

# Lecture 2

It is a small *de tour* but it is important to understand this before we move to calculus of variations.

## Necessary and Sufficient Conditions for Finite-variable Unconstrained Minimization

---

ME 256, Indian Institute of Science

**Calculus of Variations and Structural Optimization**

**G. K. Ananthasuresh**

Professor, Mechanical Engineering, Indian Institute of Science, Bangalore

[suresh@mecheng.iisc.ernet.in](mailto:suresh@mecheng.iisc.ernet.in)

# Outline of the lecture

## Necessary conditions for unconstrained optimization problem

- In one variable
- In two variables
- In multiple variables

## What we will learn:

The concept of a local minimum

The premise for writing the necessary condition

The concept of gradient of a function of  $n$  variables

And, of course, the necessary conditions of unconstrained optimization problem

# Global and local minima: definitions

Simple case:  $f(x)$ , a function of a single variable,  $x$ .

## Global minimum

$x^*$  is global minimizer of  $f(x)$  if  $f(x^*) \leq f(x) \forall x$  in the feasible interval of  $x$ .

## Local minimum

$x^*$  is a local minimizer of  $f(x)$  if  $f(x^*) \leq f(x)$  in a small neighborhood of  $x^*$  in the feasible interval of  $x$ .

$N = \text{small neighborhood} = \left\{ x \mid x \in S \text{ with } \|x - x^*\| < \delta \right\}, \delta > 0$

These are just definitions; not conditions.

Definitions of this sort do not let you check if a given value of  $x$  is a minimum or not unless you exhaustively check the entire domain of  $x$ .

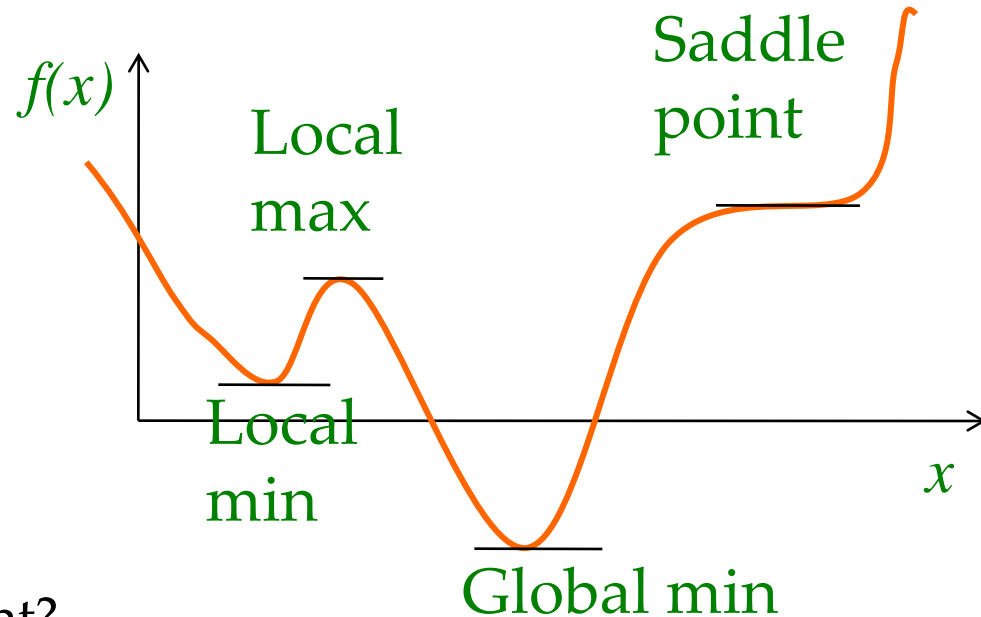
A condition would let you check this easily.

# Conditions for a local minimum of $f(x)$

If  $x^*$  is a local minimum of  $f(x)$ , then...

$$\left. \frac{df}{dx} \right|_{x^*} = 0 \quad \text{Necessary condition}$$

$$\left. \frac{df}{dx} \right|_{x^*} = 0 \quad \& \quad \left. \frac{d^2 f}{dx^2} \right|_{x^*} > 0 \quad \text{Sufficient condition}$$



Why is the necessary condition not sufficient?

