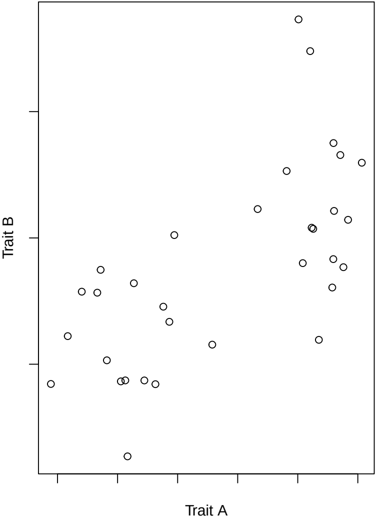
The Phylogenetic Regression

Paul Bastide[1]

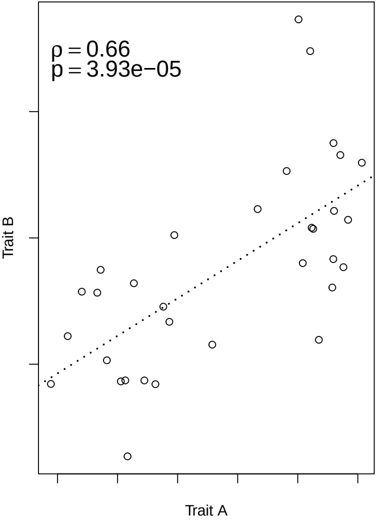
[1] IMAG, Université de Montpellier, CNRS

December 2021

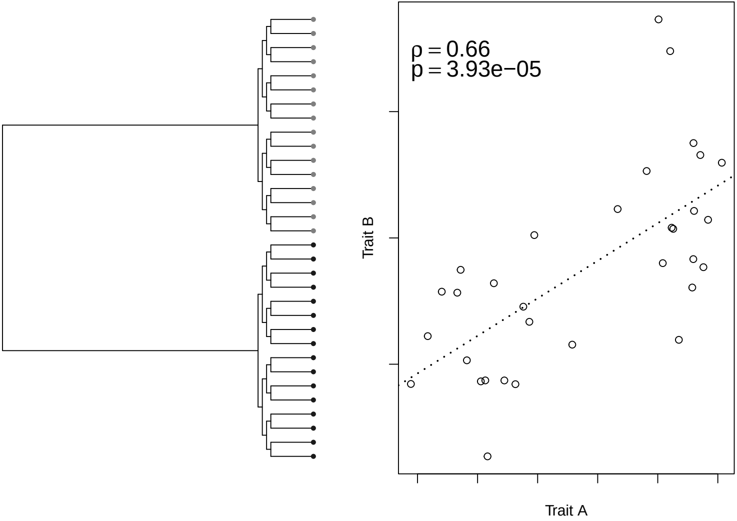
Phylogenetic Comparative Methods



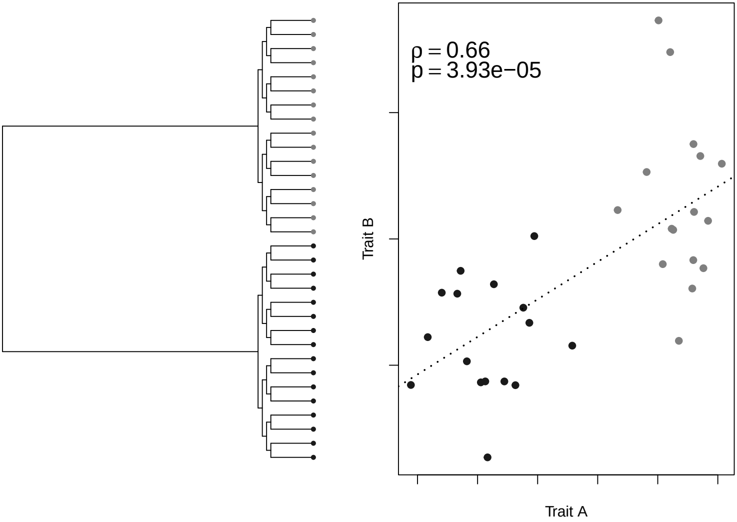
Phylogenetic Comparative Methods



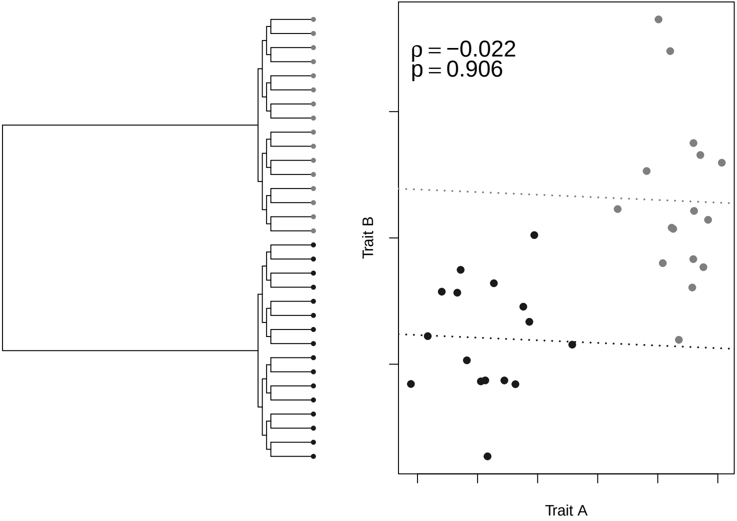
Phylogenetic Comparative Methods



Phylogenetic Comparative Methods



Phylogenetic Comparative Methods



Outline

1. Linear Regression Framework
2. Efficient Computations
3. Phylogenetic Mixed Model

Outline

Simple Gaussian Regression

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i, \quad \forall 1 \leq i \leq n$$

- y_i: quantitative response for i
- x_i: quantitative predicting variable for i
- ϵ_i: "error" for i Gaussian random variable

Simple Gaussian Regression

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i, \quad \forall 1 \leq i \leq n$$

- y_i: quantitative response for i
- x_i: quantitative predicting variable for i
- ϵ_i: "error" for i Gaussian random variable
  - Centered: $\mathbb{E}[\epsilon_i] = 0$ for all i
  - Homoscedastic: $\mathbb{V}\text{ar}[\epsilon_i] = \sigma^2$ for all i
  - Independent: $\mathbb{C}\text{ov}[\epsilon_i; \epsilon_j] = 0$ for all $i \neq j$

Simple Gaussian Regression

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i, \quad \forall 1 \leq i \leq n$$

- y_i: quantitative response for i
- x_i: quantitative predicting variable for i
- ϵ_i: "error" for i Gaussian random variable
    - Centered: $\mathbb{E}[\epsilon_i] = 0$ for all i
    - Homoscedastic: $\mathbb{V}\text{ar}[\epsilon_i] = \sigma^2$ for all i
    - Independent: $\mathbb{C}\text{ov}[\epsilon_i; \epsilon_j] = 0$ for all $i \neq j$
    - iid: $\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}_n, \sigma^2 \mathbf{I}_n)$

Simple Gaussian Regression

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i, \quad \forall 1 \leq i \leq n$$

Simple Gaussian Regression

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i, \quad \forall 1 \leq i \leq n$$

$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \beta_0 \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} + \beta_1 \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} \epsilon_1 \\ \vdots \\ \epsilon_n \end{pmatrix}$$

Simple Gaussian Regression

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i, \quad \forall 1 \leq i \leq n$$

$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \beta_0 \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} + \beta_1 \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} \epsilon_1 \\ \vdots \\ \epsilon_n \end{pmatrix}$$

$$\mathbf{Y} = \beta_0 \mathbf{1}_n + \beta_1 \mathbf{x} + \boldsymbol{\epsilon}$$

Simple Gaussian Regression

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i, \quad \forall 1 \leq i \leq n$$

$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \beta_0 \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} + \beta_1 \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} \epsilon_1 \\ \vdots \\ \epsilon_n \end{pmatrix}$$

$$\mathbf{Y} = \beta_0 \mathbf{1}_n + \beta_1 \mathbf{x} + \boldsymbol{\epsilon}$$

- $\mathbf{Y} = (y_1, \dots, y_n)^T$ **random** vector of responses
- $\mathbf{1}_n = (1, \dots, 1)^T$ vector of ones
- $\mathbf{x} = (x_1, \dots, x_n)^T$ **non random** vector of predictors
- $\boldsymbol{\epsilon} = (\epsilon_1, \dots, \epsilon_n)^T$ **random iid Gaussian** vector of errors
- β_0, β_1 **non random, unknown** coefficients

Simple Gaussian Regression

$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \beta_0 \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} + \beta_1 \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} \epsilon_1 \\ \vdots \\ \epsilon_n \end{pmatrix}$$

Simple Gaussian Regression

$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \beta_0 \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} + \beta_1 \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} \epsilon_1 \\ \vdots \\ \epsilon_n \end{pmatrix}$$

$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix} + \begin{pmatrix} \epsilon_1 \\ \vdots \\ \epsilon_n \end{pmatrix}$$

Simple Gaussian Regression

$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \beta_0 \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} + \beta_1 \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} \epsilon_1 \\ \vdots \\ \epsilon_n \end{pmatrix}$$

$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix} + \begin{pmatrix} \epsilon_1 \\ \vdots \\ \epsilon_n \end{pmatrix}$$

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$$

Simple Gaussian Regression

$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \beta_0 \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} + \beta_1 \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} \epsilon_1 \\ \vdots \\ \epsilon_n \end{pmatrix}$$

$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix} + \begin{pmatrix} \epsilon_1 \\ \vdots \\ \epsilon_n \end{pmatrix}$$

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$$

- $\mathbf{X}$ ($n \times 2$) matrix of regressors
- $\boldsymbol{\beta}$ (length 2) vector of coefficients

Multiple Gaussian Regression

$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \beta_0 \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} + \beta_1 \begin{pmatrix} x_{11} \\ \vdots \\ x_{n1} \end{pmatrix} + \cdots + \beta_p \begin{pmatrix} x_{1p} \\ \vdots \\ x_{np} \end{pmatrix} + \begin{pmatrix} \epsilon_1 \\ \vdots \\ \epsilon_n \end{pmatrix}$$

Multiple Gaussian Regression

$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \beta_0 \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} + \beta_1 \begin{pmatrix} x_{11} \\ \vdots \\ x_{n1} \end{pmatrix} + \cdots + \beta_p \begin{pmatrix} x_{1p} \\ \vdots \\ x_{np} \end{pmatrix} + \begin{pmatrix} \epsilon_1 \\ \vdots \\ \epsilon_n \end{pmatrix}$$

$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} 1 & x_{11} & \dots & x_{1p} \\ \vdots & & & \\ 1 & x_{n1} & \dots & x_{np} \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{pmatrix} + \begin{pmatrix} \epsilon_1 \\ \vdots \\ \epsilon_n \end{pmatrix}$$

Multiple Gaussian Regression

$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \beta_0 \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} + \beta_1 \begin{pmatrix} x_{11} \\ \vdots \\ x_{n1} \end{pmatrix} + \cdots + \beta_p \begin{pmatrix} x_{1p} \\ \vdots \\ x_{np} \end{pmatrix} + \begin{pmatrix} \epsilon_1 \\ \vdots \\ \epsilon_n \end{pmatrix}$$

$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} 1 & x_{11} & \dots & x_{1p} \\ \vdots & & & \\ 1 & x_{n1} & \dots & x_{np} \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{pmatrix} + \begin{pmatrix} \epsilon_1 \\ \vdots \\ \epsilon_n \end{pmatrix}$$

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$$

Multiple Gaussian Regression

$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \beta_0 \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} + \beta_1 \begin{pmatrix} x_{11} \\ \vdots \\ x_{n1} \end{pmatrix} + \cdots + \beta_p \begin{pmatrix} x_{1p} \\ \vdots \\ x_{np} \end{pmatrix} + \begin{pmatrix} \epsilon_1 \\ \vdots \\ \epsilon_n \end{pmatrix}$$

$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} 1 & x_{11} & \dots & x_{1p} \\ \vdots & & & \\ 1 & x_{n1} & \dots & x_{np} \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{pmatrix} + \begin{pmatrix} \epsilon_1 \\ \vdots \\ \epsilon_n \end{pmatrix}$$

$$\mathbf{Y} = \mathbf{X}\beta + \epsilon$$

- $\mathbf{X}$ ($n \times p$) matrix of regressors ($rg(\mathbf{X}) = p$)
- β (length p) vector of coefficients

Linear Regression

$$\mathbf{Y} = \mathbf{X}\beta + \sigma\boldsymbol{\epsilon}$$

Linear Regression

$$\mathbf{Y} = \mathbf{X}\beta + \sigma\boldsymbol{\epsilon}$$

- $\mathbf{Y}$ observed vector of traits at the tips of the tree (length n)

Linear Regression

$$\mathbf{Y} = \mathbf{X}\beta + \sigma\boldsymbol{\epsilon}$$

- $\mathbf{Y}$ observed vector of traits at the tips of the tree (length n)
- $\mathbf{X}$ matrix of regressors (size $n \times p$)
- β vector of coefficients (length p)

Linear Regression

$$\mathbf{Y} = \mathbf{X}\beta + \sigma\boldsymbol{\epsilon}$$

- $\mathbf{Y}$ observed vector of traits at the tips of the tree (length n)
- $\mathbf{X}$ matrix of regressors (size $n \times p$)
- β vector of coefficients (length p)
- $\boldsymbol{\epsilon}$ vector of errors (length n)

Linear Regression

$$\mathbf{Y} = \mathbf{X}\beta + \sigma\boldsymbol{\epsilon}$$

- $\mathbf{Y}$ observed vector of traits at the tips of the tree (length n)
- $\mathbf{X}$ matrix of regressors (size $n \times p$)
- β vector of coefficients (length p)
- $\boldsymbol{\epsilon}$ vector of errors (length n)

$$\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}_n, \mathbf{I}_n)$$

Linear Regression: Estimation

$$\mathbf{Y} = \mathbf{X}\beta + \sigma\boldsymbol{\epsilon}$$

Linear Regression: Estimation

$$\mathbf{Y} = \mathbf{X}\beta + \sigma\boldsymbol{\epsilon}$$

$$\hat{\beta} = \underset{\beta \in \mathbb{R}^p}{\operatorname{argmin}} \sum_{i=1}^{n} \left(y_i - \sum_{j=1}^{p} x_{ij}\beta_j \right)^2 = \underset{\beta \in \mathbb{R}^p}{\operatorname{argmin}} \|\mathbf{Y} - \mathbf{X}\beta\|^2$$

