第48卷第2期
2008年3月
大连理工大学学报
Journal of Dalian University of Technology
Vol. 48, No. 2
Mar. 2008

文章编号：1000-8608(2008)02-0308-05

二维装箱问题非线性规划模型和算法

于洪霞[1,2]， 张绍武[3]， 张立卫[*1]

（1.大连理工大学 应用数学系，辽宁 大连 116024；
2.上海电力学院 数理系，上海 200090；
3.大连理工大学 电子与信息工程学院，辽宁 大连 116024）

摘要：二维装箱问题是具有广泛应用背景的一类组合优化问题，这类问题是NP难问题，很难得到精确解．将二维装箱问题表示为一个非线性规划模型，用变分分析中切锥的概念建立了这一优化问题的一阶最优性条件．给出了求解这一优化问题的增广Lagrange方法，并求解了具体问题．数值实验表明增广Lagrange方法适合求解该问题，对于不超过10个物品的装箱问题可以求得精确解．

关键词：二维装箱问题；一阶最优性条件；增广Lagrange方法
中图分类号：O221.2 **文献标志码**：A

0 引 言

二维装箱问题可以分为两类：一类是给定数量不限大小固定的容器，目标是使用最少数量的容器装完所有的物品；另一类是给定一个宽度为有限数值的矩形容器，它的高度不限，要求将所有物品放入该容器中，并取得最小的放置高度，使容器利用率最高．本文只考虑第二类装箱问题．

在计算机科学和工业领域中，装箱问题有着广泛的应用背景，包括多处理器任务调度、资源分配、运输计划等，因此对其求解的研究具有广泛的应用价值．从计算复杂性理论来讲，装箱问题是NP难问题，很难精确求解，已有的大部分成果都是针对近似算法的研究．目前的近似算法有下次适应(next-fit decreasing height，NFDH)、首次适应(first-fit decreasing height，FFDH)、最佳适应(best-fit decreasing height，BFDH)等[1]．本文建立二维装箱问题的数学模型，将其转化为一个非线性规划问题，给出数值算法并对不超过10个物品的装箱问题进行求解．

1 数学模型

给定宽度为L，高度不限的矩形容器；矩形物品集合$J=\{1,\cdots,n\}$，每个矩形物品的宽度为l_i，高度为h_i．以矩形容器左下角的位置为坐标原点，每个小矩形物品左下角坐标用(x_i,y_i)表示，则每个小矩形的内部可以表示为

$$\begin{aligned}S_i(x_i,y_i)=\{(\boldsymbol{u},\boldsymbol{w})\in\mathbf{R}^{2n}\mid\ & x_i<u_i<x_i+l_i,\\ & y_i<w_i<y_i+h_i\}\end{aligned}$$

令$H=\{(i,j)\mid i\in J,j\in J,i<j\}$，则$|H|=n(n-1)/2$，任何两个矩形物品不重合可以表示为

$$S_i(x_i,y_i)\cap S_j(x_j,y_j)=\varnothing,\quad (i,j)\in H$$

从而二维装箱问题可以用下面的数学模型来表示：

广义模型

$$\begin{aligned}&\min\quad v\\ &\text{s.t.}\quad S_i(x_i,y_i)\cap S_j(x_j,y_j)=\varnothing,\quad (i,j)\in H,\\ &\qquad 0\leqslant x_i\leqslant L-l_i,\quad 0\leqslant y_i\leqslant v-h_i,\quad i\in J\end{aligned}$$

由于约束$S_i(x_i,y_i)\cap S_j(x_j,y_j)=\varnothing$，$(i,j)\in H$不是通常的由函数表示的约束，这个模型是一个非光滑数学规划问题，下面研究如何将其转化为一个光滑的数学规划问题．

令$P^X(S)$和$P^Y(S)$分别表示矩形S在x轴和y轴上的投影，则

$$P^X(S_i(x_i,y_i))=(x_i,x_i+l_i),$$
$$P^Y(S_i(x_i,y_i))=(y_i,y_i+h_i)$$

容易知道两个矩形物品重合当且仅当它们到x轴和y轴的投影都相交，经过简单的计算能得到

收稿日期：2005-12-20； **修回日期**：2007-12-17．
基金项目：国家自然科学基金资助项目(10771026)．
作者简介：于洪霞(1978-)，女，博士；张立卫*(1966-)，男，教授，博士生导师．

$$P^X(S_i(x_i,y_i)) \cap P^X(S_j(x_j,y_j)) \neq \varnothing \Leftrightarrow$$
$$\left|x_i - x_j + \frac{l_i - l_j}{2}\right| - \frac{l_i + l_j}{2} < 0,$$
$$P^Y(S_i(x_i,y_i)) \cap P^Y(S_j(x_j,y_j)) \neq \varnothing \Leftrightarrow$$
$$\left|y_i - y_j + \frac{h_i - h_j}{2}\right| - \frac{h_i + h_j}{2} < 0$$

