Detection of change points in panel data based on Bayesian MT method

Takumi Enomoto*, Yasushi Nagata

Waseda University: 3-4-1 Okubo, Shinjuku-ku, Tokyo 169-8555, Japan.
*contact author’s e-mail address: enoeno@akane.waseda.jp

Abstract:
Panel data are a sort of multi-dimensional time-series data that consist of several sets of the observation unit as an identical community over time. Panel data include more information than time-series data or cross-section data. Using panel data, we are able to estimate more accurately, increase degrees of freedom, and avoid multi-collinearity.

In this paper, we consider a change point problem in panel data. Detection methods for change points in panel data have not been extensively proposed. This paper proposes a detection method for change points in panel data using the Mahalanobis–Taguchi (MT) method (Taguchi (2002)). As the MT method does not assume panel data or time-series data, we must extend the theory of the MT method. Furthermore, when the sample size is small, it is difficult to estimate the covariance matrix precisely (Miyakawa et al. (2007)). In this paper, the MT method is extended using Bayesian inference to resolve the above-mentioned difficulties. Conducting Monte Carlo simulations and a real data analysis, we show that our proposed method is useful.

Keywords
MT method, Bayesian analysis, Panel data, Change point

1. Introduction

In this study, we use panel data, which are a sort of multi-dimensional time-series data that consist of several sets of the observation unit as an identical community over time (Cheng (2007)). In short, panel data are cross-sectional time-series data. Figure 1 shows the concept of panel data.

Panel data include more information than time-series data or cross-section data. Using panel data, we are able to estimate more accurately, increase degrees of freedom, and avoid multi-collinearity. However, in panel data, there may exist change points in time caused by changes in the structure of populations, error distributions, and other factors.

This paper considers the detection of change points in panel data using the Mahalanobis–Taguchi (MT) method (Taguchi (2002)). The MT method is an MT system procedure that is a collection of multivariate analysis procedures proposed by Dr. Gen-ichi Taguchi. Tatebayashi et al. (2008) and Woodall et al. (2003) summarized procedures of the MT system clearly.

The MT method is used to detect outliers; however, the existing method cannot treat panel data or time-series data. Furthermore it is difficult for the MT method to estimate the covariance matrix accurately when the sample size is small (Miyakawa et al. (2007)).

This paper extends the MT method using Bayesian inference to detect the change point of the panel data and deal with above problems. Namely, the proposed method enables us to detect the change points in panel data or multivariate time-series data more accurately even if the sample size is small. We call this proposed method the Bayesian Mahalanobis-Taguchi (BMT) method.

Iizuka (1982) applied the Bayes approach to determine the individual normal range of a serum chemical ingredient in the field of clinical medicine. The normal range of each chemical ingredient for the whole healthy group was determined conventionally based on the healthy group. However, the chemical ingredients have substantial individual differences. Thus, Iizuka did not set the normal range for the whole healthy group, but sought to determine a normal range for individuals. As the sample size was small for individuals, Iizuka tried to solve the problem using Bayesian inference for a type of panel data. For this inference problem, Iizuka dealt with each item separately. In other words, Iizuka repeated univariate analysis.

Radhakrishnan (1984) considered the estimation of the Mahalanobis distance using Bayesian inference for
three situations. The first situation is the estimation of the Mahalanobis distance (Euclidean distance) between two groups of univariate. The second is the estimation of the Mahalanobis distance between two groups of $p$-variates. The third is the estimation of the Mahalanobis distance between $k$ groups of $p$-variates.

In this paper, we combine the MT method and Bayesian inference based on multivariate data. In addition, we propose the BMT method, which is an extension of Iizuka (1982) to multivariate data, to detect change points for panel data or multivariate time-series data when sample size is small.

This paper consists of six sections. In Section 2, we explain the MT method. In Section 3, we propose the BMT method. In Section 4, we conduct Monte Carlo simulations to evaluate the performance of the BMT method. In Section 5, we conduct a real data analysis based on the BMT method using data from the Japan National Tax Agency (2015) and discuss the results. Finally, in Section 6, we present our conclusions and directions for future studies.

