Article

Time-Series Multivariate Analysis by Orbit Analysis and Principal Component Analysis Combined (1)[1]

ITAKI, Masahiko

Contents

Key words : GDP, Multiplier effect, Orbit analysis, Developmental principal component analysis, Multiple regression analysis of principal components, Orbit analysis of variables and principal components, Causality

Introduction

Either in natural science or social science, the fundamental purpose of research is to understand changes of an object over time on the basis of its structural analysis. For that purpose scientists devise diverse variables and observe their variations. However, those variables which are hypothetically chosen for the sake of measurement of variations rarely represent the actual substance of the object or forces that bring about its changes. We could expect much less to identify clear causality among those variables.

> It is the case for aggregates of GDP that are taken as an example in the present article: a bit surprising as it may sound to the common sense of Keynesian economics, those aggregates, i.e. fixed capital formation, household consumption, trade balance and government consumption, are all variables for observing and measuring a national economy, rather than its substance or driving forces *per se*. In order to understand this point, let us take an example of school examinations of five subjects, i.e. literature, mathematics, natural science, social science and English. Although they are five subjects for educating students or variables for measuring their academic achievements, intrinsic academic capabilities can be more properly explained by certain synthetic variables independent from each other, such as "general academic ability", "natural-science-biased ability" and "humanities-biased ability". This is the basic idea by which principal component analysis is conducted.

The purpose of this article is, as already suggested above, to examine the question of "What is the unit of change?" The *unit* can be understood as substance of a research object or a force that brings about changes, which we analyze on the basis of ordinary principal component analysis. It is followed by developmental principal component analysis in order to elucidate "multilayered structure of forces that bring about changes", in which orthogonal principal components are dialectically interpreted as "multilayered structure of unified opposites".

Conducting orbit analysis for leading-following relations (Itaki (2014)) among all variables and principal components, we construct a new theory of causality in which principal components are causes and variables are results. It is a replacement of so-called "Granger causality" that was rejected in principle (*ibid*).

In addition as a byproduct of using GDP for our example, a precise method of measuring the multiplier effect is proposed with the help of multiple regression analysis of principal components: i.e. an increment of GDP by means of one unit increase in fixed capital formation, household consumption, trade balance or government consumption.

I. Partial correction of orbit analysis

Orbit analysis will be partially corrected because some of the Microsoft EXCEL's functions in "V. Calculation of the direction of orbit rotation and leading-following relations", Itaki (2014), give either wrong results or inconsistent treatments when three coordinates necessary to determine the direction of orbit rotation are in the following three cases:

(1) Cases in which line *ab* is horizontal: if three coordinates are, for example, *a*(1, 0), *b*(2, 0) and *c*(3, 1), angle *bac* is correctly calculated to be positive, i.e. an anticlockwise rotation, and the result is correctly that variable *x* leads (X). However, if three coordinates are, for example, *a*(1, 0), *b*(2, 0) and *c*(3, -1), angle *bac* is correctly calculated to be negative, i.e. a clockwise rotation, but the result is wrongly that

variable *y* leads (YY); it should be that variable *x* leads (X).

(2) Cases in which line *abc* is horizontal: when angle *bac* is either 0, π or $-\pi$, the direction of rotation should be unidentified ("−"), but the results are wrongly that variable *x* leads for π and variable *y* leads for $-\pi$.

(3) Inconsistent treatments in various cases in which coordinates become vertical or horizontal are corrected: when three coordinates *a*, *b* and *c* become vertical, #DIV/0! comes out and when they become horizontal, #ERROR comes out, because their leading-following relations are unidentified (although when only two coordinates become vertical or horizontal, their leading-following relations can be identified. Refer to the illustrations below for details). In those cases, one ranking point should be divided between the two variables by 0.5 and 0.5.

Corrected functions are as follows in the same format as that in Itaki (2014). The shadows [2] are corrections [3]. Note that even after the corrections the global system of short-term interest rates in Itaki (2014) does not change.

Slope: (D3) =SLOPE(C2:C3,B2:B3), (D4) =SLOPE(C3:C4,B3:B4)

Rotation (in radians): (E3)

```
=IF(AND(-PI()<=IF(0<=ATAN2(B4-B2,C4-C2),ATAN2(B4-B2,C4-C2),2*PI()+ATAN2(B4-B2,C4-C2))-IF(0<=ATAN2(B3-B2,C3-C2),ATAN2(B3-B2,C3-C2),2*PI()+ATAN2(B3-B2,C3-C2)),IF(0<=ATAN2(B4-B2,C4-C2),ATAN2(B4-B2,C4-C2),2*PI()+ATAN2(B4-B2,C4-C2))-IF(0<=ATAN2(B3-B2,C3-C2),ATAN2(B3-B2,C3-C2),2*PI()+ATAN2(B3-B2,C3-C2))<=PI()),IF(0<=ATAN2(B4-B2,C4-C2),ATAN2(B4-B2,C4-C2),2*PI()+ATAN2(B4-B2,C4-C2))-IF(0<=ATAN2(B3-B2,C3-C2),ATAN2(B3-B2,C3-C2),2*PI()+ATAN2(B3-B2,C3-C2)),IF(PI()<IF(0<=ATAN2(B4-B2,C4-C2),ATAN2(B4-B2,C4-C2),2*PI()+ATAN2(B4-B2,C4-C2))-IF(0<=ATAN2(B3-B2,C3-C2),ATAN2(B3-B2,C3-C2),2*PI()+ATAN2(B3-B2,C3-C2)),IF(0<=ATAN2(B4-B2,C4-C2),ATAN2(B4-B2,C4-C2),2*PI()+ATAN2(B4-B2,C4-C2))-IF(0<=ATAN2(B3-B2,C3-C2),ATAN2(B3-B2,C3-C2),2*PI()+ATAN2(B3-B2,C3-C2))-2*PI(),2*PI()+IF(0<=ATAN2(B4-B2,C4-C2),ATAN2(B4-B2,C4-C2),2*PI()+ATAN2(B4-B2,C4-C2))-IF(0<=ATAN2(B3-B2,C3-C2),ATAN2(B3-B2,C3-C2),2*PI()+ATAN2(B3-B2,C3-C2))))
```

Rotation (in degrees): (F3) =DEGREES(E3)

Leading-following: (G3)

```
=IF(AND(ISERROR(D3),F3<0,F3<>-180),"YY",IF(AND(ISERROR(D3),0<F3,180<>F3),"-YY",IF(AND(0<=D3,0<F3,180<>F3),"X",IF(AND(0<D3,F3<0,F3<>-180),"YY",IF(AND(D3<0,0<F3,180<>F3),"-YY",IF(AND(D3<=0,F3<0,F3<>-180),"-X","-"))))))
```

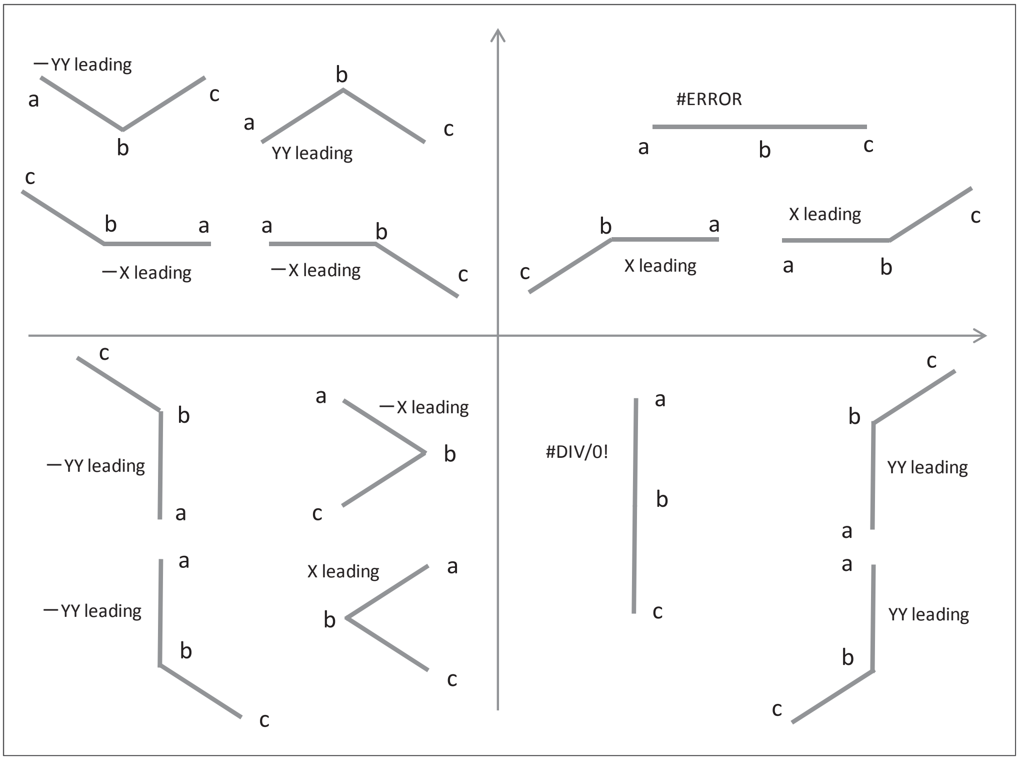
Dissolved functions for rotation:

H3 =B3-B2
H4 =B4-B2
I3 =C3-C2
I4 =C4-C2
J3 =ATAN2(H3,I3)
J4 =ATAN2(H4,I4)

K3 =IF(0<=J3,J3,2*PI()+J3)
K4 =IF(0<=J4,J4,2*PI()+J4)
L3=IF(AND(-PI()<=K4-K3,K4-K3<=PI()),K4-K3,IF(PI()<K4-K3,K4-K3-2*PI(),2*PI()+K4-K3))

Table 1: Calculation of the direction of orbit rotation and leading-following relations in EXCEL

	A	B	C	D	E	F	G	H	I	J	K	L
1		Variable x	Variable y	Slope	Rotation (in radians)	Rotation (in degrees)	Leading or following	Dissolved functions for rotation				Rotation (in radians)
2	1989	18.63	9.21									
3	1990	20.78	5.75	−1.61	−1.28	−73.45	−X	2.15	−3.46	−1.01	5.27	−1.28
4	1991	12.39	2.18	0.43	−0.81	−46.45	YY	−6.24	−7.03	−2.30	3.99	
5	1992	16.04	7.80	1.54	1.59	90.82	X					
6	1993	4.42	7.19	0.05	−1.90	−108.97	YY					
7	1994	17.23	11.96	0.37	1.96	112.46	X					
8	1995	−1.00	13.02	−0.06	2.78	159.18	-YY					
9	1996	28.65	6.84	−0.21	0.04	2.16	-YY					
10	1997	17.53	9.88	−0.27	0.47	26.98	-YY					
11	1998	6.09	2.18	0.67	−0.45	−25.97	YY					
12	1999	11.83	9.08	1.20	−0.13	−7.41	YY					
13	2000	21.34	16.32	0.76								
14												
15	Notes:	1. Positive radians or degrees signify an anticlockwise rotation, and negative ones signify a clockwise rotation.										
16		2. *X* signifies that variable *x* leads, and *YY* signifies that variable *y* leads, while negative means that slope is negative.										



