

Ordinačné metódy

Danka Némethová

28. – 30. 5. 2007

Úvod

- ◆ **Mnohorozmerné metódy:**

názov „mnohorozmerné“ – dáta sú tvorené objektami (vzorky, lokality), každý z nich je charakterizovaný viacerými parametrami (druhmi)

každý z týchto parametrov môžeme považovať za jeden rozmer objektu (vzorky)

DÁTOVA MATICA

	druh 1	druh 2	druh 3
vzorka 1			
vzorka 2			
vzorka 3			
vzorka 4			
vzorka 5			
vzorka 6			

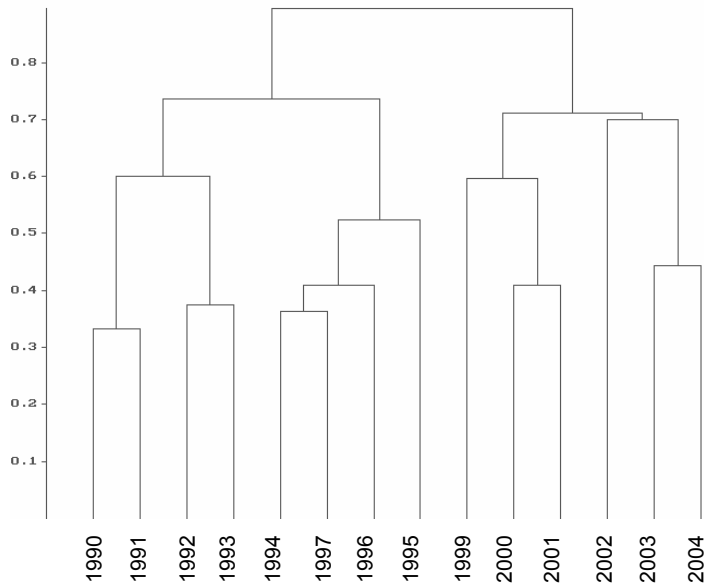
Hodnoty pre druhy (presencia/absencia;
abundancia; dominancia) pre každú vzorku

Ordinácia a zhluková analýza sú jediné možné techniky, ktoré môžeme použiť bez nameraných charakteristík prostredia.

Úvod

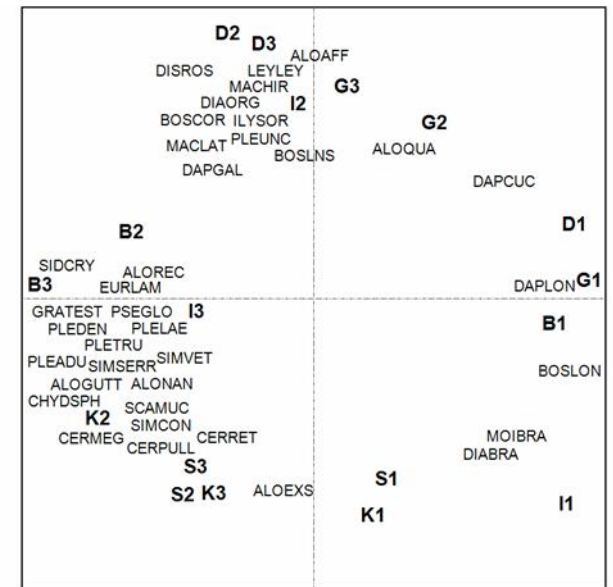
ZHLUKOVÁ ANALÝZA

- ◆ Klasifikuje vzorky (lokality), druhy alebo premenné
- ◆ Nachádza skupiny v dátach



ORDINÁCIA

- ◆ Usporiadáva objekty pozdĺž trendu v dátach

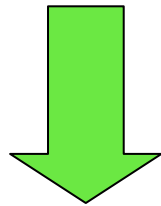


Úvod

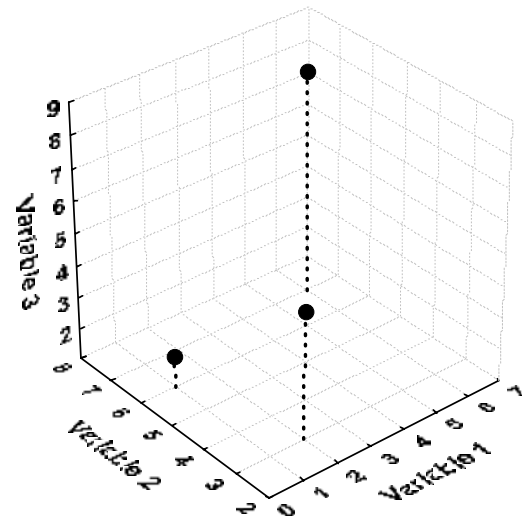
Objekty charakterizované p parametrami je možné si predstaviť ako body v p rozmernom priestore, kde každý z rozmerov predstavuje hodnoty jedného parametra. V prípade spoločenských sú objektami vzorky a parametrami druhy, prípadne charakteristiky prostredia.

Keď pracujeme len s dvoma alebo troma parametrami, je možné bez problémov sledovať v dvoj- alebo trojrozmernom grafe vzťahy medzi objektami, ich vzdialenosť a zoskupenie.

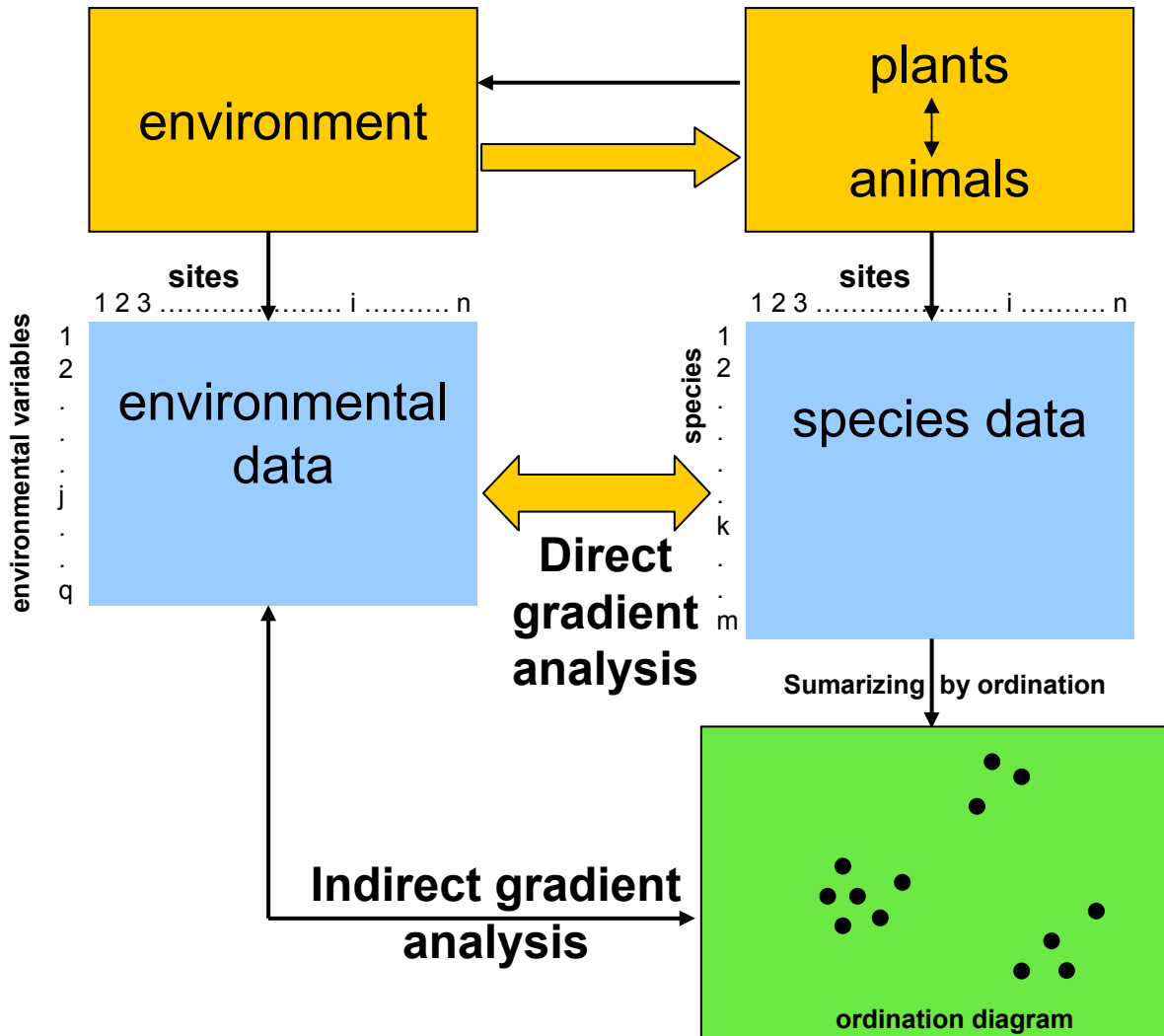
Pri väčšom počte parametrov je nutné redukovať ich počet s čo najmenšou stratou informácie.



Ordinačné metódy

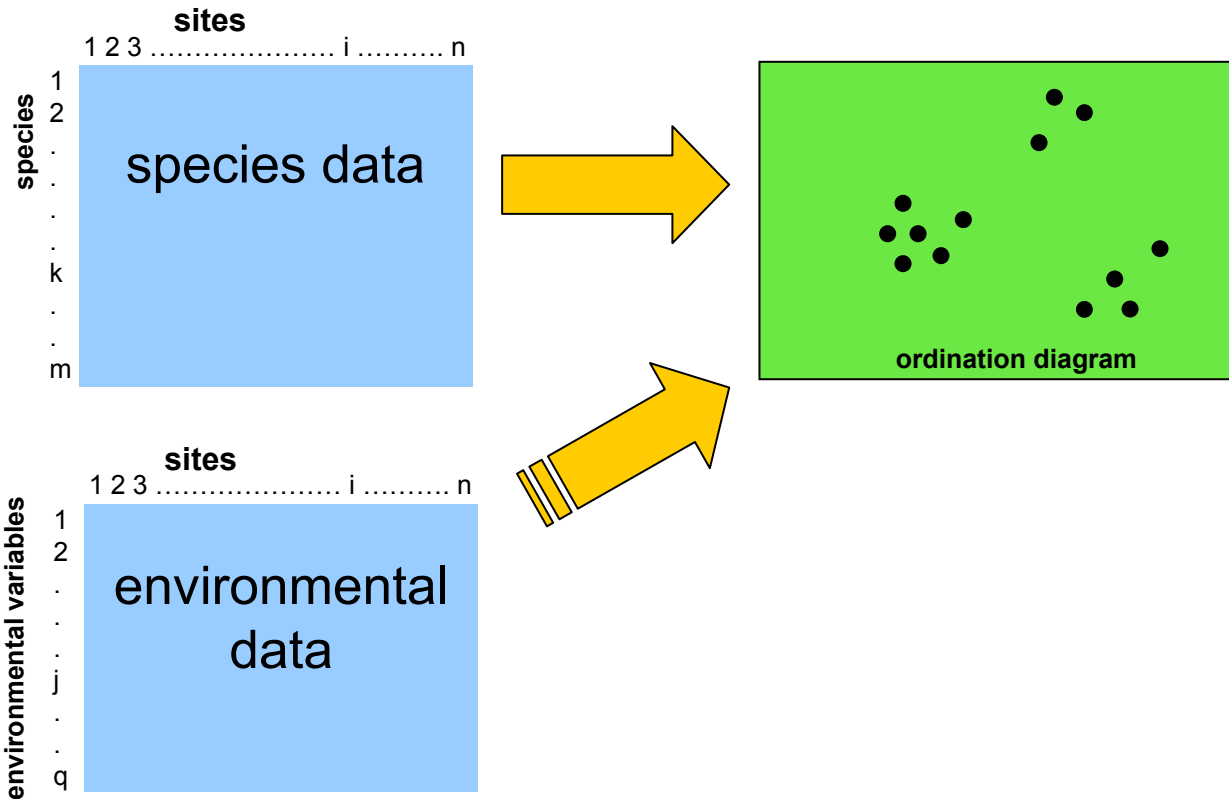


Ordinačné metódy



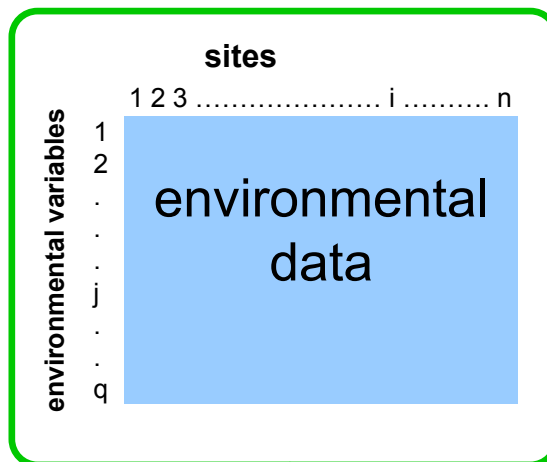
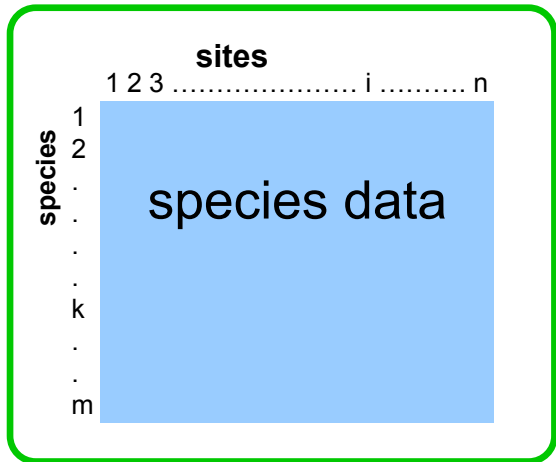
- ◆ zoradí objekty pozdĺž environmentálneho gradientu
- ◆ cieľom ordinácie je sformulovať hypotézy o vzťahu medzi druhovým zložením spoločenstva na lokalitách a základnými environmentálnymi faktormi

Ordinačné metódy = gradientová analýza



- ◆ Ordinačné metódy nepredpokladajú žiadne apriorné zoskupenie objektov.
- ◆ Ordinačné metódy patria medzi metódy, ktoré sa používajú hlavne k tvorbe hypotéz.

Ordinačné metódy: terminológia



◆ Primárne dáta:

vzorky, objekty (*samples, sites*)

Každá vzorka zahŕňa hodnoty pre viac druhov (*species*) alebo tzv. charakteristík prostredia (*environmental variables, variables*).

Vysvetľované premenné
(*response*)
druhové dáta (*species data*)

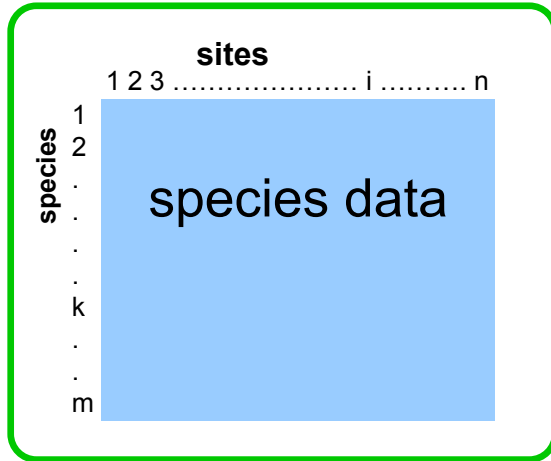
Vysvetľujúce premenné
(*explanatory*)

Charakteristiky prostredia
(*environmental variables, variables*)

Kovariáty
(*covariates, covariables*)
ich vplyv chceme oddeliť

Ordinačné metódy: typy dát

Vysvetľované premenné



Druhové dáta: akékoľvek premenné, ktorých hodnoty chceme predpovedať.

Druhové zloženie spoločenstva

je možné určovať presným kvantitatívnym spôsobom:

- ◆ počet jedincov jednotlivých druhov
- ◆ percentická pokrývnosť
- ◆ odhad biomasy

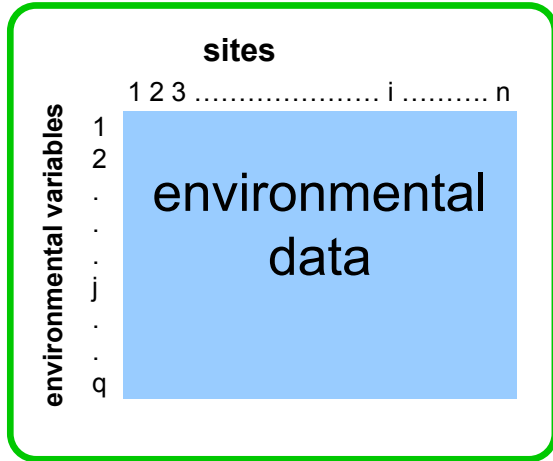
prípadne podľa semikvantitatívnej stupnice

- ◆ Braun-Blanquetová stupnica
- alebo len kvalitatívnym spôsobom
- ◆ prítomnosť či neprítomnosť

Ak by sme chceli napr. predpovedať obsah iónov kovu na základe skladby krajiny v povodí riek, v programe Canoco by boli koncentrácie jednotlivých iónov "druhy".

Ordinačné metódy: typy dát

Vysvetľujúce premenné, prediktory



Môžu byť použité k predpovedaniu hodnôt vysvetľovaných premenných (napr. predpovedanie zloženia rastlinného spoločenstva na základe znalosti pôdných charakteristík a type obhospodarovania).