Linear Regression: Estimation

$$\mathbf{Y} = \mathbf{X}\beta + \sigma\boldsymbol{\epsilon}$$

$$\hat{\beta} = \underset{\beta \in \mathbb{R}^p}{\operatorname{argmin}} \sum_{i=1}^{n} \left(y_i - \sum_{j=1}^{p} x_{ij}\beta_j \right)^2 = \underset{\beta \in \mathbb{R}^p}{\operatorname{argmin}} \|\mathbf{Y} - \mathbf{X}\beta\|^2$$

$$\hat{\beta} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{Y}$$

Linear Regression: Estimation

$$\mathbf{Y} = \mathbf{X}\beta + \sigma\boldsymbol{\epsilon}$$

$$\hat{\beta} = \underset{\beta\in\mathbb{R}^p}{\operatorname{argmin}} \sum_{i=1}^{n} \left(y_i - \sum_{j=1}^{p} x_{ij}\beta_j \right)^2 = \underset{\beta\in\mathbb{R}^p}{\operatorname{argmin}} \|\mathbf{Y} - \mathbf{X}\beta\|^2$$

$$\hat{\beta} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{Y}$$

$$\hat{\mathbf{Y}} = \mathbf{X}\hat{\beta} = \underset{\mathbf{u}\in\mathcal{M}_{\mathbf{X}}(\mathbb{R}^p)}{\operatorname{argmin}} \|\mathbf{Y} - \mathbf{u}\|^2 = \mathbf{P}_{\mathbf{X}}\mathbf{Y}$$

Linear Regression: Estimation

$$\mathbf{Y} = \mathbf{X}\beta + \sigma\boldsymbol{\epsilon}$$

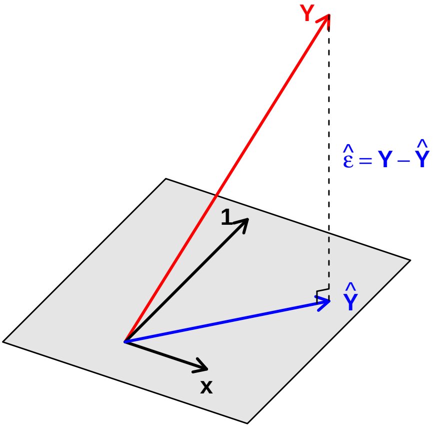
$$\hat{\beta} = \underset{\beta \in \mathbb{R}^p}{\operatorname{argmin}} \sum_{i=1}^{n} \left(y_i - \sum_{j=1}^{p} x_{ij}\beta_j \right)^2 = \underset{\beta \in \mathbb{R}^p}{\operatorname{argmin}} \|\mathbf{Y} - \mathbf{X}\beta\|^2$$

$$\hat{\beta} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{Y}$$

$$\hat{\mathbf{Y}} = \mathbf{X}\hat{\beta} = \underset{\mathbf{u} \in \mathcal{M}_{\mathbf{X}}(\mathbb{R}^p)}{\operatorname{argmin}} \|\mathbf{Y} - \mathbf{u}\|^2 = \mathbf{P}_{\mathbf{X}}\mathbf{Y}$$

$$\hat{\sigma}^2 = \frac{1}{n-p} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 = \frac{1}{n-p} \|\mathbf{Y} - \hat{\mathbf{Y}}\|^2$$

Linear Regression: Estimation



Linear Regression: Estimation

$$\mathbf{Y} = \mathbf{X}\beta + \sigma\boldsymbol{\epsilon}$$

Linear Regression: Estimation

$$\mathbf{Y} = \mathbf{X}\beta + \sigma\boldsymbol{\epsilon}$$

Estimators:

$$\hat{\beta} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{Y}$$
$$\hat{\sigma}^2 = \frac{1}{n-p}\|\mathbf{Y} - \hat{\mathbf{Y}}\|^2$$

Linear Regression: Estimation

$$\mathbf{Y} = \mathbf{X}\beta + \sigma\epsilon$$

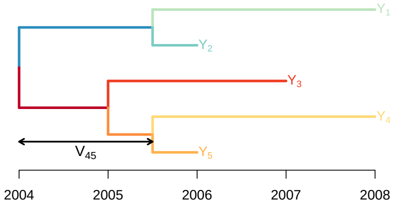
Estimators:

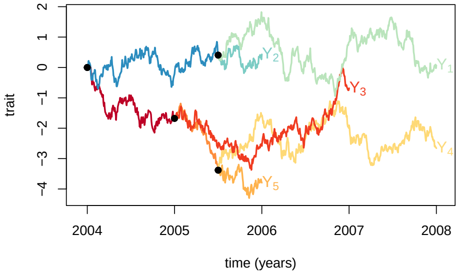
$$\hat{\beta} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{Y}$$
$$\hat{\sigma}^2 = \frac{1}{n-p}\|\mathbf{Y} - \hat{\mathbf{Y}}\|^2$$

Distribution:

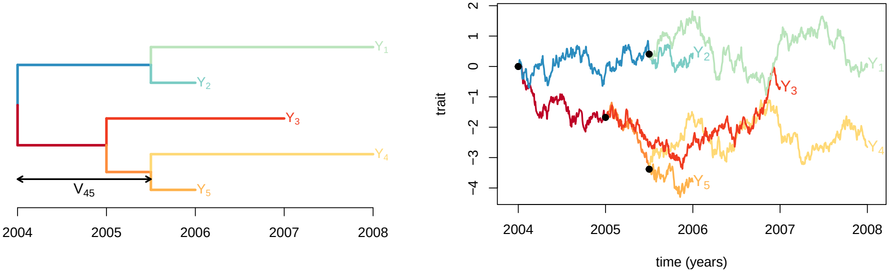
$$\frac{\hat{\beta}_k - \beta_k}{\sqrt{\hat{\sigma}^2[(\mathbf{X}^T\mathbf{X})^{-1}]_{kk}}} \sim \mathcal{T}_{n-p}$$
$$\frac{(n-p)\hat{\sigma}^2}{\sigma^2} \sim \chi^2(n-p)$$

Brownian Motion on a Tree



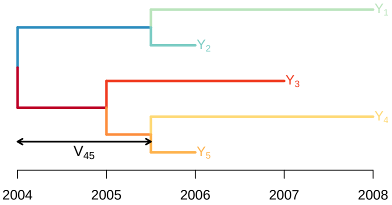


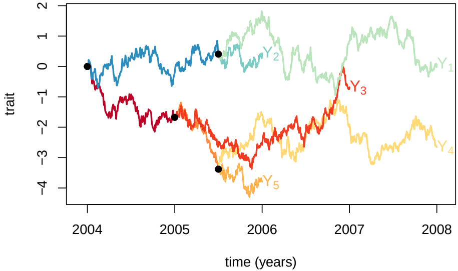
Brownian Motion on a Tree



- Variance: $\mathbb{C}\text{ov}(Y_i, Y_j) = \sigma^2 V_{ij}$
- Expectation: μ

Brownian Motion on a Tree





- Variance: $\mathbb{C}\mathrm{ov}(Y_i, Y_j) = \sigma^2 V_{ij}$
- Expectation: μ
- Distribution: $\mathbf{Y} = \mu\mathbf{1} + \sigma\mathbf{E}$, $\mathbf{E} \sim \mathcal{N}(\mathbf{0}_n, \mathbf{V})$.

Phylogenetic Linear Regression

$$\mathbf{Y} = \mathbf{X}\beta + \sigma\mathbf{E}$$

Phylogenetic Linear Regression

$$\mathbf{Y} = \mathbf{X}\beta + \sigma\mathbf{E}$$

- **Y** observed vector of traits at the tips of the tree (length *n*)

Phylogenetic Linear Regression

$$\mathbf{Y} = \mathbf{X}\beta + \sigma\mathbf{E}$$

- **Y** observed vector of traits at the tips of the tree (length n)
- **X** matrix of regressors (size $n \times p$)
- β vector of coefficients (length p)

Phylogenetic Linear Regression

$$\mathbf{Y} = \mathbf{X}\beta + \sigma\mathbf{E}$$

- **Y** observed vector of traits at the tips of the tree (length n)
- **X** matrix of regressors (size $n \times p$)
- β vector of coefficients (length p)
- **E** vector of phylogenetic errors (length n)

Phylogenetic Linear Regression

$$\mathbf{Y} = \mathbf{X}\beta + \sigma\mathbf{E}$$

- **Y** observed vector of traits at the tips of the tree (length n)
- **X** matrix of regressors (size $n \times p$)
- β vector of coefficients (length p)
- **E** vector of phylogenetic errors (length n)

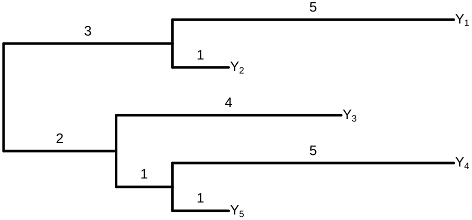
$$\mathbf{E} \sim \mathcal{N}(\mathbf{0}_n, \mathbf{V})$$

V informed by the tree and the trait model

Phylogenetic Linear Regression

$$\mathbf{Y} = \mathbf{X}\beta + \sigma\mathbf{E} \quad \mathbf{E} \sim \mathcal{N}(\mathbf{0}_n, \mathbf{V})$$

BM on a tree:

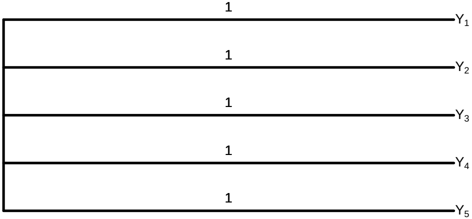


$$\mathbf{V} = \begin{array}{c} \\ Y_1 \\ Y_2 \\ Y_3 \\ Y_4 \\ Y_5 \end{array} \begin{array}{c} \begin{array}{ccccc} Y_1 & Y_2 & Y_3 & Y_4 & Y_5 \end{array} \\ \begin{pmatrix} 8 & 3 & 0 & 0 & 0 \\ 3 & 4 & 0 & 0 & 0 \\ 0 & 0 & 6 & 2 & 2 \\ 0 & 0 & 2 & 8 & 3 \\ 0 & 0 & 2 & 3 & 4 \end{pmatrix} \end{array}$$

Phylogenetic Linear Regression

$$\mathbf{Y} = \mathbf{X}\beta + \sigma\mathbf{E} \quad \mathbf{E} \sim \mathcal{N}(\mathbf{0}_n, \mathbf{V})$$

BM on a star tree:

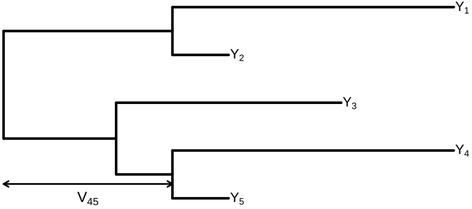


$$\mathbf{V} = \begin{array}{c} \\ Y_1 \\ Y_2 \\ Y_3 \\ Y_4 \\ Y_5 \end{array} \begin{array}{c} \begin{array}{ccccc} Y_1 & Y_2 & Y_3 & Y_4 & Y_5 \end{array} \\ \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{array}$$

Phylogenetic Linear Regression

$$\mathbf{Y} = \mathbf{X}\beta + \sigma\mathbf{E} \quad \mathbf{E} \sim \mathcal{N}(\mathbf{0}_n, \mathbf{V}(\alpha))$$

OU on a tree:

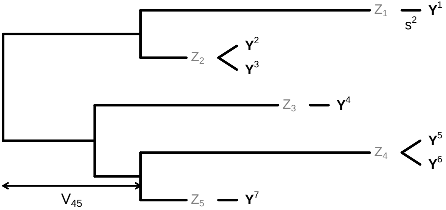


$$\mathbf{V}(\alpha)_{ij} = e^{-\alpha(V_i+V_j)}\frac{e^{2\alpha V_{ij}} - 1}{2\alpha}$$

Phylogenetic Linear Regression

$$\mathbf{Y} = \mathbf{X}\beta + \sigma \mathbf{E} \quad \mathbf{E} \sim \mathcal{N}(\mathbf{0}_n, \mathbf{V}(\phi)) \quad \phi = \frac{s^2}{\sigma^2}$$

BM with errors:



$$\mathbf{V}(\phi)_{ij} = V_{ij}$$
$$\mathbf{V}(\phi)_{ii} = V_{ii} + \phi$$

Generalized Least Squares

Assume that the tree and model ($\boldsymbol{\theta}$) is known.

$$\mathbf{Y} = \mathbf{X}\beta + \sigma\mathbf{E}, \quad \mathbf{E} \sim \mathcal{N}(\mathbf{0}_n, \mathbf{V})$$

Generalized Least Squares

Assume that the tree and model ($\boldsymbol{\theta}$) is known.

$$\mathbf{Y} = \mathbf{X}\beta + \sigma\mathbf{E}, \quad \mathbf{E} \sim \mathcal{N}(\mathbf{0}_n, \mathbf{V})$$

Cholesky Decomposition:

$$\mathbf{V} = \mathbf{L}\mathbf{L}^T$$

$$\mathbb{V}\text{ar}[\mathbf{L}^{-1}\mathbf{E}] = \mathbf{L}^{-1}\mathbf{V}[\mathbf{L}^{-1}]^T = \mathbf{I}$$

Generalized Least Squares

Assume that the tree and model ($\boldsymbol{\theta}$) is known.

$$\mathbf{Y} = \mathbf{X}\beta + \sigma\mathbf{E}, \quad \mathbf{E} \sim \mathcal{N}(\mathbf{0}_n, \mathbf{V})$$

Cholesky Decomposition:

$$\mathbf{V} = \mathbf{L}\mathbf{L}^T$$

$$\mathbb{V}\mathrm{ar}[\mathbf{L}^{-1}\mathbf{E}] = \mathbf{L}^{-1}\mathbf{V}[\mathbf{L}^{-1}]^T = \mathbf{I}$$

De-correlated Regression:

$$\mathbf{L}^{-1}\mathbf{Y} = (\mathbf{L}^{-1}\mathbf{X})\beta + \sigma\mathbf{E}', \quad \mathbf{E}' \sim \mathcal{N}(\mathbf{0}_n, \mathbf{I})$$

