



# Анализ избыточности (Redundancy analysis, RDA)

Анализ и визуализация многомерных данных с  
использованием R

Марина Варфоломеева, Вадим Хайтов

# Анализ избыточности (Redundancy analysis, RDA)

- Связь нескольких наборов переменных
- Анализ избыточности, теория и практика
- Проверка значимости ординации
- Выбор оптимальной модели
- Частный анализ избыточности и компоненты объясненной инерции
- Компоненты объясненной изменчивости

## Вы сможете

- Проводить анализ избыточности
- Оценивать долю объясненной инерции
- Интерпретировать компоненты по нагрузкам переменных
- Строить ординацию объектов в пространстве компонент
- Проверять значимость модели ординации при помощи пермутационного теста
- Разделять объясненную инерцию на компоненты, связанные с разными наборами переменных, при помощи частного анализа избыточности

# **Связь нескольких наборов переменных**

## Пример: генетика бабочек *Euphydryas editha*

Частоты разных аллелей фосфоглюкоизомеразы и данные о факторах среды для 16 колоний бабочек *Euphydryas editha* в Калифорнии и Орегоне

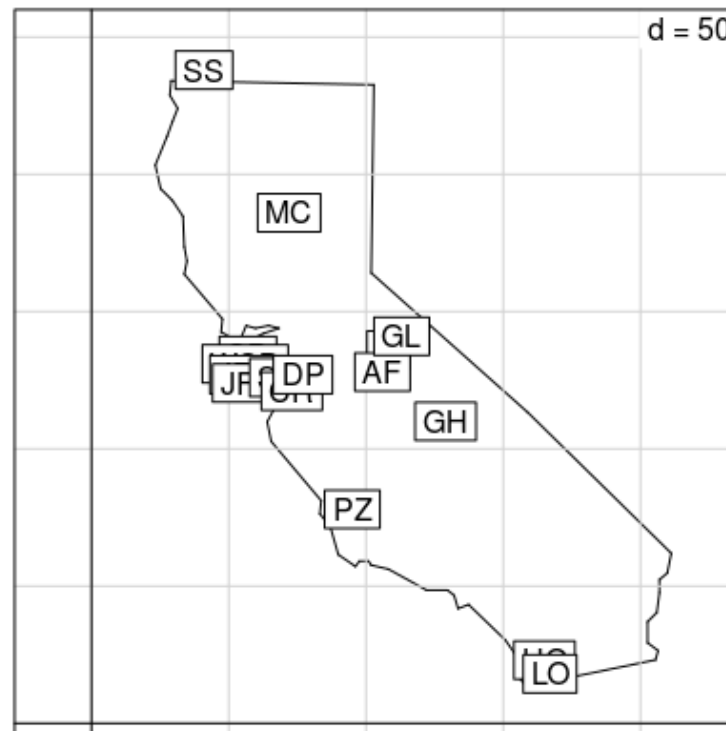


*Euphydryas editha*

Winged Wonder by [Roger Lynn on Flickr](#)

## Как вы думаете, что будет определять генетическую структуру в колониях бабочек?

```
library(ade4)
data(butterfly)
# расположение сайтов
s.label(butterfly$xy, contour = butterfly$contour, inc = FALSE)
```



# Структура данных

- \$xy - координаты колоний
- \$envir - 4 фактора среды для колоний
- \$genet - частоты 6 аллелей в колониях
- \$contour - карта Калифорнии

```
str(butterfly, max.level = 2, give.attr = FALSE, vec.len = 2)
```

```
## List of 5
## $ xy      : 'data.frame': 16 obs. of 2 variables:
## ..$ x: num [1:16] 41 57 56 57 58 ...
## ..$ y: num [1:16] 238 134 131 127 124 ...
## $ envir   : 'data.frame': 16 obs. of 4 variables:
## ..$ Altitude : num [1:16] 500 800 570 550 550 ...
## ..$ Precipitation: num [1:16] 43 20 28 28 28 ...
## ..$ Temp_Max : num [1:16] 98 92 98 98 98 ...
## ..$ Temp_Min : num [1:16] 17 32 26 26 26 ...
## $ genet   : 'data.frame': 16 obs. of 6 variables:
## ..$ 0.4 : num [1:16] 0 0 0 0 0 ...
## ..$ 0.6 : num [1:16] 3 16 6 4 1 ...
## ..$ 0.8 : num [1:16] 22 20 28 19 8 ...
## ..$ 1 : num [1:16] 57 38 46 47 50 ...
## ..$ 1.16: num [1:16] 17 13 17 27 35 ...
## ..$ 1.3 : num [1:16] 1 13 3 3 6 ...
## $ contour: 'data.frame': 59 obs. of 4 variables:
## ..$ x1: num [1:59] 29 103 ...
## ..$ y1: num [1:59] 234 233 ...
## ..$ x2: num [1:59] 103 102 ...
## ..$ y2: num [1:59] 233 164 ...
## $ Spatial:
```

## Создадим переменные с более короткими названиями для удобства

```
# частоты аллелей  
gen <- butterfly$genet  
head(gen, 3)
```

```
##      0.4 0.6 0.8 1 1.16 1.3  
## SS    0  3  22 57  17  1  
## SB    0 16  20 38  13 13  
## WSB   0  6  28 46  17  3
```

```
# переменные среды и географические координаты  
env_geo <- cbind(butterfly$envir, butterfly$xy)  
head(env_geo, 3)
```

```
##      Altitude Precipitation Temp_Max Temp_Min x  y  
## SS          500           43      98      17 41 238  
## SB          800           20      92      32 57 134  
## WSB         570           28      98      26 56 131
```

