Normalizing biproportional methods [1]

Louis de Mesnard
LATEC (UMR CNRS 5118)
Faculty of Economics and Management - University of Dijon
2 Bd Gabriel, B.P. 26611,
F-21066 Dijon Cedex,
FRANCE
E-mail: *louis.de-mesnard@u-bourgogne.fr*

ABSTRACT. Biproportional methods are used to update matrices: the projection of a matrix **Z** to give it the column and row sums of another matrix is **R Z S**, where **R** and **S** are diagonal and secure the constraints of the problem (**R** and **S** have no signification at all because they are not identified). However, normalizing **R** or **S** generates important mathematical difficulties: it amounts to put constraints on Lagrange multipliers, non negativity (and so the existence of the solution) is not guaranteed at equilibrium or along the path to equilibrium.

KEYWORDS. Biproportion, RAS, normalization, Stone, Leontief.

AMS subject classifications. 15A15, 14N05, 65Q05

JEL classification. C63, C67, D57

---

[1] Paper published in *The Annals of Regional Science,* 2002, 36, 1: 139-144.

1 Introduction

Biproportional methods (and the RAS method created by Nobel laureate Richard Stone) are well known to update matrices **Z** (that can be matrices of commodity flows, of transportation flows, of demographic flows, of physical flows, etc., or other type of homogenous matrices as contingency tables). Consider two non-negative matrices $\mathbf{Z}^0$ and $\mathbf{Z}^1$, homogenous by rows and columns (i.e., the sum of their elements can be computed by rows and columns), that can be the same matrix at two different dates, or two different but comparable matrices over space. By a biproportion, one computes a matrix $\hat{\mathbf{Z}}$ that is the nearest to another matrix $\mathbf{Z}^0$ (following a certain criteria as the minimization of information or the maximization of entropy) but with the row and column sums (the *margins*) of a third matrix $\mathbf{Z}^1$: $\sum_j \hat{z}_{ij} = z^1_{i\bullet}$, for all i and $\sum_i \hat{z}_{ij} = z^1_{\bullet j}$, for all j ($z_{i\bullet}$ and $z_{\bullet j}$ denote $\sum_j z_{ij}$ and $\sum_i z_{ij}$ respectively). The result is given by $\hat{\mathbf{Z}} = K(\mathbf{Z}^0,\ \mathbf{Z}^1) = \mathbf{R}\ \mathbf{Z}^0\ \mathbf{S}$, where $K(\ \ )$ is the operator of biproportion and where **R** and **S** are diagonal. As it was demonstrated (Mesnard, 1994) that all algorithms computing a biproportion (RAS and any other) are equivalent and give the same result, one can choose one of the most simple to handle, those of Bachem and Korte (1979). So, one is able to see factors **R** and **S** as Lagrange or Kuhn-Tucker multipliers obtained by, say, the minimization of the quantity of information: $\min_{\hat{z}_{ij}}\ \mathrm{I} = \sum_i \sum_j \hat{z}_{ij}\ \log \frac{\hat{z}_{ij}}{z^0_{ij}}$, s.t. $\sum_i \hat{z}_{ij} = z^1_{\bullet j}$ for all j and $\sum_j \hat{z}_{ij} = z^1_{i\bullet}$ for all i :

(1) $$r_i^{(k+1)} = \frac{z^1_{i\bullet}}{\sum_{j=1}^{m} s_j^{(k)}\ z^0_{ij}} \text{ for all } i, \text{ and } s_j^{(k+1)} = \frac{z^1_{\bullet j}}{\sum_{i=1}^{n} r_i^{(k+1)}\ z^0_{ij}} \text{ for all } j$$

They are transcendent, and their value is found only iteratively. After an initialization, for example by $s_j^{(0)} = 1$, for all *j*, this leads to an equilibrium:

$$(2) \qquad r_i^* = \frac{z_{i\bullet}^1}{\sum_{j=1}^{m} s_j^* \, z_{ij}^0} \text{ for all } i, \text{ and } s_j^* = \frac{z_{\bullet j}^1}{\sum_{i=1}^{n} r_i^* \, z_{ij}^0} \text{ for all } j$$

Initialization can start also by $r_i^{(0)}$ and nothing obliges to choose the same uniform value for all $s_j^{(0)}$ (or all $r_i^{(0)}$). The solution exists, is unique and convergent under some conditions of existence: when you initialize by an initial and arbitrary $\mathbf{R}^{(0)}$ or $\mathbf{S}^{(0)}$, you find two unique $\mathbf{R}^*$ and $\mathbf{S}^*$ matrices at equilibrium, after an iterative numerical computation; it was demonstrated for RAS by Bacharach (1970).

