博士学位請求論文要旨

"Theoretical Investigation on Macroeconomic Time Series"

田中　晋矢

Macroeconomic time series such as GDP, the industrial production index, and the unemployment rate reflect the state of a real economy. Since macroeconomic time series are the most important information set for economic activity, they are widely used as data for empirical analyses. However, the following two points should be considered when dealing with macroeconomic time series.

First, the integrated order (stationarity or non-stationarity) of the series should be specified correctly because the asymptotic properties of estimators crucially depend on it. For example, when the series are $I(1)$ –that is, when the series have a unit root– the regression is spurious and the OLS estimator no longer has consistency without any treatments (such as differencing) except for the special case of cointegration. Then, misspecifying a series as stationary when it is actually non-stationary is a fatal problem in parameter estimation. On the other hand, regarding a stationary series as non-stationary is also problematic because the problem of over-differencing arises. To avoid such problems, it is necessary to use tests for a unit root or stationarity with correct size and high power.

The second point concerns the treatment of multi-dimensional data. Recent development of technological advances in data collection has made large-dimensional datasets of macroeconomic time series available for empirical analysis. Although using as many macroeconomic time series as we can is widely believed to be ideal, the traditional VAR model would not be appropriate for large-dimensional data, which lead to a lack of degrees of freedom of the VAR model, causing the incidental parameters problem in theory and poor accuracy of the estimators in practice. Alternatively, approximate factor models have been used for large-dimensional data recently because they are quite useful for reducing the dimensions of data without the loss of their rich information. However, theoretical studies on approximate factor models are ongoing, and further theoretical studies on the approximate factor models are yet needed.

The aim of this dissertation is to provide new theoretical findings on testing for stationarity and approximate factor models. This dissertation is organized as follows:

Chapter 1 Overview

Chapter 2 Reducing the Size Distortion of the KPSS test

Chapter 3 Investigating Finite Sample Properties of Estimators for Approximate Factor Models When N Is Small

Chapter 4 Identification of Approximate Factor Model through Heteroskedasticity

1. Summary of Chapter 2

Chapter 2 investigates the sources of the size distortion of the test proposed by Kwiatkowski et al. (1992) (KPSS test, hereafter) and proposes the bias corrected KPSS test statistic with less size distortion.

The KPSS test is most widely used by applied econometricians to test for the stationarity of macroeconomic time series. Although there is a similar type of test called unit root test, which include the Dickey–Fuller test and the Phillips-Perron test, the KPSS test is essentially different from the unit root tests in that the null hypothesis of the KPSS test is that the series is (covariance) stationary, while that of the unit root tests is that the series is non-stationary. Thus, the KPSS test is quite useful for detecting the non-stationarity of a series from a complementary point of view.

Let $y_t = \mu_t + x_t$ where μ_t is a deterministic term, $(1-\alpha)x_t = u_t$ and u_t is a covariance stationary process with mean zero so that $E[y_t] = \mu_t$. The KPSS test statistic is defined as

$$\text{KPSS} = \frac{\frac{1}{T^2}\sum_{t=1}^{T}\left(\sum_{s=1}^{t}\hat{x}_s\right)^2}{\hat{\omega}},$$

where $\hat{x}_t$ is the regression residual of y_t on μ_t and $\hat{\omega}$ is a consistent estimator of the long-run variance $\omega = \lim_{T\to\infty} \text{Var}(T^{-1/2}\sum_{t=1}^{T} x_t)$. Then, the KPSS test statistic is constructed by two parts: (i) the squared sum of the partial sum of $\hat{x}_t$ (numerator) and (ii) the estimator of the long-run variance (denominator). Under $|\alpha| < 1$ –that is, y_t is covariance stationary– it can be shown that the KPSS test statistic weakly converges to a certain non-standard distribution that is free from nuisance parameters when T goes to infinity.

However, as Caner and Kilian (2001) and many others point out, the KPSS test has poor finite sample properties and suffers from serious size distortion of the test when the series is persistent. In particular, when α is close to 1, the KPSS test seriously overrejects the null of stationarity compared to its nominal size so that the user excessively misinterprets the stationary times series as non-stationary.