II. Some features of principal component analysis

It would be almost needless to say that we always witness in our research extremely complicated phenomena unfolding themselves in front of us and seeking explanations. So we devise certain variables (or indicators), with which we attempt to describe and understand the phenomena in terms of relations between the variables (or indicators). Here we often encounter a troublesome dilemma: if the number of variables is small, our understanding of the phenomena remains unsatisfactory; to the contrary, a large number of variables would hinder their clear-cut understanding. One possible solution to the dilemma is to limit the number of variables to an "appropriate" one from a certain viewpoint. However, if we can sum up information that a large number of variables offer and turn them into a concise number of representative synthetic variables (or comprehensive indicators), our understanding of the phenomena could improve to a great extent. The principal component analysis has developed on this basic idea. H. Hotelling, one of the original developers of principal component analysis, talked about its significance as follows (Hotelling (1933), p.1):

> "Consider n variables attaching to each individual of a population. These statistical variables $x_1, x_2, \ldots, x_n$, might for example be scores made by school children in tests of speed and skill in solving arithmetical problems or in reading; or they might be various physical properties of telephone poles, or the rates of exchange among various currencies. The x's will ordinarily be correlated. It is natural to ask whether some more fundamental set of independent variables exists, perhaps fewer in number than the x's, which determine the values the x's will take." (Hotelling (1933) p.1.)

Hotelling seems to distinguish two types, i.e. a set of more fundamental and independent variables and a set of fewer variables in number. The distinction between them will play an important role later when we proceed from ordinary principal component analysis to developmental principal component analysis. I. T. Jolliffe, who has been fascinated by principal component analysis for over 30 years, mentioned with this respect as follows:

> "The central idea of principal component analysis (PCA) is to reduce the dimensionality of a data set consisting of a large number of interrelated variables, while retaining as much as possible of the variation present in the data set. This is achieved by transforming to a new set of variables, the principal components (PCs), which are uncorrelated, and which are ordered so that the first *few* retain most of the variation present in *all* of the original variables." (Jolliffe (2002) p.1.) [4)]

N principal components that are obtained from n variables have the following properties [5)]:

(1) N principal components that are obtained by linearly combining original n variables are uncorrelated to each other: i.e. principal components are orthogonal to each other in n-dimensional space.

(2) The following inequality relations hold among the principal components: the variance of the first principal component > the variance of the second principal component > ⋯ > the variance of the nth principal component. The total variance of all the principal components is equal to that of the original n variables. The ratio of a variance of each principal component to the total variance is called a proportion. A few top principal

components can often explain much variance of all variables.

Firstly, an example of principal component analysis on non-time-series data is examined in order to understand some of its important features: i.e. examinations of 50 students in literature, mathematics, natural science, social science and English. We start our analysis with 50-students' 5-subject scores and their total, the objective of which is to measure their academic ability.

Those 50 students are ranked in order of their total scores. A simple observation with respect to total scores would reveal a variety of patterns of academic ability. Some students are very good at literature and some other students are very bad at natural science. There are some very talented students who gain high scores in all subjects, but others are less talented and get miserable scores in all subjects. These facts may suggest that the score of a specific subject, such as English or mathematics, does not correctly predict variations of total scores, and that simple total scores are not a perfect indicator of general academic ability of students. Generally speaking, however, top students are likely to acquire high scores in any subjects; in the middle, there are some characteristic students who are extremely good, for example, at English or incredibly bad, for example, at mathematics; and students in the bottom are likely to gain low scores in many subjects.

Our experiences would suggest a hypothesis that academic ability could be measured by a certain combination of subjects instead of independently by a single subject: for instance, "general academic ability" may exist which exerts positive effects on all subjects. The more "general academic ability" increases, the higher scores of mathematics and English as well as literature and others can be expected. Alternatively, "natural-science-biased ability" may exist: it has a positive correlation with scores of natural science and mathematics, but a negative correlation with English and literature. By contrast, we may think of "humanities-biased ability", which might be a reversed one to "natural-science-biased ability" and thus, we should regard them as heads and tails of the same coin, i.e. "natural-science- or humanities-biased ability" altogether. On top of that, there may be a specific ability that has a high positive correlation only with mathematics, but has nothing to do with all the other subjects, which could be named "mathematics-biased ability".

A crucial point is that those "general academic ability", "natural-science- or humanities-biased ability" and "mathematics-biased ability" are independent from each other and have no relation with each other. Ability that is only and specifically biased to natural science or humanities is different in dimension from comprehensive ability with no bias to any specific subject. That is also true to ability that is concentrated only on mathematics. It would be reasonable to assume that those components of academic ability with different dimensions scatter over 50 students and each student has them all in various proportions. By comparison, 5 subjects, i.e. literature, mathematics, natural science, social science and English, are independently carried out in examinations, whose scores more or less synchronize together thanks to overlapping effects of those components of academic ability. Therefore, those 5 different subjects are *categories of items* for measuring academic ability, rather than *categories of substance* that form academic ability.

> The relations, in our example, between subjects and components of academic ability are equivalent to those between variables and principal components. By means of principal component analysis, complicated variations in scores of all subjects are resolved into a set of independent components of academic ability, and in turn, by means of multiple regression analysis of principal components, scores of each subject and their totals are reconstructed by those academic components.

Next, imagine the stage of a kind of psychodrama, in which tense emotions are exchanged among 10 actors and actresses, in order to understand some features of principal component analysis on time-series data:

> Six out of ten of them belong to the largest group of friends; remaining four form another group of friends. Those two groups are in conflict, but two in the minor group keep some relations with members of the major group. The reason for the conflict is that the bosses of the two groups are rivals to one another. There are some factions in the major group that exercise delicate maneuvers. A woman in the major group and a man in the minor group are lovers, just like Juliet in the House of Capulet and Romeo in the House of Montague. We are already informed of those complicated human relations among them, but the audience is not. And, furthermore, the play is a pantomime and the audience has to understand those entangled relations only with the help of men's and women's actions on the stage.
>
> Now, the pantomime begins. Members of the major group start to gather in the center of the stage. But, their movements are not straight: its factions keep some tactful distances among them while assembling. Those two members in the minor group who keep somewhat friendly relations with the major group exhibit rather complicated movements: although they are mainly attracted by the minor group that gathers at an edge of the stage, they also make eyes at the major group and a certain faction as well. They are intermingled in those triple human relations. Romeo and Juliet may show the most complicated maneuvers of all. Juliet shifts towards the center of the stage, where her own major group gathers, and Romeo towards the edge, both with reluctant steps. They actually would like to spend a good time together without being bothered by others at the opposite side of the stage. They gradually come close together, but quickly separate and temporarily keep some awkward distance when someone of the opposite group approaches. Those triple human relations are reflected on their highly complex movements: i.e. the conflict between the two groups, minor conflicts among factions and love between themselves.

The reason for difficulties to comprehend the plot of the play is that one actor or actress embodies not only one human relation, but also several intertwined relations, which altogether determine each movement on the stage of an actor or actress. What would happen then, if those multilayered human relations are respectively separated from individuals, brought together and classified by human relation? Attraction of the major group, minor group, factions and lovers, and repulsion between the major and minor groups, among factions – those are all perfectly independent forces from each other. Human relations with respect to the major group probably work as the strongest force of all, and those with respect to the minor group and others would be ranked according to the strength of their attraction. Those invisible, mutually independent human relations, after being separated, could be expressed concretely as movements on the stage: the attraction of the major group would go straight to the center of the stage, that of the minor group to the edge of the stage and that of the two lovers to the opposite side of the stage, which are all independent from and not interfering with others. Those ten actors and actresses are *categories that demonstrate bearers* of human relations changing over time, rather than *their real substances*.

As exemplified by the pantomime above, the actors and actresses who combine and bear a variety of human relations are equivalent to variables, and human relations that are respectively abstracted and put together are equivalent to principal components. We later show that variables' complicated variations over time are resolved into a set of independent principal components and, by means of multiple regression analysis of principal components, the original variables are reconstructed by the principal components.

Now we proceed to explore a frontier of time-series multivariate analysis with an example of Japan's GDP, in which firstly some important features of principal component analysis are concretely demonstrated, secondly it is further developed into developmental principal component analysis and lastly it is combined with orbit analysis.

III. Analysis of Japan's GDP

The outline of our argument is as follows: (1) Basic analysis and (2) Orbit analysis deal with, as it were, "the visible world" of observers, in which we observers would be able to directly confirm, only if paying appropriate attentions to data, such statistical phenomena as value of change, rate of change and direction of orbit rotation. By contrast, (3) Principal component analysis and (4) Developmental principal component analysis deal with "the invisible world", in which principal components abstracted from data could not be detected at all with our bare eyes. It is a world of concept that we can reach only as a result of certain statistical procedures. Lastly in (5) Principal component analysis and orbit analysis combined, those "visible world" and "invisible world" will be combined together, in which we understand how the world observers see with their bare eyes is constructed and operated by invisible forces behind the scene

(1) Basic analysis

Table 2 "Japan's GDP" reveals historical development of Japan's GDP from 1956 to 2012: i.e. annual changes in value at current prices and annual growth rates, instead of nominal or real absolute value of GDP. The reason is that the analytical purpose here is to capture the driving forces of change in GDP, rather than its long-term trend.

Its graph is Fig. 1 "Japan's GDP (annual change in value at current prices)", which illustrates main four aggregates of GDP, i.e. household consumption, government consumption, trade balance and fixed capital formation. The reason why we adopt nominal values that are influenced by changes in prices due to inflation or deflation is that we later perform orbit analysis on those data: annual nominal values are

