3. FINITE DIMENSIONALITY

In any *numerical approximation* process, we deal solely with finite dimensional subspaces and with operators whose ranges are finite dimensional. In this section we study such subspaces and operators.

We start with a result concerning the closedness of the sum of two closed subspaces of a complex Banach space X. In general, such a sum need not be a closed subspace, as can be seen by considering $X = \ell^2$, F_1 = the closed linear span of $\{e_{2n} : n = 1,2,\ldots\}$, i.e., $\left\{\sum_{n=1}^{\infty} a_n e_{2n} : a_n \in \mathbb{C}, \sum_{n=1}^{\infty} |a_n|^2 < \infty\right\}$ and F_2 = the closed linear span of $\{e_{2n} + \frac{1}{n} e_{2n+1} : n = 1,2,\ldots\}$, i.e., $\left\{\sum_{n=1}^{\infty} b_n(e_{2n} + \frac{1}{n} e_{2n+1}) : b_n \in \mathbb{C}, \sum_{n=1}^{\infty} |b_n|^2 < \infty\right\}$. Then $\sum_{n=1}^{j} \frac{e_{2n+1}}{n}$ belongs to $F_1 + F_2$ for each $j = 1,2,\ldots$, but $\sum_{n=1}^{\infty} \frac{e_{2n+1}}{n}$ does not. However, if one of the summands is finite dimensional, we have the following result.

PROPOSITION 3.1 Let Y be a finite dimensional subspace and Z be a closed subspace of X. Then $Y + Z = \{y + z : y \in Y, z \in Z\}$ is a closed subspace of X. In particular, Y itself is closed in X.

Proof Assume first that Y is one dimensional, say $Y = \mathrm{span}\{y_1\}$. If $y_1 \in Z$, then $Y + Z = Z$, which is given to be closed. If $y_1 \notin Z$, let

$$d = \mathrm{dist}(y_1, Z) > 0 .$$

Consider a sequence $(\alpha_n y_1 + z_n)$ in $Y + Z$, which converges to x in X. Now, for every $z \in Z$, we have

$$|\alpha_n| d \le \|\alpha_n y_1 + z\| . \tag{3.1}$$

This is obvious if $\alpha_n = 0$, and if $\alpha_n \neq 0$, then $-z/\alpha_n$ is in Z so that $d \leq \|y_1-(-z/\alpha_n)\|$. Since $(\alpha_n y_1+z_n)$ is a Cauchy sequence, it follows from (3.1) that (α_n) is also Cauchy. Let $\alpha_n \to \alpha \in \mathbb{C}$. Then $z_n \to x - \alpha y_1$ which belongs to Z , since Z is closed. Thus,

$$x = \alpha y_1 + (x-\alpha y_1) \in Y + Z ,$$

showing that $Y + Z$ is closed.

If Y is of dimension $m < \infty$, and $\{y_1,\dots,y_m\}$ is a basis for Y , then a repeated application of the above result to Z , $\text{span}\{y_1,Z\},\dots,\text{span}\{y_1,\dots,y_{m-1},Z\}$ shows that $Y + Z$ is closed.

In particular, if we take $Z = \{0\}$, then we see that $Y + Z = Y$ is closed. //

We are now in a position to prove a result regarding the complementation of finite dimensional subspaces, which was promised in the last section.

THEOREM 3.2 Let Y be an m dimensional subspace of X , and let $x_1,\dots,x_m$ form an ordered basis for Y .

Then there exist $x_1^*,\dots,x_m^*$ in X^* such that

(3.2) $$\langle x_j^*,x_i\rangle = \delta_{i,j} , \quad i,j = 1,\dots,m .$$

The map

(3.3) $$Px = \sum_{j=1}^{m} \langle x,x_j^*\rangle x_j , \quad x \in X ,$$

is a projection on Y along $Z \equiv \bigcap_{j=1}^{m} Z(x_j^*)$, so that $X = Y \oplus Z$.

Also,

(3.4) $$P^*x^* = \sum_{j=1}^{m} \langle x^*,x_j\rangle x_j^* , \quad x^* \in X^* .$$

If $X = Y \oplus \tilde{Z}$, then there exist unique $\tilde{x}_1^*, \ldots, \tilde{x}_m^* \in \tilde{Z}^\perp$ which satisfy $\langle \tilde{x}_j^*, x_i \rangle = \delta_{i,j}, i,j = 1,\ldots,m$. They form the <u>ordered basis of $\tilde{Z}^\perp$ which is adjoint to the given ordered basis</u> $x_1,\ldots,x_m$ <u>of</u> Y.

