Research Article

Optimal Portfolio Estimation for Dependent Financial Returns with Generalized Empirical Likelihood

Hiroaki Ogata

School of International Liberal Studies, Waseda University, Tokyo 169-8050, Japan

Correspondence should be addressed to Hiroaki Ogata, hiroakiogata@aoni.waseda.jp

This paper proposes to use the method of generalized empirical likelihood to find the optimal portfolio weights. The log-returns of assets are modeled by multivariate stationary processes rather than i.i.d. sequences. The variance of the portfolio is written by the spectral density matrix, and we seek the portfolio weights which minimize it.

1. Introduction

The modern portfolio theory has been developed since circa the 1950s. It is common knowledge that Markowitz [1, 2] is a pioneer in this field. He introduced the so-called mean-variance theory, in which we try to maximize the expected return (minimize the variance) under the constant variance (the constant expected return). After that, many researchers followed, and portfolio theory has been greatly improved. For a comprehensive survey of this field, refer to Elton et al. [3], for example.

Despite its sophisticated paradigm, we admit there exists several criticisms against the early portfolio theory. One of them is that it blindly assumes that the asset returns are normally distributed. As Mandelbrot [4] pointed out, the price changes in the financial market do not seem to be normally distributed. Therefore, it is appropriate to use the nonparametric estimation method to find the optimal portfolio. Furthermore, it is empirically observed that financial returns are dependent. Therefore, it is unreasonable to fit the independent model to it.

One of the nonparametric techniques which has been capturing the spotlight recently is the empirical likelihood method. It was originally proposed by Owen [5, 6] as a method of

inference based on a data-driven likelihood ratio function. Smith [7] and Newey and Smith [8] extended it to the generalized empirical likelihood (GEL). GEL can be also considered as an alternative of generalized methods of moments (GMM), and it is known that its asymptotic bias does not grow with the number of moment restrictions, while the bias of GMM often does.

From the above point of view, we consider to find the optimal portfolio weights by using the GEL method under the multivariate stationary processes. The optimal portfolio weights are defined as the weights which minimize the variance of the return process with constant mean. The analysis is done in the frequency domain.

This paper is organized as follows. Section 2 explains about a frequency domain estimating function. In Section 3, we review the GEL method and mention the related asymptotic theory. Monte Carlo simulations and a real-data example are given in Section 4. Throughout this paper, A' and A^* indicate the transposition and adjoint of a matrix A, respectively.

2. Frequency Domain Estimating Function

Here, we are concerned with the m-dimensional stationary process $\{\mathbf{X}(t)\}_{t\in\mathbb{Z}}$ with mean vector $\mathbf{0}$, the autocovariance matrix

$$\Gamma(h) = \text{Cov}[\mathbf{X}(t+h), \mathbf{X}(t)] = E[\mathbf{X}(t+h)\mathbf{X}(t)'], \tag{2.1}$$

and spectral density matrix

$$f(\lambda) = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} \Gamma(h)e^{-ih\lambda}, \quad -\pi \leq \lambda < \pi. \tag{2.2}$$

Suppose that information of an interested p-dimensional parameter $\boldsymbol{\theta} \in \Theta \subset \mathbb{R}^p$ exists through a system of general estimating equations in frequency domain as follows. Let $\phi_j(\lambda; \boldsymbol{\theta})$, $(j = 1, \ldots, q, q \geq p)$ be $m \times m$ matrix-valued continuous functions on $[-\pi, \pi)$ satisfying $\phi_j(\lambda; \boldsymbol{\theta}) = \phi_j(\lambda; \boldsymbol{\theta})^*$ and $\phi_j(-\lambda; \boldsymbol{\theta}) = \phi_j(\lambda; \boldsymbol{\theta})'$. We assume that each $\phi_j(\lambda; \boldsymbol{\theta})$ satisfies the spectral moment condition

$$\int_{-\pi}^{\pi} \text{tr}\{\phi_j(\lambda; \boldsymbol{\theta}_0)f(\lambda)\}d\lambda = 0 \quad (j = 1, \ldots, q), \tag{2.3}$$

where $\boldsymbol{\theta}_0 = (\theta_{10}, \ldots, \theta_{p0})'$ is the true value of the parameter. By taking an appropriate function for $\phi_j(\lambda; \boldsymbol{\theta})$, (2.3) can express the best portfolio weights as shown in Example 2.1.

