

Optimization Algorithms: Theory and Software Implementation

Prof. Thirumulanathan D

Department of Mathematics

Institute of IIT Kanpur

Lecture: 4

Hello, everyone. So, this is the fourth lecture in the course. Recall that in the previous lecture, we were trying to minimize a function: $f(x_1, x_2, \dots, x_n)$, where f is a function of n independent variables. Each x_i can take any value in \mathbb{R} .

To find the minimum, we first find the set of critical points, i.e., points where the gradient of f is zero: $\nabla f = 0$

For each critical point $x^* = (x_1^*, x_2^*, \dots, x_n^*)$, we compute the Hessian matrix $\nabla^2 f(x^*)$. We have the following characterizations:

If $t^\top \nabla^2 f(x^*) t > 0$ for all $t \neq 0$, then x^* is a local minimum.

If $t^\top \nabla^2 f(x^*) t < 0$ for all $t \neq 0$, then x^* is a local maximum.

If there exists t_1, t_2 such that $t_1^\top \nabla^2 f(x^*) t_1 > 0$ and $t_2^\top \nabla^2 f(x^*) t_2 < 0$, then x^* is a saddle point.

If $t^\top \nabla^2 f(x^*) t = 0$ for some $t \neq 0$, further investigation is needed.

Examples

1. Let $n = 2$, and $f(x_1, x_2) = x_1^2 + x_2^2$.

$$\nabla f = \begin{bmatrix} 2x_1 \\ 2x_2 \end{bmatrix} \Rightarrow \nabla f = 0 \text{ at } x_1 = 0, x_2 = 0$$

Hessian:

$$\nabla^2 f(x^*) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

For any $t = (t_1, t_2)$, $t^\top \nabla^2 f(x^*) t = 2t_1^2 + 2t_2^2 > 0$ for all $t \neq 0$. So, $x^* = (0, 0)$ is a **local minimum**.

2. $f(x_1, x_2) = -x_1^2 - x_2^2$

$$\nabla f = \begin{bmatrix} -2x_1 \\ -2x_2 \end{bmatrix} \Rightarrow x^* = (0, 0)$$

$$\nabla^2 f(x^*) = \begin{bmatrix} -2 & 0 \\ 0 & -2 \end{bmatrix}$$

$$t^T \nabla^2 f(x^*) t = -2t_1^2 - 2t_2^2 < 0 \text{ for all } t \neq 0, \text{ Local maximum.}$$

3. $f(x_1, x_2) = x_1^2 - x_2^2$

$$\nabla f = \begin{bmatrix} 2x_1 \\ -2x_2 \end{bmatrix} \Rightarrow x^* = (0, 0)$$

$$\nabla^2 f(x^*) = \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix}$$

$$t^T \nabla^2 f(x^*) t = 2t_1^2 - 2t_2^2 ,$$

Saddle point since this quantity can be positive or negative.

(Refer slide time 10:21)

Lecture 4

$\min_{x \in \mathbb{R}^n} f(x_1, \dots, x_n)$

Examples: (i) $f(x_1, x_2) = x_1^2 + x_2^2$. $C = \{x: \begin{bmatrix} 2x_1 \\ 2x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}\} = \{(0, 0)\}$

$\nabla_2 f(x^*) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$. $t^T \nabla_2 f(x^*) t = [t_1 \ t_2] \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} t_1 \\ t_2 \end{bmatrix} = 2t_1^2 + 2t_2^2 > 0 \ \forall t: [t_1, t_2] \neq (0, 0)$

$\therefore x^* = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ is a local min.

(ii) $f(x_1, x_2) = -x_1^2 - x_2^2$. $C = \{x: \begin{bmatrix} -2x_1 \\ -2x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}\} = \{(0, 0)\}$

$\nabla_2 f(x^*) = \begin{bmatrix} -2 & 0 \\ 0 & -2 \end{bmatrix}$. $t^T \nabla_2 f(x^*) t = -2t_1^2 - 2t_2^2 < 0 \ \forall t: [t_1, t_2] \neq (0, 0)$

$\therefore x^* = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ is a local max.

