Week 10: Monday, Oct 26

Logistics

The midterm was to be turned in by the start of today's lecture. I sketched solutions to some of the problems on the board; more complete solutions are now available on the course CMS.

Gershgorin theory revisited

Recall from last time that the eigenvalues of a matrix A must be inside the union of the *Gershgorin disks*

$$G_j = \mathcal{B}_{\rho_j}(a_{jj}) \equiv \{z \in \mathbb{C} : |a_{jj} - z| \leq \rho_j\} \text{ where } \rho_j = \sum_{i \neq j} |a_{ij}|.$$

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 $\hat{A}$:

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

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

$$(1) \qquad \hat{G}_j \subseteq \mathcal{B}_{\rho_j + \sum_j |f_{ij}|}(a_{jj}) = \left\{ z \in \mathbb{C} : |a_{jj} - z| \leq \rho_j + \sum_i |f_{ij}| \right\}.$$

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

$$(2) \qquad \hat{G}_j \subseteq \mathcal{B}_{\rho_j + \|F\|_1}(a_{jj}).$$

The Bauer-Fike theorem

We now apply Gershgorin theory together with a carefully chosen similarity to prove a bound on the eigenvalues of $A+E$ where E 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

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

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 (2) 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 (1).

To use (1), 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 (1), we have the Bauer-Fike theorem.

Theorem 1 *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 $A + F$.