Inner Product Spaces, Orthogonal Projections, and Orthonormal Bases

Def: An *inner product* on a vector space V is a map $\langle\ ,\ \rangle : V \times V \to \mathbb{R}$ such that

(a) $\langle f,g\rangle = \langle g,f\rangle$ (symmetry)

(b) $\langle f+h,g\rangle = \langle f,g\rangle + \langle h,g\rangle$

(c) $\langle cf,g\rangle = c\langle f,g\rangle$

$\}$ LT in 1st coordinate.

(d) $\langle f,f\rangle > 0$ for all $f \neq 0$ in V (positive definite)

Note: By (c), we can see that $\langle f,f\rangle = 0$ if and only if $f = 0$.

Note: Applying the symmetry condition (a) to conditions (b) and (c), we see that

$$\langle f,g+h\rangle = \langle f,g\rangle + \langle f,h\rangle \text{ and } \langle f,cg\rangle = c\langle f,g\rangle.$$

So inner products are linear transformations in both coordinates.

Example: The dot product on $\mathbb{R}^n$ is the most common example of an inner product. Show that the dot product is an inner product.

(a) $\vec{v}\cdot\vec{w} = v_1w_1 + v_2w_2 + \cdots + v_nw_n = w_1v_1 + w_2v_2 + \cdots + w_nv_n = \vec{w}\cdot\vec{v}$

(b) $(\vec{v}+\vec{u})\cdot\vec{w} = (v_1+u_1)w_1 + \cdots + (v_n+u_n)w_n = [v_1w_1 + \cdots + v_nw_n] + [u_1w_1 + \cdots + u_nw_n]$
$= \vec{v}\cdot\vec{w} + \vec{u}\cdot\vec{w}$

(c) $(c\vec{v})\cdot\vec{w} = (cv_1)w_1 + \cdots + (cv_n)w_n = c(v_1w_1 + \cdots + v_nw_n) = c(\vec{v}\cdot\vec{w})$

(d) $\vec{v}\cdot\vec{v} = v_1^2 + \cdots + v_n^2 = 0 \Leftrightarrow v_1 = \cdots = v_n = 0$. So $\vec{v}\cdot\vec{v} > 0$ for all $\vec{v} \neq 0$.

Example: Let $V = \mathbb{R}^{n\times m}$, the space of all $n \times m$ matrices with entries in $\mathbb{R}$. Define a pairing on V by $\langle A,B\rangle = \mathrm{tr}(A^TB)$ for all matrices A, B in V. Show that $\langle\ ,\ \rangle$ is an inner product.

(a) $\mathrm{tr}(A^TB) = \mathrm{tr}((A^TB)^T) = \mathrm{tr}(B^TA) = \langle B,A\rangle$

(b) $\langle A+C,B\rangle = \mathrm{tr}((A+C)^TB) = \mathrm{tr}(A^TB + C^TB) = \mathrm{tr}(A^TB) + \mathrm{tr}(C^TB)$
$= \langle A,B\rangle + \langle C,B\rangle$

(c) $\langle cA,B\rangle = \mathrm{tr}((cA)^TB) = \mathrm{tr}(cA^TB) = c\,\mathrm{tr}(A^TB) = c\langle A,B\rangle$

(d) If $A \neq 0$, say $A = \begin{bmatrix} | & & | \\ \vec{v}_1 & \cdots & \vec{v}_m \\ | & & | \end{bmatrix}$. Then $\langle A,A\rangle = \mathrm{tr}\left(\begin{bmatrix} - & \vec{v}_1 & - \\ & \vdots & \\ - & \vec{v}_m & - \end{bmatrix}\begin{bmatrix} | & & | \\ \vec{v}_1 & \cdots & \vec{v}_m \\ | & & | \end{bmatrix}\right)$
$= \mathrm{tr}\begin{pmatrix} \vec{v}_1\vec{v}_1 & \vec{v}_1\vec{v}_2 & \cdots & \vec{v}_1\vec{v}_m \\ \vdots & & & \vdots \\ \vec{v}_m\vec{v}_1 & \cdots & \cdots & \vec{v}_m\vec{v}_m \end{pmatrix} = \|\vec{v}_1\|^2 + \cdots + \|\vec{v}_m\|^2 \neq 0.$