![Figure 1: Concept of panel data](image)

2. MT Method

The MT method was developed by Dr. Gen-ichi Taguchi as a technique for implementing pattern recognition (Taguchi (2002)). The MT method performs pattern recognition by applying the Mahalanobis distance ($MD$). The MTA method, TS method, T method, and RT method have all been proposed after the proposal of the MT method. Tatebayashi et al. (2008) and Woodall et al. (2003) summarized these procedures. The collection of these procedures is called the MT system.

The main features of the MT method are that it sets a unit space and uses a $MD$. Unlike discrimination analysis, the MT method sets only one unit space as a group. Unknown data are determined to belong to the unit space or not using the $MD$ from the center of the unit space. In the MT method, the unit space is determined as a group with the homogeneity for this purpose.

The MT method is carried out in the following steps.

Step1: Set a unit space.

Step2: Extract the features from the data (determine the variables).

Step3: Normalize each variable.

We assume that the unit space consists of $k$ variables and $n$ cases (the sample size). Let $x_{ij}$ be the value of the $j$th variable ($j=1,2,...,k$) in the $i$th case ($i=1,2,...,n$). These variables are standardized by equation (1):

$$X_{ij} = \frac{x_{ij} - \overline{x}_j}{s_j},$$

where $\overline{x}_j$ and $s_j$ are the sample mean and the standard deviation of the $j$th variable, respectively. Then, we calculate the correlation coefficient matrix $R$ and its inverse matrix $R^{-1}$.

Furthermore, if new data $y = (y_1, y_2, ..., y_k)^T$ have been obtained, they are normalized as in equation (2):

$$Y_j = \frac{y_j - \overline{x}_j}{s_j}.$$
Step 4: Define the threshold.
A threshold is determined according to the data or some standard. It has been studied for a long time about the distribution of the Mahalanobis distance. Such as Hawkins (1981) and Gnanadesikan and Kettenring (1972) discussed the distribution of the Mahalanobis distance. Furthermore, Tracy et al. (1992) derived the exact distribution of the Mahalanobis distance. In this paper, we set the 80 percentile as the threshold in the MT method using the result of Tracy et al. (1992).

Step 5: Calculate the MD for unknown data.
Using the normalized new data \( \mathbf{Y} = (Y_1, Y_2, ..., Y_k)^T \), we calculate the MD. The MD in the MT method is calculated as in equation (3). In the MT method, the squared distance is divided by the number \( k \) of variables:

\[
MD = \frac{\mathbf{Y}^T \mathbf{R}^{-1} \mathbf{Y}}{k}.
\]

Step 6: Determine whether the new data are normal or abnormal.
Step 7: Update the unit space.

3. Bayesian MT Method

In this section, we propose the Bayesian MT (BMT) method by combining the procedures proposed by Iizuka (1982) and Radhakrishnan (1984). The BMT method enables us to detect change points in panel data.

Let \( x_{ip} \) be an observed value, where the index \( i (1 \leq i \leq n) \) denotes the time-series and \( p (1 \leq p \leq k) \) denotes an observation item (note that in the MT system, a variable is called an item). Moreover, we define \( \mathbf{x}_i = (x_{i1}, x_{i2}, ..., x_{ip}, ..., x_{ik})^T \). Let us assume that

\[
\mathbf{x}_i \sim N_i(\mathbf{\mu}, \mathbf{A}^{-1}) \ i.i.d.,
\]

where \( N_i(\mathbf{\mu}, \mathbf{A}^{-1}) \) is a \( k \)-variates normal distribution whose mean vector is \( \mathbf{\mu} \) and whose covariance matrix is \( \mathbf{A}^{-1} \). Then, the likelihood function of \( \mathbf{\mu} \) and \( \mathbf{A}^{-1} \) can be expressed as in equation (5):

\[
f(\mathbf{x} | \mathbf{\mu}, \mathbf{A}^{-1}) = \prod_{i=1}^{n} \left| \mathbf{A} \right|^{-\frac{n}{2}} \exp \left( -\frac{1}{2} (\mathbf{x}_i - \mathbf{\mu})^T \mathbf{A}^{-1} (\mathbf{x}_i - \mathbf{\mu}) \right) \nonumber \]

