Change point tests in functional factor models with application to yield curves

Piotr Kokoszka
Department of Statistics, Colorado State University

Joint work with

Lajos Horváth
University of Utah

Gabriel Young
Columbia University

Patrick Bardsley
University of Texas
Scalar change point model

\[ X_i = \mu_i + u_i, \quad 1 \leq i \leq N, \quad \mathbb{E}u_i = 0. \]

Null hypothesis:

\[ H_0 : \mu_1 = \mu_2 = \ldots = \mu_N. \]

\( H_A \): the means change in some way: abruptly, gradually, epidemically ...

It is typically assumed that the errors \( u_i \) are homoskedastic, the series \( u_i \) is stationary.

There has been limited work for nonstationary \( u_i \): Cavaliere, Taylor, Dalla, Giraitis, Phillips ...
CUSUM process:

\[ Z_N(t) = N^{-1/2} \left( \sum_{\ell=1}^{\lfloor Nt \rfloor} X_\ell - \frac{\lfloor Nt \rfloor}{N} \sum_{\ell=1}^{N} X_\ell \right) . \]

Functionals of this process are used as tests statistics. If the \( u_i \) are iid, \( Z_N \) typically leads to optimal tests, it can be connected to the likelihood method.

There has been recent work for the \( u_i \) which are dependent and exhibit change/points or irregular trends in variability. Change points or trends in the variability of the \( u_i \) can confound change points in the mean of the \( X_i \).
Assumptions on the $u_i$

$$u_i = u_{i,N} = a(i/N)e_i$$

$[0, 1] \ni t \mapsto a(t)$ has regular variation.

$$N^{-1/2} \sum_{\ell=1}^{[Nt]} e_{\ell} \overset{{\cal D}[0,1]}{\rightarrow} \sigma W(t).$$

(no mixing needed). Set

$$b(t) = \sigma^2 \int_0^t a^2(u)du.$$ 

Then

$$N^{-1/2} \sum_{\ell=1}^{[Nt]} u_{\ell} \overset{{\cal D}[0,1]}{\rightarrow} W(b(t)).$$

Main difficulty: the function $b(\cdot)$ is unknown.
Under $H_0$,

$$Z_N(t) \xrightarrow{\mathcal{D}[0,1]} \Gamma(t), \text{ where } \Gamma(t) = W(b(t)) - tW(b(1)).$$

$$C(t, s) = E[\Gamma(t)\Gamma(s)] = b(t \wedge s) - tb(s) - sb(t) + tsb(1).$$

$b(t)$ is estimated by the LRV of $X_i, i \leq [Nt]$. 

$\lambda_i$ eigenvalues of $C(\cdot, \cdot)$

$$\int_0^1 Z_N^2(t)dt \xrightarrow{\mathcal{D}} \int_0^1 \Gamma^2(t)dt. \xrightarrow{\mathcal{D}} \sum_{i=1}^{\infty} \lambda_i \xi_i^2.$$ 

Weighted functionals, e.g. Anderson–Darling, are possible.
i - day (or month)

t\_j - time to maturity (1 month, 3 months, ..., 30 years)
Fractions of bonds with continuous maturities are traded.
Our theory can be in discrete or continuous time \( t \).

\[ X_i(t) \] - yield of a bond bought on day \( i \) and held for time \( t\_j \).
For the theory, we can use, the yield curves, \( X_i(t), t \in [0, T] \).
Yield curves

US yield curves over five business days around the Lehman Brothers bankruptcy filing on Sep 15 2008.

Crisis of 2008

Yield curves over five business days around the Lehman Brothers bankruptcy filing on Sep 15 2008.
Yield curves

US yield curves over $N = 100$ business days; the central time point corresponds to the Lehman Brothers collapse.
Factor models

\[ X_i(t_j) = \sum_{k=1}^{K} \beta_{i,k} f_k(t_j) + \varepsilon_i(t_j). \]

The \( f_k \) are deterministic functions, treated as known in standard finance methodology.

The \( K \) time series \( \{\beta_{i,k}, \ k = 1, 2, 3, \ldots\} \) are modeled as time series (modern dynamic models).

If \( \beta_{i,k} \equiv \beta_k \), static model.
Nelson–Siegel factors $f_1(t, \lambda), f_2(t, \lambda), f_3(t, \lambda)$. 

\[ \lambda = 0.0609 \]
\{X_i(t_j), \ i = 1, 2, 3, \ldots \} is a multivariate time series of dimension \( J = 10 \).

Why not model it as a VAR process?

The yield curves have a specific shape whose components have economic interpretation.