Class 2 ←

Def: The *norm* or *magnitude* of an element f in an inner product space V is: $\|f\| = \sqrt{\langle f,f\rangle}$

Def: We say that two elements f, g in an inner product space V are *orthogonal* or *perpendicular* if: $\langle f,g\rangle = 0$

Def: The distance between two elements f, g in an inner product space V is: $\|f-g\| \left(= \sqrt{\langle f-g,f-g\rangle}\right)$

Example: Let $V = C[0,1]$, and consider the inner product $\langle f,g\rangle = \int_0^1 f(t)g(t)\,dt$. Let $m(t) = 3t^3$ and $n(t) = -t$. Find $\|m\|$ and find the distance between m and n.

(a) $\|m\| = \sqrt{\langle m,m\rangle} = \sqrt{\int_0^1 9t^6\,dt} = \sqrt{\frac{9}{7}t^7\Big]_0^1} = \sqrt{\frac{9}{7}}$

(b) $\|m-n\| = \|3t^3+t\| = \sqrt{\int_0^1 9t^6 + 6t^4 + t^2\,dt} = \sqrt{\frac{9}{7}t^7 + \frac{6}{5}t^5 + \frac{1}{3}t^3\Big]_0^1}$

$= \sqrt{\frac{9}{7} + \frac{6}{5} + \frac{1}{3}} = \sqrt{\frac{296}{105}}$

Note: The standard inner product on $\mathbb{R}^n$ is the dot product. So in $\mathbb{R}^n$, we have the following facts.

- Two vectors $\vec{v}, \vec{u} \in \mathbb{R}^n$ are orthogonal/perpendicular if $\vec{v} \cdot \vec{u} = 0$.
- The length of a vectors $\vec{v} \in \mathbb{R}^n$ is $\|\vec{v}\| = \sqrt{\vec{v} \cdot \vec{v}}$.
- A vector $\vec{u} \in \mathbb{R}^n$ is a unit vector if the length of $\vec{u}$ is 1, i.e. if $\|\vec{u}\| = 1$.
- If $\vec{v} \in \mathbb{R}^n$, then the vector $\vec{u} = \frac{\vec{v}}{\|\vec{v}\|}$ is a unit vector in the same direction as $\vec{v}$.
- A vector $\vec{w} \in \mathbb{R}^n$ is *orthogonal to a subspace* V in $\mathbb{R}^n$ if $\vec{x}$ is orthogonal to every vector in $\vec{v}$, i.e. if $\vec{x} \cdot \vec{v} = 0$ for all $\vec{v} \in V$.

Def: We say that a collection $g_1, \ldots, g_m$ of elements in an inner product space V are *orthonormal* if they are unit vectors and if each one is orthogonal to the rest. In other words, the collection is *orthonormal* if

$$\langle g_i, g_j \rangle = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

Ex: The standard basis vectors $\vec{e}_1, \ldots, \vec{e}_n \in \mathbb{R}^n$ are orthonormal. (So they form an *orthonormal basis* of $\mathbb{R}^n$). Also, any subset of this collection is orthonormal.

class 1

Ex: Let θ be any angle. Show that the vectors $\begin{bmatrix} \sin\theta \\ \cos\theta \end{bmatrix}$ and $\begin{bmatrix} -\cos\theta \\ \sin\theta \end{bmatrix}$ are orthonormal in $\mathbb{R}^2$.

unit vectors: $\vec{v} \cdot \vec{v} = \sin^2\theta + \cos^2\theta = 1$

$\vec{w} \cdot \vec{w} = \cos^2\theta + \sin^2\theta = 1$

orthogonal: $\vec{v} \cdot \vec{w} = -\sin\theta\cos\theta + \cos\theta\sin\theta = 0$.

