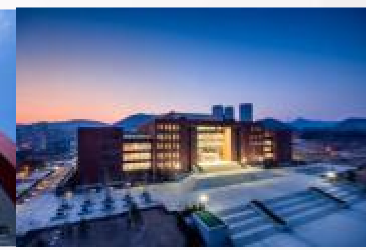
最优化方法
Optimality Methods

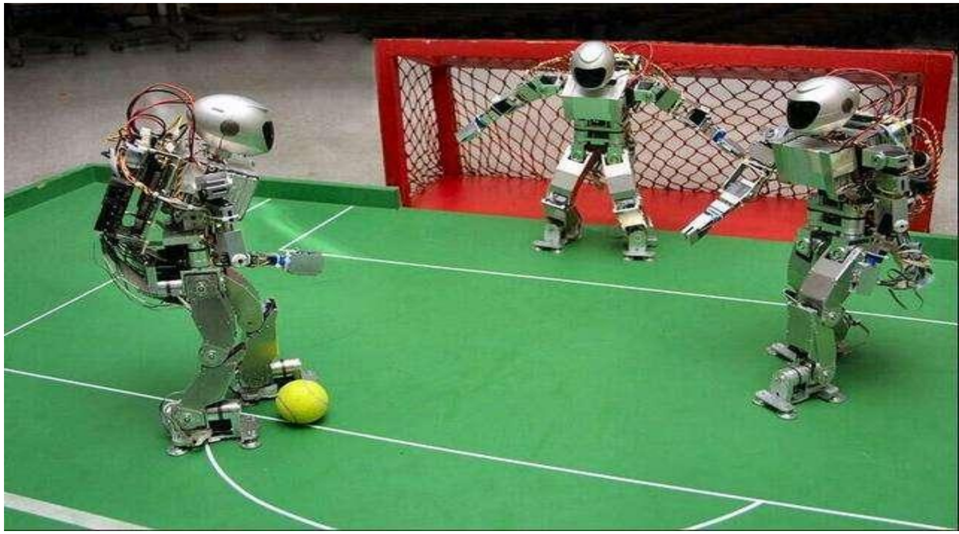


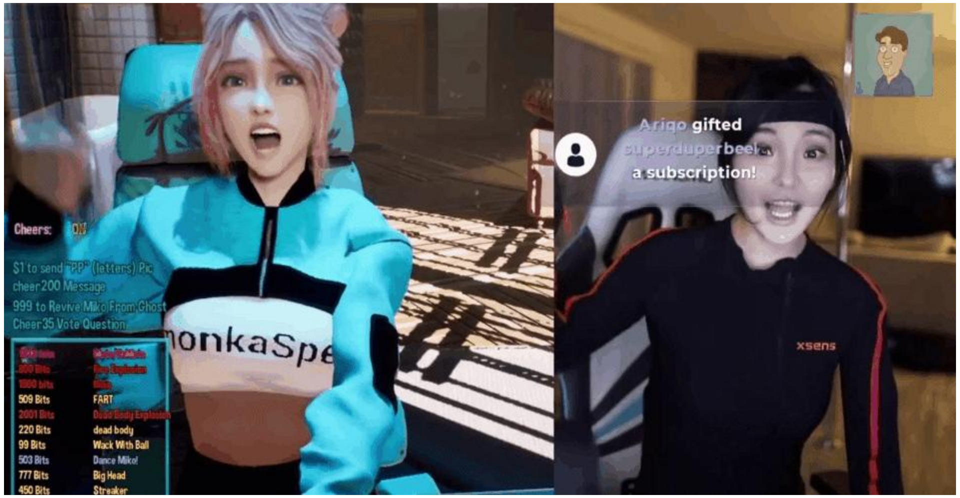




课程简介

天猫全球狂欢季 2020
双11终极狂欢
全年底价 最后一天 买1天省1年
逛主会场
喵，欢迎来天猫
请登录
免费注册
我的淘宝
购物车
收藏夹
手机版
淘宝网
商家支持
网站导航
双11买1天省1年
全年底价，每满300减40
搜索 天猫 商品/品牌/店铺
搜索
商品分类
20年双11
服饰风尚
超值囤货
手机数码
家具建材
精品家具
全球尖货
天猫超市
苏宁精选
00:05:25
成交额超过300亿
GMV has surpassed RMB 30 billion
眼睛一闭一睁，嗷~~
Blink and you'll miss it!
超值精品大聚会！

课程简介





目 录 (CONTENT)

大连理工大学
Dalian University Of Technology

01
PART ONE

课程简介
Introduction

大连理工大学 Dalian University Of Technology

■ 课程简介

授课教师　张平平未来技术学院/人工智能学院　副教授

时间　**10-17周**

地点　综合教学2号楼**B406** 周一**(7-8)**、周四**(3-4)**

联系方式

Email: zhpp@dlut.edu.cn

办公室：学生创新实践中心**243**

■ 课程简介

最优化方法2...

扫一扫二维码，加入群聊

QQ

助教 (孙珲)

QQ学习群: 752043812

课程大纲

- 引言
- 线性规划的基本性质
- 单纯形方法原理
- 对偶原理及灵敏度分析
- 运输问题
- 线性规划的内点算法
- 最优性条件
- 算法
- 一维搜索
- 使用导数的最优化方法
- 无约束最优化的直接方法
- 可行方向法
- 惩罚函数法
- 二次规划
- 整数规划简介
- 动态规划简介

■ 课程考核(32学时)

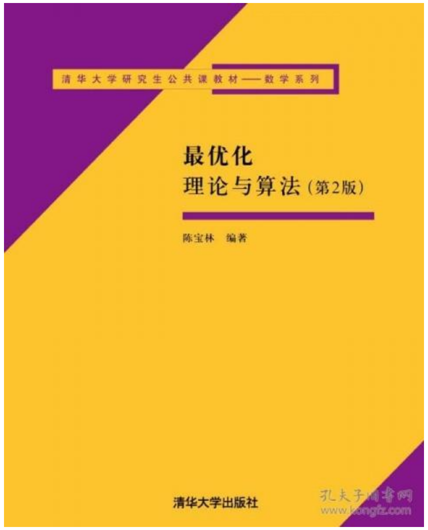
■平时作业（30%）

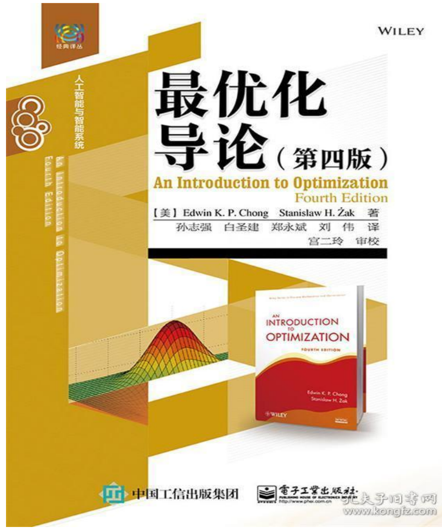
- ➢内容：最优化方法课后作业
  - ●独立完成（抄袭扣平时分10）
  - ●每人至少完成4次作业（少于4次扣10）
  - ●未按题目数量完成（扣1-5）
  - ●错误率较高（扣1-5）
- ➢提交时间：课程章节结束后，下一章节开始前，按班级提交给助教
- ➢考勤：不定时点名至少5次（每次10分）

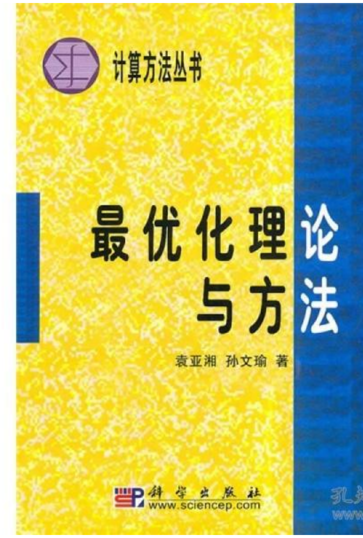
■期末考试（70%）

- ➢闭卷考试，课程过程及习题作业中选择
- ➢百分制，选择+计算
- ➢考试成绩为卷面成绩，除合分错误，不存在修改的可能

■ 参考书目







1. 陈宝林 《最优化理论与算法(第二版)》清华大学出版社，2005
2. [美]Edwin K.P. Chong Stanislaw H.Zak 著，/孙志强，白圣建，郑永斌，刘伟(译)/宫二玲(校) 《最优化导论(第四版)》电子工业出版社，2016.
3. 傅英家等主编《最优化理论与方法》国防工业出版社，2008.
4. 袁亚湘、孙文瑜 著《最优化理论与方法》科学出版社，1997.

