Penalty terms for estimation of ARMA models: A Bayesian inspiration

Helgi Tómasson

University of Iceland, Faculty of Economics,
Oddi v/Sturlugötu, IS-101 Reykjavík, Iceland

Abstract. Bayesian methods are based on combining a problem, a model, prior information, and data using Bayes rule. This paper addresses the implementation of a Bayesian approach to stationary ARMA models. The interpretation of the parameters of a ARMA models is somewhat non-intuitive. The interpretation of the spectral function is much clearer. A Bayesian expression of a prior belief in the frequency domain, i.e., stating a preference on the shape of the spectral function, may therefore be more natural than formulating a prior on the time-domain parameters. Stating a prior on a function space is non-trivial. In this paper the fact that the spectral density of an ARMA model is rational is exploited. The use of complex theory residue calculus is used to derive analytic measures of desirable features of the spectral function. The approach is equally suited for discrete- and continuous-time models.

Keywords: Residue calculus, rational spectrum, ARMA

1 Introduction

Bayesian methods are based on combining a problem, a model, prior information and data using Bayes rule. This paper addresses the implementation of a Bayesian approach to stationary ARMA (Auto-Regressive-Moving-Average), and continuous-time-ARMA (CARMA), models. The parameters of ARMA and CARMA models are essentially a parsimonious way of modelling an auto-correlation function. In general correlations, and in particular an auto-correlation function (ACF), and the ARMA-parameters, of a stationary process are hard to interpret. Therefore, defining a sensible a priori opinion about correlations is a non-trivial issue.

In time-series analysis the spectral curve, the Fourier transform of the ACF is much easier to interpret. Expressing a prior opinion in the frequency domain is therefore a more natural approach. However, the spectral density is a continuous function and operating with a probability distribution on a function space is difficult.

A natural Bayesian approach is to state the prior in the frequency domain, i.e., that *a priori*, the spectral density is of a particular type, or is in some sense "close" to a particular spectral density. The concept "close" requires some measure of distance between spectral densities. A conceivable measure of closeness between two curves is the Kullback-Leibler distance.

Calculation of a Bayes estimator can often be implemented as a frequentistic estimator with an added penalty term. The penalty term biases the classical estimator towards a prior. In this case, a prior of a smooth spectrum is illustrated. The technique is based on the fact that the spectrum of the ARMA model is rational. The residue calculus of complex analysis gives exact expressions of some integrals of rational functions. In particular a measure of smoothness, e.g., the integral of the squared second derivative of the spectral density can be calculated directly. A penalty term based on a function of this measure can then be added to an objective function, e.g., a log-likelihood function. Then this improved objective function can be used to shrink the fitted model towards a priori ideas of the spectral shape. This approach can be modified implement other forms of a priori information on the spectral function.

This paper is organized as follows. First a brief review of the continuous-time ARMA model is given. Section 3 shows an intuition of a Bayesian approach and the interpretation of the prior term in the likelihood function as a penalty term in classical estimation. Section 4 reviews mathematical results on partial fractions and residue calculus that are useful for calculation of some penalty terms of interest. In section 5 the computational machinery for a continuous-time ARMA implemented in an R-package is illustrated. Section 6 concludes with a discussion on extending these ideas to discrete-time models, other types of penalty functions, comparison with other types of penalty terms like AIC and BIC.

2 On ARMA models

A continuous-time ARMA, CARMA, process can be defined in terms of a continuous-time innovation process and a stochastic integral. A common choice of innovation process is the Wiener process, $W(t)$. A representation of a CARMA(p,q) process in terms of the differential operator D is:

$$\begin{gathered}
Y^{(p)}(t) + \alpha_1 Y^{(p-1)}(t) + \cdots + \alpha_p Y(t) = \\
\sigma d(W(t) + \beta_1 W^1(t) + \cdots + \beta_q W^{(q))}(t)), \\
\text{or } \boldsymbol{\alpha}(D)\, Y(t) = \sigma \boldsymbol{\beta}(D)\, dW(t), \\
\boldsymbol{\alpha}(z) = z^p + \alpha_1 z^{p-1} + \cdots + \alpha_p, \\
\boldsymbol{\beta}(z) = 1 + \beta_1 z + \cdots + \beta_q z^q.
\end{gathered}$$

