

The Common Principal Component (CPC) Approach to Functional Time Series (FTS) Models

Farah Yasmeen

Department of Statistics, University of Karachi, Karachi, Pakistan

Email address:

riazfarah@yahoo.com (F. Yasmeen)

To cite this article:

Farah Yasmeen. The Common Principal Component (CPC) Approach to Functional time Series (FTS) Models. *Science Journal of Applied Mathematics and Statistics*. Vol. 3, No. 1, 2015, pp. 22-26. doi: 10.11648/j.sjams.20150301.14

Abstract: The functional time series (FTS) models are used for analyzing, modeling and forecasting age-specific mortality rates. However, the application of these models in presence of two or more groups within similar populations needs some modification. In these cases, it is desirable for the disaggregated forecasts to be *coherent* with the overall forecast. The '*coherent*' forecasts are the non-divergent forecasts of sub-groups within a population. Reference [1] first proposed a coherent functional model based on product and ratios of mortality rates. In this paper, we relate some of the functional time series models to the common principal components (CPC) and partial common principal components (PCPC) models introduced by [2] and provide the methods to estimate these models. We call them common functional principal component (CFPC) models and use them for coherent mortality forecasting. Here, we propose a sequential procedure based on Johansen methodology to estimate the model parameters. We use vector approach and make use of error correction models to forecast the specific time series coefficient for each sub-group.

Keywords: Mortality, Forecast, Coherent Forecasts, Functional Data, Life Expectancy, Sex-Ratio, Cointegration

1. Introduction

Functional time series (FTS) encompasses data in the form of curves that are observed at regular intervals in time. Recently these models are applied for demographic forecasting and breast cancer mortality forecasting (see [3] and [4]).

However, the application of these models in presence of two or more groups within similar populations needs some modification. [1] Introduced a simple and interpretable method of estimating a functional linear model. This model provides coherent mortality forecasts for two or more groups, i.e. forecasts of individual groups which will not diverge in the long run. The coherent functional model is one of several extensions of the basic functional time series model proposed by [3], however, the authors did not discuss the implementation and estimation of these models.

In this paper, I relate some of the functional models described above to the common principal components (CPC) and partial common principal components (PCPC) models introduced by [2], and provide methods for estimating these models. These models will be called common functional principal component (CFPC) models and will be used for coherent mortality forecasting. As described earlier, [1]

defined a functional time series model for the entire aggregate of groups (the product term) and functional time series models for each of the ratio terms of group-specific rates to the product term. The time series coefficients of the product model were then forecast with (possibly non-stationary) ARIMA models and the ratio coefficients with stationary ARMA or ARFIMA models. In this paper, I propose a sequential procedure for estimating the model parameters. My procedure differs from the procedure described in [1] that here I will use a vector approach, and will make use of error correction models for forecasting the specific time series coefficients.

For illustrative purposes, I apply these CFPC models to the all-cause mortality data of males and females in Australia. The interest lies in developing a suite of tools that will provide more accurate forecasts than the existing approaches when applied to real data sets.

2. Common Functional Principal Component (CFPC) Models

The CPC analysis assumes that the space spanned by the eigenvectors is identical across groups, whereas the variances associated with these common principal components vary [2].

The main advantage of the CPC approach relative to the ordinary PCA approach is that it is a way of compressing the high-dimensional data of several groups into a small number of common factors across groups. We need to estimate a smaller number of parameters with CPC than with ordinary PCA applied separately to each group. The CPC framework has the advantage that it is mathematically appealing and empirically parsimonious. For the estimation procedure and the properties of the parameters of CPC models, the reader is referred to [2] and [5].

Common principal component (CPC) models have been considered particularly important in Implied Volatility (IV) dynamics [see 6-7]. Some other references in this area include [8-9]. However, these models have never been used for mortality forecasting. Here, I combine the ideas of functional data analysis and common principal components with time series models. I differ from the above approaches in the way that I use CPC models for the coherent forecasting of mortality rates.