Talbe 2: Japan's GDP

	Annual change in value (unit: 1 million yen at current prices)					Annual growth rates (unit: %)				
	GDP	Household consumption	Government consumption	Trade balance	Fixed capital formation	GDP	Household consumption	Government consumption	Trade balance	Fixed capital formation
1956	1,052,700	559,400	28,600	−88,500	524,500	12.6	10.2	3.4	−197.1	32.3
1957	1,436,100	727,100	70,000	−182,900	647,500	15.2	12.0	8.0	419.5	30.1
1958	680,000	488,700	78,000	387,700	70,400	6.3	7.2	8.3	−171.2	2.5
1959	1,652,000	873,700	94,800	−18,000	507,900	14.3	12.0	9.3	−11.2	17.7
1960	2,819,400	1,244,600	164,800	−70,900	1,261,700	21.4	15.3	14.8	−49.5	37.4
1961	3,326,800	1,635,500	202,500	−381,000	1,529,100	20.8	17.4	15.8	−527.0	33.0
1962	2,606,200	1,622,500	262,500	344,000	898,100	13.5	14.7	17.7	−111.4	14.6
1963	3,170,500	2,119,000	323,100	−240,800	864,200	14.4	16.7	18.5	−682.2	12.2
1964	4,428,100	2,255,700	281,900	153,700	1,433,000	17.6	15.3	13.6	−74.8	18.1
1965	3,324,700	2,211,100	338,300	511,400	420,300	11.3	13.0	14.4	−987.3	4.5
1966	5,304,000	2,902,900	364,300	137,200	1,779,300	16.1	15.1	13.5	29.9	18.2
1967	6,560,500	3,263,000	355,900	−496,700	2,725,900	17.2	14.7	11.7	−83.2	23.6
1968	8,244,400	3,568,500	524,000	490,100	3,279,500	18.4	14.0	15.4	489.6	23.0
1969	9,254,000	4,326,200	624,200	401,300	3,873,600	17.5	14.9	15.9	68.0	22.1
1970	11,116,000	5,032,700	896,900	−50,700	4,602,600	17.9	15.1	19.7	−5.1	21.5
1971	7,356,400	4,897,500	966,200	1,257,200	1,594,000	10.0	12.8	17.7	133.6	6.1
1972	11,693,100	6,670,900	1,115,300	−63,700	3,886,500	14.5	15.4	17.4	−2.9	14.1
1973	20,103,700	10,406,900	1,799,600	−2,104,100	9,414,700	21.8	20.9	23.9	−98.6	29.9
1974	21,745,700	12,604,200	2,903,900	−1,029,500	5,756,400	19.3	20.9	31.1	−3,408.9	14.1
1975	14,083,300	11,850,700	2,649,900	1,061,700	1,440,900	10.5	16.3	21.6	−106.2	3.1
1976	18,246,200	11,021,100	1,527,000	1,272,700	3,809,400	12.3	13.0	10.3	2,039.6	7.9
1977	19,048,700	11,292,400	1,826,000	1,705,200	4,036,800	11.4	11.8	11.1	127.7	7.8
1978	18,782,100	10,846,900	1,509,300	514,300	6,164,800	10.1	10.1	8.3	16.9	11.0
1979	17,142,500	12,154,800	1,733,700	−5,555,900	8,024,100	8.4	10.3	8.8	−156.3	12.9
1980	21,292,100	2,168,900	12,817,300	−217,800	6,890,700	9.6	1.7	59.7	10.9	9.8
1981	18,229,500	8,451,300	2,572,900	4,138,400	3,179,000	7.5	6.4	7.5	−186.5	4.1
1982	13,018,400	10,312,600	2,281,600	−70,000	703,300	5.0	7.3	6.2	−3.6	0.9
1983	10,971,700	7,980,700	2,343,400	3,017,900	−1,392,000	4.0	5.3	6.0	163.2	−1.7
1984	17,916,600	8,163,300	1,953,500	3,168,900	3,782,500	6.3	5.1	4.7	65.1	4.8
1985	22,427,000	9,552,000	1,916,500	3,003,300	6,833,300	7.4	5.7	4.4	37.4	8.2
1986	15,157,600	7,321,500	2,096,100	2,241,500	4,055,200	4.7	4.1	4.6	20.3	4.5
1987	13,610,700	8,576,000	2,038,400	−2,719,800	6,823,500	4.0	4.7	4.3	−20.5	7.2
1988	26,572,700	10,762,800	2,201,500	−2,321,900	13,702,200	7.5	5.6	4.4	−22.0	13.6
1989	29,379,300	14,162,300	3,338,400	−2,002,100	13,624,800	7.7	7.0	6.5	−24.3	11.9
1990	32,658,800	17,174,600	3,979,200	−2,063,900	13,954,700	8.0	7.9	7.2	−33.1	10.9
1991	26,640,800	11,795,400	4,049,300	3,373,800	6,778,100	6.0	5.0	6.9	80.8	4.8
1992	11,361,000	9,476,600	3,475,500	2,850,600	−2,257,300	2.4	3.8	5.5	37.8	−1.5
1993	2,929,000	5,280,000	2,800,200	367,900	−4,806,200	0.6	2.1	4.2	3.5	−3.3
1994	12,031,600	12,738,900	3,492,000	−903,900	−2,085,700	2.5	4.9	5.0	−8.4	−1.5
1995	5,963,500	3,749,300	3,351,300	−3,061,200	−728,900	1.2	1.4	4.6	−31.0	−0.5
1996	10,227,900	6,326,800	2,820,000	−4,437,400	4,842,400	2.0	2.3	3.7	−65.3	3.5
1997	11,263,500	5,910,200	1,631,700	3,263,000	159,200	2.2	2.1	2.1	138.1	0.1
1998	−10,759,700	−2,436,100	938,200	3,936,400	−11,946,100	−2.1	−0.8	1.2	70.0	−8.3
1999	−7,535,400	1,332,100	1,543,600	−1,532,600	−3,603,700	−1.5	0.5	1.9	−16.0	−2.7
2000	4,956,800	−709,900	3,181,000	−643,400	−164,400	1.0	−0.2	3.8	−8.0	−0.1
2001	−4,316,800	1,620,700	3,346,800	−4,152,200	−5,679,400	−0.8	0.6	3.9	−56.2	−4.4
2002	−6,396,200	−749,600	1,651,500	3,463,100	−8,624,500	−1.3	−0.3	1.8	107.1	−7.0
2003	−292,200	−1,524,100	37,300	1,559,800	−1,987,700	−0.1	−0.5	0.0	23.3	−1.7
2004	4,870,500	1,085,100	565,800	1,602,300	−436,900	1.0	0.4	0.6	19.4	−0.4
2005	177,700	2,533,300	558,900	−2,765,800	787,000	0.0	0.9	0.6	−28.1	0.7
2006	2,784,000	2,300,700	−501,900	−725,900	2,322,100	0.6	0.8	−0.5	−10.2	2.1
2007	6,288,200	688,700	826,700	2,305,500	885,200	1.2	0.2	0.9	36.2	0.8
2008	−11,765,900	−2,066,600	226,500	−7,700,800	−3,319,000	−2.3	−0.7	0.2	−88.8	−2.9
2009	−30,070,600	−9,113,700	800,200	754,300	−14,471,700	−6.0	−3.1	0.9	77.6	−12.9
2010	11,245,700	2,925,400	1,309,000	4,036,600	−1,559,500	2.4	1.0	1.4	233.8	−1.6
2011	−11,761,200	−1,082,800	1,074,700	−10,046,700	441,200	−2.4	−0.4	1.1	−174.3	0.5
2012	5,244,700	5,000,800	1,438,300	−5,215,700	3,948,500	1.1	1.8	1.5	121.8	4.1

Data : IMF, *International Financial Statistics.*

indispensable for examining actual pulling and being-pulled relations in quantity among those aggregates. Household consumption, the center pole of an economy, indicates that the Japanese economy could be divided into three periods: one until the early 1970s of stable growth, another in the 1970s and 1980s of rapid expansion and the last one since 1990 of shrinkage owing to the burst of an economic babble.

Figures 2, 3 and 4 "Japan's GDP (annual growth rate)" illustrate annual growth rates of those aggregates. Comparisons between household consumption and government consumption, between GDP and fixed capital formation and between GDP and trade balance make clear their particular features: household consumption and government consumption move closely along with GDP; the movement of fixed capital formation amplifies that of GDP; and trade balance, by and large, negatively correlates to GDP.

Observation of changes in value and rates of change allows us to get much insight into

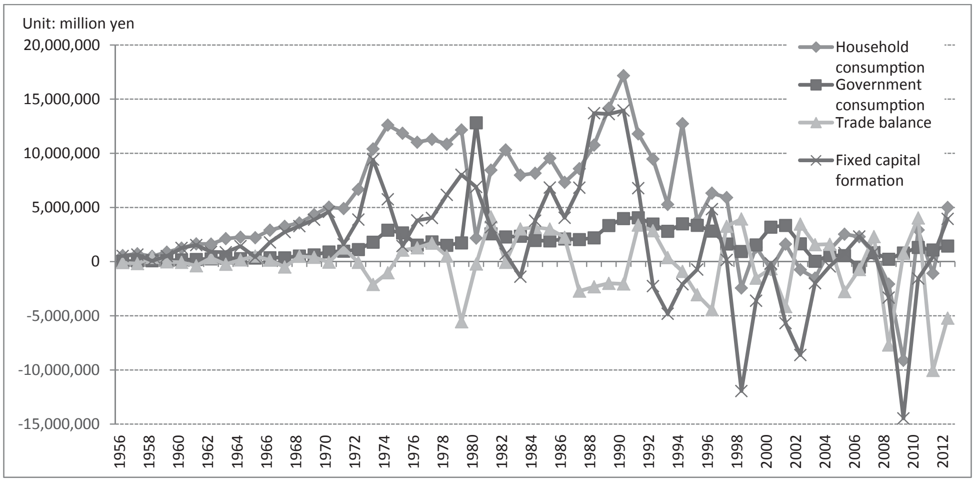


Fig. 1: Japan's GDP (annual change in value at current prices)

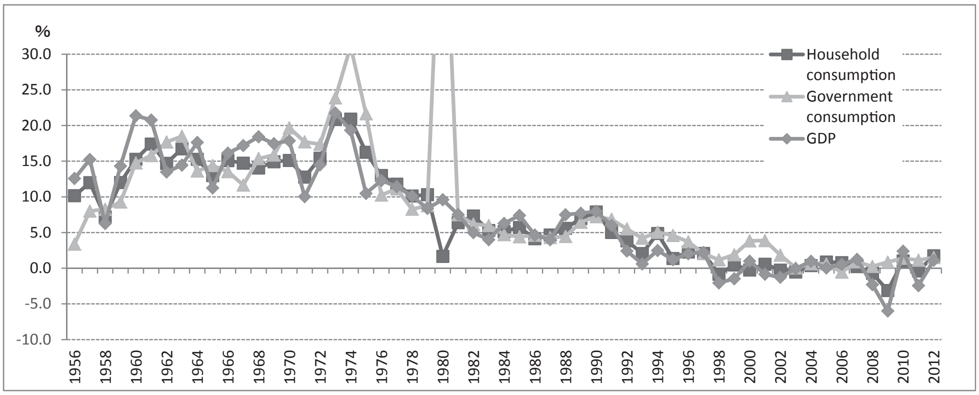


Fig. 2: Japan's GDP (annual growth rate)

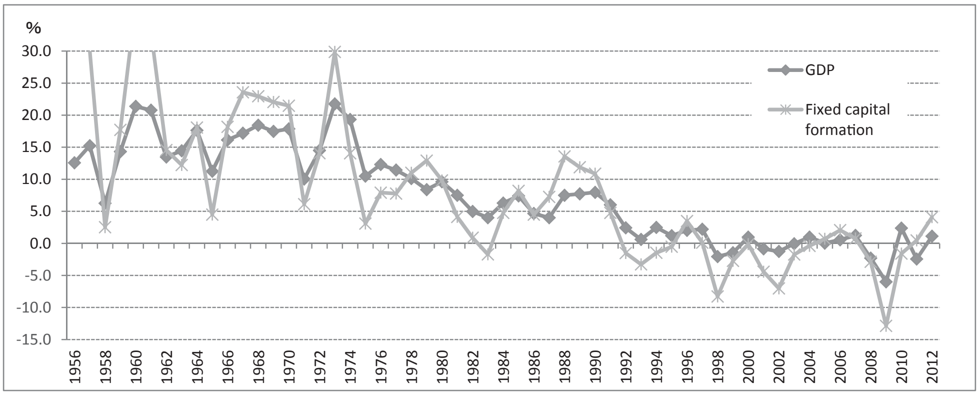


Fig. 3: Japan's GDP (annual growth rate)

