Gakushuin Economic Papers, Vol. 31, No. 3 October 1994

Mathematical Characteristics of ROXY Index (IV): ROXY Index as Compared with Correlation Coefficient

Yasushi Asami, Shino Funamoto, Noriyuki Hiraoka
Jean H.P.Paelinck and Tatsuhiko Kawashima*

Contents

Abstract

The ROXY index serves as an instrument for the investigation into the stage and speed of the process of spatial convergence and divergence observed on various socio-economic activities. This paper investigates the basic mathematical characteristics of the ROXY index, focusing upon (i) the functional relationships of the ROXY index with the correlation coefficient and regression coefficient, and (ii) the similarity and difference between the ROXY index and correlation coefficient in mathematical characteristics. The results of the investigation show, as indicated in a summary table that the ROXY index possesses several characteristics advantageous to empirical analyses on the dynamic process of the phenomena of spatial redistribution of socio-economic activities.

Key Words

Centralization, Commutativity, Convergence-divergence index, Concentration, Correlation coefficient, Half-unboundedness, Homogeneity of degree zero, Linear relation, MMMM property, Order-invariance, Permutativity, Regression coeffient, ROXY index, Scale-invariance, Spatial cycles

* Asami and Funamoto are associated with the Department of Urban Engineering at the University of Tokyo, Hiraoka with the National Infrastructure Department of Mitsubishi Research Institute in Tokyo, Paelinck with the Rotterdam School of Economics at Erasmus University of Rotterdam, and Kawashima with the Department of Economics at Gakushuin University in Tokyo. Kawashima gratefully acknowledges the research support from the Grant-in-Aid for General Scientific Research by the Ministry of Education, Science and Culture in Japan.

1 Introduction

Since the original version of the basic concept of the ROXY index had been proposed toward the end of 1970s[1], a number of studies have been carried out applying the ROXY-index approach to empirical analyses on the phenomena of spatial-cycle processes observed in both inter-metropolitan and intra-metropolitan scopes in conjunction with socio-economic activities. In parallel with these developments, mathematical characteristics peculiar to the ROXY index have been theoretically investigated for the purpose of obtaining a better insight into its structural implications. For example, Kawashima and Hiraoka (1993) have shown, for the system of the one-dimensional discrete-linear region, that there exists a straightforward functional relationship *between* the ROXY-index value for which the CBD distance is used as its weighting factor *and* the ROXY-index value for which the reversed CBD distance is used as its weighting factor. This distance-weighting operationability wakens/wekened the ROXY index on interesting tool of spatial econometrics. Hiraoka and Kawashima (1993) have also developed two types of theoretically-ideal formulations of the ROXY index; one for the system of a one-dimensional continuous-linear region, and the other for the system of a two-dimensional fan-shaped region. Each of these two formulations was examined by Hiraoka and Kawashima (1994) to specify a functional relationship *between* the ROXY-index value calculated by use of the CBD distance as its weighting factor *and* the ROXY-index value calculated by use of the reversed CBD distance as its weighting factor.

The present paper which has been prepared under such particular circumstances carries three sub-themes. First, it investigates in Section 2 the functional relationship of the value of the ROXY index with that of the correlation coefficient and with that of the regression coefficient. Secondly in Sections 3 are discussed the mathematical characteristics commonly shared by the ROXY index and the correlation coefficient. Thirdly, the mathematical characteristics of the ROXY index which are different from those of the correlation coefficient are examined in Section 4.

2 Relationship between ROXY Index and Correlation Coefficient

2.1 Definition of ROXY Index

The ROXY index is a quantitative measure which can be indicative of the stage and speed of the process of spatial convergence and divergence conceived in both of the theoretical and empirical spheres. Suppose, for example, we are interested in investigating the phenomenon of population concentration and deconcentration in a system of metropolitan areas. To this inter-metropolitan analysis, we can apply the ROXY index which is defined as follows;

$$ROXY_{t,t+1} \equiv \left(\frac{WAGR_{t,t+1}}{SAGR_{t,t+1}} - 1\right) \times 10^4$$

$$= \left(\frac{\sum_{i=1}^{n}(x_i^t \times y_i^{t,t+1})}{\sum_{i=1}^{n} x_i^t} \times \frac{n}{\sum_{i=1}^{n} y_i^{t,t+1}} - 1\right) \times 10^4, \tag{2.1}$$

where

$ROXY_{t,t+1}$: ROXY index for the period between years t and $t+1$

x_i^t : population of metropolitan area i in year t (This variable is used as ***weighting variable***[2] for the calculation of $WAGR_{t,t+1}$.)

