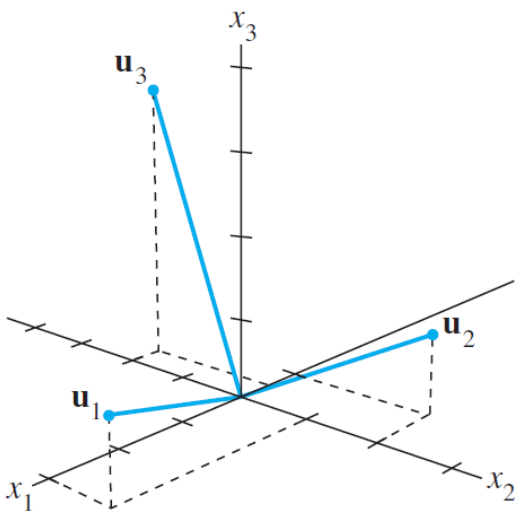
Lecture 10: Vector Algebra: Orthogonal Basis

- Orthogonal Basis of a subspace
- Computing an orthogonal basis for a subspace using Gram-Schmidt Orthogonalization Process

Orthogonal Set

- Any set of vectors that are mutually orthogonal, is a an orthogonal set.



EXAMPLE 1 Show that $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is an orthogonal set, where

$$\mathbf{u}_1 = \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}, \quad \mathbf{u}_3 = \begin{bmatrix} -1/2 \\ -2 \\ 7/2 \end{bmatrix}$$

SOLUTION Consider the three possible pairs of distinct vectors, namely, $\{\mathbf{u}_1, \mathbf{u}_2\}$, $\{\mathbf{u}_1, \mathbf{u}_3\}$, and $\{\mathbf{u}_2, \mathbf{u}_3\}$.

$$\mathbf{u}_1 \cdot \mathbf{u}_2 = 3(-1) + 1(2) + 1(1) = 0$$
$$\mathbf{u}_1 \cdot \mathbf{u}_3 = 3\left(-\tfrac{1}{2}\right) + 1(-2) + 1\left(\tfrac{7}{2}\right) = 0$$
$$\mathbf{u}_2 \cdot \mathbf{u}_3 = -1\left(-\tfrac{1}{2}\right) + 2(-2) + 1\left(\tfrac{7}{2}\right) = 0$$

Each pair of distinct vectors is orthogonal, and so $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is an orthogonal set.

Orthonormal Set

- Any set of unit vectors that are mutually orthogonal, is a an *orthonormal* set.
- In other words, any orthogonal set is an orthonormal set if all the vectors in the set are unit vectors.
- Example: $\{\widehat{\boldsymbol{u}_1}, \widehat{\boldsymbol{u}_2}, \widehat{\boldsymbol{u}_3}\}$ is an orthonormal set, where,

$$\widehat{\boldsymbol{u}_1} = \begin{bmatrix} \frac{3}{\sqrt{11}} \\ \frac{1}{\sqrt{11}} \\ \frac{1}{\sqrt{11}} \end{bmatrix}, \widehat{\boldsymbol{u}_2} = \begin{bmatrix} -\frac{1}{\sqrt{6}} \\ \frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \end{bmatrix}, \widehat{\boldsymbol{u}_3} = \begin{bmatrix} -\frac{1}{\sqrt{66}} \\ -\frac{4}{\sqrt{66}} \\ \frac{7}{\sqrt{66}} \end{bmatrix}$$

An orthogonal set is Linearly Independent

If $S = \{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ is an orthogonal set of nonzero vectors in $\mathbb{R}^n$, then S is linearly independent and hence is a basis for the subspace spanned by S.

PROOF If $\mathbf{0} = c_1\mathbf{u}_1 + \cdots + c_p\mathbf{u}_p$ for some scalars $c_1, \ldots, c_p$, then

$$\begin{aligned} 0 = \mathbf{0}\cdot\mathbf{u}_1 &= (c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_p\mathbf{u}_p)\cdot\mathbf{u}_1 \\ &= (c_1\mathbf{u}_1)\cdot\mathbf{u}_1 + (c_2\mathbf{u}_2)\cdot\mathbf{u}_1 + \cdots + (c_p\mathbf{u}_p)\cdot\mathbf{u}_1 \\ &= c_1(\mathbf{u}_1\cdot\mathbf{u}_1) + c_2(\mathbf{u}_2\cdot\mathbf{u}_1) + \cdots + c_p(\mathbf{u}_p\cdot\mathbf{u}_1) \\ &= c_1(\mathbf{u}_1\cdot\mathbf{u}_1) \end{aligned}$$