\[
\propto \left| \mathbf{A} \right|^{-\frac{n}{2}} \exp \left( -\frac{1}{2} \sum_{i=1}^{n} (\mathbf{x}_i - \mathbf{\mu})^T \mathbf{A}^{-1} (\mathbf{x}_i - \mathbf{\mu}) \right) \nonumber \]

\[
= \left| \mathbf{A} \right|^{-\frac{n}{2}} \exp \left( -\frac{1}{2} \text{tr}(\mathbf{AS}) \right),
\]

where \( \mathbf{S} \) is defined as

\[
\mathbf{S} = \sum_{i=1}^{n} (\mathbf{x}_i - \mathbf{\mu})(\mathbf{x}_i - \mathbf{\mu})^T.
\]

Moreover, we consider the prior distribution of \( \mathbf{\mu} \) and \( \mathbf{A}^{-1} \) as the within-subject distribution. When we define \( \mathbf{\alpha} \) as the within-subject mean, \( \beta \) as the quantity of prior information about the within-subject mean, \( \mathbf{W} \) as the within-subject variance, and \( \nu \) as the quantity of prior information about the within-subject variance, we assume that the joint probability density function \( f(\mathbf{\mu}, \mathbf{A}^{-1}) \) can be expressed as

\[
f(\mathbf{\mu}, \mathbf{A}^{-1}) = N_i(\mathbf{\mu}|\mathbf{\alpha}, (\beta \mathbf{A})^{-1}) \times W(\mathbf{A}|\mathbf{W}, \nu),
\]

where \( W(\mathbf{A}|\mathbf{W}, \nu) \) is the probability density function of the \( k \)-variate Wishart distribution with covariance matrix \( \mathbf{W} \) and degrees of freedom \( \nu \), which is defined as

\[
W(\mathbf{A}|\mathbf{W}, \nu) = 2^{\frac{k \nu}{2}} \pi^\frac{k(k-1)}{4} \left| \mathbf{W} \right|^{\frac{\nu}{2}} \prod_{p=1}^{k} \left( \frac{\nu + 1 - \beta}{2} \right) \left| \mathbf{A} \right|^{\nu - k - 1} \exp \left( -\frac{1}{2} \text{tr}(\mathbf{W}^{-1} \mathbf{A}) \right).
\]
Moreover, we define
\[
N_k (\mu \mid a, (\beta A)^{-1}) = \frac{1}{(2\pi)^{n_k/2}} \exp \left\{ -\frac{\beta}{2} (a - \mu)^T A (a - \mu) \right\}.
\]  

(9)

Denoting \( D \) as
\[
D = 2^{\nu/2} \pi^{-\nu/4} \prod_{p=1}^k \left( \frac{\nu + 1 - p}{2} \right),
\]

(10)
equation (8) is expressed as
\[
W(\Lambda \mid W, \nu) = D|\Lambda|^{\nu/2 - 1} \exp \left\{ -\frac{1}{2} \nu \text{tr}(W^{-1} \Lambda) \right\}.
\]

(11)

Using equations (9) and (11), equation (7) can be expressed as
\[
f(\mu, \Lambda^{-1}) = \frac{1}{(2\pi)^{n/2}} \exp \left\{ -\frac{\beta}{2} (a - \mu)^T A (a - \mu) \times D|\Lambda|^{\nu/2 - 1} \exp \left\{ -\frac{1}{2} \nu \text{tr}(W^{-1} \Lambda) \right\} \right\} \times |\Lambda|^{\nu/2} \exp \left\{ -\frac{1}{2} \nu \text{tr}(W^{-1} \Lambda) \right\} \times |\Lambda|^{\nu/2} \exp \left\{ -\frac{1}{2} \nu \text{tr}(A S) \right\}.
\]

(12)

Next, we derive the posterior distribution. We assume that the data \( x_1, x_2, \ldots, x_n \) have been obtained and denote the average as \( \bar{x} \) and the covariance matrix as \( V \). Then, the probability density function of the posterior distribution \( f(\mu, \Lambda^{-1} \mid \bar{x}, V) \) can be expressed as equation (13) using the Bayes theorem:
\[
f(\mu, \Lambda^{-1} \mid \bar{x}, V) \propto f(\mu, \Lambda^{-1} \mid \bar{x}, V) \times f(\bar{x}, V \mid \mu, \Lambda^{-1})
\]
\[
= |\Lambda|^{\nu/2} \exp \left\{ -\frac{1}{2} \nu \text{tr}(W^{-1} \Lambda) \right\} \times |\Lambda|^{\nu/2} \exp \left\{ -\frac{1}{2} \nu \text{tr}(A S) \right\}.
\]

