Hypothesis testing with a phylogeny

The challenge of accounting for phylogenetic non-independence

Guillaume Louvel

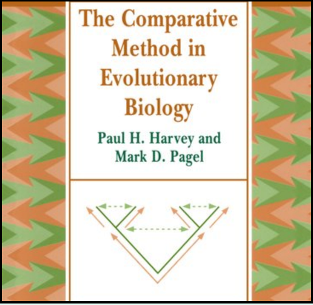
march 22, 2018

In the litterature: the **P**hylogenetic **C**omparative **M**ethod.

László Zsolt Garamszegi
Editor

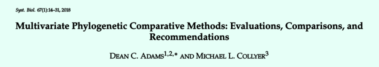
Modern Phylogenetic Comparative Methods and Their Application in Evolutionary Biology

Concepts and Practice



The Comparative Method in Evolutionary Biology

Paul H. Harvey and Mark D. Pagel



Syst. Biol. 67(1):14–31, 2018

Multivariate Phylogenetic Comparative Methods: Evaluations, Comparisons, and Recommendations

Dean C. Adams[1,2,*] and Michael L. Collyer[3]

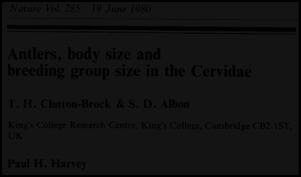
Test *evolutionary* hypotheses using observations from a set of species.

Concrete questions

2 morphological traits

- Does one appear only when the other one exists?
- Are their value correlated? (magnitude and direction of change given the other)
- What is their *rate* of evolution?

example publications



Nature Vol. 285 19 June 1980

Antlers, body size and breeding group size in the Cervidae

T. H. Clutton-Brock & S. D. Albon

King's College Research Centre, King's College, Cambridge CB2 1ST, UK

Paul H. Harvey

- Maddison (1990) (discrete binary traits)
- Social Brain Hypothesis (continuous traits) (DeCasien et al. 2017)

What's the problem with a standard linear regression?

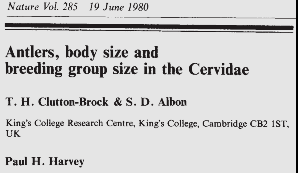
Disclaimer: strategy applicable for any *structured* data (spatial, etc).

Concrete questions

2 morphological traits

- Does one appear only when the other one exists?
- Are their value correlated? (magnitude and direction of change given the other)
- What is their *rate* of evolution?

example publications



Nature Vol. 285 19 June 1980

Antlers, body size and breeding group size in the Cervidae

T. H. Clutton-Brock & S. D. Albon

King's College Research Centre, King's College, Cambridge CB2 1ST, UK

Paul H. Harvey

- Maddison (1990) (discrete binary traits)
- Social Brain Hypothesis (continuous traits) (DeCasien et al. 2017)

What's the problem with a standard linear regression?

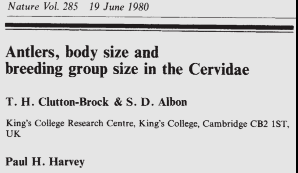
Disclaimer: strategy applicable for any *structured* data (spatial, etc).

Concrete questions

2 morphological traits

- Does one appear only when the other one exists?
- Are their value correlated? (magnitude and direction of change given the other)
- What is their *rate* of evolution?

example publications



Nature Vol. 285 19 June 1980

Antlers, body size and breeding group size in the Cervidae

T. H. Clutton-Brock & S. D. Albon

King's College Research Centre, King's College, Cambridge CB2 1ST, UK

Paul H. Harvey

- Maddison (1990) (discrete binary traits)
- Social Brain Hypothesis (continuous traits) (DeCasien et al. 2017)

What's the problem with a standard linear regression?

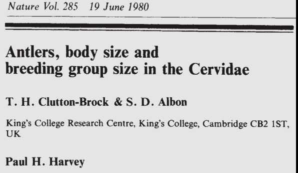
Disclaimer: strategy applicable for any *structured* data (spatial, etc).

Concrete questions

2 morphological traits

- Does one appear only when the other one exists?
- Are their value correlated? (magnitude and direction of change given the other)
- What is their *rate* of evolution?

example publications



Nature Vol. 285 19 June 1980

Antlers, body size and breeding group size in the Cervidae

T. H. Clutton-Brock & S. D. Albon

King's College Research Centre, King's College, Cambridge CB2 1ST, UK

Paul H. Harvey

- Maddison (1990) (discrete binary traits)
- Social Brain Hypothesis (continuous traits) (DeCasien et al. 2017)

What's the problem with a standard linear regression?

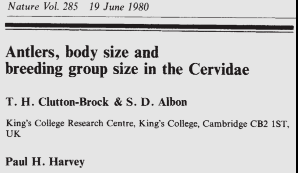
Disclaimer: strategy applicable for any *structured* data (spatial, etc).

Concrete questions

2 morphological traits

- Does one appear only when the other one exists?
- Are their value correlated? (magnitude and direction of change given the other)
- What is their *rate* of evolution?

example publications



Nature Vol. 285 19 June 1980

Antlers, body size and breeding group size in the Cervidae

T. H. Clutton-Brock & S. D. Albon

King's College Research Centre, King's College, Cambridge CB2 1ST, UK

Paul H. Harvey

- Maddison (1990) (discrete binary traits)
- Social Brain Hypothesis (continuous traits) (DeCasien et al. 2017)

What's the problem with a standard linear regression?

Disclaimer: strategy applicable for any *structured* data (spatial, etc).

Comparative data ≠ controled experiment

What explains morphological diversity? Phylogenetic inheritance, chance events, adaptation.

Phylogenetic inertia

- phylogenetic niche conservatism;
- phylogenetic time lag;
- different adaptive responses.

Losos 1994:

> *In a comparative analysis, a wrong phylogeny is better than no phylogeny at all.*

Comparative data $\neq$ controled experiment

What explains morphological diversity? Phylogenetic inheritance, chance events, adaptation.

Phylogenetic inertia

- phylogenetic niche conservatism;
- phylogenetic time lag;
- different adaptive responses.

Losos 1994:

> *In a comparative analysis, a wrong phylogeny is better than no phylogeny at all.*

Comparative data $\neq$ controled experiment

What explains morphological diversity? Phylogenetic inheritance, chance events, adaptation.

