

Calculation of Standard Errors of estimates of genetic and phenotypic parameters in DMU

Just Jensen & Per Madsen
Danish Institute of Agricultural Sciences (DIAS)
Dept. of Animal Breeding and Genetics
Research Centre Foulum
Box 50, 8830 Tjele
Denmark

Email: per.madsen@agrsci.dk or just.jensen@agrsci.dk

When estimating genetic and phenotypic parameters in the form of variance and covariance components we maximize a log-likelihood function:

$$l(\Theta | y) = c - \frac{1}{2} [\ln |V| + \ln |C| + y'Py] \quad [1]$$

The general log-likelihood function is based on the multivariate normal density and can be rewritten as:

$$l(\Theta | y) \propto -\frac{1}{2} [\ln |R| + \ln |G| + \ln |C| + y'Py] \quad [2]$$

The form in [2] is much easier to compute than [1].

In order to obtain the variance of the parameter vector ($\text{Var}(\Theta)$) we must find:

$$I_o(\Theta) = -\frac{\partial^2 l(\Theta | y)}{\partial \Theta \partial \Theta'} \quad [3]$$

Expression [3] is the observed information matrix.

DMUAI is based on a combination of Newton-Raphson and Fisher Scoring (it is based on average of observed and expected information). So in DMUAI the observed information matrix is approximated by the average of observed and expected information matrices.

June 2005

Asymptotic variance of Θ is :

$$\text{Var}(\Theta) = I_o^{-1}(\Theta) \quad [4]$$

Square roots of the diagonal elements of $\text{Var}(\Theta)$ will provide standard errors of elements of Θ .

Genetic parameters are, however, often standardized into, say, heritabilities, genetic correlations, and so on. These are generally nonlinear functions of elements in Θ .

To compute the variance of such non-linear functions of random variables in Θ we use Taylor series expansions.

Let $g = f(\Theta_1, \Theta_2, \dots, \Theta_t)$ where $\Theta_1, \Theta_2, \dots, \Theta_t$ is a subset of the elements in the Θ -vectors.

If we assume $f(\cdot)$ to be continuous around $\hat{\Theta}$ and that all derivatives exist and are continuous we can expand $f(\cdot)$ around $\hat{\Theta}$, where the $\hat{\cdot}$ indicates the maximizer of [1].

$$f(\Theta_1, \Theta_2, \dots, \Theta_t) = f(\hat{\Theta}_1, \hat{\Theta}_2, \dots, \hat{\Theta}_t) + \left[\frac{\partial f}{\partial \Theta_1} \right]_{\Theta_1=\hat{\Theta}_1} (\Theta_1 - \hat{\Theta}_1) + \dots + \left[\frac{\partial f}{\partial \Theta_t} \right]_{\Theta_t=\hat{\Theta}_t} (\Theta_t - \hat{\Theta}_t) \quad [5]$$

Taking expectations on both sides we obtain:

$$E[f(\Theta_1, \Theta_2, \dots, \Theta_t)] = f(\hat{\Theta}_1, \hat{\Theta}_2, \dots, \hat{\Theta}_t) \quad [6]$$

By subtracting [6] from [5] we get:

$$f(\Theta_1, \Theta_2, \dots, \Theta_t) - E[f(\Theta_1, \Theta_2, \dots, \Theta_t)] = \left[\frac{\partial f}{\partial \Theta_1} \right]_{\Theta_1=\hat{\Theta}_1} \sigma_{\Theta_1}^2 + \dots + \left[\frac{\partial f}{\partial \Theta_t} \right]_{\Theta_t=\hat{\Theta}_t} (\Theta_t - \hat{\Theta}_t) \quad [7]$$

Squaring both sides and take expectations

$$\begin{aligned} [E[f(\Theta_1, \Theta_2, \dots, \Theta_t)] - E[f(\Theta_1, \Theta_2, \dots, \Theta_t)]]^2 = \\ \left(\frac{\partial f}{\partial \Theta_1} \right)^2 \bigg|_{\Theta_1=\hat{\Theta}_1} \sigma_{\Theta_1}^2 + \dots + \left(\frac{\partial f}{\partial \Theta_t} \right)^2 \bigg|_{\Theta_t=\hat{\Theta}_t} \sigma_{\Theta_t}^2 + \\ 2 \sum_{i=1}^t \sum_{j=i+1}^t \left[\left(\frac{\partial f}{\partial \Theta_i} \frac{\partial f}{\partial \Theta_j} \right) \bigg|_{\Theta_i=\hat{\Theta}_i} \sigma_{\Theta_i \Theta_j} \right] \end{aligned} \quad [8]$$

In [8] the parameters $\sigma_{\Theta_i}^2$, $\sigma_{\Theta_i \Theta_j}$ can be obtained from [4]. The left hand side of [8] is the definition of a variance.

