How to forecast the future input-output table?
-An approach based on historical table series

Wen Long
(Graduate University of Chinese Academy of Sciences; Research Center on Fictitious Economy and Data Science, Chinese Academy of Sciences, Beijing 100190, China)
Huiwen Wang
(School of Economics and Management, Beihang University, Beijing 100083, China)

Abstract: Since it is necessary to take much time and resources for compiling an input-output table, there inevitably exists time lag in all the finished tables, which has greatly weakened the analysis and application function of input-output table. Though the development of techniques of adjusting input coefficient has relieved disadvantage of time lag to some extent, yet this kind of methods is still inapplicable to predictive problem. Consider that internal constraints of the table, this paper proposes a new approach based on the tables series to solve the problem of forecasting the future input-output table.

Key Words: input-output table, forecast, constant sum constraint

1. Introduction

The input-output model of economics creates a picture of a regional economy describing flows to and from industries and institutions, which uses a matrix representation of a nation's (or a region's) economy to predict the effect of changes in one industry on others and by consumers, government, and foreign suppliers on the economy [1,2]. Wassily Leontief is credited with the development of this analysis.

Nowadays this technique has been widely applied in the economic analysis [3]. But in fact, an important problem cannot be ignored while applying input-output tables. Since it is necessary to take much time and resources for compiling an input-output table, there inevitably exists time lag in all the finished tables, which has greatly weakened the analysis and application functions of input-output table. For example, compiling an input-output table in China is based on direct decomposing method, which has taken longer time than the method of indirect deduction that is often adopted in western countries [4]. When using the direct consumption coefficients of years ago to analyze the present even forecast the future economic activities, it is obvious that the results will be distorted.

Therefore how to adjust and amend the direct consumption coefficients become an important problem and attract many researchers’ attentions [5~10]. One of the most representative approaches is RAS invented by Stone, the British economist, and all kinds of modified methods based on RAS. RAS method can complete the technical adjustments with relatively fast speed and a relatively small factor of the workload, which is widely applied all over the world.

Though the development of techniques of adjusting the direct consumption coefficients such as RAS and its extension methods have offered a kind of solution and relieved disadvantage of time lag to some extent, yet this kind of methods still have many limits and inapplicable to predictive problem.

This paper focuses on the problem of how to forecast the future input-output tables. As we all know, input-output table has been balanced, and there are some constraints in the rows and columns, which
causes that cells in the table are not dependent each other. For that reason, if we forecast the cells
directly so as to obtain the future table by assembling these predicted cells, the balance in the table is
nearly not assured. Hence, keeping the internal constraints of the table is the key problem of
forecasting the future tables. Consider that point, a new approach based on the historical table series
will be proposed in the paper, which can effectively solve the problem of forecasting the future
input-output table.

This paper is organized as follows. In the second section, the economic table has been abstracted
into a numereded matrix. Based on that, the balance relationship in the table is also expressed by
mathematic constraints in the matrix. The third section presents the model of this new predictive
approach in detail. Finally, this paper is concluded by the last section.

2. Simplification and constraints of IO table

From the viewpoint of mathematics, input-output table is a typical data sheet with
two-dimensional constraints, though it offers abundant economic information. That means the IO table
satisfies the constant sum constraints in both row and column direction.

As we know, there includes thousands of data in a complete IO table and it is necessary to
simplify the forecasting problem of IO table. Combining economic problem and mathematical problem,
this paper solves the difficulty by abstracting an IO matrix from an IO table, in the meantime keeping
the most important economic and mathematical information.

