

データ変換をともしなうDEA(その2)

— 一般化平均DEAのLP解法 —

日大生産工

○篠原 正明

情報システム研究所

篠原 健

1. はじめに

データ変換関数がべき乗関数の場合 ($f(x)=x^p, g(y)=y^q$) について、数理計画法による定式化を行い、それに基づき離散評点DEAによる表計算解法、NLP解法、LP解法を適用し、各解法の計算結果の整合性を検討する。

2. 数理計画法による定式化

$p \neq q$ の一般的な場合の DMU_k の相対効率値 R_k の最大化問題は(1)~(3)となる。

目的関数: (1)

$$R_k = \frac{\frac{(u^T y_k^{(q)})^{1/q}}{(v^T x_k^{(p)})^{1/p}}}{\max_j \left\{ \frac{(u^T y_j^{(q)})^{1/q}}{(v^T x_j^{(p)})^{1/p}} \right\}}$$

変数: u, v (2) 制約条件: $u \geq 0, v \geq 0$ (3)

目的関数 R_k の分母を 1 と固定すると、(4)~(6)となる。

$$\text{目的関数: } Z_k = \frac{(u^T y_k^{(q)})^{1/q}}{(v^T x_k^{(p)})^{1/p}} \rightarrow \text{最大化} \quad (4)$$

$$\text{制約条件: } \max_j \left\{ \frac{(u^T y_j^{(q)})^{1/q}}{(v^T x_j^{(p)})^{1/p}} \right\} = 1 \quad (5)$$

$$u \geq 0, v \geq 0 \quad (6)$$

さらに、目的関数 Z_k の分母を 1 と固定し、(5)を不等式条件に置換すると、(7)~(10)となる。

$$\text{目的関数: } Z'_k = (u^T y_k^{(q)})^{1/q} \rightarrow \text{最大化} \quad (7)$$

$$\text{制約条件: } (v^T x_k^{(p)})^{1/p} = 1 \quad (8)$$

$$(u^T y_j^{(q)})^{1/q} \leq (v^T x_j^{(p)})^{1/p} \quad j = \dots, n \quad (9)$$

$$u \geq 0, v \geq 0 \quad (10)$$

定式化(7)~(10)において、一般的に $p \neq q$ ならば制約式(9)は非線形制約で、それゆえに、非線形最適問題となる。もし、 $p=q>0$ ならば、(11)~(14)となり、通常のDEA・CCRモデルと同様なLP問題に帰着する。

$$\text{目的関数: } Z'_k = (u^T y_k^{(p)}) \rightarrow \text{最大化} \quad (11)$$

$$\text{制約条件: } v^T x_k^{(p)} = 1 \quad (12)$$

$$u^T y_j^{(p)} \leq v^T x_j^{(p)} \quad j = \dots, n \quad (13) \quad u \geq 0, v \geq 0 \quad (14)$$

3. 効率性評価例

3.1 7DMU 1入力2出力CCRモデルの例

7人の学生jに対する数(y_{1j})、理(y_{2j})の成績評価データ(表1)に適用する。

表1: 1入力2出力データ (例3.1)

生徒	A	B	C	D	E	F	G
入力 授業時間	1	1	1	1	1	1	1
出力 数学	3	6	9	10	1	5	6
理科	10	9	7	3	9	5	4

入力項目数=1かつデータ値が同じ($x_j=1$)なので、任意の p について、仮想入力 $v x_j$ は省略でき、以下の R_k^q の最大化に問題(15)、(16)に帰着する($q>0$)。以下に $q=0.5, 1, 2$ の場合のDEA型成績評価問題を扱うが、成績点数データにデータ変換をほどこす理由については付録1を参照のこと。

$$\text{目的関数: } R_k(u_1, u_2)^q = \frac{u_1 y_{1k}^q + u_2 y_{2k}^q}{\max_j \{u_1 y_{1j}^q + u_2 y_{2j}^q\}} \rightarrow \text{最大化} \quad (15)$$

$$\text{変数: } u = (u_1, u_2) \quad (16)$$

[離散評点DEA・表計算解法]

