An Alternative Algorithm
for Detecting Anchor Points

F. Hosseinzadeh Lotfi \textsuperscript{a1}, G. R. Jahanshahloo\textsuperscript{a}, F. Moradi\textsuperscript{a} and F. Jannati\textsuperscript{b}

\textsuperscript{a}Department of Mathematics, Science and Research Branch
Islamic Azad University, Tehran, Iran

\textsuperscript{b}Department of Mathematics
University of Kurdistan, Sanandaj, Iran

Abstract

Bougnol and Dulá [M.L.Bougnol,J.H.Dulá, 2009. Anchor point in DEA. European Journal of Operational Research 192 (2009) 668-676] recently proposed an alternative algorithm to identify anchor points. Despite its innovative aspect, it is very expensive from computational point of view. In this paper, an algorithm which identifies a main part of anchor points without solving any model is proposed. also, the remaining anchor points are identified only by solving one model.

Keywords: Data envelopment analysis; Anchor points

1 Introduction

In a recent paper Bougnol and Dulá [5] proposed an algorithm, hereafter called BD-algorithm, for identifying anchor points which solved $m + s$ subproblems to identify that a point is anchor point or not. The aim of this short paper is to present an alternative model to that of Bougnol and Dulá [5]. It is shown that a main part of anchor points can be identified only with a simple comparison between inputs and outputs components. Also, a mixed integer linear programming is proposed for determining the remaining anchor points.

In the next section the BD-algorithm is reviewed. Section 3 contains the proposed model. There are some examples in section 4 and finally section 5 concludes the paper.

\textsuperscript{1}Corresponding author, Email, addresses: farhad@hosseinzadeh.ir
2 BD-model

Set $\mathcal{A} = \{a^1, \ldots, a^n\}$ includes the data of $n$ pints, one for each DMU in the model. A DMU’s data point is composed of two parts, the input components $0 \leq x_j = (x_{1j}, \ldots, x_{mj})^T \in \mathbb{R}^{m \times 1}$ and the output components $0 \leq y_j = (y_{1j}, \ldots, y_{sj})^T \in \mathbb{R}^{s \times 1}$ as follows:

$$a^j = \begin{bmatrix} -x_j \\ y_j \end{bmatrix} \in \mathbb{R}^{m+s \times 1}, \quad j \in J = \{1, \ldots, n\}$$

Let the set $I = \{1, \ldots, m\}$ include input index and the set $O = \{1, \ldots, s\}$ include output index. Also, we assume that $x_j \neq 0$ and $y_j \neq 0$ and that there is no duplication in $\mathcal{A}$. Our work will focus on the variable returns (VR) production possibility set ($\mathcal{P}^{VR}$):

$$\mathcal{P}^{VR} = \{z \in \mathbb{R}^{m+s \times 1} \mid \sum a^j \lambda_j \geq z; \quad \text{s.t.} \quad \sum \lambda_j = 1, \lambda_j \geq 0; \forall j\}.$$

The input oriented BCC\(^1\) model [7] and its dual are as follow, respectively:

$$\begin{align*}
\text{Min} \left\{ \theta \mid \sum a^j \lambda_j \geq a(\theta)^o, \quad \sum \lambda_j = 1, \lambda_j \geq 0; \forall j \right\} \quad \text{where} \quad a(\theta)^o = \begin{bmatrix} -\theta x_o \\ y_o \end{bmatrix} \\
\text{Max} \left\{ \pi \overline{a}^\sigma + \beta \mid \pi a^j \leq \beta, \forall j, 0 \leq \pi \in \mathbb{R}^{1 \times m+s} \right\} \quad \text{where} \quad \overline{a}^\sigma = \begin{bmatrix} 0 \\ y_o \end{bmatrix} \quad a^o = \begin{bmatrix} x_o \\ 0 \end{bmatrix}
\end{align*}$$

**Definition 1.** DMU\(_o\) is BCC efficient if $\theta^* = 1$ and in all optimal solution any slack variable equal to zero.

All DMUs are classified on the three categories as follows; F, inefficient (weak efficient and interior), E, efficient non-extreme and $E^*$, extreme-efficient. We define $J^* = \{j \mid a^j \in E^*\}$ where $J^*$ also called the frame of $\mathcal{A}$. See [1, 7, 8] for more detail.

**Definition 2.**

- The hyperplane $\mathcal{H}(\pi, \beta) \in \mathbb{R}^{m+s \times 1}$ is the set $\{z \mid \pi.z = \beta\}$. We refer to $\pi$ as the orthogonal vector of the hyperplane.

- A face of a polyhedral set is the support set of a supporting hyperplane.

