \in \mathbb{C}$, $v \in \mathbb{C}^n$ we have $\ \lambda v \ ^2 = (\lambda v)^* (\lambda v) = (\bar{\lambda} v)^\top (\lambda v)$ where $\overline{\lambda v} = \bar{\lambda} \bar{v}$ thus at the end we get $\ \lambda v \ ^2 = \lambda ^2 v^* v$. This is because $\overline{\lambda \lambda} = \lambda ^2$ where that absolute value is the modulo for a complex number.

8.2 Introduction to Eigenvectors and Eigenvalues

D 8.2.1 Eigenvector / Eigenvalue pair: Given $A \in \mathbb{R}^{n \times n}$, we say $\lambda \in \mathbb{C}$ is an <i>eigenvalue</i> of A and $\mathbf{v} \in \mathbb{C}^n \setminus \{0\}$ is an <i>eigenvector</i> of A associated with λ when $A \mathbf{v} = \lambda \mathbf{v}$. $(\lambda, \mathbf{v})$ is an <i>eigenvalue-eigenvector</i> pair. If $\lambda \in \mathbb{R}$, then we have a real eigenvalue-eigenvector pair.
L 8.2.3 Real Eigenvalues / Eigenvectors: Let $A \in \mathbb{R}^{n \times n}$. Then, $\lambda \in \mathbb{R}$ is a real (A real EW always has a real EV), eigenvalue of A if and only if $\det(A - \lambda I) = 0$. A vector $\mathbf{v} \in \mathbb{R}^n \setminus \{0\}$ is an eigenvector associated with λ if and only if $\mathbf{v} \in N(A - \lambda I)$.
P 8.2.4 Characteristic polynomial: $(-1)^n \det(A - zI) = \det(zI - A) = (z - \lambda_1)(z - \lambda_2) \dots (z - \lambda_n)$. The coefficient of the λ^n term is $(-1)^n$. Usually determined from $\det(M - \lambda I)$.
T 8.2.5 Existence of eigenvalue: Every matrix $A \in \mathbb{R}^{n \times n}$ has an eigenvalue (perhaps complex-valued).
P 8.2.7 Eigenvalue of orthogonal matrix: If $Q \in \mathbb{R}^{n \times n}$ be an orthogonal matrix. If $\lambda \in \mathbb{C}$ is an eigenvalue of Q , then $ \lambda = 1$.
Even a real valued A can have complex eigenvalues. For the 90-degree rotation matrix (orthogonal), the EWs are $i, -i$.
L 8.2.8 Complex Eigenvalues exist on conjugate pairs of real A: Let $A \in \mathbb{R}^{n \times n}$. If $(\lambda, \mathbf{v})$ is an eigenvalue-eigenvector pair, then $(\bar{\lambda}, \bar{\mathbf{v}})$ is an eigenvalue-eigenvector pair.

8.3 Properties of Eigenvalues and Eigenvectors

P 8.3.1 Eigenvalue modifications based on the type of matrix: <ul style="list-style-type: none"> If $(\lambda, \mathbf{v})$ is an eigenvalue-eigenvector pair of A, then $(\lambda^k, \mathbf{v})$ is an eigenvalue-eigenvector pair of A^k for $k \geq 1$. If $(\lambda, \mathbf{v})$ is an eigenvalue-eigenvector pair of A with $\lambda \neq 0$, then $(\frac{1}{\lambda}, \mathbf{v})$ is an eigenvalue-eigenvector pair of A^{-1}.
L 8.3.2 Linear independence: If $\lambda_1, \dots, \lambda_n$ are all distinct, the corresponding eigenvectors $\mathbf{v}_1, \dots, \mathbf{v}_n$ are linearly independent.
T 8.3.3 Existence of basis from Eigenvalues: Let $A \in \mathbb{R}^{n \times n}$ with n distinct real eigenvalues. Then there exists a basis of $\mathbb{R}^n$, $\mathbf{v}_1, \dots, \mathbf{v}_n$ made of eigenvectors of A .
D 8.3.4 Trace of a matrix: The trace of A is defined by $\text{Tr}(A) = \sum_{i=1}^n A_{ii}$.
L 8.3.5 Transposition equality of Eigenvalues: The eigenvalues of $A \in \mathbb{R}^{n \times n}$ are the same as those of $A^\top$. (Not the EVs though).
L 8.3.6 Determinant and Trace via Eigenvalues: Let $A \in \mathbb{R}^{n \times n}$ and let $\lambda_1, \dots, \lambda_n$ be its eigenvalues as they appear in the characteristic polynomial. Then, $\det(A) = \prod_{i=1}^n \lambda_i$, $\text{Tr}(A) = \sum_{i=1}^n \lambda_i$.
L 8.3.7 Cyclic invariance of the trace: For $A, B, C \in \mathbb{R}^{n \times n}$: $\text{Tr}(AB) = \text{Tr}(BA)$, then $\text{Tr}(ABC) = \text{Tr}(BCA) = \text{Tr}(CAB)$.

(i) A and $A^\top$ share eigenvalues not eigenvectors . (ii) AB and BA always have the same set of real eigenvalues. (iii) Only if $AB = BA$ (commute) then EWs of AB are those of $A + B$ (iv) Gaussian-Elimination does not preserve EWs or EVs. (v) Real symmetric Matrix $\rightarrow$ real EWs. Real skew -symmetric Matrix $\rightarrow$ imaginary EWs. (vi) $\lambda + c$ is a real EW of $M + cI$. (Ex: $1 + \lambda^2$ are the EWs of $I + A^2$) (vii) A is invertible iff no eigenvalue is zero. (viii) Eigenvalues of A^k : λ_i^k for each eigenvalue λ_i . (ix) Eigenvalues of A^{-1} : $\frac{1}{\lambda_i}$ for each eigenvalue $\lambda_i \neq 0$.
--

EW $\lambda = 0 \iff A$ not invertible.
EWs of Diagonal: The EWs of a diagonal matrix are the values on the diagonal. A complete set of eigenvectors is the set $e_1, \dots, e_n$.
EWs of Triangular: The EWs of a triangular matrix are the values on the diagonal. However, not every triangular matrix has a complete set of eigenvectors, ex: $\begin{bmatrix} 1 & \\ 0 & 1 \end{bmatrix}$.

