Math 245

James Yu

October 5, 2017

Contents

Week 1

September 8 - September 13

1.1 Inner Product

Definition 1.1.1 (Inner Product Space)**.** An *inner product space* (over $\mathbb{C}$) is a *vector space* V and a function:

$$\langle \cdot, \cdot \rangle : V \times V \to \mathbb{C}$$

satisfying:

i. $\langle \vec{x} + \vec{y}, \vec{z} \rangle = \langle \vec{x}, \vec{z} \rangle + \langle \vec{y}, \vec{z} \rangle$ for all $\vec{x}, \vec{y}, \vec{z} \in V$

ii. $\langle c\vec{x}, \vec{y} \rangle = c \langle \vec{x}, \vec{y} \rangle$ for all $\vec{x}, \vec{y} \in V$ and $c \in \mathbb{C}$

iii. $\langle \vec{x}, \vec{y} \rangle = \overline{\langle \vec{y}, \vec{x} \rangle}$ for all $\vec{x}, \vec{y} \in V$

iv. $\langle \vec{x}, \vec{x}, \in \rangle \mathbb{R}_{>0}$ if $\vec{x} \neq \vec{0}$, $\langle \vec{x}, \vec{x} \rangle = 0$ otherwise

Example 1.1.1.1 (Standard Complex Inner Product)**.**

$$V = \mathbb{C}^n \qquad\qquad \langle (x_1, \dots, x_n), (y_1, \dots, y_n) \rangle = x_1\overline{y_1} + \dots + x_n\overline{y_n}$$

Properties *i*, *ii*, and *iii* clearly hold. For *iv*, for any $\vec{x} = (a_1 + b_1 i, \dots, a_n + b_n i)$

$$\langle \vec{x}, \vec{x} \rangle = (a_1 + b_1 i)\overline{(a_1 + b_1 i)} + \dots + (a_n + b_n i)\overline{(a_n + b_n i)} \tag{1.1}$$
$$= a_1^2 + b_1^2 + \dots + a_n^2 + b_n^2 \tag{1.2}$$

This is the *standard complex inner product.* If we replace $\mathbb{C}^n$ with $\mathbb{R}^n$ then we get the *standard real inner product* (dot product).

Example 1.1.1.2 (L^2 Inner Product)**.**

$$V = C([0, 1]) \qquad\qquad \langle f, g \rangle = \int_0^1 f\overline{g}$$

This is called the L^2 inner product on V

Example 1.1.1.3 (Frobenius Inner Product)**.**

$$V = M_n(\mathbb{C}) \qquad\qquad \langle A, B\rangle = \operatorname{tr}\left(A\overline{B^\intercal}\right)$$

This is called the *Frobenius inner product* on V. It satisfies *iv* because, for $A = (a_{ij})$, $B = (b_{ij})$ we have:

$$\operatorname{tr}\left(A\overline{B^\intercal}\right) = \sum a_{ij}\overline{b_i j}$$

which is equivalent to the *standard complex inner product.*

1.2 Cauchy-Bunyakovsky-Schwarz Inequality

Definition 1.2.1 (Length)**.** If $\vec{v}$ is a vector in an *inner product space*, the *length* of $\vec{v}$ is:

$$\|\vec{v}\| = \sqrt{\langle \vec{v}, \vec{v}\rangle}$$

Theorem 1.2.2 (Cauchy-Schwarz)**.** *Let $\vec{x}, \vec{y} \in V$ be vectors in an* inner product space*, then:*

$$|\langle \vec{x}, \vec{y}\rangle| \leq \|\vec{x}\| \, \|\vec{y}\|$$

Proof. If $\vec{y} = \vec{0}$, this is trivial. Otherwise, for any $c \in \mathbb{C}$

$$\begin{aligned} 0 &\leq \|\vec{x} - c\vec{y}\|^2 && (1.3)\\ &\leq \|\vec{x}\|^2 - \bar{c}\langle \vec{x}, \vec{y}\rangle - c\langle \vec{y}, \vec{x}\rangle + c\bar{c}\,\|y\|^2 && (1.4)\end{aligned}$$

So, let $c = \dfrac{\langle \vec{x}, \vec{y}\rangle}{\|y\|^2}$:

$$\begin{aligned} 0 &\leq \|\vec{x}\|^2 - \frac{\overline{\langle \vec{x}, \vec{y}\rangle}}{\|y\|^2}\langle \vec{x}, \vec{y}\rangle - \frac{\langle \vec{x}, \vec{y}\rangle}{\|y\|^2}\langle \vec{y}, \vec{x}\rangle + \frac{\langle \vec{x}, \vec{y}\rangle\,\overline{\langle \vec{x}, \vec{y}\rangle}}{\|y\|^4}\|y\|^2 && (1.5)\\ &\leq \|\vec{x}\|^2 - \frac{|\langle \vec{x}, \vec{y}\rangle|^2}{\|y\|^2} - \frac{|\langle \vec{x}, \vec{y}\rangle|^2}{\|y\|^2} + \frac{|\langle \vec{x}, \vec{y}\rangle|^2}{\|y\|^2} && (1.6)\\ |\langle \vec{x}, \vec{y}\rangle|^2 &\leq \|\vec{x}\|^2\,\|\vec{y}\|^2 && (1.7)\\ |\langle \vec{x}, \vec{y}\rangle| &\leq \|\vec{x}\|\,\|\vec{y}\| && (1.8)\end{aligned}$$

□

Remark. We can define the angle between $\vec{x}, \vec{y}$ as $\cos^{-1}\dfrac{|\langle \vec{x}, \vec{y}\rangle|}{\|\vec{x}\|\,\|\vec{y}\|}$.

1.3 Orthogonality

Definition 1.3.1 (Orthogonality)**.** Two vectors $\vec{x}, \vec{y}$ are *orthogonal* if and only if $\langle \vec{x}, \vec{y} \rangle = 0$.

Definition 1.3.2 (Unit vector)**.** A *unit vector* is a vector of length 1.

Definition 1.3.3 (Orthogonal Set)**.** An *orthogonal set* is a set S where $\langle \vec{x}, \vec{y} \rangle = 0$ for all $\vec{x}, \vec{y} \in S$, $\vec{x} \neq \vec{y}$.

Definition 1.3.4 (Orthonormal Set)**.** An *orthonormal set* is an *orthogonal set* in which each vector is a *unit vector*.

