

Lecture 11:

Infinite-dimensional/Function-valued Traits: Covariance Functions and Random Regressions

Bruce Walsh lecture notes
Introduction to Mixed Models
SISG (Module 12), Seattle
17– 19 July 2019

Longitudinal traits

- Many classic quantitative traits are longitudinal -- measured at multiple time points --- milk yield, body size, etc.
- We have already examined the **repeated-measures design** wherein an identical trait (assumed to be unchanging) is measured multiple times.
- For most longitudinal traits, we expect the trait to change over time, such as a growth curve.
- These are function-valued traits, also called infinite-dimensional traits.
- One critical feature of such traits is that their additive variances change with t , and trait values from different time points have different correlations.

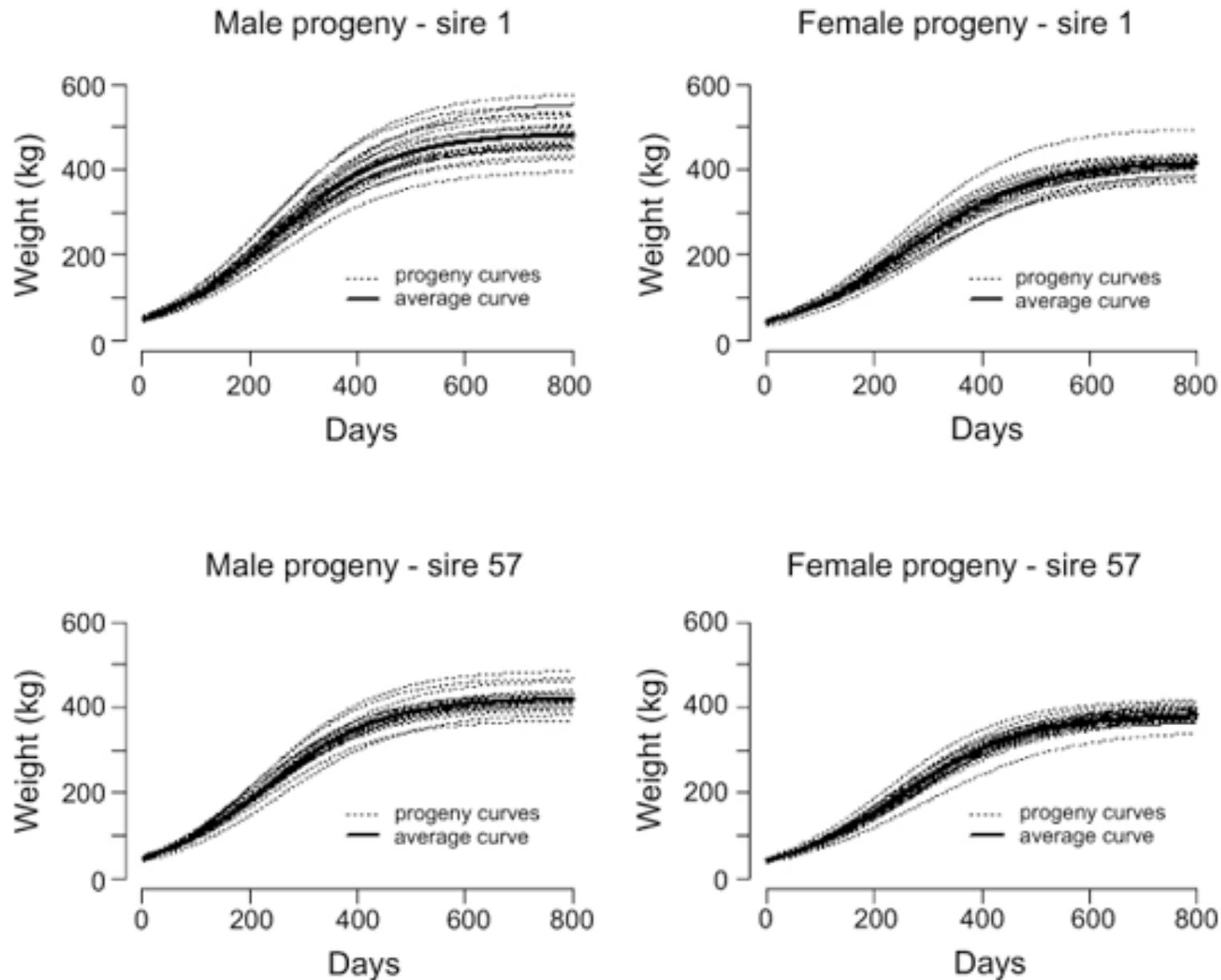


Figure 3 - Mixed logistic growth curves (---) fitted for all progeny of sire 1 (24 males and 32 females) and all progeny of sire 57 (20 males and 59 females) and associated average growth curves (—).

Norms of reaction

- The other type of function-valued trait is one indexed by some continuous environmental variable (as opposed to time), such as adult body weight as a function of temperature or grain yield as a function of total rainfall.
- The measurement of such traits generally requires replication of individuals over environments (versus the sequential evaluation of a single individual with longitudinal traits). As with $G \times E$, this can be done
 - Using clones/pure lines
 - Using family members
- Such curves are common in ecology & evolution and are called norms of reaction, and are measures of $G \times E$
 - Norms of reaction measure phenotypic plasticity --- variation that can be expressed from a fixed genotype, which is often an important adaptation in changing environments.

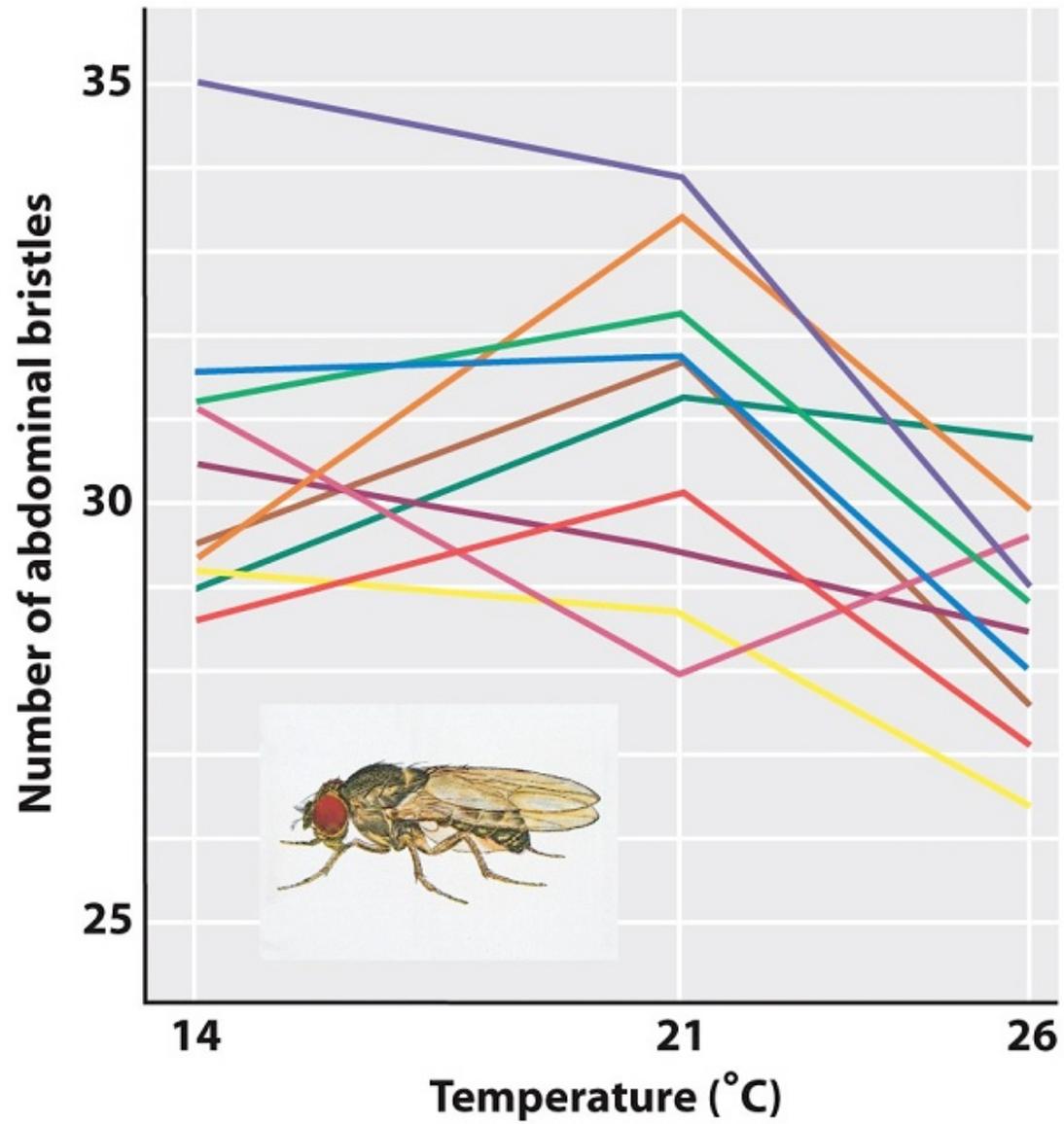


Figure 18-6
Introduction to Genetic Analysis, Ninth Edition
 © 2008 W. H. Freeman and Company

How to model such traits?

