Matematica Avanzata

Lez. 15 – **Data Envelopment Analysis**

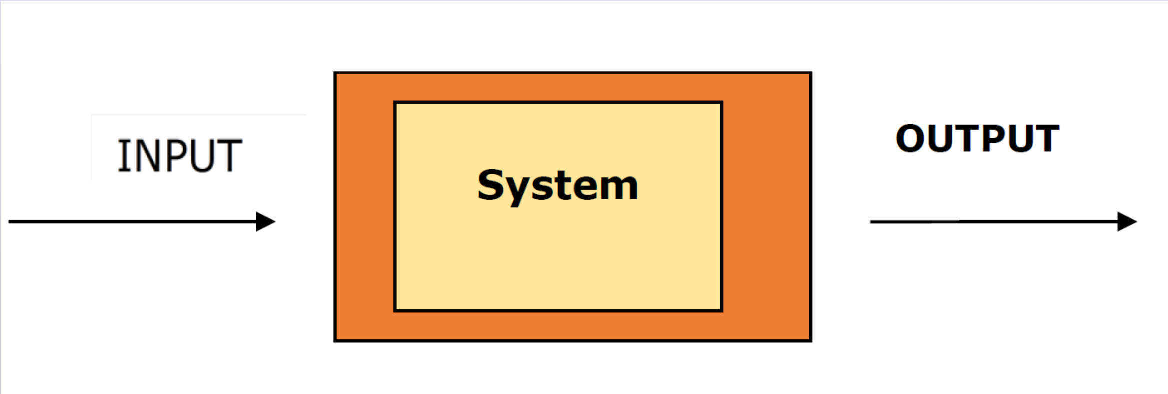
Prof. Antonio VIOLI

Benevento 03/05/2023

Introduzione

In diversi contesti applicativi può emergere l'esigenza di valutare in maniera comparativa l'efficienza di Decision Making Unit (DMU), come aziende, ospedali, dipartimenti universitari, ecc.

Da questo momento rappresenteremo la generica unità come un «sistema» del tipo *black box* che riceve determinati *input* e produce determinati *output*



Introduzione

In maniera intuitiva, l'efficienza può essere considerata come la capacità di gestire «bene» le risorse disponibili.

In maniera più formale, l'efficienza tecnica è definita come il rapporto tra la quantità totale degli output (prodotti, servizi, ecc.) realizzati e la quantità complessiva di input (risorse, tempo, ecc.) utilizzati:

$$E = \frac{\text{Output}}{\text{Input}}$$

L'efficienza può essere migliorata incrementando il livello degli output per un dato livello di input o, viceversa, riducendo il livello di input per un dato livello di output.

Esempio

Consideriamo 8 DMU ognuna caratterizzata da 1 solo input ed 1 solo output, rappresentative di 8 punti vendita di una catena di negozi.

Come input si consideri il numero di impiegati e come output il volume delle vendite.

La dirigenza vuole determinare quale DMU sia efficiente (e quale no) e stabilire quale sia la DMU *benchmark*.

L'obiettivo è quello di suggerire alle DMU non efficienti opportuni aggiustamenti, basati sul confronto con le DMU efficienti.

