Bayesian and Modern Statistics: Homework 9

Shinichiro Shirota *

September 27, 2013

I conduct missing data analysis via gibbs sampler and complete analysis. I generated true simulation data (100) from the bivariate normal distribution in the same settings of previous assignment. For the sampling parameters and missing data, I used gibbs sampling and set the burin in period (500) and samples (1000). I generate missing for $p = 0\%, 5\%, 10\%, 15\%, 20\%, 30\%, 40\%, 50\%$ and conduct missing data analysis. I also conducted the complete analysis for $p = 5\%, 50\%$ case by discarding the missing data and corresponding data. The estimation results show that even if the percentage of missing data is relatively high such as $40\%, 50\%$, 95% credible interval include true value and their variance is not so deteriorated. On the other hand, in the case of complete analysis (discarding the other vector data correspond to missing data), the variances of posterior samples are higher than those of gibbs sampling. Comparing to the full data, gibbs sampling approach for missing data show the similar results, of course, if the rate of missing is high, it will deteriorate the posterior variances.

Estimation procedures.

I conducted gibbs sampling with the following manner.

1 Set the initial value for μ, Σ and Y_{miss}

2 Generate μ and Σ form posterior distribution

Generation of μ and Σ can be conducted as the same manner of the case of no missing data because gibbs sampling of μ and Σ is conducted given the generated data which correspond to missing data.

3 Generate missing data

Given the μ and Σ, we generate the missing data as a latent variable from conditional posterior distribution which is based on the conditional distribution of Gaussian distribution.

4 Iterate 2,3 step by the chain converging.

*1st year graduate student, Department of Statistical Science, Duke University., US. E-mail:ss571@stat.duke.edu.

Estimation results.

Table 1: The statistics of posterior samples, $p = 0\%$ (full)

	Mean	SD	95%
μ_1	0.106	0.093	[-0.071, 0.289]
μ_2	0.062	0.095	[-0.139, 0.257]
Σ_{11}	0.899	0.123	[0.686, 1.176]
Σ_{12},Σ_{21}	0.690	0.114	[0.487, 0.941]
Σ_{22}	0.940	0.138	[0.702, 1.249]

Table 2: The statistics of posterior samples, $p = 5\%$

	Mean	SD	95%
μ_1	0.111	0.097	[-0.077, 0.290]
μ_2	0.064	0.097	[-0.124, 0.253]
Σ_{11}	0.904	0.125	[0.683, 1.170]
Σ_{12},Σ_{21}	0.697	0.113	[0.495, 0.935]
Σ_{22}	0.931	0.132	[0.692, 1.215]

Table 3: The statistics of posterior samples, $p = 10\%$

	Mean	SD	95%
μ_1	0.113	0.099	[-0.079, 0.296]
μ_2	0.064	0.098	[-0.128, 0.262]
Σ_{11}	0.903	0.130	[0.675, 1.193]
Σ_{12},Σ_{21}	0.687	0.117	[0.483, 0.934]
Σ_{22}	0.939	0.137	[0.709, 1.236]

Table 4: The statistics of posterior samples, $p = 15\%$

	Mean	SD	95%
μ_1	0.128	0.097	[-0.060, 0.321]
μ_2	0.059	0.095	[-0.131, 0.245]
Σ_{11}	0.912	0.129	[0.694, 1.201]
Σ_{12},Σ_{21}	0.679	0.113	[0.482, 0.927]
Σ_{22}	0.932	0.132	[0.697, 1.241]

Table 5: The statistics of posterior samples, $p = 20\%$

	Mean	SD	95%
μ_1	0.105	0.097	[-0.083, 0.299]
μ_2	0.070	0.102	[-0.121, 0.281]
Σ_{11}	0.891	0.126	[0.673, 1.184]
Σ_{12},Σ_{21}	0.670	0.118	[0.473, 0.940]
Σ_{22}	0.954	0.145	[0.696, 1.259]

Table 6: The statistics of posterior samples, $p = 30\%$

	Mean	SD	95%
μ_1	0.156	0.109	[-0.056, 0.368]
μ_2	0.051	0.097	[-0.137, 0.242]
Σ_{11}	0.973	0.143	[0.735, 1.275]
Σ_{12},Σ_{21}	0.671	0.119	[0.468, 0.940]
Σ_{22}	0.860	0.138	[0.628, 1.186]

The density plot of posterior samples of complete analysis is as follows.

