
DEA

(Data Envelopment Analysis)



FOM
Focus On Data Mining

INDEX

1.Introduction

2.What's DEA

3.Example/Practice of DEA

4.Conclusion

01 | Introduction

MCDM (다기준 의사결정)

• MCDM : Multi Criteria Decision Making

✓ 다기준 의사결정 - MCDM – If not a Roman Numeral, then What? In 1979

✓ 속성 :

정량적 요소와 정성적 요소가 혼재

상이한 scale의 혼재로 단일 지수로 통합하기 어려움

요소간 상대적 중요성 평가가 어려움

복수의 목표, 대안, 속성의 동시 고려가 어려움

✓ 접근 방법 :

Scoring 모형 -> 자료의 문제를 scaleless하거나 점수 scale로 통일하여 해결

AHP -> 비교가능성을 가정하고 intransitivity(비이행성)를 허용

Fuzzy -> 비교가능성과 intransitivity를 허용하지 않는 접근

ANP -> 속성간 inter independence를 가정

02 | What's DEA

Definition

• DEA : Data Envelopment Analysis

- ✓ 자료 포락 분석
- ✓ 1978년 처음으로 제안 -> OR/M (Operations Research/Management)에 널리 활용됨
- ✓ 선형계획법을 기반으로 DMU 사이의 상대적인 효율성을 비교하는 기법
- ✓ 상대적 효율성, 성능 측정 및 벤치마킹, 비모수적 접근하는 기법
- ✓ DMU라는 의사결정 단위에 투입 및 산출을 적용 후 개별 DMU별로 떨어진 거리 계산
- ✓ 자원 배분을 최적화 하고 생산을 최대화 하는 것

• 특징 :

- ✓ '과제 중심적 접근' 과 '중요한 과제에 초점'을 맞추어 모든 비교 대상의 성과를 과학적,정량적으로 도출하며, 나아가 개선해야 할 부분과 개선의 가능성, best practice와 benchmark 대상 등을 제시
(위에서의 비교 대상이란, DMU : 프로젝트, R&D 과제, 기관, 지점, 관리자의 업적 등 Business Operations나 Business Process)
(DMU : Decision Maker Units)
- ✓ Market Price가 없는 경우에 상대적 효율성을 측정하기 위해 고안됨
- ✓ Input,output간에 어떤 사전적 함수관계에 대한 가정이 없더라도 modeling할 수 있는 장점

02 | What's DEA

Definition

• DEA 방식

- ✓ 비교 대상인 의사결정단위를 결정 -> 각 의사결정단위의 투입물과 산출물을 추계하여 DEA 분석 진행
- ✓ 산출 요소별 가중치를 곱해 더한 가중 합계를 투입 요소별 가중치를 곱해 더한 가중 합계로 나눠줌
- ✓ 위의 연산의 결과물이 DEA 기법에 따른 효율성
- ✓ 목표 : optimizing
 - input : 효율을 올리기 위한 식에서 분모를 선형이 되려면 1이 되어야함
 - output : 모든 k개의 상대적 효율성은 1보다 작거나 같아야함

02 | What's DEA

DEA의 장점

• 장점

- ✓ 다른 성격의 여러 투입과 산출 요소를 함께 고려한 분석
- ✓ 객관적 요소의 통합
- ✓ 비효율적 DMU의 약점 부분 및 실적 개선 가능 목표 제시
- ✓ Star player와 best practice 도출
- ✓ 비효율적 DMU의 benchmark 대상 제시
- ✓ 투입과 산출사이의 정해진 함수관계를 설정하지 않고 다수의 투입물과 다수의 산출물을 내생적으로 조절
- ✓ 투입과 산출에 대한 가중치를 사전에 정할 필요가 없음
- ✓ 효율성 값이 간단한 measure로 제시, 상대적 비효율성의 원인이 되는 투입물과 산출물을 확인 가능

02 | What's DEA

Techniques

•Objective function of DMU's efficiency θ_j (CCR Model)

