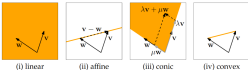


1 Vectors

D 1.4 Linear Combinations – $\sum_{i=1}^n \lambda_i \mathbf{v}_i$ scaled combinations of n vectors $\mathbf{v}_i$.

D 1.7 Combination Types –
(i) **Affine**: $\sum_{i=1}^n \lambda_i = 1$
(ii) **Conic**: $\lambda_j \geq 0$ for $j = 1, 2, \dots, n$
(iii) **Convex**: Both Affine and Conic.

D 1.9 Scalar/Dot Product –
 $\mathbf{v} \cdot \mathbf{w} = \sum_{i=1}^n v_i w_i$
Just multiply component wise, add all components together. Results are in $\mathbb{R}$.



D 1.11 Euclidean norm, squared norm, unit vector – $\|\mathbf{v}\| := \sqrt{\mathbf{v} \cdot \mathbf{v}} = \sqrt{\mathbf{v}^T \mathbf{v}}$. **Squared Norm**: $\|\mathbf{v}\|^2 := \mathbf{v}^T \mathbf{v}$. **Unit Vector**: $\|\mathbf{u}\| = 1 \Rightarrow \frac{\mathbf{u}}{\|\mathbf{u}\|} = \frac{\mathbf{v}}{\|\mathbf{v}\|}$ (for any vector $\mathbf{v} \neq 0$)

L 1.12 Cauchy-Schwarz inequality – $|\mathbf{v} \cdot \mathbf{w}| \leq \|\mathbf{v}\| \|\mathbf{w}\|$ for any two vectors $\mathbf{v}, \mathbf{w} \in \mathbb{R}^m$. Equally if $\mathbf{v} \lambda = \mathbf{w}$ or $\mathbf{w} \lambda = \mathbf{v}$.

D 1.14 Angles – $\cos(\alpha) = \frac{\mathbf{v} \cdot \mathbf{w}}{\|\mathbf{v}\| \|\mathbf{w}\|} \in [-1, 1]$ if $\mathbf{v} \perp \mathbf{w} \in \mathbb{R}^m$, then $\mathbf{v} \cdot \mathbf{w} = 0$.

D 1.16 Hyperplane through origin – Let $\mathbf{d} \in \mathbb{R}^m, \mathbf{d} \neq 0, H_{\mathbf{d}} = \{\mathbf{v} \in \mathbb{R}^m : \mathbf{v} \cdot \mathbf{d} = 0\}$

L 1.17 Triangle inequality – $\|\mathbf{v} + \mathbf{w}\| \leq \|\mathbf{v}\| + \|\mathbf{w}\|$

D 1.21 Linear (in)dependence – Vectors are linearly dependent if:
a) No vector is a linear combination of the others.
b) There are no λ_i , such that $\sum_{i=1}^n \lambda_i \mathbf{v}_i = 0$.
For matrices: $\neg \exists \mathbf{x} \neq 0 : A\mathbf{x} = 0$ (no column vector is a linear combination of the other)

L 1.22 Other definitions of linear dependence – Let $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n \in \mathbb{R}^m$. Statements are equivalent:
(i) At least one vector is a linear combination of the others.
(ii) There are scalars $\lambda_1, \lambda_2, \dots, \lambda_n$ besides $0, \dots, 0$, such that $\sum_{j=1}^n \lambda_j \mathbf{v}_j = 0$ (0 is a non-trivial linear combination of the vectors)
(iii) At least one of the vectors is a linear combination of the previous ones.

D 1.25 Span – Set of all linear combinations of a vector:
 $\text{Span}(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n) := \left\{ \sum_{j=1}^n \lambda_j \mathbf{v}_j : \lambda_j \in \mathbb{R} \text{ for all } j \in [n] \right\}$

L 1.26 – Given a set of vectors: $\mathbf{v}_1, \dots, \mathbf{v}_m \in \mathbb{R}^m$ and $\mathbf{v} \in \mathbb{R}^m$ be a linear combination of those vectors, then adding this combination does not change the span:
 $\text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_m) = \text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_m, \mathbf{v})$

Construction of vectors with standard unit vectors – Every vector in a vector space (D 1.4) can be written as: $\mathbf{u} = \sum_{i=1}^n \mathbf{u}_i \mathbf{e}_i$, where $\mathbf{e}_i$ is a standard unit vector (just one component being 1, all others 0).

L 1.28 Span of m linearly independent vectors is $\mathbb{R}^m$ – Let $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m \in \mathbb{R}^m$ be linearly independent. Then $\text{Span}(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m) = \mathbb{R}^m$.

2 Matrices

D 2.1 Matrix A – $A = [a_{ij}]_{i=1, j=1}^{m, n}$: m rows, n columns (first rows, then columns)

D 2.2 Matrix addition, scalar multiplication – Addition: $A + B = [a_{ij} + b_{ij}]_{i=1, j=1}^{m, n}$.
Scalar multiplication: $\lambda A = [\lambda a_{ij}]_{i=1, j=1}^{m, n}$

D 2.3 Square matrices – Identity matrix: quadratic matrix, diagonals 1, $A = AI = IA$;
Diagonal matrix: $a_{ij} = 0$ if $i \neq j$; Triangle matrix: lower if a_{ij} , upper else;
Symmetric matrix: $a_{ij} = a_{ji} \forall i, j, A^T = A$; Skew-symmetric matrix: $a_{ij} = -a_{ji} \forall i, j, A = -A^T$

D 2.4 Matrix-Vector Product – Rows of matrix $(m \times n)$ with vector (n) elements), i.e. $\mathbf{u}_1 = \sum_{i=1}^n a_{1,i} \mathbf{v}_i, i\mathbf{x} = \mathbf{x}$; Trace: Sum of the diagonal entries of a matrix.

O 2.5 – Let A be a $m \times n$ matrix. (i) A vector $\mathbf{b} \in \mathbb{R}^m$ is a linear combination of the columns of A iff. there's a vector $\mathbf{x} \in \mathbb{R}^n$ (of suitable scalars), such that $A\mathbf{x} = \mathbf{b}$. (ii) The columns of A are linearly independent iff. $\mathbf{x} = 0$ is the only vector such that $A\mathbf{x} = 0$.

D 2.9 Column space – $\mathbf{C}(A) := \{A\mathbf{x} : \mathbf{x} \in \mathbb{R}^n\}$, i.e.: the span (set of all linear combinations) of the column vectors.

D 2.10 Rank – $\text{rank}(A) :=$ the number of linearly independent column vectors, also sometimes called column rank.

L 2.11 Independent columns span the column space – Let A be an $m \times n$ matrix with r independent columns, and let C be the $m \times r$ submatrix containing the independent columns. Then $\mathbf{C}(A) = \mathbf{C}(C)$.

