

# DEAのモデルをめぐる——再論——

刀根 薫

## はじめに

DEAはデータオリエンテッドの効率性測定手法であり、入力と出力の間の生産関数について最小の仮定しか設けていない、いわゆるノンパラメトリックな手法である。Charnes, Cooper and Rhodes ([3])による最初の研究に始まり、その後急激に研究と応用が展開され Seiford ([8])によれば論文の数も1994年までに1000編に達しようとしている。著者は前論文「DEAのモデルをめぐる」([11])において諸モデルとその特徴について述べたが、本稿においては、モデルとデータとの適合性を中心に、DEA適用上の諸問題について考究する。併せて加法モデルにおける効率性を提案する。

## 1. DEAの基本モデル

### 1.1 CCRモデル

$n$ コの事業体(DMU)に関する $m$ コの入力データ $X \in R^{m \times n}$ と $s$ コの実出力データ $Y \in R^{s \times n}$ をもとに事業体 $DMU_o$  ( $o = 1, \dots, n$ )の効率性を測定するCCR (Charnes, Cooper and Rhodes)モデル([3])は次のように定式化される。

$$\begin{aligned} \min \quad & \theta \\ \text{st.} \quad & \theta x_o - X\lambda - s_x = 0 \\ & Y\lambda - s_y = y_o \\ & \lambda \geq 0, s_x \geq 0, s_y \geq 0, \end{aligned} \quad (1)$$

ここに、 $\lambda \in R^n$ ,  $s_x \in R^m$ ,  $s_y \in R^s$ ,  $\theta \in R$ は変数である。CCRは最初 $X > 0$ ,  $Y > 0$ を仮定した。もともと、このLPは多入力、多出力に関する比率尺

度問題から導出されたものである([3], [10])。上のLPを解くためにはまず、 $\theta$ を最小化し、次にスラックの和 $e^T s_x + e^T s_y$  (ただし $e^T = (1, \dots, 1)$ )を最大化するという2段階法を用いる。そうして得られた最適解 $(\theta^*, \lambda^*, s_x^*, s_y^*)$ とすると、 $DMU_o$ は

$$\theta^* = 1, s_x^* = 0, s_y^* = 0$$

を満たすとき、CCR効率的といい、それ以外るときCCR非効率的と呼ぶ。このLPは $DMU_o$ の現在の出力 $y_o$ を最低限保証した上で、入力を出来るだけ縮小する計画を求めている。その意味で、入力指向型モデルという。この最適解による解の改善案は次式によってなされる。

$$\begin{aligned} x_o^* &= \theta^* x_o - s_x^* \\ y_o^* &= y_o + s_y^*. \end{aligned} \quad (2)$$

このモデルでは次の生産可能集合を仮定している。

$$P = \{(x, y) | x \geq X\lambda, y \leq Y\lambda, \lambda \geq 0\}. \quad (3)$$

効率的フロンティアは断片的に平面であり、原点を通る。効率的フロンティア上の点の規模の収益性は一定(constant returns to scale = CRS)である。

上の入力指向型モデルに対して、次の出力指向型のCCRモデル(以下CCROと呼ぶ)がある。

$$\begin{aligned} \max \quad & \eta \\ \text{st.} \quad & X\lambda + s_x = x_o \\ & \eta y_o - Y\lambda + s_y = 0 \\ & \lambda \geq 0, s_x \geq 0, s_y \geq 0. \end{aligned} \quad (4)$$

明かに、最適解では $\eta^* \geq 1$ である。CCRとCCROの最適解の間には簡単な関係がある([10])。

### 1.2 BCCモデル

BCCモデルはCCRモデルに対して、次の制約を追加したものである。

とね かおる 埼玉大学大学院政策科学研究科

〒338 浦和市内大久保255

受付：95.8.23 採択：95.10.23

$$e^T \lambda = 1. \quad (5)$$

このことにより、生産可能集合は次のように、現存するDMUの凸包とそれより大きい入力と小さい出力をもつ点から構成される。

$$P = \{(x, y) | x \geq X\lambda, y \leq Y\lambda, e^T \lambda = 1, \lambda \geq 0\} \quad (6)$$

効率的フロンティアは普通原点を通らず、その上の点は規模の収益が増加型、一定型、減少型のいずれかに属する ([2], [10])。出力指向型 BCC モデル (BCCO) は、CCRO モデルに  $e^T \lambda = 1$  を追加したもので、生産可能集合は BCC モデルと同じである。

