Optimization HW 6

Michael Nameika

March 2023

Section 11.2 Problems

3. Consider the function

$$f(x_1, x_2) = 8x_1^2 + 3x_1x_2 + 7x_2^2 - 25x_1 + 31x_2 - 29$$

Find all stationary points of this function and determine whether they are local minimizers and maximizers. Does this function have a global minimizer or a global maximizer?

To find all stationary points, we must solve $\nabla f = 0$. Well,

$$\nabla f(x_1, x_2) = \begin{pmatrix} 16x_1 + 3x_2 - 25 \\ 3x_1 + 14x_2 + 31 \end{pmatrix}$$

So we wish to solve

$$\begin{pmatrix} 16 & 3 \\ 3 & 14 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 25 \\ -31 \end{pmatrix}$$

From this, we get the solution

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 443/215 \\ -571/215 \end{pmatrix}$$

Now, to determine if this stationary point is a local min/max, let us inspect the Hessian:

$$\nabla^2 f(x_1, x_2) = \begin{pmatrix} 16 & 3 \\ 3 & 14 \end{pmatrix}$$

The Hessian is positive definite since in row echelon form we have the pivots are positive, as we can see below:

$$\begin{pmatrix} 16 & 3 \\ 0 & 215/16 \end{pmatrix}$$

Since the Hessian is positive definite, we have that this point is a local minimum. Additionally, since $\lim_{x_1 \to \infty} f(x_1, x_2), \lim_{x_2 \to \infty} f(x_1, x_2) = \infty$, our point is the global minimum.

13. Consider the quadratic function

$$f(x) = \frac{1}{2} x^T Q x - c^T x.$$

(i) Write the first-order necessary condition. When does a stationary point exist?

Notice that we may rewrite the problem as

$$\frac{1}{2} x^T Q x - c^T x = \frac{1}{2} \sum_{k=1}^{n} \sum_{i=1}^{n} q_{ki} x_k x_j - \sum_{i=1}^{n} c_i x_i$$

From this, we can see that the only terms that have x_i are

$$\left[\frac{1}{2}x^T Q x - c^T\right]\Bigg|_i = \frac{1}{2}q_{ii}x_i^2 + \frac{1}{2}\sum_{\substack{k=1\\k\neq i}}^{n} q_{ki}x_k x_i + \frac{1}{2}\sum_{\substack{j=1\\j\neq i}}^{n} q_{ij}x_i x_j - c_i x_i$$

Taking the partial derivative of the above equation with respect to x_i, we find

$$\begin{aligned}
&\frac{\partial}{\partial x_i}\left(\frac{1}{2}q_{ii}x_i^2 + \frac{1}{2}\sum_{\substack{k=1\\k\neq i}}^{n} q_{ki}x_k x_i + \frac{1}{2}\sum_{\substack{j=1\\j\neq i}}^{n} q_{ij}x_i x_j - c_i x_i\right) = \\
&= q_{ii}x_i + \frac{1}{2}\sum_{\substack{k=1\\k\neq i}}^{n} q_{ki}x_k + \frac{1}{2}\sum_{\substack{j=1\\j\neq i}}^{n} q_{ij}x_j - c_i \\
&= \frac{1}{2}\sum_{k=1}^{n} q_{ki}x_k + \frac{1}{2}\sum_{j=1}^{n} q_{ij}x_j - c_i
\end{aligned}$$

From this, we can see

$$\nabla f(x) = \frac{1}{2}(Q + Q^T)x - c$$

Then for a stationary point x_* to exist, we require $\nabla f(x_*) = 0$, or equivalently,

$$\nabla f(x_*) = \frac{1}{2}(Q + Q^T)x_* - c = 0$$

$$\frac{1}{2}(Q + Q^T)x_* = c$$

That is, we need x_* to be the solution to the linear system $1/2(Q + Q^T)x_* = c$.

(ii) Under what conditions on Q does a local minimizer exist?

