Ordinačné metódy

Danka Haruštiaková

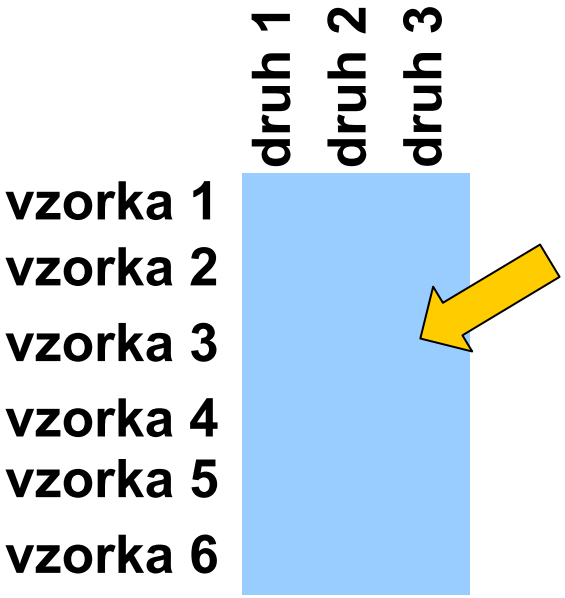
Podzim 2009

Inštitút bioštatistiky a analýz, Masarykova univerzita

Úvod

- **Mnohorozmerné metódy**:
  názov „mnohorozmerné“ – dáta sú tvorené objektami (vzorky, lokality), každý z nich je charakterizovaný viacerými parametrami (druhmi)
  každý z týchto parametrov môžme považovať za jeden rozmer objektu (vzorky)

DÁTOVA MATICA



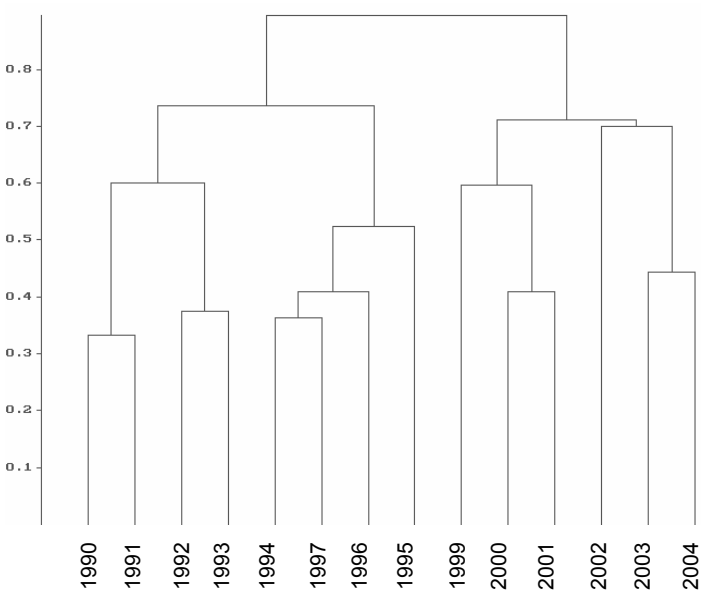
Hodnoty pre druhy (presencia/absencia; abundancia; dominancia) pre každú vzorku

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

Úvod

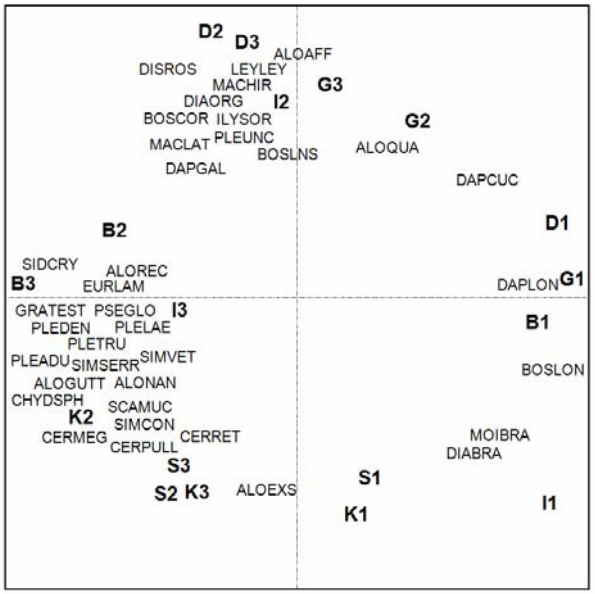
ZHLUKOVÁ ANALÝZA

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



ORDINÁCIA

- Usporadúva objekty pozdĺž trendu v dátach

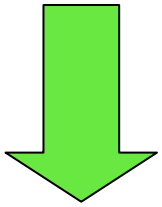


Úvod

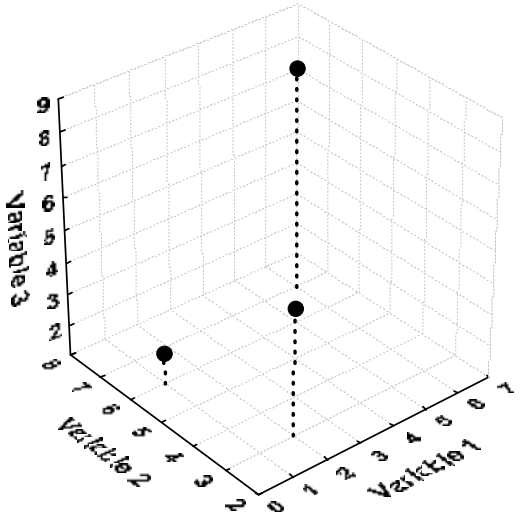
Objekty charakterizované *p* parametrami je možné si predstaviť ako body v *p* rozmernom priestore, kde každý z rozmerov predstavuje hodnoty jedného parametra. V prípade spoločenstiev sú objektami vzorky a parametrami druhy, prípadne charakteristiky prostredia.

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.

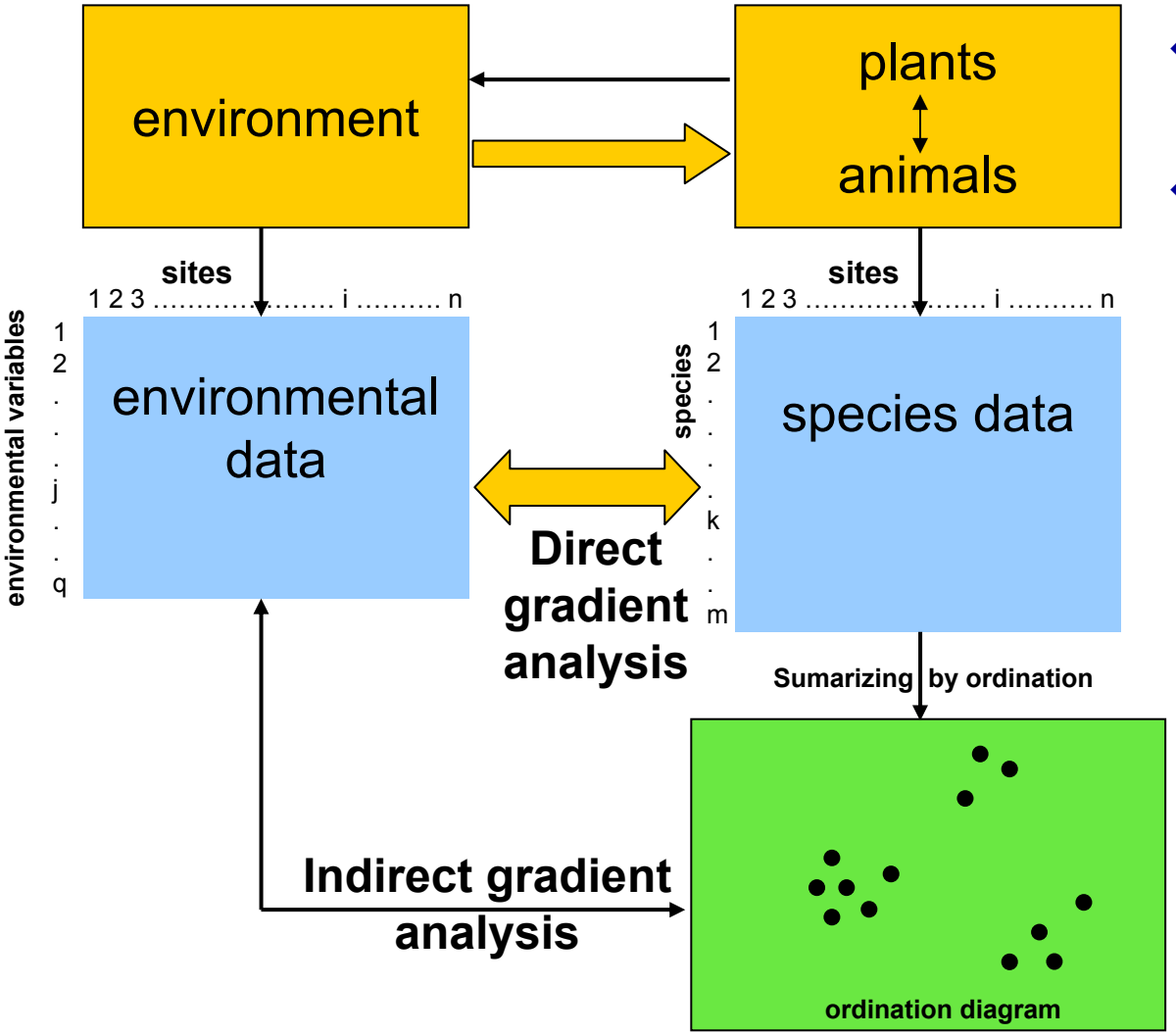
Pri väčšom počte parametrov je nutné redukovať ich počet s čo najmenšou stratou informácie.