Here, $Y^{(p)} = D^p Y(t)$, denotes the p-th derivative of $Y(t)$. The path of a Wiener process is nowhere differentiable so the symbol $D\,W(t)$, and higher derivatives, is of a purely notational nature. The spectral density of $Y(t)$ is a rational function:

$$f(\omega) = \frac{\sigma^2}{2\pi} \frac{\boldsymbol{\beta}(i\omega)\boldsymbol{\beta}(-i\omega)}{\boldsymbol{\alpha}(i\omega)\boldsymbol{\alpha}(-i\omega)}.$$

The spectral representation of CARMA is:

$$Y(t) = \int_{-\infty}^{\infty} \exp(i\omega\, t) dZ(\omega),$$
$$E(dZ(\omega)) = 0, \quad E(dZ(\omega)\overline{dZ(\omega)}) = f(\omega)d\omega,$$
$$E(dZ(\omega)\overline{dZ(\lambda)}) = 0, \quad \lambda \neq \omega.$$

In mathematics an univariate linear dynamic system can be expressed as a linear differential equation of a particular order. This can be written as a multidimensional first order differential equation. In the state space form, the AR part of CARMA represents a linear differential equation. Just as in the discrete-time ARMA has several possible state-space representations, the continuous-time CARMA has several possible state-space representations. See, e.g, Tsay (2010) for the discrete-time case, and Bergstrom (1988) and Zadrozny (1988) for the continuous-time case.

The stationarity condition of the CARMA requires the roots of the polynomial $\boldsymbol{\alpha}(z)$ to have negative real-parts and that $p > q$.

3 Intuition

A standard Bayesian approach for the normal model is to assign a normal prior for the mean, e.g,:

$$X|\mu, \sigma \sim N(\mu, \sigma^2),$$
$$\mu|\sigma \sim N(\mu_0, \sigma_0^2), \quad \sigma_0 = \tau\sigma.$$

Given data, $x_1, \ldots, x_n$, and reparameterizing, $v = 1/\sigma$, $v_0 = k_0 v$, the log-posterior is (σ known):

$$log(p(\mu, v|x_1, \ldots, x_n, \mu_0, k_0)) =$$
$$\text{constant} + \underbrace{\frac{n}{2}\log(v) - v\sum_{i=1}^{n}(x_i - \mu)^2/2}_{A} + \underbrace{\frac{1}{2}\log(v) - k_0 v(\mu - \mu_0)^2/2}_{B}.$$

The number k_0 expresses the certainty in the prior. If k_0 is set to zero and the log-posterior (as a function of μ) is maximized the result is the ML estimator and a nonzero k_0 biases the ML-estimate towards the prior (μ_0). The mode of the posterior can serve as a Bayesian point estimate. The objective function is just the likelihood function, A, with an extra "penalty term", B, penalizing for deviations from the central value of the *prior*-distribution. Penalty terms are commonly added to the log-likelihood functions as a model selection tool. Well known examples are AIC and BIC.

The statistical analysis of a ARMA is essentially estimating a correlation structure based on one observations of a particular vector $\boldsymbol{y} = (y(t_1), \ldots, y(t_n))$.

The parameters of the ARMA are, a scale parameter σ, and a set of parameters that decide the auto-correlation function, $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$. I.e:

$$\boldsymbol{y} \sim N(\boldsymbol{0}, M(\boldsymbol{\alpha}, \boldsymbol{\beta})\sigma^2).$$

The likelihood (with $v = 1/\sigma$) is then:

$$L(v, \boldsymbol{\alpha}, \boldsymbol{\beta}|\boldsymbol{y}) \propto v^{\frac{n}{2}} |M(\boldsymbol{\alpha}, \boldsymbol{\beta})|^{-1/2} e^{-v\boldsymbol{y}' M(\boldsymbol{\alpha},\boldsymbol{\beta})^{-1}\boldsymbol{y}/2}.$$

Using a gamma prior for v,

$$v \sim gamma(a, b),$$

yields an analytical form of the posterior for v,

$$p(v|\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{y}) \propto v^{\frac{n}{2}} |M(\boldsymbol{\alpha}, \boldsymbol{\beta})|^{-1/2} e^{-v\boldsymbol{y}' M(\boldsymbol{\alpha},\boldsymbol{\beta})^{-1}\boldsymbol{y}/2} v^{a-1} e^{-bv},$$

i.e. the posterior is:

$$gamma(a + n/2, b + \boldsymbol{y}' M(\boldsymbol{\alpha}, \boldsymbol{\beta})^{-1}\boldsymbol{y}/2).$$