Charakteristiky prostredia, príp. kovariáty

- ◆ kvantitatívne premenné
- ◆ semikvantitatívne premenné
- ◆ faktoriálne (kategorálne) premenné - prekódovanie do 0,1

- ◆ faktoriálne (kategorálne) – v Canoco prekódovať do tzv. **indikátorových premenných** (*dummy variables*)



vzorka	Geo
Vz 1	žula
Vz 2	žula
Vz 3	čadič
Vz 4	rula



vzorka	žula	čadič	rula
Vz 1	1	0	0
Vz 2	1	0	0
Vz 3	0	1	0
Vz 4	0	0	1



Kovariáty (*covariables, covariates*): ak určitá vysvetľujúca premenná má vplyv na druhové dáta, ale pre nás je nezaujímava, môžeme jej vplyv odpočítať => jej vplyv neinterpretujeme, chceme ho vziať do úvahy pri hodnotení vplyvu iných premenných

Ordinačné metódy: typy dát

Čo s chýbajúcimi dátami:

- ◆ **Vzorky**, v ktorých hodnoty chýbajú, môžeme **vypustiť**. Výhodné vtedy, ak sú chýbajúce dáta len v niekoľko málo vzorkách (*case-wise deletion*).
- ◆ **Premenné**, v ktorých hodnoty chýbajú, môžeme **vypustiť**, ak ich nie je veľa.
- ◆ **Doplnenie** chýbajúcich údajov:
 - ◆ doplnenie priemeru zo vzoriek, kde sú hodnoty k dispozícii
 - ◆ dopočítanie chýbajúcich hodnôt na základe mnohonásobného regresného modelu (takto ale prichádzame o stupne voľnosti) možnosť vzorkám s doplnenými hodnotami priradiť nižšiu váhu

Ordinačné metódy: príklad využitia

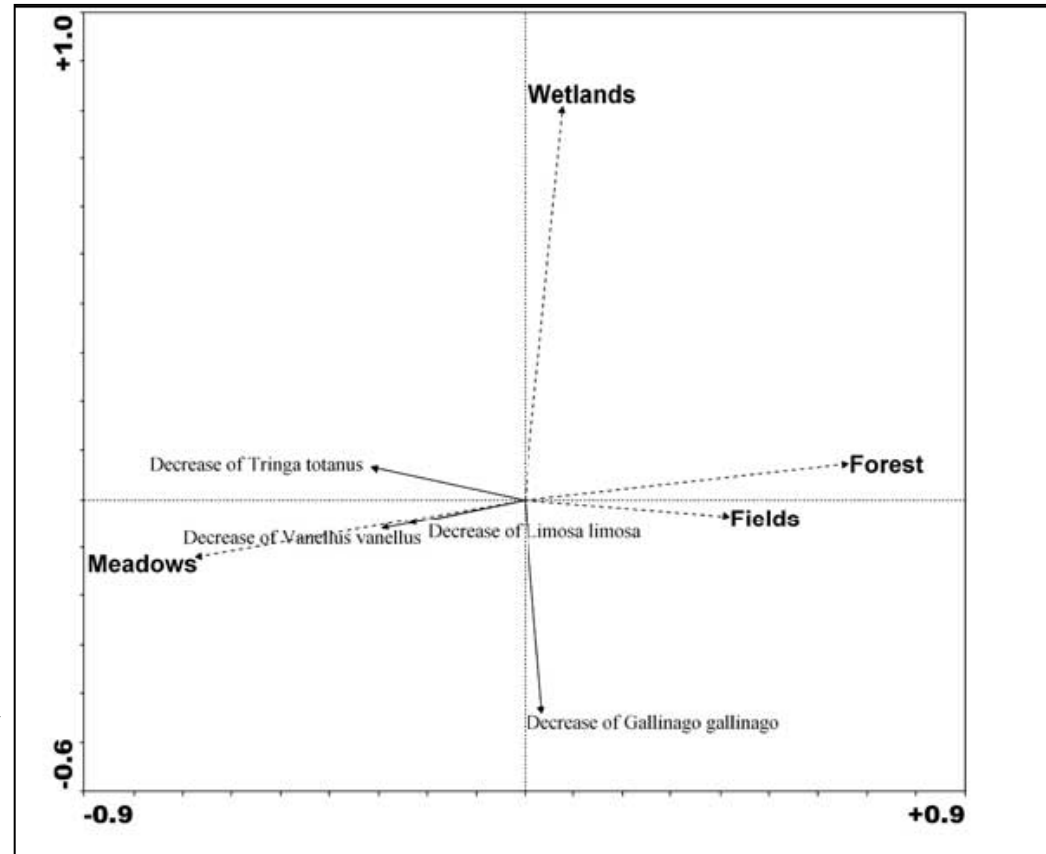
Ukážka užitočného použitia mnohorozmernej analýzy:

Otázky:

- ◆ Môžeme na základe aktuálneho stavu krajiny predpovedať zánik hniezdiska ohrozeného vodného vtáka?
- ◆ Které zo zložiek krajiny sú pre túto predpoveď najdôležitejšie?

Diagram naznačuje:

- ◆ 3 vtáčie druhy v krajine s vyšším percentom lúk svoju hniezdnú frekvenciu znižujú
- ◆ druh *Gallinago gallinago* ustúpil do krajiny, kde bolo malé percento mokradí
- ◆ otestovaním naznačených vzťahov nebola zistená štatistická významnosť žiadneho vzťahu



Ordinačný diagram znázorňuje prvú a druhú os redunďnej analýzy (RDA) údajov o hniezdných preferenciách vtákov.

Typy štatistických modelov

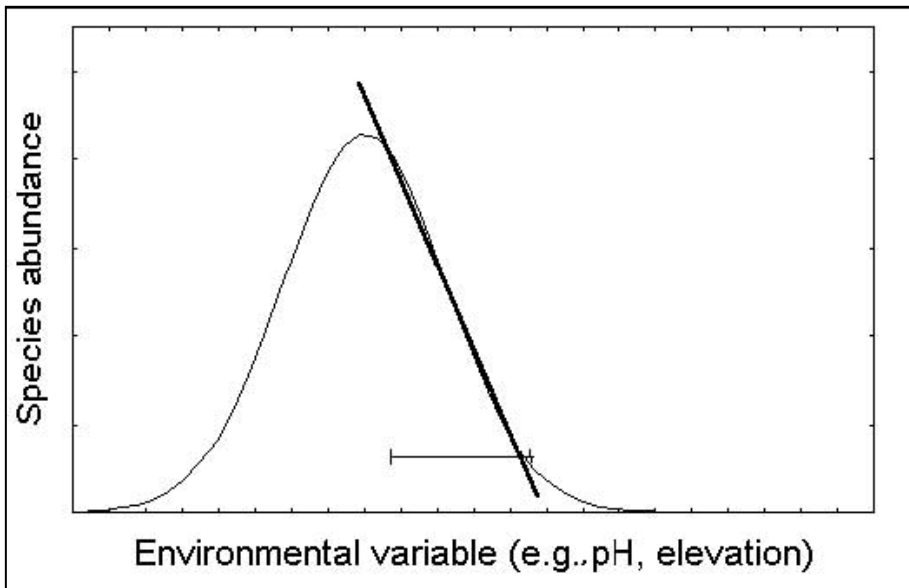
Nasledujúca tabuľka zhŕňa najdôležitejšie štatistické metódy používané v rôznych situáciách:

Vysvetľovaná premenná ...	Prediktor(y)	
	nemáme	máme
... je jedna	<ul style="list-style-type: none">◆ zhrnutie distribučných vlastností	<ul style="list-style-type: none">◆ regresný model s.l.
... je ich viac	<ul style="list-style-type: none">◆ nepriama gradientová analýza (indirect gradient analysis - PCA, DCA, NMDS)◆ zhluková analýza	<ul style="list-style-type: none">◆ priama gradientová analýza◆ obmedzená zhluková analýza◆ diskriminačná analýza (discriminant analysis - CVA)

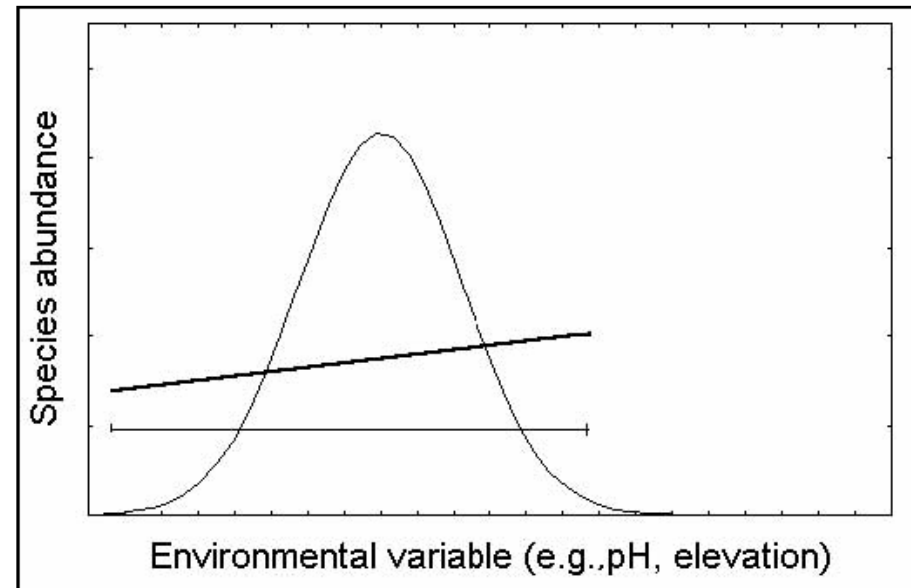
Modely odpovede druhov na gradienty prostredia

Dva typy modelu odpovede druhu na gradienty prostredia

- ◆ lineárny (*linear*) – najjednoduchší odhad (na krátkom gradiente dobre funguje lineárna aproximácia akejkoľvek funkcie)
- ◆ unimodálny (*unimodal*) – predpokladá, že druh má na gradientu prostredia svoje optimum (na dlhom gradiente je aproximácia lineárnou funkciou veľmi nevhodná)



Lineárna aproximácia unimodálnej odpovede na krátkom výseku gradientu



Lineárna aproximácia unimodálnej odpovede na dlhej časti gradientu

Základné techniky ordinačných metód

Indirect gradient analysis

- ♦ vytvorí teoretickú premennú (gradient) ktorá najlepšie charakterizuje druhové dáta na základe lineárneho alebo unimodálneho modelu (závisí od dĺžky gradientu)

linear model

Principal component analysis (PCA)
Principal coordinate analysis (PCoA)

unimodal model

Correspondence analysis (CA)
Detrended correspondence analysis (DCA)

non-metric ordination

Multidimensional scaling (NMDS)

Direct gradient analysis

- ♦ gradient je lineárnou kombináciou konkrétnych environmentálnych premenných

linear model

Redundancy analysis (RDA)
Canonical correlation analysis

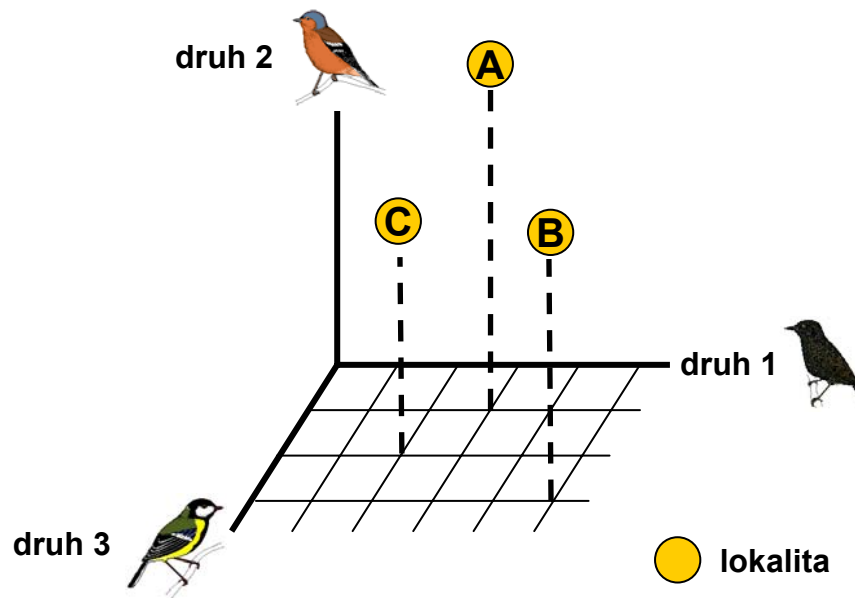
unimodal model

Canonical correspondence analysis (CCA)

Nepriame ordinačné metódy

Problém nepriamej ordinácie môžeme formulovať niekoľkými spôsobmi:

1. Nájdi také rozloženie vzoriek v ordinačnom priestore, kde vzdialenosť vzorky v ordinačnom priestore odpovedá najlepšie rozdielom v druhovom zložení. Toto explicitne robí **nemetrické** (ale aj metrické) **mnohorozmerné škálovanie** (*non-metric multidimensional scaling, NMDS*).
2. Nájdi teoretické (latentné) premenné (= ordinačné osi), pre ktoré je celková závislosť všetkých druhov najtesnejšia. Tento model vyžaduje, aby bol **typ odpovedí** druhov na premenné explicitne špecifikovaný: **lineárna** odpoveď pre lineárne metódy, **unimodálna** odpoveď pre metódy založené na vážených priemeroch. V lineárnych metódach je skóre vzorky lineárnou kombináciou (váženým súčtom) skóre druhov. V metódach váženého priemeru sa skóre vzorky vypočíta váženým priemerom druhových skóre (po určitých úpravách).
3. Keď si predstavíme vzorky ako body v mnohorozmernom priestore, kde sú druhy osami a pozícia každej vzorky odpovedá početnosti príslušného druhu. Potom je cieľom ordinácie nájsť také premietnutie tohto mnohorozmerného priestoru do priestoru s menším počtom dimenzií, ktoré spôsobí minimálne skreslenie priestorových väzieb. Výsledok závisí na tom, ako definujeme „minimálne skreslenie“.



Priame ordinačné metódy

Priame ordinačné metódy:

hľadanie najlepších vysvetľujúcich premenných.

V nepriamych ordináciách hľadáme akúkoľvek premennú, ktorá je schopná vysvetliť najlepšie druhové zloženie (a tú potom vezmeme ako ordinačnú os).

V priamych ordináciách sú ordinačnými osami vážené charakteristiky prostredia.

Čím menej týchto charakteristík máme, tým prísnejšie bude obmedzenie.

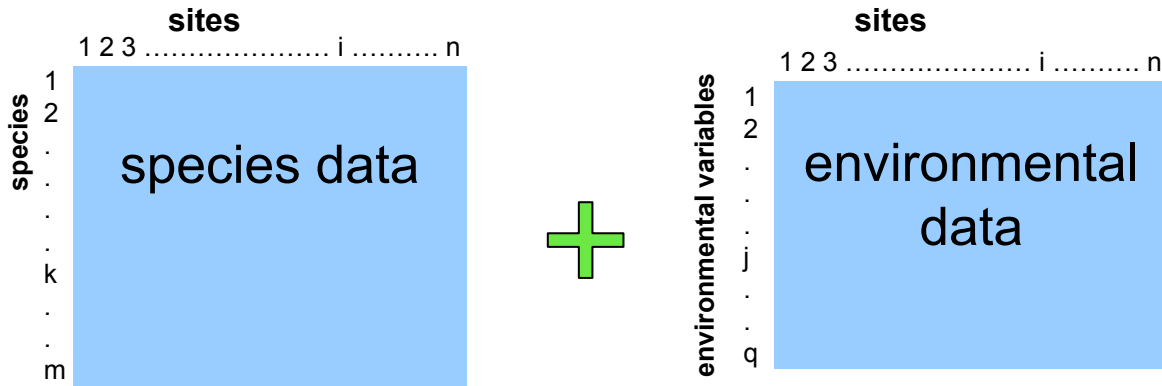
 Ak je ich počet väčší než počet vzoriek zmenšený o jednu, tak sa ordinácia stáva nepriamou.

Neobmedzené (*unconstrained*) ordinačné osy odpovedajú smeru najväčšej variability v súbore dát. **Obmedzené** (*constrained*) **ordinačné osi** odpovedajú smeru najväčšej variability v dátovom súbore, ktorá môže byť vysvetlená charakteristikami prostredia.

 Počet obmedzených osí nemôže byť väčší než počet charakteristík prostredia.

Priama či nepriama gradientová analýza?

Máme druhové dáta aj charakteristiky prostredia.



Môžeme použiť oba prístupy: priamu aj nepriamu ordináciu.

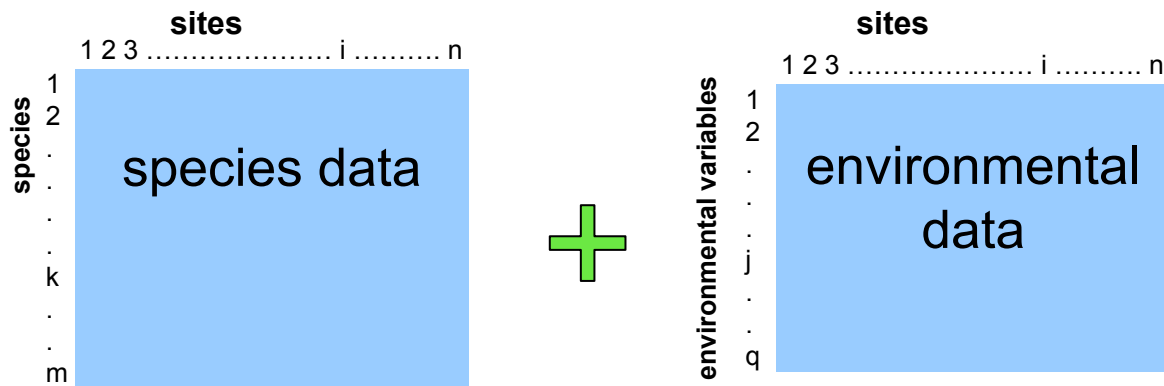
1. Spočítame najprv nepriamu ordináciu s následnou regresiou ordinačných osí na merané charakteristiky prostredia (tj. premietnutie týchto charakteristík do ordinačného diagramu)
2. Spočítame priamu (obmedzenú) ordináciu.

Tieto prístupy sú komplementárne a mali by sa použiť oba.

Je potrebné vždy uviesť metódu, ktorá bola použitá.

Hybridná gradientová analýza?

Máme druhové dáta aj charakteristiky prostredia.



Hybridná analýza: „kríženec“ medzi priamou a nepriamou ordináciou.