Proof Let $Y_j = \text{span}\{x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_m\}$ for $j = 1,\ldots,m$. Then by Proposition 3.1, Y_j is a closed subspace of X and $x_j \notin Y_j$. By Corollary 1.2, there is $x_j^* \in X^*$ such that $x_j^* \in Y_j^\perp$ and $\langle x_j^*, x_j \rangle = 1$ for each $j = 1,\ldots,m$. These $x_1^*,\ldots,x_m^*$ satisfy (3.2).

The map P given by (3.3) is clearly linear and continuous; it is a projection since $Px_j = x_j$ for $j = 1,\ldots,m$ by (3.2), so that

$$P^2x = \sum_{j=1}^{m} \langle x, x_j^* \rangle Px_j = \sum_{j=1}^{m} \langle x, x_j^* \rangle x_j = Px .$$

Also, $R(P) = \text{span}\{x_1,\ldots,x_m\} = Y$, and since $x_1,\ldots,x_m$ are linearly independent, we have $Z(P) = \bigcap_{j=1}^{m} Z(x_j^*)$.

Next, for $x^* \in X^*$ and all $x \in X$, we have

$$\begin{aligned}\langle P^*x^*, x \rangle &= \langle x^*, Px \rangle \\ &= \langle x^*, \sum_{j=1}^{m} \langle x, x_j^* \rangle x_j \rangle \\ &= \sum_{j=1}^{m} \langle x_j^*, x \rangle \langle x^*, x_j \rangle \\ &= \langle \sum_{j=1}^{m} \langle x^*, x_j \rangle x_j^*, x \rangle .\end{aligned}$$

Hence we obtain (3.4).

Now, let $X = Y \oplus \tilde{Z}$ and let $\tilde{P}$ be the projection on Y along $\tilde{Z}$. Let $\tilde{x}_j^* = \tilde{P}^* x_j^*$ for $j = 1,\ldots,m$. Then for $i,j = 1,\ldots,m$,

$$\begin{aligned}\langle x_i, \tilde{x}_j^* \rangle &= \langle \tilde{P}x_i, x_j^* \rangle = \langle x_i, x_j^* \rangle = \delta_{i,j} , \\ \langle y, \tilde{x}_j^* \rangle &= \langle \tilde{P}y, x_j^* \rangle = 0 \quad \text{for all } y \in Z(\tilde{P}) = \tilde{Z} .\end{aligned}$$

Thus, $\tilde{x}_1^*,\ldots,\tilde{x}_m^*$ form a linearly independent set in $\tilde{Z}^\perp$. Since $\tilde{Z}^\perp$ is isomorphic to Y^* (Proposition 2.2), it has the same dimension as Y , viz., m . This shows that $\tilde{x}_1^*,\ldots,\tilde{x}_m^*$ form a basis of $\tilde{Z}^\perp$. //

It follows from (3.3) and (3.4) that

(3.5) $$\|P\| = \|P^*\| \leq \sum_{j=1}^{m} \|x_j^*\| \|x_j\| .$$

If $m = 1$, then we have

$$\|Px\| = |\langle x,x_1^*\rangle| \|x_1\| , \quad x \in X ,$$

so that

$$\|P\| = \sup\{\|Px\| : x \in X, \|x\| \leq 1\}$$

$$= \|x_1^*\| \|x\| .$$

If $m > 1$, then strict inequality can hold in (3.5). This will be clear from the examples we shall soon give.

Remark 3.3 Here is a result which is 'dual' to the first part of Theorem 3.2: Let $\{x_1^*,\ldots,x_m^*\}$ be a linearly independent subset of X^* . Then there exist $x_1,\ldots,x_m$ in X such that

$$\langle x_j^*,x_i\rangle = \delta_{i,j} , \quad i,j = 1,\ldots,m .$$

This is an immediate consequence of the following: If y^* , $y_1^*,\ldots,y_n^*$ are in X^* and $Z(y^*) \supset \bigcap_{j=1}^{n} Z(y_j^*)$, then $y^* \in \text{span}\{y_1^*,\ldots,y_n^*\}$. In fact, consider the conjugate linear function $F : X \to \mathbb{C}^n$ given by

$$Fx = [\langle y_1^*,x\rangle,\ldots,\langle y_n^*,x\rangle]^t , \quad x \in X .$$