Ordinačné metódy



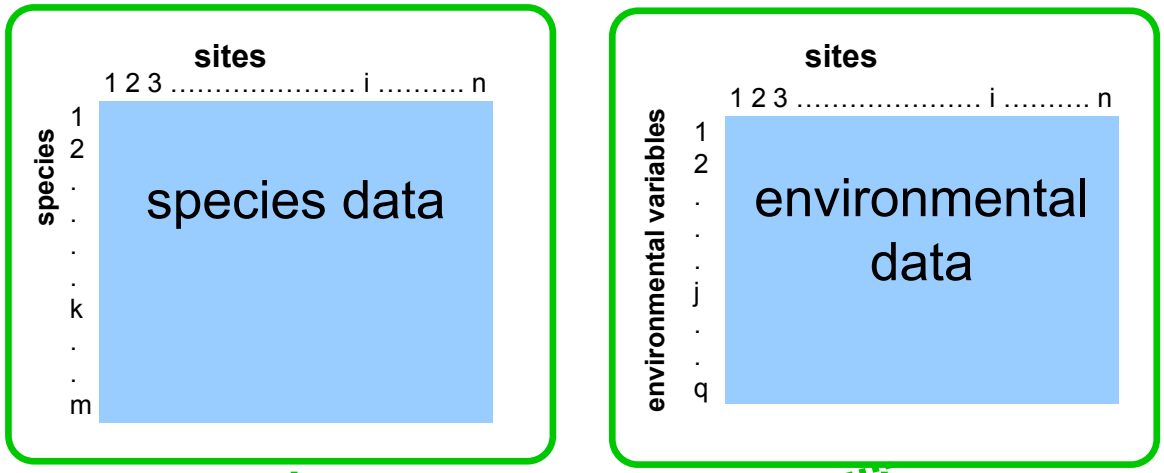
Ordinačné metódy v ekológii



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

- Ordinačné metódy nepredpokladajú žiadne apriorné zoskupenie objektov.
- Ordinačné metódy patria medzi metódy, ktoré sa pužívajú 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)
akékoľvek premenné, kt. hodnoty chceme predpovedať

Vysvetľujúce premenné
(*explanatory*)

Charakteristiky prostredia
(*environmental variables, variables*)

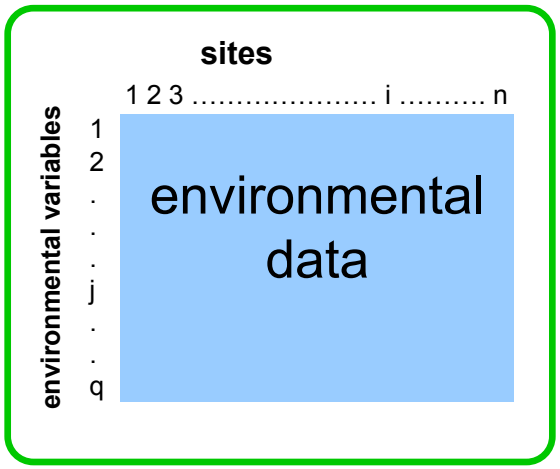
Kovariáty
(*covariates, covariables*)
ich vplyv chceme oddeliť

Napr. druhové zloženie spoločenstva

- je možné určovať presným kvantitatívnym spôsobom (počet jedincov jednotlivých druhov; percentická pokryvnosť; odhad biomasy)
- prípadne podľa semikvantitatívnej stupnice (Braun-Blanquetová stupnica)
- alebo len kvalitatívnym spôsobom (prítomnosť či neprítomnosť)

Ordinačné metódy: typy dát

Vysvetľujúce premenné, prediktory



Môžu byť použité k predpovedaniu hodnôt vysvetľovaných premenných

Charakteristiky prostredia, príp. kovariáty

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

- faktoriálne (kategoriálne) – 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 nezaujíma, môžem 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

Typy štatistických modelov

Nasledujúca tabuľka zhŕňa najdôležitejšie štatistické metódy používané v rôznych situáciách:

Vysvetľovaná premenná ...	**Prediktor(y)**	
	nemáme	**máme**
... je jedna	◆ zhrnutie distribučných vlastností	◆ regresný model s.l.
... je ich viac	◆ nepriama gradientová analýza (indirect gradient analysis - PCA, DCA, NMDS) ◆ zhluková analýza	◆ priama gradientová analýza ◆ obmedzená zhluková analýza ◆ diskriminačná analýza (discriminant analysis - CVA)

Ordinačné metódy, gradientová analýza

- Výraz **gradientová analýza** je tu používaný v širšom slova zmysle pre akúkoľvek metódu, ktorá sa pokúša dať do vzťahu druhovú skladbu a gradienty prostredia (merené alebo hypotetické).
- **Cieľom** gradientovej analýzy je **nájsť smery najväčšej variability** v zložení spoločenstva a ich závislosť na určujúcich premenných prostredia.
- Zaoberá sa vzťahom zloženia spoločenstva k (známym alebo neznámym) gradientom prostredia.

Nepriama gradientová analýza

(*indirect gradient analysis*)

- Osi variability v druhovom zložení (môžu byť a mali by byť potom vztiahnuté k nameraným charakteristikám prostredia, keď sú tieto k dispozícii)

Priama gradientová analýza

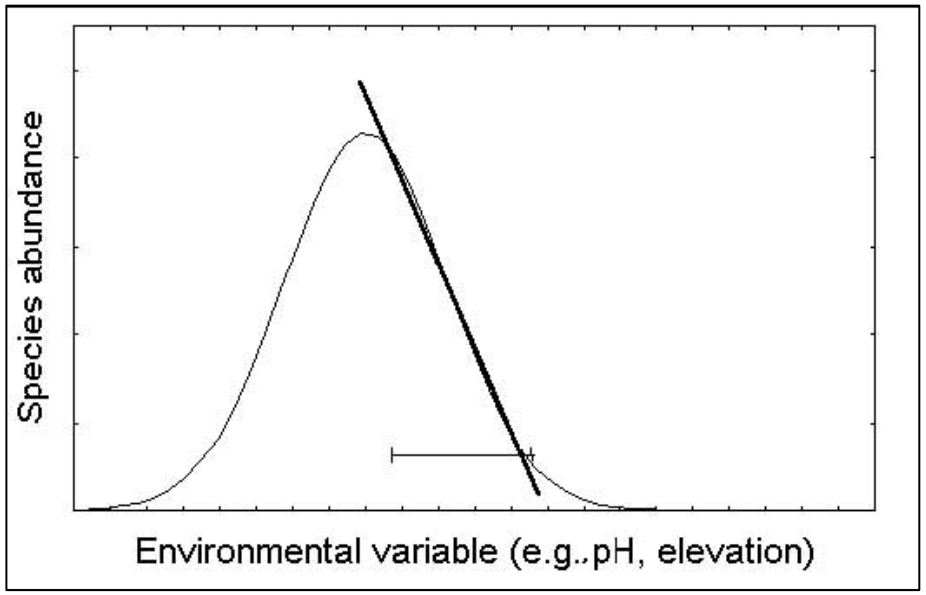
(*direct gradient analysis*)

- Variabilita v druhovom zložení vysvetlená charakteristikami prostredia.

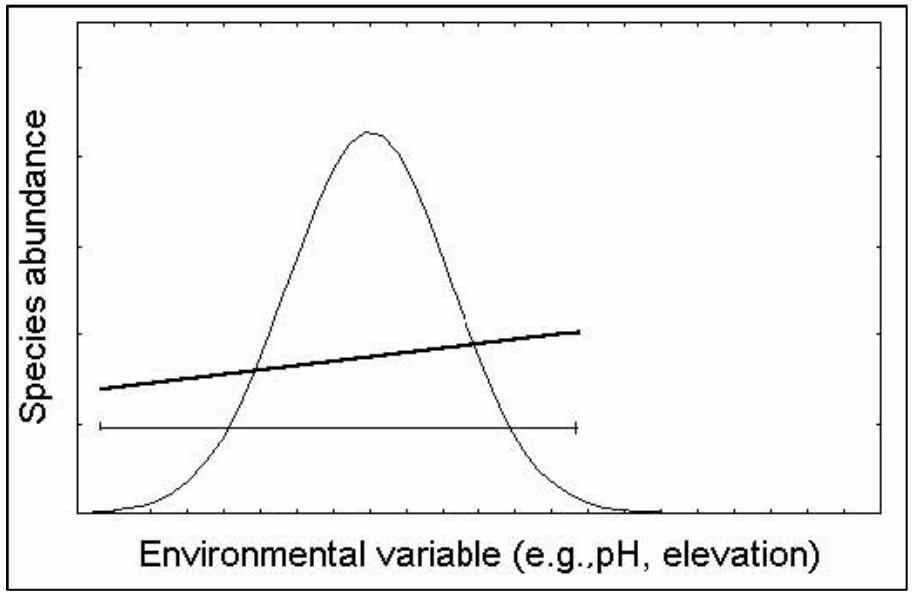
Modely odpovede druhov na gradienty prostredia

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)