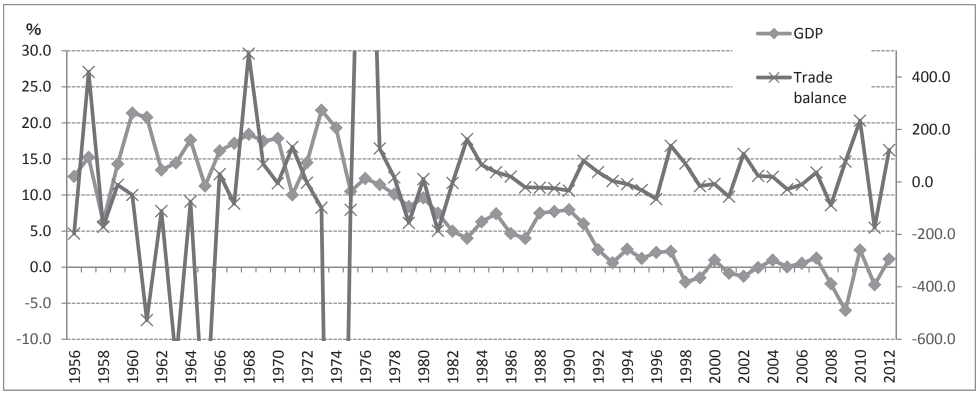


Fig. 4: Japan's GDP (annual growth rate)

Note: Trade balance is on the right axis.

variations of GDP. There are, however, two main obstacles or new problems in such basic analysis:

Firstly, those four aggregates are neither independent from each other nor move autonomously; they are rather causes and effects of each other and thus, move in a bunch, being entangled with each other. Our new task would be, therefore, to disentangle the unity and synchronism among the aggregates, and to identify leading and following variables.

Secondly, the synchronization implies, for example, that household consumption is not independent and is under the influence of government consumption, trade balance and fixed capital formation altogether. Even when you observe and describe specific variations of household consumption, you actually observe synthetic effects of those four aggregates through the window of household consumption. This holds true for government consumption, trade balance and fixed capital formation as well and thus, even if you conduct four individual observations, you actually observe the same phenomenon four times in succession. This fact suggests that those four aggregates are *four items* necessary for measuring GDP, but they are not *four categories* necessary for explaining variations of GDP. We have to identify more essential and autonomous substance, forces or motive power that is hidden under those four synchronizing items.

It should be emphasized, however, that basic statistical information, such as annual changes in value and annual growth rates, is the goal as well as the starting point of our analysis. We will come back to the basic information after overcoming those two limitations and solving the problems; then, it will be accompanied by rich additional information and shed a new explanatory light onto the structure and movement of Japan's GDP.

(2) Orbit analysis

Orbit analysis is a statistical method that extracts leading-following relations between two variables by combining coordinates of time-series data along time and calculating the slope and the direction of rotation of the "orbit" thus depicted. In the case of many variables, calculations of all combinations between pair variables produce a hierarchy of leading-following relations among all variables (Itaki(2014)), which identifies the kick-starter variable that heralds all variations among other variables and determines their order of following. The method is applied to the four aggregates of Japan's GDP in the same period in Fig. 5 "9-year moving average of ranking points for Japan's GDP (expenditure)". Due to wide annual variance of ranking points, 9-year moving average is adopted out of 5-, 7- and 9-year calculations, which seems to represent most appropriately medium-term variations.

> Annual changes in value are used for the calculations, because there have to be quantitative pulling- and being-pulled-relations among variables for them to be actually leading-following relations (Itaki (2014) p.16, pp.30-31.). Household consumption, government consumption, trade balance and fixed capital formation are apparently in quantitative pulling- and being-pulled-relations in their annual changes in value. That is not *a priori* true in their annual growth rates: although an increase in household consumption by 10 billion yen, for example, directly causes quantitative deterioration in trade balance or an increase in fixed capital formation by 1 billion yen or so, its increase by 5 %, for example, does not necessarily cause an increase or decrease in trade balance or fixed capital formation by a certain percentage point.

Fig. 5 shows that the kick-starter (i.e. the first leading variable) was fixed capital formation in 1958-67, household consumption in 1972-80, government consumption in the period of fiscal reconstruction 1983-85, fixed capital formation again in the period of an economic bubble and its burst 1988-1992 and trade balance in the period of the great depression 1996-2009. Another feature is that fixed capital formation was in the lowest rank in 1977-84, the period between the aftermath of the first worldwide recession since

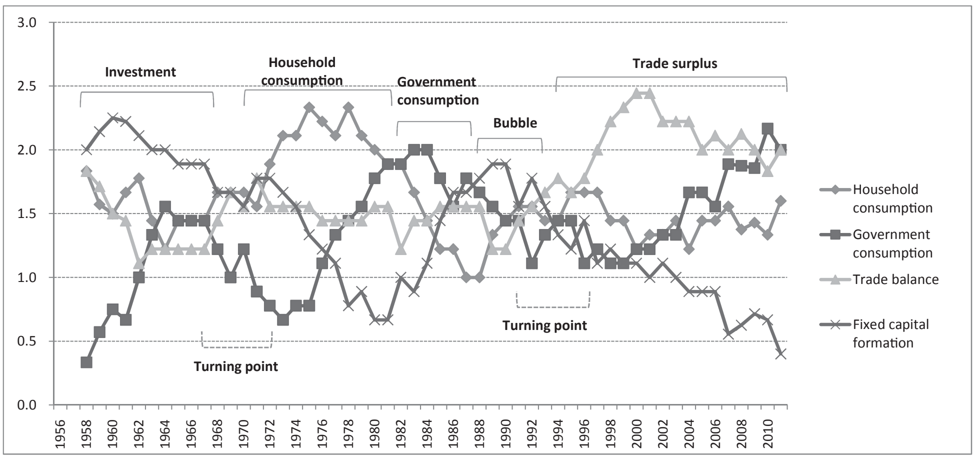


Fig. 5: 9-year moving average of ranking points for Japan's GDP (expenditure)

the end of Second World War, 1974-75, and the previous year of the Plaza Accord in 1985; and that happened again since 2000. Investment in plants and equipments does not now play the role of generating changes in GDP. On top of that, two major structural transformations of Japanese economy took place, first in the end of the 1960s and early 1970s and second in the first half of the 1990s in the aftermath of the burst of bubble, when ranks were intertwined in a bunch and changed quickly. We now know that information acquired by orbit analysis by and large corresponds to our widely shared knowledge about the post-World War history of Japanese economy, and also that it adds many other insights to our understanding.

> We ought to pay enough attention to the difference between the two concepts, "leading and following" and "preceding and lagging" [6]. The unit of our observation and analysis is one period of time, the beginning and end of which we can observe and compare, but we do not know actions and reactions that may take place among variables during the minimum one period of time. Therefore, we cannot know, in principle, which variable precedes or lags, or whether reversals in ranks repeat themselves, during the unit period, just like in a black box. What we can actually observe and measure is distinction between a variable that actively leads changes and a variable that follows the changes initiated. Those changes are produced by complicated actions and reactions (positive and negative feedbacks) among variables. The distinction is observed and recorded as if being temporal preceding-lagging relations. We should be careful enough, therefore, that leading-following relations in Fig. 5 among four aggregates are not temporal preceding-lagging relations. Their distinction reveals its importance later when orbit analysis is conducted among the principal components of those four aggregates.

An application of orbit analysis to annual variations in value in Fig. 1 exposes hidden leading-following relations among the aggregates in Fig. 5. It is certainly a big step forwards to the understanding of structure and movements of the object, which, however, touches only its surface; we next proceed to principal component analysis and further to developmental principal component analysis that allow us to investigate more essential and autonomous substance, forces or motive power of the object that is hidden under the surface of those leading-following aggregates.

(3) Principal component analysis

1. Terms

The results of principal component analysis on the four aggregates of Japan's GDP are shown in Table 3-1 "Eigenvectors and loadings (Japan's GDP)" and Table 4 "Principal component scores of Japan's GDP". Explanation of terms in these tables and some features of principal component analysis are given as follows:

> **Principal component analysis on the basis of variance and of correlation:** there are two types of principal component analysis, one that is based on original data, i.e. their variance, and another on their correlation. The latter is conducted on standardized scores (i.e. Z scores) whose standard deviation is one. Here we adopt the former because the four aggregates share the same monetary unit, i.e. the Japanese

Table 3-1: Eigenvectors and loadings (Japan's GDP)

	First PC	Second PC	Third PC	Fourth PC	
Eigenvalue	4.5806E+13	1.0081E+13	5.1814E+12	3.0719E+12	
Proportion	0.714	0.157	0.081	0.048	
Eigenvector					
Household consumption	**0.692**	0.508	**0.476**	−0.193	
Government consumption	0.097	0.079	0.172	**0.977**	
Trade balance	−0.042	**0.725**	−0.685	0.067	
Fixed capital formation	**0.714**	−0.459	−0.525	0.059	
Loading					
Household consumption	4,683,135	1,611,578	1,082,588	−338,804	
Government consumption	655,022	249,658	392,460	1,712,535	
Trade balance	−287,146	2,300,797	−1,558,252	116,867	
Fixed capital formation	4,833,451	−1,458,606	−1,194,678	103,131	
Loading in absolute value					Total
Household consumption	4,683,135	1,611,578	1,082,588	338,804	7,716,105
Government consumption	655,022	249,658	392,460	1,712,535	3,009,676
Trade balance	287,146	2,300,797	1,558,252	116,867	4,263,061
Fixed capital formation	4,833,451	1,458,606	1,194,678	103,131	7,589,866
Loading as % of total absolute value					Total
Household consumption	60.69	20.89	14.03	4.39	100.00
Government consumption	21.76	8.30	13.04	56.90	100.00
Trade balance	6.74	53.97	36.55	2.74	100.00
Fixed capital formation	63.68	19.22	15.74	1.36	100.00

Table 3-2: Square of loadings (Japan's GDP)

Square of loading	First PC	Second PC	Third PC	Fourth PC	Total
Household consumption	21,931,753,661,120	2,597,182,065,076	1,171,995,842,755	114,788,423,779	25,815,719,992,730
Government consumption	429,053,682,023	62,329,237,393	154,025,004,713	2,932,777,346,053	3,578,185,270,182
Trade balance	82,452,785,285	5,293,665,008,199	2,428,148,117,134	13,657,966,489	7,817,923,877,107
Fixed capital formation	23,362,250,192,644	2,127,532,328,119	1,427,256,322,341	10,635,914,964	26,927,674,758,067
Total	45,805,510,321,072	10,080,708,638,786	5,181,425,286,943	3,071,859,651,285	64,139,503,898,086

yen, and we would like to conduct orbit analysis and principal component analysis on the basis of monetary value. The latter is usually used in the case in which variables have different units of measurement or in which, despite sharing the same unit, difference in value does not make sense for one reason or another.

