Review: The space $L(V, W)$

$(f+g)(x) = f(x) + g(x)$

$(cf)(x) = cf(x).$

Definition

Let V and W be vector spaces. The set of all linear transformations from V to W is called $L(V, W)$.

Proposition

$L(V, W)$ is a vector space under pointwise addition of functions and pointwise multiplication of functions by real numbers.

Review

$W = \mathbb{R}$.

$L(V, \mathbb{R}) \sim M_{1 \times n}(\mathbb{R})$ $\dim W = 1$.

$m = 1$.

Theorem

Let $\dim V = n$ *and* $\dim W = m$. *Then* $L(V, W)$ *is isomorphic to the vector space* $\underline{M_{m\times n}(\mathbb{R})}$ *of* $m \times n$ *matrices.*

(In other words there exists a bijective linear transformation from $L(V, W)$ *to* $M_{m\times n}(\mathbb{R})$*).*

Review

$$\begin{tikzcd} U \arrow[r, "S"] \arrow[rr, bend right, "T \circ S"'] & V \arrow[r, "T"] & W \end{tikzcd}$$

Theorem

Let U, V and W be finite dimensional vector spaces of dimensions m, n and k respectively, having bases $\mathcal{A}, \mathcal{B}$ and $\mathcal{C}$ respectively.

Let $S \in L(U, V)$ and $T \in L(V, W)$. Then $T \circ S \in L(U, W)$ and

$$[T \circ S]_{\mathcal{A},\mathcal{C}} = [T]_{\mathcal{B},\mathcal{C}}[S]_{\mathcal{A},\mathcal{B}}$$

Linear Operators

Definition

Let V be a vector space. A linear mapping $T : V \to V$ is called a *linear operator*.

Remark

The vector space of all linear operators on a vector space is called $L(V, V)$. This is a special case of an $L(V, W)$ type space with $V = W$. The space $L(V, V)$ is isomorphic to $M_{n \times n}$, the set of all $n \times n$ square matrices having real entries.

$\dim V = n$.

$Av = \lambda v$

Definition

Let V be a vector space. Let $T \in L(V, V)$. We say $\lambda \in \mathbb{R}$ is an *eigenvalue of* T if there exists a non-zero vector $v \in V$, called an *eigenvector of* T such that

$$T(v) = \lambda v.$$

Theorem

Let V be a finite dimensional vector space. Let $\mathcal{B} = \{v_1, \ldots, v_n\}$ be a basis of V. Let $T \in L(V, V)$. Then λ is an eigenvalue of T iff λ is an eigenvalue of $[T]_{\mathcal{B}}$.

Proof: Assume λ is an eigenvalue of T.

$\therefore \exists v \in V, v \neq 0$ such that $T(v) = \lambda v$.

$$\Longrightarrow \underline{[T(v)]_\beta} = [\lambda v]_\beta = \lambda [v]_\beta \quad (1)$$

We also know that

$$[T(v)]_\beta = [T]_\beta [v]_\beta \quad \text{--- } (2)$$

(by definition of $[T]_\beta$).

From (1) & (2),

$$[T]_\beta [v]_\beta = \lambda [v]_\beta .$$

$\therefore$ λ is an eigenvalue of $[T]_\beta$.

Conversely suppose λ is an eigenvalue of $[T]_\beta$.

$\Rightarrow \exists\ v \in \mathbb{R}^n$ such that

$$[T]_\beta v = \lambda v. \quad \text{--- } 3$$

Let $w \in V$ be chosen such that $[w]_\beta = v$. --- 4

From 3 and 4, we have

$$[T]_\beta [w]_\beta = \lambda [w]_\beta .$$

$$\Rightarrow [Tw]_\beta = \lambda [w]_\beta$$

$\therefore Tw = \lambda w \Rightarrow \lambda$ is an eigenvalue of T.

Definition

Let V be a finite dimensional vector space. A linear operator $T \in L(V, V)$ is said to be *diagonalizable* if there exists a basis of V consisting of eigenvectors of T.

