中国科学技术大学
University of Science and Technology of China

GAMES102

GAMES 102在线课程

几何建模与处理基础

刘利刚

中国科学技术大学

中国科学技术大学
University of Science and Technology of China

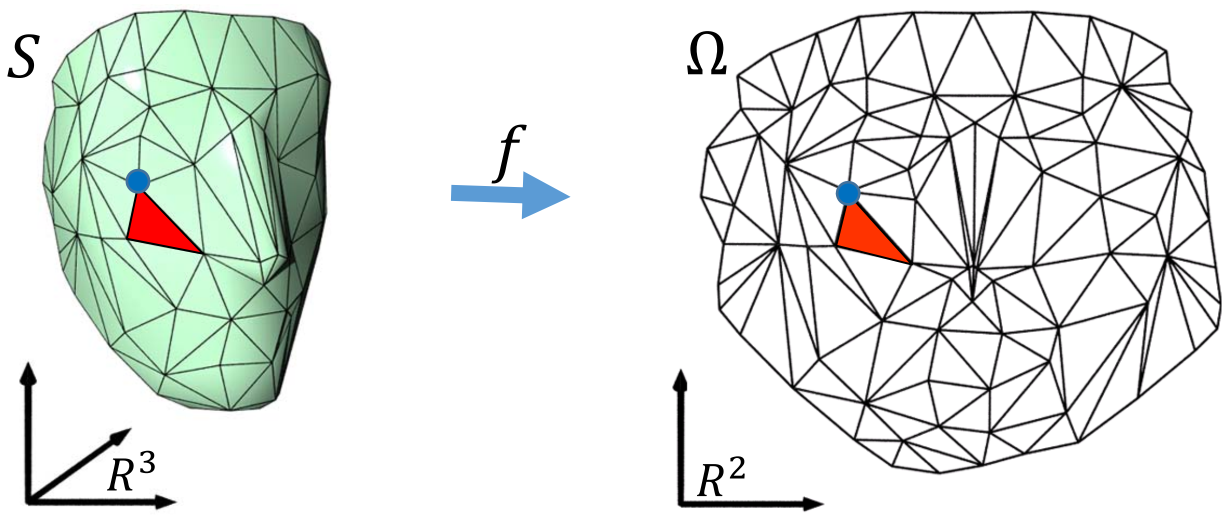
GAMES 102在线课程：几何建模与处理基础

几何优化

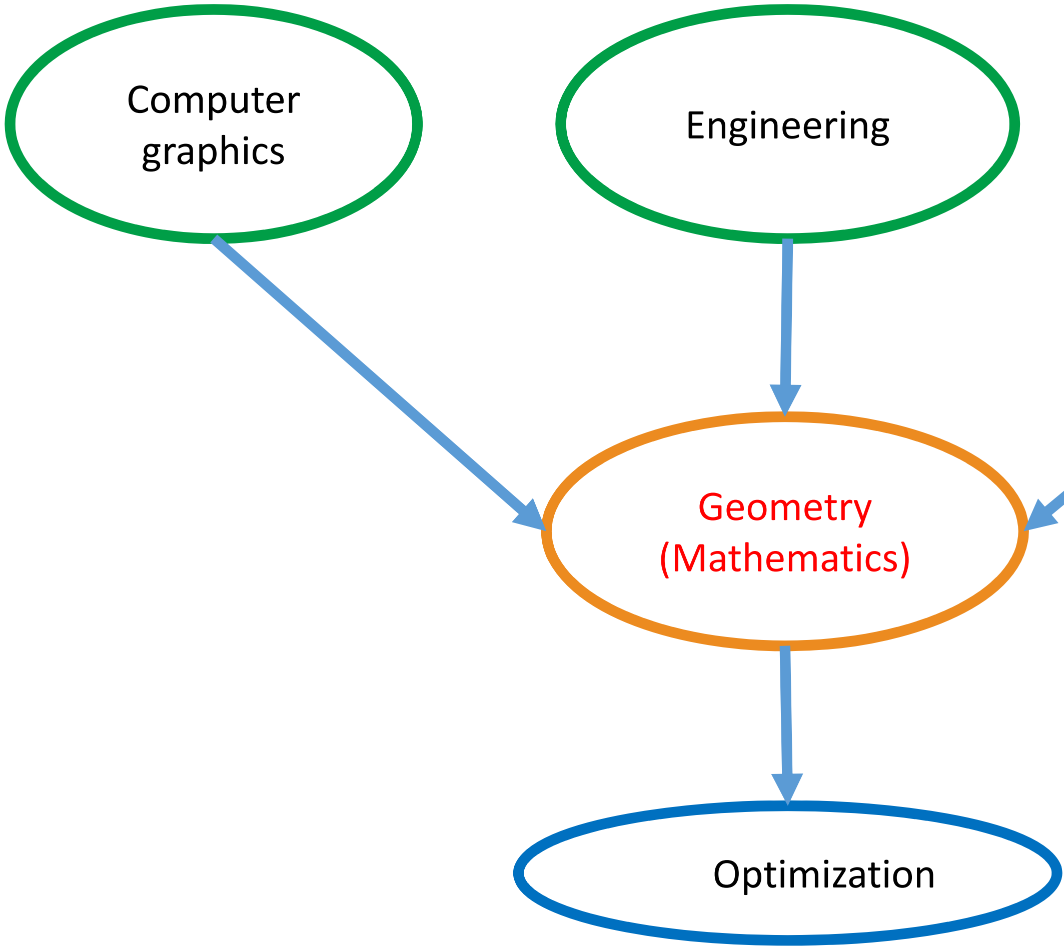
回顾：参数化中的优化问题

$$\min_{V} E(V) = \sum_{t \in T} \left(\sigma_1^2 + \frac{1}{\sigma_1^2} + \sigma_2^2 + \frac{1}{\sigma_2^2}\right)$$

s.t. $\sigma_1 \sigma_2 > 0, \quad \forall t$



回顾：几何处理中的优化问题



Geometry Problem and Modeling

1. Formulate an objective energy $E(x)$
2. Define constraints, if apply
   - Equality / Inequality
   - Linear / Nonlinear
3. Numerical optimization

$$\begin{aligned} \underset{x \in \mathbb{R}^n}{\text{minimize}}\ & E(x) \\ \text{subject to}\ & c_1(x) = d_1 \\ & c_2(x) > d_2 \end{aligned}$$

Fundamentals

优化问题的一般形式

高维实值函数：$f:\mathbb{R}^n \to \mathbb{R}$

$$\min_{x\in\mathbb{R}^n} f(x)$$
$$\text{s.t. } g(x) = 0$$
$$h(x) \geq 0$$

目标函数 or 能量函数

等式约束

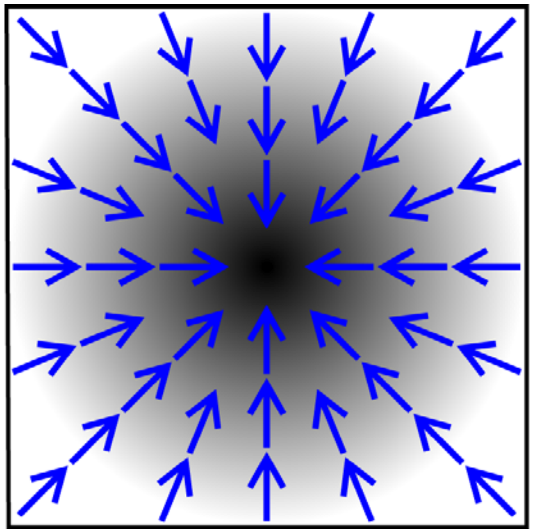
不等式约束

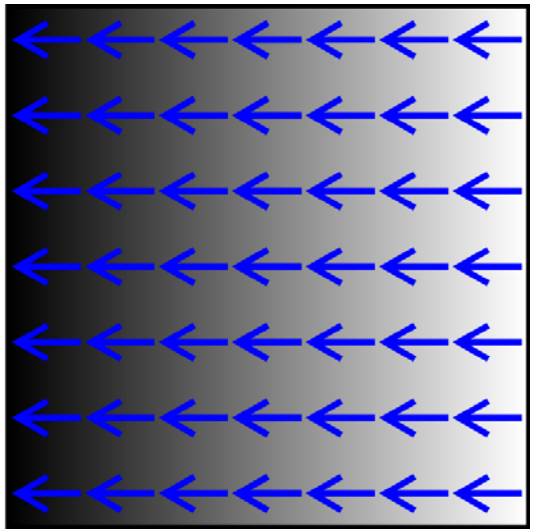
- Two roles
  - Client: Which optimization tool is relevant?
    - 不同的优化问题须用不同的优化方法
  - Designer: Can I design an algorithm for this problem?
    - 特定的优化问题需要设计特定的优化方法达到最佳性能
- Optimization is a **huge** field.

梯度 (Gradient)：一阶导数

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

$$\to \nabla f = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \dots, \frac{\partial f}{\partial x_n} \right)$$



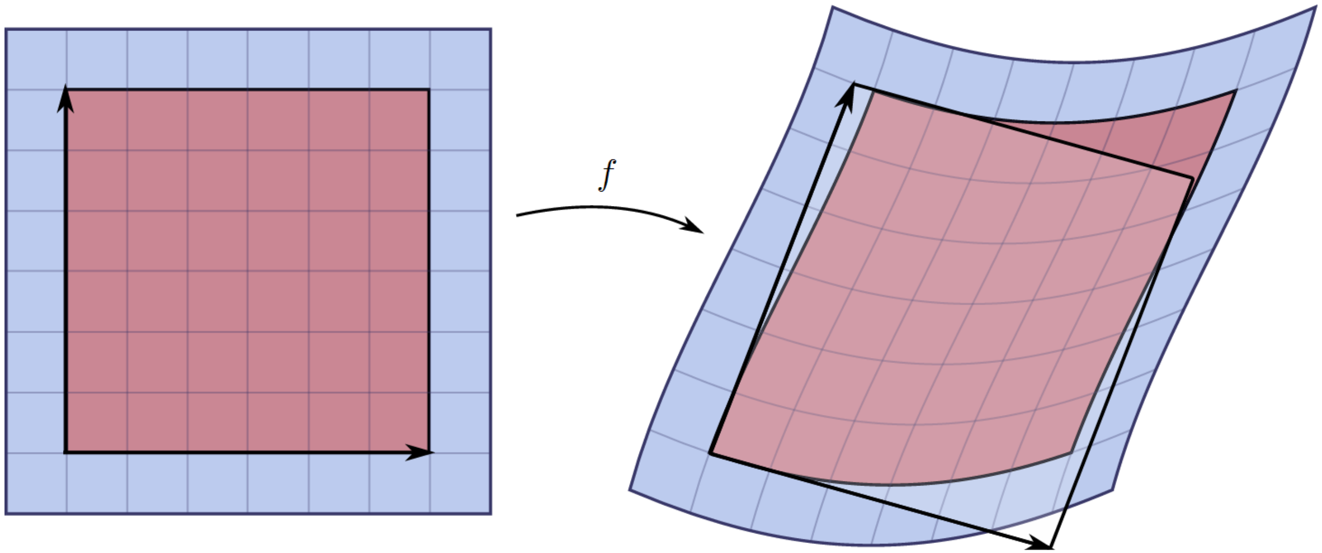


Courtesy of Justin Solomon and David Bommes

Jacobian: 一阶“导数”矩阵

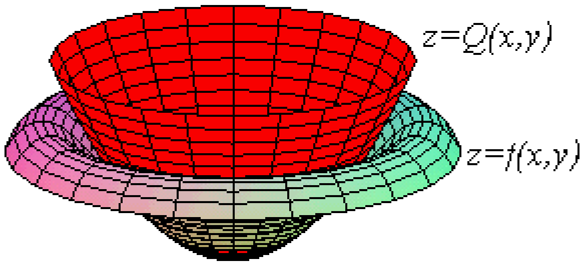
$$f : \mathbb{R}^n \to \mathbb{R}^m$$

$$\to (Df)_{ij} = \frac{\partial f_i}{\partial x_j}$$



Hessian： 二阶“导数”矩阵

$$f:\mathbb{R}^n \to \mathbb{R} \to H_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}$$

