The Derivation of the Hessian Matrix of exp(A) with a Symmetric Real Matrix A

Yoshihisa BABA

I. Introduction

The problem of the parameterization which ensures a variance-covariance matrix to be positive definite has been extensively discussed in the field of statistics. There are two approaches in estimating variance-covariance matrices—constrained optimization approach and unconstrained parameterization apporach. Pinheiro and Bates (1996) argue that constrained optimization would be generally difficult and moreover the statistical properties of constrained estimates would not be easy to characterize. Pourahmadi (1999) also discusses that issue and proposes an unconstrained parameterization.

One of unconstrained parameterizations which ensues the positive definite variance covariance matrix is the utilization of the matrix exponential since the matrix exponential of any real symmetric matrix is positive definite. Chiu (1994) discusses various issues about exponential covariance models. Linton (1993) derives analytic derivatives of a matrix exponential with respect to a symmetric matrix in the context of a multivariate exponential GARCH model. Chen and Zadrozny (2001) also derives analytical derivatives of the matrix exponential in estimating linear continuous-time models.

Those papers mentioned above derive only the first derivative of the matrix exponential but not the second derivative. Vinod (2000) emphasized the importance of using the analytical derivatives to preserve numerical accuracy in deriving numerical solutions for non-linear problems. The reliance on numerical derivatives in actual computings may induce unreliable numerical estimates. Therefore, it seems to be worthwhile to obtain the analytical expression for the Hessian matrix of the matrix exponential.

The purpose of the present paper is to derive the gradient and Hessian of the matrix exponential. The rest of the paper is as follows: Some definitions and properties of matrices including the matrix exponential are briefly reviewed in section II. The gradient and Hessian of the matrix exponential are derived in section III and section IV concludes the paper.

II. Some Definitions and Properties of Matrices

We introduce the matrices, which are frequently used throughout the present paper. Let A is a $n \times n$ symmetric real matrix. The exponential of A is defined as

$$\exp(A) = \sum_{m=0}^{\infty} \frac{A^m}{m!} \tag{2.1}$$

Chiu (1994) drives various properties of the matrix exponential including that $\exp(A)$ is positive definite. The $(n \times m)$ commutation matrix K_{nm} is defined such that

$K_{nm} vec(B) = vec(B')$ for a $n \times m$ matrix, B.

The $\left(n^2 \times \frac{1}{2} n(n+1)\right)$ duplication matrix D_n matrix is defined such that

$vec(A) = D_n vech(A)$

The following properties of the Kronocker Product and the vec operator are very useful (see Lütkepohl (1996) for example).

P1 $vec(ABC) = (C' \otimes A) vec(B)$.

P2 $(A \otimes B)(C \otimes D) = (AC \otimes BD)$ if A and C are conformable matrices and B and D are also conformable matrices.

P3 Let A and a be a $n \times n$ matrix and a p x 1 vector respectively. Then,
$K_{nn}(A \otimes a) = a \otimes A$

P4 Let A and B be $n \times n$ matrices. Then,
$vec(A \otimes B) = (I_n \otimes K_{nn} \otimes I_n)(vecA \otimes vecB)$

Let x be a k dimensional column vector and B be a $n \times n$ matrix which a function of x. Following Magnus and Neudecker (1999), we define the first and second derivatives of B with respect to x as follows and denote those matrices with J and H respectively:

$$J = \frac{\partial vecB}{\partial x'}$$

$$H = \frac{\partial}{\partial x'} vec\left(\frac{\partial vecB}{\partial x'}\right)'$$

III. Main Results

Following the definitions in the previous section, the first and second partial derivatives of $\exp(A)$ with respect to distinct elements of a matrix A are given by

$$J = \frac{\partial vec(\exp(A))}{\partial (vechA)'} \left(\text{a } n^2 \times \frac{1}{2} n(n+1) \text{ matirx}\right) \tag{3.1}$$

$$H = \frac{\partial}{\partial (vechA)'} vec\left(\frac{\partial vec(\exp(A))}{\partial (vechA)'}\right)' \left(a \ \frac{1}{2} n^3(n+1) \times \frac{1}{2} n(n+1) \text{ matrix}\right) \tag{3.2}$$