$y_i^{t,t+1}$: annual growth ratio of population in metropolitan area i during the period between years t and $t+1$ (This variable functions in our study of this paper as ***principally-concerned variable***[3] which is given as the k-th root of p_i^{t+k}/p_i^t where p_i^t and p_i^{t+k} respectively represent the population of metropolitan area i in year t and that in year $t+k$.)

n : number of metropolitan areas

$WAGR_{t,t+1}$: weighted average of the annual growth ratio $y_i^{t,t+1}$ over n metropolitan areas which is given as

$$\sum_{i=1}^{n}(x_i^t \times y_i^{t,t+1})/\sum_{i=1}^{n} x_i^t$$

$SAGR_{t,t+1}$: simple average of the annual growth ratio $y_i^{t,t+1}$ over n metropolitan areas which is given as

$$\sum_{i=1}^{n} y_i^{t,t+1}/n\ ,$$

Let us now introduce three vectors $\boldsymbol{x}$, $\boldsymbol{y}$ and $\boldsymbol{1}$ which are respectively defined as follows[4];

$$\boldsymbol{x}=\begin{bmatrix} x_1^t \\ x_2^t \\ \vdots \\ \vdots \\ x_n^t \end{bmatrix},\qquad \boldsymbol{y}=\begin{bmatrix} y_1^{t,t+1} \\ y_2^{t,t+1} \\ \vdots \\ \vdots \\ y_n^{t,t+1} \end{bmatrix},\qquad \boldsymbol{1}=\begin{bmatrix} 1 \\ 1 \\ \vdots \\ \vdots \\ 1 \end{bmatrix},$$

where $\boldsymbol{x} \geqq \boldsymbol{o}$ *and* $\boldsymbol{y} \geqq \boldsymbol{o}$.

Then, the ROXY index given by equation (2.1) can be expressed as follows;

$$ROXY_{t,t+1} \equiv \left(\frac{n\cdot(\boldsymbol{x},\boldsymbol{y})}{(\boldsymbol{1},\boldsymbol{x})(\boldsymbol{1},\boldsymbol{y})}-1\right)\times 10^4\ , \tag{2.2}$$

where the notation $(\boldsymbol{a}, \boldsymbol{b})$ indicates the inner product of vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ .That is,

$$(\boldsymbol{a},\boldsymbol{b})=\sum_{i=1}^{n}(a_i b_i)\ .$$

2.2 Functional Relationship: ROXY Index and Correlation Coefficient

For the sake of simplification, let R($\boldsymbol{x}$,$\boldsymbol{y}$) denote the ROXY index excluding the multiplier component of 10^4 in equation (2.2). It then results that

$$R(\boldsymbol{x},\boldsymbol{y})=\frac{n\cdot(\boldsymbol{x},\boldsymbol{y})}{(1,\boldsymbol{x})(1,\boldsymbol{y})}-1\,. \tag{2.3}$$

From equation (2.3), we obtain

$$\begin{aligned}R(\boldsymbol{x},\boldsymbol{y})&=\frac{n\cdot(\boldsymbol{x},\boldsymbol{y})}{(1,\boldsymbol{x})(1,\boldsymbol{y})}-1\\&=\frac{n\sum_{i=1}^{n}(x_i\cdot y_i)}{\sum_{i=1}^{n}x_i\cdot\sum_{i=1}^{n}y_i}-1\\&=\frac{\sum_{i=1}^{n}(x_i\cdot y_i)}{n\cdot\bar{x}\cdot\bar{y}}-1\,,\end{aligned} \tag{2.4}$$

where $\bar{a}$ indicates the simple average of all elements in vector $\boldsymbol{a}$, and is equal to

$$\left(\sum_{i=1}^{n}a_i\right)/n\,.$$

From equation (2.4), we have

$$\begin{aligned}R(\boldsymbol{x},\boldsymbol{y})&=\frac{\sum_{i=1}^{n}(x_i\cdot y_i)-n\cdot\bar{x}\cdot\bar{y}}{n\cdot\bar{x}\cdot\bar{y}}\\&=\frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{n\cdot\bar{x}\cdot\bar{y}}\\&=\frac{Cov(\boldsymbol{x},\boldsymbol{y})}{\bar{x}\cdot\bar{y}}\,,\end{aligned} \tag{2.5}$$

where Cov($\boldsymbol{x}$,$\boldsymbol{y}$) , which means the covariance between vectors $\boldsymbol{x}$ and $\boldsymbol{y}$, is given by

$$\left\{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})\right\}/n\,.$$

Meanwhile, the Pearson's correlation coefficient $\rho\,(\boldsymbol{x},\boldsymbol{y})$ is given by

$$\rho(\boldsymbol{x},\boldsymbol{y})=\frac{Cov(\boldsymbol{x},\boldsymbol{y})}{\sigma_x\cdot\sigma_y}\,, \tag{2.6}$$

where σ_a, which means the standard deviation of vector $\boldsymbol{a}$, is given by

$$\left\{\sum_{i=1}^{n}(a_i-\bar{a})^2/n\right\}^{1/2}.$$

From equations (2.5) and (2.6), it can be found that the following functional relationship exists between $R(\boldsymbol{x},\boldsymbol{y})$ and $\rho\,(\boldsymbol{x},\boldsymbol{y})$;

$$R(\boldsymbol{x},\boldsymbol{y}) = \frac{\sigma_x \cdot \sigma_y}{\bar{x} \cdot \bar{y}} \rho(\boldsymbol{x},\boldsymbol{y}) \tag{2.7}$$

Equation (2.7) implies that the ratio of "the ROXY-index value (after elimination of multiplier component of 10^4)" to "the value of the Pearson's correlation coefficient" is equal to the ratio of "the product of two standard deviations, one for the vector of weighting variable $\boldsymbol{x}$ and the other for the vector of principally-concerned variable $\boldsymbol{y}$," to "the product of two simple averages, one for the vector $\boldsymbol{x}$ and the other for the vector $\boldsymbol{y}$."

