ÖREBRO UNIVERSITY

Multivariate meta-analysis based on generalized random effects model

Olha Bodnar

(with Taras Bodnar)

3. November 2022, MATHMET, Paris

Multivariate random effects meta-analysis

$$\mathbf{x}_i = \boldsymbol{\mu} + \boldsymbol{\lambda}_i + \boldsymbol{\varepsilon}_i \quad \text{for} \quad i = 1, ..., n$$

- $\boldsymbol{\mu}$ – common mean vector, main parameter of the model
- $\boldsymbol{\lambda}_i \sim \mathcal{N}_p(\mathbf{0}, \boldsymbol{\Psi})$ – random effects
- $\boldsymbol{\Psi}$ – heterogeneity matrix, corresponds to the dark uncertainty
- $\boldsymbol{\varepsilon}_i \sim \mathcal{N}_p(\mathbf{0}, \mathbf{U}_i)$
- $\mathbf{U}_i$, $i = 1, ..., n$ – known covariance matrices

Gasparrini, A., Armstrong, B., and Kenward, M. (2012). Multivariate meta-analysis for non-linear and other multi-parameter associations. *Statistics in Medicine*, **31**:3821-3839.

Jackson, D. and Riley, R. D. (2014). A refined method for multivariate meta-analysis and meta-regression. *Statistics in Medicine*, **33**:541-554.

Jackson, D., White, I. R., and Thompson, S. G. (2010). Extending DerSimonian and Laird's methodology to perform multivariate random effects meta-analyses. *Statistics in Medicine*, **29**:1282-1297.

Liu, D., Liu, R. Y., and Xie, M. (2015). Multivariate meta-analysis of heterogeneous studies using only summary statistics: efficiency and robustness. *Journal of the American Statistical Association*, **110**:326-340.

Generalized multivariate random effects model

$$p(\mathbf{X}|\boldsymbol{\mu},\boldsymbol{\Psi}) = \frac{1}{\sqrt{\det(\boldsymbol{\Psi}\otimes\mathbf{I}+\mathbf{U})}} f\left(\mathrm{vec}(\mathbf{X}-\boldsymbol{\mu}\mathbf{1}^\top)^\top(\boldsymbol{\Psi}\otimes\mathbf{I}+\mathbf{U})^{-1}\mathrm{vec}(\mathbf{X}-\boldsymbol{\mu}\mathbf{1})\right)$$

$$\Longleftrightarrow \mathbf{X}|\boldsymbol{\mu},\boldsymbol{\Psi} \sim E_{p,n}(\boldsymbol{\mu}\mathbf{1}^\top, \boldsymbol{\Psi}\otimes\mathbf{I}+\mathbf{U}, f)$$

- $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_n]$ – observation matrix
- $\boldsymbol{\mu}\mathbf{1}^\top$ – location matrix with $\mathbf{1} = (\underbrace{1, ..., 1}_{n})^\top$
- $\boldsymbol{\Psi}\otimes\mathbf{I}+\mathbf{U}$ – dispersion matrix with $\mathbf{U} = \mathrm{diag}(\mathbf{U}_1, ..., \mathbf{U}_n)$
- f – density generator such that $\int_0^\infty t^{pn-1} f(t^2)dt < \infty$

Aim: to infer $\boldsymbol{\theta} = (\boldsymbol{\mu}, \boldsymbol{\Psi})$ given $\mathbf{X}$.

Noninformative priors

- Constant (diffuse) prior (Laplace, 1812)

$$\pi(\boldsymbol{\theta}) \propto 1$$

  - not invariant under re-parametrization

- Jeffrey's prior (Jeffrey, 1946, 1961)

$$\pi(\boldsymbol{\theta}) \propto \sqrt{\det(\mathbf{F}(\boldsymbol{\theta}))} \qquad \text{with} \qquad \mathbf{F}(\boldsymbol{\theta}) \text{ – Fisher information matrix}$$

  - invariant, usually good for one-group parameter problems, often difficulties for multi-group parameter problems

- Berger and Bernardo reference prior
  - One-group parameter case (Bernardo 1979): maximizes Shannon's mutual information

$$I(\mathbf{X}, \pi(\theta_1)) = E_{\mathbf{X}}\left(\log \frac{\pi(\theta_1|\mathbf{X})}{\pi(\theta_1)}\right)$$

  - Multi-group parameter case (Berger & Bernardo 1992)

Normal multivariate random effects model

Density generator

$$f(u) = K_{p,n} \exp(-u/2) \quad \text{with} \quad K_{p,n} = (2\pi)^{-pn/2},$$

- Conditional posterior for $\boldsymbol{\mu}$ given $\boldsymbol{\Psi}$

$$\boldsymbol{\mu}|\boldsymbol{\Psi}, \mathbf{X} \sim \mathcal{N}\left(\left(\sum_{i=1}^{n}(\boldsymbol{\Psi}+\mathbf{U}_i)^{-1}\right)^{-1}\sum_{i=1}^{n}(\boldsymbol{\Psi}+\mathbf{U}_i)^{-1}\mathbf{x}_i, \left(\sum_{i=1}^{n}(\boldsymbol{\Psi}+\mathbf{U}_i)^{-1}\right)^{-1}\right)$$

