Electronic Journal of Statistics
Vol. 8 (2014) 1825–1834
ISSN: 1935-7524
DOI: 10.1214/14-EJS937B

Analysis of juggling data: Registering data to principal components to explain amplitude variation*

Dominik Poss†

Department of Economics
Bonn Graduate School of Economics
Universität Bonn
Adenauerallee 24-26, D-53113
e-mail: dposs@uni-bonn.de

and

Heiko Wagner†

Department of Economics
Institut für Finanzmarktökonomik und Statistik
Universität Bonn
Adenauerallee 24-26, D-53113
e-mail: heikowagner@uni-bonn.de

Abstract: The paper considers an analysis of the juggling dataset based on registration. An elementary landmark registration is used to extract the juggling cycles from the data. The resulting cycles are then registered to functional principal components. After the registration step the paper then lays its focus on a functional principal component analysis to explain the amplitude variation of the cycles. More results about the behavior of the juggler's movements of the hand during the juggling trials are obtained by a further investigation of the principal scores.

Keywords and phrases: Continuous registration, functional principal components analysis, landmark registration, phase variation, time warping, warping function.

Received August 2013.

1. Introduction

Functional Principal Component Analysis (FPCA) approximates a sample curve $f(t)$ as a linear combination of orthogonal basis functions $\gamma_j(t)$ with coefficients θ_j:

$$f(t) \approx \sum_{j=1}^{L} \gamma_j(t)\theta_j. \tag{1}$$

*Main article 10.1214/14-EJS937.

†We thank the MBI (The Ohio State University, Columbus, OH 43210) for their support at the CTW: "Statistics of Time Warpings and Phase Variations". The financial support by the Deutsche Forschungsgemeinschaft (DFG) through the Bonn Graduate School of Economics (BGSE) and DFG Sachbeihilfe KN 567/3-1 is greatly acknowledged.

The principal components γ_j have the best basis property: for any fixed number L of orthogonal basis functions, the expected total squared lose is minimized. The choice of L is up to the operator, depending what accuracy is needed. It is often possible to describe the essential parts of the variations of functional data by looking only at a usually very small set of principal components and the corresponding principal scores θ_j.

However, if the curves have phase variation, even the most elementary tools of any data analysis like the pointwise mean or variance will not be able to describe the data adequately [3]. In such a case not only are more principal components needed to describe the same amount of variation in the data, but also further analysis based on principal components will become more difficult to interpret. In order to analyze the juggling data, we use a registration procedure introduced by [1] in which the principal components are the features which are aligned. The juggling data is a nice application, because the data set contains many problems that have to be solved using different strategies.

After registering the data in Section 2, we perform a FPCA on the individual juggling cycles in Section 3. In Section 4 we examine the evolution of the scores of the juggling cycles over the trials where we additionally take the information from the warping functions into account. Section 5 summarizes our findings.

2. Registering the juggling data

During our analysis we are especially interested in the juggling cycles. We will use the following notation: for $t \in [0, 1]$ let $f(t) = (f_x(t), f_y(t), f_z(t))$ be the spatial coordinates of a typical juggling cycle, $\mu(t) = \mathbb{E}(f(t))$ their structural mean and $\gamma_j(t) = (\gamma_{x,j}(t), \gamma_{y,j}(t), \gamma_{z,j}(t))$ be a typical principal component. We refer to chapter 8.5 of [3] for an instruction on how to calculate the principal components in our multivariate case in practice. Referred to [2], a juggling cycle is observed on the "clock time scale" which is the "juggling time" t transformed by a warping function h. As usual, we assume h to be an element of the space $\mathcal{H}$ of strictly increasing continuous functions. We hence observe

$$f[h(t)] = \mu[h(t)] + \sum_{j=1}^{\infty} \gamma_j[h(t)]\theta_j, \tag{2}$$

where $\theta_j = \int_0^1 \gamma_{x,j}(u) f_x(u) + \gamma_{y,j}(u) f_y(u) + \gamma_{z,j}(u) f_z(u)\, du$.

