目录

[基本数学 1](#_Toc28968023)

[范数 1](#_Toc28968024)

[线性规划 3](#_Toc28968025)

[单纯形法 3](#_Toc28968026)

[对偶 4](#_Toc28968027)

[灵敏度分析 7](#_Toc28968028)

[非线性规划 8](#_Toc28968029)

[最优性条件 8](#_Toc28968030)

[算法 10](#_Toc28968031)

[使用导数的最优化方法 10](#_Toc28968032)

[最速下降法 10](#_Toc28968033)

[牛顿法 10](#_Toc28968034)

[共轭梯度法(重点) 11](#_Toc28968035)

[可行方向法 11](#_Toc28968036)

[Zoutendijk可行方向法 11](#_Toc28968037)

[Rosen梯度投影法 11](#_Toc28968038)

[既约梯度法 12](#_Toc28968039)

[罚函数法 12](#_Toc28968040)

[Lagrange对偶 12](#_Toc28968041)

[凸集分离定理相关 13](#_Toc28968042)

[凸集分离定理 13](#_Toc28968043)

[Farkas定理 13](#_Toc28968044)

[其它 13](#_Toc28968045)

[没懂的 14](#_Toc28968046)

# 基本数学

## 范数

###### 直观理解

一个n维向量的每个分量经过一通瞎JB操作得到一个非负数，可以理解为是n维空间到非负1维空间的一个映射。

还有矩阵范数, 也是把矩阵经过一个映射变成一个数, 常用的有矩阵1范数, 2范数和无穷范数.

###### 数学定义

###### 常用范数

###### 凸集

###### 方向

证明d是一个方向的重要结论：

###### 极方向

###### 凸函数

###### 其它

正定矩阵可以用顺序主子式来判定。

半正定矩阵可以用主子式来判定。

矩阵相乘除了一次确定一个位置最终的值之外（即一行x一列），还有一种求法，每次确定最终结果的一部分，加起来是最终的结果（一列x一行）。

证明极点，

# 线性规划

可行域一定是凸集. 最优解可以在极点取得，但是不是一定在极点取得，比如有无限多个最优解的一条线上。

### 单纯形法

注意的顺序和基变量的顺序有关

求解线性规划问题的最优解等价于求解最优基本可行解.单纯形法的思想就是在不断找基本可行解,直到找到最优的.

单纯型法3步骤: 1 判断是否是最优解 2 选进基变量 3 选出基变量 4 转1. 看出来这里是先选进基, 再选出基

如果基变量取值, 则为非退化, 有一个基变量为0,则为退化

普通单纯形法中要保证最右端列一直, 即保持解的可行性

求解极小化问题,要找判别数最大的,如果最大的都,那么就是最优解.反过来,求解极大化问题,要找判别数最小的,如果最小的都,那么就是最优解，求极大化问题的时候，建议直接把乘以-1转化为极小化问题，用极小化问题的做法来做，最后的最优值取相反数即可

当选取某个进基变量后，在选取出基变量的过程中，发现所有的列都是负数，则原问题无界

两阶段法的基本思想: 第一阶段找基本可行解, 第二阶段就是原始的求解过程.

为了在第一阶段就找到单位矩阵I, 两阶段法和大M法大概率都要引入人工变量, 就是为了得到单位矩阵I, 然后求有人工变量的基本可行解，此时有三种情况

两阶段法和大M法可只先掌握一个

算判别数的时候，基变量的判别数都是0

两阶段法，第一阶段最后的单纯形表，把人工变量的列去掉，保留系数部分和右端的b，最后一行的判别数（基变量为0）和重新计算，然后开始第二阶段。为什么要重新算？因为目标函数变了，你原来的目标函数是min （人工变量相加），现在的目标函数变成了原问题。

在选出基变量的时候，如果有两行相等(b/系数)，则随便选一个

在两阶段法和大M法的过程中，引入的人工变量也必须是非负的，也就是说，在两阶段的第一阶段，如果最优值已经是0了，则第一阶段也就结束了

人工变量最后都是0，但是一部分是基变量，这个时候要把人工变量出基。人工变量出基的顺序任意的，在已选取某个人工变量出基的情况下，选取入基变量也是任意的，但是这个入基变量系数不为0。如果在找入基变量的时候发现所有的非基变量系数都是0，说明这个方程是多余的，直接把这个人工变量为基的那行去掉，也就是说这个基我不要了。整个换基过程不用看判别数，也不用看系数的正负(P64，例3.2.1-3.2.3)在主元消去的过程中，判别数那一行不用参与，0值的人工变量出基后，用最后的单纯型表去掉人工变量列，判别数行重新计算，开始第二阶段，跟前面的情况一样