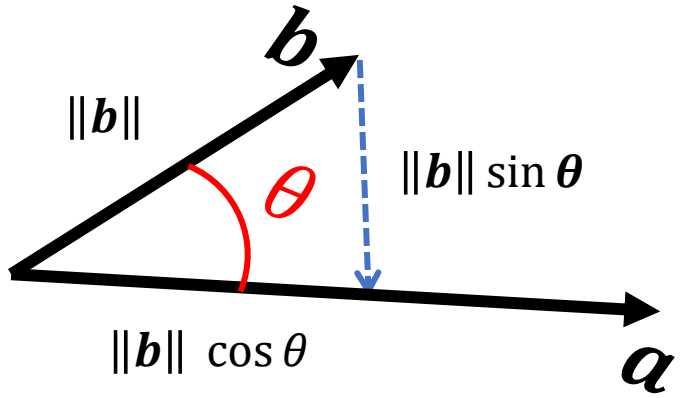
because $\mathbf{u}_1$ is orthogonal to $\mathbf{u}_2, \ldots, \mathbf{u}_p$. Since $\mathbf{u}_1$ is nonzero, $\mathbf{u}_1\cdot\mathbf{u}_1$ is not zero and so $c_1 = 0$. Similarly, $c_2, \ldots, c_p$ must be zero. Thus S is linearly independent. ■

Projection of vector $\boldsymbol{b}$ on vector $\boldsymbol{a}$

- $\boldsymbol{a} \cdot \boldsymbol{b} = \|\boldsymbol{a}\| \, \|\boldsymbol{b}\| \, \cos\theta$
- Vector $\boldsymbol{c}$ is the image/perpendicular projection of $\boldsymbol{b}$ on $\boldsymbol{a}$
- Direction of $\boldsymbol{c}$ is the same as $\boldsymbol{a}$
- Magnitude of $\boldsymbol{c}$ is $\|\boldsymbol{c}\| = \|\boldsymbol{b}\| \, \cos\theta = \frac{\boldsymbol{a}\cdot\boldsymbol{b}}{\|\boldsymbol{a}\|}$

$$\|\boldsymbol{c}\| = \hat{\boldsymbol{a}} \cdot \boldsymbol{b}$$

- If $\hat{\boldsymbol{a}}$ is the unit vector of $\boldsymbol{a}$, then
- vector $\boldsymbol{c} = \frac{\boldsymbol{a}\cdot\boldsymbol{b}}{\|\boldsymbol{a}\|}\hat{\boldsymbol{a}} = \frac{\boldsymbol{a}\cdot\boldsymbol{b}}{\|\boldsymbol{a}\|}\frac{\boldsymbol{a}}{\|\boldsymbol{a}\|} = \frac{\boldsymbol{a}\cdot\boldsymbol{b}}{\boldsymbol{a}\cdot\boldsymbol{a}}\boldsymbol{a}$



Orthogonal Basis

- An orthogonal basis for a subspace W of R^n is a basis for W that is also an orthogonal set.

- Example:
  $\left\{\begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\end{bmatrix}, \begin{bmatrix}0\\0\\1\end{bmatrix}\right\}$ is basically the x, y, and z axis. It is an orthogonal basis in $\mathbb{R}^3$, and it spans the whole $\mathbb{R}^3$ space. It is also an orthogonal set.

Orthogonal Basis

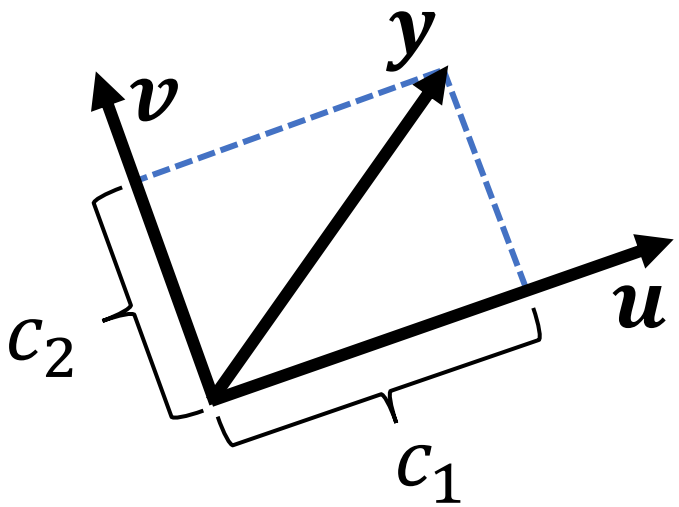
- We know that given a basis of a subspace, any vector in that subspace will be a linear combination of the basis vectors.
- For example, if $\boldsymbol{u}\ and\ \boldsymbol{v}$ are linearly independent and form the basis for a subspace S, then any vector $\boldsymbol{y}$ in S can be expressed as:

$$\boldsymbol{y} = c_1\boldsymbol{u} + c_2\boldsymbol{v}$$

- But computing c_1 and c_2 is not straight forward.

