HW 4

Friday, October 2, 2020 15:45

Jordan Saethre

Consider the Freudenstein and Roth test function

$$f(x) = f_1(x)^2 + f_2(x)^2 \qquad x \in \mathbb{R}^2$$

Where

$$f_1(x) = -13 + x_1 + ((5 - x_2)x_2 - 2)x_2$$

$$f_2(x) = -29 + x_1 + ((x_2 + 1)x_2 - 14)x_2$$

(i) Show that f has 3 stationary points. Find them then prove that one is a global min, one is a strict local min and the third is a saddle point.

$$f_1(x) = -13 + x_1 + (5x_2 - x_2^2 - 2)x_2$$

$$= -13 + x_1 + 5x_2^2 - x_2^3 - 2x_2$$

$$f_2(x) = -29 + x_1 + (x_2^2 + x_2 - 14)x_2$$

$$= -29 + x_1 + x_2^3 + x_2^2 - 14x_2$$

$$f(x) = (-13 + x_1 + 5x_2^2 - x_2^3 - 2x_2)^2$$

$$+ \left(-29 + x_1 + x_2^3 + x_2^2 - 14x_2\right)^2$$

$$\left.\begin{aligned} f_{x_1} = \; & 2(-13 + x_1 + 5x_2^2 - x_2^3 - 2x_2)\cdot(1) \\ & + 2(-29 + x_1 + x_2^3 + x_2^2 - 14x_2)\cdot(1) \end{aligned}\right\}$$

$$\begin{aligned} f_{x_1} = \; & -26 + 2x_1 + 10x_2^2 - 2x_2^3 - 4x_2 \\ & + -58 + 2x_1 + 2x_2^3 + 2x_2^2 - 28x_2 \end{aligned}$$

$$f_{x_1} = -84 + 4x_1 + 12x_2^2 - 32x_2$$

$$\begin{aligned} f_{x_2} = \; & 2(-13 + x_1 + 5x_2^2 - x_2^3 - 2x_2)(10x_2 - 3x_2^2 - 2) \\ & + 2(-29 + x_1 + x_2^3 + x_2^2 - 14x_2)(3x_2^2 + 2x_2 - 14) \end{aligned}$$

$$\begin{aligned} \Rightarrow \; & (-26 + 2x_1 + 10x_2^2 - 2x_2^3 - 4x_2)(10x_2 - 3x_2^2 - 2) \\ & + (-58 + 2x_1 + 2x_2^3 + 2x_2^2 - 28x_2)(3x_2^2 + 2x_2 - 14) \end{aligned}$$

$$\begin{aligned} \Rightarrow \; & -260x_2 + 20x_1x_2 + 100x_2^3 - 20x_2^4 - 40x_2^2 \\ & + 78x_2^2 - 6x_1x_2^2 - 30x_2^4 + 6x_2^5 + 12x_2^3 \\ & + 52 - 4x_1 - 20x_2^2 + 4x_2^3 + 8x_2 \\ & - 174x_2^2 + 6x_1x_2^2 + 6x_2^5 + 6x_2^4 - 84x_2^3 \\ & - 116x_2 + 4x_1x_2 + 4x_2^4 + 4x_2^3 - 56x_2^2 \end{aligned}$$

$-116x_2 + 4x_1x_2 + 4x_2^4 + 4x_2^3 - 56x_2^2$

$+812 - 28x_1 - 28x_2^3 - 28x_2^2 + 392x_2$

$\Rightarrow$ $24x_2$ ✓ $+24x_1x_2$ ✓ $+8x_2^3$ ✓ $-40x_2^4$ ✓ $-240x_2^2$ ✓

$+0$ ✓ $+12x_2^5$ ✓ $+864$ ✓ $-32x_1$ ✓

$$f_{x_2} = 864 - 32x_1 + 24x_2 + 24x_1x_2 - 240x_2^2 + 8x_2^3 - 40x_2^4 + 12x_2^5$$

