

DATA ENVELOPMENT ANALYSIS AND ITS RELATED LINEAR PROGRAMMING MODELS

Rolf Färe¹, Shawna Grosskopf², Giannis Karagiannis³ and Dimitris Margaritis⁴

¹ Professor, Department of Economics and Department of Agricultural & Resource Economics, Oregon State University, Corvallis, OR, USA; rolf.fare@oregonstate.edu

² Professor Emerita, Department of Economics, Oregon State University, Corvallis, OR, USA and CERE, Umea, Sweden; shawna.grosskopf@orst.edu

³ Professor, Department of Economics, University of Macedonia, 156 Egnatia Str. Thessaloniki, 54006 GREECE; karagian@uom.gr

⁴ Professor, Department of Accounting and Finance, University of Auckland Business School, 12 Grafton Rd, Auckland 1010, New Zealand; d.margaritis@auckland.ac.nz

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 use the adjoint transformation to 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 (DMUs). 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

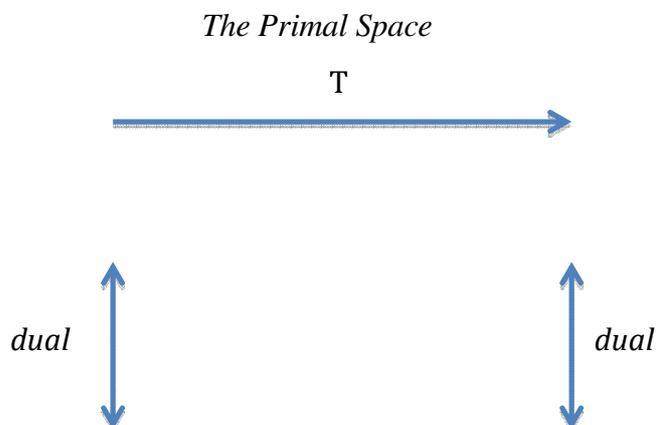
May 2015

DATA ENVELOPMENT ANALYSIS AND ITS RELATED LINEAR PROGRAMMING MODELS

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)

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



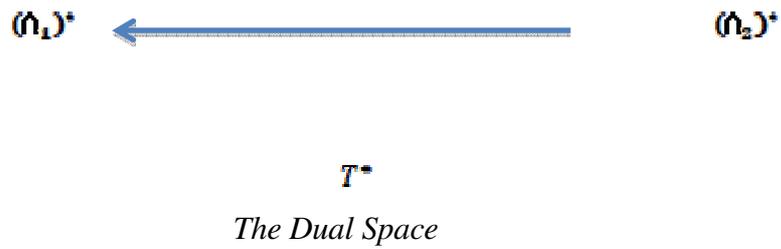


Figure 1: Linear Mapping in Primal and Dual Space

Let Ω_1 and Ω_2 be two linear spaces and T a linear mapping from Ω_2 into Ω_1 .

The dual spaces are denoted by Ω_1^* and Ω_2^* and the

adjoint operator is T^* mapping Ω_2^* into Ω_1^* . The mapping in the primal space is:

$$T\lambda_1 \geq \lambda_2, \quad \lambda_1 \in \Omega_1, \quad \lambda_2 \in \Omega_2$$

and its adjoint operator in the dual space is:

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 DEA mapping Ω_2 into Ω_1 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, Grosskopf and Margaritis (2011) relate the diet problem to a DEA version of the profit maximization problem. Färe, Grosskopf and Margaritis (2013) and Färe and Zelenyuk (in press) 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 (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-corners diagram becomes:

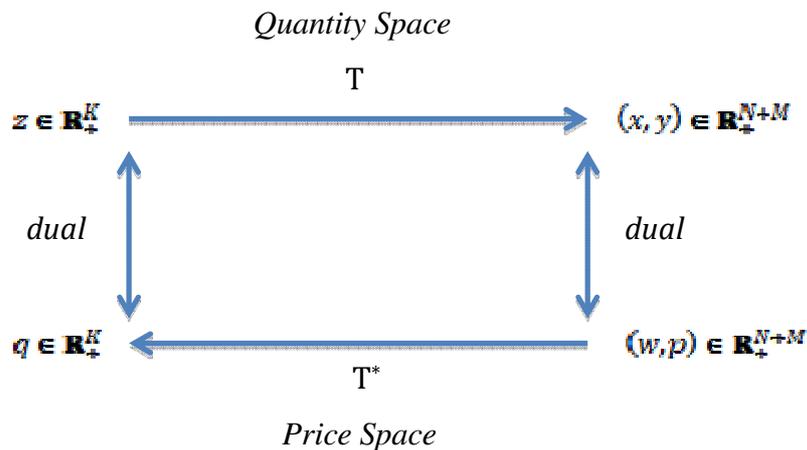


Figure 2: Adjoint Transformations in Primal and Dual Space

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:

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

and the dual mapping translates input/output prices into prices of the intensity variables:

$$T^*(w, p) \leq q$$

The aforementioned variables, namely x , y , w , p , z and q , may be related by means of four generic LP problems, which can be portrayed in our four-corners diagram as follows:

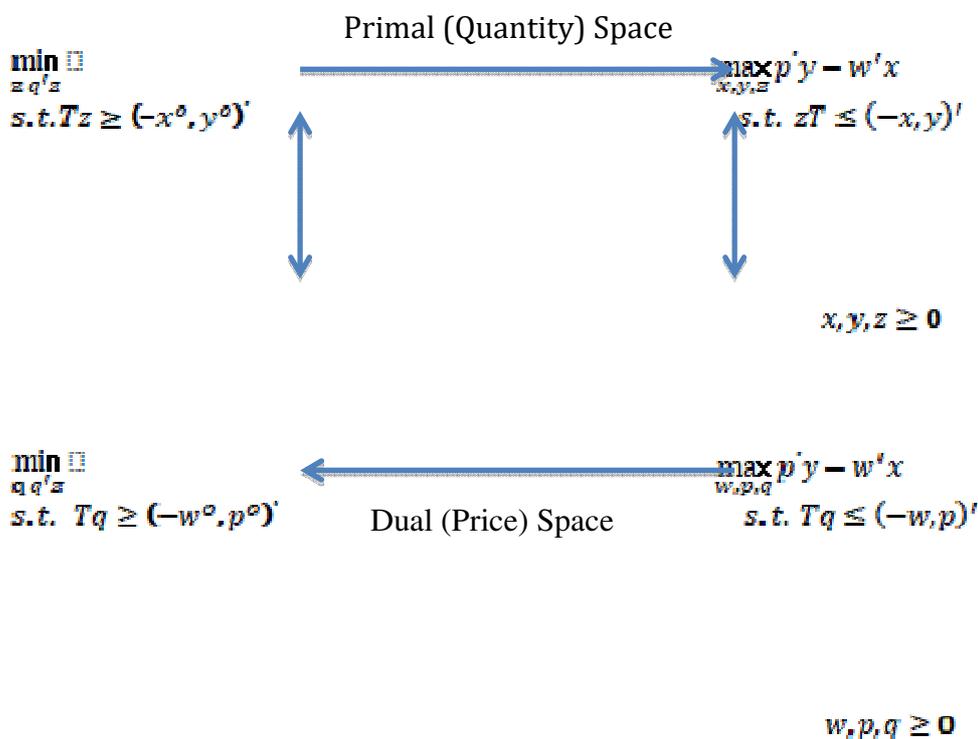


Figure 3: Adjoint Transformations - Primal and Dual Technologies

The optimization problems on the left-hand side have the intensity variables as choice variables and observed (data) variables as controls. In contrast, the LP problems in the right-hand side are augmented by optimizing over quantities. 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.

Perhaps the most interesting observation is that these otherwise generic formulations correspond to well-known LP problems. *First*, if you 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^0, y^0) being the nutritional requirements (y^0) along with limits (x^0) on inputs (food ingredients) used to produce these nutrients,³ 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, Samuelson and Solow (1958, pp. 9-24) and its extension by Färe, Grosskopf and Margaritis (2011) that allows for the outputs to be produced by inputs, we may write this as:

$$\begin{aligned}
 & \min_{z_i} \sum_{i=1}^K q_i z_i \\
 \text{s.t. } & \sum_{i=1}^K z_i x_{ii} \leq x_i^0 \quad \forall i = 1, \dots, N \\
 & \sum_{i=1}^K z_i y_{ij} \geq y_j^0 \quad \forall j = 1, \dots, M \\
 & z_i \geq 0 \quad \forall i = 1, \dots, K
 \end{aligned} \tag{1}$$

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.

