評価にあたってどのモデルを用いるべきか？

上田　徹

評価にあたってどの手法を用いるべきかは，得られているデータからある程度は決められるが，手法を決めてもその手法に関するモデル（アルゴリズム）が多数存在する場合がある．どのモデルを選択するかは，そのモデルを選択するさらに上位の基準がない限り困難であることを述べている．

キーワード：AHP，コンジョイント分析，包絡分析法

1. はじめに

最初，オープン・プロブレムに関する特集記事を書くように部会主査から依頼されたとき，とてもその任にはあらずと思い，お断りしたいと思ったが，オープン・プロブレムを広義に解釈して，評価しているときに何を悩んできたかをこの場ではご披露し，皆さんの共感が得られれば幸いと考え直し，執筆を引き受けた．これから書く内容はオープン・プロブレムとはそぐわないと思われる方も多いと思うが，ご容赦いただきたい．

昨年，「オペレーションズ・マネージメント」（上田[1]）という本を書いて，そこに階層的意思決定法(以下ではAHPと呼ぶ)，コンジョイント分析，包絡分析法（以下ではDEAと呼ぶ）の3手法を評価手法として紹介した．評価手法はそれらに限られているわけではないけれど，自分にとっては馴染みがあるので，それらに限定してどのモデルを用いるべきかを考察してみる．

2. 階層的意思決定法（AHP）について

2.1 欠落データがある場合

図1で示されるトーナメントの結果をAHPで解いてみる．勝った方に値2，負けたほうに0.5を与えることにすると，一対比較行列は

$$A=\begin{bmatrix} 1 & 2 & 2 & \times \\ 0.5 & 1 & \times & \times \\ 0.5 & \times & 1 & 2 \\ \times & \times & 0.5 & 1 \end{bmatrix}$$

うえだ　とおる
成蹊大学 理工学部情報科学科
〒180-8633 武蔵野市吉祥寺北町3-3-1

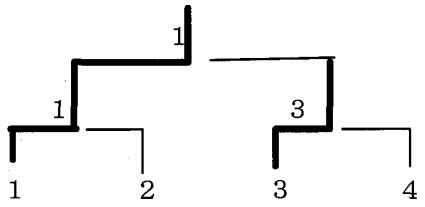

図1　トーナメントの場合

となる．

欠落値がある場合の処理法として高橋[2]ではハーカー法，二段解法，最小2乗法を取り上げている．

固有（重要度）ベクトルは要素の最大値が1になるように正規化すると，それぞれ以下のようになる．

ハーカー法：[1　0.5　0.5　0.25]

二段解法：[1　0.4422　0.6315　0.3962]

最小2乗法：[1　0.5　0.5　0.25]

この結果をみて多数決で決めればハーカー法と最小2乗法の結果を採用することになる．すなわち，直接対戦していないチーム2と3は同評価を受けていることになる．

しかし，チーム2はチーム4に負ける可能性があるのに対し，チーム3は勝っているということを考えればチーム3を上位に置く二段解法にも意味がある．

高橋[2]の二段解法(I)では行列Aの第i行の実測部のみの幾何平均をとって，重要度w_iの第1近似$\hat{w}_i$を求め，Aの欠落部a_{ij}を$\hat{w}_i/\hat{w}_j$で代用して，欠落のない行列$\hat{A}$

$$\hat{A}=\begin{bmatrix} 1 & 2 & 2 & 2.245 \\ 0.5 & 1 & 0.707 & 1 \\ 0.5 & 1.414 & 1 & 2 \\ 0.445 & 1 & 0.5 & 1 \end{bmatrix}$$

を作っている．そのため，ハーカー法の固有値が4であるのに対し，二段解法では4よりも若干大きな固有

値4.0369を取っている．

そこで対角要素以外の要素 a_{ij} をすべて $\hat{w}_i/\hat{w}_j$ で代用して，欠落のない行列 $\hat{A}$

$$\hat{A}=\begin{bmatrix} 1 & 2.245 & 1.587 & 2.245 \\ 0.445 & 1 & 0.707 & 1 \\ 0.630 & 1.414 & 1 & 1.414 \\ 0.445 & 1 & 0.707 & 1 \end{bmatrix}$$

を作った二段解法(II)（高橋[3]の例2の場合）では，固有値は4で固有ベクトルは

[1　0.4454　0.6300　0.4454]

