For $y \in W$ = a subspace of $\mathbb{R}^n$

Suppose $\mathcal{B} = \{\boldsymbol{v_1}, ..., \boldsymbol{v_i}, ...\boldsymbol{v_p}\}$ = orthogonal basis for W

$$\boldsymbol{y} = \underset{\updownarrow}{\boldsymbol{c_1}}\ \boldsymbol{v_1} + ... + \underset{\updownarrow}{\boldsymbol{c_i}}\ \boldsymbol{v_i} ... + ... + \underset{\updownarrow}{\boldsymbol{c_n}}\ \boldsymbol{v_p}$$

$$\frac{\boldsymbol{y}\cdot\boldsymbol{v_1}}{\boldsymbol{v_1}\cdot\boldsymbol{v_1}}\ \boldsymbol{v_1} + ... + \frac{\boldsymbol{y}\cdot\boldsymbol{v_i}}{\boldsymbol{v_i}\cdot\boldsymbol{v_i}}\ \boldsymbol{v_i} + ... + \frac{\boldsymbol{y}\cdot\boldsymbol{v_p}}{\boldsymbol{v_p}\cdot\boldsymbol{v_p}}\ \boldsymbol{v_p} \quad (**)$$

If it happens that

$\mathcal{B} = \{\boldsymbol{u_1}, ..., \boldsymbol{u_i}, ..., \boldsymbol{u_p}\}$ = orthonormal basis for subspace W of $\mathbb{R}^n$
(orthogonal and each $||\boldsymbol{u_i}|| = 1$)

Then, for $y \in W$:

$$\boldsymbol{y} = (\boldsymbol{y}\cdot\boldsymbol{u_1})\,\boldsymbol{u_1} + ... + (\boldsymbol{y}\cdot\boldsymbol{u_i})\,\boldsymbol{u_i} + ... + (\boldsymbol{y}\cdot\boldsymbol{u_p})\,\boldsymbol{u_p}$$

since every denominator in $(**)$ now $= 1$

Let $\boldsymbol{u} = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$ and $L = \text{Span}\{\boldsymbol{u}\}$. Write $\boldsymbol{y} = \begin{bmatrix} 5 \\ 0 \\ -3 \end{bmatrix}$ as the sum of a vector from L and a vector from $L^{\perp}$. Find the perpendicular distance from $\boldsymbol{y}$ to L.

$$\widehat{\boldsymbol{y}} = \text{proj}_L \boldsymbol{y} = \frac{\begin{bmatrix} 5 \\ 0 \\ -3 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}}{\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}} \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} = \frac{2}{6} \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{3} \\ -\frac{2}{3} \\ \frac{1}{3} \end{bmatrix}, \text{ and}$$

$$\boldsymbol{z} = \boldsymbol{y} - \widehat{\boldsymbol{y}} = \begin{bmatrix} 5 \\ 0 \\ -3 \end{bmatrix} - \begin{bmatrix} \frac{1}{3} \\ -\frac{2}{3} \\ \frac{1}{3} \end{bmatrix} = \begin{bmatrix} \frac{14}{3} \\ \frac{2}{3} \\ -\frac{10}{3} \end{bmatrix}$$

(*Note:* $\boldsymbol{z} = \begin{bmatrix} \frac{14}{3} \\ \frac{2}{3} \\ -\frac{10}{3} \end{bmatrix} \perp \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} = \widehat{\boldsymbol{y}}$, *as should be true*; *a check)*

The distance from $\boldsymbol{y}$ to L is $||\boldsymbol{y} - \widehat{\boldsymbol{y}}|| = \sqrt{\frac{196}{9} + \frac{4}{9} + \frac{100}{9}} = \sqrt{\frac{300}{9}} = \frac{10\sqrt{3}}{3}$.

$\frac{10\sqrt{3}}{3}$ is the <u>error</u> if $\widehat{\boldsymbol{y}}$ is used as an approximation for $\boldsymbol{y}$**;** if some other $\boldsymbol{v}$ on L is used to approximate $\boldsymbol{y}$, then the error $||\boldsymbol{y} - \boldsymbol{v}|| > \frac{10\sqrt{3}}{3}$, so $\widehat{\boldsymbol{y}}$ is the "best approximation" to $\boldsymbol{y}$ by a vector

from L.

