Linear Algebra 2

Jiaqi Wang

January 29, 2024

1 From Linear Algebra 1

1.1 Field

Definition 1.1.1 **(Field)** A field is a set $\mathbb{K}$ that contains two special elements, 0 and 1, together with two binary operations, addition and multiplication $+, \cdot : \mathbb{K} \times \mathbb{K} \to \mathbb{K}$, such that addition and multiplication obey the following axioms:

1. $a + b = b + a$ (addition is commutative)
2. $(a + b) + c = a + (b + c)$ (addition is associative)
3. $a + 0 = a$ (0 is the additive identity)
4. $\forall a, \exists(-a)[a + (-a) = 0]$ (existence of an inverse for addition)

   In short, i-iv express that $(\mathbb{K}, +, 0)$ is an abelian group.
5. $a \cdot b = b \cdot a$ (or simply ab) (multiplication is commutative)
6. $(ab)c = a(bc)$ (multiplication is associative)
7. $a \cdot 1 = a$ (1 is the multiplicative identity)
8. $\forall a \neq 0, \exists a^{-1}[aa^{-1} = 1]$ (existence of an inverse for multiplication)

   In short, v-viii express that $(\mathbb{K} \setminus \{0\}, \cdot, 1)$ is an abelian group.
9. $a(b + c) = ab + ac$ (multiplication is distributive over addition)

1.2 Vector Space and Subspace

Definition 1.2.1 **(Vector space)** Let $\mathbb{K}$ be a field. A $\mathbb{K}$-vector space is a set V that contains a special element, $\underline{0}$, called the zero vector, together with two operations:

1. addition: $V \times V \to V$,
2. scalar multiplication: $\mathbb{K} \times V \to V$,

such that addition and and multiplication obey the following axioms:

i) $\underline{a} + \underline{b} = \underline{b} + \underline{a}$ (addition is commutative)

ii) $(\underline{a} + \underline{b}) + \underline{c} = \underline{a} + (\underline{a} + \underline{c})$ (addition is associative)

iii) $\underline{a} + \underline{0} = \underline{a}$ ($\underline{0}$ is the additive identity)

iv) $\forall \underline{a}, \exists(-\underline{a})[\underline{a} + (-\underline{a}) = \underline{0}]$ (existence of an inverse for addition)

In short, i-iv express that $(V, +, \underline{0})$ is an abelian group.

v) $1 \cdot \underline{a} = \underline{a}$ (1 is the multiplicative identity)

vi) $\lambda(\underline{a} + \underline{b} = \lambda\underline{a} + \lambda\underline{b}$ (multiplication is distributive over vector addition)

vii) $(\lambda + \mu)\underline{a} = \lambda\underline{a} + \mu\underline{a}$ (multiplication is distributive over scalar addition)

viii) $(\lambda\mu)\underline{a} = \lambda(\mu\underline{a})$ (associativity of combining field multiplication with scalar multiplication)

Definition 1.2.2 (Subspace) A subset W of a vector space V is a (linear) subspace of V if

1. $\underline{0} \in W$.
2. $\lambda\underline{u} + \mu\underline{v} \in W$ for all $\underline{u}, \underline{v} \in W$ and $\lambda, \mu \in \mathbb{K}$.

Definition 1.2.3 (Span) Let V be a vector space over $\mathbb{K}$. Let $\underline{v}_1, \underline{v}_2, \cdots, \underline{v}_n \in V$. The span of $\underline{v}_1, \underline{v}_2, \cdots, \underline{v}_n$ is the set of all linear combinations $\lambda_1\underline{v}_1 + \lambda_2\underline{v}_2 + \cdot \lambda_n\underline{v}_n$. We denote it by $< \underline{v}_1, \underline{v}_2, \cdots, \underline{v}_n >$.

Definition 1.2.4 (Bases) A set of vectors $\{\underline{v}_1, \underline{v}_2, \cdots, \underline{v}_n\}$ is a basis of a vector space V if

1. $\{\underline{v}_1, \underline{v}_2, \cdots, \underline{v}_n\}$ is linearly independent.
2. $\{\underline{v}_1, \underline{v}_2, \cdots, \underline{v}_n\}$ spans V.

Definition 1.2.5 (Dimension) All bases of a vector space have the same size (cardinality), called the dimension of V and denoted by $\dim V$.

1.3 Inner product spaces

Definition 1.3.1 (Standard inner product of $\mathbb{K}^n$) Let $\mathbb{K}$ be a field. For two vectors $\underline{v}, \underline{u} \in \mathbb{K}^n$, the standard inner product $(\cdot, \cdot) : \mathbb{K}^n \times \mathbb{K}^n \to \mathbb{K}$ is defined by

$$(\underline{v}, \underline{u}) = v_1u_1 + v_2u_2 + \cdots + v_nu_n.$$

Let $\underline{u} = (u_1, u_2, \cdots, u_n)$ and $\underline{v} = (v_1, v_2, \cdots, v_n)$ be vectors in $\mathbb{R}^n$. The standard inner product of $\underline{u}$ and $\underline{v}$ is defined as
(Other popular notations are $\langle \underline{u}, \underline{v} \rangle$ or $\underline{u} \cdot \underline{v}$)

Definition 1.3.2 (Inner product) Let V be a vector space. We say that V together with map $(\cdot, \cdot) : V \times V \to \mathbb{K}$ is an inner product space if the following properties are satisfied:

1. $(\underline{u}, \underline{v}) = (\underline{v}, \underline{u})$ (symmetry)
2. $(\lambda\underline{u}, \underline{v} + \underline{w}) = \lambda(\underline{u}, \underline{v}) + \lambda(\underline{u}, \underline{w})$ (linearity)
3. $(\underline{v}, \underline{v}) \geq 0$ and $(\underline{v}, \underline{v}) = 0 \iff \underline{v} = \underline{0}$ (positive definite)

Definition 1.3.3 (Orthogonal) Let V be an inner product space. Two vectors $\underline{u}, \underline{v} \in V$ are orthogonal if $(\underline{u}, \underline{v}) = 0$. We write $\underline{u} \perp \underline{v}$.

Definition 1.3.4 (Length) Let V be an inner product space. The length of a vector $\underline{v} \in V$ is defined as

$$\|\underline{v}\| = \sqrt{(\underline{v}, \underline{v})}.$$

Definition 1.3.5 (Angle between vectors) Let V be an inner product space. The angle between two vectors $\underline{u}, \underline{v} \in V$ is defined as

$$\cos\theta = \frac{(\underline{u}, \underline{v})}{\|\underline{u}\|\|\underline{v}\|}.$$

1.4 Linear Map

Definition 1.4.1 **(Linear Map)** A map $\mathcal{A}: V \to W$ (V, W vector spaces) is linear if

1. $\mathcal{A}(\underline{u} + \underline{v}) = \mathcal{A}(\underline{u}) + \mathcal{A}(\underline{v})$.
2. $\mathcal{A}(\lambda \underline{u}) = \lambda \mathcal{A}(\underline{u})$.

Combined we have $\mathcal{A}(\lambda \underline{u} + \mu \underline{v}) = \lambda \mathcal{A}\underline{u} + \mu \mathcal{A}\underline{v}$.

Example 1.4.2 Reflections, rotations, projections, identity map, zero map, etc.

1.5 Multiplication with Matrices

$\mathcal{A}(\underline{v})) = A \cdot \underline{v}$, where A is a matrix.

1.6 Orthogonal projection

Definition 1.6.1 **(Orthogonal projection)** Let V be a vector space with inner product $\langle \cdot, \cdot \rangle$. Let W be a subspace of V. The orthogonal projection of V onto W is the linear map $\mathcal{P}_W : V \to W$ such that

1. $\mathcal{P}_W(\underline{v}) \in W$.
2. $\underline{v} - \mathcal{P}_W(\underline{v}) \in W^{\perp}$.

Theorem 1.6.2 **(Addition)** For $\mathcal{A}, \mathcal{B}$ linear maps, we define $(\mathcal{A} + \mathcal{B})(\underline{v}) = \mathcal{A}(\underline{v}) + \mathcal{B}(\underline{v})$.

Theorem 1.6.3 **(Scalar multiplication)** For $\mathcal{A}$ linear map, we define $(\lambda\mathcal{A})(\underline{v}) = \lambda\mathcal{A}(\underline{v})$.

Theorem 1.6.4 **(Composition)** For $\mathcal{A}, \mathcal{B}$ linear maps, we define $(\mathcal{A} \circ \mathcal{B})(\underline{v}) = \mathcal{A}(\mathcal{B}(\underline{v}))$.

