Vv417 Lecture 21

Jing Liu

UM-SJTU Joint Institute

November 19, 2019

- In $\mathbb{R}^n$, we often use the standard basis

$$\mathcal{B} = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$$

  in which vectors are orthonormal, and it is known as an orthonormal basis.

- In an inner product space $\mathcal{V}$, it is also desirable to have an orthonormal basis.

Definition

Let $\mathcal{S} = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ be a set of nonzero vectors in an inner product space. If

$$\langle \mathbf{v}_i, \mathbf{v}_j \rangle = 0 \qquad \text{for} \quad i \neq j,$$

then $\mathcal{S}$ is said to be an orthogonal set of vectors.

Theorem

An orthogonal set of nonzero vectors is linearly independent.

- If $\mathcal{B}$ is an orthogonal set that spans $\mathcal{V}$, then $\mathcal{B}$ is an orthogonal basis for $\mathcal{V}$.

Definition

An orthonormal set of vectors is an orthogonal set of unit vectors.

- The set $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ will be orthonormal if and only if

$$\langle \mathbf{u}_i, \mathbf{u}_j \rangle = \begin{cases} 1 & \text{if} \quad i = j, \\ 0 & \text{if} \quad i \neq j. \end{cases}$$

Exercise

Consider $\mathcal{C}[-\pi, \pi]$ with inner product $\langle f, g \rangle = \dfrac{1}{\pi} \displaystyle\int_{-\pi}^{\pi} f(x)g(x)\, dx$, is the set

$$\mathcal{S} = \{1,\ \cos x,\ \cos 2x,\ \ldots,\ \cos nx\}$$

an orthogonal set of vectors? Is it orthonormal?

Solution

- We need to check use the given inner product whether they are orthogonal.

Solution

- We need to check the inner products between vectors in $\mathcal{S}$, we have

$$\langle 1, \cos kx \rangle = \frac{1}{\pi}\int_{-\pi}^{\pi} \cos kx\, dx = 0 \qquad \text{for} \quad k = 1, 2, \ldots, n$$

$$\langle \cos jx, \cos kx \rangle = \frac{1}{\pi}\int_{-\pi}^{\pi} \cos jx \cos kx\, dx = 0 \qquad \text{for} \quad j \neq k$$

- We need to check the length of the vectors, we have

$$\langle \cos kx, \cos kx \rangle = \frac{1}{\pi}\int_{-\pi}^{\pi} \cos^2 kx\, dx = 1 \qquad \text{for} \quad k = 1, 2, \ldots, n$$

- However, the length of the function 1 under this inner product is not 1.

$$\|1\|^2 = \langle 1, 1 \rangle = \frac{1}{\pi}\int_{-\pi}^{\pi} 1\; dx = 2$$

- Therefore $\mathcal{S}$ is an orthogonal set, but not orthonormal.

Parseval's Formula

If $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ is an orthonormal basis for an inner product space $\mathcal{V}$, then

$$\|\mathbf{v}\|^2 = \sum_{i=1}^{n} c_i^2 \qquad \text{where} \quad \mathbf{v} = \sum_{i=1}^{n} c_i \mathbf{u}_i$$

Q: The above is easy to show but what does it mean to inner product in general

Q: Given that $\{1/\sqrt{2}, \cos 2x\}$ is an orthonormal set in $\mathcal{C}[-\pi, \pi]$ with

$$\langle f, g \rangle = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) g(x)\, dx$$

How to compute the value of the following without finding an antiderivative?

$$\int_{-\pi}^{\pi} \sin^4 x\, dx$$

Definition

An $n \times n$ matrix $\mathbf{Q}$ is said to be an orthogonal matrix if the column vectors of $\mathbf{Q}$ form an orthonormal set in $\mathbb{R}^n$.

Theorem

An $n \times n$ matrix $\mathbf{Q}$ is orthogonal if and only if

$$\mathbf{Q}^{\mathrm{T}}\mathbf{Q} = \mathbf{I}$$

And $\mathbf{Q}$ is invertible and $\mathbf{Q}^{-1} = \mathbf{Q}^{\mathrm{T}}$.

Proof

- By definition an $n \times n$ matrix $\mathbf{Q}$ is orthogonal if and only if

$$\mathbf{c}_i^{\mathrm{T}}\mathbf{c}_j = \begin{cases} 1 & \text{if} \quad i = j, \\ 0 & \text{if} \quad i \neq j. \end{cases}$$

  where $\mathbf{c}_i$ and $\mathbf{c}_j$ are ith and jth column of $\mathbf{Q}$.