$\mathcal{B} = \{\boldsymbol{v_1}, ..., \boldsymbol{v_i}, ...\boldsymbol{v_n}\}$ = orthogonal basis for $\mathbb{R}^n$

For $\boldsymbol{x} \in W$:

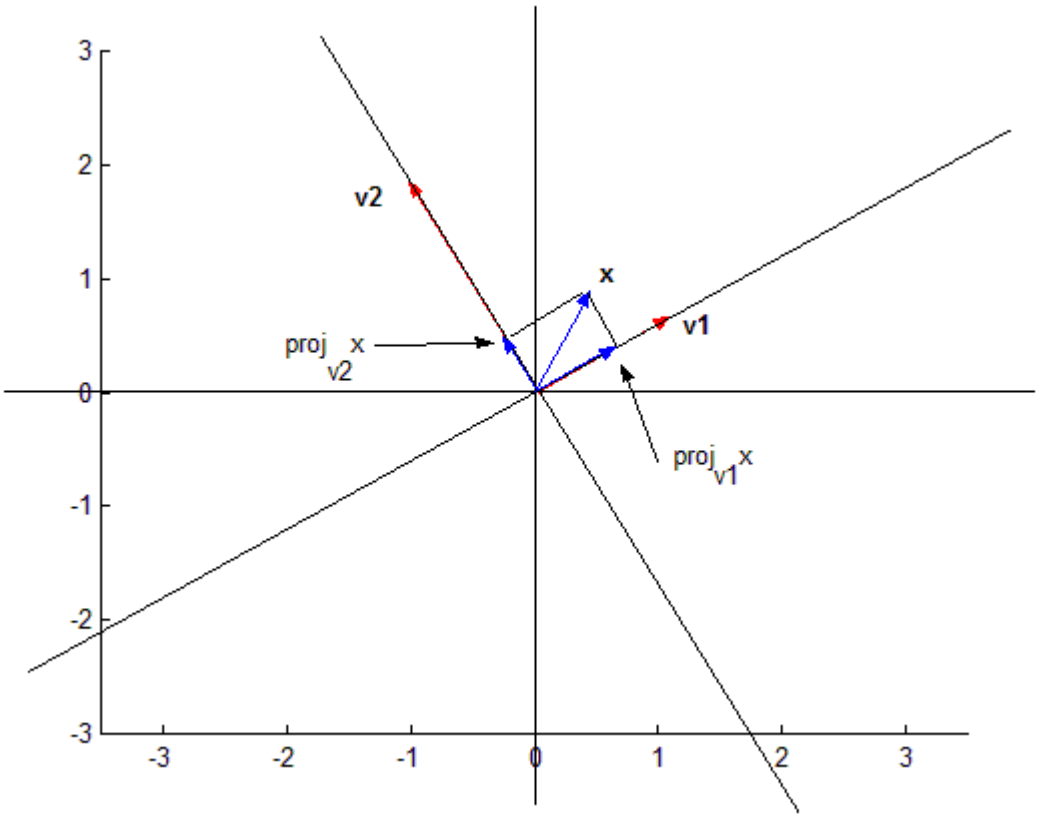
$$\boldsymbol{x} = \frac{\boldsymbol{x}\cdot\boldsymbol{v_1}}{\boldsymbol{v_1}\cdot\boldsymbol{v_1}}\boldsymbol{v_1} + ... + \frac{\boldsymbol{x}\cdot\boldsymbol{v_i}}{\boldsymbol{v_i}\cdot\boldsymbol{v_i}}\boldsymbol{v_i} + ... + \frac{\boldsymbol{x}\cdot\boldsymbol{v_n}}{\boldsymbol{v_n}\cdot\boldsymbol{v_n}}\boldsymbol{v_v}$$

$$= \text{proj}_{\boldsymbol{v_1}}\boldsymbol{x} + ... + \text{proj}_{\boldsymbol{v_i}}\boldsymbol{x} + ... + \text{proj}_{\boldsymbol{v_n}}\boldsymbol{y}$$

For an orthogonal basis, writing $\boldsymbol{x}$ in terms of the basis vectors amounts to writing $\boldsymbol{x}$ as a sum of projections onto those vectors.

An simple illustration of this: For $\mathcal{B} = \{\boldsymbol{v_1}, \boldsymbol{v_2}\}$ an orthogonal basis for $\mathbb{R}^2$:

$$\text{each } \boldsymbol{x} = \frac{\boldsymbol{x}\cdot\boldsymbol{v_1}}{\boldsymbol{v_1}\cdot\boldsymbol{v_1}}\boldsymbol{v_1} + \frac{\boldsymbol{x}\cdot\boldsymbol{v_2}}{\boldsymbol{v_2}\cdot\boldsymbol{v_2}}\boldsymbol{v_2} = \text{proj}_{\boldsymbol{v_1}}\boldsymbol{x} + \text{proj}_{\boldsymbol{v_2}}\boldsymbol{x}$$



