Data Envelopment Analysis

数据包络分析的基本理论与模型

序论

- 数据包络分析法(Data Envelopment Analysis, DEA)是以“效率”的概念作为加总模式，而效率则等于总产出除以总投入，并以效率最大化为目标 。
- DEA法的一个与其他多因素决策分析模式不同之处在于DEA不需要事先假设相对权重，而是由实证资料中推导产生，每个受评方案的效率衡量都是分别采取对该受评方案最有利的权重组合，从而避免决策者的主观判斷影响权重选取。

大纲

- 数据包络分析法简介(Data Envelopment Analysis, DEA)
- DEA基本模型
- DEA 使用步骤
- DEA的特性
- 结论
- 本人的研究工作（勿外传）

绩效评估

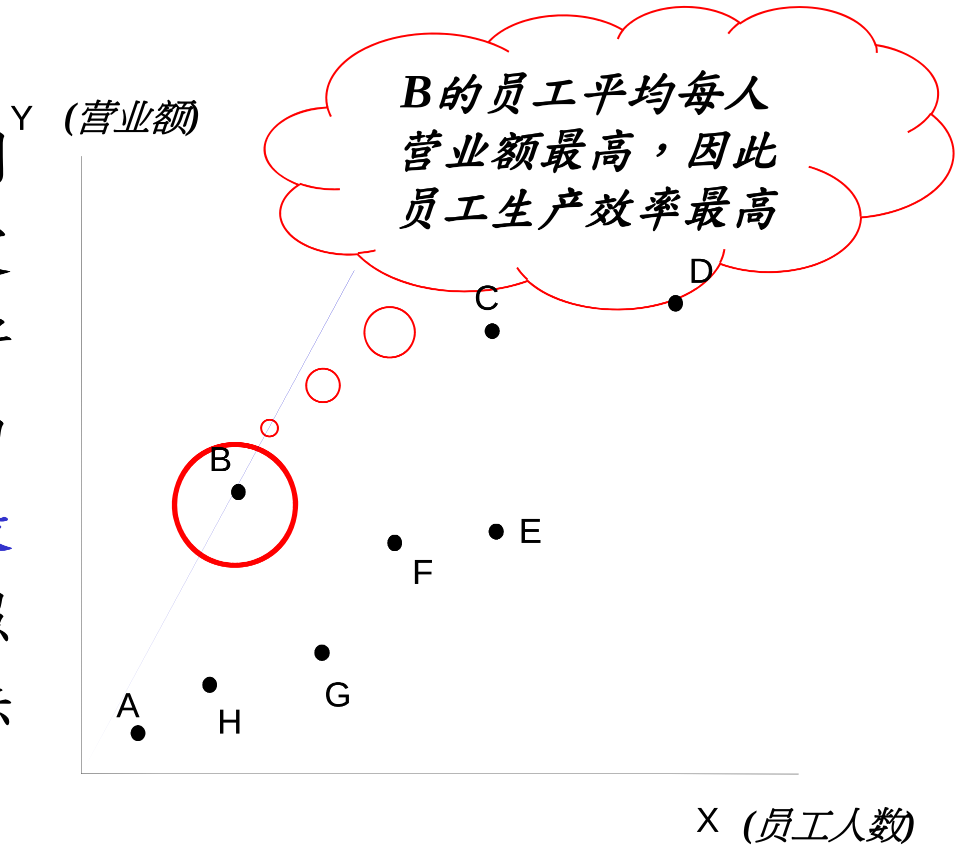
- 绩效评估(performance evaluation)是评估组织或个人如何以较少的投入资源获得较多之产出结果的多因素评估，通常使用“成本效益分析”(cost effectiveness analysis；cost benefit analysis)的“投入产出比”来同时考虑对目标分别产生负向影响的投入因子与产生正向影响的产出因子
- 评估企业的营运效率高低，可以营业額与员工人數作为评价因素，成本效益分析通常以“效率”(efficiency)的概念来分析每单位投入可产生多少产出，所以生产力可以用“员工每人营业額”的比例式表达如下

$$生产力 = \frac{营业额}{员工人数}$$

员工生产效率分析

- 生产力的目标值是越大越好，所以营业额越高越好，而员工人数越少越好

➢ 某公司分析其八個不同工厂的生产效率以决定从中选择一间绩效最好的工厂，以营业额作为产出，员工人数作为投入，可将每个工厂按照其营业额与员工人数标示于二维坐标中



数据包络分析法简介(1/2)

- 针对多个受评单位或备选方案的相对效率分析与比较，学者(Charnes et al., 1978；Banker et al., 1984)提出“数据包络分析法”(Data Envelopment Analysis, DEA)的相对比较方式。
- 将因素区分为投入项（input）与产出项（output），不预先设定权重的方式分别加总产出值和总投入值，然后总产出除以总投入的比率作为相对效率。
- DEA的应用目的为评估组织或单位的相对绩效，因此将被评估的对象称为“决策单位”(Decision Making Unit, DMU)。

数据包络分析法简介(2/2)

- 以成本效益的角度，效率等于总产出除以总投入的比率，故每一个方案的效率如式所示：

$$E = \frac{\sum_{j \in O} u_j Y_j}{\sum_{i \in I} v_i X_i}$$

  - 集合I表示结果X_i越小越好的指标
  - 集合O表示结果Y_j越大越好的指标
  - u_j与v_i分别代表集合I与集合O中每个指标对应的相对权重

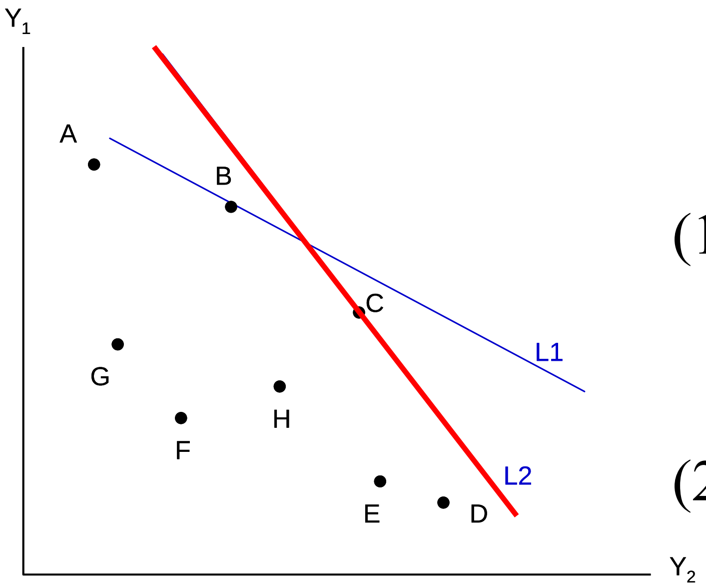
- 找出相对效率值最高的备选方案，比较不同决策单位的相对效率值，并分析效率不佳的方案应减少投入或是提高产出，提出具体的改善方向

生产边界衡量效率(1/2)

- 早期的学者以经济学观点来阐释效率，提出以生产边界(production frontier)为衡量效率的基础，主要有两种方法：
  - 参数法(parameter approach) 利用理论建构或实证推导的方式预先设定生产函数的形式。
  - 非参数法(non-parametric approach) 恰好相反，DEA法即为一种非参数的生产函数分析法：不默认投入与产出的相对权重，藉由实际投入产出的数据形成包络面(envelopment surface)，推测出生产边界