Defining v as a vector of first derivatives of f with respect to each of the t parameters in Θ , then [8] can be written in matrix notation as:

$$E[f(\Theta_1, \Theta_2, \dots, \Theta_t)] - E[f(\Theta_1, \Theta_2, \dots, \Theta_t)]^2 = \nu' \text{Var}(\Theta) \nu = \nu' I_o^{-1}(\Theta) \nu \quad [9]$$

Example:

The following bivariate model were analyzed:

$$\begin{aligned} y_{1ijk} &= F_{1i} + b_{1j} + a_{ik} + e_{1ijk} \\ y_{2ijk} &= F_{2i} + b_{2j} + a_{2k} + e_{2ijk} \end{aligned}$$

where y_{1ijk} and y_{2ijk} were records on trait 1 and 2 resp., F_{1i} were a set of fixed effects, b_{1j} were random environmental effects, a_{1k} were additive animal effects and e_{1ijk} were random residuals. Trait 1 and trait 2 were measured in disjoint environments and on different, but related animals. Due to that b_{1j} , b_{2j} , e_{1ijk} and e_{2ijk} were uncorrelated, and a_{1k} and a_{2k} were correlated.

All in all the θ -vector therefore contained seven parameters:

$$\theta' = [\sigma_{b_1}^2, \sigma_{b_2}^2, \sigma_{a_1}^2, \sigma_{a_1 a_2}, \sigma_{a_2}^2, \sigma_{e_1}^2, \sigma_{e_2}^2]$$

The θ -vector, than maximized the likelihood function were:

$$\hat{\theta} = \begin{bmatrix} \sigma_{b_1}^2 \\ \sigma_{b_2}^2 \\ \sigma_{a_1}^2 \\ \sigma_{a_1 a_2} \\ \sigma_{a_2}^2 \\ \sigma_{e_1}^2 \\ \sigma_{e_2}^2 \end{bmatrix} = \begin{bmatrix} 0.054212 \\ 0.046320 \\ 0.201997 \\ 0.015414 \\ 0.007655 \\ 0.061806 \\ 0.182703 \end{bmatrix}$$

Table 1. Asymptotic (co)variance matrix(*1000) for the θ -vector in the example.

	$\sigma_{b_1}^2$	$\sigma_{b_2}^2$	$\sigma_{a_1}^2$	$\sigma_{a_1 a_2}$	$\sigma_{a_2}^2$	$\sigma_{e_1}^2$	$\sigma_{e_2}^2$
$\sigma_{b_1}^2$.5113						
$\sigma_{b_2}^2$.0093	.3732					
$\sigma_{a_1}^2$.1155	.0203	6.8796				
$\sigma_{a_1 a_2}$	-.1082	-.1279	.0581	1.8502			
$\sigma_{a_2}^2$	-.0118	-.1087	-.0325	.2814	.8204		
$\sigma_{e_1}^2$	-.1436	-.0168	-4.310	.0246	.0272	3.0193	
$\sigma_{e_2}^2$.0058	.0330	.0217	-.1572	-.6583	-.0174	.8238

Table 2. Asymptotic correlation matrix (SE on diagonal) for the θ -vector in the example.

	$\sigma_{b_1}^2$	$\sigma_{b_2}^2$	$\sigma_{a_1}^2$	$\sigma_{a_1 a_2}$	$\sigma_{a_2}^2$	$\sigma_{e_1}^2$	$\sigma_{e_2}^2$
$\sigma_{b_1}^2$.023						
$\sigma_{b_2}^2$.021	.019					
$\sigma_{a_1}^2$.062	.013	.083				
$\sigma_{a_1 a_2}$	-.111	-.154	.016	.043			
$\sigma_{a_2}^2$	-.018	-.196	-.014	.228	.029		
$\sigma_{e_1}^2$	-.116	-.016	-.946	.010	.017	.055	
$\sigma_{e_2}^2$.009	.059	.009	-.127	-.801	-.011	.029

The following standardized population parameters could be computed:

$$h_1^2 = \frac{\sigma_{a_1}^2}{\sigma_{b_1}^2 + \sigma_{a_1}^2 + \sigma_{e_1}^2}$$

$$h_{1|b}^2 = \frac{\sigma_{a_1}^2}{\sigma_{a_1}^2 + \sigma_{e_1}^2}$$

$$r_{a_1 a_2} = \frac{\sigma_{a_1 a_2}}{\sqrt{\sigma_{a_1}^2 + \sigma_{a_2}^2}}$$

Several other population parameters could be computed, but we only use the three defined above for illustration.

