THE LONG-RUN ECONOMIC RELATIONSHIPS: AN OPTIMIZATION APPROACH TO FRACTIONAL INTEGRATED AND BILINEAR TIME SERIES

Jonas MOCKUS

Institute of Mathematics and Informatics
2600 Vilnius, Akademijos St. 4, Lithuania

Abdol S. SOOFI

Department of Economics, University of Wisconsin
Platteville, WI 53818

Abstract. One objective of this paper is to estimate the parameters p,d,q of an autoregressive fractionally integrated moving average $\text{ARFIMA}(p,d,q)$ stochastic model by minimizing the squares of the residuals using a Bayesian global optimization techniques. We consider bilinear model, too because it is the simple extension of linear model, defined by adding a bilinear term to traditional ARMA model. Therefore, the second objective of the paper is to estimate parameters of a bilinear time series.

Key words: autoregressive, fractionally integrated, bilinear, global optimization.

1. Introduction. Existence of long-run relationships among relevant economic variables is very important for at least two major reasons. First, presence of a long-run relationship among the variables of a model excludes the possibility of spurious co-movement among the variables (Granger and Newbold, 1974). Second, for policy analysis it is highly desirable to know whether there exists some exploitable connection between the instrument and target variables.

The development of co-integration methods by Engle and Granger (1987) and Johansen (1988, 1991) has given economists opportunities to explore the existence of long-run relationships among economic variables. Since the advent of these techniques, an extensive literature has emerged which deals with applications of these methods in various fields of economics. These researchers have attempted to exploit the following property of two co-integrating series.

Granger (1986) and Engle and Granger (1987) have shown that two co-integrating series tend to move in proximity of each other over the long-run. This implies that for two co-integrating series, the equilibrium error Z_t is mean-reverting in spite of wandering behavior of the series y_{1t} and y_{2t}.

Note that in the absence of the mean-reversion property, and even in cases where economic theory predicts existence of long-run equilibrium, a shock to the system can force two time series into disequilibrium permanently.

Co-integration methods, however, rest on binary choice of stationary or integrated of degree one time series in examining mean-reversion property of the residuals of the models. This approach excludes a class of long-memory stochastic processes with the fractional co-integrating property which also have mean-reverting characteristics. Due to this and other attractive properties of the ARFIMA models, empirical research using the models has flourished in recent years (Diebold and Rudebush, 1989, 1991; Cheung, 1993; Cheung and Lai, 1993; Koop *et al.*, 1994).

The coefficient estimation of the ARIFMA models is an important part of the fractional co-integrated problems. In estimating the coefficients of the ARFIMA models, three approaches have been used: Maximum likelihood (ML) (Sowell, 1992); approximate ML (Li and McLeod, 1986; Fox and Taqque, 1986); and two-step procedures (Geweke and Porter-Hudak, 1983; Janacek, 1982). Geweke and Porter-Hudak's method, unlike the ML approach, is less computationally demanding but is considered inadequate for finite sample. In all the cases local optimization techniques were used.

Recent advances in global optimization techniques (Mockus, 1994) opens new possibilities. Theoretically we may apply the global optimization algorithms to the traditional ML problem. However, for parsimony in computation, local optimization is usually applied (Sowell, 1992). In this case the optimization results depend on the initial values. It means that one cannot be sure if a global maximum is found.

Theoretically, we may get much better approximation to the global optimum using the advanced techniques of global optimization (Horst and Pardalos, 1995; Mockus, 1994), Torn and Žilinskas, 1989; and Žilinskas and Zygliavski, 1992). However, due to problems associated with optimization of polynomial-time computable real functions (Ko, 1991), the global optimization is very hard in almost all the cases.

An approximate approach in estimating the parameters of the ARFIMA models is Least Squares (LS). We minimize the sum of square residuals instead of maximizing log-likelihood. This way is less attractive theoretically: we are not sure about the asymptotics, we do not estimate the covariances, etc.

Therefore, the consensus is that the maximal likelihood estimates are, theoretically, superior to the least square ones. This opinion is based on the assumption that we may get the exact likelihood maximum. Otherwise, the only remaining advantage of likelihood maximization is the covariances estimation. Therefore, if the covariances are not to be had by other means, then we have to use the likelihood maximization. Otherwise, it would be difficult to justify not applying the simple, more convenient squared residuals minimization method (instead of likelihood maximization)[1]

One objective of this paper is to estimate the parameters p, d, q of an autoregressive fractionally integrated moving average ARFIMA(p, d, q) stochastic model by minimizing the squares of the residuals using a Bayesian global optimization method (Mockus, 1994)[2]

The linear time series models[3] describe many economic processes well enough. However there could be cases when non-linearities cannot be ignored. We think that the bilinear time series may be a useful tool in describing the long-term behavior of the processes that generate international financial data.