DEA with Data Transformation
Part 2: LP Solution to Generalized Mean DEA

Masaaki SHINOHARA and Ken SHINOHARA

u_1, u_2 を $\{1, 2, 3, 4\}$ と離散評点の組み合わせで表現し、最大化問題(15),(16)を解いた結果を表 2 に示す。又、表 3 には表計算シートの一部を示すが、新たなウェイトの組み合わせ(u_1, u_2)を追加して、よりきめ細かい包絡面を実現した。

表 2: 離散評点 DEA を適用した計算結果 (例 3.1)

学生		A	B	C	D	E	F	G
DEA効率値	P=1	1	1	1	1	0.9	0.641	0.651
	P=2	1	1	1	1	0.9	0.633	0.646
	P=0.5	1	1	1	1	0.9	0.644	0.652

表 3: 離散評点 DEA 表計算シートの一部 (例 3.1)

		A	B	C	D	E	F	G		
出力1値		3	6	9	10	1	5	6		
出力2値		10	9	7	3	9	5	4		
u1	u2									
0.05	0.05	1.53691	5.989128	0.79154	1.7424	7.426	0.932	1.7369	7.9896	-1
0.030	0.07	1.721213	7.212129	0.909169	1.7333	7.933	1	1.7333	7.933	0.963
0.05	0.75	1.736646	7.664594	0.969123	1.8183	8.183	1	1.7470	7.4705	0.9125
0.2	0.8	1.822721	8.272722	0.989565	1.8352	8.352	1	1.733	7.739	0.8837
0.087	0.33	1.437375	4.973749	0.587415	1.6333	6.633	0.633	1.6335	6.3355	-1
1	0	1	1	1	1	1	1	1	1	1
0.4	0.6	1.670807	6.708069	0.863471	1.7727	7.727	0.863	1.777	7.769	-1
0.030	0.07	1.721213	7.212129	0.909169	1.7333	7.933	1	1.7333	7.933	0.963
0.05	0.75	1.736646	7.664594	0.969123	1.8183	8.183	1	1.7470	7.4705	0.9125
0.2	0.8	1.822721	8.272722	0.989565	1.8352	8.352	1	1.733	7.739	0.8837
0.087	0.33	1.437375	4.973749	0.587415	1.6333	6.633	0.633	1.6335	6.3355	-1
1	0	1	1	1	1	1	1	1	1	1
0.4	0.6	1.670807	6.708069	0.863471	1.7727	7.727	0.863	1.777	7.769	-1
0.030	0.07	1.721213	7.212129	0.909169	1.7333	7.933	1	1.7333	7.933	0.963
0.05	0.75	1.736646	7.664594	0.969123	1.8183	8.183	1	1.7470	7.4705	0.9125
0.2	0.8	1.822721	8.272722	0.989565	1.8352	8.352	1	1.733	7.739	0.8837
0.087	0.33	1.437375	4.973749	0.587415	1.6333	6.633	0.633	1.6335	6.3355	-1
0.571	0.43	1.546904	5.469391	0.677679	1.7211	7.211	0.688	1.8112	8.1121	-1
0.45	0.55	1.649233	6.492331	0.849491	1.7693	7.693	0.849	1.7769	7.769	-1
DEA効率値		0.641	0.651	0.644	0.652	0.9	0.644	0.652	0.644	0.652

[LP 解法]

最大化問題(15),(16)は変数に $u_1 \geq 0, u_2 \geq 0$ と制約を課せば、入力データは全て $1(X=\{1\})$, 出力データを $Y=\{y_j\}$ のかわりに $Y^{(q)}=\{y_j^{(q)}\}$ とした、1 入力 2 出力 CCR モデルとして、通常の DEA ソフト(DEA ソルバー・LP 解法)で解くことができ、その結果を表 4(a),(b)に示す。表 2 よりも表 4(b)のデータが常に大きい(か等しい)が、両データは非常に近似している。