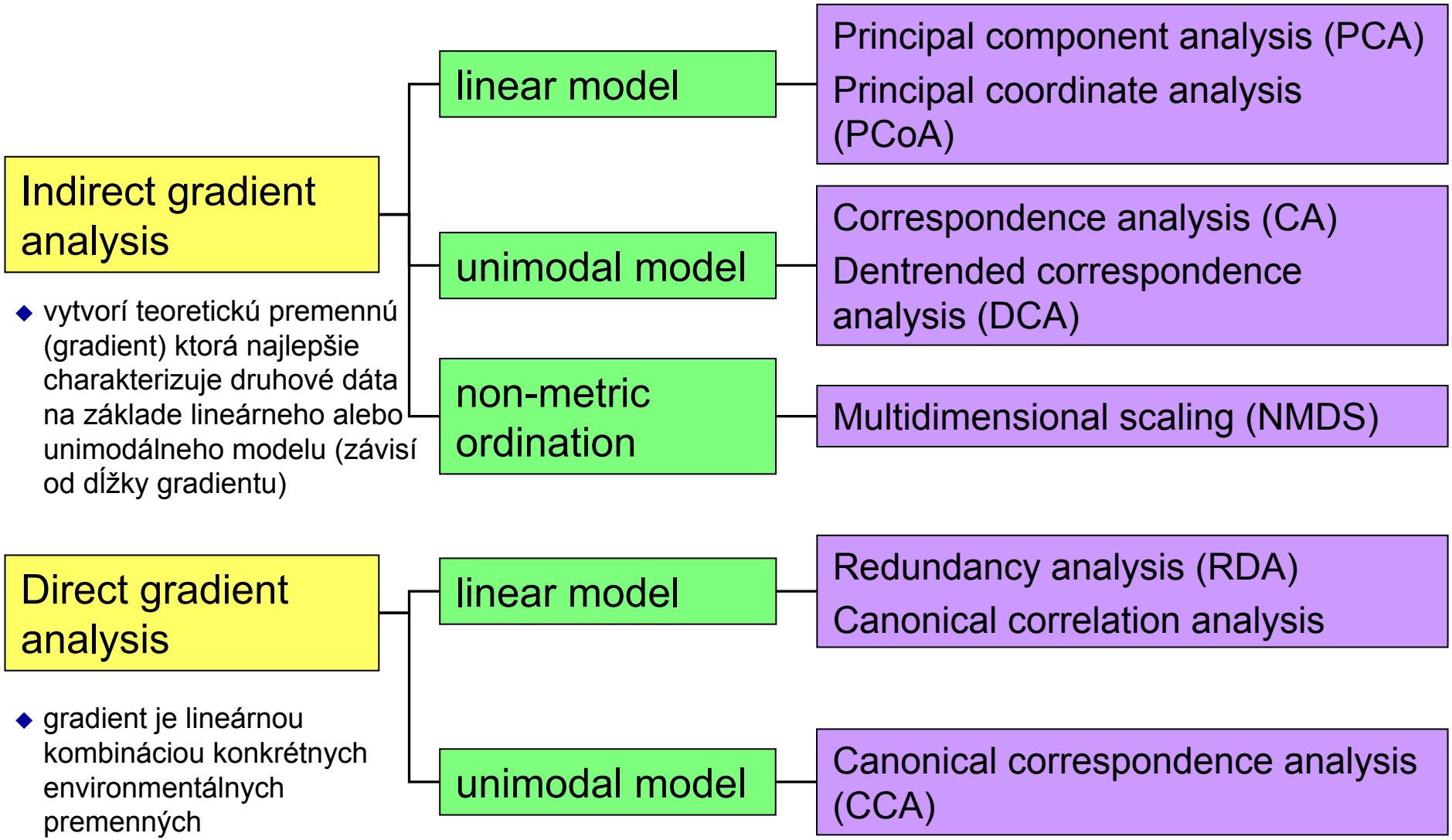


Lineárna aproximácia unimodálnej odpovede na krátkom výseku gradientu



Lineárna aproximácia unimodálnej odpovede na dlhej časti gradientu

Základné techniky ordinačných metód



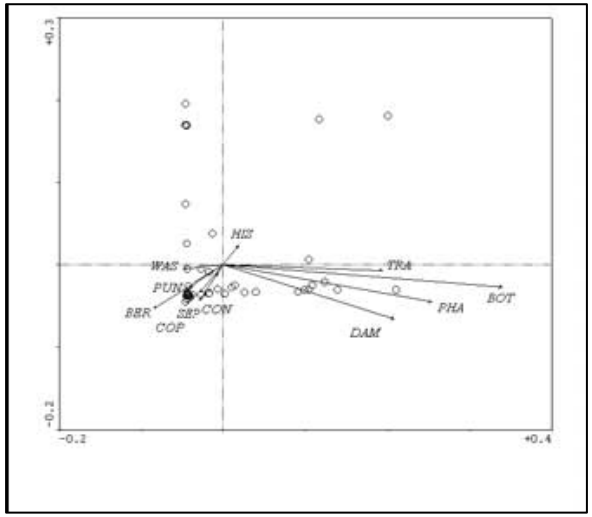
Ordinačné diagramy

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

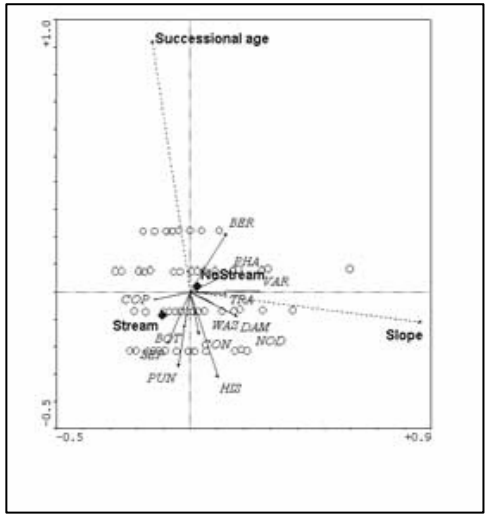
	Lineárny model	**Unimodálny model**
vzorky	◆ znázornené bodmi (symbolmi)	
druhy	◆ šípky v smere rastu abundancií	◆ body (symboly) označujúce optimum druhu
Charakteristiky prostredia kvantitatívne	◆ šípky v smere rastu hodnôt	
charakteristiky prostredia kvalitatívne	◆ centroidy pre jednotlivé kategórie	

Príklady ordinačných diagramov

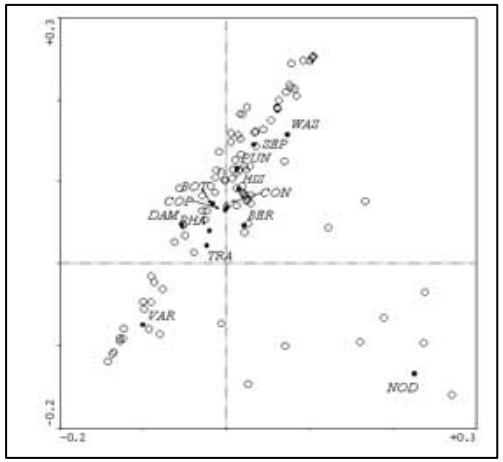
PCA



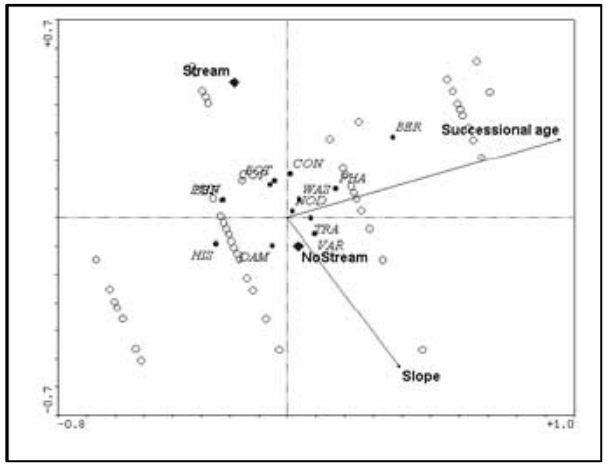
RDA



CA



CCA



Úprava dát do ordinačných metód

Danka Haruštiaková

Podzim 2009

Inštitút bioštatistiky a analýz, Masarykova univerzita

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 pokryvnosti:

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 **štandardizuje**, čiže majú nulový priemer a jednotkový rozptyl.