We consider bilinear model (16) because it seems to be the simplest extension of linear model, defined by adding term (17) to traditional ARMA model.

Therefore, the second objective of the paper is to estimate parameters of a bilinear time series.

In the way of an example, we will test whether ARFIMA and bilinear models describe the rial-to-dollar black market monthly exchange rate 1965–1988, see Fig. 1.

---

[1] If we do not need covariance estimates, we may regard it as a nuisance parameter that we ignore defining the residuals. If we do not need the expectation estimates, we may ignore those, too. Otherwise, we my include the expectations while defining the residuals.

[2] The estimation of covariance matrix Σ is not the main objective in this investigation.

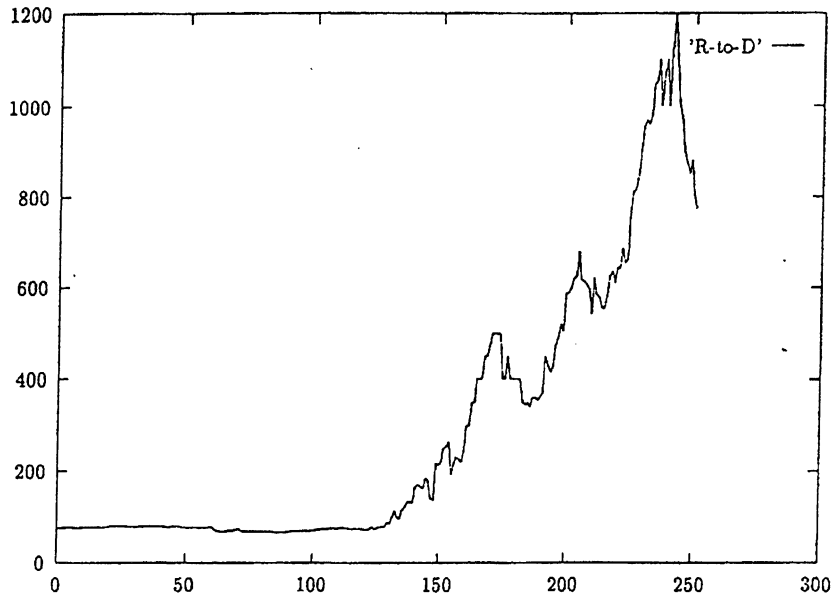[3] For example, ARFIMA(p,d,q).

Fig. 1. Black market rial-to-dollar monthly exchange rate 1965–1988.

2. Linear models

2.1. Definitions. We define ARIMA(p, d, q) process as a time series

$$A(L)(1 - L)^d z_t = B(L)\varepsilon_t. \tag{1}$$

Here

$$A(L)w_t = w_t - \sum_{i=1}^{p} a_i w_{t-i}, \tag{2}$$

and

$$B(L)\varepsilon_t = \varepsilon_t - \sum_{i=1}^{q} b_i \varepsilon_{t-i}, \tag{3}$$

where ε_t = Gaussian $(0, \sigma^2)$.

We define the transformation $(1 - L)^d$ as follows:

$$w_t = (1 - L)^d z_t = z_t - \sum_{i=1}^{\infty} d_i z_{t-i}. \tag{4}$$

Here

$$d_i = \frac{\Gamma(i - d)}{\Gamma(i + 1)\Gamma(-d)}, \tag{5}$$

where d is a fractional integration parameter.

We assume, that

$$z_{t-i} = 0, \quad w_{t-i} = 0, \quad \varepsilon_{t-i} = 0, \quad \text{if } t \leqslant i. \tag{6}$$

We truncate sequence (4)

$$d_i = 0, \quad \text{if } i > R \tag{7}$$

Here R is the truncation parameter[4]

2.2. Likelihood maximization. We define the log-likelihood function of ARIMA(p, d, g) process as follows:

$$\log \; L(z; \mu, \Sigma) = -\frac{1}{2} \; \log \; |\Sigma| - \frac{1}{2} \, (z - \mu)^T \Sigma^{-1} (z - \mu). \tag{8}$$

Here Σ is the covariance matrix and μ is expectation vector of z.