となる．二段解法(I)ではチーム2をチーム4よりも上位としているが，二段解法(II)ではチーム2と4に差をつけていない．この結果からは二段解法(II)の手順が最もよさそうに思える．

チーム2と4の対戦がないことが評価を曖昧にしていると考え，チーム2と4の対戦結果を加えてみる．チーム2がチーム4に勝った場合は，

「チーム2と3は，チーム4に勝ち，チーム1に負けた」

という同じ結論を持ち，どの方法でも同じ結論が得られる．しかし，チーム4がチーム2に勝った場合は

ハーカー法：[1　0.3536　0.7071　0.5]

二段解法(I)：[1　0.3951　0.7218　0.5680]

二段解法(II)：[1　0.3969　0.6300　0.6300]

となり，チーム3はチーム4に勝っているにもかかわらず二段解法(II)ではチーム3と4に差をつけていない．

このように，データによってよいと思われる手法が変わるのであれば，定性的モデル評価ができにくい場合にはどの手法をとるべきかが分からない．

2.2　評価尺度

重要度の求め方として，固有ベクトルによる方法と幾何平均値による方法が知られている．

幾何平均値による方法は対数二乗誤差和を最小にする解となっている．二乗誤差和の最小化はガウスの最小二乗法に代表されるようにノンパラメトリックな場合にも用いられ，回帰分析においては最良線形不偏推定量になっており，さらに誤差が正規分布に従う場合には最尤推定量であるという性質からその価値が認められている．

しかし，最小二乗法による解と最尤推定値とのどちらが望ましいかは断定できないと考えている．

もちろん分布が分かっていれば最尤推定値を求めることができるが，AICから分かるように真の分布は通常，分かっておらず，AICを比較したときに小さいほうがよいといえるだけである．最小二乗法による解がよいと主張したいわけではなく，観測値とあまり離れていない解として

「知ることのできない最良の解の代用物として使える」

といいたいのである．

観測値とあまり離れていない解としては，絶対誤差和を最小にする解にも魅力がある．絶対誤差和を最小にする解には異常値の影響を受けにくいという利点があり，線形計画問題として定式化できる場合も多い．二つの誤差和の大雑把な特徴は表1のようになる．

統計学をかじった人にはパラメトリックな推定法を信奉している人も多いように思うが，経済現象や主観的な判断値のどこに分布を特定化できる余地があるのだろうか？　あくまでも近似として既知の分布（確率密度関数）を使用できるだけであると考えている．

AHPでは個々人の評価ばかりでなく回答者全体の評価もできる．回答者 k が下したモノ i と j の一対比較値 $a_{ij}(k)$ が存在するときに，回答者全体の幾何平均値

$$b_{ij}=n\sqrt{\prod_{k=1}^{n} a_{ij}(k)}$$

をモノ i と j の一対比較行列の第 (i,j) 要素として分析することが推奨されている．

そのことを対数二乗誤差和最小の観点から正当化する論文（静岡大学の八巻教授提案の擬似整合度についても議論しており，上田，八巻教授の連名論文）を投稿したが，「対数最小二乗推定がなぜよいのか説明されていない」を主理由として採択されなかった（論文の主な内容については上田[4]に示されている．動機が不純だといわれそうだが，この原稿を引き受けた理由の一つが，「不採択に不満でどこかで述べておきたい」ということであった）．この査読者は