For the heritability corrected for the random environmental factor ($h_{1|b}^2$) we get is:

$$h_{1|b}^2 = \frac{\sigma_{a_1}^2}{\sigma_{a_1}^2 + \sigma_{e_1}^2} = \frac{.201997}{.201997 + .061806} = .7657$$

$$\frac{\partial h_{1|b}^2}{\partial \sigma_{a_1}^2} = \frac{\sigma_{e_1}^2}{(\sigma_{a_1}^2 + \sigma_{e_1}^2)^2} = \frac{.061806}{(.201997 + .061806)^2} = .8881$$

$$\frac{\partial h_{1|b}^2}{\partial \sigma_{e_1}^2} = \frac{-\sigma_{a_1}^2}{(\sigma_{a_1}^2 + \sigma_{e_1}^2)^2} = \frac{-.201997}{(.201997 + .061806)^2} = -2.9026$$

$$\begin{aligned}
\text{Var}(\mathbf{h}_{1|b}^2) &= \left(\frac{\partial \mathbf{h}_{1|b}^2}{\partial \sigma_{a_1}^2} \right)^2 \sigma_{\sigma_{a_1}^2}^2 + \left(\frac{\partial \mathbf{h}_{1|b}^2}{\partial \sigma_{e_1}^2} \right)^2 \sigma_{\sigma_{e_1}^2}^2 + 2 \left(\frac{\partial \mathbf{h}_{1|b}^2}{\partial \sigma_{a_1}^2} \right) \left(\frac{\partial \mathbf{h}_{1|b}^2}{\partial \sigma_{e_1}^2} \right) \sigma_{\sigma_{a_1 e_1}}^2 \\
&= (.8881)^2 (.0068796) + (-2.9026)^2 (.0030193) + 2(.8881)(-2.9026)(-.004310) \\
&= .04197
\end{aligned}$$

$$SE(\mathbf{h}_{1|b}^2) = \sqrt{.04197} = .02048$$

For the genetic correlation we get:

$$r_{a_1 a_2} = \frac{\sigma_{a_1 a_2}}{\sqrt{\sigma_{a_1}^2 \sigma_{a_2}^2}} = \frac{.01541}{\sqrt{.201997 + .007655}} = .39198$$

$$\frac{\partial r_{a_1 a_2}}{\partial \sigma_{a_1 a_2}} = \frac{1}{\sqrt{(\sigma_{a_1}^2 \sigma_{a_2}^2)}} = \frac{1}{\sqrt{(.201997) (.00765)}} = 25.43002$$

$$\frac{\partial r_{a_1 a_2}}{\partial \sigma_{a_1}^2} = - \frac{\sigma_{a_1 a_2}}{2 \sqrt{\sigma_{a_2}^2} (\sigma_{a_1}^2)^{\frac{3}{2}}} = - \frac{.015414}{2 \sqrt{(.007655)} (.201997)^{\frac{3}{2}}} = -.970259$$

$$\frac{\partial r_{a_1 a_2}}{\partial \sigma_{a_2}^2} = - \frac{\sigma_{a_1 a_2}}{2 \sqrt{\sigma_{a_1}^2} (\sigma_{a_2}^2)^{\frac{3}{2}}} = - \frac{.015414}{2 \sqrt{(.201997)} (.007655)^{\frac{3}{2}}} = -25.601758$$

