

文章编号: 1001-0920(2002) 03-0310-05

一种求解整数规划与 混合整数规划非线性罚函数方法

孟志青¹, 胡奇英¹, 杨晓琪²

(1. 西安电子科技大学 经济管理学院, 陕西 西安 710071; 2 香港理工大学 应用数学系, 香港)

摘 要: 证明了任何一个变量有界的整数规划问题 (IP) 和混合整数规划问题 (MIP) 都可以转化为一个等价的非整数 (或连续化) 规划问题 (NIP), 并给出一个用非线性精确罚函数法来求解该等价 NIP 的方法, 从而达到求解 IP 或 MIP 的目的, 数值实验表明了算法的可行性。该方法可广泛用于各应用领域里 IP 和 MIP 的求解, 特别是为非线性 IP 和 MIP 问题提供了一条通用的求解途径, 对解决许多实际优化问题具有重要意义。

关键词: 整数规划; 混合整数规划; 非整数规划; 非线性罚函数

中图分类号: O 122.2

文献标识码: A

A method of non-linear penalty function for solving integer programming and mixed integer programming

MENG Zhi-qing¹, HU Qi-ying¹, YANG Xiao-qi²

(1. School of Economics and Management, Xidian University, Xi'an 710071, China; 2 Department of Applied Mathematics, The Hong Kong Polytechnic University, Hong Kong, China)

Abstract: Integer programming (IP) problems and mixed integer programming (MIP) problems with bounded variables are transformed into equivalent non-integer (or continuous) programming (NIP) problems. IP and MIP are solved by using non-linear exact penalty functions optimization with the algorithm of gradient or genetic algorithm. This method can be used to solve many problems of IPs and MIPs in the application fields, including practical problems in the engineering and management. Numerical examples illustrate the feasibility of the algorithm.

Key words: integer programming; mixed integer programming; non-integer programming; non-linear penalty function

1 引 言

整数规划 (IP) 和混合整数规划 (MIP) 问题是当今国际上最优决策与应用领域里的一个极为重要的

分支^[1~10], 机械、化工、计算机、经济、生物、军事、社会等各领域里的许多优化问题均可归结为 IP 或 MIP 问题, 并且大多数的组合优化问题都可以写成一个 IP 或 MIP 问题^[1], 如背包问题、旅行商问题、

收稿日期: 2001-04-05; 修回日期: 2001-06-08

基金项目: 国家自然科学基金项目; 高等学校骨干教师资助项目

作者简介: 孟志青 (1962—), 男, 上海人, 副教授, 博士生, 从事最优决策与最优控制等研究; 胡奇英 (1965—), 男, 浙江诸暨人, 教授, 博士生导师, 从事随机决策与随机控制等研究。

最短路问题、选址问题、分配问题和生产与存储计划问题等^[1~3],因此如何求解IP和MIP问题是一个重要的研究领域。由于问题的可行解区域为离散点,所以一般不能用连续区域的求解算法,而只能用特殊方法求解,因此这方面的研究非常困难。随着计算技术的快速发展,人们已研究出许多快速求解非整数规划的算法,如变尺度法、内点法、罚函数法、信赖域法等^[1],这些算法具有良好的收敛性和收敛速度快等特点,可用于大规模问题的求解,但遗憾的是不能将这些方法用于求解IP和MIP问题。然而,如果我们能将IP和MIP转化成等价的非整数规划问题(NIP),便可以使用这些方法来求解,显然这是非常有意义的工作。如文献[4]对0-1整数规划的转化做了这方面的工作,文献[5]研究了线性0-1整数规划的转化并给出一个代数求解算法,但这方面研究国内外还很少,并未引起人们的重视。

求解非线性IP和MIP一直没有好的方法,尽管割平面法和分支定界法^[1]对小规模线性问题十分有效,但当问题规模较大时,这两种方法对于问题的求解几乎没有任何作用。近年来,人们研究了许多新的算法来求解IP或MIP,如随机搜索、模拟退火和遗传算法等^[3,6,7],在解决IP和MIP方面起到了很大作用,特别是与罚函数的结合来求解优化问题取得了成功^[4,6~8]。因此,寻找更好的方法求解IP和MIP是近年来人们一直在研究的重要课题。尽管文献[5]也给出了一个转化算法,但主要适宜0-1线性规划问题。本文证明了任何一个0-1规划问题都可以转化成等价的NIP,并且变量有界的IP和MIP都可以转化为一个等价的NIP,这种转化方式非常简单,仅仅是在原问题中对每一个整数变量增加一个约束方程;然后用非线性罚函数方法来求解这个等价的NIP。当然,也可以用现有的各种解析算法来求解等价的NIP。

