B24 July 30 Lec 2 Notes

Theorem:

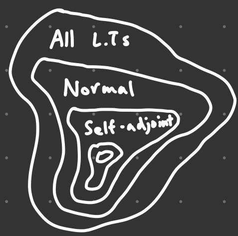
Let X be a $\mathbb{R}$-IPS, and $A: X \to X$ be a L.T. Assume A has $\dim(X)$ many eigenvalues (counting multiplicities). Then there exists an orthonormal basis $u_1, \ldots, u_n$ for X s.t.

$$[A]_{u_1,\ldots,u_n}^{u_1,\ldots,u_n}$$

is upper triangular.

Proof: Similar to the $\mathbb{C}$ case.

Definition:



If X is an IPS and $A: X \to X$, we say A is Hermitan or self-adjoint if $A^* = A$.

Theorem:

Let $A: X \to X$ be self-adjoint. Then A has $\dim(X)$ many eigenvalues (counting multiplicity), all eigenvalues are real, and there exists an orthonormal basis for X consisting of eigenvectors for A.

Proof:

We will prove any eigenvalue of A must be real. Let $\lambda \in \mathbb{F}$ be an eigenvalue with x a corresponding eigenvector.

Then:

$$\langle Ax, x \rangle = \langle \lambda x, x \rangle = \lambda \langle x, x \rangle$$

$$\langle Ax, x \rangle = \langle x, A^* x \rangle = \langle x, Ax \rangle = \langle x, \lambda x \rangle = \bar{\lambda} \langle x, x \rangle$$

So $\lambda \langle x, x \rangle = \bar{\lambda} \langle x, x \rangle \Rightarrow \lambda = \bar{\lambda}$

i.e. $\lambda \in \mathbb{R}$

In order to prove that A has $\dim(X)$ many eigenvalues, first note we already know this when $\mathbb{F} = \mathbb{C}$.

Proof (continued...):

When $\mathbb{F} = \mathbb{R}$, we consider the "complexification" of X, i.e.

$$X_{\mathbb{C}} := \{ x + iy \mid x, y \in X \}$$

with $(x_1 + iy_1) + (x_2 + iy_2) := (x_1 + x_2) + i(y_1 + y_2)$

and $\alpha(x + iy) := \alpha x + i\alpha y, \ \forall \alpha \in \mathbb{C}$.

Then $X_{\mathbb{C}}$ is a $\mathbb{C}$ v.s. and can also be given an inner product.

Define $A_{\mathbb{C}} : X_{\mathbb{C}} \to X_{\mathbb{C}}$ by $A_{\mathbb{C}}(x + iy) := A_{\mathbb{C}}(x) + i A_{\mathbb{C}}(y)$

Note that $\dim(X_{\mathbb{C}}) = \dim(X)$, and so $A_{\mathbb{C}}$ has $\dim(X)$ many eigenvalues, and $A_{\mathbb{C}}$ is self-adjoint (since A is) and so every eigenvalue of $A_{\mathbb{C}}$ is real and hence every eigenvalue of A is real (and there are $\dim(X)$ many).

Why $\dim(X_{\mathbb{C}}) = \dim(X)$?

$X = \mathbb{R}$, then $X_{\mathbb{C}} = \mathbb{C}$. If $x_1, \dots, x_n$ is a basis for X, then $x_1, \dots, x_n$ is a basis for $X_{\mathbb{C}}$.

$$\begin{aligned} x + iy &= (\alpha_1 x_1 + \dots + \alpha_n x_n) + i(\beta_1 x_1 + \dots + \beta_n x_n) \\ &= x_1 \underbrace{(\alpha_1 + \beta_1 i)}_{\mathbb{C}} + \dots + x_n \underbrace{(\alpha_n + \beta_n i)}_{\mathbb{C}} \end{aligned}$$

Lastly, we must prove that there exists an orthonormal basis for X consisting of eigenvectors for A.

