CSE 522 Notes

William Pentney

February 19, 2001

1 Deep Background

We will be concerned here primarily with complex vectors and matrices, primarily so that we may rid ourselves of them as soon as possible, although we will consider real matrices at points. We shall review some fundamental linear alegbra concepts, with a particular interest in how they pertain to complex structures.

2 Complex Vectors and Matrices

First, we shall review some basic concepts of linear algebra.

Consider two complex vectors $\mathbf{x}, \mathbf{y} \in \mathbf{C^n}$. The **inner product** of two such vectors, written as $(\mathbf{x}, \mathbf{y})$, is defined as

$$\sum_{i=1}^{n} \overline{x_i} y_i$$

where $\overline{x_i} = a - bi$ is the conjugate of $x_i = a + bi$. If $(\mathbf{x}, \mathbf{y}) = \mathbf{0}$, then $\mathbf{x}$ and $\mathbf{y}$ are **orthogonal**. Note that the following properties hold:

1. $(\mathbf{x}, \mathbf{x}) = \mathbf{x_1}^\mathbf{2} + \mathbf{x_2}^\mathbf{2} + ... + \mathbf{x_n}^\mathbf{2} \geq \mathbf{0}$
2. $(\lambda \mathbf{x}, \mathbf{y}) = \lambda(\mathbf{x}, \mathbf{y})$ for λ complex
3. $(\mathbf{x}, \mathbf{y}) = \overline{(\mathbf{y}, \mathbf{x})}$
4. $(\mathbf{x}, \lambda \mathbf{y}) = \overline{\lambda}(\mathbf{x}, \mathbf{y})$ (which follows from the previous two properties).

A set of vectors $\mathbf{x_1}, \mathbf{x_2}, ...\mathbf{x_n}$ is **linearly independent** if there do not exist nonzero constants $c_1, c_2, ..., c_n$ such that $c_1\mathbf{x_1} + \mathbf{c_2 x_2} + ... + \mathbf{c_n x_n} = \mathbf{0}$. If a set of vectors is not linearly independent, then there exists some vector $\mathbf{x_i}$ and constants $c_1, c_2, ...c_n$ such that $c_1\mathbf{x_1} + \mathbf{c_2 x_2} + ... + \mathbf{c_{i-1} x_{i-1}} = \mathbf{0}$.

Now consider the complex matrix $M \in \mathbf{C^{n\times m}}$. The **conjugate transposition** of M, M^H, is defined by

$$M^H_{j,i} = \overline{M_{i,j}}$$

for $1 \leq i \leq n$ and $1 \leq j \leq m$. (For real numbers, $M^H = M^T$, the transpose of M.) We note that $(\mathbf{x}, \mathbf{y}) = \mathbf{x^H y}$.

3 Eigenvalues, Eigenvectors, and Eigenmanifolds

Consider a matrix $A \in \mathbf{C^{n\times n}}$. If we can find a value $\lambda \in \mathbf{C}$ such that $A\mathbf{x} = \lambda\mathbf{x}$ for some $\mathbf{x} \in \mathbf{C^n}$, with $x \neq 0$, then λ is said to be an **eigenvalue** of A and $\mathbf{x}$ is a **right eigenvector** for λ. Note that $(A - \lambda I)\mathbf{x} = \mathbf{0}$, and thus det $(A - \lambda I) = 0$.

Given A and λ as defined above, we call the set $B_\lambda = \{\mathbf{x} \in \mathbf{C^n} | \mathbf{Ax} = \lambda\mathbf{x}\}$ the **right eigenmanifold** of A corresponding to λ. That is, the right eigenmanifold consists of all sets of eigenvectors for all eigenvalues of A. The dimension of B^λ is called the **geometric multiplicity** of λ; alternatively, it is simply called the **multiplicity**.

4 Properties of Eigenvalues and Eigenvectors

Let $\lambda_1, \lambda_2, \lambda_3 ... \lambda_r$ be distinct eigenvalues of matrix A, and $\mathbf{x_1}, \mathbf{x_2}, ...\mathbf{x_r}$ be arbitrary members of $B_{\lambda_1}, B_{\lambda_2}, ..., B_{\lambda_r}$, respectively.

Claim :$\{\mathbf{x_1}, \mathbf{x_2}, ..., \mathbf{x_n}\}$ is a linearly independent set.