2 整数规划或混合整数规划转化为非整数规划

设函数 $f: R^n \rightarrow R^1$, $\{ \}$, $g_i: R^n \rightarrow R^1 (i=1, \dots, m)$, $h_j: R^n \rightarrow R^1 (j=1, \dots, q)$, 其中 $I = \{1, 2, \dots, m\}$, $J = \{1, 2, \dots, q\}$ 。考虑下面的最优规划问题

IP 或 MIP

$$\begin{cases} \min f(x) \\ \text{s.t. } x \in X, g_i(x) \leq 0, h_j(x) = 0 \\ i = 1, 2, \dots, m, j = 1, 2, \dots, q \end{cases} \quad (1)$$

如果 X 是整数点集,则称IP为整数规划;如果 X 中部分分量是整数的,则称MIP为混合整数规划;如果 X 中所有元素均取0或1值,则称IP为0-1规划。设可行解集

$$X_0 = \{x \in R^n \mid x \in X, g_i(x) \leq 0, h_j(x) = 0, i = 1, 2, \dots, m, j = 1, 2, \dots, q\} \quad (2)$$

则有下面定理:

定理1 任何一个0-1规划问题(0-1IP)都可以转化为一个等价的非整数规划问题(NIP),即等价NIP的最优解也是0-1IP的最优解。

证明 在0-1IP中,设 $x = (x_1, x_2, \dots, x_n)^T \in X_0$,任何 $x_k = 0$ 或1($k = 1, 2, \dots, n$)都等价于下列方程有解: $x_k^2 - x_k = 0$ 。则有下面的非整数规划问题

$$\text{NIP} \begin{cases} \min f(x) \\ \text{s.t. } g_i(x) \leq 0, i = 1, 2, \dots, m \\ h_j(x) = 0, j = 1, 2, \dots, q \\ x_k^2 - x_k = 0, k = 1, 2, \dots, n \end{cases} \quad (3)$$

显然0-1IP与NIP具有相同的可行域,因此二者等价。

推论1 如果在 X 中取整数的变量有界,则整数规划问题或混合整数规划问题都可转化为一个等价的非整数规划问题。

证明 设 $x = (x_1, x_2, \dots, x_n)^T \in X_0$,若 x_1 要求是一个整数变量,因 x_1 有界,不失普遍性,不妨设 $0 \leq x_1 \leq L$,那么 x_1 可以表示成一个2进制数,即

$$x_1 = y_0 2^0 + y_1 2^1 + \dots + y_s 2^s \quad (4)$$

其中, $s = [\log_2 L] + 1$, $y_k = 0$ 或1($k = 1, 2, \dots, s$)。所以 x 的整数变量可以转化成仅取0或1的变量,于是IP便可以转化成等价的0-1规划问题,根据定理1知推论成立。

我们还可以给出另一种转化形式来证明推论1成立。设 $x = (x_1, x_2, \dots, x_n)^T \in X_0$,若 x_k ($k = 1, 2, \dots, n$)要求是一个有界整数变量,即要求 x_k 属于区间 $[a_k, b_k]$ 中的整数,不妨设 a_k, b_k 是整数,则有下面的非整数规划问题

$$\text{NIP} \begin{cases} \min f(x) \\ \text{s.t. } g_i(x) \leq 0, i = 1, 2, \dots, m \\ h_j(x) = 0, j = 1, 2, \dots, q \\ (x_k - a_k)(x_k - (a_k + 1)) \dots \\ (x_k - b_k) = 0, k = 1, 2, \dots, n \end{cases} \quad (5)$$

因为方程 $(x_k - a_k)(x_k - (a_k + 1)) \dots (x_k - b_k) = 0$ 的解 x_k 显然是属于区间 $[a_k, b_k]$ 中的整数,因此NIP也是IP的等价问题,故推论1成立。

