Algebra 2

Chenglong Yu

January 9, 2026

Contents

1 Introduction

The course roughly covers the following three parts. You may refer to course website of Math 371 Spring 2020 at UPenn in my personal page for related materials.

Part I: Bilinar forms. Symmetric forms, Hermitian forms, and skew-symmetric forms. Orthogonality. Spectral Theorem. Conics and Quadrics. Key examples of classical groups, and their basic properties. Lie algebra (for such groups).

Part II: Group representations. Irreducible representations and unitary representations. Characters. Schur's Lemma. Modules over principal ideal domains. Free modules. Group rings. Noetherian rings. Structure of Abelian groups. Maschke's theorem. Constructions of representations, et cetera.

Part III: Field extensions, algebraic extensions and algebraic closures, splitting fields, separable and inseparable extensions, Galois extensions, Galois correspondences, cyclotomic extensions, solvability by radicals, et cetera.

2 Bilinear Forms

2.1 Basic Definitions

Definition 2.1 (Bilinear Form)**.** *Let V be a vector space over a field $\mathbb{K}$. A map*

$$\langle \cdot, \cdot \rangle \colon V \times V \to \mathbb{K}$$

is called a ***bilinear form*** *if it is linear in both components. That is:*

1. $\langle a\boldsymbol{u} + \boldsymbol{v}, \boldsymbol{w} \rangle = a\langle \boldsymbol{u}, \boldsymbol{w} \rangle + \langle \boldsymbol{v}, \boldsymbol{w} \rangle$
2. $\langle \boldsymbol{u}, a\boldsymbol{v} + \boldsymbol{w} \rangle = a\langle \boldsymbol{u}, \boldsymbol{v} \rangle + \langle \boldsymbol{u}, \boldsymbol{w} \rangle$

for all $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in V$ and $a \in \mathbb{K}$.

Definition 2.2 (Symmetric and Skew-Symmetric Forms)**.** *A bilinear form $\langle \cdot, \cdot \rangle$ is called:*

1. ***Symmetric*** *if $\langle \boldsymbol{v}, \boldsymbol{w} \rangle = \langle \boldsymbol{w}, \boldsymbol{v} \rangle$ for all $\boldsymbol{v}, \boldsymbol{w} \in V$.*
2. ***Skew-symmetric*** *(or alternating) if $\langle \boldsymbol{v}, \boldsymbol{w} \rangle = -\langle \boldsymbol{w}, \boldsymbol{v} \rangle$ for all $\boldsymbol{v}, \boldsymbol{w} \in V$.*

There are following examples of bilinear forms.

Example 2.1 (Euclidean Space)**.** *Let $V = \mathbb{R}^n$. The* ***standard inner product*** *defined by*

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle_{st} = \sum_{i=1}^{n} x_i y_i = \boldsymbol{x}^T \boldsymbol{y}$$

is a symmetric bilinear form. It allows us to define the length of vectors and the angle between non-zero vectors.

Example 2.2 (Minkowski Space)**.** *Let $V = \mathbb{R}^{n+1}$. The* ***Lorentz form*** *is defined by*

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = x_1 y_1 + \cdots + x_{n-1} y_{n-1} - x_{n+1} y_{n+1}.$$

This is a symmetric bilinear form used in special relativity.

Example 2.3 (Matrix Space)**.** *Let $V = M_{m \times n}(\mathbb{R})$. Define*

$$\langle A, B \rangle = \operatorname{tr}(A^T B).$$

This is a symmetric bilinear form on the space of matrices.

For infinite-dimensional spaces, we have the following example.

Example 2.4. *Let V be the space of continuous real-valued functions on $[0, 1]$. Define*

$$\langle f, g \rangle = \int_0^1 f(x) g(x) \, dx.$$

This is a symmetric bilinear form on V.

2.2 Gram Matrices and Congruency

Next we consider finite-dimensional vector spaces V over a field $\mathbb{K}$ with $\dim V = n < \infty$. Let $\mathcal{B} = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ be a basis of V.

Definition 2.3 (Gram Matrix)**.** *The **Gram matrix** of a bilinear form $\langle\cdot,\cdot\rangle$ with respect to the basis $\mathcal{B}$ is the matrix $G_{\langle\cdot,\cdot\rangle,\mathcal{B}} \in M_n(\mathbb{K})$ defined by:*

$$(G_{\langle\cdot,\cdot\rangle,\mathcal{B}})_{ij} = \langle \boldsymbol{v}_i, \boldsymbol{v}_j \rangle.$$

By expansion of bilinearity, we have the following important property.

