# Phylogenetic Regressions for Multivariate Comparative Data

Julien Clavel & Hélène Morlon

08/09/2019

## Contents



---

# 1 Multivariate GLS in mvMORPH

Multivariate GLS linear model fit and multivariate hypotheses tests can be performed in mvMORPH with the `mvgls` and `manova.gls` functions respectively. These functions are the multivariate counterpart to the `gls` and `anova` functions in "nlme" with the phylogenetic `corClasses` objects from "ape". The `mvgls` function allows fitting multivariate model using either maximum likelihood (`method="LL"`) when $p<n$-$q$ or penalized likelihood for any $p$ (`method="PL-LOOCV"`, see Clavel et al. 2019). NOTE: The current version of mvMORPH (1.1.1) allows the use of the MANOVA tests in `manova.gls` with only the `RidgeArch` penalty for the penalized likelihood approach (see `?mvgls`).

The main options taken by the `manova.gls` function are described in the table below:

| Arguments | Description | Values |
|---|---|---|
| `test` | The multivariate test statistic used in the MANOVA | `"Pillai"` (default), `"Wilks"`, `"Lawley-Hotteling"`, `"Roy"` |
| `type` | The type of sums of squares (SS) and cross-products used with multiple predictors | `type="I"` (default), `"II"`, or `"III"` |
| `nperm` | The number of permutations used to assess the significance of the test* | Default is `nperm=999` |
| `nbcores` | The number of cores to use to spread the calculations and speed-up the computations | Default is 1. Note that this works only on Mac and Linux platforms |
| `permutation` | The type of optimization procedure used on the permuted datasets | `permutation="approx"` (default), or `full` |
| `L` | The contrasts coding matrix for user-specified generalized linear hypothesis test | (optional - user defined, default is NULL) |
| `rhs` | The right hand side equation of the generalized linear hypothesis test | (optional - user defined, default is `rhs=0`) |

*When the model fit is estimated by ML, parametric tests are returned by default unless otherwise specified (see below).

## 2 Maximum Likelihood or Penalized Likelihood when $p < n$?

We generally advocate the use of the PL-based method (`method="PL-LOOCV"`) to perform model fit and multivariate tests because simulations show that even a slight regularization can considerably improve the performances of the tests (Figs 3 and 4 in the MS). Moreover, with large sample size the PL approach will tend to the ML (`method="LL"`) approach given that the regularization parameter will tend to zero. This means that the PL approach is general as it leads to the conventional MANOVA tests when $n>>p$ and satisfies several optimality criteria when $p>n$ (see Clavel et al. 2019). However, there are situations when one might be interested in using the ML approach. For instance, when estimated using the ML approach, the multivariate statistics can be used to compute measures of multivariate association (the multivariate counterpart to the $R^2$). As an example, with Wilks' $\Lambda$, a simple multivariate measure of association is obtained by (see Rencher 2002, p. 172):

$$\eta^2 = 1 - \Lambda$$

Further works are needed to develop high-dimensional multivariate measures of association.

## 3 Multivariate statistics (`test`)

There are various multivariate statistics proposed in the literature. The four most common ones are the Pillai's trace (default in `manova.gls`), Wilks's $\Lambda$, Roy's largest root, and Lawley-Hotelling tests. All these statistics have the same type I error rate (i.e. the probability of rejection of the null hypothesis $H_0$ when it is true), however they differ in their probability of rejecting $H_0$ when it is false (i.e. they have different power).

The performances of the different statistics depend on the structure of the multivariate space as they depend on the eigenvalues of the matrix $E^{-1}H$ described in the paper. For instance, when the mean vectors are collinear, **Roy's** test is the more powerful since it only uses the largest eigenvalue but is outcompeted by the others when the mean vectors are more diffuse (Rencher 2002). Given that in most empirical cases mean vectors will not be collinear, Roy's test might be of least practical interest. **Pillai's trace** test is often considered as more robust because it performs better when there is heterogeneity of covariance matrices in factorial designs. However, the **Wilks** $\Lambda$ is not far behind and has a long history in the statistical literature because of its connection with likelihood ratio tests (LRTs) and well known approximations of the $F$-statistic. Pros and cons are discussed at greater length in Olson 1974 or Rencher 2002.