DMU's known as M outputs $y_1^j, y_2^j, \dots, y_m^j$ are multiplied by their respective weights $u_1^j, u_2^j, \dots, u_m^j$ and divided by the N inputs $x_1^j, x_2^j, \dots, x_m^j$ multiplied by their respective weights $v_1^j, v_2^j, \dots, v_m^j$

$$\max \theta_j = \frac{\sum_{m=1}^M y_m^j u_m^j}{\sum_{n=1}^N x_n^j v_n^j}$$

•Constraint of DMU's efficiency θ_j

No DMU_k $k=1, \dots, K$, efficiency scores exceeds one

$$\frac{\sum_{m=1}^M y_m^k u_m^j}{\sum_{n=1}^N x_n^k v_n^j} \leq 1 \quad k = 1, \dots, K,$$

► Variables

Notations	Description
E_k	The efficiency ratio of DMU k $k=1,2,\dots,K$ where K is the total number of DMUs being evaluated
u_j	A coefficient for output j – A measure of the relative decrease in efficiency with each unit reduction of output value $j=1,2,\dots,M$ where M is the total number of output types being considered
v_i	A coefficient for input i – A measure of the relative increase in efficiency with each unit reduction of input value $i=1,2,\dots,N$ where N is the total number of input types being considered
O_{jk}	The number of observed units of output j generated by DMU k
I_{jk}	The number of observed units of input i used by DMU k

02 | What's DEA

Techniques

- Objective function of DMU_k's efficiency θ_j , for $k=1,...,K$

✓ Change the objective function to more simple one

$$\max \sum_{m=1}^M y_m^j u_m^j$$

- Constraint of DMU_k's efficiency θ_j , for $k=1,...,K$

✓ Change the constraint to more simple one

$$1. \sum_{n=1}^N x_n^j v_n^j = 1$$

$$2. \sum_{m=1}^M y_m^j u_m^j \leq \sum_{n=1}^N x_n^j v_n^j$$

This new Objective and Constraints are same effect with over ones

02 | What's DEA

Techniques

- Objective function 및 Constraints Linear form

$$\max E_e = u_1 O_{1e} + u_2 O_{2e} + \cdots + u_M O_{Me}$$

$$\text{s.t. } u_1 O_{11} + u_2 O_{21} + \cdots + u_M O_{M1} - (v_1 I_{11} + v_2 I_{21} + \cdots + v_N I_{N1}) \leq 1$$

$$\vdots$$
$$\vdots$$

$$u_1 O_{1K} + u_2 O_{2K} + \cdots + u_M O_{MK} - (v_1 I_{1K} + v_2 I_{2K} + \cdots + v_N I_{NK}) \leq 1$$

$$v_1 I_{1e} + v_2 I_{2e} + \cdots + v_N I_{Ne} = 1$$

$$u_j \geq 0, \quad j = 1, 2, \dots, M$$

$$v_i \geq 0, \quad i = 1, 2, \dots, N$$

02 | What's DEA

Techniques

- Criteria

- ✓ Return to scale

- CCR : constant returns to scale
 - BCC : variable returns to scale(increasing, constant, decreasing)

- ✓ Objectives

- Input-oriented : Minimize inputs at the given level of outputs (input 지향)
 - Output-oriented : Maximize outputs at the given level of inputs (output 지향)

02 | What's DEA

Techniques

DEA의 초기 모형 : CCR모형, BCC모형

➤ DEA models

	CCR	BCC
Input-oriented	Input-oriented CCR model	Input-oriented BCC model
Output-oriented	Output-oriented CCR model	Output-oriented BCC model

03 | Example/Practice of DEA

체인호텔의 상대적 효율성 측정

• 현 상황

최근 고정비용 상승에 따른 생산성의 하락을 경험 중 -> 투입 대비 산출의 효과가 크지 않음. 호텔 건립 당시 초기 투자비용은 차지하고라도 매출이 제한적인 상황에서 인건비 상승으로 인해 수익성이 악화 됨 -> 이를 해결하고자 DEA를 수행