从而，广义模型可以等价地表示为

绝对值模型

$$\min \quad v$$
$$\text{s.t.} \quad \max\left\{\left|x_i - x_j + \frac{l_i - l_j}{2}\right| - \frac{l_i + l_j}{2},\ \left|y_i - y_j + \frac{h_i - h_j}{2}\right| - \frac{h_i + h_j}{2}\right\} \geqslant 0,\quad (i,j) \in H,$$
$$0 \leqslant x_i \leqslant L - l_i,\ 0 \leqslant y_i \leqslant v - h_i,\ i \in J$$

由于其中包含反凸约束，可行集是非凸的，进而知道绝对值模型是非凸非光滑的优化问题. 通过引入人工变量 $\boldsymbol{w}^X$、$\boldsymbol{w}^Y$、$\boldsymbol{w}^A$，可以把绝对值模型转化为一个光滑的优化问题. 令 $\boldsymbol{z} = (\boldsymbol{x}\quad \boldsymbol{y}\quad \boldsymbol{w}^X\quad \boldsymbol{w}^Y\quad \boldsymbol{w}^A\quad v)$，$M = n(n-1)/2$，定义函数 $\boldsymbol{f}_0: \mathbf{R}^n \times \mathbf{R}^n \times \mathbf{R}^M \times \mathbf{R}^M \times \mathbf{R}^M \longrightarrow \mathbf{R}$ 和 $\boldsymbol{F}: \mathbf{R}^n \times \mathbf{R}^n \times \mathbf{R}^M \times \mathbf{R}^M \times \mathbf{R}^M \times \mathbf{R} \longrightarrow \mathbf{R}^M \times \mathbf{R}^M \times \mathbf{R}^M \times \mathbf{R}^n$ 如下：

$$\boldsymbol{f}_0(\boldsymbol{z}) = v,$$
$$\boldsymbol{F}(\boldsymbol{z}) = (\boldsymbol{F}_1^T(\boldsymbol{z})\quad \boldsymbol{F}_2^T(\boldsymbol{z})\quad \boldsymbol{F}_3^T(\boldsymbol{z})\quad \boldsymbol{F}_4^T(\boldsymbol{z}))^T$$

其中

$$\boldsymbol{F}_1(\boldsymbol{z}) = \left(\left(x_1 - x_2 + \frac{l_1 - l_2}{2}\right)^2 - w_{1,2}^{X2}\quad \cdots\quad \left(x_{n-1} - x_n + \frac{l_{n-1} - l_n}{2}\right)^2 - w_{n-1,n}^{X2}\right)^T,$$

$$\boldsymbol{F}_2(\boldsymbol{z}) = \left(\left(y_1 - y_2 + \frac{h_1 - h_2}{2}\right)^2 - w_{1,2}^{Y2}\quad \cdots\quad \left(y_{n-1} - y_n + \frac{h_{n-1} - h_n}{2}\right)^2 - w_{n-1,n}^{Y2}\right)^T,$$

$$\boldsymbol{F}_3(\boldsymbol{z}) = \left(\left(w_{1,2}^A - w_{1,2}^X + \frac{l_1 + l_2}{2}\right)\times\left(w_{1,2}^A - w_{1,2}^Y + \frac{h_1 + h_2}{2}\right)\quad \cdots\quad \left(w_{n-1,n}^A - w_{n-1,n}^X + \frac{l_{n-1} + l_n}{2}\right)\times\left(w_{n-1,n}^A - w_{n-1,n}^Y + \frac{h_{n-1} + h_n}{2}\right)\right)^T,$$

$$\boldsymbol{F}_4(\boldsymbol{z}) = (v - y_1 - h_1\quad \cdots\quad v - y_n - h_n)^T$$

令 $\boldsymbol{\Omega} = \{\boldsymbol{z} \in \boldsymbol{Z} \mid \boldsymbol{F}(\boldsymbol{z}) \in \boldsymbol{D}\}$，其中 $\boldsymbol{D} = \{\boldsymbol{0}_M\} \times \{\boldsymbol{0}_M\} \times \{\boldsymbol{0}_M\} \times \mathbf{R}_+^n$，

$$\boldsymbol{Z} = \left\{\boldsymbol{z} \in \mathbf{R}^{2n+3M+1} \,\middle|\, \begin{array}{l} \boldsymbol{0}_n \leqslant \boldsymbol{x} \leqslant L\boldsymbol{e}_{[n]} - \boldsymbol{l} \\ \boldsymbol{y} \geqslant \boldsymbol{0}_n \\ \boldsymbol{w}^A \geqslant \boldsymbol{0}_M \\ \boldsymbol{w}^X \in \mathbf{R}^M, \boldsymbol{w}^Y \in \mathbf{R}^M, v \in \mathbf{R} \end{array}\right\}$$

$\boldsymbol{e}_{[n]} = (1\quad \cdots\quad 1)^T \in \mathbf{R}^n$，这样绝对值模型就可以等价地表示为一个光滑优化问题.

