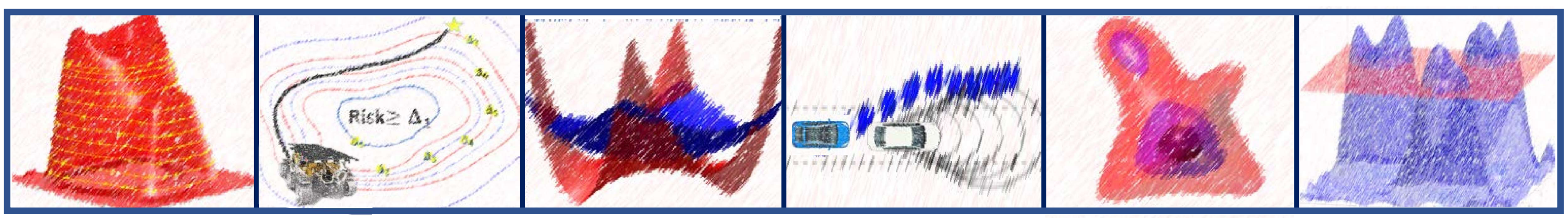
Lecture 2

Nonlinear Optimization

Overview

MIT 16.S498: Risk Aware and Robust Nonlinear Planning
Fall 2019

Ashkan Jasour
jasour@mit.edu
jasour.mit.edu

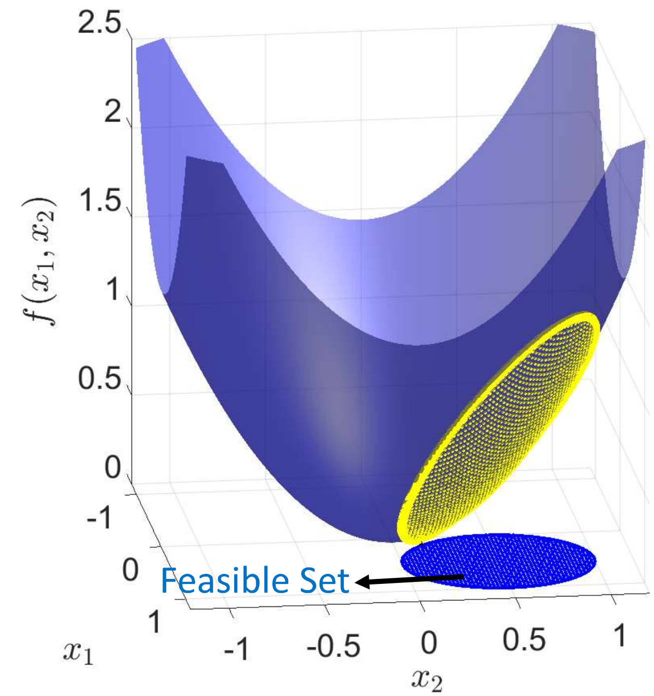


Optimization:

Minimize Objective-function(*design parameters*)
design parameters

Subject to Constraints(*design parameters*)

Design variables : $x = [x_1, ..., x_n]^T \in \mathbb{R}^n$



Real-Valued Scalar Continuous functions

$$f : \mathbb{R}^n \to \mathbb{R}$$

$$g_i : \mathbb{R}^n \to \mathbb{R}, \quad i = 1, ..., n_g$$

$$h_i : \mathbb{R}^n \to \mathbb{R}, \quad i = 1, ..., n_h$$

Nonlinear Optimization:

$$\begin{aligned} \underset{x \in \mathbb{R}^n}{\text{minimize}} \quad & f(x) \\ \text{subject to} \quad & \left.\begin{aligned} & g_i(x) \geq 0, \ i = 1, ..., n_g \\ & h_i(x) = 0, \ i = 1, .., n_h \end{aligned}\right\} \text{Feasible Set} \end{aligned}$$

← Objective function

← Inequality Constraints

← Equality Constraints

Optimization:

$$\begin{aligned} \underset{x \in \mathbb{R}^n}{\text{minimize}} \quad & f(x) \\ \text{subject to} \quad & g_i(x) \geq 0, \ i = 1, ..., n_g \\ & h_i(x) = 0, \ i = 1, .., n_h \end{aligned}$$

Step 1

Optimality Conditions

Given the objective function f and constraints g_i, h_i :

What are the conditions for x^* to be an optimal solution ?

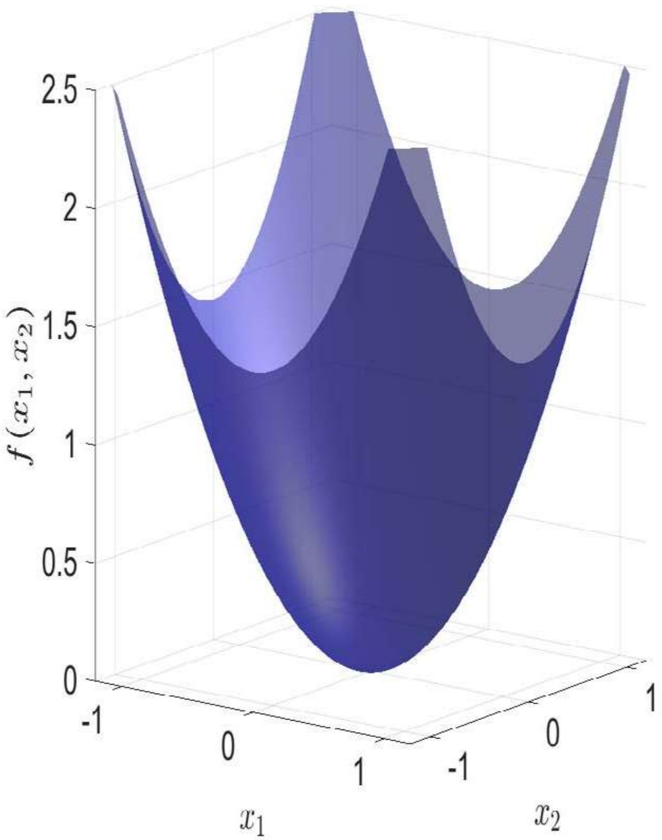
Optimality Conditions: system of nonlinear equations or inequalities

Step 2

To find x^*, we solve the system of nonlinear equations or inequities (root finding problem).

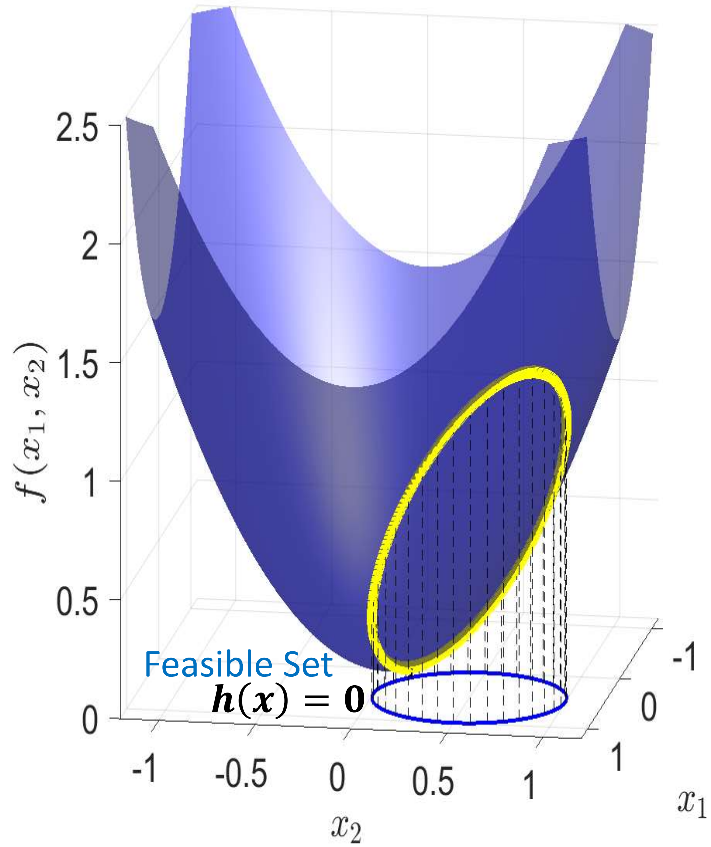
➢ Nonlinear (in general nonconvex) Optimization

Unconstrained Optimization



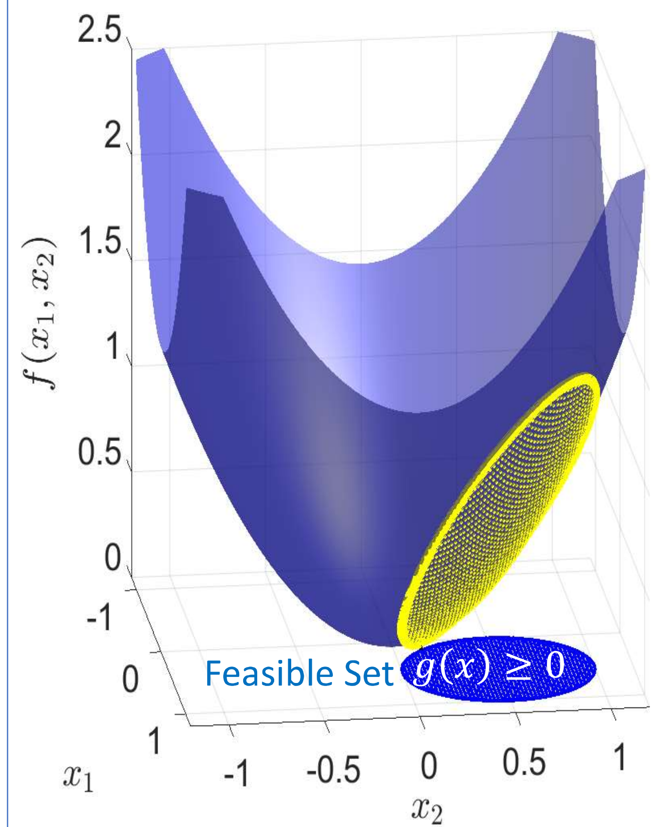
$$\underset{x \in \mathbb{R}^n}{\text{minimize}} \quad f(x)$$

Optimization with equality Constraints



$$\underset{x \in \mathbb{R}^n}{\text{minimize}} \quad f(x)$$
$$\text{subject to} \quad h_i(x) = 0, \;\; i = 1, .., n_h$$

Optimization with Inequality Constraints



$$\underset{x \in \mathbb{R}^n}{\text{minimize}} \quad f(x)$$
$$\text{subject to} \quad g_i(x) \geq 0, \;\; i = 1, ..., n_g$$

