A Primer on Phylogenetic Generalised Least Squares

Chapter · January 2014

DOI: 10.1007/978-3-662-43550-2_5

Simon Phillip Blomberg
The University of Queensland

Chapter 5
A Primer on Phylogenetic Generalised Least Squares

Matthew R. E. Symonds and Simon P. Blomberg

Abstract Phylogenetic generalised least squares (PGLS) is one of the most commonly employed phylogenetic comparative methods. The technique, a modification of generalised least squares, uses knowledge of phylogenetic relationships to produce an estimate of expected covariance in cross-species data. Closely related species are assumed to have more similar traits because of their shared ancestry and hence produce more similar residuals from the least squares regression line. By taking into account the expected covariance structure of these residuals, modified slope and intercept estimates are generated that can account for interspecific autocorrelation due to phylogeny. Here, we provide a basic conceptual background to PGLS, for those unfamiliar with the approach. We describe the requirements for a PGLS analysis and highlight the packages that can be used to implement the method. We show how phylogeny is used to calculate the expected covariance structure in the data and how this is applied to the generalised least squares regression equation. We demonstrate how PGLS can incorporate information about phylogenetic signal, the extent to which closely related species truly are similar, and how it controls for this signal appropriately, thereby negating concerns about unnecessarily 'correcting' for phylogeny. In addition to discussing the appropriate way to present the results of PGLS analyses, we highlight some common misconceptions about the approach and commonly encountered problems with the method. These include misunderstandings about what phylogenetic signal refers to in the context of PGLS (residuals errors, not the traits themselves), and issues associated with unknown or uncertain phylogeny.

M. R. E. Symonds (✉)
Centre for Integrative Ecology, School of Life and Environmental Sciences,
Deakin University, Burwood, VIC, Australia
e-mail: matthew.symonds@deakin.edu.au

S. P. Blomberg
School of Biological Sciences, The University of Queensland, St Lucia, QLD, Australia

L. Z. Garamszegi (ed.), *Modern Phylogenetic Comparative Methods and Their Application in Evolutionary Biology*, DOI: 10.1007/978-3-662-43550-2_5,

5.1 Introduction

5.1.1 The Background to PGLS

The 1980s saw a rise in appreciation of the need to take phylogeny into account when conducting analyses of trait correlations across species (Ridley 1983; Felsenstein 1985; Huey 1987; Harvey and Pagel 1991; for an entertaining overview see Losos 2011). Because of shared evolutionary history, species do not provide independent data points for analysis, thereby violating one of the fundamental assumptions of most statistical tests (Chap. 1). With appreciation of this problem came the impetus to develop statistical methods for analysing comparative data while taking phylogeny into account. Of these, phylogenetic generalised least squares (PGLS) is one of the primary methods employed.

PGLS (also called ‘phylogenetic regression’ or ‘phylogenetic general linear models’) was a method initially formulated by Grafen (1989) and subsequently developed by Martins and Hansen (1997), Pagel (1997, 1999) and Rohlf (2001). Initially, biologists were slow to incorporate phylogenetic comparative methods in their research, perhaps because methodological papers plunge quickly into mathematical formulae and statistical terminology. This chapter is intended for those without a strong statistical background as an introduction to PGLS. We explain how PGLS incorporates information about phylogeny and the strength of the phylogenetic signal: the extent to which closely related species resemble each other. We will provide advice on how to conduct analyses, and present results, and also point out areas where those new to the methods might get stuck.

5.1.2 What Kind of Analyses are PGLS Used for?

The most common type of analyses where PGLS are employed are those which seek to establish the nature of the evolutionary association between two or more biological traits—for example, the relationship between body mass and life span (Promislow and Harvey 1990). By ‘evolutionary association’, we mean evidence that traits are associated over evolutionary time. Although PGLS is frequently used to examine the association between a pair of traits, it can also handle multiple predictor variables. However, PGLS has a wider range of applications, including ancestral state estimation, assessment of mode of evolution, and identification of directionality of evolution among traits.

Analyses of coevolution among traits typically involve the estimation of regression estimates. For PGLS, the dependent (response) variable is usually a continuous variable. The predictor variable(s) may also be continuous, but PGLS can deal with pseudo-continuous ordinal data and binary discrete data. Multi-state discrete variables with non-ordinal properties (e.g. diet: insectivorous, herbivorous, piscivorous, etc.) can be dealt within a PGLS framework if they are recoded as separate binary characters (e.g. piscivory: no (0) or yes (1)).

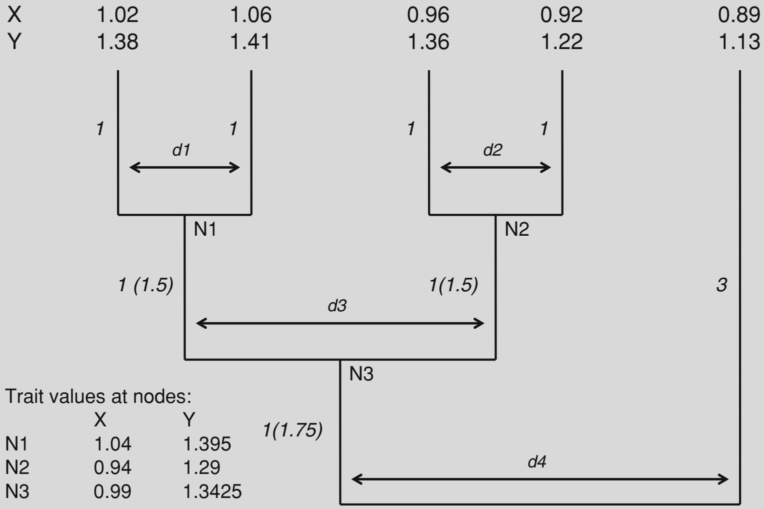
Hypothesis testing with PGLS is not appropriate for analyses with a discrete character as the response variable. Separate methods exist for dealing with discrete response variables including the concentrated changes test (Maddison 1990), pairwise comparisons (Maddison 2000), Pagel's (1994) likelihood method, and phylogenetic logistic regression (Ives and Garland 2010). Chapter 9 reviews some of these approaches.

5.1.3 PGLS and Independent Contrasts

When PGLS was first described by Grafen (1989), he described the method as a generalisation of Felsenstein's (1985) independent contrasts approach. At their heart, the two approaches have the same recognition of the problem of statistical non-independence of species data points as a result of shared ancestry. Independent contrasts resolves this problem by recognising that the differences ('contrasts') between closely related species or clades do provide independent data points for analyses, because they represent the outcome of independent evolutionary pathways (see Box 5.1 for details). PGLS likewise identifies from phylogeny the amount of expected correlation between species based on their shared evolutionary history, and weights for this in the generalised least squares regression calculation. Although couched in slightly different ways, ultimately, the results of PGLS, in their raw form, are the same as those derived from independent contrasts (Grafen 1989; Garland and Ives 2000; Rohlf 2001; Blomberg et al. 2012).

Box 5.1 Independent Contrasts

The most popular method for phylogenetic comparative analysis of continuous data has, until recently, been independent contrasts (Felsenstein 1985). The logic behind this approach is that although raw species data do not provide independent observations for analysis, differences ('contrasts') between closely related species or clades are indeed independent, because they represent the outcome of independent evolutionary pathways. By regressing the independent contrasts of one variable against the independent contrasts of another, one can estimate a regression coefficient that accounts for phylogenetic relatedness among species. Contrasts between species (or clades) are calculated downwards through the tree, with the independent variable (X) typically assigned a positive value. For the tree we discuss in this chapter (see Fig. 5.2) with 5 species, 4 independent contrasts are produced (denoted as $d1$, $d2$, $d3$, and $d4$ below).