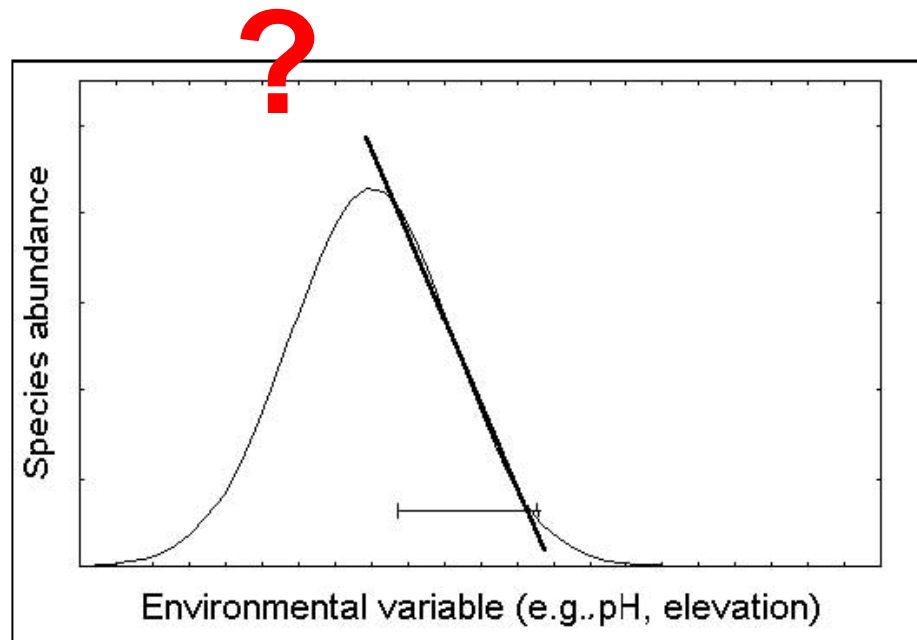
V štandardnej priamej ordinácii je toľko obmedzených (kanonických) osí, koľko je nezávislých vysvetľujúcich premenných a len ďalšie ordinačné osi sú neobmedzené.

V hybridnej analýze sa spočíta len vopred daný počet obmedzených osí a akékoľvek ďalšie ordinačné osi sú neobmedzené.

Lineárny alebo unimodálny model?

Voľba modelu: na základe dĺžky gradientu

- ◆ **unimodálny model** ak dĺžka najdlhšieho gradientu ≥ 4
(techniky váženého priemerovania sú lepšie pre heterogénne dáta)
- ◆ **lineárny model** ak dĺžka najdlhšieho gradientu < 3 (nie je to však nutnosť použiť lineárny model)
(techniky založené na modely lineárnej odpovede sú vhodné pre homogénne dátové súbory)



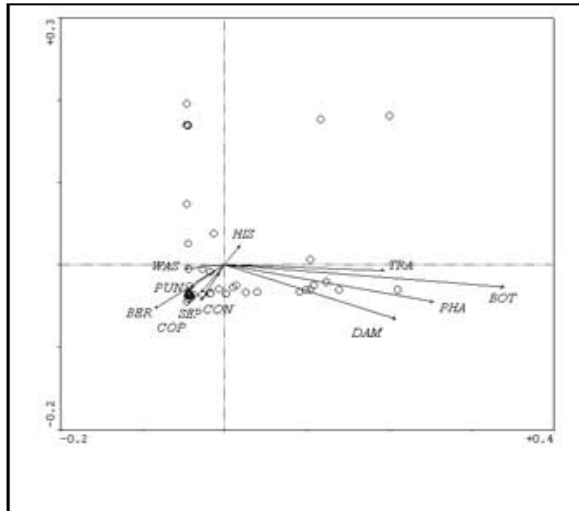
Ordinačné diagramy

Výsledky ordinácií se obvykle prezentujú ako **ordinačné diagramy**.

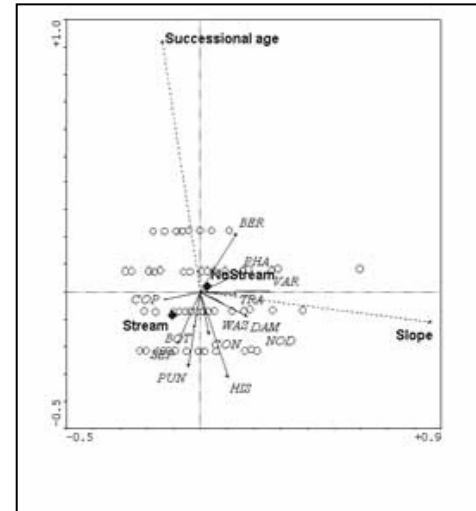
	Lineárny model	Unimodálny model
vzorky	◆ znázornené bodmi (symbolmi)	
druhy	◆ šípky v smere rastu abundancií	◆ body (symboly) označujúce optimum druhu
Charakteristiky prostredia kvantitatívne	◆ šípky v smere rastu hodnôt	
charakteristiky prostredia kvalitatívne	◆ centroidy pre jednotlivé kategórie	

Príklady ordinačných diagramov

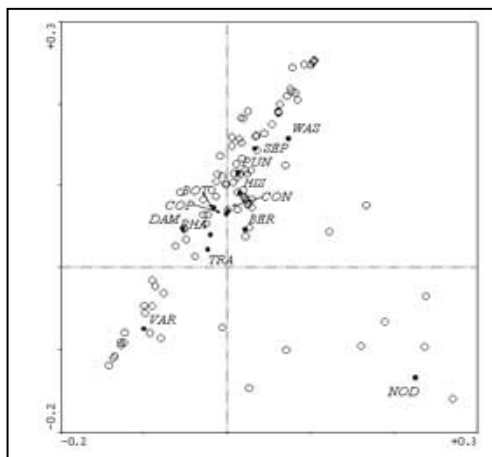
PCA



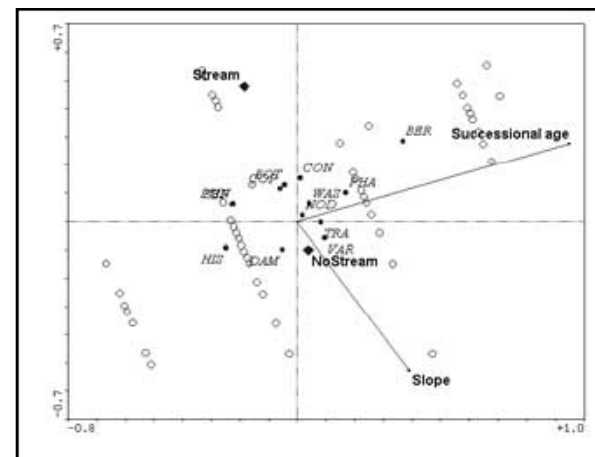
RDA



CA



CCA



Formát súborov pre program Canoco

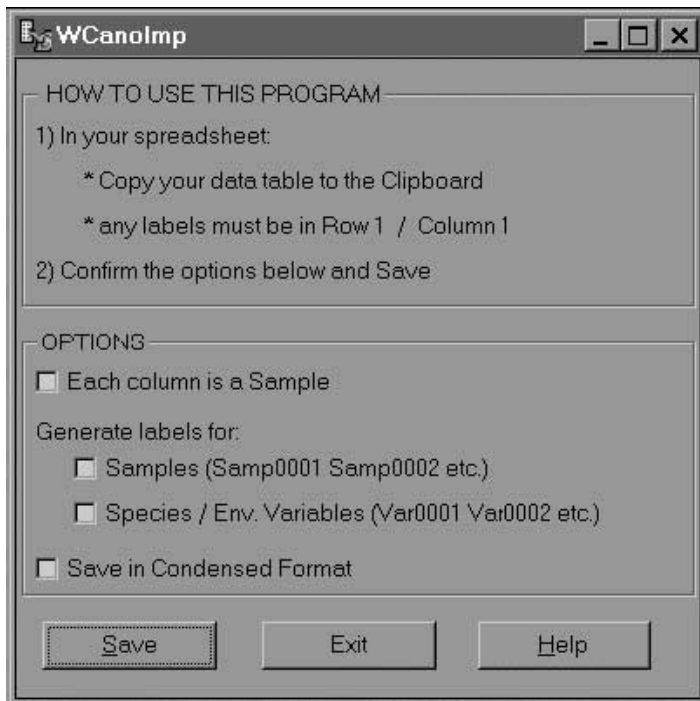
Danka Némethová

28. – 30. 5. 2007

Import súborov, program Canolmp

- ◆ Možnosť importu dát z tabuliek vytvorených v Exceli alebo z relačných databázy

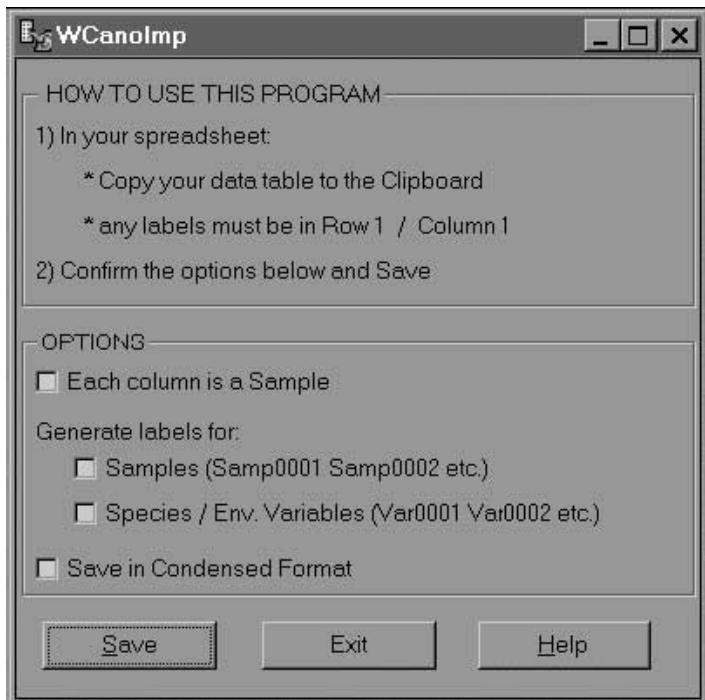
Program Canolmp



- ◆ V tabuľke usporiadať dáta do obdĺžnikovej matice (vzorky v riadkoch, premenné v stĺpcoch)
- ◆ Hlavičky riadkov a stĺpcov (možné, nie povinné)
- ◆ Mená (*labels*) musia byť dlhé najviac 8 znakov; len písmená, medzery
(ak je názov dlhší – program ho po 8 znaku odstrihne; zakázané znaky nahradí bodkou)
- ◆ Pole tabuľky – len čísla alebo prázdne polia (žiadne iné znaky)

Import súborov, program Canolmp

Program Canolmp



1. Skopírovať tabuľku v Exceli do pamäte (Windows Clipboard)
2. Spustiť program WCanolmp
3. Skontrolovať nastavenie WCanolmp
4. Ak máme maticu transponovanú (vzorky v stĺpcoch), zvoliť možnosť **Each column is a Sample**
5. Ak nemáme mená vzoriek v prvom stĺpci, zvolíme **Generate labels for: ... Samples**
6. Ak nemáme mená premenných v prvom riadku, zvolíme **Generate labels for: ... Species**
7. Možnosť **Save in Condensed Format** riadi formát, ktorý sa pri tvorbe súboru použije
8. Zvolíme **Save**: vyberieme meno súboru a uložíme
9. WCanolmp požaduje krátky popis dát; tento popis sa bude objavovať vo výstupoch z analýz. Ak popis nevložíme, je ponúkaný neutrálny text.

Dátový súbor v programe Canoco

Full format

- ① Popis súboru (max 80 znakov)
- ② Formát pre hodnoty dát
- ③ Počet premenných v matici
- ④ Dáta, začínajúc číslom vzorky
- ⑤ Nulový riadok (ukončuje časť s dátami)
- ⑥ Mená (*labels*) premenných na 8 znakov, 10 v jednom riadku
- ⑦ Mená (*labels*) vzoriek na 8 znakov, 10 v jednom riadku

```
① WCanoImp produced data file
② (I5,1X,12F3.0)
③ 12
④   1   1   1   0   1   0   1   0   0   0   0   0
     2   1   0   0   1   0   0   1   0   0   0   0
     3   0   1   0   1   0   0   0   1   0   0   0
     ...
     48   1   1   0   0   1   0   0   0   0   0   0
⑤   0   0   0   0   0   0   0   0   0   0   0   0
⑥ PhosphatBenlate Year94 Year95 Year98 B01 B02 B03 B04 B05
   B06 B07
⑦ PD01 PD02 PD03 PD04 PD05 PD06 PD07 PD08 PD09 PD10
   PD11 PD12 PD13 PD14 PD15 PD16 C01 C02 C03 C04
   ...
```

Dátový súbor v programe Canoco

Kondenzovaný formát

- 1 Popis súboru (max 80 znakov)
- 2 Formát pre hodnoty dát – formát riadku
- 3 Počet hodnôt v riadku
- 4 Dáta, začínajúc číslom vzorky
- 5 Nulový riadok (ukončuje časť s dátami)
- 6 Mená (*labels*) premenných na 8 znakov, 10 v jednom riadku
- 7 Mená (*labels*) vzoriek na 8 znakov, 10 v jednom riadku

```
1 WCanoImp produced data file
2 (I5,1X,6(I6,F3.0))
3 6
4   1      23  1      25 10      36  3      41  4      53  5      57  3
   1      89 70     100  1     102  1     115  2     121  1
   2      11  1      26  1      38  5      42 20      50  1      55 30
   2      62  2      69  1      70  5      74  1      77  1      86  7
...
   79     131 15
5  0
6 TanaVulgSeneAquaAvenPratLoliMultSalxPurpErioAnguStelPaluSphagnumCarxCaneSalxAuri
...
7 SangOffiCalaArunGlycFlui
PRESEK  SATLAV  CERLK  CERJIH  CERTOP  CERSEV  ROZ13  ROZ24  ROZC5  ROZR10
...
```

Úprava dát v programe Canoco

Danka Némethová

28. – 30. 5. 2007

Transformácia druhových dát

Logaritmická transformácia

$$y' = \log(A * y + C)$$

Čísla A a C volíme tak, aby bol výsledok vždy väčší alebo rovný 1.

Default hodnoty A a C sú rovné 1 (nulové hodnoty mení na 0, ostatné sú kladné).

Hodí sa výborne napr. na percentuálne dáta na stupnici 0-100.

Transformácia na ordinálnu škálu

Dáta o zložení rastlinného spoločenstva odhadované často na semikvantitatívnej Braun-Blanquetovej stupnici so siedmimi stupňami (r, +, 1, 2, 3, 4, 5). Takáto stupnica býva kvantifikovaná odpovedajúcimi poradovými hodnotami (od 1 do 7).

Je možné nahradiť stupne stredom intervalu pokrývnosti:

r	0.1
+	0.5
1	3
2	15
3	37.5
4	62.5
5	87.5

Transformácia druhových dát

Odmocninová transformácia

$$y' = \sqrt{y}$$

Táto transformácia môže byť vhodným riešením pre dáta vyjadrujúce počty (počet jedincov apod.). Na tieto dáta však môžeme použiť aj logaritmickú transformáciu.

Iné transformácie

Ak potrebujeme iný typ transformácie, ktorý Canoco neponúka, môžeme ju previesť v tabuľkovom procesore a transformované dáta do Canoca vyexportovať.

- ◆ Je to užitočné, ak naše „druhové“ dáta nepopisujú zloženie spoločenstva, ale niečo jako chemické či fyzikálne vlastnosti pôdy. V takom prípade mávajú premenné rôzne jednotky a pre každú z nich môže byť vhodná iná transformácia.

Transformácia vysvetľujúcich premenných

Transformácia vysvetľujúcich premenných

- ◆ U vysvetľujúcich premenných (**charakteristík prostredia a kovariát**) sa predpokladá, že nemajú jednotnú stupnicu a že pre každú z nich musíme voliť vhodnú transformáciu (vrátane častej voľby – netransformovať).
- ◆ Canoco ale taký postup neumožňuje, takže prípadnú transformáciu vysvetľujúcich premenných musíme previesť pred ich exportom do súboru v Canoco formáte.
- ◆ V každom prípade však Canoco potom, čo charakteristiky prostredia a / alebo kovariáty načíta, ich **šstandardizuje**, čiže majú nulový priemer a jednotkový rozptyl.

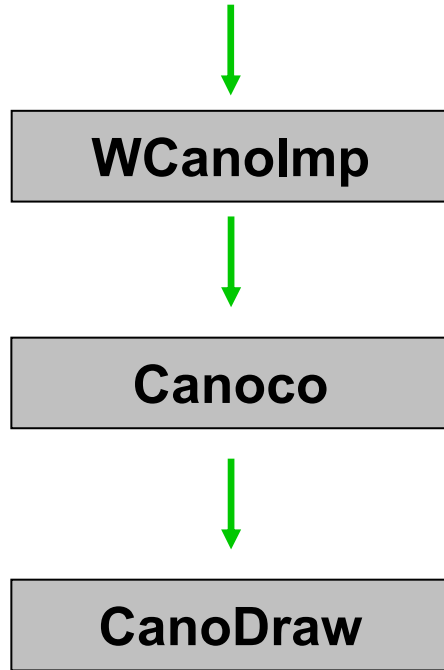
Nepriame ordinácie v Canocu

Danka Némethová

28. – 30. 5. 2007

Postup analýzy v programe Canoco

Príprava dát napr. v Exceli



Dve okná:

- ◆ **Project view** (zhrňa najdôležitejšie charakteristiky projektu)
- ◆ **Log view** (záznamy o prevedených akciách a výsledky analýzy)
- ◆ „**SOL**“ výsledky

Nastavenie projektu pomocou sprievodcu: **Options**

Po nadefinovaní projektu: **Analyze**

Po zanalyzovaní grafický výstup: **CanoDraw**

Voľba modelu – meranie dĺžky gradientu

Aby sme mohli zvoliť medzi lineárnym a unimodálnym modelom, musíme odmerať dĺžku gradientu.