Generalized Least Squares

$$\mathbf{Y} = \mathbf{X}\beta + \sigma\mathbf{E} \quad \mathbf{E} \sim \mathcal{N}(\mathbf{0}_n, \mathbf{V}(\boldsymbol{\theta}))$$

Generalized Least Squares

$$\mathbf{Y} = \mathbf{X}\beta + \sigma\mathbf{E} \quad \mathbf{E} \sim \mathcal{N}(\mathbf{0}_n, \mathbf{V}(\theta))$$

Estimators:

$$\hat{\beta} = (\mathbf{X}^T\mathbf{V}(\theta)^{-1}\mathbf{X})^{-1}\mathbf{X}^T\mathbf{V}(\theta)^{-1}\mathbf{Y}$$

$$\hat{\sigma}^2 = \frac{1}{n-p}(\mathbf{Y} - \mathbf{X}\hat{\beta})^T\mathbf{V}(\theta)^{-1}(\mathbf{Y} - \mathbf{X}\hat{\beta}) = \frac{1}{n-p}\|\mathbf{Y} - \mathbf{X}\hat{\beta}\|^2_{\mathbf{V}(\theta)^{-1}}$$

Generalized Least Squares

$$\mathbf{Y} = \mathbf{X}\beta + \sigma\mathbf{E} \quad \mathbf{E} \sim \mathcal{N}(\mathbf{0}_n, \mathbf{V}(\boldsymbol{\theta}))$$

Estimators:

$$\hat{\beta} = (\mathbf{X}^T\mathbf{V}(\boldsymbol{\theta})^{-1}\mathbf{X})^{-1}\mathbf{X}^T\mathbf{V}(\boldsymbol{\theta})^{-1}\mathbf{Y}$$

$$\hat{\sigma}^2 = \frac{1}{n-p}(\mathbf{Y} - \mathbf{X}\hat{\beta})^T\mathbf{V}(\boldsymbol{\theta})^{-1}(\mathbf{Y} - \mathbf{X}\hat{\beta}) = \frac{1}{n-p}\|\mathbf{Y} - \mathbf{X}\hat{\beta}\|^2_{\mathbf{V}(\boldsymbol{\theta})^{-1}}$$

Distribution:

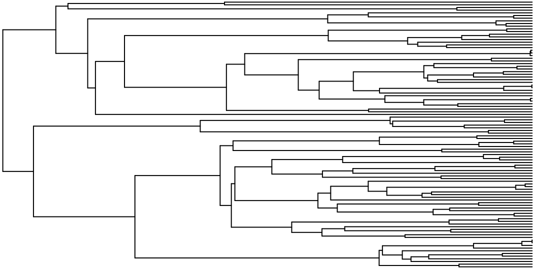
$$\frac{\hat{\beta}_k - \beta_k}{\sqrt{\hat{\sigma}^2[(\mathbf{X}^T\mathbf{V}(\boldsymbol{\theta})^{-1}\mathbf{X})^{-1}]_{kk}}} \sim \mathcal{T}_{n-p}$$

$$\frac{(n-p)\hat{\sigma}^2}{\sigma^2} \sim \chi^2(n-p)$$

Simulated Example

```r
library(ape)
library(phylolm)
```

```r
set.seed(18300718)
n <- 100
tree <- rphylo(n, birth = 0.1, death = 0)
plot(tree, show.tip.label = FALSE, no.margin = TRUE)
```



Simulated Example

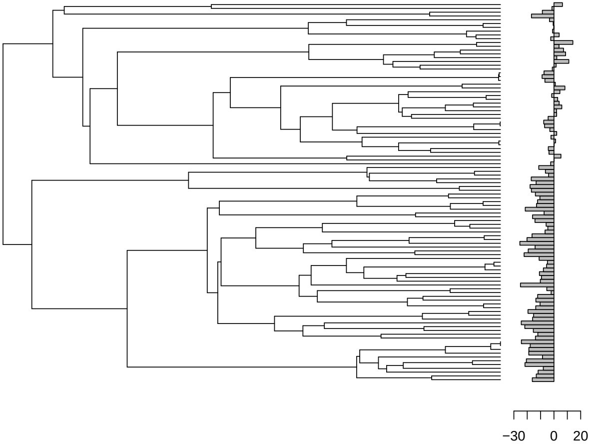
```r
# Regressor
x <- rnorm(n, mean = 0, sd = 4)

# Phylogenetic Noise
eps <- rTrait(n = 1, phy = tree, model = "BM",
              parameters = list(ancestral.state = 0, sigma2 = 2))

# Response variable
beta0 <- -10
beta1 <- 0.5
y <- beta0 + beta1 * x + eps
```

Simulated Example

```
par(mar = c(5, 0, 0, 0) + 0.1)
plot(tree, show.tip.label = FALSE, x.lim = 45)
phydataplot(y, tree, scaling = 0.1, offset = 4)
```



Simulated Example

```
# Standard linear regression
fit <- lm(y ~ x)
summary(fit)


##
## Call:
## lm(formula = y ~ x)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -20.0609  -6.4979   0.2226   7.3545  18.9895
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -7.8921     0.8888  -8.880 3.25e-14 ***
## x             0.5032     0.2168   2.321   0.0224 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 8.887 on 98 degrees of freedom
## Multiple R-squared:  0.0521,Adjusted R-squared:  0.04243
## F-statistic: 5.387 on 1 and 98 DF,  p-value: 0.02236
```

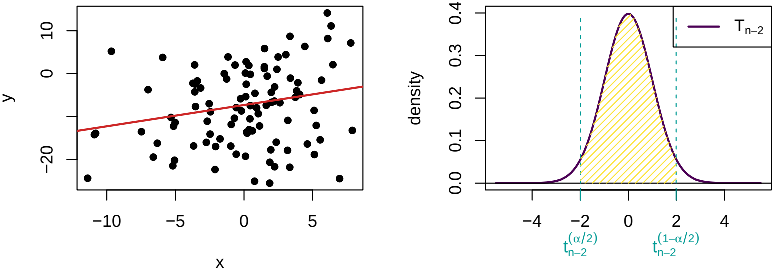
Simulated Example

```
# Phylogenetic linear regression
fitphy <- phylolm(y ~ x, phy = tree)
summary(fitphy)
```

```
##
## Call:
## phylolm(formula = y ~ x, phy = tree)
##
##    AIC logLik
##  573.2 -283.6
##
## Raw residuals:
##     Min      1Q  Median      3Q     Max
## -20.605  -7.110  -0.371   6.751  18.439
##
## Mean tip height: 37.22047
## Parameter estimate(s) using ML:
## sigma2: 1.968463
##
## Coefficients:
##              Estimate    StdErr t.value p.value
## (Intercept) -7.295577  3.075589 -2.3721 0.01964 *
## x            0.495603  0.010846 45.6953 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## R-squared: 0.9552 Adjusted R-squared: 0.9547
```

Simulated Example

$$\mathbf{y} = -10 \cdot \mathbf{1} + 0.5 \cdot \mathbf{x} + \boldsymbol{\epsilon}$$



With probability $1-\alpha$:

$$t_{n-2}(\alpha/2) \leq \frac{\hat{\beta}_1 - \beta_1}{\sqrt{\hat{\sigma}^2[(\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X})^{-1}]_{11}}} \leq t_{n-2}(1-\alpha/2)$$

Outline

Generalized Least Squares

$$\mathbf{Y} = \mathbf{X}\beta + \sigma\mathbf{E} \quad \mathbf{E} \sim \mathcal{N}(\mathbf{0}_n, \mathbf{V}(\boldsymbol{\theta}))$$

Estimators:

$$\begin{aligned}
\hat{\beta} &= (\mathbf{X}^T\mathbf{V}(\boldsymbol{\theta})^{-1}\mathbf{X})^{-1}\mathbf{X}^T\mathbf{V}(\boldsymbol{\theta})^{-1}\mathbf{Y} \\
\hat{\sigma}^2 &= \frac{1}{n-p}(\mathbf{Y} - \mathbf{X}\hat{\beta})^T\mathbf{V}(\boldsymbol{\theta})^{-1}(\mathbf{Y} - \mathbf{X}\hat{\beta}) \\
2L(\mathbf{y}|\boldsymbol{\theta}) &= n\log(2\pi) + \log\det\mathbf{V}(\boldsymbol{\theta}) + (\mathbf{Y} - \mathbf{X}\hat{\beta})^T\mathbf{V}(\boldsymbol{\theta})^{-1}(\mathbf{Y} - \mathbf{X}\hat{\beta})
\end{aligned}$$

Generalized Least Squares

$$\mathbf{Y} = \mathbf{X}\beta + \sigma\mathbf{E} \quad \mathbf{E} \sim \mathcal{N}(\mathbf{0}_n, \mathbf{V}(\boldsymbol{\theta}))$$

Estimators:

$$\begin{aligned}\hat{\beta} &= (\mathbf{X}^T\mathbf{V}(\boldsymbol{\theta})^{-1}\mathbf{X})^{-1}\mathbf{X}^T\mathbf{V}(\boldsymbol{\theta})^{-1}\mathbf{Y}\\ \hat{\sigma}^2 &= \frac{1}{n-p}(\mathbf{Y}-\mathbf{X}\hat{\beta})^T\mathbf{V}(\boldsymbol{\theta})^{-1}(\mathbf{Y}-\mathbf{X}\hat{\beta})\\ 2L(\mathbf{y}|\boldsymbol{\theta}) &= n\log(2\pi) + \log\det\mathbf{V}(\boldsymbol{\theta}) + (\mathbf{Y}-\mathbf{X}\hat{\beta})^T\mathbf{V}(\boldsymbol{\theta})^{-1}(\mathbf{Y}-\mathbf{X}\hat{\beta})\end{aligned}$$

Naive Computation: Need to invert $\mathbf{V}(\boldsymbol{\theta}) \rightarrow$ worse than $O(n^2)$.

Generalized Least Squares

$$\mathbf{Y} = \mathbf{X}\beta + \sigma\mathbf{E} \quad \mathbf{E} \sim \mathcal{N}(\mathbf{0}_n, \mathbf{V}(\boldsymbol{\theta}))$$

Estimators:

$$\hat{\beta} = (\mathbf{X}^T\mathbf{V}(\boldsymbol{\theta})^{-1}\mathbf{X})^{-1}\mathbf{X}^T\mathbf{V}(\boldsymbol{\theta})^{-1}\mathbf{Y}$$

$$\hat{\sigma}^2 = \frac{1}{n-p}(\mathbf{Y} - \mathbf{X}\hat{\beta})^T\mathbf{V}(\boldsymbol{\theta})^{-1}(\mathbf{Y} - \mathbf{X}\hat{\beta})$$

$$2L(\mathbf{y}|\boldsymbol{\theta}) = n\log(2\pi) + \log\det\mathbf{V}(\boldsymbol{\theta}) + (\mathbf{Y} - \mathbf{X}\hat{\beta})^T\mathbf{V}(\boldsymbol{\theta})^{-1}(\mathbf{Y} - \mathbf{X}\hat{\beta})$$

Naive Computation: Need to invert $\mathbf{V}(\boldsymbol{\theta}) \rightarrow$ worse than $O(n^2)$.

Three point structure:
Estimators computed in one traversal of the tree $\rightarrow O(n)$.

+

Three Point Structure Algorithm

(Ho and Ané, 2014)

$$\mathbf{Y} = \mathbf{X}\beta + \sigma\mathbf{E} \quad \mathbf{E} \sim \mathcal{N}(\mathbf{0}_n, \mathbf{V})$$

Estimators:

$$\hat{\beta} = (\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}^T\mathbf{V}^{-1}\mathbf{Y}$$

$$\hat{\sigma}^2 = \frac{1}{n-p}(\mathbf{Y} - \mathbf{X}\hat{\beta})^T\mathbf{V}^{-1}(\mathbf{Y} - \mathbf{X}\hat{\beta})$$

$$2L(\mathbf{y}|\boldsymbol{\theta}) = n\log(2\pi) + \log|\mathbf{V}| + (\mathbf{Y} - \mathbf{X}\hat{\beta})^T\mathbf{V}^{-1}(\mathbf{Y} - \mathbf{X}\hat{\beta})$$

Can be computed in linear time.

Three Point Structure Algorithm

(Ho and Ané, 2014)

$$\mathbf{Y} = \mathbf{X}\beta + \sigma\mathbf{E} \quad \mathbf{E} \sim \mathcal{N}(\mathbf{0}_n, \mathbf{V})$$

Estimators:

$$\begin{aligned}
\hat{\beta} &= (\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}^T\mathbf{V}^{-1}\mathbf{Y} \\
\hat{\sigma}^2 &= \frac{1}{n-p}(\mathbf{Y} - \mathbf{X}\hat{\beta})^T\mathbf{V}^{-1}(\mathbf{Y} - \mathbf{X}\hat{\beta}) \\
2L(\mathbf{y}|\boldsymbol{\theta}) &= n\log(2\pi) + \log|\mathbf{V}| + (\mathbf{Y} - \mathbf{X}\hat{\beta})^T\mathbf{V}^{-1}(\mathbf{Y} - \mathbf{X}\hat{\beta})
\end{aligned}$$

Can be computed in linear time.

- Works for a BM on a tree

Three Point Structure Algorithm

(Ho and Ané, 2014)

$$\mathbf{Y} = \mathbf{X}\beta + \sigma\mathbf{E} \quad \mathbf{E} \sim \mathcal{N}(\mathbf{0}_n, \mathbf{V})$$

Estimators:

$$\begin{aligned} \hat{\beta} &= (\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}^T\mathbf{V}^{-1}\mathbf{Y} \\ \hat{\sigma}^2 &= \frac{1}{n-p}(\mathbf{Y}-\mathbf{X}\hat{\beta})^T\mathbf{V}^{-1}(\mathbf{Y}-\mathbf{X}\hat{\beta}) \\ 2L(\mathbf{y}|\boldsymbol{\theta}) &= n\log(2\pi) + \log|\mathbf{V}| + (\mathbf{Y}-\mathbf{X}\hat{\beta})^T\mathbf{V}^{-1}(\mathbf{Y}-\mathbf{X}\hat{\beta}) \end{aligned}$$

Can be computed in linear time.

