

REVISIÓN DEL SOFTWARE LIBRE DISPONIBLE Y DESARROLLO DE UNA HERRAMIENTA PARA EL ANALISIS ENVOLVENTE DE DATOS (DEA)

Rescala, Carmen¹; Devincenzi, Gustavo¹; Rohde, Gricela¹; Bonaffini, Ma.Liliana¹; Bernaola, Gustavo²;
Giraudó, Marta¹

Universidad Nacional del Nordeste¹ – Universidad Tecnológica Nacional²

carmenrescala@yahoo.com.ar, gdevin@ing.unne.edu.ar; grohde@eco.unne.edu.ar;
mbonaffini@eco.unne.edu.ar; gustavo@edesycc.com.ar; marta_giraudó@yahoo.com.ar

Eje Temático: Software - Matemática Aplicada

Palabras claves: Análisis de envolverte de datos – Software – Optimización multicriterio

RESUMEN

El Análisis Envolverte de Datos se viene utilizando para analizar la eficiencia de empresas o procesos que generen uno o más resultados (outputs) a partir de uno o más insumos (inputs). Su resolución requiere en general una importante cantidad de datos y el proceso debe realizarse con un software específico. Si bien se dispone de varios paquetes que resuelven este tipo de problemas, la mayoría de los de libre disponibilidad está limitado en sus prestaciones (Nros. de DMU y/o tipos de procesos). Además en casi ninguno se conoce en detalle el proceso de resolución y/o se pueden agregar nuevas funcionalidades. En este trabajo se desarrolló una aplicación para la resolución de los modelos básicos de DEA (CCR – BCC) utilizando herramientas libres como "lp_solve" en un ambiente de amplia disponibilidad y facilidad de carga y modificación de datos tabulares como es el Microsoft Excel, programando en el lenguaje VBA del mismo.

INTRODUCCIÓN

El concepto de eficiencia está relacionado con la economía de recursos. Es frecuente definir la eficiencia de un proceso como la relación entre los resultados obtenidos (outputs o productos) y los recursos utilizados (inputs o insumos).

El Análisis Envolverte de Datos, DEA, surge como una extensión del trabajo de Farrell (1957), quién proporciona una 'medida satisfactoria de eficiencia productiva' que tiene en cuenta todos los Inputs (recursos empleados) y muestra como puede ser calculada, ilustrando su método mediante una aplicación a la producción agrícola de Estados Unidos. Charnes,

Cooper y Rhode (1978), veinte años después del trabajo de Farrell, fueron quienes propusieron esta metodología, que ha tenido gran difusión.

La misma consiste en modelos matemáticos basados en programación lineal los que realizan el cálculo de eficiencias relativas de cada entidad respecto a un conjunto definido de entidades equivalentes.

Con la misma se han abierto posibilidades para el uso en casos que han sido resistentes a otros métodos, debido a la naturaleza compleja y a menudo desconocidas de las relaciones entre las múltiples entradas y salidas que pueden participar para la evaluación de las diferentes entidades. DEA también ha sido utilizado para suministrar nuevos conocimientos sobre las actividades y entidades que han sido previamente evaluadas por otros métodos.

La formulación de los Modelos Básicos de DEA se aprecia en la tabla 1:

Table 1-1. CCR DEA Model

Input-oriented	
Envelopment model	Multiplier model
$\min \theta - \varepsilon (\sum_{i=1}^m s_i^- + \sum_{r=1}^s s_r^+)$ <p>subject to</p> $\sum_{j=1}^n x_{ij} \lambda_j + s_i^- = \theta x_{i0} \quad i = 1, 2, \dots, m;$ $\sum_{j=1}^n y_{rj} \lambda_j - s_r^+ = y_{r0} \quad r = 1, 2, \dots, s;$ $\lambda_j \geq 0 \quad j = 1, 2, \dots, n.$	$\max z = \sum_{r=1}^s \mu_r y_{r0}$ <p>subject to</p> $\sum_{r=1}^s \mu_r y_{rj} - \sum_{i=1}^m v_i x_{ij} \leq 0$ $\sum_{i=1}^m v_i x_{i0} = 1$ $\mu_r, v_i \geq \varepsilon > 0$
Output-oriented	
Envelopment model	Multiplier model
$\max \phi + \varepsilon (\sum_{i=1}^m s_i^- + \sum_{r=1}^s s_r^+)$ <p>subject to</p> $\sum_{j=1}^n x_{ij} \lambda_j + s_i^- = x_{i0} \quad i = 1, 2, \dots, m;$ $\sum_{j=1}^n y_{rj} \lambda_j - s_r^+ = \phi y_{r0} \quad r = 1, 2, \dots, s;$ $\lambda_j \geq 0 \quad j = 1, 2, \dots, n.$	$\min q = \sum_{i=1}^m v_i x_{i0}$ <p>subject to</p> $\sum_{i=1}^m v_i x_{ij} - \sum_{r=1}^s \mu_r y_{rj} \geq 0$ $\sum_{r=1}^s \mu_r y_{r0} = 1$ $\mu_r, v_i \geq \varepsilon > 0$