Theorem 1.6.5 **(Inverse)** For $\mathcal{A}$ linear map, we define $\mathcal{A}^{-1}$ such that $\mathcal{A}^{-1} \circ \mathcal{A} = \mathcal{A} \circ \mathcal{A}^{-1} = \mathcal{I}$.

1.6.6 Powers of maps

1. $\mathcal{A}^2 = \mathcal{A}\mathcal{A}$
2. $\mathcal{A}^n = \mathcal{A}\mathcal{A}^{n-1}$
3. $\mathcal{A}^{-n} = (\mathcal{A}^{-1})^n$

1.6.7 Null space and range

Definition 1.6.8 **(Null space)** The null space of a linear map $\mathcal{A}: V \to W$ is the set of vectors $\underline{v} \in V$ such that $\mathcal{A}(\underline{v}) = \underline{0}$.

Definition 1.6.9 **(Range)** The range of a linear map $\mathcal{A}: V \to W$ is the set of vectors $\underline{w} \in W$ such that $\exists \underline{v} \in V$ such that $\mathcal{A}(\underline{v}) = \underline{w}$.

Theorem 1.6.10 Let $\mathcal{A}: V \to W$ be a linear map. Then $\mathcal{A}$ is injective if and only if $\mathcal{N}(\mathcal{A}) = \{\underline{0}\}$.

Theorem 1.6.11 Let $\mathcal{A}: V \to W$ be a linear map. Then $\mathcal{A}$ is surjective if and only if $\mathcal{R}(\mathcal{A}) = W$.

Theorem 1.6.12 Let $\mathcal{A}: V \to W$ be a linear map. Then $\mathcal{A}$ is bijective if and only if $\mathcal{N}(\mathcal{A}) = \{\underline{0}\}$ and $\mathcal{R}(\mathcal{A}) = W$.

1.6.13 Null space / Range for matrix multiplication

Theorem 1.6.14 Let A be an $m \times n$ matrix. Then $\mathcal{N}(A) = \{\underline{v} \in V \mid A\underline{v} = \underline{0}\}$ and $\mathcal{R}(A) = \{\underline{w} \in V \mid \exists \underline{v} \in V$ such that $A\underline{v} = \underline{w}\}$.

1.6.15 Quotient spaces

Definition 1.6.16 (Quotient space) Let V be a vector space and W a subspace of V. The quotient space V/W is the set of cosets of W in V. I.e. $V/W = \{\underline{v} + W \mid \underline{v} \in V\}$.

Theorem 1.6.17 (Noether's fundamental theorem on homomorphisms) For any linear map $\mathcal{A} : V \to W$, there exists a linear bijection between its range $\mathcal{R}$ and the quottient space $V/\mathcal{N}$.

1.6.18 Example

Take $P : \mathbb{R}^3 \to \mathbb{R}^3, (x, a, b) \mapsto (0, a, b)$ and $\mathcal{R}(P) =< (0, a, 0), (0, 0, b) >$

Proof. ... $\bar{\mathcal{A}} : v/\mathcal{N}(\mathcal{A}) \to W, \underline{v} \mapsto \mathcal{A}\underline{v}$ Restrict target space of $\bar{\mathcal{A}}$ to $\mathcal{R}(A)$. Homework : show that $\bar{\mathcal{A}}$ is linear and injective. So $\bar{\mathcal{A}}$ is a linear bijection. □

$\bar{\mathcal{A}}^{-1}$:

2 Transition Matrices

We can find matrices for linear maps $\mathbb{K}^n \to \mathbb{K}^m$ by using the standard basis of $\mathbb{K}^n$ and $\mathbb{K}^m$. However, working with abstract vector spaces, we do not have a standard basis. For that reason we look at transition matrices.

2.1 Coordinates

Definition 2.1.1 (Coordinates) Let V be an n-dimensional vector space with basis $\alpha = \{\underline{a}_1, \ldots, \underline{a}_n\}$. Every vector $\underline{v} \in V$ can be expressed as a linear combination of the basis vectors in exactly one way:

$$\underline{v} = \sum_{i=1}^{n} \lambda_i \underline{a}_i. \tag{1}$$

The coordinates of $\underline{v}$ with respect to α are the numbers $\lambda_1, \lambda_2, \cdots, \lambda_n$.

Remark 2.1.2. Clearly, the coordinates depend on the choice of the basis α.

Example 2.1.3 Consider the vector space V of real polynomials of degree at most 2, and take the polynomial $p := 1 + 2x + 3x^2$. Let $\alpha := \{1, x, x^2\}$ be a basis of V. The α-coordinates of p are $(1, 2, 3)$

Theorem 2.1.4 Let V be an n-dimensional vector space with basis $\alpha = \{\underline{a}_1, \ldots, \underline{a}_n\}$. We will denote the map sending each vector $\underline{v}$ to its coordinates with respect to α by α.
Then α is an invertible linear map from V to $\mathbb{K}^n$.

With this notation, $\alpha(\underline{v})$ is the coordinate vector of the vector $\underline{v} \in V$ with respect to the basis α.

Definition 2.1.5 (Coordinate transformation (map)) Let α and β be bases of an n-dimensional vector space V. The map $\beta\alpha^{-1} : \mathbb{K}^n \to \mathbb{K}^n$ is called the coordinate transformation (map) from α to β.

2.2 Basis transition matrix

Definition 2.2.1 (Transition matrix) Let α and β be bases of an n-dimensional vector space V. We call the $n \times n$-matrix associated to the linear map $\beta\alpha^{-1}$ the transition matrix from basis α to basis β and denote it by ${}_\beta S_\alpha$.

The following theorem states that multiplication with the matrix ${}_\beta S_\alpha$ translates α- into β-coordinates, and gives a direct description of how the matrix looks, entry-wise.

Theorem 2.2.2 Let α and β be bases of an n-dimensional vector space V and let ${}_\beta S_\alpha$ be the basis transition matrix, i.e. the matrix of $\beta\alpha^{-1}$. Let $\underline{x} := \alpha(\underline{v})$ be the α-coordinate vector of a vector $\underline{v} \in V$. Then the β-coordinate vector of $\underline{v}$ is equal to the product ${}_\beta S_\alpha \underline{x}$. Furthermore, the columns of matrix ${}_\beta S\alpha$ are the β-coordinate vectors of the α-basis vectors.

Remark 2.2.3.

$${}_\alpha S_\beta {}_\beta S_\alpha = I, \text{ so } {}_\alpha S_\beta =_\beta S_\alpha^{-1}$$

Theorem 2.2.4 Let α, β and γ be bases of an n-dimensional vector space V, with respective basis transition matrices ${}_\beta S_\alpha$ and ${}_\gamma S_\beta$. Then the basis transition matrix from α to γ is ${}_\gamma S_\alpha = {}_\gamma S_\beta {}_\beta S_\alpha$.

Remark 2.2.5. It is important to distinguish between calculating with vectors (so elements of the vectors space V) and calculating with coordinates (so sequences of elements from $\mathbb{K}^n$).

2.3 Generalizing the map-matrix connection for spaces that aren't $\mathbb{K}^n$

Definition 2.3.1 (Matrix of a linear map) Let V and W be vector spaces with bases α and β respectively. Let $\mathscr{A} : V \to W$ be a linear map. We denote the matrix of the linear map $\beta\mathscr{A}\alpha^{-1}$ by ${}_\beta A_\alpha$ and call it the matrix of $\mathscr{A}$ with respect to the bases α and β.

Remark 2.3.2. If $V = W$ and $\alpha = \beta$, then we simplify notation by denoting the corresponding matrix by A_α. We call it the matrix of $\mathscr{A}$ with respect to the basis α.

Remark 2.3.3 (How does the matrix look?). The columns of ${}_\beta A\alpha$ are

$$(\beta\mathscr{A}\alpha^{-1})(\underline{e}_i) = \beta(\mathscr{A}\underline{a}_i), \qquad i = 1, \dots, n,$$

meaning the i-th column consists of the β-coordinates of the image $\mathscr{A}\underline{a}_i$ of the i-th basis vector $\underline{a}_i$.

Remark 2.3.4. To find the image of a vector $\underline{v} \in V$, we can:

1. Determine the coordinate vector $\alpha(\underline{v})$ of $\underline{v}$;
2. Multiply $\alpha(\underline{v})$ with the representation matrix ${}_\beta A_\alpha$, yielding the coordinate vector of $\mathscr{A}\underline{v}$;
3. Translate the coordinate vector of $\mathscr{A}\underline{v}$ back to the corresponding vector in W.