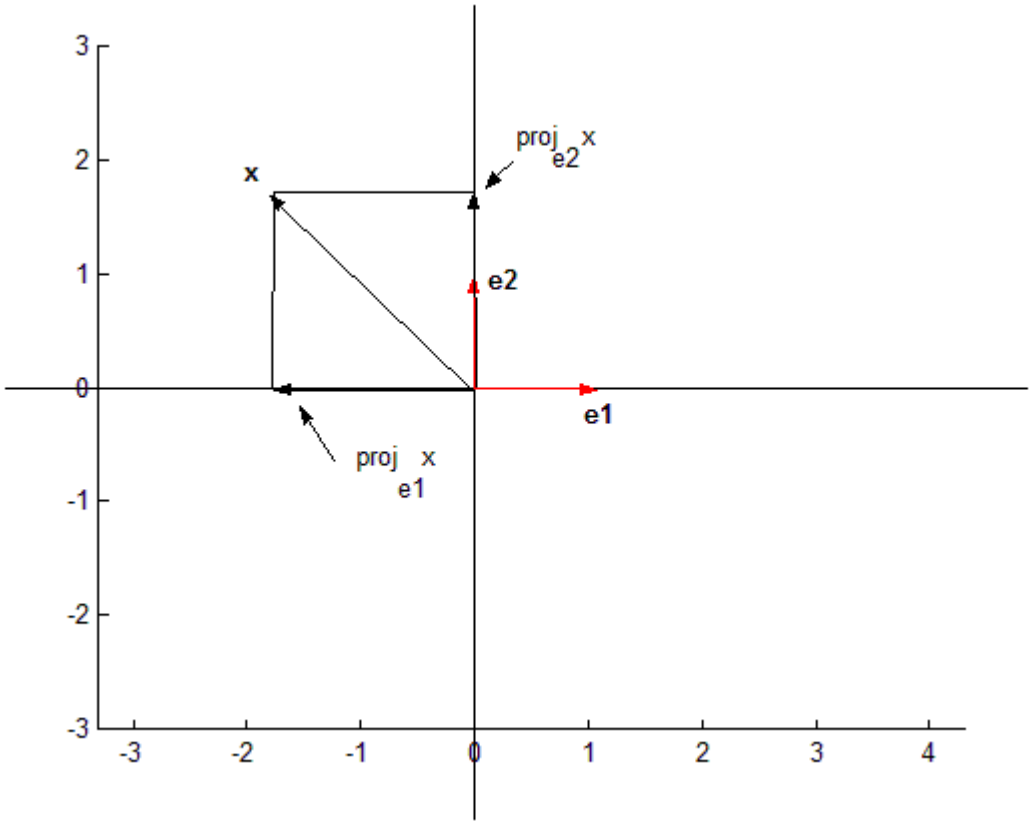
$\mathcal{B} = \{\boldsymbol{u_1}, ..., \boldsymbol{u_i}, ..., \boldsymbol{u_n}\}$ = orthonormal basis for $\mathbb{R}^n$

For $y \in W$:

$\boldsymbol{x} = (\boldsymbol{x} \cdot \boldsymbol{u_1})\,\boldsymbol{u_1} + ... + (\boldsymbol{x} \cdot \boldsymbol{u_i})\,\boldsymbol{u_i} + ... + (\boldsymbol{x} \cdot \boldsymbol{u_n})\boldsymbol{u_n}$

$= \text{proj}_{\boldsymbol{u_1}}\,\boldsymbol{x} + ... + \text{proj}_{\boldsymbol{u_i}}\,\boldsymbol{x} + ... + \text{proj}_{\boldsymbol{u_n}}\,\boldsymbol{x}$

An even simpler illustration using the orthonormal basis $\mathcal{B} - \{\boldsymbol{e_1}, \boldsymbol{e_2}\}$ for $\mathbb{R}^2$.



An $m \times n$ matrix A has <u>orthonormal columns</u>. The columns <u>must</u> be linearly independent: this means that $m \geq n$ (*why?* : A can be square or "tall and thin" $\begin{bmatrix} & \\ & \\ & \end{bmatrix}$ but <u>not</u> "short and wide" $\begin{bmatrix} & & \\ & & \end{bmatrix}$. $M \geq b$

Theorem An $m \times n$ matrix A has <u>orthonormal columns</u> if and only if $A^T A = I$