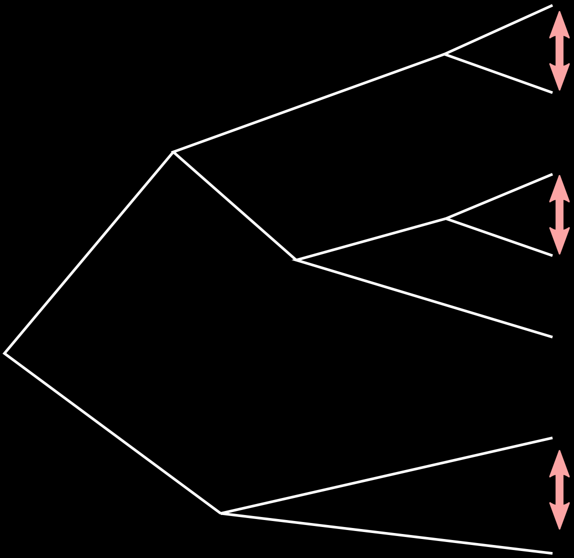
Phylogenetic inertia

- phylogenetic niche conservatism;
- phylogenetic time lag;
- different adaptive responses.

Losos 1994:

> *In a comparative analysis, a wrong phylogeny is better than no phylogeny at all.*

The sister-clade method



Question: how would you normally do?

Fisher exact test / χ^2 test

example contingence table

rows: trait 1
columns: trait 2
cell content: number of *extant* species.

Question: how would you normally do?

Fisher exact test / χ^2 test

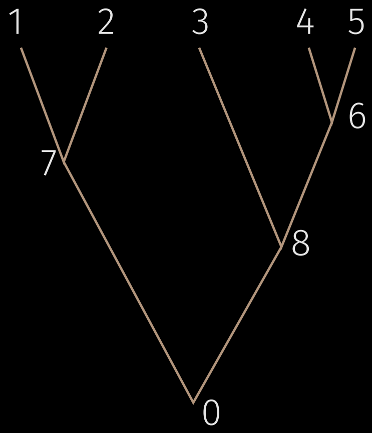
example contingence table

rows: trait 1
columns: trait 2
cell content: number of *extant* species.

- Ridley 1983: consider *branches*: count trait transitions.
- Maddison 1990: adapted to detect directionality of change

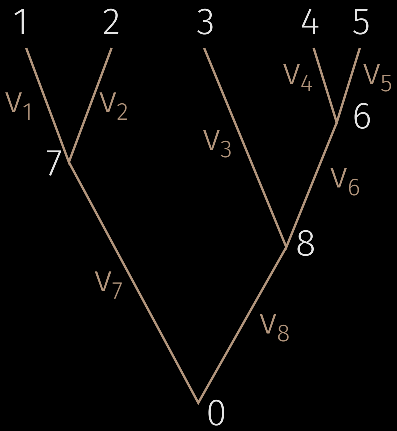
Phylogenetic Independent Contrasts

Felsenstein 1985
“Phylogenies and the Comparative Method”



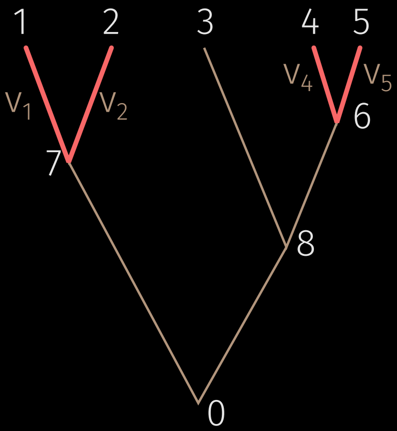
Phylogenetic Independent Contrasts

Felsenstein 1985
“Phylogenies and the Comparative Method”



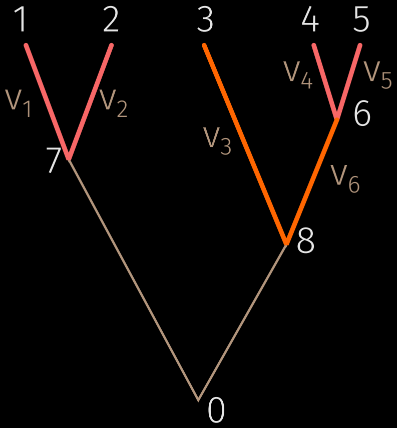
Phylogenetic Independent Contrasts

Felsenstein 1985
“Phylogenies and the Comparative Method”



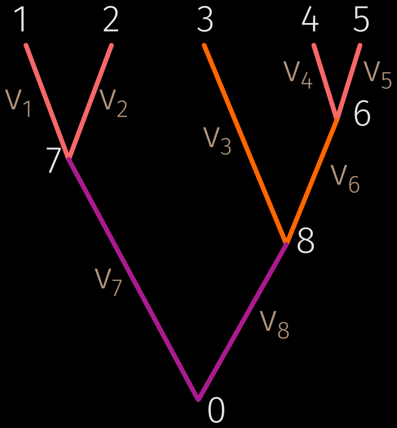
Phylogenetic Independent Contrasts

Felsenstein 1985 “Phylogenies and the Comparative Method”



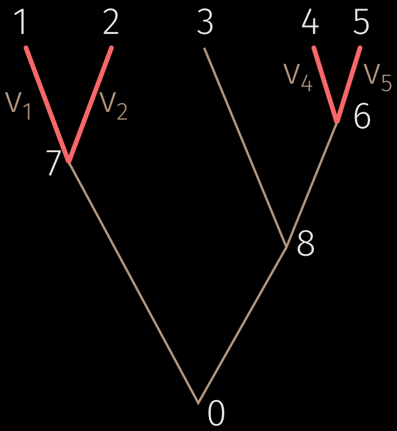
Phylogenetic Independent Contrasts

Felsenstein 1985 "Phylogenies and the Comparative Method"



Phylogenetic Independent Contrasts

Felsenstein 1985
“Phylogenies and the Comparative Method”



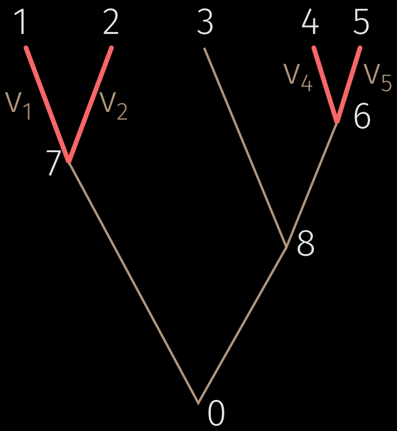
Contrast 1-2

Assuming a brownian motion:

$$(X_1 - X_2) \hookrightarrow \mathcal{N}(0, \sigma^2(v_1 + v_2))$$

Phylogenetic Independent Contrasts

Felsenstein 1985 "Phylogenies and the Comparative Method"