---

x^* is a stationary point of f if $\nabla f(x^*) = 0$

System of 2 eqns 2 unknowns:

$$\begin{cases} f_{x_1} = -84 + 4x_1 + 12x_2^2 - 32x_2 = 0 \\ f_{x_2} = 864 - 32x_1 + 24x_2 + 24x_1x_2 - 240x_2^2 + 8x_2^3 - 40x_2^4 + 12x_2^5 = 0 \end{cases}$$

$$\begin{cases} f_{x_1} = -21 + x_1 + 3x_2^2 - 8x_2 = 0 & \text{(divided by 4)} \\ f_{x_2} = 432 - 16x_1 + 12x_2 + 12x_1x_2 - 120x_2^2 + 4x_2^3 - 20x_2^4 + 6x_2^5 = 0 & \text{(divided by 2)} \end{cases}$$

(Solve f_{x_1} for x_1)

$\Rightarrow \quad x_1 = 21 + 8x_2 - 3x_2^2$

(plug into f_{x_2})

(plug into f_{x_2})

$$\Rightarrow \quad 432 - 16(21 + 8x_2 - 3x_2^2) + 12x_2$$

$$+ 12(21 + 8x_2 - 3x_2^2)x_2 - 120x_2^2 + 4x_2^3$$

$$-20x_2^4 + 6x_2^5 = 0$$

$$\Rightarrow \quad 432 - 336 - 128x_2 + 48x_2^2 + 12x_2$$

$$+ 252x_2 + 96x_2^2 - 36x_2^3 - 120x_2^2 + 4x_2^3$$

$$-20x_2^4 + 6x_2^5 = 0$$

$$\Rightarrow \quad 96 + 136x_2 + 24x_2^2 - 32x_2^3 - 20x_2^4 + 6x_2^5 = 0$$

Solve for x_2 degree 5 polynomial will have 5 roots.

Using Wolfram Alpha $\rightsquigarrow$

$$x_2 = \begin{cases} -1 + i \\ -1 - i \\ \frac{1}{3}(2 - \sqrt{22}) \\ \frac{1}{3}(2 + \sqrt{22}) \\ 4 \end{cases}$$

these are the ones we care about?

---

plug into: $x_1 = 21 + 8x_2 - 3x_2^2$

if $x_2 = 4$

$$x_1 = 21 + 8(4) - 3(4)^2$$

$= 5$

Stationary Point: $(5,4) = x_1^*$

if $x_2 = \frac{1}{3}(2 - \sqrt{22})$

$$x_1 = 21 + 8\left(\tfrac{1}{3}(2 - \sqrt{22})\right) - 3\left(\tfrac{1}{3}(2 - \sqrt{22})\right)^2$$

$$= \tfrac{1}{3}(53 - 4\sqrt{22})$$

Stationary Point: $\left(\frac{1}{3}(53 - 4\sqrt{22}), \frac{1}{3}(2 - \sqrt{22})\right) = x_2^*$

if $x_2 = \frac{1}{3}(2 + \sqrt{22})$

$$x_1 = 21 + 8\left(\tfrac{1}{3}(2 + \sqrt{22})\right) - 3\left(\tfrac{1}{3}(2 + \sqrt{22})\right)^2$$

$$x_1 = \tfrac{1}{3}(53 + 4\sqrt{22})$$

Stationary Point: $\left(\frac{1}{3}(53 + 4\sqrt{22}), \frac{1}{3}(2 + \sqrt{22})\right) = x_3^*$