The following partial derivatives are necessary to derive the expressions for matrices J and H from equation (2.1).

$$\frac{\partial vec(A^m)}{\partial (vechA)'}=\sum_{j=0}^{m-1}(A^{m-1-j}\otimes A^j)D_n \tag{3.3}$$

Since A is a symmetric matrix,

$$\left(\sum_{j=0}^{m-1}(A^{m-1-j}\otimes A^j)D_n\right)'=\sum_{j=0}^{m-1}D_n'(A^{m-1-j}\otimes A^j)$$

$$\frac{\partial}{\partial (vechA)'}vec\left(\frac{\partial vec(A^m)}{\partial (vechA)'}\right)'=\frac{\partial}{\partial (vechA)'}\left(\sum_{j=0}^{m-1}(I_{n^2}\otimes D_n')\,vec(A^{m-1-j}\otimes A^j)\right)$$

$$=\sum_{j=0}^{m-1}(I_{n^2}\otimes D_n')(I_n\otimes K_{nn}\otimes I_n)\left\{\frac{\partial vec(A^{m-1-j})}{\partial (vechA)'}\otimes vecA^j+vecA^{m-1-j}\otimes\frac{\partial vec(A^j)}{\partial (vechA)'}\right\}$$

$$=\sum_{j=0}^{m-1}(I_{n^2}\otimes D_n')(I_n\otimes K_{nn}\otimes I_n)\left[\sum_{k=0}^{m-2-j}\{(A^{m-2-j-k}\otimes A^k)D_n\}\otimes vecA^j\right.$$

$$\left.+vecA^{m-1-j}\otimes\left\{\sum_{s=0}^{j-1}(A^{j-1-s}\otimes A^s)D_n\right\}\right] \tag{3.4}$$

Define a matrix with eigenvectors of A, Q and a diagonal matirx with eigenvalues $(\lambda_1>\lambda_2>\cdots>\lambda_n)$, Λ as follows:

$A=Q\Lambda Q'$ and $vecA=(Q\otimes Q)\,vec\Lambda$ by P1.

In order to obtain the closed from of J, we need the following lemmas.

Lemma 1 For $k=0, 1, 2, \cdots$, $\sum_{j=0}^{k}(\Lambda^{k-j}\otimes\Lambda^j)=F(k)$ where $F(k)$ is a $n^2\times n^2$ diagonal matrix with $(n(l_1-1)+l_2)$ 'th diagonal $(l_1,\ l_2=1, 2, \cdots, n)$:

$(\lambda_{l_1}^k+\lambda_{l_1}^{k-1}\lambda_{l_2}+\cdots+\lambda_{l_2}^k)=\dfrac{\lambda_{l_1}^{k+1}-\lambda_{l_2}^{k+1}}{\lambda_{l_1}-\lambda_{l_2}}$ when $l_1\neq l_2$ and $(k+1)\lambda_{l_1}^k$ otherwise.

(proof)

See Linton (1993, 1995).

Lemma 2 Let $F=\sum_{k=0}^{\infty}\dfrac{1}{(k+1)!}F(k)$ where $F(k)$ is defined in Lemma 1. Then F is a $n^2\times n^2$ diagonal matrix with $(n(l_1-1)+l_2)$'th diagonal $(l_1,\ l_2=1, 2, \cdots, n)$:

$\dfrac{e^{\lambda_{l1}}-e^{\lambda_{l2}}}{\lambda_{l_1}-\lambda_{l_2}}$ if $l_1\neq l_2$ and $e^{\lambda_{l_1}}$ if $l_1=l_2$.

(proof)

Let the i'th diagonal element of F be f_{ii}. For $i=(n(l_1-1)+l_2)$ and $l_1\neq l_2(l_1,\ l_2=1, 2, \cdots, n)$,

$$f_{ii}=\frac{1}{\lambda_{l_1}-\lambda_{l_2}}\sum_{k=0}^{\infty}\frac{1}{(k+1)!}(\lambda_{l_1}^{k+1}-\lambda_{l_2}^{k+1})=\frac{e^{\lambda_{l1}}-e^{\lambda_{l2}}}{\lambda_{l_1}-\lambda_{l_2}}$$

For $i=(n(l_1-1)+l_2)$ and $l_1=l_2$,

$$f_{ii}=\sum_{k=0}^{\infty}\frac{1}{(k+1)!}(k+1)\lambda_{i_1}^{k}=\sum_{k=0}^{\infty}\frac{\lambda_{i_1}^{k}}{k!}=e^{\lambda_{i1}}$$

q. e. d.

