VIII Congreso de Ingeniería de Organización
Leganés, 9 y 10 de septiembre de 2004

Resolución de problemas DEA con entradas y salidas enteras

Gabriel Villa Caro, Jesús Racero Moreno, José Manuel García Sánchez y Marcos Calle Suárez

Área de Ingeniería de Organización. Escuela Superior de Ingenieros. Universidad de Sevilla. Avenida de los Descubrimientos, s/n. 41092. Sevilla. gvilla@esi.us.es, jrm@esi.us.es, jmgs@esi.us.es, mcalle@esi.us.es

Resumen

Este artículo estudia un escenario DEA cuyas entradas y salidas deben ser números enteros. Los modelos DEA convencionales proyectan las unidades productivas sobre puntos de operación (objetivos) que en general no respetan este tipo de restricciones. El siguiente método que se plantea resuelve este problema mediante un modelo mixto de programación entera que garantiza que los objetivos a los que tienden las unidades productivas sean números enteros.

Palabras clave: análisis por envoltura de datos, valores enteros, análisis de mejora.

1. Introducción

El análisis por envoltura de datos DEA es una conocida herramienta de programación lineal que mide la eficiencia de unidades productivas que fabrican de forma similar (DMUs). Para un análisis más extenso de DEA, se puede consultar Cooper et al (2000) y Thanassoulis (2000). Uno de los puntos fuertes de la herramienta es su carácter no paramétrico, ya que sólo es necesario conocer la cantidad de recursos que cada unidad consume (entradas) y la cantidad de producción que cada unidad genera (salidas) para medir su eficiencia relativa. Esto se consigue extrapolando de las entradas y salidas observadas un conjunto de puntos de operación admisibles que definen la tecnología. Entre las más frecuentes se encuentran la de retornos de escala constantes (CRS) y la de retornos de escala variables (VRS). Ambas consideran puntos admisibles a las combinaciones lineales de las entradas y salidas de las DMUs existentes.

Sin embargo es una situación muy habitual en los problemas DEA que algunas de las entradas (número de empleados) y/o salidas (número de clientes) sean números enteros. En estos casos, el usar las tecnologías arriba descritas puede ocasionar que algunos de los puntos eficientes sobre los que deben proyectarse las unidades existentes resulten tener para dichas entradas y salidas números fraccionarios (i.e. la solución obtenida podría indicar que habría que reducir los empleados hasta 3,45). En los casos en que los costes unitarios sean altos no es conveniente utilizar una técnica heurística de redondeo al alza o a la baja.

En el siguiente apartado se introduce un nuevo modelo DEA con variables enteras que resuelve para la tecnología CRS el caso en el que se tienen algunas entradas y/o salidas enteras. En el apartado 3 se realiza una aplicación recogida en la bibliografía, analizándose los resultados obtenidos. Por último se incluye un apartado de conclusiones del artículo.

2. Modelo de maximización de beneficios

Sean I = {1,2,...,m} y O = {1,2,…,p} los conjuntos de entradas y salidas de un problema DEA respectivamente, y sean I'⊂ I y O'⊂ O los subconjuntos de las correspondientes dimensiones que deben ser valores enteros. Evidentemente, todo los valores de x_{ij} e y_{kj} observados deben ser enteros ∀i∈ I' y ∀k∈ O'. Si asumimos tecnología CRS en el problema, se puede definir un conjunto de posibilidades de valores enteros CRS de la siguiente forma:

$$T'_{CRS} = \left\{ \begin{array}{l} (\hat{x},\hat{y}) : \exists (\lambda_1,\lambda_2,...,\lambda_n)\ \lambda_j \geq 0\ \forall j \quad \hat{x}_i \geq \sum_j \lambda_j x_{ij} \quad \hat{y}_k \leq \sum_j \lambda_j y_{kj} \\ \hat{x}_i \text{ entero } \forall i \in I' \\ \hat{y}_k \text{ entero } \forall k \in O' \end{array} \right\} \tag{1}$$