Contrast 1-2

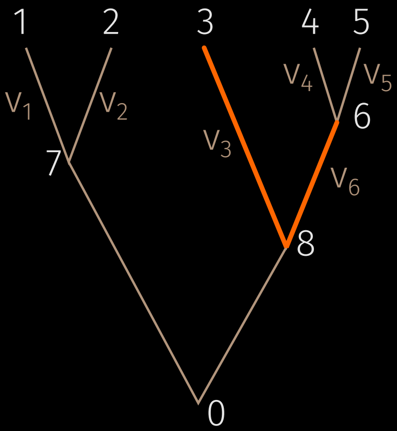
Assuming a brownian motion:

$$(X_1 - X_2) \hookrightarrow \mathcal{N}(0, \sigma^2(v_1 + v_2))$$

$$\frac{(X_1 - X_2)}{\sqrt{\sigma^2(v_1 + v_2)}} \hookrightarrow \mathcal{N}(0, 1)$$

Phylogenetic Independent Contrasts

Felsenstein 1985 "Phylogenies and the Comparative Method"



Contrast 1-2

Assuming a brownian motion:

$$(X_1 - X_2) \hookrightarrow \mathcal{N}(0, \sigma^2(v_1 + v_2))$$

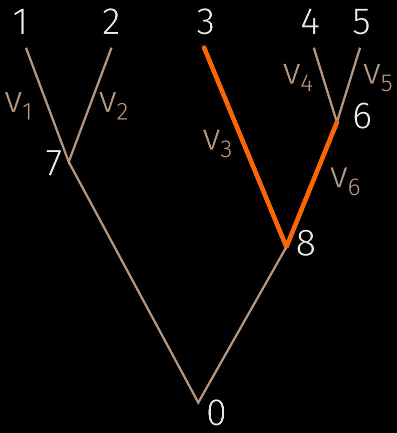
$$\frac{(X_1 - X_2)}{\sqrt{\sigma^2(v_1 + v_2)}} \hookrightarrow \mathcal{N}(0, 1)$$

Contrast 3-6

$$X_3 - X_6$$

Phylogenetic Independent Contrasts

Felsenstein 1985 “Phylogenies and the Comparative Method”



Contrast 1-2

Assuming a brownian motion:

$$(X_1 - X_2) \hookrightarrow \mathcal{N}(0, \sigma^2(v_1 + v_2))$$

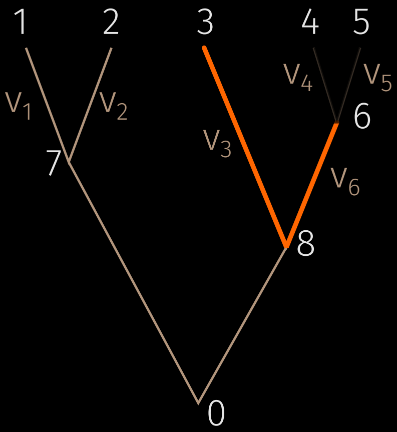
$$\frac{(X_1 - X_2)}{\sqrt{\sigma^2(v_1 + v_2)}} \hookrightarrow \mathcal{N}(0, 1)$$

Contrast 3-6

$$X_3 - ??$$

Phylogenetic Independent Contrasts

Felsenstein 1985 “Phylogenies and the Comparative Method”



Contrast 1-2

Assuming a brownian motion:

$$(X_1 - X_2) \hookrightarrow \mathcal{N}(0, \sigma^2(v_1 + v_2))$$

$$\frac{(X_1 - X_2)}{\sqrt{\sigma^2(v_1 + v_2)}} \hookrightarrow \mathcal{N}(0, 1)$$

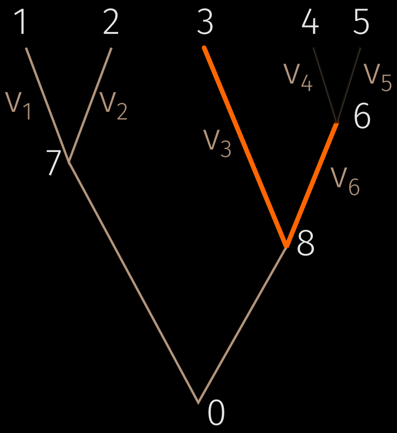
Contrast 3-6

$$X_3 - X_6$$

$$\text{where } X_6 = \frac{X_4/v_4 + X_5/v_5}{1/v_4 + 1/v_5}$$

Phylogenetic Independent Contrasts

Felsenstein 1985 "Phylogenies and the Comparative Method"



Contrast 1-2

Assuming a brownian motion:

$$(X_1 - X_2) \hookrightarrow \mathcal{N}(0, \sigma^2(v_1 + v_2))$$

$$\frac{(X_1 - X_2)}{\sqrt{\sigma^2(v_1 + v_2)}} \hookrightarrow \mathcal{N}(0, 1)$$

Contrast 3-6

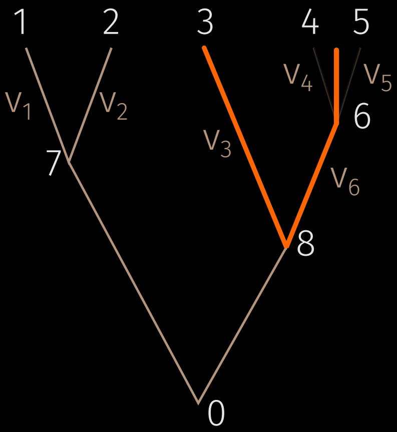
$$X_3 - X_6$$

$$\text{where } X_6 = \frac{X_4/v_4 + X_5/v_5}{1/v_4 + 1/v_5}$$

$$\mathrm{Var}(X_3 - X_6) = v_3 + v_6'$$

Phylogenetic Independent Contrasts

Felsenstein 1985 “Phylogenies and the Comparative Method”



Contrast 1-2

Assuming a brownian motion:

$$(X_1 - X_2) \hookrightarrow \mathcal{N}(0, \sigma^2(v_1 + v_2))$$

$$\frac{(X_1 - X_2)}{\sqrt{\sigma^2(v_1 + v_2)}} \hookrightarrow \mathcal{N}(0, 1)$$

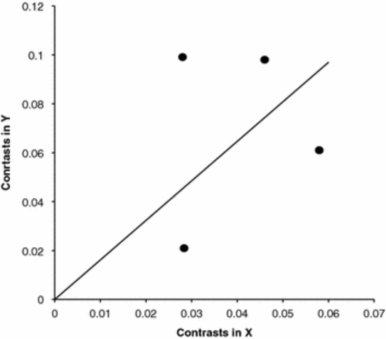
Contrast 3-6

$$X_3 - X_6$$

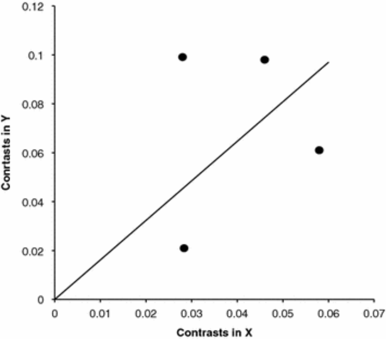
$$\text{where } X_6 = \frac{X_4/v_4 + X_5/v_5}{1/v_4 + 1/v_5}$$

