

The General Method of Moments (GMM) using MATLAB: The practical guide based on the CKLS interest rate model

*Kamil Kladívko*¹

Department of Statistics and Probability Calculus,
University of Economics, Prague
kladivk@vse.cz

Abstract

The General Method of Moments (GMM) is an estimation technique which can be used for variety of financial models. We use the CKLS class of interest rate models to demonstrate how GMM works. We discuss the practical implementation in MATLAB. We pay attention to exactly-identified versus over-identified estimation, minimization of objective function and hypothesis testing of the model. We then illustrate problem areas in implementation, specifically the importance of the variance-covariance matrix estimator of moment functions. Our technical note is supplemented with the MATLAB code of discussed topics.

1 Short-Term Interest Rate Stochastic Differential Equation

The dynamics of a short-term interest rate can be nested within the following stochastic differential equation (SDE):

$$dr_t = (\alpha + \beta r)dt + \sigma r^\gamma dZ_t, \quad (1)$$

where $\alpha, \beta, \sigma, \gamma$ are model parameters, r_t is the short-term interest rate and Z_t is the standard Brownian motion. CKLS [1] defines the parameter vector $\theta = (\alpha, \beta, \sigma^2, \gamma)$.

The SDE (1) defines a broad class of interest rate processes which include many well-known one-factor interest rate models. These models can be obtained from (1) by placing appropriate restrictions on the parameters (see Table 1).

CKLS estimated nine different specifications of the interest rate process. In this paper, we implement three of them. The implemented models are summarized in Table 1. Adding further specifications is a routine task.

Name	SDE	Restrictions
CKLS	$dr_t = (\alpha + \beta r_t)dt + \sigma r_t^\gamma dZ_t$	(unrestricted)
CIR	$dr_t = (\alpha + \beta r_t)dt + \sigma \sqrt{r_t} dZ_t$	$\gamma = \frac{1}{2}$
Vasicek	$dr_t = (\alpha + \beta r_t)dt + \sigma dZ_t$	$\gamma = 0$

Table 1: Implemented short-term interest rate processes. CIR and Vasicek are probably the most commonly used models among the models analyzed in the CKLS paper.

2 GMM Estimation of the Short-Term Interest Rate SDE

To estimate the parameters of the model, CKLS relied on the General Method of Moments (GMM). We present basic ideas of this estimation technique in this section. Details can be found in [2].

¹Supported by IG410067 and EEA/Norwegian FMP – Student mobility

2.1 Moment Conditions

To estimate parameters of a SDE using the GMM we have to discretize the SDE. CKLS applied Euler discretization scheme on (1):

$$r_{t+1} - r_t = \alpha + \beta r_t + \varepsilon_{t+1}, \quad (2)$$

$$\varepsilon_{t+1} \equiv \sigma r_t^\gamma N(0, 1), \quad (3)$$

where $N(0, 1)$ is a random shock with zero mean and unit variance. It does not have to be necessarily drawn from the standard Gaussian distribution, although Euler discretization assumes that.

From (3) we can immediately observe that

$$E[\varepsilon_{t+1}] = 0, \quad (4)$$

$$D[\varepsilon_{t+1}] = E[\varepsilon_{t+1}^2] = \sigma^2 r_t^{2\gamma}. \quad (5)$$

Employing the characteristics (4) and (5) and using the notion that the error term ε_{t+1} is uncorrelated with the explanatory variable r_t (orthogonality condition), CKLS derived a set of four moment functions:

$$f_t(\theta) = \begin{bmatrix} \varepsilon_{t+1} \\ \varepsilon_{t+1}^2 - \sigma^2 r_t^{2\gamma} \end{bmatrix} \otimes \begin{bmatrix} 1 \\ r_t \end{bmatrix} = \begin{bmatrix} \varepsilon_{t+1} \\ \varepsilon_{t+1} r_t \\ \varepsilon_{t+1}^2 - \sigma^2 r_t^{2\gamma} \\ (\varepsilon_{t+1}^2 - \sigma^2 r_t^{2\gamma}) r_t \end{bmatrix}. \quad (6)$$

The moment functions are constructed so that

$$E[f_t(\theta)] = 0. \quad (7)$$

Equations (7) are called moment conditions of moments $E[f_t(\theta)]$ in the GMM terminology. The basic idea of GMM is to pick model parameters that “fit” these moment conditions in our data set.

We could easily come up with additional moment conditions. For example, we restrict that ε_t is not autocorrelated, i.e. $E[\varepsilon_{t+1}\varepsilon_t] = 0$.

2.2 Sample Moments

The corresponding sample moments are defined as

$$g_T(\theta) = \frac{1}{T} \sum_{t=1}^T f_t(\theta), \quad (8)$$

where T is a number of available observations (sample size). Equations (8) are natural sample counterparts of population moments $E[f_t(\theta)]$. (The traditional notation used in statistics would denote these estimates $\hat{g}(\theta)$. However using T subscript is a widespread notation due to [2].)

