Time Series of Cross-sectional Distributions

Joon Y. Park

Department of Economics

Indiana University

20th International Panel Data Conference

Hitotsubashi Hall, Tokyo, Japan

9-10 July 2014

References

Main Contents

- Park and Qian (2012), Functional Regression of Continuous State Distributions, *Journal of Econometrics*.
- Chang, Kim and Park (2012), Nonstationarity in Time Series of State Densities.
- Chang, Kim and Park (2014), Common Trends in Time Series of Cross Sectional Distributions.
- Chang, Kim, Miller, Park and Park (2014), Time Series Analysis of Global Temperature Distributions: Identifying and Estimating Persistent Features in Temperature Anomalies.

Background Material

- Bosq (2000), Linear Processes in Function Spaces.

Objective

Analysis of time series of cross-sectional distributions such as

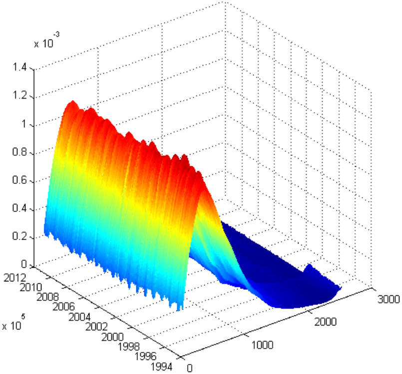
- individual earnings
- global temperatures
- household income and consumption

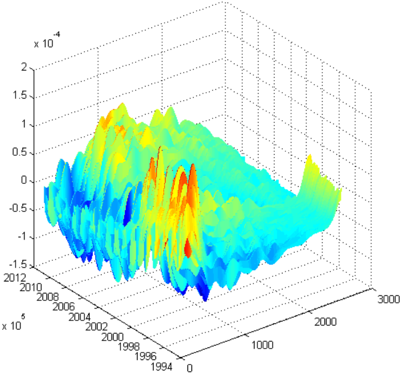
and time series of intra-period distributions such as

- daily distributions of high frequency US/UK exchange rates
- monthly distributions of high frequency S&P 500 returns

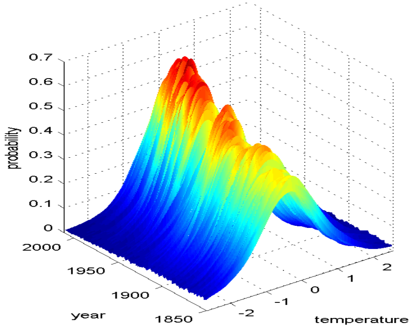
and many others.

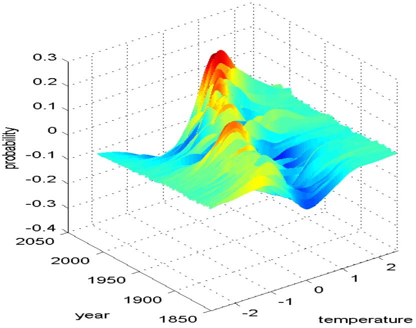
Cross-sectional Distributions of Individual Earnings



