Chapter 7: Mathematical optimization: **Unconstrained problem**

7.1) General Statement of the problem

Two classes of the optimization problem

(i) *Unconstraint optimization*

no Constraint problem

$$\max_{x_1,x_2,\dots,x_N}(\min) f(x_1, x_2, \dots, x_N)$$

Constrain-free problem.

Comments: You are free to choose combinations of "x".

Example:

$$\max_{x,y} \Pi(x,y)$$

(ii) *Constraint optimization*

$$\max_{x_1, x_2, \dots, x_N} f(x_1, x_2, \dots, x_N)$$

Equality-constraint optimization problem.

$$\boxed{\text{such that } g(x_1, x_2, \dots, x_N) = b}$$

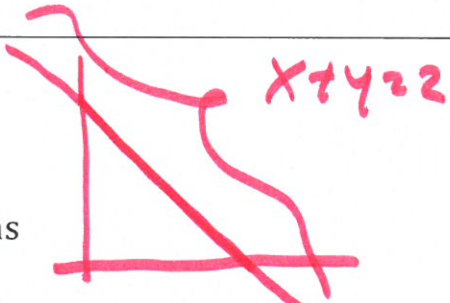
→ Constraint Set

Constraint Equation

→ EE421

⇒ M Constraints

⇒ Equality → inequality Constraint.



Comments:

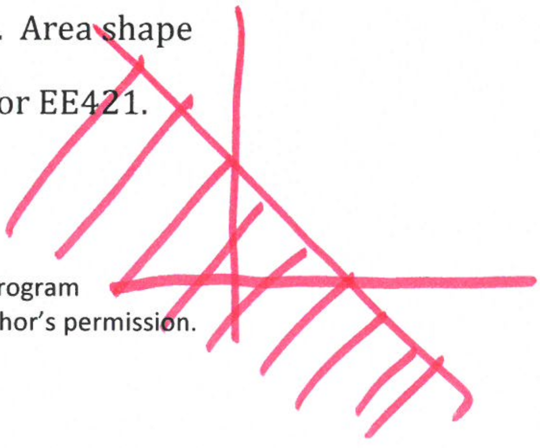
a) *M-constraint* (equations) functions

b) Constraint set could be described/represented by ***inequality***,

line / area

$x + y = 2$ $v.s.$ $x + y \leq 2$ → Line V.S. Area shape

This topics is the central discussion for EE421.



7.2) Unconstrained optimization

- This section develops toolkits to solve for the solution of the unconstraint optimization problem

7.2.1) Types of optimizer revisited!

Definition: Local optimizer

(x^*, y^*) is a ***local*** optimizer of $f(x,y)$ if and only if, for all (x', y') that are closed to (x^*, y^*),

$$f(x^*, y^*) > f(x', y') \quad \text{(local max)}$$

$$f(x^*, y^*) < f(x', y') \quad \text{(local min)}$$

How do mathematicians define the closed-to set?

- x^* ➔ x' in the neighborhood of x^*
  - $(x^* - \varepsilon, x^* + \varepsilon)$
  - Distant between x' and x^* is less than ε.