生产边界衡量效率(2/2)

- 若固定某一种生产函数关系，只有B是最佳方案，若改变生产函数关系，则只有C是最佳方案



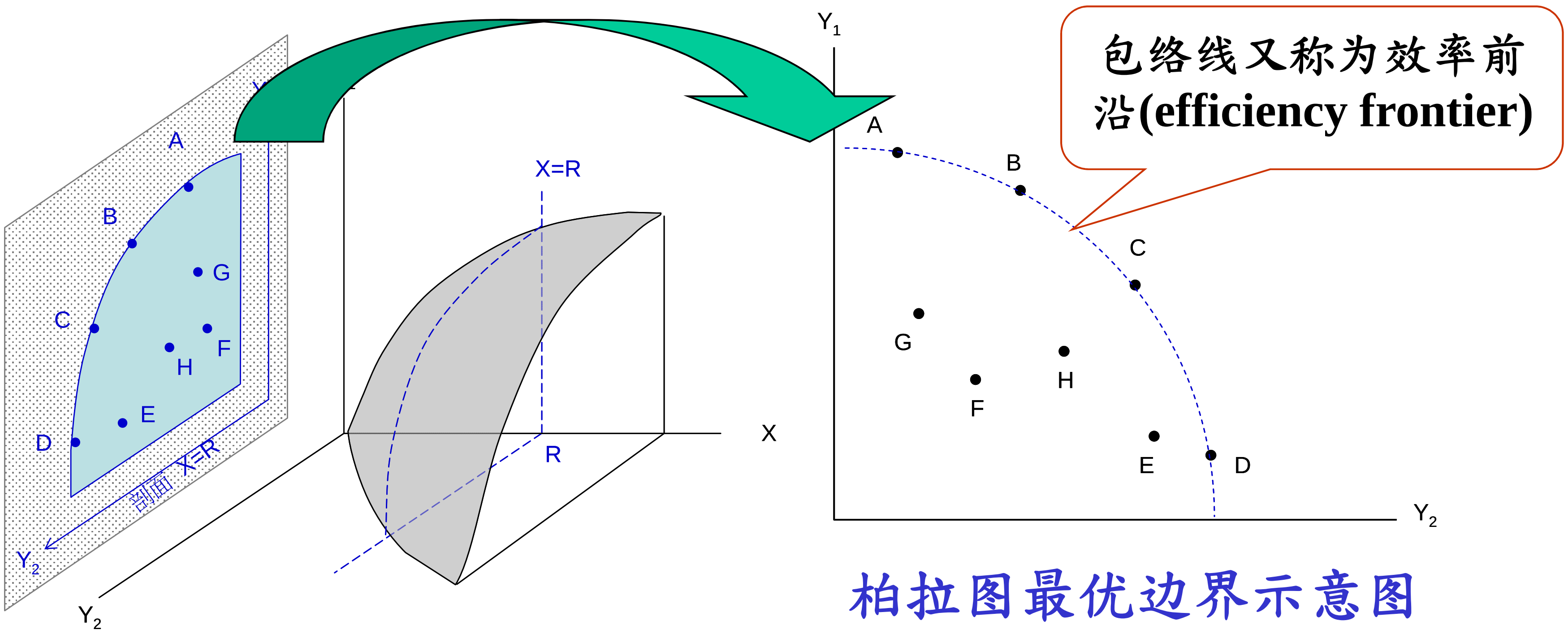
- DEA法的相对效率衡量是建立在柏拉图最优边界(Pareto optimal frontier)的效率观念上：

(1)针对某个产出项，除非增加投入资源或减少其他产出项的产量，否则该产出项之产量无法被增加

(2)针对某个投入项，除非减少产出或增加其他投入项的投入资源，否则该投入项的投入资源无法被减少

例：营运效率分析

- 以营业额和生产量作为衡量产出的产出项，以员工人数作为唯一投入项，将每个工厂按照其营业额、生产量与员工人数标示于坐标中从而构成生产曲线



柏拉图最优边界示意图

DEA基本模型

- 我们将介绍数据包络分析法中的三种主要模型：CCR模型，BCC模型
  - CCR模型
  - BCC模型

CCR模型(1/2)

- Charnes、Cooper和Rhodes (CCR)于1978年将Farrell (1957)的效率评估观念推广至多项投入和多项产出，并推导出一模型命名为CCR模型，并这种方法命名为数据包络分析法
- CCR模型假设固定规模报酬(constant return to scale)，也就是每一单位投入可得产出量是固定的，不会因规模大小而改变
- 有n个决策单元，DMU_j为n个决策单元中的一个，其效率可如下定义

s个产出项

m个投入项

$$E_j = \frac{\sum_{r=1}^{s} u_{rj} y_{rj}}{\sum_{i=1}^{m} v_{ij} x_{ij}} \qquad X_{ij}, Y_{rj} \geq 0, \quad i = 1, \cdots, m, r = 1, \cdots, s$$

$$u_{ij}, v_{rj} \geq 0, \quad i = 1, \cdots, m, r = 1, \cdots, s$$

CCR模型(2/2)

- 以向量的形式定义效率
  - Y_j为DMU_j的产出向量，X_j为DMU_j的投入向量

$$E_j = \frac{\mathbf{U}_j \cdot \mathbf{Y}_j}{\mathbf{V}_j \cdot \mathbf{X}_j} \qquad \mathbf{Y}_j = \begin{bmatrix} y_{1j} \\ y_{2j} \\ \vdots \\ y_{sj} \end{bmatrix} \qquad \mathbf{X}_j = \begin{bmatrix} x_{1j} \\ x_{2j} \\ \vdots \\ x_{mj} \end{bmatrix}$$

与上述式子相同

  - U_j为产出权重向量，V_j为投入权重向量

$$\mathbf{U}_j = [u_{1j}, u_{2j}, \cdots, u_{sj}] \qquad \mathbf{V}_j = [v_{1j}, v_{2j}, \cdots, v_{mj}]$$

数学规划模型——分式模型

- DEA法的数学规划模型是以一个决策单位DMU_p的效率E_p最大化作为目标函数，寻找最对DMU_p最有利的投入项权重组合，以及产出项权重组合，使得E_k达最大值，但所有DMU_j的效率E_j必须小于等于1，因此CCR模型的数学规划式如右所示

$$
\begin{aligned}
Max \quad & E_p = \frac{\sum_{r=1}^{s} u_{rp} y_{rp}}{\sum_{i=1}^{m} v_{ip} x_{ip}} \\
s.t. \quad & \frac{\sum_{r=1}^{s} u_{rp} y_{rj}}{\sum_{i=1}^{m} v_{ip} x_{ij}} \le 1, \quad j = 1,2,\cdots,n \qquad (1) \\
& u_{rp} \ge \varepsilon > 0, \quad r = 1,2,\cdots,s \\
& v_{ip} \ge \varepsilon > 0, \quad i = 1,2,\cdots,m
\end{aligned}
$$

non-Archimedean infinitesimal

乘数模型（multiplier model）

- 上式称为“投入导向(input oriented)模型”，E_p称为投入导向效率(input-oriented efficiency)