Now find $\nabla^2 f(x^*)$ and classify

$$\nabla^2 f(\vec{x}) = \begin{bmatrix} f_{x_1x_1} & f_{x_1x_2} \\ f_{x_2x_1} & f_{x_2x_2} \end{bmatrix}$$

$$f_{x_1x_1} = \frac{\partial}{\partial x_1}\left(-84 + 4x_1 + 12x_2^2 - 32x_2\right) = 4$$

$$f_{x_1 x_2} = \frac{\partial}{\partial x_2}\left(-84 + 4x_1 + 12x_2^2 - 32x_2\right) = 24x_2 - 32$$

$$f_{x_2 x_1} = \frac{\partial}{\partial x_1}\left(864 - 32x_1 + 24x_2 + 24x_1x_2 - 240x_2^2 + 8x_2^3 - 40x_2^4 + 12x_2^5\right)$$

$$= -32 + 24x_2$$

$$f_{x_2 x_2} = \frac{\partial}{\partial x_2}\left(864 - 32x_1 + 24x_2 + 24x_1x_2 - 240x_2^2 + 8x_2^3 - 40x_2^4 + 12x_2^5\right)$$

$$= 24 + 24x_1 - 480x_2 + 24x_2^2 - 160x_2^3 + 60x_2^4$$

$$\nabla^2 f(x^*) = \begin{bmatrix} 4 & 24x_2 - 32 \\ 24x_2 - 32 & 24 + 24x_1 - 480x_2 + 24x_2^2 - 160x_2^3 + 60x_2^4 \end{bmatrix}$$

$$x_1^* = (5, 4)$$

$$\nabla^2 f(x_1^*) = \begin{bmatrix} 4 & 64 \\ 64 & 3728 \end{bmatrix}$$

$$\text{trace}(\nabla^2 f(x_1^*)) = 4 + 3728 > 0$$

$$\det(\nabla^2 f(x_1^*)) = 4(3728) - (64)^2 > 0$$

By prop 2.20 $\nabla^2 f(x_1^*) \succ 0$ if and only if $\det(\nabla^2 f(x_1^*)) > 0$ and $tr(\nabla^2 f(x_1^*)) > 0$ and by thm 2.27 if $\nabla^2 f(x_1^*) \succ 0$ then x_1^* is a strict local minimum.

$$x_2^* = \left(\tfrac{1}{3}(53 - 4\sqrt{22}), \tfrac{1}{3}(2 - \sqrt{22})\right)$$

$$\nabla^2 f(x_2^*) = \begin{bmatrix} 4 & 24\left(\frac{1}{3}(53 - 4\sqrt{22})\right) - 32 \\ 24\left(\frac{1}{3}(53 - 4\sqrt{22})\right) - 32 & \frac{246164372}{27} - \frac{48005464\sqrt{22}}{27} \end{bmatrix}$$

$$\approx \begin{bmatrix} 4 & -53.52 \\ -53.52 & 901.89 \end{bmatrix}$$

$$\det(\nabla^2 f(x_2^*)) \approx 4(901.89) - (-53.52)^2 > 0$$
$$tr(\nabla^2 f(x_2^*)) \approx 4 + 901.89 > 0$$

$\therefore \Rightarrow$ strict local minimum

$$x_3^* = \left(\tfrac{1}{3}(53 + 4\sqrt{22}), \tfrac{1}{3}(2 + \sqrt{22})\right)$$

$$\nabla^2 f(x_3^*) = \begin{bmatrix} 4 & \frac{24}{3}(2 + \sqrt{22}) - 32 \\ \frac{24}{3}(2 + \sqrt{22}) - 32 & \frac{246164372}{27} + \frac{48005464\sqrt{22}}{27} \end{bmatrix}$$

$$\approx \begin{bmatrix} 4 & 21.52 \\ 21.52 & -643.51 \end{bmatrix}$$

By lemma 2.16 if diagonal entries have mixed signs then $\nabla^2 f(x_3^*)$ is indefinite and by thm 2.28 if $\nabla^2 f(x_3^*)$ is indefinite then x_3^* is a saddle point.

The only question left to answer is which of x_1^* and x_2^* is the global minimizer?

