Lecture 11: Orthogonality and Symmetric Matrices
(book: 6.1, 6.2, 7.1)

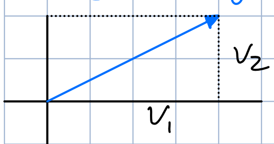
Previous episode: Diagonalization
Next episode: Old Exam (Resit 2022-2023)

---

inner / dot product $\underline{u}, \underline{v} \in \mathbb{R}^n$: $\underline{u} \bullet \underline{v} = \underline{u}^T \cdot \underline{v} = [u_1 \dots u_n] \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = u_1 v_1 + \dots + u_n v_n$

Properties:
* $\underline{u} \cdot \underline{v} = \underline{v} \cdot \underline{u}$ (commutativity)
* $(\underline{u} + \underline{v}) \cdot \underline{w} = \underline{u} \cdot \underline{w} + \underline{v} \cdot \underline{w}$ (distributivity)
* $\underline{0} \cdot \underline{u} = 0$
* $(c \cdot \underline{u}) \cdot \underline{v} = c \cdot (\underline{u} \cdot \underline{v}) = u \cdot (c \cdot \underline{v})$
* $\underline{u} \cdot \underline{u} \geq 0$
* $\underline{u} \cdot \underline{u} = 0 \iff \underline{u} = \underline{0}$.

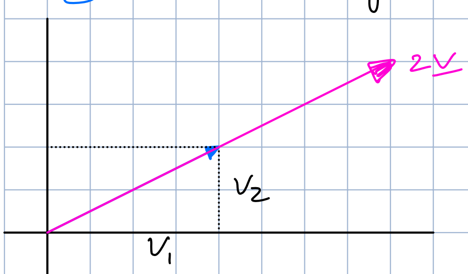
Length of a vector:



$\underline{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$ Length $= \sqrt{v_1^2 + v_2^2} = \sqrt{\underline{v} \bullet \underline{v}}$

Similar in $\mathbb{R}^3$: $\underline{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$ length $= \sqrt{v_1^2 + v_2^2 + v_3^2} = \sqrt{\underline{v} \cdot \underline{v}}$

length / norm of a vector $\underline{v} \in \mathbb{R}^n$: $\|\underline{v}\| = \sqrt{\underline{v} \cdot \underline{v}}$



$c\underline{v} = \begin{bmatrix} c \cdot v_1 \\ c \cdot v_2 \end{bmatrix}$ $\|c \underline{v}\| = |c| \cdot \|\underline{v}\|$ DIY.

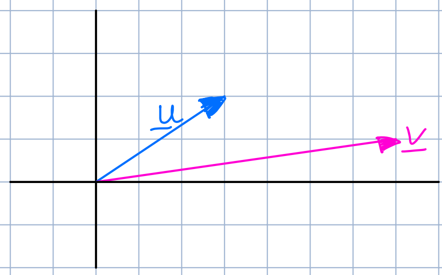
unit vector: Vector of length 1. For example, $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$

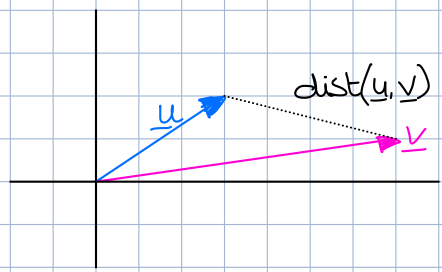
$\underline{v} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ $\|\underline{v}\| = \sqrt{2^2 + 1^2} = \sqrt{5}$.

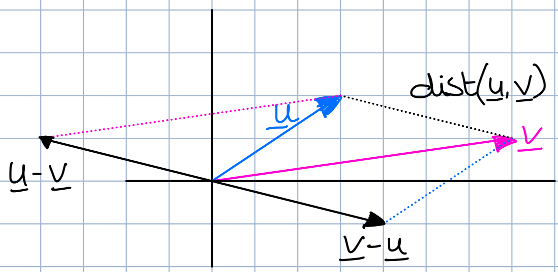
$\underline{u} = \frac{1}{\|\underline{v}\|} \underline{v} = \frac{1}{\sqrt{5}} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix}$ $\underline{w} = \begin{bmatrix} -2/\sqrt{5} \\ -1/\sqrt{5} \end{bmatrix}$

normalizing vector $\underline{v}$.