2.3 Functional Relationship: ROXY Index and Regression Coefficient

For a given pair of vectors $\boldsymbol{x}$ and $\boldsymbol{y}$, there exists an explicit functional relationship between the value of the correlation coefficient and that of the simple regression coefficient. This would mean that, in the light of what has been discussed in the above subsection, there also exists an explicit functional relationship between the value of the ROXY index and the simple regression coefficient. We will go into details below on this topic including the discussion on the standardization procedure of vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ for the ROXY index in comparison with the standardization procedure of the two vectors for the simple regression coefficient.

If we set $\boldsymbol{x}$ as the vector of explanatory variable and $\boldsymbol{y}$ as the vector of explained variable, the raw-score regression equation is generally given by the following formulation;

$$\hat{y}_i = b_{y \cdot x}(x_i - \bar{x}) + \bar{y} \,. \tag{3.1}$$

In equation (3.1), $\hat{y}_i$ means the expected value of the explained variable y_i given x_i, while $b_{y \cdot x}$ means the raw-score regression coefficient (*i.e.*, slope of the simple regression line) which is given by[5]

$$b_{y \cdot x} = \frac{Cov(\boldsymbol{x},\boldsymbol{y})}{\sigma_x^{\,2}} \,. \tag{3.2}$$

From equations (2.5), (2.6) and (3.2), the following relationships among $b_{y \cdot x}$, $R(\boldsymbol{x},\boldsymbol{y})$ and $\rho(\boldsymbol{x},\boldsymbol{y})$ can be obtained[6];

$$\begin{aligned} b_{y \cdot x} &= \frac{\bar{x} \cdot \bar{y}}{\sigma_x^{\,2}} R(\boldsymbol{x},\boldsymbol{y}) \\ &= \frac{\sigma_y}{\sigma_x} \rho(\boldsymbol{x},\boldsymbol{y}) \,. \end{aligned} \tag{3.3}$$

Equation (3.3) implies that the ratio of "the ROXY-index value (after elimination of multiplier component of 10^4)" to "the value of the raw-score regression coefficient" is equal to the ratio of "the variance of the vectors $\boldsymbol{x}$" to "the product of two simple averages, one for the vector $\boldsymbol{x}$ and the other for the vector $\boldsymbol{y}$."

Meanwhile, let us standardize x_i and y_i through the following transformation functions;

$$z_{xi}=\frac{x_i-\overline{x}}{\sigma_x} \text{ and } z_{yi}=\frac{y_i-\overline{y}}{\sigma_y}\ . \tag{3.4}$$

Then the form of the regression equation (3.1) would turn out to be

$$\hat{z}_{yi}=b_{zy\cdot zx}\cdot z_{xi} \tag{3.5}$$

where z_{yi} means the expected value of the explained variable z_{yi}, given z_{xi}, while $b_{zy\cdot zx}$ means the regression coefficient.

Under this setting, we obtain the following relationship indicating that the value of the correlation coefficient would be equal to the value of the standardized regression coefficient appearing in equation (3.5)[7];

$$b_{zy\cdot zx}=\rho\,(\boldsymbol{x},\boldsymbol{y}) \tag{3.6}$$

Accordingly, from equations (2.7) and (3.6), we get the following relationships among the standardized regression coefficient $b_{zy\cdot zx}$, ROXY index $R(\boldsymbol{x},\boldsymbol{y})$, and correlation coefficient $\rho(\boldsymbol{x},\boldsymbol{y})$;

$$\begin{aligned} b_{zy,zx} &= \frac{\overline{x}\cdot\overline{y}}{\sigma_x\cdot\sigma_y}R(\boldsymbol{x},\boldsymbol{y}) \\ &= \rho(\boldsymbol{x},\boldsymbol{y})\ . \end{aligned} \tag{3.7}$$

Be that as it may, it should be noted that equations (3.4) are not the only functions that can transform x_i and y_i in such a way that the value of the correlation coefficient would become equal to the value of the regression coefficient as shown in equation (3.6). For example, a family of the transformation functions as given below would also enable the value of the correlation coefficient to become equal to the value of the regression coefficient;

$$z_{xi}=\frac{x_i-\overline{x}}{p} \quad \text{and} \quad z_{yi}=\frac{y_i-\overline{y}}{q}\ ,$$

where p and q satisfy the condition $p/q=\sigma_x/\sigma_y$.

By analogy with this, we can transform x_i and y_i in such a way that the value of the regression coefficient would become equal to ROXY index. Among possible transformation functions of such kind would be;

$$z_{xi}=\frac{x_i-\overline{x}}{p} \quad \text{and} \quad z_{yi}=\frac{y_i-\overline{y}}{q}\ ,$$

where p and q should satisfy the condition of

$$p/q=\sigma_x^2\ /(\overline{x}\cdot\overline{y}).$$

More concrete examples of such transformation functions are given as follows for a pair of variables (Z_{xi}, Z_{yi}) where Z_{xi} and Z_{yi} are the standardized variables of x_i and y_i respectively;

$$(Z_{xi},\ Z_{yi}) = \left(\frac{x_i-\bar{x}}{\sigma_x^2\ /(\bar{x}\cdot\bar{y})},\ y_i-\bar{y}\right),$$

$$\left(\frac{x_i-\bar{x}}{\sigma_x^2\ /\bar{x}},\ \frac{y_i-\bar{y}}{\bar{y}}\right),$$

$$\left(\frac{x_i-\bar{x}}{\sigma_x^2},\ \frac{y_i-\bar{y}}{\bar{x}\cdot\bar{y}}\right), \qquad or$$

$$\left(\frac{x_i-\bar{x}}{\sigma_x/\bar{x}},\ \frac{y_i-\bar{y}}{\bar{y}/\sigma_x}\right)$$

It is to be noted here that any of the above transformation functions enables us to have the regression equation expressed by

$$\hat{Z}_{yi} = R(x,y)\cdot Z_{xi}\,.$$

Where $\hat{Z}_{yi}$ means the expected value of the explained variable Z_{yi} given the value of the explaining variable Z_{xi}.