「最適な評価尺度が存在する」

あるいは

「評価尺度を評価する尺度が存在する」

とか

「パラメトリックなモデル以外は認められない」

表1　誤差和の特徴

二乗誤差和	絶対誤差和
解が陽に求まる	解が数値的に求まる
異常値の影響を受けやすい	異常値の影響を受けにくい

と考えているのだろうか？

私は，ノンパラメトリックなモデルでは，

① モデルと観測値との何らかの距離が小さいモデルがよいモデルであり，

② その距離の定義あるいは評価尺度を変えると異なる解が得られるのが普通であり，

③ どの距離定義を用いたらよいか，どの解を用いるべきかは別の上位の評価基準を設けない限り不可能である．

といいたいのである．

例えばAICはそのような上位の評価尺度になり得ると考えられるが，

「AIC自身はKullback-Leibler情報量に依存している」

といったように，何か規準を設ければその規準に基づいて最適なものを選べるが，規準自身のよさを示すにはまた別の規準が必要になるのは必然である．

このことは数学における公理を考えてみればよいだろう．公理から出発して一つの理論体系が出来上がるが，公理自身の妥当性を証明することは無意味であろう．

3. コンジョイント分析について

上田[1]の第9章に様々なコンジョイント分析のモデルを紹介した．

コンジョイント分析を最初に実用的なものとしたMONANOVA（上田[1]の節9.2.1参照）

を初めとして多くのモデルは制約を付けられない．

大きさに応じて効用が単調に変化する場合などのように非負制約を付けたい場合は多い．非負制約のような線形制約は，線形計画法を用いるLINMAPであれば容易に対処できる．

LINMAPの定式化例を以下に示す．ただし，サービスは選好順位の低いものから順に並べられているものとする．

$$\text{目的：minimize} \quad S=\sum_{i=1}^{m-1}\gamma_i$$

$$\text{制約：}(\boldsymbol{x}_{i+1}-\boldsymbol{x}_i)^tY+\gamma_i\geq 0\,;\ i=1,2,\cdots,m-1$$

$$(\boldsymbol{x}_m-\boldsymbol{x}_1)^tY=m-1$$

$$Y\geq\mathbf{0},\ \gamma_i\geq 0\,;\ i=1,2,\cdots,m-1$$

ここで，$\boldsymbol{x}_i$はサービスiの属性値で既知の値を持ち，Yは各属性に掛かる重みで未知の（求めたい）値であり，$\boldsymbol{x}_i^tY$がサービスiの効用推定値である．制約の第1式は効用推定値が選好順位と合致しないときにγ_iを加えて選好順位と合致させるわけであるが，そのようなγ_iは小さいほど望ましいわけで，目的関数はγ_iの最小化となる．$S=0$の場合には効用推定値と選好順位とは矛盾しておらず，MONANOVAやTRADE-OFFの解にもなり得る．

しかし，この定式化では$(\boldsymbol{x}_{i+1}-\boldsymbol{x}_i)^tY$が非負であればよいので，異なるサービスが同じ効用値を持っても構わないことになってしまう．そこで，隣り合う選好順位の効用値の差をなるべく1にする規準を追加することを考えた．また，三角型ファジィ数による変数増加によって矛盾解消をする定式化も行ってきた（上田[1]の節9.2.5，節9.4参照）．

しかし，制約追加が容易だという理由だけでLINMAPを，尤度関数を用いるRANKLOGITなどの他のモデルよりも上位としてよいだろうか？

AICは

$-2\times$（モデルの最大対数尤度）

$+2\times$（モデルの自由パラメタ数）

すなわちモデルと観測値との（非）適合度を第1項で測り，モデルの複雑性を第2項で測ってその和の大きさでモデル間の比較を行っている．

モデルと観測値との差が正規分布に従っているときには第1項は最小二乗推定されたパラメタを用いる場合と一致するので，選考順位と効用値の一致度を表す（対数尤度に限定しない）関数とモデルの自由パラメタ数の関数との和で新しい評価尺度を作り，そのもとでモデル間の比較を行うことが考えられる．しかし，その場合にはなぜその評価尺度が適切かを示すKullback-Leibler情報量のような基準が必要であろう．

予測モデルでは，誤差を減らすためにモデルを複雑にすればするほど，説明変数を増やせば増やすほど予測力が落ちる傾向があり，モデルの複雑性をペナルティとするAICの意味は定性的に理解できる．

しかし，「回答を何とか解釈したい」という観点から，上田[1]では説明変数をどんどん増やしていった．説明が付かなかったら説明変数を増やせばよいと主張したいわけではない．他の，よりよい効用定式化があるかもしれないが，残念ながらその有無が分からない．説明変数を増やしていくやり方は

「回答者の選択順位を再現できる」

ということにだけ価値を置いた場合にはそれなりの意味があろう．しかし，再現できる定式化がいくつもある場合には別の規準を設けない限り，それらの定式化を比較できない．

4. 包絡分析法（DEA）について

DEA用のソフトウェアとしてDEA-Solver[5]がある．それは文献[6]に基づいており，その付録に30個のモデルの略号が示されている．モデルはこれで尽きているわけではなく，私自身も最近は距離最小化加法モデルを提唱している[7]．自分が評価したい対象に対して何を用いるべきかは大問題である．