Example 2.1 (portfolio selection). In this example, we set $m = p = q$. Let $x_i(t)$ be the log-return of ith asset $(i = 1, \ldots, m)$ at time t and suppose that the process $\{\mathbf{X}(t) = (x_1(t), \ldots, x_m(t))'\}$ is stationary with zero mean. Consider the portfolio $p(t) = \sum_{i=1}^{m} \theta_i x_i(t)$, where $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_m)'$ is a vector of weights, satisfying $\sum_{i=1}^{m} \theta_i = 1$. The process $\{p(t)\}$ is a linear combination of the

stationary process, hence $\{p(t)\}$ is still stationary, and, from Herglotz's theorem, its variance is

$$\operatorname{Var}\{p(t)\} = \boldsymbol{\theta}' \operatorname{Var}\{\mathbf{X}(t)\}\boldsymbol{\theta} = \boldsymbol{\theta}'\left\{\int_{-\pi}^{\pi} f(\lambda)d\lambda\right\}\boldsymbol{\theta}. \tag{2.4}$$

Our aim is to find the weights $\boldsymbol{\theta}_0 = (\theta_{10},\ldots,\theta_{m0})'$ that minimize the variance (the risk) of the portfolio $p(t)$ under the constrain of $\sum_{i=1}^{m}\theta_i = 1$. The Lagrange function is given by

$$L(\boldsymbol{\theta},\lambda) = \boldsymbol{\theta}'\left\{\int_{-\pi}^{\pi} f(\lambda)d\lambda\right\}\boldsymbol{\theta} + \xi(\boldsymbol{\theta}'\mathbf{e} - 1), \tag{2.5}$$

where $\mathbf{e} = (1,1,\ldots,1)'$ and ξ is Lagrange multiplier. The first order condition leads to

$$(I - \mathbf{e}\boldsymbol{\theta}_0')\left[\int_{-\pi}^{\pi}\{f(\lambda) + f(\lambda)'\}d\lambda\right]\boldsymbol{\theta}_0 = \mathbf{0}, \tag{2.6}$$

where I is an identity matrix. Now, for fixed $j = 1,\ldots,m$, consider to take

$$\phi_j(\lambda;\boldsymbol{\theta}) = \begin{cases} 2\theta_j(1 - \theta_j) & (j,j)\text{th component} \\ 1 - 2\theta_j\theta_\ell & (j,\ell)\text{th and } (\ell,j)\text{th component with } \ell = 1,\ldots,m \text{ and } \ell \neq j \\ -2\theta_k\theta_\ell & (k,\ell)\text{th component with } k,\ \ell = 1,\ldots,m \text{ and } k \neq j,\ \ell \neq j. \end{cases} \tag{2.7}$$

Then, (2.3) coincides with the first order condition (2.6), which implies that the best portfolio weights can be solved with the framework of the spectral moment condition.

Besides, we can express other important indices for time series. In what follows, several examples are given.

Example 2.2 (autocorrelation). Denote the autocovariance and the autocorrelation of the process $\{x_i(t)\}$ (ith component of the process $\{\mathbf{X}(t)\}$) with lag h by $\gamma_i(h)$ and $\rho_i(h)$, respectively. Suppose that we are interested in the joint estimation of $\rho_i(h)$ and $\rho_j(k)$. Take

$$\phi_1(\lambda;\boldsymbol{\theta}) = \begin{cases} \cos(h\lambda) - \theta_1 & (i,i)\text{th component} \\ 0 & \text{otherwise,} \end{cases}$$

$$\phi_2(\lambda;\boldsymbol{\theta}) = \begin{cases} \cos(k\lambda) - \theta_2 & (j,j)\text{th component} \\ 0 & \text{otherwise.} \end{cases} \tag{2.8}$$