- Marginal posterior for $\boldsymbol{\Psi}$

$$\begin{aligned}\pi(\boldsymbol{\Psi}|\mathbf{X}) \quad &\propto \quad \frac{\pi(\boldsymbol{\Psi})}{\sqrt{\det(\sum_{i=1}^{n}(\boldsymbol{\Psi}+\mathbf{U}_i)^{-1})}\prod_{i=1}^{n}\sqrt{\det(\boldsymbol{\Psi}+\mathbf{U}_i)}} \\ &\times \quad \exp\left(-\frac{1}{2}\sum_{i=1}^{n}(\mathbf{x}_i-\tilde{\mathbf{x}}(\boldsymbol{\Psi}))^\top(\boldsymbol{\Psi}+\mathbf{U}_i)^{-1}(\mathbf{x}_i-\tilde{\mathbf{x}}(\boldsymbol{\Psi}))\right)\end{aligned}$$

t multivariate random effects model

Density generator

$$f(u) = K_{p,n,d}(1+u/d)^{-(pn+d)/2} \quad \text{with} \quad K_{p,n,d} = (\pi d)^{-pn/2}\frac{\Gamma\left((d+pn)/2\right)}{\Gamma\left(d/2\right)}$$

- Conditional posterior of $\boldsymbol{\mu}$ given $\boldsymbol{\Psi}$

$$\begin{aligned}\pi(\boldsymbol{\mu}|\boldsymbol{\Psi},\mathbf{X}) \quad &\propto \quad \left(1+\frac{1}{pn+d-p}\frac{pn+d-p}{d+\sum_{i=1}^{n}(\mathbf{x}_i-\tilde{\mathbf{x}}(\boldsymbol{\Psi}))^\top(\boldsymbol{\Psi}+\mathbf{U}_i)^{-1}(\mathbf{x}_i-\tilde{\mathbf{x}}(\boldsymbol{\Psi}))}\right.\\ &\times \quad \left.(\boldsymbol{\mu}-\tilde{\mathbf{x}}(\boldsymbol{\Psi}))^\top\left(\sum_{i=1}^{n}(\boldsymbol{\Psi}+\mathbf{U}_i)^{-1}\right)(\boldsymbol{\mu}-\tilde{\mathbf{x}}(\boldsymbol{\Psi}))\right)^{-(pn+d)/2}\end{aligned}$$

- Marginal posterior for $\boldsymbol{\Psi}$

$$\begin{aligned}\pi(\boldsymbol{\Psi}|\mathbf{X}) \quad &\propto \quad \frac{\pi(\boldsymbol{\Psi})}{\sqrt{\det(\sum_{i=1}^{n}(\boldsymbol{\Psi}+\mathbf{U}_i)^{-1})}\prod_{i=1}^{n}\sqrt{\det(\boldsymbol{\Psi}+\mathbf{U}_i)}}\\ &\times \quad \left(1+\frac{1}{d}\sum_{i=1}^{n}(\mathbf{x}_i-\tilde{\mathbf{x}}(\boldsymbol{\Psi}))^\top(\boldsymbol{\Psi}+\mathbf{U}_i)^{-1}(\mathbf{x}_i-\tilde{\mathbf{x}}(\boldsymbol{\Psi}))\right)^{-(pn+d)/2}\end{aligned}$$

Empirical illustration

Study	$X_{i;1}$ (SBP)	$X_{i;2}$ (DBP)	$\sqrt{U_{i;11}}$ (SBP)	$\rho_{i;12} = \frac{U_{i;12}}{\sqrt{U_{i;11} U_{i;22}}}$	$\sqrt{U_{i;22}}$ (DBP)
1	-6.66	-2.99	0.72	0.78	0.27
2	-14.17	-7.87	4.73	0.45	1.44
3	-12.88	-6.01	10.31	0.59	1.77
4	-8.71	-5.11	0.30	0.77	0.10
5	-8.70	-4.64	0.14	0.66	0.05
6	-10.60	-5.56	0.58	0.49	0.18
7	-11.36	-3.98	0.30	0.50	0.27
8	-17.93	-6.54	5.82	0.61	1.31
9	-6.55	-2.08	0.41	0.45	0.11
10	-10.26	-3.49	0.20	0.51	0.04

Data collected in 10 studies about the effectiveness of hypertension treatment with the aim to reduce blood pressure. $X_{i;1}$ and $X_{i;2}$ denote the treatment effects on the systolic blood pressure (SBP) and the diastolic blood pressures (DBP) from the ith study, while $\mathbf{U}_i = (U_{i;lj})_{lj=1,2}$ is the corresponding covariance matrix.

