

解决线性规划系统识别问题的新方法

杨德权*, 王佳

(大连理工大学 系统工程研究所, 辽宁 大连 116024)

摘要: 对一类特殊的逆线性规划问题——线性规划系统识别——进行研究,即试图通过给定的输入-输出数据来估计线性规划模型的技术系数矩阵以及目标函数系数. 构建了估计技术系数矩阵的行估计模型,并对该模型进行改进得到更好的估计模型;基于 Troutt 提出的最大决策效率方法,构建了估计标准化目标函数系数的模型;通过两个数值算例说明该估计方法具有良好的表面有效性,且符合提出的后续验证准则.

关键词: 运筹学;逆线性规划;估计模型;线性规划系统识别

中图分类号: O221

文献标识码: A

doi: 10.7511/dllgxb201401021

0 引言

线性规划是运筹学中运用最广泛的技术之一,现实生活中有很多问题都被简化为线性规划问题来求解^[1],比如在复杂的热力学系统中化学平衡的计算^[2]、水电站水库优化调度^[3]等. 在建模过程中,所有的参数,包括技术系数矩阵、目标函数系数以及右端项向量均需要是已知的,但是这些参数并不易提前获知. 文献[4]分析表明,在实践中应用生产计划模型常因难以获取所需的数据而受到阻碍;文献[5]则指出,企业通常很难精确地阐明他们的目标以及技术系数矩阵的约束. 因此,通过适当的方法获取线性规划模型的技术系数矩阵以及目标函数系数就显得尤为重要.

Troutt 等^[6]提出了一类特殊的逆线性规划问题,称为线性规划系统识别(简称为 LPSI). 即通过给定的一组生产单元,包括实际决策向量以及可利用资源向量,来估计线性规划模型的技术系数矩阵以及目标函数系数的问题. 对此,文献[5-6]给出了解决方法. 在过去的几十年,已经有许多专家学者研究技术系数矩阵以及目标函数系数的估计方法. 如 Sengupta^[7]指出,利用随机系数回归法估计技术系数矩阵是值得研究的;而 Ray^[8]分析,利用随机系数回归法估计的系数可能会出现负值;进而, Dixon 等^[9]通过改进随机系

数回归法,提出了严格不等式约束的随机系数回归法以估计非负的系数矩阵. 另外,文献[10-12]针对目标函数的估计,考虑了如下逆优化问题:尽可能小地调整线性规划模型的目标函数系数,使得给定的可行解成为最优解.

然而, LPSI 问题与以往的研究不同在于:它是在给定同一个前提下,即给定输入-输出数据,同时估计模型技术系数矩阵与目标函数系数的问题. 自 LPSI 问题提出以来,有关它的研究还不多,如文献[5]给出了一个启发式算法来解决 LPSI 问题. 因此,本文希望做出一些努力使得 LPSI 问题更为人所知,也试图提出新方法来解决 LPSI 问题.

本文先概要介绍 LPSI 问题,然后针对 LPSI 问题提出解决方法:首先,构建估计技术系数矩阵的行估计模型(简称 RE 模型),并加以改进以消除 RE 模型潜在的问题,得出改进的 RE 模型(简称 MRE 模型);其次,回顾文献[5]估计目标函数系数的最大决策效率方法(简称 MDEA 模型),通过消除 MDEA 模型的不足,构建估计标准化目标函数系数的模型(简称 IMDEA 模型);再次,提出后续验证准则,以检验估计方法得到的结果是不是有效的. 接下来给出了两个数值算例以验证估计方法的准确性,并与文献[5]的结果进行比较.

1 LPSI 问题的阐述

线性规划问题通过给定输入输出来获得最优决策,其中输入是技术水平以及目标函数系数,输出是资源右端项;而 LPSI 问题给定的输入输出分别为可利用资源向量与实际决策向量,用来估计技术系数矩阵以及目标函数系数.这说明,LPSI不只是一类特殊的逆线性规划问题,还属于给定输入输出的最优实现规划问题的范畴.本文接下来给出 LPSI 的定义,并沿用文献[6]中的标记方式,除非是特别说明,向量均指的是列向量.

假设给定 T 组观测数据,包括实际决策向量和可利用资源向量.对于每组观测指数 $t(t = 1, \cdots, T)$,实际决策向量 $\mathbf{y}^t$,也称输出向量,其每个元素为 $y_r^t(r = 1, \cdots, R)$;可利用资源向量即输入向量 $\mathbf{x}^t$,其每个元素为 $x_i^t(i = 1, \cdots, I)$.假定:(1)存在非负技术系数矩阵 $\mathbf{A}$,对所有的指数 t ,满足 $\mathbf{A}\mathbf{y}^t \leq \mathbf{x}^t$; (2)存在非负目标函数系数 $\boldsymbol{\pi}$,其每个元素为 π_r ,使得 $\mathbf{y}^t$ 近似或者恰好是如下线性规划模型 P_t 的最优解:

$$\begin{aligned} P_t: \max \quad & z^{t+} = \boldsymbol{\pi}'\mathbf{y} \\ \text{s. t.} \quad & \mathbf{A}\mathbf{y} \leq \mathbf{x}^t; t = 1, \cdots, T \\ & y_r \geq 0; r = 1, \cdots, R \end{aligned}$$

即 LPSI 问题求解的是满足假定(1)的非负技术系数矩阵 $\mathbf{A}$ 以及满足假定(2)的非负目标函数系数 $\boldsymbol{\pi}$.

