

不確実性による非効率性を考慮した確率的DEAモデル

02602144 神戸大学 * 道田 英雄 MICHIDA Hideo
01604524 神戸大学 森田 浩 MORITA Hiroshi
01501824 神戸大学 藤井 進 FUJII Susumu

1. はじめに

多入力多出力システムの効率性評価の手法であるDEA (Data Envelopment Analysis) では従来、確定したデータセットに対する評価が行われてきたが、データそのものに存在する不確実性を考慮した評価を行うことができない。なぜなら、一回だけの測定ではデータの不確実性が得られないためである。不確実性を考慮した一例として、複数回のサンプリングから得られたデータに基づいた評価をする方法が考えられる。

本研究では複数回のサンプリングからのデータによる評価の方法として、確率的計画問題のリコース問題をDEAに取り入れた、DEAのリコースモデルを用いる。このモデルからは、効率値とリコース値という2つの評価値が得られる。これまでは、リコース値を定性的な特徴から、データのばらつき具合の指標として用いてきたが、リコース値の定量的な分析により、真の効率値がとらえる範囲およびその信頼性について考察する。

2. リコースモデルの定式化と解法

本研究でDEAの評価モデルとして用いるCCRモデルの双対問題を以下に示す。

$$\text{CCR} \left\{ \begin{array}{l} \text{Minimize} \quad \theta - \epsilon(s_x + s_y) \\ \text{subject to} \quad \theta X_o - X\lambda - s_x = 0 \\ \quad \quad \quad Y_o - Y\lambda + s_y = 0 \\ \quad \quad \quad \theta \geq 0, \lambda \geq 0, s_x \geq 0, s_y \geq 0 \end{array} \right.$$

入出力データ (X, Y) が不確実性を含む場合、CCRモデルでは効率性評価ができないため確率計画法のリコース問題にあてはめる。リコースとは制約式を満たさない量に対応するものであり、リコース値 $Q(\theta, \lambda, s_x, s_y, \omega)$ は

$$Q(\theta, \lambda, s_x, s_y, \omega) = \min_{z_1, z_2 \geq 0} \{q'_x(w)z_1 + q'_y(w)z_2\}$$

$$q'_x(w)z_1 = -\theta X_o(w) + X(w)\lambda + s_x(w)$$

$$q'_y(w)z_2 = Y_o(w) - Y(w)\lambda + s_y(w)$$

で表現され、その期待値 $Q(\theta, \lambda, s_x, s_y)$ を目的関数に組み込むことで、リコースモデルの定式化ができる。こ

で θ は $0 \leq \theta \leq 1$ の変数であるため、入出力の値に応じて決まる r なるウェイトにおいて、リコース値が大きくなりすぎないように正規化している。

$$\text{REC} \left\{ \begin{array}{l} \text{Minimize} \quad \theta + rQ(\theta, \lambda, s_x, s_y) - \epsilon(s_x + s_y) \\ \text{subject to} \quad \theta \geq 0, \lambda \geq 0, s_x \geq 0, s_y \geq 0 \end{array} \right.$$

このモデルにより効率性評価を行うと、効率値以外にデータのばらつきの指標を表すリコース値が得られる。観測値が $(X^l, Y^l), l = 1, \dots, L$ という離散的な値をとり、それぞれが実現する確率を p_l とすると

$$Q(\theta, \lambda, s_x, s_y) = \sum_{l=1}^L p_l Q(\theta, \lambda, s_x, s_y, \omega^l)$$

として目的関数に組み込むことができる。リコース問題の確率変数が離散的である場合、 L 型法 [1] と呼ばれる効率的な解法で解くことができる。 L 型法では、2種類のカット (feasibility cut と optimality cut) を順次加えていく切除平面法であり、有限回の繰り返しの後、最適解が得られる。

3. 不確実性を取り入れた効率性評価

データのばらつきを考慮した評価を行うためには、複数回のサンプリングから得られたデータをそのまま利用するのが有効である。このことは、リコースモデルを用いることで実現でき、評価値として効率値とリコース値が得られる。このとき、測定データは不確実性を含んでいるため、必ずしも得られた効率値が真の効率値である保証はない。そこで効率値に加えて、もう1つの評価値であるリコース値を利用することで、真の効率値が得られるかというシミュレーションを2通りの方法で行なう。なお、以下ではリコース値 Q を m 入力について $r_{x1}, r_{x2}, \dots, r_{xm}$ 、 s 出力について $r_{y1}, r_{y2}, \dots, r_{ys}$ のように分解して扱うことにする。

(1) リコース値を効率値の相当量に変換する方法

各入出力項目ごとに分解されたリコース値が、各項目でどれだけの比重を持っているかを考える。 L 回の測定で入力1について \bar{X}_1 なる期待値とリコースモデルから r_{x1} なるリコース値が得られたとき、 r_{x1}/\bar{X}_1 が入力1に