- One obvious approach is to treat the trait measured at discrete time points as a series of correlated traits.
 - Makes sense to do this for something like parity (litter number), as individuals are all measured at the same event, i.e., parity one, parity two, etc.
 - However, with a trait like a growth or some performance curve, we often expect to have different time measurements for different individuals.
 - We could either lump these into groups (reducing precision) or treat each different time/tuning variable value as a different trait (much missing data).
 - Better solution: estimate the trait covariance function, where $C(t_1, t_2) = \text{Cov}[z(t_1), z(t_2)]$ or $\text{Cov}[A(t_1), A(t_2)]$

Covariance function approach

- Kirkpatrick popularized the use of covariance functions (largely in evolutionary biology) in the mid-late 1980's.
- He noted that traits measured with respect to some continuous indexing variable (such as time or temperature) have effectively infinite dimensions, as one could (in theory) always consider finer and finer time scales.
 - Thus, rather than treat them as a (potentially) every-expanding set of discrete correlated traits, better to simply consider the covariance $C(t_1, t_2)$ between any two time points within the range of the sampled data. Note that $C(t_1, t_1)$ is the trait variance at time t_1 .
 - $C(t_1, t_2)$ is the **covariance function**, the logical extension of the covariance matrix $C(i, j)$ used for correlated traits, using continuous, rather than integer, indexes.

Covariance functions (cont)

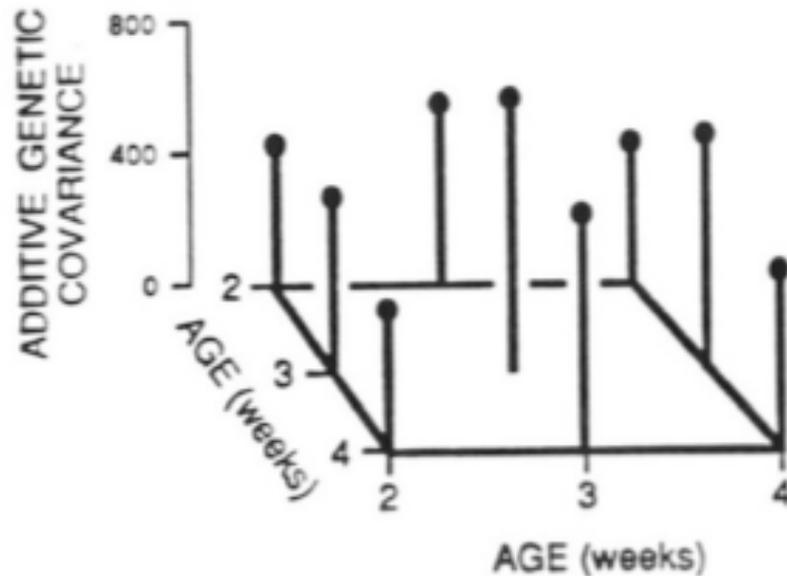
- As with any quantitative trait, the covariance between the values at two time points can be decomposed into an additive-genetic (breeding value) covariance function and a residual (or environmental) covariance function,
 - $C_Z(t_1, t_2) = C_A(t_1, t_2) + C_E(t_1, t_2)$
- The issue in the estimation of the additive covariance function is how one proceeds from an additive-covariance matrix estimate \mathbf{G} from discrete time points to a continuous function covering all possible values with the span of time sampled to estimate \mathbf{G} .
 - Basic (initial) idea: Use curve-fitting based on low-degree polynomials to use \mathbf{G} to fit a covariance function
 - This is typically done by using Legendre polynomials as the basis function.

Riska et al. (1984) data on breeding values for log(body weight)

The basic idea was illustrated by Kirkpatrick with a data set on mouse body weight measured at ages 2, 3, and 4 weeks. Riska et al. estimated the G matrix as

$$\hat{G} = \begin{matrix} & \begin{matrix} 2 & 3 & 4 \end{matrix} \\ \begin{matrix} 2 \\ 3 \\ 4 \end{matrix} & \begin{pmatrix} 436 & 522 & 424 \\ 522 & 808 & 665 \\ 424 & 665 & 558 \end{pmatrix} \end{matrix}$$

Plotting these values on a lattice at these discrete time points gives



Ideally, would like some sort of smooth curve for this data.

Towards the covariance function

- Suppose we assume the breeding value at time t (for $2 \leq t \leq 4$ weeks) is in the form of a quadratic, so that individual's i breeding value is given by
 - $A_i(t) = a_{i0} + a_{i1} t + a_{i2} t^2$.
 - Here the a_{ij} (for $0 \leq j \leq 2$) are regression coefficients unique to individual i , and are unchanging over time.
- A different individual (j) also has a quadratic regression, but with different coefficients
 - $A_j(t) = a_{j0} + a_{j1} t + a_{j2} t^2$.
 - the a_{ij} are referred to as random regression coefficients, as they are random (drawn from some distribution) OVER individuals, but constant over time WITHIN an individual.

Towards the covariance function (cont)

We can think of these random regression coefficients as being drawn from a distribution:

$$\begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} \sim \mathbf{0}, \mathbf{C}_{\mathbf{G}}, \quad \text{where} \quad \mathbf{C}_{\mathbf{G}} = \begin{pmatrix} \sigma_0^2 & \sigma_{01} & \sigma_{02} \\ \sigma_{01} & \sigma_1^2 & \sigma_{12} \\ \sigma_{02} & \sigma_{12} & \sigma_2^2 \end{pmatrix}$$

Ideally, we would like to use our estimate of \mathbf{G} to make inferences on the elements in $\mathbf{C}_{\mathbf{G}}$.

We can write the additive value in time t for individual i as $\mathbf{a}_i^T * \mathbf{t}$, where $\mathbf{a}_i^T = (a_{i0}, a_{i1}, a_{i2})$ and $\mathbf{t}^T = (1, t, t^2)$

Towards the covariance function

The regression $A_i(t) = a_{i0} + a_{i1}t + a_{i2}t^2 = \mathbf{a}_i^T \mathbf{t}$ yields the covariance function, as the value of the vector \mathbf{t} for different times are constants, giving

$$\begin{aligned}\text{Cov}[A_i(t_1), A_i(t_2)] &= \text{Cov}[\mathbf{a}_i^T \mathbf{t}_1, \mathbf{a}_i^T \mathbf{t}_2] \\ &= \mathbf{t}_1^T \text{Cov}(\mathbf{a}_i, \mathbf{a}_i) \mathbf{t}_2 \\ &= \mathbf{t}_1^T \mathbf{C}_G \mathbf{t}_2\end{aligned}$$

This is a bilinear form (the generalization of a quadratic form).

$$\begin{aligned} \text{Cov}[A(t_1), A(t_2)] &= \mathbf{t}_1^T \mathbf{C}_G \mathbf{t}_2 \\ &= (1 \quad t_1 \quad t_1^2) \begin{pmatrix} \sigma_0^2 & \sigma_{01} & \sigma_{02} \\ \sigma_{01} & \sigma_1^2 & \sigma_{12} \\ \sigma_{02} & \sigma_{12} & \sigma_2^2 \end{pmatrix} \begin{pmatrix} 1 \\ t_2 \\ t_2^2 \end{pmatrix} \end{aligned}$$

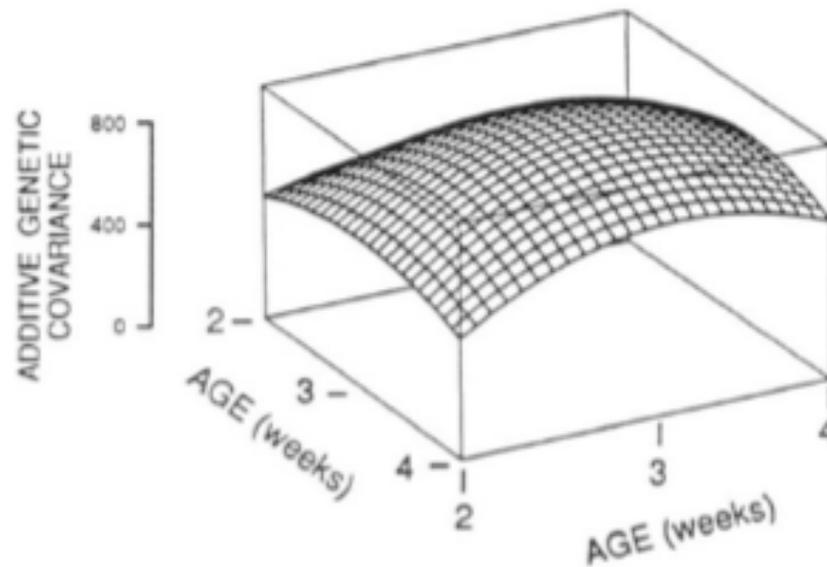
Expanding gives

$$\begin{aligned} \text{Cov}[A(t_1), A(t_2)] &= \sigma_0^2 + \sigma_{01}(t_1 + t_2) + \sigma_{02}(t_1^2 + t_2^2) \\ &\quad + \sigma_1^2 t_1 t_2 + \sigma_{12}(t_1^2 t_2 + t_1 t_2^2) + \sigma_2^2 t_1^2 t_2^2 \end{aligned}$$

More generally, fitting an m-th degree polynomial for A gives the product of two m-degree polynomials for the covariance function

$$\begin{aligned} A_i(t) &= \sum_{j=0}^m a_{ij} t^j \\ \text{Cov}[A_i(t_1), A_i(t_2)] &= \sum_{j=0}^m \sum_{k=0}^m a_{jk} t_1^j t_2^k \end{aligned}$$

Kirkpatrick estimated to covariance function for the Riska data by assuming an individual's breeding value over time can be modeled by 2nd degree polynomial. The resulting covariance function gives the following surface:



Estimated additive-genetic covariance function

Details

- Before building on these basic ideas to estimate the covariance function, some background on Legendre polynomials is required, as these are used as the basis functions (building blocks) for curve-fitting instead of the set $(1, t, t^2, \dots, t^k)$
 - Specifically, we could approximate a function $f(t)$ by the k -th degree polynomial $f(t) = \sum^k a_i t^i$.
 - Instead, we approximate it by a weighted sum of the functions $\phi_0(t), \phi_1(t), \dots, \phi_k(t)$, where $\phi_j(t)$ is a polynomial of degree j (the **Legendre polynomial of order j** , for $0 \leq j \leq k$), using $f(t) = \sum^k b_i \phi_i(t)$.