# 4 Types of sums of squares and cross-products (`type`)

When considering multiple predictors, that may be to some extent correlated, there are different ways to assess the effects of each of them. Correlations or collinearity between predictors may arise for multiple reasons. For instance, because of unbalanced grouping schemes in factorial designs, or because of autocorrelation between samples - as in comparative data. In such a situation, there is an overlap in the residual error explained by the different factors. Hence there are different ways to treat this shared part of residual error in the tests, which correspond to the computation of two broad classes of sums of squared (SS) errors [or equivalently sums of squares and cross-products for multivariate tests]: the sequential (type I) and adjusted (type II or III). For instance, the `type="I"` considers that the shared part is explained by the predictors that precede new predictors entered in the model in a sequential fashion. In contrast, adjusted tests (`type="II"` or `type="III"`) consider only the part of the variation that can be uniquely attributed to a predictor given all the predictors in the model. We briefly present these differences below.

## 4.1 Type I (sequential)

Sequential tests have the desirable feature that they are additive. That is, the total sum of squares remains the same for all possible ordering of the explanatory (or independent) variables. This means that all the variation in the model is accounted for by those tests while the adjusted tests may miss some of the residual variance or account for it twice. However, in sequential tests the importance of the factors depends on the order they are entered in the model. Hence, one may obtain a different result depending on the ordering of the predictors. With multiple predictors the sequential tests can be interesting when they are ordered theoretical or by the experimental design. The comparison of sequential tests with different ordering of the variables is also useful to assess the effect of adjusting for a given variable first on the tests significance.

## 4.2 Type II (adjusted)

Type II tests use adjusted sums of squares. In contrasts to the sequential tests, the tests performed with type II SS do not depend on the ordering of the predictors. Moreover, these adjusted sums of squares are said to obey the principle of marginality. That is, they respect the hierarchy of the predictors level (e.g., main effects and interactions). The significance of the tests for the main effects does not account for higher-order effects terms (e.g. interactions) while the tests for the higher-orders account for all of the lower-orders terms.

## 4.3 Type III (adjusted)

Type III tests also use adjusted sums of squares. The main difference with the type II is that they include higher-level terms (e.g. interactions) when making adjustments. In other words, the tests for an effect are performed after adjusting for all the other model terms (hence they violate the marginality principle since main effect tests are accounting for the interactions). This type of test is controversial because it is often assumed that interpreting main effects in the presence of significant interactions is of little interest and that

type II tests are more powerful when the interaction is nonsignificant (see e.g. Langsrud 2003). However, these tests do not depend on the differences in sample sizes for factorial designs (assuming the treatments have been correctly adjusted using the `?contrasts` argument. See also `?Anova` in `car` package and McFarquhar 2016).

### 4.4 Summary

When a model has multiple predictors (but no interaction terms), then the choice is whether a sequential (`type="I"`) or an adjusted test (either `type="II"` or `type="III"`) is needed. For instance, if we have two predictors A and B, and we are interested in testing the residual effects of B that remain after controlling for A, then a sequential test should be used (`type="I"`). When there is no such *a priori* order, then an adjusted test is less arbitrary and should be preferred (e.g. `type="II"`). In practice, if there are low covariations between the predictors, using adjusted or non-adjusted tests won't make a big difference.

When a multiple regression model includes interaction terms (i.e. higher order terms), we should look at the results of the highest-order term first. All the types of test will give the same answer for this term, and when it is significant, they indicate that we can confidently reject the null hypothesis that the main effects are additives because they are interdependent. The lower order terms will be different between all the tests. In most situations however, the tests of the main effects are of little interest when there is a significant interaction.

N.B. In base R, the default option for the conventional (i.e. non phylogenetic) ANOVA and MANOVA tests is the sequential (or type I) SS. In the `car` package the default is type II. We chose to use the `type I` test as the default to follow base R implementation.

## 5 Number of permutations (`nperm`)

The permutations are used to approximate the distribution of the multivariate statistic under the null hypothesis ($H_0$) being evaluated. The larger the better, but this has a computational cost. We think that ~999 permutations is probably enough in most situations. This can be assessed by plotting the permuted distribution(s) using the `plot()` function on the `manova.gls` output (see below). Furthermore, parallel computing (through forking) is available on Linux and Mac OS to speed-up the computations (see the example below). Note that it is also possible to use permutations rather than parametric tests on ML model fit by using the argument `"parametric=FALSE"` (see `?manova.gls`) for example if the sample size is small.