(13)

We define the updating parameters as follows:
\[
a' = n\bar{x} + \beta a.
\]

(14)
\[
\beta' = \beta + n.
\]

(15)
\[
W' = W + S + \frac{\beta a}{\beta + n} (\bar{x} - a)^T.
\]

(16)
\[
\nu' = \nu + n.
\]

(17)

From Radhakrishnan (1984), the conditional posterior distribution of the Mahalanobis distance follows the non-central chi-square distribution with non-centrality parameters
\[
(a' - a)^T \left( 1 + \frac{1}{\beta'} \right)^{-1} (a' - a) \] and degrees of freedom \( \nu' \).

Then, in the BMT method, the \( MD \) of the new sample \( y \) can be expressed as
\[
MD = (y - a')^T \left( 1 + \frac{1}{\beta'} \right)^{-1} (y - a').
\]

(18)
In this paper, we set the 90 percentile as the threshold in the BMT method using the above non-central chi-square distribution.

4. A Verification of the BMT Method

In this section, we evaluate the performance of the BMT method by Monte Carlo simulations. In subsection 4.1, we describe the simulation settings. In subsection 4.2, we show the results of the simulation and discuss the results.

4.1 Simulation Settings

In this subsection, we present the simulation conditions used to evaluate the performance of the BMT method. In this simulation, we detect change points by using the BMT method and MT method. First, we prepare four observation targets (A, B, C, D) as the data format of the unit space. Next, we set four items for each observation. Finally, we set five times for each observation item. In addition each observation target in the unit space follows the normal distribution as shown in Table 1. We assume that there are correlations, as shown in Table 2.

Then, we describe the test data. We set the observation target A as the test target. In other words, we detect the change points of the target A by the MT and BMT methods. We assume test data for 30 time series. The test data have two change points. The test data for periods 1–10 are generated in the same way as in the unit space. The test data for periods 11–20 and periods 21–30 are generated using different means, variances and correlation coefficients, which are shown in Table 3 and Table 4. It should be noted that the unit space of the MT method uses the only sample of the observation target A in the panel data. In the BMT method, we use the mean of the panel data to obtain the initial value of $\alpha$. Moreover, we set the initial value of $W$ as the covariance matrix of the unit space. In addition, we set $\beta = 4, \nu = 20$. The simulations are replicated 100,000 times.

Table 1: Distribution of each observation target in the unit space

| Target A | Item 1 | Item 2 | Item 3 | Item 4 |
|----------|--------|--------|--------|--------|
|          | N(5,1) | N(10,1) | N(0,2) | N(5,2) |
| Target B | N(2,1) | N(6,1) | N(3,2) | N(9,2) |
| Target C | N(5,1) | N(11,1) | N(2,2) | N(5,2) |
| Target D | N(6,1) | N(8,1) | N(0,2) | N(4,2) |

Table 2: Correlation coefficient matrix of each observation target in the unit space

| Target A | Item 1 | Item 2 | Item 3 | Item 4 |
|----------|--------|--------|--------|--------|
| Item 1   | 1      | -0.2   | 0.8    | -0.1   |
| Item 2   | -0.2   | 1      | -0.6   | 0.7    |
| Item 3   | 0.8    | -0.6   | 1      | -0.3   |
| Item 4   | -0.1   | 0.7    | -0.3   | 1      |

| Target B | Item 1 | Item 2 | Item 3 | Item 4 |
|----------|--------|--------|--------|--------|
| Item 1   | 1      | 0      | 0      | 0      |
| Item 2   | 0      | 1      | 0      | 0      |
| Item 3   | 0      | 0      | 1      | 0      |
| Item 4   | 0      | 0      | 0      | 1      |