Suppose there are J related functional time series corresponding to age-specific mortality rates of J groups and it is of interest to forecast them jointly rather than independently. Let $m_{t,j}(x)$ denote the mortality rate for age and year t , $t = 1, \dots, n$, for the j th group. We will model the log mortality, $y_{t,j}(x_i) = \log [m_{t,j}(x)]$, and assume that there is an underlying smooth function $g_{t,j}(x)$ that we are observing with error. Thus,

$$y_{t,j}(x_i) = g_{t,j}(x_i) + \sigma_{t,j}(x_i)\varepsilon_{t,j,i} \quad (1)$$

It is desirable to define a functional time series model for the whole aggregate of groups. Since the mean function can vary across groups, the following model termed as CFPC-I can be used:

$$g_{t,j}(x_i) = \mu_j(x) + \sum_{k=1}^K \beta_{t,k} \phi_k(x) + \varepsilon_{t,j}(x) \quad (2)$$

According to [2], the space spanned by the principal components is same in all groups but their variances can vary. Hence, we can define another model with same basis functions but different coefficients

$$g_{t,j}(x_i) = \mu_j(x) + \sum_{k=1}^K \beta_{t,k,j} \phi_k(x) + \varepsilon_{t,j}(x) \quad (3)$$

I will call this the CFPC-II model and it is of considerable importance when the age patterns (basis functions) do not differ significantly among the groups but their time series coefficients can vary across the groups.

The idea of partial common principal components (PCPC) models was also introduced by [2]. A PCPC(p) model, where 'p' denotes the order of the common eigen-vectors, is appropriate when one assumes that in each group there are 'p' joint (common) components, and the remaining 'q' are non-common or specific components. If we relate this idea to demographic models, then this type of model is proposed by [10], who used one common and one group-specific component for each of the group.

Combining these ideas and defining a functional linear model with more than one component, say 'K' common and 'L'

specific components, we have the following functional linear model

$$g_{t,j}(x_i) = \mu_j(x) + \sum_{k=1}^K \beta_{t,k} \phi_k(x) + \sum_{l=1}^L \gamma_{t,j,l} \psi_{j,l}(x) + \varepsilon_{t,j}(x) \quad (4)$$

where each $\gamma_{t,j,l}$ is a stationary time series, but $\beta_{t,k}$ may be non-stationary. This model will be termed as Partial Common Functional Principal Component (PCFPC)(p, q) model.

3. Estimation of Parameters

For estimating the parameters of a CFPC model, it is desirable to obtain a suitable matrix M of the observed log mortality rates on which SVD is to be applied. Then we will get the basis functions and corresponding estimates of the time series coefficient.

Let M be a $p \times n$ matrix with (i, t) th element, i.e.

$$M = \begin{bmatrix} m_{11} & m_{12} & \dots & m_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ m_{p1} & m_{p2} & \dots & m_{pn} \end{bmatrix}$$

For CFPC-I, first we obtain the weighted average of the log mortality rates after smoothing. Let $f_{t,j}(x_i)$ is the value of smoothed log mortality rate at age and time t for the j th group and $p_{t,j}(x_i)$ is the (i, t) th element of, the total number of exposures for the j th population, $j = 1, 2, \dots, J$.

$$P_j(x) = [P_{1,j}(x) P_{2,j}(x) \dots P_{n,j}(x)]$$

where

$$P_{t,j}(x) = \begin{bmatrix} p_{t,j}(x_1) \\ p_{t,j}(x_2) \\ \vdots \\ p_{t,j}(x_p) \end{bmatrix}$$

For CFPC-I, the matrix M can be obtained as

$$m_{i,t} = \frac{\sum_{j=1}^J [f_{t,j}(x_i)] \times p_{t,j}(x_i)}{\sum_{j=1}^J p_{t,j}(x_i)} \quad (5)$$

After obtaining M from the elements given in equation (5), we apply SVD to get the common basis functions and coefficients for model given in (2).