## 6 Permutation strategies (`permutation`)

When using a different evolutionary model than Brownian Motion (e.g. `lambda`, `OU` or `EB`), there are different ways of computing the multivariate statistics on the permuted samples. The approximate strategy (`permutation="approx"` - the default option) fixes the evolutionary model to the parameters estimated by the `mvgls` function and only tries to optimize the level of regularization/penalization on each permuted datasets. The "exact" strategy (`permutation="full"`) optimizes both the evolutionary model and regularization/penalization parameters on each permuted datasets. While this permutation scheme accounts for the estimation of all the parameters in the permuted datasets, optimization is more difficult and time consuming. We chose to use by default the approximate version since it is more stable, faster, and there were no noticeable differences under a large set of simulations.

## 7 Examples

To illustrate how to fit high-dimensional phylogenetic regressions in mvMORPH, we use here two simple simulated examples.

i) In the first example, we use the traditionnal one-way MANOVA to test for differences between two groups ($q$=2) on a $n$=36 species tree with $p$=5 traits. Although the method can be used on cases where $p$>$n$-$q$, here we used a simpler (and faster) example to illustrate the general workflow. We also fit an ANCOVA model (multiple predictors) to illustrate the differences between the types of sums of squares and cross-products (`type`).

ii) The second example is directed to the more experienced users who want to perform general linear hypothesis tests using contrast coding. General linear hypothesis testing is a powerful tool to design custom hypothesis tests based on the multivariate statistics (including pairwise and planned comparisons).

Finally, we illustrate the use of the multivariate tests on the empirical phyllostomid bat dataset ($p$>$n$-$q$) from Monteiro & Nogueira (2011).

## 7.1 Model fit and hypothesis testing: the one-way MANOVA

In the following simulated example, we want to test if the hypothetical species inhabiting the "forest" habitat ($n$=18) show differences with species from the "open" habitat ($n$=18).

```
# load the package and simulate some fake data
library(mvMORPH)
set.seed(1, sample.kind = "Rounding")     # for reproducibility with older R versions
n = 36; p=5
tree = pbtree(n=n)
R = crossprod(matrix(runif(p*p),p,p))     # random covariance matrix
theta = rep(0,p)                          # the mean of the p traits
trait <- mvSIM(tree, model="BM1",         # phenotypic traits
               param=list(sigma=R, theta=theta))
grp <-  rep(c("forest","open"), each=n/2)
pheno <-  trait + rep(c(0,10), each=n/2)  # introduce some differences (~10) on "open"
size <- rTraitCont(tree)                  # an hypothetical "body-size" for the species
```

Now that we have our dataset ("trait") and grouping categories ("grp"), we can fit the phylogenetic multivariate linear model by penalized likelihood (PL). We will consider here the Pagel's $\lambda$ model for the phylogenetic structure, and use the default parameters values of the `mvgls` function (an archetypal ridge penalty [ `penalty="RidgeArch"` ] with a target matrix proportional to the identity [`target="unitVariance"`] - See ?`mvgls` for details).

```
# First create a list object with the required data
data = list(trait=pheno, habitat=as.factor(grp), size=size)

# Fit the multivariate linear model
fit <- mvgls(trait ~ habitat, data=data, tree=tree, model="lambda")
```

Note that for a MANOVA design, we have to specify that the values of the "grp" variable are factors. This is done above by using the `as.factor()` function, as for conventional regressions in R.

We can have a look at the model fit output:

```
# Print the model fit
print(fit)
```

```
>
> Call:
> mvgls(formula = trait ~ habitat, data = data, tree = tree, model = "lambda")
>
>
> Generalized least squares fit by penalized REML
```

```
> LOOCV of the log-restricted-likelihood: -120.8996
>
>
> Parameter estimate(s):
> lambda: 1
>
> Regularization parameter (gamma): 0.0056
>
>
> Covariance matrix of size: 5 by 5
> for 36 observations
>
> Coefficients:
>              [,1]      [,2]      [,3]      [,4]      [,5]
> (Intercept)   0.15306   0.87385   0.05411  -0.13147   0.73431
> habitatopen   8.26574   7.67076   8.63673   8.97492   8.40437
```

The model fit indicates that the estimated Pagel's $\lambda$ is equal to 1 meaning that there is a Brownian motion (BM) correlation structure in the model residuals (as expected, given that we simulated under BM...). Furthermore, the regularization parameter is ~0.0056, indicating that there is no need for a strong regularization of the covariance matrix. This is also expected since $n > p$ in this example. We can further note from the estimated coefficients that there are clear differences between the two groupings [the intercept here corresponds to the "forest" species; this is because R uses by default the first factor as baseline], with the "open" dwellers having a mean value which is roughly higher by a value of 10. So far so good, we retrieved the simulated features.

Now that we have fitted our model, we can test the null hypothesis ($H_0$) that there are no differences between "forest" and "open" species using the `manova.gls` function with 999 permutations.