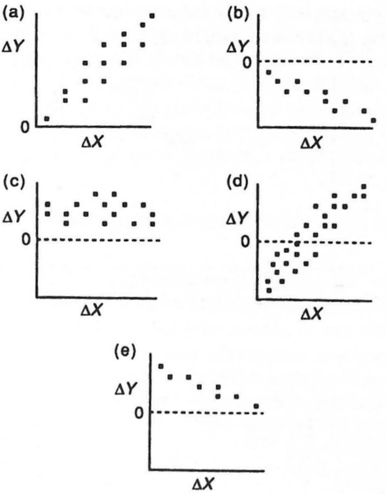
$$\mathrm{Var}(X_3 - X_6) = v_3 + v_6 + \frac{v_4 v_5}{v_4 + v_5}$$

Analysing plots of PICs



Analysing plots of PICs





From Harvey, Pagel, et al. 1991 p.160

In practice

Example

```
library(ape)
# extract data with: geiger::treedata(df, tree)
pic.X <- pic(data$X, tree)
pic.Y <- pic(data$Y, tree)
cor.test(pic.X, pic.Y)
```

Reminder: the Linear Model

2D visualisation

$$y_i = a + bx_i + e_i$$

i experimental unit;

y_i response variable;

x_i explanatory variable;

a, b regression coefficients (model parameters);

e_i residual error (variance not explained by the model).

Reminder: the Linear Model

2D visualisation

$$y_i = a + bx_i + e_i$$

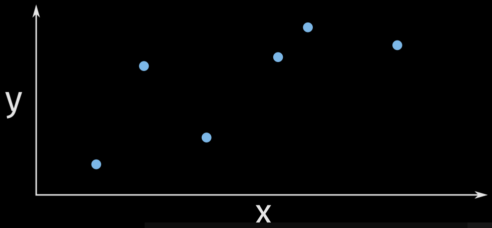
- i experimental unit;
- y_i response variable;
- x_i explanatory variable;
- a, b regression coefficients (model parameters);
- e_i residual error (variance not explained by the model).

Reminder: the Linear Model

2D visualisation

$$y_i = a + bx_i + e_i$$

i	experimental unit;
y_i	response variable;
x_i	explanatory variable;
a, b	regression coefficients (model parameters);
e_i	residual error (variance not explained by the model).



Reminder: the Linear Model

2D visualisation

$$y_i = a + bx_i + e_i$$

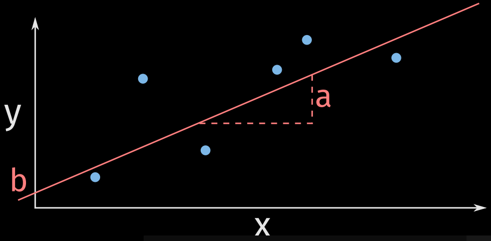
i experimental unit;

y_i response variable;

x_i explanatory variable;

a, b regression coefficients (model parameters);

e_i residual error (variance not explained by the model).

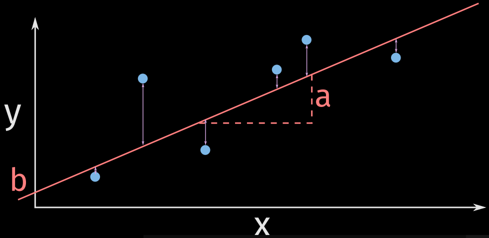


Reminder: the Linear Model

2D visualisation

$$y_i = a + bx_i + e_i$$

- i experimental unit;
- y_i response variable;
- x_i explanatory variable;
- a, b regression coefficients (model parameters);
- e_i residual error (variance not explained by the model).

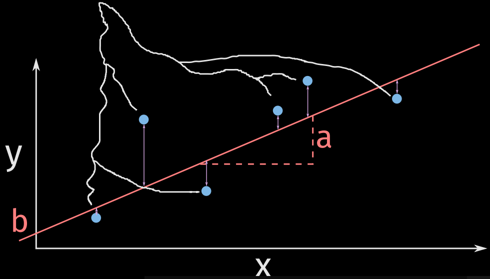


Reminder: the Linear Model

2D visualisation

$$y_i = a + bx_i + e_i$$

- i experimental unit;
- y_i response variable;
- x_i explanatory variable;
- a, b regression coefficients (model parameters);
- e_i residual error (variance not explained by the model).



Reminder: the Linear Model

With multiple explanatory variables:

$$y_i = \theta_0 + \theta_1 x_{i1} + \dots + \theta_2 x_{ip} + \dots \qquad + \quad e_i$$

$$Y = X\theta \qquad + \quad E$$

$$\begin{bmatrix}\dots\\ y_i\\ \dots\end{bmatrix} = \text{exp units} \updownarrow \overset{\text{explanatory vars}}{\overleftrightarrow{\begin{bmatrix}\dots & & \\ & x_{ip} & \\ & & \dots\end{bmatrix}}} \begin{bmatrix}\dots\\ \theta_p\\ \dots\end{bmatrix} \quad + \begin{bmatrix}\dots\\ e_i\\ \dots\end{bmatrix}$$

Linear Model Assumptions

- linearity : $\mathbb{E}(Y) = X\theta$
- e_i are:
  - *independent* and *identically* distributed (homoscedasticity) .
  - normally distributed $\mathcal{N}(0, \sigma^2)$

Reminder: the Linear Model

With multiple explanatory variables:

$$\begin{aligned} y_i &= \theta_0 + \theta_1 x_{i1} + \ldots + \theta_2 x_{ip} + \ldots &&+ \quad e_i \\ Y &= X\theta &&+ \quad E \end{aligned}$$

$$\begin{bmatrix} \cdots \\ y_i \\ \cdots \end{bmatrix} = \text{exp units} \updownarrow \overset{\text{explanatory vars}}{\overleftrightarrow{\begin{bmatrix} \cdots & & \\ & x_{ip} & \\ & & \cdots \end{bmatrix}}} \begin{bmatrix} \cdots \\ \theta_p \\ \cdots \end{bmatrix} + \begin{bmatrix} \cdots \\ e_i \\ \cdots \end{bmatrix}$$

Linear Model Assumptions

- **linearity** : $\mathbb{E}(Y) = X\theta$
- e_i **are:**
  - *independent* and *identically* distributed (homoscedasticity) .
  - normally distributed $\mathcal{N}(0, \sigma^2)$