Second, the technology matrix T can be used to trace out the production technology and in this case the right top LP problem corresponds to the traditional linear programming model of the firm (see Dorfman, Samuelson and Solow, 1958; Ch. 6), which in standard DEA formulation takes the form:

$$Tz \geq (x, y) = \left\{ (x, y) : \sum_{i=1}^K z_i x_{ii} \leq x_i \quad \forall i = 1, \dots, N, \sum_{i=1}^K z_i y_{ij} \geq y_j \quad \forall j = 1, \dots, M, z_i \geq 0 \right\}$$

The last constraint on the intensity variables imposes constant returns to scale in the underlying technology.⁴ Assuming profit maximization the corresponding DEA model is:

$$\begin{aligned} & \max_{x, y, z} p'y - w'x \\ & \text{s. t. } \sum_{i=1}^K z_i x_{ii} \leq x_i \quad \forall i = 1, \dots, N \\ & \sum_{i=1}^K z_i y_{ij} \geq y_j \quad \forall j = 1, \dots, M \\ & x_i, y_i, z_i \geq 0 \quad \forall i = 1, \dots, K \end{aligned} \quad (2)$$

In the profit maximization problem of the firm we seek to find the combination of input and output quantities that maximize the difference between revenue and cost for given input and output prices.

Even though seemingly dissimilar since the diet problem is a minimization problem while the profit DEA problem is a maximization problem and in addition the two have different choice variables, the diet and the profit maximization problems are related to each other.⁵ In particular, Färe, Grosskopf and Margaritis (2011) have shown by means of formulating a Lagrangian-Kuhn-Tucker problem that models (1) and (2) are dual in the sense that the intensity variables (z) in the DEA model (2) are shadow prices in the diet problem (1).

Third, with the same interpretation of the data as in the above case, the right bottom corner LP problem may be viewed as the profit maximization problem in the dual (price) space where the DMUs seek to maximize the difference between revenue and cost by choosing the input and output prices for given input and output quantities. In standard DEA formulation this problem is:

$$\begin{aligned} & \max_{w, p, q} p'y - w'x \\ & \text{s. t. } \sum_{i=1}^K q_i x_{ii} \leq w_i \quad \forall i = 1, \dots, N \end{aligned} \quad (3)$$

$$\sum_{i=1}^K q_i y_{ij} \geq p_j \quad \forall j = 1, \dots, M$$

$$w_l, p_l, q_l \geq 0 \quad \forall l = 1, \dots, K$$

where the q 's have the role of the intensity variables. This model may be viewed as the profit maximization analogue of the price efficiency model of Färe (1984) and Färe, Grosskopf and Nelson (1990) used to provide estimates of dual efficiency in situations where firms maximize profits with respect to shadow rather than observed prices. A practical application of this dual (price space) formulation is the pricing of DMUs (see Färe, Grosskopf and Margaritis, 2013). This approach, for example, may be used to retrieve the virtual price of an efficient DMU and in this sense provide valuable input in determining the premium (discount) an acquiring firm may be willing to pay for a target company.

Lastly, if you consider the technology matrix T as being a set of sub-indicators that are to be aggregated into a composite indicator, set the intensity vector equal to one and let the targets to be achieved be determined by the price $(w, p) > 0$ data, then the left bottom corner LP minimization problem may be viewed as the dual (envelopment) form of the BoD model.⁶ In a more explicit form it may be written as:

$$\begin{aligned} \min_{q_l} \quad & \sum_{l=1}^K q_l \mathbf{1}^l \\ \text{st} \quad & \sum_{l=1}^K q_l x_{il} \leq w_i^0 \quad \forall i = 1, \dots, N \\ & \sum_{l=1}^K q_l y_{lj} \geq p_j^0 \quad \forall j = 1, \dots, M \\ & q_l \geq 0 \quad \forall l = 1, \dots, K \end{aligned} \quad (4)$$

As can be seen from (4), the BoD model 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). By solving the above model we obtain a set of unit-specific shares, depicted by the dual intensity variable q , assigning less (more) weight to those sub-indicators for which the assessed DMU has a relatively

weak (strong) performance compared to other units in the sample and which are used to construct the composite indicator.

Moreover, Färe and Karagiannis (2014) verified that the diet problem in (1) is equivalent to the dual (envelopment) form of the BoD model in (4) as long as food prices are set equal to one, the intensity variables are set equal to food quantities and the right-hand side variables in the first two set of constraints in (4) are viewed as **given** targets to be achieved, i.e., nutritional requirements. In that sense the diet problem and the BoD model are linear programming duals. **Similarly, the BoD model in (4) may be thought of as the analogue of (3) with intensity variables whose value is unity, which in the Farrell (1957) framework refers to efficient DMUs.**⁷

