上海交通大学
SHANGHAI JIAO TONG UNIVERSITY

上海交通大学自动化系
Department of Automation
Shanghai Jiao Tong University

Lecture-1
《运筹与优化的基本概念》

Main Contents

知识点		学时
运筹与优化的基础知识		2
线性规划基础理论 14（Linear Programming）	图解法与基本性质	2
	单纯型算法及其变形	4
	对偶原理与对偶单纯性算法	4
	灵敏度分析及其应用	2
整数规划 24（Integer programming）	整数规划建模	2
	整数规划方法 分支定界、割平面等方法	4
	整数规划对偶理论 Lagrange理论	2
	运输问题应用	2
32网络流问题		8
34动态规划		2
42非线性规划（Nonlinear Programming）	无约束规划 最优化算法	4
	有约束规划 最优化算法	4
44*现代优化算法-进化计算		2
46*线性规划内点方法		2
48*大规模线性规划方法		2
考试		----

参考书：【1】最优化理论与算法，陈宝林，清华大学出版社；
【2】Dimitris Bertisimas, John N.Tsitsikils. Introduction to Linear Optimization. Athena Scientific
【3】刑文训、谢金星. 现代优化计算方法，清华大学出版社；.

Grading

- 开/闭卷考试-80%
- 作业及应用案例报告-15%
- 其他情况-5%

运筹学历史

- 运筹学（Operations Research），是一门基于应用数学的跨领域研究，利用统计学、数学模型和优化算法，去寻找复杂问题中的**最佳或近似最佳**解答。