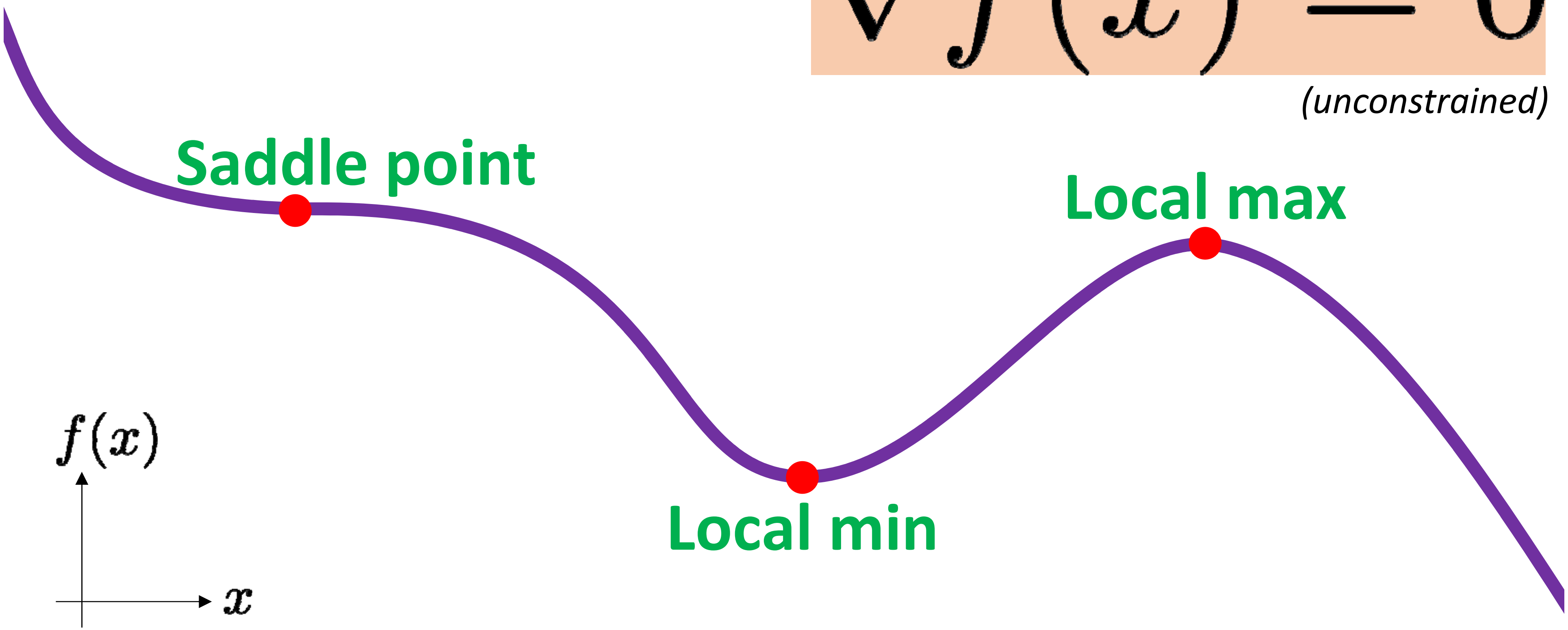


$$f(x) \approx f(x_0) + \nabla f(x_0)^\top (x - x_0) + (x - x_0)^\top H f(x_0)(x - x_0)$$

驻点（Critical point）

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

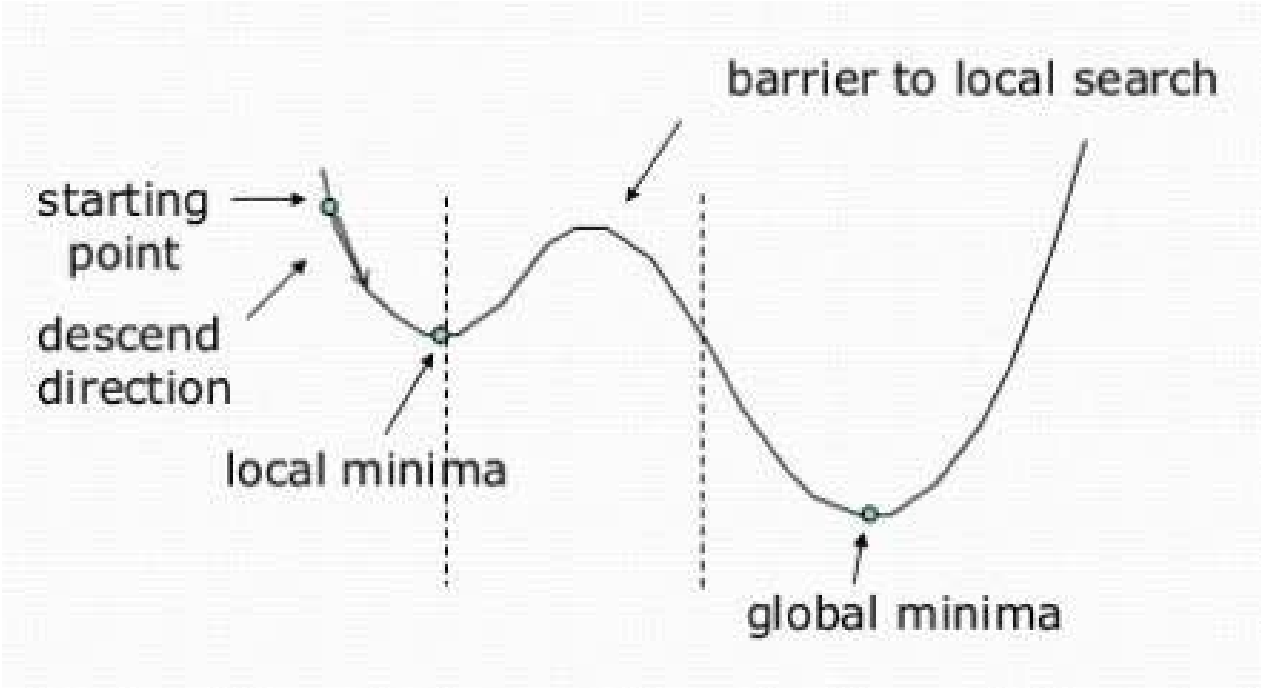
(unconstrained)



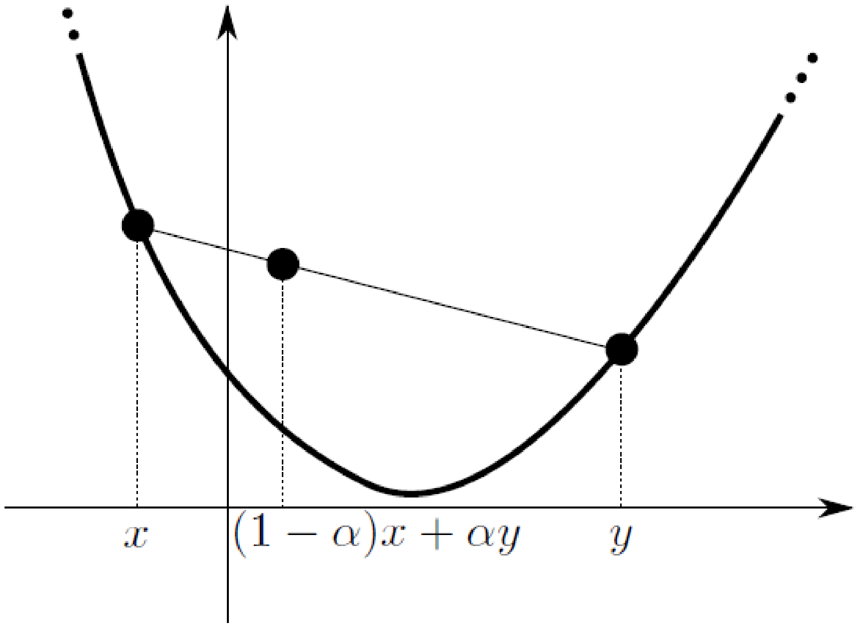
Critical points may not be minima.

一般非线性函数的最小值

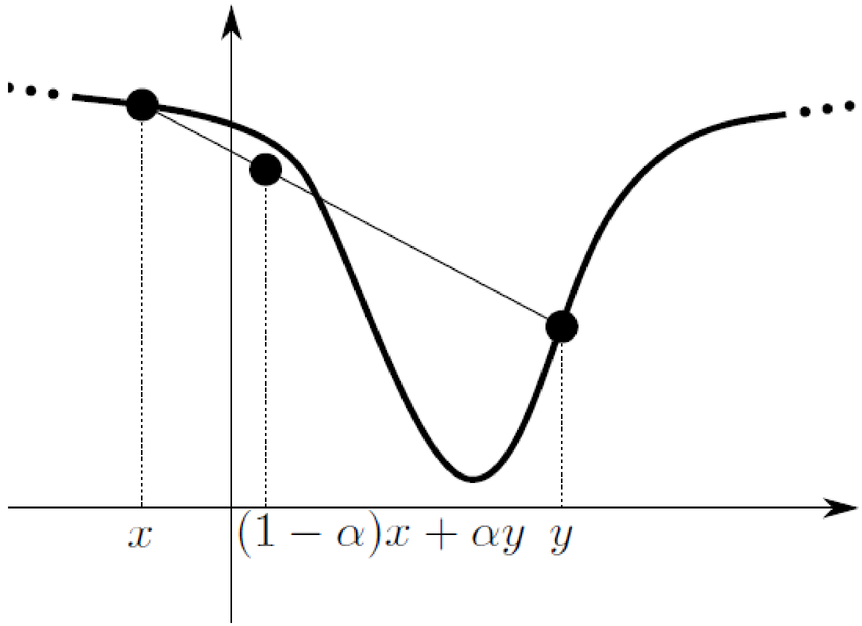
- 仍无法求解!
- 数值求解
  - 从某初值开始，逐步找其附近的极小值



凸函数的驻点就是最小值



(a) Convex



(b) Quasiconvex

Numerical Algorithms, Solomon

优化问题的类型

- Constrained / Unconstrained
- Linear / Nonlinear
- Global / Local
- Convex / Nonconvex
- Continuous / Discrete
- Stochastic / Deterministic
- Single objective / Multiple objectives

$$\text{minimize } \left(E_1(x), E_2(x), .., E_k(x)\right)$$

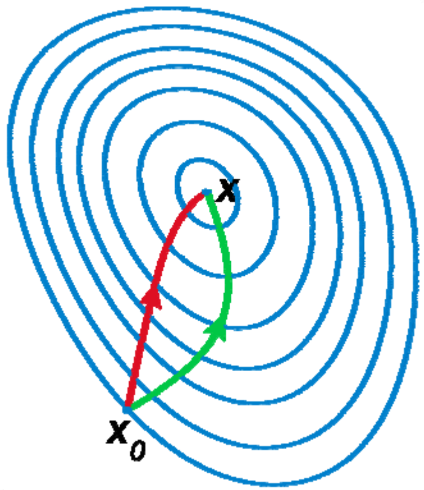
$$E = \lambda_1 E_1 + \lambda_2 E_2 + \cdots + \lambda_k E_k$$

无约束的优化问题

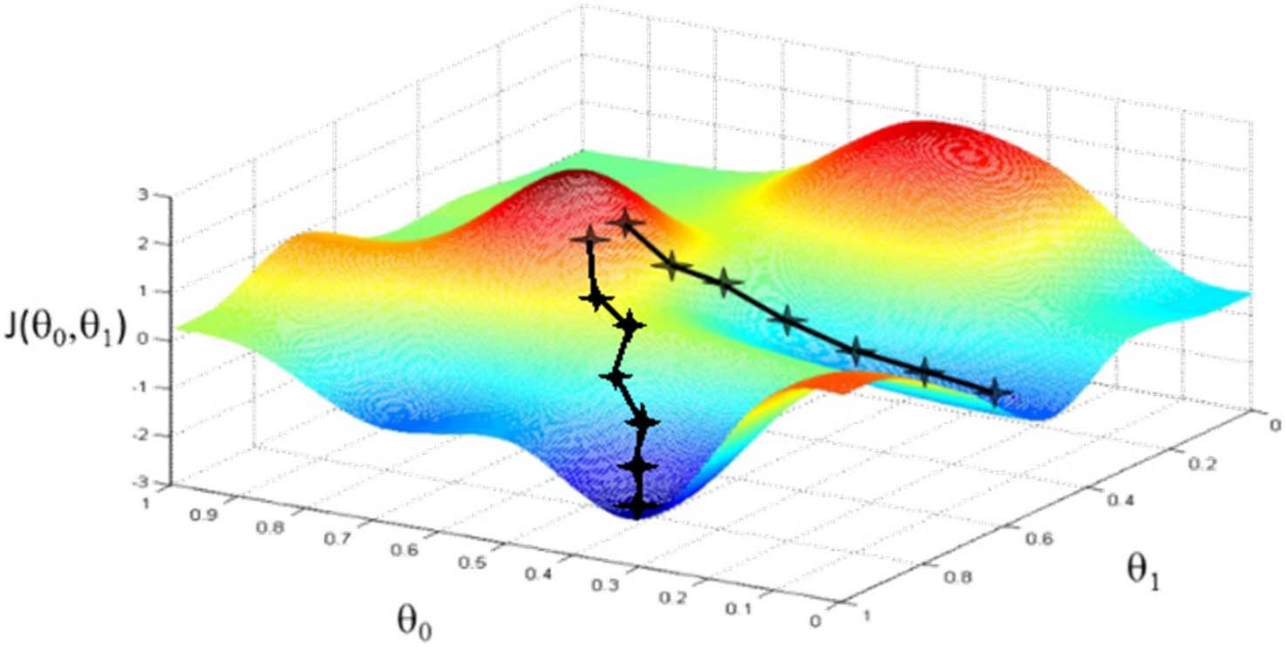
Unconstrained Optimization

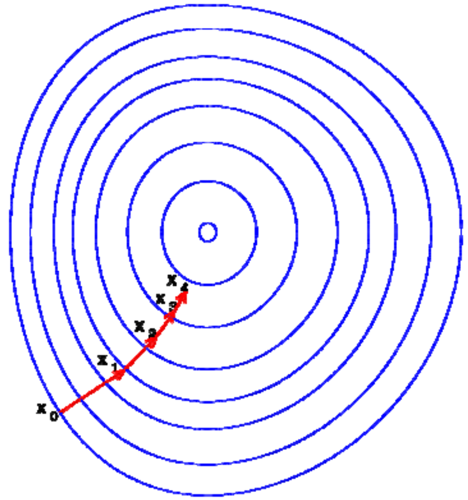
$$\min_x f(x)$$

- Gradient descent
- Newton
- Quasi-Newton
- Coordinate descent



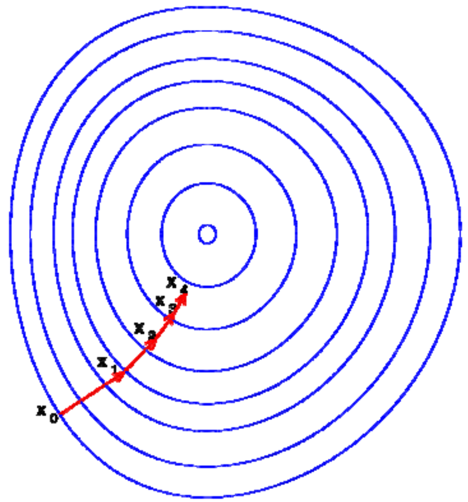
梯度下降法
(Gradient descent)