非线性规划模型

$$\min \quad \boldsymbol{f}_0(\boldsymbol{z})$$
$$\text{s.t.} \quad \boldsymbol{F}(\boldsymbol{z}) \in \boldsymbol{D}$$
$$\boldsymbol{z} \in \boldsymbol{Z}$$

为了得到非线性规划模型的一阶最优性条件，引入下列标记：

$$\boldsymbol{W}^X = \mathbf{diag}_{(i,j)\in H}(w_{i,j}^X)$$
$$\boldsymbol{W}^Y = \mathbf{diag}_{(i,j)\in H}(w_{i,j}^Y)$$
$$\boldsymbol{A} = \mathbf{diag}_{(i,j)\in H}(\alpha_{i,j})$$
$$\boldsymbol{B} = \mathbf{diag}_{(i,j)\in H}(\beta_{i,j})$$
$$\boldsymbol{C} = \mathbf{diag}_{(i,j)\in H}(\gamma_{i,j})$$

$$\boldsymbol{D}_1 = \begin{pmatrix} d_{1,2}(\boldsymbol{x}) & d_{1,3}(\boldsymbol{x}) & \cdots & d_{1,n}(\boldsymbol{x}) & 0 & \cdots & 0 & \cdots & 0 \\ -d_{1,2}(\boldsymbol{x}) & 0 & \cdots & 0 & d_{2,3}(\boldsymbol{x}) & \cdots & d_{2,n}(\boldsymbol{x}) & \cdots & 0 \\ 0 & -d_{1,3}(\boldsymbol{x}) & \cdots & 0 & -d_{2,3}(\boldsymbol{x}) & \cdots & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & \cdots & d_{n-1,n}(\boldsymbol{x}) \\ 0 & 0 & \cdots & -d_{1,n}(\boldsymbol{x}) & 0 & \cdots & -d_{2,n}(\boldsymbol{x}) & \cdots & -d_{n-1,n}(\boldsymbol{x}) \end{pmatrix}$$

$$\boldsymbol{D}_2 = \begin{pmatrix} c_{1,2}(\boldsymbol{x}) & c_{1,3}(\boldsymbol{x}) & \cdots & c_{1,n}(\boldsymbol{x}) & 0 & \cdots & 0 & \cdots & 0 \\ -c_{1,2}(\boldsymbol{x}) & 0 & \cdots & 0 & c_{2,3}(\boldsymbol{x}) & \cdots & c_{2,n}(\boldsymbol{x}) & \cdots & 0 \\ 0 & -c_{1,3}(\boldsymbol{x}) & \cdots & 0 & -c_{2,3}(\boldsymbol{x}) & \cdots & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & \cdots & c_{n-1,n}(\boldsymbol{x}) \\ 0 & 0 & \cdots & -c_{1,n}(\boldsymbol{x}) & 0 & \cdots & -c_{2,n}(\boldsymbol{x}) & \cdots & -c_{n-1,n}(\boldsymbol{x}) \end{pmatrix}$$

其中 $d_{i,j}(\boldsymbol{x})=2(x_i-x_j)+l_i-l_j$，$c_{i,j}(\boldsymbol{y})=2(y_i-y_j)+h_i-h_j$，$\alpha_{i,j}=-(w^{\Lambda}_{i,j}-w^{Y}_{i,j}+(h_i+h_j)/2)$，$\beta_{i,j}=-(w^{\Lambda}_{i,j}-w^{X}_{i,j}+(l_i+l_j)/2)$，$\gamma_{i,j}=2w^{\Lambda}_{i,j}-w^{X}_{i,j}+(l_i+l_j)/2-w^{Y}_{i,j}+(h_i+h_j)/2$. 这样 $\nabla\boldsymbol{F}_i^{\mathrm{T}}(\boldsymbol{z})(i=1,\cdots,4)$ 可以表示为

$$\nabla\boldsymbol{F}_1^{\mathrm{T}}(\boldsymbol{z})=\begin{pmatrix}\boldsymbol{D}_1\\ \boldsymbol{0}_{n\times M}\\ -2\boldsymbol{W}^{X}\\ \boldsymbol{0}_{M\times M}\\ \boldsymbol{0}_{M\times M}\\ \boldsymbol{0}_{1\times M}\end{pmatrix},\ \nabla\boldsymbol{F}_2^{\mathrm{T}}(\boldsymbol{z})=\begin{pmatrix}\boldsymbol{0}_{n\times M}\\ \boldsymbol{D}_2\\ \boldsymbol{0}_{M\times M}\\ -2\boldsymbol{W}^{Y}\\ \boldsymbol{0}_{M\times M}\\ \boldsymbol{0}_{1\times M}\end{pmatrix},$$