From the work in part (i), it is easily shows that

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

and so we must have that $Q + Q^T$ is positive semidefinite for a local minimizer to exist.

(iii) Under what conditions on Q does f have a stationary point, but no local minima nor maxima?

For f to have a stationary point but no local minima nor maxima, $Q + Q^T$ must be indefinite.

Section 11.3 Problems

2. Use Newton's method to solve

$$\text{minimize} \quad f(x) = 5x^5 + 2x^3 - 4x^2 - 3x + 2.$$

Look for a solution in the interval $-2 \leq x \leq 2$. Make sure that you have found a minimum and not a maximum. You may want to experiment with different initial guesses of the solution.

Implementing Newton's method in MATLAB, with the initial guess $x_0 = 0$, we arrive at the stationary point $x_* \approx -0.2899$, which, as we can see from the images below, is a local maximum since $f''(x) < 0$:

```
x_0 =

  -0.289897948556636
```

```
>> fprime(x_0)

ans =

    -4.440892098500626e-16
```

```
>> fprime2(x_0)

ans =

 -13.915101530718509
```

Since we are searching for a minimum, this stationary point is not optimal. Trying the initial guess $x_0 = 1$, we arrive at the stationary point $x_* \approx 0.6899$, which, from the images below, is a local minimum of f.

```
x_0 =

   0.689897948556636
```

```
>> fprime(x_0)

ans =

    -2.664535259100376e-15
```

```
>> fprime2(x_0)

ans =

  33.115101530718498
```

The corresponding value of f is

$$f(x_*) \approx -0.5354$$

3. Use Newton's method to solve

$$\text{minimize} \quad f(x_1, x_2) = 5x_1^4 + 6x_2^4 - 6x_1^2 + 2x_1x_2 + 5x_2^2 + 15x_1 - 7x_2 + 13.$$

Use the initial guess $(1, 1)^T$. Make sure that you have found a minimum and not a maximum.

Implementing Newton's method for this problem in MATLAB, with the initial guess of $(1, 1)^T$, we find the following stationary point:

$$x_* \approx \begin{pmatrix} -1.42 \\ 0.5434 \end{pmatrix}$$

And, from the images below, we can see that $\nabla^2 f(x_*)$ is positive definite (since the eigenvalues of $\nabla^2 f(x_*)$ are strictly positive), so x_* corresponds to a local minimum.

```
x_0 =

  -1.142054928369237
   0.543372481205087
```

```
gradFx_0 =

   1.0e-14 *

   0.532907051820075
                   0
```

```
hessFx_0 =

  66.257367564747781    2.000000000000000
   2.000000000000000   31.258263039830009
```

```
>> eig(hessFx_0)

ans =

  31.144345190282863
  66.371285414294931
```

The corresponding value of f is

$$f(x_*) \approx -6.496$$

7. The purpose of this exercise is to prove Theorem 11.2. Assume that the assumptions of the theorem are satisfied.

(i) Prove that

$$x_{k+1} - x_* = \nabla^2 f(x_k)^{-1}[\nabla^2 f(x_k)(x_k - x_*) - (\nabla f(x_k) - \nabla f(x_*))].$$

Proof: Assume that $\nabla^2 f(x)$ is Lipschitz continuous on an open convex set S and $x_* \in S$ is a stationary point of $f(x)$ From Newton's method, we have that

$$x_{k+1} = x_k - \left[\nabla^2 f(x_k)\right]^{-1} \nabla f(x_k)$$

Subtracting x_* from each side, we obtain

$$x_{k+1} - x_* = x_k - x_* - \left[\nabla^2 f(x_k)\right]^{-1} \nabla f(x_k)$$

and so we may rewrite the above equation as

$$x_{k+1} - x_* = \left[\nabla^2 f(x_k)\right]^{-1} \left(\nabla^2 f(x_k)(x_k - x_*) - \nabla f(x_k)\right)$$

since x_* is a stationary point, we have that $\nabla f(x_*) = 0$ and so we can add $\nabla f(x_*)$ into the above equation to obtain

$$x_{k+1} - x_* = \left[\nabla^2 f(x_k)\right]^{-1} \left(\nabla^2 f(x_k)(x_k - x_*) - (\nabla f(x_k) - \nabla f(x_*))\right)$$

Which is what we sought to show.

