MATEMATIKA, 2007, Volume 23, Number 1, 1–8

A Matrix Variance Inequality for k-Functions

[1]**Norhayati Rosli** & [2]**Wan Muhamad Amir W Ahmad**
[1]Fakulti Kejuruteraan Kimia & Sumber Asli, Kolej Universiti Kejuruteraan & Teknologi Malaysia
25000 Kuantan, Pahang, Malaysia
[2]Departments of Mathematics, Faculty of Science and Technology
Kolej Universiti Sains & Teknologi Malaysia
Mengabang Telipot, 21030 Kuala Terengganu, Terengganu, Malaysia
e-mail: [1]norhayati@kuktem.edu.my

Abstract In this paper a course of solving variational problem is considered. [2] obtained what appears to be specialized inequality for a variance, namely, that for a standard normal variable X , $Var[g(x)] \geq E[g'(x)]^2$. However both of the simplicity and usefulness of the inequality has generated a plethora of extensions, as well as alternative proofs. [5] had focused on a result of two random variables for the normal and gamma distribution. They obtained the result of normal distribution with k functions, without proving and the proof is presented here. This paper also extend the result obtained by [5] to the k functions for the gamma distribution.

Keywords Normal Distribution, Gamma Distribution, Laguerre Family, Hermite Polynomials

1 Introduction

In solving a variational problem, [2] obtained what appears to be specialized inequality for a variance. Let X be normally distributed with density $\varphi(x)$ and mean 0 and variance 1. If g is absolutely continuous and $g(X)$ has finite variance, then

$$E[g'(X)]^2 \geq Var[g(X)]. \tag{1}$$

Equality in (1) is achieved for linear functions. This inequality have arisen earlier, especially because of its use in variational problems. There are many papers that deal with inequality (1) and in many cases they relate to the single function. However, the random variables might have multivariate distributions. So, we present the study of matrix variance inequality for the normal and gamma distribution with k-functions.

2 Literature Review

Chernoff's proof is based on expanding $g(X)$ in orthonormalized Hermite polynomials with respect to the normal density

$$g(X) = a_0 + a_1 H_1(X) + a_2 H_2(X) + \cdots \tag{2}$$

with probability 1. Let

$$E[H_i(X)] = 0, \quad E[H_i(X)H_j(X)] = \delta_{ij}, \tag{3}$$

$$\frac{dH_i(x)}{dx}\sqrt{i}H_{i-1}(x)$$

$$a_i = E[g(X)H_i(X)] \tag{4}$$

and

$$Var(g(X)) = a_1^2 + a_2^2 + \cdots + a_n^2 + \cdots , \tag{5}$$

$$g'(X) = a_1 + \sqrt{2}a_2H_1(X) + \cdots + \sqrt{n}a_nH_{n-1}(X) + R_n'(X).$$

So that if $g'(X)$ has a second moment,

$$E[g'(X)]^2 = \sum ia_i^2 \geq Var[g(X)]. \tag{6}$$

And if $g'(X)$ has no second moment then $\sum ia_i^2$ is infinite. For a logconcave density $\exp[-\varphi(x)]$, [4] proved that

$$Var[g(X)] \leq E[g'(X)/\varphi''(X)]^2, \tag{7}$$

and for the normal density, (7) is reduced to (1).

[1] shows that

$$E[g'(X)] = E[Xg(X)] \tag{8}$$

has a similar flavor to that of (1). Stein's proof is essentially based on integration by parts, but can also be proven by using Hermite polynomials. [6] and [7] provide an alternative proof based on the Cauchy-Schwarz inequality. [6] prove that for the normal density $\varphi(x)$, $\varphi'(x) = -x\varphi(x)$. [7] extends (1) to the case that $X_1, \ldots, X_k$ are independent $N(0, 1)$ random variables and g is defined on R^k. Then

$$Var[g(X)] \leq E[g_1(X)]^2 + \cdots + E[g_k(X)]^2,$$

where $g_i(x) = \delta g(x)/\delta x_i$ and $X = (X_1, \ldots, X_k)$.

[10] provides other extensions, and in particular, the lower bounds in (1);

$$[Eg'(X)] \leq Var[g(X)] \leq E[g'(X)]^2,$$

$$[Eg'(X)]^2 + \frac{1}{2}[Eg(X)]^2 \leq Var[g(X)].$$

[11] improved the bounds for families other than normal. They also obtain the lower bound for the normal distribution

$$\sum_{k=1}^{n} \frac{\left|EG^{(k)}(X)\right|^2}{k!} \leq Var[G(X)],$$

with $n = 2$.