Results of Bayesian estimation

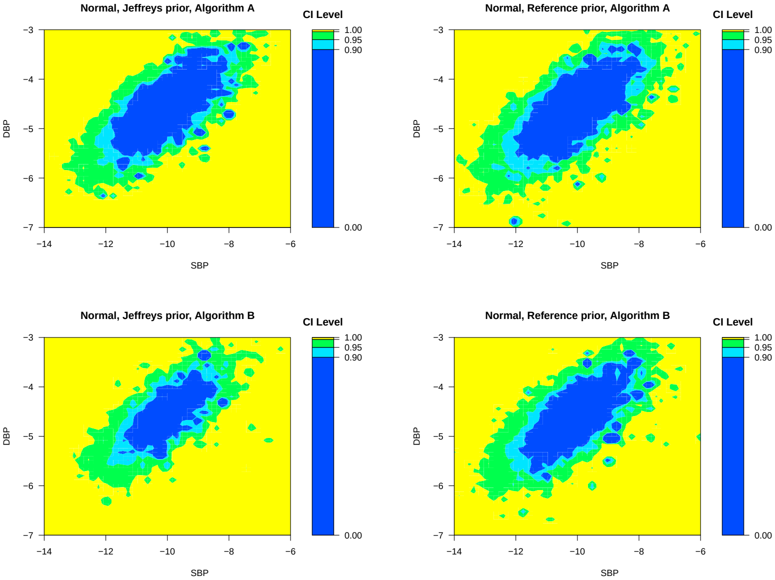
	Normal random effects model		t random effects model	
	μ_1 (SBP)	μ_2 (DBP)	μ_1 (SBP)	μ_2 (DBP)
Jeffreys prior, Algorithm A				
post. mean	-9.79	-4.05	-10.15	-4.67
post. median	-9.60	-4.27	-10.10	-4.66
post. sd.	0.88	0.93	1.08	0.60
cred. inter.	[-11.73, -8.00]	[-5.61, -2.66]	[-12.44,-8.11]	[-5.87,-3.51]
Jeffreys prior, Algorithm B				
post. mean	-9.78	-4.37	-10.03	-4.66
post. median	-9.84	-4.37	-9.97	-4.65
post. sd.	0.74	0.50	1.13	0.61
cred. inter.	[-11.46, -8.39]	[-5.38, -3.38]	[-12.46, -7.99]	[-5.90, -3.51]
Berger and Bernardo reference prior, Algorithm A				
post. mean	-9.81	-4.49	-10.11	-4.67
post. median	-9.87	-4.44	-10.06	-4.66
post. sd.	1.04	0.59	1.16	0.64
cred. inter.	[-12.06, -8.00]	[-5.78, -3.42]	[-12.58, -7.97]	[-5.96, -3.43]
Berger and Bernardo reference prior, Algorithm B				
post. mean	-9.70	-4.51	-10.08	-4.68
post. median	-9.72	-4.53	-10.03	-4.65
post. sd.	1.01	0.58	1.13	0.64
cred. inter.	[-11.88, -8.06]	[-5.67, -3.49]	[-12.50, -7.97]	[-5.94, -3.42]

Maximum likelihood, Gasparrini et al. (2012)		
estimator	-9.47	-4.41
stand. error	0.68	0.44
cred. inter.	[-10.79, -8.14]	[-5.26, -3.55]
Restrictive max. likelihood, Gasparrini et al. (2012)		
estimator	-9.51	-4.43
stand. error	0.73	0.47
cred. inter.	[-10.95, -8.07]	[-5.35, -3.51]
Method of moments, Jackson et al. (2013)		
estimator	-9.17	-4.31
stand. error	0.55	0.36
cred. inter.	[-10.26, -8.08]	[-5.02, -3.60]

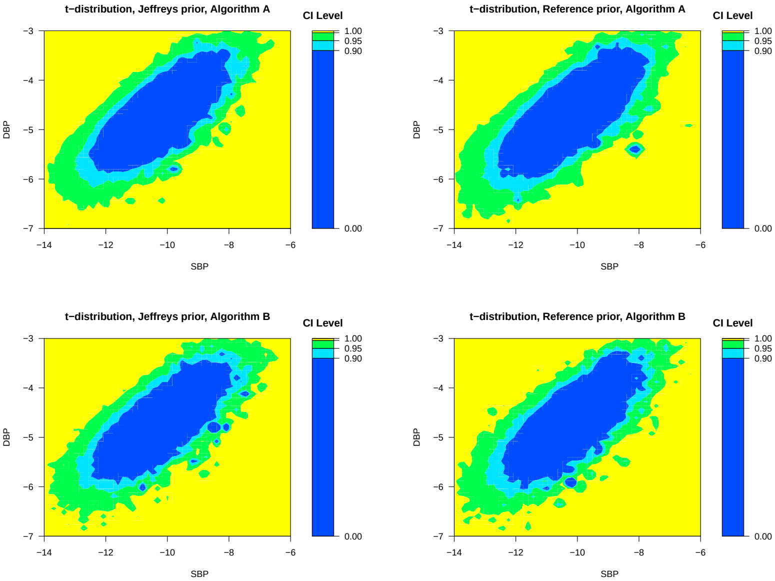
Gasparrini, A., Armstrong, B., and Kenward, M. (2012). Multivariate meta-analysis for non-linear and other multi-parameter associations. *Statistics in Medicine*, **31**:3821-3839.

Jackson, D., White, I. R., and Riley, R. D. (2013). A matrix-based method of moments for fitting the multivariate random effects model for meta-analysis and meta-regression. *Biometrical Journal*, **55**:231-245.

Credible sets for μ_1 (SBP) and μ_2 (DBP): normal multivariate REM



Credible sets for μ_1 (SBP) and μ_2 (DBP): t multivariate REM



Convergence diagnostic for the constructed Markov chains: Split-$\hat{R}$ estimates based on the rank normalization in Vehtari et al. (2021)

	μ_1 (SBP)	μ_2 (DBP)	ψ_{11} (SBP)	ψ_{21}	ψ_{22} (DBP)
Jeffreys prior, Algorithm A					
normal	1.006	1.010	1.007	1.002	1.003
t-dist.	1.003	1.004	1.005	1.001	1.004
Jeffreys prior, Algorithm B					
normal	1.078	1.044	1.025	1.017	1.026
t-dist.	1.010	1.008	1.006	1.007	1.005
Berger and Bernardo reference prior, Algorithm A					
normal	1.006	1.013	1.095	1.068	1.108
t-dist.	1.049	1.026	1.021	1.017	1.033
Berger and Bernardo reference prior, Algorithm B					
normal	1.020	1.031	1.046	1.020	1.031
t-dist.	1.014	1.010	1.051	1.032	1.023

