第一章 引论

1.1 最优化方法举例

解决问题步骤：1.建立数学模型 2.进行数学加工和求解

举例 1. 化工产品生产 2.运输问题 3. 拟定生产计划 4. 合理下料问题 5. 汽轮机叶片的优化设计 6. 船体放样 7. 投资决策问题

1.2 最优化的基本概念

分类 1. 间接最优化。将问题用数学描述，然后求解

2. 直接最优化。通过少量试验，根据实验结果的比较求最优解

1.2.1 最优化问题的提法和基本概念

在决策变量下，满足不等式约束条件和等式约束条件，使得目标函数取极小值。

最优解/全局最优解： 存在x\*使得对任意x有f(x\*)<=f(x)

严格… <

局部… 的一个邻域 <=

严格局部… <

定理2.1 若约束条件和目标函数连续，则可行集是闭集，全局最优解集合也是闭集

定理2.2 一阶必要条件 极小点梯度为零。梯度为对每个xi求偏导

梯度为零的点称稳定点或驻点

1.2.2 二维最优化问题的几何意义

f(x1,x2)是曲面 f(x1,x2)=C是等高线

图解法： 1. 画出可行集R的图形

2. 常数C取一系列值做目标函数的等值线族

3. 观察等值线族与可行集R,确定使目标函数取最小值的可行点，即为最优解。

1.2.3 最优化问题分类

1. 线性规划LP。 都是线性函数

2. 二次规划QP。 f(x)是二次函数

3. 非线性函数NLP。 至少有一个是分线性函数。 无约束非线性规划/约束非线性规划

4. 整数规划IP。 设计变量取非负的整数值。混合型整数规划问题。0-1规划问题

5. 几何规划GP。 目标函数和约束函数是设计变量的广义多项式。正项/符号

6. 多目标规划MP。 X是一个向量函数

1.3 凸集和凸函数

1.3.1 凸集

定义3.1 凸集：连接Ω内任意两点的线段上一切点都在Ω中。空集∅为凸集

定义3.2 凸组合：α>=0 Σα = 1，Σαix是凸组合

定理3.1 凸集充要条件：x的任意凸组合仍包含在Ω中

定理3.2 任意一组凸集的交集仍是凸集

定义3.3 凸包：包含集合Ω的所有凸集的交集

定义3.4 锥：对任意x和所有α，都有αx∈Ω

1.3.2 凸函数

定义3.5 凸函数：f(x)是凸集Ω上的函数，f(λx+(1-λ)y)<= λf(x)+(1-λ)f(y)对于[0,1]的λ成立

严格凸函数：… < …

几何意义：图像上任意两点的连线处处在图像上方

定理3.3 f1(x)，f2(x)是凸函数，f1(x)+f2(x)也是凸函数

定理3.4 f(x)是凸函数，则af(x)也是凸函数

推论：凸函数的非负线性组合也是凸函数

定理3.5 f(x)是凸函数，f(x)<=c时c的集合是凸集

定理3.6 充要条件：可微函数f(x)是凸集⬄f(y)>=f(x)+(▽f(x))T(y-x)

定理3.7 充要条件：f(x)为凸函数⬄f(x)的Hesse阵F(x)=▽2f(x)在整个Ω上是半正定的

定理3.8 ▽2f(x)在Ω上处处正定=>f(x)是严格凸函数

1.3.3 凸规划

前提：等式约束可以写成不等式约束

定义3.6 凸规划问题： f(x)、g(x) 均为可行集上的凸函数

定理3.9 凸规划问题的可行集是凸集

定理3.10 目标函数的任意局部极小值都是非空可行集上的全局极小值

定理3.11 f(x)在非空可行集上是严格凸函数，则全局极小点是唯一的

1.3.4 拟凸函数

定义3.7 拟凸函数：f(λx+(1-λ)y) <=max{f(x),f(y)}

严格拟凸函数 <

严格拟凸函数不一定是拟凸函数

命题： 凸函数必定为拟凸函数，反之则不然。

定理3.12 充要条件： f(x)是拟凸函数⬄Hc(f) = {x|x∈Ω，f(x)<=c}是凸集

定理3.13 充要条件： f是拟凸函数⬄若f(x(1))<=f(x(2)) 则f(x(2))T(x(1)-x(2))

