EXAMENSARBETEN I MATEMATIK

MATEMATISKA INSTITUTIONEN, STOCKHOLMS UNIVERSITET

Operator theory in finite-dimensional vector spaces

av

Kharema Ebshesh

2008 - No 12

Operator theory in finite-dimensional vector spaces

Kharema Ebshesh

Examensarbete i matematik 15 högskolepoäng, påbyggnadskurs

Handledare: Andrzej Szulkin och Yishao Zhou

2008

Abstract

The theory of linear operators is an extensive area. This thesis is about the linear operators in finite dimensional vector spaces. We study the symmetric, unitary, isometric, and normal operators, and orthogonal projection in the unitary space, the eigenvalue problem and the resolvent. We give a proof of the minimax principle in the end.

Acknowledgements

I have the pleasure to thank Prof. Andrzej Szulkin and Prof. Yishao Zhou for their help, I enjoyed my work with them, I actually learned a lot during the preparation of this thesis.
I would like to thank all the teachers that I had in the Maths Department.
I would like to thank Prof. Mikael Passare for inviting me to Sweden.
Finally I will not forget my teachers in Libya who supported me during my studies.
/Kharema

Contents

Introduction **4**

1 Operators in vector spaces **5**

1.1 Vector spaces and adjoint vector spaces 5
1.2 Linear operators 12
1.3 Projections 16
1.4 The adjoint operator 17
1.5 The eigenvalue problem 19
1.6 The resolvent 22

2 Operators in unitary spaces **26**

2.1 Unitary spaces 26
2.2 Symmetric operators 29
2.3 Unitary, isometric and normal operators 32
2.4 Orthogonal projections 33
2.5 The eigenvalue problem 35
2.6 The minimax principle 36

References **38**

Introduction

This report contains two sections. In Section 1 we study linear operators in finite dimensional vector spaces, where projection, the adjoint operators are introduced. In particular, we study the eigenvalue problem and some properties of the resolvent. In Section 2, different operators in unitary spaces, such as symmetric, unitary, isometric and normal operators are considered. We study the eigenvalue problems for these operators. Finally we prove the minimax principle for eigenvalues.

The results in this report are primarily taken from Chapter 1 of [1].

1 Operators in vector spaces

1.1 Vector spaces and adjoint vector spaces

Let X be a vector space, and $\dim X$ be the dimension of X. In this thesis we shall always assume that $\dim X < \infty$.

A subset M of X is a subspace, if M is itself a vector space. We define the codimension of M in X by setting $\mathrm{codim} M = \dim X - \dim M$.

Example 1. The set $X = C^N$ of all ordered N-tuples $u = (\xi_i) = (\xi_1, ..., \xi_N)$ of complex numbers is an N-dimensional vector space.

Let $\dim X = N$, if $x_1, ..., x_N$ are linearly independent, then they are a basis of X, and each $u \in X$ can be uniquely represented as

$$u = \sum_{j=1}^{N} \xi_j x_j \ , \tag{1}$$

and the scalars ξ_j are called the coefficients of u with respect to this basis.

Example 2. In C^N the N vectors $x_j = (0, ..., 0, 1, 0, ..., 0)$ with 1 in the *j-th* place, $j = 1, ..., N$, form a basis (the canonical basis). The coefficients of $u = (\xi_j)$ with respect to the canonical basis are the ξ_j themselves.

If $\{x'_j\}$ is another basis of X, since $u = \sum \xi_j x_j$, there is a system of linear relations

$$x_k = \sum_j \gamma_{jk} x'_j \ , \quad k = 1, ..., N. \tag{2}$$

When ξ_j,ξ'_j are coefficients of the same vector u with respect to the bases $\{x_j\}$ and $\{x'_j\}$ respectively, they are then related to each other by

$$\xi'_j = \sum_k \gamma_{jk} \xi_k \ , \quad j = 1, ..., N. \tag{3}$$

The inverse transformations to (2) and (3) are

$$x'_j = \sum_k \hat{\gamma}_{kj} x_k \ , \quad \xi_k = \sum_j \hat{\gamma}_{kj} \xi'_j \ , \tag{4}$$

where $(\hat{\gamma}_{jk})$ is the inverse of the matrix (γ_{jk}) :

$$\sum_i \hat{\gamma}_{ji}\gamma_{ik} = \sum_i \gamma_{ji}\hat{\gamma}_{ik} = \delta_{jk} = \begin{cases} 1, j = k \\ 0, j \neq k. \end{cases} \tag{5}$$

Let M_1, M_2 be subspaces of X, we define the subspace $M_1 + M_2$ by

$$M_1 + M_2 = \{x_1 + x_2 : x_1 \in M_1, x_2 \in M_2\}. \tag{6}$$

Theorem 1. If M_1, M_2 are subspaces of X, then

$$\dim(M_1 + M_2) = \dim M_1 + \dim M_2 - \dim(M_1 \cap M_2). \tag{7}$$

Proof. We can see that $M_1 \cap M_2$ is subspace of both M_1 and M_2. Suppose that $\dim M_1 = m_1, \dim M_2 = m_2$ and $\dim(M_1 \cap M_2) = m$. We have to prove that

$$\dim(M_1 + M_2) = m_1 + m_2 - m.$$

Suppose that $\{x_1, ..., x_m\}$ is a basis of $M_1 \cap M_2$. We can extend this basis to a basis of M_1 and to a basis of M_2. Let for example

$$\left\{x_1, ..., x_m, x_1^{(1)}, ..., x_{m_1-m}^{(1)}\right\}, \left\{x_1, ..., x_m, x_1^{(2)}, ..., x_{m_2-m}^{(2)}\right\}$$

be two bases of M_1 and M_2 respectively.
Let $B = \left\{x_1, ..., x_m, x_1^{(1)}, ..., x_{m_1-m}^{(1)}, x_1^{(2)}, ..., x_{m_2-m}^{(2)}\right\}$. Then B generates $M_1 + M_2$. Now we will show that the vectors in B are linearly independent. Suppose that

$$\alpha_1 x_1 + ... + \alpha_m x_m + \beta_1 x_1^{(1)} + ... + \beta_{m_1-m} x_{m_1-m}^{(1)} + \gamma_1 x_1^{(2)} + ... + \gamma_{m_2-m} x_{m_2-m}^{(2)} = 0. \tag{8}$$

Let

$$u = \alpha_1 x_1 + ... + \alpha_m x_m + \beta_1 x_1^{(1)} + ... + \beta_{m_1-m} x_{m_1-m}^{(1)}, \tag{9}$$

we have also

$$u = -\gamma_1 x_1^{(2)} - ... - \gamma_{m_2-m} x_{m_2-m}^{(2)}, \tag{10}$$

thus $u \in M_1$ and $u \in M_2$ by (9) and (10). Hence $u \in M_1 \cap M_2$. Therefore there are $\zeta_i, i = 1, ..., m$ such that $u = \sum \zeta_i x_i$, and

$$\zeta_1 x_1 + ... + \zeta_m x_m + \gamma_1 x_1^{(2)} + ... + \gamma_{m_2-m} x_{m_2-m}^{(2)} = 0.$$

Since $\left\{x_1, ..., x_m, x_1^{(2)}, ..., x_{m_2-m}^{(2)}\right\}$ is a basis of M_2, then $\gamma_1 = 0, ..., \gamma_{m_2-m} = 0$. Substituting this in (8), we obtain $\alpha_1 x_1 + ... + \alpha_m x_m + \beta_1 x_1^{(1)} + ... + \beta_{m_1-m} x_{m_1-m}^{(1)} = 0$, thus $\alpha_1 = 0, ..., \alpha_m = 0, \beta_1 = 0, ..., \beta_{m_1-m} = 0$. Hence B is a basis of $M_1 + M_2$. Since B has $m_1 + m_2 - m$ vectors, we get the required result. □

Definition 1. (Direct sum) Let $M_1, ..., M_s$ be subspaces of X. X is a direct sum of them if each $u \in X$ has a unique expression of the form

$$u = \sum_j u_j, \; u_j \in M_j, \; j = 1, ..., s \; . \tag{11}$$

Then we write

$$X = M_1 \oplus ... \oplus M_s \; .$$

Proposition 1. If $X = M_1 \oplus ... \oplus M_s$, then

1. $X = M_1 + ... + M_s$.
2. $M_i \bigcap M_j = \{0\}$, where $i \neq j$.
3. $\dim X = \sum_j \dim M_j$.

Proof. If $X = M_1 \oplus ... \oplus M_s$, then each $u \in X$ can be uniquely expressed as $u = u_1 + ... + u_s$, where $u_j \in M_j$. Hence $X = M_1 + ... + M_s$. Now let $u \in M_i \cap M_j$, where $i \neq j$, then

$$u = 0_1 + ... + 0_{i-1} + u_i + 0_{i+1} + ... + 0_j + ... + 0_s \; ,$$

and

$$u = 0_1 + ... + 0_i + ... + 0_{j-1} + u_j + 0_{j+1} + ... + 0_s \; .$$

Since the expression (11) for u is unique, $u_i = u_j = 0$ and hence $u = 0$.

To show the last equality in the proposition we assume that $\dim X = N$, $\dim M_j = m_j$, and we have to show that $N = \sum m_j$.

Let x_i^1,...,x_i^s, $i = 1, ..., m_j$ be bases of $M_j, j = 1, ..., s$ respectively. Suppose that

$$\sum_{i=1}^{m_1} \alpha_{1i} x_i^1 + \sum_{i=1}^{m_2} \alpha_{2i} x_i^2 + ... + \sum_{i=1}^{m_s} \alpha_{si} x_i^s = 0 \ . \tag{12}$$

$\sum_{i=1}^{m_j} \alpha_{ji} x_i^j \in M_j$, and since X is the direct sum of M_j, the representation (12) is unique, and so for $j = 1, ..., s$ we have $\sum_{i=1}^{m_j} \alpha_{ji} x_i^j = 0$. Since $\{x_i^j\}$ are linearly independent, $\alpha_{ji} = 0$ for $i = 1, ..., m_j, j = 1, ..., s$. Hence $x_i^j, i = 1, ..., m_j, j = 1, ..., s$ are linearly independent.
Now suppose $u \in X$, $u = u_1 + u_2 + ... + u_s, u_j \in M_j$. Since u_j is expressed in a unique way as a linear combination of $x_i^j, i = 1, ..., m_j$, then u is expressed in a unique way as a linear combination of $x_i^j, j = 1, ..., s$. Hence $x_i^j, j = 1, ..., s$ is a basis of X. Thus

$$\dim X = \sum_j \dim M_j \ .$$

$\square$

Definition 2. We call $\|u\|$ a norm of $u \in X$, if
(i) $\|u\| \geq 0$ for all $u \in X$ and $\|u\| = 0$ iff $u = 0$.
(ii) $\|\alpha u\| = |\alpha| \|u\|$ for all $\alpha \in C, u \in X$.
(iii) $\|u + v\| \leq \|u\| + \|v\|$ for all $u, v \in X$.