Principal components: four principal components are extracted from four variables. With the first principal component being Z_1, household consumption X_1, the government consumption X_2, trade balance X_3 and fixed capital formation X_4, we get $Z_1 = 0.692\ X_1 + 0.097\ X_2 - 0.042\ X_3 + 0.714\ X_4$ from the eigenvector in the table: the first principal component is composed as a synthetic function of the original four variables. So are the other principal components.

Eigenvector: eigenvectors that consist of the coefficients of the principal components have a property that the sum of squares of all elements in each row and each column is one: for example, regarding household consumption $0.692^2 + 0.508^2 + 0.476^2 + (-0.193)^2 = 1$, and regarding the first principal component $0.692^2 + 0.097^2 + (-0.042)^2 + 0.714^2 = 1$. Values and signs of elements give us an important clue when we attempt to interpret the meanings of principal components.

Loading: the sum of squares of all loadings of the first principal component, i.e. household consumption 4,683,135, government consumption 65,502, trade balance

Table 4: Principal component scores of Japan's GDP (unit: 1 million yen)

	Principal component scores (mean=0)				Principal component scores (mean≠0)			
	First PC	Second PC	Third PC	Fourth PC	First PC	Second PC	Third PC	Fourth PC
1956	−4,097,185	−1,763,137	−1,809,766	−841,648	768,182	−18,898	56,285	−55,230
1957	−3,885,290	−1,799,675	−1,722,804	−832,671	980,077	−55,436	143,248	−46,252
1958	−4,485,831	−1,241,443	−1,922,533	−774,680	379,536	502,796	−56,481	11,739
1959	−3,888,143	−1,539,685	−1,688,422	−833,996	977,224	204,554	177,629	−47,578
1960	−3,084,140	−1,730,550	−1,859,365	−796,469	1,781,227	13,689	6,687	−10,051
1961	−2,605,882	−1,876,732	−1,595,014	−840,140	2,259,485	−132,493	271,038	−53,721
1962	−3,090,468	−1,063,354	−1,755,986	−767,787	1,774,899	680,885	110,066	18,631
1963	−2,740,446	−1,214,780	−1,091,280	−845,541	2,124,921	529,459	774,772	−59,123
1964	−2,260,364	−1,124,064	−1,601,958	−852,448	2,605,003	620,175	264,094	−66,030
1965	−3,024,178	−417,821	−1,326,808	−824,456	1,841,189	1,326,418	539,243	−38,038
1966	−1,556,541	−960,126	−1,450,402	−877,767	3,308,826	784,113	415,650	−91,349
1967	−605,258	−1,672,236	−1,343,457	−942,153	4,260,109	72,003	522,595	−155,735
1968	−24,102	−1,043,185	−2,135,256	−738,584	4,841,265	701,054	−269,205	47,835
1969	937,943	−987,992	−2,008,639	−758,110	5,803,310	756,247	−142,588	28,308
1970	1,993,005	−1,270,391	−1,698,799	−615,470	6,858,372	473,848	167,253	170,948
1971	−297,967	996,368	−1,067,459	−611,444	4,567,400	2,740,607	798,593	174,974
1972	2,636,841	−102,141	−497,285	−761,751	7,502,208	1,642,098	1,368,766	24,667
1973	9,322,828	−2,170,264	−107,111	−626,081	14,188,195	−426,025	1,758,940	160,337
1974	8,291,916	1,491,224	2,312,702	−115,432	13,157,283	3,235,463	4,178,753	670,986
1975	4,575,243	4,586,740	2,743,934	−332,450	9,440,610	6,330,979	4,609,986	453,968
1976	5,575,067	3,142,164	768,252	−1,115,831	10,440,434	4,886,403	2,634,304	−329,413
1977	5,935,784	3,512,328	533,411	−833,903	10,801,151	5,256,567	2,399,463	−47,485
1978	7,167,136	1,420,697	−34,683	−1,011,425	12,032,503	3,164,936	1,831,369	−225,007
1979	9,679,250	−3,150,773	3,805,635	−1,340,344	14,544,617	−1,406,534	5,671,687	−553,926
1980	2,806,241	−2,958,923	−2,092,069	11,709,036	7,671,608	−1,214,684	−226,018	12,495,454
1981	3,326,313	4,286,280	−1,904,499	556,868	8,191,680	6,030,519	−38,448	1,343,286
1982	2,996,549	3,295,829	3,110,761	−513,853	7,861,916	5,040,068	4,976,812	272,565
1983	−238,440	5,317,313	998,209	79,913	4,626,927	7,061,552	2,864,260	866,332
1984	3,539,215	3,111,588	−1,801,321	−21,810	8,404,582	4,855,827	64,730	764,608
1985	6,682,352	2,292,011	−2,635,056	−157,936	11,547,719	4,036,250	−769,005	628,483
1986	3,204,628	1,898,193	−1,685,353	234,458	8,069,995	3,642,432	180,699	1,020,877
1987	6,254,617	−2,336,588	844,741	−232,349	11,119,984	−592,349	2,710,792	554,069
1988	12,679,214	−4,085,529	−1,969,704	−64,421	17,544,581	−2,341,290	−103,653	721,997
1989	15,072,702	−2,003,306	−335,199	406,065	19,938,069	−259,067	1,530,853	1,192,483
1990	17,457,321	−620,271	1,077,084	465,182	22,322,688	1,123,968	2,943,135	1,251,600
1991	8,385,960	3,892,270	−1,425,040	1,513,813	13,251,327	5,636,509	441,012	2,300,232
1992	295,354	6,441,908	2,473,513	834,848	5,160,721	8,186,147	4,339,565	1,621,266
1993	−4,388,862	3,630,554	3,398,521	670,719	476,504	5,374,793	5,264,572	1,457,137
1994	2,836,159	5,299,529	6,508,022	−19,907	7,701,526	7,043,768	8,374,074	766,511
1995	−2,337,350	−1,461,107	2,973,048	1,516,359	2,528,017	283,132	4,839,099	2,302,778
1996	3,431,958	−3,751,331	2,125,351	735,038	8,297,325	−2,007,093	3,991,403	1,521,457
1997	−642,600	3,675,393	−1,091,137	−107,627	4,222,767	5,419,632	774,915	678,791
1998	−15,158,732	5,433,614	712,179	160,758	−10,293,365	7,177,853	2,578,230	947,177
1999	−6,302,832	−401,777	1,974,151	150,086	−1,437,465	1,342,462	3,840,203	936,505
2000	−5,138,832	−2,245,161	−1,128,499	2,406,389	−273,465	−500,922	737,552	3,192,808
2001	−7,299,866	−1,058,232	5,304,992	1,559,392	−2,434,499	686,007	7,171,044	2,345,811
2002	−11,530,464	4,476,810	217,951	695,601	−6,665,097	6,221,048	2,084,003	1,482,019
2003	−7,402,089	−471,438	−2,609,034	−468,306	−2,536,722	1,272,801	−742,982	318,112
2004	−4,439,767	212,859	−2,120,002	−362,201	425,600	1,957,098	−253,951	424,218
2005	−2,378,953	−2,780,242	915,452	−868,136	2,486,414	−1,036,003	2,781,504	−81,718
2006	−1,632,801	−2,208,717	−1,580,187	−1,633,333	3,232,566	−464,478	285,864	−846,914
2007	−3,774,445	−65,627	−3,438,820	94,036	1,090,922	1,678,612	−1,572,769	880,454
2008	−8,317,030	−6,831,084	4,203,752	−874,398	−3,451,663	−5,086,845	6,069,803	−87,980
2009	−21,461,374	887,670	1,016,424	955,956	−16,596,007	2,631,909	2,882,476	1,742,374
2010	−3,999,435	3,485,158	−2,193,870	104,498	865,932	5,229,397	−327,819	890,917
2011	−4,769,261	−9,692,445	4,450,298	−170,964	96,106	−7,948,206	6,316,350	615,454
2012	1,775,331	−4,686,372	2,258,436	−463,190	6,640,698	−2,942,133	4,124,487	323,228

−287,146 and fixed capital formation 4,833,451 is equal to the eigenvalue 45,805,510,321,072. This holds true for the second principal component and others. Loadings represent composition of an eigenvalue of each principal component and give us another important clue, in addition to eigenvectors, when we attempt to interpret the meanings of principal components.

Eigenvalue and its proportion: the sum of eigenvalues of all principal components is equal to the variance of the original variables, which is shown in Table 3-2 "Square of loadings (Japan's GDP)": all information about the variance of the original variables is turned into that of the principal components. As in Table 3-1, the eigenvalue of the first principal component is the largest and that of the fourth principal component is the smallest. The share of the eigenvalue of each principal component in the total eigenvalue is shown as its proportion: the first principal component contains 71.4% of variance of those four variables, the second 15.7%, the third 8.1% and the fourth 4.8% in the decreasing order. Therefore, the first and second principal components with 87.1% in total can explain most of the variance of all the variables. Principal component analysis has the function of effectively summarizing the number of variables.

Generally speaking, the higher the correlations among the original variables are, the larger the proportion of the first principal component is and the more quickly the proportions of the second and other principal components decrease. To the contrary, when the correlations are low, the proportion of the first principal component becomes small and those of the second and others decrease slowly. Principal component analysis on the basis of correlation of variables with no correlations with each other will produce the same proportions among principal components: for example, 0.2 each for five principal components.

Principal component scores: they are time-series values of each principal component that are calculated by means of coefficients in an eigenvector (see Table 4 "Principal component scores of Japan's GDP"). They are usually calculated with their mean being zero; scores with their mean not being zero are also calculated in the table.

Correlations among variables and principal components: Table 5 "Correlation matrix (Japan's GDP)" provides correlation coefficients between variables, between variables and principal components and between principal components. As stated above, all principal components are orthogonal to each other and thus, their correlation coefficients are all zero. By contrast, in the case of principal component analysis on the basis of correlation, the order of correlation coefficients between variables and principal components is the same as that of coefficients in an eigenvector, and their correlation coefficients are the same as the loadings. In the case of principal component analysis on the basis of variance, however, neither property holds true.

The original four variables, i.e. household consumption X_1, government consumption X_2, trade balance X_3 and fixed capital formation X_4 are not orthogonal and thus, more or less correlated to each other. Suppose there are four axes in the four dimensional space, each of which represents each variable, they

Table 5: Correlation matrix (Japan's GDP)

	Household consumption	Government consumption	Trade balance	Fixed capital formation	First PC	Second PC	Third PC	Fourth PC
Household consumption	1.000							
Government consumption	**0.345**	1.000						
Trade balance	0.045	−0.005	1.000					
Fixed capital formation	**0.719**	0.256	−0.198	1.000				
First PC	**0.922**	0.346	−0.103	**0.931**	1.000			
Second PC	0.317	0.132	**0.823**	−0.281	−0.000	1.000		
Third PC	0.213	0.207	−0.557	−0.230	0.000	−0.000	1.000	
Fourth PC	−0.067	**0.905**	0.042	0.020	−0.000	0.000	0.000	1.000

are not orthogonal and their variations influence each other. Four principal components are synthetic variables, Z_1, Z_2, Z_3 and Z_4, made of the four variables X_1, X_2, X_3 and X_4, and set to be orthogonal in the space with their correlations being zero.

