1 绪 论

层次优化问题一般泛指双层规划、多层规划、双层变分不等式、双层均衡等具有层次结构的优化问题，该研究广泛应用于电力系统优化、交通规划、供应链管理、云市场定价策略等领域. 本章主要介绍双层规划、三层规划和双层变分不等式三类具有嵌套结构的层次优化问题的相关最优性理论与算法.

1.1 双层规划问题

双层规划（BP）模型于 1973 年被 Bracken 和 McGill 系统研究[17]，近年来有许多这方面的研究成果[13,25,31,34,60,61,67,75,88,98,100,101]. 该模型的上下层之间具有嵌套结构，可用于描述具有主从关系的两个决策者决策博弈的过程. 上层决策者（领导）首先做出决策，继而下层决策者（随从）做出最优决策并反馈给上层. 双层规划模型可描述为

$$(BP)\text{“}\min_{x}\text{”}\left\{F(x,y):x\in X,y\in\psi(x)\right\},\tag{1.1}$$

$\psi(x)$ 为下述参数优化问题的最优解集：

$$\min_{y}\left\{f(x,y):y\in K(x)\right\}.\tag{1.2}$$

这里 F， f 是 $R^n\times R^m$ 上的实值连续函数； X 和 $K(x)$ 分别代表上层和下

层约束集，可定义为

$$X := \left\{x \in R^n : G(x) \leqslant 0\right\}, K(x) := \left\{y \in R^m : g(x,y) \leqslant 0\right\}, \tag{1.3}$$

其中，$G: R^n \to R^k; g: R^n \times R^m \to R^p$. 这里“$\min\limits_x$”之所以加上引号是因为当下层最优解集 $\psi(x)$ 不止一个元素时，若 y 取 $\psi(x)$ 中的不同值，则 $F(x,y)$ 关于 x 的极小化问题面临不同选择. 也就是说，该问题是不适定的，一般可将其转化为乐观双层规划问题和悲观双层规划问题.

- 乐观双层规划模型：

$$\min_x \left\{\varphi_o(x) : x \in X\right\}, \tag{1.4}$$

其中

$$\varphi_o(x) = \min_y \left\{F(x,y) : y \in \psi(x)\right\}. \tag{1.5}$$

由乐观双层规划模型（1.4）（1.5）可知，上层决策者认为下层决策者在众多最优策略中总选择对自己最有利的决策，即下层决策者与上层决策者完全合作.

在实际应用中还存在乐观双层规划的另一种表述形式：

$$\min_{x,y} \left\{F(x,y) : x \in X, y \in \psi(x)\right\}. \tag{1.6}$$

但这两种表述形式并不总是等价的. Dempe 等在文献[28，33]中指出两种乐观双层规划模型在全局解的情形下是等价的，但在局部解的情形还需假设解集映射

$$S_0(x) := \arg\min_y \left\{F(x,y) : y \in \psi(x)\right\}.$$

在问题（1.4）-（1.5）的最优解处是内半连续的，这两种表述形式才是等价的.

对于乐观双层规划模型（1.6）-（1.2），目前常见的处理方式是将其转化为相对容易求解的单层规划的问题. 现有的单层转变方式主要有如下三种：

1. KKT 条件法

利用下层的 KKT 条件将双层规划（1.6）-（1.2）转化为如下单层规划：

$$\min_{x,y,\lambda} \ F(x,y),$$

$$\text{s. t.}\begin{cases} G(x)\leqslant 0, \\ \nabla_y f(x,y)+\lambda^{\cdot}\nabla_y g(x,y)=0, \\ g(x,y)\leqslant 0, \lambda\geqslant 0, \\ \lambda^{\cdot} g(x,y)=0. \end{cases}$$

通常称此规划为均衡约束规划（MPCC）. MPCC 问题虽被研究的时间不长，但其相关理论与算法都相对比较完善. 遗憾的是即使在下层为凸规划的情形下，这种转化方式都不总是等价的，还需下层满足某些约束规格，比如 Slater 约束规格. 基于此转化方式，不少学者对原双层规划问题的理论与算法进行了研究，参见文献[3, 34, 90].

