

基于 DEA 模型的中国足球协会超级联赛俱乐部效率研究

系统工程小组作业

潘崇林、唐溢、舒欣

商学院

2023 年 11 月 2 日



① 基本概念

② 实际应用

③ 参考文献



① 基本概念

DEA 基本介绍

CCR 模型

投入导向 BCC 模型

② 实际应用

③ 参考文献



① 基本概念

DEA 基本介绍

CCR 模型

投入导向 BCC 模型

② 实际应用

③ 参考文献



基本概念

- DMU(决策单元) 一个 DMU 就是一个将一定“输入”转换成一定“输出”的实体
- T(生产可能集) 一项经济 (生产) 活动中必然包含 x (输入) 与 y (产出) 两项活动
- 规模报酬 CCR 模型认为规模报酬不变,BCC 模型则认为可变



生产可能集

定义 1 (生产可能)

称集合 $T = \{(x, y) \mid \text{产出 } y \text{ 能用输入 } x \text{ 生产出来}\}$ 为所有可能的生产活动构成的生产可能集。

公理

- 凸性 对任意的 $(x, y) \in T$ 和 $(x', y') \in T$, 以及 $\mu \in [0, 1]$ 有 $\mu(x, y) + (1 - \mu)(x', y') \in T$
- 锥性 若 $(x, y) \in T$ 及 $k \geq 0$, 则 $k(x, y) = (kx, ky) \in T$
- 无效性 设 $(\{x\}, \{y\}) \in \{T\}$, 若 $x' \geq x$, 则 $(\{x'\}, \{y\}) \in \{T\}$; 若 $y' \leq y$ 则 $(\{x\}, \{y'\}) \in \{T\}$
- 最小性 生产可能集 $\{T\}$ 是满足上述条件 (1)、(2)、(3) 的所有集合的交集



① 基本概念

DEA 基本介绍

CCR 模型

投入导向 BCC 模型

② 实际应用

③ 参考文献



相关数据

	指标	部门 权数	1	2		j		n
投入	1	v_1	x_{11}	x_{12}		x_{1j}		x_{1n}
	2	v_2	x_{21}	x_{22}		x_{2j}	\cdots	x_{2n}
	:	\vdots	\vdots	\vdots	\cdots	\vdots	\vdots	\vdots
	m	v_m	x_{m1}	x_{m2}		x_{mj}		x_{mn}
输出	1	u_1	y_{11}	y_{12}		y_{1j}		y_{1n}
	2	u_2	y_{21}	y_{22}		y_{2j}	\cdots	y_{2n}
	:	\vdots	\vdots	\vdots	\cdots	\vdots	\vdots	\vdots
	q	u_p	y_{p1}	y_{p2}		y_{pj}		y_{pn}



技术效率

技术效率指的是一个生产单元的生产过程达到本行业技术水平的程度。第 k 个决策单元相应的效率评价指数：

$$h_k = \frac{u_1 y_{1k} + u_2 y_{2k} + \cdots + u_q y_{qk}}{v_1 x_{1k} + v_2 x_{2k} + \cdots + v_m x_{mk}} = \frac{\sum_{r=1}^q u_r y_{rk}}{\sum_{i=1}^m v_i x_{ik}}$$



线性规划形式

$$\begin{aligned}
 & \max \sum_{r=1}^q \mu_r y_{rk} \\
 & \text{s.t.} \quad \sum_{r=1}^q \mu_r y_{rj} - \sum_{i=1}^m \nu_i x_{ij} \leq 0 \\
 & \quad \sum_{i=1}^m \nu_i x_{ik} = 1 \\
 & \quad \mu \geq 0; \nu \geq 0 \\
 & \quad i = 1, 2, \dots, m; r = 1, 2, \dots, q; j = 1, 2, \dots, n
 \end{aligned}$$