Nepriame ordinačné metódy

Danka Haruštiaková

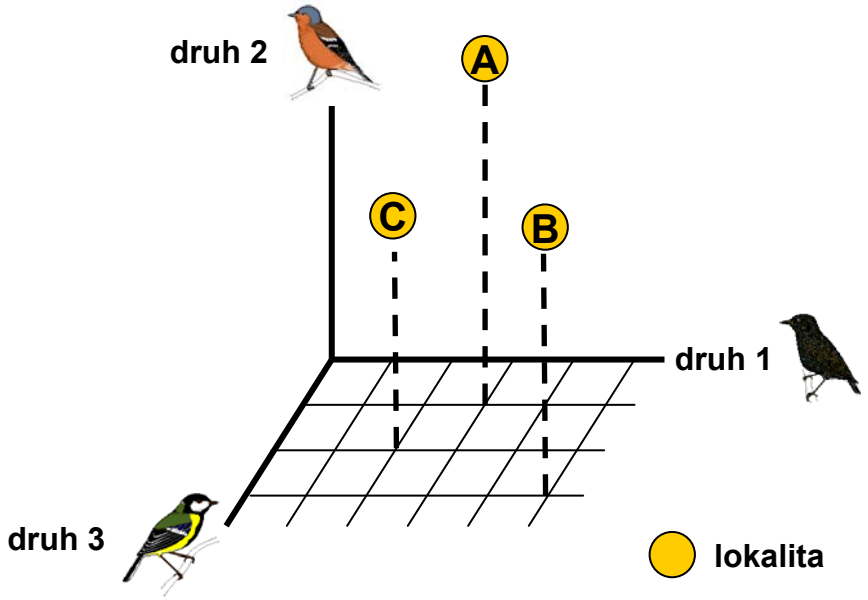
Podzim 2009

Inštitút bioštatistiky a analýz, Masarykova univerzita

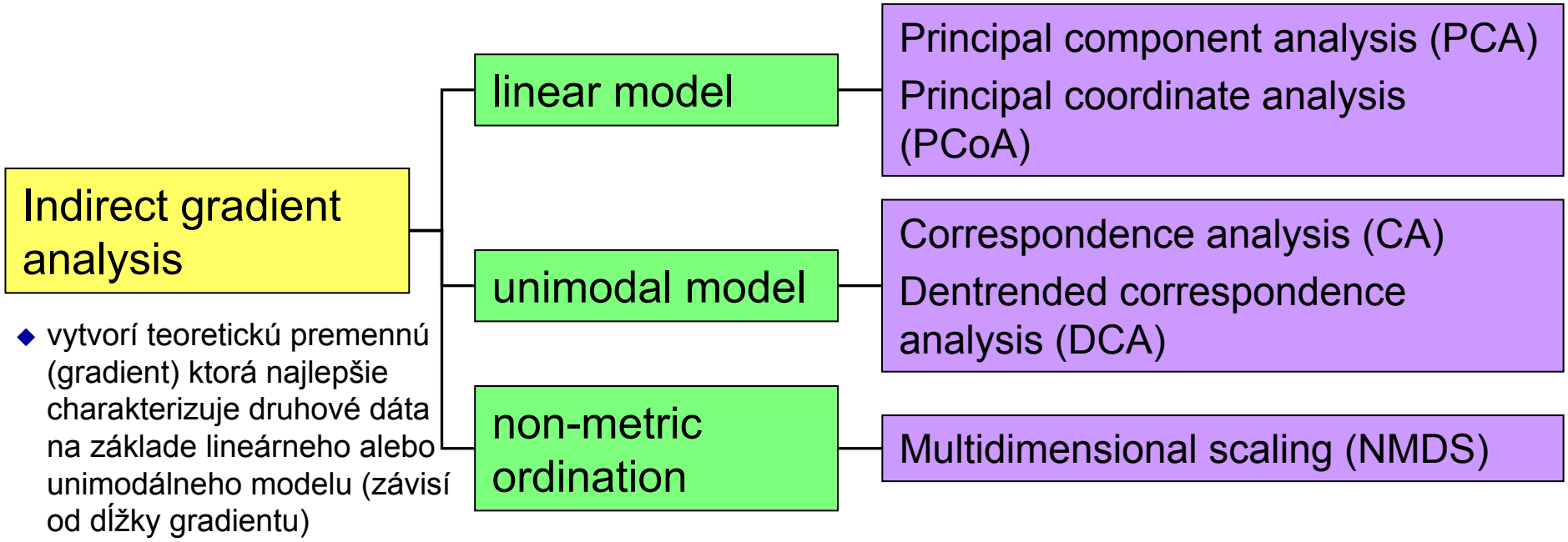
Nepriame ordinačné metódy

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“.



Základné techniky ordinačných metód



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

Voľba modelu: meranie dĺžky gradientu

Aby sme mohli zvoliť medzi lineárnym a unimodálnym modelom, musíme odmeriať 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 Haruštiaková

Podzim 2009

MU IBA Inštitút bioštatistiky a analýz, Masarykova univerzita

Analýza hlavných komponent (PCA)

Vstupné dáta

- Spojité alebo dummy premenné popisujúce jednotlivé objekty

Výstupy analýzy

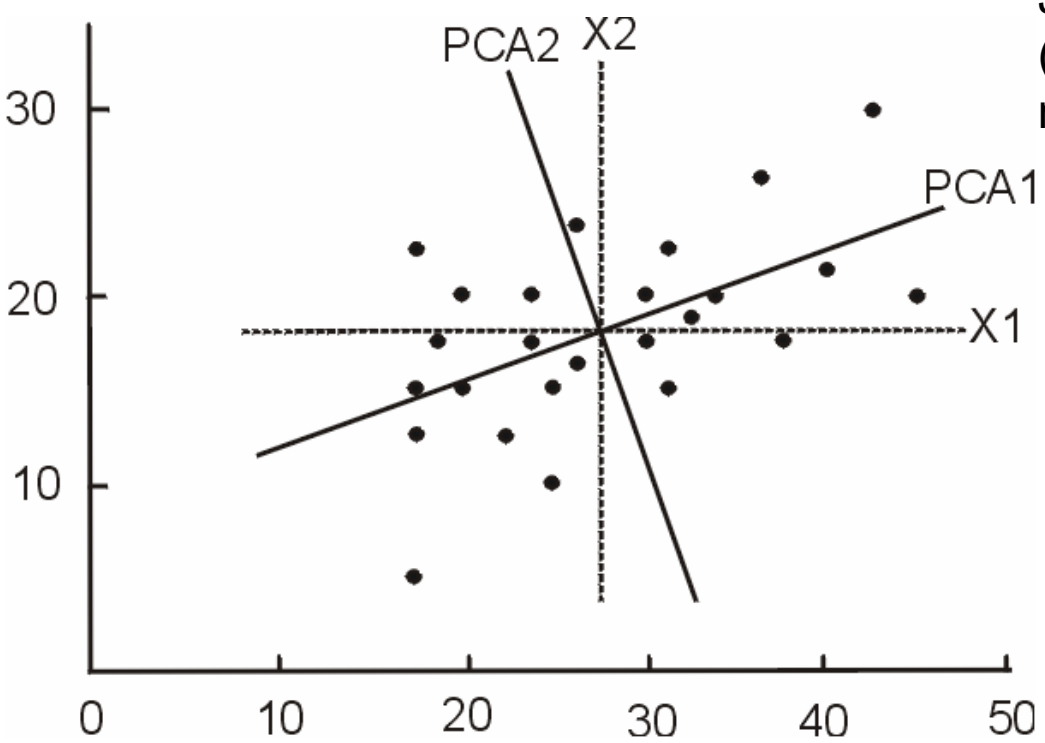
- Vzťahy všetkých pôvodných faktorov v jednoduchom xy grafe
- Pozícia objektov v priestore – jednoduchá identifikácia segmentov a vplyvov faktorov na rôzne skupiny

Kritické problémy analýzy

- Odľahlé hodnoty
- Úplne nezávislé premenné – nie je tu žiadna duplicitná informácia k vysvetleniu

Analýza hlavných komponent (PCA)

Nahrádza pôvodný súbor pozorovaných parametrov (druhová matica) 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.

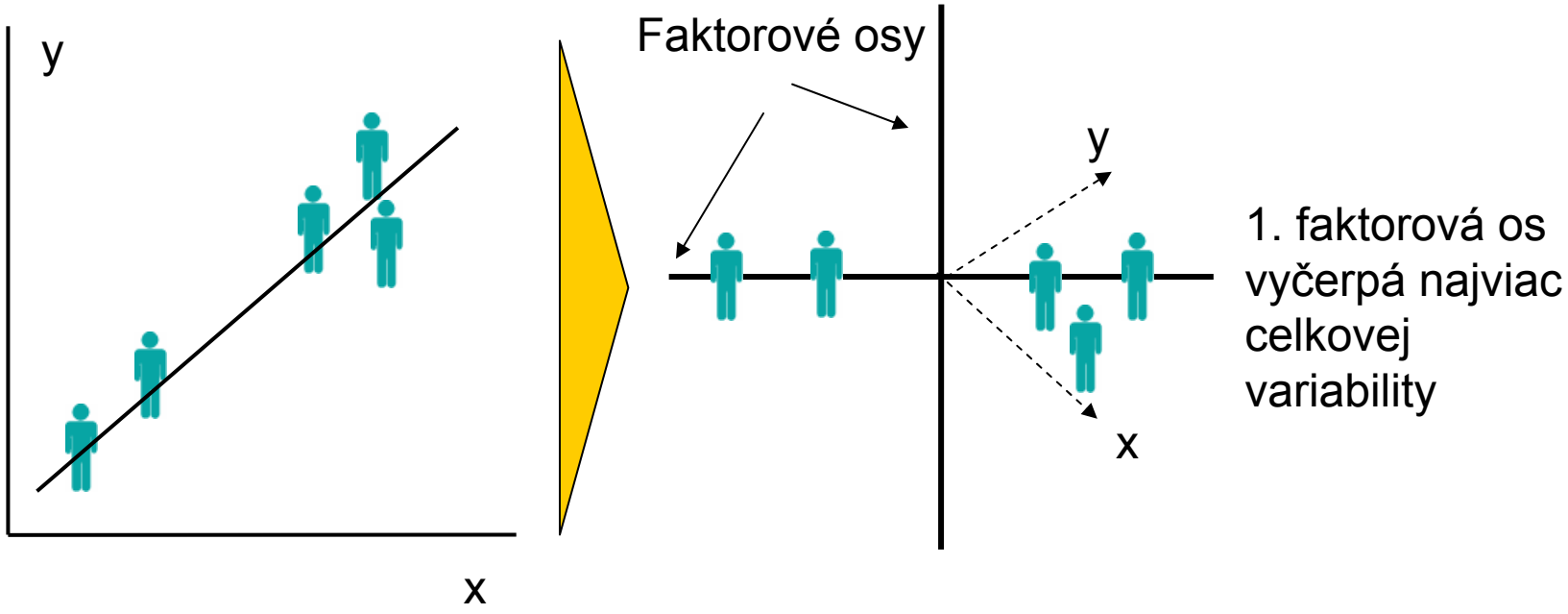


Je založená na **vlastnej analýze** (eigenanalysis) symetrických matíc (**korelačnej, kovariančnej**)

Analýza hlavných komponent (PCA)

Princíp

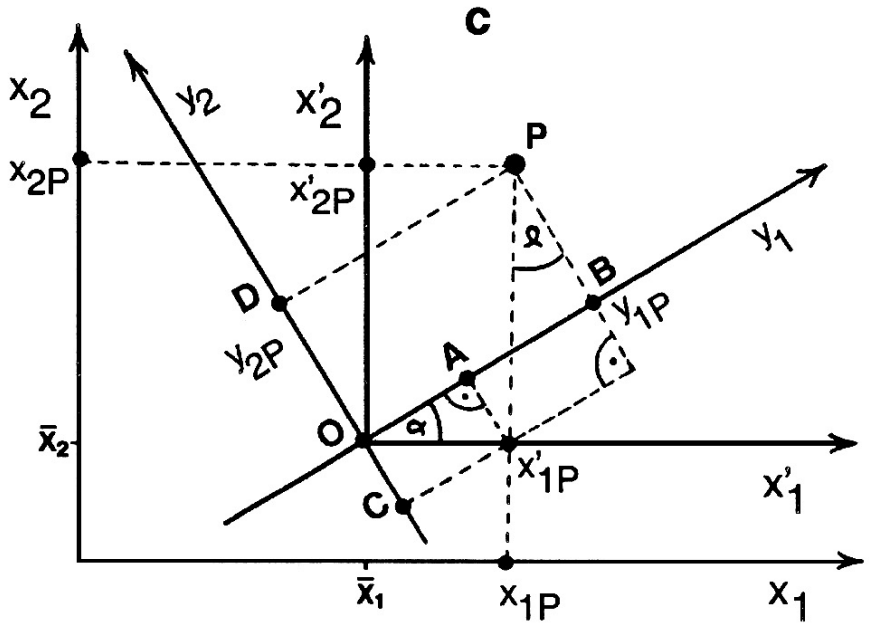
- Premenné sú navzájom korelované, teda časť informácie v súbore je duplicitná
- Analýza odstráni duplicitu z dát a zobrazí len unikátnu informáciu



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.