It is possible to give various interpretations to factors **R** and **S**. For example, RAS can be used to project matrices of technical coefficients. Technical coefficients, created by Nobel laureate Wassily Leontief (1986), are defined by $a_{ij} = \frac{z_{ij}}{x_j}$, where z_{ij} is the quantity of commodity *i* that flows from an economic agent *i* (that can be a sector of production, an industry, or any productive agent) to an economic agent *j* and x_j is the output of *j*, with the accounting identity $\sum_i z_{ij} + v_j = x_j$ for all *j* (where v_j is the value-added of *j*). Stone and Brown (1962), followed by others (Paelinck and Waelbroeck, 1963) (Snower, 1990), give an interpretation of **R** and **S** in terms of *absorption effect* and *fabrication effect*. Consider two matrices of technical coefficients $\mathbf{A}^0$ and $\mathbf{A}^1$ for two years $t = 0$ and $t = 1$ (evaluated at the prices of $t = 1$). Then $\hat{\mathbf{A}} = K(\mathbf{A}^0,\ \mathbf{A}^1) = \mathbf{R}\ \mathbf{A}^0\ \mathbf{S}$, where $\hat{\mathbf{A}}$ has the same margins than $\mathbf{A}^1$. **R** and **S** are interpreted as the absorption-substitution effect and the fabrication-transformation effect respectively. A similar explanation can be adopted when you consider transportation, demographic or physical flows are considered; for contingency matrices, one must find another interpretation.

However, as the terms r_i and s_j are **not identified**, they **cannot** be interpreted for themselves and all these interpretations fall. Not identified means that, following Bacharach (1970: 22),

$\mathbf{R\,A\,S} = (\lambda\,\mathbf{R})\,\mathbf{A}\,(\lambda^{-1}\,\mathbf{S})$ for any nonzero scalar λ; or, in (1), if you multiply $r_i^{(0)}$ by λ, s_j^* will be divided by λ and r_i^* multiplied by λ. This removes all significance of **R** and **S** in terms of fabrication or absorption effects. Remark that the products $r_i\,s_j$, for all $(i,\ j)$, are **identified**, i.e., $\hat{r}_i^*\,\hat{s}_j^* = r_i^*\,s_j^*$ for all $(i,\ j)$: it remains permissible to conduct a decomposition of matrix change over time (Mesnard, 1990, 1997) (Van der Linden and Dietzenbacher, 1995) because this one is based on the computation of the product $\mathbf{R}\,\mathbf{Z}^0\,\mathbf{S}$ to be compared to $\mathbf{Z}^1$ and not on a particular interpretation of **R** or **S**. A similar property can be found with other methods based on an iterative algorithm, for example, the bicausative method (Mesnard, 2000).

To avoid non identification, Bacharach has proposed to make a normalization (1970: 22) of factors **R** and **S**: such normalization could be simple, as $r_1 = 1$ or $\sum_i r_i = 1$, but it has always an arbitrary character, so it does not solves the problem of interpreting correctly the *r* and *s* terms. Van der Linden and Dietzenbacher (1995: 129) have taken up this idea: on the base of an economic reasoning, they normalize **R**, arguing that the global substitution effect must be equal to zero for the whole economy (x_j^1 is the output of sector *j* at year 1):

$$(3) \qquad \sum_i \sum_j r_i^*\, a_{ij}^0\, s_j^* \frac{x_j^1}{\sum_j x_j^1} - \sum_i \sum_j a_{ij}^0\, s_j^* \frac{x_j^1}{\sum_j x_j^1} = 0$$

Other type of justifications can be found to justify a normalization of **R** or **S**: it is not the aim of this note to discuss the validity of one interpretation or the other, but to demonstrate that the normalization of **R** or **S** is not valid. I will rely on the example of the above normalization.