# **Анализ избыточности, теория и практика**



# Анализ избыточности (RDA, Redundancy analysis)

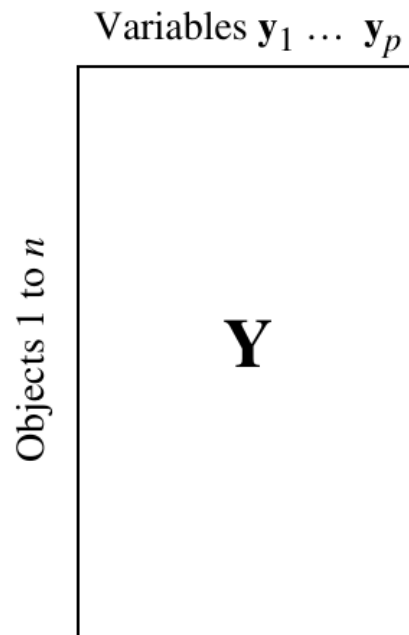
- Метод прямой ординации (ограниченной ординации = constrained ordination)
- Основан на анализе главных компонент
- Нужно две матрицы данных: матрица зависимых переменных и матрица предикторов
- Нужно найти такие компоненты матрицы зависимых переменных, которые являются линейными комбинациями предикторов и отражают максимум изменчивости.

## Условия применимости RDA

- Линейная зависимость переменных-откликов от предикторов
- Независимость наблюдений (сайтов)
- Сайтов должно быть значительно больше, чем предикторов (как в регрессии, см. проклятие размерности)

# Непрямая ординация и множественная регрессия

(a) Simple ordination of matrix **Y**:  
principal comp. analysis (PCA)  
correspondence analysis (CA)



(b) Ordination of **y** (single axis) under  
constraint of **X**: multiple regression

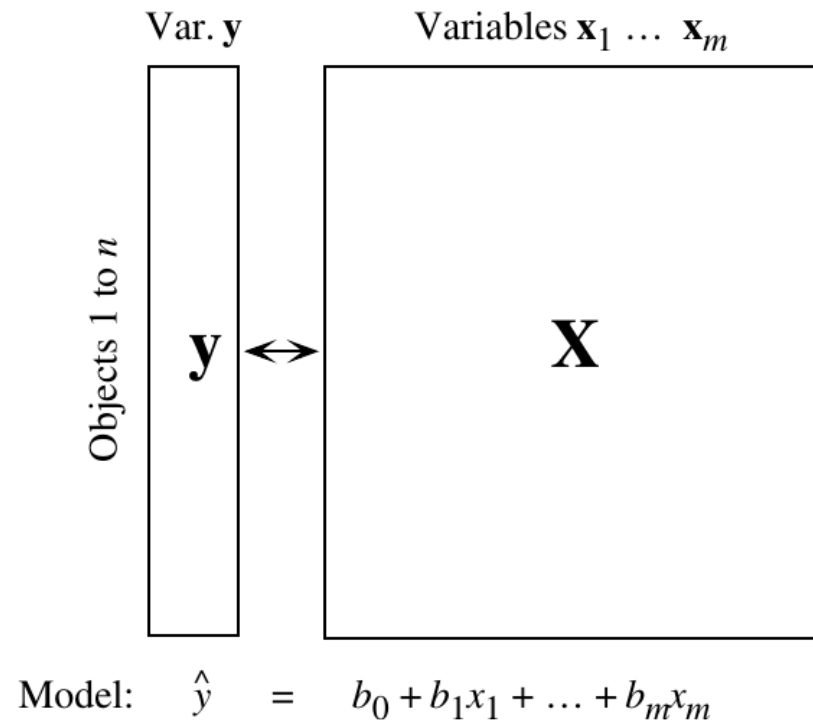


Рис. из Legendre, Legendre, 1998

## Прямая (ограниченная) ординация

- (c) Ordination of **Y** under constraint of **X**:  
redundancy analysis (RDA)  
canonical correspondence analysis (CCA)

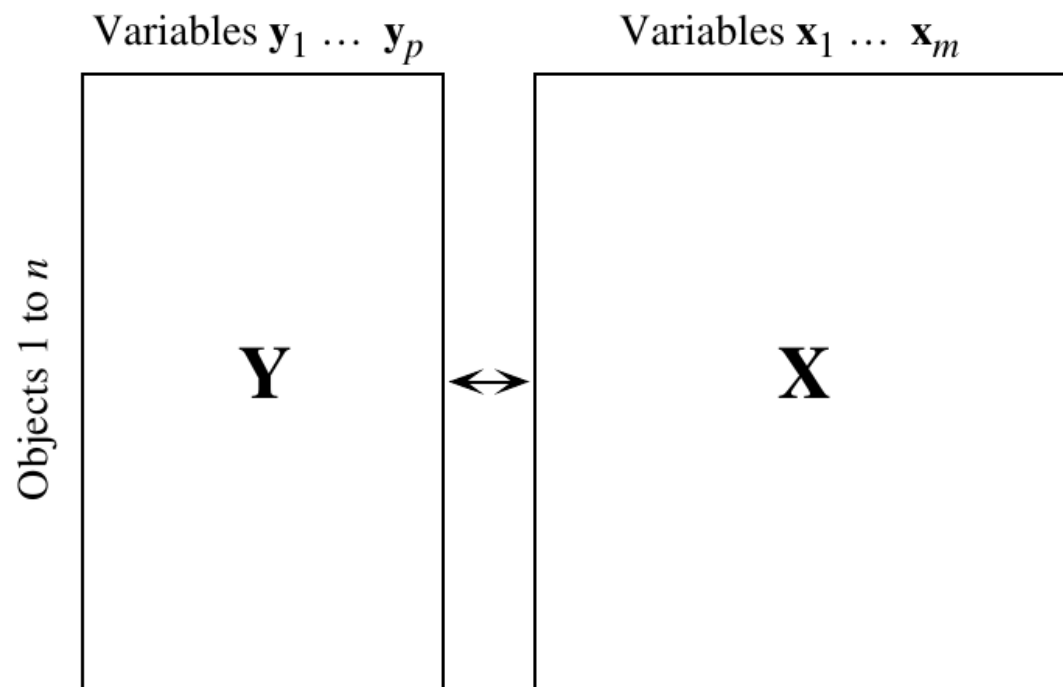


Рис. из Legendre, Legendre, 1998

## Если множественную линейную регрессию можно обобщить для нескольких зависимых переменных — получится RDA

$$\hat{y}_i = b_0 + b_1 x_{1i} + b_2 x_{2i} + \dots + b_k + \epsilon_i$$

Уравнение множественной линейной регрессии можно переписать в виде матриц.