大连理工大学 Dalian University Of Technology

■ 相关/推荐课程

- 运筹学，西安邮电大学，
- https://www.icourse163.org/course/XIYOU-1207132805

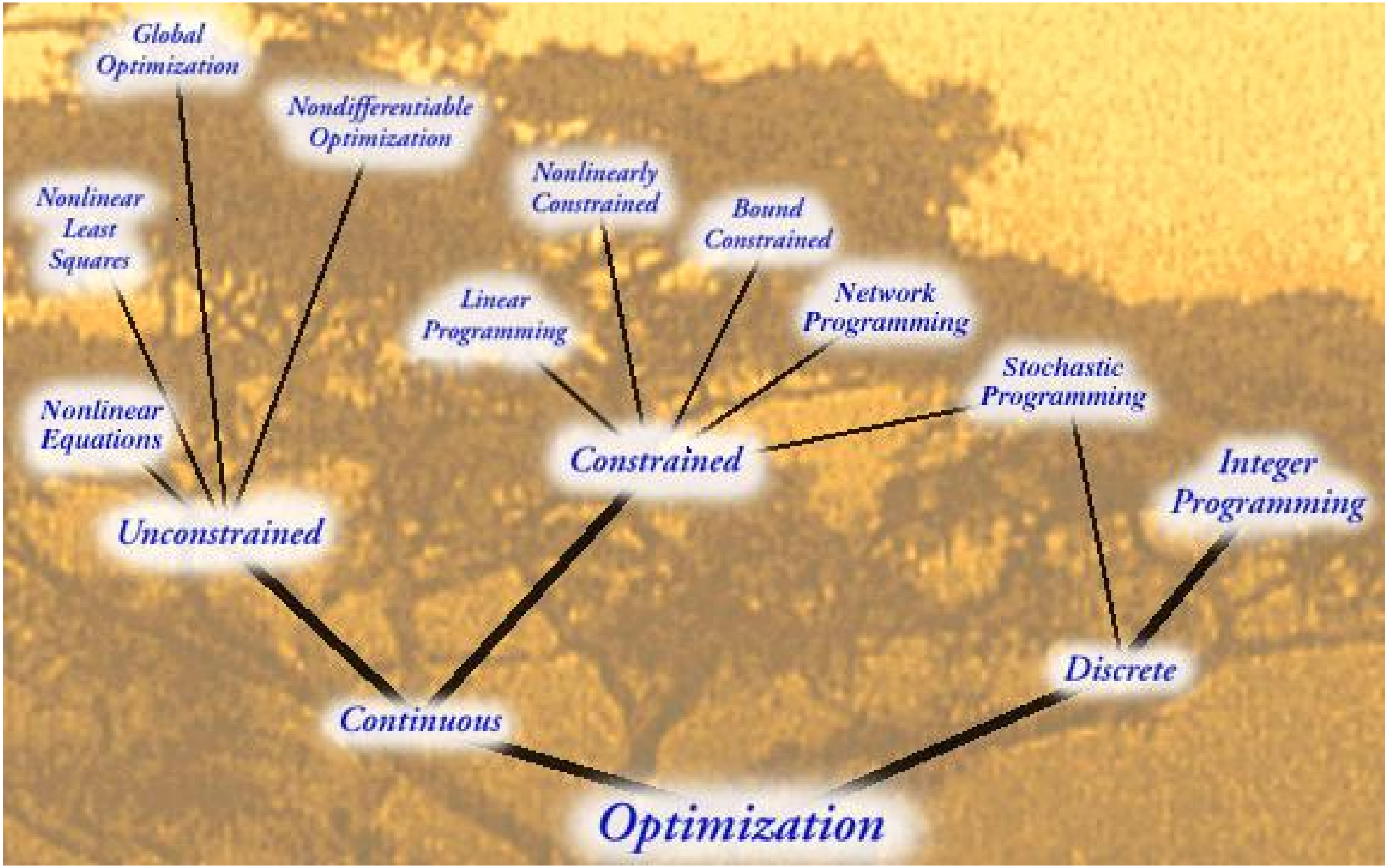
- 最优化理论，南京理工大学，
  https://www.icourse163.org/course/preview/NJUST-1003545092/?tid=1003779153

- 最优化方法，四川大学，
- https://www.icourse163.org/spoc/course/SCU-1206412801

■ 学科概述

- 最优化是从所有可能的方案中选择最合理的一种方案，以达到最佳目标的科学
- 达到最佳目标的方案是最优方案，寻找最优方案的方法就是最优化方法(算法)
- 这种方法的数学理论即为最优化理论
- 最优化方法是运筹学方法论之一
- 最优化首先是一种理念，其次才是一种方法

■ 优化树


Global Optimization
Nondifferentiable Optimization
Nonlinear Least Squares
Nonlinearly Constrained
Bound Constrained
Linear Programming
Network Programming
Stochastic Programming
Nonlinear Equations
Constrained
Integer Programming
Unconstrained
Discrete
Continuous
Optimization

■ 最优化的发展历程(理论)

- 费马(1638);牛顿(1670)

$$\min_x f(x) \quad x \in R$$

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

一元函数

- 欧拉(1755)

$$\min f(x_1, x_2, \ldots, x_n)$$

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

多元函数

- 拉格朗日(1797)

$$\min \quad f(x_1, x_2, \ldots, x_n)$$

$$\text{s.t.} \quad g_k(x_1, x_2, \ldots, x_n) = 0$$

多元约束

$$k = 1, 2, \ldots, m$$

- 欧拉，拉格朗日：无穷维问题，变分学
- 柯西：最早应用最速下降法

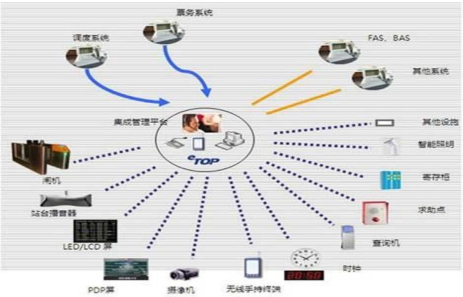
最优化的发展历程(计算机)

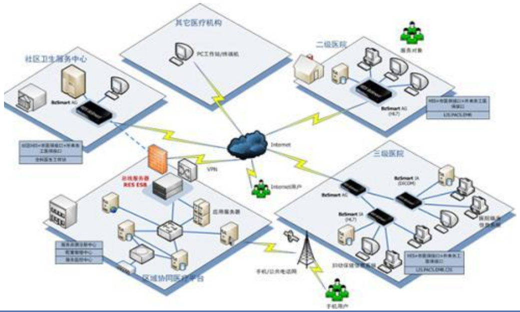
- 1930年代，康托诺维奇：线性规划
- 1940年代，Dantzig：单纯形方法
  冯 诺依曼：对策论
- 1950年代，Bellman：动态规划，最优性原理；KKT条件
- 1960年代，Zoutendijk,Rosen,Carroll等非线性规划算法
  Duffin,Zener等几何规划，
  Gomory等整数规划,
  Dantzig等随机规划
- 1970年代，Cook等复杂性理论，组合优化迅速发展

