Week 9: Wednesday, Oct 21

Logistics

Midterm out today (also up on web page) — return your solutions in class on Monday, October 26. Good luck!

Eigenvalue problems

An eigenvalue $\lambda \in \mathbb{C}$ of a matrix $A \in \mathbb{C}^{n \times n}$ is a value for which the equations $Av = v\lambda$ and $w^* A = \lambda w^*$ have nontrivial solutions (the eigenvectors w^* and v). Together, (λ, v) forms an eigenpair and (λ, v, w^*) forms an eigentriple. An eigenvector is a basis for a one-dimensional invariant subspace: that is, A maps anything multiple of v to some other multiple of v. More generally, a matrix $V \in \mathbb{C}^{n \times m}$ spans an invariant subspace if $AV = VL$ for some $L \in \mathbb{C}^{n \times m}$.

Associated with any square A, we can write a matrix Q whose columns form an orthonormal basis for nested invariant subspaces of A; that is, the first k columns of Q form a k-dimensional invariant subspace of A. This structure of nested invariant subspaces gives us that

$$AQ = QT,$$

where T is an upper triangular matrix. The factorization

$$A = QTQ^*$$

is a *Schur factorization*. Most of the next week or two will be devoted to methods to compute Schur factorizations (or parts of Schur factorizations). The Schur factorization is nearly as versatile as, and is far more numerically stable than, the *Jordan canonical form*

$$AV = VJ.$$

where J is a block diagonal matrices with *Jordan* blocks of the form

$$J_\lambda = \begin{bmatrix} \lambda & 1 & & \\ & \ddots & \ddots & \\ & & \lambda & 1 \\ & & & \lambda \end{bmatrix}.$$

The *algebraic multiplicity* of an eigenvalue λ is the number of times it appears on the diagonal of the Jordan form, or the number of times the factor $z - \lambda$ divides the characteristic polynomial $\det(A - zI)$. The *geometric multiplicity* is given by the number of Jordan blocks associated to λ, or by the dimension of the null space of $(A - \lambda I)$. In general, there is exactly one eigenvector of A for each Jordan block, and the eigenvectors form a basis iff A is diagonalizable – that is, if A has only 1-by-1 Jordan blocks and all geometric and algebraic multiplicities match. The diagonalizable matrices form a dense set in $\mathbb{C}^{n\times n}$, a fact which is often convenient in proofs (since an argument for the diagonalizable case together with a continuity argument often yields a general solution). This fact also explains part of why the Jordan canonical form is annoying for numerical work: if every matrix is an arbitrarily small perturbation of something diagonalizable, then the Jordan form is discontinuous as a function of A! Even among the diagonalizable matrices, though, the eigenvector decomposition is often overrated for computational purposes. Poor conditioning of the eigenvector basis can make diagonalization a numerically unstable business, and most computations that are naively formulated in terms of an eigenvector basis can equally well be formulated in terms of Schur basis.

In *generalized eigenvalue problems*, we ask for nontrivial solutions to

$$(A - \lambda B)v = 0.$$

There are also *nonlinear eigenvalue problems*, which show up in my research but which we will not talk about in class. In addition to these variants on the eigenvalue problem, there are also many different factors that affect the how we choose algorithms. Is the problem...

1. nonsymmetric or symmetric?
2. standard or generalized?
3. to find all eigenvalues or just a few?
4. to compute eigenvectors, invariant subspaces, or just eigenvalues?

For different answers to these questions, there are different “best” choices of algorithm. For the next week or two, we will focus specifically on the problem of computing eigenpairs, invariant subspaces, and Schur forms for nonsymmetric matrices. After that, we will move on to symmetric problems, which have so much more mathematical structure that they are treated almost entirely differently from their nonsymmetric brethren.

The 2-by-2 case: some illustrative examples

Many of the salient features that occur in general eigenvalue problems can be illustrated with the 2-by-2 matrix

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

Finding an eigenvalue is equivalent to finding a root of the characteristic polynomial:

$$\begin{aligned} p(z) &= \det(A - zI) = (a-z)(d-z) - bc \\ &= z^2 - (a+d)z + (ad - bc). \end{aligned}$$

If the roots of the characteristic polynomial are λ_1 and λ_2, then we have

$$\begin{aligned} p(z) &= (z - \lambda_1)(z - \lambda_2) \\ &= z^2 - (\lambda_1 + \lambda_2)z + \lambda_1 \lambda_2. \end{aligned}$$

We recognize the second coefficient in the characteristic polynomial as minus the trace $a + d = \lambda_1 + \lambda_2$. The constant coefficient is the determinant $ad - bc = \lambda_1 \lambda_2$. Both these coefficients can be seen as functions of the eigenvalues, but both can be computed efficiently without referring to the eigenvalues explicitly.

Now suppose we choose some fixed $\lambda \in \mathbb{C}$ and look at the 2-by-2 matrices for which λ is an eigenvalue. If we just want λ to be *an* eigenvalue, we must satisfy one scalar equation: $p(\lambda) = 0$. To find matrices for which λ is a *double* eigenvalue, we must satisfy the additional constraint $a + d = 2\lambda$. And there is only one 2-by-2 matrix for which λ is a double eigenvalue with *geometric* multiplicity 2: $A = \lambda I$. Put differently, the set of 2-by-2 matrices for which λ is an eigenvalue has codimension 1 (i.e. it is described by one scalar constraint); the set of 2-by-2 matrices for which λ is an eigenvalue with algebraic multiplicity 2 has codimension 2; and the set of 2-by-2 matrices for which λ is an eigenvalue with geometric multiplicity 2 has codimension 4.

More generally, we can say that among general complex n-by-n matrices, the existence of *some* multiple eigenvalue is a codimension 1 phenomena (somewhat rare in general); and the existence of an eigenvalue with *geometric* multiplicity greater than 1 is a codimension 3 phenomena (very rare in general). Of course, things change if we consider structured matrices. For example, in symmetric matrices the algebraic and geometric multiplicities of all eigenvalues are the same.