Math 270B: Numerical Approximation and Nonlinear Equations

Instructor: Michael Holst

Winter Quarter 2011

Homework Assignment #1
Due : Give to the class TA within two weeks if you would like it marked.

Our goal in this homework is to review some basic concepts from linear algebra and from calculus of real and multivariate functions, and to then use these ideas to develop techniques for solving systems of nonlinear equations. From linear algebra we will need matrices, eigenvalues and eigenvectors, singular values, and related concepts. From calculus, we will need to recall the definitions of continuity and differentiation, in the cases of both real-valued and vector-valued functions of one and many variables. Key tools throughout the course will be Taylor series and the Taylor Remainder Theorem.

The starred exercises are those that require the use of a tool such as Matlab.

Exercise 1.1. If x is an eigenvector of A, show that βx is also an eigenvector for any $\beta \neq 0$. What is the associated eigenvalue? Use this result to show that the unit vector $x/\|x\|$ formed from an eigenvector x is also an eigenvector of A corresponding to the same eigenvalue as that of x.

Exercise 1.2. Let $(x, y) : V \mapsto \mathbb{R}$ be an inner-product on a vector space V with associated scalar field $\mathbb{R}$. We know that (x, y) must satisfy the three properties of an inner-product:

1. $(x, x) \geq 0, \quad \forall x \in V, \quad (x, x) = 0 \text{iff} x = 0.$
2. $(x, y) = (y, x), \quad \forall x, y \in V.$
3. $(\alpha x + \beta y, z) = \alpha(x, z) + \beta(y, z), \quad \forall \alpha, \beta \in \mathbb{R}, \quad \forall x, y, z \in V.$

Use these three properties to show that the induced norm $\|x\| = (x, x)^{1/2}$ satisfies the three properties of a norm:

1. $\|\alpha x\| = |\alpha| \; \|x\|, \quad \forall \alpha \in \mathbb{R}, \quad \forall x \in V.$
2. $\|x\| \geq 0, \quad \forall x \in V, \quad \|x\| = 0 \text{iff} x = 0.$
3. $\|x + y\| \leq \|x\| + \|y\|, \quad \forall x, y \in V.$

Hint: Showing the first two properties is very easy; to show the last property (triangle inequality), assume the Cauchy-Schwarz inequality holds: $|(x, y)| \leq \|x\| \; \|y\|$.

Exercise 1.3. Let $F(x)$ denote a twice-differentiable function of one variable. Assuming only the mean-value theorem of integral calculus: $F(b) = F(a) + \int_a^b F'(t)\, dt$, derive the following variants of the Taylor-series expansion with integral remainder:

(a) $F(x + h) = F(x) + \int_x^{x+h} F'(t)\, dt.$

(b) $F(x + h) = F(x) + h \int_0^1 F'(x + \xi h)\, d\xi.$

(c) $F(x + h) = F(x) + hF'(x) + h \int_0^1 [F'(x + \xi h) - F'(x)]\, d\xi.$

(d) $F(x + h) = F(x) + hF'(x) + h^2 \int_0^1 F''(x + \xi h)(1 - \xi)\, d\xi.$ (Hint: Try expanding $F'(x + h)$ using a formula like part (b) and then differentiate with respect to h using the chain rule.)

Exercise 1.4. Let $F : \mathbb{R}^n \to \mathbb{R}^n$ be a differentiable function. Using the multi-dimensional verson of the fundamental theorem of calculus, show (i.e. derive) that the following Taylor expansion holds with the given integral remainder:

$$F(x + h) = F(x) + F'(x)h + \int_0^1 \{F'(x + \xi h) - F'(x)\}\, h \; d\xi.$$

Note that this is the multi-dimensional analogue of what you showed in Exercise 3(c). While this particular form of the remainder will be the most useful to us in the case of $\mathbb{R}^n$, derive the multi-dimensional analogues of the other three forms of the remainder you derived in Exercise 3.

Exercise 1.5. Find the gradient vector $F(x) = \nabla f(x)$ of the following functions, and then find the Jacobian matrix of $F(x)$. (The Jacobian matrix of $F(x) = \nabla f(x)$ is the same as the Hessian matrix $\nabla^2 f(x)$ of $f(x)$).