• 호텔의 상대적 효율성 측정의 중요성

많은 호텔들이 고정자산투자 증가에 따른 영업위험과 타인자본 부담에 따른 재무 위험이 증가 -> 호텔의 수익성과 재무 건전성 악화시킴, 객실과 더불어 F&B 등 각종 부대사업장의 매출 역시 침체 중 -> 차별적인 경쟁력이나 코스트를 획기적으로 절감해야 수익성이 생김

이를 위해 코스트의 구조를 파악, 다른 호텔과의 비교를 통해 비용을 어떻게 줄일지 결정 -> 다른 호텔과의 비교는 기본적으로 부분적인 생산성 정도를 이용한 평가를 수행

생산성 정도를 사용해 비교를 진행시 1:1값을 기준으로 한 성과 지표는 각 평가단위를 단순 비교함으로써 잘못을 범하기 쉬움 -> DEA는 비교대상그룹에서 benchmark를 제시, benchmark 대비 각 비교대상의 효율성을 제시

03 | Example/Practice of DEA

체인호텔의 상대적 효율성 측정

- What is input,output?

- ✓ Input :

- 종업원수, 유형자산, 객실 수, 부대사업장크기

- ✓ Output :

- 매출, 객실 점유율, RevPar, 영업이익

- ✓ DEA는 다수의 input,output의 산출을 한 번에 고려해 효율성을 측정

- ✓ 효율성을 분해해 기술효율성, 순수기술효율성, 규모효율성 측정 가능

- ✓ DEA의 초기 모형 : CCR모형, BCC모형

03 | Example/Practice of DEA

체인호텔의 상대적 효율성 측정

•What's CCR :

- ✓ CCR모형의 기본 가정 : 생산함수가 규모수익불변(Constant Return to Scale : CRS) -> 규모에 따라 산출되는 수익이 일정 비율로 증가 -> 규모에 따라 증가하므로 무한확장가능성이 있음
- ✓ 일정 비율의 수익률

- Input 방향의 CCR포락모형

$$\text{Min } h_0 = \theta - \epsilon \left(\sum_{i=1}^m S_i^- + \sum_{\gamma=1}^s S_{\gamma}^+ \right) \quad \langle \text{식 1} \rangle$$

$$\text{subject to } \theta x_{i0} - S_i^- - \sum_{j=1}^n x_{ij} \lambda_j = 0$$

$$\sum_{j=1}^n y_{\gamma j} \lambda_j - S_i^- = y_{\gamma 0}$$

$$S_i^-, S_{\gamma}^+, \lambda_j \geq 0$$

Where.

S_i^- : 투입의 여유변수

S_{γ}^+ : 투입의 여유변수

λ_j : 투입의 여유변수

ϵ : non-Archimedean 상수

03 | Example/Practice of DEA

체인호텔의 상대적 효율성 측정

•What's BCC :

- ✓ BCC모형의 기본 가정 : 규모수익가변(Variable Return to Scale : VRS) -> CCR모형의 무한확장가능성을 배제하기 위해 새로운 제약식 $\sum_{j=1}^n \lambda_j = 1$ 을 추가 -> 단순 선형 효율성 모델이 아닌 Efficient frontier(효율적 투자선)를 가지게 됨 즉, 볼록한 프론티어(경계)를 가짐

- ✓ 변수 축적에 대한 반환

- Input 방향의 CCR포락모형

$$\text{Min } h_0 = \theta - \epsilon \left(\sum_{i=1}^m S_i^- + \sum_{\gamma=1}^s S_{\gamma}^+ \right) \quad \langle \text{식 2} \rangle$$