Orthogonal relations are extremely useful for regression analysis. If we conduct multiple regression analysis with X_1, X_2, X_3 and X_4 being independent variables, a serious problem of multiple colinearity always occurs owing to mutual correlations among them. It is a rather peculiar phenomenon in which signs of some coefficients of those four variables derived from multiple regression analysis may be reversed against our theoretical expectation or in which reliability on some coefficients may substantially go down. Unfortunately, the trouble does not seem to attract enough attention in some empirical researches. However, multiple regression analysis with principal components, rather than the original variables, being independent variables, will make our research free from multiple colinearity and furthermore, reduce the number of variables despite achieving better results.

Orthogonal and non-orthogonal relations raise a rather radical question about methodology of multivariate analysis: in our example, those four aggregates are *four items* necessary to measure time-series variations of GDP, but they are not appropriate categories that represent *four substances* of GDP necessary to explain its variations.

For example, four items are more or less correlated to each other and thus, you cannot observe or describe variations of household consumption either individually or independently from others, because its variations result from comprehensive effects of all four items. This holds true for government consumption, trade balance and fixed capital formation as well; even if you conduct four individual observations, you actually observe the same phenomenon four times in succession.

More generally speaking, that would lead to the question on what *units* should be set to correctly measure, describe and explain variations of a research object. In our example, those four independent principal components are *units* of four dimensions: it is just like to adopt units of four independent dimensions, such as weight, volume, temperature and position, to record variations of an object. Only after properly setting *units of variations*, we can search essential and autonomous substance, forces and motive power of our research object.

2. Interpretation

Theoretically and empirically strict definitions are already given to GDP's four aggregates, household consumption X_1, government consumption X_2, trade balance X_3 and fixed capital formation X_4. How about those four principal components derived from the aggregates? To begin with, how can we understand the meaning of the first principal component $Z_1 = 0.692 X_1 + 0.097 X_2 - 0.042 X_3 + 0.714 X_4$? Interpretation has to be performed on the basis of signs, values and proportions of the four coefficients. I. T. Jolliffe stated about the possibility of successful interpretation as follows:

"It must be emphasized that although in many examples the PCs can be readily interpreted, this is by no means universally true. There is no reason, *a priori*, why a mathematically derived linear function of the original variables (which is what the PCs are) should have a simple interpretation. It is remarkable how often it seems to be possible to interpret the first few PCs, though it is probable that some interpretations owe a lot to the analyst's ingenuity and imagination. Careful thought should go into any interpretation and, at an earlier stage, into the choice of variables and whether to transform them. In some circumstances, transformation of variables before analysis may improve the chances of a simple interpretation. Conversely, the arbitrary inclusion of logarithms, powers, ratios, etc., of the original variables can make it unlikely that any simple interpretation will be found." (Jolliffe (2002) p.64.)

An outstanding feature of the first principal component is as shown in Table 3-1 that the coefficients of fixed capital formation and household consumption are as large as 0.714 and 0.692 respectively with substantially small government consumption 0.097 and trade balance − 0.042. Our interpretation, therefore, should start with the fact that fixed capital formation and household consumption have positive coefficients of a similar size.

The first principal component, which explain 71.4% of all variations of those four aggregates of Japan's GDP in 1956-2012, seems to represent the basic business cycle in which when fixed capital formation increases, household consumption also increases and vice versa. The basic economic process can smoothly proceed without support of government consumption or reliance on trade surplus; trade balance even has a small negative coefficient. The first principal component can be safely interpreted as a representative indicator of *normal* economic circulation of Japan.

Fig. 6 "First principal component (1)" and Fig. 7 "First principal component (2)" examine time-series data sets of fixed capital formation and household consumption multiplied by each eigenvector and their total, which is compared with the first principal component (mean ≠ 0). In Fig. 6 fixed capital formation and household consumption similarly fluctuate because of their high correlation coefficient 0.719; their total amplifies their fluctuations. As shown in Fig. 7, the total almost perfectly overlaps the first principal component.

The first principal component provides us with much information about Japanese economy. It entered a period of big turmoil after a smooth expansion until the 1970s, mainly due to large fluctuations of fixed capital formation. Stable household consumption in the 1970s made it possible, however, to keep a generally high level. This seems to be the background in Fig. 5 of household consumption being the kick-starter in the period.

The Japanese economy entered a period of stagnation in the first half of the 1980s after experiencing the second Oil Shock in 1979, and a period of the unprecedented bubble from 1987 to 1990. Although the bubble was characterized by speculation in stocks and real estate, it appeared as part of a "normal" business cycle as a result of cumulative effects of fixed capital formation and household consumption as long as GDP is concerned. Further interestingly, the cumulative effects were preceded by

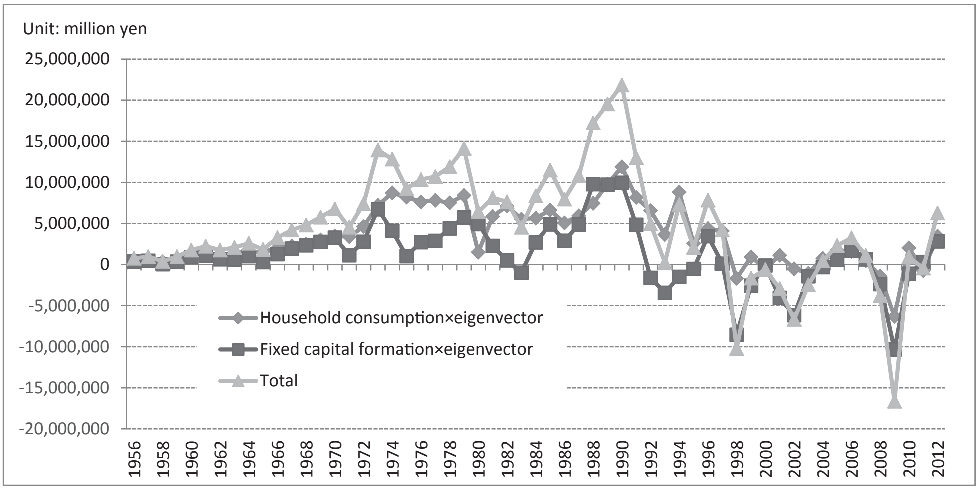


Fig. 6: First principal component (1)

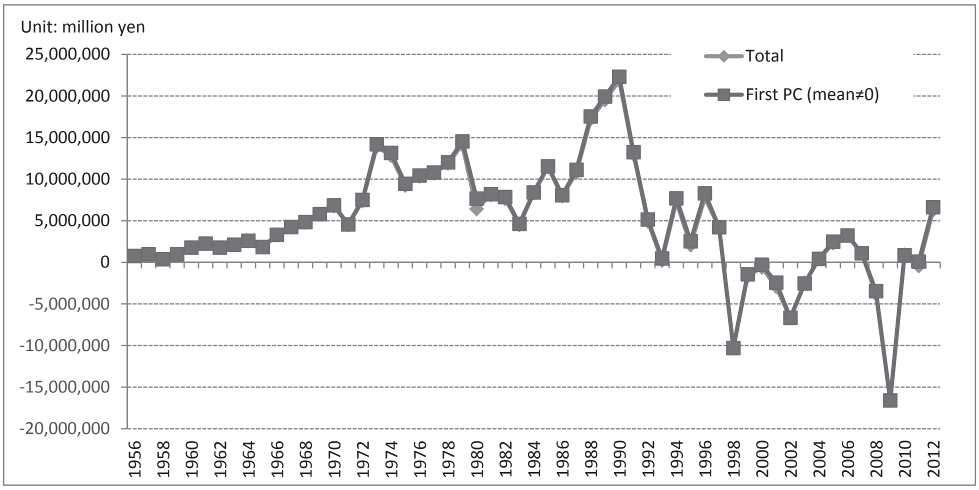


Fig. 7: First principal component (2)

fixed capital formation, followed by a sharp increase in household consumption and finally, reversed by fixed capital formation in the collapse: it is a typical case of boom and burst. The process reflects the fact in Fig. 5 that fixed capital formation after a long pause took the role of the kick-starter in the bubble period.

Since the financial crisis of 1998 [7], a "normal" business cycle completely died down and shrinkage of fixed capital formation dragged the Japanese economy down to the bottom, although household consumption barely supported it. As of 2012 since the Lehman Brothers Shock in September 2008, it began to recover the basic business cycle again.

The outstanding feature of the second principal component in Table 3-1 is the overwhelming coefficient 0.725 of trade balance in the eigenvector with positive household consumption 0.508 and negative fixed capital formation − 0.459. Let us disregard small government consumption 0.079 and interpret the second principal component as follows:

The second principal component, which explains 15.7% of all variations of four GDP aggregates, increases as trade balance increases and fixed capital formation decreases, which are actually negatively correlated with each other with correlation coefficient − 0.198. It apparently represents formation of surplus capital due to shrinkage of national economy and its external emission: i.e. "abnormal" economic shrinkage and formation of international surplus capital (Itaki (2006) Ch. 6).

As to the relation between fixed capital formation and household consumption, they are in highly positive correlation by 0.719 and thus, fluctuate almost simultaneously. In the eigenvector of the second principal component, however, household consumption takes the positive coefficient 0.508, although fixed capital formation takes the negative coefficient − 0.459; the loading of household consumption amounts to 1,611,578 and that of fixed capital formation − 1,458,606, slightly less than the former. These results suggest that in the period, in which a decrease in fixed capital formation leads to shrinkage of GDP as a whole, domestic surplus capital, emitted externally as trade surplus, sustains household consumption to some extent.

Fig. 8 "Second principal component (1)" and Fig. 9 "Second principal component (2)" illustrate in time-series what we discussed above. Fixed capital formation and household consumption are multiplied by respective coefficients in the eigenvector and the results are summed up; the total, trade balance multiplied by its coefficient and the second principal component are compared with each other in Fig. 9. Addition of trade balance to household consumption and fixed capital formation by and large amplifies variations of trade balance in the same direction; in fact, the second principal component and trade balance are in highly positive correlation by 0.823. The peak years of the second principal component, i.e. 1975, 83, 92, 98, 2002 and 10, exactly correspond to those in which surplus capital filled the domestic economy. In 1975, 83 and 92, in particular, the second principal component reached its peaks despite decreases in trade surplus; it suggests that even in the case in which the existence of surplus capital does not explicitly appear owing to various noises such as mobilization of counteracting fiscal policy, the second principal component could work as a good indicator of its existence. In other words, we can interpret the second principal component as the representative that concentrates all the variations of those four aggregates as long as surplus capital is concerned.