Is the sufficient condition also necessary?

Think about the literal meaning of “necessary” and “sufficient”.

# Why is necessary condition “necessary”?

Consider Taylor series expansion of  $f(x)$  around  $x^*$ .

$$f(x) = f(x^*) + \left. \frac{df}{dx} \right|_{x^*} (x - x^*) + \frac{1}{2} \left. \frac{d^2 f}{dx^2} \right|_{x^*} (x - x^*)^2 + O(3)$$

Perturbation  
around  $x^*$

$$\delta x^* = (x - x^*)$$

Zeroth  
order  
term

First  
order  
term

Second  
order  
term

Higher  
order  
terms

$$f(x) = f(x^*) + f'(x^*) \Delta x^* + \frac{1}{2} f''(x^*) (\Delta x^*)^2 + O(3)$$

$x^*$  is a local minimizer of  $f(x)$  if  $f(x^*) \leq f(x)$  in a small neighborhood of  $x^*$  in the feasible interval of  $x$ .

Here the small neighborhood is  $\Delta x^*$ . When it is small, it is the first order term that matters more than the second order term.

Unless  $f'(x^*)$  is zero, we cannot be sure that the definition is satisfied. More in the next slide.

# Why is necessary condition *necessary*? (contd.)

For small  $\Delta x^*$  (as small as you can imagine...)

$$f(x) = f(x^*) + \underbrace{f'(x^*)\Delta x^*}_{\text{May be positive or negative depending on the sign of } f'(x^*) \text{ as } \Delta x^* \text{ can be positive or negative.}} + \underbrace{\frac{1}{2}f''(x^*)(\Delta x^*)^2}_{\text{Negligible}} + O(3)$$

May be positive  
or negative  
depending on the  
sign of  $f'(x^*)$  as  $\Delta x^*$   
can be positive or  
negative.

Negligible

So, for  $f(x^*) \leq f(x)$  it is **necessary** to have  $f'(x^*) = 0$

# Why is necessary condition not sufficient?

Note that we only talk about the function “not being less in the small neighborhood of the minimizing point” to say that it has a local minimum.

So, the condition is that the first order term is zero for any small perturbation. This necessitates the first order derivative to be zero.

This condition is necessary, as noted in the previous slide.

But...

The necessary condition is true for a local minimum and a local maximum. So, it is not sufficient to conclude that a given value of  $x$  is a local minimum.

# Is sufficient condition also necessary?

Consider...