9 Diagonalization, Singular Value Decomposition

9.1 Diagonalization

T 9.1.1 Diagonalization Theorem, ability changing basis: $A = V \Lambda V^{-1}$, where V 's columns are its eigenvectors and Λ is a diagonal matrix with $\Lambda_{i,i} = \lambda_i$ and all other entries 0. $A \in \mathbb{R}^{n \times n}$ and has to have a complete set of real eigenvectors (Eigenbasis). Equivalently, $\Lambda = V^{-1} A V$, since V is invertible.
D 9.1.2 Diagonalizable matrix: A matrix $A \in \mathbb{R}^{n \times n}$ is called <i>diagonalizable</i> if there exists an invertible matrix V such that $V^{-1} A V = \Lambda$, where Λ is a diagonal matrix.
Being invertible does not guarantee diagonalisable (may not have complete set of eigenvectors). Being diagonalisable does not guarantee invertible (may have 0 EWs).
D 9.1.3 Complete set of Eigenvectors: If we can find eigenvectors forming a basis of $\mathbb{R}^n$ for A , we say that A has a <i>complete set of real eigenvectors</i> .
P 9.1.6 Eigenvalues and Eigenvectors of a projection matrix: Let P be the projection matrix on the subspace $U \subset \mathbb{R}^n$. Then P has two eigenvalues, 0 and 1, and a complete set of real eigenvectors.
D 9.1.7 Similar matrices: Matrices $A \in \mathbb{R}^{n \times n}$ and $B \in \mathbb{R}^{n \times n}$ are called <i>similar</i> if there exists and invertible matrix S , such that $B = S^{-1} A S$. If A similar to B then B also similar to A .
P 9.1.8 Similar matrices have the same eigenvalues: Similar matrices $A \in \mathbb{R}^{n \times n}$ and $B = S^{-1} A S \in \mathbb{R}^{n \times n}$ have the same eigenvalues. The matrix A has a complete set of real eigenvectors iff. B does. $A v = \lambda v \Leftrightarrow \lambda S^{-1} v = S^{-1} A v = S^{-1} A S S^{-1} v = B (S^{-1} v)$.
D 9.1.10 Geometric multiplicity: Let $A \in \mathbb{R}^{n \times n}$ and let λ be an eigenvalue of A . Then $\dim(\mathcal{N}(A - \lambda I))$ is called the <i>geometric multiplicity</i> of λ .
L 9.1.11 Complete set of real Eigenvectors: A matrix has a complete set of real eigenvectors iff. all its eigenvalues are real, and the geometric multiplicities equal the algebraic multiplicities for all eigenvalues.

9.2 Symmetric Matrices, Spectral Theorem

T 9.2.1 Spectral Theorem: Any symmetric matrix $A \in \mathbb{R}^{n \times n}$ has n real eigenvalues and an orthonormal basis of $\mathbb{R}^n$ consisting of eigenvectors A .
C 9.2.2 Eigendecomposition: For any symmetric matrix $A \in \mathbb{R}^{n \times n}$, there exists and 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 diagonal with diagonal entries equal to the eigenvalues of A , and $V^\top V = I$. This decomposition is called the <i>eigendecomposition</i> .
C 9.2.4 Rank of real symmetric matrices: <ul style="list-style-type: none"> If A is a real symmetric matrix, then $\text{rank}(A)$ is the number of non-zero eigenvalues of A (counting repetitions). For a general $n \times n$ matrix, $\text{rank}(A) = n - \dim(\mathcal{N}(A))$, so the geometric multiplicity of the eigenvalue $\lambda = 0$ equals $\dim(\mathcal{N}(A))$.

P 9.2.6 Rank-One Spectral Decomposition: Let $A \in \mathbb{R}^{n \times n}$ be symmetric, and let $\mathbf{v}_1, \dots, \mathbf{v}_n$ be an orthogonal basis of eigenvectors of A (the columns of V), with associated eigenvalues $\lambda_1, \dots, \lambda_n$. Then $A = \sum_{k=1}^n \lambda_k \mathbf{v}_k \mathbf{v}_k^\top$. <i>A real symmetric matrix is a weighted sum of orthogonal projections onto its eigenvector directions, with weights given by the eigenvalues.</i>
--

L 9.2.7 Orthogonality of Eigenvectors: 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 $\mathbf{v}_1, \mathbf{v}_2$. Then, $\mathbf{v}_1$ and $\mathbf{v}_2$ are orthogonal. $\lambda_1 \mathbf{v}_1^\top \mathbf{v}_2 = (A \mathbf{v}_1)^\top \mathbf{v}_2 = \mathbf{v}_1^\top A^\top \mathbf{v}_2 = \mathbf{v}_1^\top (A \mathbf{v}_2) = \lambda_2 \mathbf{v}_1^\top \mathbf{v}_2$ Thus $\mathbf{v}_1^\top \mathbf{v}_2$ must be 0.
L 9.2.8 Symmetric matrix has real Eigenvalues: A symmetric matrix $A \in \mathbb{R}^{n \times n}$ has only real eigenvalues $\lambda \in \mathbb{C} \Rightarrow \lambda \in \mathbb{R}$. If $A \mathbf{v} = \lambda \mathbf{v}$: $\lambda \ \mathbf{v} \ ^2 = \bar{\lambda} \mathbf{v}^* \mathbf{v} = (\lambda \mathbf{v})^* \mathbf{v} = (A \mathbf{v})^* \mathbf{v} = \mathbf{v}^* A^* \mathbf{v} = \mathbf{v}^* A \mathbf{v} = \mathbf{v}^* \lambda \mathbf{v} = \lambda \ \mathbf{v} \ ^2$ $\implies$ every symmetric matrix $A \in \mathbb{R}^{n \times n}$ has a real eigenvalue λ (C 9.2.9)