2.4 How do base changes affect the matrix of a linear map?

Theorem 2.4.1 (Effect of change of basis) Choose in a finite-dimensional space V two bases α and β, and suppose $\mathscr{A} : V \to V$ is linear. Then

$$A_\beta = {}_\beta S_\alpha A_\alpha {}_\alpha S_\beta.$$

3 Eigenvalues and Eigenvectors

3.1 Diagonalization of matrices

Definition 3.1.1 A square matrix A has *diagonal form* if all elements a_{ij} with $i \neq j$ are zero.

Theorem 3.1.2 Let $\mathcal{A} : V \to V$ be a linear map and let $\alpha = \{\underline{a}_1, \dots, \underline{a}_n\}$ be a basis of V. The matrix A_α has a diagonal form

$$A_\alpha = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n \end{pmatrix}$$

if and only if $\mathcal{A}\underline{a}_1 = \lambda_i \underline{a}_i$ for all $i \in \{1, \dots, n\}$.

3.2 Eigenvalues and eigenvectors

Definition 3.2.1 **(Eigenvector and eigenvalue)** Let $\mathcal{A} : V \to V$ be a linear map from a $\mathbb{K}$-vector space V to itself. A vector $\underline{v} \neq \underline{0} \in V$ is called an *eigenvector* of $\mathcal{A}$ with *eigenvalue* λ if $\mathcal{A}\underline{v} = \lambda\underline{v}$. We denote the set of all eigenvalues of $\mathcal{A}$ by spec$\mathcal{A}$ and call it the *spectrum* of $\mathcal{A}$.

Theorem 3.2.2 Let $\mathcal{A} : V \to V$ be a linear map with representation matrix A_α for a basis α. Then A_α is in diagonal form if and only if α is a basis of eigenvectors of $\mathcal{A}$. In this case, the diagonal entries of A_α are the eigenvalues of $\mathcal{A}$.

Definition 3.2.3 **(Eigenspace)** Let $\mathcal{A} : V \to V$ be a linear map. For any scalar $\lambda \in \mathbb{K}$, we denote

$$E_\lambda := \mathcal{N}(\mathcal{A} - \lambda\mathcal{I})$$

Since *null spaces* are subspaces, E_λ is a subspace, called *the eigenspace* of $\mathcal{A}$ for λ.

Remark 3.2.4. Eigenspaces *indeed are spaces of eigenvectors for a given eigenvalue*: E_λ is the null space of the linear map $\mathcal{A} - \lambda\mathcal{I}$.

- $\underline{v} \in E_\lambda \iff (\mathcal{A} - \lambda\mathcal{I})\underline{v} = \underline{0}$.
- $\underline{v} \in E_\lambda \iff \mathcal{A}\underline{v} - \lambda\underline{v} = \underline{0}$.

So any vector $\underline{v}$ lies in E_λ if and only if $(\mathcal{A} - \lambda\mathcal{I})\underline{v} = \underline{0}$, which is equivalent to $\mathcal{A}\underline{v} - \lambda\underline{v} = \underline{0}$.

Remark 3.2.5 (null space as eigenspace). We can also write the null space of $\mathcal{A}$ as an eigenspace: E_0 consists of vectors that are mapped to 0 times itself, so on $\underline{0}$.

3.3 Computing eigenvalues and eigenspaces

Theorem 3.3.1 λ is an eigenvalue if and only if $det(A - \lambda\mathcal{I}) = 0$. Let $\alpha = \{\underline{a}_1, \dots, \underline{a}_n\}$ be a basis for V, and let

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & & \vdots \\ \vdots & & \ddots & \vdots \\ a_{n1} & \dots & \dots & a_{nn} \end{pmatrix}$$

be the matrix of $\mathcal{A}$ w.r.t. this basis. Then the eigenvectors for eigenvalue λ, in α-coordinates, are the non-zero solutions of the system

$$\begin{pmatrix} a_{11}-\lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22}-\lambda & & \vdots \\ \vdots & & \ddots & \vdots \\ a_{n1} & \dots & \dots & a_{nn}-\lambda \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

3.4 Characteristic polynomial

Definition 3.4.1 (Characteristic polynomial) Let $\mathcal{A} : V \to V$ be a linear map and let A_α be the matrix of $\mathcal{A}$ w.r.t. a basis α. We call the equation $\det(A_\alpha - \lambda\mathcal{I}) = 0$ the *characteristic equation* of A_α, and the left-hand side off this equation, $\det(A_\alpha - \lambda\mathcal{I})$, the *characteristic polynomial* of A_α.

We also call them characteristic equation/polynomial of $\mathcal{A}$, and denote the characteristic polynomial by $\chi_{\mathcal{A}}$.

Theorem 3.4.2 Let $\mathcal{A} : V \to V$ be a linear map, α and β be two bases for V and let A_α/A_β be the matrix of $\mathcal{A}$ w.r.t. a basis α/β. Then $\det(A_\alpha - \lambda\mathcal{I}) = \det(A_\beta - \lambda\mathcal{I})$.

Remark 3.4.3. The characteristic polynomial is independent of the choice of basis.

Theorem 3.4.4 Let $\mathcal{A} : V \to V$ be a linear map on a vector space V of dimension n, and let

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & & \vdots \\ \vdots & & \ddots & \vdots \\ a_{n1} & \dots & \dots & a_{nn} \end{pmatrix}$$

be the matrix of $\mathcal{A}$ w.r.t. a basis α. Then the characteristic polynomial $\chi_{\mathcal{A}}$ is a polynomial of degree (exactly) n, and of the following shape:

$$\chi_{\mathcal{A}}(\lambda) = (-1)^n\lambda^n + (-1)^{n-1}(a_{11} + a_{22} + \dots + a_{nn})\lambda^{n-1} + \dots + c_1\lambda + c_0$$

for some coefficients $c_0, c_1, \dots \in \mathbb{K}$.

Definition 3.4.5 (Trace) The sum of the diagonal elements of a square matrix A is called the *trace* of the matrix A. We denote it by $\operatorname{tr}(A)$.

Theorem 3.4.6 Let $\mathcal{A} : V \to V$ be a linear map with $\dim(V) < \infty$. For every basis α, the matrix A_α

1. has the same trace, which we therefore also call the *trace* of $\mathcal{A}$, and denote it by $\operatorname{tr}(\mathcal{A})$.
2. has the same determinant, which we therefore also call the *determinant* of $\mathcal{A}$, and denote it by $\det(\mathcal{A})$. We have the identity $\det(\mathcal{A}) = c_0$, where c_0 is the constant coefficient of the characteristic equation.

Theorem 3.4.7 Let A be a square matrix with entries in $\mathbb{K}$, where $\mathbb{K} \in \{\mathbb{C}, \mathbb{C}\}$, with characteristic polynomial $\chi_A(\lambda)$. Then the

- trace of the matrix is the sum of the roots of χ_A; and the
- determinant of the matrix is the product of the roots of χ_A.

3.5 Linear independence of eigenvectors

Theorem 3.5.1 Let $\mathcal{A} : V \to V$ be a linear map and let $\underline{v}_1, \ldots, \underline{v}_n$ be eigenvectors of $\mathcal{A}$ for *mutually different* eigenvalues $\lambda_1, \ldots, \lambda_n$. Then $\underline{v}_1, \ldots, \underline{v}_n$ are linearly independent.

4 Invariant subspaces

4.1 Invariant subspace

Definition 4.1.1 (Invariant subspace) Let W be a subspace of V. W is called *invariant under linear map* $\mathcal{A} : V \to V$ if $\mathcal{A}\underline{w} \in W$ for all $\underline{w} \in W$.

Example 4.1.2 Null space and range

- The null space $\mathcal{N}$ pf a linear map $\mathcal{A}$ is always invariant: if $\underline{x} \in \mathcal{N}$, then $\mathcal{A}\underline{x} = \underline{0}$, and $\underline{0} \in \mathcal{N}$.
- The range $\mathcal{R}$ of a linear map $\mathcal{A}$ is invariant if and only if $\mathcal{A}$ is surjective: if $\underline{y} \in \mathcal{R}$, then $\underline{y} = \mathcal{A}\underline{x}$ for some $\underline{x} \in V$, and $\mathcal{A}\underline{x} \in \mathcal{R}$.

Example 4.1.3 Counterexample, rotation in two-dimension space Let $\mathcal{A}$ be a 90° rotation map. Then let $W =< e_1 >$. W is not invariant

Theorem 4.1.4 Let $\mathcal{A} : V \to V$ be linear and let $W =< \underline{a}_1, \ldots, \underline{a}_n >$. W is invariant under $\mathcal{A}$ if and only if $\mathcal{A}\underline{a}_i \in W$ for $i = 1, \ldots, n$.

