Non-square diagonal matrices

Aim lecture: We apply the spectral theory of hermitian operators to look at linear maps $T : V \longrightarrow W$ which are not necessarily endomorphisms. This gives useful factorisations of non-square matrices.

In this lecture, V, W will denote fin dim inner product spaces.

Defn

A not necessarily square matrix $D = (d_{ij})_{ij} \in M_{mn}(\mathbb{C})$ is *diagonal* if $d_{ij} = 0$ whenever $i \neq j$. It's *diagonal entries* are $d_{11}, d_{22}, \ldots$.

E.g.

Basic Question Can you find co-ord systems $C_V : \mathbb{F}^n \longrightarrow V, C_W : \mathbb{F}^m \longrightarrow W$ so that the representing matrix $C_W^{-1} \circ T \circ C_V \in M_{mn}(\mathbb{F})$ is diagonal? What if you are only allowed orthonormal co-ord systems?

New version of isomorphism thm

Let $T : V \longrightarrow W$ be linear & pick vector space complements $V' \leq V$ to ker T and $W' \leq W$ to im T. We thus have natural isomorphisms between external & internal direct sums $\Phi_V : V' \oplus \ker T \longrightarrow V, \Phi_W : \text{im } T \oplus W' \longrightarrow W$.

Prop 1

The corresponding map on external direct sums $\Phi_W^{-1} \circ T \circ \Phi_V : V' \oplus \ker T \longrightarrow \text{im } T \oplus W'$ has matrix form

$$\begin{pmatrix} T|_{V'} & 0 \\ 0 & 0 \end{pmatrix}$$

where $T|_{V'} : V' \longrightarrow \text{im } T$ is an isomorphism.

Proof. Note that $T|_{V'}$ is an isomorphism by the isom thm of lecture 18. We verify the matrix form by considering $\mathbf{v}' \in V', \mathbf{v} \in \ker T$ and computing

$$\begin{pmatrix} T|_{V'} & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \mathbf{v}' \\ \mathbf{v} \end{pmatrix} =$$

$$\Phi_W^{-1} \circ T \circ \Phi_V \begin{pmatrix} \mathbf{v}' \\ \mathbf{v} \end{pmatrix} =$$

Diagonal forms

Continue notn from last slide. Pick co-ord systems $C_{V'} : \mathbb{F}^r \longrightarrow V', C_{\text{ker}} : \mathbb{F}^n \longrightarrow \ker T$. Hence $(C_{V'}\ \ C_{\text{ker}}) = \Phi_V \circ (C_{V'} \oplus C_{\text{ker}}) : \mathbb{F}^{r+n} \longrightarrow V$ is the corresponding co-ord system on V.

Note $C_{\text{im}} = T|_{V'} \circ C_{V'} : \mathbb{F}^r \longrightarrow V' \longrightarrow \text{im } T$ is a co-ord system for im T and pick a co-ord system $C_{W'} : \mathbb{F}^m \longrightarrow W'$ for W' so $(C_{\text{im}}\ \ C_{W'})$ is a co-ord system for W.

Cor

Wrt co-ord systems $(C_{V'}\ \ C_{\text{ker}}), (C_{\text{im}}\ \ C_{W'})$, T is represented by the $(r+m) \times (r+n)$-matrix $\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$ where $r = \text{rank } T$.

Proof. Note first $r = \dim \text{im } T = \text{rank } T$. We now compute

$$(C_{\text{im}}\ \ C_{W'})^{-1} \circ T \circ (C_{V'}\ \ C_{\text{ker}}) = \begin{pmatrix} C_{\text{im}} & 0 \\ 0 & C_{W'} \end{pmatrix}^{-1} \circ \Phi_W^{-1} \circ T \circ \Phi_V \circ \begin{pmatrix} C_{V'} & 0 \\ 0 & C_{\text{ker}} \end{pmatrix} =$$

Thought experiment

Q Suppose $A \in M_{mn}(\mathbb{C})$ can be factorised as $A = U_l D U_r^*$ with U_l, U_r unitary matrices of size m, n resp & $D = (d_{ij})_{ij}$ diagonal. How do you find U_l, U_r, D in terms of A?

Trick Eliminate U_l by considering $A^*A = U_r D^* U_l^* U_l D U_r^* = U_r D^* D U_r^*$
Oberservation Note $D^*D \in M_{nn}$ is square diagonal with (ex) diagonal entries $\overline{d_{ii}} d_{ii} \geq 0$. We have thus unitarily diagonalised A^*A.

Upshot

If indeed we can factorise $A = U_l D U_r^*$ as above then our observation suggests trying

1. the columns $\mathbf{u}_i$ of U_r are an orthonormal basis of e-vectors for A^*A.
2. the diagonal entries of D have square modulus the corresp e-values of A^*A.
3. obtain U_l by solving $AU_r = U_l D$ if possible, i.e. if $\mathbf{u}'_i$ is the i-th column of U_l then comparing i-th columns gives $A\mathbf{u}_i = d_{ii}\mathbf{u}'_i$.

Singular values

Let $T : V \longrightarrow W$ be linear as usual.

