21-241 MATRICES AND LINEAR TRANSFORMATIONS
SUMMER 1 2012
COURSE NOTES
JUNE 12

PAUL MCKENNEY

Definition. If $x, y \in \mathbb{C}^n$, their *inner product* is defined to be

$$\langle x, y \rangle = \sum_{k=1}^{n} x_k \overline{y_k}$$

We say that x and y are *orthogonal*, and write $x \perp y$, if $\langle x, y \rangle = 0$.

Note that, viewing y^H and x as $1 \times n$ and $n \times 1$ size matrices respectively, we have

$$\langle x, y \rangle = y^H x$$

In the case where $x, y \in \mathbb{R}^n$, all the complex conjugates drop away, and we can write $y^\top$ instead of y^H.

Example. Let p and q be points in the plane $\mathbb{R}^2$. Show that if p and q lie on the unit circle, then $\langle p, q \rangle$ is exactly $\cos(\theta)$, where θ is the angle between p and q on the unit circle (equivalently, the length of the arc between p and q, centered at 0.) Show that $p \perp q$ if and only if the lines L and R, going through 0 and p, and 0 and q respectively, are perpendicular.

Proof. Suppose p and q lie on the unit circle; then we can write their coordinates down as $p = (\cos\varphi, \sin\varphi)$ and $q = (\cos\psi, \sin\psi)$ for some φ and ψ, and the angle between them is exactly $\theta = |\varphi - \psi|$. Now,

$$\langle p, q \rangle = \cos\varphi \cos\psi + \sin\varphi \sin\psi = \cos(\varphi - \psi)$$

Since cos is an even function, $\cos(\varphi - \psi) = \cos\theta$.

The other part of the problem follows from this part; $p \perp q$ if and only if θ is a zero of cos, if and only if θ is one of $\pi/2$ or $3\pi/2$, if and only if the lines L and R described are orthogonal. □

Before we continue we'll need to recall some facts about complex conjugation, ie the map $z \mapsto \bar{z}$ given by $a + bi \mapsto a - bi$. In short, it's as nice as you would want.

Fact 1. For all $z, w \in \mathbb{C}$,

(1) $\overline{z + w} = \bar{z} + \bar{w}$,

(2) $\overline{zw} = \bar{z}\bar{w}$,
(3) $\bar{\bar{z}} = z$.

Example. Which of the above algebraic properties hold for the hermitian operator? Ie, which of

(1) $(A+B)^H = A^H + B^H$,
(2) $(AB)^H = A^H B^H$,
(3) $(A^H)^H = A$

is true for all $A, B \in M_n(\mathbb{C})$? If one of them is wrong, what's the "right" version?

Fact 2. (1) For all $x, y, z \in \mathbb{C}^n$, $\langle x+y, z\rangle = \langle x, z\rangle + \langle y, z\rangle$.
(2) For all $x, y \in \mathbb{C}^n$ and $\lambda \in \mathbb{C}$, $\langle \lambda x, y\rangle = \lambda \langle x, y\rangle$.
(3) For all $x, y \in \mathbb{C}^n$, $\langle x, y\rangle = \overline{\langle y, x\rangle}$.

This fact is often summarized in the following way. (We often say that the inner product is linear in its first argument, and *conjugate*-linear in its second argument.)

Corollary 1. *For all* $x_1, x_2, y_1, y_2 \in \mathbb{C}^n$ *and* $\lambda_1, \lambda_2, \mu_1, \mu_2 \in \mathbb{C}$,

$$\langle \lambda_1 x_1 + \lambda_2 x_2, \mu_1 y_1 + \mu_2 y_2\rangle = \lambda_1\overline{\mu_1}\langle x_1, y_1\rangle + \lambda_1\overline{\mu_2}\langle x_1, y_2\rangle + \lambda_2\overline{\mu_1}\langle x_2, y_1\rangle + \lambda_2\overline{\mu_2}\langle x_2, y_2\rangle$$

It's useful to keep in mind that when $x, y \in \mathbb{R}^n$, all of the complex conjugates above disappear.

Fact 3. For all $x \in \mathbb{C}^n$, $\langle x, x\rangle$ is nonnegative.

