

Data envelopment analysis and its related linear programming models

Rolf Färe^{1,2} · Shawna Grosskopf¹ · Giannis Karagiannis³ ·
Dimitris Margaritis⁴

Published online: 3 November 2015
© Springer Science+Business Media New York 2015

Abstract We provide a unifying framework synthesizing the dual spaces of production and value used in DEA efficiency measurement with some well-known linear programming (LP) problems. Specifically, we make use of the technology matrix to map intensity variables into input–output space, and the adjoint transformation of the technology matrix to map input–output prices into prices of intensity variables. We show how the diet problem, a classical LP problem, is related to DEA and also use the adjoint matrix to demonstrate a procedure for pricing efficient decision-making units. We further illustrate the relationship between benefit-of-the-doubt aggregation and the diet problem.

Keywords DEA · Diet problem · Benefit-of-the-doubt · Primal · Dual · Linear programming

1 Introduction

The idea that economic models have a primal (quantity) and a dual (price) representation was developed by R.W. Shephard to address the data realities available to those pursuing applied research in economics. He noted that ‘statistical studies of cost functions are generally more accessible than corresponding empirical investigations of production functions, because economic data are most frequently in price and monetary terms’ (Shephard 1953, p. 28). Dorfman, Samuelson and Solow elaborate on this idea in their 1958 book, p. 38:

We thank two anonymous referees for their comments.

✉ Giannis Karagiannis
karagian@uom.gr

¹ Department of Economics, Oregon State University, Corvallis, OR 97331, USA

² Department of Applied Economics, Oregon State University, Corvallis, OR 97331, USA

³ Department of Economics, University of Macedonia, 156 Egnatia St., 54006 Thessaloniki, Greece

⁴ Department of Accounting and Finance, University of Auckland, Auckland, New Zealand

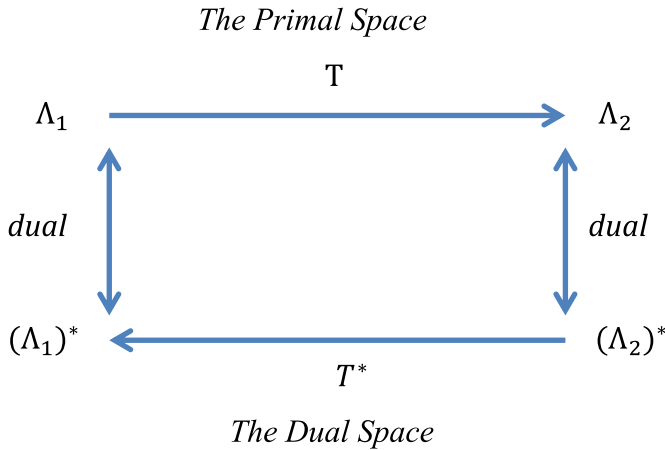


Fig. 1 Linear mapping in primal and dual space

It happens that mathematical linear-programming problems come in pairs; every mathematical linear-programming problem is intimately related to another problem called its “dual”. This statement would be no more than an interesting mathematical curiosity if it were not the fact that if an economic problem can be formulated as a linear-programming problem, then there will generally be a related economic problem that corresponds to the dual.

Linear programming (LP) and Data Envelopment Analysis (DEA)¹ models also have primal and dual formulations, which can be elucidated in the four-corner diagram in Fig. 1:²

In Fig. 1 we let Λ_1 and Λ_2 be two linear spaces and T a linear mapping from Λ_1 into Λ_2 . The dual spaces are denoted by $(\Lambda_1)^*$ and $(\Lambda_2)^*$ and the adjoint operator is T^* mapping $(\Lambda_2)^*$ into $(\Lambda_1)^*$. The mapping in the primal space is:

$$T\lambda_1 \geq \lambda_2, \lambda_1 \in \Lambda_1, \lambda_2 \in \Lambda_2$$

and its adjoint operator in the dual space is:

$$T^*\lambda_2^* \geq \lambda_1^*, \lambda_1^* \in (\Lambda_1)^*, \lambda_2^* \in (\Lambda_2)^*.$$

Shephard (1970, pp. 11–12) also noted ‘the dual problems in linear programming, one for imputing input prices and another for imputing output prices, are special forms of the first two of the three dualities introduced for determination of shadow prices, with more restrictive constraints for the prices imputed’.