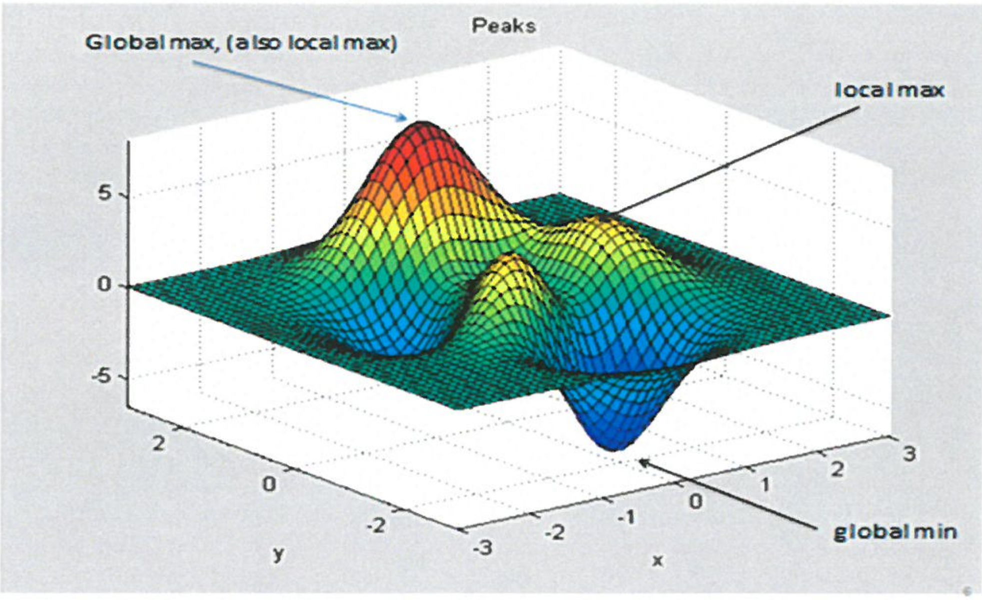
- (x^*, y^*) ➔ (x', y') in the neighborhood of (x^*, y^*)
  - Distant between (x', y') and (x^*, y^*) is less than ε.
  - $\forall (x', y')$: $\sqrt{(x' - x^*)^2 + (y' - y^*)^2} < \varepsilon$

Definition: Global optimizer

(x^*, y^*) is a ***global*** optimizer of $f(x, y)$ if and only if, for all (x', y') in the **entire domain set** that defines "f",

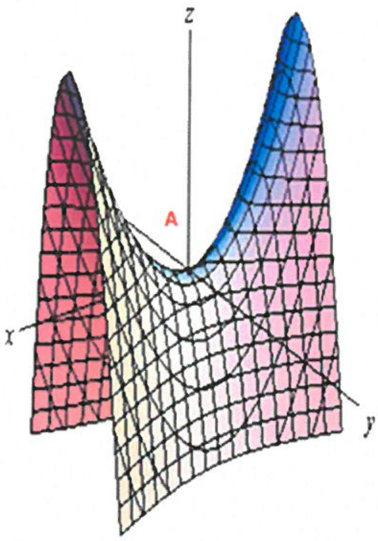
$$f(x^*, y^*) > f(x', y') \quad \text{(global max)}$$

$$f(x^*, y^*) < f(x', y') \quad \text{(global min)}$$

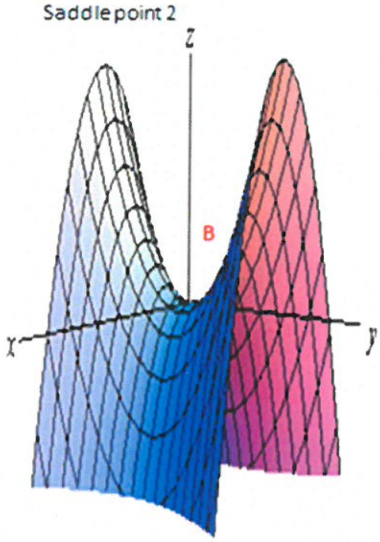


- **Comments**:
  - If the point is local max, it must be local max on both axes (directions).
  - If the point is local max on one axis (direction), but local min on the other axis (direction), the point is neither considered as local max nor local min. Instead, the point is called "***saddle point***".

Saddle point 1



Saddle point 2



Definition: Saddle point 1

(x^*, y^*) is a *saddle point* of $f(x, y)$ if and only if, for all (x', y') that are closed to (x^*, y^*),

$f(x^*, y^*) > f(x^*, y')$ (max in y)

$f(x^*, y^*) < f(x', y^*)$ (min in x)

Definition: Saddle point 2

(x^*, y^*) is a *saddle point* of $f(x, y)$ if and only if, for all (x', y') that are closed to (x^*, y^*),

$f(x^*, y^*) < f(x^*, y')$ (min in y)

$f(x^*, y^*) > f(x', y^*)$ (max in x)

7.2.2) Unconstraint (constraint-free) optimization: theorem

- We start out with a 2-varaible case. Later, we will generalize the result into N-variable case.

Theorem: *Properties of Local optimizer*

Suppose that $f(x, y)$ is a twice-differentiable function. That is, both ∇f and Hessian matrix exist. We would observe the following two properties for which the local optimizer (x^*, y^*) must hold.

