# Dot Product / Length

**Dot Product** of 2 vectors $\vec{u}$ and $\vec{v}$ ($\in R^n$) is defined as $\vec{u}\cdot\vec{v} = \vec{u}^T\vec{v} = (u_1\ u_2 \dots)\begin{pmatrix} v_1 \\ v_2 \\ \vdots \end{pmatrix} = u_1v_1 + u_2v_2 + \dots$

↳ $\vec{a}\cdot\vec{b} = \|\vec{a}\|\,\|\vec{b}\|\cos\theta$

↳ $\vec{a}$ and/or $\vec{b}$ are 0 vectors, OR $\theta = 90°$

↳ What values of $k$ is $\vec{u}\cdot\vec{v} = 0$?

$\vec{u} = \begin{pmatrix} -1 \\ k \\ 2 \end{pmatrix}, \vec{v} = \begin{pmatrix} 4 \\ 1 \\ 3 \end{pmatrix}$

$\vec{u}\cdot\vec{v} = \vec{u}^T\vec{v} = -1(4) + k + 6 = 0$

$k + 2 = 0$

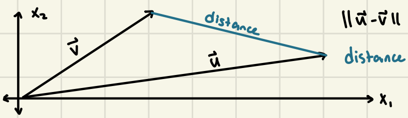
$k = -2$

**Unit vector** - if $\vec{v}$ has length 1

↳ $\vec{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $\frac{1}{\sqrt{5}}\begin{pmatrix} 1 \\ 2 \end{pmatrix}$, $\frac{1}{\sqrt{3}}\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$

**distance** between $\vec{u}$ and $\vec{v}$ is $\|\vec{u}-\vec{v}\|$

↳ Distance between $\vec{u} = \begin{pmatrix} 7 \\ 1 \end{pmatrix}$ and $\vec{v} = \begin{pmatrix} 3 \\ 2 \end{pmatrix}$




$\|\vec{u}-\vec{v}\| = \left\|\begin{pmatrix} 4 \\ -1 \end{pmatrix}\right\|$

distance $= \sqrt{17}$

If $\vec{u}, \vec{v}$, and $\vec{w}$ are in $R^n$, $c \in R$

↳ **linearity** → $(\vec{v}+\vec{w})\cdot\vec{u} = \vec{v}\cdot\vec{u} + \vec{w}\cdot\vec{u}$

↳ **Scalars** → $(c\vec{u})\cdot\vec{w} = c(\vec{u}\cdot\vec{w})$

↳ **Symmetry** → $\vec{u}\cdot\vec{w} = \vec{w}\cdot\vec{u}$

↳ **Positivity** → $\vec{u}\cdot\vec{u} \geq 0$ ← Only true b/c ALL entries are real

↳ $\vec{u}\cdot\vec{u} = 0 \xleftrightarrow{\text{if and only if}} \vec{u} = 0$

**length / magnitude** of a vector $\vec{u}$ is:

$\|\vec{u}\| = \sqrt{\vec{u}\cdot\vec{u}} = \sqrt{u_1^2 + u_2^2 + \dots u_n^2}$

↳ $\|c\vec{v}\| = |c|\,\|\vec{v}\|$ → length of $c\vec{v}$

↳ $\|\vec{u}\| = 5$, $\|\vec{v}\| = \sqrt{3}$, $\vec{u}\cdot\vec{v} = -1$

find $\|\vec{u}+\vec{v}\|$

$\|\vec{u}+\vec{v}\|^2 = \left(\sqrt{(\vec{u}+\vec{v})\cdot(\vec{u}+\vec{v})}\right)^2$

$= (\vec{u}+\vec{v})\cdot(\vec{u}+\vec{v})$

$= \underbrace{\vec{u}\cdot\vec{u}}_{\|\vec{u}\|^2} + \underbrace{\vec{v}\cdot\vec{v}}_{\|\vec{v}\|^2} + \vec{u}\cdot\vec{v} + \vec{v}\cdot\vec{u}$

$= 5^2 + \sqrt{3}^2 - 1 + -1$

$= 25 + 3 - 2$

$= 26 \rightarrow \|\vec{u}+\vec{v}\|^2 = 26$

$\sqrt{26} = \|\vec{u}+\vec{v}\|$

# Orthogonality

Two vectors are **orthogonal** if $\vec{u}\cdot\vec{v} = 0$

↳ if $\|\vec{u}+\vec{v}\|^2 = \|\vec{u}\|^2 + \|\vec{v}\|^2$ bc $\|\vec{u}+\vec{v}\|^2 = (\vec{u}+\vec{v})\cdot(\vec{u}+\vec{v})$

↳ pythagorean theorem — n-dimensional version

If $\vec{u} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $\vec{v} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$, then $\vec{u}\cdot\vec{v} = 0$, and

$\|\vec{u}+\vec{v}\|^2 = \left\|\begin{pmatrix} 2 \\ 0 \end{pmatrix}\right\|^2 = \left(\sqrt{2^2+0^2}\right)^2 = 4$

