

OHP Picture 1

Recap: 2023-03-13

- Predictors with discrete fixed levels such as "Breed", the different levels are converted to numeric codes.
In R: Assignment of factor levels to numbers is done according to the alphabetic order of the factor levels.
- This sort of predictors ('x') are also called factors. The possible values that a factor can take are called levels.
- Eg.: "Breed" is a factor
"Angus", "Limousin", "Simmental" are the levels

Model: - Show influence of breed on expect

OHP Picture 2

Model : - Show influence of breed on expected body weight:
- For animal i , the expected body weight
 $E(y_i) = b_0 + b_1 x_i$
where x_i is the numeric code for the breed of animal i .

Consequence of Model : Consider data

$$\begin{aligned} \text{Animal 1: } E(y_1) &= b_0 + b_1 \cdot 1 \\ 3: \quad E(y_3) &= b_0 + b_1 \cdot 3 \\ 10: \quad E(y_{10}) &= b_0 + b_1 \cdot 2 \end{aligned} \left. \begin{array}{l} \text{dependent} \\ \text{on assignment} \\ \text{of codes to} \\ \text{breeds} \end{array} \right\}$$

$$E(y_{s_i}) - E(y_{m_i}) = 2 \cdot b_1$$

$$E(y_{u_i}) - E(y_{m_i}) = b_1$$

Exception : Regression on discrete Factors can be used in Marker effect estimation

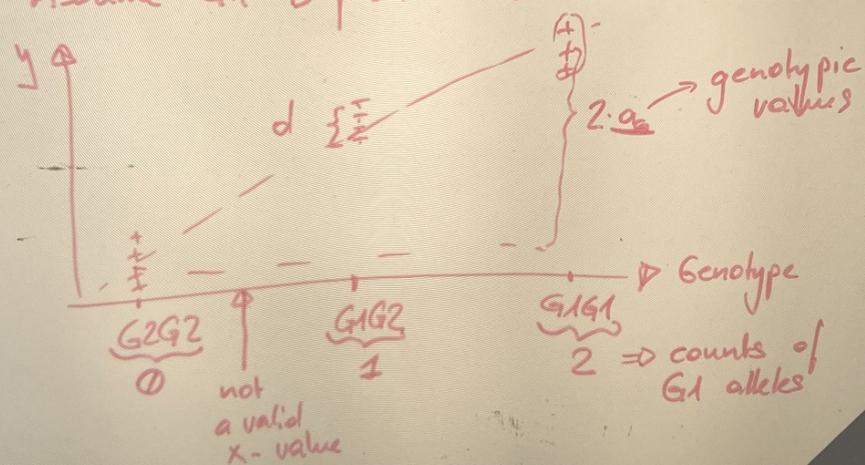
OHP Picture 3

Exception : Regression on discrete Factors can be used in Marker effect estimation.

