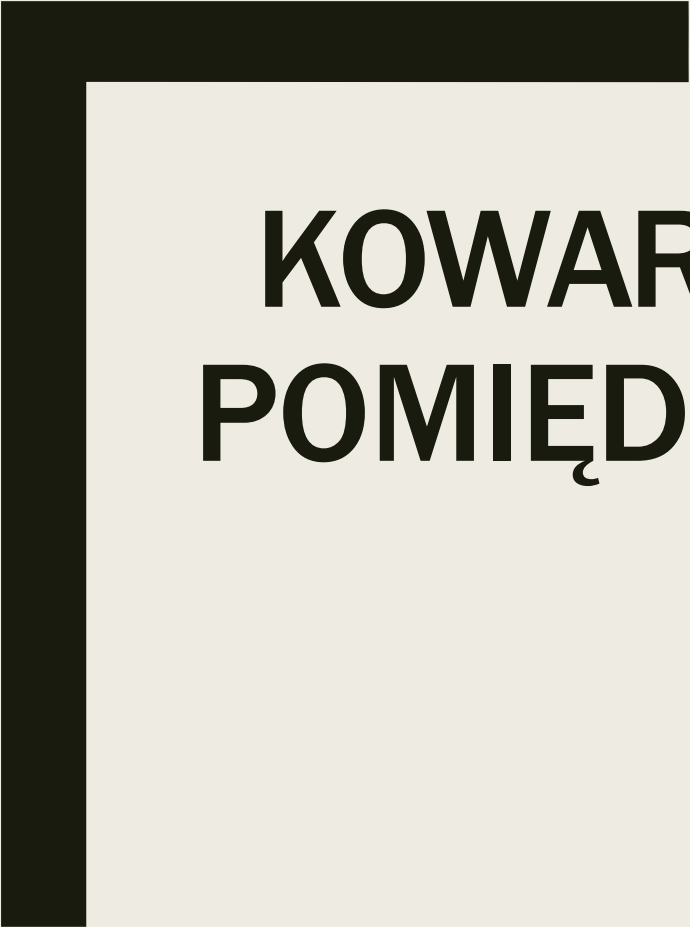
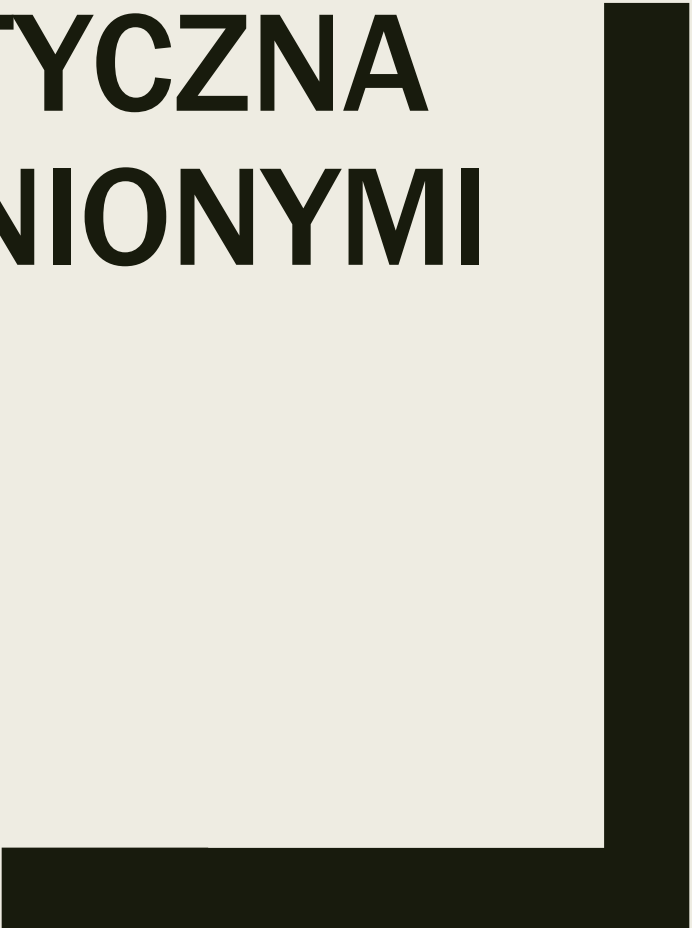




KOWARIANCJA GENETYCZNA POMIĘDZY SPOKREWNIONYMI OSOBNIKAMI

Dawid Słomian



The numerator relationship matrix A

- Coancestry / kinship coefficient → frakcja alleli identycznych przez pochodzenie (identical by descent; IBD) między parą osobników
- Additive genetic relationship coefficient / współczynnik spokrewnienia
- Numerator relationship matrix / macierz spokrewnień addytywnie genetycznych / A