- Works for a BM on a tree
- Implemented in phylolm

Three Point Structure Algorithm

(Ho and Ané, 2014)

$$\mathbf{Y} = \mathbf{X}\beta + \sigma\mathbf{E} \quad \mathbf{E} \sim \mathcal{N}(\mathbf{0}_n, \mathbf{V})$$

Estimators:

$$\begin{aligned}
\hat{\beta} &= (\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}^T\mathbf{V}^{-1}\mathbf{Y} \\
\hat{\sigma}^2 &= \frac{1}{n-p}(\mathbf{Y} - \mathbf{X}\hat{\beta})^T\mathbf{V}^{-1}(\mathbf{Y} - \mathbf{X}\hat{\beta}) \\
2L(\mathbf{y}|\boldsymbol{\theta}) &= n\log(2\pi) + \log|\mathbf{V}| + (\mathbf{Y} - \mathbf{X}\hat{\beta})^T\mathbf{V}^{-1}(\mathbf{Y} - \mathbf{X}\hat{\beta})
\end{aligned}$$

Can be computed in linear time.

- Works for a BM on a tree
- Implemented in phylolm
- What about other models ? $\mathbf{V} = \mathbf{V}(\boldsymbol{\theta})$.

Tree Transformation Strategy

(Pagel, 1999)

Strategy

- See $\mathbf{V}(\theta)$ as a BM on a modified tree.

Tree Transformation Strategy (Pagel, 1999)

Strategy

- See $\mathbf{V}(\theta)$ as a BM on a modified tree.
- For fixed $\mathbf{V}(\theta)$, use the fast algorithm.

Tree Transformation Strategy (Pagel, 1999)

Strategy

- See $\mathbf{V}(\theta)$ as a BM on a modified tree.
- For fixed $\mathbf{V}(\theta)$, use the fast algorithm.
- Optimize on $\mathbf{V}(\theta)$ numerically.

Tree Transformation Strategy (Pagel, 1999)

Strategy

- See $\mathbf{V}(\theta)$ as a BM on a modified tree.
- For fixed $\mathbf{V}(\theta)$, use the fast algorithm.
- Optimize on $\mathbf{V}(\theta)$ numerically.

Limits

- In general, only works for ultrametric trees.

Tree Transformation Strategy

(Pagel, 1999)

Strategy

- See $\mathbf{V}(\boldsymbol{\theta})$ as a BM on a modified tree.
- For fixed $\mathbf{V}(\boldsymbol{\theta})$, use the fast algorithm.
- Optimize on $\mathbf{V}(\boldsymbol{\theta})$ numerically.

Limits

- In general, only works for ultrametric trees.
- Confidence intervals and tests only valid conditionally on $\boldsymbol{\theta}$.

$$\frac{\hat{\beta}_k - \beta_k}{\sqrt{\hat{\sigma}^2[(\mathbf{X}^T\mathbf{V}(\boldsymbol{\theta})^{-1}\mathbf{X})^{-1}]_{kk}}} \sim \mathcal{T}_{n-p}$$

Pagel's Lambda

Relax the BM variance structure:

$$\mathbf{V}(\lambda)_{ii} = \mathbf{V}_{ii}$$
$$\mathbf{V}(\lambda)_{ij} = \lambda \mathbf{V}_{ij}$$

Pagel's Lambda

Relax the BM variance structure:

$$\begin{aligned}\mathbf{V}(\lambda)_{ii} &= \mathbf{V}_{ii}\\ \mathbf{V}(\lambda)_{ij} &= \lambda\mathbf{V}_{ij}\end{aligned}$$

Equivalent to running a BM on a modified tree with:

$$t(\lambda)_i = \begin{cases}\lambda t_i & \text{if } i \text{ internal node}\\ \lambda t_i + (1-\lambda)V_i & \text{if } i \text{ leaf}\end{cases}$$

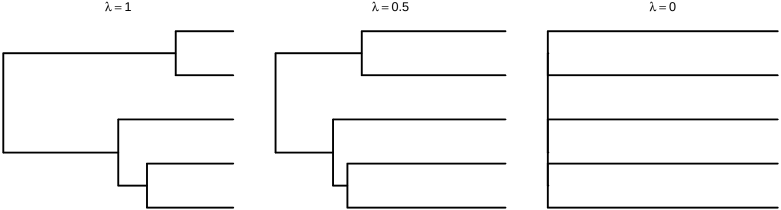
Pagel's Lambda

Relax the BM variance structure:

$$\mathbf{V}(\lambda)_{ii} = \mathbf{V}_{ii}$$
$$\mathbf{V}(\lambda)_{ij} = \lambda\mathbf{V}_{ij}$$

Equivalent to running a BM on a modified tree with:

$$t(\lambda)_i = \begin{cases} \lambda t_i & \text{if } i \text{ internal node} \\ \lambda t_i + (1-\lambda)V_i & \text{if } i \text{ leaf} \end{cases}$$



Pagel's Lambda

- Used a lot in practice.

Pagel's Lambda

- Used a lot in practice.
- Also exist Pagel's κ and δ.

Pagel's Lambda

- Used a lot in practice.
- Also exist Pagel's κ and δ.
- λ "phylogenetic heritability" (see below).

Ornstein-Uhlenbeck

Assuming a ultrametric tree with height T:

$$\mathbf{V}(\alpha)_{ij} = e^{-2\alpha T} \frac{e^{2\alpha V_{ij}} - 1}{2\alpha}$$

Ornstein-Uhlenbeck

Assuming a ultrametric tree with height T:

$$\mathbf{V}(\alpha)_{ij} = e^{-2\alpha T} \frac{e^{2\alpha V_{ij}} - 1}{2\alpha}$$

Equivalent to running a BM on a modified tree with:

$$V_i(\alpha) = e^{-2\alpha T} \frac{e^{2\alpha V_i} - 1}{2\alpha}$$

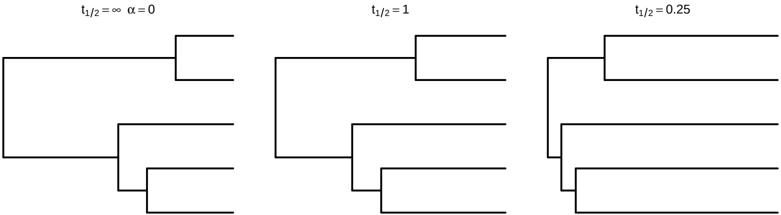
Ornstein-Uhlenbeck

Assuming a ultrametric tree with height T:

$$\mathbf{V}(\alpha)_{ij} = e^{-2\alpha T} \frac{e^{2\alpha V_{ij}} - 1}{2\alpha}$$

Equivalent to running a BM on a modified tree with:

$$V_i(\alpha) = e^{-2\alpha T} \frac{e^{2\alpha V_i} - 1}{2\alpha}$$



AC/DC

$$dX_t = \sigma_0 e^{rt/2} dB_t$$

$$\mathbf{V}(r)_{ij} = \frac{e^{rV_{ij}} - 1}{r}$$

AC/DC

$$dX_t = \sigma_0 e^{rt/2} dB_t$$

$$\mathbf{V}(r)_{ij} = \frac{e^{rV_{ij}} - 1}{r}$$

Assuming a ultrametric tree height T, equivalent to an OU with:

$$\begin{aligned} r &= 2\alpha \\ \sigma_0^2 &= \sigma^2 e^{-2\alpha T} \end{aligned}$$

AC/DC

$$dX_t = \sigma_0 e^{rt/2} dB_t$$

$$\mathbf{V}(r)_{ij} = \frac{e^{rV_{ij}} - 1}{r}$$

Assuming a ultrametric tree height T, equivalent to an OU with:

$$r = 2\alpha$$
$$\sigma_0^2 = \sigma^2 e^{-2\alpha T}$$

Case $r < 0$ (EB) is a repulsing OU.

Outline

1. Linear Regression Framework
2. Efficient Computations
3. Phylogenetic Mixed Model
    - Mixed Models and BM Heritability
    - General Heritability
    - HIV Viral Load

Phylogenetic Mixed Model

(Lynch, 1991; Housworth et al., 2004)

Phylogenetic Regression

$$\mathbf{Y} = \mathbf{X}\beta + \sigma \mathbf{E}^p \quad \mathbf{E}^p \sim \mathcal{N}(\mathbf{0}_n, \mathbf{V}(\theta))$$

Phylogenetic Mixed Model

(Lynch, 1991; Housworth et al., 2004)

Phylogenetic Regression

$$\mathbf{Y} = \mathbf{X}\beta + \sigma \mathbf{E}^p \quad \mathbf{E}^p \sim \mathcal{N}(\mathbf{0}_n, \mathbf{V}(\theta))$$

Phylogenetic Mixed Model

$$\mathbf{Y} = \mathbf{X}\beta + \sigma \mathbf{E}^p + s\mathbf{E}^e \quad \mathbf{E}^e \sim \mathcal{N}(\mathbf{0}_n, \mathbf{I})$$

Phylogenetic Mixed Model

(Lynch, 1991; Housworth et al., 2004)

Phylogenetic Regression

$$\mathbf{Y} = \mathbf{X}\beta + \sigma\mathbf{E}^p \quad \mathbf{E}^p \sim \mathcal{N}(\mathbf{0}_n, \mathbf{V}(\theta))$$

Phylogenetic Mixed Model

$$\mathbf{Y} = \mathbf{X}\beta + \sigma\mathbf{E}^p + s\mathbf{E}^e \quad \mathbf{E}^e \sim \mathcal{N}(\mathbf{0}_n, \mathbf{I})$$

Distinction between:

- The heritable part of the trait

Phylogenetic Mixed Model

(Lynch, 1991; Housworth et al., 2004)

Phylogenetic Regression

$$\mathbf{Y} = \mathbf{X}\beta + \sigma\mathbf{E}^p \quad \mathbf{E}^p \sim \mathcal{N}(\mathbf{0}_n, \mathbf{V}(\boldsymbol{\theta}))$$

Phylogenetic Mixed Model

$$\mathbf{Y} = \mathbf{X}\beta + \sigma\mathbf{E}^p + s\mathbf{E}^e \quad \mathbf{E}^e \sim \mathcal{N}(\mathbf{0}_n, \mathbf{I})$$

Distinction between:

- The heritable part of the trait
  - Variance $\sigma^2\mathbf{V}(\boldsymbol{\theta})$ given by the trait evolution on the tree

Phylogenetic Mixed Model

(Lynch, 1991; Housworth et al., 2004)

Phylogenetic Regression

$$\mathbf{Y} = \mathbf{X}\beta + \sigma \mathbf{E}^p \quad \mathbf{E}^p \sim \mathcal{N}(\mathbf{0}_n, \mathbf{V}(\boldsymbol{\theta}))$$

Phylogenetic Mixed Model

$$\mathbf{Y} = \mathbf{X}\beta + \sigma \mathbf{E}^p + s\mathbf{E}^e \quad \mathbf{E}^e \sim \mathcal{N}(\mathbf{0}_n, \mathbf{I})$$

Distinction between:

- The heritable part of the trait
  - Variance $\sigma^2\mathbf{V}(\boldsymbol{\theta})$ given by the trait evolution on the tree
- The environemental part of the trait

Phylogenetic Mixed Model

(Lynch, 1991; Housworth et al., 2004)

Phylogenetic Regression

$$\mathbf{Y} = \mathbf{X}\beta + \sigma\mathbf{E}^p \quad \mathbf{E}^p \sim \mathcal{N}(\mathbf{0}_n, \mathbf{V}(\boldsymbol{\theta}))$$

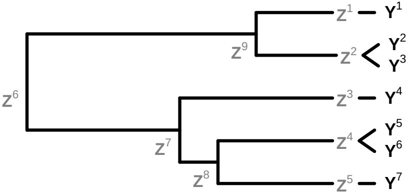
Phylogenetic Mixed Model

$$\mathbf{Y} = \mathbf{X}\beta + \sigma\mathbf{E}^p + s\mathbf{E}^e \quad \mathbf{E}^e \sim \mathcal{N}(\mathbf{0}_n, \mathbf{I})$$

Distinction between:

- The heritable part of the trait
  - Variance $\sigma^2\mathbf{V}(\boldsymbol{\theta})$ given by the trait evolution on the tree
- The environemental part of the trait
  - Variance iid s^2 given by the error model

BM on a Ultrametric Tree

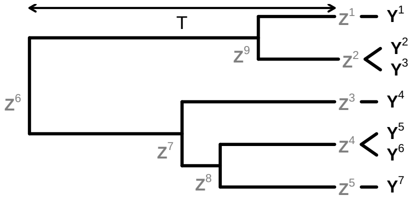


$$\mathbf{Z}^r = \boldsymbol{\mu} \qquad \text{root}$$

$$\mathbf{Z}^j \;\Big|\; \mathbf{Z}^{\mathrm{pa}(j)} \sim \mathcal{N}\left(\mathbf{Z}^{\mathrm{pa}(j)},\ \sigma^2 t_j\right) \qquad \text{nodes}$$

$$\mathbf{Y}^i \;\Big|\; \mathbf{Z}^{\mathrm{pa}(i)} \sim \mathcal{N}\left(\mathbf{Z}^{\mathrm{pa}(i)}, s^2\right) \qquad \text{obs}$$

BM on a Ultrametric Tree



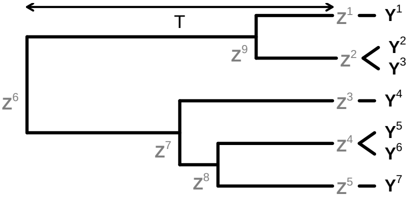
$$\mathbf{Z}^r = \boldsymbol{\mu} \qquad \text{root}$$

$$\mathbf{Z}^j \;\Big|\; \mathbf{Z}^{\text{pa}(j)} \sim \mathcal{N}\left(\mathbf{Z}^{\text{pa}(j)},\ \sigma^2 t_j\right) \qquad \text{nodes}$$

$$\mathbf{Y}^i \;\Big|\; \mathbf{Z}^{\text{pa}(i)} \sim \mathcal{N}\left(\mathbf{Z}^{\text{pa}(i)}, s^2\right) \qquad \text{obs}$$

Ultrametric tree: Tip variance is constant equal to $\sigma^2 T$

BM on a Ultrametric Tree



$$\mathbf{Z}^r = \boldsymbol{\mu} \qquad \text{root}$$

$$\mathbf{Z}^j \,\Big|\, \mathbf{Z}^{\mathrm{pa}(j)} \sim \mathcal{N}\left(\mathbf{Z}^{\mathrm{pa}(j)},\ \sigma^2 t_j\right) \qquad \text{nodes}$$

$$\mathbf{Y}^i \,\Big|\, \mathbf{Z}^{\mathrm{pa}(i)} \sim \mathcal{N}\left(\mathbf{Z}^{\mathrm{pa}(i)}, s^2\right) \qquad \text{obs}$$

Ultrametric tree: Tip variance is constant equal to $\sigma^2 T$

Phylogenetic Heritability:

$$h^2 = \frac{\sigma^2 T}{\sigma^2 T + s^2}$$

BM on a Ultrametric Tree

Assume a ultrametric tree with height T.