The parameters, $\boldsymbol{\alpha}, \boldsymbol{\beta}$, define the ACF and the shape of the spectral form. In analogy with the normal mean approach is to add a penalizing term to the likelihood. A natural choice is to multiply the likelihood with a function, h, of the distance between the spectral function, $f(\boldsymbol{\alpha}, \boldsymbol{\beta})$ and a *a priori* spectral function $f(\boldsymbol{\alpha}_0, \boldsymbol{\beta}_0)$. The posterior can be of the form:

$$p(\boldsymbol{\alpha}, \boldsymbol{\beta}, \sigma|\boldsymbol{y}) \propto L(\boldsymbol{\alpha}, \boldsymbol{\beta}, \sigma|\boldsymbol{y}) v^{a-1} e^{-bv} h(KLD(f(\boldsymbol{\alpha}, \boldsymbol{\beta}, \sigma), f^*(\boldsymbol{\alpha}_0, \boldsymbol{\beta}_0))).$$

Here KLD is the Kullback-Leibler distance from spectral density f to the function f^*.

$$KLD(f, f^*) = \int_{-\infty}^{\infty} \log(\frac{f(\omega)}{f^*(\omega)}) f(\omega) \mathrm{d}\omega.$$

Here:

$$f(\omega) = \frac{\beta(i\omega)\beta(-i\omega)}{\alpha(i\omega)\alpha(-i\omega)},$$
$$f^*(\omega) = \sigma^* \frac{\beta_0(i\omega)\beta_0(-i\omega)}{\alpha_0(i\omega)\alpha_0(-i\omega)},$$

where σ^* is chosen such that,

$$\int_{-\infty}^{\infty} f(\omega) \mathrm{d}\omega = \int_{-\infty}^{\infty} f^*(\omega) \mathrm{d}\omega.$$

Here the penalty term is based on some kind of a distance measure between the spectral shape defined by the ARMA parameters and an a priory defined spectral shape. In regression it can be based on Ridge-regression or empirical

Bayes approaches such as James-Stein. These penalty terms can be motivated with Bayesian arguments. Technically the ordinary maximum-likelihood is replaced with a penalized maximum-likelihood. The idea is to put a preference on simpler models. The AIC, BIC and the R^2-adjusted all penalize by using a function of the number of estimated parameters. The number of parameters is not always the natural way of grading complexity. In regression it seems reasonable that the model:

$$y = a + bx + e,$$

is simpler than:

$$y = \sin(\cos(ax))^a \exp(-bx)/x^b + e.$$

The ARMA(1,0) model:

$$dY + Y = dW, \quad ,$$

is actually the same as:

$$Y^{(4)} + 4Y^{(3)} + 6Y^{(2)} + 4Y^{(1)} + Y = d(W + 3W^{(1)} + 3W^{(2)} + W^{(3)}.$$

That is the ARMA(1,0) is a special case of (many) ARMA(4,3) models. Estimation of six additional parameters might result in a spectral function with an unreasonable shape. However, it might be of interest to estimate a model which is more complicated than an AR(1). One might, however, want restrict the freedom of the additional parameters.

In time-series analysis, just as in non-parametric regression a smoothness restriction may be enforced on the fitted values. That is the sharp spikes and turns are penalized. In economics a well known procedure of this type is the Hodrick-Prescott filter.

In stationary time-series analysis a natural form of a priori information might consist of a specification of the spectral function or some features of the spectral function. In analogy with the Hodrick-Prescott filter one can introduce a term that penalizes for sharp spikes and turns, e.g., a term proportional to:

$$\int_{-\infty}^{\infty} (f''(\omega))^2 d\omega.$$

For treatment of smoothness prior concepts in time-series analysis, see, e.g., Kitagawa & Gersch (1996).

