

[研究論文]

R による SFA の解法とその応用 Solving the problems of SFA based upon the R-package

譚 康融[†]

Kangrong Tan[†]

[†]久留米大学 経済学部

[†] Faculty of Economics Kurume University

1. はじめに

近年、経済・経営分析の関連分野において、マクロ・ミクロ経済モデル等に関するパラメータの統計的推定や、モデルによる試算・推計などと言った総合的な数値解析、いわば、コンピューテーションナルエコノミクスが台頭しつつある。

その1つ大きな理由としては、1990年代半ば、インターネット商用化が本格的に動き出した。以来、インターネットを利用したビジネス、例えば、電子商取引(Electronic Commerce)、モバイルビジネスが急速に発展を遂げ、日進月歩の ICT (Information Communication Technology) 技術革新によって、今日はクラウドコンピューティング(Cloud Computing)は普及しつつ、リアルタイムのトランザクション(Transaction)処理が可能となり、ビジネスのリエンジニアリングが進み、より多くのデータをより迅速に収集することができ、ビッグデータの時代に突入した。

企業の経営効率分析とは、収集されたデータに基づき、企業経営に関わる様々なデータを抽出し、企業の経営効率性を評価することである。すなわち、生産性・コストコントロールの最適化(産出、利潤の最大化、あるいは費用・コストの最小化)が実現されたかどうかを測ることである。経営効率(経営パフォーマンス)への適切な評価を行うことによって、経営における不足面を改善し、経営のより効率化を図られる。

また経済・経営活動に伴うリスクファクターに対して、リスクコントロール(適切な評価、回避等)は、欠かせないものであり、PDCA(Plan, Do, Check, Action)を実施することにより、絶えずに変化している経営環境に応じて、常に効率的経営を実現させ、企業のさらなる発展を促す。

さらに今日の就職事情を考えると、文系学生であろうかどうかにかかわらず、自ら企業選択の時に、正しく企業評価を行える能力は、自分の就職、キャリア、将来の人生に大きな影響を与えることは言うまでもない。また近年、従来の日本的経営とされる終身雇用、年功序列体制からますます能力主義への移行が見られており、企業は総合能力の高い人材を求めている傾向も否めない。文系出身者であっても、基本的な数理分析の手法を教養として身に付ける必要があると考えられる。

本稿は企業経営の効率性評価において、広く利用されている SFA(Stochastic Frontier Analysis)手法について、その基本モデルを解説し、特に R のパッケージである Benchmarking を用いて、シミュレーションや、実証分析などを含めた実例を挙げながら、SFA 効率分析を行う場

合に利用されるコマンドや、プログラミング手法を中心に展開していく。本稿は以下のように構成される。第 2 節では、R の基本的な機能を紹介する。第 3 節では、SFA の数理モデルの基本を取上げる。第 4 節では、シミュレーションや、実証分析などの例題を通じて、Benchmarking パッケージを用いた経営効率性の分析、プログラミング手法の詳細を示していく。

2. R の豊富な計算機能

フリー統計処理ソフト R は、従来の統計処理専用ソフトウェア（例えば、SAS,SPSS など）に比べて、解析機能はさほど変わらない一方、導入コストはほとんどかからないことが知られている。金融機関を始め、一般企業や大学等の研究機関も多く導入されている。実際、R は R-project という組織によって、開発、運営されている。その Website は www.r-project.org である。

R を導入するにあたって、まずはこのサイトにアクセスし、使用されているコンピュータに合うバージョンをダウンロードする。例えば、OS は Windows 系であれば、Windows バージョンの R を、Mac の iOS であれば、Mac 用のバージョンをそれぞれダウンロードする。UNIX 系のバージョンさえも用意されてある。また、ダウンロードの際に、各国の Mirror Site が設けられていて、ユーザは各自のニーズに応じて、R パッケージをサーバからダウンロードしてインストールすることができる。

R のコンソール画面で、昔風の DOS プロンプトの感覚で、コマンドを入力し、すぐ計算できる。R は簡単な計算機能から積分、微分などの高度な計算機能まで備えている。以下の例題を R のカーソル画面で確かめれば分かる。

[例題] 2000 万円を 1%の金利で 30 年間預金した場合の金額

```
> 2000*(1+0.01)^30
```

```
[1]2695.698
```

[例題] $\sin(x)$ が区間 $[0,2*3.14]$ 上での積分

```
>integrate(sin, 0,2*3.1415)
```

```
1.716938e-08 with absolute error <4.4e-14
```

[例題] 区間 $[-1,1]$ 上での関数 $f(x) = x^3$ の積分

```
>f<-function(x) x^3
```

```
>integrate(f,-1,1)
```

```
0 with absolute error < 5.6e-15
```

[例題] 正規分布の確率密度関数の積分

```
>integrate(dnorm,-Inf,Inf)
```

```
1 with absolute error < 9.4e-05
```