1.3.5 伪凸函数

定义3.8: 伪凸函数： 如果▽f(x(1))T(x(1)-x(2))<=0 必有f(x(1))<=f(x(2))

严格伪凸函数 <

可微的凸函数时伪凸函数。 严格伪凸函数必是伪凸函数

定理3.14 若f是Ω上的伪凸函数，则f是Ω上的严格拟凸函数和拟凸函数

1.3.6 广义凸规划

广义凸规划： 可行集R是凸集，f(x)是R上的(严格)拟凸函数或(严格)伪凸函数

定理3.15 可行集R是凸集，f(x)是R上的严格拟凸函数，则广义凸规划任一局部最优解也是全局最优解

定理3.16 可行集R是开凸集，f(x)是R上的伪凸函数，若▽f(x\*)=0，则x\*是全局最优解

第二章 线性规划

线性规划：目标函数为决策变量的线性函数和约束条件为线性等式或不等式组成的数学规划

2.1 引言 线性规划的标准形式

标准形式： min CTx

s.t. Ax=b, x>=0

A=(aij)m\*n )x=(x1,x2,…,xn)T rank(A)=m<=n c=(c1,c2,…,cn)T b=(b1,b2,…,bm)T b>=0

通过松弛变量和剩余变量转化为等式约束

维数：变量的个数

阶数：等式约束方程的数目

可行点/可行解/容许解：满足约束条件的解

A的秩为m，则A的n列有m列是线性无关的。Bm\*m是非奇异矩阵，BxB=b。得到一个解xB

基或基底：B

基本解：x

基本变量：与B的列对应的x的分量xi

退化的基本解：基本解中有一个或多个基本变量为零。

基本可行解：x既是基本解，又是可行解

基本可行解不超过m个正xi的可行解

费退化的基本可行解x是恰有m个正xi的可行解

m阶n维的线性规划问题，基本可行解个数不超过CmnC

最优可行解/最优解/解：使目标函数达到最大（小）值时的可行解

2.2 线性规划的基本定理

定义2.1 正闭半空间：H+={x|x∈Rn,cTx>=b}

负闭半空间：H-={x|x∈Rn,cTx<=b}

定义2.2 多面凸集/凸多胞体：有限个闭半空间的集合，即集合S={x|Ax<=b}

定义2.3 多面凸体/凸多面体：有界且非空的多面凸集

定义2.4 极点/顶点：S为凸集，x∈S，不存在x1,x2∈S,使得x=αx1+(1-α)x2成立

定理2.1 充要条件： x是凸集 ⬄ x是Ax=b的基本可行解

定理2.2 线性规划的基本定理 若线性规划由可行解，则有基本可行解。

有最优解，则有最优的基本可行解

推论1 可行集是非空的，则至少有一个极点

推论2 线性规划问题存在有限的最优解，则至少存在一个可行集R的极点是有限最优解

推论3 可行集R至多有有限多个极点

2.3 单纯形法

枚举法：求出所有基本可行解，进行比较。当m和n较大时，可行解数量过大，无法枚举。

单纯形法：使一个基本可行解转移到另一个基本可行解的规则

约束方程式的规范形式

X1 +y1m+1xm+1+…+y1nxn=y10

X2 +y2m+1xm+1+…+y2nxn=y20

…

Xm+ymm+1xm+1+…+ymnxn=ym0

简写 Xi+Σyijxj=yi0

基本解为 x = (y10,y20,…,ym0,0,…,0)T

2.3.3 基本可行解的转换

