STAT 714 fa 2023

Linear algebra review 4/6

Orthogonal subspaces, bases, projections, Gram-Schmidt

$\underset{\sim}{u}, \underset{\sim}{v}$ orth. means

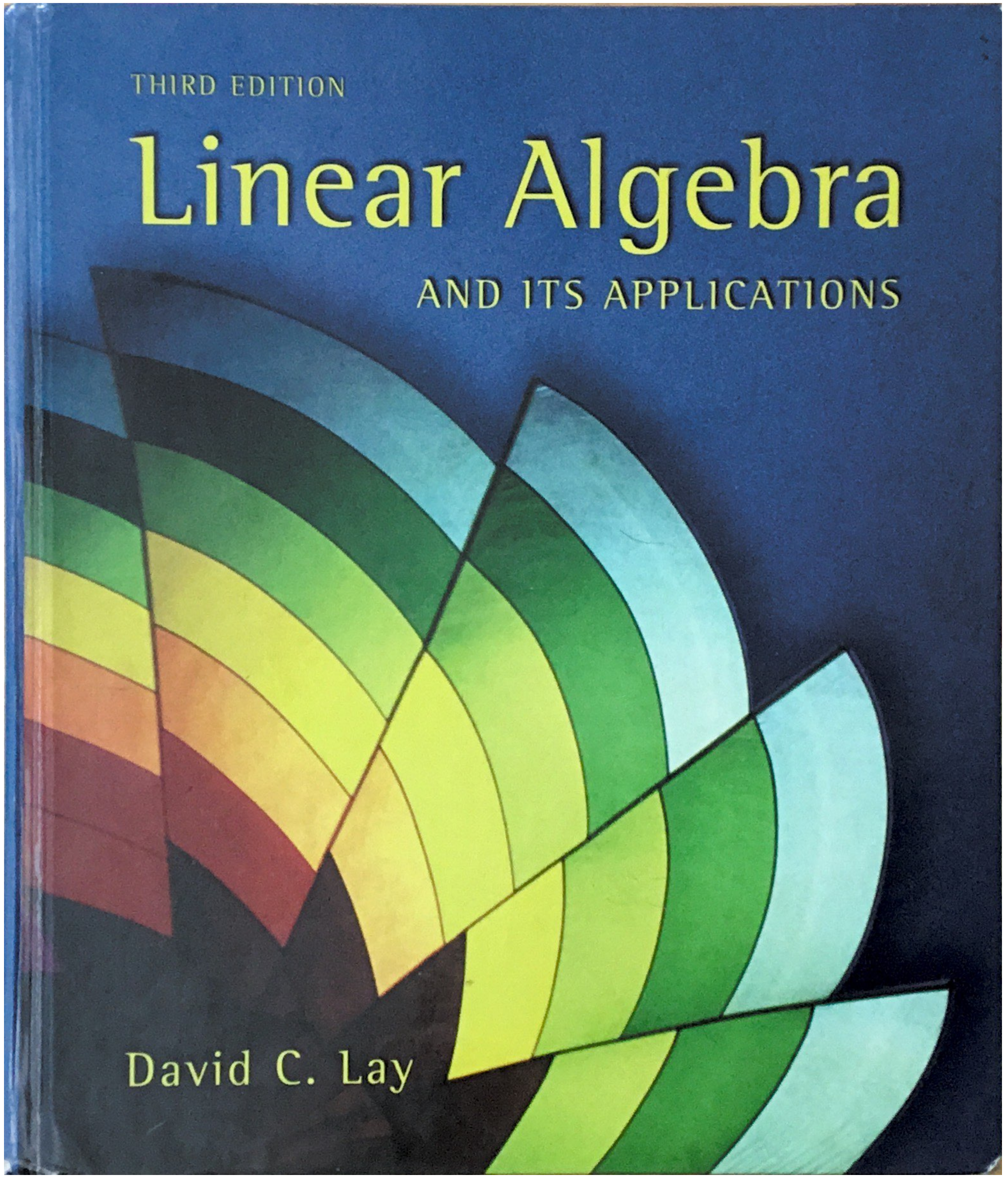
$\underset{\sim}{u} \cdot \underset{\sim}{v} = 0$

Karl B. Gregory

University of South Carolina

These slides are an instructional aid; their sole purpose is to display, during the lecture, definitions, plots, results, etc. which take too much time to write by hand on the blackboard. They are not intended to explain or expound on any material.

These notes include scanned excerpts from Lay (2003):



Orthogonal and orthonormal sets of vectors

A set of vectors $\{\mathbf{v}_1, \dots, \mathbf{v}_p\} \in \mathbb{R}^n$ is an *orthogonal set* if $\mathbf{v}_i \cdot \mathbf{v}_j = 0$ for all $i \neq j$.

If in addition $\|\mathbf{v}_i\| = 1$ for $i = 1, \dots, n$, the set is an *orthonormal set*.

Example: The *elementary vectors*

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \\ 0 \end{bmatrix}, \quad \dots \quad , \quad \mathbf{e}_{n-1} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{e}_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$$

in $\mathbb{R}^n$ are an orthonormal set; moreover $\text{Span}\{\mathbf{e}_1, \dots, \mathbf{e}_n\} = \mathbb{R}^n$.

$$\mathbf{U} = [\underset{\sim}{u}_1 \cdots \underset{\sim}{u}_n] \qquad \mathbf{U}^T\mathbf{U} = \begin{bmatrix} \underset{\sim}{u}_1^T\underset{\sim}{u}_1 & \cdots & \underset{\sim}{u}_1^T\underset{\sim}{u}_n \\ \vdots & & \\ \underset{\sim}{u}_n^T\underset{\sim}{u}_1 & & \underset{\sim}{u}_n^T\underset{\sim}{u}_n \end{bmatrix} = \begin{bmatrix} 1 & & \\ & 1 & \\ & & \ddots \\ & & & 1 \end{bmatrix}$$

$\{\underset{\sim}{u}_1, \dots, \underset{\sim}{u}_n\}$ are orthonormal

Result (Orthonormal columns)

An $m \times n$ matrix $\mathbf{U}$ has orthonormal columns if and only if $\mathbf{U}^T\mathbf{U} = \mathbf{I}$.

Prove the result.