分式规划(1)可以等价转化为左侧的线性规划

$$
\begin{aligned}
Max \quad & E_p = \sum_{r=1}^{s} u_{rp} y_{rp} \\
s.t. \quad & \sum_{i=1}^{m} v_{ip} x_{ip} = 1 \qquad\qquad (2) \\
& \sum_{r=1}^{s} u_{rp} y_{rj} - \sum_{i=1}^{m} v_{ip} x_{ij} \le 0, \quad j = 1,2,\cdots,n \\
& u_{rp} \ge \varepsilon > 0, \quad r = 1,2,\cdots,s \\
& v_{ip} \ge \varepsilon > 0, \quad i = 1,2,\cdots,m
\end{aligned}
$$

对偶模型

- 为了让求解较有效率，并可以分析松弛变量(slack variable),我们通常考虑(2)的如下对偶规划

$$Min \quad E_p = \theta_p - \varepsilon\left(\sum_{i=1}^{m} s_{ip}^{-} + \sum_{r=1}^{s} s_{rp}^{+}\right)$$

$$s.t. \quad \sum_{j=1}^{n} \lambda_{jp} x_{ij} - \theta_p x_{ip} + s_{ip}^{-} = 0, \qquad i = 1,2,\cdots,m$$

$$\sum_{j=1}^{n} \lambda_{jp} y_{rj} - s_{rj}^{+} = y_{rp}, \qquad r = 1,2,\cdots,s \quad (3)$$

$$\lambda_{jp} \geq 0, \qquad j = 1,2,\cdots,n$$

$$s_{ip}^{-}, s_{rp}^{+} \geq 0, \qquad i = 1,2,\cdots,m, r = 1,2,\cdots,s$$

有效性判断及投影分析

- CCR efficient: 若E_p=1，则DMU_p是CCR(强)有效的；
- CCR weak efficient: 若 $\theta_p = 1$，则DMU_p是CCR弱有效的；
- CCR inefficient:若$\theta_p < 1$ ，则DMU_p是CCR无效的；
- 对于非有效的DMU_o，有如下投入改进和产出改进

$$\begin{aligned}\Delta x_o &= x_o - (\theta^* x_o - s^{-*}) = (1-\theta^*) x_o + s^{-*} \\ \Delta y_o &= s^{+*}.\end{aligned}$$

- 以及CCR投影(CCR projection)

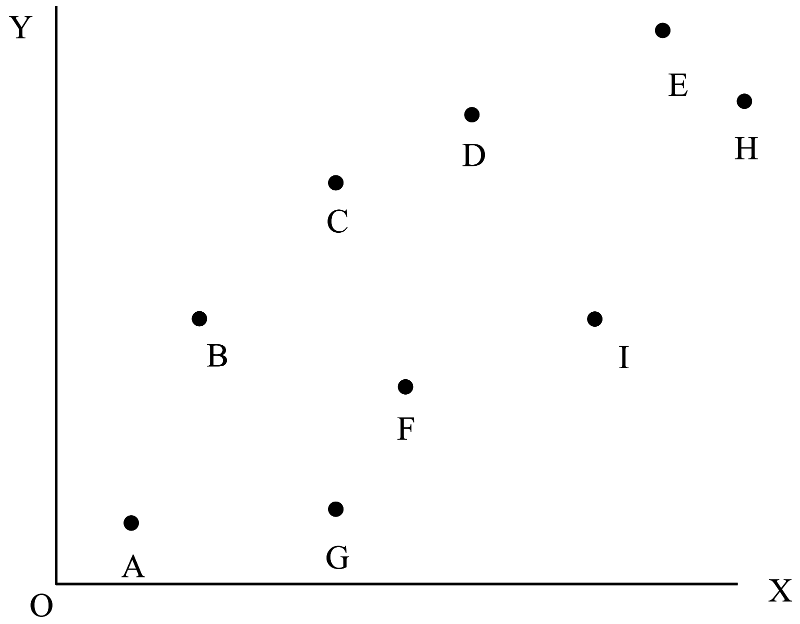
$$\begin{aligned}\hat{x}_o &= x_o - \Delta x_o = \theta^* x_o - s^{-*} \leq x_o \\ \hat{y}_o &= y_o + \Delta y_o = y_o + s^{+*} \geq y_o.\end{aligned}$$

Benchmark for improvement

A case with one input and one output

- Table 1 The inputs and outputs of 9 DMUs

DMU	A	B	C	D	E	F	G	H	I
input	1	2	4	6	9	5	4	10	8
output	1	4	6	7	8	3	1	7	4



$$
\begin{aligned}
&Max \quad E_p = u_{1p} y_{1p} \\
&s.t. \quad v_{1p} x_{1p} = 1 \\
&\qquad u_{1p} y_{1j} - v_{1p} y_{1j} \le 0, j = A, B, \cdots, I \\
&\qquad u_{1p}, v_{1p} \ge \varepsilon > 0
\end{aligned}
$$

以DMU_A与DMU_B为例进行求解

DMU_A为目标决策单元的CCR模型

$$
\begin{aligned}
Max \quad & E_A = u_{1A} \times 1 \\
s.t. \quad & v_{1A} \times 1 = 1 \\
& u_{1A} - v_{1A} \le 0 \\
& 4u_{1A} - 2v_{1A} \le 0 \\
& 6u_{1A} - 4v_{1A} \le 0 \\
& 7u_{1A} - 6v_{1A} \le 0 \\
& 8u_{1A} - 9v_{1A} \le 0 \\
& 3u_{1A} - 5v_{1A} \le 0 \\
& u_{1A} - 4v_{1A} \le 0 \\
& 7u_{1A} - 10v_{1A} \le 0 \\
& 4u_{1A} - 8v_{1A} \le 0 \\
& u_{1A}, v_{1A} \ge \varepsilon > 0
\end{aligned}
$$

$$u_{1A}^* = 0.5, v_{1A}^* = 1$$

$$E_A^* = 0.5$$

相对无效率的决策单元

DMU_B为目标决策单元的CCR模型

$$
\begin{aligned}
Max \quad & E_B = 4u_{1B} \\
s.t. \quad & v_{1B} \times 2 = 1 \\
& u_{1B} - v_{1B} \le 0 \\
& 4u_{1B} - 2v_{1B} \le 0 \\
& 6u_{1B} - 4v_{1B} \le 0 \\
& 7u_{1B} - 6v_{1B} \le 0 \\
& 8u_{1B} - 9v_{1B} \le 0 \\
& 3u_{1B} - 5v_{1B} \le 0 \\
& u_{1B} - 4v_{1B} \le 0 \\
& 7u_{1B} - 10v_{1B} \le 0 \\
& 4u_{1B} - 8v_{1B} \le 0 \\
& u_{1B}, v_{1B} \ge \varepsilon > 0
\end{aligned}
$$