Branch	A	B	C	D	E	F	G	H
Employees	2	3	3	4	5	5	6	8
Sale volume	1	3	2	3	4	2	3	5

Esempio

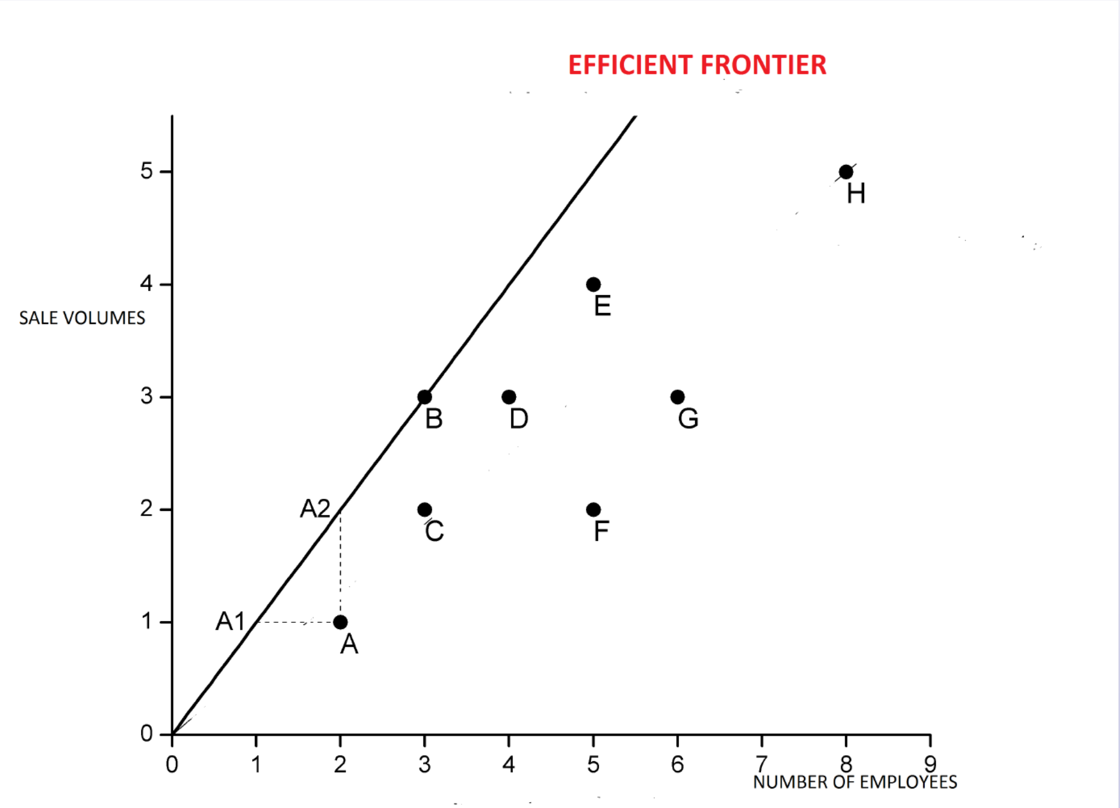
Calcolando l'efficienza come il rapporto tra output e input si determina che la DMU migliore è la ***B***, mentre la meno efficiente è la ***F***

Branch	A	B	C	D	E	F	G	H
Employees	2	3	3	4	5	5	6	8
Sale volume	1	3	2	3	4	2	3	5
Sale/Employee	0.5	1	0.66	0.75	0.8	0.4	0.5	0.625

Esempio

Dal punto di vista grafico, la semiretta che parte dall'origine degli assi e passa per il punto ***B*** costituisce una sorta di «frontiera efficiente».

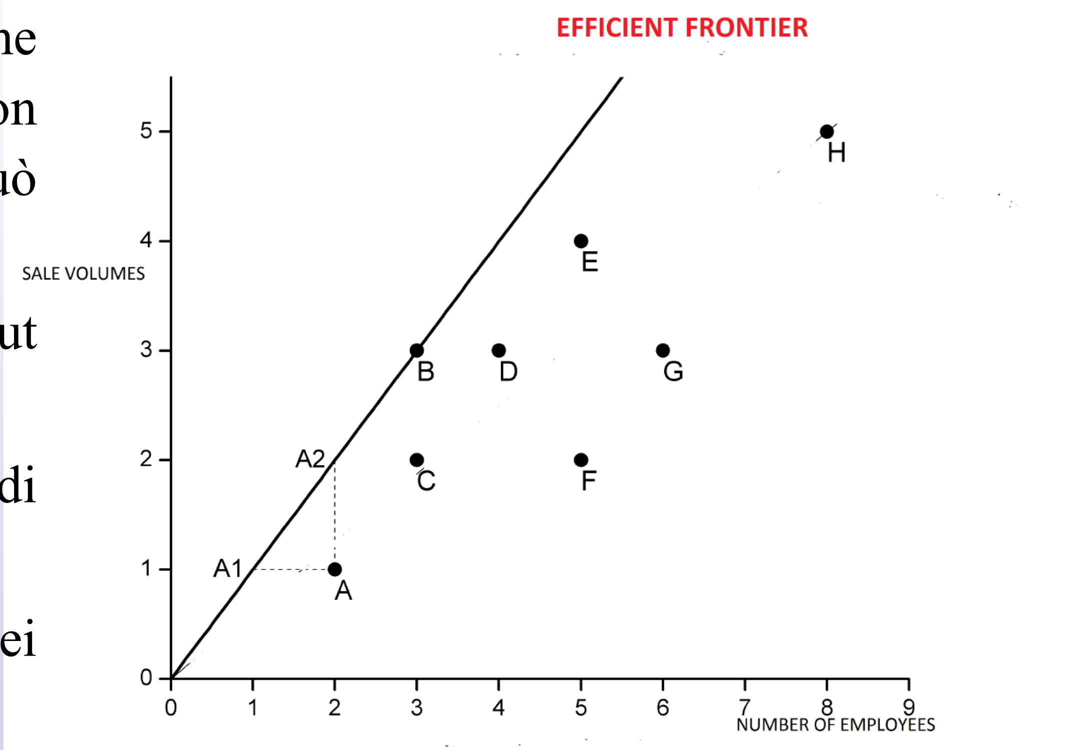
La distanza degli altri punti dalla frontiera misura la distanza delle DMU non efficienti dall'efficienza.