Table 7: The statistics of posterior samples, $p = 40\%$

	Mean	SD	95%
μ_1	0.083	0.102	[-0.113, 0.280]
μ_2	0.123	0.107	[-0.088, 0.330]
Σ_{11}	0.880	0.138	[0.664, 1.186]
Σ_{12},Σ_{21}	0.614	0.125	[0.382, 0.873]
Σ_{22}	0.942	0.158	[0.687, 1.284]

Table 8: The statistics of posterior samples, $p = 50\%$

	Mean	SD	95%
μ_1	0.053	0.103	[-0.159, 0.250]
μ_2	0.041	0.106	[-0.162, 0.234]
Σ_{11}	0.933	0.148	[0.692, 1.282]
Σ_{12},Σ_{21}	0.674	0.126	[0.443, 0.940]
Σ_{22}	0.973	0.154	[0.708, 1.334]

Table 9: The statistics of posterior samples, complete analysis $p = 5\%$

	Mean	SD	95%
μ_1	0.132	0.095	[-0.045, 0.315]
μ_2	0.061	0.097	[-0.122, 0.244]
Σ_{11}	0.889	0.130	[0.663, 1.169]
Σ_{12},Σ_{21}	0.696	0.118	[0.492, 0.951]
Σ_{22}	0.941	0.140	[0.702, 1.259]

Table 10: The statistics of posterior samples, complete analysis $p = 50\%$

	Mean	SD	95%
μ_1	0.175	0.158	[-0.148, 0.495]
μ_2	0.064	0.170	[-0.272, 0.402]
Σ_{11}	0.995	0.224	[0.643, 1.540]
Σ_{12},Σ_{21}	0.795	0.212	[0.470, 1.285]
Σ_{22}	1.163	0.259	[0.779, 1.800]

