Quantitative genetics in structured populations

Goals:

1) Estimate the amount of genetic and environmental variation in the population.
2) Estimate the genetic and environmental effects on an individual.
3) Predict the phenotype of individuals using genome, pedigree, and environment.

Intermediate problems:

1) Genetic decompositions.
2) Estimating covariance matrices.

Conventions

g_n, $g(n)$: Scalars are italicized, non-bold, lower case letters with a subscript.

$\mathbf{g}$ vector are bold lower-case letters.

$\mathbf{V}_g$ matrices are bold upper-case letters, typically with a subscript.

$\langle \mathbf{g} \rangle$ denotes an expectation of the vector $\mathbf{g}$.

$\langle \boldsymbol{\beta}, \boldsymbol{\delta} \rangle$ denotes the inner product of two vectors.

σ_β denotes the norm of a vector.

Terms

z: Phenotypic deviation.

g: Genetic deviation.

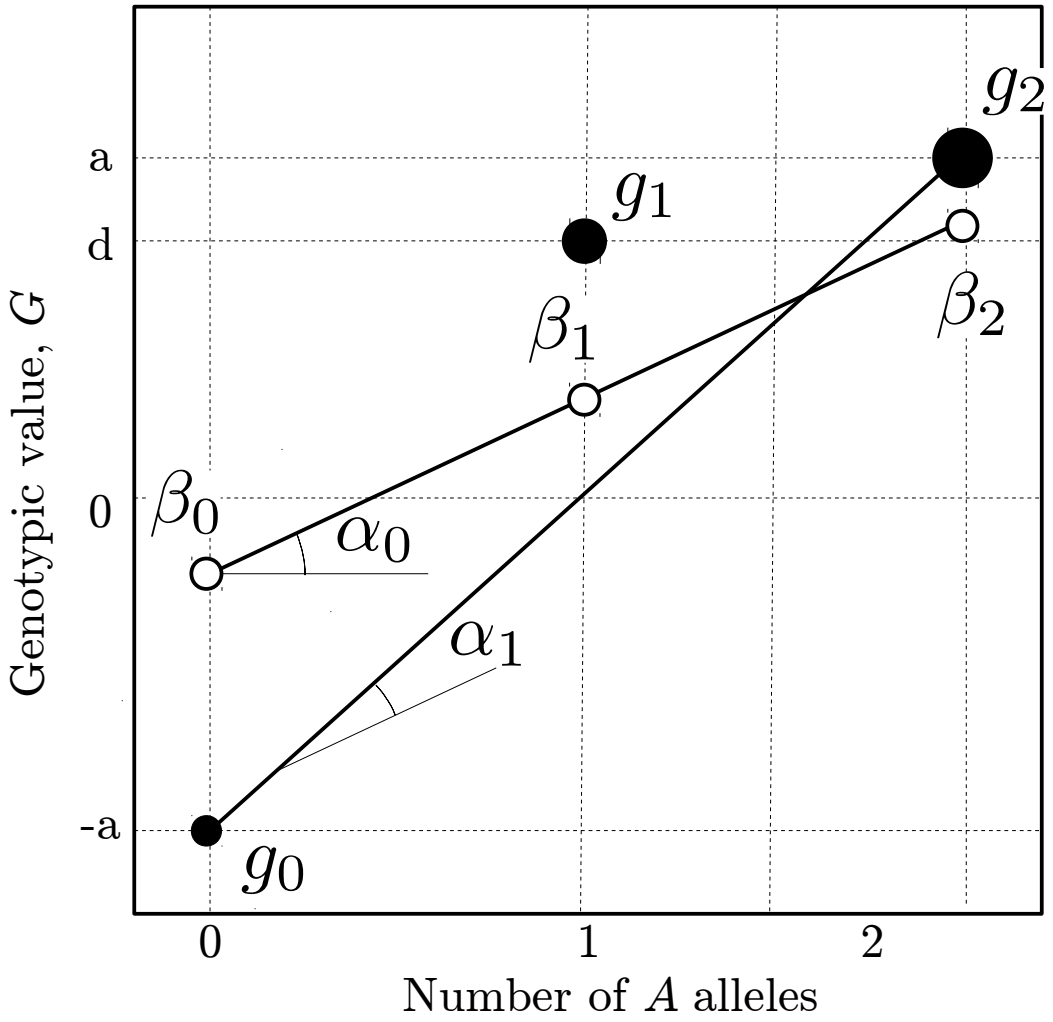
β: Additive deviation.

δ: Random mating dominance deviation

ι : Inbred dominance deviation

e: Environmental deviation.

$$\mathbf{z} = \mathbf{g} + \mathbf{e} = \boldsymbol{\beta} + \boldsymbol{\delta} + \boldsymbol{\iota} + ... + \mathbf{e}$$


g_2
a
g_1
d
β_1
β_2
β_0
0
α_0
α_1
-a
g_0
0
1
2
Genotypic value, G
Number of A alleles

The genetic deviations of an individual's phenotype at a locus.

