The University of British Columbia

CPSC 406

Computational Optimization

Homework #1

Student name:

ALI SIAHKOOHI

Student ID:

620192032

Instructor:

Professor Ascher

Date Submitted:

September 22nd, 2016

Question #1,
Part (a)

Function $f(x)$ is defined as below:

$$f(x) = \alpha x^2 + \beta x + \gamma, \tag{1 - 1}$$

1. Since $f(\underline{x})$ is twice differentiable, In order to be a convex function:

$$f''(x) \geq 0 \quad \Rightarrow \quad \alpha \geq 0. \tag{1 – 2}$$

2. In order to be a concave function:

$$f''(x) \leq 0 \quad \Rightarrow \quad \alpha \leq 0. \tag{1 – 3}$$

3. In order to be both convex and concave function, I visited the following webpage to check out the definition of both concave and convex function.

Mathematical methods for economic theory
Martin J. Osborne
URL: http://mjo.osborne.economics.utoronto.ca/index.php/tutorial/index/1/cv1/t

If every line segment joining two points on the function is never above the graph, or never below the graph, functions are concave and convex, repetitively. So the only function that is both convex and concave is affine function. Affine function is defined as below:

$$f(\underline{x}) = a\underline{x} + b, \quad a, b \in \mathbb{R}, \quad \underline{x} \in \mathbb{R}^n. \tag{1 – 4}$$

So the function is both convex and concave if:

$$\alpha = 0 \tag{1 – 5}$$

It's worth mentioning that, in number one and two, if we want the function to be strictly convex or concave, we need to have strict inequalities.

4. There is no $\alpha, \beta, \gamma \in \mathbb{R}$ such that the function is neither convex nor concave. For a polynomial to be neither convex nor concave, it's necessary to be at least third-degree polynomial.

Part (b)

i. In the following equations, it has been shown that f(x) is always less than or equal to zero. So it is a <u>concave</u> function.

$$f(x) = -\sqrt{1+x^4} \quad \Rightarrow \quad f'(x) = \frac{-1}{2}(4x^3)(1+x^4)^{\frac{-1}{2}} = -2x^3(1+x^4)^{\frac{-1}{2}},$$

$$\Rightarrow \quad f''(x) = -2\left(3x^2(1+x^4)^{\frac{-1}{2}} + x^3(\frac{-1}{2})(4x^3)(1+x^4)^{\frac{-3}{2}}\right) = -2\left(\frac{3x^2(1+x^4)}{(1+x^4)^{\frac{3}{2}}} - \frac{2x^6}{(1+x^4)^{\frac{3}{2}}}\right),$$

$$\Rightarrow \quad f''(x) = \frac{-2x^2(3+x^4)}{(1+x^4)^{\frac{3}{2}}} \le 0, \quad x \in \mathbb{R}. \qquad (1-6)$$

ii. In this part, f(x) is a linear (affine) function. Regarding to this, f(x) is <u>both convex and concave</u> on the real line.

iii. For this function, we need to find out the sign of its second derivative in order to found out what type of function it is.

$$f(x) = \frac{2x}{1+x^3} \Rightarrow f'(x) = 2\left(\frac{1}{1+x^3} - \frac{x(3x^2)}{(1+x^3)^2}\right) = 2\left(\frac{1-2x^3}{(1+x^3)^2}\right),$$

$$\Rightarrow f''(x) = 2\left(\frac{-6x^2}{(1+x^3)^2} + \frac{-2(3x^2)(1-2x^3)}{(1+x^3)^3}\right) = 12\frac{x^2(x^3-2)}{(1+x^3)^3}, \qquad (1-7)$$

$$\Rightarrow \begin{cases} f''(x) > 0, & x < -1 \\ f''(x) \le 0, & -1 < x \le \sqrt[3]{2} \\ f''(x) \ge 0, & x \ge \sqrt[3]{2} \end{cases} \qquad (1-8)$$

As it can be interpreted from equation (1 – 8), the function is convex in two intervals, and concave in between. So f(x) is <u>neither convex nor concave</u>.