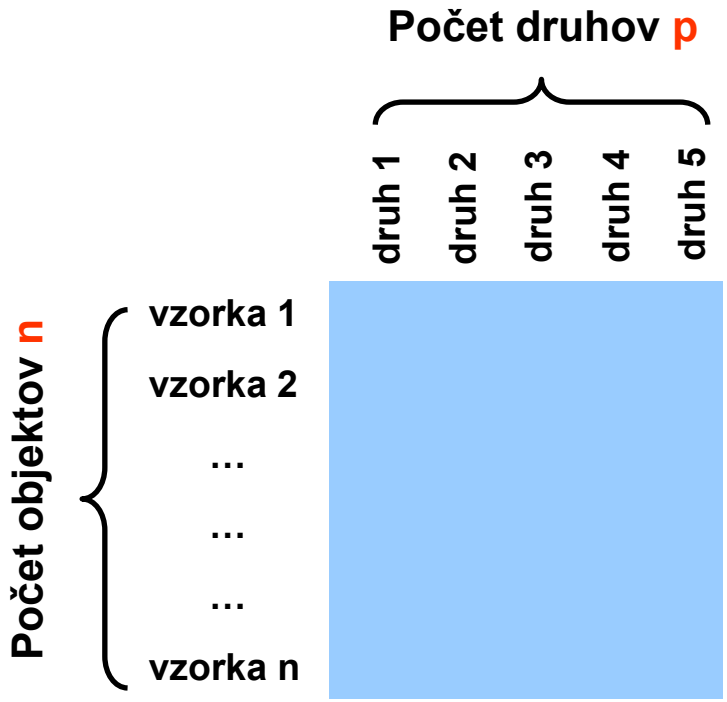


Pôvodne bola PCA navrhnutá pre kvantitatívne znaky, môže sa však použiť aj na znaky binárne a semikvantitatívne.

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

Analýza hlavných komponent (PCA)

- Počet objektov (vzoriek) pri PCA by mal byť aspoň o jeden väčší než je počet analyzovaných parametrov (druhov).
- Obvykle se však odporúča, aby sa počet objektov blížil druhej mocnine počtu parametrov (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)

Indirect gradient analysis

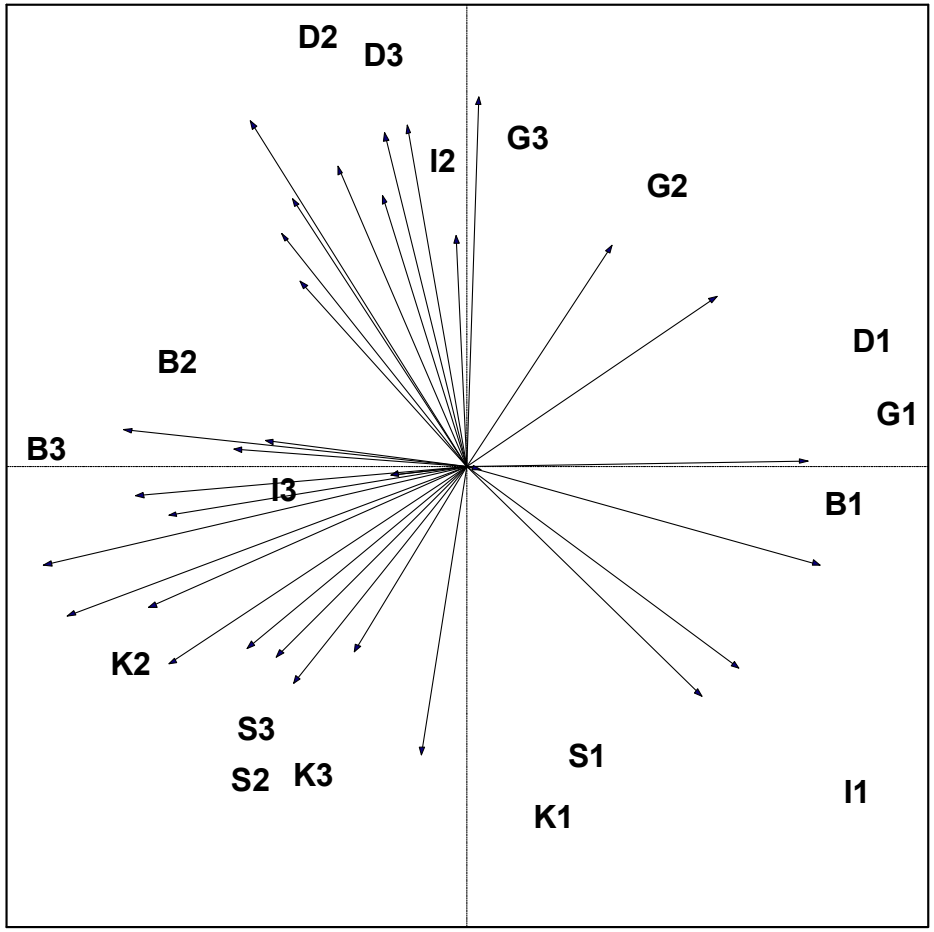
Principal component analysis

- PCA je postavená na lineárnom modely; 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

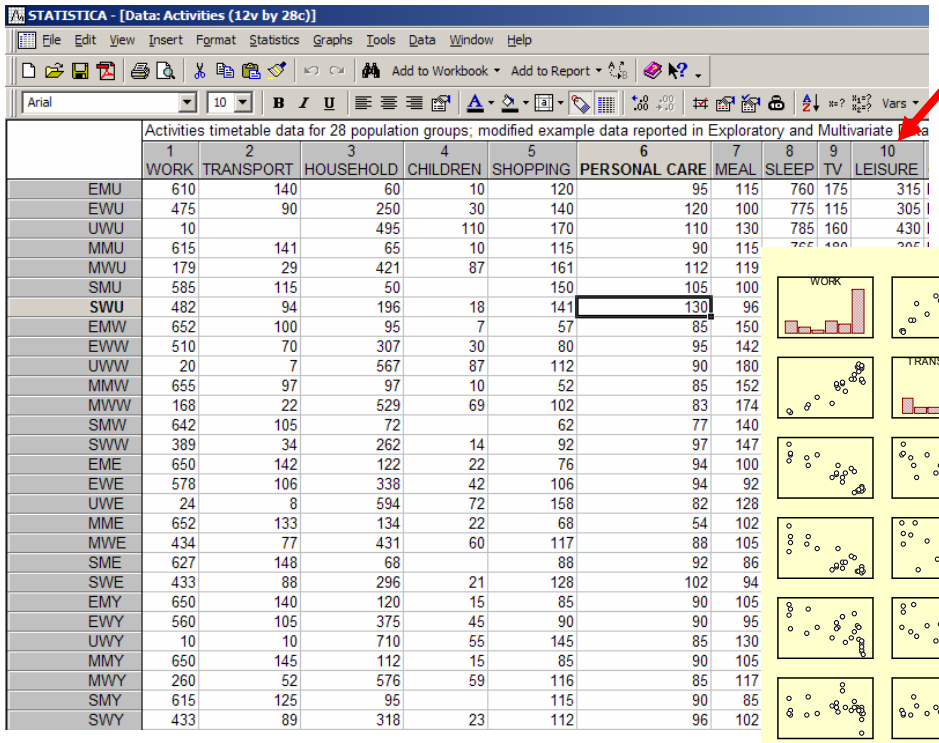
- 6 lokalít, každá lokalita sledovaná 3 obdobia
- dátová matica: 18 vzoriek x 63 plankt. druhov

  hodnoty = stupeň dominancie



PCA v Statistica

Vstupy výpočtu PCA

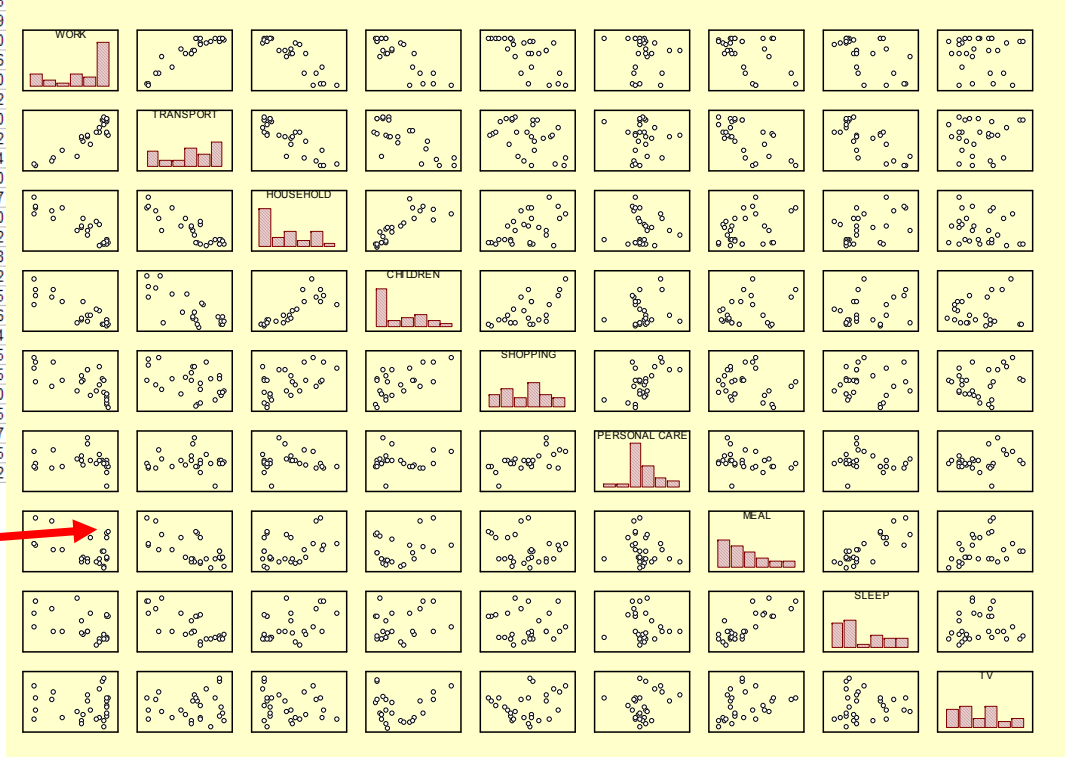


Activities timetable data for 28 population groups; modified example data reported in Exploratory and Multivariate [illegible]a

