MAS565 Numerical Analysis HW 9

20218125 최지석

#6.25. By the Geršgorin circle theorem, we know that three eigenvalues of A all lie on the union of three disks

$$D_1 = \{ z : |z - 5.2| \leq 2.8 \},$$

$$D_2 = \{ z : |z - 6.4| \leq 1.1 \},$$

$$D_3 = \{ z : |z - 4.7| \leq 2.7 \}.$$

As A is symmetric all eigenvalues of A are real. The intersection of the real axis with the disks D_1, D_2, and D_3 are $[2.4, 8]$, $[5.3, 7.5]$, and $[2, 7.4]$, respectively. Hence for any eigenvalue $\lambda(A)$ of A, we have the bounds $2 \leq \lambda(A) \leq 8$. Recall Exercise 8(a), from which we get

$$\mathrm{cond}_2(A) = \frac{\lambda_{max}(A)}{\lambda_{min}(A)} \leq \frac{8}{2} = 4$$

since A being symmetric implies A being normal.

#6.26. (a) We know nothing about off-diagonal entries except that all off-diagonal entries have modulus less than or equal to $\frac{1}{4}$. Hence all we can do is to apply Geršgorin circle theorem directly. Once we note that for each row and column the sum of the moduli of off-diagonal entries are at most 1 thus each Geršgorin disks have radius bounded by 1, we conclude that the eigenvalues $\lambda_1, \cdots, \lambda_5$ of A with

$|\lambda_1| \ge \cdots \ge |\lambda_5|$ we have

$$|\lambda_1 - 21| \le 1, \quad |\lambda_2 + 9| \le 1, \quad |\lambda_3 - 4| \le 1,$$

and since two Geršgorin disks $\{z : |z-1| \le 1\}$ and $\{z : |z-0| \le 1\}$ have nonempty intersection the best we can conclude is that

$$\{\lambda_4, \lambda_5\} \in \{z : |z-1| < 1 \text{ or } |z| < 1\}.$$

(b) Consider $D = \operatorname{diag}\{d_1, d_2, d_3\}$ then for $A = \begin{bmatrix} 1 & 10^{-3} & 10^{-4} \\ 10^{-3} & 2 & 10^{-3} \\ 10^{-4} & 10^{-3} & 3 \end{bmatrix}$. the matrix $D^{-1}AD$ has the same spectrum with A and its (i,j)-entry is $\frac{a_{ij} d_j}{d_i}$ as long as $d_1 d_2 d_3 \neq 0$. Denote $\rho_i = \sum_{\substack{k=1 \\ k \neq i}}^{3} \left| \frac{a_{ik} d_k}{d_i} \right|$. Note that the diagonal entries are invariant under this similarity transformation. So the Geršgorin disks are always centered at 1, 2, and 3, hence, the Geršgorin disks always can be written as $D_i = \{z : |z - i| \le \rho_i\}$ for $i = 1, 2, 3$.

Consider the special case where $d_1 = 1$, $d_2 = 10k$, $d_3 = k$ for positive real k. Then $\rho_1 = 10^{-2}k + 10^{-4}k$, $\rho_2 = 10^{-4}/k + 10^{-4}$, and $\rho_3 = 10^{-4}/k + 10^{-2}$. As $\rho_3 < \rho_2 + 1$, if $D_2 \cap D_1 = \emptyset$ then $D_3 \cap D_1 = \emptyset$ also. For $D_2 \cap D_1$ to be $\emptyset$ we must have $\rho_1 + \rho_2 < 1$. As $k \mapsto \rho_1 + \rho_2$ defines a convex function we have an interval that satisfies $\rho_1 + \rho_2 < 1$, and the infimum of k satisfying $\rho_1 + \rho_2 < 1$ is such that $\rho_1 + \rho_2 = 1$. Solving the equation $\rho_1 + \rho_2 = (10^{-2} + 10^{-4})k + \frac{10^{-4}}{k} + 10^{-4} = 1$ the root of this equation that is smaller is $\frac{9999 - \sqrt{99979597}}{202} < 1.0002 \times 10^{-4}$. So ρ_1 can be made as small as 1.0102×10^{-6}.

