Introduction to Optimization 236330

**Homework 2**

Pavel Rastopchin 321082026 Uri Kirstein 311137095

Question 1

Let a,b be two points on the surface C. since C is a convex set, the line that connects the two points lies entirely within C, and f1 and f2 are convex on every point of that line. We will show that g(x) is convex by definition. Let , then we need to show that

Let’s take a look at the two expressions. Since f1(x) is convex in this line:

From symmetry this is true for the other expression inside the maximum. Hence we get:

So g is convex in C.

Question 2

Let . It implies that . We will show that any point on the line between points a and b is within the set L. For every choice of , the point . Let’s take a look at the value of the function f in such a point.

Because f is convex:

Question 3 –

Let y be a vector of the appropriate size. We will show the positive semidefiniteness of the hessian directly.

Ay is a vector, and since hessian of f is positive semidefinite:

In conclusion, g has a positive semidefinite hessian which means g is convex.

Question 4 –

Jensen’s inequality for the discrete case is:

For a real convex function , n numbers in ’s domain and a series of coefficients such that and , then .

**Lemma:** For n numbers , where C is a convex set, and a series of coefficients such that and , then

Proof: by induction.

Base: For n=2, we get that . From C’s convexity, .

Assumption: For n=k, .

Step: For n=k+1,

Looking at , all coefficients are positive and sum to 1, all hence from the assumption, . We will note . We see that which belongs to C from C’s convexity.

Q.E.D

We will prove it by induction on the number of points and alpha variables.

*Base:*

We have in ’s domain. Since is convex, by definition we know that

Induction’s assumption:

For n numbers in ’s domain and a series of coefficients such that and , then .

Induction step:

Looking at n+1 numbers in ’s domain and a series of coefficients such that and .

From the lemma, , hence from convexity of we obtain:

The sum of coefficients of is 1, and so from the assumption we get that

Q.E.D

Question 5-

We know that For a real convex function , n numbers in ’s domain and a series of coefficients such that and , then .

We choose to be –log, a convex function. We choose . It satisfies the conditions on the coefficients. –log’s domain is the real plain. For , using Jensen’s Inequality:

By exponentiation of each side of the inequality:

Q.E.D

Question 6 –

1. We have seen in class that and that . By substituting we get
2. We want to minimize . We know that f is concave, so if we find a local inimum it is also a global minimum. We will find it analytically.

We require equality to zero:

If , we would have stopped at the kth step, as this expression is just the gradient of the function f at the point . Hence, we require

This is assuming the matrices are invertible.

Question 7 –

Using Taylor multivariate theorem, there exists a z such that:

Since :

Looking at the left hand side, following Hint 2, we will minimize it in regards to y.

Substituting back to the original inequality we get:

We can do the same for the right hand side and obtain

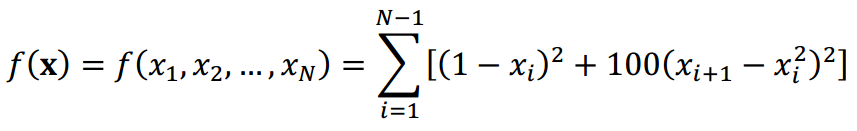
By hint 3, we can combine the two inequalities to

Q.E.D

**Question 8:**

1. **Derive the analytical expressions of gradient and Hessian of Rosenbrock function.**

Rosenbrock function:



i.e:

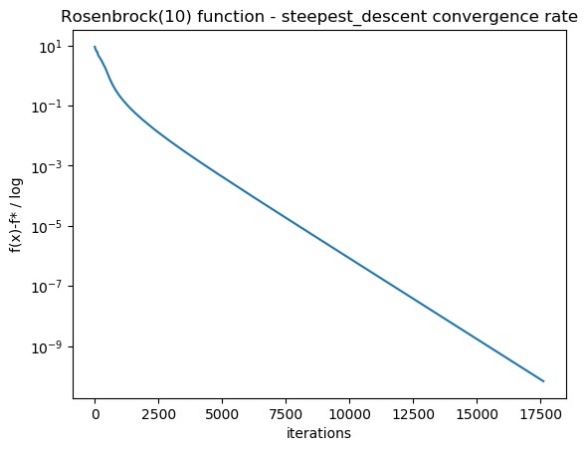
Denote the sum:

