INTERNATIONAL ECONOMIC REVIEW
Vol. 38, No. 3, August 1997

A GEOMETRIC APPROACH TO THE STOLPER–SAMUELSON THEOREM*

BY KOJI SHIMOMURA[†1]

Kobe University, Japan

This paper presents a pair of new equivalent conditions for a given $n \times n$ matrix of distributive shares to satisfy the Stolper–Samuelson criterion. Specifying n as 4 and making use of barycentric coordinates, I give geometric characterization to the equivalent conditions.

1. INTRODUCTION

The derivation of equivalent conditions for a given $n \times n$ matrix of distributive shares to satisfy the Stolper–Samuelson criterion has received the close attention of trade theorists, in particular during the 1960s and 1970s.[2] As Inada (1971) made clear, the derivation of the equivalent conditions for the S-S criterion is itself equivalent to the derivation of the equivalent conditions for a given $n \times n$ stochastic matrix to have an inverse that is Minkowski or Metzler. This paper presents a pair of new equivalent conditions, one for Minkowski and the other for Metzler matrices.

The new conditions have two virtues as follows. First, the proof of the equivalence of each condition is short and simple. Some standard knowledge about Minkowski and Metzler matrices and linear algebra is enough to understand the proof.

Second, specifying n as 4 and making use of barycentric coordinates, we can provide a geometric characterization of each equivalent condition. These characterizations make clear just how the higher-dimensional Stolper–Samuelson theorem requires the pairing of goods and factors of production.

The main theorem is stated and proved in the next section. Section 3 provides geometric visualization of Minkowski- and Metzler-properties in the 4×4 setting. Section 4 makes a concluding remark.

* Manuscript received August 1993; revised October 1996.

† E-mail: *simomura@rieb.kobe-u.ac.jp*

[1] An earlier version of this paper was based on the discussion with Ronald W. Jones, during his visit to RIEB, Kobe University, Japan, in 1993. I am grateful to him for continuously encouraging me to complete this paper. Most of the present version was completed when I was visiting the University of New South Wales, Australia, in 1996. I thank Murray C. Kemp for very helpful suggestions. I also thank Fumio Dei, Bill Ethier, Robert K. McCleery and Nobuo Minabe for valuable comments. Two anonymous referees are greatly acknowledged concerning the present version. In particular, the proof of the main theorem owes heavily to the suggestion of one of them.

[2] For example, Chipman (1969), Kemp and Wegge (1969a, b), Inada (1971), Minabe (1967), Uekawa (1971), Uekawa, Kemp and Wegge (1973). See also Jones (1993), Jones, Marjit and Mitra (1993), and Mitra and Jones (forthcoming), as recent contributions.

2. THE MAIN THEOREM

Let $\Theta \equiv [\theta_1, \ldots, \theta_n]$ be an $n \times n$ matrix of the distributive shares of primary factors of production. Each column, θ_i, represents each of n commodities, respectively. I assume that Θ is nonsingular throughout.

Inada (1971) defined the strong Stolper–Samuelson criterion in two alternative ways:

[SSS-I] Θ has an inverse that is a Minkowski matrix, that is, the inverse Θ^{-1} has positive diagonal and negative off-diagonal elements.
[SSS-II] Θ has an inverse that is a Metzler matrix, that is, Θ^{-1} has negative diagonal and positive off-diagonal elements.

First, I state a standard theorem of square matrices.

LEMMA 1. (i) *If Θ^{-1} is Minkowski or Metzler, then the principal minors of all orders of Θ^{-1} are nonsingular.*
(ii) *If Θ^{-1} is Minkowski, then any principal minor of less than n order of Θ^{-1} is a positive matrix.*
(iii) *If Θ^{-1} is Metzler, then any principal minor of less than n order of Θ^{-1} is a negative matrix.*

PROOF. The proof is standard and is left to the reader. (To prove (i) and (ii), note that any principal minor of less than n order of Minkowski and Metzler stochastic matrices is a matrix with a dominant diagonal. The proof of (iii) is in Inada 1971, Theorem 2, p. 224. Q.E.D.

Now I state the main theorem of this paper.

