Adjoint of an operator

P. Sam Johnson

Outline of the talk

We shall discuss existence and results on adjoints of the following :

- real / complex matrices
- linear operators between finite / infinite dimensional inner product spaces
- bounded linear operators between Hilbert spaces
- unbounded linear operators between Hilbert spaces
- bounded linear operators between Banach spaces
- operators between topological vector spaces
- operators between indefinite inner product spaces

Part - 1

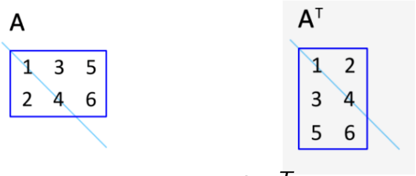
Adjoints of Matrices

Motivation

Consider a real matrix A of order $m \times n$.

The transpose of A is a matrix which flips a matrix over its diagonal, that is, it switches the row and column indices of the matrix by producing another matrix denoted as A^T. It is achieved by any one of the following equivalent actions :

- Reflect A over its main diagonal (which runs from $a_{11}, a_{22}, \ldots$) to obtain A^T.



- Write the rows of A as the columns of A^T.
- Write the columns of A as the rows of A^T.

Motivation

Formally, the i-th row, j-th column element of A^T is the j-th row, i-th column element of A:

$$[A^T]_{ij} = [A]_{ji}.$$

If A is an $m \times n$ matrix, then A^T is an $n \times m$ matrix. Note that if x and y are viewed as column vectors in $\mathbb{R}^n$, then the **Euclidean inner product (dot product)** on $\mathbb{R}^n$ is defined by

$$\langle x, y\rangle = y^T x.$$

Then we have

$$\langle Ax, y\rangle = \langle x, A^T y\rangle \quad \text{for all} \;\; x \in \mathbb{R}^n, \; y \in \mathbb{R}^m \tag{1}$$

because

$$\sum_{i=1}^{m} y_i \Big(\sum_{j=1}^{n} a_{ij} x_j\Big) = \sum_{i=1}^{n} \Big(\sum_{j=1}^{m} a_{ji} y_j\Big) x_i \quad \text{or} \quad \langle Ax, y\rangle = y^T(Ax) = (A^T y)^T x = \langle x, A^T y\rangle.$$

Motivation

Let A and B be real matrices of order $m \times n$. one can prove the following.

(a) $(A+B)^T = A^T + B^T$

(b) $(cA)^T = cA^T$ for any $c \in \mathbb{R}$

(c) $(AB)^T = B^T A^T$ (under the assumption, the product AB is compatible)

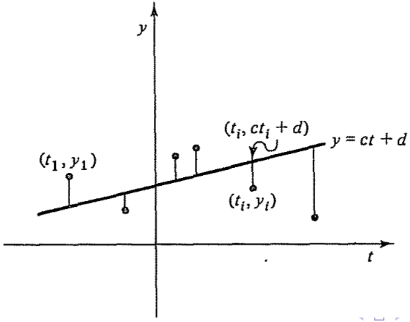
(d) $(A^T)^T = A$

(e) $I^T = I$, here I denotes the identity matrix.

We now discuss an application of transpose for fitting a best possible line to the data collected.

Consider the following problem: An experimenter collects data by taking measurements $y_1, y_2, \ldots, y_m$ at times $t_1, t_2, \ldots, t_m$, respectively. For example, the experimenter may be measuring unemployment at various times during some period.

Suppose that the data $(t_1, y_1), (t_2, y_2), \ldots, (t_m, y_m)$ are plotted as points in the plane.



Application : Least Squares Approximation

From this plot, the experimenter feels that there exists an essentially linear relationship between y and t, say $y = ct + d$, and would like to find the constants c and d so that the line $y = ct + d$ represents the best possible fit to the data collected.

One such estimate of fit is to calculate the error E that represents the sum of the squares of the verticle distances from the points to the line; that is,

$$E = \sum_{i=1}^{m} (y_i - ct_i - d)^2.$$

Thus the problem is reduced to finding the constants c and d that minimize E.

The procedure of minimizing $\sqrt{E}$ is the same as minimizing E. For calculation purpose, E has been taken in place of $\sqrt{E}$.

By minimizing E, we are going to fit a line ; for this reason, the line $y = ct + d$ is called the **least squares line.**

If we let

$$A = \begin{pmatrix} t_1 & 1 \\ t_2 & 1 \\ \vdots & \vdots \\ t_m & 1 \end{pmatrix}, \quad x = \begin{pmatrix} c \\ d \end{pmatrix} \quad \text{and} \quad y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}.$$

We shall find the **least squares line** by using the transpose of A.

If we denote the Euclidean distance between the origin and a vector $x \in \mathbb{R}^n$ by $\|x\|$, then the error E which represents the sum of the squares of the verticle distances from the points to the line is

$$E = \|y - Ax\|^2.$$

Application : Least Squares Approximation

For a given matrix $A_{m \times n}$ and an element $y \in \mathbb{R}^m$, the matrix equation

$$Ax = y \tag{2}$$

may not have a solution (so the error E is non-zero).

When the length of E is as small as possible, x_0 is a least residual norm solution (LRN solution) which is classically known as a least squares solution. That is, the least squares solution x_0 makes $E = \|Ax_0 - y\|^2$ as small as possible.

We shall see that (2) has an LRN-solution $\iff$ $A^T Ax = A^T y$ has a solution.

We shall also see that $rank(A^T A) = rank(A)$.

The equation $A^T Ax = A^T y$ is called the **normal form** of $Ax = y$.

(Classical) Adjoint

Definition 1.

Let A be a square matrix of order $n \geq 2$. The **(classical) adjoint** *of A or the* **adjugate matrix**, *or the* **adjunct matrix** *of A is the transpose of its cofactor matrix (the matrix of cofactors of A).*

Recall that $(adj\ A)_{ij} = (-1)^{i+j} det\, A(j|i)$ and $(adj\, A)A = (det\, A)I$.

$det\, A(j|i)$ is the determinant of the $(n-1) \times (n-1)$ matrix obtained by deleting the ith row and jth coloum of A. The adjugate has sometimes been called the "adjoint", but today the "adjoint" of a matrix normally refers to its corresponding (modern) adjoint, which is its conjugate transpose.

We use only the (modern) adjoint and call it simply "adjoint".

We should think of A^* as the "backwards" version of A (not to be confused with its inverse).

Let A be a complex matrix of order $m \times n$. The following example illustrates that the transpose of A does not satisfy the relation :

$$\langle Ax, y\rangle = \langle x, A^T y\rangle \quad \text{for all} \;\; x \in \mathbb{C}^n, \;\; y \in \mathbb{C}^m. \tag{3}$$

Example 2.

Let $A = \begin{pmatrix} i & 2-i \\ 3 & 4i \end{pmatrix}$, $x = \begin{pmatrix} 2 \\ 3i \end{pmatrix}$, $y = \begin{pmatrix} 4i \\ 5 \end{pmatrix}$.

Then $\langle Ax, y\rangle = -62 - 12i$ *and* $\langle x, A^T y\rangle = -44$.

Hence

$$\langle Ax, y\rangle \neq \langle x, A^T y\rangle.$$

(Modern) Adjoint

Notice that in $\mathbb{R}$ we have

$$x_1^2 + x_2^2 + \cdots + x_n^2 = 0 \implies x_1 = x_2 = \cdots = x_n = 0.$$

Similarly, in $\mathbb{C}$, we have

$$\overline{x_1}x_1 + \overline{x_2}x_2 + \cdots + \overline{x_n}x_n = 0 \implies x_1 = x_2 = \cdots = x_n = 0$$

because $\overline{x}x = |x|^2$ for any complex number x.

For $x \in \mathbb{R}^n$, we have $x^T x = 0$ implies $x = 0$. The complex analogue of the above relation is $\overline{x}^T x = 0$ implies $x = 0$. Motivated by this, the complex analogue of A^T replaces A^T by $\overline{A}^T$. It is observed that the conjugate transpose satisfies the relation

$$\langle Ax, y \rangle = \langle x, A^T y \rangle \quad \text{for all } x \in \mathbb{C}^n, \quad y \in \mathbb{C}^m.$$

and it is denoted by A^*. We denote $\overline{A}^T$ by A^* for any complex matrix A and call it the **(modern) adjoint** of A.

Adjoint of a complex matrix

Let $\mathbb{K}$ be the field of real or complex scalars.

Definition 3.

Let $A \in M_{m\times n}(\mathbb{K})$. We define the **conjugate transpose** *or* **adjoint** *of A to be the $n \times m$ matrix A^* such that*

$$(A^*)_{ij} = \overline{A_{ji}}$$

for all i, j. The symbol A^* is read "A star."

Example 4.

Let $A = \begin{pmatrix} i & 1+2i \\ 2 & 3+4i \end{pmatrix}$. *Then* $A^* = \begin{pmatrix} -i & 2 \\ 1-2i & 3-4i \end{pmatrix}$.

In the case that A has real entries, A^* is simply the **transpose** of A.

If x and y are viewed as column vectors in $\mathbb{K}^n$, then the standard inner product on $\mathbb{K}^n$ is defined by $\langle x, y \rangle_n = y^* x$.

Exercise 5.

Let $A \in M_{m \times n}(\mathbb{K})$.

(a) **Existence :** *Prove that* $\langle Ax, y \rangle_m = \langle x, A^* y \rangle_n$ *for all* $x \in \mathbb{K}^n, y \in \mathbb{K}^m$.

(b) **Uniqueness :** *Suppose that for some $B \in M_{n \times m}(\mathbb{K})$, we have*

$$\langle Ax, y \rangle_m = \langle x, By \rangle_n \quad \textit{for all } x \in \mathbb{K}^n, y \in \mathbb{K}^m.$$

Prove that $B = A^*$.

LA-2(P-30)E-23

Adjoint of a complex matrix

Exercise 6.