Legendre Polynomials

For curve-fitting, orthogonal polynomials are often used, where $\phi_k(t)$ denotes a k-th degree polynomial. The set of these building blocks $\phi_0(t), \phi_1(t), \dots, \phi_k(t)$.. are defined to be orthogonal in the sense that the integral of $\phi_i(t) \phi_j(t) = 0$ when i and j are not equal. We also assume they are scaled to have unit length, with the integral $\phi_i^2(t) = 1$.

For $-1 \leq t \leq 1$, the first five scaled Legendre polynomials are given by

$$\phi_0(t) = 0.7071$$

$$\phi_1(t) = 1.2247 t$$

$$\phi_2(t) = -0.7906 + 2.3717 t^2$$

$$\phi_3(t) = -2.8062 t + 4.6771 t^3$$

$$\phi_4(t) = 0.7955 - 7.9550 t^2 + 9.2808 t^4$$

$$\phi_5(t) = 4.2973 t - 20.5205 t^3 + 18.4685 t^5$$

For example, the curve $y = a + b t$ can be written as

$$y = a/(0.7071) \phi_0(t) + b/(1.2247) \phi_1(t) \text{ for } -1 \leq t \leq 1.$$

More generally, any k-th degree polynomial can be written as

$$\sum^k a_i \phi_i(t)$$

$$\phi_0(t) = 0.7071$$

$$\phi_1(t) = 1.2247 t$$

$$\phi_2(t) = -0.7906 + 2.3717 t^2$$

$$\phi_3(t) = -2.8062 t + 4.6771 t^3$$

$$\phi_4(t) = 0.7955 - 7.9550 t^2 + 9.2808 t^4$$

$$\phi_5(t) = 4.2973 t - 20.5205 t^3 + 18.4685 t^5$$

In matrix form, $\phi = \mathbf{M}\mathbf{t}$, where $\phi = \begin{pmatrix} \phi_0(t) \\ \phi_1(t) \\ \phi_2(t) \\ \phi_3(t) \\ \phi_4(t) \\ \phi_5(t) \end{pmatrix}$, $\mathbf{t} = \begin{pmatrix} 1 \\ t \\ t^2 \\ t^3 \\ t^4 \\ t^5 \end{pmatrix}$

j-th row of \mathbf{M} are the coefficients for the jth Legendre polynomial

Row 4 =
coefficients
for ϕ_4 .

$$\mathbf{M} = \begin{pmatrix} 0.7071 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1.2247 & 0 & 0 & 0 & 0 \\ -0.7906 & 0 & 2.3717 & 0 & 0 & 0 \\ 0 & -2.8062 & 0 & 4.5777 & 0 & 0 \\ \dots \rightarrow 0.7944 & 0 & -7.9950 & 0 & 9.2808 & 0 \\ 0 & 4.2973 & 0 & -20.5205 & 0 & 18.4685 \end{pmatrix}$$

1 t t² t³ t⁴ t⁵

How do we write the following 5th order polynomial in terms of Legendre polynomials?

$$y = 4 - 6x + 14x^2 + 26x^3 + 50x^4 - 110x^5$$

Note that $y = \mathbf{a}^T \mathbf{x}$, where

$$\mathbf{a} = \begin{pmatrix} 4 \\ -6 \\ 14 \\ 26 \\ 50 \\ -110 \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} 1 \\ x \\ x^2 \\ x^3 \\ x^4 \\ x^5 \end{pmatrix}$$

$$\begin{pmatrix} \phi_0(x) \\ \phi_1(x) \\ \phi_2(x) \\ \phi_3(x) \\ \phi_4(x) \\ \phi_5(x) \end{pmatrix} = \mathbf{M} \begin{pmatrix} 1 \\ x \\ x^2 \\ x^3 \\ x^4 \\ x^5 \end{pmatrix} \quad \text{implies} \quad \begin{pmatrix} 1 \\ x \\ x^2 \\ x^3 \\ x^4 \\ x^5 \end{pmatrix} = \mathbf{M}^{-1} \begin{pmatrix} \phi_0(x) \\ \phi_1(x) \\ \phi_2(x) \\ \phi_3(x) \\ \phi_4(x) \\ \phi_5(x) \end{pmatrix}$$

Giving $\mathbf{x} = \mathbf{M}^{-1}\boldsymbol{\phi}$. Since $y = \mathbf{a}^T \mathbf{x} = \mathbf{a}^T \mathbf{M}^{-1} \boldsymbol{\phi}$, weights on Legendre polynomials are given by $\mathbf{a}^T \mathbf{M}^{-1}$

Weights are given by $\mathbf{a}^T \mathbf{M}^{-1}$

```
R returns > M  
      [,1] [,2] [,3] [,4] [,5] [,6]  
[1,] 0.7071 0.0000 0.0000 0.0000 0.0000 0.0000  
[2,] 0.0000 1.2247 0.0000 0.0000 0.0000 0.0000  
[3,] -0.7906 0.0000 2.3717 0.0000 0.0000 0.0000  
[4,] 0.0000 -2.8062 0.0000 4.5777 0.0000 0.0000  
[5,] 0.7944 0.0000 -7.9950 0.0000 9.2808 0.0000  
[6,] 0.0000 4.2973 0.0000 -20.5205 0.0000 18.4685  
> t(a)%%solve(M)  
      [,1] [,2] [,3] [,4] [,5] [,6]  
[1,] 26.51006 -32.1633 24.06409 -21.01970 5.387467 -5.956087
```

Giving $y = 26.51006 * \phi_0(x) - 32.1633 * \phi_1(x) + 24.06409 * \phi_2(x) - 21.01970 * \phi_3(x) + 5.387467 * \phi_4(x) - 5.956087 * \phi_5(x)$

More generally, any k -degree polynomial $\mathbf{y} = \mathbf{a}^T \mathbf{x}_k$ can be expressed as a weighted series of the first $k+1$ Legendre polynomials ϕ_0, \dots, ϕ_k , where the weights are $\mathbf{a}^T \mathbf{M}^{-1}$. \mathbf{M} is $(k+1) \times (k+1)$, with the j th row being the coefficients on x for the j -th order Legendre polynomial.

The Covariance function in terms of Legendre polynomials

- Express the trait breeding value for individual i at time t_j by an m -th order polynomial,
 - $A_i(t_j) = \sum_k^m a_{ik} \phi_k(t_j)$, where $\mathbf{a}_i \sim \mathbf{0}, \mathbf{C}_G$
 - Define the vectors
 - $\boldsymbol{\phi}_m(t) = (\phi_0(t), \phi_1(t), \dots, \phi_m(t))^T$, which we often write as just $\boldsymbol{\phi}_m$ or $\boldsymbol{\phi}$ for brevity
 - $\mathbf{a}_i = (a_{i0}, a_{i1}, \dots, a_{im})^T$.
- Hence $A_i(t_j) = \boldsymbol{\phi}_m(t)^T \mathbf{a}_i = \mathbf{a}_i^T \boldsymbol{\phi}_m(t)$.
 - $\text{Cov}[A_i(t_1), A_i(t_2)] = \text{Cov}[\mathbf{a}_i^T \boldsymbol{\phi}_m(t_1), \mathbf{a}_i^T \boldsymbol{\phi}_m(t_2)]$
 - $\text{Cov}[A_i(t_1), A_i(t_2)] = \boldsymbol{\phi}_m(t_1)^T \mathbf{C}_G \boldsymbol{\phi}_m(t_2)$

Covariance function (cont)

- $\text{Cov}[A_i(t_1), A_i(t_2)] = \boldsymbol{\phi}_m(t_1)^T \mathbf{C}_G \boldsymbol{\phi}_m(t_2)$
- Recall for $\mathbf{t}_m = (1, t, t^2, \dots, t^m)^T$ that
 - $\boldsymbol{\phi}_m(t) = \mathbf{M}\mathbf{t}_m$, where \mathbf{M} is the $(m+1) \times (m+1)$ matrix of coefficients for the first $(m+1)$ Legendre polynomials
- Substituting in $\boldsymbol{\phi}(t) = \mathbf{M}\mathbf{t}$ yields
 - $\text{Cov}[A_i(t_1), A_i(t_2)] = \mathbf{t}_1^T \mathbf{M}^T \mathbf{C}_G \mathbf{M} \mathbf{t}_2$, or
 - $\text{Cov}[A_i(t_1), A_i(t_2)] = \mathbf{t}_1^T \mathbf{H} \mathbf{t}_2$, with $\mathbf{H} = \mathbf{M}^T \mathbf{C}_G \mathbf{M}$
 - This allows us to express the covariance function in terms t_1 and t_2 directly