D 2.12 Transpose – Mirror the matrix along its diagonal. $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \leftrightarrow A^T = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}$

O 2.13 Transpose twice & transposing symmetric matrices – $(A^T)^T = A$ Moreover, a square matrix A is symmetric iff. $A = A^T$

D 2.14 Row Space – $\mathbf{R}(A) := \mathbf{C}(A^T)$. With the row rank of A , being the column rank of A^T .

D 2.17 Nullspace – Nullspace contains all input vectors that lead to output vector 0. $\mathbf{N}(A) = \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} = 0\}$. It can easily be obtained from RREF by setting $\mathbf{b} = 0$ in $A\mathbf{x} = \mathbf{b}$. If there are any free variables, choose any real (or complex) number satisfying the condition. To find the basis, rewrite and apply this Lemma: $R\mathbf{x} = 0 \Leftrightarrow I \cdot x(I) + Q \cdot x(Q) = 0$ e.g for $R = \begin{bmatrix} 1 & 2 & 0 & 3 \\ 0 & 0 & 1 & -2 \end{bmatrix} \rightarrow I[x_1, x_3] + \begin{bmatrix} 2 & 3 \\ 0 & -2 \end{bmatrix} [x_2] = 0 \Leftrightarrow x(I) = -Q \cdot x(Q)$. We may then freely choose to the free variables $x(Q)$, then find basis variables $x(I)$ using the above, typically choose $\mathbf{e}_1, \dots, \mathbf{e}_k$ for $x(Q)$ to get the k th vector of basis ($k =$ number of columns of Q). Finally, combine the vectors into one.

D 2.27 Kernel and Image –
• **Kernel**: $\mathbf{N}(A) = \text{Ker}(T) := \{\mathbf{x} \in \mathbb{R}^n : T(\mathbf{x}) = 0\} \subseteq \mathbb{R}^n$ (If A is the unique $m \times n$ matrix, such that $T = T_A$)
• **Image**: $\mathbf{C}(A) = \text{Im}(T) := \{T(\mathbf{x}) : \mathbf{x} \in \mathbb{R}^n\} \subseteq \mathbb{R}^m$ (If A is the unique $m \times n$ matrix, such that $T = T_A$), the set of all outputs that T can produce.

L 2.23 – A function $T : \mathbb{R}^n \rightarrow \mathbb{R}^m / T : \mathbb{R}^n \rightarrow \mathbb{R}$ is a linear transformation / linear functional iff. these two linearity axioms hold for all $\mathbf{x}, \mathbf{x}' \in \mathbb{R}^n$ and all $\lambda \in \mathbb{R}$:
(i) $T(\mathbf{x} + \mathbf{x}') = T(\mathbf{x}) + T(\mathbf{x}')$
(ii) $T(\lambda \mathbf{x}) = \lambda T(\mathbf{x})$

Visualizing linear transformations – A matrix can be seen as a re-mapping the unit-vectors $\hat{i}, \hat{j}, \hat{k}, \dots$, scaling and re-orienting them. Each column vector can be seen as the new unit vector $\mathbf{e}_i$. For example, $R(\theta) = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}$ would be a rotation matrix, that rotates the plane counterclockwise by θ .
To prove that a transformation T is linear, use Lemma 2.23. $A\mathbf{x} = \sum_{i=1}^n x_i \mathbf{v}_i$, where $\mathbf{v}_i$ is the i th column of A .

O 2.39 Matrix multiplication – $A \times B = C, c_{ij} = \sum_{k=1}^n a_{i,k} b_{k,j}$. Dimension restrictions: A is an $m \times n$ matrix, B is $n \times p$, the result C will be $m \times p$. For each entry, multiply the i th row of A with the j th column of B . This is NOT commutative, but associative & distributive. The right-to-left order matters, not the position of any parenthesis.

L 2.40 Matrix multiplication with transposition – $(AB)^T = B^T A^T, (AT)^T = A$

D 2.44 Outer product – $\text{rank}(A) = 1 \Leftrightarrow \exists$ non-zero vectors $\mathbf{v} \in \mathbb{R}^m, \mathbf{w} \in \mathbb{R}^n$ such that A is an outer product, i.e. $A = \mathbf{v}\mathbf{w}^T$, thus $\text{rank}(\mathbf{v}\mathbf{w}^T) = 1$.

T 2.46 CR Decomposition – $A = CR$. Get R from (reduced) row echelon form, C is the columns from A where there is a pivot in R . $C \in \mathbb{R}^{m \times r}, R \in \mathbb{R}^{r \times n}$ (in RREF), $r = \text{rank}(A)$.
To find REF try to create pivots: $R_0 = \begin{bmatrix} 1 & 0 & 2 & 3 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$. Use Gauss-Jordan elimination to find it.
RREF is simply REF without any zero rows. (i.e. in R_0 , R (in RREF would be R_0 without the last row).

3 Linear Equations

Get the system of linear equations into a $A\mathbf{x} = \mathbf{b}$ form, which can then be solved by Gaussian elimination. **Gaussian elimination**: Transform A into an upper-triangle matrix by performing row operations on it (into RREF). First, augment A by $\mathbf{b}$, then perform Gaussian elimination, and finally solve by back-substitution. Gaussian elimination succeeds iff. $A \in \mathbb{R}^{m \times m}$ and the columns of A are linearly independent (Runtime $O(m^3)$).

O 2.56 Invertible matrix – Matrix A is invertible, if it is square and there exists B , such that: $AB = I \Leftrightarrow BA = I \Leftrightarrow AB = BA = I$

D 2.57 Inverse matrix – If $AB = I$ for invertible A , then B is its inverse, denoted as A^{-1} .

O 2.58 Inverse of the Inverse – $(A^{-1})^{-1} = A$

L 2.59 – If A and B are invertible $m \times m$ matrices, AB is also invertible, and $(AB)^{-1} = B^{-1} A^{-1}$

L 2.60 – If A is an invertible $m \times m$ matrix, its transpose is also invertible, and $(A^T)^{-1} = (A^{-1})^T$

D 3.13 Reduced Row Echelon Form – Let $R = [r_{ij}]_{i=1, j=1}^{m, n}$ be an $m \times n$ matrix. R is in RREF, if there are some natural number $r \leq m$ and column indices $1 \leq j_1 \leq j_2 \leq \dots \leq j_r$ ("the indices of the 'downward step' such that the following two conditions hold":
(i) For every $i \in [r]$, column j_i of R is the standard unit vector $\mathbf{e}_{j_i}$.
(ii) All entries r_{ij} "below the staircase" are 0.

L 3.14 – A matrix R in RREF $(j_1, j_2, \dots, j_r)$ has independent columns $j_1, j_2, \dots, j_r$ and therefore rank r .

