On Bayesian Inference for Generalized Multivariate Gamma Distribution

Sourish Das[a,b], Dipak K Dey[c]

[a] *Department of Statistical Science, Duke University, Durham, NC*
[b] *(SAMSI) Statistical and Applied Mathematical Science Institute, Durham, NC*
[c] *Department of Statistics, University of Connecticut, Storrs, CT*

Abstract

In this paper we define a generalized multivariate gamma (MG) distribution and develop various properties of this distribution. Then we consider a Bayesian decision theoretic approach to develop the inference technique for the related scale matrix Σ. We show that maximum posteriori (MAP) estimate is a Bayes estimator. We also develop the testing problem for Σ using Bayes factor. This approach provides a mathematically closed form solution for Σ. Only other approach to Bayesian inference for MG distribution is given on Tsionas (2004), which is based on Markov Chain Monte Carlo (MCMC) technique. Tsionas (2004) technique involves costly matrix inversion whose computational complexity increases in cubic order, hence make inference infeasible for Σ, for large dimension. In this paper, we provide an elegant closed form Bayes factor for Σ.

Key words: Autoregressive structure; Bayes estimator; Dispersion matrix; MAP estimate; Multivariate beta distribution.

---

1. Introduction

Over the last fifty years, many authors worked on the multivariate gamma (MG) distribution. Krishnamoorty and Parthasarathy (1951) and Krishnaiah and Rao (1961) discussed at length the issues related to the MG distribution. Lukacs and Laha (1964) discussed some closed form expressions for the MG type function with shape parameter α and the scale matrix Σ, when α is a positive integer. Griffiths (1984) and later Bapat (1989), proved the infinite divisibility property of the MG distribution. Mathai and Moschopoulos (1991, 1992) proposed a new form of MG distribution when the components are positively correlated. They further developed multivariate extension of three parameters univariate gamma distribution and obtain the explicit forms for the moments, moment generating function and conditional moments. Coelho (1998) presented the exact distribution of generalized Wilk's Λ statistics, where the basis distribution is generalized integer gamma distribution. Eaton (2000) presented the vector space approach to multivariate statistics and discussed Wishart distribution which is a special case of MG distribution (i.e., equation (1)) presented in section 2. Kotz et al. (2000) presented different versions of MG distribution. Recently, Furman (2008) presented ladder-type MG distribution for modeling dependent insurance losses. In this paper, we develop a Bayesian inference for MG presented in equation (1). This version takes care of wide range of covariance matrices, such as compound symmetry which has a randomization justification such as 'split-plot design'.

In this paper we further explore various properties of the MG distribution and develop analytical expressions for Bayesian decision theoretic inference for the unknown scale matrix. In Section 2, we discuss a closed form expression of the MG distribution and the definition of the generalized MG distribution. In Section 3, we discuss main properties of the generalized MG distribution including the additive property and its relationship with generalized multivariate beta distribution. Then we discuss the inverted multivariate gamma (I-MG) distribution and derive its mode. In Section 4, we discuss the posterior distribution of scale matrix and the issues for the related Bayesian inference technique. Section 5, concludes the paper with a brief discussion.

2. Generalized Multivariate Gamma Distribution

Let $\Sigma = ((\sigma_{ij}))$ be a real, symmetric positive definite matrix of order p and $\boldsymbol{Z} = ((z_{ij}))$ be a real symmetric positive definite matrix of the same order in $\frac{p(p+1)}{2}$ variables z_{ij}. Then consider the function

$$f(\boldsymbol{Z}) = K.\exp\Big\{ -\frac{1}{\beta} tr\Sigma^{-1}\boldsymbol{Z}\Big\} \mid Z \mid^{\lambda-1}, \quad \boldsymbol{Z} > 0,$$

where $Z > 0$ indicates that Z is positive definite matrix, $\mid \cdot \mid$ denote determinant, tr denote trace of a matrix, λ and β are positive real numbers and K is normalizing constant. Lukacs and Laha (1964), showed that the normalizing constant is given as

$$K = \frac{\mid \Sigma^{-1} \mid^{\lambda + \frac{p-1}{2}}}{\beta^{\lambda p + \frac{p(p-1)}{2}} \pi^{\frac{p(p-1)}{4}} \prod_{j=1}^{p} \Gamma\Big(\lambda + \frac{p-j}{2}\Big)}.$$

If we choose $\alpha = \lambda + \frac{p-1}{2}$, then

$$K = \frac{\mid \Sigma^{-1} \mid^{\alpha}}{\beta^{p\alpha} \pi^{\frac{p(p-1)}{4}} \prod_{j=1}^{p} \Gamma\big(\alpha - \frac{1}{2}(j-1)\big)}.$$