Prop-Defn

1. T^*T is an hermitian operator with e-values which are non-negative reals.
2. We may write the positive e-values as $\sigma_1^2, \ldots, \sigma_\rho^2$ for unique positive reals $\sigma_1, \ldots, \sigma_\rho$. These σ_i are called the *singular values of A*.
3. Let $E_{\sigma_i^2}$ be the σ_i^2-e-space of T^*T. Then $(\ker T)^\perp = \oplus_i E_{\sigma_i^2}$.

Proof. 1) Note T^*T is hermitian since $(T^*T)^* = T^*T^{**} = T^*T$. Let $\mathbf{v} \in V$ be an e-vector with e-value λ. Then

$$\lambda(\mathbf{v}|\mathbf{v}) = (\mathbf{v}|\lambda\mathbf{v}) = (\mathbf{v}|T^*T\mathbf{v}) = (T\mathbf{v}|T\mathbf{v}) \geq 0$$

so λ is non-negative real.

3) Note from Prop lecture 36, that $\ker T = \ker T^*T = E_0$. The result follows from the orthogonal direct sum decomposition of e-spaces for T^*T.

Example

E.g. What are the singular values of $A = \begin{pmatrix} 1 & 1 \\ 2 & 2 \\ 2 & 2 \end{pmatrix}$?

Singular value decomposition theorem

Let $A \in M_{mn}(\mathbb{C})$ and $\mathbf{u}_1, \ldots, \mathbf{u}_\rho$ be an o/n basis of e-vectors for $(\ker A)^\perp = \oplus_i E_{\sigma_i^2}$ & $\sigma_1, \ldots, \sigma_\rho$ be the corresponding singular values. Let $\mathbf{u}_{\rho+1}, \ldots, \mathbf{u}_n$ be an o/n basis $\ker A$ so $U_r = (\mathbf{u}_1 \ldots \mathbf{u}_n)$ is a unitary matrix defining a co-ord system adapted to the direct sum decomposition $\mathbb{C}^n = (\ker T)^\perp \oplus \ker T$ as in Prop 1.

Theorem (Singular decomposition)

We have the factorisation $A = U_l D U_r^*$ (called *singular value decomposition*) where

1. U_r is the unitary $n \times n$-matrix above.
2. $D = (d_{ij})_{ij} \in M_{mn}(\mathbb{R})$ is the diagonal matrix whose non-zero diagonal entries are $d_{ii} = \sigma_i, i = 1, \ldots, \rho$.
3. $U_l = (\mathbf{u}'_1 \ldots \mathbf{u}'_m)$ is unitary with columns $\mathbf{u}'_i = \sigma_i^{-1} A\mathbf{u}_i$ for $i = 1, \ldots, \rho$ & $\mathbf{u}'_{\rho+1}, \ldots, \mathbf{u}'_m$ an o/n basis for $(\text{im } A)^\perp$.

Rem There is choice in the possible unitary matrices U_l, U_r. Also U_l defines a co-ord system adapted to the direct sum decompn $\mathbb{C}^m = \text{im } A \oplus (\text{im } A)^\perp$ as in Prop 1.

Proof will follow

Example cont'd

E.g. Find a singular value decomposition for $A = \begin{pmatrix} 1 & 1 \\ 2 & 2 \\ 2 & 2 \end{pmatrix}$?

Sketch proof thm

The proof follows along the same lines as the proof of the corollary using Prop 1. That argument reduces us to showing

Lemma

1. $(\mathbf{u}'_1 \dots \mathbf{u}'_\rho) : \mathbb{F}^\rho \longrightarrow \text{im } A$ defines an o/n co-ord system.
2. The matrix representing the associated linear map $T_A|_{(\ker A)^\perp} : (\ker A)^\perp \longrightarrow \text{im } A$ wrt co-ord systems $C = (\mathbf{u}_1 \dots \mathbf{u}_\rho)$, $C' = (\mathbf{u}'_1 \dots \mathbf{u}'_\rho)$ is the square diagonal matrix $(\sigma_1) \oplus \dots \oplus (\sigma_\rho)$.

Proof. 1) Note $\{\mathbf{u}'_1 \dots \mathbf{u}'_\rho\}$ spans im A since

$\text{im } A = \text{Span}(A\mathbf{u}_1, \dots, A\mathbf{u}_n) = \text{Span}(A\mathbf{u}_1, \dots, A\mathbf{u}_\rho) = \text{Span}(\mathbf{u}'_1, \dots, \mathbf{u}'_\rho)$

To check they form an o/n basis, we need now only check orthogonality

$$(\mathbf{u}'_i|\mathbf{u}'_j) = (\sigma_i^{-1} A\mathbf{u}_i | \sigma_j^{-1} A\mathbf{u}_j) = (\sigma_i \sigma_j)^{-1}(A^* A\mathbf{u}_i|\mathbf{u}_j)$$
$$= (\sigma_i \sigma_j)^{-1}(\sigma_i^2 \mathbf{u}_i|\mathbf{u}_j) = \sigma_i \sigma_j^{-1}(\mathbf{u}_i|\mathbf{u}_j)$$

which is 0 if $i \neq j$ and 1 if $i = j$. 1) is now proved.

2) We calculate the i-th column of the representing matrix to be

$$(C')^{-1} \circ T_A|_{(\ker A)^\perp} \circ C\mathbf{e}_i = (C')^{-1} A\mathbf{u}_i = (C')^{-1} \sigma_i \mathbf{u}'_i = \sigma_i \mathbf{e}_i$$