$\|\vec{u}\|^2 = \sqrt{\vec{u}\cdot\vec{u}}^2 = \sqrt{2}^2 = 2$

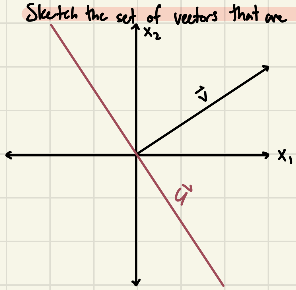
$\|\vec{v}\|^2 = \sqrt{\vec{v}\cdot\vec{v}}^2 = \sqrt{2}^2 = 2$

$\Big\}\ 4$

↳ $\vec{0}$ is **orthogonal** to every vector in $R^n$

↳ **orthogonality** usually refers to non-zero vectors

Sketch the set of vectors that are **orthogonal** to $\vec{v} = \begin{pmatrix} 3 \\ 2 \end{pmatrix}$. Is the set a subspace?




$\vec{u}\cdot\vec{v} = 0$

$(u_1\ u_2)\begin{pmatrix} 3 \\ 2 \end{pmatrix} = 0$

$\vec{u} = \begin{pmatrix} 2 \\ -3 \end{pmatrix}$ and any multiple of $\vec{u}$

$\vec{u}$ is a subspace bc the line (span) goes through the origin

# Orthogonal Compliments

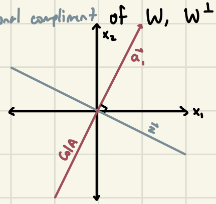
If $W$ is a subspace of $R^n$, $\vec{v}$ is **orthogonal** to $W$ if $\vec{z}$ is **orthogonal** to every vector in $W$

↳ The set of all vectors **orthogonal** to $W$ is the **orthogonal compliment** of $W$, $W^\perp$ ("W perp")

↳ $W^\perp = \{\vec{z} \in R^n$ such that $: \vec{z}\cdot\vec{w} = 0$ for all $\vec{w} \in W\}$

$A = \begin{pmatrix} 1 & 3 \\ 2 & 6 \end{pmatrix}$, find $(\text{Col}A)^\perp$ if Col A is span of $\vec{a}_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$

↳ $(\text{Col}A)^\perp$ is span $\vec{z} = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$




Find $(\text{Null } A)^\perp$

$\begin{pmatrix} 1 & 3 \\ 2 & 6 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 3 \\ 0 & 0 \end{pmatrix}$ $\begin{matrix} x_2 \text{ is free} \\ x_1 = -3x_2 \end{matrix}\Big\} \rightarrow x = x_2\begin{pmatrix} -3 \\ 1 \end{pmatrix}$ or

↳ $(\text{Null } A)^\perp$ is span $\begin{pmatrix} 1 \\ 3 \end{pmatrix}$

Line $L$ is subspace $R^3$ spanned by $\vec{v} = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}$

Space $L^\perp$ is a plane, construct $L^\perp$

if $\vec{u} \in L^\perp$ and $\vec{u} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$, then $\vec{u}\cdot\vec{v} = \begin{pmatrix} x \\ -y \\ 2z \end{pmatrix}_0$

$x - y + 2z = 0$, so if solving for $z$, $2z = \frac{y - x}{2}$

$z = \frac{y}{2} - \frac{x}{2}$

# The Four Fundamental Subspaces

**Row A (Row Space)** – space spanned by rows of matrix A

↳ Basis for Row A is given by pivot rows of A

↳ $\dim(\text{Row}A) = \dim(\text{Col}A)$

↳ $\text{Row}A = \text{Col}A^T$

↳ In general, RowA and ColA are NOT related → other than

$A = \begin{pmatrix} 1 & 3 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$, construct:

Row A: $\{(1\ 3\ 0), (0\ 0\ 1)\}$

$\text{Row}A^{\perp}$: NullA, $x_2$ is free, $\begin{matrix} x_1 = -3x_2 \\ x_3 = 0 \end{matrix} \rightarrow x = x_2\begin{pmatrix} -3 \\ 1 \\ 0 \end{pmatrix}$, $\left\{\begin{pmatrix} -3 \\ 1 \\ 0 \end{pmatrix}\right\}$

Col A: $\left\{\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}\right\}$

$\text{Col}A^{\perp}$: Null $A^T$, $\begin{pmatrix} 1 & 0 & 0 \\ 3 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ $\begin{matrix} x_1 = 0x_3 \\ x_2 = 0x_3 \\ x_3 \text{ is free} \end{matrix} \rightarrow x = x_3\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$, $\left\{\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\right\}$

If A is $\mathbb{R}^{m \times n}$, $n = \dim \text{Col}A + \dim \text{Nul}A$

↳ but, $n \overset{\text{also}}{=} \dim \text{Row}A + \dim \text{Null}A$

If $\vec{v}$ is in NulA, $A\vec{v} = 0$ ← by definition

↳ dot product of $A \cdot v = 0$, so $\vec{v}$ is orthogonal to rows of A

↳ Row A is orthogonal to NulA, $\text{Row}A^{\perp} = \text{Nul}A$ — vectors in $\mathbb{R}^n$

↳ Col A is orthogonal to $\text{Nul}A^T$, $\text{Col}A^{\perp} = \text{Nul}A^T$ — vectors in $\mathbb{R}^m$

Four Fundamental Subspaces

# Orthogonal Bases

A set of vectors $\{\vec{u}_1, \ldots \vec{u}_p\}$ are an orthogonal set of vectors if for each $j \neq k$, $\vec{u}_j \perp \vec{u}_k$ (all mutually orthogonal)

Make the orthogonal set:

$\vec{u}_1 = \begin{pmatrix} 4 \\ 0 \\ 1 \end{pmatrix}$, $\vec{u}_2 = \begin{pmatrix} -2 \\ 0 \\ 8 \end{pmatrix}$, $\vec{u}_3 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$ ← HAS to be nonzero!