Defining the multivariate gamma function (Anderson, 1984)

$$\Gamma_p(\alpha) = \pi^{\frac{p(p-1)}{4}} \prod_{j=1}^{p} \Gamma\Big(\alpha - \frac{1}{2}(j-1)\Big),$$

we obtain the general form of the pdf of a generalized multivariate gamma distribution as

$$f(\boldsymbol{Z}) = \frac{\mid \Sigma^{-1} \mid^{\alpha} \ \mid \boldsymbol{Z} \mid^{\alpha - \frac{1}{2}(p+1)}}{\beta^{p\alpha} \Gamma_p(\alpha)} \ \exp\Big\{ -\frac{1}{\beta} tr \Sigma^{-1} \boldsymbol{Z} \Big\} \quad \boldsymbol{Z} > 0 \ , \tag{1}$$

where $\alpha > \frac{(p-1)}{2}$, since $\lambda > 0$, $\beta > 0$, and Σ^{-1} is a positive definite matrix of full rank. We denote this as $\boldsymbol{Z} \sim MG_p(\alpha, \beta, \Sigma)$. This distribution generalizes many other distributions. If we choose $\alpha = \frac{n}{2}$ and $\beta = 2$, then $\boldsymbol{Z}$ has a Wishart distribution (Anderson, 1984, pp. 252) with pdf

$$f(\boldsymbol{Z}) = \frac{\mid \boldsymbol{Z} \mid^{\frac{1}{2}(n-p-1)}}{\Gamma_p(\frac{n}{2}) \mid \Sigma \mid^{\frac{n}{2}} 2^{\frac{pn}{2}}} \ \exp\Big\{ -\frac{1}{2} tr \Sigma^{-1} \boldsymbol{Z} \Big\} \ ,$$

where $\boldsymbol{Z} > 0$. Note that n is a positive integer, generally considered as the sample size. If we choose $\Sigma = I$ and $p = 1$ then $\boldsymbol{Z}$ follows a chi-square distribution with n degrees of freedom. Finally if we choose $p = 1$ and $\beta = 1$ in (1), then $\boldsymbol{Z}$ follows a univariate $Gamma(\alpha, \sigma)$ distribution.

The Laplace transform of (1) is derived from Lukacs and Laha (1964) as

$$\phi_Z(\boldsymbol{t}) = E\Big[\exp\Big\{ -tr(T\boldsymbol{Z})\Big\}\Big] = \mid I_p + \beta \Sigma T \mid^{-\alpha} . \tag{2}$$

We observe immediately that the characteristic function of (1) as

$$\psi_z(\boldsymbol{t}) = E\Big[i \exp\Big\{tr(T\boldsymbol{Z})\Big\}\Big] = \mid I_p - i\beta\Sigma T \mid^{-\alpha},$$

where T is a symmetric matrix of order p. We must note that before Griffith's paper (Griffiths, 1984) and the follow-up paper of Bapat (Bapat, 1989), all the literature for MG was developed when Z is real, symmetric matrix of order p and α is an integer. Hence, Bapat's condition is required to define a MG distribution when $\alpha > 0$. Consequently, we can develop the definition of multivariate gamma distribution for real symmetric matrix of order p as follows.

Definition 1. Let Z be a real symmetric matrix of order p with $\frac{p(p+1)}{2}$ variables z_{ij} and $\Sigma = ((\sigma_{ij}))$ be the corresponding dispersion matrix, such that for a diagonal matrix D with diagonal elements 1 or -1, $(D\Sigma D)^{-1}$ has non - positive off diagonal elements. Hence due to the Bapat's condition (Bapat, 1989), Z with Laplace transform (2) having the density function as

$$f(Z) = \frac{\mid \Sigma \mid^{\alpha}}{\Gamma_p(\alpha)\beta^{\alpha p}} \exp\Big\{ - \frac{1}{\beta} tr\Sigma^{-1}Z \Big\} \mid Z \mid^{\alpha - \frac{1}{2}(p+1)}, Z > 0 \tag{3}$$

has an infinitely divisible multivariate gamma distribution with parameters $\alpha \geq \frac{p-1}{2}$, $\beta \geq 0$, and Σ a positive definite matrix. We denote that as $Z \sim MG_p(\alpha, \beta, \Sigma)$. Note that if $0 \leq \alpha < \frac{p-1}{2}$ then Z has a degenerate distribution.

2.1. Some Patterned Dispersion Matrix

Here we discuss some popular patterned dispersion matrices which satisfy Bapat's condition. Thus validity of the multivariate gamma distributional assumptions can be justified through the appropriate covariance matrix structure. First we consider for the dispersion matrix which has compound symmetry structure.

2.1.1. Compound Symmetry

One of the first covariance pattern models used for the analysis of repeated measures data is compound symmetry. With a cmpound symmetry it is assumed that the variance is constant across occasions, say σ^2, and $Corr(Y_{ij}, Y_{ik}) = \rho$ for all j and k. That is

$$\Sigma = \sigma^2\Big((1-\rho)I + \rho J\Big)$$

with the constraint that $\rho \geq 0$, I is identity matrix and J is matrix of ones.

The compound symmetry covariance has a randomization justification in certain repeated measure designs (e.g., split-plot design). In an experiment where the within-subject factor is randomly allocated to subjects, randomization arguments can be made to show that the constant variance and constant correlation conditions hold. However, the randomization argument is not justifiable in the longitudinal data setting since the measurement occasions cannot be randomly allocated to subjects.