Example 1.3.4.1 (Standard Basis)**.** The standard basis in $\mathbb{R}^n$ is *orthonormal*

Theorem 1.3.5 (Orthonormal Coordinates)**.** *Let* $\{\vec{v}_1, \ldots, \vec{v}_n\}$ *be an* orthogonal basis *of an* inner product space V*. Then for any* $\vec{x} \in V$ *we have:*

$$\vec{x} = \sum_{i=1}^{n} \frac{\langle \vec{x}_i, \vec{v}_i \rangle}{\|\vec{v}_i\|^2} \vec{v}_i$$

Proof. Write $\vec{x} = a_1\vec{v}_1 + \ldots + a_n\vec{v}_n$. Then, for any i:

$$\begin{aligned} \langle \vec{x}, \vec{v}_i \rangle &= \langle a_1\vec{v}_1 + \ldots + a_n\vec{v}_n, \vec{v}_i \rangle && (1.9) \\ &= a_1 \langle \vec{v}_1, \vec{v}_i \rangle + \ldots + a_n \langle \vec{v}_n, \vec{v}_i \rangle && (1.10) \\ &= a_i \langle \vec{v}_i, \vec{v}_i \rangle && (1.11) \\ a_i &= \frac{\langle \vec{x}, \vec{v}_i \rangle}{\|\vec{v}_i\|^2} && (1.12) \end{aligned}$$

□

Remark. If $\{\vec{v}_1, \ldots, \vec{v}_n\}$ is *orthonormal*, then $\vec{x} = \sum_{i=1}^{n} \langle \vec{x}_i, \vec{v}_i \rangle \vec{v}_i$

Remark. The $\vec{v}_i$ coordinate of $\vec{x}$ depends only $\vec{x}$ and $\vec{v}_i$. It does not depend on any other vectors in the basis.

Remark. In finite dimensions, *inner product spaces* always have *orthnormal bases.*

Theorem 1.3.6 (Orthogonal $\implies$ Linear Independence)**.** *Let* S *be an* orthogonal set *of non-zero vectors, then* S *is linearly independent.*

Proof. For any $\vec{v}_1, \ldots, \vec{v}_n \in S$, set $a_1\vec{v}_1 + \ldots + a_n\vec{v}_n = \vec{0}$. By similar construction as 1.3.5, we can show for any i that:

$$a_i \langle \vec{v}_i, \vec{v}_i \rangle = 0$$

Since $\vec{v}_i \neq 0$ by assumption, $a_i = 0$ for all i. □

1.4 Gram-Schmidt Procedure

Given a basis $\{\vec{w}_1, \dots \vec{w}_n\}$ for a (finite dimensional) *inner product space* V, the Gram-Schmidt gives an *orthogonal basis* for V as follows:

$$\text{Step ①} \quad \text{Set } \vec{v}_1 = \vec{w}_1$$

$$\text{Step ②} \quad \text{Set } \vec{v}_2 = \vec{w}_2 - \frac{\langle \vec{w}_2, \vec{v}_1 \rangle}{\|\vec{v}_1\|^2} \vec{v}_1$$

$$\dots$$

$$\text{Step ⓘ} \quad \text{Set } \vec{v}_i = \vec{w}_i - \sum_{j=1}^{i-1} \frac{\langle \vec{w}_2, \vec{v}_i \rangle}{\|\vec{v}_i\|^2} \vec{v}_i$$

Claim 1.4.1. $\{\vec{v}_1, \dots, \vec{v}_n\}$ is an *orthogonal basis* of V

Proof. We first check that $\{\vec{v}_1, \dots, \vec{v}_n\}$ is *orthogonal*.

We proceed by induction on i. If $n = 1$, we are vacuously done.

Otherwise, assume that $\{\vec{v}_1, \dots, \vec{v}_i\}$ is *orthogonal*. For any $1 \leq j \leq i$ we have:

$$\langle \vec{v}_{i+1}, \vec{v}_j \rangle = \left\langle \vec{w}_{i+1} - \sum_{k=1}^{i} \frac{\langle \vec{w}_{i+1}, \vec{v}_k \rangle}{\|\vec{v}_k\|^2} \vec{v}_k, \vec{v}_j \right\rangle \tag{1.13}$$

$$= \left\langle \vec{w}_{i+1} - \frac{\langle \vec{w}_{i+1}, \vec{v}_j \rangle}{\|\vec{v}_j\|^2} \vec{v}_j, \vec{v}_j \right\rangle \tag{1.14}$$

$$= \langle \vec{w}_{i+1}, \vec{v}_j \rangle - \frac{\langle \vec{w}_{i+1}, \vec{v}_j \rangle}{\|\vec{v}_j\|^2} \|\vec{v}_j\|^2 \tag{1.15}$$

$$= \langle \vec{w}_{i+1}, \vec{v}_j \rangle - \langle \vec{w}_{i+1}, \vec{v}_j \rangle \tag{1.16}$$

$$= 0 \tag{1.17}$$

Furthermore, $\vec{v}_i \neq \vec{0}$. For $i = 1$ we have $\vec{v}_1 = \vec{w}_1 \neq \vec{0}$ by assumption. Otherwise, we have $\vec{v}_i = \vec{w}_i - \vec{x}$ for $x \in \operatorname{span}\{\vec{w}_1, \dots, \vec{w}_{i-1}\}$. Thus $\vec{v}_i$ is a nonzero linear combination of $\{\vec{w}_1, \dots, \vec{w}_i\}$ and is therefore non-zero.

Thus $\{\vec{v}_1, \dots, \vec{v}_n\}$ is a set of n vectors in an n-dimensional space that are orthogonal and nonzero. By 1.3.6 they are linearly independent, and thus a basis of V. □

Remark. To obtain an *orthonormal basis* of V, simply divide each $\vec{v}_i$ by its length. This is called *normalizing*.

1.5 Orthogonal Complement

Definition 1.5.1 (Orthogonal Complement)**.** Let V be an *inner product space* and $W \subset V$ a subspace. The orthogonal complement of W is:

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

Example 1.5.1.1 (Orthogonal complement of entire space is empty)**.**

$$W = V \qquad W^{\perp} = \{\vec{v} \in V : \langle \vec{v}, \vec{w} \rangle = 0 \text{ for all } \vec{w} \in V\} = \left\{\vec{0}\right\}$$
$$\text{because } \langle \vec{v}, \vec{v} \rangle = 0 \implies \vec{v} = \vec{0}$$

Example 1.5.1.2 (Orthogonal complement of empty subspace is the entire space)**.**

$$W = \left\{\vec{0}\right\} \qquad W^{\perp} = V$$

Example 1.5.1.3 (Orthogonal complelment of a line is a plane)**.**

$$V = \mathbb{R}^3 \qquad W = \{(0, 0, z) : z \in \mathbb{R}\} \qquad W^{\perp} = \{(x, y, 0) : x, y \in \mathbb{R}\}$$

Week 2

September 15 - September 20

2.1 Orthogonal Complement (continued)

Theorem 2.1.1. *Let V be a finite-dimensional inner product space, and $W \subset V$ be a subspace, then:*

$$V \simeq W \oplus W^{\perp}$$

via the transformation $T : W \oplus W^{\perp} \to V$ given by:

$$T(\vec{w}, \vec{w}') = \vec{w} + \vec{w}'$$

Proof. We prove the theorem by writing an inverse for T. Let $\{\vec{w}_1, \ldots, \vec{w}_k\}$ be an orthonormal basis of W and define:

$$\Psi : V \to W \oplus W^{\perp} \tag{2.1}$$

$$\Psi(\vec{v}) = \left(\sum_{i=1}^{k} \langle \vec{v}, \vec{w}_i \rangle \, \vec{w}_i \ , \ \vec{v} - \sum_{i=1}^{k} \langle \vec{v}, \vec{w}_i \rangle \, \vec{w}_i \right) \tag{2.2}$$