BM on a Ultrametric Tree

Assume a ultrametric tree with height T.

BM PMM total variance is:

$$\sigma^2\mathbf{V} + s^2\mathbf{I} \quad \text{with}\, \text{Diag}(\mathbf{V}) = T\mathbf{I}$$

BM on a Ultrametric Tree

Assume a ultrametric tree with height T.

BM PMM total variance is:

$$\sigma^2\mathbf{V} + s^2\mathbf{I} \quad \text{with}\,\text{Diag}(\mathbf{V}) = T\mathbf{I}$$

To have both variances on the same scale, we write:

$$\sigma^2\mathbf{V} + s^2\mathbf{I} = \sigma^2 T\mathbf{V}/T + s^2\mathbf{I} \quad \text{with}\,\text{Diag}(\mathbf{V}/T) = \mathbf{I}$$

BM on a Ultrametric Tree

Assume a ultrametric tree with height T.

BM PMM total variance is:

$$\sigma^2\mathbf{V} + s^2\mathbf{I} \quad \text{with}\,\text{Diag}(\mathbf{V}) = T\mathbf{I}$$

To have both variances on the same scale, we write:

$$\sigma^2\mathbf{V} + s^2\mathbf{I} = \sigma^2 T\mathbf{V}/T + s^2\mathbf{I} \quad \text{with}\,\text{Diag}(\mathbf{V}/T) = \mathbf{I}$$

The phylogenetic heritability is defined as:

$$h^2 = \frac{\sigma^2 T}{\sigma^2 T + s^2}$$

Pagel's lambda a Ultrametric Tree

Assume a ultrametric tree with height T.

Pagel's lambda a Ultrametric Tree

Assume a ultrametric tree with height T.

Pagel's lambda variance is:

$$\sigma_\lambda^2 \mathbf{V}(\lambda) = \sigma_\lambda^2 \left[\lambda \mathbf{V} + (1 - \lambda) T \mathbf{I}\right]$$

Pagel's lambda a Ultrametric Tree

Assume a ultrametric tree with height T.

Pagel's lambda variance is:

$$\sigma_\lambda^2 \mathbf{V}(\lambda) = \sigma_\lambda^2 \left[\lambda \mathbf{V} + (1-\lambda) T \mathbf{I}\right]$$

BM PMM total variance is:

$$\sigma^2 \mathbf{V} + s^2 \mathbf{I} = (\sigma^2 + s^2/T) \left[\frac{\sigma^2}{\sigma^2 + s^2/T} \mathbf{V} + \left(1 - \frac{\sigma^2}{\sigma^2 + s^2/T}\right) T \mathbf{I}\right]$$

Pagel's lambda a Ultrametric Tree

Assume a ultrametric tree with height T.

Pagel's lambda variance is:

$$\sigma_\lambda^2 \mathbf{V}(\lambda) = \sigma_\lambda^2 \left[\lambda \mathbf{V} + (1 - \lambda) T \mathbf{I}\right]$$

BM PMM total variance is:

$$\sigma^2 \mathbf{V} + s^2 \mathbf{I} = (\sigma^2 + s^2/T) \left[\frac{\sigma^2}{\sigma^2 + s^2/T} \mathbf{V} + \left(1 - \frac{\sigma^2}{\sigma^2 + s^2/T}\right) T \mathbf{I}\right]$$

The two models are equivalent with:

$$\lambda = \frac{\sigma^2 T}{\sigma^2 T + s^2} = h^2 \qquad \sigma_\lambda^2 = \sigma^2 + s^2/T$$

Pagel's lambda a Ultrametric Tree

Assume a ultrametric tree with height T.

Pagel's lambda variance is:

$$\sigma_\lambda^2 \mathbf{V}(\lambda) = \sigma_\lambda^2 \left[\lambda \mathbf{V} + (1-\lambda) T \mathbf{I}\right]$$

BM PMM total variance is:

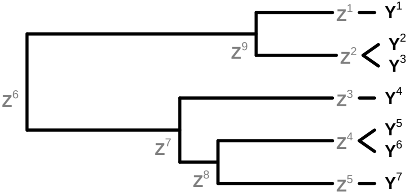
$$\sigma^2 \mathbf{V} + s^2 \mathbf{I} = (\sigma^2 + s^2/T) \left[\frac{\sigma^2}{\sigma^2 + s^2/T}\mathbf{V} + \left(1 - \frac{\sigma^2}{\sigma^2 + s^2/T}\right) T \mathbf{I}\right]$$

The two models are equivalent with:

$$\lambda = \frac{\sigma^2 T}{\sigma^2 T + s^2} = h^2 \qquad \sigma_\lambda^2 = \sigma^2 + s^2/T$$

λ is the heritability

BM on a Non-Ultrametric Tree

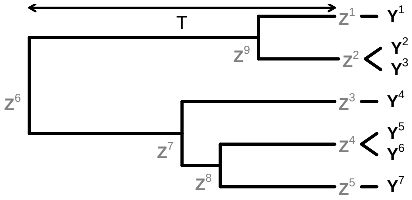


$$\mathbf{Z}^r = \boldsymbol{\mu} \qquad \text{root}$$

$$\mathbf{Z}^j \,\Big|\, \mathbf{Z}^{\text{pa}(j)} \sim \mathcal{N}\left(\mathbf{Z}^{\text{pa}(j)},\ \sigma^2 t_j\right) \qquad \text{nodes}$$

$$\mathbf{Y}^i \,\Big|\, \mathbf{Z}^{\text{pa}(i)} \sim \mathcal{N}\left(\mathbf{Z}^{\text{pa}(i)}, s^2\right) \qquad \text{obs}$$

BM on a Non-Ultrametric Tree



$$\mathbf{Z}^r = \boldsymbol{\mu} \qquad \text{root}$$

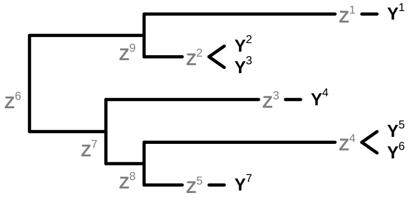
$$\mathbf{Z}^j \mid \mathbf{Z}^{\text{pa}(j)} \sim \mathcal{N}\left(\mathbf{Z}^{\text{pa}(j)},\ \sigma^2 t_j\right) \qquad \text{nodes}$$

$$\mathbf{Y}^i \mid \mathbf{Z}^{\text{pa}(i)} \sim \mathcal{N}\left(\mathbf{Z}^{\text{pa}(i)}, s^2\right) \qquad \text{obs}$$

Ultrametric tree: Tip variance is constant equal to $\sigma^2 T$

$$h^2 = \frac{\sigma^2 T}{\sigma^2 T + s^2}$$

BM on a Non-Ultrametric Tree



$$\mathbf{Z}^r = \boldsymbol{\mu} \qquad \text{root}$$

$$\mathbf{Z}^j \mid \mathbf{Z}^{\mathrm{pa}(j)} \sim \mathcal{N}\left(\mathbf{Z}^{\mathrm{pa}(j)},\ \sigma^2 t_j\right) \qquad \text{nodes}$$

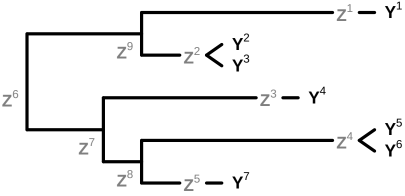
$$\mathbf{Y}^i \mid \mathbf{Z}^{\mathrm{pa}(i)} \sim \mathcal{N}\left(\mathbf{Z}^{\mathrm{pa}(i)}, s^2\right) \qquad \text{obs}$$

Ultrametric tree: Tip variance is constant equal to $\sigma^2 T$

$$h^2 = \frac{\sigma^2 T}{\sigma^2 T + s^2}$$

Non-Ultrametric tree: How to define tip variances ?

BM on a Non-Ultrametric Tree



$$\mathbf{Z}^r = \boldsymbol{\mu} \qquad \text{root}$$

$$\mathbf{Z}^j \mid \mathbf{Z}^{\text{pa}(j)} \sim \mathcal{N}\left(\mathbf{Z}^{\text{pa}(j)},\ \sigma^2 t_j\right) \qquad \text{nodes}$$

$$\mathbf{Y}^i \mid \mathbf{Z}^{\text{pa}(i)} \sim \mathcal{N}\left(\mathbf{Z}^{\text{pa}(i)}, s^2\right) \qquad \text{obs}$$

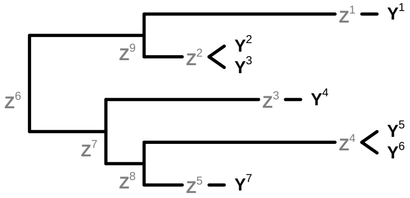
Ultrametric tree: Tip variance is constant equal to $\sigma^2 T$

$$h^2 = \frac{\sigma^2 T}{\sigma^2 T + s^2}$$

Non-Ultrametric tree: How to define tip variances ?

$$h^2 = \frac{\sigma^2 \bar{T}}{\sigma^2 \bar{T} + s^2} \quad \text{(Mitov and Stadler, 2018)}$$

BM on a Non-Ultrametric Tree



$$\mathbf{Z}^r = \boldsymbol{\mu} \qquad \text{root}$$

$$\mathbf{Z}^j \,\Big|\, \mathbf{Z}^{\mathrm{pa}(j)} \sim \mathcal{N}\left(\mathbf{Z}^{\mathrm{pa}(j)},\ \sigma^2 t_j\right) \qquad \text{nodes}$$

$$\mathbf{Y}^i \,\Big|\, \mathbf{Z}^{\mathrm{pa}(i)} \sim \mathcal{N}\left(\mathbf{Z}^{\mathrm{pa}(i)}, s^2\right) \qquad \text{obs}$$

Ultrametric tree: Tip variance is constant equal to $\sigma^2 T$

$$h^2 = \frac{\sigma^2 T}{\sigma^2 T + s^2}$$

Non-Ultrametric tree: How to define tip variances ?

$$h^2 = \frac{\sigma^2 \bar{T}}{\sigma^2 \bar{T} + s^2} \quad \text{(Mitov and Stadler, 2018)}$$

$$h^2 = \frac{\sigma^2 T_{max}}{\sigma^2 T_{max} + s^2} \quad \text{(Leventhal and Bonhoeffer, 2016)}$$

BM on a Non-Ultrametric Tree

Assume a non ultrametric tree.

BM on a Non-Ultrametric Tree

Assume a non ultrametric tree.

BM PMM total variance is:

$$\sigma^2\mathbf{V} + s^2\mathbf{I} \quad \text{with}\, \mathrm{Diag}(\mathbf{V}) = \mathrm{Diag}(T_1, \cdots, T_n)$$

BM on a Non-Ultrametric Tree

Assume a non ultrametric tree.

BM PMM total variance is:

$$\sigma^2\mathbf{V} + s^2\mathbf{I} \quad \text{with} \, \text{Diag}(\mathbf{V}) = \text{Diag}(T_1, \cdots, T_n)$$

Not a unique scale anymore.

BM on a Non-Ultrametric Tree

Assume a non ultrametric tree.

BM PMM total variance is:

$$\sigma^2\mathbf{V} + s^2\mathbf{I} \quad \text{with}\,\text{Diag}(\mathbf{V}) = \text{Diag}(T_1, \cdots, T_n)$$

Not a unique scale anymore.

Leventhal and Bonhoeffer (2016) write: $\sigma^2 T_{max}\mathbf{V}/T_{max} + s^2\mathbf{I}$.
The phylogenetic heritability is then: $h^2 = \frac{\sigma^2 T_{max}}{\sigma^2 T_{max} + s^2}$

BM on a Non-Ultrametric Tree

Assume a non ultrametric tree.

BM PMM total variance is:

$$\sigma^2\mathbf{V} + s^2\mathbf{I} \quad \text{with}\,\mathrm{Diag}(\mathbf{V}) = \mathrm{Diag}(T_1, \cdots, T_n)$$

Not a unique scale anymore.

Leventhal and Bonhoeffer (2016) write: $\sigma^2 T_{max}\mathbf{V}/T_{max} + s^2\mathbf{I}$.
The phylogenetic heritability is then: $h^2 = \frac{\sigma^2 T_{max}}{\sigma^2 T_{max} + s^2}$

Mitov and Stadler (2018) use $\bar{T}$ instead.
The phylogenetic heritability is then: $h^2 = \frac{\sigma^2 \bar{T}}{\sigma^2 \bar{T} + s^2}$

Pagel's Lambda on a Non-Ultrametric Tree

Assume a non ultrametric tree.

Pagel's Lambda on a Non-Ultrametric Tree

Assume a non ultrametric tree.

Pagel's lambda variance is:

$$\sigma_\lambda^2\mathbf{V}(\lambda) = \sigma_\lambda^2\left[\lambda\mathbf{V} + (1-\lambda)\mathbf{V}^*\right]$$

with $\mathbf{V}^* = \mathrm{Diag}(\mathbf{V}) = \mathrm{Diag}(T_1, \cdots, T_n)$.

Pagel's Lambda on a Non-Ultrametric Tree

Assume a non ultrametric tree.

Pagel's lambda variance is:

$$\sigma_\lambda^2 \mathbf{V}(\lambda) = \sigma_\lambda^2 \left[\lambda \mathbf{V} + (1-\lambda)\mathbf{V}^*\right]$$

with $\mathbf{V}^* = \mathrm{Diag}(\mathbf{V}) = \mathrm{Diag}(T_1, \cdots, T_n)$.

Can λ be used as a measure of heritability in this case ?

Pagel's Lambda on a Non-Ultrametric Tree

Assume a non ultrametric tree.

Pagel's lambda variance is:

$$\sigma_\lambda^2 \mathbf{V}(\lambda) = \sigma_\lambda^2 \left[\lambda \mathbf{V} + (1-\lambda)\mathbf{V}^*\right]$$

with $\mathbf{V}^* = \text{Diag}(\mathbf{V}) = \text{Diag}(T_1, \cdots, T_n)$.