Gauss-Jordan elimination – Makes Gaussian elimination possible for $m \times n$ matrices and works similarly. Transform the augmented matrix $[A|\mathbf{b}]$ into RREF:
1. Swap rows, so the entry with the largest absolute value is the pivot a_{ij}
2. For each row, use the pivot to clear all entries below it using $R_i \leftarrow R_i - \left(\frac{\text{target}}{\text{pivot}}\right) R_{\text{pivot}}$.
3. Normalize all pivots to 1 by dividing the entire row by the pivot value.
4. Clear all entries above the pivots using row additions.
After reaching RREF, check the last row(s) $[0 \dots 0 | c]$:
• **No Solution**: $0 = c$ (where $c \neq 0$)
• **Unique Solution**: A square identity matrix on the left, the right side are the solved variables $\mathbf{b}_i$.
• **Infinite Solutions**: Row of zeroes $[0 \dots 0 | 0]$
A general RREF Structure:

$$\begin{bmatrix} 1 & 0 & \text{free} & b_1 \\ 0 & 1 & \text{free} & b_2 \\ 0 & 0 & 0 & b_3 \end{bmatrix}$$

4 Four fundamental Subspaces

4.1 Vector Spaces

D 4.1 Vector Space – Vector space is a triple $(V, +, \cdot)$ where V is a set of vectors, satisfying the vector space axioms, commutativity, associativity, existence of zero and negative vectors and identity element (1), compatibility of $\odot$ with $\cdot$ (in $\mathbb{R}$), distributivity over $\odot$ ($\lambda(\mathbf{v} + \mathbf{w}) = \lambda\mathbf{v} + \lambda\mathbf{w}$) and distributivity over $\cdot$ ($(\lambda\mu)\mathbf{v} = \lambda(\mu\mathbf{v}) = \lambda\mu\mathbf{v}$). To define a vector space, we need to define addition and scalar multiplication for the elements in a canonical way (according to the accepted standard).

D 4.8 Subspace – Let V be a vector space. A nonempty subset $U \subseteq V$ is a subspace of V if these two axioms are true $\forall \mathbf{v}, \mathbf{w} \in U$ and $\forall \lambda \mathbf{v} \in U$:

$$\mathbf{v} + \mathbf{w} \in U \quad \lambda \mathbf{v} \in U$$

They guarantee that vector addition and scalar multiplication doesn't take us outside the subspace.

L 4.9 Subspaces always has 0 – Let $U \subseteq V$ be a subspace of V . Then $0 \in U$.

L 4.11 Column space 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, $\mathbf{R}(A) = \mathbf{C}(A^T)$ is a subspace of $\mathbb{R}^n$.

E 4.13 Nullspace is a subspace – Let $A \in \mathbb{R}^{m \times n}$. Then the nullspace $\mathbf{N}(A) = \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{v} = 0\}$ 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 : $\text{Span}(B) = V$.

L.4.19 Independent columns is a basis – Set of vectors that are linearly independent and span B , the subspace of V . For $\mathbb{R}^m$, the set of unit vectors is a basis. For a matrix, all linearly independent columns form a basis of the column space $\mathbf{C}(A)$. **Calculating**: If we have a matrix with full column / row rank, then the basis are all column / row vectors.

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 $\mathbb{R}^m$.

D 4.21 Finitely generated vector space – A vector space V is called finitely generated if there exists a finite subset B of V with $\text{Span}(B) = V$. 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: $B, B' \subseteq V \Rightarrow |B| = |B'|$.

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 $\lambda_1, \lambda_2 \in \mathbb{R}$, $T(\lambda_1 \mathbf{x}_1 + \lambda_2 \mathbf{x}_2) = \lambda_1 T(\mathbf{x}_1) + \lambda_2 T(\mathbf{x}_2)$.

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 $\dim(\mathbf{R}(A)) = r$

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

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

L 4.35 Nullspace isomorphism – $R = \text{RREF}(A)$, then $T : N(R) \rightarrow \mathbb{R}^{n-r}$ is an isomorphism between $N(R)$ and $\mathbb{R}^{n-r} \Rightarrow \dim(N(R)) = n - r$.

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

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 $\dim(\text{Sol}(A, \mathbf{0})) := \dim(N(A))$

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**.
"Typical" matrices are with $m \leq n$ and have rank $r = m$.

5 Orthogonality and Projections

5.1 Definition

Orthogonality – A geometric and algebraic tool in order to be able to decompose a space into subspaces.

D 5.1.1 Orthogonal subspaces – Two vectors are orthogonal if their scalar product is 0 : $\mathbf{v}^T \mathbf{w} = \sum_{i=1}^n \mathbf{v}_i \mathbf{w}_i = 0$. 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 $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_j$ 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 $\{\mathbf{v}_1, \dots, \mathbf{v}_k, \mathbf{w}_1, \dots, \mathbf{w}_l\}$ are 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 $V \cap W = \{\mathbf{0}\}$ and their union is $V \cup W = \{\lambda \mathbf{v} + \mu \mathbf{w} : \lambda, \mu \in \mathbb{R}, \mathbf{v} \in V, \mathbf{w} \in W\}$. If $\dim(V) = k$ and $\dim(W) = l$, then $\dim(V \cup W) = k + l \leq n$, for $V, W \subseteq \mathbb{R}^n$.

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

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

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

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

5.2 Projections

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

L 5.2.2 One-dimensional projection formula – Projection of $\mathbf{b}$ on $S = \{\lambda \mathbf{a} | \lambda \in \mathbb{R}\} = C(\mathbf{a})$: $\text{proj}_S(\mathbf{b}) = \frac{\mathbf{a} \mathbf{a}^T \mathbf{b}}{\mathbf{a}^T \mathbf{a}}$, where $\mathbf{a} \in \mathbb{R}^m \setminus \{\mathbf{0}\}$. We note, that $(\mathbf{b} - \text{proj}_S(\mathbf{b})) \perp \text{proj}_S(\mathbf{b})$, i.e. the "error-vector" is perpendicular.

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 . Let A be the matrix with column vectors $\mathbf{a}_1, \dots, \mathbf{a}_n$. The general formula: $\text{proj}_S(\mathbf{b}) = A\mathbf{x}$, where $\mathbf{x} = A^T A \mathbf{x} = A^T \mathbf{b}$.

L 5.2.4 Properties of $A^T A - A - A^T A$ is invertible $\Leftrightarrow A$ has linearly independent columns. $\Rightarrow A^T A$ is a square matrix, symmetric, and invertible.

T 5.2.5 Projection in terms of projection matrix – $\text{proj}_S(\mathbf{b}) = P\mathbf{b}$ with projection matrix $P = A(A^T A)^{-1} A^T$. A is a matrix given in a task.

6 Applications of Orthogonality and Projections

6.1 Least Squares Approximation

Least Squares – Approximate a solution to a system of equations. Find $\mathbf{x}$ for which $A\mathbf{x}$ is as close as possible to $\mathbf{b}$: $\min_{\mathbf{x} \in \mathbb{R}^n} \|\mathbf{A}\mathbf{x} - \mathbf{b}\|^2$. Using the normal equations, we get $A^T A \hat{\mathbf{x}} = A^T \mathbf{b}$.