2. 最优值函数法

利用下层最优值函数将原问题（1.6）-（1.2）转化为如下的单层规划：

$$\min_{x,y} \ F(x,y),$$

$$\text{s.t.}\begin{cases} G(x)\leqslant 0, \\ f(x,y)-\varphi(x)\leqslant 0, \end{cases}$$

其中 $\varphi(x):=\min\limits_{y}\{f(x,y);g(x,y)\leqslant 0\}$.一些学者应用这种转化方式给出了原双层规划问题的最优性条件[28, 97, 102]，并构造了不少求解算法. 然而最优值函数 φ 在一般情况下是不可微的，这为求解算法的构造带来了极大的困难. Lin 等在文献[64]中用积分熵函数近似代替 φ，再结合投影法构造了一个求解算法，并用此算法求解了不少具有实际应用

背景的下层非凸的双层规划问题，可以肯定的是用积分熵函数光滑逼近下层最优值函数是一种较好的处理方式，但由于积分熵函数自身在大规模情形下可能会失效，这就导致以它为基础构造的算法很难求解规模较大的问题.

3. KKT 条件和最优值函数的混合方法

Ye 和 Zhu 在文献[104]中，结合下层的 KKT 条件与最优值函数将原问题（1.6）-（1.2）转化为如下单层规划：

$$\min_{x,y,\lambda} \ F(x,y),$$

$$\text{s.t.}\begin{cases} G(x)\leqslant 0, \\ \nabla_y f(x,y)+\lambda^{\cdot}\nabla_y g(x,y)=0, \\ g(x,y)\leqslant 0,\ \lambda\geqslant 0, \\ \lambda^{\cdot}\ g(x,y)=0, \\ f(x,y)-\varphi(x)\leqslant 0, \end{cases}$$

并基于此转化问题给出了更弱的最优性条件，这种转化方式也为算法的构造提供了一种新的思路. 然而均衡约束和不可微的最优值函数的存在亦是构造算法需要逾越的最大障碍.

- 悲观双层规划模型：

$$\min_x\left\{\varphi_p(x):x\in X\right\}, \tag{1.7}$$

其中
$$\varphi_p(x)=\max_y\left\{F(x,y):y\in\psi(x)\right\}. \tag{1.8}$$

由悲观双层规划模型（1.7）-（1.8）可知，上层决策者认为下层决策者做出了最坏决策，即上、下层完全不合作. 对于该问题，由于上层事实上需求解 min max 问题，因此较上层只需求 min min 问题的乐观模型困难许多. 也正因如此，目前悲观双层规划问题的研究成果比较少.

有一些学者采用类似于乐观情况的下层问题的KKT条件或者最优值函数给出了该问题的最优性条件，参见文献[32, 65]. Wiesemann 在文献[91]中将悲观双层规划问题转化为一系列半无限规划问题，构造了一个求解算法.

1.2 三层规划问题

在电力系统优化[2]、交通规划[14]、供应链管理[107]等问题中经常会出现三层规划模型，对于该模型，各层决策者都希望自己的目标最优，但各层的决策也都会受其他层的决策结果所影响[106]. 有如下的决策过程，首先上层决策者做出决策，然后中间层决策者根据上层的决策信息再做出决策，最后下层决策者再根据上层和中间层决策者的决策情况做出决策，并将结果反馈给上层[89]. 该模型可描述如下：

$$\text{(TLP)} \quad \text{“}\min_{x}\text{”}\, f_1(x,y,z),$$

$$\text{s.t.}\quad g_1(x,y,z)\leqslant 0,\ (y,z)\in\psi(x),$$

这里$\psi(x)$是下述参数优化问题的最优解集：

$$\text{“}\min_{y}\text{”}\, f_2(x,y,z),$$

$$\text{s.t.}\quad g_2(x,y,z)\leqslant 0,\ z\in\omega(x,y),$$

其中$\omega(x,y)$是下述参数优化问题的最优解集：

$$\min\nolimits_z f_3(x,y,z),$$

$$\text{s.t.}\quad g_3(x,y,z)\leqslant 0,$$

这里 $f_i: R^n\times R^m\times R^p\to R, i=1,2,3$ ； $g_i: R^n\times R^m\times R^p\to R^{qi}, i=1,2,3$ （$g_i=(g_i^1,\cdots,g_i^{q_i})^{\cdot}$）；$f_1$，$f_2$和$f_3$分别是上层、中间层和下层的目标函数. 这里

