CS556 Iterative Methods Homework 1 Solution.

(1a) For the 1D finite difference problem on $[0, L] = [0, 1]$ with uniform grid-spacing $\Delta x = 1/(n+1)$, the governing system matrix is

$$A = \frac{1}{\Delta x^2} \text{tridiag}(-1, 2, -1). \quad (1)$$

Since A is SPD, the largest eigenvalue satisfies

$$\lambda_{\max} \geq \max_{\underline{z} \in \mathbb{R}^n} \frac{\underline{z}^T A \underline{z}}{\underline{z}^T \underline{z}}. \quad (2)$$

Take $\underline{z} = \underline{e}_j$, the jth column of I, the $n \times n$ identity matrix. Then $\underline{z}^T A \underline{z} = a_{jj} = 2/\Delta x^2 = 2(n+1)^2 \sim 2n^2$, so,

$$\lambda_{\max}(A) \geq \frac{2}{\Delta x^2} \sim 2n^2 = O(n^2). \quad (3)$$

This lower bound is within a factor of two of the actual asymptotic value, $\lambda_{\max}(A) \sim 4/\Delta x^2$, but it is a reasonable first cut at an approximation. As we will see below, it works well for the nonuniform mesh case where a closed-form approximation is not readily accessible.

For every useful discretization of the continuous problem, we have

$$\lambda_{\min}(A) \sim \pi^2. \quad (4)$$

Therefore, the condition number of A satisfies

$$\kappa(A) = \frac{\lambda_{\max}}{\lambda_{\min}} = \frac{O(n^2)}{\pi^2} = O(n^2). \quad (5)$$

(1b) For the same problem with nonuniform spacing, the mininum eigenvalue relationship (4) holds, so $\lambda_{\min} \sim \pi^2$. To get the maximum eigenvalue, look at the jth equation:

$$-\left.\frac{d^2u}{dx^2}\right|_{x_j} \approx -\frac{1}{\overline{\Delta x}} \left[\frac{u_{j+1} - u_j}{\Delta x_{j+1}} - \frac{u_j - u_{j-1}}{\Delta x_j} \right] = \frac{1}{\Delta x_j \Delta x_{j+1}} \left(\alpha u_{j-1} + \beta u_j + \gamma u_{j+1} \right), \quad (6)$$

with

$$\alpha = -\frac{\Delta x_{j+1}}{\overline{\Delta x}}, \quad \beta = 2, \quad \gamma = -\frac{\Delta x_j}{\overline{\Delta x}}. \quad (7)$$

Therefore, the diagonal entry is

$$a_{jj} = \frac{2}{\Delta x_j \Delta x_{j+1}} = O(\Delta x_j^{-2}), \quad (8)$$

where the last equality holds by virtue of the assumption that the variation between Δx_j and Δx_{j+1} is bounded by a constant. (For the Chebyshev point distribution that ratio turns out to be ≈ 3, rather than 2 as hinted at in the assignment.) From Part **(1a)**, we have that

$$\lambda_{\max}(A) \geq \max_j |a_{jj}| = O(\Delta x_{\min}^{-2}). \quad (9)$$

Therefore, the condition number of A satisfies

$$\kappa(A) = \frac{\lambda_{\max}}{\lambda_{\min}} = \frac{O(\Delta x_{\min}^{-2})}{\pi^2} = O(\Delta x_{\min}^{-2}). \quad (10)$$

(1c) Use Gershgorin to localize the eigenvalues of $D^{-1}A$ for both the uniform and nonuniform cases.

- *Uniform Case.* Here, $D^{-1} = \frac{h^2}{2} I$, so $D^{-1}A$ =tridiag$(-\frac{1}{2},\ 1\ -\frac{1}{2})$. Consequently, all Gershgorin discs are centered at (1,0) in the complex λ-plane. On every row, save for the first and last, the disc radius is

$$\rho_i \;=\; \sum_{j \neq i} |\alpha_{ij}| \;=\; \frac{1}{2} + \frac{1}{2} \;=\; 1, \tag{11}$$

  where α_{ij} are taken to be the entries of $D^{-1}A$. The discs associated with the first and last rows of the matrix have radius $\frac{1}{2}$, which implies that they lie within the union of the remaining discs and therefore do not enlarge potential domain of the eigenvalues. Thus, the eigenvalues of $D^{-1}A$ are in the disc of radius 1 centered at (0,1) in the complex plane. Given that $D^{-1}A$ is also real symmetric, then the eigenvalues are real and we can assert that $\lambda \in [0, 2]$

- *Nonuniform Case.* From (6)–(7) we see that $D^{-1}A$ for the nonuniform case leads to

$$\left[D^{-1}A\right]_j \;=\; \text{tridiag}(\frac{\alpha}{\beta},\ 1,\ \frac{\gamma}{\beta}). \tag{12}$$

  By construction, the Gershgorin discs are all centered at (1,0). For the interior points, the radii are

$$\rho_j \;=\; \left|\frac{\alpha}{\beta}\right| \;+\; \left|\frac{\gamma}{\beta}\right| \;+\; = \; \frac{1}{2\overline{\Delta x}} \left(|\Delta x_{j+1}| \;+\; |\Delta x_j|\right) \;=\; 1. \tag{13}$$

  As in the case of the uniform points, the discs associated with the end points are centered at (1,0) and within the unit discs of the remaining rows.