- Calculate $M = A^T A$
- Calculate $\mathbf{b}' = A^T \mathbf{b}$
- Solve resulting System of Equations $M\hat{\mathbf{x}} = \mathbf{b}'$ as usual.

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 $A = \begin{bmatrix} 1 & t_1 \\ \vdots & \vdots \\ 1 & t_n \end{bmatrix}$ and a result vector $\mathbf{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$ where t_i is the total number of data points and t_i is the slope of the i th function, where b_i is its output. The first column is all 1's because the constant element has no scalar. This comes from the following concept: $f(t) = \alpha_0 + \alpha_1 t$, so if the first data point is $(1, 2)$, we get $\alpha_1 \cdot 1 = 2$, which will then transform into a SLE with other equations.

L 6.1.2 – If A has linearly dependent columns, $t_i = t_j, \forall i \neq j$.

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

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

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 $\mathbf{x}_1 \in C(A^T A)$ 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:

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

6.3 Orthogonal Bases and Gram Schmidt

D 6.3.1 Orthogonal vectors – $\mathbf{q}_i^T \mathbf{q}_j = \delta_{ij} = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}$, with δ_{ij} being the Kronecker delta.

D 6.3.2 Orthogonal matrix – A square matrix $Q \in \mathbb{R}^{n \times m}$ is an *orthogonal matrix* when $Q^T Q = I$. If it is square, then $QQ^T = I, Q^{-1} = Q^T$, 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 $(Q\mathbf{x})^T (Q\mathbf{y}) = \mathbf{x}^T \mathbf{y}$

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 $\hat{\mathbf{x}} = Q^{-1} \mathbf{b}$ and the projection matrix is given by QQ^T .

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

D 6.3.10 QR Decomposition – $A = QR$, where $R = Q^T A$ 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, $Q Q^T A = A$, 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 $\text{proj}_{C(A)}(\mathbf{b}) = \text{proj}_{C(Q)}(\mathbf{b})$ and for least squares, we have $R\hat{\mathbf{x}} = Q^T \mathbf{b}$. This can efficiently be solved using back-substitution because R is triangular.

6.4 Pseudoinverse

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

D 6.4.7 CR Decomposition with pseudoinverses – For $A \in \mathbb{R}^{m \times n}$ with $\text{rank}(A) = r$ and a CR-decomposition $A = CR$, we define $A^\dagger = R^T C^T$. In general, $A^\dagger = R^T (C^T C)^{-1} = R^T (C^T C R R^T)^{-1} C^T = R^T (C^T A R^T)^{-1} C^T$.

L 6.4.8 Unique solution of least squares with pseudoinverses – 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 $A\hat{\mathbf{x}} = \mathbf{b}$. The solution is $\hat{\mathbf{x}} = A^\dagger \mathbf{b}$, with $A\hat{\mathbf{x}} = \mathbf{b}$, and in the general case $A^\dagger = R^T C^T = R^T (C^T A R^T)^{-1} C^T$.

P 6.4.9 TS Decomposition – For $A \in \mathbb{R}^{m \times n}$ with $\text{rank}(A) = r$, 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 Pseudoinverses – Let $A \in \mathbb{R}^{m \times n}$.

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

Moreover, $AA^\dagger = CRR^T(RR^T)^{-1}(C^T C)^{-1}C^T = C(C^T C)^{-1}C^T$, which is the projection onto $C(A)$, and $(AA^\dagger)^T = AA^\dagger$.

7 The Determinant

The determinant can be understood as a number that corresponds to *how much* the associated linear transformation scales space. For example, a 2D linear transformation with a determinant 2, will scale any area in the space up by 2 after the linear transformation has been applied.

7.1.2 times 2

D 7.1.1.2 2×2 Determinant – For $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, $\det = ad - bc$.

L 7.1.2 Multiplication of determinants – $\det(AB) = \det(A) \cdot \det(B)$

L 7.1.3 Invertibility related to the determinant – A matrix $A \in \mathbb{R}^{2 \times 2}$ is invertible iff. $\det(A) \neq 0$.

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 $\Rightarrow -1$.

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

7.2 General case

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 $A \in \mathbb{R}^{n \times n}$, $\det(A) = \det(A^T)$
- 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 $P \in \mathbb{R}^{m \times m}$ 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 $\det(Q) = -1$.
 $1 = \det(I) = \det(Q^T Q) = \det(Q^T) \det(Q) = \det(Q)^2$, 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 $C_{ij} = (-1)^{i+j} \det(A_{ij})$

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:

$$\det \begin{bmatrix} \alpha_1 \mathbf{a}_1^T & \alpha_1 \mathbf{a}_1^T \\ & \mathbf{a}_2^T \end{bmatrix} = \alpha_1 \det \begin{bmatrix} \mathbf{a}_1^T \\ \mathbf{a}_2^T \end{bmatrix} + \alpha_1 \det \begin{bmatrix} \mathbf{a}_1^T \\ \mathbf{a}_2^T \end{bmatrix}$$

8 Eigenvalues and Eigenvectors

8.1 Complex Numbers

Operations – $i^2 = -1$ (NOT $i = \sqrt{-1}$, since otherwise $1 = -1$). Complex number $z_j = a_j + b_k i$. Addition, Subtraction $(a_1 \pm a_2) + (b_1 \pm b_2)i$. Multiplication $(a_1 a_2 - b_1 b_2) + (a_1 b_2 + a_2 b_1)i$. Division $\frac{a_1 + a_2 i + b_1 i + b_2 i^2}{b_1^2 + b_2^2} + \frac{a_2 b_1 - a_1 b_2 i}{b_1^2 + b_2^2}$

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)

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 – $z = re^{i\theta}$, $z \in \mathbb{C}$, $r > 0$ is the modulus of z , $\theta \in [0, 2\pi]$.

T 8.2.1 Fundamental Theorem of Algebra – Any degree n non-constant ($n \geq 1$) polynomial $P(z) = \alpha_n z^n + \alpha_{n-1} z^{n-1} + \dots + \alpha_1 z + \alpha_0$, ($\alpha_n \neq 0$) 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 *algebraic multiplicity* of zero.

Inner product on $\mathbb{C}^n$ and Conjugate Transpose – The inner product on $\mathbb{C}^n$ is given by $\langle \mathbf{v}, \mathbf{w} \rangle = \mathbf{w}^* \mathbf{v}$.
 $A^* = A^T$

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 eigenvalue of A and $\mathbf{v} \in \mathbb{C}^n \setminus \{0\}$ is an eigenvector of A associated with λ when $A\mathbf{v} = \lambda \mathbf{v}$. $(\lambda, \mathbf{v})$ is an eigenvalue-eigenvector pair. If $\lambda \in \mathbb{R}$, then we have a real eigenvalue-eigenvector pair. *Imagine the eigenvectors to be the normalized vectors that don't change when applying a linear transformation.*

L 8.2.3 Real Eigenvalues / Eigenvectors – Let $A \in \mathbb{R}^{n \times n}$. Then, $\lambda \in \mathbb{R}$ is a real 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 $\mathbf{v} \in \mathcal{N}(A - \lambda I)$.

