Precision Assessment of Attitude Measurement Devices of Spacecrafts

Hung-Hsiang Liang, *M.S.E., Aerospace Engineering, University of Michigan*

***Abstract*—A Bayesian covariance matrix assessment method of attitude measurement devices of spacecrafts is presented in this paper. The Bayesian inference model (posterior) is established by an inverse Wishart prior and a Gaussian lieklihood which incorporates the experimental measurement data. Metropolis-Hastings Markov Chain Monte Carlo (MH-MCMC) with random walk proposal is used to sample from the posterior and obtain the estimation of the covariance matrix. Further diagnoses of MCMC show that the proposed assessment method can provide simulations with good mixing, rapid convergence, and high efficiency.**

***Index Terms*—Bayesian inference, Markov Chain Monte Carlo, MCMC, precision assessment, attitude measurements, attitude determination**

I. Introduction

ATTITUDE sensors are important for a spacecraft. Estimating their precision is important for fine-tuning the flight control system. Among all attitude sensors, star trackers provide precise attitude measurements. We wish to assess the measurement errors of star trackers using a star simulator which displays skies with known boresights $X = [\alpha, \beta]^T$. The star tracker then takes measurements of the simulated skies and returns Right Ascension (R.A.) and Declination (Dec.) coordinates, denotes $\hat{X} = [\hat{\alpha}, \hat{\delta}]^T$. To avoid projection effects on the celestial sphere, X is shifted with quaternion calculation to $X' = [0, 0]^T$ and $\hat{X}$ is shifted with the same transformation and obtain $\hat{X}' = [\hat{\alpha}', \hat{\delta}']^T$. Given that the measurements are noised by Gaussian noise ξ with zero mean and covariance Σ, that is:

$$\hat{X}' = X' + \xi, \ \xi \sim \mathcal{N}\left(\begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix}\Sigma_{11} & \Sigma_{12}\\ \Sigma_{21} & \Sigma_{22}\end{bmatrix}\right) \tag{1}$$

The likelihood of $\hat{X}'$, i.e. $f_{\hat{X}'|\theta}$, is a Gaussian distribution given θ (i.e. Σ), which needs to be further estimated carefully.

In this project, we propose to estimate the covariance matrix Σ with the Bayesian inference

$$f_{\theta|\hat{X}'} \propto f_{\hat{X}'|\theta} f_\theta \tag{2}$$

where f_θ is the prior, $f_{\hat{X}'|\theta}$ is the likelihood, and $f_{\theta|\hat{X}'}$ is the posterior. As the posterior can be known up to the normalization, the Metropolis-Hastings Markov Chain Monte Carlo (MH-MCMC) can be used to estimate the components of Σ.

The data used in this project are generated from the previous research. The boresight of the star tracker is carefully aligned to a star simulatior (with known coordinate X). If they are perfectly aligned, we should expect $\mathbb{E}(\hat{X}') = \mathbb{E}(\hat{X})$. The measurements of the simulated stars are output from the star tracker and obtain $\hat{X}'$.

This paper is arranged as follows. The Section II will discuss the method, algorithms, and data used in this project. The results and discussions are presents in Section III. And finally, we conclude this paper in Section IV.

II. Methods

A. Bayesian Inference Model

Bayesian inference is described in Equation 2, where f_θ is the prior, $f_{\hat{X}'|\theta}$ is the likelihood, and $f_{\theta|\hat{X}'}$ is the posterior. These distribution is discussed as follows.

1) Prior: The proposed prior is the inverse Wishart distribution since the covariance matrix follows inverse Wishart distribution [1] with the following form:

$$f_{\mathbf{\Sigma}} \triangleq p(\mathbf{\Sigma}) = \frac{|\mathbf{V}|^{m/2}|\mathbf{\Sigma}|^{-(m+p+1)/2}\exp\left[-tr(\mathbf{V}\mathbf{\Sigma^{-1}})/2\right]}{2^{mp/2}\Gamma(m/2)} \tag{3}$$

where $\mathbf{V}$ is the scale matrix, m is the degree of freedom for a $p \times p$ covariance matrix $\mathbf{\Sigma}$ [2].

