

Fractional Cointegration Using Frequency Domain Bootstrap Methods

La Cointegrazione Frazionaria con l'Utilizzo del Bootstrap nel Dominio delle Frequenze

Margherita Gerolimetto

Dipartimento di Scienze Statistiche – Università di Padova

Via Cesare Battisti, 241- Padova

e-mail: gerolime@stat.unipd.it

Riassunto: Lo scopo di questo lavoro è di costruire una distribuzione bootstrap per lo stimatore Narrow Band Least Squares (Robinson, 1994) del parametro di cointegrazione frazionaria, utilizzando un metodo di ricampionamento bootstrap nel dominio delle frequenze invece che nel più tradizionale dominio temporale. Il bootstrap che utilizzeremo è basato parzialmente su una proposta di Hidalgo (2003) e consiste nel ricampionamento delle trasformate discrete di Fourier dei residui della regressione di cointegrazione frazionaria utilizzando il wild bootstrap (Wu, 1986). Il lavoro include un esperimento Monte Carlo per descrivere la performance del metodo.

Keywords: Fractional cointegration, bootstrap methods

1. Introduction

Bootstrap algorithms have received much attention from the literature for, perhaps, two main reasons. First, these methods can approximate the finite sample distribution of statistics in a more effective way than techniques based on their asymptotic counterparts. Secondly, the bootstrap methods make possible to compute valid asymptotic quantiles of the limiting distribution when the limiting distribution itself is unknown or, if it is known, it is impossible to compute its quantiles.

As far as time series are concerned, the so-called parametric bootstrap can be employed if the residuals of the model are *iid* and the bootstrap samples are built from the bootstrap residuals. If the errors are not *iid* the parametric bootstrap scheme can be not valid and other methods have been proposed to circumvent this problem. For instance, in case of heteroschedastic errors the wild bootstrap (Wu, 1986) can be a proper one. It consists of resampling the residuals and, at the same time, treating the heteroschedasticity. Taking the example of a simple linear regression model $y_t = \mathbf{X}_t \boldsymbol{\beta} + \varepsilon_t$ the bootstrap residuals are $\varepsilon_t^* = e_t \cdot f(e_t)$, where e_t are drawings from an external distribution such that $E(e_t)=0$ and $E(e_t^2)=1$ and f is an opportune function (Mammen, 1993).

In this work we will consider a peculiar bootstrap approach that is based on resampling the discrete Fourier transform (hereafter DFT) of the regression residuals (Hidalgo, 2003). Indeed, one interesting property of the DFT of the regression residuals is that they are asymptotically uncorrelated although, possibly, heteroschedastic among others.

So, our idea is to resample the DFT of the residuals with the wild bootstrap, which, as previously said, is supposed to be a proper one in case of heteroschedasticity.

