INTERNATIONAL TRANSACTIONS IN OPERATIONAL RESEARCH

Intl. Trans. in Op. Res. 21 (2014) 619–625
DOI: 10.1111/itor.12074

Nonparametric cost and revenue functions under constant economies of scale: an enumeration approach for the single output or input case

Walter Briec[a], Kristiaan Kerstens[b] and Ignace Van de Woestyne[c]

[a]*LAMPS, Université de Perpignan, France*
[b]*CNRS-LEM (UMR 8179), IESEG School of Management, 3 rue de la Digue, F-59000 Lille, France*
[c]*KU Leuven, Research unit MEES, Brussels, Belgium*
E-mail: briec@univ-perp.fr [Briec]; k.kerstens@ieseg.fr [Kerstens];
ignace.vandewoestyne@kuleuven.be [Van de Woestyne]

Received 27 June 2013; received in revised form 15 December 2013; accepted 24 December 2013

Abstract

This article shows how the linear programs needed to compute cost and revenue functions under constant returns to scale and a single output or input, respectively, can be replaced with a more efficient enumeration algorithm.

Keywords: nonparametric cost and revenue functions; enumeration; linear programming

1. Introduction

Nowadays, cost and revenue functions are often estimated using nonparametric, deterministic estimators (see, e.g., Cooper et al., 2006; Hackman, 2008; Ray, 2004). This involves the computation of one linear program (LP) per observation under evaluation in the sample. Obtaining statistical inference from these extremum estimators using recent bootstrapping techniques requires again solving an LP in each draw (see, e.g., De Borger et al., 2008 for an application). This can result in a substantial computational burden.

It has gone unnoticed so far that the computation of the cost function can be simplified in the single output case for constant economies of scale. Similarly, the solution of the revenue function also simplifies in the single input case under identical economies of scale. To the best of our knowledge, this is the first contribution showing that an enumeration algorithm works for these specific convex data envelopment analysis (DEA) value based models. Soleimani-damaneh (2009) is among the sole contributions we are aware of developing a similar enumeration strategy to determine efficiency position and returns-to-scale classification for the standard DEA models.

Our contribution must be seen against the background of a small, burgeoning literature focusing on a variety of strategies to speed up the LP computations underlying DEA production frontier models. Ali (1993) is probably the first study initiating this research into the computational aspects of DEA. Following a taxonomy introduced in some early overview article of Dulá (2002), one can distinguish between preprocessors, enhanced procedures, and new algorithms. In contrast to this rather substantial literature, to our knowledge very few articles have focused on simplifying the computational burden for computing cost or revenue functions. Following up on an earlier contribution by Camanho and Dyson (2005), Jahanshahloo et al. (2008) simplify the LP formulations for the traditional convex cost functions by cutting down on the amount of constraints and decision variables. Some similarly related articles are found in the literature. Our approach continues this line of research by focusing on a specific returns-to-scale assumption on a convex technology and by restricting the number of inputs or outputs.

The purpose of this note is to prove both results regarding the use of enumeration for cost and revenue functions under constant returns to scale (CRS) and a single output or input, respectively. Section 2 introduces basic definitions. Section 3 contains the main results. A concluding section offers some further perspectives.

2. Technology, cost, and revenue functions

Deterministic, nonparametric technologies are based on activity analysis. A technology uses a vector of inputs $x \in \mathbb{R}^N_+$ to produce a vector of outputs $y \in \mathbb{R}^M_+$. This technology or production possibility set is the set of all feasible input–output vectors: $T = \{(x,y)\colon x \text{ can produce } y\}$. Alternatively, the input set $L(y)$ denotes all input vectors x producing the output vector y: $L(y) = \{x\colon (x,y) \in T\}$. Equally so, the output set $P(x)$ is defined as the set of all output vectors y that can be obtained from the input vector x: $P(x) = \{y\colon (x,y) \in T\}$.