$$\begin{bmatrix} \hat{y}_1 \\ \hat{y}_2 \\ \vdots \\ \hat{y}_n \end{bmatrix} = \begin{bmatrix} 1 & x_{1,1} & x_{1,2} & \cdots & x_{1,k} \\ 1 & x_{2,1} & x_{2,2} & \cdots & x_{2,k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n,1} & x_{n,2} & \cdots & x_{n,k} \end{bmatrix} \cdot \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ \vdots \\ b_k \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{bmatrix}$$

Сокращенная форма записи:  $\hat{\mathbf{y}} = \mathbf{X}\mathbf{b} + \epsilon$ , причем  $\mathbf{b} = (\mathbf{X}^T \mathbf{X})^{-1} (\mathbf{X}^T \mathbf{y})$ .

Во множественной линейной регрессии  $\hat{\mathbf{y}} = \mathbf{X}(\mathbf{X}^T \mathbf{X})^{-1} (\mathbf{X}^T \mathbf{y})$

В RDA зависимая переменная — матрица, т.е.  $\hat{\mathbf{Y}} = \mathbf{X}(\mathbf{X}^T \mathbf{X})^{-1} (\mathbf{X}^T \mathbf{Y})$ ,

# Принцип RDA

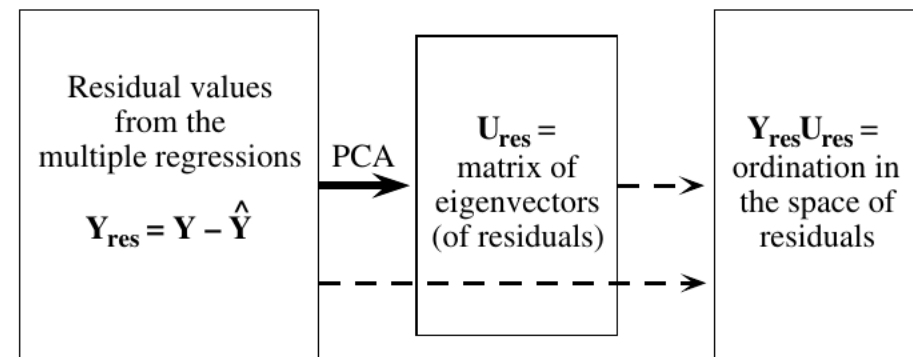
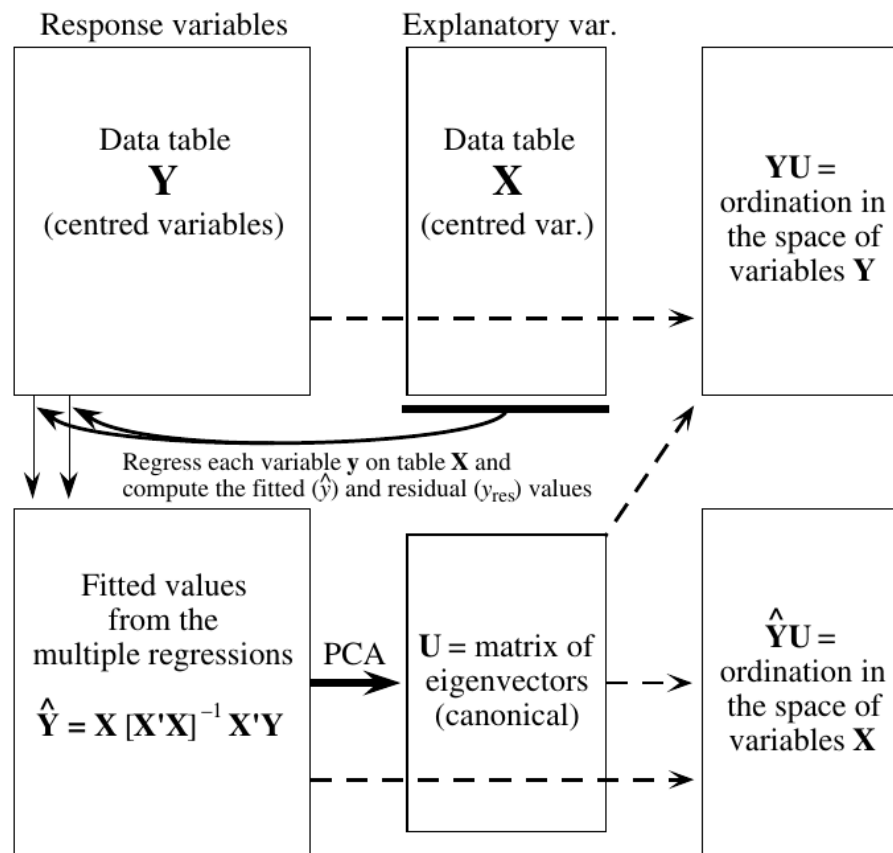


Рис. из Legendre, Legendre, 1998

**RDA <sub>B</sub> R**

## RDA в vegan

- Зависимые переменные (отклики) - генетические данные
- Независимые переменные (предикторы) - переменные среды

```
library(vegan)
bf_rda <- rda(gen ~ Altitude + Precipitation + Temp_Max + Temp_Min, data = env_geo)
summary(bf_rda)
```

```
##
## Call:
## rda(formula = gen ~ Altitude + Precipitation + Temp_Max + Temp_Min,      data = env_geo)
##
## Partitioning of variance:
##              Inertia Proportion
## Total              730      1.000
## Constrained        380      0.521
## Unconstrained      350      0.479
##
## Eigenvalues, and their contribution to the variance
##
## Importance of components:
##              RDA1      RDA2      RDA3      RDA4      PC1      PC2
## Eigenvalue      357.20  19.6391  2.30457  0.75999  211.979  101.540
## Proportion Explained  0.49  0.0269  0.00316  0.00104   0.291   0.139
## Cumulative Proportion  0.49  0.5165  0.51966  0.52070   0.811   0.950
##              PC3      PC4      PC5
## Eigenvalue      27.1344  7.9192  1.12065
## Proportion Explained  0.0372  0.0109  0.00154
## Cumulative Proportion  0.9876  0.9985  1.00000
##
## Accumulated constrained eigenvalues
## Importance of components:
##              RDA1      RDA2      RDA3      RDA4
```

