

Linear Algebra 2

Labix

January 20, 2024

Abstract

Contents

1	Linear Forms and Quadratic Forms	3
1.1	Linear Forms	3
1.2	Quadratic Forms	3
1.3	Bilinear Forms	6
1.4	Sesquilinear Forms	8
2	Inner Product Spaces	9
2.1	Norms	9
2.2	Inner Products	10
3	Orthogonality	12
3.1	Orthogonal Vectors	12
3.2	Orthonormal Basis	13
3.3	Orthogonal Complements	14
3.4	Orthogonal Maps	15
4	Orthogonality in \mathbb{R}^n	16
4.1	Reduction of Quadratic Forms over \mathbb{R}	16
4.2	Reduction of Inner Products	17
4.3	Adjoints	18
4.4	Singular Value Decomposition	21

1 Linear Forms and Quadratic Forms

1.1 Linear Forms

Definition 1.1.1: Linear Forms

A linear form on V is a linear map from V to \mathbb{F} .

Proposition 1.1.2: Dual Space

The set of all linear forms on V forms a vector space called the dual space V' .

Proof. Simply a check on the axioms of vector space. □

Lemma 1.1.3

Let V be a finite dimensional vector space. Then V' is also finite dimensional and $\dim(V') = \dim(V)$.

Definition 1.1.4: Dual Basis

Let v_1, \dots, v_n be a basis of V , then the dual basis of v_1, \dots, v_n is the list ϕ_1, \dots, ϕ_n of elements of V' , where ϕ_k is a linear functional such that

$$\phi_k(v_i) = \begin{cases} 1 & \text{if } k = i \\ 0 & \text{if } k \neq i \end{cases}$$

Proposition 1.1.5

The dual basis of a basis of V is a basis of V'

Definition 1.1.6: Dual Map

Let $T \in \mathcal{L}(V, W)$. The dual map of T is the linear map $T' \in \mathcal{L}(W', V')$ defined by $T'(\phi) = \phi \circ T$ for $\phi \in W'$.

Proposition 1.1.7

Let $S, T \in \mathcal{L}(V, W)$ and $\lambda \in \mathbb{F}$.

- $(S + T)' = S' + T'$
- $(\lambda T)' = \lambda T'$
- $(ST)' = T'S'$.

1.2 Quadratic Forms

Definition 1.2.1: Quadratic Forms

A quadratic form in n variables x_1, \dots, x_n over a field K is a polynomial

$$q(x_1, \dots, x_n) = \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j$$

Proposition 1.2.2

Every quadratic form $q(x_1, \dots, x_n) = \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j$ can be represented by a matrix multiplication, namely

$$q(x_1, \dots, x_n) = \begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix} \begin{pmatrix} a_{11} & \frac{1}{2}a_{12} & \cdots & \frac{1}{2}a_{1n} \\ \frac{1}{2}a_{21} & a_{22} & \cdots & \frac{1}{2}a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{2}a_{n1} & \frac{1}{2}a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

In particular, this matrix is symmetric with $a_{ij} = a_{ji}$ for $i, j \in \{1, \dots, n\}$.

Proof. Multiplying out the entries of the matrix multiplication gives the original quadratic form. \square

Proposition 1.2.3

A change of basis via the change of basis matrix P also changes the symmetric matrix of the quadratic form by $P^T A P$

Definition 1.2.4: Congruent Matrices

Two matrices A, B are said to be congruent if there exists some invertible matrix P such that $B = P^T A P$

Beware that congruences does not apply to only symmetric matrices. We will see more of it in action in bilinear forms.

Proposition 1.2.5

Two symmetric matrices are congruent if and only if they represent the same quadratic form with respect to different bases.

Theorem 1.2.6

Let $q(x_1, \dots, x_n)$ be a quadratic form in n variables over a field K whose characteristic is not 2. Then there exists a basis such that $q(y_1, \dots, y_n) = c_1 y_1^2 + \cdots + c_n y_n^2$ for some $c_1, \dots, c_n \in K$.

Proof. There is a shorter proof for this theorem, but for the sake of the construction of c_1, \dots, c_n , we will prove the theorem constructively. Suppose that q is represented by the symmetric matrix $A = (a_{ij})_{n \times n}$ with respect to the basis b_1, \dots, b_n . There are three steps in the construction. I use b_1, \dots, b_n to indicate the old basis and b'_1, \dots, b'_n to indicate the basis after the step.

Step 1: Arrange such that $q(b_1) \neq 0$. There are four cases here.

- If $a_{11} \neq 0$, then we are done.
- If $a_{11} = 0$ but $a_{kk} \neq 0$ for some $1 < k \leq n$. Then just set $b'_1 = b_k$ and $b'_k = b_1$. At the same time, the matrix for the quadratic form is changed by swapping rows r_1 and r_k , and then swapping the columns c_1 and c_k
- If $a_{kk} = 0$ for all $k \in \{1, \dots, n\}$, but there are some i, j such that $a_{ij} \neq 0$, then set $b'_i = b_i + b_j$ since $q(b_i + b_j) = 2a_{ij} \neq 0$ and so we reduced this case to the previous two

cases. The matrix then becomes

$$\begin{pmatrix} 2a_{1k} & a_{12} + a_{k2} & \cdots & a_{1k} & \cdots & a_{1n} + a_{kn} \\ a_{12} + a_{k2} & 0 & \cdots & a_{2k} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{1k} & a_{k2} & \cdots & 0 & \cdots & a_{kn} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{1n} + a_{kn} & a_{n2} & \cdots & a_{nk} & \cdots & 0 \end{pmatrix}$$

- If $a_{ij} = 0$ for all $i, j \in \{1, \dots, n\}$ then it is the zero function.

