LECTURE 4

SINGULAR VALUE DECOMPOSITION (SVD)

For simplicity suppose $A \in \mathbb{R}^{n\times n}$.
(A is real and square)

$\boxed{\text{SVD OF } A}$: $A = U \Sigma V^T$

* $U, V \in \mathbb{R}^{n\times n}$ s.t.
$$U^T U = V^T V = I$$
* $\Sigma \in \mathbb{R}^{n\times n}$: diagonal with nonnegative entries

EXAMPLE

$$\begin{bmatrix} 0.2 & 3.6 \\ 3.4 & 1.2 \end{bmatrix} = \underbrace{\begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}}_{\substack{U \\ (U^T U = I)}} \underbrace{\begin{bmatrix} 3\sqrt{2} & 0 \\ 0 & 2\sqrt{2} \end{bmatrix}}_{\Sigma} \underbrace{\begin{bmatrix} 0.6 & 0.8 \\ -0.8 & 0.6 \end{bmatrix}}_{\substack{V^T \\ (V^T V = I)}}$$

REMARK

$$V^T V = I \iff \underbrace{v_i^T v_j}_{\substack{(i,j)\text{th} \\ \text{entry of } V^T V}} = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}$$

SVD reveals column and null spaces.

e.g. $A \in \mathbb{R}^{2\times 2}$ with SVD

$$A = \begin{bmatrix} u_1 & u_2 \end{bmatrix} \begin{bmatrix} \sigma_1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} v_1^T \\ v_2^T \end{bmatrix}$$

$$= \begin{bmatrix} u_1 & u_2 \end{bmatrix} \begin{bmatrix} \sigma_1 v_1^T \\ 0 \end{bmatrix}$$

$$= \sigma_1 u_1 v_1^T$$

* $\{v_1, v_2\}$ is a basis for $\mathbb{R}^2$ (EXERCISE)

$$\mathrm{Col}(A) = \{ \sigma_1 u_1 v_1^T (\alpha_1 v_1 + \alpha_2 v_2) \mid \alpha_1, \alpha_2 \in \mathbb{R} \}$$

(SINCE $v_1^T v_2 = 0$ AND $v_1^T v_1 = 1$) $= \{ \alpha_1 \sigma_1 u_1 \mid \alpha_1 \in \mathbb{R} \}$

$$= \mathrm{span}\{u_1\}$$

$$\mathrm{Null}(A) = \{ \alpha_1 v_1 + \alpha_2 v_2 \mid \alpha_1, \alpha_2 \in \mathbb{R} \text{ and } \sigma_1 u_1 v_1^T (\alpha_1 v_1 + \alpha_2 v_2) = 0 \}$$

(SINCE $v_1^T v_1 \neq 0$ AND $v_1^T v_2 = 0$) $= \{ \alpha_1 v_1 + \alpha_2 v_2 \mid \alpha_1, \alpha_2 \in \mathbb{R} \text{ and } \alpha_1 = 0 \}$

$$= \mathrm{span}\{v_2\}$$

GEOMETRIC INTERPRETATION

Let $A \in \mathbb{R}^{2\times 2}$ with SVD

$$A = \begin{bmatrix} u_1 & u_2 \end{bmatrix} \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix} \begin{bmatrix} v_1^T \\ v_2^T \end{bmatrix}$$

$$= \sum_{j=1}^{2} \sigma_j u_j v_j^T$$

Consider the image of unit circle in $\mathbb{R}^2$ under the linear transformation

$$T(v) = Av$$

Unit Circle

$$S = \{ \alpha_1 v_1 + \alpha_2 v_2 \mid \alpha_1, \alpha_2 \in \mathbb{R} \text{ and } \|\alpha_1 v_1 + \alpha_2 v_2\|_2 = 1 \}$$

$$= \{ \alpha_1 v_1 + \alpha_2 v_2 \mid \alpha_1, \alpha_2 \in \mathbb{R} \text{ s.t. } \alpha_1^2 + \alpha_2^2 = 1 \}$$

Image of Unit Circle

$$AS = \{ A(\alpha_1 v_1 + \alpha_2 v_2) \mid \alpha_1, \alpha_2 \in \mathbb{R} \text{ s.t. } \alpha_1^2 + \alpha_2^2 = 1 \}$$

$$= \{ \alpha_1 \sigma_1 u_1 + \alpha_2 \sigma_2 u_2 \mid \alpha_1, \alpha_2 \in \mathbb{R} \text{ s.t. } \alpha_1^2 + \alpha_2^2 = 1 \}$$