定理1和推论1及其证明给出了IP和MIP转

化为非整数规划的两种方法,即连续化的方法。当转化后的目标函数和约束全部可微时,我们便可用现有的许多算法求解转化后的问题^[1]。下面研究一种将转化的等价问题进一步化为无约束优化问题的求解方法。

3 用非线性精确罚函数法来求解等价的NIP

近年来对非线性罚函数理论的研究取得了很大进展。非线性罚函数具有良好的特性,如光滑性和精确性^[9,10],为解决约束优化问题提供了新的途径。由上节可以将IP问题转化成等价的NIP问题,然后利用非线性精确罚函数求解,尽管约束的个数增加,但随着现代计算机性能的飞速提高,使许多大规模计算问题变得可行。这里我们总认为已将IP转化成下面的NIP

$$\text{NIP} \begin{cases} \min f(x) \\ \text{s.t. } x \in X, g_i(x) \leq 0, h_j(x) = 0 \\ i = 1, 2, \dots, m, j = 1, 2, \dots, q \end{cases} \quad (6)$$

设 $g_i^+(x) = \max\{0, g_i(x)\}$, $i = 1, \dots, m$ 。考虑下面NIP对应的非线性罚函数法

$$F(x, \rho) = \left[f(x)^2 + \rho \sum_{i=1}^m g_i^+(x)^2 + \rho \sum_{j=1}^q h_j(x)^2 \right]^{\frac{1}{2}} \quad (7)$$

其中 $\rho \geq 0$ 。如果 $g_i(\forall i = 1, \dots, m)$ 和 $h_j(\forall j = 1, \dots, q)$ 是连续可微的,显然 $F(x, \rho)$ 也是连续可微的。考虑对应的罚函数优化问题

$$P(\rho) \begin{cases} \min F(x, \rho) \\ \text{s.t. } x \in X \end{cases} \quad (8)$$

文献[9,10]已经证明罚函数 $F(x, \rho)$ 在一定条件下是一个精确罚函数,即存在一个罚参数 ρ 使得罚问题 $P(\rho)$ 的最优解也是原问题(NIP)的最优解。下面给出求解整数规划或混合整数规划的一个算法:

Step0: 首先将整数规划(IP)或混合整数规划(MIP)转化成非整数规划(NIP),给出相应的非线性罚函数优化问题;

Step1: 给定开始点 x^s 及 $N > 1$, 置 $k = 0$;

Step2: 使用一个开始点 x^s , 利用迭代算法求解罚函数优化问题: $\lim_{x \rightarrow X} F(x, \rho_k)$, 设 x^k 是对应的最优解;

Step3: 如果 x^k 是可行解则停止,得到NIP的一

个解 x^k ; 否则, 设 $\rho_{k+1} = N \rho_k$, 置 $k := k + 1$, 重复 Step2。

注1 在实际计算中为了得到最优解,一般要求目标函数 f 在 X 上保持非负。如果存在 $f(x) < 0$, 则用函数 $e^{f(x)}$ 来替代 $f(x)$ 作为新的目标函数; 如果 $f(x)$ 有界, 则找一个充分大的整数 $M > f(x)$ ($x \in R^n$), 使用函数 $f(x) + M$ 来替代 $f(x)$ 作为新的目标函数, 这样并不影响原问题的最优解。Step2 的计算可以采用梯度搜索或遗传算法计算。

定理2 假设函数 $f, g_i(\forall i = 1, \dots, m)$ 和 $h_j(\forall j = 1, \dots, q)$ 均是连续函数, $\lim_{x \rightarrow X} f(x) = +\infty$ 。设由算法得到的序列 x^k 对应的罚函数值 $F(x^k, \rho_k)$ 有界, 如果存在 x^k 是NIP的可行解, 那么 x^k 也是NIP的最优解; 否则, 序列 x^k 中没有一点是可行解, 那么序列 x^k 的任意聚点 x^* 是有界的并且是NIP的最优解。