[4] discussed a generalization to operators and use a complete orthonormal system. [8] obtain an equality similar to (1) by consider the double exponential distribution with density $\exp(-|x|)/2$.

3 Characterizations

In this paper we provide a proof of matrix variance inequality for the normal and gamma distributions of k-functions. Actually the result of matrix variance inequality for the normal distribution is obtained by [5] without proving. We extend this paper to find the matrix variance inequality for the gamma distribution with k-functions. We used the proposition obtained in [2]. This proposition later studied by [5] and shows the proof of matrix variance inequality of normal and gamma distribution for 2-functions.

Proposition. Let X be a $N(0,1)$ random variable, $g_1,\ldots,g_k$ absolutely continuous functions with finite variances. Let $H=(h_{ij})$ and $C=(c_{ij})$ be $k\times k$ matrices defined by

$$\begin{aligned} h_{ij} &= E[g_i'(X)g_j'(X)], \\ c_{ij} &= Cov[g_i(X),g_j(X)]. \end{aligned} \tag{9}$$

Then $H\geq C$ in the Loewner ordering, i.e, $H-C$ is nonnegative definite.

The proof of this proposition is discussed in Section 4.

4 A Matrix Variance Inequality for The Normal Distribution

We show the proof of a matrix variance inequality for k-functions which are normal distribution. To prove the proposition, we expand $g_1(X),g_2(X),\ldots,g_k(X)$ in orthonormalized Hermite polynomials:

$$\begin{aligned} g_1(X) &= a_0 + a_1H_1(X) + a_2H_2(X) + \cdots \\ g_2(X) &= b_0 + b_1H_1(X) + b_2H_2(X) + \cdots \\ \vdots \quad & \quad \vdots \quad \vdots \qquad \vdots \qquad\qquad \vdots \\ g_k(X) &= u_0 + u_1H_1(X) + u_2H_2(X) + \cdots \end{aligned} \tag{10}$$

with probability 1, where

$$E[\mathbf{H}_i(X)]=0, \quad E[\mathbf{H}_i(X)\mathbf{H}_j(X)]=\delta_{ij}, \tag{11}$$

$$\frac{d\mathbf{H}_i(x)}{dx}\sqrt{i}\mathbf{H}_{i-1}(x)$$

$$\begin{aligned} a_i &= E[g_1(X)\mathbf{H}_i(X)] \\ b_i &= E[g_2(X)\mathbf{H}_i(X)] \\ \vdots \quad & \vdots \qquad \vdots \\ u_i &= E[g_k(X)\mathbf{H}_i(X)]. \end{aligned}$$

Then, from (5) we have

$$Var[g_1(X)] = \sum_{i=1}^{\infty} a_i^2,\ Var[g_2(X)] = \sum_{i=1}^{\infty} b_i^2, \ldots,\ Var[g_k(X)] = \sum_{i=1}^{\infty} u_i^2 \tag{12}$$

$$Cov[g_1(X), g_2(X)] = \sum_{i=1}^{\infty} a_i b_i, ...,\ Cov[g_1(X), g_k(X)] = \sum_{i=1}^{\infty} a_i u_i, \tag{13}$$

$$Cov[g_2(X), g_k(X)] = \sum_{i=1}^{\infty} b_i u_i. \tag{14}$$

Hence, $\mathbf{H}$ and $\mathbf{C}$ are $k \times k$ matrices in the form

$$\mathbf{H} = \begin{pmatrix} \sum i a_i^2 & \sum i a_i b_i & \ldots & \sum i a_i u_i \\ \sum i a_i b_i & \sum i b_i^2 & \ldots & \sum i b_i u_i \\ \vdots & \vdots & \ddots & \vdots \\ \sum i a_i u_i & \sum i b_i u_i & \ldots & \sum i u_i^2 \end{pmatrix} \text{ and} \tag{15}$$

$$\mathbf{C} = \begin{pmatrix} \sum a_i^2 & \sum a_i b_i & \ldots & \sum a_i u_i \\ \sum a_i b_i & \sum b_i^2 & \ldots & \sum b_i u_i \\ \vdots & \vdots & \ddots & \vdots \\ \sum a_i u_i & \sum b_i u_i & \ldots & \sum u_i^2 \end{pmatrix}, \text{ respectively.} \tag{16}$$

Then, we get

$$\mathbf{H} - \mathbf{C} = \begin{pmatrix} \sum (i-1) a_i^2 & \sum (i-1) a_i b_i & \ldots & \sum (i-1) a_i u_i \\ \sum (i-1) a_i b_i & \sum (i-1) b_i^2 & \ldots & \sum (i-1) b_i u_i \\ \vdots & \vdots & \ddots & \vdots \\ \sum (i-1) a_i u_i & \sum (i-1) b_i u_i & \ldots & \sum (i-1) u_i^2 \end{pmatrix}. \tag{17}$$