Exercise: Check if these matrices have orthonormal columns:

$$\mathbf{C} = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix} \qquad \mathbf{D} = \begin{bmatrix} 1 & 1/\sqrt{2} \\ 0 & 1/\sqrt{2} \end{bmatrix}$$

has orthonormal columns

have unit normal

$$\frac{1}{\sqrt{2}} \cdot \frac{1}{\sqrt{2}} + \left(\frac{1}{\sqrt{2}}\right)\left(-\frac{1}{\sqrt{2}}\right) = 0$$

Theorem (Results for matrices with orthonormal columns)

Let $\mathbf{U}$ *be an* $m \times n$ *matrix with orthonormal columns and let* $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$*. Then*

1. $\|\mathbf{U}\mathbf{x}\| = \|\mathbf{x}\|$
2. $(\mathbf{U}\mathbf{x}) \cdot (\mathbf{U}\mathbf{y}) = \mathbf{x} \cdot \mathbf{y}$ At home
3. $(\mathbf{U}\mathbf{x}) \cdot (\mathbf{U}\mathbf{y}) = 0 \iff \mathbf{x} \cdot \mathbf{y} = 0$

Exercise: Prove the above results.

Why not call this an orthonormal matrix??

$U^T U = I$

Orthogonal matrix

An *orthogonal matrix* is a square invertible matrix $\mathbf{U}$ such that $\mathbf{U}^{-1} = \mathbf{U}^T$.

If $\mathbf{U}$ is $n \times n$ with orthonormal columns, we have $\mathbf{U}^T\mathbf{U} = \mathbf{I}_n$ and $\mathbf{U}\mathbf{U}^T = \mathbf{I}_n$ (since the left inverse is the right inverse).

So square matrices with orthonormal columns are called *orthogonal matrices*.

Orthogonal complement

Let W be a subspace of $\mathbb{R}^n$.

- If $\mathbf{z} \cdot \mathbf{x} = 0$ for all $\mathbf{x} \in W$, we say $\mathbf{z}$ is *orthogonal to* W.
- The *orthogonal complement* of W is the set of all such vectors $\mathbf{z}$.

\perp

Denote the orthogonal complement of a subspace W as $W^\perp$.

Theorem (Results about orthogonal complements)

Let W be a subspace of $\mathbb{R}^n$. Then:

1. *A vector $\mathbf{x}$ is in $W^\perp$ iff $\mathbf{x}$ is orthogonal to every vector in a set that spans W.*
2. *$W^\perp$ is a subspace of $\mathbb{R}^n$.*

Prove the result.

① "$\Rightarrow$" Let $\underset{\sim}{x} \in W^{\perp}$ and let $W = \text{Span}\{\underset{\sim}{u}_1, \dots, \underset{\sim}{u}_p\}$

For any $\underset{\sim}{y} \in W$, we have $\underset{\sim}{x} \cdot \underset{\sim}{y} = 0$.

Choose $\underset{\sim}{y} = \underset{\sim}{u}_1 \in W$. Then $\underset{\sim}{x} \cdot \underset{\sim}{u}_1 = 0$

$\vdots$

$\underset{\sim}{y} = \underset{\sim}{u}_p \in W$. Then $\underset{\sim}{x} \cdot \underset{\sim}{u}_p = 0$

"$\Leftarrow$" Suppose $W = \text{Span}\{\underset{\sim}{u}_1, \dots, \underset{\sim}{u}_p\}$ and $\underset{\sim}{x} \cdot \underset{\sim}{u}_j = 0$, $j = 1, \dots, p$.

Then for any $\underset{\sim}{y} \in W$, we can write

$$\underset{\sim}{y} = c_1 \underset{\sim}{u}_1 + \dots + c_p \underset{\sim}{u}_p$$

Then

$$\begin{aligned} \underset{\sim}{x} \cdot \underset{\sim}{y} &= \underset{\sim}{x} \cdot \left(c_1 \underset{\sim}{u}_1 + \dots + c_p \underset{\sim}{u}_p \right) \\ &= c_1 \, \underset{\sim}{x} \cdot \underset{\sim}{u}_1 + \dots + c_p \, \underset{\sim}{x} \cdot \underset{\sim}{u}_p \\ &= 0 \end{aligned}$$

$\Rightarrow \underset{\sim}{x} \in W^{\perp}$.

② (i) Let W be a subspace of $\mathbb{R}^n$.

A vector $\underset{\sim}{x}$ is in $W^{\perp}$ if $\underset{\sim}{x} \cdot \underset{\sim}{y} = 0 \;\; \forall \, \underset{\sim}{y} \in W$.

Check: $\underset{\sim}{0} \cdot \underset{\sim}{y} = 0 \;\; \forall \, \underset{\sim}{y} \in W$. $\Rightarrow \underset{\sim}{0} \in W^{\perp}$

(ii) Let $\underset{\sim}{x}_1, \underset{\sim}{x}_2 \in W^{\perp}$. Then $\underset{\sim}{x}_1 \cdot \underset{\sim}{y} = 0$ and $\underset{\sim}{x}_2 \cdot \underset{\sim}{y} = 0$ $\forall \, \underset{\sim}{y} \in W$.

$$(\underset{\sim}{x}_1 + \underset{\sim}{x}_2) \cdot \underset{\sim}{y} = \underset{\sim}{x}_1 \cdot \underset{\sim}{y} + \underset{\sim}{x}_2 \cdot \underset{\sim}{y} = 0$$

$$\Rightarrow \underset{\sim}{x}_1 + \underset{\sim}{x}_2 \in W^{\perp}$$

(iii) Let $\underset{\sim}{x} \in W^{\perp}$. Then $\underset{\sim}{x} \cdot \underset{\sim}{y} = 0 \quad \forall \underset{\sim}{y} \in W$.