表 4(a): $y_1^{(q)}-y_2^{(q)}$ 平面での LP 解法 DEA 効率値 (例 3.1)

学生		A	B	C	D	E	F	G
DEA	q=1	1	1	1	1	0.9	0.641	0.6512
効率値	q=2	1	1	1	1	0.81	0.4013	0.4181
	q=0.5	1	1	1	1	0.9487	0.8031	0.8083

表 4(b): y_1-y_2 平面での LP 解法 DEA 効率 (例 3.1)

学生		A	B	C	D	E	F	G
DEA	q=1	1	1	1	1	0.9	0.641	0.6512
効率値	q=2	1	1	1	1	0.9	0.6335	0.6466
	q=0.5	1	1	1	1	0.9	0.6449	0.6533

[$y_1^{(q)}-y_2^{(q)}$ 平面での図解法]

本例は出力データ $Y^{(q)}=\{y_j^{(q)}\}$ を持つ 1 入力 2 出力 CCR モデルと見直すことができるので、 $q=2, 0.5$ の場合について区分直

線的なフロンティア面を図 1(a),(b)に示す。この図にもとづき、各学生の原点からのユークリッド距離の比率に基づく効率値を計算した。計算結果は、表 4(a)のデータと全て一致した。

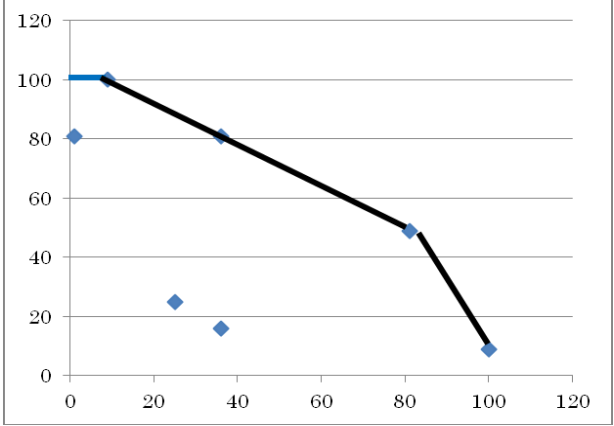


図 1(a): $y_1^{(q)}-y_2^{(q)}$ 平面での図解法($q=2$) (例 3.1)

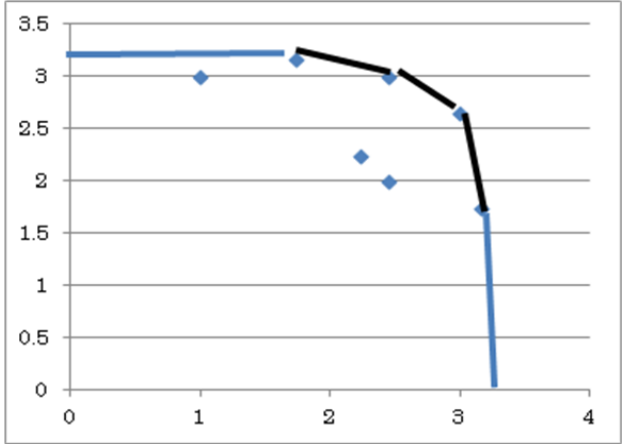


図 1 (b): $y_1^{(q)}-y_2^{(q)}$ 平面での図解法($q=0.5$) (例 3.1)