By previous two theorems, there exists a basis $u_1, \dots, u_n$ so that

$$[A]^{u_1, \dots, u_n}_{u_1, \dots, u_n} = \begin{bmatrix} A_{11} & \cdots & A_{1n} \\ & \ddots & \vdots \\ 0 & & A_{nn} \end{bmatrix}$$

and

$$\left([A]^{u_1, \dots, u_n}_{u_1, \dots, u_n}\right)^* = [A^*]^{u_1, \dots, u_n}_{u_1, \dots, u_n}$$

$$= \begin{bmatrix} \overline{A_{11}} & & 0 \\ \vdots & \ddots & \\ \overline{A_{1n}} & \cdots & \overline{A_{nn}} \end{bmatrix}$$

Since $A = A^*$, we have

$$\begin{bmatrix} A_{11} & \cdots & A_{1n} \\ & \ddots & \vdots \\ 0 & & A_{nn} \end{bmatrix} = \begin{bmatrix} \overline{A_{11}} & & 0 \\ \vdots & \ddots & \\ \overline{A_{1n}} & \cdots & \overline{A_{nn}} \end{bmatrix}$$

So $A_{jk} = 0$ for $j \neq k$, and $A_{jj} \in \mathbb{R}$ for $1 \leq j \leq n$.

Proof (Continued...):

Since $A(u_i) = A_{ii} u_i$, for $1 \leq i \leq n$, it follows that each u_i is an eigenvector, and so $u_1, \dots, u_n$ are an orthonormal basis for X consisting of eigenvectors for A. $\boxtimes$

Corollary:

Let A be a self-adjoint matrix. Then

$$A = UDU^*$$

for U a unitary matrix, and D a diagonal matrix.

Proof: Use previous theorem and the one below.

Proposition:

A $n \times n$ matrix A is unitarily equivalent to a diagonal matrix iff there is an orthonormal basis for $\mathbb{F}^n$ consisting of eigenvectors of A.

Definition:

A L.T. $N : X \to X$ (where X is IPS) is called normal if

$$N^* N = N N^*$$

Remark: Self-adjoint $\Rightarrow$ Normal, since $A^*A = AA = AA^*$

Proposition:

If $A : X \to X$ is a L.T. and has an orthonormal basis of eigenvectors, then A is normal.

Proof:

Let $u_1, \dots, u_n$ be an orthonormal basis of eigenvectors, then:

$$[AA^*]_{u_1,\dots,u_n}^{u_1,\dots,u_n} = \underbrace{[A]_{u_1,\dots,u_n}^{u_1,\dots,u_n}}_{\text{diagonal}} \underbrace{[A^*]_{u_1,\dots,u_n}^{u_1,\dots,u_n}}_{\text{diagonal}}$$

$$= [A^*]_{u_1,\dots,u_n}^{u_1,\dots,u_n} [A]_{u_1,\dots,u_n}^{u_1,\dots,u_n}$$

$$= [A^*A]_{u_1,\dots,u_n}^{u_1,\dots,u_n}$$

$$\Rightarrow AA^* = A^*A$$

Theorem:

If X is a $\mathbb{C}$-IPS, and $N: X \to X$ is normal, then there is an orthonormal basis for X consisting of eigenvectors of N.

Proof:

We know from the $\mathbb{C}$ case of $\star$ that there exists an orthogonal basis $u_1, \dots, u_n$ for X s.t.

$$[N]_{u_1,\dots,u_n}^{u_1,\dots,u_n} \text{ is upper triangular}$$

We will abbreviate $[N] = [N]_{u_1,\dots,u_n}^{u_1,\dots,u_n}$.

Since $NN^* = N^*N$, we have:

$$\begin{aligned}
[N][N]^* &= [N][N^*] \\
&= [NN^*] \\
&= [N^*N] \\
&= [N^*][N] \\
&= [N]^*[N]
\end{aligned}$$

Lemma: (IH) here

If B is an upper triangular $n \times n$ matrix, and $BB^* = B^*B$, then B is diagonal.

$[N^*] = [N]^*$ since $T: \mathbb{R}^2 \to \mathbb{R}^2$, $(x,y) \mapsto (x+y, y)$, $[T] = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$, $[T]^* = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$

$T^*: \mathbb{R}^2 \to \mathbb{R}^2$, $(x,y) \mapsto (x+y, y)$, $[T^*] = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$

i.e. $[N][N]^* = [N]^*[N]$ $\triangle$

We will show that if $\triangle$ holds, then $[N]$ must be diagonal

We will induct on $\dim [N]$ (the dimension of the matrix).

If $\dim [N] = 1$, (base case), the conclusion is trivial (any 1×1 matrix is diagonal).