Choosing $d_1 = k$, $d_2 = 10k$, and $d_3 = 1$ on the other hand now gives $\rho_1 = 10^{-4}/k + 10^{-2}$, $\rho_2 = 10^{-4}/k + 10^{-4}$, and $\rho_3 = 10^{-2}k + 10^{-4}k$, so the exact same logic in the previous paragraph except 1 and 3 exchanged applies, allowing

ρ_3 to be as small as 1.0102×10^{-6}. For each $i=1,2,$ and 3 as A is already a diagonally dominant matrix we have no ambiguity of letting λ_i to be the eigenvalue of A that is closest to i, then we get $|\lambda_1-1|<1.0102\times10^{-6}$ and $|\lambda_3-3|<1.0102\times10^{-6}$.

Finally to estimate λ_2, consider the special case where $\lambda_1=k$, $\lambda_2=1$, and $\lambda_3=k$ for some positive real k. Then we get $\rho_1=\rho_3=10^{-3}/k+10^{-4}$, and $\rho_2=2\times10^{-3}k$. Now the configuration of the disks become symmetric along the vertical line $\mathrm{Re}(z)=2$, so we get that three disks D_1, D_2, D_3 are disjoint if and only if $\rho_1+\rho_2<1$. Again $k\mapsto\rho_1+\rho_2$ is a convex function, so infimum of k occurs when $\rho_1+\rho_2=1$. Solving this equation we get $k=\frac{9999-\sqrt{99979201}}{40}<1.0002\times10^{-3}$, so ρ_2 can be made as small as 2.0004×10^{-6}. Hence $|\lambda_2-2|<2.0004\times10^{-6}$.

#6.27.(a) Let X be the block diagonal matrix $X=\begin{bmatrix}A&0\\0&B\end{bmatrix}$ then by the normality of H, from Theorem 6.9.7, we get that for any eigenvalue $\lambda(X)$ of X there exists an eigenvalue $\lambda(H)$ of H such that

$$|\lambda(X)-\lambda(H)|\le \mathrm{lub}_2(X-H). \qquad \cdots(*)$$

Now as $\det(X-\mu I)=\det\begin{bmatrix}A-\mu I&0\\0&B-\mu I\end{bmatrix}=\det(A-\mu I)\det(B-\mu I)$, any eigenvalue $\lambda(B)$ of B is also an eigenvalue of X. Thus for any eigenvalue $\lambda(B)$ of B, there exists an eigenvalue $\lambda(H)$ of H such that

$$|\lambda(B)-\lambda(H)|\le \mathrm{lub}_2(X-H).$$

If we can show that $\mathrm{lub}_2(X-H) \le \sqrt{\mathrm{lub}_2(C^H C)}$ then we are done.

Observe that $X - H = \begin{bmatrix} 0 & -C \\ -C^H & 0 \end{bmatrix}$ so for vectors x and y of suitable sizes and $\left\| \begin{bmatrix} x \\ y \end{bmatrix} \right\|_2^2 = \|x\|^2 + \|y\|^2 = 1$, let σ_{max} be the largest singular value of C then

$$
\begin{aligned}
(\mathrm{lub}_2(X-H))^2 &= \left(\max_{\left\|\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right]\right\|_2=1} \left\| \begin{bmatrix} 0 & C \\ C^H & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \right\|_2 \right)^2 \\
&= \left(\max_{\left\|\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right]\right\|_2=1} \left\| \begin{bmatrix} Cy \\ C^H x \end{bmatrix} \right\|_2 \right)^2 \\
&= \max_{\left\|\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right]\right\|_2=1} \left\| \begin{bmatrix} Cy \\ C^H x \end{bmatrix} \right\|_2^2 \qquad \left(\because \left\| \begin{bmatrix} Cy \\ C^H x \end{bmatrix} \right\|_2 \ge 0 \right) \\
&= \max_{\|x\|_2^2 + \|y\|_2^2 = 1} \|Cy\|_2^2 + \|C^H x\|_2^2 \\
&\le \max_{\|x\|_2^2 + \|y\|_2^2 = 1} \sigma_{max}^2 \|y\|^2 + \sigma_{max}^2 \|x\|^2 \\
&= \sigma_{max}^2 .
\end{aligned}
$$