From \mathbf{G} to \mathbf{C}_G

- The key component to the covariance function is the covariance matrix \mathbf{C}_G for the additive genetic random regression coefficients. How do we obtain this?
- We start with what Kirkpatrick called the “full estimate”
 - Given an estimated \mathbf{G} matrix of the trait measured at m time points, we can describe trait breeding value as an $m-1$ degree polynomial
 - This is done as a weighted combination of the first m Legendre polynomials, $\phi_0, \phi_1, \dots, \phi_{m-1}$.
 - $G_{ij} = \text{Cov}[A(t_i), A(t_j)] = \phi_m(t_i) \mathbf{C}_G \phi_m(t_j)^T$

The full estimate does an element-by-element matching of \mathbf{G} to functions of $\phi_m(t_i)$ (which are known constants) and \mathbf{C}_G .

$$\begin{aligned}
 \mathbf{G} &= \begin{pmatrix} G_{11} & \cdots & G_{1m} \\ \vdots & \ddots & \vdots \\ G_{m1} & \cdots & G_{mm} \end{pmatrix}, \quad \text{where } G_{ij} = \phi^T(t_i)\mathbf{G}_C\phi(t_j) \\
 &= \begin{pmatrix} \phi^T(t_1)\mathbf{G}_C\phi(t_1) & \cdots & \phi^T(t_1)\mathbf{G}_C\phi(t_m) \\ \vdots & \ddots & \vdots \\ \phi^T(t_m)\mathbf{G}_C\phi(t_1) & \cdots & \phi^T(t_m)\mathbf{G}_C\phi(t_m) \end{pmatrix} \\
 &= \begin{pmatrix} \phi^T(t_1) \\ \vdots \\ \phi^T(t_m) \end{pmatrix} \mathbf{G}_C \begin{pmatrix} \phi(t_1) \\ \vdots \\ \phi(t_m) \end{pmatrix} = \Phi^T \mathbf{G}_C \Phi
 \end{aligned}$$

$$\mathbf{G} = \mathbf{\Phi}^T \mathbf{G}_C \mathbf{\Phi} \quad \text{implies} \quad \mathbf{G}_C = \left(\mathbf{\Phi}^T \right)^{-1} \mathbf{G} \mathbf{\Phi}_1^{-1}$$

where

$$\mathbf{\Phi}^T = \begin{pmatrix} \phi^T(t_1) \\ \phi^T(t_2) \\ \vdots \\ \phi^T(t_m) \end{pmatrix} = \begin{pmatrix} \phi_0(t_1) & \phi_1(t_1) & \cdots & \phi_{m-1}(t_1) \\ \phi_0(t_2) & \phi_1(t_2) & \cdots & \phi_{m-1}(t_2) \\ \vdots & \vdots & \cdots & \vdots \\ \phi_0(t_m) & \phi_1(t_m) & \cdots & \phi_{m-1}(t_m) \end{pmatrix}$$

Note that $\mathbf{\Phi}$ is a matrix of constants --- the Legendre polynomials evaluated at the sample time points. Note that time points are scaled to be within $(-1, 1)$, so ordering time on the original scale as $T_1 < \dots < T_m$, scaled values are given by $t_i = 2(T_i - T_1)/(T_m - T_1) - 1$

Example: Riska's data

$$\mathbf{G} = \begin{pmatrix} 436.0 & 522.3 & 424.2 \\ 522.3 & 808.0 & 664.7 \\ 424.2 & 664.7 & 558.0 \end{pmatrix}$$

$$\begin{aligned} \Phi^T &= \begin{pmatrix} \phi_0(-1) & \phi_1(-1) & \phi_2(-1) \\ \phi_0(0) & \phi_1(0) & \phi_2(0) \\ \phi_0(1) & \phi_1(1) & \phi_2(1) \end{pmatrix} \begin{array}{l} \leftarrow \dots 2 \text{ weeks, } t = -1 \\ \leftarrow \dots 3 \text{ weeks, } t = 0 \\ \leftarrow \dots 4 \text{ weeks, } t = 1 \end{array} \\ &= \begin{pmatrix} 0.7071 & -1.2247 & 1.5811 \\ 0.7071 & 0 & -0.7906 \\ 0.7071 & 1.2247 & 1.5811 \end{pmatrix} \end{aligned}$$

$$\mathbf{G}_C = \left(\Phi^T \right)^{-1} \mathbf{G} \Phi^{-1} = \begin{pmatrix} 1348.1 & 66.6 & -111.7 \\ 66.6 & 24.2 & -14.0 \\ -111.7 & -14.0 & 14.5 \end{pmatrix}$$

```

> G<-matrix(c(436.0,522.3,424.2,522.3,808.0,664.7,424.2,664.7,558.0),nrow=3)
> G
      [,1] [,2] [,3]
[1,] 436.0 522.3 424.2
[2,] 522.3 808.0 664.7
[3,] 424.2 664.7 558.0
> Phi<-matrix(c(0.7071,0.7071,0.7071,-1.2247,0,1.2247,1.5811,-0.7906,1.5811),nrow=3)
> Phi
      [,1] [,2] [,3]
[1,] 0.7071 -1.2247 1.5811
[2,] 0.7071 0.0000 -0.7906
[3,] 0.7071 1.2247 1.5811
> solve(Phi)%*% G %*% solve(t(Phi))
      [,1] [,2] [,3]
[1,] 1348.14866 66.55166 -111.68492
[2,] 66.55166 24.26844 -14.01216
[3,] -111.68492 -14.01216 14.50677

```

The resulting covariance function becomes

$$\begin{aligned} \text{Cov}(t_1, t_2) &= \phi^T(t_1) \mathbf{G}_C \phi(t_2) \\ &= (\phi_0(t_1) \quad \phi_1(t_1) \quad \phi_2(t_1)) \begin{pmatrix} 1348.1 & 66.6 & -111.7 \\ 66.6 & 24.2 & -14.0 \\ -111.7 & -14.0 & 14.5 \end{pmatrix} \begin{pmatrix} \phi_0(t_2) \\ \phi_1(t_2) \\ \phi_2(t_2) \end{pmatrix} \end{aligned}$$

This bilinear form expresses the covariance function in terms of the Legendre polynomials. Usually we would like to express this as a polynomial in t_1 & t_2 :

One could do this by first substituting in the polynomial form for $\phi_i(t)$, expanding and collecting terms. However, much easier to do this in matrix form. Recall the coefficient matrix \mathbf{M} from earlier in the notes, where $\phi = \mathbf{M}t$. Writing the covariance function as $\phi_1^T \mathbf{G}_C \phi_2 = (\mathbf{M}t_1)^T \mathbf{G}_C (\mathbf{M}t_2) = t_1^T \mathbf{M}^T \mathbf{G}_C \mathbf{M} t_2 = t_1^T \mathbf{H} t_2$, where $\mathbf{H} = \mathbf{M}^T \mathbf{G}_C \mathbf{M}$.

The covariance function becomes $\mathbf{t}_1^T \mathbf{H} \mathbf{t}_2$, with $\mathbf{H} = \mathbf{M}^T \mathbf{C}_G \mathbf{M}$

Since the first three Legendre polynomials are used, \mathbf{M} is 3 x 3

$$\mathbf{M} = \begin{pmatrix} 0.7071 & 0 & 0 \\ 0 & 1.2247 & 0 \\ -0.7906 & 0 & 2.3717 \end{pmatrix}$$

$\mathbf{H} = \mathbf{M}^T \mathbf{C}_G \mathbf{M}$ gives

$$\mathbf{H} = \begin{pmatrix} 808.0 & 71.2 & -214.5 \\ 71.2 & 36.4 & -40.7 \\ -214.5 & -40.7 & 81.6 \end{pmatrix}$$

Expanding this out gives

$$\begin{aligned} \text{Cov}(A_1, A_2) = & 808 + 71.2(t_1 + t_2) + 36.4 t_1 t_2 \\ & - 40.7(t_1^2 t_2 + t_1 t_2^2) - 215.0(t_1^2 + t_2^2) \\ & + 81.6 t_1^2 t_2^2 \end{aligned}$$

More generally, the coefficient on $t_1^{i-1} t_2^{j-1}$ in the covariance expansion is given by H_{ij} . -- the (i,j)-th element of \mathbf{H} .

The Eigenstructure of \mathbf{C}_G

- The variance-covariance matrix \mathbf{C}_G of the random regression coefficients is extremely information on the nature of variation for the function-valued trait.
- The function-valued analogue of the eigenvector is the eigenfunction, which also has an associated eigenvalue. Akin to the eigenvector associated with the largest eigenvalue accounting for the largest single direction of variation, the eigenfunction associated with the largest eigenvalue is the functional curve associated with the most variation.
- The eigenvalues of \mathbf{C}_G are the same as those for the covariance function, while the associated eigenvectors of \mathbf{C}_G give the weights on the orthogonal polynomials that recover the eigenfunctions of the covariance function.