- 1955年我国从“运筹帷幄之中，决策千里之外”（见《史记》）这句话摘取“运筹”二字，将O.R.正式译作运筹学。

- 运筹学实际上是一门实用性强的课程，强调解决实际运营中问题。

运筹学研究内容

- 运筹学作为一门现代科学，是在第二次世界大战期间首先在英美两国发展起来的，有的学者把运筹学描述为为组织系统的各种经营作出决策的科学手段。

- 运筹学的具体内容包括：规划论（包括线性规划、非线性规划、整数规划和动态规划）、库存论、图论、决策论、对策论、排队论、博弈论、可靠性理论等。

运筹是目标，优化是手段

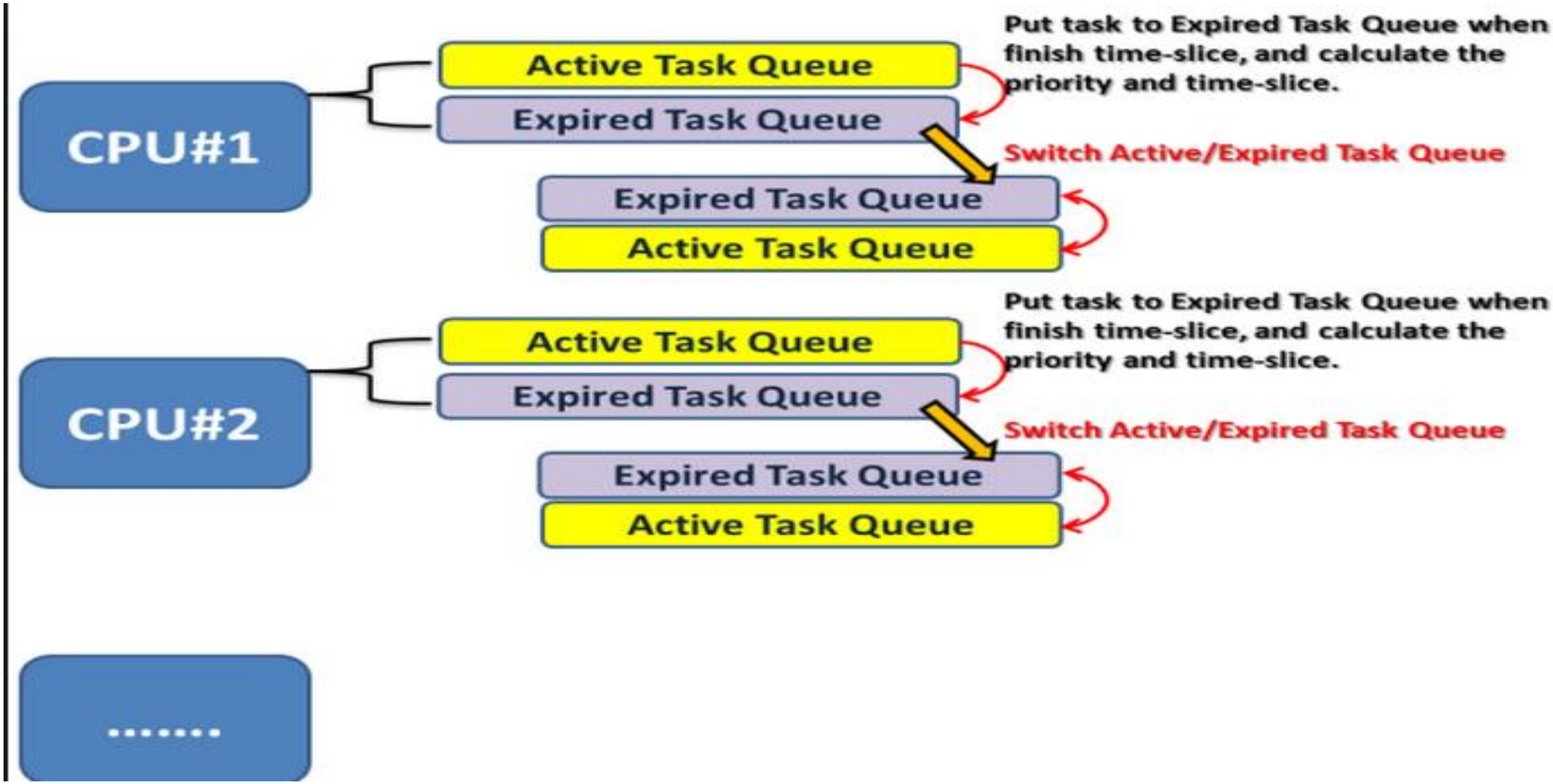
运筹学性质与特点

- 开放性
- 复杂和多样性
- 多学科交叉
- 系统与整体性
- 数学方法解决实际问题

运筹学的发展

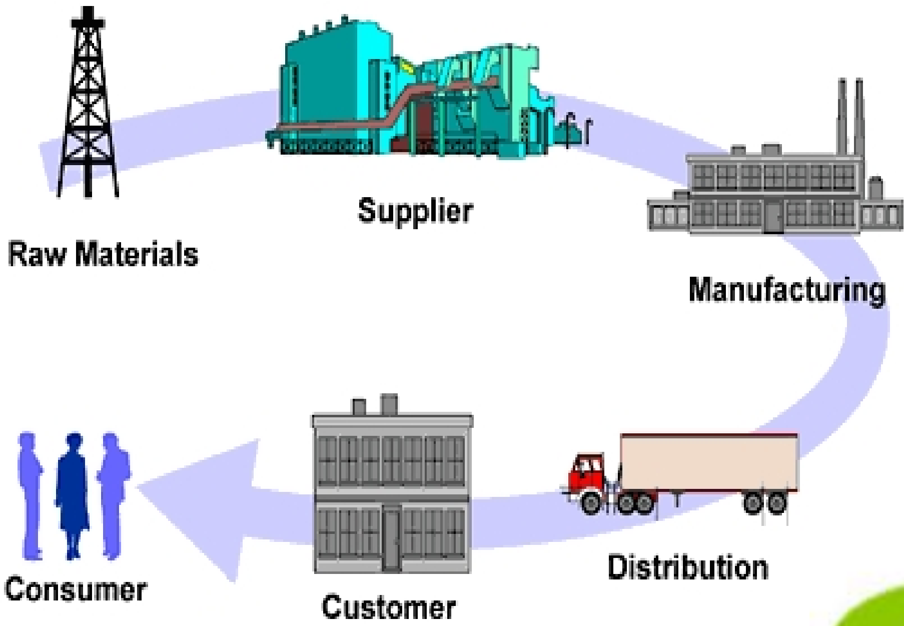
- 运筹学（Operations Research），是一应用数学和形式科学的跨领域研究，发展至今已广泛的深入到各种应用之中。

计算机架构设计-多核CPU任务调度



任务复杂度，任务优先级，用户满意度，负载均衡性

商业领域应用-供应链管理



原材料供应稳定，
制造的协同性，
运输经济性，
用户满意度，
供应链网络安全性

政治军事应用-国家间的政治博弈

- ☐ 博弈论（Game Theory）已经广泛的应用于大国之间竞争中。如美国的智库兰德公司，它的长处是进行战略研究。兰德公司组织大批专家对朝鲜战争进行评估，通过建立复杂的博弈模型，准确的在朝鲜战争前夕预测“中国将出兵朝鲜”。
- ☐ 科学家们还研究核武器的博弈模型，得出的结论是核威慑作用运大于核武器的实际使用效果。

运筹学在交通的应用效益统计

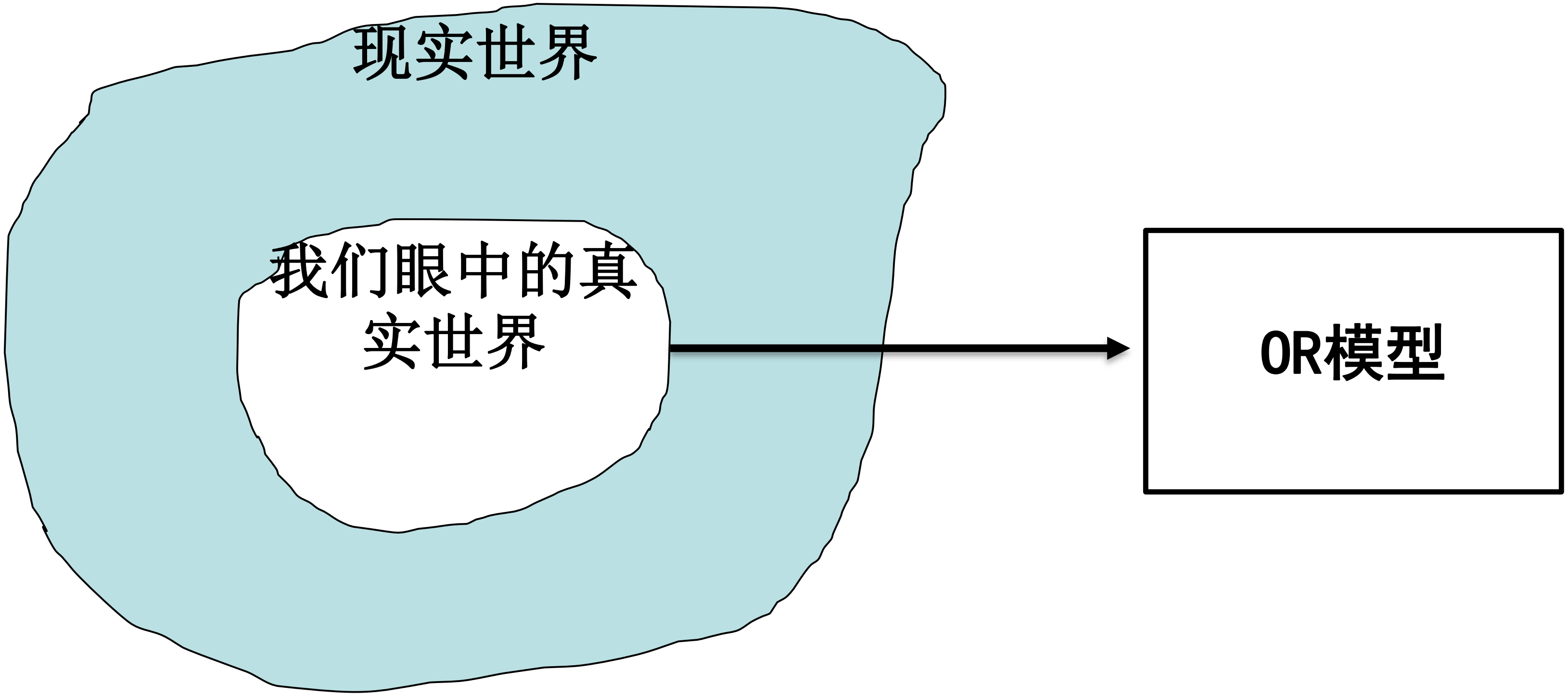
- 物流网络的运输优化，节约15%成本，提升5%交通运输效率，CO_2排放减少16.86万吨。
- 出行时间降低12%，缩短20%的应急反应时间，节约2640亿美元
- （2016年《国际空间数据产业分析报告》数据）