Gradient

Real-valued scalar function $f(x) : \mathbb{R}^n \rightarrow \mathbb{R}$

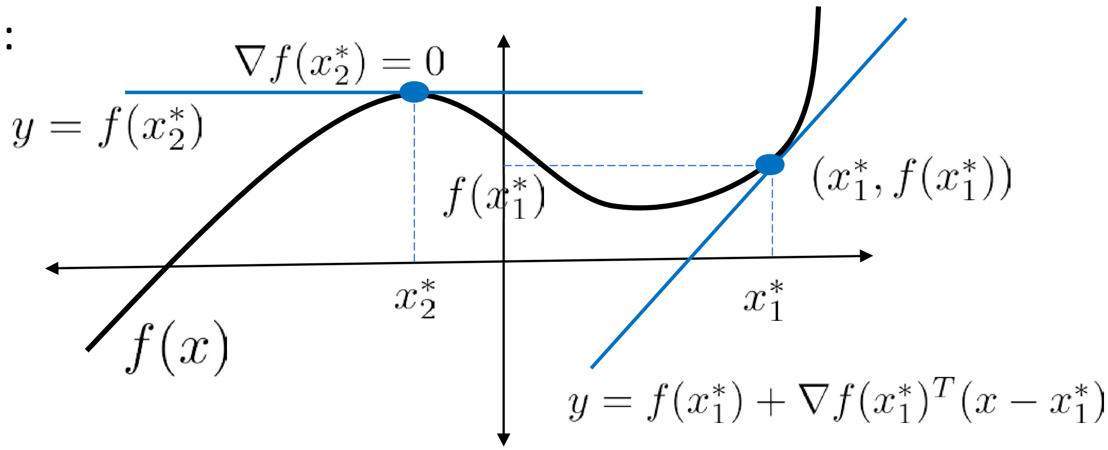
Gradient vector : $\nabla f = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \dots \\ \frac{\partial f}{\partial x_n} \end{bmatrix}$

$$f(x) \Rightarrow \boxed{\nabla} \Rightarrow \nabla f = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \dots \\ \frac{\partial f}{\partial x_n} \end{bmatrix}$$

Gradient and Tangent Line:

Tangent line (Hyperplane) at point $(x^*, f(x^*))$:

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



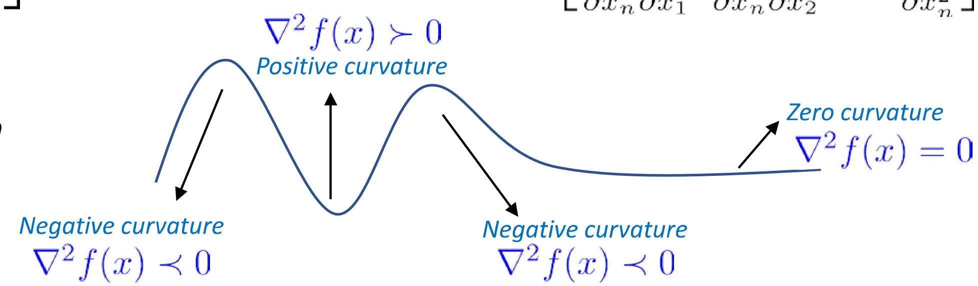
$\nabla f(x^*) = 0 \rightarrow x^* : \text{maximum/minimum}$

Hessian

Real-valued **scalar** function $f(x) : \mathbb{R}^n \rightarrow \mathbb{R}$

$$f(x) \xrightarrow{\textit{Gradient vector}} \nabla f = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \cdots \\ \frac{\partial f}{\partial x_n} \end{bmatrix} \xrightarrow{\textit{Hessian matrix}} \nabla^2 f(x) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \vdots & & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix}$$

Hessian matrix : describes the local curvature of the function



Jacobian

Real-valued **vector** function $f(x) : \mathbb{R}^n \rightarrow \mathbb{R}^m$

$$f = \begin{bmatrix} f_1(x) \\ \cdots \\ f_m(x) \end{bmatrix} \xrightarrow{\textit{Jacobian matrix}} \mathbf{J} = \nabla f(x) = \begin{bmatrix} [\nabla f_1(x)]^T \\ \vdots \\ [\nabla f_m(x)]^T \end{bmatrix} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{bmatrix}$$

Hessian matrix : Jacobian matrix of gradient vector.

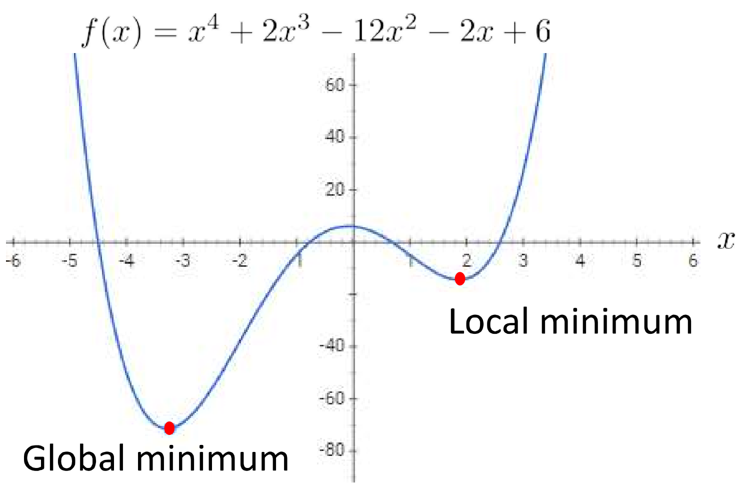
Global minimum x^*:

$$f(x^*) \leq f(x) \quad \underbrace{\forall x \in \mathbb{R}^n}_{\text{(Global)}}$$

Local minimum x^*:

$$f(x^*) \leq f(x) \quad \underbrace{\forall x \in B_{r>0}(x^*)}_{\text{Neighborhood (ball) around } x^*}$$

(Local)



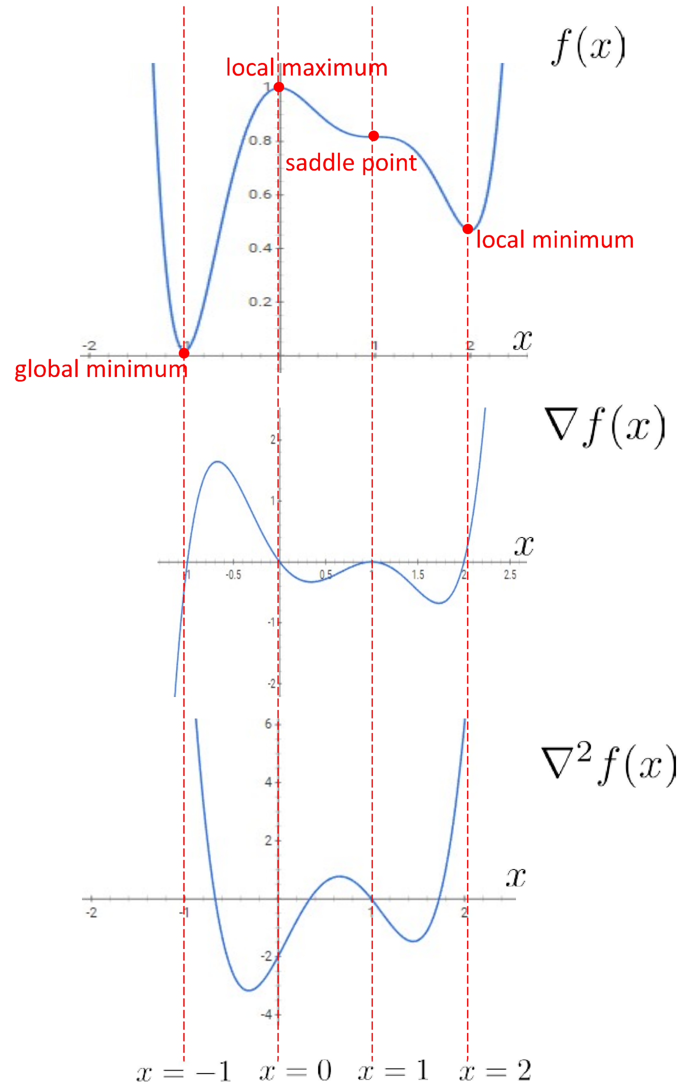
2) Optimality Conditions: Unconstrained Optimization

Unconstrained Optimization: $\underset{x \in \mathbb{R}^n}{\text{minimize}} \quad f(x)$

Necessary Condition: $\nabla f(x) = 0, \quad \nabla^2 f(x) \succeq 0$

Sufficient Condition: $\nabla f(x) = 0, \quad \nabla^2 f(x) \succ 0$

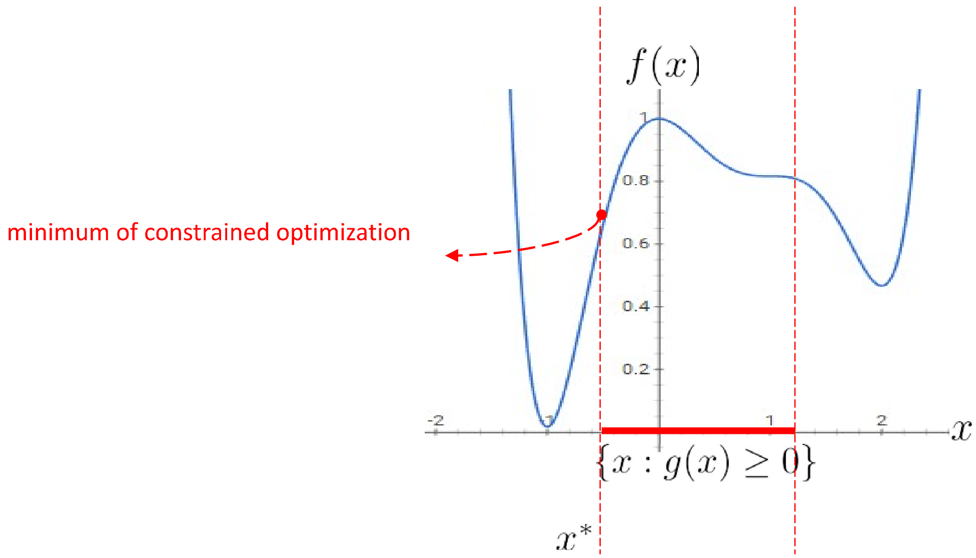
- To find x^* (argmin $f(x)$), Find the *roots* of $\nabla f(x^*)$.
- If obtained root satisfies $\nabla^2 f(x^*) \succ 0$, point x^* is (local/global) minimum.