| Target C | Item 1 | Item 2 | Item 3 | Item 4 |
|----------|--------|--------|--------|--------|
| Item 1   | 1      | -0.1   | 0.7    | -0.3   |
| Item 2   | -0.1   | 1      | -0.2   | 0.7    |
| Item 3   | 0.7    | -0.2   | 1      | -0.2   |
| Item 4   | -0.3   | 0.7    | -0.2   | 1      |

| Target D | Item 1 | Item 2 | Item 3 | Item 4 |
|----------|--------|--------|--------|--------|
| Item 1   | 1      | 0.3    | 0.5    | -0.3   |
| Item 2   | 0.3    | 1      | -0.4   | 0.7    |
| Item 3   | 0.5    | -0.4   | 1      | -0.3   |
| Item 4   | -0.3   | 0.7    | -0.3   | 1      |
Table 3: Distribution for periods 11-20 and 21-30 for the test data of the target A

|                | Item 1    | Item 2    | Item 3    | Item 4    |
|----------------|-----------|-----------|-----------|-----------|
| Periods 11-20  | \(N(7,1^2)\) | \(N(8,1^2)\) | \(N(4,2^2)\) | \(N(3,2^2)\) |
| Periods 21-30  | \(N(3,1^2)\) | \(N(12,1^2)\) | \(N(0,2^2)\) | \(N(4,2^2)\) |

Table 4: Correlation coefficient matrix for periods 11-20 and 21-30 for the test data of the target A

|                | Item 1  | Item 2  | Item 3  | Item 4  |
|----------------|---------|---------|---------|---------|
| Periods 11-20  |         |         |         |         |
| Item 1         | 1       | 0       | 0.8     | 0.2     |
| Item 2         | 0       | 1       | 0       | 0.7     |
| Item 3         | 0.8     | 0       | 1       | 0       |
| Item 4         | 0.2     | 0.7     | 0       | 1       |
| Periods 21-30  |         |         |         |         |
| Item 1         | 1       | -0.2    | 0       | -0.1    |
| Item 2         | -0.2    | 1       | -0.6    | 0       |
| Item 3         | 0       | -0.6    | 1       | -0.3    |
| Item 4         | -0.1    | 0       | -0.3    | 1       |

4.2 Simulation Results and Discussion

We show the simulation results of the MT method in Figure 2 and BMT method in Figure 3, where each the vertical axes indicates the Mahalanobis distance (MD) and each the horizontal axes indicates the time period. Each of the dotted lines represents the threshold of the each method. Note that in Figure 2 and Figure 3, box plots are different from usual box plots. In these figures, the upper whisker represents 90 percentile and the lower whisker represents 10 percentile.

First, we focus on the results of the MT method in Figure 2. The values of the MD in the MT method are low and flat in periods 1-10. However, the MD increases in periods 11–30, when the population of the test data changes. In addition, the MD is flat in periods 11–20 and 21-30.

Next, we focus on the transition curve of the BMT method in Figure 3. The MD of period 11 in the BMT method increases temporarily, and it gradually decreases afterward. At period 11, when the abnormal data are generated, the BMT method detects them, and the MD increases temporarily. Afterward, as the trend of abnormal data continues, the abnormalities gradually fade because of Bayesian learning. Thus, the successive abnormal data are gradually discriminated as the normal data and the MD becomes smaller. Moreover, BMT method detects next change point at the period 21. The MD of period 21 in the BMT method increases temporarily, but it also gradually decreases afterward. These show that learning by Bayesian inference goes well, and we can confirm the BMT method works.

Then, we show the abnormality determination rate in each period in Figure 4. The left dotted bar graph represents the abnormality determination rate of the BMT method. The right hatched bar graph represents the abnormality determination rate of the MT method.

From results of time period 1, abnormality determination rates of the BMT method and the MT method are around 10 percent. This shows that type I error rates of the BMT method and the MT method are equal. Therefore, we are able to compare the detection capability of MT method and the BMT method at period 11 and 21. The BMT method detects change points which are periods 11 and 21 more accurately than MT method. From this result indicates that type II error of the BMT method is smaller than that of the MT method.
Figure 2: Simulation results of the MT method

Figure 3: Simulation results of the BMT method
5. Data Analysis Using the BMT Method

In this section, we analyze real data. In subsection 5.1, we explain the details of the data and describe the data analysis settings. In subsection 5.2, we show the results of the data analysis and discuss the results.