Then, (2.3) leads to

$$\theta_{10} = \left\{ \int_{-\pi}^{\pi} \exp(ih\lambda) f_{ii}(\lambda) d\lambda \right\} \left\{ \int_{-\pi}^{\pi} f_{ii}(\lambda) d\lambda \right\}^{-1},$$

$$\theta_{20} = \left\{ \int_{-\pi}^{\pi} \exp(ik\lambda) f_{jj}(\lambda) d\lambda \right\} \left\{ \int_{-\pi}^{\pi} f_{jj}(\lambda) d\lambda \right\}^{-1}.$$

(2.9)

From Herglotz's theorem, $\theta_{10} = \gamma_i(h)/\gamma_i(0)$, and $\theta_{20} = \gamma_j(k)/\gamma_j(0)$. Then, $\boldsymbol{\theta_0} = (\theta_{10}, \theta_{20})'$ corresponds to the desired autocorrelations $\boldsymbol{\rho} = (\rho_i(h), \rho_j(k))'$. The idea can be directly extended to more than two autocorrelations.

Example 2.3 (Whittle estimation). In this example, we set $p = q$. Consider fitting a parametric spectral density model $f_{\boldsymbol{\theta}}(\lambda)$ to the true spectral density $f(\lambda)$. The disparity between $f_{\boldsymbol{\theta}}(\lambda)$ and $f(\lambda)$ is measured by the following criterion:

$$D(f_{\boldsymbol{\theta}}, f) \equiv \int_{-\pi}^{\pi} \left[\log \det f_{\boldsymbol{\theta}}(\lambda) + \mathrm{tr}\left\{ f_{\boldsymbol{\theta}}(\lambda)^{-1} f(\lambda) \right\} \right] d\lambda,$$

(2.10)

which is based on Whittle likelihood. The purpose here is to seek the quasi-true value $\underline{\boldsymbol{\theta}}$ defined by

$$\underline{\boldsymbol{\theta}} = \arg \min_{\theta \in \Theta} D(f_{\boldsymbol{\theta}}, f).$$

(2.11)

Assume that the spectral density model is expressed by the following form:

$$f_{\boldsymbol{\theta}}(\lambda) = \left\{ \sum_{j=0}^{\infty} B_{\boldsymbol{\theta}}(j) \exp(ij\lambda) \right\} K \left\{ \sum_{j=0}^{\infty} B_{\boldsymbol{\theta}}(j) \exp(ij\lambda) \right\}^{*},$$

(2.12)

where each $B_{\boldsymbol{\theta}}(j)$ is an $m \times m$ matrix ($B_{\boldsymbol{\theta}}(0)$ is defined as identity matrix), and K is an $m \times m$ symmetric matrix. The general linear process has this spectral form, so this assumption is not so restrictive. The key of this assumption is that the elements of the parameter $\boldsymbol{\theta}$ do not depend on K. We call such a parameter innovation-free. Imagine that you fit the VARMA process, for example. Innovation-free implies that the elements of the parameter $\boldsymbol{\theta}$ depend on only AR or MA coefficients and not on the covariance matrix of the innovation process. Now, let us consider the equation:

$$\left. \frac{\partial}{\partial \boldsymbol{\theta}} D(f_{\boldsymbol{\theta}}, f) \right|_{\theta = \underline{\theta}} = \mathbf{0},$$

(2.13)

to find quasi-true value. The Kolmogorov's formula says

$$\det K = \exp\left[\frac{1}{2\pi} \int_{-\pi}^{\pi} \log \det\{ 2\pi f_{\boldsymbol{\theta}}(\lambda) \} d\lambda \right].$$

(2.14)