Optimality Conditions: Constrained Optimization

Constrained Optimization:

$$\begin{array}{ll} \underset{x\in\mathbb{R}^n}{\text{minimize}} & f(x) \\ \text{subject to} & g(x) \geq 0 \end{array}$$



$$\nabla f(x^*) \neq 0$$

➢ Optimality conditions of *unconstrained optimization* are *not valid* for *constrained optimization*.

Optimization:

$$\begin{aligned} \underset{x\in\mathbb{R}^n}{\text{minimize}} \quad & f(x) \\ \text{subject to} \quad & g_i(x) \geq 0,\ i = 1, ..., n_g \\ & h_i(x) = 0,\ i = 1, .., n_h \end{aligned}$$

KKT (Karush-Kuhn-Tucker) Necessary Optimality Condition:

Lagrange function $L(x, \mu, \lambda) = f(x) - \sum_{i=1}^{n_g} \mu_i g_i(x) - \sum_{i=1}^{n_h} \lambda_i h_i(x)$

Lagrange multiplier

Optimality Cond.

$\nabla_x L(x, \mu, \lambda) = 0$ *Stationarity*

$(\nabla_{\lambda_i} L(x, \lambda) = 0)$ $-h_i(x^*) = 0,\ i = 1, ..., n_h$

$g_i(x^*) \geq 0, \quad i = 1, ..., n_g$ *Primal Feasibility*

$\mu_i \geq 0$ *Dual Feasibility*

$\mu_i^* g_i(x^*) = 0, \quad i = 1, ..., n_g$ *Dual Complementary Slackness*

KKT (Karush-Kuhn-Tucker) Optimality Condition

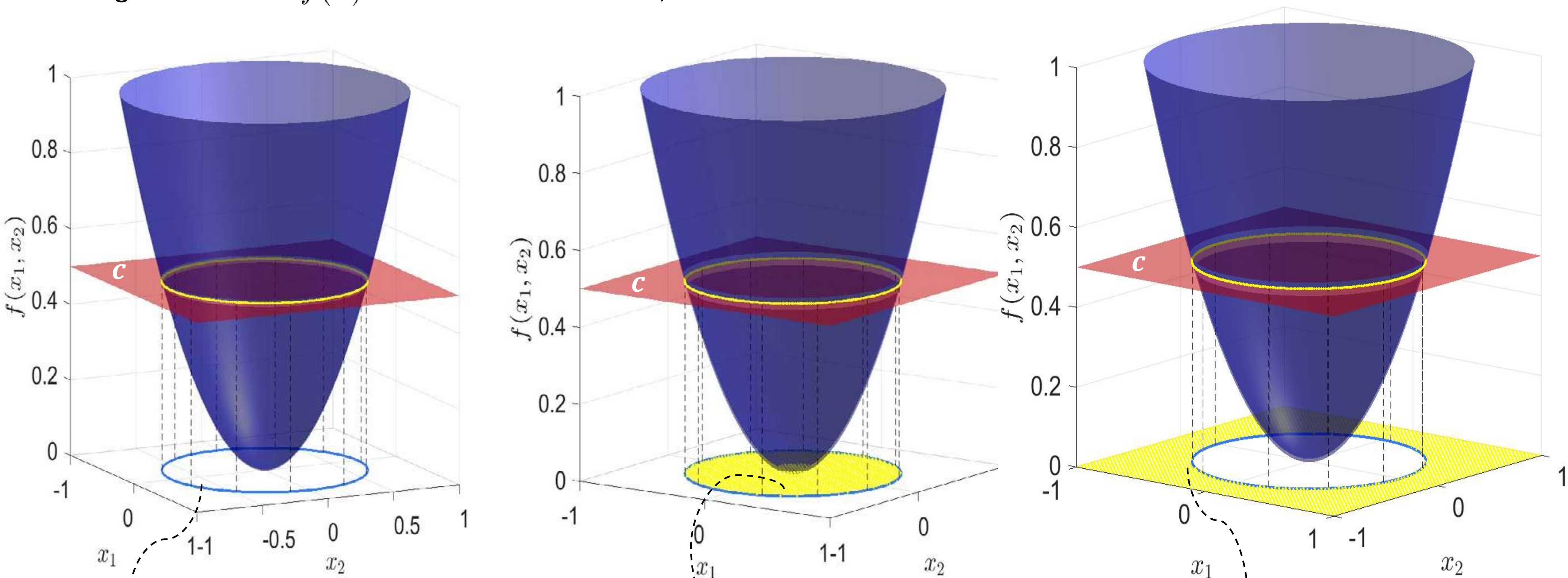
- Geometrical Interpretation

Basic Definitions

- Level Set
- Level Set and Gradient Vector
- Tangent Level Sets

Level, Sub-level, and Super-level Sets

- For a given function $f(x): \mathbb{R}^n \rightarrow \mathbb{R}$ and $c \in \mathbb{R}$, we define:



Level Set (Contour):

$$\{x \in \mathbb{R}^n : f(x) = c\}$$

Sub-Level Set :

$$\{x \in \mathbb{R}^n : f(x) \leq c\}$$

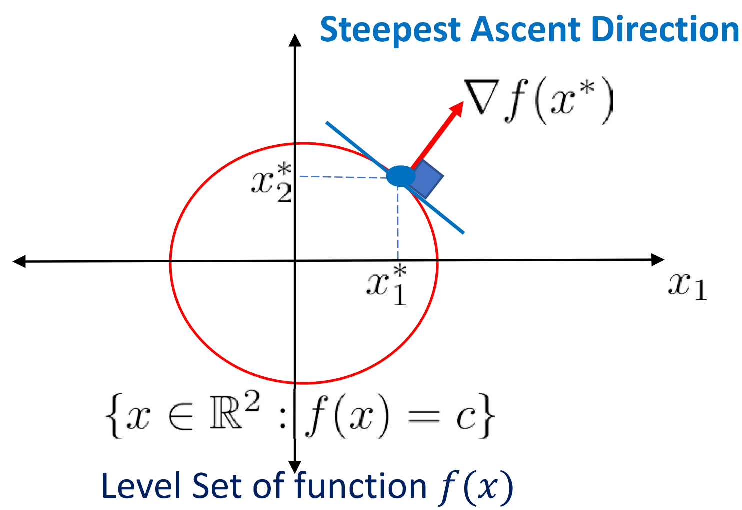
Supper-Level Set :

$$\{x \in \mathbb{R}^n : f(x) \geq c\}$$

Gradient and Level Sets

$\nabla f(x^*)$ is perpendicular vector to the level set.

$\nabla f(x^*)$ shows the direction of the steepest ascent.



Tangent Level sets

- Function $f(x) : \mathbb{R}^n \to \mathbb{R}$ and $g(x) : \mathbb{R}^n \to \mathbb{R}$

- Assume the level sets

  $\{x \in \mathbb{R}^n : f(x) = c_1\}$ and $\{x \in \mathbb{R}^n : g(x) = c_2\}$

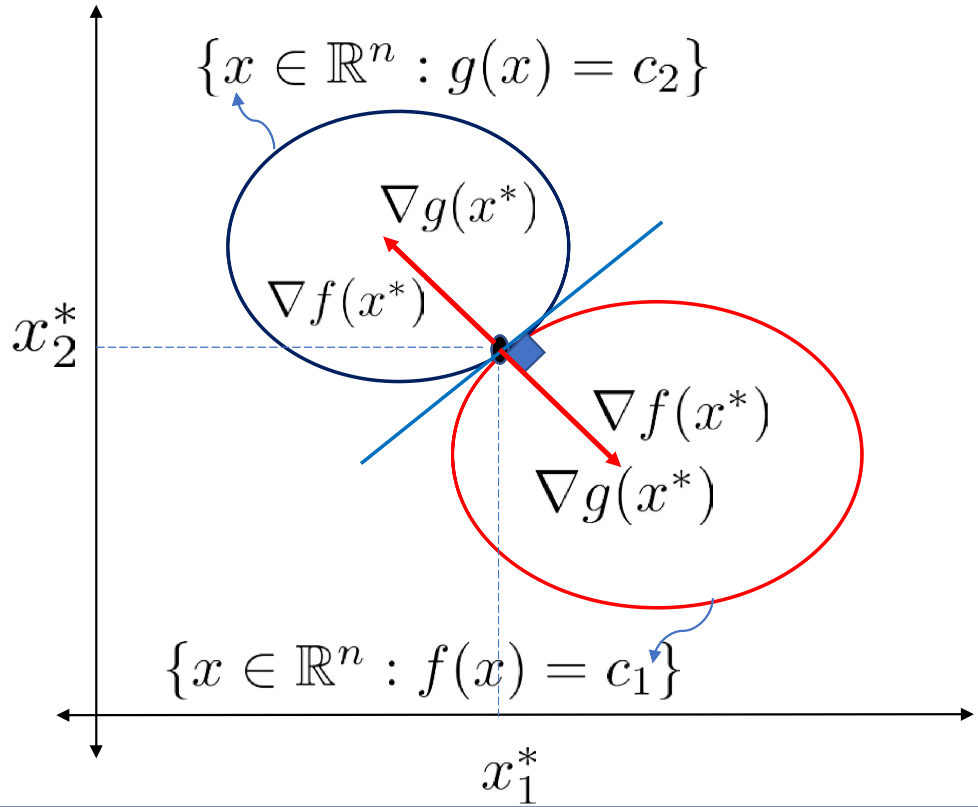
are tangent at the point x^*.

What is the relationship of the gradient vectors of $f(x)$ and $g(x)$ at point x^*?