Ψ is well defined since the first entry is in W by being a linear combination of $\vec{w}_i$, and the right entry is in $W^{\perp}$ because it is orthogonal to each $\vec{w}_i$ in our basis. It clear that $T \circ \Psi = \mathrm{id}_V$, so it remains to be shown that $\Psi \circ T = \mathrm{id}_{W \oplus W^{\perp}}$:

$$\begin{aligned}
\Psi\left(T\left(\vec{w}, \vec{w}'\right)\right) &= \Psi(\vec{w} + \vec{w}') && (2.3)\\
&= \left(\sum_{i=1}^{k} \langle \vec{w} + \vec{w}', \vec{w}_i \rangle \, w_i \ , \ \vec{w} + \vec{w}' - \sum_{i=1}^{k} \langle \vec{w} + \vec{w}', \vec{w}_i \rangle \, w_i \right) && (2.4)\\
&= \left(\sum_{i=1}^{k} \langle \vec{w}, \vec{w}_i \rangle \, w_i \ , \ \vec{w} + \vec{w}' - \sum_{i=1}^{k} \langle \vec{w}, \vec{w}_i \rangle \, w_i \right) && (2.5)\\
&= (\vec{w}, \vec{w}') && (2.6)
\end{aligned}$$

Thus T and Ψ are inverses. Since T and Ψ are linear transformations, T is an isomorphism. □

Corollary 2.1.1.1 (Extension of orthonormal basis). *Let* $\{\vec{w}_1, \ldots, \vec{w}_k\}$ *be an orthonormal basis of a subspace* W*. One can extend this to an orthornomal basis of the entire space:*

$$\{\vec{w}_1, \ldots, \vec{w}_k, \vec{v}_{k+1}, \ldots, \vec{v}_n\}$$

where $\{\vec{v}_{k+1}, \ldots, \vec{v}_n\}$ *is an orthonormal basis of* $W^\perp$*.*

Corollary 2.1.1.2 (Dimension of orthogonal complement).

$$\dim V = \dim W + \dim W^\perp$$

Corollary 2.1.1.3 (Duality of orthogonal complement).

$$\left(W^\perp\right)^\perp = W$$

Corollary 2.1.1.4 (Intersection of subspace and orthogonal complement).

$$W \cap W^\perp = (0)$$

Definition 2.1.2 (Projection onto a subspace). Let $W \subset V$ be a subspace and $\vec{v} \in V$. Then for $\Psi : V \to W \oplus W^\perp$ as defined in 2.1.1, we define the *projection of* $\vec{v}$ *onto* W to be the first coordinate $\Psi(\vec{v})$, denoted:

$$\mathrm{proj}_W(\vec{v})$$

Remark. If $\{\vec{w}_1, \ldots, \vec{w}_k\}$ is an orthonormal basis of W, then:

$$\mathrm{proj}_W(\vec{v}) = \sum_{i=1}^{k} \langle \vec{v}, \vec{w}_i \rangle \vec{w}_i$$

2.2 Adjoints

Definition 2.2.1 (Conjugate Transpose). For any matrix B, we define B^* to be the *conjugate transpose* given by taking the conjugate of each entry in $B^\intercal$, that is:

$$B^* = \overline{B^\intercal}$$

Lemma 2.2.2 (Unique inner product form of a linear transformation). *Let* $\mathcal{U} : V \to \mathbb{F}$ *be a linear transformation, then there exists some unique* $z \in V$ *such that:*

$$\mathcal{U}(\vec{v}) = \langle \vec{v}, \vec{z} \rangle \text{ for all } \vec{v} \in V$$

Proof. Let $\{\vec{v}_1, \ldots, \vec{v}_n\}$ be an orthonormal basis of V and define $\vec{z} \in V$ to be:

$$\vec{z} = \overline{\mathcal{U}(\vec{v}_1)}\vec{v}_1 + \ldots + \overline{\mathcal{U}(\vec{v}_n)}\vec{v}_n$$

Then we check that $\mathcal{U}(\vec{v}) = \langle \vec{v}, \vec{z} \rangle$ for all $\vec{v} \in V$:

$$\begin{aligned} \mathcal{U}(\vec{v}) &= \mathcal{U}\left(a_1\vec{v}_1 + \ldots + a_n\vec{v}_n\right) && (2.7)\\ &= a_1\mathcal{U}(\vec{v}_1) + \ldots + a_n\mathcal{U}(\vec{v}_n) && (2.8) \end{aligned}$$

$$\begin{aligned} \langle \vec{v}, \vec{z} \rangle &= \left\langle \vec{v}, \overline{\mathcal{U}(\vec{v}_1)}\vec{v}_1 + \ldots + \overline{\mathcal{U}(\vec{v}_n)}\vec{v}_n \right\rangle && (2.9)\\ &= \mathcal{U}(\vec{v}_1)\left\langle \vec{v}, \vec{v}_1 \right\rangle + \ldots + \mathcal{U}(\vec{v}_n)\left\langle \vec{v}, \vec{v}_n \right\rangle && (2.10)\\ &= a_1\mathcal{U}(\vec{v}_1) + \ldots + a_n\mathcal{U}(\vec{v}_n) && (2.11)\\ &= \mathcal{U}(\vec{v}) && (2.12) \end{aligned}$$

To show that $\vec{z}$ is unique, suppose that $\langle \vec{v}, \vec{z} \rangle = \langle \vec{v}, \vec{z}' \rangle$ for all $\vec{v} \in V$, then:

$$0 = \langle \vec{v}, \vec{z} \rangle - \langle \vec{v}, \vec{z}' \rangle = \langle \vec{v}, \vec{z} - \vec{z}' \rangle$$

Since this holds for all $\vec{v}$, we must have $\vec{z} - \vec{z}' = 0$ (indeed, $V^{\perp} = (0)$), we have $\vec{z}' = \vec{z}$ as required. □

Theorem 2.2.3 (Existence of unique adjoint)**.** *Let $T : V \to V$ be a linear transformation on an inner product space V. There exists a unique linear transformation $T^* : V \to V$ satisfying:*

$$\langle T\vec{x}, \vec{y} \rangle = \langle \vec{x}, T^*\vec{y} \rangle \text{ for all } \vec{x}, \vec{y} \in V$$

