Notes on Eigenvalues and Eigenvectors

Robert A. van de Geijn
Department of Computer Science
The University of Texas
Austin, TX 78712
rvdg@cs.utexas.edu

October 31, 2014

If you have forgotten how to find the eigenvalues and eigenvectors of 2×2 and 3×3 matrices, you may want to review Linear Algebra: Foundations to Frontiers - Notes to LAFF With.

1 Definition

Definition 1. *Let $A \in \mathbb{C}^{m \times m}$. Then $\lambda \in \mathbb{C}$ and nonzero $x \in \mathbb{C}^m$ are said to be an eigenvalue and corresponding eigenvector if $Ax = \lambda x$. The tuple (λ, x) is said to be an eigenpair. The set of all eigenvalues of A is denoted by $\Lambda(A)$ and is called the spectrum of A.*

The action of A on an eigenvector x is as if it were multiplied by a scalar. The direction does not change, only its length is scaled:

$$Ax = \lambda x.$$

Theorem 2. *Scalar λ is an eigenvalue of A if and only if*

$$(\lambda I - A) \left\{ \begin{array}{l} \textit{is singular} \\ \textit{has a nontrivial null-space} \\ \textit{has linearly dependendent columns} \\ \det(\lambda I - A) = 0 \\ (\lambda I - A)x = 0 \textit{ has a nontrivial solution} \\ \textit{etc.} \end{array} \right.$$

The following exercises expose some other basic properties of eigenvalues and eigenvectors:

Exercise 3. *Eigenvectors are not unique.*

Exercise 4. *Let λ be an eigenvalue of A and let $\mathcal{E}_\lambda(A) = \{x \in \mathbb{C}^m | Ax = \lambda x\}$ denote the set of all eigenvectors of A associated with λ (including the zero vector, which is not really considered an eigenvector). Show that this set is a (nontrivial) subspace of $\mathbb{C}^m$.*

Definition 5. *Given $A \in \mathbb{C}^{m \times m}$, the function $p_m(\lambda) = \det(\lambda I - A)$ is a polynomial of degree at most m. This polynomial is called the characteristic polynomial of A.*

The definition of $p_m(\lambda)$ and the fact that is a polynomial of degree at most m is a consequence of the definition of the determinant of an arbitrary square matrix. This definition is not particularly enlightening other than that it allows one to succinctly related eigenvalues to the roots of the characteristic polynomial.

Remark 6. *The relation between eigenvalues and the roots of the characteristic polynomial yield a disconcerting insight: A general formula for the eigenvalues of a $m \times m$ matrix with $m > 4$ does not exist.*

The reason is that there is no general formula for the roots of a polynomial of degree $m > 4$. Given any polynomial $p_m(\chi)$ of degree m, an $m \times m$ matrix can be constructed such that its characteristic polynomial is $p_m(\lambda)$. If

$$p_m(\chi) = \alpha_0 + \alpha_1 \chi + \cdots + \alpha_{m-1}\chi^{m-1} + \chi^m$$

and

$$A = \begin{pmatrix} -\alpha_{n-1} & -\alpha_{n-2} & -\alpha_{n-3} & \cdots & -\alpha_1 & -\alpha_0 \\ 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \end{pmatrix}$$

then

$$p_m(\lambda) = \det(\lambda I - A)$$

Hence, we conclude that no general formula can be found for the eigenvalues for $m \times m$ matrices when $m > 4$. What we will see in future "Notes on ..." is that we will instead create algorithms that *converge* to the eigenvalues and/or eigenvalues of matrices.

Theorem 7. *Let $A \in \mathbb{C}^{m \times m}$ and $p_m(\lambda)$ be its characteristic polynomial. Then $\lambda \in \Lambda(A)$ if and only if $p_m(\lambda) = 0$.*

Proof: This is an immediate consequence of Theorem 2. □

In other words, λ is an eigenvalue of A if and only if it is a root of $p_m(\lambda)$. This has the immediate consequence that A has at most m eigenvalues and, if one counts multiple roots by their multiplicity, it has exactly m eigenvalues. (One says "Matrix $A \in \mathbb{C}^{m \times m}$ has m eigenvalues, multiplicity counted.)