➢ Gradient vectors $\nabla f(x^*)$ and $\nabla g(x^*)$ are parallel at the tangent point x^*.

$$\nabla f(x^*) = \lambda \nabla g(x^*)$$

Constant (+/-)

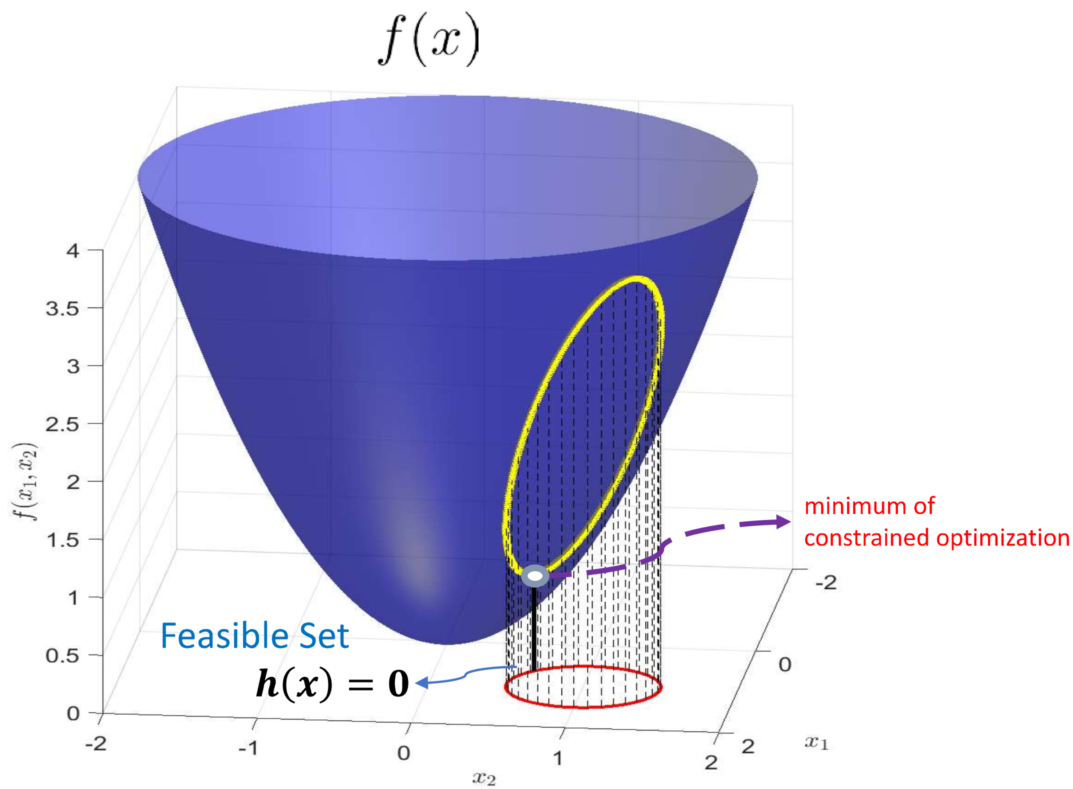


2) Optimality Conditions: Optimization with "Equality" Constraints

Optimization with "Equality" Constraint:

$$\underset{x\in\mathbb{R}^n}{\text{minimize}} \quad f(x)$$
$$\text{subject to} \quad h(x) = 0$$

- We obtain the optimality condition
  by looking at the level sets of $f(x)$, $\mathrm{h(x)}$.



2) Optimality Conditions: Optimization with "Equality" Constraints

$$\underset{x \in \mathbb{R}^n}{\text{minimize}} \quad f(x)$$
$$\text{subject to} \quad h(x) = 0$$

Optimality Condition: $h(x^*) = 0$ and $\nabla f(x^*) = \lambda \nabla h(x^*)$

Standard Format:

$$L(x, \lambda) = f(x) - \lambda h(x) \xrightarrow{\text{Optimality Cond.}} \nabla_{x,\lambda} L(x, \lambda) = 0 \begin{cases} \nabla_x L(x, \lambda) = 0 \longrightarrow \nabla f(x^*) = \lambda \nabla h(x^*) \\ \nabla_\lambda L(x, \lambda) = 0 \longrightarrow -h(x^*) = 0 \end{cases}$$

Lagrange function *Lagrange multiplier*

$$\underset{x \in \mathbb{R}^n}{\text{minimize}} \quad f(x)$$
$$\text{subject to} \quad h_i(x) = 0, \quad i = 1, \ldots, n_h$$

$$L(x, \lambda) = f(x) - \sum_{i=1}^{n_h} \lambda_i h_i(x) \xrightarrow{\text{Optimality Cond.}^1} \begin{matrix} \nabla_x L(x, \lambda) = 0 \\ \nabla_\lambda L(x, \lambda) = 0 \end{matrix}$$

Lagrange function

.[1]Note that λ could be +/-. Hence, Lagrange function could also take the following form: $L(x, \lambda) = f(x) + \sum_{i=1}^{n_h} \lambda_i h_i(x)$

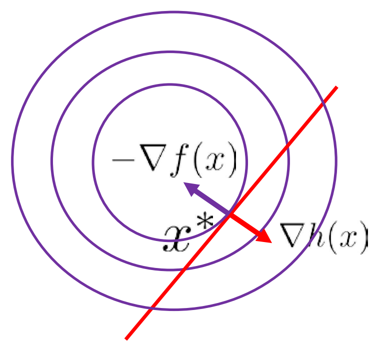
Chapter 3.1: Necessary Conditions For Equality Constraints, "Nonlinear Programming" Dimitri Bertsekas, MIT,

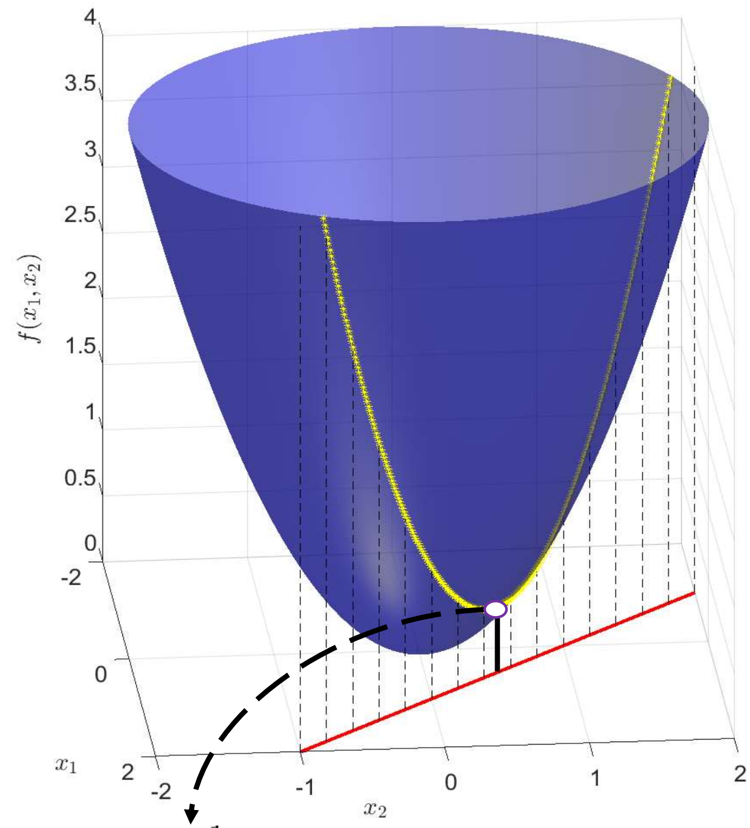
2) Optimality Conditions: Optimization with “Equality” Constraints

Example

$$\begin{aligned} \underset{x\in\mathbb{R}^2}{\text{minimize}} \quad & x_1^2 + x_2^2 \\ \text{subject to} \quad & x_1 + x_2 - 1 = 0 \end{aligned}$$

Lagrange function: $L(x,\lambda) = (x_1^2 + x_2^2) - \lambda(x_1 + x_2 - 1)$





Optimality Cond.

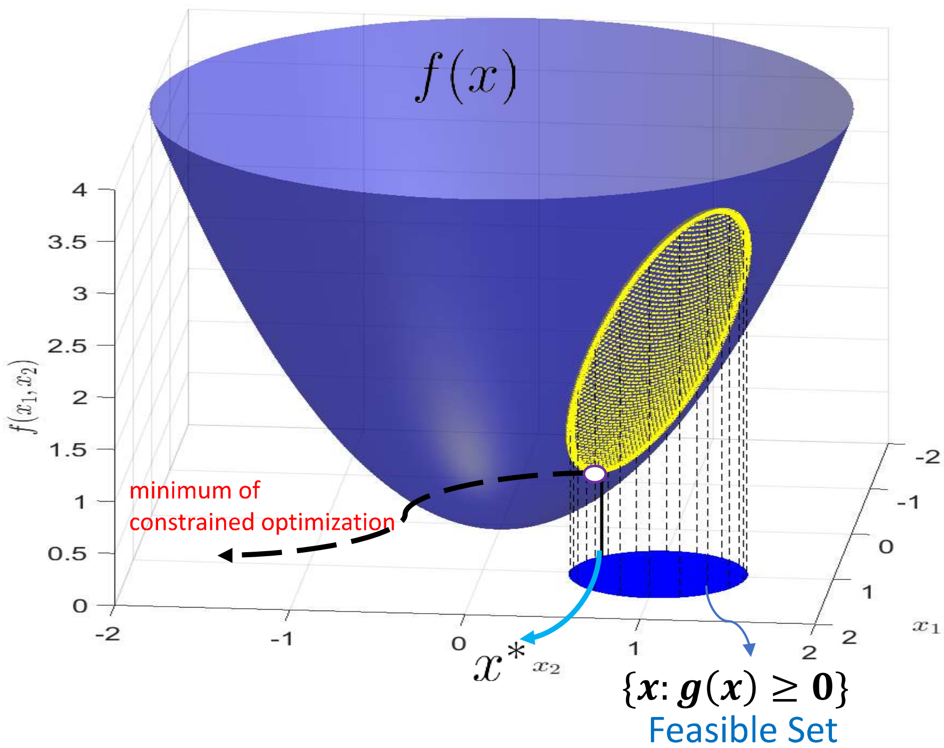
$$\nabla_{x,\lambda} L(x,\lambda) = 0 \begin{cases} \nabla_x L(x,\lambda) = 0 \longrightarrow \begin{bmatrix} 2x_1^* \\ 2x_2^* \end{bmatrix} - \lambda \begin{bmatrix} 1 \\ 1 \end{bmatrix} = 0 \\ \nabla_\lambda L(x,\lambda) = 0 \longrightarrow x_1^* + x_2^* - 1 = 0 \end{cases}$$