	1 WORK	2 TRANSPORT	3 HOUSEHOLD	4 CHILDREN	5 SHOPPING	6 PERSONAL CARE	7 MEAL	8 SLEEP	9 TV	10 LEISURE
EMU	610	140	60	10	120	95	115	760	175	315
EWU	475	90	250	30	140	120	100	775	115	305
UWU	10		495	110	170	110	130	785	160	430
MMU	615	141	65	10	115	90	115			
MWU	179	29	421	87	161	112	119			
SMU	585	115	50		150	105	100			
SWU	482	94	196	18	141	130	96			
EMW	652	100	95	7	57	85	150			
EWW	510	70	307	30	80	95	142			
UWW	20	7	567	87	112	90	180			
MMW	655	97	97	10	52	85	152			
MWW	168	22	529	69	102	83	174			
SMW	642	105	72		62	77	140			
SWW	389	34	262	14	92	97	147			
EME	650	142	122	22	76	94	100			
EWE	578	106	338	42	106	94	92			
UWE	24	8	594	72	158	82	128			
MME	652	133	134	22	68	54	102			
MWE	434	77	431	60	117	88	105			
SME	627	148	68		88	92	86			
SWE	433	88	296	21	128	102	94			
EMY	650	140	120	15	85	90	105			
EWY	560	105	375	45	90	90	95			
UWY	10	10	710	55	145	85	130			
MMY	650	145	112	15	85	90	105			
MWY	260	52	576	59	116	85	117			
SMY	615	125	95		115	90	85			
SWY	433	89	318	23	112	96	102			

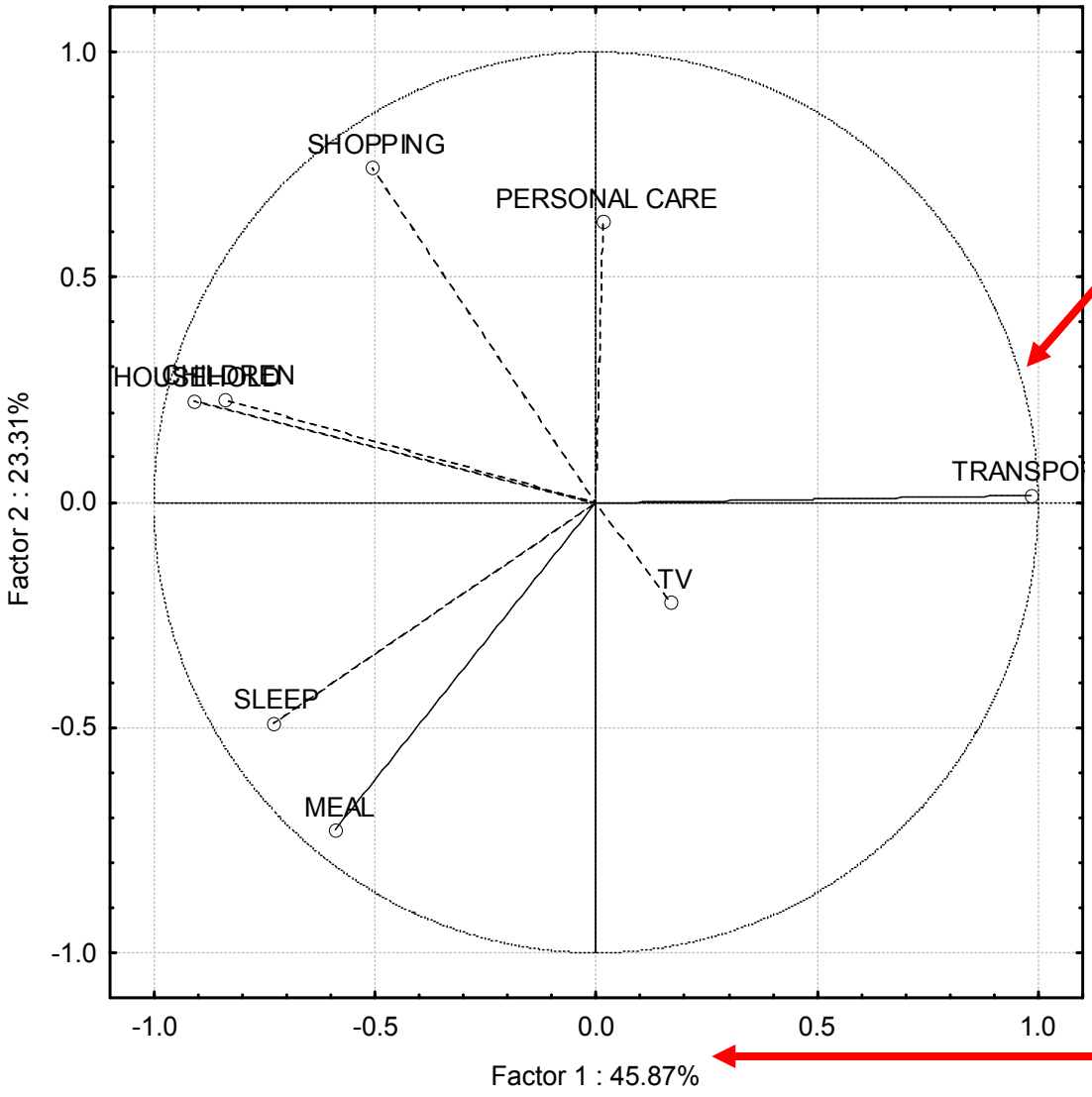
Vstupná tabuľka spojitých dát

Nutná analýza vzťahu premenných – analýza predpokladov



PCA v Statistica

Výstupy analýzy hlavných komponent



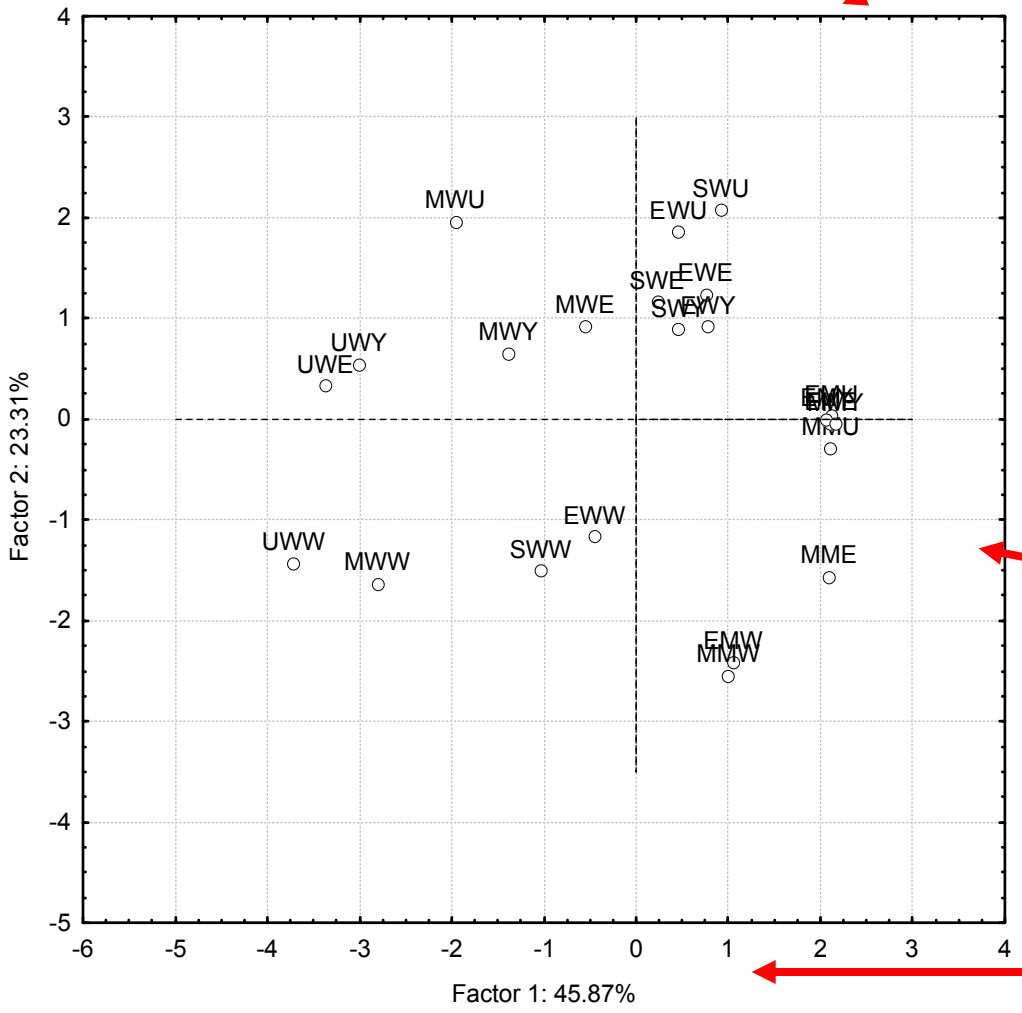
Pozícia faktora = miera väzby parametra s danou osou (-1,+1)
Dôležitá pre interpretáciu.