非退化假定：假定Ax=b的每个基本可行解都是非退化的基本可行解。

2.3.4 最优基本可行解的确定

定理3.1 若cj-zj<0 则目标函数值可以继续减小，

定理3.2 某个基本可行解，对所有j，都有cj-zj>=0，则这个解是最优解

2.3.5 单纯形法的计算步骤和例子

1. 把一般的线性规划问题化为标准形式

2. 建立初始单纯形表

3. 若所有的检验数rj=cj-zj>=0，就得到了一个最优解，运算结束；否则转到第四步

4. 当有多个r<0时，可选其中任一列的矢量aj为进基矢量，通常选最小的ak为进基矢量

5. 对所有的yik>0,计算比值yi0/yik，主元为yrk,当A=(Im,N)时，xr为离基变量

若所有的yik<=0，则可行集是无界的，目标函数是无界的

6. 以yrk为主元，进行gauss消元，求得一个新的基本可行解。

2.4 关于单纯形法的说明和补充

2.4.1 初始基本可行解的确定

1. 大M法

定理4.1 为了求标准形式的基本可行解，可以直接去求大M法的基本可行解

2. 二阶段法

定理4.2 为了求标准形式的基本可行解，可以先求辅助问题，根据它的解得到问题的基本可行解

2.4.2 单纯形法的矩阵形式

矩阵形式

2.4.3 修正单纯形法

修正单纯形法存储量和计算量都比单纯形法少，当n远大于m时更为明显

2.4.4 退化与循环

1 摄动法 2 字典序法 3 Bland法

2.5 线性规划的对偶理论与对偶单纯形法

2.5.1 对偶线性规划问题

1. 交换常矢量c,b的位置，变矢量x用λ替换

2. 改变约束不等式的不等号方向

3. 把min换为max

4. 交换A与变矢量的位置，并按需要做适当的转置

2.5.2 对偶定理

引理1 x和w是可行解，则cTx>=wTb

线性规划的对偶定理

对偶规划中若有一个有有限的最优解，另一个也有有限的最优解，且目标函数值相同

若一个目标函数值无界，则另一个问题没有可行解

2.5.3 对偶单纯形法

定理5.1 x是任一基本解，对应基为B，wT=CTBB-1,若x,w是对偶的可行解，则x,w是最优解

定义5.1 若x是基本解，对应的检验矢量r>=0,则x为对偶可行解或正则解

基本思想： 从对偶可行的基本解开始，逐次迭代，在保持对偶可行性条件下，使原始问题的基本解x的不可行性消失，直到获得一个基本可行解为止。

2.6 线性规划的多项式算法

第三章 无约束优化方法

3.1 引言 下降递推算法

无约束优化方法：n个未知数n个方程，一般是非线性的，一般求不出准确解，只能用数值方法逐步求其近似解。

迭代法 基本思想：给出初始估计，计算一系列点列，极限是最小点。

不同算法差别是如何选取方向s(k)和步长λk

下降递推算法/下降算法：极小化序列对应的函数值是逐次减少的，至少是不增的

收敛：序列中某个点是极小点或者极限是极小点

迭代过程：1.选择初始点x(0)

2. 如果x(k)已求得，且不是极小点，设法选择方向s(k)，使目标函数沿搜索方向s(k)下降，至少不增。

3. 在射线x(k)+λs(k)上选取适当补偿λk，使f(x(k)+λs(k))<=f(x(k)),确定下一点x(k+1)=x(k)+λks(k),这种确定步长的方法称一维搜索或线搜索