P 9.2.10 Rayleigh Quotient: Given a symmetric matrix $A \in \mathbb{R}^{n \times n}$, the Rayleigh Quotient, defined for $\mathbf{x} \in \mathbb{R}^n \setminus \{0\}$, as $R(\mathbf{x}) = \frac{\mathbf{x}^\top A \mathbf{x}}{\mathbf{x}^\top \mathbf{x}}$. The maximum of it is at $R(\mathbf{v}_{\max}) = \lambda_{\max}$ and the minimum correspondingly at the smallest eigenvalue, with λ and $\mathbf{v}$ being the respective minimum and maximum eigenvalue-eigenvector pairs.
--

D 9.2.11 Positive Semidefinite (PSD) and Positive definite (PD) matrices: A symmetric matrix $A \in \mathbb{R}^{n \times n}$ is said to be PSD if all its eigenvalues are non-negative. If all the eigenvalues of A are strictly positive (no eigenvalue is zero), then we say A is PD.
Any diagonal matrix (with > 0 diagonal entries) is always PD (it's invertible and symmetric).
If we take A, B PSD (or PD) then $A + B$ is also PSD (or PD).
PD always invertible: Any PD matrix is invertible since it has only non-zero eigenvalues.
P 9.2.12 Positivity of the quadratic form: A symmetric matrix $A \in \mathbb{R}^{n \times n}$ is PSD iff. $\mathbf{x}^\top A \mathbf{x} \geq 0$ for all $\mathbf{x} \in \mathbb{R}^n$. Analogously, its PD iff. $\mathbf{x}^\top A \mathbf{x} > 0$ for all $\mathbf{x} \in \mathbb{R}^n$.
D 9.2.13 Gram Matrix: Given vectors $\mathbf{x}_1, \dots, \mathbf{v}_n \in \mathbb{R}^n$, their <i>Gram matrix</i> is $G \in \mathbb{R}^{n \times n}$ defined by $G_{i,j} = \mathbf{v}_i^\top \mathbf{v}_j$. If $V = [\mathbf{v}_1 \dots \mathbf{v}_n] \in \mathbb{R}^{m \times n}$, then $G = V^\top V$. If $A = [\mathbf{a}_1 \dots \mathbf{a}_n] \in \mathbb{R}^{m \times n}$, one also calls $A A^\top$ a Gram matrix (although abuse of notation). Note, that $A A^\top = \sum_{i=1}^n \mathbf{a}_i \mathbf{a}_i^\top$, which will be an $m \times m$ matrix.
P 9.2.15 Same Eigenvalues of transposed matrices: Given a real matrix $A \in \mathbb{R}^{m \times n}$, the non-zero eigenvalues of $A^\top A \in \mathbb{R}^{n \times n}$ and $A A^\top \in \mathbb{R}^{m \times m}$ are the same. Also both are symmetric and PSD.
P 9.2.16 Cholesky Decomposition: Every symmetric PSD matrix M is a gram matrix of an upper triangular matrix C , so that $M = C^\top C$. Where $C = \Lambda^{\frac{1}{2}} V^\top$ and V is the matrix of eigenvectors.

9.3 Singular Value Decomposition

D 9.3.1 Singular Value Decomposition: Let $A \in \mathbb{R}^{m \times n}$. There exists orthogonal matrices $U \in \mathbb{R}^{m \times m}$ and $V \in \mathbb{R}^{n \times n}$ such that $A = U \Sigma V^\top$, where $\Sigma \in \mathbb{R}^{m \times n}$ is a diagonal matrix, in the sense that $\Sigma_{i,j} = 0$ when $i \neq j$, and the diagonal elements are non-negative and ordered in descending order. $U^\top U = I$ and $V^\top V = I$. The columns of U and V are called the left and right singular vectors of A , and the diagonal entries of Σ are the singular values of A , ordered as $\sigma_1 \geq \dots \geq \sigma_{\min\{m,n\}}$.
The singular values are the square-root of the eigenvalues of $A A^\top$ (or use $A^\top A$): $\sigma_i = \sqrt{\lambda_i}$.
R 9.3.2 Compact form of SVD: If A has rank r , then the SVD can be written as $A = U_r \Sigma_r V_r^\top$, where $U_r \in \mathbb{R}^{m \times r}$ and $V_r \in \mathbb{R}^{n \times r}$ have orthonormal columns, and $\Sigma_r \in \mathbb{R}^{r \times r}$. is a diagonal matrix with first r singular values. This representation stores $r(m + n + 1)$ real numbers instead of $m n$. For small r , this yields substantial savings and motivates low-rank approximations.

T 9.3.3 Every matrix has an SVD: Every matrix $A \in \mathbb{R}^{m \times n}$ has an SVD: $A = U \Sigma V^\top$. Equivalently, every linear transformation is diagonal in orthonormal bases of singular vectors and can be understood in 3 separate steps (the three composing matrices, $V^\top, \Sigma, U$).

P 9.3.4 SVD as a sum of rank-one matrices: Let $A \in \mathbb{R}^{m \times n}$ have rank r , with singular values $\sigma_1, \dots, \sigma_r$, and corresponding singular vectors $\mathbf{u}_1, \dots, \mathbf{u}_r$ and $\mathbf{v}_1, \dots, \mathbf{v}_r$.

Then $A = \sum_{k=1}^r \sigma_k \mathbf{u}_k \mathbf{v}_k^\top$. We can write any rank- r matrix $A \in \mathbb{R}^{m \times n}$ as a sum of r rank-1 matrices.

