Topic: Eigenvalues

Nick Battista

Date Created: 6/25/2014
Date Modified: 7/27/16

1 REMEMBER EIGENVALUES?

Remember how much fun calculating eigenvalues using paper and pencil was in undergraduate Linear Algebra? There was computing determinants of matrices symbollically, forming a characteristic polynomial only to find it roots, and solving some singular systems of equations. Everything that should translate swimmingly to numerically finding eigenvalues, right? Literally, none of those things are exceedingly computer friendly.

The process of finding eigenvalues on a computer is going to have a very different feel and approach, than how one finds them by hand. Before we dive into the abstract details, we will first flat out state that numerically computing eigenvalues tends to be a messy business. Unlike solving systems of equations (which also have a very much more natural transition from pencil and paper to the *in silica* implementations), we will not be able to state *a priori* the computational cost of finding eigenvalues.

We will not be able to say the explicit cost of finding eigenvalues; however, will only be able to approximate the number of operations required. This is all to say, *numerical eigenvalue solvers need must be iterative*. We can easily observe this since at the heart of each eigenvalue problem lies a characteristic polynomial, to which we must find its roots, and finding roots of equations is an iterative process. Whatever our method, it will require an iterative flavor, and hence will introduce errors, outside of round-off errors.

Before we dive right into the numerical side of things, we will first refresh a bunch of eigenvalue identities, i.e., factorizations, properties, among other things, which we will put into our numerical eigenvalue toolbox.

2 EIGENVALUE TOOLBOX

This section will more or less be a list of eigenvalue identities, definitions, factorizations, and everything that is good…in the eigen-world. Hopefully this looks familiar, if not, please indulge me, and peruse this section with a smile.

2.1 SOME EIGEN-RELATED DEFINITIONS TO KEEP AROUND FOR A RAINY DAY…

Here it goes…let's go out and define!

1. λ is an eigenvalue $\leftrightarrow$ there exists $\mathbf{x}$ such that $\lambda\mathbf{x} - A\mathbf{x} = 0 \leftrightarrow \lambda I - A$ is *singular* $\leftrightarrow \det(\lambda I - A) = 0$.

2. *Definition:* The *geometric multiplicity* of an eigenvalue, λ, is the dimension of the nullspace of $\lambda I - A$, i.e., $(\lambda I - A)$. In short, it is the number of eigenvectors associated with one particular eigenvalue, λ.

3. *Definition:* The *algebraic multiplicity* of an eigenvalue, λ, of a matrix A, is the the multiplicity of λ as a root of the characteristic polynomial, e.g., $p_A(\lambda) = \det(\lambda I - A) = 0$ Furthermore, note that the above two definitions imply that

$$\text{algebraic mult.} \geq \text{geometric mult.}\ .$$

4. *Definition:* A matrix, A, is said to be *defective* when the algebraic multiplicity is strictly greater than the geometric multiplicity, that is, not enough linearly independent eigenvectors exist for a particular eigenvalue. *Spoiler*, we will not be able to factorize A into an eigenvalue decomposition, if A is defective. On that note, I'll also drop here that all eigenvectors are linearly independent, so if A is full rank, e.g., $(A) = \{\}$, then the eigenspace of A spans the same space as A. Stronger than that, though, we can say that the eigen-space of A spans the same space as A.

5. *Definition:* The determinant of a matrix A can be written as a product of its eigenvalues, e.g.,

$$\det(A) = \prod_{k=1}^{n} \lambda_k.$$

6. *Definition:* Similarly, the *trace* of a matrix A can be written as a sum of its eigenvalues, e.g.,

$$\operatorname{tr}(A) = \sum_{k=1}^{n} \lambda_k.$$

   We note the above two definitions may be useful in some applications, but we will not harness them, often, if at all in everyday life. It's best to keep them around in the toolbox, though.

2.2 Eigenvalue Decompositions: the big guns

If the above subsection was nails, screws, and bolts of eigenvalue problems, then this section we learn about the power tools. Matrix factorizations are very useful, as demonstrated when solving linear systems. We will restrict our discussion here to solely eigenvalue-type factorizations, e.g., eigenvalue decompositions. We will mention three different eigenvalue decompositions.

1. *Diagonalization*: We can write A as a product of three matrices,

$$A = X\Lambda X^{-1}.$$

   if and only A is *non-defective*, meaning all of A's algebraic multiplicities = geometric multiplicities. The matrix X is composed of eigenvectors in their column form. The matrix Λ is a diagonal matrix, where it's diagonal is composed of the eigenvalues of A. We note if $\Lambda_{11} = \lambda_1$, then the first column of X is the eigenvector associated with λ_1, and so on.

   You are probably familiar with this eigenvalue decomposition from undergraduate Linear Algebra and for this, I apologize.

2. If $A \in \mathbb{R}^{n\times n}$ and has n orthogonal eigenvectors, then A has a *unitary diagonalization*, and we can write

$$A = Q\Lambda Q^*.$$

   Similarly to matrix X above, Q is a matrix, whose columns are all the eigenvectors of A; however, in this case they are orthonormal, hence we have $QQ^* = Q^*Q = I$. Furthermore, the matrix Λ is a diagonal matrix, which contains the eigenvalues on the diagonal.

