

Error covariance structure for the dyadic model used in *dmm()*

Neville Jackson

10 Dec 2024

Contents

1	The dyadic model equations	1
2	Dyadic model assumptions	2
2.1	The assumption of independence of columns of \mathbf{W}	2
2.2	The assumptions of independence of dyadic residuals	2
3	Covariance structure of dyadic residuals	3
3.1	Covariance structure arising from sets of products of data values	3
3.2	How does Searle's equation 2 generate the covariance structure of Table 2	7
4	What part of the variance of sets of products is explained by fitting the dyadic model	8
4.1	OLS	8
4.2	GLS	9

1 The dyadic model equations

The dyadic model is presented in Section 6.2.2 of the document *dmmOverview.pdf* [1]. It results in a set of equations (the DME's) which are given in matrix form as equation 12, which is reproduced below

$$\Psi = \mathbf{W}\Gamma + \Delta \quad (1)$$

It is important to understand each of the matrix components of these equations, so we elaborate as follows
 First, the following variables set the size of the problem and the sizes of the above matrices

n number of individuals with data

m number of individuals in pedigree

l number of traits

k number of fixed effects

c number of variance components to be estimated

Now we explain each matrix

Ψ $n^2 \times l^2$ matrix of dyadic covariances for each pair of individuals (row) and each traitpair (col). Each covariance needs to be appropriately adjusted for fixed effects. The columns of Ψ become the dependent variables in a multi-trait multiple regression.

\mathbf{W} $n^2 \times c$ matrix containing the coefficients of the dyadic model equations, which become the independent variables of a multiple regression. Each column of \mathbf{W} has the form $Vec(\mathbf{M}\mathbf{Z}_c\mathbf{R}_c\mathbf{Z}'_c\mathbf{M}')$ where Vec is an operator that vectorizes a matrix, \mathbf{M} is a matrix from the fixed effect model such that $Y - X\hat{\alpha} = \mathbf{M}Y$, \mathbf{Z}_c is an incidence matrix relating individuals with data to individuals in the pedigree, and \mathbf{R}_c is a relationship matrix relevant to component c . Note that relationship matrices are used, not their inverse.

Γ $c \times l^2$ matrix of (co)variance component parameters to be estimated, which become the partial regression coefficients of a multiple regression.

Δ $n^2 \times l^2$ matrix of dyadic model residuals. Note the variance of these is not the individual environmental variance component - that has to be explicitly fitted as one of the columns of matrix \mathbf{W} and appears as one row of matrix Γ . The Δ matrix elements are the extent to which each dyadic covariance in matrix Ψ deviates from its expectation. The (co)variances of the elements of Δ enter into the standard errors of variance component estimates, in the same way that the (co)variances of residuals enter into the standard errors of any regression coefficient estimates.

Matrices Ψ and \mathbf{W} have n^2 rows, so we are looking at solving n^2 equations in c unknowns. This is an overdetermined system. We can use least squares to obtain an approximate solution.

2 Dyadic model assumptions

Before attempting a regression fit of model (1), it is worth looking at how well the data conform to the assumptions made in using least squares to fit a multiple regression model. The critical assumptions are

- the independent variables (columns of W) are uncorrelated
- the residuals (columns of Δ) are uncorrelated with each other and with the independent variates

A least squares fit does not involve assumptions regarding the distribution of residuals, but this does become involved when using residuals to obtain standard errors of parameter estimates.

2.1 The assumption of independence of columns of W

The correlations among independent variables (columns of W) are returned by *dmm()* in the attribute *dme.correl* of the returned object. These are pairwise correlations between the columns of the W matrix. They are not quite the same thing as correlations between relationship matrix elements, because the columns of W also involve M and Z matrices, as defined in *dmmOverview.pdf* [1] Tables 5 and 6.

The Z matrix simply selects a subset of the relationship matrix corresponding to those individuals which have observations. The M matrix comes from the fixed effects model, and consists of a set of weights which adjust for the degrees of freedom involved in each of the fixed effects applicable to each individual. The M matrix is related to the *hat* matrix ($M = I - H$). So the columns of the W matrix are a weighted subset of the elements of the appropriate relationship matrix. Their correlations are therefore not the same as relationship matrix element correlations. , but are usually similar

In the case of a dataset with mean only, and all individuals in the pedigree with data, such as the *warcolak* dataset, the subset is all of the relationship matrix and the weights in M are all equal, so the correlations of the columns of W are exactly the same as the correlations of the elements of the relationship matrices, in this special case.

