“This is a post-peer-review, pre-copyedit version of an article published in the Journal of the Operational Research Society. The definitive publisher-authenticated version Färe, R. and Charles, V. (2017) ‘A directional distance function approach to void the non-Archimedean in DEA’, Journal of the Operational Research Society. doi: 10.1057/s41274-017-0264-x. is available online at: https://link.springer.com/article/10.1057/s41274-017-0264-x”

A Directional Distance Function Approach to Void the Non-Archimedean in DEA

Rolf Färe[a], Vincent Charles[b,*]

[a] *Department of Economics, Oregon State University, Corvallis, OR 97331, United States*
[a] *AREC at University of Maryland, Baltimore, MD 21201, United States*
[b] *Buckingham Business School, University of Buckingham, Buckingham MK18 1EG, United Kingdom*
[b] *CENTRUM Catolica Graduate Business School, Pontificia Universidad Catolica del Peru, Lima 33, Peru*

Abstract

Over the past years, the data envelopment analysis (DEA) methodology has registered widespread use among researchers from many fields. Furthermore, it is important to note that the non-Archimedean infinitesimal, ϵ, is a key concept in DEA models. Nevertheless, it is known that some computational difficulties arise when using ϵ in DEA. In this short communication, we show how the non-Archimedean may be voided using a directional distance function approach. Thus, our approach avoids choosing a real number (10^{-5} or 10^{-6}) as a value for ϵ or estimating the same.

Keywords: Data Envelopment Analysis, directional distance function, non-Archimedean.
JEL Classification: C02, C61.

1. Introduction

Ever since Charnes, Cooper, and Rhodes (1979)'s proposal that the weights must be strictly positive in data envelopment analysis (DEA) models, which anticipated the introduction of the non-Archimedean infinitesimal ϵ, the same has become a topic of interest for many researchers. The non-Archimedean ϵ can be observed in two places: (a) in the

*Corresponding author
Email addresses: `rolf.fare@oregonstate.edu` (Rolf Färe), `v.charles@buckingham.ac.uk` (Vincent Charles)

multiplier model, wherein ϵ is used as a lower bound for the weights and (b) in the envelopment model, wherein ϵ is used at the objective function level. The original two-stage computation of DEA involving the non-Archimedean ϵ can be found in Cooper *at al.* (2006).

It is known, however, that some computational difficulties arise when using an infinitesimal in DEA because of the finite tolerances in computer calculations. For example, setting ϵ to 10^{-5} or 10^{-6} may not just change the efficiency score, but also the rankings (Färe *et al.*, 2016).

Attempts have been made in time to determine, estimate, bound, or even eliminate ϵ. The lack of consensus over a concrete need for ϵ was elegantly stated in 1993 by Thompson et al., who drew attention to the fact that "it is not well recognized that the artificial, non-Archimedean construct is not necessarily needed to exclude zero multipliers and to identify positive slacks" (p. 379). In the same year, Tone (1993) proposed an ϵ-free DEA model, and also introduced a new DEA efficiency measure, which, however, is not invariant to the scaling of the input and output data; hence, by means of using a three-phase approach, he adjusted the measure by considering some weights corresponding to the relative importance of the inputs and outputs of the decision-making unit (DMU) of interest. Ali and Seiford (1993), on the other hand, proposed an upper bound on ϵ for feasability for the multiplier side and boundedness for the associated dual envelopment side. Later, Mehrabian, Jahanshahloo, Alirezaee, and Amin (2000) showed that Ali and Seiford's bound is invalid and presented a procedure for determining an assurance interval for ϵ.

In the following years, most of the studies have focused on estimating ϵ. Jahanshahloo and Khodabakhshi (2004) determined an assurance interval for ϵ. Furthermore, Amin and Toloo (2004) presented an algorithm for computing ϵ in DEA models, showing that this algorithm is polynomial-time of $\mathrm{O}(n)$, where n is the number of DMUs. Alirezaee (2005), on the other hand, determined the assurance interval of ϵ by means of a partition-based algorithm, which involves solving a few number of linear programs. A more recent paper is that by Podinovski and Bouzdine-Chameeva (2017), wherein the authors prove the existence of an effective bound for ϵ.