For CFPC-II, the matrix M will be

$$M = [M_1 | M_2 | \dots | M_J]$$

Where $M_k = (m_{(it)k})_{p \times n}$ and

$$m_{(i,t)j} = \frac{[f_{t,j}(x_i) - \mu_j(x_j)] \times p_{t,j}(x_i)}{\sum_{j=1}^J p_{t,j}(x_i)} \quad (6)$$

Applying SVD to $M = \psi \Lambda \psi'$ gives where ψ is of order $p \times p$ and $B = M' \psi$ is of order $Jn \times p$. The common basis function $\psi_j(x_i)$ is the (i, j) th element of ψ . The specific coefficients can be obtained from $B = M' \psi$.

Finally, for the PCFPC model, we first apply SVD to the matrix M for CFPC-I given in equation (5) so that, the terms are estimated.

$$\mu_j(x) + \sum_{k=1}^K \beta_{t,k} \phi_k(x)$$

Next, the problem is to estimate the specific coefficients and basis functions for each group. For this, we apply SVD to the residual matrix for the j th group defined as

$$R_j = [R_{1,j}(x) R_{2,j}(x) \dots R_{n,j}(x)]$$

where

$$R_{t,j}(x) = f_{t,j}(x) - [\mu_j(x) + \sum_{k=1}^K \beta_{t,k} \phi_k(x)] \quad (7)$$

Applying SVD to R_j gives $R_j = \Lambda_j \psi_j B_j$. The specific basis function $\psi_{j,l}(x)$ is the (i, l) th element of Λ_j whereas the coefficient $\gamma_{t,j,l}$ can be obtained from $R_j' \Lambda_j$, as described for the common components.

4. Forecasting the Coefficients

4.1. CFPC Models

To obtain the forecast value of mortality rates from CFPC model, we first obtain the forecast for each of the common (or non-common) components in the model.

Let $\hat{\beta}_{n,h,j}$ and $\hat{\beta}_{n,h,j,k}$ denote the h -step ahead forecast of $\beta_{n+h,j}$ and $\beta_{n+h,j,k}$ respectively. Also, let $\hat{g}_{n,h,j}(x)$ denote the h -step ahead forecast of $g_{n+h/n,j}(x)$ for $j = 1, 2, \dots, J$. Then the forecasts for CFPC models are given by are estimated.

$$\hat{g}_{n+h/n,j}(x) = \hat{\mu}_j(x) + \sum_{k=1}^K \hat{\beta}_{n,h,k} \hat{\phi}_k(x) \quad (8a)$$

$$\hat{g}_{n+h/n,j}(x) = \hat{\mu}_j(x) + \sum_{k=1}^K \hat{\beta}_{n,h,j,k} \hat{\phi}_k(x) \quad (8b)$$

$$\hat{g}_{n+h/n,j}(x) = \hat{\mu}_j(x) + \sum_{k=1}^K \hat{\beta}_{n,h,k} \hat{\phi}_k(x) + \sum_{l=1}^L \hat{\gamma}_{n,h,j,l} \hat{\psi}_{j,l}(x) \quad (8c)$$

where $\hat{\mu}_j(x)$ are the estimated values of $\mu_j(x)$, $j = 1, 2, \dots, J$ and $\hat{\phi}_k(x)$ are the estimated common basis functions. Here $\hat{\gamma}_{n,h,j,l}$ denote the h -step ahead forecast of the specific coefficient for PCFPC model given in equation (4).

To obtain the forecasts of the mortality rates from the CFPC model, I first obtain the forecast for each of the common (or non-common) components in the model. Note that the basis functions for the first two CFPC models are the same. One can use univariate time series models for the coefficients of the models given in equations (2) and (3) and the common components of the model given in equation (4).

The common coefficients $\beta_{t,k}$, $k=1,2,\dots,K$, all are independent, as they are all obtained from a principal component decomposition. Hence, there is no need to apply multivariate time series models; rather, univariate models, including ARIMA models, exponential smoothing statespace models [11], random walk with drift and other univariate time series models can be used. Here I use ARIMA models, with an automatic algorithm based on [12] for choosing the values of

the parameters.

4.2. PCFPC Models

In this section, I will discuss some important issues with using partial common functional principal component (PCFPC) models. They include forecasting the specific coefficients of these models and the cointegration in these components.