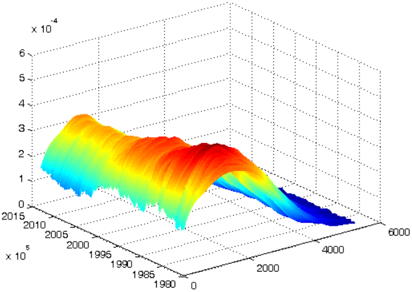


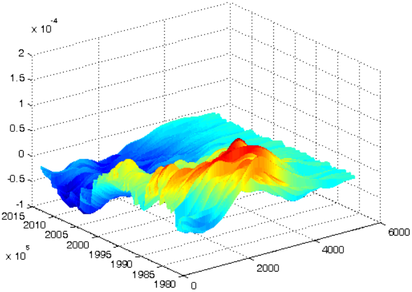
Global Temperature Distributions



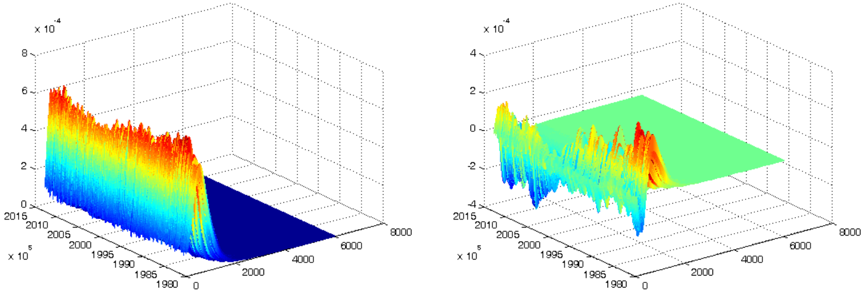


Cross-sectional Distributions of Household Income

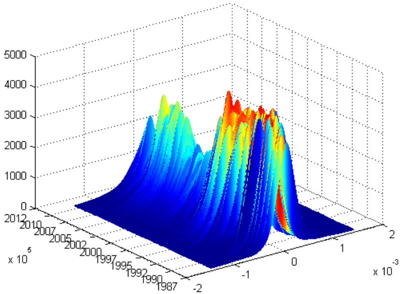


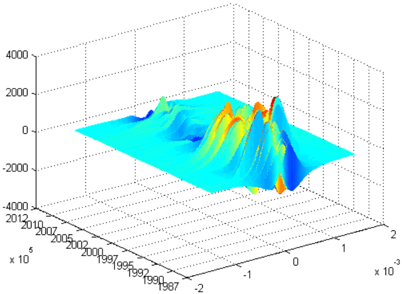


Cross-sectional Distributions of Consumption Expenditures



Monthly Distributions of High Frequency S&P 500 Returns





Objectives

We propose a framework and methodology to analyze the time series of state densities, which may represent either cross-sectional or intra-period distributions. Each state density is regarded as a realization of Hilbertian random variable, and a functional time series model is used to fit a given time series of state variables.

We identify and extract the nonstationarity in time series of state distributions. The stationary and nonstationary components of state densities are decomposed and used to model the shortrun and longrun relationships respectively. The nonstationary portions in the moments of state densities are also obtained.

Contents

Technical Background

Hilbert-Valued Random Variables

A Hilbert-valued random variable w is defined as

$$w : \ \Omega \to H,$$

where H is a Hilbert space.

Hilbert-valued random variables include

- real random variables: $H = \mathbb{R}$
- vector-valued random variables: $H = \mathbb{R}^N$
- function-valued random variables: $H = L^2(\mathbb{R})$

as special cases.

Mean

The mean μ of a H-valued random variable w is defined as a vector in H satisfying

$$\mathbb{E}\langle v, w\rangle = \langle v, \mu\rangle$$

for all $v \in H$, which exists if $\mathbb{E}\|w\| < \infty$. We write

$$\mu = \mathbb{E}w$$

following the usual convention.

Variance Operator

The variance operator Σ of a H-valued random variable w is defined as an operator on H satisfying

$$\mathbb{E}\langle v_i, w - \mathbb{E}w\rangle\langle v_j, w - \mathbb{E}w\rangle = \langle v_i, \Sigma v_j\rangle$$

for all v_i, $v_j \in H$, which exists if $\mathbb{E}\|w\|^2 < \infty$. We write

$$\Sigma = \mathbb{E}\Big((w - \mathbb{E}w) \otimes (w - \mathbb{E}w)\Big),$$

which reduces to $\Sigma = \mathbb{E}\Big((w - \mathbb{E}w)(w - \mathbb{E}w)'\Big)$ if H is finite-dimensional.

Functional AR Model

Let (w_t) be a time series of random functions. The functional AR(1) model for (w_t) is given by

$$w_t = Aw_{t-1} + \varepsilon_t,$$

where A is a linear operator transforming a function into another function, and (ε_t) is an iid sequence of random functions. We may write

$$w_t = \varepsilon_t + A\varepsilon_{t-1} + A^2\varepsilon_{t-2} + \cdots$$

by recursive substitution.

Mean Reversion in Functional AR Model

The mean reversion of (w_t) is determined by the norm $\|A\|$ of A.