Vehtari, A., Gelman, A., Simpson, D., Carpenter, B., and Bürkner, P.-C. (2021). Rank-normalization, folding, and localization: An improved $\hat{R}$ for assessing convergence of MCMC (with Discussion). *Bayesian Analysis*, **16**: 667-718

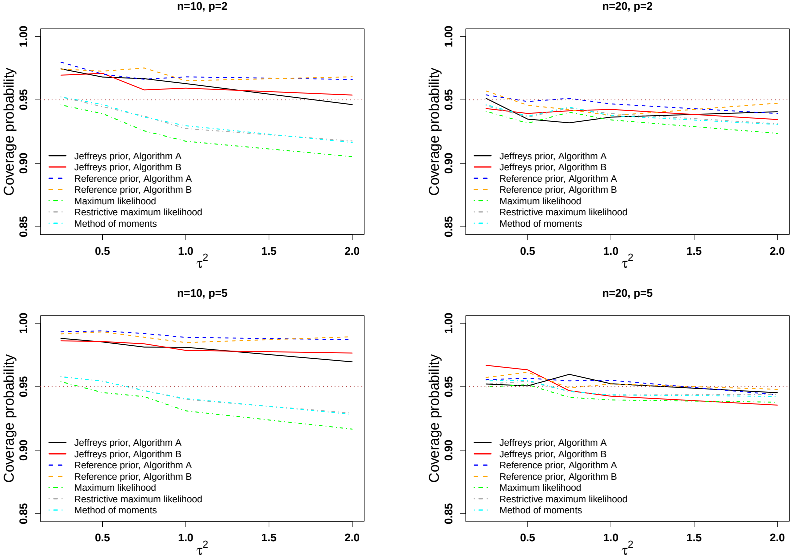
Setup of the simulation study

- $\mathbf{X}$ is drawn from the normal multivariate random effects model with the same $\boldsymbol{\mu}$, $\mathbf{\Psi} = \tau^2 \mathbf{\Xi}$, and $\mathbf{U} = diag(\mathbf{U}_1, ..., \mathbf{U}_p)$
- The elements of $\boldsymbol{\mu}$ are generated $Unif[1, 5]$
- The eigenvalues of $\mathbf{\Xi}$, $\mathbf{U}_1$, ... , $\mathbf{U}_{p-1}$, and $\mathbf{U}_p$ are generated from $Unif[1, 4]$
- The eigenvectors of $\mathbf{\Xi}$, $\mathbf{U}_1$, ... , $\mathbf{U}_{p-1}$, and $\mathbf{U}_p$ are generated from the Haar distribution
- $p \in \{2, 5\}$, $n \in \{10, 20\}$, and $\tau^2 \in \{0.25, 0.5, 0.75, 1, 2\}$

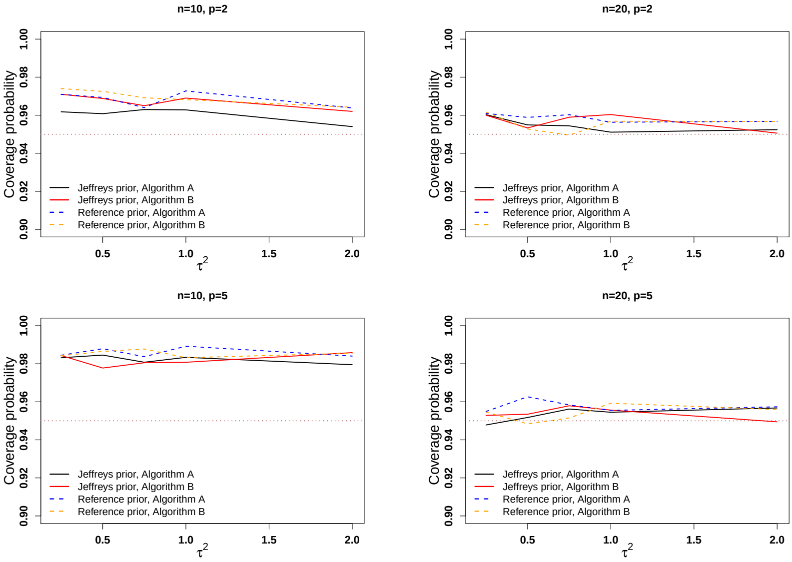
Performance measure:

empirical coverage probability of the credible interval constructed for μ_1, based on 5000 independent repetitions

Coverage probabilities of the 95% credible intervals for μ_1



Coverage probabilities of the 95% credible intervals for μ_1



Summary

- An objective Bayesian inference is proposed for the generalized multivariate random effects model.
- Both the Jeffreys and the Berger and Bernardo reference priors and corresponding posteriors are derived.
- The results are applied to data consisting of results obtained in ten studies that assess the effectiveness of hypertension treatment for reducing blood pressure.
- Coverage probabilities are investigated by simulations.

Thank you very much for your attention

1. **Initialization:** Choose the initial values $\boldsymbol{\mu}^{(0)}$ and $\boldsymbol{\Psi}^{(0)}$ for $\boldsymbol{\mu}$ and $\boldsymbol{\Psi}$ and set $b = 0$.