2 估计非负技术系数矩阵

2.1 行估计模型

设待估计的非负技术系数矩阵为 $\mathbf{A} = (a_{ir})_{I \times R}$.根据假定(1), $\mathbf{A}$ 的每一行元素均需要满足 T 组约束条件,即对于 $\mathbf{A}$ 的第 i 行元素来说,满足 $\sum_{r=1}^R a_{ir}y_r^t \leq x_i^t, t = 1, \cdots, T$.据此,构建 $\mathbf{A}$ 的第 i 行元素的行估计模型(简称 RE 模型)如下:

$$\begin{aligned} \max \quad & \sigma_{t=1}^T \left(\sum_{r=1}^R a_{ir}y_r^t \right) \\ \text{s. t.} \quad & \sum_{r=1}^R a_{ir}y_r^t \leq x_i^t \\ & a_{ir} \geq 0 \end{aligned}$$

其中 σ 为聚合函数,可选择 $\sum$ (求和)、 $\prod$ (求积)、 $\min$ (求最小)等.本文选取的聚合函数为 $\sum$,故 RE 模型的目标函数变为

$$\max \sum_{t=1}^T \left(\sum_{r=1}^R a_{ir}y_r^t \right)$$

通过构造 I 个 RE 模型来估计出 $\mathbf{A}$ 的每一行元素,从而得到矩阵 $\mathbf{A}$.很明显,估计出来的 $\mathbf{A}$ 满足假定条件(1),而且是唯一的.

RE 模型将假定(1)作为约束条件来建模,可以保证求得的矩阵 $\mathbf{A}$ 是符合要求的.另外,如果 LPSI 问题对 $\mathbf{A}$ 没有非负限制,那么将 RE 模型的约束条件 $a_{ir} \geq 0$ 去掉仍然可以求解 $\mathbf{A}$,也就是可以拓宽 RE 模型的适用范围.但是,如果存在某组观测数据 t_0 远离分布整体,即 $\mathbf{x}^{t_0}$ 与 $\mathbf{y}^{t_0}$ 是奇异值,那么这组数据对估计技术系数矩阵会产生影响,通过 RE 模型估计的结果可能不会很准确.因此有必要消除这一潜在问题.

2.2 改进的 RE 模型

在 RE 模型的基础上,对 T 组约束条件的每一组条件加上权重,即 λ_t 对应第 $t(t = 1, \cdots, T)$ 组约束条件,得到改进的 RE 模型如下:

$$\begin{aligned} \max \quad & \sum_{t=1}^T \lambda_t \left(\sum_{r=1}^R a_{ir}y_r^t \right) \\ \text{s. t.} \quad & \sum_{r=1}^R a_{ir}y_r^t \leq x_i^t \\ & \sum_{t=1}^T \lambda_t = 1 \\ & a_{ir} \geq 0 \end{aligned}$$

λ_t 可由间距法^[13]生成.间距法简述如下:

设 $U_1, U_2, \cdots, U_{T-1}$ 是 $[0, 1]$ 上的均匀随机数,对它们进行升序排列变为 $U_{[1]}, U_{[2]}, \cdots, U_{[T-1]}$.定义间距 $\lambda_1, \lambda_2, \cdots, \lambda_T$ 为 $\lambda_1 = U_{[1]}, \cdots, \lambda_{T-1} = U_{[T-1]} - U_{[T-2]}, \lambda_T = 1 - U_{[T-1]}$.由此生成了一组数据 $(\lambda_1, \lambda_2, \cdots, \lambda_T)$,满足 $\sum_{t=1}^T \lambda_t = 1$.

引入权重 λ_t 是对约束条件 $\sum_{r=1}^R a_{ir}y_r^t \leq x_i^t$ 的一种约束,而约束条件又与观测数据有关,因此该组观测数据对估计结果的作用会被影响.由于随机产生的权重 λ_{t_0} 可能近似为 0,那么即使某组观测数据 t_0 远离分布整体(比如 $\mathbf{x}^{t_0}$ 与 $\mathbf{y}^{t_0}$ 数据偏大),MRE 模型也可能避免过多地受到该组数据的影响,甚至是不受影响.相比 RE 模型,通过引入权重 λ_t 构造的 MRE 模型,对给定的数据更敏感.因此,MRE 模型将能消除 RE 模型潜在的问题,也说明 MRE 模型优于 RE 模型.

3 估计标准化目标函数系数

3.1 回顾 MDEA 模型

MDEA 模型是由文献[5]提出的,能够估计标准化目标函数系数.下文中构建的估计目标函数系数的模型是在 MDEA 模型的基础上改进得到的,因此有必要介绍一下 MDEA 模型的建模过程.