THEOREM. *Let $\Lambda[b_1, \ldots, b_n] \equiv \{\Sigma_{i=1}^n a_i b_i | a_i > 0, \Sigma_{i=1}^n a_i = 1\}$, where b_i and a_i are an n-dimensional vector and a scalar, respectively. Let e_i, $i = 1, \ldots, n$, be an n-dimensional vector with unity in the ith element and zero in all other elements.*
[1][SSS-I] holds iff:

(PI)

$$\text{for any subset of } N \equiv \{1, \ldots, n\}, \text{ say } J, \text{ and for } k \in N - J,\ \theta_k \in \Lambda[E_{N-J}, \Theta_J],$$

$$\text{where } E_{N-J} \equiv \{e_s | s \in N - J\} \text{ and } \Theta_J \equiv \{\theta_j | j \in J\}.$$

[2][SSS-II] holds iff:

(PII)

$$\text{for any subset of } N \equiv \{1, \ldots, n\}, \text{ say } J, \text{ and for } k \in N - J, \Lambda[\theta_k, E_{N-J}] \cap \Lambda[\Theta_J]$$

$$\neq \varnothing, \text{ where } \varnothing \text{ denotes the empty set.}$$

PROOF. [(PI) $\Rightarrow$ (SSS-I) and (PII) $\Rightarrow$ (SSS-II)]. First, I prove that (PI) $\Rightarrow$ (SSS-I). J is specified as $N - \{k\}$. (PI) implies that for any $k \in N$ θ_k is a positive convex combination of $\theta_1, \ldots, \theta_{k-1}, \theta_{k+1}, \ldots, \theta_n$, and e_k. It follows that there are n positive values, a_{kj}, $j = 1, \ldots, n$, such that

$$a_{kk}\theta_k + \sum_{j=1}^{n} (-a_{kj})\theta_j = e_k \text{ for any } k \in N,$$

which means that Θ^{-1} is Minkowski. By a parallel argument I prove (PII) $\Rightarrow$ (SSS-II).
[(SSS-I) $\Rightarrow$ (P-I) and (SSS-II) $\rightarrow$ (PII)]. Let $\overline{N} \equiv \{-1, -2, \ldots, -n\}$. I partition Θ and Θ^{-1} as follows.

$$(1) \qquad \Theta = \{\Theta_{n,m}, \Theta_{n,(-m)}\}, \Theta^{-1} \equiv \begin{bmatrix} \Theta^{m,m} & \Theta^{m,(-m)} \\ \Theta^{(-m),m} & \Theta^{(-m),(-m)} \end{bmatrix}, \quad 1 \le m < n$$

Thus, $\Theta_{x,y}$ (respectively, $\Theta^{x,y}$), $x, y \in N \cup \overline{N}$, is a submatrix of Θ (respectively, Θ^{-1}) that consists of (i) the first x rows of Θ (respectively, Θ^{-1}) if $x \in N$, or the last $n + y$ rows of Θ (respectively, Θ^{-1}) if $x \in \overline{N}$, and (ii) the first y columns of Θ (respectively, Θ^{-1}) if $y \in N$, or the last $n + y$ columns of Θ (respectively, Θ^{-1}) if $y \in \overline{N}$.

Let I be the n-dimensional identity matrix. $I_{x,y}$ can be defined similarly to $\Theta_{x,y}$. Define ω_x as the x-dimensional vector with all elements being unity. Lastly, it is assumed that superscript T attached to a vector indicates the transpose of the vector.

First, since $\Theta\Theta^{-1} = I$, (1) implies that

$$(2) \qquad \Theta_{n,m}\Theta^{m,m} + \Theta_{n,(-m)}\Theta^{(-m),m} = I_{n,m}$$

If Θ^{-1} is either Minkowski or Metzler, then Lemma 1(i) assures us that the inverse matrix $[\Theta^{m,m}]^{-1}$ exists for any $m \in N$. Therefore,

$$(3) \qquad \Theta_{n,m} = I_{n,m}[\Theta^{m,m}]^{-1} - \Theta_{n,(-m)}\Theta^{(-m),m}[\Theta^{m,m}]^{-1}$$

Equation (3) shows that $\Theta_{n,m}$ is a linear combination of $I_{n,m}$ and $\Theta_{n,(-m)}$ with the weights $[\Theta^{m,m}]^{-1}$ and $-\Theta^{(-m),m}[\Theta^{m,m}]^{-1}$. Now, since Θ is a stochastic matrix ($\omega_n^T\Theta = \omega_n^T$), so is the inverse ($\omega_n^T\Theta^{-1} = \omega_n^T$), which implies that $\omega_m^T\Theta^{m,m} + \omega_{(-m)}^T\Theta^{(-m),m} = \omega_n^T I_{n,m}$, or

$$(4) \qquad \omega_m^T = \omega_n^T[\Theta^{m,m}]^{-1} - \omega_{(-m)}^T\Theta^{(-m),n}[\Theta^{m,m}]^{-1},$$

which implies that the sum of the weights on the RHS of (3) is unity for each column vector of $\Theta_{n,m}$.