$4(-2) + 0 + x = 0$
$x - 8 = 0$
$x = 8$

($u_2$ and $u_3$)
$-2(0) + 0y + z = 4(0) + 0y + z$

↳ If $\vec{u}$ are all non-zero, then the orthogonal set are ALL linearly independent

If $\{\vec{u}_1 \ldots \vec{u}_p\}$ is an orthogonal basis of W, if $\vec{w} \in W$, $\vec{w} = c_1\vec{u}_1 + \ldots c_p\vec{u}_p$

↳ the scalars are $c_q = \dfrac{\vec{w} \cdot \vec{u}_q}{u_q \cdot u_q}$ (More efficient than row reductions)

↳ You can ONLY apply the theorem if you have an orthogonal basis for W

W is an orthogonal set to $\vec{x}$; confirm that an orthogonal basis is given by $\vec{u}$ and $\vec{v}$ (a), and compute the expansion of $\vec{s}$ if $\vec{s}$ is a basis for W (b):

$\vec{x} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$, $\vec{u} = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$, $\vec{v} = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$, $\vec{s} = \begin{pmatrix} 3 \\ -4 \\ 1 \end{pmatrix}$

a) $\vec{u}$ and $\vec{v}$ have to be in W, and orthogonal

in W: $\vec{u} \cdot \vec{x} = 0$, $\vec{v} \cdot \vec{x} = 0$ orthogonal: $\vec{u} \cdot \vec{v} = 0$

$1 - 2 + 1 = 0$ ✓, $-1 + 0 + 1 = 0$ ✓ $-1 + 0(-2) + 1 = 0$, $-1 + 1 = 0$ ✓

b) expansion of $\vec{s}$, $\vec{s} = c_1\vec{u} + c_2\vec{v}$ — Both are in set W, proven previously

$= \dfrac{\vec{s} \cdot \vec{u}}{\vec{u} \cdot \vec{u}}\vec{u} + \dfrac{\vec{s} \cdot \vec{v}}{\vec{v} \cdot \vec{v}}\vec{v}$

$= \dfrac{12}{6}\vec{u} + \dfrac{-2}{2}\vec{v}$

$\vec{s} = 2\vec{u} - \vec{v}$

$c_p = \dfrac{\vec{w} \cdot \vec{u}_p}{\vec{u}_p \cdot \vec{u}_p}$ ↗ $\|\vec{u}_p\| = 1$

$c_p = \vec{w} \cdot \vec{u}_p$

**Orthonormal Basis** – orthogonal basis $\{\vec{u}_1 \ldots \vec{u}_p\}$ where every vector has unit length; $\vec{w} = (\vec{w} \cdot \vec{u}_1)\vec{u}_1 + \ldots (\vec{w} \cdot \vec{u}_p)\vec{u}_p$ AND

$\|\vec{w}\| = \sqrt{(\vec{w} \cdot \vec{u}_1)^2 + \ldots (\vec{w} \cdot \vec{u}_p)^2}$

W is perpendicular to $\vec{x}$, find coefficient for orthonormal basis for W $\{\vec{u}, \vec{v}\}$

$x = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$, $u = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$, $v = \frac{1}{\sqrt{6}}\begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$

u needs to $\in W$, and $u \cdot u = 1$

$u \cdot x = 0$, $\|u\| = 1$

$\frac{1}{\sqrt{\ }}(1\ 0\ -1)\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = 0$ | $\sqrt{1^2 + (-1)^2}\left(\frac{1}{\sqrt{\ }}\right) = 1$

$\sqrt{2}\left(\frac{1}{\sqrt{2}}\right) = 1$

v needs to $\in W$, $u \cdot v = 0$, $\|v\| = 1$

$v = \frac{1}{\sqrt{k}}\begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} \rightarrow u \cdot v$, $c_1 = c_3$

$x \cdot v = 0$, $\frac{1}{\sqrt{k}}(c_1\ c_2\ c_3)\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = 0$, $c_2 = -2c_1$

Choosing $c_1 = 1$, $c_2 = -2$ and $c_3 = 1$; if $\|v\| = 1$, $k = 6$

Bases are NOT unique!

↳ any vector $w \in \mathbb{R}^3$ can be written as any linear combination of $\{e_1, e_2, e_3\}$ such as $\{e_1, -e_2, e_3\}$

# Orthogonal Projections

$\vec{y}$ and $\vec{u}$ are 2 non-zero vectors, and $\vec{u}$ is in the span of W

↳ The closest vector in W is $\hat{y}$; how do you find $\hat{y}$?