以上、生産可能集合の図表示については [10], [11] を参照されたい。

### 1.3 加法モデル

基本的加法モデルは次の LP によって示される。

$$\begin{aligned} \max \quad & e^T s_x + e^T s_y \quad (7) \\ \text{st.} \quad & X\lambda + s_x = x_o \\ & Y\lambda - s_y = y_o \\ & e^T \lambda = 1 \\ & \lambda \geq 0, s_x \geq 0, s_y \geq 0. \end{aligned}$$

このモデルの生産可能集合と効率的フロンティアは BCC モデルのそれと同一である。この LP の目的関数は当該の DMU  $(x_o, y_o)$  から効率的フロンティアへの最大  $l_1$ -ノルムの点を求める。最適解において  $s_x^* = 0, s_y^* = 0$  かつそのときのみ効率的である。加法モデルによる効率化は  $x_o^* = x_o - s_x^*, y_o^* = y_o + s_y^*$  である。このモデルにおいては入出力データの符号は自由である。

### 1.4 乗法モデル

Cobb-Douglas 型の生産関数

$$y = e^{c_0} x_1^{c_1} \dots x_m^{c_m} \quad \text{または}$$

$$\log y = c_0 + c_1 \log x_1 + \dots + c_m \log x_m$$

に対する DEA モデルとして、原データの対数を取って

$$\hat{X} = (\log x_{ij}) \in R^{m \times n}$$

$$\hat{Y} = (\log y_{ij}) \in R^{s \times n}$$

とし、 $\hat{X}, \hat{Y}$  をデータとして加法モデルを適用する。このモデルの効率的フロンティアは断片的に piecewise log-linear であるが、原データ空間では (上に) 凸ですらない。

## 2. データについて

Miller ([5]) によれば、一般にデータは次の 4 種類に分類される。

- (1) 比率尺度データ: これは原点からの距離の比で決まるデータであり、原点の位置が決定的に重要な役割をもつ。単位の取り方とは無関係である。
- (2) 間隔尺度データ: データ間の差だけが問題となる尺度。たとえば温度のように原点をどこに置くかは本質的に意味がない。
- (3) 順位尺度データ: 順位だけが問題であり、順位間の距離等は問題としない。
- (4) カテゴリカルデータ: どのカテゴリに属するかどうかだけが問題であり、カテゴリ間の優劣等は問題としない。

もともと DEA は比率尺度データをもとに展開されたが、最近では間隔尺度等のデータに対しても適用されるようになってきた。したがって、原点の位置と無関係なモデルが要求されるようになってきている。

## 3. 座標の平行移動による影響

間隔尺度を取り扱ったり、比率尺度でも負のデータがある場合には、それに適合したモデルを用いる必要がある。CCR, CCRO モデルは共に、効率的フロンティアが座標軸の原点を通るので、原点を移動させれば  $\theta^*$  の変化をもたらす。すなわち、データの原点移動 (平行移動) に対して不変ではない。それに対して、BCC モデルでは  $x$  軸の 0 点 ( $y$  軸) が  $\theta^*$  の測定の基準となり、 $y$  軸方向の平行移動に対しては不変 (invariant) である。このことは  $e^T \lambda = 1$  という制約 (凸条件) がもたらす (詳しくは Ali 等 ([1]) 参照)。すなわち、出力データをどのように底上げ (下げ) しても最適解は不変である。BCC モデルは正、ゼロ、負の出力データを処理することができる。全く同じ理由から、加法モデルは入力、出力の両方の原点移動に対して不変である。すなわち、正、ゼロ、負の入力、出力データを対象とすることができる (この性質を translation invariance という)。乗法モデルは不変ではない。

## 4. 単位の変更による影響

CCR モデルにおいて、データ  $X, Y$  を正の対角行列  $P \in R^{m \times m}, Q \in R^{s \times s}$  をもとに、 $PX, QY$  という単位の変換をしても (1) 式の最適解は不変である。ただし、第 2 段階の LP で  $\max e^T s_x + e^T s_y$  の代わりに、

$$\max \sum_{i=1}^m \frac{s_{xi}}{|x_{Mi}|} + \sum_{i=1}^s \frac{s_{yi}}{|y_{Mi}|} \quad (8)$$