Now, suppose that Θ^{-1} is Minkowski. Then, $\Theta^{(-m),m}$ is a negative matrix and Lemma 1(ii) implies that $[\Theta^{m,m}]^{-1}$ is a positive matrix. It follows that the weights on the right-hand side of (3) are all positive and that (3) means that any column vector of $\Theta_{n,m}$ is a positive convex combination of the columns of $I_{n,m}$ and $\Theta_{n,(-m)}$. Therefore, if Θ^{-1} is Minkowski, it satisfies (PI).

Next, suppose that Θ^{-1} is Metzler. Then, $\Theta^{(-m),m}$ is a positive matrix and Lemma 1(iii) implies that $[\Theta^{m,m}]^{-1}$ is a negative matrix. Rewriting (3) as

(5) $$\Theta_{n,m} - I_{n,m}[\Theta^{m,m}]^{-1} = -\Theta_{n,(-m)}\Theta^{(-m)m}[\Theta^{m,m}]^{-1},$$

one can see that a positive combination of each column of $\Theta_{n,m}$ and m column vectors of $I_{n,m}$ is equal to a positive combination of the $n-m$ column vectors of $\Theta_{n,(-m)}$. In view of (4), it follows that for any column vector of $\Theta_{n,m}$, say θ_i, there are n positive scalars a_i, $i = 1, \ldots, n$, such that

(6) $$\theta_i + \sum_{j=1}^{m} a_j e_j = \sum_{j=m+1}^{n} a_j \theta_j \qquad i = 1, \ldots, m$$

and

(7) $$1 + \sum_{j=1}^{m} a_j = \sum_{j=m+1}^{n} a_j > 0$$

Combining (6) and (7), I arrive at

(8)

$$\left(1\Big/\left(1 + \sum_{s=1}^{m} a_s\right)\right) \cdot \theta_i + \sum_{j=1}^{m}\left(a_j\Big/\left(1 + \sum_{s=1}^{m} a_s\right)\right) \cdot e_j = \sum_{j=m+1}^{n}\left(a_j\Big/\left(1 + \sum_{s=m+1}^{n} a_s\right)\right) \cdot \theta_j,$$

which means that a positive convex combination of the columns of θ_j and $I_{n,m}$ is equal to that of the columns of $\Theta_{n,(-m)}$. Therefore, if Θ^{-1} is Metzler, it satisfies (PII). Q.E.D.

REMARK 1. If $N - \{k)$ is chosen as J, $\theta_k \in \Lambda[E_{N-J}, \Theta_J]$ for any $k \in N$ is equivalent to the Kemp–Wegge condition (Kemp and Wegge 1969a). It follows that the Kemp–Wegge condition is a necessary condition for (PI) to hold. Inada (1971, pp. 222–223) presented a necessary condition for Θ^{-1} to be Metzler. The condition is dual to the Kemp–Wegge condition in the sense that it becomes the Kemp–Wegge condition if all inequalities in Inada's condition are just reversed. Hence, if $N - \{k)$ is chosen as $J, \Lambda[\theta_k, E_{N-J}] \cap \Lambda[\Theta_J] \neq \phi$ for any $k \in N$ is equivalent to Inada's condition. It follows that the condition is necessary condition for (PII) to hold.

REMARK 2. Applying a parallel argument to the row vectors of Θ, one can derive other equivalent conditions. Let $\bar{\theta}_i$ be the ith row vector of Θ and define $\bar{\Theta}_J \equiv \{\bar{\theta}_j | j \in J\}$. Let $\Lambda^*[b_1, \ldots, b_n] \equiv \{\Sigma_{i=1}^{n} a_i b_i | a_i > 0\}$. Then, one can prove the following propositions:[3]

- [1′] Θ^{-1} is Minkowski iff for any subset of $N \equiv \{1, \ldots, n\}$, say J, and for any $k \in N - J$, $\bar{\theta}_k \in \Lambda^*[E_{N-J}, \bar{\Theta}_J]$.
- [2′] Θ^{-1} is Metzler iff for any subset of $N \equiv \{1, \ldots, n\}$, say J, and for any $k \in N - J$, $\Lambda[\bar{\theta}_k, E_{N-J}] \cap \Lambda^*[\Theta_J] \neq \phi$.