(a) *Let Q be an $n \times n$ matrix whose columns form an orthonormal basis for V. Prove that $Q^* = Q^{-1}$.*

(b) *Let α be the standard ordered basis for V. Define linear operators T and U on V by $Tx = Ax$ and $Ux = A^*x$. Show that $[U]_\beta = [T]_\beta^*$, for any orthonormal basis β for V.*

LA-2(P-30)E-23

Adjoint of a complex matrix

Theorem 7.

Let A and B be matrices over $\mathbb{K}$ of order $m \times n$. Then

(a) $(A+B)^* = A^* + B^*$

(b) $(cA)^* = \overline{c}A^*$ *for any* $c \in \mathbb{K}$

(c) $(AB)^* = B^*A^*$ *(under the assumption, the product AB is compatible)*

(d) $(A^*)^* = A$

(e) $I^* = I$, *here I denotes the identity matrix.*

LA-2(P-6)T-5

The above result suggests an analogy between the conjugates of complex numbers ($z \mapsto \overline{z}$) and the transposes of complex matrices ($A \mapsto A^*$).

(a) $\overline{z_1 + z_2} = \overline{z_1} + \overline{z_2}$

(b) $\overline{cz_1} = c\overline{z_1}$ *for any* $c \in \mathbb{R}$

(c) $\overline{z_1 z_2} = \overline{z_1} \,.\, \overline{z_2}$ *(not reversed)*

(d) $\overline{(\overline{z_1})} = z_1$

In some sense, every complex matrix has real and imaginary parts.

We discuss extensions of this analogy and we mention something along these lines now. A complex number z is real if and only if $z = \bar{z}$. One might expect that the matrices A such that $A = A^*$ behave in some way like the real numbers.

Every complex number z can be written as $z = z_1 + iz_2$, where z_1 and z_2 are real. One might expect that every matrix A can be written as $A = A_1 + iA_2$, where $A_1 = A_1^*$ and $A_2 = A_2^*$. Thus, in some sense, A has a 'real part' and an 'imaginary part'.

The matrices A_1 and A_2 satisfying $A_1 = A_1^*$ and $A_2 = A_2^*$, and are given by

$$A_1 = \frac{1}{2}(A + A^*) \qquad A_2 = \frac{1}{2i}(A - A^*).$$

In some sense, every complex matrix has real and imaginary parts.

Exercises 8.

1. *Let A be an $n \times n$ matrix. Define $A_1 = \frac{1}{2}(A + A^*)$ and $A_2 = \frac{1}{2i}(A - A^*)$.*
   (a) *Prove that $A_1^* = A_1, A_2^* = A_2$, and $A = A_1 + iA_2$. Would it be reasonable to define A_1 and A_2 to be the real and imaginary parts, respectively, of the matrix A?*
   (b) *Let A be an $n \times n$ matrix. Prove that the representation in (a) is unique. That is, prove if $A = B_1 + iB_2$, where $B_1^* = B_1$ and $B_2^* = B_2$, then $B_1 = A_1$ and $B_2 = A_2$.*

LA-2(P-31)E-24

Results on adjoint of a complex matrix

Theorem 9.

For any complex matrix A, we have the following :

1. *null space* $(A) =$ *null space* (A^*A).
2. $rank(A) = rank(A^*A)$ *and* $rank(A^*) = rank(AA^*)$
3. *coloum space* $(A) =$ *column space* (AA^*)
4. *row space* $(A) =$ *row space* (A^*A)
5. $rank(A) = rank(A^*)$

LA-2(P-65,67)T-53

Using Theorem (9), we can give a left inverse of a complex matrix with full column rank and a right inverse of a complex matrix will full row rank.

Results on adjoint of a complex matrix

Theorem 10.

Let A be an $m \times n$ complex matrix.

1. *A is of full column rank iff A^*A is non-singular. Also then $(A^*A)^{-1}A^*$ is a left inverse of A, and*
2. *A is of full row rank iff AA^* is non-singular. Also then $A^*(AA^*)^{-1}$ is a right inverse of A.*

LA-2(P-??)T-??

Corollary 11.

For any complex matrix A, there exist complex matrices C and D such that $A = AA^*C$ and $A = DA^*A$.

LA-2(P-65,67)T-53

Theorem 12.

For complex matrices A, B and C we have the cancellation laws :

1. *If $AA^*B = AA^*C$, then $A^*B = A^*C$.*
2. *If $BA^*A = CA^*A$, then $BA^* = CA^*$.*

LA-2(P-65,67)T-53

In some sense, every complex matrix has real and imaginary parts.

Exercises 13.

1. *Let A be an $n \times n$ matrix. Prove that* $\det(A^*) = \overline{\det(A)}$. LA-2(P-50)E-42
2. *Suppose that A is an $m \times n$ matrix in which no two columns are identical. Prove that A^*A is a diagonal matrix if and only if every pair of columns of A is orthogonal.* LA-2(P-51)E-43

LA-2(P-65,67)T-53

Part - 2

Adjoint of a linear operator on a finite-dimensional inner product space

Adjoint of a linear operator on a finite-dimensional inner product space

Question : For a linear operator T between finite-dimensional inner product spaces V and W, what is an extension of this analogy (conjugate transpose)?

We shall first discuss for a linear operator T on a finite-dimensional inner product space V (linear operator $T : V \to V$), to define a related linear operator on V called the **adjoint of** T and it will be denoted by T^*.

We shall see that the matrix representation of T^* with respect to any orthonormal basis β for V is $[T]^*_\beta$ (the complex conjugate of $[T]_\beta$).

Adjoint of a linear operator on a finite-dimensional inner product space

Two distinct inner products on a given vector space yield two distinct inner product spaces. For instance, it can be shown that

$$\langle f(x), g(x)\rangle_1 = \int_0^1 f(t)g(t)\ dt \quad \text{and} \quad \langle f(x), g(x)\rangle_2 = \int_{-1}^1 f(t)g(t)\ dt$$

are inner products on the vector space $P_n(\mathbb{R})$ of polynomials of degree at most n with coefficients in $\mathbb{R}$.

Even though the underlying vector space is the same, however, these two inner products yield two different inner product spaces.

For example, the polynomials $f(x) = x$ and $g(x) = x^2$ are orthogonal in the second inner product space, but not in the first.

Adjoint of a linear operator on a finite-dimensional inner product space

It will be observed that the adjoint of T depends not only on T but on the inner product as well.

To see the existence of the "adjoint" of a linear operator T on V, we begin with linear functionals on an inner product space and their relation to the inner product.

The basic result is that any linear functional f on a finite-dimensional inner product space is "inner product with a fixed vector in the space," i.e., that such an f has the form

$$f(x) = \langle x, y \rangle \quad \text{for some fixed } y \in V.$$

Adjoint of a linear operator on a finite-dimensional inner product space

We use this result to prove the existence of the "adjoint" of a linear operator T on V, this being a linear operator T^* such that

$$\langle Tx, y \rangle = \langle x, T^* y \rangle$$

for all $x, y \in V$.

Through the use of an **orthonormal basis**, this adjoint operation on linear operators (passing from T to T^*) is identified with the operation of forming the conjugate transpose of a matrix.

We explore slightly the analogy between the adjoint operation and conjugation on complex numbers.

Adjoint of a linear operator on a finite-dimensional inner product space

Let V be a finite dimensional inner product space, and let y be some fixed vector in V. We define a function g_y from V into the scalar field $\mathbb{K}$ by

$$g_y(x) = \langle x, y \rangle.$$

This function g_y is a linear functional on V, because, by its very definition, $\langle x, y \rangle$ is linear as a function of x. More interesting is the fact that **if V is finite-dimensional, every linear operator from V into $\mathbb{K}$ is of this form.**

Theorem 14.

Let V be a finite-dimensional inner product space over $\mathbb{K}$, and let $g : V \to \mathbb{K}$ be a linear operator. Then there exists a unique vector $y \in V$ such that $g(x) = \langle x, y \rangle$ for all $x \in V$. LA-2(P-1)T-1

How to get y ?

Outline of the proof : Let $\{v_1, v_2, \ldots, v_n\}$ be an orthonormal basis for V. If g is any linear functional on V, then g has the form

$$g(x) = a_1 g(v_1) + \cdots + a_n g(v_n) \quad \text{for} \quad x = \sum_{i=n}^{n} a_i v_i.$$

If we wish to find a vector $y = \sum_{i=n}^{n} b_i v_i$ in V such that $\langle x, y \rangle = g(x)$ for all x, then clearly the coordinates b_i of y must satisfy $b_i = \overline{g(v_i)}$ because $\displaystyle\sum_{i=n}^{n} a_i \overline{b_i} = \sum_{i=n}^{n} a_i g(v_i)$.

Accordingly

$$y = \overline{g(v_1)} v_1 + \cdots + \overline{g(v_n)} v_n$$

is the desired vector.

Example 15.

Define $g : \mathbb{R}^2 \to \mathbb{R}$ *by* $g(a_1, a_2) = 2a_1 + a_2$; *clearly* g *is a linear operator.*

Let $\beta = \{e_1, e_2\}$ *be an orthonormal basis for* $\mathbb{R}^2$, *and let*

$$y = g(e_1)e_1 + g(e_2)e_2 = 2e_1 + e_2 = (2, 1),$$

as in the proof of Theorem 14.

Then

$$g(a_1, a_2) = \langle (a_1, a_2), (2, 1) \rangle = 2a_1 + a_2.$$

Exercise 16.

For each of the following inner product spaces V (over $\mathbb{K}$) and linear transformations $g : V \to \mathbb{K}$, find a vector y such that $g(x) = \langle x, y\rangle$ for all $x \in V$.

1. $V = \mathbb{R}^3$, $g(a_1, a_2, a_3) = a_1 - 2a_2 + 4a_3$
2. $V = \mathbb{C}^2$, $g(z_1, z_2) = z_1 - 2z_2$
3. $V = P_2(\mathbb{R})$ with $\langle f, h\rangle = \int_0^1 f(t)h(t)\,dt, \quad g(f) = f(0) + f'(1)$

LA-2(P-34)E-27

Geometric Fact : Where does the unique vector y lie?

The essential geometric fact is that y lies in the orthogonal complement of the null space of g.