Proposition 1 $J=\{(Q\otimes Q)F(Q'\otimes Q')\}D_n$ where F is defined in Lemma 2.

(proof)

By Lemma 1,

$$\sum_{j=0}^{m-1}(A^{m-1-j}\otimes A^{j})=(Q\otimes Q)\sum_{j=0}^{m-1}(\Lambda^{m-1-j}\otimes\Lambda^{j})(Q'\otimes Q')=(Q\otimes Q)F(m-1)(Q'\otimes Q')$$

From equation (3.1) and (3.3),

$$J=\sum_{m=0}^{\infty}\frac{1}{m!}\{(Q\otimes Q)F(m-1)(Q'\otimes Q')\}D_n$$

$$=\left\{(Q\otimes Q)\sum_{m=1}^{\infty}\frac{1}{m!}F(m-1)(Q'\otimes Q')\right\}D_n=\{(Q\otimes Q)F(Q'\otimes Q')\}D_n$$

by Lemma 2 in which F is defined.

q. e. d.

In order to obtain the Hessian matrix H, we need closed expressions for $\sum_{k=0}^{m-2-j}\{(A^{m-2-j-k}\otimes A^{k})D_n\}\otimes vecA^{j}$ and $vecA^{m-1-j}\otimes\left\{\sum_{s=0}^{j-1}(A^{j-1-s}\otimes A^{s})D_n\right\}$ in equation (3.4). The following lemmas give those expressions.

Lemma 3 The following equations hold:

$$\sum_{j=0}^{m-1}\sum_{k=0}^{m-2-j}\{(A^{m-2-j-k}\otimes A^{k})D_n\}\otimes vecA^{j}$$

$$=K_{n^2,n^2}(Q\otimes Q\otimes Q\otimes Q)\sum_{j=0}^{m-2}(vec\Lambda^{j}\otimes F(m-2-j))((Q'\otimes Q')D_n)K_{\frac{1}{2}n(n+1),1}$$

$$\sum_{j=0}^{m-1}vecA^{m-1-j}\otimes\left\{\sum_{s=0}^{j-1}(A^{j-1-s}\otimes A^{s})D_n\right\}$$

$$=(Q\otimes Q\otimes Q\otimes Q)\sum_{j=0}^{m-2}(vec\Lambda^{j}\otimes F(m-2-j))((Q'\otimes Q')D_n)$$

(proof)

$A^{j}=Q\Lambda^{j}Q'$ and $vecA^{j}=(Q\otimes Q)vec\Lambda^{j}$ since $QQ'=I$. Together with P1, P2, and, P3, this yields the following results.

$$\sum_{k=0}^{m-2-j}\{(A^{m-2-j-k}\otimes A^{k})D_n\}\otimes vecA^{j}$$

$$=\sum_{k=0}^{m-2-j}K_{n^2,n^2}[(vecA^{j}\otimes\{(A^{m-2-j-k}\otimes A^{k})D_n\})]$$

$$=\sum_{k=0}^{m-2-j} K_{n^2,n^2}[\{(Q\otimes Q)\, vec\Lambda^j\}\otimes\{(Q\Lambda^{m-2-j-k}Q')\otimes(Q\Lambda^k Q')D_n\}]$$

$$=K_{n^2,n^2}\sum_{k=0}^{m-2-j}[\{(Q\otimes Q)\, vec\Lambda^j\}\otimes\{((Q\otimes Q)(\Lambda^{m-2-j-k}Q'\otimes\Lambda^k Q'))D_n\}]$$

$$=K_{n^2,n^2}\sum_{k=0}^{m-2-j}[\{(Q\otimes Q)\, vec\Lambda^j\}\otimes\{((Q\otimes Q)(\Lambda^{m-2-j-k}\otimes\Lambda^k)(Q'\otimes Q'))D_n\}]$$

$$=K_{n^2,n^2}(Q\otimes Q\otimes Q\otimes Q)\sum_{k=0}^{m-2-j}(vec\Lambda^j\otimes\Lambda^{m-2-j-k}\otimes\Lambda^k)((Q'\otimes Q')D_n)$$

$$=K_{n^2,n^2}(Q\otimes Q\otimes Q\otimes Q)(vec\Lambda^j\otimes F(m-2-j))((Q'\otimes Q')D_n)$$