对偶形式

$$\begin{aligned}
 & \min \theta \\
 \text{s.t. } & \sum_{j=1}^n \lambda_j x_{ij} \leq \theta x_{ik} \\
 & \sum_{j=1}^n \lambda_j y_{rj} \geq y_{rk} \\
 & \lambda_j \geq 0 \\
 & i = 1, 2, \dots, m; r = 1, 2, \dots, q; j = 1, 2, \dots, n
 \end{aligned}$$

λ 表示 DMU 的线性组合系数, k 表示待评价的 DMU, 参数 θ^* 即为效率值, 其范围在 0 到 1 之间。



矩阵形式

$$\min 0 \cdot \lambda_1 + 0 \cdot \lambda_2 + \cdots + 0 \cdot \lambda_n + 1 \cdot \theta$$

$$\text{s.t.} \quad \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} & -x_{1k} \\ x_{21} & x_{22} & \cdots & x_{2n} & -x_{2k} \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ x_{m1} & x_{m2} & \cdots & x_{mn} & -x_{mk} \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \\ \theta \end{bmatrix} \leq \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} -y_{11} & -y_{12} & \cdots & -y_{1n} & 0 \\ -y_{21} & -y_{22} & \cdots & -y_{2n} & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ -y_{q1} & -y_{q2} & \cdots & -y_{qn} & 0 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \\ \theta \end{bmatrix} \leq \begin{bmatrix} -y_{1k} \\ -y_{2k} \\ \vdots \\ -y_{qk} \end{bmatrix}$$

$$\lambda_j \geq 0$$

$$i = 1, 2, \cdots, m; r = 1, 2, \cdots, q; j = 1, 2, \cdots, n$$



投入导向与产出导向

- 投入导向假设投入是固定的条件下求产出最大
- 产出导向假设产出是固定的条件下求投入最小
- 显然上面介绍的模型是投入导向



① 基本概念

DEA 基本介绍

CCR 模型

投入导向 BCC 模型

② 实际应用

③ 参考文献



对偶形式

$$\begin{aligned}
 & \min \theta \\
 \text{s.t. } & \sum_{j=1}^n \lambda_j x_{ij} \leq \theta x_{ij} \\
 & \sum_{j=1}^n \lambda_j y_{rj} \geq y_{rj} \\
 & \sum_{j=1}^n \lambda_j = 1 \\
 & \lambda \geq 0 \\
 & i = 1, 2, \dots, m; r = 1, 2, \dots, q; j = 1, 2, \dots, n
 \end{aligned}$$



标准形式

$$\begin{aligned}
 & \min \theta \\
 \text{s.t. } & \sum_{j=1}^n \lambda_j x_{ij} + s^{-} = \theta x_{ik} \\
 & \sum_{j=1}^n \lambda_j y_{rj} - s^{+} = y_{rk} \\
 & \sum_{j=1}^n \lambda_j = 1 \\
 & \lambda \geq 0 \\
 & i = 1, 2, \dots, m; r = 1, 2, \dots, q; j = 1, 2, \dots, n
 \end{aligned}$$



DEA 有效性

$s^- s^+$ 是松弛变量设 CCR 与 BCC 模型的最优解分别为 $\lambda^* s^{*-} s^{*+} \theta^*$, 则可对决策单元进行有效性检验. 值得注意的是, DEA 模型的有效性是“**相对**”有效性

有效性

- 若 $\theta^* < 1$, 决策单元为“非有效”
- 若 $\theta^* = 1 s^{*-} \neq 0 s^{*+} \neq 0$, 决策单元为“弱有效”
- 若 $\theta^* = 1 s^{*-} = 0 s^{*+} = 0$, 决策单元为“有效”



实际意义

参数名称	符号	含义
综合效率	e	$\theta^* e = te \times se$
技术效率	te	单独考虑技术效率，CCR 则相反
规模效率	se	规模效率表示为 λ_j
规模报酬	rs	单位产出与投入比例随规模变化，可以是 递增、递减或不变
投入冗余	γ	BCC 模型可以同时评估单位的效率和规模报酬 也可以表示为 s^+
冗余度	α	投入冗余值与投入的比值



