On Bayesian Inference for Generalized Multivariate Gamma Distribution

Sourish Das\textsuperscript{a,b}, Dipak K Dey\textsuperscript{c}

\textsuperscript{a}Department of Statistical Science, Duke University, Durham, NC
\textsuperscript{b}SAMSI Statistical and Applied Mathematical Science Institute, Durham, NC
\textsuperscript{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 $\Sigma$. We show that maximum posteriori (MAP) estimate is a Bayes estimator. We also develop the testing problem for $\Sigma$ using Bayes factor. This approach provides a mathematically closed form solution for $\Sigma$. 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 $\Sigma$, for large dimension. In this paper, we provide an elegant closed form Bayes factor for $\Sigma$.

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 Krishnan-niah 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 $\alpha$ and the scale matrix $\Sigma$, when $\alpha$ 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 multivari-
ate extension of three parameters univariate gamma distribution and obtain
the explicit forms for the moments, moment generating function and condi-
tional moments. Coelho (1998) presented the exact distribution of general-
ized 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 insur-
ance losses. In this paper, we develop a Bayesian inference for MG presented
in equation (1). This version takes care of wide range of covariance dependent
variables 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 ex-
pression 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 general-
ized 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 Σ = ((σ_{ij})) be a real, symmetric positive definite matrix of order p
and 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

\[
f(Z) = K \exp \left\{ -\frac{1}{\beta} tr \Sigma^{-1} Z \right\} \ | \ Z \ |^{\lambda-1}, \quad Z > 0,
\]

where \( Z > 0 \) indicates that \( Z \) is positive definite matrix, \( | \cdot | \) denote de-
terminant, \( tr \) denote trace of a matrix, \( \lambda \) and \( \beta \) are positive real numbers
and \( K \) is normalizing constant. Lukacs and Laha (1964), showed that the
normalizing constant is given as

\[ K = \frac{|\Sigma^{-1}|^{\lambda + \frac{n-1}{2}}}{\beta^{\lambda p} \pi^{\frac{n(p-1)}{2}} \prod_{j=1}^{p} \Gamma\left(\lambda + \frac{p-j}{2}\right)}. \]

If we choose \( \alpha = \lambda + \frac{n-1}{2} \), then

\[ K = \frac{|\Sigma^{-1}|^{\alpha}}{\beta^{p} \pi^{\frac{n(p-1)}{2}} \prod_{j=1}^{p} \Gamma\left(\alpha - \frac{1}{2}(j-1)\right)}. \]

Defining the multivariate gamma function (Anderson, 1984)

\[ \Gamma_p(\alpha) = \pi^{\frac{n(p-1)}{4}} \prod_{j=1}^{p} \Gamma\left(\alpha - \frac{1}{2}(j-1)\right), \]

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

\[ f(Z) = \frac{|\Sigma^{-1}|^{\alpha} |Z|^{\alpha - \frac{1}{2}(p+1)}}{\beta^{p} \Gamma_p(\alpha)} \exp\left\{-\frac{1}{\beta} tr\Sigma^{-1}Z\right\} \quad Z > 0, \quad (1) \]

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

\[ f(Z) = \frac{|Z|^{\frac{1}{2}(n-p-1)}}{\Gamma_p(\frac{n}{2}) |\Sigma|^{\frac{1}{2} \frac{n-p-1}{2}}} \exp\left\{-\frac{1}{2} tr\Sigma^{-1}Z\right\}, \]

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

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

\[ \phi_Z(t) = E\left[ \exp\left\{-tr(TZ)\right\}\right] = |I_p + \beta \Sigma T|^{-\alpha}. \quad (2) \]

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

\[ \psi_z(t) = E\left[i \exp\left\{tr(TZ)\right\}\right] = |I_p - i\beta \Sigma T|^{-\alpha}, \]

3
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 $\alpha$ 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

$$f(Z) = \frac{|\Sigma|^\alpha}{\Gamma_p(\alpha)\beta^{p}} \exp\left\{-\frac{1}{\beta} tr(\Sigma^{-1}Z)\right\} |Z|^ {\alpha - \frac{1}{2}(p+1)}, Z > 0 \quad (3)$$

has an infinitely divisible multivariate gamma distribution with parameters $\alpha \geq \frac{p-1}{2}, \beta \geq 0$, and $\Sigma$ 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 compound symmetry it is assumed that the variance is constant across occasions, say $\sigma^2$, and $\text{Corr}(Y_{ij}, Y_{ik}) = \rho$ for all $j$ and $k$. That is