If we choose $D = I$, then by the definition 1, all the off-diagonal elements of Σ are non-negative. Again all the off diagonal elements of Σ^{-1} are non-positive, where Σ^{-1} is given as

$$\Sigma^{-1} = \frac{1}{\sigma^2(1-\rho)}\Big(I - \frac{\rho}{1+(p-1)}J\Big).$$

Therefore dispersion matrix with compound symmetry structure, which corresponds to split-plot design, satisfies Bapat's condition of the definition 1. Hence we can argue the underlying distribution of the data follows the infinitely divisible multivariate gamma distribution. Next we discuss for the dispersion matrix which has autoregressive structure of first order.

2.1.2. Autoregressive covariance pattern model with heterogeneous variances

The Σ of the heterogeneous variances autoregressive (HVAR) model is given by

$$\Sigma = ((\sigma_{ij})) = \Big((1-\rho^{|i-j|})\sigma_i\sigma_j I + \rho^{|i-j|}\sigma_i\sigma_j J\Big),$$

which has $p+1$ parameters. All the off-diagonal element of Σ are non-negative. Again all the off diagonal elements of

$$\Sigma^{-1} = \begin{pmatrix} \frac{1}{\sigma_1^2(1-\rho^2)} & \frac{-\rho}{\sigma_1\sigma_2(1-\rho^2)} & 0 & \cdots & 0 \\ \frac{-\rho}{\sigma_1\sigma_2(1-\rho^2)} & \frac{1}{\sigma_2^2}\left(\frac{1+\rho^2}{1-\rho^2}\right) & \frac{-\rho}{\sigma_2\sigma_3(1-\rho^2)} & \cdots & 0 \\ 0 & \frac{-\rho}{\sigma_2\sigma_3(1-\rho^2)} & \frac{1}{\sigma_2^2}\left(\frac{1+\rho^2}{1-\rho^2}\right) & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \frac{1}{\sigma_p^2(1-\rho^2)} \end{pmatrix}$$

are non-positive. Therefore, we can argue that the underlying distribution is the infinite divisible multivariate gamma distribution. One nice feature of HVAR is that autoregressive model with lag 1, i.e., AR(1) is a special case of HVAR. In the next section, we discuss some important properties of the generalized multivariate gamma distribution.

3. Properties of Generalized Multivariate Gamma Distribution

We derive some properties of multivariate gamma distribution, like additive property of multivariate gamma distribution, marginal distribution and some other properties.

Lemma 1. *If Z_1 follows $MG_p(\alpha_1, \beta, \Sigma)$, Z_2 follows $MG_p(\alpha_2, \beta, \Sigma)$ and Z_1, Z_2 are independent of each other. Then $Z_1 + Z_2$ follows $MG_p(\alpha_1 + \alpha_2, \beta, \Sigma)$.*

Proof. Let $Y = Z_1 + Z_2$ and Y has the Laplace transform as

$$\phi_Y(\boldsymbol{t}) = E\Big[\exp\Big\{-tr(TY)\Big\}\Big] = E\Big[\exp\big\{-tr\,T(Z_1 + Z_2)\big\}\Big]$$

which is equivalent to

$$\phi_Y(\boldsymbol{t}) = E\Big[\exp\big\{-tr(TZ_1)\big\}\Big]E\Big[\exp\big\{-tr(TZ_2)\big\}\Big].$$

Therefore,

$$\phi_Y(\boldsymbol{t}) = \mid I + \beta\Sigma T \mid^{-\alpha_1} \ \mid I_p + \beta\Sigma T \mid^{-\alpha_2} = \ \mid I_p + \beta\Sigma T \mid^{-(\alpha_1+\alpha_2)}.$$

Hence Z follows $MG_p(\alpha_1 + \alpha_2, \beta, \Sigma)$.

Note that lemma 1 is an extension of proposition 8.4 in Eaton (2000).

Lemma 2. *If Z follows $MG_p(\alpha, \beta, \Sigma)$ and $\mid Z \mid$ denotes the determinant of Z then*

$$E\big[\mid Z \mid^h\big] = \frac{\Gamma_p(\alpha + h)}{\Gamma_p(\alpha)} . \frac{\beta^{hp}}{\mid \Sigma \mid^h} \tag{4}$$

provided $Re(h) \geq -\alpha + \frac{p-1}{2}$.

