Chapter 3: Spectral Theory in Inner Prod. Spcs.

Note Title 5/27/2013

(I) Throughout this chapter, $E_{\mathbb{K}}$ will denote a finite-dimensional inner product space E over the field $\mathbb{K}$, and

$$\mathcal{L}(E) \stackrel{\text{def.}}{=} \mathcal{L}(E,E) = \mathrm{Hom}(E,E) = \mathrm{End}(E)$$

is the algebra of $\mathbb{K}$-linear operators $E \to E$, where for any $f, g \in \mathcal{L}(E)$ and $a \in \mathbb{K}$ one has:

- $(f + a g)(x) \stackrel{\text{def.}}{=} f(x) + a \cdot g(x)$ (linear structure)
- $(fg)(x) \stackrel{\text{def.}}{=} f(g(x))$ (multiplicative structure)

...which then produces $f(g_1+g_2) = fg_1 + fg_2$ and analogously on the left!

Recall: Given bases $\alpha = (e_1, \ldots, e_n)$, $\beta = (f_1, \ldots, f_n)$ of E and $f \in \mathcal{L}(E)$:

- $[x]_\alpha \stackrel{\text{def.}}{=} [e^1(x), \ldots, e^n(x)]^T$ for all $x \in E$
- $\left[\begin{smallmatrix}\alpha\\ \beta\end{smallmatrix}\right] \in \mathbb{K}^{n \times n}$ satisfying $[x]_\beta = \left[\begin{smallmatrix}\alpha\\ \beta\end{smallmatrix}\right][x]_\alpha$ for all $x \in E$
- $[f]_\alpha \in \mathbb{K}^{n\times n}$ s.t. $[f(x)]_\alpha = [f]_\alpha [x]_\alpha$ for all $x \in E$
- incidentally, we recall:

$$[f]_\alpha = \left[\begin{smallmatrix}\beta\\ \alpha\end{smallmatrix}\right][f]_\beta\left[\begin{smallmatrix}\alpha\\ \beta\end{smallmatrix}\right] = \left[\begin{smallmatrix}\beta\\ \alpha\end{smallmatrix}\right][f]_\beta\left[\begin{smallmatrix}\beta\\ \alpha\end{smallmatrix}\right]^{-1}$$

- In particular, if $E = \mathbb{K}^n$ and $f(x) = Mx$ for $M \in \mathbb{K}^{n\times n}$, letting α be the standard basis and $P = [f_1 | \cdots | f_n]$ we have:

$$P = \left[\begin{smallmatrix}\beta\\ \alpha\end{smallmatrix}\right], \quad M = P[f]_\beta P^{-1} \quad \left(\Leftrightarrow [f]_\beta = P^{-1} M P\right)$$

Defn: two matrices $M, N \in M_n(\mathbb{K})$ are similar over $\mathbb{R}$ (or $\mathbb{C}$) if there exists an ~~invertible~~ real (resp. complex) -valued invertible matrix P such that $M = P^{-1}NP$. We write $M \sim N$ in this case.

It is easy to check that similarity (in either form) is an equivalence relation on $M_n(\mathbb{K})$.

Perhaps it is a bit early but note an interesting property of orthonormal bases of $\mathbb{K}^n$ in this context ($\mathbb{K}$ is $\mathbb{R}$ or $\mathbb{C}$!)

Recall that when $\alpha = (v_1, \dots, v_n)$ and $\beta = (u_1, \dots, u_n)$ are bases of $\mathbb{K}^n$ then
$$\begin{bmatrix}\alpha\\ \varepsilon\end{bmatrix} = \underbrace{\left[v_1 \middle| \cdots \middle| v_n\right]}_{V}, \quad \begin{bmatrix}\beta\\ \varepsilon\end{bmatrix} = \underbrace{\left[u_1 \middle| \cdots \middle| u_n\right]}_{U}$$

and so $\begin{bmatrix}\alpha\\ \beta\end{bmatrix} = U^{-1}V$.

Now, if α is orthonormal w.r.t. the std. inner product

$\Leftrightarrow \quad v_i^* v_j = \delta_{ij}$ for all $i, j \in \{1, \dots, n\}$

$\Leftrightarrow \quad V^*V = I \quad \Leftrightarrow \quad V$ is orthogonal/unitary.

Defn: Let E be a FDIPS with basis $\alpha = (v_1, \dots, v_n)$. Then the Gram matrix of the inner product w.r.t. α is the matrix
$$(G_\alpha)_{ij} \overset{\text{def.}}{=} \langle v_j, v_i \rangle.$$

We observe, for all $x, y \in E$: $\langle x, y \rangle = \sum_{i,j=1}^{n} x_i \bar{y}_j \langle v_i, v_j \rangle$,

which means: $\langle x, y \rangle \equiv [y]_\alpha^* G_\alpha [x]_\alpha$

In particular, it follows that for any bases α, β of E,

$$G_\beta = \left[{}^{\beta}_{\alpha}\right]^* G_\alpha \left[{}^{\beta}_{\alpha}\right]$$

And so we have $G_\beta = \left[{}^{\beta}_{\alpha}\right]^* \left[{}^{\beta}_{\alpha}\right]$ whenever α is an orthonormal basis.