OR问题研究原则(methodology)

- 定义问题
- 构建模型
- 模型求解
- 模型验证
- 方案实施

迭代演进过程

问题定义-模型



OR问题定义

- **需求：确定好需要优化的目标**
- **决策：确定好优化需决策变量**
- **约束：确定好优化过程中的约束条件**

问题：如何以高分通过“运筹与最优化”这门课程？

OR问题的数学建模

- 把问题定义从语言描述转换为数学关系式

- 确定数学模型的求解算法的复杂程度

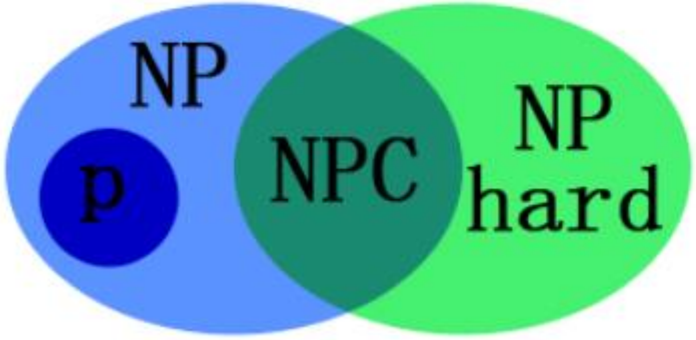
  - 空间复杂度：**输入、计算过程中的所需的存储空间。**

  - 时间复杂度：
    算法运行迭代中所需要的时间，由于算法运行时间与硬件相关的，为了便于比较，通常将时间复杂度简化为基本计算（加、减、乘、除、布尔运算等）的次数。

OR问题的复杂度

- 如n维向量计算的计算次数为：（$2n-1$）（n次乘法+n-1次加法）
- 而n维方阵相乘的计算次数为：（$2n-1$）n^2

根据求解算法的复杂度可以将OR问题分为：



P指多项式时间（Polynomial）

NP指不确定性多项式时间（nondeterministic polynomial）

更难的则是NPC和NP-hard问题

计算模型复杂度

- P/NP问题中包含了复杂度类P与NP的关系。1971年史提芬·古克（Stephen A. Cook）和Leonid Levin相对独立地提出了下面的问题，即复杂度类P和NP是否是恒等的（P=NP?）

- **P/NP问题**是在理论信息学中计算复杂度理论领域里至今未被解决的问题，也是克雷数学研究所七个千禧年大奖难题之一。

OR问题的算法求解

- 根据数学模型的复杂程度，针对问题的特性，确定求解算法

- 从应用角度的看，算法是相对单纯的一个环节

- 关键点：

  尽可能从数学模型中获得最优解的信息用于指导算法的搜索过程

OR问题的模型验证

- 验证模型是否正确？
- 答案是否有意义？
- 有没有超出我们想象的结果？

- 验证三法：
- （1）逻辑验证（包括灵敏度分析）；
- （2）历史数据验证；
- （3）仿真验证；

OR模型的好坏

- 何为“好”的优化模型？
- 模型中的冗余信息越好越好，主要外在表现：

  1、决策变量越少越好；

  2、过程约束越少越好；

OR问题的方案落地实施

- 把运筹学优化结果转变为现实中可操作的措施：
- 直接应用：

  形成软件、模块等应用工具

- 辅助决策：

  （1）为重大战略决策提供依据；

  （2）将优化结果简化，针对性提供分类指导依据（根据场景分类）；

运筹学软件应用框架

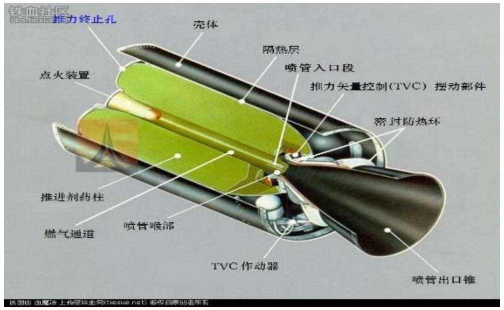
模型与算法	软件开发	场景应用
数学规划理论	算法（BOOST）	供应链设计
启发式优化	数据库（SQL.）	…
算法性能分析	可视化（UML）	流程优化
…	…	

运筹学中的若干例子

- 维生素摄入优化问题

- 物流运输优化问题

- 最优化控制问题



运筹学案例-食谱问题

- 人每天需要摄入多种维生素才能保证人的健康，假设考虑三种维生素A,B,C，对三种维生素的需要分别为60，100，80. 现可采购5种食物，这些食物的价格和每100单位中含有的维生素数量如下表：

食品	A	B	C	单价
1	3	5	4	50
2	0	6	3	30
3	4	0	2	15
4	1	4	2	18
5	11	0	4	21

- 问各食物每天采购多少量，才能既保证三种维生素的摄入，由能够使费用最低?

