CS5321 Numerical Optimization Homework 2

Due March 24

1. (20%) A set $\mathcal{S}$ convex if the straight line connecting any two points in $\mathcal{S}$ is entirely in $\mathcal{S}$. A function is called *convex* if its domain $\mathcal{S}$ is convex, and for any $\vec{x}, \vec{y} \in \mathcal{S}$,

$$f(\alpha\vec{x} + (1-\alpha)\vec{y}) \le \alpha f(\vec{x}) + (1-\alpha)f(\vec{y}),$$

for all $\alpha \in [0, 1]$.

   (a) Prove that when f is convex, any local minimizer $\vec{x}^*$ is a global minimizer of f. (Hint: Suppose there is another point $\vec{z} \in \mathcal{S}$ such that $f(\vec{z}) \le f(\vec{x}^*)$. Then $\vec{x}^*$ is not a local minimizer.)

   (b) Suppose $f(\vec{x}) = \vec{x}^T Q\vec{x}$, where Q is a symmetric positive semidefinite matrix. Show that $f(\vec{x})$ is convex. (Hint: It might be easier to show $f(\vec{y} + \alpha(\vec{x} - \vec{y})) - \alpha f(\vec{x}) - (1-\alpha)f(\vec{y}) \le 0$.)

(a) Suppose $\vec{x}^*$ is the local but not a global minimizer. We can find $\vec{z}$ such that $f(\vec{z}) < f(\vec{x}^*)$. Consider the line segment of $\vec{x}^*$ and $\vec{z}$,

$$\vec{x} = \mu\vec{z} + (1-\mu)\vec{x}^*$$

for some $\mu \in (0, 1]$. Since f is convex,

$$f(\vec{x}) \le \mu f(\vec{z}) + (1-\mu)f(\vec{x}^*) < f(\vec{x}^*)$$

for any μ, which violates the assumption that $\vec{x}^*$ is the local minimizer.

(b)

$$\begin{aligned} f(\vec{y} + \alpha(\vec{x} - \vec{y})) &= \alpha^2\vec{x}^T Q\vec{x} + (1-\alpha)^2\vec{y}^T Q\vec{y} + 2\alpha(1-\alpha)\vec{x}^T Q\vec{y} && (1) \\ \alpha f(\vec{x}) - (1-\alpha)f(\vec{y}) &= \alpha^2\vec{x}^T Q\vec{x} + (1-\alpha)^2\vec{y}^T Q\vec{y} + \alpha(1-\alpha)(\vec{x}^T Q\vec{x} + \vec{y}^T Q\vec{y}) && (2) \end{aligned}$$

(1) − (2) $\Rightarrow -\alpha(1-\alpha)(\vec{x} - \vec{y})^T Q(\vec{x} - \vec{y}) \le 0$. (because Q is symmetric positive semi-definite.) Therefore, $f(\vec{x})$ is convex.

2. (30%) For a given function $f(x) : \mathbb{R} \to \mathbb{R}$,

(a) What is the quadratic polynomial $p(x)$ satisfying $p(0) = f(0)$, $p(1) = f(1)$, and $p'(0) = f'(0)$? Express $p(x)$ by $f(0)$, $f(1)$, and $f'(0)$.
Suppose $p(x) = ax^2 + bx + c$. $p'(x) = 2ax + b$. For given conditions,

$$\begin{pmatrix} 0 & 0 & 1 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} f(0) \\ f(1) \\ f'(0) \end{pmatrix}$$

Thus, $c = f(0)$, $b = f'(0)$, and $a = f(1) - f(0) - f'(0)$.

(b) What is the minimizer of $p(x)$ for $x \in [0, 1]$? You may need to discuss different cases for different $f(0)$, $f(1)$, and $f'(0)$.
Let $z = \frac{f'(0)}{2(f(1)-f'(0)-f(0))}$.

$$x^* = \begin{cases} 0, & f'(0) \geq 0 \text{ and } f(0) \leq f(1); \\ z, & f'(0) < 0 \text{ and } z \in (0, 1); \\ 1, & \text{Otherwise.} \end{cases}$$

(c) What is the cubic polynomial $q(x)$ satisfying $q(0) = f(0)$, $q(\alpha_1) = f(\alpha_1)$, $q(\alpha_2) = f(\alpha_2)$, and $q'(0) = f'(0)$? Express $q(x)$ by $f(0)$, $f(\alpha_1)$, $f(\alpha_2)$, and $f'(0)$.
Suppose $p(x) = ax^3 + bx^2 + cx + d$. $p'(x) = 3ax^2 + bx + c$. For given conditions,

$$\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ \alpha_1^3 & \alpha_1^2 & \alpha_1 & 1 \\ \alpha_2^3 & \alpha_2^2 & \alpha_2 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \begin{pmatrix} f(0) \\ f'(0) \\ f(\alpha_1) \\ f(\alpha_2) \end{pmatrix}.$$

Solve that for a, b, c, d.

3. (50%) Let $f_1(x, y) = \frac{1}{2}x^2 + \frac{9}{2}y^2$ and $f_2(x, y) = \frac{1}{2}x^2 + y^2$.

(a) Derive the gradient g and the Hessian H of f_1 and f_2, and compute Hs' eigenvalues.

(b) Write Matlab codes to implement the steepest descent method and Newton's method with $\vec{x}_0 = (9, 1)$, and compare their convergent results. The formula of the steepest descent method is

$$\vec{x}_{k+1} = \vec{x}_k - \frac{\vec{g}_k^T \vec{g}_k}{\vec{g}_k^T H_k \vec{g}_k} \vec{g}_k,$$

and the formula of Newton's method is

$$\vec{x}_{k+1} = \vec{x}_k - H_k^{-1} \vec{g}_k,$$

where $\vec{g}_k = g(\vec{x}_k)$ and $H_k = H(\vec{x}_k)$.