Proposition 2.1 (Matrix Representation of Bilinear Forms)**.** *If $\boldsymbol{u} = \sum x_i \boldsymbol{v}_i$ and $\boldsymbol{w} = \sum y_j \boldsymbol{v}_j$ are vectors in V with coordinate vectors $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{K}^n$, then the value of the bilinear form can be computed via matrix multiplication:*

$$\langle \boldsymbol{u}, \boldsymbol{w} \rangle = \boldsymbol{x}^T G_{\langle\cdot,\cdot\rangle,\mathcal{B}} \boldsymbol{y}. \tag{1}$$

In fact the formula (1) can be used to define bilinear forms from arbitrary matrices A.

So if we consider the space of bilinear forms on V, it has a natural structure of $\mathbb{K}$-vector space stucture and it is isomorphic to the space of $n \times n$ matrices over $\mathbb{K}$. To summarize, we have the following proposition.

Proposition 2.2 (Matrix representation of bilinear forms)**.** *Let $\mathbf{Bil}(V)$ denote the space of bilinear forms on V. Then $\mathbf{Bil}(V)$ is a vector space over $\mathbb{K}$, and the map*

$$\begin{array}{rcl} \mathbf{Bil}(V) & \to & M_n(\mathbb{K}) \\ \langle\cdot,\cdot\rangle & \mapsto & G_{\langle\cdot,\cdot\rangle,\mathcal{B}} \end{array}$$

is a vector space isomorphism between the space of bilinear forms on V and the space of $n \times n$ matrices over $\mathbb{K}$.

The symmetric and skew-symmetric bilinear forms correspond to symmetric and skew-symmetric matrices, respectively.

Proposition 2.3. *Let $\langle\cdot,\cdot\rangle$ be a bilinear form on V and $A = G_{\langle\cdot,\cdot\rangle,\mathcal{B}}$ is the Gram matrix of $\langle\cdot,\cdot\rangle$ with respect to the basis $\mathcal{B}$. Then:*

1. *$\langle\cdot,\cdot\rangle$ is symmetric if and only if A is a symmetric matrix, i.e. $A = A^T$.*
2. *$\langle\cdot,\cdot\rangle$ is skew-symmetric if and only if A is a skew-symmetric matrix, i.e. $A = -A^T$.*

The dependence of Gram matrices on the choice of basis is described as follows.

Proposition 2.4 (Change of Basis)**.** *Let $\mathcal{B} = \{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n\}$ and $\mathcal{B}' = \{\boldsymbol{w}_1, \cdots, \boldsymbol{w}_n\}$ be two bases of V. Let P be the transition matrix from $\mathcal{B}$ to $\mathcal{B}'$ (i.e., $\boldsymbol{w}_j = \sum_i P_{ij} \boldsymbol{v}_i$). Then the Gram matrices are related by:*

$$G_{\langle\cdot,\cdot\rangle,\mathcal{B}'} = P^T G_{\langle\cdot,\cdot\rangle,\mathcal{B}} P.$$

Proof. By definition:

$$\begin{aligned} (G_{\langle\cdot,\cdot\rangle,\mathcal{B}'})_{jk} &= \langle \boldsymbol{w}_j, \boldsymbol{w}_k \rangle = \langle \sum_i P_{ij} \boldsymbol{v}_i, \sum_l P_{lk} \boldsymbol{v}_l \rangle \\ &= \sum_i \sum_l P_{ij} \langle \boldsymbol{v}_i, \boldsymbol{v}_l \rangle P_{lk} \\ &= \sum_i \sum_l (P^T)_{ji} (G_{\langle\cdot,\cdot\rangle,\mathcal{B}})_{il} P_{lk} \\ &= (P^T G_{\langle\cdot,\cdot\rangle,\mathcal{B}} P)_{jk}. \end{aligned}$$

$\square$

Definition 2.4 (Congruency)**.** *Two square matrices A and B are called* ***congruent*** *if there exists an invertible matrix P such that $B = P^T AP$. It is straightforward to verify that congruency is an equivalence relation on the set of square matrices.*

Since all the invertible $n \times n$ matrices can appear as the change of basis matrix for an n-dimensional vector space, so two matrices are congruent if and only if they represent the same bilinear form under two bases.

Remark 2.1. *Coordinate change of bilinear forms corresponds to matrix congruency, whereas linear operators change coordinates via similarity ($P^{-1}AP$).*

Definition 2.5 (Isometry)**.** *Let $(V_1, \langle\cdot,\cdot\rangle_1)$ and $(V_2, \langle\cdot,\cdot\rangle_2)$ be vector spaces equipped with bilinear forms. A linear map $f : V_1 \to V_2$ is called an* ***isometry*** *if*

$$\langle f(\boldsymbol{u}), f(\boldsymbol{v})\rangle_2 = \langle \boldsymbol{u}, \boldsymbol{v}\rangle_1$$

for all $\boldsymbol{u}, \boldsymbol{v} \in V_1$.