- On the other hand, if $\boldsymbol{u}\ and\ \boldsymbol{v}$ form an orthogonal basis, then

$$c_1 = \frac{\boldsymbol{y}\cdot\boldsymbol{u}}{\|\boldsymbol{u}\|} = \frac{\boldsymbol{y}\cdot\boldsymbol{u}}{\boldsymbol{u}\cdot\boldsymbol{u}} \text{ and } c_2 = \frac{\boldsymbol{y}\cdot\boldsymbol{v}}{\|\boldsymbol{v}\|} = \frac{\boldsymbol{y}\cdot\boldsymbol{v}}{\boldsymbol{v}\cdot\boldsymbol{v}}$$

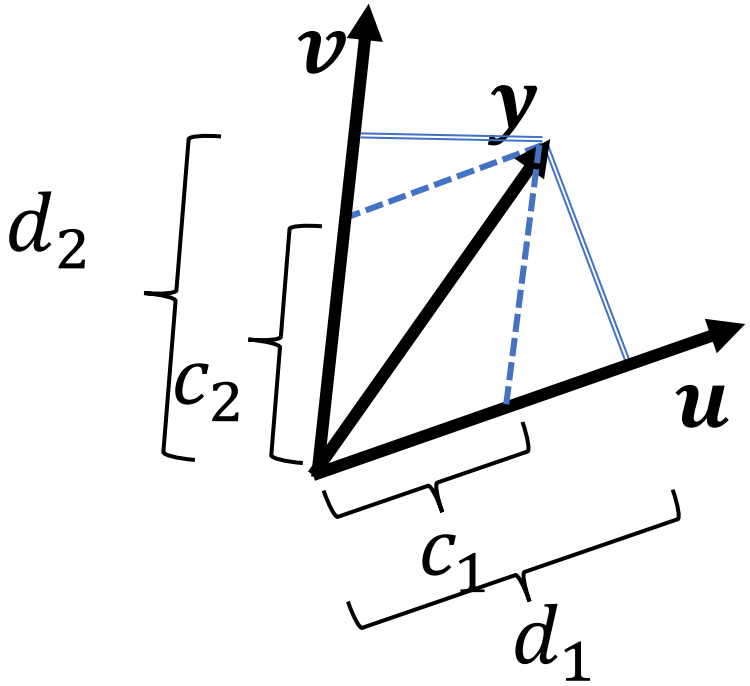


Does not work if it is not Orthogonal basis

- $\boldsymbol{y} = c_1\boldsymbol{u} + c_2\boldsymbol{v}$
- But computing c_1 and c_2 is not straight forward (yet).

What is computed is,

- $d_1 = \frac{\boldsymbol{y}\cdot\boldsymbol{u}}{\|\boldsymbol{u}\|} = \frac{\boldsymbol{y}\cdot\boldsymbol{u}}{\boldsymbol{u}\cdot\boldsymbol{u}}$
- $d_2 = \frac{\boldsymbol{y}\cdot\boldsymbol{v}}{\|\boldsymbol{v}\|} = \frac{\boldsymbol{y}\cdot\boldsymbol{v}}{\boldsymbol{v}\cdot\boldsymbol{v}}$



Orthogonal Decomposition Theorem

Given a vector u in R^n, consider the problem of decomposing a vector y in R^n into two components:

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

where $\hat{y}$ is in $span\{u\}$ and z is orthogonal to u. $\hat{y}$ is called the *orthogonal projection* of y onto u.