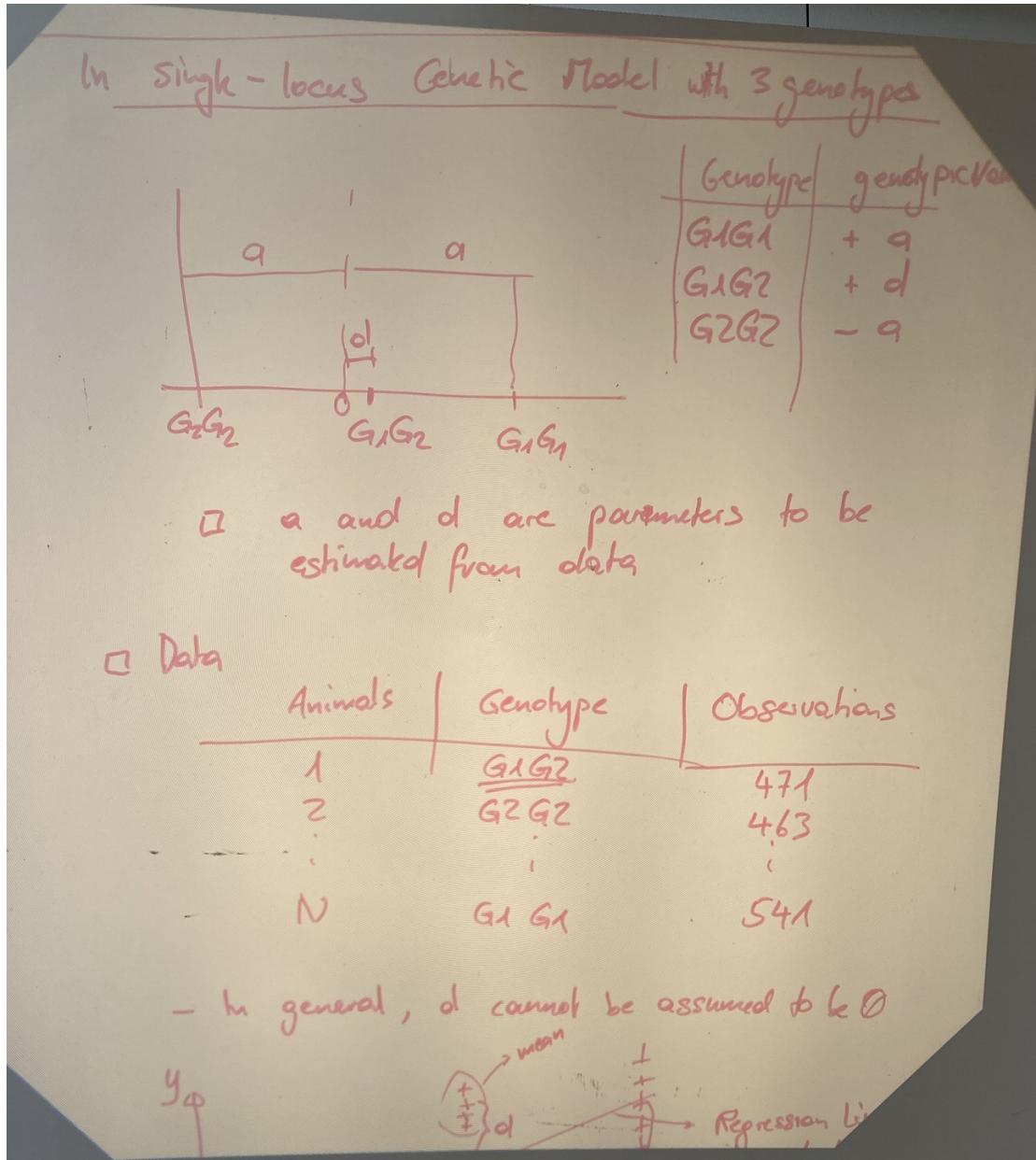
□ Data set :

Animal	Locus G	Observation
1	G1G2	471
2	G2G2	463
:		
N	G1G1	541

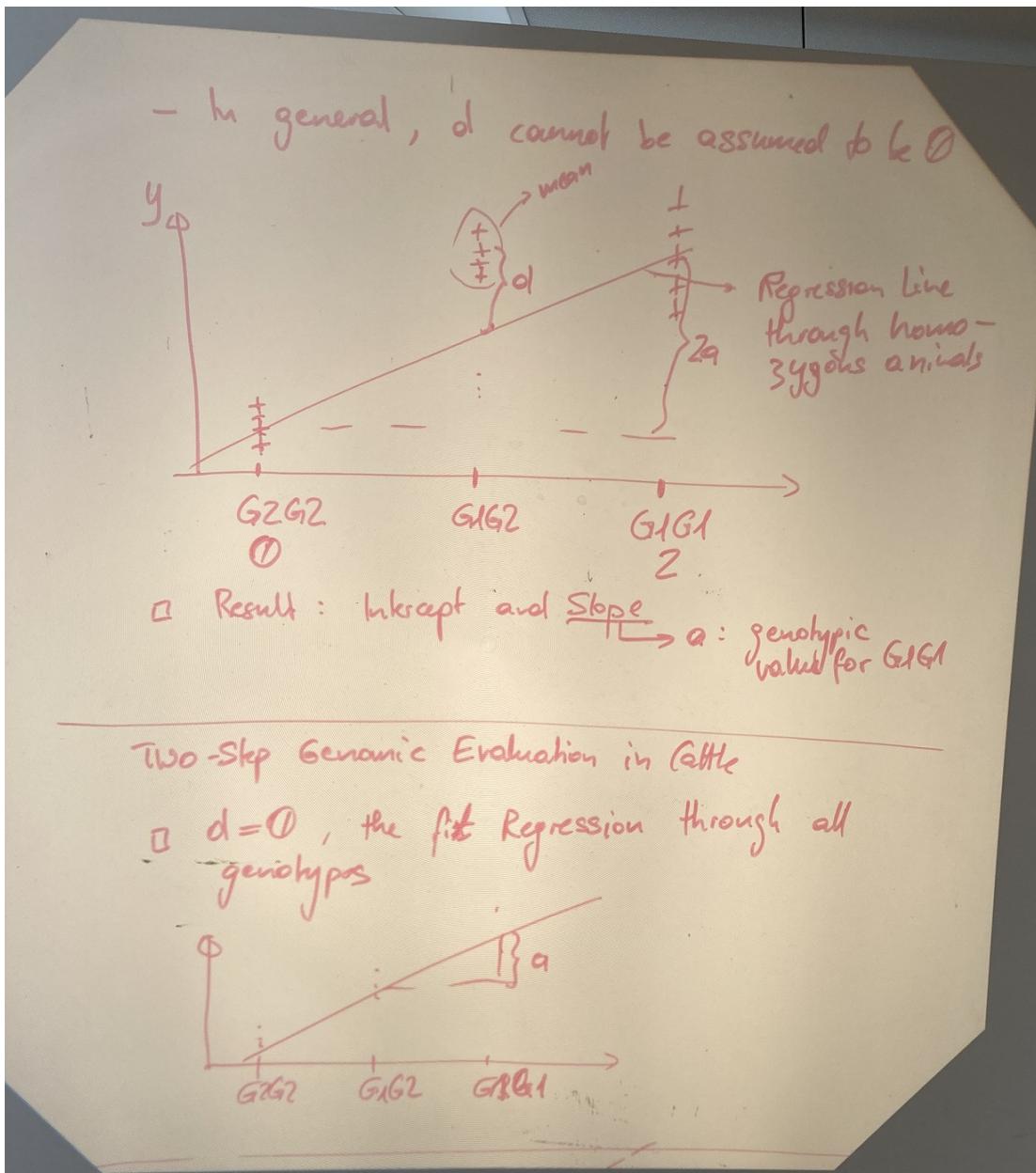
□ Assume G1 is positive Allele :