4 Some computational aspects

A characteristic feature of the spectral density for an ARMA model is that it is a rational function.

$$f(\omega) = \frac{\sigma^2}{2\pi} \frac{\boldsymbol{\beta}(i\omega)\boldsymbol{\beta}(-i\omega)}{\boldsymbol{\alpha}(i\omega)\boldsymbol{\alpha}(-i\omega)} = \frac{\sigma^2}{2\pi} \frac{\prod_{j=1}^{q}(1 + \mu_j^2\omega^2)}{\prod_{j=1}^{p}(\omega^2 + \lambda_j^2)}. \tag{1}$$

Where the λ_j's are the roots of the polynomial $\boldsymbol{\alpha}(z)$, and the μ_j's are the reciprocals of the roots of $\boldsymbol{\beta}(z)$. In the case where the roots of the polynomial $\boldsymbol{\alpha}(z)$ are distinct (and different from the roots of $\boldsymbol{\beta}(z)$, a partial fraction expression of $f(\omega)$ can be given by:

$$f(\omega) = \frac{\sigma^2}{2\pi}(\frac{a_1}{\omega - i\lambda_1} + \cdots \frac{a_p}{w - i\lambda_p} + \frac{b_1}{\omega + i\lambda_1} + \cdots \frac{b_p}{\omega + i\lambda_p}), \tag{2}$$

and another can be given by

$$f(\omega) = \frac{\sigma^2}{2\pi}(\frac{c_1}{\omega^2 + \lambda_1^2} + \cdots + \frac{c_p}{\omega^2 + \lambda_p^2}). \tag{3}$$

In the case of some roots of $\boldsymbol{\alpha}(z)$ being equal, terms of the type $1/(\omega - i\lambda_k)^{m_k}$, where m_k is the multiplicity of the root λ_k, will be present in (2). Both forms of partial fractions are convenient. e.g., the variance due to the frequency interval (ω_1, ω_2) will be given by:

$$\int_{\omega_1}^{\omega_2} f(\omega)d\omega = \frac{\sigma^2}{2\pi}\sum_{j=1}^{p} 1/|\lambda_j|(\arctan(\omega_2/|\lambda_j|) - \arctan(\omega_1/|\lambda_j|)).$$

The residue calculus of complex analysis offers a useful tool for calculating integrals of rational functions. The residue theorem states that

$$\int h(x)dx = 2\pi i \sum Res(h(z)), \quad \text{over a certain path,}$$

where the sum is evaluated over the residues of the function h. For details see e.g., Kreyszig (1999). The theoretical auto-covariance function, $\gamma(\tau) = E(Y(t)Y(t-\tau))$, can be derived by residue calculus:

$$\gamma(\tau) = \int_{-\infty}^{\infty} e^{i\tau\omega} f(\omega)d\omega. \tag{4}$$

As the realpart of λ_j is negative the term $(\omega - i\lambda_j)$ has a pole in the upper negative half-plane, each term in (4) is readily derived:

$$\int_{-\infty}^{\infty} \frac{a_j e^{i\tau\omega}}{\omega - \lambda_j} d\omega = a_j e^{-\lambda_j|\tau|}.$$

Partial fraction of the spectral density can be useful for a variety of descriptive features. E.g. one can define a measure of smoothness:

$$\int_{-\infty}^{\infty} (f''(\omega))^2 d\omega. \tag{5}$$

The expression (5) can be derived directly from (2) by observing that:

$$f''(\omega) = \frac{\sigma^2}{2\pi}(\frac{2a_1}{(\omega - i\lambda_1)^3} + \cdots \frac{2a_p}{(w - i\lambda_p)^3} + \frac{2b_1}{(\omega + i\lambda_1)^3} + \cdots \frac{2b_p}{(\omega + i\lambda_p)^3}),$$

$f''(\omega)^2$ will contain p terms of the type $a_j/(\omega - i\lambda_j)^6$ and p terms $b_k/(\omega + i\lambda_k)^6$ and $p(p-1)$ terms, $k \neq j$, of the type $a_k a_j/((\omega - i\lambda_k)^3)(\omega - i\lambda_j)^3)$ and similarly $p(p-1)$ terms, $k \neq j$, $b_k b_j/((\omega + i\lambda_j)(\omega + i\lambda_j))$. The residues in the upper half-plane of these terms sums to zero. The integral will be the sum of the $2p^2$ residue terms of the type

$$a_k b_j/((\omega - i\lambda_k)^3(\omega + i\lambda_j)^3).$$

The residues of these terms are of the form:

$$3 \cdot 4 a_k b_j/(-(i\lambda_k + i\lambda_j)^5,$$

and the integral therefore,

$$\int_{-\infty}^{\infty} (f''(\omega))^2 d\omega = 2\pi i \cdot 2 \sum_{k=1}^{p} \sum_{j=1}^{p} 3 \cdot 4 \frac{a_k b_j}{-(i\lambda_k + i\lambda_j)^5}.$$