Orthogonal Decomposition Theorem

- Decompose $y = \hat{y} + z$
- Let $\hat{y} = \alpha u$ for some scalar α. Then

$$z = y - \hat{y} = y - \alpha u$$

- Since $z \cdot u = 0$, then

$$u^T(y - \alpha u) = 0$$

  so we have $u^T y = \alpha u^T u$, and

$$\alpha = \frac{y^T u}{u^T u}$$

- $\mathrm{Proj}_u(y) = \hat{y} = \alpha u = \frac{y^T u}{u^T u} u$

Orthogonal Decomposition Theorem

Let $W = span\{u_1, \ldots, u_p\}$ is a subspace of R^n, where $\{u_1, \ldots, u_p\}$ is an orthogonal set. Decompose y into two components:

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

where $\hat{y}$ is a vector in W and z is orthogonal to W. $\hat{y}$ is called the orthogonal projection of y onto W.

- Since $\hat{y}$ is in W, write

$$\hat{y} = c_1 u_1 + c_2 u_2 + \ldots + c_p u_p$$

- $z = y - \hat{y}$ is orthogonal to W, implying that $u_i \cdot z = 0$ for every i.
- From $u_i^T(y - \hat{y}) = 0 \implies c_i = \frac{u_i^T y}{u_i^T u_i}$
- So the orthogonal project of y onto W is

$$\mathrm{Proj}_W(y) = \hat{y} = \frac{u_1^T y}{u_1^T u_1} u_1 + \ldots + \frac{u_p^T y}{u_p^T u_p} u_p$$

Orthogonal Decomposition Theorem

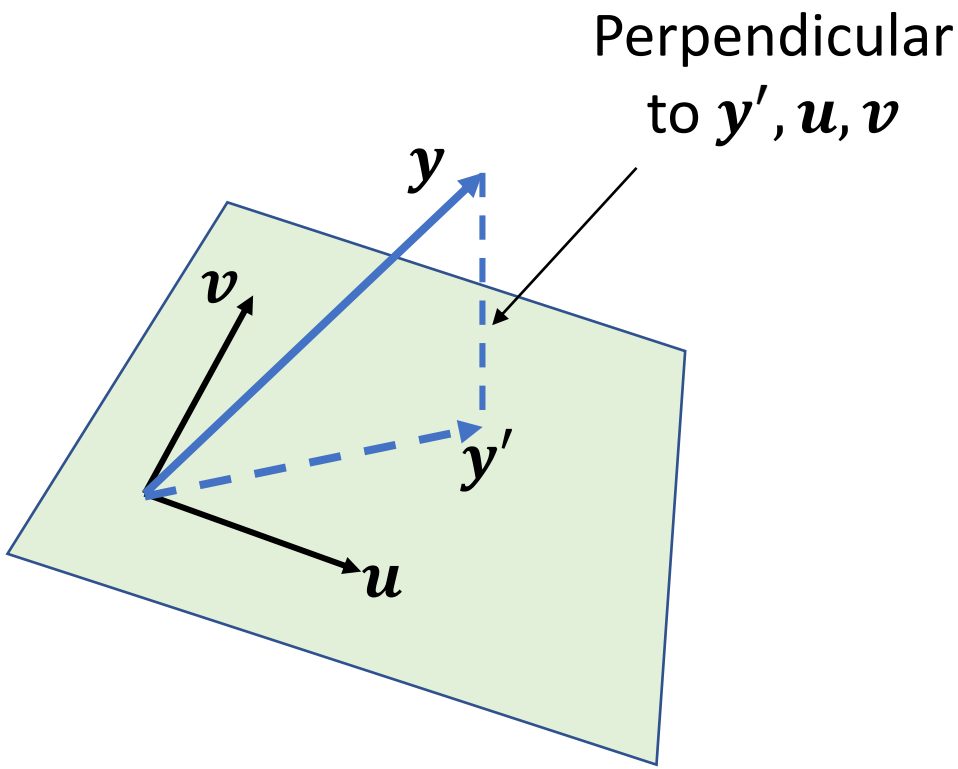
Example

Let $u_1 = \begin{bmatrix} 2 \\ 5 \\ -1 \end{bmatrix}$, $u_2 = \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}$ and $y = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$. Find the orthogonal projection of y onto $W = \mathrm{Span}\{u_1, u_2\}$.

Projection of a vector on a Subspace

- $\boldsymbol{u}$ and $\boldsymbol{v}$ are orthogonal 3D vectors.
- They span a plane (green plane) in 3D
- $\boldsymbol{y}$ is an arbitrary 3D vector out of the plane.
- $\boldsymbol{y}'$ is the projection of $\boldsymbol{y}$ onto the plane.
- $\boldsymbol{y}' = \frac{\boldsymbol{y}\cdot\boldsymbol{u}}{\boldsymbol{u}\cdot\boldsymbol{u}}\boldsymbol{u} + \frac{\boldsymbol{y}\cdot\boldsymbol{v}}{\boldsymbol{v}\cdot\boldsymbol{v}}\boldsymbol{v}$
- The "point" $\boldsymbol{y}'$ is also

the closest point to $\boldsymbol{y}$ on the plane.