In this step the change of basis matrix is just the elementary matrices.

Step 2: Now we modify b_2, \dots, b_n to make them orthogonal to b_1 . Now set $b'_k = b_k - \frac{a_{1k}}{a_{11}}b_1$. This way, the matrix entry a_{1k} becomes zero. Now the change of basis matrix becomes

$$P = \begin{pmatrix} 1 & -\frac{a_{12}}{a_{11}} & \cdots & -\frac{a_{1n}}{a_{11}} \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}$$

After this step all the change of basis matrix should be compiled and calculated so that the new matrix for the quadratic form can be formed.

Step 3: Since the matrix for the quadratic form is now

$$\begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & ? & \cdots & ? \\ \vdots & \vdots & \ddots & \vdots \\ 0 & ? & \cdots & ? \end{pmatrix}$$

We can induct on n by repeating the process of step 1 with the entry a_{22} until we reach a_{nn} . □

This main theorem of quadratic forms shows that every quadratic form is congruent to a diagonal matrix. To illustrate the process of reduction, we look an example.

Example 1.2.7

Find a nice basis for the quadratic form $q\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) = xy + 3yz - 5xz$.

Proof. Using the formula, we construct the matrix of q as

$$A = \begin{pmatrix} 0 & \frac{1}{2} & -\frac{5}{2} \\ \frac{1}{2} & 0 & \frac{3}{2} \\ -\frac{5}{2} & \frac{3}{2} & 0 \end{pmatrix}$$

We start the first change of basis. Notice that $a_{11} = a_{22} = a_{33} = 0$ in the matrix while a_{12} is not. Denote the standard basis by b_1, b_2, b_3 . We perform the first basis change by $\{b'_1 = b_1 + b_2, b'_2 = b_2, b'_3 = b_3\}$. This means that the change of basis matrix from new to old is

$$P' = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

To change A into the new matrix A' , we simply replace r_1 by $r_1 + r_2$ and replace c_1 by $c_1 + c_2$. This gives

$$A' = \begin{pmatrix} 1 & \frac{1}{2} & -1 \\ \frac{1}{2} & 0 & \frac{3}{2} \\ -1 & \frac{3}{2} & 0 \end{pmatrix}$$

We keep track of changing the basis for clarity. We now have $A' = (P')^T A P'$.

The next step is to set the new basis to $b_2'' = b_2' - \frac{1}{2}b_1'$ and $b_3'' = b_3' + b_1'$. This means that the new basis is now $\{b_1 + b_2, \frac{1}{2}(b_2 - b_1), b_1 + b_2 + b_3\}$. The change of basis from this basis back to the old one is now

$$P'' = \begin{pmatrix} 1 & -\frac{1}{2} & 1 \\ 1 & \frac{1}{2} & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

Now the new matrix A'' is formed by replacing r_2 with r_2 by $r_2 - \frac{1}{2}r_1$ and r_3 with $r_3 + r_1$. Noticing that A'' must be symmetric, we need to take the new elements and replace the remaining lower triangular elements so that A'' maintains symmetric. Also observe that the replacement of rows is exactly one changes into the new basis from the previous basis, where $b_2'' = b_2' - \frac{1}{2}b_1'$ etc. Now we have

$$A'' = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{4} & 2 \\ 0 & 2 & -1 \end{pmatrix}$$

We now have $A'' = (P'')^T A' P''$.

Now we perform the next change of basis. We set $b_3''' = b_3'' - \frac{2}{-\frac{1}{4}}b_2'' = b_3'' + 8b_2''$. Now the new basis is $\{b_1 + b_2, \frac{1}{2}(b_2 - b_1), -3b_1 + 5b_2 + b_3\}$. The change of basis matrix is now

$$P''' = \begin{pmatrix} 1 & -\frac{1}{2} & -3 \\ 1 & \frac{1}{2} & 5 \\ 0 & 0 & 1 \end{pmatrix}$$

Similar to the above, we replace r_3 with the same transformation and adjust A''' so that it remains symmetric. Thus now we have

$$A''' = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{4} & 0 \\ 0 & 0 & 15 \end{pmatrix}$$

This means that we are done with $A''' = (P''')^T A'' P'''$. □

In general, this result of diagonalization is different from that of similar matrices. One should not be confusing reduction of congruent matrices into diagonal matrices and reduction of similar matrices into JCF as well as reduction of diagonalizable matrices into diagonal matrices.

1.3 Bilinear Forms

Definition 1.3.1: Bilinear Maps

Let V, W be vector spaces over a field \mathbb{F} . A bilinear map on V and W is a map $\tau : V \times W \rightarrow \mathbb{F}$ such that

- $\tau(a_1 + a_2, w) = a_1\tau(v_1, w) + a_2\tau(v_2, w)$
- $\tau(v, b_1w_1 + b_2w_2) = b_1\tau(v, w_1) + b_2\tau(v, w_2)$

Theorem 1.3.2

Let V and W has basis e_1, \dots, e_n and f_1, \dots, f_m respectively. Let $\tau(v, w)$ be a bilinear map. Then

$$\tau(v, w) = v^T \begin{pmatrix} \tau(e_1, f_1) & \cdots & \tau(e_1, f_m) \\ \vdots & \ddots & \vdots \\ \tau(e_n, f_1) & \cdots & \tau(e_n, f_m) \end{pmatrix} w$$

Proof. Simple to see by expanding the matrix multiplication. □

Proposition 1.3.3

Let A be the matrix of the bilinear map $\tau V \times W \rightarrow \mathbb{F}$ with respect to the basis e_1, \dots, e_n and f_1, \dots, f_m of V and W . Let B be the matrix with respect to the basis of e'_1, \dots, e'_n and f'_1, \dots, f'_m of V and W . Let P and Q be the change of basis matrix where $Pv' = v$ and $Qw' = w$ for $v \in V$ and $w \in W$. Then

$$B = P^T A Q$$

Definition 1.3.4: Bilinear Forms

A bilinear form is a bilinear map that maps $V \times V$ to \mathbb{F} .