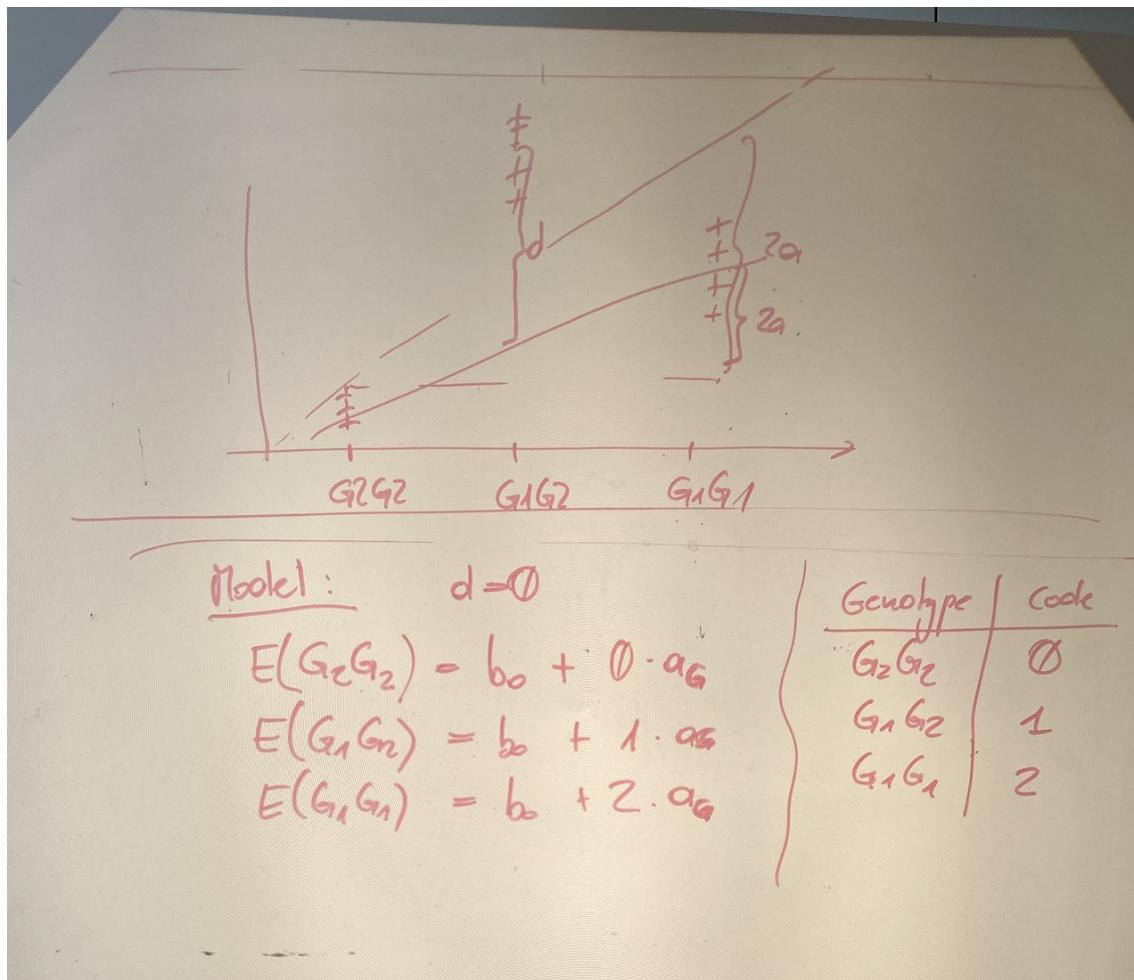
OHP Picture 4



OHP Picture 5



OHP Picture 6



OHP Picture 7

Regression on Dummy Variables

- Not one predictor (x -variable) per factor (breed)
but an additional predictor per level