“$\min\limits_{x}$”和“$\min\limits_{y}$”之所以会加上引号，是因为它与双层规划模型类似，即当 $\omega(x,y)(\psi(x))$ 不止一个元素时，中间层（上层）决策者可能不会有足够的影响力使得下层（中间层）决策者选择对中间层（上层）最有利的解 $z\in\omega(x,y)((z,y)\in\psi(x))$. 故类似于双层规划，三层规划亦具有乐观模型和悲观模型.

一类乐观三层规划模型：

$$\min_{x,y,z} f_1(x,y,z),$$

$$\text{s.t.}\quad g_1(x,y,z)\leqslant 0,\ (y,z)\in\psi(x),$$

这里 $\psi(x)$ 是下述参数优化问题的最优解集：

$$\min_{y,z} f_2(x,y,z),$$

$$\text{s.t.}\quad g_2(x,y,z)\leqslant 0,\ z\in\omega(x,y),$$

其中 $\omega(x,y)$ 是下述参数优化问题的最优解集：

$$\min_{z} f_3(x,y,z),$$

$$\text{s.t.}\quad g_3(x,y,z)\leqslant 0.$$

由上述乐观三层规划模型可知，上层决策者认为中间层决策者在众多最优策略中总选择对自己最有利的决策，而中间层决策者同样认为下层决策者会在最优策略中总选择对自己最有利的决策. 即上层与中间层、中间层与下层均为完全合作关系. 目前鲜有关于乐观三层规划的理论研究，现有成果大都集中于线性乐观三层规划的算法研究. 例如，White[89]提出了求解线性三层规划的罚函数方法. Alguacil 等[2]利用下层的 KKT 条件将线性三层规划问题转化为双层规划问题，提出了求解该问题的二阶段方法. Zhang 等[106]提出了求解线性三层规划的 Kth-Best 算法，且在该文中还给出了解存在的充分条件. Huang 等[53]提出了线性多层规划的直觉模糊交互式方法，该算法对三层规划问题求解效果甚佳.

与悲观双层规划类似，对于三层规划模型，上层与中间层、中间层与下层只要有一对为完全非合作关系，则该模型为悲观三层规划. 下面介绍三类悲观模型：

（1）中间层为悲观问题的悲观三层规划模型：

$$\min_{x} \max_{y,z} f_1(x, y, z),$$

$$\text{s.t.} \quad g_1(x) \leqslant 0, (y, z) \in \psi(x),$$

这里 $\psi(x)$ 是下述优化问题的最优解集（假设其有不止一个元素）：

$$\min_{y,z} f_2(x, y, z),$$

$$\text{s.t.} \quad g_2(x, y, z) \leqslant 0, z \in \psi(x, y),$$

这里 $\psi(x, y)$ 是下述优化问题的最优解集：

$$\min_{z} f_3(x, y, z),$$

$$\text{s.t.} \quad g_3(x, y, z) \leqslant 0.$$

上层规划需求解 min max 问题，使得该模型的理论分析和算法构造都极为困难. 因此即使是线性的情况，目前也鲜有这两方面的研究成果. 但该模型却有实际应用背景，以在云市场定价策略研究[93]中的应用为例. 云市场的三类服务商分别为：基础设施及服务（IaaS）、平台及服务（PaaS）、软件及服务（SaaS）. 这三类服务商之间分别为上层决策者、中间层决策者和下层决策者，且为递阶嵌套关系，各层决策者都希望最大化自己的利益，但各层决策者设定的价格都会影响到其他层决策者价格的设定. 若 PaaS 与 IaaS 是完全合作的，但是 PaaS 与 SaaS 为完全不合作，则该问题可以描述为中间层为悲观问题的悲观三层规划模型.