P62 例3.2.1-例3.2.3

### 对偶

对偶有两种，一种是对称对偶，一种是非对称对偶

对称对偶：

其中A的行数和列数相等。做法：目标函数的系数和b对换，min-max互换，A转置，约束大于等于变小于等于

非对称对偶：直接背



记住几个对偶定理：

*（弱对偶）*

#### 互补松弛定理

心法：自己是松的，对偶就是紧的，变量约束双向互换

应用：当知道原问题的最优解的时候，利用互补松弛定理求出对偶问题的最优解，常常伴随着写出对偶问题，所以要会写出一个问题的对偶形式

##### 对偶单纯形法

首先，为啥要用对偶单纯形法来求解原问题，而不是直接从原问题应用单纯形法来求解？是因为有些问题用对偶单纯形法要更简单，多做题应该就有体会了！

然后，对偶单纯形法不需要引入人工变量，标准模型也是极小化模型

总体思想：先得到一个对偶可行的基本解（这个最初的对偶可行基本解不一定原问题可行），然后不停改进这个对偶可行的基本解，直到它是原问题的可行解，则此时找到原问题的最优解。所谓改进对偶可行的基本解，是指改进后，对偶问题的目标函数值有增加

一开始找对偶可行的基本解的时候，引入松弛变量，如果松弛变量是负的，就乘以-1，把它转成正的。

对偶单纯形法，是先求出基变量，再求进基变量，和经典单纯型法相反。在经典单纯形法中，始终保持右端列非负，即保持原问题的可行性；在对偶单纯形法中，始终保持，即保持对偶问题的可行性，但是不要求右端列非负，所以不需要引入人工变量

在对偶单纯形法的过程中，如果化到某一次发现右端列，则现行基本解是最优基本可行解

如果，则要先选出基变量，再选进基变量，进行主元消去，进行下一次迭代。出基变量选取：为了在保持对偶可行的条件下求得原问题的可行解，应选择取最小负值的基变量作为出基变量，即的行。进基变量选取：为保持对偶可行性，用判别数来除以r行的每个负元，从里面选最小的

当迭代中出现： 但第r行无负元，无法确定下标k来进基，这说明原问题的变量取任何非负值都不能满足第r个方程，原问题无解。作为对比，在基本单纯形法中，如果第j列全,则原问题无界

综上，对偶单纯形法的计算步骤如下：

注意对偶单纯形法只是不用引进人工变量，松弛变量还是需要的

最优单纯形表中松弛变量对应的判别数的相反数是对偶问题的最优解 P146 例4.2.1

##### 对偶单纯形法的扩充问题

经典单纯形法可能一开始不好找单位基矩阵，因此引入了人工变量的概念，对偶单纯形法也一样，如果一开始找不到对偶可行的基本解，则要解一个扩充问题，从这个角度来说，好像对偶单纯形法也没什么特别大的优势= =

拿到问题以后，先把松弛变量搞了，如果松弛变量是-1，则等式两边都乘以-1把松弛变量的系数变成1，这个时候应该能得到一个单位矩阵的基，但是发现判别数没有都，说明这个基本解不是对偶可行的，这个时候就要引入扩充问题了，引入扩充问题是加一行和一列，非基变量和新引入的变量系数为1，基变量系数为0，这行的b是M（一个很大的正数）

150页，例4.2.2（）和151页例4.2.3（）

选择新增的那行中判别数最大的一列为主元，进行主元消去（注意这一步伴随着出基进基），则得到对偶可行的基本解(注意这里只有1次迭代)，从这个单纯形表开始按照对偶单纯形法的方法继续往下做，新增加的人工变量也要参与第二个步骤