① 基本概念

② 实际应用

基本介绍

中超俱乐部效率评价

效率优化

③ 参考文献



① 基本概念

② 实际应用

基本介绍

中超俱乐部效率评价

效率优化

③ 参考文献



案例简介

基于 2015-2017 赛季中国足球协会超级联赛 (以下简称“中超”) 俱乐部的面板数据, 从“投入产出”角度建立俱乐部效率评价体系, 运用数据包络分析 (DEA) 方法对中超俱乐部效率进行测度, 并通过效率优化解决其投入冗余问题

决策单元为中超的所有俱乐部, 2015-2017 赛季中超参赛球队数目固定为 16 支, 这也符合 DEA 模型对于决策单元与投入产出指标之间数量上的要求, 即决策单元数量应多于投入与产出指标数量之和的 2 倍

中超俱乐部评价体系

指标属性	变量	含义	评价目的
投入	x_1	球员身价	人工成本
产出	y_1	积分	成绩
	y_2	胜场	成绩
	y_3	进球	成绩

- 求解工具 MATLAB、DEAP 2.1



中超俱乐部效率测度

俱乐部	2015 赛季				2016 赛季				2017 赛季			
	身价	积分	胜场	进球	身价	积分	胜场	进球	身价	积分	胜场	进球
广州恒大	4570	67	19	71	4645	64	19	62	4670	64	20	69
上海上港	1703	65	19	63	3648	52	14	56	7580	58	17	72
山东鲁能	2580	59	18	66	3163	34	9	38	2553	49	13	49
北京国安	1725	56	16	46	2128	43	11	34	2578	40	11	42
河南建业	500	46	12	35	628	35	10	26	730	30	7	34
上海申花	2163	42	12	42	2965	48	12	46	2670	35	9	52
石家庄永昌	1075	39	8	34	550	30	7	28	-	-	-	-
重庆力印	788	35	9	37	1050	37	9	43	835	36	9	37
江苏苏宁	1128	35	9	39	4503	57	17	53	4385	32	7	40
长春亚泰	1038	35	8	39	708	35	10	30	1770	44	12	46
杭州绿城	510	33	8	27	480	32	8	28	-	-	-	-
辽宁开新	853	31	7	30	720	36	9	38	710	18	4	30
天津泰达	1565	31	7	39	1258	36	9	38	2920	31	8	30
广州富力	1143	31	8	35	1265	40	11	47	1385	52	15	59
贵州茅台	1103	29	7	39	-	-	-	-	-	-	-	-
上海申鑫	385	17	4	30	-	-	-	-	-	-	-	-
河北华夏	-	-	-	-	2500	40	11	34	2218	52	15	55
延边富德	-	-	-	-	490	37	10	39	533	22	5	32
天津权健	-	-	-	-	-	-	-	-	3188	54	15	46
贵州智诚	-	-	-	-	-	-	-	-	1098	42	12	39



指标数据处理

- 替换指标
- 无量纲化处理 功效系数法等
- 赋一个极小值 如 0.01、0.001 等

不变有效性

- 推论 (1) 各决策单元的同一指标数据同时扩大或缩小正的若干倍, 其 DEA 有效性 (BC^2) 不变;
- 推论 (2) 各决策单元的同一指标数据同时加上相同的正数或适当减去相同的正数使得变换后的数据 $(\bar{x}_j, \bar{y}_j) > 0$ 时, 其 DEA 有效性 (BC^2) 不变.

