CS 6210: Matrix Computations

Perturbation theory

David Bindel

2025-10-22

Building blocks

Similarity transforms

When we talked about least squares problems, we spent some time discussing the transformations that preserve the Euclidean norm: orthogonal transformations. It is worth spending a moment now to give a name to the transformations that preserve the eigenvalue structure of a matrix. These are *similarity* transformations.

Suppose $A \in \mathbb{C}^{n\times n}$ is a square matrix, and $X \in \mathbb{C}^{n\times n}$ is invertible. Then the matrix XAX^{-1} is said to be *similar* to A, and the mapping from A to XAX^{-1} is a *similarity transformation*. If A is the matrix for an operator from $\mathbb{C}^n$ onto itself, then XAX^{-1} is the matrix for the same operator in a different basis. The eigenvalues and the Jordan block structure of a matrix are preserved under similarity, and the matrix X gives a relationship between the eigenvectors of A and those of XAX^{-1}. Note that this goes both ways: two matrices have the same eigenvalues and Jordan block structure iff they are similar.

Symmetric polynomials

Usually, we think of the characteristic polynomial $p(z) = \det(zI - A)$ as a function of z. However, we can also think about it as a function of A. In particular, we can think of the coefficients in the characteristic polynomial as functions of A; some of these functions have names, like the determinant (the constant coefficient) and the trace (the coefficient at order $d-1$). Collectively, the coefficients are *elementary symmetric polynomials* of the eigenvalues of A — symmetric in this case meaning polynomials in n variables that are invariant under permutation of the arguments. In fact, the fundamental theorem of characteristic polynomials says that any symmetric polynomial in the n eigenvalues (i.e. any function that is the same under arbitrary permutations of the arguments) can be defined in terms of these coefficients of the characteristic polynomial. Thus, any nice symmetric function of the eigenvalues will be

a smooth function of the entries of the matrix A, *even if* the individual eigenvalues become rather unpleasant.

Contour integrals

Complex analysis is not a pre-requisite for this course, and this material may be considered optional. Nonetheless, it is useful as background both for understanding the perturbation theory of eigenvalues and for understanding certain numerical algorithms that we will describe in a couple weeks.

One of the beautiful, fundamental facts about complex analysis is the Cauchy residue theorem. Suppose $\Omega \subset \mathbb{C}$ is a simply connected domain, and that $f : \Omega \subset \mathbb{C} \to \mathbb{C}$ is holomorphic (aka analytic) except at finitely many points $\xi_1, \ldots, \xi_n$. If $\Gamma \subset \Omega$ is a rectifiable curve, then

$$\int_\Gamma f(z)\,dz = 2\pi i \sum_{j=1}^{n} I(\Gamma, \xi_j)\,\mathrm{Res}(f, \xi_j),$$

where $I(\Gamma, \xi_j)$ is the winding number (the number of times Γ goes in a positive direction around ξ_j) and $\mathrm{Res}(f, \xi_j)$ is the *residue* at ξ_j.

Closely related is the Cauchy integral theorem, which says that if $f : \Omega \to \mathbb{C}$ is holomorphic and Γ is a positively oriented simple closed curve that winds once around a, then

$$f(a) = \frac{1}{2\pi i} \int_\Gamma \frac{f(z)}{z-a}\,dz$$

and

$$f^{(n)}(a) = \frac{n!}{2\pi i} \int_\Gamma \frac{f(z)}{(z-a)^{(n+1)}}\,dz.$$

What if we replace the scalar a by a matrix A? In this case, we end up with integrals involving the *resolvent* $R(z) = (zI - A)^{-1}$, which turns out to be an extremely useful object. Suppose A is diagonalizable, with $A = V\Lambda V^{-1}$, and the spectrum of A is inside our curve Γ. Then we can consider the integral one eigen-direction at a time:

$$\begin{aligned} f(A) &= \frac{1}{2\pi i} \int_\Gamma (zI - A)^{-1} f(z)\,dz \\ &= V \left(\frac{1}{2\pi i} \int_\Gamma (zI - \Lambda)^{-1} f(z)\,dz \right) V^{-1} \\ &= V f(\Lambda) V^{-1}. \end{aligned}$$