Note that by stating equation (2), we met the natural assumption that time and therefore also the warping function has to be the same in all three directions by introducing a common h function for all three spatial dimensions. In contrast to [2] where the tangential velocity function is used to avoid the problem of facing three spatial dimensions at once, we will work in the original three dimensional coordinate system. By doing so we hope to find effects which are only observable within the raw data. We approach the registration of the cycles with a two stage procedure by performing what we call "macro" and "micro" warping. By macro warping we mean a very basic registration. The purpose of this registration step

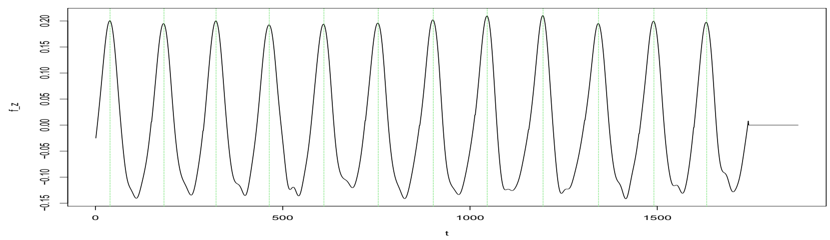

Fig 1. *A random trial along the x direction together with the chosen landmarks.*

is to normalize the overall juggling speed such that we can properly extract the cycles from each trial. We adjusted the data for the different numbers of cycles per trial by trimming each trial down to the first 10 juggling cycles. In order to preserve as much information of the cycles as possible for further analysis, we chose the simplest possible landmark registration which consists only of one landmark per cycle located at the local maxima occurring along the z-direction and a linear interpolation of the h function between. Since we only select one landmark per cycle, identifying it can be done very quickly.

The next step is to cut of all cycles at the landmarks such that we end up with a set of data consisting of a total of 100 cycles. This cropping implies that each of the cycles starts when one of the balls leaves the hand of the juggler to go up in the air in a high arc as seen in Figure 1.

During the "micro" step, we register all 100 cycles simultaneously. By doing this we perform a very precise warping on the cycles. This is in fact a more difficult task than the "macro" warping part, because a lot of different features in the cycle curves have to be taken into account. To clarify this point we displayed a random sample of 20 cycles in Figure 2.

It is seen from Figure 2 that the data needs more than just one principal component to be explained accurately. For example, by looking at the first half of this random sample along the x direction (left plot in the figure), we see variation which is obviously not induced by phase variation. Also a closer look at the

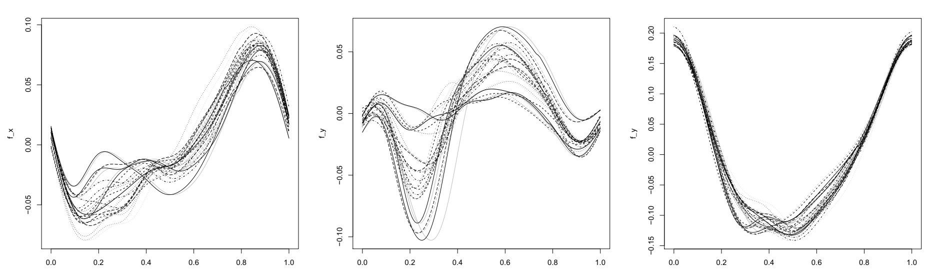

Fig 2. *The figure shows a random sample of* 20 *cycles for the x, y and z direction. Registered curves are displayed black, corresponding unregistered curves grey.*