Gene Cont. n	Gene dev. $g(n)$	$\bar{f}$	$\Pr(n)$	Additive dev., $\beta(n)$	Dominance Dev., $\delta(n)$	Dom. Dev. Inbred $\iota(n)$
0	$-a-\mu$	0	q^2	$-2\alpha p$	$-2dp^2$	0
1	$d-\mu$	0	$2pq$	$\alpha(q-p)$	$2dpq$	0
2	$a-\mu$	0	p^2	$2\alpha q$	$-2dq^2$	0
0	$-a-\mu$	1	q	$-2\alpha p$	0	$-2dp(q-p)$
1	$d-\mu$	1	0	$\alpha(q-p)$	d	$d(q-p)^2$
2	$a-\mu$	1	p	$2\alpha q$	0	$2dq(q-p)$
0	$-a-\mu$	$-q/p$	0	$-2\alpha p$	$-2d$	$-2dp(q-p)$
1	$d-\mu$	$-q/p$	$2(pq+q^2)$	$\alpha(q-p)$	0	$d(q-p)^2$
2	$a-\mu$	$-q/p$	p^2-q^2	$2\alpha q$	0	$2dq(q-p)$

$\mu = a(p-q) + 2pqd(1-\bar{f})$

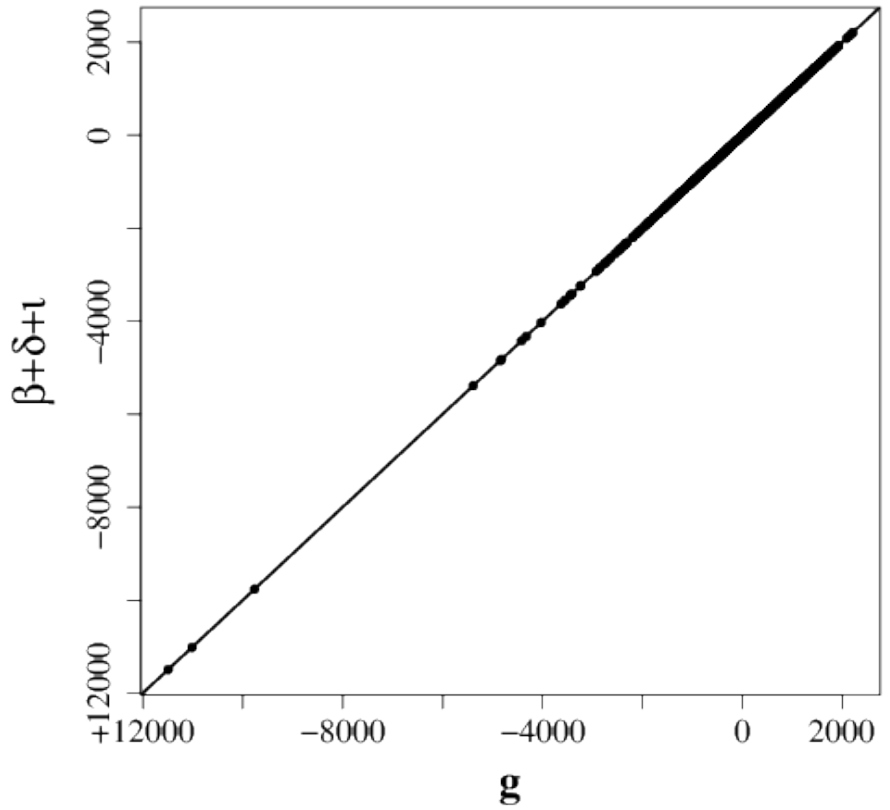
Constructing the genetic deviations

Gene Content n	Freq. (sample)	Dominance Deviation, $\delta(n)$	Inbred Dom. dev, $\iota(n)$
0	$q^2 + pq\bar{f}$	$\frac{-2dp^2}{c}\,\frac{(q-f_i q)}{(q+f_i p)}$	$\frac{-4p\alpha_1 f_i}{1+f_i}+\Delta\mu$
1	$2pq(1-\bar{f})$	$\frac{2dpq}{c}$	$\frac{2(q-p)\alpha_1 f_i}{1+f_i}+\Delta\mu$
2	$p^2 + pq\bar{f}$	$\frac{-2dq^2}{c}\,\frac{(p-f_i p)}{(p+f_i q)}$	$\frac{4q\alpha_1 f_i}{1+f_i}+\Delta\mu$

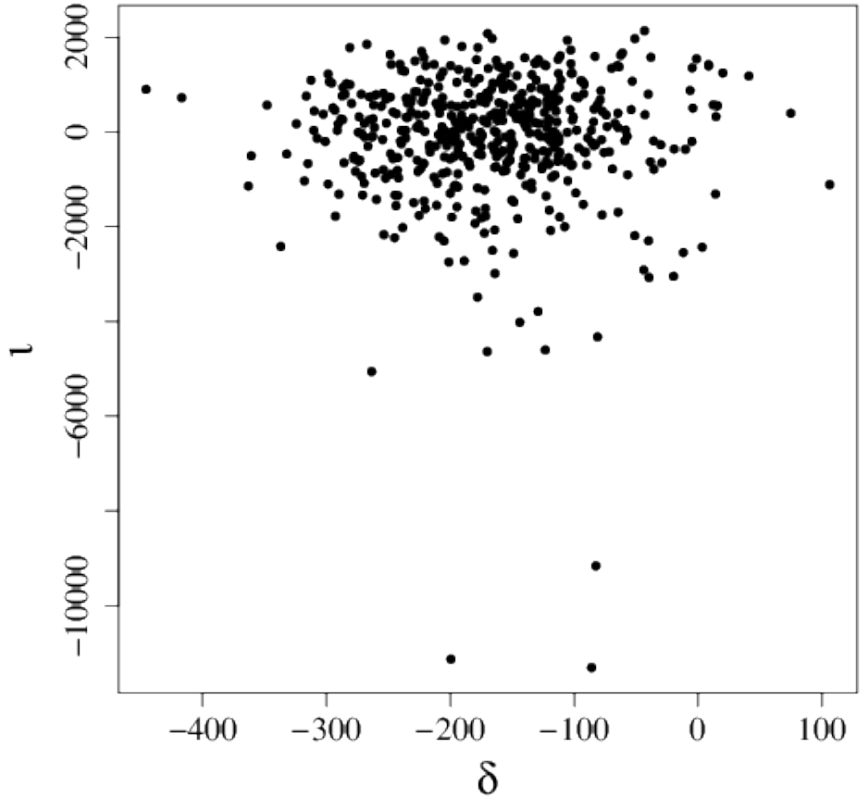
$\mathbf{n}, \mathbf{f} \rightarrow \boldsymbol{\beta}, \boldsymbol{\delta}, \boldsymbol{\iota}$