Factors lead to nontrivial and practically relevant dimension reduction and improved estimation and prediction.
Change point test in a functional factor model

\[ X_i(t) = \sum_{k=1}^{K} \beta_{i,k} f_k(t) + \varepsilon_i(t), \quad 1 \leq i \leq N. \]

(Constant time)

\[ \beta_{i,k} = \mu_{i,k} + b_{i,k}, \quad Eb_{i,k} = 0. \]

\[ \mu_i = [\mu_{i,1}, \mu_{i,2}, \ldots, \mu_{i,K}]^\top \]

\[ H_0 : \mu_1 = \mu_2 = \ldots = \mu_N. \]
Change point vs. break point

Recall: $H_0$: mean structure does not change

Error functions: $\sum_{k=1}^{K} b_{i,k} f_k(t) + \varepsilon_i(t)$.
The distribution of the error functions can change at known points:

$$1 = i_0 < i_1 < i_2 < \ldots < i_M < i_{M+1} = N.$$  

(Similar to prior information in Bayesian inference)
On each segment, error functions are realizations of potentially different weakly dependent processes in $L^2$.

Compare:
unknown \textit{change points} in mean structure
known \textit{break points} in error structure
Determining break points

Dates of substantial central bank intervention,

Dates of events of economic impact,

Exploratory analysis of the variability of the yield curves.

After the application of the test, some change points may be close to break points.
Vectors of projections:

\[ z_i = [\langle X_i, f_1 \rangle, \ldots, \langle X_i, f_K \rangle]^\top \]

Introduce the deterministic matrix

\[ C = [\langle f_k, f_j \rangle, 1 \leq k, j \leq K]. \]

and random vectors

\[ b_i = [b_{i,1}, b_{i,2}, \ldots, b_{i,K}]^\top, \quad \epsilon_i = [\langle \epsilon_i, f_1 \rangle, \langle \epsilon_i, f_2 \rangle, \ldots, \langle \epsilon_i, f_K \rangle]^\top. \]

Then

\[ z_i = C \mu_i + \gamma_i, \quad \gamma_i = C b_i + \epsilon_i, \]

Change in the vectors \( \mu_i \) is equivalent to a change in the \( z_i \) at the same change points.
Detection through projections onto factors

CUSUM process:

$$\alpha_N(x) = N^{-1/2} \left( \left\lfloor N x \right\rfloor \sum_{i=1}^{\left\lfloor N x \right\rfloor} z_i - \frac{\left\lfloor N x \right\rfloor}{N} \sum_{i=1}^{N} z_i \right), \quad 0 \leq x \leq 1.$$  

Even under $H_0$, the distribution of the $z_i$ can change at break points.

$\left\{ \gamma_i^{(m)} \right\}$ - model for the vector errors $\gamma_i$ over the interval $(i_m, i_{m+1}]$.

Long–run covariance matrices:

$$V_m = \sum_{\ell=-\infty}^{\infty} \text{cov}(\gamma_i^{(m)}, \gamma_{i+\ell}^{(m)}).$$
Detection through projections onto factors

Limit distribution of the CUSUM process:

$$\alpha_N \to G^0, \quad \text{in } D^K([0, 1]), \quad G^0(x) = G(x) - xG(1).$$

$$\{G(x), \ x \in [0, 1]\}$$ is a mean zero $\mathbb{R}^K$–valued Gaussian process with covariances

$$E[G(x)G^\top(y)] = \sum_{j=1}^m (\theta_j - \theta_{j-1})V_j + (x - \theta_m)V_{m+1}, \ \theta_m \leq x \leq \theta_{m+1}, y \geq x.$$

The covariances of the process $G^0$ can be computed explicitly (a long formula).

Cramér–von–Mises functional

$$\int_0^1 \|\alpha_N(x)\|^2 \, dx \to \int_0^1 \|G^0(x)\|^2 \, dx.$$

We must simulate the distribution of the right-hand side.
Detection through projections onto factors

1. Eigenvalue method (Eigen):

Estimate $\mathbf{R}(x, y) = E[\mathbf{G}^0(x)\mathbf{G}^0(y)^\top]$. 