If $Fx = F\tilde{x}$, then $\langle y^*, x\rangle = \langle y^*, \tilde{x}\rangle$. Hence we see that there is a linear map $A : \mathbb{C}^n \to \mathbb{C}$ such that $y^* = AF$. Let $\alpha_1, \ldots, \alpha_n \in \mathbb{C}$ be such that for all $c(1), \ldots, c(n)$ in $\mathbb{C}$,

$$A[c(1), \ldots, c(n)]^t = \alpha_1 c(1) + \ldots + \alpha_n c(n) .$$

Then for every $x \in X$,

$$\begin{aligned}\langle y^*, x\rangle &= A(Fx) \\ &= \alpha_1 \langle y_1^*, x\rangle + \ldots + \alpha_n \langle y_n^*, x\rangle \\ &= \langle \sum_{i=1}^{n} \alpha_i y_i^* , x\rangle .\end{aligned}$$

Thus, $y^* = \sum_{i=1}^{n} \alpha_i y_i^* \in \text{span}\{y_1^*, \ldots, y_n^*\}$.

Remark 3.4 Let $x_1, \ldots, x_m$ be in X and $x_1^*, \ldots, x_m^*$ be in X^* such that the matrix

$$A = [a_{i,j}] , \quad a_{i,j} = \langle x_i, x_j^*\rangle , \quad i,j = 1, \ldots, m ,$$

is invertible. Let its inverse be given by $B = [b_{i,j}]$. Then $\{x_1, \ldots, x_m\}$ is a linearly independent set in X and

(3.6) $$y_j^* = \sum_{k=1}^{m} \bar{b}_{k,j} x_k^* , \quad j = 1, \ldots, m$$

satisfies

$$\langle x_i, y_j^*\rangle = \delta_{i,j} , \quad i,j = 1, \ldots, m .$$

This can be seen as follows. Let $\alpha_1 x_1 + \ldots + \alpha_m x_m = 0$ for some $\alpha_i \in \mathbb{C}$, $i = 1, \ldots, m$. Then

$$\sum_{i=1}^{m} \alpha_i \langle x_i, x_j^*\rangle = 0 , \quad j = 1, \ldots, m .$$

Since the matrix A is invertible, the above system has a unique solution, namely $\alpha_1 = \ldots = \alpha_m = 0$. Thus, $\{x_1,\ldots,x_m\}$ is linearly independent in X . Again, since AB is the identity matrix, we see that

$$\sum_{k=1}^{m} a_{i,k}b_{k,j} = \delta_{i,j} .$$

But for $i,j = 1,\ldots,m$,

$$\sum_{k=1}^{m} a_{i,k}b_{k,j} = \sum_{k=1}^{m} \langle x_i,x_k^*\rangle b_{k,j} = \langle x_i, \sum_{k=1}^{m} \bar{b}_{k,j}x_k^*\rangle .$$

Hence the result. Also, it can be similarly seen that the set $\{x_1^*,\ldots,x_m^*\}$ is linearly independent in X^* , and since A^HB^H is the identity matrix, it follows that

(3.7) $$y_j = \sum_{k=1}^{m} b_{j,k}x_k$$

satisfies $\langle y_j,x_i^*\rangle = \delta_{i,j}$, $i,j = 1,\ldots,m$.

Examples of projections on finite dimensional subspaces

(i) Let X be an n-dimesional space and Y be the m-dimensional subspace with an ordered basis $x_1,\ldots,x_m$. Extend this basis to a basis of X by adding the elements $x_{m+1},\ldots,x_n$ to it. Let $Z = \text{span}\{x_{m+1},\ldots,x_n\}$. Then the projection P on Y along Z is represented by the matrix

$$m\left\{\left[\begin{array}{ccc:c} 1 & & 0 & \\ & \ddots & & 0 \\ 0 & & 1 & \\ \hdashline & 0 & & 0 \end{array}\right]\right.$$

with respect to the ordered basis $x_1, \ldots, x_n$. If $x_j^* \in X^*$ with $\langle x_j^*, x_i \rangle = \delta_{i,j}$, $i,j = 1,\ldots,m$, then P^* is also represented by the same matrix with respect to the ordered basis $x_1^*, \ldots, x_n^*$.