4.2.1. Forecasting the Specific Coefficients

For forecasting the specific coefficients of a PCFPC model, I use vector autoregressive (VAR) models with co-integration. Using a VAR(p) model, it is necessary for all of the variables included to have the same order of integration. We consider the following two cases:

1. If all the variables are stationary (or $I(0)$), we can use the standard case, i.e., a VAR model in levels is appropriate. The appropriate lag length is selected according to some model specification criteria e.g., AIC or BIC.
2. If the variables are non-stationary $I(d)$ with $d \geq 1$, then the simplest approach is to difference each of them 'd' times and then apply the standard VAR model in the usual manner; but in the case of *cointegration*, differencing can result in a large positive autocorrelation in residuals. Hence, we make use of cointegration and vector error correction models (VECM).

4.2.2. Cointegration in Specific Coefficients

[13] first pointed out that a linear combination of two or more non-stationary series may be 'stationary'. The stationary combination may be interpreted as 'cointegration', or an 'equilibrium relationship' between the variables. If the variables are not cointegrated, then one variable might drift above or below the other variable in the long run.

4.2.3. Cointegration when Cointegrating Vector is Pre-Specified

For the two-sex data, suppose the males and females time series coefficients are co-integrated with pre-specified vector $\beta = (1, -1)'$ i.e the long-run relationship between the male series $y_{t,m}$ and female series $y_{t,f}$ is $y_{t,m} = y_{t,f}$. We can test it for its co-integration using a unit root test for the difference series $d = y_{t,m} - y_{t,f}$, using Augmented Dickey Fuller ADF test. We have to test

$$H_0 = \beta' y_t = y_{t,m} - y_{t,f} \sim I(1) \text{ (no co integration)}$$

$$H_1 = \beta' y_t = y_{t,m} - y_{t,f} \sim I(0) \text{ (co integration)}$$

Co-integration is found if the unit-root test rejects the no-cointegration null (see [14]). For forecasting the specific components of a PCFPC model, we use Johansen Methodology [15], based on the vector error correction model (VECM).

4.2.4. Johansen Methodology

Johansen methodology starts with a VAR model. A VAR process with p lags is defined as

$$y_t = \Phi_0 + \Phi_1 y_{t-1} + \Phi_2 y_{t-2} + \dots + \Phi_p y_{t-p} + \varepsilon_t (9)$$

where Φ_i 's are $(K \times K)$ coefficient matrices for $i = 1, 2, \dots, p$ and is K -dimensional white noise i.e. $\varepsilon_t \sim N(0, \Sigma)$.

One can obtain the following form of vector error correction models (VECM):

$$\Delta y_t = \Phi_0 + \Pi y_{t-1} + \Gamma_1 \Delta y_{t-1} + \Gamma_2 \Delta y_{t-2} \dots + \Gamma_p \Delta y_{t-p} + \varepsilon_t (10)$$

with

$$\Gamma_i = - \sum_{j=i+1}^p \Phi_j \text{ for } i = 1, 2, \dots, p-1$$

and

$$\Pi = (I - \sum_{i=1}^p \Phi_i)$$

Here matrix Π is called the long run impact matrix and matrices Γ_i 's are called short-run impact matrices. The matrix Πy_{t-1} represents the co-integration relation.

4.2.5. Forecasting from VECM

To forecast from VECM, we first convert it into appropriate VAR model. e.g. in bivariate case

$$\Delta y_t = \Phi_0 + \Pi y_{t-1} + \Gamma_1 \Delta y_{t-1} + \varepsilon_t (11)$$

After estimating the matrices Π and Γ_1 , we can easily convert this model into VAR(2) using

$$\Phi_2 = -\Gamma_1 (12a)$$

$$\Phi_1 = \Gamma_1 - \Pi + I_2 (12b)$$

Once the parameters's are estimated, forecasts can be obtained in the usual manner as forecast from a VAR model.