$$f(x) = x^4$$

$$f'(x) = 4x^3 \Rightarrow x^* = 0$$

$$f''(x) = 12x^2 = 0$$

$$f'''(x) = 24x = 0 \Rightarrow x^* = 0$$

$$f^{iv}(x) = 24 > 0$$

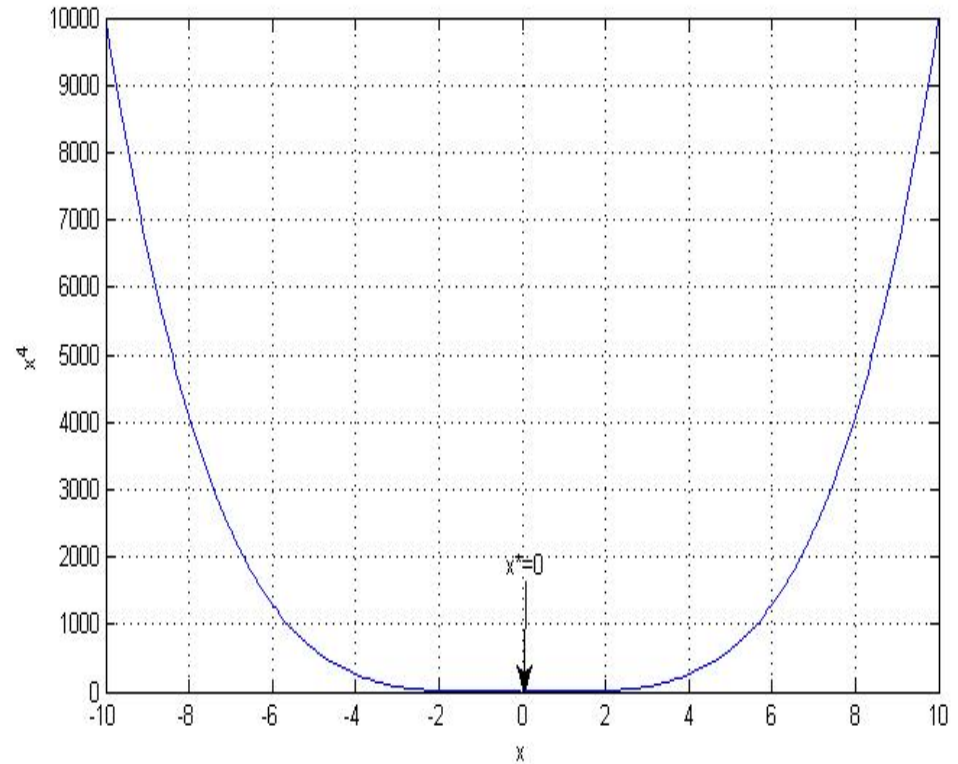
$$x^* = 0$$

$$\left. \frac{df}{dx} \right|_{x^*} = 0$$

Necessary  
condition  
is satisfied.

$$\left. \frac{d^2 f}{dx^2} \right|_{x^*} = 0$$

Sufficient  
condition is **not**  
satisfied.



But  $x^* = 0$ , is a minimizer here! So,  
sufficient condition is not  
necessary.



# Understand “necessary” and “sufficient” well.

The logic of necessary and sufficient conditions should be clearly understood.

They actually mean what they say but it can be confusing and misleading sometimes.

What is necessary may not be sufficient.

What is sufficient may not be necessary.

Sometimes, a condition can be necessary and sufficient.

Note all of this we are saying only in the context of a local minimum.

For a global minimum, there is no “operationally useful” definition or condition.

Let us move to a function of more than one variable next...

# A function of two variables...

Taylor series expansion of  $f(x,y)$  around  $(x^*,y^*)$ :

$$f(x,y) = f(x^*,y^*) + f_x(x^*,y^*)\Delta x^* + f_y(x^*,y^*)\Delta y^* + \\ + \frac{1}{2} \left\{ f_{xx}(x^*,y^*)(\Delta x^*)^2 + 2f_{xy}(x^*,y^*)\Delta x^*\Delta y^* + f_{yy}(x^*,y^*)(\Delta y^*)^2 \right\} + O(3)$$

where

$$f_x = \frac{\partial f}{\partial x}; f_y = \frac{\partial f}{\partial y}; f_{xx} = \frac{\partial^2 f}{\partial x^2}; f_{xy} = \frac{\partial^2 f}{\partial x \partial y}; f_{yy} = \frac{\partial^2 f}{\partial y^2}$$

In matrix form

$$f(x,y) = f(x^*,y^*) + \begin{Bmatrix} f_x(x^*,y^*) & f_y(x^*,y^*) \end{Bmatrix} \begin{Bmatrix} \Delta x^* \\ \Delta y^* \end{Bmatrix} + \\ + \frac{1}{2} \begin{Bmatrix} \Delta x^* & \Delta y^* \end{Bmatrix} \begin{bmatrix} f_{xx}(x^*,y^*) & f_{xy}(x^*,y^*) \\ f_{xy}(x^*,y^*) & f_{yy}(x^*,y^*) \end{bmatrix} \begin{Bmatrix} \Delta x^* \\ \Delta y^* \end{Bmatrix} + O(3)$$