3 Mathematical characteristics of ROXY Index: Common to Correlation Coefficient

The correlation coefficient $\rho(x,y)$ carries the following well-known mathematical characteristics for case $X \geqq O$ and $y \geqq O$;

(a-1) Commutativity between x and y:

$$\rho(x,y) = \rho(y,x)\,.$$

(a-2) Permutativity for elements in both x and y:

$$\rho(px,py) = \rho(x,y)\,,$$

Where p is a square-matrix operator which would permutate the elements in a vector[8] and referred to as a permutation operator.

(a-3) Hyper-homogeneity of degree zero for both x and y:

$$\rho(ax,y) = \rho(x,by) = \rho(cx,dy) = \rho(x,y)$$

where $a > 0$, $b > 0$, $c > 0$ and $d > 0$.

(a-4) Order-invariance between $\boldsymbol{x}$ and $\boldsymbol{y}$:

$$\forall(\xi,\eta)\in P^2(\boldsymbol{x},\boldsymbol{y}):\rho(\xi,\eta)=\rho(\boldsymbol{x},\boldsymbol{y})$$

Where $P^2(\boldsymbol{x},\boldsymbol{y})$ indicates the set of all pairs of vectors which we can generate by permutating the elements in vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ simultaneously by the same permutation operator, including the set of vectors $\boldsymbol{x}$ and $\boldsymbol{y}$.

(a-5) Maximum for monotonic increase and minimum for monotonic decrease (i.e., Monoinc-Max・Monodec-Min property, or MMMM property)

(i) If $\forall i\neq j:\ x_i>x_j \Longrightarrow y_i \geqq y_j$, and

if $\forall i\neq j:\ y_i>y_j \Longrightarrow x_i \geqq x_j$, then

$$(\boldsymbol{x},\boldsymbol{y})\in \arg\mathop{Max}_{\substack{\xi\in P(x)\\ \eta\in P(x)}}\rho(\xi,\eta),$$

(ii) If $\forall i\neq j:\ x_i>x_j \Longrightarrow y_i \leqq y_j$, and

if $\forall i\neq j:\ y_i>y_j \Longrightarrow x_i \leqq x_j$, then

$$(\boldsymbol{x},\boldsymbol{y})\in \arg\mathop{Min}_{\substack{\xi\in P(x)\\ \eta\in P(x)}}\rho(\xi,\eta),$$

where $P(\boldsymbol{x})$ indicates the set of all vectors which we can obtain by permutating the elements in a vector $\boldsymbol{x}$, including the vector $\boldsymbol{x}$ itself.

The ROXY index $R(\boldsymbol{x},\boldsymbol{y})$, as indicated below, carries the mathematical characteristics similar to the correlation coefficient;

(b-1) Commutativity between $\boldsymbol{x}$ and $\boldsymbol{y}$[9]:

$$\begin{aligned} R(\boldsymbol{x},\boldsymbol{y}) &= \frac{\sigma_x\cdot\sigma_y}{\bar{x}\cdot\bar{y}}\rho(\boldsymbol{x},\boldsymbol{y}) && \text{[from (2.7)]}\\ &= \frac{\sigma_y\cdot\sigma_x}{\bar{y}\cdot\bar{x}}\rho(\boldsymbol{y},\boldsymbol{x}) && \text{[from (a-1)]}\\ &= R(\boldsymbol{y},\boldsymbol{x})\,. && \text{[from (2.7)]} \end{aligned}$$

(b-2) Permutativity for elements in both $\boldsymbol{x}$ and $\boldsymbol{y}$:

$$R(p\boldsymbol{x}, p\boldsymbol{y}) = \frac{\sigma_x \cdot \sigma_y}{\bar{x} \cdot \bar{y}} \rho(p\boldsymbol{x}, p\boldsymbol{y}) \qquad \text{[from (2.7)]}$$

$$= \frac{\sigma_y \cdot \sigma_x}{\bar{y} \cdot \bar{x}} \rho(\boldsymbol{x}, \boldsymbol{y}) \qquad \text{[from (a-2)]}$$

$$= R(\boldsymbol{x}, \boldsymbol{y}) . \qquad \text{[from (2.7)]}$$

Where $\boldsymbol{p}$ is a permutation operator.