But we know that $\sigma_{max}^2 = \mathrm{lub}_2(C^H C)$. Therefore $\mathrm{lub}(X-H) \le \sqrt{\mathrm{lub}_2(C^H C)}$ indeed.

(b) Actually from (*) it follows that for any eigenvalue $\lambda(A)$ of A there exists an eigenvalue $\lambda(H)$ of H such that

$$| \lambda(A) - \lambda(H) | \le \mathrm{lub}_2(X-H) \le \sqrt{\mathrm{lub}_2(C^H C)}.$$

Furthermore as X is also Hermitian hence normal, we can change the role of X and H to conclude the reverse, that is, for any eigenvalue $\lambda(H)$ of H, there exists an eigenvalue λ of either A or B such that

$$| \lambda(H) - \lambda | \le \mathrm{lub}_2(H-X) \le \sqrt{\mathrm{lub}_2(C^H C)} .$$

In the almost reducible tridiagonal case, $C^H C$ is a matrix with $(1,1)$-entry ε^2 and all other entries 0. Being diagonal, we have $\sqrt{\mathrm{lub}_2(C^H C)} = \sqrt{\varepsilon^2} = \varepsilon$. Therefore, all eigenvalues of H are in distance at most ε from some eigenvalue of either A or B.

#6.32. In order to apply Bendixon's theorem, define

$$H_1 = \frac{1}{2}(A+A^H) = \begin{bmatrix} 3 & 1 & 0 \\ 1 & 15 & -2 \\ 0 & -2 & 19.5 \end{bmatrix} \quad \text{and} \quad H_2 = \frac{1}{2i}(A-A^H) = \begin{bmatrix} 0 & -0.1i & 0.1i \\ 0.1i & 0 & 0.1i \\ -0.1i & -0.1i & 0 \end{bmatrix}$$