The standard radial input efficiency measure is defined as

$$DF_i(x, y) = \min\{\lambda \mid \lambda \geq 0, (\lambda x) \in L(y)\}. \tag{1}$$

Its main properties are (i) $0 < DF_i(x,y) \leq 1$, with efficient production on the boundary (isoquant) of $L(y)$ represented by unity; and (ii) it has a cost interpretation (see, for instance, Hackman, 2008).

Assume that p is a vector of positive input prices ($p \in \mathbb{R}^N_+\backslash\{0\}$). Then, the cost function corresponding to a given technology is defined as follows[1]:

$$C(p, y) = \inf\{p \cdot x : x \in L(y)\}. \tag{2}$$

We now briefly elucidate the basic efficiency decomposition distinguishing technical and allocative efficiency. For a given observation, the radial input efficiency measure to the isoquant of the input set $L(y)$ represents its technical efficiency. For the same observation, the radial distance to the iso-cost line tangent to this isoquant represents a measure of cost efficiency. Finally, since cost efficiency is

[1]The radial input efficiency measure, being the inverse of the input distance function, is related to the cost function via duality relation (for details, see Hackman, 2008).

always lower or equal to technical efficiency, in case of a difference this can be attributed to allocative efficiency. The resulting basic efficiency decomposition states that cost efficiency is the product of a technical efficiency component and an allocative efficiency component (see Cooper et al., 2006, ch. 8, for further details).

Equally so, assume that r is a vector of positive output prices ($r \in \mathbb{R}_+^M$), then the revenue function corresponding to a given technology is defined by

$$R(r, x) = \max\{r \cdot y : y \in P(x)\}. \tag{3}$$

Apart from imposing traditional assumptions on technology (i.e., no free lunch and inaction, closedness, free disposal of inputs and outputs, and convexity), the sole key assumption we invoke in this contribution is CRS (i.e., when $(x,y) \in T$, then $\delta(x,y) \in T$, $\forall \delta > 0$). Several nonparametric technologies have been derived from these axioms (Banker et al., 1984, are among the earlier sources).

A convex technology based on K observations (x_k, y_k), $k = 1, \ldots, K$, satisfying the above axioms and CRS, has been defined in Charnes et al. (1978) as follows:

$$T^{CRS} = \left\{ (x, y) \mid x \in \mathbb{R}_+^N\, y \in \mathbb{R}_+^M \sum_{k=1}^{K} z_k y_k \geq y,\ \sum_{k=1}^{K} z_k x_k \leq x,\ z_k \geq 0, \quad k = 1, \ldots, K \right\}. \tag{4}$$

Introduction of this technology in this article is considered to mark the start of the DEA literature. Computing a cost (1) or revenue (2) function with respect to this CRS technology is a standard model in the DEA literature (e.g., Cooper et al., 2006) and normally requires solving one LP per observation (eventually a simplified version as elaborated by Jahanshahloo et al., 2008).

3. Main results

Minimal assumptions on observed inputs and outputs are usually formulated as follows. Summing over all observations, there is a strictly positive aggregate production of every output and a strictly positive aggregate consumption of every input. Every unit produces a positive amount of at least one output and employs a positive amount of at least one input (see, e.g., Färe et al., 1994, pp. 44–45). When considering a single output case, this implies that all observations have a strictly positive single output. Likewise, for the single input case, this implies that all observations use a strictly positive single input.

Proposition 1. *In the case of CRS and a single, strictly positive output ($M = 1$), the cost function $C^{CRS}(p, y)$ is computed as follows:*

$$C^{CRS}(p, y) = y \min_{k=1\ldots K} \left\{ \frac{1}{y_k} \cdot p \cdot x_k \right\}.$$

