Numerical analysis: Homework 4

Pratyush Sudhakar

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.

Solution: Newton's method:

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}$$

where x_k is the current approximation, $f(x_k)$ is the function value at x_k, and $f'(x_k)$ is the derivative of the function at x_k.

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

$$f(x) = x^2, \quad f'(x) = 2x$$

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

The root of $f(x) = x^2$ is $x^* = 0$.

Let $e_k = x^* - x_k = -x_k$. Then,

$$|e_{k+1}| = |x^* - x_{k+1}| = \left|0 - \frac{x_k}{2}\right| = \frac{|x_k|}{2} = \frac{|e_k|}{2}$$

Since $|e_{k+1}| = \rho|e_k|$ with $\rho = \frac{1}{2}$, the method exhibits linear convergence with rate $\rho = \frac{1}{2}$.

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

$$f(x) = \sin x + x^3, \quad f'(x) = \cos x + 3x^2$$

The root of $f(x) = \sin x + x^3$ can be found numerically. One root is $x^* \approx 0$. Using Taylor expansion around $x^* = 0$:

$$\sin x \approx x - \frac{x^3}{6} + O(x^5)$$

$$\cos x \approx 1 - \frac{x^2}{2} + O(x^4)$$

At each step of Newton's method:

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

Expanding $f(x_k)$ and $f'(x_k)$ around $x^* = 0$:

$$f(x_k) \approx x_k - \frac{x_k^3}{6} + x_k^3 = x_k + \frac{5}{6}x_k^3$$

$$f'(x_k) \approx 1 - \frac{x_k^2}{2} + 3x_k^2 = 1 + \frac{5}{2}x_k^2$$

Thus,

$$x_{k+1} \approx x_k - \frac{x_k + \frac{5}{6}x_k^3}{1 + \frac{5}{2}x_k^2}$$

Approximating the denominator using a Taylor expansion:

$$\frac{1}{1 + \frac{5}{2}x_k^2} \approx 1 - \frac{5}{2}x_k^2$$

Hence,

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

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

$$x_{k+1} \approx x_k - \left(x_k + \frac{5}{3}x_k^3\right)$$

$$x_{k+1} \approx x_k - x_k - \frac{5}{3}x_k^3$$

$$x_{k+1} \approx -\frac{5}{3}x_k^3$$

Thus, the quadratic term vanishes and the convergence is cubic. Therefore,

$$|e_{k+1}| \approx C|e_k|^3$$

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

$$f(x) = \sin\frac{1}{x}, \quad f'(x) = -\frac{\cos\frac{1}{x}}{x^2}$$

Convergence Analysis: The function $f(x) = \sin\frac{1}{x}$ has a root at $x \to \infty$, which is not practical for numerical methods. Newton's method may not converge due to the oscillatory and discontinuous nature of the function around the root.

Order of Convergence: Since the function is highly oscillatory and discontinuous at the root, the order of convergence is difficult to determine and is likely to be non-standard.

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 $i = 1, \ldots, n$).

Solution: Consider the linear system $Ax = b$, where A is a strictly diagonally dominant matrix. The Jacobi method for solving $Ax = b$ can be written in matrix form as:

$$x^{(k+1)} = D^{-1}(b - (L + U)x^{(k)})$$

where $A = D + L + U$, with D being the diagonal part of A, L the strictly lower triangular part, and U the strictly upper triangular part. Rewriting the Jacobi iteration:

$$x^{(k+1)} = D^{-1}(b - (L + U)x^{(k)}) = D^{-1}b - D^{-1}(L + U)x^{(k)}$$

Defining $T_J = -D^{-1}(L + U)$, the iteration can be expressed as:

$$x^{(k+1)} = T_J x^{(k)} + D^{-1}b$$

The Jacobi method converges if and only if the spectral radius $\rho(T_J)$ of the iteration matrix T_J is less than 1, i.e., $\rho(T_J) < 1$.
Next, we apply the Gershgorin circle theorem. For the matrix $T_J = -D^{-1}(L + U)$:

$$(T_J)_{ii} = 0 \quad \text{(since the diagonal elements are zero)}$$

$$(T_J)_{ij} = -\frac{a_{ij}}{a_{ii}} \quad \text{for } i \neq j$$

Each Gershgorin disc for T_J is centered at 0 with radius:

$$R_i = \sum_{j\neq i} |(T_J)_{ij}| = \sum_{j\neq i} \left| -\frac{a_{ij}}{a_{ii}} \right| = \frac{1}{|a_{ii}|} \sum_{j\neq i} |a_{ij}|$$

Since A is strictly diagonally dominant:

$$|a_{ii}| > \sum_{j\neq i} |a_{ij}|$$

Hence,

$$1 > \sum_{j\neq i} \left| \frac{a_{ij}}{a_{ii}} \right|$$

This implies:

$$\sum_{j\neq i} |(T_J)_{ij}| < 1$$

According to the Gershgorin circle theorem, all eigenvalues of T_J lie within the union of the Gershgorin discs, each centered at 0 with radius less than 1. Hence, all eigenvalues of T_J lie strictly within the unit circle in the complex plane:

$$\rho(T_J) < 1$$

Since the spectral radius of T_J is less than 1, the Jacobi method converges for any initial guess $x^{(0)}$. Hence, proved.

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: Given that $\nabla f(x) = 0$ and the Hessian $\nabla^2 f(x)$ is indefinite, we need to show that there exists a direction in which moving from x decreases the function value.
The second-order Taylor expansion of f around the point x is:

$$f(x+p) \approx f(x) + \nabla f(x)^T p + \frac{1}{2} p^T \nabla^2 f(x) p$$

Since $\nabla f(x) = 0$, this simplifies to:

$$f(x+p) \approx f(x) + \frac{1}{2} p^T \nabla^2 f(x) p$$

The Hessian $\nabla^2 f(x)$ being indefinite means it has both positive and negative eigenvalues. Therefore, there exists a direction p such that:

$$p^T \nabla^2 f(x) p < 0$$

To find such a direction p, consider the eigenvalues and eigenvectors of $\nabla^2 f(x)$. Since $\nabla^2 f(x)$ is indefinite, it has at least one negative eigenvalue. Let v be an eigenvector corresponding to a negative eigenvalue $\lambda < 0$ of $\nabla^2 f(x)$. Then:

$$\nabla^2 f(x) v = \lambda v$$

and

$$v^T \nabla^2 f(x) v = \lambda v^T v$$

Since $\lambda < 0$ and $v^T v > 0$ (as v is nonzero), we have:

$$v^T \nabla^2 f(x) v < 0$$

Now, let $p = \alpha v$ for some scalar $\alpha > 0$. Substituting $p = \alpha v$ into the Taylor expansion, we get:

$$f(x + \alpha v) \approx f(x) + \frac{1}{2} (\alpha v)^T \nabla^2 f(x) (\alpha v) = f(x) + \frac{1}{2} \alpha^2 v^T \nabla^2 f(x) v$$

Since $v^T \nabla^2 f(x) v < 0$, this becomes:

$$f(x + \alpha v) \approx f(x) + \frac{1}{2} \alpha^2 \lambda v^T v$$

Given that $\lambda < 0$ and $v^T v > 0$, the term $\frac{1}{2} \alpha^2 \lambda v^T v$ is negative. Therefore:

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

Thus, by choosing $p = \alpha v$ for a sufficiently small $\alpha > 0$, moving in the direction of the eigenvector v corresponding to a negative eigenvalue of the Hessian $\nabla^2 f(x)$ will decrease the function value. This implies that x is not a local minimizer, as there exists a direction p in which the function value decreases.

Hence, we can make progress in an optimization scheme by moving in directions of negative curvature.