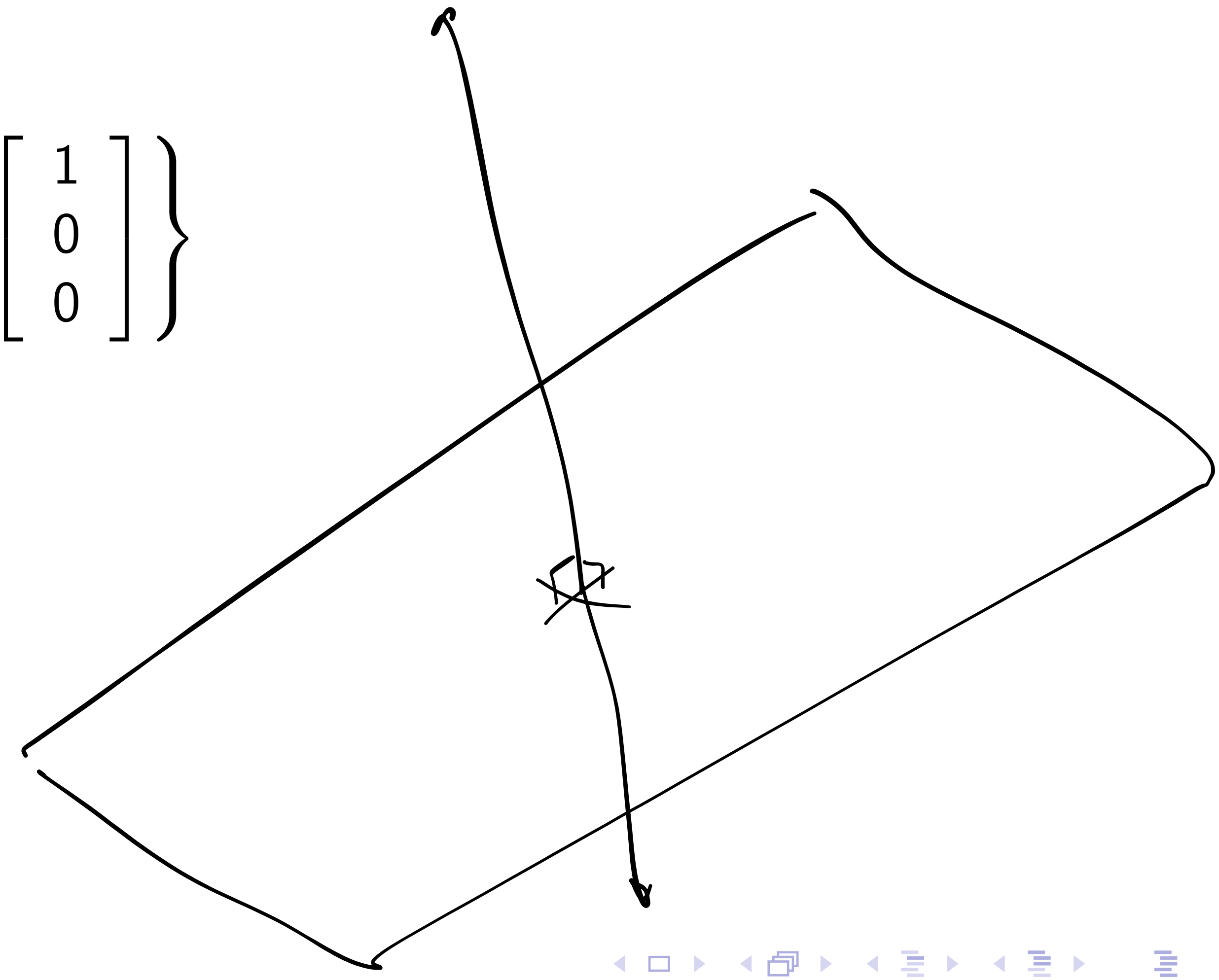
$$\Rightarrow c\underset{\sim}{x} \cdot \underset{\sim}{y} = 0 \quad \forall \underset{\sim}{y} \in W.$$

$$\Rightarrow c\underset{\sim}{x} \in W^{\perp}.$$

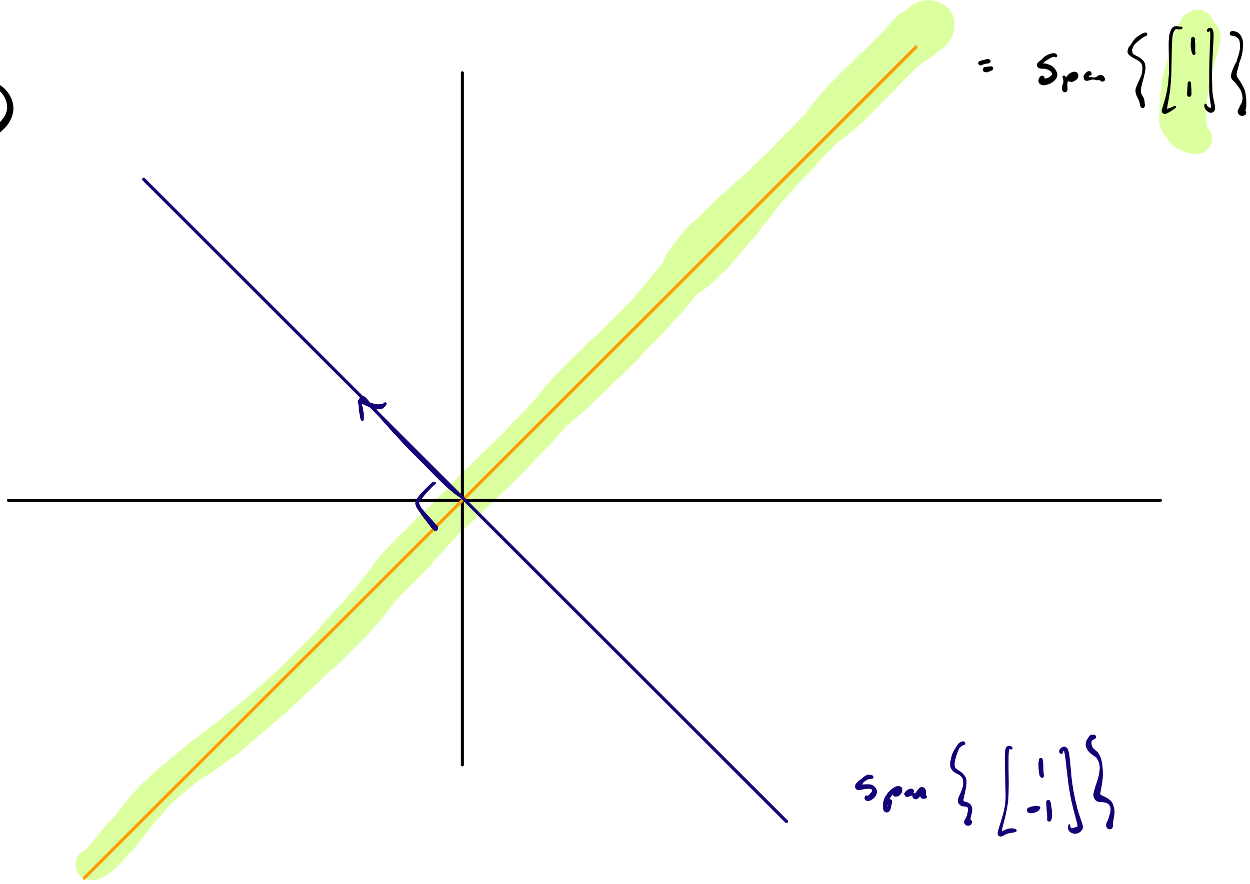
Exercise: Find the orthogonal complements of these subspaces:

1. $W_1 = \text{Span}\left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$
2. $W_2 = \text{Span}\left\{ \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right\}$

Draw pictures.



1


= span $\left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$
span $\left\{ \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right\}$

The next result tells us how to find an orthogonal complement.

Theorem (Orthogonal complement of column and row space)

Let $\mathbf{A}$ *be an* $m \times n$ *matrix. Then* $(\operatorname{Col}\mathbf{A})^{\perp} = \operatorname{Nul}\mathbf{A}^{T}$ *and* $(\operatorname{Row}\mathbf{A})^{\perp} = \operatorname{Nul}\mathbf{A}$.

Prove the result.

$\underset{m\times n}{A}$ $\qquad (\mathrm{Col}\,A)^\perp = \mathrm{Nul}\,A^T.$

"$\subseteq$" $\qquad$ Let $\underset{\sim}{x} \in (\mathrm{Col}\,A)^\perp$.

Means, for any $\underset{\sim}{y} \in \mathrm{Col}\,A$, $\underset{\sim}{x}\cdot\underset{\sim}{y} = 0$.

Write $\underset{\sim}{y} = A\underset{\sim}{b}$ for some $\underset{\sim}{b}$.

$$\mathrm{Col}\,A = \{\underset{\sim}{y} : \underset{\sim}{y} = A\underset{\sim}{b} \text{ for some } \underset{\sim}{b} \in \mathbb{R}^n\}$$

$\Rightarrow \quad \underset{\sim}{x}\cdot(A\underset{\sim}{b}) = 0 \quad$ for all $\underset{\sim}{b} \in \mathbb{R}^n$.

$\Rightarrow \quad \underset{\sim}{x}^T A\underset{\sim}{b} = 0 \quad$ for all $\underset{\sim}{b} \in \mathbb{R}^n$.

$\Rightarrow \quad \underset{\sim}{x}^T A = \underset{\sim}{0}^T$

$\Rightarrow \quad A^T\underset{\sim}{x} = \underset{\sim}{0}$

$\Rightarrow \quad \underset{\sim}{x} \in \mathrm{Nul}\,A^T.$

$$\boxed{\begin{array}{l} A\underset{\sim}{x} + \underset{\sim}{b} = 0 \\ \quad \forall\, \underset{\sim}{x} \in \mathbb{R}^n \\ \Rightarrow A = 0 \\ \quad \underset{\sim}{b} = 0 \end{array}}$$

"$\supseteq$" $\qquad$ Let $\underset{\sim}{x} \in \mathrm{Nul}\,A^T$. $\quad$ Then $A^T\underset{\sim}{x} = \underset{\sim}{0}$.

Let $\underset{\sim}{y} \in \mathrm{Col}\,A$. Want to show $\underset{\sim}{x}\cdot\underset{\sim}{y} = 0$.

Can write $\underset{\sim}{y} = A\underset{\sim}{b}$ for some $\underset{\sim}{b}$.

Then $\quad \underset{\sim}{x}\cdot\underset{\sim}{y} = \underset{\sim}{x}\cdot A\underset{\sim}{b} = \underset{\sim}{x}^T A\underset{\sim}{b} = \underset{\sim}{b}^T A^T\underset{\sim}{x} = 0$

$\Rightarrow \quad \underset{\sim}{x} \in (\text{Col } A)^{\perp}.$

Exercise: Find the orthogonal complement of $\text{Span}\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \right\}$

Let $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 0 \end{bmatrix}$. Then Col A.

Orth. comple. is $\text{Nul } A^T$. $\quad \text{Nul } A^T = \{ \underset{\sim}{x} : \underline{A\underset{\sim}{x} = \underset{\sim}{0}} \}$

$A^T = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}$

$$\begin{bmatrix} A^T & \underset{\sim}{0} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 \end{bmatrix}$$

$$\sim \begin{bmatrix} 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

$$x_1 + x_2 = 0$$

$$x_3 = 0$$

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_2 \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$$

$$\text{Nul } A^T = \text{Span} \left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \right\}$$

$$A^T \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} = \underset{\sim}{0}$$

2 $W = V$ means $W \subset V \Rightarrow V^\perp \subset W^\perp$ and $V \subset W \Rightarrow W^\perp \subset V^\perp$ means $W^\perp = V^\perp$.

Theorem (Chopping result. Cf. Monahan (2008) Result A.6)

1 *For two vector spaces* W *and* V, $W \subset V$ *implies* $V^\perp \subset W^\perp$.

2 *As a consequence, we have additionally* $\boxed{W = V \iff W^\perp = V^\perp}$.

Prove the result.

1 Let $\underset{\sim}{x} \in V^\perp$. This means for any $\underset{\sim}{y} \in V$ I have $\underset{\sim}{x} \cdot \underset{\sim}{y} = 0$.

Take any $\underset{\sim}{w} \in W$. Then $\underset{\sim}{w} \in V$ (since $W \subset V$).

So $\underset{\sim}{x} \cdot \underset{\sim}{w} = 0$, giving $\underset{\sim}{x} \in W^\perp$.

So we've show $V^\perp \subset W^\perp$.

