IIT KHARAGPUR

NPTEL ONLINE CERTIFICATION COURSES

OPTIMIZATION IN CHEMICAL ENGINEERING

Basic Concepts of Optimization – I

(WEEK – 3: LECTURE – 11)

Dr. Debasis Sarkar

Department of Chemical Engineering

Indian Institute of Technology Kharagpur

Basic Concepts of Optimization – I

Week 3:

- Continuity of functions
- Unimodal and multimodal functions
- Convex and concave functions
- Optimality criteria for unconstrained single variable functions
- Optimality criteria for unconstrained multi-variable functions

Today's Topic:

- Continuity of functions
- Unimodal and multimodal functions
- Positive/Negative definite matrix

Continuity of Functions

It is convenient to work with functions that are continuous and have continuous derivatives.

Definition: A function of a single variable x is continuous at a point x_o if

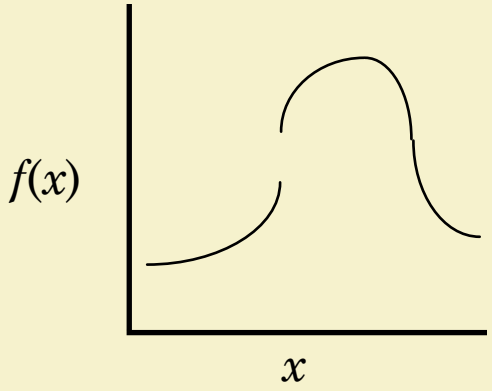
$$\text{(a)} \quad f(x_o) \quad \text{exists}$$

$$\text{(b)} \quad \lim_{x \to x_o} f(x) \quad \text{exists}$$

$$\text{(c)} \quad \lim_{x \to x_o} f(x) = f(x_o)$$

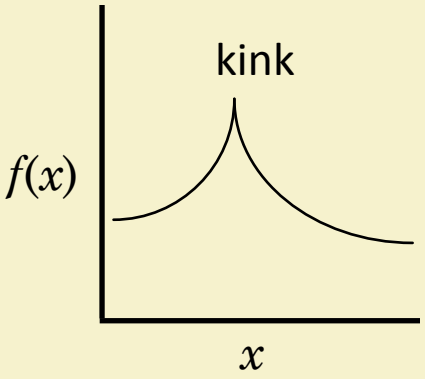
If $f(x)$ is continuous at every point in a region R, then $f(x)$ is said to be continuous throughout R.

Continuity of Functions



$f(x)$ is discontinuous

If the decision variables take only discrete values, the objective function will be discontinuous. Example: cost per unit diameter of pipe.



$f(x)$ is continuous, but

$$f'(x) \equiv \frac{df}{dx}(x)$$ is not

Presence of "kink" may cause difficulty in optimization algorithms that use derivative information.

Continuity of Functions

Is $f(x) = 1/x$ continuous? What about derivative of $f(x)$?

Solution:

$f(x) = 1/x$ is continuous except at $x = 0$ because $f(0)$ is not defined.

$$f'(x) = \frac{df(x)}{dx} = -\frac{1}{x^2}$$ is continuous except at $x = 0$

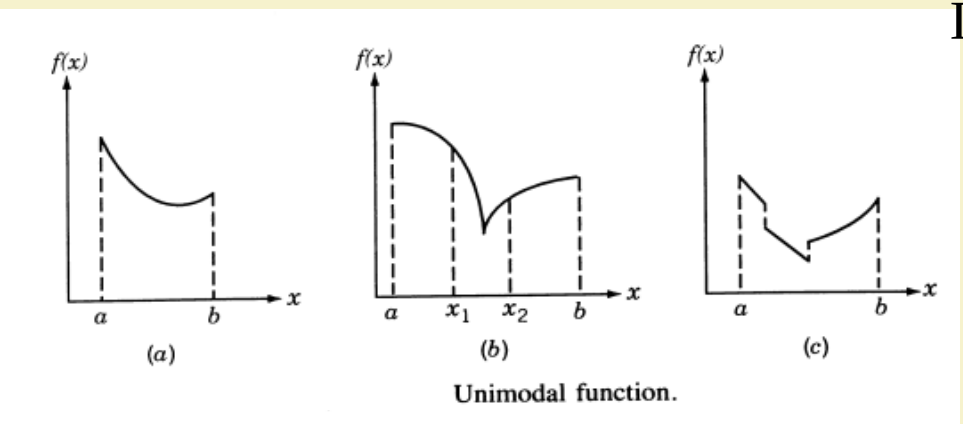
Another example: $f(x) = \ln(x)$ is continuous for $x > 0$. $\ln(x)$ is not defined for $x \leq 0$.

Unimodal and Multimodal Functions

Objective functions may be unimodal or multimodal.

Unimodal objective functions – have only one optimum in a given interval

Multimodal objective functions – have more than one optimum in the given interval

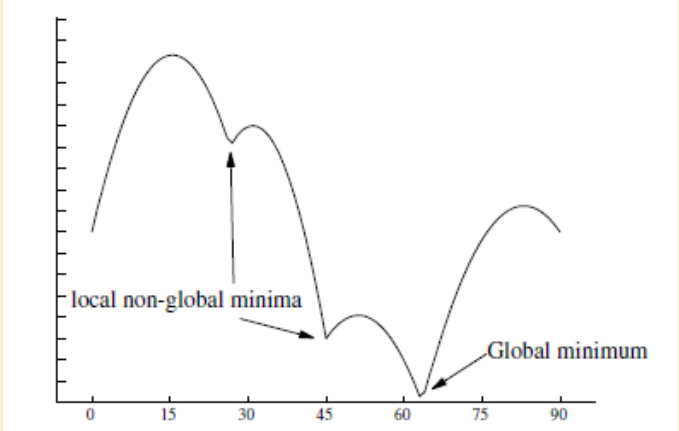


Unimodal function.

If $f'(x) = 0$ at the extremum

⇓

Stationary point



Multivariate Functions: Contour Plots

Consider a scalar function $f(\mathbf{x})$ of n variables $(x_1,...,x_n)$. The function $f(x)$ is univariate if n = 1. The function $f(x)$ is multivariate if n = 2 or higher.

For a multivariate function, the equation $z = f(\mathbf{x})$ defines a **surface** in $(n+1)$ dimensional space.

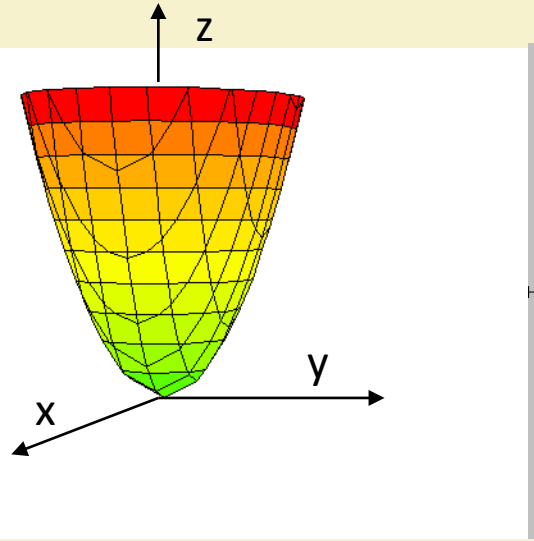
In the case $n = 2$, the points $z = f(x_1,x_2)$ represent a three dimensional surface.

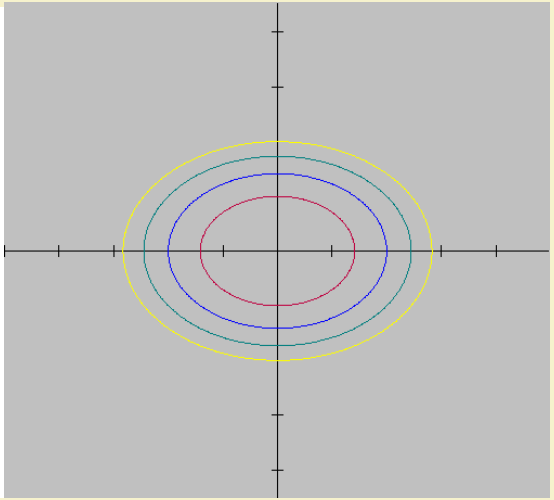
Let C be a particular value of $f(x_1,x_2)$. Then $f(x_1,x_2) = C$ defines a curve in x_1 and x_2 on the plane $z = C$. This is known as **Level Curve.**

If we consider different values of C, we obtain a family of level curves which provide a **Contour Map** of the function $z = f(x_1,x_2)$.