证明 首先, 如果 x^k 是NIP的可行解, 我们证明 x^k 也是NIP的最优解。显然 $F(x^k, \rho_k) = f(x^k)$, 由算法知 x^k 是问题 $\lim_{x \rightarrow X} F(x, \rho_k)$ 的最优解, 即

$$f(x^k) = F(x, \rho_k), \quad \forall x \in X \quad (9)$$

因此有 $f(x^k) = F(x, \rho_k) = f(x), \quad \forall x \in X_\infty$

现在证明第2个结论。根据假设, 存在某个数 $L > 0$ 使得

$$L > F(x^k, \rho_k), \quad k = 0, 1, \dots \quad (10)$$

假设 x^k 无界, 设 $x^k \rightarrow +\infty$ ($k \rightarrow \infty$), 因此由式(10)得 $L > f(x^k) = F(x^k, \rho_k)$, 这便导致一个矛盾, 因为 $\lim_{x \rightarrow X} f(x) = +\infty$ 。

最后证明 x^k 的任何极限点均属于 X_∞ 。不失一般性, 不妨设 $x^k \rightarrow x^*$, 假设 $x^* \notin X_\infty$, 那么存在某个 i 或 j 使得 $g_i(x^*) > 0$ 或 $h_j(x^*) \neq 0$ 。因为 $g_i(i = 1, \dots, m)$ 和 $h_j(j = 1, \dots, q)$ 是连续函数, 所以 $F(x, \rho_k)$ ($k = 1, 2, \dots$) 也是连续函数。由假设我们有

$$L > F(x^k, \rho_k) = \left[f(x^k)^2 + \rho_k g_i^+(x^k)^2 + \rho_k h_j(x^k)^2 \right]^{\frac{1}{2}} \quad (11)$$

设 $k \rightarrow \infty$, 则 $\rho_k \rightarrow \infty$ 得到 $L > F(x^k, \rho_k) \rightarrow +\infty$, 导致一个矛盾, 因此 $x^* \in X_\infty$ 。于是又有

$$F(x^k, \rho_k) = F(x, \rho_k) = f(x), \quad \forall x \in X_\infty \quad (12)$$

设 $k \rightarrow \infty$, 则得 $f(x^*) = f(x), \quad \forall x \in X_\infty$

表 1 N IPe 和 N IPe 的最优解计算结果

k	ρ_k	N IPe		N IPe	
		$x^k = (x_1, x_2, x_3)$	$f(x^k)$	$x^k = (x_1, x_2, x_3)$	$f(x^k)$
0	10^3	(0 263 6, 0 164 8, 2 127 2)	7. 417 5	(2 000 0, 0 002 4, 0 004 8)	20 952 4
1	10^4	(1 000, 1 002 0, 3 002 0)	1. 992 1	(2 000 0, 2 000 0, 2 000 5)	4 998 0
2	10^5	(1 000, 1 000 2, 3 000 2)	1. 999 2	(2 000 0, 2 000 0, 3 000 0)	2 000 0
3	10^6	(1 000, 1 000 0, 3 000 0)	2 000 0		

4 计算实例

例 1 求下面整数规划问题

$$\text{I Pe} \begin{cases} \min f(x) = (x_1 - 1)^2 + (x_2 - 2)^2 + (x_3 - 4)^2 \\ \text{s t } -x_1 + x_2 + x_3 \leq 3.5 \\ x_1 + x_2 - x_3 \leq 6 \\ x_1, x_2, x_3 \text{ 为 } 0 \text{ 的整数} \end{cases}$$

不难求得各变量的界分别为 $0 \leq x_1 \leq 6, 0 \leq x_2 \leq 3.5, 0 \leq x_3 \leq 3.5$, 因此令 $x_1 = y_1 + 2y_2 + 4y_3$, $x_2 = y_4 + 2y_5, x_3 = y_6 + 2y_7$.

由推论 1 知, 上面的整数规划可化为如下两个等价的非整数规划

$$\text{N IPe} \begin{cases} \min f(x) = (y_1 + 2y_2 + 4y_3 - 1)^2 + (y_4 + 2y_5 - 2)^2 + (y_6 + 2y_7 - 4)^2 \\ \text{s t } -y_1 - 2y_2 - 4y_3 + y_4 + 2y_5 + y_6 + 2y_7 \leq 3.5 \\ y_1 + 2y_2 + 4y_3 + y_4 + 2y_5 - y_6 - 2y_7 \leq 6 \\ y_i^2 - y_i = 0, \quad i = 1, 2, \dots, 7 \end{cases}$$
$$\text{N IPe} \begin{cases} \min f(x) = (x_1 - 1)^2 + (x_2 - 2)^2 + (x_3 - 4)^2 \\ \text{s t } -x_1 + x_2 + x_3 \leq 3.5 \\ x_1 + x_2 - x_3 \leq 6 \\ x_1(x_1 - 1)(x_1 - 2) \dots (x_1 - 6) = 0 \\ x_2(x_2 - 1)(x_2 - 2)(x_2 - 3) = 0 \\ x_3(x_3 - 1)(x_3 - 2)(x_3 - 3) = 0 \end{cases}$$

