Last Time (8/30/12)

**Gram-Schmidt:** - set of linearly independent vectors can yield an orthonormal set that spans the space

**Trace and Determinant:** - definitions and properties

**Eigenvalues and Eigenvectors** - definitions, the characteristic polynomial, and how eigenvalues determine the trace and determinant

**Similarity** - definition, relation to characteristic polynomial, and diagonalizability

1 Further Review

1.2 Warm-up/Correction

1.2.11 **Remark.** Last time, we wrote that, as $\det(AA^{-1}) = \det(A)\det(A^{-1})$, and as $\det(AA^{-1}) = \det(I) = 1$, we have that if $A$ is invertible then $\det(A) \neq 0$. We can also prove the converse using previous insights.

Suppose $\det(A) \neq 0$. Because $\det(A)$ is the product of the eigenvalues of $A$, we know that none of the eigenvalues of $A$ can be zero. However, if $A$ were not invertible, then we would have that $\ker(A) \neq \{0\}$, so for some $x \in \mathbb{C}^n$, $x \neq 0$, we have that $Ax = 0 = 0x$.

Then $x$ is an eigenvector of $A$ corresponding to eigenvalue zero. This is a contradiction, so it must be the case that $A$ is invertible.
1.4 Similarity (cont’d)

1.4.20 Theorem. A matrix $A \in M_n$ is diagonalizable $\iff \exists$ a set of $n$ linearly independent eigenvectors of $A$.

Proof. Given a diagonalizable matrix $A \in M_n$, let $S \in M_n$ be an invertible matrix such that $S^{-1}AS = D$, where $D \in M_n$ is diagonal. Left-multiplying by $S$ gives us $AS = SD$. We write $S = [x_1 \ x_2 \ \ldots \ x_n]$, and note that by properties of left/right multiplication, we get

$$[Ax_1 \ Ax_2 \ \ldots \ Ax_n] = AS = SD = [d_{1,1}x_1 \ d_{2,2}x_2 \ \ldots \ d_{n,n}x_n],$$

where the $d_{k,k}$ are the diagonal entries of $D$. From this we see that each $x_k$ is an eigenvector with corresponding eigenvalue $d_{k,k}$. Then by the invertibility of $S$, we know that it has full rank, so the $x_k$ are linearly independent.

Conversely, suppose that there is a linearly independent set $\{x_1, x_2, \ldots, x_n\}$ of eigenvectors of $A$. Then we can form $S = [x_1 \ x_2 \ \ldots \ x_n]$. Then $AS = [\lambda_1 x_1 \ \lambda_2 x_2 \ \ldots \ \lambda_n x_n]$, where $\lambda_k$ is the eigenvalue corresponding to $x_k$, so if we let $D$ be the diagonal matrix with entries $d_{k,k} = \lambda_k$ for $1 \leq k \leq n$, then $AS = SD$. As $S$ is of full rank, it is invertible. Thus left-multiplying by $S^{-1}$ gives us $S^{-1}AS = D$.

Unfortunately, not every matrix is diagonalizable. For example, the only eigenvalue of the matrix

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

is $\lambda = 0$, and solving the equation

$$A \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

gives us that $x_2 = 1$. Thus the eigenspace corresponding to zero has dimension one, so we can only have a linearly independent set of eigenvectors of size one, which can never form a basis for $\mathbb{C}^2$.

1.4.21 Question. Can we find easier conditions for diagonalizability?

1.4.22 Theorem. If $A \in M_n$, $p_A(t) = \prod_{j=1}^{n} (t - \lambda_j)$, and $\lambda_j \neq \lambda_k$ for all $j \neq k$, then $A$ is diagonalizable.

Proof. We will show that $A$ admits a linearly independent set of $n$ eigenvectors. For each $1 \leq j \leq n$, let $x_j$ be such that $Ax_j = \lambda_j x_j$. Suppose that the $x_j$ are not linearly independent. Then there exists a nontrivial linear combination $\alpha_1 x_{j_1} + \alpha_2 x_{j_2} + \ldots + \alpha_r x_{j_r} = 0$ where $r \leq n$ and $\alpha_j \neq 0$ for all $1 \leq j \leq r$. We can choose a linear combination of the $x_j$ that has the minimal possible value for $r$. Assume, without loss of generality, that $j_1 = 1$, $j_2 = 2$, etc. Otherwise, we can renumber. Then

$$A(\alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_r x_r) = 0 \iff \alpha_1 \lambda_1 x_1 + \alpha_2 \lambda_2 x_2 + \ldots + \alpha_r \lambda_r x_r = 0.$$
If we subtract from this second equation the equation
\[ \lambda_r (\alpha_1 x + \alpha_2 x + \ldots + \alpha_r x) = 0, \]
we obtain
\[ \alpha_1 (\lambda_1 - \lambda_r) x_1 + \alpha_2 (\lambda_2 - \lambda_r) x_2 + \ldots + \alpha_{r-1} (\lambda_{r-1} - \lambda_r) x_1 = 0. \]
However, this contradicts the minimality of \( r \): we have found a nontrivial, vanishing linear combination of the \( x_j \) that involves only \( r - 1 \) vectors. Thus the \( x_j \) are linearly independent. \( \square \)