- $\nabla f\,(x^*, y^*) = 0$. **(First-order necessary condition)**
- If (x^*, y^*) is the local maximizer (minimizer) of $f(x, y)$, then $f(x, y)$ is concave (convex) at (x^*, y^*). **(Second-order necessary condition)**

Theorem: *Necessary and Sufficient condition*

This theorem is commonly known as the two-step cooked procedure for locating local optimizers. Given that f(x,y) is twice differentiable function, (x^*, y^*) must be a stationary point on both x and y axis.

$$(x^*, y^*) \rightarrow \begin{bmatrix} f_x(x^*, y^*) \\ f_y(x^*, y^*) \end{bmatrix} = 0$$

(i) **First-order condition:** Locate (x^*, y^*) such that Gradient of the function is equal to zero.

(ii) **Second-order condition:** Distinguishing the type of local optimizers by checking for the type of curvatures at the point located from (i).
  a. Concave → local maximizer
  b. Convex → local minimizer

N-variable $\Rightarrow \nabla f = 0$; $\nabla f = \underset{N \times 1}{\begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_N \end{bmatrix}}$

N-unknown variables ~~fom~~ from N-equations

Single variable $\Rightarrow \dfrac{d^2y}{dx^2} > 0 \rightarrow$ Convex $\rightarrow$ min
$\quad\quad < 0 \rightarrow$ concave $\rightarrow$ max

Concavity/convexity of a function

In single variable case, we check for the concavity/convexity of a function using the second-order derivative.

- This is fine, because we only have one axis/direction of change to consider.

For the multivariate case, it is **no longer** appropriate to simply look at the second-order partial derivative. Mathematically, we check for the concavity/convexity of a multivariate function using the **second-order total differential**.

- $f(x \quad, y \quad)$ is said to be *convex* at the point (x^*, y^*) if and only if

$$d^2 f(x^*, y^*) > 0$$

- $f(x \quad, y \quad)$ is said to be *concave* at the point (x^*, y^*) if and only if

$$d^2 f(x^*, y^*) < 0$$

- Second-order total differential is either less than or greater than zero.
- No; Not just looking at second-order partial derivative.

Example 7.A Derive the expression for the second-order differential of f(x,y)

$f(x,y)$

first-order total differential of f.

$$d(df) = d\left(\overbrace{f_x(x,y)}^{\frac{\partial f}{\partial x}} \cdot dx + \overbrace{f_y(x,y)}^{\frac{\partial f}{\partial y}} \cdot dy\right)$$

Second order total differential of f

$$\Rightarrow d(df)$$

; $z = x_1 + x_2$

$dz = dx_1 + dx_2$

$$d^2 f = d f_x(x,y) \cdot dx + d f_y(x,y) \cdot dy.$$

$$d f_x(x,y) = f_{xx}(x,y) \cdot dx + f_{xy}(x,y) \cdot dy.$$

$$d f_y(x,y) = f_{yx}(x,y) \cdot dx + f_{yy}(x,y) dy$$

- All the sub-matrices of "A", indexed by A_i, each of which is expanded along the diagonal of the original matrix A.

Three leading Principal minors of $\underline{\underline{A}}$

$$A = \begin{bmatrix} 1 & -4 & 3 \\ -4 & 5 & 0 \\ 3 & 0 & 9 \end{bmatrix} \Rightarrow A_1 = [1]; \quad A_2 = \begin{bmatrix} 1 & -4 \\ -4 & 5 \end{bmatrix}; \quad A_3 = A$$

A is Positive / negative | sign $|A_i|$

Note: There exists "N" leading principal minors for an N-by-N matrix.