# Two-variable function (contd.)

$$\begin{aligned} f(x, y) &= f(x^*, y^*) + \begin{Bmatrix} f_x(x^*, y^*) & f_y(x^*, y^*) \end{Bmatrix} \begin{Bmatrix} \Delta x^* \\ \Delta y^* \end{Bmatrix} + \\ &+ \frac{1}{2} \begin{Bmatrix} \Delta x^* & \Delta y^* \end{Bmatrix} \begin{bmatrix} f_{xx}(x^*, y^*) & f_{xy}(x^*, y^*) \\ f_{xy}(x^*, y^*) & f_{yy}(x^*, y^*) \end{bmatrix} \begin{Bmatrix} \Delta x^* \\ \Delta y^* \end{Bmatrix} + O(3) \\ &= f(x^*, y^*) + \nabla f(x^*, y^*)^T \begin{Bmatrix} \Delta x^* \\ \Delta y^* \end{Bmatrix} + \frac{1}{2} \begin{Bmatrix} \Delta x^* & \Delta y^* \end{Bmatrix} \mathbf{H}(x^*, y^*) \begin{Bmatrix} \Delta x^* \\ \Delta y^* \end{Bmatrix} + O(3) \end{aligned}$$

where  $\nabla f(x^*, y^*) = \begin{Bmatrix} f_x(x^*, y^*) \\ f_y(x^*, y^*) \end{Bmatrix}$  and  $\mathbf{H}(x^*, y^*) = \begin{bmatrix} f_{xx}(x^*, y^*) & f_{xy}(x^*, y^*) \\ f_{xy}(x^*, y^*) & f_{yy}(x^*, y^*) \end{bmatrix}$

**Gradient** **Hessian**

# Necessary conditions for the minimum of $f(x,y)$

Second order term is negligible and the first order term should be zero for small  $(\delta x^*, \delta y^*)$

So,

$$\nabla f(x^*, y^*) = \begin{Bmatrix} f_x(x^*, y^*) \\ f_y(x^*, y^*) \end{Bmatrix} = \begin{Bmatrix} 0 \\ 0 \end{Bmatrix} \quad \text{This has two scalar equations in it.}$$

Two variables  $(x^*, y^*)$  to be found using two scalar equations!

$$f_x(x^*, y^*) = 0$$

$$f_y(x^*, y^*) = 0$$

# Sufficient condition for the minimum of $f(x,y)$

$$\frac{1}{2} \begin{Bmatrix} \Delta x^* & \Delta y^* \end{Bmatrix} \begin{bmatrix} f_{xx}(x^*, y^*) & f_{xy}(x^*, y^*) \\ f_{xy}(x^*, y^*) & f_{yy}(x^*, y^*) \end{bmatrix} \begin{Bmatrix} \Delta x^* \\ \Delta y^* \end{Bmatrix} > 0 \text{ for any } (\Delta x^*, \Delta y^*)$$

$$\frac{1}{2} \begin{Bmatrix} \Delta x^* & \Delta y^* \end{Bmatrix} \mathbf{H}(x^*, y^*) \begin{Bmatrix} \Delta x^* \\ \Delta y^* \end{Bmatrix} > 0 \text{ for any } (\Delta x^*, \Delta y^*)$$



A matrix that has this property is said to be **positive definite**.

# Positive definite, and other definite things...

For any  $(\Delta x^*, \Delta y^*)$

$$\frac{1}{2} \begin{Bmatrix} \Delta x^* & \Delta y^* \end{Bmatrix} \mathbf{H}(x^*, y^*) \begin{Bmatrix} \Delta x^* \\ \Delta y^* \end{Bmatrix} > 0$$

Positive definite  $\mathbf{H}$ ; minimum.

$$\frac{1}{2} \begin{Bmatrix} \Delta x^* & \Delta y^* \end{Bmatrix} \mathbf{H}(x^*, y^*) \begin{Bmatrix} \Delta x^* \\ \Delta y^* \end{Bmatrix} \geq 0$$

Positive semi-definite  $\mathbf{H}$ ; minimum or flat.

$$\frac{1}{2} \begin{Bmatrix} \Delta x^* & \Delta y^* \end{Bmatrix} \mathbf{H}(x^*, y^*) \begin{Bmatrix} \Delta x^* \\ \Delta y^* \end{Bmatrix} < 0$$

Negative definite  $\mathbf{H}$ ; maximum.

$$\frac{1}{2} \begin{Bmatrix} \Delta x^* & \Delta y^* \end{Bmatrix} \mathbf{H}(x^*, y^*) \begin{Bmatrix} \Delta x^* \\ \Delta y^* \end{Bmatrix} \leq 0$$

Negative semi-definite  $\mathbf{H}$ ; maximum or flat.