To find an Eigenvalue and Eigenvector of a matrix $M \in \mathbb{R}^{n \times n}$, simply calculate the eigenvalue first, using the zeros of the polynomial obtained from calculating $\det(M - \lambda I)$, which is obtained from **L 8.2.3** $\det(M - \lambda I) = 0$. This means, we simply need to calculate the determinant of $M - \lambda I$, which is fairly straightforward. We can then try to find the eigenvectors $\mathbf{v}$ such that $M\mathbf{v} = \lambda \mathbf{v}$, or in other words a non-zero element of $N(M - \lambda I) \setminus \{0\}$, i.e. the null space of $M - \lambda I$. This means we try to find a solution such that $0 = (m - \lambda I)\mathbf{v}$, where $\mathbf{v}$ is not the zero vector.

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 $|\lambda| = 1$.

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 –

- 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^T .

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 $\text{Tr}(ABC) = \text{Tr}(BCA) = \text{Tr}(CAB)$.

Change of basis – With the linear transformation $L: \mathbb{R}^n \rightarrow \mathbb{R}^m$ (given by e.g. $\mathbf{x} \in \mathbb{R}^n \mapsto A\mathbf{x} \in \mathbb{R}^m$), for which we want to find a matrix B that maps it from a basis $\mathbf{v}_1, \dots, \mathbf{v}_n$ to another one $\mathbf{w}_1, \dots, \mathbf{w}_n$. Now that B helps us map a vector α to a vector β , which has the different basis. We now define U as the matrix whose columns are the B is first basis and V as the matrix whose columns are the second basis. Now, if $L(\mathbf{x}) = V\beta$ and $\mathbf{x} = U\alpha$, so $\beta = V^{-1}AU\alpha$, now $\beta = V^{-1}AU$.

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 $\Lambda_{ii} = \lambda_i$ and all other entries 0. $A \in \mathbb{R}^{n \times n}$ and has to have a complete set of real eigenvectors (Eigenbasis). Equivalently, $\Lambda = V^{-1}AV$, since V is invertible.

D 9.1.2 Diagonalizable matrix – 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.

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 *complete set of real eigenvectors*.

P 9.1.6 Eigenvalues and Eigenvectors of a projection matrix – Let P be the projection matrix on the subspace $U \subseteq \mathbb{R}^n$. Then

$$P^2 = P$$

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 *similar* if there exists and invertible matrix S , such that $B = S^{-1}AS$.

P 9.1.8 Similar matrices have the same eigenvalues – Similar matrices $A \in \mathbb{R}^{n \times n}$ and $B = S^{-1}AS \in \mathbb{R}^{n \times n}$ have the same eigenvalues. The matrix A has a complete set of real eigenvectors iff. B does.

D 9.1.10 Geometric multiplicity – Let $A \in \mathbb{R}^{n \times n}$ and let λ be an eigenvalue of A . Then $\dim(N(A - \lambda I))$ is called the *geometric multiplicity* 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 orthogonal 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 $A = V\Lambda V^T$, where $\Lambda \in \mathbb{R}^{n \times n}$ is diagonal with diagonal entries equal to the eigenvalues of A , and $V^T V = I$. This decomposition is called the *eigendecomposition*.

C 9.2.4 Rank of real symmetric matrices –
• 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, $\text{rank}(A) = n - \dim(N(A))$, so the geometric multiplicity of the eigenvalue $\lambda = 0$ equals $\dim(N(A))$.

R 9.2.5 – For general $n \times n$ (non-symmetric) matrices, the rank is n minus the dimension of the nullspace, so it is n minus the geometric multiplicity of $\lambda = 0$. Since symmetric matrices always have a complete set of eigenvalues and eigenvectors, the geometric multiplicities are always the same as algebraic multiplicities.

$$\dim(N(A)) + \text{rank}(A) = n$$

P 9.2.6 Rank-One Spectral Decomposition – Let $A \in \mathbb{R}^{m \times n}$ 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 $\lambda_1, \dots, \lambda_n$.

$$\text{Then } A = \sum_{k=1}^n \lambda_k \mathbf{v}_k \mathbf{v}_k^T.$$

A real symmetric matrix is a weighted sum of orthogonal projections onto its eigenvector directions, with weights given by the eigenvalues.

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.

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 $A\mathbf{v} = \lambda\mathbf{v}$:

$$\lambda \|\mathbf{v}\|^2 = \lambda \mathbf{v}^T \mathbf{v} = (\lambda \mathbf{v})^T \mathbf{v} = (A\mathbf{v})^T \mathbf{v} = \mathbf{v}^T A^T \mathbf{v} = \mathbf{v}^T A \mathbf{v} = \mathbf{v}^T \lambda \mathbf{v} = \lambda \|\mathbf{v}\|^2$$

$\Rightarrow$ 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 $R(\mathbf{x}) = \frac{\mathbf{x}^T A \mathbf{x}}{\mathbf{x}^T \mathbf{x}}$. 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.

P 9.2.12 Positivity of the quadratic form – A symmetric matrix $A \in \mathbb{R}^{n \times n}$ is PSD iff. $\mathbf{x}^T A \mathbf{x} \geq 0$ for all $\mathbf{x} \in \mathbb{R}^n$. Analogously, its PD iff. $\mathbf{x}^T A \mathbf{x} > 0$ for all $\mathbf{x} \in \mathbb{R}^n$.

D 9.2.13 Gram Matrix – Given vectors $\mathbf{x}_1, \dots, \mathbf{x}_n \in \mathbb{R}^n$, their *Gram matrix* is $G \in \mathbb{R}^{n \times n}$ defined by $G_{ij} = \mathbf{v}_i^T \mathbf{v}_j$. We have $i, j \leq n$ because $G \in \mathbb{R}^{n \times n}$. If $V = [\mathbf{v}_1 \dots \mathbf{v}_n] \in \mathbb{R}^{m \times n}$, then $G = V^T V$. If $A = [\mathbf{a}_1 \dots \mathbf{a}_n] \in \mathbb{R}^{m \times n}$, one also calls AA^T a Gram matrix (although abuse of notation). Note, that $AA^T = \sum_{i=1}^n \mathbf{a}_i \mathbf{a}_i^T$, which will be an $n \times n$ 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 $A^T A \in \mathbb{R}^{n \times n}$ and $AA^T \in \mathbb{R}^{m \times m}$ 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 $M = C^T C$.

