MAXIMUM LIKELIHOOD ESTIMATION FOR DISCRETE MULTIVARIATE VASICEK PROCESSES

Michael Pokojovy, Ebenezer Nkum and Thomas M. Fullerton Jr.

This supplementary documents contains additional figure and proofs referred to in the paper.

S1 Supplemental Figure

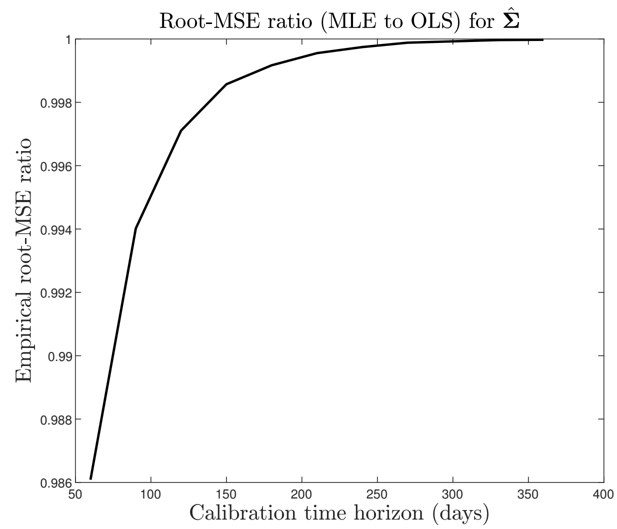

Figure S(1): Simulated root-MSE ratio (MLE to OLS)

S2 Auxiliary Results

We briefly present and prove some auxiliary results for vector and matrix differentiation below. Recall $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \boldsymbol{x}'\boldsymbol{y} = \boldsymbol{y}'\boldsymbol{x}$ for $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^p$.

Lemma S2.1. *Suppose $\boldsymbol{A} \in \mathbb{R}^{p \times p}$. Then the function*

$$f(\boldsymbol{x}) = \langle \boldsymbol{A}\boldsymbol{x}, \boldsymbol{x} \rangle \text{ is smooth with } \tfrac{\partial}{\partial \boldsymbol{x}} f(\boldsymbol{x}) = 2\boldsymbol{A}\boldsymbol{x}. \tag{S2.1}$$

Proof. We observe

$$f(\boldsymbol{x}+\boldsymbol{h}) = \langle \boldsymbol{A}(\boldsymbol{x}+\boldsymbol{h})\big), \boldsymbol{x}+\boldsymbol{h} \rangle = \langle \boldsymbol{A}\boldsymbol{x}, \boldsymbol{x} \rangle + 2\langle \boldsymbol{A}\boldsymbol{x}, \boldsymbol{h} \rangle + \langle \boldsymbol{A}\boldsymbol{h}, \boldsymbol{h} \rangle \tag{S2.2}$$

$$= f(\boldsymbol{x}) + 2\langle \boldsymbol{A}\boldsymbol{x}, \boldsymbol{h}\rangle + o\big(|\boldsymbol{h}|\big) \text{ as } \boldsymbol{h} \to \boldsymbol{0}, \tag{S2.3}$$

whence the claim trivially follows with Frechét definition of the derivative. □

Lemma S2.2. *Suppose $\boldsymbol{\Sigma} \in \mathbb{R}^{p\times p}$ is symmetric, $\boldsymbol{A} \in \mathbb{R}^{p\times p}$ and $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^p$, we consider the following functions:*

$$f(\boldsymbol{A}) = \langle \boldsymbol{\Sigma}^{-1}\boldsymbol{u}, \boldsymbol{A}\boldsymbol{v}\rangle \quad \textit{and} \quad g(\boldsymbol{A}) = \big(\boldsymbol{A}\boldsymbol{v}\big)'\boldsymbol{\Sigma}^{-1}\big(\boldsymbol{A}\boldsymbol{v}\big) = \big|\boldsymbol{\Sigma}^{-1/2}(\boldsymbol{A}\boldsymbol{v})\big|^2. \tag{S2.4}$$

Then $f(\cdot)$ and $g(\cdot)$ are smooth with

$$\tfrac{\partial}{\partial \boldsymbol{A}} f(\boldsymbol{A}) = (\boldsymbol{\Sigma}^{-1}\boldsymbol{u})\boldsymbol{v} \quad \textit{and} \quad \tfrac{\partial}{\partial \boldsymbol{A}} g(\boldsymbol{A}) = 2\boldsymbol{\Sigma}^{-1}\boldsymbol{A}(\boldsymbol{v}\boldsymbol{v}'). \tag{S2.5}$$

Proof. For $\boldsymbol{H}, \boldsymbol{A} \in \boldsymbol{R}^{p\times p}$, as $\boldsymbol{H} \to 0$, we consider $f(\boldsymbol{A}+\boldsymbol{H})$ and $g(\boldsymbol{A}+\boldsymbol{H})$, and employ Einstein's summation convention to obtain

