Estimating equations for a class of time-irreversible multi-factor models

Nina Boyarchenko and Sergei Levendorskiĭ

The University of Chicago; the University of Texas at Austin

October 26, 2007

Abstract

The standard operator approach to the identification problem of diffusions and more general Markov processes relies on variational principles for self-adjoint operators. If the process is not time reversible, equivalently, the infinitesimal operator of the process is not self-adjoint, these principles are not applicable. We develop the spectral decomposition for multi-factor time-irreversible Ornstein-Uhlenbeck processes, Affine Term Structure Models (ATSMs) and Quadratic Term Structure Models (QTSMs), and use it to construct new estimating equations.

Operator method

Short bibliography

Hansen and Scheinkman (1995),

Hansen et al (1998),

Kessler and Sørensen (1999),

Chen et al. (2005),...

Rationale

- Need to estimate parameters of a continuous time process but observations are available at discrete time moments only.
- Equivalently, need to find $\mathcal{L}$, the infinitesimal generator of a process in continuous time, but can observe transition operators $\exp(\tau\mathcal{L})$ for non-vanishing $\tau > 0$

Resulting estimation equations

$$e^{\tau\lambda_j} = E[u_j(X_{t+\tau})u_j(X_t) \mid X_t], \tag{1}$$

where λ_j are eigenvalues of $\mathcal{L}$.

How to apply?

Normalize $\tau = 1$ and introduce notation

- q– the stationary distribution of the process
- e^{λ_0}– the maximal eigenvalue of $e^{\mathcal{L}}$
- ϕ_0–the corresponding eigenfunction

Time-reversible processes

$\mathcal{L}$ is a self-adjoint operator in $L_2(q(x)dx)$, therefore e^{λ_0} and ϕ_0 can be found as a unique solution of the maximization problem

$$e^{\tau\lambda_0} := \max \frac{\left(e^{\tau\mathcal{L}}u, u\right)_q}{(u,u)_q} = \max \frac{E[u(X_t+\tau)u(X_t) \mid X_t]}{(u,u)_q}, \tag{2}$$

and the next eigenvalues (possibly, multiple ones) and eigenfunctions can be found using the same maximization problem (2) but with the maximization over a subspace orthogonal to the subspace generated by the already found eigenfunctions.

An infinite series of estimating equations results. In many cases, it becomes possible to find parameters of the process

Time-irreversible processes

Reduction to a self-adjoint operator

The infinitesimal generator is not self-adjoint, therefore the procedure described above is not directly applicable. In some cases, one can find a similarity transformation, which reduces the infinitesimal generator of the process to a self-adjoint operator. Then variational principles can be applied.

Multi-factor case

Typically, such a transformation does not exist or is difficult to find

In Economics and Finance

there are many popular families of time-irreversible multi-factor models, which, for realistic parameter values, are not reducible to models with self-adjoint infinitesimal generators

Examples of (typically) time-irreversible models

- multi-factor Ornstein-Uhlenbeck processes
- Affine Term Structure Models (ATSMs)
- Quadratic Term Structure Models (QTSMs)

Applications

pricing interest rate derivatives and stocks under stochastic interest rates.

Recently, ATSMs and QTSMs have also been applied to model default intensities in credit risk models.

Our contribution I.

A general method for the spectral decomposition of QTSMs, ATSMs of type $A_0(n)$ and OU processes, processes in random time including.

- construction of finite-dimensional subspaces $\mathcal{V}_j$ invariant under $\mathcal{L}$
- calculation of the matrices of $\mathcal{L}^j$, the restrictions $\mathcal{L}^j$ of $\mathcal{L}$ to $\mathcal{V}_j$, $j = 0, 1, \dots$
- decomposition $\mathcal{L} = \oplus_{j=0}^{\infty} \mathcal{L}^j$
- in the general case, matrices of operators $\mathcal{L}^j$ can be calculated solving systems of ODE of order 1, which is a stable and accurate procedure if the matrix size is not large
- in the two-factor case, explicit spectral decomposition of $\mathcal{L}^j$

Our contribution II.

derivation of new estimating equations based on invariants such as traces and determinants of restrictions of $\exp(\tau\mathcal{L})$ on finite-dimensional subspaces

second group of estimating equation uses the variational principles for the self-adjoint operators associated with the restrictions of the quadratic form $(e^{\tau\mathcal{L}}u, u)_q$ to finite-dimensional invariant subspaces