Theorem 2.1. *Two finite-dimensional $\mathbb{K}$-vector spaces with bilinear forms are isometric if and only if their Gram matrices (under any chosen bases) are congruent.*

3 Symmetric Forms

Throughout this subsection, assume $\langle\cdot,\cdot\rangle$ is a **symmetric** bilinear form on a vector space V over a field $\mathbb{K}$ (where $\mathrm{char}(\mathbb{K}) \neq 2$, i.e. 2 is invertible in $\mathbb{K}$).

3.1 Diagonalization of Gram Matrices

The assumption on the characteristic is necessary for the following polarization identity.

Proposition 3.1 (Polarization Identity)**.** *The bilinear form is completely determined by its quadratic form $q(\boldsymbol{v}) = \langle \boldsymbol{v}, \boldsymbol{v}\rangle$. Specifically:*

$$\langle \boldsymbol{v}, \boldsymbol{w}\rangle = \frac{1}{2}\left(\langle \boldsymbol{v}+\boldsymbol{w}, \boldsymbol{v}+\boldsymbol{w}\rangle - \langle \boldsymbol{v}, \boldsymbol{v}\rangle - \langle \boldsymbol{w}, \boldsymbol{w}\rangle\right).$$

This implies that if $\langle\cdot,\cdot\rangle$ is not identically zero, there must exist some vector $\boldsymbol{v}$ such that $\langle \boldsymbol{v}, \boldsymbol{v}\rangle \neq 0$.

Theorem 3.1 (Diagonalization / Orthogonal Basis)**.** *There exists a basis $\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_n\}$ of V such that the Gram matrix of $\langle\cdot,\cdot\rangle$ is diagonal. That is, $\langle \boldsymbol{v}_i, \boldsymbol{v}_j\rangle = 0$ for $i \neq j$.*

Proof. We proceed by induction on $n = \dim V$.

1. **Base case:** If $\langle\cdot,\cdot\rangle \equiv 0$, any basis works. If $n = 1$, any basis works.
2. **Inductive step:** Assume the statement holds for dimensions $< n$. If $\langle\cdot,\cdot\rangle \equiv 0$, we are done. Otherwise, by the polarization identity, there exists $\boldsymbol{v}_1 \in V$ such that $\langle \boldsymbol{v}_1, \boldsymbol{v}_1\rangle \neq 0$ (such a vector is called non-isotropic).

   Define $W = \{\boldsymbol{w} \in V \mid \langle \boldsymbol{v}_1, \boldsymbol{w}\rangle = 0\}$. This is the orthogonal complement of the line spanned by $\boldsymbol{v}_1$. Consider the map $\phi : V \to \mathbb{K}$ given by $\boldsymbol{w} \mapsto \langle \boldsymbol{v}_1, \boldsymbol{w}\rangle$. Since $\langle \boldsymbol{v}_1, \boldsymbol{v}_1\rangle \neq 0$, the map is non-zero, hence surjective. Thus $\dim W = \dim(\ker\phi) = n - 1$.

We claim $V = \text{span}(\boldsymbol{v}_1) \oplus W$. For any $\boldsymbol{v} \in V$, let

$$\boldsymbol{w} = \boldsymbol{v} - \frac{\langle \boldsymbol{v}, \boldsymbol{v}_1 \rangle}{\langle \boldsymbol{v}_1, \boldsymbol{v}_1 \rangle} \boldsymbol{v}_1.$$

Then a direct check shows $\langle \boldsymbol{w}, \boldsymbol{v}_1 \rangle = 0$, so $\boldsymbol{w} \in W$. Thus $\boldsymbol{v} \in \text{span}(\boldsymbol{v}_1) + W$. The intersection is zero because $\boldsymbol{v}_1$ is not in W.

By the induction hypothesis, W admits an orthogonal basis $\{\boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$. Then $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ is an orthogonal basis for V.

□

Under this orthogonal basis, the Gram matrix is diagonal:

$$G = \begin{pmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n \end{pmatrix},$$

where $d_i = \langle \boldsymbol{v}_i, \boldsymbol{v}_i \rangle$.

Remark 3.1. *The proof of the diagonalization can also be obtained from the theory of nondegeneracy criterion on subspaces (see Problem 3).*

3.2 Sylvester's Law of Inertia

When $\mathbb{K} = \mathbb{R}$, we can scale the basis vectors to normalize the diagonal entries coefficients to be 1, -1, or 0. More precisely, we choose the basis vectors as follows:

- If $d_i > 0$, replace $\boldsymbol{v}_i$ by $\frac{1}{\sqrt{d_i}} \boldsymbol{v}_i$.
- If $d_i < 0$, replace $\boldsymbol{v}_i$ by $\frac{1}{\sqrt{-d_i}} \boldsymbol{v}_i$.
- If $d_i = 0$, leave $\boldsymbol{v}_i$ unchanged.