Tabla 1 - Fuente: COOPER, W. W.; SEIFORD, L. M. y ZHU, J.^[3], pág 13.

Para la resolución de estos modelos se hace necesario el uso de herramientas computacionales. Son varios los desarrollos de software que se conocen. Una búsqueda que ha realizado el grupo de investigación por Internet nos permitió armar la tabla 2. La colocación de la "L" como supra-índice luego del nombre indica que dicho paquete es de libre uso.

Programa	Autor(es) y/o Página Web
BSFM ^L	Arickx, Broeckhove, Dejonghe & van den Broeck (1997)
DEA Excel Solver ^L	Zhu (2003); www.deafontier.com/deasolver.html
DEAP (v 2.1) ^L	Coelli (1996); www.uq.edu.au/economics/cepa
DEAQual ^L	http://faculty.fuqua.duke.edu/%7Ekamakura/bio/WagnerKamakuraDownloads.htm
DEA-Solver-Pro	www.saitech-inc.com/Products/Prod-DSP.asp
EMS ^L	www.wiso.uni-dortmund.de/lsg/or/scheel/ems
FEAR ^L	Wilson (2008); www.clemson.edu/economics/faculty/wilson/Software/FEAR/fear.html
Frontier (v 4.1) ^L	Coelli (1996); www.uq.edu.au/economics/cepa
Frontier Analyst	www.banxia.com
LIMDEP	www.limdep.com
OnFront	www.emq.com/software.html
PIM-DEAssoft	www.deasoftware.co.uk
ISYDS (SIAD) ^L	Angula-Meza et al. (2005); http://www.uff.br/decisao/
StoNED ^L	Kuosmanen (2006, 2007); Kuosmanen & Kortelainen (2007); www.nomepre.net/stoned
TFPIP (v 1.0) ^L	Coelli (1997); www.uq.edu.au/economics/cepa
xIDEA	www.prodtools.com/Products.html#
WinBUGS ^L	Griffin & Steel (2007); www2.warwick.ac.uk/fac/sci/statistics/staff/academic/steel/steel_homepage/software/
Programa Online	Página Web
DEA Solver Online ^L	www.dea.uni-hohenheim.de/index.php

Tabla 2

En un trabajo Barr, R. S.^[1] presenta un análisis del software para DEA, que se reseña en la tabla 3.

	COMMERCIAL				NON-COMMERCIAL			
	DEA Solver Pro	Frontier Analyst	OnFront	Warwick DEA	DEA Excel Solver	DEAP	EMS	Pioneer
	Version 4.0	3.1.5	2.02	1.0	1.0	2.1	1.3.0	2.0
From	SAITECH	Barwa Software	EMQC	Warwick Univ.	Zhu	Colletti	Scheel	Barr, McLeod
A Models								
1	CCR/CRS	●	●	●	●	●	●	●
2	BCC/VRS	●	●	●	●	●	●	●
3	NIRS, NDRS, GRS	●	.	●	.	●	●	●
4	Additive/slack-based method	●	.	●	.	●	●	.
5	Malmquist	●	●	.	.	●	●	.
6	Non-convex	●	.	.	.	●	●	.
7	Non-radial	●	.	.	●	.	.	.
8	Preference-structure	●	.	.
9	Undesirable-measure	●	.	.
10	Context-dependent	●	.	.
11	Free-disposal hull (FDH)	●	.	.	.	●	●	●
12	Cost efficiency	●	.	●	.	●	.	.
13	Revenue efficiency	●	.	●	.	●	.	.
14	Profit, revenue/cost efficiency	●	.	●	.	●	.	.
15	Target, mixed improvement	●	.	.	●	.	.	.
16	Capacity utilization	.	.	●
17	Variable-benchmark	●	.	.
18	Fixed-benchmark	●	.	.
19	Minimum-efficiency	●	.	.
20	Value chain	●	.	.
21	Weak disposability	.	.	●	.	●	.	.
22	New cost, revenue, profit	●
23	Congestion	●	.	●	.	●	.	.
24	Scale elasticity	●

Tabla 3

No obstante, para ninguno de ellos se dispone del código fuente, por lo que no se conoce en detalle la rutina de optimización utilizada ni se pueden realizar otras incorporaciones o funcionalidades a los modelos definidos.