Definición: Una DMU_J es CRS eficiente-entero si ningún otro punto de operación con valores enteros le domina, i.e. si.

$$\forall (\hat{x},\hat{y}) \in T'_{CRS}\ \left[\hat{x} \leq \overline{x}_J\right] \cap \left[\hat{y} \geq \overline{y}_J\right] \Leftrightarrow (\hat{x},\hat{y}) = (\overline{x}_J,\overline{y}_J) \tag{2}$$

Definición: La frontera de la tecnología CRS eficiente-entero es el conjunto de los puntos de operación enteros no dominados, i.e.

$$\left(T'_{CRS}\right)^{ef} = \left\{ (\hat{x},\hat{y}) \in T'_{CRS} : \forall (\overline{x},\overline{y}) \in T'_{CRS} \quad \left[\overline{x} \leq \hat{x}\right] \cap \left[\overline{y} \geq \hat{y}\right] \Leftrightarrow (\overline{x},\overline{y}) = (\hat{x},\hat{y}) \right\} \subset T'_{CRS} \tag{3}$$

Proposición: si una DMU_J existente es CRS eficiente, entonces también es CRS eficiente-entero, i.e.

$$\left(\overline{x}_J,\overline{y}_J\right) \in T^{ef}_{CRS} \quad \Rightarrow \quad \left(\overline{x}_J,\overline{y}_J\right) \in \left(T'_{CRS}\right)^{ef} \tag{4}$$

Demostración: es una consecuencia directa de las definiciones de ambos conjuntos y del hecho de que:

a) Las DMUs existentes son valores enteros y pertenecen por tanto a T'_{CRS}.

b) El conjunto de posibilidades de producción de valores enteros CRS está incluido en la tecnología CRS□

Hay que hacer notar que, aunque las DMUs CRS eficientes son también CRS eficiente-entero, lo contrario no es cierto. Es decir, puede haber unidades CRS eficiente-entero que cuando son relajadas las restricciones de integridad podría ser dominada por algún otro punto de operación.

El siguiente modelo MILP mide la mejora que pueden desarrollar cada una de las unidades existentes en esta nueva tecnología.

$$
\begin{aligned}
&\text{Min} \quad \theta_J \; - \; \varepsilon \, (\sum_i s_i^- \; + \; \sum_k s_k^+ \,) \\
&\text{s.a.} \\
&\sum_j \lambda_j x_{ij} = x_i \qquad \forall i \\
&x_i \; = \; \theta_J \, x_{iJ} \; - \; s_i^- \qquad \forall i \\
&\sum_j \lambda_j \, y_{kj} = y_k \qquad \forall k \\
&y_k \; = \; y_{kJ} \; + \; s_k^+ \qquad \forall k \\
&\lambda_j \geq 0 \;\; \forall j \quad s_i^- \, , x_i \geq 0 \;\; \forall i \quad s_k^+ \, , y_k \geq 0 \;\; \forall k \\
&\theta_J \text{ libre} \quad x_i \text{ entera } \forall i \in I' \quad y_k \text{ entera } \forall k \in O'
\end{aligned}
\tag{5}
$$

Este modelo tiene n+2*(m+p)+1 variables, de las cuales una es libre y |I'| + |O'| son enteras. El número de restricciones es de 2*(m+p). Es el típico modelo de dos fases con orientación de entrada que busca la máxima reducción radial de todas las entradas y la subsiguiente apuración de entradas y salidas mediante la maximización de las holguras. Lo que lo hace diferente es la integridad de las restricciones impuestas sobre las dimensiones de entrada y salida que son enteras. De la solución óptima del modelo resulta:

a) La reducción radial CRS entera θ_J^*.

b) Las holguras adicionales que resultan s_i^{-*}, $\forall i$; s_k^{+*}, $\forall k$.

c) Un punto de operación objetivo con dimensiones enteras $(x^*_1,\ldots,x^*_m,y^*_1,\ldots,y^*_p)$.