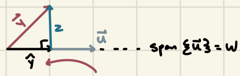
↳ $\vec{y} = \hat{y} + z$, $\vec{z} \in W^{\perp}$; how do you find $\hat{y}$ and $z$?

↳ $0 = \vec{z} \cdot \vec{u}$, $\vec{z} = \vec{y} - \underbrace{k\vec{u}}_{\hat{y}}$

↳ $0 = (\vec{y} - k\vec{u}) \cdot \vec{u}$

$= \vec{y} \cdot \vec{u} - k\vec{u} \cdot \vec{u} \to k = \frac{\vec{y} \cdot \vec{u}}{\vec{u} \cdot \vec{u}}$, $\vec{u} \neq \vec{0}$

$\hat{y} = \frac{\vec{y} \cdot \vec{u}}{\vec{u} \cdot \vec{u}} \vec{u}$




Orthogonal Projection

The orthogonal projection of $\vec{y}$ onto $\vec{u}$ is the vector in the span of $\vec{u}$ closest to $\vec{y}$, $\hat{y}$

$\text{proj}_{\vec{u}} \vec{y} = \frac{\vec{y} \cdot \vec{u}}{\vec{u} \cdot \vec{u}} \vec{u}$, $\vec{y} = \hat{y} + z$ where $\vec{z} \in W^{\perp}$

↳ $\|\vec{y}\|^2 = \|\hat{y}\|^2 + \|z\|^2$

If $\vec{u}$ is in 1D subspace S, and $S^{\perp}$ is also a 1D subspace, projection of $\vec{u}$ onto $S^{\perp}$ is $\vec{0}$

↳ If $\vec{x} \in S^{\perp}$, $\vec{x} \neq \vec{0}$, then $\text{proj}_{\vec{x}} \vec{u} = \frac{\vec{u} \cdot \vec{x}}{\vec{x} \cdot \vec{x}} \vec{x} \to \frac{0}{\vec{x} \cdot \vec{x}} \vec{x} = \vec{0}$ TRUE

L is spanned by $\vec{u}$, $\vec{u} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$, $\vec{y} = \begin{pmatrix} 3 \\ 4 \\ 2 \end{pmatrix}$; Calculate projection $\vec{y}$ onto L; What is distance btw $\vec{y}$ / L

$\text{proj}_{\vec{u}} \vec{y} = \frac{\vec{u} \cdot \vec{y}}{\vec{u} \cdot \vec{u}} \vec{u} = \frac{9}{3} \vec{u} = \begin{pmatrix} 3 \\ 3 \\ 3 \end{pmatrix} \longrightarrow \vec{y} = \vec{z} + \hat{y} \to \begin{pmatrix} 3 \\ 4 \\ 2 \end{pmatrix} = \vec{z} + \begin{pmatrix} 3 \\ 3 \\ 3 \end{pmatrix}$, $\vec{z} = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$,

distance $= \|\vec{z}\|$, $\sqrt{0^2 + 1^2 + (-1)^2}$

$= \sqrt{2}$

# Matrices with Orthonormal Columns

An $m \times n$ matrix $U$ has orthonormal columns if and only if $U^T U = I_n$

↳ $n$ cannot be $> m$, BUT $m$ can be $\geq n$

(orthonormal matrix)

Orthogonal Matrix – square matrix with orthonormal columns

↳ If $U$ is an orthogonal matrix, $U^{-1} = U^T$

↳ $\det U = 1$ or $-1$

Proof: $1 = \det I = \det(A^T A) = \det A^T \cdot \det A = 1 \cdot 1$ or $-1 \cdot -1$

1) $\|U\vec{x}\| = \|\vec{x}\|$

2) $(U\vec{x}) \cdot (U\vec{y}) = \vec{x} \cdot \vec{y}$

3) $(U\vec{x}) \cdot (U\vec{y}) = 0 \leftrightarrow \vec{x} \cdot \vec{y} = 0$

$\vec{x} \to U\vec{x}$ preserves length + orthogonality

Proof: $\|Ux\|^2 = (U\vec{x}) \cdot (U\vec{x})$

$= x^T U^T U \vec{x}$

$= x^T I \vec{x}$

$= x^T x = \|\vec{x}\|^2$, $\|\vec{x}\| = \|Ux\|$

True or False? 1) If $U$ is orthogonal, its columns are linearly independent true!

2) If the determinant of a matrix is 1, then the matrix must be orthogonal false!

# The Best Approximation Theorem

W is a subspace of $R^n$, $\vec{y} \in R^n$, $\hat{y}$ is the orthogonal projection of $\vec{y}$ onto W $\to$ Then for ANY vector $\vec{v} \neq \hat{y}$, $\vec{v} \in W$:

↳ $\|\vec{y} - \hat{y}\| < \|\vec{y} - \vec{v}\| \to \hat{y}$ is the unique vector in W closest to $\vec{y}$

↳ $\vec{y} - \vec{v} = \vec{y} - \vec{v} + (\hat{y} - \hat{y}) = \underset{\in W^{\perp}}{(\vec{y} - \hat{y})} + \underset{\in W}{(\hat{y} + \vec{v})}$

If $v$ is a vector in $R^n$ and W is a subspace, then $\text{proj}_W(\text{proj}_W \vec{v}) = \text{proj}_W \vec{v}$ true!