Model bazowy → hodowlany → notacja skalarna

- fenotyp = środowisko + efekt addytywnie genetyczny + błąd
- $y_{ij} = \beta_i + a_i + e_{ij}$
- y_{ij} wydajność j osobnika i
- β_i efekt środowiskowy osobnika i np. płeć, rok urodzenia, stado
- a_i wartość hodowlana osobnika i = suma efektów addytywnych wszystkich genów
- e_{ij} sumaryczny efekt innych czynników środowiskowych oraz czynników genetycznych nieaddytywnych warunkujących wydajność j osobnika i

Model bazowy $\Rightarrow$ hodowlany $\Rightarrow$ notacja skalarna

- $y_{ij} = \beta_i + a_i + e_{ij}$
- $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{a} + \mathbf{e}$
- $\mathbf{y} \sim N(0, \mathbf{V}\sigma_y^2)$
- $\mathbf{e} \sim N(0, \mathbf{I}\sigma_e^2)$
- $\mathbf{a} \sim N(0, \mathbf{A}\sigma_a^2)$
- $\mathbf{V} = \mathbf{Z}\mathbf{G}\mathbf{Z}^T + \mathbf{R}$
- $\mathbf{R} = \mathbf{I}\sigma_e^2$
- $\mathbf{G} = \mathbf{A}\sigma_a^2$

The numerator relationship matrix A

A

- Kwadratowa, symetryczna
- Elementy przekątnej (a_{ii})
 - $1+F_i$
 - F_i współczynnik inbredu osobnika „ i ”
 - prawdopodobieństwo, że 2 losowo wybrane geny osobnika „ i ” będą miały allele identyczne przez pochodzenie

The numerator relationship matrix A

A

- Kwadratowa, symetryczna
- Elementy poza-przekątne (a_{ij})
 - *licznik (numerator) współczynnika spokrewnienia pomiędzy osobnikiem „i”, a osobnikiem „j”*
 - *współczynnik spokrewnienia (Wright 1922), korelacja pomiędzy wartościami hodowlanymi osobników i, j*
 - $\frac{cov(a_i, a_j)}{\sqrt{a_{ii}a_{jj}}} = \frac{a_{ij}}{\sqrt{a_{ii}a_{jj}}}$

The numerator relationship matrix A

$$A\sigma_a^2$$

Macierz kowariancji pomiędzy wartościami hodowlanymi osobników

The numerator relationship matrix A

- Tworzenie macierzy A na podstawie pliku z pochodzeniem osobników (Henderson 1976)

```
dawid.slomian@iz-pib.local@srvmixblup02:~/Single_step_final_stature/5th_gen
HOLPOLF005450026655 HOLDEUM000123058567 HOLPOLF005322177027
HOLPOLF005450026754 HOLDEUM000123058567 HOLPOLF005326154604
HOLPOLF005450026808 HOLDEUM000123058567 HOLPOLF005326153782
HOLPOLF005450026822 HOLDEUM000123058567 HOLPOLF005344919520
HOLPOLF005457080568 HOLFRAM005368780749 HOLPOLF005378231940
HOLPOLF005458858104 HOL840M003200824712 HOLNLDF000562148329
HOLPOLF005458858111 HOL840M003200824712 HOLNLDF000658145827
HOLPOLF005459839614 HOLFRAM005368780749 HOLPOLF005352217359
HOLPOLF005459839621 HOLFRAM005368780749 HOLPOLF005352217359
HOLPOLF005460067136 HOLDEUM000123058567 HOLPOLF005518486681
HOLPOLF005460325700 HOLDEUM000123058567 HOLPOLF005395910873
HOLPOLF005462753006 HOLFRAM002921633811 HOLPOLF005351866848
HOLPOLF005471577662 HOLFRAM008501201725 HOLPOLF005332355521
HOLPOLF005471577747 HOLFRAM008501201725 HOLPOLF005400870390
HOLPOLF005472842615 HOLITAM017992167089 HOLPOLF005474488491
HOLPOLF005477529764 HOLFRAM005368780749 HOLPOLF005323626906
HOLPOLF005477530128 HOLFRAM005368780749 HOLPOLF005323629020
```


The numerator relationship matrix A

- Tworzenie macierzy A na podstawie pliku z pochodzeniem osobników (*Henderson 1976*)

1. Posortować osobniki rosnąco wg roku urodzenia

2. Obliczać elementy macierzy A rozpoczynając od najstarszego osobnika

$$a_{ji} = a_{ij} = 0.5 \cdot (a_{js} + a_{jd}); a_{ii} = 1 + 0.5 \cdot a_{sd} \text{ (oboje rodzice znani)}$$

$$a_{ji} = a_{ij} = 0.5 \cdot (a_{js}); a_{ii} = 1 \text{ (jeden rodzic nieznany)}$$

$$a_{ji} = a_{ij} = 0; a_{ii} = 1 \text{ (oboje rodzice nieznani)}$$

The numerator relationship matrix A

Table 2.1. Pedigree for six animals.