   We will note a few things.

   - This counts as both an eigenvalue decomposition and SVD, i.e.,

$$A = Q\Lambda Q^* = U\sum V^{-1},$$

     where we let $U = Q$, $\sum = \Lambda$, and $V^{-1} = \text{sign}(\Lambda)Q^*$.
   - Hermitian matrices, i.e., $A = A^*$, are unitary diagonalizable; e.g., all of their eigenvectors are orthogonal (orthonormal) and eigenvalues are real-valued! Let's prove these claims!

     Proof. We will now prove that all eigenvalues are real and that their associated eigenvectors are orthogonal.

a) First we will prove that all the eigenvalues of a Hermitian matrix are real. To do this we will recall the infamous eigenvalue equation,

$$A\mathbf{x} = \lambda\mathbf{x}.$$

To isolate λ by itself we multiply both sides by $\mathbf{x}^*$, and solve for λ, e.g.,

$$\lambda = \frac{\mathbf{x}^* A\mathbf{x}}{\mathbf{x}^*\mathbf{x}}.$$

Now we are ready for the final step, ,

$$\lambda^* = \left(\frac{\mathbf{x}^* A\mathbf{x}}{\mathbf{x}^*\mathbf{x}}\right)^* = \frac{\mathbf{x}^* A^*\mathbf{x}}{\mathbf{x}^*\mathbf{x}} = \frac{\mathbf{x}^* A\mathbf{x}}{\mathbf{x}^*\mathbf{x}} = \lambda,$$

and hence $\lambda \in \mathbb{R}$.

b) We now will prove all the eigenvectors are orthogonal. Assume thatλ and μ are distinct eigenvalues associated eigenvectors $\mathbf{x}$ and $\mathbf{y}$, respectively. Hence we will now show that $<\mathbf{x},\mathbf{y}>=0$

To begin we consider,

$$\lambda <\mathbf{x},\mathbf{y}>=<\lambda\mathbf{x},\mathbf{y}>=<A\mathbf{x},\mathbf{y}>.$$

Now using properties of inner products, we can write

$$<A\mathbf{x},\mathbf{y}>=<\mathbf{x},A^*\mathbf{y}>.$$

However, since $A = A^*$, we get

$$<\mathbf{x},A^*\mathbf{y}>=<\mathbf{x},A\mathbf{y}>=<\mathbf{x},\mu\mathbf{y}>$$

Using another property of inner products, we can pull out a scalar by

$$<\mathbf{x},\mu\mathbf{y}>=\mu^* <\mathbf{x},\mathbf{y}>=\mu <\mathbf{x},\mathbf{y}>.$$

Hence we have

$$\lambda <\mathbf{x},\mathbf{y}>=\mu <\mathbf{x},\mathbf{y}>,$$

but since we assumed $\mu \neq \lambda$, this implies that $<\mathbf{x},\mathbf{y}>=0$.

I have to admit this is a more long winded way to show this property of Hermitian matrices. Since we are assuming A is full rank, we can diagonalize A, e.g., let

$$A = X\Lambda X^{-1},$$

and now since $A = A^*$, we have

$$A^* = (X\Lambda X^{-1})^* = (X^{-1})^*\Lambda^* X^* = (X\Lambda X^{-1})^* = (X^{-1})^*\Lambda X^*,$$

since all the eigenvalues are real. However this must be equal to the original diagonalization of A because A is Hermitian. Therefore we have that

$$X = (X^{-1})^* \quad \text{and} \quad X^{-1} = X^*.$$

The only way this above statement could be true is if $X^* = X^{-1}$, which makes X an orthogonal matrix, and our result falls out.

□

In fact, we can extend this to all matrices that are normal!

Theorem 1. *A matrix is unitary diagonalizable if and only if it is normal.*

Recall that a matrix, A, is *normal* if $AA^* = A^*A$.

3. Regardless of the spectral properties of A, we can write it in its Schur form. The Schur form of a matrix is where A is decomposed as,

$$A = QTQ^*.$$

where Q is an orthogonal matrix and T is an upper-triangular matrix. Please note that Q is not, in general, composed of eigenvectors, as in the unitary diagonalization case. Moreover, upper-triangular matrix T is in the *Jordan form*, but still maintains all the eigenvalues of A along its main diagonal. This is because A and T are *similar*. We now define what that means.

- *Definition*: Two matrices, A and B, are similar if and only if there exists non-singular matrix P, such that we can write $A = PBP^{-1}$. Furthermore if two matrices are similar, then they have the same eigenvalues. A clear example of this is the diagonalization of a matrix.

Moreover, because of the *Jordan form* of T, every matrix has a Schur Factorization. That's important enough to write as a theorem, give me a second.

Theorem 2. *Every square matrix has a Schur Factorization.*

Boom. Why is this so important? Well as we approach numerically finding eigenvalues, we will almost exclusively use the Schur Factorization. The only time we won't, is typically when A is Hermitian, and we see the Schur Factorization becomes a Unitary Diagonalization of A, that is, T is diagonal. Also, whenever we have orthogonal matrices involved, computations tend to have nice stability properties, from the backward stability of how one will compute Q, i.e., using Householder Reflectors or Givens Rotations.