In regression analysis it is generally considered that if there are collinearities among the independent variates amounting to correlations greater than around 0.5 then the estimates of regression coefficients are suspect. Translating this to our dyadic model, the variance component estimates are not likely to achieve a realistic separation if the columns of W have correlations exceeding 0.5. The option of using principal component regression (*dmeopt="pcr"*) has been developed as an experimental approach to dealing with serious collinearities among the components.

If we want to plot these W matrix column correlations (eg as scatteplots) we need argument *dmekeep=T* in calling *dmm()*. The default for *dmm()* is not to save the DME's in its return object. This can be overridden with argument *dmekeep=T*. This results in 2 attributes *dme.psi* and *dme.wmat* being added to the returned object. Caution, this can result in a very large returned object. The attribute *dme.wmat* is the W matrix, and its columns can be plotted with the standard *plot()* routine.

2.2 The assumptions of independence of dyadic residuals

To check the dyadic residuals Δ , we can use the S3 *plot()* method included in the *dmm* package. This will output histograms, qqnorm plots, and scatterplots of residuals against fitted and observed values. Plots tend to be more informative for datasets with smaller numbers of individuals.

Dyadic residuals are usually not far from normally distributed, but may be leptokurtic and slightly skewed to the right.

If dyadic residuals are correlated with fitted values of the components, then there is something wrong with the model, probably some extra component should be fitted.

One can expect dyadic residuals to be correlated with observed dyadic covariances. It is normal for the fitted components in a dyadic model to explain only a small fraction of the total variation, (ie the dyadic residuals Δ are large) and this of course leads to the observed covariances being highly correlated with residuals.

The question of patterns of correlation among the dyadic residuals themselves is a seriously difficult area. The covariances (or correlations) among the dyadic residuals for one trait form an $n^2 \times n^2$ matrix - ie n^4 elements. Too many to compute and cant be stored in R, except for very small datasets. This issue is discussed in the document *dmmOverview.pdf* [1] in relation to why *dmm()* is only able to do REML estimates. REML estimates require that the covariance (or correlation) matrix of the dyadic residuals be constructed and used to compute a GLS rather than an OLS solution to the DME's. That is only computationally feasible for very small datasets.

The covariance structure of dyadic residuals is actually known. It is derived in Searle et al (1992) [2] on pages 407-413. It involves fourth moments of the observations.

3 Covariance structure of dyadic residuals

Searle et al (1992) [2] give the following for the covariance matrix of the observations, and therefore of the errors, of a dispersion-mean model.

$$\mathcal{V} = cov(\Psi) \tag{2}$$

$$= (\mathbf{B} \otimes \mathbf{B})(\mathbf{I} + \mathbf{S}) \tag{3}$$

$$= (\mathbf{I} + \mathbf{S})(\mathbf{B} \otimes \mathbf{B}) \tag{4}$$

where

\mathbf{I} is an identity matrix of order n^2

\mathbf{S} is a commutation matrix $n^2 \times n^2$

\mathbf{B} is the covariance matrix all the observations, after removing fixed effects. The observations after fixed effects removed by least squares are $\mathbf{M}\mathbf{y}$ and the set of covariances (or products) of these residuals is $\mathbf{M}\mathbf{y} \otimes \mathbf{M}\mathbf{y}$. Now the covariance matrix of $\mathbf{M}\mathbf{y}$ is $\mathbf{B} = \mathbf{M}\mathbf{V}\mathbf{M}$, where \mathbf{V} is $cov(\mathbf{y})$.

The term $\mathbf{B} \otimes \mathbf{B}$ gives the covariance matrix of all the covariances of products of residuals.

So \mathcal{V} is the expected value of the covariance of dyadic residuals Δ . Searle's equation 2 only applies if \mathbf{y} is normally distributed. There are additional terms involving the fourth moments of \mathbf{y} if it is not normal.

Searle's presentation is univariate. There is only one observed trait in the vector \mathbf{y} . The term $(\mathbf{I} + \mathbf{S})$ is not trait dependent, but \mathbf{B} depends on \mathbf{V} which depends on the variance component estimates for a trait as well as on relationship matrices.

We need to look at each of the components of these errors of the dyadic model in order to understand why a set of products of data values has correlated errors and why a fit of data to the dyadic model often has a large amount of unexplained variance. The dyadic model residuals, Δ , need to be understood .