Multivariate Functions: Level Curves: Contour Plots

Sketch a contour map of the function using the level curves at $C = 0, 2, 4$, etc.





$$z = f(x,y) = x^2 + 2y^2$$

$C = 0$: $0 = x^2 + 2y^2 \rightarrow$ Point at $(0,0)$

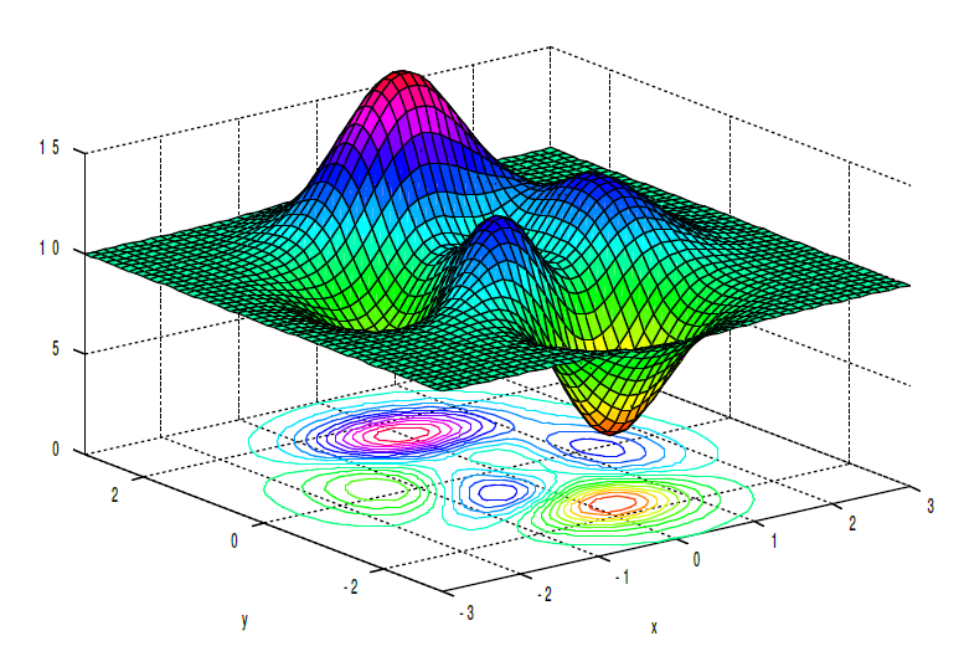
$C = 2$: $2 = x^2 + 2y^2 \rightarrow$ Ellipse: $\frac{x^2}{2} + \frac{y^2}{1} = 1$

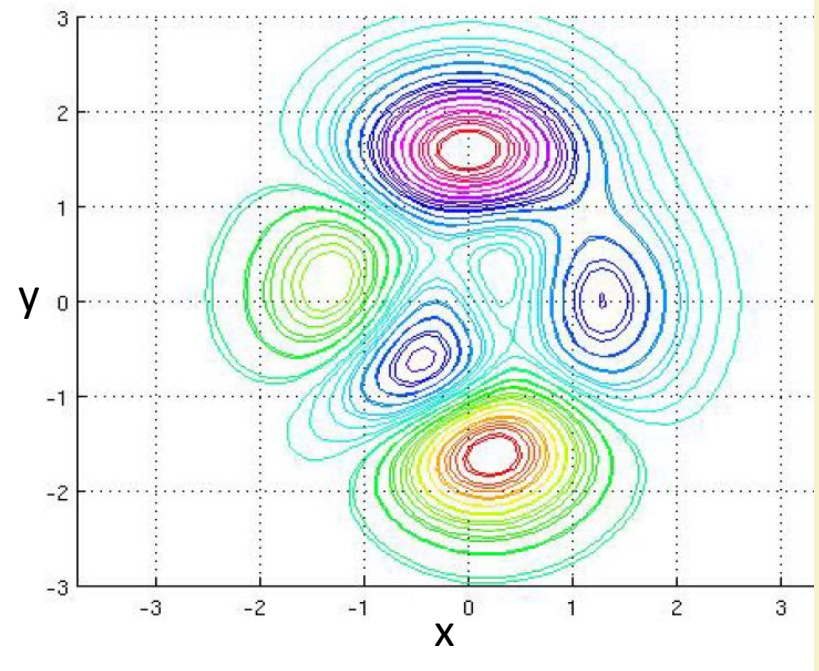
$C = 4$: $4 = x^2 + 2y^2 \rightarrow$ Ellipse: $\frac{x^2}{4} + \frac{y^2}{2} = 1$

Contour map with point (0, 0) for $C = 0$ and then ellipses expanding out from the center for the remaining values of C.

Multivariate Functions: Contour Plots

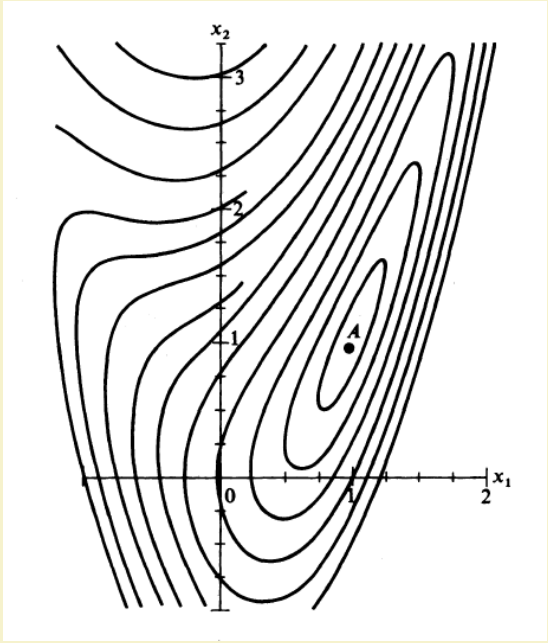
A *contour plot* is a nice way to visualize the graph of a function of two variables

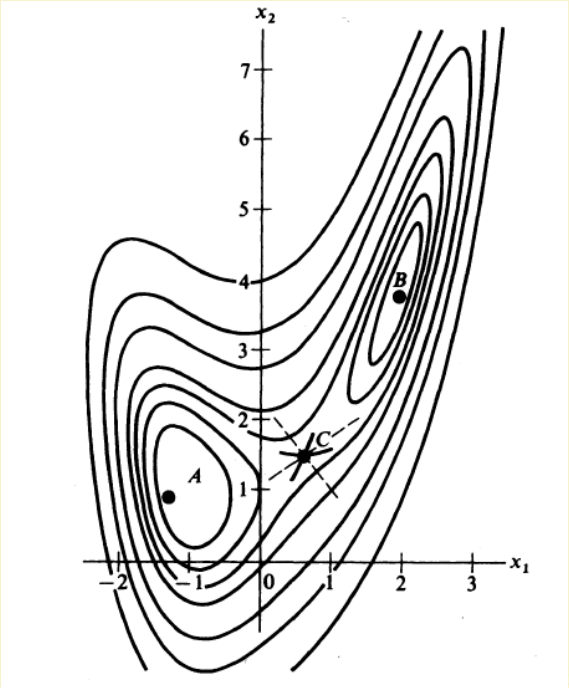




Multivariate Functions: Contour Plots

A *contour plot* is a nice way to visualize the graph of a function of two variables





Positive Definite Matrix

A real symmetric matrix **A** is **positive definite** if

$$\mathbf{x}^T\mathbf{A}\mathbf{x} > 0, \ \forall\ \mathbf{x} \neq \mathbf{0}$$

Equivalent conditions

(a) $\mathbf{x}^T\mathbf{A}\mathbf{x} > 0, \ \forall\ \mathbf{x} \neq \mathbf{0}$.
(b) **A** has only positive eigenvalues.
(c) Leading principal sub-matrices of **A** have positive determinants.
(d) **A** has only positive pivots.

Test for positive definiteness

Given a symmetric two by two matrix

$$\begin{bmatrix} a & b \\ b & c \end{bmatrix}$$

$x^T Ax$ is positive except when $x = 0$

Eigenvalue test: $\lambda_1 > 0, \lambda_2 > 0$.

Determinants test: $a > 0, ac - b^2 > 0$.

Pivot test: $a > 0, \dfrac{ac - b^2}{a} > 0$.

Negative Definite Matrix

A is **negative definite** if

$$\mathbf{x}^T\mathbf{A}\mathbf{x} < 0, \ \forall\ \mathbf{x} \neq \mathbf{0}$$

Equivalent conditions

(a) $\mathbf{x}^T\mathbf{A}\mathbf{x} < 0, \ \forall\ \mathbf{x} \neq \mathbf{0}$.

(b) **A** has only negative eigenvalues.

(c) Leading principal sub-matrices have alternating determinants.

(d) **A** has only negative pivots. $|A_1| < 0, \ |A_2| > 0, \ |A_3| < 0, \ etc.$

Positive, Negative, Semi-definite, Indefinite Matrix

A symmetric matrix A is

1. positive definite if $x'Ax > 0$ for all $x \neq 0$
2. negative definite if $x'Ax < 0$ for all $x \neq 0$
3. positive semidefinite if $x'Ax \geq 0$ for all x
4. negative semidefinite if $x'Ax \leq 0$ for all x
5. indefinite if it is neither positive nor negative semidefinite (i.e. if $x'Ax > 0$ for some x and $x'Ax < 0$ for some x).

Example

Let $A = \begin{bmatrix} 5 & 2 \\ 2 & 3 \end{bmatrix}$

Its eigenvalues are the solutions to:

$$|A - \lambda I| = \lambda^2 - 8\lambda + 11 = 0,$$

$$4 \pm \sqrt{5}.$$

A is a positive definite matrix.

The pivots of this matrix are 5 and $(\det A)/5 = 11/5$.

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 3 & 0 \\ 0 & 0 & 5 \end{pmatrix}$$

$$\begin{vmatrix} 1-\lambda & 1 & 0 \\ 1 & 3-\lambda & 0 \\ 0 & 0 & 5-\lambda \end{vmatrix} = (5-\lambda)(\lambda^2 - 4\lambda + 2) = 0$$

$\therefore \lambda = 5$ and $\lambda = 2 \pm \sqrt{2}$.

A is a positive definite matrix.

Example

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$$

Eigenvalues are 3 and -1

Matrix A is indefinite

$$A = \begin{pmatrix} 2 & -1 & b \\ -1 & 2 & -1 \\ b & -1 & 2 \end{pmatrix}$$

For what values of b, matrix A will be positive semidefinite?