（2）下层为悲观问题的悲观三层规划模型：

$$\min_{x,y,z} f_1(x, y, z),$$

$$\text{s.t.}\quad g_1(x) \leqslant 0, (y,z) \in \psi(x),$$

这里 $\psi(x)$ 是下述优化问题的最优解集：

$$\min_y \max_z f_2(x,y,z),$$

$$\text{s.t.}\quad g_2(x,y,z) \leqslant 0, z \in \psi(x,y),$$

其中 $\psi(x,y)$ 是下述优化问题的最优解集（假设其有不止一个元素）：

$$\min_z f_3(x,y,z),$$

$$\text{s.t.}\quad g_3(x,y,z) \leqslant 0.$$

（3）中间层和下层均为悲观问题的悲观三层规划模型：

$$\min_x \max_{x,y} f_1(x,y,z),$$

$$\text{s.t.}\quad g_1(x) \leqslant 0, (y,z) \in \psi(x),$$

这里 $\psi(x)$ 是下述优化问题的最优解集（假设其有不止一个元素）：

$$\min_y \max_z f_2(x,y,z),$$

$$\text{s.t.}\quad g_2(x,y,z) \leqslant 0, z \in \psi(x,y),$$

其中 $\psi(x,y)$ 是下述优化问题的最优解集（假设其有不止一个元素）：

$$\min_z f_3(x,y,z),$$

$$\text{s.t.}\quad g_3(x,y,z) \leqslant 0.$$

三者之间两两博弈时常会出现完全非合作的情形，但是中间层需求解 min max 问题又使得该模型的求解难度进一步加大，目前还没有关于后面这两种模型的研究成果.

1.3 双层变分不等式问题

一些数学规划问题可以转化为变分不等式问题来求解. 为了更好地研究双层优化模型，很有必要研究双层变分不等式问题. 近十年来有许多关于双层变分不等式问题的研究成果，参见文献[6,7,20,74,76,79]. 这些模型有很多应用，但不足之处在于上下层之间没有很好的嵌套，而有时候实际问题抽象出来的双层变分不等式模型需要上下层之间有相互嵌套关系，故研究具有嵌套结构的双层变分不等式是很有必要的.

在文献[85]中，Wan 和 Chen 提出了具有嵌套结构的双层变分不等式（BVI），给出了当下层问题的解集为单点集时的最优性条件并构造了求解适定情形的算法. 下面我们先回忆文献[85]中的模型. 令 K 为 R^n 的子集，$H: R^m \times R^n \to R^n$, $P: R^m \to R^m$, $T: K$ Ã R^m 为集值映射. 考虑下面的双层变分不等式问题：求 $(x,z) \in K \times R^m$，使得

$$\langle H(z,x), x-y \rangle \leqslant 0, \qquad \forall y \in K, \tag{1.9}$$

其中 $z \in M(x)$. $M(x)$ 是下述参数变分不等式的最优解集. 求 $z \in T(x)$，使得

$$\langle P(z), z-v \rangle \leqslant 0, \qquad \forall v \in T(x). \tag{1.10}$$

不等式（1.9）和（1.10）分别称为上层变分不等式（UVI）和下层变分不等式（LVI）. x 为上层决策变量，z 为下层决策变量. 问题（BVI）的最优解集合为 Θ. 显然，（BVI）包括两个变分不等式. 下层变分不等式的约束域 $T(x)$ 受上层变量 x 所影响，上层变分不等式也包含下层的决策变量 z. 因此该模型上下两层有很好的嵌套关系，也正因如此使得该问题很难求解. 从广义的角度来看，该问题类似于拟变分不等式[39,57]. 若 $M(x)$ 为单点集，我们称（BVI）为适定的变分不等式；若 $M(x)$ 有不止一个元素，则称之为不适定的双层变分不等式. 对于这

类具有嵌套结构的双层变分不等式问题，无论是适定的情形，还是不适定的情形，其研究均尚处于初级阶段，如何构造一个对中等规模数据有效的方法，目前还是一极富挑战性的问题.