$$\text{subject to } \theta x_{i0} - S_i^- - \sum_{j=1}^n x_{ij} \lambda_j = 0$$

$$\sum_{j=1}^n y_{\gamma j} \lambda_j - S_{\gamma}^- = y_{\gamma 0}$$

$$\sum_{j=1}^n \lambda_j = 1$$

$$S_i^-, S_{\gamma}^+, \lambda_j \geq 0$$

03 | Example/Practice of DEA

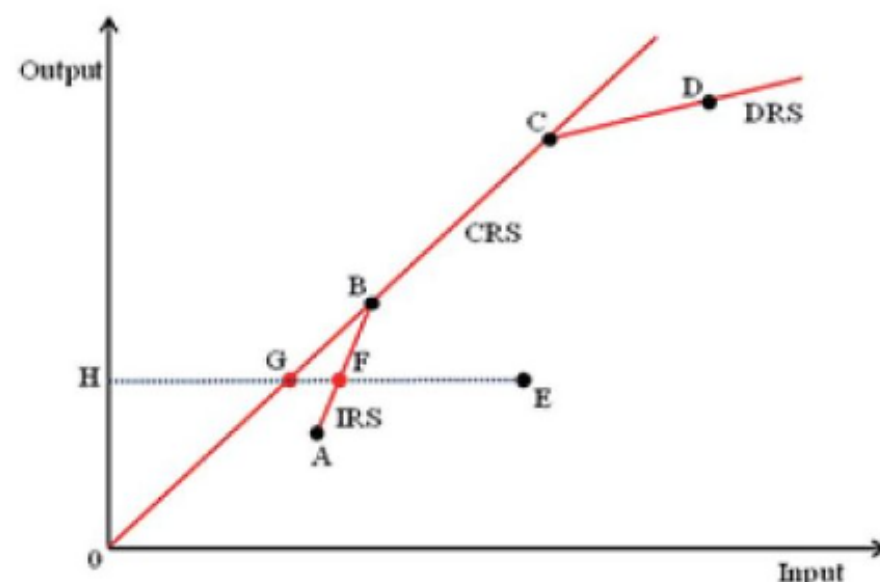
체인호텔의 상대적 효율성 측정

• 위 식의 산출물

- ✓ 기술 효율성 -> 기술적 효율성 = 규모 효율성 * 순수기술 효율성
- ✓ 규모 효율성 = 기술 효율성 / 순수기술 효율성 -> 얼마나 작은 규모 측면에서 효율적으로 투입-산출 활동을 하는가를 측정하기 위함
- ✓ 규모 효율성 = 1 -> 기술 효율성 = 순수기술 효율성 -> 규모 비효율성이 나타나지 않음

$$\text{기술적 효율성} \left(\frac{HE}{HG} \right) = \text{순수기술적 효율성} \left(\frac{HE}{HF} \right) \times \text{규모의 효율성} \left(\frac{HE}{HF} \right)$$

$$\text{규모의 효율성} \left(\frac{HE}{HF} \right) = \text{기술적 효율성} \left(\frac{HE}{HG} \right) / \text{순수기술적 효율성} \left(\frac{HE}{HF} \right)$$