$$u_{1B}^* = 0.25, v_{1B}^* = 0.5$$

$$E_B^* = 1$$

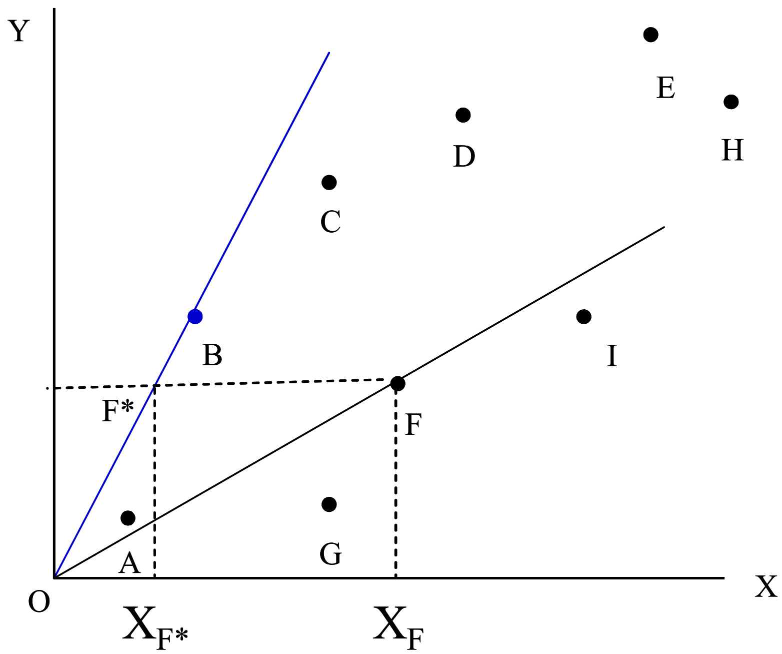
相对有效率的决策单元

九个决策单元的CCR模型解

决策单元	相对效率	投入权重	产出权重
A	0.500	1.000	0.050
B	1.000	0.500	0.250
C	0.750	0.250	0.125
D	0.583	0.167	0.083
E	0.444	0.111	0.056
F	0.300	0.200	0.100
G	0.125	0.250	0.125
H	0.350	0.100	0.050
I	0.250	0.125	0.063

在X-Y二维平面上寻找一条通过原点的包络线

寻找的方向从Y轴开始顺时钟旋轉，直到与第一个决策单元相交才停止，即为有效率的决策单元。



- 通过的蓝线即为效率前沿（efficiency frontier）

以DMU_F为例

- 对于相对无效率的决策单元，可以找出其在有效前沿的投影点，作为改善的参考方向，并与投影点的效率比较可得相对效率。
- 上图DMU_F为例，在有效前沿的投影点为F^*=(1.5,3)，就是在相同的产出水平下，欲达有效率，需减少水平达F^*，故DMU_F的相对效率为F点与$F*$点之效率比，其中点与点的效率分别为

$$E_F = \frac{\overline{OY}}{\overline{\overline{OX_F}}} \qquad E_{F*} = \frac{\overline{OY}}{\overline{\overline{OX_{F*}}}}$$

- DMU_F的相对效率为0.3

与数学规划模型求解的结果相同

$$h_F = \frac{E_F}{E_{F*}} = \frac{\overline{OY}}{\overline{\overline{OX_F}}} \times \frac{\overline{OX_{F*}}}{\overline{\overline{OY}}} = \frac{\overline{OX_{F*}}}{\overline{\overline{OX_F}}} = \frac{1.5}{5} = 0.3$$

以七个决策单元为例

决策单元	A	B	C	D	E	F	G
投入一	9	8.3	4.6	1.75	1.5	2.5	4
投入二	2.5	3	6	7	5	2.5	2
产出	20	18	16	14	12	14	16

$$Max \qquad E_p = u_{1p} y_{1p}$$

$$s.t. \qquad v_{1p} x_{1p} + v_{2p} x_{2p} = 1$$

$$u_{1p} y_{1j} - \left(v_{1p} x_{1j} + v_{2p} x_{2j} \right) \leq 0, \qquad j = A, B, \cdots, G$$

$$u_{1p}, v_{1p}, v_{2p} \geq \varepsilon > 0$$

以DUM_B为例进行求解

$$Max \quad E_B = 18u_{1B}$$

$$
\begin{aligned}
s.t. \quad & 8.3v_{1B} + 3v_{2B} = 1 \\
& 20u_{1B} - \left(9v_{1B} + 2.5v_{2B}\right) \le 0 \\
& 18u_{1B} - \left(8.3v_{1B} + 3v_{2B}\right) \le 0 \\
& 16u_{1B} - \left(4.6v_{1B} + 6v_{2B}\right) \le 0 \\
& 14u_{1B} - \left(1.75v_{1B} + 7v_{2B}\right) \le 0 \\
& 12u_{1B} - \left(1.5v_{1B} + 5v_{2B}\right) \le 0 \\
& 14u_{1B} - \left(2.5v_{1B} + 2.5v_{2B}\right) \le 0 \\
& 16u_{1B} - \left(4v_{1B} + 2v_{2B}\right) \le 0 \\
& u_{1B}, v_{1B}, v_{2B} \ge \varepsilon > 0
\end{aligned}
$$

最优权重

$$u_{1B}^* = 0.0417, v_{1B}^* = 0.000, v_{2B}^* = 0.333$$

效率

$$h_B^* = 0.75$$

相对无效率的决策单元

七个决策单元CCR模型解

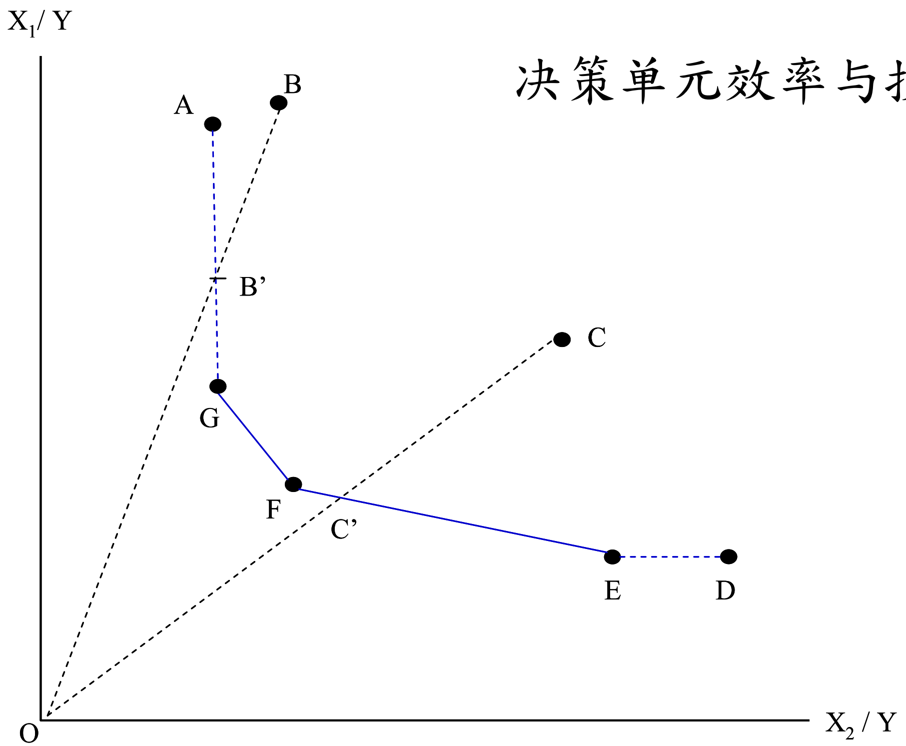
DMU	相对效率	投入X_1权重	投入X_2权重	产出Y权重
A	1.0000	0.0000	0.4000	0.0050
B	0.7500	0.0000	0.3330	0.0417
C	0.5882	0.1681	0.0378	0.0368
D	1.0000	0.0517	0.0000	0.0714
E	1.0000	0.3889	0.0833	0.0833
F	1.0000	0.3265	0.0735	0.0714
G	1.0000	0.1500	0.2000	0.0625