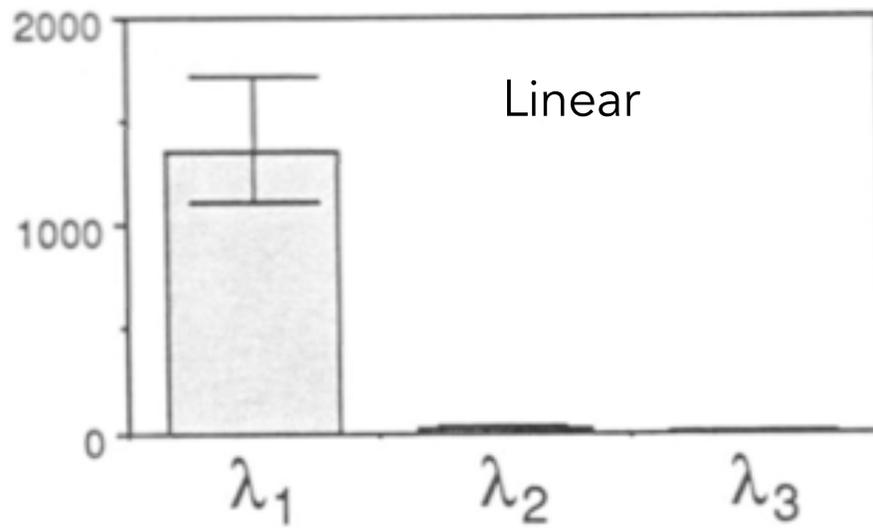
Back to Riska's data

$$\mathbf{G}_C = \left(\Phi^T\right)^{-1} \mathbf{G} \Phi^{-1} = \begin{pmatrix} 1348.1 & 66.6 & -111.7 \\ 66.6 & 24.2 & -14.0 \\ -111.7 & -14.0 & 14.5 \end{pmatrix}$$

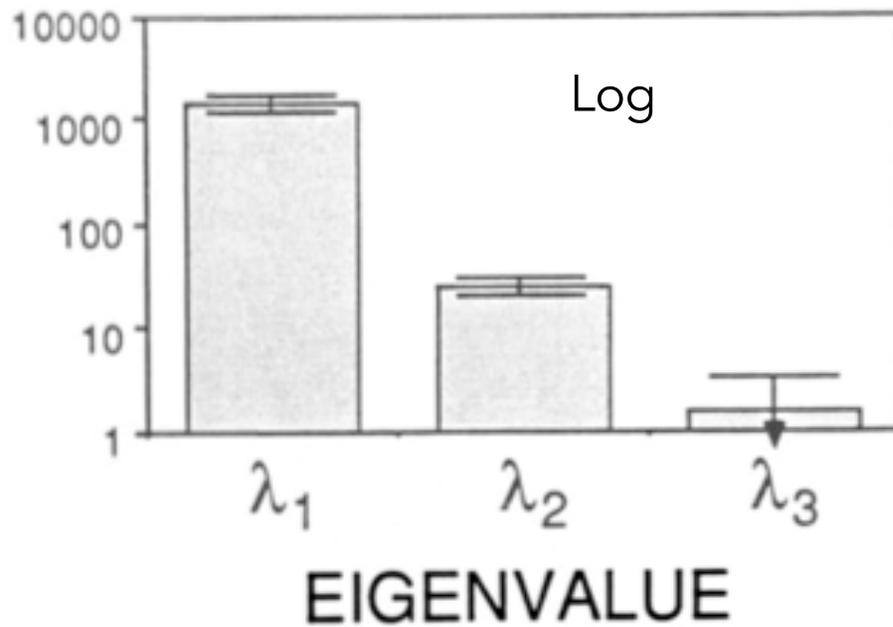
```
> eigen(CG)
$values
[1] 1360.844364  24.544765  1.534744

$vectors
      [,1]      [,2]      [,3]
[1,] -0.99526560  0.07934234 -0.05613532
[2,] -0.05042796 -0.91529538 -0.39961406
[3,]  0.08308671  0.39489133 -0.91496308
```

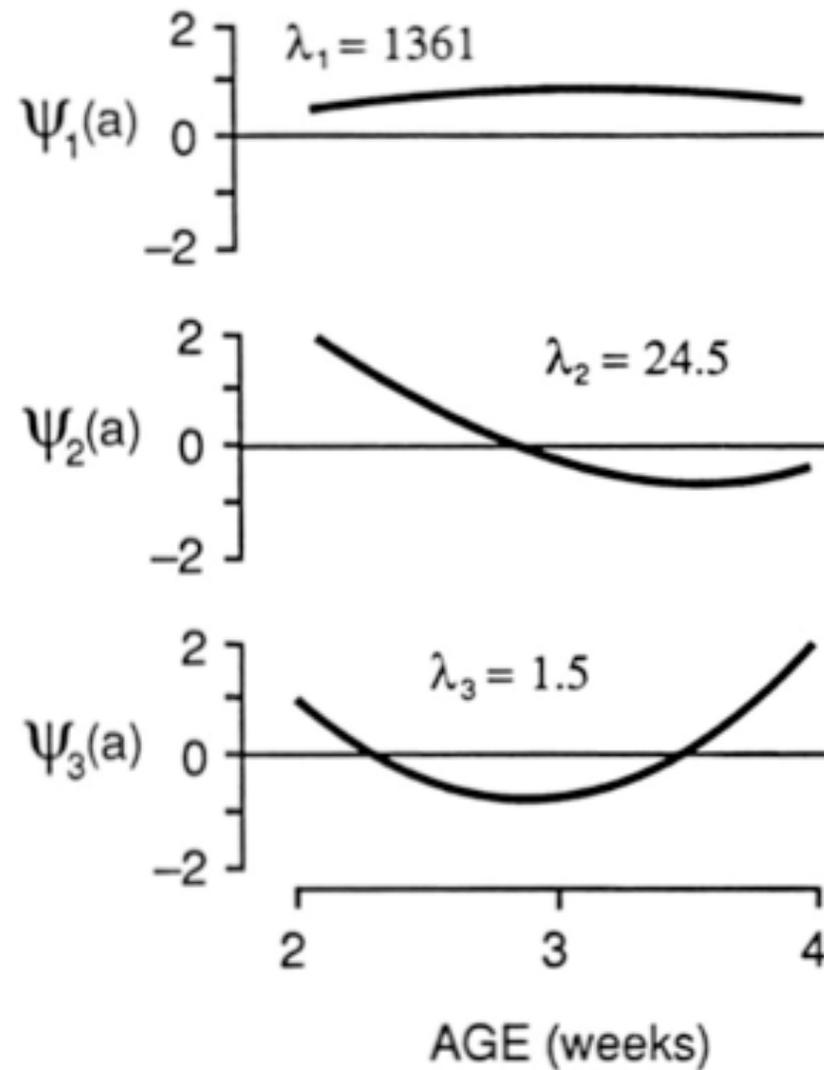
First eigenvector



CG has a dominant eigenvalue --- most of the variation in the breeding value for growth is along one curve



Associated eigenfunctions for C_G for the Riska dataset



Eigenfunctions of \mathbf{C}_G

- If \mathbf{e}_i denotes the eigenvector associated with the i th eigenvalue λ_i of \mathbf{C}_G , then for the covariance function
 - λ_i is the i th eigenvalue
 - associated eigenfunction is $\phi_m(t)^T \mathbf{e}_i$
 - $= e_{i1}\phi_0(t) + e_{i2}\phi_1(t) + \dots + e_{im}\phi_{m-1}(t)$
 - Since $\phi = \mathbf{M}t$, we have $(\mathbf{M}t)^T \mathbf{e}_i = t^T (\mathbf{M}^T \mathbf{e}_i)$, giving the weights on $(1, t, t^2, \dots, t^{m-1})$ as $\mathbf{M}^T \mathbf{e}_i$
 - For Riska's data, the leading eigenfunction is
 - $\psi_1(t) = 0.7693 - 0.0617 t - 0.1971 t^2$