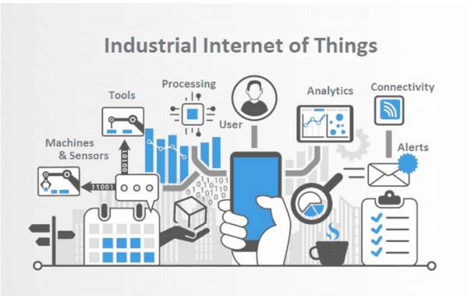
最优化应用举例

具有广泛的实用性

- 运输问题，车辆调度，员工安排，空运控制等
- 工程设计，结构设计等
- 资源分配，生产计划等
- 通信：光网络、无线网络、互联网、物联网等
- 制造业：钢铁生产，车间调度等
- 医药生产，化工处理等
- 电子工程，集成电路VLSI 等
- 电子排版（TEX,Latex等）







1.食谱问题

我每天要求一定量的两种维生素，V_c和V_b。
假设这些维生素可以分别从牛奶和鸡蛋中得到。

维生素	奶中含量	蛋中含量	每日需求
V_c(mg)	2	4	40
V_b(mg)	3	2	50
单价($)	3	2.5	

需要确定每天喝奶和吃蛋的量，
目标以便以最低可能的花费购买这些食物，
而**满足**最低限度的维生素需求量。

■ 1.食谱问题

令x表示要买的奶的量，y为要买的蛋的量。食谱问题可以写成如下的数学形式：

$\text{Min } 3x + 2.5y$	极小化目标函数
$\text{s.t. } 2x + 4y \geq 40$	
$3x + 2y \geq 50$	可行区域
$x, y \geq 0.$	可行解

建立关于何时出现最小费用(或者最大利润)的排序，或者计划，早期被标示为**规划**。

求最优安排或计划的问题，称作**规划问题**。

■ 2.运输问题

设某种物资有m个产地$A_1,A_2,\ldots,A_m$,各产地的产量是$a_1,a_2,\ldots,a_m$;有n个销地$B_1,B_2,\ldots,B_n$.各销地的销量是$b_1,b_2,\ldots,b_n$.假定从产地$A_i(i=1,2,\ldots,m)$到销地$B_j(j=1,2,\ldots,n)$运输单位物品的运价是c_{ij}.

问怎样调运这些物品才能使总运费最小?

如果运输问题的总产量等于总销量，即有

$$\sum_{i=1}^{m} a_i = \sum_{j=1}^{n} b_j$$

则称该运输问题为产销平衡问题；反之，称产销不平衡问题。

■ 2.运输问题

令x_{ij}表示由产地A_i运往销地B_j的物品数量，则产销平衡问题的数学模型为：

$$\min z=\sum_{i=1}^{n}\sum_{j=1}^{m}c_{ij}x_{ij}$$

$$s.t\begin{cases}\sum_{j=1}^{n}x_{ij}=a_i & i=1,2,\cdots,m \\ \sum_{i=1}^{m}x_{ij}=b_j & j=1,2,\cdots n \\ x_{ij}\geq 0 & \begin{matrix} i=1,2,\cdots,m \\ j=1,2,\cdots n\end{matrix}\end{cases}$$

■ 基本概念

在上述例子中,

1. 目标函数和约束函数都是线性的, 称之为**线性规划问题**,
2. 模型中含有非线性函数, 称之为**非线性规划问题**.

在线性与非线性规划中,

1. 满足约束条件的点称为**可行点**,
2. 全体可行点组成的集合称为**可行集或可行域**.
3. 如果一个问题的可行域是整个空间, 则称此问题为**无约束问题**.

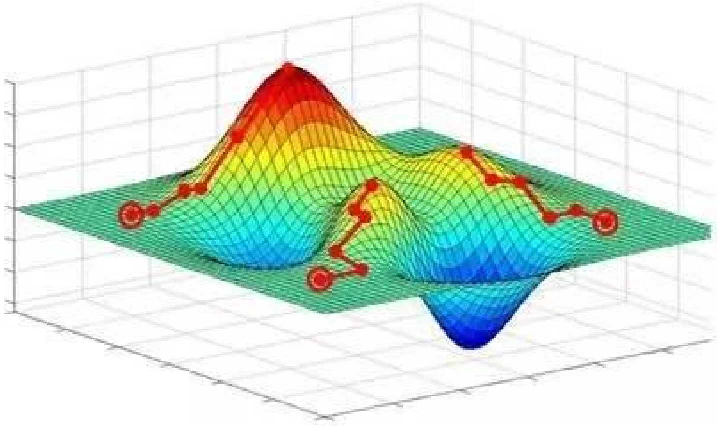
■ 基本概念

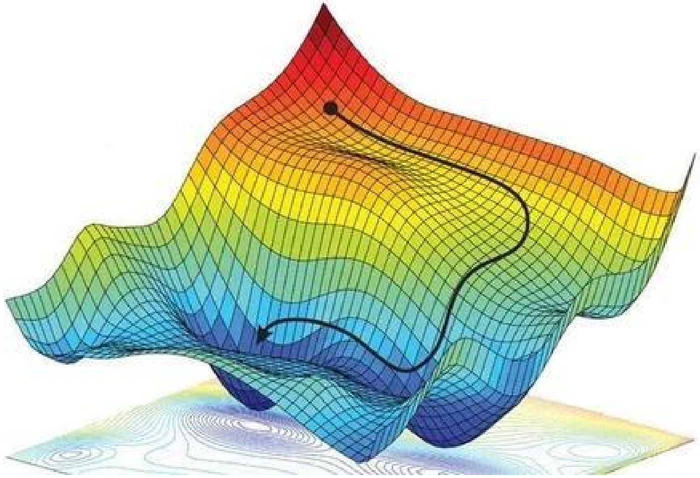
最优化问题可归结成如下数学形式：

$$\min_{x\in R^n} f(x) \qquad \text{---目标函数}$$

$$s.t. \quad g_i(x)\ge 0, i\in I$$

$$h_j(x)=0, j\in E$$





■ 基本概念

定义 1.1 设f(x)为目标函数，S为可行域，$x_0 \in S$,

若对每一个$x \in S$, 成立$f(x) \geq f(x_0)$, 则称 x_0 为极小化问题min f(x), $x \in S$的**最优解(整体最优解)**

(全局极小点)

定义 1.2 设f(x)为目标函数，S为可行域，

若存在x_0的ε邻域

$$N_\varepsilon(x_0) = \{x \mid \|x - x_0\| < \varepsilon, \varepsilon > 0\}$$

使得对每个$x \in S \cap N_\varepsilon(x_0)$, 成立$f(x) > f(x_0)$

则称 x_0为极小化问题min f(x), $x \in S$的**局部最优解**

(局部极小点)

■ 基本概念

- 对于极大化问题，可类似定义全局极大点和局部极大点;
- 全局极小点也是局部极小点;
- 局部极小点不一定是全局极小点;
- 某些条件下(如凸规划)，局部极小点也是全局极小点

1.*预备知识

- 线性空间
- 范数
- 集合与序列
- 矩阵的分解与校正
- 函数的可微性与展开

■ 1.1 线性空间

定义 1.3 给定一非空集合G以及在G上的一种代数运算 $+:G\times G\to G$(称为加法),若下述条件成立:

$$(1)\forall a,b,c\in G,\text{有}\,a+(b+c)=(a+b)+c$$

$$(2)\exists 0\in G,\text{使得}\,\forall a\in G,\text{有}\,a+0=0+a=a$$

$$(3)\forall a\in G,\exists -a\in G\,\text{使得}\,a+(-a)=(-a)+a=0$$

则<G,+>称为一个群.若还满足对任意的a,b∈G,有a+b=b+a,则<G,+>称为一个阿贝尔群(交换群)

1.1 线性空间

定义 1.4 给定一非空集合V和一个域F,并定义两种运算

加+:V×V→V以及数乘·: F×V→V.