2. **Generating new values of $\boldsymbol{\mu}^{(w)}$ and $\boldsymbol{\Psi}^{(w)}$ from the proposal:**

Algorithm A:

(i) For given data $\mathbf{X} = (\mathbf{x}_1, \ldots, \mathbf{x}_n)$, generate $\boldsymbol{\mu}^{(w)}$ from $\boldsymbol{\mu}|\mathbf{X} \sim t\left(n-p, \bar{\mathbf{x}}, \dfrac{(n-1)\mathbf{S}}{n(n-p)}\right)$;

(ii) Using data $\mathbf{X}$ and the drawn in step (i) $\boldsymbol{\mu}^{(w)}$, generate $\boldsymbol{\Psi}^{(w)}$ from

$$\boldsymbol{\Psi}|\boldsymbol{\mu}^{(w)}, \mathbf{X} \sim IW_p\left(n+p+1, \sum_{i=1}^{n}(\mathbf{x}_i - \boldsymbol{\mu}^{(w)})(\mathbf{x}_i - \boldsymbol{\mu}^{(w)})^\top\right).$$

Algorithm B:

(i) For given data $\mathbf{X} = (\mathbf{x}_1, \ldots, \mathbf{x}_n)$, generate $\boldsymbol{\Psi}^{(w)}$ from $\boldsymbol{\Psi}|\mathbf{X} \sim IW_p(n+p, (n-1)\mathbf{S})$;

(ii) Using data $\mathbf{X}$ and the drawn in step (i) $\boldsymbol{\Psi}^{(w)}$, generate $\boldsymbol{\mu}^{(w)}$ from

$$\boldsymbol{\mu}|\boldsymbol{\Psi}^{(w)}, \mathbf{X} \sim N_p\left(\bar{\mathbf{x}}, \frac{\boldsymbol{\Psi}^{(w)}}{n}\right).$$

3. **Computation of the Metropolis-Hastings ratio:**

$$MH^{(b)} = \frac{\pi(\boldsymbol{\mu}^{(w)}, \boldsymbol{\Psi}^{(w)}|\mathbf{X})q_R(\boldsymbol{\mu}^{(b-1)}, \boldsymbol{\Psi}^{(b-1)}|\mathbf{X})}{\pi(\boldsymbol{\mu}^{(b-1)}, \boldsymbol{\Psi}^{(b-1)}|\mathbf{X})q_R(\boldsymbol{\mu}^{(w)}, \boldsymbol{\Psi}^{(w)}|\mathbf{X})}.$$

4. **Moving to the next state of the Markov chain:**
   (i) Generate $U^{(b)}$ from the uniform distribution on $[0,1]$;
   (ii) If $U^b < \min\left\{1, MH^{(b)}\right\} \pi(\boldsymbol{\mu}^{(b)}$, then set $\boldsymbol{\mu}^{(b)} = \boldsymbol{\mu}^{(w)}$ and $\boldsymbol{\Psi}^{(b)} = \boldsymbol{\Psi}^{(w)}$ (Markov chain moves to the new state). Otherwise, set $\boldsymbol{\mu}^{(b)} = \boldsymbol{\mu}^{(b-1)}$ and $\boldsymbol{\Psi}^{(b)} = \boldsymbol{\Psi}^{(b-1)}$ (Markov chain stays in the previous state).

5. Return to step (2), increase b by 1, and repeat until the sample of size B is accumulated.

Theorem 1 (Fisher information matrix)

The Fisher information matrix is given by

$$\mathbf{F} = \begin{pmatrix} \mathbf{F}_{11} & \mathbf{O} \\ \mathbf{O} & \mathbf{F}_{22} \end{pmatrix} \quad \text{with}$$

$$\begin{aligned}
\mathbf{F}_{11} &= \frac{4J_1}{pn}\sum_{i=1}^{n}(\boldsymbol{\Psi}+\mathbf{U}_i)^{-1}, \\
\mathbf{F}_{22} &= \mathbf{G}_p^\top \Bigg[\left(\frac{J_2}{2pn+p^2n^2}-\frac{1}{4}\right)\text{vec}\left(\sum_{i=1}^{n}(\boldsymbol{\Psi}+\mathbf{U}_i)^{-1}\right)\text{vec}\left(\sum_{j=1}^{n}(\boldsymbol{\Psi}+\mathbf{U}_j)^{-1}\right)^\top \\
&+ \frac{2J_2}{2pn+p^2n^2}\sum_{i=1}^{n}\left((\boldsymbol{\Psi}+\mathbf{U}_i)^{-1}\otimes(\boldsymbol{\Psi}+\mathbf{U}_i)^{-1}\right)\Bigg]\mathbf{G}_p
\end{aligned}$$

where

$$J_i = \mathrm{E}\left((R^2)^i\left(\frac{f'(R^2)}{f(R^2)}\right)^2\right) \quad \text{with} \quad R^2=\text{vec}(\mathbf{Z})^\top\text{vec}(\mathbf{Z}), \quad \mathbf{Z}\sim E_{p,n}(\mathbf{O}_{p,n},\mathbf{I}_{p\times n},f)$$

and $\mathbf{G}_p$ stands for the duplication matrix.