The preceding analysis *shows that diagonal scaling successfully localizes the eigenvalues for both the uniform and nonuniform grid spacing cases.*

Unfortunately, this is a necessary condition but in this case not sufficient to prove convergence because the largest Gershgorin disc touches both endpoints of the interval [0,2], which means that the spectral radius of $E = I - D^{-1}A$ could be as large as 1, rather than strictly < 1. This case is unlike the Helmholtz problem considered in class, $H = I + \gamma^2 A$, where Gershgorin would immediately localize the discs of E to have radii < 1, independent of n.

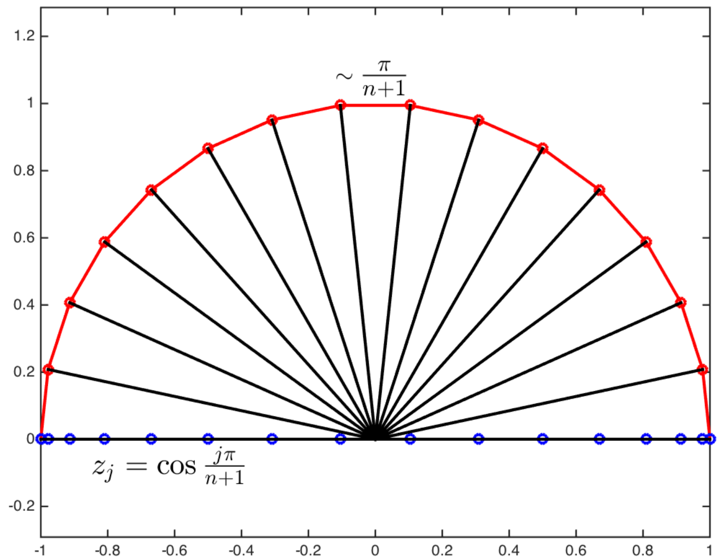

Figure 1: Chebyshev points on [-1,1].

(1d) Figure 1 shows the Chebyshev points, $z_j = cos(\theta_j)$ for uniformly-distributed angles $\theta_0, \ldots, \theta_{n+1}$ on $[-\pi, \pi]$. (Our points are translated and scaled to $[0, 1]$, $x_j = \frac{1}{2}(1 + z_j)$.) Here, we discuss the relative spacings, per the HW1 questions.

- As illustrated in Fig. 1, the maximum spacing $\Delta z_j \sim \frac{\pi}{n+1}$, which corresponds to the arclength of each segment. As $\Delta\theta_j \longrightarrow 0$ the distance between the center points tends towards this arclength. Similarly, max $\Delta x_j \sim \frac{\pi}{2(n+1)}$.
- From Fig. 1 scales as

$$\Delta z_{\min} \quad \sim \quad 1 - cos\frac{\pi}{n+1} \ \sim \ \frac{\theta^2}{2}, \tag{14}$$

  where $\theta := \pi/(n + 1)$. Consequently, $\Delta x_{\min} \sim \theta^2/4$.
- The maximum ratio occurs near the ends of the domain. After the first point, the second grid point is located at $x_2 \sim (2\theta)^2/4$, so that the second space is of size $\sim 3\theta^2/4 \sim 3\Delta x_{\min}$. Thus, the slowly varying condition holds (adjacent spacings vary by at most a factor of 3).

  In fact, a closer inspection reveals that

$$a_{jj} \quad = \quad \frac{2}{\Delta x_j \Delta x_{j+1}}. \tag{15}$$

  Since the size ratio for the smallest Chebyshev points is 3, we can anticipate that $\max_j a_{jj} \sim 2/(3\Delta x_{min}^2)$, which is plotted in Fig. 2, along with the actual eigenvalues.

Summary. The maximum eigenvalue and condition number of A for the uniform case scale like $O(n^2)$ while those for the Chebyshev case scale like $O(n^4)$, which is much more severe. Diagonal scaling, however, brings all the eigenvalues into the unit disc for both cases and yields a condition number that scales as $O(n^2)$, as illustrated in Fig. 1.