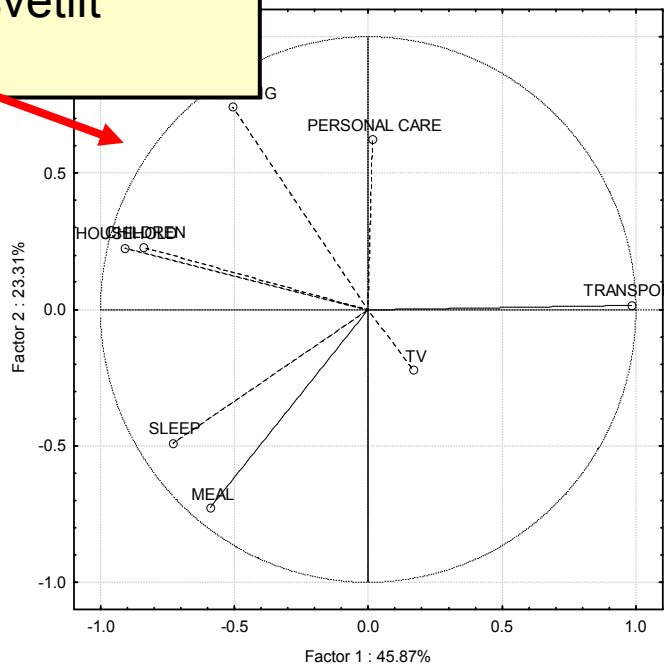
Množstvo vyčerpanej variability (informačná hodnota osi)

PCA v Statistica

Výstupy PCA

Pozíciu objektu možno vysvetliť pomocou grafu faktorov.



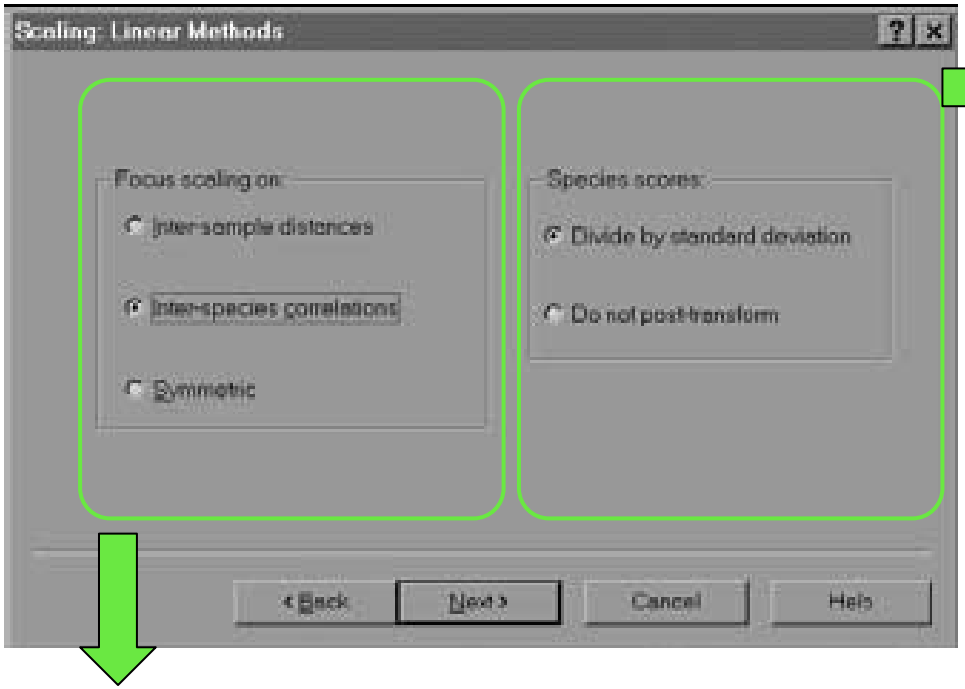


Pozícia objektu vo faktorovom priestore

Množstvo vyčerpanej variability (informačná hodnota osi)

PCA v Canoco

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**.

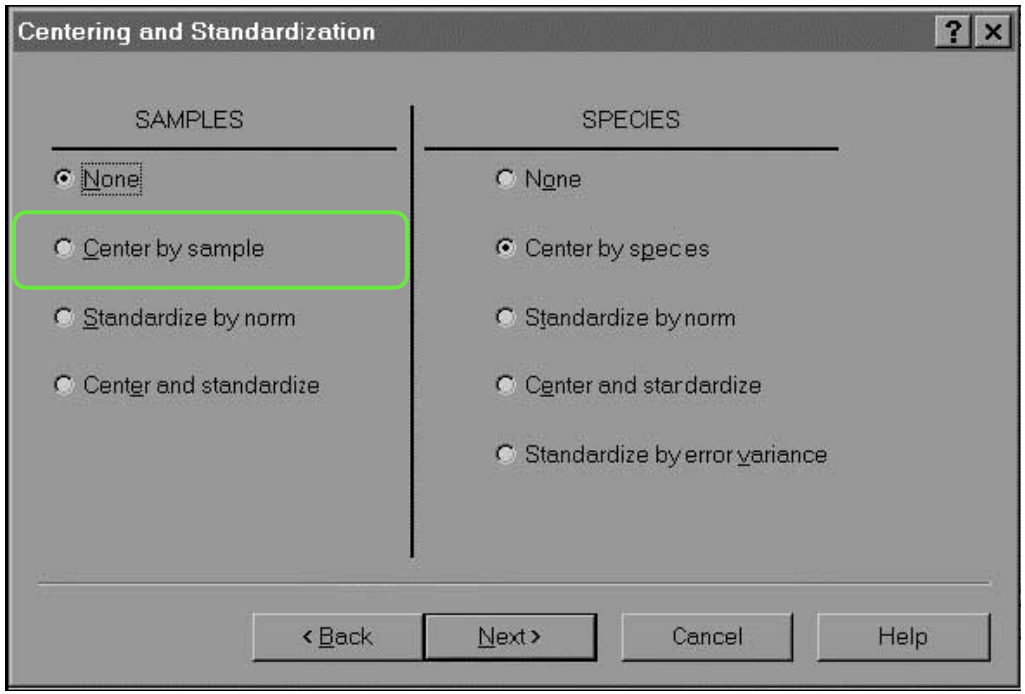
Ak máme charakteristiky prostredia, prípadne kovariáty, *species scaling* umožňuje charakterizovať korelácie medzi charakteristikami prostredia.

PCA v Canoco

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

Centrovanie

Priemer každého riadku bude rovný nule.

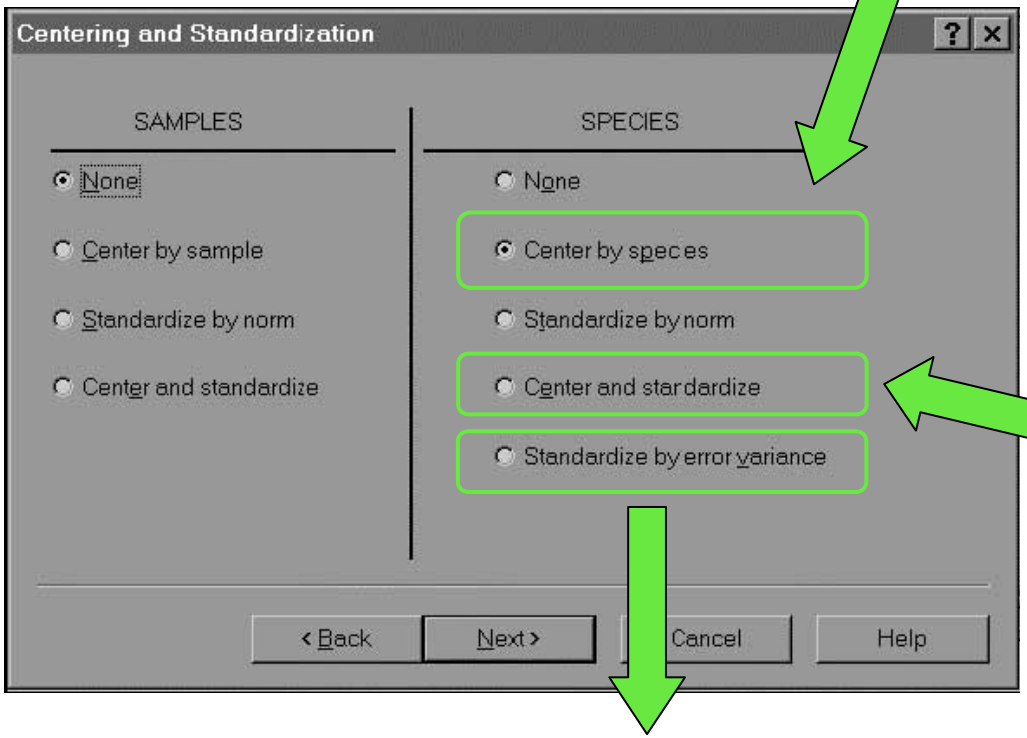


Centrovanie druhov nutné pre lineárne metódy s obmedzením (**RDA**) alebo pre parciálnu lineárnu ordináciu (tj. pri použití **kovariát**)

PCA v Canoco

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.