Calf	Sire	Dam
3	1	2
4	1	Unknown
5	4	3
6	5	2

	1	2	3	4	5	6
1	1.00	0.0	0.50	0.50	0.50	0.25
2	0.00	1.0	0.50	0.00	0.25	0.625
3	0.50	0.50	1.00	0.25	0.625	0.563
4	0.50	0.00	0.25	1.00	0.625	0.313
5	0.50	0.25	0.625	0.625	1.125	0.688
6	0.25	0.625	0.563	0.313	0.688	1.125

■ $a_{11} = 1$; $a_{12} = a_{21} = 0$; $a_{13} = a_{31} = 0.5(a_{11} + a_{12}) = 0.5(1 + 0) = 0.5$

■ $a_{22} = 1$; $a_{23} = a_{32} = 0.5(a_{21} + a_{22}) = 0.5(0 + 1) = 0.5$

■ ...

■ $a_{33} = 1$; $a_{34} = a_{43} = 0.5(a_{31} + a_{32}) = 0.5(0.5 + 0) = 0.25$

■ ...

■ $a_{66} = 1 + 0.5(a_{25}) = 1 + 0.5(0.25) = 1.125$

$a_{ji} = a_{ij} = 0.5 \cdot (a_{js} + a_{jd})$; $a_{ii} = 1 + 0.5 \cdot a_{sd}$ (oboje rodzice znani)

$a_{ji} = a_{ij} = 0.5 \cdot (a_{js})$; $a_{ii} = 1$ (jeden rodzic nieznany)

$a_{ji} = a_{ij} = 0$; $a_{ii} = 1$ (oboje rodzice nieznani)

Dekompozycja macierzy A

- $A = TDT^T$ (*Thompson 1977*)
- Dekompozycja służy do odwrócenia macierzy A
- T macierz trójkątna zawiera współczynniki spokrewnienia par osobników

$$t_{ii} = 1$$

$$t_{ij} = 0.5 \cdot (t_{sj} + t_{dj}) \text{ (oboje rodzice znani)}$$

$$t_{ij} = 0.5 \cdot (t_{sj}) \text{ (jeden rodzic nieznany)}$$

$$t_{ij} = 0 \text{ (oboje rodzice nieznani)}$$

Dekompozycja macierzy A

- $A = TDT^T$ (Thompson 1977)
- **D** macierz diagonalna

$$d_{ii} = 0.5 - 0.25 \cdot (F_s + F_d) \text{ (oboje rodzice znani)}$$

$$d_{ii} = 0.75 - 0.25 \cdot (F_s) \text{ (jeden rodzic nieznany)}$$

$$d_{ii} = 1 \text{ (oboje rodzice nieznani)}$$

$$F_i = 0.5 \cdot a_{sd}$$

Odwrotność macierzy A

- A^{-1}
- model mieszany
- predyktor BLUP (Best Linear Unbiased Predictor)
- rozwiązanie układu równań modelu mieszanego MME (Mixed Model Equations)
- obliczenie A^{-1} jest szybsze niż A
- tworzenie i odwracanie macierzy A dla dużych zbiorów danych niemożliwe obliczeniowo
- *Henderson (1976)* → bezpośrednie stworzenie macierzy A^{-1}

Odwrotność macierzy A

- $A^{-1} = (T^{-1})^T D^{-1} T^{-1}$
- $D^{-1} = 1/d_{ii}$
- $T^{-1} = I - M$
- M macierz opisująca spokrewnienie w obrębie jednego pokolenia: dzieci-rodzice → każdy wiersz zawiera 0.5 w kolumnach rodziców

$$T^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.5 & 0.5 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.5 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.5 & 0.5 & 0.0 & 0.0 \\ 0.0 & 0.5 & 0.0 & 0.0 & 0.5 & 0.0 \end{bmatrix}$$

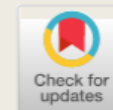
$(I) \qquad \qquad (M)$

Software



ORIGINAL RESEARCH article

Front. Genet., 28 July 2021 | <https://doi.org/10.3389/fgene.2021.655638>



Download A

Alternative Ways of Computing the Numerator Relationship Matrix



Mohammad Ali Nilforooshan^{1*},



Dorian Garrick² and



Bevin Harris¹

¹Livestock Improvement Corporation, Hamilton, New Zealand

²AL Rae Centre for Genetics and Breeding, School of Agriculture, Massey University, Palmerston North, New Zealand