There is a question of where the sources of the size distortion of the KPSS test come from. The primary source of the distortion comes from (ii): the estimation bias of ω. Since we can see that x_t has an AR(∞) representation given by $x_t = \phi(L)\varepsilon_t$ under moderate assumptions on u_t where ε_t is a sequence of white noise with $E[\varepsilon_t] = \sigma_\varepsilon^2$, it is natural to employ the autoregressive spectral density estimator of ω defined by $\hat{\sigma}_\varepsilon^2/\hat{\phi}^2(1)$ where $\hat{\phi}^2(1)$ and $\hat{\sigma}_\varepsilon^2$ are the OLS estimates of AR(p) lag polynomials with fixed p and σ_ε^2 respectively. However, the KPSS test statistic with the autoregressive spectral density estimator suffers from a lack of power because it is not consistent, although it effectively mitigates the overrejection problem.

To avoid inconsistency of the test, Chapter 2 applies the data dependent boundary rule suggested by Sul et al. (2005) to the autoregressive spectral density estimator. The alternative autoregressive spectral density estimator is defined as

$$\tilde{\omega}_{AR} = \frac{\hat{\sigma}_\varepsilon^2}{(1-\tilde{\phi})^2} \quad \text{where} \quad \tilde{\phi} = \min\left(\sum_{j=1}^{p} \hat{\phi}_j,\ 1 - \frac{c}{\sqrt{T}}\right),$$

with c being some positive constant.

Furthermore, Chapter 2 sheds light on the source of the size distortion that comes from (i). Based on the AR(∞) representation of x_t, Chapter 2 derives the bias term b_T up to $O\left(\frac{1}{T}\right)$. Specifically, when x_t is an AR(1) process given by $x_t = \phi x_{t-1} + \varepsilon_t$, we will see that b_T can be represented by

$$b_T = -\frac{b_0}{T}\frac{\sigma_\varepsilon^2 \phi_1}{(1-\phi_1)^2(1-\phi_1^2)},$$

where b_0 is a fixed positive constant that varies depending on the specification of μ_t. Note that since the macroeconomic time series has a positive autocorrelation in general, $\phi_1 > 0$ so that $b_T < 0$. Then, in contrast to the bias from (ii), the bias from (i) turns out to be the source of underrejection and leads to diminish the power of the test. Taking into account these findings, Chapter 2 proposes the following bias corrected KPSS test statistic:

$$KPSS_{BC} = \frac{\frac{1}{T^2}\sum_{t=1}^{T}\left(\sum_{s=1}^{t}\hat{x}_s\right)^2 - \hat{b}_T}{\tilde{\omega}_{AR}},$$

where $\hat{b}_T$ is the OLS estimator of b_T. Although the bias corrected KPSS test statistic has the same limiting distribution as the original KPSS test statistic under the null hypothesis, it is expected that the bias corrected KPSS test statistic has much less size distortion than the existing ones based on the arguments above. We will see that Monte Carlo experiments provide the evidence that the bias corrected KPSS test statistic effectively reduces the size distortion as expected.

2. Summary of Chapter 3

Chapter 3 examines the finite sample properties of estimators for approximate factor models when the dataset is low-dimensional. Let x_{it} be the ith macroeconomic time series at time t for $i = 1, 2, \ldots, N$ and $t = 1, 2, \ldots, T$. The factor models assume that x_{it} is driven by r latent factors that are common among the series and an idiosyncratic error that is unique for i:

$$x_{it} = \lambda_{i1}F_{1t} + \lambda_{i2}F_{2t} + \cdots + \lambda_{ir}F_{rt} + \varepsilon_{it},$$

where F_{kt} is the kth latent factor, ε_{it} is the idiosyncratic error of i at t and λ_{ik} is the coefficient of F_{kt} called factor loading in the literature. We sometimes represent the factor model in multivariate form as