Example: 2-factor OU model

The knowledge of the stationary distribution, q, and the trace of the restriction of $e^{\tau \mathcal{L}}$ to the invariant subspace of dimension 2 suffices to infer the parameters of the model up to the sign of the off-diagonal elements of the mean-reverting matrix in a basis of eigenvectors e^1, e^2 of the quadratic form $-\log q$

If the mean-reverting matrix is symmetric, the stationary distribution and the trace of this restriction can be viewed as a natural parametrization of the two-factor OU process with the unit variance-covariance matrix. This identification is similar in the spirit to (albeit different from) the identification of scalar diffusions in terms of the stationary density and a conveniently chosen eigenvalue-eigenfunction pair of the conditional expectation operator over a finite unit of time in Hansen et al. (1998).

If the process is not time reversible

For the complete identification, the sign of the off-diagonal elements of the mean-reverting matrix in a basis of eigenvectors e^1, e^2 of the quadratic form $-\log q$ is needed. Equivalently, we can use the sign of the covariance function

$$E[\phi_{01}(X_{t+\tau})\phi_{10}(X_t)] = (e^{\tau\mathcal{L}}\phi_{01}, \phi_{10})_q,$$

where ϕ_{01}, ϕ_{10} are the coordinate functions w.r.t. the basis e^1, e^2. Similarly to Hansen et al. (1998), this identification method is robust to the presence of possibly temporarily dependent randomization of the sample time interval, and, because infinitely many estimation equations can be derived considering the other blocks, there is an extensive list of overidentifying restrictions.

Example: more details

The SDE: $dX_t = \kappa(\theta - X_t)dt + dW_t$

Assume that the stationary distribution, q, is known. Then, choosing an appropriate orthonormal coordinate system, we may assume that

$q(x) = \exp[-(\mu_1^2 x_1^2 + \mu_2^2 x_2)]$ and $\theta = 0$.

Need to identify κ.

Step 1. From the spectral decomposition procedure

$$\begin{bmatrix} 1/\mu_1 & 0 \\ 0 & 1/\mu_2 \end{bmatrix} \cdot \begin{bmatrix} \kappa_{11} & \kappa_{21} \\ \kappa_{12} & \kappa_{22} \end{bmatrix} + \begin{bmatrix} \kappa_{11} & \kappa_{12} \\ \kappa_{21} & \kappa_{22} \end{bmatrix} \cdot \begin{bmatrix} 1/\mu_1 & 0 \\ 0 & 1/\mu_2 \end{bmatrix} = 2I$$

Example, cont-d

From this equation

- $\kappa_{jj} = \mu_j, j = 1, 2,$
- $\mu_1\kappa_{12} + \mu_2\kappa_{21} = 0.$

Need one more equation

2–factor model

$\dim \mathcal{V}^m = m + 1, m = 0, 1, 2, \ldots$

OU process

$\mathcal{V}^m$ is the space of homogenous polynomials of degree m.

Example: the role of new estimating equations

Take any orthonormal basis $\{g_{10}, g_{01}\}$ in $\mathcal{V}^1 \subset L_2(q(x)dx)$

assume that the trace of the the restriction of $e^{\tau\mathcal{L}}$ on $\mathcal{V}^1$ is observable:

$$\begin{aligned} k &:= \mathrm{Tr}\, e^{\tau\mathcal{L}}|_{\mathcal{V}^1} \\ &= E[g_{10}(X_{t+\tau})g_{10}(X_t) \mid X_t] + E[g_{01}(X_{t+\tau})g_{01}(X_t) \mid X_t]. \end{aligned}$$

The eigenvalues of $\mathcal{L}^1$ are

$$\begin{aligned}\lambda_{10} &= -\Lambda/2 - (\mu_1 + \mu_2) + \frac{1}{2}\mathrm{Tr}\,\kappa = -\Lambda/2 - (\mu_1 + \mu_2)/2,\\ \lambda_{01} &= \Lambda/2 - (\mu_1 + \mu_2) + \frac{1}{2}\mathrm{Tr}\,\kappa = \Lambda/2 - (\mu_1 + \mu_2)/2,\end{aligned}$$

where $\Lambda = \sqrt{\alpha^2 - \beta^2}$, $\alpha = \mu_1 - \mu_2, \beta = B_{12}/2$, and B is the matrix explicitly calculated in terms of $\mu_j, j = 1, 2$, and κ. Therefore, we can calculate the trace

$$\begin{aligned}\mathrm{Tr}\, e^{\tau\mathcal{L}}|_{\mathcal{V}^1} &= e^{\tau\lambda_{10}} + e^{\tau\lambda_{01}}\\ &= e^{-\tau[(\mu_1+\mu_2)/2+\Lambda/2]} + e^{-\tau[(\mu_1+\mu_2)/2-\Lambda/2]},\end{aligned}$$

and obtain the second equation for κ_{12} and κ_{21}.