4.2.6. Choosing among the Best CFPC Model

We define

K: the number of common components and

L: the number of non-common or specific components

$PV_{1,k}$: Percentage variance explained by the first K common components;

$PV_{2,j,l}$: Percentage variance explained by the first L non-common components in jth group, here

$$PV_{1,k} = \frac{\lambda_k}{\sum_{k=1}^K \lambda_k} PV_{2,j,l} = \frac{\lambda_{j,l}}{\sum_{l=1}^L \lambda_{j,l}} (13)$$

Where λ_k and $\lambda_{j,l}$ are the eigenroots corresponding to the kth common and lth non-common factor for the jth group respectively. Also, one can define the cumulative variation as follows:

CP_1 : Total variation explained by the first K common components;

$CP_{2,j}$: Total variation explained by the first L non-common components in jth group; with

$$CP_1 = \frac{\sum_{k=1}^K \lambda_k}{\sum_{k=1}^K \lambda_k} CP_{2,j} = \frac{\sum_{l=1}^L \lambda_{j,l}}{\sum_{l=1}^L \lambda_{j,l}} (14)$$

4.2.7. Coefficient of Explanation

To measure the performance of a CFPC model, we define

$$CE_j = 1 - \frac{\sum_{t=1}^n \int_x [y_{t,j}(x) - \hat{g}_{t,j}(x)]^2 dx}{\sum_{t=1}^n \int_x [y_{t,j}(x) - \bar{\mu}_j(x)]^2 dx} (15)$$

where $y_{t,j}(x)$ are the observed value of log mortality rate and $\hat{g}_{t,j}(x)$ are fitted values obtained from a CFPC model.

5. Empirical Application

In this section, we will illustrate the procedure of fitting and forecasting through CFPC models using an application to the age-sex specific data of Australia. The data are obtained from [16].

Table 1 shows the coefficient of explanation for different CFPC models, each with six common components. CFPC-I gives considerably higher values of and the model is explaining about 96.82% variability in females and 96.38% in males.

The values of for CFPC-II are slightly lower in both groups (74.4% in female and 73.88% in males).

Table 1. Coefficient of explanation for Australian Sex Data.

Model	Female	Male
CFPC-I (6)	0.9682	0.9638
CFPC-II (6)	0.7439	0.7388
PCFPC(6,6)	0.9931	0.9904
Independent	0.9886	0.9857

If we apply PCFPC model, we will get the best fitted model as now it is capturing about 99% variability in both groups. We use $K=6$ and $L=6$ and the value of CP_1 is 99.4% for common components and the values of $CP_{2,j}$ are 83.2% and 79.9% for females and males respectively, which are sufficiently high.

Table 2 represents the p-value of Augmented Dickey Fuller (ADF) unit root test for the specific coefficients for males and females. For the first two coefficients, we are unable to reject the hypothesis about the presence of unit-root. The next step is to determine whether the individual series are cointegrated. For this, we can use Engle-Granger cointegration test [13] based on the residuals from OLS regression of one variable (say $y_{1,t}$) on other variable ($y_{2,t}$).

$$y_M = \alpha_M + \beta_M y_F + z_{1t} (16a)$$

$$y_F = \alpha_F + \beta_F y_M + z_{2t} (16b)$$

Table 2. p-value for the Augmented Dickey-Fuller unit root test for specific coefficient $\gamma_{t,j,l}$. Specific Coefficients.

Series	1	2	3	4	5	6
Male	0.6099	0.9566	0.0100	0.0118	0.0317	0.0259
Female	0.5156	0.9004	0.0308	0.0215	0.0265	0.0154

Table 3. *p*-values for augmented Dickey Fuller unit root test for Engle-Granger Cointegration Method.

Coefficient No.	1	2	3	4	5	6
Residuals (male)	0.017	0.020	0.025	0.018	0.047	0.045
Residuals (female)	0.010	0.039	0.010	0.010	0.031	0.028

Table 3 represents the *p*-value of ADF test for the static residuals $z_{1,t}$ and $z_{2,t}$. It is interesting that for both males and females, the *p*-values are smaller than 0.05. After confirming the cointegration between the first two coefficients of males and females, the next step is to determine the nature of cointegration relation. For bivariate time series, the cointegrating vector β will be $\beta = (1, -\beta^*)'$

In our case, suppose that the demographers assume that there is a long-run equilibrium in the male and female series with pre-specified cointegrating vector $\beta = (1, -1)'$. For this, we first check the presence of unit-root in the differences of observed series. Table 4 shows the *p*-values of Augmented Dickey Fuller test for the difference of specific coefficients

$$d = y_m - y_f$$

The results in tables (2) and (4) confirm that the original series were I(1) and their differences are stationary.