Theorem (Lemma 2.1 and Results 2.2 and 2.4 of Monahan)

For an $n \times p$ matrix $\mathbf{X}$, we have

1. $\text{Nul}\, \mathbf{X}^T\mathbf{X} = \text{Nul}\, \mathbf{X}$
2. $\text{Col}\, \mathbf{X}^T\mathbf{X} = \text{Col}\, \mathbf{X}^T$
3. $\mathbf{X}^T\mathbf{XA} = \mathbf{X}^T\mathbf{XB} \iff \mathbf{XA} = \mathbf{XB}$.

Prove the results.

1. $\text{Nul}\, X^TX = \text{Nul}\, X$

" $\subset$ " Let $\underset{\sim}{x} \in \text{Nul}\, X^TX$. $\underbrace{\text{Then } X^TX\underset{\sim}{x} = \underset{\sim}{0}}$

$\Rightarrow \underset{\sim}{x}^T X^T X \underset{\sim}{x} = 0$

$\Rightarrow \|X\underset{\sim}{x}\|^2 = 0 \iff X\underset{\sim}{x} = \underset{\sim}{0}. \Rightarrow \underset{\sim}{x} \in \text{Nul}\, X$

"$\supseteq$" Let $\underset{\sim}{x} \in \text{Nul } X$. Then $X\underset{\sim}{x} = \underset{\sim}{0}$.

$$\Rightarrow \quad X^T X \underset{\sim}{x} = \underset{\sim}{0}$$

$$\Rightarrow \quad \underset{\sim}{x} \in \text{Nul } X^T X. \quad \square$$

② $\text{Col } X^T X = \text{Col } X^T \iff \underbrace{(\text{Col } X^T X)^\perp = (\text{Col } X^T)^\perp}_{\text{Just need to show this}}$

$$(\text{Col } X^T X)^\perp = \text{Nul } (X^T X)^T = \text{Nul } X^T X$$

$$(\text{Col } X^T)^\perp = \text{Nul } (X^T)^T = \text{Nul } X$$

$$\boxed{\text{Col } A = \text{Nul } A^T}$$

Since $\text{Nul } X^T X = \text{Nul } X$, we have $\text{Col } X^T X = \text{Col } X^T$.

③ $X^T X A = X^T X B \iff XA = XB$ "Cool Result"

"$\Leftarrow$" Let $XA = XB$. Then premultiply both sides by X^T to get

$$X^T X A = X^T X B.$$

"$\Rightarrow$" Let $X^T X A = X^T X B$.

Write $X^T X (A - B) = 0$.

$\Rightarrow$ Columns of $A - B$ are in $\text{Nul } X^T X$.

$\Rightarrow$ Columns of $A - B$ are in $\text{Nul } X$ (since $\text{Nul } X^T X = \text{Nul } X$)

$\Rightarrow \quad X(A-B)=0 \quad \Rightarrow \quad XA = XB.$

Theorem (An orthogonal set of nonzero vectors is a basis)

If $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ *is an orthogonal set of nonzero vectors in* $\mathbb{R}^n$*, then it is linearly independent and therefore a basis for* Span$\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$.

An *orthogonal basis* is a basis which is an orthogonal set.

Prove the result.

Want to show $\quad c_1 \underset{\sim}{u}_1 + \cdots + c_p \underset{\sim}{u}_p = \underset{\sim}{0} \iff c_1 = \cdots = c_p = 0.$

Let $\quad \underset{\sim}{x} = c_1 \underset{\sim}{u}_1 + \cdots + c_p \underset{\sim}{u}_p = \underset{\sim}{0}.$

$\Rightarrow \quad 0 = \|\underset{\sim}{x}\|^2 = \underset{\sim}{x}^T \underset{\sim}{x} = (c_1 \underset{\sim}{u}_1 + \cdots + c_p \underset{\sim}{u}_p)^T (c_1 \underset{\sim}{u}_1 + \cdots + c_p \underset{\sim}{u}_p)$

$$= c_1^2 \underset{\sim}{u}_1^T \underset{\sim}{u}_1 + c_1 \underset{\sim}{u}_1^T c_2 \underset{\sim}{u}_2 + \cdots$$

$$+ c_2^2 \underset{\sim}{u}_2^T \underset{\sim}{u}_2 \quad \cdots$$

$$\vdots$$

$$+ c_p^2 \underset{\sim}{u}_p^T \underset{\sim}{u}_p$$

$$= c_1^2 \|\underset{\sim}{u}_1\|^2 + \cdots + c_p^2 \|\underset{\sim}{u}_p\|^2 \iff c_1 = \cdots = c_p = 0.$$

Example: Here are two bases for the same space—one orthogonal, one not:

$$\text{Span}\left\{\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}\right\} = \text{Span}\left\{\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}\right\}$$

How to find a vector's "coordinates" with respect to an orthogonal basis:

Result (Find a vector's coordinates wrt an orthogonal basis)

Let $\{\mathbf{u}_1, \dots, \mathbf{u}_p\}$ *be an orth. basis for a subspace* W *of* $\mathbb{R}^n$*. Then for each* $\mathbf{y} \in W$*,*

$$\mathbf{y} = c_1\mathbf{u}_1 + \cdots + c_p\mathbf{u}_p, \quad \text{where } c_j = \frac{\mathbf{y} \cdot \mathbf{u}_j}{\mathbf{u}_j \cdot \mathbf{u}_j}, \quad j = 1, \dots, p.$$

Prove the result.

$\underset{\sim}{y} = c_1 \underset{\sim}{u}_1 + \cdots + c_p \underset{\sim}{u}_p$.

Find c_1: write

$\underset{\sim}{u}_1 \cdot \underset{\sim}{y} = \underset{\sim}{u}_1 \cdot (c_1 \underset{\sim}{u}_1 + \cdots + c_p \underset{\sim}{u}_p) = c_1 \underset{\sim}{u}_1 \cdot \underset{\sim}{u}_1$

$\iff c_1 = \dfrac{\underset{\sim}{u}_1 \cdot \underset{\sim}{y}}{\underset{\sim}{u}_1 \cdot \underset{\sim}{u}_1}$.