---

\(^1\)Banker, Charnes and Cooper
• A facet of an $m$-dimensional polyhedral set is an $m - 1$ dimensional face.

**Definition 3.** The point $a^j \in E^* \subset A$ is an anchor point if it belongs to an unbounded face of $\mathcal{P}^{VR}$.

For more detail, you can refer to [1,2,6,7].

**Definition 4.** The $i$th simple projection of $\mathcal{P}^{VR}$, $\mathcal{P}^{VR}_i$ is the $VR$ production possibility set of the $n$ data points where the $i$th component has been omitted.

The proof of the two following lemmas could be found in Bougnol and Dulá [5].

**Lemma 1.** The point $a^j \in E^*$ is an anchor point if and only if it belongs to the support set of a supporting hyperplane $H(\hat{\pi}, \hat{\beta}) \in \mathbb{R}^{m+s \times 1}$ such that the orthogonal vector, $\hat{\pi}$, contains at least one zero.

**Lemma 2.** Consider the frame of the data set, $E^*$. The data point $a^j \in E^*$ is an anchor point if and only if it belongs to the boundary of at least one simple projection.

3 An alternate algorithm

3.1 A property of anchor points

In this section it is shown that anchor points have an intrusting property which helps us to identify a main part of anchor points without solving any model. Let $K_i = \{j \in J | a^i_j = \text{Max}_{j \in J} a^i_j\}$ for $i \in I \cup O$.

**Lemma 3.** If the point $a^\hat{j} \in E^*$ where $\hat{j} \in K_i$ for at least one $i$, then $a^\hat{j}$ is an anchor point.

**Proof.** Let $\hat{j} \in K_{i_1}, i_1 \in I$. With regard to the definition of $K_{i_1}$ if $\lambda \in \mathbb{R}^{m+s \times 1}$, $\lambda \geq 0$ and $\sum_j \lambda_j = 1$, then $\sum_j \hat{\lambda}x_{i_1,j} \geq x_{i_1,j}$. Hence $a^\hat{j}$ belongs to the boundary of $\mathcal{P}^{VR}_i$ for each $i \in I \cup O$, $i \neq i_1$ and then $a^\hat{j}$ is an anchor point considering Lemma 2. It can be proved in a similar way for $i_1 \in O$.

It means that if a point $a^j \in E^*$ has the best operation at least in one index, the minimum consumption in one input or the maximum production in one output, then it is an anchor point. The inverse of the above lemma may be invalid. See examples 1 and 2 in Bougnol and Dulá [5]. Units $a^5$ in example 1 and $a^{10}$ in example 2 are anchor points but they aren’t in any $K_i$. 
Let $K = \bigcup_{i=1}^{m+s} K_i$. The set $K \cap J^*$, according to Lemma 3, is a subset of anchor point set.

Suppose that $W_o$ contains optimal solutions of (2):

$$W_o = \{ \pi \mid \pi a^o + \beta = 1, \pi a^o = 1, \pi a^j \leq \beta, j \in J, \pi_i \geq 0, i \in I \cup O \}$$

According to Lemma 1, it is sufficient that we search among all optimal dual solutions for a solution which contains at least one zero. For this reason, we purpose the following model:

$$b^*_o = \min \left\{ \sum_{i \in O} b_i \mid \pi_i \leq Mb_i, b_i \in \{0, 1\}, i \in I \cup O, \pi \in W_o \right\} \quad (3)$$

$b_i$ is binary variable and $M$ is a large positive number. If $b_i = 0$, then $\pi_i = 0$ otherwise $\pi_i \leq M$ is a redundant constrain. Moreover $\sum_{I \cup O} b_i$ equals to the number of nonzero elements of $\pi$. It should be mentioned that $W_o$, considering its normalization constraint, is bounded.

**Lemma 4.** Let $a^o \in E^*$. $a^o$ is an anchor point if and only if $b^*_o < m + s$.

**Proof.** If $b^*_o < m + s$, then there is an optimal solution $\pi^* = (\pi^*_1, \ldots, \pi^*_m, \pi^*_{m+s})$, $b^* = (b^*_1, \ldots, b^*_m, b^*_{m+s})$ such that at least one $b^*_i$ equals to zero. Consequently, at least one $\pi^*_i$ equals zero. Therefor, $a^o$ is an anchor point. Conversely let $a^o \in E^*$ is an anchor point. According to Lemma 1 there is an optimal solution $\hat{\pi} \in W_o$ where $\hat{\pi}_i = 0$ for at least one $i$. Define $\hat{b}_i = 0$ if $\hat{\pi}_i = 0$ otherwise $\hat{b}_i = 1$. It is obvious that $(\hat{\pi}, \hat{b})$ is a feasible solution to (3) where its objective value is less than $m + s$. It implies that $b^*_o < m + s$. \hfill \Box

### 3.2 Proposed algorithm

The set $A = \{a^1, \ldots, a^n\}$

**Phase 1.**

- Step 1. Find the VR fram, $E^*$, of $A$
- Step 2. Find $K = \bigcup_{i=1}^{m+s} K_i$
- Step 3. For each $j \in J^* \cap K$ classify $a^j$ as an anchor point.

