运筹学

玩原神玩的

封面日期：2025 年 4 月 18 日

前言标题

这是一个基于LaTeX的模板，用于撰写学习笔记。

模板旨在提供一个简单、易用的框架，以便你能够专注于内容，而不是排版细节，如不是专业者，不建议使用者在模板细节上花费太多时间，而是直接使用模板进行笔记撰写。遇到问题，再进行调整解决。

前言页显示日期：2025 年 4 月 18 日

目录

第零章　对偶理论

Theorem 0.0.1. 对偶问题的性质

1. 弱对偶性：$CX \leq b^T Y$
2. 最优性：$CX = b^T Y \Rightarrow X, Y$ 均为最优解
3. 强对偶性：设 X^0, Y^0 分别是原始问题和对偶问题的可行解
   则必存在最优解 X^*, Y^*，且有 $c^T X^* = b^T Y^*$
4. 互补松弛性：X^*, Y^* 为最优解的充要条件

$$(AX - b)^T Y = 0 \text{ 和 } X^T(A^T Y - C^T) = 0$$

$$X_s^T Y = 0 \text{ 和 } X^T Y_s = 0$$

Proof. 弱对偶性:

设 **X**、**Y** 分别是原始问题和对偶问题的可行解，有

$$z = c^T \mathbf{X} \leq \mathbf{Y}^T \mathbf{A} \mathbf{X} \leq \mathbf{Y}^T \mathbf{b} = w$$

□

Remark. 一个问题无界解时，另一个问题无可行解

互补松弛性

可行解 $\mathbf{X}^0$、$\mathbf{Y}^0$ 分别是原始问题和对偶问题的最优解，充要条件是

$$(\mathbf{A}\mathbf{X}^0 - \mathbf{b})^T \mathbf{Y}^0 = 0 \quad \text{和} \quad \mathbf{X}^0(\mathbf{A}^T \mathbf{Y}^0 - \mathbf{C}^T) = 0$$

或

$$\mathbf{X}_s^{0T} \mathbf{Y}^0 = 0 \quad \text{和} \quad \mathbf{X}^{0T} \mathbf{Y}_s^0 = 0$$

Proof. 满足互补松弛性时有最优解的证明

充分性:

原问题

$$\max\ z = CX$$

$$\text{s.t. } AX + X_s = b$$

$$X, X_s \geq 0$$

对偶问题

$$\min\ w = b^T Y$$

$$\text{s.t. } A^T Y - Y_s = C^T$$

$$Y, Y_s \geq 0$$

可得：

$$z = CX^0 = X^{0T}(A^T Y^0 - Y_s^0) = X^{0T} A^T Y^0 - X^{0T} Y_s^0$$

$$w = b^T Y^0 = (X^{0T} A^T + X_s^{0T}) Y^0 = X^{0T} A^T Y^0 + X_s^{0T} Y^0$$

故有 $w = z$。由最优性知均为最优解。

必要性:

若 X^0, Y^0 是最优解，有： $CX^0 = b^T Y^0 = z^* = w^*$

原问题

$$\max z = CX$$

$$s.t. \quad AX \leq b$$

$$X, X_s \geq 0$$

对偶问题

$$\min w = b^T Y$$

$$s.t. \quad A^T Y \geq C^T$$

$$Y, Y_s \geq 0$$

可得： $X^0 T A^T Y^0 \leq b^T Y^0 = w^* = z^* = CX^0 \leq X^0 T A^T Y^0$

即： $z^* = w^* = X^0 T A^T Y^0 \Rightarrow X_s^0 T Y^0 = 0$ 和 $X^0 T Y_s^0 = 0$。 □

原问题	对偶问题
$\max\ z = CX$ s.t. $AX \le b$ $X \ge 0$	$\min\ w = b^T Y$ s.t. $A^T Y \ge C^T$ $Y \ge 0$
$AX \ge b$	$Y \le 0$
$AX = b$	Y unr
$X \le 0$	$A^T Y \le C^T$
X unr	$A^T Y = C^T$

Remark. 1. 松弛变量的检验数是对偶问题的变量的相反数

2. 实际上可以理解为，松弛变量是工厂生产的多余材料，检验数是增加某方案的利润，松弛变量相当于只增加材料的消耗，它的检验数就是材料的价格，正好对应对偶问题的变量的相反数

Definition 0.0.1. 影子价格影子价格（Shadow Price）是利润最大化生产下资源的边际利润，反映了资源的利润价值。

可以理解为在当前方案下的资源的心理价格，即在当前方案下，买资源一方愿意出的价格。

在利润最大化的生产计划中

1. 边际利润大于0的资源没有剩余（影子价格越大说明资源越紧缺）
2. 有剩余的资源，边际利润等于0（购进该种资源不能增加利润）
3. 安排生产的产品，差额成本等于0（出售的机会成本等于生产利润）
4. 差额成本大于0，不安排生产（出售机会成本大于生产利润，不如卖资源）

Remark. 对偶问题解与检验数的关系

1. 单纯形表上原问题的检验数对应了对偶问题的一个基解
2. 单纯形表上原问题的和检验数对应了对偶问题具有相同的目标函数值

第一章　运输问题、目标规划、整数规划

1.1　运输问题

Definition 1.1.1. 生产平衡问题的一般模型

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

$$\text{s.t.} \quad \sum_{j=1}^{n} x_{ij} = a_i, \quad i = 1,2,...,m$$

$$\sum_{i=1}^{m} x_{ij} = b_j, \quad j = 1,2,...,n$$

$$x_{ij} \geq 0, \quad i = 1,2,...,m; \quad j = 1,2,...,n$$

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

rank(A) = m+n-1 (约束有1个冗余)的解释:

m个产地,n个销地，只要满足m+n-1，最后一个自动满足

1. 西北角法：优先选择西北角的运输方案
2. 最小元素法：优先选择运费最小的方案
3. Vogel法：优先考虑次小运费和最小运费差额大的方案。

前两个省略，伏格尔法的基本思想是，如果次小运费和最小运费的差额很大，**不早点占住小运费的地方，就会导致差可行解。**

何老师的两个问题: 1、闭回路是否一定存在？ 2、闭回路是否唯一？

对于第一个问题，闭回路是肯定存在的，闭回路实际上是单纯形法出基入基在运输问题的特殊体现，单纯形法显然总能出基入基，所以一定可以找到闭回路。

对于第二个问题，闭回路是唯一的。

Proof. 假设存在两个环路，对于同一个非基变量，那么可以合并为一个只含基变量的环路

等价于一个基变量对应的方案可以由其他基变量对应的方案表出，这和基变量的定义矛盾。 □

Remark. 除了用闭回路来求检验数和出基入基，实际上还可以用位势法来求检验数，而且快的多，还是解析解法。

位势法求检验数:

$$\sigma_j = c_j - C_B B^{-1} P_j = c_j - Y^T P_j$$

$$\sigma_{ij} = c_{ij} - [u_1, ..., u_m, v_1, ..., v_n] P_{ij} = c_{ij} - (u_i + v_j)$$

对应m+n-1个基变量，有 $\sigma_{ij} = 0$，则:

$$u_i + v_j = c_{ij}$$

共有m+n个变量，m+n-1个等式，故解不唯一，称为位势。根据位势求非基变量对应的 σ_{ij}。

这里我们常常令$u_1 = 0$

1.2 目标规划

Definition 1.2.1. 目标规划的思想和方法

思想：将定量技术和定性技术结合，承认矛盾、冲突的合理性，强调通过协调，达到总体和谐

方法：软约束+优先级

Example 1.2.1. 电视机厂装配25寸和21寸两种彩电，每台电视机需装备时间1小时，每周装配线计划开动40小时，预计每周25寸彩电销售24台，每台可获利80元，每周21寸彩电销售30台，每台可获利40元。

该厂目标：

1. 避免开工不足。 2. 允许装配线加班，但尽量不超过10小时。 3. 尽量满足市场需求，尤其是25寸彩电。

解：设 x_1, x_2 分别表示25寸，21寸彩电产量，目标函数为：

$$\min Z = P_1 d_1^- + P_2 d_2^+ + P_3 (W_{33}^- d_3^- + W_{34}^- d_4^-)$$

其中，P_1, P_2, P_3 是目标函数的权重系数。

约束条件为：

$$x_1 + x_2 + d_1^- - d_1^+ = 40 \quad （上班时间约束）$$

$$x_1 + x_2 + d_2^- - d_2^+ = 50 \quad （加班时间约束）$$

$$x_1 + d_3^- - d_3^+ = 24 \quad （25寸市场需求）$$

$$x_2 + d_4^- - d_4^+ = 30 \quad （21寸市场需求）$$

其中，$x_1, x_2, d_i^-, d_i^+ \geq 0$，并且有互补松弛条件：

$$d_i^- \cdot d_i^+ = 0 \quad (i = 1, 2, 3, 4)$$

Remark. 优先因子：P表明第i个目标的重要程度

权重系数：W表明相同优先因子目标的权重

何老师的两个问题:

1. 高一级的目标没有满足？低一级的目标是否还有满足机会？

2. $d_i^- \cdot d_i^+ = 0$ 为什么可不考虑？

对于第一个，高一级满足了才能满足低一级，这是优先级不可逆性。

对于第二个，从直接上讲，生产不可能既过剩又短缺，所以不需要多此一举。同时，从数学上讲，对于相同的X,如果其一不为0，说明没优化到最优解，还可以再优化。

何老师的第三个问题:对于目标规划问题，整体求解和序贯求解哪种更好？

从优先级的不可逆性上讲，二者是等价的。

1.3 整数规划问题

Definition 1.3.1. 分支定界法

用于求解整数规划问题，思想是对(极大化的)整数规划问题进行线性松弛，求得最大值，然后在分支的整数规划问题的可行解里求得最小值，直到最后收敛。

1. **求解整数规划问题：**

   - 对于整数规划问题 A，首先解其相应的线性规划问题 B。

- **若** B 没有可行解，**则** A 也没有可行解，停止。
- **若** B 有最优解且符合整数条件，**则** B 此时的最优解即为 A 的最优解。
- **若** B 有最优解但不符合整数条件，**则** 记此时 B 的目标函数值为 Z（上界），那么 A 的目标函数值 $z^* \leq Z$。

2. **寻找整数可行解：**
   - 使用观察法或试探法找到一个整数可行解，通常取简单的组合数进行试探，如：$(0,1),(0,2)$ 等。
   - 计算该解的目标函数值，并记其为 Z，那么此时有 $Z \leq z^* \leq Z$。
3. **迭代过程：**
   - 使用分支定界法的迭代过程进行求解，不断对问题进行分支，选择潜在的解空间进行进一步求解。
   - 每次通过分支操作将问题分解为子问题
生成新的约束并更新当前的上界和下界。
   - 在每一分支时，若子问题的下界大于当前的上界，则可舍弃该分支，继续探索其他分支。
   - 最终通过迭代找到整数解 z^*。

Definition 1.3.2. 割平面法

割平面的思想是构造可行割，让原来整数规划问题的整数最优解一定满足，但是切割掉了原来可行域的平面，缩小了搜索范围。

Example 1.3.1. 求解

$$\max \quad z = x_1 + x_2$$

$$\text{s.t.} \begin{cases} -x_1 + x_2 \leq 1 \\ 3x_1 + x_2 \leq 4 \\ x_1, x_2 \geq 0, \text{整数} \end{cases}$$

不考虑整数条件

增加松弛变量

$$\max \quad z = x_1 + x_2 + 0x_3 + 0x_4$$

$$\text{s.t.}\begin{cases} -x_1+x_2+x_3=1 \\ 3x_1+x_2+x_4=4 \\ x_1,x_2,x_3,x_4\geq 0 \end{cases}$$

这时对于这个线性规划问题，我们可以进行单纯形法的求解，得到变换后的A。

(为什么不一开始就用添加可行割的方法?一开始你也不知道是哪个变量不是整数)

$$x_1-\frac{1}{4}x_3+\frac{1}{4}x_4=-\frac{3}{4}$$
$$x_2+\frac{3}{4}x_3+\frac{1}{4}x_4=-\frac{7}{4}$$