$$\nabla\boldsymbol{F}_3^{\mathrm{T}}(\boldsymbol{z})=\begin{pmatrix}\boldsymbol{0}_{n\times M}\\ \boldsymbol{0}_{n\times M}\\ \boldsymbol{A}\\ \boldsymbol{B}\\ \boldsymbol{C}\\ \boldsymbol{0}_{1\times M}\end{pmatrix},\ \nabla\boldsymbol{F}_4^{\mathrm{T}}(\boldsymbol{z})=\begin{pmatrix}\boldsymbol{0}_{n\times n}\\ -\boldsymbol{I}_{n\times n}\\ \boldsymbol{0}_{M\times n}\\ \boldsymbol{0}_{M\times n}\\ \boldsymbol{0}_{M\times n}\\ \boldsymbol{1}_{1\times n}\end{pmatrix}$$

函数 $\boldsymbol{F}$ 的 Jacobi 转置矩阵为

$$\nabla\boldsymbol{F}^{\mathrm{T}}(\boldsymbol{z})=(\nabla\boldsymbol{F}_1^{\mathrm{T}}(\boldsymbol{z})\quad\nabla\boldsymbol{F}_2^{\mathrm{T}}(\boldsymbol{z})\quad\nabla\boldsymbol{F}_3^{\mathrm{T}}(\boldsymbol{z})\quad\nabla\boldsymbol{F}_4^{\mathrm{T}}(\boldsymbol{z}))$$

为求得 $\boldsymbol{N}_{\boldsymbol{\Omega}}(\bar{z})$，$\bar{z}\in\boldsymbol{\Omega}$，给出 $\boldsymbol{N}_{\boldsymbol{Z}}(\bar{z})$ 和 $\boldsymbol{N}_{\boldsymbol{D}}(\boldsymbol{F}(\bar{z}))$ 的计算公式.

令

$$\boldsymbol{Z}_x=\{\boldsymbol{x}\in\mathbf{R}^n\mid\boldsymbol{0}_n\leqslant\boldsymbol{x}\leqslant L\boldsymbol{e}[n]-\boldsymbol{l}\},$$
$$\boldsymbol{Z}_y=\{\boldsymbol{y}\in\mathbf{R}^n\mid\boldsymbol{y}\geqslant\boldsymbol{0}_n\}$$

其中 $Z_x^i=[0,L-l_i]$，$Z_y^i=[0,\infty)$. 则

$$\boldsymbol{N}_{\boldsymbol{Z}}(\bar{z})=\boldsymbol{N}_{\boldsymbol{Z}_x}(\bar{\boldsymbol{x}})\times\boldsymbol{N}_{\boldsymbol{Z}_y}(\bar{\boldsymbol{y}})\times\boldsymbol{N}_{\mathbf{R}^M}(\bar{\boldsymbol{w}}^{X})\times\boldsymbol{N}_{\mathbf{R}^M}(\bar{\boldsymbol{w}}^{Y})\times\boldsymbol{N}_{\mathbf{R}_+^M}(\bar{\boldsymbol{w}}^{\Lambda})\times N_{\mathbf{R}}(\bar{v})$$

其中

$$\boldsymbol{N}_{\boldsymbol{Z}_x}(\bar{\boldsymbol{x}})=N_{Z_x^1}(\bar{x}_1)\times\cdots\times N_{Z_x^n}(\bar{x}_n)$$

$$N_{Z_x^i}(\bar{x}_i)=\begin{cases}\mathbf{R}_-; & \bar{x}_i=0\\ \{\boldsymbol{0}\}; & 0<\bar{x}_i<L-l_i\\ \mathbf{R}_+; & \bar{x}_i=L-l_i\end{cases}$$

$$\boldsymbol{N}_{\boldsymbol{Z}_y}(\bar{\boldsymbol{y}})=N_{Z_y^1}(\bar{y}_1)\times\cdots\times N_{Z_y^n}(\bar{y}_n)$$

$$N_{Z_y^i}(\bar{y}_i)=\begin{cases}\mathbf{R}_-; & \bar{y}_i=0\\ \{\boldsymbol{0}\}; & \bar{y}_i>0\end{cases}$$

$$\boldsymbol{N}_{\mathbf{R}^M}(\bar{\boldsymbol{w}}^{X})=\boldsymbol{N}_{\mathbf{R}^M}(\bar{\boldsymbol{w}}^{Y})=\{\boldsymbol{0}_M\}$$

$$\boldsymbol{N}_{\mathbf{R}_+^M}(\bar{\boldsymbol{w}}^{\Lambda})=N_{\mathbf{R}_+}(\bar{w}^{\Lambda}_{1,2})\times\cdots\times N_{\mathbf{R}_+}(\bar{w}^{\Lambda}_{n-1,n})$$

$$N_{\mathbf{R}_+}(\bar{w}^{\Lambda}_{i,j})=\begin{cases}\mathbf{R}_-; & \bar{w}^{\Lambda}_{i,j}=0\\ \{\boldsymbol{0}\}; & \bar{w}^{\Lambda}_{i,j}>0\end{cases}$$