It is stationary if $\|A\| < 1$, i.e., $\|Av\| < \|v\|$ for all v. Mean reversion in all directions. Deviates from mean only temporarily, and randomly fluctuates around the mean in all directions.

It has a unit root in the direction of v if $\|Av\| = \|v\|$. Persistent, and non mean reverting due to the presence a stochastic trend with no mean reversion in the direction of v.

It is explosive in the direction of v if $\|Av\| > \|v\|$. No mean reversion in the direction of v.

Coordinate Process

We assume that there exists an orthonormal basis (v_i) of H, so that we may write

$$w_t = \sum_{i=1}^{\infty} \langle v_i, w_t \rangle v_i,$$

and define

$$\langle v_i, w_t \rangle$$

to be the i-th coordinate process for $i = 1, 2, \ldots$. Each coordinate process may be stationary, have a unit root, or be explosive.

Cross-Sectional Moment

Let

$$\iota_\kappa(s) = s^\kappa$$

for $\kappa = 1, 2, \ldots$. For a time series (f_t) of probability densities, we have

$$\langle \iota_\kappa, f_t \rangle = \int_K \iota_\kappa(s) f_t(s) ds,$$

which we call the κ-th cross-sectional moment of f_t for $\kappa = 1, 2, \ldots$. For $\kappa = 1$, in particular, we have the cross-sectional mean $\langle \iota_\kappa, f_t \rangle$ of f_t, which is in contrast with the mean $\mathbb{E} f_t$ of f_t.

Distributional Autoregression

Distributional AR(1) Model

The distributional AR(1) model for the time series (f_t) of probability densities is given by

$$f_t = \mu + Af_{t-1} + \varepsilon_t,$$

where (ε_t) is an iid random functions with mean zero and finite variance operator. We may write the model more compactly as

$$w_t = Aw_{t-1} + \varepsilon_t,$$

where $w_t = f_t - \mathbb{E}f_t$ is the centered density function for all $t = 1, 2, \ldots$. Note that $\int_K w_t(s)ds = 0$, where $K \subset \mathbb{R}$ is the support of (f_t).

Ill-Posed Inverse Problem

We may easily deduce the functional Yule-Walker equation

$$\mathbb{E}(w_t \otimes w_{t-1}) = A\mathbb{E}(w_{t-1} \otimes w_{t-1}),$$

which we may write as $N = AM$. However, we cannot define

$$A = NM^{-1}.$$

The operator M is self-adjoint, positive and infinite-dimensional, so we have

$$M = \sum_{i=1}^{\infty} \lambda_i (v_i \otimes v_i)$$

by the spectral representation theorem, where the spectrum (λ_i), $\lambda_i > 0$, has 0 as its limit point. Therefore, M necessarily has an ill-posed inverse problem.

Estimation of Distributional AR(1) Model

We define

$$N_T = \sum_{t=2}^{T} (f_t - \bar{f}_T) \otimes (f_{t-1} - \bar{f}_T), \;\; M_T = \sum_{t=1}^{T} (f_t - \bar{f}_T) \otimes (f_t - \bar{f}_T)$$

with $\bar{f}_T = \sum_{t=1}^{T} f_t / T$, and let

$$M_T^+ = \sum_{i=1}^{m} \frac{1}{\lambda_i^T} (v_i^T \otimes v_i^T),$$

where (λ_i^T, v_i^T) is eigenvalue and eigenvector pair of M_T, and m is an appropriately chosen truncation number. Then we may estimate A by $A_T = N_T M_T^+$.

Density Prediction

Once an estimator A_T of A is obtained, we may readily predict the density f_{T+1} at time $T+1$ as

$$f_{T+1} - \bar{f}_T = A_T \left(f_T - \bar{f}_T \right),$$

which may be useful in many different contexts.