$\Rightarrow A_1 = [-1]$ $|A_1| = -1$ < 0

Example 7.B: $A = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}$; $A = \begin{bmatrix} -1 & 2 \\ 2 & -5 \end{bmatrix}$ $A_2 = \begin{bmatrix} -1 & 2 \\ 2 & -5 \end{bmatrix}$

Find all the sub-matrices of A and check for definiteness of A. $|A_2| = 1$ > 0

$\Rightarrow$ Negative definite

$A_1 = [2] = 2$

$A_2 = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} = 2^2 - (-1)(-1) = 3$

$A_3 = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} \begin{matrix} 2 & -1 \\ -1 & 2 \\ 0 & -1 \end{matrix} = 8 + 0 + 0 - 0 - 2 - 2 = 4$

$\hookrightarrow$ A is positive definite how to check for the definiteness of the Hessian Matrix!

(H)

- If Hessian matrix is *positive definite*, then $d^2f(x^*,y^*) > 0$ for all possible combinations of dx and dy.
    - Implies that "f" is convex!

- If Hessian matrix is *negative definite*, then $d^2f(x^*,y^*) < 0$ for all possible combinations of dx and dy.
    - Implies that "f" is concave!

Theorem: general theorem for checking definiteness of a matrix

Suppose that A is a ***symmetric*** matrix.

- A is positive definite if the determinants of all the ***leading principal minors*** are positive.
- A is negative definite if the determinants of all the ***leading principal minors*** have alternate signs, *with* the negative value for the determinant of the initial leading principal minor A_1.

Comments: Since "H" is symmetric, we can apply the theorem to check for the definiteness of "H" using the above theorem. An important question remained is the definition of what so called leading principle minors.

Definition: Leading principal minors

$$d^2f = \left(f_{xx}(x,y)dx + f_{xy}\cdot dy\right)dx$$

$$+\left(f_{yx}(x,y)\cdot dx + f_{yy}(x,y)\cdot dy\right)\cdot dy$$

$$= f_{xx}(dx)^2 + f_{yy}(dy)^2 + f_{xy}dxdy$$

$$+ f_{yx}dxdy$$

$$= f_{xx}(dx)^2 + 2f_{xy}\cdot dxdy + f_{yy}(dy)^2$$

$$d^2f = [dx \quad dy]\begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix}\begin{bmatrix} dx \\ dy \end{bmatrix}$$

$$d^2f = [dx \quad dy]\, H \begin{bmatrix} dx \\ dy \end{bmatrix}$$

H is heassion Matrixx

Theorem:

$H = \begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix} \Rightarrow |H_1| = f_{xx}$

$|H_2| = f_{xx} \cdot f_{yy} - (f_{xy})(f_{yx})$

From the theorems stated above, we can conclude that, *under the 2-variable case*, second-order condition requires that $|H_2| = f_{xx} \cdot f_{yy} - (f_{xy})^2$

- For the minimizer, H is positive definite = "f" is convex → min point
  - $\det(H_1) = f_{xx} > 0$
  - $\det(H_2) = f_{xx} * f_{yy} - f_{xy}^2 > 0$

- For the maximizer, H is negative definite = "f" is concave
  - $\det(H_1) = f_{xx} < 0$ $d^2f < 0$ alternate sign
  - $\det(H_2) = f_{xx} * f_{yy} - f_{xy}^2 > 0$

Conclusion:

Two-step procedures:

(1), Find all possible pairs of (x^*, y^*) that make $f_x = f_y = 0$
(search for all critical points, i.e. point in which slope is equal to 0.)

(2), Check for concavity/convexity at your sol^n (x^*, y^*)

look at the sign of $|H_i|$

sign of the determinant of all the sub-matrices of Hessian Matrix max (min) $f(x,y)$

$f(x,y)$: objective function

$\rightarrow$ x, y

$\left.\begin{matrix} x \\ y \end{matrix}\right\}$ choice variables that you can choose.

$\rightarrow$ 2-step procedure.

1-st step $\Rightarrow$ First-derivative test

$(x^*, y^*) \rightarrow \dfrac{\partial f}{\partial x} = \dfrac{\partial f}{\partial y} = 0$

"Stationartize"

"critical point"

$(x_1, x_2, x_3, \ldots, x_N) \rightarrow \dfrac{\partial f}{\partial x_1} = \dfrac{\partial f}{\partial x_2} = \ldots \dfrac{\partial f}{\partial x_n} = 0$

$\Downarrow$

maximum / minimum point

then solve for 2-unknown (N-unknown) from 2-equation (N-equation)