日期: /

Net ID: 24112456g Name: Zhong Qiaoyang

1. Discuss whether the following functions are convex.
   (a) (5 points) $f : \mathbb{R}^2 \to \mathbb{R}$ defined by $f(x) = \cos^2 x_1 + \sin^2 x_2$.
   (b) (5 points) $f : \mathbb{R} \to \mathbb{R}$ defined by

$$f(x) = \begin{cases} |x| - \pi/2 & \text{if } |x| > \pi/2, \\ -\cos x & \text{if } -\pi/2 \le x \le \pi/2. \end{cases}$$

   (c) (10 points) $f : \mathbb{R}^3 \to \mathbb{R}$ defined by

$$f(x) = \sum_{i=1}^{3} \ln(4 + (x_i)^2) + \frac{1}{4}\|x\|_2^2 + |8x_1 + 3x_2 - 6x_3 + 1| + |4x_1 - x_2 + 5x_3 - 12|$$

1. (a) $\nabla^2 f(x) = \begin{pmatrix} -2\cos 2x_1 & 0 \\ 0 & 2\cos 2x_2 \end{pmatrix}$

$\nabla^2 f(x)$ is not positive definite on $\mathbb{R}^2$, so $f(x)$ is not convex.

(b) on $(-\frac{\pi}{2}, \frac{\pi}{2})$, $f''(x) = \cos x > 0$.

on $|x| > \frac{\pi}{2}$, $f''(x) = 0$.

when $x = \frac{\pi}{2}$, we have $f'(x) = 1$, $\lim_{x \to \frac{\pi}{2}^+} \frac{f(x) - f(\frac{\pi}{2})}{x - \frac{\pi}{2}} = \lim_{x \to \frac{\pi}{2}^-} \frac{f(x) - f(\frac{\pi}{2})}{x - \frac{\pi}{2}}$.

and $\lim_{x \to \frac{\pi}{2}^+} \frac{f'(x) - f'(\frac{\pi}{2})}{x - \frac{\pi}{2}} = \lim_{x \to \frac{\pi}{2}^-} \frac{f'(x) - f'(\frac{\pi}{2})}{x - \frac{\pi}{2}} = 0$.

same as $x = -\frac{\pi}{2}$, so we have $f(x) \in C^2(\mathbb{R})$

and $f''(x) \ge 0$ for $x \in \mathbb{R}$, so $f(x)$ is convex.

(c) set $f_1(x) := \sum_{i=1}^{3} \ln(4 + (x_i)^2)$, $f_2(x) := \frac{1}{4}\|x\|_2^2$

$f_3(x) := |8x_1 + 3x_2 - 6x_3 + 1|$, $f_4(x) := |4x_1 - x_2 + 5x_3 - 12|$

set $\nabla_x^2 f_1(x) = \mathrm{diag}\left(\frac{8 - 2x_1^2}{(x_1^2+4)^2}, \frac{8 - 2x_2^2}{(x_2^2+4)^2}, \frac{8 - 2x_3^2}{(x_3^2+4)^2}\right)$

$\nabla_x^2 f_2(x) = \frac{1}{2} I$,

since $\frac{8 - 2x^2}{(x^2+4)^2} + \frac{1}{2} = \frac{8 - 2x^2 + (x^2+4)^2 \cdot \frac{1}{2}}{(x^2+4)^2} \ge \frac{8 + 6x^2}{(x^2+4)^2} > 0$

hence $\nabla_x^2 f_1(x) + \nabla_x^2 f_2(x) \succ 0$. $\Rightarrow f_1 + f_2$ is convex and assume $y := 8x_1+3x_2-6x_3+1$. y is linear transform of x.

we have $|\lambda y_1| + |(1-\lambda)y_2| \geq |\lambda y_1 + (1-\lambda)y_2|$, triangular inequality.

so f_3 and f_4 are convex.