DMU_E、DMU_F、DMU_G为有效的决策单元

其余均为无效的决策单位

效率分析及求解

无效率决策单元的相对效率可经由其在有效前沿上的投影点两相比较后求得



效率前沿与效率分析

决策单元效率与投影点效率的比值 $h_C = \dfrac{E_C}{E_{C*}}$

$$h_C = \frac{\overline{OC^*}}{\overline{OC}}$$

$$h_C = \frac{\overline{OC^*}}{\overline{OC}} = 0.5882$$

产出导向(output-oriented)型 CCR模型

- 若效率比例改为如下定义：

$$E'_j = \frac{\sum_{i=1}^{m} v_{ij} x_{ij}}{\sum_{r=1}^{s} u_{rj} y_{rj}} \qquad Min \quad E'_p = \frac{\sum_{i=1}^{m} v_{ip} x_{ip}}{\sum_{r=1}^{s} u_{rp} y_{rp}}$$

- 相应的CCR分式模型为：

$$s.t. \quad \frac{\sum_{i=1}^{m} v_{ip} x_{ij}}{\sum_{r=1}^{s} u_{rp} y_{rj}} \geq 1, \quad j = 1,2,\cdots,n \qquad (4)$$

$$u_{rp} \geq \varepsilon > 0, \quad r = 1,2,\cdots,s$$

$$v_{ip} \geq \varepsilon > 0, \quad i = 1,2,\cdots,m$$

Output-oriented multiplier model

- multiplier model:

$$
\begin{aligned}
&Min \qquad E'_p = \sum_{i=1}^{m} v_{ip} x_{ip} \\
&s.t. \quad \sum_{r=1}^{s} u_{rp} y_{rp} = 1 \qquad\qquad (5) \\
&\qquad \sum_{r=1}^{s} u_{rp} y_{rj} - \sum_{i=1}^{m} v_{ip} x_{ij} \le 0, \quad j = 1,2,\cdots,n \\
&\qquad u_{rp} \ge \varepsilon > 0, \qquad r = 1,2,\cdots,s \\
&\qquad v_{ip} \ge \varepsilon > 0, \qquad i = 1,2,\cdots,m
\end{aligned}
$$

The dual form of output-oriented CCR model

- Dual form

$$Max \quad E'_p = \eta_p + \varepsilon\left(\sum_{i=1}^{m} t_{ip}^{-} + \sum_{r=1}^{s} t_{rp}^{+}\right)$$

$$s.t. \quad \sum_{j=1}^{n} \mu_{jp} x_{ij} - x_{ip} + t_{ip}^{-} = 0, \qquad i = 1, 2, \cdots, m$$

$$\sum_{j=1}^{n} \mu_{jp} y_{rj} - t_{rj}^{+} = \eta_p y_{rp}, \qquad r = 1, 2, \cdots, s \quad (6)$$

$$\mu_{jp} \geq 0, \qquad j = 1, 2, \cdots, n$$

$$t_{ip}^{-}, t_{rp}^{+} \geq 0, \qquad i = 1, 2, \cdots, m, r = 1, 2, \cdots, s$$

有效性判断及投影分析

- CCR efficient: iff. $\eta_p = 1$ and all $t_{ip}^-, t_{rp}^+ = 0$;
- CCR weak efficient: if $\eta_p = 1$;
- CCR inefficient: if $\eta_p > 1$;

- Output oriented CCR projection for inefficient DMU p

$$\hat{x}_{ip} = x_{ip} - t_{ip}^{-*}, i = 1, \cdots, m;$$

$$\hat{y}_{rp} = \eta_p^* y_{rp} + t_{rp}^{+*}, r = 1, \cdots, s;$$

The efficiency frontier of output oriented CCR model is consistent with that of input oriented CCR model, but their projections for inefficient DMUs are different .

BCC模型(1/3)

- 在不同的生产规模下，规模报酬将会随之改变，于初创期生产规模小时，投入产出比会随着规模增加而提升，称为规模报酬递增(Increasing Returns to Scale, IRS)
- 达到高峰期时，产出与规模成正比而达到最佳生产规模，称为规模报酬固定
- 当生产规模过于庞大时，产出减缓，则称为规模报酬递减(Decreasing Returns to Scale, DRS)，也就是投入增加时，产出增加的比例会少于投入增加的比例
- Banker等人在CCR模型的基础上进行变动规模报酬(Variable Returns to Scale, VRS)假设，并以此为基础衡量决策单元的相对效率，称为BCC模型(Banker et al., 1984)

BCC模型(2/3)

- 此模型将决策单元是否达到有效的生产规模也纳入评估，可同时衡量规模效率(scale efficiency)与技术效率(technical efficiency)
- 投入导向的BCC模式如右

$$Max \quad E_p = \frac{\sum_{r=1}^{s} u_{rp} Y_{rp}}{\sum_{i=1}^{m} v_i^k X_i^k}$$

$$s.t. \quad \frac{\sum_{r=1}^{s} u_{rp} Y_{rj}}{\sum_{i=1}^{m} v_{ip} X_{ij}} \leq 1, \quad j = 1, 2, \cdots, n$$

$$u_{rp} \geq \varepsilon > 0, \quad r = 1, 2, \cdots, s$$

$$v_{ip} \geq \varepsilon > 0, \quad i = 1, 2, \cdots, m$$

$$Max \quad t_p = \frac{\sum_{r=1}^{s} u_{rp} Y_{rp} - u_{0p}}{\sum_{i=1}^{m} v_i^k X_i^k}$$