上述方法处理的依据是来自 DEA 数据各决策单元同一指标经平移、放缩不改变其 DEA 有效性, 但其具体效率值可能改变



软件求解

```

clc,clear;
% 导入数据
data = xlsread('zuqiu_data.xlsx',2);
X= data(:,1)'; % 投入指标数据, 每一列代表每个决策单元的投入数据
Y= data(:,2:4)'; % 产出指标数据, 每一列代表每个决策单元的产出数据
n=size(X,2); % 决策单元数
m=size(X,1); % 投入指标数
q=size(Y,1); % 产出指标数
% 投入导向数据
w = [];
for i = 1:n
    f = [zeros(1,n) 1]; % 定义目标函数
    A = [X -X(:,i); -Y zeros(q,1)]; % 指定不等式约束
    b = [zeros(1,m) -Y(:,i)'];
    Aeq = [ones(1,n) 0]; % 定义等式约束
    beq = 1;
    LB = [zeros(n+1,1)]; % 指定下界
    UB = [];
    w(:,i) = linprog(f,A,b,Aeq,beq,LB,UB); % 模型求解
end
theta_BCC_input = w(n+1,:); % 结果输出

```

```

zq2.txt      DATA FILE NAME
zq2-out.txt  OUTPUT FILE NAME
20          NUMBER OF FIRMS
1           NUMBER OF TIME PERIODS
3           NUMBER OF OUTPUTS
1           NUMBER OF INPUTS
0           0=INPUT AND 1=OUTPUT ORIENTATED
1           0=CRS AND 1=VRS
4           0=DEA(MULTI-STAGE), 1=COST-DEA, 2=MALMQUIST-DEA, 3=DEA(1-
STAGE), 4=DEA(2-STAGE)

```

DEAP 2.1 求解

MATLAB 求解



① 基本概念

② 实际应用

基本介绍

中超俱乐部效率评价

效率优化

③ 参考文献



效率及规模收益

俱乐部	2015 赛季				2016 赛季				2017 赛季			
	身价 e	积分 te	胜场 se	进球 rs	身价 e	积分 te	胜场 se	进球 rs	身价 e	积分 te	胜场 se	进球 rs
广州恒大	0.214	1.000	0.214	drs	0.200	1.000	0.200	drs	0.393	1.000	0.393	drs
上海上港	0.518	1.000	0.518	rs	0.193	0.903	0.214	drs	0.21	1.000	0.21	rs
山东鲁能	0.353	1.000	0.353	drs	0.151	0.155	0.976	irs	0.47	0.502	0.936	drs
北京国安	0.386	0.688	0.561	rs	0.268	0.664	0.403	drs	0.394	0.394	1.000	crs
河南建业	1.000	1.000	1.000	crs	0.780	0.780	1.000	crs	0.964	0.967	0.997	irs
上海申花	0.270	0.370	0.729	drs	0.214	0.736	0.291	rs	0.347	0.436	0.795	rs
石家庄永昌	0.442	0.444	0.995	rs	0.722	0.873	0.828	irs				
重庆力帆	0.638	0.744	0.859	rs	0.515	0.836	0.616	drs	1.000	1.000	1.000	crs
江苏苏宁	0.467	0.596	0.784	drs	0.185	0.826	0.224	rs	0.173	0.183	0.944	drs
长春亚泰	0.506	0.647	0.782	rs	0.692	0.692	1.000	crs	0.626	0.627	0.998	drs
杭州绿城	0.747	0.879	0.849	irs	0.883	1.000	0.883	irs				
辽宁开新	0.483	0.516	0.936	irs	0.663	0.679	0.976	irs	0.704	0.751	0.938	irs
天津泰达	0.331	0.429	0.771	drs	0.380	0.389	0.976	irs	0.253	0.26	0.974	irs
广州富力	0.413	0.437	0.945	rs	0.467	1.000	0.467	drs	1.000	1.000	1.000	crs
贵州茅台	0.466	0.609	0.765	rs								
上海申菱	1.000	1.000	1.000	crs								
河北华夏					0.216	0.381	0.566	rs	0.622	0.624	0.996	drs
延边富德					1.000	1.000	1.000	crs	1.000	1.000	1.000	crs
天津权健									0.432	0.606	0.712	drs
贵州智诚									1.000	1.000	1.000	crs
均值	0.515	0.710	0.754		0.471	0.745	0.664		0.599	0.709	0.868	