$\Rightarrow f$ is convex. $\square$

2. (10 points) For the following function, find all the stationary points and determine their nature based on their Hessian, if possible.

$$f(x) = \frac{1}{2}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}^T \begin{bmatrix} 3 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} - 6x_1 + 4x_2 - 3x_3^3 + 2024.$$

2. $\nabla f(x) = \begin{pmatrix} 3 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} + \begin{pmatrix} -6 \\ 4 \\ -9x_3^2 \end{pmatrix} = 0$

$\Rightarrow x = \begin{pmatrix} -1 \\ 3 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 3 \\ 1/3 \end{pmatrix}, \begin{pmatrix} -1 \\ 3 \\ -1/3 \end{pmatrix}$, 3 stationary points

$\nabla^2 f(x) = \begin{pmatrix} 3 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -18x_3 \end{pmatrix}$

if $x = \begin{pmatrix} -1 \\ 3 \\ 0 \end{pmatrix}$, eigen values of $\nabla^2 f(x)$ is $1, 2\pm\sqrt{2}$, $\nabla^2 f(x) \succ 0$,

$\begin{pmatrix} -1 \\ 3 \\ 0 \end{pmatrix}$ is local minimizer.

if $x = \begin{pmatrix} -1 \\ 3 \\ 1/3 \end{pmatrix}$, eigen values is $-5, 2\pm\sqrt{2}$,

$\begin{pmatrix} -1 \\ 3 \\ 1/3 \end{pmatrix}$ is not minimizer or maximizer locally.

if $x = \begin{pmatrix} -1 \\ 3 \\ -1/3 \end{pmatrix}$ eigen values is $7, 2\pm\sqrt{2}$. $\nabla^2 f(x) \succ 0$.

$\begin{pmatrix} -1 \\ 3 \\ -1/3 \end{pmatrix}$ is local minimizer. $\square$

3. (15 points) Consider the function

$$f(x) = \frac{1}{2}x^T\begin{bmatrix} 6 & -3 \\ -3 & 8 \end{bmatrix}x + 2x_2^2 + x_1^4 - 2x_1^3.$$

Perform two iterations of the BFGS method with $\alpha_k \equiv 1.5$, $x^0 = (1,0)$ and $H_0 = I$. Write down x^1 and x^2. You may correct your answers to 4 d.p.

3. $H_{k+1} = \left(I - \frac{s^k {y^k}^T}{{y^k}^T s^k}\right) H_k \left(I - \frac{y^k {s^k}^T}{{y^k}^T s^k}\right) + \frac{s^k {s^k}^T}{{y^k}^T s^k}$

1. $d^k = -H^k \cdot \nabla f(x^k)$
2. $x^{k+1} = x^k + \alpha^k \cdot d^k$
3. $y^k = \nabla f(x^{k+1}) - \nabla f(x^k)$, $s^k = x^{k+1} - x^k$
4. H^{k+1}.

日期: /

$H_0 = I, \quad x^0 = (1, 0), \quad \alpha_k \equiv 1.5$

$$\nabla f(x) = \begin{pmatrix} 6 & -3 \\ -3 & 8 \end{pmatrix} x + \begin{pmatrix} 4x_1^3 - 6x_1^2 \\ 4x_2 \end{pmatrix}$$

(1) $d^0 = -H^0 \cdot \nabla f(x^0) = \begin{pmatrix} -4 \\ 3 \end{pmatrix}, \quad \nabla f(x^0) = \begin{pmatrix} 4 \\ -3 \end{pmatrix}$

$x^1 = x^0 + \alpha \cdot d^0 = \begin{pmatrix} -5 \\ 4.5 \end{pmatrix}, \quad \nabla f(x^1) = \begin{pmatrix} -30 - 13.5 \\ 15 + 36 \end{pmatrix} + \begin{pmatrix} -500 - 150 \\ 18 \end{pmatrix} = \begin{pmatrix} -693.5 \\ 69 \end{pmatrix}$