## Структура общей изменчивости

О структуре изменчивости можно судить по суммам собственных чисел ординационных осей (ограниченных и неограниченных)

## Partitioning of variance:

##	Inertia	Proportion
## Total	730	1.000
## Constrained	380	0.521
## Unconstrained	350	0.479

- Total - всех осей - общая изменчивость исходной матрицы откликов (генетич. структуры в разных сайтах)
- Constrained - осей, кот. являются комбинациями факторов среды - изменчивость объясненная средой
- Unconstrained - необъясненная изменчивость



## Важность различных компонент

Можно более подробно оценить, как распределяется изменчивость между осями

```
## Eigenvalues, and their contribution to the variance
```

```
## Importance of components:
```

```
##          RDA1      RDA2      RDA3      RDA4      PC1      PC2
## Eigenvalue    357.20  19.6391  2.30457  0.75999  211.979  101.540
## Proportion Explained    0.49   0.0269  0.00316  0.00104    0.291    0.139
## Cumulative Proportion    0.49   0.5165  0.51966  0.52070    0.811    0.950
##          PC3      PC4      PC5
## Eigenvalue    27.1344  7.9192  1.12065
## Proportion Explained    0.0372  0.0109  0.00154
## Cumulative Proportion    0.9876  0.9985  1.00000
```

- Много изменчивости объяснено, но много осталось необъясненной. Первые две ограниченных оси объясняют 51% изменчивости, но первые две неограниченных объясняют еще 43%

## Распределение изменчивости, потенциально объяснимой факторами

## Accumulated constrained eigenvalues

## Importance of components:

##	RDA1	RDA2	RDA3	RDA4
## Eigenvalue	357.20	19.6391	2.30457	0.760
## Proportion Explained	0.94	0.0517	0.00607	0.002
## Cumulative Proportion	0.94	0.9919	0.99800	1.000

- Первая ограниченная ось объясняет большую часть потенциально объяснимой изменчивости. Остальные оси почти ничего не объясняют.

## Собственные векторы, нагрузки переменных = "species scores"

```
scores(bf_rda, display = "species", choices = 1:5)
```

```
##          RDA1      RDA2      RDA3      RDA4      PC1
## 0.4    0.637 -0.6604  0.19728 -0.2423  0.968
## 0.6    0.894 -0.7525 -0.00662  0.1207  2.244
## 0.8    2.618 -0.0949  0.14157  0.1699  2.768
## 1      -6.350  0.1208  0.05462  0.0470 -3.337
## 1.16   1.554  1.3370  0.11772 -0.0419 -2.355
## 1.3    0.647  0.0501 -0.50457 -0.0534 -0.289
## attr(,"const")
## [1] 10.2
```

.

## Корреляции между откликами и предикторами

- Сильная корреляция между генетической структурой и средой только для первой ограниченной оси. Для других - умеренные или слабые.

```
spenvcor(bf_rda)
```

```
## RDA1 RDA2 RDA3 RDA4  
## 0.836 0.353 0.332 0.183
```

## Визуализация ординации

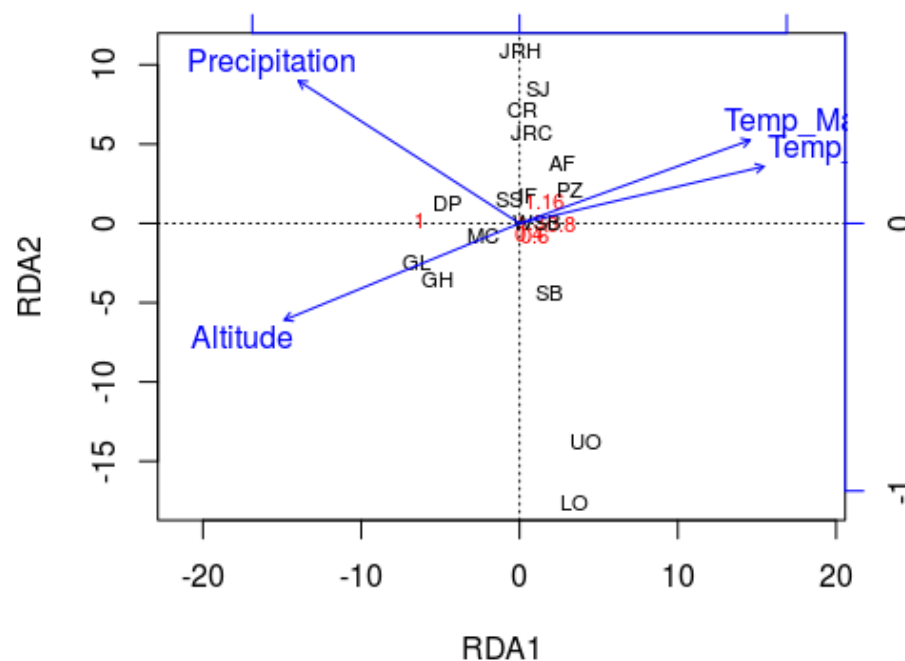
- Какие предикторы важнее всего?
- Какими факторами определяется значение зависимых переменных?