P_i 对应的对偶模型 Q_i 为

$$\begin{aligned} \min \quad & z^{\prime-} = \xi^{\prime} x^t \\ \text{s. t.} \quad & A^{\prime} \xi^t \geq \pi; \\ & \xi^t \geq 0; t = 1, \dots, T \end{aligned}$$

设 y^{t*} 为模型 P_i 的最优解, ξ^{t*} 为模型 Q_i 的最优解.由假定(1)可知, y^t 为 P_i 的可行解,则有 $\pi^{\prime} y^t \leq \pi^{\prime} y^{t*}$.由对偶原理可知, $\pi^{\prime} y^{t*} = \xi^{t* \prime} x^t$,故 $\pi^{\prime} y^t \leq \xi^{t* \prime} x^t$.引入决策效率 $v^t = \pi^{\prime} y^t / \pi^{\prime} y^{t*}$,则 $v^t = \pi^{\prime} y^t / \xi^{t* \prime} x^t$,且 $0 \leq v^t \leq 1, t = 1, \dots, T$.

接下来,引入一个标准化的约束条件 $\pi^{\prime} y^{t_0} = 1, t_0$ 为某一特定的观测指数,这并不会影响决策效率值.因为如果模型 P_i 中的目标函数系数 π 用 $k\pi (k > 0)$ 替换,那么相对应的对偶模型 Q_i 中的对偶变量同样乘上了 k ,这并没有改变决策效率 v^t 的大小.

最后,利用文献[14]中的最大决策效率原则,以各组观测数据的决策效率之和作为目标函数,文献[5]提出了估计标准化目标函数系数 π 的 MDEA 模型:

$$\begin{aligned} \max \quad & \sum_{t=1}^T \pi^{\prime} y^t / \xi^{t \prime} x^t \\ \text{s. t.} \quad & A^{\prime} \xi^{t \prime} \geq \pi; \\ & \pi^{\prime} y^t \leq \xi^{t \prime} x^t; \\ & \sum_{r=1}^R \pi_r y_r^1 = 1; \\ & \pi_r, \xi_i^t \geq 0; t=1, \dots, T; r, i=1, \dots, R \end{aligned}$$

3.2 MDEA 模型的潜在问题

通过已估计的技术系数矩阵 A ,可以得到实际利用资源向量 Ay^t ,用 x^t 表示.很明显,可利用资源向量 x^t 不一定与实际利用资源向量 x^t 相等,除非构造出的技术系数矩阵 A 能够满足所有的观测数据.但这样的 A 不仅与构造方法有关,还与数据本身有一定的关系.因此,可能存在某些 t_0 ,使得 $Ay^{t_0} < x^{t_0}$.此时 y^{t_0} 显然不是模型 P_{t_0} 可行域的顶点,即不是模型 P_{t_0} 的最优解,所以 $\pi^{\prime} y^{t_0} <$

$\pi^{\prime} y^{t_0*}$.又由于 MDEA 模型的约束条件 $\pi^{\prime} y^{t_0} \leq \xi^{t_0* \prime} x^{t_0}$ 会导致解出的 $\xi^{t_0* \prime} x^{t_0}$ 可能与 $\pi^{\prime} y^{t_0}$ 很接近,其中 $\xi^{t_0* \prime} x^{t_0}$ 为模型 Q_{t_0} 的最优值,故可能出现模型 Q_{t_0} 的最优值不等于甚至是小于模型 P_{t_0} 的最优值的情况.显然这是不合理的.利用 MDEA 模型估计出的目标函数系数求解原模型与对偶模型时发现,它们的最优值不相等,说明 MDEA 模型在建模时存在问题,这也正是本文接下来提出改进 MDEA 模型的出发点.

3.3 改进 MDEA 模型

首先,用实际利用资源向量 x^t 代替模型中的可利用资源向量 x^t .因此原模型 P_i 变为 P_i' :

$$\begin{aligned} \max \quad & z^{\prime+} = \pi^{\prime} y \\ \text{s. t.} \quad & Ay \leq x^t; t = 1, \dots, T \\ & y_r \geq 0; r = 1, \dots, R \end{aligned}$$

对偶模型 Q_i 变为 Q_i' :

$$\begin{aligned} \min \quad & z^{\prime-} = \xi^{t \prime} x^t \\ \text{s. t.} \quad & A^{\prime} \xi^t \geq \pi; t = 1, \dots, T \\ & \xi_i \geq 0; i = 1, \dots, R \end{aligned}$$

对于模型 P_i' 来说,由于 $Ay \leq Ay^t$,故 y^t 更加近似或者就是其最优解;从而 $\pi^{\prime} y^t$ 近似或者等于最优值.这将会降低模型 P_i' 与 Q_i' 的最优值不等情况发生的概率.

其次,本文采用的目标函数是线性函数

$\sum_{t=1}^T \pi^{\prime} y^t - \xi^{t \prime} x^t$,代替 MDEA 模型的非线性函数 $\sum_{t=1}^T \pi^{\prime} y^t / \xi^{t \prime} x^t$,目的是使得模型 P_i' 与 Q_i' 的最优值相差尽可能的小.原因如下:由弱对偶性可知,模型 P_i', Q_i' 的任一可行解 $\bar{y}^t, \bar{\xi}^t$ 满足 $\pi^{\prime} \bar{y}^t \leq \bar{\xi}^t x^t$.采用线性函数 $\sum_{t=1}^T \pi^{\prime} y^t - \xi^{t \prime} x^t$ 作为目标函数,由于 $\pi^{\prime} y^t - \xi^{t \prime} x^t$ 是非正数,那么通过取最大值就可以使得 $\pi^{\prime} y^t, \xi^{t \prime} x^t$ 这两个函数值的差值尽可能地小,即两个函数值更接近甚至相等,同时也能满足后文提出的后续验证准则.