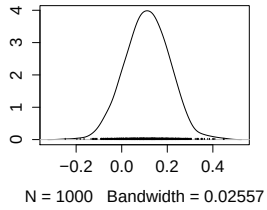

Figure 1: The density plot of posterior samples of θ_1, $p = 5\%$

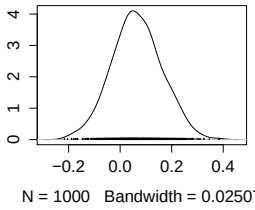

Figure 2: The density plot of posterior samples of θ_2, $p = 5\%$

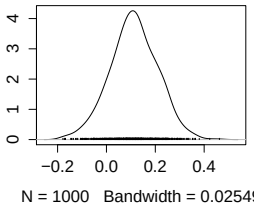

Figure 3: The density plot of posterior samples of θ_1, $p = 10\%$

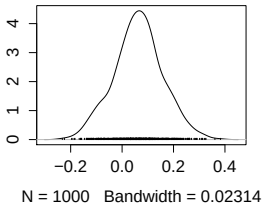

Figure 4: The density plot of posterior samples of θ_2, $p = 10\%$

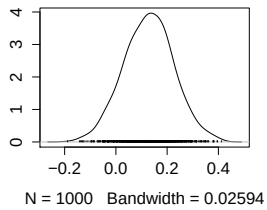

Figure 5: The density plot of posterior samples of θ_1, $p = 15\%$

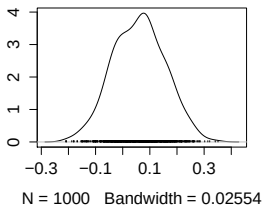

Figure 6: The density plot of posterior samples of θ_2, $p = 15\%$

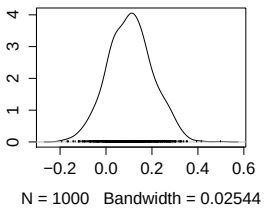

Figure 7: The density plot of posterior samples of θ_1, $p = 20\%$

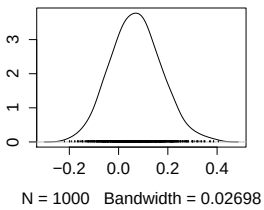

Figure 8: The density plot of posterior samples of θ_2, $p = 20\%$

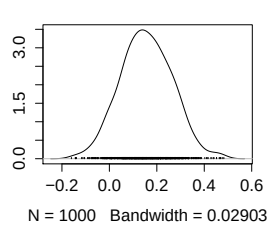

Figure 9: The density plot of posterior samples of θ_1, $p = 30\%$

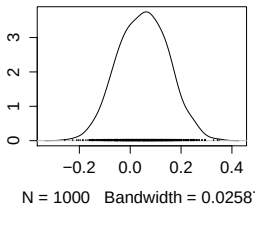

Figure 10: The density plot of posterior samples of θ_2, $p = 30\%$

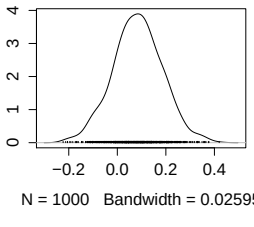

Figure 11: The density plot of posterior samples of θ_1, $p = 40\%$

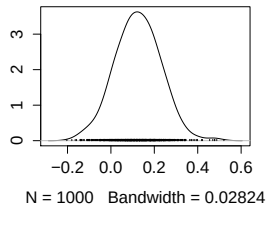

Figure 12: The density plot of posterior samples of θ_2, $p = 40\%$

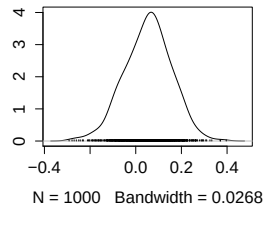

Figure 13: The density plot of posterior samples of θ_1, $p = 50\%$

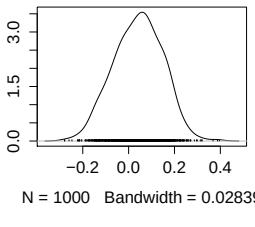

Figure 14: The density plot of posterior samples of θ_2, $p = 50\%$

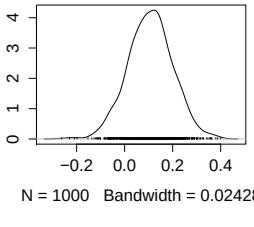

Figure 15: The density plot of posterior samples of full data, θ_1, $p = 0\%$

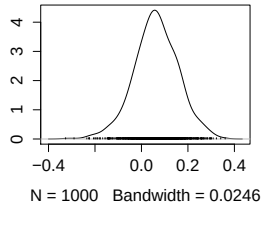

Figure 16: The density plot of posterior samples of full data, θ_2, $p = 0\%$

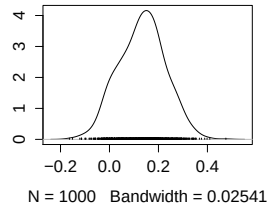

Figure 17: The density plot of posterior samples of complete analysis, θ_1, $p = 5\%$

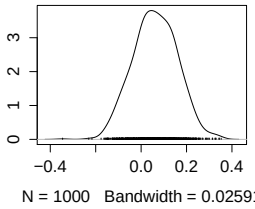

Figure 18: The density plot of posterior samples of complete analysis, θ_2, $p = 5\%$

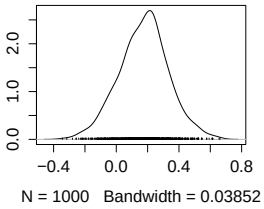

Figure 19: The density plot of posterior samples of complete analysis, θ_1, $p = 50\%$

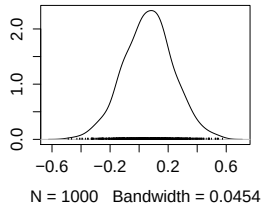

Figure 20: The density plot of posterior samples of complete analysis, θ_2, $p = 50\%$