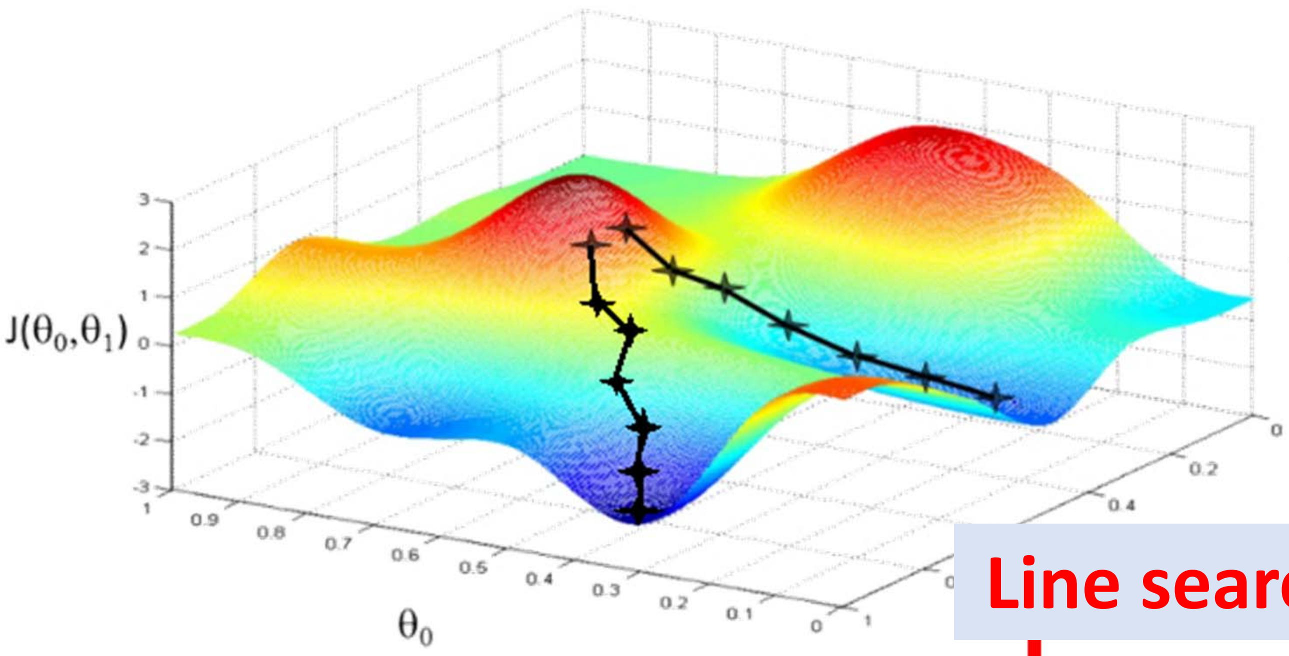




$$x_{k+1} = x_k - \alpha_k \nabla f(x_k)$$

梯度下降法 (Gradient descent)





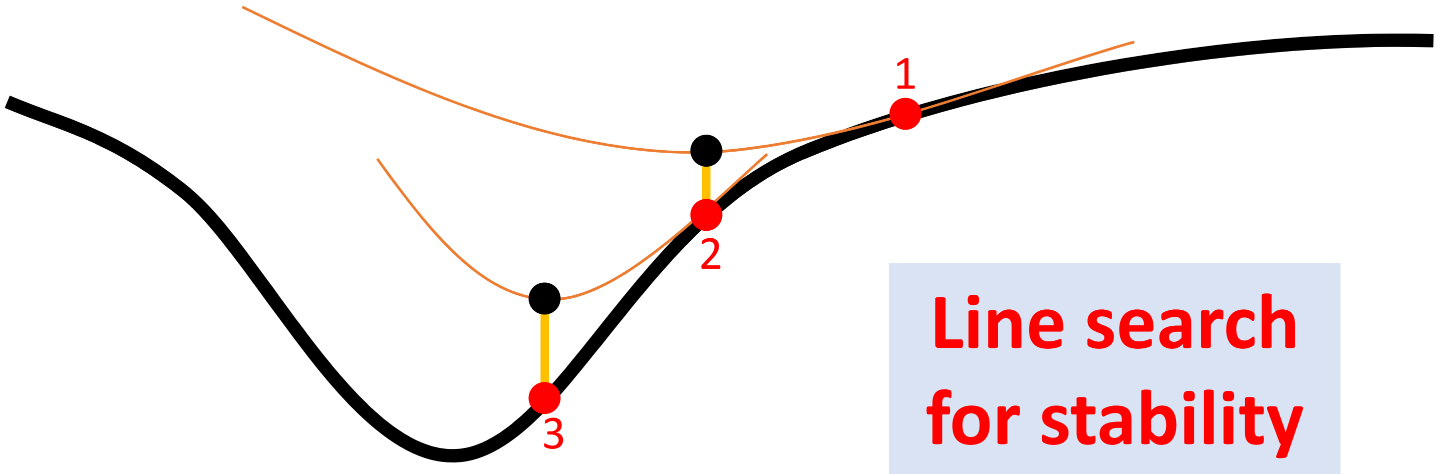
Line search

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

Gradient descent

牛顿法 (Newton’s method)

$$x_{k+1} = x_k - \boxed{[Hf(x_k)]^{-1}} \nabla f(x_k)$$



拟牛顿法 (Quasi-Newton)

$$x_{k+1} = x_k - \underbrace{M_k^{-1}}_{\text{Hessian approximation}} \nabla f(x_k)$$

- **Estimate the Hessian based on previous gradients**
- **Recursively inverse Hessian**

- BFGS (Broyden–Fletcher–Goldfarb–Shanno algorithm)
- L-BFGS

坐标下降法 (Coordinate descent)

Obj: $\text{minimize}_{x,y}\ E(x,y)$

- Alternating variables

Repeat

1. $y_{k+1} = \min_{y} E(x_k, y)$
2. $x_{k+1} = \min_{x} E(x, y_{k+1})$

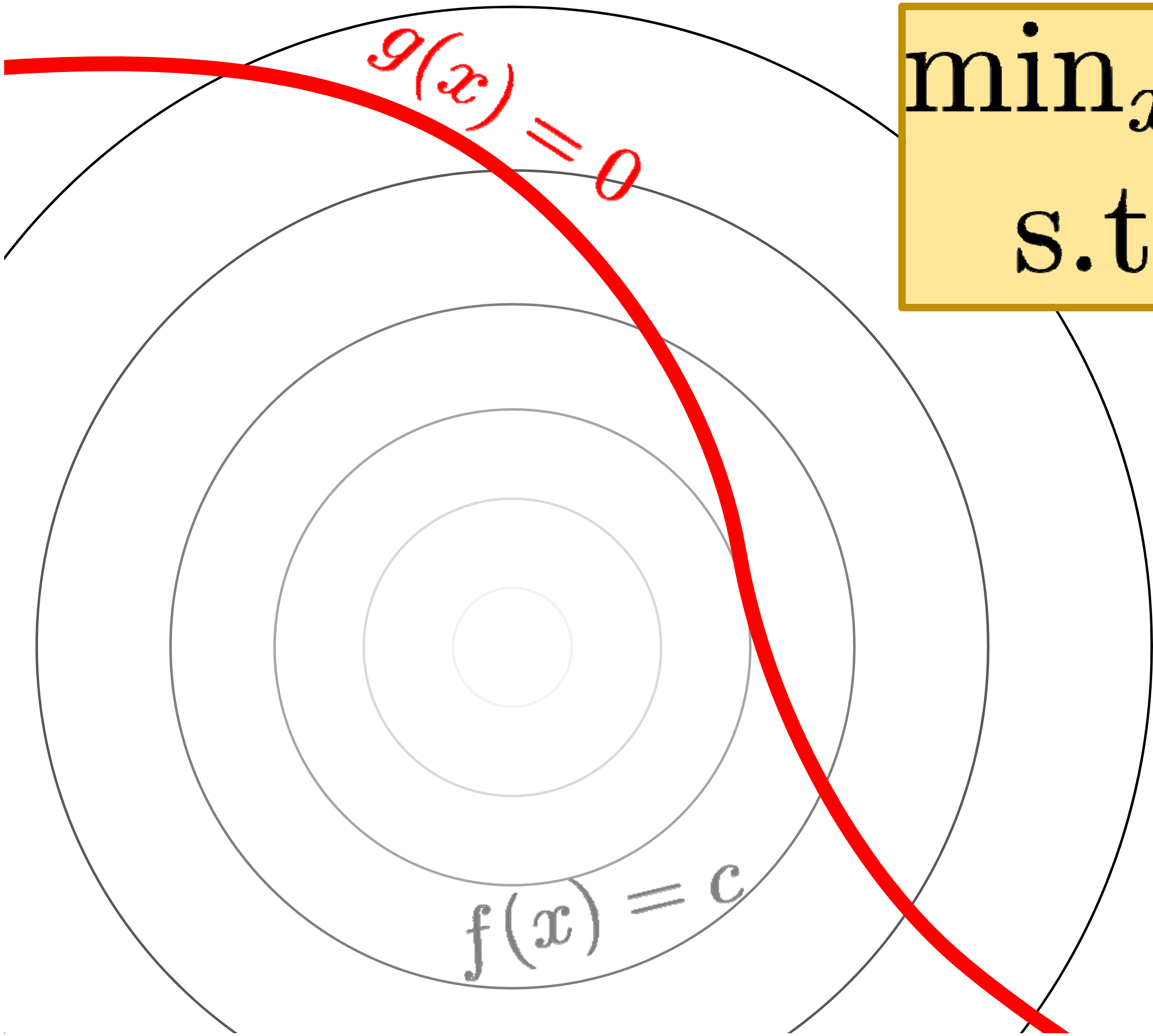
Software

- **Matlab**: `fminunc` or `minfunc`
- **C++**: `libLBFGS`, `dlib`, others

Typically provide functions for function and gradient (and optionally, Hessian).

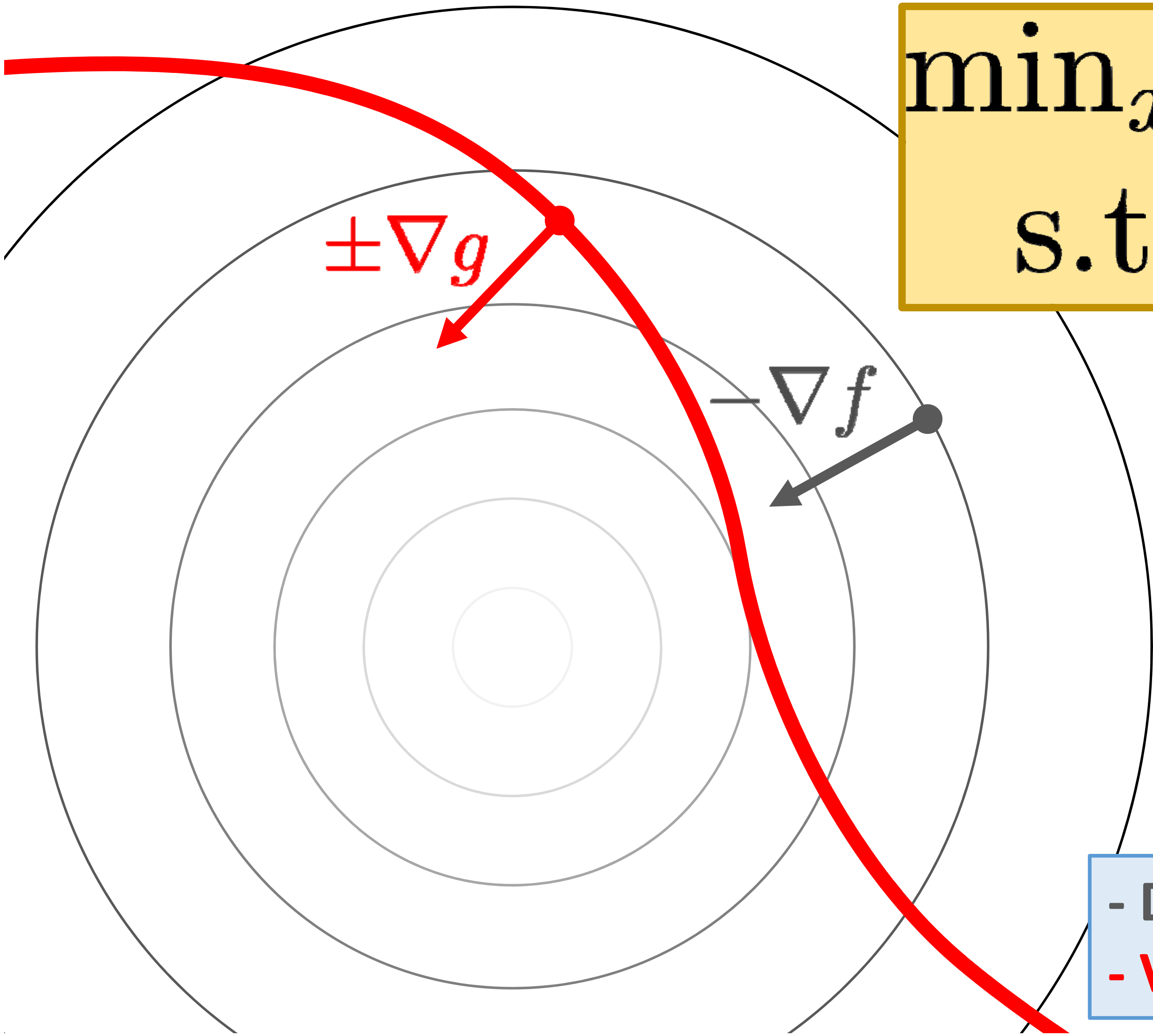
等式约束的优化问题

Lagrange Multipliers: Idea



$$
\begin{aligned}
\min_x \quad & f(x) \\
\text{s.t.} \quad & g(x) = 0
\end{aligned}
$$