Free in sign

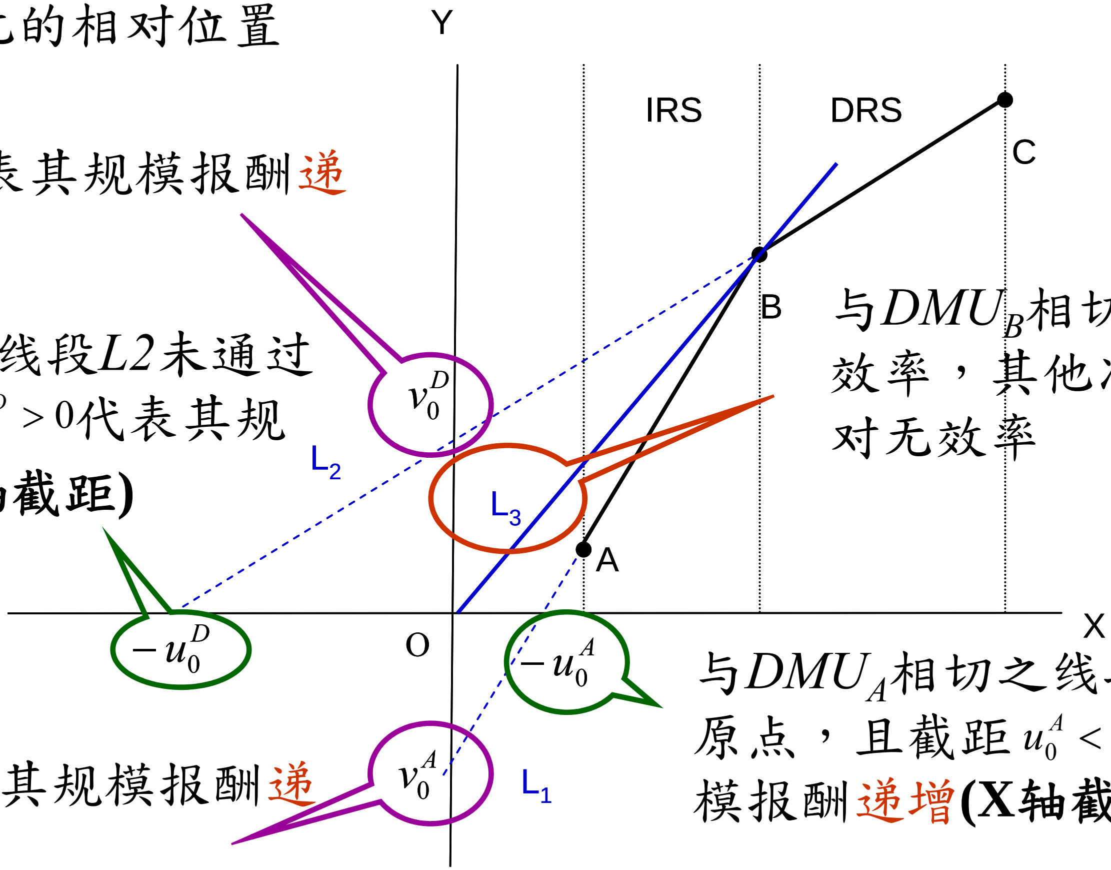
$$s.t. \quad \frac{\sum_{r=1}^{s} u_{rp} Y_{rj} - u_{0p}}{\sum_{i=1}^{m} v_{ip} X_{ij}} \leq 1, \quad j = 1, 2, \cdots, n$$

$$u_{rp} \geq \varepsilon > 0, \quad r = 1, 2, \cdots, s$$

$$v_{ip} \geq \varepsilon > 0, \quad i = 1, 2, \cdots, m$$

BCC模型(3/3)

- 显示一投入一产出下三个决策单元的相对位置



截距 $v_0^D > 0$代表其规模报酬递减 **(Y轴截距)**

与DMU_C相切的线段$L2$未通过原点，且截距 $u_0^D > 0$代表其规模报酬递减**(X轴截距)**

与DMU_B相切的线段$L3$达到有效率，其他决策单元都是相对无效率

截距 $v_0^A < 0$代表其规模报酬递增 **(Y轴截距)**

与DMU_A相切之线段$L1$未通过原点，且截距 $u_0^A < 0$代表其规模报酬递增**(X轴截距)**

BCC不同模型

一般线性规划模型

$$\begin{aligned}
Max \quad & t_p = \sum_{r=1}^{s} u_{rp} y_{rp} - u_{0p} \\
s.t. \quad & \sum_{i=1}^{m} v_{ip} x_{ip} = 1 \\
& \sum_{r=1}^{s} u_{rp} y_{rj} - \sum_{i=1}^{m} v_{ip} x_{ij} - u_{0p} \le 0, \quad j = 1,2,\cdots,n \\
& u_{rp} \ge \varepsilon > 0, \quad r = 1,2,\cdots,s \\
& v_{ip} \ge \varepsilon > 0, \quad i = 1,2,\cdots,m
\end{aligned}$$

分式规划模型

$$\begin{aligned}
Max \quad & t_p = \frac{\sum_{r=1} u_{rp} Y_{rp} - u_{0p}}{\sum_{i=1}^{m} v_i^k X_i^k} \\
s.t. \quad & \frac{\sum_{r=1}^{s} u_{rp} Y_{rj} - u_{0p}}{\sum_{i=1}^{m} v_{ip} X_{ij}} \le 1, \quad j = 1,2,\cdots,n \\
& u_{rp} \ge \varepsilon > 0, \quad r = 1,2,\cdots,s \\
& v_{ip} \ge \varepsilon > 0, \quad i = 1,2,\cdots,m
\end{aligned}$$

对偶问题

$$\begin{aligned}
Min \quad & t_p = \theta_p - \varepsilon \left(\sum_{i=1}^{m} s_{ip}^- + \sum_{r=1}^{s} s_{rp}^+ \right) \\
s.t. \quad & \sum_{j=1}^{n} \lambda_{jp} x_{ij} - \theta_p x_{ip} + s_{ip}^- = 0, \quad i = 1,2,\cdots,m \\
& \sum_{j=1}^{n} \lambda_{jp} y_{rj} - s_{rj}^+ = y_{rp}, \quad r = 1,2,\cdots,s \\
& \sum_{j=1}^{n} \lambda_{jp} = 1 \\
& \lambda_{jp}, s_{ip}^-, s_{jp}^+ \ge 0, \quad j = 1,\cdots,n, \quad i = 1,\cdots,m, \quad r = 1,\cdots,s \\
& \theta \quad free \quad in \quad sign
\end{aligned}$$

总效率与技术效率关系

- 决策单元未达相对有效率可能是技术无效率或者是规模无效率，然而CCR模型中并未考虑决策单位的规模差异，因此求得的相对效率E_p为总效率(overall efficiency)
- BCC模型求得的相对效率t_p为技术效率，和CCR模型总效率的差异为各个决策单元调整至相同生产规模后的规模效率

- 一个决策单元的总效率=技术效率*规模效率
- 当决策单元在最优生产力的规模下运作，已达规模效率，则其总效率与技术效率相等

一投入一产出之BCC模型

- 沿用CCR模型中一投入一产出九个决策单元的例子
- BCC模型化简为

$$
\begin{aligned}
Max \quad & t_p = \sum_{r=1}^{s} u_{rp} y_{rp} - u_{0p} \\
s.t. \quad & \sum_{i=1}^{m} v_{ip} x_{ip} = 1 \\
& \sum_{r=1}^{s} u_{rp} y_{rj} - \sum_{i=1}^{m} v_{ip} x_{ij} - u_{0p} \le 0, \quad j = 1, 2, \cdots, n \\
& u_{rp} \ge \varepsilon > 0, \qquad r = 1, 2, \cdots, s \\
& v_{ip} \ge \varepsilon > 0, \qquad i = 1, 2, \cdots, m
\end{aligned}
$$

$$
\begin{aligned}
Max \quad & t_p = u_{1p} y_{1p} - u_{0p} \\
s.t. \quad & v_{1p} x_{1p} = 1 \\
& u_{1p} y_{1j} - v_{1p} x_{1j} - u_{0p} \le 0, \qquad j = A, B, \cdots, I \\
& u_{1p}, v_{1p} \ge \varepsilon > 0 \\
& u_{0p} \ free
\end{aligned}
$$