6. Conclusion

In this paper, we introduced a new class of functional linear models for coherent mortality forecasting. We developed the methods for estimating their parameters and forecasting from these models.

It is found that the specific time series coefficients among different subgroups are highly correlated; hence we used vector autoregressive (VAR) and vector error correction models (VECM) to forecast them. For the purpose of illustration, the models are applied to the two-sex data of Australia. It is found that all new methods work well and mortality rates and life expectancy can be forecast in a coherent way. Also, PCFPC model provides some extra information about the age-groups that are most responsible for the difference. We found that the difference of male and female series $d = y_m - y_f$ is a stationary process. It means that there is long run equilibrium among male and female coefficients.

References

- [1] Hyndman, R.J., Booth, H. and Yasmeen, F. (2013). Coherent mortality forecasting: the product-ratio method with functional time series models. *Demography* 50(1), 261-283.

- [2] Flury, B. (1988). *Common Principal Components & Related Multivariate Models*, Wiley Series in Probability and Mathematical Statistics.
- [3] Hyndman, R.J. and Ullah, M.S. (2007). Robust forecasting of mortality and fertility rates: a functional data approach. *Computational Statistics and Data Analysis*, 51, 4942-4956.
- [4] Yasmeen, F., Hyndman, R.J. and Erbas, B. (2010). Forecasting age-related changes in breast cancer mortality among white and black US women: A functional data approach. *Cancer Epidemiology*, 34(5), 542-549.
- [5] Flury, B. D., Nel, D. G. & Pienaar, I. (1995), 'Simultaneous detection of shift in means and variances', *Journal of the American Statistical Association* 90(432), 1474-1481.
- [6] Fengler, M., Hardle, W. & Mammen, E. (2005), 'A dynamic semiparametric factor model for implied volatility string dynamics'. SFB 649 Discussion Paper No. 2005-20, SFB 649 Humboldt-Universität zu Berlin.
- [7] Fengler, M. R., Hardle, W. K. & Villa, C. (2003), 'The dynamics of implied volatilities: A common principal component approach', *Review of Derivatives Research* 6, 179-202.
- [8] Trendafilov (2010). Stepwise estimation of common principal components. *Computational Statistics & Data Analysis*, 54(12), 3446-3457. doi:10.1016/j.csda.2010.03.010
- [9] Neuenschwander B. E and Flury, B.D (2000). Common Principal Components for Dependent Random Vectors, *Journal of Multivariate Analysis*, 75(2), 163-183
- [10] Li, N. & Lee, R. (2005), Coherent mortality forecasts for a group of populations: An extension of the Lee-Carter method, *Demography* 42(3), 575-594.
- [11] Hyndman, R. J., Koehler, A. B., Ord, J. K. & Snyder, R. D. (2008), *Forecasting with exponential smoothing: the state space approach*, Springer-Verlag, Berlin. URL: www.exponentialsmoothing.net
- [12] Hyndman, R. & Khandakar, Y. (2008), 'Automatic time series forecasting: the forecast package for R', *Journal of Statistical software* 27(3) URL: <http://www.jstatsoft.org/v27/i03>
- [13] Engle, R. and Granger, C. (1987). Co-integration and error correction: representation, estimation and testing. *Econometrica*, 55, 251-276.
- [14] Hamilton, J.D. (1994). *Time Series Analysis*, Princeton University Press, Princeton, NJ.
- [15] Johansen, S. (1988). Statistical analysis of cointegration vectors. *Journal of Economics Dynamics and Control*, 12, 231-254.
- [16] Hyndman, R.J. (2008). *adbb: Australian Demographic Data Bank*. R package version 3.222. URL: robjhyndman.com/software/adbb