食谱问题成本优化模型

■ **决测变量：**

X1，*X2*，*X3*，*X4*，*X5*，分别表示每种食物购买的数量（以100为单位）？

■**优化目标：成本最低**

$$\min = 50X1 + 30X2 + 15X3 + 18X4 + 21X5$$

目标参数

■**约束条件：满足维生素摄入需求**

$$3X1 + 0X2 + 4X3 + 1X4 + 11X5 >= 60$$

$$5X1 + 6X2 + 0X3 + 4X4 + 0X5 >= 100$$

$$4X1 + 3X2 + 2X3 + 2X4 + 4X5 >= 80$$

约束参数

***X1*，*X2*，*X3*，*X4*，*X5* 均需要大于等于0，也可以具体需求设定某种食物采购的最低值。**

运输问题实例

- 设要从甲地调出物质2000吨，从乙地调出物质1100吨，分别供给A地1700吨、B地1100吨、C地200吨、D地100吨。已知每吨运费(单位百元)如表所示：

	A	B	C	D
甲地	21	25	7	15
乙地	51	51	37	15

假定运费与运量成正比，问怎样才能找出运费最省的调拨计划？

运输问题优化建模

■ **决测变量：**

x_{ij} 表示从第 i 个调货地运输至第 j 个目的地的数量；i={1,2}两个调货地，j ={1,2,3,4}四个货物的目的地；

■**优化目标：成本最低**

min= $21x_{11}+25x_{12}+7x_{13}+15x_{14}+51x_{21}+51x_{22}+37x_{23}+15x_{24}$

■**约束条件：**

$$\begin{cases} x_{11}+x_{12}+x_{13}+x_{14}\le 2000 \\ x_{21}+x_{22}+x_{23}+x_{24}\le 1100 \\ x_{11}+x_{21}\ge 1700 \\ x_{12}+x_{22}\ge 1100 \\ x_{13}+x_{23}\ge 200 \\ x_{14}+x_{24}\ge 100 \end{cases}$$

过程约束

$$x_{ij}\ge 0$$

变量约束

运输优化问题的特征

目标：线性目标函数

min= $21x_{11}+25x_{12}+7x_{13}+15x_{14}+51x_{21}+51x_{22}+37x_{23}+15x_{24}$

约束：线性不等式约束

$$
\begin{cases}
x_{11}+x_{12}+x_{13}+x_{14} \le 2000 \\
x_{21}+x_{22}+x_{23}+x_{24} \le 1100 \\
x_{11}+x_{21} \ge 1700 \\
x_{12}+x_{22} \ge 1100 \\
x_{13}+x_{23} \ge 200 \\
x_{14}+x_{24} \ge 100
\end{cases}
$$

线性规划：

Linear Programming

控制问题的最优化模型

- 动态系统：
- X(t+1)=AX(t)+BU(t)
- Y(t)　=CX(t)

决策变量：U(0)…U(T-1)

使得输出Y(t)能够快速趋于一个目标值（“0”）

控制问题的最优化建模

$$Min\ Max_{t=0,\ldots,T-1}|y(t)|$$

非线性模型

Subject to

- X(t+1)=AX(t)+BU(t)
- y(t) =CX(t)
- DU(t)<= d
- X(T)=0
- X(0)=given

控制问题的最优化建模

$$Min\ z$$

线性模型

Subject to

$-z \leq y(t) \leq z$ 对每一个 $t = 1 \ldots T - 1$

- X(t+1)=AX(t)+BU(t)
- y(t)　=CX(t)
- DU(t)<= d
- X(T)=0
- X(0)=given
- $z \geq 0$

优化算法发展历史

优化算法有着悠久的历史，每一类优化算法都有针对型的问题，优化算法一直都在完善和发展过程中

Fermat, 1638; Newton, 1670

$$\min \; f(x) \qquad x: \text{ scalar}$$

$$\frac{df(x)}{dx} = 0$$

Euler, 1755

$$\min \; f(x_1, \ldots, x_n)$$

$$\nabla f(\mathbf{x}) = 0$$

Lagrange, 1797

$$\min \quad f(x_1, \ldots, x_n)$$

$$\text{s.t.} \quad g_k(x_1, \ldots, x_n) = 0 \qquad k = 1, \ldots, m$$

Euler, Lagrange Problems in infinite dimensions, calculus of variations.

现代数学优化算法发展

George Dantzig, 1947 Simplex method.

1950s Applications.

1960s Large Scale Optimization.

1970s Complexity theory.