Proof : By contradiction. Say that $\{\mathbf{x_1}, \mathbf{x_2}, ..., \mathbf{x_n}\}$ is a linearly dependent set, i.e. there exists a positive integer j and c_1 through c_j such that

$$c_1\mathbf{x_1} + \mathbf{c_2 x_2} + ... + \mathbf{c_j x_j} = \mathbf{0}.$$

Let us say that j is the smallest such integer. We can then apply A to the right side of the equation, and apply λ_i to each term i on the left side, so that we will have

$$\lambda_1 c_1\mathbf{x_1} + \lambda_2\mathbf{c_2 x_2} + ...\lambda_\mathbf{j}\mathbf{c_j x_j} = \mathbf{A0} = \mathbf{0}.$$

All the λ_i must be nonzero; otherwise we have a linear relationship involving $j - 1$ elements, which contradicts our claim that j is the smallest possible integer for which this is true. Let us then note that $\lambda_j \neq 0$; if we multiply both sides of $c_1\mathbf{x_1} + \mathbf{c_2x_2} + ... + \mathbf{c_jx_j} = \mathbf{0}$ by λ_j and subtract from the previous equation, we have

$$(\lambda_1 - \lambda_j)c_1\mathbf{x_1} + (\lambda_\mathbf{2} - \lambda_\mathbf{j})\mathbf{c_2x_2} + ... + (\lambda_\mathbf{j} - \mathbf{1} - \lambda_\mathbf{j})\mathbf{c_{j-1}x_{j-1}} = \mathbf{0}$$

which also contradicts our definition of j. Thus $\{\mathbf{x_1}, \mathbf{x_2}, ...\mathbf{x_n}\}$ forms a linearly independent set. It also then follows that, since the vectors in each eigenfold are linearly independent from the vectors of every other eigenmanifold, A may only have a finite number of eigenmanifolds (and thus eigenvalues).

Previously it was noted that det $(A - \lambda I) = 0$. The **characteristic polynomial** of A, $p(\lambda)$, is equal to det $(A - \lambda I)$, or $(\lambda_1 - \mathbf{x})(\lambda_\mathbf{2} - \mathbf{x})...(\lambda_\mathbf{n} - \mathbf{x})$, where n is the number of eigenvalues of A. (Note that for $\mathbf{x} = \mathbf{0}$, $p(\mathbf{x}) = \lambda_\mathbf{1}\lambda_\mathbf{2}...\lambda_\mathbf{n}$.

We will take the *multiset* - i.e., a set whose members are not necessarily all distinct - of the roots of the characteristic polynomial, $\{\lambda_1, \lambda_2, ..., \lambda_n\}$, and declare it to be the **spectrum** of A. (We permit duplicates in the set so that every root of the characteristic polynomial may appear in the set as many times as it appears in the factored representation of the polynomial).

If $det(A - \lambda I) = 0$, then the rows of $A - \lambda I$ are linearly dependent, and there exists some $\mathbf{x} \neq \mathbf{0}$ such that $A\mathbf{x} - \lambda\mathbf{x} = (\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}$.

One important property of eigenvalues is that, given a matrix

$$A = \begin{bmatrix} A_1 & B \\ 0 & A_2 \end{bmatrix}.$$

where $A_1 \in \mathbf{C^{i \times i}}$ and $A_2 \in \mathbf{C^{n-i \times n-i}}$, the eigenvalues of A are the union of the eigenvalues of A_1, A_2, since det $(A - \lambda I) =$ det $(A_1 - \lambda I)$det$(A_2 - \lambda I)$.

5 Complex Values and Eigenvectors

Previously we defined the concept of right eigenvectors. A **left eigenvector** is a vector $\mathbf{x} \in \mathbf{C^n}$ such that, for some $\lambda \in \mathbf{C}$, $\mathbf{x^H A} = (\lambda\mathbf{x})^\mathbf{H} = \overline{\lambda}\mathbf{x^H}$. Note that a right eigenvector for $\overline{\lambda}$ is a left eigenvector for λ. Additionally, note that since $\mathbf{x^H}(\mathbf{A} - \lambda\mathbf{I}) = \mathbf{0}$ if $\mathbf{x}$ is a left eigenvector, det $(a - \overline{\lambda}I) = 0$; thus we arrive at the same characteristic polynomial for A. (If A is a real matrix, then the left and right eigenvectors will be identical; the distinction between the two is only significant in complex space, where eigenvalues occur in complex conjugate pairs.)

Similarly, we may define the **left eigenmanifold** to be the set of all $\mathbf{x} \in \mathbf{C^n}$ such that $A\mathbf{x^H} = \lambda\mathbf{x^H}$.

6 Similar Matrices

If we have $AX = XB$ for some $A \in \mathbf{C^{n\times n}}$, $X \in \mathbf{C^{n\times k}}$, and $B \in \mathbf{C^{k\times k}}$, then the linear combinations of the columns of X, range(X), is invariant for A. That is, if $x \in$ range(X), then $Ax \in$ range(X). As proof, we note that if $x \in$ range(X), then $\exists v \in \mathbf{C^n}, \mathbf{x} = \mathbf{Xv}, \mathbf{Ax} = \mathbf{AXv} = \mathbf{XBv} = \mathbf{Xv'} \in$ range(X). You may think of the columns of X as the eigenvectors of A corresponding to nonzero eigenvalues.