Let $W = N(g)$, the null space of g. Then $V = W \oplus W^\perp$, and g is completely determined by its values on $W^\perp$.

If g is zero, the zero element in V does this.

Suppose $g \neq 0$. Then g is of rank 1 and dim $(W^\perp) = 1$.

In fact, if P is the orthogonal projection of V on $W^\perp$, then

$$g(x) = g(Px)$$

for all x in V.

Geometric Fact : Where does the unique vector y lie?

If z is any non-zero vector in $W^{\perp}$, it follows that

$$Px = \left\langle x, \frac{z}{\|z\|} \right\rangle \frac{z}{\|z\|} = \frac{\langle x, z \rangle}{\|z\|^2} z$$

for all x in V. Thus

$$g(x) = \langle x, z \rangle . \frac{g(z)}{\|z\|^2} = \left\langle x, \frac{\overline{g(z)}}{\|z\|^2} z \right\rangle$$

for all x, and

$$y = \frac{\overline{g(z)}}{\|z\|^2} z$$

is the desired vector.

Example

We now see an example showing that Theorem 14 is not true without the assumption that V is finite dimensional.

Let V be the vector space of polynomials over the field of complex numbers, with the inner product

$$\langle f, g \rangle = \int_0^1 f(t)\overline{g(t)}\, dt.$$

This inner product can also be defined algebraically. If $f = \sum a_k x^k$ and $g = \sum b_k x^k$, then

$$\langle f, g \rangle = \sum_{j,k} \frac{1}{j+k+1} a_j \bar{b}_k.$$

Let z_0 be a fixed complex number, and let L be the linear functional "evaluation at z_0": $L(f) = f(z_0)$.

Example (contd...)

Question : Is there a polynomial g such that $\langle f, g\rangle = L(f)$ for every $f \in V$? **The answer is 'NO'.**

Suppose that there is a polynomial $g \in V$ such that $\langle f, g\rangle = L(f)$ for every f. Then we have $f(z_0) = \int_0^1 f(t)\overline{g(t)}\, dt$ for every $f \in V$. Let $h(x) = x - z_0$, so that for any f we have $(hf)(z_0) = 0$. Then

$$0 = \int_0^1 h(t)f(t)\overline{g(t)}\, dt \quad \text{for all } f.$$

In particular this holds when $f = \bar{h}g$ so that

$$\int_0^1 |h(t)|^2 |g(t)|^2\, dt = 0$$

and so $hg = 0$. Since $h \neq 0$, it must be that $g = 0$. But L is not the zero functional; hence, no such g exists.

Example (Generalized)

One can generalize the example somewhat, to the case where L is a linear combination of point evaluations.

Suppose we select fixed complex numbers $z_1, \ldots, z_n$ and scalars $c_1, \ldots, c_n$ and let

$$L(f) = c_1 f(z_1) + \cdots + c_n f(z_n).$$

Then L is a linear functional on V, but there is no g with $L(f) = \langle f, g \rangle$, unless $c_1 = c_2 = \cdots = c_n = 0$.

Just repeat the above argument with $h = (x - z_1) \cdots (x - z_n)$.

We have seen an example showing that Theorem 17 is not true without the assumption that V is finite dimensional. **But we shall prove that Theorem 17 holds true when V is a Hilbert space.**

The Adjoint of a Linear Operator

Theorem 17.

Let V be a finite-dimensional inner product space, and let T be a linear operator on V. Then there exists a unique function $T^* : V \to V$ such that

$$\langle Tx, y\rangle = \langle x, T^*y\rangle$$

for all $x, y \in V$. Furthermore, T^* is linear. LA-2(P-2)T-2

The linear operator T^* described in Theorem 17 is called the **adjoint** of the operator T.

Thus T^* is the unique linear operator on V satisfying

$$\langle Tx, y\rangle = \langle x, T^*y\rangle$$

for all $x, y \in V$.

The Adjoint of a Linear Operator

We may view these equations symbolically as adding " $*$ " to T when shifting its position inside the inner product symbol, as shown below.

We have

$$\langle Tx, y\rangle = \langle x, T^*y\rangle \quad \text{for all } x, y \in V.$$

and also

$$\langle x, Ty\rangle = \langle T^*x, y\rangle \quad \text{for all } x, y \in V$$

because $\langle x, Ty\rangle = \overline{\langle Ty, x\rangle} = \overline{\langle y, T^*x\rangle} = \langle T^*x, y\rangle$.

The Adjoint of a Linear Operator

Definition 18.

Let T be a linear operator on an inner product space V. Then we say that T **has an adjoint on** *V if there exists a linear operator T^* on V such that*

$$\langle Tx, y\rangle = \langle x, T^*y\rangle$$

for all x and y in V.

For an **infinite-dimensional inner product space***, the adjoint of a linear operator T may be defined to be the function T^* such that*

$$\langle Tx, y\rangle = \langle x, T^*y\rangle$$

for all $x, y \in V$, provided it exists.

Although the uniqueness and linearity of T^* follow as before, the existence of the adjoint is not guaranteed. We shall see some examples.

The Adjoint of a Linear Operator

We now recall **Theorem 17** : Let V be a finite-dimensional inner product space, and let T be a linear operator on V. Then there exists a unique function $T^* : V \to V$ such that

$$\langle Tx, y\rangle = \langle x, T^*y\rangle$$

for all $x, y \in V$. Furthermore, T^* is linear.

By Theorem 17 every linear operator on a finite-dimensional inner product space V has an adjoint on V.

In the infinite-dimensional case this is not always true. But in any case there is at most one such operator T^*; when it exists, we call it the **adjoint** of T.

Example of a linear operator on an infinite-dimensional inner product space which does NOT have an adjoint

Let V be the vector space of all sequences σ in $\mathbb{K}$ such that $\sigma(n) \neq 0$ for only finitely many positive integers n. For $\sigma, \mu \in V$, we define $\langle \sigma, \mu \rangle = \sum\limits_{n=1}^{\infty} \sigma(n)\overline{\mu(n)}$. Since all but a finite number of terms of the series are zero, the series converges.

For each positive integer n, let e_n be the sequence defined by $e_n(k) = \delta_{n,k}$, where $\delta_{n,k}$ is the Kronecker delta. We proved that $\{e_1, e_2, \dots\}$ is an orthonormal basis for V. Define $T : V \to V$ by

$$T(\sigma)(k) = \sum_{i=k}^{\infty} \sigma(i) \quad \text{for every positive integer k.}$$

LA-2(P-42)E-38

Example of a linear operator on an infinite-dimensional inner product space which does NOT have an adjoint

Notice that the infinite series in the definition of T converges because $\sigma(i) \neq 0$ for only finitely many i.

1. T is a linear operator on V.
2. For any positive integer n, $T(e_n) = \displaystyle\sum_{i=1}^{n} e_i$.
3. T has no adjoint.

LA-2(P-42)E-38

Another example of a linear operator on an infinite-dimensional inner product space which does NOT have an adjoint

Let V be the space of polynomials over the field of complex numbers, with the inner product

$$\langle f, g \rangle = \int_0^1 f(t)\overline{g(t)}\, dt.$$

Let D be the differentiation operator on V. Integration by parts shows that

$$\langle Df, g \rangle = f(1)g(1) - f(0)g(0) - \langle f, Dg \rangle.$$

Let us fix g and inquire when there is a polynomial D^*g such that

$$\langle Df, g \rangle = \langle f, D^*g \rangle$$

for all f.

Example (contd...)

If such a D^*g exists, we shall have

$$\langle f, D^*g\rangle = f(1)g(1) - f(0)g(0) - \langle f, Dg\rangle$$

or

$$\langle f, (D^* + D)g\rangle = f(1)g(1) - f(0)g(0).$$

With g fixed, the map L_g defined by

$$L_g(f) = f(1)g(1) - f(0)g(0) \quad \text{(difference of evaluations at 1 and 0)}$$

is a linear functional on V and cannot be of the form

$$L(f) = \langle f, h\rangle$$

unless $L = 0$.

Example (contd...)

If D^*g exists, then with $h = D^*g + Dg$ we do have $L(f) = \langle f, h \rangle$, and so

$$g(0) = g(1) = 0.$$

The existence of a suitable polynomial D^*g implies $g(0) = g(1) = 0$.

Conversely, if $g(0) = g(1) = 0$, the polynomial $D^*g = -Dg$ satisfies

$$\langle Df, g \rangle = \langle f, D^*g \rangle$$

for all f.

If we choose any g for which $g(0) \neq 0$ or $g(1) \neq 0$, we cannot suitably define D^*g, and so we conclude that D has no adjoint.

How to find adjoint of a linear operator?

The following result is helpful to find the adjoint of a linear operator on a finite dimensional inner product space.

Theorem 19.

Let V be a finite-dimensional inner product space, and let $\beta = \{v_1, v_2, \dots, v_n\}$ be an (ordered) orthonormal basis for V. Let T be a linear operator on V and let A be the matrix of T in the ordered basis β. Then

$$A_{ij} = \langle T\alpha_j, \alpha_i \rangle.$$

Moreover, in any orthonormal basis β for V, the matrix of T^* is the conjugate transpose of the matrix of T. That is,

$$[T^*]_\beta = [T]_\beta^*.$$

LA-2(P-46)E-39

Example

The following example shows that in an arbitrary ordered basis γ, the relation between $[T]_\gamma$ and $[T^*]_\gamma$ is more complicated than that given in the Theorem 19.

Example 20.

Let V be a finite-dimensional inner product space and E the orthogonal projection of V on a subspace W. Then for any vectors x and y in V.

$$\begin{aligned} \langle Ex, y\rangle &= \Big\langle Ex, Ey + (1-E)y \Big\rangle \\ &= \langle Ex, Ey\rangle = \langle Ex + (1-E)x, Ey\rangle \\ &= \langle x, Ey\rangle. \end{aligned}$$

From the uniqueness of the operator E^* it follows that $E^* = E$. Now consider the projection E defined by

$$E(x_1, x_2, x_3) = \left(\frac{3x_1 + 12x_2 - x_3}{154}\right)(3, 12, -1).$$

Example (contd...)