We employed this bootstrap approach in the more and more popular context of the cointegration analysis of long memory time series, also called fractional cointegration. Fractional cointegration is a generalization of the concept of cointegration, very popular in the econometric literature of the last twenty years. Suggested intuitively in the seminal work of Engle and Granger (1987), fractional cointegration has been formally proposed for the first time by Cheung and Lai (1993) in an applied work about exchange rates time series. Given a $(N+1) \times 1$ vector $\mathbf{X}_t^{N+1} = (y_t, \mathbf{X}_t)$, whose i -th element $x_{it} \sim I(d_i)$, $d_i > 0$, $i = 1, 2, \dots, N+1$, we say that $\mathbf{X}_t^{N+1}$ is fractionally cointegrated if there exists a $N \times 1$ vector $\boldsymbol{\beta} \neq \mathbf{0}$ such that $y_t = \boldsymbol{\beta}' \mathbf{X}_t + z_t$ with $z_t \sim I(d_z)$ and $0 < d_z < \min_{1 \leq i \leq N} d_i$. This property is possible if and only if $d_i = d_j$ for some $i \neq j$, moreover a necessary condition for $\boldsymbol{\beta}$ to be a cointegrating vector is that its i -th component be equal to zero if $d_i > d_j$ for all $j \neq i$. Therefore, the traditional cointegration scheme $I(1)/I(0)$ is generalized to the fractional cointegration scheme, $I(d_i)/I(d_z)$, with $d_z < d_i$. Intuitively, fractional cointegration appears to be much more flexible than traditional cointegration, but at the same time also more complicated. Some of the major contributions to estimate consistently the fractional cointegration parameter come from Robinson (1994) and Marinucci and Robinson (2001). They considered a Narrow Band Frequency Domain Least Squares estimate (hereafter NBLSE) of the cointegration parameter, whose expression for a general fractional cointegration regression is $\beta_{im} = [\hat{F}_{x_i x_i}(m)]^{-1} [\hat{F}_{x_i y}(m)]$ where, in general, $\hat{F}_{ab}(m)$ is the averaged cross-periodogram (for further formal details, see for example Marinucci and Robinson (2001), pg. 230, formula 7). NBLSE is the ratio of the real part of average cross-periodogram of y, x_i to the average periodogram of x_i , averaging across the m lowest Fourier frequencies and Robinson proved (1994) that if m tends to infinite, but m/n tends to zero as n tends to infinite, NBLSE is consistent. Moreover, he proved that, according to the value d_i and d_z , NBLSE converges to specific non degenerate random variables with different convergence rates. All the distributions he proved the convergence to are very complicated functionals of brownian motions and so very difficult to deal with in case one wants to make inference. For this reason, in this paper we want to propose an empirical distribution as a valid alternative to these asymptotic counterparts. In the second section of this paper we are describing the bootstrap method we employed and in the third section we are presenting some results of the Monte Carlo experiment to show the performance of the bootstrap distribution in finite sample inference.

2. The Bootstrap scheme

As partially introduced in the previous section, the bootstrap scheme we are presenting follows from an idea of Hidalgo (2003). He proved the asymptotic validity of resampling the DFT's under quite mild conditions, including long memory time series. In this paper we resample the DFT of the fractional cointegration regression residuals using the wild bootstrap. We consider a simple bivariate case, where the cointegration regression is $y_t = \beta x_t + z_t$. Then the DFT form of this regression is

$$w_y(\lambda_j) = \beta w_x(\lambda_j) + w_z(\lambda_j) \quad j = 1, \dots, [n/2]. \quad (1)$$

To compute NBLSE we need to consider only the m Fourier frequencies closest to the origin. So we can also write the following

$$\sum_{j=1}^m w_y(\lambda_j) = \beta \sum_{j=1}^m w_x(\lambda_j) + \sum_{j=1}^m w_z(\lambda_j). \quad (2)$$

We are interested in testing for the null hypothesis $\beta = \beta_0$. The statistics to test for this null hypothesis are slightly different according to the value of the long memory parameters of the series. Generally speaking, we call τ the statistics we are using for each specific case and the following is the bootstrap scheme to obtain the empirical distribution for τ :

- 1) Generate the fractional cointegrated system and write down the fractional cointegration regression in DFT form as in (1).
- 2) Resample with the wild bootstrap the w_z , so

$$w_z^*(\lambda_j) = w_z(\lambda_j) \cdot \frac{e_j}{\sqrt{1-h_j}} \quad \text{and} \quad h_j = w_x(\lambda_{j-1})^2 \left(\sum_j w_x(\lambda_{j-1})^2 \right)^{-1} \quad (3)$$

where e_j is the external distribution required by the wild bootstrap and $j = 2, \dots, [n/2]$.

- 3) Obtain the bootstrap regression

$$w_y^*(\lambda_j) = \beta w_x(\lambda_j) + w_z^*(\lambda_j) \quad j = 1, \dots, [n/2]. \quad (4)$$

- 4) Compute the bootstrap NBLSE (called NBLSE*) and the statistics τ for the bootstrap sample (called τ^*).
- 5) Repeat the previous steps a large number of times and obtain the desired quantile that is the critical value we need.