Rosenbrock function gradient:

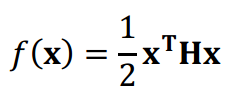
Rosenborck Hessian:

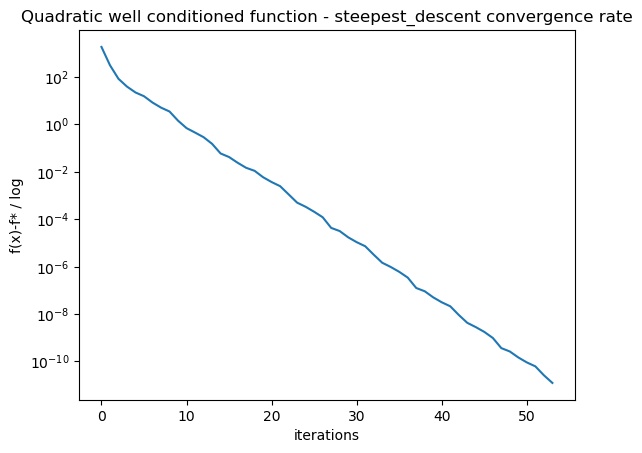
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  | 0 | 0 | … | 0 |
|  |  |  | 0 | … | 0 |
| 0 |  |  |  | … | 0 |
| 0 | 0 |  |  | … | 0 |
| 0 | 0 | 0 | … |  |  |
| 0 | 0 | 0 | … |  |  |

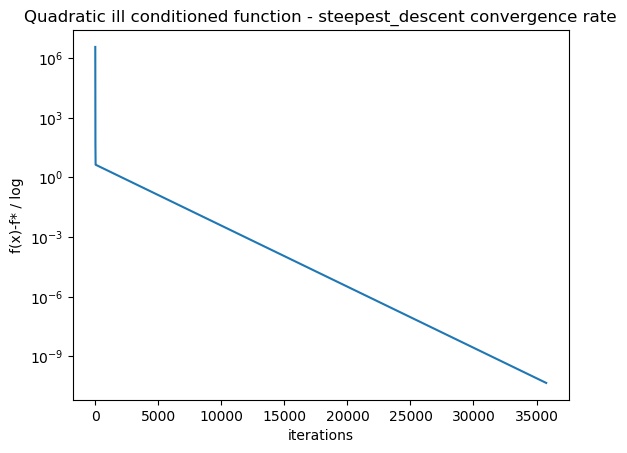
1. **Use the gradient descend method to find the optimal point of the Rosenbrock function:**

****

1. **Use the gradient descend method to optimize the following quadratic function:**



****

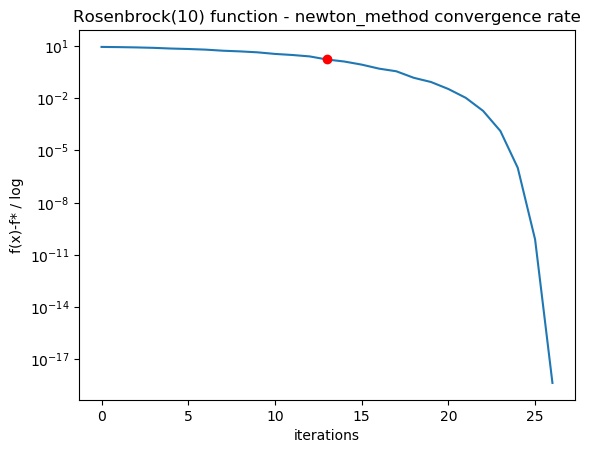
****

**Question 9:**

1. **Repeat the previous task but now use the Newton method:**

On the plots, mark the point where the convergence rate changes from linear to quadratic. Do all the graphs have this transition?

**Answer:** No, because Quadratic function converges to optimal point in single step.

****