Jeffreys and reference priors

Theorem 2

1. Berger and Bernardo reference prior: $\pi_R(\boldsymbol{\mu}, \boldsymbol{\Psi}) = \pi_R(\boldsymbol{\Psi}) \propto \sqrt{\det(\mathbf{F}_{22})}$
2. Jeffreys prior: $\pi_J(\boldsymbol{\mu}, \boldsymbol{\Psi}) = \pi_J(\boldsymbol{\Psi}) \propto \sqrt{\det(\mathbf{F})} = \sqrt{\det(\mathbf{F}_{11})}\sqrt{\det(\mathbf{F}_{22})}$

Corollary 1

For the multivariate normal random effects model (i.e. $f(u) = \exp(-u/2)/(2\pi)^{n/2}$)

1. Berger and Bernardo reference prior:

$$\pi_R(\boldsymbol{\mu}, \boldsymbol{\Psi}) = \pi_R(\boldsymbol{\Psi}) \propto \sqrt{\det\left(\mathbf{G}_p^\top \left[\sum_{i=1}^{n} ((\boldsymbol{\Psi} + \mathbf{U}_i)^{-1} \otimes (\boldsymbol{\Psi} + \mathbf{U}_i)^{-1})\right] \mathbf{G}_p\right)},$$

2. Jeffreys prior:

$$\pi_J(\boldsymbol{\mu}, \boldsymbol{\Psi}) = \pi_J(\boldsymbol{\Psi}) \propto \pi_R(\boldsymbol{\Psi})\sqrt{\det\left(\sum_{i=1}^{n} (\boldsymbol{\Psi} + \mathbf{U}_i)^{-1}\right)}.$$

Jeffreys and reference priors

Theorem 2

1. Berger and Bernardo reference prior: $\pi_R(\boldsymbol{\mu}, \boldsymbol{\Psi}) = \pi_R(\boldsymbol{\Psi}) \propto \sqrt{\det(\mathbf{F}_{22})}$
2. Jeffreys prior: $\pi_J(\boldsymbol{\mu}, \boldsymbol{\Psi}) = \pi_J(\boldsymbol{\Psi}) \propto \sqrt{\det(\mathbf{F})} = \sqrt{\det(\mathbf{F}_{11})}\sqrt{\det(\mathbf{F}_{22})}$

Corollary 2

Assume that $\mathbf{U}_1 = ... = \mathbf{U}_n = \mathbf{V}$. Then

1. Berger and Bernardo reference prior:
$$\pi_R(\boldsymbol{\mu}, \boldsymbol{\Psi}) = \pi_R(\boldsymbol{\Psi}) \propto \det\left(\boldsymbol{\Psi} + \mathbf{V}\right)^{-(p+1)/2},$$
2. Jeffreys prior:
$$\pi_J(\boldsymbol{\mu}, \boldsymbol{\Psi}) = \pi_J(\boldsymbol{\Psi}) \propto \det\left(\boldsymbol{\Psi} + \mathbf{V}\right)^{-(p+2)/2}.$$

Conditional posterior

Theorem 3

The conditional posterior $\pi(\boldsymbol{\mu}|\boldsymbol{\Psi}, \mathbf{X})$ is given by

$$\pi(\boldsymbol{\mu}|\boldsymbol{\Psi}, \mathbf{X}) \quad \propto \quad f_{\boldsymbol{\Psi},\mathbf{X}}\left((\boldsymbol{\mu} - \tilde{\mathbf{x}}(\boldsymbol{\Psi}))^\top \left(\sum_{i=1}^{n} (\boldsymbol{\Psi} + \mathbf{U}_i)^{-1} \right) (\boldsymbol{\mu} - \tilde{\mathbf{x}}(\boldsymbol{\Psi})) \right) ,$$

where

$$f_{\boldsymbol{\Psi},\mathbf{X}}(u) = f\left(\sum_{i=1}^{n} (\mathbf{x}_i - \tilde{\mathbf{x}}(\boldsymbol{\Psi}))^\top (\boldsymbol{\Psi} + \mathbf{U}_i)^{-1} (\mathbf{x}_i - \tilde{\mathbf{x}}(\boldsymbol{\Psi})) + u \right) \qquad u \geq 0 ,$$

with

$$\tilde{\mathbf{x}}(\boldsymbol{\Psi}) = \left(\sum_{i=1}^{n} (\boldsymbol{\Psi} + \mathbf{U}_i)^{-1} \right)^{-1} \sum_{i=1}^{n} (\boldsymbol{\Psi} + \mathbf{U}_i)^{-1} \mathbf{x}_i .$$

Theorem 3 shows that conditional posterior for $\boldsymbol{\mu}$ given $\boldsymbol{\Psi}$ belongs to the family of elliptically contoured distributions $\rightsquigarrow$

$$\mathrm{E}(\boldsymbol{\mu}|\boldsymbol{\Psi}, \mathbf{X}) = \left(\sum_{i=1}^{n} (\boldsymbol{\Psi} + \mathbf{U}_i)^{-1} \right)^{-1} \sum_{i=1}^{n} (\boldsymbol{\Psi} + \mathbf{U}_i)^{-1} \mathbf{x}_i$$