Do this for all c_j, $j = 1, \dots, p$

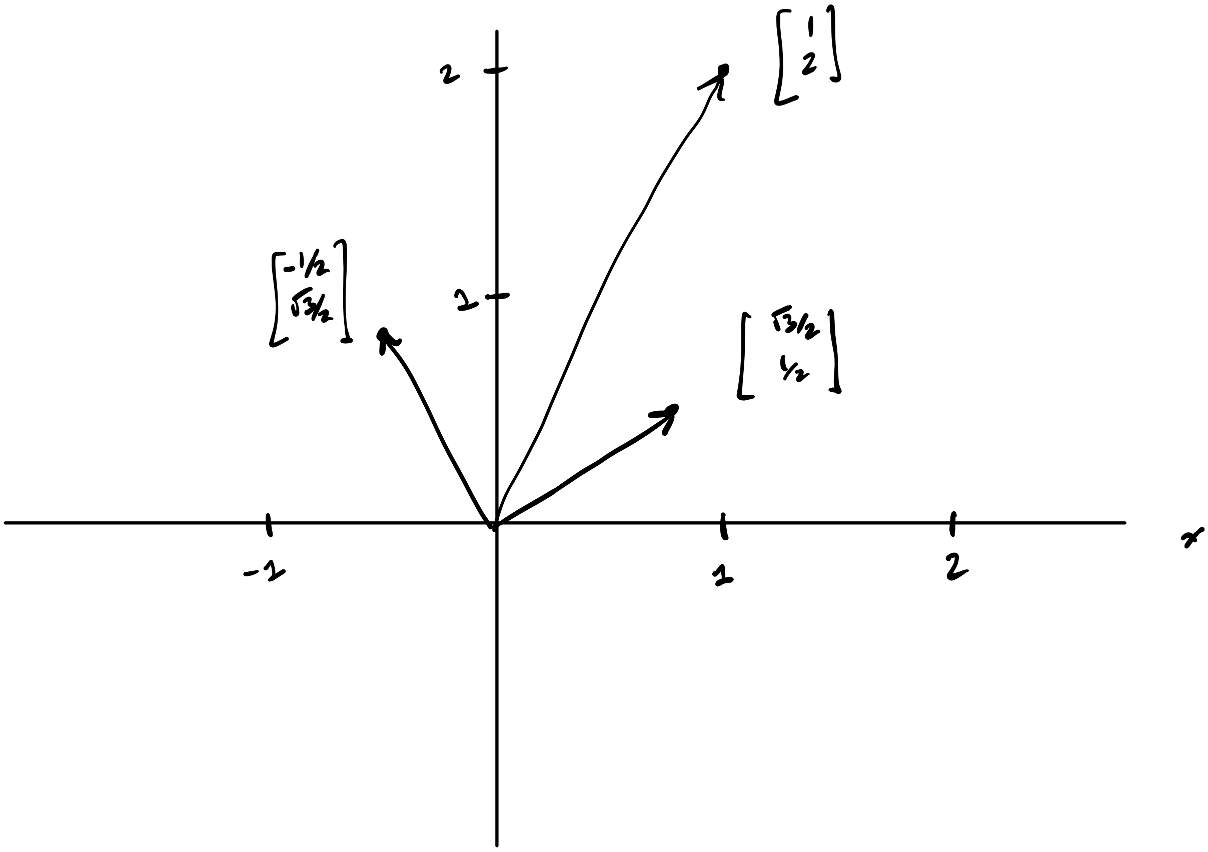
Exercise: Find the coefficients to construct $\mathbf{y} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ from the orthogonal basis

$$\left\{ \begin{bmatrix} \sqrt{3}/2 \\ 1/2 \end{bmatrix}, \begin{bmatrix} -1/2 \\ \sqrt{3}/2 \end{bmatrix} \right\}$$

$$\underset{\sim}{y} = c_1 \underset{\sim}{u}_1 + c_2 \underset{\sim}{u}_2 \quad , \quad c_1 = \frac{\underset{\sim}{u}_1 \cdot \underset{\sim}{y}}{\underset{\sim}{u}_1 \cdot \underset{\sim}{u}_1} \qquad c_2 = \frac{\underset{\sim}{u}_2 \cdot \underset{\sim}{y}}{\underset{\sim}{u}_2 \cdot \underset{\sim}{u}_2}$$

$$c_1 = \begin{pmatrix} \sqrt{3}/2 \\ 1/2 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \sqrt{3}/2 + 1 = \frac{2+\sqrt{3}}{2}$$

$$c_2 = \begin{bmatrix} -1/2 \\ \sqrt{3}/2 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = -\frac{1}{2} + \sqrt{3}$$



$$\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \frac{2+\sqrt{3}}{2}\begin{pmatrix} \sqrt{3}/2 \\ 1/2 \end{pmatrix} + \left(-\frac{1}{2} + \sqrt{3}\right)\begin{pmatrix} -\frac{1}{2} \\ \sqrt{3}/2 \end{pmatrix}$$

$$= \begin{pmatrix} \frac{(2+\sqrt{3})\sqrt{3}}{4} + \left(\frac{1}{4} - \sqrt{3}/2\right) \\ \frac{2+\sqrt{3}}{4} + \frac{3}{2} - \frac{\sqrt{3}}{4} \end{pmatrix}$$

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

Theorem (Orthogonal decomposition theorem)

Let W be a subspace of $\mathbb{R}^n$. Then we can decompose any $\mathbf{y} \in \mathbb{R}^n$ uniquely as

$$\mathbf{y} = \underbrace{\hat{\mathbf{y}}}_{\text{fitted values}} + \overbrace{\hat{\mathbf{e}}}^{\text{residuals}}, \quad \text{where} \quad \hat{\mathbf{y}} \in W \quad \text{and} \quad \hat{\mathbf{e}} \in W^{\perp}.$$

Moreover, for any orthogonal basis $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ for W, we have

$$\hat{\mathbf{y}} = \frac{\mathbf{y} \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1 + \cdots + \frac{\mathbf{y} \cdot \mathbf{u}_p}{\mathbf{u}_p \cdot \mathbf{u}_p}\mathbf{u}_p \quad \text{and} \quad \hat{\mathbf{e}} = \mathbf{y} - \hat{\mathbf{y}}.$$

The vector $\hat{\mathbf{y}}$ is called the *orthogonal projection of $\mathbf{y}$ onto W*, denoted by $\text{proj}_W \mathbf{y}$.

Note that if $\mathbf{y} \in W$ then $\text{proj}_W \mathbf{y} = \mathbf{y}$.

Prove the theorem.