以DMU_A与DMU_B为例进行求解

DMU_A为目标决策单元的BCC模型

$$
\begin{aligned}
Max \quad & E_A = u_{1A} \times 1 - u_{0A} \\
s.t. \quad & v_{1A} \times 1 = 1 \\
& u_{1A} - v_{1A} - u_{0A} \le 0 \\
& 4u_{1A} - 2v_{1A} - u_{0A} \le 0 \\
& 6u_{1A} - 4v_{1A} - u_{0A} \le 0 \\
& 7u_{1A} - 6v_{1A} - u_{0A} \le 0 \\
& 8u_{1A} - 9v_{1A} - u_{0A} \le 0 \\
& 3u_{1A} - 5v_{1A} - u_{0A} \le 0 \\
& u_{1A} - 4v_{1A} - u_{0A} \le 0 \\
& 7u_{1A} - 10v_{1A} - u_{0A} \le 0 \\
& 4u_{1A} - 8v_{1A} - u_{0A} \le 0 \\
& u_{1A}, v_{1A} \ge \varepsilon > 0 \\
& u_{0A} \ \ free
\end{aligned}
$$

DMU_B为目标决策单元的BCC模型

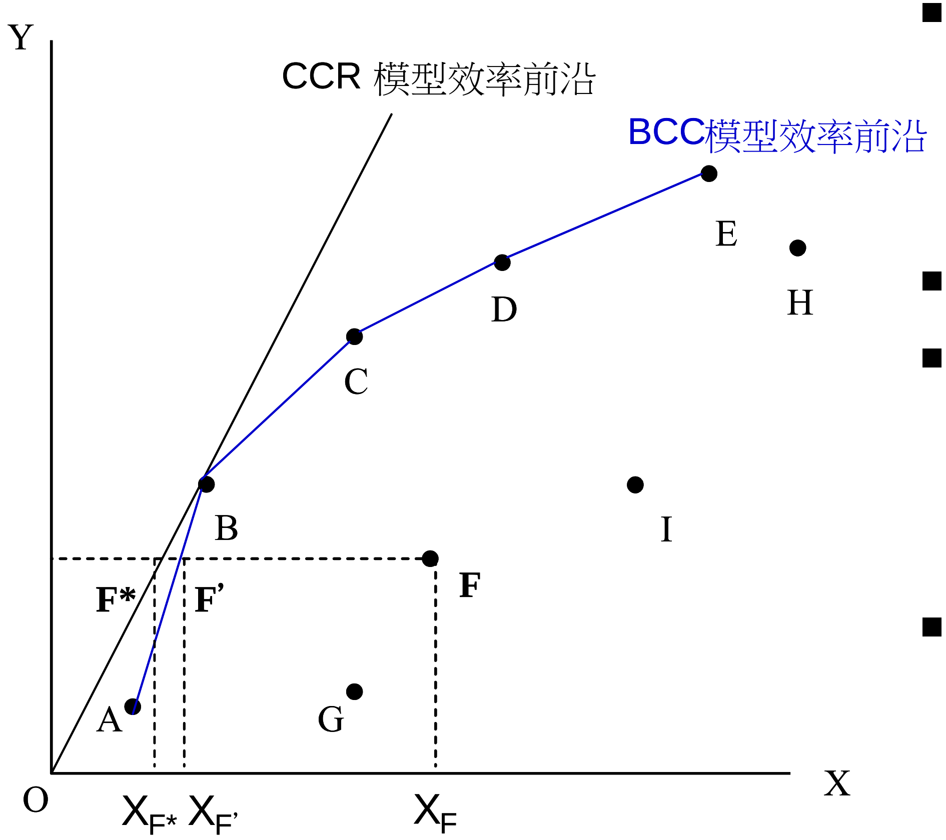
$$
\begin{aligned}
Max \quad & E_B = 4u_{1B} - u_{0B} \\
s.t. \quad & v_{1B} \times 2 = 1 \\
& u_{1B} - v_{1B} - u_{0B} \le 0 \\
& 4u_{1B} - 2v_{1B} - u_{0B} \le 0 \\
& 6u_{1B} - 4v_{1B} - u_{0B} \le 0 \\
& 7u_{1B} - 6v_{1B} - u_{0B} \le 0 \\
& 8u_{1B} - 9v_{1B} - u_{0B} \le 0 \\
& 3u_{1B} - 5v_{1B} - u_{0B} \le 0 \\
& u_{1B} - 4v_{1B} - u_{0B} \le 0 \\
& 7u_{1B} - 10v_{1B} - u_{0B} \le 0 \\
& 4u_{1B} - 8v_{1B} - u_{0B} \le 0 \\
& u_{1B}, v_{1B} \ge \varepsilon > 0 \\
& u_{0B} \ \ free
\end{aligned}
$$

只有决策单元B达到相对有效率
其他决策单位均为相对无效效率

九个决策单元的BCC模型解

决策单元	BCC模型 相对效率	投入权重	产出权重	截距	CCR模型 相对效率
A	1.000	0.000	1.000	-1.000	0.500
B	1.000	0.250	0.500	0.000	1.000
C	1.000	0.250	0.250	0.500	0.750
D	1.000	0.333	0.167	1.333	0.583
E	1.000	0.333	0.111	1.677	0.444
F	0.333	0.067	0.200	-0.133	0.300
G	0.250	0.000	0.250	-0.250	0.125
H	0.600	0.200	0.100	0.800	0.350
I	0.250	0.063	0.125	0.000	0.250

BBC模型示意图



- DMU_F的CCR模型效率为0.3，也就是总效率OE=0.3，而BBC模型效率为0.333，也就是技术效率TE=0.3
- 规模效率 $SE = 0.3 \div 0.333 = 0.901$
- 截距$u_0 = -0.133 < 0$，代表DMU_F是规模报酬递增，可再增加其生产规模而达到最优规模
- F'在$\overline{AB}$之间，也就是生产规模相当时，技术有效率的决策单元仅需投入$X_{F'}$即可获得与DMU_F相同的产出水平

DEA 使用步骤

- Roll与Golany (1989)所提出的DEA方法使用程序
- 可划分为四个大步骤
  - (1) 决策单元的选取
  - (2)投入产出项的选取
  - (3)DEA模型的选取
  - (4)评估结果的分析

问题定义与决策单元的选取

- 应用DEA于多属性评估问题时
  - 必须先定义问题了解问题本质，分清楚相对绩效评估的目的，建立设定评估准则
  - 取DEA模型构建效率边界，并将决策单元的实际投入产出与效率边界比较以衡量其效率，再将评估之结果加以分析，检查决策单元是否有效率
  - 相对无效率的决策单元，则分析其未达最优效率的原因，提出努力方向和矫正行动以改善其效率

投入项与产出项的选取

- DEA方法评估方案的相对效率时，是建立在各决策单位在各个投入或产出的相对表现数据的基础上，因此若选择了不适当的投入或产出，就如同选错目标以及决策努力方向一般，將扭曲相对效率的评估结果。
- DEA方法中，评估因子的决定和受评的方案之间是互动的;细节步骤如下：
  - 目标架构与对应的投入与产出因子
  - 区分投入因子或产出因子
  - 决定哪些投入与产出因子

决定哪些投入与产出项(1/2)