Reminder: the Linear Model

With multiple explanatory variables:

$$
\begin{aligned}
y_i &= \theta_0 + \theta_1 x_{i1} + \ldots + \theta_2 x_{ip} + \ldots && + \quad e_i \\
Y &= X\theta && + \quad E
\end{aligned}
$$

$$
\begin{bmatrix} \ldots \\ y_i \\ \ldots \end{bmatrix} = \text{exp units} \updownarrow \overset{\overset{\text{explanatory vars}}{\longleftrightarrow}}{\begin{bmatrix} \ldots & & \\ & x_{ip} & \\ & & \ldots \end{bmatrix}} \begin{bmatrix} \ldots \\ \theta_p \\ \ldots \end{bmatrix} + \begin{bmatrix} \ldots \\ e_i \\ \ldots \end{bmatrix}
$$

Linear Model Assumptions

- linearity : $\mathbb{E}(Y) = X\theta$
- e_i are:
  - *independent* and *identically* distributed (homoscedasticity) .
  - normally distributed $\mathcal{N}(0, \sigma^2)$

Reminder: the Linear Model

With multiple explanatory variables:

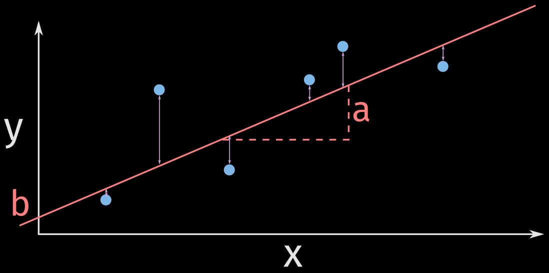
$$\begin{aligned} y_i &= \theta_0 + \theta_1 x_{i1} + \ldots + \theta_2 x_{ip} + \ldots & + \quad & e_i \\ Y &= X\theta & + \quad & E \end{aligned}$$

$$\begin{bmatrix}\ldots\\ y_i\\ \ldots\end{bmatrix} = \text{exp units}\updownarrow \overset{\text{explanatory vars}}{\overleftrightarrow{\begin{bmatrix}\ldots & & \\ & x_{ip} & \\ & & \ldots\end{bmatrix}}} \begin{bmatrix}\ldots\\ \theta_p\\ \ldots\end{bmatrix} + \begin{bmatrix}\ldots\\ e_i\\ \ldots\end{bmatrix}$$

Linear Model Assumptions

- **linearity :** $\mathbb{E}(Y) = X\theta$
- e_i **are:**
  - ***independent* and *identically* distributed (homoscedasticity) .**
  - **normally distributed** $\mathcal{N}(0, \sigma^2)$

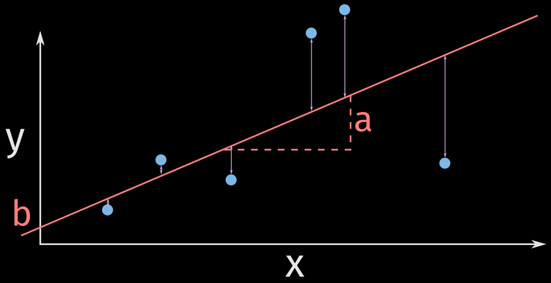
Reminder: the Linear Model



Linear Model Assumptions

- linearity : $\mathbb{E}(Y) = X\theta$
- e_i are:
  - *independent* and *identically* distributed (homoscedasticity) .
  - normally distributed $\mathcal{N}(0, \sigma^2)$

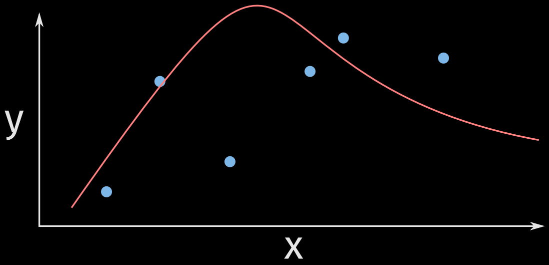
Reminder: the Linear Model



General Linear Model Assumptions

- linearity : $\mathbb{E}(Y) = X\theta$
- e_i are:
  - *independent* ~~and *identically* distributed (homoscedasticity)~~ (heteroscedacity).
  - normally distributed $\mathcal{N}(0, \lambda_i\sigma^2)$

Reminder: the Linear Model



Generalized Linear Model Assumptions

- linearity + link function: $\mathbb{E}(Y) = g(X\theta)$
- e_i are:
  - *independent* ~~and *identically* distributed (homoscedasticity)~~ (heteroscedacity).
  - normally distributed $\mathcal{N}(0, \lambda_i \sigma^2)$

Fitting a Linear model with Least Squares

Ordinary Least Squares

Minimizing the sum of squared errors:

$$\mathrm{argmin}_{a,b} S(a,b)$$

$$S(a,b) = \sum_{i=1}^{n} e_i^2$$
$$= \sum_{i=1}^{n} (y_i - (a + bx_i)^2)$$

Generalized Least Squares

Residues *covariate*:

so we use a covariance matrix $\mathrm{Cov}(E) = \Omega$

$$\hat{\theta} = (X^t \Omega^{-1} X)^{-1} X^t \Omega^{-1} Y$$

Fitting a Linear model with Least Squares

Ordinary Least Squares

Minimizing the sum of squared errors:

$$\hat{\theta} = \mathrm{argmin}_{\theta} ||Y - X\theta||^2$$

We assume:
$\mathbb{E}(E) = 0$ and $\mathbf{Cov}(E) = \sigma^2 I$

There is a solution:

$$\hat{\theta} = (X^t X)^{-1} X^t Y$$

Generalized Least Squares

Residues *covariate*:

so we use a covariance matrix $\mathrm{Cov}(E) = \Omega$

$$\hat{\theta} = (X^t \Omega^{-1} X)^{-1} X^t \Omega^{-1} Y$$

Fitting a Linear model with Least Squares

Ordinary Least Squares

Minimizing the sum of squared errors:

$$\hat{\theta} = \mathbf{argmin}_\theta ||Y - X\theta||^2$$

We assume:
$\mathbb{E}(E) = 0$ and $\mathbf{Cov}(E) = \sigma^2 I$

There is a solution:

$$\hat{\theta} = (X^t X)^{-1} X^t Y$$

Generalized Least Squares

Residues *covariate*:

so we use a covariance matrix $\mathrm{Cov}(E) = \Omega$

$$\hat{\theta} = (X^t \Omega^{-1} X)^{-1} X^t \Omega^{-1} Y$$

Fitting a Linear model with Least Squares

Ordinary Least Squares

Minimizing the sum of squared errors:

$$\hat{\theta} = \mathrm{argmin}_\theta ||Y - X\theta||^2$$

We assume:
$\mathbb{E}(E) = 0$ and $\mathbf{Cov}(E) = \sigma^2 I$

There is a solution:

