Numerical analysis: Homework 4 Solutions

Instructor: Anil Damle
Due: April 12, 2024

Policies

You may discuss the homework problems freely with other students, but please refrain from looking at their code or writeups (or sharing your own). Ultimately, you must implement your own code and write up your own solution to be turned in. Your solution, including plots and requested output from your code should be typeset and submitted via the Gradescope as a pdf file. This file must be self contained for grading. Additionally, please submit any code written for the assignment as zip file to the separate Gradescope assignment for code.

Question 1:

Implement Newton's method for root finding. For each of the following compute a root of the function and illustrate the order of convergence (i.e., q if $|e_{k+1}| = \rho|e_k|^q$, where $e_k = x^* - x_k$) and, if linear, the rate (i.e., ρ if $|e_{k+1}| = \rho|e_k|$) exhibited by the method. Discuss if you observe what you expect.

(a) $f(x) = x^2$

Solution: The Newton iteration is:

$$x_k = \frac{1}{2}x_{k-1},$$

so the convergence is clearly linear with rate 1/2. This is expected since $x^* = 0$ is not a simple root; it has multiplicity 2. Your convergence plot should clealry illustrate this.

(b) $f(x) = \sin x + x^3$

Solution: The Newton iteration is

$$x_k = x_{k-1} - \frac{\sin(x_{k-1}) + x_{k-1}^3}{\cos(x_{k-1}) + 3x_{k-1}^2}.$$

There is a simple root for $x^* = 0$, and we can in fact show that the convergence is actually cubic close to this root. This follows from Taylor expanding the numerator and the denominator

$$\begin{aligned} x_k &\approx x_{k-1} - \frac{x_{k-1} - \frac{1}{6}x_{k-1}^3 + x_{k-1}^3}{1 - \frac{1}{2}x_{k-1}^2 + 3x_{k-1}^2} \\ &= x_{k-1} - \frac{x_{k-1} + \frac{5}{6}x_{k-1}^3}{1 + \frac{5}{2}x_{k-1}^2} \\ &\approx x_{k-1} - \left(x_{k-1} + \frac{5}{6}x_{k-1}^3\right)\left(1 - \frac{5}{2}x_{k-1}^2\right) \\ &\approx \frac{5}{3}x_{k-1}^3 \end{aligned}$$

so the quadratic term vanishes and the convergence is cubic. Your convergence plot should clealry illustrate this.

(c) $f(x) = \sin\frac{1}{x}$ for $x \neq 0$

Solution: This function has infinitely many roots of the form $1/(k\pi)$ for $k \in \mathbb{Z}$, so converging to a specific root is challenging. Starting to far away from zero makes either iteration try to converge to $\pm\infty$. All roots are simple, but it is a bit challenging to converge to a specific root, as the neighborhood around which convergence is guaranteed get smaller with increasing k. Newton converges at a quadratic rate and your convergence plot should clealry illustrate this.

Question 2:

Show that given any initial guess the Jacobi method for solving $Ax = b$converges for any strictly diagonally dominant matrix A (i.e., $|a_{ii}| > \sum_{j\neq i}|a_{ij}|$ for $i = 1, \ldots, n$).

You may find the following (restricted) version of the Gershgorin circle theorem useful for this question: Given a matrix A, each of its eigenvalues lies within at least one of the discs $\{z \in \mathbb{C} \mid |z - a_{ii}| < R_i\}$ for $i = 1, \ldots, n$, where $R_i = \sum_{j\neq i}|a_{ij}|$. In other words, all of the eigenvalues of A lie within the union of a set of discs centered around the diagonal entries of A, each of whose radius is the sum of the magnitudes of the off diagonal entries in that row.

Solution: If we decompose A into its diagonal, strictly upper triangular, and strictly lower triangular parts as $A = D + L + U$ have that the Jacboi iteration matrix is $G = -D^{-1}(L + U)$. Since the diagonal elements of G are zero we have that all eigenvalues of G lie within a disk of radius $\max_i \sum_{j\neq i}|g_{ij}|$.

For all i

$$\begin{aligned} \sum_{j\neq i}|g_{ij}| &= \sum_{j\neq i}\left|\frac{a_{ij}}{a_{ii}}\right| \\ &= |a_{ii}|^{-1}\sum_{j\neq i}|a_{ij}| \\ &< 1, \end{aligned}$$

where the final inequality follows from the strict diagonal dominance of A. Therefore, all of the eigenvalues of G have magnitude less that 1 and we are guaranteed convergence for any initial guess.

Question 3:

Prove that if $\nabla f(x) = 0$ but the Hessian $\nabla^2 f(x)$ is indefinite (i.e., has positive and negative eigenvalues), then there is a direction we can move from x that decreases the function value provided we take a small enough step. (I.e., show that x is not a local minimizer and that we can make progress when running an optimization scheme using so-called directions of negative curvature.)

Solution: A Taylor expansion of f yields

$$f(x+p) = f(x) + \frac{1}{2}p^T \nabla^2 f(x) p + \mathcal{O}(\|p\|^3).$$

If we let $p = \alpha v$ for some $\alpha > 0$ where v is an eigenvector of $\nabla^2 f(x)$ with eigenvalue $\lambda < 0$, then we have that

$$f(x + \alpha v) = f(x) + \alpha^2 \frac{\lambda}{2} + \mathcal{O}(\alpha^3).$$

Since $\lambda < 0$ we have that for all small enough α (where "small enough" depends on the constant in the $\mathcal{O}(\alpha^3) term$)

$$f(x + \alpha v) < f(x).$$

Question 4 (an interesting ungraded problem):

Suppose we have a function $f : \mathbb{R}^n \to \mathbb{R}$ with a local minimizer x^* such that for any direction $p \in \mathbb{R}^n$ (say with $\|p\|_2 = 1$) there exists an $\epsilon > 0$ such that $f(x^* + \alpha p) > f(x^*)$ for all $\alpha \in (-\epsilon, \epsilon)$. Does this guarantee that x^* is a strict local minimizer of $f(x)$? Why or why not?