1979 The ellipsoid algorithm.

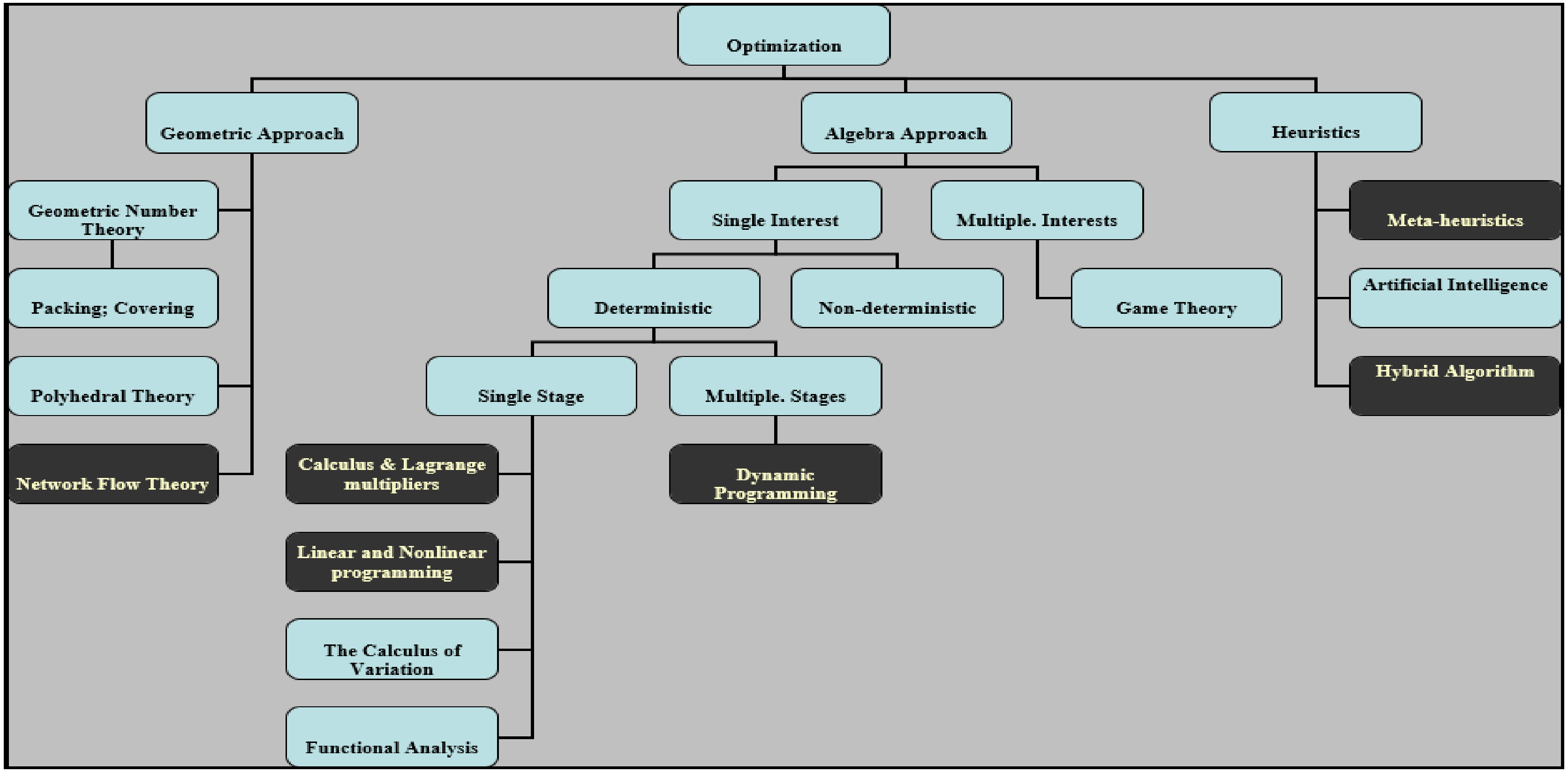
1980s Interior point algorithms.

1990s Semidefinite and conic optimization.

2000s Robust Optimization.

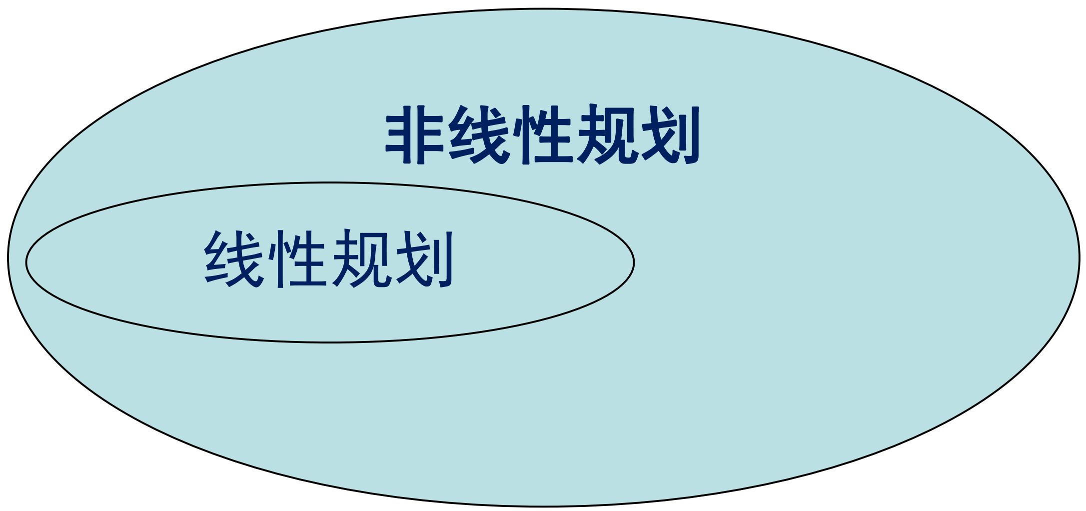
最优化算法分类

- 代数方法
- 几何方法
- 启发式方法



数学规划基础知识

数学规划是解决运筹学问题的一类重要研究分支



线性规划（Linear programming，简称LP），是运筹学中研究较早、发展最快、应用广泛、方法最成熟的一个重要分支

线性规划（Linear Programming）

目标：线 性 函 数：

$$f(x_1, x_2, \dots, x_n)\text{线性函数} \overset{if\ and\ only\ if}{\Longleftrightarrow} f(x_1, x_2, \dots, x_n) = C_1x_1 + C_2x_2 + \cdots C_nx_n$$

约束：线性不等式：

$$g(x_1, x_2, \dots, x_n)\text{线性函数；}$$
$$g(x_1, x_2, \dots, x_n) \geq 0\text{或者}$$
$$g(x_1, x_2, \dots, x_n) \leq 0\text{或者}$$
$$g(x_1, x_2, \dots, x_n) = 0$$

线性规划：

- 目标函数是线性函数
- 决策变量满足的约束是线性方程或者线性不等式
- 决策变量要求>=0 或者=0或者<=0

线性规划的一般描述

- Max 或者 Min $\sum_{i=1}^{n} c_i x_i$

- S.t.(Subjected to)

$$g(x_1, x_2, \dots, x_n)\text{线性函数；}$$
$$g(x_1, x_2, \dots, x_n) \geq 0\text{或者}$$
$$g(x_1, x_2, \dots, x_n) \leq 0$$
$$x_i \geq 0, \forall i$$

线性规划的几个延伸概念

- 非线性规划：目标函数为非线性函数或者约束中存在非线性等式或者不等式
- 可行解：在线性或者非线性规划中，满足约束条件的解
- 可行集：在线性或者非线性规划中，所有可行解的集合。
- 全局极小点：该可行解为全可行集中最小。
- 局部极小点：存在可行集的某个范围，该可行解最小