Триплоты:- переменные-отклики ("species"),- объекты ("sites")- переменные-предикторы (непрерывные в виде векторов, дискретные в виде центроидов)

Биплоты:- отклики + предикторы- объекты + предикторы

## Триплот корреляций (scaling = 2): Какие переменные среды важнее всего?

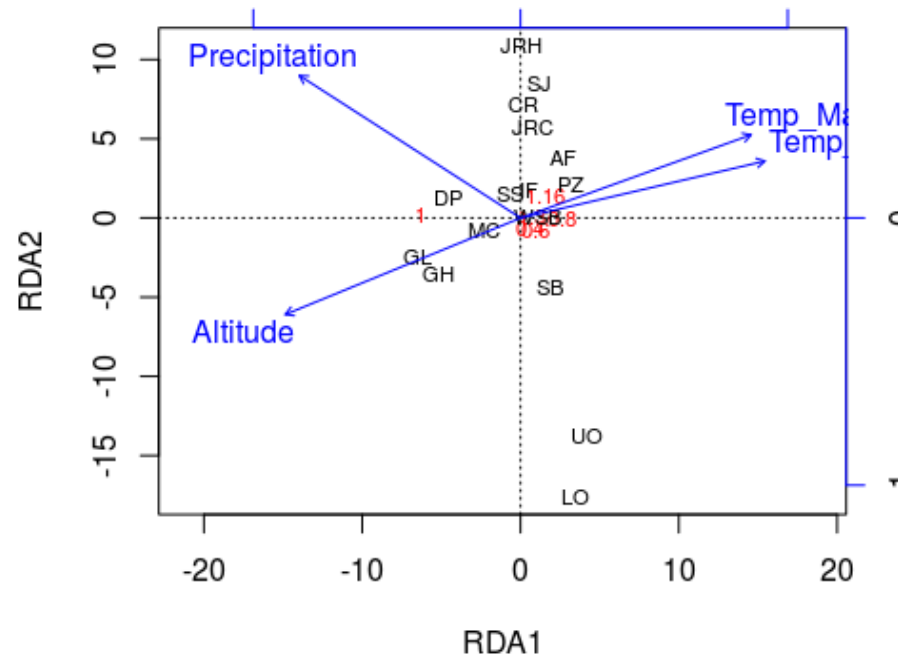
```
plot(bf_rda, scaling = 2)
```



- Векторы - независимые переменные, факторы среды
- Надписи - объекты (сайты, особи, популяции и пр.)

# Пример интерпретации триплота корреляций

```
plot(bf_rda, scaling = 2)
```

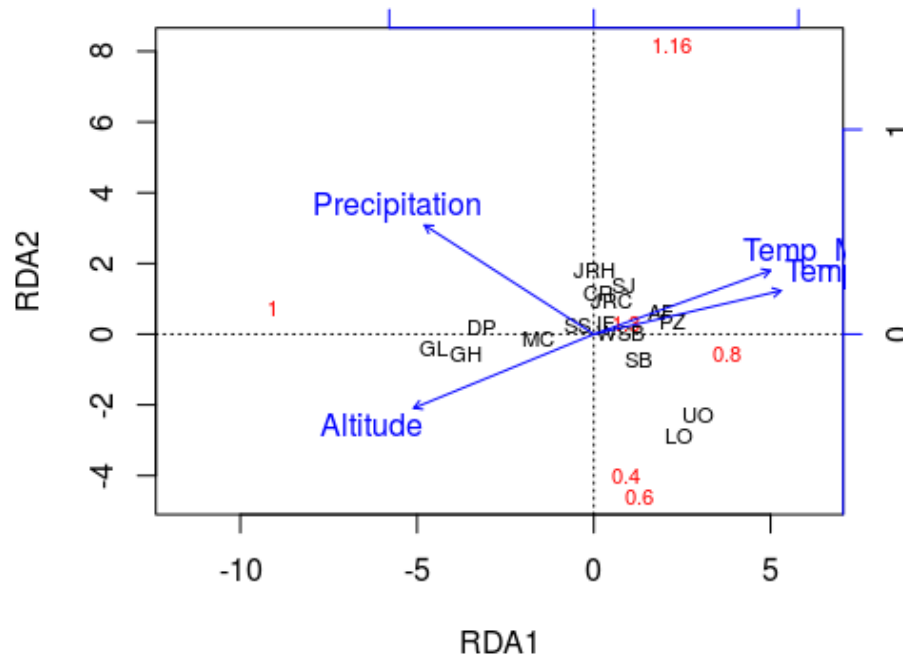


- Вдоль первой оси изменяется температура, высота и осадки
- Вдоль второй оси - немного меняется уровень осадков

# Триплот расстояний (scaling = 1)

```
plot(bf_rda, scaling = 1)
```

- Надписи - объекты (сайты, особи, популяции и пр.)
- Красные надписи - зависимые

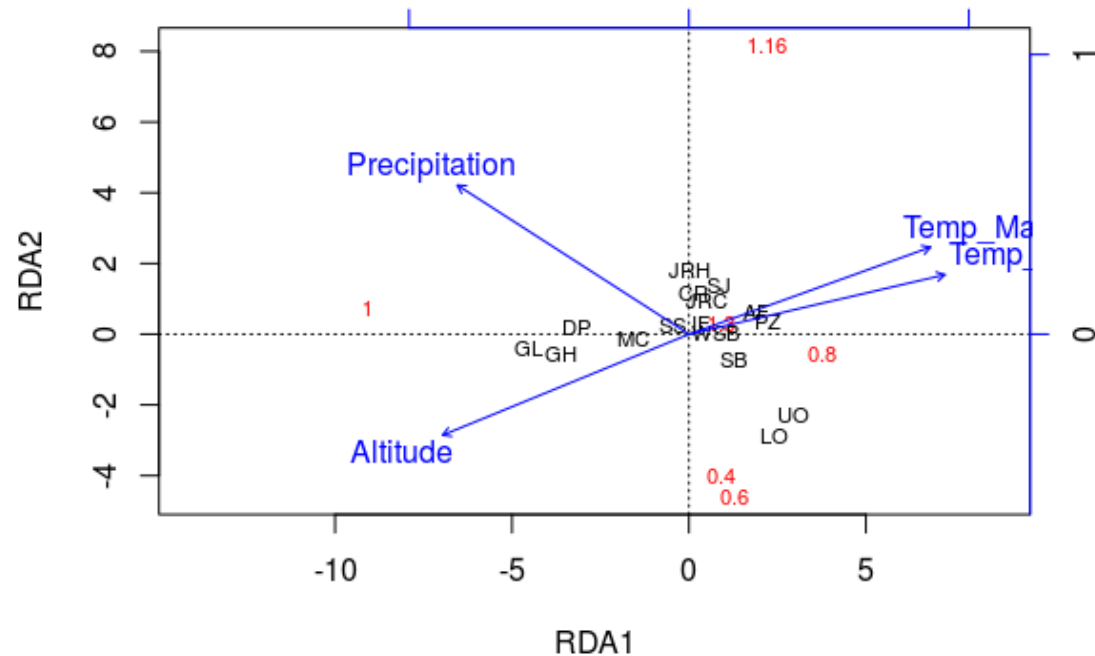


```
# s.label(butterfly$xy, contour =  
butterfly$contour, inc = FALSE) # карта
```