[3] The proof is in Shimomura (1994). It is available from the author on request.

[1′] can be rewritten as

[1″] Θ^{-1} is Minkowski iff for any pair of partition (I, J), $I \cup J = N$, $I \cap J = \phi$, $I, J \neq \phi$, and any $k \in I$, there exists positive x_j, $j \in J$, such that

$$\theta_{ki} > \sum_{j \in J} \theta_{ji} x_j \quad \text{for any } i \in I$$

$$\theta_{ki} = \sum_{j \in J} \theta_{ji} x_j \quad \text{for any } i \in J$$

[1″] is almost the same equivalent condition as Uekawa (1971) derived; see his condition IV. (An important difference between his condition IV and the above [1″] is that the part in his condition which corresponds to the above equality holds with (weak) inequality " $\leq$." Note that the economic interpretation he gave to Condition IV can be directly applied to [1″].)

3. GEOMETRY OF (PI) AND (P-II) IN THE 4 × 4 SETTING

In this section I specify n as 4 and visualize (PI) and (PII) by using barycentric coordinates. Consider a tetrahedron Ω with four equal equilateral triangles as its sides. Choose a point in or on the tetrahedron and draw four perpendicular segments to the four sides. Denote the length of the segments by θ_{*j}, $j = 1, 2, 3, 4$. It is a well-known fact that the sum $\Sigma_{j=1}^{4} \Theta_{*j}$ takes on a common value for any such point. Thus without loss the common value can be assumed to be unity.

Next denote by the space of the 4-dimensional nonnegative vectors with components that sum to unity, one can map any element $x \equiv (x_1, x_2, x_3, x_4)$ in Ω^* to a point in Ω by making x_j equal to θ_{*j} of the point. The mapping, say f, is clearly one-to-one. Moreover the mapping has the following property.

LEMMA 2. *Take any x, x^* in Ω^* and consider a nonnegative convex combination, $x(a) \equiv ax + (1-a)x^*$, $0 \leq a \leq 1$. Then the point $f(x(a))$ in Ω is on the segment connecting $f(x)$ and $f(x^*)$ and separates it with the ratio $(1-a)$: a. Similarly, any nonnegative convex combination of three vectors x, x^* and x^{**} in Ω^* corresponds to a point in Ω that is somewhere on the triangle that has $f(x)$, $f(x^*)$ and $f(x^{**})$ as its vertices. Furthermore, any nonnegative convex combination of four vectors in Ω^* corresponds to a point in or on the tetrahedron that has four points corresponding to the vectors.*

PROOF. See Figure 1, where x_j and x_j^* are the length of the segments perpendicular to a common side of Ω. Clearly the length of AB in the figure is $ax_j + (1-a)x_j^*$. Note that any point on a triangle is a nonnegative convex combination of one vertex and a nonnegative convex combination of the other two vertices with an appropriate value of weight a. Then the second and last parts of the present lemma will be clear. Q.E.D.

Now take a look at Figure 2. Four points corresponding to four share vectors $f(\theta_i)$, $i = 1, 2, 3, 4$, make up a small tetrahedron in Ω, while the four points $f(e_i)$,