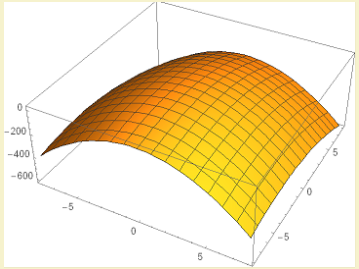
Solution: If A is positive semidefinite, $\det(A) \geq 0$

$$\det(A) = (-2b^2 + 2b + 4) \geq 0$$

$$\Rightarrow (b-2)(b+1) \leq 0$$

$$\Rightarrow -1 \leq b \leq 2$$

Optimization in Chemical Engineering



Thank You

IIT KHARAGPUR

NPTEL ONLINE CERTIFICATION COURSES

OPTIMIZATION IN CHEMICAL ENGINEERING

Basic Concepts of Optimization – I

(WEEK – 3: LECTURE – 12)

Dr. Debasis Sarkar

Department of Chemical Engineering

Indian Institute of Technology Kharagpur

Basic Concepts of Optimization – I

Week 3:

- ✓ Continuity of functions
- ✓ Unimodal and multimodal functions
- Convex and concave functions
- Optimality criteria for unconstrained single variable functions
- Optimality criteria for unconstrained multi-variable functions

Today's Topic:

- Gradient vector, Hessian matrix
- Convex and concave functions

Gradient Vector

Consider a function of n variables: $f(\mathbf{x}) = f(x_1, x_2, \ldots, x_n)$. Here, $\mathbf{x} = [x_1, x_2, \ldots\ldots, x_n]^T$

The slope of $f(\mathbf{x})$ at a point $\mathbf{x} = \overline{\mathbf{x}}$ is: $\left.\frac{\partial f(\mathbf{x})}{\partial x_i}\right|_{\mathbf{x}=\overline{\mathbf{x}}}$

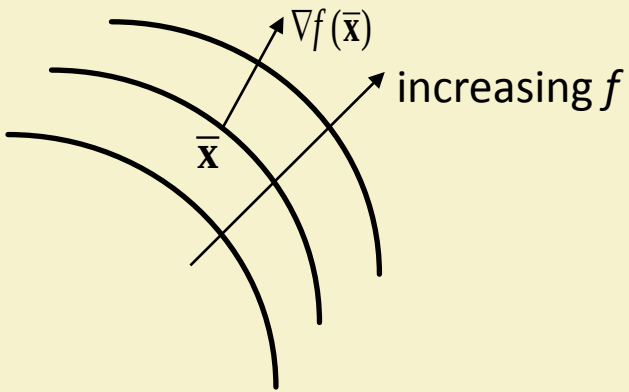
The n-vector of these partial derivatives is called the gradient vector of function f and denoted by ∇f

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

$$\nabla f(\mathbf{x}) \equiv \begin{bmatrix} \frac{\partial f(\mathbf{x})}{\partial x_1} \\ \vdots \\ \frac{\partial f(\mathbf{x})}{\partial x_n} \end{bmatrix}$$ a column vector

Gradient Vector

The gradient vector of function $f(\mathbf{x})$ at a point $\mathbf{X} = \overline{\mathbf{X}}$ is normal to the contour through that point in the direction of increasing f.



At every point, the gradient vector points in a direction where the function value increases most rapidly.

A point $\mathbf{x}^*$ is called a stationary point or critical point if $\nabla f(\mathbf{x}^*) = 0$ at $\mathbf{x}^*$.

These stationary or critical points may be extremum. In other words, these are the likely candidates for maximum and minimum.

Example

Find the gradient vector of $f(x,y) = 3x^2 - 5y^2$ at the point P(2, -3).

Solution:

$$\nabla f(x,y) \equiv \begin{bmatrix} \frac{\partial f(x,y)}{\partial x} \\ \\ \frac{\partial f(x,y)}{\partial y} \end{bmatrix} = \begin{bmatrix} 6x \\ \\ -10y \end{bmatrix}$$

The gradient at P (2, -3) is:

$$\nabla f(x,y)_{2,-3} = \begin{bmatrix} 6x \\ \\ -10y \end{bmatrix}_{2,-3} = \begin{bmatrix} 12 \\ \\ 30 \end{bmatrix}$$

NOTE: If $f(\mathbf{x})$ is linear function, $\nabla f(\mathbf{x})$ is constant. $f(x) = c^T x + \alpha \Rightarrow \nabla f(x) = c$

Hessian Matrix

Consider any n-variable function. The following n^2 second partial derivatives can be written:

$$\frac{\partial}{\partial x_i}\left(\frac{\partial f(\mathbf{x})}{\partial x_j}\right), \quad i=1,\cdots,n; \quad j=1,\cdots,n \quad \Rightarrow \quad \frac{\partial^2 f(\mathbf{x})}{\partial x_i \partial x_j}, \quad i \neq j; \quad \frac{\partial^2 f(\mathbf{x})}{\partial x_i^2}, \quad i = j.$$

Hessian matrix for an n-variable function, $f(x_1, x_2, \ldots, x_n)$:

$$\mathbf{H} = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix}$$

These n^2 second partial derivatives are represented by a square, symmetric matrix, known as Hessian matrix, H.

Hessian Matrix

Hessian is always a symmetric matrix. Since $f(\mathbf{x})$ is assumed to be continuously differentiable, the cross partial derivatives are equal:

$$\frac{\partial^2 f(\mathbf{x})}{\partial x_i \partial x_j} = \frac{\partial^2 f(\mathbf{x})}{\partial x_j \partial x_i},\ i = 1,\ldots,n;\ j = 1,\ldots,n$$

$$\mathbf{H}(\mathbf{x}) \equiv \nabla^2 f(\mathbf{x}) \equiv \begin{bmatrix} \frac{\partial^2 f(\mathbf{x})}{\partial x_1^2} & \cdots & \frac{\partial^2 f(\mathbf{x})}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f(\mathbf{x})}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 f(\mathbf{x})}{\partial x_n^2} \end{bmatrix}$$

For a quadratic function $f(\mathbf{x})$, the Hessian matrix is a constant matrix:

$$f(\mathbf{x}) = \tfrac{1}{2}\mathbf{x}^T\mathbf{H}\mathbf{x} + \mathbf{c}^T\mathbf{x} + \alpha$$

$$\Rightarrow \nabla f(\mathbf{x}) = \mathbf{H}\mathbf{x} + \mathbf{c}$$

$$\Rightarrow \nabla^2 f(\mathbf{x}) = \mathbf{H}$$

Gradient Vector: Hessian Matrix: Example

Find the Gradient Vector and Hessian Matrix for: $f(x_1, x_2) = (x_1 - 2)^2 + 2(x_2 - 1)^2$

Solution:

Gradient Vector:

$$\nabla f(\mathbf{x}) \equiv \begin{bmatrix} \dfrac{\partial f(\mathbf{x})}{\partial x_1} \\ \\ \dfrac{\partial f(\mathbf{x})}{\partial x_2} \end{bmatrix} = \begin{bmatrix} 2x_1 - 4 \\ \\ 4x_2 - 4 \end{bmatrix}$$

Hessian Matrix:

$$\mathbf{H}(\mathbf{x}) \equiv \nabla^2 f(\mathbf{x}) \equiv \begin{bmatrix} \dfrac{\partial^2 f(\mathbf{x})}{\partial x_1^2} & \dfrac{\partial^2 f(\mathbf{x})}{\partial x_1 \partial x_2} \\ \\ \dfrac{\partial^2 f(\mathbf{x})}{\partial x_1 \partial x_2} & \dfrac{\partial^2 f(\mathbf{x})}{\partial x_2^2} \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ \\ 0 & 4 \end{bmatrix}$$

Gradient Vector: Hessian Matrix: Example

Find the Gradient Vector and Hessian Matrix for: $f(x, y) = x^3 - y^3 + 9xy$

Solution:

Gradient Vector:

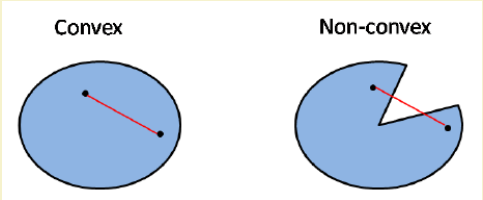
$$\nabla f(x, y) = \begin{pmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{pmatrix} = \begin{pmatrix} 3x^2 + 9y \\ -3y^2 + 9x \end{pmatrix}$$

Hessian Matrix:

$$\nabla^2 f(x, y) = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{pmatrix} = \begin{pmatrix} 6x & 9 \\ 9 & -6y \end{pmatrix}$$

Convex Set

A set of points (or a region) is defined as a convex set in *n*-dimensional space if, for all pairs of points x_1, and x_2, in the set, the straight-line segment joining them is also entirely in the set.



Mathematical statement of a convex set:

We say a set C is convex if for any two points x_1 and x_2 in the set C, the line segment given by $\alpha x_1 + (1-\alpha)x_2$, $0 \le \alpha \le 1$ also lies in the set C.

- The convex region may be closed (bounded) by a set of functions, or may be open (unbounded).
- The intersection of any number of convex set is a convex set.

Convex and Concave Functions

A function is called **concave** over a given region R if

$$f\left(\theta \mathbf{x}_a + (1-\theta)\mathbf{x}_b\right) \geq \theta f(\mathbf{x}_a) + (1-\theta) f(\mathbf{x}_b)$$

$$\text{where: } \mathbf{x}_a, \mathbf{x}_b \in R, \text{ and } 0 \leq \theta \leq 1.$$

The function is **strictly concave** if $\geq$ is replaced by $>$.

