1. December 3, 2013

**Lemma 1.** Let $A$ be a real symmetric $n \times n$ matrix, then the roots of its characteristic polynomial $\det(\lambda I - A)$ are all real.

**Proof.** We consider $A$ as a hermitian linear transformation from $\mathbb{C}^n$ to $\mathbb{C}^n$ with respect to the standard hermitian product on $\mathbb{C}^n$. A root of $\det(\lambda I - A)$ corresponds to an eigenvalue $\lambda_1$. Pick an eigenvector $x \neq 0$ belonging to $\lambda_1$ and $Ax = \lambda_1 x$. However, we know that any eigenvalue of a hermitian linear transformation must be real (use $\langle Ax, x \rangle = \langle x, Ax \rangle$).

\[ \square \]

**Theorem 1.** Let $W$ be an $n$-dimensional real vector space, $\langle \langle \cdot, \cdot \rangle \rangle$ be a real inner product, and $L : W \to W$ be a symmetric linear transformation, i.e. $\langle \langle Lx, y \rangle \rangle = \langle \langle x, Ly \rangle \rangle$. Then $L$ has an orthonormal basis of eigenvectors.

**Proof.** We prove the statement by induction on $n$. $n = 1$ case can be easily verified. Let’s assume the theorem is true for $n - 1$. We take an orthonormal basis $[e_1, \cdots, e_n]$ for $W$ and write $Le_i = \sum_{j=1}^{n} a_{ij}e_j$. Since $L$ is symmetric, we have $a_{ji} = a_{ij}$ or $A = [a_{ij}]$ is a symmetric matrix. From the lemma, any root of the characteristic polynomial of $A$ is real. Take $\lambda_1$ to be a root, then $A - \lambda_1 I$ is singular. Therefore, the linear transformation $L - \lambda_1 I$ is singular, and there exists a $u_1 \in W$ such that $Lu_1 = \lambda_1 u_1$. Now we look at the orthogonal complement $S$ of $u_1$ in $W$ with respect to $\langle \langle \cdot, \cdot \rangle \rangle$. It can be shown that, as in the proof of the spectral theorem, $L$ maps $S$ into itself. We can then apply the induction hypothesis to $S$ which is an $n - 1$ dimensional space.

\[ \square \]

We discuss the real version of the spectral theorem in the following.

**Definition 1.** Let $V$ be a finite dimensional complex vector space with hermitian product $\langle \cdot, \cdot \rangle$ and $[e_1, \cdots, e_n]$ be an orthonormal basis for $V$. 

\[ \text{Date: December 5, 2013.} \]
A vector \( v \) is said to be “real” (with respect to \([e_1, \cdots, e_n]\)) if it is a linear combination of \( e_1, \cdots, e_n \) with real coefficients.

Note that each vector \( v \) in \( V \) can be uniquely written as \( v = \Re v + \sqrt{-1} \Im v \) where both \( \Re v \) and \( \Im v \) are real.

**Theorem 2.** Let \( V \) be a finite dimensional complex vector space with hermitian product \( \langle \cdot, \cdot \rangle \) and \([e_1, \cdots, e_n]\) be an orthonormal basis for \( V \). Let \( T : V \to V \) be a hermitian linear transformation such that \( T(e_i) \) is real for each \( i = 1, \cdots, n \). Then there exists a real orthonormal basis of eigenvectors \( u_i \) of \( T \), i.e. \( \langle u_i, u_j \rangle = \delta_{ij} \) and \( Tu_i = \lambda_i u_i \), \( i = 1, \cdots, n \).

Of course, all \( \lambda_i \)'s are real because \( T \) is hermitian. This is a special case of the spectral theorem: when \( T \) is “real” in the above sense, all the eigenvectors can be chosen to be “real”.

**Proof.** Consider the set \( W \) of all real vectors in \( V \). \( W \) is a real vector space and \([e_1, \cdots, e_n]\) is a basis. \( \langle \cdot, \cdot \rangle \) also defines a real inner product \( \langle \cdot, \cdot \rangle \) on \( W \) (check that it satisfies all axioms). The restriction \( L = T|_W \) of \( T \) to \( W \) is a real linear transformation from \( W \) to \( W \) since each \( T(e_i) \) is real. That \( T \) is hermitian implies \( L \) is symmetric with respect to \( \langle \cdot, \cdot \rangle \) or \( Lx, y \rangle = \langle x, Ly \rangle \). This now follows from the previous theorem. \( \square \)

2. DECEMBER 5, 2013

We shall prove one more diagonalization theorem. Recall a linear transformation \( T : V \to V \) for a finite dimensional complex vector space \( V \) with hermitian product \( \langle \cdot, \cdot \rangle \) is unitary if \( \langle Tx, Ty \rangle = \langle x, y \rangle \).

**Theorem 3.** Any unitary linear transformation is diagonalizable over an orthonormal basis.

**Proof.** Suppose \( T \) is a unitary linear transformation.

Step 1. Any eigenvalue \( \lambda \) of \( T \) satisfies \( |\lambda| = 1 \). Suppose \( Tx = \lambda x \), then \( \langle x, x \rangle = \langle Tx, Tx \rangle = |\lambda|^2 \langle x, x \rangle \).