若<V,+>构成一交换群,且两种运算满足下面性质:

$\forall a,b \in V, \forall \lambda,\mu \in F$以及单位元$1 \in F$,有

$$1 \cdot a = a$$

$$\lambda \cdot (\mu \cdot a) = (\lambda\mu) \cdot a$$

$$(\lambda + \mu) \cdot a = \lambda \cdot a + \mu \cdot a$$

$$\lambda \cdot (a + b) = \lambda \cdot a + \lambda \cdot b$$

则称V在域F上关于加法和数乘运算构成一**线性空间**,简称V为F上的线性空间.记为**V(F)**.若V的非空子集合S关于加法和数乘运算在F上也构成一线性空间,则S称为F上的**线性子空间**.

■ 1.1 线性空间

例子

1, R^n是实数域R上的一线性空间.

2, $R[x]_n$是系数在实数域R上次数小于n的全体多项式组成的集合, 则$R[x]_n$关于多项式的加法以及数与多项式的乘法构成一线性空间.

3, $R^{m\times n}$是实数域R上所有m×n矩阵组成的集合, 则其关于矩阵加法和数乘运算构成一线性空间.

大连理工大学 Dalian University Of Technology

■ 1.2 范数

定义 1.5 若函数$\|\cdot\|:\mathrm{R}^n\to\mathrm{R}$满足：

(1)正定性：$\forall x\in R^n,\|x\|\geq 0,\|x\|=0\Leftrightarrow x=0;$

(2)三角不等式：$\forall x,y\in R^n,\|x+y\|\leq\|x\|+\|y\|;$

(3)齐次性：$\forall x\in R^n,\forall\alpha\in R,\|\alpha x\|=|\alpha|\|x\|.$

则$\|\cdot\|$称为R^n上的范数

例子：$x=(x_1,x_2,\ldots,x_n)^T\in R^n$

2－范数 $\|x\|_2=(\sum_{i=1}^{n}x_i^2)^{1/2}$；1－范数 $\|x\|_1=\sum_{i=1}^{n}|x_i|$；

∞－范数 $\|x\|_\infty=\max_i|x_i|$；p－范数 $\|x\|_p=(\sum_{i=1}^{n}x_i^p)^{1/p},1\leq p<+\infty;$

■ 1.3 集合与序列

$x^0 \in \mathrm{R}^n$的$\varepsilon-$邻域：

$$N_\varepsilon(x^0)=\{x \mid \|x-x^0\|<\varepsilon\}$$

$x^0 \in S \subseteq R^n$称为S的内点，若$\exists \varepsilon>0$,使得$N_\varepsilon(x^0)\subseteq S$.

集合S的所有内点的全体称为S的内部,记为int(S)或$\mathring{\mathrm{S}}$.

若集合$S=\mathrm{int}(S)$,则S称为一个开集；

集合S的补集$\mathrm{S}^{\mathrm{C}}=\{x \mid x \notin S . x \in R^n\}$;

若集合S的补集S^{C}是开集，则称S为闭集；

集合S的闭包是指包含它的最小闭集，记为clS或$\overline{\mathrm{S}}$

集合S的边界定义为集合$\partial \mathrm{S}=\mathrm{clS} \cap cl\mathrm{S}^{\mathrm{C}}$

1.3 集合与序列

集合$S \subseteq R^n$称为有界集，若$\exists M > 0$,使得$\forall x \in S, \|x\| \le M$

R^n中的有界闭集也称为紧集。

非空集合$S \subseteq R^n$，若存在一实数$\alpha \in R$，使得$\forall x \in S, x \le \alpha$成立

则α称为S的一个上界，若存在一实数$\beta \in R$，使得$\forall x \in S, x \ge \beta$成立

则β称为S的一个下界。

S的上确界$\sup(S)$是指它的最小上界,其下确界$\inf(S)$是指它的最大下界.

给定有界的实数序列$\{x_k\}$,对$\forall n \in N$,令

$\alpha_n = \sup\{x_k \mid k \ge n\}, \beta_n = \inf\{x_k \mid k \ge n\}$,

则序列$\{\alpha_n\}$和$\{\beta_n\}$,分别是单调递减和单调递增的

实数序列，且$\alpha_n \ge \beta_n$.从而都有极限，分别称为

上极限和下极限，记为

$$\limsup_{k\to\infty} x_k = \lim_{n\to\infty} \alpha_n, \liminf_{k\to\infty} x_k = \lim_{n\to\infty} \beta_n$$

■ 1.4 矩阵的分解与校正

*Th*1.4 给定矩阵A $\in R^{n\times n}$，A的顺序主子式

$$\begin{pmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & & \vdots \\ a_{k1} & \cdots & a_{kk} \end{pmatrix}$$

都可逆,则存在唯一的单位下三角矩阵L (主对角元全为1)和可逆的上三角矩阵U，使得A=LU (称为A的LU-分解)

Th1.5 若n阶矩阵A可逆，则存在一个排列矩阵P，单位下三角矩阵L和上三角矩阵U，使得PA=LU

■ 1.4 矩阵的分解与校正

定理 1.3 设A为对称正定矩阵，则

（1）矩阵A可唯一的分解成$A=LDL^T$, 其中L为单位下三角矩阵，D为对角矩阵

（2）存在可逆的下三角矩阵L，使得$A=LL^T$. 当L的对角元素为正时，分解是唯一的。(**Cholesky分解**)

*Th*1.7　设$A\in R^{n\times n}$可逆, $B\in R^{m\times m}$, $U,V\in R^{n\times m}$.若矩阵$B-V^TA^{-1}U$

是可逆的，则分快矩阵$\begin{pmatrix} A & U \\ V^T & B \end{pmatrix}$可逆，且

$$\begin{pmatrix} A & U \\ V^T & B \end{pmatrix}^{-1} = \begin{pmatrix} A^{-1}+A^{-1}UGV^TA^{-1} & -A^{-1}UG \\ -GV^TA^{-1} & G \end{pmatrix}$$

其中$G=(B-V^TA^{-1}U)^{-1}$.

■ 1.5 函数的可微性与展开

设$f:R^n\to R$是连续函数,对于$x^0\in R^n$和单位向量

$e^i=(\delta_1^i,\delta_2^i,\dots,\delta_n^i)^T,\delta_i^i=1,\delta_j^i=0(i\neq j)$,一元函数$f(x^0+te^i)$

在$t=0$的导数(若存在的话)称为f在点x^0关于x_i的一阶偏导数.