In this short communication, we propose a way to void the non-Archimedean ϵ through a directional distance function approach. This is an alternative to Cooper *et al.* (2006).

This approach, also, allows to differentiate between efficiency (perfect) and efficiency in the second phase; hence, its robustness and practicality lie within.

2. Directional distance function-based approach

Assume there are $k = 1, 2, ..., K$ decision making units (DMUs) using $x^k \in \Re_+^N$ inputs to produce $y^k \in \Re_+^M$ outputs. A DEA technology created from the data is:

$$T = \left\{ (x, y) : \sum_{k=1}^{K} z_k x_{kn} \leqq x_n, \forall n, \sum_{k=1}^{K} z_k y_{km} \geqq y_m, \forall m, z_k \geqq 0, \forall k \right\} \tag{1}$$

where $z_k \geqq 0, k = 1, 2, ..., K$ are the intensity variables. System (1) meets constant returns to scale and has inputs and outputs, which are freely (strongly) disposable.

The input-oriented two-stage DEA model containing the non-Archimedean ϵ can be written as:

$$\min \lambda - \epsilon \sum_{n=1}^{N} s_n$$

subject to

$$\sum_{k=1}^{K} z_k x_{kn} = \lambda x_{k'n} - s_n, \quad n = 1, 2, \ldots, N, \tag{2}$$

$$\sum_{k=1}^{K} z_k y_{km} \geqq y_{k'm}, \quad m = 1, 2, \ldots, M,$$

$$z_k \geqq 0, \quad k = 1, 2, \ldots, K,$$

where s_n is the n^{th} slack variable.

To void the ϵ, we set it to $+1$, then take $s_n = \beta_n g_n$, where $g_n = (g_1, ..., g_n, ..., g_N)$ is the directional vector with the unit of measurement for g_n as s_n.Then, we set these to $+1_n$.This gives us the following new problem, which includes no non-Archimedean:

$$
\begin{aligned}
&\min \ \lambda - \sum_{n=1}^{N} \beta_n 1_n \\
&\text{subject to} \\
&\sum_{k=1}^{K} z_k x_{kn} = \lambda x_{k'n} - \beta_n 1_n, \quad n = 1, 2, \ldots, N, \\
&\sum_{k=1}^{K} z_k y_{km} \geqq y_{k'm}, \quad m = 1, 2, \ldots, M, \\
&z_k \geqq 0, \quad k = 1, 2, \ldots, K.
\end{aligned} \tag{3}
$$

System (3) is a merger between Farrell's (1957) input-oriented model and Färe and Grosskopf's (2010a, b) slacks-based model of directional distance function with $\beta_n, \forall n$.

Proposition: *β_n is independent of unit of measurement.*
Proof: Consider the following n^{th} constraint from System(3), $\sum_{k=1}^{K} z_k x_{kn} = \lambda x_{k'n} - \beta_n 1_n$. Note that $1_n = g_n$ has the same unit of measurement as x_n. Let us change x_{kn} to $\hat{x}_{kn} = (x_{kn}\mu)$, kilograms to pounds by μ. Then, $\sum_{k=1}^{K} z_k(x_{kn}\mu) = \lambda(x_{k'n}\mu) - \beta_n(1_n\mu)$, which implies $\mu \sum_{k=1}^{K} z_k x_{kn} = \mu(\lambda x_{k'n} - \beta_n 1_n)$. Therefore, $\beta_n 1_n$ did not change; hence, it is independent of unit of measurement. ■

The slacks-based measure System (2) has the drawback that slacks of different units are added. To remove this problem from our model, *i.e.*, System (3), we just drop the 1_n in the objective function, so that we have min $\lambda - \sum_{n=1}^{N} \beta_n$. Since β_n is independent of unit of measurement, so is the objective function.