In the standard orthonormal basis $\gamma = \{e_1, e_2, e_3\}$, the matrix representation of the linear operator E is

$$A = \frac{1}{154}\begin{bmatrix} 9 & 36 & -3 \\ 36 & 144 & -12 \\ -3 & -12 & 1 \end{bmatrix}.$$

Since $E = E^*$, A is also the matrix of E^*, and

$$A = [E^*]_\gamma = [E]^*_\gamma = A^*.$$

LA-2(P-62)E-50

Example

On the other hand, suppose $v_1 = (154, 0, 0)$, $v_2 = (145, -36, 3)$, $v_3 = (-36, 10, 12)$. Then $\{v_1, v_2, v_3\}$ is a basis, and $Ev_1 = (9, 36, -3), Ev_2 = (0, 0, 0), Ev_3 = (0, 0, 0)$.

Since $(9, 36, -3) = -(154, 0, 0) - (145, -36, 3)$, the matrix B of E in the basis $\{v_1, v_2, v_3\}$ is given by $B = \begin{bmatrix} -1 & 0 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$.

In this case $B \neq B^*$, and B^* is not the matrix of $E^* = E$ in the basis $\{\alpha_1, \alpha_2, \alpha_3\}$. Hence

$$B = [E^*]_\beta \neq [E]_\beta^* = B^*.$$

Applying the corollary, we conclude that $\{\alpha_1, \alpha_2, \alpha_3\}$ is not an orthonormal basis.

Example 21.

Let T be the linear operator on $\mathbb{C}^2$ defined by

$$T(a_1, a_2) = (2ia_1 + 3a_2, a_1 - a_2).$$

If β is the standard ordered basis for $\mathbb{C}^2$, then

$$[T]_\beta = \begin{pmatrix} 2i & 3 \\ 1 & -1 \end{pmatrix}.$$

So

$$[T^*]_\beta = [T]_\beta^* = \begin{pmatrix} -2i & 1 \\ 3 & -1 \end{pmatrix}.$$

Hence

$$T^*(a_1, a_2) = (-2ia_1 + a_2, 3a_1 - a_2).$$

Examples : Adjoints of some linear operators

Let A be an $m \times n$ matrix with entries from a field F. We denote by L_A the mapping

$$L_A : F^n \to F^m \quad \text{defined by} \quad L_A(x) = Ax$$

for each column vector $x \in F^n$.

We call L_A a **left-multiplication transformation.** Then

(a) $L_{A^*} = (L_A)^*$.
(b) $[(L_A)^*]_\beta = [L_A]_\beta^*$ for any orthonormal basis β for V.

Examples : Adjoints of some linear operators

Example 22.

Let V be $\mathbb{C}^{n\times 1}$, the space of complex $n \times 1$ matrices, with inner product

$$\langle X, Y\rangle = Y^* X.$$

If A is an $n \times n$ matrix with complex entries, the adjoint of the linear operator $X \to AX$ is the operator $X \to A^* X$

because $\langle AX, Y\rangle = Y^* AX = (A^* Y)^* X = \langle X, A^* Y\rangle$.

Note that this is really a special case of the Theorem 19.

Example 23.

This is similar to above Example 22. Let V be $M_{n\times n}(F)$ with the inner product $\langle A, B\rangle = tr(B^*A)$. Let M be a fixed $n \times n$ matrix over F. The inner product on $M_{n\times n}(F)$ is called the **Frobenius inner product***.*

The adjoint of left multiplication by M is left multiplication by M^*. Of course, "left multiplication by M" is the linear operator L_M defined by $L_M(A) = MA$.

$$\begin{aligned} \langle L_M\langle A\rangle, B\rangle &= tr(B^*(MA)) = tr(MAB^*) \\ &= tr(AB^*M) = tr(A(M^*B)^*) \\ &= \langle A, L_M^*(B)\rangle. \end{aligned}$$

Thus $(L_M)^* = L_{M^*}$.
In the computation above, we twice used the characteristic property of the trace function: $tr(AB) = tr(BA)$.

Examples : Adjoints of some linear operators

Example 24.

Let V be the space of polynomials over the field of complex numbers, with the inner product

$$\langle f, g\rangle = \int_0^1 f(t)\overline{g(t)}\, dt.$$

If f is a polynomial, $f = \sum a_k x^k$, we let $\bar{f} = \sum \bar{a}_k x^k$. That is, $\bar{f}$ is the polynomial whose associated polynomial function is the complex conjugate of that for f:

$$\bar{f}(t) = \overline{f(t)}, \quad t \text{ real}.$$

Consider the operator "multiplication by f"; that is, the linear operator M_f, defined by $M_f(g) = fg$. Then this operator has an adjoint, namely, multiplication by $\bar{f}$. For

$$\begin{aligned} \langle M_f g, h\rangle &= \langle fg, h\rangle = \int_0^1 f(t)g(t)\overline{h(t)}\, dt = \int_0^1 g(t)[\overline{\overline{f(t)}h(t)}]\, dt \\ &= \langle g, \bar{f}h\rangle = \langle g, M_{\bar{f}}h\rangle \qquad \text{and so } (M_{\bar{f}})^* = M_f. \end{aligned}$$

We see that the adjoint operation, passing from T to T^*, behaves somewhat like conjugation on complex numbers. The following theorem suggests the analogy between the conjugates of complex numbers and the adjoints of linear operators.

Theorem 25.

Let V be a finite-dimensional inner product space, and let T and U be linear operators on V. Then

(a) $(T+U)^* = T^* + U^*$;

(b) $(cT)^* = \bar{c}T^*$ *for any* $c \in F$;

(c) $(TU)^* = U^*T^*$;

(d) $T^{**} = T$;

(e) $I^* = I$.

The same proof works in the infinite-dimensional case, provided that the existence of T^* and U^* is assumed.

The Adjoint of a Linear Operator

The above result is often phrased as follows : The mapping $T \to T^*$ is a conjugate-linear anti-isomorphism of period 2.

The analogy with complex conjugation which we mentioned above is, of course, based upon the observation that complex conjugation has the properties

$$\overline{z_1 + z_2} = \bar{z}_1 + \bar{z}_2, \quad \overline{(z_1 z_2)} = \bar{z}_1 \bar{z}_2, \quad \bar{\bar{z}} = z.$$

One must be careful to observe the reversal of order in a product, which the adjoint operation imposes :

$$(UT)^* = T^* U^*.$$

We shall discuss extensions of this analogy as we continue our study of linear operators on an inner product space. We might mention something along these lines now. A complex number z is real if and only if $z = \bar{z}$. One might expect that the linear operators T such that $T = T^*$ behave in some way like the real numbers. This is in fact the case.

For example, if T is a linear operator on a finite-dimensional **complex inner product space**, then

$$T = U_1 + iU_2$$

where $U_1 = U_1^*$ and $U_2 = U_2^*$. Thus, in some sense, T has a 'real part' and an 'imaginary part'. The operators U_1 and U_2 satisfying $U_1 = U_1^*$ and $U_2 = U_2^*$, and are given by

$$U_1 = \frac{1}{2}(T + T^*) \qquad U_2 = \frac{1}{2i}(T - T^*).$$

Exercise 26.

Let $T : V \to W$ be a linear transformation, where V and W are finite-dimensional inner product spaces with inner products $\langle .,. \rangle_1$ and $\langle .,. \rangle_2$, respectively.

Prove the following results.

1. *There is a unique adjoint T^* of T, and T^* is linear.*
2. *If β and γ are orthonormal bases for V and W respectively, then $[T^*]_\gamma^\beta = ([T]_\beta^\gamma)^*$.*
3. *$rank(T^*) = rank(T)$.*
4. *$\langle T^*x, y \rangle_1 = \langle x, Ty \rangle_2$ for all $x \in W$ and $y \in V$.*
5. *For all $x \in V$, $T^*T(x) = 0$ if and only if $Tx = 0$.*

LA-2(P-48)E-40

Definition 27.

Let

$$T : V \to W$$

be a linear transformation, where V and W are finite-dimensional inner product spaces with inner products $\langle ., . \rangle_1$ and $\langle ., . \rangle_2$, respectively.

A function

$$T^* : W \to V$$

is called an **adjoint** *of T if*

$$\langle Tx, y \rangle_2 = \langle x, T^* y \rangle_1 \quad \text{for all } x \in V \text{ and } y \in W.$$

Relation between T and T^*

Let us now relate the kernel and range of a linear operator to those of its adjoint.

Theorem 28.

Let $T : V \to W$ be a linear transformation, where V and W are finite-dimensional inner product spaces. Then

1. $N(T^*) = R(T)^{\perp}$.
2. $R(T^*) = N(T)^{\perp}$.
3. *T is injective if and only if T^* is surjective.*
4. *T is surjective if and only if T^* is injective.*
5. $N(T^*T) = N(T)$.
6. $N(TT^*) = N(T^*)$.
7. $R(T^*T) = R(T^*)$.
8. $R(TT^*) = R(T)$.
9. *If P is a projection onto $R(P)$ along $N(P)$, then P^* is projection onto $N(P)^{\perp}$ along $R(P)^{\perp}$.*

LA-2(P-??)T-??

Exercises 29.

1. *For each of the following inner product spaces V and linear operators T on V, evaluate T^* at the given vector in V.*
   (a) $V = \mathbb{R}^2$, $T(a,b) = (2a+b, a-3b)$, $x = (3,5)$.
   (b) $V = \mathbb{C}^2$, $T(z_1,z_2) = (2z_1 + iz_2, (1-i)z_1)$, $x = (3-i, 1+2i)$.
   (c) $V = P_1(\mathbb{R})$ *with* $\langle f, g\rangle = \int_{-1}^{1} f(t)g(t)\,dt, \quad T(f) = f' + 3f,$
   $f(t) = 4-2t$.

   LA-2(P-35)E-28

2. *Let V be a finite-dimensional inner product space, and let T be a linear operator on V. Prove that if T is invertible, then T^* is invertible and $(T^*)^{-1} = (T^{-1})^*$.* *LA-2(P-36)E-31*

Exercises 30.