Testing for Moment Dynamics

We may test for moment dynamics such as ARCH-M. For instance, it follows directly from the distributional AR(1) model that

$$\begin{aligned}\langle \iota_1, f_t \rangle &= \langle \iota_1, \mu \rangle + \langle \iota_1, A f_{t-1} \rangle + \langle \iota_1, \varepsilon_t \rangle \\ &= \langle \iota_1, \mu \rangle + \langle A^* \iota_1, f_{t-1} \rangle + \langle \iota_1, \varepsilon_t \rangle,\end{aligned}$$

where A^* is the adjoint of A. The variance in the previous period does not affect the mean in the current period if in particular ι_2 is orthogonal to $A^*\iota_1$.

Distributional Unit Root

Unit Root in Distribution

We assume that there exists an orthonormal basis (v_i) of H such that the i-th coordinate process

$$\langle v_i, w_t \rangle$$

has a unit root for $i = 1, \ldots, n$, while it is stationary for all $i \geq n + 1$.

By convention, we set $n = 0$ if all the coordinate processes are stationary.

Unit Root and Stationarity Subspaces

Using the symbol $\bigvee$ to denote span, we let

$$H_N = \bigvee_{i=1}^{n} v_i \quad \text{and} \quad H_S = \bigvee_{i=n+1}^{\infty} v_i$$

so that $H = H_N \oplus H_S$. In what follows, H_N and H_S will respectively be referred to as the unit root and stationarity subspaces of H.

Accordingly, we let Π_N and Π_S be the projections on H_N and H_S, respectively. Moreover, we define

$$w_t^N = \Pi_N w_t \quad \text{and} \quad w_t^S = \Pi_S w_t.$$

Note that

$$w_t = w_t^N + w_t^S,$$

since $\Pi_N + \Pi_S = 1$.

Functional Principal Component Analysis

Our procedure to estimate H_N and test for its dimension M is based on the FPCA on the unnormalized sample variance operator of (w_t)

$$M_T = \sum_{t=1}^{T} w_t \otimes w_t$$

where T is the sample size.

Denote the pairs of eigenvalues and eigenvectors of M_T by

$$(\lambda_i^T, v_i^T), \quad i = 1, \ldots, T$$

and order (λ_i^T) so that $\lambda_1^T \geq \cdots \geq \lambda_T^T$.

Sample Unit Root and Stationarity Subspaces

Assuming $T > n$, we define sample unit root space as the subspace

$$H_N^T = \bigvee_{i=1}^{n} v_i^T$$

spanned by the eigenvectors corresponding to n largest eigenvalues of M_T. Denote by Π_N^T the projection on H_N^T.

The sample stationarity subspace is defined by $\Pi_S^T = 1 - \Pi_N^T$, so that we have $\Pi_N^T + \Pi_S^T = 1$ analogously as the relationship $\Pi_N + \Pi_S = 1$.

Asymptotics for Sample Projections

Under very general regularity conditions, we have

$$\Pi_N^T = \Pi_N + O_p(T^{-1})$$
$$\Pi_S^T = \Pi_S + O_p(T^{-1})$$

for all large T.

Decomposition of Sample Variance Operator

To develop our asymptotics, we decompose M^T as

$$M^T = T^2 M^T_{NN} + T M^T_{NS} + T M^T_{SN} + T M^T_{SS},$$

where

$$M^T_{NN} = \frac{1}{T^2}\Pi_N\left(\sum_{t=1}^T w_t \otimes w_t\right)\Pi_N = \frac{1}{T^2}\sum_{t=1}^T w_t^N \otimes w_t^N$$

$$M^T_{NS} = \frac{1}{T}\Pi_N\left(\sum_{t=1}^T w_t \otimes w_t\right)\Pi_S = \frac{1}{T}\sum_{t=1}^T w_t^N \otimes w_t^S$$

$$M^T_{SS} = \frac{1}{T}\Pi_S\left(\sum_{t=1}^T w_t \otimes w_t\right)\Pi_S = \frac{1}{T}\sum_{t=1}^T w_t^S \otimes w_t^S$$

and M^T_{SN} is the adjoint of M^T_{NS}, i.e., $M^T_{SN} = M^{T*}_{NS}$.

Asymptotics for Sample Variance Operators

Under some regularity conditions, we have

$$M_{NN}^T \ \to_d \ M_{NN} = \int_0^1 (W \otimes W)(r)dr,$$

where W is Brownian motion on H_N. Also, it follows that

$$M_{SS}^T \ \to_p \ M_{SS}.$$

Moreover, we have

$$M_{NS}^T, M_{SN}^T = O_p(1)$$

for all large T.