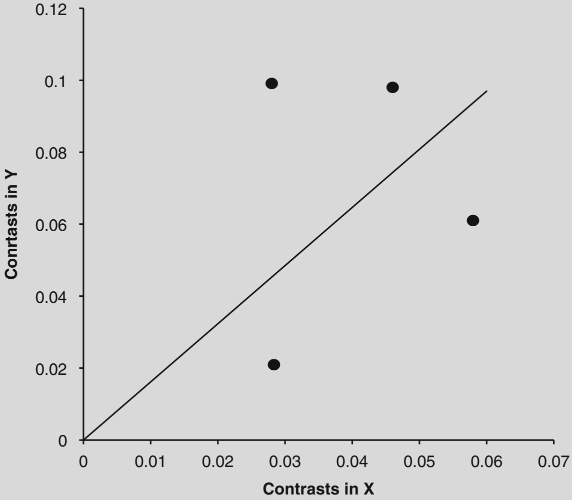
Trait values at nodes:

	X	Y
N1	1.04	1.395
N2	0.94	1.29
N3	0.99	1.3425

For *d*1 and *d*2, the calculation of the raw contrast values is relatively straightforward (it is just the difference between the species trait values). Calculation of the contrast values for *d*3 requires estimation of trait values for the nodes *N*1 and *N*2. These can be estimated as the means of the daughter species weighted by the daughter branch lengths to reflect amount of time over which divergence has occurred (in our example, the daughter branch lengths are the same length, so the weighted means are the same as the raw means). In order to reflect uncertainty with these estimates, the branch lengths leading to these ancestral nodes are modified by lengthening them by an amount equal to (daughter branch length 1 × daughter branch length 2)/(daughter branch length 1 + daughter branch length 2). These modified branch lengths are shown in the brackets after the raw branch lengths on the figure. The trait values for node *N*3 can likewise be estimated as the phylogenetically weighted mean of the estimated trait values at nodes *N*1 and *N*2, and the raw contrast *d*4 subsequently calculated. As before, the branch length between the base of tree and *N*3 must be lengthened using the formula above, using the (modified) daughter branch lengths. While the raw contrast values are now statistically independent, they do not conform to another statistical requirement of having been drawn from a normal distribution with the same expected variance. Hence, they must be standardised by dividing by their standard deviation: the square root of the sum of the branch lengths leading to the two taxa in the contrast (remembering to use the modified branch lengths for internal branches in the tree). For our example, the four contrasts can now be calculated:

Contrast	Raw contrasts		Standard deviation	Standardised contrasts	
	X	Y		X	Y
$d1$	0.04	0.03	$\sqrt{(1+1)} = \sqrt{2}$	0.028	0.021
$d2$	0.04	0.14	$\sqrt{(1+1)} = \sqrt{2}$	0.028	0.099
$d3$	0.1	0.105	$\sqrt{(1.5+1.5)} = \sqrt{3}$	0.058	0.061
$d4$	0.1	0.2125	$\sqrt{(3+1.75)} = \sqrt{4.75}$	0.046	0.098

These standardised contrasts can now be plotted in a normal bivariate scatterplot.



Note that for the independent contrasts, the regression line must be forced through the origin (i.e. have a zero intercept) (Garland et al. 1992). To understand why, consider that for species A, the predicted value of Y (Y_{A}) is

$$Y_{\mathrm{A}} = b_0 + b_1 X_{\mathrm{A}}$$

where b_0 is the intercept and b_1 is the slope value. Likewise, for species B

$$Y_{\mathrm{B}} = b_0 + b_1 X_{\mathrm{B}}$$

For the contrast $Y_{\mathrm{A}} - Y_{\mathrm{B}}$, therefore,

$$Y_{\mathrm{A}} - Y_{\mathrm{B}} = (b_0 + b_1 X_{\mathrm{A}}) - (b_0 + b_1 X_{\mathrm{B}}) = b_0 + b_1 X_{\mathrm{A}} - b_0 - b_1 X_{\mathrm{B}}$$

Notice that the intercept b_0 terms cancel out in this equation and therefore are removed from the calculation of the regression of the contrasts:

$$Y_A - Y_B = b_1 X_A - b_1 X_B = b_1(X_A - X_B)$$

where $X_A - X_B$ is the contrast in X. For our example, the regression coefficient for the standardised contrasts of Y on X is 1.616.

In practice, however, most statistical packages for PGLS have an advantage over those that employ independent contrasts, because they do not automatically rely on the assumption that closely related species will necessarily be similar because of their shared phylogenetic history. In their most basic formulation, both methods assume that continuous traits evolve according to a random walk process, i.e. Brownian motion, such that the change in the value of a trait over a given period of time is given by a random number drawn from a normal distribution with a given standard deviation and mean of 0 (i.e. the value is equally likely to go up or down). Under this model, species that share a more recent common ancestor should have more similar trait values than more distantly related species because their traits have had less time to diverge (see Fig. 5.1).

However, there are many situations in which traits are evolutionarily labile, where closely related species are not necessarily more similar (Blomberg et al. 2003). Criticism that phylogenetic comparative methods might 'over-correct' for phylogeny when applied in such circumstances has been levelled for some years (e.g. Westoby et al. 1995; Björklund 1997; Rheindt et al. 2004; see also Chap. 14). In some circumstances, therefore, a traditional non-phylogenetically controlled analysis might be statistically more appropriate, not least if phylogenies are in extreme error (Abouheif 1998; Symonds 2002; Blomberg et al. 2012). Proposed solutions include presenting the results of both non-phylogenetic and phylogenetic analyses, but this does not resolve the issue of which analysis to base inference on, and it is unclear how one should proceed should the analyses produce conflicting results (see Freckleton 2009; and 'Misconceptions, problems, and pitfalls' later). Additionally, this still presents results based on two very contrasting scenarios—one which assumes no phylogenetic effect on the data and the other which assumes a strong effect. In many cases, the true effect of phylogeny is intermediate, in which case, both types of analysis would be invalid.

This problem can be overcome with PGLS, because it allows one to incorporate information on the extent of phylogenetic signal in the data (see 'Incorporating phylogenetic signal into PGLS' later). If there is no phylogenetic signal in the data, then PGLS will return estimates identical to an ordinary least squares regression analysis. If phylogenetic signal is intermediate, then PGLS can correct for phylogeny to the appropriate degree. While independent contrasts can also be adapted to deal with this issue (as in fact Felsenstein explicitly flagged in his original 1985 paper), in practice, the statistical packages which calculate independent contrasts do not automatically do so and therefore assume that the phylogeny does

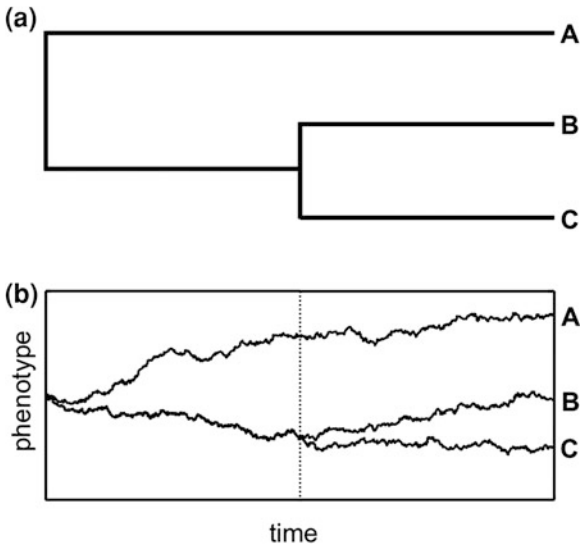


Fig. 5.1 **a** Three-species phylogeny and **b** illustration of possible phenotypic divergence over time (i.e. evolutionary history) in those three species by a Brownian motion model of evolution. Note how the traits gradually diverge such that, typically, species *B* is most similar to *C*. Figure reproduced from Revell et al. (2008) with permission of Liam Revell and Oxford University Press

accurately describe the error structure in the data (i.e. the way species values deviate from least squares regression line—closely related species having similar errors).

PGLS and independent contrasts also present their output in slightly different ways. PGLS calculates an intercept value in the regression equation, whereas independent contrasts force the intercept through the origin (see Box 5.1 and Garland et al. 1992) and the intercept must be subsequently deduced by noting that the line goes through the phylogenetic mean (the estimated ancestral value for the response variable at the root of the phylogeny). Plots of independent contrasts also differ from plots of PGLS (which present the actual species values, rather than contrasts: see 'How to present a PGLS analysis' below). That said, contrast plots can be very informative for detecting outlier clades that are strongly influencing regression estimates.

5.2 Requirements for a PGLS Analysis

The two requirements for a PGLS analyses are a set of comparative species data and a phylogeny for those species. Chapters 2 and 3 provide greater discussion on preparing phylogenies for comparative analysis, but we provide here a quick reminder. The phylogeny may be produced de novo from phylogenetic analysis of DNA sequence data, for example. Alternatively, it may be taken from an already published source and pruned to the relevant species, or augmented as a composite phylogeny using other sources. The phylogeny should ideally include branch lengths and be fully resolved. If not fully resolved, then some determination must be made as to whether the polytomies (when more than two species descend from a node) represent known or unknown phylogeny (i.e. the true evolutionary process—in which case, we call them 'hard' polytomies—or just uncertainty about the true pattern of relationships—'soft' polytomies). We shall discuss later (in Misconceptions, problems, and pitfalls) methods for dealing with polytomies in PGLS.