また、正規乱数を用いたシミュレーションは経済・経営の分野でもよく利用されている。以

下の `morm(N,μ,σ)` コマンドで簡単に正規乱数を生成できることは有難い。標準正規分布に従う乱数を生成する場合は、コマンドの中の平均、標準偏差を省略し、`morm(乱数の個数)` だけ書けば良い。本文の後半で、この正規乱数を用いた SFA 解析のシミュレーションを示す予定である。

```
>y<-morm(100);      #標準正規分布に従う乱数 100 個を生成  
  
>y  
[1]-0.26471635    0.06836203    1.04438384    0.73242955-0.03757372  
省略  
[97] -0.21872425    1.30495947    0.21969349    0.51006633  
  
>plot(y,type="b");  #データのプロット  
  
>hist(y);          #y のヒストグラムを描く
```

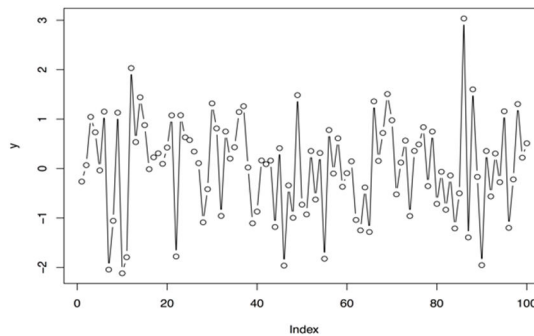


図1 正規乱数の時系列データプロット

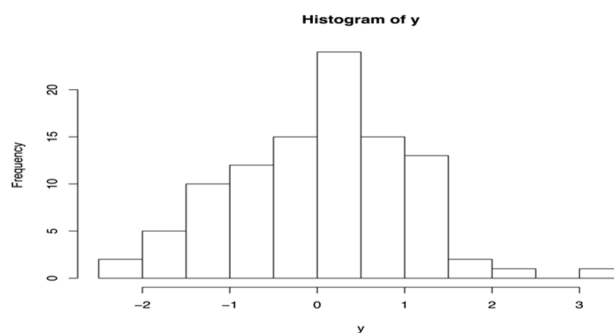


図2 正規乱数のヒストグラム

図1, 2 はそれぞれ, 生成した正規乱数のプロット及びヒストグラムを示している。

3.1 SFA 手法とは

Stochastic Frontier Analysis(SFA)の和訳は確率フロンティア分析法となる。生産者である企業の経営効率分析には、ミクロ経済理論、あるいは伝統的な経済理論分析がある[1]。

例えば、Cobb-Douglas 生産関数を用いた場合は、従来の方法では、ランダムノイズの影響はあまり考えられなかった上、多くの先行研究は、最小二乗法 (Least Squares Method)を用いられ、ノイズのタイプは平均ゼロの対称正規分布がほとんどのケースで仮定されていた。

しかしながら、現実問題として、各々の企業・生産者は必ずしも効率的ではなく、すなわち、経営資源投入の最小化、あるいは産出の最大化が実現されているとは言えない状態がある。そこで、経営効率評価を行うことによって、各企業経営の優劣がはっきり判明できるようになる。

また、経済全体、あるいはグローバル経済の景気局面はどういう状況であるか、これは企業経営に大きな影響を与える一つ無視できない要因である。一方、企業内のマネジメントや、従業員のモラルなどは、もう一つ企業経営に与える大きな要因である。SFA 手法では、企業が置かれている経営環境は対称的分布を持つ確率的要素が設定されている一方、各企業の内部経営状況について、もう一つの非対称分布を持つ確率的要素が導入される[2][3][4][5][6][7]。

DEA(Data Envelopment Analysis)手法に比べて、SFA は確率的なノイズを導入した評価方法である。確率変数を導入することによって、ノイズへの対応ができる一方、多出力ケースへの対応が難しくなり、また生産関数のタイプを予め仮定しないとイケないのは、DEA と大きく異なる点であろう[8][9]。

3.2 SFA の概略

まず、Cobb-Douglas 型の生産関数(Production Function)を思い出させてみよう。ここで、 K, L はそれぞれ生産要素の資本と労働であり、 Y は産出であり、両辺を対数取ると、

$$Y = CK^{\alpha}L^{\beta} \quad (1)$$

$$\log Y = \log C + \alpha \log K + \beta \log L \quad (2)$$

となる。もっと一般的に書くと、 $\log y_j = \beta_0 + \sum_{i=1}^m \beta_i \log x_{ij}$ と表せ、ここで X_i は i 番目の生産要素である。また、上述した2つの確率要因をそれぞれ確率変数 v と u として導入し、ここで、 $v \sim N(0, \sigma^2)$ を観測誤差項、 $u \sim N^+(0, \sigma^2)$ を非効率性とする、

$$\log Y_j = \beta_0 + \sum_{i=1}^m \beta_i \log X_{ij} + v_j - u_j \quad (3)$$

となり、 $j(=1 \dots n)$ は j 番目の企業を意味する。ここで、 $\log Y_j = y_j$, $\log X_{ij} = x_{ij}$ とする。