Esempio

Dall'analisi della figura si possono ricavare utili informazioni su come migliorare l'efficienza di DMU non efficienti. Ad esempio, la DMU ***A*** può migliorare:

- Riducendo l'input a parità di output (dal punto **A** al punto **A1**)
- Incrementando l'output a parità di input (dal punto **A** al punto **A2**)
- Più in generale raggiungendo uno dei punti compresi tra **A1** e **A2**



Data Envelopment Analysis

Quando nel processo di valutazione si tiene conto di più input e più output, l'efficienza non è immediatamente valutabile come nell'esempio precedente.

Poiché l'efficienza tecnica è definita come il rapporto tra output e input, in generale si considerano come funzioni di input ed output delle somme pesate. La scelta dei «pesi» associati a ciascun fattore di input e di output è cruciale per la valutazione dell'efficienza.

Data Envelopment Analysis: metodologia non parametrica

La DEA consente di determinare, mediante l'utilizzo di modelli di PL, l'insieme di pesi «ottimale» per ciascuna DMU e una valutazione della sua efficienza relativa.

Data Envelopment Analysis

Si consideri un insieme <u>omogeneo</u> di K DMU. Per ogni DMU k siano noti i valori di m fattori di input e n fattori di output.

Siano:

- x_i^k livello di input i utilizzato dalla DMU k
- y_j^k livello di output j prodotto dalla DMU k

Siano inoltre:

- w_i^k peso associato all'input i per la DMU k
- t_j^k peso associato all'output j per la DMU k

Si assume che tali valori siano non negativi.

Data Envelopment Analysis

L'efficienza per la DMU *k* è definita come

$$E^k = \frac{\sum_{j=1}^{n} y_j^k t_j^k}{\sum_{i=1}^{m} x_i^k w_i^k}$$

Si potrà determinare per ogni DMU il valore dei pesi che massimizzi il suo livello di efficienza.

Data Envelopment Analysis

Per ogni DMU *k* si definisce un modello di ottimizzazione per la determinazione dei pesi ottimali, assumendo che il massimo valore possibile per l'efficienza sia 1.

$$\max \qquad E^k = \frac{\sum_{j=1}^{n} y_j^k t_j^k}{\sum_{i=1}^{m} x_i^k w_i^k}$$

$$\frac{\sum_{j=1}^{n} y_j^l t_j^k}{\sum_{i=1}^{m} x_i^l w_i^k} \leq 1 \quad l = 1, \ldots, K$$

$$t_j^k \geq 0 \quad j = 1, \ldots, n$$

$$w_i^k \geq 0 \quad i = 1, \ldots, m$$

Aspetti critici

- Valori dei pesi
  - I pesi sono definiti come variabili non negative
  - Onde evitare di associare peso nullo a qualche fattore di input e/o di output è preferibile impostare una soglia minima δ
  - Tale scelta assicura che tutti i fattori siano presi in considerazione
- Natura del modello
  - La presenza di un rapporto tra funzioni delle variabili rende il modello <u>non</u> lineare
  - È preferibile utilizzare formulazioni alternative