规模收益分析

效率	含义	俱乐部		
		2015 赛季	2016 赛季	2017 赛季
[0, 0.6)	无效率	广州恒大、上海上港、山东鲁能、北京国安、上海申花、石家庄永昌、江苏苏宁、长春亚泰、辽宁开新、天津泰达、广州富力、贵州茅台	广州恒大、上海上港、山东鲁能、北京国安、上海申花、重庆力帆、江苏苏宁、天津泰达、广州富力、河北华夏	广州恒大、上海上港、山东鲁能、北京国安、上海申花、江苏苏宁、天津泰达、天津权健
[0.6, 0.8)	较低效率	重庆力帆、杭州绿城	河南建业、石家庄永昌、长春亚泰、辽宁开新	长春亚泰、辽宁开新、河北华夏
[0.8, 1)	中等效率		杭州绿城	河南建业
1	高效率	河南建业、上海申鑫	延边蓄德	重庆力帆、广州富力、延边富德、贵州智诚



中超俱乐部规模收益观察

- 2015-2017 赛季的 48 个规模收益观测值中，有 28 个为规模收益递减，共占总体的 58.3%。
- 这些递减的观测值分别在各赛季中为 12 个、8 个和 8 个。
- 这表明中超大多数俱乐部的投入都处于“饱和型”阶段，投入相对冗余，但由于配置不合理导致浪费现象严重，这也制约了俱乐部效率的提升。
- 值得注意的是，广州恒大、上海上港、上海申花和江苏苏宁这四支俱乐部在 2015-2017 赛季均为规模收益递减。
- 这一方面印证了这四支俱乐部雄厚的资金实力，但另一方面也反映了它们在资金投入的配置上仍存在一定问题。
- 另外，有 10 个观测值为规模收益递增，占总体的 20.8%。



① 基本概念

② 实际应用

基本介绍

中超俱乐部效率评价

效率优化

③ 参考文献



效率优化

俱乐部	2015 赛季		2016 赛季		2017 赛季	
	r	α	r	α	r	α
广州恒大	0.000	0.000%	0.000	0.000%	0.000	0.000%
上海上港	0.000	0.000%	355.000	9.731%	0.000	0.000%
山东鲁能	0.000	0.000%	2673.909	84.537%	1271.125	49.789%
北京国安	537.571	31.164%	714.667	33.584%	1562.778	60.620%
河南建业	0.000	0.000%	138.000	21.975%	24.429	3.346%
上海申花	1362.250	62.980%	782.222	26.382%	1505.889	56.400%
石家庄永昌	598.000	55.628%	70.000	12.727%		
重庆力帆	202.071	25.644%	172.500	16.429%	0.000	0.000%
江苏苏宁	456.143	40.438%	781.333	17.351%	3580.579	81.655%
长春亚泰	366.143	35.274%	218.000	30.791%	660.000	37.288%
杭州绿城	61.552	12.069%	0.000	0.000%		
辽宁开新	412.483	48.357%	230.909	32.071%	177.000	24.930%
天津泰达	893.143	57.070%	768.909	61.122%	2160.500	73.990%
广州富力	643.000	56.255%	0.000	0.000%	0.000	0.000%
贵州茅台	431.143	39.088%				
上海申金	0.000	0.000%				
河北华夏			1548.333	61.933%	833.000	37.556%
延边富德			0.000	0.000%	0.000	0.000%
天津权健					1255.500	39.382%
贵州智诚					0.000	0.000%



中超俱乐部投入冗余情况

- 2015-2017 赛季中，中超俱乐部普遍存在投入冗余情况
- 各赛季投入冗余的俱乐部分别达到 11、12、10 支，占总体的 68.75%、75%、62.5%
- 这表明超过一半的俱乐部资金投入未能得到充分发挥，投入效能较低
- 此外，俱乐部投入冗余度也较大，排名前三的俱乐部投入冗余度均超过 50%。
- 结合表 4 可以发现，这些俱乐部的效率均落入无效率区间
- 这与它们高投入冗余度的结果是一致的



① 基本概念

② 实际应用

③ 参考文献



参考文献 I



Thanks!

