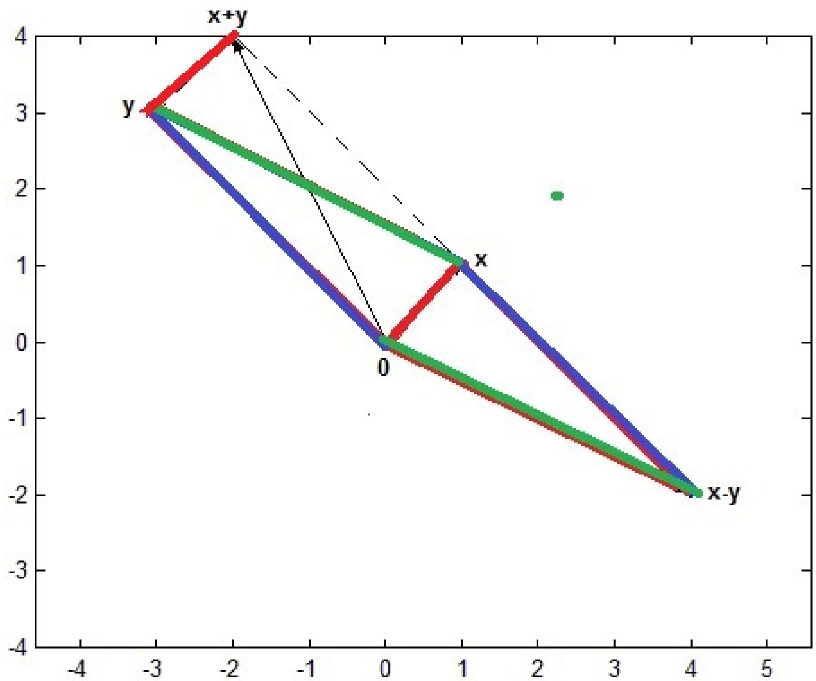
Pythagorean Theorem in $\mathbb{R}^n$

Suppose $\boldsymbol{x}, \boldsymbol{y}$ are in $\mathbb{R}^n$ and that $\boldsymbol{x} \perp \boldsymbol{y}$.

Then $||\,\boldsymbol{x}\,||^2 + ||\,\boldsymbol{y}\,||^2 \quad \begin{cases} = ||\,\boldsymbol{x} - \boldsymbol{y}\,||^2 \\ = ||\,\boldsymbol{x} + \boldsymbol{y}\,||^2 \end{cases}$ and also



Proof $||\,\boldsymbol{x} - \boldsymbol{y}\,||^2$

$$
\begin{aligned}
||\,\boldsymbol{x} - \boldsymbol{y}\,||^2 &= (\boldsymbol{x} - \boldsymbol{y}) \cdot (\boldsymbol{x} - \boldsymbol{y}) \\
&= \boldsymbol{x} \cdot \boldsymbol{x} - 2\,\boldsymbol{x} \cdot \boldsymbol{y} + \boldsymbol{y} \cdot \boldsymbol{y} \\
&= ||\,\boldsymbol{x}\,||^2 - 0 + ||\,\boldsymbol{y}\,||^2 \\
&= ||\,\boldsymbol{x}\,||^2 + ||\,\boldsymbol{y}\,||^2
\end{aligned}
$$

and, $||\,\boldsymbol{x} + \boldsymbol{y}\,||^2$

$$
\begin{aligned}
||\,\boldsymbol{x} + \boldsymbol{y}\,||^2 &= ||\,\boldsymbol{x} - (-\boldsymbol{y})\,||^2 \\
&= ||\,\boldsymbol{x}\,||^2 + ||\,-\boldsymbol{y}\,||^2 \\
&= ||\,\boldsymbol{x}\,||^2 + ||\,\boldsymbol{y}\,||^2
\end{aligned}
$$

Orthogonal Decomposition Theorem

W a subspace of $\mathbb{R}^n$. Each $\boldsymbol{y} \in \mathbb{R}^n$ can be written in a unique way as

$$\boldsymbol{y} = \widehat{\boldsymbol{y}} + \boldsymbol{z} \text{ where } \begin{cases} \widehat{\boldsymbol{y}} & \text{is in } W \\ \boldsymbol{z} & \text{is in } W^{\perp} \end{cases}$$