This T^* is called the adjoint *of T.*

Proof. For any $\vec{y} \in V$, define $g_{\vec{y}} : V \to \mathbb{F}$ (where $\mathbb{F}$ is $\mathbb{C}$ or $\mathbb{R}$), by:

$$g_{\vec{y}}(\vec{v}) = \langle T\vec{v}, \vec{y} \rangle$$

Then $g_{\vec{y}}$ is a linear transformation because:

$$\begin{aligned} g_{\vec{y}}\left(\vec{v} + \vec{w}\right) &= \left\langle T\left(\vec{v} + \vec{w}\right), \vec{y} \right\rangle && (2.13)\\ &= \langle T\vec{v} + T\vec{w}, \vec{y} \rangle && (2.14)\\ &= \langle T\vec{v}, \vec{y} \rangle + \langle T\vec{w}, \vec{y} \rangle && (2.15)\\ &= g_{\vec{y}}(\vec{v}) + g_{\vec{y}}(\vec{w}) && (2.16) \end{aligned}$$

$$\begin{aligned} g_{\vec{v}}\left(c\vec{v}\right) &= \left\langle T\left(c\vec{v}\right), \vec{y} \right\rangle && (2.17)\\ &= c\left\langle T\vec{v}, \vec{y} \right\rangle && (2.18)\\ &= cg_{\vec{y}}(\vec{v}) && (2.19) \end{aligned}$$

Then we can define $T^* : V \to V$ by the map from $\vec{y} \in V$ to the unique $\vec{z}$ generated by 2.2.2 for $g_{\vec{y}}$. Then, by definition of $\vec{z}$ we have:

$$\langle T\vec{x}, \vec{y} \rangle = g_{\vec{y}}(\vec{x}) = \langle \vec{x}, \vec{z} \rangle = \langle \vec{x}, T^*\vec{y} \rangle$$

By uniqueness of $\vec{z}$, this mapping T^* is unique. Thus it remains only to show that T^* is linear. For all $\vec{x}$, $\vec{y}$, $\vec{z} \in V$ and $c \in \mathbb{F}$:

$$\begin{aligned} \langle \vec{x}, T^*(c\vec{y}) \rangle &= \langle T\vec{x}, c\vec{y} \rangle && (2.20) \\ &= \bar{c} \langle T\vec{x}, \vec{y} \rangle && (2.21) \\ &= \bar{c} \langle \vec{x}, T^*\vec{y} \rangle && (2.22) \\ &= \langle \vec{x}, cT^*\vec{y} \rangle && (2.23) \end{aligned}$$

By the same orthogonal complement dimension argument as in 2.2.2, since this holds for all $\vec{x}$, $T^*(c\vec{y}) = cT^*\vec{y}$ as required. Similarly:

$$\begin{aligned} \langle \vec{x}, T^*(\vec{y} + \vec{z}) \rangle &= \langle T\vec{x}, \vec{y} + \vec{z} \rangle && (2.24) \\ &= \langle T\vec{x}, \vec{y} \rangle + \langle T\vec{x}, \vec{z} \rangle && (2.25) \\ &= \langle \vec{x}, T^*\vec{y} \rangle + \langle vecx, T^*\vec{z} \rangle && (2.26) \\ &= \langle \vec{x}, T^*\vec{y} + T^*\vec{z} \rangle && (2.27) \end{aligned}$$

Again, by the argument used in 2.2.2, since this holds for all $\vec{x}$, we have $T^*(\vec{y} + \vec{z}) = T^*\vec{y} + T^*\vec{z}$ as required. Thus T^* is unique and linear as required. □

Theorem 2.2.4 (Equivalence of conjugate transpose and adjoint)**.** *If B is an orthonormal basis of V, then:*

$$[T]_B^* = [T^*]_B$$

Proof. Let $B = \{\vec{v}_1, \ldots, \vec{v}_n\}$, $[T]_B = (a_{ij})$ and $[T^*]_B = (b_{ji})$. Then for any i, j:

$$b_{ij} = \langle T^*\vec{v}_j, \vec{v}_i \rangle = \langle \vec{v}_j, T\vec{v}_i \rangle = \overline{\langle T\vec{v}_i, \vec{v}_j \rangle} = \overline{a_{ji}}$$

□

2.3 Least Squares (example)

Say $\{(x_1, y_1), \ldots, (x_m, y_n)\}$ is a set of points in $\mathbb{R}^2$ and we want to find the line that best fits the data. More precisely, we want to find $a, b \in \mathbb{R}$ such that the line $y = ax + b$ minmizes the quantity:

$$E = \sum_{i=1}^{m} |y_i - (ax_i + b)|^2$$

Define the following:

$$A = \begin{bmatrix} x_1 & 1 \\ \vdots & \vdots \\ x_m & 1 \end{bmatrix} \qquad \vec{x} = \begin{bmatrix} a \\ b \end{bmatrix} \qquad \vec{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix}$$

Then we can rewrite the error E as:

$$E = \|A\vec{x} - \vec{y}\|^2$$

This is minimized when $A\vec{x} = \operatorname{proj}_{\operatorname{im} A}(\vec{y})$, so we just need to find $\vec{x}$ given $A\vec{x}$.

Remark (Author's Note)**.** In the following section we will take the adjoint of A even though $A : \mathbb{R}^2 \to \mathbb{R}^n$ and we only defined the adjoint for endofunctions. You can justify this to yourself by imagining the extensions of A and $\vec{x}$ given by:

$$A' = \begin{bmatrix} x_1 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_m & 1 & 0 & \dots & 0 \end{bmatrix} \qquad \vec{x} = \begin{bmatrix} a \\ b \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

And convincing yourself that we have not fundamentally changed the problem. In general though, there is something called the Hermitian adjoint which properly describes this property for $A : H_1 \to H_2$ where H_1 and H_2 are Hilbert spaces that have an appropriate inner product. I don't know anything about Hilbert spaces, this just happens to be a thing that exists.

We know that $\langle A\vec{x}, \vec{y}\rangle = \langle \vec{x}, A^*\vec{y}\rangle$. Thus, if $A^*A\vec{x} = \vec{0}$ we have:

$$A^*A\vec{x} = \vec{0} \implies \langle A^*A\vec{x}, \vec{x}\rangle = 0 \tag{2.28}$$

$$\implies \langle A\vec{x}, A\vec{x}\rangle = 0 \tag{2.29}$$

$$\implies A\vec{x} = \vec{0} \tag{2.30}$$

This tells us that if $\ker A = (\vec{0})$, then $\ker(A^*A) = (\vec{0})$ meaning A^*A is invertible. In any practical case $\ker A = (\vec{0})$ since, otherwise, that would mean all of our x_is are equal, so our line doesn't represent anything interesting. Thus, if $\vec{x}$ is our best fit line, then:

$$A\vec{x} - \vec{y} \in (\operatorname{im} A)^\perp \tag{2.31}$$

$$\implies \langle A\vec{z}, A\vec{x} - \vec{y}\rangle = 0 \text{ for all } \vec{z} \in \mathbb{R}^2 \tag{2.32}$$

$$\implies \langle \vec{z}, A^*\left(A\vec{x} - \vec{y}\right)\rangle = 0 \tag{2.33}$$

$$\implies A^*\left(A\vec{x} - \vec{y}\right) = 0 \tag{2.34}$$

$$\implies A^*A\vec{x} = A^*\vec{y} \tag{2.35}$$

$$\implies \vec{x} = \left(A^*A\right)^{-1} A^*\vec{y} \tag{2.36}$$

Week 3

September 22 - September 27

3.1 Normal Operators

Definition 3.1.1. Let $T : V \to V$ be a linear transformation on an inner product space V. We say T is *normal* if:

$$T^*T = TT^*$$

Remark. If there exists an orthonormal basis B such that $[T]_B$ is diagonal, then $[T^*]_B = [T]_B^*$ is also diagonal thus:

$$[T]_B^*[T]_B = [T]_B[T]^*B \tag{3.1}$$

$$T^*T = TT^* \tag{3.2}$$

So T is *normal*.