Ex: Show that the vectors $\begin{bmatrix} 1/2 \\ 1/2 \\ 1/2 \\ 1/2 \end{bmatrix}$ and $\begin{bmatrix} 1/2 \\ -1/2 \\ 1/2 \\ -1/2 \end{bmatrix}$ are orthonormal in $\mathbb{R}^4$.

unit: $\vec{v} \cdot \vec{v} = \frac{1}{4} + \frac{1}{4} + \frac{1}{4} + \frac{1}{4} = 1$

$\vec{w} \cdot \vec{w} = \frac{1}{4} + \frac{1}{4} + \frac{1}{4} + \frac{1}{4} = 1$

orthogonal: $\vec{v} \cdot \vec{w} = \frac{1}{4} - \frac{1}{4} + \frac{1}{4} - \frac{1}{4} = 0$.

Thm: Orthogonal vectors are linearly independent. So any set of n orthogonal vectors in a vector space V of dimension n is a basis for V.

Recall: (Ch. 2) Let L be any line in $\mathbb{R}^2$. Any vector $\vec{x} \in \mathbb{R}^2$ can be written uniquely as $\vec{x} = \vec{x}_p + \vec{x}_o$ where $\vec{x}_p$ is parallel to L and $\vec{x}_o$ is orthogonal to L. Here $\vec{x}_p$ is also denoted $\text{proj}_L(\vec{x})$.

Thm: Suppose $g_1, \ldots, g_m$ is an orthonormal basis of a subspace W of an inner product space V. Then the projection of an element f in V onto the subspace W is:

$$\text{proj}_W f = \langle g_1, f \rangle g_1 + \cdots + \langle g_m, f \rangle g_m .$$

Fact: Let W be a subspace of $\mathbb{R}^n$. Let $\{\vec{u}_1, \ldots, \vec{u}_m\}$ be an orthonormal basis of V. Then for any $\vec{x} \in \mathbb{R}^n$,

$$\operatorname{proj}_W(\vec{x}) = \vec{x}_p = (\vec{u}_1 \cdot \vec{x})\vec{u}_1 + \cdots + (\vec{u}_m \cdot \vec{x})\vec{u}_m = \sum_{i=1}^{m} (\vec{u}_i \cdot \vec{x})\vec{u}_i$$

Obs: Let V be the (xy)-plane in $\mathbb{R}^3$. Then $\{\vec{e}_1, \vec{e}_2\}$ is an orthonormal basis for V. Let $L_1 = \operatorname{span}\{\vec{e}_1\}$ be the x-axis an let $L_2 = \operatorname{span}\{\vec{e}_2\}$ be the y-axis. Then for any $\vec{x} \in \mathbb{R}^3$, $\operatorname{proj}_V(\vec{x}) = \operatorname{proj}_{L_1}(\vec{x}) + \operatorname{proj}_{L_2}(\vec{x})$.

This is true in general: Let V be a subspace of $\mathbb{R}^n$. Let $\{\vec{u}_1, \ldots, \vec{u}_m\}$ be an orthonormal basis of V. Let $L_i = \operatorname{span}(\vec{u}_i)$ for each i. Then

$$\operatorname{proj}_V(\vec{x}) = \sum_{i=1}^{m} (\vec{u}_i \cdot \vec{x})\vec{u}_i = \sum_{i=1}^{m} \operatorname{proj}_{L_i}(\vec{x}).$$

Ex: Let $A = \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 1 \\ 1 & -1 \end{bmatrix}$ and let $V = \operatorname{Im}(A)$. Find $\operatorname{proj}_V(\vec{x})$ for $\vec{x} = \begin{bmatrix} 1 \\ 3 \\ 2 \\ 4 \end{bmatrix}$.

① Find an orthonormal basis for $V = \operatorname{Im}(A)$: $\vec{u}_1 = \frac{\vec{v}_1}{\|\vec{v}_1\|} = \frac{\vec{v}_1}{2} = \begin{bmatrix} 1/2 \\ 1/2 \\ 1/2 \\ 1/2 \end{bmatrix}$

$\vec{u}_2 = \frac{\vec{v}_2}{\|\vec{v}_2\|} = \frac{\vec{v}_2}{2} = \begin{bmatrix} 1/2 \\ -1/2 \\ 1/2 \\ -1/2 \end{bmatrix}$