If $W = \{\boldsymbol{u_1}, \dots, \boldsymbol{u_p}\}$ is an orthogonal basis for W, then formulas are

$$\widehat{\boldsymbol{y}} = \frac{\boldsymbol{y} \cdot \boldsymbol{u_1}}{\boldsymbol{u_1} \cdot \boldsymbol{u_1}} \boldsymbol{u_1} + \dots + \frac{\boldsymbol{y} \cdot \boldsymbol{u_p}}{\boldsymbol{u_p} \cdot \boldsymbol{u_p}} \boldsymbol{u_p}$$

and $\boldsymbol{z} = \boldsymbol{y} - \widehat{\boldsymbol{y}}$

$\widehat{\boldsymbol{y}}$ is called the orthogonal projection of $\boldsymbol{y}$ on W : $\widehat{\boldsymbol{y}} = \text{proj}_W \boldsymbol{y}$

$\boldsymbol{z}$ is called the component of $\boldsymbol{y}$ orthogonal to W

Best Approximation Theorem

$\widehat{\boldsymbol{y}}$ is the point in W closest to $\boldsymbol{y}$ meaning that:

if $\boldsymbol{v} \in W$ and $\boldsymbol{v} \neq \widehat{\boldsymbol{y}}$, then

$$|\,|\boldsymbol{y} - \widehat{\boldsymbol{y}}|| < ||\boldsymbol{y} - \boldsymbol{v}||$$

(therefore $\widehat{\boldsymbol{y}}$ is the vector in W that "best approximates" $\boldsymbol{y}$

Let $W = \text{Span}\{ \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ -1 \\ 1 \end{bmatrix} \} \subseteq \mathbb{R}^4$.

$$\boldsymbol{y} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \notin W$$

Example

Find the point in W closest to $\boldsymbol{y}$ and find the distance from $\boldsymbol{y}$ to W.

$$\widehat{\boldsymbol{y}} = \frac{\boldsymbol{y}\cdot\boldsymbol{u}_1}{\boldsymbol{u}_1\cdot\boldsymbol{u}_1}\boldsymbol{u}_1 + \frac{\boldsymbol{y}\cdot\boldsymbol{u}_2}{\boldsymbol{u}_2\cdot\boldsymbol{u}_2}\boldsymbol{u}_2 + \frac{\boldsymbol{y}\cdot\boldsymbol{u}_3}{\boldsymbol{u}_3\cdot\boldsymbol{u}_3}\boldsymbol{u}_3$$

$$= \ldots + \frac{\begin{bmatrix}1\\1\\1\\1\end{bmatrix}\begin{bmatrix}2\\0\\1\\1\end{bmatrix}}{\begin{bmatrix}2\\0\\1\\1\end{bmatrix}\begin{bmatrix}2\\0\\1\\1\end{bmatrix}}\begin{bmatrix}2\\0\\1\\1\end{bmatrix} + \ldots$$

so

$$\widehat{\boldsymbol{y}} = \frac{1}{1}\begin{bmatrix}0\\1\\0\\0\end{bmatrix} + \frac{4}{6}\begin{bmatrix}2\\0\\1\\1\end{bmatrix} + \frac{0}{2}\begin{bmatrix}0\\0\\-1\\1\end{bmatrix} = \begin{bmatrix}\frac{4}{3}\\1\\\frac{2}{3}\\\frac{2}{3}\end{bmatrix}$$

$= \text{point in W closest to } \boldsymbol{y}.$

Distance from $\boldsymbol{y}$ to $W = ||\boldsymbol{z}|| = ||\boldsymbol{y} - \widehat{\boldsymbol{y}}||$

$$= \left|\left|\begin{bmatrix}-\frac{1}{3}\\0\\\frac{1}{3}\\\frac{1}{3}\end{bmatrix}\right|\right| = \sqrt{\frac{1}{9}+\frac{1}{9}+\frac{1}{9}} = \frac{1}{\sqrt{3}} = \frac{\sqrt{3}}{3}.$$

$\begin{bmatrix}\frac{4}{3}\\1\\\frac{2}{3}\\\frac{2}{3}\end{bmatrix}$ is "best approximation" to $\boldsymbol{y} = \begin{bmatrix}1\\1\\1\\1\end{bmatrix}$ by a point from W.

$||\boldsymbol{y} - \widehat{\boldsymbol{y}}||$ is the size of <u>error</u> when $\widehat{\boldsymbol{y}}$ is used as an approximation for $\boldsymbol{y}$. If $\boldsymbol{v} \in W$ and $\boldsymbol{v} \neq \widehat{\boldsymbol{y}}$, then $||\boldsymbol{y} - \boldsymbol{v}|| > ||\boldsymbol{y} - \widehat{\boldsymbol{y}}|| = \frac{\sqrt{3}}{3}$.