Asymptotics for Eigenvalues and Eigenvectors

In unit root subspace H_N, eigenvectors and appropriately normalized eigenvalues of sample variance operator M_T of (w_t) converge in distribution, and their distributional limits are given by the distributions of eigenvalues and eigenvectors of random operator M_{NN}, i.e.,

$$\left(T^{-2}\lambda_i^T, v_i^T\right) \to_d (\lambda_i(M_{NN}), v_i(M_{NN}))$$

jointly for $i = 1, \ldots, n$, under some regularity conditions.

In stationarity subspace H_S, eigenvectors and appropriately normalized eigenvalues of sample variance operator M_T of (w_t) converge in probability to their population counterparts, i.e.,

$$\left(T^{-1}\lambda_{n+i}^T, v_{n+i}^T\right) \to_p (\lambda_i, v_i)$$

for $i = 1, 2, \ldots$, under very general regularity conditions.

Testing for Distributional Unit Roots

To determine the number of unit roots in (w_t), we consider the test of the null hypothesis

$$\mathrm{H}_0: \ \dim(H_N) = n$$

against the alternative hypothesis

$$\mathrm{H}_1: \ \dim(H_N) \leq n - 1$$

successively downward. We estimate the number of unit roots in (w_t) by the smallest value of n for which we fail to reject the null hypothesis H_0.

Intuitive but Infeasible Test

We expect that the eigenvalue λ_n^T would have a discriminatory power for the test of null against the alternative, since it has different orders of stochastic magnitudes under the null and alternative hypotheses.

However, it cannot be used directly as a test statistic, since its limit distribution is dependent upon nuisance parameters.

Therefore, we need to modify it appropriately to get rid of its nuisance parameter dependency problem.

A Feasible Test for Unit Root Dimension

We let

$$z_t^T = (\langle v_1^T, w_t \rangle, \ldots, \langle v_n^T, w_t \rangle)'$$

for $t = 1, \ldots, T$, and define the product sample moment $M_n^T = \sum_{t=1}^T z_t^T z_t^{T\prime}$, and the long-run variance estimator $\Omega_n^T = \sum_{|k| \leq \ell} \varpi_\ell(k) \Gamma_T(k)$ of (z_t^T), where ϖ_ℓ is the weight function with bandwidth parameter ℓ and Γ_T is the sample autocovariance function defined as $\Gamma_T(k) = T^{-1} \sum_t \Delta z_t^T \Delta z_{t-k}^{T\prime}$.

Our test statistic is defined as

$$\tau_n^T = T^{-2} \lambda_{\min} \left(M_n^T, \Omega_n^T \right),$$

where $\lambda_{\min} \left(M_n^T, \Omega_n^T \right)$ is the smallest generalized eigenvalue of M_n^T with respect to Ω_n^T.

Asymptotics for Distributional Unit Root Test

Under very general conditions, we show that

$$\tau_n^T \to_d \lambda_{\min}\left(\int_0^1 W_n(r)W_n(r)'dr - \int_0^1 W_n(r)dr \int_0^1 W_n(r)'dr\right)$$

under the null, as $T \to \infty$, where W_n is n-dimensional standard vector Brownian motion and $\lambda_{\min}(\cdot)$ denotes the smallest eigenvalue of its matrix argument.

On the other hand, we have $\tau_n^T \to_p 0$ under the alternative hypothesis as $T \to \infty$. Therefore, we reject the null in favor of the alternative if the test statistic τ_n^T takes small values.

Critical Values for Distributional Unit Root Test

Critical values of the test τ_n^T are calculated and tabulated as

n	1	2	3	4	5
1%	0.0274	0.0175	0.0118	0.0103	0.0085
5%	0.0385	0.0223	0.0154	0.0127	0.0101
10%	0.0478	0.0267	0.0175	0.0139	0.0111

for $n = 1, \ldots, 5$.