② $\operatorname{proj}_V(\vec{x}) = (\vec{u}_1 \cdot \vec{x})\vec{u}_1 + (\vec{u}_2 \cdot \vec{x})\vec{u}_2 = \left(\frac{1}{2} + \frac{3}{2} + \frac{2}{2} + \frac{4}{2}\right)\vec{u}_1 + \left(\frac{1}{2} - \frac{3}{2} + \frac{2}{2} - \frac{4}{2}\right)\vec{u}_2$

$= 5\vec{u}_1 - 2\vec{u}_2 = \begin{bmatrix} 5/2 \\ 5/2 \\ 5/2 \\ 5/2 \end{bmatrix} - \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 3/2 \\ 7/2 \\ 3/2 \\ 7/2 \end{bmatrix}.$

Thm: Let $\vec{u}_1, \ldots, \vec{u}_n$ be an orthonormal basis for $\mathbb{R}^n$. Then for any $\vec{x} \in \mathbb{R}^n$, $\vec{x} = (\vec{u}_1 \cdot \vec{x})\vec{u}_1 + \cdots + (\vec{u}_n \cdot \vec{x})\vec{u}_n$. $(= \operatorname{proj}_{\mathbb{R}^n}(\vec{x}))$

Note: This means that if $\vec{u}_1, \ldots, \vec{u}_n$ is an orthonormal basis for $\mathbb{R}^n$, then $\vec{x} = \sum_{i=1}^{n} \operatorname{proj}_{\operatorname{span}\{\vec{u}_i\}}(\vec{x})$ for any $\vec{x} \in \mathbb{R}^n$.

Note: The theorem means that if $\mathcal{B}$ is an *orthonormal* basis for $\mathbb{R}^n$, then we now have a formula for finding the $\mathcal{B}$-coordinates of any $\vec{x} \in \mathbb{R}^n$.

Ex: Consider the orthonormal basis $\mathcal{B} = \left\{ \begin{bmatrix} 1/2 \\ 1/2 \\ 1/2 \\ 1/2 \end{bmatrix}, \begin{bmatrix} -1/2 \\ 1/2 \\ 1/2 \\ -1/2 \end{bmatrix}, \begin{bmatrix} 1/2 \\ -1/2 \\ 1/2 \\ -1/2 \end{bmatrix}, \begin{bmatrix} 1/2 \\ 1/2 \\ -1/2 \\ -1/2 \end{bmatrix} \right\}$ of $\mathbb{R}^4$. Find the $\mathcal{B}$-coordinates of $\vec{x} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}$.

Recall: $[x]_{\mathcal{B}} = \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}$ where $\vec{x} = a\vec{u}_1 + b\vec{u}_2 + c\vec{u}_3 + d\vec{u}_4$

old method: solve these four linear eq'ns.

new method: $a = \vec{u}_1 \cdot \vec{x} = \frac{1}{2} + \frac{2}{2} + \frac{3}{2} + \frac{4}{2} = 5$

$b = \vec{u}_2 \cdot \vec{x} = -\frac{1}{2} + \frac{2}{2} + \frac{3}{2} - \frac{4}{2} = 0$

$c = \vec{u}_3 \cdot \vec{x} = \frac{1}{2} - \frac{2}{2} + \frac{3}{2} - \frac{4}{2} = -1$

$d = \vec{u}_4 \cdot \vec{x} = \frac{1}{2} + \frac{2}{2} - \frac{3}{2} - \frac{4}{2} = -2$

$\Rightarrow [x]_{\mathcal{B}} = \begin{bmatrix} 5 \\ 0 \\ -1 \\ -2 \end{bmatrix}$

Recall: Let V be a subspace of $\mathbb{R}^n$. The *orthogonal complement* of V is the set $V^\perp$ of all vectors $\vec{x} \in \mathbb{R}^n$ which are orthogonal to V. (Note that $V^\perp$ is the kernel of the orthogonal projection proj_V onto V).

Properties: Let V be a subspace of $\mathbb{R}^n$.

1. Then $V^\perp$ is a subspace of $\mathbb{R}^n$.
2. $V \cap V^\perp = \{\vec{0}\}$.
3. $\dim(V) + \dim(V^\perp) = n$.
4. $(V^\perp)^\perp = V$.