It may be that no branch length information is available for the phylogeny, in which case, one may either set all branch lengths as equal (Purvis et al. 1994), or use an algorithm such as that used by Grafen (1989) where the depth of each node in the tree is related to the number of daughter species derived from that node (see also Pagel 1992, for an alternative approach). Once compiled, the phylogeny should be formatted so that it can be read by the computer package being used for analysis. Typically, this will be a Nexus file with the stored tree presented in that file in Newick format (Maddison et al. 1997). Trees can be saved in this format by most phylogenetic analysis and tree manipulation packages.

There are several computing packages that perform PGLS: COMPARE (Martins 2004) is an online interface that will conduct PGLS and other functions including independent contrasts, but users should note that COMPARE is no longer being supported or updated. BayesTraits (Pagel and Meade 2013) implements PGLS through its package Continuous (Pagel 1997; Pagel 1999). Finally, several packages within the R statistical framework can derive PGLS estimations very quickly and efficiently, including *ape* (Paradis et al. 2004), *picante* (Kembel et al. 2010), *caper* (Orme et al. 2012), *phytools* (Revell 2012), *nlme* (Pinheiro et al. 2013), and *phyreg* (Grafen 2014).

5.3 Calculation of PGLS

5.3.1 Calculation of Parameter Estimates

The simplest way to think of PGLS is as a weighted regression. In a standard regression, each independent data point contributes equally to the estimation of the regression line. By contrast, PGLS 'downweights' points that derive from species with shared phylogenetic history. These PGLS calculations are automatically done using the appropriate statistical package (see above). Nevertheless, some knowledge of the basic approach involved in this statistical method may be informative.

In an ordinary least squares (OLS) regression model, the relationship of a response variable Y to a predictor variable X_1 can be given using the regression equation:

$$Y = b_0 + b_1 X_1 + \varepsilon \tag{5.1}$$

where b_0 is the intercept value of the regression equation, b_1 is the parameter estimate (the slope value) for the predictor, and ε is the residual error (i.e. for a given point, how far it falls off the regression line). Of course, there may also be other predictor variables in the model—X_2, X_3, etc., with associated regression slope estimates (b_2, b_3, etc.), but for simplicity, we shall focus on the simplest version of linear regression.

To illustrate our discussion, we use a simple example (Fig. 5.2). Fiddler crabs of the genus *Uca* are well known for their enlarged claws, which are used in competition between males for access to females (Crane 1975). As a sexually selected trait, we might expect these claws to show positive allometry (i.e. the parameter estimate b_1 of the regression of log(claw size) on log(body size) should be greater than 1; see Rosenberg (2002) for discussion of fiddler crab claw allometry, and Bonduriansky (2007) for explanation and analysis of the idea more generally). To test this idea, we collated data on body size (carapace breadth) and claw size (propodus length) for five species from Crane (1975). We also obtained a phylogenetic topology for the group (Rosenberg 2001).

For a simple regression with one predictor (X), the slope of the regression line *b1* is given by

$$b_1 = \frac{\sum_{i=1}^{n} (X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^{n} (X_i - \bar{X})(X_i - \bar{X})} \tag{5.2}$$

where n is the sample size, X_i is the ith value of X (up to the last value X_n), and $\bar{X}$ represents the mean value of X (0.97). Likewise for Y_i and $\bar{Y}$ (1.30). The intercept b_0 then simply follows:

$$b_0 = \bar{Y} - b_1 \bar{X} \tag{5.3}$$

For our fiddler crab data, the OLS estimate of the allometric equation is log(propodus length) = −0.229 + 1.577 × log(carapace breadth), with the b_1 term appearing to support the idea of positive allometry in claw length. The parameter estimates b_0, b_1, b_2, and so on (collectively denoted as the vector $\boldsymbol{\beta}$) are the values which minimise the residual variation from the least squares regression line.

For generalised least squares, we need to consider an additional element of the regression equation, in the form of the variance–covariance matrix, which represents the expected covariance structure of the residuals from the regression equation (see Appendix A for a more technical description of the mathematical formulation involved). In the case of OLS, the implicit assumption is that there is

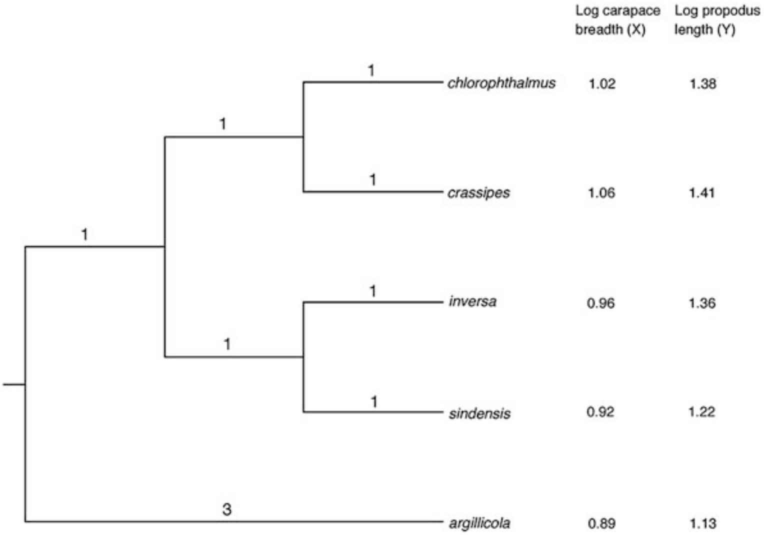


Fig. 5.2 Phylogeny of five *Uca* fiddler crab species, with morphometric data. Numbers on the phylogeny represent branch lengths

no covariance between residuals (i.e. all species are independent of each other, and residuals from closely related species are not more similar on average than residuals from distantly related species). This ($n \times n$) variance–covariance matrix is denoted as **C**, and for five species under the assumption of no phylogenetic effects on the residuals, it looks like:

$$\mathbf{C} = \begin{bmatrix} \boldsymbol{\sigma}_\varepsilon^2 & 0 & 0 & 0 & 0 \\ 0 & \boldsymbol{\sigma}_\varepsilon^2 & 0 & 0 & 0 \\ 0 & 0 & \boldsymbol{\sigma}_\varepsilon^2 & 0 & 0 \\ 0 & 0 & 0 & \boldsymbol{\sigma}_\varepsilon^2 & 0 \\ 0 & 0 & 0 & 0 & \boldsymbol{\sigma}_\varepsilon^2 \end{bmatrix}$$

The first row and first column represent values from comparisons with the first species (in our case *Uca chlorophthalmus*, see Fig. 5.2), the second row and column with *Uca crassipes*, and so on. Hence, the diagonal elements (the line of values from top left to bottom right) represent the variance of the residuals, while the other off-diagonal elements equal zero, meaning there is no covariation among the residuals. When this variance–covariance structure is assumed, the results of GLS are the same as those of OLS (the contribution of **C** to the regression calculation essentially drops out).

Recall that the key statistical issue with cross-species analyses is that species data points are non-independent because of their shared phylogenetic history. Consequently, the errors may also be non-independent or autocorrelated (residuals from closely related species may be similar). Hence, there will be covariation in residuals, which we must account for in our variance–covariance matrix, **C**.

Estimation of the expected covariance structure was a key insight by Felsenstein (1973) that Grafen (1989) used in his phylogenetic regression. Like all good insights, it is elegantly simple: the expected covariance will be related to the amount of shared evolutionary history between the species. Hence, the diagonal elements (i.e. the variance elements) of the matrix are the total length of branches from the root of the tree to the tips. This will be the same for each cell if the phylogeny is ultrametric (i.e. all tips are the same distance from the root of the phylogeny), as it is in the case of our example (distance = 3, see Fig. 5.2). The off-diagonal covariance elements represent the total shared branch length of the evolutionary history of the two species being compared. Hence, for *U. chlorophthalmus* and *U. crassipes,* we see that each species has independent (non-shared) branch lengths of 1. Conversely, the two species share 2 branch lengths in their evolutionary history back to the root of the tree. Consequently, the value entered into column 1–row 2 (and column 2–row 1) of the matrix is 2. We can repeat this for all the other species comparisons (e.g. *U. sindensis* and *U. argillicola* do not share any evolutionary history, so their expected covariance is 0) and produce the new expected variance–covariance matrix:

$$\mathbf{C}_{\mathbf{phyl}} = \begin{bmatrix} 3 & 2 & 1 & 1 & 0 \\ 2 & 3 & 1 & 1 & 0 \\ 1 & 1 & 3 & 2 & 0 \\ 1 & 1 & 2 & 3 & 0 \\ 0 & 0 & 0 & 0 & 3 \end{bmatrix}$$

When this new version of **C** is applied to the GLS calculation (see Appendix A), we eventually end with the PGLS solution: log(propodus length) = − 0.276 + 1.616 × log(carapace breadth). Note, as we said earlier, that the regression slope coefficient is the same here as derived from independent contrasts (see Box 5.1). In this case, our final PGLS regression is not so different from the OLS regression, but there can easily be circumstances where this is not the case. We can plot and compare the two regression slopes for our data (Fig. 5.3).