After this scaling, the Gram matrix becomes diagonal with entries in $\{1, -1, 0\}$. In fact, these numbers are determined by the bilinear form itself, independent of the choice of basis.

Theorem 3.2 (Sylvester's Law of Inertia)**.** *Let $\langle \cdot, \cdot \rangle$ be a real symmetric bilinear form on V. There exists a basis under which the Gram matrix is diagonal with entries in $\{1, -1, 0\}$. Usually, the basis is ordered such that:*

$$G = \begin{pmatrix} I_p & 0 & 0 \\ 0 & -I_q & 0 \\ 0 & 0 & 0_{n-p-q} \end{pmatrix}.$$

Furthermore, the integers p (index of positivity) and q (index of negativity) are invariants depending only on $\langle \cdot, \cdot \rangle$, not on the choice of basis.

The triple $(p, q, n - p - q)$ is called the **signature** of the form.

Proof of Uniqueness. The coordinate transformation allows us to write any symmetric matrix A as congruent to a diagonal matrix with diagonal entries $d_i \in \{1, -1, 0\}$. Suppose we have two such decompositions yielding indices (p, q) and (p', q'). Consider the subspaces corresponding to the basis vectors:

- Basis $\mathcal{B} = \{\boldsymbol{v}_1, \dots, \boldsymbol{v}_n\}$ gives p positive, q negative terms. Let $V^+ = \operatorname{span}(\boldsymbol{v}_1, \dots, \boldsymbol{v}_p)$. Then $\dim V^+ = p$.
- Basis $\mathcal{B}' = \{\boldsymbol{w}_1, \dots, \boldsymbol{w}_n\}$ gives p' positive, q' negative terms. Let $W^{\leq 0} = \operatorname{span}(\boldsymbol{w}_{p'+1}, \dots, \boldsymbol{w}_n)$. Then $\dim W^{\leq 0} = n - p'$.

If $p > p'$, then $\dim V^+ + \dim W^{\leq 0} > n$. So the dimension of the intersection

$$\dim V^+ \cap W^{\leq 0} = \dim V^+ \dim W^{\leq 0} - \dim(V^+ + W^{\leq 0}) \geq \dim V^+ \dim W^{\leq 0} - \dim V > 0$$

Then there is an nonzero vector $\boldsymbol{x} \in V^+ \cap W^{\leq 0}$. Write $\boldsymbol{x} = \sum_{i=1}^{p} a_i v_i = \sum_{j=p'+1}^{n} b_j w_j$. Then $\langle \boldsymbol{x}, \boldsymbol{x} \rangle = \sum a_i^2 > 0$ (from V^+) and $\langle \boldsymbol{x}, \boldsymbol{x} \rangle = -\sum_{j=p'+1}^{p\prime+q'} -b_j^2 \leq 0$ (from $W^{\leq 0}$). This is a contradiction. Thus $p \leq p'$. By symmetry, $p' \leq p$, so $p = p'$. A similar argument shows $q = q'$. $\square$

Corollary 3.1. *Two real symmetric matrices A and B of order n are congruent if and only if they have the same positive index of inertia and negative index of inertia.*

Remark 3.2. *Even though the signature $(p, q, n-p-q)$ is unique, the specific orthogonal basis achieving this signature is not unique. The subspaces V^+ contributing the positive part (or the subspaces for the negative part) are not unique. But the subspace corresponding to the zero part is unique. Try to define this subspace intrinsically. It is called the **radical** of the form.*

3.3 Positive Definite Forms

In the proof of Sylvester's law of inertia, we have constructed subspaces where the quadratic form is shows positive and negative properties.

Definition 3.1 (Definiteness)**.** *Let V be a real vector space and $\langle \cdot, \cdot \rangle$ be a symmetric bilinear form.*

1. *$\langle \cdot, \cdot \rangle$ is **positive definite** (denoted $\langle \cdot, \cdot \rangle > 0$) if $\langle \boldsymbol{v}, \boldsymbol{v} \rangle > 0$ for all $\boldsymbol{v} \neq \mathbf{0}$.*
2. *$\langle \cdot, \cdot \rangle$ is **negative definite** (denoted $\langle \cdot, \cdot \rangle < 0$) if $\langle \boldsymbol{v}, \boldsymbol{v} \rangle < 0$ for all $\boldsymbol{v} \neq \mathbf{0}$.*
3. *$\langle \cdot, \cdot \rangle$ is **positive semi-definite** (denoted $\langle \cdot, \cdot \rangle \geq 0$) if $\langle \boldsymbol{v}, \boldsymbol{v} \rangle \geq 0$ for all $\boldsymbol{v} \in V$.*
4. *$\langle \cdot, \cdot \rangle$ is **negative semi-definite** (denoted $\langle \cdot, \cdot \rangle \leq 0$) if $\langle \boldsymbol{v}, \boldsymbol{v} \rangle \leq 0$ for all $\boldsymbol{v} \in V$.*