〈그림 1〉 규모의 효율성

*윤지현, 최규원(2008), 회계정보를 이용한 호텔기업의 효율성분석, 관광학연구 참조

-> 기울기가 1보다
큰지 작은지에 따라 달라짐

03 | Example/Practice of DEA

Input-Oriented CCR Model Example

- Tourist hotel

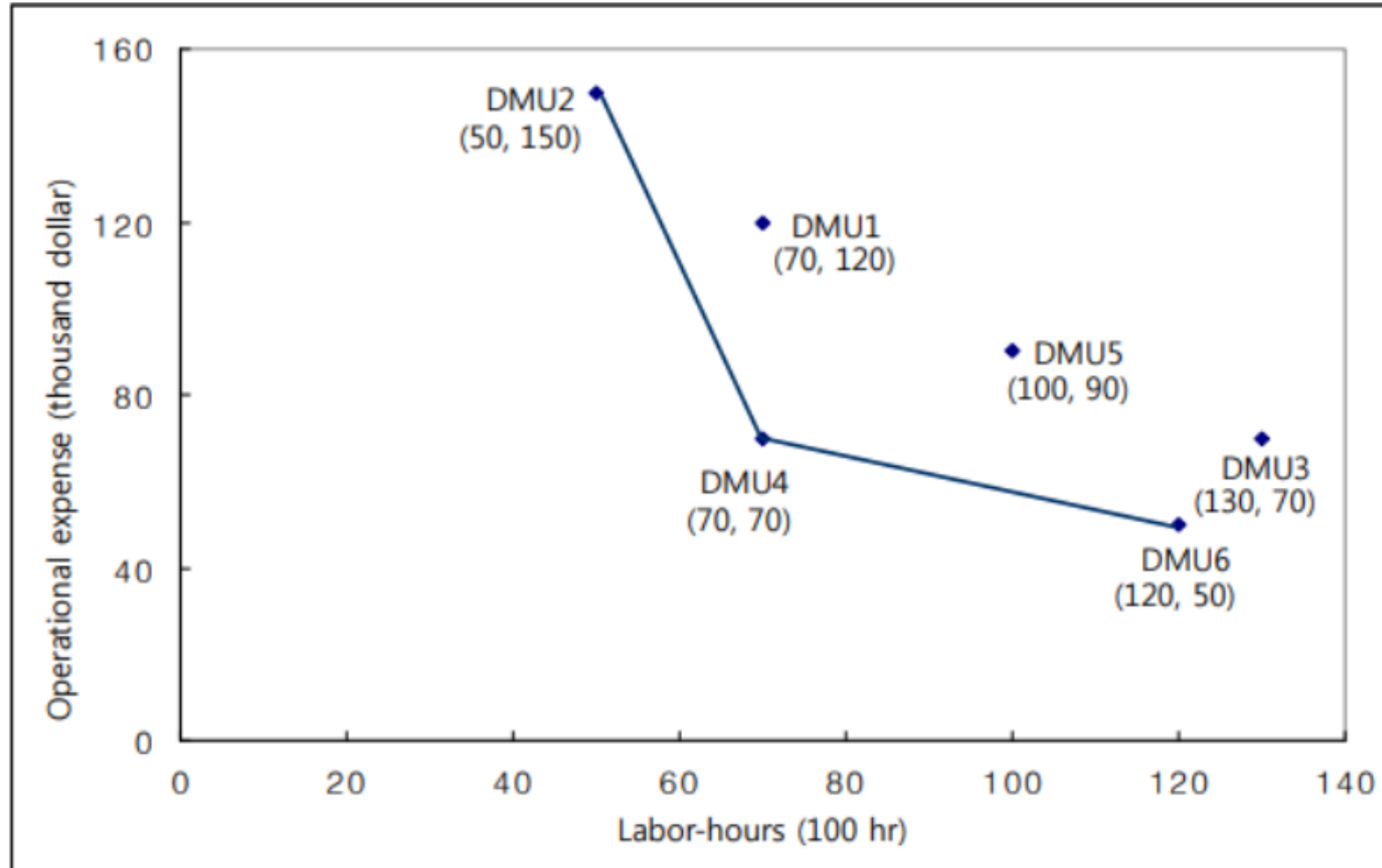
DMUs	Output	Input	
	Number of guests	Labor-hours (100 hr)	Operational cost (thousand dollar)
DMU1	500	70	120
DMU2	500	50	150
DMU3	500	130	70
DMU4	500	70	70
DMU5	500	100	90
DMU6	500	120	50

DMU1~6들의 효율성 계산하기

03 | Example/Practice of DEA

Input-Oriented CCR Model Example

- Efficient frontier plot



Efficient frontier -> 효율적 투자선 그리기
Input 값들을 사용하여 그림

03 | Example/Practice of DEA

Input-Oriented CCR Model Example

- LP form으로 변형 후 최적화

DMU 1) $\max E_1 = 500 u_1$

$s.t. \quad 500u_1 - 70v_1 - 120v_2 \leq 0$

$500u_1 - 50v_1 - 150v_2 \leq 0$

$500u_1 - 130v_1 - 70v_2 \leq 0$

$500u_1 - 70v_1 - 70v_2 \leq 0$

$500u_1 - 100v_1 - 90v_2 \leq 0$

$500u_1 - 120v_1 - 50v_2 \leq 0$

$70v_1 + 120v_2 = 1$

$u_1, v_1, v_2 \geq 0$

DMU 2)

⋮

DMU 6)