よって、 u と v の確率密度関数はそれぞれ

$$f(v) = \frac{1}{\sqrt{2\pi}\sigma_v} \exp\left\{-\frac{v^2}{2\sigma_v^2}\right\} \quad (4)$$

$$f(u) = \frac{2}{\sqrt{2\pi}\sigma_u} \exp\left\{-\frac{u^2}{2\sigma_u^2}\right\} \quad (5)$$

となり, ただし, $u < 0$ の時は, $f(u) = 0$ となる。それらの同時密度関数は

$$f(u, v) = \frac{1}{\pi\sigma_u\sigma_v} \exp\left\{-\frac{u^2}{2\sigma_u^2} - \frac{v^2}{2\sigma_v^2}\right\} \quad (6)$$

ここで $\xi = v - u$ とおくと,

$$f(u, \xi) = \frac{1}{\pi\sigma_u\sigma_v} \exp\left\{-\frac{u^2}{2\sigma_u^2} - \frac{(u+\xi)^2}{2\sigma_v^2}\right\} \quad (7)$$

$$f(\xi) = \int_0^\infty f(u, \xi) du = \int_0^\infty \frac{1}{\pi\sigma_u\sigma_v} \exp\left\{-\frac{u^2}{2\sigma_u^2} - \frac{(u+\xi)^2}{2\sigma_v^2}\right\} du \quad (8)$$

また

$$\lambda = 2\sigma_v^2 \frac{\sqrt{\sigma_u^2}}{\sqrt{\sigma_v^2}} \quad (9)$$

$$\sigma^2 = \sigma_u^2 + \sigma_v^2 \quad (10)$$

すると,

$$f(\xi) = \frac{\sqrt{2}}{\sigma\sqrt{\pi}} \left(1 - \Phi\left(\frac{\xi\lambda}{\sigma}\right)\right) \exp\left\{-\frac{\xi^2}{2\sigma^2}\right\} \quad (11)$$

あるいは

$$f(\xi) = \frac{\sqrt{2}}{\sigma\sqrt{\pi}} \left(\Phi\left(-\frac{\xi\lambda}{\sigma}\right)\right) \exp\left\{-\frac{\xi^2}{2\sigma^2}\right\} \quad (12)$$

故に尤度関数 $L(\beta, \lambda, \sigma^2)$ は

$$L(\beta, \lambda, \sigma^2) = \prod_{j=1}^n f(\xi_j) \quad (13)$$

となる。すなわち, 対数尤度関数は

$$\text{Log}(L(\beta, \lambda, \sigma^2)) = \sum_{j=1}^n \log(f(\xi_j)) \quad (14)$$

となる。すなわち,

$$\log(L(\beta, \lambda, \sigma^2)) = \log(f(y_1 - f(x_1; \beta), \dots, y_n - f(x_n; \beta), \sigma^2, \lambda)) \quad (15)$$

$$\log(L(\beta, \lambda, \sigma^2)) = -\frac{n}{2} \log\left(\frac{\pi}{2}\right) - \frac{n}{2} \log \sigma^2 + \sum_{j=1}^n \log\left(-\frac{\lambda(y_j - f(x_j; \beta))}{\sqrt{\sigma^2}}\right) - \frac{1}{2\sigma^2} \sum_{j=1}^n (y_j - f(x_j; \beta))^2 \quad (16)$$

パラメータを推定するために最尤法 (Maximum Likelihood) を利用する。最尤法の数値計算は, 色々なケースで, 計算手法が工夫されているが, Newton 法を用いた逐次代入法により, 尤度の最大値を求める方法が広く応用されている。

また, 生産関数はトランスログ, あるいは CES 生産関数を採用されている場合もあり, さら

に生産性よりも費用・コストを重視する場合は、費用関数(Cost Function)を考慮し、ある一定の生産量に対してその費用・コストを最小化することに設定するモデルもある。

さらに $\exp(-u)$ は非効率性を表して、 $\lim_{u \rightarrow \infty} \exp(-u) = 0$ のとき、効率値 $D=0$ 、一方、 $u=0$ のとき、 $D=1$ となる。しかし何らかの手法で u 値を推定しないと、各企業の経営効率値が得られないわけである。これまで効率性に関わる u の推定方法は

- 1) TeBC: $E[e^{-u}|\epsilon]$
- 2) TeJ: $E[u|\epsilon]$
- 3) TeMode: $e^{pmin(0, -\mu^*)}$

と言った推定方法が考案されている。すべての効率値に関する詳細の導出過程はやや煩雑の為、ここで割愛するが、知りたい読者は参考文献に参照されたい[2][3][4][5][6][7]。