$y^0 = \begin{pmatrix} -697.5 \\ 72 \end{pmatrix}, \quad s^0 = x^1 - x^0 = \begin{pmatrix} -6 \\ 4.5 \end{pmatrix}$

$$H^{k+1} = \left(I - \frac{s^k y^{kT}}{y^{kT} s^k}\right) H_k \left(I - \frac{y^k s^{kT}}{y^{kT} s^k}\right) + \frac{s^k s^{kT}}{y^{kT} s^k} \qquad H^1 = \left(I - \frac{s^0 y^{0T}}{y^{0T} s^0}\right)\left(I - \frac{y^0 s^{0T}}{y^{0T} s^0}\right) + \frac{s^0 \cdot s^{0T}}{y^{0T} \cdot s^0}$$

$$= \begin{pmatrix} 1.0125 & 0.0125 \\ 0.0125 & 1.0125 \end{pmatrix}$$

(2) $d^1 = -H^1 \cdot \nabla f(x^1) = \begin{pmatrix} 701.2906 \\ -61.2093 \end{pmatrix},$

? $x^2 = x^1 + \alpha \cdot d^1 = \begin{pmatrix} 1046.9360 \\ -87.3139 \end{pmatrix}$. □

$x_1 = \begin{pmatrix} -5 \\ 4.5 \end{pmatrix}$

4. (15 points) Let $a \in \mathbb{R}$ and consider the following linear programming problem.

$$\begin{array}{ll} \text{Maximize} & ax_1 - 3x_2 - 2x_3 + 3x_4 \\ \text{subject to} & 5x_1 - x_2 + x_4 \geq 6, \\ & 3x_2 + 2x_3 - x_4 = 8, \\ & x_1 + x_2 + x_3 + x_4 \leq 1, \\ & x_1, x_2, x_3, x_4 \text{ unrestricted.} \end{array}$$

By considering its dual or otherwise, find the optimal value of the above problem for all values of a.

4.

	Maximize $c^T x$ subject to $Ax♣b$, $x◇$	Minimize $b^T y$ subject to $A^T y◇c$, $y♣$
♣	ith constraint $\leq$	ith variable ≥ 0
	ith constraint $\geq$	ith variable ≤ 0
	ith constraint $=$	ith variable unrestricted
◇	jth variable ≥ 0	jth constraint $\geq$
	jth variable ≤ 0	jth constraint $\leq$
	jth variable unrestricted	jth constraint $=$

$c := (a \ \ -3 \ \ -2 \ \ 3)^T, \quad A := \begin{pmatrix} 5 & -1 & 0 & 1 \\ 0 & 3 & 2 & -1 \\ 1 & 1 & 1 & 1 \end{pmatrix}$

$b := (6 \ \ 8 \ \ 1)^T$

dual: $\text{Min } b^T y$

s.t. $A^T y = c$

$y_1 \leq 0, \ y_3 \geq 0.$

solve the constraints $A^T y = c$, we get $y = \begin{pmatrix} -1 \\ -2 \\ 2 \end{pmatrix}$, if $a = -3$.

if $a \neq -3$, feasible set is ϕ.

① if $a = -3$, optimal value is -20.

② if $a \neq -3$, $\exists\, x = (2, 0, 7/3, -10/3)^T \in$ primal feasible set, so optimal value is $+\infty$. □

日期:

5. Consider the function $f : \mathbb{R}^3 \to \mathbb{R}$ defined by

$$f(x) = \ln(4 + (x_1 - x_2 - 3)^2) + \ln(4 + (x_1 + 2x_3 + 1)^2) + \sum_{i=1}^{3} \sqrt{7 + x_i^2}.$$

Define

$$B := \begin{bmatrix} 1 & -1 & 0 \\ 1 & 0 & 2 \end{bmatrix} \quad \text{and} \quad h(y) = \ln(4 + (y_1 - 3)^2) + \ln(4 + (y_2 + 1)^2).$$

(a) (5 points) Write down the first and second-order derivatives of the function $g(t) = \ln(4 + t^2)$.

(b) (5 points) Compute $\nabla h(y)$, $\nabla^2 h(y)$ and BB^T.

(c) (15 points) Consider an iterate of the following form:

$$x^{k+1} = x^k - 0.6602 \cdot \nabla f(x^k).$$

Argue that any accumulation point of $\{x^k\}$ is a stationary point of f.