# Пример интерпретации триплота расстояний

```
plot(bf_rda, scaling = 1)
```



- Генетическая структура в LO и UO похожа, но не похожа на остальные места
- GL и GH - более высокогорные сайты, чем LO и UO

# Проверка значимости ординации

## Общий тест на значимость ординации

- тестируем гипотезу о том, что отношения между генотипом и средой значимы.

$H_0$ : значения предикторов в пробах не зависят от переменных среды (генетическая структура не зависит от среды)

- основан на пермутациях: проверяем, насколько наблюдаемая связь сильнее, чем если случайно переставить данные
- статистика - сумма всех соб. чисел ограниченных осей

# Общий тест: Влияют ли факторы на зависимые переменные?

Есть ли связь генетики со средой?

```
anova(bf_rda)
```

```
## Permutation test for rda under reduced model
## Permutation: free
## Number of permutations: 999
##
## Model: rda(formula = gen ~ Altitude + Precipitation + Temp_Max + Temp_Min, data =
env_geo)
##           Df Variance      F Pr(>F)
## Model      4      380  2.99  0.025 *
## Residual  11      350
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

- связь генетической структуры и среды значима

# Тест факторов, type I эффекты: Какие факторы влияют на зависимые переменные?

```
anova(bf_rda, by = "term")
```

```
## Permutation test for rda under reduced model
## Terms added sequentially (first to last)
## Permutation: free
## Number of permutations: 999
##
## Model: rda(formula = gen ~ Altitude + Precipitation + Temp_Max + Temp_Min, data =
env_geo)
##           Df Variance      F Pr(>F)
## Altitude    1      279 8.79  0.007 **
## Precipitation 1       72 2.28  0.120
## Temp_Max     1       20 0.62  0.536
## Temp_Min     1        8 0.26  0.806
## Residual    11      350
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

- Генетическая структура популяций бабочек достоверно зависит от высоты, если в модель включены др. факторы.
- Но это Type I эффекты - они зависят от порядка включения факторов в модель. Т.е. после включения высоты в модель другие факторы уже не влияют.

## Тест факторов, type III эффекты: Какие факторы влияют на зависимые переменные?

```
anova(bf_rda, by = "mar")
```

```
## Permutation test for rda under reduced model
## Marginal effects of terms
## Permutation: free
## Number of permutations: 999
##
## Model: rda(formula = gen ~ Altitude + Precipitation + Temp_Max + Temp_Min, data =
env_geo)
##           Df Variance      F Pr(>F)
## Altitude    1         3 0.08  0.95
## Precipitation 1        34 1.08  0.32
## Temp_Max     1        25 0.79  0.46
## Temp_Min     1         8 0.26  0.78
## Residual    11       350
```

- Если протестировать каждый из факторов отдельно, при условии, что все остальные включены в модель, то получится, что ни один из них не влияет.

## Тест значимости осей, ограниченных факторами:

- $H_0$ : значения переменных-откликов для объектов не зависят от переменных-предикторов
- пермутационный: выбирает оси, которые объясняют больше изменчивости, чем из др. матриц, полученных путем перестановок

## Тест значимости осей, ограниченных факторами: Вдоль какой из осей значимо меняется генетическая структура?

```
anova(bf_rda, by = "axis")

## Permutation test for rda under reduced model
## Marginal tests for axes
## Permutation: free
## Number of permutations: 999
##
## Model: rda(formula = gen ~ Altitude + Precipitation + Temp_Max + Temp_Min, data =
env_geo)
##      Df Variance      F Pr(>F)
## RDA1    1      357 11.24  0.001 ***
## RDA2    1       20  0.62  0.540
## RDA3    1        2  0.07  0.962
## RDA4    1         1  0.02  0.999
## Residual 11      350
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

- Генетическая структура значимо меняется вдоль первой главной оси



# Выбор оптимальной модели

## Выбор оптимальной модели

У нас проблема. Если мы тестируем любой из факторов, после включения остальных в модель - он не влияет. Это значит, что модель не оптимальна.

Как подобрать оптимальную модель?

- Можно использовать пошаговый выбор модели: добавляем в модель лучшие переменные и снова исключаем те, что потеряли значимость. (Вспомните, как это было для регрессионных моделей.)

Какой можно использовать тест для сравнения моделей?

- Модели с разным числом предикторов можно сравнить при помощи пермутационного теста (AIC для ограниченных ординаций не существует!)
- **Осторожно!** В `vegan` факторы включенные в модель обозначаются "-", а факторы, исключенные из модели - "+"

# Пошаговый выбор оптимальной модели

Для пошагового выбора нам понадобятся полная и нулевая модели

```
m1 <- rda(gen ~ Altitude + Precipitation + Temp_Max + Temp_Min, data = env_geo)
m0 <- rda(gen ~ 1, data = env_geo)
```

Запускаем пошаговый выбор