Similarly one can use residue calculus to create of measure of steep hills in the spectrum,

$$\int_{-\infty}^{\infty} (f'(\omega))^2 d\omega,$$

or weighing $(f'(\omega))^2$ or $(f''(\omega))^2$ with a rational function.

The partial fraction trick can also be applied to calculate the Kullback-Leibner (KL) metric, as a measure of the distance between two functions, f and f_0 (e.g. a prior).

$$D(f_1; f_0) = \int f_1(\omega) \log(f_1(\omega)) d\omega - \int f_1(\omega) \log(f_0(\omega)) d\omega. \tag{6}$$

Using a partial fraction formulation of the spectral density the terms in (6) that need to be integrated will be of the form:

$$-\frac{c_{1,k}}{(\omega^2 + \lambda_{1,k}^2)} \log(w^2 + \lambda_{1,j}^2), \text{ and } \frac{c_{1,j}}{(\omega^2 + \lambda_{1,k}^2)} \log(1 + \mu_{1,j}^2 w^2).$$

Gradshteyn & Ryzhik (2007, eq 1, page 560) give the result:

$$\int_0^{\infty} \log(1 + \mu^2 x^2) \frac{dx}{x^2 + \lambda^2} = \frac{\pi}{\sqrt{\lambda^2}} \log(\sqrt{\lambda^2 \mu^2} + 1). \tag{7}$$

Here the square-root is take such that the real part of the square-root is positive. By use of partial fractions the KL distance can be written as:

$$\int \sum_{k=1}^{p_1} \frac{c_{1,k}}{\omega^2 + \lambda_{1,k}^2} (\sum_{j=1}^{q_1} \log(1 + \mu_{1,j}^2 \omega^2)) d\omega - \int \sum_{k=1}^{p_1} \frac{c_{1,k}}{\omega^2 + \lambda_{1,k}^2} (\sum_{j=1}^{p_1} \log(\omega^2 + \lambda_{1,j}^2)) d\omega -$$

$$\int \sum_{k=1}^{p_1} \frac{c_{1,k}}{\omega^2 + \lambda_{1,k}^2} (\sum_{j=1}^{q_0} \log(1 + \mu_{0,j}^2 \omega^2)) d\omega + \int \sum_{k=1}^{p_1} \frac{c_{1,k}}{\omega^2 + \lambda_{1,k}^2} (\sum_{j=1}^{p_0} \log(\omega^2 + \lambda_{0,j}^2)) d\omega.$$

This integral consists of $p_1 \times q_1 + p_1^2 + p_1 \times q_0 + p_1 \times p_0$ terms and each of them can be calculated by the use of (7),

$$\pi(\sum_{k=1}^{p_1}\sum_{j=1}^{q_1}\frac{c_{1,k}\log(1+\lambda_{1,k}\mu_{1,j})}{\lambda_{1,k}} - \sum_{k=1}^{p_1}\sum_{j=1}^{p_1}\frac{c_{1,k}\log(1+\lambda_{1,j}\lambda_{1,k})}{\lambda_{1,j}} +$$
$$\sum_{k=1}^{p_1}\sum_{j=1}^{q_0}\frac{c_{1,k}\log(1+\lambda_{1,k}\mu_{0,j})}{\lambda_{0,j}} - \sum_{k=1}^{p_1}\sum_{j=1}^{p_0}\frac{c_{1,k}\log(1+\lambda_{1,k}\lambda_{0,j})}{\lambda_{0,j}}).$$

Given the partial fractions of (3) these terms are all known. An analytical result of Kullback-Leibner distance, (6), can be obtained by applying (7) to the $p \times (q + p)$ terms in the formula above. All that is needed are the roots of the polynomials.