将系数和常数项均分解成整数与非负真分数之和移项

$$x_1-x_3=\frac{3}{4}-\left(\frac{3}{4}x_3+\frac{1}{4}x_4\right)$$
$$x_2-1=\frac{3}{4}-\left(\frac{3}{4}x_3+\frac{1}{4}x_4\right)$$

Remark. 左边起到整数约束的作用

如果你检测下就会发现，添加的可行割让所有整数解都满足了，但是约束到了原来的线性规划的解。

Example 1.3.2. 如果m个互相排斥的约束条件（≤型）：

$$a_{i1}x_1+a_{i2}x_2+\cdots+a_{in}x_n\leq b_i, \text{其中} i=1,2,\cdots,m$$

为了保证这m个约束条件只有一个起作用，引入m个0-1变量$y_i(i=1,2,\cdots,m)$和一个充分大的数M，从而约束条件可以变为：

$$\begin{cases} y_1+y_2+\cdots+y_m=m-1 \\ a_{i1}x_1+a_{i2}x_2+\cdots+a_{in}x_n\leq b_i+y_iM, \text{其中} i=1,2,\cdots,m \end{cases}$$

其中，M这个充分大的数保证了当$y_i=1$时，对应的约束条件是多余的。

Remark. 0-1型整数规划

这类问题我们一般使用隐枚举法(可以大致理解为特殊的分支定界法)

何老师的问题:隐枚举法这样操作是否还能提速?

可以的兄弟!只要调换枚举顺序就可以了。先枚举$X = [0, 0, \cdots, 0]^T$，再从价格系数高的开始枚举就可以了。

Definition 1.3.3. 指派问题标准形式

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

$$\text{s.t.} \quad \sum_{j=1}^{n} x_{ij} = 1 \quad \text{每个人有且只有一项工作}$$

$$\sum_{i=1}^{n} x_{ij} = 1 \quad \text{每项工作有且只有一个人}$$

$$x_{ij} = 0\text{或}1 \quad i, j = 1, 2, \ldots, n;$$

指派问题是特殊的运输问题

对于指派问题，我们使用匈牙利算法来求解。基本思想是，每一行对应一个产地，每一列对应一个销地，先对每一行进行减法操作，然后对每一列进行减法操作，最后得到一个矩阵。

这个新的成本矩阵和原矩阵最优解一样(每条路都打折，等于没打折)

Definition 1.3.4. 独立零元素

位于不同行、不同列的零元素，称为独立零元素。

系数矩阵C中独立零元素的个数最多等于能覆盖所有零元素的最少直线数。

我们可以用独立零元素(实际上只要有一个基准值就行了，不用是零)，如果独立零元素个数等于n，说明我们找到了一个最优解。

Remark. 可能有多个最优解

第二章　非线性凸规划

Definition 2.0.1. 非线性规划的数学模型

$$\min f(X)$$

$$\text{s.t.} \quad h_i(X) = 0 \quad i = 1, 2, \ldots, m$$

$$g_j(X) \geq 0 \quad j = 1, 2, \ldots, l$$

$$X \in R^n$$

何老师的问题:非线性规划找的是局部最小点还是全局最小点?

实际上对于非线性凸规划，我们找到的局部最小点就是全局最小点。对于复杂的非线性规划，实际上我们不确定找到的是局部最小点还是全局最小点。

Definition 2.0.2. 凸函数与凹函数

凸函数为满足下列条件的函数:

$$f[\lambda X_1 + (1-\lambda)X_2] \leq \lambda f(X_1) + (1-\lambda)f(X_2) \quad 0 < \lambda < 1 \quad X_1, X_2 \in R^n$$

Proof. 证明凸函数的局部最小点就是全局最小点

对于任意 $\theta \in [0, 1]$，验证：

$$f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta)f(y)$$

展开左边并整理后，比较两边的差值为：

$$\frac{\theta(1-\theta)}{2}(x-y)^T A(x-y) \leq 0$$

由于 A 半正定且 $\theta(1-\theta) \geq 0$，不等式成立，故 f 是凸函数。 □

凸函数的性质

1. 凸函数非负线性组合仍为凸函数

$$\gamma_1 f_1(X) + \gamma_2 f_2(X) \quad \gamma_i \geq 0$$

2. 若 $f(X)$ 是定义在凸集 R_C 上的凸函数，则其 β 水平集 S_β 为凸集。

$$S_\beta = \{X \mid f(X) \leq \beta, X \in R_C\} \quad \beta\text{水平集}$$

3. 对于凸函数 $f(X)$，若存在 $X^* \in R^n$ 满足

$$\nabla f(X^*)^T (X - X^*) \geq 0 \quad \forall X \in R^n$$

则 X^* 为 $f(X)$ 的全局最小点。充要条件!

(1)(3)显然,(3)先证明是局部最小点，然后证明的全局最小点。

对于(2)，直接使用定义是可以证明的。

Definition 2.0.3. 凸规划问题

$$\min f(X)$$

$$\text{s.t. } h_i(X) = 0 \quad i = 1, 2, ..., m \quad \text{线性函数}$$

$$g_j(X) \geq 0 \quad j = 1, 2, ..., l \quad \text{凹函数}$$

$$X \in R^n$$

实际上第二个凹函数的要求，如果是左右同乘一个负号，实际上是一个凸函数。所以这个规划问题的可行域也是个凸集(线性函数是凸函数也是凹函数)。

Theorem 2.0.1. 若目标函数为严格凸函数，则如果全局最优解存在必为唯一全局最优解。（反证法）

如果不是严格，我们用定义可以说明凸函数的最优解解集一定是各个最优解的凸包

Definition 2.0.4. 海森矩阵

$$\nabla^2 f(X) = \begin{bmatrix} \frac{\partial^2 f(X)}{\partial x_1^2} & \frac{\partial^2 f(X)}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f(X)}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f(X)}{\partial x_2 \partial x_1} & \frac{\partial^2 f(X)}{\partial x_2^2} & \cdots & \frac{\partial^2 f(X)}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f(X)}{\partial x_n \partial x_1} & \frac{\partial^2 f(X)}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f(X)}{\partial x_n^2} \end{bmatrix}$$

$$f(X+\Delta X)=f(X)+\nabla f(X)^T\Delta X+\frac{1}{2}\Delta X^T\nabla^2 f(X)\Delta X+O(||\Delta X||^2)$$