Part (c)

Let f be a function of many variables with continuous partial derivatives and cross partial derivatives on the convex open set S and denote the Hessian of f at the point x by H(x). Then:

a. f is concave if and only if H(x) is negative semi definite for all $x \in S$
b. f is convex if and only if H(x) is positive semi definite for all $x \in S$

Source: Sydsaeter, Knut and Hammon, Peter J., Mathematics for economic analysis, Prentice Hall, Upper Saddle River, New Jersey, 1995

$$f(x_1, x_2) = 2x_1^2 - 3x_1x_2 + 5x_2^2 + 17x_1 - 17x_2, \quad x \in \mathbb{R}^2,$$

$$\frac{\partial f}{\partial x_1} = 4x_1 - 3x_2 + 17, \quad \frac{\partial f}{\partial x_2} = -3x_1 + 10x_2 - 17,$$

$$\frac{\partial^2 f}{\partial x_1^2} = 4, \quad \frac{\partial^2 f}{\partial x_1 \partial x_2} = -3, \quad \frac{\partial^2 f}{\partial x_2 \partial x_1} = -3, \quad \frac{\partial^2 f}{\partial x_2^2} = 10,$$

$$\Rightarrow H(x) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} \end{bmatrix} = \begin{bmatrix} 4 & -3 \\ -3 & 10 \end{bmatrix},$$

$$\Rightarrow H(x) = USV^T = \begin{bmatrix} -0.3827 & 0.9239 \\ 0.9239 & 0.3827 \end{bmatrix} \begin{bmatrix} 11.2426 & 0 \\ 0 & 2.7574 \end{bmatrix} \begin{bmatrix} -0.3827 & 0.9239 \\ 0.9239 & 0.3827 \end{bmatrix}^T, \quad (1-9)$$

Since the eigenvalues of the Hessian are both positive, the Hessian matrix is positive definite. According to the above mentioned characterizations, the function $f(x_1, x_2)$ is a <u>convex</u> function for $x \in \mathbb{R}^2$.

Question #2

Part (a)

We are going prove that the given function is <u>convex.</u> It is worth mentioning that, the sum of two convex functions is a convex function on the intersection of domains of the two functions. Suppose:

$$f(x) = f_1(x) + f_2(x), \text{ which } f_1(x) = \frac{1}{2}(x-b)^2, \quad f_2(x) = \lambda|x|. \quad (2-1)$$

First, by computing the second derivative of $f_1(x)$ we are going to show that this is a convex function:

$$f_1(x) = \frac{1}{2}(x-b)^2 \Rightarrow f_1''(x) = 1, \quad (2-2)$$

This means that $f_1(x)$ is a convex function over real line. For $f_2(x)$ we are going to use the following statement:

A differentiable function $f:\mathbb{R}\to\mathbb{R}$ is convex, if and only if:

$$f(y) \ge f(x) + f'(x)(y - x). \qquad (2-3)$$

Source: Boyd S, Vandenberghe L., Convex optimization, Cambridge university press; 2004 March 8.

We are going to start proving $f_2(x)$ is convex by computing its first derivative. In fact we know that absolute function is differentiable everywhere on real line, except x=0:

$$f_2(x) = \lambda|x| \Rightarrow \frac{df_2(x)}{dx} = \lambda\frac{|x|}{x}, \quad \text{for } x \in \mathbb{R} - \{0\}. \qquad (2-4)$$

Now we should verify that $f_2(x)$ and its first derivative satisfy equation (2 – 3):

$$f_2(y) \ge f_2(x) + f_2'(x)(y - x), \quad \text{which means:}$$

$$\lambda|y| \ge \lambda|x| + \lambda(y - x)\frac{|x|}{x} \Rightarrow |y| - |x| \ge (y - x)\frac{|x|}{x}$$

$$\text{If } x > 0: \quad \Rightarrow |y| - x \ge y - x \Rightarrow |y| \ge y \ \text{ which is trivial!} \qquad (2-5)$$

$$\text{If } x < 0: \quad \Rightarrow |y| + x \ge x - y \Rightarrow |y| \ge -y \ \text{ which is also trivial for all } y \in \mathbb{R}. \qquad (2-6)$$