4 Rパッケージの利用

フリー統計解析ソフト R 言語には、いろんなパッケージが開発され、公開されている。ここで利用するのは Benchmarking パッケージである[7]。前述した DEA 問題もこのパッケージで解析することができる。パッケージを利用する前に、まず、パッケージのインストールが必要である。次に R コンソール、あるいは R の Script を用いて、`library(Benchmarking)` というコマンドで、パッケージを読み込む。その次、解析対象となるデータセットの読み込み、そして解析コマンドを書くという流れになる。

[例題] 正規乱数を用いた SFA 法のシミュレーション解析

```
>library(Benchmarking)
>n<- 50 #企業の数の設定
>x1<- 1:50+rnorm(50,0,10) #x の値を与える
>x2<-100+ rnorm(50,0,10)
>x<- cbind(x1,x2)
>y<-0.5+1.5*x1+2*x2+rnorm(n,0,1)-pmax(0,rnorm(n,0,1)) #y の値を与える
>sfa(x,y) #コマンド sfa を用いた解析
```

Coefficients:

(Intercept)	xx1	xx2
-1.182	1.487	2.018

```
> summary(sfa(x,y))    #結果をまとめる
```

```

                Parameters  Std.err  t-value  Pr(>|t|)
(Intercept)    -1.18179    2.06179  -0.573188    0.569
xx1             1.48742    0.01085  137.032837    0.000
xx2             2.01810    0.01198  168.518940    0.000
lambda          0.01764    2.31251   0.007629    0.993
sigma2          1.1236
sigma2v = 1.12327 ; sigma2u = 0.0003495753
log likelihood = -73.85586

```

```
Convergence = 4; number of evaluations of likelihood function
```

```
19 Max value of gradien: 0.09430119
```

```
Length of last step: 0
```

```
Final maximal allowed step length:0.33075
```

```
# 各企業の効率性の推定及び表示
```

```
>o<-sfa(x,y)
```

```
>eff(o)
```

```

[1] 0.9851398 0.9852687 0.9849763 0.9852490 0.9851328 0.9852987 0.9852842
[8] 0.9851902 0.9851721 0.9854656 0.9853176 0.9852493 0.9852833 0.9851533
[15] 0.9853590 0.9852599 0.9852954 0.9851633 0.9851880 0.9850228 0.9850797
[22] 0.9853206 0.9852294 0.9853957 0.9853537 0.9854235 0.9851720 0.9854137
[29] 0.9856682 0.9853887 0.9852047 0.9852534 0.9853002 0.9852757 0.9851795
[36] 0.9852526 0.9852021 0.9851725 0.9851930 0.9851842 0.9850920 0.9853251
[43] 0.9851687 0.9853471 0.9852984 0.9852939 0.9851905 0.9852193 0.9854166
[50] 0.9852696

```

```
#異なる定義の効率値の計算 (te=teBC)
```

```
>te <-te.sfa(o) #equal to eff(o), teBC
```

```
>teBC<-teBC.sfa(o)
```

```
>teM<-teMode.sfa(o)
```

```
>teJ<-teJ.sfa(o)
```

```
>cbind(eff(o),te,teBC,Mode=eff(o,type="Mode"),teJ)[1:10,]
```

```

                eff(o)      te      teBC      Mode      teJ
[1,] 0.9851398 0.9851398 0.9851398 0.9996716 0.9850768
[2,] 0.9852687 0.9852687 0.9852687 1.0000000 0.9852065
[3,] 0.9849763 0.9849763 0.9849763 0.9992207 0.9849124

```

```
[4,] 0.9852490 0.9852490 0.9852490 0.9999765 0.9851866
[5,] 0.9851328 0.9851328 0.9851328 0.9996522 0.9850698
[6,] 0.9852987 0.9852987 0.9852987 1.0000000 0.9852366
[7,] 0.9852842 0.9852842 0.9852842 1.0000000 0.9852220
[8,] 0.9851902 0.9851902 0.9851902 0.9998120 0.9851275
[9,] 0.9851721 0.9851721 0.9851721 0.9997615 0.9851093
[10,] 0.9854656 0.9854656 0.9854656 1.0000000 0.9854045
```

```
> sigma2.sfa(o) # Estimated variances
```

```
sigma2
```

```
1.123619
```

```
> lambda.sfa(o) # Estimated lambda
```

```
lambda
```

```
0.01764121
```

```
> ymax<-max(te)
```

```
> ymax
```

```
[1] 0.9856682
```

```
> which(te==ymax)
```

```
[1] 29
```

```
> plot(te,type="b")
```

```
> abline(v=29)
```