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. That is,

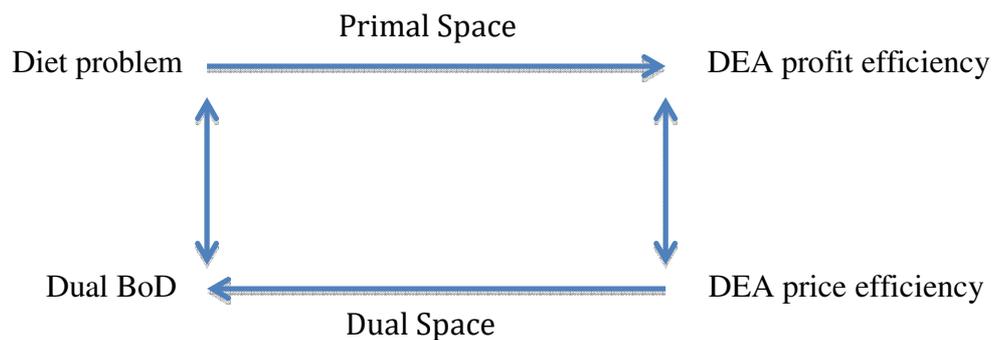


Figure 4: The Four LP Problems

3. Conclusion

We have provided via the means of a four-corner schematic representation 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 space. Specifically, we used the adjoint transformation to show how the Diet Problem, a classical LP problem, is related to DEA profit efficiency for given input and output 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 problem may be viewed as the profit maximization problem in the dual (price) space where the DMUs maximize the difference between revenue and cost by choosing the input and 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. and T. van Puyenbroeck. An Introduction to “Benefit of the Doubt” Composite Indicators, *Social Indicators Research*, 2007, 82, 111-145.
- Dorfman, R., Samuelson, P.A. and R.M. Solow. *Linear Programming and Economic Analysis*, McGraw-Hill, N.Y., 1958.
- Färe, R. The Dual Measure of Efficiency, *Journal of Economics*, 1984, 44, 283-288.
- Färe, R. and S. Grosskopf. Two Perspectives on DEA: Unveiling the Link between CCR and Shephard, *Journal of Productivity Analysis*, 2002, 17, 41-47.
- Färe, R. and G. Karagiannis. Benefit-of-the-doubt Aggregation and the Diet Problem, *Omega*, 2014, 47, 33-35.
- Färe, R. and V. Zelenyuk. Pricing of Decision Making Units under Non-constant Returns to Scale, *Journal of Operational Research Society*, (in press).
- Färe, R., Grosskopf, S. and D. Margaritis. Pricing Decision-Making Units, *Journal of Operational Research Society*, 2013, 64, 619-612.
- Färe, R., Grosskopf, S. and D. Margaritis. The Diet Problem and DEA, *Journal of Operational Research Society*, 2011, 62, 1420-1422.
- Färe, R., Grosskopf, S. and J. Nelson. On Price Efficiency, *International Economic Review*, 1990, 31, 709-720.
- Farrell, M.J. The Measurement of Productive Efficiency, *Journal of Royal Statistical Society Series A*, 1957, 120, 253-90.
- Lovell, C.A.K and J.T. Pastor. Radial DEA Models without Inputs or without Outputs, *European Journal of Operational Research*, 1999, 118, 46-51.
- Liu, W.B., Zhang, D.Q., Meng, W., Li, X.X. and F.A. Xu. A Study of DEA Models without Explicit Inputs, *Omega*, 2011, 472-480.
- Magill, M. and M. Quinzii. *Theory of Incomplete Markets*, MIT Press, Cambridge MA, 1996.
- OECD. *Handbook on Constructing Composite Indicators: Methodology and User Guide*, 2008.

Shephard, R.W. *Cost and Production Functions*, Princeton University Press, Princeton, 1953.

Shephard, R.W. *Theory of Cost and Production Functions*, Princeton University Press, Princeton, 1970.

Stigler, G.J. The Cost of Subsistence, *Journal of Farm Economics*, 1945, 27, 303-314.

Endnotes

¹ Charnes, Cooper and Rhodes (1978).

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

³ Such limits may also be imposed by increased risks such as toxins, food tolerance or other health requirements.

⁴ Other forms of returns to scale may be imposed on the technology by adding restrictions to the intensity variables.

⁵ Note that in DEA profit maximization problem we are solving for the optimal (x,y) vector, that is, the ‘best’ vectors of the right-hand side of the inequality constraints in (2), while in the diet problem the right-hand side vector in (1) is given and we are solving for the ‘best’ combination of the foods (by choosing the z ’s or intensity variables) minimizing the diet cost. We do solve for the intensity variables as well in the profit maximization problem; however, they are not included in the objective function.

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

⁷ The BoD formulation given in (4) is a special case of the inner product of $q \in \mathbb{R}_+^K$ and $z \in \mathbb{R}_+^K$ shown in Figure 3, where for any $z_i \geq 0$, setting it equal to one and the other elements equal to zero, shows that q_i is the unit price of that DMU.