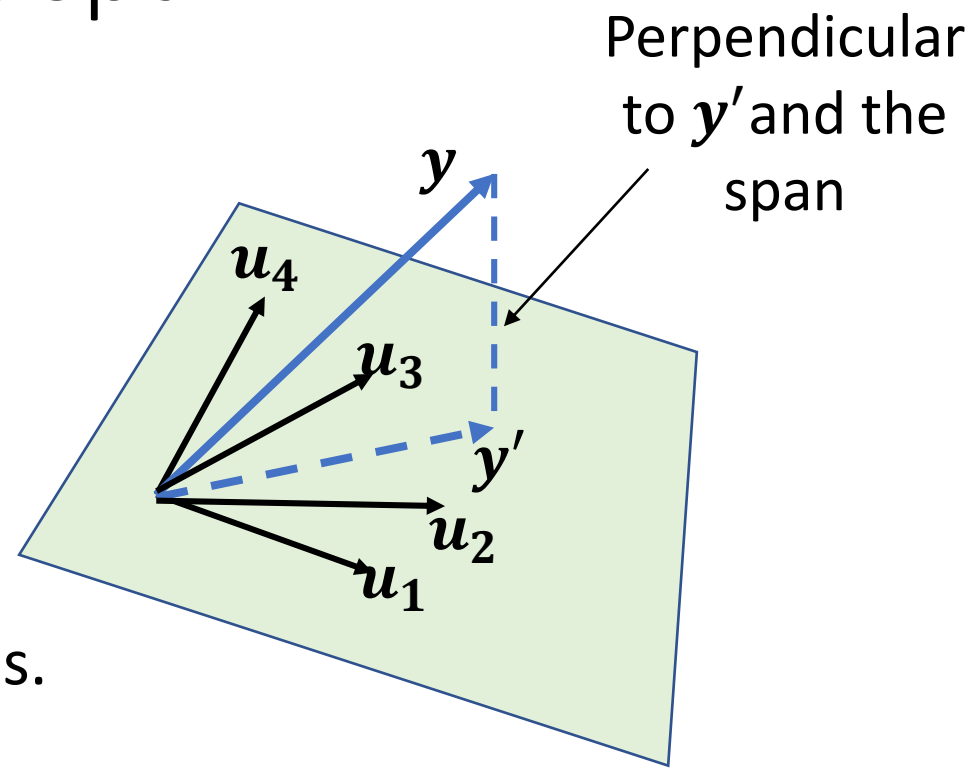
$\boldsymbol{y} - \boldsymbol{y}'$ is perpendicular to $\boldsymbol{y}'$, *Span*$\{\boldsymbol{u}, \boldsymbol{v}\}$, and hence $\boldsymbol{u}$ *and* $\boldsymbol{v}$

Closest point of a vector to a span does not depend on the basis of that span

$\boldsymbol{y}$ is an arbitrary 3D vector out of the plane.

- $\boldsymbol{y}'$ is the projection of $\boldsymbol{y}$ onto the plane.
- $\boldsymbol{y}' = x_1\boldsymbol{u_1} + x_2\boldsymbol{u_2}$
- $\boldsymbol{y}' = x_3\boldsymbol{u_3} + x_4\boldsymbol{u_4}$
- $\boldsymbol{y}' = x_5\boldsymbol{u_1} + x_6\boldsymbol{u_3}$
- …
- Coordinates of $\boldsymbol{y}'$ change if the basis changes.
- But the vector $\boldsymbol{y}'$ itself does not change.
- Hence the **the closest point** to $\boldsymbol{y}$ on the plane does not change.

Even here, $\boldsymbol{y} - \boldsymbol{y}'$is perpendicular to $\boldsymbol{y}'$ and *Span*{.,.}