Example 3.

$$\|u\| = \max_j |\xi_j|,$$

$$\|u\| = \sum_j |\xi_j|$$

$$\|u\| = \left(\sum_j |\xi_j|^2 \right)^{\frac{1}{2}} ,$$

where ξ_j are coefficients of u with respect to the basis $\{x_j\}$ in X, are three different norms.

We show only that the last expression is a norm. Let ξ_j,η_j be the coefficients of u, v respectively, then

-$\|u\| = \left(\sum |\xi_j|^2\right)^{\frac{1}{2}} \geq 0$ and $\|u\| = 0$ iff $\xi_j = 0$, $\forall_j$.

-$\|\alpha u\| = \left(\sum |\alpha\xi_j|^2\right)^{\frac{1}{2}} = \left(|\alpha|^2 \sum |\xi_j|^2\right)^{\frac{1}{2}} = |\alpha| \left(\sum |\xi_j|^2\right)^{\frac{1}{2}}$.

-It follows from the Schwarz inequality (see e.g. [2]) that

$$(\sum |\xi_j\eta_j|)^2 \leq \sum |\xi_j|^2 \sum |\eta_j|^2,$$

therefore

$$2\sum |\xi_j\eta_j| \leq 2\left(\sum |\xi_j|^2\right)^{\frac{1}{2}} \left(\sum |\eta_j|^2\right)^{\frac{1}{2}}.$$

Adding $\sum |\xi_j|^2 + \sum |\eta_j|^2$ to both sides gives

$$\begin{aligned}\sum(|\xi_j|^2 + 2|\xi_j\eta_j| + |\eta_j|^2) &= \sum |\xi_j|^2 + 2\sum |\xi_j\eta_j| + \sum |\eta_j|^2 \\ &\leq \sum |\xi_j|^2 + 2(\sum |\xi_j|^2)^{\frac{1}{2}}(\sum |\eta_j|^2)^{\frac{1}{2}} + \sum |\eta_j|^2 \\ &= [(\sum |\xi_j|^2)^{\frac{1}{2}} + (\sum |\eta_j|^2)^{\frac{1}{2}}]^2.\end{aligned}$$

Hence

$$(\sum |\xi_j + \eta_j|^2)^{\frac{1}{2}} \leq (\sum |\xi_j|^2)^{\frac{1}{2}} + (\sum |\eta_j|^2)^{\frac{1}{2}},$$

that is

$$\|u + v\| \leq \|u\| + \|v\|.$$

-The adjoint space

Definition 3. (Linear forms and semilinear forms): A complex-valued function $f[u]$ defined for $u \in X$ is called a linear form if

$$f[\alpha u + \beta v] = \alpha f[u] + \beta f[v] \tag{13}$$

for all $u, v \in X$, and all scalars α, β, and a semilinear form if

$$f[\alpha u + \beta v] = \bar{\alpha} f[u] + \bar{\beta} f[v]. \tag{14}$$

Example 4. Let $x_1, ..., x_N$ be a fixed basis in X. It follows from (13) that a linear form on X can be expressed in the form

$$f[u] = \sum_{j=1}^{N} \alpha_j\xi_j \ , \text{ where } u = (\xi_j) \ \text{ and } f[x_j] = \alpha_j \ ,$$

and similarly by (14) a semilinear form on X can be expressed in the form

$$f[u] = \sum_j \alpha_j \bar{\xi_j} \ .$$

$f[u]$ is a semilinear form if and only if $\overline{f[u]}$ is a linear form.

Definition 4. The set of all semilinear forms on X is a vector space, called the adjoint (or conjugate) space of X and is denoted by X^*.

Let us denote $f[u]$ by (f, u) where f is a semilinear form. It follows from the definition that (f, u) is linear in f and semilinear in u:

$$(\alpha f + \beta g, u) = \alpha(f, u) + \beta(g, u), \tag{15}$$

$$(f, \alpha u + \beta v) = \bar{\alpha}(f, u) + \bar{\beta}(f, v). \tag{16}$$

Example 5. For $X = C^N, X^*$ may be regarded as the set of all row vectors $f = (\alpha_j)$ whereas X is the set of all column vectors $u = (\xi_j)$, and

$$(f, u) = \sum \alpha_j \bar{\xi_j}.$$

The adjoint basis The principal content in this part is the following theorem:

Theorem 2. Suppose $\{x_j\}$ is a basis of X, and let $e_1, ..., e_N$ be vectors in X^* defined by

$$(e_j, x_k) = \delta_{jk} = \begin{cases} 1, j = k \\ 0, j \neq k \end{cases} , \tag{17}$$

then $\{e_j\}$ is a basis of X^*.

Proof. First we show that e_j satisfying (17) exist. Define $e_j, j = 1, ..., N$ by $(e_j, u) = \bar{\xi_j}$. Then this corresponds to $\alpha_k = \delta_{jk}, k = 1, ..., N$, in the formula above. Next we shall show that $\{e_j\}$ generate X^*. Let $f \in X^*$, and suppose $(f, x_1) = \alpha_1, (f, x_2) = \alpha_2, ..., (f, x_N) = \alpha_N$. Put $g = \sum \alpha_j e_j$, then $(g, x_1) = (\sum \alpha_j e_j, x_1) = \alpha_1(e_1, x_1) + \alpha_2(e_2, x_1) + ... + \alpha_N(e_N, x_1) = \alpha_1$, and similarly for $j = 2, ..., N, (g, x_j) = \alpha_j$ so that $f(x_j) = g(x_j), j = 1, ..., N$. Since f, g are equal on vectors of the basis $\{x_j\}$, then $f = g = \alpha_1 e_1 + ... + \alpha_N e_N$, i.e., f is a linear combination of $e_1, ..., e_N$. It remains to show $e_1, ..., e_N$ are linearly independent. Suppose that $\alpha_1 e_1 + ... + \alpha_N e_N = 0$. Then

$$\begin{aligned} 0 = (0, x_1) &= (\alpha_1 e_1 + ... + \alpha_N e_N, x_1) \\ &= \alpha_1 (e_1, x_1) + ... + \alpha_N (e_N, x_1) \\ &= \alpha_1 \ , \end{aligned}$$

and similarly for $k = 2, ..., N$, so that we have $\alpha_1 = ... = \alpha_N = 0$. Thus $e_1, ..., e_N$ are linearly independent. Hence $\{e_j\}$ is a basis of X^*. □

Let $\{x_j\}$ be a basis of X, and let $\{e_1, ..., e_N\}, \{e'_1, ..., e'_N\}$ be vectors in X^* satisfying (17). Then by Theorem 2 $\{e_j\}$ and $\{e'_j\}$ are bases of X^* and

$$e'_i = \sum_{j=1}^{N} \alpha_j^{(i)} e_j, \ \ i = 1, ..., N \ ,$$

so that

$$\begin{aligned} (e'_1, x_1) &= \alpha_1^{(1)} (e_1, x_1) + \alpha_2^{(1)} (e_2, x_1) + ... + \alpha_N^{(1)} (e_N, x_1) \\ &= \alpha_1^{(1)} \cdot 1 + \alpha_2^{(1)} \cdot 0 + ... + \alpha_2^{(1)} \cdot 0 \\ &= \alpha_1^{(1)} \ , \end{aligned}$$

and $(e'_1, x_2) = \alpha_2^{(1)}, ..., (e'_1, x_N) = \alpha_N^{(1)}$. Hence $\alpha_j^{(1)} = \delta_{j1}$ and $e'_1 = e_1$. Similarly for $j = 2, ..., N$ we obtain $e'_j = e_j$. Hence the basis $\{e_j\}$ of X^* that satisfies (17) is unique. It is called the basis adjoint to the basis $\{x_j\}$ of X. Theorem 2 shows that

$$\dim X^* = \dim X. \tag{18}$$

Let $\{x_j\}$ and $\{x'_j\}$ be two bases of X related to each other by (2). Then the corresponding adjoint bases $\{e_j\}$ and $\{e'_j\}$ of X^* are related to each other by the formulas

$$e'_j = \sum_k \bar{\gamma}_{jk} e_k \ , \quad e_k = \sum_j \bar{\bar{\gamma}}_{kj} e'_j \ . \tag{19}$$

Furthermore we have

$$\bar{\gamma}_{jk} = (e'_j, x_k) \ , \quad \bar{\bar{\gamma}}_{kj} = (e_k, x'_j) \ . \tag{20}$$

Definition 5. Let $f \in X^*$. The norm $\|f\|$ is defined by

$$\|f\| = \sup_{0 \neq u \in X} \frac{|(f, u)|}{\|u\|} = \sup_{\|u\|=1} |(f, u)| \ . \tag{21}$$

1.2 Linear operators

Definition 6. Let X, Y be two vector spaces. A function T that sends every vector u of X into a vector $v = Tu$ of Y is called a linear transformation or a linear operator on X to Y if

$$T(\alpha_1 u_1 + \alpha_2 u_2) = \alpha_1 T u_1 + \alpha_2 T u_2 \tag{22}$$

for all $u_1, u_2 \in X$ and all scalars α_1, α_2 .