$$\hat{\theta} = (X^t X)^{-1} X^t Y$$

Generalized Least Squares

Residues *covariate*:

so we use a covariance matrix $\mathbf{Cov}(E) = \Omega$

$$\hat{\theta} = (X^t \Omega^{-1} X)^{-1} X^t \Omega^{-1} Y$$

Fitting a Linear model with Least Squares

Ordinary Least Squares

Minimizing the sum of squared errors:

$$\hat{\theta} = \mathrm{argmin}_\theta ||Y - X\theta||^2$$

We assume:
$\mathbb{E}(E) = 0$ and $\mathrm{Cov}(E) = \sigma^2 I$

There is a solution:

$$\hat{\theta} = (X^t X)^{-1} X^t Y$$

Generalized Least Squares

Residues *covariate*:

so we use a covariance matrix $\mathrm{Cov}(E) = \Omega$

$$\hat{\theta} = (X^t \Omega^{-1} X)^{-1} X^t \Omega^{-1} Y$$

Some terminology

General Linear Model: t-test, multiple regression, ANOVA, ANCOVA

⚠ **Linearity:** $y = \alpha + \beta x_1 + \gamma x_2 + e$

⚠ **"Generalized":**

Generalized Linear Model Assumptions

- linearity *+ link function*: $\mathbb{E}(Y) = g(X\theta)$
- e_i are:
  - *independent* and *identically* distributed.

Generalized Least Squares Assumptions

- linearity: $\mathbb{E}(Y) = X\theta$
- e_i are:
  - ~~*independent* and *identically* distributed.~~

Some terminology

General Linear Model: t-test, multiple regression, ANOVA, ANCOVA

⚠ Linearity: $y = \alpha + \beta x_1 + \gamma x_2 + e$

⚠ “Generalized”:

Generalized Linear Model Assumptions

- linearity *+ link function*: $\mathbb{E}(Y) = g(X\theta)$
- e_i are:
  - *independent* and *identically* distributed.

Generalized Least Squares Assumptions

- linearity: $\mathbb{E}(Y) = X\theta$
- e_i are:
  - ~~*independent* and *identically* distributed.~~

Some terminology

General Linear Model: t-test, multiple regression, ANOVA, ANCOVA

⚠ Linearity: $y = \alpha + \beta x_1 + \gamma x_2 + \delta x_1^2 + e$

⚠ “Generalized”:

Generalized Linear Model Assumptions

- linearity *+ link function*: $\mathbb{E}(Y) = g(X\theta)$
- e_i are:
  - *independent* and *identically* distributed.

Generalized Least Squares Assumptions

- linearity: $\mathbb{E}(Y) = X\theta$
- e_i are:
  - ~~*independent* and *identically* distributed.~~

Some terminology

General Linear Model: t-test, multiple regression, ANOVA, ANCOVA

⚠ Linearity: $y = \alpha + \beta x_1 + \gamma x_2 + \delta x_1^2 + \zeta x_1 x_2 + e$

⚠ “Generalized”:

Generalized Linear Model Assumptions

- linearity *+ link function*: $\mathbb{E}(Y) = g(X\theta)$
- e_i are:
  - *independent* and *identically* distributed.

Generalized Least Squares Assumptions

- linearity: $\mathbb{E}(Y) = X\theta$
- e_i are:
  - ~~*independent* and *identically* distributed.~~

Some terminology

General Linear Model: t-test, multiple regression, ANOVA, ANCOVA

⚠ Linearity: $y = \alpha + \beta x_1 + \gamma x_2 + \delta x_1^2 + \zeta x_1 x_2 + \eta \log(x_1) + e$

⚠ “Generalized”:

Generalized Linear Model Assumptions

- linearity *+ link function*: $\mathbb{E}(Y) = g(X\theta)$
- e_i are:
  - *independent* and *identically* distributed.

Generalized Least Squares Assumptions

- linearity: $\mathbb{E}(Y) = X\theta$
- e_i are:
  - ~~*independent* and *identically* distributed.~~

Some terminology

General Linear Model: t-test, multiple regression, ANOVA, ANCOVA

⚠ Linearity: $y = \alpha + \beta x_1 + \gamma x_2 + \delta x_1^2 + \zeta x_1 x_2 + \eta \log(x_1) + e$

⚠ “Generalized”:

Generalized Linear Model Assumptions

- linearity *+ link function*: $\mathbb{E}(Y) = g(X\theta)$
- e_i are:
  - *independent* and *identically* distributed.

Generalized Least Squares Assumptions

- linearity: $\mathbb{E}(Y) = X\theta$
- e_i are:
  - ~~*independent* and *identically* distributed.~~

Some terminology

General Linear Model: t-test, multiple regression, ANOVA, ANCOVA

⚠ Linearity: $y = \alpha + \beta x_1 + \gamma x_2 + \delta x_1^2 + \zeta x_1 x_2 + \eta \log(x_1) + e$

⚠ “Generalized”:

Generalized Linear Model Assumptions

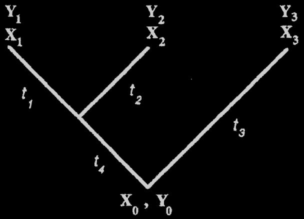
- linearity *+ link function*: $\mathbb{E}(Y) = g(X\theta)$
- e_i are:
  - *independent* and *identically* distributed.

Generalized Least Squares Assumptions

- linearity: $\mathbb{E}(Y) = X\theta$
- e_i are:
  - ~~*independent* and *identically* distributed.~~

Phylogenetic Generalized Least Squares

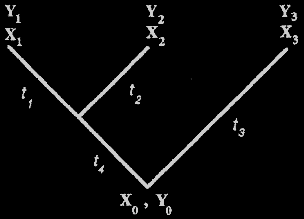
Grafen 1989
Getting the covariance matrix from the phylogeny:



Phylogenetic Generalized Least Squares

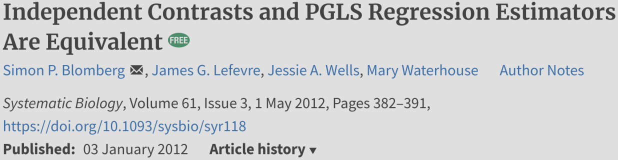
Grafen 1989

Getting the covariance matrix from the phylogeny:



$$\begin{bmatrix} \sigma^2(t_1 + t_4) & \sigma^2 t_4 & 0 \\ \sigma^2 t_4 & \sigma^2(t_2 + t_4) & 0 \\ 0 & 0 & \sigma^2 t_3 \end{bmatrix}$$

Phylogenetic Independant Constrasts VS Phylogenetic Generalized Least Squares



Independent Contrasts and PGLS Regression Estimators Are Equivalent FREE

Simon P. Blomberg ✉, James G. Lefevre, Jessie A. Wells, Mary Waterhouse Author Notes