$$\begin{aligned}
\text{Var}(\mathbf{r}_{a_1 a_2}) &= \left(\frac{\partial \mathbf{r}_{a_1 a_2}}{\partial \sigma_{a_1 a_2}} \right)^2 \sigma_{\sigma_{a_1 a_2}}^2 + \left(\frac{\partial \mathbf{r}_{a_1 a_2}}{\partial \sigma_{a_1}^2} \right)^2 \sigma_{\sigma_{a_1}^2}^2 + \left(\frac{\partial \mathbf{r}_{a_1 a_2}}{\partial \sigma_{a_2}^2} \right)^2 \sigma_{\sigma_{a_2}^2}^2 \\
&\quad + 2 \left(\frac{\partial \mathbf{r}_{a_1 a_2}}{\partial \sigma_{a_1 a_2}} \right) \left(\frac{\partial \mathbf{r}_{a_1 a_2}}{\partial \sigma_{a_1}^2} \right) \sigma_{\sigma_{a_1 a_2} \cdot \sigma_{a_1}^2} + 2 \left(\frac{\partial \mathbf{r}_{a_1 a_2}}{\partial \sigma_{a_1 a_2}} \right) \left(\frac{\partial \mathbf{r}_{a_1 a_2}}{\partial \sigma_{a_2}^2} \right) \sigma_{\sigma_{a_1 a_2} \cdot \sigma_{a_2}^2} + 2 \left(\frac{\partial \mathbf{r}_{a_1 a_2}}{\partial \sigma_{a_1}^2} \right) \left(\frac{\partial \mathbf{r}_{a_1 a_2}}{\partial \sigma_{a_2}^2} \right) \sigma_{\sigma_{a_1}^2 \cdot \sigma_{a_2}^2} \\
&= (25.4300)^2 (.001850) + (-.97026)^2 (.0068796) + (-25.60176)^2 (.0008204) \\
&\quad + 2(25.43001)(-.97026)(.00005811) + 2(25.43001)(-25.60176)(.00028) \\
&\quad + 2(-.97026)(-25.60176)(-.000033) \\
&= 1.36972
\end{aligned}$$

$$SE(\mathbf{r}_{a_1 a_2}) = \sqrt{1.36972} = 1.17035$$

Finally for the heritability (h_1^2) we get:

$$h_1^2 = \frac{\sigma_{a_1}^2}{\sigma_{b_1}^2 + \sigma_{a_1}^2 + \sigma_{e_1}^2} = \frac{201997}{.054212 + .201997 + .061806} = .6352$$

$$\frac{\partial h_1^2}{\partial \sigma_{a_1}^2} = \frac{\sigma_{b_1}^2 + \sigma_{e_1}^2}{(\sigma_{b_1}^2 + \sigma_{a_1}^2 + \sigma_{e_1}^2)^2} = \frac{.054212 + .061806}{(.054212 + .201997 + .061806)^2} = 1.147177$$

$$\frac{\partial h_1^2}{\partial \sigma_{b_1}^2} = \frac{-\sigma_{a_1}^2}{(\sigma_{b_1}^2 + \sigma_{a_1}^2 + \sigma_{e_1}^2)^2} = \frac{-.201997}{(.054212 + .201997 + .061806)^2} = -1.99733$$

$$\frac{\partial h_1^2}{\partial \sigma_{e_1}^2} = \frac{-\sigma_{a_1}^2}{(\sigma_{b_1}^2 + \sigma_{a_1}^2 + \sigma_{e_1}^2)^2} = \frac{-.201997}{(.054212 + .201997 + .061806)^2} = -1.99733$$

$$\begin{aligned} \text{Var}(h_1^2) &= \left(\frac{\partial h_1^2}{\partial \sigma_{a_1}^2} \right)^2 \sigma_{\sigma_{a_1}^2}^2 + \left(\frac{\partial h_1^2}{\partial \sigma_{b_1}^2} \right)^2 \sigma_{\sigma_{b_1}^2}^2 + \left(\frac{\partial h_1^2}{\partial \sigma_{e_1}^2} \right)^2 \sigma_{\sigma_{e_1}^2}^2 \\ &\quad + 2 \left(\frac{\partial h_1^2}{\partial \sigma_{a_1}^2} \right) \left(\frac{\partial h_1^2}{\partial \sigma_{b_1}^2} \right) \sigma_{\sigma_{a_1}^2 \sigma_{b_1}^2} + 2 \left(\frac{\partial h_1^2}{\partial \sigma_{a_1}^2} \right) \left(\frac{\partial h_1^2}{\partial \sigma_{e_1}^2} \right) \sigma_{\sigma_{a_1}^2 \sigma_{e_1}^2} + 2 \left(\frac{\partial h_1^2}{\partial \sigma_{b_1}^2} \right) \left(\frac{\partial h_1^2}{\partial \sigma_{e_1}^2} \right) \sigma_{\sigma_{b_1}^2 \sigma_{e_1}^2} \\ &= (1.147177)^2 (.0068796) + (-1.99733)^2 (.005113) + (-1.99733)^2 (.003019) \\ &\quad + 2(1.147177)(-1.99733)(.000116) + 2(1.147177)(-1.99733)(-.004310) \\ &\quad + 2(-1.99733)(-1.99733)(-.001436) \end{aligned}$$

$$= .03901$$

$$\text{SE}(h_1^2) = \sqrt{.03901} = .19751$$