如果某处梯度为0，海森矩阵为半正定，实际上这只是局部极小值的必要条件。除非加强为领域梯度均大于等于0。

Definition 2.0.5. 拉格朗日乘数法(针对等式约束):

定义Lagrange函数

$$\min L(X,\lambda)=f(X)+\lambda^T h(X)=f(X)+\sum_{i=1}^{m}\lambda_i h_i(X)$$

$$\lambda=\begin{bmatrix}\lambda_1 & \lambda_2 & \cdots & \lambda_m\end{bmatrix}^T$$

求无约束极值问题：

$$\left.\frac{\partial L(X,\lambda)}{\partial X}\right|_{X^*}=0 \quad \Rightarrow \quad \nabla f(X^*)+\sum_{i=1}^{m}\lambda_i\nabla h_i(X^*)=0$$

$$\left.\frac{\partial L(X,\lambda)}{\partial \lambda}\right|_{\lambda^*}=0 \quad \Rightarrow \quad h_i(X^*)=0 \quad i=1,2,\ldots,m$$

不等式约束极值问题:

$$\min f(X)$$

$$\text{s.t.} \quad g_j(X)\geq 0 \quad j=1,2,\ldots,l$$

$$X\in R^n$$

Definition 2.0.6. 定义：令目标函数值 $f(X^{(0)})$ 下降的方向，满足

$$\exists\lambda_0>0,\ \text{对于}\ \forall\lambda\in(0,\lambda_0],\ \text{有}\ f(X^{(0)}+\lambda P)<f(X^{(0)})$$

判别条件: $\nabla f(X^{(0)})^T P<0$

Theorem 2.0.2. Gordan 引理

A_j 为一组已知向量，不存在向量 P，使

$$A_j^T P<0 \quad j=1,2,...,l$$

同时成立的充要条件：

> 存在不全为零的非负实数 $\mu_j \geq 0$，使
>
> $$\sum_{j=1}^{l} \mu_j A_j = 0$$
>
> 几何含义：A_j 不可能分布在任何超平面的同一侧。

Proof. 用Farkas引理证明Gordan引理

Farkas引理：对于矩阵$A \in \mathbb{R}^{m\times n}$和向量$b \in \mathbb{R}^m$，以下两个系统恰有一个有解：

1. 存在$x \geq 0$使得$Ax = b$;
2. 存在$y \in \mathbb{R}^m$使得$y^\top A \geq 0^\top$且$y^\top b < 0$。

证明Gordon引理：

1. **构造辅助系统：**

   将系统A加强为存在满足归一化条件的解。令 $e = (1, 1, \ldots, 1)^\top$，定义新系统：

   $$\exists x \geq 0, \quad \begin{cases} Ax = 0 \\ e^\top x = 1 \end{cases}$$

   若此系统有解，则原系统A有非零解；反之则无。

2. **应用Farkas引理：**

   将上述系统写作矩阵形式 $\begin{bmatrix} A \\ e^\top \end{bmatrix} x = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$。根据Farkas引理，恰有以下之一成立：

   (a) 存在 $x \geq 0$ 满足方程

   (b) 存在 $\begin{bmatrix} y \\ \lambda \end{bmatrix} \in \mathbb{R}^{m+1}$ 使得：

   $$y^\top A + \lambda e^\top \geq 0^\top \quad 且 \quad y^\top 0 + \lambda \cdot 1 < 0$$

   第二个不等式简化为 $\lambda < 0$。

3. **分析对偶条件：**

   若原系统无解，则存在 y 和 $\lambda < 0$ 使得：

   $$y^\top A_j + \lambda \geq 0 \quad (\forall j = 1, \ldots, n)$$

   令 $\mu = -\lambda > 0$，则有：

   $$y^\top A_j \geq \mu > 0 \quad (\forall j)$$

即 $A^\top y > 0$，故系统B成立。

4. **反方向证明：**

若系统B有解（存在 y 使 $A^\top y > 0$），假设系统A也有解（存在非零 $x \geq 0$ 使 $Ax = 0$），则：

$$y^\top Ax = (A^\top y)^\top x > 0 \quad (\text{因 } A^\top y > 0 \text{ 且 } x \geq 0, \neq 0)$$

但左边等于 $y^\top 0 = 0$，矛盾。故系统B存在时系统A无解。

综上，系统A与系统B恰有一个成立，Gordon引理得证。 □

Theorem 2.0.3. Fritz John条件

若X*是局部极小点，存在不全为零的非负实数 $\mu_j \geq 0, j \in J(X*) \cup \{0\}$，使

$$\mu_0 \nabla f(X*) - \sum_{j \in J} \mu_j \nabla g_j(X*) = 0$$

Proof. 证明：Gordan引理的直接推论(不存在一个可行的下降方向)

$$\nabla f(X*)^T P < 0$$

$$-\nabla g_j(X*)^T P < 0$$

□

Remark. 等价Fritz John条件(将J显式化)

若X^*是局部极小点，则存在不全为零的非负实数 $\mu_j \geq 0$，$j = 0, 1, \ldots, l$，使下列条件成立：

$$\begin{cases} \mu_0 \nabla f(X^*) - \sum_{j=1}^{l} \mu_j \nabla g_j(X^*) = 0 & \text{Lagrange驻点条件} \\ \mu_j g_j(X^*) = 0 & j = 1, 2, \ldots, l \quad \text{互补松弛条件} \\ \mu_j \geq 0 & \sum_{j=0}^{m} \mu_j \neq 0 \quad j = 0, 1, \ldots, l \quad \text{强非负条件} \end{cases}$$

上述条件隐含了如下事实：$\mu_j = 0 \quad j \notin J$

Definition 2.0.7. Kuhn-Tucker条件

若Fritz John条件的$\mu_0 \geq 0$，则可推出KKT条件的数学形式：

$$\begin{cases} \nabla f(X^*) - \sum_{j=1}^{l} \mu_j^* \nabla g_j(X^*) = 0 & \text{Lagrange驻点条件} \\ \mu_j^* g_j(X^*) = 0 & j = 1, 2, \ldots, l \quad \text{互补松弛条件} \\ \mu_j^* \geq 0 & j = 1, 2, \ldots, l \quad \text{非负条件} \end{cases}$$

Karush(1939), Kuhn-Tucker (1951)

何老师的问题:什么条件下$\mu_0 \geq 0$?