But since we have not considered the zero point, I am going to add the proof below in order to be more accurate, in order to prove absolute function is convex, we need to show the following:

$$|\alpha x + \beta y| \le |\alpha x| + |\beta y|, \quad \text{for } \alpha+\beta=1 \qquad (2-7)$$

i) $\alpha x \ge 0,\ \beta x \le 0:\ |\alpha x + \beta y| = \alpha x + \beta y = \alpha x + \beta y = |\alpha x| + |\beta y|$

ii) $\alpha x \ge 0,\ \beta x \le 0:\ |\alpha x + \beta y| = |\alpha x - |\beta y|| \le \alpha x - \beta y = |\alpha x| + |\beta y|$

iii) $\alpha x \le 0,\ \beta x \ge 0:\ |\alpha x + \beta y| = |-|\alpha x| + \beta y| \le -\alpha x + \beta y = |\alpha x| + |\beta y|$

iv) $\alpha x \ge 0,\ \beta x \le 0:\ |\alpha x + \beta y| = -(\alpha x + \beta y) = -\alpha x - \beta y = |\alpha x| + |\beta y|$

So $f(x) = f_1(x) + f_2(x)$ is <u>convex</u> for all $x \in \mathbb{R}$.

Part (b)

Since the function is convex, the only thing needed is to find where the first derivative of function hits the x-axis! Since the function is not derivative at x = 0, In addition to the zero first derivation point, there is a possibility that x = 0, could be the optimal point.

$$f(x) = \frac{1}{2}(x-b)^2 + \lambda|x|,$$

$$\frac{df(x)}{dx} = x - b + \lambda\frac{|x|}{x} = 0,$$

$$\Rightarrow x^*(b,\lambda) = \begin{cases} b - \lambda\frac{|b|}{b}, & b > \sqrt{3}\lambda,\ b\in\mathbb{R},\ \lambda\geq 0 \\ 0, & b \leq \sqrt{3}\lambda,\ b\in\mathbb{R},\ \lambda\geq 0 \end{cases} \qquad (2-7)$$

#Question 3

We want to prove that for the one variable function and given assumptions, there is a constant M, such that, for k large enough:

$$|x^* - x_{k+1}| \leq M|x^* - x_k|^2. \qquad (3-1)$$

Starting from the Taylor's series expansion of function r(x) at $x = x_k$:

$$0 = r(x^*) = r\left(x_k + (x^* - x_k)\right) = r(x_k) + (x^* - x_k)r'(\xi), \qquad (3-2)$$

Which ξ is between x^* and x_k. By adding a term to both sides of the equation:

$$\Rightarrow r(x_k) + (x^* - x_k)r'(x_k) = (x^* - x_k)\left(r'(x_k) - r'(\xi)\right),$$
$$\Rightarrow |r(x_k) + (x^* - x_k)r'(x_k)| = |x^* - x_k||r'(x_k) - r'(\xi)|, \qquad (3-3)$$

Since ξ is between x^* and x_k:

$$|x_k - \xi| \leq |x^* - x_k|, \qquad (3-4)$$

And considering Lipschitz continuously assumption:

$$|r'(x_k) - r'(\xi)| \leq \gamma|x_k - \xi|, \qquad (3-5)$$

$$\Rightarrow |r(x_k) + (x^* - x_k)r'(x_k)| \leq \gamma|x^* - x_k||x_k - \xi| \leq \gamma|x^* - x_k||x^* - x_k|,$$
$$\Rightarrow |r(x_k) + (x^* - x_k)r'(x_k)| \leq \gamma|x^* - x_k|^2. \qquad (3-6)$$

Now, using Newton's method relation:

$$x_{k+1} = x_k - \frac{r(x_k)}{r'(x_x)}, \qquad (3-7)$$

$$\Rightarrow \left|x^* - x_{k+1}\right| = x^* - x_k + \frac{r(x_k)}{r'(x_k)} = \left|\frac{1}{r'(x_k)}\right|\left(r(x_k) + (x^* - x_k)r'(x_k)\right), \qquad (3-8)$$

Considering the fact that $r'(x_k) \neq 0$ and using the continuity assumptions of the function, one can say that there is an open set (a ball) around x^* which:

$$\left|\frac{1}{r'(x_k)}\right| \leq \beta, \qquad (3-9)$$

Using the fact above, equation (3 – 8), and (3 – 6), we can conclude that:

$$\Rightarrow \left|x^* - x_{k+1}\right| \leq \left|\frac{1}{r'(x_k)}\right|\left|r(x_k) + (x^* - x_k)r'(x_k)\right| \leq \beta\gamma\left|x^* - x_k\right|^2,$$

$$\Rightarrow \left|x^* - x_{k+1}\right| \leq \beta\gamma\left|x^* - x_k\right|^2. \qquad (3-10)$$

#Question 4

I have implemented a function which first, it performs a Gaussian elimination process on the sparse matrix A, and second, it solves the inverse problem by back substituting method.