2) Likelihood: As the measured data $\hat{X}' = [\hat{X_1}', \hat{X_2}', ..., \hat{X_p}']^T$ are noised by Gaussian noise, we have

$$\mathbf{X}|\mathbf{\Sigma} \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma}) \tag{4}$$

then the density function is

$$f_{\hat{X}'|\Sigma} = (2\pi)^{-p/2}|\mathbf{\Sigma}|^{-1/2}\exp\left(-\frac{1}{2}\mathbf{x}^T\mathbf{\Sigma}^{-1}\mathbf{x}\right) \tag{5}$$

If we are given a set of sample $\mathbf{D} = (\mathbf{x_1}, ..., \mathbf{x_n})$ with n sample size, the likelihood function for $\mathbf{\Sigma}$ is

$$\begin{aligned}\mathcal{L}(\mathbf{\Sigma}|\mathbf{D}) &\propto f_{\mathbf{D}|\Sigma} \propto |\mathbf{\Sigma}|^{-n/2}\exp\left(-\frac{1}{2}\sum_{i=1}^{n}\mathbf{x}_i^T\mathbf{\Sigma}^{-1}\mathbf{x}_i\right)\\ &= |\mathbf{\Sigma}|^{-n/2}\exp\left[-\frac{1}{2}tr\left(\sum_{i=1}^{n}\mathbf{x}_i\mathbf{x}_i^T\mathbf{\Sigma}^{-1}\right)\right]\\ &= |\mathbf{\Sigma}|^{-n/2}\exp\left[-\frac{n}{2}tr\left(\mathbf{S}\mathbf{\Sigma}^{-1}\right)\right]\end{aligned} \tag{6}$$

where $\mathbf{S} = \sum_{i=0}^{n}\mathbf{x}_i\mathbf{x}_i^T/n$ is the biased sample covariance matrix and x_i are centered at 0 [2].

3) posterior: The posterior can be obtained using Equation 2 (here we modify the subscript to meet the notations):

$$f_{\mathbf{\Sigma}|\mathbf{D}} \propto \mathcal{L}(\mathbf{\Sigma}|\mathbf{D}) f_{\mathbf{\Sigma}} \tag{7}$$

substitude Equation 7 with Equations 6 and 3, respectively, we then can obtain the following posterior:

$$\begin{aligned} f_{\mathbf{\Sigma}|\mathbf{D}} \propto & |\mathbf{\Sigma}|^{-n/2} \exp\left[-\frac{n}{2} tr\left(\mathbf{S}\mathbf{\Sigma}^{-1}\right)\right] \\ & |\mathbf{\Sigma}|^{-(m+p+1)/2} \exp\left[-tr(\mathbf{V}\mathbf{\Sigma}^{-1})/2\right] \\ = & |\mathbf{\Sigma}|^{-(n+m+p+1)/2} \exp\{-\frac{1}{2} tr\left[(n\mathbf{S}+\mathbf{V})\,\mathbf{\Sigma}^{-1}\right]\} \end{aligned} \tag{8}$$

With known $\mathbf{V}$, m, n, and p. Comparing Equation 8 to Equation 3, we can easily find that $f_{\mathbf{\Sigma}|\mathbf{D}}$ is also an inverse Wishart distribution.

B. Algorithms

To sample from Equation 8, MH-MCMC is proposed. According to [3], the main steps of MH-MCMC are as follows:

1) Set initial state $\mathbf{\Sigma}^{(0)}$.
2) Sample from a proposal distribution to obtain new proposed state:

$$\vec{v}_{\mathbf{\Sigma}^{(k+1)-}} \sim q(\vec{v}_{\mathbf{\Sigma}^{(k+1)-}} | \vec{v}_{\mathbf{\Sigma}^{(k)}}) \tag{9}$$

where

$$\vec{v}_{\mathbf{\Sigma}^{(k)}} \triangleq \left[\Sigma_{11}^{(k)}, \Sigma_{12}^{(k)}, \Sigma_{22}^{(k)}\right]^T \tag{10}$$

Equation 10 is valid due to the fact that $\Sigma_{12}^{(k)} = \Sigma_{21}^{(k)}$.