DMU 1 ~ 6 들의 효율성을 각각 최적화

→ input, output들을 weight를 곱한 후 최적화 진행

03 | Example/Practice of DEA

Input-Oriented CCR Model Example

• Optimized efficiency result

➤ Results

DMU	Efficiency (E)	Efficiency Reference Set	u_1	v_1	v_2
DMU1	0.8750	DMU2 (0.4375) DMU4 (0.5625)	0.0018	0.0100	0.0025
DMU2	1	---	0.0020	0.0114	0.0029
DMU3	0.8033	DMU4 (0.3115) DMU6 (0.6885)	0.0016	0.0033	0.0082
DMU4	1	---	0.0020	0.0114	0.0029
DMU5	0.7539	DMU4 (0.8923) DMU6 (0.1077)	0.0015	0.0031	0.0077
DMU6	1	---	0.0020	0.0041	0.0102

↓
*DMUs corresponding to the binding constraints
(shadow price)*

앞 장에서의 LP form의 E_n ($n=1,2,...,6$)를 최적화

-> input, output 값들을 활용하여

03 | Example/Practice of DEA

Input-Oriented CCR Model Example

• Interpretation of Result

➤ Improvement - How to make DMU1 efficient (1/2)

▪ Arithmetic calculation

✓ Benchmark any efficient DMU

✓ Reduce each input → increase efficiency ratings by 0.125

– Reduce labor-hours by 12.5($=0.125 \div 0.01$)

– Reduce operational cost by 50($=0.125 \div 0.0025$)

✓ Benchmark HCU(Hypothetical Composite Unit)1

	Reference Set		HCU1	DMU1	Excess Inputs Used
	DMU2 (0.4375)	DMU4 (0.5625)			
Labor-hours	$0.4375 \times 50 +$	$0.5625 \times 70 =$	61.25	70	8.75 ($=70-61.25 = 70 \times 0.125$)
Operational cost	$0.4375 \times 150 +$	$0.5625 \times 70 =$	105	120	15 ($=120-105 = 120 \times 0.125$)