5 Implementation in R

A practical partial-fraction algorithm has been implemented in a R-package `ctarmaRcpp`, which is a `Rcpp` version of the `ctarma` packages used for the computations described in Tómasson (2015). For calculation of the partial fractions in (2) and (3) a algorithm based on Chen & Leung (1981) was implemented in the R function `partfrac1`. For each estimated model the roots of the AR and MA part are found and then various measures can be calculated. E.g.

$$1/(6+11x+6x^2+x^3) = \frac{1}{2(x+3)} - \frac{1}{x+2} + \frac{1}{2(x+1)},$$

here the roots are -1,-2,-3, and the function `partfrac1` gives the coefficients in the partial fraction (all roots distinct).

```
 partfrac1(c(6,11,6,1),1,c(-1,-2,-3),1)
 [1]  0.5 -1.0  0.5
```

The partial fraction in (3) enables the calculation of the Kullback-Leibler distance between two spectral shapes. A data set on the Earth's temperature for the past 800.000 years is used as an illustration on an unevenly sampled time series. The `ctarmaRcpp` package bundles data and model into a R object. The maximized log-likelihood of a continuous-time ARMA(2,1) is contained in `m2e`. The log-likelihood of `m2e` is calculate by:

```
> ctarma.loglik(m2e)
[1] -5701.584
```

An ARMA(4,3) gives log-likelihood of -5664.627, and an ARMA(6,5) a log-likelihood of -5660.819. The coefficients of the estimated ARMA(2,1), are

```
[1] 1792.32808   13.39429
> m2e$bhat
[1] 1.00000000 0.02315723
> m2e$sigma
[1] 1331.322
```

Similarly the estimated coefficients of the ARMA(4,3) are:

```
> m4e$ahat
[1] 1497.15420 3410.91710 2328.64602   28.11924
> m4e$bhat
[1] 1.0000000 1.2087125 0.3772288 0.0128648
> m4e$sigma
[1] 2239.939
```

The Kullback-Leibler distance is calculated with the function `kullbackDist` (here the implementation is between spectral shapes).

```
> kullbackDist(m4e$ahat,m4e$bhat,m4e$sigma,m2e$ahat,m2e$bhat)
[1] 1.172553
```

and for the ARMA(6,5)

```
> kullbackDist(m6e$ahat,m6e$bhat,m6e$sigma,m2e$ahat,m2e$bhat)
[1] 3.706201
```

The generalization to more complicated models is straightforeward.

6 Discussion

In this paper has shown an application of partial fractions and residue calculus to calculate measures of complexity of the spectral functions. The motivation of this measures is that the number of parameters, such as AIC and BIC, may not have the desired properties. The approach offers a way to measure the distance between functions, and a measure of features such as smoothness of a spectral function. The approach is based on the fact that the spectral function of an ARMA model is a rational function. The approach in this paper is based on continous-time ARMA but the arguments carry directly over to the discrete-time ARMA. The derivation of measures of the spectral features boil down to calculation of roots of polynomials. The ideas described allow the expression of many forms of penalty terms, e.g., a priori formulations of the spectral function.

Bibliography

Bergstrom, A. R. (1988). The history of continuous-time econometric models. *Econometric theory*, *4*, 365–383.

Chen, C. & Leung, K. (1981). A new look at partial fraction expansion from a high-level language viewpoint. *Computers & Mathematics with Applications*, *7*(5), 361 – 367.

Gradshteyn, I. S. & Ryzhik, I. M. (2007). *Table of integrals, series, and products* (Seventh ed.). Elsevier/Academic Press, Amsterdam. Translated from the Russian, Translation edited and with a preface by Alan Jeffrey and Daniel Zwillinger, With one CD-ROM (Windows, Macintosh and UNIX).

Kitagawa, G. & Gersch, W. (1996). *Smoothness Priors Analysis of Time Series.* Springer-Verlag New York.

Kreyszig, E. (1999). *Advanced Engineering Mathematics* (8 ed.). John Wiley & Sons. Residue theorem, page 723-724.

Tómasson, H. (2015). Some computational aspects of gaussian CARMA modelling. *Statistics and Computing*, *25*(2), 375–387.

Tsay, R. S. (2010). *Analysis of Financial Time Series* (Third ed.). John Wiley & Sons.

Zadrozny, P. (1988). Gaussian likelihood of continuous-time ARMAX models when data are stocks and flows at different frequencies. *Econometric Theory*, *4*(1), 108–124.