(ii) Let X be a Hilbert space and Y a subspace with an ordered basis $x_1, \ldots, x_m$. Then $X = Y \oplus Y^\perp$, and it follows from Theorem 3.2 that there exist $y_1, \ldots, y_m$ in $(Y^\perp)^\perp = Y$ such that

$$\langle y_j, x_i \rangle = \delta_{i,j} , \quad i,j = 1,\ldots,m .$$

It is clear that $y_1, \ldots, y_m$ are linearly independent and hence form a basis of Y . The sets $x_1, \ldots, x_m$ and $y_1, \ldots, y_m$ are said to form <u>a biorthogonal family in</u> Y . Given $x_1, \ldots, x_m$, the y_j's can be found as follows. Since $y_j \in Y$, we have

$$y_j = \alpha_{1,j} x_1 + \ldots + \alpha_{m,j} x_m , \quad \alpha_{1,j}, \ldots, \alpha_{m,j} \text{ in } \mathbb{C} .$$

Hence for $i = 1,\ldots,m$,

$$\delta_{i,j} = \langle y_j, x_i \rangle = \sum_{k=1}^{m} \alpha_{k,j} \langle x_k, x_i \rangle .$$

Thus, $\alpha_{1,j}, \ldots, \alpha_{m,j}$ can be obtained as the unique solution of the above system of m equations in m unknowns.

Note that the set $\{x_1, \ldots, x_m\}$ is orthonormal iff $y_j = x_j$ for $j = 1,\ldots,m$. Often it is convenient to have an orthonormal basis $\{u_1, \ldots, u_m\}$ of Y such that $\operatorname{span}\{u_1, \ldots, u_k\} = \operatorname{span}\{x_1, \ldots, x_k\}$ for each $k = 1,\ldots,m$. Such a set can be obtained by the famous *Gram-Schmidt orthonormalization process* ([L], 22.3).

Note that the projection P on Y along $Y^\perp$ is given by

$$Px = \sum_{j=1}^{m} \langle x, y_j \rangle x_j , \quad x \in X .$$

Since P is an orthogonal projection and $P \neq 0$, we see by Proposition 2.3 that $P^* = P$ and $\|P\| = 1$. Since

$$1 = |\langle x_j, y_j \rangle| \leq \|x_j\| \, \|y_j\| ,$$

we see that the upper bound for $\|P\|$ given in (3.5), namely $\sum_{j=1}^{m} \|x_j\| \, \|y_j\|$, is very rough when m is large.

As a concrete case, consider $X = L^2([0,1])$ and $x_j(t) = t^j$, $0 \leq t \leq 1$, for $j = 0,\ldots,m-1$. Then $Y = \text{span}\{x_0,\ldots,x_{m-1}\}$ is the space of all polynomials of degree $\leq m-1$. To find $y_j \in Y$ with $\langle y_j, x_i \rangle = \delta_{i,j}$, we consider

$$y_j(t) = \alpha_{0,j} + \alpha_{1,j} t + \ldots + \alpha_{m-1,j} t^{m-1} , \quad 0 \leq t \leq 1 .$$

Since

$$\langle x_k, x_i \rangle = \int_0^1 t^{k+i} dt = \frac{1}{k+i+1} ,$$

we see that $[\alpha_{i,j}]$ is the inverse of the $m \times n$ <u>Hilbert matrix</u> $\left[\frac{1}{i+j+1}\right]$, $i,j = 0,\ldots,m-1$. This matrix is, however, known to be numerically intractable. It is, therefore, advisable to orthonormalize the set $\{x_0,\ldots,x_{m-1}\}$ to obtain the *Legendre polynomials* and work with them.

(iii) Let $X = C([a,b])$ with the supremum norm. Consider a partition

$$a = t_0 \leq t_1 < \ldots < t_m \leq t_{m+1} = b$$

of $[a,b]$. The points $t_1,\ldots,t_m$ will be called the *nodes*. Let

$$Y = \{x \in X : x \text{ is linear on } [t_{i-1},t_i] , \ i = 1,\ldots,m+1 ,$$
$$x(a) = x(t_1) \text{ and } x(t_m) = x(b)\} .$$

Every element x of Y is piecewise linear; the linearity of x can break down only at the nodes $t_1,\dots,t_m$. Let $e_i \in Y$ be such that

$$e_i(t_j) = \delta_{i,j} , \quad i,j = 1,\dots,m .$$

The functions $e_1,\dots,e_m$ form a basis of Y ; their graphs are shown in Figure 3.1.