5.3.2 Hypothesis Testing and Goodness of Fit

After calculating the intercept and slopes using GLS, it is common to ask questions about the magnitude of these quantities. In particular, we may be interested in whether the intercept and/or slopes are significantly different from zero. A Wald *t*-test can be conducted for each parameter in the model simply by dividing the parameter estimate by its associated standard error (i.e. the square root of the estimated variance of the parameter) and then comparing the result to a standard

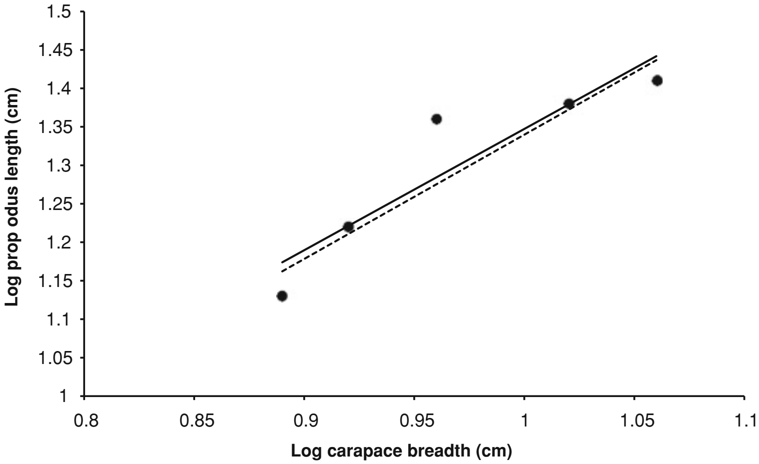


Fig. 5.3 Comparison of OLS (*solid*) and PGLS (*dashed*) regression lines for the fiddler crab claw allometry data

t distribution, using the residual degrees of freedom from the model, to calculate a P value. To test the null hypothesis that $b1 = 0$, the t statistic will therefore be

$$t = \frac{b_1}{\sqrt{\mathrm{Var}(b_1)}} \tag{5.4}$$

Calculation of the degrees of freedom can be non-trivial. In particular, the residual degrees of freedom may need to be reduced if there are soft polytomies in the tree (Purvis and Garland 1993; see also below). F tests for multiple variables can be similarly designed. An alternative test is the likelihood ratio chi-squared test, which has the advantage that it depends only on the likelihood of a general model (which includes the parameter) compared to a restricted model without the parameter of interest. Popular software (such as *nlme* for R) will carry out all of these procedures.

In OLS regression, it is often useful to consider how much of the total variance is explained by the model using the coefficient of variation (R^2). Unfortunately, the OLS definition of R^2 does not carry over easily into GLS. Several definitions of 'pseudo R^2' have been proposed (Menard 2000), but none of them are correct in all situations. It is therefore important to bear this issue in mind when using R^2 for PGLS regressions. Indeed, some authors prefer not to report R^2 statistics at all (e.g. Bates 2000; Lumley 2009).

A more important issue is the estimation of effect sizes and associated confidence intervals from GLS models. The parameter estimates of slopes (for continuous predictors) and the intercept and differences between means (for categorical predictors) are the most important results of the analyses. Confidence intervals for parameters can be constructed in the usual way by multiplying the standard deviation of the parameter estimate by 1.96 to derive the 95 % confidence interval, if the sample is large (roughly >30 residual degrees of freedom), or by relating to the *t* distribution if the sample is smaller.

5.4 Phylogenetic Signal

5.4.1 Phylogenetic Signal and Pagel's λ

Up to now, we have assumed that the expected phylogenetic variance–covariance matrix accurately describes the error structure of the data. In other words, we assume the phylogeny is accurate (but see 'Misconceptions, problems, and pitfalls' later) and that species trait values have evolved via a Brownian motion model of gradual evolution, with the amount of evolutionary change along a branch being proportional to the branch length. However, if the phylogeny or evolutionary model is not accurate and there is in reality less or no phylogenetic covariance in the residuals (the OLS expectation), then using the phylogeny as estimated may be inappropriate. What we need is a way of determining the extent of phylogenetic autocorrelation in the data. This can be achieved by estimating phylogenetic signal.

Phylogenetic signal is the extent to which trait values are statistically related to phylogeny. In other words, phylogenetic signal indicates the extent to which closely related species tend to resemble each other (Blomberg et al. 2003). Estimation of phylogenetic signal can provide some insight into how particular traits have evolved. Thus, traits exhibiting strong phylogenetic signal (e.g. body size and morphology; Freckleton et al. 2002) have most likely evolved by gradual changes over time (e.g. a Brownian motion model of evolution). Alternatively, traits with no phylogenetic signal (e.g. many social behaviours, Blomberg et al. 2003) may either be extremely labile (they change around very much) on the time scale of phylogeny or conversely extremely stable (they do not change at all) (Revell et al. 2008).

Our interest here lies in the application of phylogenetic signal to PGLS, so we will not provide extensive discussion of the biological significance of phylogenetic signal. For interested readers, we recommend two excellent papers on the subject of phylogenetic signal (Revell et al. 2008; Kamilar and Cooper 2013).

We shall concentrate on one of the most commonly used quantitative measures of phylogenetic signal: Pagel's λ (Pagel 1997, 1999), because this measure can be directly implemented in PGLS calculations. However, there are numerous other measures of phylogenetic signal that can be employed dependent on the statistical framework and the model of evolution assumed. Each, in some way, measures the extent to which common descent of species describes the pattern of traits across species. Examples include Moran's *I* (Gittleman and Kot 1990), Abouheif's test for serial independence (Abouheif 1999), Grafen's ρ (Grafen 1989), the

Ornstein-Uhlenbeck model parameter α (Martins and Hansen 1997), Hansen's phylogenetic half-time (Hansen 1997), Blomberg et al.'s K (Blomberg et al. 2003), Ives and Garland's 'a' and 'd' (Chap. 9), and Fritz and Purvis's D metric (Fritz and Purvis 2010). Some of these are compatible with the PGLS framework (e.g. Grafen's ρ). For more detailed reviews, see Blomberg and Garland (2002), Münkemüller et al. (2012), and Chaps. 9, 11 and 14.

We have already introduced the expected variance–covariance matrix, $\mathbf{C_{phyl}}$, that is calculated based on the phylogenetic relationships of the species in the analysis (see above). This is the expected covariance structure, but what is the actual covariance structure? We can estimate this for a single trait or, as is the case for PGLS, the residual errors (an important distinction as we shall see later). To get one of the individual off-diagonal elements, the covariance (cov) for a pair of species (*i* and *j*) and a given trait (*X*) is the product of the deviation of each species from the mean of the trait:

$$cov\left(X_i, X_j\right) = \left(X_i - \bar{X}\right)\left(X_j - \bar{X}\right)$$

For our fiddler crab X values (log carapace breadth), the observed matrix is

$$\mathbf{C_{obs}} = \begin{bmatrix} 0.0025 & 0.0045 & -0.0005 & -0.0025 & -0.0040 \\ 0.0045 & 0.0081 & -0.0009 & -0.0045 & -0.0072 \\ -0.0005 & -0.0009 & 0.0001 & 0.0005 & 0.0008 \\ -0.0025 & -0.0045 & 0.0005 & 0.0025 & 0.0040 \\ -0.0040 & -0.0072 & 0.0008 & 0.0040 & 0.0064 \end{bmatrix}$$

We might ask which is the better 'fit' to this $\mathbf{C_{obs}}$ matrix, $\mathbf{C_{phyl}}$, or $\mathbf{C_{non\text{-}phyl}}$? It is also possible that there is intermediate phylogenetic signal in the data. Might this be a more likely scenario? We can establish this by estimating λ, which is a multiplier of the off-diagonal elements of the expected variance–covariance matrix. If λ is less than 1, this has the effect of shortening the internal branches and extending the terminal branches of the tree (see Fig. 5.4). At its extremes, $\lambda = 0$ sets the off-diagonal elements to zero producing the non-phylogenetic covariance matrix, whereas $\lambda = 1$ is identical to the expected phylogenetic covariance matrix under a Brownian motion model of evolution. Values greater than 1 are not valid because the off-diagonal values in the covariance matrix cannot exceed the diagonals in GLS (species cannot be more similar to other species than they are to themselves).

λ is not calculated through the GLS formula itself. Rather, its value is estimated through maximum likelihood estimation. A λ value of 0 is consistent with no phylogenetic signal in the trait, whereas a value of 1 is consistent with strong phylogenetic signal. Intermediate values of λ indicate intermediate phylogenetic signal. Many of the R packages cited earlier can estimate λ for individual traits. In the case of our example, the maximum likelihood value of λ for carapace breadth is 1, and for claw length, it is 0.888.