综上所述,改进的 MDEA 模型(简称 IMDEA 模型)如下:

$$\begin{aligned} \max \quad & \sum_{t=1}^T \pi^{\prime} y^t - \xi^{t \prime} x^t \\ \text{s. t.} \quad & A^{\prime} \xi^{t \prime} \geq \pi; \\ & \pi^{\prime} y^t \leq \xi^{t \prime} x^t; \end{aligned}$$

$$\sum_{r=1}^R \pi_r y_r^1 = 1;$$
$$\pi_r, \xi_i^t \geqslant 0; t=1, \cdots, T; r, i=1, \cdots, R$$

4 判定结果的合理性

4.1 后续验证准则

对估计方法的准确性进行验证是必要的,但是文献[5]并没有这么做. 本文提出了一个后续验证准则: 利用估计出的目标函数系数 π 与技术系数矩阵 A , 求解模型 P'_i 与 Q'_i 的最优值, 如果解出的两个最优值相等, 那么估计出的 A 与 π 也就是有效的; 反之则说明估计方法不准确. 在最优值相等的情况下, 称 A 与 π 分别为有效技术系数矩阵与有效标准化目标函数系数.

4.2 最优估计结果的选择方法

本文针对 LPSI 问题提出的解决思路是, 先利用 RE 模型或 MRE 模型估计出技术系数矩阵, 然后再利用 IMDEA 模型以及估计的技术系数矩阵来估计目标函数系数, 并称这两个估计量为一组估计结果. 由于 MRE 模型可以估计出任意多个技术系数矩阵, 对应地能够估计出多个目标函数系数, 即得到多组估计结果. 因此, 有必要选出一组最优的估计结果. 选择方法如下: 首先,

需要保证估计结果符合后续验证准则, 即要求有效的技术系数矩阵与目标函数系数; 其次, 每一组估计结果均需求解 T 个模型 P'_i , 从而得到 T 个最优值, 其最大值为该估计结果所能达到的最优值; 最后, 对多个估计结果所对应的最优值进行比较, 其最大值为 MRE 模型所能达到的最优值, 而且对应的 A 与 π 为最优的估计结果.

5 数值算例

5.1 15 个医院的测试数据

本文采用的例子与文献[5]中的一致, 是来自于 Sherman^[15] 的 15 个医院的测试数据, 见表 1. 本文数据均采用软件 Matlab R2009a 编程所得.

如果需要详细了解这些数据的来源可以查看文献[16], 在这里仅作简要描述: 给定一个初始的技术系数矩阵

$$A^0 = \begin{pmatrix} 0.004 & 0.005 & 0.003 \\ 7 & 9 & 0 \\ 20 & 30 & 200 \end{pmatrix}$$

以及 15 个输出向量 y (表中数据), 再利用 $x = A^0 y$ 求出输入向量.

表 1 15 个医院的测试数据

Tab. 1 15-Hospital test datasets

医院	输入向量			输出向量		
	医院医务人员的数量	医院可接纳病人的床位数	每年医院设备的成本	医学院学生人数	普通病患	严重病患
H1	23.5	41 050	130 000	50	3 000	2 000
H2	24.5	43 160	140 000	50	2 000	3 000
H3	26.0	43 160	150 000	100	2 000	3 000
H4	25.0	41 050	140 000	100	3 000	2 000
H5	28.5	50 530	160 000	50	3 000	3 000
H6	36.0	62 105	210 000	100	2 000	5 000
H7	51.5	92 630	270 000	50	10 000	2 000
H8	25.0	49 475	140 000	100	3 000	2 000
H9	24.5	43 160	165 000	50	2 000	3 000
H10	77.0	92 630	340 000	100	10 000	2 000
H11	44.5	65 260	265 000	50	5 000	3 000
H12	30.0	60 000	170 000	100	10 000	2 000
H13	43.5	81 110	245 000	50	4 000	5 000
H14	30.0	60 000	170 000	100	3 000	3 000
H15	26.5	47 370	160 000	50	3 000	2 000

注: 数据来自文献[5-6](下同)

5.1.1 验证 MDEA 模型估计结果存在的问题

文献 [5] 得到的技术系数矩阵 $\mathbf{A} = \begin{pmatrix} 0.004 & 0.005 & 0.002 \\ 7.391 & 9.453 & 0.188 \\ 21.020 & 23.268 & 265.267 \end{pmatrix}$, 目标函数系数 $\boldsymbol{\pi} =$

$(0.000\ 179\ 7\ 0.000\ 230\ 4\ 0.000\ 004\ 6)$. 通过验证发现,模型 $\mathbf{P}_i$ 与 $\mathbf{Q}_i$ 的最优值并不相等,最大偏差达到 0.15,甚至 H10 的对偶模型的最优值小于原模型的最优值,因此不满足后续验证准则,说明估计的结果不是有效的. 详细数据参见表 2.