Proof. Assume there is a single, strictly positive output ($M = 1$). Consider the technology T^{CRS} enveloping the sample $S = \{(x_1, y_1), \ldots, (x_K, y_K)\}$. For $k = 1, \ldots, K$, denote $\delta_k = \frac{y}{y_k}$. Now, define the transformed sample $S' = \{(\delta_1 x_1, \delta_1 y_1), \ldots, (\delta_K x_K, \delta_K y_K)\} = \{(\delta_1 x_1, y), \ldots, (\delta_K x_K, y)\}$

realizing the same technology T^{CRS}. The conical hull (Cc) of a finite set is the conical hull of its convex hull (Co). Consequently,

$$T^{CRS} = (Cc(S) + \Theta) \cap \mathbb{R}_+^{N+1} = (Cc(S') + \Theta) \cap \mathbb{R}_+^{N+1} = (Cc(Co(S')) + \Theta) \cap \mathbb{R}_+^{N+1},$$

with $\Theta = \mathbb{R}_+^N \times (-\mathbb{R}_+)$. It follows that

$$\begin{aligned} L^{CRS}(y) &= \left\{x : (x, y) \in (Cc(Co(S')) + \Theta) \cap \mathbb{R}_+^{N+1}\right\} \\ &= \left\{x : (x, y) \in \left(Co\left(\left\{\delta_1 x_1, \ldots, \delta_K x_K\right\}\right) + \mathbb{R}_+^N\right) \times \{y\}\right\} \\ &= Co\left(\left\{\delta_1 x_1, \ldots, \delta_K x_K\right\}\right) + \mathbb{R}_+^N. \end{aligned}$$

Since $Co(\{\delta_1 x_1, \ldots, \delta_K x_K\})$ is a convex polyhedron by definition, the minimum of any nondecreasing linear function (e.g., the cost function) is achieved at some vertex point (see Eremin, 2002). Thus,

$$C^{CRS}(p, y) = \inf\left\{p \cdot x : x \in L^{CRS}(y)\right\} = \min_k \left\{p \cdot \delta_k \cdot x_k\right\} = \min_k \left\{\frac{y}{y_k} \cdot p \cdot x_k\right\}. \qquad \square$$

Proposition 2. *In the case of CRS and a single, strictly positive input ($N = 1$), the revenue function $R^{CRS}(r, x)$ is computed as follows:*

$$R^{CRS}(r, x) = x \max_{k=1\ldots K} \left\{\frac{1}{x_k} \cdot r \cdot y_k\right\}.$$

Proof. Assume there is a single, strictly positive input ($N = 1$). Consider a technology T^{CRS} enveloping the sample $S = \{(x_1, y_1), \ldots, (x_K, y_K)\}$. For $k = 1, \ldots, K$, denote $\mu_k = \frac{x}{x_k}$. Now, define the transformed sample $S' = \{(\mu_1 x_1, \mu_1 y_1), \ldots, (\mu_K x_K, \mu_K y_K)\} = \{(x, \mu_1 y_1), \ldots, (x, \mu_K y_K)\}$ realizing the same technology T^{CRS}. Using similar arguments as in Proposition 1, we obtain $P^{CRS}(x) = (Co(\{\mu_1 y_1, \ldots, \mu_K y_K\}) + (-\mathbb{R}_+^M)) \cap \mathbb{R}_+^M = \{y \in \mathbb{R}^M : 0 \leq y \leq \sum_{k=1}^K z_k \mu_k y_k, \sum_{k=1}^K z_k = 1, z_k \geq 0\}$.

Since the price vector $r \in \mathbb{R}_{++}^M$ is nonnegative, we have

$$\begin{aligned} R^{CRS}(r, x) &= \max\left\{r \cdot y : y \in P^{CRS}(x)\right\} \\ &= \max\left\{r \cdot y : y \in \left(Co\left(\left\{\mu_1 y_1, \ldots, \mu_K y_K\right\}\right) + (-\mathbb{R}_+^M)\right) \cap \mathbb{R}_+^M\right\} \\ &= \max\left\{r \cdot y : y \in \left(Co\left(\left\{\mu_1 y_1, \ldots, \mu_K y_K\right\}\right) + (-\mathbb{R}_+^M)\right)\right\} \\ &= \max\left\{r \cdot y : y \in Co\left(\left\{\mu_1 y_1, \ldots, \mu_K y_K\right\}\right)\right\}. \end{aligned}$$