Note that for a matrix B, if $B\mathbf{y} = \lambda\mathbf{y}$, then $A(X\mathbf{y}) = \mathbf{XBy} = \mathbf{X}\lambda\mathbf{y} = \lambda\mathbf{Xy}$. Thus if $\mathbf{y}$ is a right eigenvector for B, $X\mathbf{y}$ is a right eigenvector for A with some eigenvalue of A. If X has full column rank - that is, if there is no nonzero $\mathbf{y}$ such that $X\mathbf{y} = \mathbf{0}$ - then the eigenvalues of B are a subset of the eigenvalues of A.

Given a non-singular matrix $X \in \mathbf{C^{n\times n}}$, $AX = XB$ iff $X^{-1}A = BX^{-1}$, and thus the eigenvalues of the similar matrices A and $B = X^{-1}AX$ are the same, since the eigenvalues of A are a subset of the eigenvalues of B and vice versa.

7 Spectral Representation of Real Matrices

Consider complex matrix $A \in \mathbf{C^{n\times n}}$. If $A\mathbf{x_i} = \lambda_\mathbf{i}\mathbf{x_i}$ and $\mathbf{y_j}^H A = \overline{\lambda_j}\mathbf{y_j}^H$, then $\overline{\lambda_j}\mathbf{y_j}^H\mathbf{x_i} = \mathbf{y_j}^\mathbf{H}\mathbf{Ax_i} = \mathbf{y_j}^\mathbf{H}\lambda_\mathbf{i}\mathbf{x_i} = \lambda_\mathbf{i}\mathbf{y_j}^\mathbf{H}\mathbf{x_i}$. If A is real, then $A\overline{\mathbf{x}} = \overline{\lambda}\overline{\mathbf{x}}$ if $A\mathbf{x} = \lambda\mathbf{x}$. The eigenvalues of a real matrix A are either real or occur in complex conjugate pairs. Furthermore, if a pair of left and right eigenmanifolds are either orthogonal or have conjugate corresponding eigenvalues.

For a real $n \times n$ matrix A, there are vectors $\{\psi_i\}_{i=1}{}^n$ and $\{\phi_i{}^H\}_{i=1}{}^n$, which form a biorthogonal basis pair (note that $\phi_i{}^H\psi_j = 0$ if $i \neq j$, $\phi_i{}^H\psi_i = 1$), consisting only of (left or right) eigenvectors where

$$A_{\psi i} = \lambda_i\psi_i, \phi_i{}^H A = \lambda_i\phi_i{}^H.$$

Thus we only need to concern ourselves with a biorthogonal basis for the right eigenmanifold for λ and the left eigenmanifold for $\overline{\lambda}$.

We can use the the information gathered thus to introduce a new means of representing a real matrix with complex eigenvectors and eigenvalues. Take $A \in \mathbf{R^{n\times n}}$. Let $\{\psi_i\}_{i=1}{}^n$ and $\{\phi_i{}^H\}_{i=1}{}^n$ be a biorthogonal basis pair of eigenvectors, where $A_{\psi_i} = \lambda_i\psi_i, \phi_i{}^H A = \lambda_i\phi_i{}^H$.

We then can represent A as $A = \Psi\Lambda\Phi$, where

$$\Psi = \begin{bmatrix} \vdots & \vdots & & \vdots \\ \psi_1 & \psi_2 & \cdots & \psi_n \\ \vdots & \vdots & & \vdots \end{bmatrix}$$

$$\Lambda = \begin{bmatrix} \lambda_1 & 0 & .. & & 0 \\ 0 & \lambda_2 & 0 & ... & 0 \\ 0 & 0 & \ddots & & \\ \vdots & \vdots & & & \\ 0 & 0 & \cdots & & \lambda_n \end{bmatrix}$$

$$\Phi = \begin{bmatrix} \cdots & {\phi_1}^H & \cdots \\ \cdots & {\phi_2}^H & \cdots \\ \cdots & \vdots & \cdots \\ \cdots & {\phi_n}^H & \cdots \end{bmatrix}$$

Since the ψ_i and ϕ_i vectors are biorthogonal, $\Psi\Phi = I$. Additionally, since $A\psi_i = \lambda_i\psi_i$, and the ψ_i form a basis for the whole space, $A = \Psi\Lambda\Phi$. Finally, it follows that $A^m = \Psi\Lambda^m\Phi$. Since Λ has nonzero values only along its diagonals, it is relatively easy to compute Λ^m.