



Fundamentals of Genomic Prediction and Data-Driven Crop Breeding (August 4-8, 2025)

Relationship Matrices in Genetics

Module 2
August 5, 2025

Waseem Hussain and Mahender Anumalla
Rice Breeding Innovations Platform
IRRI

Covariance Between Individuals

$$Cov(X,Y) = 2f_{xy}\sigma_a^2 (\text{assuming Dominance} = 0)$$

Where f_{xy} is the coefficient of co-ancestry between two individuals and σ_a^2 is the additive genetic variance

Example of Rice

Genotypes	1	2	3	4
1	Var (1)	cov(1,2)	cov(1,3)	cov(1,4)
2		Var (2)	cov(2,3)	cov(2,4)
3			Var (1)	cov(3,4)
4				Var (4)

$$v(u) = A * \sigma_a^2$$

$$v(u) = A * \sigma_a^2 = (1 + F_i)\sigma_a^2$$

Elements of the **Numerator relationship matrix (A)** has the elements given by the $u(x,y) = 2f_{xy}$

- The probability of identical genes by descent occurring in two individuals is termed as **Coancestry or coefficient of Kinship** (Falconer and Mackay, 1996).
- The matrix that indicates the additive genetic relationship among individuals is called the **numerator relationship matrix (A)**.
- **A** is symmetric and its diagonal for an individual ***i*** (**a_{ii}**) is equal to **$1 + F_i$** , where **F_i** is the inbreeding coefficient of animal ***i*** (Wright, 19th to 22nd). Whether genotype is inbred or not
- When **A** is multiplied by additive genetic variance σ_a^2 , it results into the covariance among breeding values.

Covariance Between Individuals

$$Cov(X, Y) = 2f_{xy}\sigma_a^2 (\text{assuming Dominance} = 0)$$

Where f_{xy} is the coefficient of co-ancestry between two individuals and σ_a^2 is the additive genetic variance

Example of Rice

Genotypes	1	2	3	4
1	Var (1)	cov(1,2)	cov(1,3)	cov(1,4)
2		Var (2)	cov(2,3)	cov(2,4)
3			Var (1)	cov(3,4)
4				Var (4)

$$v(u) = A * \sigma_a^2$$

Elements of **the Numerator relationship matrix (A)** has the elements given by the $u(x, y) = 2f_{xy}$

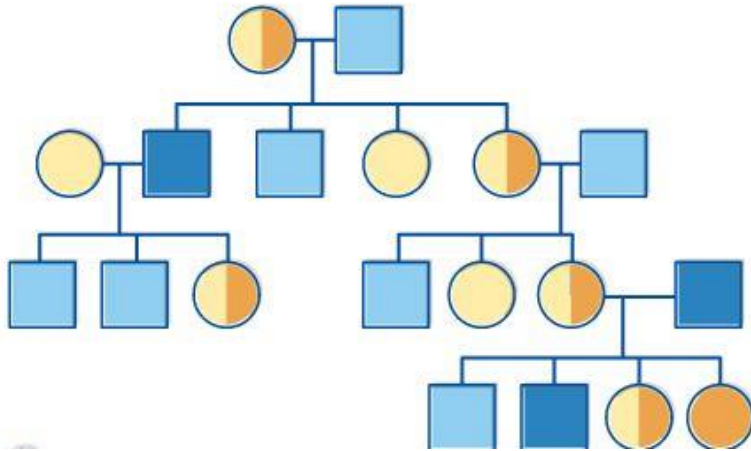
$$\text{Mixed Model } Y = \mu + X\beta + Zu$$

variance of random effect term u is given by **A**

Note:

- The diagonal represents variance and off-diagonal covariance
- If genotypes are independent off-diagonal elements=0
- Non-zero elements of off-diagonal reflect use of **relative information for BLUP estimation**

Algorithm to Construct A matrix



DESIGNATIO	Female	Male
IR64	IR5657-33-2-1	IR2061-465-1-5-5
IRRI154	IR73012-137-2-2-2	PSBRC10(IR50404-57-2-2-3)
IRRI193	IR68077-82-2-2-2-3	IR00A117
IR05N412	IR72875-94-3-3-2	IR73707-45-3-2-3
IR05N419	IR72887-34-2-1-3	IR73707-45-3-2-3
IR06N155	IR72158-11-5-2-3_IR73707-45	IR72875-94-3-3-2
IR09A220	IR72903-121-2-1-2	IR71606-1-1-4-2-3-1-2(NSIC110)
IR10A231	IRRI143_IR73718-23-2-1-3	IR00A110
IR10F559	IR80410-B-197-4_IRRI149	NSICRC158
IR10N237	IR01N111_IRRI164	IR72890-81-3-2-2
IR10N271	IR01W106	IR71676-90-2-2
IR11A282	IR04A427	BR29
IR11A303	IR04A427	IR72875-94-3-3-2
IR11A306	IR04A427	IR73006-12-3-3-2
IR11N121	IR05N341_IR64680-81-2-2-1-3	PSBRC10(IR50404-57-2-2-3)
IR11N202	IR05N173	BR29
IR12N135	IR01N149_IR64680-81-2-2-1-3	FEDEARROZ50
IR12F111	IR44004-74-3-2-3-3-3	IR70181-32-PMI1-1-5-1
BRRIDHAN55	BR10(BR51-46-5)_BR23	BR847-76-1-1
BRRIDHAN55		0