$$x_1^* = x_2^* = \tfrac{1}{2}, \ \ \lambda^* = 1$$

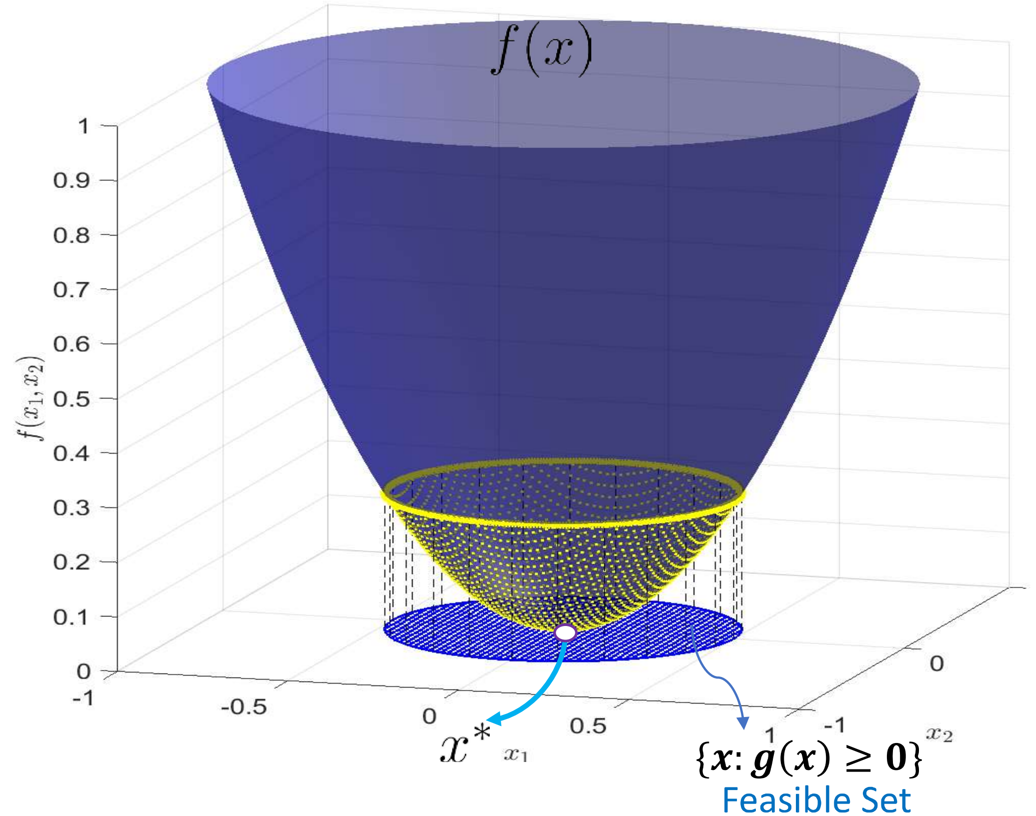
3) Optimality Conditions: Optimization with "Inequality" Constraints

$$\begin{aligned} \underset{x\in\mathbb{R}^n}{\text{minimize}} \quad & f(x) \\ \text{subject to} \quad & g(x) \geq 0 \end{aligned}$$

Case 1: x^* is on the boundary of the feasible region.



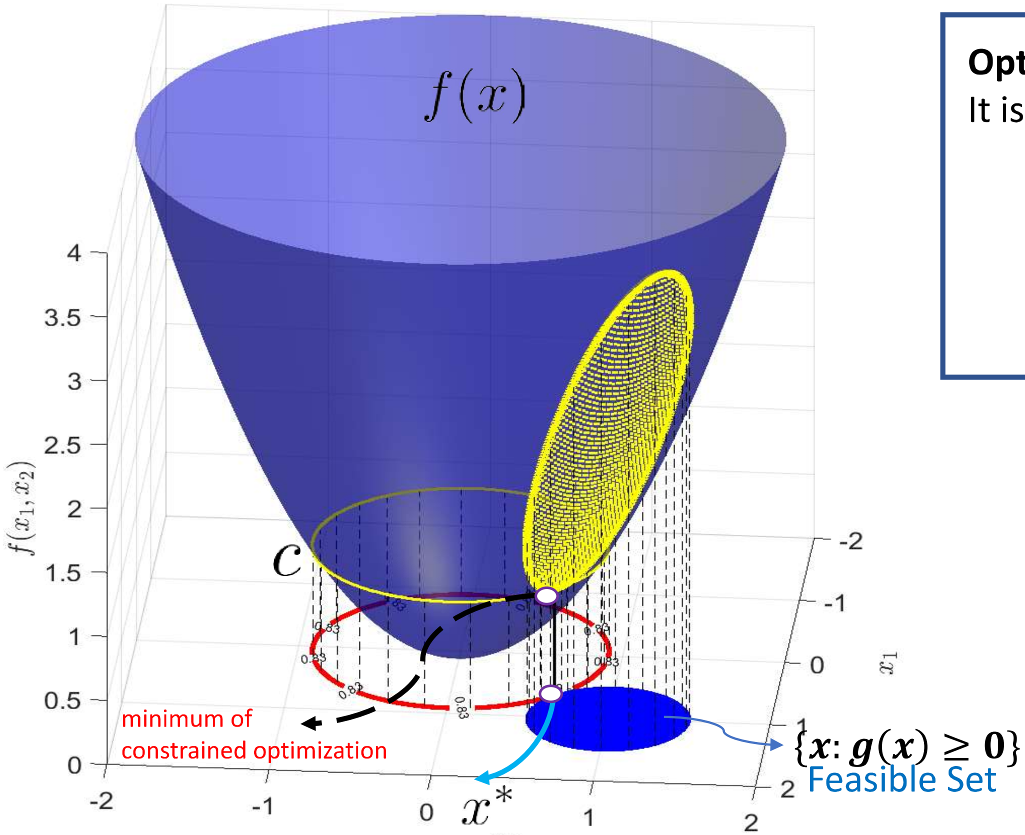
Case 2: x^* is inside the feasible region.



3) Optimality Conditions: Optimization with "Inequality" Constraints

$$\underset{x\in\mathbb{R}^n}{\text{minimize}} \quad f(x)$$
$$\text{subject to} \quad g(x) \geq 0$$

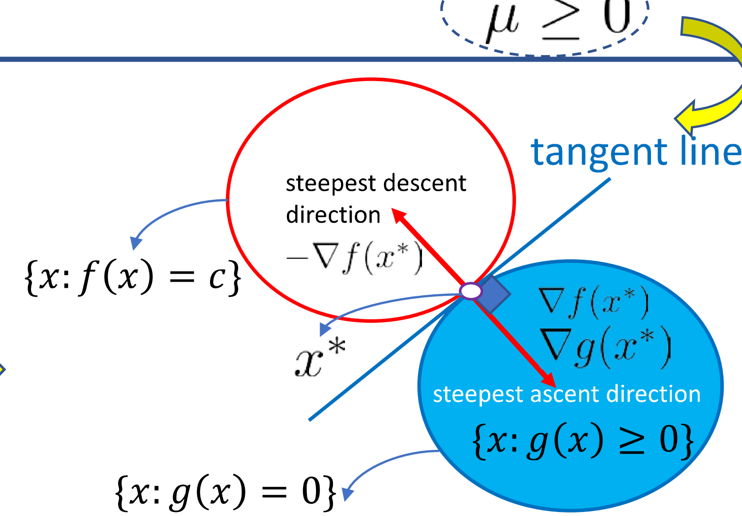
Case 1: x^* is on the boundary of the feasible region.



Optimality Condition:
It is similar to constrained optimization with equality constraint.

$$g(x^*) = 0 \quad \text{and} \quad \nabla f(x^*) = \mu \nabla g(x^*)$$

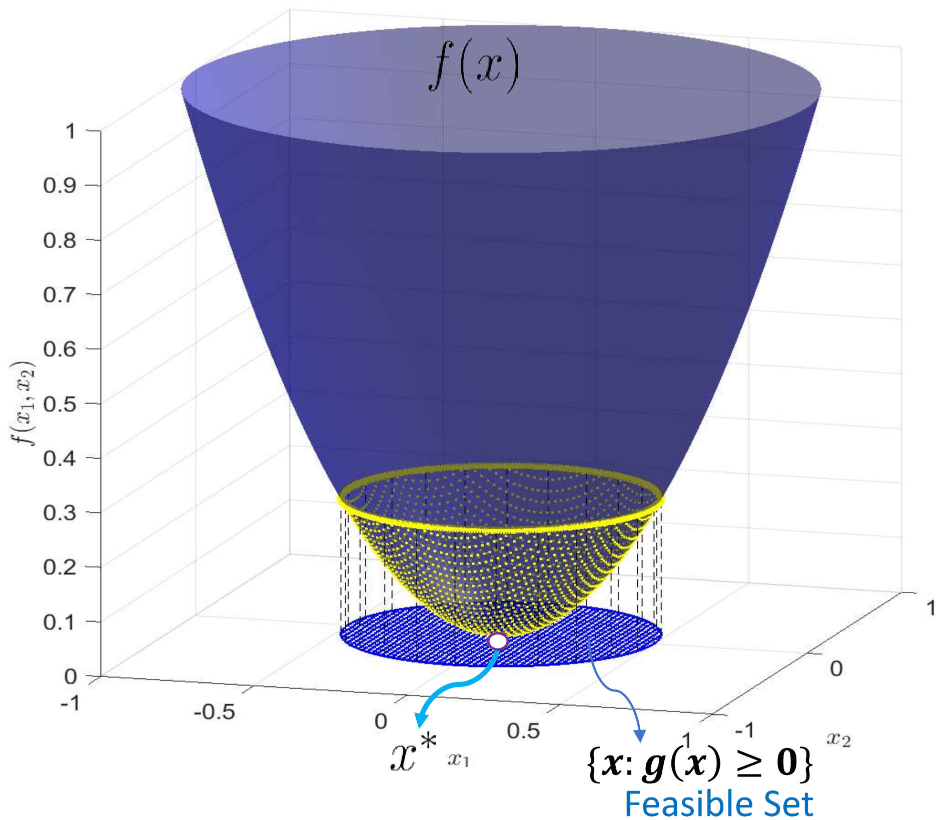
$$\mu \geq 0$$



3) Optimality Conditions: Optimization with "Inequality" Constraints

$$\underset{x \in \mathbb{R}^n}{\text{minimize}} \quad f(x)$$
$$\text{subject to} \quad g(x) \geq 0$$

Case 2: x^* is inside the feasible region.



Optimality Condition:
It is similar to *"unconstrained"* optimization.

$$g(x^*) > 0 \quad \text{and} \quad \boxed{\nabla f(x^*) = 0}$$

$$\nabla f(x^*) = \mu^* \nabla g(x^*)$$
$$\mu^* = 0$$

3) Optimality Conditions: Optimization with “Inequality” Constraints

$$\underset{x\in\mathbb{R}^n}{\text{minimize}} \quad f(x)$$
$$\text{subject to} \quad g(x) \geq 0$$

Case 1) Optimality Condition:
It is similar to constrained optimization with equality constraint.

$g(x^*) = 0$ and $\mu^* > 0$ and $\nabla f(x^*) = \mu^* \nabla g(x^*)$

Case 2) Optimality Condition:
It is similar to unconstrained optimization.