For $A = U \Sigma V^T$

$$A^T A = V \Sigma^2 V^T \quad \& \quad A A^T = U \Sigma^2 U^T \tag{1}$$
$$A^\dagger = V \Sigma^\dagger U^T \tag{2}$$
$$A^T = V \Sigma U^T = A^T A A^\dagger = A^\dagger A A^T \tag{3}$$

9.4 Applications

Inverse of A using SVD: A is invertible and has an SVD $A = U \Sigma V^\top$ then it's inverse can be easily computed using SVD as $A^{-1} = V \Sigma^{-1} U^\top$.
As A invertible all singular values are non-zero and thus Σ can be easily inverted (square diagonal matrix).

Pseudoinverse of A using SVD: Σ^+ is obtained from Σ by taking the reciprocal ($\frac{1}{\sigma_i}$) of each non-zero singular value, leaving the zeros in place, and transposing the matrix.
The pseudoinverse acts like an inverse in the "active" subspace (where singular values are nonzero) and projects to zero in the null space.

10 Strategies

10.1 Systems of Equations

General Solution:

- Form the augmented matrix $[A|\mathbf{b}]$.
- Perform **Gauss-Jordan Elimination** to get to RREF.
- Consistency Check:** If any row looks like $[0 \dots 0 | \text{non-zero}]$, there is **no solution**.

Calculate Inverse Matrix A^{-1} :

- Form the augmented matrix $[A|I]$.
- Perform **Gauss-Jordan Elimination**.
- 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).

Calculate CR^T Decomposition: Begin by computing the RREF of A to get R :

- C is the matrix obtained from the independent columns of A (take pivot columns from A)
- R' is R without any zero-rows.

Private Non-zero entry: The private non-zero argument (only vector in set that has that entry non-zero) is useful for proving linear independence.

Rank 1 Matrices: $\text{rank}(A) = 1$ equivalent to $\exists v \in \mathbb{R}^m \setminus \{\mathbf{0}\}, w \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ s.t. $A = v w^T$

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

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

Linear Dependence Equivalence:

- $v_i = \sum_{j \neq i} \lambda_j v_j$ for some $i \in [n]$
- $v_i = \sum_{j=1}^{i-1} \lambda_j v_j$ for some $i \in [n]$
- $Ax = 0$ for some $x \neq 0$
- A is not invertible

- $\det(A) = 0$
- Gauss elimination doesn't succeed

- 0 is an eigenvalue of A

- $Ax = b$ doesn't have a unique solution for all $b \in \mathbb{R}^n$
- n vectors in $\mathbb{R}^m$ with $n > m$ are dependent**

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 $n \times n$):**
 - Use row operations to convert A to an upper triangular matrix U .
 - 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 .
 - $\det(A) = (\text{corresponding factors}) \times \prod \mathbf{u}_{ii}$.
- Method D (Cofactors):** Choose the best cofactors (maximum number of 0s in row/column) and use the cofactor formula to get determinant.

Eigenvalues: Get the polynomial for $\det(A - \lambda I) = 0$ (set the diagonals to $A_{ii} - \lambda$) using determinant formula

- Factorise or use abc formula to find roots (the EWs)
- Check your results using the sum (**trace** formula) and the product (**determinant** formula).

Eigenvectors: Find the RREF R of $A - \lambda_i I$.

- Try to guess a vector, sometimes it's easy to see
- If that didn't work, find the nullspace $N(A - \lambda_i I)$ and choose any $v \neq 0$
- Check the result by multiplying with original A .

$x^\top x$ in sum form: Given $x = c_1 v_1 + c_2 v_2 + \dots + c_n v_n = \sum_{i=1}^n c_i v_i$

$$x^\top x = \left(\sum_{i=1}^n c_i v_i \right)^\top \left(\sum_{j=1}^n c_j v_j \right) = \left(\sum_{i=1}^n c_i v_i^\top \right) \left(\sum_{j=1}^n c_j v_j \right) = \sum_{i=1}^n \sum_{j=1}^n c_i c_j v_i^\top v_j$$

We can also expand the off-diagonal and diagonal seperately: $x^\top x = \sum_{i=1}^n c_i^2 v_i^\top v_i + 2 \sum_{i < j} c_i c_j v_i^\top v_j$

10.2 Fundamental Spaces

Basis for Column Space $C(A)$:

- Perform **Gaussian Elimination** on A to get R (no need for full RREF).
- Identify indices of the **pivot columns** in R (e.g. col 1, 3, 4).
- Result:** Select the corresponding columns from the **original** matrix A .

Basis for Row Space $R(A)$: $R(A) = C(A^\top)$

- Perform **Gaussian Elimination** on A to get R .
- Result:** The non-zero rows of R (transposed to be column vectors) from the basis.

Basis for Nullspace $N(A)$:

- Solve $Ax = 0$ using **Gauss-Jordan** to get RREF.
- Remove the zero-rows because $0^\top x = 0$
- Solve for $Rx = 0$
- Choose independent "special solutions" for the free variables.
- Compute the values for the others.
- Put together to get the nullspace basis.

Unique Decomposition into Row-Space and Nullspace: Because $C(A^\top) = N(A)^\perp$ we can decompose any $x \in \mathbb{R}^n$ into $x = x_r + x_n$ (unique $x_r \in R(A)$ and $x_n \in N(A)$). Then $Ax = Ax_r + Ax_n = Ax_r + 0$.

Quick Rank Reference: For an $m \times n$ matrix A with rank r :

- $\text{rank}(A) = \text{rank}(A^\top)$ (row rank = column rank)
- $\text{rank}(A) \leq \min(m, n)$ (limited by dimensions)
- $\text{rank}(A) = r \Rightarrow \dim(\mathbf{C}(A)) = r, \dim(\mathbf{R}(A)) = r$
- $\text{rank}(A) = r \Rightarrow r = n - \dim(\mathbf{N}(A))$ (nullity)
- $\text{rank}(A^\top) = r \Rightarrow r = m - \dim(\mathbf{N}(A^\top))$ (nullity)
- Full column rank: $\text{rank}(A) = n \Rightarrow$ columns are linearly independent, $\mathbf{N}(A) = \{0\}$
- Full row rank: $\text{rank}(A) = m \Rightarrow Ax = b$ is solvable for all b
- Full rank: $\text{rank}(A) = m = n \Rightarrow A$ is invertible

10.3 Orthogonality, Projections & Least Squares

Projection on 2d Plane: $P = I - nn^\top$ with n orthogonal to the plane. (Note that $I - 2nn^\top$ is then the reflection).