Let

$$\alpha_i = \sqrt{i-1}a_i, \quad \beta_i = \sqrt{i-1}b_i, ..., \quad \gamma_i = \sqrt{i-1}u_i$$

where

$$\tau_{(1)} = (\alpha_1, \alpha_2, \alpha_3, \ldots), \quad \tau_{(2)} = (\beta_1, \beta_2, \beta_3, \ldots), ..., \text{ and } \tau_{(k)} = (\gamma_1, \gamma_2, \gamma_3, \ldots). \tag{18}$$

Actually (17) is in the form of

$$\mathbf{H} - \mathbf{C} = \begin{pmatrix} \sum \sqrt{i-1}a_i\sqrt{i-1}a_i & \sum \sqrt{i-1}a_i\sqrt{i-1}b_i & \ldots & \sum \sqrt{i-1}a_i\sqrt{i-1}u_i \\ \sum \sqrt{i-1}a_i\sqrt{i-1}b_i & \sum \sqrt{i-1}b_i\sqrt{i-1}b_i & \ldots & \sum \sqrt{i-1}b_i\sqrt{i-1}u_i \\ \vdots & \vdots & \ddots & \vdots \\ \sum \sqrt{i-1}a_i\sqrt{i-1}u_i & \sum \sqrt{i-1}b_i\sqrt{i-1}u_i & \ldots & \sum \sqrt{i-1}u_i\sqrt{i-1}u_i \end{pmatrix}$$

It is obvious that if $\alpha_i = \sqrt{i-1}a_i$, then $\alpha_i' = \sqrt{i-1}a_i$. Consequently we can rewrite (17) as

$$\mathbf{H}-\mathbf{C} = \begin{pmatrix} \sum\alpha_i\alpha_i' & \sum\alpha_i\beta_i' & \dots & \sum\alpha_i\gamma_i' \\ \sum\beta_i\alpha_i' & \sum\beta_i\beta_i' & \dots & \sum\beta_i\gamma_i' \\ \vdots & \vdots & \ddots & \vdots \\ \sum\gamma_i\alpha_i' & \sum\gamma_i\beta_i' & \dots & \sum\gamma_i\gamma_i' \\ \end{pmatrix}.$$

Hence, we obtained

$$\mathbf{H}-\mathbf{C} = \begin{pmatrix} \tau_{(1)}\tau_{(1)}' & \tau_{(1)}\tau_{(2)}' & \cdots & \tau_{(1)}\tau_{(k)}' \\ \tau_{(2)}\tau_{(1)}' & \tau_{(2)}\tau_{(2)}' & \cdots & \tau_{(2)}\tau_{(k)}' \\ \vdots & \vdots & \ddots & \vdots \\ \tau_{(k)}\tau_{(k)}' & \tau_{(k)}\tau_{(2)}' & \cdots & \tau_{(k)}\tau_{(k)}' \end{pmatrix} = \begin{pmatrix} \tau_{(1)} \\ \tau_{(2)} \\ \vdots \\ \tau_{(k)} \end{pmatrix} \begin{pmatrix} \tau_{(1)}' & \tau_{(2)}' & \cdots & \tau_{(k)}' \end{pmatrix} \geq 0. \tag{19}$$

5 A Matrix Inequality for Gamma Distribution

The gamma density function is defined by

$$g(X) = \frac{x^\alpha e^{-x}}{\Gamma(\alpha+1)}, \quad \alpha > -1.$$

[9] uses the Laguerre family of orthogonal family to obtain the inequality

$$Var[g(X)] \leq EX[g'(X)]^2, \tag{20}$$

with equality if and only if $g(x)$ is linear. The key features of the Laguerre family are

$$E[L_n^{(\alpha)}(X)L_n^{(\alpha)}(X)] = \begin{pmatrix} n+\alpha \\ n \end{pmatrix}\delta_{ij}, \quad \frac{dL_n^{(\alpha)}}{dx} = -L_{n-1}^{(\alpha+1)}(x). \tag{21}$$

Let say

$$\begin{aligned} g_1(x) &= \sum a_n L_n^{(\alpha)}(x), \\ g_2(x) &= \sum b_n L_n^{(\alpha)}(x), \\ g_3(x) &= \sum c_n L_n^{(\alpha)}(x), \\ \vdots \quad & \vdots \quad \vdots \\ g_k(x) &= \sum u_k L_k^{(\alpha)}(x). \end{aligned} \tag{22}$$