Proof.

$$\begin{aligned} E\big[\mid Z \mid^h\big] &= \frac{\mid \Sigma \mid^\alpha}{\Gamma_p(\alpha)\beta^{\alpha.p}} \int_{Z>0} \exp\Big\{-\frac{1}{\beta} tr\Sigma^{-1}Z\Big\} \mid Z \mid^{h+\alpha-\frac{1}{2}(p+1)} dZ \\ &= \frac{\mid \Sigma \mid^\alpha}{\Gamma_p(\alpha)\beta^{\alpha.p}} . \frac{\Gamma_p(\alpha + h)\beta^{(\alpha+h)p}}{\mid \Sigma \mid^{\alpha+h}} \\ &= \frac{\Gamma_p(\alpha + h)}{\Gamma_p(\alpha)} . \frac{\beta^{hp}}{\mid \Sigma \mid^h}. \end{aligned}$$

□

Note that, if $h = 1$, then

$$E\big[\mid Z \mid\big] = \frac{\beta}{\mid \Sigma \mid}.\frac{\Gamma\big(\alpha+1\big)\Gamma\big(\alpha+\frac{1}{2}\big)}{\Gamma\big(\alpha+\frac{1}{2}(2-p)\big)\Gamma\big(\alpha+\frac{1}{2}(1-p)\big)}. \tag{5}$$

This follows from (4) as

$$\begin{aligned} E\big[\mid Z \mid\big] &= \frac{\beta}{\mid \Sigma \mid}.\frac{\Gamma_p(\alpha+1)}{\Gamma_p(\alpha)} \\ &= \frac{\beta}{\mid \Sigma \mid}.\frac{\pi^{\frac{p(p-1)}{4}}\prod_{j=1}^{p}\Gamma\big(\alpha+1-\frac{1}{2}(j-1)\big)}{\pi^{\frac{p(p-1)}{4}}\prod_{j=1}^{p}\Gamma\big(\alpha-\frac{1}{2}(j-1)\big)} \\ &= \frac{\beta}{\Sigma}.\frac{\Gamma\big(\alpha+1\big)\Gamma\big(\alpha+\frac{1}{2}\big)}{\Gamma\big(\alpha+\frac{1}{2}(2-p)\big)\Gamma\big(\alpha+\frac{1}{2}(1-p)\big)}. \end{aligned}$$

Also note that if we choose, $h = 1, p = 1$, and $\Sigma = I$, then Z follows univariate $Gamma(\alpha, \beta)$ and (5) implies $E(Z) = \alpha\beta$, which is the mean of a univariate gamma distribution. Next we develop a result which shows the relation between the generalized multivariate gamma and generalized multivariate beta distributions.

Lemma 3. *If Z_1 follows $MG_p(\alpha_1, \beta, \Sigma)$, Z_2 follows $MG_p(\alpha_2, \beta, \Sigma)$ and Z_1, Z_2 are independent of each other. Then $U = Z_1(Z_1+Z_2)^{-1}$ follows the multivariate generalized beta distribution with parameters α_1 and α_2 with the density as,*

$$f(U) = \frac{\Gamma_p(\alpha_1+\alpha_2)}{\Gamma_p(\alpha_1)\Gamma_p(\alpha_2)} \mid U \mid^{\alpha_1-\frac{1}{2}(p+1)} \mid I - U \mid^{\alpha_2-\frac{1}{2}(p+1)} .$$

Proof. Consider the transformations $U = Z_1(Z_1+Z_2)^{-1}$ and $V = (Z_1+Z_2)$. We can argue that U is ancillary to Σ and V is complete sufficient. Therefore, by Basu's theorem, U and V are independent of each other. Now the joint distribution of Z_1 and Z_2 is

$$\begin{aligned} f(Z_1, Z_2) =& \frac{\mid \Sigma \mid^{\alpha_1}}{\Gamma_p(\alpha_1)\beta^{\alpha_1 p}} \exp\Big\{-\frac{1}{\beta}tr\Sigma^{-1}Z_1\Big\} \mid Z_1 \mid^{\alpha_1-\frac{1}{2}(p+1)} . \\ & \frac{\mid \Sigma \mid^{\alpha_2}}{\Gamma_p(\alpha_2)\beta^{\alpha_2 p}} \exp\Big\{-\frac{1}{\beta}tr\Sigma^{-1}Z_2\Big\} \mid Z_2 \mid^{\alpha_2-\frac{1}{2}(p+1)} \\ =& \frac{\mid \Sigma \mid^{\alpha_1+\alpha_2}}{\Gamma_p(\alpha_1+\alpha_2)\beta^{(\alpha_1+\alpha_2)p}} \exp\Big\{-\frac{1}{\beta}tr\Sigma^{-1}(Z_1+Z_2)\Big\} \mid Z_1+Z_2 \mid^{\alpha_1+\alpha_2-\frac{1}{2}(p+1)} \\ . & \frac{\Gamma_p(\alpha_1+\alpha_2)}{\Gamma_p(\alpha_1)\Gamma_p(\alpha_2)} \mid Z_1 \mid^{\alpha_1-\frac{1}{2}(p+1)} \mid Z_2 \mid^{\alpha_2-\frac{1}{2}(p+1)} . \end{aligned}$$