Since E has an orthonormal basis we conclude:

Thm: Let E be a FDIPS over $\mathbb{K}$ and let β be a basis of E. Then there exists an invertible matrix $P \in M_n(\mathbb{K})$ such that $G_\beta = P^* P$.

Example: Consider $E = \mathcal{P}_m$, the space of polynomials of degree $\leq m$ with coefficients in $\mathbb{R}$ and

$$\langle f, g \rangle \stackrel{\text{def}}{=} \int_0^1 f(x) g(x)\, dx.$$

Then the Gram matrix of the std basis $\mathcal{E} = (1, x, \ldots, x^m)$ is

$$G_\mathcal{E} = \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \cdots & \frac{1}{m+1} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \cdots & \frac{1}{m+2} \\ \vdots & & \ddots & & \vdots \\ \frac{1}{m+1} & \cdots & & & \frac{1}{2m+1} \end{bmatrix}$$

Now, if we apply Gram-Schmidt to produce an ONB α of E, then $P = \left[{}^{\mathcal{E}}_{\alpha}\right]$ will produce a decomposition of $G_\mathcal{E}$ in the form $P^* P$. ∎

Try doing this for, say, $m = 4$.

II Enough about the tools, let's deal with the problems at hand:

__Recall:__ An operator $f \in \mathcal{L}(E)$ is triangulable/diagonalizable, if there exists a basis α of E such that $[f]_\alpha$ is an upper-triangular/diagonal matrix

__Defn:__ We say that an operator $f \in \mathcal{L}(E)$ is orthogonally (or unitarily) triangulable/diagonalizable if there is an __orthonormal__ basis α of E s.t. $[f]_\alpha$ is upper-triangular/diagonal matrix.

__What does it mean to be triangulable?__

$f: E \to E$ is triangulable $\iff$ there is a strictly increasing sequence

$$\{0\} < W_1 < \dots < W_{\dim E} = E$$

of subspaces such that

$$f(W_i) \subseteq W_i \text{ for all } i.$$

$$\alpha = (e_1, \dots, e_{k+l})$$

$$\overbrace{\begin{bmatrix} k\times k & * \\ 0 & l\times l \end{bmatrix}}^{[f]_\alpha} \begin{bmatrix} * \\ \hline 0 \end{bmatrix} = \begin{bmatrix} * \\ \hline 0 \end{bmatrix}$$

$$f(\mathrm{Span}(e_1, \dots, e_k)) \subseteq \mathrm{Span}(e_1, \dots, e_k)$$

__Defn:__ Let E be a vec. spc. and $f \in \mathcal{L}(E)$. We say that a subspace $W < E$ is f-invariant if $f(W) \subseteq W$.

__Defn:__ An increasing sequence $\{0\} = W_0 < W_1 < \dots < W_{\dim E} = E$ of subspaces in E is called a __flag__. A flag of f-invariant subspaces is called an f-invariant flag.

What does it mean to be diagonalizable?

$[f]_\alpha = \mathrm{diag}(d_1, \dots, d_n) \iff f(e_i) = d_i e_i \iff \alpha$ is a basis of eigenvectors of f.

$(\alpha = (e_1, \dots, e_n))$

Defns: • λ is an eigenvalue of $f \overset{\text{def.}}{\iff} \lambda \in \mathrm{Spec}(f)$

$\overset{\text{def.}}{\iff} \lambda I - f$ is singular

• $E_\lambda(f) \overset{\text{def}}{=} \mathrm{Ker}(I - \lambda f)$ the λ-th eigenspace of f

• $0 \neq v \in E_\lambda(f)$ is called an eigenvector of f

• $\det(tI - f) \overset{\text{def}}{=} \chi_f(t)$ is the characteristic poly. of f

Thus: $\lambda \in \mathrm{Spec}(f) \iff \chi_f(\lambda) = 0$

HOWEVER, THIS REALLY DEPENDS ON THE FIELD FROM WHICH λ IS DRAWN. FOR US, THIS IS ALWAYS $\mathbb{C}$, BUT THEN EIGENVECTORS FOR EIGENVALUES MAY ONLY APPEAR IN $\mathbb{C} \otimes E$!