Proposición: Una DMU$_J$ existente es eficiente-entera si y sólo si $\theta_J^*=1$, $s_i^{-*}=0$, $\forall i$, $s_k^{+*}0$, $\forall k$.

Demostración: probaremos que es condición necesaria y suficiente.

Respecto a la necesaria, hay que hacer notar que si $\theta_J^*<1$, ó $\exists s_i^{-*}>0$, ó $\exists s_k^{+*}>0$, entonces el punto de operación objetivo obtenido pertenecería a T'_{CRS}, no coincidiría con DMU$_J$ pero le dominaría, con lo que por definición no sería CRS eficiente-entero.

Que es condición suficiente se demuestra por reducción al absurdo, asumiendo que la proposición no se cumple, y viendo que se llega a una contradicción. Si existe una DMU$_J$ que cumple que $\theta_J^*=1$, $s_i^{-*}=0$, $\forall i$, $s_k^{+*}0$, $\forall k$ y no es CRS eficiente-entero, por definición debe haber algún punto de operación entero $(\hat{x},\hat{y}) \in T'_{CRS}$ con $(\hat{x},\hat{y}) \neq (\overline{x}_J, \overline{y}_J)$ que lo domine, i.e.

$$\hat{x} \leq \overline{x}_J \;\text{ e }\; \hat{y} \geq \overline{y}_J \tag{6}$$

Como $(\hat{x},\hat{y}) \in T'_{CRS}$, debe existir un vector de landas $(\lambda_1, \lambda_2,\ldots,\lambda_n)$ con $\lambda_j \geq 0$, $\forall j$ tal que:

$$\hat{x}_i \geq \sum_j \lambda_j x_{ij} \;\; \forall i \;\text{ y }\; \hat{y}_k \leq \sum_j \lambda_j y_{kj} \;\; \forall k \tag{7}$$

Como se cumple la condición (6), debe haber un vector $(\hat{\lambda}_1,\hat{\lambda}_2,...,\hat{\lambda}_n)$ $\hat{\lambda}_j \geq 0\ \forall\, j$ tal que:

$$\hat{x}_i = \sum_j \hat{\lambda}_j x_{ij} \quad \forall i \quad y \quad \hat{y}_k = \sum_j \hat{\lambda}_j y_{kj} \quad \forall k \tag{8}$$

La solución al modelo propuesto es:

$$\begin{aligned} \hat{x}_i &= \hat{\theta}_J\, x_{iJ} - \hat{s}_i^- \leq x_{iJ} \quad \forall i \\ \hat{y}_k &= y_{kJ} + \hat{s}_k^+ \geq y_{kJ} \quad \forall k \end{aligned} \tag{9}$$

Como $(\hat{x},\hat{y}) \in T'_{CRS}$, debe ser un valor entero en las dimensiones I' y O', y como $(\hat{x},\hat{y}) \neq (\overline{x}_J,\overline{y}_J)$, al menos una de las desigualdades debe ser estricta, es decir, ó $\theta_J^*<1$, y/ó $\exists s_i^{-*}>0$, y/ó $\exists s_k^{+*}>0$, lo que es seguro es que:

$$\hat{\theta}_J - \varepsilon\left(\sum_i \hat{s}_i^- + \sum_k \hat{s}_k^+\right) < \theta_J^* - \varepsilon\left(\sum_i \left(\hat{s}_i^-\right)^* + \sum_k \left(\hat{s}_k^+\right)^*\right) = 1 \tag{10}$$

Lo que significa que hemos encontrado una solución admisible para el modelo propuesto que posee una mejor función objetivo que la supuestamente solución óptima, llegándose a una contradicción.

Proposición: El valor de eficiencia proporcionada por el modelo CRS entero nunca es menor que el proporcionado por el modelo CRS tradicional, i.e. $\theta_J^*<(\theta_J^*)^{CRS}$.