Thus,

$$\sum_{j=0}^{m-1}\sum_{k=0}^{m-2-j}\{(A^{m-2-j-k}\otimes A^k)D_n\}\otimes vecA^j$$

$$=K_{n^2,n^2}(Q\otimes Q\otimes Q\otimes Q)\sum_{j=0}^{m-1}(vec\Lambda^j\otimes F(m-2-j))((Q'\otimes Q')D_n)$$

$$=K_{n^2,n^2}(Q\otimes Q\otimes Q\otimes Q)\sum_{j=0}^{m-2}(vec\Lambda^j\otimes F(m-2-j))((Q'\otimes Q')D_n)$$

since $F(m-2-j)$ can be defined only for $m-2-j\geq 0$.

$$vecA^{m-1-j}\otimes\left\{\sum_{s=0}^{j-1}(A^{j-1-s}\otimes A^s)D_n\right\}$$

$$=\{(Q\otimes Q)\, vec\Lambda^{m-1-j}\}\otimes\left\{\sum_{s=0}^{j-1}((Q\Lambda^{j-1-s}Q')\otimes(Q\Lambda^s Q'))D_n\right\}$$

$$=\{(Q\otimes Q)\, vec\Lambda^{m-1-j}\}\otimes[\{(Q\otimes Q)\sum_{s=0}^{j-1}(\Lambda^{j-1-s}\otimes\Lambda^s)(Q'\otimes Q')\}D_n]$$

$$=(Q\otimes Q\otimes Q\otimes Q)(vec\Lambda^{m-1-j}\otimes F(j-1))((Q'\otimes Q')D_n)$$

$$\sum_{j=0}^{m-1}vecA^{m-1-j}\otimes\{\sum_{s=0}^{j-1}(A^{j-1-s}\otimes A^s)D_n\}$$

$$=(Q\otimes Q\otimes Q\otimes Q)\sum_{j=0}^{m-1}(vec\Lambda^{m-1-j}\otimes F(j-1))((Q'\otimes Q')D_n)$$

Let $j'=m-1-j$, and so $j-1=m-2-j'$ and $F(j-1)$ can be defined only for $j-1\geq 0$.

$$\sum_{j=0}^{m-1}(vec\Lambda^{m-1-j}\otimes F(j-1))$$

$$=\sum_{j'=0}^{m-1}(vec\Lambda^{j'}\otimes F(m-2-j'))$$

$$=\sum_{j'=0}^{m-2}(vec\Lambda^{j'}\otimes F(m-2-j'))$$

since $m-2-j'\geq 0$.

Thus,

$$\sum_{j=0}^{m-1}vecA^{m-1-j}\otimes\left\{\sum_{s=0}^{j-1}(A^{j-1-s}\otimes A^s)D_n\right\}$$

$$=(Q\otimes Q\otimes Q\otimes Q)\sum_{j=0}^{m-2}(vec\Lambda^j\otimes F(m-2-j))((Q'\otimes Q')D_n)$$

q. e. d.

Let $G(k)=\sum_{j=0}^{k} vec\Lambda^{j}\otimes F(k-j)$. Lemma 4 and 5 give the expressions for $G(k)$ and $\sum_{k=0}^{\infty}\frac{1}{(k+2)!}G(k)$.