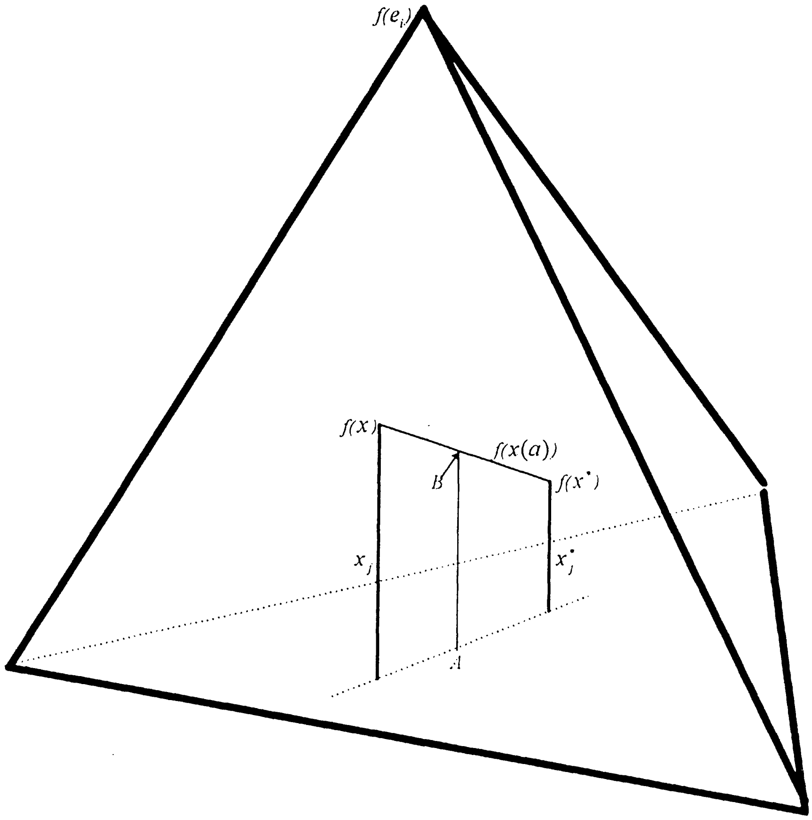

FIGURE 1

$i = 1, 2, 3, 4$, are vertices of Ω. Then observe that there are two cones that share $f(\theta_i)$ as the common vertex; one that is *opposite* to the small tetrahedron and the other that contains it. I call the former one the *outside cone* with vertex $f(\theta_i)$ and the other the *inside cone* with vertex $f(\theta_i)$. For example, the outside cone with vertex $f(\theta_1)$ is AFGE and the inside cone with the same vertex is AHIJ. If, as is depicted in Figure 2, each outside cone contains just one of $f(e_i)$, then the straight line connecting $f(\theta_i)$ and $f(e_i)$ hits the triangle that has three points $f(\theta_j)$, $j \neq i$, as its vertices. One can also observe in Figure 2 that any $f(\theta_i)$ is in the two tetrahedrons, one that has point $f(e_i)$ and three points $f(\theta_j)$, $j \neq i$, as its vertices and the other that has points $f(e_i)$, $f(e_k)$ and $f(\theta_j)$, $i \neq k \neq j \neq i$, as its vertices. In view of Lemma 2, it follows that (PI) is established. Moreover, one may verify that if any $f(\theta_i)$ is in the tetrahedron that has $f(e_i)$ and three points $f(\theta_i)$, $j \neq i$, as its vertices, then each outside cone contains just one of $f(e_i)$. Hence I arrive at the geometric proposition.

PROPOSITION 1. *(PI) holds in the* 4×4 *setting iff each outside cone contains just one of* $f(e_i)$, $i = 1, 2, 3, 4$.

