APM462 Nonlinear Optimization, Winter 2014 Midterm solutions

1. (10 pts) Find a point satisfying the first-order conditions for the function

$$f(x, y) = x^2 + 2xy + 4y^2 - 2y + x$$

and prove that this point is a global minimum of f.

solution.

The first-order conditions are

$$2x + 2y + 1 = 0, \qquad 2x + 8y - 2 = 0.$$

It is straightforward to solve this and find that $(x, y) = (-1, \frac{1}{2})$.

To prove that it is a global minimum point, note that the matrix of second derivatives of f is

$$\begin{pmatrix} 2 & 2 \\ 2 & 8 \end{pmatrix}$$

this matrix is positive definite, since both diagonal entries are positive and the determinant is positive. So f is convex, and any critical point must be a global minimum point.

2. (10 pts) Let $f(x) = x \ln x$, for $x > 0$.

Suppose that you want to find the minimum of f by Newton's method. Give the formula for x_{n+1} in terms of x_n. Simplify your answer as much as possible.

solution.

By using the general formula

$$x_{n+1} = x_n - \frac{f'(x_n)}{f''(x_n)}$$

and doing a little calculus, one finds that

$$x_{n+1} = -x_n \ln(x_n).$$

3. (10 pts) Assume that f is a convex function on E^n, and that $L : E^m \to E^n$ is affine, which means that it has the form

$$L(x) = Bx + c$$

where B is an $n \times m$ matrix and $c \in E^n$.

Let $g(x) = f(L(x))$, and prove that g is a convex function on E^m.

solution.

First note that for any $\theta \in [0,1]$,

$$L(\theta x + (1-\theta)y) = B(\theta x + (1-\theta)y) + \theta c + (1-\theta)c = \theta L(x) + (1-\theta)L(y).$$

Since f is convex, it follows that

$$f\Big(L(\theta x + (1-\theta)y)\Big) = f\Big(\theta L(x) + (1-\theta)L(y)\Big) \le \theta f(L(x)) + (1-\theta)f(L(y)).$$

This says exactly that $g = f \circ L$ is convex.

4. (10 pts) Let $f(x,y) = \frac{1}{2}x^2 + y$, which is a convex function on E^2.

Assume you want to minimize f by the method of steepest descent.

a. If you start at the point $\vec{x}_0 = (1,1)$, find $\vec{x}_1$ and $\vec{x}_2$ (that is, the first two iterates of the method of steepest descent.)

b. Will the method of steepest descent converge to a global minimum for this problem? Explain your answer.

solution. By using the definition of the method of steepest descent and doing some calculations, one find that

$$\vec{x}_1 = (-1,-1), \qquad \vec{x}_2 = (1,-3).$$

The method of steepest descent will not converge because there is no global minimum point for the problem. This is clear because there is no point satisfying the first-order necessary conditions. It is also clear that $f(0,-k) = k$, so that f is not bounded from below.

In fact, it is not hard to check that $\vec{x}_k = ((-1)^k, 1-2k)$, so that the y values become increasingly negative. This is natural, in view of the fact tht we can always decrease the function by decreasing y.

5. (10 pts) Let f be a continuous function on a compact set $K \subset E^n$,

Consider the following algorithm for minimizing f: starting from a point x_0, iteratively define

$$x_{n+1} = \text{ any point in the set } A(x) := \{y \in K : f(y) < f(x)\}.$$

Does this algorithm satisfy the hypotheses of the global convergence theorem? Justify your answer.

solution 1.

This algorithm does not satisfy the hypotheses of the global convergence theorem because it is not closed.

To see this, suppose for concreteness that $f(x) = x^2$ for $x \in (-\infty, \infty)$, let $x_k = 1$ for all positive integers k, and let $y_k = 1 - \frac{1}{k}$.

Then

$$y_k \in A(x_k) \text{ for every } k, \qquad x_k \to x = 1 \qquad y_k \to y = 1$$

but $y \notin A(x)$.

solution 2.

A different way to see that the algorithm does not satisfy the hypotheses of the global convergence theorem is to see that it does not always converge to a minimizer.

For example, let $f(x) = x^2$ as above, and let $x_k = 1 + \frac{1}{k}$. Then $x_{k+1} \in A(x_k)$, so this sequence is indeed generated by the algorithm. But $x_k \to 1$, which is certainly not the global minimum point of the function.

Note that if the algorithm did converge to a minimum point, it would say that to find a minimizer of a function f, all that you need to do is to choose a sequence of points x_k such that $f(x_{k+1}) < f(x_k)$ for all k. This is clearly too good to be true!