Also: f is diagonalizable $\iff E = \bigoplus_{\lambda \in \mathbb{K}} E_\lambda(f)$

$\iff \dim E = \sum_{\lambda \in \mathbb{K}} \dim \mathrm{Ker}(\lambda I - f)$

AS IT IS KNOWN THAT $E_\lambda(f) \cap \sum_{\mu \neq \lambda} E_\mu(f) = \{0\}$ FOR ALL $f \in \mathcal{L}(E)$.

One wonderful property of $\mathbb{C}$ is:

Theorem (Schur): Let $f \in \mathcal{L}(E)$, where E is a finite-dimensional vector space over $\mathbb{C}$. Then f is triangulable. Moreover, if E is an inner product space, then f is unitarily triangulable.

Pf: We note that orthogonal triangulability is directly obtained from triangulability via Gram-Schmidt process.

Suppose the first assertion of the theorem is false and find a counter-example $f \in \mathcal{L}(E)$ with $\dim E$ minimal.

By the fundamental theorem of Algebra, $\chi_f(t)$ has a root $\lambda_0 \in \mathbb{C}$ and hence there is an eigenvector $z \in E$: $f(x) = \lambda_0 x$.

We look at $\bar{f}: \begin{cases} E/_{\mathbb{C}z} \longrightarrow E/_{\mathbb{C}z} \\ x + \mathbb{C}z \longmapsto f(x) + \mathbb{C}z \end{cases}$

This is well-defined, as $f(x + az) = f(x) + a\lambda_0 z$ for all $x \in E, a \in \mathbb{C}$.

Now, $\dim E/_{\mathbb{C}z} < \dim E \Rightarrow \bar{f}$ is triangulable.

$\Rightarrow$ there is an $\bar{f}$-invariant flag $o < \overline{W}_1 < \dots < \overline{W}_d$ in $E/_{\mathbb{C}z}$

$\Rightarrow$ there is a f-invariant flag $o < \mathbb{C}z < W_1 < \dots < W_d$ in E where $W_i = \pi^{-1}(\overline{W}_i)$ and $\pi: \begin{cases} E \to E/_{\mathbb{C}z} \\ x \mapsto x + \mathbb{C}z. \end{cases}$

$\boxtimes$

Some consequences of Schur's theorem:

(1) An operator $f \in E_{\mathbb{R}} \to E_{\mathbb{R}}$ represented by a matrix $A \in M_n(\mathbb{R})$ in an (orthonormal) basis α of E over $\mathbb{R}$ extends to an operator see (2)

$$\mathbb{C}\otimes f : \mathbb{C}\otimes E \to \mathbb{C}\otimes E$$

represented by A in the corresponding basis $\mathbb{C}\otimes\alpha$ of $\mathbb{C}\otimes E$

$\Longrightarrow$ $\mathbb{C}\otimes f$ is triangulable over $\mathbb{C}$

$\Longrightarrow$ $A \sim$ a triangular matrix B over $\mathbb{C}$

$\Longrightarrow$ $\mathrm{Spec}(A) = \{\text{diagonal entries of } B\}$

$\Longrightarrow$ $\begin{cases} tr(f) = tr(A) = \text{sum of eigenvalues of } f \text{ over } \mathbb{C} \\ \qquad = -(\text{coefficient of } t^{n-1} \text{ in } \chi_f(t)) \\ det(f) = det(A) = (\text{product of eigenvalues of } f \text{ over } \mathbb{C}) \\ \qquad = (-1)^n (\text{free coefficient of } \chi_f(t)) \end{cases}$

(2) when $A = [f]_\alpha$, α is orthonormal, we have $A^* = [f^*]_\alpha$ but moreover:

$$(\mathbb{C}\otimes f)^* = \mathbb{C}\otimes f^*$$ and both are represented by A^* with respect to $\mathbb{C}\otimes\alpha$.

Thus, WE MAY REASON ABOUT f, f^* AND THEIR REPRESENTING MATRICES ALWAYS ASSUMING ANY COMPLEX EIGENVALUE HAS AN EIGENVECTOR.

③ A word of CAUTION: If $A=[f]_\alpha$ is upper triangular, then A^* is lower-triangular

$\Longrightarrow$ writing $\alpha=(v_1,\dots,v_n)$ and $W_k=\mathrm{Span}(v_1,\dots,v_k)$, $(W_0=\{0\})$

results in a f-invariant flag $W_1 \subsetneq W_2 \subsetneq \dots \subsetneq W_n = E$

but then this flag is not necessarily invariant under f^*, but

$W'_{n-k}=\mathrm{Span}(v_{k+1},\dots,v_n)$ is an f^*-invariant flag.

To understand what is going on:

Lemma: Suppose $W<E$ is f-invariant. Then $W^\perp$ is f^*-invariant.

Proof: Suppose that $f(x)\in W$ for all $x\in W$.