Modello CCR – Input Oriented

- Formulazione proposta da Charnes, Cooper e Rhodes nel 1978 (CCR)
- Per semplicità si omette l'apice *k* nelle variabili di decisione

$$
\begin{aligned}
\max z = \quad & \sum_{j=1}^{n} y_j^k t_j \\
& \sum_{i=1}^{m} x_i^k w_i = 1 \\
& \sum_{j=1}^{n} y_j^l t_j - \sum_{i=1}^{m} x_i^l w_i \leq 0 \quad l = 1, \ldots, K \\
& w_i \geq \delta \quad i = 1, \ldots, m \\
& t_j \geq \delta \quad j = 1, \ldots, n
\end{aligned}
$$

Modello CCR – Input Oriented

- Il valore ottimo della funzione obiettivo rappresenta il livello di efficienza della DMU *k*. Se è $\underline{E^k = 1}$, allora è <u>efficiente</u>.
- Se la DMU *k* risulta essere <u>non efficiente</u> ($\underline{E^k < 1}$)è possibile individuare l'insieme di riferimento (*Reference Set*), ossia l'insieme delle DMU risultate efficienti con la combinazione di pesi ottimale per *k*.
- Tale insieme contiene i *benchmark* per la DMU *k* (inefficiente)

$$RF(k) = \left\{ l \mid \sum_{j=1}^{n} y_j^l t_j^* - \sum_{i=1}^{m} x_i^l w_i^* = 0 \right\}$$

Modello CCR – Output Oriented

- Formulazione alternativa
- L'efficienza è misurata dal reciproco della f.o.

$$\min z = \sum_{i=1}^{m} x_i^k w_i$$

$$\sum_{j=1}^{n} y_j^k t_j = 1$$

$$\sum_{j=1}^{n} y_j^l t_j - \sum_{i=1}^{m} x_i^l w_i \leq 0 \quad l = 1, \ldots, K$$

$$w_i \geq \delta \quad i = 1, \ldots, m$$

$$t_j \geq \delta \quad j = 1, \ldots, n$$

Esempio

L'azienda SuperLab produce e commercializza apparecchi elettronici in tutto il mondo, mediante 5 divisioni territoriali (Nord America, Sud America, Europa, Asia, Africa), ognuna delle quali costituisce una DMU. I prodotti di punta sono smartphone, tablet e laptop. All'inizio del 2020 il management intende valutare l'efficienza delle divisioni territoriali mediante l'utilizzo della DEA, utilizzando come input gli investimenti nel *Product Development* e nel *Marketing* e come output il volume di vendita di ciascun prodotto.

La tabella che segue riporta i valori misurati nell'anno precedente per input e output.

Esempio

	INPUTS($ in millions)		OUTPUTS (numbers in '000)		
DMU	Product Development	Marketing	Smart Phones	Tablets	Laptops
N. America	5	14	9	4	16
Europe	10	18	8	2	9
Asia	9	16	9	4	10
Africa	7	12	6	1	8
S. America	9	15	10	4	14

Esempio

Per ogni DMU è possibile formulare e risolvere un modello *ad hoc*, avente come variabili di decisione i pesi ottimali per la DMU in esame.

Per la divisione Nord America il modello è il seguente:

$$
\begin{aligned}
\max z \quad & = 9t_1 + 4t_2 + 16t_3 \\
& 5w_1 + 14w_2 = 1 \\
& 9t_1 + 4t_2 + 16t_3 - 5w_1 - 14w_2 \leq 0 \\
& 8t_1 + 2t_2 + 9t_3 - 10w_1 - 18w_2 \leq 0 \\
& 9t_1 + 4t_2 + 10t_3 - 9w_1 - 16w_2 \leq 0 \\
& 6t_1 + 1t_2 + 8t_3 - 7w_1 - 12w_2 \leq 0 \\
& 10t_1 + 4t_2 + 14t_3 - 9w_1 - 15w_2 \leq 0 \\
& w_i \geq \delta \quad i = 1, 2 \\
& t_j \geq \delta \quad j = 1, 2, 3
\end{aligned}
$$