Conditional posterior

Theorem 3

The conditional posterior $\pi(\boldsymbol{\mu}|\boldsymbol{\Psi},\mathbf{X})$ is given by

$$\pi(\boldsymbol{\mu}|\boldsymbol{\Psi},\mathbf{X}) \quad \propto \quad f_{\boldsymbol{\Psi},\mathbf{X}}\left((\boldsymbol{\mu}-\tilde{\mathbf{x}}(\boldsymbol{\Psi}))^\top\left(\sum_{i=1}^{n}(\boldsymbol{\Psi}+\mathbf{U}_i)^{-1}\right)(\boldsymbol{\mu}-\tilde{\mathbf{x}}(\boldsymbol{\Psi}))\right),$$

where

$$f_{\boldsymbol{\Psi},\mathbf{X}}(u) = f\left(\sum_{i=1}^{n}(\mathbf{x}_i-\tilde{\mathbf{x}}(\boldsymbol{\Psi}))^\top(\boldsymbol{\Psi}+\mathbf{U}_i)^{-1}(\mathbf{x}_i-\tilde{\mathbf{x}}(\boldsymbol{\Psi}))+u\right) \qquad u\geq 0,$$

with

$$\tilde{\mathbf{x}}(\boldsymbol{\Psi}) = \left(\sum_{i=1}^{n}(\boldsymbol{\Psi}+\mathbf{U}_i)^{-1}\right)^{-1}\sum_{i=1}^{n}(\boldsymbol{\Psi}+\mathbf{U}_i)^{-1}\mathbf{x}_i.$$

Bayesian estimate: posterior mean vector

$$\hat{\boldsymbol{\mu}} = \mathrm{E}(\boldsymbol{\mu}|\mathbf{X}) = \mathrm{E}(\mathrm{E}(\boldsymbol{\mu}|\boldsymbol{\Psi},\mathbf{X})|\mathbf{X}) = \mathrm{E}(\tilde{\mathbf{x}}(\boldsymbol{\Psi})|\mathbf{X}) = \mathrm{E}\left(\left(\sum_{i=1}^{n}(\boldsymbol{\Psi}+\mathbf{U}_i)^{-1}\right)^{-1}\sum_{i=1}^{n}(\boldsymbol{\Psi}+\mathbf{U}_i)^{-1}\mathbf{x}_i\middle|\mathbf{X}\right)$$

Marginal reference posterior

Theorem 4

The marginal posterior $\pi(\mathbf{\Psi}|\mathbf{X})$ is given by

$$\begin{aligned}
\pi(\mathbf{\Psi}|\mathbf{X}) \quad &\propto \quad \frac{\pi(\mathbf{\Psi})}{\sqrt{\det(\sum_{i=1}^{n}(\mathbf{\Psi}+\mathbf{U}_i)^{-1})}\prod_{i=1}^{n}\sqrt{\det(\mathbf{\Psi}+\mathbf{U}_i)}} \\
&\times \quad \int_0^\infty u^{p-1} f\left(u^2+\sum_{i=1}^{n}(\mathbf{x}_i-\tilde{\mathbf{x}}(\mathbf{\Psi}))^\top(\mathbf{\Psi}+\mathbf{U}_i)^{-1}(\mathbf{x}_i-\tilde{\mathbf{x}}(\mathbf{\Psi}))\right)\mathrm{d}u\,,
\end{aligned}$$

Marginal reference posterior

Theorem 4

The marginal posterior $\pi(\mathbf{\Psi}|\mathbf{X})$ is given by

$$\pi(\mathbf{\Psi}|\mathbf{X}) \propto \frac{\pi(\mathbf{\Psi})}{\sqrt{\det(\sum_{i=1}^n (\mathbf{\Psi}+\mathbf{U}_i)^{-1})}\prod_{i=1}^n \sqrt{\det(\mathbf{\Psi}+\mathbf{U}_i)}}$$
$$\times \int_0^\infty u^{p-1} f\left(u^2 + \sum_{i=1}^n (\mathbf{x}_i - \tilde{\mathbf{x}}(\mathbf{\Psi}))^\top (\mathbf{\Psi}+\mathbf{U}_i)^{-1} (\mathbf{x}_i - \tilde{\mathbf{x}}(\mathbf{\Psi}))\right) du\,,$$

Propriety and moment existence

Theorem 5

Consider the generalized multivariate random effects model with $\mathbf{U} = \mathrm{diag}(\mathbf{U}_1, ..., \mathbf{U}_n)$. Let $f(u)$ be a non-increasing function in $u \geq 0$ and $\frac{J_2}{2pn+p^2n^2} - \frac{1}{4} \leq 0$ where J_2 is defined in Theorem 1.

1. If $n \geq p$, then the posterior $\pi(\boldsymbol{\mu}, \mathbf{\Psi}|\mathbf{X})$ derived under the Jeffreys prior $\pi_J(\mathbf{\Psi})$ is proper.
2. If $n \geq p+1$, then the posterior $\pi(\boldsymbol{\mu}, \mathbf{\Psi}|\mathbf{X})$ derived under the Berger and Bernardo reference prior $\pi_R(\mathbf{\Psi})$ is proper.