1. Spravíme skúšobný projekt – nastavíme detrendovanú korešpondenčnú analýzu (DCA), prípadne jej kanonickú formu (DCCA).
2. Použijeme metódu odstránenia trendu po segmentoch (čo v sebe zahŕňa tiež Hillovo škálovanie ordinačných skóre)
3. Zvolíme aj ostatné nastavenia rovnaké ako v záverečných analýzach
4. Spustíme analýzu
5. V okne Log view prezrieme výsledky – na konci výpisu je súhrnná tabuľka (Summary table), v nej riadok začínajúci slovami „Lengths of gradient“

```
Lengths of gradient : 2.990 1.324 .812 .681
```

- ◆ **unimodálny model** ak dĺžka najdlhšieho gradientu ≥ 4
- ◆ **lineárny model** ak dĺžka najdlhšieho gradientu < 3 (nie je to však nutnosť použiť lineárny model)

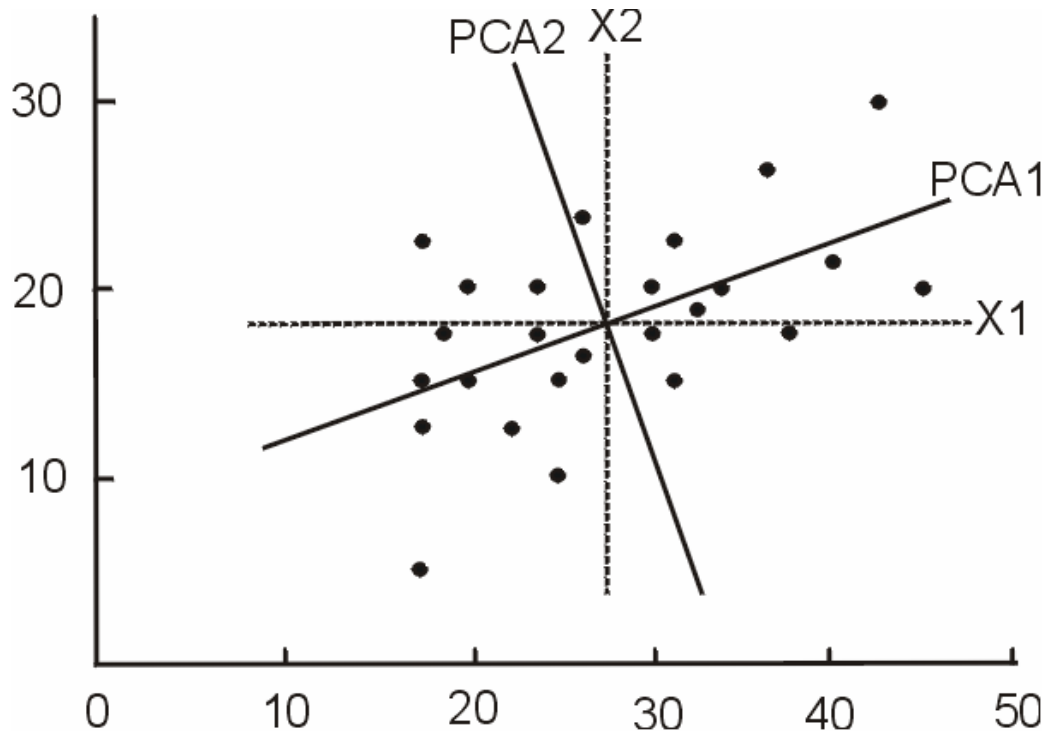
Analýza hlavných komponent (PCA)

Danka Némethová

28. – 30. 5. 2007

Analýza hlavných komponent (PCA)

Nahrádza pôvodný súbor pozorovaných parametrov (druhovú maticu) súborom nových (hypotetických), vzájomne nekorelovaných premenných tak, že prvá nová os (prvá hlavná komponenta, PC1, prvý nový parameter) je vedená v smere najväčšej variability medzi objektami, druhá os (druhá hlavná komponenta, PC2, druhý nový parameter) je vedená v smere najväčšej variability, ktorý je kolmý na smer prvej komponenty, atď.

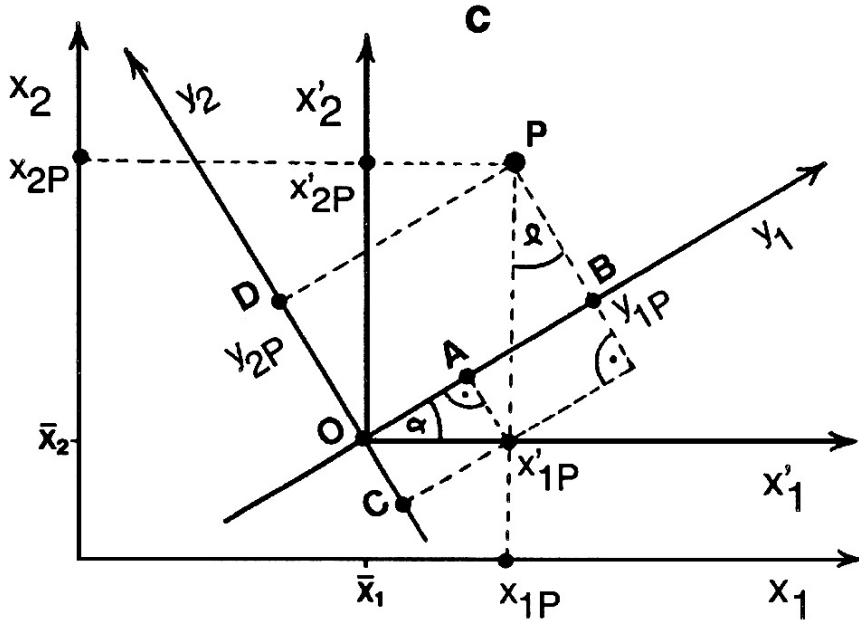


Je založená na **vlastnej analýze** (eigenanalysis) symetrických matíc (**korelačnej, kovariančnej**)

Analýza hlavných komponent (PCA)

Cieľ PCA: určenie uhlov medzi pôvodnými a novými osami súradnicovej sústavy, súradnice objektov v novom systéme súradnic.

Nové osy (komponenty) nie sú vzájomne korelované.

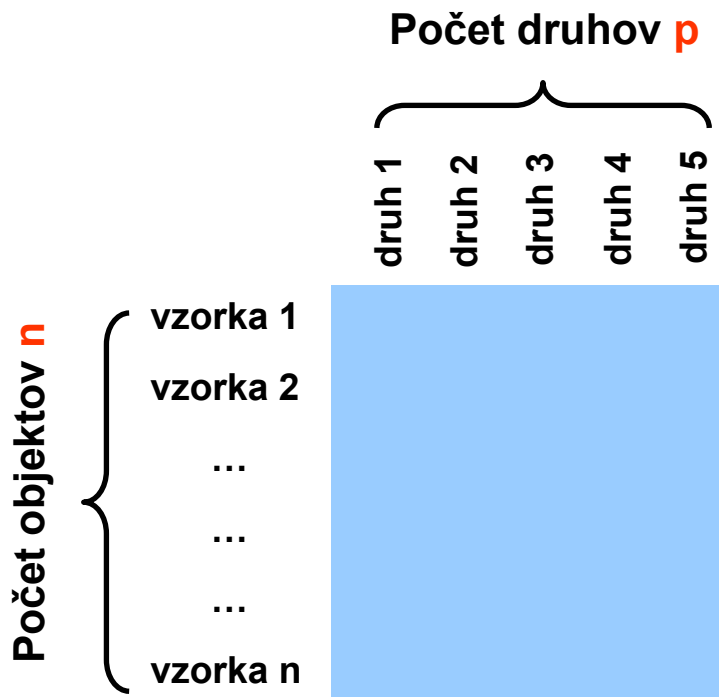


Pôvodne bola PCA navrhnutá pre kvantitatívne znaky, môže sa však použiť aj na znaky binárne a semikvantitatívne.

Vlastné čísla matice $\lambda_1, \lambda_2, \dots, \lambda_p$ sú interpretovateľné ako miery rozptylu zachytené komponentami y_1, \dots, y_p .

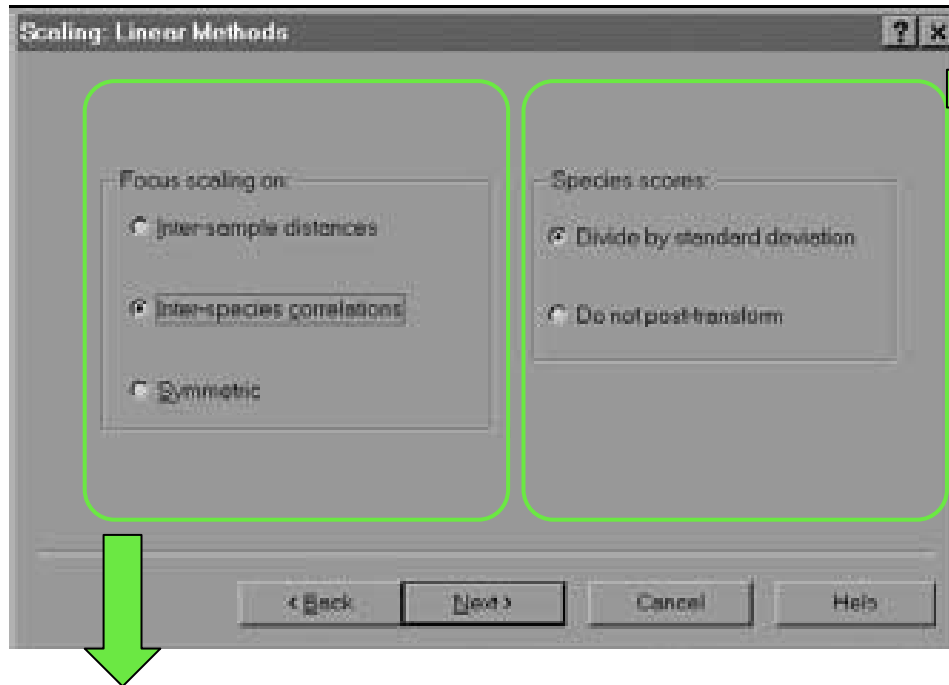
Analýza hlavných komponent (PCA)

- ◆ Počet objektov (vzoriek) pri PCA by mal byť aspoň o jeden väčší než je počet analyzovaných druhov.
- ◆ Obvykle sa však odporúča, aby sa počet objektov blížil druhej mocnine počtu druhov (súvisí s počtom stupňov voľnosti).
- ◆ V prípade, že $n \leq p$, výsledná matica (korelačná alebo kovariančná) rádu p má len $n - 1$ nezávislých riadkov alebo stĺpcov. V takom prípade príslušná matica má $p - (n - 1)$ nulových vlastných čísiel (na umiestnení n objektov podľa ich vzájomných vzdialeností je potrebných len $n - 1$ rozmerov).



Analýza hlavných komponent (PCA)

Nastavenie škálovania



Početnosti jednotlivých druhov sa môžu odrážať v dĺžke ich šípok (dominantné druhy budú mať potom šípky dlhšie než druhy vzácnejšie). (*species scores: do not post-transform*)

Každý druh môže byť zrelativizovaný (*divide by standard deviation* – vhodné pre tzv. korelačné projekčné diagramy).

Presnosť záveru o podobnosti druhov, vzťahov medzi druhmi a / alebo charakteristikami prostredia závisí z časti na škálach na jednotlivých ordinačných osiach.

V prvom rade sa rozhodneme, či sa pri interpretácii zameriame na vzorky (porovnanie tried vzoriek, apod.) alebo druhy.

Ak máme charakteristiky prostredia, prípadne kovariáty, *species scaling* umožňuje charakterizovať korelácie medzi charakteristikami prostredia.

Analýza hlavných komponent (PCA)

Pred vlastným počítaním ordinácie je nutné nastaviť možnosti manipulácie s tabuľkou druhových dát

Centrovanie

Priemer každého riadku bude rovný nule.



The dialog box is titled "Centering and Standardization" and is divided into two columns: "SAMPLES" and "SPECIES".

SAMPLES:

- None
- Center by sample (highlighted with a green box)
- Standardize by norm
- Center and standardize

SPECIES:

- None
- Center by species
- Standardize by norm
- Center and standardize
- Standardize by error variance

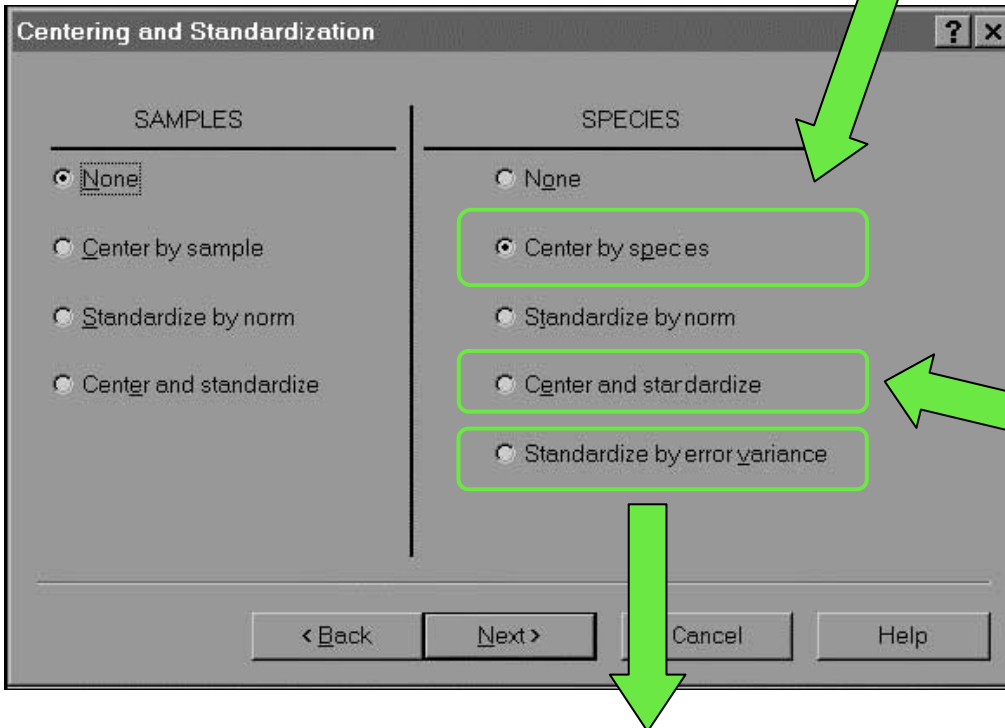
Buttons at the bottom: < Back, Next >, Cancel, Help.

Centrovanie druhov nutné pre lineárne metódy s obmedzením (**RDA**) alebo pre parciálnu lineárnu ordináciu (tj. pri použití **kovariát**)

Analýza hlavných komponent (PCA)

Štandardizácia

Priemer každého stĺpca bude rovný nule.



Štandardizácia (vzoriek alebo druhov) spôsobí, že norma každého riadku alebo stĺpca bude rovná jednej. Táto **norma** je odmocnina zo sumy štvorcov hodnôt v riadku alebo stĺpci. Ak použijeme centrovanie aj štandardizáciu, prevedie sa centrovanie ako prvé.

Po vycentrovaní a štandardizácii budú v stĺpcoch premenné s nulovým priemerom a jednotkovým rozptylom. PCA na druhových dátach bude odpovedať „**PCA na matici korelácií**“.

- ◆ Ak máme charakteristiky prostredia (v RDA a v PCA externe), môžeme zvoliť štandardizáciu chybovým rozptylom (error variance).
- ◆ Tu Canoco odhaduje pre každý druh zvlášť rozptyl v druhových dátach, ktorý zostane nevysvetlený po fitovaní závislosti hodnôt tohto druhu na vybraných charakteristikách prostredia (a kovariátach, ak ich máme).
- ◆ Prevrátená hodnota tohto rozptylu sa potom použije ako váha druhu.
- ◆ Čím lepšie bude druh popísaný charakteristikami prostredia, tým vyššiu bude mať váhu.

Analýza hlavných komponent (PCA)

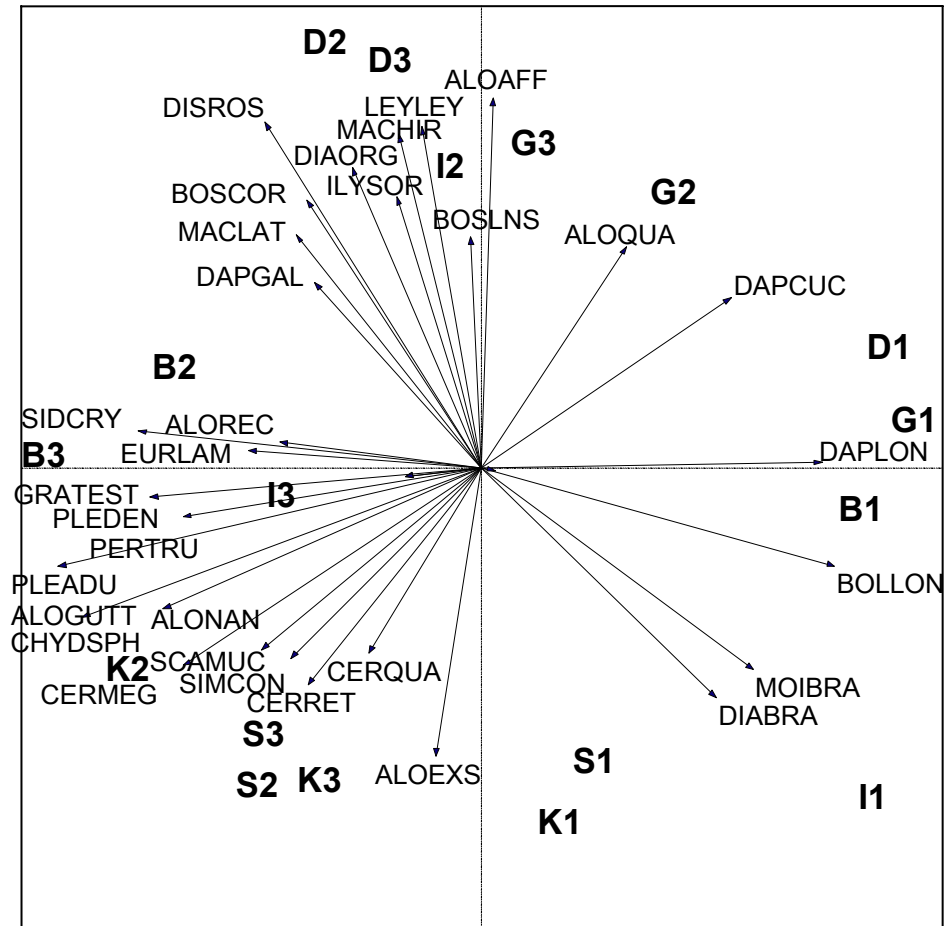
Indirect gradient analysis

Principal component analysis

- ◆ PCA je postavená na lineárnom modeli; abundancia každého druhu buď narastá alebo klesá s hodnotou každého environmentálneho gradientu
- ◆ PCA je definovaná pre kovariančnú a pre korelačnú maticu
- ◆ PCA nie je vhodná pre dátovú maticu s veľa nulami

REÁLNE DÁTA

- ◆ 6 lokalít, každá lokalita sledovaná 3 obdobia
- ◆ dátová matica: 18 vzoriek x 63 plankt. druhov
hodnoty = stupeň dominancie



Analýza hlavných koordinát (PCoA)

Danka Némethová

21. – 22. 5. 2007

Analýza hlavných koordinát v Canoco

Principal coordinates analysis (PCoA, PCO):

klasické, metrické škálovanie

Vstupom je matica nepodobností alebo podobností medzi vzorkami, z ktorej sa počíta ordinácia.