Assume (IH) the result holds when $\dim [N] = n$ and suppose $\dim [N] = n+1$

So

$$[N] = \left[\begin{array}{c|ccc} a_{11} & a_{12} & \cdots & a_{1,n+1} \\ \hline 0 & & & \\ \vdots & & \underbrace{N_1}_{\substack{n \times n \text{ upper} \\ \text{triangular} \\ \text{matrix}}} & \\ 0 & & & \end{array}\right]$$

and

$$[N][N]^* = \left[\begin{array}{c|ccc} a_{11} & a_{12} & \cdots & a_{1,n+1} \\ \hline 0 & & & \\ \vdots & & N_1 & \\ 0 & & & \end{array}\right] \left[\begin{array}{c|ccc} \overline{a_{11}} & 0 & \cdots & 0 \\ \hline \overline{a_{12}} & & & \\ \vdots & & N_1^* & \\ \overline{a_{1,n+1}} & & & \end{array}\right]$$

Proof (continued...):

$$= \left[\begin{array}{c|ccc} \sum_{i=1}^{n} |a_{1i}|^2 & * & \cdots & * \\ \hline * & & & \\ \vdots & & N_1 N_1^* & \\ * & & & \end{array}\right]$$

Consider

$$[N]^*[N] = \left[\begin{array}{c|ccc} \overline{a_{11}} & 0 & \cdots & 0 \\ \hline \overline{a_{12}} & & & \\ \vdots & & N_1^* & \\ \overline{a_{1,n+1}} & & & \end{array}\right] \left[\begin{array}{c|ccc} a_{11} & a_{12} & \cdots & a_{1,n+1} \\ \hline 0 & & & \\ \vdots & & N_1 & \\ 0 & & & \end{array}\right]$$

$$= \left[\begin{array}{c|ccc} |a_{11}|^2 & * & \cdots & * \\ \hline * & & & \\ \vdots & & N_1^* N_1 & \\ * & & & \end{array}\right]$$

So $\sum_{i=1}^{n} |a_{1i}|^2 = |a_{11}|^2 + \dots + |a_{nn}|^2 = |a_{11}|^2$

$\Rightarrow \; a_{22} = \dots = a_{nn} = 0$

So

$$[N] = \left[\begin{array}{c|ccc} a_{11} & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & N_1 & \\ 0 & & & \end{array}\right]$$

and $N_1 N_1^* = N_1^* N_1 \overset{\text{I.H.}}{\Rightarrow} N_1$ is diagonal

So $[N]_{u_1,\dots,u_n}^{u_1,\dots,u_n}$ is diagonal, hence $u_1,\dots,u_n$ is an orthonormal basis for X consisting of eigenvectors for N $\boxtimes$

Proposition:

A L.T. $N : X \to X$ is normal iff

$$\|N^* x\| = \|Nx\| \quad , \forall x \in X$$

Proof ($\Rightarrow$):

Assume $N : X \to X$ is normal. Then:

$$
\begin{aligned}
\|N^*x\|^2 &= \langle N^*x, N^*x \rangle \\
&= \langle NN^*x, x \rangle \\
&= \langle N^*Nx, x \rangle \\
&= \langle Nx, Nx \rangle \\
&= \|Nx\|^2
\end{aligned}
$$

$\Rightarrow \|N^*x\| = \|Nx\|, \ \forall x \in X$

Proof ($\Leftarrow$):

Assume $\|N^*x\| = \|Nx\|, \ \forall x \in X$

Let $x, y \in X$

We will show

$$\langle N^*Nx, y \rangle = \langle NN^*x, y \rangle$$

which implies $N^*N = NN^*$

Indeed

$$
\begin{aligned}
\langle N^*Nx, y \rangle = \langle Nx, Ny \rangle
&= \frac{1}{4} \sum_{\alpha = \pm 1, \pm i} \alpha \|Nx + \alpha Ny\|^2 && \text{Polarization identity (assume } X \text{ is a } \mathbb{C}\text{-IPS)} \\
&= \frac{1}{4} \sum_{\alpha = \pm 1, \pm i} \alpha \|N(x + \alpha y)\|^2 && \text{By linearity} \\
&= \frac{1}{4} \sum_{\alpha = \pm 1, \pm i} \alpha \|N^*(x + \alpha y)\|^2 \\
&= \frac{1}{4} \sum_{\alpha = \pm 1, \pm i} \alpha \|N^*x + \alpha N^*y\|^2 \\
&= \langle N^*x, N^*y \rangle \\
&= \langle NN^*x, y \rangle \quad \blacksquare
\end{aligned}
$$

Definition:

A self-adjoint L.T. $A: X \to X$ is called positive definite if

$$\langle Ax, x\rangle > 0, \ \forall x \in X,$$

and $A: X \to X$ is called positive semidefinite if

$$\langle Ax, x\rangle \geq 0, \ \forall x \in X$$