3) Compute the acceptance probability of propagation from $\mathbf{\Sigma}^{(k)}$ to $\mathbf{\Sigma}^{(k+1)-}$:

$$\begin{aligned} & a(\mathbf{\Sigma}^{(k)}, \mathbf{\Sigma}^{(k+1)-}) = \\ & \min\left(1, \frac{f_{\mathbf{\Sigma}|\mathbf{D}}(\mathbf{\Sigma}^{(k+1)-})}{f_{\mathbf{\Sigma}|\mathbf{D}}(\mathbf{\Sigma}^{(k)})} \frac{q(\vec{v}_{\mathbf{\Sigma}^{(k)}} | \vec{v}_{\mathbf{\Sigma}^{(k+1)-}})}{q(\vec{v}_{\mathbf{\Sigma}^{(k+1)-}} | \vec{v}_{\mathbf{\Sigma}^{(k)}})}\right) \end{aligned} \tag{11}$$

4) Accept or reject proposed state:

$$\mathbf{\Sigma}^{(k+1)} = \begin{cases} \mathbf{\Sigma}^{(k+1)-} & \text{with probability } a(\mathbf{\Sigma}^{(k)}, \mathbf{\Sigma}^{(k+1)-}) \\ \mathbf{\Sigma}^{(k)} & \text{otherwise} \end{cases} \tag{12}$$

5) Repeat step 2) to step 4).

The random walk proposal is selected, which is defined as follows

$$\vec{v}_{\mathbf{\Sigma}^{(k+1)-}} = \vec{v}_{\mathbf{\Sigma}^{(k)}} + \xi,\ \ \xi \sim \mathcal{N}(0, \mathbf{\Gamma}) \tag{13}$$

The proposal distribution is:

$$\begin{aligned} & q(\vec{v}_{\mathbf{\Sigma}^{(k+1)-}} | \vec{v}_{\mathbf{\Sigma}^{(k)}}) \\ & = \frac{1}{2\pi^{3/2}|\Gamma|^{1/2}} \exp\left(-\frac{1}{2}\|\vec{v}_{\mathbf{\Sigma}^{(k+1)-}} - \vec{v}_{\mathbf{\Sigma}^{(k)}}\|_{\Gamma}^{2}\right) \end{aligned} \tag{14}$$

To improve the efficiency of the MH-MCMC, we implement Laplace approximation to approximate the posterior with a Gaussian distribution and generate an initial sample and the covariance. The Taylor expansion around some point $\vec{v}^{*}_{\mathbf{\Sigma}}$ so that any $\vec{v}_{\mathbf{\Sigma}} = \vec{v}^{*}_{\mathbf{\Sigma}} + \delta\vec{v}_{\mathbf{\Sigma}}$

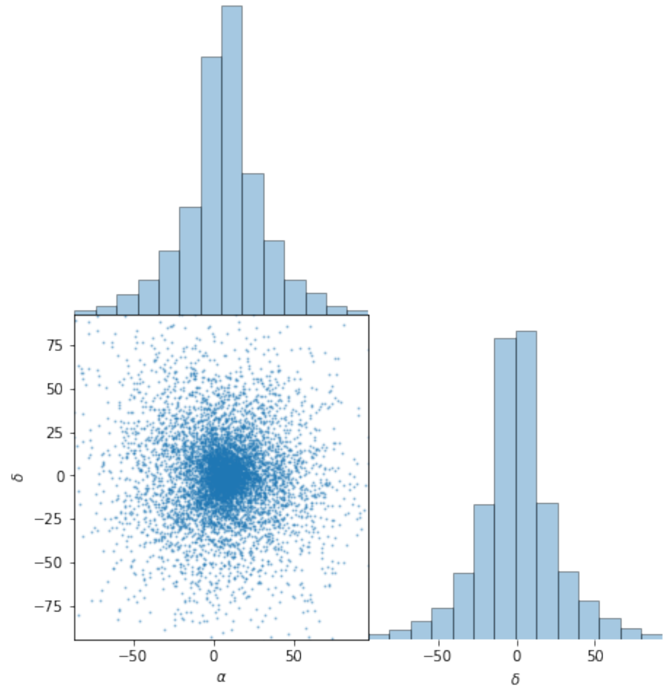

Fig. 1. The data ($\mathbf{M}\hat{X}'$ where $\mathbf{M}$ is the transformation matrix, see text) used in this project. 7439 data points are plotted.