4.2 Restriction unto an invariant subspace

Definition 4.2.1 (Restriction unto an invariant subspace) If W is invariant under $\mathcal{A}$, then all image vectors $\mathcal{A}\underline{w}$ with $\underline{w} \in W$ are again in W. So if we restrict $\mathcal{A}$ to W, we obtain a well-defined linear map $W \to W$, the *restriction of the map* $\mathcal{A}$ unto W, which we denote by $\mathcal{A}|_W$.

Invariant spaces give us a simpler matrix shape, because the matrix contains a block of the restriction:

Theorem 4.2.2 Suppose $\alpha = \{\underline{a}_1, \ldots, \underline{a}_2\}$ is a basis for V such that $W =< \underline{a}_1, \ldots, \underline{a}_m >$ is invariant under $\mathcal{A}$. Then the matrix A_α has the following form:

$$\begin{pmatrix} & & & * & \ldots & * \\ & M_1 & & \vdots & & \vdots \\ 0 & \ldots & 0 & \vdots & & \vdots \\ & \vdots\vdots & & \vdots & & \vdots \\ 0 & \ldots & 0 & * & \ldots & * \end{pmatrix},$$

The $m \times m$-matrix M_1 is the matrix of the restriction $\mathcal{A}_{|W} : W \to W$ w.r.t. the basis $\{\underline{a}_1, \ldots, \underline{a}_m\}$.

Example 4.2.3 Proving invariance and analysing a map without even knowing its full map description Consider in $\mathbb{R}^4$ the (independent) vectors

$$\underline{a} = (1, -1, 1, -1) \text{ and } \underline{b} = (1, 1, 1, 1).$$

Say we have a linear map $\mathcal{A} : \mathbb{R}^4 \to \mathbb{R}^4$ of which we only know that

$$\mathcal{A}\underline{a} = (4, -6, 4, -6) \text{ and } \mathcal{A}\underline{b} = (4, 6, 4, 6).$$

Even without knowing the full description of $\mathcal{A}$, we will now show that $W =< \underline{a}, \underline{b} >$ is invariant and determine a matrix of the restriction unto W - $\mathcal{A}_{|W} : W \to W$.

To show the invariance of $<\underline{a}, \underline{b}>$, we must verify that $\mathcal{A}\underline{a}$ and $\mathcal{A}\underline{b}$ are linear combinations of $\underline{a}$ and $\underline{b}$. We do this by simultaneously solving the systems of equations with columns $\underline{a}, \underline{b}, \mathcal{A}\underline{a}$ and $\mathcal{A}\underline{b}$:

$$\left(\begin{array}{cc|cc} 1 & 1 & 4 & 4 \\ -1 & 1 & -6 & -6 \\ 1 & 1 & 4 & 4 \\ -1 & 1 & -6 & -6 \end{array}\right)$$

After row reduction and deleting zero rows, the system reduces to

$$\left(\begin{array}{cc|cc} 1 & 0 & 5 & -1 \\ 0 & 1 & -1 & 5 \end{array}\right),$$

which tells us that $\mathcal{A}\underline{a} = 5\underline{a} - \underline{b}$ and $\mathcal{A}\underline{b} = -\underline{a} + 5\underline{b}$, so W is invariant under $\mathcal{A}$. This also tells us how the matrix of the restriction $A_{|W} W \to W$ w.r.t. the basis $\{\underline{a}, \underline{b}\}$ looks like:

$$\begin{pmatrix} 5 & -1 \\ -1 & 5 \end{pmatrix}.$$

Using the restriction matrix, we can now even *determine some eigenvectors without knowing the full map*. The characteristic polynomial of the restriction is $\chi_{A_{|W}}(\lambda) = (5-\lambda)^2 - 1 = 25 - 10\lambda + \lambda^2 - 1 = (\lambda-4)(\lambda-6)$. We find that the matrix has eigenvalues 4 and 6. In coordinates, we compute the respective $E_4 =< (1,1) >$ and $E_6 =< (1,-1 >$. In this basis, the restriction map is simply the diagonal map with the eigenvalues on the diagonal:

$$\begin{pmatrix} 4 & 0 \\ 0 & 6 \end{pmatrix}.$$

We transform the coordinate vectors back into elements of $\mathbb{R}^4$: $\underline{a} + \underline{b} = (2,0,2,0)$ and $\underline{a} - \underline{b} = (0,2,0,2)$. So the eigenvector basis of W is $\{(2,0,2,0),(0,2,0,2)\}$.

We now can simplify the representation of the full map: if we pick any basis α of $\mathbb{R}^4$ such that the first two basis vectors are the eigenvectors $(2,0,2,0)$ and $(0,2,0,2)$, then the full matrix has the shape

$$A_\alpha = \begin{pmatrix} & * & \dots & * \\ \begin{pmatrix} 4 & 0 \\ 0 & 6 \end{pmatrix} & \vdots & & \vdots \\ 0 \dots 0 & \vdots & & \vdots \\ \vdots\vdots & \vdots & & \vdots \\ 0 \dots 0 & * & \dots & * \end{pmatrix}$$

Remark 4.2.4. The characteristic polynomial of a restriction always divides the characteristic polynomial of the larger map.

Theorem 4.2.5 If W is an invariant subspace for the linear map $\mathcal{A}V \to V$, then $\chi_{A_{|W}}$, the characteristic polynomial of $\mathcal{A}$'s restriction unto W, $\mathcal{A}_{|W} : W \to W$, is a factor of $\chi_{\mathcal{A}}$, the characteristic polynomial of the map $\mathcal{A} : V \to V$.

Lemma 4.2.6 Let A be a $p \times p$-matrix, and let B be a $q \times q$-matrix. Then

$$\det \begin{pmatrix} A & * \\ O & B \end{pmatrix} = \det(A) \cdot \det(B),$$

where $*$ stands for na arbitrary $(p \times q)$-matrix and O for the $q \times p$-zero matrix.

4.3 Nice results for combinations of invariant subspaces

Theorem 4.3.1 Let $\alpha = \{\underline{a}_1, \dots, \underline{a}_n\}$ be a basis for V such that $W_1 =< \underline{a}_1, \dots, \underline{a}_m >$ and $W_2 =< \underline{a}_{m+1}, \dots, \underline{a}_n$ are invariant under $\mathcal{A} : V \to V$. Then the matrix A_α has the following form:

$$A_\alpha = \begin{pmatrix} & & 0 & \dots & 0 \\ M_1 & & \vdots & & \vdots \\ & & 0 & \dots & 0 \\ 0 \dots 0 & & & & \\ \vdots\vdots & & & M_2 & \\ 0 \dots 0 & & & & \end{pmatrix}$$

Here M_1 and M_2 are the $m \times m$ and $(n-m) \times (n-m)$ matrices of the two restrictions $\mathcal{A}_{|W_1} : W_1 \to W_1$ and $\mathcal{A}_{|W_2} : W_2 \to W_2$.
In addition we have that

$$\det(A_\alpha) = \det(M_1) \det(M_2),$$

and that the characteristic polynomial of $\mathcal{A}$ is the product of the characteristic polynomials of the two restrictions:

$$\chi_{\mathcal{A}} = \chi_{A_{|W_1}} \chi_{A_{|W_2}}.$$

Remark 4.3.2. We remark that this result can be generalised further such that it holds for an arbitrary number of invariant subspaces: if V can be broken down into invariant subspaces $W_1, \dots, W_p$, we can pick a basis α whose i-th section is a basis for W_i. Let $A_i : W_i \to W_i$ denote the restriction of $\mathcal{A}$ unto the subspace W_i.

Then the matrix A_α has the form.

$$A_\alpha = \begin{pmatrix} M_1 & & \\ & \ddots & \\ & & M_p \end{pmatrix},$$

where M_i is the matrix of the respective restriction $\mathcal{A}_i$ w.r.t. the respective basis.

5 Jordan Normal Form

Definition 5.0.1 **(Jordan blocks)** We denote by $J_n(\lambda)$ the $n \times n$ matrix

$$J_n(\lambda) = \begin{bmatrix} \lambda & 1 & & \\ & \lambda & \dots & \\ & & \dots & 1 \\ & & & \lambda \end{bmatrix}$$

that has λ's on the diagonal, 1's directly above the diagonal, and 0's everywhere else.
We call a matrix J a *Jordan block* if it is of that shape, i.e. $J = J_n(\lambda)$ for some field element $\lambda \in \mathbb{K}$ and some integer n.

Example 5.0.2 Jordan block of dimension 3 for value 2

$$J_3(2) = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}.$$

Definition 5.0.3 **(Jordan Blocks and Jordan Normal Form)** Let A be a quadratic matrix with entries in $\mathbb{K}$. We say that A is in *Jordan normal form* if it is of a diagonal block shape

$$A = \begin{bmatrix} J_1 & & \\ & \cdots & \\ & & J_p \end{bmatrix},$$

with each block J_i being a Jordan block, so each block being $J_i = J_{n_i}(\lambda_i)$ for some field element $\lambda_i \in \mathbb{K}$ and some integer n_i.

