Pooling estimates of covariance components using a penalized likelihood approach

Karin Meyer
Animal Genetics and Breeding Unit, University of New England, Armidale, Australia

Fourth International Congress on Quantitative Genetics, Edinburgh, Scotland, June 17–22, 2012

The Task
- Estimate covariance matrices
  - MANY traits
  - Multiple sources of variation
- Combine results from part analyses

Problems
- ‘Full’ multivariate analysis not feasible
  → Use overlapping subsets of traits
  → Multiple estimates for individual \( \sigma_{ij} \)
- Pooled matrices
  → Must be positive definite
  → ‘Preserve’ phenotypic variances
  → Weight partial results differently
- Sampling variation

Pool estimates of covariance components using a penalized likelihood approach

Penalized pseudo likelihood

\[
-2 \log L \propto \sum_i d_i \left[ \log |V_i| + \text{tr}(V_i^{-1}M_i) \right] + \psi P
\]

with

- \( M_i = \sum_x C_x \otimes S_i \)
- \( V_i = \sum_x C_x \otimes \Sigma_i \)

\( \psi P \) = ‘Weight’ for subset \( i \)

\( \text{Cov(} RE \text{ in expect.}) \)

\( \text{Estimates for } RE \text{ for subset } i \)

\( \text{Expectation of } S_i \)

Solution
- Likelihood approach:
  - Treat estimates from part analyses as ‘data’
    → matrices of corrected MS/CP
  - Pool matrices for all random effects simultaneously
    → Assume pseudo pedigree structure
  - Impose penalty
    → borrow strength from phenotypic covariance
    → reduce sampling variation

Results
1. Matrix eigenvalues
   - truncate at \( \approx 0 \)
   - shrink to \( \lambda_{min} \approx 0 \)
2. Canonical eigenvalues
   - truncate at \( \approx 0 \)
   - ‘bend’ to \( \lambda_{min} \approx 0 \)
   - ‘bend’ further
3. ML approach
   - individual matrices
   - with pseudo pedigree
   - with pseudo ped. & penalty

Reduction in average loss (%)†

- Genetic
- Residual
- Phenotypic

-30 0 30 60
-30 0 30 60
0 30 60

† compared to multivariate analysis

Conclusions
- Analysis by parts ⇒ good estimates of large covariance matrices
  → some loss in efficiency, esp. for \( \hat{E}_G \)
- Likelihood approach performs well
  → flexible alternative to current methods
  → pool matrices for all sources of variation simultaneously
  → pseudo pedigree structure approximates sampling correlation between \( \Sigma_i \) for different sources ⇒ \( \Sigma_p \approx \text{constant} \)
- Penalization can improve estimates substantially
  → recommend ‘mild’ penalty

Software
Add-on to WOMBAT
→ run option – ‘pool’
→ general form of input
not restricted to WOMBAT estimates!

http://didgeridoo.une.edu.au/km/pool.php

Fourth International Congress on Quantitative Genetics, Edinburgh, Scotland, June 17–22, 2012