Demostración: Es una consecuencia directa del hecho de que el conjunto de posibilidades de producción CRS con valores enteros está contenido en el conjunto CRS, $T'_{CRS} \subset T_{CRS}$. Por lo tanto, la región de admisibilidad del modelo propuesto está incluida en el proporcionado por el CCR-I. Por lo tanto, como CCR-I es una relajación del modelo DEA MILP propuesto, el valor de su función objetivo debe ser menor que la obtenida por el modelo DEA MILP□

3. Ilustración

Se ha usado un caso encontrado en la bibliografía, concretamente en Shafer y Bradford (1995), donde se mide la eficiencia relativa de 47 configuraciones de producción celular. Se consideran como recursos el número de trabajadores y de máquinas y como salidas una medida de la media del producto en curso ("work in process, WIP"), tiempo medio que pasan las piezas en el sistema de producción ("flow time") y una medida media del nivel de utilización de mano de obra ("worker utilization"). Las salidas fueron obtenidas mediante simulación, y en el caso de las dos primeras se optó por utilizar su inversa para que aumentar la magnitud implicara una mejora en el desempeño de la unidad. Más concretamente, los valores usados de estas salidas son las inversas multiplicadas por 10^{-5} y 10^{-3} respectivamente.

En este caso son las entradas las que hay que considerar como enteras. Esta aplicación fue resuelta con un modelo CCR-I tradicional admitiendo valores fraccionarios en los recursos procediendo posteriormente a un redondeo heurístico. Esta técnica no la consideramos adecuada debido al elevado coste de los recursos. En este artículo se soluciona el problema con el modelo propuesto.

En la Tabla 1 se muestran los resultados obtenidos mediante el modelo CCR-I tradicional.

Tabla 1. Resultados del modelo CCR-I tradicional

DMU	Eficiencia	Puntos de operación objetivo				
		# trabajadores	# máquinas	WIP^{-1}	Flow Time^{-1}	Utilización
1	0.939	9.88	17.84	451.06	151.79	37.47
2	0.744	6.29	14.88	97.95	37.43	39.04
3	0.831	9.02	16.62	381.53	129.22	37.67
4	0.943	12.16	18.86	505.56	169.63	35.54
5	0.711	5.90	15.64	6.60	8.39	43.08
6	0.862	10.36	18.97	448.03	151.45	42.11
7	0.938	12.97	20.64	551.92	185.19	39.40
8	0.850	14.03	18.70	508.65	170.65	32.20
9	**1.000**	**6**	**24**	**1.93**	**2.72**	**63.55**
10	**1.000**	**9**	**24**	**3.53**	**10.75**	**66.23**
11	0.750	6.83	18.01	12.43	11.21	49.49
12	0.883	7.99	21.20	8.01	11.07	58.41
13	0.976	10.61	23.43	250.63	89.91	59.42
14	**1.000**	**13**	**24**	**546.45**	**185.19**	**54.72**
15	0.960	12.60	23.04	549.45	185.61	50.74
16	0.948	22.75	22.75	639.79	214.59	30.42
17	0.816	5.71	20.40	1.97	3.95	54.56
18	0.793	7.52	19.83	14.32	12.55	54.48
19	0.853	8.04	21.34	8.70	11.34	58.76
20	0.909	9.68	22.73	161.03	60.84	59.37
21	0.925	11.65	23.12	408.16	140.40	55.21
22	**1.000**	**14**	**25**	**662.25**	**222.22**	**50.36**
23	0.955	13.74	23.89	634.08	212.77	47.55
24	0.707	14.40	17.67	484.73	162.61	28.66
25	0.610	4.88	15.87	1.95	4.31	42.83
26	0.697	6.89	18.12	16.21	12.47	49.71
27	0.753	7.39	19.57	9.09	10.76	53.88
28	0.864	11.10	22.48	366.30	126.70	54.33
29	0.860	12.03	22.37	497.51	168.81	51.25
30	0.875	12.62	22.75	578.03	194.46	47.57
31	0.871	13.43	22.64	602.41	202.13	44.44
32	0.854	22.20	22.20	624.22	209.37	29.68
33	0.584	5.26	15.77	1.97	5.31	42.92
34	0.614	6.28	16.58	10.66	10.07	45.59
35	0.676	6.88	18.24	8.19	9.94	50.22
36	0.725	7.43	19.57	14.59	12.53	53.76
37	0.805	11.93	21.74	526.31	177.62	47.32
38	0.854	13.66	23.05	613.49	205.85	45.28
39	0.744	11.14	20.10	510.21	171.65	42.06
40	0.801	21.64	21.64	608.45	204.08	28.93
41	0.559	5.59	16.22	3.93	6.50	44.27
42	0.530	5.83	15.37	11.46	9.84	42.23
43	0.582	6.38	16.89	8.68	9.55	46.47
44	0.621	7.52	18.01	107.41	41.73	47.47
45	0.611	8.13	17.72	204.08	72.67	44.62
46	0.729	12.40	21.15	562.30	188.68	41.75
47	**1.000**	**29**	**24**	**693.48**	**232.56**	**24.39**