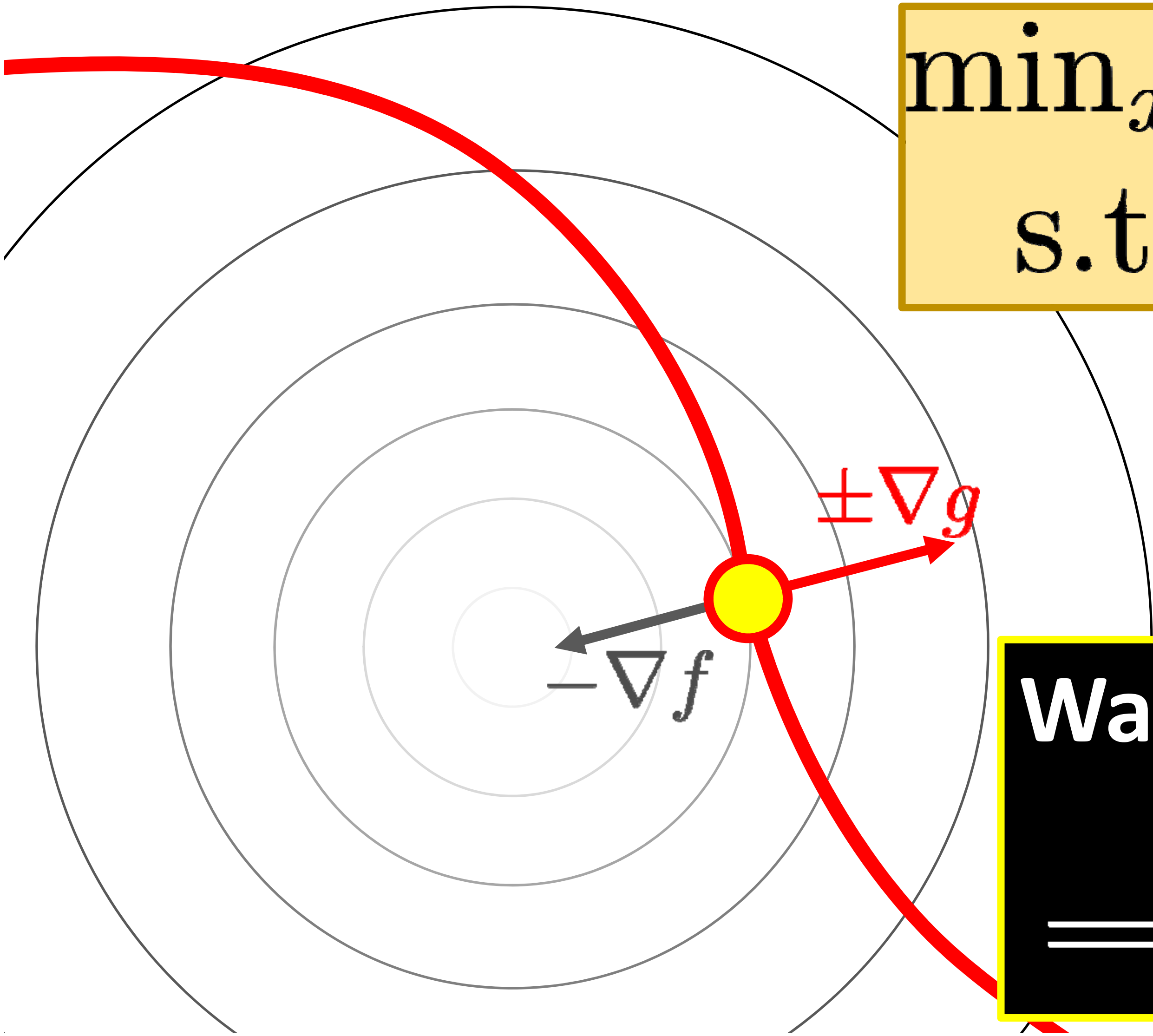
Lagrange Multipliers: Idea



$$\begin{aligned} \min_x \quad & f(x) \\ \text{s.t.} \quad & g(x) = 0 \end{aligned}$$

- Decrease f: $-\nabla f$
- **Violate constraint: $\pm\nabla g$**

Lagrange Multipliers: Idea



$$\begin{aligned} \min_x \quad & f(x) \\ \text{s.t.} \quad & g(x) = 0 \end{aligned}$$

Want:

$$\nabla f \parallel \nabla g$$

$$\implies \nabla f = \lambda \nabla g$$

Use of Lagrange Multipliers

Turns constrained optimization into

unconstrained root-finding.

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

$$g(x) = 0$$

Many Options

- **Reparameterization**

  Eliminate constraints to reduce to unconstrained case

- **Newton's method**

  Approximation: quadratic function with linear constraint

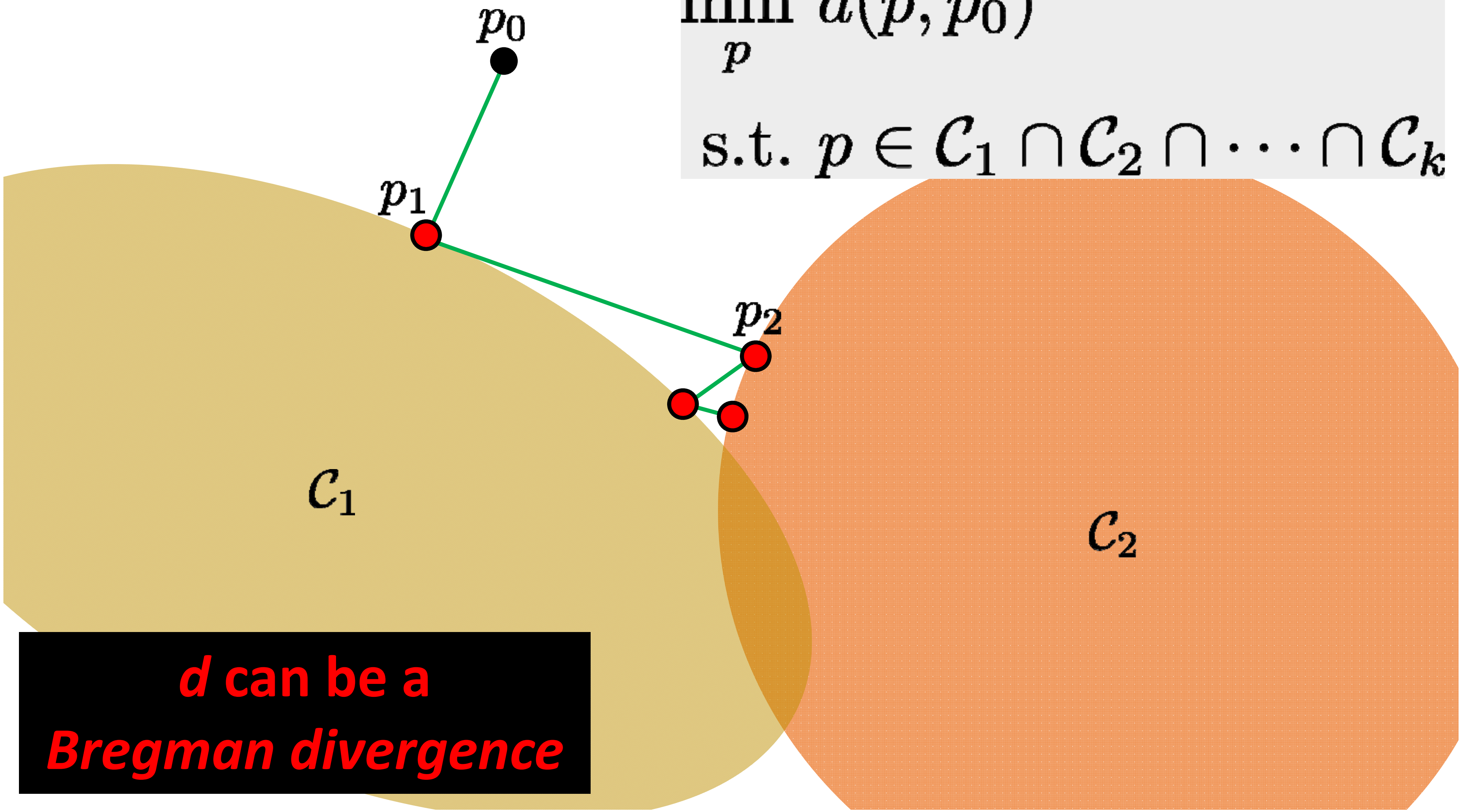
- **Penalty method**

  Augment objective with barrier term, e.g. $\boldsymbol{f(x) + \rho|g(x)|}$

Alternating Projection

$$\min_{p} \; d(p, p_0)$$

$$\text{s.t. } p \in \mathcal{C}_1 \cap \mathcal{C}_2 \cap \cdots \cap \mathcal{C}_k$$



d* can be a *Bregman divergence

Augmented Lagrangians

$$\begin{aligned} \min_x \quad & f(x) \\ \text{s.t.} \quad & g(x) = 0 \end{aligned}$$

$$\downarrow$$

$$\begin{aligned} \min_x \quad & f(x) + \frac{\rho}{2}\|g(x)\|_2^2 \\ \text{s.t.} \quad & g(x) = 0 \end{aligned}$$

Does nothing when constraint is satisfied

Add constraint to objective

Alternating Direction Method of Multipliers (ADMM)

$$\begin{array}{rl} \min_{x,z} & f(x) + g(z) \\ \text{s.t.} & Ax + Bz = c \end{array}$$

$$\Lambda_\rho(x, z; \lambda) = f(x) + g(z) + \lambda^\top (Ax + Bz - c) + \frac{\rho}{2}\|Ax + Bz - c\|_2^2$$

$$x \leftarrow \arg\min_x \Lambda_\rho(x, z, \lambda)$$

$$z \leftarrow \arg\min_z \Lambda_\rho(x, z, \lambda)$$

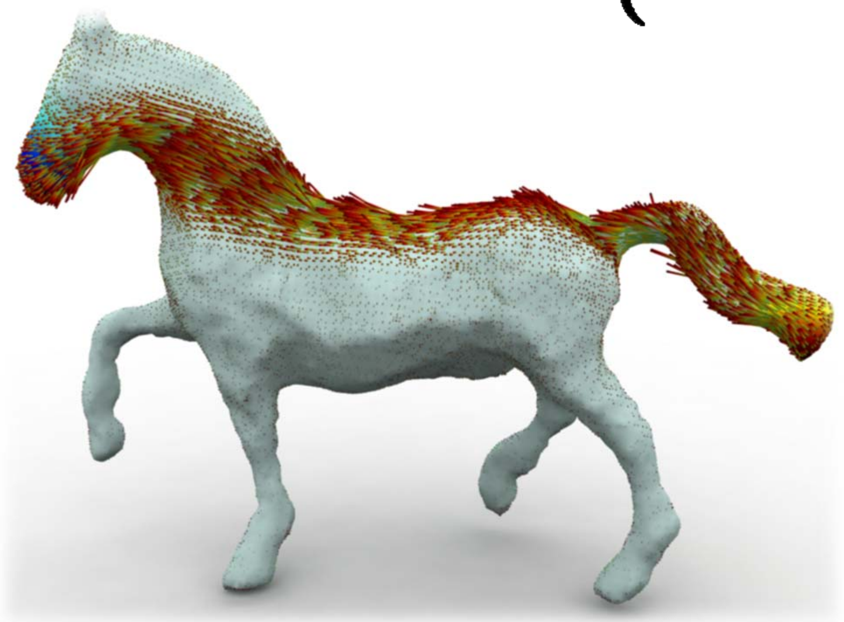
$$\lambda \leftarrow \lambda + \rho(Ax + Bz - c)$$

https://web.stanford.edu/~boyd/papers/pdf/admm_slides.pdf

The Art of ADMM "Splitting"

$$\left\{\begin{array}{rl}\min_J & \sum_i \|J_i\|_2 \\ \text{s.t.} & MJ = b\end{array}\right\} \longrightarrow \left\{\begin{array}{rl}\min_{J,\bar{J}} & \sum_i \left(\|J_i\|_2 + \frac{\rho}{2}\|J_i - \bar{J}_i\|_2^2\right) \\ \text{s.t.} & M\bar{J} = b \\ & J = \bar{J}\end{array}\right\}$$

Augmented part



Takes some practice!