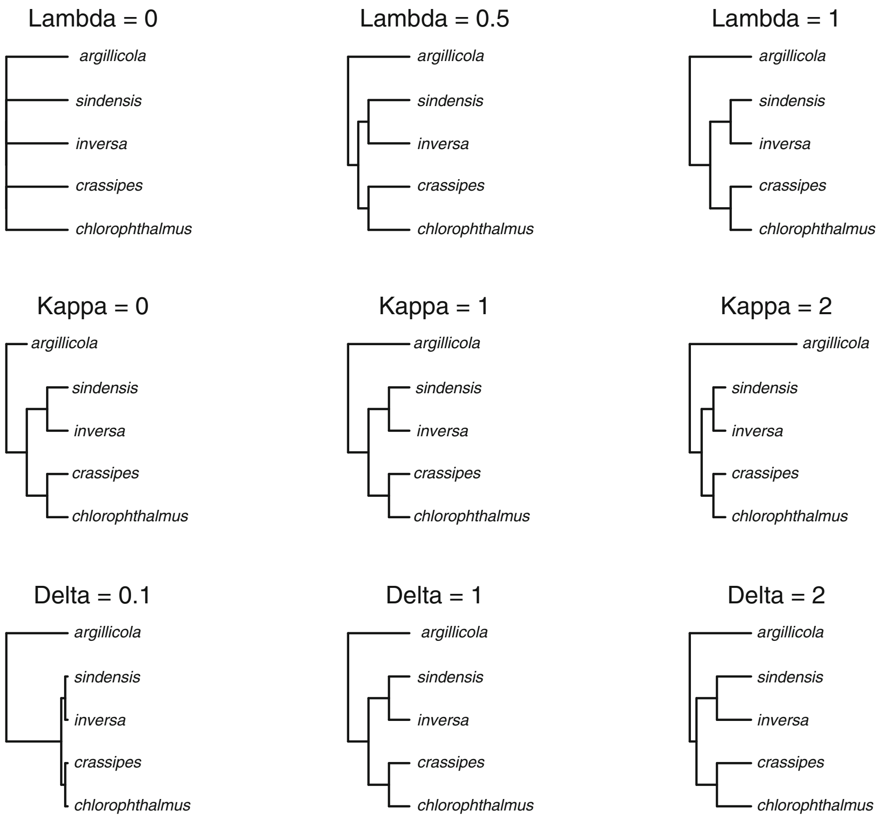


Fig. 5.4 Pagel's branch length transformations applied to the *Uca* fiddler crab phylogeny under different values of λ, δ, and κ. The $\lambda = 1$, $\delta = 1$, and $\kappa = 1$ phylogenies are identical to Fig. 5.2. Note that $\lambda = 0$ phylogeny is the same evolutionary assumption as used by traditional OLS regression (each species has independently evolved and shares no phylogenetic history)

There is no clear-cut interpretation of whether intermediate values of λ indicate 'weak' or 'strong' phylogenetic signal because it depends on the likelihood profile of λ for the specific data set (Kamilar and Cooper 2013). However, one can use likelihood ratio (LR) tests and calculate P values to assess whether the estimated maximum likelihood value of λ differs significantly from 0 or 1. As a brief aside, some authors (e.g. Pinheiro and Bates 2000) have pointed out that such likelihood ratio tests where the null value cannot exceed a certain value (such as less than 0 or more than 1) will be inherently conservative. Note that simulation studies have demonstrated that the significance of λ is also very sensitive to the number of species, and λ may perform poorly as a measure of phylogenetic signal at small sample sizes (Münkemüller et al. 2012).

It is worth pointing out that Pagel (1997, 1999) developed two other measures, related to λ, that are also branch length modifiers and are calculated through maximum likelihood estimation. The first of these, δ, is a power transformation of the summed branch lengths from the root to the tips of the tree, and the second, κ, is a power transformation of the individual branch lengths themselves. As with λ, both can be used to infer something about the evolutionary process. δ is a measure of whether trait evolution has sped up ($\delta > 1$) or slowed down ($\delta < 1$) over evolutionary time. κ is a measure of mode of evolution, with $\kappa = 0$ depicting evolutionary change that is independent of branch length—indicating a punctuated model of evolution. Figure 5.4 illustrates the effect of different values of these parameters. As with λ, both δ and κ can also be applied to PGLS calculation (see below), although they are not as commonly utilised as λ in that context.

5.4.2 Incorporating Phylogenetic Signal into PGLS

One of the principal advantages of PGLS is that one can control for the amount of phylogenetic signal in the data by altering the properties of the variance–covariance matrix **C**. In the case of independent contrasts, the usual assumption is that the phylogeny accurately describes the error structure of the data. PGLS, however, can account for intermediate levels of phylogenetic signal. With Pagel's λ, one simply multiplies the off-diagonal elements of **C** by λ and uses this new version of the matrix $\mathbf{C}_\lambda$ in the PGLS calculation. Note that the lambda multiplier can also be used to generate the modified phylogeny for use in an independent contrasts analysis, with identical results.

It is key to recognise that, in PGLS, λ applies to the residual errors from the regression model, not the strength of signal in the response variable or predictor variables. Consequently, the λ values for the PGLS regression may vary from those for the individual traits themselves (see 'Misconceptions, problems, and pitfalls' below). This is actually demonstrated by our example: the maximum likelihood estimate of λ for the regression is 0, as opposed to the individual trait values of 1 and 0.888. Thus, even though there is strong phylogenetic signal in our individual traits, there is no signal when claw length is regressed against body width, and the actual phylogenetic regression estimates will be identical to the ordinary least squares regression estimates ($b_0 = -0.229$, $b_1 = 1.577$). Note that this only applies if your phylogeny is ultrametric (all tips being the same distance from the root of the tree).

5.5 How to Present a PGLS Analysis

One advantage of PGLS is that, for the graphical presentation of relationships, one can simply plot the species data points on the relevant axes as you would do for a standard regression plot (see Fig. 5.3). The main difference is that the plot should include the PGLS regression line, rather than the standard OLS regression line.

When it comes to the presentation of the statistical analysis itself, again the presentation does not differ from what you would do with a standard regression analysis—present the PGLS estimates and standard errors, and, if appropriate, r, *t*, or F values and associated P values. The one key difference is that it is usual to present the estimate (such as λ) of the phylogenetic signal associated with the regression, along with its confidence intervals, as this provides an indication to the reader as to the extent that phylogeny is affecting the error structure of the data (remember that this is the signal associated with residual errors, not the individual variables).

Finally, because there is increasing appreciation of statistical approaches that are not based on frequentist thinking (i.e. traditional null-hypothesis significance testing with P values) (Garamszegi et al. 2009), it should be noted that PGLS is compatible with other methods of statistical inference, such as information-theoretic (e.g. using Akaike's Information Criterion) or Bayesian approaches (see Chaps. 10 and 12).

5.6 Misconceptions, Problems, and Pitfalls

As with any statistical technique, problems may arise with PGLS in practice, primarily due to violations of basic assumptions of the method. There are also several general misconceptions about phylogenetic comparative methods that apply to PGLS. For readers interested in these issues, we recommend Freckleton's (2009) review of the 'seven deadly sins of comparative analysis'. Many of these concern basic statistical assumptions, and these will be covered in the next chapter. Here, though, we address several other common practical issues.

5.6.1 Reporting Both PGLS and OLS

It is not necessary to report both PGLS and OLS (i.e. phylogenetically and non-phylogenetically controlled analyses). Unless your aim is specifically to compare the results of the two analyses (and perhaps infer the effects of phylogeny on the relationship between traits), then it is not necessary or desirable to carry out both types of analysis. The tendency to use both sets of results stemmed from concerns about the appropriateness of accounting for phylogeny in certain analyses, and perhaps a desire to 'cover one's bases' in the consequent interpretation. However, as we have seen, PGLS can explicitly take into account phylogenetic signal and hence control for it appropriately. If there is no signal in the residual structure (as we saw with our fiddler crab example), then the results of PGLS will be the same as OLS. By contrast, if there is phylogenetic signal, then PGLS will control for it, and a raw-data analysis would be statistically flawed in any event.

5.6.2 The Assumptions of the Evolutionary Model

The version of PGLS we have presented here stems from perhaps the simplest evolutionary model, the Brownian motion model (see earlier), as described by Felsenstein (1985). However, as Felsenstein (1985, p. 13) himself commented ‘there are certainly many reasons for being skeptical (*sic*) of its validity’. Of course, in the absence of other knowledge, this is perhaps a reasonable starting point. However, there are other implementations of PGLS that invoke alternative evolutionary models, such as the Ornstein–Uhlenbeck model, where there is semi-random walk evolution with a tendency towards trait optima reflecting different selective regimes (see Chaps. 14 and 15; Martins and Hansen 1997; Butler and King 2004; Hansen et al. 2008). Part III of this book examines alternative evolutionary models in detail.