5.1 Data Background

We obtained data from the Japan National Tax Agency (JNTA) (2015) for the annual average alcoholic beverage consumption per capita by prefecture in Japan for 2001–2013. Okinawa is not included in this data set, and so there are data for 46 prefectures. The observation items used as explanatory variables are shown in Table 5 and the part of the data for 2001 are given in Table 6 as an example.

Table 5: Observation items used as explanatory variables

| Explanatory variable No. | Observed items          |
|--------------------------|-------------------------|
| 1                        | Shochu                  |
| 2                        | Beer and low-malt beer  |
| 3                        | Fruit liqueur           |
| 4                        | Liqueur                 |

Table 6: Data example (part of the data for 2002)

|                | Shochu | Beer and low-malt beer | Fruit liqueur | Liqueur |
|----------------|--------|------------------------|---------------|--------|
| Hokkaido       | 10.4   | 71.2                   | 3.2           | 3.3    |
| Aomori         | 11     | 71.8                   | 2.1           | 4.2    |
| Iwate          | 9.8    | 64                     | 2.1           | 3.9    |
| Miyagi         | 8      | 63.4                   | 2.4           | 4.2    |
Malt-free beer is an alcoholic beverage that is similar to conventional beer, although it is made using different raw materials and a different manufacturing process from beer and low-malt beer. Malt-free beer is classified in the category of alcoholic beverages that includes liqueurs and other brews rather than beer.

Malt-free beer was developed by various beverage companies beginning in 2005 to avoid high taxes on beer; malt-free beer was then in high demand because of its low price. However, in 2007, the tax on malt-free beer was increased, resulting in a rise in price. Nevertheless, malt-free beer has enjoyed long-standing popularity in Japan. In 2010, imports increased, and company X (the beverage company in Japan) developed a private brand of malt-free beer.

We show a summary of above-mentioned important events in the history of the beer market in Table 7. We investigate whether the MD becomes large or not in 2007 and 2010 by changing the demands on the basis of the events listed in Table 7.

| Year | Event |
|------|-------|
| 2007 | Increase in the price of malt-free beer with revision of the liquor tax law |
| 2010 | Increase in imports |
|      | Private brand created by the company X becomes popular |

5.2 Analysis Assumptions and Results

In this analysis, we use data for 8 years for Kagoshima, Hyogo, Yamanashi, Tokyo and Shimane prefectures to test the targets. We choose these prefectures for the following reasons. Kagoshima consumes a great deal of Shochu, Hyogo consumes a great deal of liqueur, Yamanashi consumes a great deal of fruit liquor, and Tokyo consumes a great deal of beer. We also choose Shimane to test the targets because Shimane has no outstanding characteristics about explanatory valuables. The purpose of this data analysis is to analyze the transition curve of the MD of each prefecture. We set the unit space of the BMT method as follows. In the BMT method, the initial unit space is composed of the data for all the prefectures for 2001-2005. However, in the case of testing 2007, for example, we calculate the MD using the Bayes estimator, which is obtained using the data for all the prefectures for 2001-2005 and the data for the target prefectures for 2001-2006.

It is also necessary to set the initial value of each parameter for the BMT method. We use the initial value of $\alpha$ for the mean of the data for all prefectures for 2001-2005. Moreover, we set the initial value of $W$ as the covariance matrix of the unit space. In addition, we set $\beta = 4$, $\nu = 5$.

Due to limitations of space, we show results about Kagoshima, Shimane and Tokyo prefecture. We show the transition curve of the MD of above prefectures in Figures 5-7. In addition, the simulated threshold values for each year for the BMT method are shown in Figures 5-7, with 80 percentile obtained from random numbers generated from non-central chi-square distributions whose parameters are calculated for each year.

In Figure 5, the MD is large at 2010 and exceeds a threshold. From this result, the BMT method detects change point in 2010 about Kagoshima. In Figure 5, the MD is also large at 2007. From these results, Kagoshima was affected by events of the malt-free beer. Although Kagoshima consumes a great deal of Shochu, supply and demand of Shochu in Kagoshima had changed under the influence of the events of malt-free beer.