V ordinačnom diagrame sú vzorky rozmiestnené tak, že podobné vzorky sú blízko seba, kým vzorky nepodobné sú od seba vzdialené.

Možnosť spočítať PCoA v Canoco:

1. zvoliť analýzu hlavných komponent (PCA)
2. ako druhové dáta je pripravená matica podobností alebo nepodobností (avšak s opačným znamienkom) – táto matica je teda štvorcová
3. Centered by samples
4. Centered by species
5. Symetrické škálovanie ordinačných skóre; species score nie sú nijak transformované

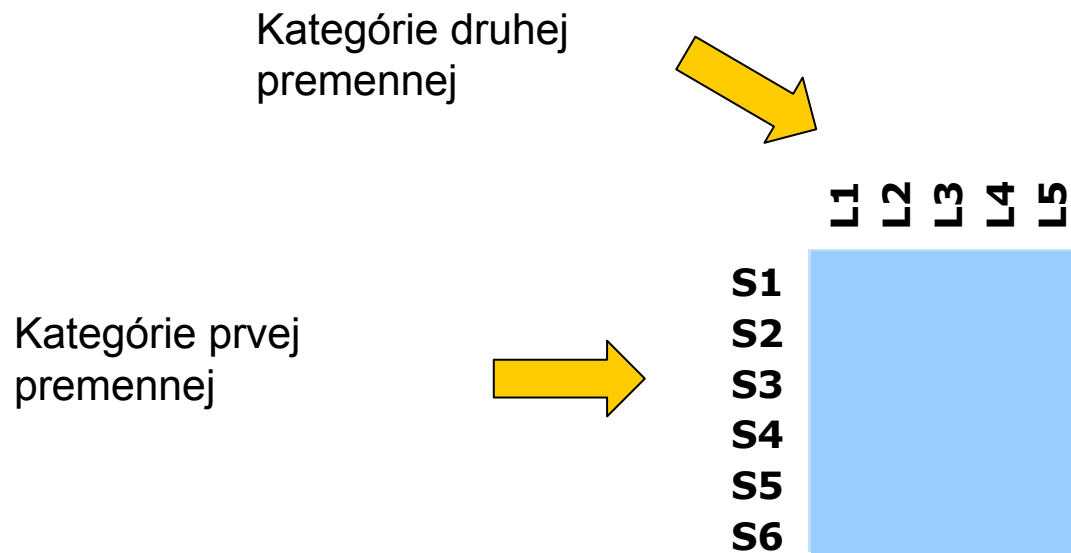
Korešpondenčná analýza (CA) a detrendovaná korešpondenčná analýza (DCA)

Danka Némethová

28. – 30. 5. 2007

Korešpondenčná analýza

- ◆ Korešpondenčná analýza - nástroj pre analýzu vzťahov medzi riadkami a stĺpcami kontingenčných tabuliek
- ◆ Skúmanie vzťahov medzi dvoma premennými
- ◆ Kontingenčná tabuľka: frekvenčná tabuľka (dvojvstupná), ktorá zaznamenáva kumulatívne počtenosti dvoch nominálnych (kategoriálnych) premenných. Každý stĺpec a každý riadok tabuľky reprezentuje jednu kategóriu danej premennej.



Korešpondenčná analýza

Korešpondenčná analýza všeobecne:

- ◆ Základnou myšlienkou metódy korešpondenčnej analýzy je vytvoriť či odvodiť indexy (pokiaľ možno „jednoduché“), ktoré budú nejakým spôsobom označovať (kvantifikovať) vzťahy medzi riadkovými a stĺpcovými kategóriami. Z týchto indexov potom budeme schopní odvodiť, ktorá stĺpcová kategória má väčšiu či menšiu váhu v danom riadku a naopak.
- ◆ Korešpondenčná analýza sa tiež vzťahuje k otázke zníženia dimenzionality dát podobne ako napr. analýza hlavných komponentov (principal component analysis: PCA) a k snahe o dekompozíciu tabuľky na faktory.
- ◆ Grafické znázornenie vzťahov, ktoré obdržíme z korešpondenčnej analýzy, je založené na myšlienke reprezentovať všetky stĺpce a riadky a interpretovať relatívne pozície bodov ako váhy prislúchajúce danému stĺpcu a riadku. Systém indexov, ktorý si pomocou tejto metódy odvodíme, nám teda bude poskytovať súradnice každého stĺpca a riadku. Tieto súradnice zakreslíme do grafu, z ktorého môžeme poznať, ktoré stĺpcové kategórie sú viac dôležité v riadkových kategóriách a naopak.

Korešpondenčná analýza

Korešpondenčná analýza v synekológii:

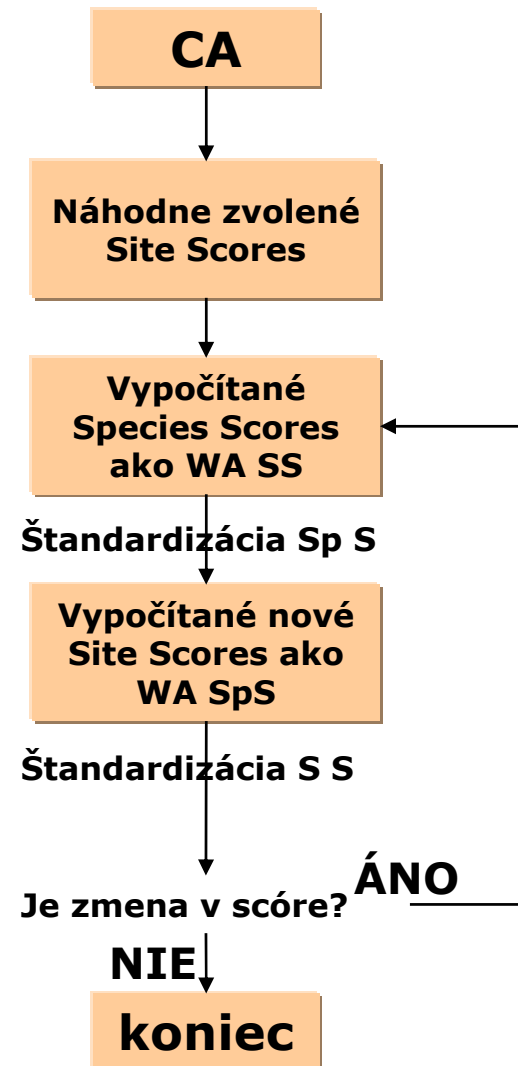
- ◆ Najjednoduchšou cestou ako odhadnúť optimum druhu pre unimodálny model je spočítať vážený priemer tých hodnôt charakteristík prostredia, pri ktorých sa druh vyskytuje.
- ◆ Ako váha sa pri výpočte používa početnosť ci iná dôležitostná hodnota druhu.
- ◆ Pri váženom priemerovaní je implicitne zahrnutá štandardizácia po vzorkách aj po druhoch.

Korešpondenčná analýza v ekológii spoločenských

Korešpondenčná analýza:

reciprocal averaging or eigenanalysis

	Samp1	Samp2	Samp3	WA1	WA2	WA3	WA4
Cirsium	0	0	3	13.000	10.000	10.000	10.000
Glechoma	5	2	1	4.625	1.363	1.312	1.310
Rubus	6	2	0	3.250	0.113	0.062	0.060
Urtica	8	1	0	2.556	0.050	0.028	0.027
<i>initial value</i>	2	7	13				
WA1	3.319	3.661	10.906				
WA1resc.	0.000	0.450	10.000				
WA2	0.415	0.600	7.841				
WA2resc.	0.000	0.249	10.000				
WA3	0.377	0.555	7.828				
WA3resc.	0.000	0.240	10.000				
WA4	0.375	0.553	7.827				
WA4resc.	0.000	0.239	10.000				



Korešpondenčná analýza v ekológii spoločenskíev

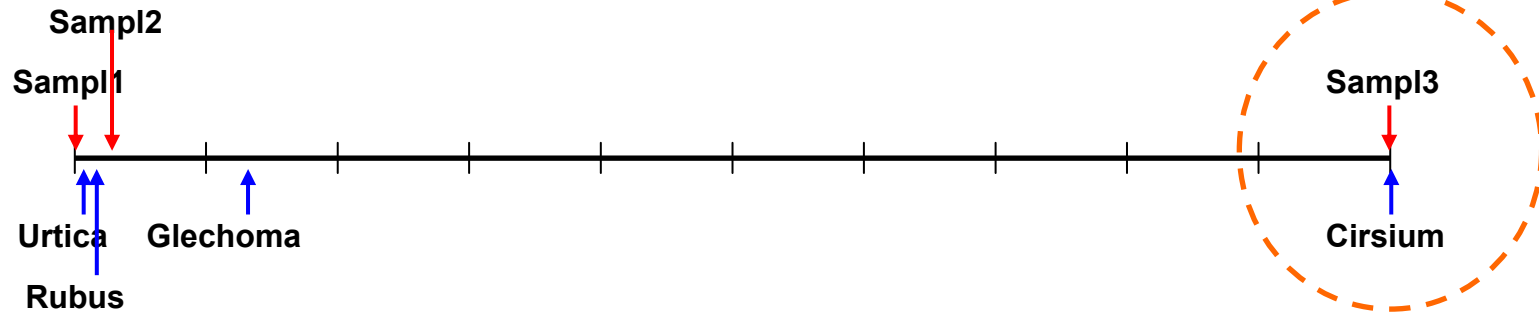
Korešpondenčná analýza:

reciprocal averaging or eigenanalysis

	Samp1	Samp2	Samp3	WA4
Cirsium	0	0	3	10.000
Glechoma	5	2	1	1.310
Rubus	6	2	0	0.060
Urtica	8	1	0	0.027

WA4resc. 0.000 0.239 10.000

odľahlá hodnota
(outlier)



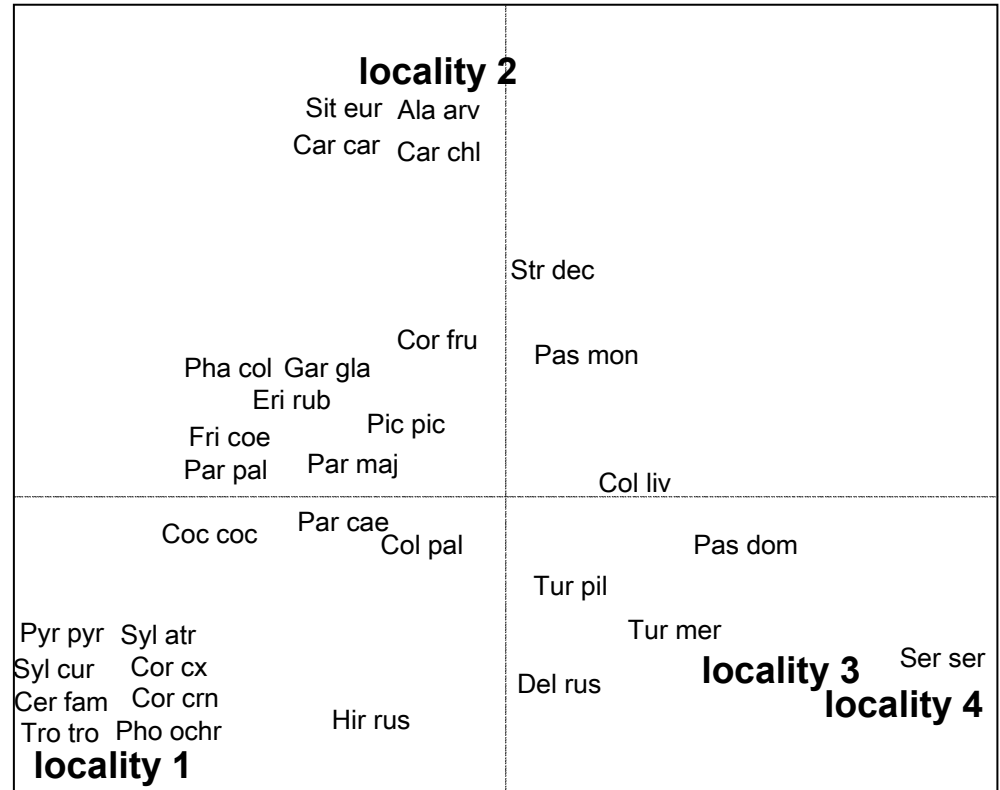
Korešpondenčná analýza: výsledky

- ◆ Ordinačný diagram
- ◆ Skóre druhov a lokalít (riadkov a stĺpcov)
- ◆ Charakteristické vektory a charakteristické čísla matice (eigenvalues, eigenvector)

 Vysoké skóre: druh s nízkou frekvenciou

Charakteristické číslo (eigenvalue) odpovedá časti variability súboru vysvetlenej danou osou.

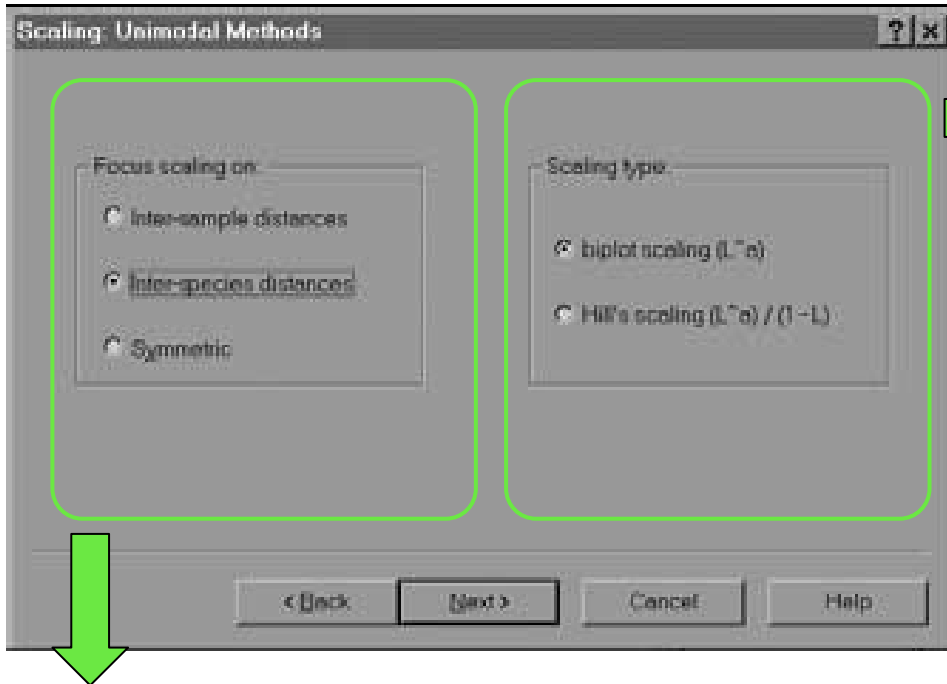
Väčšinou používame prvé dva – tri charakteristické vektory = ordinačné osi.



Ordinačné osi sú na sebe lineárne nezávislé.

Korešpondenčná analýza (CA)

Nastavenie škálovania



Typ škálovania určuje, ako sa pozerať na druhové dáta pri diagrame druhy+vzorky.

Biplot scaling je vhodnejší pre kratšie gradienty.

Hillovo škálovanie zjednocuje šírky ník pre všetky osi.

V prvom rade sa rozhodneme, či sa pri interpretácii zameriame na vzorky (porovnanie tried vzoriek, apod.) alebo druhy.

Ak máme charakteristiky prostredia, prípadne kovariáty, *species scaling* umožňuje charakterizovať korelácie medzi charakteristikami prostredia.