4.1 CCRモデルとBCCモデル

入力行列を X，出力行列を Y，着目DMU（Decision Making Unit）o の入力ベクトル $\boldsymbol{x}_o$，出力ベクトルを $\boldsymbol{y}_o$ とすると，最も基本的なCCRモデルでは次のように主問題を定式化する．

$$\text{目的：}\theta^*=\max \boldsymbol{u}^t\boldsymbol{y}_o \tag{1}$$

$$\text{制約：}\boldsymbol{v}^t\boldsymbol{x}_o=1$$

$$-\boldsymbol{v}^tX+\boldsymbol{u}^tY\leq\boldsymbol{0} \tag{2}$$

$$\boldsymbol{v}\geq\boldsymbol{0},\ \boldsymbol{u}\geq\boldsymbol{0}$$

BCCモデルでは次のように主問題を定式化する．

$$\text{目的：}\max \boldsymbol{u}^t\boldsymbol{y}_o+u_0 \tag{3}$$

$$\text{制約：}\boldsymbol{v}^t\boldsymbol{x}_o=1$$

$$-\boldsymbol{v}^tX+\boldsymbol{u}^tY+u_0\boldsymbol{e}\leq\boldsymbol{0} \tag{4}$$

$$\boldsymbol{v}\geq\boldsymbol{0},\ \boldsymbol{u}\geq\boldsymbol{0}$$

u_0：符号制約なし

（$\boldsymbol{e}$ はすべての要素が1の列ベクトル）

すなわちCCRモデルとBCCモデルの違いは定数項 u_0 を入れているかどうかだけである．

しかし，効率値には大きな差が出る場合がある．総資産，従業員数を入力，売上高，経常利益を出力として企業の経営効率性を評価した場合であったが，CCRモデルの効率値は0.638であったのに対し，BCCモデルの効率値は1という企業Aがあった．なお，乗数の値の比較のために入出力は観測値/（最大値−最小値）という正規化を実施している．

この企業Aは入出力がすべて最大値を採る企業で，どちらのモデルでも従業員数，経常利益は無視することが自分に有利ということになっている．その結果，CCRモデルでは売上高は少ないが，売上高/総資産の大きな企業Bと比べて非常に劣ることになる．一方，BCCモデルでは絶対値の大きな負の定数 u_0 をとることにより企業Bの仮想的出力（$\boldsymbol{u}^t\boldsymbol{y}_B+u_0$）も負の値にすることができて，企業Aの効率値は1にできた．

u_0 が負の場合には

「大きな規模を達成するためには小さな企業を維持するよりも余分な費用が掛かるため，それを考慮してやらなければならない」

といった理屈を付けられるかもしれない．

しかし，仮想的出力を負にするほどの u_0 を取ることはおかしいと考えてすべての企業 i に（$\boldsymbol{u}^t\boldsymbol{y}_i+u_0\geq 0$）の条件をBCCモデルに追加すると企業Aの効率値は0.752にしかならなかった．

また，企業Aの効率性検討では従業員数，経常利益は無視することとなったが，無視しないように乗数に関する制約を課す領域限定法のようなモデルを用いると，さらに異なる効率値が得られる．

このようにDEAのもととなる線形計画法では線形制約の追加がいくらでもでき，そのためにいくらでも新しいモデルが生み出され得る．

4.2 加法モデル

BCCモデルの双対問題は次のように定式化される．

$$\text{目的：}\min \theta \tag{5}$$

$$\text{制約：}\theta\boldsymbol{x}_o-X\boldsymbol{\lambda}\geq\boldsymbol{0},$$

$$Y\boldsymbol{\lambda}\geq\boldsymbol{y}_o \tag{6}$$

$$\boldsymbol{e}^t\boldsymbol{\lambda}=\boldsymbol{1},\ \boldsymbol{\lambda}\geq\boldsymbol{0}$$

この問題を第1段階として解く．制約(6)の第1式，第2式にスラック変数 $\boldsymbol{s}_x$，$\boldsymbol{s}_y$ を追加すると

$$\theta\boldsymbol{x}_o-X\boldsymbol{\lambda}=\boldsymbol{s}_x,\ Y\boldsymbol{\lambda}-\boldsymbol{y}_o=\boldsymbol{s}_y$$