$g(x^*) > 0$ and $\mu^* = 0$ and $\nabla f(x^*) = \mu^* \nabla g(x^*)$

$g(x^*) \geq 0$ and $\mu^* \geq 0$ and $\mu^* g(x^*) = 0$ and $\nabla f(x^*) = \mu^* \nabla g(x^*)$

Standard Format:

$L(x,\mu) = f(x) - \mu g(x)$ **Optimality Cond.** $\Rightarrow$ $\nabla_x L(x^*,\mu^*) = 0$

$g(x^*) \geq 0 \qquad \mu^* \geq 0 \qquad \mu^* g(x^*) = 0$

Lagrange function *Lagrange multiplier*

$$\underset{x\in\mathbb{R}^n}{\text{minimize}} \quad f(x)$$
$$\text{subject to} \quad g_i(x) \geq 0, \ i = 1, \ldots, n_g$$

Lagrange function

$L(x,\lambda) = f(x) - \sum_{i=1}^{n_g} \mu_i g_i(x)$ **Optimality Cond.** $\Rightarrow$

$\nabla_x L(x^*,\mu^*) = 0 \qquad g_i(x^*) \geq 0$

$\mu_i^* g_i(x^*) = 0 \qquad \mu_i^* \geq 0$

Optimization:

$$\begin{aligned} \underset{x \in \mathbb{R}^n}{\text{minimize}} \quad & f(x) \\ \text{subject to} \quad & g_i(x) \geq 0, \; i = 1, ..., n_g \\ & h_i(x) = 0, \; i = 1, .., n_h \end{aligned}$$

Optimization:

$$\begin{aligned} \underset{x\in\mathbb{R}^n}{\text{minimize}} \quad & f(x) \\ \text{subject to} \quad & g_i(x) \geq 0,\ i = 1, ..., n_g \\ & h_i(x) = 0,\ i = 1, .., n_h \end{aligned}$$

KKT (Karush-Kuhn-Tucker) Necessary Optimality Condition:

Lagrange function $L(x,\mu,\lambda) = f(x) - \sum_{i=1}^{n_g} \mu_i g_i(x) - \sum_{i=1}^{n_h} \lambda_i h_i(x)$

Lagrange multiplier

Note: we might also see the following form of the Lagrange function:

$$L(x,\mu,\lambda) = f(x) + \sum_{i=1}^{n_g} \mu_i g_i(x) + \sum_{i=1}^{n_h} \lambda_i h_i(x)$$

i) Because inequality constraints of the original problem is in he form of $g_i(x) \leq 0$

ii) $\lambda : +/-$

Chapter 3.3: Inequality Constraints, "Nonlinear Programming" Dimitri Bertsekas, MIT,

Optimization

$$\underset{x \in \mathbb{R}^n}{\text{minimize}} \quad f(x)$$
$$\text{subject to} \quad g_i(x) \geq 0, \; i = 1, ..., n_g$$
$$h_i(x) = 0, \; i = 1, .., n_h$$

Optimality

Constrained Optimization:

$$\nabla_x L(x, \mu, \lambda) = 0$$
$$-h_i(x^*) = 0, \; i = 1, ..., n_h$$
$$g_i(x^*) \geq 0, \quad i = 1, ..., n_g$$
$$\mu_i \geq 0$$
$$\mu_i^* g_i(x^*) = 0, \quad i = 1, ..., n_g$$

system of nonlinear equations and inequalities

Unconstrained Optimization:

$$\nabla_x f(x) = 0$$

system of nonlinear equations

To find x^*, we need to solve a system of nonlinear equations and inequalities.

Newton's Method: Solving a system of nonlinear equations $F(x) = 0$

- We **can not** solve arbitrary "*nonlinear equations*" very easily.
- We **can** solve "*linear equations*" using the techniques from linear algebra.
- We linearize the system of "*nonlinear equations*" with a first-order Taylor and solve the obtained "*linear system*", and then iterate.

(0) *k=0,*

(1) *Linear approximation at point* x_k:

$$y = F(x_k) + \nabla F(x_k)^T (x - x_k)$$

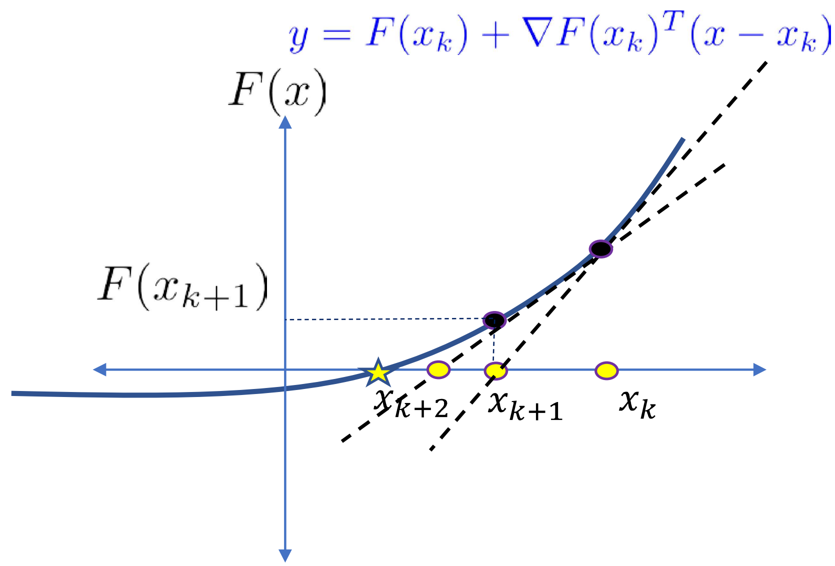
(2) *Solve linear system:*

$$y = 0 \longrightarrow F(x_k) + \nabla F(x_k)'(x - x_k) = 0$$

⇨ (1) *Linear approximation at point* x_k:

⇨ $$\boxed{x_{k+1} = x_k - (\nabla F(x_k))^{-1} F(x_k)}$$

(3) *Go to Step* (1)



system of nonlinear equations $F(x) = 0$ ***Update at each iteration:*** $x_{k+1} = x_k - \mathbf{J}_F^{-1} F(x_k)$

Jacobian matrix of $F(x)$

1) Newton's Method for Unconstrained optimization

$$\underset{x \in \mathbb{R}^n}{\text{minimize}} \quad f(x)$$

- **Optimality Condition:** $F(x) = 0 \longrightarrow \nabla f(x) = 0$

  system of nonlinear equations

- **Newton's method:** $\longrightarrow$

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

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

$$\boxed{x_{k+1} = x_k - \mathbf{H}_f^{-1} \nabla f(x_k)}$$

2) Newton's Method for Constrained optimization with Equality Constraints

$$\begin{aligned} &\underset{x\in\mathbb{R}^n}{\text{minimize}} && f(x) \\ &\text{subject to} && h_i(x)=0,\ i=1,\dots,n_h \end{aligned}$$

- **Optimality Condition:** $F(x,\lambda)=\begin{bmatrix}\nabla_x L(x,\lambda)\\ \nabla_\lambda L(x,\lambda)\end{bmatrix}=0$

system of nonlinear equations
($(n+n_h)$equations, $(n+n_h)$variables)

Lagrange function

$$L(x,\lambda)=f(x)-\sum_{i=1}^{n_h}\lambda_i h_i(x)=f(x)-\lambda^T h(x)$$

$$h(x)=[h_1(x),\dots,h_{n_h}(x)]^T$$

- **Newton's method:** $\begin{bmatrix}x_{k+1}\\ \lambda_{k+1}\end{bmatrix}=\begin{bmatrix}x_k\\ \lambda_k\end{bmatrix}-\mathbf{J}_F^{-1}F(x_k,\lambda_k)$

$$F(x,\lambda)=\begin{bmatrix}\nabla_x L(x,\lambda)\\ \nabla_\lambda L(x,\lambda)\end{bmatrix}=\begin{bmatrix}\nabla_x f(x)-\mathbf{J}_h^T\lambda\\ -h(x)\end{bmatrix}$$

$$\mathbf{J}_F=\begin{bmatrix}[\nabla F_1(x)]^T\\ [\nabla F_2(x)]^T\end{bmatrix}=\begin{bmatrix}[\nabla(\nabla_x L(x,\lambda))]^T\\ [\nabla(-h(x))]^T\end{bmatrix}=\begin{bmatrix}[\nabla^2_{xx}L(x,\lambda) & -\mathbf{J}_h^T]\\ [-\mathbf{J}_h & 0]\end{bmatrix}$$

$$\boxed{\begin{bmatrix}x_{k+1}\\ \lambda_{k+1}\end{bmatrix}=\begin{bmatrix}x_k\\ \lambda_k\end{bmatrix}-\begin{bmatrix}\nabla^2_{xx}L(x,\lambda) & -\mathbf{J}_h^T\\ -\mathbf{J}_h & 0\end{bmatrix}^{-1}\begin{bmatrix}\nabla_x L(x,\lambda)\\ -h(x)\end{bmatrix}}$$

3) Newton's Method for Optimization with Inequality Constraints

- Newton's method deal with solving a system of nonlinear equations $F(x) = 0$

- Optimality condition for **constrained optimization** (involving inequality constraint) is a "*system of nonlinear equations and inequalities*".

$$\begin{aligned} \underset{x \in \mathbb{R}^n}{\text{minimize}} \quad & f(x) \\ \text{subject to} \quad & \boxed{g_i(x) \geq 0, \;\; i = 1, ..., n_g} \\ & h_i(x) = 0, \;\; i = 1, .., n_h \end{aligned}$$

KKT:

$$\begin{aligned} & \nabla_x L(x, \mu, \lambda) = 0 \\ & -h_i(x^*) = 0, \;\; i = 1, ..., n_h \\ & \boxed{\begin{aligned} & g_i(x^*) \geq 0, \quad i = 1, ..., n_g \\ & \mu_i \geq 0 \end{aligned}} \\ & \mu_i^* g_i(x^*) = 0, \quad i = 1, ..., n_g \end{aligned}$$

Interior Point Method

- Reformulate optimization problem with Inequality Constraints as optimization with only equality constraints.
- Apply Newton's method to optimality *system of nonlinear equations* .