Therefore we can simplify this as,

$$
\begin{aligned}
f(Z_1, Z_2) =& \frac{|\ \Sigma\ |^{\alpha_1+\alpha_2}}{\Gamma_p(\alpha_1+\alpha_2)\beta^{(\alpha_1+\alpha_2)p}} \exp\Big\{-\frac{1}{\beta}tr\Sigma^{-1}(Z_1+Z_2)\Big\}\ |\ Z_1+Z_2\ |^{\alpha_1+\alpha_2-\frac{1}{2}(p+1)} \\
&\cdot\quad \frac{\Gamma_p(\alpha_1+\alpha_2)}{\Gamma_p(\alpha_1)\Gamma_p(\alpha_2)}\ |\ Z_1(Z_1+Z_2)^{-1}\ |^{\alpha_1-\frac{1}{2}(p+1)}|\ I - Z_1(Z_1+Z_2)^{-1}\ |^{\alpha_2-\frac{1}{2}(p+1)} \\
&= f(V).f(U) = f(V,U)\ ,
\end{aligned}
$$

where $V = (Z_1 + Z_2)$ follows $MG_p(\alpha_1 + \alpha_2, \beta, \Sigma)$ by the previous result and $U = Z_1(Z_1 + Z_2)^{-1}$ follows multivariate beta distribution with parameters α_1 and α_2 having density as

$$
f(U) = \frac{\Gamma_p(\alpha_1+\alpha_2)}{\Gamma_p(\alpha_1)\Gamma_p(\alpha_2)}\ |\ U\ |^{\alpha_1-\frac{1}{2}(p+1)}|\ I-U\ |^{\alpha_2-\frac{1}{2}(p+1)}\ .
$$

□

In the next result we develop the inverted generalized multivariate gamma distribution when Z is a real symmetric random matrix of order p having a generalized multivariate gamma distribution.

Lemma 4. *Let Z be a real symmetric random matrix of order p which follows $MG_p(\alpha, \beta, \Sigma)$, then $W = Z^{-1}$ follows an inverted generalized multivariate gamma distribution, denoted as $InvMG_p(\alpha, \beta, \Sigma^{-1})$ with the density function as*

$$
f(W) = \frac{|\ \Sigma^{-1}\ |^{\alpha}}{\Gamma_p(\alpha)\beta^{\alpha p}} \exp\Big\{-\frac{1}{\beta}tr\Sigma^{-1}W^{-1}\Big\}\ |\ W\ |^{-\alpha-\frac{1}{2}(p+1)}, \tag{6}
$$

where W and Σ^{-1} are positive definite, $\alpha \geq \frac{p-1}{2}$, $\beta \geq 0$.

Proof. We know the Jacobian of transformation of $W = Z^{-1}$ is

$$
J(Z \to W) =|\ W\ |^{-(p+1)}\ .
$$

Hence, we have the density of the W as in (6). □

Note that similar result for inverse Wishart distribution is available in Anderson (1984) and Gupta and Nagar (2000). If we choose $\alpha = \frac{m}{2}$, where m is a positive integer, generally consider as the sample size and $\beta = 2$ then W

follows inverted Wishart distribution with $\Psi = \Sigma^{-1}$ precision matrix and m degrees of freedom with density

$$f(W) = \frac{\mid \Psi \mid^{\frac{m}{2}} \mid W \mid^{-\frac{1}{2}(m+p+1)} \exp\left\{ -\frac{1}{2} tr \Psi W^{-1} \right\}}{2^{\frac{mp}{2}} \Gamma_p(\frac{m}{2})}.$$

Again if $p = 1$ and $\Sigma = I$ then W has a univariate inverse gamma distribution with parameter as α, β and density as

$$f(w) = \frac{1}{\Gamma(\alpha)\beta^{\alpha}} \; e^{-\frac{1}{\beta w}} \left(\frac{1}{w}\right)^{\alpha+1},$$

where $0 \leq w < \infty$, $\alpha \geq 0$ and $\beta \geq 0$. Next we derive the mode of inverse multivariate gamma distribution which will be needed for the Bayesian inference.

Lemma 5. *Suppose W is a real symmetric random matrix of order p which follows $InvMG_p(\alpha, \beta, \Sigma)$, then the mode of W is*

$$\widehat{W} = \frac{\Sigma^{-1}}{\beta(\alpha + \frac{1}{2}(p+1))}. \tag{7}$$

Proof. Taking logarithm on both sides of the density of *inverse multivariate gamma* (6), we have

$$\ln f(W) = Const. - \frac{1}{\beta} tr\big(\Sigma^{-1} W^{-1}\big) - \big(\alpha + \frac{1}{2}(p+1)\big) \ln \mid W \mid .$$

Now differentiating with respect to W and setting that with equate to $\mathbf{0}$, we have

$$\begin{aligned}
-\frac{1}{\beta} \frac{\partial tr\big(\Sigma^{-1} W^{-1}\big)}{\partial W} - \big(\alpha + \frac{1}{2}(p+1)\big) \frac{\partial \ln \mid W \mid}{\partial W} &= \mathbf{0}, \\
\frac{W^{-1}\big(\Sigma^{-1}\big) W'^{-1}}{\beta} - \big(\alpha + \frac{1}{2}(p+1)\big) \big(W'\big)^{-1} &= \mathbf{0}.
\end{aligned}$$

Hence the mode of W is

$$\widehat{W} = \frac{\Sigma^{-1}}{\beta\big(\alpha + \frac{1}{2}(p+1)\big)}.$$

□

In the next section, we develop the posterior distribution and the Bayesian inference.