理解概念的物理意义

- 向量范数：一种衡量向量长度的定义，满足三个条件。如欧氏范数
- 矩阵范数：一种衡量矩阵长度的定义，满足四个条件。如谱范数
- 梯度，Hesse矩阵，Taylor展开式

向量范数

- 向量范数：定义向量的尺度

(1) n维实值函数：

$$\|\blacksquare\|: R^n \to R:$$

(2) 满足三个条件：

(1)$\|\boldsymbol{x}\| \geq 0, \forall x \in R^n$; $\|\boldsymbol{x}\|$=0当且仅当$x = 0$;

(2) $\|a\boldsymbol{x}\| = |a|\|\boldsymbol{x}\|, \forall \boldsymbol{x} \in R^n, \forall a \in R$;

(3) $\|\boldsymbol{x} + \boldsymbol{y}\| \leq \|\boldsymbol{x}\| + \|\boldsymbol{y}\|, \forall \boldsymbol{x}, \boldsymbol{y} \in R^n$

示例：L_1, L_2, L_∞

向量范数

设$x = (x_1, x_2, \ldots, x_n)^T \in R^n$，对于$1 \le p < \infty$，$L_p$范数定于如下：

$$\|x\|_p = \left(\sum_{j=1}^{n} |x_j|^p\right)^{1/p}$$

常用的向量范数有：

$$\|x\|_1 = \sum_{j=1}^{n} |x_j|$$

$$\|x\|_2 = \left(\sum_{j=1}^{n} |x_j^2|\right)^{1/2}$$

？？？

Question：L_∞如何证明？

矩阵范数

- 定义矩阵的尺度
- 现有向量范数定义-》矩阵范数定义
- 诱导矩阵范数定义：设A为$m \times n$矩阵:

$$\|\boldsymbol{A}\|_p \stackrel{\text{def}}{=} \max_{\|\boldsymbol{x}\|_{\boldsymbol{p}}=1} \|\boldsymbol{A}\boldsymbol{x}\|_p$$

矩阵范数性质

矩阵范数的性质：

单位矩阵$\boldsymbol{I}$, $\|\boldsymbol{I}\|=1$

$\|\boldsymbol{Ax}\| \le \|\boldsymbol{A}\| \ \|\boldsymbol{x}\|$ 相容性

$\|\lambda \boldsymbol{A}\| = |\lambda| \ \|\boldsymbol{A}\|$ 齐次性

$\|\boldsymbol{A}+\boldsymbol{B}\| \le \|\boldsymbol{A}\| + \|\boldsymbol{B}\|$ 三角不等式

$\|\boldsymbol{AD}\| \le \|\boldsymbol{A}\| \ \|\boldsymbol{D}\|$ 相容性

常用矩阵范数：

$\|\boldsymbol{A}\|_1 = \max_j \sum_{i=1}^m |a_{ij}|$ 最大的列和

$\|\boldsymbol{A}\|_2 = \sqrt{\lambda(\boldsymbol{A}^T\boldsymbol{A})}$ 谱范数

$\|\boldsymbol{A}\|_\infty = \max_i \sum_{j=1}^n |a_{ij}|$ 最大的行和

F范数$\|\boldsymbol{A}\|_F = \sqrt{\sum\sum(a_{ij})^2}$ **非诱导范数**

梯度（Gradient）和Hesse矩阵

- 梯度：函数f 在x的梯度（多元函数）：

- $$\nabla f(x)=\begin{bmatrix}\frac{\partial f(x)}{\partial x_1}\\ \frac{\partial f(x)}{\partial x_2}\\ \vdots\\ \frac{\partial f(x)}{\partial x_n}\end{bmatrix}$$

- 表征函数的趋势

- Hesse矩阵：
- $\nabla^2 f(x)=$

$$\begin{bmatrix}\frac{\partial^2 f(x)}{\partial x_1\partial x_1},\cdots,\frac{\partial^2 f(x)}{\partial x_1\partial x_n}\\ \frac{\partial^2 f(x)}{\partial x_2\partial x_1},\cdots,\frac{\partial^2 f(x)}{\partial x_2\partial x_n}\\ \vdots\\ \frac{\partial^2 f(x)}{\partial x_n\partial x_1},\cdots,\frac{\partial^2 f(x)}{\partial x_n\partial x_n}\end{bmatrix}$$

表征函数趋势的变化

Taylor展开

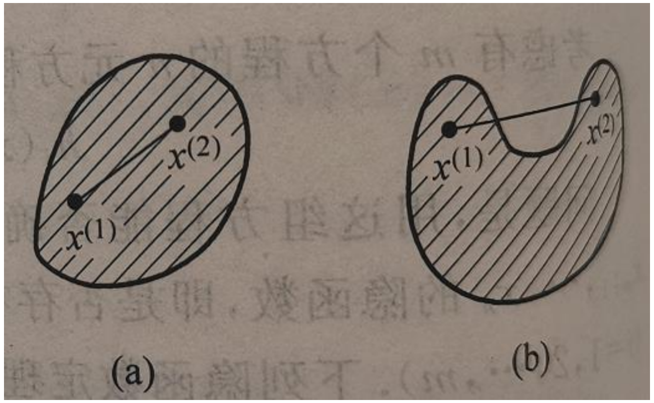
- 在数学中，泰勒公式是一个用函数在某点的信息描述其附近取值的公式。
- 泰勒公式可以用这些导数值做系数构建一个多项式来近似函数在这一点的邻域中的值。泰勒公式还给出了这个多项式和实际的函数值之间的偏差。

- 18世纪早期英国牛顿学派最优秀代表人物之一的英国数学家泰勒（Brook Taylor），于1685年初出生（清康熙年间），他在1712年当选为英国皇家学会会员，并于两年后获法学博士学位。同年（即1714年）出任英国皇家学会秘书，四年后因健康理由辞退职务。
- 1717年，他以泰勒定理求解了数值方程。
- 最后在1731年12月29日于伦敦逝世。