$$N_{\mathbf{R}}(\bar{v})=\{\boldsymbol{0}\}$$

$\boldsymbol{D}$ 在 $\boldsymbol{F}(\bar{z})$ 处的法锥 $\boldsymbol{N}_{\boldsymbol{D}}(\boldsymbol{F}(\bar{z}))$ 为

$$\boldsymbol{N}_{\boldsymbol{D}}(\boldsymbol{F}(\bar{z}))=\boldsymbol{N}_{\{\boldsymbol{0}_M\}}(\boldsymbol{F}_1(\bar{z}))\times\boldsymbol{N}_{\{\boldsymbol{0}_M\}}(\boldsymbol{F}_2(\bar{z}))\times\boldsymbol{N}_{\{\boldsymbol{0}_M\}}(\boldsymbol{F}_3(\bar{z}))\times\boldsymbol{N}_{\mathbf{R}_+^n}(\boldsymbol{F}_4(\bar{z}))$$

其中

$$\boldsymbol{N}_{\{\boldsymbol{0}_M\}}(\boldsymbol{F}_1(\bar{z}))=\boldsymbol{N}_{\{\boldsymbol{0}_M\}}(\boldsymbol{F}_2(\bar{z}))=\boldsymbol{N}_{\{\boldsymbol{0}_M\}}(\boldsymbol{F}_3(\bar{z}))=\mathbf{R}^M$$

$$\boldsymbol{N}_{\mathbf{R}_+^n}(\boldsymbol{F}_4(\bar{z}))=\boldsymbol{N}_{\mathbf{R}_+}(\boldsymbol{F}_4^1(\bar{z}))\times\cdots\times\boldsymbol{N}_{\mathbf{R}_+}(\boldsymbol{F}_4^n(\bar{z}))$$

$$\boldsymbol{N}_{\mathbf{R}_+}(\boldsymbol{F}_4^i(\bar{z}))=\begin{cases}\mathbf{R}_-; & \boldsymbol{F}_4^i(\bar{z})=\boldsymbol{0}\\ \{\boldsymbol{0}\}; & \boldsymbol{F}_4^i(\bar{z})>\boldsymbol{0}\end{cases}$$

引理1（$\boldsymbol{\Omega}$在$\bar{z}$的法锥） 设 $\bar{z}\in\boldsymbol{\Omega}$满足 $\bar{w}^{X}_{i,j}\neq 0$，$\bar{w}^{Y}_{i,j}\neq 0$，$\bar{w}^{X}_{i,j}-(l_i+l_j)/2\neq\bar{w}^{Y}_{i,j}-(h_i+h_j)/2$，$(i,j)\in H$，$\bar{\boldsymbol{w}}^{\Lambda}>\boldsymbol{0}$，则 $\boldsymbol{\Omega}$ 在 $\bar{z}$ 的法锥可以表示为 $\boldsymbol{N}_{\boldsymbol{\Omega}}(\bar{z})=\{\nabla\boldsymbol{F}^{\mathrm{T}}(\bar{z})\boldsymbol{p}+\boldsymbol{q}\mid\boldsymbol{p}\in\boldsymbol{N}_{\boldsymbol{D}}(\boldsymbol{F}(\bar{z})),\boldsymbol{q}\in\boldsymbol{N}_{\boldsymbol{Z}}(\bar{z})\}$.

证明 根据文献[5]的定理6.14，只需证明下面的约束规范成立：

$$-\nabla\boldsymbol{F}^{\mathrm{T}}(\bar{z})\boldsymbol{p}\in\boldsymbol{N}_{\boldsymbol{Z}}(\bar{z}),\quad \boldsymbol{p}\in\boldsymbol{N}_{\boldsymbol{D}}(\boldsymbol{F}(\bar{z}))\Rightarrow\boldsymbol{p}=\boldsymbol{0}\tag{1}$$

令 $\boldsymbol{p}=(\boldsymbol{p}^{1\mathrm{T}}\ \cdots\ \boldsymbol{p}^{4\mathrm{T}})^{\mathrm{T}}\in\boldsymbol{N}_{\boldsymbol{D}}(\boldsymbol{F}(\bar{z}))$，则 $-\nabla\boldsymbol{F}^{\mathrm{T}}(\bar{z})\boldsymbol{p}\in\boldsymbol{N}_{\boldsymbol{Z}}(\bar{z})$ 可以写成 $\nabla\boldsymbol{F}_1^{\mathrm{T}}(\bar{z})(-\boldsymbol{p}^1)+\cdots+\nabla\boldsymbol{F}_4^{\mathrm{T}}(\bar{z})(-\boldsymbol{p}^4)$，或者写成矩阵的形式：