2 Normalization: discussion

The *r* and *s* terms are not independent, but linked by (2), so, in (3) terms s_j^* have to be replaced by their expression in (2), that gives a nonlinear relation between the r_i^* only:

(4) $$\sum_j a^1_{\bullet j}\, x^1_j = \sum_j a^1_{\bullet j}\, x^1_j \left(\sum_i a^0_{ij}\right)\left(\sum_i a^0_{ij}\, r^*_i\right)^{-1}$$

The left member of this expression simplifies as a constant without r^*_i, while the right member is hyperbolic. However, (4) does not define exactly terms r^*_i: $f\,(r^*_1,\ r^*_2,\ \ldots,\ r^*_n) = 0$.

2.1 Constraints on Lagrange multipliers

For any type of normalization the r_i and s_j terms are a transformation of Lagrange multipliers, λ_i and μ_j, of an optimization process when the quantity of information is minimized:

min $\sum_i \sum_j \hat{a}_{ij} \log \dfrac{\hat{a}_{ij}}{a^0_{ij}}$, s.t. $\sum_j \hat{a}_{ij} = a^1_{i\bullet}$ (multiplier: λ_i) for all i, and $\sum_i \hat{a}_{ij} = a^1_{\bullet j}$ for all j (multiplier: μ_j). After a changing of variables, this leads to $r_i = \exp{-(1+\lambda_i)}$ for all i and $s_j = \exp{-\mu_j}$ for all j. As the r and s are simply a transformation of the multipliers, normalization appears to be a constraint on the Lagrange multipliers, which is unusual and in contradiction with the concept of Lagrange or Kuhn-Tucker multipliers.

2.2 Negativity of r and s terms

Normalization of substitution effects could bring some additional difficulties because if the global substitution effect is zero, (4) shows that some terms could be positive and some negative. This implies that some $\hat{z}_{ij}$ could be negative, what violates the hypotheses (r and s are exponentials but as they are not identified, *they* can be all positive or all negative without any problem).

Example. The following example indicates the computation of the constraint r terms at equilibrium in a 2x2 case. Consider:

$$\mathbf{A}^0 = \begin{bmatrix} 0.1 & 0.2 \\ 0.3 & 0.1 \end{bmatrix}, \mathbf{A}^1 = \begin{bmatrix} 0.4 & 0.9 \\ 0.5 & 0.8 \end{bmatrix}, x_1^1 = 20, x_2^1 = 25.$$

In the normalization expression,

$$r_1^* \; 0.1 \; s_1^* \; 20 + r_1^* \; 0.2 \; s_2^* \; 25 + r_2^* \; 0.3 \; s_1^* \; 20 + r_2^* \; 0.1 \; s_2^* \; 25$$
$$= 0.1 \; s_1^* \; 20 + 0.2 \; s_2^* \; 25 + 0.3 \; s_1^* \; 20 + 0.1 \; s_2^* \; 25$$

the terms $s_1^* = \dfrac{0.5}{0.1 \; r_1^* + 0.3 \; r_2^*}$ and $s_2^* = \dfrac{0.8}{0.2 \; r_1^* + 0.1 \; r_2^*}$ have to be inserted:

$$0.1 \; r_1^* \; \frac{0.5}{0.1 \; r_1^* + 0.3 \; r_2^*} \; 20 + 0.2 \; r_1^* \; \frac{0.8}{0.2 \; r_1^* + 0.1 \; r_2^*} \; 25$$
$$+ \; 0.3 \; r_2^* \; \frac{0.5}{0.1 \; r_1^* + 0.3 \; r_2^*} \; 20 + 0.1 \; r_2^* \; \frac{0.8}{0.2 \; r_1^* + 0.1 \; r_2^*} \; 25$$
$$= 0.1 \; \frac{0.5}{0.1 \; r_1^* + 0.3 \; r_2^*} \; 20 + 0.2 \; \frac{0.8}{0.2 \; r_1^* + 0.1 \; r_2^*} \; 25$$
$$+ 0.3 \; \frac{0.5}{0.1 \; r_1^* + 0.3 \; r_2^*} \; 20 + 0.1 \; \frac{0.8}{0.2 \; r_1^* + 0.1 \; r_2^*} \; 25$$
$$\Leftrightarrow 30 = \frac{4}{0.1 \; r_1^* + 0.3 \; r_2^*} + \frac{6}{0.2 \; r_1^* + 0.1 \; r_2^*}$$