Theorem

Let V be a finite dimensional vector space. Let $\mathcal{B} = \{v_1, \ldots, v_n\}$ be a basis of V. Let $T \in L(V, V)$. Then T is diagonalizable iff $[T]_{\mathcal{B}}$ is diagonalizable.

Proof: Assume T is diagonalizable.

Let $\mathcal{C} = \{w_1, \ldots, w_n\}$ be a basis of V consisting of eigenvectors of T, i.e.

$$T(w_j) = \lambda_j w_j, \text{ for } j = 1, \ldots, n$$

where $\lambda_1, \ldots, \lambda_n$ are eigenvalues of T (not necessarily distinct).

$$[T(w_j)]_{\varphi} = [\lambda w_j]_{\varphi}$$

$$= \lambda [w_j]_{\varphi} . \quad \text{— } (1)$$

$$[T(\omega_j)]_{\varphi} = [T]_{\varphi}[\omega_j]_{\varphi} \quad \text{--- (2)}$$

From (1) & (2) $[T]_{\varphi}[\omega_j] = \lambda_j[\omega_j]$. $\forall j = 1, \dots, n$.

Let $\{e_1, \ldots, e_n\}$ be the standard basis of $\mathbb{R}^n$.

Then $[T]_{\mathcal{E}} e_j = \lambda_j e_j$

$$\therefore \quad [T]_{\mathcal{E}} = [\lambda_1 e_1 \ \ldots \ \lambda_n e_n]$$

$\therefore$ $[T]_{\mathcal{C}}$ is diagonal.

Since $[T]_{\beta}$ is similar to $[T]_{\mathcal{C}}$, $[T]_{\beta}$ is diagonalizable.

Conversely, assume that

$[T]_\beta$ is diagonalizable.

Then there exists a basis $\{w_1, \ldots, w_n\}$ of $\mathbb{R}^n$ consisting of eigenvectors of $[T]_\beta$.

there exist $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ such that

i.e. $\wedge [T]_\beta \underline{w_j} = \lambda_j w_j$ — (1)

for every $j = 1, \ldots, n$.

Let $u_j \in V$ be chosen such that $[u_j]_\beta = w_j$ — (2) for each $j = 1, \ldots, n$.

From 1 and 2 we get

$$[T]_\beta [u_j]_\beta = \lambda_j [u_j]_\beta .$$

$$\forall j = 1, \ldots, n.$$

$$\Rightarrow [T(u_j)]_\beta = \lambda_j [u_j]_\beta .$$

$$\Rightarrow \quad T(u_j) = \lambda_j u_j$$

for $j = 1, \ldots, n$.

Clearly $\{u_1, \ldots, u_n\}$ is a basis of V.

$\therefore$ T is diagonalizable.

Review - March 11th and March 22nd

Proposition

Let V be a finite dimensional vector space. Let $\mathcal{B}, \mathcal{C}$ be bases for V. Let $T : V \to V$ be a linear transformation. Then the matrix of T with respect to $\mathcal{B}$ and $\mathcal{C}$ are similar to each other. If $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ is the change-of-basis matrix from $\mathcal{B}$ to $\mathcal{C}$, then

$$[T]_{\mathcal{B}} = \underset{\mathcal{C}\leftarrow\mathcal{B}}{P}^{-1} [T]_{\mathcal{C}} \underset{\mathcal{C}\leftarrow\mathcal{B}}{P}.$$

Theorem

Let A be an $n \times n$ matrix. Then A is diagonalizable if and only if there exists a basis of $\mathbb{R}^n$ consisting of eigenvectors of A.