then a direct application of the Geršgorin theorem on H_2 we get $-0.2 \le \lambda(H_2) \le 0.2$ for any eigenvalue $\lambda(H_2)$ of H_2, so it follows that $-0.2 \le \operatorname{Im} \lambda \le 0.2$ for any eigenvalue λ of A. Also applying Geršgorin theorem directly to H_1 we see that $2 \le \lambda_{min}(H_1) \le \lambda_{max}(H_1) \le 21.5$, but using similarity transformations as in #6.26(b) we can improve this bound. If we use a similarity transform $A \mapsto D^{-1}AD$ with an invertible diagonal matrix D then the diagonal entries do not change, and since A and H_1 have the same main diagonals, for convenience (with abuse of notation) we let D_1, D_2, and D_3 the Geršgorin disks each centered at 3, 15, and 19.5, and let their radii be ρ_1, ρ_2, and ρ_3 respectively. As, for any matrix A, A and A^T share the same spectrum, the Geršgorin disks can be constructed columnwise, that is, we may take ρ_i's to be the sum of the moduli of off-diagonal entries in the i^{th} column. This will be the case in this solution. Consider $D^{-1}H_1D$ with $D = \operatorname{diag}\{\lambda\mu, \mu, 1\}$, $\lambda > 0$, $\mu > 0$, then $\rho_1 = \lambda$, $\rho_2 = \frac{1}{\lambda} + 2\mu$, and $\rho_3 = \frac{2}{\mu}$. Finding an exact solution of a nonconvex minimization problem minimizing $\min\{3-\rho_1, 15-\rho_2, 19.5-\rho_3\}$ is a difficult task, so we just take $\lambda = \frac{1}{11}$, $\mu = \frac{1}{2}$ so that $\min\{3-\rho_1, 15-\rho_2, 19.5-\rho_3\} = 3 - \frac{1}{11}$. Similarly, maximizing $\max\{3+\rho_1, 15+\rho_2, 19.5+\rho_3\}$ is also a difficult task, so we take $\mu = \frac{9}{4}$, $\lambda = \frac{9}{8}$ so that $\max\{3+\rho_1, 15+\rho_2, 19.5+\rho_3\} = \frac{367}{18}$. From Bendixson theorem we have a rectangular region $R = \{z : \frac{32}{11} \le \operatorname{Re} z \le \frac{367}{18}, -0.2 \le \operatorname{Im} z \le 0.2\}$ which includes all eigenvalues of A. However we can reduce the region further by applying Geršgorin theorem to $D^{-1}AD$ for a suitable diagonal matrix D. Take $D = \mathrm{diag}\{k, 1, 1\}$ for positive k. Then it becomes $\rho_1 = 0.9k + 0.1k = k$, $\rho_2 = 1.9 + \frac{1.1}{k}$, and $\rho_3 = 2.1 + \frac{0.1}{k}$. For D_1 to be as small as possible while not intersecting D_2 nor D_3, we need $\rho_1 + \rho_2 < 12$, and if $k < 5$ then $\rho_2 > \rho_3$ hence $D_2 \cap D_1 = \emptyset$ will imply $D_3 \cap D_1 = \emptyset$. Using an argument similar to #6.26(b) the optimal k is the one satisfying $\rho_1 + \rho_2 = k + 1.9 + \frac{1.1}{k} = 12$, which is $k^* = \frac{1}{20}(101 - \sqrt{9761}) \approx 0.1101$. That is, the eigenvalue of A with the smallest modulus is in the disk $D_1^* = \{z : |z-3| \le \frac{1}{20}(101 - \sqrt{9761})\}$. On the other hand, take $D = \mathrm{diag}\{1, k, 1\}$, then it becomes $\rho_1 = \frac{0.9}{k} + 0.1$, $\rho_2 = 1.1k + 1.9k = 3k$, and $\rho_3 = 0.1 + \frac{2.1}{k}$. For $D_1 \cap D_2 = \emptyset$ we need $\rho_1 + \rho_2 = \frac{0.9}{k} + 0.1 + 3k < 12$, which gives the interval approximately $(0.07713, 3.8895)$. For $D_1 \cap D_3 = \emptyset$ we need $\rho_1 + \rho_3 = \frac{3}{k} + 0.2 < 16.5 \Leftrightarrow k > \frac{30}{163} \approx 0.1840$. That is, for D_1 not to intersect $D_2 \cup D_3$ we need $\frac{30}{163} < k < 3.8895$. Now what we want is to have $D_2 \cup D_3$ as small as possible, but we already have an upper bound for $\mathrm{Re}\,\lambda$ so we need a lower bound of $\{\mathrm{Re}(z) : z \in D_2 \cup D_3\}$. Then we have to maximize $\min\{15 - \rho_2, 19.5 - \rho_3\} = \min\{15 - 3k, 19.4 - \frac{2.1}{k}\}$. As $15 - 3k$ is a decreasing function while $19.4 - \frac{2.1}{k}$ is increasing on $k > 0$, the maximum is when $15 - 3k = 19.4 - \frac{2.1}{k}$. Solving this for k we get $k^{**} = \frac{\sqrt{1114} - 22}{30} \approx 0.3792$, and the maximum becomes $15 - 3k^{**} \approx 13.862$. Especially, this happens when D_2 is inscribed in D_3, so $D_3^* = \{z : |z - 19.5| \le 0.1 + \frac{63}{\sqrt{1114} - 22}\}$ is the region containing the remaining two eigenvalues of A.

In conclusion we obtain the region $R \cap (D_1^* \cup D_3^*)$ which contains all three eigenvalues of A. This region is shaded in the complex plane as the following figure