Step 2. Eigenvectors belonging to distinct eigenvalues of \( T \) are orthogonal. Suppose \( Tx = \lambda x, Ty = \mu y, \langle x, y \rangle = \langle Tx, Ty \rangle = \lambda \mu \langle x, y \rangle \). Suppose \( \langle x, y \rangle \neq 0 \), we have \( 1 = \lambda \mu \) but \( 1 = \lambda \lambda \) and thus \( \mu = \lambda \), a contradiction.

Step 3. Induction on the dimension of \( V \). \( n = 1 \) easily verified. Suppose the theorem is true for \( n - 1 \). Take an eigenvalue \( \lambda_1 \) and an eigenvector \( u_1 \) such that \( Tu_1 = \lambda_1 u_1 \). Again, we look at the orthogonal complement \( S \) of \( u_1 \) and we claim \( T(S) \subset S \) or \( T \) restricts to a linear transformation on \( S \). Suppose \( x \in S \), suffices to prove \( \langle Tx, u_1 \rangle = 0 \) Write \( \langle Tx, u_1 \rangle = \langle Tx, T(u_1/\lambda_1) \rangle = \langle x, u_1/\lambda_1 \rangle = 0 \). \( \square \)
To summarize, the following linear transformations are diagonalizable:

1) Hermitian, skew hermitian, or unitary linear transformation on a finite-dimensional complex linear space with hermitian product. Diagonalizable over orthonormal basis.

2) Symmetric linear transformation on a finite-dimensional real linear space with real inner product. Diagonalizable over orthonormal basis.

3) Linear transformation with distinct eigenvalues.

2.1. Cayley-Hamilton Theorem. Q: Given an $n \times n$ matrix $A$, can we find a polynomial $f$ such that $f(A) = O$, the zero matrix?

Example 1. $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, then $A^2 = O$. We can take the polynomial to be $\lambda^2$.

$B = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$, then $(B - I)^2 = O$. We can take the polynomial to be $\lambda^2 - 2\lambda + 1$.

Consider $A \in \mathcal{M}_{n \times n}$. Since $dim(\mathcal{M}_{n \times n}) = n^2$ and $I, A, A^2, \cdots A^n$ are $n^2 + 1$ vectors in $\mathcal{M}_{n \times n}$, they must be linearly dependent and there exist $c_0, \cdots c_{n^2}$ such that

$$c_0 I + c_1 A + \cdots c_{n^2} A^n = O.$$

In section 7.1 Cayley-Hamilton theorem says that we can do better than this. In fact, we can find a polynomial of degree $n$ that works and it is exactly the characteristic polynomial of $A$.

**Theorem 4.** Let $A$ be an $n \times n$ matrix and $f(\lambda) = \det(\lambda I - A) = \lambda^n + c_{n-1}\lambda^{n-1} + \cdots + c_0$ be its characteristic polynomial. Then $f(A) = O$ in the sense that $A^n + c_{n-1}A^{n-1} + \cdots + c_0 I = O$.

Example 2. $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$. $\det(\lambda I - A) = \lambda^2 - 5\lambda - 2$. By the theorem $A^2 - 5A - 2I = O$. This is very useful in calculating higher power of $A$, for example $A^3 = 5A^2 + 2A = 5(5A + 2I) + 2A = 27A + 10I$.

See Apostol 7.11 for the proof of Cayley-Hamilton. Here is a heuristic proof. First, suppose $A$ is diagonalizable, $A = SAS^{-1}$ for a diagonal matrix $\Lambda$ with diagonal entries $\lambda_1, \cdots \lambda_n$. We compute $A - \lambda_i I = S(\Lambda - \lambda_i I)S^{-1}$ and $(A - \lambda_1 I) \cdots (A - \lambda_n I) = S(\Lambda - \lambda_1 I) \cdots (\Lambda - \lambda_n I)S^{-1}$. It is now easy to check that $(\Lambda - \lambda_1 I) \cdots (\Lambda - \lambda_n I) = O$.

We now turn to the general case. Recall that if $A$ has distinct eigenvalues, then $A$ is diagonalizable. We claim that a generic $n \times n$ matrix has distinct eigenvalues.
Consider the case $n = 3$. In the space of eigenvalues, $\lambda_1 = \lambda_2$, $\lambda_1 = \lambda_3$, and $\lambda_2 = \lambda_3$ represent planes. As long as the eigenvalues of $A$ avoid those planes, $A$ is diagonalizable. Even if $A$ has multiple eigenvalues, say $\lambda_1 = \lambda_2$. We say find a sequence of $n \times n$ matrices $A_i$, away from those planes, such that $A_i \to A$ as $i \to \infty$. Let $f_i$ be the characteristic polynomial of $A_i$ and $f$ be the characteristic polynomial of $A$, we have $f_i(A_i) = 0$. We claim that $f_i \to f$ as $i \to \infty$ as well and $f(A) = \lim_{i \to \infty} f_i(A_i) = O$.

To make this argument rigorous, we need to study the notions of limit, continuity, etc in the space $\mathcal{M}_{n \times n}$, which is isomorphic to $\mathbb{R}^{n^2}$. These are topics of multivariable calculus which will be covered next semester.