Esempio

Il modello per la divisione Europa è il seguente:

$$
\begin{aligned}
\max z \quad & = 8t_1 + 2t_2 + 9t_3 \\
& 10w_1 + 18w_2 = 1 \\
& 9t_1 + 4t_2 + 16t_3 - 5w_1 - 14w_2 \leq 0 \\
& 8t_1 + 2t_2 + 9t_3 - 10w_1 - 18w_2 \leq 0 \\
& 9t_1 + 4t_2 + 10t_3 - 9w_1 - 16w_2 \leq 0 \\
& 6t_1 + 1t_2 + 8t_3 - 7w_1 - 12w_2 \leq 0 \\
& 10t_1 + 4t_2 + 14t_3 - 9w_1 - 15w_2 \leq 0 \\
& w_i \geq \delta \quad i = 1, 2 \\
& t_j \geq \delta \quad j = 1, 2, 3
\end{aligned}
$$

Esempio

DMU	Efficiency
N. America	1
Europe	0.67
Asia	0.875
Africa	0.75
S. America	1

Dall’analisi dei risultati solo 2 DMU risultano efficienti.

È possibile avere indicazioni su come intervenire sulle DMU non efficienti?

Come migliorare l'efficienza

- Per ottenere utili informazioni finalizzate al miglioramento dell'efficienza, possiamo considerare il modello duale.
- Riscriviamo il modello CCR input oriented

$$
\begin{aligned}
\max z = \quad & \sum_{j=1}^{n} y_j^k t_j \\
& \sum_{i=1}^{m} x_i^k w_i = 1 \qquad (\theta) \\
& \sum_{j=1}^{n} y_j^l t_j - \sum_{i=1}^{m} x_i^l w_i \leq 0 \quad \forall l \quad (\lambda^l) \\
& -w_i \leq -\delta \quad \forall i \quad (\sigma_i^-) \\
& -t_j \leq -\delta \quad \forall j \quad (\sigma_j^+)
\end{aligned}
$$

Modello CCR input oriented Duale

$$\min \quad z = \theta - \delta\left(\sum_{i=1}^{m} \sigma_i^- + \sum_{j=1}^{n} \sigma_j^+\right)$$

$$\sum_{l=1}^{K} y_j^l \lambda^l - \sigma_j^+ = y_j^k \quad \forall j$$

$$\sum_{l=1}^{K} x_i^l \lambda^l + \sigma_i^- = \theta x_i^k \quad \forall i$$

$$\lambda^l \geq 0 \quad \forall l$$

$$\sigma_j^+ \geq 0 \qquad \forall j$$

$$\sigma_i^- \geq 0 \qquad \forall i$$

$$\theta \qquad \text{free}$$

Osservazioni

- Il modello primale ha $(m+n+K+1)$ vincoli, mentre il modello duale ne ha $(m+n)$.
- Solitamente il numero di DMU osservate K è di molto superiore rispetto a $(m+n)$, per cui la risoluzione del problema duale è più semplice.
- Le variabili del duale λ' sono associate ai vincoli sul valore di efficienza delle singole DMU.
- Dalla teoria della dualità, se un vincolo del primale è soddisfatto per uguaglianza, la corrispondente variabile del duale λ' è positiva, altrimenti è pari a 0 (vedi condizioni di complementarietà).

Osservazioni

- Se nella soluzione del primale uno dei vincoli sul valore di efficienza, associato ad una DMU diversa da quella che si sta valutando, risulta soddisfatto per uguaglianza, allora la DMU in esame sarà inefficiente.
- Le variabili del duale λ' positive sono pertanto associate alle DMU efficienti, diverse da quella in esame, ossia il Reference Set.
- Tale insieme contiene i *benchmark* per la DMU *k* (inefficiente).
- Riepilogando:
  - Se $z^* = 1$ (ossia $\theta^* = 1, \sigma_j^{+*} = 0, \sigma_j^{-*} = 0$), la DMU è efficiente.
  - Se $z^* < 1$ (ossia se esiste qualche variabile σ_j^{+*} o σ_j^{-*} strettamente positiva), la DMU è inefficiente. Si può individuare il Reference Set per migliorare l'efficienza

Osservazioni

- In quest'ultimo caso (qualche variabile σ_j^{+*} o σ_j^{-*} strettamente positiva), dall'analisi dei vincoli del problema duale è possibile trarre indicazioni su come migliorare l'efficienza.

$$\sum_{l=1}^{K} y_j^l \lambda^{l*} = y_j^k + \sigma_j^{+*} \quad \forall j$$

$$\sum_{l=1}^{K} x_i^l \lambda^{l*} = \theta^* x_i^k - \sigma_i^{-*} \quad \forall i$$

Osservazioni

- L'idea è quella di creare una DMU «ideale» che risulti migliore della DMU (inefficiente) in esame.
- Si utilizzano a tal proposito le DMU del Reference Set, ossia quelle associate a valori positivi di λ'.
- La nuova DMU ideale utilizzerà per ogni fattore di input i una quantità non maggiore di quella utilizzata dalla DMU in esame.
- Analogamente, produrrà per ogni fattore di output j una quantità non inferiore a quella prodotta dalla DMU in esame.