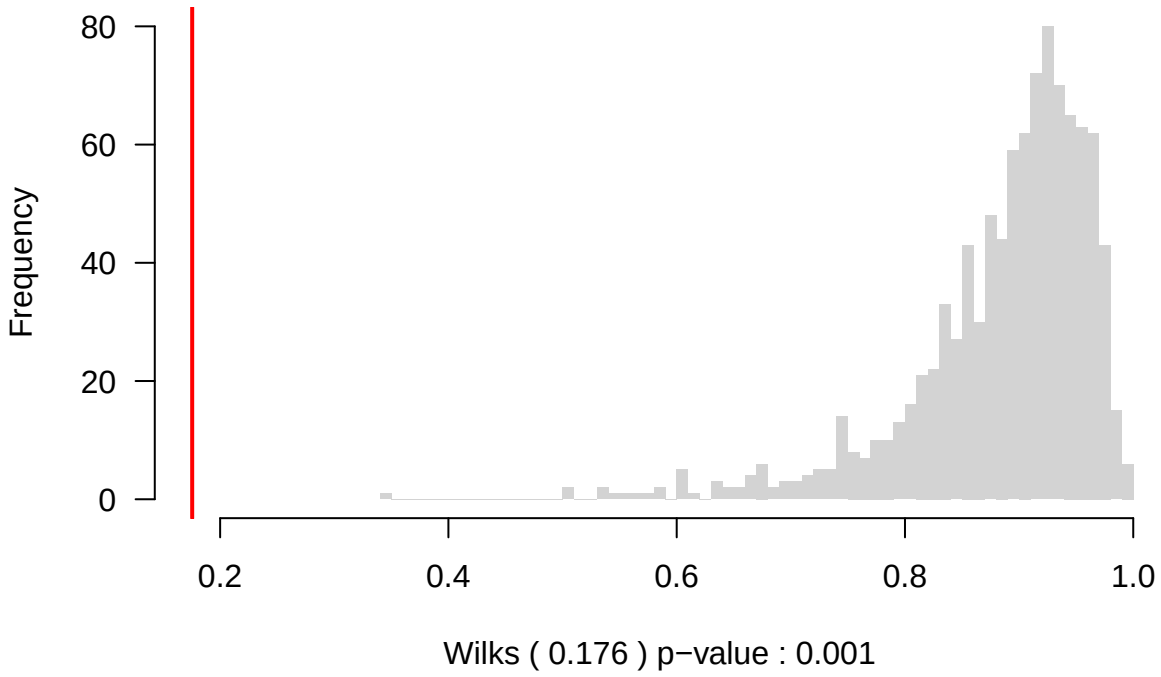
```
# The overall MANOVA test:
(aov <- manova.gls(fit,
                   nperm=999, test="Wilks", verbose=TRUE))
```

```
> Sequential MANOVA Tests with 999 permutations: Wilks test statistic
>         Test stat Pr(>Stat)
> habitat    0.1755      0.001 **
> ---
> Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The output of the function shows that the estimated Wilks's statistic is ~0.17 and that the $p$-value is significant so that we can confidently reject the null hypothesis.

It is possible to visualise the distribution of the test statistic under the null hypothesis obtained by permutations using the `plot()` function on the "manova" object:

```
# display the test statistic
plot(aov)
```



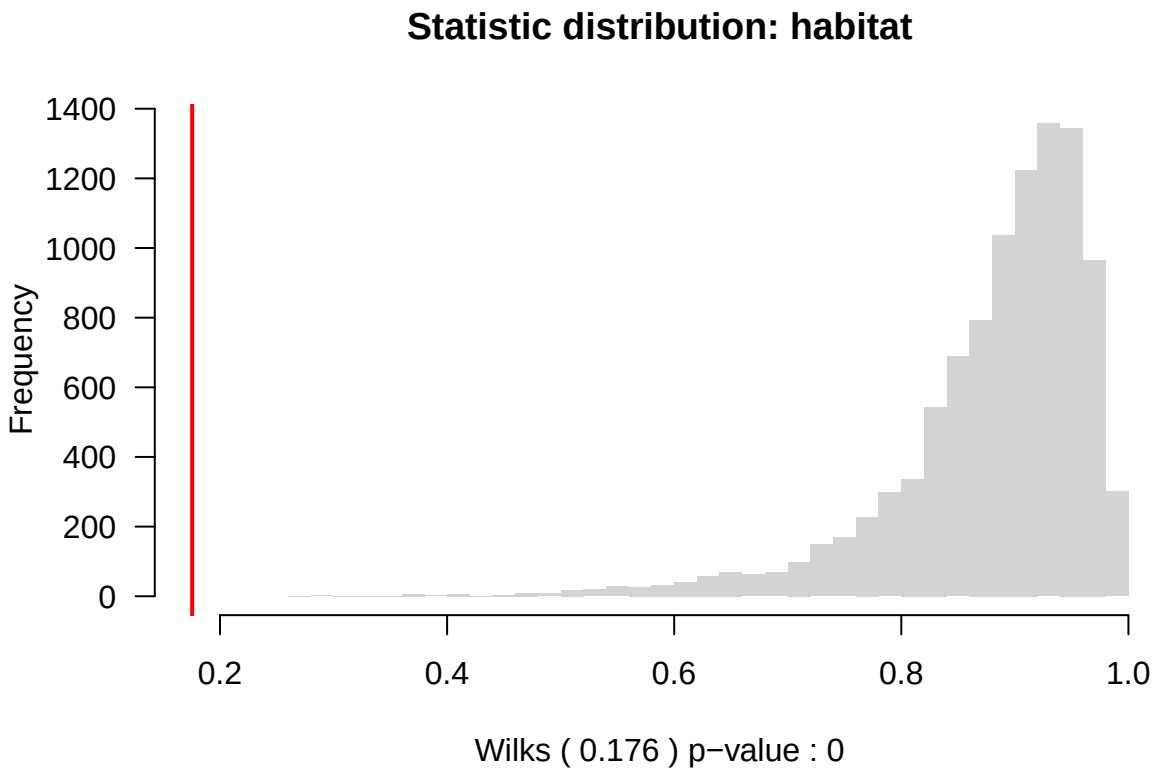

Of course, if we increase the number of permutations we will obtain a better (smoother) distribution of the test statistic under $H_0$ - although it seems that it's not necessary here.

To reduce the computational burden of the permutation procedure, we can use "parallel" computing (on Linux and Mac only) through the `nbcores` argument and re-run the analysis with an increased number of permutations:

```
# Fit with parallel computing on 4 cores for 9999 permutations
(aov2 <- manova.gls(fit,
                nperm=9999, nbcores=4L, test="Wilks", verbose=TRUE))
```

```
> Sequential MANOVA Tests with 9999 permutations: Wilks test statistic
>         Test stat Pr(>Stat)
> habitat    0.1755     1e-04 ***
> ---
> Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# display the test statistic
plot(aov2)
```




## 7.2 Multivariate multiple regression tests (multiple predictors, MANCOVA, etc.)

In this new example we reuse the previous data and we add a supplementary predictor ("size") to the model. This model with a continuous covariate is known as an Analysis of Covariance (ANCOVA or MANCOVA):

```
# Fit the MANCOVA model
fit2 <- mvgls(trait ~ size + habitat, data=data, tree=tree, model="lambda")
```

Again, the effects can be tested using the `manova.gls` function:

```
# The overall MANOVA test:
(aov3 <- manova.gls(fit2,
                nperm=999, test="Wilks", verbose=TRUE))
```

```
> Sequential MANOVA Tests with 999 permutations: Wilks test statistic
>         Test stat Pr(>Stat)
> size       0.8559     0.355
> habitat    0.1742     0.001 **
> ---
> Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The MANOVA table shows that only the "habitat" factor is significant, and that the "size" effect is not (this is expected since the "size" has been simulated independently of the trait data). We also note that this analysis required the permutations to be performed for each predictor. The default test used here is the "sequential" test also known as "type I" sum of squares (and cross-products). This test assesses the effect of each predictor in a sequential order after accounting for the effect of the previous ones. This means that different results might be obtained depending on the order of the predictors (see above).

We can use instead an adjusted or "type II" sum of squares (and cross products) test to avoid this issue (if we want to know the marginal contribution of each factor):

```
# The overall (type II) MANOVA test:
(aov4 <- manova.gls(fit2,
               nperm=999, test="Wilks", type="II", verbose=TRUE))
```

```
> Type II MANOVA Tests with 999 permutations: Wilks test statistic
>         Test stat Pr(>Stat)
> size       0.8862     0.538
> habitat    0.1742     0.001 **
> ---
> Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

In this example the results differ only slightly which is expected given that the predictors were simulated with no interdependences.

It is also possible to assess the effect of an interaction between both predictors using the `*` operator as with the classical `lm` function:

```
# Fit the MANCOVA model with interaction term
fit3 <- mvgls(trait ~ size + habitat + size*habitat,
              data=data, tree=tree, model="lambda")

# The overall (type II) MANOVA test:
(aov5 <- manova.gls(fit3,
               nperm=999, test="Wilks", type="II", verbose=TRUE))
```

```
> Type II MANOVA Tests with 999 permutations: Wilks test statistic
>              Test stat Pr(>Stat)
> size            0.8852     0.491
> habitat         0.1738     0.001 **
> size:habitat    0.9041     0.686
> ---
> Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Once again, the results show that only the habitat effect is significant. This is totally expected since we simulated a large difference (~10) between both groups but no relationship to the variable "size".