Although in DEA the mapping Λ_1 into Λ_2 is widely used in applied work, only recently have there been attempts to relate its structural formulation to other known LP modes. Färe and Grosskopf (2002) explore the relation between the primal DEA model and Shephard’s distance function. Färe et al. (2011) relate the diet problem to a DEA version of the profit maximization problem. Färe et al. (2013) and Färe and Zelenyuk (2015) study the relationship between linear programming theory and DEA-based pricing of decision-making units (DMU). Färe and Karagiannis (2014) explore the relation between benefit-of-the-doubt

¹ Charnes et al. (1978).

² This is one of our interpretations of Magill and Quinzii (1996), Figure 13.1.

(BoD) aggregation and the diet problem. As we will show in this paper, all these problems can be integrated into our four-corner diagram. To do this we employ the adjoint transformation of technology, a generalization of the more familiar primal and dual representations of technology developed by Shephard (1953, 1970).

2 Main results

Denote inputs by $x \in \mathbb{R}_+^N$ and outputs by $y \in \mathbb{R}_+^M$ and their corresponding prices by $w \in \mathbb{R}_+^N$ and $p \in \mathbb{R}_+^M$, respectively. In addition, let the intensity vectors be $z \in \mathbb{R}_+^K$ and their corresponding prices denoted by $q \in \mathbb{R}_+^K$. By noticing that the dual space of the real numbers is the set of real numbers, i.e., $\mathbb{R} = \mathbb{R}^*$, our four-corner diagram becomes Fig. 2.

In the finite dimension case, the adjoint operator T^* is just the transpose of T . Now with T being the technology matrix, the primal mapping transforms intensity variables into input/output vectors:

$$Tz \geq (x, y)$$

where the DEA technology is of the form:

$$\begin{aligned} \{(x, y) : \sum_{k=1}^K z_k x_{kn} \leq x_n, \quad n = 1, \dots, N, \\ \sum_{k=1}^K z_k y_{km} \geq y_m, \quad m = 1, \dots, M, \\ z_k \geq 0, k = 1, \dots, K\} \end{aligned} \tag{1}$$

using the input–output data for the $k = 1, \dots, K$ DMU:s. The intensity variables are $z_k, k = 1, \dots, K$, one for each DMU.

On the other hand, the dual mapping T^* transforms input and output prices into the prices of the intensity variables. The DEA formulation of the adjoint mapping is given by

$$\left\{ (q_1, \dots, q_k) : \sum_{m=1}^M p_m y_{km} - \sum_{n=1}^N w_n x_{kn} \leq q_k, k = 1, \dots, K, q_k \geq 0 \right\}. \tag{2}$$

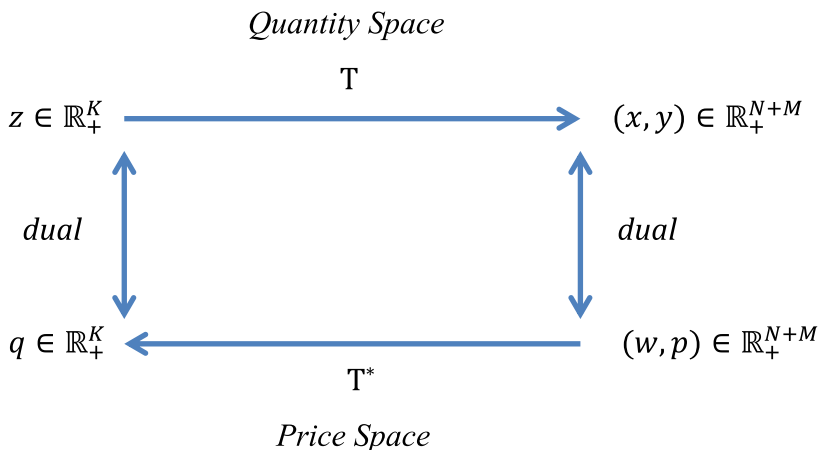


Fig. 2 Adjoint transformations in primal and dual space

With the primal technology from (1) we may introduce a number of optimization problems including:

Cost Minimization:

$$\min_x wx \text{ s.t. (1)}$$

Revenue Maximization:

$$\max_y py \text{ s.t. (1)}$$

and Profit Maximization³

$$\max_{x,y} py - wx \text{ s.t. (1)}$$

In this paper the primal problem we choose to study is the DEA formulation of cost minimization, thus for DMU k' , the individual DMU under evaluation, we have

$$\begin{aligned} \min_{x,z} wx \text{ s.t. } & \sum_{k=1}^K z_k x_{kn} \leq x_n, \quad n = 1, \dots, N. \\ & \sum_{k=1}^K z_k y_{km} \geq y_{k'm}, \quad m = 1, \dots, M. \\ & z_k \geq 0, k = 1, \dots, K \end{aligned} \tag{3}$$

Note that we are solving for the cost minimizing input vector, hence there is no k' on x on the RHS of the input constraint.

Associated with this problem there is a dual linear programming problem for each DMU, referred to by Färe and Grosskopf (2002) as the Shephard output pricing model, see also Shephard (1970, p. 290):

$$\max_p py_{k'} \text{ s.t. } \sum_{m=1}^M p_m y_{k'm} - \sum_{n=1}^N w_n x_{k'n} \leq 0, \tag{4}$$

where the solution yields p^* , which is the shadow price vector for the outputs y of the DMU under evaluation.

The dual mapping in (2) is a model for pricing DMUs.⁴ For example, if input and output prices are known then the ‘price’ of DMU k' equals

$$q_{k'} = \sum_{m=1}^M p_m y_{k'm} - \sum_{n=1}^N w_n x_{k'n}. \tag{5}$$

This approach for pricing DMUs may be used to retrieve the virtual price of an efficient DMU and in this sense provide useful input in determining the premium (discount) an acquiring firm may be willing to pay for a target company being considered under mergers and acquisitions.

In Fig. 3 we have summarized the optimization problems associated with the four corners. The optimization problems on the left-hand side of the box have the intensity variables or their prices as choice variables and observed (data) variables as controls. In contrast, the LP problems on the right-hand side are augmented by optimizing over quantities or their prices. Specifically, the optimization problems at the top have quantity variables as choice variables while those on the bottom have price variables (either observed or virtual) as choice variables.

³ In the profit maximization problem we may want to restrict the intensity variables to satisfy $\sum_{k=1}^K z_k \leq 1$ or $= 1$ in order for the maximum to exist.

⁴ Originally developed by Färe et al. (2013).

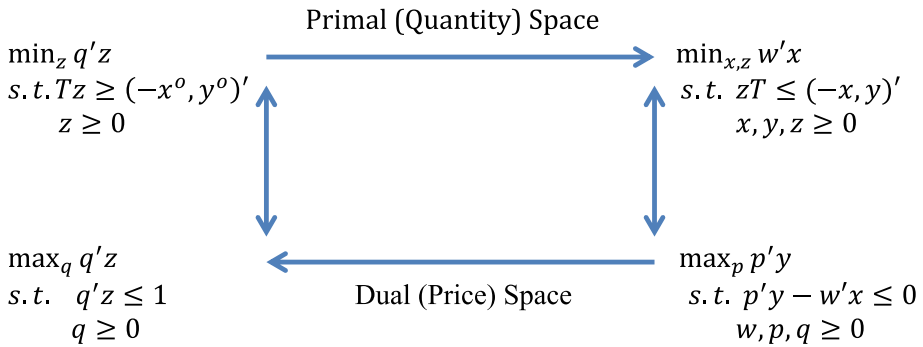


Fig. 3 Adjoint transformations—primal and dual technologies

Beginning with the upper left hand side in Fig. 3 we have the diet problem. We think of the technology matrix T as being nutrient contents of the food items included in a diet (and their associated inputs), the intensity variables z being the quantities of food items consumed, q being their market prices, and the targeted data vector (x^o, y^o) being the nutritional requirements (y^o) along with limits (x^o) on inputs (food ingredients) used to produce these nutrients. Hence the left top corner LP minimization problem may be viewed as a compact form of one of the earliest LP problems, namely the diet problem as formulated by Stigler (1945). Following Dorfman et al. (1958, pp. 9–24) and its extension by Färe et al. (2011) which allows for the outputs to be produced by inputs, we may write this as:

$$\begin{aligned}
 & \min_{z_k} \sum_{k=1}^K q_k z_k & (6) \\
 \text{s. t. } & \sum_{k=1}^K z_k x_{kn} \leq x_{k'n}, \quad n = 1, \dots, N \\
 & \sum_{k=1}^K z_k y_{km} \geq y_{k'm}, \quad m = 1, \dots, M \\
 & z_k \geq 0, \quad k = 1, \dots, K
 \end{aligned}$$

In the diet problem we seek to find the minimum cost diet for a typical adult that satisfies certain nutritional requirements. These reflect health standards determined by the minimum amount of vitamins and other nutrients (e.g., carbohydrates, protein, minerals, dietary fat) that an adult requires during a particular period of time. The nutritional component of food items, i.e., the elements of the T matrix, give the constant amount of each nutritional element contained in every unit of any given food item independent of the other food items that may be consumed simultaneously. Then solving the above problem will result in the least cost combination of food quantities that satisfy a number of nutritional requirements for given price and nutritive values of each food item included in the diet.⁵

⁵ Färe et al. (2011) have shown by means of formulating a Lagrangian–Kuhn–Tucker problem that the diet problem can be considered to be dual to the DEA profit maximization problem in the sense that the intensity variables in the profit maximization model are shadow prices in the diet problem.

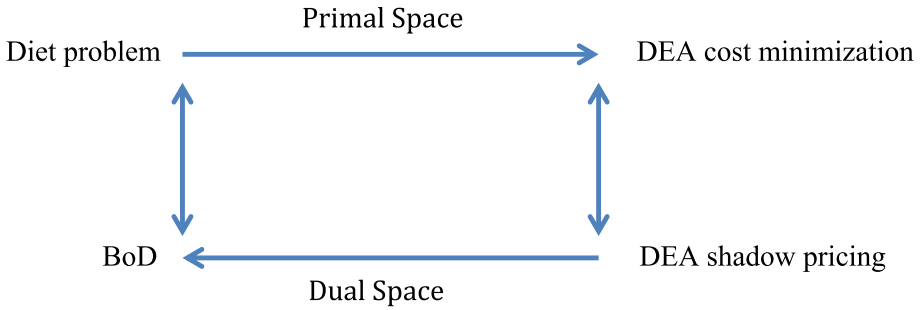


Fig. 4 The four LP problems

The LP dual of the original diet problem (3) is written for observation k' as

$$\begin{aligned}
 & \max_{p,w} \sum_{m=1}^M p_m y_{k'm} - \sum_{n=1}^N w_n x_{k'n} & (7) \\
 & \text{s.t. } \sum_{m=1}^M p_m y_{km} - \sum_{n=1}^N w_n x_{kn} \leq q_k, \quad k = 1, \dots, K \\
 & \quad p_m \geq 0, \quad m = 1, \dots, M \\
 & \quad w_n \geq 0, \quad n = 1, \dots, N
 \end{aligned}$$

which gives the maximum value of nutrient requirements subject to the condition that the imputed marginal value of each food item is less than or equal to its market price.

Färe and Karagiannis (2014) have shown that the dual formulation of the diet problem (7) is equivalent to the primal formulation of the benefit-of-the-doubt (BoD) model as long as food prices are set equal to one, i.e., $q_k = 1, k = 1, \dots, K$ and the aggregation weights in the BoD model are equal to input and output prices. This implies that for DMU k'

$$\begin{aligned}
 & \max_{s_{ki}} \sum_{m=1}^M s_m y_{k'm} - \sum_{n=1}^N s_n x_{k'n} & (8) \\
 & \text{s.t. } \sum_{m=1}^M s_m y_{k'm} - \sum_{n=1}^N s_n x_{k'n} \leq 1, \quad k = 1, \dots, K \\
 & \quad s_n, s_m \geq 0, \quad n = 1, \dots, N, \quad m = 1, \dots, M
 \end{aligned}$$

where s_m, s_n refer to aggregation weights. In the BoD framework, (x, y) should be interpreted as sub-indicators to be aggregated into a composite indicator.⁶ By solving the above model we obtain a set of unit-specific shares assigning less (more) weight to those sub-indicators for which the assessed DMU has relatively weak (strong) performance compared to other units in the sample and which are used to construct the composite indicator.