ただし、

$$|x_{Mi}| = \max\{|x_{ij}| : j = 1, \dots, n\} \quad (i = 1, \dots, m)$$

$$|y_{Mi}| = \max\{|y_{ij}| : j = 1, \dots, n\} \quad (i = 1, \dots, s)$$

としておけば、こちらも単位の取り方にたいして不変である。

同様の工夫を施せば、CCRO, BCC, BCCOの各モデルはデータの単位の変更の影響を受けない(この性質を units invariance という)。

ただし、基本的な加法モデルの最適解はデータの単位の影響を受ける。しかし、加法モデルの目的関数を(8)のようにすれば、units invariance になる(translation invariance ではなくなるが)。乗法モデルは不変である。

Pastor ([7]) は(8)の代わりに次の式を用いることを提案している。この式は unit invariant かつ translation invariant である。

$$\max \sum_{i=1}^m \frac{s_{xi}}{\sigma_{xi}} + \sum_{i=1}^s \frac{s_{yi}}{\sigma_{yi}} \quad (9)$$

ただし、 $\sigma_{xi}$  ( $\sigma_{yi}$ ) は入力(出力)項目  $x_i$  ( $y_i$ ) の標本標準偏差である。 $\sigma_{xi}$  ( $\sigma_{yi}$ ) が0に近い場合には注意を要する。

## 5. モデルとデータの適合性

以上の考察を基に、表1に、基本的なDEAモデルとデータの適合性その他の重要な関連を記した。この表の中で、Semi-positive とあるのは各DMUにつき当該のデータが半正(非負でゼロでない)であることを意味し、Free は正、ゼロ、負を許す。例えば、CCRモデルでは、 $\lambda$ の非負性と入力データ  $X$  が半正であるために、 $\theta^*$  は  $0 \leq \theta^* \leq 1$  になることが保証される。

表1: モデルの特徴

モデル		CCR	CCRO	BCC	BCCO	加法	乗法
データ	X	Semi-positive	Semi-positive	Semi-positive	Free	Free	Positive
	Y	Free	Free	Free	Semi-positive	Free	Positive
平行移動	X	No	No	No	Yes	Yes	No
	Y	No	No	Yes	No	Yes	No
単位の変更		Yes	Yes	Yes	Yes	No	Yes
$\theta^*, \eta^*$		$0 \leq \theta^* \leq 1$	$1 \leq \eta^*$	$0 < \theta^* \leq 1$	$1 \leq \eta^* < \infty$	7節参照	7節参照
規模の収益		CRS	CRS	VRS	VRS	VRS	VRS

このように、CCRモデルのデータ領域は、従来の正領域から拡張することができるのである。「規模の収益性」の中のCRSは一定(constant returns to scale)であることを、VRSは可変(variable returns to scale)であることを示す。

## 6. モデルの選択をめぐって

DEAの適用に当たってモデルの選択は最も重要な問題である。モデルの選択によって効率値が異なるからである。領域限定法、コーンレリショ法、制御不能変数、カテゴリ変数等の採用も現実的に必要であるが、ここでは基本的なモデル選定に当たって考慮すべきことをあげる。

### 6.1 生産可能集合の形状

CCRとCCROモデルでは、効率的フロンティアにおいては規模の収益性が一定(constant returns to scale=CRS)であると仮定している。それに対してBCC, BCCO, 加法, 乗法モデルでは規模の収益性が増加(increasing=IRS), 一定(constant=CRS), 減少(decreasing=DRS)でありえるという仮定を用いている。入力と出力の対応に関する予備的調査、例えば回帰分析、Cobb-Douglas型分析、エキスパートの知識等によって生産可能集合の形状が推定されるならば、それに最も適合したDEAモデルを採用する。ただし、普通の回帰型の分析では1出力対多入力の関係を調べるのに対して、DEAは多出力対多入力の対応の問題にしていることに注意したい。

### 6.2 入力指向か出力指向か

DEAの一つの目的は非効率的な事業体を効率的フロンティアに投影することである。そのために、入力指向型—現状の出力レベルを最低限保証しながら入力をできるだけ縮小する—と、出力指向型—現状の入