Can λ be used as a measure of heritability in this case ?

BM PMM with scaled variance is:

$$\sigma^2 \mathbf{V} + s_s^2 \mathbf{V}^*$$

Error accumulate in time (Leventhal and Bonhoeffer, 2016)
a tip further from the root has a larger error.

Pagel's Lambda on a Non-Ultrametric Tree

Pagel's lambda variance:

$$\sigma_\lambda^2 \mathbf{V}(\lambda) = \sigma_\lambda^2 \left[\lambda \mathbf{V} + (1-\lambda)\mathbf{V}^*\right]$$

BM PMM with scaled variance:

$$\sigma^2 \mathbf{V} + s_s^2 \mathbf{V}^* = (\sigma^2 + s_s^2) \left[\frac{\sigma^2}{\sigma^2 + s_s^2}\mathbf{V} + (1 - \frac{\sigma^2}{\sigma^2 + s_s^2})\mathbf{V}^*\right]$$

Pagel's Lambda on a Non-Ultrametric Tree

Pagel's lambda variance:

$$\sigma_\lambda^2 \mathbf{V}(\lambda) = \sigma_\lambda^2 \left[\lambda \mathbf{V} + (1-\lambda)\mathbf{V}^*\right]$$

BM PMM with scaled variance:

$$\sigma^2 \mathbf{V} + s_s^2 \mathbf{V}^* = (\sigma^2 + s_s^2)\left[\frac{\sigma^2}{\sigma^2 + s_s^2}\mathbf{V} + (1 - \frac{\sigma^2}{\sigma^2 + s_s^2})\mathbf{V}^*\right]$$

The two models are equivalent with:

$$\lambda = \frac{\sigma^2}{\sigma^2 + s_s^2} \qquad \sigma_\lambda^2 = \sigma^2 + s_s^2$$

Pagel's Lambda on a Non-Ultrametric Tree

Pagel's lambda variance:

$$\sigma_\lambda^2 \mathbf{V}(\lambda) = \sigma_\lambda^2 \left[\lambda \mathbf{V} + (1-\lambda)\mathbf{V}^*\right]$$

BM PMM with scaled variance:

$$\sigma^2 \mathbf{V} + s_s^2 \mathbf{V}^* = (\sigma^2 + s_s^2)\left[\frac{\sigma^2}{\sigma^2 + s_s^2}\mathbf{V} + (1 - \frac{\sigma^2}{\sigma^2 + s_s^2})\mathbf{V}^*\right]$$

The two models are equivalent with:

$$\lambda = \frac{\sigma^2}{\sigma^2 + s_s^2} \qquad \sigma_\lambda^2 = \sigma^2 + s_s^2$$

λ is the heritability if the independent errors are scaled.

BM Heritability

Ultrametric Tree

$$h^2 = \frac{\sigma^2 T}{\sigma^2 T + s^2} = \lambda$$

Non Ultrametric Tree

$$h^2 = \frac{\sigma^2 T_{max}}{\sigma^2 T_{max} + s^2}$$

$$h^2 = \frac{\sigma^2 \bar{T}}{\sigma^2 \bar{T} + s^2}$$

$$h^2 = \frac{\sigma^2}{\sigma^2 + s_s^2} = \lambda$$

General Mixed Model

(Blanquart et al., 2017)

$$\mathbf{Y} = \mathbf{c} + \mathbf{m}(\boldsymbol{\theta}) + \mathbf{E}^p(\boldsymbol{\theta}) + \mathbf{E}^e$$

with:

$\mathbf{c}$	vector of covariables
$\mathbf{m}(\boldsymbol{\theta})$	expectation of the phylogenetic model
$\mathbf{E}^p(\boldsymbol{\theta}) \sim \mathcal{N}(\mathbf{0}_n, \boldsymbol{\Sigma}(\boldsymbol{\theta}))$	variance of the phylogenetic model
$\mathbf{E}^e \sim \mathcal{N}(\mathbf{0}_n, s^2\mathbf{I})$	envorinemental variance

General Mixed Model

(Blanquart et al., 2017)

$$\mathbf{Y} = \mathbf{c} + \mathbf{m}(\boldsymbol{\theta}) + \mathbf{E}^p(\boldsymbol{\theta}) + \mathbf{E}^e$$

with:

$\mathbf{c}$	vector of covariables
$\mathbf{m}(\boldsymbol{\theta})$	expectation of the phylogenetic model
$\mathbf{E}^p(\boldsymbol{\theta}) \sim \mathcal{N}(\mathbf{0}_n, \boldsymbol{\Sigma}(\boldsymbol{\theta}))$	variance of the phylogenetic model
$\mathbf{E}^e \sim \mathcal{N}(\mathbf{0}_n, s^2\mathbf{I})$	envorinemental variance

Example: BM

$$\mathbf{m}(\boldsymbol{\theta})_i = \mu$$
$$\boldsymbol{\Sigma}(\boldsymbol{\theta})_{ij} = \sigma^2 V_{ij}$$

General Mixed Model

(Blanquart et al., 2017)

$$\mathbf{Y} = \mathbf{c} + \mathbf{m}(\boldsymbol{\theta}) + \mathbf{E}^p(\boldsymbol{\theta}) + \mathbf{E}^e$$

with:

$\mathbf{c}$	vector of covariables
$\mathbf{m}(\boldsymbol{\theta})$	expectation of the phylogenetic model
$\mathbf{E}^p(\boldsymbol{\theta}) \sim \mathcal{N}(\mathbf{0}_n, \boldsymbol{\Sigma}(\boldsymbol{\theta}))$	variance of the phylogenetic model
$\mathbf{E}^e \sim \mathcal{N}(\mathbf{0}_n, s^2\mathbf{I})$	envorinemental variance

Example: OU

$$\mathbf{m}(\boldsymbol{\theta})_i = \mu e^{-\alpha V_i} + \beta(1 - e^{-\alpha V_i})$$

$$\boldsymbol{\Sigma}(\boldsymbol{\theta})_{ij} = \frac{\sigma^2}{2\alpha} e^{-\alpha(V_i+V_j)}(e^{2\alpha V_{ij}} - 1)$$

Empirical Phylogenetic Heritability (Blanquart et al., 2017)

$$\mathbf{Y} = \mathbf{c} + \mathbf{m}(\boldsymbol{\theta}) + \mathbf{E}^p(\boldsymbol{\theta}) + \mathbf{E}^e$$

Empirical heritability

$$H^2_{emp} = \frac{\mathbb{S}[\mathbf{m}(\boldsymbol{\theta}) + \mathbf{E}^p(\boldsymbol{\theta})]}{\mathbb{S}[\mathbf{Y}]}$$

with $\mathbb{S}$ the *empirical* variance: $\mathbb{S}[\mathbf{x}] = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \frac{1}{n}\sum_{i=1}^{n} x_i\right)^2$

Empirical Phylogenetic Heritability (Blanquart et al., 2017)

$$\mathbf{Y} = \mathbf{c} + \mathbf{m}(\boldsymbol{\theta}) + \mathbf{E}^p(\boldsymbol{\theta}) + \mathbf{E}^e$$

Empirical heritability

$$H^2_{emp} = \frac{\mathbb{S}[\mathbf{m}(\boldsymbol{\theta}) + \mathbf{E}^p(\boldsymbol{\theta})]}{\mathbb{S}[\mathbf{Y}]}$$

with $\mathbb{S}$ the *empirical* variance: $\mathbb{S}[\mathbf{x}] = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \frac{1}{n}\sum_{i=1}^{n} x_i\right)^2$

H^2_{emp} is a stochastic quantity.

Empirical Phylogenetic Heritability (Blanquart et al., 2017)

$$\mathbf{Y} = \mathbf{c} + \mathbf{m}(\boldsymbol{\theta}) + \mathbf{E}^p(\boldsymbol{\theta}) + \mathbf{E}^e$$

Empirical heritability

$$H^2_{emp} = \frac{\mathbb{S}[\mathbf{m}(\boldsymbol{\theta}) + \mathbf{E}^p(\boldsymbol{\theta})]}{\mathbb{S}[\mathbf{Y}]}$$

with $\mathbb{S}$ the *empirical* variance: $\mathbb{S}[\mathbf{x}] = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \frac{1}{n}\sum_{i=1}^{n} x_i\right)^2$

H^2_{emp} is a stochastic quantity.

Can be computed through simulations.

Mean Empirical Heritability (Blanquart et al., 2017)

$$\mathbf{Y} = \mathbf{c} + \mathbf{m}(\boldsymbol{\theta}) + \mathbf{E}^p(\boldsymbol{\theta}) + \mathbf{E}^e$$

Empirical heritability

$$H^2_{emp} = \frac{\mathbb{S}[\mathbf{m}(\boldsymbol{\theta}) + \mathbf{E}^p(\boldsymbol{\theta})]}{\mathbb{S}[\mathbf{Y}]}$$

Mean Empirical heritability

$$h^2_{emp} = \mathbb{E}\left[\frac{\mathbb{S}[\mathbf{m}(\boldsymbol{\theta}) + \mathbf{E}^p(\boldsymbol{\theta})]}{\mathbb{S}[\mathbf{Y}]}\right] \approx \frac{\mathbb{E}\left[\mathbb{S}[\mathbf{m}(\boldsymbol{\theta}) + \mathbf{E}^p(\boldsymbol{\theta})]\right]}{\mathbb{E}\left[\mathbb{S}[\mathbf{Y}]\right]}$$

Mean Empirical Heritability

(Blanquart et al., 2017)

$$\mathbf{Y} = \mathbf{c} + \mathbf{m}(\boldsymbol{\theta}) + \mathbf{E}^p(\boldsymbol{\theta}) + \mathbf{E}^e$$

Empirical heritability

$$H^2_{emp} = \frac{\mathbb{S}[\mathbf{m}(\boldsymbol{\theta}) + \mathbf{E}^p(\boldsymbol{\theta})]}{\mathbb{S}[\mathbf{Y}]}$$

Mean Empirical heritability

$$h^2_{emp} = \mathbb{E}\left[\frac{\mathbb{S}[\mathbf{m}(\boldsymbol{\theta}) + \mathbf{E}^p(\boldsymbol{\theta})]}{\mathbb{S}[\mathbf{Y}]}\right] \approx \frac{\mathbb{E}\left[\mathbb{S}[\mathbf{m}(\boldsymbol{\theta}) + \mathbf{E}^p(\boldsymbol{\theta})]\right]}{\mathbb{E}\left[\mathbb{S}[\mathbf{Y}]\right]}$$

Expectation must be analytically computed for each model.

BM Mean Empirical Heritability

(Hassler et al., 2020)

$$\mathbf{Y} = \mu\mathbf{1} + \sigma\mathbf{E}^p + s\mathbf{E}^e$$

Mean Empirical Variance

$$\mathbb{E}\left[\mathbb{S}[\mathbf{Y}]\right] = \frac{n}{n-1}s^2 + \left(\frac{1}{n}\operatorname{tr}(\mathbf{V}) - \frac{1}{n^2}\mathbf{1}^T\mathbf{V}\mathbf{1}\right)\sigma^2$$

Mean Empirical heritability

$$h^2_{emp} \approx \frac{\left(\frac{1}{n}\operatorname{tr}(\mathbf{V}) - \frac{1}{n^2}\mathbf{1}^T\mathbf{V}\mathbf{1}\right)\sigma^2}{\frac{n}{n-1}s^2 + \left(\frac{1}{n}\operatorname{tr}(\mathbf{V}) - \frac{1}{n^2}\mathbf{1}^T\mathbf{V}\mathbf{1}\right)\sigma^2}$$

BM Mean Empirical Heritability (Hassler et al., 2020)

$$\mathbf{Y} = \mu \mathbf{1} + \sigma \mathbf{E}^p + s \mathbf{E}^e$$

Mean Empirical Variance

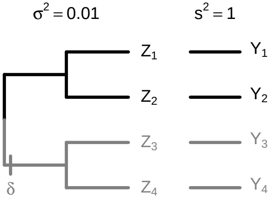
$$\mathbb{E}\left[\mathbb{S}[\mathbf{Y}]\right] = \frac{n}{n-1} s^2 + \left(\frac{1}{n} \operatorname{tr}(\mathbf{V}) - \frac{1}{n^2} \mathbf{1}^T \mathbf{V} \mathbf{1} \right) \sigma^2$$

Mean Empirical heritability

$$h^2_{emp} \approx \frac{\left(\frac{1}{n} \operatorname{tr}(\mathbf{V}) - \frac{1}{n^2} \mathbf{1}^T \mathbf{V} \mathbf{1} \right) \sigma^2}{\frac{n}{n-1} s^2 + \left(\frac{1}{n} \operatorname{tr}(\mathbf{V}) - \frac{1}{n^2} \mathbf{1}^T \mathbf{V} \mathbf{1} \right) \sigma^2}$$

Yet again an other definition of the BM heritability.