Since $Co\{(\mu_1 y_1, \ldots, \mu_K y_K)\}$ is a convex polyhedron by definition, the maximum of any nondecreasing linear function (e.g., the revenue function) is achieved at some extreme point. Thus,

$$R^{CRS}(r,x) = \max\left\{r\cdot y : y \in P^{CRS}(x)\right\} = \max\left\{r\cdot y : y \in Co\left(\left\{\mu_1 y_1,\ldots,\mu_K y_K\right\}\right)\right\}$$
$$= \max_k\left\{r\cdot\mu_k\cdot y_k\right\} = \max_k\left\{\frac{x}{x_k}\cdot r\cdot y_k\right\}. \qquad \square$$

Remark. We are grateful to a referee for explicitly outlining alternative ways of proving Propositions 1 and 2. For Proposition 1, we start from the LP formulation of the cost function $C^{CRS}(p,y)$ in Camanho and Dyson (2005) (their formula (3)):

$$\begin{aligned}
&\min \sum_{i=1}^{N} p_i x_i^0 \\
&\text{s.t.} \quad \sum_{k=1}^{K} x_{ik} z_k = x_i^0, \quad i = 1,\ldots,N, \\
&\qquad \sum_{k=1}^{K} y_k z_k \geq y, \\
&\qquad z_k \geq 0, \quad k = 1,\ldots,K, \quad x_i^0 \geq 0,\ i = 1,\ldots,N.
\end{aligned}$$

Obviously, for satisfying the constraints, all variables x_i^0 must be strictly positive. Consequently, for having a basic feasible solution, only one variable $z_{k'}$ ($k' \in \{1,\ldots,K\}$) is nonzero while all others are zero. The constraints now simplify to $x_{ik'}z_{k'} = x_i^0,\ i = 1,\ldots,N$ and $y_{k'}z_{k'} \geq y$. Solving the equality constraints for $z_{k'}$ and substituting in the inequality constraint shows that $x_i^0 \geq \frac{y}{y_{k'}} x_{ik'}$ for all $i = 1,\ldots,N$. Thus, $\sum_{i=1}^{N} p_i x_i^0 \geq y \sum_{i=1}^{N} \frac{1}{y_{k'}} p_i x_{ik'}$ from which the desired result follows. A proof using similar arguments can be obtained for Proposition 2 and is left as an exercise to the reader.

We include an algorithm for computing C^{CRS} for all observations:

Algorithm 1.

For $i = 1,\ldots,K$ do:

(1) Select the ith observation $(x_i, y_i) = (x_{i1},\ldots,x_{iN}, y_{i1})$ and its input price vector $p_i = (p_{i1},\ldots,p_{iN})$.
(2) Put $C = \infty$.
(3) For $k = 1,\ldots,K$ do:
   (a) $C1 = \frac{y_{i1}}{y_{k1}} \cdot \sum_{j=1}^{N} p_{ij} x_{kj}$
   (b) If $C1 < C$ then $C = C1$.
(4) The variable C holds the value of $C^{CRS}(p_i, y_i)$ for the ith observation.

A similar algorithm could be formulated for R^{CRS}.

Having proven the two main results, we spell out the computational consequences in the next corollary.

Corollary 1. *In the case of CRS and a single output, the cost function can be computed by enumeration in a smaller number of operations compared to LP. The same applies to the revenue function in the case of CRS and a single input.*