Proof

- Since $\mathbf{c}_i^{\mathrm{T}}\mathbf{c}_j$ is the (i,j) element of $\mathbf{Q}^{\mathrm{T}}\mathbf{Q}$. thus $\mathbf{Q}$ is orthogonal if and only if

$$\mathbf{Q}^{\mathrm{T}}\mathbf{Q} = \mathbf{I}$$

Exercise

Show for any fixed θ, the matrix is orthogonal

$$\mathbf{Q} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

Solution

- Consider the product,

$$\mathbf{Q}^{\mathrm{T}}\mathbf{Q} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} = \mathbf{I}$$

- So $\mathbf{Q}$ is orthogonal.

- In $\mathbb{R}^n$, inner products are preserved under orthonormal change of variables

$$\langle \mathbf{Qx}, \mathbf{Qy} \rangle = (\mathbf{Qy})^{\mathrm{T}}\mathbf{Qx} = \mathbf{y}^{\mathrm{T}}\mathbf{Q}^{\mathrm{T}}\mathbf{Qx} = \mathbf{y}^{\mathrm{T}}\mathbf{x} = \langle \mathbf{x}, \mathbf{y} \rangle$$

- In particular, if $\mathbf{x} = \mathbf{y}$, then

$$\|\mathbf{Qx}\|^2 = \|\mathbf{x}\|^2 \implies \|\mathbf{Qx}\| = \|\mathbf{x}\|$$

- So multiplication by an orthogonal matrix preserves the lengths of vectors.

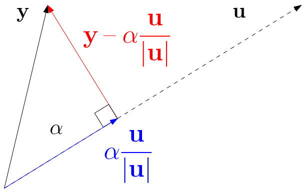
Properties of Orthogonal Matrices

If $\mathbf{Q}$ is an $n \times n$ orthogonal matrix, then

1. $\mathbf{Q}^{\mathrm{T}}\mathbf{Q} = \mathbf{I}$
2. $\mathbf{Q}^{\mathrm{T}} = \mathbf{Q}^{-1}$
3. $\langle \mathbf{Qx}, \mathbf{Qy} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle$
4. $\|\mathbf{Qx}\| = \|\mathbf{x}\|$
5. $\mathbf{QQ}^{\mathrm{T}} = \mathbf{I}$ and thus the rows of $\mathbf{Q}$ form an orthonormal basis for $\mathbb{R}^n$.

- In $\mathbb{R}^n$, we often want to decompose a vector $\mathbf{y}$ into two vector components,

parallel and orthogonal to a vector $\mathbf{u}$.



Q: How to find the two vectors in $\mathbb{R}^n$? What if it is an inner product space?

Definition

For non-zero $\mathbf{y}$ and $\mathbf{u}$ in an inner product space, the vector projection of $\mathbf{y}$ onto $\mathbf{u}$

$$\operatorname{proj}_{\mathbf{u}} \mathbf{y} = \|\mathbf{y}\| \cos\theta \frac{\mathbf{u}}{\|\mathbf{u}\|} = \frac{\mathbf{u}}{\|\mathbf{u}\|} \left(\frac{\langle \mathbf{u}, \mathbf{y} \rangle}{\|\mathbf{u}\|} \right) = \frac{\langle \mathbf{u}, \mathbf{y} \rangle}{\langle \mathbf{u}, \mathbf{u} \rangle} \mathbf{u}$$

The vector $(\mathbf{y} - \operatorname{proj}_{\mathbf{u}} \mathbf{y})$ is called the vector component of $\mathbf{y}$ orthogonal to $\mathbf{u}$.

Definition

A vector $\mathbf{x}$ in an inner product $\mathcal{V}$ is said to be orthogonal to a subspace $\mathcal{W}$ of $\mathcal{V}$ if $\mathbf{x}$ is orthogonal to all the vectors $\mathbf{v}$ in $\mathcal{W}$, that is,

$$\langle \mathbf{x}, \mathbf{v} \rangle = 0 \qquad \text{for all vectors } \mathbf{v} \text{ in } \mathcal{W}.$$

- It is clear if $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ is a basis of $\mathcal{W}$, then $\mathbf{x}$ is orthogonal to $\mathcal{W}$

  if and only if $\mathbf{x}$ is orthogonal to all the vectors in $\mathcal{B}$.