↳ $\text{proj}_W \vec{v} = \hat{v}$, $\hat{v} \in W$, $\text{proj}_W \hat{v} = \hat{v}$, so $\text{proj}_W \hat{v} = \hat{v}$

What is the distance between $\vec{y}$ and subspace $W = \text{span}\{\vec{u}_1, \vec{u}_2\}$? $y = \begin{pmatrix} 4 \\ 0 \\ 3 \end{pmatrix}$, $\vec{u}_1 = \begin{pmatrix} 2 \\ 2 \\ 0 \end{pmatrix}$, $\vec{u}_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$, $\hat{y} = \begin{pmatrix} 2 \\ 2 \\ 3 \end{pmatrix}$

$y = z + \hat{y} \to \begin{pmatrix} 4 \\ 0 \\ 3 \end{pmatrix} = z + \begin{pmatrix} 2 \\ 2 \\ 3 \end{pmatrix} \to z = \begin{pmatrix} 2 \\ -2 \\ 0 \end{pmatrix} \to \|z\|^2 = 2^2 + (-2)^2 + 0^2 \longrightarrow z = \sqrt{8} \to z = 2\sqrt{2}$

# Orthogonal Decomposition Theorem

$\vec{u}_1 \ldots \vec{u}_s$ is an orthonormal basis for $R^S$ – Let W span $\{\vec{u}_1, \vec{u}_s\}$

↳ For any vector $\in R^S$, can we construct $\hat{y}$ and $z$ so that $\vec{y} = \hat{y} + z$, $z \in W^{\perp}$ and $\hat{y} \in W$

↳ $u_i$ spans $R^S$, so $\vec{y} = \overset{\text{linear combination}}{\sum_{i=1}^{s} c_i \vec{u}_i}$

$= \underbrace{c_1\vec{u}_1 + c_2\vec{u}_2}_{\text{in } W} + \underbrace{c_3\vec{u}_3 + c_4\vec{u}_4 + c_5\vec{u}_5}_{W^{\perp}\text{, bc the set is orthonormal}}$

$= \hat{y} + z$

Let W be a subspace of $R^n$ – Then $\vec{y} \in R^n$ has the unique decomposition $\vec{y} = \hat{y} + z$, $\hat{y} \in W$, $z \in W^{\perp}$

And, if $\vec{u}_1 \ldots \vec{u}_p$ is any orthogonal basis for W,

$\hat{y} = \frac{\vec{y} \cdot \vec{u}_1}{\vec{u}_1 \cdot \vec{u}_1} \vec{u}_1 + \ldots + \frac{\vec{y} \cdot \vec{u}_p}{\vec{u}_p \cdot \vec{u}_p} \vec{u}_p$ ↳ $\hat{y}$ is the orthogonal projection onto W

Construct the decomposition $\vec{y} = \hat{y} + z$, where $\hat{y}$ is the orthogonal projection of $\vec{y}$ onto $W = \text{Span}\{\vec{u}_1, \vec{u}_2\}$, and $z \in W^\perp$

↳ $\vec{y} = \begin{pmatrix} 4 \\ 0 \\ 3 \end{pmatrix}, \vec{u}_1 = \begin{pmatrix} 2 \\ 2 \\ 0 \end{pmatrix}, \vec{u}_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$

$\hat{y} = \frac{8}{8}\begin{pmatrix} 2 \\ 2 \\ 0 \end{pmatrix} + \frac{3}{1}\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$

$\hat{y} = \begin{pmatrix} 2 \\ 2 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 3 \end{pmatrix}$

$\hat{y} = \begin{pmatrix} 2 \\ 2 \\ 3 \end{pmatrix}$

$\vec{y} = \hat{y} + z$

$\begin{pmatrix} 4 \\ 0 \\ 3 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \\ 3 \end{pmatrix} + z$

$z = \begin{pmatrix} 2 \\ -2 \\ 0 \end{pmatrix}$

$\vec{y} = \begin{pmatrix} 2 \\ 2 \\ 3 \end{pmatrix} + \begin{pmatrix} 2 \\ -2 \\ 0 \end{pmatrix}$

# Grand Schmidt Process

MUST be linearly independent!

Grand Schmidt Process – if given a set of vectors $\{\vec{x}_1, \vec{x}_2\}$ in $\mathbb{R}^n$, we can construct an orthogonal basis, $\{\vec{v}_1, \vec{v}_2\}$ for the space they span

$\vec{v}_1 = x_1, \quad \vec{v}_2 = \vec{x}_2 - \frac{\vec{x}_2 \cdot \vec{x}_1}{\vec{x}_1 \cdot \vec{x}_1}\vec{x}_1$

↳ IF linearly dependent, $v_2$ is 0!