$V \longrightarrow \mathbb{R}$

Definition

When $W = \mathbb{R}$, the set $L(V, W) = L(V, \mathbb{R})$ is called the *dual* of V, denoted by V^*. Each element of $L(V, \mathbb{R})$ is called a *linear functional*.

isomorphism. $V^* = L(V, \mathbb{R})$

$$V^* \cong M_{1\times n} \cong \mathbb{R}^n$$

(KA)

Theorem

Let V be a finite dimensional vector space and let $\langle ., . \rangle$ be an inner product on V. If $T \in V^*$, then there exists a unique vector $w_T \in V$ such that

$$T(v) = \langle v, w_T \rangle, \ \forall v \in V.$$

__Proposition__: Let $(V, \langle\,\rangle)$ be an inner product space. Let $w \in V$. The mapping

$$T(v) = \langle w, v \rangle, \forall v \in V$$

is a linear functional.

Proof: Let $v_1, v_2 \in V$.

$$T(v_1 + v_2) = \langle w, v_1 + v_2 \rangle$$

$$= \langle w, v_1 \rangle + \langle w, v_2 \rangle$$

$$= T(v_1) + T(v_2). \quad \text{— (1)}$$

Let $v \in V$, $c \in \mathbb{R}$.

$$T(cv) = \langle w, cv \rangle$$

$$= c\langle w, v \rangle$$

$$= c\,T(v)$$

From 1 & 2, $T \in L(V, \mathbb{R}) = V^*$ 2

Proof of Theorem $(\star\star)$:

By the rank-nullity theorem, $\dim \ker T$

$$= n - \dim \operatorname{range} T$$
$$\leq n - \dim(\mathbb{R}) = n-1$$

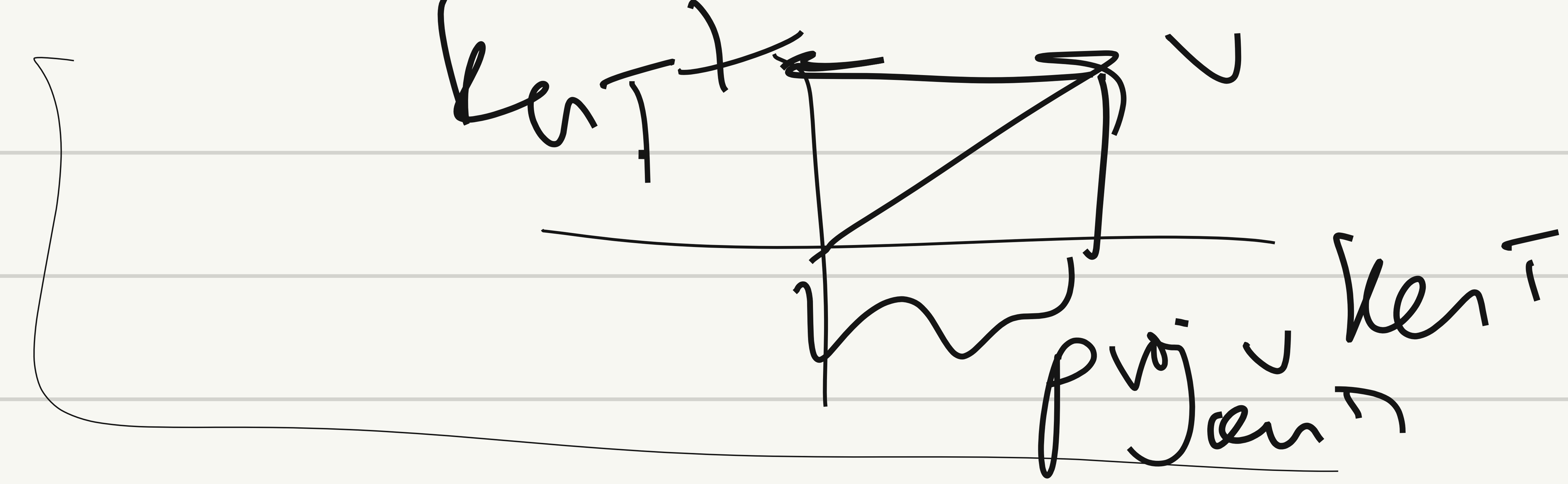
$$\therefore \quad \dim\left((\operatorname{Ker} T)^{\perp} \right) = n - (n-1)$$

$$= 1$$

Since $(\operatorname{Ker} T)^{\perp}$ is one-dimensional, it contains exactly two unit vectors.
Let a be a unit vector in $(\operatorname{Ker} T)^{\perp}$.