To compute our new model in which λ is to be minimized and $\sum_{n=1}^{N} \beta_n$ should be maximized (we need to make $\lambda x_{k'n} - \beta_n 1_n$ small), one may solve a non-linear problem as follows:

$$\max \lambda + \sum_{n=1}^{N} \beta_n 1_n$$

subject to

$$\sum_{k=1}^{K} z_k x_{kn} = \frac{x_{k'n}}{\lambda} - \beta_n 1_n, \quad n = 1, 2, \ldots, N,$$
$$\sum_{k=1}^{K} z_k y_{km} \geqq y_{k'm}, \quad m = 1, 2, \ldots, M, \tag{4}$$
$$z_k \geqq 0, \quad k = 1, 2, \ldots, K.$$

or solve the problem in two phases, following Zieschang (1984):

Phase I: $\lambda^* = \min \lambda$

subject to

$$\sum_{k=1}^{K} z_k x_{kn} \leqq \lambda x_{k'n}, \quad n = 1, 2, \ldots, N,$$
$$\sum_{k=1}^{K} z_k y_{km} \geqq y_{k'm}, \quad m = 1, 2, \ldots, M, \tag{5}$$
$$z_k \geqq 0, \quad k = 1, 2, \ldots, K.$$

Phase II: $\max \sum_{n=1}^{N} \beta_n 1_n$

subject to

$$\sum_{k=1}^{K} z_k x_{kn} = \lambda^* x_{k'n} - \beta_n 1_n, \quad n = 1, 2, \ldots, N,$$
$$\sum_{k=1}^{K} z_k y_{km} \geqq y_{k'm}, \quad m = 1, 2, \ldots, M, \tag{6}$$
$$z_k \geqq 0, \quad k = 1, 2, \ldots, K.$$

Definition 1: *The DMU of interest (k') is said to be efficient (perfectly) if and only if* $\lambda^* = 1$ *and all* $\beta_n = 0, n = 1, 2, ..., N$.
Definition 2: *The DMU of interest (k') is said to be efficient if and only if* $\lambda^* = 1$ *and at least one* $\beta_n \neq 0, n = 1, 2, ..., N$.

The DMU of interest is said to be perfectly efficient when there is no room for improvement. It is to be noted that our paper cannot find weakly efficient units in line with Cooper *et al*. (2006). We numerically illustrate our two-phase procedure in the following section.

3. Numerical Example

Let us consider five DMUs with two inputs and one unique output (= 1), as follows: $A(1, 2, 1), B(1, 1, 1), C(2, 1, 1), D(2, 2, 1)$, and $E(2, 4, 1)$. Figure 1 depicts the DMUs' position in the production possibility set with the x-axis being the first input over the output and the y-axis being the second input over the output.

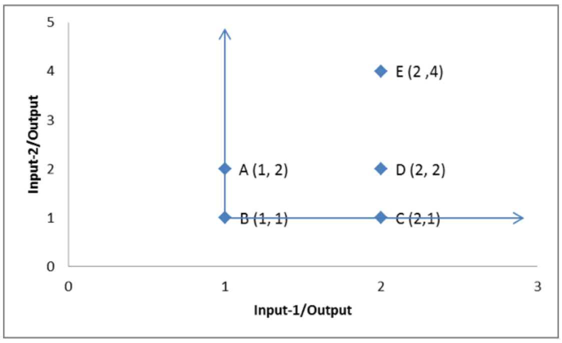

Figure 1: Data representation.

One can visually infer from Figure 1 that DMUs A, B, and C form the efficient frontier and DMUs D and E are inefficient. It can also be observed that though DMUs A and C are efficient, still there is room for improvement by means of reducing their respective inputs. We shall further detail these observations.

Solving System (3) in two phases using the given inputs and output data yields the following results in Tables 1-2, where Table 1 is obtained using System (5).

