EAAP 2011

Flexible prior specification for the genetic covariance matrix via the generalized inverted Wishart distribution

(Sebastián **Munilla** and Rodolfo J. C. **Cantet**)

Departamento de Producción Animal
Facultad de Agronomía – UBA
CONICET
BUENOS AIRES - ARGENTINA

Scenario

- Under a conjugated Bayesian approach, genetic covariance matrix **Σ** is usually assumed to follow an Inverted Wishart (IW) distribution

$$\boldsymbol{\Sigma} \sim IW\left(\upsilon, \mathbf{S}\right)$$

- However:

1. The formulation lacks flexibility
2. Convergence is slow as at each iteration the samples of all parameters in the covariance matrix are correlated

S => full set of complementary hyperparmeters
υ => a single scalar parameter models the uncertainty attached to it

Maternal animal model

- Take for instance the MAM

$$\boldsymbol{y} = \boldsymbol{Xb} + \boldsymbol{Za} + \boldsymbol{Z}_p\boldsymbol{e}_p + \boldsymbol{e}_o$$

- With

$$\text{Cov}\begin{bmatrix}\boldsymbol{a}\\ \boldsymbol{e}_p\\ \boldsymbol{e}_o\end{bmatrix} = \begin{bmatrix}\boldsymbol{\Sigma}\otimes\boldsymbol{A} & \boldsymbol{0} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{I}_d\sigma^2_{e_m} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{I}_n\sigma^2_{e_o}\end{bmatrix}$$

- Where

$$\boldsymbol{\Sigma} = \begin{bmatrix}\Sigma_{11} & \Sigma_{12}\\ \Sigma_{21} & \Sigma_{22}\end{bmatrix} \equiv \begin{bmatrix}\sigma^2_{a_o} & \sigma_{a_oa_m}\\ \sigma_{a_oa_m} & \sigma^2_{a_m}\end{bmatrix} \sim IW\left(\upsilon, \mathbf{S}\right)$$

Objectives

Motivation

- Instead, setting a generalized inverted Wishart (GIW) distribution provides a much more flexible approach

Objectives

1. Introduce the GIW distribution
2. Describe a Bayesian updating method to elicit prior genetic parameters for a maternally influenced trait

The GIW distribution

Details

- Originally introduced by Brown *et al*. (1994)
- Based on the Bartlett decomposition: $\boldsymbol{\Sigma} \rightarrow (\Sigma_{11}, \tau, \Gamma)$

$$\boldsymbol{\Sigma} = \begin{bmatrix} \Sigma_{11} & \Sigma_{11}\tau \\ \tau\Sigma_{11} & \Gamma + \tau^2\Sigma_{11} \end{bmatrix}, \text{ where } \begin{cases} \tau = \Sigma_{21}\Sigma_{11}^{-1} \\ \Gamma = \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12} \end{cases}$$

- Consider next the following partition of BV vector under the MAM

$$\boldsymbol{a}_o \sim N(\boldsymbol{0}, \Sigma_{11} \times \boldsymbol{A})$$
$$\boldsymbol{a}_m \mid \boldsymbol{a}_o \sim N(\boldsymbol{a}_o\tau, \Gamma \times \boldsymbol{A})$$

The GIW distribution

Definition

- Assume that a priori

$$\Sigma_{11} \sim S_0\chi^{-2}_{\upsilon_0}$$
$$\tau \mid \Gamma \sim N(\tau_0, \Gamma \times H)$$
$$\Gamma \sim S_0\chi^{-2}_{\upsilon_1+1}$$

- Then $\boldsymbol{\Sigma}$ follows a $GIW(\upsilon_0, \upsilon_1, S_0, S_1, \tau_0, H)$

Main advantages

- Flexibility
- Conditional conjugacy

The GIW distribution

The IW could be regarded a special case of the GIW on defining the following set of hyperparameters:

$$S_0 = \Sigma^*_{11}$$
$$S_1 = \Sigma^*_{22} - \Sigma^{*2}_{12}\Sigma^{*-1}_{11}$$
$$\tau_0 = \Sigma^*_{12}\Sigma^{*-1}_{11}$$
$$H = \Sigma^{*-1}_{11}$$

($\boldsymbol{\Sigma}^*$ prior scale matrix for $\boldsymbol{\Sigma}$)