当你看到变量个数比较多，大概4-6个，然后约束方程比较少，大概2个，应该差不多就是要用对偶单纯形法了

### 灵敏度分析

主要研究参数的变化对最优解和最优值造成的影响，做到有的放矢，心中不慌，参数包括

c，A，b，增加新约束

灵敏度分析基本是搞清楚两个问题，第一，谁变？第二，变化有没有影响基？

题型：1 参数改变后，最优解和最优值是否改变，改变成什么？ 2 参数在什么变化范围内最优解/最优值保持不变？

WARNING：如果是把max化成min求解最优化问题，则最后的单纯形表中，把最后一行都乘以-1，再进行灵敏度分析

##### c变

1 非基变量的系数变了（只影响自己的判别数）

变化后的判别数=

2 基变量系数变了（影响所有非基变量的判别数）

系数变的那个基变量的判别数还是0不变。其它（包括）=原数+（变化后-变化前）\*第k行系数，第k行是基变量那一行。加完后继续用单纯形法做下去。

P160 例4.4.1

##### b变

设最优解对应的基矩阵为B，则后（注意最优基是当前基变量对应的原始问题的矩阵中的列，而不是化来化去化过之后的，但是最优基的逆在最优单纯形表中一开始的基对应的列上）

*，用新的替代即可*

##### A变（不考）

非基列改变：只会影响这一列的判别数和最优表中的列系数，如果新的判别数还是，则最优解不变，如果，最优表中的列变为，判别数变为，进基，继续往下迭代

基列改变：直接从头当作一个新问题开始计算

##### 增加新约束

1. 如果最优解满足新添加的约束，则自然还是最优解，这个很好理解
2. 如果最优解不满足新添加的约束，则把新约束加到最优表的最后一行（注意添加松弛变量入基），把基化为单位矩阵，再用对偶单纯形法继续往下做即可。见例4.4.3 P166

# 非线性规划

拿到非线性规划的题，不管三七二十一，先把求出来，注意在特定的点，哪些是起作用的，哪些是不起作用的，

## 最优性条件

研究非线性规划的最优解要满足的必要条件和充分条件，分为两种来讨论，无约束问题和有约束问题

#### 无约束（简单）

把高中和大学学的低维的情况扩展到高维。模型：

一阶必要条件：

二阶必要条件：

二阶充分条件：是局部极小点

推论：正定二次函数有唯一极小点

如果，

#### 有约束（重点）

一般不能用无约束的方法来解决有约束的问题，因为自变量的范围受到限制

下降方向：

可行方向锥：。在局部极小点处，有（几何最优性条件）

#### 一阶最优性条件（必要条件）

，则称为约束和的正则点

再扩展一下，把扩展到所有的，则有

将，则一般情形的一阶必要条件可以表述为

如果是凸规划（），

#### 二阶条件

对于有约束问题，即使是KKT点且在这个点的Hessian矩阵正定，也不一定是最优解

二阶必要条件：

二阶充分条件：

如果题目求出来d=0，则，自然满足要求（P237，例7.2.7和例7.2.8，这两个例子多看看，很有代表性）

一阶在不是凸规划的情况下只能给出必要条件，在凸规划情况下

二阶在不是凸规划的情况下能给出局部最优的充分条件

一般解题只用一阶必要和二阶充分，几乎不会用二阶必要（目前我还没见过用二阶必要来做的）

验证一个点是不是最优解，先看它是否满足约束（是否是可行点），如果连约束都不满足，下面的就不用求了。

拿到一个非线性规划问题，先看是不是凸规划，如果是凸规划，直接求KKT点就好了，二阶充分条件不必用，麻烦。

## 算法

闭映射：

会判断一个算法在某个点是不是闭的。P258，8.1.4 B(x)的例题好好看

# 使用导数的最优化方法

这一章和下一章都是用数值方法（迭代法）求解无约束问题，只能求近似解。

## 最速下降法

就是梯度下降法，方向：，如果当前的，则算法停止迭代，是自己认为的一个可以接受的阈值

## 牛顿法

二阶Taylor展开的一个应用，迭代公式，如果，则停止迭代

二次终止性：算法作用于二次凸函数经有限次迭代必达到极小点

最速下降方法没有二次终止性

牛顿法有二次终止性

共轭方向法有二次终止性

牛顿方向，牛顿方向不一定是下降方向，只有Hessian矩阵正定才是下降的

## 共轭梯度法(重点)

共轭：

注意这里二次项有个，所以在代数转化成矩阵的时候对角线要乘以2。

关于A共轭的个非零共轭方向线性无关。因此，这个方向可以生成一个子空间，叫做扩张子空间。在扩张子空间的条件下，有，也就是说这一步迭代点的导数和以前的搜索方向全部正交。