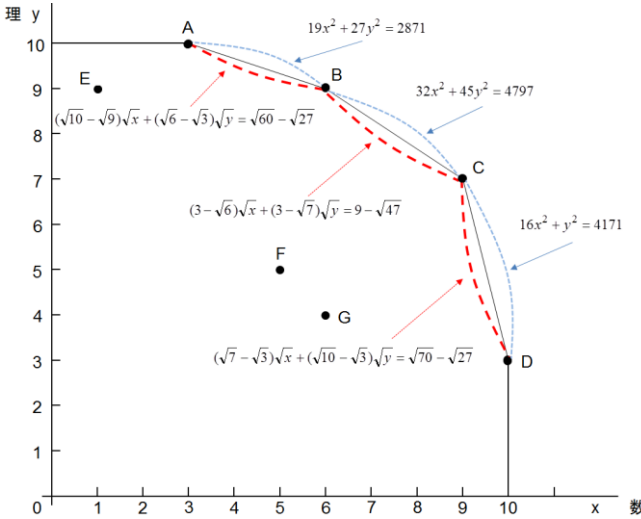


図 2: y_1-y_2 平面での図解法($q=1$ [黒実線], 2 [赤破線], 0.5 [青点線])
[y_1-y_2 平面での図解法]

図 1 においては、変換後のデータ($y_1^{(q)}, y_2^{(q)}$)平面上で区分線形なフロンティアを描写したが、 $q=1, 2, 0.5$ の場合について、変換前の原データ(y_1, y_2)平面上でのフロンティアを図 2 に示

した。これにより、図2において非線形フロンティアが実現することがわかる(なお、図的には、そうは見えないが、 $q=2$ [青点線]の各非線形区分フロンティアは全 DMU を包絡している)。各学生の原点からのユークリッド距離の比率にもとづく効率値を計算した結果、表4(b)のデータと全て一致した。

3.2 8 DMU 2 入力 2 出力 CCR モデルの例 ($p=q$)

医師数 x_1 、看護師数 x_2 を入力、外来患者数 y_1 、入院患者数 y_2 を出力する 2 入力 2 出力データ (表 5) について、 $p(=q)=0.5$, 1, 2 の場合について、効率性評価問題(17),(18)を検討する。

$$\text{目的関数: } R_k(u_1, u_2) = \frac{u_1 y_{1k}^p + u_2 y_{2k}^p}{v_1 x_{1k}^p + v_2 x_{2k}^p} \rightarrow \text{最大化(17)}$$

$$\max_j \left\{ \frac{u_1 y_{1j}^p + u_2 y_{2j}^p}{v_1 x_{1j}^p + v_2 x_{2j}^p} \right\}$$

$$\text{変数: } u=(u_1, u_2), v=(v_1, v_2) \quad (18)$$

[離散評点 DEA・表計算解法]

u_1, u_2, v_1, v_2 を $\{1, 2, 3, 4\}$ と離散評点の組み合わせで表現し、最大化問題(17),(18)を解いた結果を表6に示す。

[LP 解法]

最大化問題(17),(18)は変数に $u \geq 0, v \geq 0$ と制約を課せば、入力データ $X=\{x_{ij}\}$ 、出力データ $Y=\{y_{ri}\}$ のかわりに $X^{(p)}=\{x_{ij}^p\}$, $Y^{(p)}=\{y_{ri}^p\}$ とした 2 入力 2 出力 CCR モデルとして、通常の DEA ソフト (DEA ソルバー・LP 解法) で解くことができ、その結果を表7(a),(b)に示す。表6よりも表7(b)のデータが常に大きい(か等しい)が、両データは非常に近似している。

4. おわりに

一般化平均 DEA に対する LP 解法を検討し、その妥当性を示した。一般のデータ変換をとともなう DEA に対する LP 解法、NLP 解法の検討が今後の課題である。

参考文献

[1] 富山広也:DEAの幾何学、日本大学生産工学研究科平成23年度修士論文 (2012. 3)

付録1：成績点数のデータ変換法

例えば難易度が高い100点満点のテスト等において、点数分布を全体的に高得点に変換する方法として「 $y=10\sqrt{x}$ 変換」が知られている。ただし、 x は点数原データ (0~100)、 y は変換後データ (0~100) である。この変換では、 $x=0$ 点は $y=0$ 点へ、 $x=100$ 点(満点)は $y=100$ 点(満点)へ、さらに、例えば、 $x=36$ は $y=60$ へ、 $x=64$ は $y=80$ へと、上限と下限は固定されて、それ以外の点数がなめらかに増加する変換が実現できる。この「 $y=10\sqrt{x}$ 変換」を、 $\sqrt{x}$ を x^p とし、さらに上下限を考慮した変換法として一般化し、以下に提案する。