Closest point of a vector to a span does not depend on the basis of that span

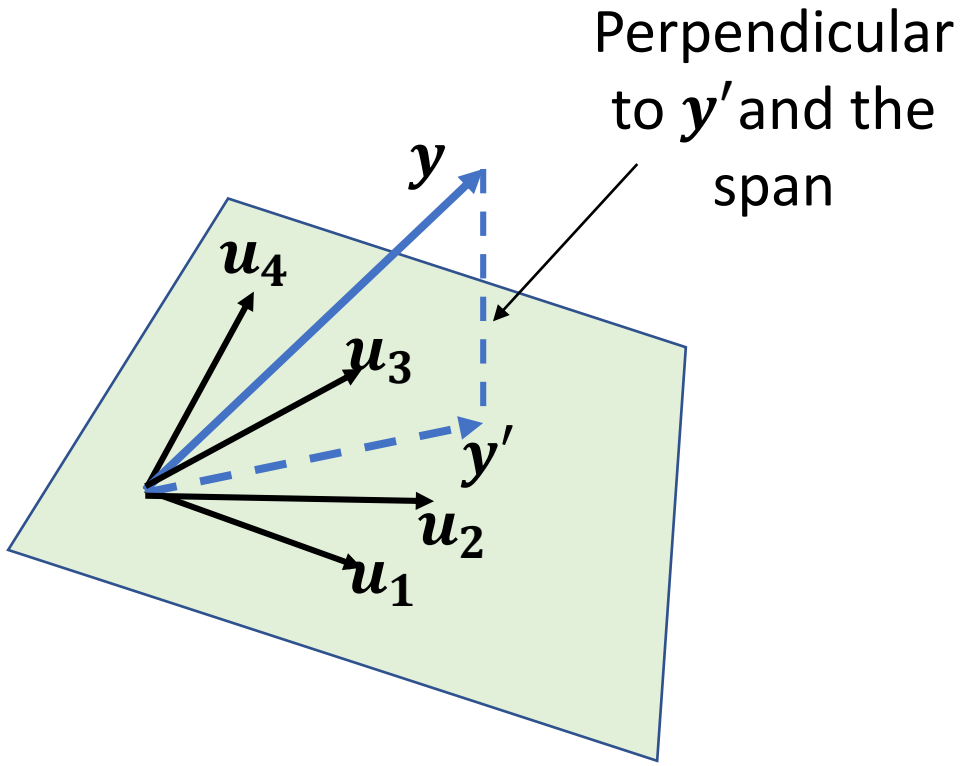
- FINDING THE CLOSEST POINT OF A VECTOR TO A SPAN means :

"Finding the coordinates of the projection of the vector"

- So, if you want to compute the closest point of a vector to a span, then find an appropriate basis with which you can compute the coordinates of the projection easily.

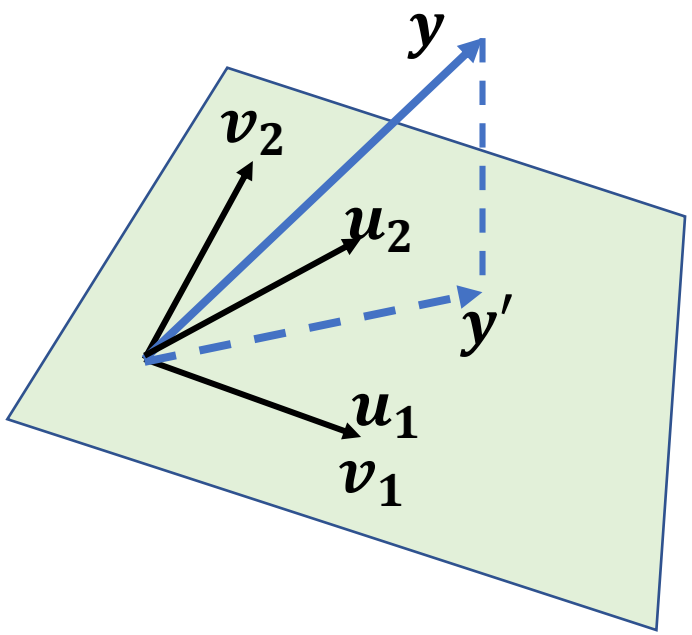
What would be that basis?

Answer: An orthogonal basis!



Projection on a span of non-orthogonal vectors

- How to find projection of any arbitrary 3D vector onto the span of two non-orthogonal, linearly independent vectors?
- $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$ are not orthogonal, but linearly independent vectors in 3D.
- $\boldsymbol{y}$ is an arbitrary 3D vector.
- Find the projection of $\boldsymbol{y}$ in the space spanned by $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$.
  - a) First, find the orthogonal set of vectors $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ that span the same subspace as $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$. In other words, find an orthogonal basis.
  - b) Project $\boldsymbol{y}$ onto the space spanned by orthogonal $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ vectors, as we earlier.



How to find an orthogonal basis?