(b-3) Hyper-homogeneity of degree zero for both $\boldsymbol{x}$ and $\boldsymbol{y}$[10]:

(i) $$R(a\boldsymbol{x}, \boldsymbol{y}) = \frac{(a\sigma_x) \cdot \sigma_y}{(a\bar{x}) \cdot \bar{y}} \rho(a\boldsymbol{x}, \boldsymbol{y}) \quad \text{[from (2.7) and } Var(a\boldsymbol{x}) = a^2 Var(x)\text{]}$$

$$= \frac{\sigma_x \cdot \sigma_y}{\bar{x} \cdot \bar{y}} \rho(\boldsymbol{x}, \boldsymbol{y}) \qquad \text{[from (a-3)]}$$

$$= R(\boldsymbol{x}, \boldsymbol{y}) \qquad \text{[from (2.7)]}$$

(ii) $$R(\boldsymbol{x}, b\boldsymbol{y}) = \frac{\sigma_x \cdot (b\sigma_y)}{\bar{x} \cdot (b\bar{y})} \rho(\boldsymbol{x}, b\boldsymbol{y}) \qquad \text{[from (2.7)]}$$

$$= \frac{\sigma_x \cdot \sigma_y}{\bar{x} \cdot \bar{y}} \rho(\boldsymbol{x}, \boldsymbol{y}) \qquad \text{[from (a-3)]}$$

$$= R(\boldsymbol{x}, \boldsymbol{y}) \qquad \text{[from (2.7)]}$$

(iii) $$R(c\boldsymbol{x}, d\boldsymbol{y}) = R(\boldsymbol{x}, d\boldsymbol{y}) \qquad \text{[from (i)]}$$

$$= R(\boldsymbol{x}, \boldsymbol{y}) \qquad \text{[from (ii)]}$$

Where $a > 0$, $b > 0$, $c > 0$ and $d > 0$

(b-4) Order-invariance between $\boldsymbol{x}$ and $\boldsymbol{y}$: [from (2.7) and (a−4)]

$$\forall (\xi, \eta) \epsilon P^2(\boldsymbol{x}, \boldsymbol{y}) : R(\xi, \eta) = R(\boldsymbol{x}, \boldsymbol{y})$$

(b-5) Maximum for monotonic increase and minimum for monotonic decrease[11] (i.e., Monoinc-Max·Monodec-Min property, or MMMM property)

(i) If $\forall i \neq j: x_i > x_j \Longrightarrow y_i \geqq y_j$, and

if $\forall i \neq j: y_i > y_j \Longrightarrow x_i \geqq x_j$, then

$$(\boldsymbol{x}, \boldsymbol{y}) \epsilon \underset{\substack{\xi \epsilon P(\boldsymbol{x}) \\ \eta \epsilon P(\boldsymbol{x})}}{arg\ Max}\ R(\xi, \eta), \qquad \text{[from (2.7) and (a − 5)]}$$

(ii) If $\forall i \neq j$: $x_i > x_j \Longrightarrow y_i \leqq y_j$, and

if $\forall i \neq j$: $y_i > y_j \Longrightarrow x_i \leqq x_j$, then

$$(\boldsymbol{x}, \boldsymbol{y}) \in \underset{\substack{\boldsymbol{\xi} \in P(\boldsymbol{x}) \\ \boldsymbol{\eta} \in P(\boldsymbol{x})}}{arg\ Min}\ R(\boldsymbol{\xi}, \boldsymbol{\eta}), \qquad \text{[from (2.7) and (a − 5)]}$$

4 Mathematical Characteristics of ROXY Index : Different from Correlation Coefficient

The ROXY index has been developed as an analytical instrument to investigate the stage and speed of the process of spatial convergence and divergence, while the correlation coefficient has been developed as an analytical tool to measure the intensity of linear association between two variables. Therefore, some mathematical characteristics of the ROXY index are different from those of the coefficient. We discuss in this section two examples on the difference between the ROXY index and correlation coeficient.

4.1 Dependency on Parallel-shift Transformation

As to the correlation coefficient, even if we add (or subtract) a given constant to (or from) every element of vectors $\boldsymbol{x}$ and another given constant to (or from) every element of vector $\boldsymbol{y}$, the value of the correlation coefficient would remain the same. This characteristic of independency of the parallel-shift transformation on vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ can be expressed as

$$\rho(\boldsymbol{x} + \boldsymbol{a}, \boldsymbol{y} + \boldsymbol{b}) = \rho(\boldsymbol{x}, \boldsymbol{y}) \tag{5.1}$$

where all elements of $\boldsymbol{a}$ are a and all elements of $\boldsymbol{b}$ are b.

The parallel-shift transformation would, however, affect the ROXY-index value as shown in what follows;

$$\begin{aligned}
R(\boldsymbol{x} + \boldsymbol{a}, \boldsymbol{y} + \boldsymbol{b}) &= \frac{\sigma_{x+a} \cdot \sigma_{y+b}}{\overline{x+a} \cdot \overline{y+b}} \rho(\boldsymbol{x} + \boldsymbol{a}, \boldsymbol{y} + \boldsymbol{b}) && \text{[from (2.7)]} \\
&= \frac{\sigma_x \cdot \sigma_y}{\overline{x+a} \cdot \overline{y+b}} \rho(\boldsymbol{x}, \boldsymbol{y}) && \text{[from (5.1) and } Var(\boldsymbol{x} - a) = Var(\boldsymbol{x})\text{]} \\
&= \frac{\overline{x} \cdot \overline{y}}{\overline{x+a} \cdot \overline{y+b}} \cdot \frac{\sigma_x \cdot \sigma_y}{\overline{x} \cdot \overline{y}} \rho(\boldsymbol{x}, \boldsymbol{y}) \\
&= \frac{\overline{x} \cdot \overline{y}}{\overline{x+a} \cdot \overline{y+b}} R(\boldsymbol{x}, \boldsymbol{y}) && \text{[from (2.7)]}
\end{aligned}$$