基本概念-凸集

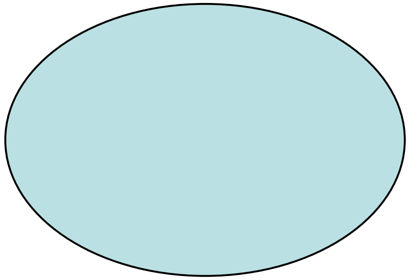
凸集的定义：

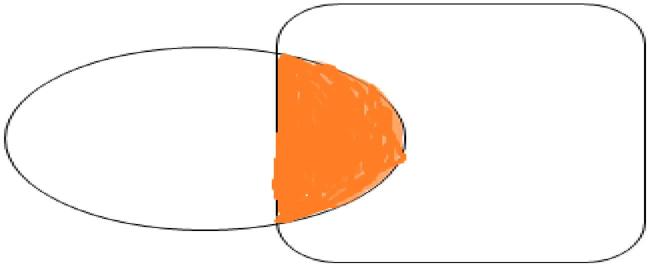
$$\lambda x^{(1)} + (1-\lambda)x^{(2)} \in S$$
$$x^{(1)} \in S$$
$$x^{(2)} \in S$$
$$\lambda \in [0,1]$$



几个凸集的例子

$$D = \{x \mid Ax = b, x \geq 0\}$$

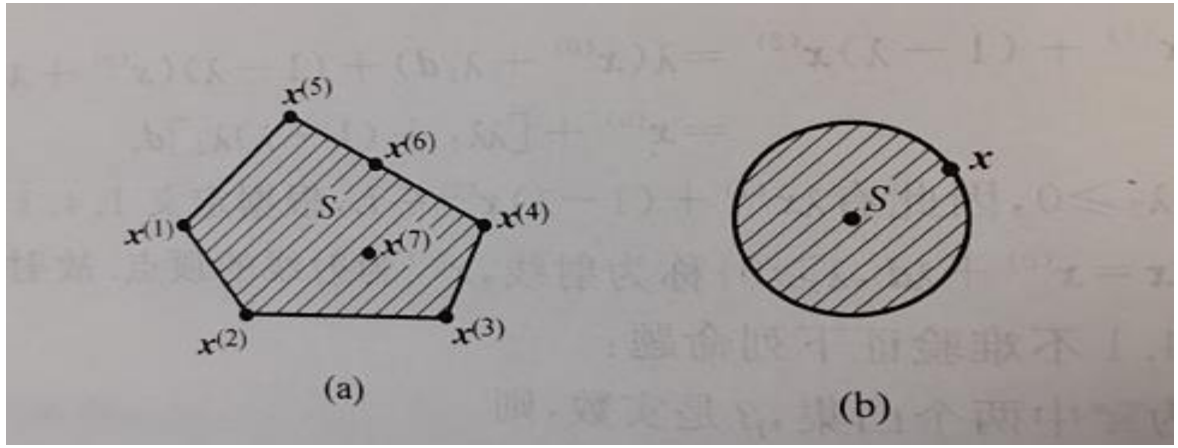




凸集的极点与极方向

基本概念-凸集

凸集的极点

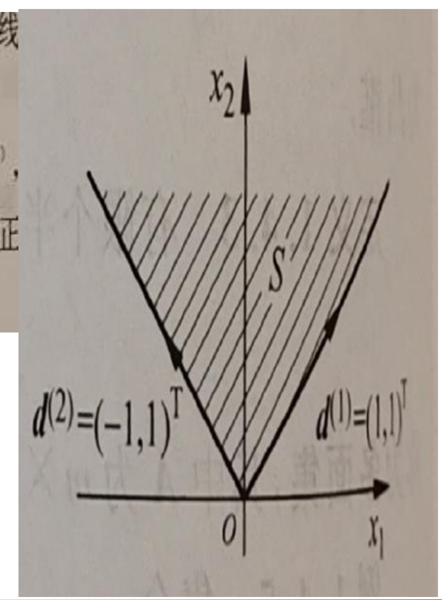


方向与极方向

定义 1.4.5 设 S 为 $\mathbb{R}^n$ 中的闭凸集，d 为非零向量，如果对 S 中的每一个 x，都有射线

$$\{x+\lambda d \mid \lambda \geqslant 0\} \subset S,$$

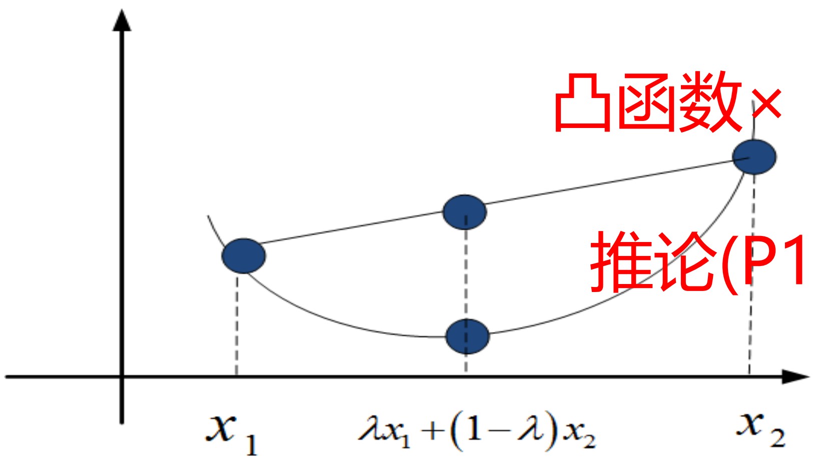
则称向量 d 为 S 的**方向**. 又设 $d^{(1)}$ 和 $d^{(2)}$ 是 S 的两个方向，若对任何正数 λ，有 $d^{(1)} \neq \lambda d^{(2)}$，则称 $d^{(1)}$ 和 $d^{(2)}$ 是两个不同的方向. 若 S 的方向 d 不能表示成该集合的两个不同方向的正的线性组合，则称 d 为 S 的**极方向**.