Solomon et al. "Earth Mover's Distances on Discrete Surfaces." SIGGRAPH 2014.

Want two *easy* subproblems

不等式约束的优化问题

一般形式

objective function

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

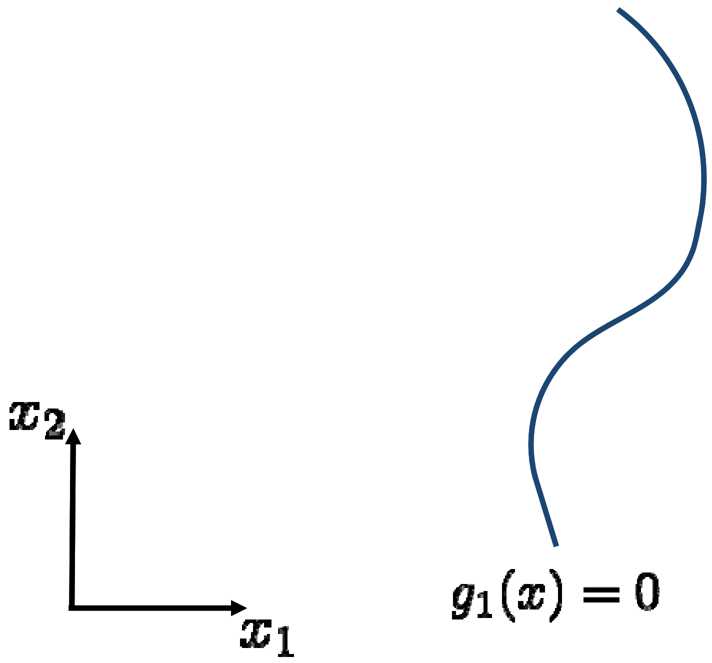
$$\begin{array}{lll} \text{minimize} & f(x) & \\ \text{subject to} & g_i(x) \le 0 & i = 1 \ldots m \end{array}$$

constraint functions

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

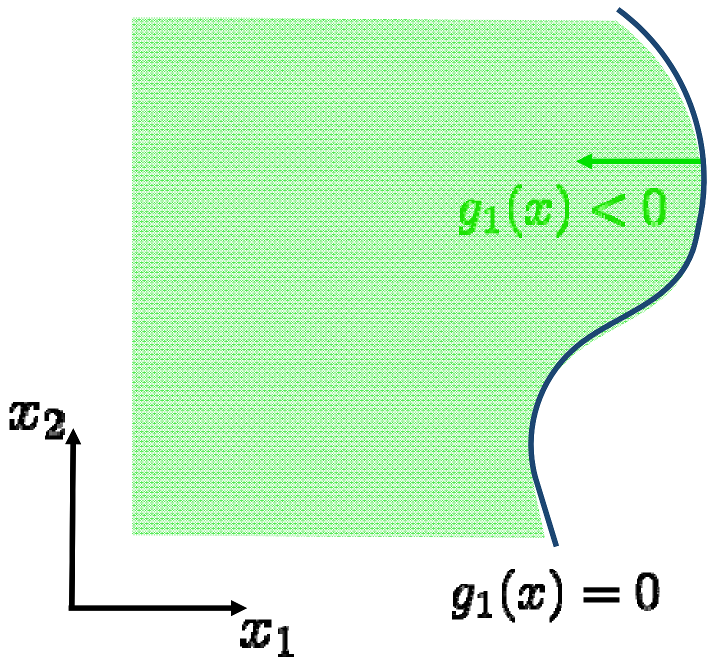
几何解释

$$
\begin{array}{lll}
\text{minimize} & f(x) & \\
\text{subject to} & g_i(x) \leq 0 & i = 1 \ldots m
\end{array}
$$



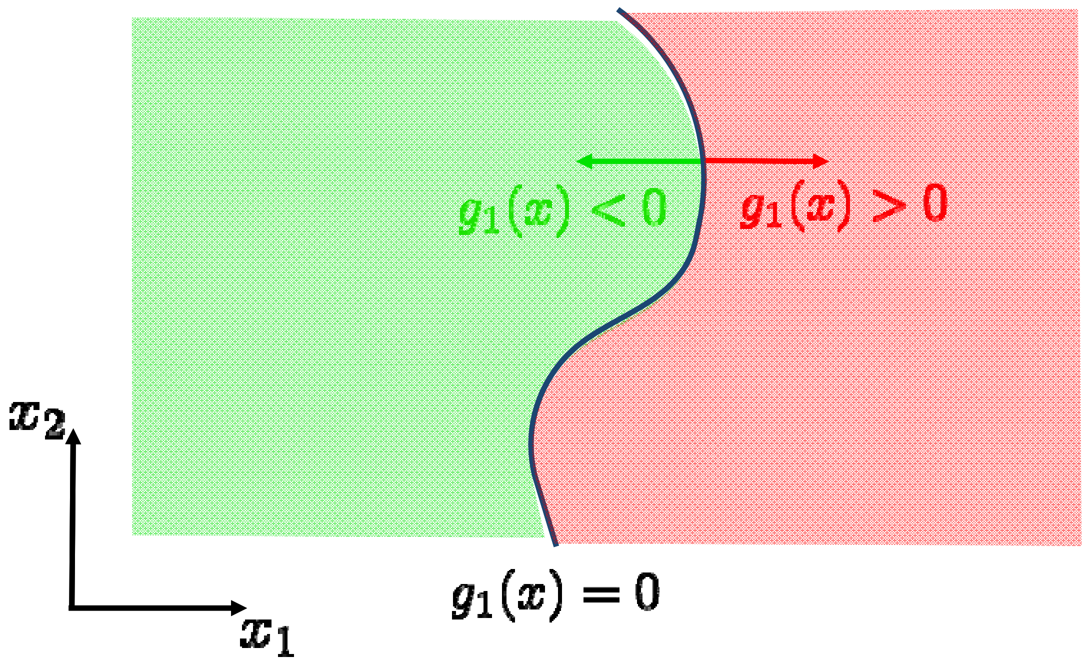
几何解释

$$\begin{array}{lll} \text{minimize} & f(x) & \\ \text{subject to} & g_i(x) \leq 0 & i = 1 \ldots m \end{array}$$



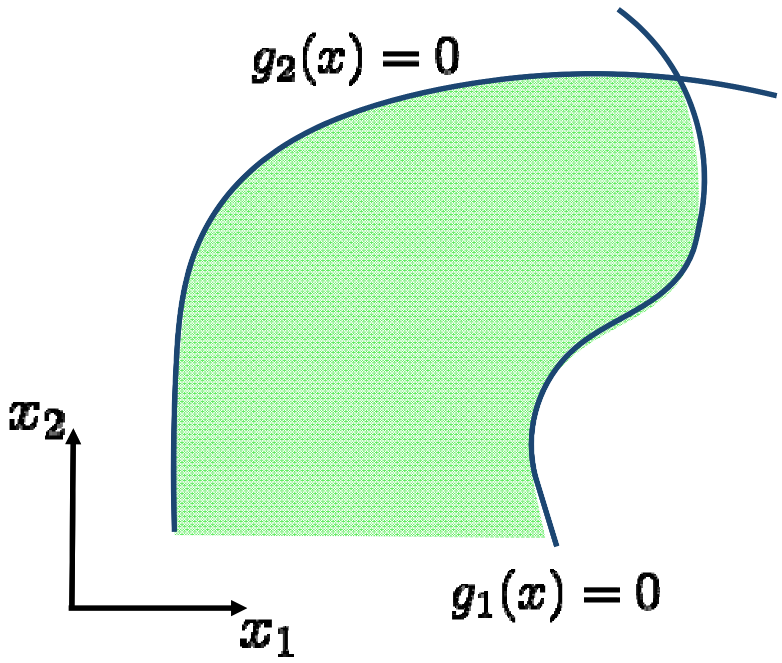
几何解释

$$\begin{array}{lll} \text{minimize} & f(x) & \\ \text{subject to} & g_i(x) \leq 0 & i = 1 \ldots m \end{array}$$



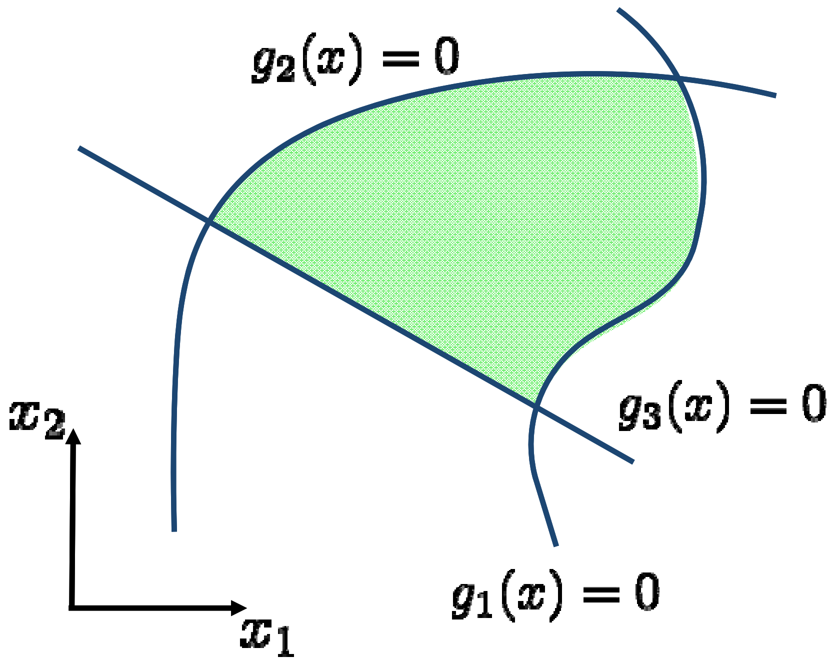
几何解释

$$\begin{array}{llll} \text{minimize} & f(x) & & \\ \text{subject to} & g_i(x) & \leq 0 & i = 1 \ldots m \end{array}$$



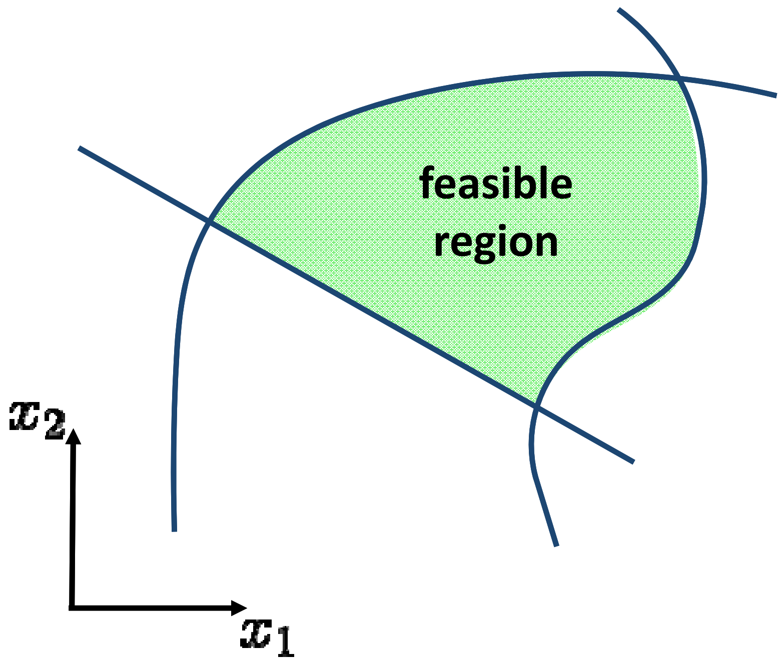
几何解释

$$
\begin{array}{lll}
\text{minimize} & f(x) & \\
\text{subject to} & g_i(x) \leq 0 & i = 1 \ldots m
\end{array}
$$