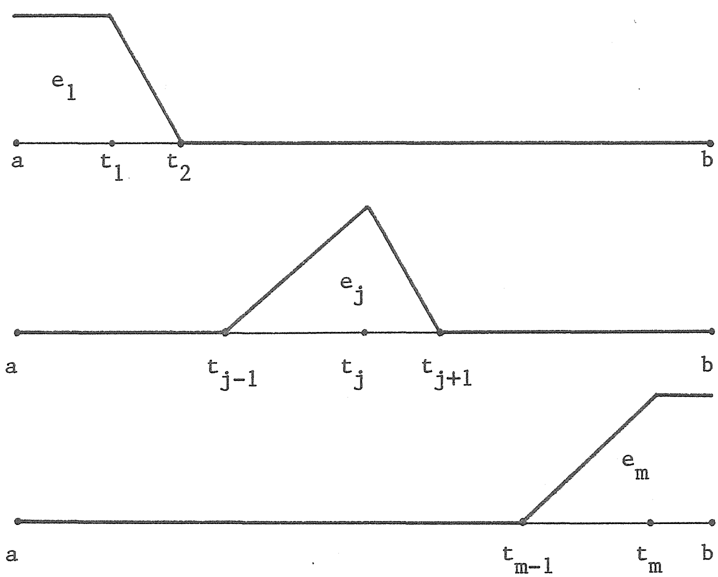

Figure 3.1

We give explicit formulae for these <u>piecewise linear hat functions</u> for later computational use:

$$e_1(t) = \begin{cases} 1 , & \text{if} \quad a \le t < t_1 \\ (t_2-t)/(t_2-t_1) , & \text{if} \quad t_1 \le t < t_2 , \\ 0 , & \text{if} \quad t_2 \le t \le b \end{cases}$$

$$e_m(t) = \begin{cases} 0 , & \text{if} \quad a \le t < t_{m-1} \\ (t_{m-1}-t)/(t_{m-1}-t_m) , & \text{if} \quad t_{m-1} \le t < t_m , \\ 1 , & \text{if} \quad t_m \le t \le b \end{cases}$$

and for $j = 2,\dots,m-1$,

$$e_j(t) = \begin{cases} 0 & , \text{ if } \quad a \le t < t_{j-1} \\ (t_{j-1}-t)/(t_{j-1}-t_j) & , \text{ if } \quad t_{j-1} \le t < t_j \\ (t_{j+1}-t)/(t_{j+1}-t_j) & , \text{ if } \quad t_j \le t < t_{j+1} \\ 0 & , \text{ if } \quad t_{j+1} \le t \le b \end{cases} .$$

It can be easily checked that for $j = 1,\dots,m$, $e_j(t) \ge 0$ for all $t \in [a,b]$, e_j vanishes outside $[t_{j-1}, t_{j+1}]$, at any fixed $t \in [a,b]$ at most two of the functions $e_1,\dots,e_m$ are nonzero, and for all t ,

$$e_1(t) + \dots + e_m(t) = 1 .$$

Because of such very nice properties, these so-called *hat functions* $e_1,\dots,e_m$ prove to be very useful in numerical calculations.

For $j = 1,\dots,m$, define $e_j^* \in X^*$ by

$$e_j^*(x) = x(t_j) \ , \quad x \in X .$$

Then $\langle e_j^*, e_i \rangle = \delta_{i,j}$. Consider for $x \in X$,

$$Px = \sum_{j=1}^{m} x(t_j)e_j .$$

Then P is a projection on Y along

$$Z = \{x \in X : x(t_j) = 0 \text{ for each } j = 1,\dots,m\} .$$

Note that for $t \in [a,b]$,

$$\begin{aligned} |Px(t)| &\le \sum_{j=1}^{m} |x(t_j)| \, |e_j(t)| \\ &\le \|x\|_\infty \sum_{j=1}^{m} e_j(t) = \|x\|_\infty . \end{aligned}$$

Hence $\|P\| = 1$. Also, it is easy to see that $\|e_j\| = \|e_j^*\| = 1$ for $j = 1,\ldots,m$. Again, we see that the bound for $\|P\|$ given in (3.5) need not be sharp.

Finite rank operators

We now consider operators whose ranges are finite dimensional. An operator $T \in BL(X)$ is said to be <u>of finite rank</u> if the dimension of its range $R(T)$ is finite; this dimension is called the <u>rank</u> of the operator.