4.2 Half-unboundedness of Value Range and Implications as Convergence-divergence Index

The range of the value of the correlation coefficient, $\rho(\boldsymbol{x},\boldsymbol{y})$, remains between -1 and $+1$. Concerning the range of ROXY index R $(\boldsymbol{x},\boldsymbol{y})$, consider the case when the ROXY-index method is applied to an inter-metropolitan analysis. In this situation, generally speaking, the sign of y_i which indicates the annual growth ratio of the population of metropolitan area i, is positive for all i, and so is the sign of x_i which indicates the population of metoropolitan area i. In order to obtain the minimum and maximum values of the ROXY index for this case, let y_i approach to infinity with y_i (for all j but i) remaining the original fined values[12]. Then, we have ;

$$\lim_{y_i\to\infty} R(x,y) = \lim_{y_i\to\infty}\frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{n\bar{x}\cdot\bar{y}} \qquad \text{[from (2.5)]}$$

$$= \frac{(x_i-\bar{x})-\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})}{n\cdot\frac{1}{n}\cdot\bar{x}}$$

$$= \frac{x_i}{\bar{x}}-1 \qquad (5.1)$$

Equation (5.1) that, if x_i takes its maximum value at $i=imax$, and minimum value at $i=imin$, then the ROXY index has (i) *its maximum value of* $(x_{imax}/\bar{x})-1$ when $y_{imin}>0$ and other y_i*'s are zero, and* (ii) *its minimum value of* $(x_{imin}/\bar{x})-1$, when $y_{imin}>0$ and other y_i*'s are zero. Accordingly, it can be noticed that in the general situation ;*

$$\frac{x_{imin}}{\bar{x}}-1 < ROXY(x,y) < \frac{x_{imax}}{\bar{x}}-1.$$

Therefore, since x_{imin} *can be close to* $+0$, *we have the possible range of the value of ROXY* (x, y) as ;

$$-1 < ROXY(x,y) < +\infty.$$

This implication was, as a matter of fact, already pointed out in Hiraoka and Kawashima (1993), for the case where the CBD distance is employed as the weighting factor in an intra-metropolitan analysis[13].

Let us now examine the derivative of the ROXY index $R(x,y)$ with respect to y_i in order to obtain a better understamding on how the marginal change of y_i would contribute to the change in the ROXY-index value. Differentiating $R(x,y)$ by y_i, for this purpose, we have;

$$\frac{d}{dy_i}R(x,y) = \frac{1}{n}\cdot\frac{1}{(\bar{x}\cdot\bar{y})^2}\cdot\left[\left\{(x_i-x)-\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})\right\}(\bar{x}\cdot\bar{y}) - \left\{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})\right\}\frac{\bar{x}}{n}\right]$$

$$= \frac{\bar{x}}{n(\bar{x}\cdot\bar{y})^2}[\{(x_i-\bar{x})\cdot\bar{y}-Cov(x,y)].$$

Therefore, when y_i increases, the followings can be pointed out :

(i) The value of the ROXY index increases if

$$x_i > \bar{x} + \frac{1}{\bar{y}} Cov(\boldsymbol{x}, \boldsymbol{y}) .$$

(ii) The value of the ROXY index does not change if

$$x_i = \bar{x} + \frac{1}{\bar{y}} Cov(\boldsymbol{x}, \boldsymbol{y}) .$$

(iii) The value of ROXY index decreases if

$$x_i < \bar{x} + \frac{1}{\bar{y}} Cov(\boldsymbol{x}, \boldsymbol{y}) .$$

Therefore, it can be seen (1) that the rate of increase in the value of the ROXY index would be maximized when y_{imax} marginally increases as compared with the marginal increase in the value of y_i (for $i \neq imax$), and (2) that the rate of decrease in the value of the ROXY index would be maximized when y_{imin} marginally increases as compared with the marginal increase in the value of y_i (for $i \neq imin$). From what has been discussed above, can be derived the following implications ;

(i) If $\forall i$: $x_i \geqq x_{i+1}$ and $y_1 > x_1$, and $\forall i \neq 1$: $y_i = z_i$, then $R(\boldsymbol{x}, \boldsymbol{y}) > R(\boldsymbol{x}, \boldsymbol{z})$.

(ii) If $\forall i$: $x_i \geqq x_{i+1}$ and $y_n > z_n$, *and* $\forall i \neq n$: $y_i = z_i$, then $R(\boldsymbol{x}, \boldsymbol{y}) < R(\boldsymbol{x}, \boldsymbol{z})$.

In one aspect, the aforementioned in this section can perhaps be interpreted as the reasonably desirable and necessarily unavoidable characteristics of the ROXY index which we employ to measure the magnitude of the of spbtial convergence-divergence.

For the purpose of providing a somewhat more concrete idea on a part of what we have discussed in this subsection, let us look at the numerical examples given in Table 1, where we have Cases 1 and 2. Note that the difference between the two cases is only on the value of the third element of vector $\boldsymbol{y}$; 3 for Case 1, and 6 for Case 2. In Case 1, there exists a strict linear relationship of $\boldsymbol{x} = 2\boldsymbol{y}$ between the two vectors $\boldsymbol{x}$ and $\boldsymbol{y}$. In Case 2, however there exists no strict linear relationship between the two vectors $\boldsymbol{x}$ and $\boldsymbol{y}$. Instead, in Case 2, the set of the two vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ reflects the phenomenon of a relative convergence toward each vector's third element which is larger than any other element in each vector. From Table 1, we can see the followings:

(1) The value of the correlation coefficient (1.000) for Case 1 is greater than that (0.945) for Case 2. It is because the correlation coefficient is constructed to show the strength of the linear relationship between $\boldsymbol{x}$ and $\boldsymbol{y}$.