Korešpondenčná analýza (CA)

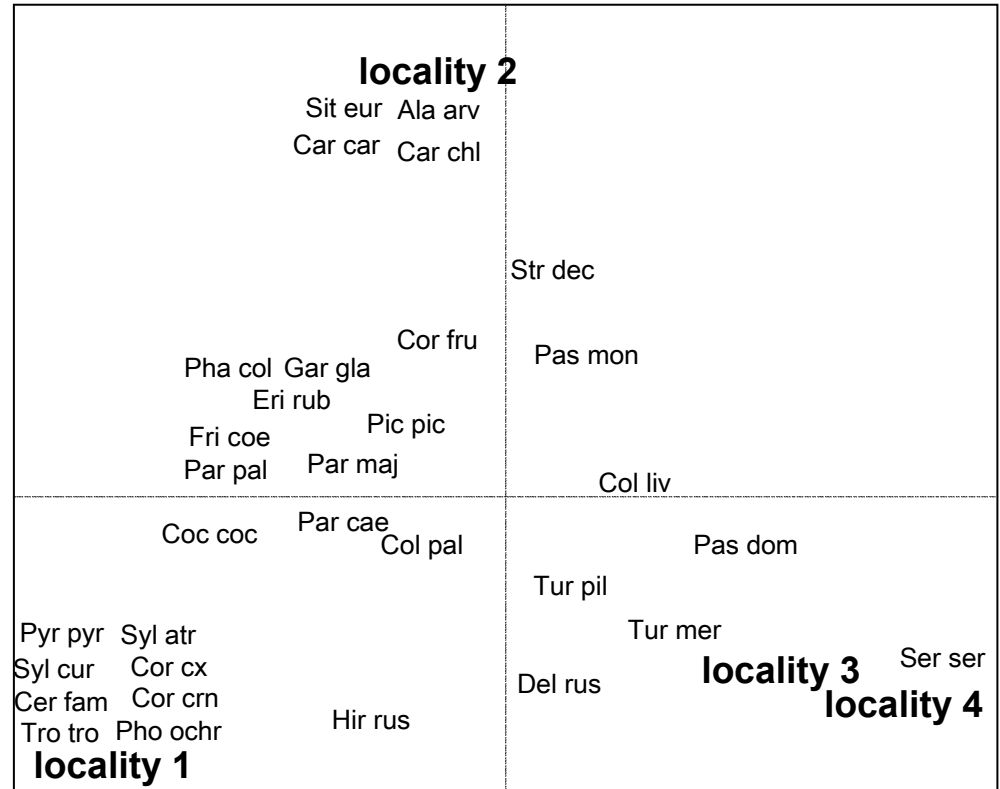
Indirect gradient analysis

Correspondence analysis

- ◆ CA je postavená na unimodálnom modeli; každý druh sa vyskytuje v ohraničenom rozsahu hodnôt každého environmentálneho gradientu
- ◆ CA je odporúčaná pre druhové dáta, ktoré obsahujú mnoho nulových hodnôt

REÁLNE DÁTA

- ◆ vtáacie druhy na 4 lokalitách
- ◆ dátová matica: 4 lokality x 38 dr. vtákov
hodnoty = priemerná abundancia



Korešpondenčná analýza: „arch effect“

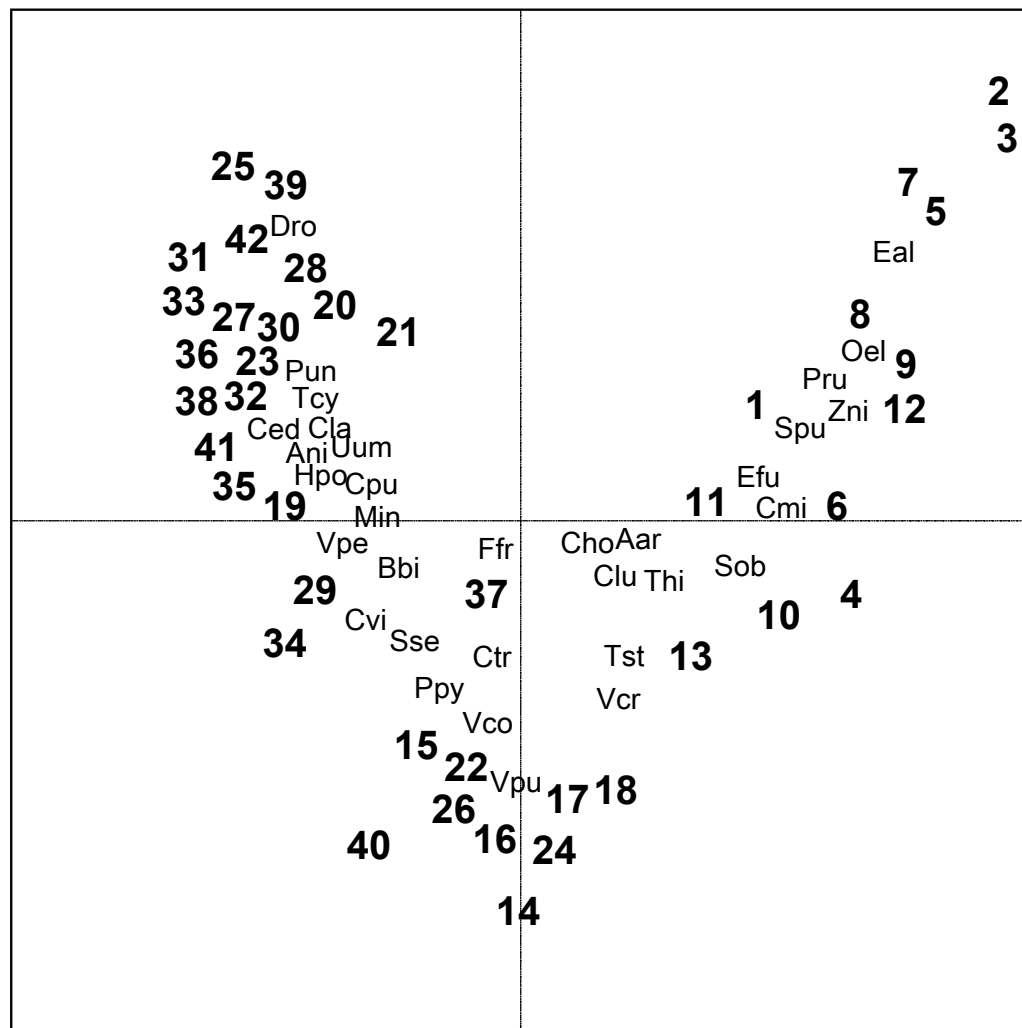
Indirect gradient analysis

Correspondence analysis

- ◆ CA je postavená na unimodálnom modeli
- ◆ pri silnej unimodálnej odozve sa v ordinačnom diagrame CA zvykne ukázať tzv. „arch effect“
- ◆ „arch effect“ môžeme odstrániť použitím detrendovanej formy CA

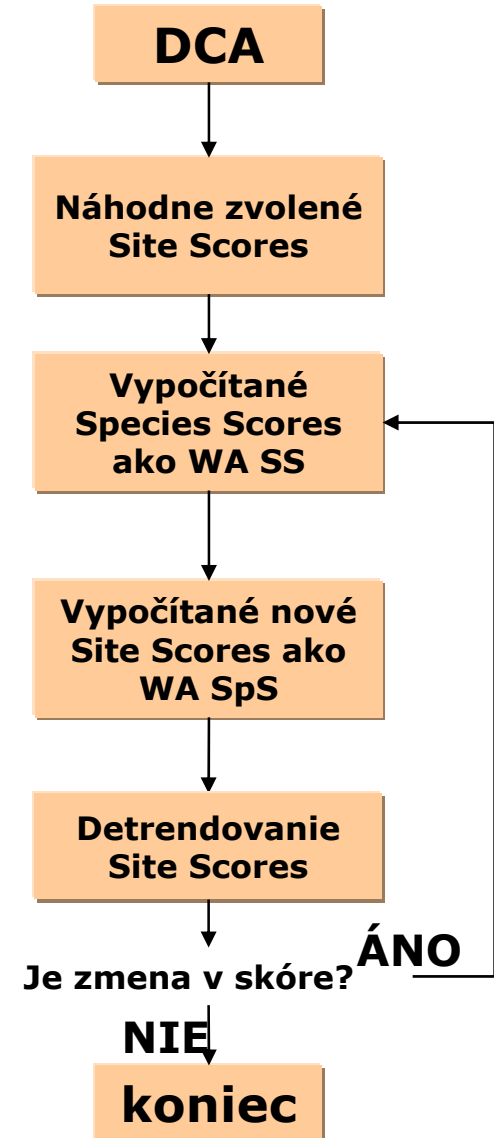
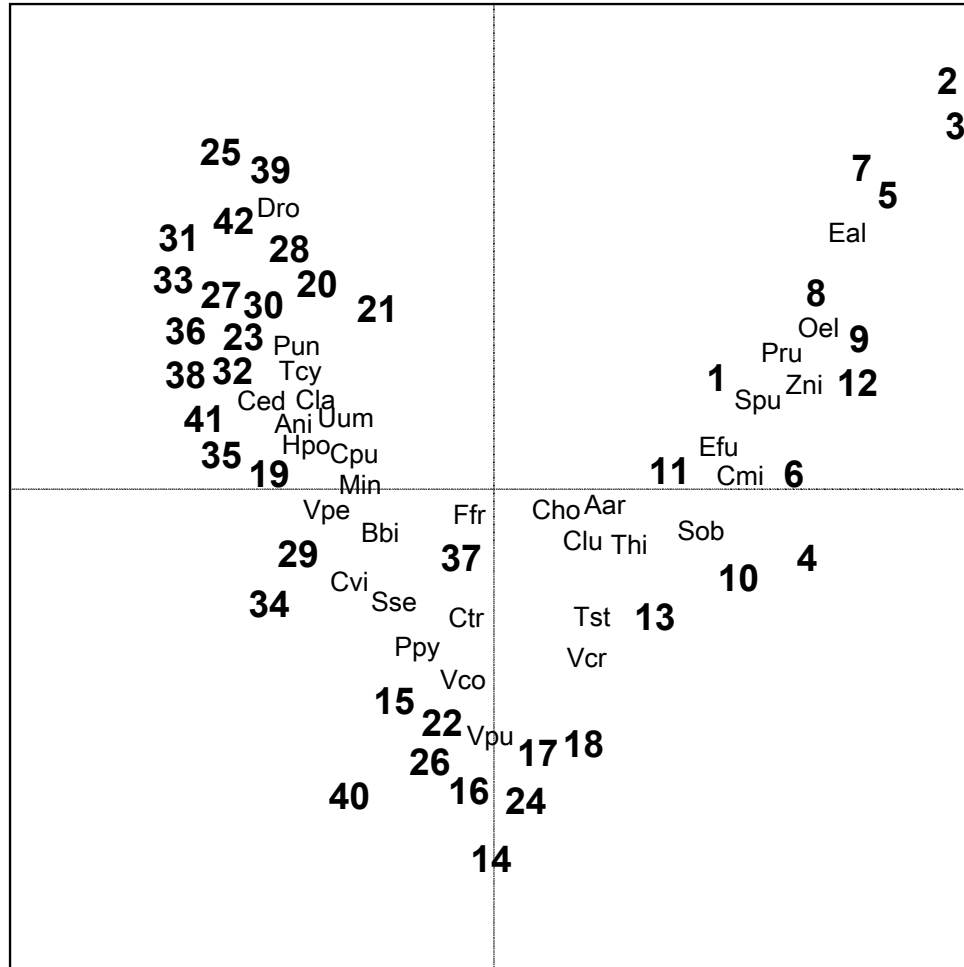
REÁLNE DÁTA

- ◆ suchozemské slimáky
- ◆ dátová matica: 42 lokalít x 33 dr. slimákov
hodnoty = stupeň dominancie



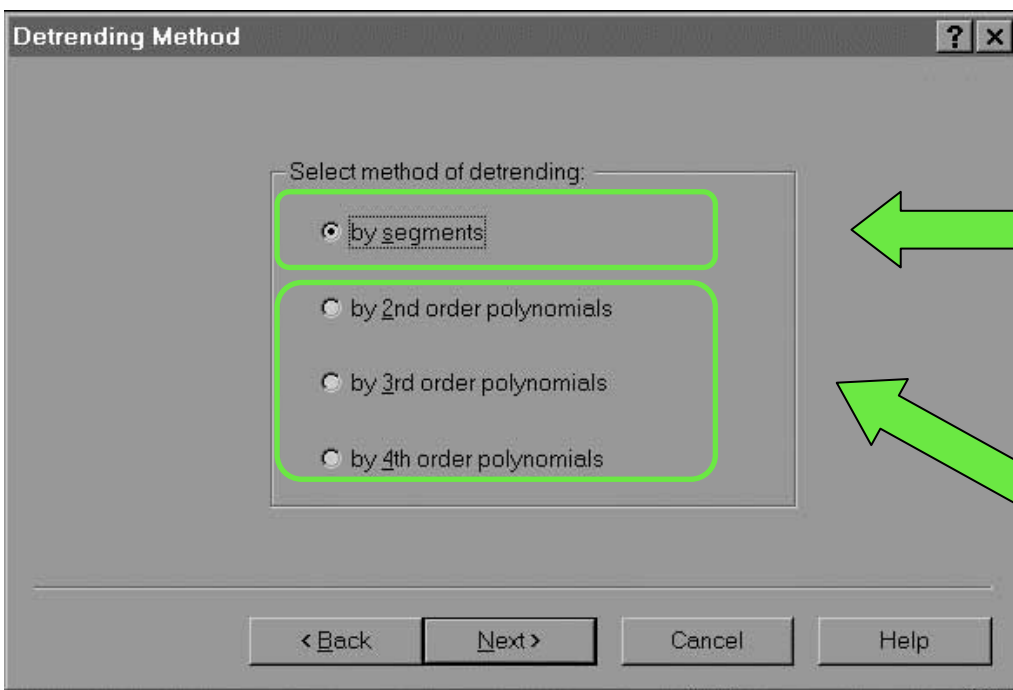
Korešpondenčná analýza: „arch effect“

„arch effect“, „horse shoe effect“



Detrendovaná korešpondenčná analýza (DCA)

Odstraňovanie trendu



odstraňovanie trendu po segmentoch

- ◆ neodporúča sa pre unimodálne ordinačné metódy, kde sú používané kovariáty alebo charakteristiky prostredia

odstraňovanie trendu polynómami

- ◆ keď sú používané kovariáty alebo charakteristiky prostredia a je potrebné odstániť trend

- ◆ Pre unimodálne ordinácie s obmedzením (CCA) obvykle nie je detrendovanie nutné. Ak sa v CCA oblúkový efekt objaví, je to známkou nadbytočnosti v súbore zvolených charakteristík prostredia.
- ◆ Doporučuje sa vylúčiť silne korelované premenné. Výber charakteristík prostredia, ktoré sú medzi sebou korelované len minimálne, sa dá previesť postupnou selekciou charakteristík prostredia (*forward selection of environmental variables*).

Detrendovaná korešpondenčná analýza (DCA)

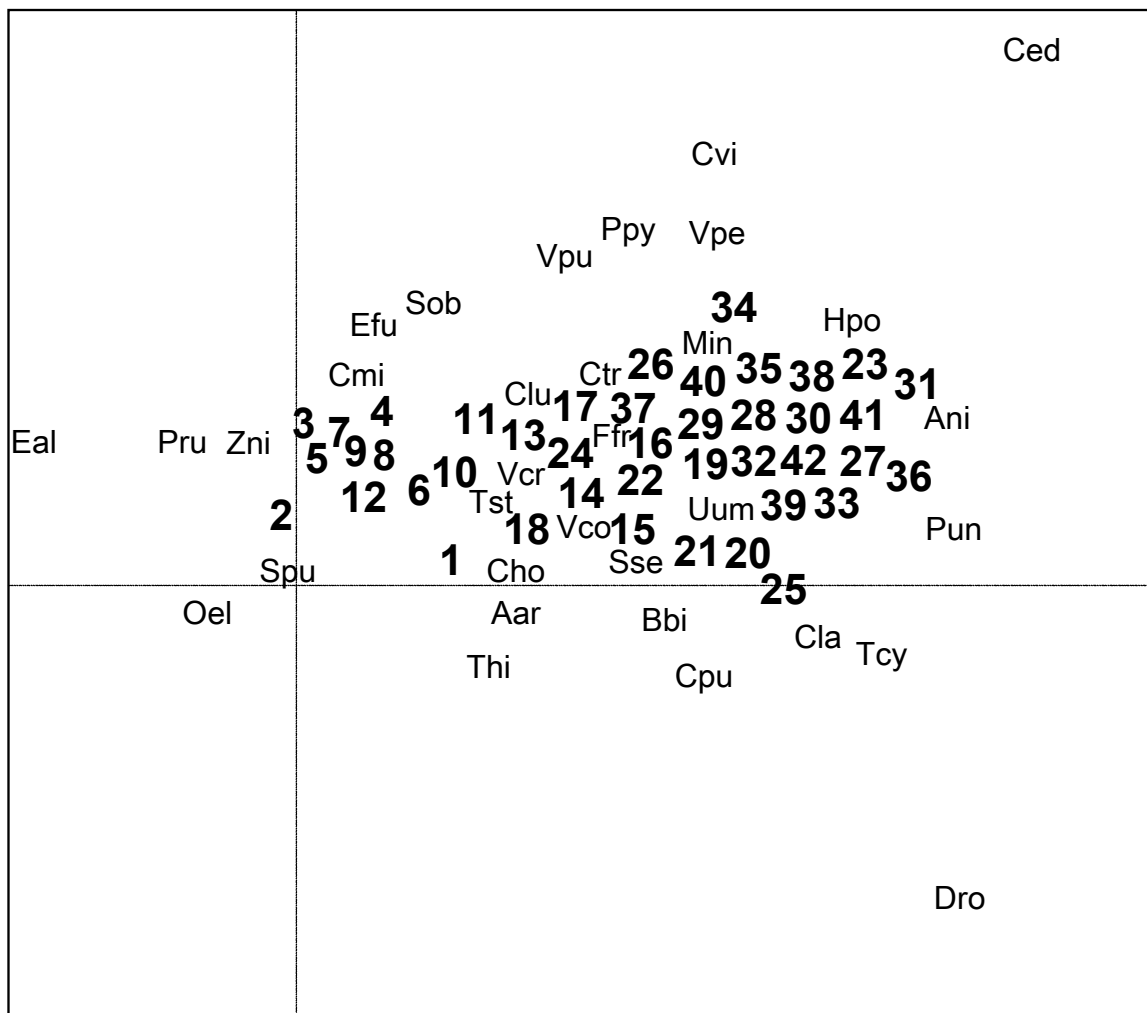
Indirect gradient analysis

Detrended correspondence analysis

- ◆ DCA je postavená na unimodálnom modeli
- ◆ DCA odstraňuje „arch effect“ niekoľkými možnými spôsobmi

REÁLNE DÁTA

- ◆ suchozemské slimáky
 - ◆ dátová matica: 42 lokalít x 33 druhov slimákov
- hodnoty = stupeň dominancie



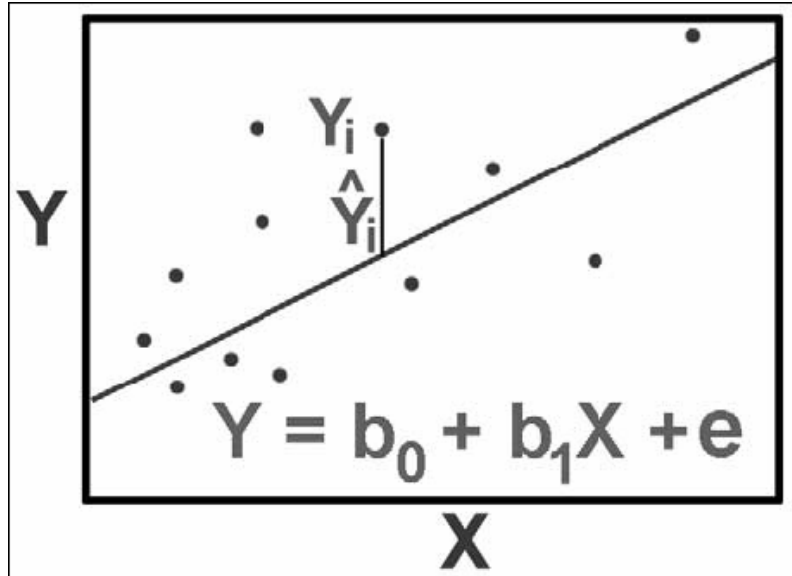
Priame ordinácie – ordinácie s obmedzením

Danka Némethová

28. – 30. 5. 2007

Priama gradientová analýza

Grafické znázornenie jednoduchého lineárneho regresného modelu



Y závislá premenná (vysvetľovaná) nezávislá
 X premenná (vysvetľujúca)

regresný reziduál, označený ako e : rozdiel medzi (pozorovanými) hodnotami vysvetľovanej premennej Y a hodnotami predpovedanými modelom (očakávané hodnoty, Y so strieškou).