Distance between two vectors:



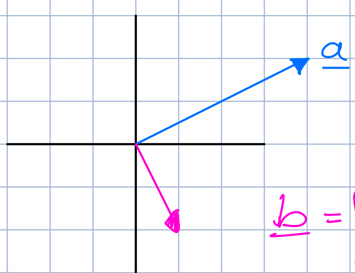




Hence, $\text{dist}(\underline{u},\underline{v}) = \|\underline{v}-\underline{u}\| = \|\underline{u}-\underline{v}\|$.

Two vectors are orthogonal $(\underline{u} \perp \underline{v})$ if $\underline{u} \cdot \underline{v} = 0$.
"perpendicular"

$\iff \|\underline{u}+\underline{v}\|^2 = \|\underline{u}\|^2 + \|\underline{v}\|^2$.



$\underline{a} = \begin{bmatrix} 4 \\ 2 \end{bmatrix}$

$\underline{b} = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$

$\underline{a} \cdot \underline{b} = 4 \cdot 1 + 2 \cdot (-2) = 4 - 4 = 0$

$\underline{a}$ is perpendicular to $\underline{b}$.

Applications:
* it relates to the correlation coefficient in statistics
* it is important for matching & feature detection in signal and image processing.

Recall the Null space of a matrix A.

$$\mathrm{Nul}(A) = \{\underline{x} \in \mathbb{R}^n : A\underline{x} = \underline{0}\}$$

$$\begin{bmatrix} \text{---} & \underline{r}_1 & \text{---} \\ \text{---} & \underline{r}_2 & \text{---} \\ & \vdots & \\ \text{---} & \underline{r}_m & \text{---} \end{bmatrix} \begin{bmatrix} \\ \underline{x} \\ \\ \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$\begin{cases} \underline{r}_1^T \cdot \underline{x} = 0 \\ \underline{r}_2^T \cdot \underline{x} = 0 \\ \quad \vdots \\ \underline{r}_m^T \cdot \underline{x} = 0 \end{cases}$$

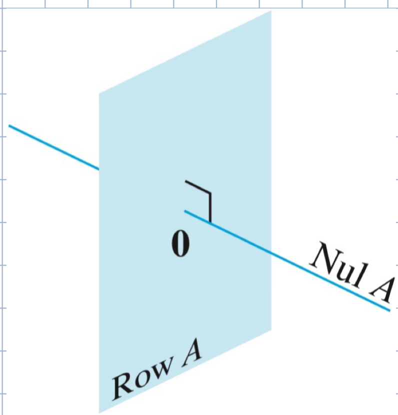
So, every $\underline{x} \in \mathrm{Nul}(A)$ is orthogonal to each of the rows of A.

$$(c_1 \cdot \underline{r}_1^T + c_2 \cdot \underline{r}_2^T + \dots + c_m \cdot \underline{r}_m^T) \cdot \underline{x} = c_1(\underline{r}_1^T \cdot \underline{x}) + c_2 (\underline{r}_2^T \cdot \underline{x}) + \dots + c_m(\underline{r}_m^T \cdot \underline{x})$$
$$= c_1 \cdot 0 + c_2 \cdot 0 + \dots + c_m \cdot 0 = 0.$$

So, every vector in $\mathrm{Nul}(A)$ is orthogonal to every vector in $\mathrm{Row}(A)$

$\Rightarrow$ $\mathrm{Nul}(A) \perp \mathrm{Row}(A)$.

Recall: $\mathrm{Nul}(A)$ in $\mathbb{R}^n$, dim = # free vars; $\mathrm{Row}(A)$ in $\mathbb{R}^n$, dim = # pivot rows. $\sum$ dims $= n$.



And, since $\mathrm{Col}(A) = \mathrm{Row}(A^T)$, we also have $\mathrm{Col}(A) \perp \mathrm{Nul}(A^T)$.

W: subspace of $\mathbb{R}^n$.

$W^\perp$ ("W perpendicular"): orthogonal complement of W.
↳ all vectors in $\mathbb{R}^n$ that are orthogonal to W.

$(\mathrm{Row}(A))^\perp = \mathrm{Nul}(A)$. $\quad ((\mathrm{Col}(A))^\perp = \mathrm{Nul}(A^T)$.

$W^\perp$ is also a subspace of $\mathbb{R}^n$ (exc. 38 Ch 6.1).
$(W^\perp)^\perp = W$.

In general, $\dim(W) + \dim(W^\perp) = n$.

$\{\underline{v}_1, \dots, \underline{v}_k\}$ is an orthogonal set if $\underline{v}_i \cdot \underline{v}_j = 0$ for all $i \neq j$.