Definition 1.3.5: Congruent Matrices

Two matrices are congruent if there exists an invertible matrix P such that $B = P^T A P$.

Note that this definition of congruence in matrices coincides with the definition given with symmetric matrices.

Definition 1.3.6: Types of Bilinear Forms

A bilinear form is said to be

- symmetric if $\tau(v, w) = \tau(w, v)$
- antisymmetric if $\tau(v, w) = -\tau(w, v)$
- positive definite if $\tau(v, v) > 0$ for all $v \in V$.

Proposition 1.3.7

Let $q : V \rightarrow K$ be a function. Then the following are equivalent.

- q is a quadratic form
- $q(cv) = c^2 v$ for $c \in K$ and $v \in V$ and $\tau(v, w) = \frac{1}{2}(q(v+w) - q(v) - q(w))$ is a bilinear form on V
- $q(v) = \tau(v, v)$ is a symmetric bilinear form on V

Proof. Let $q : V \rightarrow K$ be a function.

- (1) \implies (2): Since every term in a quadratic form is quadratic, $q(cv) = c^2 q(v)$ is natural. The fact that $\tau(v, w)$ is bilinear is also easy to check.

- (2) \implies (3): From (2) we know that $\tau(v, v) = q(v)$ by substituting v in the position of w and thus it clearly is a bilinear form. The position of w and v are also interchangeable and thus it is symmetric.
- (3) \implies (1): If $\tau(v, v)$ is a symmetric bilinear form then the matrix of τ is symmetric since $a_{ij} = \tau(e_i, f_j) = \tau(f_j, e_i) = a_{ji}$. Thus $q(v) = \tau(v, v)$ defines a quadratic form.

□

1.4 Sesquilinear Forms

2 Inner Product Spaces

2.1 Norms

Throughout this section, \mathbb{F} means either \mathbb{R} or \mathbb{C} . In general normed vector spaces only perform well in these two fields.

Definition 2.1.1: Norm

Let V be a vector space. A norm on V is a function $\|\cdot\| : V \rightarrow \mathbb{F}$ such that

- $\|x\| \geq 0$ for all $x \in V$ with equality if and only if $x = 0$
- $\|\lambda x\| = |\lambda| \|x\|$ for all $x \in V$ and $\lambda \in \mathbb{F}$
- $\|x + y\| \leq \|x\| + \|y\|$ for all $x, y \in V$

Definition 2.1.2: Normed Vector Space

A normed vector space is a pair $(V, \|\cdot\|)$ where V is a vector space and $\|\cdot\|$ is a norm on V .

Proposition 2.1.3

Every normed vector space is a metric space.

Proof. Can easily be seen that setting $d(x, y) = \|x - y\|$ allows the norm to become a metric. □

Definition 2.1.4: Convex Set

Let V be a vector space. A subset K of V is convex if $x, y \in K$ implies

$$\lambda x + (1 - \lambda)y \in K$$

for $0 \leq \lambda \leq 1$.

Lemma 2.1.5

For every normed vector space, the unit ball $B_1(0) = \{v \in V \mid \|v\| \leq 1\}$ is convex.

Proposition 2.1.6

Let $N : V \rightarrow \mathbb{R}^+$ be a function that satisfies the first two requirements of a norm. If N also satisfies the fact that $\{x \in V \mid N(x) \leq 1\}$ is convex, then N is a norm.

Definition 2.1.7: Isometries

If (X, d_1) and (Y, d_2) are metric spaces, then a distancing preserving map between X and Y is a map

$$f : X \rightarrow Y$$

such that for any $P, Q \in X$, we have

$$d(f(P), f(Q)) = d(P, Q)$$

A bijective distancing preserving map is called an isometry.

2.2 Inner Products

Inner products are only properly defined for vector spaces over \mathbb{R} and \mathbb{C} . From this point onwards we will limit our discussions with $V = \mathbb{R}^n$ or \mathbb{C}^n and $K = \mathbb{R}$ or \mathbb{C} .

Definition 2.2.1: Inner Products

An inner product on V is a function that takes each ordered pair (x, y) of a vector space V to a number $\langle x, y \rangle \in K$ such that

- $\langle x, x \rangle \geq 0$ for all $x \in V$ with equality if and only if $x = 0$.
- $\langle x + z, y \rangle = \langle x, y \rangle + \langle z, y \rangle$ for all x for all $x, y, z \in V$
- $\langle \lambda x, y \rangle = \lambda \langle x, y \rangle$ for all $x, y \in V$ and $\lambda \in K$
- $\langle x, y \rangle = \overline{\langle y, x \rangle}$ for all $x, y \in V$

In this case V is called an inner product space.

Proposition 2.2.2

Let $u, v, w \in V$. Let $\lambda \in K$. Then the following are true.