Všetky štatistické modely majú dve dôležité zložky:

- 1. systematická** – časť variability vysvetľovaných premenných, ktorú môžeme vysvetliť vysvetľujúcimi premennými (prediktormi) pomocou zvolenej parametrickej funkcie.
- 2. stochastická** – ostávajúca časť variability hodnôt vysvetľovanej premennej, ktorú nemožno predpovedať systematickou časťou modelu. Definuje sa pomocou predpokladaných pravdepodobnostných a distribučných vlastností.

Priama gradientová analýza

Regresný model

- ◆ Kvalitu modelu posudzujeme podľa množstva variability popísanej systematickou zložkou (obvykle v pomere k stochastickej zložke).

Regresný model s viacerými premennými

- ◆ Možnosť postupného výberu významných premenných
- ◆ Začíname s nulovým modelom bez prediktorov, predpokladáme, že variabilitu vysvetľovanej premennej nejde predpovedať, a popisuje ju len stochastická zložka. Potom vyberieme z dostupných premenných jediný prediktor – ten, ktorý v regresnom modeli vysvetľuje najviac variability.
- ◆ Aj keď zvolíme ten najlepší prediktor, môže byť jeho príspevok len náhodný => testovanie (prehádzanie hodnôt tohto prediktoru ...)
- ◆ Postupné testovanie všetkých premenných; končíme keď „najlepší“ z ostávajúcich kandidátov už nie je „dostatočne dobrý“.

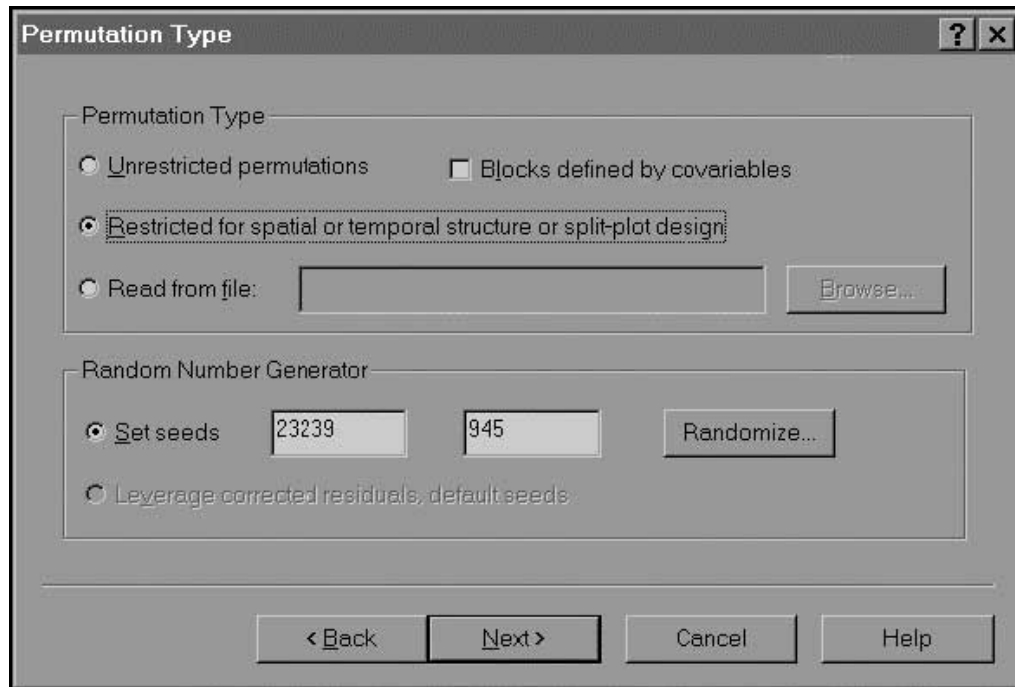
Priama gradientová analýza

Priama gradientová analýza (*direct gradient analysis; constrained, canonical ordination methods*) – kombinácia ordinácie a regresie

- ◆ Nepriame gradientové analýzy hľadali teoretické gradienty, ktoré boli „optimálnymi“ prediktormi v regresných modeloch lineárnej či unimodálnej odpovedi druhov.
- ◆ Metódy priamej gradientovej analýzy sa snažia o to isté, ale gradienty, ktoré je týmto metódam „dovolené nájsť“, sú viac obmedzené. Tieto gradienty sú lineárnou kombináciou predložených vysvetľujúcich premenných (charakteristík prostredia). Abundanciu jednotlivých druhov sa snažíme vysvetliť pomocou zložených premenných, ale tieto premenné sú definované na základe hodnôt pozorovaných charakteristík.
- ◆ Metódy priamej gradientovej analýzy sa podobajú mnohorozmernej násobnej regresii.
- ◆ V priamej gradientovej analýze: vplyv prediktorov na vysvetľované premenné cez niekoľko „zprostredkujúcich“ gradientov – kanonických ordinačných osí (*canonical axes, constrained axes*).
- ◆ Existuje tu toľko kanonických osí, koľko je nezávislých vysvetľujúcich premenných.

Výber štatisticky významných premenných

Výber štatisticky významných premenných: permutačný test



Monte-Carlo permutačný test: testuje štatistickú významnosť obmedzených ordinačných modelov

H₀: primárne (druhové) dáta sú nezávislé na vysvetľujúcich premenných

- ◆ rôzne spôsoby nastavenia testu pre dáta s určitou priestorovou, časovou a logickou vnútornou štruktúrou, v závislosti na usporiadaní pokusu a odbere vzorky

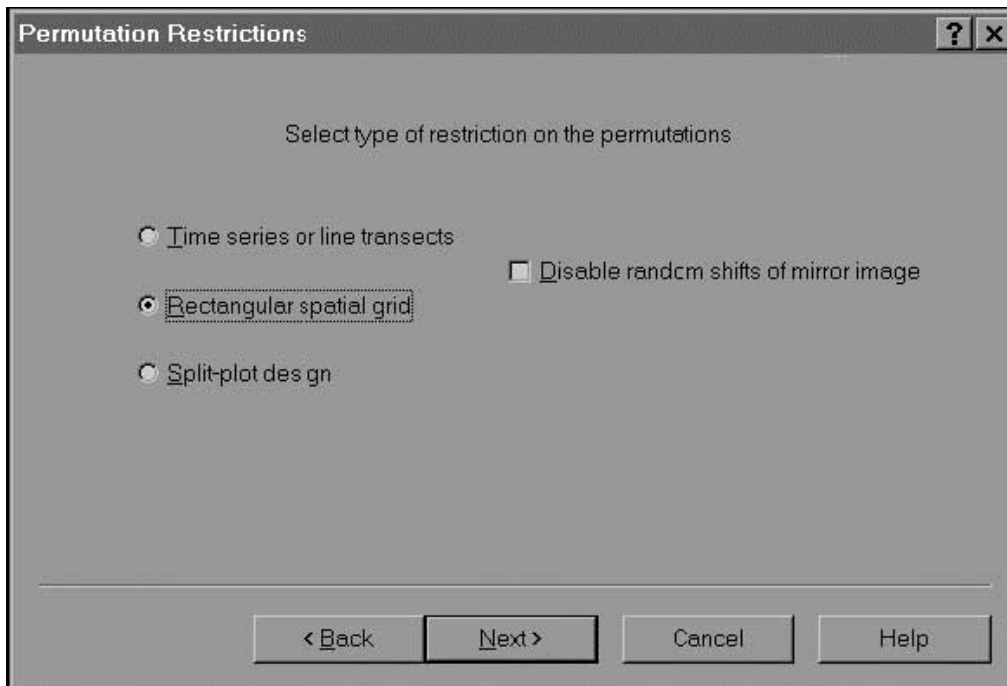
Výber štatisticky významných premenných

Permutačný test

- ◆ permutované hodnoty premennej – vytvorenie niekoľkých permutácií (náhodné prehodenie hodnôt premennej medzi vzorkami) – testovanie rozdielu od pôvodnej premennej

Priestorové a časové obmedzenia

- ◆ ak je v dátach vnútorná štruktúra použijeme pri permutáciach obmedzenie

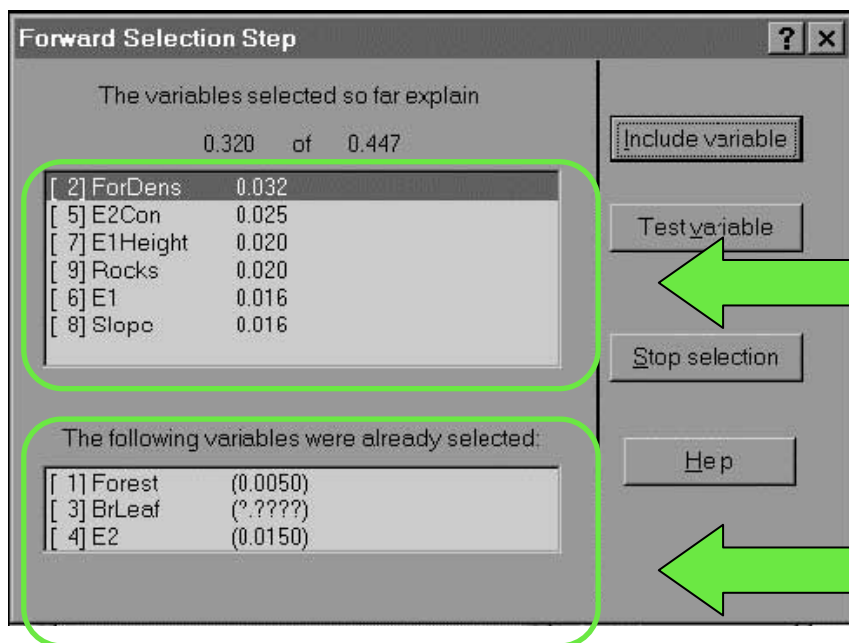


- ◆ vzorky pozdĺž časového alebo lineárneho transektu => permutácie „rotovaním“
- ◆ **split-plot design**
permutácie v rámci bloku – ten je charakterizovaný niekoľkými nominálnymi premennými

Výber štatisticky významných premenných

Permutačný test

- ◆ Ak použijeme manuálne permutačné testovanie – vidíme priebeh testovania po krokoch.



Kandidáti na prediktory

Vybrané charakteristiky prostredia

Testovanie významnosti priamej ordinácie

Permutačný test

- ◆ Testovanie významnosti prvej kanonickej ordinačnej osi: Monte-Carlo permutačný test
- ◆ Vhodný typ permutácií je určený typom experimentálneho designu a designu vzorkovania (možnosti permutačných testov pre split-plot designs a iné multi-level designs)
- ◆ **Global permutation test – Both above tests**
vykonajú sa dva Monte-Carlo testy:
 1. test významnosti prvej kanonickej osi
 2. test významnosti všetkých kanonických osí
- ◆ Testovať významnosť ordinačnej osi v nepriamych analýzach nie je možné.
- ◆ Testovať môžeme aj vplyv environmentálnych premenných po odčítaní kovariátov (parciálny test)

Redundančná analýza (RDA)

Danka Némethová

28. – 30. 5. 2007

Redundačná analýza (RDA)

Direct gradient analysis

Redundancy analysis

- ◆ RDA nie je vhodná pre druhové dáta, v ktorých sa vyskytuje mnoho nulových hodnôt

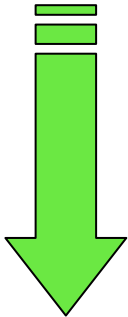
RDA je kanonická forma analýzy hlavných komponent (PCA)

- ◆ V obmedzenej metóde (RDA) podliehajú **skóre objektov** (vzoriek) obmedzujúcej podmienke: definujú sa ako **lineárna kombinácia vysvetľujúcich premenných**

Redundačná analýza (RDA)

Principal component analysis (PCA)

PCA ... regresia



Redundancy analysis (RDA)

RDA ... mnohonásobná regresia

- ◆ Abundancia každého druhu je modelovaná lineárnou regresiou podľa vysvetľujúcej premennej, ktorej hodnoty sú neznáme (neznáme x ; teoretický gradient, prvá hlavná komponenta).
- ◆ RDA obmedzuje hodnoty tak, že požaduje, aby x bolo lineárnou kombináciou meraných charakteristík prostredia.
- ◆ RDA je mnohonásobnou regresiou pre všetky druhy súčasne s lineárnym obmedzením regresných koeficientov.

Supplementary species, samples, variables

- ◆ Tzv. **suplementárne** druhy, vzorky, charakteristiky prostredia (v staršej verzii Canoca označované ako **pasívne**) sa odlišujú od aktívnych tým, že neovplyvňujú tvorbu ordinačných osí.
- ◆ Môžu byť však pridané do existujúcej ordinácie (napr. regresným modelovaním ich dát na existujúce ordinačné osi).
- ◆ Druhy a vzorky, ktoré majú byť pasívne, musia byť pripravené v matici druhových dát.
- ◆ Charakteristiky prostredia, ktoré majú byť pasívne, musia byť pripravené v samostatnom súbore.

Kanonická korešpondenčná analýzy (CCA)

Danka Némethová

28. – 30. 5. 2007

Kanonická korešpondenčná analýza (CCA)

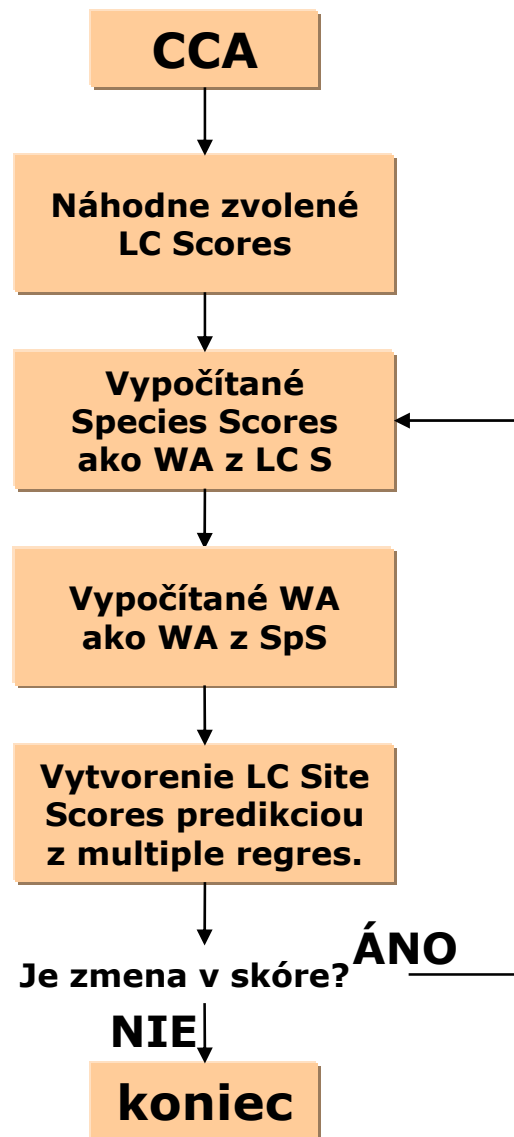
CCA je obmedzená ordinácia

- ◆ druhové dáta + vysvetľujúce premenné
- ◆ len „zmysluplné“ vysvetľujúce premenné

- ◆ Forward selection:

Permutačný test H_0 :

Vysvetľovacia sila skupiny environmentálnych premenných sa pridaním danej premennej nezvýši viac, než keby sme pridali takú premennú, ktorá má rovnaké distribučné vlastnosti ako uvažovaná premenná, ale nemá žiadny vzťah k druhovým dátam.