- Factor breed

Levels	Predictors	
Angus	x_1	b_1
Limousin	x_2	b_2
Simmental	x_3	b_3

In matrix-vector notation
additional columns in X -matrix
↓ unknowns

$$\text{Animal 1: } y_{11} = b_0 + 1 \cdot b_1 + 0 \cdot b_2 + 0 \cdot b_3 + e_{11}$$

$$\text{Animal 2: } y_{21} = b_0 + 1 \cdot b_1 + 0 \cdot b_2 + 0 \cdot b_3 + e_{21}$$

$$3: y_{33} = b_0 + 0 \cdot b_1 + 0 \cdot b_2 + 1 \cdot b_3 + e_{33}$$

$$10: y_{102} = b_0 + 0 \cdot b_1 + 1 \cdot b_2 + 0 \cdot b_3 + e_{102}$$

In Matrix-Vector Notation

$$\begin{pmatrix} y_{11} \\ y_{21} \\ y_{33} \\ \vdots \\ y_{102} \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \end{pmatrix} + \begin{pmatrix} e_{11} \\ e_{21} \\ e_{33} \\ \vdots \\ e_{102} \end{pmatrix}$$

Group animals according to

OHP Picture 8

In Matrix - Vector Notation :

vector $y = \begin{bmatrix} y_{11} \\ y_{21} \\ y_{41} \\ y_{72} \\ y_{82} \\ y_{92} \\ y_{102} \\ y_{83} \\ y_{53} \\ y_{63} \end{bmatrix} = \begin{bmatrix} 471 \\ 463 \\ 470 \\ \vdots \\ \vdots \\ \vdots \\ 491 \end{bmatrix}$

Group animals according to breed
All Angus Animals first, then all Limousin animals, last all Simmental animals

vector : $b = \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \end{bmatrix}$

$b_0 \rightarrow$ Intercept
 $b_1 \rightarrow$ effect of Angus on BW
 $b_2 \rightarrow$ effect of Limousin on BW
 $b_3 \rightarrow$ effect of Simmental on BW

vector $e = \begin{bmatrix} e_{11} \\ e_{21} \\ e_{41} \\ e_{63} \end{bmatrix}$

Matrix $X = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}$

OHP Picture 9

Matrix $X = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & 1 & 0 \end{bmatrix}$

} design matrix
incidence

Model : $\underline{y} = \underline{X} \cdot \underline{b} + \underline{e} \rightarrow \text{Unknown}$

Known Known Unknown, estimated from data

Estimates of b using least squares

Task

- Find \hat{b} such that $e^T e$ is minimal
- Result : $X^T X \hat{b}^{(0)} = X^T y$ (NEq)

Reg : $\hat{b} = \underbrace{(X^T X)^{-1}}_{\text{is only possible, if } X \text{ has full rank}} X^T y$

OHP Picture 10

□ Solution to NEq:

$$\overset{(0)}{b} = (X^T X)^{-1} X^T y$$

where $(X^T X)^{-1}$ is
a generalized inverse
of $(X^T X)$

Generalized Inverse

For matrix A , the matrix G is a
generalized inverse, if

$$AGA = A$$

$$(AGA)^T = A^T$$

Special case
where A^{-1}
 $\Rightarrow G = A^{-1}$

$$A \cdot G \cdot A = A \cdot A^{-1} \cdot A \\ = A$$

System of equations : $Ax = y \rightarrow$ right hand side
↓
unknowns

(Normal equations : $\underbrace{(X^T X)}_{A} \overset{(0)}{b} = \underbrace{X^T y}_{y}$)

with G being a
generalised inverse of $(X^T X)$

OHP Picture 11

For $Ax = y$ with G being a generalized inverse of $A \Leftrightarrow A^TGA = A$
 $x = Gy$ is a solution of $Ax = y$.

Pre-multiply with $\Rightarrow Ax = AGy$
replace y with $Ax \Rightarrow Ax = \underbrace{AGA}_{=A}Ax$

Normal equations : $(X^T X) b^* = X^T y$
 G being a generalized inverse of $(X^T X)$
 $\Rightarrow b^* = G X^T y$