↳ If $\{\vec{x}_1, \dots \vec{x}_p\}$ is a subspace $W$ of $\mathbb{R}^n$, $W_1 = \text{Span}\{\vec{v}_1\}$ $\quad v_1 = \vec{x}_1$

$W_2 = \text{Span}\{\vec{v}_1, \vec{v}_2\} \quad v_2 = \vec{x}_2 - \text{proj}_{W_1}\vec{x}_2$

$\vdots$

$W_p = \text{Span}\{\vec{v}_1 \dots \vec{v}_p\} \quad \vec{v}_p = \vec{x}_p - \text{proj}_{W_{p-1}}\vec{x}_p$

Let $W$ be the subspace of $\mathbb{R}^3$ spanned by $\vec{x}_1$ and $\vec{x}_2$ - Construct an orthogonal basis for $W$.

$x_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad x_2 = \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix}$

↳ Set $v_1 = x_1$, $v_2 = x_2 - \text{proj}_{x_1}x_2$

$= x_2 - \frac{x_2 \cdot x_1}{x_1 \cdot x_1}x_1$

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

$v_2 = \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}$

orthogonal basis to span $\{x_1, x_2\}$ is $\left\{\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}\right\}$

Vectors span a subspace $W$ for $\mathbb{R}^4$ - construct an orthogonal basis:

$x_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix}, x_2 = \begin{pmatrix} 0 \\ 1 \\ 2 \\ 1 \end{pmatrix}, x_3 = \begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix}$

$v_1 = x_1$, $v_2 = x_2 - \text{proj}_{W_1}x_2$

$= x_2 - \frac{x_2 \cdot x_1}{x_1 \cdot x_1}x_1$

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

$v_2 = \begin{pmatrix} -1 \\ 0 \\ 1 \\ 1 \end{pmatrix}$

$v_3 = x_3 - \text{proj}_{W_2}x_3$

$= x_3 - \left(\frac{x_3 \cdot v_1}{v_1 \cdot v_1}v_1 + \frac{x_3 \cdot v_2}{v_2 \cdot v_2}v_2\right)$

$= x_3 - \left(\frac{0}{3}v_1 + \frac{-1}{3}v_2\right)$

$= x_3 + \frac{1}{3}v_2$

$v_3 = \frac{1}{3}\begin{pmatrix} -1 \\ 3 \\ -2 \\ 1 \end{pmatrix}$

Orthogonal basis for $W$ is $\{v_1, v_2, v_3\}$

$\left\{\begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \\ 1 \end{pmatrix}, \frac{1}{3}\begin{pmatrix} -1 \\ 3 \\ -2 \\ 1 \end{pmatrix}\right\}$

# QR Factorization

A set of vectors form an orthonormal basis if the vectors are mutually orthogonal and have unit length

↳ Make an orthonormal basis from the orthogonal basis of $v_1 = \begin{pmatrix} 3 \\ 2 \\ 0 \end{pmatrix}, v_2 = \begin{pmatrix} -2 \\ 3 \\ 1 \end{pmatrix}$

↳ Length of $v_1 = \sqrt{13}$, so $\frac{1}{\sqrt{13}}v_1 = \hat{v}_1$, $\frac{1}{\sqrt{14}}v_2 = \hat{v}_2$

orthonormal basis

Any $m \times n$ matrix with linearly independent columns has the QR Factorization → $A = QR$

matrix

$m \times n$, columns are orthonormal basis for Col A

$n \times n$, upper triangular (positive entries on diagonal)

↳ $Q$ is obtained by Grand-Schmidt Process

↳ $R$ is obtained by $R = Q^TA$, because $Q$ is an orthonormal basis ($Q^TQ = I$)

$Q^T = Q^{-1}$

↳ length of the $j^{th}$ column of $R$ = length of $j^{th}$ column of $A$

Construct a QR factorization for $A = \begin{pmatrix} 3 & -2 \\ 2 & 3 \\ 0 & 1 \end{pmatrix}$

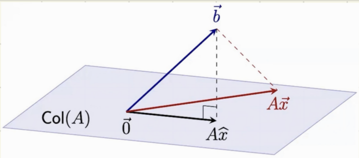
↳ Columns of Q form an orthonormal basis for Col A, but columns of A are already orthogonal (★ Special Case ★)

↳ $Q = \begin{pmatrix} 3/\sqrt{13} & -2/\sqrt{14} \\ 2/\sqrt{13} & 3/\sqrt{14} \\ 0 & 1/\sqrt{14} \end{pmatrix}$, $R = Q^T A$, $R = \begin{pmatrix} 13/\sqrt{13} & 0 \\ 0 & 14/\sqrt{14} \end{pmatrix}$

## Least-Squares Solution to Inconsistent Systems

Let A be an $m \times n$ matrix – a least-squares solution to $A\vec{x} = \vec{b}$ is the solution $\hat{x}$ for which $\|\vec{b} - A\hat{x}\| \leq \|\vec{b} - A\vec{x}\|$ for all $\vec{x} \in \mathbb{R}^n$

↳ Identify $\vec{x}$ that minimizes $\|\vec{b} - A\vec{x}\|$, denoted as $\hat{x}$




The closest vector in Col$A$ to $\vec{b}$ is $A\hat{x}$.