Eigenfunctions: $\psi_i(t) = t^T (M^T e_i)$

$$M = \begin{pmatrix} 0.7071 & 0 & 0 \\ 0 & 1.2247 & 0 \\ -0.7906 & 0 & 2.3717 \end{pmatrix}$$

$$e_1 = \begin{pmatrix} 0.995 \\ 0.050 \\ -0.083 \end{pmatrix}, \quad e_2 = \begin{pmatrix} -0.079 \\ 0.915 \\ -0.395 \end{pmatrix}, \quad e_3 = \begin{pmatrix} 0.056 \\ 0.400 \\ 0.915 \end{pmatrix}$$

$$M^T e_1 = \begin{pmatrix} 0.769 \\ 0.062 \\ -0.197 \end{pmatrix}, \quad M^T e_2 = \begin{pmatrix} 0.256 \\ 1.121 \\ -0.937 \end{pmatrix}, \quad M^T e_3 = \begin{pmatrix} -0.684 \\ 0.490 \\ 2.170 \end{pmatrix}$$

$$\psi_2(t) = 0.256 + 1.121*t - 0.937*t^2$$

$$\psi_3(t) = -0.684 + 0.490*t + 2.170*t^2$$

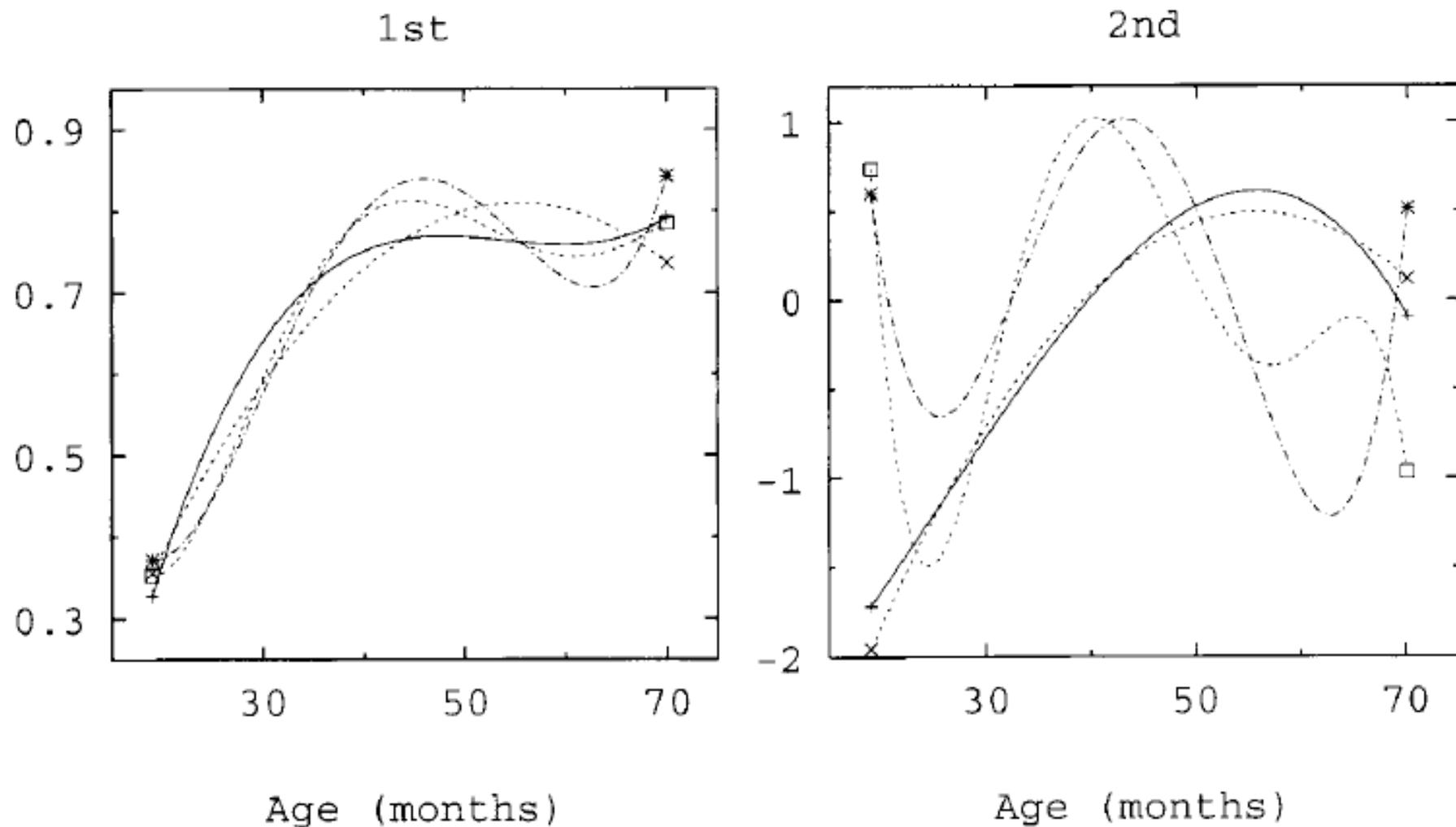


Figure 3. Estimated first and second eigenfunction of the genetic covariance function, for orders of polynomial fit of 3 (\times), 4 ($+$), 5 ($*$) and 6 (\square), respectively (rank 3 estimates of the coefficient matrices).

Over-fitting G_c ?

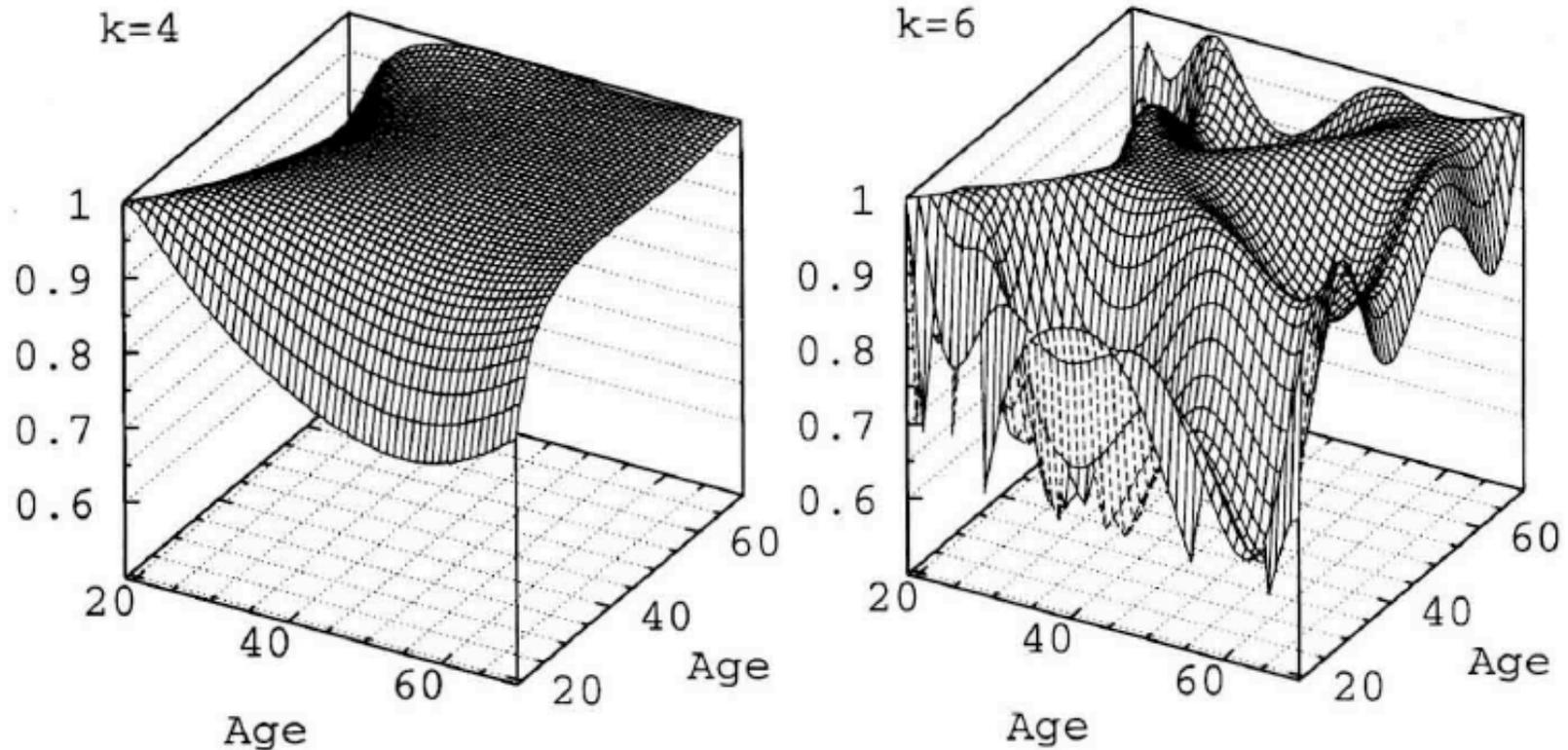


Figure 5. Estimates of genetic correlations for orders of polynomial fit (k) of 4 and 6.

Meyer's data showing how increasing the degree of polynomial used results in over-fitting. In her words: "surfaces become 'wiggly' "

Reduced estimation of C_G

- While the full estimate ($\text{rank } C_G = \text{rank of observed } G$) is (relatively) straightforward, this likely results in an overfit of the data, as the covariance function is forced to exactly fit the observed values for all t_1, t_2 , some of which are sampling noise
 - Results in a less smooth covariance function than one based on using a reduced dimension.
 - Kirkpatrick originally suggested a least-squares approach, while Meyer & Hill suggested a REML-based approach
 - Key breakthrough, first noticed by Goddard, and fully developed by Meyer, is the connection between covariance functions and random regressions.
 - This should not be surprising given that we started with random regressions to motivate covariance functions.
 - The key is that standard BLUP approaches (for multivariate traits) can be used for random regressions.