The feature of the third principal component in Table 3-1 is that household consumption takes coefficient 0.476 in its eigenvector; by contrast, those of trade balance and fixed capital formation are − 0.685 and − 0.525 respectively. The sign of trade balance turns negative here, though positive in the second principal component. We should note that the combination in the eigenvector of household consumption 0.476, trade balance − 0.685 and fixed capital formation −0.525 stands for exactly the same as the sign-reversed

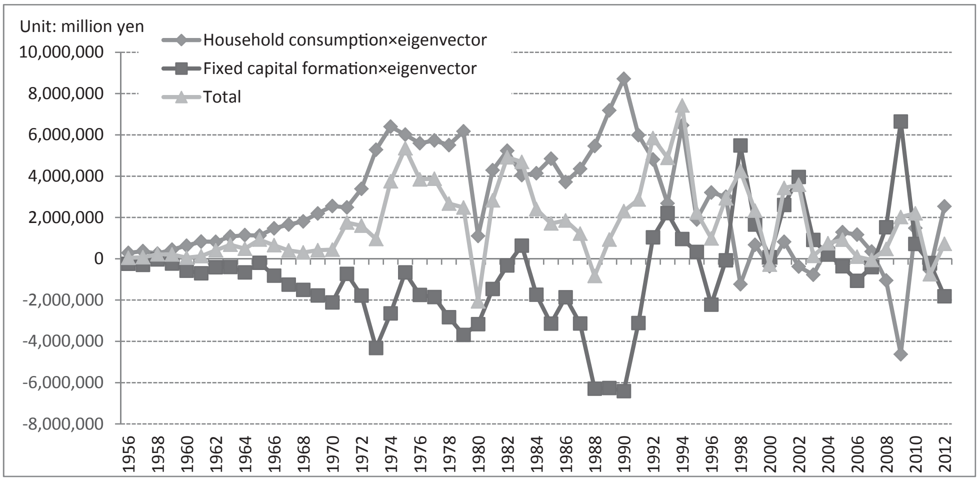


Fig. 8: Second principal component (1)

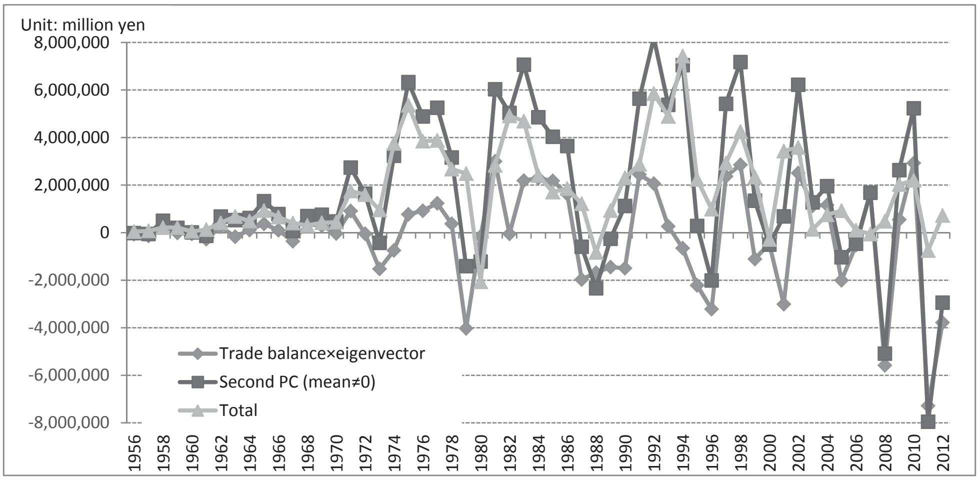


Fig. 9: Second principal component (2)

combination of household consumption − 0.476, trade balance 0.685 and fixed capital formation 0.525; in one case the component increases in positive value and in the other it does in negative value. Now, let us temporarily disregard government consumption with relatively small coefficient 0.172 and start our interpretation of the third principal component.

Despite its small proportion 8.1%, the third principal component carries a great significance. As in Table 5, household consumption and fixed capital formation are in the highest positive correlation among those four aggregates; nevertheless, they show opposite signs to each other in the eigenvector. How can we interpret it? And although household consumption and trade balance are also in slightly positive correlation, they show opposite signs, too. It means that the third principal component combines those three aggregates in the perfectly reversed manner to their original correlations. Interpretation about that is extremely difficult.

The third principal component increases as household consumption increases. But, by contrast to the first principal component, it increases as fixed capital formation decreases. Strictly speaking, it increases as discrepancy between household consumption and fixed capital formation stretches: i.e. it increases when fixed capital formation is sluggish but, household consumption autonomously expands or to the contrary, when household consumption does not shrink very much but, fixed capital formation autonomously declines quite sharply. These situations are exactly the opposite to those of the first principal component in which household consumption and fixed capital formation go up and down hand in hand. In this sense, we can safely say that the third principal component is the opposite moment to the first.

Next, the third principal component increases as discrepancy between household consumption and trade balance expands. In other words, it increases when trade balance is sluggish but, household consumption autonomously increases or to the contrary, when household consumption does not shrink very much but, trade balance autonomously deteriorates precipitously. These situations are exactly the opposite to those of the second principal component in which household consumption and trade balance simultaneously fluctuate in the same direction. In this sense, the third principal component is also the opposite moment to the second.

Fig. 10 “Third principal component” examines in time-series what we discussed above. Rather complicated as it is, we find that the peak years of the third principal component came in 1975, 79, 82, 90, 94, 99, 2001, 05, 08 and 11. Pay an attention to negative signs of fixed capital formation and trade balance in the eigenvector and compare them with the principal component, and we can identify

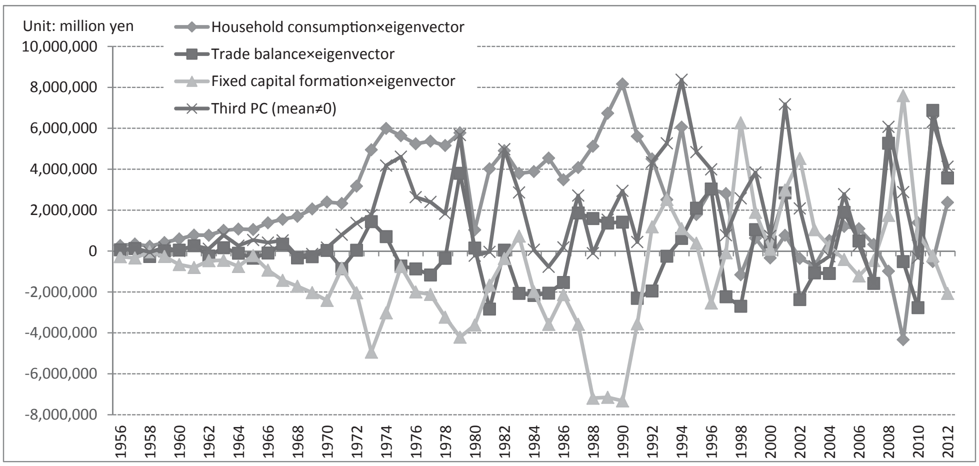


Fig. 10: Third principal component

following factors that caused peak years: in 1975 a sharp decrease in fixed capital formation (i.e. a simultaneous global recession); in 1979 a sudden decrease in trade balance (i.e. the second Oil Shock); in 1982 a decrease in both fixed capital formation and trade balance (i.e. a simultaneous global recession); in 1987 and 90 a decrease in trade balance despite a sharp increase in household consumption and fixed capital formation (i.e. the Heisei bubble economy); in 1994 an increase in household consumption and a decrease in trade balance (i.e. a slight recovery after the bust of the bubble economy); in (1998-) 1999 a precipitous decrease in household consumption and a collapse of fixed capital formation despite a jump of trade balance (i.e. a financial crisis); in 2001 a decrease in both fixed capital formation and trade balance despite stagnant household consumption (i.e. the bust of the global IT bubble); in 2005 a slight recovery of household consumption and a decrease in trade balance (i.e. a recovery after the bust of the IT bubble); in 2008 (-09) a collapse of both household consumption and fixed capital formation and a jump of trade balance (i.e. the Lehman Brothers Shock); and in 2011 stagnant household consumption and fixed capital formation and a decline of trade balance (i.e. the Great East Japan Earthquake). These facts suggest that an autonomous increase in household consumption exceptionally took place only in the Heisei bubble period of 1987 and 1990, in a recovery period of 1994 after the bust of the bubble and in another recovery period of 2005 after the bust of the IT bubble. And thus, many cases of an increase in the third principal component took place when fixed capital formation decreased or trade balance deteriorated due to an adverse situation of the world economy as a whole, with household consumption being kept rather stable.

Let us divide what the third principal component represents into positive aspects and negative aspects. As to its negative aspects, they are all “abnormal situations” for smooth development of capitalistic economy such as an expansion of household consumption due to an economic babble, an unexpected decline of fixed capital formation and a sharp decrease in trade balance due to a simultaneous global recession, an oil shock or an earthquake disaster. It is, therefore, an opposite moment to the first principal component that represents basic and “normal” business cycle of the Japanese economy, and also an opposite moment to the second principal component that allows household consumption to sustain or even increase because of surplus capital turning into trade surplus. As to its positive aspects, household consumption resists an unexpected decline of fixed capital formation and trade balance and keeps a certain level despite those adverse effects. Although the first principal component stands for a smooth development of the Japanese economy, it is just that of “capitalistic Japanese economy”, in which a decline of fixed capital formation forces people’s consumption to shrink that should be kept intact whatever happens. And the second principal component, when the first principal component malfunctions, merely transfers a domestic problem abroad and keeps household consumption to a certain level. Therefore, the third principal component embodies an economy’s intrinsic nature, i.e. economic activities ultimately for people’s consumption, rather than its capitalistic nature that are symbolized by the first and second principal components; or it expresses an economy's trans-historical nature, rather than its specifically historical nature. In this sense as well, the third principal component is an opposite moment to both the first and second principal components.

The most noticeable feature of the fourth principal component in Table 3-1 is that the coefficient of government consumption in its eigenvector is extremely high, 0.977. By contrast, trade balance and fixed capital formation are almost negligible while household consumption has a slightly large coefficient − 0.193. So let us interpret the meaning of the fourth principal component by the relations between government and household consumption. It carries important conceptual significance despite its substantially small proportion 4.8%.

As in Table 5, government consumption and household consumption are positively correlated with coefficient 0.345, which are combined with reverse coefficients 0.977 and − 0.193 respectively in the eigenvector. We should not underestimate the significance of − 0.193 since variations of household consumption in their absolute terms overwhelm those of government consumption and thus, exert considerable negative effects on it. In general, the movements of government consumption synchronizes those of household consumption so that when government consumption increases more than household consumption does and when the former decreases more than the latter, the fourth principal component fluctuates up and down: the discrepancy between them is expressed in the fourth principal component.