$$c = \frac{(1+f_i)pq}{(f_i p + q)(f_i q + p)} \qquad \Delta\mu = 2pqd(f - f_i)$$

Components sum to one; inbreeding and dom. are almost orthogonal.

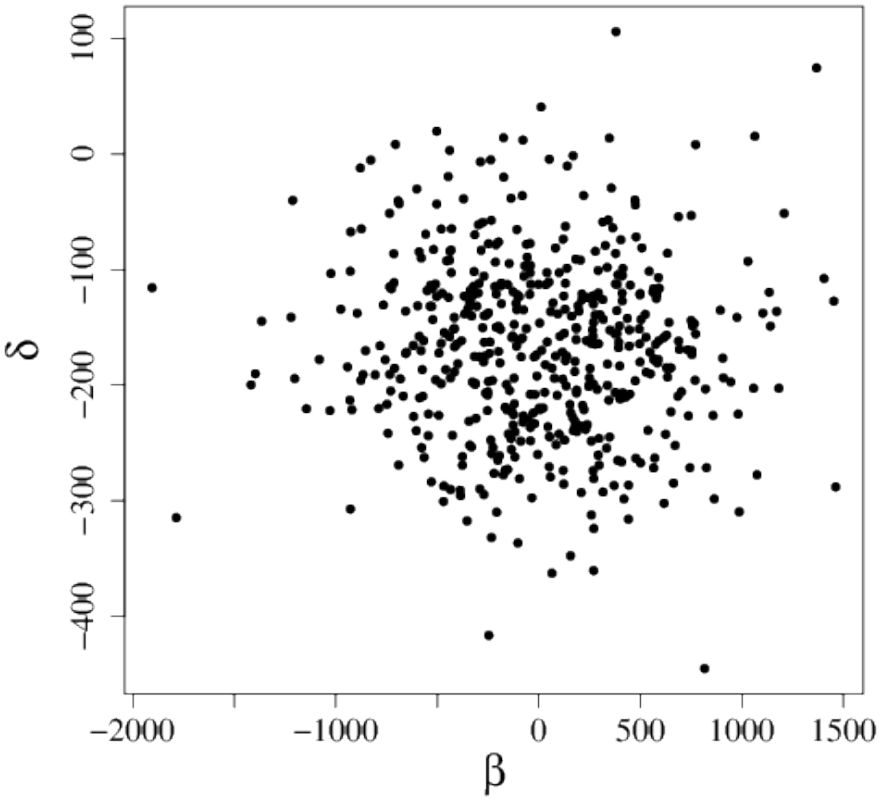


$$\mathbf{g} = \boldsymbol{\beta} + \boldsymbol{\delta} + \boldsymbol{\iota}$$

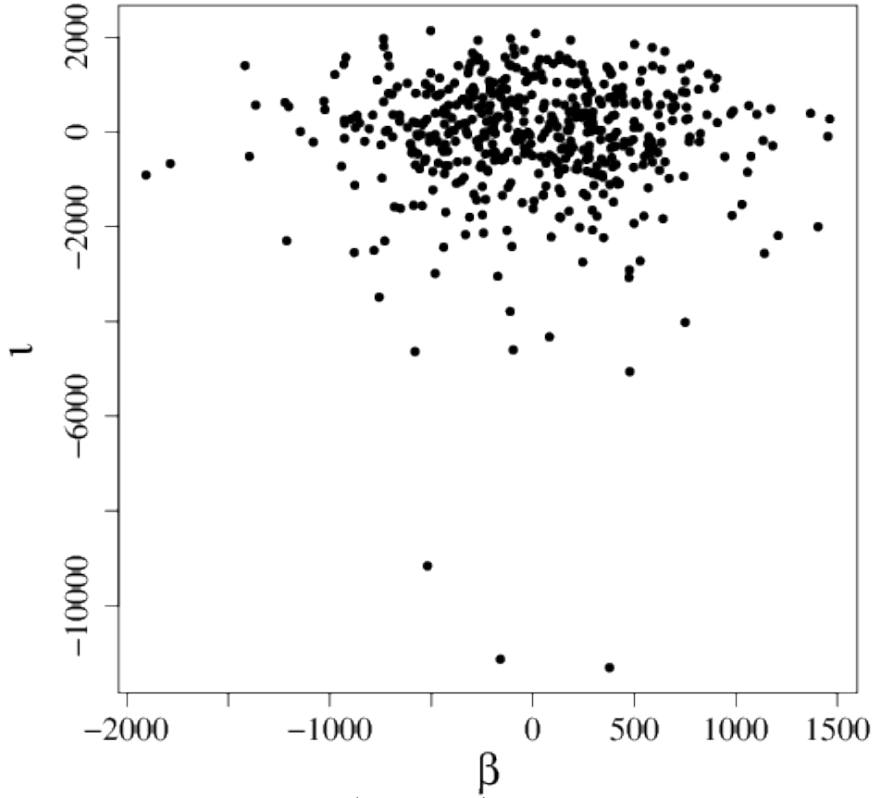


$$\langle \boldsymbol{\iota}, \boldsymbol{\delta} \rangle \approx 0$$

Additive dev. is nearly orthogonal to dominance dev. and inbred dev.