Shifted BM Mean Empirical Heritability

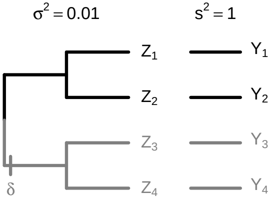


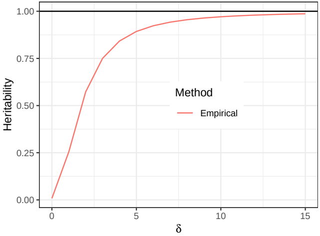
$$\mathbf{Y} = \mu\mathbf{1} + \delta\mathbf{1}_{n/2} + \sigma\mathbf{E}^{p} + s\mathbf{E}^{e}$$

Mean Empirical heritability

$$h^2_{emp} \approx \frac{\left(\frac{1}{n}\operatorname{tr}(\mathbf{V}) - \frac{1}{n^2}\mathbf{1}^T\mathbf{V}\mathbf{1}\right)\sigma^2 + \frac{\delta^2}{4}}{\frac{n}{n-1}s^2 + \left(\frac{1}{n}\operatorname{tr}(\mathbf{V}) - \frac{1}{n^2}\mathbf{1}^T\mathbf{V}\mathbf{1}\right)\sigma^2 + \frac{\delta^2}{4}}$$

Shifted BM Mean Empirical Heritability





$$\mathbf{Y} = \mu \mathbf{1} + \delta \mathbf{1}_{n/2} + \sigma \mathbf{E}^p + s \mathbf{E}^e$$

Mean Empirical heritability

$$h^2_{emp} \approx \frac{\left(\frac{1}{n}\operatorname{tr}(\mathbf{V}) - \frac{1}{n^2}\mathbf{1}^T\mathbf{V}\mathbf{1}\right)\sigma^2 + \frac{\delta^2}{4}}{\frac{n}{n-1}s^2 + \left(\frac{1}{n}\operatorname{tr}(\mathbf{V}) - \frac{1}{n^2}\mathbf{1}^T\mathbf{V}\mathbf{1}\right)\sigma^2 + \frac{\delta^2}{4}}$$

Population Heritability

(Bastide et al., 2020)

$$\mathbf{Y} = \mathbf{c} + \mathbf{m}(\boldsymbol{\theta}) + \mathbf{E}^p(\boldsymbol{\theta}) + \mathbf{E}^e$$

Mean Empirical heritability

$$h^2_{emp} \approx \frac{\mathbb{E}\left[\mathbb{S}[\mathbf{m}(\boldsymbol{\theta}) + \mathbf{E}^p(\boldsymbol{\theta})]\right]}{\mathbb{E}\left[\mathbb{S}[\mathbf{Y}]\right]} \quad \text{with:} \quad \mathbb{S}[\mathbf{x}] = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

Population Heritability

(Bastide et al., 2020)

$$\mathbf{Y} = \mathbf{c} + \mathbf{m}(\boldsymbol{\theta}) + \mathbf{E}^p(\boldsymbol{\theta}) + \mathbf{E}^e$$

Mean Empirical heritability

$$h^2_{emp} \approx \frac{\mathbb{E}\left[\mathbb{S}[\mathbf{m}(\boldsymbol{\theta}) + \mathbf{E}^p(\boldsymbol{\theta})]\right]}{\mathbb{E}\left[\mathbb{S}[\mathbf{Y}]\right]} \quad \text{with:} \quad \mathbb{S}[\mathbf{x}] = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

Population heritability

$$h^2_{pop} = \frac{\mathbb{E}\left[\mathbb{S}_{pop}[\mathbf{m}(\boldsymbol{\theta}) + \mathbf{E}^p(\boldsymbol{\theta})]\right]}{\mathbb{E}\left[\mathbb{S}_{pop}[\mathbf{Y}]\right]}$$

with: $\mathbb{S}_{pop}[\mathbf{x}] = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mathbb{E}[x_i])^2$ $\qquad$ $\mathbb{E}\left[\mathbb{S}_{pop}[\mathbf{x}]\right] = \frac{1}{n}\sum_{i=1}^{n}\mathbb{V}\mathrm{ar}[x_i]$

BM Population Heritability

(Bastide et al., 2020)

$$\mathbf{Y} = \mu \mathbf{1} + \sigma \mathbf{E}^p + s\mathbf{E}^e$$

Mean Population Variance

$$\mathbb{E}\left[\mathbb{S}_{pop}[\mathbf{Y}]\right] = \frac{1}{n}\sum_{i=1}^{n} \mathbb{V}\mathrm{ar}\left[Y_i\right] = \frac{1}{n}\sum_{i=1}^{n}(\sigma^2 T_i + s^2) = \sigma^2 \bar{T} + s^2$$

BM Population Heritability

(Bastide et al., 2020)

$$\mathbf{Y} = \mu\mathbf{1} + \sigma\mathbf{E}^p + s\mathbf{E}^e$$

Mean Population Variance

$$\mathbb{E}\left[\mathbb{S}_{pop}[\mathbf{Y}]\right] = \frac{1}{n}\sum_{i=1}^{n}\mathbb{V}\mathrm{ar}\left[Y_i\right] = \frac{1}{n}\sum_{i=1}^{n}(\sigma^2 T_i + s^2) = \sigma^2\bar{T} + s^2$$

Population heritability

$$h^2_{pop} = \frac{\sigma^2\bar{T}}{\sigma^2\bar{T} + s^2}$$

BM Population Heritability

(Bastide et al., 2020)

$$\mathbf{Y} = \mu\mathbf{1} + \sigma\mathbf{E}^p + s\mathbf{E}^e$$

Mean Population Variance

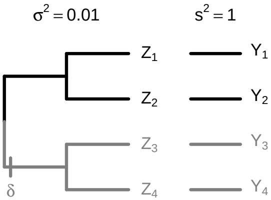
$$\mathbb{E}\left[\mathbb{S}_{pop}[\mathbf{Y}]\right] = \frac{1}{n}\sum_{i=1}^{n}\mathbb{V}\mathrm{ar}\left[Y_i\right] = \frac{1}{n}\sum_{i=1}^{n}(\sigma^2 T_i + s^2) = \sigma^2\bar{T} + s^2$$

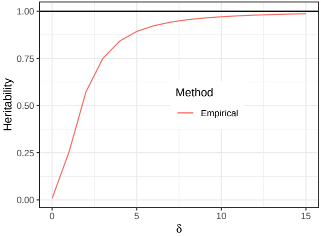
Population heritability

$$h^2_{pop} = \frac{\sigma^2\bar{T}}{\sigma^2\bar{T} + s^2}$$

Same definition as Mitov and Stadler (2018).

Shifted BM Population Heritability

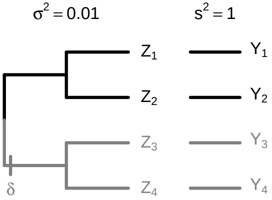


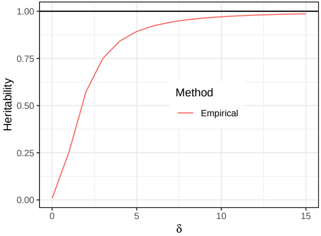


Mean Empirical heritability

$$h^2_{emp} \approx \frac{\left(\frac{1}{n}\operatorname{tr}(\mathbf{V}) - \frac{1}{n^2}\mathbf{1}^T\mathbf{V}\mathbf{1}\right)\sigma^2 + \frac{\delta^2}{4}}{\frac{n}{n-1}s^2 + \left(\frac{1}{n}\operatorname{tr}(\mathbf{V}) - \frac{1}{n^2}\mathbf{1}^T\mathbf{V}\mathbf{1}\right)\sigma^2 + \frac{\delta^2}{4}}$$

Shifted BM Population Heritability





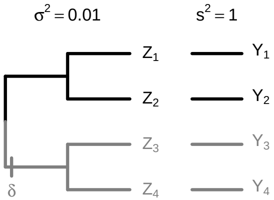
Mean Empirical heritability

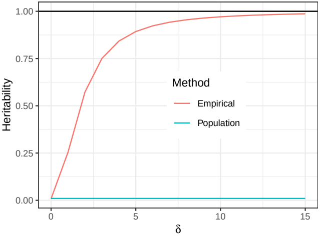
$$h^2_{emp} \approx \frac{\left(\frac{1}{n}\operatorname{tr}(\mathbf{V}) - \frac{1}{n^2}\mathbf{1}^T\mathbf{V}\mathbf{1}\right)\sigma^2 + \frac{\delta^2}{4}}{\frac{n}{n-1}s^2 + \left(\frac{1}{n}\operatorname{tr}(\mathbf{V}) - \frac{1}{n^2}\mathbf{1}^T\mathbf{V}\mathbf{1}\right)\sigma^2 + \frac{\delta^2}{4}}$$

Mean Population Variance unchanged by the shift

$$\mathbb{E}\left[\mathbb{S}_{pop}[\mathbf{Y}]\right] = \frac{1}{n}\sum_{i=1}^{n}\mathbb{V}\mathrm{ar}\left[Y_i\right] = \frac{1}{n}\sum_{i=1}^{n}(\sigma^2 T_i + s^2) = \sigma^2 \bar{T} + s^2$$

Shifted BM Population Heritability





Mean Empirical heritability

$$h^2_{emp} \approx \frac{\left(\frac{1}{n}\operatorname{tr}(\mathbf{V}) - \frac{1}{n^2}\mathbf{1}^T\mathbf{V}\mathbf{1}\right)\sigma^2 + \frac{\delta^2}{4}}{\frac{n}{n-1}s^2 + \left(\frac{1}{n}\operatorname{tr}(\mathbf{V}) - \frac{1}{n^2}\mathbf{1}^T\mathbf{V}\mathbf{1}\right)\sigma^2 + \frac{\delta^2}{4}}$$

Population heritability unchanged

$$h^2_{pop} = \frac{\sigma^2\,\bar{T}}{\sigma^2\,\bar{T} + s^2}$$

Heritability

$$\mathbf{Y} = \mathbf{c} + \mathbf{m}(\boldsymbol{\theta}) + \mathbf{E}^p(\boldsymbol{\theta}) + \mathbf{E}^e$$

Empirical heritability: stochastic, affected by expectation

$$H^2_{emp} = \frac{\mathbb{S}[\mathbf{m}(\boldsymbol{\theta}) + \mathbf{E}^p(\boldsymbol{\theta})]}{\mathbb{S}[\mathbf{Y}]} \quad \text{with:} \quad \mathbb{S}[\mathbf{x}] = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

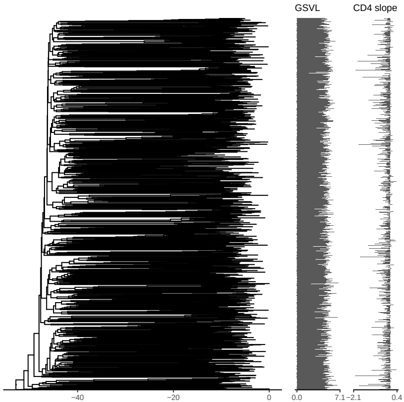
Mean Empirical heritability: affected by expectation

$$h^2_{emp} = \frac{\mathbb{E}\left[\mathbb{S}[\mathbf{m}(\boldsymbol{\theta}) + \mathbf{E}^p(\boldsymbol{\theta})]\right]}{\mathbb{E}\left[\mathbb{S}[\mathbf{Y}]\right]}$$

Population heritability: unaffected by expectation.

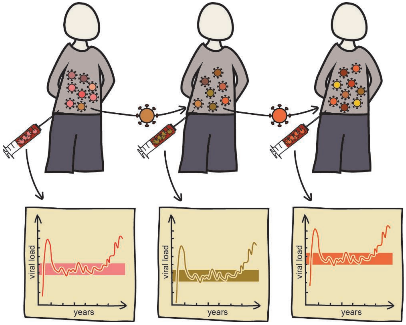
$$h^2_{pop} = \frac{\mathbb{E}\left[\mathbb{S}_{pop}[\mathbf{m}(\boldsymbol{\theta}) + \mathbf{E}^p(\boldsymbol{\theta})]\right]}{\mathbb{E}\left[\mathbb{S}_{pop}[\mathbf{Y}]\right]} \quad \text{with:} \quad \mathbb{E}\left[\mathbb{S}_{pop}[\mathbf{x}]\right] = \frac{1}{n}\sum_{i=1}^{n}\mathbb{V}\text{ar}\left[x_i\right]$$

HIV Virulence Heritability (Alizon et al., 2010; Blanquart et al., 2017)



HIV-1, Sub-type B, MSM, 1171 tips

CD4: T cells decline rate
GSVL: viral load



Fraser et al. (2014)

Questions: Is virulence “heritable”? Which model of trait evolution?

HIV Virulence Heritability

Alizon et al. (2010)
Data: Swiss HIV Cohort (MSM, $n = 134$)
Trait: spVL
Method: Pagel's lambda
Heritability: 51% [±27%]

HIV Virulence Heritability

Alizon et al. (2010)
Data: Swiss HIV Cohort (MSM, $n = 134$)
Trait: spVL
Method: Pagel's lambda
Heritability: 51% [$\pm$27%]

Vrancken et al. (2015)
Data: Swiss HIV Cohort (MSM, $n = 134$)
Trait: spVL
Method: Pagel's lambda
Heritability: 50.1% $[16.5\%, 85.7\%]$

HIV Virulence Heritability

Alizon et al. (2010)
Data: Swiss HIV Cohort (MSM, $n = 134$)
Trait: spVL
Method: Pagel's lambda
Heritability: 51% [±27%]

Hodcroft et al. (2014)
Data: UK HIV Drug Resistance Database ($n = 8,483$)
Trait: spVL
Method: BM with $\bar{T}$ (population)
Heritability: 5.8% $[2.9\%, 8.7\%]$

Vrancken et al. (2015)
Data: Swiss HIV Cohort (MSM, $n = 134$)
Trait: spVL
Method: Pagel's lambda
Heritability: 50.1% $[16.5\%, 85.7\%]$

HIV Virulence Heritability

Alizon et al. (2010)
Data: Swiss HIV Cohort (MSM, $n = 134$)
Trait: spVL
Method: Pagel's lambda
Heritability: 51% [$\pm$27%]

Hodcroft et al. (2014)
Data: UK HIV Drug Resistance Database ($n = 8,483$)
Trait: spVL
Method: BM with $\bar{T}$ (population)
Heritability: 5.8% [2.9%, 8.7%]

Vrancken et al. (2015)
Data: Swiss HIV Cohort (MSM, $n = 134$)
Trait: spVL
Method: Pagel's lambda
Heritability: 50.1% [16.5%, 85.7%]

Blanquart et al. (2017)
Data: BEEHIVE (MSM, $n = 1,196$)
Trait: GSVL, CD4
Method: OU-BM, empirical
Heritability GSVL: 30% [8%, 44%] (OU)
Heritability CD4: 11% [0%, 19%] (BM)

HIV Virulence Heritability

Alizon et al. (2010)
Data: Swiss HIV Cohort (MSM, $n = 134$)
Trait: spVL
Method: Pagel's lambda
Heritability: 51% [±27%]

Hodcroft et al. (2014)
Data: UK HIV Drug Resistance Database ($n = 8,483$)
Trait: spVL
Method: BM with $\bar{T}$ (population)
Heritability: 5.8% $[2.9\%, 8.7\%]$

Hassler et al. (2020)
Data: BEEHIVE (MSM, $n = 1,171$)
Trait: GSVL, CD4
Method: BM, mean empirical
Heritability GSVL: 21% $[11\%, 30\%]$
Heritability CD4: 16% $[7\%, 35\%]$

Vrancken et al. (2015)
Data: Swiss HIV Cohort (MSM, $n = 134$)
Trait: spVL
Method: Pagel's lambda
Heritability: 50.1% $[16.5\%, 85.7\%]$