3.1 Covariance structure arising from sets of products of data values

Let there be 4 observations (x, y, u, v) and form the two products xy and uv . The covariance between these two products is

$$Cov(xy, uv) = E(x)E(u)Cov(y, v) + E(x)E(v)Cov(y, u) + \tag{5}$$

$$E(y)E(u)Cov(x, v) + E(y)E(v)Cov(x, u) + \tag{6}$$

$$Cov(x, u)Cov(y, v) + Cov(x, v)Cov(y, u) \tag{7}$$

In our case the data values (x, y, u, v) are all adjusted for fixed effects, so the expected values are all zero and the formula reduces to

$$Cov(xy, uv) = Cov(x, u)Cov(y, v) + Cov(x, v)Cov(y, u) \tag{8}$$

Table 1: Covariance matrix of products of two independent observations y_1 and y_2

Product	y_1y_1	y_1y_2	y_2y_1	y_2y_2
y_1y_1	$2Var^2(y_1)$	0	0	0
y_1y_2	0	$Var(y_1)Var(y_2)$	$Var(y_1)Var(y_2)$	0
y_2y_1	0	$Var(y_1)Var(y_2)$	$Var(y_1)Var(y_2)$	0
y_2y_2	0	0	0	$2Var^2(y_2)$

We now look at what this implies, using a small example of just 2 data values y_1 and y_2 leading to a set of 4 products y_1y_1, y_1y_2, y_2y_1 and y_2y_2 . Applying the above formula we have, for the case where y_1 and y_2 are independent

$$Cov(y_1y_1, y_1y_1) = 2Cov^2(y_1, y_1) \quad (9)$$

$$= 2Var^2(y_1) \quad (10)$$

$$Cov(y_1y_2, y_1y_1) = 0 \quad (11)$$

$$Cov(y_2y_2, y_1y_1) = 0 \quad (12)$$

$$Cov(y_1y_2, y_1y_2) = Cov(y_1, y_1)Cov(y_2, y_2) + Cov(y_1, y_2)Cov(y_2, y_1) \quad (13)$$

$$= Var(y_1)Var(y_2) \quad (14)$$

$$Cov(y_1y_2, y_2y_1) = Cov(y_1, y_2)Cov(y_2, y_1) + Cov(y_1, y_1)Cov(y_2, y_2) \quad (15)$$

$$= Var(y_1)Var(y_2) \quad (16)$$

This is a sample of the 16 possible combinations, forming the covariance matrix of all the products of 2 data values. The pattern should be obvious. Table 1 shows the full pattern for the case of y_1 and y_2 independent. If $Var(y_1) = Var(y_2)$ the entries reduce to $2V_2$ for the covariances of self products with themselves, and V_2 for the remaining non-zero block diagonal covariances.

This is the pattern specified by the $(\mathbf{I} + \mathbf{S})$ part of equation 2. If we construct a commutation matrix of size 4 and add an identity matrix to it we get

```
kmat <- function(m,n)
# commutation matrix or order m x n
{
  K <- matrix(0,m*n,m*n)
  for (i in 1 : m){
    for (j in 1 : n){
      K[i + m*(j - 1), j + n*(i - 1)] <- 1
    }
  }
  return(K)
}
```

```
> kmat(2,2)
  [,1] [,2] [,3] [,4]
[1,]  1  0  0  0
[2,]  0  0  1  0
[3,]  0  1  0  0
[4,]  0  0  0  1
>
> kmat(2,2) + diag(4)
  [,1] [,2] [,3] [,4]
[1,]  2  0  0  0
[2,]  0  1  1  0
[3,]  0  1  1  0
```

```
[4,] 0 0 0 2  
>
```

So the term $(\mathbf{I} + \mathbf{S})$ specifies the pattern of covariances of products of observations that are independent and have unit variance. If these assumptions are removed, we have the set of covariances of products shown in Table 2

Table 2: Covariance matrix of products of two observations y_1 and y_2 , without the assumption of independence