Let T be of finite rank. Consider a subspace Y of X containing $R(T)$, and let $x_1,\ldots,x_m$ be an ordered basis of Y . By Theorem 3.2, find $x_j^* \in X^*$, $j = 1,\ldots,m$ such that $\langle x_j^*, x_i\rangle = \delta_{i,j}$. For $x \in X$, we have

$$Tx = \alpha_1 x_1 + \ldots + \alpha_m x_m , \quad \alpha_j \in \mathbb{C} .$$

By applying x_j^* on both sides, we see

$$\text{(3.8)} \qquad Tx = \sum_{j=1}^{m} \langle Tx, x_j^*\rangle x_j = \sum_{j=1}^{m} \langle x, T^* x_j^*\rangle x_j .$$

Now, T maps Y into Y and $x_j^*|_Y$, $j = 1,\ldots,m$ form the ordered basis of Y^* which is adjoint to the ordered basis $x_1,\ldots,x_m$ of Y . Hence T_Y is represented by the matrix $(t_{i,j})$ with respect to the basis $x_1,\ldots,x_m$, where $t_{i,j} = \langle Tx_j, x_i^*\rangle$, $i,j = 1,\ldots,m$. The operator $(T_Y)^* : Y^* \to Y^*$ is then represented by the matrix $[\bar{t}_{j,i}]$ with respect to the basis $x_1^*|_Y,\ldots,x_m^*|_Y$. See Example (i) at the end of Section 1.

The sum of the diagonal entries $\langle Tx_j, x_j^*\rangle$, $j = 1,\ldots,m$ of the above matrix $(t_{i,j})$ is called the <u>trace of the finite rank operator</u> T :

$$\text{(3.9)} \qquad \mathrm{tr}(T) = \sum_{j=1}^{m} \langle Tx_j, x_j^*\rangle .$$

We now show that the trace of T does not depend on the choice of the finite dimensional subspace Y which contains $R(T)$, or the ordered basis $x_1,\dots,x_m$ of Y, or its adjoint basis $x_1^*,\dots,x_m^*$.

PROPOSITION 3.6 Let Y_0 be a finite dimensional subspace of X containing $R(T)$, $y_1,\dots,y_n$ an ordered basis of Y_0, and $y_1^*,\dots,y_n^*$ an adjoint basis. Then

$$\mathrm{tr}(T) = \sum_{j=1}^{n} \langle Ty_j, y_j^* \rangle .$$

Proof We can assume without loss of generality that Y_0 contains Y. For, otherwise we can consider the subspace Y_1 spanned by Y and Y_0 and argue in a similar manner twice.

Now, if necessary, extend the linearly independent set $\{x_1,\dots,x_m\}$ in Y to an ordered basis $x_1,\dots,x_n$ of Y_0 in such a way that for $i = m+1,\dots,n$, we have $\langle x_j^*, x_i \rangle = 0$, $j = 1,\dots,m$. For $j = m+1,\dots,n$, find $x_j^* \in X^*$ such that

$$\langle x_j^*, x_i \rangle = \delta_{i,j} , \quad i = 1,\dots,n .$$

For $m+1 \leq j \leq n$, we have $Tx_j \in Y$ so that

$$Tx_j = \alpha_1 x_1 + \dots + \alpha_m x_m , \quad \alpha_j \in \mathbb{C} .$$

Hence for $j = m+1,\dots,n$ we have

$$\langle Tx_j, x_j^* \rangle = \sum_{i=1}^{m} \alpha_i \langle x_i, x_j^* \rangle = 0 .$$

Thus,

$$\sum_{j=1}^{n} \langle Tx_j, x_j^* \rangle = \sum_{j=1}^{m} \langle Tx_j, x_j^* \rangle = \mathrm{tr}(T) . \tag{3.10}$$

Since $\{x_1,\dots,x_n\}$ is a basis of Y_0, we have

$$y_i = \sum_{k=1}^{n} \alpha_{k,i} x_k , \quad \alpha_{k,i} \in \mathbb{C} , \quad i = 1,\dots,n .$$

Similarly, since $\{x_k^*|_{Y_0} : k = 1,\dots,n\}$ is a basis of Y_0^*, we have

$$y_i^*|_{Y_0} = \sum_{k=1}^{n} \beta_{k,i} x_k^*|_{Y_0} , \quad \beta_{k,i} \in \mathbb{C} , \; i = 1,\dots,n .$$

Now,

$$\begin{aligned} \delta_{i,j} = \langle y_j^*, y_i \rangle &= \sum_{k=1}^{n} \beta_{k,j} \langle x_k^*, \sum_{p=1}^{n} \alpha_{p,i} x_p \rangle \\ &= \sum_{k=1}^{n} \beta_{k,j} \bar{\alpha}_{k,i} , \quad i,j = 1,\dots,n . \end{aligned}$$