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 $U^T A = U^T \Sigma V^T$, where $\Sigma \in \mathbb{R}^{m \times n}$ is a diagonal matrix, in the sense that $\Sigma_{ij} = 0$ when $i \neq j$, and the diagonal elements are non-negative and ordered in descending order. $U^T U = I$ and $V^T V = I$. 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 $\sigma_1 \geq \dots \geq \sigma_{\min\{m,n\}}$.

R 9.3.2 Compact form of SVD – If A has rank r , then the SVD can be written as $A = U_r \Sigma_r V_r^T$, 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 mn . 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: $A = U\Sigma V^T$. 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^T, Σ, 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

$$A = \sum_{k=1}^r \sigma_k \mathbf{u}_k \mathbf{v}_k^T$$

We can write any rank- r matrix $A \in \mathbb{R}^{m \times n}$ as a sum of r rank-1 matrices.

10 Strategies

10.1 Systems of Equations

General Solution –

- Form the augmented matrix $[A|b]$.
- Perform **Gauss-Jordan Elimination** to get to RREF.
- Consistency Check**: If any row looks like $[0 \dots 0]non-zero$, there is **no solution**.
- Identify variables: **Pivot variables**: Columns with leading 1s; **Free variables**: Columns without leading 1s.
- General Solution**: $\mathbf{x} = \mathbf{x}_p + \mathbf{x}_1 \mathbf{s}_1 + \mathbf{x}_2 \mathbf{s}_2$: Set free variables to 0, solve for pivots; $\mathbf{x}_h$: Write pivot variables in terms of free variables. Extract free variables as coefficients.

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).

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*.

Calculating the Determinant –

- Method A** (2×2): $\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc$.
- 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 ($kR_i - kR_j$): det does not change; Scalar multiplication (kR_i): multiply det by k .
- $\det(A) = (\text{corresponding factors}) \times \prod \mathbf{u}_{ii}$.

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 $R(A) = R(A) = C(A^T)$

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

Basis for Nullspace $N(A)$ –

- Solve $A\mathbf{x} = 0$ using **Gauss-Jordan** to get RREF.
- Express pivot variables in terms of free variables.
- Result**: The vectors multiplying the free variables form the basis.
- Dimension**: $\dim(N(A)) = n - r$ (Columns minus Rank).

10.3 Orthogonality, Projections & Least Squares

Least Squares Approximation – Problem: $A\mathbf{x} = \mathbf{b}$ has no solution ($m > n$). Find $\hat{\mathbf{x}}$ that minimizes $\|A\mathbf{x} - \mathbf{b}\|^2$.

- Calculate matrix $M = A^T A$.
 - Calculate vector $d = A^T \mathbf{b}$.
 - Solve the system $M\hat{\mathbf{x}} = d$ (using Gaussian elimination).
- Note: If columns of A are independent, $\hat{\mathbf{x}} = (A^T A)^{-1} A^T \mathbf{b}$.*

Projection of $\mathbf{b}$ onto Subspace S –

- Find a basis for S and put them as columns in matrix A .
- Calculate $\hat{\mathbf{x}}$ using *Least Squares*.
- Result**: The projection $p = A\hat{\mathbf{x}}$.
- Projection Matrix**: $P = A(A^T A)^{-1} A^T$.

- **Input:** Independent vectors $\mathbf{a}_1, \dots, \mathbf{a}_r$.
- **Output:** Orthonormal vectors $\mathbf{q}_1, \dots, \mathbf{q}_r$.
- Follow **Definition 6.3.8**

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

1. Find eigenvalues $\lambda_1, \dots, \lambda_n$.
2. Find n independent eigenvectors $\mathbf{v}_1, \dots, \mathbf{v}_n$ (If you cannot find n independent vectors, A is not diagonalizable).
3. **Construct Matrices:**
 - Λ : Diagonal matrix with λ_i on diagonal.
 - V : Matrix with eigenvectors $\mathbf{v}_i$ as columns (order must match λ_i).

1. Eigenvalues will be *real*
2. Eigenvectors for different λ are automatically orthogonal.
3. **Important:** Normalize the eigenvectors to length 1.
4. **Result:** $A = Q\Lambda Q^T$ (where Q is orthogonal matrix of normalized eigenvectors).

1. Compute $M = A^T A$.
2. Find eigenvalues of M : $\lambda_1, \dots, \lambda_r$ (sorted high to low).
3. **Singular Values:** $\sigma_i = \sqrt{\lambda_i}$, and place these in diagonal Σ .
4. **Find Singular Vectors (V):** Calculate orthonormal vectors of $A^T A$. These are columns of V .
5. **Left Singular vectors (U):** For non-zero σ_i , calculate $u_i = \frac{1}{\sigma_i} A v_i$.
6. (if needed) compute U to be an orthonormal basis using Gram-Schmidt if A is not square / full-rank.

- **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 : $Q^T Q = I$. Determinant is ± 1 . Preserves lengths ($\|Q\mathbf{x}\| = \|\mathbf{x}\|$)

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 $u, v \in U$ and $\lambda \in \mathbb{R}$. Show $\lambda u + v \in U$.

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

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).

- **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 = $\dim(W)$).
- **Bijjective:** Show both (or if $\dim(V) = \dim(W)$, just one is enough).

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

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):

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

Change of Basis – Let $B_{old} = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ and $B_{new} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$. Let T be a transformation with matrix A in standard basis.

- Matrix of T relative to B_{new} is: $D = V^{-1}AV$
- Where $V = [\mathbf{v}_1 \dots \mathbf{v}_n]$ (Cols are basis vectors).
- If B_{new} are eigenvectors, $D = \Lambda$ (Diagonal).

- 1. Right Singular Vectors (V):** Compute $M = A^T A$. 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^T).

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).

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 $a \pm bi$).

1. All eigenvalues $\lambda_i > 0$.
2. $\mathbf{x}^T \mathbf{A} \mathbf{x} > 0$ for all $\mathbf{x} \neq \mathbf{0}$.
3. All pivots (from Gaussian Elim without swap) are > 0 .
4. **Sylvester's Criterion:** All upper-left sub-determinants > 0 .

Note: For Positive Semidefinite, replace $>$ with $\geq$.

- Product: $E_{ij}E_{kl} = (\mathbf{e}_i\mathbf{e}_j^\top)(\mathbf{e}_k\mathbf{e}_l^\top) = \mathbf{e}_i(\mathbf{e}_j^\top\mathbf{e}_k)\mathbf{e}_l^\top$.
- $\mathbf{e}_j^\top\mathbf{e}_k = \delta_{jk}$ (1 if $j = k$, else 0).
- So $E_{ij}E_{kl} = \delta_{jk}E_{il}$. (Zero unless "inner indices" match).