Pythagorean Theorem: Let $\vec{x}, \vec{y} \in \mathbb{R}^n$. Then $\vec{x} \perp \vec{y}$ if and only if $||\vec{x} + \vec{y}||^2 = ||\vec{x}||^2 + ||\vec{y}||^2$.

Proof:

$$\|\vec{x}+\vec{y}\|^2 = (\vec{x}+\vec{y})\cdot(\vec{x}+\vec{y}) = \vec{x}\cdot\vec{x} + \vec{x}\cdot\vec{y} + \vec{y}\cdot\vec{x} + \vec{y}\cdot\vec{y}$$

$$= \|\vec{x}\|^2 + 2\vec{x}\cdot\vec{y} + \|\vec{y}\|^2$$

$$= \|\vec{x}\|^2 + 0 + \|\vec{y}\|^2$$

b/c $\vec{x}\cdot\vec{y} = 0$ (b/c $\vec{x} \perp \vec{y}$).

Section 5.2: Gram-Schmidt Process and QR Factorization

Method: (The Gram-Schmidt Process) Let $\vec{v}_1, \ldots, \vec{v}_m$ be a basis for a subsapce V in $\mathbb{R}^n$. Let

$$\begin{aligned} \vec{w}_1 &= \vec{v}_1 \\ \vec{w}_2 &= \vec{v}_2 - \frac{\vec{v}_2 \cdot \vec{w}_1}{\vec{w}_1 \cdot \vec{w}_1}\vec{w}_1 \\ \vec{w}_3 &= \vec{v}_3 - \frac{\vec{v}_3 \cdot \vec{w}_1}{\vec{w}_1 \cdot \vec{w}_1}\vec{w}_1 - \frac{\vec{v}_3 \cdot \vec{w}_2}{\vec{w}_2 \cdot \vec{w}_2}\vec{w}_2 \\ &\vdots \end{aligned}$$

Then $\vec{w}_1, \ldots, \vec{w}_m$ is an orthogonal basis for V.

Let $\vec{u}_1 = \dfrac{\vec{w}_1}{||\vec{w}_1||}, \vec{u}_2 = \dfrac{\vec{w}_2}{||\vec{w}_2||}, \ldots$. Then $\vec{u}_1, \ldots, \vec{u}_m$ is an orthonormal basis for V.

Ex: Let $B = \left\{ \vec{v}_1 = \begin{bmatrix} -3 \\ 4 \end{bmatrix}, \vec{v}_2 = \begin{bmatrix} 1 \\ 7 \end{bmatrix} \right\}$. Then B is a basis for $\mathbb{R}^2$. Use the Gram-Schmidt process to find an orthonormal basis $\mathcal{U} = \{\vec{u}_1, \vec{u}_2\}$ of $\mathbb{R}^2$.

did this in class 2

① $\vec{w}_1 = \vec{v}_1 = \begin{bmatrix} -3 \\ 4 \end{bmatrix}$

$\vec{w}_2 = \vec{v}_2 - \dfrac{\vec{v}_2 \cdot \vec{w}_1}{\vec{w}_1 \cdot \vec{w}_1}\vec{w}_1$

$= \begin{bmatrix} 1 \\ 7 \end{bmatrix} - \frac{25}{25}\begin{bmatrix} -3 \\ 4 \end{bmatrix}$

$= \begin{bmatrix} 4 \\ 3 \end{bmatrix}$

this is an orthogonal basis for $\mathbb{R}^2$.

② normalize:

$\vec{u}_1 = \vec{w}_1 / ||\vec{w}_1|| = \dfrac{\vec{w}_1}{5} = \begin{bmatrix} -3/5 \\ 4/5 \end{bmatrix}$

$\vec{u}_2 = \vec{w}_2 / ||\vec{w}_2|| = \dfrac{\vec{w}_2}{5} = \begin{bmatrix} 4/5 \\ 3/5 \end{bmatrix}$

Then $\mathcal{U} = \{\vec{u}_1, \vec{u}_2\}$ is an orthonormal basis for $\mathbb{R}^2$ (can check).

Ex: Let $V = \text{Im}(M)$ where $M = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & -1 \\ 1 & 0 \end{bmatrix}$.