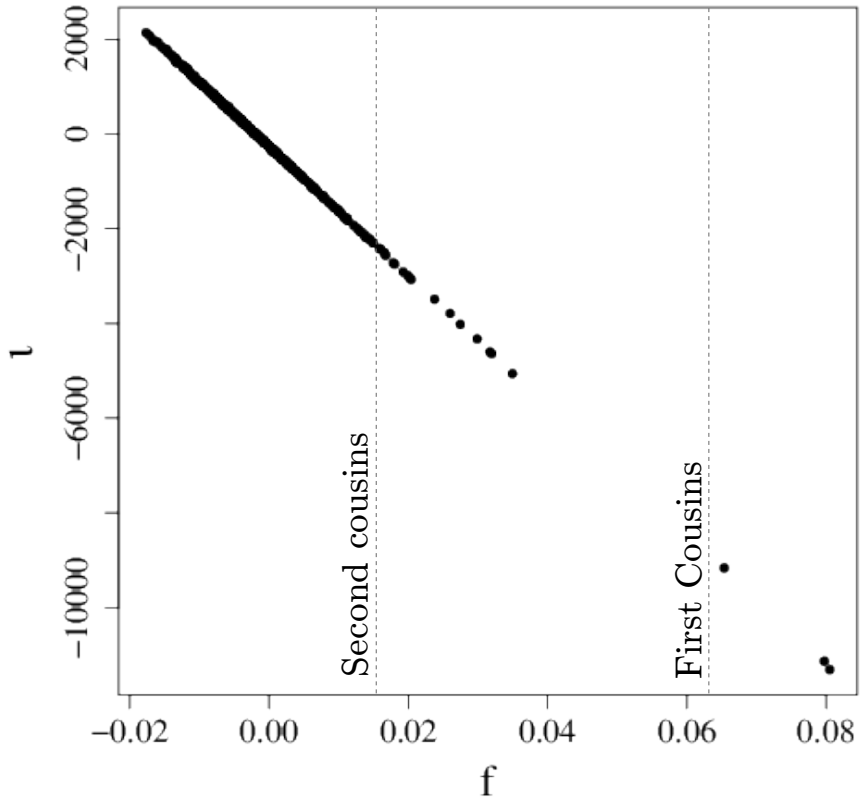
$\langle \boldsymbol{\beta}, \boldsymbol{\delta} \rangle \approx 0$



$\langle \boldsymbol{\beta}, \boldsymbol{\iota} \rangle \approx 0$

Inbred dom. dev. are a projection of dominance dev. onto inbreeding.