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

- $\text{rank}(AB) \leq \min(\text{rank}(A), \text{rank}(B))$.
- $\text{rank}(A + B) \leq \text{rank}(A) + \text{rank}(B)$.
- **Sylvester:** $\text{rank}(A) + \text{rank}(B) - n \leq \text{rank}(AB)$ (where $A : m \times n, B : n \times k$).

$$= (z - \lambda_1)(z - \lambda_2)(z - \lambda_3) \dots (z - \lambda_n)$$

Strategies:

3x3: $\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \rightarrow \begin{bmatrix} a & b & c & b \\ d & e & f & c \\ g & h & i & a \end{bmatrix} \mid 2 \times 2 \det \begin{bmatrix} c & d \\ e & f \end{bmatrix} = ad - bc$

$\det(A) = a \cdot \det(b, c, d) - b \cdot \det(a, c, d) + c \cdot \det(a, b, d)$

1. Calc/row with most 0's: 10^3 (if 1, det is 1)
 All: row ops $\rightarrow$ easier Mat; sign change; row $\rightarrow$ nothing

Inverses: $[A|I] \rightarrow [I|B]$, $B = A^{-1}$; Gauss $[B|I]$; Multiply row to $\log_2$ best 5

Expand EV: 1 EV $\rightarrow \det(A - \lambda I) = 0$

2. Find $\vec{v} \neq 0$; $\vec{v} \in N(A - \lambda I) \rightarrow (A - \lambda I) \cdot \vec{v} = \vec{0}$

Cramer's rules: $\begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$ det is small + around 1
 is a_{ij}

$\Rightarrow \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \rightarrow \det(A) \cdot x_j = \det(A) \cdot x_j \Rightarrow x_j = \frac{\det(A)}{\det(A)}$ (if $\det(A) \neq 0$)

Gram Schmidt: $q_1 = \frac{a_1}{\|a_1\|}$, $q_2 = \frac{a_2}{\|a_2\|}$, $q_3 = \frac{a_3}{\|a_3\|}$

Linear Regression/LS: Find $a, b \in \mathbb{R}$ st $\{(t_i) \mapsto at_i + b \mid \min \sum (t_i - y_i)^2\}$; $x_1 = -1, x_2 = 0, x_3 = 1$
 $y_1 = 2, y_2 = 1, y_3 = 3 \Rightarrow$ use L. 5.2.3 with $A = \begin{bmatrix} -1 & 1 \\ 0 & 1 \\ 1 & 1 \end{bmatrix} \Rightarrow$ model const. b , $\hat{b} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$

4. Find Computations: $C(A) = \{ \text{nulls } I \text{ ident } A: R(A)^n \text{ of } A^T; N(A) = \text{nulls}; \text{ what goes to } 0$

SVD: 1. Calculate $A^T A \in \mathbb{R}^{n \times n}$ 2. Calculate EV of $A^T A$ (same as AA^T) and
Evec $AA^T \Rightarrow$ cols of U $A^T A$ Evec $\Rightarrow$ cols of V

3. Singular values of Σ : decreasing $\sqrt{E}$ of $A^T A / A A^T$
 $\hookrightarrow$ If A square: $\Sigma = R^{nn}$ with $\text{rankings} = 0$

Gauss-Jordan: 1. Multiply row with a scalar 2) Subtract row from another
3. Swap rows | GJ: Non-zero rows above 0 rows, Pivot strict right at pivot in prev. row

to make pivot 1 3) Use row ops to make entries above 1 below pivot (pivot col) 0
4) Repeat until in RREF

↳ no sol if 0...0 | b ≠ 0 ↳ unique if every variable corresponds to pivot ↳ variables don't correspond to pivot = free variables ⇒ unlimited solutions
Gauss Solutions: unique solution $[I_n | b]$ square + LI columns
infinite sols $[I_r | b]$ rank < # variables = # cols
 $[I_r | b]$ if ≠ 0 rank coefficient matrix < rank extended M

Solution Space: only a subspace, if $\vec{b} = \vec{0}$

skipped copy of Nullspace

With LU we can solve with $L\vec{y} = \vec{b}$, where $U\vec{x} = \vec{y}$ 1 $A\vec{x} = \vec{b}$

1. $[A|0]$ Augmented matrix

2. $A \rightarrow$ RREF/REF (or GE if less time)

3. Identify free / pivot vars (free if no pivot corresp.)

4. Express sol in terms of free vars (parametr. vars)

5. Choose fixed value for parameter

6. Add to $A\vec{x} = \vec{0}$ solution ⇒

Sol(A, b) = span {Set of vectors}

(If 0 row ⇒ $[0 \dots 0 | b_i]$, if not no sol; If no free vars ⇒ unique sol.)

Eigenvale Equation implies $A\vec{v} = \lambda\vec{v}$, where λ is EV. Thus, $A\vec{v} = \vec{0} \Rightarrow \lambda = 0$, $A\vec{v} = \vec{v} \Rightarrow \lambda = 1$

Invertibility Hack:

- mult. diag. matrix with A does not change rank(A)

- A inv $\Leftrightarrow A^T A$ inv.

- $[D]$ 1 square inv $\Leftrightarrow$ all diag. entries $\neq 0$

- $A = [D]$, if A inv. $\Rightarrow A^{-1} = [D^{-1}]$, also for upper

Special Properties:

- Product of full rank M is full rank, if compatible

- Mult. of diag. M is comm. (Also for Δ -M)

- $AB = BA$ has same EV

- every non-fill rank M has 0 as EV

- nilpotent M has 0 as only EV + Evcc set

PQ-Formel: $x_2 + px + q \Rightarrow -\frac{p}{2} \pm \sqrt{\left(\frac{p}{2}\right)^2 - q}$

Linear Transformation ⇒ Show $T(\vec{x} + \vec{y}) = T(\vec{x}) + T(\vec{y})$
and $T(\lambda\vec{x}) = \lambda T(\vec{x})$ a.k.a. $T(\lambda\vec{x} + \vec{y}) = \dots = \lambda T(\vec{x}) + T(\vec{y})$
LT als Matrix: Wo werden e_i abgebildet? Was ist $e_i = [e_i]$?
 $T\left[\begin{pmatrix} x_1 \\ x_2 \\ \vdots \end{pmatrix}\right] = T(x_1 e_1 + \dots + x_n e_n) = x_1 T(e_1) + \dots + x_n T(e_n)$
 $= x_1 (Ae_1) + \dots + x_n (Ae_n) = x_1 e_1 + \dots + x_n e_n = A \cdot \begin{bmatrix} x_1 \\ x_2 \\ \vdots \end{bmatrix}$
Dann ist $A = [e_1 \dots e_n]$.