2.1 Simplification of IO table

For an IO table with \(m\) sectors, we can denote it using a following matrix:

\[
\begin{array}{ccc|ccc}
    a_{11} & \cdots & a_{1(m-1)} & a_{1m} & c_{11} & \cdots & c_{1k} \\
    \vdots & & \vdots & \vdots & \vdots & & \vdots \\
    a_{m1} & \cdots & a_{m(m-1)} & a_{mn} & c_{m1} & \cdots & c_{mk} \\
    \hline
    g_{11} & \cdots & g_{1(m-1)} & g_{1m} & \quad & \quad & \quad \\
    \vdots & & \vdots & \vdots & \quad & \quad & \quad \\
    g_{m1} & \cdots & g_{m(m-1)} & g_{mn} & \quad & \quad & \quad \\
\end{array}
\]

Fig. 1 Observed IO table

Obviously, there are 3 quadrants in an IO table. \(a_{ij}(i=1,\cdots,m; j=1,\cdots,m)\) in Quadrant I
represents intermediate input and use of every sector. \(c_{ij}(i=1,\cdots,m; j=1,\cdots,k)\) in Quadrant II
represents final use and import of product. \(g_{ij}(i=1,\cdots,l; j=1,\cdots,m)\) in Quadrant III represents
added value of every sector.

According to general simplification principle, the most important information should be
maintained and others can be eliminated. Therefore, we keep the whole of quadrant I and the total of
quadrant II and III, for quadrant I is the most important to IO table. So we get the following simplified IO table:

\[
\begin{array}{cccc}
 a_{11} & \cdots & a_{1(m-1)} & a_{1m} \\
 \vdots & \ddots & \vdots & \vdots \\
 a_{(m-1)1} & \cdots & a_{(m-1)(m-1)} & a_{(m-1)m} \\
 a_{m1} & \cdots & a_{m(m-1)} & a_{mm} \\
 \end{array}
\]

\[
\begin{array}{c}
 C_1 \\
 \vdots \\
 C_{m-1} \\
 C_m \\
 \end{array}
\]

Fig. 2 Simplified IO table

Comparing simplified and original IO table, we find that quadrant I is the selfsame and it is quadrant II and III that have been simplified. We denote

\[
C_i = \sum_{j=1}^{i} c_j, \quad G_j = \sum_{i=1}^{j} g_{ij}.
\]

The 3 quadrants constitute a complete IO table. There are close relationship in quadrant I, II and III, so they must be treated as a whole while forecasting IO table. According to integrity principle, we combine the 3 quadrants to a whole matrix, and the blank quadrant IV can be denoted as 0. We obtain IO matrix \( A \):

\[
A = \begin{bmatrix}
 a_{11} & \cdots & a_{1(m-1)} & a_{1m} \\
 \vdots & \ddots & \vdots & \vdots \\
 a_{(m-1)1} & \cdots & a_{(m-1)(m-1)} & a_{(m-1)m} \\
 a_{m1} & \cdots & a_{m(m-1)} & 0 \\
 \end{bmatrix}
\]

Here, \( a_{ij} (i = 1, \ldots, n-1, j = 1, \ldots, n-1) \) represents intermediate input and use of \( m \) sectors, \( a_m (i = 1, \ldots, n-1) \) represents total of final uses and \( a_{ij} (j = 1, \ldots, n-1) \) represents total of added values. Hence \( n = m + 1 \).

Perform the same process to every IO table, and a series of IO matrix can be obtained, denoted as \( A^1, A^2, \ldots, A^T \). In that way, the problem of forecasting IO table has been transformed into the one of forecasting future IO matrix \( A^{t+1} \) based on the known \( A^1, A^2, \ldots, A^T \).

### 2.2 Constraints of IO table

According to the economic characters of IO table, there exist the following constraints:

1. Total input = total output
2. Total intermediate input = total intermediate use
3. Total input of sector \( i \) = total output of sector \( i \)

Correspondingly, we denote these constraints in the IO matrix as follows:

\[
\begin{align}
\sum_{j=1}^{m} a_{ij} &= \sum_{i=1}^{n} a_{ij} \\
\sum_{j=1}^{m} a_{ij} &= \sum_{i=1}^{n} a_{ij} \\
\sum_{j=1}^{m} a_{ij} &= \sum_{i=1}^{n} a_{ij} \\
\sum_{j=1}^{m} a_{ij} &= \sum_{j=1}^{m} a_{ij}, \quad i = j, \quad i = 1, \ldots, n-1
\end{align}
\]
Obviously, constraint (2-2) is always satisfied.