Proposition 3.2. *Let V be a finite-dimensional real vector space with a symmetric form $\langle \cdot, \cdot \rangle$. The form $\langle \cdot, \cdot \rangle$ is positive definite if and only if its index of positivity p equals $\dim(V)$.*

The proof for the uniqueness of signature also shows the following **intrinsic** characterization.

Proposition 3.3 (Characterization of Signature)**.** *The positive index of inertia p of a symmetric form $\langle \cdot, \cdot \rangle$ on V can be characterized by:*

$$p = \max\{\dim W \mid W \subseteq V \textit{ is a subspace where } \langle \cdot, \cdot \rangle|_W \textit{ is positive definite}\}.$$

Similarly, the negative index q is the maximal dimension of a subspace where $\langle \cdot, \cdot \rangle$ is negative definite.

When the symmetric form is positive definite, we also call the corresponding Gram matrix a **positive definite matrix**. More properties of positive definite matrices are summarized in the exercises.

3.4 Euclidean Spaces

Definition 3.2 (Euclidean Space)**.** *A real vector space V equipped with a positive definite symmetric bilinear form $\langle\cdot,\cdot\rangle$ is called a* ***Euclidean space*** *or an* ***inner product space****.*

See Problem 3 for the definition of nondegeneracy and that if $(V,\langle\cdot,\cdot\rangle)$ is an inner product space, then it is non-degenerate and all its subspaces are also non-degenerate.

Proposition 3.4. *Every Euclidean space of dimension n is isometric to the standard Euclidean space $(\mathbb{R}^n,\langle\cdot,\cdot\rangle_{st})$.*

Example 3.1 (Polynomial Space)**.** *Let $V = P_n(\mathbb{R}) = \{f(x)\in\mathbb{R}[x] \mid \deg f \le n\}$ (sometimes denoted $\mathbb{R}[x]_{\le n}$). Define the bilinear form:*

$$\langle f,g\rangle = \int_0^1 f(x)g(x)\,dx.$$

Since $\int_0^1 (f(x))^2\,dx > 0$ for any non-zero polynomial, $\langle\cdot,\cdot\rangle$ is positive definite. Consider the standard basis $\{1,x,x^2,\dots\}$. The Gram matrix G under this basis has entries:

$$G_{ij} = \langle x^{i-1},x^{j-1}\rangle = \int_0^1 x^{i+j-2}\,dx = \frac{1}{i+j-1}.$$

This matrix is known as the ***Hilbert matrix****.*

3.5 Gram-Schmidt process and QR Decomposition

While Sylvester's theorem guarantees an orthogonal basis, in Euclidean spaces we can construct an **orthonormal** basis algorithmically from any given basis.

Definition 3.3 (Orthonormal Basis)**.** *A basis $\{\boldsymbol{w}_1,\dots,\boldsymbol{w}_n\}$ of a Euclidean space $(V,\langle\cdot,\cdot\rangle)$ is* ***orthonormal*** *if*

$$\langle \boldsymbol{w}_i,\boldsymbol{w}_j\rangle = \delta_{ij} = \begin{cases} 1 & \text{if } i=j \\ 0 & \text{if } i\neq j \end{cases}.$$

In terms of matrices, an orthonormal basis corresponds to orthogonal matrices.

Definition 3.4 (Orthogonal Matrix)**.** *A square real matrix Q is called an* ***orthogonal matrix*** *if its column vectors form an orthonormal basis of $\mathbb{R}^n$ under the standard inner product. Equivalently, Q is orthogonal if and only if $Q^TQ = I$ or $QQ^T = I$, or all the row vectors of Q form an orthonormal basis of $\mathbb{R}^n$.*

Theorem 3.3 (Gram-Schmidt Process)**.** *Let $\{\boldsymbol{v}_1,\dots,\boldsymbol{v}_n\}$ be an arbitrary basis of a Euclidean space V. One can construct an orthonormal basis $\{\boldsymbol{w}_1,\dots,\boldsymbol{w}_n\}$ such that*

$$\operatorname{span}(\boldsymbol{w}_1,\dots,\boldsymbol{w}_k) = \operatorname{span}(\boldsymbol{v}_1,\dots,\boldsymbol{v}_k)$$

for all $k = 1,\dots,n$.