Fig. 11 "Fourth principal component" examines in time-series what we discussed above. Except in 1980 when government consumption shot up enormously, the fourth principal component moves by and large in line with government consumption: the correlation coefficients between the fourth principal component and government consumption and between the fourth principal component and

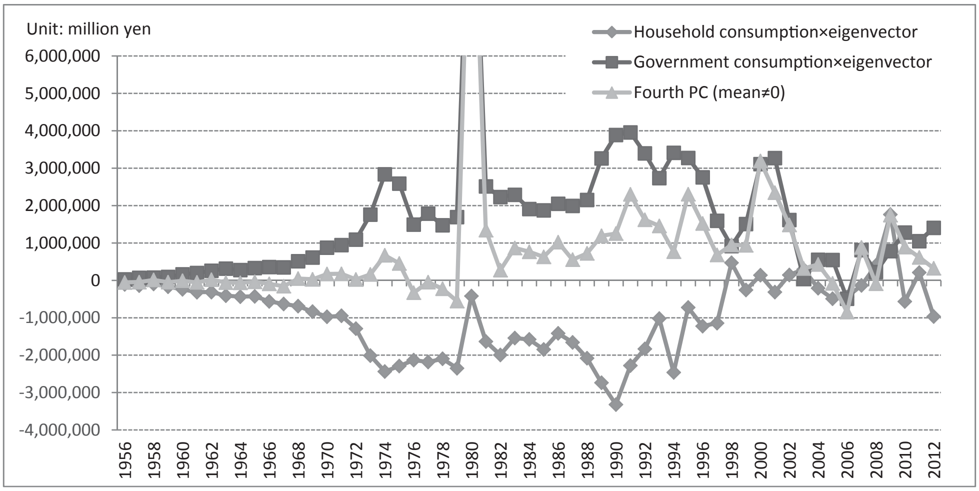


Fig. 11: Fourth principal component

household consumption are 0.905 and − 0.067 respectively. Peak years of the fourth principal component came in 1974, 80, 91, 95, 2000 and 09, which almost perfectly correspond to those of government consumption. You may suppose that there is virtually no difference between the fourth principal component and government consumption. However, as we see later, not only the fourth principal component determines government consumption, but also does the first principal component that represents normal business cycle, the second principal component that represents external emission of domestic surplus capital and the third principal component that represents various "abnormalities".

We can understand, therefore, that the fourth principal component expressed autonomous variations of government consumption, which affect the economy as a whole, apart from those of household consumption, trade balance and fixed capital formation. It is certain, of course, that all variations of government consumption are not autonomous; and all the other components also affect government consumption. To the contrary, such autonomous variations do not have influences only on government consumption: coefficients in the eigenvector of trade balance and fixed capital formation are, however small they may be, 0.067 and 0.059 and thus, they exert some positive effects on part of trade balance and fixed capital formation. The fourth principal component factors in all those effects.

Our interpretation above allows us to understand that the fourth principal component stands on a completely different dimension from those of the first, second and third principal components: the first, i.e. the "normal" business cycle, the second, i.e. the "extraordinary" business cycle with surplus capital and the third, i.e. the "abnormal" situations that disturb business cycles or trans-historical economic situations, are all representative of Japanese economy's internal mechanism. By contrast, the fourth principal component is considered to be external policy mechanism, exerted on the internal mechanism, which is an opposite moment in a completely different dimension.

Policy mechanism as it is, fiscal policies regarding "normal" business cycles are included in the first principal component, those regarding "extraordinary" business cycles in the second and those regarding "abnormal" situations in the third: they are 0.097, 0.079 and 0.172 respectively in the row of government consumption in the eigenvector and 655,022, 249,658 and 392,460 in loadings. The loading of the fourth principal component is 1,712,535 and thus, the following unequal relations hold true: the second < the third < the first < the fourth. The autonomous expenditure, not directly related to the economic mechanism, records the largest and is followed by the expenditure for ordinary economic fluctuations. They are in proportion 8.30% < 13.04% < 21.76% < 56.90%.

(To be continued.)

Notes

1）The tentative Japanese translation of this paper is available in *Working Paper Series*, IR2014-4, “ 軌道分析と主成分分析の組み合わせによる時系列多変量解析 ”, published in March, 2015, College of International Relations, Ritsumeikan University (http://www.ritsumei.ac.jp/acd/cg/ir/college/bulletin/workingpaper/IR2014-4.pdf).

The author highly appreciates detailed and encouraging comments that Professor Mitsunobu TAKAHAMA kindly gave me on my presentation of orbit analysis in the spring conference of the Japan Society of International Economics on June 7th, 2014 in the Tama campus of Hose University.

2）“<>” stands for “ ≠ ” in EXCEL.

3）EXCEL template for orbit analysis is available. Have a contact with itaki@ir.ritsumei.ac.jp if necessary.

4）It is said that Peason (1901) and Hotelling (1933) originated principal component analysis and Hotelling first used the term “principal components”. Refer to Jolliffe (2002) pp.6-9 and Leeuw (2013) as for the history of its development.

5）Refer to Jolliffe (2002), Ueda (2003) and Uchida (2013) as for mathematical properties of principal component analysis and computation of principal components.

6）See “VII. Conceptual framework of orbit analysis” and “VIII. Leading-following relations and the Granger causality”, Itaki (2014).

7）In November 1997 the Hokkaido Takushoku Bank went bankrupt and the Yamaichi Securities Company went out of business. In October 1998 the Japan Long-term Credit Bank (now the Shinsei Bank) and in December the Japan Securities Credit Bank (now the Aozora Bank) were temporarily nationalized. The situation was exactly that on the eve of a financial crisis. See Itaki (2006) p.233.

References

Granger, Clive W. J. (2003), “Nobel Lecture: Time series analysis, cointegration, and applications”, December 8, 2003.

Hegel, G. W. F. (1991) [1817], *The Encyclopaedia Logic*, Indianapolis/ Cambridge: Hackett Publishing Company, Inc., translated by T. F. Geraets, W. A. Suchting and H. S Harris.

Hotelling, Harold (1933), *Analysis of a Complex of Statistical Variables into Principal Components*, 1933, Warwick & York, reprinted from “Analysis of a complex of statistical variables into principal components”, *Journal of Educational Psychology*, vol. 24 (6), September 1933, pp.417-441 and vol. 24 (7), October 1933, pp.498-520.

Itaki, Masahiko (2002a), “Formulation of dialectics as social science methodology: representations, forms of existence and essence (1)”, *The Ritsumeikan Journal of International Studies*, vol. 14 no. 4, March 2002, pp.1-25, in Japanese.（板木雅彦「社会科学方法論としての弁証法の定式化――表象および存在形態と本質（上）」『立命館国際研究』第 14 巻 4 号，2002 年 3 月，1－25 ページ）

Itaki, Masahiko (2002b), “Formulation of dialectics as social science methodology: representations, forms of existence and essence (2)”, *The Ritsumeikan Journal of International Studies*, vol. 15 no. 1, June 2002, pp.53-72, in Japanese.（板木雅彦「社会科学方法論としての弁証法の定式化――表象および存在形態と本質（中）」『立命館国際研究』第 15 巻 1 号，2002 年 6 月，53－72 ページ）

Itaki, Masahiko (2002c), “Formulation of dialectics as social science methodology: representations, forms of existence and essence (3)”, *The Ritsumeikan Journal of International Studies*, vol.

15 no. 2, October 2002, pp.1-17, in Japanese.（板木雅彦「社会科学方法論としての弁証法の定式化――表象および存在形態と本質（下）」『立命館国際研究』第 15 巻 2 号，2002 年 10 月，1－17 ページ）

Itaki, Masahiko (2003a), “Formulation of dialectics as social science methodology: forms of motion in space and their functions and barriers”, *The Ritsumeikan Journal of International Studies*, vol. 15 no. 3, March 2003, pp.249-268, in Japanese.（板木雅彦「社会科学方法論としての弁証法の定式化――空間的運動形態と機能・制限」『立命館国際研究』第 15 巻 3 号，2003 年 3 月，249－268 ページ）

Itaki, Masahiko (2003b), “Formulation of dialectics as social science methodology: forms of motion in time and their functions and barriers”, *The Ritsumeikan Journal of International Studies*, vol. 16 no.1, June 2003, pp.31-47, in Japanese. 板木雅彦「社会科学方法論としての弁証法の定式化――時間的運動形態と機能・制限」『立命館国際研究』第 16 巻 1 号，2003 年 6 月，31－47 ページ）

Itaki, Masahiko (2006), *The Birth of International Surplus Capital*, Kyoto, Japan: Minerva Shobou, in Japanese.（板木雅彦『国際過剰資本の誕生』ミネルヴァ書房）

Itaki, Masahiko, (2014), “Orbit analysis of leading-following relations among multiple variables”, *The Ritsumeikan Journal of International Studies*, vol. 27 no. 1, June 2014, pp.1-33.（『立命館国際研究』第 27 巻 1 号，2014 年 6 月）
(http://www.ritsumei.ac.jp/acd/cg/ir/college/bulletin/Vol.27-1/27_1_01_Itaki.pdf)

Jolliffe, I. T. (2002), *Principal Component Analysis*, second edition, New York: Springer-Verlag.

Leeuw, Jan De (2013), “History of nonlinear principal component analysis” (http://escholarship.org/uc/item/1vp9f9kz#page-2)

Pearson, K (1901), “On lines and planes of closest fit to systems of points in space”, *Philosophical Magazine*, July-December 1901, pp.559-572.

Stone, Richard (1947), “On the interdependence of blocks of transactions”, *Supplement to the Journal of the Royal Statistical Society*, vol. IX, no. 1.

Uchida, Osamu (2013), *The Basics and Applications of Principal Component Analysis*, Nichigikaren, in Japanese.（内田治『主成分分析の基本と活用』日技科連）

Ueda, Shoichi (2003), *Principal Component Analysis*, Asakura Shoten, in Japanese.（上田尚一『主成分分析』朝倉書店）

（ITAKI, Masahiko, Professor, College of International Relations,
Ritsumeikan University, itaki@ir.ritsumei.ac.jp）

軌道分析と主成分分析の組み合わせによる 時系列多変量解析（上）

自然科学においても社会科学においても，分析の基本的な課題は，分析対象の構造をおさえた上でその変化をとらえることにあるといってよかろう。そのために科学者は，さまざまな変数を工夫してその変動を観察する。しかし，計測の便宜のために仮説的に設定されたそれら変数が，そのまま分析対象の変化をもたらす力であったり実体であったりすることは稀である。ましてや，それら変数の間に明瞭な因果関係を特定できることはほとんど期待できない。

本稿の課題は，「変化をとらえる単位」は何かを考察することである。この単位とは，変化をもたらす力・実体ということができる。これを通常の主成分分析に依拠しながら考察した後に，「変化をもたらす諸力の重層構造」をとらえるために発生的主成分分析が提起されている。ここで，互いに直交関係にある主成分が，弁証法的な「対立物の統一の重層構造」に読み替えられることになる。

そして，変数間の先導・追従関係をとらえる軌道分析（板木（2014））を，全変数・主成分に施すことで, 主成分を原因, 変数を結果とする新しい因果関係論が構築される。いわゆる「グレンジャーの因果性」の成立は原理的に否定されたが（同上），それに代わる新たな因果関係論が提起されている。

なお，国内総生産を例証として用いた副産物として，固定資本形成，民間消費，貿易収支，政府消費それぞれが1単位増大した場合の国内総生産の増大分—すなわち，乗数効果を主成分重回帰分析によって正確に計測する手法が示されている。

（板木 雅彦，立命館大学国際関係学部教授）