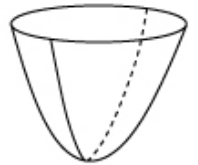
$$\frac{1}{2} \begin{Bmatrix} \Delta x^* & \Delta y^* \end{Bmatrix} \mathbf{H}(x^*, y^*) \begin{Bmatrix} \Delta x^* \\ \Delta y^* \end{Bmatrix} = 0$$

Null-definite  $\mathbf{H}$ ; just flat; neither minimum nor maximum.

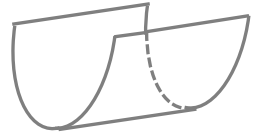
# Peaks, valleys, folds, ridges, and flat planes...

The surface represented by  $f(x,y)$  locally looks like this at  $(x^*,y^*)$ .

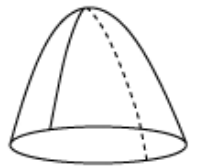
Positive definite  $\mathbf{H}$ ; minimum; **Bottom of a valley**



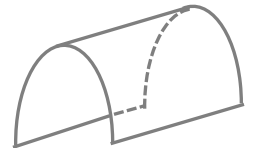
Positive semi-definite  $\mathbf{H}$ ; minimum or flat; **A valley fold**



Negative definite  $\mathbf{H}$ ; maximum; **Peak of a hill**



Negative semi-definite  $\mathbf{H}$ ; maximum or flat; **A ridge**



Null-definite  $\mathbf{H}$ ; neither minimum nor maximum; **Just flat**



# Indefiniteness of a matrix

$$\frac{1}{2} \begin{Bmatrix} \Delta x^* \\ \Delta y^* \end{Bmatrix} \mathbf{H}(x^*, y^*) \begin{Bmatrix} \Delta x^* \\ \Delta y^* \end{Bmatrix} \begin{matrix} > \\ ? 0 \text{ for any } (\Delta x^*, \Delta y^*) \\ < \end{matrix}$$

It is positive sometimes and negative sometimes... it is **indefinite**. Then,  $(x^*, y^*)$  is a **saddle point**.

<http://explore.org/photos/2238/horse-saddle>



Minimum one way  
and maximum the  
other way!

<http://www.pringles.com/products>





# Function of $n$ variables

Taylor's series expansion...

$$\begin{aligned} f(\mathbf{x}) &= f(\mathbf{x}^*) + \sum_{i=1}^n \frac{\partial f(\mathbf{x}^*)}{\partial x_i} (x_i - x_i^*) + \\ &\frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2 f(\mathbf{x}^*)}{\partial x_i \partial x_j} (x_i - x_i^*) (x_j - x_j^*) + O(3) \\ &= f(\mathbf{x}^*) + \nabla f^T(\mathbf{x}^*) \Delta \mathbf{x}^* + \frac{1}{2} \Delta \mathbf{x}^{*T} \mathbf{H}(\mathbf{x}^*) \Delta \mathbf{x}^* + O(3) \end{aligned}$$

# Gradient of an $n$ -variable function

$$\nabla f(\mathbf{x}^*) = \left\{ \begin{array}{c} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \vdots \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{array} \right\} = \left\{ \begin{array}{c} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{array} \right\}$$