$$\begin{aligned} \log f_{\mathbf{\Sigma}|\mathbf{D}}(\mathbf{\Sigma}) = \log f_{\mathbf{\Sigma}|\mathbf{D}}(\mathbf{\Sigma}^*) & + \nabla \log f_{\mathbf{\Sigma}|\mathbf{D}}(\mathbf{\Sigma})|_{\mathbf{\Sigma}=\mathbf{\Sigma}^*}\delta\mathbf{\Sigma} \\ & + \frac{1}{2}\Delta \log f_{\mathbf{\Sigma}|\mathbf{D}}(\mathbf{\Sigma})|_{\mathbf{\Sigma}=\mathbf{\Sigma}^*}\delta\mathbf{\Sigma}^2 \\ & + H.O.T. \end{aligned} \tag{15}$$

If we choose $\vec{v}^{*}_{\mathbf{\Sigma}} = \vec{v}^{MAP}_{\mathbf{\Sigma}}$, then the second term of the equation above vanishes and we obtain

$$\begin{aligned} \log f_{\mathbf{\Sigma}|\mathbf{D}}(\mathbf{\Sigma}) \approx & \log f_{\mathbf{\Sigma}|\mathbf{D}}(\mathbf{\Sigma}^{M.A.P}) \\ & + \frac{1}{2}\Delta \log f_{\mathbf{\Sigma}|\mathbf{D}}(\mathbf{\Sigma})(\mathbf{\Sigma} - \mathbf{\Sigma}^{M.A.P})^2 \end{aligned} \tag{16}$$

C. Data

The data used in this project are generated from the previous research. The boresight of the star tracker is carefully aligned to a star simulator (which generates skies with known coordinate X). Therefore, we can expect $\mathbb{E}(\hat{X}') = \mathbb{E}(\hat{X})$ and the dispersion is namely the contribution of the noise. The $\hat{X}$'s are transformed by $\mathbf{M}$ to the original point, i.e. $\mathbf{M}\hat{X} = 0$. The measurement data $\hat{X}'$ are transformed by the identical $\mathbf{M}$ as well. Then, we obtain the distribution of $\mathbf{M}\hat{X}'$ (Figure 1) and the corresponding heat map (Figure 2). In total, 7439 $\hat{X}'$'s are processed and used in this project.

III. Results and Discussions

A. Results

With the posterior function defined in Equation 8 and the proposal distribution defined in Equation 14, a 200000-sample

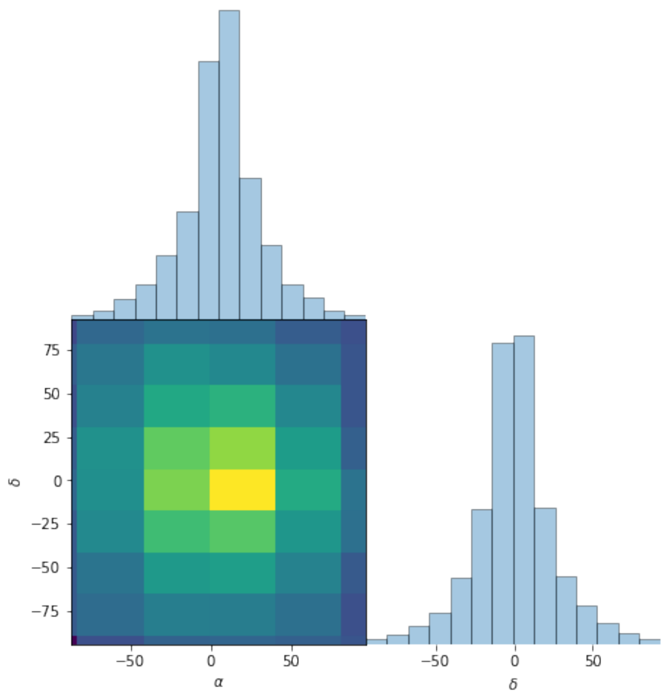

Fig. 2. The heat map of the data ($\mathbf{M}\hat{X}'$ where $\mathbf{M}$ is the transformation matrix, see text) used in this project.