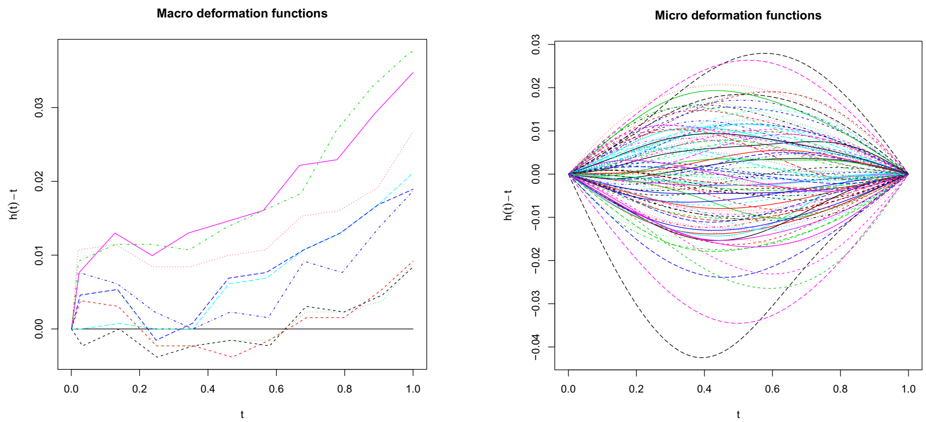

FIG 3. *The deformation functions estimate during the macro- and microwarping.*

middle part in the z direction (right plot) reveals a lot of variation which can not be explained by amplitude variation of a single component. Situations where we encounter more complex amplitude variations are well suited for the registration method presented in [1]. This procedure has another advantage because it allows to control the intensity of the micro warping due to the smoothing parameter in equation (16) of [1].

The method can be easily adapted to the multivariate case. Let D be the derivative operator, then a straightforward modification of equation (15) of [1] now becomes

$$SSE(\widetilde{h}) = \int_0^1 \sum_{k=(x,y,z)} \{f_k(u) - f_k[h^{-1}(u)] - Df_k[h^{-1}(u)]\widetilde{h}(u)\}^2 \, du \qquad (3)$$

which has to be minimized over $\widetilde{h} \in \mathcal{H}$. Finding a common warping function for multivariate data can easily be handled by using (3) for the SSE part occurring in the procedure of [1].

The result of our alignment is shown as the black curves in Figure 2 where we registered the curves to 3 principal components. We observe that after the warping procedure the main features along all directions are well aligned. By looking at the first half of the left plot of Figure 2 one can observe the complexity of the juggling cycles along the x direction: If the cycles would belong to a one dimensional space (i.e. all cycles were random shifts from a mean curve), then all features would have been aligned. However, a more complex model underlies the data along this direction and any attempt to force the data to fit in a simpler model will destroy the intrinsic features of the data; the alleged shift we are observing after the registration is in fact a part of the data. The warping functions for our alignment are displayed in Figure 3 through the deformation functions $h(t) - t$ obtained from the macro and micro step. Note that the deformation functions for the macro step do not end at a value of 0 since we only

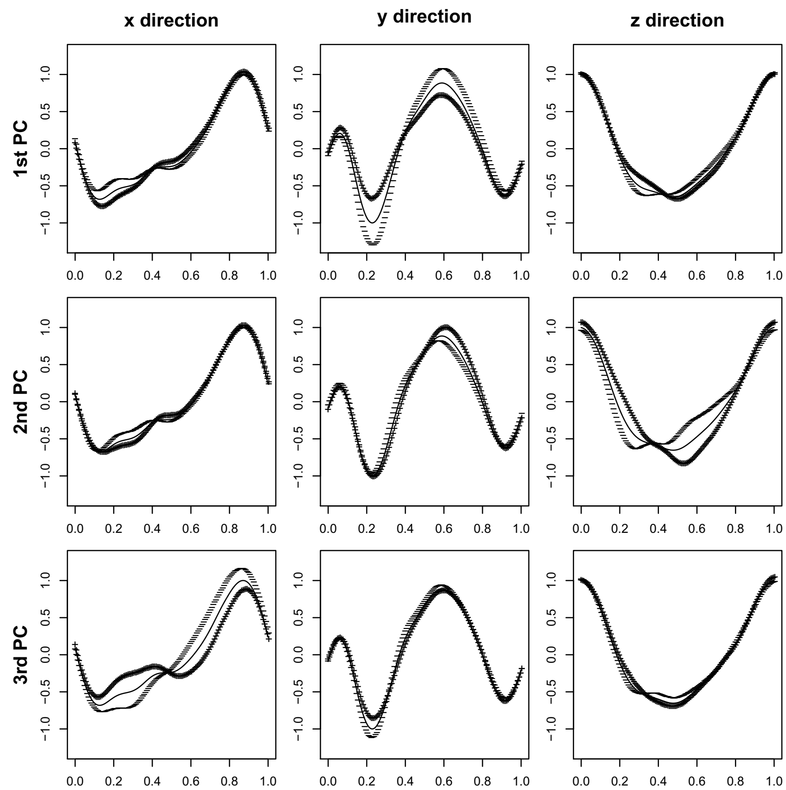

FIG 4. *The Figure shows the effect of adding or subtracting a multiple of each of the principal components to the scaled mean curves. The columns are the spatial directions x,y,z and the rows represent the first, second and third principal component respectively.*

displayed the part of the warping functions corresponding to the first 10 cycles within the trials.