1. *Let T be a linear operator on a finite-dimensional inner product space V. Prove the following results.*
   (a) $N(T^*T) = N(T)$. *Deduce that* $rank(T^*T) = rank(T)$.
   (b) $rank(T) = rank(T^*)$. *Deduce from (a) that* $rank(TT^*) = rank(T)$.
   (c) *For any* $n \times n$ *matrix* A, $rank(A^*A) = rank(AA^*) = rank(A)$.

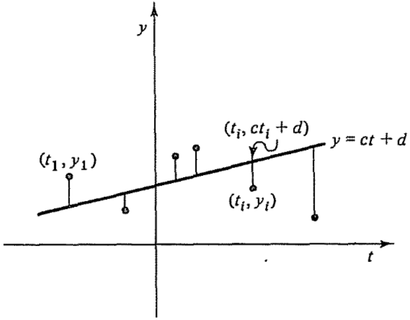
   LA-2(P-40)E-36

2. *Let $T : V \to W$ be a linear transformation, where V and W are finite-dimensional inner product spaces. Prove that* $(R(T^*))^\perp = N(T)$. *LA-2(P-50)E-41*

Least Squares Approximation

Consider the following problem: An experimenter collects data by taking measurements $y_1, y_2, \ldots, y_m$ at times $t_1, t_2, \ldots, t_m$, respectively.

For example, he or she may be measuring unemployment at various times during some period. Suppose that the data $(t_1, y_1), (t_2, y_2), \ldots, (t_m, y_m)$ are plotted as points in the plane.



Least Squares Approximation

From this plot, the experimenter feels that there exists an essentially linear relationship between y and t, say $y = ct + d$, and would like to find the constants c and d so that the line $y = ct + d$ represents the best possible fit to the data collected.

One such estimate of fit is to calculate the error E that represents the sum of the squares of the verticle distances from the points to the line; that is,

$$E = \sum_{i=1}^{m} (y_i - ct_i - d)^2.$$

Least Squares Approximation

Thus the problem is reduced to finding the constants c and d that minimize E. For this reason the line $y = ct + d$ is called the **least squares line.**

If we let

$$A = \begin{pmatrix} t_1 & 1 \\ t_2 & 1 \\ \vdots & \vdots \\ t_m & 1 \end{pmatrix}, \quad x = \begin{pmatrix} c \\ d \end{pmatrix} \quad \text{and} \quad y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}.$$

Then it follows that $E = \|y - Ax\|^2$.

Least Squares Approximation

We develop a general method for finding an explicit vector $x_0 \in \mathbb{K}^n$ that minimizes E; that is, given an $m \times n$ matrix A, we find $x_0 \in \mathbb{K}^n$ such that

$$\|y - Ax_0\| \leq \|y - Ax\|$$

for all vectors $x \in \mathbb{K}^n$.

This method not only allows us to find the linear function that best fits data, but also, for any positive integer n, the best fit using a polynomial of degree at most n.

Least Squares Approximation

First, we need some notation and two simple lemmas. For $x, y \in \mathbb{K}^n$, let $\langle x, y \rangle_n$ denote the **standard inner product** of x and y in $\mathbb{K}^n$.

Recall that if x and y are regarded as column vectors, then $\langle x, y \rangle_n = y^* x$.

Lemma 31.

Let $A \in M_{m \times n}(\mathbb{K}), x \in \mathbb{K}^n$, and $y \in \mathbb{K}^m$. Then

$$\langle Ax, y \rangle_m = \langle x, A^* y \rangle_n.$$

Lemma 32.

Let $A \in M_{m \times n}(\mathbb{K})$. Then rank $(A^* A) =$ rank (A).

Corollary 33.

If A is an $m \times n$ matrix such that rank $(A) = n$, then $A^* A$ is invertible.

Recall

We just recall the following results on an orthonormal basis of a finite-dimensional subspace of an inner product space.

Theorem 34.

Let W be a finite-dimensional subspace of an inner product space V. If $\{x_1, x_2, \dots, x_k\}$ is an orthonormal basis of W and $y \in V$, then

$$y = \sum_{i=1}^{k} \langle y, x_k \rangle x_i + z,$$

where $z \in W^\perp$.
Furthermore, this representation of y is unique. That is, if also $y = u + z'$, where $u \in W$ and $z' \in W^\perp$, then $u = \sum_{i=1}^{k} \langle y, x_i \rangle x_i$ and $z' = z$.

Corollary 35.

In the notation of the above theorem, the vector

$$y_1 = \sum_{i=1}^{k} \langle y, x_i \rangle x_i$$

is the unique vector in W *that is "closest" to* y *; that is, if* $u \in W$*, then*

$$\|y - u\| \geq \|y - y_1\|.$$

In addition,

$$y - y_1 \in W^{\perp}.$$

Least Squares Approximation

Now let A be an $m \times n$ matrix and $y \in F^m$. Define $W = \{Ax : x \in F^n\}$; that is, $W = R(L_A)$. By the Corollary 35, there exists a unique vector in W that is closest to y.

Call this vector Ax_0 where $x_0 \in F^n$. Then $\|Ax_0 - y\| \leq \|Ax - y\|$ for all $x \in F^n$; so x_0 has the property that $E = \|Ax_0 - y\|$ is minimal, as desired.

To develop a practical method for finding such an x_0, we note from Theorem 34 and its corollary that $Ax_0 - y \in W^{\perp}$; so $\langle Ax, Ax_0 - y \rangle_m = 0$ for all $x \in F^n$.

Thus, by Lemma 31, we have that $\langle x, A^*(Ax_0 - y) \rangle_n = 0$ for all $x \in F^n$; that is, $A^*(Ax_0 - y) = 0$.

So we need only find a solution x_0 to $A^*Ax = A^*y$. If, in addition, we assume that rank $(A) = n$, then by Lemma 32 we have $x_0 = (A^*A)^{-1}A^*y$.

We summarize this discussion in the following theorem.

Theorem 36.

Let $A \in M_{m\times n}(F)$ and $y \in F^m$. Then there exists $x_0 \in F^n$ such that

$$(A^*A)x_0 = A^*y$$

and $\|Ax_0 - y\| \leq \|Ax - y\|$ for all $x \in F^n$.

Furthermore, if rank $(A) = n$, then

$$x_0 = (A^*A)^{-1}A^*y.$$

LA-2(P-52)T-44

Least Squares Approximation

To return to our problem, let us suppose that the data collected are (1,2),(2,3), (3,5) and (4,7). Then $A = \begin{pmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \\ 4 & 1 \end{pmatrix}$ and $y = \begin{pmatrix} 2 \\ 3 \\ 5 \\ 7 \end{pmatrix}$. Hence

$A^T A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \\ 4 & 1 \end{pmatrix} = \begin{pmatrix} 30 & 10 \\ 10 & 4 \end{pmatrix}$. Therefore

$(A^T A)^{-1} = \frac{1}{20} \begin{pmatrix} 4 & -10 \\ -10 & 30 \end{pmatrix}$. Thus

$\begin{pmatrix} c \\ d \end{pmatrix} = x_0 = \frac{1}{20} \begin{pmatrix} 4 & -10 \\ -10 & 30 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ 3 \\ 5 \\ 7 \end{pmatrix} = \begin{pmatrix} 1.7 \\ 0 \end{pmatrix}$.

It follows that the line $y = 1.7t$ is the least squares line. The error E may be computed directly as $\|Ax_0 - y\|^2 = 0.3$.

Least Squares Approximation

Suppose that the experimenter chose the times t_i $(1 \leq i \leq m)$ to satisfy

$$\sum_{i=1}^{m} t_i = 0.$$

Then the two columns of A would be orthogonal. so A^*A would be diagonal matrix. In this case, the computations are greatly simplified.

In practice, the $m \times 2$ matrix A in our least squares application has rank equal to two, and hence A^*A is invertible.

For otherwise, the first column of A is a multiple of the second column. which consists only of ones. But this would occur only if the experimenter collects all the data at exactly one time.

Least Squares Approximation

Finally, the method above may also be applied if, for some k, the experimenter wants to fit a polynomial of degree at most k to the data.

For instance, if a polynomial $y = ct^2 + dt + e$ of degree at most 2 is desired. The appropriate model is

$$x = \begin{pmatrix} c \\ d \\ e \end{pmatrix}, \quad y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix},$$

and

$$A = \begin{pmatrix} t_1^2 & t_1 & 1 \\ \vdots & \vdots & \vdots \\ t_m^2 & t_m & 1 \end{pmatrix}.$$

Minimal Solutions to Systems of Linear Equations

Even when a system of linear equations $Ax = b$ is consistent, there may be no unique solution.

In such cases, it may be desirable to find a solution of minimal norm.

A solution s to $Ax = b$ is called a **minimal solution** of

$$\|s\| \leq \|u\|$$

for all other solutions u.

The next theorem assures that every consistent system of linear equations has a unique minimal solution and provides a method for computing it.

Theorem 37.

Let $A \in M_{m\times n}(F)$ and $b \in F^m$. Suppose that $Ax = b$ is consistent. Then the following statements are true.

(a) *Existence : There exists exactly one minimal solution s of $Ax = b$, and $s \in R(L_{A*})$.*

(b) *Computation : If u satisfies $(AA^*)u = b$, then $s = A^*u$. That is, the vector s is the only solution to $Ax = b$ that lies in $R(L_{A^*})$.*

LA-2(P-56)T-45

To find the minimal solution to this system, we must first find some solution u to $AA^*x = b$. Then we can find s from $s = A^*u$.

Example

Consider the system

$$
\begin{aligned}
x + 2y + z &= 4 \\
x - y + 2z &= -11 \\
x + 5y &= 19.
\end{aligned}
$$

Let $A = \begin{pmatrix} 1 & 2 & 1 \\ 1 & -1 & 2 \\ 1 & 5 & 0 \end{pmatrix}$ and $b = \begin{pmatrix} 4 \\ -11 \\ 19 \end{pmatrix}$.