A function is called **convex** over a given region R if

$$f\left(\theta \mathbf{x}_a + (1-\theta)\mathbf{x}_b\right) \leq \theta f(\mathbf{x}_a) + (1-\theta) f(\mathbf{x}_b)$$

$$\text{where: } \mathbf{x}_a, \mathbf{x}_b \in R, \text{ and } 0 \leq \theta \leq 1.$$

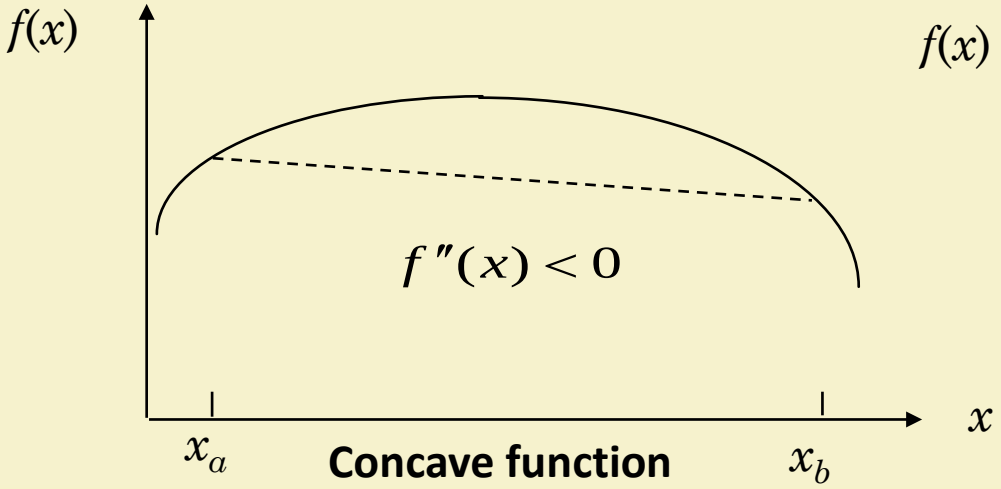
A function is called **strictly convex** if $\leq$ is replaced by $<$.

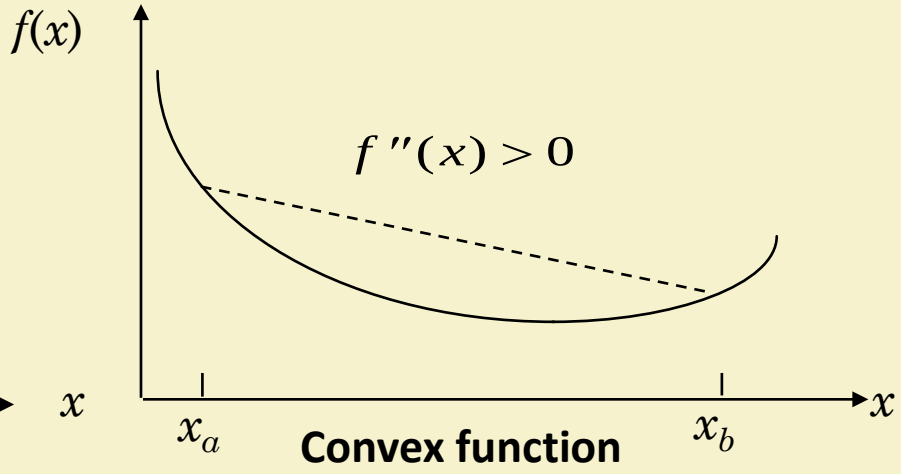
Convex and Concave Functions: Single Variable

If $f''(x) = \frac{\partial^2 f}{\partial x^2} \leq 0$ then $f(x)$ is concave.

If $f''(x) = \frac{\partial^2 f}{\partial x^2} \geq 0$ then $f(x)$ is convex.

Note: A straight line is both convex and concave.





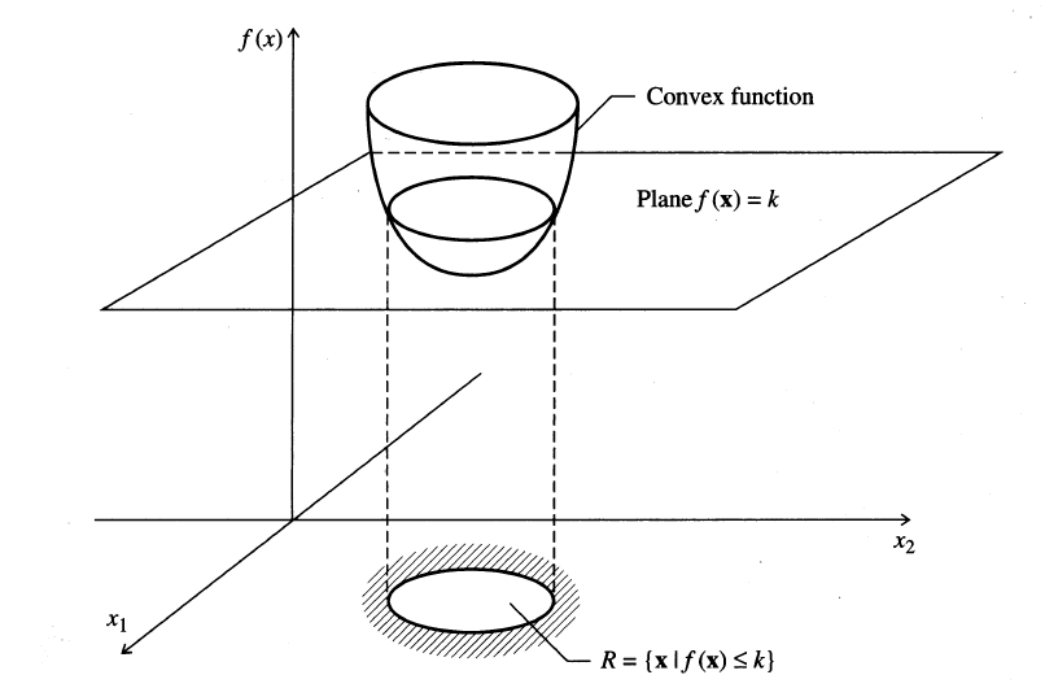
Convex and Concave Functions: Multivariate Functions

For a multivariate function $f(x)$, the condition for convexity are as follows:

Function $f(x)$	Hessian Matrix, H	$x^T Hx$ $(x \neq 0)$	λ of H
Strictly convex	Positive definite	> 0	> 0
Convex	Positive semi-definite	≥ 0	≥ 0
Strictly concave	Negative definite	< 0	< 0
Concave	Negative semi-definite	≤ 0	≤ 0

λ= Eigenvalues

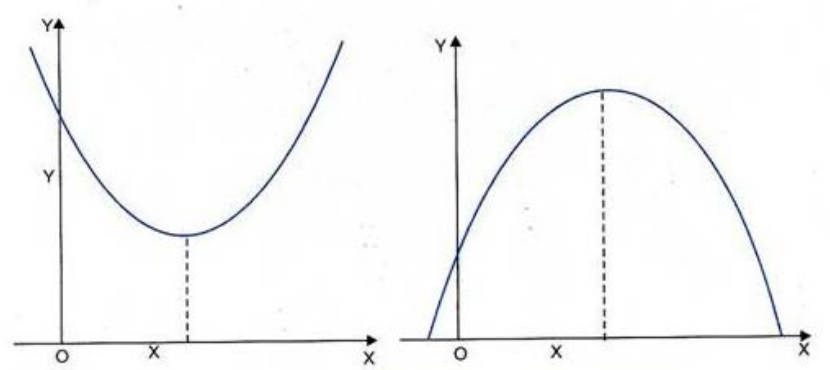
Convex Set and Convex Functions



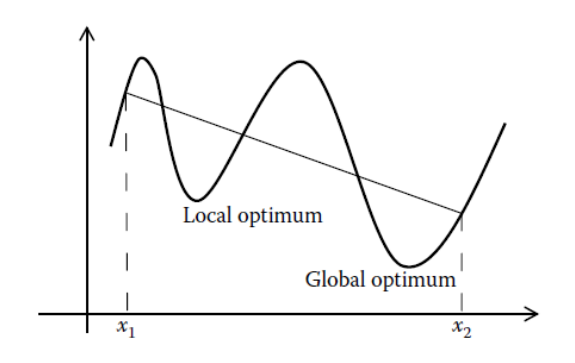
A convex set is formed by a plane cutting a convex function

Convex and Concave Functions

Local optima = Global optima for convex functions



convex function Concave function



Non-convex function

Convex and Concave Functions: Example

Consider: $f(x_1, x_2) = x_1 x_2$

Hessian matrix, H: $\nabla^2 f(x_1, x_2) = H(x_1, x_2) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

Eigen values are (1, -1). H is indefinite. The function is neither convex nor concave.

Convex and Concave Functions: Example

Consider: $f(x_1, x_2) = x_1^3 + x_2^4$

Hessian matrix, H: $\nabla^2 f(x_1, x_2) = H(x_1, x_2) = \begin{bmatrix} 6x_1 & 0 \\ 0 & 12x_2^2 \end{bmatrix}$

Eigenvalues of H: $\lambda_1 = 6x_1, \quad \lambda_2 = 12x_2^2$

Only one eigenvalue is always positive $\left(x_2^2\right)$. The other eigenvalue may be either positive or negative ($6x_1$).

The Hessian matrix is positive semi-definite for $x_1 \geq 0$ but indefinite for $x_1 < 0$

Thus the function is neither convex nor concave over $\Re^n$

Example

$$f(x_1, x_2, x_3) = 3x_1^2 + 2x_2^2 + x_3^2 - 2x_1x_2 - 2x_1x_3 + 2x_2x_3 - 6x_1 - 4x_2 - 2x_3$$

$$\nabla f(x_1, x_2, x_3) = \begin{pmatrix} 6x_1 - 2x_2 - 2x_3 - 6 \\ 4x_2 - 2x_1 + 2x_3 - 4 \\ 2x_3 - 2x_1 + 2x_2 - 2 \end{pmatrix}$$

- H is a symmetric matrix.
- All diagonal elements are positive.