4. 检验x(k+1)是否为极小点或者满足精度要求的近似极小点，迭代结束，否则继续迭代

定义1.1 线性收敛：存在β∈(0,1)使得当k>k0时，||x(k+1)-x\*||<=β||x(k)-x\*||

定义1.2 α阶收敛：存在β∈(0,1)和α>1使得当k>k0时，||x(k+1)-x\*||<=β||x(k)-x\*||α

a=2时 二阶收敛 1<a<2时超线性收敛 阶数越大，收敛越快

结束条件： ||x(k)-x\*||<ε x\*未知 用||x(k+1)-x(k)||<ε

3.2 一维搜索

(精确)一维搜索：步长λk由φ(λ)=f(x(k)+ λs(k))的极小值确定，即f(x(k)+λks(k))=min(f(x(k)+ λs(k)))

3.2.1 牛顿法

基本思想 用φ(x)在x0处的二阶泰勒展开式来近似代替φ(x)

即φ(x)≈g(x)= φ(x0)+ φ’(x0)(x-x0)+1/2φ’’(x0)(x-x0)2

即x1 = x0-φ’(x0)/ φ’’(x0)

优点：收敛速度快，至少是二阶收敛

缺点：需计算二阶导数，初始点选不好可能不收敛

3.2.2 抛物线法

利用f(x)在三个点x0,x1,x2处的函数值来构造一个二次函数y=g(x)≈a0+a1x+a2x2 使它满足…

X拔=-a1/2a2 =1/2 [x1+x0-b1/a2]

优点：一定条件下是超线性收敛的，收敛阶为1.3

缺点：不保证一定收敛，存在退化情况

3.2.3 三次差值法

用a，b两点处的函数值f(a),f(b)和导数值f’(a),f’(b)来构造三次差值多项式g(x)

g(x)=α(x-a)3+β(x-a)2+γ(x-a)+δ 满足…

α=0 x=a-γ/2β

α≠0 x=a+(-β+sqrt(β2-3αγ))/3α

β=(3u-v)/(b-a)

α=(v-2u)/(b-a)2

3.2.4 平分法

找到一个区间[a,b]有f’(a)\*f’(b)<=0 通过二分法来查找

优点：计算量小，程序简单。总能收敛

缺点：收敛速度慢

3.2.5 “成功-失败”法

1. 选定初始点x0，搜索步长h>0及精度ε>0

2. 计算x1=x0+h，f(x1)

3. 若f(x1)<f(x0) 搜索成功，下一次用x1代替x0，2h代替h

否则搜索失败，若|h|<ε，则计算阶数

否则，用-h/4代替h返回第2步

3.2.6 0.618法

1. 计算λ1=a1+(1-α)(b1-a1), φ(λ1)

μ1=a1+α(b1 -a1), φ(μ1)

2. 若bk-ak<ε，计算结束。否则若φ(λ1)> φ(μ1),转3，否则转4

3. 令ak+1=λk，bk+1=bk，λk+1=μk，μk+1=ak+1+α(bk+1-ak+1) 计算φ(λk+1)。转5

4. 令ak+1=ak，bk+1=μk，μk+1=λk，λk+1=ak+1+(1-α)(bk+1-ak+1) 计算φ(λk+1)。转5

5. 令k=k+1,返回2

3.2.7 确定初始搜索区间和初始点的方法

1. 选定初始点的估计值t0，初始步长h>0，计算φ(t0)

2. 令t2=t0+h,计算φ(t2)

3. 若φ(t2)<= φ(t0)，转4，否则h=-h，转4

4. t1=t0+h，计算φ(t1)

5. 若φ(t1)<= φ(t0)，令h=2h,t2=t0,t0=t1,转4，否则，转6

6. 令a=min(t1,t2),b=max(t1,t2),则[a,b]是所求初始搜索区间，t0=(b+a)/2作为初始点

3.2.8 不精确一维搜索

不精确一维搜索：f(x(k+1))比f(x(k))下降一定的数量，x(k+1)处导数值比x(k)大一定数量

Wolfe-Powell准则：

1. f(x(k))-f(x(k+1))>=-c1λk▽f(x(k))Ts(k)
2. ▽f(x(k+1))Ts(k)>=c2▽f(x(k))Ts(k) s

常取c1=0.1 c2=0.5

3.3 求多变量函数极值的基本下降法

3.3.1 最速下降法

基本思想：沿负梯度方向下降最快

步骤： 1. 给定初始点x(0),精度ε>0，令k=1

2. 计算▽f(x(k))