Theorem

For a vector $\mathbf{x}$ in $\mathcal{V}$ and a subspace $\mathcal{W}$ of $\mathcal{V}$, we can write

$$\mathbf{x} = \mathbf{x}^{\parallel} + \mathbf{x}^{\perp}$$

where $\mathbf{x}^{\parallel}$ is in $\mathcal{W}$ and $\mathbf{x}^{\perp}$ is orthogonal to $\mathcal{W}$, and this representation is unique.

Proof

- Let $\mathbf{x} \in \mathcal{V}$ and $\mathcal{B} = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m\}$ be an orthogonal basis of $\mathcal{W}$, then

$$\mathbf{x}^{\parallel} = \alpha_1 \mathbf{u}_1 + \alpha_2 \mathbf{u}_2 + \cdots + \alpha_m \mathbf{u}_m$$

is in $\mathcal{W}$, we need to show there is a set of α_i's such that

$$\mathbf{x}^{\perp} = \mathbf{x} - \mathbf{x}^{\parallel}$$

is orthogonal to $\mathcal{W}$, that is,

$$\begin{aligned} \langle \mathbf{x}^{\perp}, \mathbf{u}_i \rangle &= 0 \qquad \text{for} \quad i = 1, 2, \ldots, m \\ \langle \mathbf{x} - \mathbf{x}^{\parallel}, \mathbf{u}_i \rangle &= 0 \\ \langle \mathbf{x}, \mathbf{u}_i \rangle - \langle \mathbf{x}^{\parallel}, \mathbf{u}_i \rangle &= 0 \\ \Longrightarrow \ \alpha_i &= \frac{\langle \mathbf{x}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \end{aligned}$$

- This shows it is always possible to find such α_i's. $\square$

Definition

The vector $\mathbf{x}^{\parallel}$ is called the orthogonal projection of $\mathbf{x}$ onto $\mathcal{W}$, denoted by

$$\mathbf{x}^{\parallel} = \operatorname{proj}_{\mathcal{W}}(\mathbf{x})$$

- The transformation $T(\mathbf{x}) = \operatorname{proj}_{\mathcal{W}}(\mathbf{x})$ is linear.

Formula for the orthogonal projection

If $\mathcal{W}$ is a subspace of $\mathcal{V}$ with an orthonormal basis $\mathcal{B} = \{\mathbf{u}_1, \mathbf{u}_2 \dots, \mathbf{u}_m\}$, then

$$\mathbf{x}^{\parallel} = \operatorname{proj}_{\mathcal{W}}(\mathbf{x}) = \langle \mathbf{u}_1, \mathbf{x}\rangle \mathbf{u}_1 + \langle \mathbf{u}_2, \mathbf{x}\rangle \mathbf{u}_2 + \cdots + \langle \mathbf{u}_m, \mathbf{x}\rangle \mathbf{u}_m \qquad \text{for all } \mathbf{x} \text{ in } \mathcal{V}.$$

Exercise

Given $\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & -1 \\ 1 & 1 \end{bmatrix}$ and $\mathbf{x} = \begin{bmatrix} 1 \\ 3 \\ 1 \\ 7 \end{bmatrix}$. Find orthogonal projection of $\mathbf{x}$ onto $\operatorname{col}(\mathbf{A})$.

$$\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & -1 \\ 1 & 1 \end{bmatrix} \quad \text{and} \quad \mathbf{x} = \begin{bmatrix} 1 \\ 3 \\ 1 \\ 7 \end{bmatrix}$$

Solution

- Note the two columns happen to be orthogonal but not orthonormal, hence

$$\begin{aligned} \mathbf{x}^{\parallel} &= \frac{\langle \mathbf{u}_1, \mathbf{x} \rangle}{\langle \mathbf{u}_1, \mathbf{u}_1 \rangle} \mathbf{u}_1 + \frac{\langle \mathbf{u}_2, \mathbf{x} \rangle}{\langle \mathbf{u}_2, \mathbf{u}_2 \rangle} \mathbf{u}_2 \\ &= \frac{(\mathbf{u}_1 \cdot \mathbf{x})}{(\mathbf{u}_1 \cdot \mathbf{u}_1)} \mathbf{u}_1 + \frac{(\mathbf{u}_2 \cdot \mathbf{x})}{(\mathbf{u}_2 \cdot \mathbf{u}_2)} \mathbf{u}_2 \\ &= \frac{12}{4} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} + \frac{4}{4} \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 2 \\ 2 \\ 4 \end{bmatrix} \end{aligned}$$

Exercise

For an inner product space $\mathcal{V}$, consider the orthogonal projection

$$T(\mathbf{x}) = \mathrm{proj}_{\mathcal{W}}(\mathbf{x}) \qquad \text{where} \quad \mathbf{x} \in \mathcal{V}$$

onto a subspace $\mathcal{W}$ of $\mathcal{V}$. Describe how the range and kernel of T relate to $\mathcal{W}$.