① Prove that $\underset{\sim}{\hat{y}} + \underset{\sim}{\hat{e}}$ exists

② Prove that it's unique.

(1) Let $\{\underset{\sim}{u}_1, \dots, \underset{\sim}{u}_p\}$ be an orth. basis for W.

Let $\hat{\underset{\sim}{y}} = \frac{\underset{\sim}{y} \cdot \underset{\sim}{u}_1}{\underset{\sim}{u}_1 \cdot \underset{\sim}{u}_1} \underset{\sim}{u}_1 + \dots + \frac{\underset{\sim}{y} \cdot \underset{\sim}{u}_p}{\underset{\sim}{u}_p \cdot \underset{\sim}{u}_p} \underset{\sim}{u}_p \in W$.

and let $\hat{\underset{\sim}{e}} = \underset{\sim}{y} - \hat{\underset{\sim}{y}}$.

Then $\underset{\sim}{y} = \hat{y} + \hat{\underset{\sim}{e}} = \hat{\underset{\sim}{y}} + (\underset{\sim}{y} - \hat{\underset{\sim}{y}})$.

Need to show $\hat{\underset{\sim}{y}} \in W$ and $\hat{\underset{\sim}{e}} \in W^\perp$.

To show $\hat{\underset{\sim}{e}} \in W^\perp$, must show $\hat{\underset{\sim}{e}} \cdot \underset{\sim}{u}_j = 0$ for all $j = 1, \dots, p$.

Take
$$\begin{aligned}
\underset{\sim}{u}_j \cdot \hat{\underset{\sim}{e}} &= \underset{\sim}{u}_j \cdot (\underset{\sim}{y} - \hat{\underset{\sim}{y}}) \\
&= \underset{\sim}{u}_j \cdot \left(y - \left(\frac{\underset{\sim}{y} \cdot \underset{\sim}{u}_1}{\underset{\sim}{u}_1 \cdot \underset{\sim}{u}_1} \underset{\sim}{u}_1 + \dots + \frac{\underset{\sim}{y} \cdot \underset{\sim}{u}_p}{\underset{\sim}{u}_p \cdot \underset{\sim}{u}_p} \underset{\sim}{u}_p \right)\right) \\
&= \underset{\sim}{u}_j \cdot \underset{\sim}{y} - \frac{\underset{\sim}{y} \cdot \underset{\sim}{u}_j}{\cancel{\underset{\sim}{u}_j \cdot \underset{\sim}{u}_j}} \cancel{\underset{\sim}{u}_j \cdot \underset{\sim}{u}_j} \\
&= \underset{\sim}{u}_j \cdot \underset{\sim}{y} - \underset{\sim}{y} \cdot \underset{\sim}{u}_j \\
&= 0.
\end{aligned}$$

Can do for all $j = 1, \dots, p$, so $\hat{\underset{\sim}{e}} \in W^\perp$.

(2) Suppose
$$\underset{\sim}{y} = \underbrace{\overset{\in W}{\hat{\underset{\sim}{y}}_1} + \overset{\in W^\perp}{\hat{\underset{\sim}{e}}_1} = \overset{\in W}{\hat{\underset{\sim}{y}}_2} + \overset{\in W^\perp}{\hat{\underset{\sim}{e}}_2}}$$

$$\iff \underbrace{\hat{\underset{\sim}{y}}_1 - \hat{\underset{\sim}{y}}_2}_{\in W} = \underbrace{\hat{\underset{\sim}{e}}_2 - \hat{\underset{\sim}{e}}_1}_{\in W^\perp} = \underset{\sim}{0} \quad \text{because } W \cap W^\perp = \{\underset{\sim}{0}\}$$

$W \cap W^\perp = \{\underset{\sim}{0}\}$

Theorem (Best approximation theorem)

Let W be a subspace of $\mathbb{R}^n$, $\mathbf{y}$ any vector in $\mathbb{R}^n$, and $\hat{\mathbf{y}} = \text{proj}_W\,\mathbf{y}$. Then

$$\underbrace{\|\mathbf{y} - \hat{\mathbf{y}}\|}_{\text{distance between } y \text{ and } \hat{y}} < \|\mathbf{y} - \mathbf{v}\| \quad \textit{for all } \mathbf{v} \in W \textit{ distinct from } \hat{\mathbf{y}}.$$

Prove the result.

$$\|y - v\|^2 = \|y - \hat{y} + \hat{y} - v\|^2$$

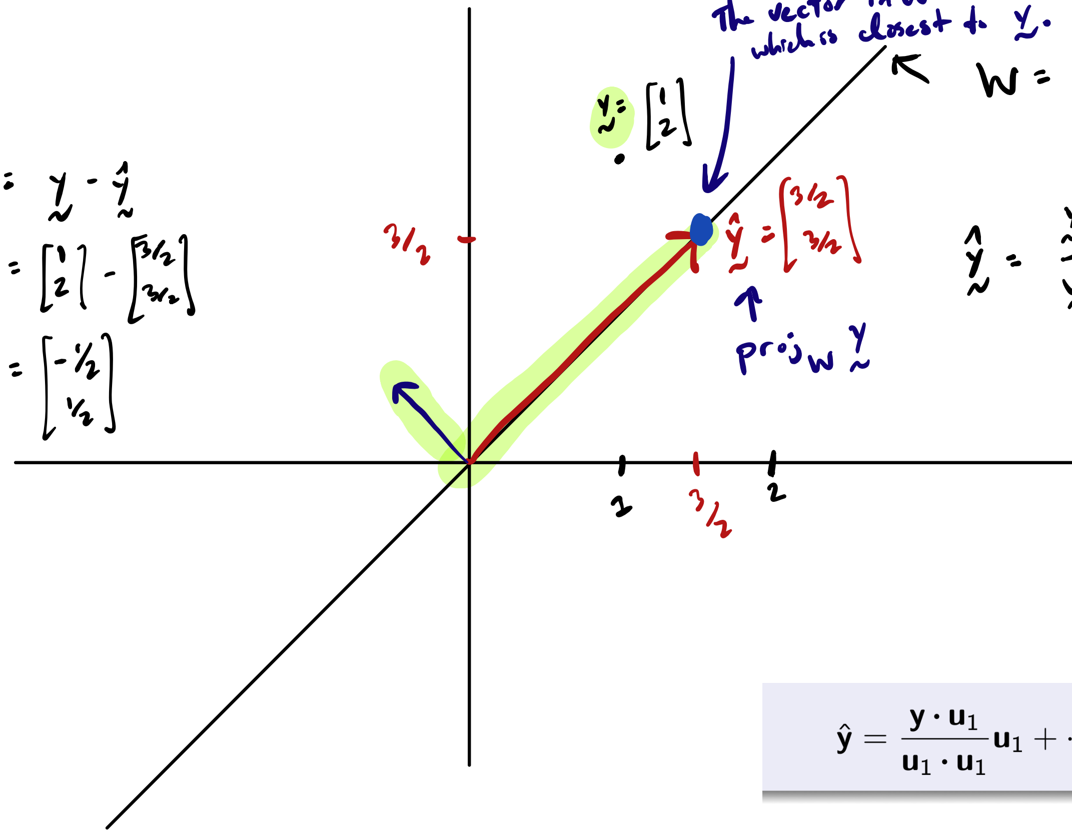
$$= ((y - \hat{y}) + (\hat{y} - v))^T((y - \hat{y}) + (\hat{y} - v))$$