Con esta situación, se decidió considerar el desarrollo de una rutina propia.

Se analizaron:

- Lenguaje / herramienta a utilizar
- Interfase
- Rutinas / librerías de resolución de los problemas de programación Lineal (PL)

En virtud a la forma tabular en que se suelen disponer los datos para estos procesos, se eligió como plataforma para el desarrollo la planilla MS-Excel. En este entorno disponemos

de facilidad de carga y modificación de los datos y un lenguaje de programación suficientemente potente (VBA).

Como interfase se diseñaron formularios donde se definen los principales parámetros del problema a resolver. Se toma la información de base desde celdas de hojas que también se indican.

Respecto a la rutina para PL, se analizó en principio el Solver que viene con el Excel.

El número máximo de celdas ajustables y las restricciones que puede utilizar un modelo de Solver aparecen en la tabla siguiente.

Microsoft Excel	Cant. celdas ajustables	Nº de restricciones
3.x	200	299
4.x	200	500
5.0	200	500
7.0	200	500
97 y sup.	200	Lineal ilimitado/No Lineal 500

Si bien para un importante número de problemas puede ser satisfactorio, se realizó el análisis de otras alternativas para no tener esta limitación.

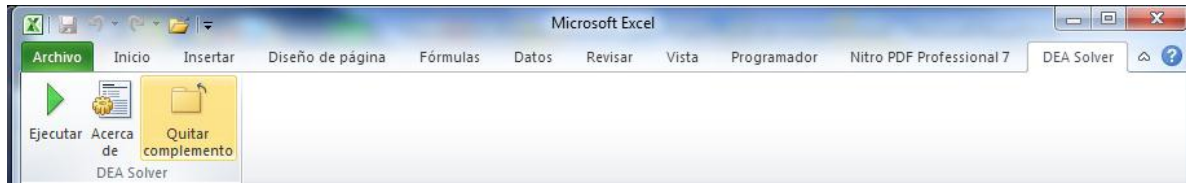
Se buscaron librerías de libre uso de PL. En las siguientes se hacía necesaria una compilación y/o adecuación para usarla en este entorno: BLAS (Basic Linear Algebra Subprograms), LAPACK (Linear Algebra Package), ATLAS (Automatically Tuned Linear Algebra Software), GotoBLAS2, entre otras.

La que cumplió las expectativas fue lp_solve. Esta librería implementa el método Simplex para variables reales (puede calcular también problemas con variables enteras). Dispone de una muy buena documentación de su uso y de su amplia API (Application Programming Interface)

Con estas definiciones se procedió al desarrollo de la aplicación, implementando en principio los modelos básicos.

A continuación se muestran los menús y formularios de la aplicación, la que fue desarrollada como complemento para las versiones 2007 o superior, a los efectos guardar solamente las hojas de datos y disponer de una sola versión de la misma.

Al cargarse el complemento se agrega una nueva solapa a la cinta de opciones, que se aprecia a la derecha de la imagen. Dicha solapa dispone de las opciones que se observan en la parte inferior izquierda de la imagen.



Es necesario tener una planilla cargada con los datos a procesar.

La información debe disponerse de la siguiente forma:

Una columna donde se identifica con nombre o denominación las DMU's a procesar (unidades de análisis o unidades de decisión). Se utiliza una fila por cada DMU.

Una columna por cada Input (definiendo el valor medido o estimado para cada DMU en su fila correspondiente)

Una columna por cada Output (definiendo también el valor medido o estimado para cada DMU en su fila correspondiente)

Un ejemplo se muestra a continuación:

DMU'S	(I)ACTIVO CTE	(I)Bienes de Uso	(O)Ventas
2005	863409.41	82779.71	3579183.28
2006	564408.14	135783.59	1694539.43
2007	1181301.01	225719.44	2615370.07
2008	1241476.71	332756.04	2274175.11

La opción Ejecutar  es la que inicia el proceso.