Solution

- First of all, both the range and kernel are subspace of $\mathcal{V}$, and the range is a subspace of $\mathcal{W}$, since if $\mathcal{S} = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ is an orthogonal basis of $\mathcal{W}$, then

$$T(\mathbf{x}) = \frac{\langle \mathbf{u}_1, \mathbf{x} \rangle}{\langle \mathbf{u}_1, \mathbf{u}_1 \rangle}\mathbf{u}_1 + \frac{\langle \mathbf{u}_2, \mathbf{x} \rangle}{\langle \mathbf{u}_2, \mathbf{u}_2 \rangle}\mathbf{u}_2 + \cdots + \frac{\langle \mathbf{u}_n, \mathbf{x} \rangle}{\langle \mathbf{u}_n, \mathbf{u}_n \rangle}\mathbf{u}_n$$

- The kernel is a subspace of $\mathcal{V}$ such that

$$T(\mathbf{x}^*) = \mathbf{0} \implies \langle \mathbf{x}^*, \mathbf{u}_j \rangle = 0 \qquad \text{for} \quad j = 1, 2, \ldots n.$$

- So the kernel is a vector space in which all vectors in it is orthogonal to $\mathcal{W}$.

Q: Given an $m \times n$ matrix $\mathbf{A}$ and

$$\mathbf{x} \in \text{null}(\mathbf{A})$$

is there a vector in the row space of $\mathbf{A}$ that is orthogonal to $\mathbf{x}$?

Q: Given an $m \times n$ matrix $\mathbf{A}$ and

$$\mathbf{x} \in \text{null}(\mathbf{A})$$

is there any vector in the row space of $\mathbf{A}$ that is parallel to $\mathbf{x}$?

Definition

Two subspaces $\mathcal{X}$ and $\mathcal{Y}$ of an inner product space $\mathcal{V}$ are said to be orthogonal if $\langle \mathbf{x}, \mathbf{y} \rangle = 0$ for every $\mathbf{x} \in \mathcal{X}$ and every $\mathbf{y} \in \mathcal{Y}$. If $\mathcal{X}$ and $\mathcal{Y}$ are orthogonal, we write

$$\mathcal{X} \perp \mathcal{Y}$$

- Therefore, we have the following according to this definition and notation

$$\text{null}(\mathbf{A}) \perp \text{col}(\mathbf{A}^{\mathrm{T}})$$

Definition

Consider a subspace $\mathcal{W}$ of an inner product space $\mathcal{V}$. The orthogonal complement $\mathcal{W}^{\perp}$ of $\mathcal{W}$ is the set of all vectors $\mathbf{x} \in \mathcal{V}$ that are orthogonal to all vectors in $\mathcal{W}$:

$$\mathcal{W}^{\perp} = \{\mathbf{x} \in \mathcal{V} : \langle \mathbf{v}, \mathbf{x} \rangle = 0, \text{ for all } \mathbf{v} \in \mathcal{W}\}$$

- Note that $\mathcal{W}^{\perp}$ is the kernel of the orthogonal projection onto $\mathcal{W}$.

Properties of the orthogonal complement

Consider a subspace $\mathcal{W}$ of an inner product space $\mathcal{V}$.

