Orthogonality and the Adjoint of a Matrix

Remember that if A is an $m \times n$ real matrix, $A^t A$ is, among other things, only the zero matrix if every entry of A is zero.

This doesn't hold for complex numbers!

Example 1: Let

$$A = \begin{bmatrix} 1 & 1 \\ i & i \end{bmatrix}$$

$$A^t A = \begin{bmatrix} 1 & i \\ 1 & i \end{bmatrix} \begin{bmatrix} 1 & 1 \\ i & i \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

The replacement for A^t in $M_n(\mathbb{C})$ is called the adjoint of A.

Definition: If A is an $m \times n$ complex matrix, define the adjoint of A to be the $n \times m$ matrix A^* with

$$\boxed{(A^*)_{j,i} = \overline{A}_{i,j}}$$

for all $1 \leq i \leq m$, $1 \leq j \leq n$.

Matlab Calling Command

" $*$ " is already taken for multiplication, so

$\boxed{A'}$ gives the adjoint of A

Example 2:

If again $A = \begin{bmatrix} 1 & 1 \\ i & i \end{bmatrix}$,

$$A^* = \begin{bmatrix} 1 & -i \\ 1 & -i \end{bmatrix} \text{ and }$$

$$A^*A = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}$$

One can check that if A is an $m \times n$ complex matrix and B is an $n \times k$ complex matrix,

$$\boxed{(AB)^* = B^* A^*}$$

If

$x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_n) \in \mathbb{C}^n$

we may define an inner product

on $\mathbb{C}^n$ by

$$\boxed{x^* y = \sum_{i=1}^{n} \bar{x}_i y_i}$$

(physicists inner product)

Proposition: (inner-product properties)

Let $x_1, x_2, y \in \mathbb{C}^n$ and let $\alpha \in \mathbb{C}$.

1) $(x_1 + x_2)^* y = x_1^* y + x_2^* y$

2) $y^*(x_1 + x_2) = y^* x_1 + y^* x_2$

3) $(\alpha x)^* y = \bar{\alpha}\, x^* y$

4) $x^*(\alpha y) = \alpha\, x^* y$

You may also see $x^* y$ denoted by $\langle x, y \rangle$.

The inner product yields the 2-norm on $\mathbb{C}^n$ given by

$$\boxed{\|x\|_2 = (x^* x)^{1/2}}$$

The inner product and the properties of the adjoint give that for $x, y \in \mathbb{C}^n$ and $A \in M_n(\mathbb{C})$,

$$
\begin{aligned}
& x^*(A^*y) \\
&= (x^*A^*)y \\
&= (Ax)^*y
\end{aligned}
$$

Using the bracket notation,

$$\boxed{\begin{aligned} \langle Ax, y\rangle &= (Ax)^* y \\ &= x^*(A^* y) \\ &= \langle x, A^* y\rangle \end{aligned}}$$

This is a very important property and appears implicitly in many calculations.

Definition: (orthogonality)

If $x, y \in \mathbb{C}^n$, we say

x and y are orthogonal

if

$$\boxed{x^* y = 0}.$$

Orthogonal = "right-angled"

A set of vectors $S \subseteq \mathbb{C}^n$ is orthogonal if $x^* y = 0$ $\forall$ $x, y \in S$, $x \neq y$.

Example 3: Let $x = \begin{bmatrix} 3 \\ i \end{bmatrix} \in \mathbb{C}^2$.

Then if $y = \begin{bmatrix} i \\ 3 \end{bmatrix}$

$$x^* y = \begin{bmatrix} 3 & -i \end{bmatrix} \begin{bmatrix} i \\ 3 \end{bmatrix}$$

$$= 3i - 3i = 0,$$

so x and y are orthogonal.

In general, if $x = \begin{bmatrix} \alpha \\ \beta \end{bmatrix} \in \mathbb{C}^2$,

$y = \begin{bmatrix} -\bar{\beta} \\ \bar{\alpha} \end{bmatrix}$, then y

is orthogonal to x.