If Γ only encloses part of the spectrum, then only those eigenvalues inside Γ are represented. In particular, we can use this to compute a *spectral projector*:

$$P_\Gamma = \frac{1}{2\pi i} \int_\Gamma (zI - A)^{-1}\,dz = \sum_{\lambda_i \text{ inside } \Gamma} v_i w_i^*$$

where $w_i^* = e_i^T V^{-1}$ is a row eigenvector for the eigenvalue λ_i. The trace of the spectral projector gives a count of the number of eigenvalues inside of the contour. We can also compute the sum of the eigenvalues inside the contour as

$$\frac{1}{2\pi i} \operatorname{tr}\left(\int_\Gamma z(zI - A)^{-1}\, dz\right).$$

All of these contour integrals are continuously defined up to the point where the contour intersects a pole. For example, using what we know about the distance to singularity, this means that if that if $\|E\| < \min_{z\in\Gamma} \|(zI - A)^{-1}\|^{-1}$, then the trace of the spectral projector for $A + sE$ remains continuously defined for $0 \le s \le 1$ – which means that A and $A + E$ have the same number of eigenvalues. This is essentially the same argument behind *Rouché's theorem*: if f and f are holomorphic on Ω and $|f(z)| \le |g(z)|$ for all z on a simple rectifiable closed contour Γ, then f and g have the same number of zeros inside Γ.

Eigenvalue perturbations

Consider the matrix

$$A(\epsilon) = \begin{bmatrix} \lambda & 1 \\ \epsilon & \lambda \end{bmatrix}.$$

The characteristic polynomial of $A(\epsilon)$ is $p(z) = z^2 - 2\lambda z + (\lambda^2 - \epsilon)$, which has roots $\lambda \pm \sqrt{\epsilon}$. These eigenvalues are *continuous* functions of ϵ at $\epsilon = 0$, but they are not differentiable functions. This is a more general phenomenon: an $O(\epsilon)$ perturbation to a matrix with an eigenvalue with multiplicity m usually splits the eigenvalue into m distinct eigenvalues, each of which is moved from the original position by $O(\epsilon^{1/m})$. We expect, then, that it will be difficult to accurately compute multiple eigenvalues of general nonsymmetric matrices in floating point. If we are properly suspicious, we should suspect that *nearly* multiple eigenvalues are almost as troublesome — and indeed they are. On the other hand, while we usually lose some accuracy when trying to compute nearly multiple eigenvalues, we should not always expect to lose *all* digits of accuracy.

The next lecture or two will be spent developing the perturbation theory we will need in order to figure out what we can and cannot expect from our eigenvalue computations.

First-order perturbation theory

Suppose $A \in \mathbb{C}^{n\times n}$ has a simple[1] eigenvalue λ with corresponding column eigenvector v and row eigenvector w^*. We would like to understand how λ changes under small perturbations to A. If we formally differentiate the eigenvalue equation $Av = v\lambda$, we have

$$(\delta A)v + A(\delta v) = (\delta v)\lambda + v(\delta\lambda).$$

[1] An eigenvalue is simple if it is not multiple.

If we multiply this equation by w^*, we have

$$w^*(\delta A)v + w^*A(\delta v) = \lambda w^*(\delta v) + w^*v(\delta\lambda).$$

Note that $w^*A = \lambda w^*$, so that we have

$$w^*(\delta A)v = w^*v(\delta\lambda),$$

which we rearrange to get

$$\delta\lambda = \frac{w^*(\delta A)v}{w^*v}. \tag{1}$$

This formal derivation of the first-order sensitivity of an eigenvalue only goes awry if $w^*v = 0$, which we can show is not possible if λ is simple.