This implies that if $\boldsymbol{\theta}$ is innovation-free, the quantity $\int_{-\pi}^{\pi} \log \det\{f_{\boldsymbol{\theta}}(\lambda)\}d\lambda$ is independent of $\boldsymbol{\theta}$ and (2.13) leads to

$$\frac{\partial}{\partial \boldsymbol{\theta}} \int_{-\pi}^{\pi} \mathrm{tr}\left\{f_{\boldsymbol{\theta}}(\lambda)^{-1}f(\lambda)d\lambda\right\}\bigg|_{\boldsymbol{\theta}=\underline{\boldsymbol{\theta}}} = \mathbf{0}. \tag{2.15}$$

This corresponds to (2.3), when we set

$$\phi_j(\lambda; \boldsymbol{\theta}) = \frac{\partial f_{\boldsymbol{\theta}}(\lambda)^{-1}}{\partial \theta_j} \quad (j = 1, \ldots, p), \tag{2.16}$$

so the quasi-true value can be expressed by the spectral moment condition.

Based on the form of (2.3), we set the estimating function for $\boldsymbol{\theta}$ as

$$\mathbf{m}(\lambda_t; \boldsymbol{\theta}) = \left(\mathrm{tr}\{\phi_1(\lambda_t; \boldsymbol{\theta})I_n(\lambda_t)\}, \ldots, \mathrm{tr}\{\phi_q(\lambda_t; \boldsymbol{\theta})I_n(\lambda_t)\}\right)', \tag{2.17}$$

where $I_n(\lambda)$ is the periodogram, defined by

$$I_n(\lambda) = (2\pi n)^{-1}\left\{\sum_{t=1}^{n}\mathbf{X}(t)\exp(it\lambda)\right\}\left\{\sum_{t=1}^{n}\mathbf{X}(t)\exp(it\lambda)\right\}^{*}, \tag{2.18}$$

where $\lambda_t = (2\pi t)/n$, $(t = -[(n-1)/2], \ldots, [n/2])$. Then, we have

$$\frac{2\pi}{n}\sum_{-[(n-1)/2]}^{[n/2]}E[\mathbf{m}(\lambda_t; \boldsymbol{\theta})] \longrightarrow \left[\int_{-\pi}^{\pi}\mathrm{tr}\{\phi_j(\lambda; \boldsymbol{\theta}_0)f(\lambda)\}d\lambda\right]_{j=1,\ldots,q} = \mathbf{0}. \tag{2.19}$$

3. Generalized Empirical Likelihood

Once we construct the estimating function, we can make use of the method of GEL as in the work by Smith [7] and Newey and Smith [8]. GEL is introduced as an alternative to GMM and it is pointed out that its asymptotic bias does not grow with the number of moment restrictions, while the bias of GMM often does.

To describe GEL let $\rho(v)$ be a function of a scalar v which is concave on its domain, an open interval $\mathcal{V}$ containing zero. Let $\widehat{\Lambda}_n(\boldsymbol{\theta}) = \{\boldsymbol{\lambda} : \boldsymbol{\lambda}'\mathbf{m}(\lambda_t; \boldsymbol{\theta}) \in \mathcal{V}, \ t = 1, \ldots, n\}$. The estimator is the solution to a saddle point problem

$$\widehat{\boldsymbol{\theta}}_{\mathrm{GEL}} = \arg\min_{\boldsymbol{\theta}\in\Theta}\sup_{\boldsymbol{\lambda}\in\widehat{\Lambda}_n(\boldsymbol{\theta})}\sum_{t=1}^{n}\rho(\boldsymbol{\lambda}'\mathbf{m}(\lambda_t; \boldsymbol{\theta})). \tag{3.1}$$

The empirical likelihood (EL) estimator (cf. [9]) the exponential tilting (ET) estimator (cf. [10]), and the continuous updating estimator (CUE) (cf. [11]) are special cases with $\rho(v) = \log(1-v)$, $\rho(v) = -e^v$ and $\rho(v) = -(1+v)^2/2$, respectively. Let $\Omega = E[\mathbf{m}(\lambda_t; \boldsymbol{\theta}_0)\mathbf{m}(\lambda_t; \boldsymbol{\theta}_0)']$,

Table 1: Estimated autocorrelations for two-dimensional-AR(1) model.