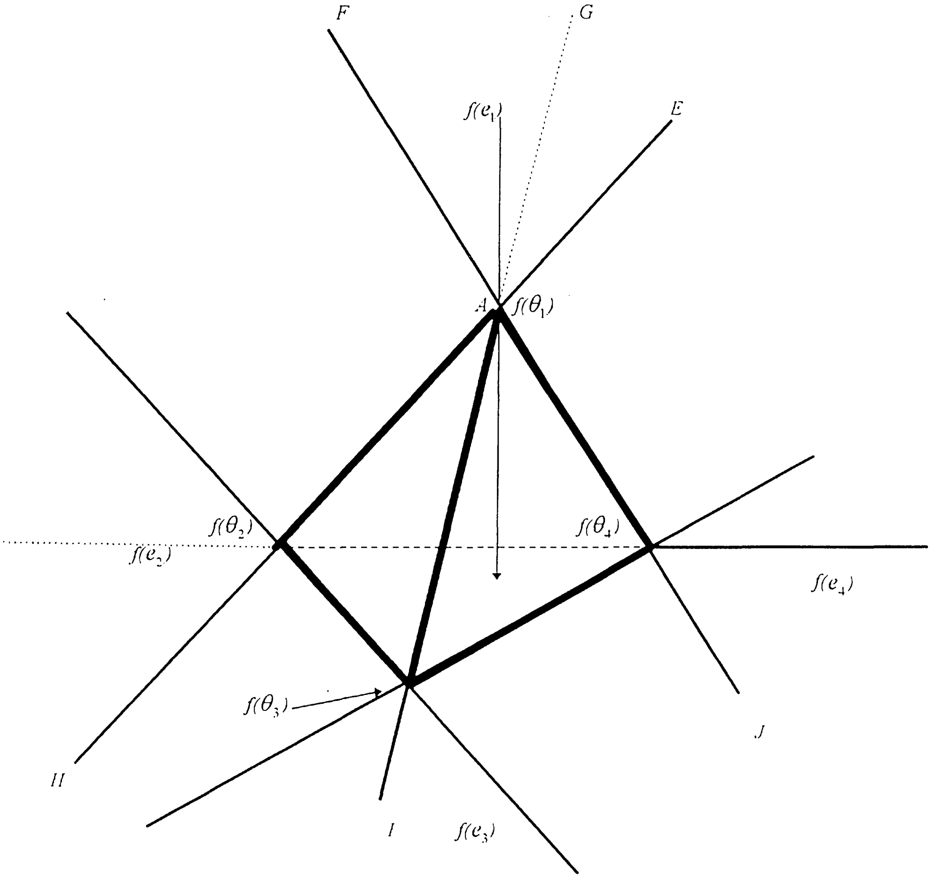

FIGURE 2

Next suppose that each point $f(e_i)$ is in the inside cone with vertex $f(\theta_i)$. Of course, $f(e_i)$ is not in the tetrahedron that has $f(\theta_j)$, $j = 1, 2, 3, 4$, as its vertices. Thus as is clear from Figure 3, the segment connecting $f(\theta_i)$ and $f(e_i)$ must intersect the triangle that has $f(\theta_j)$, $j \neq i$, as its three vertices. One can also observe from Figure 3 that any triangle that has $f(\theta_i)$, $f(e_i)$ and $f(e_j)$, $i \neq j$, as its vertices intersects the segment connecting two points $f(\theta_s)$ and $f(\theta_k)$, where $\{s, k\} \cap \{i, j\} = \varnothing$, and that any tetrahedron that has $f(\theta_i)$, $f(e_i)$, $f(e_j)$ and $f(e_s)$ as its vertices contains $f(\theta_k)$, where $\{i, j, s, k\} = \{1, 2, 3, 4\}$. In view of Lemma 2, it follows that (PII) is established. Moreover, one may verify that if for each $f(\theta_i)$ there is a point $f(e_i)$ such that the segment connecting $f(\theta_i)$ and $f(e_i)$ must intersect the triangle that has $f(\theta_j)$, $j \neq i$, as its vertices, then each inside cone contains just one $f(e_i)$. Thus I arrive at our second geometric proposition.

PROPOSITION 2. *(PII) holds in the* 4×4 *setting iff each inside cone contains just one of* $f(e_i)$, $i = 1, 2, 3, 4$.