(a) $f(x) = 2\big(x_2 - x_1^2\big)^2 + \big(x_1 - 3\big)^2$.

(b) $f(x) = (2x_1 + x_2)^2 + 4(x_2 - x_3)^4$.

Exercise 1.6. Find $f'(x)$, $\nabla f(x)$ and $\nabla^2 f(x)$ for the following functions of n variables.

(a) $f(x) = \frac{1}{2}x^T H x$, where H is an $n \times n$ constant matrix.

(b) $f(x) = b^T A x - \frac{1}{2}x^T A^T A x$, where A is an $m \times n$ constant matrix and b is a constant m-vector.

(c) $f(x) = \|x\| = \left(\sum_{i=1}^n x_i^2\right)^{1/2}$.

Exercise 1.7.* Create a MATLAB m-file of the form:

```
function [F,J] = D(x)
  F = [ a ; b ];
  J = [ c d ; e f ];
```

where the expressions for a, b, c, d, e, f are chosen so that the function returns the 2×1-vector-valued function $F(x)$ and the 2×2 Jacobian matrix $J(x)$ for the function $F(x)$ from part (a) of Problem 1.5. Use this m-file to compute F and J at $x = (1,\, 0)^T$; and $x = (1,\, 1)^T$. Capture the output from the computation and turn it in with the homework.

Exercise 1.8. Give (precisely) the following definitions:

1. A direction of decrease for a function $f(x) : \mathbb{R}^n \mapsto \mathbb{R}$.
2. A descent direction for a function $f(x) : \mathbb{R}^n \mapsto \mathbb{R}$.
3. Lipschitz continuity of a function $F(x) : \mathbb{R}^n \mapsto \mathbb{R}^m$.

Exercise 1.9. Prove that a descent direction is always a direction of decrease.

Exercise 1.10. Calculate the following derivatives (show your work):

1. The derivative $f'(x)$, the gradient $\nabla f(x)$, and the Hessian matrix $\nabla^2 f(x)$ of the following real-valued function of three real variables:
$$f(x) = x_1^3 - 2x_2^2 + x_3$$
2. The jacobian matrix $F'(x)$ of the following function which maps $\mathbb{R}^n$ into $\mathbb{R}^m$:
$$F(x) = Ax - b, \quad \text{where}: A \in \mathbb{R}^{m \times n},\ x \in \mathbb{R}^n,\ b \in \mathbb{R}^m.$$

Exercise 1.11. Using a one-dimensional Taylor series, derive Newton's method for solving $f(x) = 0$, where $f : \mathbb{R} \mapsto \mathbb{R}$. Draw a detailed picture for the case of $f(x) = x^2 - 9$, illustrating how the iterates behave, beginning with $x_0 = 10$. (The picture doesn't have to be pretty, but it must be logically correct and labeled correctly.)

Exercise 1.12. Let $F(x) : D \subset \mathbb{R}^n \mapsto \mathbb{R}^n$, and assume that the jacobian $F'(x) : D \subset \mathbb{R}^n \mapsto \mathbb{R}^{n \times n}$ is Lipschitz continuous with Lipschitz constant γ. Show that the error in the linear model

$$L_k(x) = F(x^k) + F'(x^k)(x - x^k)$$

of $F(x)$ can be bounded as follows:

$$\|F(x) - L_k(x)\| \leq \frac{1}{2}\gamma\|x - x^k\|^2.$$

Hint: I did this in class; just use the (multi-dimensional) Taylor formula you derived in Problems 3 and 4.

Exercise 1.13.* Write a MATLAB m-file implementing Newton's method, using the MATLAB routines you wrote for problem 7. Apply Newton's method to part (a) of Problem 1.5 and collect the iteration information, using first the initial guess of $x = (1, 0)^T$, and then using the initial guess of $x = (1, 1)^T$.

Exercise 1.14. State and prove a basic convergence theorem for Newton's method for solving $F(x) = 0$, where $F\colon \mathbb{R}^n \to \mathbb{R}^n$. (See the notes from class regarding the minimal assumptions you will need to use.) First show that Newton's method converges superlinearly without assuming that $F'(x)$ is Lipschitz, and then use the result in Exercise 12 to show that Newton's method converges quadratically when $F'(x)$ is Lipschitz.