$$x_t = \Lambda F_t + \varepsilon_t,$$

where $x_t = [x_{1t}, x_{2t}, \ldots, x_{Nt}]'$ $(N \times 1)$, $\Lambda = [\lambda_1', \lambda_2', \ldots, \lambda_N']'$ $(N \times r)$ with $\lambda_i = [\lambda_{i1}, \lambda_{i2}, \ldots, \lambda_{ir}]'$ $(r \times 1)$ and $\varepsilon_t = [\varepsilon_{1t}, \varepsilon_{2t}, \ldots, \varepsilon_{Nt}]'$ $(N \times 1)$. Let $E[\varepsilon_t \varepsilon_t'] = \Omega$. The model is called the "approximate" factor model when the idiosyncratic errors are cross-sectionally dependent, i.e. Ω is non-diagonal positive definite matrix while the model is called the "exact" factor model if Ω is diagonal.

Since it is natural to consider that macroeconomic time series and their shocks are mutually dependent, the approximate factor model is suitable for economic analysis. However, we confront a problem in estimating the factors by approximate factor models: we need large dimensional datasets to obtain the estimators because their asymptotic behavior crucially depends on the assumption $N \to \infty$. Then, when we cannot obtain large dimensional datasets because of data limitations, must we give up employing approximate factor model in empirical analysis? Chapter 3 tries to answer this question by investigating the finite sample properties of the estimators for approximate factor models when the dataset is low-dimensional via Monte Carlo simulations.

Chapter 3 investigates the finite sample properties of the following three estimators when $N = 5, 7$ and 10: (*i*) the principal component analysis (PCA) estimator, (*ii*) the quasi-maximum likelihood (QML) estimator, and (*iii*) the state-space subspace (SSS) estimator. Chapter 3 yields the following key findings. First, the PCA and QML estimators perform very well even when N is small, while the SSS estimator does not. This result overturns the rule-of-thumb that the exact factor model should be employed when the data is low-dimensional. Second, the PCA estimator is suitable for strongly dependent approximate factor models, while the QML estimator is appropriate for exact or weakly dependent approximate factor models. Third, the PCA estimator is more robust to the degree of cross sectional dependence of the series than is the QML estimator.

3. Summary of Chapter 4

Chapter 4 investigates the identification problem of factor models and proposes a new identification scheme for large-dimensional factor models through heteroskedasiticity of factors.

It is well known that factor models have a serious problem in estimating factors: the "rotation problem." Consider the factor model

$$x_t = \Lambda^0 F_t^0 + \varepsilon_t, \qquad t = 1, 2, \ldots, T,$$

which is the same as the one in Chapter 3. Note that 0 indicates the true values. Let H^0 be a non-singular $r \times r$ matrix and insert $(H^0)^{-1}H^0$ $(= I_r)$ between Λ^0 and F_t^0 in this model. We thus obtain

$$x_t = \Lambda^0 (H^0)^{-1} H^0 F_t^0 + \varepsilon_t,$$
$$\equiv \mathcal{L}^0 \mathcal{F}_t^0 + \varepsilon_t,$$

where $\mathcal{L}^0 = \Lambda^0(H^0)^{-1}$ and $\mathcal{F}_t^0 = H^0 F_t^0$. Then, it is obvious that we cannot distinguish $\left(\Lambda^0, F_t^0\right)$ from $\left(\mathcal{L}^0, \mathcal{F}_t^0\right)$ without imposing additional assumptions on the model. That is, we cannot consistently estimate (true) Λ^0 and F_t^0 without imposing any identifying restrictions on the factors and factor loadings in general. We call H^0, $\mathcal{L}^0$ and $\mathcal{F}_t^0$ rotation matrix, rotated factor loadings (matrix) and rotated factors, respectively.

In econometric literature, previous studies such as Bernanke et al. (2005), Stock and Watson (2005) and Yamamoto (2011) propose identification schemes for factors based on the structural factor model. These are quite similar to that of the structural VAR model and we can estimate the true factor and factor loadings from rotated factors and factor loadings under their restrictions. However, as in the traditional structural VAR model, their identification schemes basically rely on zero restrictions such as Sims' (1980) recursive restriction and Blanchard and Quah's (1988) long-run restriction, which seem to be quite restrictive and ad hoc because they impose zeros on the structural parameters *a priori*.