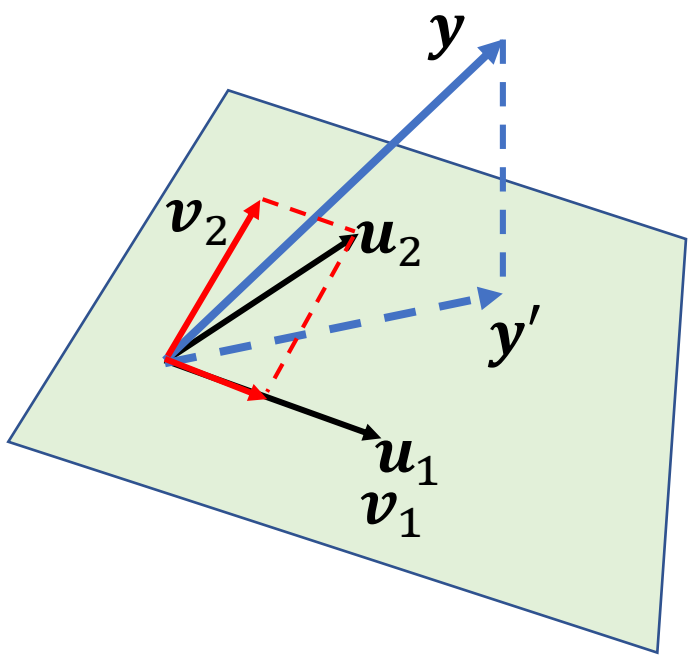
Given a basis $\{x_1, \ldots, x_p\}$ for a subspace W of R^n, find an orthogonal basis $\{v_1, \ldots, v_p\}$ for W such that for any $i = 1, \cdots, p$

$$span\{v_1, \cdots, v_i\} = span\{x_1, \cdots, x_i\}$$

How to find an orthogonal basis?

- Assume that the first vector $\boldsymbol{u}_1$ is in the orthogonal basis. Other vector(s) of the basis are computed that are perpendicular to $\boldsymbol{u}_1$
- Let $\boldsymbol{v}_1 = \boldsymbol{u}_1$
- Let $\boldsymbol{v}_2 = \boldsymbol{u}_2 - \frac{\boldsymbol{u}_2\boldsymbol{v}_1}{\boldsymbol{v}_1\boldsymbol{v}_1}\boldsymbol{v}_1$
- We know that $\boldsymbol{v}_2$ is perpendicular to $\boldsymbol{v}_1$.
- $\boldsymbol{v}_2$ is in the *Span*$\{\boldsymbol{u}_1, \boldsymbol{u}_2\}$ (Why?)
- So *Span*$\{\boldsymbol{u}_1, \boldsymbol{u}_2\}$ = *Span*$\{\boldsymbol{v}_1, \boldsymbol{v}_2\}$
- *And* $\{\boldsymbol{v}_1, \boldsymbol{v}_2\}$ *is an orthogonal basis*
- Projection of $\boldsymbol{y}$ on to the *Span*$\{\boldsymbol{u}_1, \boldsymbol{u}_2\}$

$$\boldsymbol{y}' = \frac{\boldsymbol{y}\cdot\boldsymbol{v}_1}{\boldsymbol{v}_1\cdot\boldsymbol{v}_1}\boldsymbol{v}_1 + \frac{\boldsymbol{y}\cdot\boldsymbol{v}_2}{\boldsymbol{v}_2\cdot\boldsymbol{v}_2}\boldsymbol{v}_2$$



Gram-Schmidt Orthogonalization Process

Given a basis $\{x_1, \ldots, x_p\}$ for a subspace W of R^n, find an orthogonal basis $\{v_1, \ldots, v_p\}$ for W such that for any $i = 1, \cdots, p$

$$span\{v_1, \cdots, v_i\} = span\{x_1, \cdots, x_i\}$$

1. $v_1 = x_1$
2. $v_2 = x_2 - Proj_{v_1}(x_2) = x_2 - \frac{x_2^T v_1}{v_1^T v_1} v_1$
3. $v_3 = x_3 - Proj_{span\{v_1, v_2\}}(x_3) = x_3 - \frac{x_3^T v_1}{v_1^T v_1} v_1 - \frac{x_3^T v_2}{v_2^T v_2} v_2$

   $\vdots$
4. $v_p = x_p - Proj_{span\{v_1, \ldots v_{p-1}\}}(x_p)$

   $= x_p - \frac{x_p^T v_1}{v_1^T v_1} v_1 - \frac{x_p^T v_2}{v_2^T v_2} v_2 + \ldots + \frac{x_p^T v_{p-1}}{v_{p-1}^T v_{p-1}} v_{p-1}$

Gram-Schmidt Orthogonalization Process

Example:

Let $x_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$, $x_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}$ and $x_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}$. Construct an orthogonal basis for $W = \mathrm{Span}\{x_1, x_2, x_3\}$.