El primer formulario que se presenta permite definir cuestiones como:

- Lenguaje en que se muestran los diálogos (Español / Inglés)
- Tolerancia para el cálculo
- Título del trabajo
- Valor de la variable Epsilon
- Si se desea normalización de los datos.

La siguiente opción nos permite definir:

- Etiquetas de los DMU's
- Etiquetas de los Inputs
- Etiquetas de los Outputs
- Matriz de Inputs
- Matriz de Outputs
- Orientación del modelo (Inputs u Outputs)
- Tipo de modelo (Multiplicativo o Envolvente)
- Tipo de frontera (CCR o BCC)
- Si se aplica Súper eficiencia
- Tipo de resultados (resumido o detallado) y cantidad de decimales para los cálculos.

Especificación de los datos y modelo

Etiquetas de los DMU's:

Etiquetas de los Inputs:

Etiquetas de los Outputs:

Matriz de inputs:

Matrix de outputs:

Orientación del modelo

Orientado a los inputs Orientado a los outputs

Tipo de modelo

Multiplicativo Envolvente

Tipo de frontera

CCR BCC

Super-eficiencia - Si Resultados - detallado

Super-eficiencia - No Resultados - resumido

Cant. dec. Lambdas 4

Cant. dec. Inputs 4

Cant. dec. Outputs 4

Haciendo clic en el botón “Resolver”, se genera una nueva Hoja con los resultados conforme la elección realizada.

Una parte de la misma, con el ejemplo y selecciones antes indicados, se muestra a continuación:

	A	B	C	D
1	Resultados - detallado CCR_Mul_I			
2	DMU - 2005			
3	Resultado de eficiencia	1.000		
4	Pesos Input (Vi)	0.000	0.000	
5	Pesos Output (Ui)	0.000		
6	Inputs	ACTIVO CTE	Bienes de Uso	
7	Inputs real	863.409	82.780	
8	Contribución	100.000%	0.000%	
9	Outputs	Ventas		
10	Outputs real	3579.183		
11	Contribución	100.000%		
12				
13	DMU - 2006			
14	Resultado de eficiencia	0.724		
15	Pesos Input (Vi)	0.000	0.000	
16	Pesos Output (Ui)	0.000		
17	Inputs	ACTIVO CTE	Bienes de Uso	
18	Inputs real	564.408	135.784	
19	Contribución	100.000%	0.000%	
20	Outputs	Ventas		
21	Outputs real	1694.539		
22	Contribución	100.000%		
23				
24	DMU - 2007			
25	Resultado de eficiencia	0.534		
26	Pesos Input (Vi)	0.000	0.000	
27	Pesos Output (Ui)	0.000		
28	Inputs	ACTIVO CTE	Bienes de Uso	
29	Inputs real	1181.301	225.719	
30	Contribución	100.000%	0.000%	
31	Outputs	Ventas		
32	Outputs real	2615.370		
33	Contribución	100.000%		
34				
35	DMU - 2008			
36	Resultado de eficiencia	0.442		
37	Pesos Input (Vi)	0.000	0.000	
38	Pesos Output (Ui)	0.000		
39	Inputs	ACTIVO CTE	Bienes de Uso	
40	Inputs real	1241.477	332.756	
41	Contribución	100.000%	0.000%	
42	Outputs	Ventas		
43	Outputs real	2274.175		
44	Contribución	100.000%		
45				

REFERENCIAS BIBLIOGRÁFICAS

- [1] BARR, R.S. (2004); "DEA Software Tools and Technology: A State-of-the-Art Survey" in: W.W. Cooper, L.M. Seiford, J. Zhu; Handbook on Data Envelopment Analysis, Kluwer, Boston, pp. 539-566.
- [2] CHARNES, A.; COOPER, W.W. y RHODES, E. (1978); "Measuring the Efficiency of Decision Making Units"; European Journal of Operational Research, vol. 2, nº 6, pp. 429-444.
- [3] COOPER, W. W.; SEIFORD, L. M. y ZHU, J. (2004); "Handbook On Data Envelopment Analysis"; Kluwer Academic Publisher; ISBN: 1402077971.
- [4] FARRELL, M.J. (1957); "The Measurement of Productive Efficiency"; Journal of the Royal Statistics Society, Serie A, 120, pp. 253-281.
- [5] lp_solve reference guide menú; en <http://lpsolve.sourceforge.net/5.0/> (accedido en abril de 2012)