Chapter 4 proposes a new identification scheme for approximate factor models through heteroskedasticity of factors, following the ideas of Rigobon (2003). It is worth noting that although our identification scheme is based on the structural model, it does not require zero restrictions such as the orthogonality of factors or recursive restriction of factor loadings.

The main assumptions of our model are that the sample period is divided into two regimes and that the variance of the factors changes depending on the regime while the factor loadings are invariant through regimes. Let S_1 and S_2 be the first and second halves of all sample periods, respectively. Then the process of factors is assumed as follows:

$$F_t^0 = \Phi_1^0(L) F_{t-1}^0 + e_t, \quad E[e_t e_t'] = I_r \qquad t \in S_1$$
$$F_t^0 = \Phi_2^0(L) F_{t-1}^0 + e_t, \quad E[e_t e_t'] = \Sigma_2^{e0} \qquad t \in S_2$$

where $\Sigma_2^{e0} = diag(\sigma_{10}^2, \sigma_{20}^2, \ldots, \sigma_{r0}^2)$. On the other hand, Chapter 4 assumes that the factor model remains $x_t = \Lambda^0 F_t^0 + \varepsilon_t$ in both regimes. Although it seems that the invariance of factor loadings is unrealistic given that the process of factors varies depending on regimes, we see in Chapter 4 that the Great Moderation is one example of this assumption, so it is not an unrealistic one.

Based on the assumptions and asymptotic properties of the PCA estimator, we construct identifying restrictions that link reduced form parameters with structural parameters. Let $\tilde{F}^s$ and $\tilde{\Lambda}^s = [\tilde{\lambda}_{1,s}, \tilde{\lambda}_{2,s}, \ldots, \tilde{\lambda}_{N,s}]'$ be the PCA estimators of $F^{0,s}$ and Λ, respectively, where $F^{0,s}$ is a set of the true factors in S_s. Furthermore, assume that H_s^0 is a rotation matrix in S_s. Since $\tilde{\lambda}_{i,s} - (H_s^{0\prime})^{-1}\lambda_i^0 \xrightarrow{p} 0$ holds for each i and s in a standard approximate factor model, essentially we have the restrictions[1]

$$
\begin{aligned}
&\tilde{\Lambda}^1 - \Lambda(H_1^0)^{-1} = 0, \\
&\tilde{\Lambda}^2 - \Lambda(H_2^0)^{-1} = 0.
\end{aligned}
$$

Note that the first terms are obtained from the data and the second terms are structural parameters. Then, we can estimate the structural parameters by minimum distance estimation and obtain the estimators of rotation matrices H_1^0 and H_2^0. Since $\widetilde{F}_t^s - H_s^0 F_t^0 \xrightarrow{p} 0$ holds for $s = 1, 2$ from Bai (2003), we yield the estimator of true factors by

$$
\widehat{F}_t = \begin{cases} (\hat{H}_1^0)^{-1}\,\widetilde{F}_t & \text{if } t \in S_1 \\ (\hat{H}_2^0)^{-1}\,\widetilde{F}_t & \text{if } t \in S_2 \end{cases},
$$

where $\hat{H}_1$ and $\hat{H}_2$ are the minimum distance estimators of H_1 and H_2 respectively. It should be noted that since the total number of restrictions depends on N and $N \to \infty$ is assumed in standard approximate factor models, the asymptotic properties of the minimum distance estimator are not trivial. Chapter 4 investigates the asymptotic properties of the minimum distance estimators and derives the consistency of $\hat{H}_1$ and $\hat{H}_2$.

Monte Carlo simulations confirm that our identification scheme works well and give encouraging evidences that we can precisely estimate the true factors with our estimator as expected.

[1] See also Chapter 4, which considers more restrictions.