Miglioramento dell'efficienza

- Per rendere efficiente una DMU non efficiente si può pertanto intervenire sui livelli dei fattori di input e di output nel seguente modo

  - **Radial input reduction**

$$x_i^k \Longrightarrow \quad \theta^* x_i^k - \sigma_i^{-*}$$

  - **Output increase**

$$y_j^k \Longrightarrow \quad y_j^k + \sigma_j^{+*}$$

- Tali correzioni equivalgono a «spostare» la DMU in esame sulla frontiera efficiente.

Esempio

- Come migliorare l'efficienza della divisione Europa?

DMU	Efficiency
N. America	1
Europe	0.67
Asia	0.875
Africa	0.75
S. America	1

- Dalla soluzione del problema duale si ottiene:

$$\lambda_1 = 0.158 \qquad \lambda_5 = 0.658$$

Esempio

- I valori delle variabili λ rappresentano i pesi da utilizzare per costruire la DMU ideale per la DMU Europa.
- La soluzione del duale fornisce anche indicazioni su quanto migliorare i valori dei singoli fattori di input e output.

Esempio

- Per gli output si registra

$$\sigma_2^+ = 1.263 \quad \sigma_3^+ = 2.737$$

- I nuovi livelli saranno

$$y_1^2 = 8$$

$$y_2^2 = 2 + 1.263$$

$$y_3^2 = 9 + 2.737$$

Esempio

- Per gli input, visto che tutte le variabili $\sigma_j^{-*} = 0$, i livelli di input possono essere ridotti (radialmente) mediante la seguente espressione:

$$x_i^2 = \theta^* * x_i^2$$

dove $\theta^* = 0.671$.

Esempio

- Per risultati completi vedi file GAMS

Modello CCR output oriented duale

$$\max \quad z = \phi + \delta\left(\sum_{i=1}^{m} \sigma_i^- + \sum_{j=1}^{n} \sigma_j^+\right)$$

$$\sum_{l=1}^{K} y_j^l \lambda^l = \phi y_j^k + \sigma_j^+ \quad \forall j$$

$$\sum_{l=1}^{K} x_i^l \lambda^l = x_i^k - \sigma_i^- \quad \forall i$$

$$\lambda^l \geq 0 \quad \forall l$$

$$\sigma_j^+ \geq 0 \qquad \forall j$$

$$\sigma_i^- \geq 0 \qquad \forall i$$

$$\phi \qquad \text{free}$$

Rendimenti di scala

- Limite del modello CCR: rendimenti di scala costanti
- Si assume che ad un incremento degli input utilizzati corrisponda un incremento proporzionale degli output prodotti
- Tale ipotesi non è sempre verosimile: in numerosi contesti applicativi si assiste al fenomeno dell’economia (o diseconomia) di scala.
- Pertanto, è plausibile che gli output crescano in modo più (o meno) che proporzionale rispetto all’aumento degli input.
- L’ipotesi di rendimenti di scala variabili è più generale e comporta un minor numero di DMU efficienti e valori di efficienza più bassi.
- Modello BCC (Banker, Charnes & Cooper, 1984): consente di modellare rendimenti di scala variabili

$$E^k = \frac{\sum_{j=1}^{n} y_j^k t_j^k + \mu}{\sum_{i=1}^{m} x_i^k w_i^k}$$

Modello BCC input oriented

$$\max z = \sum_{j=1}^{n} y_j^k t_j + \mu$$

$$\sum_{i=1}^{m} x_i^k w_i = 1$$

$$\sum_{j=1}^{n} y_j^l t_j + \mu - \sum_{i=1}^{m} x_i^l w_i \le 0 \quad l = 1, \ldots, K$$