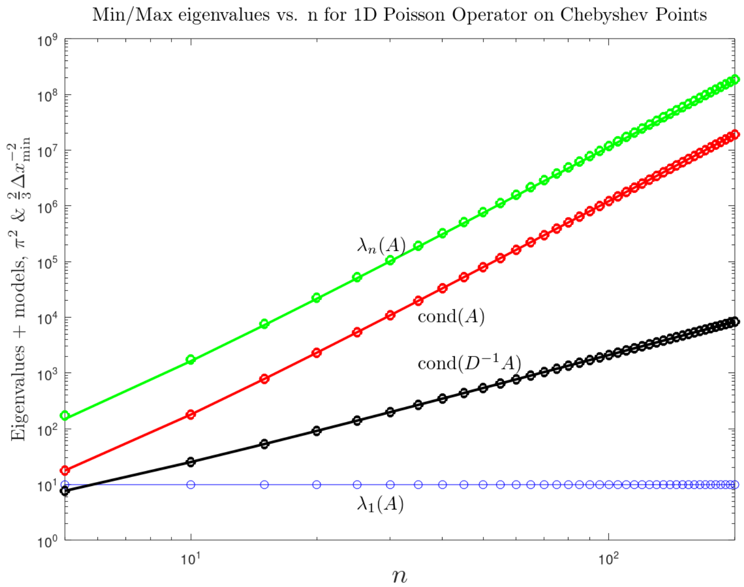

Figure 2: Computed (symbols) and modeled (lines) eigenvalues for Poisson operator on Chebyshev points: green=$\lambda_{\max}$, blue=$\lambda_{\min} \sim \pi^2$. Also shown is the condition number for A_{cheb}, which scales as $O(n^4)$, and for $D^{-1}A_{\text{cheb}}$, which scales as $O(n^2)$.

2. With each power of the tridiagonal matrix A the number of diagonals increases by 2. So we start with $\sim 3n$ nonzeros, then $\sim 5n$ for $k = 2$, and so forth. Neglecting end effects (i.e., for $k \ll n$) the number of nonzeros for A^k is $\sim n(1 + 2k)$.

3. Consider a vector $\underline{e}$ that is zero everywhere except for entries e_j, $j = j_0 : j_1$. The matrix-vector product $\underline{w} = A\underline{e}$ will be a linear combination of columns $\underline{a}_{j_0} : \underline{a}_{j_1}$ of A, which will have nonzeros in locations $j = j_0' : j_1'$, where $j_0' = \max(1, j_0 - 1)$ and $j_1' = \min(n, j_1 + 1)$. Thus, if we start with $\underline{e} = \underline{e}_j$, the jth column of the identity matrix, the number of nonzeros will grows by 1 on either side of j for each matrix-vector product. If $j = n/2$ (assuming n even), then it will take $k = n/2$ matrix-vector products to produce a nonzero in every location of $A^k\underline{e}_j$.

4. Figure 3 shows the columns of $Z := A^{-1}$. We see that all the entries are strictly positive and that A^{-1} is consequently completely full.

If we were to solve $A\underline{u} = \underline{f}$ problem in parallel, with $\underline{f}$ a distributed data vector and $\underline{u}$ a distributed solution vector then we would need an all-to-all communication because *each* nonzero in $\underline{f}$ has a nontrivial influence on every value of $\underline{u}$. This assertion is a direct consequence of the fact that every column vector of A^{-1} is completely full.

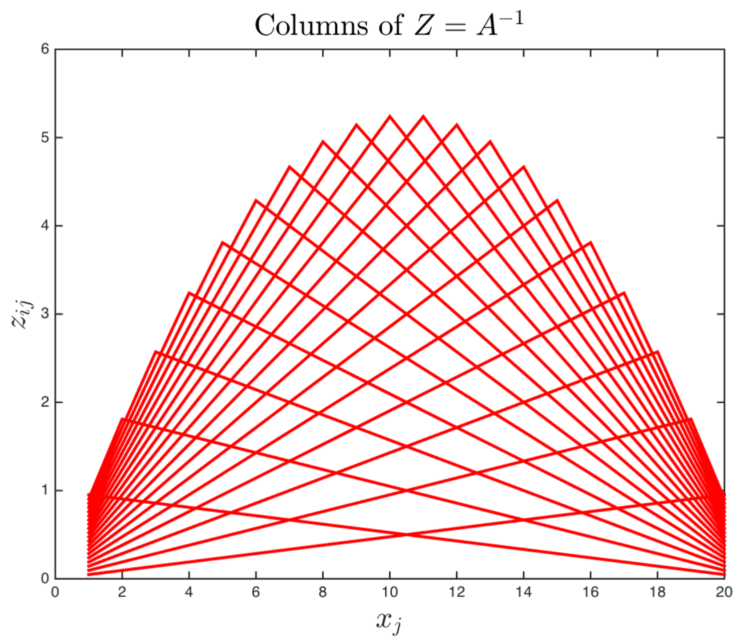

Figure 3: Columns of A^{-1} for n=20, where A is the 1D Poisson operator on a uniform grid.