Product	y_1y_1	y_1y_2	y_2y_1	y_2y_2
y_1y_1	$2\text{Var}^2(y_1)$	$2\text{Var}(y_1)\text{Cov}(y_1, y_2)$	$2\text{Var}(y_1)\text{Cov}(y_1, y_2)$	$2\text{Cov}^2(y_1, y_2)$
y_1y_2	$2\text{Var}(y_1)\text{Cov}(y_2, y_1)$	$\text{Var}(y_1)\text{Var}(y_2) + \text{Cov}(y_1, y_2)\text{Cov}(y_2, y_1)$	$\text{Var}(y_1)\text{Var}(y_2) + \text{Cov}(y_1, y_2)\text{Cov}(y_2, y_1)$	$2\text{Var}(y_2)\text{Cov}(y_1, y_2)$
y_2y_1	$2\text{Var}(y_1)\text{Cov}(y_2, y_1)$	$\text{Var}(y_1)\text{Var}(y_2) + \text{Cov}(y_2, y_1)\text{Cov}(y_1, y_2)$	$\text{Var}(y_1)\text{Var}(y_2) + \text{Cov}(y_2, y_1)\text{Cov}(y_1, y_2)$	$2\text{Var}(y_2)\text{Cov}(y_1, y_2)$
y_2y_2	$2\text{Cov}^2(y_2, y_1)$	$2\text{Var}(y_2)\text{Cov}(y_2, y_1)$	$2\text{Var}(y_2)\text{Cov}(y_2, y_1)$	$2\text{Var}^2(y_2)$

[h]

In interpreting Table 2 we have to remember that covariances like $Cov(y_1, y_2)$ between observations are covariances between individuals arising from genetic relationships between individuals. So information from the relationship matrix enters into the dyadic errors via the covariances in Table 2. Also inbreeding information enters into the dyadic errors via the variances in Table 2. For example terms such as the covariance of $y_1 y_1$ with itself contain only variances, and are the same in the independent case of Table 1.

3.2 How does Searle's equation 2 generate the covariance structure of Table 2

Start with 2 observations on related individuals. After fixed effects, the covariance matrix among individuals is

$$\mathbf{V} = \mathbf{A}\sigma_A^2 + \mathbf{I}\sigma_E^2 \quad (17)$$

$$= \begin{bmatrix} 1 & r \\ r & 1 \end{bmatrix} \sigma_A^2 + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \sigma_E^2 \quad (18)$$

$$= \begin{bmatrix} (\sigma_A^2 + \sigma_E^2) & r\sigma_A^2 \\ r\sigma_A^2 & (\sigma_A^2 + \sigma_E^2) \end{bmatrix} \quad (19)$$

where

r is the additive relationship coefficient between the two individuals

σ_A^2 is additive genetic variance

σ_E^2 is environmental variance

\mathbf{A} is additive relationship matrix

\mathbf{I} is identity matrix

\mathbf{V} is covariance matrix among individuals

We are interested in matrix $\mathbf{B} = \mathbf{M}\mathbf{V}\mathbf{M}$. Here we ignore \mathbf{M} for simplicity which is equivalent to there being no fixed effects. So \mathbf{B} is just \mathbf{V} here, and we can form $\mathbf{B} \otimes \mathbf{B}$ as follows

$$\mathbf{B} \otimes \mathbf{B} = \begin{bmatrix} (\sigma_A^2 + \sigma_E^2) & r\sigma_A^2 \\ r\sigma_A^2 & (\sigma_A^2 + \sigma_E^2) \end{bmatrix} \otimes \begin{bmatrix} (\sigma_A^2 + \sigma_E^2) & r\sigma_A^2 \\ r\sigma_A^2 & (\sigma_A^2 + \sigma_E^2) \end{bmatrix} \quad (20)$$

$$= \begin{bmatrix} (\sigma_A^2 + \sigma_E^2)^2 & (\sigma_A^2 + \sigma_E^2)r\sigma_A^2 & (\sigma_A^2 + \sigma_E^2)r\sigma_A^2 & r^2\sigma_A^4 \\ (\sigma_A^2 + \sigma_E^2)r\sigma_A^2 & (\sigma_A^2 + \sigma_E^2)^2 & r^2\sigma_A^4 & (\sigma_A^2 + \sigma_E^2)r\sigma_A^2 \\ (\sigma_A^2 + \sigma_E^2)r\sigma_A^2 & r^2\sigma_A^4 & (\sigma_A^2 + \sigma_E^2)^2 & (\sigma_A^2 + \sigma_E^2)r\sigma_A^2 \\ r^2\sigma_A^4 & (\sigma_A^2 + \sigma_E^2)r\sigma_A^2 & (\sigma_A^2 + \sigma_E^2)r\sigma_A^2 & (\sigma_A^2 + \sigma_E^2)^2 \end{bmatrix} \quad (21)$$

$\mathbf{B} \otimes \mathbf{B}$ is the covariance matrix among all product combinations of the individual y values, allowing for relationship matrices among individuals and the variance components estimates, but not allowing for the covariance structure of sets of products. We can see how this reduces to a diagonal structure if the additive relationship coefficient (r) is zero.