Matrix A is constructed using **path coefficients** and a **recursive method** (Henderson, 1976)

$$A = TD\hat{T} \text{ (Decomposition)}$$

Where, T is lower triangular matrix

And D is diagonal matrix

	IR64	IRRI154	IRRI193	IR05N412	IR05N419	IR06N155	IR09A220
IR64	1.07666	0.13361	0.01675	0.08966	0.03874	0.08966	0.31168
IRRI154	0.13361	1.03618	0.03124	0.04905	0.05164	0.05101	0.09983
IRRI193	0.01675	0.03124	1.00000	0.01019	0.01580	0.01019	0.02137
IR05N412	0.08966	0.04905	0.01019	1.00000	0.29292	0.39858	0.04587
IR05N419	0.03874	0.05164	0.01580	0.29292	1.00000	0.17475	0.04730
IR06N155	0.08966	0.05101	0.01019	0.39858	0.17475	1.00000	0.04587
IR09A220	0.31168	0.09983	0.02137	0.04587	0.04730	0.04587	1.02126
IR10A231	0.22386	0.05424	0.01408	0.02707	0.02324	0.02707	0.09506
IR10F559	0.22908	0.08668	0.02460	0.05484	0.04891	0.05313	0.09350
IR10N237	0.06116	0.09207	0.02596	0.04761	0.10362	0.04761	0.09354
IR10N271	0.03119	0.04701	0.01812	0.01327	0.01976	0.02108	0.03659
IR11A282	0.00000	0.00391	0.00000	0.00000	0.00000	0.03125	0.00000
IR11A303	0.08966	0.04905	0.01019	0.25503	0.03120	0.25503	0.04587

Final A matrix

The inverse of A matrix

$$\text{cov}(u) = A^{-1} * \frac{1}{\sigma_u^2}$$

A^{-1} the inverse of numerator relationship matrix

Why to Inverse of the matrix

- Simple structure than A matrix
- A big matrix with huge dimensions is computationally intense ; an inverse matrix works better.



Dissecting Relationship matrices

Additive and Dominance Model

$$y_i = \mu + \sum_{j=1}^n X_{ij}a_j + \sum_{j=1}^n W_{ij}d_j + e_i$$

Matrix Notation: $y = 1\mu + Xa + Wd + e$

Where y_i is the phenotypic value of the individual i

μ is the population mean

a_j is the additive effect

d_j is the dominant affect of each marker

X_{ij} is the design marker matrix = **1** ($A_1 A_1$), **0** ($A_1 A_2$), **-1** ($A_2 A_2$) for genotypes

W_{ij} is the design marker matrix = **0** ($A_1 A_1$), **1** ($A_1 A_2$), **0** ($A_2 A_2$) for genotypes



Dominance and Additive Relationship

$$\text{Phenotype}(P) = \text{Genotype}(G) + \text{Environment}(E) + \text{Residual}(e)$$

$$V_P = V_g + V_e + V_{ge}$$

$$\sigma_G^2 = \sigma_A^2 + \sigma_D^2$$

$$\sigma_G^2 = 2pq[a + d(q - p)]^2 + [2pqd]^2$$

$$\text{Additive variance } (\sigma_A^2) = 2pq\alpha^2$$

$$\text{Dominance variance } (\sigma_D^2) = [2pqd]^2$$



Dissecting Relationship matrices

$$\sigma_A^2 = 2pq\alpha^2 \text{ and } \sigma_D^2 = [2pqd]^2$$

If a and d effects are considered to be random and the covariance between them equal to zero the covariance of additive individual affect (u) is

$$Cov(u) = XX'\sigma_a^2 + X\dot{X}'(q-p)^2\sigma_d^2$$

Where σ_a^2 and σ_d^2 are SNP variances for additive and dominant components, respectively

$$\text{Now, } \sigma_A^2 = 2pq\sigma_a^2 + 2pq(q-p)^2\sigma_d^2$$

Assuming Zero covariance between a and d ,

$$Cov(u) = \frac{X\dot{X}'}{2pq}\sigma_A^2 = G\sigma_A^2$$

which is the classical G matrix of GBLUP ([VanRaden 2008](#)).

As for the dominant deviations, its covariance is

$$Cov(v) = \frac{W\dot{W}}{(2pq)^2}\sigma_D^2 = D\sigma_D^2$$

Proposed by Vitezica et al. 2013



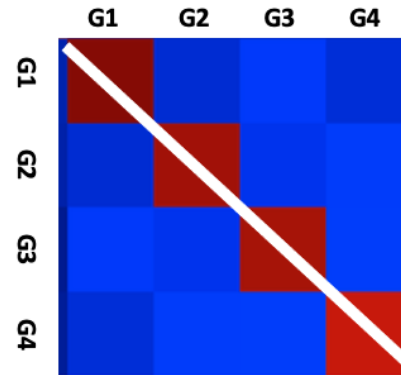
$$\frac{W\dot{W}}{2pq(1-2pq)}\sigma_D^2 \text{ (Proposed by Su et al. 2012)}$$

Construction of G matrix

3 steps

1. Create a centered Z matrix
2. Create the cross product
3. Divide it by $\sum_{i=1}^m 2p_j(1-p_j)$

$$G = \frac{Z\hat{Z}}{\sum_{i=1}^m 2p_j(1-p_j)}$$



First G matrix, VanRaden (2008)

Note:

- If genotypes are independent off-diagonal elements=0
- Non-zero elements of off-diagonal reflect use of **relative information** for **BLUP estimation**

$\sigma^2 a$ is additive variance = $\sum_{i=1}^m 2p_j(1-p_j)$

Additional Read

- <https://academic.oup.com/genetics/article/195/4/1223/5935377>
- <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0045293>
- <https://www.sciencedirect.com/science/article/pii/S0022030209707933>
- [https://pism.in.com/10.3168/jds.S0022-0302\(75\)84776-X](https://pism.in.com/10.3168/jds.S0022-0302(75)84776-X)
- <http://morotalab.org/Mrode2005/relmat/relmat.html#section0008000000000000>