To find the minimal solution to this system, we must first find some solution u to $AA^*x = b$. Now $AA^* = \begin{pmatrix} 6 & 1 & 11 \\ 1 & 6 & -4 \\ 11 & -4 & 26 \end{pmatrix}$.

Example (contd...)

So we consider the system

$$\begin{aligned} 6x + y + 11z &= 4 \\ x + 6y - 4z &= -11 \\ 11x - 4y + 26z &= 19 \end{aligned}$$

for which one solution is $u = \begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix}$ (Any solution will suffice.) Hence

$s = A^* u = \begin{pmatrix} -1 \\ 4 \\ -3 \end{pmatrix}$ is the minimal solution to the given system.

Exercises 38.

1. *Let V be the space $\mathbb{C}^2$, with the standard inner product. Let T be the linear operator defined by $T(1,0) = (1,-2)$, $T(0,1) = (i,-1)$. If $\alpha = (x_1, x_2)$, find $T^*\alpha$.*
2. *For each of the sets of data that follows, use the least squares approximation to find the best fits with both*
   (i) a linear function and
   (ii) a quadratic function.
   Compute the error E in both cases.
   (a) $\{(-3,9),(-2,6),(0,2),(1,1)\}$
   (b) $\{(1,2),(3,4),(5,7),(7,9),(9,12)\}$
   (c) $\{(-2,4),(-1,3),(0,1),(1,-1),(2,-3)\}$

LA-2(P-58)E-46

Exercise 39.

1. *In physics, Hooke's law states that (within certain limits) there is a linear relationship between the length x of a spring and the force y applied to (or exerted by) the spring. That is, $y = cx + d$, where c is called the* **spring constant**. *Use the following data to estimate the spring constant (the length is given in inches and the force is given in pounds).*

Length x	Force y
3.5	1.0
4.0	2.2
4.5	2.8
5.0	4.3

LA-2(P-60)E-47

Exercise 40.

1. *Find the minimal solution to each of the following systems of linear equations.*

a) $x + 2y - z = 12$

b) $$\begin{aligned} x + 2y - z &= 1 \\ 2x + 3y + z &= 2 \\ 4x + 7y - z &= 4 \end{aligned}$$

c) $$\begin{aligned} x + y - z &= 0 \\ 2x - y + z &= 3 \\ x - y + z &= 2 \end{aligned}$$

d) $$\begin{aligned} x+y+z-w &= 1 \\ 2x - y + w &= 1 \end{aligned}$$

LA-2(P-60)E-48

Exercise 41.

Consider the problem of finding the least squares line $y = ct + d$ corresponding to the m observations $(t_1, y_1), (t_2, y_2), \ldots, (t_m, y_m)$.

We recall the result : Let $A \in M_{m\times n}(F)$ and $y \in F^m$. Then there exists $x_0 \in F^n$ such that $(A^*A)x_0 = A^*y$ and $\|Ax_0 - y\| \leq \|Ax - y\|$ for all $x \in F^n$. Furthermore, if rank$(A) = n$, then $x_0 = (A^*A)^{-1}A^*y$.

Show that the equation $(A^*A)x_0 = A^*y$ takes the form of the normal equations:

$$\left(\sum_{i=1}^{m} t_i^2\right) c + \left(\sum_{i=1}^{m} t_i\right) d = \sum_{i=1}^{m} t_i y_i \quad \text{and} \left(\sum_{i=1}^{m} t_i\right) c + md = \sum_{i=1}^{m} y_i.$$

These equations may also be obtained from the error E by setting the partial derivatives of E with respect to both c and d equal to zero.

Use the second normal equation of (a) to show that the least squares line must pass through the center of mass, $(\bar{t}, \bar{y})$, where

$$\bar{t} = \frac{1}{m}\sum_{i=1}^{m} t_i \quad \text{and} \quad \bar{y} = \frac{1}{m}\sum_{i=1}^{m} y_i.$$

LA-2(P-61)E-49

Exercises 42.

1. *Let T be the linear operator on $\mathbb{C}^2$ defined by $T(1,0) = (1+i,2)$, $T(0,1) = (i,i)$. Using the standard inner product, find the matrix of T^* in the standard ordered basis. Does T commute with T^*?*
2. *Let V be $\mathbb{C}^3$ with the standard inner product. Let T be the linear operator on V whose matrix in the standard ordered basis is defined by*
$$A_{ik} = i^{i+k}, \quad (i^2 = -1).$$
Find a basis for the null space of T^*.

Exercise 43.

Let V be an inner product space and β, γ fixed vectors in V. Show that $T\alpha = \langle \alpha, \beta \rangle \gamma$ defines a linear operator on V. Show that T has an adjoint, and describe T^* explicitly.

Now suppose V is $\mathbb{C}^n$ with the standard inner product, $\beta = (y_1, \ldots, y_n)$, and $\gamma = (x_1, \ldots, x_n)$. What is the j, k entry of the matrix of T in the standard ordered basis? What is the rank of this matrix?

Exercise 44.

Let V be the space of the polynomials over $\mathbb{R}$ of degree less than or equal to 3, with the inner product

$$\langle f, g \rangle = \int_0^1 f(t)g(t)dt.$$

If t is a real number, find the polynomial g, in V such that

$$\langle f, g_1 \rangle = f(t)$$

for all f in V.

Exercise 45.

Let V be the space of the polynomials over R of degree less than or equal to 3, with the inner product

$$\langle f, g \rangle = \int_0^1 f(t)g(t)dt.$$

Let D be the differentiation operator on V. Find D^*.

Exercise 46.

Let V be the space of $n \times n$ matrices over the complex numbers, with the inner product

$$\langle A, B \rangle = tr(AB^*).$$

Let P be a fixed invertible matrix in V, and let T_P be the linear operator on V defined by

$$T_P(A) = P^{-1}AP.$$

Find the adjoint of T_P.

Exercises 47.

1. *Let* $H = \mathbb{R}^2$ *and* T *be defined by* $T(1,0) = (1,0)$ *and* $T(0,1) = 2(0,1)$. *Find the matrix* A *representing* T *with respect to* $\{(1,0),(0,1)\}$. *Show that* $A^* = A$ *even though* M *is not symmetric.*
2. *Suppose that a linear operator* $T : \mathbb{K}^n \to \mathbb{K}^m$ *is given by*

$$(Tx)(i) = \sum_{j=1}^{n} k_{ij}x(j), \quad i = 1,2,\ldots,m, \; x \in \mathbb{K}^n.$$

Show that $T^* : \mathbb{K}^m \to \mathbb{K}^n$ *is given by*

$$(T^*y)(j) = \sum_{j=1}^{m} \overline{k_{ij}}y(i), \quad j = 1,2,\ldots,n, \; y \in \mathbb{K}^m.$$

LA-2(P-19)E-16

Exercise 48.

Let $H = \mathbb{K}^n$ and $A \in B(H)$. Let $x_1, x_2, \ldots, x_n$ be an algebraic basis for H and A be represented by the matrix (k_{ij}) with respect to $x_1, x_2, \ldots, x_n$. Show that $(\overline{k_{ji}})$ may not represent A^* with respect to $x_1, x_2, \ldots, x_n$.

LA-2(P-23)E-19

Part - 3

Adjoint of an operator between Hilbert spaces

Adjoint of an operator between Hilbert spaces

One of the fundamental properties of a Hilbert space H is the fact that there is a natural correspondence between the vectors in H and the functionals in H^*.

Let y be a fixed vector in H. Consider the function f_y defined on H by

$$f_y(x) = \langle x, y \rangle.$$

It is easy to see that f_y is linear, continuous and $\|f_y\| = \|y\|$. Thus $y \to f_y$ is a norm-preserving mapping of H into H^*.

LA-2(P-63)N-51

This observation would be of no more than passing interest if it were not for the fact that every functional in H^* arises in just this way.

Theorem 49.

Let H be a Hilbert space, and let f be an arbitrary functional in H^*. Then there exists a unique vector y in H such that

$$f(x) = \langle x, y \rangle \quad \text{for every } x \in H.$$

LA-2(P-63)T-52

This result tells us that the norm-preserving mapping of H into H^* defined by

$$y \to f_y \quad \text{where } f_y(x) = \langle x, y \rangle \quad \text{for every } x \in H \tag{4}$$

is actually a mapping of H onto H^*.

Adjoint of an operator between Hilbert spaces

It would be pleasant if (4) were also a linear mapping. This is not quite true, however, for

$$f_{y_1+y_2} = f_{y_1} + f_{y_2} \quad \text{and} \quad f_{\alpha y} = \overline{\alpha} f_y. \tag{5}$$

it is an easy consequence of (5) that the mapping (4) is an isometry, for

$$\|f_x - f_y\| = \|f_{x-y}\| = \|x - y\|.$$

Adjoint of an operator between Hilbert spaces

Exercises 50.

1. *Let H be a Hilbert space, and show that H^* is also a Hilbert space with respect to the inner product defined by*

$$\langle f_x, f_y \rangle = \langle y, x \rangle.$$

In just the same way, the fact that H^* is a Hilbert space implies that H^{**} is a Hilbert space whose inner product is given by

$$\langle F_f, F_g \rangle = \langle g, f \rangle.$$

2. *Let H be a Hilbert space. We have two natural mappings of H into H^{**}, the second of which is onto : the Banach space natural imbedding $x \to F_x$, where $F_x(f) = f(x)$, and the product mapping $x \to f_x \to F_{f_x}$, where $f_x(y) = \langle y, x \rangle$ and $F_{f_x}(f) = \langle f, f_x \rangle$. Show that these mappings are equal, and conclude that H is reflexive. Show also that $\langle F_x, F_y \rangle = \langle x, y \rangle$.*

Adjoint of an operator between Hilbert spaces

Let H be a Hilbert space and $h \in H$ be fixed. It is clear that the expression $\langle Tx, y \rangle$ is a scalar-valued continuous linear function of x. By Theorem (49), there exists a unique vector z such that

$$\langle Tx, y \rangle = \langle x, z \rangle \quad \text{for all } x.$$

We now write $z = T^* y$, and since y is arbitrary, we have

$$\langle Tx, y \rangle = \langle x, T^* y \rangle \quad \text{for all } x, y. \tag{6}$$

The fact that T^* is uniquely determined by (6).