Estimation of Nonstationarity Subspace

Once we determine the number of unit roots n in (w_t), we may estimate the nonstationarity subspace H_N by

$$H_N^T = \bigvee_{i=1}^{n} v_i^T,$$

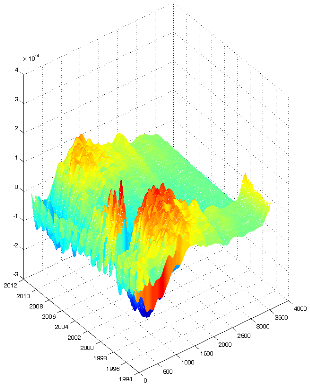
i.e., the span of the n orthonormal eigenvectors of M^T associated with n largest eigenvalues of M_T.

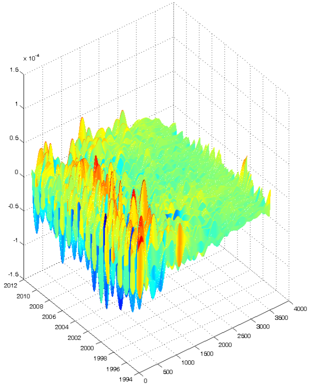
Under general regularity conditions, we have

$$H_N^T \to_p H_N$$

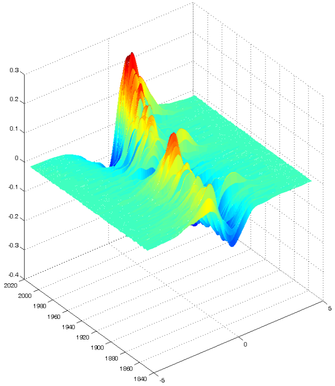
as $T \to \infty$. Recall that λ_i^T does not converge in probability to λ_i individually for $i = 1, \ldots, n$.

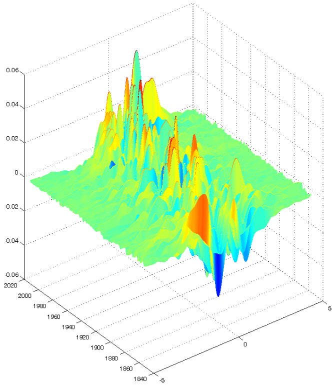
Cross-sectional Distributions of Income



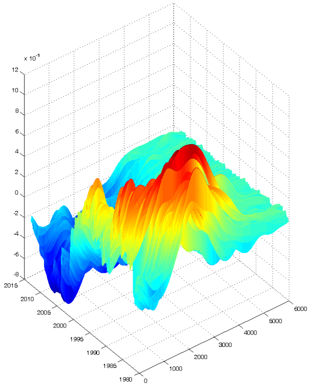


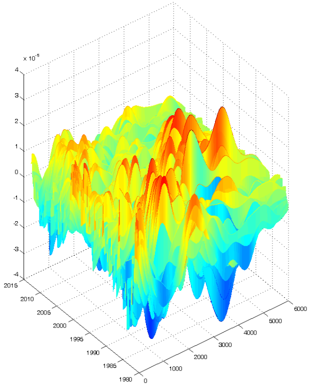
Cross-sectional Distributions of Global Temperature



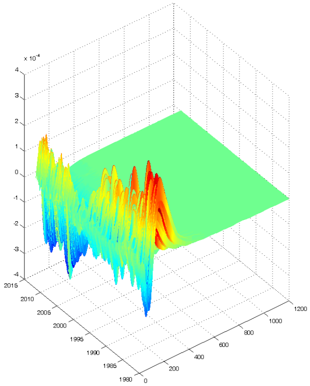


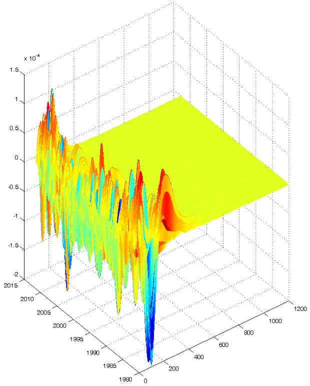
Cross-sectional Distributions of Household Income





Cross-sectional Distributions of Consumption Expenditures





Degree of Persistency in Cross-sectional Moments

We may now find how much nonstationarity proportion exists in each cross-sectional moments. By convention, we assume that ι_κ is redefined as $\iota_\kappa - \int_K \iota_\kappa(s)ds$, so that we may consider it with (w_t) in the same Hilbert space H.