10.4 Eigenvalues & Decomposition

Find Eigenvalues and Eigenvectors:

- Eigenvalues** (λ): Solve characteristic equation $\det(A - \lambda I) = 0$.
- Eigenvectors** ($\mathbf{v}$): For each found λ :
 - Form matrix $(A - \lambda I)$
 - Find the Nullspace basis of this matrix (solve $(A - \lambda I)\mathbf{v} = 0$).

- Singular Value Decomposition (SVD): Goal:** $A = U \Sigma V^\top$.
- Compute $M = A^\top A$.
 - Find eigenvalues of M : $\lambda_1, \dots, \lambda_r$ (sorted high to low).
 - Singular Values:** $\sigma_i = \sqrt{\lambda_i}$, and place these in diagonal Σ .
 - Find Right Singular Vectors** (V): Calculate orthonormal vectors of $A^\top A$. These are columns of V .
 - Left Singular vectors** (U): For non-zero σ_i , calculate $\mathbf{u}_i = \frac{1}{\sigma_i} A \mathbf{v}_i$.
 - (if needed) compute U to be an orthonormal basis using Gram-Schmidt if A is not square / full-rank.

Recover A from a complete set of eigenvectors: Let V be the 3×3 matrix with columns $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$. Since the three vectors are orthonormal, V must be an orthogonal matrix, i.e. we have $V V^\top = V^\top V = I$.
We want to find the matrix A just from the information about its eigenvectors and eigenvalues. In particular, we know that $A \mathbf{v}_1 = \mathbf{v}_1$, $A \mathbf{v}_2 = -\mathbf{v}_2$, and $A \mathbf{v}_3 = 0$. We can summarize this in the equation $AV = VD$ where $D \in \mathbb{R}^{3 \times 3}$ is the diagonal matrix $D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

Multiplying from the right by $V^{-1} = V^\top$ yields $A = V D V^\top$

SVD for A^{-1} : For A invertible, the singular values are $1/\sigma$.

Diagonalisation with shared EW or EV:

- A, B same EVs: different Λ_A, Λ_B in diagonalisation
- A, B same EWs: **same** Λ different V, W matrices of eigenvectors.

Spell for Complete Set independent: n distinct real EWs $\implies$ all EVs are linearly independent $\implies$ complete set of real EVs $\implies$ forms a basis of $\mathbb{R}^n$.

B invertible then $Bv_1, \dots, Bv_n$ indep.: Since B is invertible, since $v_1, \dots, v_n$ linearly independent, $Bv_1, \dots, Bv_n$ are also linearly independent.

10.5 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 $|\lambda| = 1$. Preserves lengths ($\|Q\mathbf{x}\| = \|\mathbf{x}\|$)
- $\text{Tr}(A + B) = \text{Tr}(A) + \text{Tr}(B)$.
- $\text{Tr}(kA) = k \text{Tr}(A)$.
- $\text{Tr}(ABC) = \text{Tr}(BCA) = \text{Tr}(CAB)$ (Cyclic property).
- $\text{Tr}(A^\top A) = \sum_{i,j} a_{ij}^2$ (Sum of all squared entries).

10.6 Proof Toolkit (Standard Strategies)

How to prove U is a Subspace (D 4.8): To prove $U \subseteq V$ is a subspace:

- Check 1 (Zero):** Show $0 \in U$. (Usually easy, if fails $\rightarrow$ not a subspace).
- Check 2 (Closure):** Let $\mathbf{u}, \mathbf{v} \in U$ and $\lambda \in \mathbb{R}$. Show $\lambda \mathbf{u} \in U$ and $\mathbf{u} + \mathbf{w} \in U$.
Counter-example: To disprove, find specific vectors where closure fails or show $0 \notin U$.

How to prove Surjectivity / Injectivity: Let $T: V \rightarrow W$ be linear (matrix A).

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

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 $\mathbf{x} \neq 0$.

10.7 Spectral Theory & Properties

Algebraic vs Geometric Multiplicity: For eigenvalue λ :

- Alg. Mult. (n_a):** Number of times λ is root of $\det(A - \lambda I)$.
- Geo. Mult. (n_g):** $\dim(N(A - \lambda I))$ (Num. of independent eigenvectors).
- Property:** $1 \leq n_g \leq n_a$.
- Diagonalizable:** iff $\sum n_g = n$ (i.e. $n_g = n_a$ for all λ).

How to cast the Diagonalisation Spell: Since $\lambda_1, \dots, \lambda_n$ are distinct, we know from Lemma 8.3.2 that any n corresponding real eigenvectors $\mathbf{v}_1, \dots, \mathbf{v}_n \in \mathbb{R}^n$ of A are linearly independent. In other words, A has a complete set of real eigenvectors $\mathbf{v}_1, \dots, \mathbf{v}_n$. By assumption, B also has this complete set of real eigenvectors $\mathbf{v}_1, \dots, \mathbf{v}_n$. Let V be the matrix with columns $\mathbf{v}_1, \dots, \mathbf{v}_n$. From Theorem 9.1.1 it follows that $A = V \Lambda_A V^{-1}$ for some diagonal matrix $\Lambda_A \in \mathbb{R}^{n \times n}$. Analogously, it follows that $B = V \Lambda_B V^{-1}$ for some diagonal matrix $\Lambda_B \in \mathbb{R}^{n \times n}$. Now observe that since both Λ_A and Λ_B are diagonal matrices, we have $\Lambda_A \Lambda_B = \Lambda_B \Lambda_A$

10.8 Matrix Algebra Hacks

Block Matrices: If $M = \begin{bmatrix} A & B \\ 0 & D \end{bmatrix}$ (Block triangular):

- $\det(M) = \det(A) \cdot \det(D)$.
- Eigenvalues of M are eigenvalues of $A \cup$ eigenvalues of D .