While

$$
\sum_{j=1}^{n} \sum_{i=1}^{n} a_{ij} = \sum_{i=1}^{n} a_{ij},
$$

(3)

According to constraint (2-1) and (3), we get

$$
\sum_{i=1}^{n} a_{ii} = \sum_{j=1}^{n} a_{ij}.
$$

(4)

Equation (4) has given the economic information that total added values of all the sectors equal total final uses of all the sectors. Combine constraint (2-3) and (4), obtain

$$
\sum_{i=1}^{n} a_{ij} = \sum_{j=1}^{n} a_{ij}, \quad i = j, \ i = 1, \cdots, n.
$$

(5)

We simplify the equations (2) into equation (5), that means sum of row equals to sum of corresponding column in the IO table, that is, the constant sum constraints exist in both two directions. Therefore IO table is a kind of typical datasheet with two-dimensional constraints from the view of mathematics.

3. Forecast of IO table

Through the above process, we abstract an IO table with \((n-1)\) sectors into an \(n \times n\) dimensional IO matrix with constraints, defined as:

$$
A = \begin{bmatrix}
    a_{11} & \cdots & a_{1(n-1)} & a_{1n} \\
    \vdots & \ddots & \vdots \\
    a_{(n-1)1} & \cdots & a_{(n-1)(n-1)} & a_{(n-1)n} \\
    a_{nn} & \cdots & a_{n(n-1)} & 0
\end{bmatrix}_{n \times n}
$$

Here, \(\sum_{j=1}^{n} a_{ij} = \sum_{j=1}^{n} a_{ij}, \quad i = j, \ i = 1, \cdots, n\).

Similar with compositional data [11], the elements in matrix \(A\) vary under the constant constraints so that they are not permitted to build predictive model directly. Therefore, some transformation should be performed to matrix \(A\) to eliminate the influence of constraints.

First we discuss the problem of forecasting IO matrix on condition that \(n=3\).

Step 1. Release constraints in IO matrix and forecast the transformed matrix.

For any time \(t = 1, 2, \cdots, T\), define a IO matrix as \(X^t\):

$$
X^t = \begin{bmatrix}
x_{11}^t & x_{12}^t & x_{13}^t \\
x_{21}^t & x_{22}^t & x_{33}^t \\
x_{31}^t & x_{32}^t & 0
\end{bmatrix}_{3 \times 3}
$$

According to formula (5), we get the following constraints equations:

$$
\begin{aligned}
x_{11}^t + x_{12}^t + x_{13}^t &= x_{11} + x_{12}^t + x_{31}^t & (6-1) \\
x_{31}^t + x_{32}^t + x_{33}^t &= x_{12}^t + x_{22}^t + x_{13}^t & (6-2) \\
x_{33}^t + x_{32}^t + 0 &= x_{13}^t + x_{33}^t + 0 & (6-3) \\
x_{ij}^t & \geq 0, \quad i = 1, \cdots, n-1, \ j = 1, \cdots, n-1 & (6-4)
\end{aligned}
$$
Define
\[ y'_y = \frac{x'_y}{x_{i1}} \]  

(7)

Here \( x'_{yi} \neq 0 \) \( (i = 1, \ldots, 3) \).

Perform transformation as function (7) to matrix \( X' \); then IO matrix \( X' \) has been converted into a new matrix \( Y' \) with no constraints:

\[
Y' = \begin{bmatrix}
1 & x'_{i2} & x'_{i1} \\
x'_{i1} & x'_{i2} & x'_{i3} \\
x'_{i1} & x'_{i3} & 0 \\
\end{bmatrix} = \begin{bmatrix}
1 & y'_{i2} & y'_{i3} \\
y'_{i2} & y'_{i3} & 0 \\
y'_{i3} & 0 & 0 \\
\end{bmatrix}
\]

(8)

Obviously \( y'_y \in (-\infty, +\infty) \), \( t = 1, \ldots, T \), and it is convenient for modeling to \( y'_y \).

After time series of matrix \( X' \) have been transformed into time series of matrix \( Y' \), each cell \( y'_y \) in matrix \( Y' \) will form a temporal sequence \( y'_y = (y'_{1y}, \ldots, y'_{Ty})' \). In order to forecast the values of all the cells in matrix \( Y \) at \( T+l \) time \( \hat{Y}^{T+l} \), we can refer the approach of predictive modeling of large-scale sequential curves [12].