REMARK

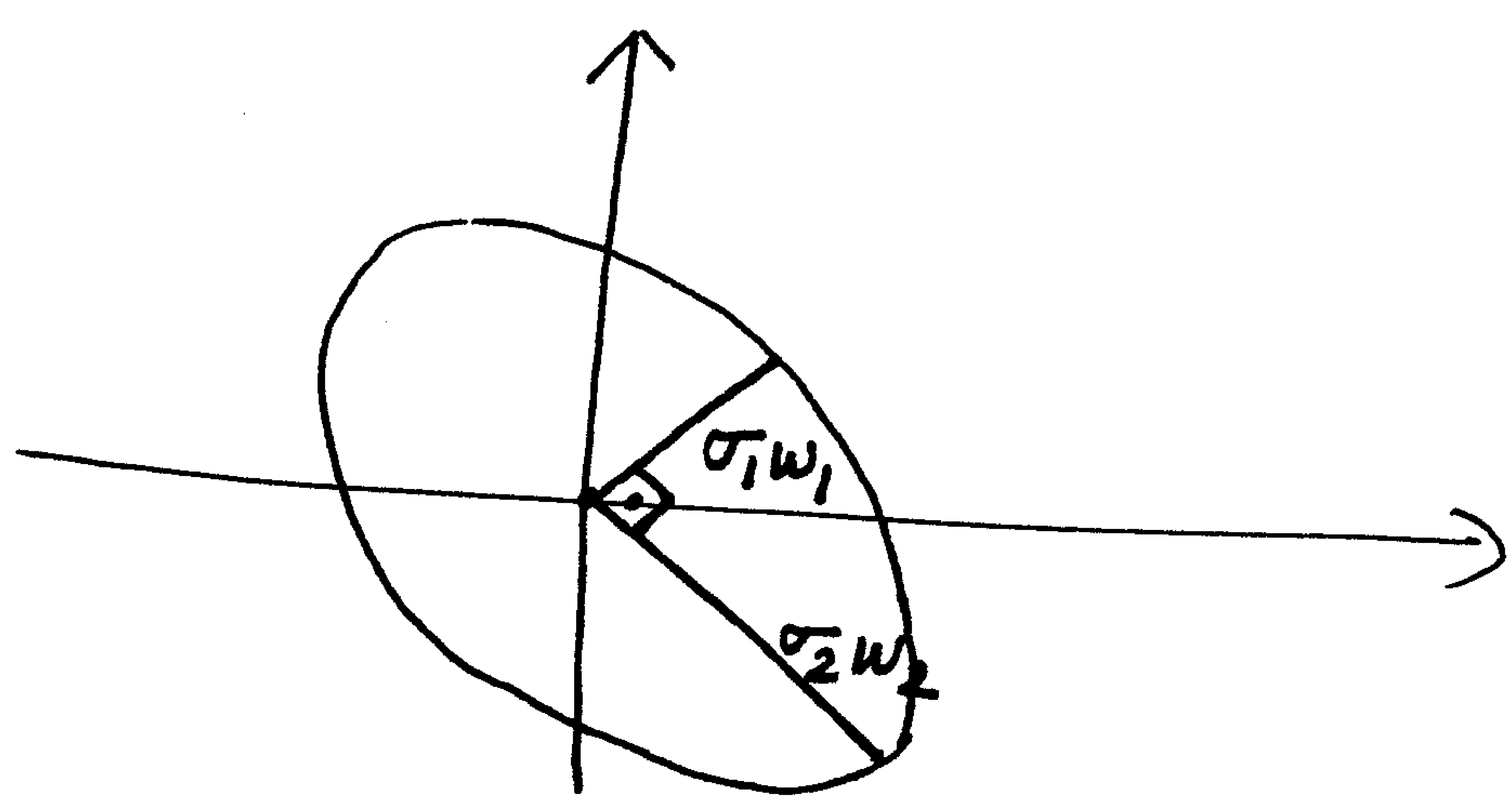
$$A v_1 = (\sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T) v_1 \overset{\left(\substack{v_1^T v_1 = 1 \\ v_2^T v_1 = 0}\right)}{=} \sigma_1 u_1$$

$$A v_2 = \sigma_2 u_2$$

AS is an ellipse.