(iii) $f(x_1, x_2) = x_1^2 - x_2^2$. $C = \{x: \begin{bmatrix} 2x_1 \\ -2x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}\} = \{(0, 0)\}$

$\nabla_2 f(x^*) = \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix}$. $t^T \nabla_2 f(x^*) t = 2t_1^2 - 2t_2^2 \begin{cases} > 0 & |t_1| > |t_2| \\ < 0 & |t_1| < |t_2| \end{cases}$

$\therefore x^* = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ is a saddle point.

Lec-04 Dr.Thiru

4. $f(x_1, x_2) = x_1^2 - x_2^4$

$$\nabla f = \begin{bmatrix} 2x_1 \\ -4x_2^3 \end{bmatrix} \Rightarrow x^* = (0, 0)$$

$$\nabla^2 f(x^*) = \begin{bmatrix} 2 & 0 \\ 0 & -8x_2^2 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}$$

$$t^T \nabla^2 f(x^*) t = 2t_1^2$$

This is ≥ 0 , but equals 0 when $t_1 = 0, t_2 \neq 0$. So, **further probe is needed.**

Consider:

$$f(0, \epsilon) = -\epsilon^4 < 0, \quad f(\epsilon, 0) = \epsilon^2 > 0, f(0, 0) = 0, \text{ Saddle point.}$$

5. $f(x_1, x_2) = x_1^2 + x_2^4$

$$\nabla f = \begin{bmatrix} 2x_1 \\ 4x_2^3 \end{bmatrix} \Rightarrow x^* = (0, 0)$$

$$\nabla^2 f(x^*) = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}$$

$$t^T \nabla^2 f(x^*) t = 2t_1^2 \geq 0$$

Further probe needed. But since $f(x_1, x_2) \geq 0 = f(0, 0)$, **local and global minimum.**

6. $f(x_1, x_2) = -x_1^2 - x_2^4$, Same analysis:

$$\nabla^2 f(x^*) = \begin{bmatrix} -2 & 0 \\ 0 & 0 \end{bmatrix}, \quad t^T \nabla^2 f(x^*) t = -2t_1^2 \leq 0$$

Further probe needed, but since $f(x_1, x_2) \leq 0 = f(0, 0)$, **local maximum.**

(Refer slide time 20:11)

(v) $f(x) = x_1^2 - x_2^4$. $C = \{x: \begin{bmatrix} 2x_1 \\ -4x_2^3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}\} = \{ \begin{bmatrix} 0 \\ 0 \end{bmatrix} \}$
 $\nabla_2 f(x^*) = \begin{bmatrix} 2 & 0 \\ 0 & -12x_2^2 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}$
 $t^T \nabla_2 f(x^*) t = 2t_1^2 \geq 0, \forall t \in \mathbb{R}^2$
 If $t = (0, 1)$, then $t^T \nabla_2 f(x^*) t = 2(0)^2 = 0$.
 \Rightarrow Further probe is needed.
 $f(0, 0) = 0, f(\epsilon, 0) = \epsilon^2 > 0, f(0, \epsilon) = -\epsilon^4 < 0$.
 $\Rightarrow \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ is a saddle point.

(vi) $f(x) = x_1^2 + x_2^4$. $C = \{x: \begin{bmatrix} 2x_1 \\ 4x_2^3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}\} = \{ \begin{bmatrix} 0 \\ 0 \end{bmatrix} \}$
 $\nabla_2 f(x^*) = \begin{bmatrix} 2 & 0 \\ 0 & 12x_2^2 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}$. \Rightarrow Further probe needed.
 $f(x_1, x_2) = x_1^2 + x_2^4 \geq 0 = f(0, 0) \Rightarrow \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ is a local min

(vii) $f(x) = -x_1^2 - x_2^4$. $C = \{x: \begin{bmatrix} -2x_1 \\ -4x_2^3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}\} = \{ \begin{bmatrix} 0 \\ 0 \end{bmatrix} \}$
 $\nabla_2 f(x^*) = \begin{bmatrix} -2 & 0 \\ 0 & -12x_2^2 \end{bmatrix} = \begin{bmatrix} -2 & 0 \\ 0 & 0 \end{bmatrix}$. $t^T \nabla_2 f(x^*) t = -2t_1^2 \leq 0, \forall t \in \mathbb{R}^2$.
 If $t = (0, 1)$, then $t^T \nabla_2 f(x^*) t = -2(0)^2 = 0 \Rightarrow$ Further probe needed.
 $f(x_1, x_2) = -x_1^2 - x_2^4 \leq 0 = f(0, 0) \Rightarrow \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ local max.