- 就个别因子而言，为确保纳入DEA分析的投入或产出项，均能一定程度地解释其对相对效率衡量结果的影响，可利用「后向消去法」(backward elimination)逐一检查并消去对效率较无影响之投入或产出因子
- DEA模型结果的相对权重值越小，通常表示该投入项或产出项对效率的影响越小(以线性模型而言)，后向消去法是从影响最小的投入或产出项开始删除
- 也可采用「前向选择法」(forward selection)，也就是由影响最大的投入或产出因子开始逐一检查并选入模型当中

决定哪些投入与产出项(2/2)

- 就因子总数而言，DEA方法在处理多项投入多项产出时虽有其优越处，但其所能处理之投入产出项个数并非毫无限制
- 每增加一项投入或产出因子，则会新增数个投入产出比率，导致DEA的鉴别力(discriminating power)降低
- 根据Golany与Roll(1989)提出的经验法则，投入因子与产出因子相加的总个数不能超过受评决策单元个数的二分之一
- Dyson等(2001)则提出更严格的标准，认为受评决策单位的个数不能低于投入因子个数与产出因子个数乘积的两倍

评估结果分析与诠释(1/2)

- DEA计算结果的分析与诠释，可以由下列四个方向来讨论：
  - 效率分析：由DEA的评估相对效率的结果，除了可以利用CCR模型计算总效率及BCC模型计算技术效率，并可以推導规模效率，同样是相对有效率的决策单元可比较其被其他决策单元参考的次数，被参考的次数越多，则表示该决策单位为相对有效率的衡量稳健度(robustness)越高
  - 差额变量分析：针对无效率的决策单元，可以通过DEA模型的差额变量分析(slack variable analysis)了解投入资源使用状况，找出无效率之来源及对应的因子应该改善的大小程度

评估结果分析与诠释(2/2)

- 规模报酬分析：决策单元的无效率可能是源自于技术效率或不同规模报酬的规模效率，在规模报酬变动时分析BCC模型
- 敏感度分析：因为DEA模式的结果容易受到考虑的投入产出以及决策单元的数据影响，为了提高研究的效度，使评估结果更具效果，因此必须进一步进行敏感度分析，这可以由三方面着手
  - ✓一是投入与产出因子增减变化对评估结果的影响
  - ✓二是决策单元变化对评估结果的影响
  - ✓三是各单元的投入与产出值变化对评估结果的影响。

DEA的特性(1/2)

- DEA和其他多目标决策模型一样，可以同时考虑多项的投入与产出因子，来评估决策单元的相对效率
- 投入或产出因子的度量单位改变，尽管会造成该因子值的改变，但并不会影响DEA的评估结果
- DEA是一种非参数方法，因此不需要事先知道投入与产出属性间的生产函数形式。然而，可以根据决策者主观偏好而建构相应的模型
- DEA可处理衡量单位不需统一的数据，使数据处理上更具弹性

DEA的特性(2/2)

- DEA是一种采用实证数据的基准比较(empirical benchmarking)，由于不和理论上的绝对标准比较，因此评估结果是相对效率而非绝对效率，因此效率为1并不代表没有改进之处
- DEA不仅可以评估相对效率，指出效率有待改进的DMU，也可以利用差额变量分析和敏感度分析，提供决策者各种改进效率值的可行途径
- DEA可以同时处理不同决策单元的多个投入与产出项而用单一总体衡量指标来表达相对效率
- DEA可以让各DMU找到对自己最有利的权重以尽可能的提升该DMU的效率，因此DEA所推导的权重不含人为主观因素

结论

- DEA是一种可以同时衡量多项产出与多项投入的多目标评估方法，也是一种非参数分析法，能够在不需默认生产函数型式且不需事先决定投入产出相对权重的条件下，求取各单元的相对效率。相关研究证明DEA不仅可用于非营利组织的绩效衡量，对于营利组织的绩效衡量，乃至于多目标决策问题均有其适用的地方。

本人的研究

- 基于DEA的基金绩效评价模型及其应用
- 基于DEA的环境能源效率评测

Murthi et al. (1997) DPEI

$$\max \ \mathrm{DPEI}_{j_0} = \frac{R_{j_0}}{\sum_{i=1}^{I} w_i c_{ij_0} + \upsilon \sigma_{j_0}}$$

$$\text{s.t.} \ \frac{R_j}{\sum_{i=1}^{I} w_i c_{ij} + \upsilon \sigma_j} \leq 1, \quad j = 1, \cdots, J,$$

$$w_i \geq \varepsilon, \ \upsilon \geq \varepsilon,$$

- J is the number of funds in the category,
- I is the number of different transaction costs,
- R_j is the return rate of the jth fund,
- σ_j is the standard deviation of the return for the jth fund,
- c_{ij} is the ith transaction cost for the jth fund, such as expense ratio, loads and turnover

Basso and Funari (2001)

- IDEA–1 index Compared with DPEI, it are only the investment costs which directly weigh on the investors, i.e., subscription and redemption fees; its risk measures include σ_j, $\sqrt{HV_j}$, β_j .

IDEA–2 index:

$$\max I_{j_0,\mathrm{DEA}-2} = \frac{u_1 R_{j_0} + u_2 d_{j_0}}{\sum_{i=1}^{h} v_{j_0} q_{ij_0} + \sum_{i=1}^{I} w_i c_{ij_0}}$$

$$\text{s.t.} \quad \frac{u_1 R_j + u_2 d_j}{\sum_{i=1}^{h} v_i q_{ij} + \sum_{i=1}^{I} w_i c_{ij}} \leq 1, \; j = 1, \cdots, J,$$

$$u_r \geq \varepsilon, \qquad r = 1, 2,$$

$$v_i \geq \varepsilon, \qquad i = 1, \cdots, h,$$

$$w_i \geq \varepsilon, \qquad i = 1, \cdots, I,$$

$q_{1j},\cdots,q_{hj}$ are h risk measures; d_j is the stoch·astic dominance indicator

IDEA–g index: add a few traditional performance indexes in output

Chen et al. (2006)

$$\max \frac{u_1 o_{j_0} + u_2 d_{j_0} + \sum_{r=1}^{p} \omega_r I_{rj_0}}{\sum_{i=1}^{h} \upsilon_i q_{ij_0} + \sum_{k=1}^{K} \omega_k c_{kj_0}}$$

$$\text{s.t.} \quad \frac{u_1 o_j + u_2 d_j + \sum_{r=1}^{p} \omega_r I_{rj}}{\sum_{i=1}^{h} \upsilon_i q_{ij} + \sum_{k=1}^{K} \omega_k c_{kj}} \leq 1, \; j = 1, \cdots, n,$$

$$u_r \geq \varepsilon, \qquad r = 1, 2,$$

$$\omega_r \geq \varepsilon, \qquad r = 1, \cdots, p,$$

$$\upsilon_i \geq \varepsilon, \qquad i = 1, \cdots, m,$$

$$\omega_k \geq \varepsilon, \qquad k = 1, \cdots, K,$$

Irj are the traditional performance indexes,. Risk measures include Var, CVaR and traditional risk measures.

Highlights:

- to reflect the skewness and leptokurtosis in return distributions of funds, VaR and CVaR are introduced into inputs of the existing DEA models;
- to fairly evaluate the relative performance of the same fund during different time periods, we treat the same fund during different periods as different DMUs

结束

Thank you ！