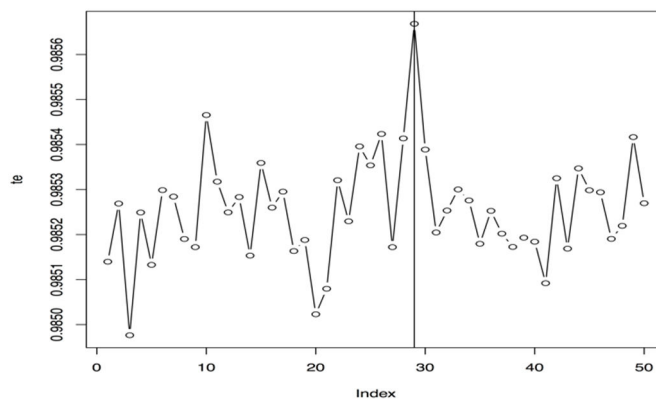


図3 50社の経営効率のプロット

図3は正規乱数を用いてシミュレートされた50社企業への経営効率評価値を示している。横軸は会社の番号で、図から29番目の企業の経営効率値が一番高いことは一目瞭然であろう。

[例題]デンマークの乳業生産者のデータセットを用いた経営効率分析

これはRdocumentation.orgが提供しているデータセットである。その内容:108ヶ所の農場観測値(農場の番号, ミルクの生産量, エネルギー支出, 獣医への支出, 乳牛の頭数)。図4には, 各変数間の関係が示されている。

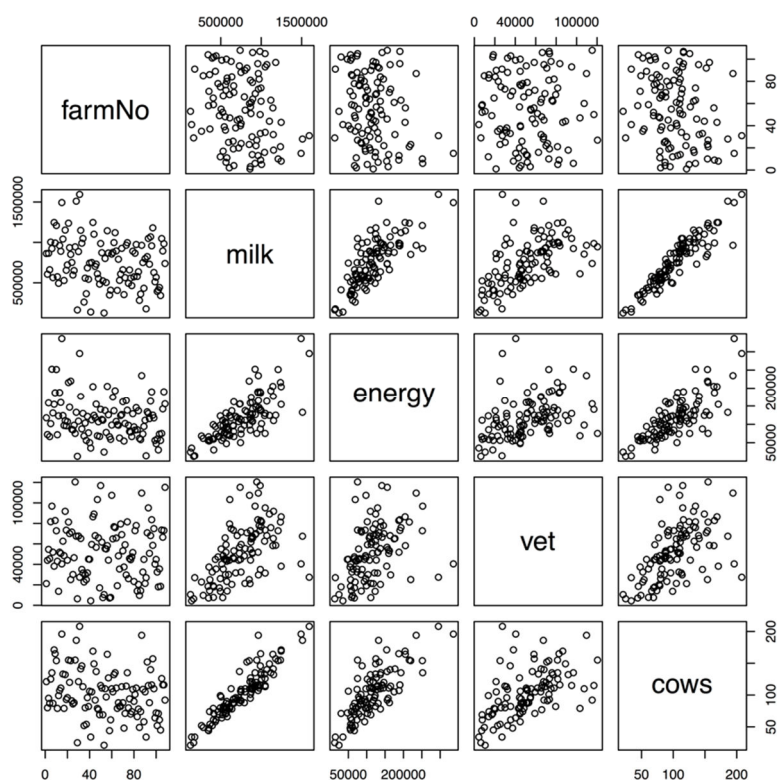


図4 各乳業農場のデータ

前例と同じく, 以下の手順で Benchmarking パッケージを利用する。

- 1) library(Benchmarking)でパッケージを読み込み。
- 2) 牛乳生産者のデータセットを読み込み。
- 3) 解析コードを書き, プログラムを実行する。
- 4) 得られた結果をまとめる。

```
>library(Benchmarking)
>milkProd <- read.csv( "milkProd2.csv")
>head(milkProd) #最初の6行のデータの表示
```

	farmNo	milk	energy	vet	cows
1	1	862533	117894	21186	121
2	2	605764	72049	43910	80
3	3	865658	158466	54583	95
4	4	662331	66783	45469	87
5	5	1003444	101714	81625	125
6	6	923512	252408	96807	135

```
> tail(milkProd) # 後ろからの6行のデータの表示
```

	farmNo	milk	energy	vet	cows
103	103	441835	73450	35536	61
104	104	342816	56077	18511	45
105	105	1047603	75114	73498	126
106	106	894828	142008	66516	116
107	107	983645	190440	73142	116
108	108	738916	156109	115209	92

```
#投入各要素をまとめて x に入れる
```

```
>x<-with(milkProd, cbind(vet, energy, cows))
```

```
>head(x)
```

```
vet energy cows
```

```
[1,]
```

```
21186 117894 121
```

```
省略
```

```
[6,]
```

```
96807 252408 135
```

```
>tail(x)
```

```
vet energy cows
```

```
[103,] 35536 73450 61
```

```
省略
```

```
[108,] 115209 156109 92
```

```
#産出値をまとめて y に入れる
```

```
>y<-matrix(milkProd$milk)
>head(y)
      [,1]
[1,] 862533
省略
[6,] 923512
>tail(y)
      [,1]
[103,] 441835
省略
[108,] 738916
>milkSfa <- sfa(log(x), log(y))
>ols <- lm(log(y)~log(x))
>cbind(ols=coef(ols), sfa=coef(milkSfa)) #OLS との比較
```

	ols	sfa
(Intercept)	7.10341187	7.52014421
log(x)vet	0.09551563	0.06281416
log(x)energy	0.12132193	0.12156101
log(x)cows	0.85907831	0.87878814

```
>milkSfa
Coefficients:
(Intercept)  xvet      xenergy  xcows
7.52014      0.06281  0.12156  0.87879

#あるいは
>sfa=coef(milkSfa)
>sfa
```