Lemma 4 Let e_i be a $n\times 1$ unit vector with i'th elment being one and others zero and $\underset{l\times m}{\mathbf{0}}$ be the $l\times m$ zero matrix. The expression for $G(k)$ is given by

$$G(k) = \begin{bmatrix} \sum_{j=0}^{k}\lambda_1^j e_1\otimes F(k-j) \\ \sum_{j=0}^{k}\lambda_2^j e_2\otimes F(k-j) \\ \cdot \\ \cdot \\ \cdot \\ \sum_{j=0}^{k}\lambda_n^j e_n\otimes F(k-j) \end{bmatrix} = \begin{bmatrix} G(k,\ 1) \\ G(k,\ 2) \\ \vdots \\ G(k,\ n) \end{bmatrix}$$

where

$$G(k,\ i) = \begin{bmatrix} \underset{(i-1)n^2\times n^2}{\mathbf{0}} \\ \sum_{j=0}^{k}\lambda_i^j F(k-j) \\ \underset{(n-i)n^2\times n^2}{\mathbf{0}} \end{bmatrix}$$

(proof)

$$G(k)=\sum_{j=0}^{k} vec\Lambda^{j}\otimes F(k-j).$$

$G(k)$ can be partitioned as

$$G(k) = \begin{bmatrix} G(k,\ 1) \\ G(k,\ 2) \\ \vdots \\ G(k,\ n) \end{bmatrix}$$

where

$$G(k,\ i) = \begin{bmatrix} \underset{(i-1)n^2\times n^2}{\mathbf{0}} \\ \sum_{j=0}^{k}\lambda_i^j F(k-j) \\ \underset{(n-i)n^2\times n^2}{\mathbf{0}} \end{bmatrix}$$

Let $\Gamma(k,\ i)=\sum_{j=0}^{k}\lambda_i^j F(k-j)$ and $\Gamma(k,\ i)_{ll}$ be its l' diagonal element ($l=1,\ 2,\ \cdots,\ n^2$). The expressions for $\Gamma(k,\ i)_{ll}(l=n(l_1-1)+l_2,\ l_1,\ l_2=1,\ 2,\ \cdots,\ n)$, are given below:

(i) if $i\neq l_1$, $i\neq l_2$ and $l_1\neq l_2$,

$$\Gamma(k,\ i)_{ll}=\sum_{j=0}^{k}\lambda_i^j\frac{\lambda_{l_1}^{k+1-j}-\lambda_{l_2}^{k+1-j}}{\lambda_{l_1}-\lambda_{l_2}}$$

$$=\frac{1}{\lambda_{l_1}-\lambda_{l_2}}\left\{\sum_{j=0}^{k}\lambda_i^j\lambda_{l_1}^{k+1-j}-\sum_{j=0}^{k}\lambda_i^j\lambda_{l_2}^{k+1-j}\right\}$$

$$=\frac{1}{\lambda_{l_1}-\lambda_{l_2}}\left\{\frac{\lambda_{l_1}^{k+2}-\lambda_i^{k+2}}{\lambda_{l_1}-\lambda_i}-\lambda_i^{k+1}-\frac{\lambda_{l_2}^{k+2}-\lambda_i^{k+2}}{\lambda_{l_2}-\lambda_i}-\lambda_i^{k+1}\right\}$$

$$=\frac{\lambda_{l_1}^{k+2}}{(\lambda_{l_1}-\lambda_{l_2})(\lambda_{l_1}-\lambda_i)}-\frac{\lambda_{l_2}^{k+2}}{(\lambda_{l_1}-\lambda_{l_2})(\lambda_{l_2}-\lambda_i)}+\frac{\lambda_i^{k+2}}{(\lambda_{l_1}-\lambda_i)(\lambda_{l_2}-\lambda_i)}$$

(ii) if $i=l_1$ and $i\neq l_2$,

$$\Gamma(k,\ i)_{ll}=\sum_{j=0}^{k}\lambda_i^j\frac{\lambda_i^{k+1-j}-\lambda_{l_2}^{k+1-j}}{\lambda_i-\lambda_{l_2}}$$

$$=\frac{1}{\lambda_i-\lambda_{l_2}}\{(\lambda_i^{k+1}-\lambda_{l-2}^{k+1})+\lambda_i(\lambda_i^k-\lambda_{l-2}^k)+\cdots+\lambda_i^{k-1}(\lambda_i^2-\lambda_{l-2}^2)+\lambda_i^k(\lambda_i-\lambda_{l-2})\}$$

$$=\frac{1}{\lambda_i-\lambda_{l_2}}\left\{(k+2)\lambda_i^{k+1}-\frac{\lambda_{l_2}^{k+2}-\lambda_i^{k+2}}{\lambda_{l_2}-\lambda_i}\right\}$$

$$=\frac{(k+2)\lambda_i^{k+1}}{\lambda_i-\lambda_{l_2}}+\frac{\lambda_{l_2}^{k+2}-\lambda_i^{k+2}}{(\lambda_i-\lambda_{l_2})^2}$$