Systematic Biology, Volume 61, Issue 3, 1 May 2012, Pages 382–391,
https://doi.org/10.1093/sysbio/syr118

Published: 03 January 2012 **Article history** ▾

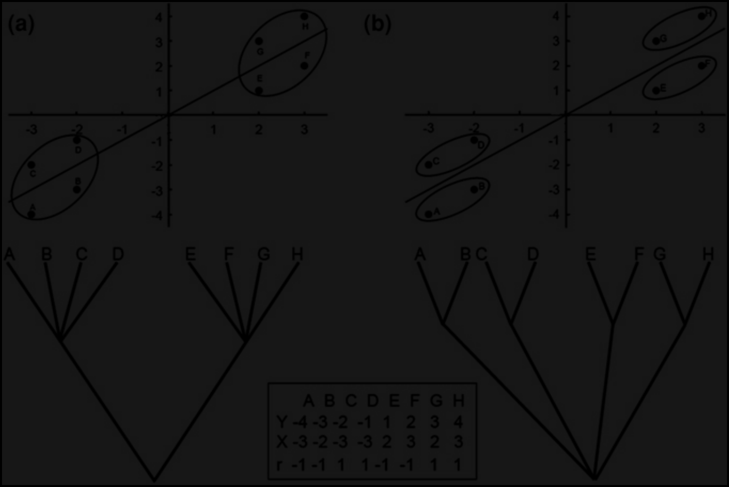
- PIC do not provide the intercept directly;
- Both can be generalized to other evolution processes (Ornstein-Uhlenbeck);
- PGLS easier with partially unresolved phylogenies.

The problem with phylogenetic relatedness
The sister-clade method
2 binary traits
Continuous traits
PIC
LM vs GLM
OLS vs GLS
PGLS
PIC vs PGLS
Should you correct for phylogenetic structure?
PGLS example in R
Summary
Bibliography
References

Should you correct for phylogenetic structure?

Yes if there is *phylogenetic signal* in the *residuals*.

⚠ in residuals $\neq$ in the response variable !

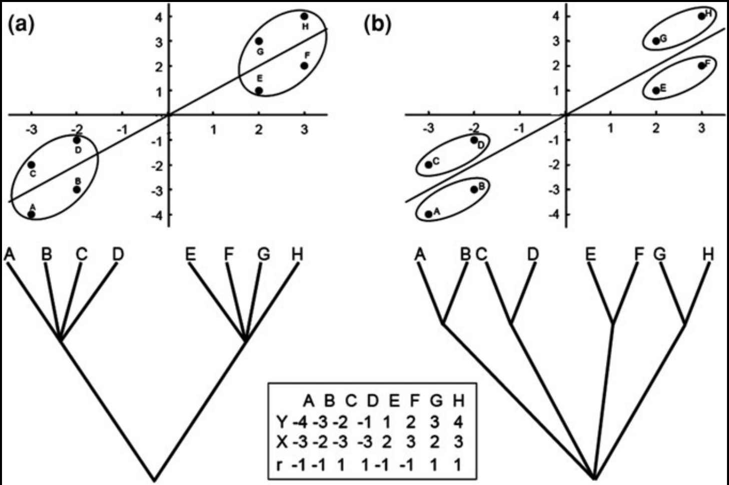


	A	B	C	D	E	F	G	H
Y	-4	-3	-2	-1	1	2	3	4
X	-3	-2	-3	-3	2	3	2	3
r	-1	-1	1	1	-1	-1	1	1

From Labra et al 2009

Should you correct for phylogenetic structure?

Yes if there is *phylogenetic signal* in the *residuals*.
⚠ in residuals $\neq$ in the response variable !



	A	B	C	D	E	F	G	H
Y	-4	-3	-2	-1	1	2	3	4
X	-3	-2	-3	-3	2	3	2	3
r	-1	-1	1	1	-1	-1	1	1

From Labra et al 2009

Brief tutorial in R

Example from http://www.mpcm-evolution.org/practice.

Some packages

```r
library(ape)  # tree handling
library(nlme) # regression modelling
# or
library(caper) # pgls() function
```

data

```r
library(ade4); data(lizards)
tree <- read.tree(text = lizards$hprA)
dat <- lizards$traits[tree$tip.label, ] # sort data
    according to tree
plot(tree, main = "Phylogeny for 18 Lizard Species",
    direction = "up", srt = -90, label.offset = 1)
```

Brief tutorial in R

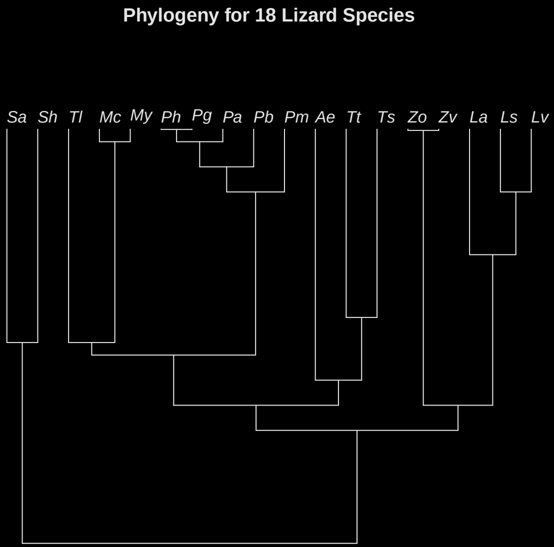
Example from http://www.mpcm-evolution.org/practice.

Some packages

```r
library(ape)  # tree handling
library(nlme) # regression modelling
# or
library(caper) # pgls() function
```

data

```r
library(ade4); data(lizards)
tree <- read.tree(text = lizards$hprA)
dat <- lizards$traits[tree$tip.label, ] # sort data
    according to tree
plot(tree, main = "Phylogeny for 18 Lizard Species",
    direction = "up", srt = -90, label.offset = 1)
```


Phylogeny for 18 Lizard Species
Sa
Sh
Tl
Mc
My
Ph
Pg
Pa
Pb
Pm
Ae
Tt
Ts
Zo
Zv
La
Ls
Lv