Table 1: Analytics of DMUs (Phase I)

DMU	λ^*	z_A	z_B	z_C	z_D	z_E
A	1	1	0	0	0	0
B	1	0	1	0	0	0
C	1	0	1	0	0	0
D	0.5	0	1	0	0	0
E	0.5	0	1	0	0	0

Based on Table 1, one can infer that, indeed, DMUs A, B, and C are forming a frontier as the respective λ^* attain unity. DMUs D and E, on the other hand, are inefficient, with the respective λ^* being not equal to unity. Furthermore, from the z_B column in Table 1, one can observe that DMU B is the peer for DMUs C, D, and E, but not for DMU A. Table 2 has been generated by solving System (6) with the additional information λ^* that was obtained from System (5).

Table 2: Analytics of DMUs (Phase II)

DMU	λ^*	z_A	z_B	z_C	z_D	z_E	β_1	β_2	Status
A	1	0	1	0	0	0	0	1	E
B	1	0	1	0	0	0	0	0	E_p
C	1	0	1	0	0	0	1	0	E
D	0.5	0	1	0	0	0	0	0	IE
E	0.5	0	1	0	0	0	0	1	IE

Note: E_p – Efficient (perfectly); E – Efficient; IE – Inefficient.

For DMU B, $\lambda^* = 1$ and β_1 and β_2 are zero; hence, by definition, B is an efficient (perfectly) DMU. For DMUs A and C, $\lambda^* = 1$ and β_2 and β_1, respectively, are non-zero; hence, by definition, DMUs A and C are efficient; nevertheless, there is still room for improvement by means of reducing their respective inputs. For DMUs D and E, $\lambda^* \neq 1$; hence, as previously mentioned, DMUs D and E are inefficient. It is evident from the z_B column in Table

2 that DMU B is the peer for the efficient and inefficient DMUs. In the case of the efficient DMUs A and C, β_1 and β_2 can be interpreted as the number of units in $\lambda^* x_1$ and $\lambda^* x_2$, respectively, that can be decreased to push DMUs A and C to be efficient (perfectly) like DMU B at its input level. Similarly, in the case of the inefficient DMUs D and E, the corresponding βs can be interpreted as the number of units in $\lambda^* x_1$ and $\lambda^* x_2$, respectively, that can be decreased to push DMUs D and E to be efficient (perfectly) like DMU B at its input level.

4. Concluding remarks

The non-Archimedean infinitesimal, ϵ, was introduced in order to avoid the value of zero for the weights in DEA models. In this short communication, we have shown a way to void ϵ using a directional distance function approach. The limitation of the present study consists in the fact that the proposed method still solves the problem in two phases. The advantage, however, is that by voiding ϵ, one can avoid the complexities posed by the many computational efforts.

Acknowledgements

The authors would like to thank both the editors and five anonymous reviewers for their constructive comments on the previous version of this short communication.

References

[1] Ali, A. I., and Seiford, L. M. (1993). Computational accuracy and infinitesimals in data envelopment analysis. Information Systems and Operational Research, 31(4), 290–297.

[2] Alirezaee, M. R. (2005). The overall assurance interval for the non-Archimedean epsilon in DEA models: A partition base algorithm. Applied Mathematics and Computation, 164(3), 667–674.

[3] Amin, G. R., and Toloo, M. (2004). A polynomial-time algorithm for finding ϵ in DEA models. Computers and Operations Research, 31(5), 803–805.

[4] Charnes, A., Cooper, W. W., and Rhodes, E. (1979). Short communication: measuring the efficiency of decision-making units. European Journal of Operational Research, 3(4), 339.

[5] Cooper, W. W., Seiford, L. M., and Tone, K. (2006). Introduction to Data Envelopment Analysis and Its Uses: with DEA-Solver Software and References. New York, NY: Springer-Verlag.

[6] Färe, R., and Grosskopf, S. (2010a). Directional distance functions and slacks-based measures of efficiency. European Journal of Operational Research, 200(1), 320-322.