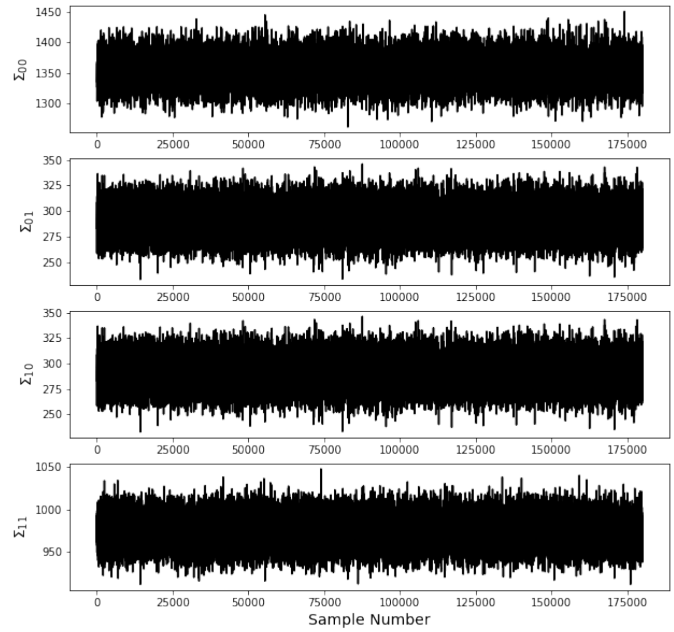

Fig. 3. The traceplot of all components in $\mathbf{\Sigma}$. A white noise like properties can be seen from each of the panels. The reader should also notify that the Σ_{01} and Σ_{10} are exactly identical.

TABLE I
THE PARAMETERS USED IN MCMC SIMULATION.

Parameter	Value
$\mathbf{V}$	$\mathbf{I}_{3\times 3}$
m	3
n	7439
p	2
$\mathbf{\Gamma}$	$500 \times \mathbf{I}_{3\times 3}$
$\vec{v}_{\mathbf{\Sigma}(0)}$	$[1353.16, 290.80, 973.02]^T$
Burn-in	20000

MH-MCMC is presented and the parameters used are listed in Table I. We should note that the first 20000 MCMC samples are removed to prevent the interference from these potential unstable samples.

The resulted covariance matrix simulated by MH-MCMC simulation is:

$$\mathbf{\Sigma}_{MCMC} = \begin{bmatrix} 1354.26 & 291.02 \\ 291.02 & 973.86 \end{bmatrix} \tag{17}$$

The Equation 17 implies that the one-σ errors of both directions of the star tracker are:

$$\begin{cases} \sqrt{1354.26} = 36.80\text{” in R.A direction} \\ \sqrt{973.86} = 31.21\text{” in Dec. direction} \end{cases} \tag{18}$$

B. Diagnosis

Now, we need to examine whether the simulation is in the good condition.

1) Visual inspection of mixing: First, whether the proposed samples have fully explored the posterior distribution and the mixing condition should be inspected. Figure 3 is the traceplot of the posterior space and each panels represent $\Sigma_{11}, \Sigma_{12}, \Sigma_{21}, \Sigma_{22}$ respectively. From this figure, the white noise like pattern indicates that the MCMC performs a good mixing and has generated samples moving around and covering the posterior.

2) Autocorrelations: The autocorrelation describes the correlation the between iterations with l (distance) apart. Autocorrelation converges to 0 when with larger l. A rapid convergence is desired such that two samples are more conditional independence at fixed l. Autocorrelation $R(l)$ is defined as

$$\begin{aligned} R(l) &= corr(\vec{v}_{\mathbf{\Sigma}(k)}, \vec{v}_{\mathbf{\Sigma}(k+1)}) \\ &= \frac{\sum_{i=0}^{n-l}(\vec{v}_{\mathbf{\Sigma}(i)} - \bar{\vec{v}}_{\mathbf{\Sigma}})(\vec{v}_{\mathbf{\Sigma}(i+1)} - \bar{\vec{v}}_{\mathbf{\Sigma}})}{\sum_{i=0}^{n}(\vec{v}_{\mathbf{\Sigma}(i)} - \bar{\vec{v}}_{\mathbf{\Sigma}})^2} \end{aligned} \tag{19}$$

To reduce computation time, we compute $l \in [0, 1000]$ cases using the fact that the autocorrelation converges to 0 at larger l. We can observe from Figure 4 that the autocorrelation of all components converge quickly to 0, which means that our proposal is preferable outcome.

3) Integrated autocorrelation values & Acceptance ratios: Now, we are here to assess the performance of the MCMC. The acceptance ratio is one of the hint of the performance. If the acceptance ratio is too low, the proposal need to be refined as the proposed points may be stocked at some points. The rule of thumb of a good proposal is that the acceptance ratio of a newly proposed sammple is around 20% - 30%.