5.1 Annihilating matrices using minimal polynomials

Theorem 5.1.1 **(Cayley-Hamilton theorem for $\mathbb{K} = \mathbb{C}$)** A complex quadratic matrix's characteristic polynomial always annihilates it.
More formally: Let A be a complex quadratic matrix with characteristic polynomial χ_A, and let $\chi_A(A)$ be the matrix obtained by applying χ_A to A. Then $\chi_A(A) = 0$ (zero matrix).

Definition 5.1.2 **(Minimal polynomial)** Let A be a quadratic square matrix. The minimal polynomial of A is the (unique) polynomial m_A for which

- m_A annihilates A, i.e. $m_A(A) = 0$ (zero matrix)
- m_A is minimal, i.e. any other polynomial p that annihilates A is a multiple of m_A.
- m_A is monic, i.e. the leading coefficient of m_A is 1.

Theorem 5.1.3 Let m be a monic polynomial that annihilates a complex square matrix A. Then the following two statements are equivalent:

1. m divides all other polynomials that annihilate A.
2. The degree of m is minimal, i.e. the degree of all other polynomials annihilating A is at least as high as m's degree.

Corollary 5.1.4 The minimal polynomial m_A of a complex square matrix A always divides the characteristic polynomial χ_A.

Theorem 5.1.5 Let $\mathcal{A} : V \to V$ be a linear map on a finite-dimensional complex vector space V, with distinct eigenvalues $\lambda_1, \dots, \lambda_n$.
Then $(\lambda - \lambda_1) \cdot (\lambda - \lambda_2) \dots (\lambda - \lambda_n)$ divides the minimal polynomial m_A.

Corollary 5.1.6 Let $\mathcal{A} : V \to V$ be a linear map on a finite-dimensional complex vector space V, with distinct eigenvalues $\lambda_1, \dots, \lambda_n$ and characteristic polynomial $\chi_{\mathcal{A}}(\lambda) = \Pi_{i=1}^n (\lambda - \lambda_i)^{m_i}$. Then the minimal polynomial of $\mathcal{A}$ is

$$m_{\mathcal{A}}(\lambda) = \Pi_{i=1}^n (\lambda - \lambda_i)^{e_i}$$

with each multiplicity e_i being somewhere between 1 and m_i, the algebraic multiplicity of λ_i.

5.2 Generalized eigenvectors and spaces

Definition 5.2.1 **(Generalized eigenvector)** Let $\mathcal{A} : V \to V$ be a linear map on a finite-dimensional vector space, and let μ be an eigenvalue of $\mathcal{A}$. We call $\underline{x}$ a *generalized eigenvector of rank* m of $\mathcal{A}$ for eigenvalue μ if

- $(\mathcal{A} - \mu\mathcal{I})^m \underline{x} = \underline{0}$, but
- $(\mathcal{A} - \mu\mathcal{I})^{m-1} \neq \underline{0}$

we denote the collection of all generalized eigenvectors of rank up to m by $E_\mu^{(m)}$, so

$$E_\mu^{(m)} := \mathcal{N}((\mathcal{A} - \mu\mathcal{I})^m),$$

and call it the *generalized eigenspace of rank* m of $\mathcal{A}$ for eigenvalues μ.

Theorem 5.2.2 Generalized eigenspaces are invariant.

Theorem 5.2.3 Let $\mathcal{A} : V \to V$ be a linear map. For any polynomial f with coefficients in V's scalar field, we have

$$f(\mathcal{A})(\mathcal{A}\underline{v}) = \mathcal{A}(f(\mathcal{A})\underline{v})$$

Theorem 5.2.4 **(Primary Decomposition Theorem)** For any linear map $\mathcal{A} : V \to V$ on a finite-dimensional complex vector space V, we can find a basis of generalized eigenvectors. V splits into the (invariant) generalized eigenspaces, with the ranks corresponding to the respective eigenvalue's degree in the minimal polynomial.
More formally, let $\mathcal{A} : V \to V$ be a linear map on a finite-dimensional complex vector space V, with distinct eigenvalues $\lambda_1, \ldots, \lambda_n$ and minimal polynomial $m_{\mathcal{A}}(\lambda) = \Pi_{i=1}^n (\lambda - \lambda_i)_i^e$.
Then for each i, we find a basis α_i of the generalized eigenspace $E_{\lambda_i}^{(e_i)}$ in a way such that

- the bases $\alpha_1, \ldots, \alpha_n$, are all linearly independent
- combining the bases, so taking the union $\alpha := \bigcup_i \alpha_i$, spans the full vector space V.

Remark 5.2.5. Since we are decomposing the vector space along the map's eigenvalues, and since the collection of the map's eigenvalues is called the spectrum, this is also sometimes called *spectral decomposition*.

5.3 Finding a basis that creates the Jordan Normal Form

Theorem 5.3.1 Let $\mathcal{A} : V \to V$ be a linear map on a finite-dimensional complex vector space, and let μ be an eigenvalue such that $m_{\text{geo}}(\mu) < m_{\text{alg}}(\mu)$.
Then there exists a vector $\underline{x}$ such that $(\mathcal{A} - \mu\mathcal{I})\underline{x} \neq \underline{0}$, but $(\mathcal{A} - \mu\mathcal{I})^2\underline{x} = \underline{0}$. In other words, $\underline{x}$ is a generalized eigenvector of rank 2.

Theorem 5.3.2 Let $\mathcal{A} : V \to V$ be a linear map on a finite-dimensional vector space. If the algebraic multiplicity of 0 is larger than its geometric multiplicity, then the intersection of $\mathcal{A}$'s range and null space contains more than the zero vector.

Theorem 5.3.3 **(2-dimensional Jordan form from rank 2-eigenvector, general case)** In general, taking a generalized eigenvector $\underline{x}$ of rank 2 for an eigenvalue μ and the vector $\underline{x}' := (\mathcal{A} - \mu\mathcal{I})\underline{x}$ always creates an invariant, two-dimensional subspace of $E_\mu^{(2)}$, and for the basis $\{\underline{x}', \underline{x}\}$, the respective representation matrix for this subspace is the 2-dimensional Jordan block for eigenvalue μ, so

$$J_2(\mu) = \begin{pmatrix} \mu & 1 \\ 0 & \mu \end{pmatrix}.$$

Definition 5.3.4 **(Jordan chain)** Let μ be an eigenvalue of the matrix A whose degree in the minimal polynomial m_A is e, and let $\underline{x}$ be a generalized eigenvector of rank e for eigenvalue μ. We set

- $\underline{x}_e := \underline{x}$
- $\underline{x}_{e-1} := (A - \mu\mathcal{I})\underline{x}_e$

- $\underline{x}_{e-2} := (A - \mu\mathcal{I})\underline{x}_{e-1} = (A - \mu\mathcal{I})^2\underline{x}$
- ...
- $\underline{x}_1 := (A - \mu\mathcal{I})\underline{x}_2 = (A - \mu\mathcal{I})^{e-1}\underline{x}$

and call $(x_1, \ldots, x_e)$ a *Jordan chain* for eigenvalue μ.

Theorem 5.3.5 (Jordan form from generalized eigenvectors of higher ranks) A Jordan chain as defined creates an invariant, e-dimensional subspaces of $E_\mu^{(e)}$, and when picking the Jordan chain as its basis, the respective representation matrix for this subspace is the e-dimensional Jordan block for eigenvalue μ.

Theorem 5.3.6 Let *mu* be an eigenvalue of a matrix A, and let the degree of the linear term $\lambda - \mu$ in the matrix's minimal polynomial m_A be e.
Then the length of the longest Jordan chain for μ is e.

Theorem 5.3.7 The dimension of a generalized eigenspace that appears in the vector space in the vector space decomposition matches the algebraic multiplicity of the respective eigenvalue.
More formally, let μ be an eigenvalue of a linear map $\mathcal{A} : V \to V$ on a finite-dimensional vector space V, and let e denote the multiplicity of μ in $\mathcal{A}$'s minimal polynomial $m_{\mathcal{A}}$.
Then the dimension of the generalized eigenspace $E_\mu^{(e)}$ is equal to the algebraic multiplicity of μ, i.e. the multiplicity of μ in the characteristic polynomial.

Theorem 5.3.8 The number of Jordan blocks for an eigenvalue μ equals the dimension of the corresponding eigenspace E_μ, so the dimension of the null space $\mathcal{N}(A - \mu\mathcal{I})$.

Remark 5.3.9 (Takeaway).