Definition 3.1.2. Let $T : V \to V$ be a linear transformation on a vector space V, and let W be a subspace of V. We say W is T-invariant if, for all $\vec{w} \in W$, $T\vec{w} \in W$.

Lemma 3.1.3 (Schur)**.** *Let $T : V \to V$ be a linear transformation on an inner product space V. If the characteristic polynomial of T splits completely, then there is an orthonormal basis B of V such that $[T]_B$ is upper triangular.*

Proof. We induce on $\dim V$. The case $\dim V = 1$ is trivial since all 1×1 matrices are upper triangular. So we assume the lemma holds for all inner product spaces W with $\dim W < \dim V$. Since the characteristic polynomial splits completely, there is some eigenvector $\vec{v} \in V$ and corresponding eigenvalue λ satisfying:

$$T\vec{v} = \lambda\vec{v}$$

Thus, for any $\vec{x} \in V$:

$$0 = \langle (T - \lambda I)\,\vec{v}, \vec{x} \rangle \tag{3.3}$$

$$= \left\langle \vec{v}, \left(T^* - \overline{\lambda} I\right) \vec{x} \right\rangle \tag{3.4}$$

Which means that $\vec{v} \in \left(\operatorname{im}\left(T^* - \overline{\lambda} I\right)\right)^\perp$. Thus $\left(T^* - \overline{\lambda} I\right)$ is not surjective, so by rank-nullity theorem, there is some nonzero $\vec{z} \in \ker\left(T^* - \overline{\lambda} I\right)$, giving:

$$\left(T^* - \overline{\lambda} I\right)\vec{z} = 0 \tag{3.5}$$
$$T^*\vec{z} = \overline{\lambda}\vec{z} \tag{3.6}$$

Without loss of generality, assume that $\|\vec{z}\| = 1$, since the equality holds under scalar multiplication of $\vec{z}$. Let $W = \operatorname{span}\{\vec{z}\}$, then W is T^*-invariant. Then, for all $\vec{y} \in W^\perp$:

$$\langle T\vec{y}, c\vec{z}\rangle = \overline{c}\langle \vec{y}, T^*\vec{z}\rangle \tag{3.7}$$
$$= \overline{c}\lambda\langle \vec{y}, \vec{z}\rangle \tag{3.8}$$
$$= 0 \text{ by choice of } \vec{y} \tag{3.9}$$

Thus $W^\perp$ is T-invariant. This means $T\big|_{W^\perp} : W^\perp \to W^\perp$ is a linear transformation (whose characteristic polynomial splits completely, proof omitted in this class but this follows from the fact that T splits), and $\dim W^\perp = \dim V - 1$. Thus, by our inductive hypothesis there exists an orthonormal basis $\beta = \{\vec{v}_1, \dots, \vec{v}_{n-1}\}$ of $W^\perp$ such that $\left[T\big|_{W^\perp}\right]_P$ is upper triangular. Thus:

$$[T^*]_B = \begin{bmatrix} \left[T^*\big|_{W^\perp}\right]_\beta & 0 \\ * & \overline{\lambda} \end{bmatrix} \tag{3.10}$$
$$[T]_B = \begin{bmatrix} \left[T\big|_{W^\perp}\right]_\beta & * \\ 0 & \lambda \end{bmatrix} \tag{3.11}$$

which is upper triangular. □

Theorem 3.1.4 (Orthonormal Diagonalizability of Complex Linear Transformations)**.** *If $T : V \to V$ is a* normal *linear transformation on a complex inner product space V, then there exists an orthonormal basis B such that $[T]_B$ is diagonal.*

Proof. Since all polynomials split over $\mathbb{C}$, by 3.1.3, there is an orthonormal basis:

$$B = \{\vec{v}_1, \dots, \vec{v}_n\}$$

such that $[T_B]$ is upper triangular. We will show that $[T]_B$ is also diagonal. Let $[T]_B = (a_{ij})$, we will show that $a_{ij} = 0$ if $i \neq j$ by induction on j. If $j = 1$, this is immediate from upper triangularity, so if the claim holds for all $j' < j$. If $i < j$ then:

$$0 = \|T\vec{v}_i - \lambda\vec{v}_i\|^2 \text{ for } \lambda = a_{ii} \tag{3.12}$$
$$= \langle T\vec{v}_i - \lambda\vec{v}_i, T\vec{v}_i - \lambda\vec{v}_i\rangle \tag{3.13}$$
$$= \langle (T - \lambda I)\vec{v}_i, (T - \lambda I)\vec{v}_i\rangle \tag{3.14}$$
$$= \langle \vec{v}_i, (T - \lambda I)^*(T - \lambda I)\vec{v}_i\rangle \tag{3.15}$$
$$= \langle \vec{v}_i, (T - \lambda I)(T - \lambda I)^*\vec{v}_i\rangle \tag{3.16}$$
$$= \left\langle (T^* - \overline{\lambda} I)\vec{v}_i, (T^* - \overline{\lambda} I)\vec{v}_i\right\rangle \tag{3.17}$$
$$= \|T^*\vec{v}_i - \lambda\vec{v}_i\| \tag{3.18}$$

Thus $T^*\vec{v}_i = \overline{\lambda}\vec{v}_i$. Then consider:

$$T\vec{v}_j = a_{1j}\vec{v}_1 + \cdots + a_{jj}\vec{v}_j$$

By orthonormality of our basis, it follows that:

$$\begin{aligned} a_{ij} &= \langle T\vec{v}_j, \vec{v}_i\rangle & (3.19)\\ &= \langle \vec{v}_j, T^*\vec{v}_i\rangle & (3.20)\\ &= \left\langle \vec{v}_j, \overline{\lambda}\vec{v}_i\right\rangle & (3.21)\\ &= 0 & (3.22) \end{aligned}$$

As required, each entry a_{ij} with $i < j$ is 0, and entries with $i > j$ follow from upper triangularity. □

Corollary 3.1.4.1. *If $T : V \to V$ is a linear transformation on a complex inner product space V, then there exists an othonormal basis B such that $[T]_B$ is diagonal if and only if T is* normal.