Interior Point Method

$$\begin{aligned} \underset{x\in\mathbb{R}^n}{\text{minimize}} \quad & f(x) \\ \text{subject to} \quad & g_i(x) \geq 0, \ \ i = 1, ..., n_g \\ & h_i(x) = 0, \ \ i = 1, ..., n_h \end{aligned}$$

Slack variables "s_i"

$$\begin{aligned} \underset{x\in\mathbb{R}^n, s\in\mathbb{R}^{n_g}}{\text{minimize}} \quad & f(x) \\ \text{subject to} \quad & g_i(x) - s_i = 0, \ i = 1, ..., n_g \\ & s_i \geq 0, \ i = 1, ..., n_g \\ & h_i(x) = 0, \ \ i = 1, ..., n_h \end{aligned}$$

$z = [s_i|_{i=1}^{n_g}, x_i|_{i=1}^{n}]^T$

$\bar{f}(z) = f(x)$

$\bar{h}(z) = [g(x)_i - s_i \ |_{i=1}^{n_g}, h_i(x) \ |_{i=1}^{n_h}]^T$

$$\begin{aligned} \underset{z\in\mathbb{R}^{n+n_g}}{\text{minimize}} \quad & \bar{f}(z) \\ \text{subject to} \quad & \bar{h}(z) = 0 \\ & z_i \geq 0, \ \ i = 1, ..., n_g \end{aligned}$$

Optimization with equality constraints

$$\begin{aligned} \underset{z\in\mathbb{R}^{n+n_g}}{\text{minimize}} \quad & \bar{f}(z) + \text{Penalty Function}(z_i) \\ \text{subject to} \quad & \bar{h}(z) = 0 \end{aligned}$$

Use Newton's Method

Interior Point Method

Optimization with equality constraints

$$\underset{z\in\mathbb{R}^{n+n_g}}{\text{minimize}} \quad \bar{f}(z) - \sum_{i=1}^{n_g} \frac{1}{t} ln(z_i)$$
$$\text{subject to} \quad \bar{h}(z) = 0$$

Use Newton's Method

(0) *k=0, feasible* z_k , t_k

(1) Inner Loop: *Compute* $z^*\ (k+1)$ *by solving constrained optimization using newton's method.*

Lagrange function $L(z,\lambda) = \bar{f}(z) - \frac{1}{t}\sum_{i=1}^{n_g} ln(z_i) - \lambda^T \bar{h}(z)$ *Lagrange multipliers*

Optimality Condition: $F(z,\lambda) = \begin{bmatrix} \nabla_z L(z,\lambda) \\ \nabla_\lambda L(z,\lambda) \end{bmatrix} = 0 \longrightarrow \boxed{\begin{bmatrix} z_{k+1} \\ \lambda_{k+1} \end{bmatrix} = \begin{bmatrix} z_k \\ \lambda_k \end{bmatrix} - \mathbf{J}_F^{-1} F(z_k,\lambda_k)}$

(2) *Outer Loop:* *Increase* $t_{k+1} = \beta t_k$

(3) $k \leftarrow k+1$, Go to step (1)

- Every limit point of a sequence $\{z_k\}$ generated by a barrier method solves the original constrained problem.
- For more information and convergence analysis see:

- Stephen J. Wright, "On the convergence of the Newton/log-barrier method", Mathematical Programming , Volume 90, Issue 1, pp 71–100, 2001.
- A. Forsgren, P. E. Gill, M. H. Wright "Interior Methods for Nonlinear Optimization", SIAM REVIEW, Vol. 44, No. 4, pp. 525–597, 2002.
- Chapter 4.1: Barrier and Interior Point Methods, "Nonlinear Programming" Dimitri Bertsekas, MIT.

1

$$\underset{x\in\mathbb{R}^n}{\text{minimize}} \quad f(x)$$

Optimality Condition

$$\nabla_x f(x) = 0$$

Newton's Method Step

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

2

$$\underset{x\in\mathbb{R}^n}{\text{minimize}} \quad f(x)$$

$$\text{subject to} \quad h_i(x) = 0,\ i = 1, ..., n_h$$

Optimality Condition

$$F = \begin{cases} \nabla_x L(x,\mu,\lambda) = 0 \\ -h_i(x^*) = 0,\ i = 1,...,n_h \end{cases}$$

Lagrange Function

$$L(x,\lambda) = f(x) - \lambda^T h(x)$$

Newton's Method Step

$$\begin{bmatrix} x_{k+1} \\ \lambda_{k+1} \end{bmatrix} = \begin{bmatrix} x_k \\ \lambda_k \end{bmatrix} - \mathbf{J}_F^{-1} F(x_k, \lambda_k)$$

$$\Downarrow$$

$$\begin{bmatrix} x_{k+1} \\ \lambda_{k+1} \end{bmatrix} = \begin{bmatrix} x_k \\ \lambda_k \end{bmatrix} - \begin{bmatrix} \nabla^2_{xx} L(x,\lambda) & -\mathbf{J}_h^T \\ -\mathbf{J}_h & 0 \end{bmatrix}^{-1} \begin{bmatrix} \nabla_x L(x,\lambda) \\ -h(x) \end{bmatrix}$$

3

$$\underset{x\in\mathbb{R}^n}{\text{minimize}} \quad f(x)$$

$$\text{subject to} \quad g_i(x) \geq 0,\ i = 1, ..., n_g$$

$$h_i(x) = 0,\ i = 1, ..., n_h$$

***Optimality Condition* (KKT)**

$$\nabla_x L(x,\mu,\lambda) = 0 \qquad -h_i(x^*) = 0,\ i = 1,...,n_h$$

$$g_i(x^*) \geq 0, \quad i = 1,...,n_g$$

$$\mu_i^* \geq 0 \quad \mu_i^* g_i(x^*) = 0, \quad i = 1,...,n_g$$

Optimization with equality constraints

$$z = [s_i|_{i=1}^{n_g}, x_i|_{i=1}^{n}]^T \qquad \bar{f}(z) = f(x)$$

$$\bar{h}(z) = [g(x)_i - s_i\ |_{i=1}^{n_g}, h_i(x)\ |_{i=1}^{n_h}]^T$$

$$\underset{z\in\mathbb{R}^{n+n_g}}{\text{minimize}} \quad \bar{f}(z) - \sum_{i=1}^{n_g} \frac{1}{t} ln(z_i)$$

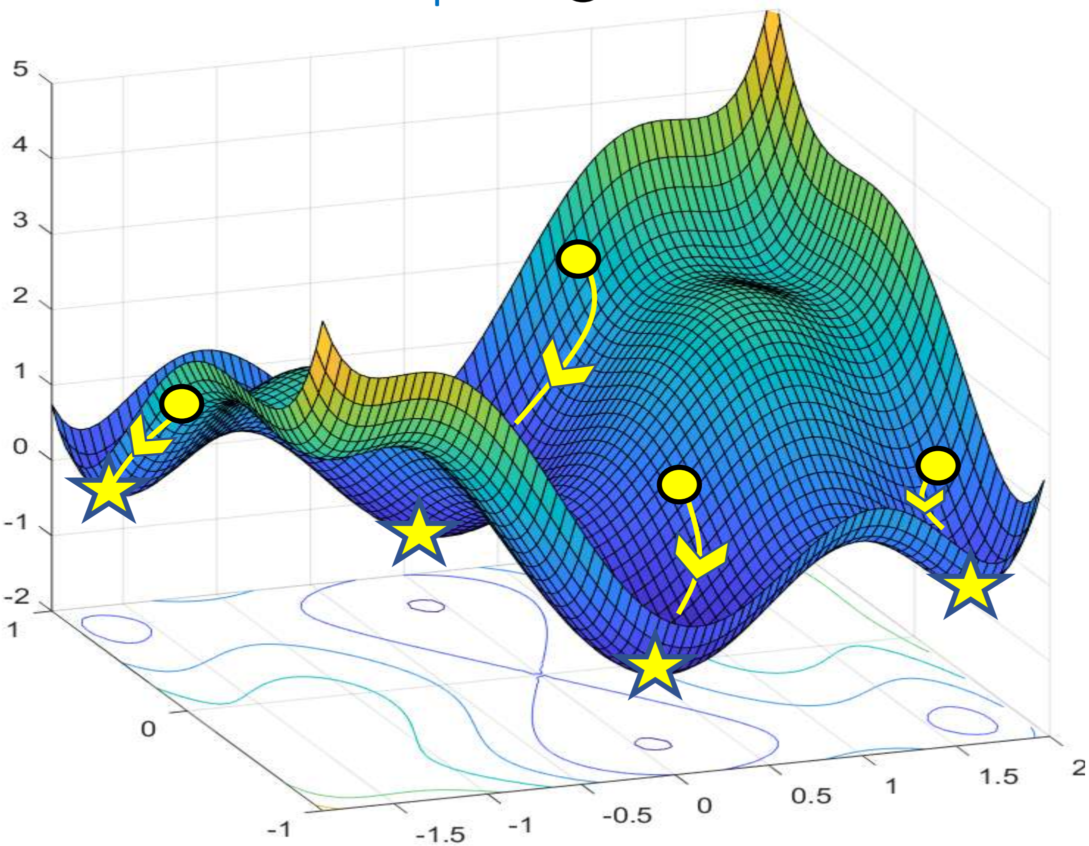
$$\text{subject to} \quad \bar{h}(z) = 0$$