1. The orthogonal complement $\mathcal{W}^{\perp}$ of $\mathcal{W}$ is a subspace of $\mathcal{V}$.
2. The intersection of $\mathcal{W}^{\perp}$ and $\mathcal{W}$ consists of the zero vector alone:

$$\mathcal{W}^{\perp} \cap \mathcal{W} = \{\mathbf{0}\}.$$

3. $\dim \mathcal{W} + \dim \mathcal{W}^{\perp} = n$
4. $\left(\mathcal{W}^{\perp}\right)^{\perp} = \mathcal{W}$

Proof

1. Let $T(\mathbf{x}) = \mathrm{proj}_{\mathcal{W}}(\mathbf{x})$, then $\mathcal{W}^{\perp} = \mathrm{kernel}(T)$. Thus it is a vector space.
2. If a vector $\mathbf{x}$ is in $\mathcal{W}$ as well as $\mathcal{W}^{\perp}$, then $\mathbf{x}$ is orthogonal to itself

$$\langle \mathbf{x}, \mathbf{x} \rangle = 0 \implies \mathbf{x} = \mathbf{0}$$

Q: Suppose $\mathbf{e}_1, \mathbf{e}_2 \in \mathbb{R}^2$ and let

$$\mathcal{X} = \mathrm{span}(\mathbf{e}_1) \qquad \text{and} \qquad \mathcal{Y} = \mathrm{span}(\mathbf{e}_2)$$

are $\mathcal{X}$ and $\mathcal{Y}$ orthogonal complement of each other?

Q: What if $\mathbf{e}_1, \mathbf{e}_2 \in \mathbb{R}^n$, where $n \geq 3$?

Fundamental Subspaces Theorem

If $\mathbf{A}$ is an $m \times n$ matrix, then

$$\mathrm{null}(\mathbf{A}) = \left(\mathrm{col}(\mathbf{A}^{\mathrm{T}})\right)^{\perp} \qquad \text{and} \qquad \mathrm{null}(\mathbf{A}^{\mathrm{T}}) = \left(\mathrm{col}(\mathbf{A})\right)^{\perp}$$

Proof

- We know that

$$\operatorname{null}(\mathbf{A}) \perp \operatorname{col}\left(\mathbf{A}^{\mathrm{T}}\right) \implies \operatorname{null}(\mathbf{A}) \subset \left(\operatorname{col}\left(\mathbf{A}^{\mathrm{T}}\right)\right)^{\perp}$$

- Let $\mathbf{x} \in \left(\operatorname{col}(\mathbf{A}^{\mathrm{T}})\right)^{\perp}$, then $\mathbf{x}$ must be orthogonal to each row of $\mathbf{A}$,

$$\begin{aligned} \mathbf{A}\mathbf{x} = \mathbf{0} \implies \mathbf{x} \in \operatorname{null}(\mathbf{A}) &\implies \left(\operatorname{col}\left(\mathbf{A}^{\mathrm{T}}\right)\right)^{\perp} \subset \operatorname{null}(\mathbf{A}) \\ &\implies \left(\operatorname{col}\left(\mathbf{A}^{\mathrm{T}}\right)\right)^{\perp} = \operatorname{null}(\mathbf{A}) \end{aligned}$$

- Let $\mathbf{B}^{\mathrm{T}} = \mathbf{A}$, then

$$\left(\operatorname{col}\left(\mathbf{B}\right)\right)^{\perp} = \operatorname{null}(\mathbf{B}^{\mathrm{T}}) \implies \left(\operatorname{col}\left(\mathbf{A}\right)\right)^{\perp} = \operatorname{null}(\mathbf{A}^{\mathrm{T}}) \quad \square$$

- Similarly things can be said about the kernel and range of a linear map.

Definition

Suppose $\mathcal{U}$ and $\mathcal{V}$ are subspaces of an inner product space $\mathcal{W}$ and each $\mathbf{w} \in \mathcal{W}$ can be written uniquely as a sum

$$\mathbf{u} + \mathbf{v} \qquad \text{where } \mathbf{u} \in \mathcal{U} \text{ and } \mathbf{v} \in \mathcal{V},$$

then $\mathcal{W}$ is called the direct sum of $\mathcal{U}$ and $\mathcal{V}$, and we write

$$\mathcal{W} = \mathcal{U} \oplus \mathcal{V}$$

Q What is the difference between the sum $\mathcal{U} + \mathcal{V}$ and the direct sum of $\mathcal{U} \oplus \mathcal{V}$?

Theorem

If $\mathcal{S}$ is a subspace of $\mathcal{W}$, then

$$\mathcal{W} = \mathcal{S} \oplus \mathcal{S}^{\perp}$$

Q: How do the four fundamental spaces of $\mathbf{A}$ relate to each other?