1. Find an orthonormal basis for V.

2. Find the change of basis matrix R from the basis $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -1 \\ 0 \end{bmatrix} \right\}$ to the basis $\mathcal{U}$ found in 1.

1. ① $\vec{w}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$

$\vec{w}_2 = \begin{bmatrix} 0 \\ 1 \\ -1 \\ 0 \end{bmatrix} - \frac{0}{4}\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ -1 \\ 0 \end{bmatrix}$

(so $\vec{v}_1$ & $\vec{v}_2$ were already orthogonal).

did this in class 1

② $\vec{u}_1 = \dfrac{\vec{w}_1}{2} = \begin{bmatrix} 1/2 \\ 1/2 \\ 1/2 \\ 1/2 \end{bmatrix}$

$\vec{u}_2 = \dfrac{\vec{w}_2}{\sqrt{2}} = \begin{bmatrix} 0 \\ 1/\sqrt{2} \\ -1/\sqrt{2} \\ 0 \end{bmatrix}$

$\Rightarrow \mathcal{U} = \left\{ \begin{bmatrix} 1/2 \\ 1/2 \\ 1/2 \\ 1/2 \end{bmatrix}, \begin{bmatrix} 0 \\ 1/\sqrt{2} \\ -1/\sqrt{2} \\ 0 \end{bmatrix} \right\}$

2. $R = \begin{bmatrix} [\vec{v}_1]_{\mathcal{U}} & [\vec{v}_2]_{\mathcal{U}} \end{bmatrix}$

$[\vec{v}_1]_{\mathcal{U}} = (\vec{u}_1 \cdot \vec{v}_1)\vec{u}_1 \overset{0}{+ (\vec{u}_2 \cdot \vec{v}_1)\vec{u}_2} = (\frac{1}{2}+\frac{1}{2}+\frac{1}{2}+\frac{1}{2})\vec{u}_1 = 2\vec{u}_1 + 0\vec{u}_2$

$= \cancel{2\vec{u}_1 + 0\vec{u}_2} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$

$[\vec{v}_2]_{\mathcal{U}} = (\vec{u}_1 \cdot \vec{v}_2)\vec{u}_1 + (\vec{u}_2 \cdot \vec{v}_2)\vec{u}_2$

$= 0\vec{u}_1 + \frac{2}{\sqrt{2}}\vec{u}_2 = 0\vec{u}_1 + \sqrt{2}\vec{u}_2$

$= \begin{bmatrix} 0 \\ \sqrt{2} \end{bmatrix}$

$\Rightarrow R = \begin{bmatrix} 2 & 0 \\ 0 & \sqrt{2} \end{bmatrix}.$

Thm: (QR-factorization) Let M be an $n \times m$ matrix with linearly independent columns $\vec{v}_1, \ldots, \vec{v}_m$. Then there exists and $n \times m$ matrix Q whose columns $\vec{u}_1, \ldots, \vec{u}_m$ are orthonormal, and an upper triangular matrix R with positive diagonal entries, such that

$$M = QR$$

Furthermore, $r_{11} = \|\vec{v}_1\|$, $r_{jj} = \|\vec{v}_j^{\perp}\|$ for $2 \leq j \leq m$, and $r_{ij} = \vec{u}_i \cdot \vec{v}_j$ for $i < j$.

Ex: Verify the QR-factorization theorem for the last example.

Let $Q = \begin{bmatrix} 1/2 & 0 \\ 1/2 & 1/\sqrt{2} \\ 1/2 & -1/\sqrt{2} \\ 1/2 & 0 \end{bmatrix}$, $R = \begin{bmatrix} 2 & 0 \\ 0 & \sqrt{2} \end{bmatrix}$.

Then $QR = \begin{bmatrix} 1+0 & 0+0 \\ 1+0 & 0+1 \\ 1+0 & 0-1 \\ 1+0 & 0+0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & -1 \\ 1 & 0 \end{bmatrix} = M$ ✓.

Ex: Find the QR-factorization of $M = \begin{bmatrix} 2 & 2 \\ 1 & 7 \\ -2 & 1 \end{bmatrix}$. $\vec{v}_1 = \begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix}$, $\vec{v}_2 = \begin{bmatrix} 2 \\ 7 \\ 1 \end{bmatrix}$.