5.6.3 Phylogenetic Signal in the Context of PGLS

Phylogenetic signal for traits should not be used as justification for using (or not using) PGLS. As we discussed earlier, when one has a measure of phylogenetic signal for a trait, it is possible to use likelihood ratio tests to examine whether the observed value of signal differs significantly from 0 or 1. It has become quite common to argue that if one of the traits being investigated does not display any significant phylogenetic signal, then it is unnecessary to perform a phylogenetically controlled analysis (see Revell 2010 for further discussion of this issue). However, with PGLS, the assumptions regarding phylogenetic non-independence concerns the residual errors of the regression model, not the individual traits themselves. As our fiddler crab example demonstrates, it is quite possible to have strong phylogenetic signal in the traits when examined individually but not in the residual errors (and the converse is also true).

5.6.4 Dealing with Phylogenetic Inaccuracy and Uncertainty

With any phylogenetic comparative method, a fundamental assumption is that the phylogeny being used as the basis for analysis is accurate and known without error (Harvey and Pagel 1991, p. 121). Clearly, it is highly unlikely that this will be the case, and therefore, one should bear in mind that any phylogenetic comparative analysis is naturally contingent on the particular phylogeny being used. Fortunately, simulation studies have generally found that independent contrasts and PGLS are fairly robust to errors in both phylogenetic topology and branch lengths (Díaz-Uriarte and Garland 1998; Symonds 2002; Martins and Housworth 2002; Stone 2011). However, there are several points surrounding the issue of phylogenetic uncertainty that bear consideration.

First, any phylogenetic information is better than none at all (Symonds 2002). It may be that there is not a convenient single phylogeny available, in which case inference can still be based on composite trees (i.e. when phylogenetic information from several trees is fitted together), or from supertrees (Chap. 3; Bininda-Emonds 2004). Alternatively, practitioners may attempt to produce a phylogeny themselves using published DNA sequence data (e.g. from GenBank.) There are a number of phylogenetic packages available that enable use of this approach relatively quickly (e.g. phyloGenerator: Pearse and Purvis 2013). In the complete absence of any phylogenetic information or means to construct a phylogeny, the taxonomic information may suffice (indeed the original version of PGLS as described by Grafen 1989 was based around a taxonomic 'phylogeny').

Second, sometimes, there are multiple phylogenetic hypotheses for the study species, in which case the approach advocated by Harvey (1991) of conducting analyses over multiple phylogenies can be employed. For example, Symonds and Elgar (2002) demonstrated how estimation of the metabolic scaling coefficient in mammals differs depending on which of 32 phylogenies was used as the basis for analysis. Often, phylogenetic analysis itself presents hundreds of most probable trees, and it is possible to carry out PGLS using each of these phylogenetic hypotheses. De Villemereuil et al. (2012) have developed one such approach and demonstrated that by generating regression estimates across a range of candidate trees, one improves estimation of the model parameters and associated confidence intervals. Such an approach can be combined with multimodel inference (see Chap. 12). De Villemereuil et al. (2012) argue that this approach is superior to basing analysis on a single consensus tree.

Finally, one must often deal with polytomies where more than 2 branches descend from a node. These polytomies may be an actual representation of the true evolutionary branching process, or simply a lack of knowledge of that process (so-called hard and soft polytomies, respectively, Purvis and Garland 1993). Although the original formulation of PGLS (Grafen 1989) explicitly allowed for phylogenetic uncertainty in the form of polytomies, there have been ongoing issues associated with polytomies in PGLS analyses (see discussion in Rohlf 2001), including the loss of degrees of freedom in the statistical analysis. Some packages (e.g. COMPARE, Martins 2004) do not permit polytomies at all. There are three principal recommendations for dealing with polytomies in a PGLS framework. One (usually argued in the case of 'hard' polytomies) is to arbitrarily resolve the polytomies into a fully resolved bifurcating phylogeny, but to assign zero or minimal branch length (say 0.0001) to the resolved internal branches (Felsenstein 1985). The second, more appropriate for soft polytomies, is to carry out analyses on all (or at least many) possible resolutions of the phylogenetic tree in a manner analogous to the methods above for comparing across multiple phylogenies, using the Grafen (1989) algorithm to assign branch lengths (see Chap. 12). The third is simply to reduce the degrees of freedom by making them equal to 1 for soft polytomies (Purvis and Garland 1993). A final approach, based on generalised estimating equations, has also been proposed by Paradis and Claude (2002).

5.6.5 Dealing with Intraspecific Variation

In this chapter, we have considered only variation between species and therefore used species average values as our data points. Indeed, the majority of published phylogenetic comparative analyses ignore variation within species, despite its potential impact on results (see meta-analysis by Garamszegi and Møller 2010). There are methods (Chap. 7; Ives et al. 2007; Revell and Reynolds 2012) for dealing with intraspecific variation and measurement error in the PGLS framework that have been implemented in some computer packages. In short, while obtaining detailed information on intraspecific variation might not be possible for some comparative analyses, it is recommended that it be taken into account when it is possible to do so.

Acknowledgments We are grateful to László Zsolt Garamszegi for his advice and encouragement during the writing of this chapter. Alan Grafen provided insightful comments on an earlier draft.

A.1 Appendix

A.1.1 Further Mathematical Details of the Calculation of OLS and PGLS Using Our Worked Example

An alternative way of expressing the ordinary least squares regression formula that is quicker and more effective for analysis with more than one predictor is using matrix algebra. Here, the equation to obtain regression estimates is given as

$$\boldsymbol{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$$

In this case, $\boldsymbol{\beta}$ is the vector consisting of the parameter estimates (b_0, b_1, and so on if more than one predictor variable). **X** is a matrix consisting of n rows and ($m + 1$) columns (m is the number of predictor variables), where the first column represents a constant (given the value 1 on each row), and the subsequent columns are the X values for each predictor variable. In the matrix formulation, the term $\mathbf{X}'$ denotes the 'transpose' of **X**—simply put, the rows become columns, and the columns become rows.

$$\mathbf{X} = \begin{bmatrix} 1 & 1.02 \\ 1 & 1.06 \\ 1 & 0.96 \\ 1 & 0.92 \\ 1 & 0.89 \end{bmatrix} \quad \mathbf{X}' = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1.02 & 1.06 & 0.96 & 0.92 & 0.89 \end{bmatrix}$$

When multiplied together, these become $\mathbf{X}'\mathbf{X}$, calculated as follows:

$$\mathbf{X}'\mathbf{X} = \begin{bmatrix} 5 & 4.85 \\ 4.85 & 4.724 \end{bmatrix}$$

Here, the value in row i, column j of $\mathbf{X}'\mathbf{X}$ equals the sum total of row i elements of $\mathbf{X}'$ multiplied by their respective column j elements of $\mathbf{X}$. So for example, row 2, column 2 of $\mathbf{X}'\mathbf{X}$ is $(1.02 \times 1.02) + (1.06 \times 1.06) + (0.96 \times 0.96) + (0.92 \times 0.92) + (0.89 \times 0.89) = 4.724$.

Finally, the suffix $^{-1}$ applied to $\mathbf{X}'\mathbf{X}$ indicates the 'inverse' matrix. The way the inverse matrix is calculated is somewhat complex but it is the matrix that when multiplied by it original form ($\mathbf{X}'\mathbf{X}$) produces a matrix with 1s in the diagonal elements, and 0s in the off-diagonals (this is known as the identity matrix—see below).

$$(\mathbf{X}'\mathbf{X})^{-1} = \begin{bmatrix} 48.21 & -49.49 \\ -49.49 & 51.02 \end{bmatrix}$$

$\mathbf{y}$ is the vector of n rows, containing the values of Y.

$$\mathbf{y} = \begin{bmatrix} 1.38 \\ 1.41 \\ 1.36 \\ 1.22 \\ 1.13 \end{bmatrix}$$

As with $\mathbf{X}'\mathbf{X}$, for the $\mathbf{X}'\mathbf{y}$ vector, the row i value is the overall total of each of the row i elements of $\mathbf{X}'$ multiplied by their respective counterparts in the column of $\mathbf{y}$ (i.e. row $2 = (1.02 \times 1.38) + (1.06 \times 1.41) + (0.96 \times 1.36) + (0.92 \times 1.22) + (0.89 \times 1.13) = 6.336$.

$$\mathbf{X}'\mathbf{y} = \begin{bmatrix} 6.5 \\ 6.336 \end{bmatrix}$$

Hence, when $(\mathbf{X}'\mathbf{X})^{-1}$is then multiplied by $\mathbf{X}'\mathbf{y}$, we get the OLS solution for $\boldsymbol{\beta}$

$$\boldsymbol{\beta} = \begin{bmatrix} (48.21 \times 6.5) + (-49.49 \times 6.336) \\ (-49.49 \times 6.5) + (51.02 \times 6.336) \end{bmatrix} = \begin{bmatrix} -0.229 \\ 1.577 \end{bmatrix}$$

where the first value (−0.229) is the intercept (b_0) and the second value is the slope estimate (b_1).