Step 2. Return to IO matrix \( \hat{X}^{T+l} \) from forecasted \( \hat{Y}^{T+l} \)

It is very important that how to return to the original IO matrix \( \hat{X}^{T+l} \) after we have forecasted the future transformed matrix \( \hat{Y}^{T+l} \).

Define the predicted transformed matrix at \( T+l \) time as follows:

\[
\hat{Y}^{T+l} = \begin{bmatrix}
1 & \hat{y}^{T+l}_{i2} & \hat{y}^{T+l}_{i1} \\
\hat{y}^{T+l}_{i1} & \hat{y}^{T+l}_{i2} & \hat{y}^{T+l}_{i3} \\
\hat{y}^{T+l}_{i1} & \hat{y}^{T+l}_{i3} & 0 \\
\end{bmatrix} = \begin{bmatrix}
1 & \hat{x}^{T+l}_{i2} & \hat{x}^{T+l}_{i1} \\
\hat{x}^{T+l}_{i1} & \hat{x}^{T+l}_{i2} & \hat{x}^{T+l}_{i3} \\
\hat{x}^{T+l}_{i1} & \hat{x}^{T+l}_{i3} & 0 \\
\end{bmatrix}
\]

According to constraints (6), we get the following equations:

\[
\begin{cases}
\hat{x}^{T+l}_{i2} = \hat{y}^{T+l}_{i2} - \hat{y}^{T+l}_{i1} \\
\hat{x}^{T+l}_{i3} = \hat{y}^{T+l}_{i3} - \hat{y}^{T+l}_{i1} \\
\hat{x}^{T+l}_{i2} = \hat{y}^{T+l}_{i2} - \hat{y}^{T+l}_{i1} \\
\hat{x}^{T+l}_{i3} = \hat{y}^{T+l}_{i3} - \hat{y}^{T+l}_{i1} \\
\hat{x}^{T+l}_{i2} + \hat{x}^{T+l}_{i1} + \hat{x}^{T+l}_{i3} + \hat{x}^{T+l}_{i2} + \hat{x}^{T+l}_{i3} = \hat{x}^{T+l}_{i1} + \hat{x}^{T+l}_{i3} \\
\hat{x}^{T+l}_{i2} + \hat{x}^{T+l}_{i1} + \hat{x}^{T+l}_{i3} = \hat{x}^{T+l}_{i2} + \hat{x}^{T+l}_{i3} \\
\hat{x}^{T+l}_{i1} + \hat{x}^{T+l}_{i3} + \hat{x}^{T+l}_{i2} + \hat{x}^{T+l}_{i3} = \hat{x}^{T+l}_{i1} + \hat{x}^{T+l}_{i3} \\
\hat{x}^{T+l}_{i1} + \hat{x}^{T+l}_{i3} + \hat{x}^{T+l}_{i2} + \hat{x}^{T+l}_{i3} = \hat{x}^{T+l}_{i1} + \hat{x}^{T+l}_{i3} \\
\end{cases}
\]
Simplify equations (9) then obtain equations (10) with 3 known variables which are \( x_{11}^{T+l} \), \( x_{21}^{T+l} \) and \( x_{31}^{T+l} \) respectively.

\[
\begin{align*}
(1 + y_{12}^{T+l} + y_{13}^{T+l}) x_{11i}^{T+l} &= x_{11i}^{T+l} + x_{21i}^{T+l} + x_{31i}^{T+l} \quad (10-1) \\
(1 + y_{22}^{T+l} + y_{23}^{T+l}) x_{21i}^{T+l} &= y_{12}^{T+l} x_{11i}^{T+l} + y_{22}^{T+l} x_{21i}^{T+l} + y_{32}^{T+l} x_{31i}^{T+l} \quad (10-2) \\
(1 + y_{32}^{T+l}) x_{31i}^{T+l} &= y_{13}^{T+l} x_{11i}^{T+l} + y_{23}^{T+l} x_{21i}^{T+l} \quad (10-3)
\end{align*}
\]

It can be proved that there exists redundant equation in constraints (10), so that it is necessary to add an extra constraint to equation (10).