$$= \underline{(y - \hat{y})^T(y - \hat{y})} + \underbrace{2(\overbrace{y - \hat{y}}^{\in W^\perp})^T(\overbrace{\hat{y} - v}^{\in W})}_{=0} + \underline{(\hat{y} - v)^T(\hat{y} - v)}$$

$$= \|y - \hat{y}\|^2 + \underline{\|\hat{y} - v\|^2}$$

$$\Leftarrow \quad \|\underset{\sim}{y} - \hat{\underset{\sim}{y}}\|^2 = \|\underset{\sim}{y} - \underset{\sim}{v}\|^2 - \underbrace{\|\hat{\underset{\sim}{y}} - \underset{\sim}{v}\|^2}_{>0} < \|\underset{\sim}{y} - \underset{\sim}{v}\|^2$$

Exercise: Let $\mathbf{v} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\mathbf{y} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$. Give the orthogonal projection $\hat{\mathbf{y}}$ of $\mathbf{y}$ on $\text{Span}\{\mathbf{v}\}$. Draw pictures! Check orthogonality of $\mathbf{y} - \hat{\mathbf{y}}$ and $\mathbf{y}$.



$$\hat{e} = y - \hat{y} = \begin{bmatrix}1\\2\end{bmatrix} - \begin{bmatrix}3/2\\3/2\end{bmatrix} = \begin{bmatrix}-1/2\\1/2\end{bmatrix}$$

$$W = \text{span}\left\{ \underbrace{\begin{bmatrix}1\\1\end{bmatrix}}_{v} \right\}$$

$$\hat{y} = \frac{y \cdot v}{v \cdot v} v = \frac{\begin{pmatrix}1\\2\end{pmatrix}^T \begin{pmatrix}1\\1\end{pmatrix}}{\begin{pmatrix}1\\1\end{pmatrix}^T \begin{pmatrix}1\\1\end{pmatrix}} \begin{pmatrix}1\\1\end{pmatrix} = \frac{3}{2}\begin{pmatrix}1\\1\end{pmatrix} = \begin{pmatrix}3/2\\3/2\end{pmatrix}$$

$$\hat{\mathbf{y}} = \frac{\mathbf{y}\cdot\mathbf{u}_1}{\mathbf{u}_1\cdot\mathbf{u}_1}\mathbf{u}_1 + \cdots + \frac{\mathbf{y}\cdot\mathbf{u}_p}{\mathbf{u}_p\cdot\mathbf{u}_p}\mathbf{u}_p \quad a$$

The Gram-Schmidt process turns a basis into an orthogonal basis.

The Gram–Schmidt Process

Given a basis $\{\mathbf{x}_1, \dots, \mathbf{x}_p\}$ for a subspace W of $\mathbb{R}^n$, define

$$\mathbf{v}_1 = \mathbf{x}_1$$

$$\mathbf{v}_2 = \mathbf{x}_2 - \frac{\mathbf{x}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 \leftarrow \overbrace{x_2 - \text{proj}_{\text{span}\{\underset{\sim}{v_1}\}} x_2}^{\text{orth to } \underset{\sim}{v_1}}$$

$$\mathbf{v}_3 = \mathbf{x}_3 - \frac{\mathbf{x}_3 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 - \frac{\mathbf{x}_3 \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2}\mathbf{v}_2 \leftarrow x_3 - \text{proj}_{\text{span}\{\underset{\sim}{v_1}, \underset{\sim}{v_2}\}} x_3$$

$$\vdots$$

$$\mathbf{v}_p = \mathbf{x}_p - \frac{\mathbf{x}_p \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 - \frac{\mathbf{x}_p \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2}\mathbf{v}_2 - \cdots - \frac{\mathbf{x}_p \cdot \mathbf{v}_{p-1}}{\mathbf{v}_{p-1} \cdot \mathbf{v}_{p-1}}\mathbf{v}_{p-1}$$

Then $\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ is an orthogonal basis for W. In addition

$$\text{Span}\{\mathbf{v}_1, \dots, \mathbf{v}_k\} = \text{Span}\{\mathbf{x}_1, \dots, \mathbf{x}_k\} \quad \text{for } 1 \le k \le p \tag{1}$$

Prove the result.

Exercise: Use the Gram-Schmidt process to orthogonalize the basis

$$\left\{ \overset{\underset{\sim}{x}_2}{\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}}, \overset{\underset{\sim}{x}_1}{\begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}}, \overset{\underset{\sim}{x}_3}{\begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}} \right\}.$$

$$\underset{\sim}{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

$$\underset{\sim}{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} - \frac{\begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \end{pmatrix}^T \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}}{\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} - \frac{3}{4} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -3/4 \\ 1/4 \\ 1/4 \\ 1/4 \end{bmatrix}.$$

$$\underset{\sim}{v}_3 = \cdots = \begin{bmatrix} 0 \\ -2/3 \\ 1/3 \\ 1/3 \end{bmatrix}$$

Lay, D. C. (2003). *Linear algebra and its applications. Third edition.* Pearson Education.

Monahan, J. F. (2008). *A primer on linear models*. CRC Press.