Then

$$
\begin{aligned}
&v_{11} \equiv Var[g_1(X)] = \sum a_n^2 \begin{pmatrix} n+\alpha \\ n \end{pmatrix}, \\
&v_{22} \equiv Var[g_2(X)] = \sum b_n^2 \begin{pmatrix} n+\alpha \\ n \end{pmatrix}, \\
&v_{33} \equiv Var[g_3(X)] = \sum c_n^2 \begin{pmatrix} n+\alpha \\ n \end{pmatrix}, \\
&v_{12} \equiv Cov[g_1(X), g_2(X)] = \sum a_n b_n \begin{pmatrix} n+\alpha \\ n \end{pmatrix}, \\
&v_{13} \equiv Cov[g_1(X), g_3(X)] = \sum a_n c_n \begin{pmatrix} n+\alpha \\ n \end{pmatrix}, \\
&v_{23} \equiv Cov[g_2(X), g_3(X)] = \sum b_n c_n \begin{pmatrix} n+\alpha \\ n \end{pmatrix}, \\
&\vdots \\
&v_{kk} \equiv Var[g_{kn}(X)] = \sum u_n^2 \begin{pmatrix} n+\alpha \\ n \end{pmatrix}, \\
&v_{1k} \equiv Cov[g_1(X), g_k(X)] = \sum a_n u_n \begin{pmatrix} n+\alpha \\ n \end{pmatrix}, \\
&v_{2k} \equiv Cov[g_2(X), g_k(X)] = \sum b_n u_n \begin{pmatrix} n+\alpha \\ n \end{pmatrix}, \\
&v_{3k} \equiv Cov[g_3(X), g_k(X)] = \sum c_n u_n \begin{pmatrix} n+\alpha \\ n \end{pmatrix}
\end{aligned} \tag{23}
$$

and following development in [9], we let

$$
\begin{aligned}
&c_{11} = EX[g_1'(X)]^2 = \sum a_n^2 n \begin{pmatrix} n+\alpha \\ n \end{pmatrix}, \\
&c_{22} = EX[g_2'(X)]^2 = \sum b_n^2 n \begin{pmatrix} n+\alpha \\ n \end{pmatrix}, \\
&c_{33} = EX[g_3'(X)]^2 = \sum c_n^2 n \begin{pmatrix} n+\alpha \\ n \end{pmatrix}, \\
&c_{12} = EX[g_1'(X) g_2'(X)]^2 = \sum a_n b_n n \begin{pmatrix} n+\alpha \\ n \end{pmatrix}, \\
&c_{13} = EX[g_1'(X) g_3'(X)]^2 = \sum a_n c_n n \begin{pmatrix} n+\alpha \\ n \end{pmatrix}, \\
&\vdots \\
&c_{kk} = EX[g_k'(X)]^2 = \sum u_n^2 n \begin{pmatrix} n+\alpha \\ n \end{pmatrix}, \\
&c_{1k} = EX[g_1'(X) g_k'(X)]^2 = \sum a_n u_n n \begin{pmatrix} n+\alpha \\ n \end{pmatrix}, \\
&c_{2k} = EX[g_2'(X) g_k'(X)]^2 = \sum b_n u_n n \begin{pmatrix} n+\alpha \\ n \end{pmatrix}, \\
&c_{3k} = EX[g_3'(X) g_k'(X)]^2 = \sum c_n u_n n \begin{pmatrix} n+\alpha \\ n \end{pmatrix}.
\end{aligned} \tag{24}
$$

If $\mathbf{V} = (v_{ij})$, $\mathbf{C} = (c_{ij})$, then we have

$$\mathbf{V} = \begin{pmatrix} \sum a_n^2 \binom{n+\alpha}{n} & \sum a_n b_n \binom{n+\alpha}{n} & \dots & \sum a_n u_n \binom{n+\alpha}{n} \\ \sum a_n b_n \binom{n+\alpha}{n} & \sum b_n^2 \binom{n+\alpha}{n} & \dots & \sum b_n u_n \binom{n+\alpha}{n} \\ \vdots & \vdots & \ddots & \vdots \\ \sum a_n b_n \binom{n+\alpha}{n} & \sum a_n b_n \binom{n+\alpha}{n} & \dots & \sum u_n^2 \binom{n+\alpha}{n} \end{pmatrix}$$

$$\mathbf{C} = \begin{pmatrix} \sum a_n^2 n \binom{n+\alpha}{n} & \sum a_n b_n n \binom{n+\alpha}{n} & \dots & \sum a_n u_n n \binom{n+\alpha}{n} \\ \sum a_n b_n n \binom{n+\alpha}{n} & \sum b_n^2 n \binom{n+\alpha}{n} & \dots & \sum b_n u_n n \binom{n+\alpha}{n} \\ \vdots & \vdots & \ddots & \vdots \\ \sum a_n u_n n \binom{n+\alpha}{n} & \sum a_n u_n n \binom{n+\alpha}{n} & \dots & \sum u_n^2 n \binom{n+\alpha}{n} \end{pmatrix}.$$