5. (a) $g'(t) = \dfrac{2t}{4+t^2}$, $g''(t) = \dfrac{8-2t^2}{(4+t^2)^2}$

(b) $\nabla^2 h(y) = \begin{pmatrix} \dfrac{8-2(y_1-3)^2}{(4+(y_1-3)^2)^2} & 0 \\ 0 & \dfrac{8-2(y_2+1)^2}{(4+(y_2+1)^2)^2} \end{pmatrix}$

$\nabla h(y) = \left(\dfrac{2(y_1-3)}{4+(y_1-3)^2}, \dfrac{2(y_2+1)}{4+(y_2+1)^2} \right)$

$B \cdot B^T = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 0 & 2 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -1 & 0 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 5 \end{pmatrix}$

(c) set $L := \|\nabla^2 f(x)\|_2$

set $y := Bx$, $f_1(x) := \ln(4+(x_1-x_2-3)^2) + \ln(4+(x_1+2x_3+1)^2)$,

$f_2(x) := \sum_{i=1}^{3} \sqrt{7+x_i^2}$.

$\nabla_x^2 f_1(x) = B^T \nabla_y^2 h(y) B$

$\nabla_x^2 f_2(x) = \operatorname{diag}\left(\dfrac{7}{(7+x_1^2)^{3/2}}, \dfrac{7}{(7+x_2^2)^{3/2}}, \dfrac{7}{(7+x_3^2)^{3/2}} \right)$

$\|\nabla^2 f(x)\|_2 \le \|\nabla_x^2 f_1(x)\|_2 + \|\nabla_x^2 f_2(x)\|_2$

$\le \|B\|_2^2 \|\nabla_y^2 h(y)\|_2 + \|\nabla_x^2 f_2(x)\|_2$

$= \left(\dfrac{7+\sqrt{13}}{2}\right)^2 \cdot \dfrac{1}{2} + \dfrac{1}{\sqrt{7}}$

$\dfrac{8-2t^2}{(4+t^2)^2} \le \dfrac{8}{(4+t^2)^2} \le \dfrac{1}{2}$

$\dfrac{7}{(7+x^2)^{3/2}} \le \dfrac{1}{\sqrt{7}}$

≤ 19 ? $\dfrac{2}{19} < 0.6602$?

if step size $< \dfrac{2}{L}$, accumulation points is stationary.

(d) Consider an iterate of the following form:

$$x^{k+1} = x^k + \alpha_k d^k.$$

Let $d^k = -\nabla f(x^k)$ and α_k be chosen to satisfy the Wolfe's condition. Suppose that x^k is nonstationary for all k.

i. (5 points) Show that the sequence $\{x^k\}$ is bounded.

ii. (10 points) Show that any accumulation point of $\{x^k\}$ is stationary. You should specify the ℓ that can be used in Zoutendijk's theorem.

(d) (i) $\|x^{k+1}\|_2 \le \sum_{i=1}^{3} \sqrt{7+(x_i^{(k+1)})^2} \le f(x^{k+1}) < f(x^k) + \alpha_k \cdot d^k \cdot \nabla f(x^k) < f(x^k)$

$\le \cdots \le f(x^0)$

hence, $\{x^k\}_{k=1}^{\infty}$ is bounded.

(ii) since ① $f \in C^1(\mathbb{R}^3)$, $f > 0$.

② $\forall\ x, y \in \mathbb{R}^3$, $\exists$ $\|\nabla f(x) - \nabla f(y)\|_2 \leq \ell \cdot \|x-y\|_2$

where $\ell = \sup_t \|\nabla^2 f(t)\|_2$, since $f \in C^2(\mathbb{R}^3)$, we can apply this theorem.

and we know $\|\nabla^2 f(t)\|_2$ has upper bound from 5.(c).

$\Rightarrow$ Applying Zoutendijk's theorem, we have

$$\sum_{k=0}^{\infty} \cos^2\theta_k \cdot \|\nabla f(x_k)\|_2^2 < \infty.$$

where $\cos\theta_k = \dfrac{-\nabla f(x^k)^T d^k}{\|\nabla f(x^k)\| \cdot \|d^k\|} = 1$,

hence we have $\|\nabla f(x_k)\|_2 \to 0$ as $k \to +\infty$.

$\Rightarrow$ any accumulation point is stationary. $\square$

日期:　　　/