In Figure 6, the BMT method detects change point in 2010 about Shimane but MD is not large at 2007. These results indicate that when the malt-free beer price had risen in 2007, the BMT method cannot detect the change point in 2007 because the malt-free beer has enjoyed long-standing popularity in Japan. However, the BMT method detects change point in 2010. This result indicates that increase in imports and private brand created by the company X made the consumption of the malt free beer rise so that the demand structure of alcoholic beverage were changed in 2010.

In Figure 7, the BMT method detects change point in 2011 about Tokyo and MD is not large at 2007. In Tokyo, beer had been more popular than malt-free beer previously so that Tokyo was not affected so much when tax of malt-free beer increased in 2007. Afterwards, malt-free beer gradually became popular and took demand of beer in Tokyo because of events in 2010.
Figure 5: Transition curve of the $MD$ for Kagoshima

Figure 6: Transition curve of the $MD$ for Shimane
6. Conclusions and Future Studies

In this paper, we proposed the BMT method, which combines the MT method and Bayesian inference in order to detect change points in panel data when sample size is small. The results of our analysis show that the $MD$ becomes larger at the change point, enabling us to prove the validity of the BMT method. In addition, we analyzed JNTA (2015) data using the BMT method. In these data, there are two change points in the consumption structure of beer and liqueur. According to the results of the data analysis, increase in imports and private brand created by company X made the consumption of the malt free beer rise so that the demand structure of alcoholic beverage was changed in 2010.

We mention three directions for future study. The first is to increase the number of types of panel data used for the real data analysis to show the versatility of the BMT method. The second is to extend the BMT method by grouping. In this paper, we used prefectures with particular characteristics about certain items to test the target. However, it would be beneficial to increase the sample size and show the generality of the analysis result by grouping. The third is the verification of the BMT method using multi-variables time-series data.

References:

Cheng, H. (2007), *Method of micro-econometrics: Panel data analysis*, Iwanamisyoten. (In Japanese)
Gnanadesikan, R. and Kettenring, J. R. (1972), “Robust estimates, residuals, and outlier detection with multiresponse data”, *Biometrics*, Vol. 28, pp.81-124.
Hawkins, D. M. (1981), “A new test for multivariate normality and homoscedasticity”, *Technometrics*, Vol.23, pp. 37-44.
Iizuka, Y. (1982), “Bayesian approach to setting the individual normal range of serum chemical components”, *Japanese Journal of Applied Statistic*, Vol. 11, No.3, pp.113-124. (In Japanese)
Japan National Tax Agency (2015), "Liquor sales per adult (consumption) quantity table"
https://www.nta.go.jp/shiraberu/senmonjoho/sake/shiori-gaikyo/shiori/01.html (View Date: 2015.01.30).
Miyakawa, M., Tanaka, K., Iwasawa, T. and Nakanishi, H. (2007), “The error rate in the Mahalanobis-Taguchi System”, *Quality*, Vol. 37, No.1, pp.101-106. (In Japanese)
Radhakrishnan, R. (1984), “Estimating Mahalanobis’s distance using Bayesian analysis”, *Communications in Statistics-Theory and Methods*, 13(20), pp. 2583-2600.
Taguchi, G. (2002), *Technological development in the MT System*, Japan Standards Association. (In Japanese)
Tatebayashi, K., Tezima, S. and Hasegawa, R. (2008), *Introduction of the MT system*, JUSE Press Ltd. (In Japanese)
Tracy, N. D., Young, J. C. and Mason, R. L. (1992), “Multivariate control charts for individual observations”,

[DOI : 10.17929/tqs.2.36] Copyright © 2016 Journal of the Japanese Society for Quality Control. All rights reserved.
Acknowledgement:

We would like to thank the anonymous referees for their valuable comments. This work was partly supported by JSPS Grants-in-Aid for Scientific Research Grant Number 15K00056

Author’s biographical notes

Takumi Enomoto is a graduate student in the Department of Industrial and Management System Engineering of the Graduate School of Creative Science and Engineering at Waseda University.

Yasushi Nagata is a professor in the Department of Industrial and Management System Engineering of the School of Creative Science and Engineering at Waseda University.

[DOI : 10.17929/tqs.2.36]

Received March 3, 2015
Revised November 27, 2015
Accepted April 7, 2016