Kanonická korešpondenčná analýza (CCA)

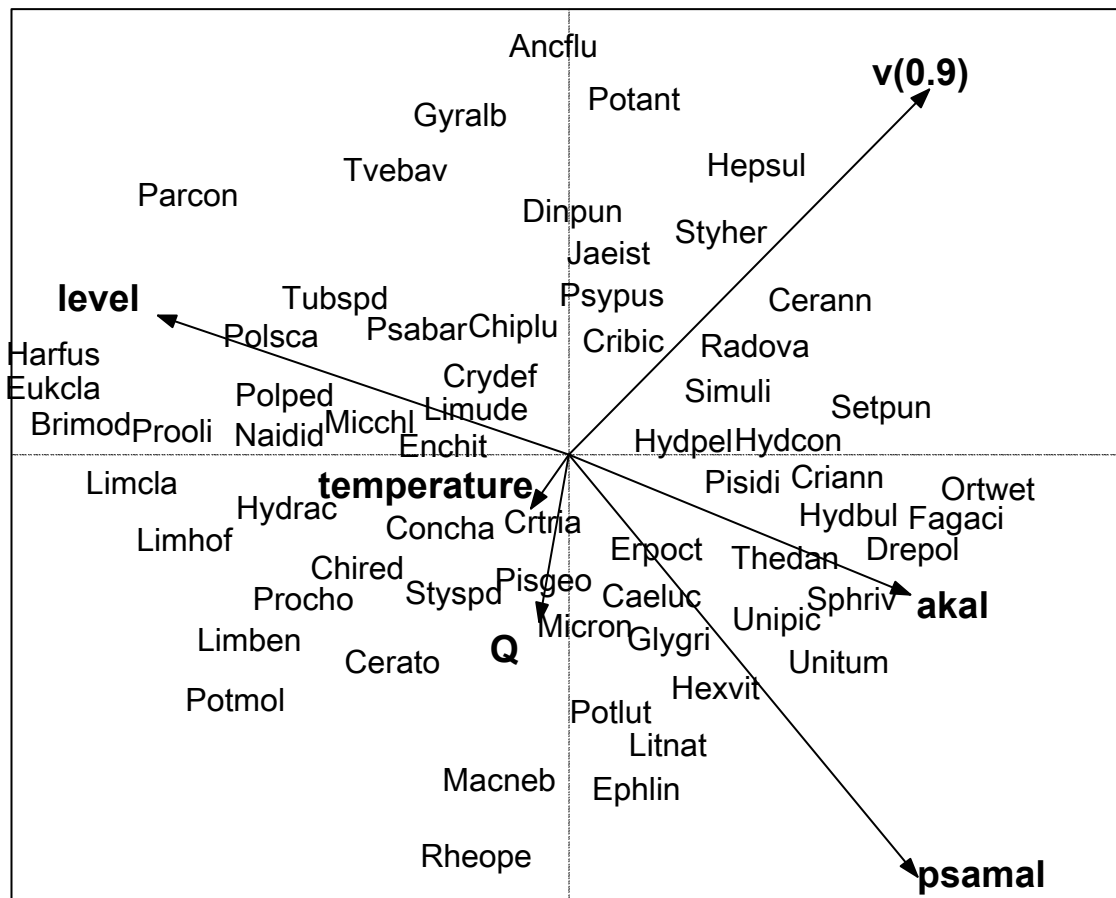
Direct gradient analysis

Canonical correspondence analysis

- ◆ CCA je kanonická forma CA
- ◆ CCA sa odporúča pre druhové dáta s veľkým výskytom nulových hodnôt

REÁLNE DÁTA

- ◆ spoločenstvá makrozoobentosu
- ◆ dátové matice:
60 lok. x 63 tax. (stupeň dominancie)
60 lok. x 13 environm. faktorov (fs)



Parciálne ordinácie

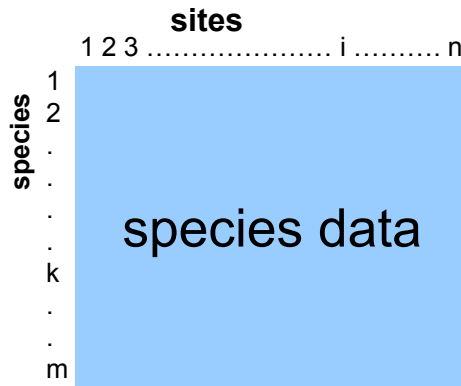
Danka Némethová

28. – 30. 5. 2007

Parciálna ordinácia

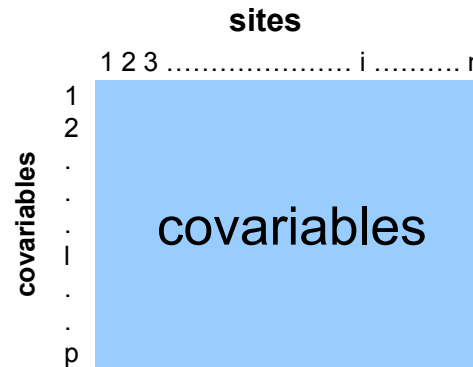
Indirect gradient analysis

Druhové dáta



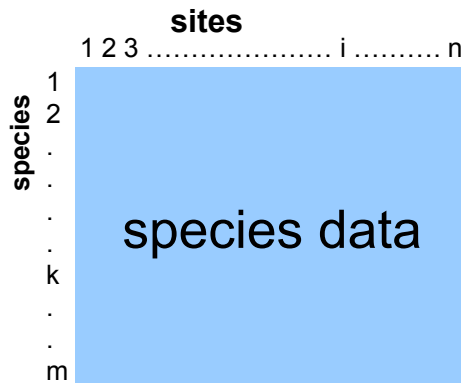
+

Kovariáty



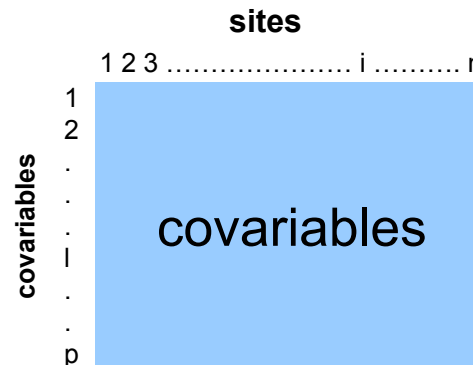
Direct gradient analysis

Druhové dáta



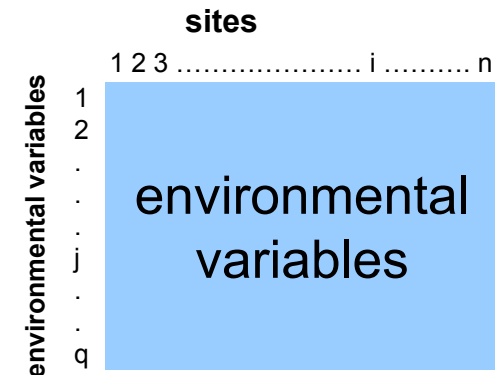
+

Kovariáty



+

Charakteristiky prostredia



Parciálna ordinácia

Parciálne ordinácie

Pre všetky metódy je možné použiť dielčie (parciálne) analýzy. V parciálnych analýzach je najprv oddelený vplyv kovariát a analýza je potom prevedená len na zostávajúcej variabilite.

Dátové zdroje:

Principal component analysis (PCA)

Correspondence analysis (CA)

Dentrended correspondence analysis (DCA)

druhovú dáta + kovariáty

Redundancy analysis (RDA)

Canonical correspondence analysis (CCA)

druhovú dáta + charakteristiky prostredia + kovariáty

Diskriminačná analýza (CVA, DFA)


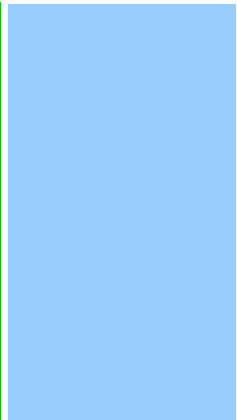
Danka Némethová

28. – 30. 5. 2007

Diskriminačná analýza

Diskriminačná analýza

- ◆ sa snaží zjednodušiť viacrozmernú štruktúru dát výpočtom súhrnných osí (diskriminačnej funkcie).
- ◆ Je jednou z metód **ordinácie**.
- ◆ Logika osí v diskriminačnej analýze je, že maximálne diskriminuje skupiny.

	skupina	deskriptor 1	deskriptor 2	deskriptor 3
vzorka 1				
vzorka 2				
vzorka 3				
vzorka 4				
vzorka 5				
vzorka 6				

- ◆ Zaoberá sa závislosťou jednej **kvalitatívnej premennej** na niekoľkých kvantitatívnych premenných.
- ◆ Objekty charakterizované sériou deskriptorov (parametrov) – kvantitatívne parametre. Známa príslušnosť objektov do skupín.

Diskriminačná analýza

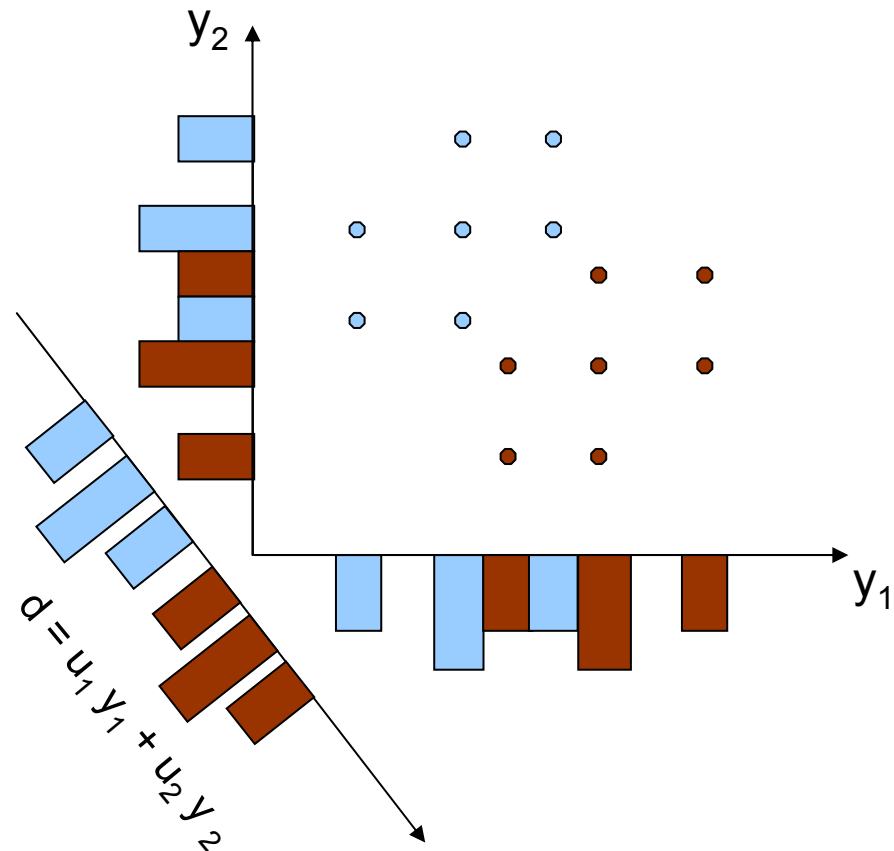
- ◆ Analýza nachádza takú kombináciu vstupných parametrov, ktorá oddeľuje od seba skupiny objektov

Skupina	y_1	y_2
A	3	5
A	3	7
A	5	5
A	5	7
A	5	9
A	7	7
A	7	9
B	6	2
B	6	4
B	8	2
B	8	4
B	8	6
B	10	4
B	10	6



Kvalitatívna
premenná

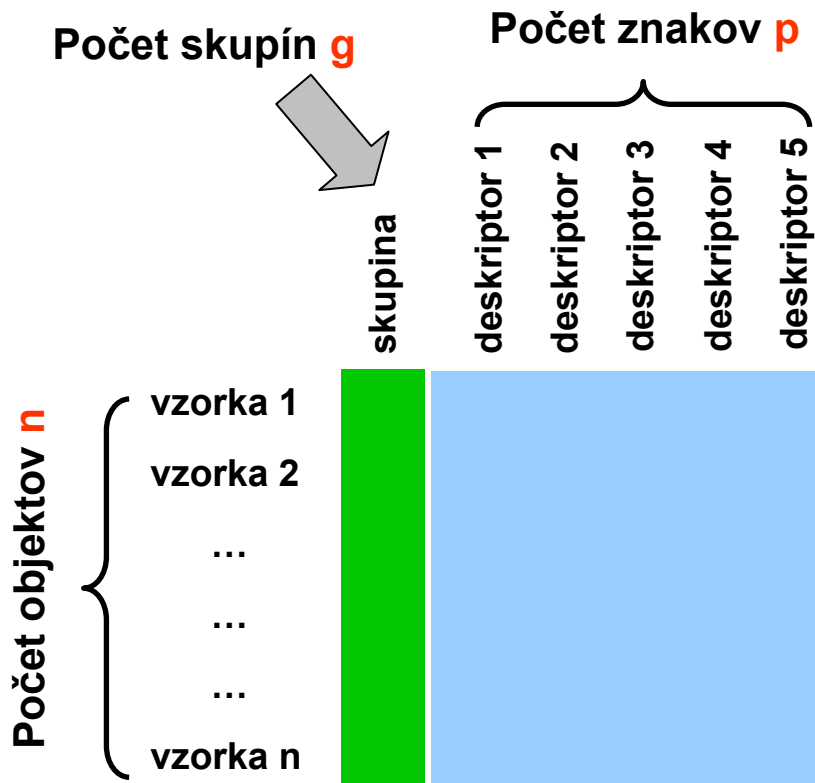
Kvantitatívne
premenné
(deskriptory)



Diskriminačná analýza

Pre počty skupín (g), znakov (p) a objektov (n) musí platiť:

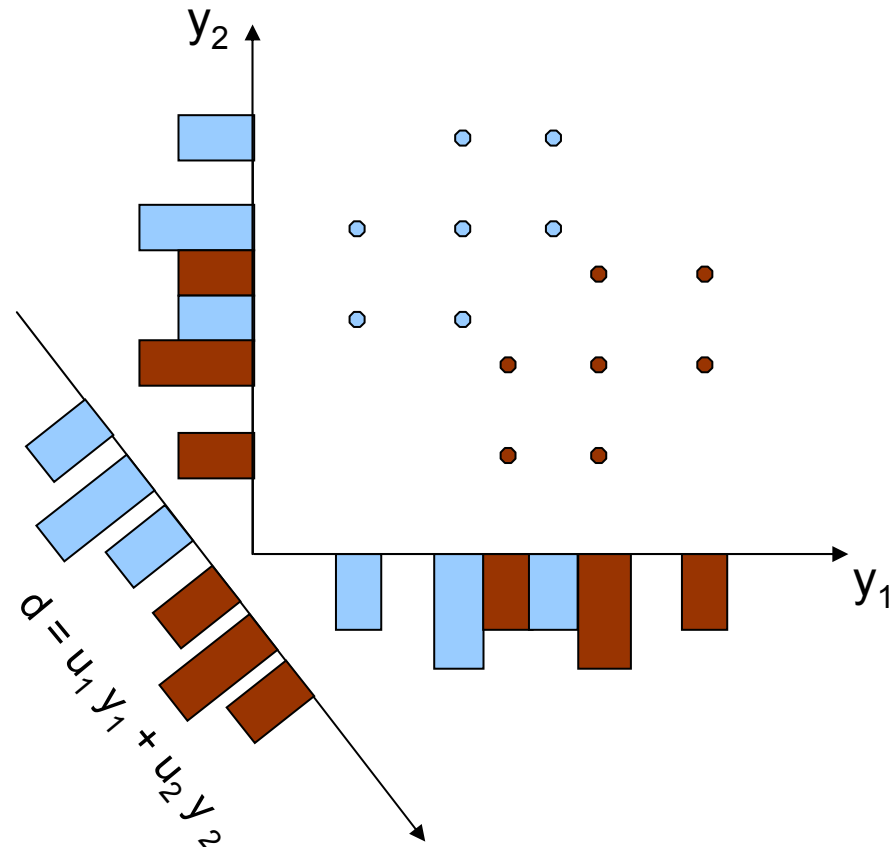
1. Musia byť aspoň dve skupiny objektov: $g \geq 2$
2. V každej skupine musia byť najmenej 2 objekty
3. Počet znakov použitých v analýze musí byť menší než počet objektov znížený o počet skupín: $0 < p < (n-g)$
4. Žiadny znak by nemal byť v niektorej skupine konštantný



Diskriminačná analýza

- ◆ Výsledkom diskriminačnej analýzy je diskriminačná funkcia (koeficienty deskriptorov).
- ◆ Premenné s najväčšími (štandardizovanými) koeficientami najviac prispievajú k predikcii príslušnosti do skupín.

Skupina	Raw coefficients	Standardized coefficients
y_1	-0.6124	-1.0
y_2	0.6124	1.0
konštanta	0.6124	
vlastná hodnota	3.9375	3.9375



- ◆ Počet diskriminačných funkcií je rovný počtu skupín znížený o jednu (prípadne počtu premenných, ak je tento nižší ako $g-1$)

Diskriminačná analýza

Vstup diskriminačnej analýzy:

- ◆ Tabuľka objektov charakterizovaných niekoľkými kvantitívnymi parametrami a jednou kvalitatívnou premennou (ktorá priraduje objektom príslušnosť ku skupine)

Výstup diskriminačnej analýzy:

- ◆ Ordinačný diagram (osami sú korene, čiže diskriminačné funkcie)
- ◆ Korene diskriminačnej analýzy (koeficienty diskriminačných funkcií)

Pri použití diskriminačnej analýzy je potrebné pamätať na obmedzenia:

- ◆ parametrická metóda; vyžaduje normálne rozdelenie deskriptorov
- ◆ problém odľahlých hodnôt
- ◆ závislé na rozložení premenných
- ◆ výsledky udáva v pravdepodobnostiach
- ◆ nie je schopná zachytiť nelineárne vzťahy medzi prediktormi
- ◆ nedá sa použiť na silne korelované prediktory

Diskriminačná analýza v Canoco

Canonical Variates Analysis (CVA), discriminant analysis (DFA)

Možnosť spočítať CVA v Canoco:

1. zvoliť kanonickú korešpondenčnú analýzu (CCA)
2. rozdelenie vzoriek do skupín vo forme druhových dát, ktoré sú binárne a charakterizujú príslušnosť vzorky ku skupine
3. charakteristiky prostredia ako environmentálne dáta
4. zvoliť Hillovo škálovanie so zameraním na inter-species distances

V súbore **.sol**:

- ◆ species scores sú stredmi zhhlukov v CVA ordinačnom diagrame
- ◆ sample scores, ktoré sú lineárnou kombináciou charakteristík prostredia sú škálované tak, že rozptyl v rámci skupín sa rovná 1

Permutačný test môže byť použitý na hodnotenie rozdielov medzi skupinami.

Je možné špecifikovať aj kovariáty => parciálna CVA = one-way Multivariate Analysis of Covariance (MANOCO).

Priame vs. nepriame ordinačné metódy

Danka Némethová

21. – 22. 5. 2007

Nepriama vs. priama gradientová analýza

Indirect gradient analysis

+ Druhové zloženie je ľahko determinovateľné a tak je lepším indikátorom prostredia ako akákoľvek kombinácia meraných environmentálnych premenných.

- Environmentálny gradient je možné charakterizovať len na základe druhových dát.

Direct gradient analysis

+ Priama gradientová analýza poskytuje súhrn vzťahov druh-prostredie.
Gradient je charakterizovaný pomocou env. premenných.

- Predpokladáme, že všetky druhy reagujú na zložený gradient env. premenných podľa rovnakého modelu odozvy.

- Environmentálne podmienky nie je možné vždy charakterizovať úplne – môže sa stať, že prehliadneme nejaký dôležitý faktor.

Neparametrická ordinácia (NMDS)

Danka Némethová

21. – 22. 5. 2007

Neparametrická ordinácia (NMDS)

Indirect gradient analysis

Multidimensional scaling

- ◆ mnohonásobné škálovanie sa používa ako prieskumná metóda
- ◆ cieľom analýzy je zobraziť pozorované podobnosti alebo nepodobnosti (vzdialenostiach) medzi skúmanými objektami v euklidovskom priestore
- ◆ pomocou NMDS môžeme analyzovať nielen korelačné matice (ako v PCA) ale aj hocikajú inú maticu podobnosti/nepodobnosti

+ neparametrická ordinácia je robustnejšia k vychýleným hodnotám (napr. druh s výnimočne vysokou abundanciou na lokalite v jednom roku)

dá sa použiť pred použitím nehierarchického zhukovania K-means (v prípadoch keď nie je možné použiť euklidovské vzdialenosti)

- počet dimenzií musí byť určený vopred
ťažko interpretovateľné výsledky