Drawing samples: Metropolis-Hastings algorithm

Idea: Put $\mathbf{U}_1 = \ldots = \mathbf{U}_n = \mathbf{O}$ in $\pi(\boldsymbol{\mu}, \boldsymbol{\Psi}|\mathbf{X})$ and let

$$\bar{\mathbf{x}} = \frac{1}{n}\sum_{i=1}^{n}\mathbf{x}_i \quad \text{and} \quad \mathbf{S} = \frac{1}{n-1}\sum_{i=1}^{n}(\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})^\top,$$

- Berger and Bernardo reference prior:

$$q_R(\boldsymbol{\mu}, \boldsymbol{\Psi}|\mathbf{X}) = \det(\boldsymbol{\Psi})^{-(n+p+1)/2} f\left(\mathrm{tr}\left(\boldsymbol{\Psi}^{-1}\sum_{i=1}^{n}(\mathbf{x}_i - \boldsymbol{\mu})(\mathbf{x}_i - \boldsymbol{\mu})^\top\right)\right)$$

$$\propto \quad \left(1 + \frac{1}{n-p}\frac{n(n-p)}{n-1}(\boldsymbol{\mu} - \bar{\mathbf{x}})^\top \mathbf{S}^{-1}(\boldsymbol{\mu} - \bar{\mathbf{x}})\right)^{-n/2}$$

$$\times \quad \det(\boldsymbol{\Psi})^{-(n+p+1)/2}\det(\sum_{i=1}^{n}(\mathbf{x}_i - \boldsymbol{\mu})(\mathbf{x}_i - \boldsymbol{\mu})^\top)^{n/2} f\left(\mathrm{tr}\left(\boldsymbol{\Psi}^{-1}\sum_{i=1}^{n}(\mathbf{x}_i - \boldsymbol{\mu})(\mathbf{x}_i - \boldsymbol{\mu})^\top\right)\right)$$

- Jeffreys prior:

$$q_J(\boldsymbol{\mu}, \boldsymbol{\Psi}|\mathbf{X}) = \det(\boldsymbol{\Psi})^{-(n+p+2)/2} f\left(\mathrm{tr}\left(\boldsymbol{\Psi}^{-1}\sum_{i=1}^{n}(\mathbf{x}_i - \boldsymbol{\mu})(\mathbf{x}_i - \boldsymbol{\mu})^\top\right)\right)$$

$$\propto \quad \left(1 + \frac{1}{n-p+1}\frac{n(n-p+1)}{n-1}(\boldsymbol{\mu} - \bar{\mathbf{x}})^\top \mathbf{S}^{-1}(\boldsymbol{\mu} - \bar{\mathbf{x}})\right)^{-(n+1)/2}$$

$$\times \quad \det(\boldsymbol{\Psi})^{-(n+p+2)/2}\det(\sum_{i=1}^{n}(\mathbf{x}_i - \boldsymbol{\mu})(\mathbf{x}_i - \boldsymbol{\mu})^\top)^{(n+1)/2} f\left(\mathrm{tr}\left(\boldsymbol{\Psi}^{-1}\sum_{i=1}^{n}(\mathbf{x}_i - \boldsymbol{\mu})(\mathbf{x}_i - \boldsymbol{\mu})^\top\right)\right)$$

Algorithm A: drawing samples from $\pi(\boldsymbol{\mu}, \boldsymbol{\Psi}|\mathbf{X})$ under reference prior

1. **Initialization:** Choose the initial values $\boldsymbol{\mu}^{(0)}$ and $\boldsymbol{\Psi}^{(0)}$ for $\boldsymbol{\mu}$ and $\boldsymbol{\Psi}$ and set $b = 0$.
2. **Generating new values of $\boldsymbol{\mu}^{(w)}$ and $\boldsymbol{\Psi}^{(w)}$ from the proposal:**
   1. For given data $\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_n)$, generate $\boldsymbol{\mu}^{(w)}$ from $t_p\left(n-p, \bar{\mathbf{x}}, \frac{(n-1)\mathbf{S}}{n(n-p)}\right)$ with $\bar{\mathbf{x}}$ and $\mathbf{S}$ as in (21);
   2. Using data $\mathbf{X}$ and the drawn in step (i) $\boldsymbol{\mu}^{(w)}$, generate $\boldsymbol{\Psi}^{(w)}$ from $\boldsymbol{\Psi}|\boldsymbol{\mu} = \boldsymbol{\mu}^{(w)}, \mathbf{X} \sim GIW_p(n+p+1, \sum_{i=1}^n (\mathbf{x}_i - \boldsymbol{\mu}^{(w)})(\mathbf{x}_i - \boldsymbol{\mu}^{(w)})^\top, f)$.
3. **Computation of the Metropolis-Hastings ratio:**

$$MH^{(b)} = \frac{\pi(\boldsymbol{\mu}^{(w)}, \boldsymbol{\Psi}^{(w)}|\mathbf{X}) q_R(\boldsymbol{\mu}^{(b-1)}, \boldsymbol{\Psi}^{(b-1)}|\mathbf{X})}{\pi(\boldsymbol{\mu}^{(b-1)}, \boldsymbol{\Psi}^{(b-1)}|\mathbf{X}) q_R(\boldsymbol{\mu}^{(w)}, \boldsymbol{\Psi}^{(w)}|\mathbf{X})}.$$

4. **Moving to the next state of the Markov chain:**
   1. Generate $U^{(b)}$ from the uniform distribution on $[0, 1]$;
   2. If $U^b < \min\left\{1, MH^{(b)}\right\} \pi(\boldsymbol{\mu}^{(b)}$, then set $\boldsymbol{\mu}^{(b)} = \boldsymbol{\mu}^{(w)}$ and $\boldsymbol{\Psi}^{(b)} = \boldsymbol{\Psi}^{(w)}$ (Markov chain moves to the new state). Otherwise, set $\boldsymbol{\mu}^{(b)} = \boldsymbol{\mu}^{(b-1)}$ and $\boldsymbol{\Psi}^{(b)} = \boldsymbol{\Psi}^{(b-1)}$ (Markov chain stays in the previous state).
5. Return to step (2), increase b by 1, and repeat until the sample of size B is accumulated.