For generalised least squares, an additional element is added to the regression equation, in the form of the variance–covariance matrix, which represents the expected covariance structure of the residuals from the regression equation. In the case of OLS regression, the assumption is that there is no covariance between residuals (i.e. all species are independent of each other, and residuals from closely related species are not more similar on average than residuals from distantly related species). This ($n \times n$) variance–covariance matrix is denoted as **C**, and the regression equation becomes

$$\boldsymbol{\beta} = \left(\mathbf{X}'\mathbf{C}^{-1}\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{C}^{-1}\mathbf{y}$$

Under the assumption that there is no covariance among the residuals and they are normally distributed, with mean = 0 and standard deviation σ_ε, then

$$\mathbf{C} = \begin{bmatrix} \sigma_\varepsilon^2 & 0 & 0 & 0 & 0 \\ 0 & \sigma_\varepsilon^2 & 0 & 0 & 0 \\ 0 & 0 & \sigma_\varepsilon^2 & 0 & 0 \\ 0 & 0 & 0 & \sigma_\varepsilon^2 & 0 \\ 0 & 0 & 0 & 0 & \sigma_\varepsilon^2 \end{bmatrix}$$

The diagonal elements (the line of values from top left to bottom right) therefore represent the variance of the residuals, while the other off-diagonal elements = 0, meaning there is no covariation among the residuals. The inverse of this matrix, $\mathbf{C}^{-1}$, has essentially the same properties (all the off-diagonal elements remain as 0) except the diagonal elements now equal $1/\sigma_\varepsilon^2$. When this variance–covariance structure is assumed, the results of GLS are the same as those of OLS (the **C** part of the regression equation essentially drops out). On the other hand, if the variances are not equal, then you have a standard weighted least squares regression.

For phylogenetic generalised least squares, our expected variance–covariance matrix is $\mathbf{C_{phyl}}$ (see main text), and its inverse

$$\mathbf{C_{phyl}} = \begin{bmatrix} 3 & 2 & 1 & 1 & 0 \\ 2 & 3 & 1 & 1 & 0 \\ 1 & 1 & 3 & 2 & 0 \\ 1 & 1 & 2 & 3 & 0 \\ 0 & 0 & 0 & 0 & 3 \end{bmatrix}$$

$$\mathbf{C_{phyl}^{-1}} = \begin{bmatrix} 0.619 & -0.381 & -0.048 & -0.048 & 0 \\ -0.381 & 0.619 & -0.048 & -0.048 & 0 \\ -0.048 & -0.048 & 0.619 & -0.381 & 0 \\ -0.048 & -0.048 & -0.381 & 0.619 & 0 \\ 0 & 0 & 0 & 0 & 0.333 \end{bmatrix}$$

Taking apart the components of the GLS regression equation, we first calculate the product $\mathbf{X}'\mathbf{C}^{-1}$whose row i and column j values are the total of the ith row of

$\mathbf{X}'$ multiplied by the *j*th column of $\mathbf{C}^{-1}$. So, for example, row 2, column 3 of $\mathbf{X}'\mathbf{C}^{-1}$ is (1.02 × -0.048) + (1.06 × -0.048) + (0.96 × 0.619) + (0.92 × -0.381) + (0.89 × 0) = 0.144

$$\mathbf{X}'\mathbf{C}^{-1} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1.02 & 1.06 & 0.96 & 0.92 & 0.89 \end{bmatrix} \times \begin{bmatrix} 0.619 & -0.381 & -0.048 & -0.048 & 0 \\ -0.381 & 0.619 & -0.048 & -0.048 & 0 \\ -0.048 & -0.048 & 0.619 & -0.381 & 0 \\ -0.048 & -0.048 & -0.381 & 0.619 & 0 \\ 0 & 0 & 0 & 0 & 0.333 \end{bmatrix}$$
$$= \begin{bmatrix} 0.142 & 0.142 & 0.142 & 0.142 & 0.333 \\ 0.137 & 0.177 & 0.144 & 0.104 & 0.296 \end{bmatrix}$$

In similar fashion $\mathbf{X}'\mathbf{C}^{-1}\mathbf{X}$is therefore

$$\mathbf{X}'\mathbf{C}^{-1}\mathbf{X} = \begin{bmatrix} 0.142 & 0.142 & 0.142 & 0.142 & 0.333 \\ 0.137 & 0.177 & 0.144 & 0.104 & 0.296 \end{bmatrix} \times \begin{bmatrix} 1 & 1.02 \\ 1 & 1.06 \\ 1 & 0.96 \\ 1 & 0.92 \\ 1 & 0.89 \end{bmatrix}$$
$$= \begin{bmatrix} 0.901 & 0.859 \\ 0.859 & 0.825 \end{bmatrix}$$

The inverse of which is

$$(\mathbf{X}'\mathbf{C}^{-1}\mathbf{X})^{-1} = \begin{bmatrix} 130.141 & -135.389 \\ -135.389 & 142.060 \end{bmatrix}$$

The second component of the GLS regression equation $\mathbf{X}'\mathbf{C}^{-1}\mathbf{y}$ follows likewise as

$$\mathbf{X}'\mathbf{C}^{-1}\mathbf{y} = \begin{bmatrix} 0.142 & 0.142 & 0.142 & 0.142 & 0.142 \\ 0.137 & 0.177 & 0.144 & 0.104 & 0.296 \end{bmatrix} \times \begin{bmatrix} 1.38 \\ 1.41 \\ 1.36 \\ 1.22 \\ 1.13 \end{bmatrix} = \begin{bmatrix} 1.139 \\ 1.097 \end{bmatrix}$$

where, for example, the first row value (1.139) = (0.142 × 1.38) + (0.142 × 1.41) + (0.142 × 1.36) + (0.142 × 1.22) + (0.142 × 1.13).

Finally, we can combine our two products to obtain the PGLS solution for $\boldsymbol{\beta}$.

$$\boldsymbol{\beta}_{\mathbf{PGLS}} = \left(\mathbf{X}'\mathbf{C}^{-1}\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{C}^{-1}\mathbf{y} = \begin{bmatrix} 130.141 & -135.389 \\ -135.389 & 142.060 \end{bmatrix} \times \begin{bmatrix} 1.139 \\ 1.097 \end{bmatrix}$$
$$= \begin{bmatrix} -0.276 \\ 1.616 \end{bmatrix}$$

where $b_0 = -0.276$ and $b_1 = 1.616$.

References

Abouheif E (1998) Random trees and the comparative method: a cautionary tale. Evolution 52:1197–1204

Abouheif E (1999) A method for testing the assumption of phylogenetic independence in comparative data. Evol Ecol Res 1:895–909

Bates D (2000) fortunes: R fortunes. R package version 1.5-0, http://CRAN.R-project.org/package=fortunes

Bininda-Emonds ORP (ed) (2004) Phylogenetic supertrees: combining information to reveal the tree of life. Kluwer Academic Publishers, Dordrecht

Björklund M (1997) Are ‘comparative methods’ always necessary? Oikos 80:607–612

Blomberg SP, Garland T Jr (2002) Tempo and mode in evolution: phylogenetic inertia, adaptation and comparative methods. J Evol Biol 15:899–910

Blomberg SP, Garland T Jr, Ives AR (2003) Testing for phylogenetic signal in comparative data: behavioral traits are more labile. Evolution 57:717–745

Blomberg SP, Lefevre JG, Wells JA, Waterhouse M (2012) Independent contrasts and PGLS regression estimators are equivalent. Syst Biol 61:382–391

Bonduriansky R (2007) Sexual selection and allometry: a critical reappraisal of the evidence and ideas. Evolution 61:838–849

Butler MA, King AA (2004) Phylogenetic comparative analysis: a modelling approach for adaptive evolution. Am Nat 164:683–695

Crane J (1975) Fiddler crabs of the Wworld: ocypodidae: genus Uca. Princeton University Press, Princeton

De Villemereuil P, Wells JA, Edwards RD, Blomberg SP (2012) Bayesian models for comparative analysis integrating phylogenetic uncertainty. BMC Evol Biol 12:102