Algorithm. The construction proceeds inductively:

1. Set $\tilde{\boldsymbol{w}}_1 = \boldsymbol{v}_1$. Since basis vector are not zero, the inner product $\langle\tilde{\boldsymbol{w}}_1,\tilde{\boldsymbol{w}}_1\rangle > 0$. Normalize it to obtain $\boldsymbol{w}_1$:
$$\boldsymbol{w}_1 = \frac{\tilde{\boldsymbol{w}}_1}{\sqrt{\langle\tilde{\boldsymbol{w}}_1,\tilde{\boldsymbol{w}}_1\rangle}}.$$

2. Set $\tilde{\boldsymbol{w}}_2 = \boldsymbol{v}_2 - \langle \boldsymbol{v}_2, \boldsymbol{w}_1 \rangle \boldsymbol{w}_1$. This vector is satisfies $\langle \tilde{\boldsymbol{w}}_2, \boldsymbol{w}_1 \rangle = 0$. From the construction, we also know that
$$\operatorname{span}(\tilde{\boldsymbol{w}}_2, \boldsymbol{w}_1) = \operatorname{span}(\boldsymbol{v}_1, \boldsymbol{v}_2).$$
So $\tilde{\boldsymbol{w}}_2 \neq 0$. Normalize it:
$$\boldsymbol{w}_2 = \frac{\tilde{\boldsymbol{w}}_2}{\sqrt{\langle \tilde{\boldsymbol{w}}_2, \tilde{\boldsymbol{w}}_2 \rangle}}.$$

3. In general, for step k, define
$$\tilde{\boldsymbol{w}}_k = \boldsymbol{v}_k - \sum_{j=1}^{k-1} \langle \boldsymbol{v}_k, \boldsymbol{w}_j \rangle \boldsymbol{w}_j$$
Then
$$\langle \tilde{\boldsymbol{w}}_k, \boldsymbol{w}_i \rangle = 0 \text{ for } i < k,$$
Inductively we have
$$\operatorname{span}(\boldsymbol{w}_1, \cdots, \boldsymbol{w}_{k-1}), \tilde{\boldsymbol{w}}_k = \operatorname{span}(\boldsymbol{v}_1, \cdots, \boldsymbol{v}_{k-1}), \tilde{\boldsymbol{w}}_k = \operatorname{span}(\boldsymbol{v}_1, \cdots, \boldsymbol{v}_{k-1}, \boldsymbol{v}_k)$$
so $\tilde{\boldsymbol{w}}_k \neq 0$. Then normalize:
$$\boldsymbol{w}_k = \frac{\tilde{\boldsymbol{w}}_k}{\sqrt{\langle \tilde{\boldsymbol{w}}_k, \tilde{\boldsymbol{w}}_k \rangle}}.$$

□

To summarize, the Gram-Schmidt process constructs an orthonormal basis $\{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n\}$ from any given basis $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ such that the transition matrix from $\{\boldsymbol{w}_i\}$ to $\{\boldsymbol{v}_i\}$ is upper triangular with positive diagonal entries $\frac{1}{\sqrt{\langle \tilde{\boldsymbol{w}}_i, \tilde{\boldsymbol{w}}_i \rangle}}$. Upper triangularity follows from the fact that the subspaces spanned by the first k basis vectors are preserved. So the change of basis matrix has the form:
$$P = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ 0 & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & p_{nn} \end{pmatrix},$$
where $p_{ii} = \frac{1}{\sqrt{\langle \tilde{\boldsymbol{w}}_i, \tilde{\boldsymbol{w}}_i \rangle}} > 0$. The transition matrix from the orthonormal basis $\{\boldsymbol{v}_i\}$ back to the original basis $\{\boldsymbol{w}_i\}$ is then given by P^{-1}, which is also upper triangular with positive diagonal entries.
$$(\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n) = (\boldsymbol{w}_1, \boldsymbol{w}_2, \cdots, \boldsymbol{w}_n) P^{-1}$$

In terms of matrices, this is called the **QR decomposition** of a matrix.

Definition 3.5 (QR Decomposition)**.** *If V is the standard Euclidean space $\mathbb{R}^n$, then any basis $(\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n)$ gather together to form an invertible matrix A. The orthonormal basis vectors $Q = (\boldsymbol{w}_1, \boldsymbol{w}_2, \cdots, \boldsymbol{w}_n)$ form an orthogonal matrix. The decomposition above can be rewritten as*
$$A = QR,$$
*where $R = P^{-1}$ is an upper triangular matrix with positive diagonal entries. This is called the **QR decomposition** of the matrix A.*

The uniquess of the QR decomposition is stated in Problem 2

4 Exercises

4.1 Useful practices

Please submit solutions to the following problems in this section. Some problems help you to review the material we have learned, and some problems introduce useful concepts and theorems not covered in class.