- $\langle 0, u \rangle = \langle u, 0 \rangle = 0$
- $\langle u, v + w \rangle = \langle u, v \rangle + \langle u, w \rangle$
- $\langle u, \lambda v \rangle = \bar{\lambda} \langle u, v \rangle$

Proof. Let $\langle \cdot, \cdot \rangle$ be an inner product over V .

- $\langle 0, u \rangle = \langle 0, u \rangle + \langle 0, u \rangle$ thus $\langle 0, u \rangle = 0$. $\langle u, 0 \rangle = 0$ can be proven using the below property.
- Let $u, v, w \in V$. Then

$$\begin{aligned} \langle u, v + w \rangle &= \overline{\langle u, v + w \rangle} \\ &= \overline{\langle v + w, u \rangle} \\ &= \overline{\langle v, u \rangle + \langle w, u \rangle} \\ &= \overline{\langle v, u \rangle} + \overline{\langle w, u \rangle} \\ &= \langle u, v \rangle + \langle u, w \rangle \end{aligned}$$

- Applying the same technique as above gives the desired result.

□

Proposition 2.2.3

Every inner product induces a norm.

Proof. Define the norm to be $\|x\| = \sqrt{\langle x, x \rangle}$.

□

Proposition 2.2.4: Cauchy-Schwartz Inequality

For all $x, y \in V$,

$$|\langle x, y \rangle| \leq \|x\| \cdot \|y\|$$

with equality if and only if $y = \lambda x$ for some $\lambda \in \mathbb{F}$.

Proof. Let $z = x - \frac{|\langle x, y \rangle|}{\|y\|^2} y$. We have $\|z\| \geq 0$.

$$\begin{aligned} \|z\|^2 &= \left\langle x - \frac{|\langle x, y \rangle|}{\|y\|^2} y, x - \frac{|\langle x, y \rangle|}{\|y\|^2} y \right\rangle \\ &= \langle x, x \rangle - 2 \left\langle x, \frac{|\langle x, y \rangle|}{\|y\|^2} y \right\rangle + \left\langle \frac{|\langle x, y \rangle|}{\|y\|^2} y, \frac{|\langle x, y \rangle|}{\|y\|^2} y \right\rangle \\ &= \langle x, x \rangle - 2 \frac{|\langle x, y \rangle|^2}{\|y\|^2} + \frac{|\langle x, y \rangle|^2}{\|y\|^4} \langle y, y \rangle \\ &= \langle x, x \rangle - 2 \frac{|\langle x, y \rangle|^2}{\|y\|^2} + \frac{|\langle x, y \rangle|^2}{\|y\|^2} \\ &= \|x\|^2 - \frac{|\langle x, y \rangle|^2}{\|y\|^2} \end{aligned}$$

Thus we have

$$\begin{aligned} \|x\|^2 &\geq \frac{|\langle x, y \rangle|^2}{\|y\|^2} \\ \|x\| &\geq \frac{|\langle x, y \rangle|}{\|y\|} && \text{(Since they are all positive)} \\ \|x\| \cdot \|y\| &\geq |\langle x, y \rangle| \end{aligned}$$

Note that we have equality if and only if $\|z\| = 0$, meaning $x = \frac{|\langle x, y \rangle|}{\|y\|^2} y$. We are done by taking $\lambda = \frac{\|y\|^2}{|\langle x, y \rangle|}$. \square

3 Orthogonality

3.1 Orthogonal Vectors

Definition 3.1.1: Orthogonal Vectors

Two vectors $u, v \in V$ an inner product space are called orthogonal if $\langle u, v \rangle = 0$.

Corollary 3.1.2

Let V be an inner product space. Let $u, v \in V$. Then the following are true.

- 0 is orthogonal to every vector in V
- 0 is the only vector in V that is orthogonal to itself.

Proof. Easy check involving properties of the inner product. □

Theorem 3.1.3: Pythagorean Theorem

Suppose that $u, v \in V$ an inner product space and u, v are orthogonal. Then

$$\|u + v\|^2 = \|u\|^2 + \|v\|^2$$

Proof. The norm here is induced by the inner product and thus $\|x\| = \sqrt{\langle x, x \rangle}$. We have that

$$\begin{aligned} \|u + v\|^2 &= \langle u + v, u + v \rangle \\ &= \langle u, u \rangle + \langle u, v \rangle + \langle v, u \rangle + \langle v, v \rangle \\ &= \langle u, u \rangle + \langle v, v \rangle \\ &= \|u\|^2 + \|v\|^2 \end{aligned}$$

□

Theorem 3.1.4: Orthogonal Decomposition

Let $u, v \in V$ and $v \neq 0$. Set $c = \frac{\langle u, v \rangle}{\|v\|^2}$ and $w = u - \frac{\langle u, v \rangle}{\|v\|^2}v$. Then $\langle w, v \rangle = 0$ and $u = cv + w$.

Proof. The fact that $u = cv + w$ is natural so we only have to prove that $\langle w, v \rangle = 0$. We have that

$$\begin{aligned} \langle w, v \rangle &= \left\langle u - \frac{\langle u, v \rangle}{\|v\|^2}v, v \right\rangle \\ &= \langle u, v \rangle - \left\langle \frac{\langle u, v \rangle}{\|v\|^2}v, v \right\rangle \\ &= \langle u, v \rangle - \frac{\langle u, v \rangle}{\|v\|^2} \langle v, v \rangle \\ &= \langle u, v \rangle - \frac{\langle u, v \rangle}{\|v\|^2} \|v\|^2 \\ &= 0 \end{aligned}$$

Thus we are done. □

Proposition 3.1.5

Every orthonormal list of vectors are linearly independent.