Use equality in distribution:

$$\int_0^1 \| \mathbf{G}^0(x) \|^2 dx \sim \sum_{j=1}^{\infty} \lambda_j Z_j^2.$$ 

The $Z_j$ are independent standard normal;

$$\int_0^1 \mathbf{R}(x, y) \phi_j(y) dy = \lambda_j \phi_j(x).$$
2. Direct simulation method (\textbf{Sim}): 

Generate replications of the process $G^0$. Enough to generate 

$$G(x) = \sum_{j=1}^{m} (\theta_j - \theta_{j-1})^{1/2} G_j(1) + (\theta_{m+1} - \theta_m)^{1/2} G_{m+1} \left( \frac{x - \theta_m}{\theta_{m+1} - \theta_m} \right),$$

$$x \in (\theta_m, \theta_{m+1}].$$

$$E[G_j(x)G_j(y)^\top] = \min(x, y)V_j,$$

$$G_j(x) = L_jW(x), \quad L_jL_j^\top = V_j,$$

$$W = [W_1, W_2, \ldots, W_K]^\top,$$

$K$ independent standard Wiener processes.
Nonparametric functional approach

Motivation: the factor model can be rewritten as

\[ X_i(t) = \tau_i(t) + \eta_i(t), \quad 1 \leq i \leq N, \]

where

\[ \tau_i(t) = \sum_{k=1}^{K} \mu_{i,k} f_k(t), \quad \eta_i(t) = \sum_{k=1}^{K} b_{ik} f_k(t) + \varepsilon_i(t). \]

The first line is a valid model without the second line.

We can test

\[ H_0 : \quad \tau_1 = \tau_2 = \ldots = \tau_N. \]

A version of the **Eigen** method can be derived, it has to be custom coded, one needs to solve

\[ \int \int U^0(x, y; t, s) \varphi_j(y, s) dyds = \lambda_j \varphi_j(x, t). \]

The 4D kernel is expressed in terms suitably defined LRV kernels.
Application to yield curves

Case (1)
Application to yield curves

Case (2)
Application to yield curves

Case (3)
## Application to yield curves

<table>
<thead>
<tr>
<th>Sampling Period</th>
<th>Method</th>
<th>Break Point</th>
<th>P–value</th>
</tr>
</thead>
<tbody>
<tr>
<td>03/20/2008 – 03/19/2009</td>
<td>ProjSim</td>
<td>yes</td>
<td>1.5%</td>
</tr>
<tr>
<td></td>
<td>ProjEigen</td>
<td>yes</td>
<td>1.7%</td>
</tr>
<tr>
<td></td>
<td>NFEigen</td>
<td>yes</td>
<td>0.1%</td>
</tr>
<tr>
<td></td>
<td>ProjSim</td>
<td>no</td>
<td>87.9%</td>
</tr>
<tr>
<td></td>
<td>ProjEigen</td>
<td>no</td>
<td>85.2%</td>
</tr>
<tr>
<td></td>
<td>NFEigen</td>
<td>no</td>
<td>26.2%</td>
</tr>
<tr>
<td>(1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>06/30/2005 - 06/29/2006</td>
<td>ProjSim</td>
<td>yes</td>
<td>0.1%</td>
</tr>
<tr>
<td></td>
<td>ProjEigen</td>
<td>yes</td>
<td>0.0%</td>
</tr>
<tr>
<td></td>
<td>NFEigen</td>
<td>yes</td>
<td>0.2%</td>
</tr>
<tr>
<td></td>
<td>ProjSim</td>
<td>no</td>
<td>56.7%</td>
</tr>
<tr>
<td></td>
<td>ProjEigen</td>
<td>no</td>
<td>50.5%</td>
</tr>
<tr>
<td></td>
<td>NFEigen</td>
<td>no</td>
<td>57.2%</td>
</tr>
<tr>
<td>(2)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>02/16/2012 – 02/14/2013</td>
<td>ProjSim</td>
<td>yes</td>
<td>68.1%</td>
</tr>
<tr>
<td></td>
<td>ProjEigen</td>
<td>yes</td>
<td>66.9%</td>
</tr>
<tr>
<td></td>
<td>NFEigen</td>
<td>yes</td>
<td>55.8%</td>
</tr>
<tr>
<td></td>
<td>ProjSim</td>
<td>no</td>
<td>80.4%</td>
</tr>
<tr>
<td></td>
<td>ProjEigen</td>
<td>no</td>
<td>77.7%</td>
</tr>
<tr>
<td></td>
<td>NFEigen</td>
<td>no</td>
<td>76.3%</td>
</tr>
<tr>
<td>(3)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Conclusions based on similar data examples and simulations:

1. If a possible break point in the error structure is not taken into account in any of the testing procedures, an existing change point in the mean structure may not to be detected.

2. Break points can be misplaced, ±50 days is fine.

3. The tests are generally well calibrated if \( N = 500 \). If \( N = 250 \), all tests have a tendency to overreject by some 2 percent.

4. If the intelligible factors (Lengwiller and Lenz (2010) are used, the empirical size of projSim test improves at the 5% nominal level, but the size of the ProjEigen test deteriorates.

5. We recommend ProjSim with intelligible factors and NFEigen.

The paper contains complete asymptotic theory and details of numerical implementation.