表 2 应用 MDEA 模型估计的详细数据

Tab. 2 Detailed data of MDEA model estimation

医院	$\mathbf{P}_i$ 最优值	$\mathbf{Q}_i$ 最优值	$\mathbf{Q}_i$ 与 $\mathbf{P}_i$ 最优值差
H1	0.999 999 999 999 949 8	1.000 526 464 595 344	$5.264\ 645\ 953\ 941\ 827 \times 10^{-4}$
H2	1.050 745 337 376 003	1.051 954 256 076 444	$1.208\ 918\ 700\ 441\ 153 \times 10^{-3}$
H3	1.050 974 446 968 397	1.051 954 256 076 307	$9.798\ 091\ 079\ 100\ 679 \times 10^{-4}$
H4	1.000 229 109 592 344	1.000 526 464 595 290	$2.973\ 550\ 029\ 461\ 425 \times 10^{-4}$
H5	1.230 401 380 507 093	1.231 585 925 846 469	$1.184\ 545\ 339\ 375\ 953 \times 10^{-3}$
H6	1.511 777 207 982 683	1.513 707 578 165 430	$1.930\ 370\ 182\ 746\ 888 \times 10^{-3}$
H7	2.257 592 301 917 579	2.257 704 419 377 115	$1.121\ 174\ 595\ 359\ 875 \times 10^{-4}$
H8	1.000 229 109 592 344	1.152 006 902 535 736	$1.517\ 777\ 929\ 433\ 921 \times 10^{-1}$
H9	1.050 745 337 376 003	1.051 954 256 076 311	$1.208\ 918\ 700\ 308\ 148 \times 10^{-3}$
H10	2.257 821 411 509 974	2.257 821 411 507 480	$-2.494\ 005\ 002\ 517\ 952 \times 10^{-12}$
H11	1.589 713 466 769 273	1.590 605 531 778 117	$8.920\ 650\ 088\ 439\ 786 \times 10^{-4}$
H12	1.230 630 490 099 487	1.382 408 283 042 885	$1.517\ 777\ 929\ 433\ 979 \times 10^{-1}$
H13	1.870 860 184 652 468	1.976 923 301 908 404	$1.060\ 631\ 172\ 559\ 361 \times 10^{-1}$
H14	1.230 630 490 099 487	1.382 408 283 042 915	$1.517\ 777\ 929\ 434\ 279 \times 10^{-1}$
H15	0.999 999 999 999 949 8	1.154 566 105 429 773	$1.545\ 661\ 054\ 298\ 231 \times 10^{-1}$

5.1.2 验证本文所提方法的准确性 利用文献 [5]估计的技术系数矩阵,通过 IMDEA 模型估计目标函数系数 $\boldsymbol{\pi} = (0.000\ 176\ 4\ 0.000\ 212\ 0\ 0.000\ 934\ 1)$. 通过验证发现,模型 $\mathbf{P}'_i$ 与 $\mathbf{Q}'_i$ 的最优值近似相等,详细数据参见表 3. 考虑到计算机的误差,可以近似地认为二者是相等的. 故判断 IMDEA 模型是有效的.

利用 RE 模型得到的技术系数矩阵,以及利用 IMDEA 模型得到的目标函数系数如下:

$$\mathbf{A}^1 = \begin{pmatrix} 0.004 & 0.005 & 0.003 \\ 7.367 & 9.474 & 0 \\ 20 & 30 & 200 \end{pmatrix}$$

$\boldsymbol{\pi}^1 = (0.000\ 168\ 3\ 0.000\ 225\ 1\ 0.000\ 896\ 0)$, 最优值 $z = 2.223\ 2$,而且模型 $\mathbf{P}'_i$ 与 $\mathbf{Q}'_i$ 的最优值近似相等,详细数据参见表 4.

为了方便起见,本文仅利用 MRE 模型构造 30 个 $\mathbf{A}$. 按照 4.2 给出的选择最优估计结果的方法,得到最优的技术系数矩阵以及目标函数系数如下:

$$\mathbf{A}^2 = \begin{pmatrix} 0.004 & 0.005 & 0.003 \\ 7.369 & 9.472 & 0 \\ 20 & 30 & 200 \end{pmatrix}$$

$\boldsymbol{\pi}^2 = (0.000\ 168\ 4\ 0.000\ 225\ 1\ 0.000\ 896\ 0)$, 最优值 $z = 2.223\ 3$,而且模型 $\mathbf{P}'_i$ 与 $\mathbf{Q}'_i$ 的最优值近似相等,详细数据参见表 5.

比较最优值可知,MRE 模型的结果比 RE 模型的结果要好一些. 而且,RE 模型与 MRE 模型估计的最优有效技术系数矩阵与初始的 $\mathbf{A}^0$ 相比都很接近,相比文献[5]估计的结果要精确很多,说明本文提出的模型在处理该问题上取得了令人满意的结果.

