$\begin{pmatrix} 1 & 0.999 \\ 0.999 & 1 \end{pmatrix}$

Z. Wei

Weakly Informative Prior for Covariance Matrices v2_20230911

Boundary Estimation Problem

ML and REML $\hat{\mathbf{\Sigma}}$ are close to being non-positive definite (degenerate, singular).

- Most variables have zero variances or some variables are a linear combination of others. p. 140

```
> fit2u.alt <- lme4::lmer(LYMP ~ Phase * Time + (Phase * Time | ID), , REML=F,
                          control=lme4::lmerControl(optCtrl=list(maxfun=1e5)))
boundary (singular) fit: see help('isSingular')

> fit2u <- nlme::lme(fixed=LYMP ~ Phase * Time, random=~Phase * Time | ID, data=, method="ML",
                     control=nlme::lmeControl(opt="optim", optCtrl=list(maxit=1e5)))
```

In Bayesian statistics, the Wishart distribution is the conjugate prior for the inverse covariance matrix (i.e., the precision matrix $\mathbf{\Omega} = \mathbf{\Sigma}^{-1}$) of a multivariate normal random vector.

$\mathbf{\Omega} \sim \mathscr{W}(\mathbf{\Psi}^{-1}, \nu)$ Degrees of freedom $\nu > p - 1$ $\mathbf{\Omega} \mid \text{Data} \sim \mathscr{W}\left((n\mathbf{S} + \mathbf{\Psi})^{-1}, n + \nu\right)$

Scale matrix $\mathbf{\Psi}^{-1}_{p\times p} \succ 0$

$\mathbb{E}[\mathbf{\Omega}] = \nu \mathbf{\Psi}^{-1}$ `Omega ~ dwish(Psi, nu) #JAGS` `Omega ~ wishart(nu, InvPsi); //Stan`

Alternatively, the inverse-Wishart distribution is the conjugate prior for the covariance matrix.

$\mathbf{\Sigma} \sim \mathscr{W}^{-1}(\mathbf{\Psi}, \nu)$ Degrees of freedom $\nu > p - 1$ $\mathbf{\Sigma} \mid \text{Data} \sim \mathscr{W}^{-1}(n\mathbf{S} + \mathbf{\Psi}, n + \nu)$

Scale matrix $\mathbf{\Psi}_{p\times p} \succ 0$

$\mathbb{E}[\mathbf{\Sigma}] = \dfrac{\mathbf{\Psi}}{\nu - p - 1}$ for $\nu > p + 1$ `Sigma ~ inv_wishart(nu, Psi); //Stan`

Remark: If $\mathbf{\Sigma} \sim \mathscr{W}\left(\dfrac{\mathbf{I}_{p\times p}}{2\theta}, \nu\right)$, then $p(\mathbf{\Sigma}) \propto \displaystyle\prod_{r=1}^{p} f_\Gamma\left(\lambda_r\,; \frac{\nu - p + 1}{2}, \theta\right)$. p. 142-143, eq. 6

```
> dgamma(.001, shape=1.5, rate=10^-4) #≈0
```

$$(u_{0i}, u_{1i})^\top \overset{\text{i.i.d.}}{\sim} \mathcal{N}_2(\mathbf{0}, \mathbf{\Sigma}) \quad \text{and} \quad \mathbf{\Sigma} \equiv \begin{pmatrix} \sigma_0^2 & \rho\sigma_0\sigma_1 \\ \rho\sigma_0\sigma_1 & \sigma_1^2 \end{pmatrix} \text{ for } p = 2 \tag{1}$$

$\mathbf{\Sigma} \sim \mathscr{W}^{-1}\left((\nu_0 - 3)\cdot\mathbf{S}, \nu_0\right)$	Both prior and posterior means are the sample covariance matrix.	(2.1)
$\mathbf{\Sigma} \sim \mathscr{W}^{-1}(\mathbf{I}_{2\times 2}, 2)$	A multivariate generalization of the inverse-gamma distribution.	(2.2)
$\rho \sim U[-1, 1]$, $\sigma_0^2 \sim \text{Inv-Gamma}(10^{-4}, 10^{-4})$, $\sigma_1^2 \sim \text{Inv-Gamma}(10^{-4}, 10^{-4})$		(2.3)
$\rho \sim U[-1, 1]$, $\sigma_0 \sim U[0, \infty)$, $\sigma_1 \sim U[0, \infty)$	Unconstrained?	(2.4)
$\rho \sim U[-1, 1]$, $\sigma_0 \sim HC(0, 25)$, $\sigma_1 \sim HC(0, 25)$		(2.5)
$\mathbf{\Sigma} \sim \mathscr{W}(5 \times 10^3 \cdot \mathbf{I}_{2\times 2}, 4)$	Let θ be sufficiently small and $\nu = p + 2$ by default.	(2.6)

