Last Time (8/28/12)

Course info: website - math.uh.edu/~bgb

Matrix multiplication: left (premultiplication) and right (postmultiplication)

Nullity and rank: "dimension counting"

Dot product and orthogonality: inner product, orthogonal projection, least squares property

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

1 Further Review

1.1 Gram-Schmidt (cont’d)

1.1.9 Proposition (Gram-Schmidt). Given a linearly independent set \( \{x_1, x_2, \ldots, x_n\} \) in \( \mathbb{C}^m \), then there exists an orthonormal system \( \{z_1, z_2, \ldots, z_n\} \) such that for each \( j \leq n \),

\[
\text{span}\{x_l : l \leq j\} = \text{span}\{z_l : l \leq j\}.
\]

Proof. Since orthonormality implies linear independence, the dimension of both sides is equal. It is enough to show that we can find inductively for each \( j \leq n \) a vector \( z_j \) which forms an orthonormal system \( \{z_1, z_2, \ldots, z_j\} \) with the preceding ones and

\[
z_j = u_{j,1}x_1 + u_{j,2}x_2 + \cdots + u_{j,j}x_j,
\]

with appropriate coefficients \( u_{j,k} \), so \( z_j \in \text{span}\{x_l : l \leq j\} \). As a consequence, \( z_k \in \text{span}\{x_l : l \leq j\} \) for \( k \leq j \) and thus

\[
\text{span}\{z_l : l \leq j\} \subseteq \text{span}\{x_l : l \leq j\}
\]

but since the dimension on both sides is equal, the two spans must be the same.
The inductive choice of $z_j$ is as follows:

$$z_1 = \frac{x_1}{\|x_1\|}.$$ 

Given an orthonormal system $\{z_1, z_2, \ldots, z_{k-1}\}$ with the same span as $\{x_1, x_2, \ldots, x_{k-1}\}$, then we let

$$y_k = x_k - \langle x_k, z_{k-1} \rangle z_{k-1} - \langle x_k, z_{k-2} \rangle z_{k-2} - \cdots - \langle x_k, z_1 \rangle z_1.$$ 

We have by the assumed orthogonality of $\{z_1, z_2, \ldots, z_{k-1}\}$ that $\langle y_k, z_j \rangle = \langle x_k, z_j \rangle - \langle x_k, z_j \rangle = 0$ for $j < k$. On the other hand $y_k \neq 0$ because $x_k$ is not in the span of the preceding $x_j$’s, which is by induction assumption equal to the span of $z_j$’s. Thus we normalize

$$z_k = \frac{y_k}{\|y_k\|}$$

such that $\{z_1, z_2, \ldots, z_k\}$ is orthonormal and has the same span as $\{x_1, x_2, \ldots, x_{k-1}\}$. 

\[ \square \]

1.2 Trace and Determinant

1.2.10 Definition. For $A \in M_n$, where the entries of $A$ are represented by $A = (a_{i,j})_{i,j=1}^n$, we let

$$tr[A] = \sum_{j=1}^n a_{j,j}$$

and

$$det[A] = \sum_{\sigma \in S_n} sgn(\sigma)a_{1,\sigma(1)}a_{2,\sigma(2)}\cdots a_{n,\sigma(n)}$$

where the sum runs over all $n!$ permutations $\sigma$ of the $n$ items $\{1, 2, \ldots, n\}$, and $sgn(\sigma)$ is 1 if $\sigma$ is an even permutation and $-1$ if it is odd.

1.2.11 Remark.
- If $A = [a_1|a_2|\ldots|a_n]$, then $det(A) = f(a_1, a_2, \ldots, a_n)$ and it is the only function $f$ which is linear in each column vector, alternating in the columns,

$$f(a_1, a_2, \ldots, a_j, \ldots, a_k, \ldots, a_n) = -f(a_1, a_2, \ldots, a_k, \ldots, a_j, \ldots, a_n)$$

and it is normalized, so that $det(I) = 1$.