在加上约束规格限制时$\mu_0 \geq 0$，比如说正则条件

- **Fritz John点**：满足Fritz John条件（允许目标函数乘子为零）
- **KKT点**：满足KKT条件（需约束规格，乘子非负）
- **正则KKT点**：满足LICQ或严格互补的KKT点
- **极值点**：局部最优解
  - **包含关系**：

    $$\text{Fritz John点} \supsetneq \text{KKT点} \supsetneq \text{正则KKT点}$$

  - **极值点归属**：
    - 若满足约束规格：极值点 $\subset$ KKT点
    - 若不满足约束规格：极值点 $\subset$ Fritz John点
  - **非极值点存在性**：
    - 存在非极值的KKT点（如鞍点）
    - 存在非极值的Fritz John点（目标乘子为零）

等价不等式约束极值问题（等式约束改写成不等式）

$$\min f(X)$$

$$\text{s.t.} \quad h_i(X) \geq 0 \quad i = 1, 2, \ldots, m$$

$$-h_i(X) \geq 0 \quad i = 1, 2, \ldots, m$$

$$g_j(X) \geq 0 \quad j = 1, 2, \ldots, l$$

$$X \in R^n$$

问题：此时约束是否满足正则条件？

显然不满足，等式约束的梯度函数线性相关

一般约束的KKT条件的推导

$$\mu_0^* \nabla f(X^*) - \sum_{j=1}^{l} \mu_j^* \nabla g_j(X^*) - \sum_{i=1}^{m} \mu_i'^* \nabla h_i(X^*) + \sum_{i=1}^{m} \mu_i''^* \nabla h_i(X^*) = 0$$

$$\mu_0^* \nabla f(X^*) - \sum_{j=1}^{l} \mu_j^* \nabla g_j(X^*) - \sum_{i=1}^{m} (\mu_i'^* - \mu_i''^*) \nabla h_i(X^*) = 0$$

$$\mu_0^* \nabla f(X^*) - \sum_{j=1}^{l} \mu_j^* \nabla g_j(X^*) - \sum_{i=1}^{m} \gamma_i^* \nabla h_i(X^*) = 0$$

$$\mu_i'^* \geq 0 \quad \mu_i''^* \geq 0 \quad \Rightarrow \quad \gamma_i^* = \mu_i'^* - \mu_i''^* \text{无符号约束} \quad i = 1, 2, \ldots, p$$

注意 μ_0, μ_j, γ_i 不可同时为0!

Definition 2.0.8. 一般约束下的Fritz John条件

若X*是局部极小点，存在不全为零的μj (j=0,1,2,...,m)和γi (i=0,1,2,...,p)，满足

$$\mu_0^* \nabla f(X^*) - \sum_{j=1}^{l} \mu_j^* \nabla g_j(X^*) - \sum_{i=1}^{m} \gamma_i^* \nabla h_i(X^*) = 0 \quad \text{Lagrange驻点条件}$$

$$\mu_j^* g_j(X^*) = 0 \quad j = 1, 2, ..., m \quad \text{互补松弛条件}$$

$$\mu_j^* \geq 0 \quad j = 0, 1, ..., m$$

$$\sum_{j=0}^{l} \mu_j^* + \sum_{i=1}^{m} |\gamma_i^*| \neq 0 \quad \text{强非负条件}$$

$$\gamma_i^* h_i(X^*) = 0 \quad i = 1, 2, ..., p \quad \text{等式互补松弛条件}$$

Definition 2.0.9. KKT条件的矩阵形式

$$\nabla f(X^*) - \nabla h(X^*) \gamma^* - \nabla g(X^*) \mu^* = 0$$

$$\mu^* \odot g(X^*) = 0 \quad \Leftrightarrow \quad \mu_j^* g_j(X^*) = 0 \quad j = 1, 2, \ldots, l$$

其中

$$g(X^*) \geq 0$$

$$h(X^*) = 0$$

$$\mu^* \geq 0$$

Theorem 2.0.4. 凸规划下的KKT条件为最优解的充要条件

Proof. KKT条件的充分性证明

$$\nabla f(X^*) - \sum_{i=1}^{m} \gamma_i^* \nabla h_i(X^*) - \sum_{j=1}^{l} \mu_j^* \nabla g_j(X^*) = 0 \quad \mu_j^* g_j(X^*) = 0 \quad \mu_j^* \geq 0$$

$f(X)$是凸函数 $\Rightarrow f(X) \geq f(X^*) + \nabla f(X^*)^T(X - X^*)$

$$= f(X^*) + \sum_{i=1}^{m} \gamma_i^* \nabla h_i(X^*)^T(X - X^*) + \sum_{j=1}^{l} \mu_j^* \nabla g_j(X^*)^T(X - X^*)$$

$h_i(X)$线性函数且 $h_i(X) = 0 \Rightarrow \nabla h_i(X^*)^T(X - X^*) = h_i(X) - h_i(X^*) = 0$ $g_j(X)$是凹函数 $\Rightarrow \nabla g_j(X^*)^T(X - X^*) \geq g_j(X) - g_j(X^*)$ 根据互补松弛性条件 $\mu_j^* g_j(X^*) = 0$ 且 $g_j(X) \geq 0$，有

$$f(X) \geq f(X^*) + \sum_{j=1}^{l} \mu_j^* \left[g_j(X) - g_j(X^*)\right] = f(X^*) + \sum_{j=1}^{l} \mu_j^* g_j(X) \geq f(X^*)$$