# 可行方向法

可行方向法研究的是有约束问题的最优化方法。基本思想也是梯度下降法，即找方向和步长，搜索方向的选择方式不同形成不同的可行方向法

### Zoutendijk可行方向法

模型：

设是可行解，，则非零向量d是。下降可行方向：

因此，求解下降可行方向归结为求解下面这个问题

*（用Farkars定理证明的，这个证明我服了，可以看看P371，例12.1.2）*

### Rosen梯度投影法

投影矩阵：

投影矩阵的性质：

思路：当可行点在区域内部，沿着最速下降方向搜索。当可行点在区域的边界，沿着最速下降方向(负梯度方向)在M的零空间的投影向量方向搜索，M是起作用约束的梯度为行构成的矩阵。

可以看它一些定理的证明过程，但是我觉得定理本身都没有去背的必要，太长了= =

### 既约梯度法

# 罚函数法

一种有约束的最优化方法，基本思想：借助惩罚函数把约束问题转化为无约束问题。要会求罚函数和障碍函数

##### 外点罚函数法

##### 内点罚函数法

内点罚函数只适用于不等式约束，如果优化问题有等式约束，则不适用内点罚函数法

### Lagrange对偶

非线性规划的对偶问题，要能够写出一个非线性规划的对偶问题。

1. 写出Lagrange函数(注意集约束不要搞到Lagrange函数里面，这一点跟普通的Lagrange函数有所不同)
2. 求下确界
3. 把分开
4. 计算
5. 得到对偶函数

注意 （P243, 例7.3.1）

# 凸集分离定理相关

开集：所有点都是内点。内点：这个点邻域内的所有点还在集合内

闭集：补集是开集

紧集：有界闭集

凸集分离定理

Farkas定理

证明过程看老师课件，很厉害，[9-凸集分离定理 P15](../课件和书/最优化/9-凸集分离定理.pdf)，核心是点与凸集的严格分离

**Gordan定理**

[9-凸集分离定理 P17](../课件和书/最优化/9-凸集分离定理.pdf)，证明套路跟Farkars定理一模一样

# 其它

填空+判断(10x3=30)+计算（20线性+20非线性=40）+证明（10x3=30）

不管是考互补松弛还是单纯形，答案最后把最优解和最优值一并给出，不要只给出一个，有可能扣分，吧。

要会用Gordan和Farkas定理证明相关结论，且会这两个定理的证明方法。作业里面有类似的，多看下。

老师经常会把和互换，阴险，一定要注意标准模型和定义里面的符号！然后定理的使用前提也要注意，比如什么可微，二次可微等，不满足定理的使用前提的不能用。

迭代超过5次，分母上百，则单纯形法你就做错了，老师不会出这么变态的题。

不能用初等方法和画图求解，一定要用最优化方法（即这门课学过的方法），不然没分

对偶可行的基本解的定义

灵敏度分析：矩阵A的列改变不考察，只考察。不仅要会求新基和新最优解，而且要会求最优解或者最优基不变的情况下

要会判断234阶矩阵的正定性，还是用主子式，用特征方程要解3次方程，烦得很

局部最优解不一定是KKT点，因为可能这个点不是正则点

算法：会判断算法在某个点是不是闭的

牛顿方向不一定是下降方向，只有Hessian矩阵正定才是下降的

Rosen梯度投影法和既约梯度法多看看，要考证明

会写外罚函数和内罚函数

算法的迭代公式不用记，如果要考，肯定会把迭代过程写在题目里，然后考其它的，不会让你手算数值分析。

正交投影矩阵怎么写的不用记，题目里面会给。

看到要对那种的命题进行证明的，首先想到数学归纳法。

# 没懂的

第二次作业第二题证明它是极方向的第一种解法，用线性相关和线性无关来证明的，没看懂，去问下陆老师

Gordan定理有像Farkas定理那样直观的几何解释么？

Gordan定理的证明过程我觉得挺正确的。。。但是我记得老师上课的时候说这个证明方法不对，我有点懵逼。。。[9-凸集分离定理 P17](file:///C:\Users\hl\Desktop\课程git\课件和书\最优化\9-凸集分离定理.pdf)

凸集，凸函数，极点，极方向，既约梯度法，Farkas，Gordan定理，可行方向法那一章的内容和定理证明，作业