If $Y = X$ we say T is a linear operator in X .
If M is a subspace of X, then $T(M)$ is a subspace of Y, the subspace $T(X)$ of Y is called the range of T and is denoted by $R(T)$, $\dim(R(T))$ is called the rank of T. The codimension of $R(T)$ with respect to Y is called the deficiency of T and is denoted by def T, hence

$$\text{rank } T + \text{def } T = \dim Y \ . \tag{23}$$

The set of all $u \in X$ such that $Tu = 0$ is a subspace of X and is called the kernel or null space of T and is denoted by $N(T)$. $\dim(N(T))$ is denoted by nul T, and we have

$$\text{rank } T + \text{nul } T = \dim X \ . \tag{24}$$

If both nul T and def T are zero, then T is one-to-one. In this case the inverse operator T^{-1} is defined.

Let $\{x_k\}$ be a basis of X. Each $u \in X$ has the expansion (1) so that

$$Tu = \sum_{k=1}^{N} \xi_k T x_k \ , \quad N = \dim X \ . \tag{25}$$

Thus an operator T on X to Y is determined by giving the values of $Tx_k, \quad k = 1, ..., N$. If $\{y_j\}$ is a basis of Y, each Tx_k has the expansion

$$Tx_k = \sum_{j=1}^{M} \tau_{jk} y_j \ , \quad M = \dim Y \ . \tag{26}$$

Substituting (26) into (25), the coefficients η_j of $v = Tu$ are given by

$$\eta_j = \sum_{k=1}^{N} \tau_{jk} \xi_k, \quad j = 1, ..., M \ . \tag{27}$$

In this way an operator T on X to Y is represented by an $M \times N$ matrix (τ_{jk}) with respect to the bases $\{x_k\}$,$\{y_j\}$ of X, Y respectively.
When (τ'_{jk}) is the matrix representing the same operator T with respect to a new pair of bases $\{x'_k\}, \{y'_j\}$, we can find the relationship between the matrices (τ'_{jk}) and (τ_{jk}) by combining (26) and a similar expression for Tx'_k in terms of $\{y'_j\}$ with the formulas (2),(4):

$$\begin{aligned}
Tx'_k &= T(\sum_h \hat{\gamma}_{hk} x_h) \\
&= \sum_h \hat{\gamma}_{hk} T x_h \\
&= \sum_h \hat{\gamma}_{hk} \sum_{i=1}^{M} \tau_{ih} y_i \\
&= \sum_i \sum_h \tau_{ih} \hat{\gamma}_{hk} y_i \\
&= \sum_i \sum_h \tau_{ih} \hat{\gamma}_{hk} \sum_j \gamma'_{ji} y'_j \\
&= \sum_j \sum_i \sum_h \gamma'_{ji} \tau_{ih} \hat{\gamma}_{hk} y'_j \\
&= \sum_j \sum_{i,h} \gamma'_{ji} \tau_{ih} \hat{\gamma}_{hk} y'_j \\
&= \sum_j \tau'_{jk} y'_j \ ,
\end{aligned}$$

where $x_h = \sum_k \gamma_{kh} x'_k$ and $x'_k = \sum_h \hat{\gamma}_{hk} x_h,\ y_i = \sum_j \gamma'_{ji} y'_j$ and $y'_j = \sum_i \gamma''_{ij} y_i$.
Thus the matrix τ'_{jk} is the product of three matrices (γ'_{jk}), (τ_{jk}) and $(\hat{\gamma}_{jk})$,

$$(\tau'_{jk}) = (\gamma'_{jk})(\tau_{jk})(\hat{\gamma}_{jk}) \ . \tag{28}$$

Thus when T is an operator on X to itself det (τ_{jk}) and the trace of (τ_{jk}), i.e., $\sum \tau_{jj}$ are determined by the operator T itself. More precisely, we shall show det (τ_{jk}) and trace (τ_{jk}) are the same for each choice of the basis for X. (28) becomes

$$(\tau'_{jk}) = (\gamma_{jk})(\tau_{jk})(\hat{\gamma}_{jk}) \ . \tag{29}$$

We show tr $(\gamma\tau)$ = tr $(\tau\gamma)$. Let $\gamma\tau = (a_{jk})$ and $\tau\gamma = b_{jk}$, then tr $(\gamma\tau) = \sum_j a_{jj} = \sum_j \sum_k \gamma_{jk}\tau_{kj} = \sum_k \sum_j \tau_{kj}\gamma_{jk} = \sum_k b_{kk} =$ tr $(\tau\gamma)$. Since $(\hat{\gamma}_{jk})$ is the inverse of the matrix (γ_{jk}), and we know that

$$\det(\gamma\tau\hat{\gamma}) = \det(\gamma)\det(\tau)\det(\hat{\gamma}),$$

we have

$$\det(\tau') = \det(\tau) \ , \ \text{and} \ \ \mathrm{tr}\ (\tau') = \mathrm{tr}\ (\gamma\tau\hat{\gamma}) = \mathrm{tr}\ (\tau\gamma\hat{\gamma}) = \mathrm{tr}\ (\tau)\ . \tag{30}$$

Example 6. If $\{f_j\}$ is the basis of Y^* adjoint to $\{y_j\}$ then

$$\tau_{jk} = (f_j, Tx_k)\ . \tag{31}$$

Proof. Since $Tx_k = \sum_i \tau_{ik} y_i$, then

$$\begin{aligned}
(f_j, Tx_k) &= (f_j, \sum_i \tau_{ik} y_i) \\
&= \sum_i \tau_{ik}(f_j, y_i) \\
&= \sum_i \tau_{ik}\delta_{ij} \\
&= \tau_{jk}\ .
\end{aligned}$$

□

Example 7. Let $\{x_j\}$ and $\{e_j\}$ be the bases of X and X^*, respectively, which are adjoint to each other. If T is an operator on X to itself, we have

$$\mathrm{tr}\ T = \sum_j (e_j, Tx_j)\ . \tag{32}$$

Proof. Similarly as in the last example, $Tx_j = \sum_i \tau_{ij} x_i$, therefore

$$\begin{aligned}
\sum_j (e_j, Tx_j) &= \sum_j (e_j, \sum_i \tau_{ij} x_i) \\
&= \sum_j (\sum_i \tau_{ij}(e_j, x_i)) \\
&= \sum_j \sum_i \tau_{ij}\delta_{ij} \\
&= \sum_j \tau_{jj}\ .
\end{aligned}$$

□

If T and S are two linear operators on X to Y, then we define:

$$(\alpha S + \beta T)u = \alpha(Su) + \beta(Tu) \ .$$

If S maps X to Y and T maps Y to Z, then we set

$$(TS)u = T(Su) \ .$$

Example 8. 1. rank $(S+T) \leq$ rank $S +$ rank T.

2. rank $(TS) \leq \max(\text{rank } T, \text{rank } S)$.

Proof. (1) Let $R(S) = M_1, R(T) = M_2, R(S+T) = M$. Since each $v \in M$ can be expressed in the form $v = v_1 + v_2, v_1 \in M_1, v_2 \in M_2$, $M = M_1 + M_2$, thus $\dim M = \dim M_1 + \dim M_2 - \dim(M_1 \cap M_2) \leq \dim M_1 + \dim M_2$. Hence rank $(S+T) \leq$ rank $(S) +$ rank (T).
(2) Let $S : X \to Y$, $T : R(S) \to Z$; $TS : X \to Z$. Then we have rank $S +$ nul $S = \dim X$ and rank $(TS) +$ nul $(TS) = \dim X$. Since nul $S \leq$ nul TS, rank $TS \leq$ rank S. Let $T : Y \to Z$, $TS : X \to Z$. Since $T : Y \to Z$, rank $T+$def $T = \dim Z$, and rank $(TS)+$def $(TS) = \dim Z$. Since def $T \leq$ def (TS), rank $(TS) \leq$ rank T. Thus rank $(TS) \leq \max(\text{rank } T, \text{rank } S)$. □

Let us denote by $L(X,Y)$ the set of all operators on X to Y. It is a vector space. Let $L(X) = L(X,X)$, then we have:

- $T0 = 0T = 0$.
- $T1 = 1T = T$; (1 is the identity operator) .
- $T^mT^n = T^{m+n}, (T^m)^n = T^{mn}, m,n = 0,1,...,$.
- If $S,T \in L(X)$ are nonsingular, then $T^{-1}T = 1$, $T^{-n} = (T^{-1})^n$, and $(TS)^{-1} = S^{-1}T^{-1}$.

For any polynomial $P(z) = \alpha_0 + \alpha_1 z + ... + \alpha_n z^n$ in the indeterminate z, we define the operator

$$P(T) = \alpha_0 + \alpha_1 T + ... + \alpha_n T^n.$$

Example 9. If $S \in L(X,Y)$ and $T \in L(Y,X)$, then $ST \in L(Y)$ and $TS \in L(X)$.

1.3 Projections

Let M, N be two complementary subspaces of X, $X = M \oplus N$. Thus each $u \in X$ can be uniquely expressed in the form $u = u' + u'', u' \in M, u'' \in N$. u' is called the projecion of u on M along N. If $v = v' + v''$, then $\alpha u + \beta v$ has the projection $\alpha u' + \beta v'$ on M along N. If we set $u' = Pu$, it follows that P is a linear operator in X called the projection operator or projection on M along N. $1 - P$ is the projection on N along M, and we have $Pu = u$ if and only if $u \in M, Pu = 0$ if and only if $u \in N$, that is $R(P) = N(1-P) = M, N(P) = R(1-P) = N$. Furthermore, $PPu = Pu$, that is P is idempotent, i.e.,

$$P^2 = P.$$

Remark 1. Any idempotent operator P is a projection.
To show, let $M = R(P)$ and $N = R(1-P)$. If $u' \in M$, there is u such that $Pu = u'$ and therefore $Pu' = P^2u = Pu = u'$. Similarly if $u'' \in N$. Now let $u \in M \cap N$. Then $u = Pu = 0$. So $M \cap N = \{0\}$. Thus each $u \in X$ has the expression $u = u' + u''$ with $u' = Pu \in M$ and $u'' = (1-P)u \in N$, proving that P is the projection on M along N .

