Lecture Notes for Math 104: Fall 2010 (Week 3)

Jacob Bernstein

CHAPTER 1

Seventh Lecture and Eighth Lecture

In this lecture we introduced the hermitian inner product and adjoint. These are generalizations to the complex settings of the dot product and transpose which are defined for real vectors and matrices. This is an amalgamation of the seventh and eigth lectures.

1. Notation

We remind our selves of some notation before preceding further

Let $A \in \mathbb{C}^{m\times n}$ be a $m \times n$ matrix. We've often expressed A as a set of columns:

$$A = \begin{bmatrix} \mathbf{a}_1 | & \dots & | \mathbf{a}_n \end{bmatrix}$$

here $\mathbf{a}_i$ is also a $m \times 1$ matrix We can also write A in terms of its rows

$$A = \begin{bmatrix} a_1' \\ \vdots \\ a_m' \end{bmatrix}$$

where a_j' is a $1 \times n$ matrix. As we've seen multiplying A on the right by a vector

$$\mathbf{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$$

is the same as getting a linear combination of the columns, that is

$$A\mathbf{v} = \sum_{j=1}^{n} v_j \mathbf{a}_j$$

In a similar manner multiplying A on left by a row $q = \begin{bmatrix} q_1 & \dots & q_m \end{bmatrix}$ gives a linear combination of the rows:

$$q\mathbf{v} = \sum_{i=1}^{m} q_i a_i' \in \mathbb{C}^{1\times n}$$

We then have rules for matrix multiplication if $B \in \mathbb{C}^{n\times k}$ and

$$B = \begin{bmatrix} \mathbf{b}_1 | & \dots & | \mathbf{b}_k \end{bmatrix}$$

then

$$AB = \begin{bmatrix} A\mathbf{b}_1 | & \dots & | A\mathbf{b}_n \end{bmatrix}$$

Similarly, if we write

$$B = \begin{bmatrix} b_1' \\ \vdots \\ b_m' \end{bmatrix}$$

then

$$AB = \begin{bmatrix} a'_1 B \\ \vdots \\ a'_m B \end{bmatrix}$$

In either case this yields:

$$AB = \begin{bmatrix} a'_1\mathbf{b}_1 & \cdots & a'_1\mathbf{b}_k \\ \vdots & a'_i\mathbf{b}_j & \vdots \\ a'_m\mathbf{b}_1 & \cdots & a'_m\mathbf{b}_k \end{bmatrix}$$

2. Adjoints

We now introduce an important formal operation on matrices. Let

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \in \mathbb{C}^{m\times n}, B = \begin{bmatrix} b_{11} & \cdots & b_{1m} \\ \vdots & \ddots & \vdots \\ b_{n1} & \cdots & b_{nm} \end{bmatrix} \in \mathbb{C}^{n\times m}$$

We say that B is the *hermitian conjugate* or *adjoint* of A when and only when $b_{ij} = \bar{a}_{ji}$ and write $B = A^*$. Another way to think about this is: Let $\mathbf{v} \in \mathbb{C}^m$ be the vector $\mathbf{v} = (v_1, \dots, v_m)$. we define $\mathbf{v}^* = [\bar{v}_1, \dots, \bar{v}_m]$. For a matrix $A \in \mathbb{C}^{m\times n}$ we write $A = \begin{bmatrix}\mathbf{a}_1 | & \cdots & | \mathbf{a}_n\end{bmatrix}$ and define

$$A^* = \begin{bmatrix} \mathbf{a}_1^* \\ \vdots \\ \mathbf{a}_n^* \end{bmatrix}$$

When $A \in \mathbb{R}^{m\times n}$ is a real matrix we have that $\bar{a}_{ij} = a_{ij}$ and so all are doing when taking the adjoint is flipping. You've probaby see this before and called it the transpose $A^\top$.

An important class of matrices are the *hermitian* and *symmetric* matrices. A hermitian matrix is one so that $A^* = A$. A symmetric matrix is just a hermitian matrix that has real entries (and so $A^\top = A$). Notice any such matrix (in either case) is square.