3. 若||▽f(x(k))||< ε,迭代结束，取x\*=x(k) ,否则转到4

4. 用精确一维搜索求φ(λ)=f(x(k)-λ▽f(x(k)))的极小点λk,使f(x(k)-λk▽f(x(k)))>f(x(k))

5. 令x(k+1)=x(k)-λk▽f(x(k))，k=k+1

优点：方法简单。从不好的初始点也能保证算法收敛

缺点：极小点附近搜索慢。收敛速度与变量的尺度关系很大。关于小的扰动不稳定

3.3.2 Newton法

f(x)≈φ(x)=f(x(k))+(▽f(x(k)))T(x-x(k))+1/2(x-x(k))T▽2f(x(k))(x-x(k))

x(k+1)=x(k)-(▽2f(x(k)))-1▽f(x(k))

优点：收敛快，至少二阶收敛

缺点：要求f(x)二阶可微。计算二阶梯度的-1困难。初始点距离x\*太远可能不收敛

3.3.3 阻尼Newton法

步长λk不总取1，取使得f(x(k)+λks(k))=min(f(x(k))+λs(k))

x(k+1)=x(k)- λk (▽2f(x(k)))-1▽f(x(k))

收敛定理：{f(x(k))}是严格单调序列，{x(k)}极限点必为极小点

3.4 共轭方向法与共轭梯度法

比最速下降法收敛速度快，又不用求海赛矩阵

3.4.1 共轭方向和它的一些性质

定义4.1 A是n阶对称矩阵，p，q是n维向量，若pTAq=0，则p和q关于A正交/共轭

pTq=0 p和q正交

有p1,p2,…,pm，pTiApj=0(i≠j，j=1,2,…,m)成立，则p是A-正交/共轭向量组。称A共轭方向

定理4.1 pi是线性无关的，若q与pi都正交，则q=0

定理4.2 A是n阶正定矩阵，p1,p2,…,pn是A-共轭的n维非零向量组，则向量组线性无关

定理4.3 H是n阶对称正定矩阵，s是一组H共轭方向，若从任意初始点出发，依次经过s进行精确一维搜索，则至多n次迭代，即可求得f(x)最小点

3.4.2 共轭方向的生成与共轭梯度法

共轭方向法：选取一组H共轭的方向，就可以在n步之内求得极小点

共轭梯度法：利用每次一维最优化所得到的点处的梯度来生成共轭方向

3.4.3 共轭梯度法的计算步骤和程序框图

3.4.4 应用举例

3.5 变尺度法

3.5.1 基本思想

最速下降法计算简单，但搜索太慢

牛顿法收敛速度快，但计算复杂

拟牛顿方程 G△xk=△gk G是n阶对称正定矩阵

3.5.2 对称秩1算法

公式： Hk+1=Hk+(△xk-Hk△gk)(△xk-Hk△gk)T/△gTk(△xk-Hk△gk)

缺点：Hk+1不一定正定。分母可能为0

3.5.3 DFP算法

一种秩2对称算法

Hk+1=Hk+(△xk△xkT)/(△xkT△gk)-(Hk△gk(Hk△gk)T)/( △gkTHk△gk)

计算步骤

1. 给定初始x(0)、计算精度ε>0和初始矩阵H0=I
2. 计算s(k)=-Hkgk,沿s(k)进行精确一维搜索，求出步长λk,使f(x(k)+λks(k))=minf(x(k)+λs(k))

令x(k+1)=x(k)+λks(k)

1. 若||gk+1||<ε,则取x\*=x(k+1)，结束。否则由公式计算Hk+1,若k≠n-1，则k=k+1返回2. 否则，令x(0)=x(k+1),k=0，返回2