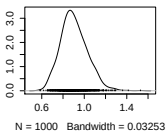

Figure 21: The density plot of posterior samples of Σ_{11}, $p = 5\%$

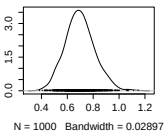

Figure 22: The density plot of posterior samples of Σ_{12}, $p = 5\%$

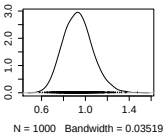

Figure 23: The density plot of posterior samples of Σ_{22}, $p = 5\%$

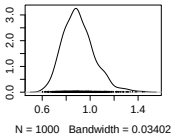

Figure 24: The density plot of posterior samples of Σ_{11}, $p = 10\%$

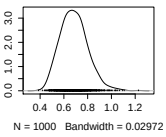

Figure 25: The density plot of posterior samples of Σ_{12}, $p = 10\%$

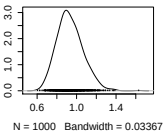

Figure 26: The density plot of posterior samples of Σ_{22}, $p = 10\%$

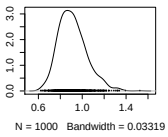

Figure 27: The density plot of posterior samples of Σ_{11}, $p = 15\%$

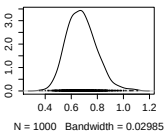

Figure 28: The density plot of posterior samples of Σ_{12}, $p = 15\%$

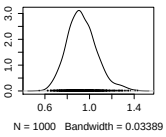

Figure 29: The density plot of posterior samples of Σ_{22}, $p = 15\%$

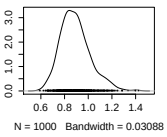

Figure 30: The density plot of posterior samples of Σ_{11}, $p = 20\%$

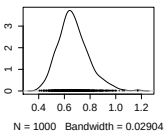

Figure 31: The density plot of posterior samples of Σ_{12}, $p = 20\%$

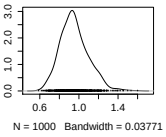

Figure 32: The density plot of posterior samples of Σ_{22}, $p = 20\%$

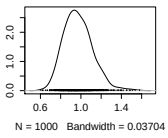

Figure 33: The density plot of posterior samples of Σ_{11}, $p = 30\%$

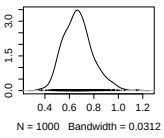

Figure 34: The density plot of posterior samples of Σ_{12}, $p = 30\%$

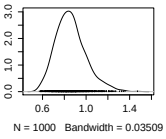

Figure 35: The density plot of posterior samples of Σ_{22}, $p = 30\%$

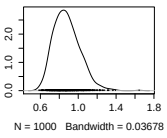

Figure 36: The density plot of posterior samples of Σ_{11}, $p = 40\%$

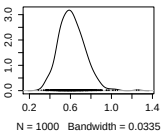

Figure 37: The density plot of posterior samples of Σ_{12}, $p = 40\%$

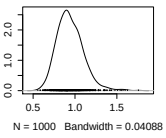

Figure 38: The density plot of posterior samples of Σ_{22}, $p = 40\%$

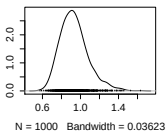

Figure 39: The density plot of posterior samples of Σ_{11}, $p = 50\%$

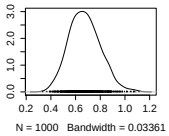

Figure 40: The density plot of posterior samples of Σ_{12}, $p = 50\%$

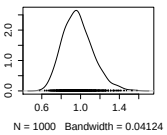

Figure 41: The density plot of posterior samples of Σ_{22}, $p = 50\%$

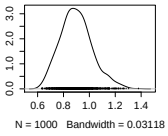

Figure 42: The density plot of posterior samples of full data, Σ_{11}, $p = 0\%$

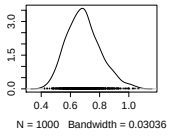

Figure 43: The density plot of posterior samples of full data, Σ_{12}, $p = 0\%$

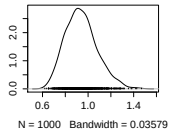

Figure 44: The density plot of posterior samples of full data, Σ_{22}, $p = 0\%$

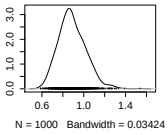

Figure 45: The density plot of posterior samples of complete analysis, Σ_{11}, $p = 5\%$

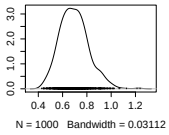

Figure 46: The density plot of posterior samples of complete analysis, Σ_{12} and Σ_{21}, $p = 5\%$

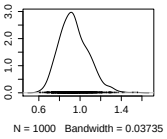

Figure 47: The density plot of posterior samples of complete analysis, Σ_{22}, $p = 5\%$

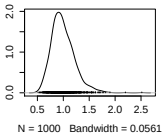

Figure 48: The density plot of posterior samples of complete analysis, Σ_{11}, $p = 50\%$

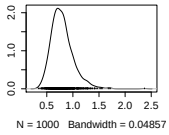

Figure 49: The density plot of posterior samples of complete analysis, Σ_{12} and Σ_{21}, $p = 50\%$

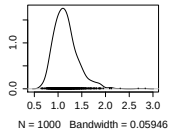

Figure 50: The density plot of posterior samples of complete analysis, Σ_{22}, $p = 50\%$