(iii) Prove that for large enough k,

$$\|x_{k+1} - x_*\| \leq L \left\|\nabla^2 f(x_k)^{-1}\right\| \|x_k - x_*\|^2.$$

and from here prove the results of the theorem.

Proof: Suppose $\nabla^2 f(x)$ is Lipschitz continuous on an open convex set S and begin by noticing that

$$\nabla f(x_k) = \nabla f(x_* + (x_k - x_*))$$

And so by Taylor's theorem,

$$\nabla f(x_* + x_k - x_*) = \nabla f(x_*) + \nabla^2 f(\xi)(x_k - x_*)$$

Where R is the remainder term. Using this in combination with our result from part (i), we have the following:

$$\begin{aligned}
x_{k+1} - x_* &= \left[\nabla^2 f(x_k)\right]^{-1} \left(\nabla^2 f(x_k)(x_k - x_*) - (\nabla f(x_k) - \nabla f(x_*))\right) \\
&= \left[\nabla^2 f(x_k)\right]^{-1} \left(\nabla^2 f(x_k)(x_k - x_*) - (\nabla f(x_*) + \nabla^2 f(\xi)(x_k - x_*) - \nabla f(x_*))\right) \\
&= \left[\nabla^2 f(x_k)\right]^{-1} \left((\nabla^2 f(x_k) - \nabla^2 f(\xi))(x_k - x_*)\right)
\end{aligned}$$

and so

$$\begin{aligned}
\|x_{k+1} - x_*\| &= \left\| \left[\nabla^2 f(x_k)\right]^{-1} \left((\nabla^2 f(x_k) - \nabla^2 f(\xi))(x_k - x_*)\right) \right\| \\
&\leq \left\| \left[\nabla^2 f(x_k)\right]^{-1} \right\| \left\| \left(\nabla^2 f(x_k) - \nabla^2 f(\xi)\right) \right\| \|x_k - x_*\| \\
&\leq \left\| \left[\nabla^2 f(x_k)\right]^{-1} \right\| \|x_k - \xi\| \, \|x_k - x_*\|
\end{aligned}$$

Since $\xi \in (x_k, x_*)$ or $\xi \in (x_*, x_k)$, $\|x_k - \xi\| \leq \|x_k - x_*\|$ so

$$\|x_{k+1} - x_*\| \leq L \left\| \left[\nabla^2 f(x_k)\right]^{-1} \right\| \|x_k - x_*\|^2$$

Which is what we sought to show.

Now we must prove the main result of Theorem 11.2: that $\{x_k\}$ converges to x_* quadratically.
Proof: Let $f(x)$ be defined on an open convex set S be such that $\nabla^2 f(x)$ is positive definite and Lipschitz continuous and consider the sequence $\{x_k\}$ generated by

$$x_{k+1} = x_k - \left[\nabla^2 f(x_k)\right]^{-1} \nabla f(x_k)$$

and further assume that x_* is a minimizer of f. If $\|x_0 - x_*\|$ is sufficiently small, we wish to show that $\{x_k\}$ converges to x_* quadratically. That is, we wish to show

$$\lim_{k\to\infty} \frac{\|x_{k+1} - x_*\|}{\|x_k - x_*\|^2} = C < \infty$$

Well from the result of part (iii) above, we have

$$\begin{aligned}
\lim_{k\to\infty} \frac{\|x_{k+1} - x_*\|}{\|x_k - x_*\|^2} &\leq \lim_{k\to\infty} \frac{L \left\|\nabla^2 f(x_k)^{-1}\right\| \|x_k - x_*\|^2}{\|x_k - x_*\|^2} \\
&= \lim_{k\to\infty} L \left\|\nabla^2 f(x_k)^{-1}\right\| \\
&= L \left\|\nabla^2 f(x_*)^{-1}\right\|
\end{aligned}$$

And so we have $\frac{\|x_{k+1} - x_*\|}{\|x_k - x_*\|}$ converges, and by definition, $\{x_k\}$ converges to x_* quadratically.