Example 10. If P is a projection, then we have

$$\text{tr } P = \dim R(P) \ .$$

Proof. Since P is an operator in X, it can be represented by an $n \times n$ ($n = \dim X$) matrix (τ_{jk}) with respect to the basis $\{x_j\}$ of X, and $Pu = u$ when $u \in M$. This basis can be chosen so that $x_1, ..., x_m \in M$ and $x_{m+1}, ..., x_n \in N$, where $N = (1-P)X$. Then $P(\alpha_1 x_1 + ... + \alpha_n x_n) = \alpha_1 x_1 + ... + \alpha_m x_m$. Hence (τ_{jk}) is diagonal with $\tau_{11} = ... = \tau_{mm} = 1$, $\tau_{m+1,m+1} = ... = \tau_{nn} = 0$. So tr $P = m = \dim R(P)$.

□

In general, if $X = M_1 \oplus ... \oplus M_s$, then each $u \in X$ can be uniquely expressed in the form $u = u_1 + ... + u_s,\ u_j \in M_j,\ j = 1, ..., s$. Then the operator P_j defined by $P_j u = u_j \in M_j$, is a projection on M_j along $N_j = M_1 \oplus ... \oplus M_{j-1} \oplus M_{j+1} \oplus ... \oplus M_s$. And we have

$$\sum P_j = 1 \ , \tag{33}$$

for $(\sum_j P_j)(\sum_i u_i) = \sum_j P_j(\sum_i u_i) = \sum P_j u_j = \sum u_j$, and

$$P_k P_j = \delta_{jk} P_j \ , \tag{34}$$

because $P_k P_j(\sum_i u_i) = P_k u_j = \delta_{kj} u_j = \delta_{kj} P_j u$.

Note that, if we have (33) and (34), then X is the direct sum of subspaces $R(P_j)$. To show that, let $M_j = R(P_j)$. For $u \in X$, by (33) we have $u = \sum P_j u = \sum u_j \in M_1 + ... + M_s$. Moreover $M_i \cap M_j = \{0\}$ for $i \neq j$, because by (34) if $u \in M_i \cap M_j$, then $u = P_i u_1 = P_j u_2$ and $u = P_i u_1 = P_i^2 u_1 = P_i P_j u_2 = 0$. Hence $X = M_1 \oplus ... \oplus M_s$. Since P_j is idempotent, it follows from Remark 1 that P_j is the projection on M_j along N_j.

1.4 The adjoint operator

Definition 7. Let $T \in L(X, Y)$, a function T^* on Y^* to X^* is called the adjoint operator of T if :

$$(T^* g, u) = (g, Tu), \ \forall g \in Y^*, \ \forall u \in X. \tag{35}$$

Then $(T^*(\alpha_1 g_1 + \alpha_2 g_2), u) = (\alpha_1 g_1 + \alpha_2 g_2, Tu) = \alpha_1(g_1, Tu) + \alpha_2(g_2, Tu) = \alpha_1(T^* g_1, u) + \alpha_2(T^* g_2, u)$ so that $T^*(\alpha_1 g_1 + \alpha_2 g_2) = \alpha_1 T^* g_1 + \alpha_2 T^* g_2$. Therefore T^* is a linear operator on Y^* to X^*, that is, $T^* \in L(Y^*, X^*)$.

The operation * has the following properties:

1. $(\alpha S + \beta T)^* = \bar{\alpha} S^* + \bar{\beta} T^*$, for $S, T \in L(X, Y)$, and $\alpha, \beta \in C$.
2. $(TS)^* = S^* T^*$, for $T \in L(Y, Z)$ and $S \in L(X, Y)$.

Note that $S^* \in L(Y^*, X^*)$ and $T^* \in L(Z^*, Y^*)$ so that $S^* T^* \in L(Z^*, X^*)$. Then $((TS)^* h, u) = (h, TSu) = (T^* h, Su) = (S^* T^* h, u), \forall h \in Z^*, \forall u \in X$. Hence 2. holds.

Example 11. If $T \in L(X)$, we have $0^* = 0, 1^* = 1$.

If $\{x_k\}, \{y_j\}$ are bases in X, Y respectively, and $T \in L(X, Y)$ is represented by a matrix (τ_{jk}) in these bases, and $\{e_k\}, \{f_j\}$ are the adjoint bases of X^*, Y^* respectively, the operator $T^* \in L(Y^*, X^*)$ can be represented by a matrix (τ^*_{kj}). These matrices are given by $\tau_{jk} = (f_j, Tx_k)$ according to (31) and $\bar{\tau}^*_{kj} = (T^* f_j, x_k) = (f_j, Tx_k)$ (see the argument below), thus

$$\tau^*_{kj} = \bar{\tau}_{jk}, \quad k = 1, ..., N = \dim X, \ j = 1, ..., M = \dim Y. \tag{36}$$

To show that $(T^* f_j, x_k) = \bar{\tau}^*_{kj}$, we first write $T^* f_j = \sum\limits_i \tau^*_{ij} f_i$ and then compute:
$(T^* f_j, x_k) = (\sum\limits_i \tau^*_{ij} f_i, x_k) = \sum\limits_i \bar{\tau}^*_{ij}(f_i, x_k) = \sum\limits_i \bar{\tau}^*_{ij}\delta_{ik} = \bar{\tau}^*_{kj}$.

Example 12. If $T \in L(X)$, we have

$$\det T^* = \overline{\det T},\ \mathrm{tr}\ T^* = \overline{\mathrm{tr}\ T} \tag{37}$$

and

$$(T^*)^{-1} = (T^{-1})^*. \tag{38}$$

Since det (τ_{jk}) and tr (τ_{jk}) are the same for each choice of the basis for X and similarly with (τ^*_{kj}), (37) is satisfied according to (36). To prove (38) note that $T^*(T^{-1})^* = (T^{-1}T)^* = 1^* = 1$.

Definition 8. (Norm of T) The norm of T is defined by

$$\|T\| = \sup_{0\neq u \in X} \frac{\|Tu\|}{\|u\|} = \sup_{\|u\|=1} \|Tu\|,\ T \in L(X,Y). \tag{39}$$

Example 13.

$$\|T\| = \sup_{\substack{0\neq u\in X\\ 0\neq f\in Y^*}} \frac{|(f,Tu)|}{\|f\|\|u\|} = \sup_{\substack{\|u\|=1\\ \|f\|=1}} |(f,Tu)|. \tag{40}$$

We first prove that the expression for $\|T\|$ given by (40) is a norm.

Proof. -$\|T\| = \sup\limits_{\substack{\|u\|=1\\ \|f\|=1}} |(f,Tu)| \geq 0$, and $= 0$ iff $T = 0$.

-$\|\alpha T\| = \sup\limits_{\substack{\|u\|=1\\ \|f\|=1}} |(f,\alpha Tu)| = \sup\limits_{\substack{\|u\|=1\\ \|f\|=1}} |\alpha(f,Tu)| = |\alpha| \sup\limits_{\substack{\|u\|=1\\ \|f\|=1}} |(f,Tu)| = |\alpha|\|T\|$.

-$\|T+S\| = \sup\limits_{\substack{\|u\|=1\\ \|f\|=1}} |(f,Tu+Su)| = \sup\limits_{\substack{\|u\|=1\\ \|f\|=1}} |(f,Tu)+(f,Su)| \leq \sup\limits_{\substack{\|u\|=1\\ \|f\|=1}} (|(f,Tu)|+$
$|(f,Su)|) \leq \sup\limits_{\substack{\|u\|=1\\ \|f\|=1}} |(f,Tu)| + \sup\limits_{\substack{\|u\|=1\\ \|f\|=1}} |(f,Su)| = \|T\| + \|S\|$. Hence $\|T\|$ defined in (40) is a norm. □

We have to show that (39) and (40) are equivalent. To see this we recall (21). Note that $|(f,u)| \le \|f\|\|u\|$. This implies that $\|u\| = \sup\limits_{0\neq f\in X^*} \dfrac{|(f,u)|}{\|f\|} = \sup\limits_{\|f\|=1} |(f,u)|$ (see Section I.2.5 in [1]). It follows that $\|Tu\| = \sup\limits_{\|f\|=1} |(f,Tu)|$ and $\|T\| = \sup\limits_{\|u\|=1} \|Tu\| = \sup\limits_{\substack{\|u\|=1\\ \|f\|=1}} |(f,Tu)|$.

Since $\|T\| = \sup\limits_{\|u\|=1} \dfrac{\|Tu\|}{\|u\|}$, then $\dfrac{\|Tu\|}{\|u\|} \le \|T\|$, so $\|Tu\| \le \|T\|\|u\|$. Hence $\|TSu\| \le \|T\|\|Su\| \le \|T\|\|S\|\|u\|$, thus

$$\|TS\| \le \|T\|\|S\| \tag{41}$$

for $T \in L(Y,Z)$ and $S \in L(X,Y)$.

If $T \in L(X,Y)$, then $T^* \in (Y^*, X^*)$ and

$$\|T^*\| = \|T\|. \tag{42}$$

This follows from (40) according to which $\|T^*\| = \sup |(T^*f,u)| = \sup |(f,Tu)| = \|T\|$ where $u \in X$, $\|u\| = 1$ and $f \in X^*$, $\|f\| = 1$.

1.5 The eigenvalue problem

Definition 9. Let $T \in L(X)$. A complex number λ is called an eigenvalue of T if there is a non-zero vector $u \in X$ such that

$$Tu = \lambda u. \tag{43}$$

u is called an eigenvector of T belonging to the eigenvalue λ. The set N_λ of all $u \in X$ such that $Tu = \lambda u$ is a subspace of X called the eigenspace of T for the eigenvalue λ and $\dim N_\lambda$ is called the multiplicity of λ .

Example 14. λ is an eigenvalue of T if and only if $\lambda - \xi$ is an eigenvalue of $T - \xi$.

Proposition 2. The eigenvectors of T belonging to different eigenvalues are linearly independent.