5.2 IMDEA 模型的进一步验证

由于 5.1 中的数值算例并没有给出初始的目标函数系数,这就导致不能直观地判断 IMDEA 模型估计的精度. 接下来,本文根据上一个算例,构造出一组新的数据来验证本文提出的 IMDEA 模型的准确性,并与 MDEA 模型进行比较. 数据构造的具体步骤如下:

表 3 应用 IMDEA 模型估计的详细数据

Tab.3 Detailed data of IMDEA model estimation

医院	P'_i 最优值	Q'_i 最优值	Q'_i 与 P'_i 最优值差
H1	0.999 999 999 999 996 0	1.000 000 000 007 065	$7.069\ 012\ 042\ 393\ 297\times10^{-12}$
H2	1.035 554 422 628 086	1.035 554 422 633 982	$5.895\ 950\ 394\ 574\ 356\times10^{-12}$
H3	1.082 257 702 188 163	1.082 257 702 195 341	$7.178\ 035\ 943\ 411\ 487\times10^{-12}$
H4	1.046 703 279 560 073	1.046 703 279 565 977	$5.903\ 944\ 000\ 351\ 658\times10^{-12}$
H5	1.211 991 997 664 834	1.211 991 997 670 996	$6.162\ 181\ 875\ 879\ 469\times10^{-12}$
H6	1.506 241 697 517 838	1.506 241 697 524 354	$6.515\ 898\ 931\ 525\ 044\times10^{-12}$
H7	2.235 063 025 257 232	2.235 063 025 262 960	$5.727\ 862\ 628\ 646\ 108\times10^{-12}$
H8	1.046 703 279 560 073	1.046 703 279 565 973	$5.899\ 947\ 197\ 463\ 007\times10^{-12}$
H9	1.035 554 422 628 086	1.035 554 422 633 985	$5.899\ 059\ 019\ 043\ 307\times10^{-12}$
H10	2.281 766 304 817 309	2.281 766 304 824 028	$6.719\ 069\ 745\ 031\ 447\times10^{-12}$
H11	1.564 867 147 738 330	1.564 867 147 744 232	$5.901\ 945\ 598\ 907\ 332\times10^{-12}$
H12	1.258 695 277 224 911	1.258 695 277 231 551	$6.640\ 021\ 865\ 678\ 136\times10^{-12}$
H13	1.812 413 568 031 257	1.812 413 568 037 287	$6.030\ 065\ 335\ 949\ 075\times10^{-12}$
H14	1.258 695 277 224 911	1.258 695 277 231 543	$6.632\ 028\ 259\ 900\ 835\times10^{-12}$
H15	0.999 999 999 999 996 0	1.000 000 000 007 046	$7.049\ 916\ 206\ 369\ 744\times10^{-12}$

表 4 应用 RE 和 IMDEA 模型估计的详细数据

Tab.4 Detailed data of RE and IMDEA model estimation

医院	P'_i 最优值	Q'_i 最优值	Q'_i 与 P'_i 最优值差
H1	1.000 000 000 000 001	1.000 000 000 187 916	$1.879\ 147\ 948\ 358\ 195\times10^{-10}$
H2	1.056 748 987 537 059	1.056 748 987 539 184	$2.124\ 966\ 869\ 132\ 550\times10^{-12}$
H3	1.101 547 760 067 539	1.101 547 760 073 645	$6.106\ 004\ 590\ 833\ 436\times10^{-12}$
H4	1.044 798 772 530 482	1.044 798 773 316 174	$7.856\ 921\ 779\ 847\ 426\times10^{-10}$
H5	1.225 089 638 016 139	1.225 089 638 114 207	$9.806\ 799\ 816\ 658\ 440\times10^{-11}$
H6	1.551 727 036 099 817	1.551 727 036 250 685	$1.508\ 679\ 847\ 717\ 076\times10^{-10}$
H7	2.178 384 553 353 567	2.178 384 553 355 468	$1.901\ 145\ 907\ 368\ 118\times10^{-12}$
H8	1.044 798 772 530 482	1.044 798 773 316 174	$7.856\ 921\ 779\ 847\ 426\times10^{-10}$
H9	1.056 748 987 537 059	1.056 748 987 539 184	$2.124\ 966\ 869\ 132\ 550\times10^{-12}$
H10	2.223 183 325 884 048	2.223 183 328 665 452	$2.781\ 403\ 996\ 010\ 567\times10^{-9}$
H11	1.561 770 938 974 301	1.561 770 939 002 334	$2.803\ 290\ 932\ 718\ 028\times10^{-11}$
H12	1.269 888 410 546 620	1.269 888 410 554 036	$7.416\ 067\ 759\ 891\ 121\times10^{-12}$
H13	1.843 609 564 527 498	1.843 609 564 656 478	$1.289\ 801\ 598\ 858\ 276\times10^{-10}$
H14	1.269 888 410 546 620	1.269 888 410 554 036	$7.416\ 067\ 759\ 891\ 121\times10^{-12}$
H15	1.000 000 000 000 001	1.000 000 000 187 916	$1.879\ 147\ 948\ 358\ 195\times10^{-10}$