Rank Properties:

- $\text{rank}(AB) \leq \min(\text{rank}(A), \text{rank}(B))$.
- $\text{rank}(A + B) \leq \text{rank}(A) + \text{rank}(B)$.

EV $\neq 0$: When working with eigenvalues in an exercise, especially things like $A(Bv) = \lambda(Bv)$ you have to consider the case where $Bv = 0$.
If $(Bv) = 0$ then Bv cannot be an eigenvector, as those are $\neq 0$. Use a case distinction.

11 Exercises

11.1 Calculations II

Find the closed form solution of $f_n = f_{n-1} + f_{n-2}$:

- First express the formula as a matrix equation: $\begin{pmatrix} f_n \\ f_{n-1} \end{pmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{pmatrix} f_{n-1} \\ f_{n-2} \end{pmatrix}$ where $M = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$.
- Find the EW-EV pairs of M which should form a basis of $\mathbb{R}^n$: The eigenvalues $\lambda_1 = \frac{1+\sqrt{5}}{2}$ (golden ratio ϕ) and $\lambda_2 = \frac{1-\sqrt{5}}{2}$ are found, along with their eigenvectors $v_1 = \begin{pmatrix} \lambda_1 \\ 1 \end{pmatrix}$ and $v_2 = \begin{pmatrix} \lambda_2 \\ 1 \end{pmatrix}$.
- Write the initial state $\mathbf{g}_1$ is written as a linear combination of eigenvectors with coefficients $\pm \frac{1}{\sqrt{5}}$:
 $g_0 = \frac{1}{\sqrt{5}}v_1 - \frac{1}{\sqrt{5}}v_2$.
- Write the formula in terms of EW-EVs: Since $M^n \mathbf{v}_i = \lambda_i^n \mathbf{v}_i$, we get the closed form: $F_n = \frac{1}{\sqrt{5}} \left[\left(\frac{1+\sqrt{5}}{2} \right)^n - \left(\frac{1-\sqrt{5}}{2} \right)^n \right]$.
- Watch out for the correct indexing! $g_n = \begin{pmatrix} f_n \\ f_{n-1} \end{pmatrix}$ then $g_n = M^{n-1}g_1$ where $g_1 = \begin{pmatrix} f_1 \\ f_0 \end{pmatrix}$.

Find matrix with specific EWs (not diagonal): Let A be the diagonal matrix with 0, 1, 2 on it's diagonals and let $V = \begin{bmatrix} e_1 & e_2 & e_3 \\ e_1 & e_2 & e_3 \end{bmatrix}$ then $B = VAV^{-1}$.

Note that V is a basis.

11.2 Multiple Choice

Use SVD and $A^T A$ to show Singular Values: We have $A = vv^T$ Then $A^T A = (vv^T)^T(vv^T) = ||v||^2 vv^T$. We show that w is an eigenvector using $A^T A w = ||v||^2 vv^T w = ||v||^2 ||w||^2 w$ with eigenvalue $\lambda = ||v||^2 ||w||^2$. We know the singular values are the square-root of the eigenvalues of $A^T A$ (or AA^T). As $A^T A$ has rank 1 (it's $||v||^2 vv^T$), this is the only eigenvalue. Thus the only and biggest singular value of A is $\sqrt{||v||^2 ||w||^2} = ||v|| ||w||$.

Rank of Pseudoinverse: Let $A \in \mathbb{R}^{7 \times 5}$ be arbitrary and $A^\dagger$ the pseudoinverse. We know $AA^\dagger$ is the projection matrix for projection onto $C(A)$. Thus $C(AA^\dagger) = C(A)$. Therefore $\dim(C(AA^\dagger)) = 7 - \dim(N(AA^\dagger))$. We also have $N(A^\dagger) \subseteq N(AA^\dagger)$ thus $\text{rank}(A) = \dim(C(AA^\dagger)) = n - \dim(N(AA^\dagger)) \leq n - \dim(N(A^\dagger)) = \text{rank}(A^\dagger)$ We show the other direction $\text{rank}(A^\dagger) \leq \text{rank}(A)$ in a symmetric argument. Thus $\text{rank}(A) = \text{rank}(A^\dagger)$.

Eigenvalues of AB and BA : We have $ABv = \lambda v$ for $v \neq 0$.

- If $Bv = 0$ then B does not have full rank thus BA also then BA also has a 0 EW.
- If $Bv \neq 0$ then $BA(Bv) = \lambda(Bv)$ thus λ an EW of BA .

Note however one of AB and BA may have more zero eigenvalues as their sizes might differ if they are $m \times n$ and $n \times m$.

11.3 Proofs

Proof of linear Transformation: Giving a matrix that models the transformation is valid. "Observe that we can rewrite $f(x) = Ax$. From the lecture, we know that any function of the form $f(x) = Ax$ is a linear transformation.