$$\begin{pmatrix}\boldsymbol{D}_1 & \boldsymbol{0}_{n\times M} & \boldsymbol{0}_{n\times M} & \boldsymbol{0}_{n\times n}\\ \boldsymbol{0}_{n\times M} & \boldsymbol{D}_2 & \boldsymbol{0}_{n\times M} & -\boldsymbol{I}_{n\times n}\\ -2\boldsymbol{W}^{X} & \boldsymbol{0}_{M\times M} & \boldsymbol{A} & \boldsymbol{0}_{M\times n}\\ \boldsymbol{0}_{M\times M} & -2\boldsymbol{W}^{Y} & \boldsymbol{B} & \boldsymbol{0}_{M\times n}\\ \boldsymbol{0}_{M\times M} & \boldsymbol{0}_{M\times M} & \boldsymbol{C} & \boldsymbol{0}_{M\times n}\\ \boldsymbol{0}_{1\times M} & \boldsymbol{0}_{1\times M} & \boldsymbol{0}_{1\times M} & \boldsymbol{1}_{1\times n}\end{pmatrix}\begin{pmatrix}-\boldsymbol{p}^1\\ -\boldsymbol{p}^2\\ -\boldsymbol{p}^3\\ -\boldsymbol{p}^4\end{pmatrix}\in\boldsymbol{N}_{\boldsymbol{Z}}(\bar{z})$$

即

$$-\boldsymbol{D}_1\boldsymbol{p}^1\in\boldsymbol{N}_{\boldsymbol{Z}_x}(\bar{\boldsymbol{x}})$$
$$-\boldsymbol{D}_2\boldsymbol{p}^2+\boldsymbol{p}^4\in\boldsymbol{N}_{\boldsymbol{Z}_y}(\bar{\boldsymbol{y}})$$
$$2\boldsymbol{W}^{X}\boldsymbol{p}^1-\boldsymbol{A}\boldsymbol{p}^3\in\boldsymbol{N}_{\mathbf{R}^M}(\bar{\boldsymbol{w}}^{X})\tag{2}$$
$$2\boldsymbol{W}^{Y}\boldsymbol{p}^2-\boldsymbol{B}\boldsymbol{p}^3\in\boldsymbol{N}_{\mathbf{R}^M}(\bar{\boldsymbol{w}}^{Y})\tag{3}$$
$$-\boldsymbol{C}\boldsymbol{p}^3\in\boldsymbol{N}_{\mathbf{R}_+^M}(\bar{\boldsymbol{w}}^{\Lambda})$$
$$\sum_{i=1}^{n}(-\boldsymbol{p}_i^4)\in N_{\mathbf{R}}(\bar{v})$$

由 $\boldsymbol{p}^4\in N_{\mathbf{R}_+^n}(\boldsymbol{F}_4(\bar{z}))$ 知 $\boldsymbol{p}^4\leqslant\boldsymbol{0}_n$，此外由 $N_{\mathbf{R}}(\bar{v})=\{0\}$ 知 $\sum_{i=1}^{n}(-\boldsymbol{p}_i^4)=\boldsymbol{0}$，从而可得出 $\boldsymbol{p}^4=\boldsymbol{0}_n$. 因为 $\bar{\boldsymbol{w}}^{\Lambda}>\boldsymbol{0}$，可知 $N_{\mathbf{R}_+^M}(\bar{\boldsymbol{w}}^{\Lambda})=\{\boldsymbol{0}_M\}$，从而 $\boldsymbol{C}\boldsymbol{p}^3=\boldsymbol{0}_M$. 其中 $\boldsymbol{C}$ 非奇异（否则，至少有一个 $\gamma_{i,j}=0$，这意味着 $\bar{w}^{X}_{i,j}-(l_i+l_j)/2=\bar{w}^{Y}_{i,j}-(h_i+h_j)/2$，与假设矛盾），因此得到 $\boldsymbol{p}^3=\boldsymbol{0}_M$. 由式(2)和(3)得到 $\boldsymbol{W}^{X}\boldsymbol{p}^1=\boldsymbol{0}_M$，$\boldsymbol{W}^{Y}\boldsymbol{p}^2=\boldsymbol{0}_M$（因为 $\boldsymbol{N}_{\mathbf{R}^M}(\bar{\boldsymbol{w}}^{X})=\boldsymbol{N}_{\mathbf{R}^M}(\bar{\boldsymbol{w}}^{Y})=$

$\{\mathbf{0}_M\})$. 又因为 $\mathbf{W}^X$ 和 $\mathbf{W}^Y$ 是非奇异的，得出 $\boldsymbol{p}^1 = \boldsymbol{p}^2 = \mathbf{0}_M$. 至此证明了式(1)，从而计算 $\mathbf{N}_{\boldsymbol{\Omega}}(\bar{z})$ 的公式可以从文献[5]得到.