```
m <- ordistep(m0, scope = formula(m1))
```

```
##
## Start: gen ~ 1
##
##           Df AIC      F Pr(>F)
## + Temp_Min    1 100  9.75  0.005 **
## + Temp_Max    1 101  8.12  0.005 **
## + Precipitation 1 102  7.30  0.005 **
## + Altitude    1 101  8.69  0.010 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Step: gen ~ Temp_Min
##
##           Df AIC      F Pr(>F)
## - Temp_Min  1 106  9.75  0.005 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##           Df AIC      F Pr(>F)
## + Precipitation 1 100  1.41  0.21
## + Temp_Max      1 101  1.08  0.33
## + Altitude      1 102  0.23  0.89
```

## Оптимальная модель, отобранная при помощи пошагового алгоритма

m\$anova

```
##           Df AIC      F Pr(>F)
## + Temp_Min  1 100  9.75  0.005 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

- Оптимальная модель содержит только один предиктор - минимальную температуру

# Частный анализ избыточности и компоненты объясненной инерции

## Зачем нужен частный анализ избыточности?

Мы уже обнаружили тесную связь генотипов со средой, и даже знаем, с какими переменными.

Но генотипы в близких местах могут быть похожи по разным причинам.

И теперь у нас два вопроса:

### Кто виноват?

- сходный климат в близких локациях
- поток генов между близкими колониями облегчен

### Что делать?

- Нужно удалить влияние географического положения, чтобы сделать корректный вывод о связи генотипов со средой.

## Частный анализ избыточности

- зависимость от одного набора переменных (предикторов), когда влияние другого (ковариат) исключено.

Техника:

1. Множественная регрессия зависимости предикторов от ковариат.
2. Остатки от этой регрессии — это то, что от ковариат не зависит — можно использовать в РСА в качестве предикторов (вместо исходных переменных среды).

## Делаем частный RDA: зависимость генетической структуры от среды с учетом географического положения

```
bf_prda_1 <- rda(gen ~ Temp_Min + Condition(x + y), data = env_geo)
anova(bf_prda_1) ## Пермутационный тест
```

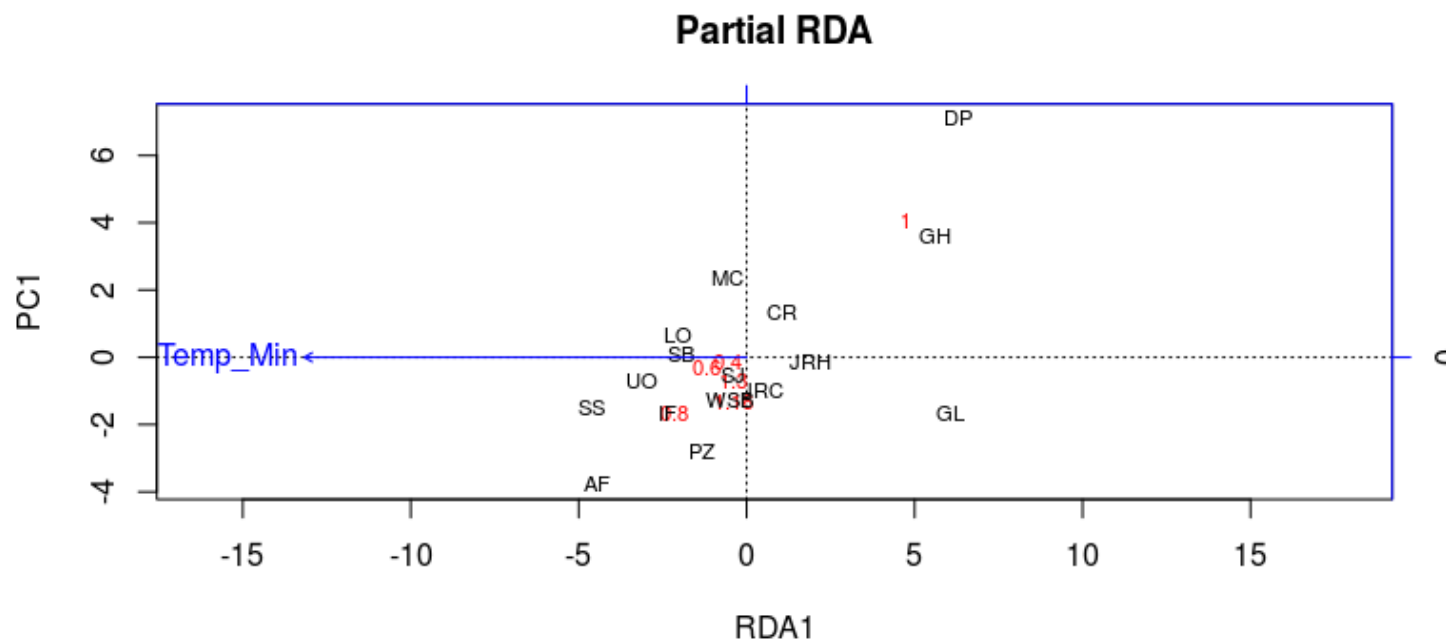
```
## Permutation test for rda under reduced model
## Permutation: free
## Number of permutations: 999
##
## Model: rda(formula = gen ~ Temp_Min + Condition(x + y), data = env_geo)
##           Df Variance      F Pr(>F)
## Model      1      202 9.01  0.001 ***
## Residual  12      268
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

- Климат объясняет генетическую изменчивость, даже после удаления влияния географических координат



# График ординации