1

T

Ar



frontiers

3.4. Henderson's Method

Using the concepts of Cholesky decomposition, $\mathbf{A}$ can be written as $\mathbf{A} = \mathbf{L}\mathbf{L}'$, where $\mathbf{L}$ is a lower triangle matrix with non-zero diagonal elements. [Henderson \(1975b\)](#) introduced a method for deriving $\mathbf{L}$ indirectly to calculate $\mathbf{A}$. If the definition of s and d change from sire and dam of j to parents of j , where $s < d$, calculation of $\mathbf{L}$ can be formulated as:

1. $\mathbf{L} = \mathbf{I}$

2. If only one parent (e.g., s) is known:

a. $L_{ij|i \leq s} = L_{si}/2$

b. $L_{jj} = \left(1 - \sum_{i=1}^s L_{ij}^2\right)^{\frac{1}{2}}$

3. If both parents are known:

a. $L_{ij|i \leq s} = (L_{si} + L_{di})/2$

b. $L_{ij|s < i \leq d} = L_{di}/2$

c. $L_{jj} = \left(1 + \frac{1}{2} \sum_{i=1}^s L_{si}L_{di} - \sum_{i=1}^s L_{ij}^2\right)^{\frac{1}{2}}$

Calculation of $\mathbf{L}'$ is coded in R function `getLp` (Data Availability Statement), and it requires a pedigree data frame as input.



Software

3.2. Thompson's Method

Applying the LDL factorization, $\mathbf{A} = \mathbf{T}\mathbf{D}\mathbf{T}'$, where $\mathbf{T}$ is a lower triangle matrix, and $\mathbf{D}$ is a diagonal matrix. In this method (Thompson, 1977), $\mathbf{T}$ and $\mathbf{D}$ matrices required to calculate $\mathbf{A}$ are derived indirectly (*i.e.*, no direct factorization or decomposition). For individual j , off-diagonal elements that were initialized with 0 are changed to $T_{ij} = (T_{si} + T_{di})/2$ if both parents s and d are known, and to $T_{ij} = T_{si}/2$ if one parent (*e.g.*, s) is known. The diagonal elements of $\mathbf{D}$ that were initiated with 1 change to $D_{jj} = (2 - F_s - F_d)/4$ if both parents are known, and to $D_{jj} = (3 - F_s)/4$ if one parent (*e.g.*, s) is known. This method is coded in R function `TDTp` (Data Availability Statement). This function requires the pedigree data frame and a corresponding vector of inbreeding coefficients as input.

Data Availability Statement

The data, scripts and functions that support the findings of this study are openly available in the figshare repository <https://doi.org/10.6084/m9.figshare.13497939>.



American
Genetic
Association

Journal of Heredity, 2016, 686–690

doi:10.1093/jhered/esw057

Computer Note

Advance Access publication October 11, 2016

OXFORD

Computer Note

Numericware N: Numerator Relationship Matrix Calculator

Bongsong Kim, William D. Beavis, and Jens Léon

From the Department of Agronomy, Iowa State University, Ames, IA 50011 (Kim and Beavis); and INRES—Plant Breeding, University of Bonn, Katzenburgweg 5, 53115 Bonn, Germany (Léon).

Address correspondence to Bongsong Kim, Noble Foundation, 2510 Sam Noble Parkway, Ardmore, Oklahoma 73401, or e-mail: bkim@noble.org.

Received May 30, 2016; First decision July 16, 2016; Accepted August 23, 2016.

Corresponding editor: C. Titus Brown

Abstract

We present the generalized numerator relationship matrix (GNRM) algorithm and Numericware N as a software tool for calculating the numerator relationship matrix (NRM). The GNRN algorithm

Package ‘pedigree’

February 20, 2015

Type Package
Title Pedigree functions
Version 1.4
Date 2012-11-27
Depends Matrix, HaploSim (>= 1.8.4), reshape
Description Pedigree related functions
License GPL (>= 2)
Author Albart Coster [aut, cre]
Maintainer Albart Coster <albart@dairyconsult.nl>
Repository CRAN
Date/Publication 2013-11-03 18:33:22
NeedsCompilation yes

R topics documented:

pedigree-package	2
add.Inds	2
blup	3
calcG	4
calcInbreeding	4
countGen	5
countOff	6
gblup	6
makeA	7
makeAinv	8
orderPed	9
trimPed	10

Software

INRAE

AgroParisTech
INSTITUT DES SCIENCES ET INNOVATIONS EN AGRICULTURE ET EN FORÊT
PARIS-SACLAY

université
PARIS-SACLAY



[Home](#) • [Welcome](#) • [Support & Expertise](#) • [Software](#) • [Pedig](#)

Pedig – software for the analysis of genealogies of large populations

Pedig Is a Fortran application specifically developed for the analysis of large populations

April 26th 2020

GRain2.2: this version replaces GRain1.0 and fixes a bug described in Doeke et al, 2020 (see paper in Publications directory) - Author: Ino Curik et al.

Dziękuję za uwagę