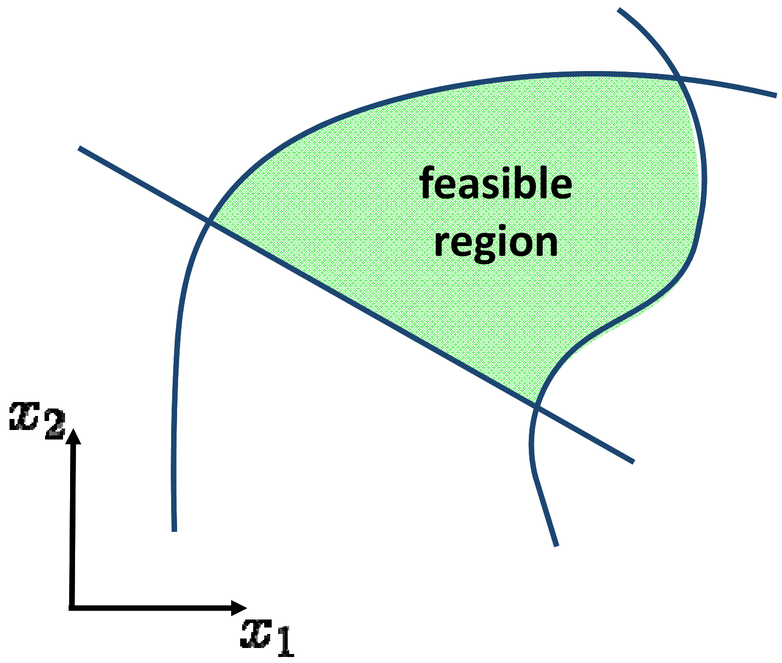
几何解释

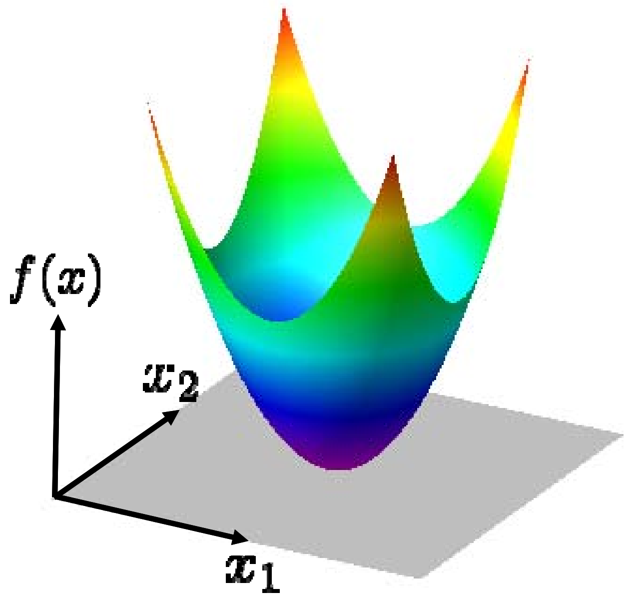
$$
\begin{array}{lll}
\text{minimize} & f(x) & \\
\text{subject to} & g_i(x) \le 0 & i = 1 \ldots m
\end{array}
$$



几何解释

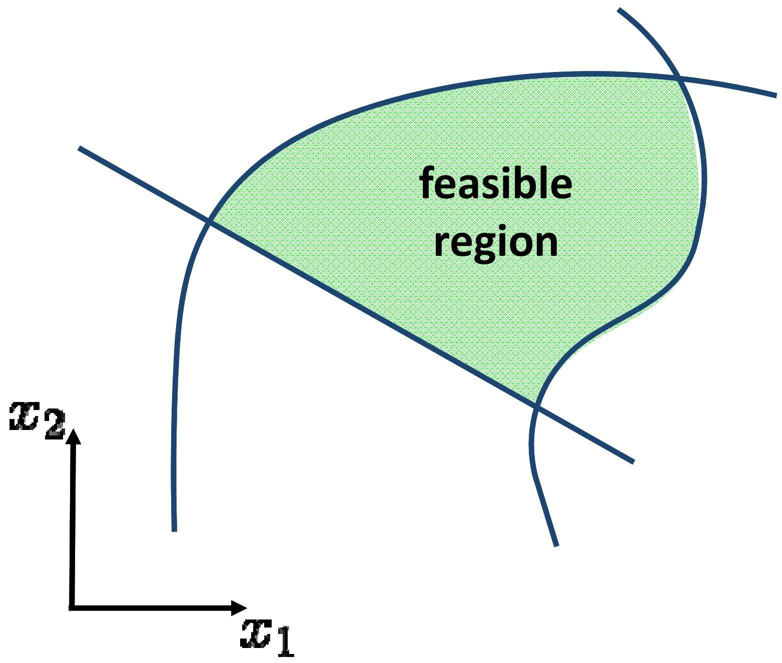
$$\begin{aligned} &\text{minimize} && f(x) \\ &\text{subject to} && g_i(x) \le 0 \qquad i = 1 \ldots m \end{aligned}$$

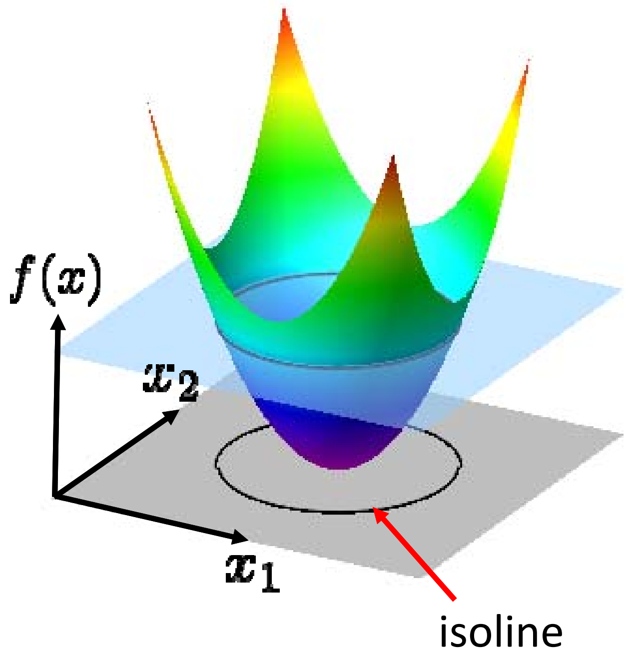




几何解释

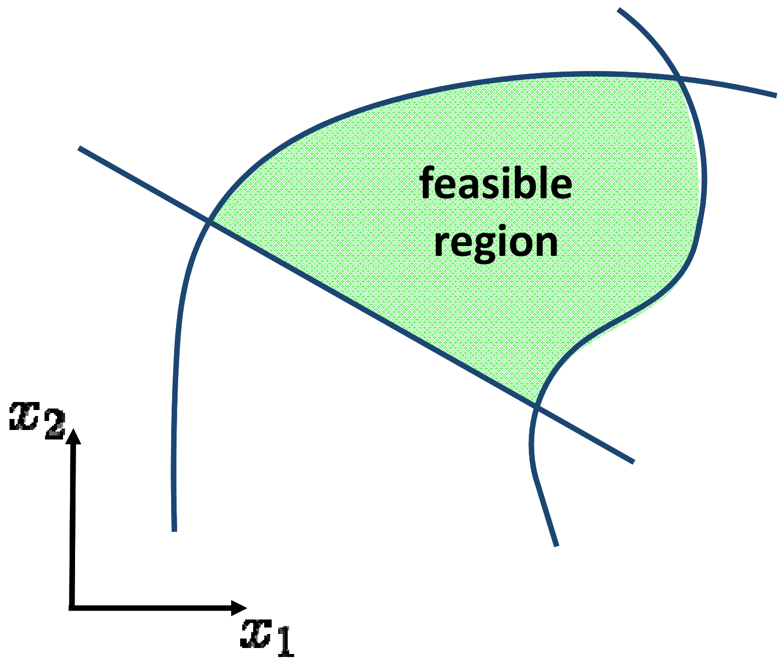
$$\begin{array}{lll} \text{minimize} & f(x) & \\ \text{subject to} & g_i(x) \le 0 & i = 1 \ldots m \end{array}$$

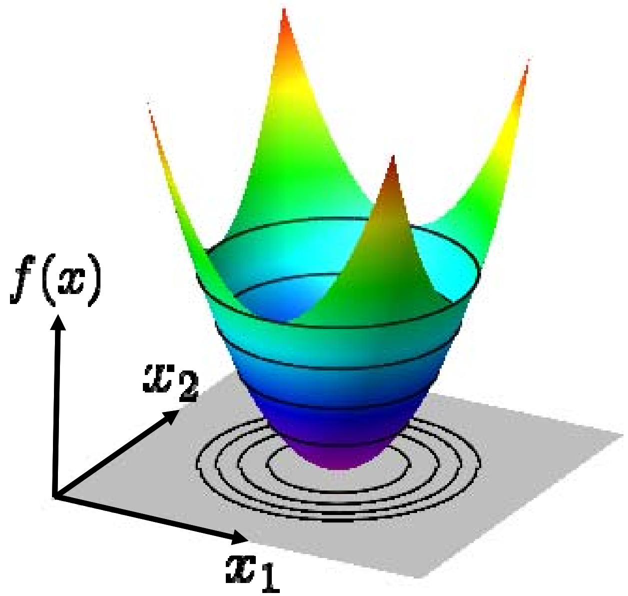




几何解释

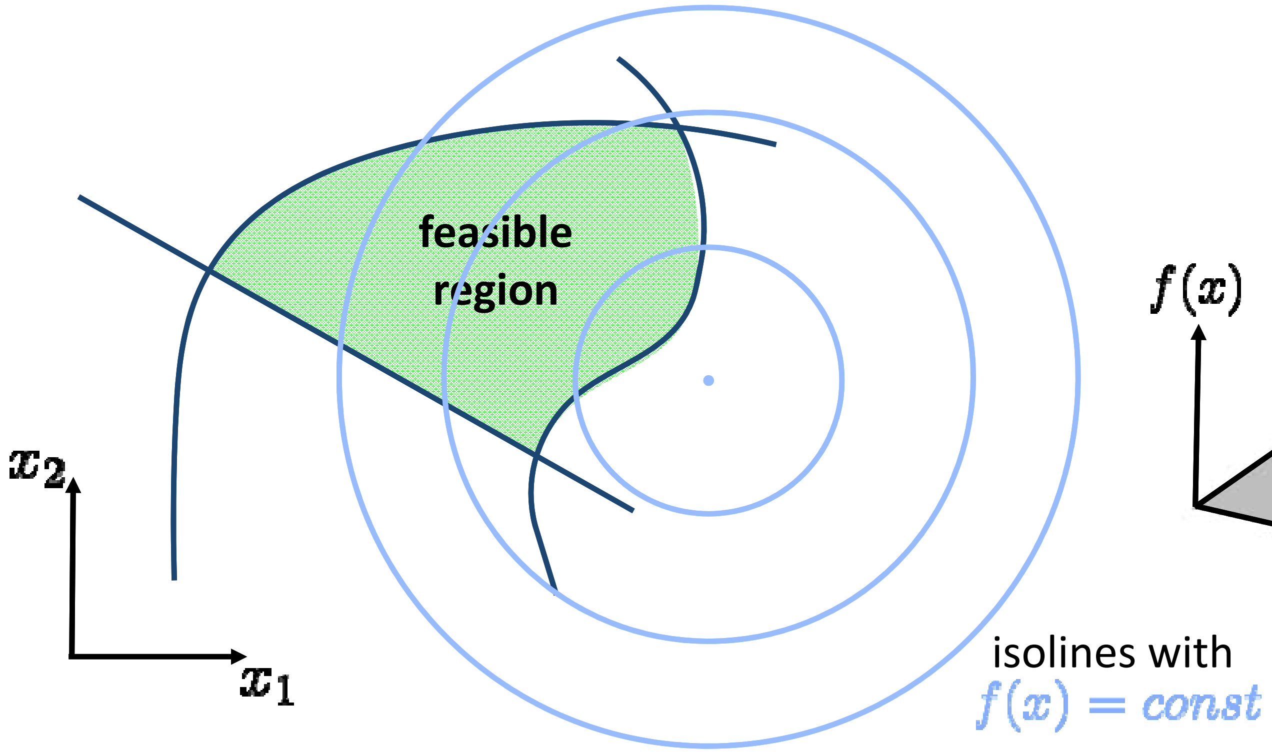
$$\begin{aligned} &\text{minimize} && f(x) \\ &\text{subject to} && g_i(x) \le 0 \qquad i = 1 \dots m \end{aligned}$$