If $A = [\alpha_{i,j}]$ and $B = [\beta_{i,j}]$, then we see that $A^H B = I$. Hence the matrix A is nonsingular and $A^{-1} = B^H$, so that $AB^H = I$, i.e.,

$$\sum_{i=1}^{n} \alpha_{k,i} \bar{\beta}_{j,i} = \delta_{k,j} , \quad k,j = 1,\dots,n .$$

Hence

$$\begin{aligned} \sum_{i=1}^{n} \langle Ty_i, y_i^* \rangle &= \sum_{i=1}^{n} \sum_{k=1}^{n} \alpha_{k,i} \langle Tx_k, \sum_{j=1}^{n} \beta_{j,i} x_j^* \rangle \\ &= \sum_{k=1}^{n} \sum_{j=1}^{n} \langle Tx_k, x_j^* \rangle \sum_{i=1}^{n} \alpha_{k,i} \bar{\beta}_{j,i} \\ &= \sum_{k=1}^{n} \langle Tx_k, x_k^* \rangle . \end{aligned}$$

Now (3.10) shows that $\sum_{i=1}^{n} \langle Ty_i, y_i^* \rangle = \mathrm{tr}(T)$. //

Let us now consider the adjoint of a finite rank operator.

THEOREM 3.7 If T is of rank $m < \infty$, then so is T^*. In fact, if

$x_1,\dots,x_m$ is an ordered basis of $R(T)$ and $x_j^* \in X^*$ with $\langle x_j^*, x_i\rangle = \delta_{i,j}$, then $T^*x_1^*,\dots,T^*x_m^*$ form a basis of $W = R(T^*)$, and $(T^*)_W : W \to W$ is represented by the matrix $[\bar{t}_{j,i}]$ with respect to this basis, where $t_{i,j} = \langle Tx_j, x_i^*\rangle$. For $x^* \in X^*$, we have

$$(3.11)\qquad T^*x^* = \sum_{j=1}^{m} \langle x^*, x_j\rangle T^*x_j^* .$$

Moreover, we have

$$(3.12)\qquad R(T^*) = Z(T)^{\perp} .$$

Proof For $x^* \in X^*$ and $x \in X$, we have

$$\begin{aligned}\langle T^*x^*, x\rangle = \langle x^*, Tx\rangle &= \langle x^*, \sum_{j=1}^{m} \langle x, T^*x_j^*\rangle x_j\rangle \\ &= \sum_{j=1}^{m} \langle T^*x_j^*, x\rangle\langle x^*, x_j\rangle \\ &= \langle \sum_{j=1}^{m} \langle x^*, x_j\rangle T^*x_j^*, x\rangle .\end{aligned}$$

Hence (3.11) follows. This shows that

$$W = R(T^*) = \operatorname{span}\{T^*x_1^*,\dots,T^*x_m^*\} .$$

Since $x_j \in R(T)$, let $x_j = Tu_j$, $u_j \in X$ for $j = 1,\dots,m$. Then

$$(3.13)\qquad \langle T^*x_j^*, u_i\rangle = \langle x_j^*, Tu_i\rangle = \langle x_j^*, x_i\rangle = \delta_{i,j} .$$

Hence $T^*x_1^*,\dots,T^*x_m^*$ are linearly independent as well. Thus, T^* has rank m . Let

$$T^*(T^*x_j^*) = s_{1,j}T^*x_1^* + \dots + s_{m,j}\,T^*x_m^* , \quad s_{i,j} \in \mathbb{C} .$$

Then (3.13) shows that

$$\begin{aligned}s_{i,j} &= \langle T^*(T^*x_j^*), u_i\rangle = \langle T^*x_j^*, Tu_i\rangle \\ &= \langle T^*x_j^*, x_i\rangle = \langle x_j^*, Tx_i\rangle .\end{aligned}$$

Hence $T^*|_W$ is represented by the matrix $(\bar{t}_{j,i})$, where $t_{i,j} = \langle Tx_j, x_i^* \rangle$.

Finally, it is easy to see that $R(T^*)$ is contained in $Z(T)^\perp$. On the other hand, let $y^* \in Z(T)^\perp$. Since for $x \in X$, we have $Tx = \sum_{j=1}^{m} \langle Tx, x_j^* \rangle x_j$ and $x_j = Tu_j$, we see that

$$x - \sum_{j=1}^{m} \langle Tx, x_j^* \rangle u_j \in Z(T) .$$

Hence

$$\begin{aligned}\langle y^*, x \rangle &= \langle y^*, \sum_{j=1}^{m} \langle Tx, x_j^* \rangle u_j \rangle \\ &= \langle \sum_{j=1}^{m} \langle y^*, u_j \rangle T^* x_j^*, x \rangle .\end{aligned}$$

This shows that $y^* = \sum_{j=1}^{m} \langle y^*, u_j \rangle T^* x_j^* \in R(T^*)$, and proves (3.12). //

Before we conclude this longish section, we give a characterization of bounded operators of finite rank.