```
plot(bf_prda_1, main = "Partial RDA")
```



- Смысл графика остался прежним, изменились нюансы
- Первая ось - мин температура.

# **Компоненты объясненной изменчивости**

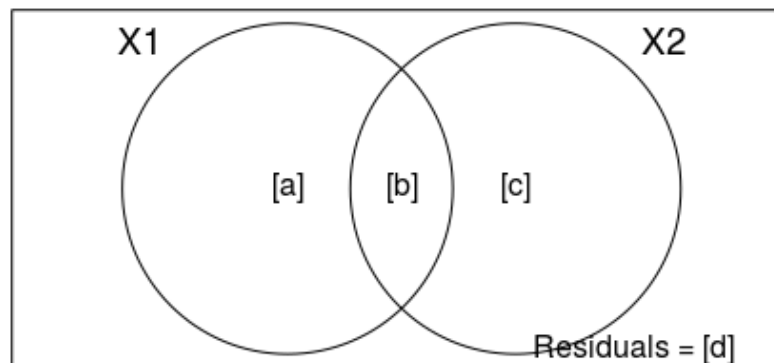
## Компоненты изменчивости

К этому моменту мы знаем, что климат объясняет генетическую изменчивость, даже после удаления влияния географических координат.

Но какая часть изменчивости генетической структуры объясняется в чистом виде географической близостью, а какая — общим действием климата и географии?

## Общую изменчивость делим на части:

showvarparts(2)



1.  $a + b + c$  - вся потенциально объяснимая средой и географией изменчивость
2.  $a$  - изменчивость, объясненная климатом
3.  $c$  - изменчивость, объясненная географией
4.  $b$  - изменчивость, совместно объясненная средой и географией

## Чтобы выделить компоненты изменчивости нам нужно несколько элементов

У нас уже есть частный RDA № 1: зависимость генетики от среды с учетом географии (для а)

Нам нужен частный RDA №2: генетика от географии с учетом свойств среды (для с)

```
bf_prda_2 <- rda(gen ~ x + y + Condition(Temp_Min), data = env_geo)
```

И полная модель RDA генетика от среды и географического положения (для а + b + с)

```
bf_rda_full <- rda(gen ~ x + y + Temp_Min, data = env_geo)
```

## Задание: Найдите компоненты инерции

1. изменчивость, потенциально объяснимую средой и географией
2. изменчивость, связанную только со средой, но не с географией
3. изменчивость, связанную только с географией, но не со средой
4. изменчивость, объясненную одновременно средой и географией

### Подсказка

Смотрите на результаты разных RDA

## Решение: 1) Сколько изменчивости потенциально объясняется средой и географией?

```
sum_full <- summary(bf_rda_full)
```

```
> sum_full
```

```
## Partitioning of variance:
```

```
##           Inertia Proportion
## Total           730      1.000
## Constrained      461      0.632
## Unconstrained     268      0.368
```

Изменчивость, объясненная вместе средой и географией, здесь достаточно велика — 461.24

```
(I_total <- sum_full$constr.chi)
```

```
## [1] 461
```

В отличие от нее, доля изменчивости, объясненной ограниченной матрицей, может быть довольно малой по отношению к общей изменчивости. Некоторые советуют сосредоточиться на доле от потенциально объяснимой изменчивости (от `sum_full$constr.chi`)

## Решение: 2) Изменчивость, объясненная климатом

```
sum_prda_1 <- summary(bf_prda_1)
```

```
> sum_prda_1
```

```
partit(sum_prda_1)
```

```
## Partitioning of variance:
```

##	Inertia	Proportion
## Total	730	1.000
## Conditioned	260	0.356
## Constrained	202	0.276
## Unconstrained	268	0.368

- Среда без географии объясняет 201.5

```
(I_env <- sum_prda_1$constr.chi)
```

```
## [1] 201
```



## Решение 3) Изменчивость, объясненная географией

```
sum_prda_2 <- summary(bf_prda_2)
```

```
> sum_prda_2
```

```
partit(sum_prda_2)
```

```
## Partitioning of variance:
```

##	Inertia	Proportion
## Total	730	1.000
## Conditioned	299	0.410
## Constrained	162	0.222
## Unconstrained	268	0.368

География без среды объясняет 161.79

```
(I_geo <- sum_prda_2$constr.chi)
```

```
## [1] 162
```

## Решение: 4) Изменчивость, совместно объясненная средой и географией

```
(I_env_geo <- I_total - I_env - I_geo)
```

```
## [1] 98
```

## Компоненты изменчивости - сводим результаты вместе

```
comp <- data.frame(Inertia = c(I_env, I_geo, I_env_geo, I_total))
rownames(comp) <- c('Только среда', 'Только география', 'Среда и география вместе', 'Общая объяснимая инерция')
comp$Proportion <- comp$Inertia/sum(comp$Inertia[1:3]) * 100
colnames(comp) <- c('Инерция', '%')
comp
```

##	Инерция	%
## Только среда	201	43.7
## Только география	162	35.1
## Среда и география вместе	98	21.2
## Общая объяснимая инерция	461	100.0

Среда объясняет 50% общей изменчивости генетической структуры - очень много, но и география объясняет 30%. И только 21% объясняется совместным влиянием среды и географии

## Take home messages

- Анализ избыточности помогает установить связь между несколькими наборами переменных. Один из наборов считается зависимым, другой считается объясняющим
- Для анализа необходимо, чтобы зависимости переменных-откликов от предикторов были линейными
- В ходе анализа выделяют два типа осей - ограниченные (объясненные) переменными-предикторами, и неограниченные (необъясненные) ими
- Частный анализ избыточности позволяет описать зависимость двух наборов переменных с поправкой на влияние дополнительных переменных (ковариат)
- При помощи частного анализа избыточности можно выделить компоненты изменчивости связанные с несколькими (2-4) наборами переменных-предикторов

## Дополнительные ресурсы

- Borcard, D., Gillet, F., Legendre, P., 2011. Numerical ecology with R. Springer.
- Legendre, P., Legendre, L., 2012. Numerical ecology. Elsevier.
- Oksanen, J., 2011. Multivariate analysis of ecological communities in R: vegan tutorial. R package version 2–0.
- The Ordination Web Page URL <http://ordination.okstate.edu/> (accessed 10.21.13).
- Quinn, G.G.P., Keough, M.J., 2002. Experimental design and data analysis for biologists. Cambridge University Press.