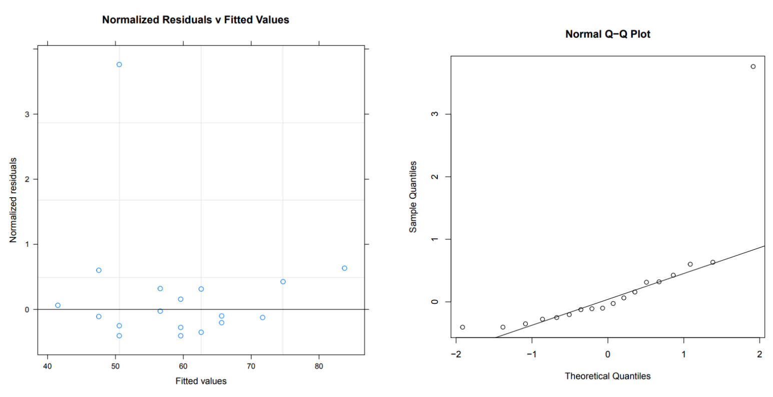
```
       mean.L matur.L max.L hatch.L hatch.m clutch.S age.mat clutch.F
Sa     69.2   58      82    27.8    0.572   6.0      13      1.5
Sh     48.4   42      56    22.9    0.310   3.2      5       2.0
Tl     168.4  132     190   42.8    2.235   16.9     19      1.0
Mc     66.1   56      72    25.0    0.441   7.2      11      1.5
My     70.1   60      81    26.6    0.550   5.4      10      1.0
Ph     49.6   39      57    23.8    0.310   2.1      8       2.0
```

```
fit <- gls(matur.L ~ age.mat, correlation=corBrownian
    (tree), data=dat)

# Custom correlation matrix:
mymat <- vcv(tree, corr=TRUE) # construct correlation
     matrix
corSymm(mymat[lower.tri(mat)],fixed=TRUE)
```

```
        mean.L  matur.L  max.L  hatch.L  hatch.m  clutch.S  age.mat  clutch.F
Sa      69.2    58       82     27.8     0.572    6.0       13       1.5
Sh      48.4    42       56     22.9     0.310    3.2       5        2.0
Tl      168.4   132      190    42.8     2.235    16.9      19       1.0
Mc      66.1    56       72     25.0     0.441    7.2       11       1.5
My      70.1    60       81     26.6     0.550    5.4       10       1.0
Ph      49.6    39       57     23.8     0.310    2.1       8        2.0
```

```
fit <- gls(matur.L ~ age.mat, correlation=corBrownian
    (tree), data=dat)

# Custom correlation matrix:
mymat <- vcv(tree, corr=TRUE) # construct correlation
     matrix
corSymm(mymat[lower.tri(mat)],fixed=TRUE)
```

```
plot(fit2, resid(., type="n")~fitted(.), main="
    Normalized Residuals v Fitted Values", abline=c
    (0,0))
res <- resid(fit2, type="n")
qqnorm(res)
qqline(res)
```



phylogenetic signal

```
fitPagel <- gls(matur.L ~ age.mat, correlation=
    corPagel(value=0.8, phy=tree3),data=dat3)
intervals(fitPagel, which="var-cov")
## Approximate 95% confidence intervals
##
## Correlation structure:
## lower est. upper
## lambda 0.49 0.899 1.308
## attr(,"label")
## [1] "Correlation structure:"
##
## Residual standard error:
## lower est. upper
## 11.76 21.88 40.72
```

The problem with phylogenetic relatedness
The sister-clade method
2 binary traits
Continuous traits
Should you correct for phylogenetic structure?
PGLS example in R
Summary
Bibliography
References

```
fitPagel0 <- gls(matur.L ~ age.mat, correlation =
    corPagel(value = 0, phy = tree3, fixed = TRUE),
    data = dat3) # independence
fitPagel1 <- gls(matur.L ~ age.mat, correlation =
    corPagel(value = 1, phy = tree3, fixed = TRUE),
    data = dat3) # Brownian motion

anova(fitPagel, fitPagel0)
## Model df AIC BIC logLik Test L.Ratio p-value
## fitPagel 1 4 140.2 143.0 -66.08
## fitPagel0 2 3 137.8 139.9 -65.91 1 vs 2 0.3439
    0.5576
```

References to get started

Felsenstein 1985 “Phylogenies and the Comparative Method”

Garamszegi 2014 *Modern Phylogenetic Comparative Methods and Their Application in Evolutionary Biology*

www.mpcm-evolution.org/practice

Harvey, Pagel, et al. 1991 *The comparative method in evolutionary biology*

That's not all!

- phylogenetic signal (Münkemüller et al. 2012)
- phylogenetic ANOVA, ANCOVA, multivariate analysis...
- inferring causality
- The seven deadly sins of comparative analysis (Freckleton 2009)

Bibliography I

DeCasien, Alex R. et al. (2017). "Primate brain size is predicted by diet but not sociality". In: *Nature Ecology & Evolution* 1, p. 0112. DOI: 10.1038/s41559-017-0112.

Felsenstein, Joseph (1985). "Phylogenies and the Comparative Method". In: *The American Naturalist* 125.1, pp. 1–15. DOI: 10.2307/2461605.

Freckleton, R. P. (2009). "The seven deadly sins of comparative analysis". In: *Journal of Evolutionary Biology* 22.7, pp. 1367–1375. DOI: 10.1111/j.1420-9101.2009.01757.x.

Garamszegi, László Zsolt, ed. (2014). *Modern Phylogenetic Comparative Methods and Their Application in Evolutionary Biology*. Berlin, Heidelberg: Springer Berlin Heidelberg. DOI: 10.1007/978-3-662-43550-2.

Bibliography II

Grafen, A. (1989). “The Phylogenetic Regression”. In: *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 326.1233.

Harvey, Paul H, Mark D Pagel, et al. (1991). *The comparative method in evolutionary biology*. Oxford University Press, Oxford.

Maddison, Wayne P. (1990). “A Method for Testing the Correlated Evolution of Two Binary Characters: Are Gains or Losses Concentrated on Certain Branches of a Phylogenetic Tree?” In: *Evolution* 44.3, p. 539. DOI: 10.2307/2409434.

Münkemüller, Tamara et al. (2012). “How to measure and test phylogenetic signal”. In: *Methods in Ecology and Evolution* 3.4, pp. 743–756. DOI: 10.1111/j.2041-210X.2012.00196.x.

Bibliography III

Ridley, Mark (1983). *The explanation of organic diversity: the comparative method and adaptations for mating.* Oxford University Press, USA.

Ordinary and Generalized Least Squares

Least Squares Estimator:

$$\hat{\theta} = \mathrm{argmin}_{\theta} \|Y - X\theta\|^2$$

It verifies:

$$X^t X \hat{\theta} = X^t Y$$

So when $X^t X$ is invertible (*i.e.* the matrix is of full rank, there are more data points than explanatory variables, and explanatory variables are independent):

$$\hat{\theta} = (X^t X)^{-1} X^t Y$$

When there is covariance between residuals (Generalized Least Squares):

$$\hat{\theta} = (X^t \Omega^{-1} X)^{-1} \Omega^{-1} X^t Y$$

(corresponds to minimizing the squared Mahalanobis length of the residual vector)