Remark. 3.1.4.1 does not apply to real inner product spaces. Consider $V : \mathbb{R}^2$ and $T : \mathbb{R}^2 \to R^2$ given by the rotation matrix:

$$\begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$$

Then T^* describes the opposite rotation, thus $T^*T = TT^* = I$ so T is *normal*, however if $\theta \notin \pi\mathbb{Z}$, T has no real eigenvectors and is thus not diagonalizable.

3.2 Self-Adjoint Operators

Definition 3.2.1. A linear transformation T is *self-adjoint* (or Hermitian) if $T = T^*$.

Remark. If $T : V \to V$ is a linear transformation on a real inner product space V, and there exists an orthonormal basis B for which $[T]_B$ is diagonal, then:

$$[T]_B = [T]_B^*$$

So T is *self-adjoint*

Remark. If $T = T^*$, then $T^*T = TT^*$ so T is normal.

Theorem 3.2.2 (Orthonormal Diagonalizability of Real Linear Transformations)**.** *If $T : V \to V$ is a linear transformation on a real inner product space V, then T is* self-adjoint *if and only if there is an orthonormal basis B such that $[T]_B$ is diagonal.*

Proof. Note that the characteristic polynomial of T must split over $\mathbb{C}$, so consider any eigenvector $\vec{x}$ and eigenvalue $\lambda \in \mathbb{C}$ such that $T\vec{x} = \lambda\vec{x}$, then:

$$(T - \lambda I)\vec{x} = \vec{0} \implies (T^* - \overline{\lambda} I)\vec{x} = 0 (\text{see proof of 3.1.4}) \tag{3.23}$$

$$\implies T^*\vec{x} = \overline{\lambda} x \tag{3.24}$$

So if T is *self-adjoint*:

$$\overline{\lambda}\vec{x} = T^*\vec{x} = T\vec{x} = \lambda\vec{x} \tag{3.25}$$

$$\overline{\lambda} = \lambda \tag{3.26}$$

Thus $\lambda \in \mathbb{R}$, so all eigenvalues of T are real. Thus the characteristic polynomial of T splits completely over $\mathbb{R}$, so invoking 3.1.3, there must exist an orthonormal basis B such that $[T]_B$ is upper triangular. However $[T]_B^* = [T^*]_B$ which must be lower triangular, so $[T]_B$ is both upper and lower triangular, meaning $[T]_B$ is diagonal. □

Corollary 3.2.2.1 (Orthonormal Diagonalizability of Symmetric Real Matrices)**.** *A real matrix is orthogonally diaganalizable if and only if it's symmetric.*

Proof. A real matrix that is *self-adjoint* is just a symmetric matrix, so this follows immediately from 3.2.2. □

3.3 Isometries

Definition 3.3.1. A linear transformation $T : V \to W$ from an inner product space V to an inner product space W is an *isometry* if $\langle T\vec{x}, T\vec{y}\rangle = \langle \vec{x}, \vec{y}\rangle$ for all $\vec{x}$, $\vec{y} \in V$.

Definition 3.3.2. An *isometry* T is *unitary* if T is surjective.

Definition 3.3.3. Let V, W be inner product spaces. If there exists a *unitary isometry* $T : V \to W$, we say V and W are *isometric.*

Remark. Every *isometry* T is injective because:

$$T\vec{x} = \vec{0} \implies \langle \vec{x}, \vec{x}\rangle = \langle T\vec{x}, T\vec{x}\rangle = 0 \tag{3.27}$$

$$\implies \vec{x} = 0 \tag{3.28}$$

Thus $\ker T = (0)$.

Remark (Author's Note)**.** Again, in this section, we will use the adjoint of T even if T is not an endomorphism. In finite dimensional vector spaces, this exists, and the conjugate transpose of the matrix representation still works, you'll just have to convince yourself.

Remark. For every *isometry* $T : V \to W$, $T^*T = I$ since, for all $\vec{x}, \vec{y} \in V$:

$$\langle T\vec{x}, T\vec{y}\rangle = \langle \vec{x}, \vec{y}\rangle \tag{3.29}$$

$$\langle \vec{x}, T^*T\vec{y}\rangle = \langle \vec{x}, \vec{y}\rangle \text{ for any } \vec{x} \tag{3.30}$$

Remark. If T is *unitary*, T is invertible so $TT^* = I = T^*T$, so T is also normal.

Lemma 3.3.4. *Let $\mathcal{U} : V \to V$ be a self-adjoint linear transformation, and $\langle \vec{x}, \mathcal{U}\vec{x} \rangle = 0$ for all $\vec{x} \in V$, then $\mathcal{U} = 0$.*

Proof. Suppose $\vec{x}$ is an eigenvector of $\mathcal{U}$ and λ be its corresponding eigenvalue, then:

$$0 = \langle \vec{x}, \mathcal{U}\vec{x} \rangle \tag{3.31}$$
$$= \langle \vec{x}, \lambda\vec{x} \rangle \tag{3.32}$$
$$= \overline{\lambda} \langle \vec{x}, \vec{x} \rangle \tag{3.33}$$

But $\vec{x} \neq \vec{0}$ by choice of $\vec{x}$ being an eigenvector, so $\lambda = 0$. Since all eigenvalues of $\mathcal{U}$ are 0 and $\mathcal{U}$ is diagonalizable (since it is self-adjoint), $\mathcal{U} = 0$. □

Theorem 3.3.5. *Let $T : V \to V$ be a surjective linear transformation on a finite dimensional inner product space V, then the following are equivalent:*

i. $TT^* = T^*T = I$

ii. $\langle T\vec{x}, T\vec{y} \rangle = \langle \vec{x}, \vec{y} \rangle$ *for all* $\vec{x}, \vec{y} \in V$

iii. *If B is an orthonormal basis, then so is $T(B)$*

iv. *There exists an orthonormal basis B such that $T(B)$ is also orthonormal*

v. $\|T\vec{x}\| = \|\vec{x}\|$ *for all* $\vec{x} \in V$.