So, it is obvious,

$$\mathbf{C} - \mathbf{V} =$$
$$\begin{pmatrix} \sum a_n^2 \binom{n+\alpha}{n} (n-1) & \sum a_n b_n \binom{n+\alpha}{n} (n-1) & \dots & \sum a_n u_n \binom{n+\alpha}{n} (n-1) \\ \sum a_n b_n \binom{n+\alpha}{n} (n-1) & \sum b_n^2 \binom{n+\alpha}{n} (n-1) & \dots & \sum b_n u_n \binom{n+\alpha}{n} (n-1) \\ \vdots & \vdots & \ddots & \vdots \\ \sum a_n u_n \binom{n+\alpha}{n} (n-1) & \sum b_n u_n \binom{n+\alpha}{n} (n-1) & \dots & \sum u_n^2 \binom{n+\alpha}{n} (n-1) \end{pmatrix}$$

We let

$$\alpha_n = a_n \sqrt{n-1} \sqrt{\binom{n+\alpha}{n}}, \quad \beta_n = b_n \sqrt{n-1} \sqrt{\binom{n+\alpha}{n}}, ...,$$
$$\gamma_n = u_n \sqrt{n-1} \sqrt{\binom{n+\alpha}{n}},$$

where

$$\tau_{(1)} = (\alpha_1, \alpha_2, \alpha_3, \dots), \quad \tau_{(2)} = (\beta_1, \beta_2, \beta_3, \dots), ..., \text{ and } \tau_{(k)} = (\gamma_1, \gamma_2, \gamma_3, \dots). \tag{25}$$

Hence,

$$\mathbf{C} - \mathbf{V} = \begin{pmatrix} \tau_{(1)} \\ \tau_{(2)} \\ \vdots \\ \tau_{(k)} \end{pmatrix} \begin{pmatrix} \tau'_{(1)} & \tau'_{(2)} & \dots & \tau'_{(k)} \end{pmatrix}$$

which is nonnegative definite.

6 Conclusion

In this paper we have shown how to proof the matrix variance inequality for k-functions which are normally distributed. We also addressed the case of proving for k-functions which are gamma distribution. Hence, we can conclude that if X is either normally distributed or gamma distribution, then $H \geq C$ in the Loewner ordering which is $H - C$ is nonnegative definite.

References

[1] C. Stein, *Estimation of the mean of a multivariate normal distribution*, Ann. Statist. 9 (1981), 1135-1151.

[2] H. Chernoff, *A note on an inequality involving the normal distribution*, Ann. Probab. 9 (1981), 533-535.

[3] H.J. Branscamp & E.H. Lieb, *On extensions of the Brunn-Minkowski and Prekopa-Leinder theorems, including inequalities for log concave functions and with an application to the diffusion equation,* J. Funct. Anal. 22 (1976), 366-389.

[4] Hwang, Chii-Ruey, Sheu, Shueen-Jyi, *A generalization of Chernoff-inequality via stochastic analysis,* Probab. Theory Related Fields 75 (1987), 149-157.

[5] I. Olkin & L. Shepp, *A matrix variance inequality,* J. Statist. Plan. Inference 130 (2005), 351-358.

[6] L.H.Y. Chen, *An inequality for the multivariate normal distribution,* J. Multivariate Anal. 12 (1982), 306-315.

[7] L.H.Y. Chen, *The central limit theorem and Poincare-type inequalities,* Ann. Probab. 16 (1988), 300-304.

[8] S. Borovkov & C. Houdre, *Converse Poincare-type inequalities for convex functions,* Satistic Probab. Lett. 42 (1997), 34-37.

[9] S. Karlin, *A general class of variance inequalities,* Multivariate Analysis: Future Directions. Elsevier Science Publishers, New York, 1993.

[10] T. Cacoullos, *On upper and lower bounds for the variance of a function of a random variable,* Ann. Probab. 10 (1982), 799-809.

[11] T. Cacoullos & V. Papathanasiou, *Characterizations of distributions by generalizations of variance bounds and simple proofs of the C. L. T. J.,* Statist. Plann. Inference 63 (1997), 157-171.