- We compute the distinct eigenvalues $\lambda_1, \ldots, \lambda_n$ of the matrix.
- Per eigenvalue, we compute the following:
  - Number of Jordan blocks for λ_i: is equal to the number of linearly independent usual eigenvectors
  - Dimension of the largest Jordan block belonging to λ_i: is equal to the degree of the factor $\lambda - \lambda_i$ in the minimal polynomial.
  - Dimension of the generalized eigenspace for λ_i, so size of all λ_i-blocks combined: is equal to the degree of the factor $\lambda - \lambda_i$ in the characteristic polynomial.

6 Orthogonal and symmetric maps

6.1 Orthogonal maps

Definition 6.1.1 (Orthogonal map) Let v be a real inner product space. A linear map $\mathcal{A} : V \to V$ is called *orthogonal* if

$$\|\mathcal{A}\underline{x}\| = \|\underline{x}\|$$

for all vectors $\underline{x} \in V$. In other words, a linear map $\mathcal{A} : V \to V$ is orthogonal if the *length is invariant* under $\mathcal{A}$.

Theorem 6.1.2 (Polarization formula) In a real inner product space V, we always have

$$(\underline{x}, \underline{y}) = \frac{1}{2}\left((\underline{x} + \underline{y}, \underline{x} + \underline{y}) - (\underline{x}, \underline{x}) - (\underline{y}, \underline{y})\right)$$

As a consequence, we can express *inner products between vectors in terms of vector lengths*:

$$(\underline{x}, \underline{y}) = \frac{1}{2}\left(\|\underline{x}+\underline{y}\| - \|x\| - \|y\|\right).$$

Theorem 6.1.3 Let V be a finite real inner product space, and let $\mathcal{A} : V \to V$ be linear. Then the following are equivalent:

1. $\mathcal{A}$ is orthogonal.
2. $\|\mathcal{A}\underline{x}\| = \|\underline{x}\|$ for all $\underline{x} \in V$.
3. $(\mathcal{A}\underline{x}, \mathcal{A}\underline{y}) = (\underline{x}, \underline{y})$ for all $\underline{x}, \underline{y} \in V$.
4. For every orthonormal system $\underline{a}_1, \ldots, \underline{a}_n$ in V, the system $\mathcal{A}\underline{a}_1, \ldots, \mathcal{A}\underline{a}_n$ is again orthonormal.
5. For every orthonormal basis α of V, the basis $\mathcal{A}\alpha$ is again orthonormal.

Theorem 6.1.4 Let V be a finite real inner product space, and let $\mathcal{A} : V \to V$ and $\mathcal{B} : v \to V$ be orthogonal linear maps.

1. The composition $\mathcal{AB} : V \to V$ is orthogonal.
2. $\mathcal{A}$ is invertible and $\mathcal{A}^{-1}$ is orthogonal.

Remark 6.1.5. As a consequence, powers of orthogonal maps are orthogonal. However, in infinite dimensional spaces, there are orthogonal maps that are not invertible.

6.2 Orthogonal matrices

Corollary 6.2.1 We now consider $\mathbb{R}^n$ with the standard inner product. A linear map $\mathcal{A} : \mathbb{R}^n \to \mathbb{R}^n$ is orthogonal if and only if the matrix $\mathcal{A}e_1, \ldots, \mathcal{A}e_n$ is an orthonormal system.

Definition 6.2.2 (Orthogonal matrix) A real $n \times n$-matrix A is called *orthogonal* if the columns of A form an orthonormal system in $\mathbb{R}^n$.

Theorem 6.2.3 Let $\mathcal{A} : \mathbb{R}^n \to \mathbb{R}^n$ be a linear map with representation matrix A. The following statements are equivalent:

1. $\mathcal{A}$ is orthogonal.
2. A is orthogonal.
3. $A^\top A = \mathcal{I}_n$. In other words, the transpose the inverse.
4. The rows of A form an orthonormal system in $\mathbb{R}^n$.

Lemma 6.2.4 Let V be an n-dimensional real inner product space, with its inner product denoted as $(\cdot,\cdot)_V$, and α an orthonormal basis of V. Let $\mathcal{A} : V \to V$ be an orthogonal map. We denote $\|\cdot\|_{\text{st}}$ the standard length in $\mathbb{R}^n$ and by $\|\cdot\|_V$ the length implied by V's inner product.

1. $\|\alpha\mathcal{A}\underline{v}\|_{\text{st}} = \|\underline{v}\|_V$
2. $\left\|\mathcal{A}\alpha^{-1}\underline{x}\right\| = \|\underline{x}\|_{\text{st}}$
3. $\alpha\mathcal{A}\alpha^{-1} : \mathbb{R}^n \to \mathbb{R}^n$ is orthogonal.
4. If $\mathcal{B} : \mathbb{R}^n \to \mathbb{R}^n$ is orthogonal, the so is $\alpha^{-1}\mathcal{B}\alpha : V \to V$.

Theorem 6.2.5 If α, β are two *orthonormal* bases of in a real inner product space, then the transition matrix $\beta S\alpha$ is orthogonal.

Theorem 6.2.6 If α and β are two orthonormal bases in a real inner product space, then

$$\alpha S\beta = \beta S\alpha^{-1} = \beta S\alpha^{\top}$$

Remark 6.2.7. In a nutshell: we can switch between

- the vector space level and the coordinate level or
- coordinate systems for different basis

without disrupting orthogonality (as long as we are working with an orthonormal basis).

Theorem 6.2.8 Let α be an orthonormal basis for a finite-dimensional real inner product space V, and let $\mathcal{A} : V \to V$ be a linear map and A_α the matrix of $\mathcal{A}$ (with respect to basis α). Then the map $\mathcal{A}$ is orthogonal if and only if its matrix A_α is orthogonal.

6.3 Classification of orthogonal maps

Theorem 6.3.1 Let α be an orthonormal basis for a finite-dimensional real inner product space V, and let $\mathcal{A} : V \to V$ be a linear map and A_α the matrix of $\mathcal{A}$
Then $\det(A_\alpha) = \pm 1$. We say that an orthogonal map is

- *directly orthogonal* if $\det(A_\alpha) = 1$
- *indirectly orthogonal* if $\det(A_\alpha) = -1$

Theorem 6.3.2 Let α be an orthonormal basis for a real inner product space V of dimension n, and let $\mathcal{A} : V \to V$ be an orthogonal map, A_α the matrix of $\mathcal{A}$, and $\chi_{\mathcal{A}} = \det(A_\alpha - \lambda\mathcal{I})$ the characteristic polynomial of $\mathcal{A}$. Then

- every real root of $\chi_{\mathcal{A}}$ is either 1 or -1
- for any non-real root μ of $\chi_{\mathcal{A}}$, the complex conjugate $\overline{\mu}$ is also a root of $\chi_{\mathcal{A}}$
- if $\det(A_\alpha) = -1$, then -1 is an eigenvalue of $\mathcal{A}$then

Theorem 6.3.3 (Classification for dimension 1) On real vector spaces of dimension 1,

- the only map with $\det(A\alpha) = 1$ is $\mathcal{I}$
- the only map with $\det(A\alpha) = -1$ is $-\mathcal{I}$

Theorem 6.3.4 (Classification for dimension 2) If V is a real inner product space of dimension 2, then

- if $\det(A_\alpha) = 1$, then $\mathcal{A}$ is a rotation around the origin by some angle φ, and we can always find an orthonormal basis α such that

$$A_\alpha = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}$$

- if $\det(A_\alpha) = -1$, then $\mathcal{A}$ is a reflection with a line $\ell =< \underline{a} >$ as its reflection axis, with eigenvalues

± 1 and eigenspace $E_1 = \ell$ and $E_{-1} = \ell^\perp$. We can always find an orthonormal basis α such that

$$A_\alpha = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Remark 6.3.5. When looking for a basis that brings directly orthogonal maps on $\mathbb{R}^2$ into rotation matrix shape, *any* orthonormal basis will do. In fact, any orthonormal basis will lead to almost the same rotation matrix

Remark 6.3.6 ($\mathbb{R}^2$: Rotation matrix looks mostly the same for different orthonormal bases). We are now looking at a directly orthogonal map $\mathcal{A}$ acting on $\mathbb{R}^2$, and take the standard basis ε. We always find an angle φ such that the respective matrix A_ε of $\mathcal{A}$ is

$$A_\varepsilon = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}$$

To analyze orthogonal maps on spaces of dimension larger than two, we will simplify the task by using an additional trick: we will pick an invariant subspace W and split V into W and its orthogonal complement $W^\perp$.