$$w_i \ge \delta \quad i = 1, \ldots, m$$

$$t_j \ge \delta \quad j = 1, \ldots, n$$

$$\mu \quad \text{free}$$

Modello BCC input oriented - duale

$$\min \quad z = \theta - \delta\left(\sum_{i=1}^{m} \sigma_i^- + \sum_{j=1}^{n} \sigma_j^+\right)$$

$$\sum_{l=1}^{K} y_j^l \lambda^l - \sigma_j^+ = y_j^k \quad \forall j$$

$$\sum_{l=1}^{K} x_i^l \lambda^l + \sigma_i^- = \theta x_i^k \quad \forall i$$

$$\sum_{l=1}^{K} \lambda^l = 1$$

$$\lambda^l \geq 0 \quad \forall l$$

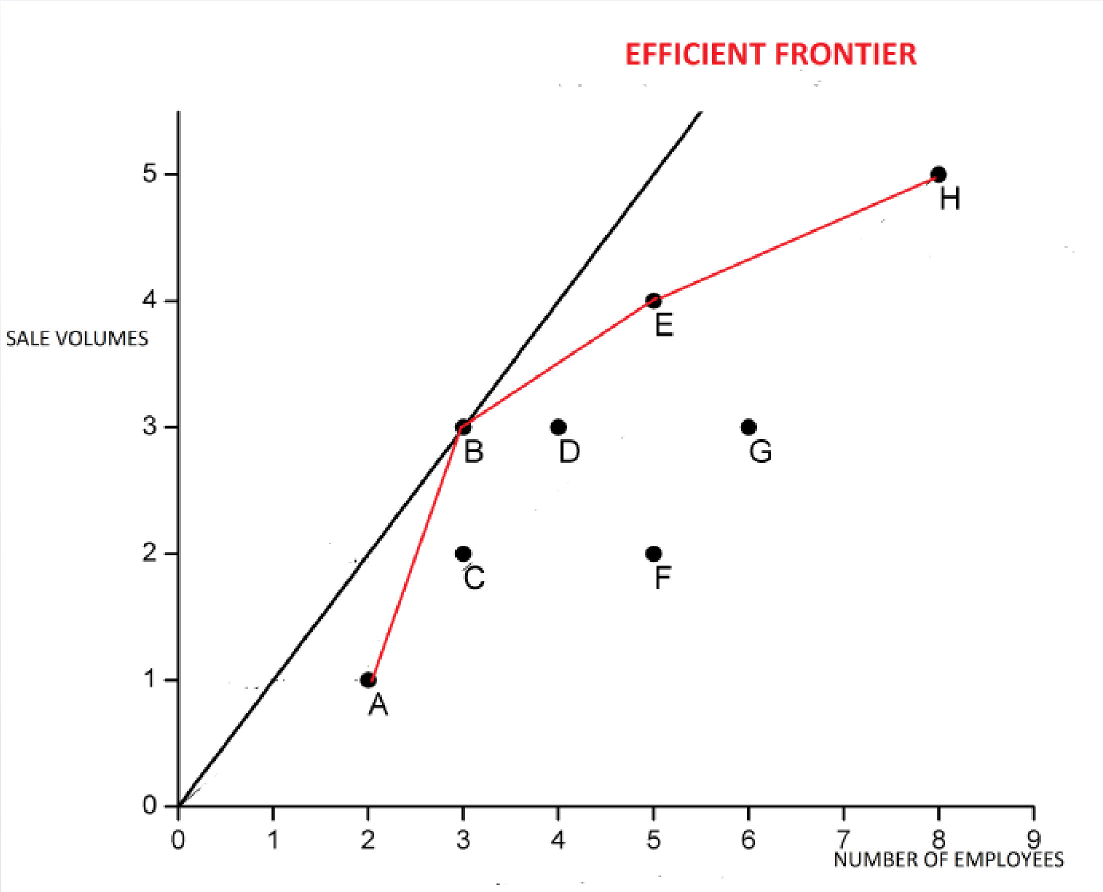
$$\sigma_j^+ \geq 0 \qquad \forall j$$

$$\sigma_i^- \geq 0 \qquad \forall i$$

$$\theta \quad \text{free}$$

Rendimenti di scala

- Si osserva che la regione ammissibile di un modello BCC è più piccola di quella di un modello CCR
- Un maggior numero di DMU risultano efficienti



Approfondimenti

- http://www.deazone.com/