Proof. Suppose that v_1, \dots, v_n are orthonormal. We want to show that $v_n = \sum_{k=1}^{n-1} a_k v_k$ implies $a_1 = \dots = a_{n-1} = 0$. Then

$$\langle v_n, v_i \rangle = \sum_{k=1}^{n-1} a_k (v_k \cdot v_i) = a_i \|v_i\|^2$$

for $i \in \{1, \dots, n-1\}$ since v_1, \dots, v_{n-1} are orthonormal. But since $\langle v_n, v_k \rangle = 0$ we must have $a_i = 0$. This means that $a_1 = \dots = a_{n-1} = 0$ and thus v_1, \dots, v_n are linearly independent. \square

3.2 Orthonormal Basis**Definition 3.2.1: Orthonormal Basis**

A basis v_1, \dots, v_n of an inner product space V with $\dim(V) = n$ is called orthonormal if

- $\|b_i\| = 1$ for $i \in \{1, \dots, n\}$
- $\langle b_i, b_j \rangle = \delta_{ij}$ for $i, j \in \{1, \dots, n\}$

Proposition 3.2.2

The orthonormal basis is indeed a basis for an inner product space V .

Proof. Since lists of orthonormal vectors are linearly independent and there are n vectors, they must also span V and thus is a basis. \square

Theorem 3.2.3

Let b_1, \dots, b_n be an orthonormal basis and $v = \sum_{k=1}^n a_k b_k$. Then

$$\|v\|^2 = \sum_{k=1}^n |a_k|^2$$

Proof. We have that

$$\begin{aligned} \|v\|^2 &= \left\langle \sum_{k=1}^n a_k b_k, \sum_{k=1}^n a_k b_k \right\rangle \\ &= \sum_{i=1}^n \sum_{j=1}^n a_i a_j (b_i \cdot b_j) \\ &= \sum_{i=1}^n \sum_{j=1}^n a_i a_j \delta_{ij} \\ &= \sum_{k=1}^n |a_k|^2 \end{aligned}$$

and we are done. \square

Proposition 3.2.4

Let b_1, \dots, b_n be an orthonormal basis of V . Then

$$v = \sum_{k=1}^n \langle v, b_k \rangle b_k$$

Proof. Applying the inner product with b_i for each $i \in \{1, \dots, n\}$ gives $a_i = \langle v, b_i \rangle$ since $\langle b_i, b_k \rangle = 0$ for any $k \neq i$. Thus if $v = \sum_{k=1}^n a_k b_k$ then $v = \sum_{k=1}^n \langle v, b_k \rangle b_k$ and we are done. \square

Theorem 3.2.5: Gram-Schmidt Procedure

Let v_1, \dots, v_m be a list of linearly independent vectors of V . Let $b_1 = \frac{v_1}{\|v_1\|}$. For $i = 2, \dots, m$. Define

$$b_i = \frac{v_i - \sum_{k=0}^{i-1} \langle v_i, b_k \rangle b_k}{\|v_i - \sum_{k=0}^{i-1} \langle v_i, b_k \rangle b_k\|}$$

Then b_1, \dots, b_m are orthonormal and has the same span as v_1, \dots, v_m .

Theorem 3.2.6

Every finite dimensional inner product space has an orthonormal basis.

Proof. By the Gram-Schmidt procedure, every basis can be transformed into an orthonormal basis. \square

3.3 Orthogonal Complements**Definition 3.3.1: Orthogonal Complement**

Let U be a subset of an inner product space V . The orthogonal complement of U is defined as

$$U^\perp = \{v \in V \mid \langle v, u \rangle = 0 \text{ for all } u \in U\}$$

Proposition 3.3.2

Let U be a subset of V .

- U is a subspace of V if and only if U^\perp is a subspace of V .
- $\{0\}^\perp = V$
- $V^\perp = \{0\}$
- $U \cap U^\perp = \{0\}$
- If $W \subseteq U$, then $U^\perp \subseteq W^\perp$

Proposition 3.3.3

Let U be a finite dimensional subspace of V . Then

$$U = (U^\perp)^\perp$$

Theorem 3.3.4

Suppose U is a finite dimensional subspace of V . Then

$$V = U \oplus U^\perp$$

and

$$\dim(U) + \dim(U^\perp) = \dim(V)$$

3.4 Orthogonal Maps**Definition 3.4.1: Orthogonal Maps**

A linear map $T : V \rightarrow V$ is said to be orthogonal if

$$\langle T(v), T(w) \rangle = \langle v, w \rangle$$

for all $v, w \in V$.

One can think of orthogonal maps as orthogonality preserving maps. If $\langle v, w \rangle = 0$ then $\langle T(v), T(w) \rangle = 0$ which means orthogonality is preserved.

Proposition 3.4.2

Let $T : V \rightarrow V$ be a linear map over an inner product space V . Let A represent the linear map T . Then the following are equivalent.

- T is orthogonal
- A is orthogonal
- T maps orthonormal bases to orthonormal bases

Proof. Suppose that $T : V \rightarrow V$ is represented by A .

- (1) \iff (2): We have that $\langle T(v), T(w) \rangle = v^T A^T A w$. Thus it is equal to $v^T w$ if and only if $A^T A = I$.
- (1) \iff (3): Suppose that T is orthogonal. Suppose that $\{b_1, \dots, b_n\}$ is orthonormal. Then $\langle T(b_i), T(b_j) \rangle = \langle b_i, b_j \rangle = 0$ for $i, j \in \{1, \dots, n\}$. Thus $\{T(b_1), \dots, T(b_n)\}$ is orthogonal. But they are also orthonormal since $\|T(b_i)\|^2 = \langle T(b_i), T(b_i) \rangle = \langle b_i, b_i \rangle = \|b_i\|^2 = 1$. This means that $\|T(b_i)\| = 1$ for $i \in \{1, \dots, n\}$.