Mixed-Models (BLUPs) for Longitudinal traits

- Simplest setting is the repeatability model, the trait breeding and residual (permanent environmental) values are assumed constant over time. The j th observation on i is
 - $y_{ij} = u + a_i + pe_i + e_{ij}$
 - $a \sim 0, \text{Var}(A)A$
- At the other extreme is the multiple-trait approach, where each sampled time point is considered as a separate, but correlated, trait. Here y_{ij} is the j th “trait” (sampled time point) for individual i .
 - $y_{ij} = u + a_{ij} + e_{ij}$
 - $a \sim 0, G \times A$
- In the middle are random-regressions, where for the j th observation (time t_j) on individual i is
 - $y_{ij} = u + \sum_k^n a_{ik} \phi_k(t_j) + \sum_k^m pe_{ik} \phi_k(t_j) + e_{ij}$
 - $a_i \sim 0, C_G$ and $p_i \sim 0, C_E$

The repeatability model

- The repeatability model assumes that the trait is unchanging between observations, but multiple observations (records) are taken over time to smooth out sampling noise (e)
- Such a record for individual k has three components
 - Breeding value a_k
 - Common (**permanent**) environmental value p_k
 - Residual value for i th observation e_{ki}
- Resulting observation is thus
 - $z_{ki} = \mu + a_k + p_k + e_{ki}$
- The **repeatability** of a trait is $r = (\sigma_A^2 + \sigma_p^2) / \sigma_z^2$
- Resulting variance of the residuals is $\sigma_e^2 = (1-r) \sigma_z^2$

Mixed-model

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{a} + \mathbf{Z}\mathbf{p} + \mathbf{e}$$

$$\begin{pmatrix} \mathbf{a} \\ \mathbf{p} \\ \mathbf{e} \end{pmatrix} \sim \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \end{pmatrix}, \begin{pmatrix} \sigma_A^2 \cdot \mathbf{A} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \sigma_p^2 \cdot \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \sigma_e^2 \cdot \mathbf{I} \end{pmatrix}$$

Mixed-model equations

$$\begin{pmatrix} \mathbf{X}^T \mathbf{X} & \mathbf{X}^T \mathbf{Z} & \mathbf{X}^T \mathbf{Z} \\ \mathbf{Z}^T \mathbf{X} & \mathbf{Z}^T \mathbf{Z} + \lambda_A \mathbf{A}^{-1} & \mathbf{Z}^T \mathbf{Z} \\ \mathbf{Z}^T \mathbf{X} & \mathbf{Z}^T \mathbf{Z} & \mathbf{Z}^T \mathbf{Z} + \lambda_u \mathbf{I} \end{pmatrix} \begin{pmatrix} \hat{\boldsymbol{\beta}} \\ \hat{\mathbf{a}} \\ \hat{\mathbf{p}} \end{pmatrix} = \begin{pmatrix} \mathbf{X}^T \mathbf{y} \\ \mathbf{Z}^T \mathbf{y} \\ \mathbf{Z}^T \mathbf{y} \end{pmatrix}$$

where

$$\lambda_A = \frac{\sigma_e^2}{\sigma_A^2} = \frac{1-r}{h^2} \quad \text{and} \quad \lambda_u = \frac{\sigma_e^2}{\sigma_p^2} = \frac{1-r}{r-h^2}$$

The multiple-trait model

- With a clearly discrete number of stages (say k), a longitudinal trait could be modeled as k correlated traits, so that individual i has values $y_{i1}, y_{i2}, \dots, y_{ik}$.
- In this case, there is no need for permanent environmental effects, as these now appear in correlations among the residuals, the within-individual environmental correlations (which are estimated by REML).
- This can be put into standard Mixed Model equations by simply “stacking” the vectors for each trait to create one vector for each random effect.

For trait j ($1 \leq j \leq k$), the mixed model becomes

$$\mathbf{y}_j = \mathbf{X}_j\boldsymbol{\beta}_j + \mathbf{Z}_j\mathbf{a}_j + \mathbf{e}_j$$

$$\begin{pmatrix} \mathbf{a}_j \\ \mathbf{e}_j \end{pmatrix} \sim \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \end{pmatrix}, \begin{pmatrix} \sigma_{A_j}^2 \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \sigma_{e_j}^2 \mathbf{I} \end{pmatrix}$$

We can write this as $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{a} + \mathbf{e}$, where

$$\begin{pmatrix} \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_k \end{pmatrix} = \begin{pmatrix} \mathbf{X}_1 & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{X}_k \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}_1 \\ \vdots \\ \boldsymbol{\beta}_k \end{pmatrix} + \begin{pmatrix} \mathbf{Z}_1 & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{Z}_k \end{pmatrix} \begin{pmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_k \end{pmatrix} + \begin{pmatrix} \mathbf{e}_1 \\ \vdots \\ \mathbf{e}_k \end{pmatrix}$$

Again, the BLUP for the vector of all EBVs is given by

$$\hat{\mathbf{u}} = \mathbf{GZ}^T\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})$$

With \mathbf{V} the covariance structure for this model

For trait j ($1 \leq j \leq k$), the mixed model becomes

$$\mathbf{y}_j = \mathbf{X}_j\boldsymbol{\beta}_j + \mathbf{Z}_j\mathbf{a}_j + \mathbf{e}_j$$

$$\begin{pmatrix} \mathbf{a}_j \\ \mathbf{e}_j \end{pmatrix} \sim \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \end{pmatrix}, \begin{pmatrix} \sigma_{A_j}^2 \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \sigma_{e_j}^2 \mathbf{I} \end{pmatrix}$$

We can write this as $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{a} + \mathbf{e}$, where

$$\begin{pmatrix} \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_k \end{pmatrix} = \begin{pmatrix} \mathbf{X}_1 & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{X}_k \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}_1 \\ \vdots \\ \boldsymbol{\beta}_k \end{pmatrix} + \begin{pmatrix} \mathbf{Z}_1 & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{Z}_k \end{pmatrix} \begin{pmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_k \end{pmatrix} + \begin{pmatrix} \mathbf{e}_1 \\ \vdots \\ \mathbf{e}_k \end{pmatrix}$$

Again, the BLUP for the vector of all EBVs is given by

$$\hat{\mathbf{u}} = \mathbf{GZ}^T\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})$$

With \mathbf{V} the covariance structure for this model

An extremely convenient notation for \mathbf{R} utilizes the Kronecker product. For a $k \times l$ matrix \mathbf{A} and a $m \times n$ matrix \mathbf{B} ,

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1l} \\ a_{21} & a_{22} & \cdots & a_{2l} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kl} \end{pmatrix} \quad \text{and} \quad \mathbf{B} = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{pmatrix}$$

the Kronecker product of \mathbf{A} and \mathbf{B} , denoted $\mathbf{A} \otimes \mathbf{B}$, is the $(km) \times (ln)$ matrix

$$\mathbf{A} \otimes \mathbf{B} = \begin{pmatrix} \mathbf{B} a_{11} & \mathbf{B} a_{12} & \cdots & \mathbf{B} a_{1l} \\ \mathbf{B} a_{21} & \mathbf{B} a_{22} & \cdots & \mathbf{B} a_{2l} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{B} a_{k1} & \mathbf{B} a_{k2} & \cdots & \mathbf{B} a_{kl} \end{pmatrix}$$

where each element is itself a matrix (of order $m \times n$) with

$$\mathbf{B} a_{ij} = \begin{pmatrix} a_{ij} b_{11} & a_{ij} b_{12} & \cdots & a_{ij} b_{1n} \\ a_{ij} b_{21} & a_{ij} b_{22} & \cdots & a_{ij} b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{ij} b_{m1} & a_{ij} b_{m2} & \cdots & a_{ij} b_{mn} \end{pmatrix}$$

Covariance structure for EBVS

The resulting covariance structure for the stacked vector of breeding values is

$$\sigma \begin{pmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_k \end{pmatrix} = \begin{pmatrix} \sigma^2(A_1)\mathbf{A} & \cdots & \sigma(A_1, A_k)\mathbf{A} \\ \vdots & \ddots & \vdots \\ \sigma(A_k, A_1)\mathbf{A} & \cdots & \sigma^2(A_k)\mathbf{A} \end{pmatrix} = \mathbf{G} \otimes \mathbf{A}$$

where \otimes denotes the Kronecker (or direct) product (LW Chapter 26) and

$$\mathbf{G} = \begin{pmatrix} \sigma^2(A_1) & \cdots & \sigma(A_1, A_k) \\ \vdots & \ddots & \vdots \\ \sigma(A_k, A_1) & \cdots & \sigma^2(A_k) \end{pmatrix}$$

is the matrix of genetic covariances of interest.

The genetic variance-covariance matrix \mathbf{G} accounts for the genetic covariances among traits. \mathbf{G} has k variances and $k(k-1)/2$ covariances, which must be estimated (REML) from the data.

Covariance structure for residuals

Similarly, the covariance structure for the stacked vectors of residuals is

$$\sigma \begin{pmatrix} \mathbf{e}_1 \\ \vdots \\ \mathbf{e}_k \end{pmatrix} = \mathbf{E} \otimes \mathbf{I}, \quad \text{where } \mathbf{E} = \begin{pmatrix} \sigma^2(e_1) & \cdots & \sigma(e_1, e_k) \\ \vdots & \ddots & \vdots \\ \sigma(e_k, e_1) & \cdots & \sigma^2(e_k) \end{pmatrix}$$

Finally, we need to specify any covariances between \mathbf{a} and \mathbf{e} . By construction $\sigma(a_z, e_z) = \sigma(a_w, e_w) = 0$, while the standard assumption is $\sigma(A_z, e_w) = \sigma(A_w, e_z) = 0$, giving the covariance structure as

$$\sigma \begin{pmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_k \\ \mathbf{e}_1 \\ \vdots \\ \mathbf{e}_k \end{pmatrix} = \begin{pmatrix} \mathbf{G} \otimes \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{E} \otimes \mathbf{I} \end{pmatrix}$$

Here the matrix \mathbf{E} accounts for within-individual correlations in the environmental (or residual) values.