We now need to multiply this with the $(\mathbf{I} + \mathbf{S})$ matrix, as in equation 2.

$$\mathcal{V} = (\mathbf{I} + \mathbf{S})(\mathbf{B} \otimes \mathbf{B}) \quad (22)$$

$$= \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} (\sigma_A^2 + \sigma_E^2)^2 & (\sigma_A^2 + \sigma_E^2)r\sigma_A^2 & (\sigma_A^2 + \sigma_E^2)r\sigma_A^2 & r^2\sigma_A^4 \\ (\sigma_A^2 + \sigma_E^2)r\sigma_A^2 & (\sigma_A^2 + \sigma_E^2)^2 & r^2\sigma_A^4 & (\sigma_A^2 + \sigma_E^2)r\sigma_A^2 \\ (\sigma_A^2 + \sigma_E^2)r\sigma_A^2 & r^2\sigma_A^4 & (\sigma_A^2 + \sigma_E^2)^2 & (\sigma_A^2 + \sigma_E^2)r\sigma_A^2 \\ r^2\sigma_A^4 & (\sigma_A^2 + \sigma_E^2)r\sigma_A^2 & (\sigma_A^2 + \sigma_E^2)r\sigma_A^2 & (\sigma_A^2 + \sigma_E^2)^2 \end{bmatrix} \quad (23)$$

$$= \begin{bmatrix} 2(\sigma_A^2 + \sigma_E^2)^2 & 2(\sigma_A^2 + \sigma_E^2)r\sigma_A^2 & 2(\sigma_A^2 + \sigma_E^2)r\sigma_A^2 & 2r^2\sigma_A^4 \\ 2(\sigma_A^2 + \sigma_E^2)r\sigma_A^2 & (\sigma_A^2 + \sigma_E^2)^2 + r^2\sigma_A^4 & (\sigma_A^2 + \sigma_E^2)^2 + r^2\sigma_A^4 & 2(\sigma_A^2 + \sigma_E^2)r\sigma_A^2 \\ 2(\sigma_A^2 + \sigma_E^2)r\sigma_A^2 & (\sigma_A^2 + \sigma_E^2)^2 + r^2\sigma_A^4 & (\sigma_A^2 + \sigma_E^2)^2 + r^2\sigma_A^4 & 2(\sigma_A^2 + \sigma_E^2)r\sigma_A^2 \\ 2r^2\sigma_A^4 & 2(\sigma_A^2 + \sigma_E^2)r\sigma_A^2 & 2(\sigma_A^2 + \sigma_E^2)r\sigma_A^2 & 2(\sigma_A^2 + \sigma_E^2)^2 \end{bmatrix} \quad (24)$$

\mathcal{V} is the covariance matrix of residuals for the dyadic model, allowing for correlated residuals due to both the covariance structure of products and the relationships between individuals and the variance component estimates. This structure agrees with Table 2. For example the term $r^2\sigma_A^4$ is the equivalent of $Cov(y_1, y_2)Cov(y_2, y_1)$, because only σ_A^2 contributes to covariance between individuals. So equation 2 is based on the expected covariances of products with a commutation matrix added to allow for the structure of a set of products.

It would be possible to dissect \mathcal{V} into

- part due to the variance components and relationships only $B \otimes B$
- part due to the covariance structure of a set of all product combinations in addition to that due to variance components and relationships $(I + S)(B \otimes B)$. This is the REML method. REML is GLS allowing for *all* correlated errors, regardless of the cause of the correlations.
- part ignoring relationships (set A to I in calculating B)

So what the covariances of product combinations do is add a term to some of the elements of $B \otimes B$

4 What part of the variance of sets of products is explained by fitting the dyadic model

Our dyadic model ($\Psi = W\Gamma + \Delta$) is a multiple regression equation. We need to ask how much of the variance of Ψ is explained by the regression fit, and what is the source of the residuals Δ .