Theorem: If $S = \{\underline{v}_1, \dots, \underline{v}_k\}$ is an orthogonal set and $\underline{0} \notin S$, then S is linearly independent and thus S forms a basis for $\mathrm{Span}\{\underline{v}_1, \dots, \underline{v}_k\}$.

Proof: Book Thm 4.

$\{\underline{v}_1, \dots, \underline{v}_k\}$ is an orthonormal set if it is an orthogonal set of unit vectors.
↳ vectors of length 1.

How to test whether $\{\underline{v}_1, \dots, \underline{v}_k\}$ is orthogonal/orthonormal?

Create $A = \begin{bmatrix} | & & | \\ \underline{v}_1 & \cdots & \underline{v}_k \\ | & & | \end{bmatrix}$

Compute $A^TA = \begin{bmatrix} - & v_1^T & - \\ & \vdots & \\ - & v_k^T & - \end{bmatrix} \begin{bmatrix} | & & | \\ v_1 & \cdots & v_k \\ | & & | \end{bmatrix} = \begin{array}{c c} & j \\ i & \begin{bmatrix} & & \\ & \underline{v}_i \cdot \underline{v}_j & \\ & & \end{bmatrix} \end{array}$

$\{\underline{v}_1, \dots \underline{v}_k\}$ is orthogonal $\iff A^TA$ is diagonal.

$\{\underline{v}_1, \dots \underline{v}_k\}$ is orthonormal $\iff A^TA$ is identity matrix.

A square matrix A is an orthogonal matrix $\iff A^TA = I_n \iff A^{-1} = A^T$.

Watch out the terminology: an orthogonal matrix has orthonormal columns

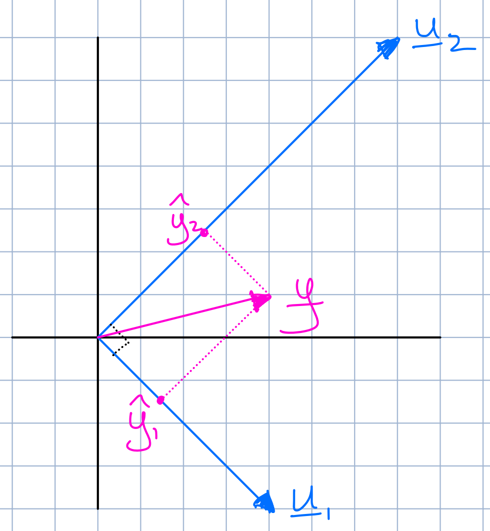
Orthogonal basis for a subspace W of $\mathbb{R}^n$: it is a basis of W, where the vectors form an orthogonal set.

Let $\{\underline{u}_1, \dots, \underline{u}_k\}$ be an orthogonal basis of W.
Let $\underline{y} \in W$.
Then, $\underline{y} = c_1 \cdot \underline{u}_1 + \dots + c_k \cdot \underline{u}_k$.
What are the **weights** $c_1, \dots, c_k$?

$$\underline{y} \bullet \underline{u}_1 = (c_1 \cdot \underline{u}_1 + \dots + c_k \cdot \underline{u}_k) \bullet \underline{u}_1 = c_1 \underline{u}_1 \bullet \underline{u}_1 + \cancel{c_2 \underline{u}_2 \bullet \underline{u}_1} + \cancel{\dots} + \cancel{c_k \underline{u}_k \bullet \underline{u}_1}$$
$$= c_1 \underline{u}_1 \bullet \underline{u}_1$$

$$\Rightarrow c_1 = \frac{\underline{y} \bullet \underline{u}_1}{\underline{u}_1 \bullet \underline{u}_1} \quad \dots\dots \quad c_k = \frac{\underline{y} \bullet \underline{u}_k}{\underline{u}_k \bullet \underline{u}_k}$$

So, it's easy to find the weights :)
(non-orthogonal basis: solving an SLE).



$$\underline{u}_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \quad \underline{y} = \begin{bmatrix} 3 \\ 5 \end{bmatrix}.$$

$$\frac{3 \cdot 2 + 5 \cdot 1}{2 \cdot 2 + 1 \cdot 1}$$

$$\underline{y} = \hat{y}_1 + \hat{y}_2 = \underbrace{\frac{\underline{y} \bullet \underline{u}_1}{\underline{u}_1 \bullet \underline{u}_1} \underline{u}_1}_{\text{orthogonal projection of } \underline{y} \text{ onto } \underline{u}_1.} + \underbrace{\frac{\underline{y} \bullet \underline{u}_2}{\underline{u}_2 \bullet \underline{u}_2} \underline{u}_2}_{\text{orthogonal projection of } \underline{y} \text{ onto } \underline{u}_2.}$$

Recall from a previous episode:
A is diagonalizable $\iff$ the sum of the dimensions of the eigenspaces equals n.