表 5 应用 MRE 和 IMDEA 模型估计的详细数据

Tab.5 Detailed data of MRE and IMDEA model estimation

医院	P'_i 最优值	Q'_i 最优值	Q'_i 与 P'_i 最优值差
H1	1.000 000 000 000 007	1.000 000 000 190 240	$1.902\ 329\ 405\ 112\ 369\times10^{-10}$
H2	1.056 721 396 310 162	1.056 721 396 312 631	$2.468\ 913\ 962\ 161\ 423\times10^{-12}$
H3	1.101 506 238 026 792	1.101 506 238 033 088	$6.296\ 074\ 772\ 649\ 263\times10^{-12}$
H4	1.044 784 841 716 637	1.044 784 842 410 155	$6.935\ 179\ 097\ 666\ 833\times10^{-10}$
H5	1.225 075 869 442 775	1.225 075 869 542 487	$9.971\ 201\ 642\ 144\ 933\times10^{-11}$
H6	1.551 657 976 912 328	1.551 657 977 047 386	$1.350\ 579\ 648\ 118\ 355\times10^{-10}$
H7	2.178 481 311 928 300	2.178 481 311 930 006	$1.705\ 746\ 655\ 034\ 091\times10^{-12}$
H8	1.044 784 841 716 637	1.044 784 842 410 155	$6.935\ 179\ 097\ 666\ 833\times10^{-10}$
H9	1.056 721 396 310 162	1.056 721 396 312 631	$2.468\ 913\ 962\ 161\ 423\times10^{-12}$
H10	2.223 266 153 644 930	2.223 266 156 444 689	$2.799\ 759\ 091\ 232\ 090\times10^{-9}$
H11	1.561 784 815 708 002	1.561 784 815 735 921	$2.791\ 900\ 044\ 485\ 374\times10^{-11}$
H12	1.269 860 711 159 405	1.269 860 711 166 536	$7.131\ 184\ 531\ 772\ 306\times10^{-12}$
H13	1.843 582 081 460 925	1.843 582 081 588 516	$1.275\ 910\ 488\ 374\ 166\times10^{-10}$
H14	1.269 860 711 159 405	1.269 860 711 166 536	$7.131\ 184\ 531\ 772\ 306\times10^{-12}$
H15	1.000 000 000 000 007	1.000 000 000 190 240	$1.902\ 329\ 405\ 112\ 369\times10^{-10}$

步骤 1 给定一个目标函数系数 π^* .

例如 $\pi^* = (0.000\ 167\ 9\ \ 0.000\ 225\ 0\ 0.000\ 922\ 8)$,它是利用 IMDEA 模型以及初始的 A^0 求解出来的,作为已知的目标函数系数.

步骤 2 根据初始的 A^0 以及表 1 中的 x' ,再加上步骤 1 的 π^* ,解模型 P_i 可得最优解 y'^* 来替换表 1 的输出向量 y' . y'^* 的具体数据参见表 6.

步骤 3 为了满 足 IMDEA 模 型 以 及

MDEA 模型中的标准化约束 $\pi^{*'}y^1 = 1$,需要将给定的目标函数系数乘上一个因子 k ,即 $k \cdot \pi^{*'}y^1 = 1$. 求出 $k = \frac{1}{\pi^{*'}y^1}$,进而得到 $\pi^0 = \frac{1}{\pi^{*'}y^1}\pi^{*'}$,作为初始的目标函数系数.

步骤 4 最后,通过 IMDEA 模型以及 MDEA 模型估计目标函数系数,并与初始的目标函数系数进行比较,得出优劣结论.

表 6 输出向量
Tab. 6 Output vector

医院	普通病患	严重病患	医学院学生人数
H1	2.363 438 442 818 936	2.722 881 211 140 828	5.223 974 046 982 366
H2	1.803 636 363 315 986	3.392 727 272 944 763	10.727 272 723 144 12
H3	2.988 277 943 879 172	1.101 122 890 629 986	6.284 470 053 509 128 $\times 10^{-27}$
H4	2.813 636 364 049 030	2.372 727 272 369 520	62.727 272 723 724 75
H5	3.931 853 653 767 086	2.554 517 076 986 331	1.006 903 596 445 011 $\times 10^{-23}$
H6	3.041 729 501 591 740	4.342 707 978 025 737	70.651 403 450 145 17
H7	8.519 363 108 615 782	3.320 424 594 256 147	4.280 976 546 124 529 $\times 10^{-30}$
H8	2.499 999 999 999 929	3.000 000 000 000 047	1.070 545 743 534 063 $\times 10^{-14}$
H9	3.545 360 227 984 923 $\times 10^{-11}$	4.795 555 555 432 014	17.407 407 421 920 25
H10	3.397 530 776 346 849 $\times 10^{-4}$	10.291 957 969 828 51	156.172 329 217 960 9
H11	3.394 772 518 116 223	4.610 732 485 909 604	262.241 583 266 236 2
H12	3.152 270 887 915 659	3.478 183 289 667 472	2.972 274 467 890 750 $\times 10^{-23}$
H13	4.000 000 000 004 928	5.499 999 999 995 391	7.991 905 729 474 602 $\times 10^{-12}$
H14	3.152 270 887 915 659	3.478 183 289 667 472	2.972 274 467 890 750 $\times 10^{-23}$
H15	1.775 618 167 900 986	3.879 505 465 679 211	3.364 268 213 467 164 $\times 10^{-22}$

利用 IMDEA 模型与 MDEA 模型估计出来的目标函数系数分别为 (0.000 166 3 0.000 221 7 0.000 650 6), (0.000 170 5 0.000 219 2 0). 将估计出的目标函数系数的各个元素与初始目标函数系数的各个元素进行比较,得到绝对偏差百分比为(0.994% 1.483% 29.503%), (1.547% 2.568% 100%). 显然 IMDEA 模型估计的目标函数系数比 MDEA 模型更接近初始值,因此 IMDEA 模型更有效.