Further: $\upsilon_0 = \upsilon + 1$, and $\upsilon_1 = \upsilon$

Sampling algorithm

1. Define υ_0 and υ_1
2. Form $\boldsymbol{\Sigma}$
3. Compute $\tilde{\upsilon}_0$ and $\tilde{\upsilon}_1$
4. Form $\boldsymbol{Q}^* = \boldsymbol{Q} + \boldsymbol{\Sigma}^*$
5. Sample $\Sigma_{11} \mid else \sim Q^*_{11}\chi^{-2}_{\tilde{\upsilon}_0}$
6. Sample $\Gamma \mid else \sim \left(Q^*_{22} - Q^{*-1}_{11}Q^{*2}_{12}\right)\chi^{-2}_{\tilde{\upsilon}_1+1}$
7. Sample $\tau \mid \Gamma, else \sim N\left(Q^{*-1}_{11}Q^*_{12}, Q^{*-1}_{11}\Gamma\right)$
8. Retrieve matrix $\boldsymbol{\Sigma}$

$$\boldsymbol{\Sigma} = \begin{bmatrix} \Sigma_{11} & \Sigma_{11}\tau \\ \tau\Sigma_{11} & \Gamma + \tau^2\Sigma_{11} \end{bmatrix}$$

Elicitation strategy

- The strategy arise naturally given the standard practice of genetic evaluations…
  1. Genetic parameters are usually re-estimated as data accrues over the years
  2. Then, we can use previous estimations to set hyperparms for the subsequent one
- How? Setting…

$$\upsilon_0 = \frac{2 \times \left[\hat{m}(\Sigma_{11})\right]^2}{\hat{v}(\Sigma_{11})} + 4 \text{ and } \upsilon_1 = \frac{2 \times \left[\hat{m}(\Gamma)\right]^2}{\hat{v}(\Gamma)} + 4$$

Simulation study

- A stochastic simulation study was carried out:
  1. MAM as the DGP
  2. Bayesian CVC estimation via Gibbs Sampler
  3. 39 Replicates – 10 Generations – Around 5,000 individuals per data set.
  4. Up to the 8th generation => Data subset
  5. Analyses: REML, Diffuse, IW100 and GIW
  6. Pmeans, PSD and autocorrelations for the genetic parameters used to compare results

Results

	Direct heritability (True value = 0.25)		Maternal heritability (True value = 0.15)		Dir-mat correlation (True value = -0.70)	
*Analyses**	*Estimate*	*SE*	*Estimate*	*SE*	*Estimate*	*SE*
REML	0.24 ± 0.04	0.05 ± 0.01	0.15 ± 0.03	0.04 ± 0.00	-0.69 ± 0.09	0.09 ± 0.02
Diffuse	0.24 ± 0.05	0.04 ± 0.01	0.14 ± 0.04	0.03 ± 0.01	-0.72 ± 0.13	0.09 ± 0.03
IW100_1	0.24 ± 0.04	0.03 ± 0.00	0.15 ± 0.03	0.02 ± 0.00	-0.69 ± 0.09	0.04 ± 0.01
IW100_2	0.29 ± 0.05	0.03 ± 0.00	0.20 ± 0.03	0.02 ± 0.00	-0.72 ± 0.06	0.04 ± 0.01
IW100_3	0.17 ± 0.04	0.02 ± 0.00	0.08 ± 0.04	0.01 ± 0.00	-0.60 ± 0.16	0.06 ± 0.02
GIW	0.23 ± 0.04	0.03 ± 0.00	0.14 ± 0.04	0.02 ± 0.01	-0.70 ± 0.12	0.06 ± 0.02

Table 1. Estimates and standard errors under different strategies with regard to prior opinion on the CVC

Results

	Direct heritability		Maternal heritability		Dir-mat correlation	
Analyses	*Lag10*	*Lag200*	*Lag10*	*Lag200*	*Lag10*	*Lag200*
Diffuse	0.92	0.48	0.95	0.62	0.96	0.55
IW100_1	0.83	0.12	0.84	0.12	0.83	0.09
IW100_2	0.82	0.11	0.82	0.10	0.80	0.07
IW100_3	0.83	0.15	0.87	0.20	0.85	0.15
GIW	0.85	0.19	0.91	0.33	0.90	0.29

Table 2. Lag10 and lag200 autocorrelations among samples

Conclusion

Take home messages

1. Differential uncertainty regarding prior knowledge on the genetic parameters can be accounted for through a GIW prior specification
2. As conditional conjugacy holds, parameter estimation can be readily accomplished via the GS

Coming soon…

Full paper: Munilla, S. and Cantet, R. J. C. 2011. JABG (in press).

Funding

- CONICET (PIP 0833/10)
- Universidad de Buenos Aires (UBACyT G042/08)
- Agencia Nacional de Ciencia y Tecnología (PICT 1863/06)

Thanks for your time