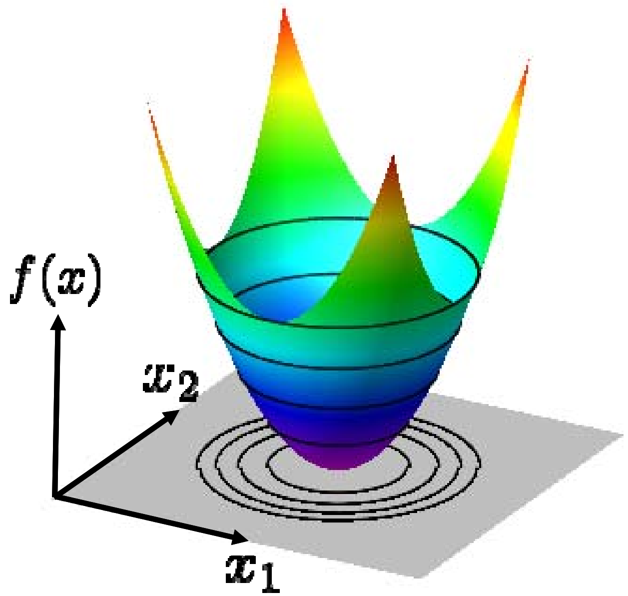




几何解释

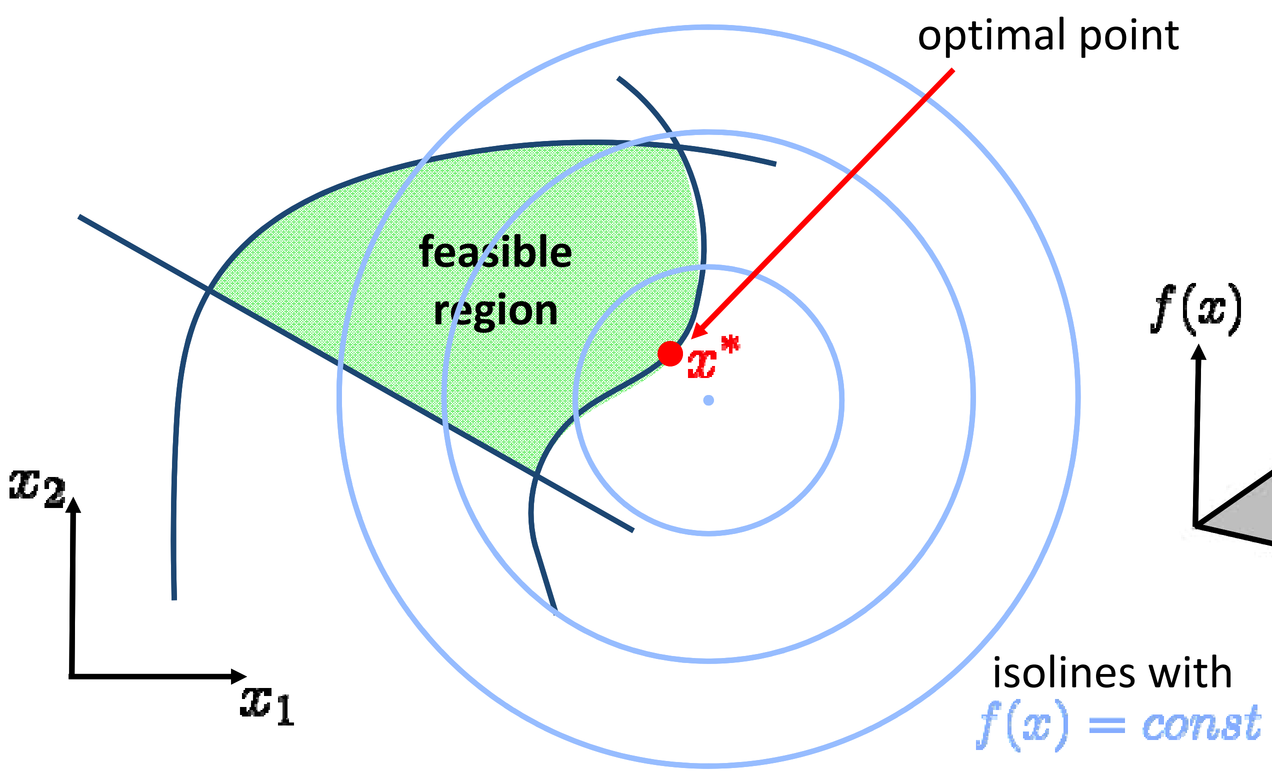
$$\begin{array}{lll} \text{minimize} & f(x) & \\ \text{subject to} & g_i(x) \le 0 & i = 1 \dots m \end{array}$$

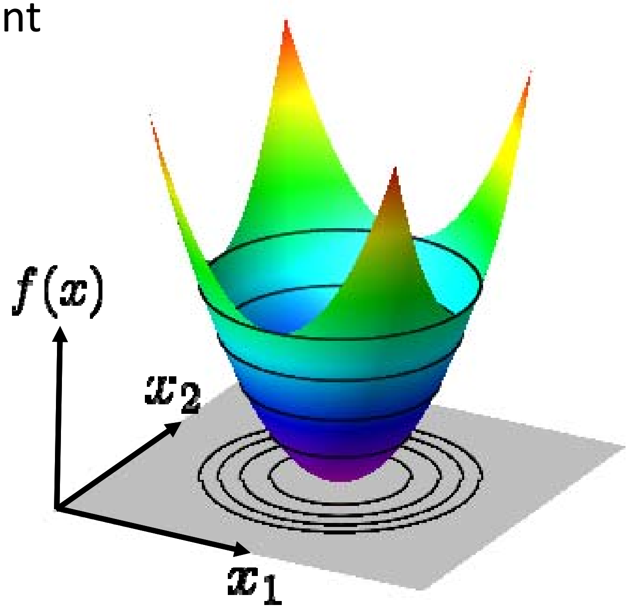




几何解释

$$\begin{array}{lll} \text{minimize} & f(x) & \\ \text{subject to} & g_i(x) \leq 0 & i = 1 \ldots m \end{array}$$





First-Order Optimality Conditions

- **Necessary condition for minimum of**

$$\begin{aligned} \text{minimize} \quad & f(x) \\ \text{subject to} \quad & g_i(x) \leq 0 \qquad i = 1 \ldots m \end{aligned}$$

- **Langrangian**: $L(x, \lambda) = f(x) + \sum_{i=1}^{m} \lambda_i g_i(x)$

- **Karush-Kuhn-Tucker (KKT)**

  conditions for **minimum** x^*

  1. Stationarity: $\nabla f(x^*) + \sum_{i=1}^{m} \lambda_i \nabla g_i(x^*) = 0$
  2. Primal feasibility: $g_i(x^*) \leq 0$
  3. Dual feasibility: $\lambda_i \geq 0$
  4. Complementary slackness: $\lambda_i g_i(x^*) = 0$

without constraints just

$\nabla f(x) = 0$

First-Order Optimality Conditions

- **Necessary condition for minimum of**

$$\begin{array}{lll} \text{minimize} & f(x) & \\ \text{subject to} & g_i(x) \leq 0 & i = 1 \ldots m \end{array}$$

- **Langrangian**: $L(x, \lambda) = f(x) + \sum_{i=1}^{m} \lambda_i g_i(x)$

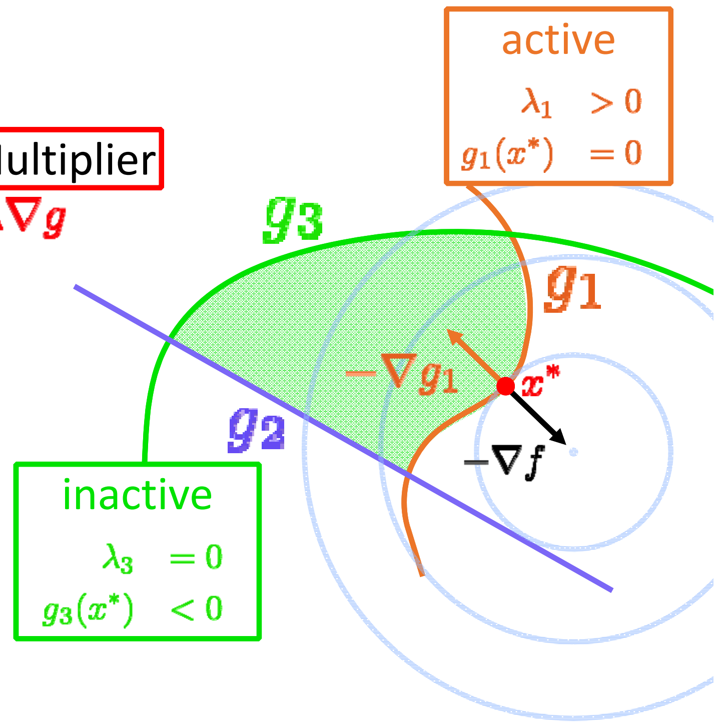
- **Karush-Kuhn-Tucker (KKT)**

  conditions for **minimum** x^*

  1. Stationarity: $\nabla f(x^*) + \sum_{i=1}^{m} \lambda_i \nabla g_i(x^*) = 0$
  2. Primal feasibility: $g_i(x^*) \leq 0$
  3. Dual feasibility: $\lambda_i \geq 0$
  4. Complementary slackness: $\lambda_i g_i(x^*) = 0$

Recall Lagrange Multiplier

$\nabla f = -\lambda \nabla g$



优化方法

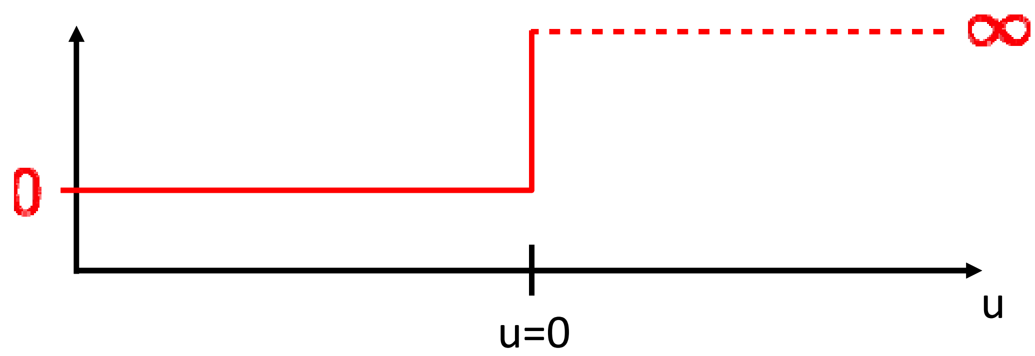
- Active set method

Repeated update active set

1. active set define equality constraints
2. *compute* the Lagrange multipliers of the active set
3. *remove* constraints with **negative** Lagrange multipliers
4. *search* for infeasible constraints

- Barrier method (Interior Point Method)

$$I_-(u) = \begin{cases} 0 & \text{if } u \le 0 \\ \infty & \text{if } u > 0 \end{cases}$$



constrained

$$\begin{array}{lll} \text{minimize} & f(x) & \\ \text{subject to} & g_i(x) \le 0 & i = 1 \dots m \end{array}$$

$\Leftrightarrow$

equivalent

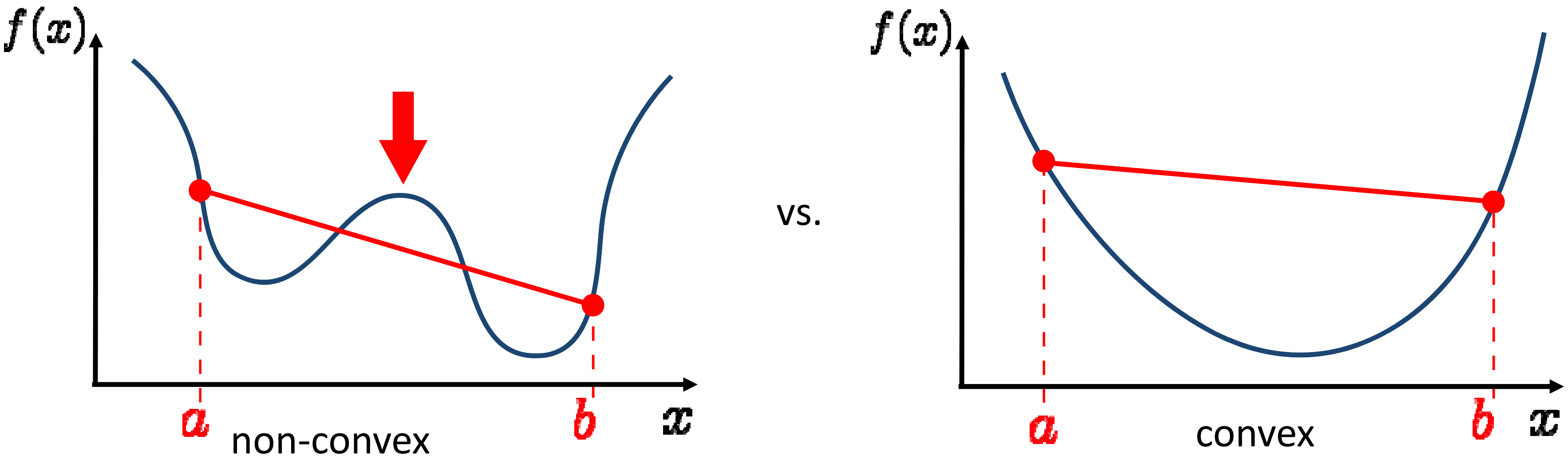
unconstrained

$$\text{minimize} \quad f(x) + \sum_{i=1}^{m} I_-(g_i(x))$$

Convex Optimization

凸函数能保证找到全局最小值

- Searching globally optimal solutions usually requires **convexity**!
- f convex if: $f((1-t)a+tb) \leq (1-t)f(a)+tf(b) \quad t \in [0,1]$



凸优化问题

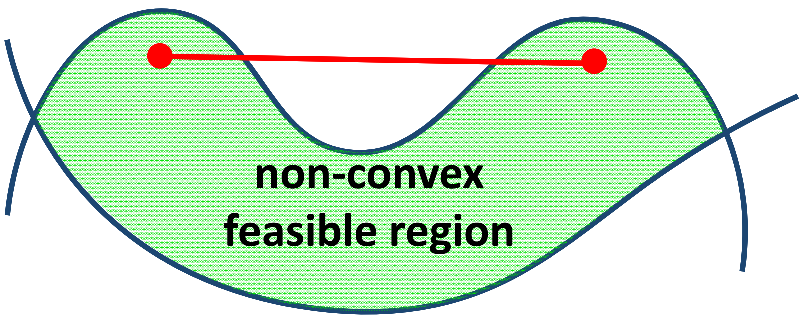
$$\begin{array}{lll} \text{minimize} & f(x) & \\ \text{subject to} & g_i(x) \leq 0 & i = 1 \ldots m \end{array}$$