Symmetric matrix: $A = A^T$.

$$\begin{bmatrix} -1 & 6 & -4 \\ 6 & 2 & 0 \\ -4 & 0 & 3 \end{bmatrix}$$

For an $n \times n$ symmetric matrix:
* All eigenvalues are real numbers.
* Eigenvectors from different eigenspaces are orthogonal
* A is diagonalizable. !

Proof: DIY..

A is called orthogonally diagonalizable if there is orthogonal matrix P and a diagonal matrix D such that $A = PDP^{-1}$
or $A = PDP^T$

A is orthogonally diagonalizable $\iff$ A is symmetric.

Example: $A = \begin{bmatrix} 3 & -2 & 4 \\ -2 & 6 & 2 \\ 4 & 2 & 3 \end{bmatrix}$ with eigenvalues -2 and 7.

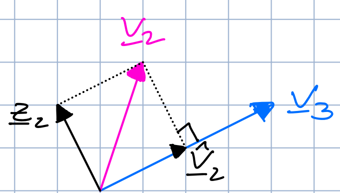
Orthogonally diagonalize A.

$$A-(-2)I = \begin{bmatrix} 5 & -2 & 4 \\ -2 & 8 & 2 \\ 4 & 2 & 5 \end{bmatrix} \sim \dots \sim \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1/2 \\ 0 & 0 & 0 \end{bmatrix} \quad \underline{x} = x_3 \begin{bmatrix} -1 \\ -1/2 \\ 1 \end{bmatrix} \quad \underline{v}_1 = \begin{bmatrix} -1 \\ -1/2 \\ 1 \end{bmatrix}$$

$$A - 7 \cdot I = \begin{bmatrix} -4 & -2 & 4 \\ -2 & -1 & 2 \\ 4 & 2 & -4 \end{bmatrix} \sim \dots \sim \begin{bmatrix} 1 & 1/2 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \underline{x} = x_2 \begin{bmatrix} -1/2 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

$$\underline{v}_2 = \begin{bmatrix} -1/2 \\ 1 \\ 0 \end{bmatrix} \qquad \underline{v}_3 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

We need to make $\underline{v}_2$ and $\underline{v}_3$ orthogonal.



Projection of $\underline{v}_2$ onto $\underline{v}_3$: $\hat{\underline{v}}_2 = \dfrac{\underline{v}_2 \bullet \underline{v}_3}{\underline{v}_3 \bullet \underline{v}_3} \underline{v}_3 = \dfrac{-1/2}{2} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -1/4 \\ 0 \\ -1/4 \end{bmatrix}$

Component of $\underline{v}_2$ orthogonal to $\underline{v}_3$:

$$\underline{z}_2 = \underline{v}_2 - \hat{\underline{v}}_2 = \begin{bmatrix} -1/2 \\ 1 \\ 0 \end{bmatrix} - \begin{bmatrix} -1/4 \\ 0 \\ -1/4 \end{bmatrix} = \begin{bmatrix} -1/4 \\ 1 \\ 1/4 \end{bmatrix}$$

Note: $\underline{z}_2$ is also an eigenvector because it's a linear combination of $\underline{v}_2$ and $\underline{v}_3$.

Moreover $\underline{z}_2 \perp \underline{v}_3$

So $\{\underline{z}_2, \underline{v}_3\} = \left\{ \begin{bmatrix} -1/4 \\ 1 \\ 1/4 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right\}$ forms an orthogonal basis for the eigenspace.

Normalize $\underline{v}_1, \underline{z}_2, \underline{v}_3$: $\underline{u}_1 = \begin{bmatrix} -2/3 \\ -1/3 \\ 2/3 \end{bmatrix}$, $\underline{u}_2 = \begin{bmatrix} -1/\sqrt{18} \\ 4/\sqrt{18} \\ 1/\sqrt{18} \end{bmatrix}$, $\underline{u}_3 = \begin{bmatrix} 1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{bmatrix}$

↳ make the length equal to 1.

Then, $P = [\underline{u}_1 \ \underline{u}_2 \ \underline{u}_3]$ and $D = \begin{bmatrix} -2 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & 7 \end{bmatrix}$. $A = PDP^{-1}$.

↳ orthogonal matrix

$A = PDP^T$

$P^T P = I_n$.