(iii) if $i=l_2$ and $i\neq l_1$,

$$\Gamma(k,\ i)_{ll}=\sum_{j=0}^{k}\lambda_i^j\frac{\lambda_{l_1}^{k+1-j}-\lambda_i^{k+1-j}}{\lambda_{l_1}-\lambda_i}$$

$$=\frac{1}{\lambda_{l_1}-\lambda_i}\left\{-(k+1)\lambda_i^{k+1}+\sum_{j=0}^{k}\lambda_i^j\lambda_{l_1}^{k+1-j}\right\}$$

$$=\frac{1}{\lambda_{l_1}-\lambda_i}\left\{-(k+1)\lambda_i^{k+1}+\frac{\lambda_{l_1}^{k+2}-\lambda_i^{k+2}}{\lambda_{l_1}-\lambda_i}-\lambda_i^{k+1}\right\}$$

$$=\frac{-(k+2)\lambda_i^{k+1}}{\lambda_{l_1}-\lambda_i}+\frac{\lambda_{l_1}^{k+2}-\lambda_i^{k+2}}{(\lambda_i-\lambda_{l_1})^2}$$

(iv) if $i\neq l_1$ and $l_1=l_2$,

$$\Gamma(k,\ i)_{ll}=\sum_{j=0}^{k}\lambda_i^j(k-j+1)\lambda_{l_1}^{k-j}$$

$$=(k+1)\lambda_{l_1}^k+k\lambda_i\lambda_{l-1}^{k-1}+(k-1)\lambda_i^2\lambda_{l-1}^{k-2}+\cdots+2\lambda_i^{k-1}\lambda_{l-1}+\lambda_i^k(=S)$$

$$\frac{\lambda_{l_1}}{\lambda_i}S=(k+1)\frac{\lambda_{l_1}^{k+1}}{\lambda_i}+k\lambda_{l-1}^k+(k-1)\lambda_i\lambda_{l-1}^{k-1}+\cdots+2\lambda_i^{k-2}\lambda_{l-1}^2+\lambda_i^{k-1}\lambda_{l_1}$$

$$S-\frac{\lambda_{l_1}}{\lambda_i}S=-(k+1)\frac{\lambda_{l_1}^{k+1}}{\lambda_i}+\sum_{j=0}^{k}\lambda_{l_1}^{k-j}\lambda_i^j=-(k+1)\frac{\lambda_{l_1}^{k+1}}{\lambda_i}+\frac{\lambda_i^{k+1}-\lambda_{l_1}^{k+1}}{\lambda_i-\lambda_{l_1}}$$

$$(\lambda_i-\lambda_{l_1})S=-(k+1)\lambda_{l_1}^{k+1}+\frac{\lambda_i(\lambda_i^{k+1}-\lambda_{l_1}^{k+1})}{\lambda_i-\lambda_{l_1}}$$

$$=-(k+2)\lambda_{l_1}^{k+1}+\lambda_{l_1}^{k+1}+\frac{\lambda_i(\lambda_{l_1}^{k+1}-\lambda_i^{k+1})}{\lambda_i-\lambda_{l_1}}$$

$$=-(k+2)\lambda_{l_1}^{k+1}+\frac{\lambda_i^{k+2}-\lambda_i^{k+2}}{\lambda_i-\lambda_{l_1}}$$

$$\Gamma(k,\ i)_{ll}=-\frac{(k+2)\lambda_{l_1}^{k+1}}{\lambda_i-\lambda_{l_1}}+\frac{\lambda_i^{k+2}-\lambda_{l_1}^{k+2}}{(\lambda_i-\lambda_{l_1})^2}$$

(v) if $i=l_1=l_2$,

$$\Gamma(k,\ i)_{ll}=\sum_{j=0}^{k}\lambda_i^j(k-j+1)\lambda_i^{k-j}$$

$$=\sum_{j=0}^{k}(k+1-j)\lambda_{l_1}^k=\lambda_i^k\sum_{i=1}^{k}i=\frac{(k+1)(k+2)}{2}\lambda_i^k$$

q. e. d.