3. The Monte Carlo experiment

As previously said, in this section we present some preliminary results of the Monte Carlo experiment we are carrying out to focus on the performance of the empirical distribution of NBLSE, obtained with this frequency domain bootstrap method.

Following Marinucci and Robinson (2001), the system we generate is

$$x_{1t} = \beta x_{2t} + z_t \quad \beta = 2 \quad (5)$$

$$x_{2t} = (1 - B)^{-d_{21}} u_{1t} \quad \text{and} \quad z_t = (1 - B)^{-d_z} u_{2t} \quad (6)$$

where $U_t = (u_{1t}, u_{2t})$ is a sequence of independent bivariate normal such that u_{1t} and u_{2t} have zero mean, unit variance and correlation 0.5. By the time being we have generated 2000 series with sample size $n=64,128$ for some pairs of values for (d_i, d_z) : (0,0.8); (0.2,0.7); (0.3,0.6); (0.1,0.35); (0.2,0.4). The first three pairs include non stationary variables ($d_i > 1/2$), while the second two pairs include only stationary variables ($d_i < 1/2$). The value of the bandwidth parameter m in averaging the cross-periodogram is set to

4,6,8. We are testing for the null hypothesis $\beta = 2$ and so in the following table we can see the percentages of rejection of the null hypothesis using the bootstrap scheme we presented in the previous section. The nominal level is 0.05.

Table 1: Percentages of rejections of the null hypothesis $\beta = 2$

d_x, d_z	$n=64$			$n=128$		
	$m=4$	$m=6$	$m=8$	$m=4$	$m=6$	$m=8$
(0,0.8)	0.06	0.062	0.072	0.058	0.059	0.071
(0.2,0.7)	0.064	0.065	0.074	0.055	0.06	0.069
(0.3,0.6)	0.061	0.062	0.08	0.06	0.057	0.063
(0.1,0.35)	0.073	0.056	0.082	0.067	0.054	0.07
(0.2,0.4)	0.068	0.058	0.079	0.06	0.055	0.072

At the moment these are only preliminary results because many other simulations need to be done. However, we can already see some important points. The discrepancies between the nominal level and the percentages of rejection are reasonably small and they tend to decrease with the increasing of the sample size. In particular, we can notice that the empirical probabilities of rejection are very close to the nominal level even at the lowest sample size ($n=64$). As the bandwidth parameter m is concerned, we can notice a better performance when $m=4,6$ for the first three pairs of series and when $m=6$ the second two pairs of series. The differences are much less evident when $n=128$, even though the results are still slightly better when $m=4,6$. These preliminary results encourage us to go on with the experiment. Indeed, the obtained percentages appear to be close enough to the nominal level to probably justify the use of the empirical distribution (especially at the lowest sample size) instead of the asymptotic ones which are non standard and very difficult to derive.

References

- Cheung Y., Lai K. (1993) A fractional cointegration analysis of purchasing power parity. *Journal of Business and Economic Statistics*, 11, 103-112.
- Granger C.W.J., Joyeux R. (1980) An introduction to long memory time series models and fractional differencing. *Journal of time series analysis*, 1,15-39.
- Engle R.F., Granger C.W.J. (1987) Co-integration and error correction representation, estimation and testing. *Econometrica*, 55, 251-276.
- Hidalgo J. (2003) An alternative bootstrap to moving blocks for time series regression models. STICERD econometrics publication n. EM/2003/452.
- Marinucci D., Robinson P.M. (2001) Semiparametric fractional cointegration analysis. *Journal of econometrics*, 105, 225-247.
- Robinson P.M. (1994) Semiparametric analysis of long memory time series. *The Annals of Statistics*, 22, 515-539.
- Wu C.F.J. (1986) Jackknife, bootstrap and other resampling methods in regression analysis (with discussion). *Annals of Statistics*, 14,1261-1295.
- Mammen E. (1993) Bootstrap and wild bootstrap for high - dimensional linear models. *Annals of Statistics*, 21, 255 - 285.