Now suppose that T maps orthonormal bases to orthonormal bases. Then if $\{b_1, \dots, b_n\}$ is orthonormal, we have

$$\begin{aligned} \langle T(v), T(w) \rangle &= \left\langle \left(\sum_{k=1}^n v_k T(b_k) \right), \left(\sum_{k=1}^n w_k T(b_k) \right) \right\rangle \\ &= \sum_{k=1}^n (v_k w_k) \langle T(b_k), T(b_k) \rangle && (\langle T(b_i), T(b_j) \rangle = 0 \text{ if } i \neq j) \\ &= \sum_{k=1}^n v_k w_k \\ &= \langle v, w \rangle \end{aligned}$$

Thus we are done. □

4 Orthogonality in \mathbb{R}^n

4.1 Reduction of Quadratic Forms over \mathbb{R}

Orthogonality is interesting for real matrices because the notion of similarity and congruence coincide under orthogonality. Notice that being similar and congruent to a diagonal matrix at the same time means that there exists an invertible P such that $P^T A P = P^{-1} A P = D$.

In the remaining sections we treat the adjugate in the case of \mathbb{R}^n and save the case for \mathbb{C}^n in another chapter. Then V in the remaining sections will only denote \mathbb{R}^n .

Definition 4.1.1: Euclidean Vector Space

An Euclidean vector space is \mathbb{R}^n equipped with an inner product.

Proposition 4.1.2

A function $b : V \times V \rightarrow \mathbb{R}$ is an inner product over \mathbb{R} if and only if b is bilinear and positive definite.

Lemma 4.1.3: Polarization Identity

For $x, y \in \mathbb{R}^n$,

$$\langle x, y \rangle = \frac{1}{4} \|x + y\|^2 - \frac{1}{4} \|x - y\|^2$$

Proof. Simple proof using the fact that $\|x\|^2 = \langle x, x \rangle$. □

Proposition 4.1.4: Sylvester's Law of Inertia

A quadratic form q over \mathbb{R} has the form

$$q(x_1, \dots, x_n) = \sum_{k=1}^t x_k^2 - \sum_{k=1}^u x_k^2$$

where $t + u = \text{rank}(q)$. The pair (t, u) is called the signature of q . This reduced quadratic form is also unique in the sense that the number of positives and number of negatives of any two reduced forms are the same.

Moreover, every symmetric matrix is congruent to a matrix of the form

$$\begin{pmatrix} I_t & 0 & 0 \\ 0 & -I_u & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

where (t, u) is the signature of the quadratic form.

Proof. We saw in theorem 1.2.7 that every quadratic form can be expressed as

$$q(y_1, \dots, y_n) = \sum_{k=1}^n c_k y_k^2$$

By doing a basis change with $b'_k = \frac{1}{\sqrt{c_k}} b_k$ whenever $c_k \neq 0$ will give us the above sum. For those that have $c_k = 0$, the terms vanish and are exactly in the kernel of the quadratic form thus $t + u = \text{rank}(q)$.

The second part is direct from the fact that same quadratic forms with different matrix representations imply their representations are similar. □

4.2 Reduction of Inner Products

Definition 4.2.1: Dot Product

The dot product in \mathbb{R}^n is defined to be the inner product given by

$$x \cdot y = x_1y_1 + \cdots + x_ny_n$$

in standard basis.

Theorem 4.2.2

Let $\langle \cdot, \cdot \rangle$ be an inner product on a real vector space V . Then there exists an basis b_1, \dots, b_n of V such that the inner product, when represented in the orthonormal basis, takes the form of exactly the dot product.

Proof. Let $\langle \cdot, \cdot \rangle$ be our inner product in question. Define a quadratic form by

$$q(x) = \langle x, x \rangle = \|x\|^2$$

We know that this quadratic form can be reduced to

$$q(x_1, \dots, x_n) = x_1^2 + \cdots + x_t^2 - x_{t+1}^2 - \cdots - x_{t+u}^2$$

Now we must have $u = 0$ since if $u > 0$, then the basis vector b_{t+1} satisfies $q(b_{t+1}) = -1$ and $q(b_{t+1}) = \langle b_{t+1}, b_{t+1} \rangle$ which is a contradiction since inner products are positive definite. Also $t = n$ since if $t < n$, then $\langle b_{t+1}, b_{t+1} \rangle = 0$ which is again a contradiction.

Using polarization, we see that $\langle x, y \rangle = x_1y_1 + \cdots + x_ny_n$ in that basis and we are done. \square

With this theorem, we know that any inner product can be expressed in the dot product as long as it is under a suitable basis. Thus we now reduce our discussion to only the dot product, as our standard inner product in \mathbb{R}^n .

The below theorem, while unrelated to the reduction of inner products, is a result of Gram-schmidt process that is only true for real matrices.

Theorem 4.2.3: QR Decomposition

Let A be an $n \times n$ matrix over \mathbb{R} . Then we can write $A = QR$ where Q is orthogonal and R is upper triangular.