From IO matrix, we can see the economic implications that row \( n \) represents the added values of every sector. Denote 
\[
\sum_{j=1}^{n} x_{ij}^{t} = g^{t}
\]

Here, the economic meaning of \( g^{t} \) is the total of added values, which is GDP. Add a constraint if we regard GDP as an exogenous variable:

\[
x_{11}^{T+l} + x_{21}^{T+l} = g^{T+l}.
\] (11)

In equation (11), \( g^{T+l} \) is the predicted value of every sector at \( T+l \) time, which can be calculated out based on historical data \( g = (g^{1}, \ldots, g^{T})' \).

Combine equations (10) and (11) and get the following new constraint equations:

\[
\begin{align*}
(1 + y_{12}^{T+l} + y_{13}^{T+l}) x_{11i}^{T+l} &= x_{11i}^{T+l} + x_{21i}^{T+l} + x_{31i}^{T+l} \quad (12-1) \\
(1 + y_{22}^{T+l} + y_{23}^{T+l}) x_{21i}^{T+l} &= y_{12}^{T+l} x_{11i}^{T+l} + y_{22}^{T+l} x_{21i}^{T+l} + y_{32}^{T+l} x_{31i}^{T+l} \quad (12-2) \\
(1 + y_{32}^{T+l}) x_{31i}^{T+l} &= g^{T+l} \quad (12-3)
\end{align*}
\]

According to constraint (12-3), we get

\[
x_{31i}^{T+l} = \frac{g^{T+l}}{1 + y_{32}^{T+l}}.
\] (13)

Define \( x_{31} = d_{1} \), \( x_{32} = y_{12} x_{11} + y_{22} x_{21} + d_{1} = d_{2} \), then equations (12) has been transformed into equations (14).

\[
\begin{align*}
(1 + y_{12}^{T+l} + y_{13}^{T+l}) x_{11i}^{T+l} &= x_{11i}^{T+l} + y_{12}^{T+l} x_{11i}^{T+l} + d_{1} \quad (14-1) \\
(1 + y_{22}^{T+l} + y_{23}^{T+l}) x_{21i}^{T+l} &= y_{12}^{T+l} x_{11i}^{T+l} + y_{22}^{T+l} x_{21i}^{T+l} + d_{2} \quad (14-2)
\end{align*}
\]

Here, \( x_{11i}^{T+l} \) and \( x_{21i}^{T+l} \) are unknown while \( y_{12}^{T+l} \), \( y_{13}^{T+l} \), \( y_{22}^{T+l} \), \( y_{23}^{T+l} \) and \( d_{1} \), \( d_{2} \) are known.

Depict the equations (14) with matrix, we get

\[
\begin{pmatrix}
1 + y_{12}^{T+l} + y_{13}^{T+l} & 0 \\
0 & 1 + y_{22}^{T+l} + y_{23}^{T+l}
\end{pmatrix}
\begin{pmatrix}
x_{11i}^{T+l} \\
x_{21i}^{T+l}
\end{pmatrix}
\begin{pmatrix}
1 + y_{12}^{T+l} + y_{13}^{T+l} \\
1 + y_{22}^{T+l} + y_{23}^{T+l}
\end{pmatrix}
\begin{pmatrix}
x_{11i}^{T+l} \\
x_{21i}^{T+l}
\end{pmatrix}
= \begin{pmatrix} d_{1} \\ d_{2} \end{pmatrix}.
\] (15)
Here \( d_i = \frac{\hat{g}_{Ti}^{T1}}{1 + \hat{y}_{T2}^{T1}} \), \( d_i = \hat{y}_{T2}^{T1} \cdot d_i \).

It can be proved that equations (15) have a unique solution. Get the solution of (15), then combined them with equation (13) and (9), the values of other cells in the future IO matrix can be calculated out. Reorganize all the predicted cells, the IO matrix \( X_{Ti}^{T1} \) at \( T+1 \) time is obtained, that is exactly what we need.