**Phase 2.**

For each $j \in J^* \setminus K$ solve (3): If $b^*_o < m + s$, then $a^j$ is an anchor point.

### 4 Example

In order to understand how the algorithm works and also to compare it with the BD-algorithm, we consider two examples which appeared in Bougnol and
Dulá [5]. The data comes in the Table 1:

<table>
<thead>
<tr>
<th>DMU</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attribute 1</td>
<td>4</td>
<td>10</td>
<td>13</td>
<td>14</td>
<td>5</td>
<td>10</td>
<td>3</td>
<td>6</td>
<td>3</td>
<td>9</td>
<td>12</td>
<td>7</td>
<td>9.144</td>
</tr>
<tr>
<td>Attribute 2</td>
<td>12</td>
<td>10</td>
<td>6</td>
<td>1</td>
<td>9</td>
<td>6</td>
<td>7</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>10.5</td>
<td>6.757</td>
</tr>
<tr>
<td>Attribute 3</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>5</td>
<td>5.5</td>
<td>4</td>
<td>6</td>
<td>5</td>
<td>8</td>
<td>7</td>
<td>3.5</td>
<td>5.579</td>
</tr>
</tbody>
</table>

For the first example, the data in Table 1 are used for treating all three attributes as outputs. The frame, $E^*$, consists eight elements: $a_1, a_2, a_3, a_4, a_5, a_9, a_{10}, a_{13}$. All of them are anchor points except $a_{13}$. The point $a_1, a_4, a_9$ were detected in Phase 1 and others in Phase 2. It can be seen in Fig. 1 that showing simple projections (2-D), points $a_1, a_4, a_9$ in output 2, output 4 and output 3, respectively, have the best operation.

The second example is resulted from the same data in Table 1 by changing the roles of first two attributes to inputs and keeping the third as an output. The frame set, $E^*$, consists: $a_4, a_9, a_{10}$, and $a_{11}$. All of these are anchor points. The point $a_4, a_9$, and $a_{11}$ were detected in Phase 1 and $a_{10}$ in Phase 2.

Both examples have confirmed the superiority of the proposed algorithm over the BD-algorithm. As you can see in Table 2 about half of anchor points were detected in Phase 1. In addition, in comparison with the BD-algorithm
our algorithm, only through solving one model, can identify whether or not a frame point is anchor point while BD-algorithm does it by solving $m + s$ subproblems. From a computational point of view if we deal with a real case having numerous DMUs, it will be a remarkable achievement.

We also tested the proposed algorithm on several real data which have been obtained from A.Emrozinejhad’s home page. The results are displayed in Table 2. The percent of anchor points which were detected in Phase 1 and the Percent of anchor points of frame are shown with PAP1 and PAP, respectively.

Table 2: Results got from real data

<table>
<thead>
<tr>
<th>Data file</th>
<th>n</th>
<th>m</th>
<th>s</th>
<th>percent of VR efficient</th>
<th>PAP1</th>
<th>PAP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dyson data</td>
<td>62</td>
<td>1</td>
<td>4</td>
<td>12.9</td>
<td>50</td>
<td>100</td>
</tr>
<tr>
<td>Scheel data</td>
<td>63</td>
<td>4</td>
<td>2</td>
<td>7.93</td>
<td>80</td>
<td>100</td>
</tr>
<tr>
<td>Charnes data</td>
<td>49</td>
<td>5</td>
<td>3</td>
<td>20.40</td>
<td>50</td>
<td>100</td>
</tr>
</tbody>
</table>

5 Conclusion

Almost all frame points are found to be anchor points; therefore, they play a major role in the geometry of the production possibility set. In this paper, we studied a new property of anchor points based on which we proposed a new algorithm which detected anchor points much efficiently. Actually, we saw that almost about half of anchor points are detectable only by a simple comparison. In addition, in comparison with the BD-algorithm our algorithm, only through solving one model, can identify whether or not a frame point is anchor point while BD-algorithm does it by solving $m + s$ subproblems.

References


Received: June, 2010