3. Inner Products

We may think of a vector $\mathbf{v} \in \mathbb{C}^n$ as a $n \times 1$ matrix. Hence we can write $\mathbf{v}^*$ to get a $1 \times n$ matrix. We define the innerproduct of two vectors $\mathbf{v}$ and $\mathbf{w} \in \mathbb{C}^n$ as:

$$\langle \mathbf{v}, \mathbf{w} \rangle = \mathbf{v}^*\mathbf{w}$$

Notice that when the vectors have real entries (i.e. $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$) this is just the usual dot product. This suggests we use the inner product to define a notion of "length" of a vector. That is set:

$$||\mathbf{v}||_2 = \sqrt{\mathbf{v}^*\mathbf{v}} \geq 0.$$

We point out that $\mathbf{v}^*\mathbf{v} \geq 0$ for any vector so the squareroot is okay. We included complex conjugation precisely to achieve this. When $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$ we can really geometrically think of $||\mathbf{v}||_2$ as the length of $\mathbf{v}$ and $\mathbf{v}^*\mathbf{w} = ||\mathbf{v}||_2||\mathbf{w}||_2 \cos\alpha$ where α is the angle between $\mathbf{v}$ and $\mathbf{w}$. For complex vectors this geometric interpretation doesn't make as much sense, but is still useful for intuition.

Some useful properties (left as an exercise). Bilinearity of the innerproduct:

$$(\mathbf{v}_1+\mathbf{v}_2)^*\mathbf{w} = \mathbf{v}_1^*\mathbf{w}+\mathbf{v}_2^*\mathbf{w}$$
$$\mathbf{v}^*(\mathbf{w}_1+\mathbf{w}_2) = \mathbf{v}^*\mathbf{w}_1+\mathbf{v}^*\mathbf{w}_2$$
$$(\alpha\mathbf{v})^*(\beta\mathbf{w}) = \bar{\alpha}\beta\mathbf{v}^*\mathbf{w}$$

These all follow from properties of matrix multiplication. Note this implies $||\lambda\mathbf{v}||_2 = |\lambda|||\mathbf{v}||_2$. The innerproduct satisfies the following inequality known as the Cauchy-Schwarz inequality:

$$|\mathbf{v}^*\mathbf{w}| \leq ||\mathbf{v}||_2||\mathbf{w}||_2.$$

with equality if and only if the $\mathbf{v}$ and $\mathbf{w}$ are collinear. The length also satisfies the so called *triangle inequality*:

$$||\mathbf{v}+\mathbf{w}||_2 \leq ||\mathbf{v}||_2+||\mathbf{w}||_2.$$

Note that both of these are easily shown for real vectors and have nice geometric intrepretations, but they also hold for complex vectors.

For general matrices $A \in \mathbb{C}^{m\times n}$ and $B \in \mathbb{C}^{n\times k}$ one has $(AB)^* = B^*A^*$. In particular if $\mathbf{v} \in \mathbb{C}^m$ and $\mathbf{w} \in \mathbb{C}^n$ then $\mathbf{v}^*(A\mathbf{w}) = (A^*\mathbf{v})^*\mathbf{w}$. That is

$$\langle A\mathbf{v},\mathbf{w}\rangle = \langle \mathbf{v}, A^*\mathbf{w}\rangle$$

This last fact is really key and we will return to it soon.

4. Orthogonality

One useful thing to use the inner product for is to tell if two vectors are *orthogonal*. We say $\mathbf{v}$ and $\mathbf{w}$ are orthogonal if $\mathbf{v}^*\mathbf{w} = 0$. If the vectors are real then this means geometrically that they are perpendicular. An example $\mathbf{e}_i$ and $\mathbf{e}_j$ are orthogonal when $i \neq j$.

We say that two sets of vectors E and F (not neccesarily vector spaces) are *orthogonal* if whenever $\mathbf{v} \in E$ and $\mathbf{w} \in F$ one has $\mathbf{v}^*\mathbf{w} = 0$. A set of non-zero vectors S is *orthogonal* if for any $\mathbf{v} \in S, \mathbf{w} \in S$ with $\mathbf{v} \neq \mathbf{w}$ one has $\mathbf{v}^*\mathbf{w} = 0$. This set is orthonormal if in addition $||\mathbf{v}||_2 = 1$ for all $\mathbf{v} \in S$. Examples include the standard basis and the vectors

$$1/\sqrt{2}\begin{bmatrix} I \\ -I \end{bmatrix}, 1/\sqrt{2}\begin{bmatrix} 1 \\ 1 \end{bmatrix} \in \mathbb{C}^2.$$

Probably the most important thing about orthogonality for our purposes is that it is a simple condition that ensures linear independence.