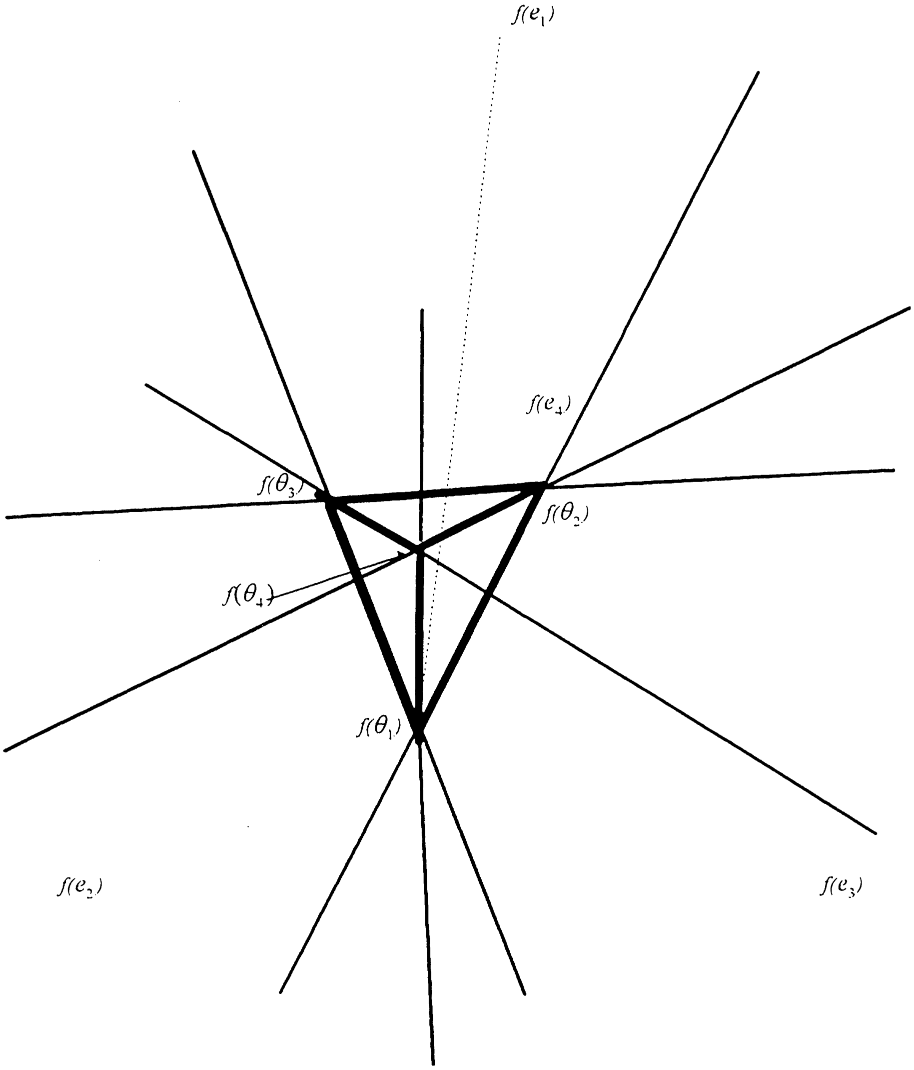

FIGURE 3

4. A CONCLUDING REMARK

The geometric technique involving barycentric coordinates has been used in economic analysis since McKenzie (1955). Recent notable examples of using the technique are Leamer (1987), Jones and Marjit (1991), Jones (1992), and Jones and Mitra (1995).

One novel point in this paper is that the technique has been extended to analyze 4-dimensional cases, while all of the above literature applied it to the analysis of 3-dimensional cases. A byproduct of this extension is that one can visualize the