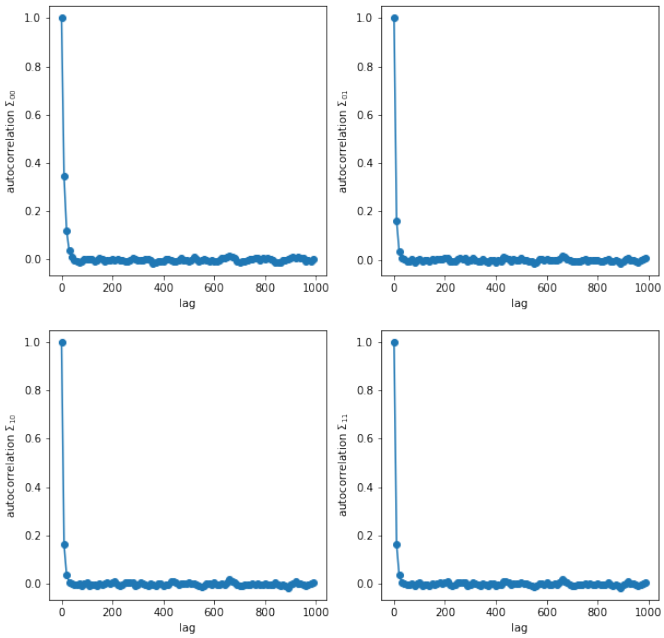

Fig. 4. The autocorrelation of $\Sigma_{11}, \Sigma_{12}, \Sigma_{21}, \Sigma_{22}$. We can see that the curves converge to 0 rapidly, which means that the sample of current step is not very correlated to the sample of few steps away.

TABLE II
TABLE OF ACCEPTANCE RATIO, IAC, AND ESS.

Acceptance ratio	0.30		
IAC Σ_{11}	3.81	ESS Σ_{11}	47193.31
IAC Σ_{12}	3.15	ESS Σ_{12}	57189.20
IAC Σ_{22}	3.44	ESS Σ_{22}	52353.52

The integrated autocorrelation (IAC) is defined as

$$IAC = 1 + 2\sum_{l=1}^{\infty} corr(f_{\mathbf{\Sigma}|\mathbf{D}}(\vec{v}_{\mathbf{\Sigma}^{(0)}}), f_{\mathbf{\Sigma}|\mathbf{D}}(\vec{v}_{\mathbf{\Sigma}^{(l)}})) = \theta \quad (20)$$

The greater the IAC, the more inefficient the MCMC is. We can define effective sample size (ESS):

$$ESS = \frac{n}{\theta} \quad (21)$$

The numbers of acceptance ration, IAC, and ESS of four algorithms are listed in Table II. We can see from Table II, the acceptance ratio of is 30%, meaning that the random walk used here is a good proposal. For IAC, we can observe that the IACs of all components in $\mathbf{\Sigma}$ are low, which means that the simulation is efficient according to the discussion of the previous paragraph and Section 16.1.3 of [3]. Low IACs result in high ESSs, which is straightforward according to Equation 21. Such high efficiency implies that the covariance matrix estimation can be achieved using relatively small amount of MCMC samples.

IV. CONCLUSIONS

In this project, we have presented a MCMC method to estimate the covariance matrix given a set of star tracker measurements. We begin establishing Bayesian inference posterior with inverse Wishart prior and the Gaussian likelihood. Then, we propose a random walk proposal to run MCMC simulation and obtain the estimated covariance matrix Σ. To ensure the quality of MCMC simulation, we perform several diagnoses including examining visual inspection of mixing, calculating autocorrelations, computing integrated autocorrelation values (IAC), and checking acceptance ratio. These diagnoses show that this MCMC method is effective and can estimate a covariance matrix using relatively small amount of samples.

REFERENCES

[1] C.-F. Chen, “Bayesian inference for a normal dispersion matrix and its application to stochastic multiple regression analysis,” *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 41, no. 2, pp. 235–248, 1979. [Online]. Available: http://www.jstor.org/stable/2985038

[2] Z. Zhang, “A note on wishart and inverse wishart priors for covariance matrix,” *Journal of Behavioral Data Science*, vol. 1, no. 2, p. 119–126, Aug. 2021. [Online]. Available: https://jbds.isdsa.org/index.php/jbds/article/view/19

[3] A. Gorodetsky, “Inference, estimation, and learning,” 2019.