

## DEA における変数選択について

01001600 成蹊大学 \*上田 徹 UEDA Tohru

## 1. まえがき

野球の打撃成績を評価したい場合に、まず、評価項目として考えられるのは3冠王で代表されるホームラン数、打点、打率であろう。しかし、DEA の入力として試合数をとった場合には、打数はほぼ試合数に比例するため、打率は安打数を打数で割った後、それと同様の試合数でまた割るということになる。また、後半のほんの少しの打席でヒットを打ったことで高打率となった選手に高い評価を下すことは躊躇われるであろう。

しかし、打率はまだましで、例えば、打率の高い打者は一般に試合数も多くなり、試合数とある程度の単調増加関係を想定することもできるが、企業評価では、そのような単調増加関係どころか規模と単調減少となるような項目もあり得る。ここでは、鉄道会社を例として、項目選定に正準相関分析がどのように寄与し得るかを示す。

## 2. 鉄道会社の評価項目

入力を駅職員数（入力 1）、車両数（入力 2）とする。出力を輸送人数（出力 1）、運輸収入（出力 2）、平均通過数量（出力 3）、平均運賃（出力 4）とする。ただし、平均通過数量（人／日キロ）とは一日当たり、キロ当りの旅客の収容数を意味する。平均運賃（円／人キロ）とは一年間の一人当たり、キロ当りの運賃を意味する。すなわち、出力 3 と出力 4 は会社規模を表現するキロを分母に持ち、出力 4 は更に人を分母に持っている。

表 1 CCR モデルにおける効率値と乗数

DMU	効率値	DMU	効率値
東武鉄道	0.566	京阪電気鉄道	0.623
西武鉄道	0.672	阪急電鉄	0.806
京成電鉄	0.719	阪神電気鉄道	0.827
京王帝都電鉄	0.873	神戸電気鉄道	0.55
東京急行電鉄	1	山陽電気鉄道	0.566
小田急電鉄	0.852	京福電気鉄道	1
京浜急行電鉄	0.797	広島電鉄	0.304
相模鉄道	1	西日本鉄道	0.558
名古屋鉄道	0.423	JR 東日本	0.765
近畿日本鉄道	0.491	JR 東海	1
南海電気鉄道	0.516	JR 西日本	0.569

## 3. CCR モデルによる評価

節 2 の入出力データの場合の CCR モデルによる効率性評価値を表 1 に示す。この表では極度に規模の小さな京福電気鉄道の効率値が 1 となってしまう。

入力の乗数の比と出力の乗数の比に上下限を設定する領域限定法を用いても事態は改善されなかった。

## 4. 正準相関分析法の利用

文献[1]で正準相関分析法を用いた乗数制約導入法を提案し、それが入出力の削除に利用できる可能性を示唆した。そこで、本鉄道会社の例に対して、正準相関分析法の適用の可能性を検討してみる。

通常の正準相関分析では

表2 正準相関の乗数の値

	通常	非負制約
駅職員	-1.728	0
車両数	2.685	1
輸送人数	1.389	0.580
運輸収入	-0.507	0.457
平均通過数量	-0.064	0
平均輸送キロ	0.317	0
相関係数	0.993	0.987

$$\sum_{i,j} v_i S_{ij} v_j = 1, \quad \sum_{i,j} u_i T_{ij} u_j = 1$$

の条件下で第1正準相関の最大化、すなわち、

$$\max \sum_{i,j} v_i R_{ij} u_j \text{の最大化}$$

を行うこととし、解(係数、DEAでの乗数  $v_i, u_j$ )は固有値問題を解くことにより求めている。ただし、 $S_{ij}$ は入力変数  $x_i$ と  $x_j$ の相関係数、 $T_{ij}$ は出力変数  $y_i$ と  $y_j$ の相関係数、 $R_{ij}$ は入力変数  $x_i$ と出力変数  $y_j$ の相関係数である。

ここでは、DEAでの使用を考慮して通常と異なり、乗数に非負制約を課した。その結果、表2に示すように通常の正準相関分析では負の乗数が多く出現しているのに対し、非負制約を課すことでもちろん負の乗数は出ないばかりでなく、第1正準相関(正準変量間の相関係数)の低下もさほどでない。

相関係数を見ると、駅職員と車両数の相関は0.991と非常に高く、2変数を用いることの無駄さが浮き彫りになっており、一方、入力と平均輸送キロとの相関は0.378、0.311と非常に小さく、平均通過数量とでは負の相関(-0.261,-0.211)にさえなっている。

入力と出力の関係を考えると、入力を増やせば出力も増える関係を捉えるべきであり、その関

係の中で、入力の割には出力が小さいと非効率的と判断したい。また、平均通過数量、平均輸送キロについては規模と関わり大きい量で割り算しているため、基本的に規模の大きい鉄道会社には不利な項目である。このような考え方をとれば、評価する際に「駅職員、平均通過数量、平均輸送キロを削除すべき」との非負制約時の正準相関の結果は尤もであると考えられる。

車両数を入力、輸送人数・運輸収入を出力としてCCRモデルを適用した場合の効率値を表3に示す。これまで悩んできた京福の効率値も低く、CCRモデルで効率的であった東急、JR東海は効率値1を維持しており、こちらの方が無理のない評価ではないかと考えている。

すなわち、変数の選択に正準相関分析法を用いることは有効であったと考える。

[1]上田徹(1994):「包絡分析法(DEA)における乗数に関する客観的制約」、成蹊大学工学研究報告、Vol.37、No.1

表3 1入力2出力時の効率値

東武	0.566	京阪	0.525
西武	0.603	阪急	0.603
京成	0.666	阪神	0.673
京王	0.805	神戸	0.434
東急	1.000	山陽	0.499
小田急	0.800	京福	0.349
京浜急行	0.699	広島	0.282
相模	0.608	西日本	0.422
名古屋	0.401	JR東日本	0.765
近鉄	0.490	JR東海	1.000
南海	0.504	JR西日本	0.569