4. Bayesian Inference Technique

Recently Tsionas (2004) developed the Bayesian inference for the multivariate gamma distribution, using Gibbs sampling and data augmentation approach. This technique involves costly matrix inversion whose computational complexity increase in cubic order, hence make inference infeasible for Σ with large p. Also MCMC based solution yields either 'posterior mean' or 'posterior median', which are Bayes estimator under 'squared error loss' and 'absolute error loss' respectively. In this paper, we provided 'posterior mode' as solution. So it is required to show that 'posterior mode' (or MAP) is Bayes estimator under certain loss function.

In this section we take Bayesian decision theoretic approach and as an auxiliary result we show that maximum posteriori (MAP) estimator (or posterior mode), which has a closed form expression, is a Bayes estimator under certain loss function. Then we present the marginal distribution of Z which has generalized multivariate gamma distribution. The following theorem shows that MAP estimator is a Bayes estimator.

Theorem 1. *Suppose A is an estimate for the unknown scale matrix Σ, where $\pi(A \mid \mathcal{D}_n)$ and $\pi(\Sigma \mid \mathcal{D}_n)$ are the corresponding posterior probability density functions over $\mathbb{R}^p$ respectively, where $\mathcal{D}_n$ indicates sample data of size n. Now the posterior expected loss of A, when the posterior distribution is $\pi(\Sigma \mid \mathcal{D}_n)$, is*

$$\rho(\Sigma, A) = \int_{\Sigma>0} L(\Sigma, A) dF^{\pi(\Sigma|\mathcal{D})}(\Sigma),$$

where the loss function is $L(\Sigma, A) = \log\left(\frac{\pi(\Sigma \mid \mathcal{D})_n}{\pi(A \mid \mathcal{D})}\right)$. Then the Bayes rule is

$$\delta^\pi(\mathcal{D}) = \operatorname*{argmax}_{\Sigma>0} \pi(\Sigma \mid \mathcal{D}_n),$$

which is the MAP estimator or mode of posterior distribution of $\pi(\Sigma \mid \mathcal{D}_n)$.

Proof. The value of A which minimizes the $\rho(\Sigma, A)$ can be obtained as follows. Assuming all integrals are finite,

$$\begin{aligned}
& \min_A \int_{\Sigma>0} \log\left(\frac{\pi(\Sigma|\mathcal{D}_n)}{\pi(A|\mathcal{D}_n)}\right) \pi(\Sigma|\mathcal{D}_n) d\Sigma \\
= & \min_A \left[\int_{\Sigma>0} \log \pi(\Sigma|\mathcal{D}_n) \cdot \pi(\Sigma|\mathcal{D}_n) d\Sigma - \int_{\Sigma>0} \log \pi(A|\mathcal{D}_n) \cdot \pi(\Sigma|\mathcal{D}_n) d\Sigma \right] \\
= & \max_A \log \pi(A|\mathcal{D}_n)
\end{aligned}$$

here $\log \pi(A \mid \mathcal{D}_n) : \mathbb{R}^{p\times p} \to \mathbb{R}$ is the log of density with respect to A. Hence solving for A gives the required results. $\square$

Remark 1. The posterior expected loss in Theorem 1, $\rho(\Sigma, A) = E\Big[\log\Big(\frac{\pi(\Sigma|\mathcal{D}_n)}{\pi(A|\mathcal{D}_n)}\Big)\Big]$, can be interpreted as Kullback Leibler divergence of the posterior distribution evaluated under action A from the true posterior distribution of unknown parameter Σ. Therefore, posterior expected loss or Kullback Leibler divergence is minimum, if we choose our action as posterior mode or MAP estimator.

Remark 2. Though Theorem 1, is proved for scale matrix Σ; this result is true for any proper posterior density.

4.1. Estimation with MAP Estimator.

Theorem 2. *Suppose Z is a real symmetric random matrix of order p which follows $MG_p(\alpha, \beta, \Sigma)$, further Σ has a prior distribution as $InvMG_p(a, b, \Psi)$, then Z has marginal distribution with density as*

$$m(Z) = \frac{\Gamma_p(a+\alpha)}{\Gamma_p(a)+\Gamma_p(\alpha)} \mid Z \mid^{\alpha-\frac{1}{2}(p+1)} \mid \Psi \mid^{\alpha} \mid Z+\Psi \mid^{-(a+\alpha)} . \tag{8}$$

Note that Z and Ψ are positive definite, $\alpha \geq \frac{p-1}{2}$ and $a \geq \frac{p-1}{2}$.