We find it convenient to reparametrise (2) to keep the implementation code neat. Further, to get the numerical results in accordance with the CKLS paper, we have to employ time step Δt . Equations (2) and (3) turn into

$$r_{t+1} = a + b r_t + \varepsilon_{t+1}, \quad (9)$$

$$\varepsilon_{t+1} \equiv \sigma r_t^\gamma \sqrt{\Delta t} N(0, 1), \quad (10)$$

where $a = \alpha \Delta t$ and $b = 1 + \beta \Delta t$. Then the sample moments look as follows:

$$g_T(\theta) = \frac{1}{T} \begin{bmatrix} \sum_{t=1}^T (r_{t+1} - a - b r_t) \\ \sum_{t=1}^T ((r_{t+1} - a - b r_t)^2 - \sigma^2 r_t^{2\gamma} \Delta t) \\ \sum_{t=1}^T (r_{t+1} - a - b r_t) r_t \\ \sum_{t=1}^T ((r_{t+1} - a - b r_t)^2 - \sigma^2 r_t^{2\gamma} \Delta t) r_t \end{bmatrix}. \quad (11)$$

2.3 Objective Function

The GMM objective function is defined as

$$J_T = g_T'(\theta)W_T g_T(\theta), \quad (12)$$

where W_T is positive-definite weighting matrix. We find our parameter estimates by solving

$$\hat{\theta} = \arg \min_{\theta} J_T. \quad (13)$$

We can see that the GMM is a minimum distance estimator. In other words, we are setting the weighted sum of squared sample moments as close to zero as possible. As we know, their population counterparts are constructed to be equal to zero. The weighting matrix tells us how much attention to pay to each moment.

For the unrestricted (CKLS) model we have four moment conditions and four parameters. We say that the CKLS model is exactly identified and we can pick an arbitrary positive-definite W_T . We can set all sample moments equal exactly zero.

2.4 Optimal Weighting Matrix W_T

The CIR and Vasicek models have three parameters, these models are overidentified. It is not possible (with probability one) to set all sample moment conditions equal to zero. To obtain asymptotically efficient estimates we set

$$W_T = \hat{S}^{-1}, \quad (14)$$

where \hat{S} is an estimate of the spectral density matrix of population moment functions. This choice of the weighting matrix secures the smallest asymptotic covariance matrix of the estimate of θ . This is important result is due to Hansen [2]. We call this the efficient GMM.

The spectral density or long run covariance matrix is defined

$$S = \sum_{j=-\infty}^{\infty} E[f_t(\theta)f_{t-j}'(\theta)]. \quad (15)$$

The spectral density matrix allows for serial correlation and heteroskedasticity in the observations of the moments function. A popular consistent estimate of the spectral density matrix is the Newey-West estimator [3]:

$$\hat{S} = \hat{S}_0 + \sum_{j=1}^k \left(1 - \frac{j}{k+1}\right) (\hat{S}_j + \hat{S}_j'), \quad (16)$$

where

$$\hat{S}_j = \frac{1}{T} \sum_{t=j+1}^T f_t(\theta)f_{t-j}'(\theta).$$

In case of overidentified models, GMM is a two stage estimator. We usually proceed in the following way:

1. We minimize (12) using identity weighting matrix ($W_T = I$). This means that we consider all moments equally important. We plug estimated parameter vector $\hat{\theta}$ into (16) and inverse to get W_T .
2. Again, we minimize (12), but this time using the W_T from the first step.

2.5 Testing the Significance of Individual Parameters

Under the efficient GMM (and under the exactly identified GMM), we have asymptotic normality for the estimated parameters:

$$\sqrt{T}(\hat{\theta} - \theta) \xrightarrow{d} N(0, (d'S^{-1}d)^{-1}), \quad (17)$$

where d is the Jacobian of the population moments vector $E[f_t(\theta)]$.

For constructing the test statistics, we substitute in the corresponding sample moments and evaluate them at the estimated parameters. For S , we substitute in our estimate (16) of the spectral density matrix, i.e. our weighting matrix W_T .²

For example, the estimate of the Jacobian matrix of the CKLS model based on the sample moments (11) looks as follows:

$$\hat{d}(\hat{\theta}) = -\frac{1}{T} \begin{bmatrix} T & \sum r_t & 0 & 0 \\ 2 \sum (r_{t+1} - a - br_t) & 2 \sum (r_{t+1} - a - br_t)r_t & \Delta t \sum r_t^{2\gamma} & 2\sigma^2 \Delta t \sum \ln(r_t)r_t^{2\gamma} \\ \sum r_t & \sum r_t^2 & 0 & 0 \\ 2 \sum (r_{t+1} - a - br_t)r_t & 2 \sum (r_{t+1} - a - br_t)r_t^2 & \Delta t \sum r_t^{2\gamma+1} & 2\sigma^2 \Delta t \sum \ln(r_t)r_t^{2\gamma+1} \end{bmatrix}. \quad (18)$$

2.6 Overall Test of Model Specification (Model Adequacy)

If the model is overidentified, GMM offers an overall test of the model by testing whether the “extra” sample moments are sufficiently close to zero relative to their distribution.