4.1 OLS

Under the assumptions of ordinary least squares (OLS), ie that errors are $N(0, \sigma^2 I)$, the sum of squares of Ψ is $\Psi' \Psi$. Note this is a different thing from the covariance structure of Ψ discussed in the previous sections, which is $\Psi \Psi'$.

The normal equation for the dyadic model considered as a regression is

$$W' W \Gamma = W' \Psi$$

which if we multiply both sides by Γ' is

$$\begin{aligned} \Gamma' W' W \Gamma &= \Gamma' W' \Psi \\ (W\Gamma)'(W\Gamma) &= \Gamma' W' \Psi \\ &= W(W'W)^{-1}W' \end{aligned}$$

The left hand side is the variance of $W\Gamma$ (with $1/n$ omitted), so this is the part of the total sum of squares $\Psi' \Psi$ which is explained by the regression fit. The residual sum of squares ($\Delta' \Delta$) is then the difference $\Psi' \Psi - \Gamma' W' \Psi$.

For our simple example with 2 observations y_1 and y_2 forming four products $y_1y_1, y_1y_2, y_2y_1, y_2y_2$, the dyadic model is

$$\begin{aligned} \Psi &= \text{Vec} \begin{bmatrix} 1 & r \\ r & 1 \end{bmatrix} \sigma_A^2 + \text{Vec} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \sigma_E^2 + \text{Vec} \Delta \\ \begin{bmatrix} y_1y_1 \\ y_1y_2 \\ y_2y_1 \\ y_2y_2 \end{bmatrix} &= \begin{bmatrix} 1 & 1 \\ r & 0 \\ r & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \sigma_A^2 \\ \sigma_E^2 \end{bmatrix} + \begin{bmatrix} \delta_{11} \\ \delta_{12} \\ \delta_{21} \\ \delta_{22} \end{bmatrix} \end{aligned}$$

and the residual sum of squares is

$$\begin{aligned}\Delta' \Delta &= \Psi' \Psi - (\mathbf{W}\Gamma)' (\mathbf{W}\Gamma) \\ &= \begin{bmatrix} y_1 y_1 & y_1 y_2 & y_2 y_1 & y_2 y_2 \end{bmatrix} \begin{bmatrix} y_1 y_1 \\ y_1 y_2 \\ y_2 y_1 \\ y_2 y_2 \end{bmatrix} - \\ &\quad \begin{bmatrix} \sigma_A^2 + \sigma_E^2 & r\sigma_A^2 & r\sigma_A^2 & \sigma_A^2 + \sigma_E^2 \end{bmatrix} \begin{bmatrix} \sigma_A^2 + \sigma_E^2 \\ r\sigma_A^2 \\ r\sigma_A^2 \\ \sigma_A^2 + \sigma_E^2 \end{bmatrix}\end{aligned}$$

Note that

$$\begin{aligned}\Psi' \Psi &= \text{tr}(\Psi \Psi') \\ &= \text{tr}(\mathcal{Y})\end{aligned}$$

so comparison with equation 24 is valid.

We can substitute for $\Psi' \Psi$ and subtract $(\mathbf{W}\Gamma)' (\mathbf{W}\Gamma)$ to get

$$\begin{aligned}\Delta' \Delta &= 6(\sigma_A^2 + \sigma_E^2)^2 + 2(r\sigma_A^2)^2 \\ &= -2(\sigma_A^2 + \sigma_E^2)^2 - 2(r\sigma_A^2)^2 \\ &= 4(\sigma_A^2 + \sigma_E^2)^2\end{aligned}$$

So the dyadic residual sum of squares clearly depends on what is in the sum of squares of products of y_1 and y_2 that is not explained by the model and its regression fit. The trace of \mathcal{Y} (equation 24) contains all of the SS due to regression plus an additional $4((\sigma_A^2 + \sigma_E^2)^2)$. So the regression fit of the dyadic model is not removing all of the systematic effects. An amount $4(\sigma_A^2 + \sigma_E^2)^2$ is ending up in the residual sum of squares, presumably in addition to random residuals such as those due to measurement errors in the Ψ values. This is puzzling. Perhaps the expectation matrix \mathbf{W} should contain more terms.

One can see why the R^2 values for a dyadic model fit are typically less than 10 percent.