## 7.3 General Linear Hypothesis Testing through the `L` argument (advanced users)

In this last example, we show how to use general linear hypothesis testing using contrast coding on a simulated dataset.

Contrast coding allows testing general linear hypotheses of the form:

$$L\beta = \Theta$$

Where $\beta$ is the $q \times p$ matrix of parameters estimated by the `mvgls` function, $L$ is a $k \times q$ matrix of full row rank, $rank(L) = k \leq q$, called the contrast coding matrix, and $\Theta$ is a $k \times p$ constant matrix, usually full of zeros (default in `manova.gls`), called the right-hand-side (`rhs`) matrix, that together specify the hypotheses to be tested using $k$ linear combinations of the parameters $\beta$ (see details in Fox 2015; and also Rencher 2002 - p. 180-183).

For instance, consider a multivariate dataset with $p$=5 variables and $q$=3 groups (on the 36 species tree simulated above):

```r
grp2 = rep(1:3, each=n/3)
pheno2 = trait + rep(c(0,0,10), each=n/3) # introduce some differences on the third group
data2 = list(pheno=pheno2, grp=as.factor(grp2))

# fit the model
fit4 <- mvgls(pheno~grp, data=data2, tree=tree, model="lambda")

# The overall MANOVA test:
(manova.gls(fit4, verbose=TRUE))
```

```
> Sequential MANOVA Tests with 999 permutations: Pillai test statistic
>     Test stat Pr(>Stat)
> grp    0.8118     0.001 **
> ---
> Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The test is significant, meaning that there are differences between (at least one of) the groups. We can assess more specific hypotheses. Recall that in this dataset we have introduced differences for the third group (+10). We can test if the third group and the second group have the same mean vector. This is done by testing a simple linear combination of the estimated parameters which can be expressed as:

$$0\beta_1 + 1\beta_2 - 1\beta_3 = \Theta$$

Where $\Theta = 0$. Indeed, we test that the difference between $\beta_2$ and $\beta_3$ is equal to 0 - i.e. we ask if they are the same. The contrast vector $L$ (or matrix) for this hypothesis is therefore:

$$L = [0, 1, -1]$$

We note that $\beta_1$ is not involved in this hypothesis and thus has a contrast code of `0`. Now this linear hypothesis can be evaluated with the `manova.gls` function:

```r
L1 <- matrix(c(0,1,-1), ncol=3) # Contrasts (vector or matrix)

# Test the first contrast:
(manova.gls(fit4, L=L1, nperm=999, verbose=TRUE))
```

```
> General Linear Hypothesis Test with 999 permutations: Pillai test statistic
>             Test stat Pr(>Stat)
> Contrasts L    0.7617     0.001 **
> ---
> Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

As expected the test is significant.

Let's test now if there are differences between the first ($\beta_1$) and the second ($\beta_2$) group by defining a new contrast matrix. Note that by default, the regression formula in **R** assumes that the first grouping factor (in alphabetical order) is the baseline group - called `intercept` in the model fit output - to which the others are compared (the other factors coefficients are estimated as deviance from the baseline; see `?contr.treatment`). This means that if we want to compare the two first groups, we need to test if $\beta_2$ is different from $\Theta = 0$ using the simple following contrasts:

$$L = [0, 1, 0]$$

```r
L2 <- matrix(c(0,1,0), ncol=3) # Contrasts (vector or matrix)

# Test the second contrast:
(manova.gls(fit4, L=L2, nperm=999, verbose=TRUE))
```

```
> General Linear Hypothesis Test with 999 permutations: Pillai test statistic
>             Test stat Pr(>Stat)
> Contrasts L   0.08306     0.681
> ---
> Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
# No differences as expected...
```

Note that we covered the three scenarios as the third one is implicit.