3. Analyzing the principal components

After the preprocessing steps we get suitable data to perform a FPCA. We chose to use three components to represent the data, which explain more than 80 percent of the total variance. The impact of the three principal components on each of the spatial directions of the data is displayed in Figure 4 where we also pictured the effect of adding and subtracting a multiple of each of the principal components to max-normalized mean curves. A closer look at Figure 4 reveals that the first component mainly explains the amplitude variation of the y direction while the second component explains mainly the z direction and the third component the x direction. While the effect of the first component of the movement of the jugglers hand along the x and z direction only accounts for a small shift in the beginning of the movement (the catch phase) it has an

TABLE 1
Variation of the j-th principal component due to the l-th spatial direction

	Spatial direction		
Principal Component	x	y	z
1st	0.117	**0.793**	0.091
2nd	0.053	0.185	**0.762**
3rd	**0.851**	0.100	0.049

important impact for the variation across the y direction. By looking at the impact of the first component along the y direction we can see that, if the ball coming in at low arch during the catch phase is juggled right in front of the juggler, then he will overcompensate for this movement by throwing the next ball from a much greater distance to himself. Such an compensation effect can also be seen for the second component along the z direction and for the the third component along the x direction. While for the y direction the latter two components mainly adjust for the two bumps, which are influenced by the first component, individually.

The importance of the components for the three directions is summarized in Table 1, where we capture the variability in the j-th principal component which is accounted for by the variation in the l-th direction. More formally: for a typical principal component γ we necessarily have $\int_0^1 \gamma_x^2(u)\,du + \int_0^1 \gamma_y^2(u)\,du + \int_0^1 \gamma_z^2(u)\,du = 1$. And hence each of the summands can be interpreted to give the proportion of the variability of the component which is accounted for by the spatial direction. It is seen from the table that the y direction contributes 80% of the variation of the first component while the z and x direction can be accounted for the variation of the second and third component respectively. These values reveal that the directions are somewhat independent in the way that each principal component represents mainly a single direction. These observations where only possible by keeping the data multivariate and not analyzing the tangential velocity function.

4. Analyzing the principal scores

If we perform activities like juggling several times, we expect something like a learning effect to happen. For a juggler this effect could be measured by the behavior of his hands along the directions, i.e. as the juggler gets more and more used to the juggling, one would expect the movements to be more efficient or at least the executions of the movements should become more homogeneous. By performing a FPCA we prepare our data for further statistical analysis which support us to answer such claims. This analysis will be performed on the scores.

Figure 5 shows the evolution of the scores corresponding to the second and third principal component over the ten trials. A typical principal score θ can be modeled as a function depending on trial $k = 1, \ldots, 10$ and number of cycle $i = 1, \ldots, 10$. Figure 5 suggests that a polynomial regression model can capture

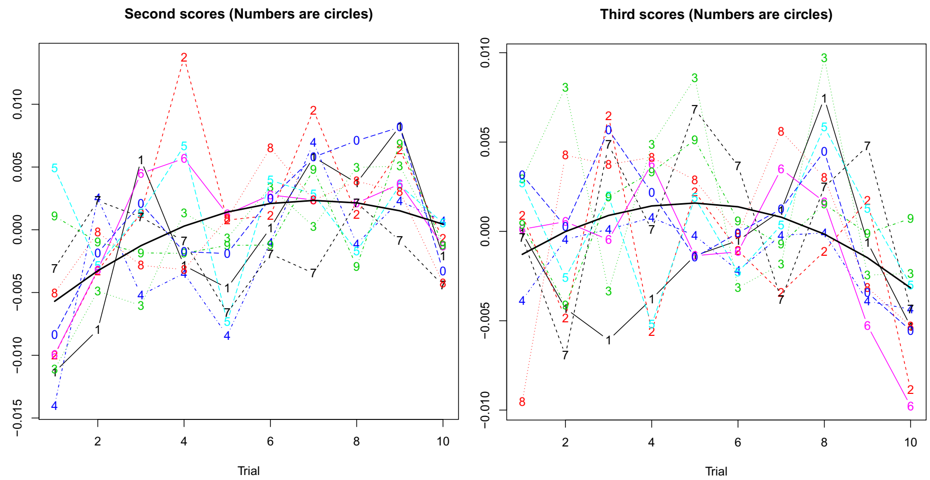

FIG 5. *The figure shows the evolution of the scores for the cycles corresponding to the second and third principal component over the ten trials. The solid line represents the estimated regression function when we impose a quadratic model.*