(2) The value of the ROXY index (0.167) after the elimination of munltiplier component of 10^4 for Case 1, is smaller than that (0.278) for Case 2. It is because the ROXY index is constructed to show the magnitude of the convergence-divergence processes.

6 Conclusion

In the present paper, we have tried to illuminate the basic mathematical characteristics of the ROXY index developed in the field of regional science. In doing so, we have compared both common and different characteristics between the value of the ROXY index and that of the correlation coefficient. The major results of our investigation in this paper is summarized in Table 2. This study would hopefully be helpful for us to obtain a better insight into the potential applicability of the ROXY index to the empirical studies on the spatial dynamic changes.

Table 1 Numerical Examples

Items	Case 1	Case 2
Vectors	$x=\begin{pmatrix}2\\4\\6\end{pmatrix},\quad y=\begin{pmatrix}1\\2\\3\end{pmatrix}$	$x=\begin{pmatrix}2\\4\\6\end{pmatrix},\quad y=\begin{pmatrix}1\\2\\6\end{pmatrix}$
Averages	$\bar{x}=4,\quad \bar{y}=2$	$\bar{x}=4,\quad \bar{y}=3$
Standard deviations	$\sigma_x{}^2=\frac{8}{3},\quad \sigma_y{}^2=\frac{2}{3}$	$\sigma_x{}^2=\frac{8}{3},\quad \sigma_y{}^2=\frac{14}{3}$
Covariances	$Cov(x,y)=\frac{4}{3}$	$Cov(x,y)=\frac{10}{3}$
Correlation coefficients	$\rho(x,y)=1$	$\rho(x,y)=\frac{5\sqrt{7}}{14}$ $=0.945$
ROXY indices without multiplier of 10^4	$R(x,y)=\frac{1}{6}$ $=0.167$	$R(x,y)=\frac{5}{18}$ $=0.278$

Table 2 ROXY Index and Correlation Coefficient

Items	ROXY index $R(x, y)$	Correlation coefficient $\rho(x, y)$
What to measure	Index to measure magnitude of convergence and divergence	Index to measure intensity of linear relationship
Mathematical relationships	$R(x,y) = \frac{\sigma_x \cdot \sigma_y}{\bar{x} \cdot \bar{y}} \rho(x,y)$ $b_{y \cdot x} = \frac{\bar{x} \cdot \bar{y}}{\sigma_x^2} R(x,y) = \frac{\sigma_y}{\sigma_x} \rho(x,y)$	
Similarities	$R(x,y) = R(y,x)$	$\rho(x,y) = \rho(y,x)$
	$R(px,py) = R(x,y)$	$\rho(px,py) = \rho(x,y)$
	$R(ax,y) = R(x,by)$ $= R(cx,dy)$ $= R(x,y)$ $(a, b, c, and\ d > 0)$	$\rho(ax,y) = \rho(x,by)$ $= \rho(cx,dy)$ $= \rho(x,y)$ $(a, b, c, and\ d > 0)$
	Order-invariance	Order-invariance
	MMMM property	MMMM property
Differences	$R(x+a, y+b) = \frac{\bar{x} \cdot \bar{y}}{\overline{x+a} \cdot \overline{x+b}} R(x,y)$	$\rho(x+a, y+b) = \rho(x,y)$
	$\frac{x_{imin}}{\bar{x}} - 1 \leq R(x,y) \leq \frac{x_{imax}}{\bar{x}} - 1$, $(x, y > 0)$	$-1 \leq \rho(x,y) \leq +1$,
	Larger value for more convergence	Larger absolute value for more linearity

Notes (1) $R(x,y)$ indicates the value of the ROXY index after elimination of the multiplier component of 10^4.

(2) p : permutation operator

(3) For order-invariance and MMMM property, see details in the text.

Notes

1) The basic concept of the ROXY index was originally constructed and applied in an empirical study by Kawashima (1978, pp.9, 13 and 14) in which the index was called ROX instead of ROXY. The ROX stands for the *R*atio *O*f "weighted average to simple average" (*X*), while ROXY stands for the *R*atio *O*f "weighted average" (*X*) to "simple average" (*Y*).
2) The *weighting variable* is the variable that is used in the ROXY-index analysis as weighting factor necessary for the calculation of the weighted average over the value of the *principally-concerned variable.* In case we are interested in investigating the phenomena of population centralization and decentralization (*i.e.*, phenomena of urbanization and suburbanization) within a metropolitan area, we often employ as weighing variable the physical distance from the CBD (central business district) to each locality in the metropolitan area. This distance is referred to as physical CBD distance, or simply CBD-distance. In this instance, we need not have the upper suffix t of the notation x_i^t since the CBD distance remains fixed over time. On the other hand, if we use the time-distance from the CBD to each locality, then we have to stick to the present notation x_i^t since the time-distance from the CBD tends to change as time goes on.
3) The *principally-concerned variable* is the variable for which we would be intererted in calculating the ROXY-index value.
4) Vector $\boldsymbol{x}$ is the one whose elements represent values of the weighting variable, while vector $\boldsymbol{y}$ is the one whose elements preprerent values of the principally-concerned variable. The vector $\boldsymbol{1}$ is the unit vector.
5) For the derivation of this relationship, see Hays (1963, p.603).
6) The relationship between $R(\boldsymbol{x},\boldsymbol{y})$ and $b_{y.x}$ was first discussed in Kawashima (1986).
7) For the derivation of this relationship, see Hays (1963, p.609).
8) As to the permutation operator $\boldsymbol{p}$, in case $n=3$ for example, the set $\boldsymbol{\Sigma}_3$ which consists of all possible permutation operators can be expressed as follows;