Färe and Karagiannis (2014) verify that the diet problem is equivalent to the dual formulation of the BoD as long as food prices are equal to one and the intensity variables are equal to z_k . From this one can verify that the BoD is equivalent to the input-oriented DEA model when there is a single constant input that takes the value of one for all DMUs (see Lovell and Pastor 1999; Liu et al. 2011).

This completes the picture of how the four LP programs are related to each other and how each one of them corresponds under certain circumstances to a well-known LP problem, which is summarized in Fig. 4.

⁶ The importance of the BoD model in constructing composite performance indicators is emphasized by the OECD (2008).

3 Conclusion

Via the means of a four-corner schematic representation we have provided a unifying framework that compares and contrasts duality familiar from linear programming to that of production theory. We used the technology matrix to map intensity variables into the input–output space, and the adjoint transformation, a linear programming type of mapping of observed inputs and outputs of DMUs into price space, as the dual approach to the quantity problem. Specifically, we used the technology matrix transformation to show how the diet problem, a classical LP problem, may be related to DEA cost minimization for given input prices. And we also used the adjoint matrix to demonstrate a procedure for obtaining valuation or pricing of DMUs. We have further shown that the diet problem and the benefit-of-the-doubt aggregation are linear programming duals. Finally, we have shown how the LP dual to the cost minimization problem may be viewed as the revenue maximization problem in the dual (price) space where the DMUs maximize revenue by choosing the output prices for given input and output quantities.

References

- Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research*, 2, 429–444.
- Cherchye, L., Moesen, W., Rogge, N., & van Puyenbroeck, T. (2007). An introduction to “benefit of the doubt” composite indicators. *Social Indicators Research*, 82, 111–145.
- Dorfman, R., Samuelson, P. A., & Solow, R. M. (1958). *Linear programming and economic analysis*. New York: McGraw-Hill.
- Färe, R., & Grosskopf, S. (2002). Two perspectives on DEA: Unveiling the link between CCR and Shephard. *Journal of Productivity Analysis*, 17, 41–47.
- Färe, R., & Karagiannis, G. (2014). Benefit-of-the-doubt aggregation and the diet problem. *Omega*, 47, 33–35.
- Färe, R., & Zelenyuk, V. (2015). Pricing of decision making units under non-constant returns to scale. *Journal of Operational Research Society*, 66, 172–173.
- Färe, R., Grosskopf, S., & Margaritis, D. (2013). Pricing decision-making units. *Journal of Operational Research Society*, 64, 612–619.
- Färe, R., Grosskopf, S., & Margaritis, D. (2011). The diet problem and DEA. *Journal of Operational Research Society*, 62, 1420–1422.
- Farrell, M. J. (1957). The measurement of productive efficiency. *Journal of Royal Statistical Society Series A*, 120, 253–290.
- Lovell, C. A. K., & Pastor, J. T. (1999). Radial DEA models without inputs or without outputs. *European Journal of Operational Research*, 118, 46–51.
- Liu, W. B., Zhang, D. Q., Meng, W., Li, X. X., & Xu, F. A. (2011). A study of DEA models without explicit inputs. *Omega*, 39, 472–480.
- Magill, M., & Quinzii, M. (1996). *Theory of incomplete markets*. Cambridge, MA: MIT Press.
- OECD. (2008). *Handbook on constructing composite indicators: Methodology and user guide*. Paris: OECD.
- Shephard, R. W. (1953). *Cost and production functions*. Princeton: Princeton University Press.
- Shephard, R. W. (1970). *Theory of cost and production functions*. Princeton: Princeton University Press.
- Stigler, G. J. (1945). The cost of subsistence. *Journal of Farm Economics*, 27, 303–314.