Proof. We will prove a ring of implications:

i. $\implies$ ii. For all $\vec{x}, \vec{y} \in V$:

$$\langle T\vec{x}, T\vec{y} \rangle = \langle \vec{x}, T^*T\vec{y} \rangle \tag{3.34}$$
$$= \langle \vec{x}, I\vec{y} \rangle \tag{3.35}$$
$$= \langle \vec{x}, \vec{y} \rangle \tag{3.36}$$

ii. $\implies$ iii. Let B be a basis $\{\vec{v}_1, \dots, \vec{v}_n\}$, then for any $\vec{v}_i$, $\vec{v}_j \in B$:

$$\langle T\vec{v}_i, T\vec{v}_j \rangle = \langle \vec{v}_i, \vec{v}_j \rangle = \delta_{ij} = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}$$

Thus $T(B)$ is orthonormal. Recall from 1.3.6 that this is sufficient to show $T(B)$ is linearly independent and thus a basis.

iii. $\implies$ iv. Immediate from the fact that V is finite dimensional so an orthonormal basis exists.

iv. $\implies$ v. For any $\vec{x} \in V$ write:

$$\vec{x} = a_1\vec{v}_1 + \cdots + a_n\vec{v}_n$$

where $\{\vec{v}_1, \ldots, \vec{v}_n\}$ are a subset of the orthonormal basis B provided by the assumption. Then:

$$\begin{aligned}
\|T\vec{x}\|^2 &= \|T(a_1\vec{v}_1 + \cdots + a_n\vec{v}_n)\|^2 && (3.37)\\
&= \|a_1T\vec{v}_1 + \cdots + a_nT\vec{v}_n\|^2 && (3.38)\\
&= |a_1|^2 + \cdots + |a_n|^2 && (3.39)\\
&= \|\vec{x}\|^2 && (3.40)
\end{aligned}$$

By non-negativity of the norm, v. holds.

v. $\implies$ i. From our assumption, for all $\vec{x}$:

$$\begin{aligned}
\|\vec{x}\| &= \|T\vec{x}\| && (3.41)\\
&= \langle T\vec{x}, T\vec{x}\rangle && (3.42)\\
&= \langle \vec{x}, T^*T\vec{x}\rangle && (3.43)
\end{aligned}$$

We have $\langle \vec{x}, (T^*T - I)\vec{x}\rangle = 0$ for all $\vec{x}$. Note that $(T^*T - I)$ is self-adjoint because $(T^*T - I)^* = T^*T - I$. Thus by 3.3.4, $T^*T - I = 0$ so $T^*T = I$. Thus T^* is a left inverse of T, so since T is an endomorphism, $T^*T = I = TT^*$.

□

Corollary 3.3.5.1. *Let V, W be* isometric *finite dimensional inner product spaces, then* $\dim V = \dim W$.

Corollary 3.3.5.2. *If* $\dim V = \dim W$ *for finite dimensional inner product spaces V, W, then V, W are* isometric.

Proof. Since V and W are finite dimensional, they have orthonormal bases $\{\vec{v}_1, \ldots, \vec{v}_n\}$, $\{\vec{w}_1, \ldots, \vec{w}_n\}$. Then we can define a linear transformation $T : V \to W$ given by $T(\vec{v}_i) = \vec{w}_i$ for all i. By 3.3.5, T is an *isometry.* □

Corollary 3.3.5.3. *Any n-dimensional inner product space is* isometric *to $\mathbb{R}^n$ or $\mathbb{C}^n$ with the standard inner product.*

Corollary 3.3.5.4. *If $T : V \to W$ is* unitary*, then its eigenvalues all have absolute value 1.*

Proof. For all $\vec{x} \in T$:

$$\|\vec{x}\| = \|T\vec{x}\| = \|\lambda\vec{x}\| = |\lambda|\,\|\vec{x}\|$$

Thus for any eigenvalue λ, $|\lambda| = 1$. □

Week 4

September 29 - October 4

4.1 Orthogonal Matrices

Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be an orthogonal linear transformation:

Definition 4.1.1. We say B is an *eigenbasis* for T if B is an orthonormal basis of eigenvectors of T.

Remark. If $n = 1$, T is one of the following:

$$[1] \qquad\qquad [-1]$$

Remark. If $n = 2$, and A is the matrix for T, A must be a real matrix satisfying:

$$AA^{\intercal} = A^{\intercal}A = I$$

and since $\{(1,0),(0,1)\}$ is an orthonormal basis we must have:

$$\left\| A \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\| = \left\| A \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\| = 1$$

thus A must be of the following form for some $\theta \in [0, 2\pi)$:

$$A = \begin{pmatrix} \cos\theta & \pm\sin\theta \\ \sin\theta & \mp\cos\theta \end{pmatrix}$$

Remark. If $n = 2$, by lifting T to being a unitary transformation $\mathbb{C}^n \to \mathbb{C}^n$, we can distinguish between rotations and reflections from the eigenvectors and eigenvalues of A. We know the eigenvalues must be complex numbers of length 1, so if they are real, they are ± 1. So let A be the matrix of T under an eigenbasis, it must be of the form:

$$\underset{\text{rotation}}{\begin{bmatrix} \pm 1 & 0 \\ 0 & \pm 1 \end{bmatrix}} \quad \text{or} \quad \underset{\text{reflection}}{\begin{bmatrix} \pm 1 & 0 \\ 0 & \mp 1 \end{bmatrix}}$$

Otherwise, if the eigenvalues are not real, they are of the form $\cos\theta + i\sin\theta$ for some $\theta \in [0, 2\pi)$ Matrices of the form:

$$A = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

have eigenvalues $\cos\theta \pm i\sin\theta$ and are rotations.

Remark. An orthogonal 2x2 matrix can be the composition of a rotation and a reflection.

Theorem 4.1.2. *Let A be a real, orthogonal, $n \times n$ matrix. Then A is block diagonal with blocks of size* 0 *or* 1*.*

Proof. Lift A to a $n \times n$ complex, unitary matrix. Then, since the entries are real

$$A\vec{x} = \lambda\vec{x} \text{ for } \vec{x} \neq 0 \implies A\vec{x} = \overline{\lambda\vec{x}}$$

Thus non-real eigenvalues come in conjugate pairs. Since A is unitary as a complex matrix, we can find an *eigenbasis* B of $\mathbb{C}^n$ for A. Then consider an arbitrary pair of eigenvalues $\vec{v}$ and $\vec{w} = \overline{\vec{v}}$. We want to find two real vectors $\vec{x}, \vec{y} \in \mathbb{R}^n$ such that $\operatorname{span}\{\vec{x}, \vec{y}\} = \operatorname{span}\{\vec{v}, \vec{w}\}$ over $\mathbb{C}$. So define

$$\vec{x} = \vec{v} + \vec{w} \qquad (= 2\Re(\vec{v})) \tag{4.1}$$

$$\vec{y} = i\vec{v} + i\vec{w} \qquad (= -2\Im(\vec{v})) \tag{4.2}$$

Clearly, by definition, $\vec{x}, \vec{y} \in \operatorname{span}\{\vec{v}, \vec{w}\}$, and furthermore we have:

$$\vec{v} = \frac{1}{2i}(i\vec{x} + \vec{y}) \tag{4.3}$$

$$\vec{w} = \frac{1}{2i}(i\vec{x} - \vec{y}) \tag{4.4}$$

$$\tag{4.5}$$

Thus $\vec{w}, \vec{w} \in \operatorname{span}\{\vec{v}, \vec{w}\}$. Applying Gram-Schmidt allows us to turn $\{\vec{x}, \vec{y}\}$ into a real orthonormal basis of $\operatorname{span}\{\vec{v}, \vec{w}\}$. Doing this for every conjugate pair of non-real $\vec{v}_i \in B$ gives us a new, real orthonormal basis B' such that:

$$[A]_{b'} = \begin{pmatrix} (2\times 2) & 0 & \dots & 0 \\ 0 & (2\times 2) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & (1\times 1) \end{pmatrix}$$

where each block is also orthogonal matrix. $\square$

Remark. This means that any orthogonal transformation T, when viewed under the right basis, is a collection of pairwise orthogonal rotations (2×2 blocks) together with some fixed and reflected lines (± 1 eigenvalues).