THEOREM 4.1. *(2.1 in T-B) The vectors in an orthogonal set S are linearly independent.*

PROOF. Suppose one has $\mathbf{v}_i \in S$ $(i = 1..l)$ that are linearly dependent. Then we can write $\mathbf{v}_k = \sum_{i=1}^{l} c_i\mathbf{v}_i$ where $c_k = 0$ and $\mathbf{v}_k$. Then $\mathbf{v}_k^*\mathbf{v}_k = \mathbf{v}_k^*\sum_{i=1}^{l} c_i\mathbf{v}_i = \sum_{i=1}^{l} c_i\mathbf{v}_k^*\mathbf{v}_i$ by the bilinearity . The orthogonality says that the left hand side equls $c_k\mathbf{v}_k^*\mathbf{v}_k = 0$. However, this implies $\mathbf{v}_k = 0$ which is impossible. □

One consequence is that if $S \subset \mathbb{C}^m$ is an orthogonal set then S is a basis of *span*(S). We refer to such a basis as an *orthonormal basis* if in addition for $\mathbf{v} \in S$ $||\mathbf{v}||_2 = 1$. If $S \subset \mathbb{C}^m$ and there are m vectors in S then S is a basis of $\mathbb{C}^m$. Note by normalizing one can always go from an orthogonal set to an orthonormal set. One good property of an orthonormal basis is that it is easy to find the coefficients of a vector with respect to the basis using the inner product. Namely, if $\mathbf{v}_1, \ldots, \mathbf{v}_k$ are an orthormal basis of $E \subset \mathbb{C}^m$ then We can write $\mathbf{v} \in \mathbb{C}^m$ as

$$\mathbf{v} = \sum_{i=1}^{k} \langle \mathbf{v}_i, \mathbf{v} \rangle \mathbf{v} = \sum_{i=1}^{k} (\mathbf{v}_i^* \mathbf{v}) \mathbf{v}$$

i.e. the coefficients are $\mathbf{v}_i^* \mathbf{v}$.

5. Unitary Matrices

We now introduce an important class of matrices that are related to what we just discussed. The idea is that while a basis corresponds to a non-singular matrix, an *Orthonormal basis* corresponds to a *Unitary* matrix. Since expanding components in an orthonormal basis is easier than doing it for a generic basis, so finding inverses for unitary matrices is easier than for non-singular matrices.

We say $Q \in \mathbb{C}^{m \times m}$ is *unitary* if $Q^* = Q^{-1}$ (if $Q \in \mathbb{R}^{m \times m}$ say is it is *orthogonal*). That is if $Q^*Q = QQ^* = I$. It straight forward to check that if $Q = \begin{bmatrix} q_1 & | \ldots | & q_m \end{bmatrix}$ then $Q^* = \begin{bmatrix} q_1^* \\ \vdots \\ q_m^* \end{bmatrix}$ so

$$Q^*Q = \begin{bmatrix} q_1^* q_1 & \ldots & q_1^* q_m \\ \vdots & \ddots & \vdots \\ q_m^* q_1 & \ldots & q_m^* q_m \end{bmatrix}$$

So Q unitary if and only if the columns form an orthonormal basis. In particular $Q^*\mathbf{b}$ gives the coefficients of $\mathbf{b}$ in the orthonormal basis given by the columns of Q. Two important additional properties of unitary matrices are that they preserve innerproduct and $2-norm$ i.e. $(Q\mathbf{v})^*(Q\mathbf{w}) = \mathbf{v}^* Q^* Q \mathbf{w} = vV^* \mathbf{w}$. and So $||Q\mathbf{v}||_2 = ||\mathbf{v}_2||$.

For real matrices this has the geometric interpretation that Q is given by a rigid motion fixing the origin. For instance rotation about the origin or reflection through any line (or plane etc) through the origin.

Example: we check that rotation is Unitary.

$$A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

then $A^*A = A^\top A = Id$ is a straight forward computation.