That is, $r_2^* = 1.2222 - 1.1667 \; r_1^* + 5.5556 \times 10^{-2} \sqrt{\left(484 - 420 \; r_1^* + 225 \; (r_1^*)^2\right)}$

Figure 1 about here

This negative slope curve can produce negative values for r_2^* as soon as $r_1^* > 2.3332$. Note an ordinary normalization as $r_1^* + r_2^* = \delta$, a line with a slope equal to -1, would generate a similar problem.■

So, positivity has to be added as two additional constraints: $r_i^* \geq 0$ for all *i* and $s_j^* \geq 0$ for all *j*. One has to choose a normalization such as no *r* or *s* terms become negative at equilibrium: normalization is no more arbitrary. Note that this is imposed at equilibrium and not at initialization: it is not sure that normalization of (4) gives a result respecting the new constraints. This leads to the next section.

2.3 Iteration and the path to equilibrium

The equilibrium value of *r* and *s* terms must be found only iteratively: condition (3), i.e., $\sum_i r_i^* \sum_j a_{ij}^0 s_j^* x_j^1 - \sum_i \sum_j a_{ij}^0 s_j^* x_j^1 = 0$, has to be set not only at equilibrium but also at each step of the iterative computation of *r* and *s* terms, that is:

$$\text{(5)} \qquad \sum_i r_i^{(k+1)} \sum_j a_{ij}^0 s_j^{(k)} x_j^1 = \sum_i \sum_j a_{ij}^0 s_j^{(k)} x_j^1$$

Terms $s_j^{(k)}$ can be replaced by $s_j^{(k)} = \dfrac{a_{\bullet j}^1}{\sum_{i=1}^n r_i^{(k)} a_{ij}^0}$ from (1):

$$\text{(6)} \qquad \sum_j a_{\bullet j}^1 x_j^1 \left(\sum_i a_{ij}^0 r_i^{(k+1)}\right)\left(\sum_{i=1}^n a_{ij}^0 r_i^{(k)}\right)^{-1} = \sum_j a_{\bullet j}^1 x_j^1 \left(\sum_i a_{ij}^0\right)\left(\sum_{i=1}^n a_{ij}^0 r_i^{(k)}\right)^{-1}$$

The left member of (6) is no more a constant as in (4). One has to demonstrate that there is a path that respects the above constraints of positivity at each step *k* from the initialization to equilibrium for *r* and *s* terms. Even when it is assumed that the solution exists for a particular matrix, **A**, one must demonstrate that this solution can be reached without passing by some negative values of $r_i^{(k)}$ or $s_j^{(k)}$; else it is necessary to impose the two additional constraints $r_i^{(k)} \geq 0$ and $s_j^{(k)} \geq 0$ for all (i, j). Knowing an acceptable positive solution, it could seem attractive to find a correct initialization of the process that corresponds to this rule, but it is impossible because the problem is transcendent.

<u>Example</u>. Iterative computation of the constraint *r* terms. The normalization expression is:

$$\left[0.1\, r_1^{(k+1)} + 0.3\, r_2^{(k+1)}\right] s_1^{(k)}\, 20 + \left[0.2\, r_1^{(k+1)} + 0.1\, r_2^{(k+1)}\right] s_2^{(k)}\, 25 = 0.4\, s_1^{(k)}\, 20 + 0.3\, s_2^{(k)}\, 25$$

The iterative terms are:

$$r_1^{(k+1)} = \frac{0.4}{0.1\, s_1^{(k)} + 0.2\, s_2^{(k)}}, \qquad r_2^{(k+1)} = \frac{0.9}{0.3\, s_1^{(k)} + 0.1\, s_2^{(k)}},$$

$$s_1^{(k+1)} = \frac{0.5}{0.1\, r_1^{(k+1)} + 0.3\, r_2^{(k+1)}}, \quad s_1^{(k+1)} = \frac{0.8}{0.2\, r_1^{(k+1)} + 0.1\, r_2^{(k+1)}}.$$