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

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 $\Sigma$ are non-negative. Again all the off diagonal elements of $\Sigma^{-1}$ are non-positive, where $\Sigma^{-1}$ is given as

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

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 $\Sigma$ of the heterogeneous variances autoregressive (HVAR) model is given by

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

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

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

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

$$\phi_Y(t) = E\left[ \exp \left\{ -tr(TY) \right\} \right] = E\left[ \exp \left\{ -tr(T(Z_1 + Z_2)) \right\} \right]$$

which is equivalent to

$$\phi_Y(t) = E\left[ \exp \left\{ -tr(TZ_1) \right\} \right] E\left[ \exp \left\{ -tr(TZ_2) \right\} \right].$$

Therefore,

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

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 $|Z|$ denotes the determinant of $Z$ then

$$E\left[ |Z|^h \right] = \frac{\Gamma_p(\alpha + h) \beta^{hp}}{\Gamma_p(\alpha) \beta^{h p}} \left| \Sigma \right|^h$$

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

**Proof.**

$$E\left[ |Z|^h \right] = \frac{\left| \Sigma \right|^\alpha}{\Gamma_p(\alpha) \beta^{p \alpha \beta}} \int_{Z>0} \exp \left\{ -\frac{1}{\beta} tr\Sigma^{-1}Z \right\} |Z|^h \left| \Sigma \right|^{\alpha - \frac{1}{2} (p+1)} dZ$$

$$= \frac{\left| \Sigma \right|^\alpha \Gamma_p(\alpha + h) \beta^{(\alpha + h)p}}{\Gamma_p(\alpha) \beta^{hp} \left| \Sigma \right|^{\alpha + h}}$$

$$= \frac{\Gamma_p(\alpha + h) \beta^{hp}}{\Gamma_p(\alpha) \beta^{hp} \left| \Sigma \right|^{\alpha + h}}.$$
Consider the transformations 

\[
\Gamma
\]

univariate. Also note that if we choose,

This follows from (4) as

Note that, if

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 \( \alpha_1 \) and \( \alpha_2 \) with the density as,

\[
f(U) = \frac{\Gamma_p(\alpha_1 + \alpha_2)}{\Gamma_p(\alpha_1)\Gamma_p(\alpha_2)} \cdot \left| U \right|^{\alpha_1-\frac{1}{2}(p+1)} \cdot \left| I - U \right|^{\alpha_2-\frac{1}{2}(p+1)}.
\]

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 \( \Sigma \) 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

\[
f(Z_1, Z_2) = \frac{\left| \Sigma \right|^{\alpha_1}}{\Gamma_p(\alpha_1)^{\beta^{\alpha_1}}} \cdot \exp \left\{-\frac{1}{\beta} tr \Sigma^{-1} Z_1 \right\} \cdot \left| Z_1 \right|^{\alpha_1-\frac{1}{2}(p+1)}.
\]

\[
= \frac{\left| \Sigma \right|^{\alpha_2}}{\Gamma_p(\alpha_2)^{\beta^{\alpha_2}}} \cdot \exp \left\{-\frac{1}{\beta} tr \Sigma^{-1} Z_2 \right\} \cdot \left| Z_2 \right|^{\alpha_2-\frac{1}{2}(p+1)}.
\]

\[
= \frac{\left| \Sigma \right|^{\alpha_1+\alpha_2}}{\Gamma_p(\alpha_1 + \alpha_2)^{\beta^{\alpha_1+\alpha_2}}} \cdot \exp \left\{-\frac{1}{\beta} tr \Sigma^{-1} (Z_1 + Z_2) \right\} \cdot \left| Z_1 + Z_2 \right|^{\alpha_1+\alpha_2-\frac{1}{2}(p+1)}.
\]