Proof. We can derive the marginal distribution of Z by integrating out the Σ from the joint distribution of Z and Σ.
That is,

$$m(Z) = \int_{\Sigma>0} f(Z \mid \Sigma)\pi(\Sigma) d\Sigma$$

which implies

$$m(Z) = \frac{\mid Z \mid^{\alpha-\frac{1}{2}(p+1)} \mid \Psi \mid^{a}}{\beta^{(a+\alpha)p}\Gamma_p(a)\Gamma_p(\alpha)} \int_{\Sigma>0} e^{-\frac{1}{\beta}tr(Z+\Psi)\Sigma^{-1}} \mid \Sigma \mid^{-\alpha-a-\frac{1}{2}(p+1)} d\Sigma,$$

which is equivalently,

$$m(Z) = \frac{\mid Z \mid^{\alpha-\frac{1}{2}(p+1)} \mid \Psi \mid^{a}}{\beta^{(a+\alpha)p}\Gamma_p(a)\Gamma_p(\alpha)} \Gamma_p(a+\alpha)\beta^{(a+\alpha)p} \mid Z+\Psi \mid^{-(a+\alpha)} .$$

Hence we have

$$m(Z) = \frac{\Gamma_p(a+\alpha)}{\Gamma_p(a)+\Gamma_p(\alpha)} \mid Z \mid^{\alpha-\frac{1}{2}(p+1)} \mid \Psi \mid^{\alpha} \mid Z+\Psi \mid^{-(a+\alpha)} .$$

$\square$

Remark 3. If we choose $\alpha = \frac{n_1}{2}$ and $a = \frac{n_2}{2}$ in (8), then the marginal distribution of Z has a generalized multivariate F-distribution (9) with n_1 and n_2 degrees of freedom and scale matrix Ψ, having density as

$$m(Z) = \frac{\Gamma_p\left(\frac{n_1+n_2}{2}\right)}{\Gamma_p\left(\frac{n_1}{2}\right) + \Gamma_p\left(\frac{n_2}{2}\right)} \mid Z \mid^{\frac{n_1-p-1}{2}} \mid \Psi \mid^{-\frac{n_1}{2}} \mid I + \Psi^{-1} Z \mid^{-\frac{n_1+n_2}{2}}, \qquad (9)$$

where in general, n_1 and n_2 are the sample sizes.

In the next result, we derive the posterior distribution of Σ, when the prior distribution has inverse multivariate gamma distribution.

Theorem 3. *Let Z be a real symmetric random matrix of order p, which follows $MG_p(\alpha, \beta, \Sigma)$,. Further Σ has prior distribution as $InvMG_p(\alpha, \beta, \Psi)$, then the posterior distribution of Σ is $InvMG_p(\alpha + a, \beta, Z + \Psi)$.*

Proof. We know from the previous result that the joint distribution of Σ and Z is

$$f(Z, \Sigma) = \frac{\mid Z \mid^{\alpha - \frac{1}{2}(p+1)} \mid \Psi \mid^{a}}{\beta^{(a+\alpha)p} \Gamma_p(a) \Gamma_p(\alpha)} e^{-\frac{1}{\beta} tr(Z+\Psi)\Sigma^{-1}} \mid \Sigma \mid^{-\alpha - a - \frac{1}{2}(p+1)},$$

and Z has marginal distribution in (8). Therefore the posterior distribution of Σ given Z follows $InvMG_p(\alpha + a, \beta, Z + \Psi)$ with density as

$$\pi(\Sigma \mid Z) = \frac{\mid Z + \Psi \mid^{a+\alpha}}{\Gamma_p(a+\alpha) \beta^{(a+\alpha)p}} e^{-\frac{1}{\beta} tr(Z+\Psi)\Sigma^{-1}} \mid \Sigma \mid^{-(a+\alpha) - \frac{1}{2}(p+1)},$$

where Z and Ψ are positive definite, $\alpha \geq \frac{p-1}{2}$, $a \geq \frac{p-1}{2}$, and $\beta \geq 0$. $\square$

Theorem 4. *The MAP estimate (or Posterior mode) for Σ is*

$$\widehat{\Sigma}_{MAP} = \frac{\left(Z + \Psi\right)^{-1}}{\beta\left\{\left(\alpha + a\right) + \frac{1}{2}(p+1)\right\}}, \qquad (10)$$

where, the posterior distribution of Σ is $InvMG_p(\alpha + a, \beta, Z + \Psi)$, $\alpha + a \geq \frac{p-1}{2}$, $\beta \geq 0$, $(Z + \Psi)$ is positive definite.

Remark 4. Proof of Theorem 4 follows directly from Lemma 5.

4.2. Test using Bayes factor

According to classical Bayesian hypothesis testing, a null hypothesis H_0 : $\Sigma = \Sigma_0$ and $H_1 : \Sigma = \Sigma_1$ are specified. Then the Bayes factor in favor of Σ_0 can be defined as

$$B = \frac{\int_{\Sigma \in \Sigma_0} f(Z \mid \Sigma) dF^{\pi_0}(\Sigma)}{\int_{\Sigma \in \Sigma_1} f(Z \mid \Sigma) dF^{\pi_1}(\Sigma)}, \tag{11}$$

where Z is an action or estimator based on data.