(Intercept)	xvet	xenergy	xcows
7.52014421	0.06281416	0.12156101	0.87878814

```
# teJ 効率値の計算
>e <- residuals(milkSfa)
```

```

>s2 <- sigma2.sfa(milkSfa)

>lambda <- lambda.sfa(milkSfa)
>mustar <- -e*lambda^2/(1+lambda^2)
>sstar <- lambda/(1+lambda^2)*sqrt(s2)
>teJ <- exp(-mustar-sstar*( dnorm(mustar/sstar)/pnorm(mustar/sstar) ))
>teJ

      [,1]
[1,] 0.8976377
[2,] 0.9143635
省略

[100,] 0.9545376
[101,] 0.9065084
[102,] 0.8265653
[103,] 0.8659444
[104,] 0.9272760
[105,] 0.9636482
[106,] 0.8823452
[107,] 0.9194875
[108,] 0.8563217

```

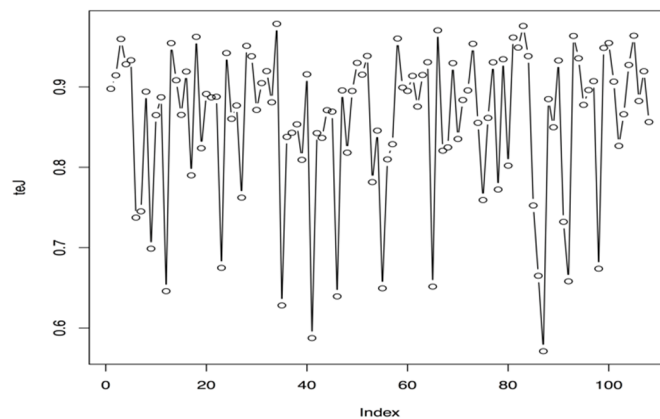


図 5 teJ 効率性

図 5 には、108 ヶ所の牛乳生産農場の teJ 効率値が描かれている。

5 SFA手法による各都道府県の生産効率評価

内閣府の統計資料により, 日本全国各都道府県の県内総生産および所得要素の表から, 47 都道府県のデータを抽出し, 各都道府県の生産効率性を推計し, SFA 手法を用いた実証分析を試みる [10].

表 1 各県の詳細データ (金額単位は 100 万円)

県番	県名	県内総生産	固定資本減耗	県内雇用者報酬
01	北海道	15,399,031	2,944,467	7,405,649
省略				
47	沖縄県	3,167,354	636,219	1,382,856

具体的には, 平成 25 年度 (2013 年度) の各都道府県内の総生産, 中間投入, 固定資本減耗, 及び雇用者報酬を集計データから抽出した。そこで, 各県の資本ストックが公表されていない為, ある程度固定資本を反映している固定資本減耗を代替的に利用した。データのフォーマットは表 1 に示されている。推計に用いた R のプログラムは以下ようになる。

```
prefecture<-read.csv("prefecture-SFA20161203.csv",header=T)
yy<-prefecture$a #県内総生産
x1<-prefecture$b #中間投入
x2<-prefecture$d #固定資本減耗
x3<-prefecture$h #県内雇用者報酬
xx<-matrix(cbind(x1,x2,x3),47,3) oo<-sfa(log(xx),log(yy))
te2<-te.sfa(oo)
```

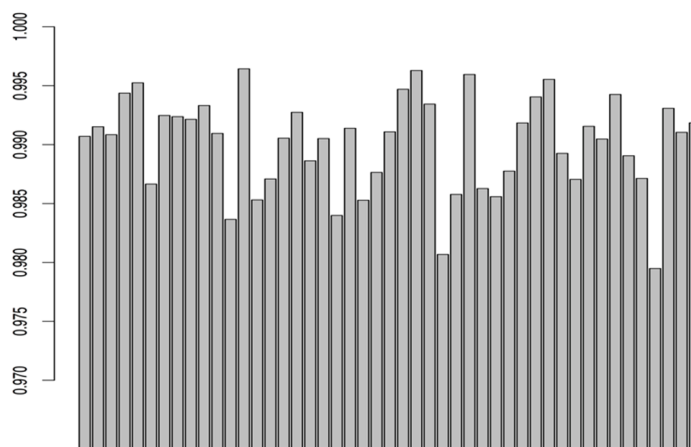


図 6 47 都道府県に関する生産効率値 (teBC)