Take $y\in W^\perp$ and we need to show that $f^*(y)\in W^\perp$. For this we need to verify that $x\perp f^*(y)$ for all $x\in W$:

$$\langle x, f^*(y)\rangle = \langle f(x), y\rangle = 0 .$$

since $f(x)\in W$
and $y\in W^\perp$

Examples of invariant subspaces: Fix $f\in\mathcal{L}(E)$ and start looking:

(1) $\mathrm{Im} f$: Indeed, $\mathrm{Im} f\subseteq E \Rightarrow f(\mathrm{Im} f)\subseteq f(E)=\mathrm{Im} f$.

(2) $\mathrm{Ker} f$: $f(\mathrm{Ker} f)=\{0\}\subseteq \mathrm{Ker} f$

(3) $\mathrm{Ker}\, p(f)$ for $p(t)\in\mathbb{K}[t]$: $p(f)(x)=0 \Rightarrow f(p(f)(x))=0 \Rightarrow p(f)(f(x))=0$.

(4) $E_\lambda(f) = \operatorname{Ker}(\lambda I - f) = \operatorname{Ker} p(f)$, $f(t) = \lambda - t$.

(5) $g \in \mathcal{L}(E)$, $gf = fg \Rightarrow E_\lambda(f)$ is g-invariant for all $\lambda \in \mathbb{C}$.

Let's check this: take $x \in E_\lambda(f) \Rightarrow f(x) = \lambda x$

and compute $f(g(x)) = g(f(x)) = g(\lambda x) = \lambda g(x) \Rightarrow g(x) \in E_\lambda(f)$ ☐

The following is a major goal for us:

Theorem: Suppose $f \in \mathcal{L}(E_{\mathbb{C}})$ is normal, that is: $f^* f = f f^*$. Then f is orthogonally diagonalizable.

Let us do some preparatory work.

Lemma: If f is normal, then $\|fx\| = \|f^*x\|$ for all $x \in E$.

Proof:

$$\|fx\|^2 = \langle fx, fx\rangle = \langle x, f^*fx\rangle = \langle x, ff^*x\rangle$$

$$= \langle f^*x, f^*x\rangle = \|f^*x\|^2. \quad ☐$$

Corollary: If f is normal, then $E_\lambda(f) = E_{\bar\lambda}(f^*)$ for all $\lambda \in \mathbb{C}$.

Proof: we note that $g = \lambda I - f$ is also normal, as $g^* = \bar\lambda I - f^*$.

Thus

$$\begin{array}{ccccccc} x \in \operatorname{Ker} g & \Leftrightarrow & \|gx\| = 0 & \Leftrightarrow & \|g^*x\| = 0 & \Leftrightarrow & x \in \operatorname{Ker} g^* \\ \Updownarrow & & & & & & \Updownarrow \\ x \in E_\lambda(f) & & & & & & x \in E_{\bar\lambda}(f^*) \end{array}$$

☐

Proof of the theorem: We proceed by induction, having noted that the theorem holds true when $\dim(E) = 0, 1$.

Suppose $d \geq 1$ is such that every normal operator on a FDIPS of dimension $\leq d$ has an orthonormal basis of eigenvectors, and suppose $f \in \mathcal{L}(E)$ is normal and $\dim E = d+1$. We need to produce an ONB for E consisting of eigenvectors of f. Remember: we are working over $\mathbb{C}$!

Then there is $\lambda \in \mathbb{C}$ s.t. $\{0\} \neq E_\lambda(f) = E_{\bar{\lambda}}(f^*) := W$.

Now, W is both f & f^*-invariant $\Rightarrow W^\perp$ is both f^* & f-invariant.

$\dim W \geq 1 \Rightarrow \dim W^\perp < \dim E \Rightarrow f$ is a normal operator on $W^\perp$ and hence there is an ONB α of $W^\perp$ consisting of f-eigenvectors.

Lemma: If W is both f & f^* invariant, then $(f|_W)^* = f^*|_W$

$\Rightarrow$ Pick an ONB α_λ of $E_\lambda(f)$, then $\alpha_\lambda \cup \alpha$ is the basis we seek. ☒

Corollary (Spectral Theorem for Normal Operators): Let f be a linear operator on a finite-dimensional unitary space E. Then f is normal if and only if it can be written as a linear combination

$$(\star) \quad f = \sum_{\lambda \in \mathbb{C}} \lambda \cdot P_\lambda$$

SPECTRAL DECOMPOSITION

where (1) $P_\lambda^2 = P_\lambda$ and $P_\lambda^* = P_\lambda$ (the P_λ are orthogonal projections)
(2) $P_\lambda P_\mu = 0$ if $\lambda \neq \mu$
(3) all but finitely many of the P_λ are zero, and $\sum_{\lambda \in \mathbb{C}} P_\lambda = I$.

Moreover, this decomposition is unique, and each P_λ is a polynomial in f. ☒