となり，第1段階から得られた最適な θ を θ^* に固定して $(\boldsymbol{s}_x^t, \boldsymbol{s}_y^t)\boldsymbol{e}$ を最大化する問題を第2段階として解く2段階方式を採っている．

ここで，$\boldsymbol{s}_x$ は θ を介しているために $(\boldsymbol{s}_x^t, \boldsymbol{s}_y^t)\boldsymbol{e}$ は当該DMU$(\boldsymbol{x}_o^t, \boldsymbol{y}_o^t)$ と目標とする（参照）点（$\boldsymbol{\lambda}^tX^t$, $\boldsymbol{\lambda}^tY^t$）との L_1 距離（各座標の差の絶対値の和）になっていない．

そこで，加法モデルは第1段階を経ずに直接 $(\boldsymbol{s}_x^t, \boldsymbol{s}_y^t)\boldsymbol{e}$ の最大化を目的としている．すなわち，BCCモデルの第2段階と加法モデルでは当該DMUからできるだけ遠い参照点を探していることになる．

BCCモデルは第1段階を経ているため極端な参照点を求めることにはなりにくいが，加法モデルでは当該DMUにとって改善目標とするには遠すぎる点を参照点としてしまう欠点がある．

それならば

「効率的フロンティア上のなるべく近い点を求める」

方が改善を勧告する場合にも説得力を持つであろう．ここで，図2から分かるように非効率的DMU D からできるだけ遠い効率的フロンテイア上の点は効率的DMUの線形結合の中から求められるが，最も近い点

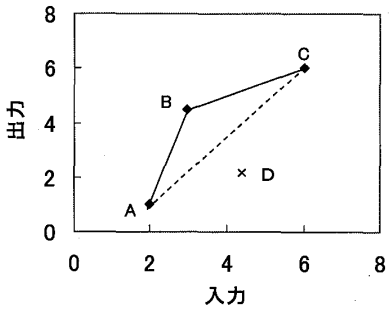

図2　効率的フロンティア A-B-C

を効率的DMUの線形結合の中から求めると線分AC上の点を求めることになってしまう．そこでなるべく近い点をもとめるには線分AB上とBC上とに分けて考えるなど上田[7]に示すような手順が必要である．

距離最大の点を求めるのか，距離最小の点を求めるのかに関しては「見習いやすい」という別の規準を持ち込んで距離最小の点を求めることを正当化した．しかし，「目標はなるべく高く掲げるべきだ」として距離最大の点を求めるべきだとの考え方もあるかもしれない．

5.　むすび

どんなモデルを使うべきかは，それを決めるべき上位の規準がない限り決められないと考えている．決められたモデルはあくまでもその上位規準のもとでのみ正当化されるものであり，その上位規準そのものの正当性はさらに上位の規準がない限り決められない．その上位の規準は他人に対する説得性を持つ必要があり，結局，定性的なよさでしか言えないものなのかもしれない．もし定量的に言おうとすると，その定量的尺度のよさをやはり定性的なよさで表現せざるを得ないのではないだろうか？

線形計画法では制約追加が容易なので，結果を見て望ましい結果が得られない場合にはさらに制約を追加すれば望ましい結果が得られるかもしれない．しかし，「望ましい」といった途端にどういう観点から望ましいといっているのかを明らかにしなければいけないし，その観点が他への説得力を持つものでなければならないだろう．

参考文献

[1] 上田徹："オペレーションズ・マネージメント，" 牧野書店 (2006)
[2] 高橋磐郎："AHPからANPへの諸問題II，" オペレーションズ・リサーチ, Vol. 43, No. 2 (1998), 100-104
[3] 高橋磐郎："AHPからANPへの諸問題III，" オペレーションズ・リサーチ, Vol. 43, No. 3 (1998), 160-163
[4] 上田徹："AHPにおけるグループ（集団）評価について，" 2003年日本OR学会春季研究発表会, 2-F-4, (2003), 226-227
[5] DEA-Solver-Pro, SAITECH (現在, Version 6.0)
[6] W. W. Cooper, L. M. Seiford and K. Tone："Data Envelopment Analysis—A Comprehensive Text with Models, Applications, References and DEA-Solver Software，" 2 nd edition, Springer (2007)
[7] 上田徹："距離最小化加法モデルadd-minの検討，" 2007年日本OR学会春季研究発表会, 1-C-1, (2007), 32-33