Proof. To prove that we will use induction.
We shall first show that any two eigenvectors of T belonging to different

eigenvalues are linearly independent. Assume next this is true for k eigenvectors, and we shall prove it for $k+1$ eigenvectors. Let $Tu_1 = \lambda_1 u_1, Tu_2 = \lambda_2 u_2, \lambda_1 \neq \lambda_2, \lambda_1 \neq 0$ and

$$\alpha_1 u_1 + \alpha_2 u_2 = 0,$$

we have

$$\alpha_1 \lambda_1 u_1 + \alpha_2 \lambda_2 u_2 = 0.$$

By multiplying the first equation by λ_1 and subtracting it from the second we obtain $(\lambda_1 - \lambda_2)\alpha_2 u_2 = 0$. Hence $\alpha_2 u_2 = 0$, but $u_2 \neq 0$, so $\alpha_2 = 0$. Now $\alpha_1 u_1 = 0$ which implies $\alpha_1 = 0$ since $u_1 \neq 0$. Hence $\alpha_1 = \alpha_2 = 0$, that is u_1, u_2 are linearly independent.
Now assume $Tu_1 = \lambda_1 u_1, ..., Tu_k = \lambda_k u_k,\ \lambda_i \neq \lambda_j$ for $i \neq j$ and $u_1, ..., u_k$ are linearly independent. We shall show that $u_1, ..., u_k, u_{k+1}$ are linearly independent where $Tu_{k+1} = \lambda_{k+1} u_{k+1}$. We have two cases: $\lambda_{k+1} = 0$ and $\lambda_{k+1} \neq 0$.
If $\lambda_{k+1} = 0$, then we have $\lambda_i \neq 0, i = 1, ..., k$. If $u_1, ..., u_k, u_{k+1}$ are linearly dependent then

$$u_1 = \alpha_2 u_2 + ... + \alpha_k u_k + \alpha_{k+1} u_{k+1},$$

and

$$\lambda_1 u_1 = Tu_1 = \alpha_2 \lambda_2 u_2 + ... + \alpha_k \lambda_k u_k,$$

thus $u_1, ..., u_k$ are linearly dependent, and this contradicts our assumption.
If $\lambda_{k+1} \neq 0$ suppose

$$u_{k+1} = \alpha_1 u_1 + ... + \alpha_k u_k,$$

where $\alpha_1 \neq 0$, then we have

$$\begin{aligned} \lambda_{k+1} u_{k+1} &= Tu_{k+1} = \alpha_1 \lambda_1 u_1 + ... + \alpha_k \lambda_k u_k \\ &= \lambda_{k+1}\left(\frac{\alpha_1 \lambda_1}{\lambda_{k+1}} u_1 + ... + \frac{\alpha_k \lambda_k}{\lambda_{k+1}} u_k\right). \end{aligned}$$

Since $u_1, ..., u_k$ are linearly independent, we obtain $\lambda_1 = \lambda_{k+1}, ..., \lambda_k = \lambda_{k+1}$, and this is also a contradiction. □

It follows from this proposition that there are at most N eigenvalues of T, where N is the dimension of X.

Proposition 3. $\lim_{n\to\infty} \|T^n\|^{\frac{1}{n}}$ exists and is equal to $\inf_{n=1,2,...} \|T^n\|^{\frac{1}{n}}$.

Proof. It follows from (41) that

$$\|T^m T^n\| \le \|T^m\| \|T^n\|, \quad \|T^n\| \le \|T\|^n, \quad m,n = 0,1,... \tag{44}$$

Set $a_n = \log \|T^n\|$, what is to be proved is that

$$\lim_{n\to\infty} \frac{a_n}{n} = \inf_{n=1,2,\cdots} \frac{a_n}{n}. \tag{45}$$

The inequality (44) gives

$$a_{m+n} \le a_m + a_n.$$

Let $n = mq + r$, where q, r are nonnegative integers with $0 \le r < m$, then the last inequality gives

$$a_n \le a_{mq} + a_r.$$

Let $a_{mq} = \log \|T^{mq}\|$. By (44) $a_{mq} \le \log \|T^m\|^q = q \log \|T^m\| = qa_m$, hence

$$a_n \le qa_m + a_r$$

and

$$\frac{a_n}{n} \le \frac{q}{n} a_m + \frac{1}{n} a_r \ .$$

Therefore

$$\lim_{n\to\infty} \sup \frac{a_n}{n} \le \lim_{n\to\infty} \sup \frac{q}{n} a_m + \lim_{n\to\infty} \sup \frac{a_r}{n} \ .$$

Since $\lim_{n\to\infty} \sup \frac{q}{n} = \frac{1}{m}$ and $\lim_{n\to\infty} \sup \frac{a_r}{n} = 0$,

$$\lim_{n\to\infty} \sup \frac{a_n}{n} \le \frac{a_m}{m} \ .$$

Since this holds for all fixed m,

$$\lim_{n\to\infty} \sup \frac{a_n}{n} \le \inf \frac{a_m}{m} \ .$$

Obviously

$$\lim_{n\to\infty} \sup \frac{a_n}{n} \ge \inf \frac{a_m}{m} \ ,$$

hence

$$\lim_{n\to\infty} \sup \frac{a_n}{n} = \inf_{n=1,2,...} \frac{a_n}{n} \ .$$

□

Now we define:

Definition 10. (Spectral radius of T) spr$T$ =lim$\|T^n\|^{\frac{1}{n}}$ =inf$\|T^n\|^{\frac{1}{n}}$.

1.6 The resolvent

Let $T \in L(X)$ and consider the equation

$$(T-\xi)u = v,$$

where ξ is a given complex number, $v \in X$ is given and $u \in X$ is to be found. This equation has a solution u for every v if and only if $T-\xi$ is nonsingular, that is ξ is different from any eigenvalue λ_k of T. Then the inverse $(T-\xi)^{-1}$ exists and the solution u is given by

$$u = (T-\xi)^{-1}v.$$

The operator

$$R(\xi) = R(\xi, T) = (T-\xi)^{-1} \tag{46}$$

is called the resolvent of T. The complementary set of the spectrum $\Sigma(T)$ is called the resolvent set of T and will be denoted by $P(T)$. The resolvent $R(\xi)$ is thus defined for $\xi \in P(T)$.

Example 15. $R(\xi)$ commutes with T. And $R(\xi)$ has exactly the eigenvalues $(\lambda_h-\xi)^{-1}$ where λ_h are eigenvalues of T.
We show that $R(\xi)$ commutes with T:

$$\begin{aligned} T &= T \cdot 1 \\ &= T(T-\xi)(T-\xi)^{-1} \\ &= (T-\xi)T(T-\xi)^{-1}. \end{aligned}$$

Hence

$$\begin{aligned} (T-\xi)^{-1}T &= (T-\xi)^{-1}(T-\xi)T(T-\xi)^{-1} \\ &= T(T-\xi)^{-1} \ . \end{aligned}$$

Now we show that if λ is an eigenvalue of T, that is $Tu = \lambda u$, then $(\lambda-\xi)^{-1}$ is an eigenvalue of $R(\xi)$.
Clearly we have

$$(T-\xi)u = (\lambda-\xi)u \ ,$$

or equivalently

$$(\lambda-\xi)^{-1}(T-\xi)u = u\ .$$

Then

$$(T-\xi)^{-1}((\lambda-\xi)^{-1}(T-\xi)u) = (T-\xi)^{-1}u\ ,$$

i.e.

$$(\lambda-\xi)^{-1}((T-\xi)^{-1}(T-\xi)u) = (T-\xi)^{-1}u\ .$$

Hence

$$(\lambda-\xi)^{-1}u = (T-\xi)^{-1}u\ .$$

Note that the resolvent satisfies the (first) resolvent equation

$$R(\xi_1) - R(\xi_2) = (\xi_1-\xi_2)R(\xi_1)R(\xi_2) \tag{47}$$

since

$$\begin{aligned}(T-\xi_1)^{-1} - (T-\xi_2)^{-1} &= (T-\xi_1)^{-1}(T-\xi_2)^{-1}(T-\xi_2)(T-\xi_1)\cdot \\ &\quad [(T-\xi_1)^{-1} - (T-\xi_2)^{-1}] \\ &= (T-\xi_1)^{-1}(T-\xi_2)^{-1}[(T-\xi_2)-(T-\xi_1)] \\ &= (\xi_1-\xi_2)(T-\xi_1)^{-1}(T-\xi_2)^{-1} \\ &= (\xi_1-\xi_2)R(\xi_1)R(\xi_2).\end{aligned}$$

Here we have used the identity $(T-\xi_2)(T-\xi_1) = (T-\xi_1)(T-\xi_2)$.

We shall show that for each $\xi_0 \in P(T)$ $R(\xi)$ is holomorphic in some disk around ξ_0.

Proposition 4. $R(\xi) = \sum(\xi-\xi_0)^n R(\xi_0)^{n+1}$ is absolutely convergent for $|\xi-\xi_0| < (\text{spr } R(\xi_0))^{-1}$ where ξ_0 is a given complex number.

To prove this we have to study the following Lemmas.

Lemma 1. (Neumann series) The series $\sum_{n=0}^{\infty} T^n$ is absolutely convergent if $\|T\| < 1$. Moreover,

$$(1-T)^{-1} = \sum_{n=0}^{\infty} T^n,\ \text{and } \|(1-T)^{-1}\| \le (1-\|T\|)^{-1}, \text{where } T \in L(X)\ . \tag{48}$$

Proof. This series is absolutely convergent because $\lim_{n\to\infty} \|T^n\|^{\frac{1}{n}} \leq \|T\| < 1$. Set $\sum_{n=0}^{\infty} T^n = S$, then

$$\begin{aligned} TS &= \sum_{n=0}^{\infty} T^{n+1} \\ &= \sum_{n=1}^{\infty} T^n \\ &= \sum_{n=0}^{\infty} T^n - 1 \\ &= S - 1\ , \end{aligned}$$

so that $TS = ST = S - 1$. Hence $(1 - T)S = S(1 - T) = 1$ and $S = (1 - T)^{-1}$. Now we have $\|(1 - T)^{-1}\| = \|\sum_{n=0}^{\infty} T^n\| \leq \sum_{n=0}^{\infty} \|T^n\| \leq \sum_{n=0}^{\infty} \|T\|^n = (1 - \|T\|)^{-1}$. $\square$

Lemma 2. The series (48) is absolutely convergent if $\|T^m\| < 1$ for some positive integer m, or equivalently, if spr $T < 1$, and the sum is again equal to $(1 - T)^{-1}$ (see proof of Proposition 3).