8. Let $\{x_k\}$ be a sequence that converges superlinearly to x_*. Prove that

$$\lim_{k\to\infty} \frac{\|x_{k+1} - x_k\|}{\|x_k - x_*\|} = 1$$

Proof: Let $\{x_k\}$ be a sequence that converges superlinearly to x_*. That is,

$$\lim_{k\to\infty} \frac{\|x_{k+1} - x_*\|}{\|x_k - x_*\|} = 0$$

We wish to show

$$\lim_{k\to\infty} \frac{\|x_{k+1} - x_k\|}{\|x_k - x_*\|} = 1$$

Well, notice, by the triangle inequality

$$\begin{aligned} \|x_{k+1} - x_k\| &= \|x_{k+1} - x_* + x_* - x_k\| \leq \|x_{k+1} - x_k\| + \|x_* - x_k\| \\ &= \|x_{k+1} - x_k\| + \|x_k - x_*\| \end{aligned}$$

By the reverse triangle inequality, we have

$$\begin{aligned} \|x_{k+1} - x_k\| &= \|x_{k+1} - x_* + x_* - x_k\| \\ &= \|x_{k+1} - x_* - (x_k - x_*)\| \\ &\geq \left| \|x_{k+1} - x_*\| - \|x_k - x_*\| \right| \end{aligned}$$

From this, we have

$$\lim_{k\to\infty} \frac{\left| \|x_{k+1} - x_*\| - \|x_k - x_*\| \right|}{\|x_k - x_*\|} \leq \lim_{k\to\infty} \frac{\|x_{k+1} - x_k\|}{\|x_k - x_*\|} \leq \lim_{k\to\infty} \frac{\|x_{k+1} - x_*\| + \|x_k - x_*\|}{\|x_k - x_*\|}$$

Evaluating the lower and upper bound limits, we find

$$\begin{aligned} \lim_{k\to\infty} \frac{\left| \|x_{k+1} - x_*\| - \|x_k - x_*\| \right|}{\|x_k - x_*\|} &= \left| \lim_{k\to\infty} \frac{\|x_{k+1} - x_*\|}{\|x_k - x_*\|} - \lim_{k\to\infty} \frac{\|x_k - x_*\|}{\|x_k - x_*\|} \right| \\ &= \left| 0 - 1 \right| \\ &= 1 \end{aligned}$$

and

$$\begin{aligned} \lim_{k\to\infty} \frac{\|x_{k+1} - x_*\| + \|x_k - x_*\|}{\|x_k - x_*\|} &= \lim_{k\to\infty} \frac{\|x_{k+1} - x_*\|}{\|x_k - x_*\|} + \lim_{k\to\infty} \frac{\|x_k - x_*\|}{\|x_k - x_*\|} \\ &= 0 + 1 \\ &= 1 \end{aligned}$$

Then we have

$$1 \leq \lim_{k\to\infty} \frac{\|x_{k+1} - x_k\|}{\|x_k - x_*\|} \leq 1$$

and by the squeeze theorem, we have

$$\lim_{k\to\infty} \frac{\|x_{k+1} - x_k\|}{\|x_k - x_*\|} = 1$$

which is what we sought to show.

9. Let f be a real-valued function of n variables and assume that f, ∇f, and $\nabla^2 f$ are continuous. Suppose that $\nabla^2 f(\bar{x})$ is nonsingular for some point $\bar{x}$. Prove that there exists constants $\epsilon > 0$ and $\beta > \alpha > 0$ such that

$$\alpha\|x - \bar{x}\| \leq \|\nabla f(x) - \nabla f(\bar{x})\| \leq \beta\|x - \bar{x}\|$$

for all x satisfying $\|x - \bar{x}\| \leq \epsilon$.