Proof. We split the matrices into two cases. Firstly consider the case where A is invertible. We can treat A as a change of basis matrix from the basis $\{a_1, \dots, a_n\}$ where a_k is the column of A for $k \in \{1, \dots, n\}$. This change of basis matrix takes $\{a_1, \dots, a_n\}$ to $\{e_1, \dots, e_n\}$ which is the standard basis. Apply the Gram-schmidt process to $\{a_1, \dots, a_n\}$ to get $\{b_1, \dots, b_n\}$ which is an orthonormal basis. Let Q be the change of basis matrix from $\{b_1, \dots, b_n\}$ to $\{e_1, \dots, e_n\}$. Let R be the change of basis matrix from $\{a_1, \dots, a_n\}$ to $\{b_1, \dots, b_n\}$. Then clearly $A = QR$. We just have to show that Q is orthogonal and R is upper triangular.

Q being orthonormal is trivial since columns of Q are just b_1, \dots, b_n . Using the Gram-schmidt process, we can see that the change of basis from $\{b_1, \dots, b_n\}$ to $\{a_1, \dots, a_n\}$, each b_k is only affected by a_1, \dots, a_k from the old basis. This means that the change of basis matrix must be upper triangular and its inverse must also be upper triangular.

Now we also have the case when A is not invertible. \square

We now give an example of QR decomposition, in conjunction with the Gram-schmidt procedure.

Example 4.2.4

Find the QR decomposition of

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

Proof. A quick check shows that A is invertible. Let a_1, a_2, a_3 be the columns of A . We start the Gram-schmidt process by taking the new basis $f_1 = \frac{a_1}{\|a_1\|} = \begin{pmatrix} -\frac{1}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} \\ 0 \end{pmatrix}$. We also need to keep track on the change of basis matrix. We have that $a_1 = \sqrt{5}f_1$.

For the next step, we find that

$$\begin{aligned} f_2 &= \frac{a_2 - (a_2 \cdot f_1)f_1}{\|a_2 - (a_2 \cdot f_1)f_1\|} \\ &= \frac{a_2}{\|a_2\|} \\ &= \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix} \end{aligned}$$

This means that $a_2 = 2f_2$.

Finally, we have that

$$\begin{aligned} f_3 &= \frac{a_3 - (a_3 \cdot f_1)f_1 - (a_3 \cdot f_2)f_2}{\|a_3 - (a_3 \cdot f_1)f_1 - (a_3 \cdot f_2)f_2\|} \\ &= \frac{a_3 - 2f_2}{\|a_3 - 2f_2\|} \\ &= \frac{a_3 - 2f_2}{\sqrt{5}} \\ &= \begin{pmatrix} -\frac{2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \\ 0 \end{pmatrix} \end{aligned}$$

We have that $a_3 = 2f_2 + \sqrt{5}f_3$. Combining everything together, we have that

$$\begin{pmatrix} -1 & 0 & -2 \\ 2 & 0 & -1 \\ 0 & -2 & -2 \end{pmatrix} = \begin{pmatrix} -\frac{1}{\sqrt{5}} & 0 & -\frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & 0 & -\frac{1}{\sqrt{5}} \\ 0 & -1 & 0 \end{pmatrix} \begin{pmatrix} \sqrt{5} & 0 & 0 \\ 0 & 2 & 2 \\ 0 & 0 & \sqrt{5} \end{pmatrix}$$

□

4.3 Adjoints**Proposition 4.3.1**

Let V be an inner product space and $T : V \rightarrow V$ be a linear map. Then there exists a unique linear map $T^* : V \rightarrow V$ such that

$$\langle T(v), w \rangle = \langle v, T^*(w) \rangle$$

for all $v, w \in V$.

Proof. Let T be a linear map. Then the function $\tau(v, w) = \langle T(v), w \rangle$ is a bilinear form since the inner product is bilinear. But we know that bilinear forms can be represented by a matrix multiplication, namely $\tau(v, w) = v^T A w$ where A is defined as in theorem 1.3.2. Then treating $A w$ as the linear map $T^*(w)$ and since $v^T w = v \cdot w$, we have that $\tau(v, w) = v \cdot T^*(w)$ thus proving existence. Uniqueness follows naturally by construction of the matrix A . \square

Definition 4.3.2: Adjoint of a Linear Map

T^* in the above case is called the adjoint of T .

Definition 4.3.3: Self-Adjoint

A linear map $T : V \rightarrow V$ is said to be self-adjoint if $T^* = T$

Proposition 4.3.4

Let T be a linear map represented by a matrix A . Then the following are true.

- T is self-adjoint if and only if A is symmetric.
- T is orthogonal if and only if $T^* = T^{-1}$.

Proof. Let T be self-adjoint. Then $Av \cdot w = v \cdot Aw$ for all $v, w \in V$. \square

Proposition 4.3.5

Let $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ be self-adjoint. Then every eigenvalues of T are real.

Proof. Suppose that $T(v) = Av$ where A is a representation of T . Suppose that λ is an eigenvalue of T . Then $Av = \lambda v$ for some eigenvector $v \in \mathbb{R}^n$. Then taking complex conjugates give

$$\overline{Av} = \overline{\lambda v}$$

$$A\bar{v} = \bar{\lambda}\bar{v}$$

Taking the transpose of $Av = \lambda v$ gives $v^T A^T = \lambda v^T$ and $v^T A = \lambda v^T$. Multiplying \bar{v} on both sides give

$$v^T A\bar{v} = \lambda v^T \bar{v}$$

$$\bar{\lambda} v^T \bar{v} = \lambda v^T \bar{v}$$

But $v^T \bar{v} = v_1 \bar{v}_1 + \cdots + v_n \bar{v}_n = |v_1|^2 + \cdots + |v_n|^2$ which is 0 if and only if v is 0. Since eigenvectors are taken to be nonzero, we must have $\lambda = \bar{\lambda}$ and thus λ is real. \square

Theorem 4.3.6

Let $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ be a linear map on the inner product space \mathbb{R}^n that is self-adjoint. Then there exists an orthonormal basis consisting of eigenvectors of T .