The test statistic is asymptotically χ^2 distributed:

$$T \times J_T(\hat{\theta}) = T \times g'_T(\hat{\theta})W_T(\hat{\theta})g_T(\hat{\theta}) \xrightarrow{d} \chi_{m-p}^2, \quad (19)$$

where m is a number of moment conditions and p is a number of parameters.

If the test statistic rejects, then the underlying model that generated the system of moment conditions is declared invalid.

3 MATLAB Implementation

The estimation routine is executed by running the `RunEstimation.m` (uses `GMMestimation.m`, `GMMobjective.m`, `GMMweightsNW.m` and `MomentsJacobia.m`). See code comments for details. The results are displayed in the MATLAB’s command window. We rely on `fminsearch` routine when minimizing the objective function.

The Treasury bill yield data used by CKLS are one-month yields from the Center for Research in Security Prices (CRSP). The data are monthly and cover period from June 1964 to December 1989. The data set is saved in `ckls.dat`.

4 Importance of the Spectral Density Matrix Estimation

One of the main results of CKLS is concerned with the value of γ . They showed that the value of γ is the most important feature differentiating interest rate models. Models which assume $\gamma < 1$, e.g. CIR and Vasicek, are rejected by the overall χ^2 test of model adequacy. And those which assume $\gamma \geq 1$ are not rejected.

²In case of the the exactly identified model (CKLS), we still calculate the spectral density matrix and do the inference.

4.1 Spectral Density Matrix

By running the estimation routine with different number of lags (different k) used in the spectral density matrix estimation (16), we can deduce that CKLS used an ordinary sample covariance matrix in their inference ($k = 0$).

By relying on the ordinary sample covariance matrix, they implicitly assume that the moment functions can be considered as independent identically distributed variables.

4.2 Spectral Density Matrix for the CIR Model

Table 2 reports sample autocorrelations of the four moment functions for the CIR model. The autocorrelations of the second and the fourth moment functions are obviously significantly different from zero.

lag	autocorrelation of $g_T(\hat{\theta})$			
1	0.04	0.18	-0.01	0.13
2	-0.06	0.15	-0.09	0.11
3	-0.10	0.16	-0.09	0.09
4	-0.06	0.12	-0.08	0.07
5	-0.01	0.05	-0.03	0.02

Table 2: This table reports sample autocorrelations (first five lags) of the four moment functions for the CIR model.

Thus we should use e.g. the Newey-West estimator (16) of the spectral density matrix S . If we do so, and set the number of lags of the Newey-West estimator greater or equal to 12, we no longer can reject the CIR model on the 5% significance level. The results are reported in Table 3. Also note that the parameter estimates differ considerably, which questions the robustness of

#lags	α	β	σ^2	χ^2	p-value
0	0.0190	-0.2281	0.0073	5.82	0.0158
12	0.0328	-0.5145	0.0057	3.73	0.0533

Table 3: This table reports GMM estimation results for the CIR model when using different number of lags in the Newey-West estimator of the spectral density matrix. First row reports results with zero lags in the Newey-West estimator, i.e. results obtained with an ordinary sample covariance matrix.

the GMM for the CIR model estimation. Qualitatively similar results can be obtained for the Vasicek model.

4.3 GMM versus Maximum Likelihood for the CIR Model

We will compare the GMM to the Maximum Likelihood (ML). The transition density function has a closed form solution for the CIR model and thus the ML estimator is available.

For easily noticeable comparison, we find it convenient to rewrite the CIR model into the following form:

$$dr_t = \lambda(\mu - r_t)dt + \sqrt{r_t}\sigma dW_t, \quad (20)$$

where interest rate r_t moves in the direction of its mean μ at speed λ . We get back to the former notation by setting $\alpha = \lambda\mu$ and $\beta = -\lambda$.

The estimation results of the ML versus the GMM are reported in Table 4. It is striking that the 12-lags Newey-West GMM estimate of the drift (λ, μ) is much closer to the ML estimate compared to the GMM estimate without using the Newey-West. This result emphasizes the importance of the optimal weighting matrix estimation.

Method	#lags	λ	μ	σ
GMM	0	0.228	0.083	0.086
GMM	12	0.515	0.064	0.075
ML		0.599	0.069	0.098

Table 4: This table reports the GMM versus the ML estimation results for the CIR model. The sample mean of interest rate time series is equal to 0.067.

Conclusion

In this technical note we have outlined, how to implement the General Method of Moments estimator for one-factor interest rate process. We have demonstrated that one should pay close attention to the weighting matrix estimation.

References

- [1] Chan, K.C., Karolyi, Andrew G., Longstaff, Francis A. and Sanders, Anthony B., 1992, “An Empirical Comparison of Alternative Models of the Short-Term Interest Rate,” *The Journal of Finance* 47, 1209–1227.
- [2] Hansen, Lars Peter, 1982, “Large Sample Properties of Generalized Method of Moments Estimators,” *Econometrica* 50, 1029–1054.
- [3] Newey, Whitney, K. and West, Kenneth, D., 1987, “A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix,” *Econometrica* 55, 703–708.