力レベルでできるだけ出力を増加させる一がある。さらに、加法型や乗数型のように、入力と出力の不足を同時に対象とするモデルもある。

### 6.3 座標系に依存するかしないか

表1から見られるように、効率測定値  $\theta^*$  ( $\eta^*$ ) をもつかもたないかで、モデルを分類することができる。効率値  $\theta^*$  はデータのもっている座標系によって決まることに先ず注意したい。例えば、CCRモデルでは効率のフロンティアは断片的に線形な多様体からなり、それらはすべて原点を通る。したがって、原点の位置は効率値  $\theta^*$  を決める上で決定的な役割を持つ。BCCモデルの効率のフロンティアは座標系とは独立であるが、効率値  $\theta^*$  は入力データ ( $X$ ) の原点に依存する。同じように、効率値  $\theta^*$  を採用するモデルは本質的に座標系に依存している。確かに、このタイプのモデルは効率値  $\theta^*$  の値の順に事業体を並べることができるという利点をもつ。しかしながら、効率値  $\theta^*$  は必ずしも完全な指標ではない、そこには入力の過剰や出力の不足が反映されていないからである。これらの要素は管理上重要な意味をもつものである。(この点に関する解決策としては ([9]) を参照。)

他方、 $\theta^*$  をもたない加法モデルやその展開としての乗法モデルは本質的に座標系から独立している。そのため、座標系の平行移動(乗法モデルでは単位変換)に対して不変である。これらのモデルでは事業体の効率性を効率のフロンティアまでの  $l_1$ -ノルムの最大値によって測定するからである。これらのモデルは効率のフロンティアへの射影としての改善案を提示するが、 $\theta^*$  のように、事業体を1次元にランクづける便利な指標をもたない。このことが、歴史的な由来は別に、CCR, BCC といった  $\theta^*$  をもつモデルが多用される一つの大きな理由である。しかし後で述べるように新しい1次元的な効率性が提案されている。

### 6.4 データに基づくモデルの選択

表1にモデルとデータの適合性が示した。この表から加法モデルが最も広いデータ域をもっていることが分かる。更に、加法モデルは、どの入力(出力)項目をも出力(入力)項目に変換できるという柔軟性をもっていることが次のようにして分かる。例えば、出力1が小さいほど好ましいという性質のものとする。(消防署の効率性を評価するさいに出力に出火件数を用いる場合等。) このようなとき、 $y_{1j}$  の符号を変え

て  $-y_{1j}$  ( $j = 1, \dots, n$ ) とし、対応する制約式を

$$\sum_{j=1}^n (-y_{1j}) \lambda_j - s_{y1} = -y_{10}$$

とする。これより、

$$\sum_{j=1}^n y_{1j} \lambda_j + s_{y1} = y_{10}$$

となり、この制約は入力制約の一部と見なすことができる。

### 6.5 入出力項目の選択

一般的に言えば、事業体の数 ( $n$ ) に対して、入力数 ( $m$ ) と出力数 ( $s$ ) の和 ( $m + s$ ) が少ないとき、多くのDMUが効率性となって、効率性の判別ができなくなる。そのようなときは、領域限定法やコーンレリション法等を用いて効率性のDMUを限定することもできるが、 $n$  は  $m + s$  より数倍であることが望ましい。入力と出力の選択はDEAを成功させるための重要な要素である。この点に関しては、次のような手順をすすめたい。最初は、重要と判断される比較的少数の入出力項目を用いて分析する。次に、次第に項目を増やしていき、追加項目の影響を確かめながら項目の増減を行い、結論に達する。

### 6.6 多くのモデルを試みよう

予備的調査によっても、入力対出力の対応関係が確認できない場合に、ただ一つのモデルに固執することは危険である。そのようなときには、できるだけ多くのモデルを試み、その結果を比較検討し、エキスパートの意見を参考にして結論を出すことが望ましい。

## 7. 加法モデルの展開

既に述べたように、加法モデルは広いデータ領域 (translation invariance) と規模の収益性が可変であるという意味で、適用性の広いモデルである。このモデルの展開について述べる。

### 7.1 重みつき加法モデル

(7)の目的関数の代わりに、重み  $w_x \geq 0$ ,  $w_y \geq 0$  をもとに次の目的関数を設定する。

$$\max w_x^T s_x + w_y^T s_y \tag{10}$$

仮に、 $w_y = 0$  とすれば、入力指向型加法モデル(第1目的関数  $\max w_x^T s_x$ , 第2目的関数  $\max e^T s_y$ )