- If $C = AB$, where $A, B \in M_n$, then

$$det(C) = det(A)det(B)$$

An instance where this becomes useful is where $C$ could be a difficult matrix, whereas the factors $A$ or $B$ might be triangular, unitary, etc. Here, multilinearity is key in getting the factorization property.
• If \( A = \begin{bmatrix} B & C \\ 0 & D \end{bmatrix} \), where \( B \in M_m, D \in M_{n-m} \), then

\[
\det(A) = \det(B)\det(D)
\]

We also have

\[
\det(A) = \sum_{j=1}^{n} (-1)^{i+j}a_{i,j}\det(A_{i,j})
\]

where the matrix \( A_{i,j} \) is obtained by deleting the \( i^{th} \) row and \( j^{th} \) column from matrix \( A \).

• By \( AA^{-1} = I \), we have

\[
\det(AA^{-1}) = \det(I) = 1
\]

Since \( \det(AA^{-1}) = \det(A)\det(A^{-1}) \), we have that \( A \) is invertible \( \Rightarrow \) \( \det(A) \neq 0 \).
We will see further below that the converse is true as well.

### 1.3 Eigenvales and Eigenvectors

#### 1.3.12 Definition. If \( A \in M_n \) and \( \exists \lambda \in \mathbb{C}^m \) and \( x \in \mathbb{C}^n \setminus \{0\} \) such that \( Ax = \lambda x \), then, \( \lambda \) is an eigenvalue of \( A \) and \( x \) is the corresponding eigenvector.

Given any such \( \lambda \), the set \( \{ y : Ay = \lambda y \} \) is called the eigenspace corresponding to \( \lambda \) and the set of eigenvalues is called the spectrum of \( A \).

#### 1.3.13 Remark. \( \lambda \) is an eigenvalue \( \iff \exists x \neq 0, Ax = \lambda x \).

• If we write the last portion as \( (\lambda I - A)x = 0 \), we can see that \( (\lambda I - A) \) is not invertible and therefore, \( \det(\lambda I - A) = 0 \) (Note that the latter relation is also necessary and sufficient).

#### 1.3.14 Definition. For \( A \in M_n \), the characteristic polynomial is defined by:

\[
P_A(t) = \det(tI - A)
\]

#### 1.3.15 Remark. With the insight from above, \( \lambda \) is an eigenvalue \( \iff P_A(\lambda) = 0 \)

• Over \( \mathbb{C} \), \( P_A \) can be factorized so that,

\[
P_A(t) = (t - \lambda_1)(t - \lambda_2) \ldots (t - \lambda_n)
\]

where some \( \lambda_j \)'s may be repeated and \( 1 \leq j \leq n \).

Comparing the definition of \( P_A \) with the factorized form gives,

\[
P_A(t) = t^n + (-1)\sum_{j=1}^{n} \lambda_j t^{n-1} + \cdots + (-1)^n \prod_{j=1}^{n} \lambda_j
\]
From the definition of the determinant, we see that
\[
det(tI - A) = t^n - tr[A]t^{n-1} + \cdots + (-1)^n \det(A)
\]
By comparing terms in the polynomial, we also notice that
\[
tr[A] = \sum_{j=1}^{n} \lambda_j
\]
and
\[
det[A] = \prod_{j=1}^{n} \lambda_j
\]

1.4 Similarity

1.4.16 Definition. A matrix \( B \in M_n \) is similar to \( A \in M_n \) if \( \exists S \in M_n \), where \( S \) is invertible, such that:
\[
B = S^{-1}AS
\]
This defines an equivalence relation, as can be verified.

1.4.17 Theorem. If \( A, B \in M_n \) are similar, then \( P_A = P_B \)

Proof.

\[
P_B(t) = \det(tI - B) = \det(t(S^{-1}S) - S^{-1}AS) = \det(S^{-1}(tI - A)S) = \det(S^{-1})\det(tI - A)\det(S) = \det(S^{-1}S)\det(tI - A) = \det(tI - A) = P_A(t)
\]

We conclude \( A, B \) share same eigenvalues with the same multiplicity.

1.4.18 Remark. Note: The characteristic polynomial is not characteristic up to similarity because,
\[
A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \text{ and } B = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}
\]
are not similar (they do not have the same rank), but \( P_A(t) = t^2 \) and \( P_B(t) = t^2 \).

1.4.19 Definition. A matrix \( A \in M_n \) is diagonalizable if it is similar to a diagonal matrix.

1.4.20 Theorem. A matrix \( A \in M_n \) is diagonalizable \( \iff \exists \) a set of \( n \) linearly independent eigenvectors of \( A \). (This will be proved during the next class).