Hay que hacer notar que, excepto para las unidades eficientes (9, 10, 14, 22 y 47), el resto tiene recursos fraccionarios. Respecto a los resultados que proporciona el modelo propuesto, expuestos en la Tabla 2:

Tabla 2. Resultados del modelo CRS entero

DMU	Eficiencia	Puntos de operación objetivo				
		# trabajadores	# máquinas	WIP^{-1}	Flow $Time^{-1}$	Utilización
1	0.947	10	18	459.83	154.64	37.47
2	0.750	6	15	98.81	36.91	39.04
3	0.850	10	17	396.86	133.99	37.67
4	0.950	12	19	508.35	170.57	36.18
5	0.727	6	16	40.84	18.82	43.08
6	0.909	11	20	499.38	168.07	42.11
7	0.955	13	21	560.89	188.20	40.39
8	0.864	14	19	515.78	173.04	33.10
9	**1.000**	**6**	**24**	**1.93**	**2.72**	**63.55**
10	**1.000**	**9**	**24**	**3.53**	**10.75**	**66.23**
11	0.792	7	19	104.21	38.90	49.49
12	0.917	9	22	110.15	44.11	58.41
13	**1.000**	**11**	**24**	**312.81**	**109.42**	**59.42**
14	**1.000**	**13**	**24**	**546.45**	**185.19**	**54.72**
15	**1.000**	**14**	**24**	**596.66**	**200.73**	**50.74**
16	0.958	23	23	646.81	216.95	30.75
17	0.857	6	21	60.98	22.22	54.56
18	0.800	8	20	26.77	16.49	54.48
19	0.880	9	22	92.48	38.49	58.76
20	0.920	10	23	195.17	71.83	59.37
21	0.960	12	24	475.96	161.33	55.21
22	**1.000**	**14**	**25**	**662.25**	**222.22**	**50.36**
23	0.960	14	24	637.41	213.88	47.55
24	0.720	15	18	495.07	166.08	28.68
25	0.625	5	16	17.24	9.16	42.83
26	0.731	8	19	128.88	48.89	49.71
27	0.769	8	20	63.11	28.27	53.88
28	0.885	11	23	408.43	138.91	54.33
29	0.885	12	23	504.77	170.59	51.25
30	0.885	13	23	590.34	198.40	47.57
31	0.885	14	23	613.43	205.83	44.61
32	0.885	23	23	646.81	216.95	30.75
33	0.593	5	16	14.01	8.18	42.92
34	0.630	7	17	58.53	25.54	45.59
35	0.704	8	19	103.04	40.67	50.22
36	0.741	8	20	70.38	30.63	53.76
37	0.815	12	22	533.28	179.76	47.32
38	0.889	14	24	637.84	214.03	47.48
39	0.778	12	21	556.36	186.69	42.06
40	0.815	22	22	618.69	207.52	29.41
41	0.586	5	17	52.48	19.29	44.27
42	0.552	7	16	80.29	32.04	42.23
43	0.621	8	18	148.42	54.83	46.47
44	0.655	9	19	231.04	81.77	47.47
45	0.621	8	18	224.50	78.38	44.62
46	0.765	13	22	585.31	196.40	43.27
47	**1.000**	**29**	**24**	**693.48**	**232.56**	**24.39**

Se puede observar cómo los valores de ambos recursos son enteros para todas las DMUs.