$(i=1,2,\dots,n)$

对于任意$i=1,2,\dots,n$,若f在点x关于x_i的一阶偏导数存在,则称

$f(x)$在x点存在一阶偏导数此时,$f(x)$在x点的梯度定义为

$$\nabla f(x)=\left(\frac{\partial f(x)}{\partial x_1},\dots,\frac{\partial f(x)}{\partial x_n}\right)^T.$$

对于$x^0\in R^n,p\in R^n,p\neq 0$,函数$f$在点$x^0$关于方向$p$的方向导数定义为:

$$\frac{\partial f(x^0)}{\partial p}=\lim_{t\to 0^+}\frac{f(x^0+tp)-f(x^0)}{t}.$$

■ 1.5 函数的可微性与展开

我们也用$Df(x^0;p)$表示f在点x^0关于方向p的方向导数.当f的一阶偏导连续时有 $Df(x^0;p)=\nabla f(x^0)^T p$

当f(x)在x点存在二阶偏导时,函数f在点x的**Hesse矩阵**定义为

$$\nabla^2 f(x)=\begin{pmatrix} \dfrac{\partial^2 f(x)}{\partial x_1^2} & \cdots & \dfrac{\partial^2 f(x)}{\partial x_1 x_n} \\ \dfrac{\partial^2 f(x)}{\partial x_2 x_1} & \cdots & \dfrac{\partial^2 f(x)}{\partial x_2 x_n} \\ \vdots & & \vdots \\ \dfrac{\partial^2 f(x)}{\partial x_n x_1} & \cdots & \dfrac{\partial^2 f(x)}{\partial x_n^2} \end{pmatrix} \quad \left(\text{其中}\frac{\partial^2 f(x)}{\partial x_i x_j}=\frac{\partial}{\partial x_j}\left(\frac{\partial f(x)}{\partial x_i}\right)\right)$$

1.5 函数的可微性与展开

$Th1.9$ $(Taylor)$ 设 $f: R^n \to R$ 连续可微,向量 $p \in R^n$,则

$$f(x+p) = f(x) + \int_0^1 \nabla f(x+tp)^T p dt = f(x) + \nabla f(x+\xi p)^T p, \xi \in (0,1)$$

$$= f(x) + \nabla f(x)^T p + o(\|p\|)$$

进而,若 f 二阶连续可微,则

$$\nabla f(x+p) = \nabla f(x) + \int_0^1 \nabla^2 f(x+tp)^T p dt = \nabla f(x) + \nabla^2 f(x+\xi p) p, \xi \in (0,1)$$

$$= \nabla f(x) + \nabla^2 f(x) p + o(\|p\|)$$

$$f(x+p) = f(x) + \nabla f(x)^T p + \int_0^1 (1-t) p^T \nabla^2 f(x+tp)^T p dt$$

$$= f(x) + \nabla f(x)^T p + \frac{1}{2} p^T \nabla^2 f(x+\xi p) p, \xi \in (0,1)$$

$$= f(x) + \nabla f(x)^T p + \frac{1}{2} p^T \nabla^2 f(x) p + o(\|p\|^2)$$

■ 1.5 函数的可微性与展开

对向量值函数$F=(f_1,f_2,\ldots,f_n):R^n\to R^m$,若每个分量函数$f_i$是(连续)可微的，则称函数$F$是(连续)可微的。向量函数$F$在$x$的导数$F'\in R^{m\times n}$是指它在$x$点的*Jacobi*矩阵，记为$F'(x)$或$J_F(x)$.为与标量函数梯度对应，定义*Jacobi*矩阵的转置为F在x点的梯度，记为

$$\nabla F(x)=J_F(x)^T=(\nabla f_1,\nabla f_2,\ldots,\nabla f_n)$$

类似地，设$F:R^n\to R^m$连续可微,$\forall x,p\in R^n$,则

$$F(x+p)=F(x)+\int_0^1 J_F(x+tp)^T\,p\,dt$$

■ 1.5 函数的可微性与展开

*Df*1.12 给定映射$G:R^n\to R^m$,点$x\in R^n$,若存在常数L>0, 使对任意$y\in R^n$,下式成立： $\|G(x)-G(y)\|\le L\|x-y\|$

则称G在x是Lipschitz连续的，L称为Lipschitz常数。若上式对$\forall x, y\in R^n$成立，则称G在R^n内Lipschitz连续。

*Th*1.10设向量值函数$F:R^n\to R^m$连续可微, 则对$\forall$ u, v,$x\in R^n$, 有

$$\|F(u)-F(v)-J_F(x)(u-v)\|\le \sup_{t\in[0,1]}\|J_F(v+t(u-v))-J_F(x)\|\|u-v\|$$

进而，若Jacobi矩阵映射J_F在R^n内是Lipschitz 连续的，记L 为Lipschitz常数，则

$$\|F(u)-F(v)-J_F(x)(u-v)\|\le L\frac{\|u-x\|+\|v-x\|}{2}\|u-v\|$$

2.*凸集与凸函数

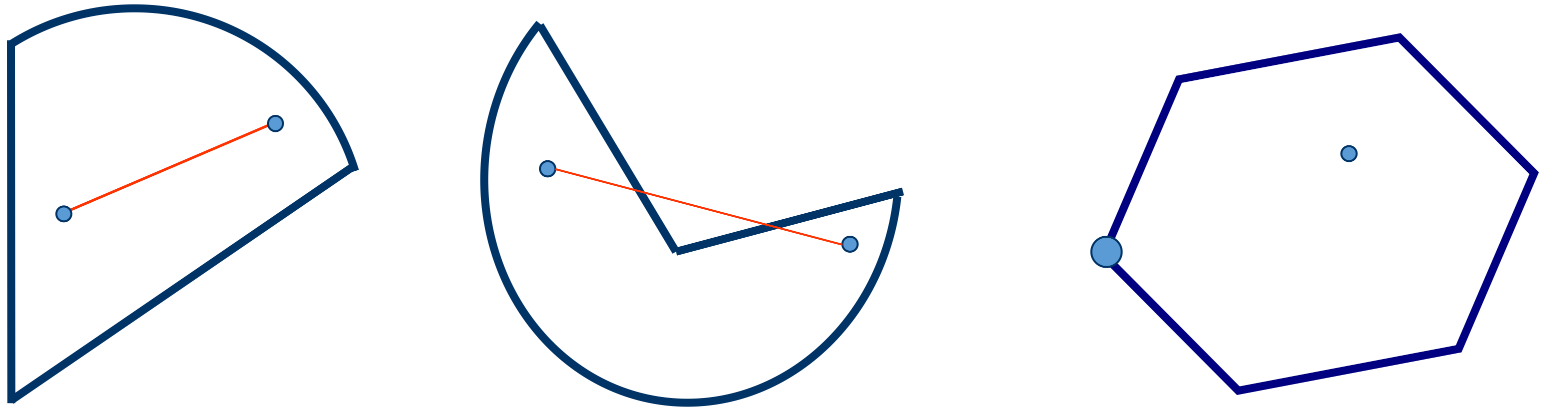
- 凸集与锥
- 凸函数
- 凸规划

■ 2.1 凸集与锥

Df 2.1 设S为n维欧氏空间R^n中的一个集合。若对任意两点$x^{(1)}, x^{(2)} \in S$及每个实数$\lambda \in [0,1]$,有

$$\lambda x^{(1)} + (1-\lambda)x^{(2)} \in S$$

则称S为凸集。$\lambda x^{(1)} + (1-\lambda)x^{(2)}$称为$x^{(1)}$和$x^{(2)}$的凸组合。



2.1 凸集与锥

例2.1 超平面$H=\{x|p^T x=\alpha\}$为凸集，其中p为n维列向量，α为实数。此外,下面相对于法向量p的半空间都是凸集:

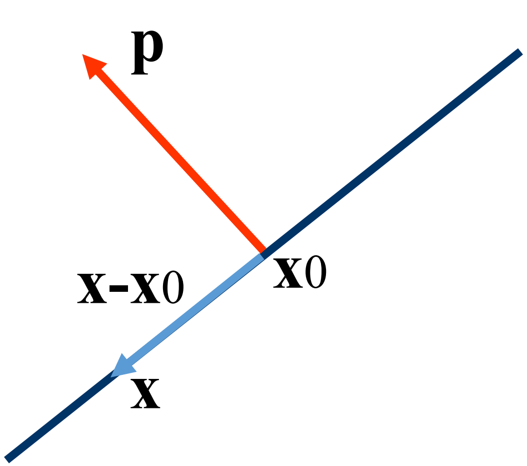
正的闭半空间$H^+=\{x|p^T x\ge\alpha\}$

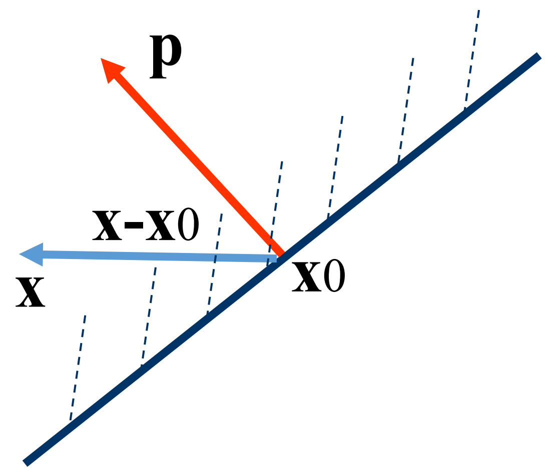
负的闭半空间$H^-=\{x|p^T x\le\alpha\}$

正的开半空间$\dot{H}^+=\{x|p^T x>\alpha\}$

负的开半空间$\dot{H}^-=\{x|p^T x<\alpha\}$

■ 2.1 凸集与锥





例2.2 集合$L=\{x \mid x=x^{(0)}+\lambda d，\ \lambda \geq 0\}$为凸集，其中 d为给定的非零向量，$x^{(0)}$为定点。

集合$L=\{x \mid x=x^{(0)}+\lambda d, \lambda \geq 0\}$称为射线，$x^{(0)}$为射线的顶点

大连理工大学 Dalian University Of Technology

■ 2.1 凸集与锥

Df 2.2 给定m个向量，$x^1,\dots,x^m\in R^n$，以及满足$\sum_{i=1}^{m}\lambda_i=1$的非负实数$\lambda_i\in R, i=1,..,m$，称向量$\lambda_1 x^1+\dots+\lambda_m x^m$为$\{x^1,\dots,x^m\}$的凸组合.

*Th*2.1 集合$S\subseteq R^n$是凸集，当且仅当S包含其中任意有限个元素的凸组合，即对$\forall m\in R^+=\{1,2,\dots\}$，任意的$x^1,\dots,x^m\in R^n$，有$\lambda_1 x^1+\dots+\lambda_m x^m\in S$，其中$\sum_{i=1}^{m}\lambda_i=1, \lambda_i\geq 0\in R, i=1,..,m$.

2.1 凸集与锥

运用定义不难验证如下命题：

命题2.1 设S_1和S_2为R^n中两个凸集,β是实数,则

1，$\beta S_1=\{\beta x \mid x\in S_1\}$为凸集。

2，$S_1\cap S_2$为凸集

3，$S_1+S_2=\{x^{(1)}+x^{(2)} \mid x^{(1)}\in S_1, x^{(2)}\in S_2\}$ 为凸集

4，$S_1-S_2=\{x^{(1)}-x^{(2)} \mid x^{(1)}\in S_1, x^{(2)}\in S_2\}$ 为凸集

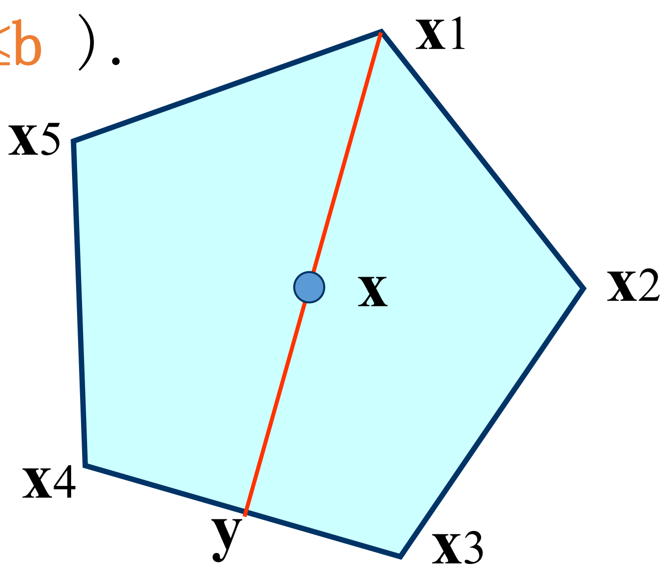
■ 2.1 凸集与锥

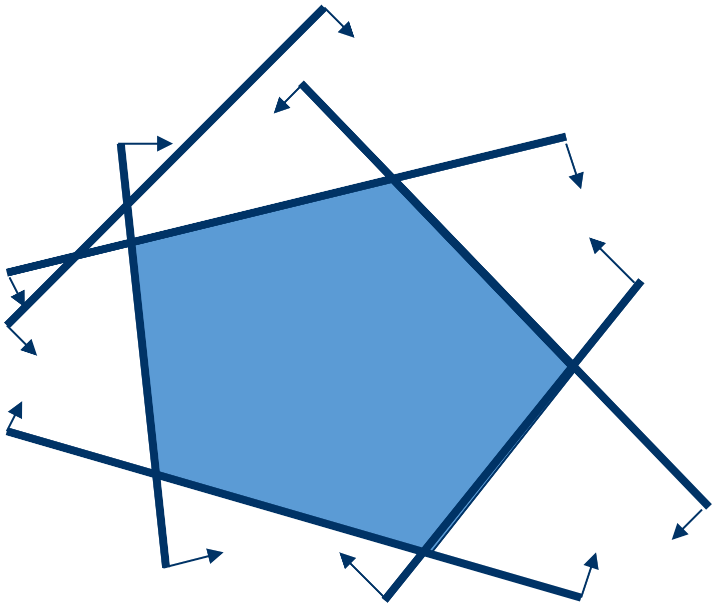
有限点集$\{x^0,x^1,\dots,x^m\}\subset R^n$的凸包称为多胞形。

若$\{x^0,x^1,\dots,x^m\}$仿射无关时，对应的凸包称为m维单纯形。

向量x^i称为该单纯形的顶点。

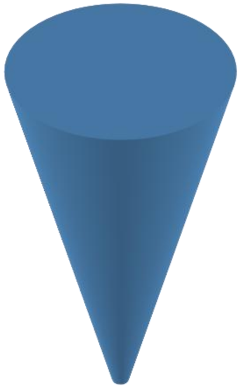
多面体(polyhedral set)是有限闭半空间的交. (可表为$Ax\le b$).





■ 2.1 凸集与锥

Df 2.4设有集合$C \subset R^n$,若对每一点$x \in C$,当λ取任何非负数时，都有$\lambda x \in C$,称C为锥,又若C为凸集，则称C为凸锥.



例2.3，向量集$\alpha^{(1)},\alpha^{(2)},\ldots,\alpha^{(k)}$的所有非负线性组合构成的集合 $\{\sum_{i=1}^{k}\lambda_i\alpha^{(i)} \mid \lambda_i \geq 0, i=1,2,\ldots,k\}$为凸锥。

■ 2.1 凸集与锥

由定义可知, 锥关于正的数乘运算封闭, 凸锥关于加法和正的数乘封闭, 一般的, 对于凸集S, 集合

$$K(S)=\{\lambda x \mid \lambda>0,\ x\in S\}$$

是包含S的最小凸锥.