Example. We can assume Σ_0 has $AR(1)$ structure and Σ_1 has a compound symmetry structure. Then the Bayes factor B in (11) can be interpreted as summary of evidence provided by the data in favor of one scientific theory, represented by the statistical model which assumes $AR(1)$ structure as scale matrix, as opposed to another which assumes compound symmetry structure.

The following theorem formulates the Bayes factor for the two competing scale matrices of the generalized multivariate gamma distribution.

Theorem 5. *Suppose Z is a real symmetric random matrix of order p which follows $MG_p(\alpha, \beta, \Sigma)$. Consider the testing $H_0 : \Sigma = \Sigma_0$ versus $H_1 : \Sigma = \Sigma_1$. Further we assume Σ_0 has a priori distribution as $InvGamma(a_0, b_0, \Psi_0)$ and Σ_1 has a priori distribution as $InvGamma(a_1, b_1, \Psi_1)$. Then the Bayes Factor in favor of H_0 is*

$$B = \frac{\Gamma_p(a_0 + \alpha)}{\Gamma_p(a_1 + \alpha)} . \frac{\Gamma_p(a_1) + \Gamma_p(\alpha)}{\Gamma_p(a_0) + \Gamma_p(\alpha)} . \left[\frac{\mid \Psi_0 \mid}{\mid \Psi_1 \mid} \right]^{\alpha} . \frac{\mid Z + \psi_0 \mid^{-(a_0+\alpha)}}{\mid Z + \psi_1 \mid^{-(a_1+\alpha)}}.$$

Proof. Under H_0, the marginal $m_0(Z)$ can be obtained from (8) as

$$m_0(Z) = \frac{\Gamma_p(a_0 + \alpha)}{\Gamma_p(a_0) + \Gamma_p(\alpha)} \mid Z \mid^{\alpha - \frac{1}{2}(p+1)} \mid \Psi_0 \mid^{\alpha} \mid Z + \Psi_0 \mid^{-(a_0+\alpha)} .$$

Under H_1, the marginal $m_1(Z)$ can be obtained from (8) as

$$m_1(Z) = \frac{\Gamma_p(a_1 + \alpha)}{\Gamma_p(a_1) + \Gamma_p(\alpha)} \mid Z \mid^{\alpha - \frac{1}{2}(p+1)} \mid \Psi_1 \mid^{\alpha} \mid Z + \Psi_1 \mid^{-(a_1+\alpha)} .$$

Now the Bayes factor B can be obtained as $B = \frac{m_0(Z)}{m_1(Z)}$. □

5. Concluding Remarks

In this paper we present the generalized multivariate gamma distribution and develop important properties of this distribution. We presented the additive property and establish its relationship with generalized multivariate beta distribution. We presented the inverted multivariate gamma distribution and its derived its mode in closed form expression. We presented a general results that shows that MAP estimator is Bayes estimator under certain loss function. Using this result we presented the Bayesian inference for the related scale matrix Σ. We also developed the testing problem for Σ using Bayes factor.

Acknowledgment

Authors would like to thank anonymous referee for many constructive comments that have improved the quality of the manuscript. First author's research was partially supported by NSF Grant DMS-0635449 to Statistical and Applied Mathematical Science Institute (SAMSI).

References

E. Tsionas, Bayesian Inference for multivariate gamma distribution, Statistics and Computing 14 (2004) 223–233.

A. S. Krishnamoorty, M. Parthasarathy, A Multivariate Gamma Type Distribution, Annals of Mathematical Statistics 22 (1951) 549–557, correction, ibid, (1960) 31, p.229.

P. R. Krishnaiah, M. M. Rao, Remarks on Multivariate Gamma Distribution, American Mathematical Monthly 68 (1961) 342–346.

E. Lukacs, R. G. Laha, Application of Characteristic Functions, Charles Griffin & Co. Ltd., 1964.

R. C. Griffiths, Characterization of infinite divisible multivariate gamma distribution, Journal of Multivariate Analysis 15 (1984) 13–20.

R. B. Bapat, Infinite Divisibility of Multivariate Gamma Distribution and M-Matrices, Sankhya-A 51 (1989) 73–78.

A. Mathai, P. Moschopoulos, On A Multivariate Gamma, Journal of Multivariate Analysis 39 (1991) 135 – 153.

A. Mathai, P. Moschopoulos, A form of multivariate gamma distribution, Annals of the Institute of Statistical Mathematics 44 (1992) 97 – 103.

C. A. Coelho, The Generalized Integer Gamma Distribution - A Basis for Distributions in Multivariate Statistics, Journal of Multivariate Analysis 64 (1998) 86–102.

M. Eaton, Multivariate statistics: a vector space approach., IMS Lecture Notes Monograph Series, Volume 53, 2000.

S. Kotz, N. Balakrishnan, N. Johnson, Continuous multivariate distributions, Wiley, 2000.

E. Furman, On Multivariate Gamma Distribution, Statistics and Probability Letters 78 (2008) 2353–2360.

T. Anderson, An Introduction to Multivariate Statistical Analysis, Wiley, second edn., 1984.

A. Gupta, D. Nagar, Matrix variate distributions, Chapman & Hall, 2000.