Random regressions

- Random regression models are basically a hybrid between repeated records models and multiple-trait models.
 - The basic structure of the model is that the trait at time t is the sum of potentially time-dependent fixed effects $\mu(t)$, a time-dependent breeding value $a(t)$, a time-dependent permanent environmental effect $p(t)$, and a residual error e . These last three are random effects
 - $y(t) = \mu(t) + a(t) + p(t) + e$
 - $a(t)$ and $p(t)$ are both approximated by random regressions, of order n and m , respectively (usually $n = m$)
 - $a_i(t_j) = \sum_k^n a_{ik} \phi_k(t_j)$ and $p_i(t_j) = \sum_k^m b_{ik} \phi_k(t_j)$
 - The vectors \mathbf{a}_i and \mathbf{b}_i for individual i are handled in a multiple-trait framework, with covariance matrices \mathbf{C}_G and \mathbf{C}_E for the within-individual vectors of additive and permanent environmental effects.

To build up the random regression model, consider the q_i observations from different times for individual i

$$\mathbf{y}_i = \begin{pmatrix} y(t_{i1}) \\ \vdots \\ y(t_{iq_i}) \end{pmatrix} = \mathbf{X}_i \boldsymbol{\beta}_i + \mathbf{Z}_{i1} \mathbf{a}_i + \mathbf{Z}_{i2} \mathbf{p}_i + \mathbf{e}_i$$

$$\mathbf{a}_i = \begin{pmatrix} a_{i0} \\ \vdots \\ a_{im} \end{pmatrix}, \quad \mathbf{p}_i = \begin{pmatrix} p_{i0} \\ \vdots \\ p_{im} \end{pmatrix}, \quad \mathbf{e}_i = \begin{pmatrix} e_{i0} \\ \vdots \\ e_{im} \end{pmatrix}$$

Here are fitting m -degree polynomials ($m < q_i$) for both the breeding value and permanent environmental value regressions. We also assume that any fixed-effects are not time dependent. Both of these assumptions are easily relaxed.

Model & covariance structure for vector \mathbf{y}_i of observations from individual i

$$\mathbf{y}_i = \begin{pmatrix} y(t_{i1}) \\ \vdots \\ y(t_{iq_i}) \end{pmatrix} = \mathbf{X}_i \boldsymbol{\beta}_i + \mathbf{Z}_{i1} \mathbf{a}_i + \mathbf{Z}_{i2} \mathbf{p}_i + \mathbf{e}_i$$

$$\mathbf{a}_i = \begin{pmatrix} a_{i0} \\ \vdots \\ a_{im} \end{pmatrix}, \quad \mathbf{p}_i = \begin{pmatrix} p_{i0} \\ \vdots \\ p_{im} \end{pmatrix}, \quad \mathbf{e}_i = \begin{pmatrix} e_{i0} \\ \vdots \\ e_{im} \end{pmatrix}$$

Covariance structure

$$\begin{pmatrix} \mathbf{a}_i \\ \mathbf{p}_i \\ \mathbf{e}_i \end{pmatrix} \sim \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \end{pmatrix}, \quad \begin{pmatrix} \mathbf{C}_G & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_E & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \sigma_e^2 \mathbf{I} \end{pmatrix}$$

The design matrix for the regression coefficients on the breeding values is very information

$$\mathbf{y}_i = \mathbf{X}_i\boldsymbol{\beta}_i + \mathbf{Z}_{i1}\mathbf{a}_i + \mathbf{Z}_{i2}\mathbf{p}_i + \mathbf{e}_i$$

$$\mathbf{Z}_{i1} = \begin{pmatrix} \phi_0(t_{i1}) & \cdots & \phi_m(t_{i1}) \\ \phi_0(t_{i2}) & \cdots & \phi_m(t_{i2}) \\ \vdots & \ddots & \vdots \\ \phi_0(t_{iq_i}) & \cdots & \phi_m(t_{iq_i}) \end{pmatrix}$$

\mathbf{Z}_{i1} is a $q_i \times (m+1)$ matrix of fixed constants that depend on the values of order zero through m Legendre polynomials, where the j th row represents these evaluated at time t_{ij} .

A KEY FEATURE is that this set of times could be different for each individual, yet the mixed model does all the bookkeeping to fully account for this.

As with the multiple trait model, stacking the individual vectors allows us to put this model in standard form. Note that while the vectors stacked for the multiple trait model represented the vectors for each trait separately, here the stacked vectors are the observations for each individual.

$$\mathbf{y} = \begin{pmatrix} \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_n \end{pmatrix}, \quad \mathbf{a} = \begin{pmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_n \end{pmatrix}, \quad \mathbf{p} = \begin{pmatrix} \mathbf{p}_1 \\ \vdots \\ \mathbf{p}_n \end{pmatrix}, \quad \mathbf{e} = \begin{pmatrix} \mathbf{e}_1 \\ \vdots \\ \mathbf{e}_n \end{pmatrix}$$

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}_1\mathbf{a} + \mathbf{Z}_2\mathbf{p} + \mathbf{e}$$

$\mathbf{Z}_1, \mathbf{Z}_2$ Block diagonal

$$\mathbf{Z}_1 = \begin{pmatrix} \mathbf{Z}_{11} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{Z}_{12} & \cdots & \mathbf{0} \\ \vdots & & \ddots & \mathbf{0} \\ \mathbf{0} & \cdots & \cdots & \mathbf{Z}_{1n} \end{pmatrix}$$

Full Model & covariance structure

$$\mathbf{y} = \begin{pmatrix} \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_n \end{pmatrix}, \quad \mathbf{a} = \begin{pmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_n \end{pmatrix}, \quad \mathbf{p} = \begin{pmatrix} \mathbf{p}_1 \\ \vdots \\ \mathbf{p}_n \end{pmatrix}, \quad \mathbf{e} = \begin{pmatrix} \mathbf{e}_1 \\ \vdots \\ \mathbf{e}_n \end{pmatrix}$$

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}_1\mathbf{a} + \mathbf{Z}_2\mathbf{p} + \mathbf{e}$$

Covariance structure

$$\begin{pmatrix} \mathbf{a} \\ \mathbf{p} \\ \mathbf{e} \end{pmatrix} \sim \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \end{pmatrix}, \quad \begin{pmatrix} \mathbf{A} \otimes \mathbf{C}_G & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \otimes \mathbf{C}_E & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \sigma_e^2 \mathbf{I} \end{pmatrix}$$

More generally, we can replace $\sigma_e^2 \mathbf{I}$ by \mathbf{R} .

Mixed-model equations (slightly more generalized covariance structure)

$$\mathbf{H} \begin{pmatrix} \hat{\mathbf{b}} \\ \hat{\mathbf{a}} \\ \hat{\mathbf{p}} \end{pmatrix} = \begin{pmatrix} \mathbf{X}^T \mathbf{R}^{-1} \mathbf{y} \\ \mathbf{Z}_1^T \mathbf{R}^{-1} \mathbf{y} \\ \mathbf{Z}_2^T \mathbf{R}^{-1} \mathbf{y} \end{pmatrix} \quad \begin{pmatrix} \mathbf{a} \\ \mathbf{p} \\ \mathbf{e} \end{pmatrix} \sim \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \end{pmatrix}, \quad \begin{pmatrix} \mathbf{A} \otimes \mathbf{C}_G & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \otimes \mathbf{C}_E & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{R} \end{pmatrix}$$

where

$$\mathbf{H} = \begin{pmatrix} \mathbf{X}^T \mathbf{R}^{-1} \mathbf{X} & \mathbf{X}^T \mathbf{R}^{-1} \mathbf{Z}_1 & \mathbf{X}^T \mathbf{R}^{-1} \mathbf{Z}_2 \\ \mathbf{Z}_1^T \mathbf{R}^{-1} \mathbf{X} & \mathbf{Z}_1^T \mathbf{R}^{-1} \mathbf{Z}_1 + \mathbf{A}^{-1} \otimes \mathbf{C}_G^{-1} & \mathbf{Z}_1^T \mathbf{R}^{-1} \mathbf{Z}_2 \\ \mathbf{Z}_2^T \mathbf{R}^{-1} \mathbf{X} & \mathbf{Z}_2^T \mathbf{R}^{-1} \mathbf{Z}_1 & \mathbf{Z}_2^T \mathbf{R}^{-1} \mathbf{Z}_2 + \mathbf{I} \otimes \mathbf{C}_E^{-1} \end{pmatrix}$$

Model-fitting issues

- A central issue is what degree m of polynomials to use.
- Standard likelihood tests can be used (compare $m = k$ with $m = k + 1$).
- Meyer suggests that tests should be comparing k with $k + 2$, as often going from odd to even does not improve fit, but going from even to even ($k+2$) does, and vice-versa.

Response to selection

- Standard BLUP selection can be used, based on some criteria for an optimal functional value (curve) in the offspring.
- The expected response in the offspring is simply obtained by substituting the average of the parental breeding values into the polynomial regression for the breeding value to generate an expected offspring curve.