Proof. Since spr $T = \inf_{n=1,2,\ldots} \|T^n\|^{\frac{1}{n}} = \lim_{n\to\infty} \|T^n\|^{\frac{1}{n}}$ and $\|T^m\|^{\frac{1}{m}} < 1$, it follows that $\lim_{n\to\infty} \|T^n\|^{\frac{1}{n}} < 1$ and the series is absolutely convergent. The proof that the sum is equal to $(1-T)^{-1}$ is the same as above (see the proof of Proposition 3). $\square$

Lemma 3.

$$S(t) = (1 - tT)^{-1} = \sum_{n=0}^{\infty} t^n T^n, \tag{49}$$

where t is a complex number. The convergence radius r of (49) is exactly equal to $1/\text{spr } T$.

Proof. By Lemma 2, (49) holds if $\text{spr}(tT) < 1$, i.e., $|t| < 1/\text{spr}T$, so the convergence radius $r \geq 1/\text{spr}T$. If $|t| > 1/\text{spr}T$, then $\text{spr}(tT) > 1$, so $\lim_{n\to\infty} \|t^n T^n\|^{\frac{1}{n}} > 1$ and the series diverges. Hence $r = 1/\text{spr}T$. $\square$

Now we can complete the proof of Proposition 4. By (47), with $\xi_1 = \xi$ and $\xi_2 = \xi_0$, we have $R(\xi) = R(\xi_0)(1 - (\xi - \xi_0)R(\xi_0))^{-1}$. Let $tT = (\xi - \xi_0)R(\xi_0)$ in Lemma 3. Then

$$R(\xi) = \sum_{n=0}^{\infty} (\xi - \xi_0)^n R(\xi_0)^{n+1}.$$

By Proposition 4 we obtain:

Proposition 5. $R(\xi)$ is holomorphic at ξ_0 in the disk $|\xi - \xi_0| < (\text{spr } T)^{-1}$.

Proposition 6. $R(\xi)$ is holomorphic at ∞.

Proof. $R(\xi)$ has the expansion

$$R(\xi) = -\xi^{-1}(1 - \xi^{-1}T)^{-1} = -\sum_{n=0}^{\infty} \xi^{-n-1}T^n, \tag{50}$$

which is convergent if and only if $|\xi| > \text{spr } T$, thus $R(\xi)$ is holomorphic at infinity. □

Example 16. $\|R(\xi)\| \leq (|\xi| - \|T\|)^{-1}$ and $\|R(\xi) + \xi^{-1}\| \leq |\xi|^{-1}(|\xi| - \|T\|)^{-1}\|T\|$, for $|\xi| > \|T\|$.

Proof. It follows from (50) that

$$\begin{aligned} \|R(\xi)\| &= \| - \sum \xi^{-n-1}T^n\| \\ &= |\xi|^{-1}\| \sum (T/\xi)^n\| \\ &\leq |\xi|^{-1} \sum |\xi|^{-n}\|T^n\| \\ &\leq |\xi|^{-1} \sum |\xi|^{-n}\|T\|^n \\ &= |\xi|^{-1}(1 - \|\xi^{-1}T\|)^{-1} \\ &= (|\xi| - \|T\|)^{-1}, \end{aligned}$$

and

$$\begin{aligned}
\|R(\xi)+\xi^{-1}\| &= \| -\sum_{n=0}^{\infty}\xi^{-n-1}T^n+\xi^{-1}\| \\
&= \|\sum_{n=1}^{\infty}\xi^{-n-1}T^n\| \\
&\le |\xi|^{-1}\sum_{n=1}^{\infty}|\xi|^{-n}\|T\|^n \\
&= |\xi|^{-1}[(1-|\xi|^{-1}\|T\|)^{-1}-1] \\
&= \frac{1}{|\xi|-\|T\|}-\frac{1}{|\xi|} \\
&= \frac{|\xi|-|\xi|+\|T\|}{|\xi|(|\xi|-\|T\|)} \\
&= |\xi|^{-1}(|\xi|-\|T\|)^{-1}\|T\|.
\end{aligned}$$

□

The spectrum $\Sigma(T)$ is never empty; T has at least one eigenvalue. Otherwise $R(\xi)$ would be an entire function such that $R(\xi)\to 0$ for $|\xi|\to\infty$, then we must have $R(\xi)=0$ by Liouville's theorem (see [3]). But this results in the contradiction that $1=(T-\xi)R(\xi)=0$.

We can see that each eigenvalue of T is a singularity of the analytic function $R(\xi)$. Since there is at least one singularity of $R(\xi)$ on the convergence circle $|\xi|=\text{spr}\ T$ according to (50), spr T coincides with the largest (in absolute value) eigenvalue of T:

$$\text{spr}\ T=\max|\lambda_h|. \tag{51}$$

This shows that spr T is independent of the norm used in its definition.

2 Operators in unitary spaces

2.1 Unitary spaces

A normed space X is a special case of a linear metric space in which the distance between any two points is defined by $\|u-v\|$, where u and v belong to X.

Definition 11. (The complex inner product) Let $u, v \in X$, and let (u, v) be a complex number, then we say that the function $(\ ,\ )$ is a complex inner product if it satisfies:

-$(\alpha u_1 + \beta u_2, v) = \alpha(u_1, v) + \beta(u_2, v)$.
-$(u, v) = \overline{(v, u)}$.
-$(u, u) > 0$, if $u \neq 0$.

From the second condition in the last definition we can obtain

$$(u, kv) = \bar{k}(u, v),$$

since $(u, kv) = \overline{(kv, u)} = \bar{k}\overline{(v, u)} = \bar{k}(u, v)$.

Definition 12. A normed space H is called a unitary space if an inner product (u, v) is defined for all vectors $u, v \in H$.

Definition 13. In a unitary space the function

$$\|u\| = (u, u)^{\frac{1}{2}} \tag{52}$$

is a norm which is called the unitary norm.

We shall show the conditions in the Definition 2:

-the first condition follows directly from the definition of the inner product.

-$\|\alpha u\| = (\alpha u, \alpha u)^{\frac{1}{2}} = [\alpha(u, \alpha u)]^{\frac{1}{2}} = [\alpha\bar{\alpha}(u, u)]^{\frac{1}{2}} = |\alpha|\|u\|$.

-$\|u + v\| = (u + v, u + v)^{\frac{1}{2}} = [(u, u + v) + (v, u + v)]^{\frac{1}{2}} = [(u, u) + (u, v) + (v, u) + (v, v)]^{\frac{1}{2}} \leq [(u, u) + |(u, v)| + |(v, u)| + (v, v)]^{\frac{1}{2}}$. By the Schwarz inequality

$$|(u, v)| \leq \|u\|\|v\| \ , \tag{53}$$

hence $\|u+v\| \leq [\|u\|^2 + |(u, v)| + |(v, u)| + \|v\|^2]^{\frac{1}{2}} \leq [\|u\|^2 + 2\|u\|\|v\| + \|v\|^2]^{\frac{1}{2}} = \|u\| + \|v\|$.

Example 17. For numerical vectors $u = (\xi_1, ..., \xi_N)$ and $v = (\eta_1, ..., \eta_N)$ set

$$(u, v) = \sum \xi_j \overline{\eta_j}, \quad \|u\|^2 = \sum |\xi_j|^2,$$

with this inner product the space C^N becomes a unitary space.

Remark 2. A characteristic property of a unitary space H is that the adjoint space H^* can be identified with H itself.
To show that, let $f, u \in H$, we have the form (f, u) is linear in f and semilinear in u by Definition 11. Then $f \in H^*$ by (15) and (16). Hence f can be considered as a vector in H or a vector in H^*. Thus H and H^* can be identified.

Definition 14. (Orthogonal) If $(u, v) = 0$ we write $u \perp v$ and say that u, v are mutually orthogonal.

If S, S' are subsets of H we say

$$u \perp S \quad \text{if } u \perp v,\ \forall v \in S \text{ (where } u \in H),$$

$$S \perp S' \quad \text{if } u \perp v,\ \forall u \in S,\ \forall v \in S'.$$

The set of all $u \in H$ such that $u \perp S$ is denoted by $S^{\perp}$.

Example 18. $u \perp S$ implies $u \perp M$ where M is the span of S.
Let $v \in M$, then there are $v_1, \cdots, v_k \in S$ and $\alpha_1, ..., \alpha_k \in C$ such that $v = \alpha_1 v_1 + \cdots + \alpha_k v_k$. So

$$\begin{aligned}
(u, v) &= (u, \alpha_1 v_1 + \cdots + \alpha_k v_k) \\
&= (u, \alpha_1 v_1) + \cdots + (u, \alpha_k v_k) \\
&= \bar{\alpha}_1 (u, v_1) + \cdots + \bar{\alpha}_k (u, v_k) \\
&= \bar{\alpha}_1 \cdot 0 + \cdots + \bar{\alpha}_k \cdot 0 \\
&= 0,
\end{aligned}$$

thus $u \perp S$.

Let $\dim H = N$. If $x_1, ..., x_N \in H$ have the property

$$(x_j, x_k) = \delta_{jk} \ , \tag{54}$$

then they form a basis of H, called an orthonormal basis, for $\alpha_1 x_1 + ... + \alpha_N x_N = 0$, implies $(\alpha_1 x_1 + ... + \alpha_N x_N, x_j) = 0$ and $\alpha_j (x_j, x_j) = 0$ for all $j = 1, ..., N$, hence $\alpha_j = 0$, showing that $x_1, ..., x_N$ are linearly independent.

2.2 Symmetric operators

Definition 15. (Sesquilinear form) Let H, H' be two unitary spaces. A complex-valued function $t[u, u']$ defined for $u \in H$ and $u' \in H'$ is called a sesquilinear form on $H \times H'$ if it is linear in u and semilinear in u'.

If $H' = H$ we speak of a sesquilinear form on H.

Let T be a linear operator on H to H', the function

$$t[u, u'] = (Tu, u') \tag{55}$$

is a sesquilinear form on $H \times H'$. Conversely, an arbitrary sesquilinear form $t[u, u']$ on $H \times H'$ can be expressed in this form by a suitable choice of an operator T on H to H'. Since $t[u, u']$ is a semilinear form on H' for a fixed u, there exists a unique $w' \in H'$ such that $t[u, u'] = (w', u')$ for all $u' \in H'$. Since w' is determined by u, we define a function T by setting $w' = Tu$. T is a linear operator on H to H'. In the same way, $t[u, u']$ can also be expressed in the form

$$t[u, u'] = (u, T^*u'). \tag{56}$$

Since H^*, H'^* can be identified with H, H' respectively, T^* can be considered as the adjoint of T on H' to H.