$$\begin{aligned}
f(\boldsymbol{A}+\boldsymbol{H}) &= \big\langle \boldsymbol{\Sigma}^{-1}\boldsymbol{u}, (\boldsymbol{A}+\boldsymbol{H})\boldsymbol{v}\big\rangle = \big\langle \boldsymbol{\Sigma}^{-1}\boldsymbol{u}, \boldsymbol{A}\boldsymbol{v}\big\rangle + \big\langle \boldsymbol{\Sigma}^{-1}\boldsymbol{u}, \boldsymbol{H}\boldsymbol{v}\big\rangle && \text{(S2.6)}\\
&= f(\boldsymbol{A}) + \big((\boldsymbol{\Sigma}^{-1})_{ij}u_j\big)\big(\mathbf{H}_{ik}\boldsymbol{v}_k\big) && \text{(S2.7)}\\
&= f(\boldsymbol{A}) + \big((\boldsymbol{\Sigma}^{-1})_{ij}\boldsymbol{u}_j\boldsymbol{v}_k\big)\big(\mathbf{H}_{ik}\big), && \text{(S2.8)}\\
g(\boldsymbol{A}+\boldsymbol{H}) &= \big\langle \boldsymbol{\Sigma}^{-\frac{1}{2}}(\boldsymbol{A}+\boldsymbol{H})\boldsymbol{v}, \boldsymbol{\Sigma}^{-\frac{1}{2}}(\boldsymbol{A}+\boldsymbol{H})\boldsymbol{v}\big\rangle && \text{(S2.9)}\\
&= \big\langle \boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{A}\boldsymbol{v}, \boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{A}\boldsymbol{v}\big\rangle + 2\big\langle \boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{A}\boldsymbol{v}, \boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{H}\boldsymbol{v}\big\rangle + o\big(\|\boldsymbol{H}\|_{\mathcal{F}}\big) && \text{(S2.10)}\\
&= g(\boldsymbol{A}) + 2\big\langle \boldsymbol{\Sigma}^{-1}\boldsymbol{A}\boldsymbol{v}, \boldsymbol{H}\boldsymbol{v}\big\rangle + o\big(\|\boldsymbol{H}\|_{\mathcal{F}}\big) && \text{(S2.11)}\\
&= g(\boldsymbol{A}) + 2\big\langle \boldsymbol{\Sigma}^{-1}\boldsymbol{A}\boldsymbol{v}\boldsymbol{v}', \boldsymbol{H}\big\rangle + o\big(\|\boldsymbol{H}\|_{\mathcal{F}}\big), && \text{(S2.12)}
\end{aligned}$$

whence the claim trivially follows from the definition of Frechét derivative. □

Chain rule and Jacobi's theorem yield the following result:

Lemma S2.3. *For any non-singular matrix $\boldsymbol{H} \in \mathbb{R}^{p\times p}$, we have*

$$\tfrac{\partial}{\partial \boldsymbol{H}}|\boldsymbol{H}| = |\boldsymbol{H}|(\boldsymbol{H}^{-1})' \quad \textit{and} \quad \tfrac{\partial}{\partial \boldsymbol{H}}\log|\boldsymbol{H}| = (\boldsymbol{H}^{-1})'. \tag{S2.13}$$

If $\boldsymbol{H}$ is symmetric, we additionally have

$$\tfrac{\partial}{\partial \boldsymbol{H}}|\boldsymbol{H}| = |\boldsymbol{H}|\boldsymbol{H}^{-1} \quad \textit{and} \quad \tfrac{\partial}{\partial \boldsymbol{H}}\log|\boldsymbol{H}| = \boldsymbol{H}^{-1}. \tag{S2.14}$$

Lemma S2.4. *Let $\boldsymbol{\Sigma}$ be symmetric non-singular, $M \in \mathbb{R}^{p\times p}$ be arbitrary and*

$$f(\boldsymbol{\Sigma}) = \langle M, \boldsymbol{\Sigma}^{-1}\rangle_{\mathcal{F}}, \tag{S2.15}$$

where the Frobenius scalar product is defined as $\langle \boldsymbol{A}, \boldsymbol{B} \rangle_{\mathcal{F}} = \mathrm{trace}(\boldsymbol{A}'\boldsymbol{B})$. *Then* $f(\cdot)$ *is smooth with*

$$\tfrac{\partial}{\partial \boldsymbol{\Sigma}} f(\boldsymbol{\Sigma}) = -\boldsymbol{\Sigma}^{-1}\boldsymbol{M}\boldsymbol{\Sigma}^{-1}. \tag{S2.16}$$

Proof. For $\boldsymbol{H} \in \mathbb{R}^{p\times p}$ symmetric, as $\boldsymbol{H} \to 0$, von Neumann series yields:

$$f\big(\boldsymbol{\Sigma}+\boldsymbol{H}\big) = \big\langle \boldsymbol{M}, (\boldsymbol{\Sigma}+\boldsymbol{H})^{-1}\big\rangle_{\mathcal{F}} = \big\langle \boldsymbol{M}, \big(\boldsymbol{\Sigma}(\boldsymbol{I}+\boldsymbol{\Sigma}^{-1}\boldsymbol{H})\big)^{-1}\big\rangle_{\mathcal{F}} \tag{S2.17}$$

$$= \big\langle \boldsymbol{M}, \big(\boldsymbol{I}+\boldsymbol{\Sigma}^{-1}\boldsymbol{H}\big)^{-1}\boldsymbol{\Sigma}^{-1}\big\rangle_{\mathcal{F}} = \big\langle \boldsymbol{M}, \big(\boldsymbol{I}-\boldsymbol{\Sigma}^{-1}\boldsymbol{H} + o\big(\|\boldsymbol{H}\|_{\mathcal{F}}\big)\big)\boldsymbol{\Sigma}^{-1}\big\rangle_{\mathcal{F}} \tag{S2.18}$$

$$= \big\langle \boldsymbol{M}, \boldsymbol{\Sigma}^{-1}\big\rangle - \big\langle \boldsymbol{M}, \boldsymbol{\Sigma}^{-1}\boldsymbol{H}\boldsymbol{\Sigma}^{-1}\big\rangle_{\mathcal{F}} \tag{S2.19}$$

$$= f(\boldsymbol{\Sigma}) - \big\langle \boldsymbol{\Sigma}^{-1}\boldsymbol{M}\boldsymbol{\Sigma}^{-1}, \boldsymbol{H}\big\rangle_{\mathcal{F}} + o\big(\|\boldsymbol{H}\|_{\mathcal{F}}\big). \tag{S2.20}$$

□