COL 726 Homework 4

28 February – 13 March, 2020

1. Let $\mathbf{A} \in \mathbb{C}^{m\times n}$, $m > n$ be a tall full-rank matrix with all singular values distinct. Consider the $(m+n)\times(m+n)$ Hermitian matrix $\mathbf{B} = \begin{bmatrix} \mathbf{0} & \mathbf{A} \\ \mathbf{A}^* & \mathbf{0} \end{bmatrix}$.

   (a) Prove that λ and $\begin{bmatrix} \mathbf{u} \\ \mathbf{v} \end{bmatrix}$ form an eigenpair of $\mathbf{B}$ with $\lambda > 0$ and $\|\mathbf{u}\|_2 = \|\mathbf{v}\|_2 = 1$ <u>if and only if</u> λ is a singular value of $\mathbf{A}$ with corresponding singular vectors $\mathbf{u}, \mathbf{v}$.

   (b) The above only accounts for n of the $m+n$ eigenpairs of $\mathbf{B}$. What are the remaining m eigenpairs? Give an explicit characterization in terms of the singular values and singular vectors of $\mathbf{A}$.

2. Let $\mathbf{A} \in \mathbb{R}^{m\times m}$ be a real square matrix. Suppose instead of the Rayleigh quotient, we estimate eigenvalues simply via the ratio of norms, $r_2(\mathbf{x}) = \|\mathbf{A}\mathbf{x}\|_2/\|\mathbf{x}\|_2$.

   (a) Assume $\mathbf{A}$ is symmetric, and let $\lambda, \mathbf{v}$ be an eigenpair with $\lambda \neq 0$. Do we still have $r_2(\mathbf{v}) = \lambda$ and $r_2(\mathbf{v}+\mathbf{x}) = r_2(\mathbf{v}) + O(\|\mathbf{x}\|^2)$? Give a proof or an explicit counterexample for each statement.

      <u>Hint</u>: You may find it useful to first prove that $\|\mathbf{v}+\mathbf{x}\|_2 = \|\mathbf{v}\|_2 + \mathbf{v}^T\mathbf{x}/\|\mathbf{v}\|_2 + O(\|\mathbf{x}\|^2)$.

   (b) Do the same for the case of nonsymmetric $\mathbf{A}$.

3. Let's investigate what happens if we perform inverse iteration with a not-very-accurate linear solver. Let $\mathbf{A} \in \mathbb{R}^{m\times m}$ be a real symmetric matrix, and let μ be an estimate of its jth eigenvalue, i.e. $|\mu - \lambda_j| < |\mu - \lambda_i|$ for all $i \neq j$. Suppose that at the kth iteration, we solve $(\mathbf{A}-\mu\mathbf{I})\mathbf{w} = \mathbf{v}^{(k)}$ using an iterative linear solver with a user-specified residual tolerance ϵ, that is, we find an approximate solution $\tilde{\mathbf{w}}$ such that $\|(\mathbf{A}-\mu\mathbf{I})\tilde{\mathbf{w}} - \mathbf{v}^{(k)}\|_2 \le \epsilon$. Let $\tilde{\mathbf{v}}^{(k+1)} = \tilde{\mathbf{w}}/\|\tilde{\mathbf{w}}\|$.

   Assume that $|\mathbf{q}_j^T\mathbf{v}^{(k)}| = c \gg \epsilon$. Find upper bounds on:

   (a) $(\|\tilde{\mathbf{w}}\|_2 - \|\mathbf{w}\|_2)/\|\mathbf{w}\|_2$,

   (b) $\|\tilde{\mathbf{v}}^{(k+1)} - \mathbf{v}^{(k+1)}\|$, and

   (c) $\|\tilde{\mathbf{v}}^{(k+1)} - (\pm\mathbf{q}_j)\|_2$ in terms of $\|\tilde{\mathbf{v}}^{(k)} - (\pm\mathbf{q}_j)\|_2$.

   <u>Hint:</u> It is easier to work in the eigenbasis, i.e. let $\mathbf{v}^{(k)} = \mathbf{Q}\mathbf{u}^{(k)}$, $\mathbf{w} = \mathbf{Q}\mathbf{x}$, etc. where $\mathbf{Q}$ is the matrix of eigenvectors.

4. Recall that to compute all the eigenvalues of a real symmetric matrix, we first reduce it to tridiagonal form $\mathbf{A}$, then apply the QR algorithm.

   (a) Show that each iteration of the pure QR algorithm preserves the tridiagonal form of $\mathbf{A}$.

   (b) I claim that each iteration of the pure QR algorithm can be performed in $O(n)$ flops when $\mathbf{A}$ is tridiagonal. Justify this claim by giving all the steps of the Householder algorithm for computing $\mathbf{Q}^{(k+1)}\mathbf{R}^{(k+1)} = \mathbf{A}^{(k)}$, and the matrix-matrix multiplication for computing $\mathbf{A}^{(k+1)} = \mathbf{R}^{(k+1)}\mathbf{Q}^{(k+1)}$, each taking $O(n)$ flops.