6 结 语

文献[6]提出的 LPSI 问题,即通过给定的输入-输出数据来估计技术系数矩阵以及目标函数系数,具有较大的理论意义和实用价值. 但是,文献[5]提出的解决方法存在两个明显的不足:一是估计技术系数矩阵的结果不够准确;二是估计标准化目标函数系数的 MDEA 模型在建模时存在缺陷,而且并没有验证估计结果是否符合建模的要求. 本文针对这两方面提出了改进方法. 首先,

提出了 RE 模型以及 MRE 模型,可以更好地估计技术系数矩阵. 其次,通过消除 MDEA 模型潜在的问题,得到了更好的 IMDEA 模型. 同时提出了后续验证准则,以验证所得的结果是否满足建模的要求. 最后,两个数值算例验证了本文提出的方法比文献[5]的方法有明显的优势.

参考文献：

[1] Erdoğmus P. Particle swarm optimization performance on special linear programming problems [J]. **Scientific Research and Essays**, 2010, 5(12):1506-1518.

[2] Belov G. On linear programming approach for the calculation of chemical equilibrium in complex thermodynamic systems [J]. **Journal of Mathematical Chemistry**, 2010, 47(1):446-456.

[3] CHENG Chun-tian, WANG Wen-chuan, XU Dong-mei, *et al.* Optimizing hydropower reservoir operation using hybrid genetic algorithm and chaos

- [J]. **Water Resources Management**, 2008, **22**(7): 895-909.
- [4] Troutt M D, Pang Wan-kai, Hou Shui-hung. Behavioral estimation of mathematical programming objective function coefficients [J]. **Management Science**, 2006, **52**(3):422-434.
- [5] Troutt M D, Brandyberry A A, Sohn C, *et al.* Linear programming system identification: the general nonnegative parameters case [J]. **European Journal of Operational Research**, 2008, **185**(1):63-75.
- [6] Troutt M D, Tadisina S K, Sohn C, *et al.* Linear programming system identification [J]. **European Journal of Operational Research**, 2005, **161**(3):663-672.
- [7] Sengupta J K. Regression and programming: overview of linkages [J]. **Arthaniti; a Bi-annual Journal of the Department of Economics, Calcutta University**, 1975, **17**:1-17.
- [8] Ray S. Methods of estimating the input coefficients for linear programming models [J]. **American Journal of Agricultural Economics**, 1985, **67**(3): 660-665.
- [9] Dixon B L, Hornbaker R H. Estimating the technology coefficients in linear programming models [J]. **American Journal of Agricultural Economics**, 1992, **74**(4):1029-1039.
- [10] HUANG Si-ming, LIU Zhen-hong. On the inverse problem of linear programming and its application to minimum weight perfect k -matching [J]. **European Journal of Operational Research**, 1999, **112**(2):421-426.
- [11] Sokkalingam P T, Ahuja R K, Orlin J B. Solving inverse spanning tree problems through network flow techniques [J]. **Operations Research**, 1999, **47**(2):291-298.
- [12] Ahuja R K, Orlin J B. Inverse optimization [J]. **Operations Research**, 2001, **49**(5):771-783.
- [13] Devroye L. **Non-uniform Random Variate Generation** [M]. New York:Springer-Verlag, 1986.
- [14] Troutt M D. A maximum decisional efficiency estimation principle [J]. **Management Science**, 1995, **41**(1):76-83.
- [15] Sherman H D. Measurement of hospital technical efficiency: A comparative evaluation of data envelopment analysis and other approaches for locating inefficiency in health care organizations [D]. Boston:Harvard University, 1981.
- [16] Bowlin W F, Charnes A, Cooper W W, *et al.* Data envelopment analysis and regression approaches to efficiency estimation and evaluation [J]. **Annals of Operations Research**, 1984, **2**(1):113-138.

A new method to solve problem of linear programming system identification

YANG De-quan*, WANG Jia

(Institute of Systems Engineering, Dalian University of Technology, Dalian 116024, China)

Abstract: A special kind of inverse linear programming called linear programming system identification, which seeks to estimate the technological coefficient matrix and the objective function coefficient vector with the given input-output data, is considered. The row estimation model and the modified row estimation model are constructed to estimate the technological coefficient matrix, and based on the maximum decision-making efficiency approach proposed by Troutt, an estimation model is constructed to estimate the objective function coefficient vector. It is found that the estimation method exhibits excellent face validity for two test datasets, and corresponds with the proposed subsequent validation criteria.

Key words: operational research; inverse linear programming; estimation model; linear programming system identification