Exercise 8. *The eigenvalues of a diagonal matrix equal the values on its diagonal. The eigenvalues of a triangular matrix equal the values on its diagonal.*

Corollary 9. *If $A \in \mathbb{R}^{m \times m}$ is real valued then some or all of its eigenvalues may be complex valued. In this case, if $\lambda \in \Lambda(A)$ then so is its conjugate, $\bar{\lambda}$.*

Proof: It can be shown that if A is real valued, then the coefficients of its characteristic polynomial are all real valued. Complex roots of a polynomial with real coefficients come in conjugate pairs. □

It is not hard to see that an eigenvalue that is a root of multiplicity k has at most k eigenvectors. It is, however, not necessarily the case that an eigenvalue that is a root of multiplicity k also has k linearly independent eigenvectors. In other words, the null space of $\lambda I - A$ may have dimension less then the algebraic multiplicity of λ. The prototypical counter example is the $k \times k$ matrix

$$J(\mu) = \begin{pmatrix} \mu & 1 & 0 & \cdots & 0 & 0 \\ 0 & \mu & 1 & \ddots & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ddots & \mu & 1 \\ 0 & 0 & 0 & \cdots & 0 & \mu \end{pmatrix}$$

where $k > 1$. Observe that $\lambda I - J(\mu)$ is singular if and only if $\lambda = \mu$. Since $\mu I - J(\mu)$ has $k - 1$ linearly independent columns its null-space has dimension one: all eigenvectors are scalar multiples of each other. This matrix is known as a *Jordan block*.

Definition 10. *A matrix $A \in \mathbb{C}^{m\times m}$ that has fewer than m linearly independent eigenvectors is said to be defective. A matrix that does have m linearly independent eigenvectors is said to be nondefective.*

Theorem 11. *Let $A \in \mathbb{C}^{m\times m}$. There exist nonsingular matrix X and diagonal matrix Λ such that $A = X\Lambda X^{-1}$ if and only if A is nondefective.*

Proof:
($\Rightarrow$). Assume there exist nonsingular matrix X and diagonal matrix Λ so that $A = X\Lambda X^{-1}$. Then, equivalently, $AX = X\Lambda$. Partition X by columns so that

$$\begin{aligned} A\left(\begin{array}{c|c|c|c} x_0 & x_1 & \cdots & x_{m-1} \end{array} \right) &= \left(\begin{array}{c|c|c|c} x_0 & x_1 & \cdots & x_{m-1} \end{array} \right) \left(\begin{array}{c|c|c|c} \lambda_0 & 0 & \cdots & 0 \\ \hline 0 & \lambda_1 & \cdots & 0 \\ \hline \vdots & \ddots & \ddots & \vdots \\ \hline 0 & 0 & \cdots & \lambda_{m-1} \end{array} \right) \\ &= \left(\begin{array}{c|c|c|c} \lambda_0 x_0 & \lambda_1 x_1 & \cdots & \lambda_{m-1} x_{m-1} \end{array} \right). \end{aligned}$$

Then, clearly, $Ax_j = \lambda_j x_j$ so that A has m linearly independent eigenvectors and is thus nondefective.
($\Leftarrow$). Assume that A is nondefective. Let $\{x_0, \cdots, x_{m-1}\}$ equal m linearly independent eigenvectors corresponding to eigenvalues $\{\lambda_0, \cdots, \lambda_{m-1}\}$. If $X = \left(\begin{array}{c|c|c|c} x_0 & x_1 & \cdots & x_{m-1} \end{array} \right)$ then $AX = X\Lambda$ where $\Lambda = \text{diag}(\lambda_0, \ldots, \lambda_{m-1})$. Hence $A = X\Lambda X^{-1}$. $\square$

Definition 12. *Let $\mu \in \Lambda(A)$ and $p_m(\lambda)$ be the characteristic polynomial of A. Then the algebraic multiplicity of μ is defined as the multiplicity of μ as a root of $p_m(\lambda)$.*