4.2 GLS

Under the assumptions of generalised least squares, ie that the errors are $N(0, V)$, the total sum of squares is $\Psi' R^{-1} \Psi$, where R is the correlation matrix corresponding to V

So we have for our simple example

$$\begin{aligned}V &= \begin{bmatrix} (\sigma_A^2 + \sigma_E^2) & r\sigma_A^2 \\ r\sigma_A^2 & (\sigma_A^2 + \sigma_E^2) \end{bmatrix} \\ R &= \begin{bmatrix} 1 & \frac{r\sigma_A^2}{(\sigma_A^2 + \sigma_E^2)} \\ \frac{r\sigma_A^2}{(\sigma_A^2 + \sigma_E^2)} & 1 \end{bmatrix} \\ R^{-1} &= \begin{bmatrix} \frac{1}{1 - (\frac{r\sigma_A^2}{(\sigma_A^2 + \sigma_E^2)})^2} & -\frac{\frac{r\sigma_A^2}{(\sigma_A^2 + \sigma_E^2)}}{1 - (\frac{r\sigma_A^2}{(\sigma_A^2 + \sigma_E^2)})^2} \\ -\frac{\frac{r\sigma_A^2}{(\sigma_A^2 + \sigma_E^2)}}{1 - (\frac{r\sigma_A^2}{(\sigma_A^2 + \sigma_E^2)})^2} & \frac{1}{1 - (\frac{r\sigma_A^2}{(\sigma_A^2 + \sigma_E^2)})^2} \end{bmatrix}\end{aligned}$$

A small numerical example is illuminating.

```
> rmat
      [,1] [,2]
[1,]  1.0  0.4
```

```

[2,] 0.4 1.0
> rmatn
      [,1] [,2]
[1,] 1.0 -0.4
[2,] -0.4 1.0
> psi
[1] 1 2
> t(psi) %*% psi
      [,1]
[1,] 5
> t(psi) %*% solve(rmat) %*% psi
      [,1]
[1,] 4.047619
> t(psi) %*% solve(rmatn) %*% psi
      [,1]
[1,] 7.857143
>

```

So a positive correlation *reduces* the total sum of squares, and a negative correlation *increases* it. That makes sense, positively correlated variates contribute less total information than independent variates, negatively correlated variates contribute more.

The normal equations for the dyadic model considered as a regression fitted by GLS are

$$\mathbf{W}'\mathbf{V}^{-1}\mathbf{W}\boldsymbol{\Gamma} = \mathbf{W}'\mathbf{V}^{-1}\boldsymbol{\Psi}$$

which if we multiply both sides by $\boldsymbol{\Gamma}'$ is

$$\begin{aligned} \boldsymbol{\Gamma}'\mathbf{W}'\mathbf{V}^{-1}\mathbf{W}\boldsymbol{\Gamma} &= \boldsymbol{\Gamma}'\mathbf{W}'\mathbf{V}^{-1}\boldsymbol{\Psi} \\ (\mathbf{W}\boldsymbol{\Gamma})'\mathbf{V}^{-1}(\mathbf{W}\boldsymbol{\Gamma}) &= \boldsymbol{\Gamma}'\mathbf{W}'\mathbf{V}^{-1}\boldsymbol{\Psi} \\ &= \mathbf{W}(\mathbf{W}'\mathbf{V}^{-1}\mathbf{W})^{-1}\mathbf{W}'\mathbf{V}^{-1}\boldsymbol{\Psi} \end{aligned}$$

The left hand side is the GLS sum of squares of $\mathbf{W}\boldsymbol{\Gamma}$, so this is the part of the total GLS sum of squares $\boldsymbol{\Psi}'\mathbf{V}^{-1}\boldsymbol{\Psi}$ which is explained by the regression fit. The residual GLS sum of squares $\boldsymbol{\Delta}'\boldsymbol{\Delta}$ is the difference $\boldsymbol{\Psi}'\mathbf{V}^{-1}\boldsymbol{\Psi} - \boldsymbol{\Gamma}'\mathbf{W}'\mathbf{V}^{-1}\boldsymbol{\Psi}$.

References

- [1] Jackson, N. (2015) An Overview of the R package dmm. From <http://cran.r-project.org/package=dmm>
Or <https://github.com/cran/dmm>
- [2] Searle, S.R., Casella, G., and McCulloch, C.E. (1992) Variance Components. John Wiley and Sons, New York.
- [3] Wolak, M.E. (2014) nadiv: an R package to create relatedness matrices for estimating non-additive genetic variances in animal models. *Methods in Ecology and Evolution* 3:792-796.