- Recall the row space and the column space have the same dimension.

$$\dim(\mathrm{col}(\mathbf{A})) = \dim(\mathrm{col}(\mathbf{A}^{\mathrm{T}})) = \mathrm{rank}(\mathbf{A})$$

- Actually, $\mathbf{A}$ can be used to establish a one-to-one correspondence between

$$\mathrm{col}(\mathbf{A}) \qquad \text{and} \qquad \mathrm{col}(\mathbf{A}^{\mathrm{T}})$$

- We can treat an $m \times n$ matrix $\mathbf{A}$ as a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$:

$$\mathbf{x} \in \mathbb{R}^n \to \mathbf{A}\mathbf{x} \in \mathbb{R}^m$$

- Since $\mathrm{col}(\mathbf{A}^{\mathrm{T}})$ and $\mathrm{null}(\mathbf{A})$ are orthogonal complements in $\mathbb{R}^n$,

$$\mathbb{R}^n = \mathrm{col}(\mathbf{A}^{\mathrm{T}}) \oplus \mathrm{null}(\mathbf{A})$$

- Each vector $\mathbf{x} \in \mathbb{R}^n$ can be written as a sum

$$\mathbf{x} = \mathbf{y} + \mathbf{z} \qquad \text{where} \quad \mathbf{y} \in \mathrm{col}(\mathbf{A}^{\mathrm{T}}) \quad \text{and} \quad \mathbf{z} \in \mathrm{null}(\mathbf{A})$$
$$\mathbf{A}\mathbf{x} = \mathbf{A}\mathbf{y} + \mathbf{A}\mathbf{z} = \mathbf{A}\mathbf{y}$$

- Therefore

$$\mathrm{col}(\mathbf{A}) = \{\mathbf{A}\mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n\} = \{\mathbf{A}\mathbf{y} \mid \mathbf{y} \in \mathrm{col}(\mathbf{A}^{\mathrm{T}})\}$$

- Thus, if we restrict the domain of $\mathbf{A}$ to $\mathrm{col}(\mathbf{A}^{\mathrm{T}})$, then

$$\mathbf{A} \text{ maps } \mathrm{col}(\mathbf{A}^{\mathrm{T}}) \text{ onto } \mathrm{col}(\mathbf{A})$$

- Furthermore, the mapping is one-to-one. If $\mathbf{x}_1$ and $\mathbf{x}_2 \in \mathrm{col}(\mathbf{A}^{\mathrm{T}})$,

$$\mathbf{A}\mathbf{x}_1 = \mathbf{A}\mathbf{x}_2 \implies \mathbf{A}(\mathbf{x}_1 - \mathbf{x}_2) = \mathbf{0} \implies (\mathbf{x}_1 - \mathbf{x}_2) \in \mathrm{col}(\mathbf{A}^{\mathrm{T}}) \cap \mathrm{null}(\mathbf{A})$$

- Since $\mathrm{col}(\mathbf{A}^{\mathrm{T}}) \cap \mathrm{null}(\mathbf{A}) = \{\mathbf{0}\}$, it follows that

$$\mathbf{x}_1 = \mathbf{x}_2$$

- So every $m \times n$ matrix $\mathbf{A}$ is invertible when viewed as a linear transformation

$$T_{\mathbf{A}} : \mathrm{col}(\mathbf{A}^{\mathrm{T}}) \to \mathrm{col}(\mathbf{A})$$

and the matrix $\mathbf{A}^+$ corresponds to $T_{\mathbf{A}}^{-1}$ is known as the pseudoinverse of $\mathbf{A}$.

- For example, consider the following matrix

$$\mathbf{A} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \end{bmatrix}$$

- This matrix is clearly not invertible since it is not even a square matrix, so the corresponding matrix transformation is not invertible.
- However, it is also clear that,

$$\operatorname{col}(\mathbf{A}^{\mathrm{T}}) = \operatorname{span}\{\mathbf{e}_1, \mathbf{e}_2\} \qquad \text{and} \qquad \operatorname{null}(\mathbf{A}) = \operatorname{span}\{\mathbf{e}_3\}$$

- So if we restrict ourselves to $\mathbf{y} \in \operatorname{col}(\mathbf{A}^{\mathrm{T}})$, that is, $\mathbf{y}$ must be of $\begin{bmatrix} y_1 \\ y_2 \\ 0 \end{bmatrix}$, then

$$T_{\mathbf{A}} : \operatorname{col}(\mathbf{A}^{\mathrm{T}}) \to \operatorname{col}(\mathbf{A})$$

  is an invertible linear transformation, and must have an inverse.
- We will come back to this and find the pseudoinverse towards the end!

- I will leave it to you to add the following to the vector/dual spaces figure.