Theorem 6.3.7 Let $\mathcal{A} : V \to V$ be an orthogonal map on a real finite-dimensional inner product space V, and let W be a linear subspace such that W is invariant under $\mathcal{A}$.
Then $W^\perp$ is also invariant under $\mathcal{A}$.

Theorem 6.3.8 Let $\mathcal{A}$ be an orthogonal map in a finite-dimensional inner product space V, and W be an invariant subspace of V. Let α_W be an orthonormal basis of W, $\alpha_{W^\perp}$ be an orthonormal basis of $W^\perp$, and set $\alpha := \alpha_W \cup \alpha_{W^\perp}$. Then

$$A_\alpha = \begin{pmatrix} M_1 & O_1 \\ O_2 & M_2 \end{pmatrix}$$

where M_1, M_2 are the orthogonal matrices of the restrictions $\mathcal{A} : W \to W$ and $\mathcal{A} : W^\perp \to W^\perp$ and O_1, O_2 are the zero matrices of the right size. In addition, $\det(A_\alpha) = \det(M_1) \cdot \det(M_2)$.

Theorem 6.3.9 (Classification for dimension 3) If V is a real inner product space of dimension 3, then

- if $\det(A) = 1$, then $\mathcal{A}$ is a rotation around a line $\ell =< \underline{a}_1 >$ (called the axis of rotation) by some angle φ. We can always find an orthonormal basis $\alpha = \{\underline{a}_1, \underline{a}_2, \underline{a}_3\}$ of V such that the respective matrix A_α of $\mathcal{A}$ has the following (rotation matrix) shape:

$$A_\alpha = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\varphi & -\sin\varphi \\ 0 & \sin\varphi & \cos\varphi \end{pmatrix}$$

  A_α has the trace $\operatorname{tr}(A_\alpha) = 1 + 2\cos\varphi$

- if $\det(A) = -1$, then $\mathcal{A}$ is a rotoreflection, i.e. a rotation around a line $\ell =< \underline{a}_1 >$ by some angle φ, together with a reflection with mirror plane $< \underline{a}_1 >^\perp$. We can always find an orthonormal basis $\alpha = \{\underline{a}_1, \underline{a}_2, \underline{a}_3\}$ of V such that the respective matrix A_α of $\mathcal{A}$ has the following shape:

$$A_\alpha = \begin{pmatrix} -1 & 0 & 0 \\ 0 & \cos\varphi & -\sin\varphi \\ 0 & \sin\varphi & \cos\varphi \end{pmatrix}$$

  A_α has the trace $\operatorname{tr}(A_\alpha) = -1 + 2\cos\varphi$.

Remark 6.3.10 (Takeaway).

1. Compute $\det(A)$ to determine whether $\mathcal{A}$ is a rotation or a rotoreflection.
2. The determinant also tells us the signs of the eigenvalue.
3. We determine the rotation axis by computing an eigenvector $\underline{a}_1$ for the eigenvalue.
4. Once we know $\det(A)$, we can read out the rotation angle φ from the trace.
5. We determine the basis $\alpha = \{\underline{a}_1, \underline{a}_2, \underline{a}_3\}$ for which the representation matrix has the wanted shape. To enforce diagonal block shape, we split V into the line spanned by $\underline{a}_1$ and its orthogonal complement $< \underline{a}_1 >^{\perp}$. Concretely:
   - We set the first basis vector to $\underline{a}_1$.
   - To complete the basis, we need to fill up $\{\underline{a}_1\}$ with some orthonormal vectors $\underline{a}_2, \underline{a}_3$ that are orthogonal to $\underline{a}_1$.
   - We can find $\underline{a}_2, \underline{a}_3$, e.g. by applying Gram-Schmidt to the 'prebasis' $\{\underline{a}_1, \underline{e}_2, \underline{e}_3\}$.
6. Rotoreflection case: the mirror plane is the plane which is orthogonal to the rotation axis, $< \underline{a}_1 >$, so we can determine the mirror plane as the plane spanned by the two bases vectors $\underline{a}_2, \underline{a}_3$.

Theorem 6.3.11 (Classification for dimensions > 3) If V is a real inner product space of dimension $n \geq 3$ and $\mathcal{A} : V \to V$ is an orthogonal map, we can always find an orthonormal basis α of V and a number $m \geq 0$ such that the respective matrix A_α of $\mathcal{A}$ has the following form:

$$A_\alpha = \begin{pmatrix} 1 & & & & & & & \\ & \ddots & & & & & & \\ & & 1 & & & & & \\ & & & -1 & & & 0 & \\ & & & & \ddots & & & \\ & & & & & R_1 & & \\ & & & 0 & & & \ddots & \\ & & & & & & & R_m \end{pmatrix}$$

where $R_1, \ldots, R_m$ are 2×2 rotation matrices for some rotation angles $\varphi_1, \ldots, \varphi_m$.

6.4 Symmetric maps

Definition 6.4.1 (Symmetric map) Let V be a real inner product space. A linear map $\mathcal{A} : V \to V$ is called *symmetric* if $(\mathcal{A}\underline{x}, \underline{y}) = (\underline{x}, \mathcal{A}\underline{y})$ for all $\underline{x}, \underline{y} \in V$.

Example 6.4.2 Orthogonal projection The orthogonal projection $\mathcal{P}$ on a line ℓ is symmetric:
If $\ell =< \underline{a} >, \|\underline{a}\| = 1$, then $\mathcal{P}\underline{x} = (\underline{x}, \underline{a})\underline{a}$. For all $\underline{x}, \underline{y} \in V$ we have

$$(\mathcal{P}\underline{x}, \underline{y}) = ((\underline{x}, \underline{a})\underline{a}, \underline{y}) = (\underline{x}, \underline{a})(\underline{a}, \underline{y})$$

$$(\underline{x}, \mathcal{P}\underline{y}) = (\underline{x}, (\underline{y}, \underline{a})\underline{a}) = (\underline{y}, \underline{a})(\underline{x}, \underline{a})$$

meaning that $(\mathcal{P}\underline{x}, \underline{y}) = (\underline{x}, \mathcal{P}\underline{y})$. Similarly, one can show that the orthogonal projection unto a subspace W is a symmetric linear map.

6.4.3 Symmetric matrices

Theorem 6.4.4 For a linear map $\mathcal{A} : V \to V$, the following are equivalent:

1. $\mathcal{A}$ is symmetric
2. for every orthonormal system $\underline{a}_1, \ldots, \underline{a}_m$ in V, we have $(\mathcal{A}\underline{a}_i, \underline{a}_j) = (\underline{a}_i, \mathcal{A}\underline{a}_j)$ for all i, j
3. there is an orthonormal basis $\{\underline{a}_1, \ldots \underline{a}_n\}$ of V satisfying $(\mathcal{A}\underline{a}_i, \underline{a}_j) = (\underline{a}_i, \mathcal{A}\underline{a}_j)$ for all i, j

Proposition 6.4.5 Let V be a finite-dimensional real inner product space and α be an orthonormal basis of V. The linear map $\mathcal{A} : V \to V$ is symmetric if and only if its matrix A_α satisfies $A_\alpha = A_\alpha^\top$.

This motivates how we define symmetric matrices:

Definition 6.4.6 **(Symmetric matrix)** A real $n \times n$ matrix A is called *symmetric* if $A = A^\top$.

We can always diagonalize symmetric matrices

Theorem 6.4.7 Let $\mathcal{A} : V \to V$ be a symmetric map on a real finite-dimensional inner product space V. Then all roots of the characteristic polynomial $\chi_{\mathcal{A}}$ are real (and hence an eigenvalue).

Theorem 6.4.8 Let $\mathcal{A} : V \to V$ be a symmetric map on a real finite-dimensional inner product space V, and let W be an invariant subspace.
Then $W^\perp$ is also invariant under $\mathcal{A}$.

Theorem 6.4.9 Let $\mathcal{A} : V \to V$ be a symmetric linear map on V with $\dim(V) < \infty$. Then there exists an orthonormal basis of eigenvectors of $\mathcal{A}$.

For symmetric maps, we can show that eigenspaces for different eigenvalues will always be orthogonal (so the space completely splits into orthogonal eigenspaces).

Theorem 6.4.10 Let $\mathcal{A} : V \to V$ be a symmetric linear map on V with $\dim(V) < \infty$, and λ_1, λ_2 be two distinct eigenvalues. Then all vectors in the eigenspace E_{λ_1} are orthogonal to all vectors in the eigenspace E_{λ_2}. In other words, $E_{\lambda_1} \perp E_{\lambda_2}$.

Corollary 6.4.11 A symmetric matrix is diagonalizable by changing to an orthonormal basis of eigenvectors, so by means of an orthogonal coordinate transformation and hence an orthogonal transition matrix.