- Click to add Text

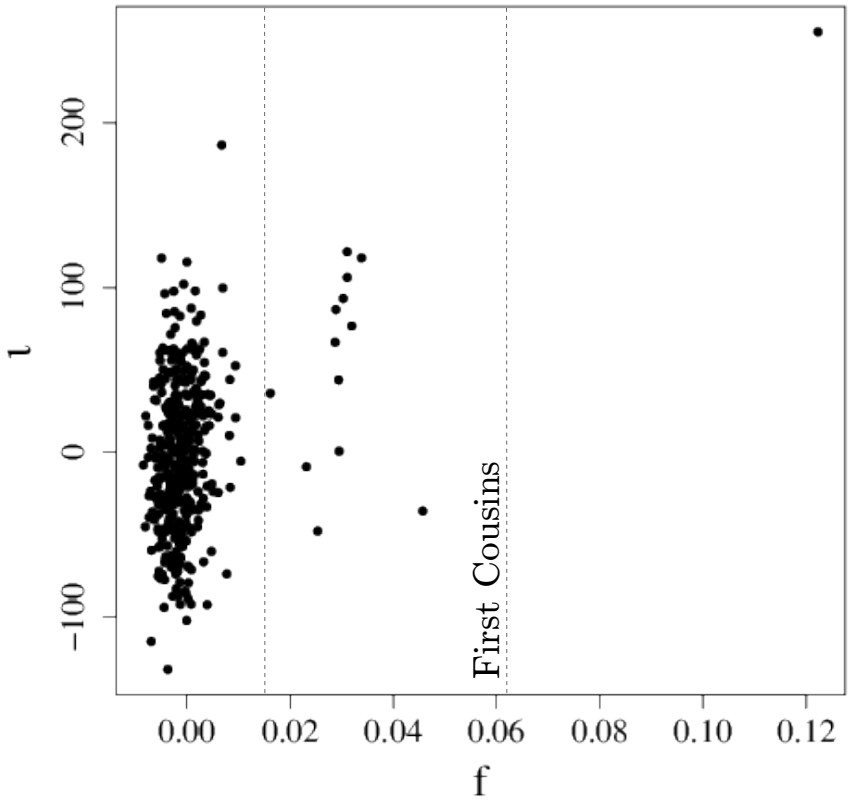


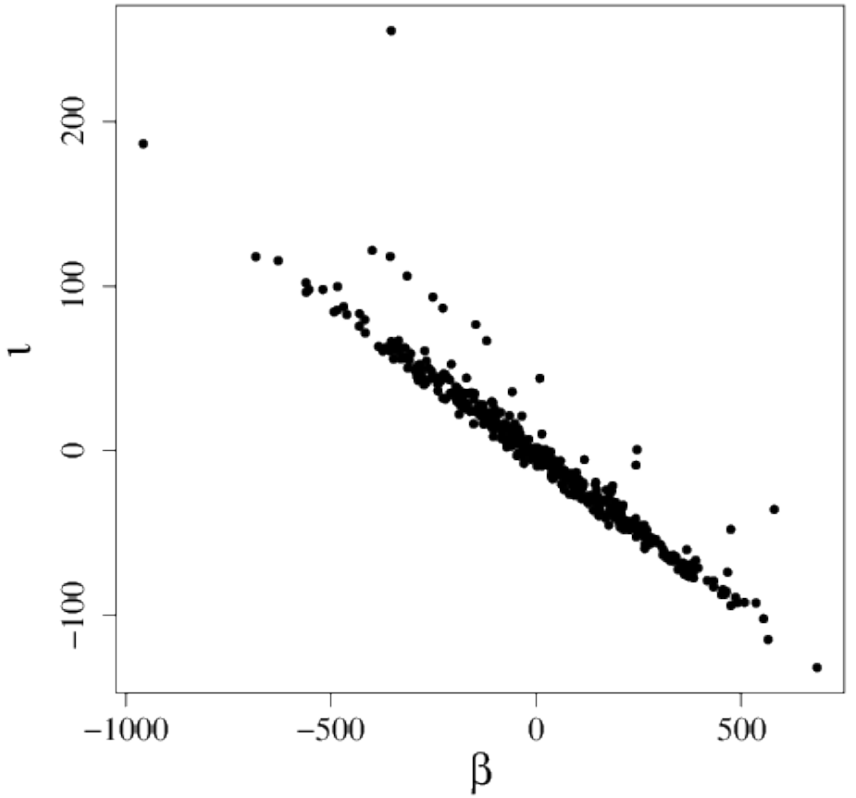
- Click to add Text

$$\frac{\langle \boldsymbol{\delta}', \mathbf{f} \rangle}{\sigma_f^2} \mathbf{f} = \boldsymbol{\iota}$$

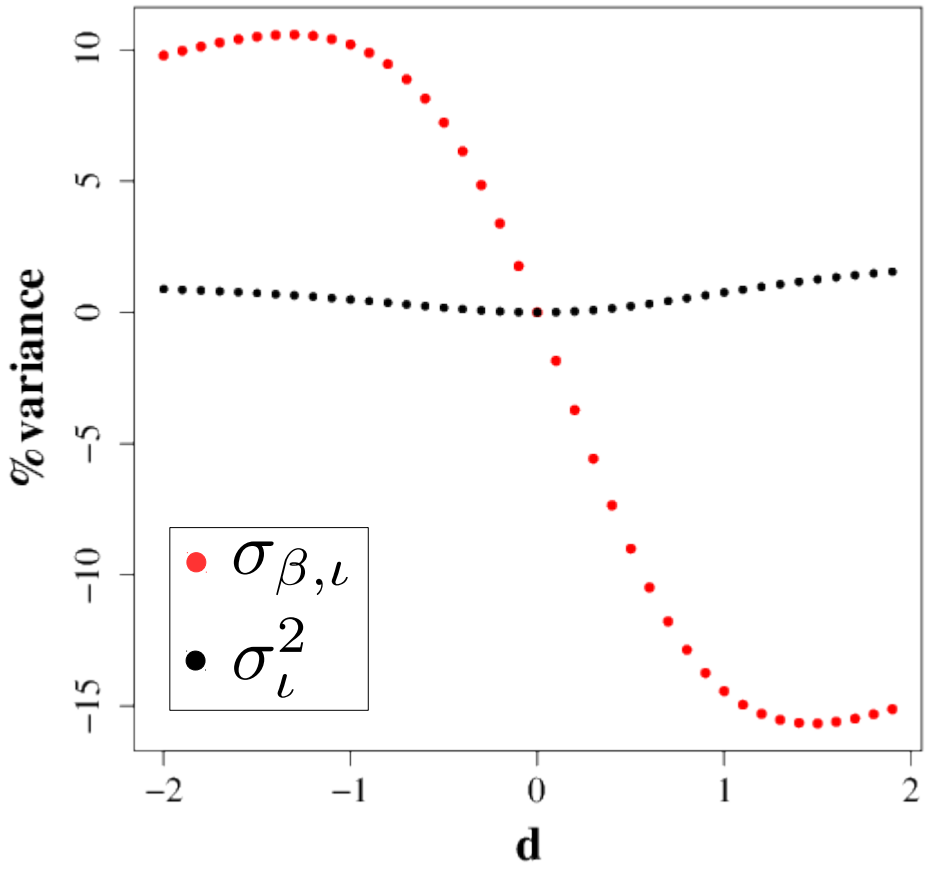
$$\boldsymbol{\delta}' - \boldsymbol{\iota} = \boldsymbol{\delta}$$