最も一般的には、(A1) の変換写像であるが、(A2) の「 x^p プラス定数項」型に限定する。

$$y=f(x) \quad (A1) \quad y=ax^p+b \quad (A2)$$

(A2)において a と b が未知パラメタで、 p は所与パラメタである。

x の上下限値を x_U, x_L 、 y の上下限値を y_U, y_L とするならば、

(A3), (A4)が成立する必要がある、従って、 a, b は(A5), (A6)となる。

$$y_U=ax_U^p+b \quad (A3) \quad y_L=ax_L^p+b \quad (A4)$$

$$a=\frac{y_U-y_L}{x_U^p-x_L^p} \quad (A5) \quad b=\frac{x_U^p y_L - x_L^p y_U}{x_U^p - x_L^p} \quad (A6)$$

[例A1] $p=0.5$, $x_L=y_L=0$, $x_U=y_U=100$ とすると、 $a=10, b=0$ で(A7)の

「 $y=10\sqrt{x}$ 変換」を得る。

$$y=10\sqrt{x} \quad (A7)$$

[例A2] $p=2$, $x_L=y_L=0$, $x_U=y_U=100$ とすると、 $a=0.01, b=0$ で(A8)を得る。

$$y=0.01x^2 \quad (A8)$$

[例A3] $p=1$, $x_L=40, x_U=80, y_L=0, y_U=100$ とすると、 $a=5/2, b=-100$

で(A9)を得る。

$$y=\frac{5}{2}x-100 \quad (A9)$$

[例A4] $p=0.5$, $x_L=40, x_U=80, y_L=0, y_U=100$ とすると、 $a=40/11, b=60$

で(A10)を得る。

$$y=\frac{40}{11}\sqrt{x}+60 \quad (A10)$$

[付録1 終わり]

表5：「医師と看護婦人数・外来患者と入院患者延数」データ（例3.2）

病院	A	B	C	D	E	F	G	H
医師 x_1	20	19	25	27	22	55	33	31
看護師 x_2	151	131	160	168	158	255	235	206
外来患者 y_1	100	150	160	180	94	230	220	152
入院患者 y_2	90	50	55	72	66	90	88	80

表6：離散評点DEAを適用した計算結果（例3.2）

病院		A	B	C	D	E	F	G	H
DEA 効率値	$p=1$	1	1	0.88	1	0.7587	0.8134	0.8915	0.7907
	$p=2$	1	1	0.881	0.9965	0.7475	0.8296	0.8733	0.7783
	$p=0.5$	1	1	0.881	1	0.7685	0.8347	0.9044	0.8009

表7 (a)： $X^{(p)}=\{x_{ij}^{(p)}\}-Y^{(p)}=\{y_{rj}^{(p)}\}$ 空間でのLP解法DEA効率値（例3.2）

病院		A	B	C	D	E	F	G	H
DEA 効率値	$p=1$	1	1	0.883	1	0.7634	0.8347	0.9019	0.7964
	$p=2$	1	1	0.778	0.9974	0.5634	0.6946	0.7875	0.6113
	$p=0.5$	1	1	0.94	1	0.8776	0.9137	0.9539	0.8963

表7 (b)： $X-Y$ 空間でのLP解法DEA効率値（例3.2）

病院		A	B	C	D	E	F	G	H
DEA 効率値	$p=1$	1	1	0.883	1	0.7634	0.8347	0.9019	0.7963
	$p=2$	1	1	0.882	0.9987	0.7506	0.8334	0.8874	0.7818
	$p=0.5$	1	1	0.883	1	0.7703	0.8349	0.9099	0.8034