□

第三章　动态规划

Definition 3.0.1. 阶段：

按时间、空间的特征分解成若干顺序联系的阶段。

Definition 3.0.2. 状态：

k阶段开始（或结束）时的客观条件，记为$s_k \in S_k$。S_k为k阶段状态集合。

Definition 3.0.3. 决策（同时依赖于阶段和状态）:

依据状态做出的决定，记为$u_k(s_k) \in D_k(s_k)$，$D_k(s_k)$为状态s_k的允许决策集合。

Definition 3.0.4. 状态的转移方程:

$$s_{k+1} = T_k(s_k, u_k(s_k))$$

Definition 3.0.5. 策略

各阶段决策依次构成的决策序列。记为

$$p_{1,n} = \{u_1(s_1), u_2(s_2), \ldots, u_n(s_n)\} \in P$$

P为允许策略集合。

Remark. 动态规划要求问题具有无后效性：

给定某阶段的状态s_k，则以后各阶段的状态s_l（$l > k$）都只受s_k的影响，与之前的状态无关。

Definition 3.0.6. 子过程与子策略

后部子过程策略，从k阶段开始到终了阶段的决策子序列，记为

$$p_{k,n}(s_k) = \{u_k(s_k), u_{k+1}(s_{k+1}), \ldots, u_n(s_n)\} \in P_{k,n}(s_k)$$

Definition 3.0.7. 子策略的指标函数

评价沿子策略$p_{k,n}$过程性能优劣的函数，记为$V_{k,n}(s_k, p_{k,n})$。

为了实现动态规划的递推结构，要求指标函数具有可分离性

$$V_{k,n}(s_k, p_{k,n}) = \phi_k(s_k, u_k, V_{k+1,n}(s_{k+1}, p_{k+1,n}))$$

ϕ_k 是 $V_{k+1,n}(s_{k+1}, p_{k+1,n})$ 的严格单调函数。

Remark. 最优化原理:

最优策略的子策略是相应子问题的最优策略（否则由策略的无后效性，这一定不是最优）。

Theorem 3.0.1. 最优化定理

策略 $p^*_{1,n}$ 是最优策略的充要条件是，对于所有的 k，都有：

$$V_{1,n}(s_1, p^*_{1,n}) = \operatorname*{opt}_{p_{1,k-1}\in P_{1,k-1}} \left(V_{1,k-1}(s_1, p_{1,k-1}) + \operatorname*{opt}_{p_{k,n}\in P_{k,n}} V_{k,n}(s_k, p_{k,n}) \right)$$

第四章　网络优化理论

Definition 4.0.1. 网络

G(图)=（V(节点)，E(弧)）

Definition 4.0.2. 次

$d(v)$:以顶点v为端点的边的个数

Definition 4.0.3. 树时一个连通无圈无向的图

树T = (V, E)，$|V| = n$,$|E| = m$，则下列说法等价：

(1) T是一个树

(2) T无圈，且m=n-1

(3) T连通，且m=n-1

(4) T无圈，但每加一个新边，可得唯一的圈

(5) T连通，但每舍去一条边即不连通

(6) T中任意两点，有唯一的链相连

实际上可以管这个叫最小生成树，因为你找不出更小一个连通无圈无向的图了。对于一个图的最小生成树，我们可以采用破圈法。

Definition 4.0.4. 生成子图：给定图 $G = (V, E)$，其生成子图 $G' = (V', E')$ 满足以下条件：

- $V' \subseteq V$：子图的顶点集合是原图顶点集合的子集；
- $E' \subseteq E$：子图的边集合是原图边集合的子集；
- E' 仅包含连接 V' 中顶点的边。

Remark. 几个概念的辨析

1. 链:无方向要求，可以反向链接
2. 道路:有方向要求，不能反向链接
3. 圈:顾名思义，但是可以反向链接
4. 回路:有方向要求的圈，不能反向链接

1. 对于所有权值都非负的图，最短路可用dijkstra算法求解
2. 对于含负权值的图，最短路可用逐次逼近法求解
3. 对于任意两点最短距离，可以使用Floyd算法求解

何老师的问题:对于逐次逼近法，最多需要几次迭代。

如果不能提早收敛的话，得要n-1次（n是节点个数）

求解最短路的两种思路:

思路1：两点间用逐次逼近法求解（这也解释了为什么迭代某次发现不变之后就是收敛了）

$$d_{ij}^{(k)} = \min_{l}\{d_{il}^{(k-1)} + l_{lj}\} \quad l = 1, 2, \ldots, n \quad 1 \leq k \leq n$$

思路2：l_{ij}用上一步迭代结果替代

$$d_{ij}^{(k)} = \min_{l}\{d_{il}^{(k-1)} + d_{lj}^{(k-1)}\} \quad l = 1, 2, \ldots, n \quad 1 \leq k \leq n$$

这就是Floyd算法的思路

第五章　对策论

Definition 5.0.1. **博弈**是指在一个特定的环境中，多个参与者（称为博弈者）通过选择策略来影响结果的过程。博弈论研究这些博弈的性质、策略选择和结果。

博弈模型 $G = \{I, II; S_1, S_2; A\}$

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

如果存在

$$\max_i \min_j a_{ij} = \min_j \max_i a_{ij} = a_{i^*j^*} = V_G$$

则 $(\alpha_{i^*}, \beta_{j^*})$ 为矩阵博弈的最优纯策略对，也称为最优局势。V_G 称为博弈值。

矩形博弈最优纯策略对存在的充要条件是存在鞍点。

$$a_{ij^*} \leq a_{i^*j^*} \leq a_{i^*j} \quad \text{(鞍点条件)}$$

$$\max_i a_{ij^*} \leq a_{i^*j^*} \leq \min_j a_{i^*j}$$

Proof. **必要性**

$$\max_i \min_j a_{ij} = \min_j \max_i a_{ij}$$

$$\exists i^*, j^* \quad \min_j a_{i^*j} = \max_i \min_j a_{ij} = \min_j \max_i a_{ij} = \max_i a_{ij^*}$$

$$a_{i^*j^*} \geq \min_j a_{i^*j} = \max_i \min_j a_{ij} = \min_j \max_i a_{ij} = \max_i a_{ij^*} \geq a_{i^*j^*}$$

$$\min_j a_{i^*j} = \max_i a_{ij^*} = a_{i^*j^*} \quad \Rightarrow \quad a_{ij^*} \leq a_{i^*j^*} \leq a_{i^*j}$$