5. Suppose the Jacobian of a function $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^n$ is too hard to compute, but we know *a priori* a matrix $\mathbf{D} \in \mathbb{R}^{n\times n}$ that approximates the Jacobian in the neighbourhood of the desired root $\mathbf{x}_*$. We run Newton's method with this constant matrix in place of the Jacobian. Under what conditions on $\mathbf{D}$ is this method locally convergent, and what is the convergence rate? For what value(s) of $\mathbf{D}$, if any, will it be quadratically convergent? Justify your answers.

6. In the following, all matrices are assumed to be real and symmetric.

   For any matrix $\mathbf{M}$, define $\mathbf{w}(\mathbf{M}) = (\lambda_{\min}(\mathbf{M}), \lambda_{\max}(\mathbf{M}))$ to be the range of its eigenvalues, interpreted as a point in $\mathbb{R}^2$. Suppose we have a two-parameter family of matrices $\mathbf{F}(\mathbf{x}) = \mathbf{A}_0 + x_1\mathbf{A}_1 + x_2\mathbf{A}_2$, and we want to find a member of this family $\mathbf{F}(\mathbf{x})$ with specified $\mathbf{w}$.

   (a) Show that if $\mathbf{A}_0$ is positive definite and $\mathbf{A}_1$ is negative definite, then for any $y_1 < \lambda_{\min}(\mathbf{A}_0)$ there exists a real number $x_1 > 0$ such that $\lambda_{\min}(\mathbf{A}_0 + x_1\mathbf{A}_1) = y_1$. Derive an upper bound for x_1 in terms of y_1 and the eigenvalue ranges of $\mathbf{A}_0$ and $\mathbf{A}_1$.

   Hint: First prove that $\lambda_{\min}(\mathbf{M}) \leq \mathbf{v}^T\mathbf{M}\mathbf{v} \leq \lambda_{\max}(\mathbf{M})$ for any $\mathbf{M}$ and unit vector $\mathbf{v}$.

   Using the above bound for an initial bracket, implement a bisection search procedure `x1 = findRoot1(A0, A1, y1, etol)` that finds x_1 to within the tolerance $|x_1 - x_1^*| \leq e_{\text{tol}}$.

   (b) Prove that if a matrix $\mathbf{M}$ with distinct eigenvalues has an eigenpair $(\lambda, \mathbf{q})$ with $\|\mathbf{q}\|_2 = 1$, then the perturbed matrix $\mathbf{M} + \delta\mathbf{M}$ has an eigenvalue $\lambda + \mathbf{q}^T(\delta\mathbf{M})\mathbf{q}$ to first order. Consequently, state the Jacobian of the map $\mathbf{f} : \mathbb{R}^2 \to \mathbb{R}^2$ given by $\mathbf{f}(\mathbf{x}) = \mathbf{w}(\mathbf{F}(\mathbf{x}))$.

   (c) Implement a function `x = findRoot2(A0, A1, A2, y, x0, rtol, maxiter)` that uses Newton's method to solve $\mathbf{f}(\mathbf{x}) = \mathbf{y}$ starting from an initial guess $\mathbf{x}^{(0)}$. Terminate when either $\|\mathbf{f}(\mathbf{x}^{(k)}) - \mathbf{y}\|_2 \leq r_{\text{tol}}$ or the specified maximum number of iterations is reached.

   Use the built-in function `[scipy.linalg.]eig` to compute eigenvalues and eigenvectors, but of course do not use any built-in functions for solving nonlinear equations.

Try your code on the following 10×10 matrices:

$$\mathbf{A}_0 = \begin{bmatrix} 1 & & & \\ & 2 & & \\ & & \ddots & \\ & & & 10 \end{bmatrix}, \quad \mathbf{A}_1 = \begin{bmatrix} -2 & -1 & & \\ -1 & -2 & \ddots & \\ & \ddots & \ddots & -1 \\ & & -1 & -2 \end{bmatrix}, \quad \mathbf{A}_2 = \begin{bmatrix} \sin(1) & \sin(2) & \cdots & \sin(10) \\ \sin(2) & \sin(4) & \cdots & \sin(20) \\ \vdots & \vdots & \ddots & \vdots \\ \sin(10) & \sin(20) & \cdots & \sin(100) \end{bmatrix},$$

with right-hand sides $y_1 = 0$ and $y_2 = \lambda_{\max}(\mathbf{A}_0) = 10$.

For bisection, run your code with $e_{\text{tol}} = 10^{-6}$ and report a table of iteration number, bracket endpoints $[a, b]$, function values $[f(a), f(b)]$, and error bound $|b - a|$. For Newton's method, run it with $r_{\text{tol}} = 10^{-12}$ starting from $\mathbf{x}^{(0)} = \mathbf{0}$, and report the iteration number k, iterate $\mathbf{x}^{(k)}$, function value $\mathbf{f}(\mathbf{x}^{(k)})$, and residual norm $\|\mathbf{f}(\mathbf{x}^{(k)}) - \mathbf{y}\|_2$. Both tables should include iteration 0 (the initial bracket/guess), and should be included in your PDF and not simply output by the program.