\[
\cdot \frac{\Gamma_p(\alpha_1 + \alpha_2)}{\Gamma_p(\alpha_1)\Gamma_p(\alpha_2)} \cdot \left| Z_1 \right|^{\alpha_1-\frac{1}{2}(p+1)} \cdot \left| Z_2 \right|^{\alpha_2-\frac{1}{2}(p+1)}.
\]
Therefore we can simplify this as,
\[
f(Z_1, Z_2) = \frac{|\Sigma|^{\alpha_1 + \alpha_2}}{\Gamma_p(\alpha_1 + \alpha_2 )\beta^{(\alpha_1 + \alpha_2)p}} \exp \left\{ -\frac{1}{\beta} tr\Sigma^{-1}(Z_1 + Z_2) \right\} |Z_1 + Z_2|^{\alpha_1 + \alpha_2 - \frac{1}{2}(p+1)} \cdot \frac{\Gamma_p(\alpha_1 + \alpha_2)}{\Gamma_p(\alpha_1)\Gamma_p(\alpha_2)} |Z_1(Z_1 + Z_2)^{-1}|^{\alpha_1 - \frac{1}{2}(p+1)} |I - Z_1(Z_1 + Z_2)^{-1}|^{\alpha_2 - \frac{1}{2}(p+1)} \]
\[
= f(V).f(U) = f(V, U),
\]
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 \( \alpha_1 \) and \( \alpha_2 \) 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
\[
f(W) = \frac{\Gamma_p(\alpha_1 + \alpha_2)}{\Gamma_p(\alpha_1)\Gamma_p(\alpha_2)} |W|^{-\frac{1}{2}(p+1)} |I - W|^{\alpha_2 - \frac{1}{2}(p+1)} .
\]
where \( W \) and \( \Sigma^{-1} \) are positive definite, \( \alpha \geq \frac{p-1}{2} \), \( \beta \geq 0 \).

**Proof.** We know the Jacobian of transformation of \( W = Z^{-1} \) is
\[
J(Z \rightarrow W) = |W|^{-(p+1)} .
\]
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

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

Again if $p = 1$ and $\Sigma = I$ then $W$ has a univariate inverse gamma distribution with parameter as $\alpha$, $\beta$ 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 $Inv\mathcal{M}G_p(\alpha, \beta, \Sigma)$, then the mode of $W$ is

$$
\hat{W} = \Sigma^{-1} \beta \left(\alpha + \frac{1}{2}(p + 1)\right).
$$

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

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

Now differentiating with respect to $W$ and setting that with equate to 0, we have

$$
-\frac{1}{\beta} \frac{\partial tr(\Sigma^{-1}W^{-1})}{\partial W} - (\alpha + \frac{1}{2}(p + 1)) \frac{\partial \ln |W|}{\partial W} = 0,
$$

$$
\frac{W^{-1}(\Sigma^{-1})W^{-1}}{\beta} - (\alpha + \frac{1}{2}(p + 1)) (W')^{-1} = 0.
$$

Hence the mode of $W$ is

$$
\hat{W} = \Sigma^{-1} \beta \left(\alpha + \frac{1}{2}(p + 1)\right).
$$

\[\square\]

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

9
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 \(\Sigma\) 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 \(\Sigma\), where \(\pi(A \mid D_n)\) and \(\pi(\Sigma \mid D_n)\) are the corresponding posterior probability density functions over \(\mathbb{R}^p\) respectively, where \(D_n\) indicates sample data of size \(n\). Now the posterior expected loss of \(A\), when the posterior distribution is \(\pi(\Sigma \mid D_n)\), is

\[
\rho(\Sigma, A) = \int_{\Sigma > 0} L(\Sigma, A) dF^{\pi(\Sigma \mid D_n)}(\Sigma),
\]

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

\[
\delta^\pi(D) = \arg\max_{\Sigma > 0} \pi(\Sigma \mid D_n),
\]

which is the MAP estimator or mode of posterior distribution of \(\pi(\Sigma \mid D_n)\).