図6は各都道府県の中で、13番目の東京都の効率値が一番高いことを示している。また各県の効率値はすべて0.97以上の為、効率値の範囲を区間[0.97,1]で拡大して見ることにより、各県の効率値の僅差は、はっきり見えてくる。さらにteMode,teJの効率性を計算して以下の結果が得られた。

```
>teBC2<-teBC.sfa(oo)
>teMode2<-teMode.sfa(oo)
>teJ2<-teJ.sfa(oo)
>plot(teMode2,type="b")
>efficiency2<-cbind(te2,teBC2,teMode2,teJ2)
>efficiency2
```

	te2	teBC2	teMode2	teJ2
[1,]	0.9907024	0.9907024	0.9986190	0.9906786
[2,]	0.9915165	0.9915165	1.0000000	0.9914954
[3,]	0.9908413	0.9908413	0.9989816	0.9908180
[4,]	0.9943678	0.9943678	1.0000000	0.9943562
[5,]	0.9952441	0.9952441	1.0000000	0.9952353
[6,]	0.9866530	0.9866530	0.9902676	0.9866170
[7,]	0.9924714	0.9924714	1.0000000	0.9924535
[8,]	0.9923737	0.9923737	1.0000000	0.9923554
[9,]	0.9921506	0.9921506	1.0000000	0.9921316
[10,]	0.9933177	0.9933177	1.0000000	0.9933027
[11,]	0.9909525	0.9909525	0.9992768	0.9909296
[12,]	0.9836590	0.9836590	0.9856834	0.9836157
[13,]	0.9964334	0.9964334	1.0000000	0.9964280
[14,]	0.9853093	0.9853093	0.9881020	0.9852699
[15,]	0.9870870	0.9870870	0.9910144	0.9870522
[16,]	0.9905457	0.9905457	0.9982186	0.9905214
[17,]	0.9927426	0.9927426	1.0000000	0.9927256
[18,]	0.9886267	0.9886267	0.9939017	0.9885964
[19,]	0.9905163	0.9905163	0.9981444	0.9904919
[20,]	0.9839892	0.9839892	0.9861496	0.9839467
[21,]	0.9913886	0.9913886	1.0000000	0.9913670
[22,]	0.9852818	0.9852818	0.9880596	0.9852423
[23,]	0.9876425	0.9876425	0.9920100	0.9876093
[24,]	0.9910865	0.9910865	0.9996392	0.9910640
[25,]	0.9946961	0.9946961	1.0000000	0.9946855
[26,]	0.9962892	0.9962892	1.0000000	0.9962834

[27,]	0.9934385	0.9934385	1.0000000	0.9934238
[28,]	0.9806816	0.9806816	0.9817855	0.9806327
[29,]	0.9857749	0.9857749	0.9888294	0.9857366
[30,]	0.9959537	0.9959537	1.0000000	0.9959470
[31,]	0.9862728	0.9862728	0.9896336	0.9862358
[32,]	0.9855864	0.9855864	0.9885323	0.9855477
[33,]	0.9877541	0.9877541	0.9922159	0.9877212
[34,]	0.9918402	0.9918402	1.0000000	0.9918202
[35,]	0.9940466	0.9940466	1.0000000	0.9940340
[36,]	0.9955363	0.9955363	1.0000000	0.9955283
[37,]	0.9892593	0.9892593	0.9952199	0.9892309
[38,]	0.9870491	0.9870491	0.9909480	0.9870142
[39,]	0.9915556	0.9915556	1.0000000	0.9915347
[40,]	0.9904764	0.9904764	0.9980442	0.9904519
[41,]	0.9942571	0.9942571	1.0000000	0.9942451
[42,]	0.9890542	0.9890542	0.9947827	0.9890252
[43,]	0.9871305	0.9871305	0.9910906	0.9870958
[44,]	0.9794906	0.9794906	0.9803446	0.9794399
[45,]	0.9930816	0.9930816	1.0000000	0.9930657
[46,]	0.9910427	0.9910427	0.9995200	0.9910201
[47,]	0.9918486	0.9918486	1.0000000	0.9918286