PROPOSITION 3.8 $T \in BL(X)$ is of finite rank if and only if T is compact and $R(T)$ is closed in X.

Proof Let T be a bounded operator of finite rank. If (x_n) is a bounded sequence in X, then (Tx_n) is also a bounded sequence in $R(T)$, which is finite dimensional. Hence by the Heine-Borel theorem, (Tx_n) has a convergent subsequence. This shows that T is compact. Also, being finite dimensional, $R(T)$ is closed in X by Proposition 3.1.

Conversely, let T be compact and $R(T)$ be closed in X. Then $T : X \to R(T)$ is a continuous map from the Banach space X onto the

Banach space $R(T)$. By the open mapping theorem ([L], 11.1), there is $\delta > 0$ such that $y \in R(T)$ and $\|y\| \leq \delta$ imply $y = Tx$ for some $x \in X$ with $\|x\| < 1$, i.e.,

$$\{y \in R(T) : \|y\| \leq \delta\} \subset \{Tx : \|x\| < 1\} ,$$

Since T is compact, the closure of the set $\{Tx : \|x\| < 1\}$ is compact. This shows that the closed ball of radius δ in $R(T)$ is compact. Hence $R(T)$ is finite dimensional ([L], 6.9). //

COROLLARY 3.9 Let $P \in BL(X)$ be a projection. Then P is of finite rank if and only if P is compact.

Proof Since $R(P) = Z(I-P)$ is closed, the result is immediate from Proposition 3.8. //

Problems

3.1 If Y is an m dimensional subspace of X , then there is a basis $\{y_1,\ldots,y_m\}$ of Y such that $\|y_j\| = 1$ and $\mathrm{dist}(y_j,\mathrm{span}\{y_1,\ldots,y_{j-1}\}) = 1$. Can we have, in fact, $\mathrm{dist}(y_j,\mathrm{span}\{y_1,\ldots,y_{j-1},y_{j+1},\ldots,y_m\}) = 1$?

3.2 Let $X = Y \oplus Z$, and let $x_1,\ldots,x_m$ (resp., $x_1^*,\ldots,x_m^*$) form an ordered basis of Y (resp., $Z^\perp$) such that $\langle x_i,x_j^*\rangle = \delta_{i,j}$. Then

$$\|x_j^*\| = 1/\mathrm{dist}(x_j,X_j) ,$$

where $X_j = Y_j \oplus Z$, with $Y_j = \mathrm{span}\{x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_m\}$.

3.3 Let $t_1,\ldots,t_m$ be the nodes in $[a,b]$ for the piecewise linear hat functions $e_1,\ldots,e_m$. Let $\tilde{X} = NBV([a,b])$, and for $\tilde{x} \in \tilde{X}$

$$Q\tilde{x} = \sum_{j=1}^{m} \left[\int_a^b e_j(t)d\tilde{x}(t)\right] f_j ,$$

where f_j is the characteristic function of the set $[t_j,b]$, $j = 1,\dots,m$. (In case $t_1 = a$, we take $f_1(a) = 0$.) Then Q is the projection on $\text{span}\{f_1,\dots,f_m\}$ along $\{\tilde{x} \in \tilde{X} : \int_a^b e_j(t)d\tilde{x}(t) = 0 , \ j = 1,\dots,m\}$ and has norm 1. It can be identified with the adjoint of the projection P of Example (iii).

3.4 $T \in BL(X)$ is of finite rank if and only if there exist $x_1,\dots,x_n$ in X and $x_1^*,\dots,x_n^*$ in X^* such that

$$Tx = \sum_{j=1}^{n} \langle x,x_j^* \rangle x_j \ , \quad x \in X \ ,$$

and in that case,

$$T^*x^* = \sum_{j=1}^{n} \langle x^*,x_j \rangle x_j^* \ , \quad x^* \in X^* \ .$$

One may assume without loss of generality that the sets $\{x_1,\dots,x_n\}$ and $\{x_1^*,\dots,x_n^*\}$ are linearly independent.

3.5 If $T \in BL(X)$ is of finite rank and $A \in BL(X)$, then TA and AT are of finite rank, and $tr(TA) = tr(AT)$.