锥C称为尖锥, 若0∈S. 尖锥称为突出的, 若它不包含一维子空间.

多面集 $\{x \mid Ax\le 0\}$ 也是凸锥，称为**多面锥**。

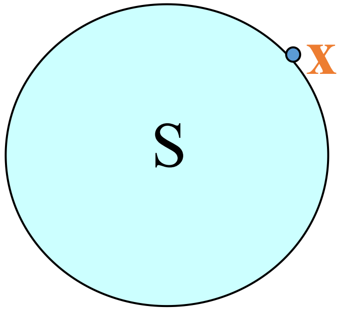
约定：非空集合S生成的凸锥, 是指可以表示成S中有限个元素的非负线性组合(称为凸锥组合)的所有点所构成的集合, 记为coneS. 若S凸, 则

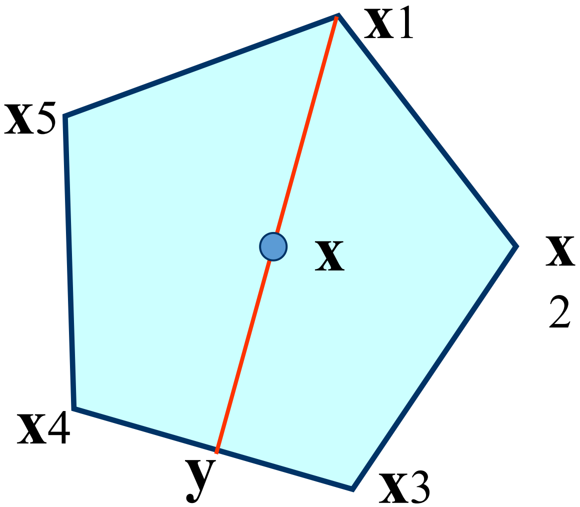
$$\text{cone}S=K(S)\cup\{0\}$$

■ 2.1 凸集与锥

Df 2.5 非空凸集中的点 $\mathbf{x}$ 称为**极点**,若 $\mathbf{x}=\lambda\mathbf{x}_1+(1-\lambda)\mathbf{x}_2$, $\lambda\in(0,1)$, $\mathbf{x}_1, \mathbf{x}_2 \in \mathbf{S}$, 则 $\mathbf{x}=\mathbf{x}_1=\mathbf{x}_2$.

换言之,x不能表示成S中两个不同点的凸组合.





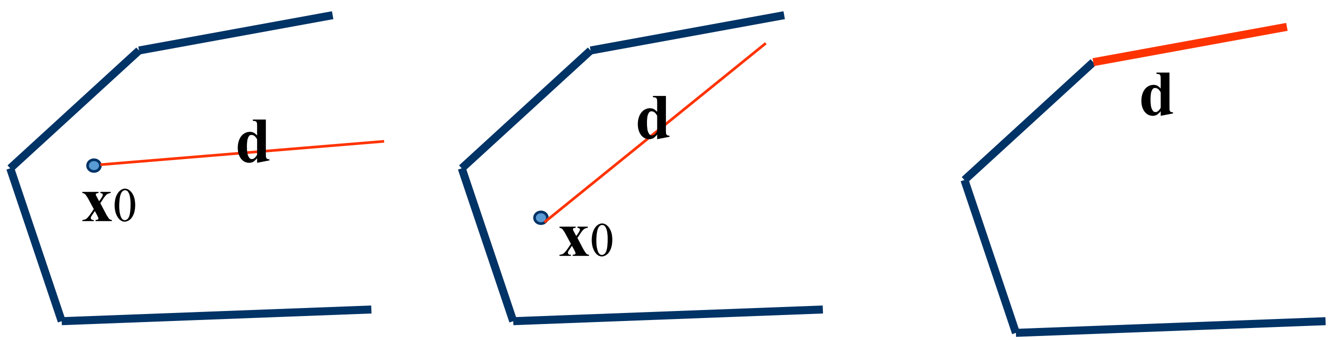
由上可知, 任何有界凸集中任一点都可表成极点的凸组合.

■ 2.1 凸集与锥

Def 2.6. 设非空凸集$S\subset R^n$，R^n中向量$d\neq 0$ 称为S的一个**回收方向(方向)**，若对每一 $x\in S$，$R(x.d)=\{x+\lambda d \mid \lambda\geq 0\}\subset S$. S的所有方向构成的尖锥称为S的回收锥，记为$0^+S$

方向d_1和d_2 称为S的两个**不同的方向**，若对任意$\lambda>0$，都有$d_1\neq\lambda d_2$；方向d称为S的极方向extreme direction，若$d=\lambda d_1+(1-\lambda)d_2$，$\lambda\in(0,1)$，$d_1$，$d_2$ 是S的两个方向，则有$d=d_1=d_2$.

换言之d不能表成它的两个不同方向的凸锥组合

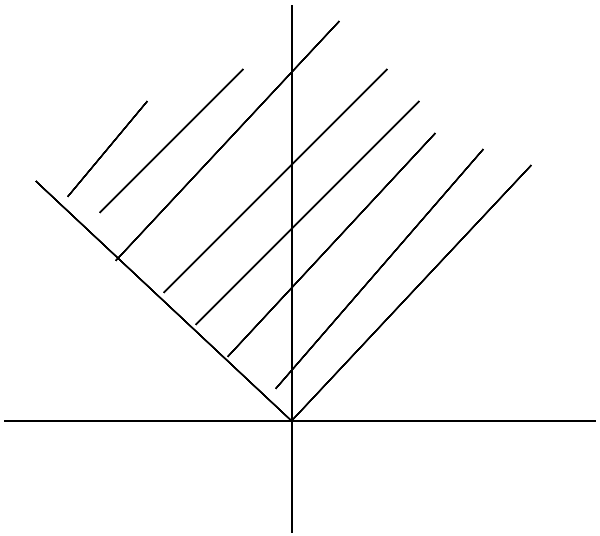


大连理工大学 Dalian University Of Technology

■ 2.1 凸集与锥

例2.4 集合$S=\{(x_1,x_2)\,|\,x_2\geq|x_1|\}$

凡是与向量$(0,1)^T$夹角$\leq 45°$的向量都是它的方向。$(1,1)^T,(-1,1)^T$是其仅有的两个极方向



例2.5 设$S=\{x\,|\,Ax=b,x\geq 0\}\neq\phi$, d是非零向量。证明, d是S的方向$\Leftrightarrow d\geq 0$且$Ad=0$.

Th2.3 若多面体P的极点(极方向)存在的话, 则极点(极方向)的数目一定有限.