	EL		ET		CUE	
	Mean	s.d.	Mean	s.d.	Mean	s.d.
$\rho_1 (1)$	0.3779	0.0890	0.3760	0.0900	0.3797	0.0911
$\rho_2 (1)$	0.4680	0.0832	0.4660	0.0817	0.4650	0.0855

$G = E[\partial \mathbf{m}(\lambda_t; \boldsymbol{\theta}_0)/\partial \boldsymbol{\theta}]$, and $\Sigma = (G'\Omega^{-1}G)^{-1}$. The following assumptions and theorems are due to Newey and Smith [8].

Assumption 3.1. (i) $\boldsymbol{\theta}_0 \in \Theta$ is the unique solution to (2.3).

(ii) Θ is compact.

(iii) $\mathbf{m}(\lambda_t; \boldsymbol{\theta})$ is continuous at each $\boldsymbol{\theta} \in \Theta$ with probability one.

(iv) $E[\sup_{\boldsymbol{\theta} \in \Theta} \|\mathbf{m}(\lambda_t; \boldsymbol{\theta})\|^{\alpha}] < \infty$ for some $\alpha > 2$.

(v) Ω is nonsingular.

(vi) $\rho(v)$ is twice continuously differentiable in a neighborhood of zero.

Theorem 3.2. *Let Assumption 3.1 hold. Then $\widehat{\boldsymbol{\theta}}_{\mathrm{GEL}} \xrightarrow{p} \boldsymbol{\theta}_0$.*

Assumption 3.3. (i) $\boldsymbol{\theta}_0 \in \mathrm{int}(\Theta)$.

(ii) $\mathbf{m}(\lambda_t; \boldsymbol{\theta})$ is continuously differentiable in a neighborhood $\mathcal{N}$ of $\boldsymbol{\theta}_0$ and $E[\sup_{\boldsymbol{\theta} \in \mathcal{N}} \|\partial \mathbf{m}(\lambda_t; \boldsymbol{\theta})/\partial \boldsymbol{\theta}'\|] < \infty$.

(iii) $\mathrm{rank}(G) = p$.

Theorem 3.4. *Let Assumptions 3.1 and 3.3 hold. Then $\sqrt{n}(\widehat{\boldsymbol{\theta}}_{\mathrm{GEL}} - \boldsymbol{\theta}_0) \xrightarrow{p} N(\mathbf{0}, \Sigma)$.*

4. Monte Carlo Studies and Illustration

In the first part of this section, we summarize the estimation results of Monte Carlo studies of the GEL method. We generate 200 observations from the following two-dimensional-AR(1) model

$$\begin{bmatrix} X_1(t) \\ X_2(2) \end{bmatrix} = \begin{bmatrix} 0.2 & -0.5 \\ -0.5 & 0.3 \end{bmatrix} \begin{bmatrix} X_1(t-1) \\ X_2(t-2) \end{bmatrix} + \begin{bmatrix} U_1(t) \\ U_2(t) \end{bmatrix}, \tag{4.1}$$

where $\{\mathbf{U}(t) = (U_1(t), U_2(t))'\}_{t \in \mathbb{Z}}$ is an i.i.d. innovation process, distributed to two-dimensional t-distribution whose correlation matrix is identity, and degree of freedom is 5. The true autocorrelations with lag 1 of this process are $\rho_1(1) = 0.3894$ and $\rho_2(1) = 0.4761$, respectively. As described in Example 2.2, we estimate $\rho_1(1)$ and $\rho_2(1)$ by using three types of frequency domain GEL method (EL, ET, and CUE). Table 1 shows the mean and standard deviation of the estimation results whose repetition time is 1000. All types work properly.