Alternatively, we can fit a model for which each group mean is explicitly estimated - rather than compared to a baseline - by specifying a factorial model without "intercept" using `+0` (or `-1`) in the regression formula:

```r
# fit the model without "intercept"
fit4bis <- mvgls(pheno ~ grp + 0, data=data2, tree=tree, model="lambda")
```

We can now test a similar hypothesis as above using the more intuitive linear relationship:

$$1\beta_1 - 1\beta_2 + 0\beta_3 = 0$$

which can be expressed by the contrasts matrix:

$$L = [1, -1, 0]$$

```r
L2bis <- matrix(c(1,-1,0), ncol=3) # Contrasts (vector or matrix)

# Test the second contrast:
(manova.gls(fit4bis, L=L2bis, nperm=999, verbose=TRUE))
```

```
> General Linear Hypothesis Test with 999 permutations: Pillai test statistic
>             Test stat Pr(>Stat)
> Contrasts L   0.08306     0.699
> ---
> Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
# The result is identical and the test shows no differences as expected...
```

Note that the slight difference in $p$-values between the two multivariate tests above is due to the randomness of the permutation scheme.

CAUTION: take care about global MANOVA tests on a regression model without an intercept (i.e. when using `-1` or `+0`) because the null hypothesis being tested might be different than for the regression model with the intercept term.

## 7.4 Testing hypothesis on the values of estimated parameters through the `rhs` argument (advanced users)

We can also test more specific hypotheses like "is the difference between the baseline $\beta_1$ and $\beta_3$ equal to 2?" For instance if we are interested in testing the estimated parameters against theoretical or published values. This is formally expressed by the following linear hypothesis:

$$0\beta_1 + 0\beta_2 + 1\beta_3 = \Theta = 2$$

To do that we can specify our own $\Theta$ matrix in the formula above through the `rhs` argument and use the appropriate contrast coding matrix `L`:

```
L3 <- matrix(c(0,0,1), ncol=3) # define coding
rhs <- 2   # is the difference between the first and the third group equal to 2?

# Test the first contrast:
(manova.gls(fit4, L=L3, rhs=rhs, nperm=999, verbose=TRUE))
```

```
> General Linear Hypothesis Test with 999 permutations: Pillai test statistic
>             Test stat Pr(>Stat)
> Contrasts L     0.534     0.001 **
> ---
> Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# indeed it is rejected (recall that beta3=10)... try with rhs=10 instead:

(manova.gls(fit4, L=L3, rhs=10, nperm=999, verbose=TRUE))
```

```
> General Linear Hypothesis Test with 999 permutations: Pillai test statistic
>             Test stat Pr(>Stat)
> Contrasts L    0.1616     0.314
> ---
> Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# it is not different from 10, as expected... Neat!
```

We can also use the model fit without intercept `fit4bis` to specifically test the values of the estimated mean for each of the groups using similar coding contrasts.

Note: if a single value is provided to `rhs`, then it is assumed that the $\Theta$ matrix contains only this value.

## 7.5 Phyllostomid bats example

Finally, we reuse here the bats dataset from Monteiro & Nogueira (2011) to illustrate how to perform multivariate analyses on an empirical dataset.

```
# Load the data
data <- get(data(phyllostomid))
# head(data$mandible) # the dataset is high-dimensional - i.e. p>(n-q)

# Fit the mandible data to the four diet and feeding modes schemes
fit_grp1 <- mvgls(mandible~grp1, data=data, data$tree, model="lambda", method="PL-LOOCV")
fit_grp2 <- mvgls(mandible~grp2, data=data, data$tree, model="lambda", method="PL-LOOCV")
fit_grp3 <- mvgls(mandible~grp3, data=data, data$tree, model="lambda", method="PL-LOOCV")
fit_grp4 <- mvgls(mandible~grp4, data=data, data$tree, model="lambda", method="PL-LOOCV")

# Then we perform the MANOVAs as in the paper
nbcores=4L # to speed up the calculations
(manova.gls(fit_grp1, test="Wilks", nbcores=nbcores, verbose=TRUE)) # Grouping 1
```

```
## Sequential MANOVA Tests with 999 permutations: Wilks test statistic
##      Test stat Pr(>Stat)
## grp1    0.1287     0.001 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
(manova.gls(fit_grp2, test="Wilks", nbcores=nbcores, verbose=TRUE)) # Grouping 2
```

```
## Sequential MANOVA Tests with 999 permutations: Wilks test statistic
##      Test stat Pr(>Stat)
## grp2  0.006345     0.001 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
(manova.gls(fit_grp3, test="Wilks", nbcores=nbcores, verbose=TRUE)) # Grouping 3
```

```
## Sequential MANOVA Tests with 999 permutations: Wilks test statistic
##      Test stat Pr(>Stat)
## grp3 0.0004306     0.001 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
(manova.gls(fit_grp4, test="Wilks", nbcores=nbcores, verbose=TRUE)) # Grouping 4
```

```
## Sequential MANOVA Tests with 999 permutations: Wilks test statistic
##      Test stat Pr(>Stat)
## grp4 1.768e-05     0.001 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

We can also perform the three linear hypotheses tests presented in the Supplementary Material using contrast coding through the `L` argument. The specific hypotheses being tested are: 1) difference between the sanguivorous and the nectivorous species, 2) between the frugivorous and the animalivorous, and 3) between the carnivorous and the insectivorous. These three tests can be expressed through the following linear hypotheses:

$$0\beta_{Carnivory} + 0\beta_{Frugivory} + 0\beta_{Insectivory} + 1\beta_{Nectarivory} - 1\beta_{Sanguivory} = \Theta_1 = 0$$

$$0.5\beta_{Carnivory} - 1\beta_{Frugivory} + 0.5\beta_{Insectivory} + 0\beta_{Nectarivory} - 0\beta_{Sanguivory} = \Theta_2 = 0$$

$$1\beta_{Carnivory} + 0\beta_{Frugivory} - 1\beta_{Insectivory} + 0\beta_{Nectarivory} - 0\beta_{Sanguivory} = \Theta_3 = 0$$

The respective contrast codes for these hypotheses are:

```
# First define the set of contrasts
L <- rbind(c(0,0,0,1,-1),      # Nectarivory vs Sanguivory
           c(0.5,-1,0.5,0,0),  # Frugivory vs. Animalivory
           c(1,0,-1,0,0))      # Insectivory vs. Carnivory
```

Note that R considers the factors in alphabetical order. As already noticed in the examples above, the first factor is recognized as an `intercept` by default in the regression formula. This implies that `Carnivory` is considered the baseline group to which the others are compared in the model `fit_grp4` above. Here, we specify a factorial model without "intercept" using `+0` in the regression formula to estimate the values for each group instead and make sensible linear hypotheses:

```
fit_glh <- mvgls(mandible~grp4+0, data=data, data$tree, model="lambda", method="PL-LOOCV")

L1 <- L[1,,drop=FALSE]  # Nectarivory vs Sanguivory
(manova.gls(fit_glh,
                    nbcores=nbcores, test="Wilks", L=L1, verbose=TRUE))
```

```
## General Linear Hypothesis Test with 999 permutations: Wilks test statistic
##             Test stat Pr(>Stat)
## Contrasts L   0.02798     0.001 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
L2 <- L[2,,drop=FALSE]  # Frugivory vs. Animalivory
(manova.gls(fit_glh,
                    nbcores=nbcores, test="Wilks", L=L2, verbose=TRUE))
```

```
## General Linear Hypothesis Test with 999 permutations: Wilks test statistic
##             Test stat Pr(>Stat)
## Contrasts L    0.0521     0.001 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
L3 <- L[3,,drop=FALSE]  # Insectivory vs. Carnivory
(manova.gls(fit_glh,
                    nbcores=nbcores, test="Wilks", L=L3, verbose=TRUE))
```

```
## General Linear Hypothesis Test with 999 permutations: Wilks test statistic
##             Test stat Pr(>Stat)
## Contrasts L    0.1331     0.005 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

# 8 References


Clavel J., Aristide L., Morlon H. 2019. A Penalized Likelihood Framework for High-Dimensional Phylogenetic Comparative Methods and an Application to New-World Monkeys Brain Evolution. *Systematic Biology*. 68:93-116.

Fox J. 2015. Applied Regression Analysis and Generalized Linear Models. *SAGE Publications*.

Langsrud O. 2003. Anova for unbalanced data: use type II instead of type III sums of squares. *Statistics and Computing*. 13:163-167.

McFarquhar M. 2016. Testable hypotheses for unbalanced neuroimaging data. *Frontiers in Neuroscience*. 10:1-13.

Monteiro L.R., Nogueira M.R. 2011. Evolutionary patterns and processes in the radiation of phyllostomid bats. *BMC Evolutionary Biology*. 11:1-23.

Olson C.L. 1974. Comparative robustness of six tests in multivariate analysis of variance. *Journal of the American Statistical Association*. 69:894-908.

Rencher A.C. 2002. Methods of Multivariate Analysis. *John Wiley & Sons*.