定理 1(非线性规划模型的一阶最优性条件) 设 $\bar{z} \in \boldsymbol{\Omega}$ 是非线性规划模型的一个局部最优解，满足 $\bar{\varpi}^X_{i,j} \neq 0, \bar{\varpi}^Y_{i,j} \neq 0, \bar{\varpi}^X_{i,j} - (l_i + l_j)/2 \neq \bar{\varpi}^Y_{i,j} - (h_i + h_j)/2, (i,j) \in H, \bar{\boldsymbol{w}}^\wedge > \mathbf{0}$，则存在一个向量 $\boldsymbol{p} \in \mathbf{N}_{\mathbf{D}}(\mathbf{F}(\bar{z}))$，使得 $-[\nabla \boldsymbol{f}_0(\bar{z}) + \nabla \mathbf{F}^{\mathrm{T}}(\bar{z})\boldsymbol{p}] \in \mathbf{N}_{\mathbf{Z}}(\bar{z})$.

证明　由文献[5]的定理 6.12 知 $-\nabla \boldsymbol{f}_0(\bar{z}) \in \mathbf{N}_{\boldsymbol{\Omega}}(\bar{z})$，再由引理 1 得证.

2　数值算法

在非线性规划模型中，约束 $z \in \mathbf{Z}$ 可以用简单的不等式来表示，进而非线性规划模型可以表述为

$$\begin{aligned} \min \quad & f_0(z) \\ \text{s.t.} \quad & c_i(z) = 0, \quad i = 1,\cdots,3M, \\ & c_i(z) \geq 0, \quad i = 3M+1,\cdots,4M+4n \end{aligned}$$

对于这个约束优化问题可以应用增广 Lagrange 方法通过求解一序列无约束优化问题得到解决，

$$\begin{aligned} \min \quad & P(z, \boldsymbol{\lambda}, \sigma) = \\ & f_0(z) + \sum_{i=1}^{3M} [-\lambda_i c_i(z) + 0.5\sigma_i c_i^2(z)] + \\ & \sum_{i=3M+1}^{4n+4M} \begin{cases} -\lambda_i c_i(z) + 0.5\sigma_i c_i^{\,2}(z); & c_i(z) < \lambda_i/\sigma_i \\ -0.5\lambda_i^{\,2}/\sigma_i; & \text{否则} \end{cases} \end{aligned}$$

其中 $\sigma > 0$ 是参数，$\boldsymbol{\lambda}$ 是 Lagrange 乘子.

算法 ALM：

步 1　给定初始值 $z_1 \in \mathbf{R}^{2n+3M+1}, \boldsymbol{\lambda}^{(1)} \in \mathbf{R}^{4M+4n}$，其中 $\lambda_i^{(1)} > 0 (i = 3M+1,\cdots,4M+4n)$；$\sigma_i^{(1)} > 0, i = 1,\cdots,4n+4M$；$\varepsilon \geq 0, k := 1$.

步 2　求解 $\min P(z, \boldsymbol{\lambda}^{(k)}, \sigma^{(k)})$ 得到 z_{k+1}；$\| c^{(-)}(z_{k+1}) \|_\infty \leq \varepsilon$(其中 $c_i^{(-)}(z_k) = \min\{0, c_i(z_k)\}$)，则停止.

步 3　对 $i = 1,\cdots,4M+4n$，令

$$\sigma_i^{(k+1)} = \begin{cases} \sigma_i^{(k)}; & |c_i^{(-)}(z_{k+1})| \leq 1/4 \, |c_i^{(-)}(z_k)| \\ \max[10\sigma_i^{(k)}, k^2]; & \text{否则} \end{cases}$$

步 4　计算

$\lambda_i^{(k+1)} = \lambda_i^{(k)} - \sigma_i^{(k)} c_i(z_{k+1}),\ i = 1,\cdots,3M$

$\lambda_i^{(k+1)} = \max\{\lambda_i^{(k)} - \sigma_i^{(k)} c_i(z_{k+1}), 0\}$，

$i = 3M+1,\cdots,4M+4n$，

$k := k+1$；转步 2.

大量的数值实验表明，对于中小规模问题，可以求得精确解. 本文选出 3 个例子说明，图 1(a)～(c)给出了对应于求得解的矩形排列.

在图 1(a)中，$n = 8$，

$\boldsymbol{l} = (3\ \ 3\ \ 5\ \ 6\ \ 8\ \ 4\ \ 5\ \ 7)$，

$\boldsymbol{h} = (20\ \ 10\ \ 6\ \ 4\ \ 6\ \ 5\ \ 5\ \ 4), L = 12$.

解　$v = 25$.

在图 1(b)中，$n = 9$，

$\boldsymbol{l} = (3\ \ 5\ \ 6\ \ 8\ \ 4\ \ 5\ \ 8\ \ 7\ \ 3)$，

$\boldsymbol{h} = (10\ \ 6\ \ 4\ \ 6\ \ 5\ \ 5\ \ 4\ \ 3\ \ 20), L = 12$.

解　$v = 28$.