用上节给出的算法分别求解 N IPe 和 N IPe, 取罚参数值 $\rho_0 = 1\,000, N = 10$, N IPe 的开始点是 $(y_1, y_2, \dots, y_7) = (8, 8, \dots, 8)$, N IPe 的开始点是 $(x_1, x_2, x_3) = (8, 8, 8)$. 在算法的 Step2 中, 用改进下降法求得对应的罚优化问题的最优解(见表 1).

由表 1 容易得出原问题(I Pe) 的最优目标值是 $f(x^*) = 2$, 最优解 $x^* = (1, 1, 3)$ 或 $(2, 2, 3)$. 可以

看出, 两种转化得到了原问题的两个不同最优解.

例 2 为验证算法的可行性和效率, 我们使用 VC++ 6.0 在 Pentium III (CPU) (128M 内存) 机上对下面 0-1 整数规划模型进行了数值实验.

$$\begin{cases} \max f(x) = x_1 + x_2 + \dots + x_n \\ \text{s t } a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \leq 10 \\ x_1, x_2, \dots, x_n = 0 \text{ 或 } 1 \end{cases}$$

其中 $a_{ij} (i = 1, 2, \dots, m; j = 1, 2, \dots, n)$ 是 $0 \sim 2$ 之间的随机数. 该模型的特征是线性约束系数的主要是在区间 $[0, 2]$ 之间, 多数在 $[0.5, 1.5]$ 中, 因此其最优目标值一般是 7, 8 或 9, 这样构造模型便于我们判断最优解是否能得到. 这里 Step2 采用的是一种改进下降算法, 每步迭代采用不同的起始点计算, 以避免收敛到局部最优. 表 2 给出了 6 个模型的计算结果.

表 2 各模型计算结果

模型大小	步数 k	目标值	小步率 1 (步 /s)	小步率 2 (步 /s)
30 × 30	2	9 000 000	92	92
60 × 60	17	8 000 000	67	67
90 × 90	9	8 000 001	41	41
100 × 100	4	7 000 000	39	39
200 × 200	12	8 000 000	7. 2	7. 5
300 × 300	2	7 000 001	4. 0	4. 3

说明 1) 表中目标值是对应步数 k 得到的, 由于模型规模和系数值不同, 所以迭代次数也不同.

2) 表中小步率 1 是 Step2 算法——改进下降法中的每一步迭代的速度, 即每秒计算一步目标值的次数, 表明了程序执行速度. 计算结果表明, 随着问题规模的增大, 执行速度大约按相应的倍数减慢, 如 100×100 与 300×300 规模相差 9 倍, 但前者执行速度仅快 9 倍多.

3) 表中小步率 2 是问题对应非整数规划(去掉 0-1 约束) 的每秒计算一步目标值的次数. 从表中结果看, 与小步率 1 相比, 将 0 或 1 转化所增加的约束个数, 对使用罚函数法计算的执行速度在规模 100

$\times 100$ 以下几乎没有影响,随着问题规模的增大影响将逐渐增加。

4) 对于大规模的该模型用普通的分支定界法或割平面法求解,显然是很麻烦的。

5 结 语

本文方法对于目标和约束具有很好的光滑性时,采用经典的算法更好些,可以较好地保证收敛性和速度。算法的复杂性问题与 Step 2 中采用的算法有关,数值例子表明,转化所增加的约束对中等规模问题程序的执行速度影响并不明显。另外我们容易将上述算法与遗传算法结合起来计算,这时的主要问题是收敛速度会变得很慢,有关这方面的内容将另文给出。