Lec-04 Dr.Thiru
 36°C Partly sunny
 Search
 ENG IN
 9:02 AM 5/29/2025

Summary: Unconstrained Optimization

To find local solutions:

1. Find critical points: $\nabla f = 0$

Evaluate the Hessian $\nabla^2 f(x^*)$

3. Analyze $t^\top \nabla^2 f(x^*) t$:

$> 0 \Rightarrow$ local min

$< 0 \Rightarrow$ local max

< 0 and $> 0 \Rightarrow$ saddle point

$= 0$ for some $t \Rightarrow$ further analysis required

Constrained Optimization

We now consider:

$$\text{Min } f(x) \quad x \in S$$

If $S \neq \mathbb{R}^n$, this is a **constrained optimization**.

For example, if $S = Z$, standard calculus-based methods do not work. We consider:

$$\text{Min } f(x) \quad \text{subject to } g_i(x) \leq 0 \text{ for } i = 1, \dots, p \quad \text{and } h_j(x) = 0 \text{ for } j = 1, \dots, m$$

Example from Lecture 1: Consumer Utility Maximization

Max $u(x_1, \dots, x_n)$ subject to

$$p_1 x_1 + \dots + p_n x_n \leq w, \quad x_i \geq 0 \text{ for all } i$$

We rewrite this as: min $-u(x_1, \dots, x_n)$.

Subject to:

$$\begin{aligned} p_1 x_1 + \dots + p_n x_n - w &\leq 0 \\ -x_1 &\leq 0 \\ &\dots \\ -x_n &\leq 0 \end{aligned}$$

So, now you can see that this problem is actually equivalent to what you have written here, where $f(x)$.

So, in this problem, f is actually $-u$, and $g_1 = p_1x_1 + \dots + p_nx_n - w$, so that is less than or equal to 0, and

$$g_2 = -x_1, \quad g_3 = -x_2, \quad \dots, \quad g_{n+1} = -x_n.$$

So, this is a problem with $p = n+1$ and $m = 0$, because you do not have any equality constraints. So yeah, this actually defines—as you all know—it actually defines a constraint set.

(Refer slide time 29:49)

Constrained Optimization

$\min_{x \in S} f(x)$

"Constrained" if $S \neq \mathbb{R}^n$.

$\min_x f(x)$

s.t. $g_i(x) \leq 0 \quad \forall i = 1, 2, \dots, p$

$h_j(x) = 0 \quad \forall j = 1, 2, \dots, m$

$\max_x u(x_1, x_2, \dots, x_n)$

s.t. $p_1x_1 + p_2x_2 + \dots + p_nx_n \leq w$

$x_i \geq 0 \quad \forall i = 1, \dots, n.$

$\min_x -u(x_1, x_2, \dots, x_n)$

s.t. $p_1x_1 + p_2x_2 + \dots + p_nx_n \leq w$

$-x_1 \leq 0$

$-x_2 \leq 0$

\vdots

$-x_n \leq 0$

\Rightarrow

$f = -u$

$g_1 = p_1x_1 + \dots + p_nx_n - w$

$g_2 = -x_1$

$g_3 = -x_2$

\vdots

$g_{n+1} = -x_n$

In the next lecture, we will learn how to define a constraint set and find the solution of a constrained optimization problem when it is of this particular form:

Min $f(x)$

subject to $g_i(x) \leq 0$ for $i = 1, \dots, p$, and $h_j(x) = 0$ for $j = 1, \dots, m$.

Thankyou