优点： 目标函数是严格凸函数，初始矩阵取I，最多经过n步收敛，即二次收敛。 目标函数是严格凸函数，是全局收敛

若Hk是对称正定函数，Hk+1也是对称正定函数

缺点： 存储量大，需要O(n2)个存储的单位

计算的稳定性不高

计算效果不好

3.5.4 吴桂变尺度算法类

3.5.5 Huang类算法

3.5.6 自调节变尺度算法类

3.5.7 BFGS变尺度算法

Hk+1=Hk-(Hk△gk△gkTHk)/(△gkTHk△gk)+(△xk△xkT)/(△xkT△gk)+(△gkTHk△gk)vkvkT

3.6 直接搜索法

3.6.1 Hooke-Jeeve方法

探测搜索：探求下降的有利方向

模式搜索：沿着有利方向进行加速

3.6.2 单纯形法

对n维空间的n+1个点上的函数值进行比较，丢掉最坏的点，代以新点。

3.6.3 Powell方法

第四章 约束优化方法

1. 用线性规划或二次规划来逐次逼近非线性规划，如SLP法，SQP法

2. 把约束优化问题转化为无约束优化问题来求解的方法，如SUMT法，SUMT内点法，乘子法等

3. 对约束优化问题不预先做转换，直接进行处理的分析方法，如可行方向法，梯度投影法，既约梯度法等

4. 对约束优化问题不预先做转换的直接搜索方法，如复形法，随机试验法

4.1 引言 Kuhn-Tucker条件

紧约束/起作用约束：i∈I(x0)时，对应的约束gi(x)<=0称为在x0处的紧约束

定理1.1 最优性的一阶必要条件

定理1.2 最优性的充分条件

4.2 惩罚函数法

基本思想：把约束问题转化为一个或一系列无约束问题来求解，

也称序列无约束极小化技术SUMT

惩罚函数法：SUMT外点法

碰壁函数法：SUMT内点法

4.2.1 惩罚函数的性质和构造

F(x,M)=f(x)+Mp(x)

p(x)是连续的，且>=0,x∈S时=0

4.2.2 计算方法和程序框图

1. 选取M1>0,精度ε>0,c>=2，初始点x(0)，令k=1

2. 以x(k-1)为初始点，求解无约束优化问题minF(x,M)=MkΣgi+(x),设其最优解为x(k)=x(Mk)

3. 令τ1=max{|hi(x(k))|},τ2=max{g(x(k))},τ=max{τ1，τ2}

4. 若τ<ε，结束。否则Mk+1=cMk,k=k+1,转2

4.2.3 收敛性分析

如果Mk0对应的无约束问题的最优价，则是原问题的最优解

否则得到一个无穷点列，在某些条件下，极限点都是最优解

定理2.1 f，gi，hj是由惩罚函数法产生的一个序列，则序列极限点是约束优化问题

的最优解

4.2.4 算法应用举例

4.3 碰壁函数法

健集：不包含孤立点和孤立的线段。即不包含等式约束

4.3.1 SUMT内点法的计算步骤

4.3.2 算法应用举例

4.3.3 收敛性分析

4.3.4 混合罚函数法

4.4 可行方向法

4.4.1 可行方向与改进方向

4.4.2 Zoutendijk的可行方向法

4.5 梯度投影法

4.5.1 理论基础

4.5.2 计算步骤

4.5.3 计算例子

4.6 既约梯度法

4.6.1 线性约束情形

4.6.2 非线性约束情形

4.7 乘子法

4.7.1 Hestenes的乘子方法

4.7.2 Powell的乘子方法

4.7.3 Rockafellar的乘子方法

4.7.4 方法应用举例

4.8 二次逼近法

4.8.1 二次规划

4.8.2 二次逼近法

4.9 极大熵方法

4.9.1 引言

4.9.2 极大熵方法

4.9.3 数值结果

4.9.4 结束语