Proof. In the case of a single output, enumeration requires $O(LK(1+N)^2)$ arithmetic operations, where L is a measure of data storage for a given precision. Ignoring the worst-case exponential complexity of the simplex method in LP, the Kamarkar interior point (IP) method needs $O(L(n)^{3.5})$ operations (with n being the number of decision variables) while the most successful IP method known so far (i.e., primal-dual Newton step IP method) has a complexity of $O(L(n)^3)$ (for details, see Chong and Zak, 2001; Eiselt and Sandblom, 2007). Transposed to our models, one thus needs at best $O(L(K+N)^3)$ operations for LP. Since in general $K > N \geq 1$, it follows that $K + N > 1 + N$ and consequently $(K + N)^2 > (1 + N)^2$. Also $K + N > K$, which combined with the previous inequality leads to $(K + N)^3 > K(1 + N)^2$. Hence, enumeration of the cost function under CRS and a single output is always quicker compared to LP. The same argument applies to Proposition 2. □

4. Conclusions

This article is the first to prove that an enumeration algorithm can be employed to solve certain specific convex DEA type value-based models. Hitherto, enumeration has solely been applied to the specific structure of a nonconvex production model (see, e.g., Ray, 2004).

Obviously, we do not claim that enumeration is a viable solution strategy for convex DEA type of production- and value-based models in general (see Soleimani-damaneh, 2009). However, it cannot be excluded that enumeration could be applied to some other specific convex DEA models. For instance, to the extent that one is willing to select an efficiency measure that always projects onto a vertex point, the same procedure could probably be applied to production models under CRS and a single output or input with the measurement orientation along this single dimension (see, e.g., Russell and Schworm, 2011, for some of the more recent choices). This could be a promising avenue for future research.

Acknowledgments

We thank two referees and an associate editor for their most constructive comments.

References

Ali, I.A., 1993. Streamlined computation for data envelopment analysis. *European Journal of Operational Research* 64, 1, 61–67.

Banker, R., Charnes, A., Cooper, W., 1984. Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management Science* 30, 9, 1078–1092.

Camanho, A.S., Dyson, R.G., 2005. Cost efficiency measurement with price uncertainty: a DEA application to bank branch assessments. *European Journal of Operational Research* 161, 2, 432–446.

Charnes, A., Cooper, W., Rhodes, E., 1978. Measuring the efficiency of decision making units. *European Journal of Operational Research* 2, 6, 429–444.

Chong, E.K.P., Zak, S.H., 2001. *Introduction to Optimization* (2nd edn). Wiley, New York.

Cooper, W.W., Seiford, L.M., Tone, K., 2006. *Data Envelopment Analysis: A Comprehensive Text with Models, Applications, References and DEA-Solver Software* (2nd edn). Springer, Berlin.

De Borger, B., Kerstens, K., Staat, M., 2008. Transit costs and cost efficiency: bootstrapping non-parametric frontiers. *Research in Transportation Economics* 23, 1, 53–64.

Dulá, J.H., 2002. Data envelopment analysis. In Pardalos, P.M., Resende, M.G.C. (eds) *Handbook of Applied Optimization*. Oxford University Press, Oxford, pp. 531–543.

Eiselt, H.A., Sandblom, C.-L., 2007. *Linear Programming and Its Applications*. Springer, Berlin.

Eremin, I.I., 2002. *Theory of Linear Optimization*. VSP, Utrecht.

Färe, R., Grosskopf, S., Lovell, C.A.K., 1994. *Production Frontiers*. Cambridge University Press, Cambridge.

Hackman, S.T., 2008. *Production Economics: Integrating the Microeconomic and Engineering Perspectives*. Springer, Berlin.

Jahanshahloo, G.R., Soleimani-damaneh, M., Mostafaee, A., 2008. A simplified version of the DEA cost efficiency model. *European Journal of Operational Research* 184, 2, 814–815.

Ray, S.C., 2004. *Data Envelopment Analysis: Theory and Techniques for Economics and Operations Research*. Cambridge University Press, Cambridge.

Russell, R.R., Schworm, W., 2011. Properties of inefficiency indexes on <Input, Output> space. *Journal of Productivity Analysis* 36, 2, 143–156.

Soleimani-damaneh, M., 2009. A fast algorithm for determining some characteristics in DEA. *Journal of the Operational Research Society* 60, 11, 1528–1534.