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

\[
\min_A \int_{\Sigma > 0} \log \left( \frac{\pi(\Sigma \mid D_n)}{\pi(A \mid D_n)} \right) \pi(\Sigma \mid D_n) d\Sigma = \min_A \left[ \int_{\Sigma > 0} \log \pi(\Sigma \mid D_n) \cdot \pi(\Sigma \mid D_n) d\Sigma - \int_{\Sigma > 0} \log \pi(A \mid D_n) \cdot \pi(\Sigma \mid D_n) d\Sigma \right] = \max_A \log \pi(A \mid D_n)
\]
here $\log \pi(A \mid D_n) : \mathbb{R}^{p \times p} \to \mathbb{R}$ is the log of density with respect to $A$. Hence solving for $A$ gives the required results.

**Remark 1.** The posterior expected loss in Theorem 1, $\rho(\Sigma, A) = E\left[ \log \left( \frac{\pi(\Sigma | D_n)}{\pi(A | D_n)} \right) \right]$, can be interpreted as Kullback Leibler divergence of the posterior distribution evaluated under action $A$ from the true posterior distribution of unknown parameter $\Sigma$. 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 $\Sigma$, 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 $\Sigma$ has a prior distribution as $\text{InvMG}_p(a, b, \Psi)$, then $Z$ has marginal distribution with density as

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

**Proof.** We can derive the marginal distribution of $Z$ by integrating out the $\Sigma$ from the joint distribution of $Z$ and $\Sigma$. That is,

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

which implies

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

which is equivalently,

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

Hence we have

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

□
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 \( \Psi \), having density as

\[
m(Z) = \frac{\Gamma_p\left(\frac{n_1+n_2}{2}\right)}{\Gamma_p\left(\frac{n_1}{2}\right) + \Gamma_p\left(\frac{n_2}{2}\right)} |Z|^{-\frac{n_1-p}{2}} |\Psi|^{-\frac{n_1}{2}} |I + \Psi^{-1}Z|^{-\frac{n_1+n_2}{2}},
\]

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

In the next result, we derive the posterior distribution of \( \Sigma \), 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 \( \Sigma \) has prior distribution as \( InvMG_p(\alpha, \beta, \Psi) \), then the posterior distribution of \( \Sigma \) is \( InvMG_p(\alpha + a, \beta, Z + \Psi) \).

**Proof.** We know from the previous result that the joint distribution of \( \Sigma \) and \( Z \) is

\[
f(Z, \Sigma) = \frac{|Z|^{a-p+1}}{\beta(a+\alpha)p \Gamma_p(a) \Gamma_p(\alpha)} \cdot e^{-\frac{1}{2} \text{tr}(Z+\Psi)\Sigma^{-1}} |\Sigma|^{-a-\frac{1}{2}(p+1)},
\]

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

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

where \( Z \) and \( \Psi \) 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 \( \Sigma \) is

\[
\hat{\Sigma}_{MAP} = \frac{(Z + \Psi)^{-1}}{\beta\{ \alpha + a \} + \frac{1}{2}(p + 1)},
\]

where, the posterior distribution of \( \Sigma \) 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 $\Sigma_0$ can be defined as

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

where $Z$ is an action or estimator based on data.

**Example.** We can assume $\Sigma_0$ has AR(1) structure and $\Sigma_1$ has a compound symmetry structure. Then the Bayes factor $B$ in (11) can be interpreted as a 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 $\Sigma_0$ has a priori distribution as $\text{InvGamma}(a_0, b_0, \Psi_0)$ and $\Sigma_1$ has a priori distribution as $\text{InvGamma}(a_1, b_1, \Psi_1)$. Then the Bayes Factor in favor of $H_0$ is

$$B = \frac{\Gamma_p(a_0 + \alpha)}{\Gamma_p(a_1 + \alpha)} \frac{\Gamma_p(a_1 + \alpha)}{\Gamma_p(a_0 + \alpha)} \frac{\Gamma_p(\alpha)}{\Gamma_p(\alpha)} \left| \begin{array}{c} \Psi_0 \\ \Psi_1 \end{array} \right|^{-\alpha} \left| Z + \psi_0 \right|^{-(a_0 + \alpha)} \left| Z + \psi_1 \right|^{-(a_1 + \alpha)}.$$

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

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

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

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

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

\hfill \Box
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 $\Sigma$. We also developed the testing problem for $\Sigma$ 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