If $\vec{b} \in \text{Col}\,A$, then $A\hat{x} = \vec{b}$ is consistent

↳ We need $\hat{x}$ so that $A\hat{x}$ is as close to $\vec{b}$ as possible

Least-Squares Error – distance from $\vec{b}$ to $A\hat{x}$

## Normal Equations

orthogonal decomposition theorem – if $A\hat{x}$ is the closest vector in Col A to be $\vec{b}$, then $\vec{b} - Ax$ is $\text{Col}\,A^{\perp}$ → $\text{Col}\,A^{\perp} = \text{Null}\,A^T$

↳ $A^T(\vec{b} - A\hat{x}) = \vec{0} \Rightarrow A^T A\hat{x} = A^T\vec{b}$

Normal Equations → $A^T A\hat{x} = A^T\vec{b}$

↳ Theorems:

→ Columns are linearly independent

→ Matrix $A^T A$ is invertible

→ Equation $A\vec{x} = \vec{b}$ has a unique least-squares solution for each $\vec{b} \in \mathbb{R}^m$

} If all of these are true, then least-squares is $\hat{x} = (A^T A)^{-1} A^T\vec{b}$

Compute least-squares for $A\vec{x} = \vec{b}$

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

$A^T A\hat{x} = A^T b$, $\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & -1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & -1 \end{pmatrix}\hat{x} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & -1 \end{pmatrix}\begin{pmatrix} 2 \\ 2 \\ -1 \end{pmatrix}$

$\begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix}\hat{x} = \begin{matrix} 2+2+-1 \\ 0+2+1 \end{matrix}$

$\begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix}\hat{x} = \begin{pmatrix} 3 \\ 3 \end{pmatrix}$

$\begin{pmatrix} 3 & 0 & 3 \\ 0 & 2 & 3 \end{pmatrix} \rightarrow \hat{x} = \begin{pmatrix} 1 \\ 3/2 \end{pmatrix}$

## QR and Least-Squares

If $A \in \mathbb{R}^{m \times n}$ has linearly independent columns, then $A = QR$, and for every $\vec{b} \in \mathbb{R}^m$, $A\vec{x} = \vec{b}$ has the unique least-squares solution $\underset{\text{invertible + upper triangular}}{R}\hat{x} = Q^T\vec{b}$

↳ 1. construct QR decomposition of A → get Q, orthonormal basis of Col A (using Gram-Schmidt), get $R = Q^T A$

2. Solve $R\hat{x} = Q^T b$ to get $\hat{x}$

Given data:

| X | -2 | -1 | 0 | 1 | 2 |
|---|---|---|---|---|---|
| Y | -2 | -1 | 1 | 2 | 2 |

, Find $y = c_1 + c_2 x$

$A\vec{x} = \begin{pmatrix} 1 & -2 \\ 1 & -1 \\ 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} -2 \\ -1 \\ 1 \\ 2 \\ 2 \end{pmatrix} = \vec{b} \rightarrow QR = \begin{pmatrix} 1/\sqrt{5} & -2/\sqrt{10} \\ \vdots & \vdots \\ 1/\sqrt{5} & 2/\sqrt{10} \end{pmatrix}\begin{pmatrix} \frac{5}{\sqrt{5}} & 0 \\ 0 & \frac{10}{\sqrt{10}} \end{pmatrix}$, $Q^T\vec{b} = \begin{pmatrix} 2\sqrt{5} \\ 11/\sqrt{10} \end{pmatrix}$, $R\hat{x} = Q^T\vec{b}$, $\hat{x} = \begin{pmatrix} 2/5 \\ 11/10 \end{pmatrix}$

# Residuals and Least-Squares

| $x_i$ | 2 | 5 | 7 | 8 |
|---|---|---|---|---|
| $y_i$ | 1 | 1 | 4 | 3 |

Find $c_0$ and $c_1$ so that $y = c_0 + c_1 x$

↳ $A\hat{x} = \vec{b}$

$$\begin{pmatrix} 1 & 2 \\ 1 & 5 \\ 1 & 7 \\ 1 & 8 \end{pmatrix}\begin{pmatrix} c_0 \\ c_1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 4 \\ 3 \end{pmatrix}$$

→ Normal, $A^TA\hat{x} = A^T\vec{b}$

→ QR, $R\hat{x} = Q^T\vec{b}$

→ $y = c_0 + c_1\vec{x} = \frac{-5}{21} + \frac{19}{42}x$

↳ Model Fit? → $\hat{y}_i = c_0 + c_1 x_i = \frac{-5}{21} + \frac{19}{42}x_i$ (for estimates)

| $x_i$ | 2 | 5 | 7 | 8 |
|---|---|---|---|---|
| $y_i$ | 1 | 1 | 4 | 3 |
| $\hat{y}_i$ | .67 | 2.02 | 2.93 | 3.38 |

Residuals — $\vec{x}$ that minimized $\|A\vec{x} - \vec{y}\|$ over all possible $\vec{x} \in \mathbb{R}^n$ → equal to minimizing $\|A\vec{x} - \vec{y}\|^2$