T^*T is a linear operator on H to itself. The relation

$$(u, T^*Tv) = (T^*Tu, v) = (Tu, Tv) \tag{57}$$

shows that T^*T is the operator associated with the sesquilinear form (Tu, Tv) on H. Note that the first two members of (57) are the inner product in H while the last is that in H'.

It follows from (57),(40) that $\|T^*T\| = \sup \dfrac{|(Tu, Tv)|}{\|u\|\|v\|} \geq \sup \dfrac{\|Tu\|^2}{\|u\|^2} = \|T\|^2$. By (41) and (42) we have $\|T^*T\| \leq \|T^*\|\|T\| = \|T\|^2$. Hence

$$\|T^*T\| = \|T\|^2 \ . \tag{58}$$

Example 19. If T is an operator on H to itself, $(Tu, u) = 0$ for all u implies $T = 0$.

Proof. We have $(T(u+v), u+v) = 0$. On the other hand, $(T(u+v), u+v) = (Tu, u) + (Tu, v) + (Tv, u) + (Tv, v) = (Tu, v) + (Tv, u) = 0$. Since v is any vector in H, then $(Tu, iv) + (Tiv, u) = -i(Tu, v) + i(Tv, u) = 0$, that is $-(Tu, v) + (Tv, u) = 0$. Combining this equality with $(Tu, v) + (Tv, u) = 0$ yields $(Tv, u) = 0, \forall u, \forall v$. Hence $T = 0$.

□

Remark 3. This property is not true when T is defined on a real space. To show that we can take T represented by the matrix

$$\tau_{jk} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},$$

then $(Tu, u) = 0$ for all u, but $T \neq 0$.

Definition 16. A sesquilinear form $t[u, v]$ (or t in short) on a unitary space H is said to be symmetric if

$$t[v, u] = \overline{t[u, v]}, \text{ for all } u, v \in H. \tag{59}$$

If t is symmetric, $t[u, u]$ is real-valued, and is denoted by $t[u]$. t is nonnegative if $t[u] \geq 0$ for all u, and positive if $t[u] > 0$ for all $u \neq 0$.

The operator T associated with a symmetric form $t[u, v]$ according to (55),(56) has the property that

$$T^* = T. \tag{60}$$

Indeed, $t[v, u] = (Tv, u) = (v, T^*u)$ and $\overline{t[u, v]} = \overline{(Tu, v)} = (v, Tu)$. Hence $T^* = T$.

Definition 17. An operator T on H to itself satisfying (60) is said to be symmetric.

(Tu, u) is real for all $u \in H$ if and only if T is symmetric. Indeed, if (Tu, u) is real, then $(Tu, u) = \overline{(u, Tu)} = (u, Tu) = (T^*u, u)$, hence by Example 19 $T^* = T$. Conversely if $T^* = T$, then $(Tu, u) = (T^*u, u) = \overline{(u, T^*u)} = \overline{(Tu, u)}$, thus (Tu, u) is real.

A symmetric operator T is nonnegative (positive) if the associated form is nonnegative (positive), and we write $T \geq 0$ to denote that T is nonnegative symmetric. More generally, we write

$$T \geq S \text{ or } \ S \leq T$$

if S, T are symmetric operators such that $T - S \geq 0$.

Example 20. If T is symmetric, $P(T)$ is symmetric for any polynomial P with real coefficients.

Since $(\alpha TS)^* = \overline{\alpha} S^* T^*$, it follows that $T^{n*} = T^{*n} = T^n$, thus

$$(\alpha_n T^n + ... + \alpha_0)^* = \bar{\alpha}_n T^{n*} + ... + \bar{\alpha}_0 = \alpha_n T^n + ... + \alpha_0.$$

Example 21. For any linear operator T on H to H', T^*T and TT^* are nonnegative symmetric operators in H and H', respectively.
We have

$$(T^*T)^* = T^*(T^*)^* = T^*T, \ (T^*Tu, u) = (Tu, Tu) \geq 0$$

and

$$(TT^*)^* = (T^*)^*T^* = TT^*, \ (TT^*u', u') = (T^*u', T^*u') \geq 0.$$

Example 22. If T is symmetric, then $T^2 \geq 0$; $T^2 = 0 \Leftrightarrow T = 0$.
$(T^2u, u) = (Tu, T^*u) = (Tu, Tu) \geq 0$; $T^2 = 0$, that is $(T^2u, u) = (Tu, Tu) = 0$ according to Example 19, and this is equivalent to $T = 0$.

Example 23. $R \leq S$ and $S \leq T$ imply $R \leq T$. $S \leq T$ and $S \geq T$ imply $S = T$.

$R \leq S$ and $S \leq T$ is equivalent to $S - R \geq 0$ and $T - S \geq 0$, therefore $((S - R)u, u) \geq 0$ and $((T - S)u, u) \geq 0$, hence $(Su, u) - (Ru, u) \geq 0$ and $(Tu, u) - (Su, u) \geq 0$. Adding the last two inequalities, we obtain $(Tu, u) - (Ru, u) \geq 0$. Hence $((T - R)u, u) \geq 0$, that is, $T - R \geq 0$.
$S \leq T$ and $S \geq T$ is equivalent to $S - T \leq 0$ and $S - T \geq 0$, and this implies that $S - T = 0$, that is $(Su, u) - (Tu, u) = 0$ for all $u \in H$. Thus $S = T$ by Example 19.

2.3 Unitary, isometric and normal operators

Definition 18. Let H and H' be unitary spaces. An operator T on H to H' is said to be isometric if

$$\|Tu\| = \|u\| \quad \text{for every} \quad u \in H. \tag{61}$$

This is equivalent to $(T^*Tu, u) = (Tu, Tu) = (u, u)$, thus

$$T^*T = 1. \tag{62}$$

This implies that

$$(Tu, Tv) = (u, v) \text{ for every } u, v \in H. \tag{63}$$

Definition 19. An isometric operator T is said to be unitary if the range of T is the whole space H'.

Example 24. $T \in L(H, H')$ is unitary if and only if $T^{-1} \in L(H', H)$ exists and

$$T^{-1} = T^* \ . \tag{64}$$

If T is unitary, $\|Tu\| = \|u\|$ implies that the mapping T is one-to-one, that is, T^{-1} exists. Since $T^*T = 1$, $T^{-1} = T^*$. If T^{-1} exists and $T^{-1} = T^*$, then $T^*T = 1$, hence T is unitary.

Example 25. T is unitary $\Leftrightarrow T^*$ is.

We have $T^{-1*} = T^{*-1}$ and by Example 24 this is equivalent to $(T^*)^* = (T^*)^{-1}$, which is equivalent to T^* being unitary.

Example 26. If $T \in L(H', H'')$ and $S \in L(H, H')$ are isometric, $TS \in L(H, H'')$ is isometric. The same is true if "isometric" is replaced by "unitary".

We have $S^*S = 1$ and $T^*T = 1$, therefore $(TS)^*TS = S^*(T^*T)S = S^*S = 1$, and if the range of S is the whole space H', and the range of T is the whole space H'', this implies that the range of TS is the whole of H''.

Definition 20. $T \in L(H)$ is said to be normal if T and T^* commute, i.e.

$$T^*T = TT^*. \tag{65}$$

Symmetric operators and unitary operators on a unitary space into itself are special cases of normal operators.

Example 27. Let T_1, T_2, T_3 be operators represented by the matrices

$$\begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & i \\ 0 & 1 & 0 \\ i & 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & i & 0 \\ i & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

respectively, then T_1 is normal, symmetric but not unitary, T_2 is normal, not symmetric but unitary, T_3 is normal, neither symmetric nor unitary.

An important property of a normal operator T is that

$$\|T^n\| = \|T\|^n, \ n = 1, 2, ... \tag{66}$$

This implies that

$$\text{spr } T = \|T\|. \tag{67}$$

We shall prove (66). If T is symmetric, by (58) $\|T^2\| = \|T\|^2$. Since T^2 is symmetric we have $\|T^4\| = \|T^2\|^2 = \|T\|^4$. Proceeding in this manner we obtain $\|T^n\| = \|T\|^n$ for $n = 2^m, m = 1, 2, \cdots$. If T is normal but not necessarily symmetric, again by (58) we have $\|T^n\|^2 = \|T^{n*}T^n\|$. Since $T^{n*}T^n = (T^*T)^n$ because T is normal, and T^*T is symmetric, by (58) $\|T^n\|^2 = \|(T^*T)^n\| = \|T^*T\|^n = \|T\|^{2n}$ for $n = 2^m$. Now if $2^m - n = r \geq 0$, then $\|T\|^n\|T\|^r = \|T\|^{n+r} = \|T^{n+r}\| \leq \|T^n\|\|T^r\| \leq \|T^n\|\|T\|^r$. Thus $\|T\|^n \leq \|T^n\|$. The opposite inequality is obvious. This proves (66).

Example 28. If T is normal then (i) $T^n = 0$ for some integer n implies $T = 0$. (ii)If T is nonsingular then T^{-1} is normal.
(i) $T^n = 0 \Rightarrow \|T^n\| = 0 = \|T\|^n \Rightarrow \|T\| = 0 \Rightarrow T = 0$.
(ii) We have $T^*T = TT^*$ therefore $(T^*T)^{-1} = (TT^*)^{-1}$ therefore $T^{-1}T^{*-1} = T^{*-1}T^{-1}$, since $T^{*-1} = T^{-1*}$, therefore $T^{-1}T^{-1*} = T^{-1*}T^{-1}$.

2.4 Orthogonal projections

Definition 21. Let M be a subspace of H and $H = M \oplus M^{\perp}$, the projection operator $P = P_M$ on M along $M^{\perp}$ is called the orthogonal projection on M.