[7] Färe, R., and Grosskopf, S. (2010b). Directional distance functions and slacks-based measures of efficiency: Some clarifications. European Journal of Operational Research, 206(3), 702.

[8] Färe, R., Grosskopf, S., and Margaritis, D. (2016). Advances in Data Envelopment Analysis. Singapore: World Scientific Publishing Co.

[9] Farrell, M. J. (1957). The measurement of productive efficiency. Journal of the Royal Statistical Society: Series A, 120(3), 253-281.

[10] Jahanshahloo, G. R., and Khodabakhshi, M. (2004). Determining assurance interval for non-Archimedean element in the improving outputs model in DEA. Applied Mathematics and Computation, 151(2), 501- 506.

[11] Mehrabian, S., Jahanshahloo, G. R., Alirezaee, M. R., and Amin, G. R. (2000). An assurance interval of the non-Archimedean epsilon in DEA models. Operations Research, 48(2), 344–347.

[12] Podinovski, V. V., and Bouzdine-Chameeva, T. (2017). Solving DEA models in a single optimization stage: Can the non-Archimedean infinitesimal be replaced by a small finite epsilon? European Journal of Operational Research, 257(2), 412-419.

[13] Thompson, R. G., Dharmapala, P. S., and Thrall, R. M. (1993). Importance for DEA of zeros in data and multipliers. Journal of Productivity Analysis, 4(4), 379–390.

[14] Tone, K. (1993). An Epsilon-Free DEA and a New Measure of Efficiency. Journal of the Operations Research Society of Japan, 36(3), 167-174.

[15] Zieschang, D. K. (1984). An extended Farrell technical efficiency measure. Journal of Economic Theory, 33(2), 387-396.

From: journals@theorsociety.com [journals@theorsociety.com]
Sent: 02 June 2017 09:36
To: V Charles
Subject: JORS-15-0292-PR2 Decision Letter

Dear Prof. Charles:

Re: tracking number JORS-15-0292-PR2: article title "A Directional Distance Function Approach to Void the Non-Archimedean in DEA"

With reference to your paper submitted to Journal of the Operational Research Society, it is my pleasure to be able to inform you that the paper has been accepted for publication, subject to standard editing.

You have the opportunity to have your accepted manuscript published Open Access as part of the Palgrave Open initiative. On payment of an Article Processing Charge your article would be made available to non-subscribers immediately upon publication. Palgrave Open articles are made freely available online and can be freely redistributed under a Creative Commons license. Full details about Palgrave Open are available at www.palgrave-journals.com/palgraveopen/index.html.

If you want to consider this option the publisher will need to know as soon as possible and ideally within 2 working days of receipt of this message. Please contact forms@palgrave.com for more information.

If you are not requesting that your article is published Open Access you will need to sign a standard author Copyright Assignment Form instead before it can be published. Please download the form from www.palgrave-journals.com/pal/authors/cafs/JORS.pdf.

Please read and sign the Copyright Assignment form for your article transferring copyright to the OR Society. Copyright forms should be returned promptly to the publisher at journalscopyright@palgrave.com

The Editors reserve the right not to publish articles or to require further revisions if copyright forms are not returned within three months of acceptance.

Once the manuscript has been copy edited and typeset you will receive an email directing you to our online proofing site where you will be able to check the PDF proofs of your article. Please read and follow the instructions carefully and send back any changes as directed. This will be your last chance to check the paper.

Thank you once again for your contribution to Journal of the Operational Research Society. I look forward to seeing it in print, and to receiving further submissions from you in the future.

You might also be interested to know that the Operational Research Society has a book series:http://www.palgrave.com/it/series/14725

If you would like to discuss how to submit a proposal, please email:
Josephine.Taylor@palgrave.com

With kind regards,

Thomas Archibald and Jonathan Crook
Editors, Journal of the Operational Research Society
journals@theorsociety.com