↳ $\vec{r} = A\vec{x} - \vec{y}$, $\|A\vec{x} - y\|^2 = \|\vec{r}\|^2 = \sum_{i=1}^{n} r_i^2$ ← entries of $\vec{r}$

(trying to minimize sum)

↳ $\|r\|^2$ is the squared distance between $\vec{y}$ and Col A

# Mean-Deviation Form

Common practice when using model $y = c_0 + c_1 x$ → compute average, $\bar{x}$ of the x-values to introduce a new variable $x_* = x - \bar{x}$

mean-deviation form

| $x_i$ | 2 | 5 | 7 | 8 |
|---|---|---|---|---|
| $y_i$ | 1 | 1 | 4 | 3 |

→ average value of $x_i$ is $\bar{x} = 5.5$, x-values $- \bar{x}$ → $y = c_0 + c_1 x_*$

↳ $\begin{pmatrix} 1 & -3.5 \\ 1 & -0.5 \\ 1 & 1.5 \\ 1 & 2.5 \end{pmatrix}\begin{pmatrix} c_0 \\ c_1 \end{pmatrix} = A\hat{x} = \begin{pmatrix} 1 \\ 1 \\ 4 \\ 3 \end{pmatrix} = \vec{b}$ → $A^TA = \underbrace{\begin{pmatrix} 4 & 0 \\ 0 & 21 \end{pmatrix}}_{\text{diagonal}}$, $A^T\vec{y} = \begin{pmatrix} 9 \\ 9.5 \end{pmatrix}$ → $y = \frac{-5}{21} + \frac{19}{42}x_i$

# General Linear Model

↳ Useful to fit data points with curves / non-straight lines

↳ Compute least-squares solution with $A^TA\hat{x} = A^T\vec{y}$

↳ Model data with $\vec{y} = A\vec{x} + \vec{r}$ ← minimize $\vec{r}$ length

↳ Least Squares "fit" → $y = c_0 + c_1 f_1(x) + c_2 f_2(x) + \dots c_k f_k(x)$

↳ If $f_i$ is known, then $c_i$ variables are a <u>linear</u> problem

Second Order Polynomial → Find coefficients $c_1$ and $c_2$ for $y = c_1 x + c_2 x^2$ that <u>best fits</u> the data

| x | -1 | 0 | 0 | 1 |
|---|---|---|---|---|
| y | 2 | 1 | 0 | 6 |

→ $y = c_1 x + c_1 x^2$ → 

$$\begin{aligned} -c_1 + c_2 &= 2 \\ 0c_1 + 0c_2 &= 1 \\ 0c_1 + 0c_2 &= 0 \\ c_1 + c_2 &= 6 \end{aligned}$$

→ $\begin{pmatrix} -1 & 1 \\ 0 & 0 \\ 0 & 0 \\ 1 & 1 \end{pmatrix}\vec{x} = \begin{pmatrix} 2 \\ 1 \\ 0 \\ 6 \end{pmatrix}$ → $A^TA\hat{x} = A^T\vec{b}$, $\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}\hat{x} = \begin{pmatrix} 4 \\ 8 \end{pmatrix}$, $\hat{x} = \begin{array}{c} c_1 \\ \begin{pmatrix} 2 \\ 4 \end{pmatrix} \\ c_2 \end{array}$ → $y = 2x + 4x^2$

Multiple Regression → $z = c_0 + c_1 f_1(x,y) + c_2 f_2(x,y) + \dots c_k f_k(x,y)$

↳ If $f_i$ is known, it is another linear problem in the $c_i$ variables

Model in form $z = c_0 + c_1 x + c_2 y$

| x | -2 | -1 | 0 | 0 | 1 | 2 |
|---|---|---|---|---|---|---|
| y | 1 | 1 | -3 | -1 | 1 | 1 |
| z | 2 | 1 | 1 | 0 | -2 | -2 |

→

$$\begin{aligned} c_0 - 2c_1 + c_2 &= 2 \\ c_0 - c_1 + c_2 &= 1 \\ c_0 + 0c_1 - 3c_2 &= 1 \\ c_0 + 0c_1 - c_2 &= 0 \\ c_0 + c_1 + c_2 &= -2 \\ c_0 + 2c_1 + c_2 &= -2 \end{aligned}$$

→ $\begin{pmatrix} 1 & -2 & 1 \\ 1 & -1 & 1 \\ 1 & 0 & -3 \\ 1 & 0 & -1 \\ 1 & 1 & 1 \\ 1 & 2 & 1 \end{pmatrix}\vec{x} = \begin{pmatrix} 2 \\ 1 \\ 1 \\ 0 \\ -2 \\ -2 \end{pmatrix}$ → $A^TA\hat{x} = A^T\vec{b}$ → $\begin{pmatrix} 6 & 0 & 0 \\ 0 & 10 & 0 \\ 0 & 0 & 14 \end{pmatrix}\hat{x} = \begin{pmatrix} 0 \\ -11 \\ 4 \end{pmatrix}$ → $z = -\frac{11}{10}x - \frac{2}{7}y$