Proof: Let $f : \mathbb{R}^n \to \mathbb{R}$ and suppose $f, \nabla f$, and $\nabla^2 f$ are continuous and suppose $\nabla^2 f(\bar{x})$ is nonsingular for some $\bar{x}$. Further suppose that $\|x - \bar{x}\| \leq \epsilon$ for some $\epsilon > 0$. Notice that

$$\nabla f(x) = \nabla f(x - \bar{x} + \bar{x})$$

and by Taylor's theorem, we have

$$\begin{aligned}
\nabla f(x) &= \nabla f(\bar{x}) + (x - \bar{x})^T \nabla^2 f(\xi) \\
\nabla f(x) - \nabla f(\bar{x}) &= (x - \bar{x})^T \nabla^2 f(\xi) \\
\|\nabla f(x) - \nabla f(\bar{x})\| &= \|\nabla^2 f(\xi)^T (x - \bar{x})\| \\
&\leq \|\nabla^2 f(\xi)\| \|(x - \bar{x})\|
\end{aligned}$$

Let $\beta = \|\nabla^2 f(\xi)\| > 0$. Now we have

$$\|\nabla f(x) - \nabla f(\bar{x})\| \leq \beta\|x - \bar{x}\|$$

Additionally, from above,

$$\begin{aligned}
\nabla f(x) &= \nabla f(\bar{x}) + \nabla^2 f(\xi)(x - \bar{x})^T \\
\nabla f(x) - \nabla f(\bar{x}) &= \nabla^2(x)(x - \bar{x})^T \\
\nabla^2 f(x)^{-1}(\nabla f(x) - \nabla f(\bar{x})) &= (x - \bar{x})^T \\
\|\nabla^2 f(x)^{-1}(\nabla f(x) - \nabla f(\bar{x}))\| &= \|x - \bar{x}\|
\end{aligned}$$

Then we have

$$\begin{aligned}
\|x - \bar{x}\| &\leq \|\nabla f(x) - \nabla f(\bar{x})\| \|\nabla^2 f(x)^{-1}\| \\
\frac{1}{\|\nabla^2 f(x)^{-1}\|}\|x - \bar{x}\| &\leq \|\nabla f(x) - \nabla f(\bar{x})\|
\end{aligned}$$

Let $\alpha = \frac{1}{\|\nabla^2 f(x)^{-1}\|} > 0$. Then we have

$$\alpha\|x - \bar{x}\| \leq \|\nabla f(x) - \nabla f(\bar{x})\| \leq \beta\|x - \bar{x}\|$$

Section 11.4 Problems

1. Find a diagonal matrix E so that $A + E = LDL^T$ where

$$A = \begin{pmatrix} 1 & 4 & 3 \\ 4 & 2 & 5 \\ 3 & 5 & 3 \end{pmatrix}$$

Notice that $a_{11} = 1 > 0$ in this (initial) stage, so we'll leave it alone. Now pivot the first column with the following operations:

$$R_2 - 4R_1 \to R_2$$
$$R_3 - 3R_1 \to R_3$$

Then A becomes

$$\begin{pmatrix} 1 & 4 & 3 \\ 0 & -14 & -7 \\ 0 & -7 & -6 \end{pmatrix}$$

Replace a_{22} in this stage with 7. That is, add 21 to a_{22}. Then pivoting the second column, by adding row two to the third row, A becomes

$$\begin{pmatrix} 1 & 4 & 3 \\ 0 & 7 & -7 \\ 0 & 0 & -13 \end{pmatrix}$$

Now replace a_{33} in this stage with 1. That is, add 14 to a_{33}. Then we have

$$E = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 21 & 0 \\ 0 & 0 & 14 \end{pmatrix}$$

Finally, we have

$$A + E = LDL^T$$

with

$$LDL^T = \begin{pmatrix} 1 & 0 & 0 \\ 4 & 1 & 0 \\ 3 & -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 4 & 3 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}$$