Extend the approach of forecasting IO matrix from 3 dimensions to general situation. For \( n \)-dimensional IO matrix, obtain the following \((n-1)\) constraints equations:

\[
\begin{align*}
(1+\sum_{j=2}^{n} y_{T1j}^{T1}) \hat{x}_{T11}^{T1} &= \hat{x}_{T11}^{T1} + \hat{x}_{T21}^{T1} + \cdots + \hat{x}_{Tn1}^{T1} \\
(1+\sum_{j=2}^{n} y_{T2j}^{T1}) \hat{x}_{T21}^{T1} &= \hat{y}_{T12}^{T1} + \hat{y}_{T22}^{T1} + \cdots + \hat{y}_{Tn2}^{T1} \\
& \vdots \\
(1+\sum_{j=2}^{n} y_{T(n-1)j}^{T1}) \hat{x}_{T(n-1)1}^{T1} &= \hat{y}_{T1(n-1)}^{T1} + \hat{y}_{T2(n-1)}^{T1} + \cdots + \hat{y}_{Tn(n-1)}^{T1} \\
(1+\sum_{j=2}^{n} y_{Tn1}^{T1}) \hat{x}_{Tn1}^{T1} &= \hat{y}_{T1n}^{T1} + \hat{y}_{T2n}^{T1} + \cdots + \hat{y}_{Tnn}^{T1}
\end{align*}
\]

(16)

According to equation (16-n), we have

\[
\hat{x}_{T1i}^{T1} = \frac{\hat{g}_{Ti}^{T1}}{1 + \sum_{j=2}^{n} \hat{y}_{T1j}^{T1}} = d_i.
\]

Define

\[
d_j = \hat{y}_{T1j}^{T1} \cdot d_i, \quad j = 1, \ldots, n-1,
\]

Then denote equations (16) as follows:

\[
\begin{pmatrix}
1 + \sum_{j=2}^{n} \hat{y}_{T1j}^{T1} & 0 & \cdots & 0 \\
0 & 1 + \sum_{j=2}^{n} \hat{y}_{T2j}^{T1} & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & 1 + \sum_{j=2}^{n} \hat{y}_{T(n-1)j}^{T1}
\end{pmatrix}
\begin{pmatrix}
\hat{x}_{T11}^{T1} \\
\hat{x}_{T21}^{T1} \\
\vdots \\
\hat{x}_{T(n-1)1}^{T1} \\
\hat{x}_{Tn1}^{T1}
\end{pmatrix}
= \begin{pmatrix}
d_1 \\
d_2 \\
\vdots \\
d_{n-1}
\end{pmatrix}
\]

(17)

Solve the equations and get all the predicted values of cells in the future IO matrix \( \hat{x}_{T1i}^{T1} (i = 1, \ldots, n; j = 1, \ldots, n) \), then the future IO matrix \( \hat{X}_{Ti}^{T1} \) will be obtained after cells reconstruction.

4. Conclusions

This paper focuses on how to forecast the future IO table based on the historical data. At present similar researches mainly discuss the problem of technical coefficients adjustment, which is to give a revision of technical coefficients of the target year based on the base table in a certain economic assumptions, aiming to reduce the impact of time lag. These methods has made great progress and been widely applied, especially in compiling the adjusted IO table. However, these
methods have showed a certain degree of applicability when solving the forecasting problem. Because these methods require part of the future economic data are known in advance, while in fact everything of future is unknown in addition to historical data.

Accordingly, this paper invents a new approach for prediction. On the basic of keeping the internal balance of IO table, using existing sequential IO table, we abstract IO table into a matrix with constraints. Solve all the elements within a new IO table by algebraic method through a series of transform and inverse transform process, and then rearrange these forecasted elements to reconstitute the future IO table.

The proposed approach in this paper not only maintains the structural integrity and validity of constraints, but also does not require additional data known in advance (in addition to the historical sequence of IO tables). Moreover, the algebraic calculations can be realized easily via some common software such as Matlab. Therefore, this method has provided a workable program for forecasting IO table, and further offered analysis basis for forecasting the future economic activities.

References