McLeod and Hipel (1978) used direct maximization of (8) to estimate the parameters of a fractional Gaussian model. To simplify the calculations, Hosking (1984) used the approximate optimization techniques, maximizing the log-likelihood function (8).

2.3. Residuals minimization. One of the advantages of residual minimization, as compared with log-likelihood maximization, is that we may see directly how the objective depends on the unknown parameters. We define residuals by recurrent expressions:

$$\varepsilon_1 = w_1 \tag{9}$$

$$\varepsilon_2 = w_2 - a_1 w_1 + b_1 \varepsilon_1$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$\varepsilon_t = w_t - a_1 w_{t-1} - \ldots - a_p w_{t-p} + b_1 \varepsilon_{t-1} + \ldots + b_q \varepsilon_{t-q}. \tag{10}$$

We minimize the sum

$$f_m(x) = \sum_{t=1}^{T} \varepsilon_t^2. \tag{11}$$

---

[4] A number of non-zero components.

Here the objective $f_m(x)$ depends on $m = p + q + 1$ unknown parameters that we represent as the m-dimensional vector $x = (x_k, \ k = 1, \ldots, m) = (a_i, \ i = 1, \ldots, p, \ b_j, \ j = 1, \ldots, q, d)$.

We see from (10), (4), and (2) that residuals ε_t are linear functions of parameters a_t. It means that the minimum conditions

$$\frac{\partial f_m(x)}{\partial a_i} = 0, \ i = 1, \ldots, p \qquad (12)$$

are given by a system of linear equations that defines the estimates of parameters $a_i = a_i(b_i, \ i = 1, \ldots, q, d)$ as a function of parameters $b_i, \ i = 1, \ldots, q, d$. It reduces the number of parameters of nonlinear optimization to $n = q + 1$.

The system

$$\frac{\partial f_m(x)}{\partial b_i} = 0, \ i = 1, \ldots, q \qquad (13)$$

may have a multiple solution, because the residuals ε_t depends on b_i as polynomials of degree $T - 1$.

The equation

$$\frac{\partial f_m(x)}{\partial d} = 0 \qquad (14)$$

also may have multiple solutions, because the residuals depend on d as a polynomial of degree R, where R is a truncation parameter.

It means that, in general, the objective $f_m(x)$ is a multi modal function of parameters d and $b_i, \ i = 1, \ldots, q$. [5] Therefore, we have to consider the methods of global optimization (see, for example Mockus, 1994).

We denote

$$f(x) = f_m(x), \quad x_i = a_i(b_i, \ i = 1, \ldots, q, d), \quad i = 1, \ldots, p. \qquad (15)$$

It means that we define by condition (12) those x-components that represent the parameters $a_i, \ i = 1, \ldots, p$.

We see that the variance σ^2 is not in residuals expressions (15) and (10). If required, we have to estimate the variance by some other well known techniques.

3. Bilinear models. It is well known that for the adequate description of some phenomena, one needs nonlinear time series. A large number of empirical

---

[5] The same reasoning applies to log-likelihood function, too.

works have tested for the nonlinearity of macroeconomics and financial time series (Brock and Potter, 1993). The simple case of the nonlinear models is bilinear time series (Subba Rao and Gabr, 1984; Liu, 1989).

An example of bilinear time series follows:

$$A(L)z_t = B(L)\varepsilon_t + C(L)z_t\varepsilon_t, \tag{16}$$

where

$$C(L)z_t\varepsilon_t = \sum_{i=1}^{s}\sum_{j=1}^{r} c_{ij}z_{t-i}\varepsilon_{t-j}. \tag{17}$$

A bilinear extension of linear ARIMA model (1) follows

$$A(L)(1 - L)^d z_t = B(L)\varepsilon_t + C(L)z_t\varepsilon_t. \tag{18}$$

We expect that this model may represent both the persistence and nonlinearity well enough.

Consider an example of bilinear time series (16) with $p = 2$, $q = 1$, $s = 2$, $r = 1$:

$$z_t - a_1 z_{t-1} a_2 z_{t-2} + \varepsilon_t + c_{1,1}z_{t-1}\varepsilon_{t-1} + c_{2,1}z_{t-2}\varepsilon_{t-1}, \quad t = 1, \ldots, T. \tag{19}$$

We consider a simple case when $T = 12$, ε_t are Gaussian $0, 1$, and $a_1 = 0.8$, $a_2 = -0.4$, $c_{1,1} = 0.6$, $c_{2,1} = 0.7$. The mean square deviation $f(x) = \sum_{i=1}^{12}\varepsilon_t$.