在图 1(c)中，$n = 10$，

$\boldsymbol{l} = (4\ \ 5\ \ 4\ \ 9\ \ 4\ \ 8\ \ 7\ \ 8\ \ 9\ \ 8)$，

$\boldsymbol{h} = (21\ \ 10\ \ 8\ \ 8\ \ 2\ \ 7\ \ 6\ \ 5\ \ 4\ \ 3), L = 13$.

解　$v = 40$.

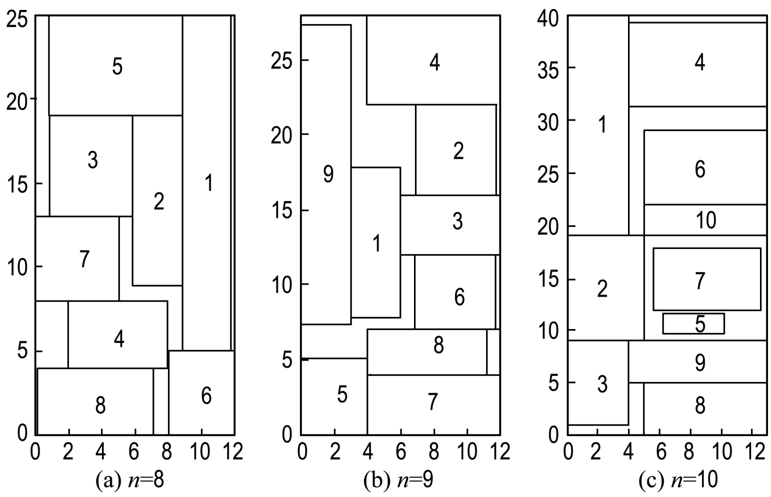

图 1　三种情况解的图示

Fig. 1　Illustrations of solutions for three cases

与经典的启发式算法如下次适应、首次适应和最佳适应相比，增广 Lagrange 算法在计算的精确度方面有很大提高，表1给出了针对上面3个例子各个算法的计算结果和增广 Lagrange 算法的计算时间.

表1 与经典的启发式方法的数值比较

Tab.1 Numerical comparisons with classical heuristic methods

n	$h_{\min}$				t_{ALM}/min
	NFDH 算法	FFDH 算法	BFDH 算法	ALM 算法	
8	35	35	37	25	3
9	38	38	41	28	9
10	54	54	53	40	35

注：n 为箱子数量；$h_{\min}$ 为最小高度；t_{ALM} 为 ALM 算法计算时间

3 结 语

本文给出二维装箱问题的非线性规划模型，以变分分析为工具建立了这一非线性规划问题的最优性条件，并用增广 Lagrange 方法对它求解.数值结果表明，对不超过10个物品的装箱问题，用该方法可以求得精确解，从计算时间上看，对这样规模的问题，该方法比经典的启发式方法要优越.如何把非线性规划方法用于有效求解大规模二维装箱问题是以后值得研究的课题.

参考文献：

[1] LODI A, MARTELLO S, VIGO D. Recent advances on two-dimensional bin packing problems [J]. **Discrete Applied Mathematics**, 2002, **123/124**: 373-390

[2] LODI A, MARTELLO S, VIGO D. Approximation algorithms for the oriented two-dimensional bin packing problem [J]. **European Journal of Operational Research**, 1999, **112**: 158-166

[3] 武晓今，朱仲英. 二维装箱问题的一种实现方法[J]. 微型电脑应用，2003，**19**(4)：20-23

[4] MARTELLO S, MONACI M, VIGO D. An exact approach to the strip-packing problem [J]. **INFORMS Journal on Computing**, 2003, **15**: 310-319

[5] ROCKAFELLAR R T, WETS R J B. **Variational Analysis** [M]. New York: Springer-Verlag, 1998

A nonlinear programming model for two-dimensional strip-packing problem and its numerical method

YU Hong-xia[1,2], ZHANG Shao-wu[3], ZHANG Li-wei*[1]

(1. Department of Applied Mathematics, Dalian University of Technology, Dalian 116024, China;
2. Department of Mathematics and Physics, Shanghai University of Electric Power, Shanghai 200090, China;
3. School of Electronic and Information Engineering, Dalian University of Technology, Dalian 116024, China)

Abstract: As a class of combinatorial optimization problems with many important applications, two-dimensional strip-packing problems are quite difficult to be solved accurately as they are NP hard. A two-dimensional strip-packing problem is formulated as a nonlinear programming (NLP) model and the notion of tangent cones in variational analysis is employed to establish the first-order optimality conditions for the NLP problem. The augmented Lagrange method is presented to solve this NLP problem and specific problems are solved by it. Numerical results show that the augmented Lagrange method is suitable for solving this NLP problem and it is able to find exact solutions to strip-packing problems involving up to 10 items.

Key words: two-dimensional strip-packing problem; first-order optimality conditions; augmented Lagrange method