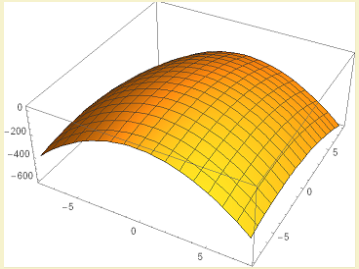
$$H(x_1, x_2, x_3) = \begin{bmatrix} 6 & -2 & -2 \\ -2 & 4 & 2 \\ -2 & 2 & 2 \end{bmatrix}$$

- The leading principal determinants are

$$|6| > 0 \qquad \begin{vmatrix} 6 & -2 \\ -2 & 4 \end{vmatrix} = 20 > 0 \qquad |H| = 16 > 0$$

H is a positive-definite matrix , which implies *f* is convex function

Optimization in Chemical Engineering



Thank You

IIT KHARAGPUR

NPTEL ONLINE CERTIFICATION COURSES

OPTIMIZATION IN CHEMICAL ENGINEERING

Basic Concepts of Optimization – I

(WEEK – 3: LECTURE – 13)

Dr. Debasis Sarkar

Department of Chemical Engineering

Indian Institute of Technology Kharagpur

Basic Concepts of Optimization – I

Week 3:

- Continuity of functions
- Unimodal and multimodal functions
- Convex and concave functions
- Optimality criteria for unconstrained single variable functions
- Optimality criteria for unconstrained multi-variable functions

Today's Topic:

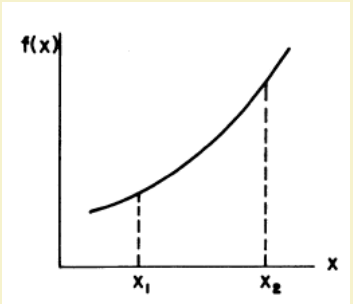
- Optimality criteria for unconstrained single variable functions

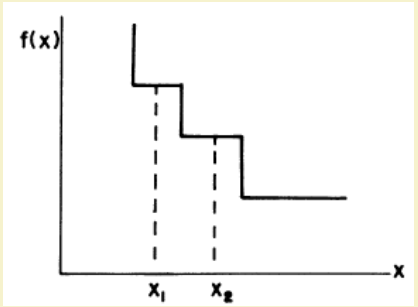
Monotonic Functions

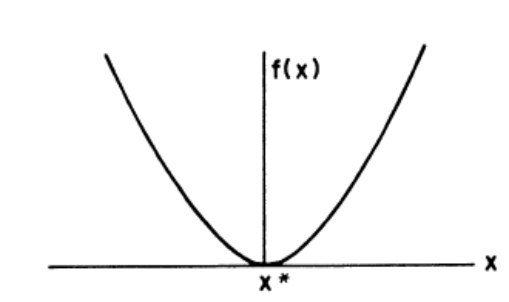
A function $f(x)$ is monotonic (either increasing or decreasing) if for any two points x_1 and x_2, with $x_1 \leq x_2$, it follows that:

$$f(x_1) \leq f(x_2) \quad \text{Monotonically increasing}$$

$$f(x_1) \geq f(x_2) \quad \text{Monotonically decreasing}$$







An unimodal function

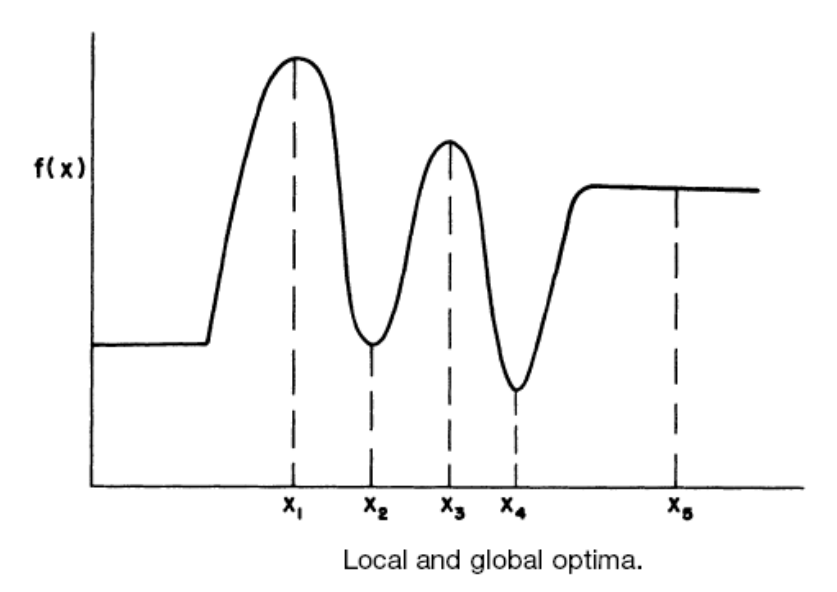
The function $f(x)$ is *unimodal* on the interval $a \leq x \leq b$ if and only if it is monotonic on either side of the single optimal point x^* in the interval.

Global and Local Minimum (Maximum)

- A function is said to have a *relative* or *local* minimum at $x = x^*$ if $f(x^*) \leq f(x+h)$ for all sufficiently small positive and negative values of *h,* i.e. in the near vicinity of the point *x*.
- Similarly, a point x^* is called a *relative* or *local* maximum if $f(x^*) \geq f(x+h)$ for all values of *h* sufficiently close to zero.

- A function is said to have a *global* or *absolute* minimum at $x = x^*$ if $f(x^*) \leq f(x+h)$ for all *x* in the domain over which *f*(*x*) is defined.
- Similarly, a function is said to have a *global* or *absolute* maximum at $x = x^*$ if $f(x^*) \geq f(x+h)$ for all *x* in the domain over which *f* (*x*) is defined.

Minimum/Maximum



Local and global optima.

Weierstrass Theorem – Existence of a Global Minimum:

If f(x) is continuous on a nonempty feasible set S that is closed and bounded, then f(x) has a global minimum.

Concepts of Necessary and Sufficient Conditions

Necessary Conditions:

- The conditions that must be satisfied at the optimum point are called necessary.
- If a point does not satisfy the necessary conditions, it cannot be optimum.
- Satisfaction of necessary conditions does not guarantee optimality of the point.

Sufficient Conditions:

- If a candidate optimum point satisfies the sufficient condition, then it is indeed an optimum point.
- If the sufficient condition is not satisfied (or can not be used), we may not be able to draw any conclusion about the optimality of the point.

Unconstrained Single Variable Problems: First-Order Necessary Conditions

- If a function $f(x)$ is defined in the interval $a \leq x \leq b$ and has a relative minimum at $x = x^*$, where $a < x^* < b$, and if the derivative $df(x)/dx = f'(x)$ exists at $x = x^*$, then $f'(x^*)=0$

- The theorem does not say that the function necessarily will have a minimum or maximum at every point where the derivative is zero.
- In general, a point x^* at which $f'(x^*)=0$ is called a *stationary/critical point.*

- A stationary point that is neither minimum nor maximum is known as *inflection point* or *saddle point*.

Unconstrained Single Variable Problems: Second-Order Necessary/Sufficient Conditions

If a function $f(x)$ is defined in the interval $a \leq x \leq b$ and has a stationary point at $x = x^*$, where $a < x^* < b$, then

Second order necessary condition for a local minimum $\frac{d^2 f}{dx^2} \geq 0$

Second order necessary condition for a local maximum $\frac{d^2 f}{dx^2} \leq 0$

Second order sufficient condition for a local minimum $\frac{d^2 f}{dx^2} > 0$

Second order sufficient condition for a local maximum $\frac{d^2 f}{dx^2} < 0$

Unconstrained Single Variable Problems: Higher Order Sufficient Conditions

Suppose at point x^* the first derivative is zero and the first nonzero higher order derivative is denoted by n.

Let us consider: $f'(x^*)=f''(x^*)=\ldots=f^{(n-1)}(x^*)=0$, but $f^{(n)}(x^*) \neq 0$.

Then x^* is:

- a **local minimum** of $f(x)$ if $f^{(n)}(x^*) > 0$ and n is **even** [+ve derivative]
- a **local maximum** of $f(x)$ if $f^{(n)}(x^*) < 0$ and n is **even** [-ve derivative]
- neither a minimum nor a maximum (**point of inflection**) if n is **odd**

Example - 1

Find the optimum value of the function $f(x) = x^2 + 5x - 2$

Does the function attain a maximum or a minimum?

Solution

Take derivative and set to zero:

$$\frac{df}{dx} = f'(x) = 2x + 5 = 0 \qquad \Rightarrow x^* = -\frac{5}{2}$$

Is x* = -5/2 minimum or maximum? Take second-order derivative.

$$\frac{d^2 f}{dx^2} = 2 > 0$$

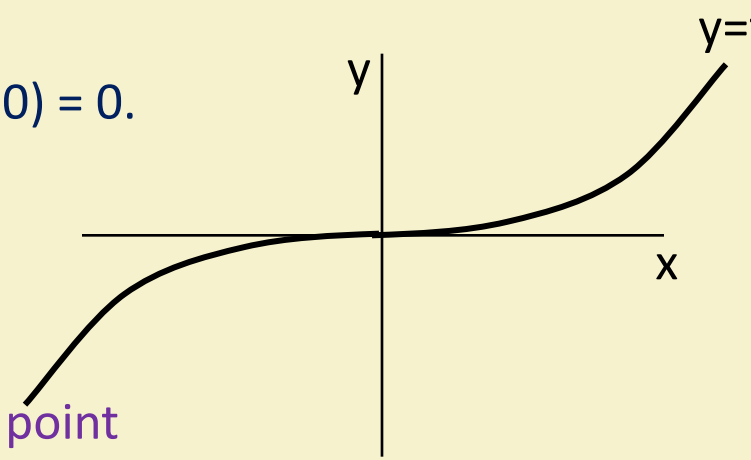
Since the second derivative is positive, the point x^* = -5/2 is a minimum point and the function attains a minimum value of -33/4 at this point.

Example - 2

Let us consider the function: $y = f(x) = x^3$. Is the origin (x = 0) a maximum or minimum point?

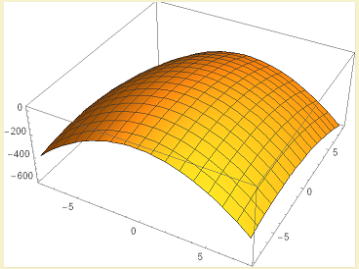
Here, $f'(x) = 3x^2$ and $f''(x) = 6x$.
Thus, at $x^* = 0$,
$f'(x^*=0) = 0$ and $f''(x^*=0) = 0$.