$$\Downarrow$$

2

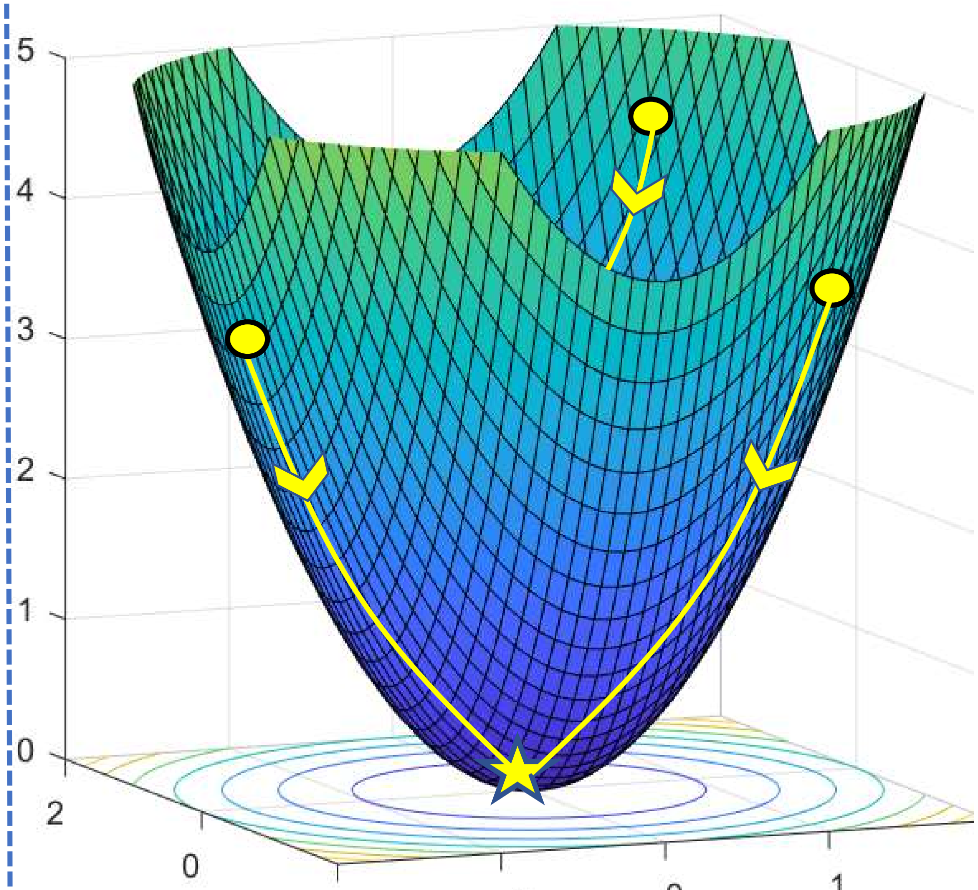
Nonconvex Optimization

- Multiple local minima
- Sensitive to initial point



Convex Optimization

- Unique minimum: global/local



Convex Optimization

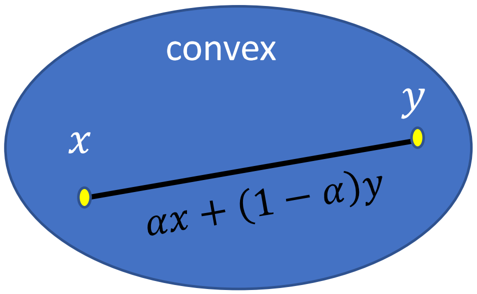
Convex Optimization:

$$\begin{aligned} \underset{x \in \mathbb{R}^n}{\text{minimize}} \quad & f(x) \\ \text{subject to} \quad & g_i(x) \leq 0, \;\; i = 1, ..., n_g \\ & h_i(x) = 0, \;\; i = 1, .., n_h \end{aligned}$$

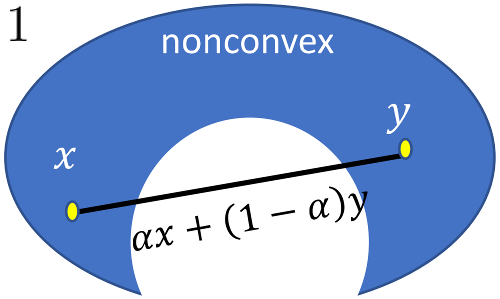
$$\left.\begin{aligned} & f : \mathbb{R}^n \to \mathbb{R} \\ & g_i : \mathbb{R}^n \to \mathbb{R}, \quad i = 1, ..., n_g \\ & h_i : \mathbb{R}^n \to \mathbb{R}, \quad i = 1, ..., n_h \end{aligned}\right\} \text{Convex Functions}$$

Convex Set:

Set $C \in \mathbb{R}^n$ is a convex set if a line segment joining any two elements lies entirely in the set.

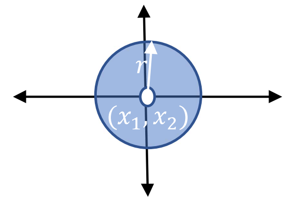


$$x, y \in C \rightarrow \alpha x + (1-\alpha)y \in C, \quad \forall 0 \leq \alpha \leq 1$$



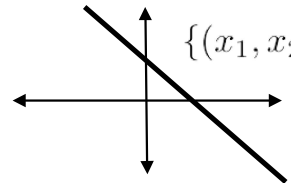
Examples:

Norm ball $\{x : ||x|| \leq r\}$ $||.||$:norm r:radius



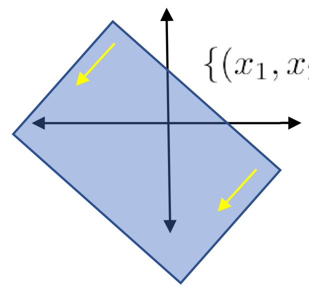
$\{(x_1, x_2) : \sqrt{x_1^2 + x_2^2} \leq r\}$

Hyperplane $\{x : a^T x = b\}$ a, b :vectors



$\{(x_1, x_2) : \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 1\}$

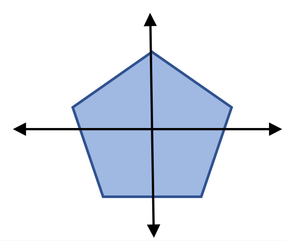
Halfspace $\{x : a^T x \leq b\}$ a, b :vectors



$\{(x_1, x_2) : \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \leq 1\}$

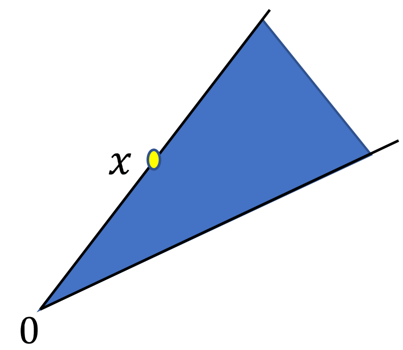
Affine Space $\{x : Ax = b\}$ A :matrix b:vectors

Polyhedron $\{x : Ax \leq b\}$



Cone:

A cone $C \in \mathbb{R}^n$ is a set such that $x \in C \rightarrow \alpha x \in C, \quad \forall \alpha \geq 0$



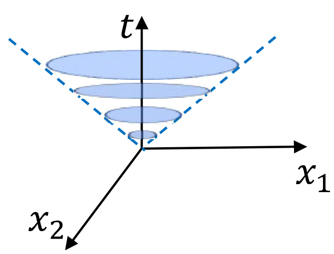
Convex Cone: $x, y \in C \rightarrow \alpha x + (1-\alpha) y \in C, \quad \forall 0 \leq \alpha \leq 1$

Examples:

Norm Cone $\{x : ||x|| \leq t\}$

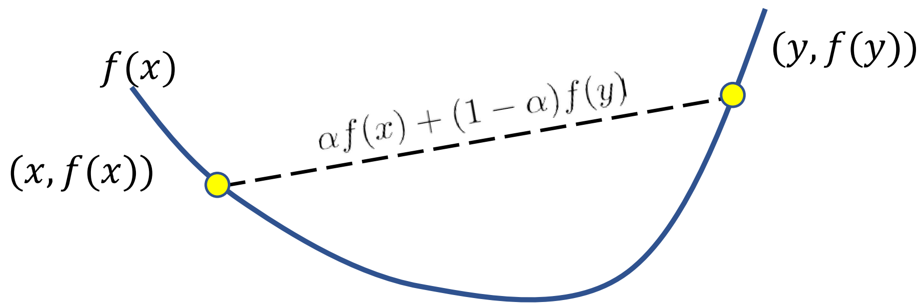
Under the l_2 norm, this is called a second-order cone.

$$\{(x_1, x_2) : \sqrt{x_1^2 + x_2^2} \leq t\}$$



Convex Function:

Function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is convex if: $f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha)f(y), \quad \forall 0 \leq \alpha \leq 1$

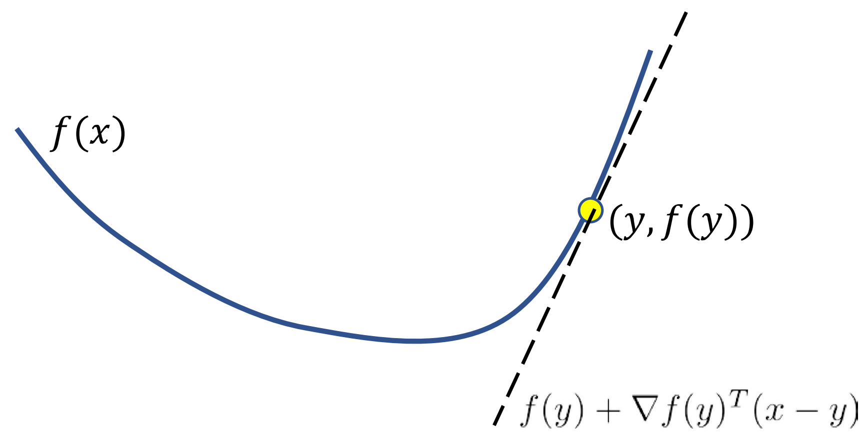


f lies below the line segment joining $f(x)$ and $f(y)$

- First order convexity condition:

$$f(x) \geq f(y) + \nabla f(y)^T(x-y), \quad \forall x, y \in dom(f)$$

- Second order convexity condition: $\nabla^2 f(x) \geq 0, \quad \forall x \in dom(f)$



- Domain of a convex function is a convex set.

Class of Convex Optimizations

Linear program (LP)

$$
\begin{aligned}
\underset{x}{\text{minimize}} \quad & c^T x \\
\text{subject to} \quad & Ax = b \\
& x \geq 0.
\end{aligned}
$$

Quadratic program (QP)

$$
\begin{aligned}
\underset{x}{\text{minimize}} \quad & \frac{1}{2} x^T P_0 x + q_0^T x + c_0 \\
\text{subject to} \quad & Ax = b \\
& x \geq 0.
\end{aligned}
$$

where $P_0 \succcurlyeq 0$

Quadratically Constrained Quadratic Program (QCQP)

$$
\begin{aligned}
\underset{x}{\text{minimize}} \quad & \frac{1}{2} x^T P_0 x + q_0^T x + c_0 \\
\text{subject to} \quad & \frac{1}{2} x^T P_i x + q_i^T x + c_i, \quad i = 1, ..., m \\
& Ax = b \\
& x \geq 0.
\end{aligned}
$$

where $P_i \succcurlyeq 0, \ i = 0, ..., m$

Second-Order Cone Program (SOCP)

$$
\begin{aligned}
\underset{x}{\text{minimize}} \quad & c^T x \\
\text{subject to} \quad & Ax = b \\
& ||C_i x + d_i||_2 \leq e_i^T x + f_i, \ i = 1, ..., m
\end{aligned}
$$

Semidefinite Program (SDP)

$$
\begin{aligned}
\underset{X}{\text{minimize}} \quad & C \bullet X \\
\text{subject to} \quad & A_i \bullet X = b_i \quad i = 1, \ldots, m. \\
& X \succcurlyeq 0.
\end{aligned}
$$

Cone Program (CP)

$$
\begin{aligned}
\underset{x}{\text{minimize}} \quad & c^T x \\
\text{subject to} \quad & Ax = b \\
& x \in K
\end{aligned}
$$