We can use Equation 1 to get a condition number for the eigenvalue λ as follows:

$$\frac{|\delta\lambda|}{|\lambda|} = \frac{|w^*(\delta A)v|}{|w^*v||\lambda|} \leq \frac{\|w\|_2\|v\|_2}{|w^*v|}\frac{\|\delta A\|_2}{|\lambda|} = \sec\theta\frac{\|\delta A\|_2}{|\lambda|}.$$

where θ is the acute angle between the spaces spanned by v and by w. When this angle is large, very small perturbations can drastically change the eigenvalue.

Gershgorin theory

The first-order perturbation theory outlined in the previous section is very useful, but it is also useful to consider the effects of *finite* (rather than infinitesimal) perturbations to A. One of our main tools in this consideration will be Gershgorin's theorem.

Here is the idea. We know that diagonally dominant matrices are nonsingular, so if $A - \lambda I$ is diagonally dominant, then λ cannot be an eigenvalue. Contraposing this statement, λ can be an eigenvalue only if $A - \lambda I$ is *not* diagonally dominant. The set of points where $A - \lambda I$ is not diagonally dominant is a union of sets $\cup_j G_j$, where each G_j is a *Gershgorin disk*:

$$G_j = B_{\rho_j}(a_{jj}) = \left\{ z \in \mathbb{C} : |a_{jj} - z| \leq \rho_j \text{ where } \rho_j = \sum_{i\neq j} |a_{ij}| \right\}.$$

Our strategy now, which we will pursue in detail next time, is to use similarity transforms based on A to make a perturbed matrix $A+E$ look "almost" diagonal, and then use Gershgorin theory to turn that "almost" diagonality into bounds on where the eigenvalues can be.

We now argue that we can extract even more information from the Gershgorin disks: we can get *counts* of how many eigenvalues are in different parts of the union of Gershgorin disks.

Suppose that $\mathcal{G}$ is a connected component of $\cup_j G_j$; in other words, suppose that $\mathcal{G} = \cup_{j\in S} G_j$ for some set of indices S, and that $\mathcal{G} \cap G_k = \emptyset$ for $k \notin S$. Then the number of eigenvalues of A

in $\mathcal{G}$ (counting eigenvalues according to multiplicity) is the same as the side of the index set S.

To sketch the proof, we need to know that eigenvalues are continuous functions of the matrix entries. Now, for $s \in [0,1]$, define

$$H(s) = D + sF$$

where D is the diagonal part of A and $F = A - D$ is the off-diagonal part. The function $H(s)$ is a *homotopy* that continuously takes us from an easy-to-analyze diagonal matrix at $H(0) = D$ to the matrix we care about at $H(1) = A$. At $s = 0$, we know the eigenvalues of A are the diagonal elements of A; and if we apply the first part of Gershgorin's theorem, we see that the eigenvalues of $H(s)$ always must live inside the union of Gershgorin disks of A for any $0 \leq s \leq 1$. So each of the $|S|$ eigenvalues that start off in the connected component $\mathcal{G}$ at $H(0) = D$ can move around continuously within $\mathcal{G}$ as we move the matrix continuously to $H(1) = A$, but they cannot "jump" discontinuously across the gap between $\mathcal{G}$ and any of the other Gershgorin disks. So at $s = 1$, there will still be $|S|$ eigenvalues of $H(1) = A$ inside $\mathcal{G}$.