Possible cases

Spectral analysis of $\mathcal{L}^1$: three cases are possible:

(i) λ_{10} and λ_{20} are different negative reals, equivalently, Λ is real;

(ii) λ_{10} and λ_{20} are complex-adjoint (and not real), equivalently, $\Lambda = i\sqrt{\beta^2 - \alpha^2}$, where $i = \sqrt{-1}$;

(iii) $\lambda_{10} = \lambda_{01}$ are real, and $\Lambda = 0$.

In terms of the observed $k = \mathrm{Tr}\, e^{\tau\mathcal{L}}|_{\mathcal{V}^1}$ and $\mu_j, j = 1, 2$,

these three cases are:

(i) $k > 2e^{-\tau(\mu_1+\mu_2)/2}$;

(ii) $k < 2e^{-\tau(\mu_1+\mu_2)/2}$;

(iii) $k = 2e^{-\tau(\mu_1+\mu_2)/2}$,

Additional equation is

$$k = e^{-\tau(\mu_1+\mu_2)/2}\left(e^{\tau\Lambda/2} + e^{-\tau\Lambda/2}\right), \tag{3}$$

which we consider as a quadratic equation w.r.t. $y = e^{\tau\Lambda/2}$. Solving (3), we find β^2, whereupon the absolute value of $b = B_{12} = 2\beta$ can be found, and then, taking $\mu_1\kappa_{12} + \mu_2\kappa_{21} = 0$ into account, we obtain that either

$$\kappa_{12} = 2\beta\sqrt{\frac{\mu_1}{\mu_2}}, \quad \kappa_{21} = -2\beta\sqrt{\frac{\mu_2}{\mu_1}}, \tag{4}$$

or

$$\kappa_{12} = -2\beta\sqrt{\frac{\mu_1}{\mu_2}}, \quad \kappa_{21} = 2\beta\sqrt{\frac{\mu_2}{\mu_1}}. \tag{5}$$

It remains to find the sign of β

Introduce

$$g_{10}(x) = \sqrt{\frac{2\mu_1}{\pi}} x_1, \quad g_{01}(x) = \sqrt{\frac{2\mu_2}{\pi}} x_2,$$

the coordinate functions in the basis of eigenvectors of $\log q$. Then the sign of $E[g_{10}(X_{t+\tau})g_{01}(X_t) \mid X_t]$ coincides with the sign of the covariance function $E[X_{t+\tau,1}X_{t,2} \mid X_t]$. Assuming that the sign of the latter is observable, we can find the sign of β calculating $(e^{\tau \mathcal{L}_q^1} g_{10}, g_{01})_q$. This can be done using the spectral decomposition.

Second group of estimating equations

Non-selfadjoint case

- no variational principle for eigenvalues and eigenfunctions of $e^{\tau\mathcal{L}}$ except for the leading ones
- can use variational principles for the self-adjoint operators F^j associated with the quadratic form $(e^{\tau\mathcal{L}}u, u)_q$ restricted to the invariant subspaces $\mathcal{V}_j$
- since the matrix of $e^{\tau\mathcal{L}^j}$ has been calculated, we can calculate the matrix of F^j and find its eigenfunctions and eigenvalues in terms of the parameters of the model
- the second group of estimating equations results

General scheme of the spectral decomposition for QTSMs

Infinitesimal generator

$$\mathcal{L} = (\kappa(\theta - x), \partial_x) + \frac{1}{2}(\Sigma\partial_x, \Sigma\partial_x) - r(x),$$

where $(A, B) = \sum_{j=1}^{n} A_j B_j$,

κ is anti-stable matrix,

Σ is invertible

$\theta \in \mathbf{R}^n$

$r(x) = d_0 + (d_1, x) + \frac{1}{2}(\Gamma x, x)$

where $\Gamma \geq 0$.

Special cases

$\Gamma = 0$: ATSM of class $A_0(n)$;

$\Gamma = 0, d = 0, d_0 = 0$: OU process

Step I. Reduction to a model operator

Affine change of variables

and reduction to an operator of the form

$$\mathcal{L} = (\kappa(\theta - x), \partial_x) + \frac{1}{2}\partial_x^2 - d_0 - \frac{1}{2}(\Gamma x, x),$$

with new θ, d_0, Γ.

