Homework 6 Numerical Analysis Fall 2023

Instructions:

- Due 12/1 at 5:00pm on Gradescope.
- You must follow the submission policy in the syllabus
- This homework is worth 40 points (less than the usual 50 points)

Problem 1. Spend at least two hours working on your project prior to the 20th. Answer the following:

(a) What is the current status of your project?

(b) What are the big tasks you have left to do before your project is done?

(c) What is your plan for completing the project in a timely manner?

Problem 2. This problem will illustrate that solving the normal equations is less stable than other approaches.

(a) For each $\kappa = 10^1, 10^2, 10^3 \ldots, 10^8$, construct a 500×100 matrix $\mathbf{A}$ whose condition number is κ. A simple way to do this is to generate $\mathbf{U}$ and $\mathbf{V}$ as random orthogonal matrices of size $m \times n$ and $n \times n$ and define $\mathbf{\Sigma}$ as a $n \times n$ diagonal matrix manually.

The following code gets you started, you just need to modify the line for the singular values `s = ...`.

```
m,n = 500,100
U,_ = np.linalg.qr(np.random.randn(m,n))
V,_ = np.linalg.qr(np.random.randn(n,n))
s = #TODO

A = U@np.diag(s)@V.T
```

Let $\mathbf{b}$ be the all ones vector, and compute the "true" solution to the least squares problem $\min_{\mathbf{x}} \|\mathbf{b} - \mathbf{A}\mathbf{x}\|_2$ by setting $\mathbf{x}_{\text{true}} = \mathbf{V}\mathbf{\Sigma}^{-1}\mathbf{U}^\mathsf{T}\mathbf{b}$. This can be done with the following code:

```
b = np.ones(m)
x_true = V@np.diag(1/s)@U.T@b
```

Now, compute the least squares solution via:

- numpy's least squares solver `np.linalg.lstsq`
- a QR based approach with numpy's `np.linalg.qr` and `np.linalg.solve` or `sp.linalg.solve_triangular`
- Solving the normal equations with `np.linalg.solve`

For each of these three methods and each value of κ, record the relative error $\|\mathbf{x}_{\text{method}} - \mathbf{x}_{\text{true}}\|_2 / \|\mathbf{x}_{\text{true}}\|_2$, where $\mathbf{x}_{\text{method}}$ is the solution obtained by the given method.

(b) Make a log-log plot with the following five (labeled) curves:
- κ vs $10^{-16}\kappa$
- κ vs $10^{-16}\kappa^2$
- κ vs relative error (for each of the three methods above)

Comment on what you observe about the plots. In particular, discuss how each method depends on κ and what the relative errors would be if we did everything in exact arithmetic

Problem 3. Suppose $\mathbf{A}$ has eigenvalue decomposition:

$$\mathbf{A} = \mathbf{V}\begin{bmatrix} -4 & & & \\ & -1 & & \\ & & 2 & \\ & & & 3 \end{bmatrix}\mathbf{V}^{-1}, \qquad \mathbf{V} = \begin{bmatrix} | & | & | & | \\ \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 & \mathbf{v}_4 \\ | & | & | & | \end{bmatrix}$$

where the $\mathbf{v}_i$ are all orthonormal.

Suppose we run inverse power method with shift c with $\mathbf{x} = \mathbf{v}_1 + \mathbf{v}_2 + \mathbf{v}_3 + \mathbf{v}_4$; that is, power method on $(\mathbf{A} - c\mathbf{I})^{-1}$.

If $c \in (0.5, 2.5)$, then we will converge to $\mathbf{v}_3$, the eigenvector corresponding to eigenvalue 2. The rate of converge is

$$\rho = \left|\frac{\lambda_2((\mathbf{A} - c\mathbf{I})^{-1})}{\lambda_1((\mathbf{A} - c\mathbf{I})^{-1})}\right|,$$

where $\lambda_1((\mathbf{A} - c\mathbf{I})^{-1})$ and $\lambda_2((\mathbf{A} - c\mathbf{I})^{-1})$ are the largest and second largest eigenvalues of $(\mathbf{A} - c\mathbf{I})^{-1}$ in magnitude respectively.

(a) Plot ρ as a function of c for c in the range $(0.5, 2.5)$.

(b) Let $\mathbf{y}_k$ be the output of k-steps of the power method, and assume $\|\mathbf{v}_3 - \mathbf{y}_k\|_2 \leq \rho^k$. For $\epsilon = 10^{-1}$, make a plot showing how large k has to be so that $\|\mathbf{v}_3 - \mathbf{y}_k\|_2 < \epsilon$ for the values of c in the range $(0.5, 2.5)$. Add more a new line to this plot for each $\epsilon = 10^{-2}, 10^{-5}, 10^{-10}$ plot. Label all the lines.

Problem 4. For the same matrix as in Problem 3, suppose we run power method with a starting vector:

$$\mathbf{x} = \mathbf{V}\begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}.$$

(a) Find a vector $\mathbf{z}$ so that $\mathbf{A}^k\mathbf{x} = \mathbf{V}\mathbf{z}$.

(b) What vector does $\mathbf{z}/\|\mathbf{z}\|$ converge to as $k \to \infty$?

(c) What vector does $\mathbf{A}^k\mathbf{x}/\|\mathbf{A}^k\mathbf{x}\|$ converge to as $k \to \infty$?

(d) Why did we get something different than on worksheet 8, where power method converged to a multiple of $\mathbf{v}_1$?