充分性

$$a_{ij*} \le a_{i*j*} \le a_{i*j}$$

$$\max_i a_{ij*} \le a_{i*j*} \le \min_j a_{i*j}$$

$$\min_j \max_i a_{ij} \le a_{i*j*} \le \max_i \min_j a_{ij}$$

又根据

$$\max_i \min_j a_{ij} \le \min_j \max_i a_{ij}$$

$$\max_i \min_j a_{ij} = \min_j \max_i a_{ij}$$

□

当我们找到这样一组决策时,对于任何其他决策和双方决策者,我们都不能得到更好的结果,这就是鞍点的意思。

Remark. 纳什均衡解 $\leftrightarrow$ 最优纯策略对

多个均衡解的性质:

- 无差别性
- 可交换性
- 存在无策略解，此时选择随机策略

Definition 5.0.2. 博奕模型 $G^* = \{S_1^*, S_2^*; E\}$

▷ 混合策略集

$$S_1^* = \left\{ x \in R^m \mid x_i \ge 0,\ \ i = 1, 2, \cdots, m; \sum_{i=1}^{m} x_i = 1 \right\}$$

x_i 为局中人I执行纯策略 α_i 的概率

$$S_2^* = \left\{ y \in R^n \mid y_j \ge 0,\ \ j = 1, 2, \cdots, n; \sum_{j=1}^{n} y_j = 1 \right\}$$

y_j 为局中人II执行纯策略 β_i 的概率

▷ 局中人I的赢得函数: $E(x, y) = x^T Ay = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} x_i y_j$

▷ 局中人II的赢得函数: $-E(x, y)$

如果双方都是理性决策

- 局中人I的最大预期赢得:

$$\max_{x \in S_1} \min_{y \in S_2} E(x, y)$$

- 局中人II的最小预期损失:

$$\min_{y \in S_2} \max_{x \in S_1} E(x, y)$$

两者关系:

$$\max_{x \in S_1} \min_{y \in S_2} E(x, y) \leq \min_{y \in S_2} \max_{x \in S_1} E(x, y)$$

混合策略

$$\mathbf{x} = [x_1, x_2, \cdots, x_m]^T \quad \mathbf{y} = [y_1, y_2, \cdots, y_n]^T$$

Theorem 5.0.1. 最优混合策略对/均衡解存在的充要条件：存在鞍点

$$E(x, y^*) \leq E(x^*, y^*) \leq E(x^*, y)$$

Proof. 一定存在混合策略对意义下的矩阵博弈均衡解

∃x*,y*满足:

$$\sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} y_j^* x_i = E(x, y^*) \leq E(x^*, y^*) \leq E(x^*, y) = \sum_{j=1}^{n} \sum_{i=1}^{m} a_{ij} x_i^* y_j^*$$

鞍点条件

$$\sum_{j=1}^{n} a_{ij} y_j^* \leq E(x^*, y^*) \leq \sum_{i=1}^{m} a_{ij} x_i^* \quad \forall i = 1, 2, \cdots, m \quad \forall j = 1, 2, \cdots, n$$

一方采用纯策略，一方采用最优混合策略

$$\exists v = E(x^*, y^*)$$

使下面2组不等式均有解

$$\sum_{i=1}^{m} a_{ij} x_i^* \geq v \quad j = 1, 2, \cdots, n$$

$$\sum_{i=1}^{m} x_i^* = 1 \quad x_i^* \geq 0 \quad i = 1, 2, \cdots, m$$

$$\sum_{j=1}^{n} a_{ij} y_j^* \le v \quad i = 1, 2, \cdots, m$$

$$\sum_{j=1}^{n} y_j^* = 1 \quad y_j^* \ge 0 \quad j = 1, 2, \cdots, n$$

使得对偶问题均有可行解，且 $w^* = v^* = E(x^*, y^*)$

可以验证

$$\mathbf{x} = [1 \quad 0 \quad \cdots \quad 0] \quad w = \min a_{1j}$$

$$\mathbf{y} = [1 \quad 0 \quad \cdots \quad 0] \quad v = \max a_{i1}$$

原问题有可行解同时对偶问题也有可行解

故有

$$\sum_{j=1}^{n} a_{ij} y_j^* = v^* = w^* = \sum_{i=1}^{m} a_{ij} x_i^* \quad \text{强对偶性}$$

又根据：

$$E(\mathbf{x}^*, \mathbf{y}^*) = \sum_{j=1}^{n} \sum_{i=1}^{m} a_{ij} x_i^* y_j^* = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij} y_j^* \right) x_i^* \le v^* \sum_{i=1}^{m} x_i^* = v^*$$

$$E(\mathbf{x}^*, \mathbf{y}^*) = \sum_{j=1}^{n} \sum_{i=1}^{m} a_{ij} x_i^* y_j^* = \sum_{j=1}^{n} \left(\sum_{i=1}^{m} a_{ij} x_i^* \right) y_j^* \ge w^* \sum_{j=1}^{n} y_j^* = w^*$$