Proof *(Just look at* 3×3 *– only to make the notation look simpler)*

Suppose $A = [\boldsymbol{a_1}\ \boldsymbol{a_2}\ \boldsymbol{a_3}]$,

$$A^T = \begin{bmatrix} \leftarrow \boldsymbol{a_1^T} \rightarrow \\ \leftarrow \boldsymbol{a_2^T} \rightarrow \\ \leftarrow \boldsymbol{a_3^T} \rightarrow \end{bmatrix}, \quad A^T A = \begin{bmatrix} \boldsymbol{a_1^T} \\ \boldsymbol{a_2^T} \\ \boldsymbol{a_3^T} \end{bmatrix} [\boldsymbol{a_1}\ \boldsymbol{a_2}\ \boldsymbol{a_3}]$$

$$= [\begin{bmatrix} \boldsymbol{a_1^T} \\ \boldsymbol{a_2^T} \\ \boldsymbol{a_3^T} \end{bmatrix} \boldsymbol{a_1} \quad \begin{bmatrix} \boldsymbol{a_1^T} \\ \boldsymbol{a_2^T} \\ \boldsymbol{a_3^T} \end{bmatrix} \boldsymbol{a_2} \quad \begin{bmatrix} \boldsymbol{a_1^T} \\ \boldsymbol{a_2^T} \\ \boldsymbol{a_3^T} \end{bmatrix} \boldsymbol{a_3}\ ]$$

Notice

$$\begin{bmatrix} \boldsymbol{a_1^T} \\ \boldsymbol{a_2^T} \\ \boldsymbol{a_n^T} \end{bmatrix} \boldsymbol{a_1} = \begin{bmatrix} a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \\ a_{13} & a_{23} & a_{m3} \end{bmatrix} \begin{bmatrix} a_{11} \\ a_{21} \\ a_{31} \end{bmatrix} = \begin{bmatrix} \boldsymbol{a_1^T a_1} \\ \boldsymbol{a_2^T a_1} \\ \boldsymbol{a_3^T a_1} \end{bmatrix}$$

and <u>in general</u>

$$A^T A = \begin{bmatrix} \boldsymbol{a_1^T a_1} & \boldsymbol{a_1^T a_2} & \boldsymbol{a_1^T a_3} \\ \boldsymbol{a_2^T a_1} & \boldsymbol{a_2^T a_2} & \boldsymbol{a_2^T a_3} \\ \boldsymbol{a_3^T a_1} & \boldsymbol{a_3^T a_2} & \boldsymbol{a_3^T a_3} \end{bmatrix}$$

$$= \begin{bmatrix} \boldsymbol{a_1 \cdot a_1} & \boldsymbol{a_1 \cdot a_2} & \boldsymbol{a_1 \cdot a_3} \\ \boldsymbol{a_2 \cdot a_1} & \boldsymbol{a_2 \cdot a_2} & \boldsymbol{a_2 \cdot a_3} \\ \boldsymbol{a_3 \cdot a_1} & \boldsymbol{a_3 \cdot a_2} & \boldsymbol{a_3 \cdot a_3} \end{bmatrix}$$

so $A^T A = I$ if and only if $\begin{cases} \boldsymbol{a_1 \cdot a_1 = a_2 \cdot a_2 = a_3 \cdot a_3 = 1} \quad) \\ \boldsymbol{a_i \cdot a_j} = 0 \ \text{ if } \ i \neq j \end{cases}$

Definition If $\begin{cases} A \text{ has orthonormal columns, and} \\ A \text{ is square} \end{cases}$

then A is called an orthogonal matrix

By the preceding result, A is orthogonal if and only if A is square, invertible and $A^T = A^{-1}$

Suppose A is $m \times n$, so that $T(\boldsymbol{x}) = A\boldsymbol{x}$ is a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$.

If A has orthonormal columns, then the transformation "preserves orthogonality":

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

$$A\boldsymbol{x} \perp A\boldsymbol{y} \quad \Leftrightarrow \quad (A\boldsymbol{x}) \cdot (A\boldsymbol{y}) = 0 \quad \Leftrightarrow \ (A\boldsymbol{x})^T (A\boldsymbol{y}) = 0$$

$$\Leftrightarrow \ \boldsymbol{x}^T A^T A\boldsymbol{y} = 0 \quad \Leftrightarrow \ \boldsymbol{x}^T I\, \boldsymbol{y} = \boldsymbol{0} \quad \Leftrightarrow \boldsymbol{x} \cdot \boldsymbol{y} = 0 \ \Leftrightarrow \quad \boldsymbol{x} \perp \boldsymbol{y}.$$

Example $\boldsymbol{A} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ is an orthogonal matrix; $\boldsymbol{x} \mapsto A\boldsymbol{x}$ represents a rotation of the plane through $90°$. Notice that if $x \perp \boldsymbol{y}$, then $A\boldsymbol{x} \perp A\boldsymbol{y}$.

Similarly, any rotation matrix $\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ is orthogonal.

Can you find a 2×2 matrix A where A is an orthogonal matrix, but the transformation $\boldsymbol{x} \mapsto A\boldsymbol{x}$ is not a rotation ?