Bess Berg
March 8, 2010

Linear Operators: some basics

These notes were thrown together somewhat quickly so they will surely contain typos and perhaps even some more fundamental mistakes, though hopefully not too many of the latter!

1 Definitions

1. The *transpose* of the $m \times n$ matrix A is denoted $A^\top$ and is the $n \times m$ matrix whose rows are the columns of A (and thus whose columns are the rows of A).

2. The *complex conjugate* of $z = a + bi \in \mathbb{C}$ is $z^* = a - bi$.

3. The *adjoint* of the $m \times n$ matrix A is denoted $A^\dagger$ and is the transpose of the complex conjugate of A, $A^\dagger = A^{*\top}$.

4. An *eigenvalue* with corresponding *eigenvector* of the (square) matrix $A \in \mathbb{C}^{n\times n}$ is any $\lambda \in \mathbb{C}$ with non-zero $x \in \mathbb{C}^n$ that satisfy the equation $Ax = \lambda x$.

5. The *characteristic polynomial* of a matrix $A \in \mathbb{C}^{n\times n}$ is the monic (leading coefficient is 1), degree-n polynomial in λ given by $det(A-\lambda I)$. Note that setting this polynomial to zero and solving for λ gives values of λ such that the matrix $A-\lambda I$ is not invertible. For each such distinct root λ the equation $(A - \lambda I)x = 0$, and thus $Ax = \lambda x$, has at least one solution x.

6. A *unitary matrix* is a square matrix with complex entries, $A \in \mathbb{C}^{n\times n}$, whose columns (and thus rows) are orthogonal unit vectors. That is, the columns of A form an orthonormal basis for $C^{n\times n}$. Note A is unitary if and only if $A^\dagger = A^{-1}$.

7. An *orthogonal matrix* is a square matrix with real entries, $A \in \mathbb{R}^{n\times n}$, whose columns (and thus rows) are orthogonal unit vectors. Note that an orthogonal matrix is the special case of a unitary matrix in which all entries are real. We have $A^\top = A^{-1}$, and the columns of A form a (real) orthonormal basis of $\mathbb{R}^{n\times n}$ (and $\mathbb{C}^{n\times n}$).

8. A *diagonalizable matrix* is a square matrix $A \in \mathbb{C}^{n\times n}$ for which there exist matrices $P, D \in \mathbb{C}^{n\times n}$ such that P is invertible, D is diagonal, and $A = PDP^{-1}$. Note that columns of P are necessarily eigenvectors of A with corresponding eigenvalues given by D.

9. A *normal matrix* is a square matrix $A \in \mathbb{C}^{n\times n}$ such that A commutes with its adjoint: $A^\dagger A = AA^\dagger$. A matrix is normal if and only if it can be diagonalized by a unitary matrix (this is known as the Spectral Theorem); equivalently, a matrix is normal if and only if its eigenvectors form a full orthogonal basis.

10. A *Hermitian* matrix is a square matrix that is self-adjoint: $A^\dagger = A$. Note that Hermitian matrices are special cases of normal matrices. In particular, a matrix is Hermitian if and only if it is normal and has all real eigenvalues.

11. A *symmetric matrix* is a square matrix with real entries, $A \in \mathbb{R}^{n\times n}$ such that $A^\top = A$. Note that A is a real-valued special case of a Hermitian (self-adjoint) matrix. Furthermore, not only does A have real eigenvalues, but one can show that the eigenvectors of A will be real too, i.e. A can be diagonalized by an orthogonal matrix.

2 Facts

Fact 1. *Eigenvectors with distinct eigenvalues are linearly independent though not necessarily orthogonal.*

Proof. The first part should be somewhat obvious, but in case not: Assume v and w are eigenvectors of A with distinct eigenvalues $\lambda \neq \gamma$ respectively. If v and w *are* linearly dependent, then there is some $\alpha \in \mathbb{C}$ such that $v = \alpha w$. Note that neither v nor w is the zero-vector, so $\alpha \neq 0$. We have $Av = \lambda v = \lambda\alpha w$; but we also have $Av = A(\alpha w) = \gamma\alpha w$, implying $\lambda\alpha w = \gamma\alpha w$ and thus $\lambda = \gamma$, a contradiction. So v and w must be linearly independent. To see that a matrix A can have non-orthogonal eigenvectors consider the matrix

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}$$

which has eigenvectors $(1,0)^\top$ (with corresponding eigenvalue 1) and (unnormalized) $(1,1)^\top$ (with corresponding eigenvalue 2).

Fact 2. *Every square $n \times n$ matrix has at least one eigenvector.*

Proof. Let $A \in \mathbb{C}^{n\times n}$. The characteristic polynomial of A is given by $det(A - \lambda I) = 0$. This polynomial is monic (leading coefficient is 1) and has degree n. By the Fundamental Theorem of Algebra it has n complex roots (counted with multiplicities). Therefore there is at least one $\lambda \in \mathbb{C}$ such that $A - \lambda I$ is not an invertible matrix, i.e. its columns are linearly dependent. So there is some linear combination of the columns equal to 0; that is, there is an $x \in \mathbb{C}^n$ such that $(A - \lambda I)x = 0$ and thus $Ax = \lambda x$. This x is an eigenvector of A, with eigenvalue λ.

Fact 3. *The geometric multiplicity of an eigenvalue is at least* 1 *but may be less than the algebraic multiplicity of that eigenvalue.*

Proof. First some definitions. The *algebraic multiplicity* of an eigenvalue λ for matrix $A \in \mathbb{C}^n$ is the multiplicity of λ as a root of the characteristic polynomial of A. The *geometric multiplicity* of λ is the dimension of its corresponding *eigenspace*, which is the span of all eigenvectors of A having eigenvalue λ. We know that there is at least one eigenvector corresponding to λ, and so the dimension of the corresponding eigenspace is at least 1. To see that the geometric multiplicity may be less than the algebraic multiplicity, consider the matrix

$$A = \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix}$$

for any $k \neq 0$. A has characteristic polynomial $(1 - \lambda)^2$. The root $\lambda = 1$ has algebraic multiplicity 2, but A only has one corresponding eigenvector, $(1, 0)^\top$.

Fact 4. Let $A \in \mathbb{C}^{n\times n}$. The following are equivalent:

1. A is diagonalizable.
2. The sum of the geometric multiplicities of the eigenvalues of A is n.
3. A has n linearly independent eigenvectors.

Fact 5. *Every matrix $A \in \mathbb{C}^{m\times n}$ has a singular value decomposition $A = UDV^{-1}$ with unitary U and V and diagonal D; only some square matrices $A \in \mathbb{C}^{n\times n}$ may be diagonalized $A = PDP^{-1}$ for invertible P and diagonal D, and only some of these A may be diagonalized by unitary matrices.*