Prove $A = I + B$ invertible: Prove that for $A = I + B$ where B is a complete 1 matrix, we have A invertible:
If A is positive definite, then A is invertible. Observe that for all nonzero $x \in \mathbb{R}^n$: $x^T A x = x^T (I + B)x = x^T I x + x^T B x = ||x||^2 + (\sum_{i=1}^n x_i)^2 > 0$

Proofs: Assume first that $\mathbf{u}$ and $\mathbf{v}$ are orthogonal. From the lecture (Lemma 5.2.2), we know that this implies $\text{proj}_U(\mathbf{v}) = 0$ and $\text{proj}_V(\mathbf{v}) = 0$. We also know that projections are linear transformations and that $\text{proj}_V(\mathbf{v}) = \mathbf{v}$ and $\text{proj}_U(\mathbf{u}) = \mathbf{u}$. Let now $\mathbf{x} \in \text{Span}(\mathbf{u}, \mathbf{v})$ be arbitrary. By definition, $\mathbf{x}$ can be written as $\mathbf{x} = \alpha\mathbf{u} + \beta\mathbf{v}$ for some $\alpha, \beta \in \mathbb{R}$. We therefore get $\text{proj}_U(\mathbf{x}) + \text{proj}_V(\mathbf{x}) = \text{proj}_U(\alpha\mathbf{u} + \beta\mathbf{v}) + \text{proj}_V(\alpha\mathbf{u} + \beta\mathbf{v}) = \alpha(\text{proj}_U(\mathbf{u}) + \text{proj}_V(\mathbf{u})) + \beta(\text{proj}_U(\mathbf{v}) + \text{proj}_V(\mathbf{v})) = \alpha\mathbf{u} + \beta\mathbf{v}$

For the reverse direction, we give an indirect proof. For this, assume that $\mathbf{u}, \mathbf{v}$ are not orthogonal. This implies (by Lemma 5.2.2) that $\text{proj}_U(\mathbf{v}) \neq 0$. However, we still have $\text{proj}_V(\mathbf{v}) = \mathbf{v}$. We thus get $\text{proj}_U(\mathbf{v}) + \text{proj}_V(\mathbf{v}) \neq \mathbf{v}$ which means that the identity $\text{proj}_U(\mathbf{x}) + \text{proj}_V(\mathbf{x}) \neq \mathbf{x}$ does not hold for all $\mathbf{x} \in \text{Span}(\mathbf{u}, \mathbf{v})$, as desired.

Find Basis of Subspace: To prove that B_1, B_2, B_3 indeed form a basis of U , we first prove that they span all of U . To this end, let $A \in U$ be arbitrary with $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ Then A can be obtained from the linear combination $A = bB_1 + cB_2 + aB_3$ since by assumption we must have $\text{Tr}(A) = 0$ and thus $a = -d$. We conclude that B_1, B_2, B_3 indeed span all of U and it remains to show that they are linearly independent. Assume that there exist scalars $c_1, c_2, c_3 \in \mathbb{R}$ with $c_1 B_1 + c_2 B_2 + c_3 B_3 = 0$. Observe that $c_1 B_1 + c_2 B_2 + c_3 B_3 = \begin{bmatrix} c_3 & c_1 \\ c_2 & -c_3 \end{bmatrix}$ and thus we must have $c_1 = c_2 = c_3 = 0$. This means that there is no non-trivial linear combination of B_1, B_2, B_3 that yields 0. Therefore, the three matrices are linearly independent.

Certificate Proof: If we get told a system like $Ax = b$ is unsolvable, we now know that the system $A^T z = 0$ and $b^T z = 1$ is solvable. We use this to transform those two equations into $\begin{bmatrix} A^T \\ b^T \end{bmatrix} z = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ which we know is solvable. And by the Lemma 6.2.2 this has a unique solution

with $z \in R(\begin{bmatrix} A^T \\ b^T \end{bmatrix})$. As $R(\begin{bmatrix} A^T \\ b^T \end{bmatrix}) = C(\begin{bmatrix} A^T \\ b^T \end{bmatrix}^T) = C([A \quad b])$.

Show Vectors are independent: We know from the lecture that any three linearly independent vectors in $\mathbb{R}^3$ are a basis of $\mathbb{R}^3$. Hence, it remains to prove that $\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3$ are linearly independent. Consider the matrices

$$W := \begin{bmatrix} | & | & | \\ \mathbf{w}_1 & \mathbf{w}_2 & \mathbf{w}_3 \\ | & | & | \end{bmatrix}, \quad V := \begin{bmatrix} | & | & | \\ \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 \\ | & | & | \end{bmatrix}, \quad \text{and} \quad M := \begin{bmatrix} 1 & -1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}.$$

Observe that we have chosen M such that by definition of $\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3$, we have $W = VM$. Also observe that V has rank 3 since its columns span all of $\mathbb{R}^3$. Next, we compute the rank of M . From the lecture, we know that the rank of a matrix is equal to the number of pivots after using Gauss elimination on the matrix. We use this on M : subtracting the first row of M once from its second row, we get the triangular matrix

$$\begin{bmatrix} 1 & -1 & 1 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

which means that M has rank 3 as well. We conclude that both V and M have rank 3 (and are invertible by the Inverse Theorem). From the lecture, we know that the product of two square invertible matrices is invertible again. This implies that W has rank 3 as well which again implies that the columns of W are linearly independent (Inverse Theorem) which is what we wanted to prove.

Proof that scalars exist: Let $\mathbf{u}, \mathbf{v}, \mathbf{w}, \mathbf{z} \in \mathbb{R}^2$ be arbitrary with $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$, $\mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}$, $\mathbf{z} = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}$

Prove that there exist scalars $c_u, c_v, c_w, c_z \in \mathbb{R}$ with at least one of c_u, c_v, c_w, c_z being non-zero, such that $c_u \mathbf{u} + c_v \mathbf{v} + c_w \mathbf{w} + c_z \mathbf{z} = 0$ and $c_u \mathbf{u} + c_v \mathbf{v} + c_w \mathbf{w} + c_z \mathbf{z} = 0$.