Problem 1. *Practice the Gram-Schmidt process and the QR decomposition. You can choose either one of the following two problems to solve.*

1. *Let $V = P_{\leq 2}(\mathbb{R})$ be the vector space of real polynomials with degree at most 2. Define an inner product on V by*

$$\langle f, g\rangle = \int_0^1 f(x)g(x)\,dx$$

   Given basis $1, x, x^2$ for V, use the Gram-Schmidt process to find an orthonormal basis for V.

2. *Calculate the QR decomposition for the matrix*

$$A = \begin{pmatrix} 3 & 2 & 100 \\ 4 & 0 & 0 \\ 0 & 0 & -5 \end{pmatrix}.$$

Problem 2. *Prove the uniqueness of the QR decomposition: if A is an $n \times n$ invertible real matrix then there exists a unique $n \times n$ orthogonal matrix Q and a unique $n \times n$ upper triangular matrix R with positive diagonal entries such that $A = QR$. (You only need to prove the uniqueness part; the existence part is given by the Gram-Schmidt process.)*

Problem 3. *Let V be a finite-dimensional vector space over field F. Define a symmetric form on V to be a bilinear form $\langle \cdot, \cdot \rangle : V \times V \to F$ similar as the real case. We call the symmetric form non-degenerate if for any $v \in V$, $\langle v, w\rangle = 0$ for all $w \in V$ implies $v = 0$.*

1. *Show that the symmetric form $\langle \cdot, \cdot \rangle$ is non-degenerate if and only if for some (and hence any) basis $\{v_1, \dots, v_n\}$ of V, the Gram matrix $A = (\langle v_i, v_j\rangle)_{1\leq i,j\leq n}$ is invertible.*

2. *Assume V has a nondegenerate symmetric form $\langle \cdot, \cdot \rangle$. Let W be a subspace of V. Define the orthogonal complement of W to be*

$$W^{\perp} = \{v \in V : \langle v, w\rangle = 0 \text{ for all } w \in W\}.$$

   Prove that $W \bigoplus W^{\perp} = V$ if and only if the restriction of $\langle \cdot, \cdot \rangle$ on W is non-degenerate.

3. *When $\mathbb{R}$ is the base field, show that a positive definite symmetric form is non-degenerate.*

4. *For an inner product space V over $\mathbb{R}$, show that for any subspace W of V, $W \bigoplus W^{\perp} = V$.*

5. *In four-dimensional Minkowski space with the Lorentz form, find an one-dimensional subspace W such that W and $W^{\perp}$ do not form a direct sum of the whole space.*

Problem 4. *In this problem, you will prove the Cauchy-Schwarz inequality in an inner product space V. The norm of an inner product space is defined by $\|u\| = \sqrt{\langle u, u\rangle}$. The Cauchy-Schwarz inequality states that for any $u, v \in V$,*

$$|\langle u, v\rangle| \leq \|u\| \cdot \|v\|.$$

Moreover, equality holds if and only if u and v are linearly dependent. You can choose any one of the following two methods to prove it.

1. *Assume $v \neq 0$. Consider the quadratic function of λ*

$$f(\lambda) = \langle u - \lambda v, u - \lambda v \rangle$$

   Show that this function is non-negative and deduce the Cauchy-Schwarz inequality from this.Show that equality holds if and only if u and v are linearly dependent.

2. *There is another method to reduce the Cauchy-Schwarz inequality to two dimensional case. Assume u and v are linearly independent (otherwise the inequality is trivial). Let $W = \operatorname{span}\{u, v\}$. Show that the Cauchy-Schwarz inequality holds in V if it holds in W. Then prove the Cauchy-Schwarz inequality in two-dimensional inner product space by directly considering the standard inner product on $\mathbb{R}^2$.*

Problem 5. *We call a symmetric matrix A positive definite if it is the Gram matrix of any positive definite symmetric form.*

1. *Prove that a symmetric matrix A is positive definite if and only if there exists an invertible matrix P such that $A = P^T P$.*
2. *Show that if A is positive definite, then its determinant is positive.*
3. *Prove that a two by two symmetric matrix is positive definite if and only if it has positive trace and positive determinant.*

Problem 6. *In the following, you will prove the criterion for positive definiteness using principal minors. A principal minor of a matrix A is the determinant of a square submatrix obtained by deleting certain rows and the corresponding columns. A leading principal minor is a principal minor obtained by deleting the last $n - k$ rows and columns for some k. In the following, show that a symmetric matrix A is positive definite if and only if all its leading principal minors are positive.*