Let $G=\sum_{k=0}^{\infty}\frac{G(k)}{(k+2)!}$ and it will be partitioned as follows:

$$G=\begin{bmatrix}\sum_{k=0}^{\infty}\frac{G(k,\ 1)}{(k+2)!}\\ \sum_{k=0}^{\infty}\frac{G(k,\ 2)}{(k+2)!}\\ \cdot\\ \cdot\\ \sum_{k=0}^{\infty}\frac{G(k,\ n)}{(k+2)!}\end{bmatrix}=\begin{bmatrix}G_1\\ G_2\\ \cdot\\ \cdot\\ G_n\end{bmatrix}$$

Lemma 5 Let Γ_i be a $n^2\times n^2$ diagonal matrix with $(n(l_1-1)+l_2)$'th diagonal elements defined as:

(i) when $i\neq l_1$, $i\neq l_2$ and $l_1\neq l_2$,

$$\frac{e_{l_1}^{\lambda}}{(\lambda_{l_1}-\lambda_{l_2})(\lambda_{l_1}-\lambda_i)}-\frac{e_{l_2}^{\lambda}}{(\lambda_{l_1}-\lambda_{l_2})(\lambda_{l_2}-\lambda_i)}+\frac{e_i^{\lambda}}{(\lambda_{l_1}-\lambda_i)(\lambda_{l_2}-\lambda_i)}$$

(ii) when $i=l_1$ and $i\neq l_2$,

$$\frac{e^{\lambda_i}}{\lambda_i-\lambda_{l_2}}-\frac{e^{\lambda_{l2}}-e^{\lambda_i}}{(\lambda_i-\lambda_{l_2})^2}$$

(iii) when $i=l_2$ and $i\neq l_1$,

$$-\frac{e^{\lambda_i}}{\lambda_{l_1}-\lambda_i}+\frac{e^{\lambda_{l1}}-e^{\lambda_i}}{(\lambda_i-\lambda_{l_1})^2}$$

(iv) when $i\neq l_1$ and $l_1=l_2$,

$$-\frac{e^{\lambda_{l1}}}{\lambda_i-\lambda_{l_1}}+\frac{e^{\lambda_i}-e^{\lambda_{l1}}}{(\lambda_i-\lambda_{l_1})^2}$$

(v) when $i=l_1=l_2$,

$$\frac{e^{\lambda_i}}{2}$$

The expression for G is given by

$$G = \begin{bmatrix} G_1 \\ G_2 \\ \vdots \\ G_n \end{bmatrix}$$

where

$$G_i = \begin{bmatrix} \underset{(i-1)n^2 \times n^2}{\mathbf{0}} \\ \Gamma_i \\ \underset{(n-i)n^2 \times n^2}{\mathbf{0}} \end{bmatrix}$$

(proof)

$$G_i = \sum_{k=0}^{\infty} \frac{G(k,\ i)}{(k+2)!}$$

and

$$G(k,\ i) = \begin{bmatrix} \underset{(i-1)n^2 \times n^2}{\mathbf{0}} \\ \sum_{j=0}^{k} \lambda_i^j F(k-j) \\ \underset{(n-i)n^2 \times n^2}{\mathbf{0}} \end{bmatrix}$$

By Lemma 4,

$$\sum_{k=0}^{\infty} \frac{1}{(k+2)!} \Gamma(k,\ i) = \Gamma_i$$

$$G_i = \begin{bmatrix} \underset{(i-1)n^2 \times n^2}{\mathbf{0}} \\ \Gamma_i \\ \underset{(n-i)n^2 \times n^2}{\mathbf{0}} \end{bmatrix}$$

The result follows.

q. e. d.