A) Gram-Schmidt to find columns $\vec{u}_i$ of Q:

① $\vec{w}_1 = \vec{v}_1 = \begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix}$, $\vec{w}_2 = \vec{v}_2 - \frac{\vec{v}_2 \cdot \vec{w}_1}{\vec{w}_1 \cdot \vec{w}_1} \vec{w}_1 = \begin{bmatrix} 2 \\ 7 \\ 1 \end{bmatrix} - \frac{9}{9} \begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix} = \begin{bmatrix} 0 \\ 6 \\ 3 \end{bmatrix}$

② $\vec{u}_1 = \frac{w_1}{\|\vec{w}_1\|} = \frac{w_1}{3} = \begin{bmatrix} 2/3 \\ 1/3 \\ -2/3 \end{bmatrix}$, $\vec{u}_2 = \frac{\vec{w}_2}{\|\vec{w}_2\|} = \frac{\vec{w}_2}{\sqrt{45}} = \begin{bmatrix} 0 \\ 6/\sqrt{45} \\ 3/\sqrt{45} \end{bmatrix} = \begin{bmatrix} 0 \\ 2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix}$.

So $Q = \begin{bmatrix} 2/3 & 0 \\ 1/3 & 2/\sqrt{5} \\ -2/3 & 1/\sqrt{5} \end{bmatrix}$.

B) $R = \left[[\vec{v}_1]_{\mathcal{U}} \ [\vec{v}_2]_{\mathcal{U}} \right]$:

$[\vec{v}_1]_{\mathcal{U}} = (\vec{u}_1 \cdot \vec{v}_1)\vec{u}_1 + (\vec{u}_2 \cdot \vec{v}_1)\vec{u}_2 = 3\vec{u}_1 + 0\vec{u}_2 = \begin{bmatrix} 3 \\ 0 \end{bmatrix}$.

$[\vec{v}_2]_{\mathcal{U}} = (\vec{u}_1 \cdot \vec{v}_2)\vec{u}_1 + (\vec{u}_2 \cdot \vec{v}_2)\vec{u}_2 = 3\vec{u}_1 + \sqrt{45}\,\vec{u}_2 = \begin{bmatrix} 3 \\ \sqrt{45} \end{bmatrix}$

$\Rightarrow R = \begin{bmatrix} 3 & 3 \\ 0 & \sqrt{45} \end{bmatrix}$.

C) Check: $QR = M$ ✓.

Section 5.3: Orthogonal Transformations and Orthogonal Matrices

Def: A linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ is called an *orthogonal linear transformation* if T preserves the lengths of vectors, i.e. if

$$\|T(\vec{x})\| = \|\vec{x}\| \text{ for all } \vec{x} \in \mathbb{R}$$

Ex: The rotation by θ matrix $T(\vec{x}) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \vec{x}$ of $\mathbb{R}^2$ is an orthogonal transformation.

Ex: Let V be a subspace of $\mathbb{R}^n$. Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be the transformation which reflects each vector across the subspace V. Then for all $\vec{x} \in \mathbb{R}^n$, $T(\vec{x}) = \mathrm{ref}_V(\vec{x}) = \vec{x}_p - \vec{x}_o$. Show that T is orthogonal.

$$\begin{aligned} \|T(\vec{x})\|^2 = \|\vec{x}_p - \vec{x}_o\|^2 &= \|\vec{x}_p\|^2 + \|-\vec{x}_o\|^2 \\ &= \|\vec{x}_p\|^2 + \|\vec{x}_o\|^2 \\ &= \|\vec{x}_p + \vec{x}_o\|^2 \quad \text{by Pyth. Thm since } \vec{x}_p \perp \vec{x}_o \\ &= \|\vec{x}\|^2 \quad \text{since } \vec{x} = \vec{x}_p + \vec{x}_o. \end{aligned}$$

Thm:
1. A linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ is orthogonal if and only if the vectors $T(\vec{e}_1), \ldots, T(\vec{e}_n)$ form an orthonormal basis for $\mathbb{R}^n$.
2. An $n \times n$ matrix A is orthogonal if and only if its columns form an orthonormal basis of $\mathbb{R}^n$.