P is a symmetric, nonnegative and idempotent operator, for

$$(Pu, u) = (u', u' + u'') = (u', u') \geq 0 \tag{68}$$

where

$$u = u' + u'',\ u \in H,\ u' \in M,\ u'' \in M^{\perp}, \tag{69}$$

by $u' \perp u''$. (68) shows also that (Pu, u) is real . Hence P is symmetric, i.e. $P^* = P$. Idempotent follows from P being a projection.

Example 29. If P is an orthogonal projection, then $1 - P$ is a symmetric, nonnegative and idempotent, and we have

$$0 \leq P \leq 1,\ \|P\| = 1 \text{ if } \ P \neq 0.$$

$(1 - P)^* = 1^* - P^* = 1 - P, ((1 - P)u, u) = (u - u', u) = (u'', u' + u'') = (u'', u'') \geq 0$, and $(1 - P)^2 = 1 - 2P + P^2 = 1 - 2P + P = 1 - P$. Now $((1 - P)u, u) \geq 0$, therefore $1 - P \geq 0$, and $(Pu, u) \geq 0$, therefore $P \geq 0$. Hence $0 \leq P \leq 1$. Since $P^2 = P$ and P is symmetric, $\|P\| = \|P^2\| = \|P\|^2$ by (58), or equivalently $\|P\|(1 - \|P\|) = 0$. But $\|P\| \neq 0$ for $P \neq 0$, therefore $\|P\| = 1$.

Example 30. $\|(1 - P_M)u\| =\text{dist}(u, M), u \in H$.
$\text{dist}(u, M) = \inf_{v \in M} \|u - v\|$. Since $\|u - v\|^2 = \|(u' - v') + u''\|^2 = \|u' - v'\|^2 + \|u''\|^2$, the infimum is attained as $v' = u'$, and is equal to $\|u''\| = \|(1 - P_M)u\|$.

Example 31. 1. $M \perp N \Leftrightarrow P_M P_N = 0$.

2. The following three conditions are equivalent:
(i)$M \supset N$,
(ii)$P_M \geq P_N$,
(iii)$P_M P_N = P_N$.

Proof. $H = M \oplus M^{\perp}$.
(1) We have $M \perp N$, therefore for each $u \in H, 0 = (P_M u, P_N u) = (u, P_M^* P_N u) = (u, P_M P_N u)$. Hence $P_M P_N = 0$. Conversely if $P_M P_N = 0$, then $P_M u = P_M P_N u = 0$ for all $u \in N$. Thus $M \perp N$.
(2) Now we will show that (i)$\Leftrightarrow$(ii), and (i)$\Leftrightarrow$(iii).
If $M \supset N$, then $H = N \oplus N_1 \oplus M^{\perp}$, where N_1 is the orthogonal complement of N in M. Let $u = u' + u'' + u''', u' \in N, u'' \in N_1, u''' \in M^{\perp}$.

(i)⇒(ii): $((P_M - P_N)u, u) = (P_M u, u) - (P_N u, u) = (u' + u'', u) - (u', u) = (u'', u) = (u'', u'') \geq 0$. (ii)⇒(i): Let $u \in N$, then $(P_M u, u) \geq (P_N u, u) = (u, u)$, hence $((I - P_M)u, u) \leq 0$. Since $I - P_M$ is the orthogonal projection on $M^\perp, I - P_M \geq 0$ so $((I - P_M)u, u) = 0$ which means that $u \in M$.
(i)⇒(iii) Since $P_N u \in N \subset M, P_M P_N u = P_N u$. (iii)⇒(i) Let $u \in N$. Then $P_M P_N u = P_N u$ by (iii), and $P_N u = u$. So $P_M u = u$, i.e., $u \in M$. □

2.5 The eigenvalue problem

Example 32. A symmetric operator has only real eigenvalues.
We have proved that $T^* = T$ is equivalent to (Tu, u) being real, and this implies that $(\lambda u, u)$ is real, for an eigenvector u with eigenvalue λ. Thus λ is real.

Example 33. Each eigenvalue of a unitary operator has absolute value one. A normal operator with this property is unitary.
Since $\|Tu\| = \|u\| \; \forall u \in H$, $Tu = \lambda u$ implies $\|\lambda u\| = \|u\|$, that is $|\lambda| = 1$.
T being normal is equivalent to $(T^*Tu, u) = (TT^*u, u)$. This is equivalent to $(Tu, Tu) = (T^*u, T^*u)$, therefore we have $(T^*u, T^*u) = (\lambda u, \lambda u) = |\lambda|^2(u, u) = (u, u)$. On the other hand $(T^*u, T^*u) = (TT^*u, u)$, hence $TT^* = 1$. Since $T \in L(H)$ and is nonsingular, the range of T is the whole space H.

Example 34. If T is normal, then we have

1. $Tu = 0$ if and only if $T^*u = 0$.
2. $(T - \lambda I)$ is a normal.
3. The eigenvalues of T^* are the conjugates of the eigenvalues of T.
4. Any two eigenvectors that belong to different eigenvalues are orthogonal.

Proof. 1. We will show that $(Tu, Tu) = (T^*u, T^*u)$.
$(Tu, Tu) = (u, T^*Tu) = (u, TT^*u) = (T^*u, T^*u)$, so that $Tu = 0$ is equivalent to $T^*u = 0$.

2. We have to show that $T - \lambda I$ commutes with its adjoint.

$$\begin{aligned}(T - \lambda I)(T - \lambda I)^* &= (T - \lambda I)(T^* - \bar{\lambda} I) = TT^* - \lambda T^* - \bar{\lambda} T + \lambda\bar{\lambda} I \\ &= T^*T - \bar{\lambda} T - \lambda T^* + \lambda\bar{\lambda} I = (T^* - \bar{\lambda} I)(T - \lambda I) \\ &= (T - \lambda I)^*(T - \lambda I),\end{aligned}$$

hence $T - \lambda I$ is a normal.

3. If $Tu = \lambda u$, $(T - \lambda I)u = 0$. Since $T - \lambda I$ is normal we have $(T - \lambda I)^* u = 0$, so $(T^* - \bar{\lambda} I)u = 0$. Thus $T^* u = \bar{\lambda} u$. Hence any eigenvector of T is also an eigenvector of T^*, and the corresponding eigenvalues are conjugate to each other.

4. Let $Tu_1 = \lambda_1 u_1, Tu_2 = \lambda_2 u_2$ and $\lambda_1 \neq \lambda_2$, then

$$\begin{aligned}
\lambda_1(u_1, u_2) &= (\lambda_1 u_1, u_2) \\
&= (Tu_1, u_2) \\
&= (u_1, T^* u_2) \\
&= (u_1, \bar{\lambda}_2 u_2) \\
&= \lambda_2(u_1, u_2).
\end{aligned}$$

Since $\lambda_1 \neq \lambda_2$, $(u_1, u_2) = 0$.

$\square$

2.6 The minimax principle

Let T be a symmetric operator in H. T is diagonalizable (see [2]) and has only real eigenvalues. For a subspace M of H, set

$$\mu[M] = \mu[T, M] = \min_{\substack{u \in M \\ \|u\|=1}} (Tu, u) = \min_{0 \neq u \in M} \frac{(Tu, u)}{\|u\|^2} \ . \tag{70}$$

Arrange the eigenvalues of T in the following order

$$\mu_1 \leq \mu_2 \leq ... \leq \mu_N,$$

where μ_i can be repeated. The minimax (or maximin) principle is

$$\mu_n = \max_{\text{codim} M \ = n-1} \mu[M] = \max_{\text{codim} M \ \leq n-1} \mu[M]. \tag{71}$$

This is equivalent to the following two propositions:

$$\mu_n \geq \mu[M] \text{ for any } \ M \text{ with codim} M \ \leq n - 1; \tag{72}$$

$$\mu_n \leq \mu[M_0] \text{ for some } \ M_0 \text{ with codim} M \ = n - 1. \tag{73}$$

Let us prove these separately.

Let $\{\varphi_n\}$ be an orthonormal basis with the property

$$T\varphi_n = \mu_n \varphi_n, \; n = 1, ..., N. \tag{74}$$

Each u has the expansion

$$u = \sum_{n=1}^{N} \xi_n \varphi_n, \; \xi_n = (u, \varphi_n), \; \|u\|^2 = \sum_{n=1}^{N} |\xi_n|^2,$$

in this basis. Then

$$Tu = \sum_{n=1}^{N} \xi_n T\varphi_n = \sum_{n=1}^{N} \mu_n \xi_n \varphi_n, \; (Tu, u) = \sum_{n=1}^{N} \mu_n |\xi_n|^2.$$

Let M be any subspace with codim$M \leq n-1$. The n-dimensional subspace M' spanned by $\varphi_1, ..., \varphi_n$ contains a nonzero vector u in common with M by (7). This u has the coefficients $\xi_{n+1}, \xi_{n+2}, ...$ equal to zero, so that

$$\begin{aligned}(Tu, u) &= \sum_{1}^{N} \mu_k |\xi_k|^2 \\ &= \sum_{1}^{n} \mu_k |\xi_k|^2 \\ &\leq \mu_n \sum_{1}^{n} |\xi_k|^2 \\ &= \mu_n \|u\|^2.\end{aligned}$$

Hence $\mu[M] = \min\limits_{0 \neq u \in M} \dfrac{(Tu, u)}{\|u\|^2} \leq \mu_n$. This proves (72).

Let M_0 be the subspace consisting of all vectors orthogonal to $\varphi_1, ..., \varphi_{n-1}$, so that codim$M_0 = n-1$. Each $u \in M_0$ has the coefficients $\xi_1, ..., \xi_{n-1}$ zero. Hence

$$\begin{aligned}(Tu, u) &= \sum_{k=n}^{N} \mu_k |\xi_k|^2 \\ &\geq \sum_{k=n}^{N} \mu_n |\xi_k|^2 \\ &= \mu_n \|u\|^2,\end{aligned}$$

which implies (73).

References

[1] T. Kato, *Perturbation Theory for Linear Operators*, Springer, Berlin, 1976.

[2] G. Strang, *Linear Algebra and its Applications*, Academic Press, New York, 1970.

[3] G. Jones, D. Singerman, *Complex Functins*, University Press, New York, 2002.