$$w^* = v^* = E(\mathbf{x}^*, \mathbf{y}^*)$$

□

Theorem 5.0.2. 从两个互为对偶的线性规划问题里，我们有互补松弛性

$$x_i^* > 0 \implies \sum_{j=1}^{n} a_{ij} y_j^* = v^* = E(x^*, y^*)$$

$$y_j^* > 0 \implies \sum_{i=1}^{m} a_{ij} x_i^* = w^* = E(x^*, y^*)$$

$$\sum_{j=1}^{n} a_{ij} y_j^* < v^* = E(x^*, y^*) \implies x_i^* = 0$$

$$\sum_{i=1}^{m} a_{ij} x_i^* > w^* = E(x^*, y^*) \implies y_j^* = 0$$

如果某条纯策略在均衡解中有被选择的可能，则对手的最优混合策略在该纯策略下的赢得值不会比 V_G 更好。

如果某条纯策略下对手的最优混合策略的赢得值比 V_G 更好，则该纯策略在均衡解中无被选择可能。

解集T(G)：博弈G的均衡解集合。

赢得矩阵严格单调变换下的解集不变性

博弈：

$$G_1 = \{S_1, S_2; A_1\}$$

$$G_2 = \{S_1, S_2; A_2\}$$

$$A_2 = A_1 + L * 1_{m\times n} \quad \Longrightarrow \quad T(G_1) = T(G_2) \quad V_{G_1} = V_{G_2} + L$$

$$A_2 = aA_1 \quad a > 0 \quad \Longrightarrow \quad T(G_1) = T(G_2) \quad V_{G_1} = aV_{G_2}$$

Proof. 证明：上述变换只改变了赢得矩阵元素的数值，不改变相对大小关系。 □

如果博弈问题具有如下对称性：

$$\mathbf{A} = -\mathbf{A}^T$$

自身角度的赢得矩阵相同

$$T_{\mathrm{I}}(G) = T_{\mathrm{II}}(G) \quad a_{ij} = \begin{cases} -a_{ji} & i \neq j \\ 0 & i = j \end{cases}$$

$$V_G = E(\mathbf{x}^*, \mathbf{x}^*) = \sum_{j=1}^{n}\sum_{i=1}^{m} a_{ij}x_i^*x_j^* = -V_G = 0$$

这场游戏没有赢家，可以近似的理解为双方的损失相同。

互补松驰性方程组

$$x_i^* > 0 \quad \Rightarrow \quad \sum_{j=1}^{n} a_{ij}y_j^* = v^* = E(x^*, y^*) \quad i \in \{1, 2, \cdots, m\}$$

$$y_j^* > 0 \quad \Rightarrow \quad \sum_{i=1}^{m} a_{ij}x_i^* = w^* = E(x^*, y^*) \quad j \in \{1, 2, \cdots, n\}$$

$$\sum_{i=1}^{m} x_i^* = 1$$

$$\sum_{j=1}^{n} y_j^* = 1$$

可以理解为对于均衡解，每条纯策略在对方的混合策略下都是相等且最优的。

Example 5.0.1. 田忌赛马

$$\mathbf{A} = \begin{bmatrix} -3 & -1 & 1 & -1 & -1 & -1 \\ -1 & -3 & -1 & 1 & -1 & -1 \\ -1 & -1 & -3 & -1 & 1 & -1 \\ -1 & -1 & -1 & -3 & -1 & 1 \\ 1 & -1 & -1 & -1 & -3 & -1 \\ -1 & 1 & -1 & -1 & -1 & -3 \end{bmatrix}$$

任何一条纯策略均是对方特定策略下的唯一赢得策略，**也就是说都有赢的可能**，故有

$$x_i^* > 0 \quad i = 1, 2, \cdots, m$$

$$y_j^* > 0 \quad j = 1, 2, \cdots, n$$

Remark. **利用线性规划求解博弈问题(需要先假设w,v≥0)**

$$\begin{aligned} \max \quad & w \\ \text{s.t.} \quad & \sum_{i=1}^{m} a_{ij} x_i \geq w \quad j = 1, 2, \cdots, n \\ & \sum_{i=1}^{m} x_i = 1 \\ & x_i \geq 0 \quad i = 1, 2, \cdots, m \end{aligned} \tag{5.1}$$

$$x_i' = \frac{x_i}{w} \tag{5.2}$$

$$
\begin{aligned}
\min \quad & \sum_{i=1}^{m} x_i' = \frac{1}{w} \\
\text{s.t.} \quad & \sum_{i=1}^{m} a_{ij} x_i' \geq 1 \quad j = 1, 2, \cdots, n \\
& x_i' \geq 0 \quad i = 1, 2, \cdots, m
\end{aligned}
\tag{5.3}
$$

$$
\begin{aligned}
\min \quad & v \\
\text{s.t.} \quad & \sum_{j=1}^{n} a_{ij} y_j \leq v \quad i = 1, 2, \cdots, m \\
& \sum_{j=1}^{n} y_j = 1 \\
& y_j \geq 0 \quad j = 1, 2, \cdots, n
\end{aligned}
\tag{5.4}
$$

$$
y_j' = \frac{y_j}{v}
\tag{5.5}
$$

$$
\begin{aligned}
\max \quad & \sum_{j=1}^{n} y_j' = \frac{1}{v} \\
\text{s.t.} \quad & \sum_{j=1}^{n} a_{ij} y_j' \leq 1 \quad i = 1, 2, \cdots, m \\
& y_j' \geq 0 \quad j = 1, 2, \cdots, n
\end{aligned}
\tag{5.6}
$$

囚徒困境的纯策略解

$$
A = \begin{bmatrix} -9 & 0 \\ -15 & -1 \end{bmatrix}, \quad B = \begin{bmatrix} -9 & -15 \\ 0 & -1 \end{bmatrix}
$$

局中人I选择:

$$
j_i^*(B) = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \Rightarrow \quad i_j^*(A) = 1 \quad \Rightarrow \quad (i^*, j^*) = (1, 1)
$$

局中人II选择:

$$
i_j^*(A) = \begin{bmatrix} 1 & 1 \end{bmatrix} \quad \Rightarrow \quad j_i^*(B) = 1 \quad \Rightarrow \quad (i^*, j^*) = (1, 1)
$$

Definition 5.0.3. Nash均衡点

满足以下条件的策略对 $(\alpha_{i*}, \beta_{j*})$

$$a_{i*j*} \geq a_{ij*} \quad i = 1, 2, \cdots, m$$

$$b_{i*j*} \geq b_{i*j} \quad j = 1, 2, \cdots, n$$

没有一个局中人愿意单方面改变策略