Ellipse in $\mathbb{R}^2$ $\left(\begin{array}{l} w_1, w_2 \text{ s.t. } w_1^T w_2 = 0 \\ \text{are semi-axes} \end{array} \right)$

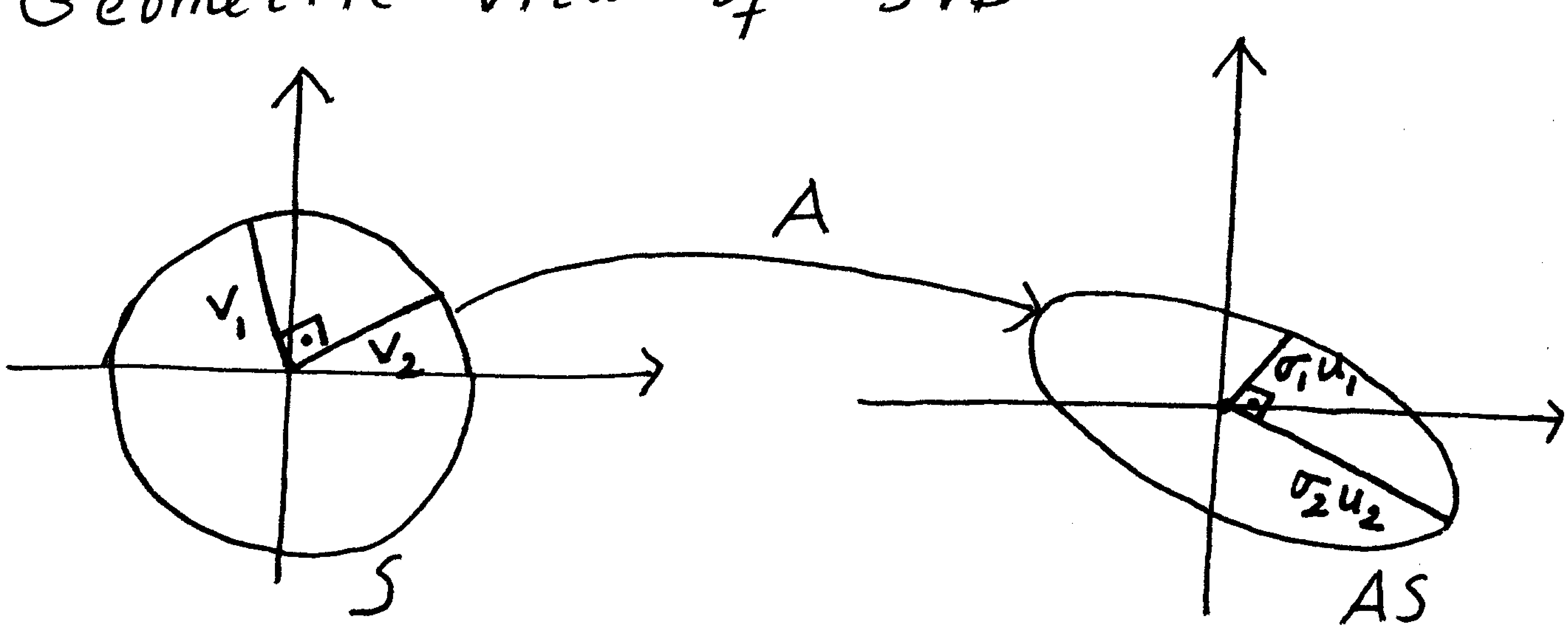
$$\left\{ \alpha_1 w_1 + \alpha_2 w_2 \;\middle|\; \frac{\alpha_1^2}{\sigma_1^2} + \frac{\alpha_2^2}{\sigma_2^2} = 1 \right\}$$



Consider AS again

$$AS = \left\{ \beta_1 u_1 + \beta_2 u_2 \;\middle|\; \frac{\beta_1^2}{\sigma_1^2} + \frac{\beta_2^2}{\sigma_2^2} = 1 \right\}$$