Consider the four vectors $\mathbf{u}', \mathbf{v}', \mathbf{w}', \mathbf{z}' \in \mathbb{R}^3$ defined as

$$\mathbf{u}' := \begin{bmatrix} u_1 \\ u_2 \\ 1 \end{bmatrix}, \mathbf{v}' := \begin{bmatrix} v_1 \\ v_2 \\ 1 \end{bmatrix}, \mathbf{w}' := \begin{bmatrix} w_1 \\ w_2 \\ 1 \end{bmatrix}, \mathbf{z}' := \begin{bmatrix} z_1 \\ z_2 \\ 1 \end{bmatrix}.$$

Since these are four vectors living in $\mathbb{R}^3$, we know from the lecture that they need to be linearly dependent. This means that there exist scalars $c_u, c_v, c_w, c_z \in \mathbb{R}$ with $c_u \mathbf{u}' + c_v \mathbf{v}' + c_w \mathbf{w}' + c_z \mathbf{z}' = 0$, with at least one of c_u, c_v, c_w, c_z non-zero. In particular, the last coordinate of the equations $c_u \mathbf{u}' + c_v \mathbf{v}' + c_w \mathbf{w}' + c_z \mathbf{z}' = 0$ implies that $c_u + c_v + c_w + c_z = 0$. The first two coordinates imply $c_u \mathbf{u} + c_v \mathbf{v} + c_w \mathbf{w} + c_z \mathbf{z} = 0$. We conclude that the scalars c_u, c_v, c_w, c_z that we found satisfy all constraints.

PD proof: S, T two PD matrices and assume that exists $v \neq 0$ s.t. $STv = \lambda v$. Prove that $\lambda > 0$:

From the lecture, we know that a positive definite matrix $M \in \mathbb{R}^{n \times n}$ satisfies $\mathbf{x}^T M \mathbf{x} > 0$ for all $\mathbf{x} \in \mathbb{R}^n \setminus \{0\}$. Moreover, any positive definite matrix is invertible since 0 is not one of its eigenvalues. Consider now $\mathbf{v}$ and λ as given by the exercise. Observe that $T\mathbf{v} \neq 0$ since T is invertible and $\mathbf{v}$ is nonzero. Using positive definiteness of S , we hence get $(T\mathbf{v})^T S(T\mathbf{v}) > 0$. Moreover, by positive definiteness of T , we have $\mathbf{v}^T T\mathbf{v} > 0$. Recall that we are given $ST\mathbf{v} = \lambda \mathbf{v}$. Multiplying by $(T\mathbf{v})^T$ from the left, we get $(T\mathbf{v})^T S(T\mathbf{v}) = \lambda \mathbf{v}^T T\mathbf{v}$. Since both $(T\mathbf{v})^T S(T\mathbf{v})$ and $\mathbf{v}^T T\mathbf{v}$ are strictly positive, λ needs to be strictly positive as well.

Certificate: $Ax + By = c$ only has solution for every c iff $\mathbf{N}(A^T) \cap \mathbf{N}(B^T) \neq \{0\}$:
Consider the system (EQ) $Ax + By = c$ and write it as a single augmented matrix:

$$\begin{bmatrix} A & B \end{bmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = c. \text{ By contraposition, it is enough to show that the system (EQ) is not solvable}$$

for some $c \in \mathbb{R}^m$ if and only if $\mathbf{N}(A^T) \cap \mathbf{N}(B^T) \neq \{0\}$

' $\Leftarrow$ ': Suppose $\mathbf{N}(A^T) \cap \mathbf{N}(B^T) \neq \{0\}$, then there exists $z \neq 0$ satisfying

$$A^T z = 0 = B^T z \implies \begin{bmatrix} A^T \\ B^T \end{bmatrix} z = 0 \implies z^T \begin{bmatrix} A & B \end{bmatrix} = 0$$

We now pick $c = \frac{z}{\|z\|^2}$ which gives $c^T z = 1$. We just showed that for such right-hand side c , $\{z \in \mathbb{R}^m \mid z^T \begin{bmatrix} A & B \end{bmatrix} = 0, z^T c = 1\} \neq \emptyset$ so by Theorem 6.2.4 the system (EQ) is not solvable for the chosen c .
' $\Rightarrow$ ': Suppose there exists some $c \in \mathbb{R}^m$ for which (EQ) has no solution. Then, by Theorem 6.2.4, there exists a vector $z \in \mathbb{R}^m$ such that $z^T \begin{bmatrix} A & B \end{bmatrix} = 0$, $z^T c = 1$ Writing the first equality in block form gives $z^T A = 0$, $z^T B = 0$, $z^T c = 1$ The first two equalities mean $z \in \mathbf{N}(A^T) \cap \mathbf{N}(B^T)$, and the last one implies that $z \neq 0$. Therefore, $\mathbf{N}(A^T) \cap \mathbf{N}(B^T) \neq \{0\}$

Certificate Calculation: By Exercise 9.2 we know that the system of linear equations is solvable for all $c \in \mathbb{R}^3$ if and only if $\mathbf{N}(A^T) \cap \mathbf{N}(B^T) = \{0\}$ In particular we want to find $b_1, b_2, b_3 \in \mathbb{R}$ so that the only solution to $\begin{bmatrix} A^T \\ B^T \end{bmatrix} z = 0$ is $z = 0$. Plugging in the values for A and B we get

$$\begin{bmatrix} 2 & 1 & 0 \\ 1 & 3 & 2 \\ b_1 & b_2 & b_3 \end{bmatrix} \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} = 0$$

This gives the system

$$\begin{cases} 2z_1 + z_2 = 0, \\ z_1 + 3z_2 + 2z_3 = 0, \\ b_1 z_1 + b_2 z_2 + b_3 z_3 = 0 \end{cases}$$

From the first two equations we obtain $z_1 = -\frac{1}{2}z_2$, $z_3 = -\frac{5}{4}z_2$
Substituting these into the third equation gives

$$b_1 \left(-\frac{1}{2}z_2\right) + b_2 z_2 + b_3 \left(-\frac{5}{4}z_2\right) = 0 \implies \left(-\frac{1}{2}b_1 + b_2 - \frac{5}{4}b_3\right) z_2 = 0.$$

Therefore, a nontrivial solution $z \neq 0$ exists if and only if $-2b_1 + 4b_2 - 5b_3 = 0$. In conclusion, the system (EQ) is solvable for any $c \in \mathbb{R}^3$ for all values of $b_1, b_2, b_3 \in \mathbb{R}$ that satisfy $-2b_1 + 4b_2 - 5b_3 \neq 0$.