Theorem: If $S \subseteq \mathbb{C}^n$ is a set of nonzero, orthogonal vectors, then S is linearly independent.

proof: Let $v_1, \dots, v_k \in S$ and suppose $\exists\ \alpha_1, \dots, \alpha_k \in \mathbb{C}$ with

$$\sum_{i=1}^{k} \alpha_i v_i = \vec{0}.$$

Then fixing j, $1 \le j \le k$,

$$0 = v_j^* \left(\sum_{i=1}^{k} \alpha_i v_i \right)$$

$$= \sum_{i=1}^{k} v_j^* (\alpha_i v_i)$$

$$= \sum_{i=1}^{k} \alpha_i v_j^* v_i$$

$$= \alpha_j v_j^* v_j$$

Since $v_j \neq 0$, $v_j^* v_j \neq 0$,

and so dividing by

$v_j^* v_j$, we get

$$\boxed{\alpha_j = 0}$$

Since $1 \leq j \leq n$ is arbitrary,

$\alpha_j = 0$ for all $1 \leq j \leq n$,

which shows S is linearly

independent.

Definition: A set of vectors $S \subseteq \mathbb{C}^n$ is called orthonormal if whenever $v, w \in S$,

$$v^\dagger w = \begin{cases} 0, & w \neq v \\ 1, & w = v. \end{cases}$$

By the previous theorem, any orthonormal set is linearly independent. If $S \subseteq \mathbb{C}^n$ has exactly n elements and is orthonormal, then we say S is an orthonormal basis for $\mathbb{C}^n$.

If $S \subseteq \mathbb{C}^n$ is an orthogonal set, then by Gram-Schmidt (more on this later), we may express any $x \in \mathbb{C}^n$ uniquely as $x = y + z$ with $y \in \operatorname{span}(S)$ and $z \in S^\perp$ where

$$S^\perp = \{ w \in \mathbb{C}^n : w^* v = 0 \ \forall v \in S \}.$$

__Example 4:__ If $x = \begin{bmatrix} 1 \\ -1 \\ i \end{bmatrix} \in \mathbb{C}^3$

and $S = \left\{ \underbrace{\begin{bmatrix} i \\ 0 \\ 3 \end{bmatrix}}_{v_1}, \underbrace{\begin{bmatrix} 3 \\ 0 \\ i \end{bmatrix}}_{v_2} \right\}$,

then $x = y + z$ where

$z = x - y$ (cheap!) and

$$y = \frac{x^* v_1}{v_1^* v_1} v_1 + \frac{x^* v_2}{v_1^* v_2} v_2$$

Calculating,

$$y = \begin{bmatrix} 1 \\ 0 \\ i \end{bmatrix}$$

$$z = \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix}$$

Note $v_1^* z = v_2^* v = 0$.

Definition: (unitary) A matrix $Q \in M_n(\mathbb{C})$ is called **unitary** if $\boxed{Q^* Q = I_n}$

where I_n is the $n \times n$ identity matrix:

$$(I_n)_{i,j} = \begin{cases} 0, & i \neq j \\ 1, & i = j \end{cases}$$

$$= \delta_{i,j}$$

(Kronecker delta)

This is equivalent to, writing the columns of Q as $q_1, q_2, \dots, q_n$,

$S = \{q_i\}_{i=1}^{n}$ is an orthonormal basis for $\mathbb{C}^n$.

Example Let $x = \begin{bmatrix} 2-i \\ 5+6i \end{bmatrix}$.

Then $\|x\|_2 = \sqrt{66}$,

and if $y = \begin{bmatrix} -5+6i \\ 2+i \end{bmatrix}$,

$$Q = \frac{1}{\sqrt{66}} \begin{bmatrix} x & y \end{bmatrix}$$

$$= \frac{1}{\sqrt{66}} \begin{bmatrix} 2-i & -5+6i \\ 5+6i & 2+i \end{bmatrix}$$

is unitary

Properties of Unitary Matrices

Let $Q \in M_n(\mathbb{C})$. Then the following properties are equivalent:

1) Q is unitary

2) $Q^* = Q^{-1}$

3) $\|Qx\|_2 = \|x\|_2 \quad \forall x \in \mathbb{C}^n$,

4) $QQ^* = I_n$