$$\boldsymbol{\Sigma}_3=\left\{\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix},\begin{bmatrix}1&0&0\\0&0&1\\0&1&0\end{bmatrix},\begin{bmatrix}0&1&0\\1&0&0\\0&0&1\end{bmatrix},\begin{bmatrix}0&1&0\\0&0&1\\1&0&0\end{bmatrix},\begin{bmatrix}0&0&1\\1&0&0\\0&1&0\end{bmatrix},\begin{bmatrix}0&0&1\\0&1&0\\1&0&0\end{bmatrix}\right\}.$$

9) This characteristics would imply that, even though we interchange the vector $\boldsymbol{x}$ for the weighting variable and the vector $\boldsymbol{y}$ for the principally-concerned variable, the value of the ROXY index would remain unchanged. Consequently, when we are interested in the intra-metropolitan analysis mentioned in the above note 2), we have the identical value of the ROXY index for the following two cases; (i) the case in which we have the CBD distance as the weighting variable and the growth ratio of each locality as the principally-concerned variable, and (ii) the case in which we have the growth ratio as the weighting variable and the CBD distance of each locality as the principally-concerned variable. This fact provides us with a desirable flexibility for the conceptualization of

the meanings of the ROXY index when we examine the process of the spatial centralization and decentralization.

10) This characteristics together with the above characteristics of (a−3), would indicate that the ROXY index and the correlation coefficient are both *scale-invariant (i.e.,* physically dimensionless). That is, their values are independent of the scale-unit employed to measure x and y. In contrast with this, the non-standardized regression coefficient is not scale-invariant since its value varies depending on the *scale-unit* applied to x and y. See Kawashima (1985,1986) for a partial discussion on the this subject.

11) In the context of the ROXY index, this characteristics can be paraphrased as follows: the ROXY-index value would be maximized when the most intensive form of spatial concentration, in terms of growth ratio, would take place for a given set of a vector $\boldsymbol{x}$ for the weighting variable and a vector $\boldsymbol{y}$ for the principally-concerned variable, while it would be minimized when the most intensive form of deconcentration would take place.

12) This operation is equivalent to the operation that makes y_j / y_i approach to zero for all j but i.

13) In case we employ the CBD distance d_i as the weighting factor in an intra-metropolitan analysis, the maximum and minimum values of the ROXY index are respectively given as follows;

$$\text{maximum value} = \frac{d_{max}}{\bar{d}} - 1, \text{ and}$$

$$\text{minimum value} = \frac{d_{min}}{\bar{d}} - 1,$$

where

d_{max} : maximum value of CBD distance,

d_{min} : minimum value of CBD distances, and

$\bar{d}$: average of CBD distance which is defined as $\frac{\sum_{i=1}^{n} d_i}{n}$.

References

Hays, L.W., 1963, *Statistics*, Fifth Edition, Harcourt Brace College Publishers, Texas, USA.

Hiraoka, N., and T. Kawashima, 1993, "Mathematical Characteristics of ROXY Index (Ⅱ): Periods of Intra-metropolitan Spatial-cycle Paths and Theoretically-ideal Formulations of ROXY Index," *Gakushuin Economic Papers*, Vol.30, No.3, Gakushuin University, Tokyo, November, pp.317-422.

Hiraoka, N., and T. Kawashima, 1994, "Mathematical Characteristics of ROXY Index (Ⅲ): Functional Relationship between 'Theoretically-ideal ROXY Index with CBD Distance Used as Weighing Factor' and 'That with Reversed CBD Distance,'" *Gakushuin Economic Papers*, Vol.30, No.4, Gakushuin University, Tokyo, February, pp.451-478.

Kawashima, T., 1978, "Recent Urban Evolution Processes in Japan: Analysis of Functional Regions," Mimeograph, Presented at the 25th North American Meetings of the Regional Science Association, Chicago, Illinois, USA, November.

Kawashima, T., 1985, "ROXY Index : An Indicative Instrument to Measure The Speed of Spatial Concentration and Deconcentration of Population," *Gakushuin Economic Papers*, Vol. 22, No. 2, Gakushuin University, Tokyo, September, pp.183-213.

Kawashima, T., 1986, "Speed of Suburbanjization: ROXY Index Analysis for Intra-metoropolitan Spatial Redistribution of Population in Japan,"*Gakushuin Economic Papers*, Vol. 22, No. 3, Gakushuin University, Tokyo, March, pp.243-304.

Kawashima, T., and N. Hiraoka, 1993, "Mathematical Characteristics of ROXY Index (I): Distance and Reversed Distance Used as Weighing Factors,"*Gakushuin Economic Papers,*, Vol. 30, No. 2, Gakushuin University, Tokyo, July, pp.255-297.