By theorem 2.32 (attainment under coerciveness) if f is a continuous coercive function on a nonempty closed set $S \subseteq \mathbb{R}^n$ then f has a global minimum.

The function f is coercive if (f continuous)

$$\lim_{\|x\| \to \infty} f(x) = \infty$$

$$f(x) = [f_1, f_2] \begin{bmatrix} f_1 \\ f_2 \end{bmatrix} = f_1^2 + f_2^2 = \|f\|^2 \to \infty$$

as $\|x\| \to \infty$.

$\therefore$ which ever of x_1^* or x_2^* is smaller in function value will be the global min

$f(x_1^*) \approx 0$ ← global minimizer

$f(x_2^*) \approx 48.98$ ← Strict local min

See jupyter notebook HW4.ipynb for problem 1 (ii).

---

2) Let f be a twice continuously differentiable function satisfying

$$L I \succeq \nabla^2 f(x) \succeq m I$$

for some $L > m > 0$ and let x^* be the unique minimizer of f over $\mathbb{R}^n$.

(i) Show that

$$f(x) - f(x^*) \geq \frac{m}{2} \|x - x^*\|^2$$

for all $x \in \mathbb{R}^n$.

Proof:

Expand $f(x)$ about the point x^*

$$f(x) = f(x^*) + \nabla f(x^*)^t (x - x^*)$$

$$+ \frac{1}{2}(x-x^*)^t \nabla^2 f(\xi)(x-x^*)$$

where $\xi \in [x, x^*]$.

Since $\nabla f(x^*) = 0$ this implies

$$f(x) - f(x^*) = \frac{1}{2}(x-x^*)^t \nabla^2 f(\xi)(x-x^*)$$

and since $\nabla^2 f(x) \succeq mI$

$$f(x) - f(x^*) \geq \frac{m}{2}\|x-x^*\|^2$$

$\square$

(ii) Let $\{x_k\}_{k\geq 0}$ be the sequence generated by damped Newton's method w/ $t_k = \frac{m}{L}$. Show that

$$f(x_k) - f(x_{k+1}) \geq \frac{m}{2L} \nabla f(x_k)^t (\nabla^2 f(x_k))^{-1} \nabla f(x_k)$$

Proof:

Taylor expand $f(x)$ about x_k and evaluate at $x = x_{k+1}$:

$$f(x_{k+1}) = f(x_k) + \nabla f(x_k)^t (x_{k+1} - x_k)$$

$$+ \frac{1}{2}(x_{k+1} - x_k)^t \nabla^2 f(x_k)(x_{k+1} - x_k)$$

$$\leq f(x_k) + \nabla f(x_k)^t (x_{k+1} - x_k)$$

Notice: $(x_{k+1} - x_k) = ((x_k + t_k d_k) - x_k)$

$$= t_k d_k$$

$$= -\frac{m}{L} \left(\nabla^2 f(x_k)\right)^{-1} \nabla f(x_k)$$

Therefore, after rearranging terms,

$$f(x_k) - f(x_{k+1}) \geq \frac{m}{L} \nabla f(x_k)^t \left(\nabla^2 f(x_k)\right)^{-1} \nabla f(x_k)$$

$\square$

(iii) Show that $x_k \to x^*$ as $k \to \infty$

Proof:

By (ii) we know $\{f(x_k)\}_{k \geq 0}$ converges since it is nonincreasing and bounded below.

$$f(x_k) - f(x_{k+1}) \longrightarrow 0 \quad \text{as} \quad k \to \infty$$

$$f(x_k) - f(x_{k+1}) \geq \frac{m}{L} \nabla f(x_k)^t \left(\nabla^2 f(x_k)\right)^{-1} \nabla f(x_k) \geq 0$$

$$f(x_k) - f(x^*) \geq \frac{m}{2} \| x_k - x^* \|^2 \geq 0$$

Since x^* is a unique minimizer it must be that $x_k \to x^*$ as $k \to \infty$

$\square$