本文方法的优点是开辟了一条通用的求解 IP 的途径,可以结合不同的求解无约束问题方法进行计算。根据推论 1,任何变量有界的 IP 或 MIP 转化成等价的 NIP 或 NIP 的方法,对于 0-1 规划 NIP 和 NIP 是一致的,同样对于非 0-1 整数规划问题, NIP 和 NIP 对于求解也是等价的,但究竟哪种形式更有利于求解还有待于进一步研究。在实际中,许多问题的变量都是有界的,但有时变量有界明确给出,因此,我们可以通过约束来求出或估计变量的上下界,一般很容易求出线性整数规划的变量的界。对于无界变量的 IP 和 MIP 问题还有待于进一步研究。

参考文献(References):

- [1] 徐光辉,刘彦佩,程侃. 运筹学基础手册[M]. 北京: 科学出版社, 1999.
- [2] Eiselt H A, Sandblom C L. Integer programming and network models[M]. Berlin: Springer Press, 2000.
- [3] 邢文训,谢金星. 现代优化计算方法[M]. 北京: 清华大学出版社, 1999.
- [4] 高峰,张连生. 多项式 0-1 整数规划的两个连续化途径[J]. 上海大学学报(J of Shanghai Univ), 1999, 5(2): 95-98.
- [5] Dimitris Bertsimas, Georgios Perakis, Sriheer Tayeer. A new algebraic geometry algorithm for integer programming[J]. Management Science, 2000, 46(7): 999-1008.
- [6] 吴志远,邵惠鹤,吴新余. 基于遗传算法的退火精确罚函数非线性约束优化方法[J]. 控制与决策(Control and Decision), 1998, 13(2): 136-140.
- [7] Yang X Q, Mees A I, Campbell K. Simulated annealing and penalty methods for binary multicommodity flow problems[J]. Progress in Optimization, 2000, 6(1): 93-105.
- [8] 赵蔚. 两层多目标规划罚函数法[J]. 自动化学报(Acta Automatica Sinica), 1998, 24(3): 331-337.
- [9] Rubinov A M, Glover B M, Yang X Q. Extended Lagrange and penalty functions in continuous optimization[J]. Optimization, 1999, 46(3): 327-351.
- [10] Rubinov A M, Glover B M, Yang X Q. Decreasing functions with applications to penalization[J]. SIAM J Optim, 1999, 10(1): 289-313.

(上接第 309 页)

参考文献(References):

- [1] Yan Shi, Masaharu M Iamamoto. A new approach of neuro-fuzzy learning algorithm for tuning fuzzy rules[J]. Fuzzy Sets and Systems, 2000, 112: 99-116.
- [2] Mauricio Figueiredo, Fernando Gomide. Design of fuzzy systems using neurofuzzy networks[J]. IEEE Trans on Neural Networks, 1999, 10(4): 815-827.
- [3] Wang L X, Mendel J M. Generating fuzzy rules by learning from examples[J]. IEEE Trans on Systems, Man and Cybernet, 1992, 22(6): 1414-1422.
- [4] Milligan G W, Cooper M C. An examination of procedure for detecting the number of clusters in a data set[J]. Psychometrika, 1985, (50): 159-179.
- [5] Michio Sugeno, Takahiro Yasukawa. A fuzzy-logic-based approach to qualitative modeling[J]. IEEE Trans on Fuzzy Systems, 1993, 1(1): 7-31.
- [6] D E Gustafson, W C Kessel. Fuzzy clustering with a fuzzy covariance matrix[A]. Proc IEEE Int Conf on Fuzzy Systems[C]. San Diego, 1979. 761-766.
- [7] A F Gomez-Skarmeta, M Delgado, M A Vila. About the use of fuzzy clustering techniques for fuzzy model identification[J]. Fuzzy Sets and Systems, 1999, (106): 179-188.
- [8] J C Bezdek. Pattern recognition with fuzzy objective function algorithms[M]. New York: Plenum, 1981.
- [9] D A Linkens, Min-You Chen. Input selection and partition validation for fuzzy modeling using neural network[J]. Fuzzy Sets and Systems, 1999, (107): 299-308.
- [10] Lin C T, Lee C S G. Neural-network-based fuzzy logic control and decision system[J]. IEEE Trans on Computer, 1991, 40(12): 1320-1336.
- [11] Yin Wang, Gang Rong. A self-organizing neural network based fuzzy system[J]. Fuzzy Set and Systems, 1999, (103): 1-11.