But $f'''(x^*=0) = 6$
Third order derivative is positive (n = 3).
Thus $x^*=0$ represents a point of inflection.



$x^* = 0$ is point of inflection (saddle point) where the point is neither a maximum nor a minimum

Optimization in Chemical Engineering



Thank You

IIT KHARAGPUR

NPTEL ONLINE CERTIFICATION COURSES

OPTIMIZATION IN CHEMICAL ENGINEERING

Basic Concepts of Optimization – I

(WEEK – 3: LECTURE – 14)

Dr. Debasis Sarkar

Department of Chemical Engineering

Indian Institute of Technology Kharagpur

Basic Concepts of Optimization – I

Week 3:

- ✓ Continuity of functions
- ✓ Unimodal and multimodal functions
- ✓ Convex and concave functions
- Optimality criteria for unconstrained single variable functions
- Optimality criteria for unconstrained multi-variable functions

Today's Topic:

- Optimality criteria for unconstrained single variable functions (Cont'd)

Optimality Criteria for Unconstrained Single Variable Functions: Review

If a function $f(x)$ is defined in the interval $a \leq x \leq b$ and has a stationary point at $x = x^*$, where $a < x^* < b$, then at $x = x^*$

First-order necessary condition: $\left.\frac{df}{dx}\right|_{x^*} = 0$

Second order necessary condition for a local minimum $\left.\frac{d^2 f}{dx^2}\right|_{x^*} \geq 0$

Second order necessary condition for a local maximum $\left.\frac{d^2 f}{dx^2}\right|_{x^*} \leq 0$

Second order sufficient condition for a local minimum $\left.\frac{d^2 f}{dx^2}\right|_{x^*} > 0$

Second order sufficient condition for a local maximum $\left.\frac{d^2 f}{dx^2}\right|_{x^*} < 0$

Optimality Criteria for Unconstrained Single Variable Functions: Review

Suppose at point x^* the first derivative is zero and the first nonzero higher order derivative is denoted by n.

Let us consider: $f'(x^*) = f''(x^*) = \ldots = f^{(n-1)}(x^*) = 0$, but $f^{(n)}(x^*) \neq 0$.

Then x^* is:

- a local minimum of $f(x)$ if $f^{(n)}(x^*) > 0$ and n is even [+ve derivative]
- a local maximum of $f(x)$ if $f^{(n)}(x^*) < 0$ and n is even [-ve derivative]
- neither a minimum nor a maximum (point of inflection) if n is odd

Optimality Criteria for Unconstrained Single Variable Functions - Necessary Conditions: Proof

Let $f(x)$ be a function of single variable x defined in the interval $a \le x \le b$ and $f(x)$ is differentiable to the n-th order over the interval. If x^* is a point in the interval, then we can write the following using Taylor's theorem:

$$f(x^*+\varepsilon) - f(x^*) = \varepsilon \left.\frac{df}{dx}\right|_{x=x^*} + \left.\frac{\varepsilon^2}{2!}\frac{d^2 f}{dx^2}\right|_{x=x^*} + \ldots\ldots + \left.\frac{\varepsilon^n}{n!}\frac{d^n f}{dx^n}\right|_{x=x^*} + O_{n+1}(\varepsilon)$$

If x^* is a local minimum of f on [a, b], then in the ε neighborhood, $f(x) \ge f(x^*)$

The above inequality implies: $f(x^*+\varepsilon) - f(x^*) \ge 0$

$$\Rightarrow \varepsilon \left.\frac{df}{dx}\right|_{x=x^*} + \left.\frac{\varepsilon^2}{2!}\frac{d^2 f}{dx^2}\right|_{x=x^*} + \ldots\ldots + \left.\frac{\varepsilon^n}{n!}\frac{d^n f}{dx^n}\right|_{x=x^*} + O_{n+1}(\varepsilon) \ge 0$$

Optimality Criteria for Unconstrained Single Variable Functions - Necessary Conditions: Proof

$$\varepsilon \left.\frac{df}{dx}\right|_{x=x^*} + \left.\frac{\varepsilon^2}{2!}\frac{d^2 f}{dx^2}\right|_{x=x^*} + \ldots\ldots + \left.\frac{\varepsilon^n}{n!}\frac{d^n f}{dx^n}\right|_{x=x^*} + O_{n+1}(\varepsilon) \geq 0$$

For ε sufficiently small, the first term will dominate the others. Since ε can be chosen both positive and negative, the inequality will hold only when $\left.\frac{df}{dx}\right|_{x=x^*} = 0$

Continuing the same argument, we can write: $\left.\frac{d^2 f}{dx^2}\right|_{x=x^*} \geq 0$ for x^* to be local minimum.

If x^* be a local maximum, the same argument will hold – but with the inequality sign reversed.

$$f(x) \leq f(x^*) \text{ and } f(x^*+\varepsilon) - f(x^*) \leq 0 \quad \Rightarrow \quad \left.\frac{df}{dx}\right|_{x=x^*} = 0 \qquad \left.\frac{d^2 f}{dx^2}\right|_{x=x^*} \leq 0$$

Optimality Criteria for Unconstrained Single Variable Functions – Sufficient Conditions: Proof

The first non-zero higher order derivative is n, we can write (x^* is a stationary point)

$$f(x^*+\varepsilon)-f(x^*)=\frac{\varepsilon^n}{n!}\frac{d^n f}{dx^n}\bigg|_{x=x^*}+O_{n+1}(\varepsilon)$$

If n is odd, then the RHS can be made positive (negative) by taking ε positive (negative).

This implies that : $f(x^*+\varepsilon)-f(x^*)\le 0 \quad \text{or} \quad f(x^*+\varepsilon)-f(x^*)\ge 0$

depending on the sign of ε . Hence, the function does not attain a minimum or maximum at x^* and x^* is a point of inflection.

Optimality Criteria for Unconstrained Single Variable Functions – Sufficient Conditions: Proof

$$f(x^*+\varepsilon) - f(x^*) = \frac{\varepsilon^n}{n!}\frac{d^n f}{dx^n}\bigg|_{x=x^*} + O_{n+1}(\varepsilon)$$

If n is even, then ε^n is always positive. For sufficiently small ε the sign of above equation will be dominated by the first term: $\frac{\varepsilon^n}{n!}\frac{d^n f}{dx^n}\Big|_{x=x^*}$

If $\frac{\varepsilon^n}{n!}\frac{d^n f}{dx^n}\Big|_{x=x^*}$ is positive, i.e., $\frac{d^n f}{dx^n}\Big|_{x=x^*} > 0$ then $f(x^*+\varepsilon) - f(x^*) > 0$ and x^* is a local minimum.

Similar argument applies for local maximum: $f(x^*+\varepsilon) - f(x^*) < 0$ when $\frac{d^n f}{dx^n}\Big|_{x=x^*} < 0$

Example - 1

Determine the maximum and minimum values of the function:

$$f(x) = -x^3 + 3x^2 + 9x + 10$$

Solution:

Use first-order condition to find stationary points.

$$\frac{df}{dx} = -3x^2 + 6x + 9 = 0$$

$$\Rightarrow (x-3)(x+1) = 0 \Rightarrow x^* = 3, -1$$

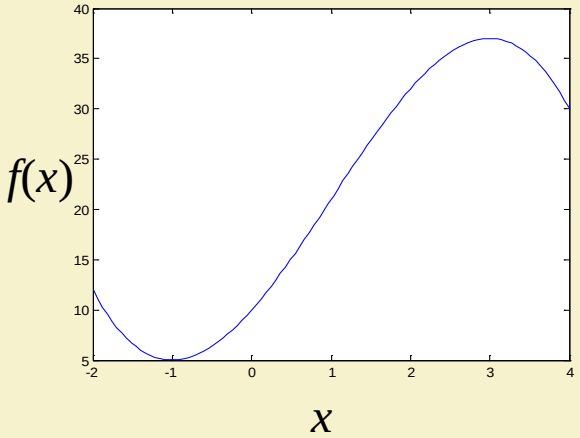
We have 2 stationary points. We have to evaluate both for minimum/maximum.

The second derivative is: $\frac{d^2 f}{dx^2} = -6x + 6$

Example – 1 (Cont'd)

$$\frac{d^2 f}{dx^2} = -6x + 6$$

Stationary point, x*	2nd Derivative	Conclusion
-1	12 > 0	x = -1 is minimum
3	-12 < 0	x = 3 is maximum



$f(3)=37, \quad f(-1)=5$

$$f(x) = -x^3 + 3x^2 + 9x + 10$$

Example - 2

Determine the maximum and minimum values of the function:

$$f(x) = 12x^5 - 45x^4 + 40x^3 + 5$$

Solution:

Here $f'(x) = 60(x^4 - 3x^3 + 2x^2) = 60x^2(x - 1)(x - 2) = 0$

This implies that $f'(x) = 0$ at $x=0$, $x=1$, and $x=2$. We have 3 stationary points.

We need to evaluate all the 3 stationary points for minimum/maximum.