Perturbing Gershgorin

Now, let us consider the relation between the Gershgorin disks for a matrix A and a matrix $\hat{A} = A + F$. It is straightforward to write down the Gershgorin disks $\hat{G}_j$ for $\hat{A}$:

$$\hat{G}_j = \mathcal{B}_{\hat{\rho}_j}(\hat{a}_{jj}) = \left\{ z \in \mathbb{C} : |a_{jj} + e_{jj} - z| \leq \hat{\rho}_j \right\} \text{ where } \hat{\rho}_j = \sum_{i \neq j} |a_{ij} + f_{ij}|.$$

Note that $|a_{jj} + e_{jj} - z| \geq |a_{jj} - z| - |f_{jj}|$ and $|a_{ij} + f_{ij}| \leq |a_{ij}| + |f_{ij}|$, so

$$\hat{G}_j \subseteq \mathcal{B}_{\rho_j + \sum_j |f_{ij}|}(a_{jj}) = \left\{ z \in \mathbb{C} : |a_{jj} - z| \leq \rho_j + \sum_i |f_{ij}| \right\}. \tag{2}$$

We can simplify this expression even further if we are willing to expand the regions a bit:

$$\hat{G}_j \subseteq \mathcal{B}_{\rho_j + \|F\|_1}(a_{jj}). \tag{3}$$

The Bauer-Fike theorem

We now apply Gershgorin theory together with a carefully chosen similarity to prove a bound on the eigenvalues of $A + F$ where F is a finite perturbation. This will lead us to the *Bauer-Fike* theorem.

The basic idea is as follows. Suppose that A is a diagonalizable matrix, so that there is a complete basis of column eigenvectors V such that

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

Then we $A + F$ has the same eigenvalues as

$$V^{-1}(A + F)V = \Lambda + V^{-1}FV = \Lambda + \tilde{F}.$$

Now, consider the Gershgorin disks for $\Lambda + \tilde{F}$. The crude bound Equation 3 tells us that all the eigenvalues live in the regions

$$\bigcup_j \mathcal{B}_{\|\tilde{F}\|_1}(\lambda_j) \subseteq \bigcup_j \mathcal{B}_{\kappa_1(V)\|F\|_1}(\lambda_j).$$

This bound really is crude, though; it gives us disks of the same radius around all the eigenvalues λ_j of A, regardless of the conditioning of those eigenvalues. Let's see if we can do better with the sharper bound in Equation 2.

To use Equation 2, we need to bound the absolute column sums of $\tilde{F}$. Let e represent the vector of all ones, and let e_j be the jth column of the identity matrix; then the jth absolute column sums of $\tilde{F}$ is $\phi_j \equiv e^T|\tilde{F}|e_j$, which we can bound as $\phi_j \leq e^T|V^{-1}||F||V|e_j$. Now, note that we are free to choose the normalization of the eigenvector V; let us choose the normalization so that each row of $W^* = V^{-1}$. Recall that we defined the angle θ_j by

$$\cos(\theta_j) = \frac{|w_j^* v_j|}{\|w_j\|_2 \|v_j\|_2},$$

where w_j and v_j are the jth row and column eigenvectors; so if we choose $\|w_j\|_2 = 1$ and $w_j^* v_j = 1$ (so $W^* = V^{-1}$), we must have $\|v_j\|_2 = \sec(\theta_j)$. Therefore, $\||V|e_j\|_2 = \sec(\theta_j)$. Now, note that $e^T|V^{-1}|$ is a sum of n rows of Euclidean length 1, so $\|e^T|V^{-1}|\|_2 \leq n$. Thus, we have

$$\phi_j \leq n\|F\|_2 \sec(\theta_j).$$

Putting this bound on the columns of $\tilde{F}$ together with Equation 2, we have the Bauer-Fike theorem.

Theorem 0.1 (Bauer-Fike)**.** *Suppose $A \in \mathbb{C}^{n\times n}$ is diagonalizable with eigenvalues $\lambda_1, \ldots, \lambda_n$. Then all the eigenvalues of $A + F$ are in the region*

$$\bigcup_j \mathcal{B}_{n\|F\|_2 \sec(\theta_j)}(\lambda_j),$$

where θ_j is the acute angle between the row and column eigenvectors for λ_j, and any connected component $\mathcal{G}$ of this region that contains exactly m eigenvalues of A will also contain exactly m eigenvalues of $A + F$.