

MA5630

Problem Set 2

Due 19 February 2003

Instructions: Do Alternative I or Alternative II.

1 Alternative I

1. Write a program that performs Broyden's method for solving $F(x) = 0$, where $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$. The program should accept as input
 - a starting vector $x^{(0)}$;
 - a function that, given n and x , computes $F(x)$ and J , the Jacobian of F at x ;
 - an error tolerance ϵ ;
 - an iteration limit N .

The program should use the exact Jacobian $J(x^{(0)})$ as A_0 , and then update using Broyden's formula to produce Jacobian approximations A_1, A_2, \dots . The program should stop with $x^{(i)}$ satisfying either

$$\|x^{(i)} - x^{(i-1)}\| \leq \epsilon$$

or $i = N$.

Be sure to test your code on simple problems before you go on to the next part.

2. Write a similar program implementing Newton's method. See the directions for Problem Set 1 for details. (If you already did this for Problem Set 1, you can use that code.)
3. Obtain a copy of the paper *Testing Unconstrained Optimization Software* by Moré, Garbow, and Hillstom (ACM TRANSACTIONS ON MATHEMATICAL SOFTWARE, Vol. 7, No. 1, 1981, 17–41). Choose five test problems (for nonlinear equations) for which the exact solution is known and compare the performance of Newton's method and Broyden's method. Write a short report discussing the efficiency of the two methods for these test problems. (You might ask such questions as "Which method can produce a low-accuracy (high-accuracy) solution more efficiently?" There are two ways to compare efficiency: total execution time and number of iterations. Be sure to discuss both measures and the appropriateness of each. Ask me if you need suggestions.)

2 Alternative 2

1. Suppose $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is continuously differentiable.

(a) Prove that if f is convex, then

$$(\nabla f(x) - \nabla f(y)) \cdot (x - y) \geq 0 \text{ for all } x, y \in \mathbb{R}^n.$$

Also prove that the inequality is strict if f is strictly convex.

(b) Prove or disprove: If

$$(\nabla f(x) - \nabla f(y)) \cdot (x - y) \geq 0 \text{ for all } x, y \in \mathbb{R}^n,$$

then f is convex.

(c) Suppose $x, y \in \mathbb{R}^n$ satisfy

$$(\nabla f(x) - \nabla f(y)) \cdot (x - y) < 0.$$

Prove that $\phi : [0, 1] \rightarrow \mathbb{R}$ defined by

$$\phi(\alpha) = f(x + \alpha(y - x))$$

is not convex.

2. (a) Prove that if $H \in \mathbb{R}^n$ is symmetric and positive definite, then so is $H + uu^T$ for any $u \in \mathbb{R}^n$.

(b) Suppose $H \in \mathbb{R}^{n \times n}$ is a symmetric positive definite matrix and $s, y \in \mathbb{R}^n$ are given vectors satisfying $s \cdot y > 0$. Show that it may not be possible to find a vector u such that

$$(H + uu^T)s = y.$$

(c) Is it possible to find $u \in \mathbb{R}^n$ and $\sigma \in \mathbb{R}$ such that $(H + \sigma uu^T)s = y$ holds if $H + \sigma uu^T$ is not required to be positive definite?

3. Suppose $A \in \mathbb{R}^{n \times n}$ is nonsingular and $\|\cdot\|$ is the operator norm on $\mathbb{R}^{n \times n}$. Prove that there exists $\epsilon > 0$ such that

$$\|A - B\| < \epsilon \Rightarrow B^{-1} \text{ exists.}$$

Hints:

(a) First prove that if $E \in \mathbb{R}^{n \times n}$ and $\|E\| < 1$, then $I - E$ is nonsingular and

$$(I - E)^{-1} = \sum_{n=0}^{\infty} E^n. \tag{1}$$

(b) Then notice that

$$B = A - (A - B) = A^{-1} (I - A^{-1}(A - B)),$$

and so B is invertible if and only if $I - A^{-1}(A - B)$ is invertible. Take $E = A^{-1}(A - B)$ in the previous part.

4. Suppose that $A : \mathbb{R}^n \rightarrow \mathbb{R}^{n \times n}$ is a matrix-valued function that is continuous. Assume that, for a certain $x_0 \in \mathbb{R}^n$, $A(x_0)$ is nonsingular. According to the previous exercise, the mapping

$$x \mapsto A(x)^{-1}$$

is well-defined for all x near x_0 . Prove that this mapping is continuous at x_0 . Hints:

(a) Using (1), prove that

$$\|E\| < 1 \Rightarrow \|(I - E)^{-1}\| \leq \frac{1}{1 - \|E\|}.$$

(b) Now take $E = A^{-1}(A - B)$ and use the previous inequality.

5. Explain how the results of the previous two problems are needed in the proof of quadratic convergence of Newton's method.