Además, hay que hacer notar que:

a) La eficiencia CCR tradicional es siempre menor que la eficiencia entera.

b) Las unidades eficientes en CRS tradicional también los son en el modelo CRS entero, pero aparecen unidades CRS eficiente-entero que no lo son en el modelo CRS tradicional.

c) La desviación absoluta media entre ambas medidas de eficiencia es 0.019, la máxima desviación absoluta es 0.047 (obtenida por DMU_6) y el coeficiente de correlación entre ambas medidas es de 0.996.

d) En algunos casos, (p.ej. DMU_1 y DMU_2) la solución obtenida por el modelo propuesto coincide con el resultado de redondear los valores fraccionarios al entero más próximo. Sin embargo no es así en otros caso como en la DMU_{15}, en donde se redondea hacia arriba (desde 23.04 a 24 para la segunda de las entradas) y a veces no al entero inmediatamente superior (desde 12.60 a 14, en el caso de la primera entrada).

e) El valor medio de las diferencias entre los valores objetivos de las entradas entre los modelos resueltos es de 0.410 para el número de trabajadores y de 0.477 para el número de máquinas. Las máximas diferencias son de 1.622 y 1.113 respectivamente, lo que significa que, en media, las proyecciones con el modelo entero añaden casi media unidad para cada entrada.

f) Para las salidas, la media de las diferencias entre el objetivo de salidas para cada modelo son de 36.39 para el WIP^{-1}, 11.53 para el flow-time^{-1}, y 0.18 para la utilización de mano de obra. Las máximas diferencias fueron de 139.74, 45.28 y 2.20, respectivamente. Esto significa que, en consonancia con los incrementos generados en las entradas, las proyecciones generadas por el modelo propuesto también incrementan las salidas.

Respecto de los tiempos de computación, hay que decir que el modelo CCR-I tradicional requirió de un tiempo despreciable, mientras que el modelo MILP utilizó 28.51 segundos para su resolución, usando el paquete de optimización XA en un computador AMD-K6 a 500 MHz y con 128 Mbytes de RAM. Se han resuelto con este último modelo casos más grandes (específicamente con 100 DMUs, 5 entradas y 5 salidas), y nunca se han necesitado más de unos pocos minutos en ser resueltos.

4. Conclusiones

En este artículo se ha desarrollado un modelo DEA para el caso en que existan entradas y salidas enteras que garantiza que las unidades sobre las que se proyectan las existentes son CRS eficiente-entero. La aplicación incluida en este artículo indica que se obtienen mejores resultados que la heurística consistente en el simple redondeo hacia los valores enteros más cercanos a la proyección obtenida por el modelo CRS tradicional. La diferencia entre dichos resultados se hace más importante cuanto mayores sean los costes de las dimensiones enteras del problema. Se ha comprobado cómo el modelo no requiere de mucho tiempo de computación, incluso para grandes problemas. Además es muy versátil, ya que es posible plantear otro tipo de casos como tecnología de retornos de escala variables, orientación de salida, o métrica aditiva con pocas modificaciones en la formulación.

Referencias

Cooper, W.W.; Seiford, L.M.; Tone, K. (2000). *Data Envelopment Analysis*, Kluwer.

Shafer, S.M.; Bradford, J.W. (1995) Efficiency Measurement of Alternative Machine Component Grouping Solutions Via Data Envelopment Analysis. *IEEE Transactions on Engineering Management,* Vol. 42, pp. 159-165.

Thanassoulis, E. (2000). *Introduction to the Theory and Application of Data Envelopment Analysis – A Foundation Text with Integrated Software*, Kluwer.