Next, we apply the proposed method to the returns of market index data. The sample consists of 7 weekly indices (S&P 500, Bovespa, CAC 40, AEX, ATX, HKHSI, and Nikkei) having 800 observations each: the initial date is April 30, 1993, and the ending date is August 22, 2008. Refer to Table 2 for the market of each index.

Table 2: Market.

Index	Market
S&P 500	NYSE
Bovespa	São Paulo Stock Exchange
CAC 40	Bourse de Paris
AEX	Amsterdam Stock Exchange
ATX	Wiener Börse
HKHSI	Hong Kong Exchanges and Clearing
Nikkei	Tokyo Stock Exchange

Table 3: Estimated portfolio weights.

	EL	ET	CUE
S&P 500	0.0759	0.0617	0.0001
Bovespa	0.6648	0.6487	0.6827
CAC 40	0.0000	0.0000	0.0000
AEX	0.0000	0.0000	0.0000
ATX	0.2593	0.2558	0.3168
HKHSI	0.0000	0.0000	0.0000
Nikkei	0.0000	0.0338	0.0004

As shown in Example 2.1, we use frequency domain GEL method to estimate the optimal portfolio weights, and the results are shown in Table 3. Bovespa and ATX account for large part in the optimal portfolio.

Acknowledgment

References

[1] H. Markowitz, "Portfolio selection," *Journal of Finance*, vol. 7, no. 1, pp. 77–91, 1952.

[2] H. M. Markowitz, *Portfolio Selection: Efficient Diversification of Investments*, John Wiley & Sons, New York, NY, USA, 1959.

[3] E. J. Elton, M. J. Gruber, S. J. Brown, and W. N. Goetzmann, *Modern Portfolio Theory and Investment Analysis*, John Wiley & Sons, New York, NY, USA, 7th edition, 2007.

[4] B. Mandelbrot, "The variation of certain speculative prices," *Journal of Business*, vol. 36, pp. 394–419, 1963.

[5] A. B. Owen, "Empirical likelihood ratio confidence intervals for a single functional," *Biometrika*, vol. 75, no. 2, pp. 237–249, 1988.

[6] A. Owen, "Empirical likelihood ratio confidence regions," *The Annals of Statistics*, vol. 18, no. 1, pp. 90–120, 1990.

[7] R. J. Smith, GEL criteria for moment condition models. Working paper, 2004, http://cemmap.ifs.org.uk/wps/cwp0419.pdf.

[8] W. K. Newey and R. J. Smith, "Higher order properties of GMM and generalized empirical likelihood estimators," *Econometrica*, vol. 72, no. 1, pp. 219–255, 2004.

[9] J. Qin and J. Lawless, "Empirical likelihood and general estimating equations," *The Annals of Statistics*, vol. 22, no. 1, pp. 300–325, 1994.

[10] Y. Kitamura and M. Stutzer, "An information-theoretic alternative to generalized method of moments estimation," *Econometrica*, vol. 65, no. 4, pp. 861–874, 1997.

[11] L. P. Hansen, J. Heaton, and A. Yaron, "inite-sample properties of some alternative GMM estimators," *Journal of Business and Economic Statistics*, vol. 14, no. 3, pp. 262–280, 1996.

Advances in
Operations Research

Advances in
Decision Sciences

Journal of
Applied Mathematics

Algebra

Journal of
Probability and Statistics

The Scientific
World Journal

International Journal of
Differential Equations

International Journal of
Combinatorics

Submit your manuscripts at
http://www.hindawi.com

Hindawi

Advances in
Mathematical Physics

Journal of
Complex Analysis

Journal of
Mathematics

Mathematical Problems
in Engineering

Abstract and
Applied Analysis

Discrete Dynamics in
Nature and Society

International
Journal of
Mathematics and
Mathematical
Sciences

Journal of
Discrete Mathematics

Journal of
Function Spaces

International Journal of
Stochastic Analysis

Journal of
Optimization