the main message of the data. i.e. we assume

$$\theta(i, k) = \alpha_0 + \alpha_1 k + \alpha_2 k^2 + \epsilon_i. \tag{4}$$

Table 2 contains the coefficients resulting from this regression. Before we interpret the results, recall that the first component explains mostly the y direction which is on one hand less complex in terms of its variability and on the other hand is less important for a juggler. Indeed, one could imagine a perfect juggling machine which would keep this direction constant such that a juggling cycle could be described by looking solely at the x and z directions. Now, the non-significant coefficients in the first row of Table 2 indicate that the movement across the y direction can not be explained by the trials. This is reasonable as one would expect that an experienced juggler mainly focuses about the movement in the other two directions and any variation of his movement along the y direction from a constant value should be random.

By the significance of the coefficients of the regressions for the scores corresponding to the second and third principal component, we can conclude that

TABLE 2
*Least squares coefficients obtained from a quadratic regression of the scores on the trials. Significance codes are added in parentheses where 0 '***'; 0.001 '**'; 0.01 '*'; 1 ' '*

Scores	Parameter Estimates		
	α_0	α_1	α_2
1st	-0.0040 ( )	0.0014 ( )	-0.0001 ( )
2nd	-0.0086 (***)	0.0031 (***)	-0.0002 (***)
3rd	-0.0029 (*)	0.0018 (**)	-0.0002 (***)

there exists indeed an evolution of the scores over the trials which can essentially be described by our regression. This evolution can be regarded as some kind of a "learning effect". For example, in Figure 5 we can see that the scores will have a small value at the peak of our regression function, implying that in this area the variation of the movement of the jugglers hand is not very high and has to be close to the mean curve. This can be seen as an improvement in his juggling skills. Interestingly, the slope of the regression function decreases at the end. While this effect is subsidiary for the second principal score and could be seen as a nuisance from the simple quadratic model, it is apparent in the evolution of the scores corresponding to the third component.

Recall that the second component mainly quantifies the variation of the jugglers hand movement along the z-direction, which captures the up- and downwards movement of his hand. A negative score in the beginning of the trials indicates that he lunges out too far before throwing the ball up in the air. As the regression function for the scores of the second component approaches values close to zero, the "learning effect" becomes visible: getting used to the juggling in the later trials, he performs almost identical movements along this direction.

If we take a more precise look at the regression function of the scores corresponding to the third component, an interpretation is somewhat more complicated as we experience a significant downward slope at the last trials. Maybe the juggler gets fatigued or the behavior is caused by some kind of a psychological effect, i.e. the concentration of the juggler decreases as he knows that he only has to perform a few more trials and gets more impatient.

Taking a look at the time frame around 0.2–0.5 of the the bottom left panel of Figure 4, we see that a particular small value of the third component implies that his hand for catching the ball coming in from a low arch is comparable moved towards the other hand. Possibly e is learning to simplify the process of catching the ball coming in from low arch. Unfortunately this implies that he has to wind up more in order to throw the ball leaving in high arch.

We were further interested in an analysis of the warping functions themselves which was the reason to perform only a very basic "macro" warping. In this special kind of data set it is not reasonable to assume that the warping function is only a nuisance parameter because the speed of juggling might have an effect on the manner of the juggling.