凸集的方向可能会有无数，极方向有限的

凸函数和凹函数

- 凸函数定义： $f(\lambda x_1+(1-\lambda)x_2)\le \lambda f(x_1)+(1-\lambda)f(x_2)$
- 严格凸函数定义： $f(\lambda x_1+(1-\lambda)x_2)< \lambda f(x_1)+(1-\lambda)f(x_2)$



凸函数×（-1）就是凹函数

推论(P18)页如何证明?

凸函数的几条基本定理

$$f(x)\text{是凸函数}\Rightarrow \lambda f(x)\text{也是凸函数}$$

$$f_1(x),\ f_2(x)\text{是凸函数}\Rightarrow f_1(x)+f_2(x)\text{也是凸函数}$$

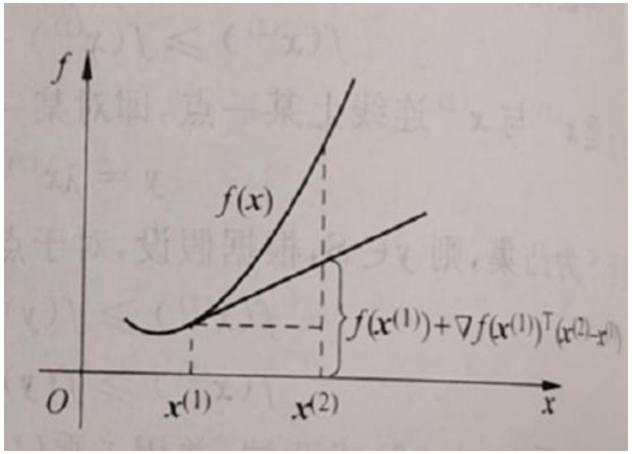
凸函数的判别

- 1、根据定义判别

- 2、具有一次可微的多元函数判别，$f(x)$是凸函数的充分必要条件是：

对任意互不相同的两点x_1，x_2，有

$$f(x_2) \geq f(x_1) + \nabla f(x_1)^{\mathrm{T}}(x_2 - x_1)$$



凸函数的判别

3 具有二次可微的多元函数判别

二次可微的多元函数判别， $f(x)$是凸函数的**充分必要条件是：**

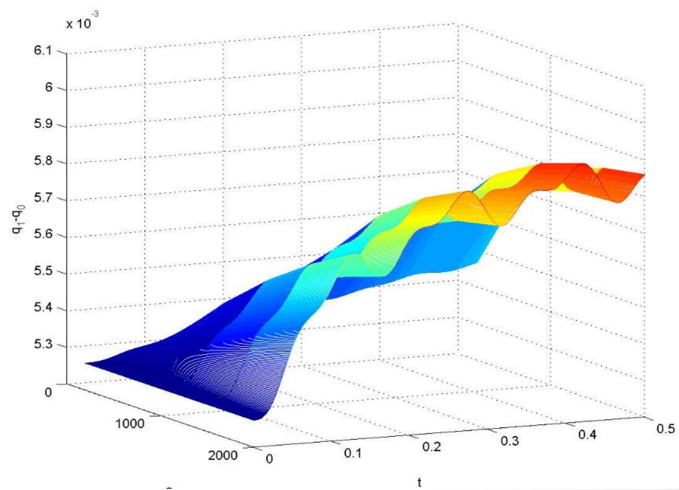
在每一点处（在定义域范围内）Hesse矩阵半正定。

若Hesse矩阵正定，说明$f(x)$是严格凸函数，

若$f(x)$是严格凸函数，则Hesse矩阵半正定

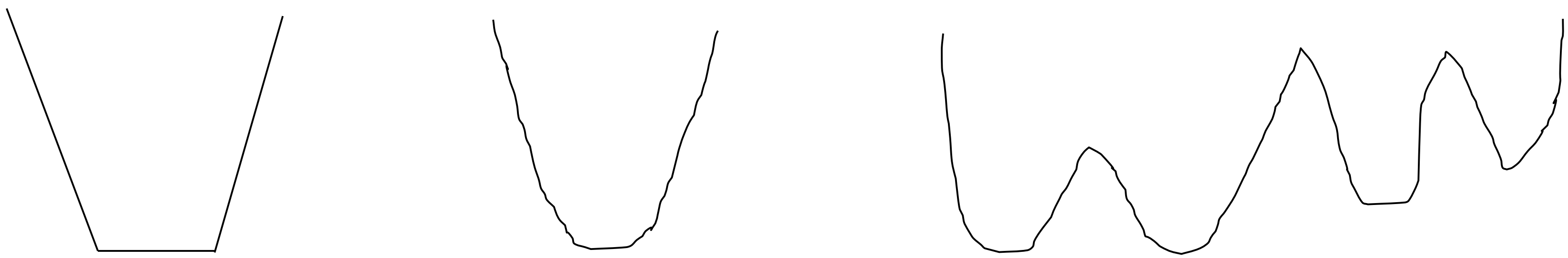
凸规划定义

- 凸规划满足两个条件：
- 1. 目标函数是凸函数
- 2. 可行解集是凸集



凸规划

线性规划是最简单的凸规划



凸规划的性质：

- 凸规划的局部极优点是全局极优点，且极优点的集合是凸集；
- 进一步，如果凸规划的目标函数是严格凸函数，又存在极优解，那么极优解是唯一的。

凸规划的一般形式

Why?

$$\min f(x)$$
$$\text{s.t.}\quad \begin{aligned} &g_i(x) \geq 0, i = 1, \dots, m \\ &h_j(x) = 0, j = 1, \dots, l \end{aligned}$$

$f(x)$是凸函数，$\boxed{g_i(x)\text{是凹函数，}}$ $h_j(x)$是线性函数

知识点回顾

- OR建模5步骤
- 凸集、凸函数、凸规划
- 数学证明法：1、推理法；2构造法；3、反证法

课后作业

P23页～P24页：
第1大题；
第2大题；
第3大题；
第4大题；
第9.1题，9.3题

本次课到此结束，谢谢！