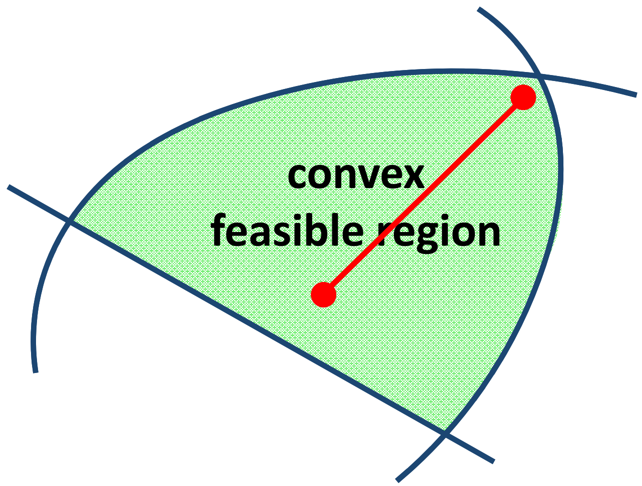
is **convex optimization problem** if

$f(x)$ and all $g_i(x)$ are convex functions

<u>consequences</u>

- feasible region is convex set

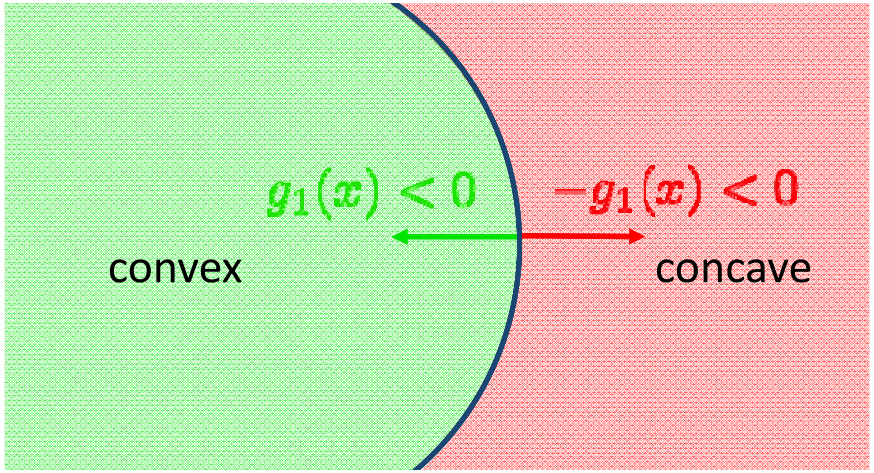




凸优化问题

$$\begin{array}{lll} \text{minimize} & f(x) & \\ \text{subject to} & g_i(x) \le 0 & i = 1 \dots m \end{array}$$

is **convex optimization problem** if $f(x)$ and all $g_i(x)$ are convex functions



<u>consequences</u>

- feasible region is convex set
- equality constraints can only be affine, i.e. $g_i(x) = a^T x + b$ since

$$g_i(x) = 0 \iff \begin{cases} g_i(x) & \le 0 \\ -g_i(x) & \le 0 \end{cases}$$

凸优化的主要方法

- Linear Programming

$$\underset{x\in\mathbb{R}^n}{\text{minimize}}\, q^T x \qquad \text{subject to} \quad \begin{array}{l} Ax = a \\ Bx < b \end{array}$$

- Quadratic Programming

$$\underset{x\in\mathbb{R}^n}{\text{minimize}}\, \frac{1}{2}x^T Q x + q^T x \qquad \text{subject to} \quad \begin{array}{l} Ax = a \\ Bx < b \end{array}$$

- Conic Programming

$$\sqrt{\sum_{i=1}^{n} x_i^2} \le x_0$$

- Semidefinite Programming (SDP)

$$\underset{A\in\mathbb{R}^{n\times n}}{\text{minimize}}\, E(A) \qquad \text{subject to} \quad A \ge 0$$

A is SPD
$\lambda_i(A) \ge 0, \forall i$

其他优化问题

Nonlinear Least Squares

Obj: minimize $\sum_i e_i^2(x)$

- Gauss-Newton

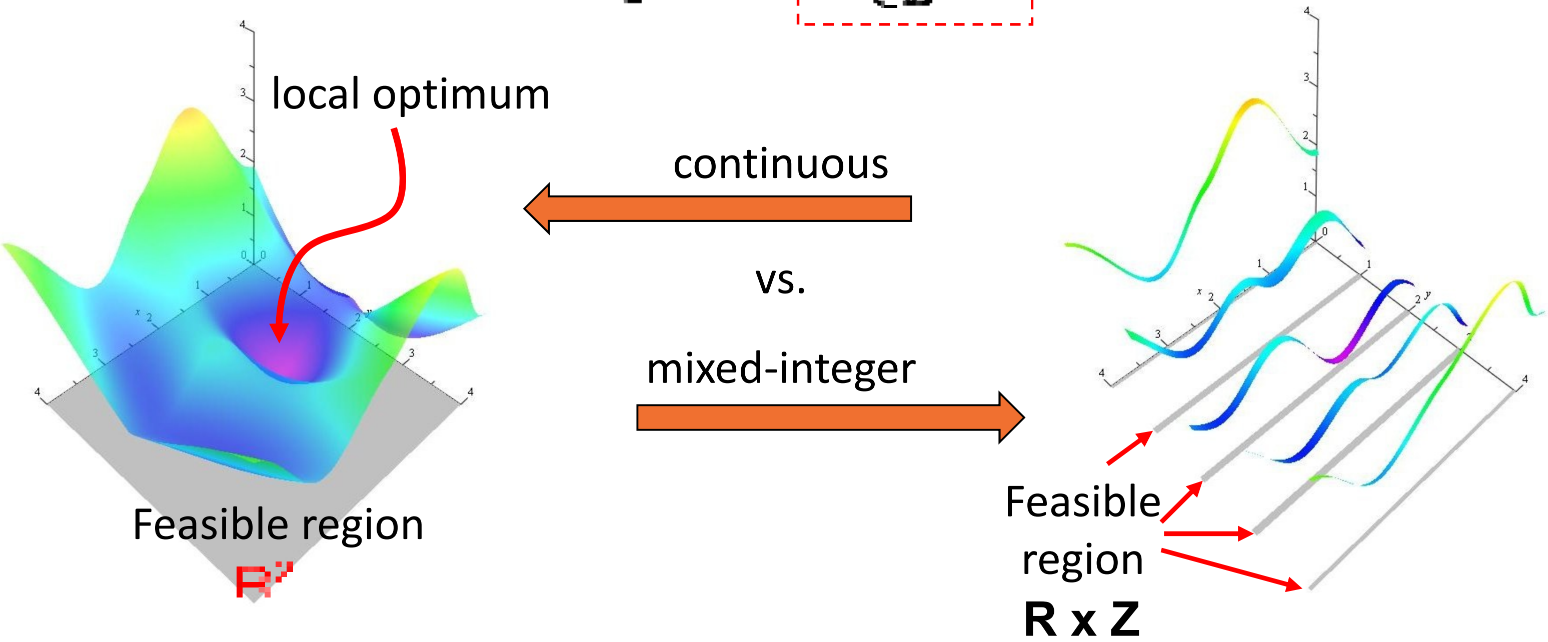
- Levenberg-Marquardt

$$\nabla^2 e_i^2 \approx 2(\nabla e_i)^T \nabla e_i$$

$$\nabla^2 \approx J^T J$$

Mixed-Integer Optimization

$$f(\underbrace{x_1 \ldots x_n}_{\in \mathbb{R}^n}, \underbrace{y_1 \ldots y_d}_{\in \mathbb{Z}^d})$$

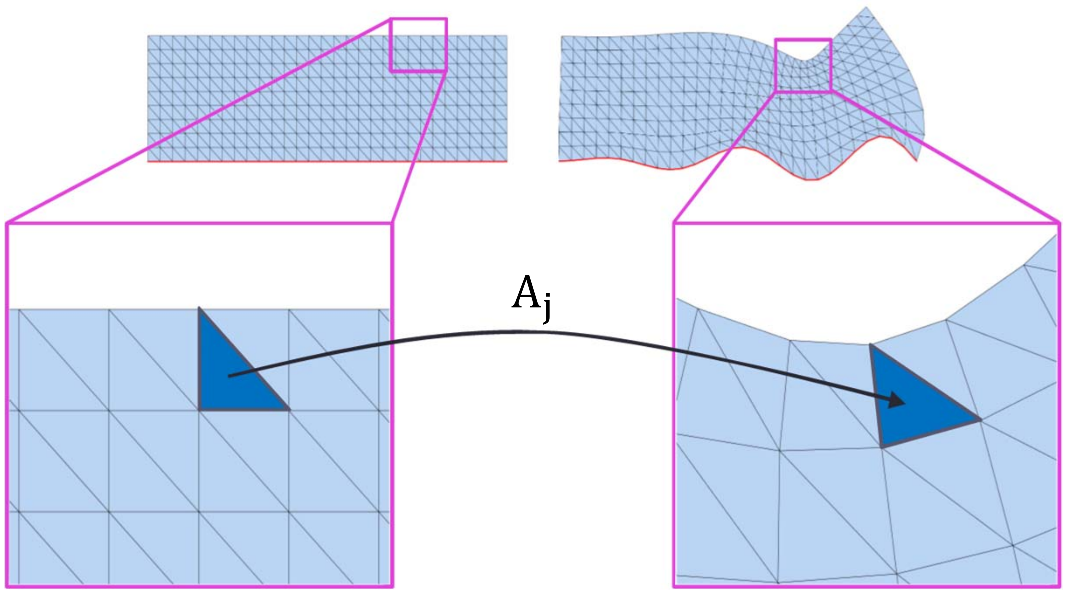


几何处理中的优化问题

- 具有特殊的几何结构，往往能有特殊的优化方法
  - 比如：见“曲面参数化”和“几何映射”两节课

$$\operatorname{argmin} \sum_j f(A_j)$$

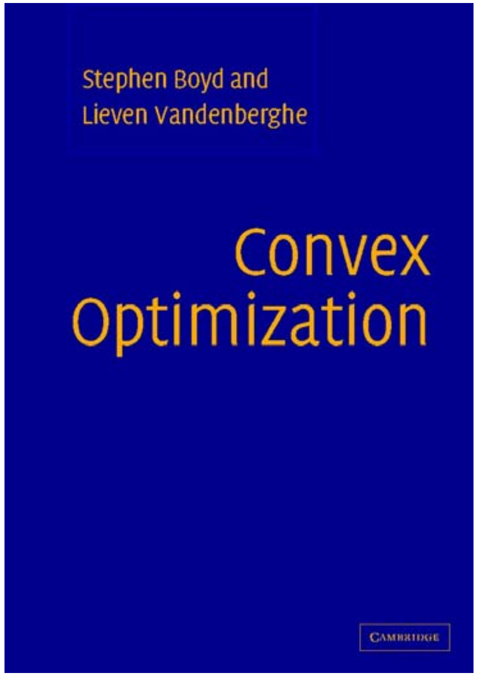
Separable



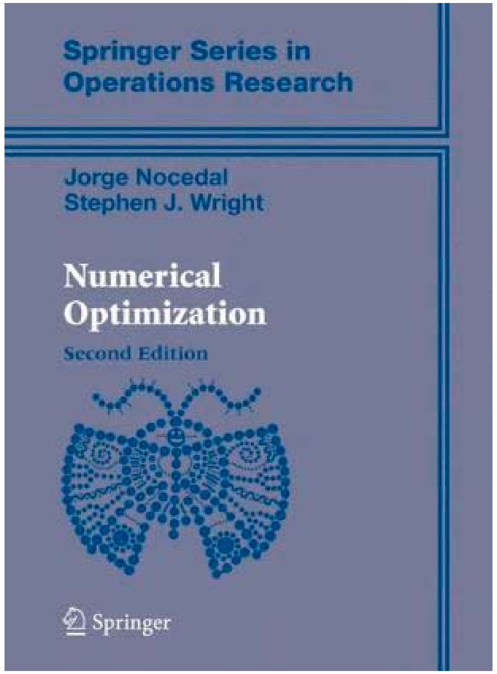
优化相关的软件

- **Eigen** — linear algebra
- **IPOPT** — fast opensource C++ interior point method
- **Mosek** — commercial (convex) optimization in C, Java, Python…
- **Gurobi** — commercial mixed-integer optimization
- **CPLEX** — commercial mixed-integer optimization
- **Matlab** — many algorithms, good for prototyping
- **CVX** — prototyping for convex optimization
- **CoMISo** — unified interface to above algorithms

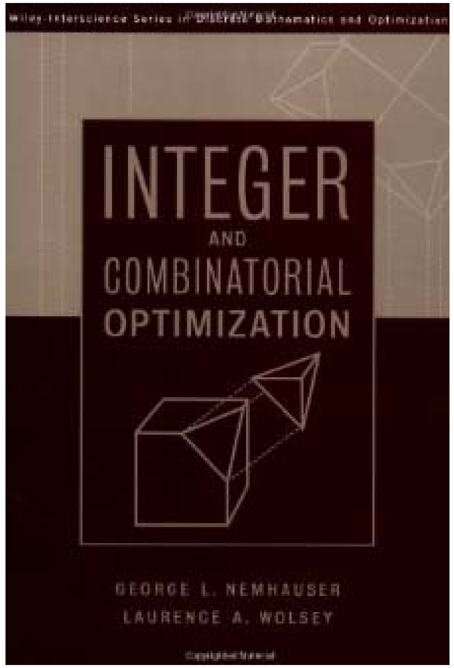
参考书目 Optimization is a **huge** field!



S. Boyd and L. Vandenberghe
Convex Optimization
Cambridge University Press, 2004.



J. Nocedal and S. J. Wright
Numerical Optimization
Springer, 2006.



G. L. Nemhauser and L. A. Wolsey
Integer and Combinatorial Optimization
John Wiley & Sons, 1999.

Get PDF online:
http://stanford.edu/~boyd/cvxbook/

中国科学技术大学
University of Science and Technology of China

谢 谢 ！