Define $w_T = \underbrace{T(a)}\, a$

Let $v \in V$.



$$v = \underbrace{\operatorname{proj}_{\ker T} v}_{\in \ker T} + \underbrace{\left(v - \operatorname{proj}_{\ker T} v\right)}$$

$$T(v) = \cancel{T\left(\operatorname{proj}_{\ker T} v\right)} + T\left(v - \operatorname{proj}_{\ker T} v\right)$$

$$= T\left(v - \operatorname{proj}_{\ker T} v\right)$$

Since $v - \operatorname{proj}_{\ker T} v \in (\ker T)^{\perp}$

$\Rightarrow v - \operatorname{proj}_{\ker T} v = \lambda u,$

for some $\lambda \in \mathbb{R}$.

$\Rightarrow T(v - \operatorname{proj}_{\ker T} v) = \lambda T(u).$

$$\Rightarrow \quad T(v) = \lambda T(u). \quad \text{— } (1)$$

$$\langle v, w_T \rangle = \langle v - \mathrm{proj}_{\ker T} v + \mathrm{proj}_{\ker T} v, w_T \rangle$$

$$= \langle v - \mathrm{proj}_{\ker T} v, w_T \rangle$$

$$= \langle \lambda u, w_T \rangle$$

$$= \langle \lambda u, T(u)u \rangle$$

$$= \lambda T(u) \underline{\langle u, u \rangle}$$

$$= \lambda T(u) \quad \text{--- } (2).$$

From (1) & (2), $T(v) = \langle v, w_T \rangle$.

Singular Values of an $m \times n$ Matrix

$\{v_1, \ldots, v_n\}$

Let $A \in M_{m \times n}(\mathbb{R})$.

Since $A^T A$ is symmetric, it is orthogonally diagonalizable. In other words, there exists an orthonormal basis of $\mathbb{R}^n$ consisting of eigenvectors of $A^T A$.

Claim: The eigenvalues of $A^T A$ are non-negative.

Idea behind proof: Each eigenvalue λ is the square of the length of Av, where v is a unit norm eigenvector corresponding to λ.

The length of Av has a special name - it's called a singular value of A.

Av_j

$$\|Av_j\|^2 = (Av_j)\cdot(Av_j)$$

$$= v_j^T A^T \cdot A v_j$$

$$= v_j^T \underbrace{(A^T A) v_j}$$

$$= \lambda_j v_j^T \cdot v_j = \lambda_j.$$

Definition

The singular values of A are the square roots of the eigenvalues of $A^T A$, denoted by $\sigma_1, \ldots, \sigma_n$, and they are arranged in decreasing order.

Singular Value Decomposition

Theorem

Let A be an $m \times n$ matrix with rank r. Then there exists an $m \times n$ matrix Σ of the form

$$\Sigma = \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix}$$

where D is an $r \times r$ diagonal matrix having as entries the first r singular values of A, $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$, and there exist an $m \times m$ orthogonal matrix U and an $n \times n$ orthogonal matrix V such that

$$A = U \Sigma V^T$$

Algorithm for Finding SVD

Step 1. Find an orthogonal diagonalization of $A^T A$.

Step 2. Arrange eigenvalues of $A^T A$ in decreasing order. The corresponding unit eigenvectors are the columns of V. The decreasing singular values are the entries of D in Σ.

Step 3. The first r columns of U are obtained by normalizing the vectors $A\mathbf{v}_1, \ldots, A\mathbf{v}_r$, where

$$V = [\mathbf{v}_1 \ldots \mathbf{v}_n]$$

and the remaining columns are obtained by extending to an orthonormal basis of $\mathbb{R}^m$.