Replacing the terms *s* in the normalization expression (5) gives:

$$10\,\frac{0.1\, r_1^{(k+1)} + 0.3\, r_2^{(k+1)}}{0.1\, r_1^{(k)} + 0.3\, r_2^{(k)}} + 20\,\frac{0.2\, r_1^{(k+1)} + 0.1\, r_2^{(k+1)}}{0.2\, r_1^{(k)} + 0.1\, r_2^{(k)}} = \frac{4}{0.1\, r_1^{(k)} + 0.3\, r_2^{(k)}} + \frac{6}{0.2\, r_1^{(k)} + 0.1\, r_2^{(k)}}. \blacksquare$$

3 Conclusion

R and **S**, the equilibrating factors of biproportional methods have no signification at all by themselves because they are not identified. The interesting attempt at correction consisting of a normalization of **R** or **S** generates some important mathematical difficulties that prevent its consideration as acceptable: it amounts to putting constraints on Lagrange multipliers, positivity (and so the existence of the solution) is not achieved at equilibrium and along the path to equilibrium.

4 Bibliographical references

BACHARACH, Michael, 1970. *Biproportional Matrices and Input-Output Change*. Cambridge University Press, Cambridge.

BACHEM, Achim and Bernhard KORTE, 1979. "On the RAS-Algorithm," *Computing* 23: 189-198.

LEONTIEF W., 1986. Input-Output Economics. Oxford University Press, New-York and Oxford, 2nd Edition.

Van der LINDEN J. and E. DIETZENBACHER, 1995. "The nature of changes in the EU cost structure of production 1965-85: an RAS approach", in H.W. ARMSTRONG and R.W. VICKERMAN, Eds, *Convergence and divergence among European Regions*, Pion Limited, London, pp. 124-139.

de MESNARD, Louis, 1990. "Biproportional Method for Analysing Interindustry Dynamics: the case of France", *Economic Systems Research*, Vol. 2, No 3: 271-293.

______ 1994. "Unicity of Biproportion", *SIAM Journal on Matrix Analysis and Applications* 2, 15: 490-495.

______ 1997. "A biproportional filter to compare technical and allocation coefficient variations", *Journal of Regional Science*, 37, 4: 541-564.

______ 2000. "Bicausative matrices to measure structural change: are they a good tool?", *The Annals of Regional Science*, 34, 3: 421-449.

PAELINCK, Jean and J. WAELBROECK, 1963. "Etude empirique sur l'évolution de coefficients «input-output». Essai d'application de la procédure «RAS» de Cambridge au tableau interindustriel belge", *Economie Appliquée*, 6: 82-111.

SNOWER, Dennis J., 1990. "New methods of updating input-output matrices", *Economic Systems Research*, 2, 1: 27-37.

STONE R. and J.A.C. BROWN, 1962. "A long-term growth model for the British economy", in: R.C. GEARY (ed.), *Europe's future in figures*, North-Holland, Amsterdam.

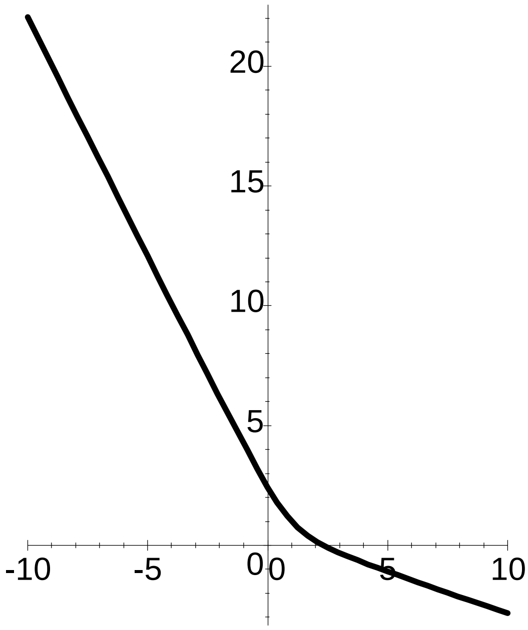

Figure 1. Function $r_2^* = \mathrm{f}(r_1^*)$ after normalization, at equilibrium