Properties of Adjoints

1. T^* is an operator on H (it maps H into itself)
2. T^* is linear
3. T^* is continuous
4. $\|T^*\| = \|T\|$.

These facts tell us that $T \to T^*$ is a mapping of $B(H)$ into itself. This mapping is called the **adjoint operation** on $B(H)$.

Exercise 51.

Show that the adjoint operation is one-to-one onto as a mapping of $B(H)$ into itself.

The presence of the adjoint operation is what distinguishes the theory of the operators on H from the more general theory of the operators on a reflexive Banach space.

Kakutani

Adjoint of an operator between Hilbert spaces

Theorem 52.

Let H and K be Hilbert spaces and $T \in B(H, K)$. Then there is a unique $T^* \in B(K, H)$ such that

$$\langle Tx, y\rangle_K = \langle x, T^*y\rangle_H$$

for all $x \in H, y \in K$. The operator T^* is called the **adjoint** *of T.*

LA-2(P-15)T-14

Adjoint of an operator between Hilbert spaces

Theorem 53.

Let $S \in B(H, K)$ and $T \in B(K, H)$. Then

1. $(S + T)^* = S^* + T^*$
2. $(kT)^* = \overline{k}T^*$
3. $(T^*)^* = T$
4. $\|T^*\| = \|T\|$
5. $\|T^*T\| = \|TT^*\| = \|T\|^2$
6. $T^*T = 0 \iff T = 0.$
7. *Let $H = K$. Then $(ST)^* = T^*S^*$, $I^* = I$, and $0^* = 0$, where I is the identity on H.*
8. *If T^{-1} exists and bounded, then $(T^*)^{-1}$ exists and is equal to $(T^{-1})^*$.*

From (1), (2) and (3), the mapping $T \to T^*$ is a conjugate-linear anti-isomorphism of period 2.

Adjoint of an operator between Hilbert spaces

In general, a bounded linear operator on an inner product space need not have an adjoint. The following example illustrates the fact the completeness is essential in Theorem 52.

Example 54.

Let $X = c_{00}$ with the inner product on ℓ^2. If $x \in X$, then

$$\sum_{j-1}^{\infty} \left| \frac{x(j)}{j} \right| \leq \Big(\sum_{j-1}^{\infty} |x(j)|^2 \Big)^{1/2} \Big(\sum_{j-1}^{\infty} \frac{1}{j^2} \Big)^{1/2} = \|x\| \Big(\frac{\pi^2}{6} \Big)^{1/2}.$$

This shows that $f(x) = \sum_{j-1}^{\infty} \frac{x(j)}{j}$, $x \in X$ defines a bounded liear functional on X. Let $A(x) = (f(x), 0, 0, \dots), x \in X$. Then $A : X \to X$ is a bounded linear operator. But A^* does not exist.

LA-2(P-18)E-15

Exercise 55.

Let H be a separable Hilbert space and $\{u_n\}$ be an orthonormal basis for H. Suppose that A and B in $B(H)$ are represented with respect to $\{u_n\}$ by the matrices (a_{ij}) and (b_{ij}) respectively. Show that

(a) *the rows and columns of these matrices (a_{ij}) and (b_{ij}) are square-summable,*

(b) *AB and A^* are represented by (c_{ij}) and (d_{ij}), where*

$$c_{ij} = \sum_{k=1}^{\infty} a_{ik} b_{kj}, \quad d_{ij} = \overline{a_{ji}}.$$

LA-2(P-20)E-17

Exercise 56.

Let H be a separable Hilbert space and $\{u_n\}$ be an infinite orthonormal basis for H. If $A \in B(H)$ is such that

$$A(u_n) = u_{n+1} \quad \text{for} \quad n = 1, 2, \dots,$$

show that

$$A^*(u_1) = 0 \quad \text{and} \quad A^*(u_n) = u_{n-1} \quad \text{for} n = 2, 3, \dots.$$

Show also that $A^*A = I \neq AA^*$, where I is the identity on H. LA-2(P-22)E-18

The adjoint of the left shift operator is the right shift operator.

Exercise 57.

Let H and K be Hilbert spaces and $T \in B(H, K)$.

(a) *If $M \subseteq H$ and $T(M) \subseteq N$, show that $M^{\perp} \supseteq T^*(N^{\perp})$.*

(b) **Converse of (a) :** *If M and N are closed subspaces of H and K respectively, show that $T(M) \subseteq N$ iff $M^{\perp} \supseteq T^*(N^{\perp})$.*

(c) *Prove that $R(T^*) \subseteq N(T)^{\perp}$ and $N(T) = R(T^*)^{\perp}$.*

LA-2(P-25)E-20

Exercises 58.

1. *Let H be a Hilbert space and $A : H \to H$ be a linear map. Suppose that there is a linear map $B : H \to H$ such that*

$$\langle Ax, y \rangle = \langle x, By \rangle \quad \text{for all } x, y \in H.$$

Show that A and B are bounded and $B = A^*$. LA-2(P-27)E-21

2. *Let H be a Hilbert space and $A \in B(H)$. Show that A is bounded below iff A^* is onto.* LA-2(P-28)E-22

Exercises 59.

1. *Let T be a linear operator on an inner product space V. Let $U_1 = T + T^*$ and $U_2 = TT^*$. Prove that $U_1 = U_1^*$ and $U_2 = U_2^*$.*
   LA-2(P-36)E-29
2. *Give an example of a linear operator T on an inner product space V such that $N(T) \neq N(T^*)$.* LA-2(P-36)E-30
3. *Let V be an inner product space, and let $y, z \in V$. Define $T : V \to V$ by*
   $$T(x) = \langle x, y \rangle z$$
   for all $x \in V$. First prove that T is linear. Then show that T^* exists, and find an explicit expression for it. LA-2(P-42)E-37

Exercises 60.

1. *Prove that if* $V = W \oplus W^{\perp}$ *and* T *is the projection on* W *along* $W^{\perp}$*, then* $T = T^*$*. Hint: Recall that* $N(T) = W^{\perp}$*.* LA-2(P-37)E-32
2. *Let* T *be a linear operator on an inner product space* V*. Prove that* $\|Tx\| = \|x\|$ *for all* $x \in V$ *if and only if* $\langle Tx, Ty \rangle = \langle x, y \rangle$ *for all* $x, y \in V$*.* LA-2(P-37)E-33
3. *For a linear operator* T *on an inner product space* V*, prove that* $T^*T = T_0$ *implies* $T = T_0$*. Is the same result true if we assume that* $TT^* = T_0$*?* LA-2(P-38)E-34

Exercises 61.

1. *Let V be an inner product space, and let T be a linear operator on V. Prove the following results.*
   (a) $R(T^*)^{\perp} = N(T)$.
   (b) $R(T^*) \subseteq N(T)^{\perp}$.
   (c) *If V is finite-dimensional, then* $R(T^*) = N(T)^{\perp}$.

LA-2(P-39)E-35

Part - 4

Adjoint of an operator between Banach spaces

If two normed spaces X and Y are isometrically isomorphic, then they are in a sense identical as normed spaces and so should have identical duals. More properly, the spaces X^* and Y^* should also be isometrically isomorphic. LA-2(P-??)E-??

Theorem 62.

Suppose that X and Y are normed spaces such that there is an isomorphism T from X onto Y. Then the map $T^* : T^* \to X^*$ given by the formula $T^*(Y^*) = y^*T$, where y^*T is the usual product of y^* and T, is an isomorphism from Y^* onto X^*, and $\|T^*\| = \|T\|$. If T is an isometric isomorphism, then so is T^*.

Adjoint operators between normed spaces

Notation : We use $\langle x, f\rangle$ for fx when f is a linear functional on X and $x \in X$. The statement $y^*Tx = Sy^*x$ then becomes $\langle Tx, y^*\rangle = \langle x, Sy^*\rangle$, which is easier to grasp.

Proposition 63.

Let X and Y be normed spaces.

1. *If $T \in L(X, Y)$, then T is bounded if and only if*

$$\sup\left\{|\langle Tx, y^*\rangle| : x \in B_X, y^* \in B_{Y^*}\right\}$$

is finite. If T is bounded, then its norm equals this supermum.

2. *If $T \in L(X, Y^*)$, then T is bounded if and only if*

$$\sup\left\{|\langle y, Tx\rangle| : x \in B_X, y \in B_Y\right\}$$

is finite. If T is bounded, then its norm equals this supermum.

Adjoint operators between normed spaces

The map T^* defined in the preceding Theorem 62 is called the **adjoint of** T and is very important in the study of linear operators.

We shall see that every bounded linear operator T from a normed space X into a normed space Y has a bounded linear adjoint T^* defined as in Theorem 62, and T^* is an isomorphism (respectively, an isometric isomorphism) from Y^* onto X^* if and only if T is an isomorphism (respectively, an isometric isomorphism) from X onto Y.

However, it is not true that X and Y must even be isomorphic when there is some isometric isomorphism from X^* onto Y^*.

Exercise 64.

Let c be the Banach space of all convergent sequences of scalars with the vector space operations and norm as given for ℓ_∞ and let c_0 be the collection of all sequences of scalars that converge to 0. Then c_0 is a Banach space since it is a closed subspace of ℓ_∞.

Prove that c^* is isometrically isomorphic to ℓ_1. Notice that c and c_0 are not isometrically isomorphic, even though dual spaces are.

Exercise 65.

Let Y be a dense subspace of a normed space X. Prove that X^* and Y^* are isometrically isomorphic.

Use this to give an example of two normed spaces X and Y such that X^* and Y^* are isometrically isomorphic, even though X and Y are not even isomorphic.

Algebraic adjoint between vector spaces

Suppose that X and Y are vector spaces and that $T \in L(X, Y)$. Then the **algebraic adjoint** of T is the linear operator $T^\#$ from $Y^\#$ into $X^\#$ given by the formula

$$T^\#(y^\#) = y^\# T.$$

That is, the algebraic adjoint $T^\#$ is defined by letting

$$\langle x, T^\# y^\# \rangle = \langle Tx, y^\# \rangle$$

whenever $x \in X$ and $y^\# \in Y^\#$.