Solving Inconsistent Systems

- Suppose $Ax = b$ has no solutions. Can we still find a solution x such that Ax is "closest" to b?
- Most common cases: A is an mxn matrix with $m > n$. The system $Ax = b$ has more equations than variables. So in general there is no solution.
- "Best solution" in the following sense: Find $\hat{x}$ such that $A\hat{x}$ is the **closest point** to b. That is,

$$\|A\hat{x} - b\| \leq \|Ax - b\|$$

  for all x in R^n.
- $\hat{x}$ is called the **least square solution**.

Solving Inconsistent Systems

- Problem: Find $\hat{x}$ such that $A\hat{x}$ is closest to b.
- The problem is equivalent to finding a point $\hat{b}$ in Col A that is closest to b.
- From the best approximation theorem, the point in Col A closest to b is the orthogonal projection of b onto Col A:

$$A\hat{x} = \hat{b} = \mathrm{proj}_{Col\ A}\ b$$

Solving Inconsistent Systems

Example 1:

A trader buys and/or sells tomatoes and potatoes. (Negative number means buys, positive number means sells.) In the process, he either makes profit (positive number) or loss (negative number). A week's transaction is shown; find the approximate cost of tomatoes and potatoes.

Tomatoes (tons)	Potatoes (tons)	Profit/Loss (in thousands)
1	-6	-1
1	-2	2
1	1	1
1	7	6

Solving Inconsistent Systems

- $\mathbf{1}\boldsymbol{t}$ **- 6**$\boldsymbol{p} = \mathbf{-1}$
- $\mathbf{1}\boldsymbol{t}$ **- 2**$\boldsymbol{p} = \mathbf{2}$
- $\mathbf{1}\boldsymbol{t}$ **+ 1**$\boldsymbol{p} = \mathbf{1}$
- $\mathbf{1}\boldsymbol{t}$ **+ 7**$\boldsymbol{p} = \mathbf{6}$
- $\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} t + \begin{bmatrix} -6 \\ -2 \\ 1 \\ 7 \end{bmatrix}$p = $\begin{bmatrix} -1 \\ 2 \\ 1 \\ 6 \end{bmatrix}$

The above equation might not have a solution (values of t and p that would satisfy that equation). So the best we can do is to find the values of t and p that would result in a vector on the right hand side that is as close as possible to the desired right hand side vector.

Solving Inconsistent Systems

Example 2:

EXAMPLE 4 Find a least-squares solution of $A\mathbf{x} = \mathbf{b}$ for

$$A = \begin{bmatrix} 1 & -6 \\ 1 & -2 \\ 1 & 1 \\ 1 & 7 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} -1 \\ 2 \\ 1 \\ 6 \end{bmatrix}$$

SOLUTION Because the columns $\mathbf{a}_1$ and $\mathbf{a}_2$ of A are orthogonal, the orthogonal projection of $\mathbf{b}$ onto Col A is given by

$$\hat{\mathbf{b}} = \frac{\mathbf{b}\cdot\mathbf{a}_1}{\mathbf{a}_1\cdot\mathbf{a}_1}\mathbf{a}_1 + \frac{\mathbf{b}\cdot\mathbf{a}_2}{\mathbf{a}_2\cdot\mathbf{a}_2}\mathbf{a}_2 = \frac{8}{4}\mathbf{a}_1 + \frac{45}{90}\mathbf{a}_2 \tag{5}$$

$$= \begin{bmatrix} 2 \\ 2 \\ 2 \\ 2 \end{bmatrix} + \begin{bmatrix} -3 \\ -1 \\ 1/2 \\ 7/2 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \\ 5/2 \\ 11/2 \end{bmatrix}$$

Now that $\hat{\mathbf{b}}$ is known, we can solve $A\hat{\mathbf{x}} = \hat{\mathbf{b}}$. But this is trivial, since we already know what weights to place on the columns of A to produce $\hat{\mathbf{b}}$. It is clear from (5) that

$$\hat{\mathbf{x}} = \begin{bmatrix} 8/4 \\ 45/90 \end{bmatrix} = \begin{bmatrix} 2 \\ 1/2 \end{bmatrix}$$

■