Definition. The *norm* of a vector $x \in \mathbb{C}^n$ is defined to be

$$\|x\| = \sqrt{\langle x, x\rangle} = \left(\sum_{k=1}^{n} |x_k|^2\right)^{\frac{1}{2}}$$

The *distance between* two vectors $x, y \in \mathbb{C}^n$ is $\|x - y\|$.

Fact 4. (1) If $x \in \mathbb{C}^n$ and $\lambda \in \mathbb{C}$, then $\|\lambda x\| = |\lambda|\,\|x\|$,
(2) If $x, y \in \mathbb{C}^n$ and $x \perp y$, then $\|x+y\|^2 = \|x\|^2 + \|y\|^2$.

Fact 5. (Cauchy-Schwartz Inequality) For any $x, y \in \mathbb{C}^n$, we have

$$|\langle x, y\rangle|^2 \le \|x\|\,\|y\|$$

Moreover, the equality above holds if and only if x and y are linearly independent.

Proof. If $y = 0$ then we're done. (Why?) So assume $y \neq 0$; then $\|y\| \neq 0$. Let

$$z = x - \frac{\langle x, y\rangle}{\langle y, y\rangle} y = x - \frac{\langle x, y\rangle}{\|y\|^2} y$$

Then,

$$\langle z, y \rangle = \left\langle x - \frac{\langle x, y \rangle}{\langle y, y \rangle} y, y \right\rangle = \langle x, y \rangle - \frac{\langle x, y \rangle}{\langle y, y \rangle} \langle y, y \rangle = 0$$

So $z \perp y$, and it follows that $\frac{\langle x,y \rangle}{\langle y,y \rangle} y$ and z are orthogonal. So by the above fact,

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

Finally, note that

$$x = \frac{\langle x, y \rangle}{\langle y, y \rangle} y + z$$

So we have

$$\|x\|^2 \geq \frac{|\langle x, y \rangle|^2}{\|y\|^2}$$

Moving stuff around and taking a root, we get

$$\|x\| \|y\| \geq \langle x, y \rangle$$

as desired. □

Theorem 1. *Suppose $x_1, \ldots, x_k \in \mathbb{C}^n$ are pairwise orthogonal, ie, $x_i \perp x_j$ for $i \neq j$, and each x_i is nonzero. Then $\{x_1, \ldots, x_k\}$ is linearly independent.*

Proof. Suppose

$$\lambda_1 x_1 + \cdots + \lambda_k x_k = 0$$

Now take the inner product with x_i;

$$0 = \langle 0, x_i \rangle = \left\langle \sum_{j=1}^{k} \lambda_j x_j, x_i \right\rangle = \sum_{j=1}^{k} \lambda_j \langle x_j, x_i \rangle = \lambda_i \|x_i\|^2$$

Since x_i is nonzero, $\|x_i\|^2 \neq 0$. Then $\lambda_i = 0$. □

Theorem 2. *Suppose $A \in M_n(\mathbb{C})$, and $x, y \in \mathbb{C}^n$. Then $\langle Ax, y \rangle = \langle x, A^H y \rangle$.*

Proof. The ith entry of Ax is

$$(Ax)_i = \sum_{j=1}^{n} A_{ij} x_j$$

Similarly, the jth entry of $A^H y$ is

$$(A^H y)_j = \sum_{i=1}^{n} (A^H)_{ji} y_i = \sum_{i=1}^{n} \overline{A_{ij}} y_i$$

So,

$$\begin{aligned}
\langle Ax, y\rangle &= \sum_{i=1}^{n}\left(\sum_{j=1}^{n} A_{ij}x_j\right)\overline{y_i} \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} A_{ij}x_j\overline{y_i} \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} x_j\overline{\overline{A_{ij}}y_i} \\
&= \sum_{j=1}^{n} x_j\sum_{i=1}^{n}\overline{\overline{A_{ij}}y_i} \\
&= \sum_{j=1}^{n} x_j\overline{(A^H y)_j} \\
&= \left\langle x, A^H y\right\rangle
\end{aligned}$$

□