Fig. 2 shows how $f(x)$ depends on $c_{2,1}$. In Fig. 2 $d(x^4)$ denotes a mean square deviation $f(x)$, and x^4 denotes variable parameter $c_{2,1}$.

4. Optimization and fractional co-integration.

Define two $I(d)$ series y_{1t} and y_{2t} as $y_t = (y_{1t}, y_{2t})$. We say that series y_{1t} and y_{2t} are co-integrated of order (d, b) if there exists a vector α such that

$$z_t = \alpha y_t \tag{20}$$

is $I(d - b)$ with $b > 0$.

Considering the fractional co-integration problem, we may estimate unknown parameters d and b at different fixed values of the vector α by the residual minimization techniques. Then we may pick up such α that maximizes the difference b. That is one way to apply optimization in the fractional co-integration problems. We expect to investigate different ways.
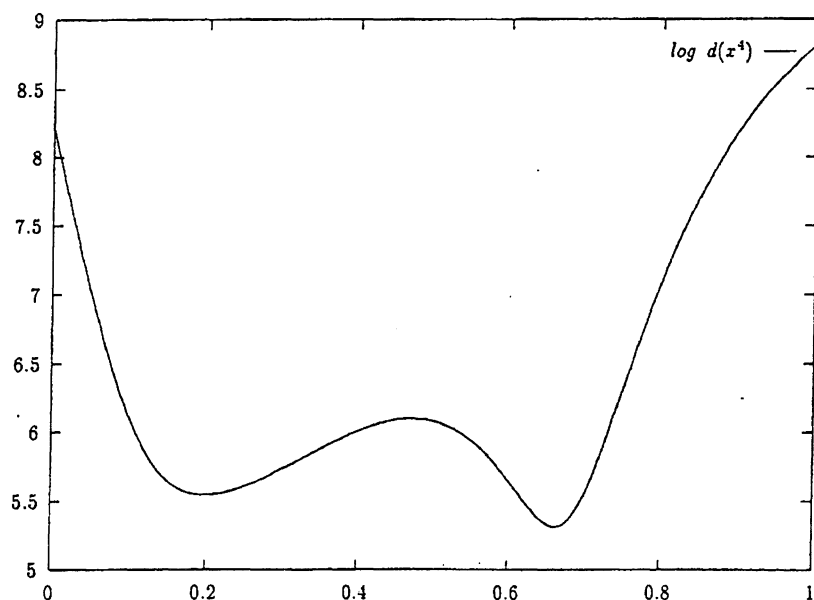
The long-run economic relationships

Fig. 2. Bilinear time series, residuals minimization.

REFERENCES

Brock, W.A., and S.M. Potter (1993). Macro econometrics. In Maddala, G.S., C.R.Rao, and H. D. Vinod (Eds.), *Handbook of Statistics*, Vol. 11. Elsevier Science Publishers, Amsterdam.

Granger, C.W.J, and R. Joyeux (1980). An introduction to long-memory time series models and fractional differencing. *Journal of Time Series Analysis*, 1, 15–39.

Cheung Yin-Wong (1993). Long memory in foreign-exchange rates. *Journal of Business and Economic Statistics*, 1, 93–101.

Cheung Yin-Wong and K. Lai (1993a). Long-run purchasing power parity during the recent float. *Journal of International Economics*, 34, 181–192.

Cheung Yin-Wong and K. Lai (1993b). Fractional co-integration analysis of purchasing power parity. *Journal of Business and Economic Statistics*, 1, 103–112.

Diebold, F.X., and G.D. Rudebusch (1989). Long memory and persistence in aggregate output. *Journal of Monetary Economics*, 24, 189–209.

Diebold, F.X., and G.D. Rudebusch (1991). Is consumption too smooth? Long memory and the deaton paradox. *Review of Economics and Statistics*, 73, 1–9.

Engle, R.F., and W.A. Granger (1987). Co-integration and error correction: representation, estimation, and Testing. *Econometrica*, 55, 251–276.

Fox, R., and M. Taqqu (1986). Large-sample properties of parameter estimates for strongly dependend stationary gaussian time series. *Annals of Statistics*, 14, 517–532.

Geweke, J., and S. Porter-Hudak (1983). The estimation and application of long memory time series models. *Journal of Time Series Analysis*, 4, 221–238.

Granger, C.W.J. (1986). Developments in the study of co-integrated economic variables. *Oxford Bulletin of Economics and Statistics*, 48, 213–228.