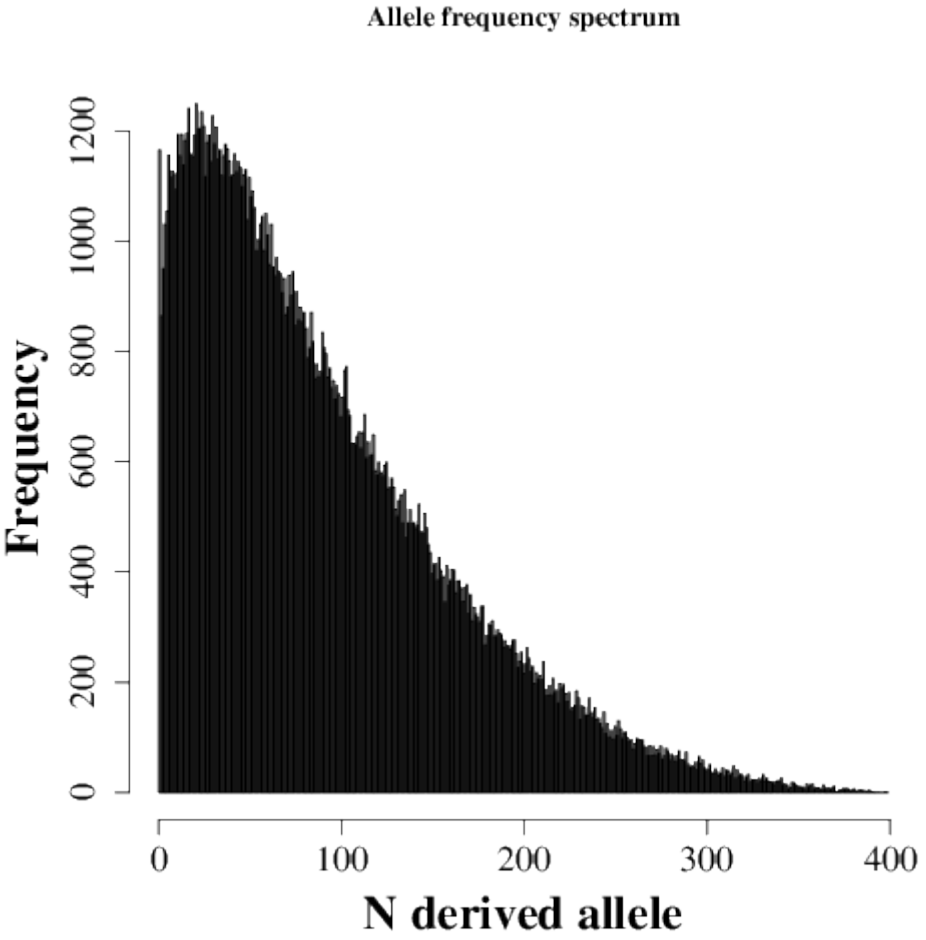
Not quite right....





Inbred dominance deviation in a randomly mating population





Covariance Matrices

$$V_z(x,y) = 2\sigma_\beta^2 V_\beta(x,y) + \sigma_\delta^2 V_\delta(x,y) + \sigma_{\beta\iota}\left(V_{\beta\iota}(x,y) + V_{\iota\beta}(x,y)\right) + \sigma_\iota^2 V_\iota(x,y)$$

$$V_\beta(x,y) = \Delta_1 + \frac{\Delta_3 + \Delta_5 + \Delta_7}{2} + \frac{\Delta_8}{4}$$

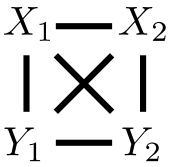
$$V_\delta(x,y) = \Delta_7 + \Delta_2$$

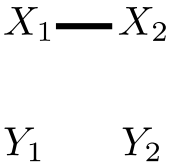
$$V_{\beta\iota}(x,y) = \Delta_1 + \Delta_3$$

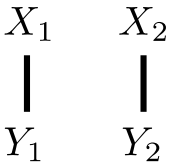
$$V_{\iota\beta}(x,y) = \Delta_1 + \Delta_5$$

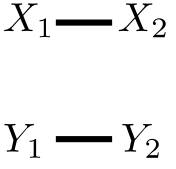
$$V_\iota(x,y) = \Delta_1$$

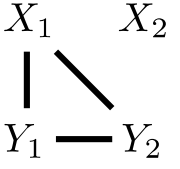
Diagonal isn't variance???