また以下の図 7,8,9 では、それぞれの効率値の関連を示している。各効率値は、多少異なるが、概ね一致していることが読み取れよう。

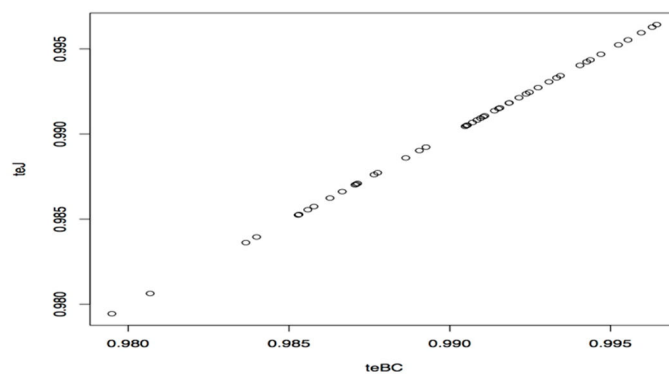


図7 teBC 効率性と teJ 効率性との比較

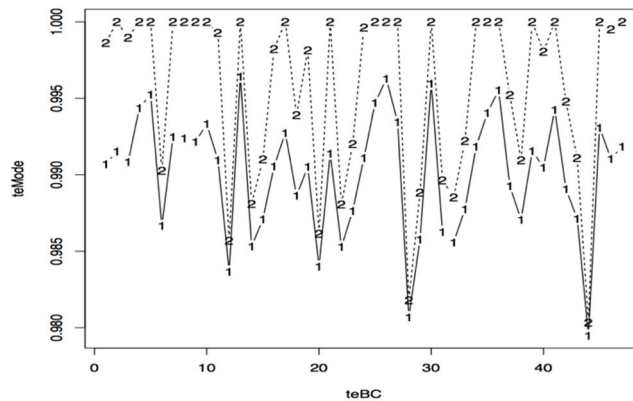


図8 teBC 効率性(1の実線)とteMode 効率性(2の点線)との比較1

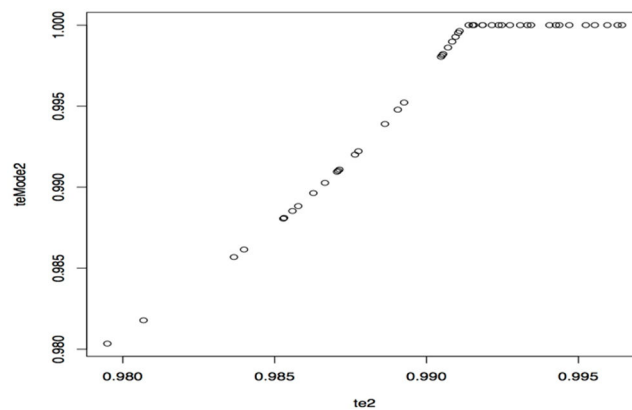


図9 teBC 効率性(1の実線)とteMode 効率性(2の点線)との比較2

6. むすび

本稿は、まず、Rの簡単な計算機能を紹介した。次にSFA(Stochastic Frontier Analysis)手法の数理モデルの導出を行い、SFA手法の基本的な考え方を示した。さらにRパッケージであるBenchmarkingを用いた解析手順、プログラミング方法について、シミュレーションを始め、いくつかの例を挙げながら、パッケージの利用法を示した。また内閣府のサイトより、実際の統計データを入手し、日本全国の47都道府県の生産効率性について、実証分析を行い、各県の生産効率性を示した。そのうえ、推計結果による定義の異なる諸効率値間の比較を行い、多少の違いがあったにもかかわらず、評価結果は概ねに一致していることが確認された。

今後の課題は、将来ビッグデータ（株価、為替レート、決算情報等）を用いた大規模企業経営効率性分析を行った場合、数値計算の迅速化を計る並列計算のアルゴリズムの開発・利用などが挙げられる。これらの課題については、今後の研究に取り込んでいく予定である。

参考文献

- [1] Farrell, M. J., The Measurement of Productive Efficiency, *Journal of the Royal Statistical Society*, 120, 253-281, 1957.
- [2] Greene, W. H., The Econometric Approach to Efficiency Analysis. In Fried, H. O., Knox Lovell, C. A., and Schmidt, P., editors, *The Measurement of Productive Efficiency*. Oxford University Press, New York and Oxford, 2008.
- [3] Aigner, D.J., Lovell, C.A.K., Schmidt, P., Formulation and Estimation of Stochastic Frontier Production Functions, *Journal of Econometrics*, 6, 21-37, 1977.
- [4] Baltas, G., Exploring Consumer Differences in Food Demand: A Stochastic Frontier Approach. *British Food Journal*, 107(9), 685-692, 2005.
- [5] Coelli, T.J., Rao, D.S.P., O'Donnell, C.J., Battese, G.E., *An Introduction to Efficiency and Productivity Analysis*, 2nd Edition, Springer, 2005.
- [6] Polachek, S. W. ; Yoon, B. J., A two-tiered Earnings Frontier Estimation of Employer and Employee Information in the Labor Market. *Review of Economics and Statistics*, 69(2), 296-302, 1987.
- [7] Bogetoft, P., Otto, L., *Benchmarking with DEA, SFA, and R*, Springer, 2000.
- [8] Banker, R. D., A., Charnes, and W. W. Cooper, Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis, *Management Science*, 1078-1092, 30(9), 1984.
- [9] 譚康融, データ包絡分析法を用いた企業評価及びRの活用, コンピュータジャーナル, Vol.,29, 2-13, 久留米大学情報教育センター, 2014
- [10] <http://www.esri.cao.go.jp>