Díaz-Uriarte R, Garland T Jr (1998) Effects of branch lengths errors on the performance of phylogenetically independent contrasts. Syst Biol 47:654–672

Felsenstein J (1973) Maximum-likelihood estimation of evolutionary trees from continuous characters. Am J Human Genet 25:471–492

Felsenstein J (1985) Phylogenies and the comparative method. Am Nat 125:1–15

Freckleton RP (2009) The seven deadly sins of comparative analysis. J Evol Biol 22:1367–1375

Freckleton RP, Harvey PH, Pagel M (2002) Phylogenetic analysis and comparative data: a test and review of evidence. Am Nat 160:712–726

Fritz SA, Purvis A (2010) Selectivity in mammalian extinction risk and threat types: a new measure of phylogenetic signal strength in binary traits. Conserv Biol 24:1042–1051

Garamszegi LZ, Møller AP (2010) Effects of sample size and intraspecific variation in phylogenetic comparative studies: a meta-analytic review. Biol Rev 85:797–805

Garamszegi LZ, Calhim S, Dochtermann N, Hegyi G, Hurd PL, Jørgensen C, Kutsukake N, Lajeunesse MJ, Pollard KA, Schielzeth H, Symonds MRE, Nakagawa S (2009) Changing philosophies and tools for statistical inferences in behavioral ecology. Behav Ecol 20:1363–1375

Garland T Jr, Ives AR (2000) Using the past to predict the present: confidence intervals for regression equations in phylogenetic comparative methods. Am Nat 155:346–364

Garland T Jr, Harvey PH, Ives AR (1992) Procedures for the analysis of comparative data using phylogenetically independent contrasts. Syst Biol 41:18–32

Gittleman JL, Kot M (1990) Adaptation: statistics and a null model for estimating phylogenetic effects. Syst Zool 39:227–241

Grafen A (1989) The phylogenetic regression. Phil Trans R Soc B 326:119–157

Grafen A (2014) phyreg: Implements the phylogenetic regression of Grafen (1989). http://cran.r-project.org/web/packages/phyreg/index.html

Hansen TF (1997) Stabilizing selection and the comparative analysis of adaptation. Evolution 51:1341–1351

Hansen TF, Pienaar J, Orzack SH (2008) A comparative method for studying adaptation to a randomly evolving environment. Evolution 62:1965–1977

Harvey PH (1991) Comparing uncertain relationships: the Swedes in revolt. Trends Ecol Evol 6:38–39

Harvey PH, Pagel MD (1991) The comparative method in evolutionary biology. Oxford University Press, Oxford

Huey RB (1987) Phylogeny, history and the comparative method. In: Feder ME, Bennett AF, Burggren WW, Huey RB (eds) New directions in ecological physiology. Cambridge University Press, Cambridge, pp 76–101

Ives AR, Garland T Jr (2010) Phylogenetic logistic regression for binary dependent variables. Syst Biol 59:9–26

Ives AR, Midford PE, Garland T Jr (2007) Within-species variation and measurement error in phylogenetic comparative methods. Syst Biol 56:252–270

Kamilar JM, Cooper N (2013) Phylogenetic signal in primate behaviour, ecology and life history. Phil Trans R Soc B 368:20120341

Kembel SW, Cowan PD, Helmus MR, Cornwell WK, Morlon H, Ackerly DD, Blomberg SP, Webb CO (2010) Picante: R tools for integrating phylogenies and ecology. Bioinformatics 26:1463–1464

Lumley T (2009) fortunes: R fortunes. R package version 1.5-0, http://CRAN.R-project.org/package=fortunes

Losos JB (2011) Seeing the forest for the trees: the limitations of phylogenies in comparative biology. Am Nat 177:709–727

Maddison DR, Swofford DL, Maddison WP (1997) Nexus: an extensible file format for systematic information. Syst Biol 46:590–621

Maddison WP (1990) A method for testing the correlated evolution of two binary characters: are gains or losses concentrated on certain branches of a phylogenetic trees? Evolution 44:539–557

Maddison WP (2000) Testing character correlation using pairwise comparisons on a phylogeny. J Theor Biol 202:195–204

Martins EP (2004) COMPARE. Version 4.6b. Computer programs for the statistical analysis of comparative data. Department of Biology, Indiana University, Bloomington. http://compare.bio.indiana.edu/

Martins EP, Hansen TF (1997) Phylogenies and the comparative method: a general approach to incorporating phylogenetic information into the analysis of interspecific data. Am Nat 149:646–667

Martins EP, Housworth EA (2002) Phylogeny shape and the phylogenetic comparative method. Syst Biol 51:873–880

Menard S (2000) Coefficients of determination for multiple logistic regression analysis. Am Stat 54(1):17–24

Münkemüller T, Lavergne S, Bzeznik B, Dray S, Jombart T, Schiffers K, Thuiller W (2012) How to measure and test phylogenetic signal. Methods Ecol Evol 3:743–756

Orme D, Freckleton R, Thomas G, Petzoldt T, Fritz S, Isaac N, Pearse W (2012) caper: comparative analysis of phylogenetics and evolution in R. http://CRAN.R-project.org/package=caper

Pagel MD (1992) A method for the analysis of comparative data. J Theor Biol 156:431–442

Pagel M (1994) Detecting correlated evolution on phylogenies: a general method for the comparative analysis of discrete characters. Proc R Soc B 255:37–45

Pagel M (1997) Inferring evolutionary processes from phylogenies. Zool Scripta 26:331–348

Pagel M (1999) Inferring the historical patterns of biological evolution. Nature 401:877–884

Pagel M, Meade A (2013) BayesTraits version 2.0 (Beta). University of Reading. http://www.evolution.rdg.ac.uk/BayesTraits.html

Paradis E, Claude J (2002) Analysis of comparative data using generalized estimating equations. J Theor Biol 218:175–185

Paradis E, Claude J, Strimmer K (2004) APE: analysis of phylogenetics and evolution in R language. Bioinformatics 20:289–290

Pearse WD, Purvis A (2013) phyloGenerator: an automated phylogeny generation tool for ecologists. Methods Ecol Evol 4:692–698

Pinheiro JC, Bates DM (2000) Mixes-effects models in S and S-PLUS. Springer, Berlin

Pinheiro J, Bates D, DebRoy S, Sarker D, R Development Core Team (2013) nlme: linear and nonlinear mixed effects models. R package version 3.1-111. http://cran.r-project.org/web/packages/nlme/index.html

Promislow DEL, Harvey PH (1990) Living fast and dying young: a comparative analysis of life-history variation among mammals. J Zool 220:417–437

Purvis A, Garland T Jr (1993) Polytomies in comparative analysis of continuous characters. Syst Biol 42:569–575

Purvis A, Gittleman JL, Luh H-K (1994) Truth or consequences: effects of phylogenetic accuracy on two comparative methods. J Theor Biol 167:293–300

Revell LJ (2010) Phylogenetic signal and linear regression on species data. Methods Ecol Evol 1:319–329

Revell LJ (2012) phytools: an R package for phylogenetic comparative biology (and other things). Methods Ecol Evol 3:217–223

Revell LJ, Reynolds RG (2012) A new Bayesian method for fitting evolutionary models to comparative data with intraspecific variation. Evolution 66:2697–2707

Revell LJ, Harmon LJ, Collar DC (2008) Phylogenetic signal, evolutionary process and rate. Syst Biol 57:591–601

Rheindt FE, Grafe TU, Abouheif E (2004) Rapidly evolving traits and the comparative method: how important is testing for phylogenetic signal? Evol Ecol Res 6:377–396

Ridley M (1983) The explanation of organic diversity. Oxford University Press, Oxford

Rohlf FJ (2001) Comparative methods for the analysis of continuous variables: geometric interpretations. Evolution 55:2143–2160

Rosenberg MS (2001) The systematics and taxonomy of fiddler crabs: a phylogeny of the genus Uca. J Crust Biol 21:839–869

Rosenberg MS (2002) Fiddler crab claw shape variation: a geometric morphometric analysis across the genus Uca (Crustacea: Brachyura: Ocypodidae). Biol J Linn Soc 75:147–162

Stone EA (2011) Why the phylogenetic regression appears robust to tree misspecification. Syst Biol 60:245–260

Symonds MRE (2002) The effects of topological inaccuracy in evolutionary trees on the phylogenetic comparative method of independent contrasts. Syst Biol 51:541–553

Symonds MRE, Elgar MA (2002) Phylogeny affects estimation of metabolic scaling in mammals. Evolution 56:2330–2333

Westoby M, Leishman MR, Lord JM (1995) On misinterpreting the ‘phylogenetic correction’. J Ecol 83:531–534