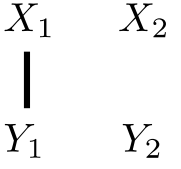
Δ_1 

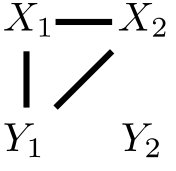
Δ_4 

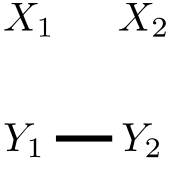
Δ_7 

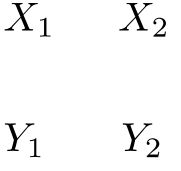
Δ_2 

Δ_5 

Δ_8 

Δ_3 

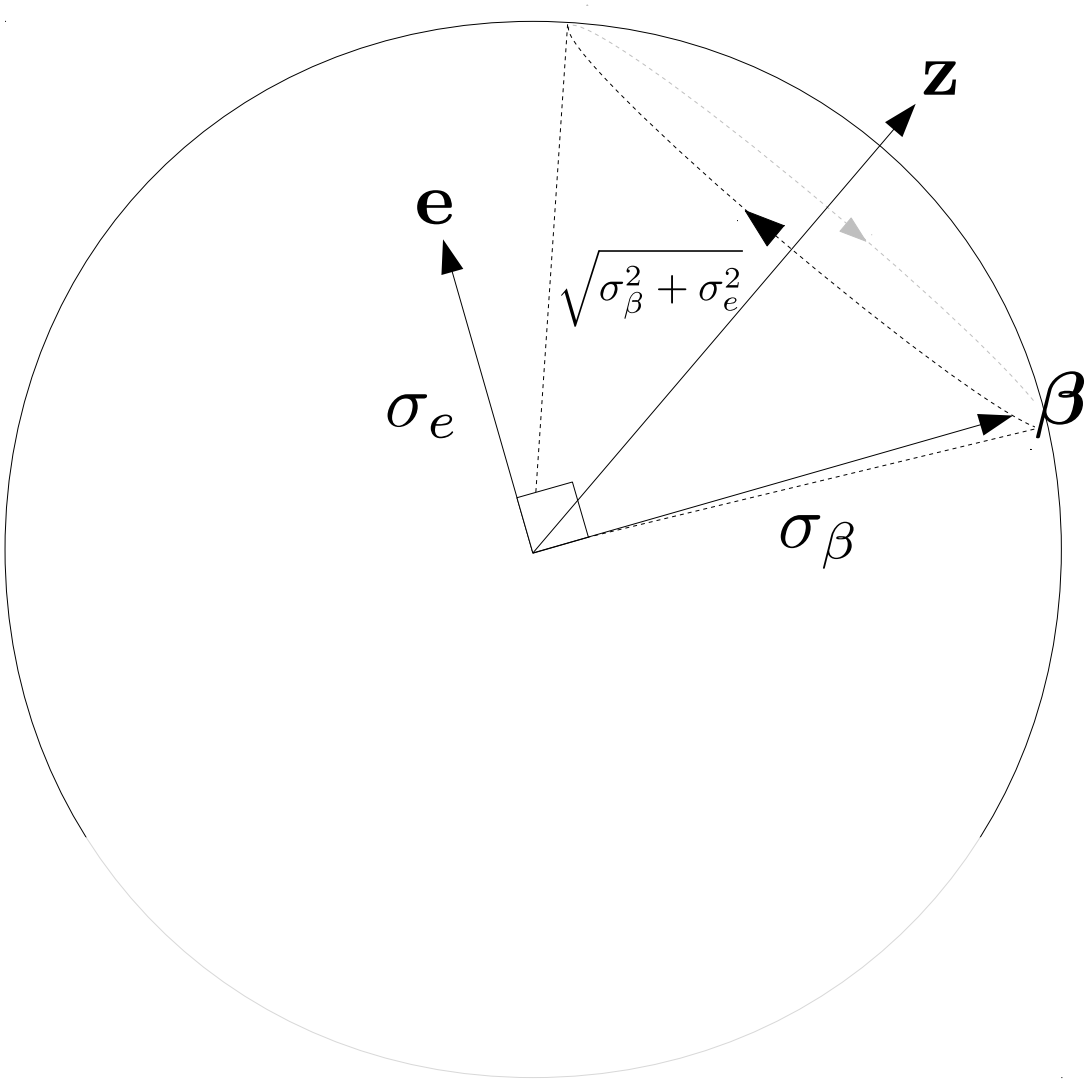
Δ_6 

Δ_9 

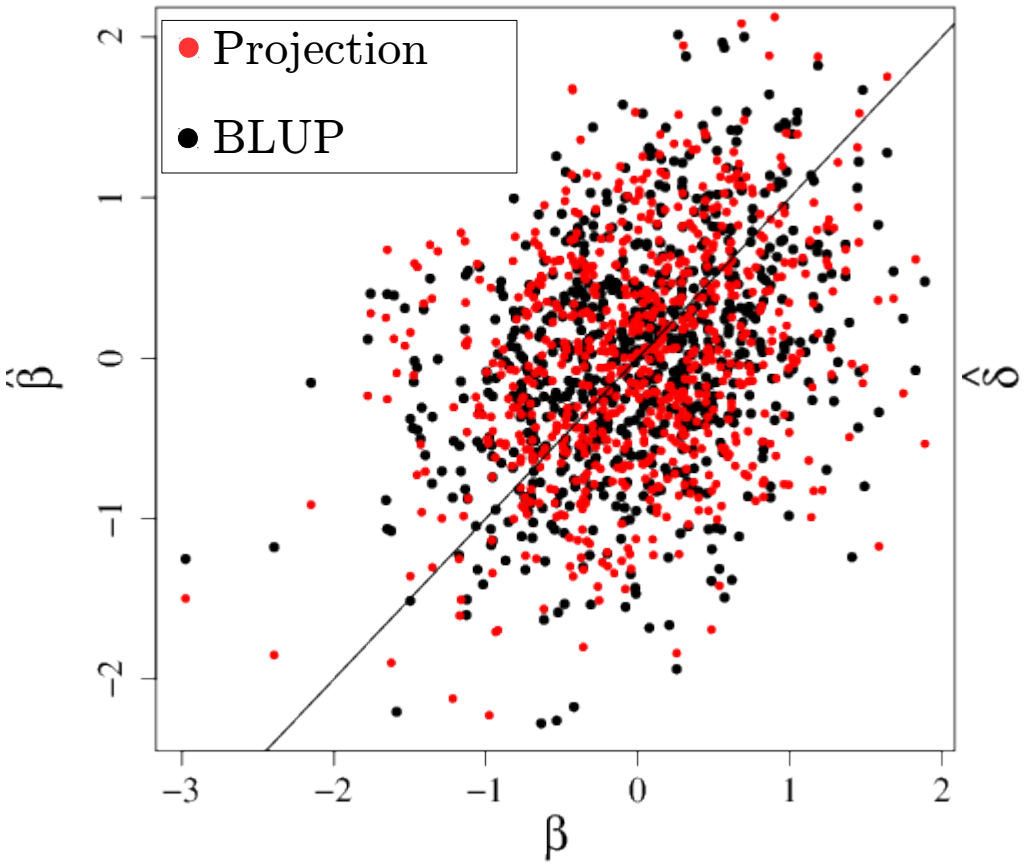
How are deviations shared between individuals?