Ex: Is the matrix $A = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$ orthogonal?

No: the columns are orthogonal, but they aren't unit vectors.

Ex: Is the matrix $A = \frac{1}{2}\begin{bmatrix} -1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 1 \\ -1 & -1 & -1 & 1 \end{bmatrix}$ orthogonal?

Yes: the columns are orthogonal and they are unit vectors.
So the columns are orthonormal $\Rightarrow$ A is orthogonal.

Thm: Suppose A and B are orthogonal matrices. Then A^{-1} and AB are orthogonal. In other words, inverses and products of orthogonal matrices are orthogonal.

Def: An $n \times n$ matrix A is *symmetric* if $A^T = A$, and *skew-symmetric* if $A^T = -A$.

Ex: The matrix A below is symmetric and the matrix B below is skew-symmetric.

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 5 & 2 & 3 \\ 3 & 2 & 2 & 1 \\ 4 & 3 & 1 & 7 \end{bmatrix}, \qquad B = \begin{bmatrix} 0 & 1 & 2 \\ -1 & 0 & -3 \\ -2 & 3 & 0 \end{bmatrix}$$

Ex: Find all 2×2 skew-symmetric matrices.

If $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is skew-symm., then $A^T = -A$, i.e. $\begin{bmatrix} a & c \\ b & d \end{bmatrix} = \begin{bmatrix} -a & -b \\ -c & -d \end{bmatrix}$

$a = -a \Rightarrow a = 0$
$d = -d \Rightarrow d = 0$ $\Rightarrow A = \begin{bmatrix} 0 & b \\ -b & 0 \end{bmatrix}$.
$b = -c$

Thm: An $n \times n$ matrix A is orthogonal if and only if $A^T A = I_n$, i.e. if and only if $A^{-1} = A^T$.

Summary: Let A be an $n \times n$ matrix. Then the following are equivalent:

1. A is an orthogonal matrix.
2. The transformation $L(\vec{x}) = A\vec{x}$ preserves length: $||A\vec{x}|| = ||\vec{x}||$ for every $\vec{x} \in \mathbb{R}^n$. (def)
3. The columns of A form an orthonormal basis of $\mathbb{R}^n$. (thm 1)
4. $A^T A = I_n$. (thm 2)
5. $A^{-1} = A^T$. (thm 2)

Properties: of A^T and A^{-1}: Let A and B be $n \times n$ matrices.

1. $(AB)^T = B^T A^T$.
2. $(AB)^{-1} = B^{-1} A^{-1}$.
3. $(A^T)^{-1} = (A^{-1})^T$

Ex: If A is symm., B is skew-symm., & A & B commute, then AB is skew-symm:
$(AB)^T = B^T A^T = (-B)A = -(AB)$

Thm: Let V be a subspace of $\mathbb{R}^n$ with orthonormal basis $\vec{u}_1, \ldots, \vec{u}_m$. The linear transformation $T(\vec{x}) = \text{proj}_V(\vec{x})$ has standard matrix QQ^T where $Q = [\vec{u}_1 \quad \vec{u}_2 \quad \cdots \quad \vec{u}_m]$.

Ex: Find the matrix of orthogonal projection onto the subspace of $\mathbb{R}^4$ spanned by the vectors

$$\vec{u}_1 = \frac{1}{2}\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \qquad \vec{u}_2 = \frac{1}{2}\begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix}.$$

$\mathcal{B} = \{\vec{u}_1, \vec{u}_2\}$ is an orthonormal basis of this subspace.

So proj_V has mtx $QQ^T = \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & -1/2 \\ 1/2 & 1/2 \\ 1/2 & -1/2 \end{bmatrix} \begin{bmatrix} 1/2 & 1/2 & 1/2 & 1/2 \\ 1/2 & -1/2 & 1/2 & -1/2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}$.

i.e. $\text{proj}_V(\vec{x}) = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix} \vec{x}$ for all $\vec{x}$ in $\mathbb{R}^4$.