Granger, C.W.J., and P. Newbold (1974). Spurious regressions in econometrics. *Journal of Econometrics*, 2, 111–120.

Granger, C.W.J., and R. Joyeux (1980). An introduction to long-memory time series models and fractional differencing. *Journal of Time Series Analysis*, 1, 15–39.

Horst, R., and P. Pardalos (1995). *Handbook of Global Optimization*. Kluwer, Dordrecht/Boston/London.

Hosking, J.R.M. (1981). Fractional differencing. *Biometrika*, 68, 165–176.

Hosking, J.R.M. (1984). Modeling persistence in hydrological time series using fractional differencing. *Water Resources Research*, 20, 1898–1908.

Janacek, G. (1982). Determining the degree of differencing for time series via thelLong spectrum. *Journal of Time Series Analysis*, 3, 177–188.

Johansen Soren (1988). Statistical analysis of co-integration vectors. *Journal of Economic Dynamics and Control*, 12, 231–254.

Johansen Soren (1991). Estimation and hypothesis testing of co-integration vectors in Gaussian vector autoregressive models. *Econometrica*, 39, 1551–1580.

Koop G., *et al.* (1994). Bayesian analysis of long memory and persistence using ARFIMA models. *Working Paper*, Department of Economics, University of Toronto.

Li, W.K., and A.I. McLeod (1986). Fractional time series modelling. *Biometrika*, 73, 217–221.

Liu, J. (1989). On the existence of general multiple bilinear time series. *Journal of Time Series Analysis*, 10, 341–355.

Lutkepohl Helmut (1991). *Introduction to Multiple Time Series Analysis*. Springer-Verlag, New York.

Ko Ker-I. (1991). *Complexity Theory of Real Functions*. Birkhauser, Basel.

McLeod, A.I., and K.W. Hippel (1978). Preservation of the rescaled adjusted range. A reassessment of the hurst phenomenon. *Water Resource Research*, 14, 491–508.

Mockus A., J. Mockus and L. Mockus (1994). Bayesian approach adapting approximate, stochastic, and heuristic optimization methods. *Informatica*, 5(1–2), 123–166.

Mockus, J. (1994). Application of Bayesian approach to numerical methods of global and stochastic optimization. *Journal of Global Optimization*, 4, 347–365.

Sowell, F. (1992). Maximum likelihood estimation of stationary univariate fractionally integrated models. *Journal of Econometrics*, 53, .165–188.

Subba Rao, T., and M.M. Gabr (1984). *An Introduction to Bispectral Analysis and Bilinear Time Series Models*. Springer-Verlag, New York.

Torn, A., and A. Zilinskas (1989). *Global optimization*. Lecture Notes in Computer Science # 350 Springer-Verlag, Berlin.

Zilinskas, A., and A. Zygliavsky (1992). *Methods of Global Optimization*. Nauka, Moscow (in Russian)..

J. Mockus graduated Kaunas Technological University, Lithuania in 1952. He got Doctor habilitus degree in the Institute of Computers and Automation, Latvia, in 1967. He is a head of Optimal Decision Theory Department, Institute of Mathematics and Informatics, Vilnius, Lithuania and professor of Kaunas Technology University.

A.S. Soofi is an associate Professor of Economics at the University of Wisconsin-Platteville. He received his Ph.D. from the Department of Economics the University of California, Riverside in 1980. His areas of research interest include excange-rate economics, time series analysis, and input-output modeling.

ILGALAIKIAI EKONOMINIAI SANTYKIAI: OPTIMIZAVIMO POŽIŪRIS TRUPMENINIAI INTEGRUOTOSE IR DVITIESINĖSE LAIKO EILUTĖSE

Jonas MOCKUS, Abdol S. SOOFI

Nagrinėjami trupmeniniai ko-integruoti uždavinai, kurie svarbūs tiriant ilgalaikius ryšius tarp susijusių ekonominių kintamųjų. Numatoma panaudoti naujus globalinio optimizavimo metodus įvertinti tiesinių ARIFMA modelių parametrus.

Nagrinejami taip pat ir kai kurie netiesiniai modeliai, konkrečiai bi-tiesiniai modeliai. Šių modelių pavyzdžiu parodoma, kad vidutinis kvadratinis nukrypimas gali turėti daug minimumų.

Kaip pirmas pavyzdys nagrinėjama Irano rialo ir JAV dolerio kurso laiko eilutė nuo 1965 iki 1988 metų.