Definition 13. *Let $\mu \in \Lambda(A)$. Then the geometric multiplicity of μ is defined to be the dimension of $\mathcal{E}_\mu(A)$. In other words, the geometric multiplicity of μ equals the number of linearly independent eigenvectors that are associated with μ.*

Theorem 14. *Let $A \in \mathbb{C}^{m\times m}$. Let the eigenvalues of A be given by $\lambda_0, \lambda_1, \cdots, \lambda_{k-1}$, where an eigenvalue is listed exactly n times if it has geometric multiplicity n. There exists a nonsingular matrix X such that*

$$A = X \left(\begin{array}{c|c|c|c} J(\lambda_0) & 0 & \cdots & 0 \\ \hline 0 & J(\lambda_1) & \cdots & 0 \\ \hline \vdots & \vdots & \ddots & \vdots \\ \hline 0 & 0 & \cdots & J(\lambda_{k-1}) \end{array} \right).$$

For our discussion, the sizes of the Jordan blocks $J(\lambda_i)$ are not particularly important. Indeed, this decomposition, known as the Jordan Canonical Form of matrix A, is not particularly interesting in practice. For this reason, we don't discuss it further and do not we give its proof.

2 The Schur and Spectral Factorizations

Theorem 15. *Let $A, Y, B \in \mathbb{C}^{m \times m}$, assume Y is nonsingular, and let $B = Y^{-1}AY$. Then $\Lambda(A) = \Lambda(B)$.*

Proof: Let $\lambda \in \Lambda(A)$ and x be an associated eigenvector. Then $Ax = \lambda x$ if and only if $Y^{-1}AYY^{-1}x = Y^{-1}\lambda x$ if and only if $B(Y^{-1}x) = \lambda(Y^{-1}x)$. $\square$

Definition 16. *Matrices A and B are said to be similar if there exists a nonsingular matrix Y such that $B = Y^{-1}AY$.*

Given a nonsingular matrix Y the transformation $Y^{-1}AY$ is called a similarity transformation of A.

It is not hard to expand the last proof to show that if A is similar to B and $\lambda \in \Lambda(A)$ has algebraic/geometric multiplicity k then $\lambda \in \Lambda(B)$ has algebraic/geometric multiplicity k.

The following is the fundamental theorem for the algebraic eigenvalue problem:

Theorem 17. Schur Decomposition Theorem *Let $A \in \mathbb{C}^{m \times m}$. Then there exist a unitary matrix Q and upper triangular matrix U such that $A = QUQ^H$. This decomposition is called the Schur decomposition of matrix A.*

In the above theorem, $\Lambda(A) = \Lambda(U)$ and hence the eigenvalues of A can be found on the diagonal of U.

Proof: We will outline how to construct Q so that $Q^H AQ = U$, an upper triangular matrix.

Since a polynomial of degree m has at least one root, matrix A has at least one eigenvalue, λ_1, and corresponding eigenvector q_1, where we normalize this eigenvector to have length one. Thus $Aq_1 = \lambda_1 q_1$. Choose Q_2 so that $Q = \left(\begin{array}{c|c} q_1 & Q_1 \end{array} \right)$ is unitary. Then

$$\begin{aligned} Q^H AQ &= \left(\begin{array}{c|c} q_1 & Q_2 \end{array} \right)^H A \left(\begin{array}{c|c} q_1 & Q_2 \end{array} \right) \\ &= \left(\begin{array}{c|c} q_1^H A q_1 & q_1^H A Q_2 \\ \hline Q_2^H A q_1 & Q_2^H A Q_2 \end{array} \right) = \left(\begin{array}{c|c} \lambda_1 & q_1^H A Q_2 \\ \hline \lambda Q_2^H q_1 & Q_2^H A Q_2 \end{array} \right) = \left(\begin{array}{c|c} \lambda_1 & w^T \\ \hline 0 & B \end{array} \right), \end{aligned}$$

where $w^T = q_1^H AQ_2$ and $B = Q_2^H AQ_2$. This insight can be used to construct an inductive proof. $\square$

One should not mistake the above theorem and its proof as a constructive way to compute the Schur decomposition: finding an eigenvalue and/or the eigenvalue associated with it is difficult.