I have not used MATLAB’s built in functions, plus I have defined the matrix A as a sparse matrix. In the Gaussian elimination and back substituting process, the function only works with the non-zero values of A and only performs operations on them.

There is also another Mfile which generates matrix A, assumes an initial value for x, finds the Ax=b and tries to use block_trid.m function to solve to for x again.

MATLAB code used is as follow:

The Function:

```
function [x] = block_trid(A, b, m, n)

%Gaussian Elimination
x = zeros(size(b));

for i = 1:n-1
    for k = 1:m
```

```
        b((i-1)*m + k) = b((i-1)*m + k) / A((i-1)*m + k, ((i-1)*m + k));

        A((i-1)*m + k, ((i-1)*m + k):((i-1)*m + 2*m)) = ...
            A((i-1)*m + k, ((i-1)*m + k):((i-1)*m + 2*m)) / ...
            A((i-1)*m + k, ((i-1)*m + k));

        for j = (k+1):2*m

            b((i-1)*m + j) = b((i-1)*m + j) - A((i-1)*m + j, ((i-1)*m + k)) *
...
                b((i-1)*m + k);

            A((i-1)*m + j, ((i-1)*m + k):((i-1)*m + 2*m)) = ...
                A((i-1)*m + j, ((i-1)*m + k):((i-1)*m + 2*m)) - ...
                A((i-1)*m + j, ((i-1)*m + k)) * A((i-1)*m + k, ((i-1)*m +
k):((i-1)*m + 2*m));

        end
    end
end


i = n;
for k = 1:m-1

    b((i-1)*m + k) = b((i-1)*m + k) / A((i-1)*m + k, ((i-1)*m + k));

    A((i-1)*m + k, ((i-1)*m + k):((i-1)*m + m)) = ...
        A((i-1)*m + k, ((i-1)*m + k):((i-1)*m + m)) / ...
        A((i-1)*m + k, ((i-1)*m + k));

    for j = (k+1):m

        b((i-1)*m + j) = b((i-1)*m + j) - A((i-1)*m + j, ((i-1)*m + k)) * ...
                b((i-1)*m + k);

        A((i-1)*m + j, ((i-1)*m + k):((i-1)*m + m)) = ...
            A((i-1)*m + j, ((i-1)*m + k):((i-1)*m + m)) - ...
            A((i-1)*m + j, ((i-1)*m + k)) * ...
            A((i-1)*m + k, ((i-1)*m + k):((i-1)*m + m));
    end
end

b(m*n) = b(n*m) / A(m*n, m*n);
A(m*n, m*n) = 1;


% Backward substitution

x(n*m) = b(n*m);

for j = (m-1):-1:1
    x((n-1)*m+ j) = b((n-1)*m+ j) - ...
        sum(A((n-1)*m+ j, (n-1)*m+ ((j+1):m))'.*x((n-1)*m+ ((j+1):m)));
end
```

```
for i = (n-1):-1:1
    for j = (m):-1:1
    x((i-1)*m+ j) = b((i-1)*m+ j) - ...
        sum(A((i-1)*m+ j, ((i-1)*m+ j+1):((i-1)*m+ 2*m))'.* ...
        x(((i-1)*m+ j+1):((i-1)*m+ 2*m)));
    end
end


end
```

The script which actually generates a sparse matrix and runs block_trid.m

```
clear; clc; close all;

n = 10000;
m = 3;

A = sparse(m*n, m*n);

mblock = [1, -.1, -.1; -.1, 1, -.1; -.1, -.1 ,1];
lblock = [.1, .5, .5; -.5, .1, .5; -.5, -.5, .1];
ublock = [.1, -.5, -.5; .5, .1, -.5; .5, .5, 1];

for k = 1:n
    A((k-1)*m+1:k*m, (k-1)*m+1:k*m) = mblock;
    if k > 1, A((k-1)*m+1:k*m, (k-2)*m+1:(k-1)*m) = lblock; end
    if k < n , A((k-1)*m+1:k*m, k*m+1:(k+1)*m) = ublock; end
end
A = A*10;
x0 = zeros(m*n, 1);
x0(1) = 1;
b = A*x0;

tic
[x] = block_trid(A, b, m, n);
toc
fprintf('Norm of error =  %g \n', norm(A*x - b))
```