1. *Show that if A is positive definite, then all its principal minors are positive. (Hint: consider the restriction of the corresponding symmetric form on the subspace spanned by the first k basis vectors.)*
2. *Use induction to show that if all leading principal minors of A are positive, then matrix A is positive definite. (Hint: use problem 3 and induction.)*

Problem 7. *Let A be a real symmetric matrix where the diagonal elements are all 2, the elements on the two sub-diagonals are all -1, and all other elements are 0. Prove that A is positive definite.*

Problem 8. *Artin chapter 8 1.1. Show that a bilinear form $\langle , \rangle$ on a real vector space V is a sum of a symmetric form and a skew-symmetric form. (skew-symmetric means alternating)*

Problem 9. *Let g be a bilinear form on a real vector space V. Prove that if g satisfies $g(x, y) = 0$ if and only if $g(y, x) = 0$, then g is either symmetric or alternating.*

4.2 Optional problems

If you would like to try some additional problems, you can find them here and you do not need to submit them.

Problem 10. *Prove that the Hilbert matrix of order n,*

$$H_n = (\frac{1}{i+j-1})_{n\times n}$$

is a positive definite matrix. (Hint: Use the symmetric form in Problem 1 (1).)

Problem 11. *Prove the reversed Cauchy-Schwarz inequality in Minkowski space:*

For all $v = (v_0, v_1, \ldots, v_n), w = (w_0, w_1, \ldots, w_n) \in \mathbb{R}^{n+1}$ satisfying

$$v_0^2 - v_1^2 - \cdots - v_n^2 > 0 \text{ and } w_0^2 - w_1^2 - \cdots - w_n^2 > 0,$$

prove the following inequality

$$(v_0^2 - v_1^2 - \cdots - v_n^2)(w_0^2 - w_1^2 - \cdots - w_n^2) \le (v_0w_0 - v_1w_1 - \cdots - v_nw_n)^2$$

and determine the necessary and sufficient condition for equality to hold.

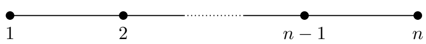
In terms of the Lorentz form $\langle v, w\rangle = v_0w_0 - v_1w_1 - \cdots - v_nw_n$, the inequality can be rewritten as

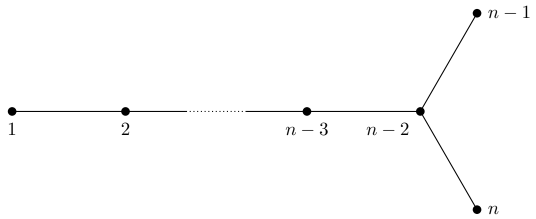
$$\langle v, v\rangle\langle w, w\rangle \le \langle v, w\rangle^2.$$

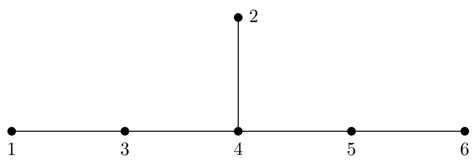
when $\langle v, v\rangle > 0$ and $\langle w, w\rangle > 0$ (in physics v and w are called time-like vectors and this implies two time-like vectors have positive product under the Lorentz form).

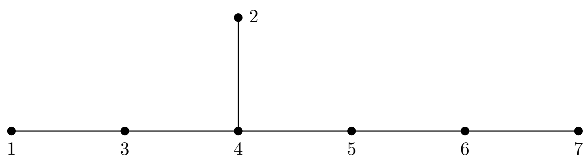
Hint: Use similar method as in Problem 4 (2).

Problem 12 (Challenge)**.** *For a graph Γ with vertices $\{v_1, \cdots, v_n\}$, consider the $n \times n$ real symmetric matrix defined by $A_\Gamma = (a_{ij})_{n\times n}$, where $a_{ij} = 2$ when $i = j$, $a_{ij} = -1$ when $i \ne j$ and v_i, v_j are adjacent (connected by an edge), and $a_{ij} = 0$ otherwise. Prove that for the following graphs Γ, A_Γ is positive definite:*

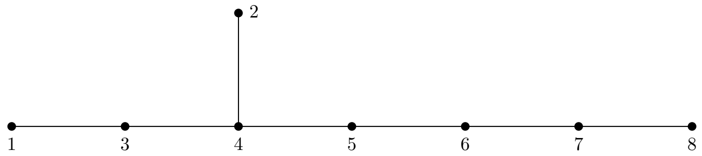
A_n:


D_n:


E_6:


E_7:


E_8:



In fact, these graphs are exactly those connected and whose corresponding matrices are positive definite.