With X, Y, and T in the preceding definition, it is immediate $T^\# y^\#$ really is a linear functional on X whenever $y^\#$ in $Y^\#$, and it is easy to check that $T^\#$ is itself linear.

Proposition 66.

Suppose that X and Y are normed spaces and that $T \in L(X, Y)$. Then T is bounded if and only if $T^{\#}(Y^*) \subseteq X^*$. If T is bounded, then the restriction T^* of $T^{\#}$ to Y^*, viewed as a member of $L(Y^*, X^*)$, is itself bounded, and $\|T^*\| = \|T\|$.

Definition 67.

Suppose that X and Y are normed spaces and that $T \in B(X, Y)$. Then the (normed-space) adjoint of T is the restriction of $T^{\#}$ to Y^*, that is, the linear operator T^* from Y^* to X^* given by the formula $T^*(y^*) = y^*T$.

Thus, with all notation as in the preceding definition, the normed-space adjoint T^* of T is defined by letting

$$\langle x, T^*y^* \rangle = \langle Tx, y^* \rangle \quad \text{whenever } x \in X \text{ and } y^* \in Y^*.$$

Adjoint operators between normed spaces

Definition 67 is due to Banach[1] for the case in which X and Y are arbitrary Banach spaces, but adjoints of operators on special spaces were used long before the appearance of Banach's paper in 1929.

Frederic Riesz made use of adjoints in published work appearing in 1910[2] and 1913[3] concerning linear operators on ℓ_2 and certain other L_p spaces.

The idea of an adjoint actually has its roots in matrix theory. LA-2(P-??)E-??

[1] Stefan Banach, Sur les fonctionelles lineaires, Studia Math, 1 (1929), 211-216, 223-239.

[2] Frederic Riesz, Untersuchungen uber systeme integrierbarer, Funktionen, Math. Ann. 69 (1910), 449-497.

[3] Frederic Riesz, Les systemes d'equations linearies a une infinite d'inconnues, Gauthier-Villars, Paries, 1913.

Corollary 68.

If X and Y are normed spaces, then the map $T \mapsto T^*$ is an isometric isomorphism from $B(X, Y)$ into $B(Y^*, X^*)$.

The isometric isomorphism of the preceding corollary does not have to map $B(X, Y)$ onto $B(Y^*, X^*)$, even when both X and Y are Banach spaces.

Exercises 69.

1. *Show that there is no continuous map from $B(\mathbb{K}, c_0)$ onto $B(c_0^*, \mathbb{K}^*)$. Suppose that X and Y are normed spaces.*
   (a) *Prove that if $X \neq \{0\}$ and Y is not reflexive, then some member of $B(Y^*, X^*)$ is not weak*-to-weak* continuous.*
   (b) *Prove that the isometric isomorphism $T \mapsto T^*$ maps $B(X, Y)$ onto $B(Y^*, X^*)$ if and if either $X = \{0\}$ or Y is reflexive.*

Example 70.

Let I be the identity operator on a normed space X. For each x in X and each x^* in X^*,

$$\langle x, I^* x^* \rangle = \langle Ix, x^* \rangle = \langle x, x^* \rangle.$$

It follows that $I^* x^* = x^*$ for each x^* in X^*, that is, that I^* is the identity operator on X^*.

Example 71.

In this example, members of ℓ_1 will also be treated as members of c_0 and ℓ_∞, so to avoid confusion, a subscript of $0, 1$, or ∞ will indicate whether a particular sequence of scalars is being treated as a member of ℓ_0, ℓ_1, or ℓ_∞ respectively.

Let T be the map from ℓ_1 into c_0 given by the formula $T((\alpha_n)_1) = (\alpha_n)_0$. Then T is clearly linear and bounded and is easily seen to have norm 1. Consider ℓ_1^* and c_0^* to be identified with ℓ_∞ and ℓ_1 respectively in the usual way. Then for each pair of elements $(\alpha_n)_1$ and $(\beta_n)_1$ of ℓ_1,

$$\langle (\alpha_n)_1, T^*(\beta_n)_1 \rangle = \langle T(\alpha_n)_1, (\beta_n)_1 \rangle = \langle (\alpha_n)_0, (\beta_n)_1 \rangle = \sum_n \beta_n \alpha_n,$$

so the element $T^*(\beta_n)_1$ of ℓa_1^* can be identified with the element $(\beta_n)_\infty$ of ℓ_∞. In short, tha adjoint of the "identity " map from ℓ_1 into c_0 is the "identity" map from ℓ_1 into ℓ_∞. It is clear that the norm of T^* is 1.

Adjoint operators between normed spaces

The preceding two examples might give one the idea that the adjoint of a one-to-one bounded linear operator between normed spaces must itself be one-to-one. The next example shows, this is not the case.

Example 72.

Let X be any nonreflexive Banach space and let Q_X be the natural map from X into X^{**}, an isometric isomorphism from X onto a closed subspace of X^{**}. Then Q_X^* maps X^{***} into X^*. Let Q_{X^*} be the natural map from X^* into X^{***}. Then for each x in X and each x^* in X^*,

$$\langle x, Q_X^* Q_{X^*} x^* \rangle = \langle Q_X x, Q_{X^*} x^* \rangle = \langle x^*, Q_X x \rangle = \langle x, x^* \rangle,$$

which implies that $Q_X^* Q_{X^*}$ is the identity map on X^* and therefore that Q_X^* maps X^{***} onto X^*. If Q_X^* were also one-to-one, then Q_X^* would have to map X^* onto X^{***}, contradicting the fact that X^* is not reflexive. The isometric isomorphism Q_X therefore does not have a one-to-one ajdoint.

The actual relationships between such properties as being one-to-one and being onto for bounded linear operators and their adjoints will be settled by the following theorems.

Theorem 73.

Suppose that X and Y are normed spaces and that $T \in B(X, Y)$.

(a) *The operator T is one-to-one if and only if $R(T^*)$ is weakly* dense in X^*.*

(b) *The operator T^* is one-to-one if and only if $R(T)$ is dense in Y.*

Theorem 74.

Suppose that X is a Banach space, that Y is a normed space, and that $T \in B(X, Y)$. Then T is an isomorphism from X onto Y if and only if T^* is an isomorphism from Y^* onto X^*. The same is true if "isomorphism" is replaced by "isometric isomorphism."

Theorem 75 (Closed Range Theorem).

Suppose that X and Y are Banach spaces and that $T \in B(X, Y)$. Then the following are equivalent.

(a) *The set $R(T)$ is closed.*

(b) *The set $R(T^*)$ is closed.*

(c) *The set $R(T^*)$ is weakly* closed.*

Theorem 76.

Suppose that X and Y are Banach spaces and that $T \in B(X, Y)$.

(a) *The operator T maps X onto Y if and only if T^* is an isomorphism from Y^* onto a subspace of X^*.*

(b) *The operator T^* maps Y^* onto X^* if and only if T is an isomorphism from X onto a subspace of Y.*

Adjoint operators between normed spaces

If T is a linear operator between normed spaces X and Y, the adjoint of T is a linear operator between duals in reverse order as explained below.

Let X and Y be normed spaces. Let $T : X \to Y$ be a linear operator. Then we have a natural map $T^* : Y^* \to X^*$ defined by

$$T^* f = f \circ T.$$

That is,

$$T^* f(x) = f(Tx), \text{ for all } f \in Y^* \text{ and } x \in X.$$

The mapping $T^* : Y^* \to X^*$ so defined is called the adjoint (or conjugate) of T.

Exercise 77.

Prove the following :

1. *If* $T \in B(X, Y)$*, then* $T^* \in B(Y^*, X^*)$ *and*

$$\|T^*\|_{B(X,Y)} = \|T\|_{B(Y^*,X^*)}.$$

2. *The mapping* $T \mapsto T^*$ *(via* $B(X, Y) \to B(Y^*, X^*)$*) is linear, injective and isometry.*
3. *Let* Z *be a normed space. If* $T \in B(X, Y)$ *and* $U \in B(Y, Z)$*, then* $(UT)^* = T^*U^*$*.*

An Extension of T

Theorem 78.

Let X, Y be normed spaces and $T \in B(X, Y)$. Then the second adjoint

$$T^{**} := (T^*)^* : X^{**} \to Y^{**}$$

is an extension of T.

If X is reflexive, then

$$T^{**} = T.$$

Relation between adjoint and inverse of T

Does the inverse of adjoint of T equal to the adjoint of inverse of T ?

Theorem 79.

Let X be a Banach space. Let Y be a normed space and $T \in B(X, Y)$. If T has a bounded inverse T^{-1} (with domain Y), then T^* has a bounded inverse $(T^*)^{-1}$ (with domain X^*).

Moreover,

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

Converse part

Exercises 80.

1. *Prove the converse part of Theorem (79) in the case $Y = X$ and it is a Banach space.*
2. *Prove the converse part of Theorem (79).*

References

- Stephen H. Friedberg, Arnold J. Insel and Lawrence E. Spence, *Linear Algebra - Fourth Edition*, Prentice-Hall, 2014 (pages mainly from 331 to 369).
- Kenneth Hoffman and Ray Kunze, *Linear Algebra - Second Edition*, Prentice-Hall of India Private Limited, New Delhi, 2017 (pages mainly from 290 to 299).
- A. Ramachandra Rao and P. Bhimasankaram, *Linear Algebra - Second Edition*, Hindustan Book Agency, New Delhi, 2000 (pages mainly from 143 to 146).
- Robert E. Megginson, *An Introduction to Banach Space Theory*, Springer, 1991.
- G.F. Simmons, *Introduction to Topology and Modern Analysis*, McGraw-Hill Book Company, New Delhi, 1963. (pages mainly from 262 to 266, 324, 325).
- Kakutani. S., and G.W. Makey, *Ring and Lattice Characterizations of Complex Hilbert Spaces*, Bulletin of AMS, 52 (1946), 727-733.