The second derivative is: $f''(x) = 60(4x^3 - 9x^2 + 4x)$

Example – 2 (Cont'd)

The second derivative is: $f''(x) = 60(4x^3 - 9x^2 + 4x)$

At *x=1, f''(x) = -60* and hence *x=1* is a relative maximum. $f_{max} = f(x=1) = 12$

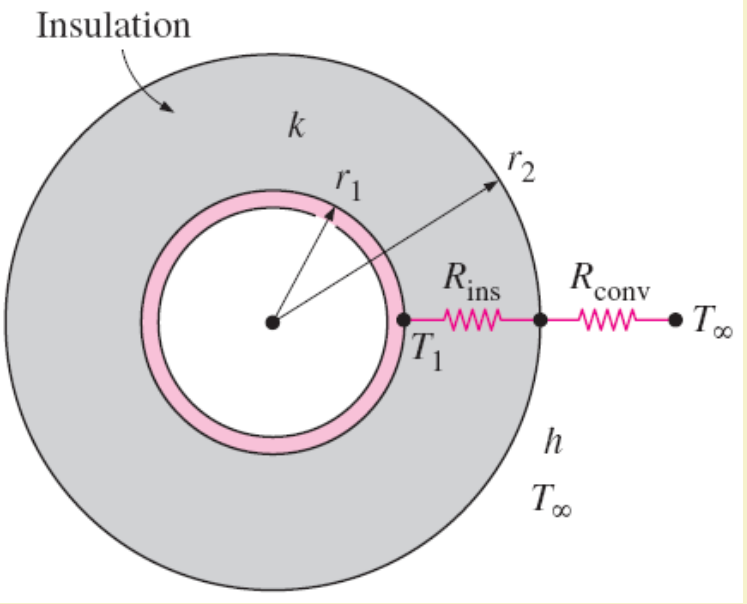
At *x=2, f''(x) = 240* and hence *x=2* is a relative minimum. $f_{min} = f(x=2) = -11$

At *x=0, f''(x) = 0* and hence we must investigate the next higher derivatives.

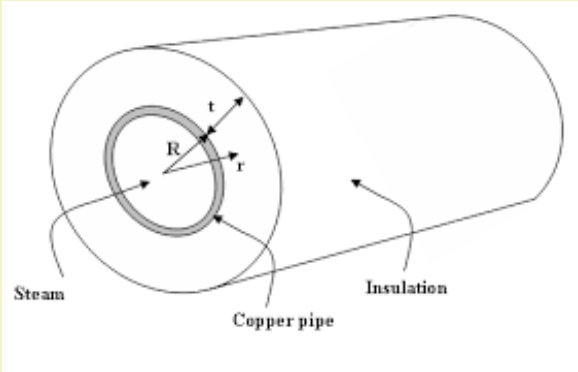
$$f'''(x) = 60(12x^2 - 18x + 4) = 240 > 0 \text{ at } x = 0$$

Since f''' ≠ 0 at *x=0, the point x=0* is neither a maximum nor a minimum, and it is an inflection point (Third derivative, n=3, is positive).

Example – 3: Critical Insulation Thickness



$$\dot{Q} = \frac{T_1 - T_\infty}{R_{ins} + R_{conv}} = \frac{T_1 - T_\infty}{\frac{\ln(r_2/r_1)}{2\pi Lk} + \frac{1}{h(2\pi r_2 L)}}$$



At critical insulation thickness, total thermal resistance will be minimum.

$$R_{tot} = \frac{\ln\frac{r}{r_1}}{2\pi kL} + \frac{1}{2\pi rLh}$$

Example – 3: Critical Insulation Thickness (Cont'd)

$$R_{tot} = \frac{\ln \frac{r}{r_1}}{2\pi kL} + \frac{1}{2\pi rLh}$$

Single variable function of r.
Use first-order condition.

$$\Rightarrow \frac{dR_{tot}}{dr} = 0$$

$$\Rightarrow \frac{1}{2\pi kLr} - \frac{1}{2\pi r^2 Lh} = 0$$

$$\Rightarrow r_c = \frac{k}{h}$$

Now check that $r_c = k/h$ is indeed minimum point.
Use second-order condition.

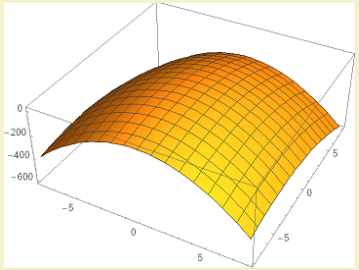
$$\frac{d^2R_{tot}}{dr^2} = -\frac{1}{2\pi kLr^2} + \frac{1}{\pi r^3 Lh}$$

At $r = k/h$:

$$\frac{d^2R_{tot}}{dr^2} = \frac{1}{\pi L\left(\frac{k}{h}\right)^2}\left(\frac{1}{k} - \frac{1}{2k}\right) = \frac{1}{2\pi L\frac{k^3}{h^2}} > 0$$

Thus at the insulation thickness $r = k/h$, the total thermal resistance is indeed minimum.

Optimization in Chemical Engineering



Thank You

IIT KHARAGPUR

NPTEL ONLINE CERTIFICATION COURSES

OPTIMIZATION IN CHEMICAL ENGINEERING

Basic Concepts of Optimization – I

(WEEK – 3: LECTURE – 15)

Dr. Debasis Sarkar

Department of Chemical Engineering

Indian Institute of Technology Kharagpur

Basic Concepts of Optimization – I

Week 3:

- ✓ Continuity of functions
- ✓ Unimodal and multimodal functions
- ✓ Convex and concave functions
- ✓ Optimality criteria for unconstrained single variable functions
- Optimality criteria for unconstrained multi-variable functions

Today's Topic:

- Optimality criteria for unconstrained multi-variable functions

Gradient Vector: Hessian Matrix: Review

$$\nabla f(\mathbf{x}) \equiv \begin{bmatrix} \frac{\partial f(\mathbf{x})}{\partial x_1} \\ \vdots \\ \frac{\partial f(\mathbf{x})}{\partial x_n} \end{bmatrix}$$

$$\mathbf{H}(\mathbf{x}) \equiv \nabla^2 f(\mathbf{x}) \equiv \begin{bmatrix} \frac{\partial^2 f(\mathbf{x})}{\partial x_1^2} & \cdots & \frac{\partial^2 f(\mathbf{x})}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f(\mathbf{x})}{\partial x_1 \partial x_n} & \cdots & \frac{\partial^2 f(\mathbf{x})}{\partial x_n^2} \end{bmatrix}$$

Hessian matrix for a three variable function: $f(x_1,x_2,x_3)$

$$\mathbf{H}(\mathbf{x}) \equiv \begin{bmatrix} \frac{\partial^2 f(\mathbf{x})}{\partial x_1^2} & \frac{\partial^2 f(\mathbf{x})}{\partial x_1 \partial x_2} & \frac{\partial^2 f(\mathbf{x})}{\partial x_1 \partial x_3} \\ \frac{\partial^2 f(\mathbf{x})}{\partial x_2 \partial x_1} & \frac{\partial^2 f(\mathbf{x})}{\partial x_2^2} & \frac{\partial^2 f(\mathbf{x})}{\partial x_2 \partial x_3} \\ \frac{\partial^2 f(\mathbf{x})}{\partial x_3 \partial x_1} & \frac{\partial^2 f(\mathbf{x})}{\partial x_3 \partial x_2} & \frac{\partial^2 f(\mathbf{x})}{\partial x_3^2} \end{bmatrix}$$

Convex and Concave Functions: Multivariate Functions: Review

For a multivariate function $f(\mathbf{x})$, the condition for convexity are as follows:

Function f(x)	Hessian Matrix, H	$x^T Hx$ $(x \neq 0)$	λ of H
Strictly convex	Positive definite	> 0	> 0
Convex	Positive semi-definite	≥ 0	≥ 0
Strictly concave	Negative definite	< 0	< 0
Concave	Negative semi-definite	≤ 0	≤ 0

λ= Eigenvalues

Optimality Criteria for Unconstrained Multi-Variable Functions: Necessary Conditions

The necessary conditions for $\mathbf{x}^*$ to be a stationary point of a multivariable function $f(\mathbf{x})$ at $\mathbf{x} = \mathbf{x}^*$:

1. The function $f(\mathbf{x})$ is twice differentiable
2. At stationary point, each partial derivative is equal to zero. Thus each element of the gradient vector of $f(\mathbf{x})$ at $\mathbf{x} = \mathbf{x}^*$ must be equal to zero.
3. The Hessian matrix H is:
   - positive semi-definite when $\mathbf{x}^*$ is a local or relative minimum point
   - negative semi-definite when $\mathbf{x}^*$ is a local relative maximum point.

Optimality Criteria for Unconstrained Multi-Variable Functions: Sufficient Conditions

For a stationary point **x*** to be an extreme point (maximum or minimum), the Hessian matrix H of $f(\mathbf{x})$ evaluated at **x*** must be:

- positive definite when **x*** is a local or relative minimum point
- negative definite when **x*** is a local relative maximum point.

- If the Hessian at **x*** is neither positive semi-definite nor negative semi-definite, the stationary point, **x***, is neither a local minimum nor a local maximum. It is a saddle point.

Optimality Criteria for Unconstrained Multi-Variable Functions

Consider the Taylor series expansion about stationary point x^* of the multivariable function $f(x)$:

$$f(x) = f(x^*) + \nabla^T f(x^*)\Delta x + \frac{1}{2}(\Delta x)^T \nabla^2 f(x^*)\Delta x + O_3 + \ldots..$$

x, x^* are n-vector

$\Delta x = x - x^*$ perturbation

$\nabla^2 f(x^*) = H(x^*)$

If $f(x^*)$ is a local minimum, $f(x) - f(x^*) \geq 0$

Since Δx is arbitrary (can be positive or negative), it is necessary that $\nabla f(x^*) = 0$

Also, $(\Delta x)^T \nabla^2 f(x^*)\Delta x \geq 0$ for $f(x) - f(x^*) \geq 0$ Thus, H should be positive semi-definite for x^* to become local minimum (Necessary condition).

Optimality Criteria for Unconstrained Multi-Variable Functions (Cont'd)

If we have, $(\Delta x)^T \nabla^2 f(x^*)\Delta x > 0$ then $f(x) - f(x^*) > 0$

That is, if H is positive definite, then x^* is local minimum (Sufficient condition)

Similarly, for local maximum: $f(x) - f(x^*) \leq 0$

Consider: $f(x) = f(x^*) + \nabla^T f(x^*)\Delta x + \frac{1}{2}(\Delta x)^T \nabla^2 f(x^*)\Delta x + O_3 + ...$

Necessary Conditions for local maximum: $\nabla f(x^*) = 0$ and $(\Delta x)^T \nabla^2 f(x^*)\Delta x \leq 0$

H is negative semi-definite for local maximum.