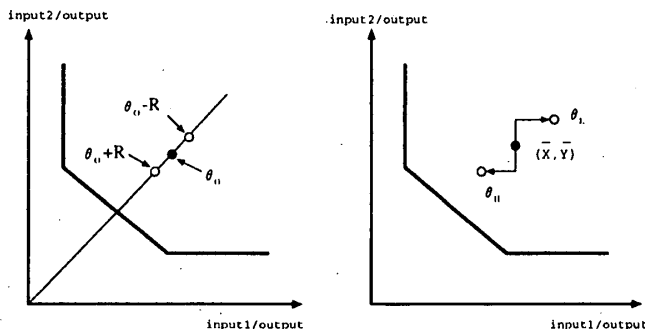


図1,(a)方法(1)の概念図 (b)方法(2)の概念図

おける不確実性の比率を示している。各入出力に対して、この比率の二乗和の平方 R を考える。すなわち

$$R = \sqrt{\left(\frac{r_{x1}}{\bar{X}_1}\right)^2 + \dots + \left(\frac{r_{xm}}{\bar{X}_m}\right)^2 + \left(\frac{r_{y1}}{\bar{Y}_1}\right)^2 + \dots + \left(\frac{r_{ys}}{\bar{Y}}\right)^2}$$

とすると、図1(a)のように観測点 θ_0 をはさんで真の効率値 θ^* の予測区間は次のようになる。

$$\theta_0 - R \leq \theta^* \leq \theta_0 + R \quad (A)$$

(2) 分解されたりコース値を各項目に加減する方法
 (1)の方法では比率から直接効率値の変動の範囲を決定したが、図1(b)のように L 回の測定で得られた入出力の期待値 (\bar{X}_0, \bar{Y}_0) に分解されたりコース値 (r_x, r_y) をそれぞれ加減することで入出力値の範囲を規定し、そこで真の効率値 θ^* の上限 θ_H と下限 θ_L を測定することで

$$\theta_L \leq \theta^* \leq \theta_H \quad (B)$$

と与えられる。

4. 数値例

2入力1出力の5つのDMU群を考える。ここではあらかじめ真の値が分かっているDMU群に、誤差の大きさの異なる乱数を付加した繰り返し測定を行ない、上記の2方法でデータ処理した場合の検証的実験を行なう。表1に5つのDMUの入出力値と効率値を示す。

各DMUの入出力に対して、標準偏差が入出力値の1, 3, 5%をもつ正規乱数を発生させ、3回の繰り返しデータを生成する。リコースモデルで効率性評価を行ったところ、例えばDMU2について、ある3回の繰り返し測定に対する評価値が $\theta_0 = 0.760$ 、 $Q = 1.334$ と得られた。 Q を各項目に分解すると、 $r_{x1} = 0.436$ 、 $r_{x2} = 0.414$ 、 $r_y = 0.484$ と得られ、上記の2つの方法に適用すると、方法(1)では $0.742 \leq \theta^* \leq 0.778$ 、方法(2)では

表1, 入出力値と効率値

DMU	input1	input2	output	θ^*
1	71	60	80	1.000
2	56	35	42	0.766
3	32	15	28	1.000
4	40	25	35	0.893
5	26	16	23	0.909

$0.760 \leq \theta^* \leq 0.774$ という予測区間が得られる。

各誤差の大きさに対してそれぞれ100回の実験をリコースモデルを用いて行い、効率値とリコース値を求めた。そして上記の2つの方法により真の効率値 θ^* の予測区間を算出し、 θ^* が予測区間に入っている割合を求めた結果を表2に示す。

表2, 条件(A), (B)を満たしている割合

DMU	Method(1)			Method(2)		
	1%	3%	5%	1%	3%	5%
誤差の大きさ						
1	1	1	1	1	1	1
2	0.88	0.90	0.87	0.94	0.92	0.93
3	1	1	1	1	1	1
4	0.89	0.89	0.85	0.97	0.98	0.96
5	0.89	0.88	0.87	0.95	0.97	0.95

方法(1)、(2)のどちらを用いてもDMU1,3は常に条件(A)、(B)を満たしている。つまり、リコースモデルで効率的と判定されたDMUにはリコース値が0となるため等号成立するからである。一方、リコースモデルで非効率的と判定されたDMUを見ると、条件(A)より条件(B)を満足する割合の方が高くなっている。つまりリコース値を単純に距離換算するよりも、入出力がとることのできる変化量と考えて効率値の範囲を限定した方が信頼性の高い評価ができることがわかる。

参考文献

- [1] J. R. Birge, and F. Louveaux (1997), "Introduction to stochastic programming", Springer.
- [2] 刀根 薫, "経営効率性の測定と改善-包絡分析法DEAによる-", 日科技連, 1993.
- [3] 森田, 道田, 藤井, "リコース付きDEAによる効率性評価", 日本OR学会秋季発表会アブスト集, 1998.