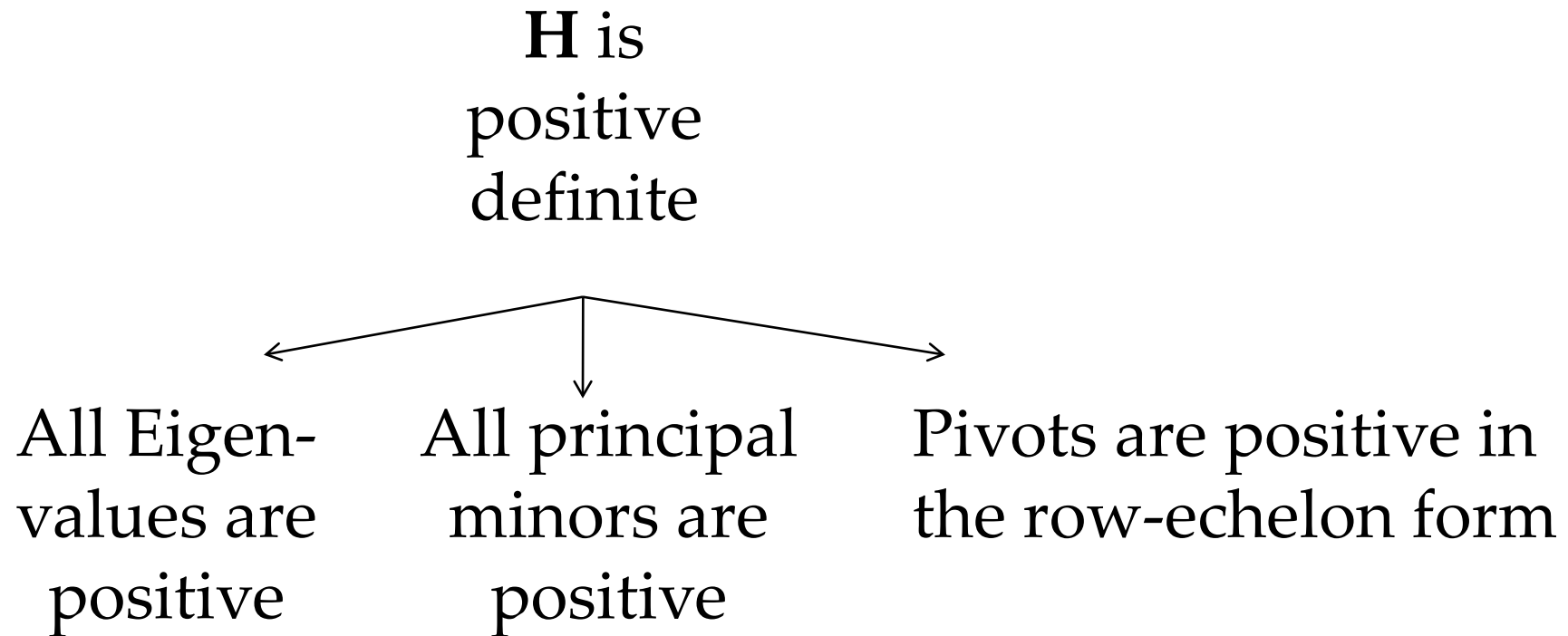
Necessary  
condition:  
 $n$  variables;  
 $n$  equations.

# Hessian of an $n$ -variable function

$$\mathbf{H}(\mathbf{x}) = \left\{ \begin{array}{cccc} \frac{d^2 f}{dx_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \dots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{d^2 f}{dx_2^2} & \dots & \vdots \\ \vdots & \dots & \dots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \dots & \dots & \frac{\partial^2 f}{\partial x_n \partial x_n} \end{array} \right\}$$

Sufficient  
condition:  
**H** should be  
positive  
definite for a  
minimum of  
 $f(\mathbf{x})$

# Rules of checking positive definiteness



# For other “definitenesses” ...

Quadratic form $\delta \mathbf{x}^{*T} \mathbf{H}(\mathbf{x}^*) \delta \mathbf{x}^*$	$\mathbf{H}$	Eigen values of $\mathbf{H}$	Nature of $\mathbf{x}^*$
Positive	Positive definite	All are positive	Local min
Negative	Negative definite	All are negative	Local max
Non-negative	Positive semi-definite	Some zero, others positive	A valley fold
Non-positive	Negative semi-definite	Some zero, others negative	A ridge
Any sign	Indefinite	Mixed signs	Saddle point

# The end note

Unconstrained finite-variable optimization

Only local minimum can have conditions that can easily checked.  
Global minimum does not have “operationally useful” definition or conditons.

Necessary condition: first order derivative is zero.

Sufficient condition: second order derivative is positive (or positive definite)

Gradient  
Hessian

In two variables, we have peaks, valleys, fold, ridges, and flat planes...

Rules for checking positive definiteness of a matrix.

Thanks