DISCUSSIONS

$$Y_{it} = (\underbrace{\beta_0}_{\text{fixed intercept}} + \underbrace{u_{0i}}_{\text{random intercept}}) + (\underbrace{\beta_1}_{\text{fixed slope}} + \underbrace{u_{1i}}_{\text{random slope}}) \cdot t + \epsilon_{it}$$ p. 144

1. What are fixed effects and random effects?
   - assumed to be constant for all trials - a subset of the entire population of treatments

2. What is restricted maximum likelihood (REML)? p. 139
   ML is generally preferred when comparing models with different fixed effects.

3. Does $|\hat{\boldsymbol{\Sigma}}| > 0$ imply that $\hat{\boldsymbol{\Sigma}} \succ 0$ and $\hat{\boldsymbol{\Sigma}}^{-1}$ exists? *https://rpubs.com/sherloconan/1015445*

```
# one negative eigenvalue
> issue <- matrix(c(1.1295254,0.4658979,-0.8727553,-0.2521035,
                    0.4663502,1.1937169,-0.9350007,-0.4105077,
                    -0.8733574,-0.9346202,1.1152784,0.1664068,
                    -0.2525762,-0.4106099,0.1664256,0.4964147), ncol=4, byrow=T)

# eigenvalue close to zero
> singular <- matrix(c(0.7573508, 0.6297603, -0.78943891, -0.28205123,
                       0.6305914, 1.2464388, -1.23028009, -0.53885046,
                       -0.7887137, -1.2304128, 1.65066378, 0.08047224,
                       -0.2813410, -0.5389635, 0.08109029, 0.78802955), ncol=4, byrow=T)

# two negative eigenvalues
> error <- matrix(c(1.0466132, 1.1008902, -1.3268045, -0.7891338,
                    1.1003529, 0.9572582, -1.2419223, -0.7404087,
                    -1.3267065, -1.2419124, 1.2020524, 0.7253173,
                    -0.7893596, -0.7404482, 0.7249186, 0.7322508), ncol=4, byrow=T)

> error.sym <- (error + t(error)) / 2
> CholWishart::dWishart(error.sym, 5, diag(4), log=T) #ERROR
```

   R only has a default function `stats::rWishart` to generate random Wishart distributed matrices. Other packages:

   `CholWishart::dWishart` can compute the density of a random Wishart distributed matrix.
   `MCMCpack::riwish` can generate a random inverse-Wishart distributed matrix.
   `LaplacesDemon::rinvwishart` also generates a random inverse-Wishart distributed matrix.

4. Why Bayes modal (BM) estimation? p. 141, eq. 4
   - viewed as maximum penalized likelihood estimation where the prior is a penalty function

5. Constrained data type and variable declaration in `Stan`

```
parameters {
  cov_matrix[4] Sigma; // 4×4 covariance matrix
}
```

Chung, Y., Gelman, A., Rabe-Hesketh, S., Liu, J., & Dorie, V. (2015). Weakly informative prior for point estimation of covariance matrices in hierarchical models. *Journal of Educational and Behavioral Statistics*, *40*, 136-157. https://doi.org/10.3102/1076998615570945
Liu, H., Zhang, Z., & Grimm, K. J. (2016). Comparison of inverse Wishart and separation-strategy priors for Bayesian estimation of covariance parameter matrix in growth curve analysis. *Structural Equation Modeling: A Multidisciplinary Journal*, *23*, 354-367. https://doi.org/10.1080/10705511.2015.1057285
Zhang, Z. (2021). A note on wishart and inverse wishart priors for covariance matrix. *Journal of Behavioral Data Science*, *1*, 119-126. https://doi.org/10.35566/jbds/v1n2/p2