Conjugation with the exponential

of an appropriate quadratic function $\exp(Wx, x)/2$:

$$\mathcal{L}_0 := e^{-(Wx,x)/2}\mathcal{L}e^{-(Wx,x)/2}$$

and reduction to the infinitesimal generator of the special form

Step I, cont-d

Special form

$$\mathcal{L}_0 = \sum_{j=1}^{n} \mu_j \mathcal{L}_j - \sum_{j>k} b_{jk} J_{jk} - \tilde{d}_0, \qquad (6)$$

where $J_{jk} = x_j \partial_k - x_k \partial_j$, $\mathcal{L}_j = \frac{1}{2}(\partial^2_{x_j} - x_j^2)$, $\tilde{d}_0$ is a constant.

Construction of the matrix W

$W = Y - Z$, where Y is the solution of the continuous algebraic Riccati equation

$$Y^2 + Y\kappa + \kappa^T Y - \Gamma = 0 \qquad (7)$$

s.t. $\kappa + Y$ is anti-stable; Z^{-1} is the positive-definite solution of Lyapunov equation

$$(\kappa + Y)X + X(\kappa + Y)^T Z - 2I_n = 0. \qquad (8)$$

Step I, finish

- $\{\psi_j\}_{j=1}^n$– an orthogonal basis of eigenvectors of Z
- $\{\mu_j\}$ – the corresponding eigenvalues
- y– the coordinates of x in the basis $\{\psi_j\}$
- C– the transformation matrix: $x = Cy$

In y–coordinates

$$\mathcal{L}_0 = \frac{1}{2}\sum_{j=1}^{n}\left(\partial_{y_j}^2 - \mu_j^2(y_j)^2\right) - (By, \partial_y) - \tilde{d}_0,$$

where

$$B := C^T Z^{1/2}(\kappa + Y)Z^{-1/2}C$$

is skew-symmetric

Step II. Block diagonalization

Creation and annihilation operators

$$z_j = \frac{1}{\sqrt{2}}(x_j - \partial_{x_j}),$$

$$z_j^* = \frac{1}{\sqrt{2}}(x_j + \partial_{x_j}).$$

In terms of z_j and z_j^*

$$\mathcal{L}_0 = -\sum_{j=1}^{n} \mu_j (z_j z_j^* + 1/2) - \sum_{j>k} b_{jk} J_{jk} - \tilde{d}_0,$$

where $J_{jk} = z_j z_k^* - z_k z_j^*$

Step II cont-d.

Commutation relations

$$[z_j, z_k] = [z_j^*, z_k^*] = 0, \quad [z_j^*, z_k] = \delta_{jk}.$$

Notation

- $H_n(x)$–Hermite polynomials
- $w_n(x) = 2^{-n/2} H_n(x) e^{-x^2/2}$
- $f_m = (m!\sqrt{\pi})^{-1/2} w_m, m = 0, 1, \ldots$
- for a multi-index $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbf{Z}_+^n$, set

$$f_\alpha := f_{\alpha_1} \otimes \cdots \otimes f_{\alpha_n}$$

- $V_m = \mathrm{span}\,\{f\alpha, |\alpha| = m\}$

Step II cont-d.

Facts

- $w_k(x) = z^k w_0, k = 0, 1, \ldots,$ are eigenfunctions of operator

$$\frac{1}{2}\left(\frac{d^2}{dx^2} - x^2\right) = -zz^* - 1/2$$

  with eigenvalues $\nu_k = -k - 1/2$

- $\{f_m\}_{m=0}^{+\infty}$ is an orthonormal basis in $L_2(\mathbf{R})$
- $\{f_\alpha\}_{\alpha \in \mathbf{Z}_+^n}$ is an orthonormal basis in $L_2(\mathbf{R}^n)$
- V_m is invariant under $\mathcal{L}_j = -z_j z_j^* - 1/2$ and J_{jk}, hence, under $\mathcal{L}_0$

Step II cont-d. Block-diagonalization

Notation:

$\mathcal{L}_m$ – the restriction of $\mathcal{L}_0$ on V^m.

Theorem

- $L_2(\mathbf{R}^n) = \oplus_{m=0}^{\infty} V_m$
- $\mathcal{L}_0 = \oplus_{m=0}^{\infty} \mathcal{L}^m$