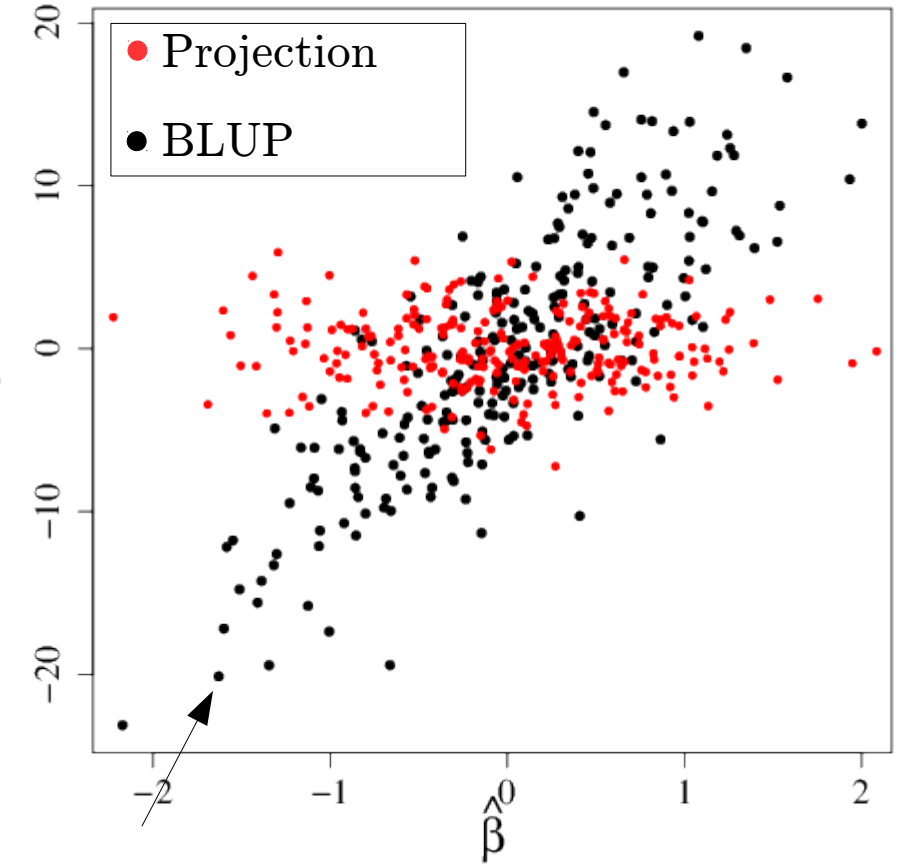
$\Pr(n_x, n_y)$	$\sum$	p_x	p_y	f_x	f_y	$\mathrm{V}_\beta(x,y)$
$\Pr(0,0)$	1	0	0	$\frac{1}{2(1-p_x)}$	$\frac{1}{2(1-p_y)}$	
$\Pr(1,0)$	1	0.5	0	$-\frac{1}{2p_x(1-p_x)}$	$\frac{1}{2(1-p_y)}$	
$\Pr(2,0)$	1	1	0	$\frac{1}{2p_x}$	$\frac{1}{2(1-p_y)}$	
$\Pr(0,1)$	1	0	0.5	$\frac{1}{2(1-p_x)}$	$-\frac{1}{2p_y(1-p_y)}$	
$\Pr(1,1)$	1	0.5	0.5	$-\frac{1}{2p_x(1-p_x)}$	$-\frac{1}{2p_y(1-p_y)}$	
$\Pr(2,1)$	1	1	0.5	$\frac{1}{2p_x}$	$-\frac{1}{2p_y(1-p_y)}$	
$\Pr(0,2)$	1	0	1	$\frac{1}{2(1-p_x)}$	$\frac{1}{2p_y}$	
$\Pr(1,2)$	1	0.5	1	$-\frac{1}{2p_x(1-p_x)}$	$\frac{1}{2p_y}$	
$\Pr(2,2)$	1	1	1	$\frac{1}{2p_x}$	$\frac{1}{2p_y}$	

Slide outlining problem



Similar performance.





BLUP variance way too high!?!?

Predicting phenotypes from genotypic deviations.

$$\begin{pmatrix} \mathbf{v}_\beta \\ \mathbf{v}_\delta \\ \mathbf{v}_\iota \end{pmatrix}^T \begin{pmatrix} \lambda_\beta \mathbf{V}_\beta^{-1} & \mathbf{0} & \lambda_{\beta,\iota} \mathbf{V}_{\beta,\iota}^{-1} \\ \mathbf{0} & \lambda_\delta \mathbf{V}_\delta^{-1} & \mathbf{0} \\ \lambda_{\beta,\iota} \mathbf{V}_{\beta,\iota}^{-1} & \mathbf{0} & \lambda_\iota \mathbf{V}_\iota^{-1} \end{pmatrix} \begin{pmatrix} \hat{\boldsymbol{\beta}} \\ \hat{\boldsymbol{\delta}} \\ \hat{\boldsymbol{\iota}} \end{pmatrix} = z_i$$

TO DO LIST

Formally derive covariance matrices.

Estimate variances/covariances.

Get the BLUP right.

Implement model of DFE.

Implement population structure in simulations.

Check against GCTA in simulations.

Check against GCTA on human data.