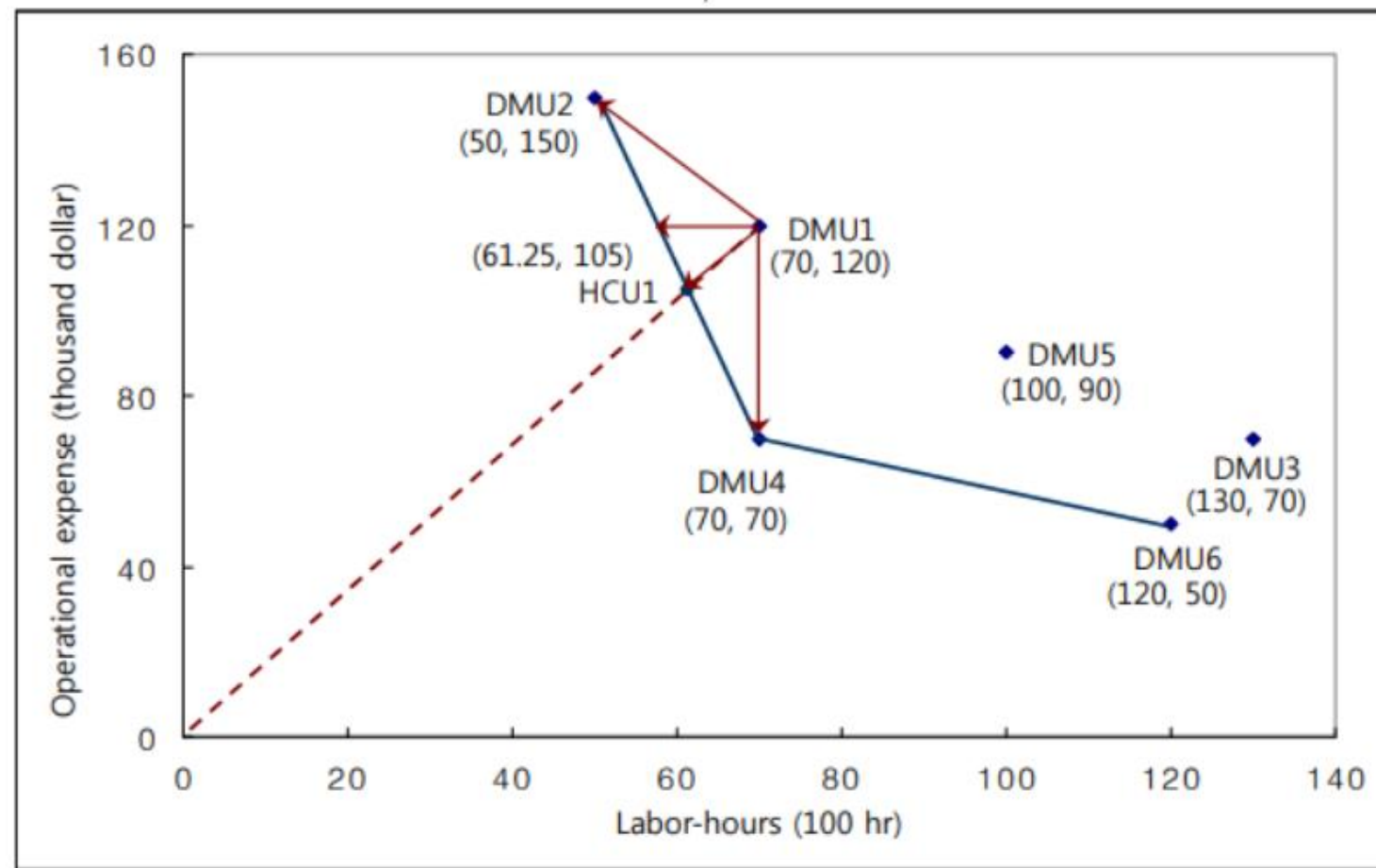
03 | Example/Practice of DEA

Input-Oriented CCR Model Example

• Interpretation of Result

➤ Improvement - How to make DMU1 efficient (2/2)

▪ Graphical representation



➤ Correspondence between MCDM and DEA



- MCDM과 DEA의 연관 관계

MCDM에서의 의사 결정 = DEA에서의 DMUs

MCDM에서의 Positive criteria = DEA에서의 Outputs

MCDM에서의 Negative criteria = DEA에서의 Inputs

MCDM에서의 Priority = DEA에서의 efficiency score

03 | Example/Practice of DEA

Input-Oriented CCR Model Example

- 은행 지점의 성과평가 사례

Branch	Personal transactions ('000s)	Number of staff
Croydon	125	18
Dorking	44	16
Redhill	80	17
Reigate	23	11

– 이 자료로 어떻게 성과를 평가할 수 있을까? -> 상대적 비율

- Ratio

Branch	Personal transactions per staff member ('000s)
Croydon	6.94
Dorking	2.75
Redhill	4.71
Reigate	2.09

간단하지만, DEA 핵심 아이디어의 전부

Branch	Relative efficiency
Croydon	$100(6.94/6.94) = 100\%$
Dorking	$100(2.75/6.94) = 40\%$
Redhill	$100(4.71/6.94) = 68\%$
Reigate	$100(2.09/6.94) = 30\%$

100%의 효율성 - Best?
40%의 효율성 - Not good?
68%의 효율성 - Good?
30%의 효율성 - Not good?

04 | Conclusion

Advantages of DEA for MCDM

- DEA의 장점

- ✓ **Data-oriented** endogenous weighting method
- ✓ Allowing each DMU to choose the **optimal weights of variables** which maximize its efficiency
- ✓ **Usefulness under the context of uncertainty** about, and **lack of consensus** on an appropriate weighting scheme

- 느낀 점

- ✓ 1979년에 발간된 방법론이지만 현대까지도 유용하게 사용되고 성능이 좋다는 것이 놀랍다.
- ✓ 이 보다 더 최근에 발간되고 나온 아이디어 및 방법론은 어떤 것이 있는지 궁금해졌다.



Q & A

감사합니다