6. (10 pts) Let f be a function on E^n defined by $f(x) = \frac{1}{2}x^TQx - b^Tx$ where Q is a positive semidefinite $n \times n$ matrix.

Assume also that $d_1, \ldots, d_n$ are Q-orthogonal vectors, and for $y = (y_1, \ldots, y_n) \in E^n$, define $g(y) = f(\sum_{i=1}^n y_i d_i)$. Find an expression for g of the form

$$g(y) = \frac{1}{2}y^TRy - c^Ty$$

for some $n \times n$ matrix R and a vector $c \in E^n$, with formulas for R and c, simplified as much as possible.

solution

$$\begin{aligned}
g(y) &= f(\sum_{i=1}^n y_i d_i) \\
&= \frac{1}{2}(\sum_{i=1}^n y_i d_i)^T Q (\sum_{j=1}^n y_j d_j) - b^T(\sum_{i=1}^n y_i d_i) \\
&= \frac{1}{2}\sum_{i,j=1}^n y_i d_i^T Q d_j y_j - \sum_{i=1}^n b^T d_i y_i \\
&= \frac{1}{2}\sum_{i,j=1}^n y_i \left(d_i^T Q d_j\right) y_j - \sum_{i=1}^n b^T d_i y_i.
\end{aligned}$$

This is the same as

$$\frac{1}{2}y^TRy - c^Ty$$

for

$$\boxed{R_{ij} = d_i^TQd_j = \begin{cases} d_i^TQd_i & \text{if } i = j \\ 0 & \text{if } i \neq j, \end{cases} \qquad c_i = b^Td_i.}$$

7. (10 pts) Assume that f is a strictly convex function on E^2, and that f attains its global minimum at the point $(-5, 2)$ and nowhere else.

For this function f, consider the minimization problem

$$\text{minimize } f \text{ in the set } \Omega := \{(x, y) \in E^2 : x \geq 0 \text{ and } y \geq 0\},$$

and prove that the minimum (x^*, y^*) must satisfy $x^* = 0$.

solution 1. Let (x, y) by any point in the set Ω such that $x > 0$.

Consider the line segment joining (x, y) to $(-5, 2)$, *i.e.* the set of points

$$\theta(x, y) + (1 - \theta)(-5, 2) \qquad 0 \leq \theta \leq 1.$$

Geometrically, since (x, y) is in the first quadrant and $(-5, 2)$ is in the second quadrant, it is clear that there is some value $\theta^* \in (0, 1)$ such that the point $\theta^*(x, y) + (1 - \theta^*)(-5, 2)$ lies on the positive y axis. Indeed, we can see that this is the case if $\theta^* x + (1 - \theta^*)(-5) = 0$, which happens when $\theta^* = \frac{5}{x+5}$, which is clearly between 0 and 1. Then by strict convexity, and since $f(x, y) > f(-5, 2)$ (the global minimum),

$$\begin{aligned} f(\theta^*(x, y) + (1 - \theta^*)(-5, 2)) &< \theta^* f(x, y) + (1 - \theta^*) f(-5, 2) \\ &< \theta^* f(x, y) + (1 - \theta^*) f(x, y) = f(x, y). \end{aligned}$$

So for every point in Ω with $x > 0$, we can find a point in Ω such that $x = 0$, and with a smaller value of f.

Thus the minimum cannot occur where $x > 0$, and so it must occur where $x = 0$.

solution 2. We consider 3 cases.

case 1 $x^* > 0$ and $y^* > 0$. Then the necessary condition is $\nabla f = 0$. But the ony critical point for a convex function is the global minimum point, which we know does not satisfy $x^* > 0$, so this is impossible.

case 2 $x^* > 0$ and $y^* = 0$. Then the necessary condition is

$$\frac{\partial f}{\partial x} = 0, \frac{\partial f}{\partial y} \geq 0.$$

Also, since f is convex, we know that at any other point (x, y)

$$f(x, y) \geq f(x^*, y^*) + \nabla f(x^*, y^*)(x - x^*, y - y^*).$$

Plug $(x, y) = (-5, 2)$ into this inequality and rewrite to find that

$$f(-5, 2) - f(x^*, y^*) \geq \frac{\partial f}{\partial x}(-5 - x^*) + \frac{\partial f}{\partial y}(-y^*)$$

But the necessary conditions imply that the right-hand side equals 0. Thus it follows that $f(-5, 2) \geq f(x^*, y^*)$, which we know is impossibe. So case 2 cannot occur.

Thus the only remaining possibility is **case 3**: $x^* = 0$.