■ 2.2 凸集分离定理

Df 2.7，设S_1和S_2是R^n中两个非空集合，$H=\{x \mid p^T x=\alpha\}$为超平面。若对$\forall x\in S_1$,有$p^T x\geq\alpha$，对于$\forall x\in S_2$,有$p^T x\leq\alpha$（或情形恰好相反），则称超平面$H$分离集合$S_1$和$S_2$.若$S_1\cup S_2\not\subset H$,则称$H$正常分离$S_1$和$S_2$。若$S_1\subseteq\dot{H}^+$,$S_2\subseteq\dot{H}^-$,则称$H$严格分离$S_1$和$S_2$。若$S_1\subseteq H(\varepsilon)^+=\{x \mid p^T x\geq\alpha+\varepsilon\},\varepsilon>0,S_2\subseteq\dot{H}^-$,则称$H$强分离$S_1$和$S_2$。

■ 2.2 凸集分离定理

$Df\,2.8$，设$S(\neq\phi)\subset R^n$, $p\in R^n$，$p\neq 0, \bar{x}\in\partial S$

若$S\subseteq H^+=\{x\,|\,p^T(x-\bar{x})\geq 0\}$或者$S\subseteq H^-=\{x\,|\,p^T(x-\bar{x})\leq 0\}$

则称$H=\{x\,|\,p^T(x-\bar{x})=0\}$是$S$在$\bar{x}$处的支撑超平面，若 $S\not\subset H$,

则称H为S在$\bar{x}$处的正常支撑超平面。

$Df\,2.9$　设$S\subseteq R^n$非空，$y\in R^n$,则点y与集合S之间的距离$dist(y,S)$定义为

$$dist(y,S)=\inf_{x\in S}\|y-x\| \qquad (2.4)$$

Th2.5设S为R^n中的闭凸集，$y\notin S$, 则存在唯一的点$\bar{x}\in S$,使得$\|y-\bar{x}\|=\inf_{x\in S}\|y-x\|$

■ 2.2 凸集分离定理

$Th2.6$ 设$S \subseteq R^n$的非空闭凸集，$y \notin S$,则点$\bar{x} \notin S$为极小化问题(2.4)的最优解当且仅且当$(y-\bar{x})^T(\bar{x}-x) \geq 0$

设S为闭凸集，$y \notin S, H=\{x \mid p^T x=\alpha\}$为超平面。

H分离点$y \Rightarrow$若$p^T y>\alpha$, 则$p^T x \leq \alpha$，$\forall x \in S$.

令$p^T y-\alpha=\varepsilon$，则y与S分离可表为

$$p^T y \geq \varepsilon+p^T x, \forall x \in S.$$

$Th2.7$ 设$S(\neq \Phi)$为R^n中的闭凸集，$y \notin S$, 则存在$p \neq 0$及实数$\varepsilon>0$，使得对点$x \in S$, 有 $p^T y \geq \varepsilon+p^T x$。

2.2 凸集分离定理

Th2.8 设$S(\neq \Phi)$为R^n中的凸集，$y\in \partial S$, 则存在$p\neq 0$，使得对点$x\in clS$, 有 $p^T y \geq p^T x$。

Th2.9. 设$S_1, S_2(\neq \Phi)$为R^n中的凸集, $S_1 \cap S_2 = \Phi$, 则存在$p\neq 0$使 $\inf\{ p^T x \mid x\in S_1\} \geq \operatorname{Sup}\{p^T x \mid x\in S_2\}$

(换言之，存在超平面H, 使得$S_1 \subseteq H^+$, $S_2 \subseteq H^-$)。

Th2.10. 设$S_1, S_2(\neq \Phi) \subseteq R^n$中的闭凸集, S_1有界，$S_1 \cap S_2 = \Phi$, 则存在超平面H强分离S_1和S_2，即存在$p\neq 0$，$\varepsilon > 0$使

$$\inf\{ p^T x \mid x\in S_1\} \geq \varepsilon + Sup\{p^T x \mid x\in S_2\}$$

■ 2.3 择一定理

作为凸集分离定理的应用，下面介绍两个择一定理：Farkas定理和Gordan定理，它们在最优化理论中是很有用的。

定理2.11（*Farkas*定理）

设A为$m\times n$矩阵，c为n维向量，则

$Ax\le 0$，$c^T x>0$有解的充要条件是，$A^T y=c, y\ge 0$无解.

2.3 择一定理

定理2.11′（Farkas置换定理）

设$A\in R^{m\times n}$，$c\in R^n$，则c为A的行向量的凸锥组合，

即$c\in \text{cone}A \Leftrightarrow$ 对任意满足$Ax\le 0$的向量$x\in R^n$，都有$c^T x\le 0$。

推论（$Gale$置换定理）设A为$m\times n$矩阵，b为m维向量，则线性系统$Ax\le b$有解$\Leftrightarrow$ 对任意满足$A^T w=0, w\ge 0$，的向量$w\in R^m$,都有$b^T w\ge 0$.

定理2.12（Gordan定理）

设A为m×n矩阵，那么，$Ax<0$有解的充要条件是不存在非零向量$y\ge 0$，使$A^T y=0$.

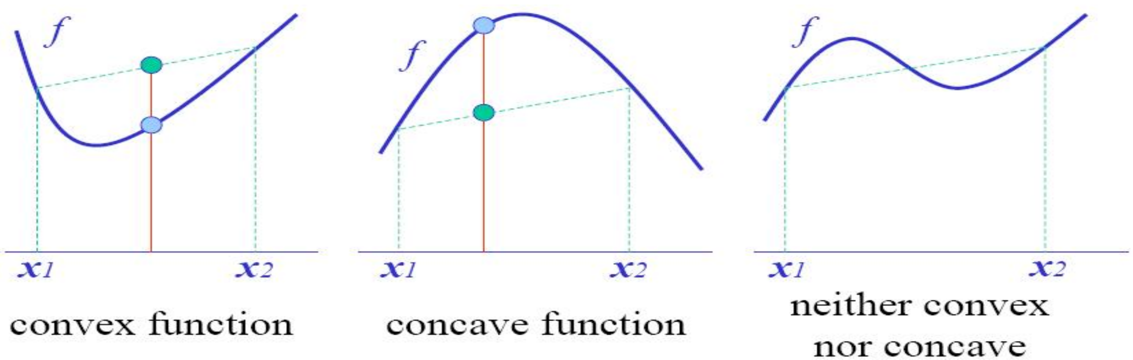
■ 2.4 凸函数

Df 2.10 设S⊂R^n是非空凸集,函数f:S→R,若对任意$\mathbf{x}_1$, $\mathbf{x}_2$∈**S**,和每一λ∈(0, 1)都有

$$f(\lambda \mathbf{x}_1+(1-\lambda)x_2) \le \lambda f(\mathbf{x}_1)+(1-\lambda)f(\mathbf{x}_2)$$

则称f是S上的**凸函数**.若上面的不等式对于x≠y严格成立,则称f是S上的**严格凸函数**.

若**-f**是**S**上的凸函数,则称f是S上的**凹函数**.若**-f**是**S**上的严格凸函数,则称**f**是**S**上的**严格凹函数**.



■ 2.4 凸函数

命题2.3 设f是定义在凸集S上的凸函数，则

(1)所有凸函数f的集合关于凸锥组合运算是封闭的,

即(a)实数$\lambda \geq 0$，则λf也是定义在S上的凸函数(b)设f_1和f_2是定义在凸集S上的凸函数，则f_1+f_2也是定义在S上的凸函数

(2)函数f在开集intS内是连续的.

(3)函数f的水平集$L(f,\alpha)=\{x|x\in S, f(x) \leq \alpha\}, \alpha \in R$

和上镜图$epi(f)=\{(x,y)|x\in S, y\in R, y\geq f(x)\}$都是凸集

■ 2.4 凸规划

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & g_i(x) \geq 0, i = 1,2,\ldots,m \\ & h_j(x) = 0, j = 1,2,\ldots,l \end{aligned}$$

$f(x)$是凸函数，$g_i(x)$是凹函数，$h_j(x)$是线性函数。可行域

$$S = \left\{ x \middle| \begin{array}{l} g_i(x) \geq 0, i = 1,2,\ldots,m \\ h_j(x) = 0, j = 1,2,\ldots,l \end{array} \right\}$$

小结和作业

小结

- 课程简介
- 预备知识
- 凸集和凸函数
- 最优化的应用

■ 作业

- 习题 2、4、9、10、12、14

- 线性规划的应用
  https://haokan.baidu.com/v?vid=3203660548000343532&pd=bjh&fr=bjhauthor&type=video

- 常见的几个凸函数与凹函数