Lemma 18. *Let $A \in \mathbb{C}^{m \times m}$ be of form $A = \left(\begin{array}{c|c} A_{TL} & A_{TR} \\ \hline 0 & A_{BR} \end{array} \right)$. Assume that Q_{TL} and Q_{BR} are unitary "of appropriate size". Show that*

$$A = \left(\begin{array}{c|c} Q_{TL} & 0 \\ \hline 0 & Q_{BR} \end{array} \right)^H \left(\begin{array}{c|c} Q_{TL} A_{TL} Q_{TL}^H & Q_{TL} A_{TR} Q_{BR}^H \\ \hline 0 & Q_{BR} A_{BR} Q_{BR}^H \end{array} \right) \left(\begin{array}{c|c} Q_{TL} & 0 \\ \hline 0 & Q_{BR} \end{array} \right).$$

Exercise 19. *Prove Lemma 18. Then generalize it to a result for block upper triangular matrices:*

$$A = \left(\begin{array}{c|c|c|c} A_{0,0} & A_{0,1} & \cdots & A_{0,N-1} \\ \hline 0 & A_{1,1} & \cdots & A_{1,N-1} \\ \hline 0 & 0 & \ddots & \vdots \\ \hline 0 & 0 & \cdots & A_{N-1,N-1} \end{array} \right).$$

Corollary 20. *Let $A \in \mathbb{C}^{m\times m}$ be of for* $A = \left(\begin{array}{c|c} A_{TL} & A_{TR} \\ \hline 0 & A_{BR} \end{array}\right)$. *Then* $\Lambda(A) = \Lambda(A_{TL}) \cup \Lambda(A_{BR})$.

Exercise 21. *Prove Corollary 20. Then generalize it to a result for block upper triangular matrices.*

A theorem that will later allow the eigenvalues and vectors of a real matrix to be computed (mostly) without requiring complex arithmetic is given by

Theorem 22. *Let $A \in \mathbb{R}^{m\times m}$. Then there exist a unitary matrix $Q \in \mathbb{R}^{m\times m}$ and* quasi *upper triangular matrix $U \in \mathbb{R}^{m\times m}$ such that $A = QUQ^T$.*

A quasi upper triangular matrix is a block upper triangular matrix where the blocks on the diagonal are 1×1 or 2×2. Complex eigenvalues of A are found as the complex eigenvalues of those 2×2 blocks on the diagonal.

Theorem 23. Spectral Decomposition Theorem *Let $A \in \mathbb{C}^{m\times m}$ be Hermitian. Then there exist a unitary matrix Q and diagonal matrix $\Lambda \in \mathbb{R}^{m\times m}$ such that $A = Q\Lambda Q^H$. This decomposition is called the Spectral decomposition of matrix A.*

Proof: From the Schur Decomposition Theorem we know that there exist a matrix Q and upper triangular matrix U such that $A = QUQ^H$. Since $A = A^H$ we know that $QUQ^H = QU^HQ^H$ and hence $U = U^H$. But a Hermitian triangular matrix is diagonal with real valued diagonal entries. □

What we conclude is that a Hermitian matrix is nondefective and its eigenvectors can be chosen to form an orthogonal basis.

Exercise 24. *Let A be Hermitian and λ and μ be distinct eigenvalues with eigenvectors x_λ and x_μ, respectively. Then $x_\lambda^H x_\mu = 0$. (In other words, the eigenvectors of a Hermitian matrix corresponding to distinct eigenvalues are orthogonal.)*

3 Relation Between the SVD and the Spectral Decomposition

Exercise 25. *Let $A \in \mathbb{C}^{m\times m}$ be a Hermitian matrix, $A = Q\Lambda Q^H$ its Spectral Decomposition, and $A = U\Sigma V^H$ its SVD. Relate Q, U, V, Λ, and Σ.*

Exercise 26. *Let $A \in \mathbb{C}^{m\times m}$ and $A = U\Sigma V^H$ its SVD. Relate the Spectral decompositions of A^HA and AA^H to U, V, and Σ.*