Blanquart et al. (2017)
Data: BEEHIVE (MSM, $n = 1,196$)
Trait: GSVL, CD4
Method: OU-BM, empirical
Heritability GSVL: 30% $[8\%, 44\%]$ (OU)
Heritability CD4: 11% $[0\%, 19\%]$ (BM)

HIV Virulence Heritability

Alizon et al. (2010)
Data: Swiss HIV Cohort (MSM, $n = 134$)
Trait: spVL
Method: Pagel's lambda
Heritability: 51% [±27%]

Hodcroft et al. (2014)
Data: UK HIV Drug Resistance Database ($n = 8,483$)
Trait: spVL
Method: BM with $\bar{T}$ (population)
Heritability: 5.8% [2.9%, 8.7%]

Hassler et al. (2020)
Data: BEEHIVE (MSM, $n = 1,171$)
Trait: GSVL, CD4
Method: BM, mean empirical
Heritability GSVL: 21% [11%, 30%]
Heritability CD4: 16% [7%, 35%]

Vrancken et al. (2015)
Data: Swiss HIV Cohort (MSM, $n = 134$)
Trait: spVL
Method: Pagel's lambda
Heritability: 50.1% [16.5%, 85.7%]

Blanquart et al. (2017)
Data: BEEHIVE (MSM, $n = 1,196$)
Trait: GSVL, CD4
Method: OU-BM, empirical
Heritability GSVL: 30% [8%, 44%] (OU)
Heritability CD4: 11% [0%, 19%] (BM)

Bastide et al. (2020)
Data: BEEHIVE (MSM, $n = 1,171$)
Trait: GSVL, CD4
Method: OU-BM, population
Heritability GSVL: 43% [28%, 59%] (OU)
Heritability CD4: 27% [15%, 38%] (BM)

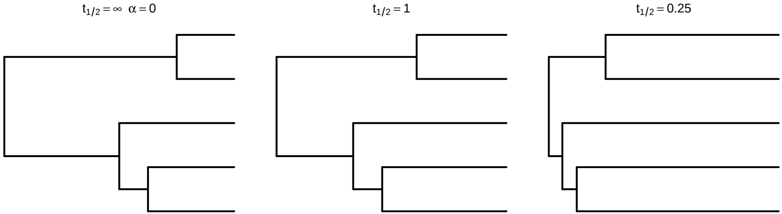
HIV Virulence Heritability

- Estimates vary a lot.
- Data is difficult to interpret (spVL, heritability).
- Various definitions of heritability are used.
- OU always give higher heritability estimates.

BM vs OU Heritability

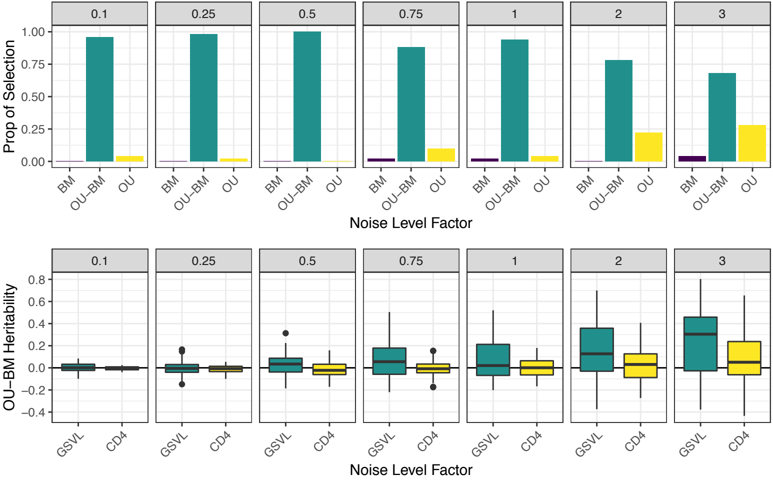
$$\mathbf{Y} = \mathbf{c} + \mathbf{m}(\theta) + \mathbf{E}^p(\theta) + \mathbf{E}^e$$

OU Phylogenetic Model allows for more individual variations.



→ heritability is higher for the OU.

OU-BM Simulations



Bibliography

Alizon, von Wyl, Stadler, et al. 2010. *PLoS Pathogens*. 6.

Bastide, Ho, Baele, et al. 2020. *Annals of Applied Statistics (in press)*. .

Blanquart, Wymant, Cornelissen, et al. 2017. *PLoS Biology*. 15:1–26.

Fraser, Lythgoe, Leventhal, et al. 2014. *Science*. 343:1243727–1243727.

Hassler, Tolkoff, Allen, et al. 2020. *Journal of the American Statistical Association*. pp. 1–15.

Ho, Ané. 2014. *Systematic Biology*. 63:397–408.

Hodcroft, Hadfield, Fearnhill, et al. 2014. *PLoS Pathogens*. 10.

Housworth, Martins, Lynch. 2004. *The American Naturalist*. 163:84–96.

Leventhal, Bonhoeffer. 2016. *Trends in Microbiology*. 24:687–698.

Lynch. 1991. *Evolution*. 45:1065–1080.

Mitov, Stadler. 2018. *Molecular Biology and Evolution*. 35:756–772.

Pagel. 1999. *Nature*. 401:877–884.

Vrancken, Lemey, Rambaut, et al. 2015. *Methods in Ecology and Evolution*. 6:67–82.

Thank you for listening

Institut Montpelliérain Alexander Grothendieck

Appendices

Three Point Structure Algorithm (Ho and Ané, 2014)

$$\mathbf{Y} = \mathbf{X}\beta + \sigma\mathbf{E} \quad \mathbf{E} \sim \mathcal{N}(\mathbf{0}_n, \mathbf{V})$$

Estimators:

$$\hat{\beta} = (\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}^T\mathbf{V}^{-1}\mathbf{Y}$$

$$\hat{\sigma}^2 = \frac{1}{n-p}(\mathbf{Y} - \mathbf{X}\hat{\beta})^T\mathbf{V}^{-1}(\mathbf{Y} - \mathbf{X}\hat{\beta})$$

$$2L(\mathbf{y}|\boldsymbol{\theta}) = n\log(2\pi) + \log|\mathbf{V}| + (\mathbf{Y} - \mathbf{X}\hat{\beta})^T\mathbf{V}^{-1}(\mathbf{Y} - \mathbf{X}\hat{\beta})$$

Three Point Structure Algorithm (Ho and Ané, 2014)

$$\mathbf{Y} = \mathbf{X}\beta + \sigma\mathbf{E} \quad \mathbf{E} \sim \mathcal{N}(\mathbf{0}_n, \mathbf{V})$$

Estimators:

$$\begin{aligned}
\hat{\beta} &= (\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}^T\mathbf{V}^{-1}\mathbf{Y} \\
\hat{\sigma}^2 &= \frac{1}{n-p}(\mathbf{Y} - \mathbf{X}\hat{\beta})^T\mathbf{V}^{-1}(\mathbf{Y} - \mathbf{X}\hat{\beta}) \\
2L(\mathbf{y}|\boldsymbol{\theta}) &= n\log(2\pi) + \log|\mathbf{V}| + (\mathbf{Y} - \mathbf{X}\hat{\beta})^T\mathbf{V}^{-1}(\mathbf{Y} - \mathbf{X}\hat{\beta})
\end{aligned}$$

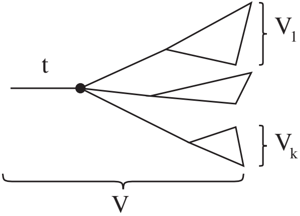
We need to compute:

- $\log|\mathbf{V}|$
- $\mathbf{Q} = \mathbf{X}^T\mathbf{V}^{-1}\mathbf{Y}$
- $p = \mathbf{1}^T\mathbf{V}^{-1}\mathbf{1}$, $\hat{\mu}_Y = \mathbf{1}^T\mathbf{V}^{-1}\mathbf{Y}/p$ and $\hat{\boldsymbol{\mu}}_X^T = \mathbf{X}^T\mathbf{V}^{-1}\mathbf{1}/p$

Three Point Structure Algorithm (Ho and Ané, 2014)

Three Points Structure:

$$\mathbf{V} = t\mathbf{1}^T\mathbf{1} + \begin{pmatrix} V_1 & & 0 \\ & \ddots & \\ 0 & & V_k \end{pmatrix}$$



Ho and Ané (2014)

Three Point Structure Algorithm (Ho and Ané, 2014)

Initialization : tree with a single tip of length t

- $\log |\mathbf{V}| = t$
- $\mathbf{Q} = \mathbf{x}^T y/t$
- $p = 1/t$, $\hat{\mu}_Y = y$ and $\hat{\boldsymbol{\mu}}_X^T = \mathbf{x}^T$

Three Point Structure Algorithm (Ho and Ané, 2014)

Propagation:

Woodbury and Sylvester formulas for $\mathbf{M} = \mathbf{A} + \mathbf{UCV}$

$$\mathbf{M}^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{U}(\mathbf{C}^{-1} + \mathbf{V}\mathbf{A}^{-1}\mathbf{U})^{-1}\mathbf{V}\mathbf{A}^{-1}$$

$$|\mathbf{M}| = |\mathbf{A}|\,|\mathbf{C}|\,\left|\mathbf{C}^{-1} + \mathbf{V}\mathbf{A}^{-1}\mathbf{U}\right|$$

Three Point Structure Algorithm (Ho and Ané, 2014)

Propagation:
Woodbury and Sylvester formulas for $\mathbf{M} = \mathbf{A} + \mathbf{UCV}$

$$\mathbf{M}^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{U}(\mathbf{C}^{-1} + \mathbf{V}\mathbf{A}^{-1}\mathbf{U})^{-1}\mathbf{V}\mathbf{A}^{-1}$$

$$|\mathbf{M}| = |\mathbf{A}|\,|\mathbf{C}|\,\left|\mathbf{C}^{-1} + \mathbf{V}\mathbf{A}^{-1}\mathbf{U}\right|$$

Applied on $\mathbf{V} = \mathbf{A} + t\mathbf{1}^T\mathbf{1}$, with $\mathbf{A} = \begin{pmatrix} V_1 & & 0 \\ & \ddots & \\ 0 & & V_k \end{pmatrix}$:

$$\mathbf{V}^{-1} = \mathbf{A}^{-1} - \frac{t}{1 + tp_A}\mathbf{A}^{-1}\mathbf{1}\mathbf{1}^T\mathbf{A}^{-1}, \qquad p_A = \mathbf{1}^T\mathbf{A}^{-1}\mathbf{1} = \sum_{s=1}^{k} p_s$$

Three Point Structure Algorithm (Ho and Ané, 2014)

Propagation:

Woodbury and Sylvester formulas for $\mathbf{M} = \mathbf{A} + \mathbf{UCV}$

$$\mathbf{M}^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{U}(\mathbf{C}^{-1} + \mathbf{VA}^{-1}\mathbf{U})^{-1}\mathbf{VA}^{-1}$$

$$|\mathbf{M}| = |\mathbf{A}|\,|\mathbf{C}|\,\left|\mathbf{C}^{-1} + \mathbf{VA}^{-1}\mathbf{U}\right|$$

Gives:

- $\log|\mathbf{V}| = \sum_{s=1}^{k} \log|\mathbf{V}_s| + \log(1 + tp_A)$
- $\mathbf{Q} = \sum_{s=1}^{k} \mathbf{Q}_s - \frac{tp_A^2}{1+tp_A} \hat{\boldsymbol{\mu}}_X^T \hat{\mu}_Y$
- $p = \frac{p_A}{1+tp_A}$, $\hat{\mu}_Y = \sum_{s=1}^{k} \frac{p_s}{p_A} \hat{\mu}_{Y,s}$ and $\hat{\boldsymbol{\mu}}_X^T = \sum_{s=1}^{k} \frac{p_s}{p_A} \hat{\boldsymbol{\mu}}_{X,s}$

Three Point Structure Algorithm (Ho and Ané, 2014)

Propagation:

Woodbury and Sylvester formulas for $\mathbf{M} = \mathbf{A} + \mathbf{UCV}$

$$\mathbf{M}^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{U}(\mathbf{C}^{-1} + \mathbf{VA}^{-1}\mathbf{U})^{-1}\mathbf{VA}^{-1}$$

$$|\mathbf{M}| = |\mathbf{A}|\,|\mathbf{C}|\left|\mathbf{C}^{-1} + \mathbf{VA}^{-1}\mathbf{U}\right|$$

Gives:

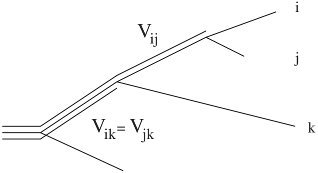
- $\log|\mathbf{V}| = \sum_{s=1}^{k} \log|\mathbf{V}_s| + \log(1 + tp_A)$
- $\mathbf{Q} = \sum_{s=1}^{k} \mathbf{Q}_s - \frac{tp_A^2}{1+tp_A}\hat{\boldsymbol{\mu}}_X^T \hat{\mu}_Y$
- $p = \frac{p_A}{1+tp_A}$, $\hat{\mu}_Y = \sum_{s=1}^{k} \frac{p_s}{p_A}\hat{\mu}_{Y,s}$ and $\hat{\boldsymbol{\mu}}_X^T = \sum_{s=1}^{k} \frac{p_s}{p_A}\hat{\boldsymbol{\mu}}_{X,s}$

Finalization : At the root of the tree.

Three Point Structure (Ho and Ané, 2014)

Three Points Structure: A matrix $\mathbf{V}$ has the 3-Point Structure if:

- it is symmetric non-negative
- for any i, j, k, the two smallest of V_{ij}, V_{ik} and V_{jk} are equal.



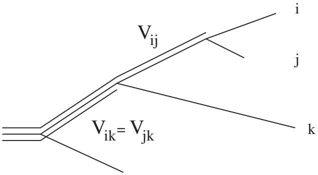
Ho and Ané (2014)

back

Three Point Structure (Ho and Ané, 2014)

Three Points Structure: A matrix $\mathbf{V}$ has the 3-Point Structure if:

- it is symmetric non-negative
- for any i, j, k, the two smallest of V_{ij}, V_{ik} and V_{jk} are equal.



Ho and Ané (2014)

Theorem: $\mathbf{V}$ is 3-point structured if and only if it is the covariance matrix of a random variable at the tips of some rooted tree under a BM model.

back