If $T(0) \neq 0 \Rightarrow$ no LT (breaks ii))

T_A is LT, if LC of cols of A and scalar of $\vec{x}$

↳ $\vec{0} \in U$

Subspace ⇒ 1) Show $U \neq \emptyset$ 2) Show for $\vec{v}, \vec{w} \in U$

arb.: $\vec{v} + \vec{w} \in U$ 3) arb. $\vec{v} \in U, \lambda \in \mathbb{R}: \lambda\vec{v} \in U$

1, 2, 3 $\Rightarrow U$ is Subspace (can also for VS)

Basis 1) Show B LI 2) Show Span(B) = V

Dim(n) 1) Find Basis 2) Show |B| = n

↳ Prove it's a basis and that $B \subseteq V$, if not shown

Span(B) = V ⇒ show for arb. $\vec{v} \in V, \vec{v} \in B$

Symmetry Use constraints and show $A^T = A$

LI: show if $\lambda\vec{v} + \mu\vec{w} + \beta\vec{u} = \vec{0}$, then $\lambda = \mu = \beta = 0$

- Alt.: Contradiction assume one comb. of others

- Alt. 2: GE on Matrix

Orthogonality (M): Product of OM is Orth., $Q^T = Q^{-1}$

$QQ^T = QQ^T = I, v^T w = 0$

Scalar Product: Prove all 4 axioms (0. 1. 10)

a) Consider the three linearly independent vectors

$$u = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, v = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, w = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

in $\mathbb{R}^3$. Since the three vectors are linearly independent, there is no fourth non-zero vector that is orthogonal to all of u, v, w . An interesting problem would be to find a fourth non-zero vector that is as close as possible to being orthogonal to all three vectors u, v, w . This is a bit outside of the scope of this exam, so instead we ask you to write down a fourth vector $x \in \mathbb{R}^3$ that satisfies the conditions

$$\sum_{i=1}^3 x_i = 2 \text{ and } u^T x = v^T x = w^T x = 0$$

Official Solution:

$$x = \begin{bmatrix} 2/3 \\ 2/3 \\ 2/3 \end{bmatrix}$$

Explanation:

Let $x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$. We now have to solve this to build a 4 × 4 system with variables x_1, x_2, x_3 . Alternatively, we could observe that

$$3x = u^T x + v^T x + w^T x = 2 \Rightarrow \sum_{i=1}^3 x_i = 2$$

and hence $x = \frac{2}{3} \cdot \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$. We can now write down the linear system

$$\begin{bmatrix} 1 & 1 & 1 \\ -1 & -1 & -1 \\ -1 & -1 & -1 \\ -1 & -1 & -1 \end{bmatrix} x = \begin{bmatrix} 2/3 \\ 2/3 \\ 2/3 \\ 2/3 \end{bmatrix}$$

which we want to solve for x . This can be done with elimination and back substitution. Since the system matrix has rank 3, there is either a unique solution or no solution at all. As it turns out, in this example there is a unique solution. We do not give all the elimination steps here. Instead, let us check that our x is indeed a solution to the above system:

$$\begin{bmatrix} 1 & 1 & 1 \\ -1 & -1 & -1 \\ -1 & -1 & -1 \\ -1 & -1 & -1 \end{bmatrix} \cdot \begin{bmatrix} 2/3 \\ 2/3 \\ 2/3 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \\ 2 \\ 2 \end{bmatrix}$$

MC: Set of LI Vectors ⇒ Unit set is dep.?

a) $\{v_1, -2v_2, v_3\}$ b) $\{v_2, v_4\}$ c) $\{v_1, 2v_1 + v_2, v_3 + v_4\}$

d) $\{2v_1 + v_2, v_1 - 2v_2, v_4, 7v_1 - 4v_2\} \rightarrow$ LD ⇒ build LES:

$$1 a v_1 + b \cdot 7 v_1 = 2 v_1, \quad 11 a \cdot (-2 v_2) + b \cdot (-4 v_2) = v_2$$

$$\begin{bmatrix} 1 & 7 \\ 2 & -4 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 7 \\ 0 & -18 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 2 \\ -5 \end{bmatrix} \Rightarrow a = -\frac{3}{2}, b = \frac{1}{2} \Rightarrow$$

LD

Extra:

Nilpotent M: $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, $B = \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix}$ at B^3 , all $\begin{bmatrix} a \\ b \end{bmatrix} \Delta B^3, \frac{1}{n}$

0-Trace basis: $2 \times 2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

2×2 Symm. basis: $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

Permutation Matrices: Swap 2nd/3rd col. $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$

$A = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \rightarrow$ stretched by 2 in vert axis

$A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \rightarrow$ shearing transformation

$A = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \rightarrow$ rotate by $\frac{\pi}{4}$

$A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ reflect by diagonal $x_2 = x_1$

LT Facts: 1) $U \rightarrow V$ LT, $k \in \mathbb{N} \forall v_i \in U$

$\forall a_i \in \mathbb{R} T(a_1 v_1 + \dots + a_n v_n) = a T(v_1) + \dots + a T(v_n)$

2) value of T in basis of U fully determines T

- rank of Δ -M, $[Q]$ -M is # non

zero entries on $[Q]$

- rank-nullity-Theorem: $n = \text{rank}(A) + \dim(N(A))$

- PD M have $\lambda_i > 0 \Leftrightarrow$ always inv.

Functions that are not linear transformations:

(1) For a vector $v \in \mathbb{R}^n$ (with $\|v\| = 1$) the map $T: \mathbb{R}^n \rightarrow \mathbb{R}$ given by $T(v) = \|v\|$.

(2) The map $T: \mathbb{R}^n \rightarrow \mathbb{R}$ given by $T(v) = \|v\|^2$.

(3) The map $T: \mathbb{R}^n \rightarrow \mathbb{R}$ given by $T(v) = \|v\|^2$.

(4) The map $T: \mathbb{R}^n \rightarrow \mathbb{R}$ given by $T(v) = \|v\|^2$.

Functions that are linear transformations:

(1) The identity map $T: \mathbb{R}^n \rightarrow \mathbb{R}^n$ given by $T(v) = v$.

(2) For any matrix A, the map $T: \mathbb{R}^n \rightarrow \mathbb{R}^n$ given by $T(v) = Av$.

(3) For a vector $v \in \mathbb{R}^n$ the map $T: \mathbb{R}^n \rightarrow \mathbb{R}$ given by $T(v) = v^T v$.

(4) The map $T: \mathbb{R}^n \rightarrow \mathbb{R}^n$ given by $T(v) = 0$.

(5) The map $T: \mathbb{R}^n \rightarrow \mathbb{R}^n$ given by $T(v) = 0$.

(6) The map $T: \mathbb{R}^n \rightarrow \mathbb{R}^n$ given by $T(v) = 0$.

(7) The map $T: \mathbb{R}^n \rightarrow \mathbb{R}^n$ given by $T(v) = 0$.

(8) The map $T: \mathbb{R}^n \rightarrow \mathbb{R}^n$ given by $T(v) = 0$.

(9) The map $T: \mathbb{R}^n \rightarrow \mathbb{R}^n$ given by $T(v) = 0$.

(10) The map $T: \mathbb{R}^n \rightarrow \mathbb{R}^n$ given by $T(v) = 0$.