Sufficient condition for local maximum: H is negative definite

Example – 1: Necessary Conditions Are Not Enough

Consider the function $f(x,y)=x^3+y^3$. Does the function attain minimum at (0,0)?

Solution:

$$\nabla f(x,y) = \left[3x^2, 3y^2\right]^T \qquad \nabla^2 f(x,y) = \begin{pmatrix} 6x & 0 \\ 0 & 6y \end{pmatrix}$$

At the point (0,0):

$$f(x,y) = 0 \qquad \nabla f(0,0) = [0,0]^T \qquad \nabla^2 f(0,0) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Hessian is positive semi-definite

Note that (x = 0, y = 0) is certainly not minimum. For example, f(-0.1, -0.1) < f(0,0).

Thus necessary conditions are not enough for optimality.

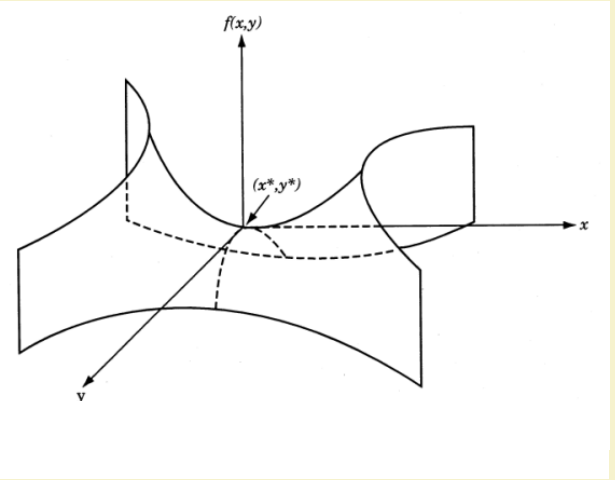
Example – 2: Saddle Point

Analyze the function $f(x,y)=x^2-y^2$ at $(x=0,y=0)$.

Apply optimality conditions: $\frac{\partial f}{\partial x} = 2x$ and $\frac{\partial f}{\partial y} = -2y$

These first derivatives are zero at $x^* = 0$ and $y^* = 0$.

The Hessian matrix of f at (x^*,y^*) is given by: $H = \begin{vmatrix} 2 & 0 \\ 0 & -2 \end{vmatrix}$



The eigenvalues of H are 2 and -2.
Since this matrix is neither positive definite nor negative definite, the point ($x^*=0$, $y^*=0$) is a saddle point.

Example – 3: A Function with 3 Variables

Consider the function $f(x) = -x_1^2 - x_2^2 - x_3^2 + 2x_1x_2 + 2x_1x_3 + 4x_1 - 5x_3 + 5$

Analyze the stationary points (maxima, minima and point of inflection).

Solution:

$$\nabla f = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \frac{\partial f}{\partial x_3} \end{bmatrix} = \begin{bmatrix} -2x_1 + 2x_2 + 2x_3 + 4 \\ -2x_2 + 2x_1 \\ -2x_3 + 2x_1 - 5 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Solve these 3 equations simultaneously to find the stationary point:

$$x_1 = \frac{1}{2}, x_2 = \frac{1}{2}, x_3 = -2$$

Example – 3 (Cont'd)

$$f(x) = -x_1^2 - x_2^2 - x_3^2 + 2x_1x_2 + 2x_1x_3 + 4x_1 - 5x_3 + 5$$

$$\nabla f = \begin{bmatrix} -2x_1 + 2x_2 + 2x_3 + 4 \\ -2x_2 + 2x_1 \\ -2x_3 + 2x_1 - 5 \end{bmatrix}$$

$$\mathbf{H} \equiv \begin{bmatrix} \frac{\partial^2 f(\mathbf{x})}{\partial x_1^2} & \frac{\partial^2 f(\mathbf{x})}{\partial x_1 \partial x_2} & \frac{\partial^2 f(\mathbf{x})}{\partial x_1 \partial x_3} \\ \frac{\partial^2 f(\mathbf{x})}{\partial x_2 \partial x_1} & \frac{\partial^2 f(\mathbf{x})}{\partial x_2^2} & \frac{\partial^2 f(\mathbf{x})}{\partial x_2 \partial x_3} \\ \frac{\partial^2 f(\mathbf{x})}{\partial x_3 \partial x_1} & \frac{\partial^2 f(\mathbf{x})}{\partial x_3 \partial x_2} & \frac{\partial^2 f(\mathbf{x})}{\partial x_3^2} \end{bmatrix}$$

$$H = \begin{bmatrix} -2 & 2 & 2 \\ 2 & -2 & 0 \\ 2 & 0 & -2 \end{bmatrix}$$

$$\frac{\partial^2 f}{\partial x_1^2} = -2, \quad \frac{\partial^2 f}{\partial x_2^2} = -2, \quad \frac{\partial^2 f}{\partial x_3^2} = -2$$

$$\frac{\partial^2 f}{\partial x_1 \partial x_2} = \frac{\partial^2 f}{\partial x_2 \partial x_1} = 2, \quad \frac{\partial^2 f}{\partial x_1 \partial x_3} = \frac{\partial^2 f}{\partial x_3 \partial x_1} = 2$$

$$\frac{\partial^2 f}{\partial x_2 \partial x_3} = \frac{\partial^2 f}{\partial x_3 \partial x_2} = 0$$

Eigenvalues of H are: -2, -2 $\sqrt{8}$.

H has both positive and negative eigenvalues. Thus H is indefinite and the stationary point is a saddle point.

Example – 4: A Function with 5 Stationary Points

Find the stationary points and analyze: $f(x_1, x_2) = (x_1^2 + x_2^2 - 1)^2 + (x_2^2 - 1)^2.$

The gradient:

$$\nabla f(\mathbf{x}) = 4\begin{pmatrix} (x_1^2 + x_2^2 - 1)x_1 \\ (x_1^2 + x_2^2 - 1)x_2 + (x_2^2 - 1)x_2 \end{pmatrix}.$$

There are 5 stationary points:

$(0,0), (1,0), (-1,0), (0,1), (0,-1).$

The Hessian matrix:

$$\nabla^2 f(\mathbf{x}) = 4\begin{pmatrix} 3x_1^2 + x_2^2 - 1 & 2x_1x_2 \\ 2x_1x_2 & x_1^2 + 6x_2^2 - 2 \end{pmatrix}.$$

We have to analyze all five stationary points.

Example – 4 (Cont'd)

At (0, 0): $\nabla^2 f(0,0) = 4\begin{pmatrix} -1 & 0 \\ 0 & -2 \end{pmatrix}$

Eigenvalues are: -4, -8. H is negative definite. (0,0) is local maximum. $f(0,0) = 2$

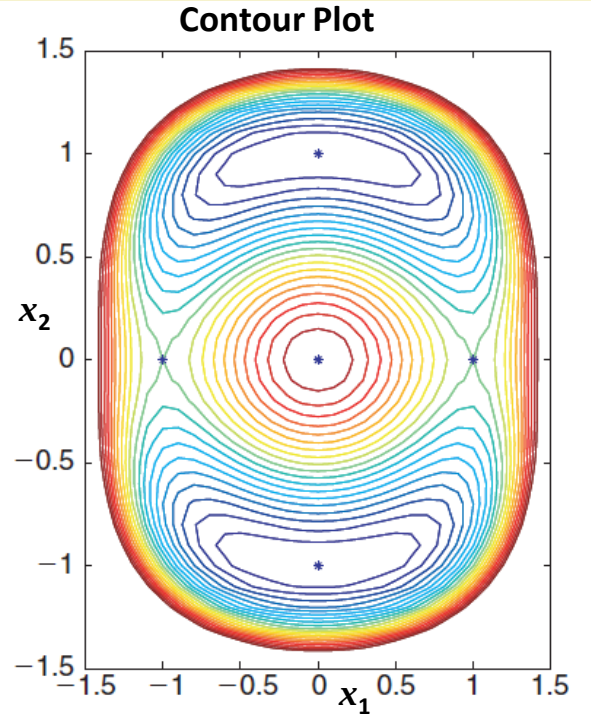
At (1, 0) and (-1,0): $\nabla^2 f(1,0) = \nabla^2 f(-1,0) = 4\begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix}$

Eigenvalues are: -4, 8. H is indefinite. (1,0) and (-1,0) are saddle points.
$f(1,0) = f(-1,0) = 1$

At (0,1) and (0,-1): $\nabla^2 f(0,1) = \nabla^2 f(0,-1) = 4\begin{pmatrix} 0 & 0 \\ 0 & 4 \end{pmatrix}$

Eigenvalues are: 0, 16. H is positive semi-definite. (0,1) and (0,-1) may or may not be local minimum. ONLY necessary conditions are met.

Example – 4 (Cont'd)



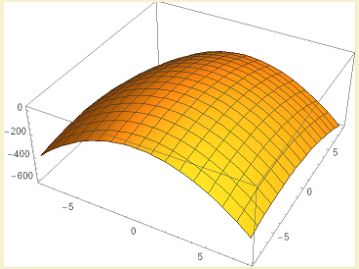
At (0,1) and (0,-1), the H is positive semi-definite. So we can not conclude if these are local minima or not.

However, here it can be seen that (0,1) and (0,-1) are actually global minimum. Function value is zero for both the cases: $f(0,1) = f(0,-1) = 0$. Note that the minimum value for the function can be zero only.

$$f(x_1, x_2) = (x_1^2 + x_2^2 - 1)^2 + (x_2^2 - 1)^2.$$

Summary: Local/Global minimum: (0,1), (0,-1)
Maximum: (0,0)
Saddle point: (1,0), (-1,0)

Optimization in Chemical Engineering



Thank You