Equivalently, for every quadratic form q on V , there exists an orthonormal basis b_1, \dots, b_n such

that

$$q(x_1, \dots, x_n) = \sum_{k=1}^n c_k x_k^2$$

for some $c_1, \dots, c_n \in \mathbb{R}$.

Proof. Notice that the two statements are exactly the same and I will omit the reason.

We prove by induction on n . Suppose that the theorem holds for $n - 1$. Let T be the linear map. By the above we know that T has at least one eigenvalue in \mathbb{R} , say λ_1 . Let f_1 be the corresponding eigenvector with magnitude 1.

Consider the orthogonal complement $W = \{w \in V \mid w \cdot f_1 = 0\}$ of f_1 . Since W is the kernel of the linear map $S : V \rightarrow \mathbb{R}$ defined by $S(v) = v \cdot f_1$, it is a subspace of V dimension $n - 1$. I claim that $T(W) \subseteq W$.

Let $w \in W$. We have

$$T(w) \cdot f_1 = w \cdot T(f_1) = w \cdot \lambda_1 f_1 = 0$$

by self-adjoint. Thus we have shown that $T(W) \subseteq W$.

Applying the induction hypothesis on W , we have an orthonormal basis f_2, \dots, f_n of W consisting of eigenvectors of T . By definition, f_1, \dots, f_n is an orthonormal basis and we are done. \square

Notice that the above two statements are also equivalent to saying that every real symmetric matrix is congruent and similar to a diagonal matrix.

Proposition 4.3.7

If $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is self-adjoint, and λ, μ are distinct eigenvalues of T with eigenvectors v, w , then $v \cdot w = 0$.

Proof. We have that

$$\begin{aligned} v^T A w &= v \cdot A w \\ &= v^T \mu w \\ &= \mu(v \cdot w) \end{aligned}$$

and

$$\begin{aligned} v^T A w &= v^T A^T w \\ &= (A v)^T w \\ &= (\lambda v)^T w \\ &= \lambda v^T w \\ &= \lambda(v \cdot w) \end{aligned}$$

Comparing the two results, we have that $(\mu - \lambda)(v \cdot w) = 0$ and thus $v \cdot w = 0$. \square

The proposition will prove itself to be useful in finding an orthonormal basis for self-adjoint linear maps.

4.4 Singular Value Decomposition

Theorem 4.4.1: Singular Value Decomposition for Linear Maps

Let $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ be a linear map of rank k . Then there exists unique positive numbers $\gamma_1 \geq \gamma_2 \geq \dots \geq \gamma_k \geq 0$ and orthonormal bases of \mathbb{R}^n and \mathbb{R}^m such that the matrix of T with respect to these bases is

$$\begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix}$$

where $D = \text{diag}(\gamma_1, \dots, \gamma_k)$. In fact, the γ_i are exactly the nonzero eigenvalues of T^*T , each one appearing as many times as the dimension of the corresponding eigenspace.

Theorem 4.4.2: Singular Value Decomposition for Matrices

Let $A_{m \times n}$ be a matrix. There exists unique singular values $\gamma_1 \geq \gamma_2 \geq \dots \gamma_k \geq 0$ where $k = \text{rank}(A)$, and orthogonal matrices P, Q such that

$$\begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix} = P^T A Q$$

where $D = \text{diag}(\gamma_1, \dots, \gamma_k)$.

We present an example of singular value decomposition for illustration.

Example 4.4.3

Find the singular value decomposition of the matrix

$$A = \begin{pmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{pmatrix}$$

Proof. Step 1: We compute the singular values of A , which is just the squareroot of the eigenvalues of $A^T A$. Now

$$A^T A = \begin{pmatrix} 80 & 100 & 40 \\ 100 & 170 & 140 \\ 40 & 140 & 200 \end{pmatrix}$$

We have that $c_{A^T A}(x) = x(360 - x)(90 - x)$. This means that the singular values are $\gamma_1 = \sqrt{360} = 6\sqrt{10}$ and $\gamma_2 = \sqrt{90} = 3\sqrt{10}$. Now we want P and Q such that

$$P^T A Q = \begin{pmatrix} 6\sqrt{10} & 0 & 0 \\ 0 & 3\sqrt{10} & 0 \end{pmatrix}$$

Step 2: We find the orthonormal eigenvectors of $A^T A$ so that it forms the matrix Q . This gives

$$Q = \begin{pmatrix} \frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & -\frac{1}{3} & -\frac{2}{3} \\ \frac{2}{3} & \frac{2}{3} & \frac{1}{3} \end{pmatrix}$$

Step 3: We now calculate P by finding the image of the above basis under A , and dividing it with the nonzero singular values. This gives

$$\begin{aligned} P &= \left(\frac{1}{6\sqrt{10}} A b_1 \quad \frac{1}{3\sqrt{10}} A b_2 \right) \\ &= \begin{pmatrix} \frac{3}{\sqrt{10}} & \frac{1}{\sqrt{10}} \\ \frac{1}{\sqrt{10}} & -\frac{3}{\sqrt{10}} \end{pmatrix} \end{aligned}$$

This means that we are done. □