となり、ある意味で BCC モデルに対応する。なお、 $w_x, w_y$  の中に (8) の重みを組み込んでおけば、units invariance である。

## 7.2 制御不能変数の処理

もしある入力(出力)項目が制御不能であれば、その項目に対するスラック  $s_{xi} (s_{yi})$  の値を初めからゼロに設定しておけばよい。

同様に、上下限つきの変数も自由に設定することができる。

## 7.3 効率性尺度の導入

加法モデルには 1 次元の効率性尺度が欠けているという難点があるが、これを解決するためにいくつかの方法が提案されている。

(1) 著者は [9] の展開として次の 2 つの尺度を提案する。

$$\theta_1^* = \left[ \left( 1 - \frac{e^T s_x^*}{e^T (x_o - x)} \right) \left( 1 - \frac{e^T s_y^*}{e^T (\bar{y} - y_o)} \right) \right]^{1/2}$$

$$\theta_2^* = 1 - \left( \frac{e^T s_x^*}{e^T (x_o - x)} + \frac{e^T s_y^*}{e^T (\bar{y} - y_o)} \right) / 2$$

ただし、

$x = \{(x_i) : x_i = \min(x_{ij} : j = 1, \dots, n), i = 1, \dots, m\}$ ,  $\bar{y} = \{(\bar{y}_i) : \bar{y}_i = \max(y_{ij} : j = 1, \dots, n), i = 1, \dots, s\}$  である。なお、 $e^T (x_o - x) = 0$ ,  $e^T (\bar{y} - y_o) = 0$  のときは、対応する分数項はゼロとする。 $DMU_o$  が効率的のときのみ  $\theta = 1$  となり、 $0 \leq \theta \leq 1$  であり、この尺度も translation invariant である。また、[7] のような工夫をすれば unit invariance となる。

(2) Pastor ([6]) は次の尺度を提案している。

$$\theta^* = 1 - \left( \sum_i \frac{s_{xi}^*}{x_{io} - x_i} + \sum_i \frac{s_{yi}^*}{\bar{y}_i - y_{io}} \right) / (m + s)$$

ただし、総和は  $s_{xi}^* > 0, s_{yi}^* > 0$  の項のみについて取る。この尺度も  $DMU_o$  が効率的のときのみ 1 となり、translation invariant かつ unit invariant である。

乗法モデルに対しても同様な 1 次元尺度を定義することができる。このような尺度の妥当性と現実的な意義付けは今後の研究課題である。

## 参考文献

- [1] Ali, I.A. and L.M. Seiford (1990), "Translation Invariance in Data Envelopment Analysis," *Operations Research Letters*, **9**, 403-406.
- [2] Banker, R.D. and R.M. Thrall (1992), "Estimation of Returns-to Scale using Data Envelopment Analysis," *European Journal of Operational Research*, **62**, 74-84.
- [3] Charnes, A., W.W. Cooper and E. Rhodes (1978), "Measuring the Efficiency of Decision Making Units," *European Journal of Operational Research*, **3**, 429-444.
- [4] Cooper, W.W. and K. Tone (1995), "A Survey of Some Recent Developments in Data Envelopment Analysis," 14th European Conference on Operational Research, Semi-Plenary Papers ed. by R. Slowinski, 149-168.
- [5] Miller, D.C. *Handbook of Research Design and Social Measurement*, New York, Longman, 1983.
- [6] Pastor, J.T. (1994), "New Additive DEA Models for Handling Zero and Negative Data," Departamento de Estadística e Investigación Operativa, Universidad de Alicante, Alicante, Spain.
- [7] Pastor, J.T. (1995), "Improving the New DEA Efficiency Measure of Tone," *ibid*.
- [8] Seiford, L.M. (1994), "A Bibliography of Data Envelopment Analysis," Technical Report, Department of Industrial Engineering and Operations Research, University of Massachusetts.
- [9] Tone, K. (1993), "An  $\varepsilon$ -free DEA and a New Measure of Efficiency," *Journal of the Operations Research Society of Japan*, **36**, 167-174.
- [10] 刀根 薫「経営効率性の測定と改善」日科技連, 1993.
- [11] 刀根 薫 (1993), "DEA のモデルをめぐって," オペレーションズ・リサーチ, **38**, 34-40.