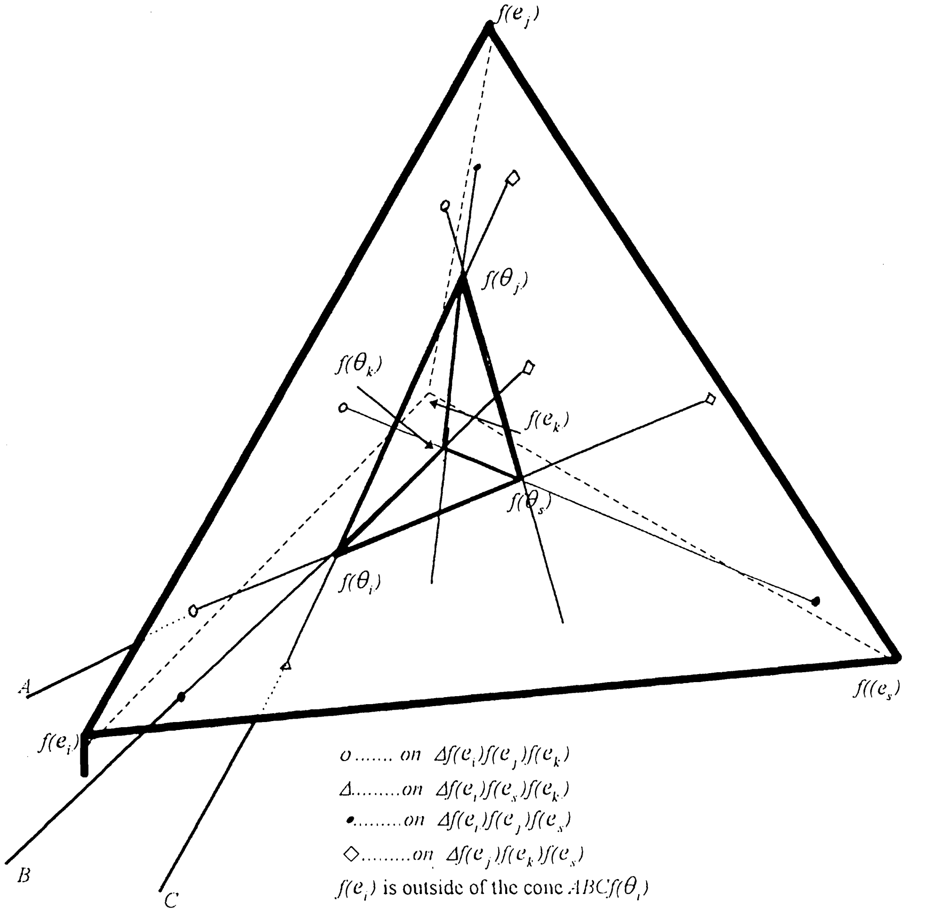

FIGURE 4

well-known case in which the Kemp–Wegge condition (Kemp and Wegge 1969) is not sufficient for a stochastic matrix to have an inverse which is Minkowski when the number of commodities and factors is not less than 4 (see Figure 4). The vertex $f(e_i)$ is outside the cone ABC$f(\theta_i)$. Thus (PI) does not hold. However, one can observe from Figure 4 that for any $f(\theta_i)$ all extended lines of segments $f(\theta_i)f(\theta_j)$, $f(\theta_i)f(\theta_s)$, and $f(\theta_i)f(\theta_k)$ hit the triangle that has $f(e_j)$, $f(e_s)$, and $f(e_k)$ as its vertices, which is the geometric expression of the Kemp–Wegge condition.[4]

[4] We can make a parallel argument concerning Inada's necessary condition; see Shimomura (1994).

REFERENCES

CHIPMAN, J., "Factor Price Equalization and the Stolper–Samuelson Theorem," *International Economic Review* 10 (1969), 399–406.

INADA, K., "The Production Coefficient Matrix and the Stolper–Samuelson Condition," *Econometrica* 39 (1971), 219–240.

JONES, R.W., "Reflections on the Stolper–Samuelson Theorem," in H. Herberg and N.V. Long, eds., *Trade, Welfare and Economic Policies* (Ann Arbor: University of Michigan Press 1993, pp. 21–36).

——, "Factor Scarcity, Factor Abundance and Attitude towards Protection: the 3×3 Model," *Journal of International Economic Integration* 7 (1992), 1–19.

—— AND S. MARJIT, "The Stolper–Samuelson Theorem, the Leamer Triangle and the Produced Mobile Factor Structure," in A. Takayama, M. Ohyama and H. Ohta, eds., *Trade, Policy and International Adjustments* (San Diego: Academic Press 1991, pp. 95–107).

—— AND T. MITRA, "Share Ribs and Income Distribution," *Review of International Economics* 3 (1995), 36–52.

——, S. Marjit, and T. Mitra, "The Stolper–Samuelson Theorem: Links to Dominant Diagonals," R. Becker, M. Boldrin, R. Jones and W. Thomson, eds., in *General Equilibrium, Growth and Trade II: The Legacy of Lionel McKenzie* (San Diego: Academic Press 1993, pp. 429–441).

KEMP, M.C. AND L. WEGGE, "On the Relation between Commodity Prices and Factor Rewards," *International Economic Review* 10 (1969a), 407–413.

—— AND ——, "Generalization of the Stolper–Samuelson and Samuelson–Rybczynski Theorems in Terms of Conditional Input-Output Coefficients," *International Economic Review* 10 (1969b), 414–425.

LEAMER, E., "Paths of Development in the Three-Factor n-Good General Equilibrium Model," *Journal of Political Economy* 95 (1987), 961–999.

MCKENZIE, L., "Equality of Factor Prices in World Trade," *Econometrics* 23 (1955), 239–257.

MINABE, N., "The Stolper-Samuelson Theorem, the Rybczynski Effect, and the Heckscher–Ohlin Theory of Trade Pattern and Factor Price Equalization: The Case of a Many-Commodity, Many-Factor Country," *Canadian Journal of Economics and Political Science* 33 (1967), 401–419.

MITRA, T. AND R.W.JONES, "Factor Shares and the Chipman Condition," in J. Melvin, J. Moore and R. Rietzman, eds., in *Trade, Welfare, and Econometrics: Essays in Honor of John Chipman*, forthcoming.

SHIMOMURA, K., "A New Set of Equivalent Conditions for a Given $N \times N$ Stochastic Matrix to Have an Inverse Which is Minkowski or Metzler," *RIEB Discussion Paper* No. 43, Kobe University, December 1994.

UEKAWA, Y., "Generalization of the Stolper–Samuelson Theorem," *Econometrica* 39 (1971), 197–217.

——, M.C. Kemp and L. Wegge, "P and PN Matricies, Minkowski and Metzler Matricies, and the Generalizations of the Stopler–Samuelson and Samuelson–Rybczynski Theorems," *Journal of International Economics* 3 (1973), 53–76.