Proposition 2 The Hessian matrix H is given by

$$H = (I_{n^2} \otimes D_n')(I_n \otimes K_{nn} \otimes I_n)(I_{n^4} + K_{n^2,n^2})(Q \otimes Q \otimes Q \otimes Q)\, G\, (Q' \otimes Q') D_n$$

(proof)

From equations (3.2) and (3.4),

$$H = \sum_{m=0}^{\infty} \frac{1}{m!} (I_{n^2} + D_n')(I_n \otimes K_{nn} \otimes I_n) \sum_{j=0}^{m-1} \{K_{n^2,n^2} (Q \otimes Q \otimes Q \otimes Q)$$
$$(vec\Lambda^j \otimes F(m-2-j)(Q' \otimes Q') D_n + (Q \otimes Q \otimes Q \otimes Q)(vec\Lambda^j \otimes F(m-2-j)(Q' \otimes Q') D_n\}$$

By Lemma 3 and 4,

$$H=\sum_{m=0}^{\infty}\frac{1}{m!}(I_{n^2}+D_n')(I_n\otimes K_{nn}\otimes I_n)\{(K_{n^2,n^2}+I_{n^4})(Q\otimes Q\otimes Q\otimes Q)$$

$$\sum_{j=0}^{m-2}(vec\Lambda^j\otimes F(m-2-j))(Q'\otimes Q')D_n\}$$

$$=(I_{n^2}+D_n')(I_n\otimes K_{nn}\otimes I_n)\{(K_{n^2,n^2}+I_{n^4})(Q\otimes Q\otimes Q\otimes Q)\sum_{m=0}^{\infty}\frac{1}{m!}G(m-2)(Q'\otimes Q')D_n$$

Thus,

$$H=(I_{n^2}\otimes D_n')(I_n\otimes K_{nn}\otimes I_n)(I_{n^4}+K_{n^2,n^2})(Q\otimes Q\otimes Q\otimes Q)G(Q'\otimes Q')D_n$$

q. e. d.

V. Concluding Remarks

The matrix exponential of a real symmetric matrix is positive definite. The use of the matrix exponential function ensures positive definite variance-covariance without any constraint. The present paper derives the first and second derivatives of the matrix exponetial which are expected to provide better numerical accuracy than numerical drivatives in nonlinear optimization problems.

In modeling and estimating time-varying conditional variance-covariance, the matrix exponential may be used to parameterize several multivariate GARCH models. For example, conditional variance-covariance matrices of VECH(1,1) models are defined as

$$H_t=\Omega+A\circ H_{t-1}+B\circ\varepsilon_{t-1}\varepsilon_{t-1}'$$

where ε_{t-1} is a $n\times 1$ vector and H_t and H_{t-1} are conditional variance-covariance matrices, so should be positive definite ($\circ$ represents the Hadamard product).

The Schur product theorem says that if A, B are positive semi-definite matrices, then $A\circ B$ is also positive semi-definite. Thus, the above conditional variance covariance matrices can be reparameterized as follows:

$$H_t=\exp(C)+\exp(F)\circ H_{t-1}+\exp(G)\circ\varepsilon_{t-1}\varepsilon_{t-1}'$$

In order to obtain numerically sensible maximum lilelihood estimates, analytical derivatives of likelihood functions rather than numerical ones should be utilized. The application of results in the present paper to model and parameterize multivariate GARCH will be a task in the future.

References

[1] Chiu, Yiu-Ming (1994), *Exponential Covariance Model*, (unpublished dissertation, University of Wisconsin-Madison).

[2] Chen, Baoline and Peter A. Zadrozny (2001), "Analytic Derivatives of the Matrix Exponential for Estimation of Linear Continuous-time Models", *Journal of Economic Dynamics and Control*, 25, 1867

-1879.

[3] Linton. Oliver (1993), "An Alternative Exponential GARCH Model with Multivariate Extension", (mimeo., Nuffield College, Oxford).

[4] Linton, Oliver (1995), "Differentiation of an Exponential Matrix Function—Solutions", *Econometric Theory*, 11, 1182-1183.

[5] Lütkepohl, Helmut (1996), *Handbook of Matrices*, Wiley, New York, U. S. A.

[6] Magnus, J. R. and H. Neudecker (1999), *Matrix Differential Calculus with Applications in Statistic and Econometrics*, Revised edition, Wiley, Chichster, U. K.

[7] Pinheiro, Jose C. and Bates D. M. (1996), "Unconstrained Parameterizations for Variance-Covariance Matrices", *Statistics and Computing*, 6, 289-296.

[8] Vinod, H. D. (2000), "Review of GAUSS for Windows, Incluing its Numerical Accuracy", *Journal of Applied Econometrics*, 15, 211-220.