Remark 6.4.12 (Takeaway).
1. We use the characteristic polynomial $\det(A - \lambda\mathcal{I})$ to find eigenvalues $\lambda_1, \ldots, \lambda_n$ with their respective eigenvectors $\underline{a}_1, \ldots, \underline{a}_n$.
2. This gives us an eigenvector basis $\alpha := \{\underline{a}_1, \ldots, \underline{a}_n\}$
3. We turn α into an orthonormal eigenvector basis by normalizing all eigenvectors
4. The representation matrix for basis α is the eigenvalue diagonal matrix:
$$A_\alpha = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & \lambda_n \end{pmatrix}$$
5. To transform A into A_α via $A_\alpha =_\alpha S_\varepsilon A_{\varepsilon\varepsilon} S_\alpha$, we need the basis transition matrices:

- ${}_{\varepsilon}S_{\alpha}$ is the matrix with the eigenvectors as columns
- ${}_{\alpha}S_{\varepsilon} = {}_{\varepsilon}S_{\alpha}^{\top}$

6.5 Quadratic forms and analyzing curves

Definition 6.5.1 (Quadratic forms) Quadratic forms over $\mathbb{R}$ are homogeneous polynomials of degree 2 with coefficients in $\mathbb{R}$, i.e. polynomials $p(x_1, \ldots, x_n)$ of the form

$$p(x_1, \ldots, x_n) = \sum_{i=1} a_{ii}x_i^2 + \sum_{i,j>i} b_{ij}x_ix_j$$

for some scalars $a_1, \ldots, a_n, b_{12}, \ldots, b_{n-1,n} \in \mathbb{R}$.

Remark 6.5.2. Any quadratic form $p(x_1, \ldots, x_n)$ over $\mathbb{R}$ ca be written as a vector-matrix product $\underline{x}^{\top}A\underline{x}$ for a symmetric matrix A:

We set the diagonal elements of A to the coefficients of $x_1^2, \ldots, x_n^2$ and for the non-diagonal elements $i \neq j$, we set $a_{ij} = a_{ji}$ to the coefficient of x_ix_j divided by 2.

Remark 6.5.3. If D is a diagonal matrix with diagonal $(\lambda_1, \ldots, \lambda_n)$, then the corresponding quadratic form $p(x_1, \ldots, x_n)$ is

$$\underline{x}^{\top}D\underline{x} = \underline{x}^{\top}\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{pmatrix} = \sum_i \lambda_i x_i^2.$$

Theorem 6.5.4 For any quadratic form $p(x_1, \ldots, x_n)$ over the reals, there exists a substitution rule $\underline{x} \mapsto \underline{y}$ that brings $p(x_1, \ldots, x_n)$ into a form without mixed products.
More formally, there exists a basis transformation map $\alpha : \mathbb{R}^n \to \mathbb{R}^n$ and a quadratic form $p'(y_1, \ldots, y_n)$ over the reals such that

- for $(y_1, \ldots, y_n) := \alpha(x_1, \ldots, x_n)$, we always get $p(x_1, \ldots, x_n) = p'(y_1, \ldots, y_n)$ (same form up to coordinate substitution)
- $p'(y_1, \ldots, y_n) = \sum_i \lambda_i y_i^2$ for some coefficients $\lambda_1, \ldots, \lambda_n$ (form has no mixed products)

Remark 6.5.5. As soon as we know the eigenvalues (including their multiplicities), we can already write down the mixed-term-free 'substituted' form $p'(y_1, \ldots, y_n) = \sum_i \lambda_i y_i^2$ without any further computations.

Equations with additional linear terms

Definition 6.5.6 (Quadratic hyper-surface) A *quadratic hyper-surface H in* $\mathbb{R}$ is the set of solutions for some quadratic equation:

$$H := \{\underline{x} \in \mathbb{R}^n \mid \underline{x}^{\top}A\underline{x} + \underline{b}^{\top}\underline{x} = 0\},$$

where A is a symmetric real $n \times n$ matrix, $\underline{b} \in \mathbb{R}^n$. (and d??? is some real number).

Theorem 6.5.7 For any quadratic hyper-surface H in $\mathbb{R}^n$, there exists a substitution $\underline{x} \mapsto \underline{y}$ that brings the hyper-surface's equation into a form containing all variables just once, and without mixed products.
More formally: let $\underline{x}^{\top}A\underline{x} + \underline{b}^{\top}\underline{x} = 0$ be the defining equation of H. Then there exists a 'coordinate substitution map' $T : \mathbb{R}^n \to \mathbb{R}^n$ and a quadratic form $p'(z_1, \ldots, z_2)$ over the reals such that

- for $(z_1, \ldots, z_n) := T(x_1, \ldots, x_n)$, we always get $\underline{x}^{\top}A\underline{x} + \underline{b}^{\top}\underline{x} = p'(z_1, \ldots, z_n)$ (same form up to coordinate substitution)
- $p'(z_1, \ldots, z_n)$ only contains each z_i once, either as a linear term z_i or as a quadratic term z_i^2 (variables appear only once and not in mixed products).

7 Linear differential equations

7.1 General solution

Definition 7.1.1 **(Derivative map)** We define the **derivative map on** n **many differentiable functions** $\mathbb{R} \to \mathbb{C}$, by defining

$$\mathcal{D}\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} := \begin{pmatrix} \dot{x_1} \\ \vdots \\ \dot{x_n} \end{pmatrix}$$

or, shorter,

$$\mathcal{D}\underline{x} = \underline{\dot{x}}.$$

Definition 7.1.2 **(System of linear differential equations)** A *system of linear differential equations with constant coefficients* and in n many unknown functions $x_1, \dots, x_n$ of a real variable t has the form

$$\underline{\dot{x}} = A\underline{x} + \underline{f},$$

where A is an $n \times n$ matrix and $\underline{f}$ is a vector of n many functions of t.
We call the system *homogeneous* if $\underline{f} = \underline{0}$ and *inhomogeneous* otherwise.

We define a linear map $\mathcal{A}\underline{x} = A\underline{x}$ and use the fact that $\mathcal{D}$ and $\mathcal{A}$ are linear maps on V to rewrite the system as

$$\begin{aligned} \underline{\dot{x}} &= \mathcal{A}\underline{x} + \underline{f} \\ \mathcal{D}\underline{x} - \mathcal{A}\underline{x} &= \underline{f} \\ (\mathcal{D} - \mathcal{A})\underline{x} &= \underline{f} \end{aligned}$$

Remark 7.1.3 (Takeaway). We find all solutions to the system by:

1. Finding the general solution to the homogeneous system

$$(\mathcal{D} - \mathcal{A})\underline{x} = \underline{0}.$$

   i.e.

$$\mathcal{N}(\mathcal{D} - \mathcal{A})$$

2. Finding a particular solution to the inhomogeneous system, i.e. a vector $\underline{p}$, such that

$$\underline{\dot{p}} = A\underline{p} + \underline{f}$$

3. Adding the general solution to the particular solution.

$$\underline{p} + \mathcal{N}(\mathcal{D} - \mathcal{A})$$

7.2 Homogeneous system: diagonalization

Theorem 7.2.1 If A is an $n \times n$ diagonal matrix with diagonal entries $\lambda_1, \dots, \lambda_n$, then the set of solutions for the respective homogeneous system $\underline{\dot{x}} = A\underline{x}$ is

$$\{x \in V \mid \underline{\dot{x}} = \mathcal{A}\underline{x}\} = \{c_1 e^{\lambda_1 t}\underline{e}_1 + \dots c_n e^{\lambda_n t}\underline{e}_n \mid c_1, \dots, c_n \in \mathbb{C}\}$$

Theorem 7.2.2 If A can be brought into real diagonal form via an eigenvector basis $\{\underline{a}_1, \dots, \underline{a}_n\}$, then the set of solutions for the respective homogeneous system $\underline{\dot{x}} = A\underline{c}$

$$\{\underline{x} \in V \mid \underline{\dot{x}} = \mathcal{A}\underline{x}\} = \{c_1 e^{\lambda_1 t}\underline{a}_1 + \cdots + c_n e^{\lambda_n t}\underline{a}_n \mid c_1, \ldots, c_n \in \mathbb{C}\}$$

where λ_i is the eigenvalue for the respective eigenvector $\underline{a}_i$.

7.3 Particular solution

Remark 7.3.1. If the additional term $\underline{f}$ is a vector multiple of e^{at}, try $\underline{p} = \underline{u}e^{at}$. But if $\underline{p} = \underline{u}e^{at}$ did not work, try bigger linear combinations. E.g. $((\underline{u} + \underline{v}t)e^{at}, (\underline{u} + \underline{v}t + \underline{w}t^2)e^{at}, \ldots)$