1.4.23 Remark. Unfortunately, this sufficient condition for diagonalizability excludes such matrices as
\[ A = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}. \]
We need to invoke properties of eigenvectors to fully characterize diagonalizability.

1.4.24 Definition. If for \( A \in M_n \), \( p_A(t) = (t - \lambda_1)^{m_1} (t - \lambda_2)^{m_2} \ldots (t - \lambda_r)^{m_r} \), then we say that \( \lambda_j \) has algebraic multiplicity \( m_j \). We call \( \text{null}(\lambda_j I - A) \) the geometric multiplicity of \( \lambda_j \).

1.4.25 Lemma. If \( A \in M_n \) has eigenvalue \( \lambda \), and \( p_A(t) = (t - \lambda)^m q(t) \) where \( q(\lambda) \neq 0 \), then the geometric multiplicity of \( \lambda \) is less than or equal to \( m \); that is, \( \text{null}(\lambda I - A) \leq m \).

Proof. Let \( E_\lambda = \{ x \in \mathbb{C}^n \mid Ax = \lambda x \} \). We choose a basis for \( E_\lambda \) denoted by \( \{ x_1, x_2, \ldots, x_r \} \). We complement this to a basis of \( \mathbb{C}^n \) given by \( \{ x_1, x_2, \ldots, x_r, x_{r+1}, \ldots, x_n \} \). If we let \( S = [x_1 \ x_2 \ldots \ x_n] \), then
\[ AS = [\lambda x_1 \ \lambda x_2 \ldots \ \lambda x_r \ y_{r+1} \ y_{r+2} \ldots \ y_n], \]
where \( y_j = Ax_j \) for all \( r + 1 \leq j \leq n \). Recalling that \( S^{-1}S = I \), we see that
\[ S^{-1}AS = [\lambda e_1\lambda e_2\ldots\lambda e_r S^{-1}y_{r+1} S^{-1}y_{r+2} \ldots S^{-1}y_n], \]
where \( e_j \) is the \( j \)-th column of the \( r \times r \) identity matrix. In block notation,
\[ S^{-1}AS = \begin{bmatrix} \lambda I & B \\ 0 & C \end{bmatrix}, \]
where \( I \) is the \( r \times r \) identity matrix, \( B \in M_{r,n-r} \), and \( C \in M_{n-r,n-r} \). By properties of the determinant,
\[ p_A(t) = p_{S^{-1}AS}(t) = p_M(t)p_C(t) = (t - \lambda)^r \det(tI - C). \]

We conclude that the algebraic multiplicity of \( \lambda \) is at least \( r \). \( \square \)

1.4.26 Remark. By factorization of \( p_A \), we have \( p_A(t) = \prod_{j=1}^r (t - \lambda_j)^{r_j} \), so \( \sum_{j=1}^r m_j = n \), and \( \sum_{j=1}^r \dim(E_{\lambda_j}) \leq n \). The only way we can have equality is if \( \dim(E_{\lambda_j}) = m_j \) for each \( j \).

1.4.27 Theorem. A matrix \( A \in M_n \) is diagonalizable if and only if the geometric and algebraic multiplicities of each eigenvalue are equal.

Proof. We begin by noting that if \( \lambda_j \neq \lambda_k \), where \( \lambda_j \) and \( \lambda_k \) are eigenvalues of \( A \), then the corresponding eigenspaces intersect trivially; that is, \( E_{\lambda_j} \cap E_{\lambda_k} = \{0\} \). Thus if \( \{v_1, v_2, \ldots, v_{r_1}\} \) and \( \{u_1, u_2, \ldots, u_{r_2}\} \) form bases for \( E_{\lambda_j} \) and \( E_{\lambda_k} \) respectively, then \( \{v_1, v_2, \ldots, v_{r_1}, u_1, u_2, \ldots, u_{r_2}\} \) is linearly independent. (To be continued). \( \square \)