Geometric view of SVD



In $\mathbb{R}^2$

$$S: \begin{array}{l}\text{Unit Circle} \\ \text{with Axes } v_1, v_2\end{array} \xrightarrow{A \in \mathbb{R}^{2\times 2}} \begin{array}{l}\text{Ellipse with: } AS \\ \text{Semi-axes } \sigma_1 u_1, \sigma_2 u_2\end{array}$$

In $\mathbb{R}^n$

$$A = \begin{bmatrix} u_1 & \cdots & u_n \end{bmatrix} \begin{bmatrix} \sigma_1 & & & 0 \\ & \sigma_2 & & \\ & 0 & \ddots & \\ & & & \sigma_n \end{bmatrix} \begin{bmatrix} v_1^T \\ \vdots \\ v_n^T \end{bmatrix}$$

$$= \sum_{j=1}^{n} \sigma_j u_j v_j^T$$

$$\begin{array}{l}\text{Unit } n\text{-Sphere} \\ \text{with axes } v_1, \dots, v_n\end{array} \xrightarrow{A \in \mathbb{R}^{n\times n}} \begin{array}{l}\text{Ellipsoid} \\ \text{with semi-axes} \\ u_1, \dots, u_n\end{array}$$

ORTHOGONALITY

Scalar product of $u, v \in \mathbb{R}^n$

$$u^T v = \sum_{j=1}^{n} u_j v_j$$

Scalar product of $u, v \in \mathbb{C}^n$

$$u^* v = \sum_{j=1}^{n} \overline{u_j}\, v_j$$

NOTE: u^* is the complex conjugate transpose of u

e.g. $u = \begin{bmatrix} 1+2i \\ 1-3i \end{bmatrix}$

$u^* = \begin{bmatrix} 1-2i & 1+3i \end{bmatrix}$

Properties of Scalar Product

(1) Commutes

$\mathbb{R}^n : \quad u^T v = v^T u$

$\mathbb{C}^n : \quad u^* v = \overline{v^* u}$

(2) Bilinear

$$(u_1 + u_2)^T v = u_1^T v + u_2^T v$$

$$(u_1 + u_2)^* v = u_1^* v + u_2^* v$$

(3) Homogeneity

$$(\alpha u)^T v = \alpha (u^T v)$$

(4) Positivity

$$v^T v = v_1^2 + \dots + v_n^2$$
$$= \|v\|_2^2 \geqslant 0 \quad (\text{and } 0 \text{ iff } v=0)$$

$$v^* v = |v_1|^2 + \dots + |v_n|^2$$
$$= \|v\|_2^2 \geqslant 0$$

<u>DEFN</u> (Orthogonality)
Vectors $u, v \in \mathbb{R}^n$ (or $\mathbb{C}^n$) are called orthogonal if $u^T v = 0$ (or $u^* v = 0$).

e.g. $u = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$ and $v = \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}$

are orthogonal in $\mathbb{R}^2$ since $u^T v = 0$.

DEFN (Orthogonal Set)

A set S in $\mathbb{R}^n$ (or $\mathbb{C}^n$) is called orthogonal if

$$v^T w = 0 \qquad \forall v, w \in \mathbb{R}^n \text{ s.t. } v \neq w$$
$$(\text{or } v^* w = 0) \qquad (\forall v, w \in \mathbb{C}^n \text{ s.t. } v \neq w)$$

e.g.

$$\left\{ \underbrace{\begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}}_{v_1}, \underbrace{\begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix}}_{v_2}, \underbrace{\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}}_{v_3} \right\} \text{ is orthogonal.}$$

$$v_1^T v_2 = v_1^T v_3 = v_2^T v_3 = 0$$

$$\left\{ \underbrace{\begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}}_{w_1}, \underbrace{\begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix}}_{w_2}, \underbrace{\begin{bmatrix} 1 \\ 0 \\ +2 \end{bmatrix}}_{w_3} \right\} \text{ is not orthogonal,}$$

$$w_1^T w_3 \neq 0$$

DEFN (Orthogonal Basis)

A basis B in $\mathbb{R}^n$ (or $\mathbb{C}^n$) is called an orthogonal basis if it is also orthogonal.

e.g. $\left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right\}$ is an orthogonal basis for $\mathbb{R}^2$

Orthonormal basis — an orthogonal basis consisting of unit vectors.

e.g. $\left\{ \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}, \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix} \right\}$ is an orthonormal basis for $\mathbb{R}^2$