To check this hypothesis we performed some further analysis on the warping functions. Note that we can not perform a FPCA on the warping functions directly, because we can not guarantee that the resulting curves are still elements of $\mathcal{H}$, i.e. strictly monotonic functions. Instead we pursue the following way out. It is well known from [3] that any function $h \in \mathcal{H}$ can be represented as

$$h(t) = \int_0^t e^{W(u)}\, du,$$

where $W(t) = log[Dh(t)]$ itself is an unrestricted function. In order to analyze the warping functions h appropriately, we can use the unrestricted functions $W(t)$. We approximate $W(t)$ by using the first two principal components which

TABLE 3
The table shows the correlation between the scores corresponding to the first two components of W and the scores corresponding to the first three components of the juggling cycles

Scores of W	Scores of the cycles		
	θ_1	θ_2	θ_3
$\theta_{W,1}$	-0.0120	0.3044	-0.2351
$\theta_{W,2}$	-0.0122	0.0355	0.0013

explain more than 95 Percent of the variations in $W(t)$ and define by $\theta_{W,1}$, $\theta_{W,1}$ a typical scores corresponding to these two components. In Table 3 we computed the correlation between the scores of W and θ.

We can determine that the speed a juggling cycle is performed with has nearly no influence on the first component of a cycle. But this speed does have an effect on the second and third component which explain mostly the x and z direction. Obviously, this effect is occurs mainly through the first component of W.

Another interesting result occurs by computing the correlation between the scores of the principal components of W and and the residuals resulting from the polynomial regression in (4). It reveals a significant amount of correlation between these variables, i.e. a not negligible part of the residuals from (4) can be explained by the juggling speed of the cycles. Moreover, running a regression of the scores of the warping function W on the trials showed no significant coefficient. From this we can conclude that, what we identified as a learning effect, has no significant impact on the warping for a specific cycle. We hence can identify two effects which influence the scores of a juggling cycle. The first is due to learning and the second is a result which is related to the specific warping. The effects are modeled by augmenting equation (4) by

$$\theta(i,k) = \alpha_0 + \alpha_1 k + \alpha_2 k^2 + \beta_1 \theta_{W,1,i} + \beta_2 \theta_{W,2,i} + \epsilon_i, \tag{5}$$

where $\theta_{W,j,i}$ is the score of the i-th cycle corresponding to the j-th principal component of the function W. Estimated coefficients are given in Table 4, from where it can be seen that neither the speed the juggling cycles are performed with, nor the trials have an impact on the movement of the jugglers hand along the y direction. Moreover, it can be seen that there is a connection between the scores of a juggling cycles and the speed of the juggling.

TABLE 4
*The table shows the results from an Regression of the cycle scores on the trial number, squared trial number as well as the scores from W with corresponding coefficients β_1 and β_2. Significance codes are added in parentheses where 0 '***'; 0.001 '**'; 0.01 '*'; 1 ' '*

Scores	Parameter Estimates				
	α_0	α_1	α_2	β_1	β_2
1st	-0.0042 ()	0.0014 ()	-0.0001 ( )	-0.0009 ()	0.0000()
2nd	-0.0081 (***)	0.0030 (***)	-0.0002 (***)	0.0034 (**)	0.0025 ()
3rd	-0.0033 (*)	0.0019 (***)	-0.0002 (***)	-0.0027 (*)	-0.0009 ()

5. Summary

We analyzed the juggling data by combining two registration methods. First we used an elementary landmark registration in order to crop the individual juggling cycles, which were the focus of our analysis. In order to perform a refined warping of the juggling cycles in a second step, we generalized the registration method from [1] to the multivariate nature of the data. We analyze the registered data by performing a FPCA using three principal components where we observed that each of the components essentially quantified the variation across a single spatial direction.

More specific information about the behavior of the juggler is contained in the scores which we studied in dependence on the trials. By doing so, we were able to identify some kind of learning effect over the trials. The movement of the jugglers hand for throwing a ball up in the air levels out over the trials. After applying an alignment procedure one should not forget about the warping functions. Interpreting the warping functions can not only be a very interesting task for itself, but they can contain important additional information which can be helpful to analyze the data.

References

[1] Kneip, A. and Ramsay, J. O. (2008). Combining registration and fitting for functional models. *J. Amer. Statist. Assoc.* **103**, 1155–1165. MR2528838

[2] Ramsay, J. O., Gribble, P. and Kurtek, S. (2014). Description and processing of functional data arising from juggling trajectories. *Electron. J. Statist.* **8** 1811–1816, Special Section on Statistics of Time Warpings and Phase Variations.

[3] Ramsay, J. O. and Silverman, B. W. (2005). *Functional Data Analysis.* 2nd edition. Springer. MR2168993