Example 4.1.2.1. In $\mathbb{R}^3$, an orthogonal matrix A may look like:

$$A = \begin{bmatrix} \cos\frac{\pi}{3} & -\sin\frac{\pi}{3} & 0 \\ \sin\frac{\pi}{3} & \cos\frac{\pi}{3} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

4.2 Rigid Motions

Definition 4.2.1. A *rigid motion* is a function $f : \mathbb{R}^n \to \mathbb{R}^n$ such that

$$\|\vec{x} - \vec{y}\| = \|f(\vec{x}) - f(\vec{y})\| \text{ for all } \vec{x}, \vec{y} \in \mathbb{R}^n$$

That is, f preserves distances.

Example 4.2.1.1. A translation $\{\vec{x} \mapsto \vec{x} + \vec{a}\}$ is a *rigid motion*

Example 4.2.1.2. An orthogonal linear transformation is a *rigid motion*

Theorem 4.2.2. *Any* rigid motion $f : \mathbb{R}^n \to \mathbb{R}^n$ *can be written uniquely as*

$$f = g \circ T$$

where g is a translation and T is an orthogonal linear transformation

Proof. Define $T : \mathbb{R}^n \to \mathbb{R}^n$ by

$$T(\vec{x}) = f(\vec{x}) - f(\vec{0})$$

T is clearly a *rigid motion*, and $T(\vec{0}) = f(\vec{0}) - f(\vec{0}) = \vec{0}$. Also $f = g \circ T$ where g is the translation $g(\vec{x}) = \vec{x} + f(\vec{0})$. We will prove that T is linear. First observe that, for any $\vec{x} \in \mathbb{R}^n$

$$\|T\vec{x}\| = \left\|T\vec{x} - T\vec{0}\right\| = \left\|\vec{x} - \vec{0}\right\| = \|\vec{x}\|$$

Next we will show that, for any $\vec{x}, \vec{y} \in \mathbb{R}^n$

$$\langle T\vec{x}, T\vec{y}\rangle = \langle \vec{x}, \vec{y}\rangle$$

this is true because

$$\begin{aligned}
\|T\vec{x} - T\vec{y}\|^2 &= \langle T\vec{x} - T\vec{y}, T\vec{x} - T\vec{y}\rangle && (4.6)\\
&= \langle T\vec{x}, T\vec{x}\rangle - \langle T\vec{x}, T\vec{y}\rangle - \langle T\vec{y}, T\vec{x}\rangle + \langle T\vec{y}, T\vec{y}\rangle && (4.7)\\
&= \|T\vec{x}\|^2 + \|T\vec{y}\|^2 - 2\langle T\vec{x}, T\vec{y}\rangle \qquad = \|\vec{x}\|^2 + \|\vec{y}\|^2 - 2\langle T\vec{x}, T\vec{y}\rangle && (4.8)
\end{aligned}$$

but also

$$\begin{aligned}
\|T\vec{x} - T\vec{y}\|^2 &= \|\vec{x} - \vec{y}\|^2 && (4.9)\\
&= \|\vec{x}\|^2 + \|\vec{y}\|^2 - 2\langle \vec{x}, T\vec{y}\rangle && (4.10)
\end{aligned}$$

subtracting these two equations from each other we obtain

$$\langle T\vec{x}, T\vec{y}\rangle = \langle \vec{x}, \vec{y}\rangle$$

Using this fact we can show the two properties of linearity:

i. For any $a \in \mathbb{R}$, $\vec{x} \in \mathbb{R}^n$:

$$\begin{aligned}
\|T(a\vec{x}) - aT(\vec{x})\|^2 &= \|T(a\vec{x})\|^2 + \|aT(\vec{x})\|^2 - 2\langle T(a\vec{x}), aT(\vec{x})\rangle && (4.11)\\
&= \|a\vec{x}\|^2 + a^2\|\vec{x}\|^2 - 2a\langle a\vec{x}, \vec{x}\rangle && (4.12)\\
&= 2a^2\|\vec{x}\|^2 - 2a^2\|\vec{x}\|^2 = 0 && (4.13)
\end{aligned}$$

Thus, by positive definiteness of the norm $T(a\vec{x}) = aT(\vec{x})$.

ii. For any $\vec{x}, \vec{y} \in \mathbb{R}^n$:

$$\begin{aligned}
\|T(\vec{x}+\vec{y}) - T(\vec{x}) - T(\vec{y})\|^2 &= \|T(\vec{x}+\vec{y})\|^2 + \|T(\vec{x})\|^2 + \|T(\vec{y})\|^2 && (4.14)\\
&\quad - 2\langle T(\vec{x}+\vec{y}), T(\vec{x})\rangle - 2\langle T(\vec{x}+\vec{y}), T(\vec{y})\rangle\\
&\quad + 2\langle T(\vec{x}), T(\vec{y})\rangle\\
&= \|\vec{x}+\vec{y}\|^2 + \|\vec{x}\|^2 + \|\vec{y}\|^2 && (4.15)\\
&\quad - 2\langle \vec{x}+\vec{y}, \vec{x}\rangle - 2\langle \vec{x}+\vec{y}, \vec{y}\rangle + 2\langle \vec{x}, \vec{y}\rangle\\
&= \|\vec{x}\|^2 + \|\vec{y}\|^2 + 2\langle \vec{x}, \vec{y}\rangle - \|\vec{x}+\vec{y}\|^2 = 0 && (4.16)
\end{aligned}$$

Thus, by positive definiteness of the norm $T(\vec{x}+\vec{y}) = T(\vec{x}) + T(\vec{y})$.

Thus T is a linear *rigid motion*, so T is orthogonal. It remains only to be shown that $f = g \circ T$ is unique. Suppose $f = g' \circ T'$ for a translation g' and orthogonal transformation T', then:

$$\begin{aligned}
f(\vec{0}) &= (g' \circ T')\vec{0} = g'(\vec{0}) && (4.17)\\
&= (g \circ T)\vec{0} = g(\vec{0}) && (4.18)
\end{aligned}$$

Thus $g(\vec{x}) = g'(\vec{x}) = \vec{x} + f(\vec{0})$. But then

$$T'\vec{x} = \left(g^{-1} \circ f\right)\vec{x} = f(\vec{x}) - f(\vec{0}) = T\vec{x}$$

so $g' = g$ and $T' = T$ as required for uniqueness. $\square$