We decompose ι_κ as $\iota_\kappa = \Pi_N \iota_\kappa + \Pi_S \iota_\kappa$, from which it follows that

$$\|\iota_\kappa\|^2 = \|\Pi_N \iota_\kappa\|^2 + \|\Pi_S \iota_\kappa\|^2 = \sum_{i=1}^{n} \langle \iota_\kappa, v_i \rangle^2 + \sum_{i=n+1}^{\infty} \langle \iota_\kappa, v_i \rangle^2,$$

where (v_i), $i = 1, 2, \ldots$, is an orthonormal basis of H such that $(v_i)_{1 \leq i \leq n}$ and $(v_i)_{i \geq n+1}$ span H_N and H_S, respectively.

Nonstationarity Proportion of Cross-sectional Moments

To measure the proportion of ι_κ lying in H_N, we define

$$\pi_\kappa = \frac{\|\Pi_N \iota_\kappa\|}{\|\iota_\kappa\|} = \sqrt{\frac{\sum_{i=1}^{n} \langle \iota_\kappa, v_i \rangle^2}{\sum_{i=1}^{\infty} \langle \iota_\kappa, v_i \rangle^2}}.$$

Note that we have $\pi_\kappa = 1$ and $\pi_\kappa = 0$, respectively, if ι_κ is entirely in H_N and H_S.

We use π_κ to represent the proportion of unit root component in the κ-th cross-sectional moment of (w_t). Clearly, it has more dominant unit root component as π_κ tends to unity, whereas it becomes more stationary as π_κ approaches to zero.

Sample Nonstationarity Proportion

Obviously, we may use the sample unit root proportion

$$\pi_\kappa^T = \sqrt{\frac{\sum_{i=1}^{n} \langle \iota_\kappa, v_i^T \rangle^2}{\sum_{i=1}^{T} \langle \iota_\kappa, v_i^T \rangle^2}}$$

to measure the unit root proportion of the κ-th cross-sectional moment of (w_t). As is well expected, the sample version π_κ^T is a consistent estimator for the original π_κ, under very general conditions.

Distributional Cointegration

Two Cross-Sectional Distributions with Unit Roots

Let (f_t) and (g_t) be two time series of densities representing cross-sectional distributions of some economic variables. Assume that they have distributional unit roots, and denote respectively by $H_N(f)$ and $H_N(g)$ the nonstationary subspaces of H for (f_t) and (g_t).

By definition, the coordinate processes

$$\langle v, f_t \rangle \quad \text{and} \quad \langle w, g_t \rangle$$

have unit roots for all $v \in H_N(f)$ and $w \in H_N(g)$.

Distributional Cointegration

Given that the coordinate processes $\langle v, f_t \rangle$ and $\langle w, g_t \rangle$ are integrated for all $v \in H_N(f)$ and $w \in H_N(g)$, it is natural to consider the possibility that some of their coordinate processes are cointegrated.

In fact, for some $v \in H_N(f)$ and $w \in H_N(g)$, we may have

$$\langle v, f_t \rangle + \langle w, g_t \rangle = \pi + u_t$$

with some constant π, where (u_t) is a general stationary process with mean zero, in which case we say that (f_t) and (g_t) are distributionally cointegrated.

More Precise Formulation

Let (f_t) and (g_t) have p- and q-unit roots, and $H_N(f)$ and $H_N(g)$ be p- and q-dimensional, respectively. Moreover, define

$$H_N(f) = \bigvee_{i=1}^{p} v_i \quad \text{and} \quad H_N(g) = \bigvee_{j=1}^{q} w_j,$$

so that $\langle v_i, f_t \rangle$ and $\langle w_j, g_t \rangle$ are unit root processes for $i = 1, \ldots, p$ and $j = 1, \ldots, q$.

Distributional Cointegrating Functions

If the $(p+q)$-dimensional process (z_t) defined as

$$z_t = \big(\langle v_1, f_t\rangle, \ldots, \langle v_p, f_t\rangle, \langle w_1, g_t\rangle, \ldots, \langle w_q, g_t\rangle\big)'$$

is cointegrated with the cointegrating vector

$$c = (\alpha_1, \ldots, \alpha_p, \beta_1, \ldots, \beta_q)',$$

then the distributional cointegration holds with

$$v = \alpha_1 v_1 + \cdots + \alpha_p v_p \quad \text{and} \quad w = \beta_1 w_1 + \cdots + \beta_q w_q.$$

The functions v and w are called the distributional cointegrating functions of (f_t) and (g_t).

Remark

There are at most r-number of linearly independent distributional cointegrating relationships,

$$r \leq \min(p, q).$$

Otherwise we would have a cointegrating vector c of the form $c = (\alpha_1, \ldots, \alpha_p, 0, \ldots, 0)'$ or $c = (0, \ldots, 0, \beta_1, \ldots, \beta_q)'$, which implies that there is a linear combination of $v_1, \ldots, v_p$ or $w_1, \ldots, w_q$ whose inner product with (f_t) or (g_t) becomes stationary, contradicting the assumption that $v_1, \ldots, v_p$ and $w_1, \ldots, w_q$ are linearly independent functions that span $H_N(f)$ and $H_N(g)$ respectively.

Distributional Cointegrating Subspace

If $r > 1$, we use the notations (v_k^C) and (w_k^C), $k = 1, \ldots, r$, for the distributional cointegrating functions of (f_t) and (g_t). However, in this case, the distributional cointegrating functions (v_k^C) and (w_k^C), $k = 1, \ldots, r$, of (f_t) and (g_t) are not individually identified, unless we impose some specific restrictions on their normalization.

The subspaces of $H_N(f)$ and $H_N(g)$ spanned by them are nevertheless well identified, which we denote by $H_C(f)$ and $H_C(g)$ and call the distributional cointegrating subspaces of (f_t) and (g_t), respectively.

Testing for Distributional Cointegration

We let

$$x_t = f_t - \bar{f}_T \quad \text{and} \quad y_t = g_t - \bar{g}_T$$

and define

$$z_t^T = (\langle v_1^T, x_t\rangle, \ldots, \langle v_p^T, x_t\rangle, \langle w_1^T, y_t\rangle, \ldots, \langle w_q^T, y_t\rangle)'$$

for $t = 1, \ldots, T$, where $v_1^T, \ldots, v_p^T$ and $w_1^T, \ldots, w_q^T$ are estimated eigenvectos spanning the unit root subspaces $H_N(f)$ and $H_N(g)$ of (f_t) and (g_t). We may simply use our statistic for distributional unit root with the redefined (z_t) as above and find the distributional cointegration between (f_t) and (g_t).

An Example

In our empirical study, we consider household income (f_t) and consumption expenditure (g_t). We find two unit roots in (f_t) and one unit root in (g_t), and one cointegrating relationship between them. Therefore, we have three unit root coordinate processes, say,

$$\langle v_1, f_t\rangle, \langle v_2, f_t\rangle \quad \text{and} \quad \langle w, g_t\rangle,$$

and one cointegrating relationship

$$\langle w, g_t\rangle = \pi + \alpha_1\langle v_1, f_t\rangle + \alpha_2\langle v_2, f_t\rangle + u_t,$$

where we assume without loss of generality that the coefficient of $\langle w^C, g_t\rangle$ is unity.

Interpretation

In our example, $\langle w, g_t \rangle$ represents the unique longrun component of household consumption expenditure and its coordinate process is given by $\langle w, g_t \rangle$. On the other hand, household income has two longrun components, whose coordinate processes are given by $\langle v_1, f_t \rangle$ and $\langle v_2, f_t \rangle$.

We have a stochastic common trend in

$$\langle w^C, g_t \rangle \quad \text{and} \quad \langle v^C, f_t \rangle,$$

where $w^C = w$ and $v^C = \alpha_1 v_1 + \alpha_2 v_2$. Moreover, we define v^C to be the longrun response of income to consumption expenditure.

Unit Root Coordinate Processes and Common Trend