Faktorová analýza (FA)

Danka Haruštiaková

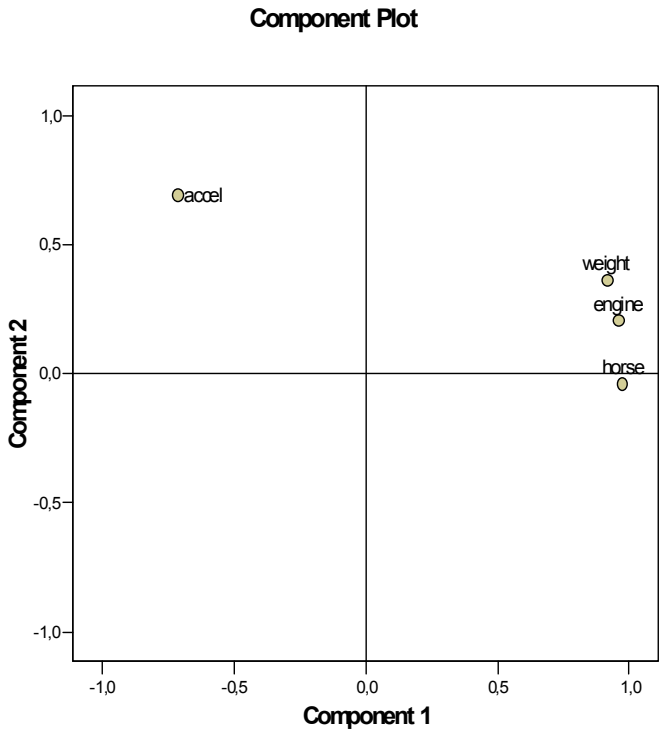
Podzim 2009

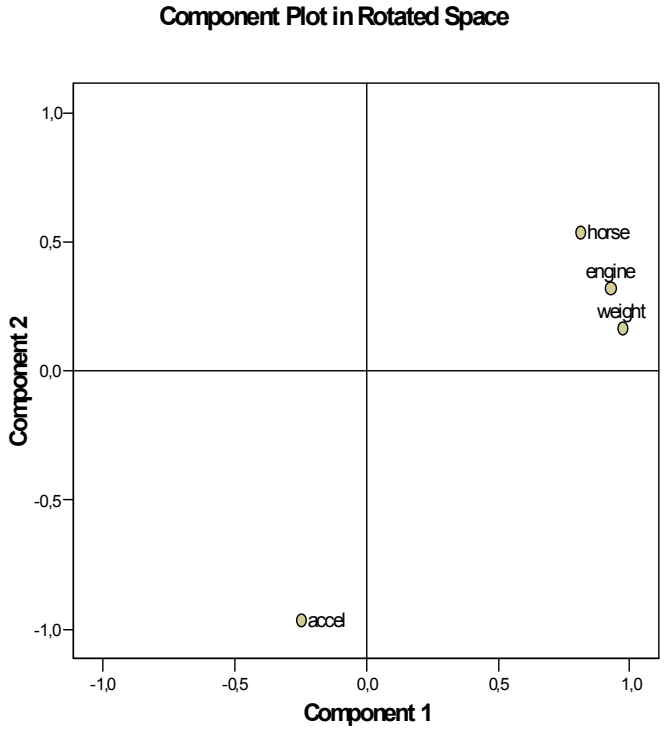
Inštitút bioštatistiky a analýz, Masarykova univerzita

Faktorová analýza (FA)

Čím sa líši od analýzy hlavných komponent?

- Jediným rozdielem je rotácia premenných tak aby sa vytvorené faktorové osi dali dobre interpretovať
- Výhodou je lepšia interpretácia vzťahu pôvodných premenných
- Nevýhodou je priestor pre subjektívny názor analytika





Analýza hlavných koordinát (PCoA)

Danka Haruštiaková

Podzim 2009

Inštitút bioštatistiky a analýz, Masarykova univerzita

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 Haruštiaková

Podzim 2009

Inštitút bioštatistiky a analýz, Masarykova univerzita

Korešpondenčná analýza

Vstupné dáta

- Tabuľka obsahujúca súhrny premenných (počty, priemery) za skupiny objektov

Výstupy analýzy

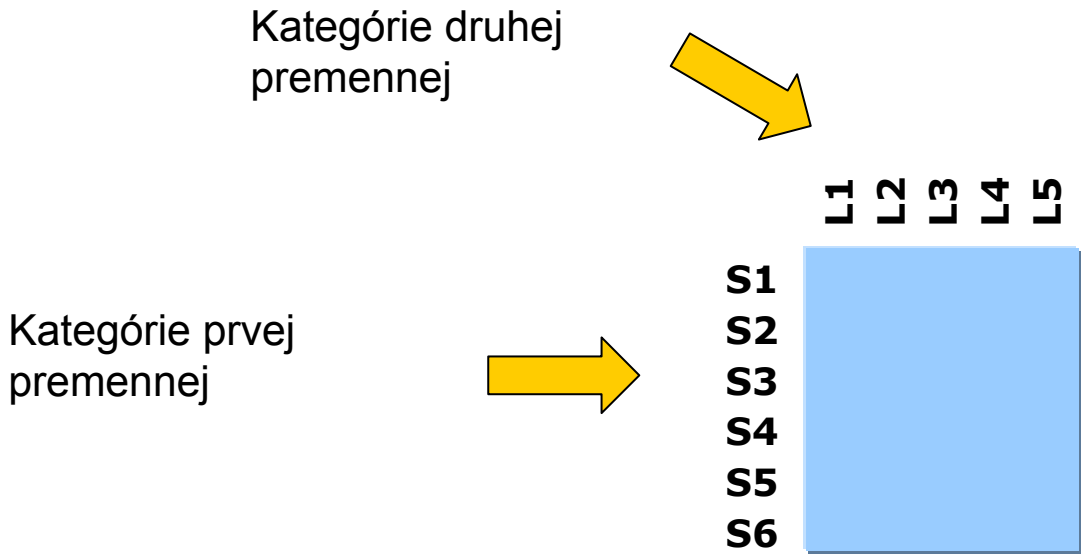
- Vzťahy všetkých pôvodných faktorov a/alebo skupín objektov v jednoduchom xy grafe

Kritické problémy analýzy

- Skupiny s malým počtom hodnôt môžu byť zaťažené značným šumom a náhodnou chybou
- Obtiažna interpretácia veľkého množstva malých skupín objektov

Korešpondenčná analýza

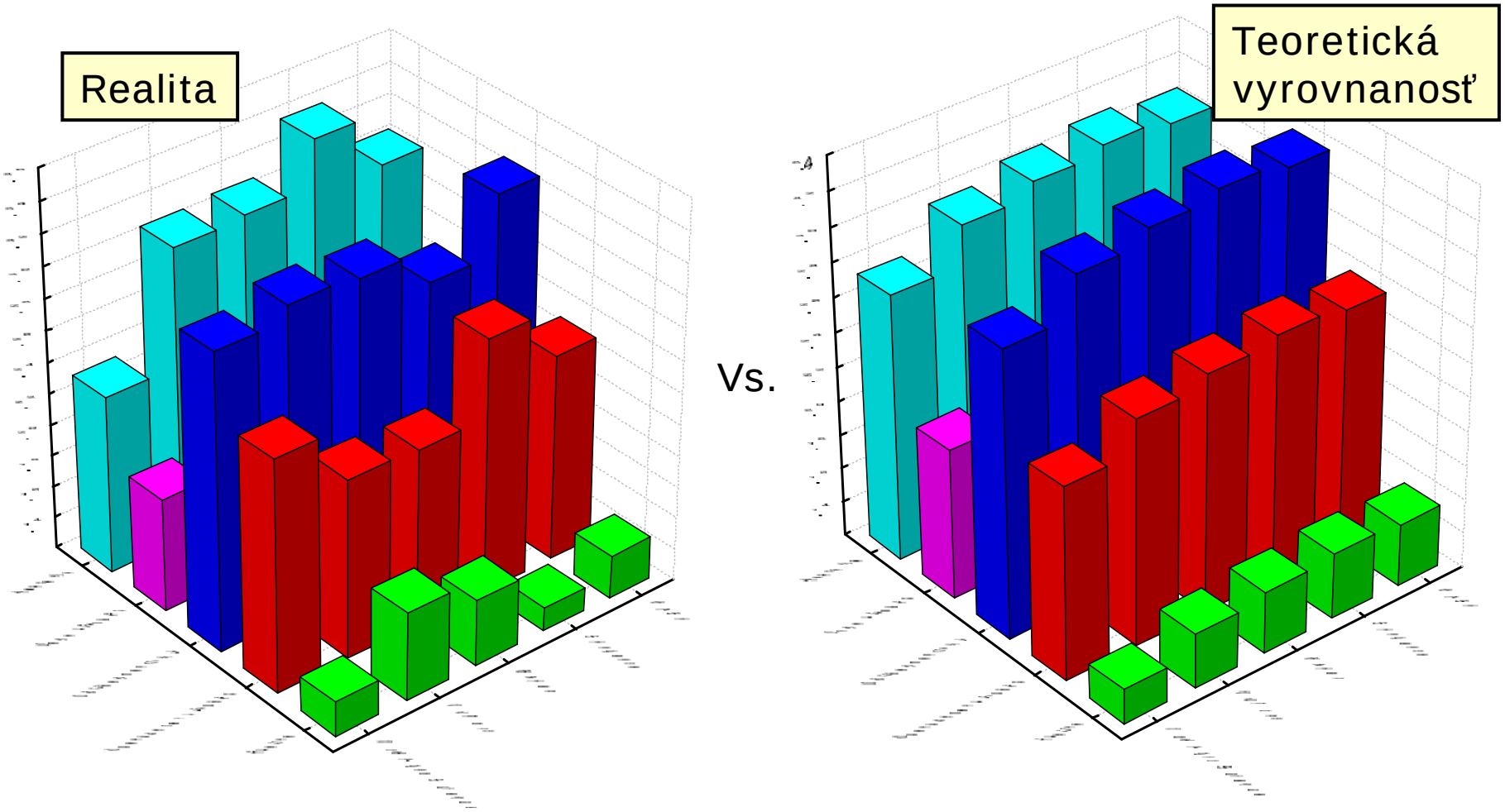
- Korešpondenčná analýza - nástroj pre analýzu vzťahov medzi riadkami a stĺpcami kontingenčných tabuliek
- Skúmanie vzťahov medzi dvoma premennými
- Kontingenčná tabuľka: frekvenčná tabuľka (dvojvstupná), ktorá zaznamenáva kumulatívné početnosti dvoch nominálnych (kategoriálnych) premenných. Každý stĺpec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



Korešpondenčná analýza

Princíp

Korešpondenčná analýza hľadá, které kombinácie riadkov a stĺpcov hodnotenej tabuľky najviac prispievajú k jej variabilite.



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.
- Korešpondenčný analýza se 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:

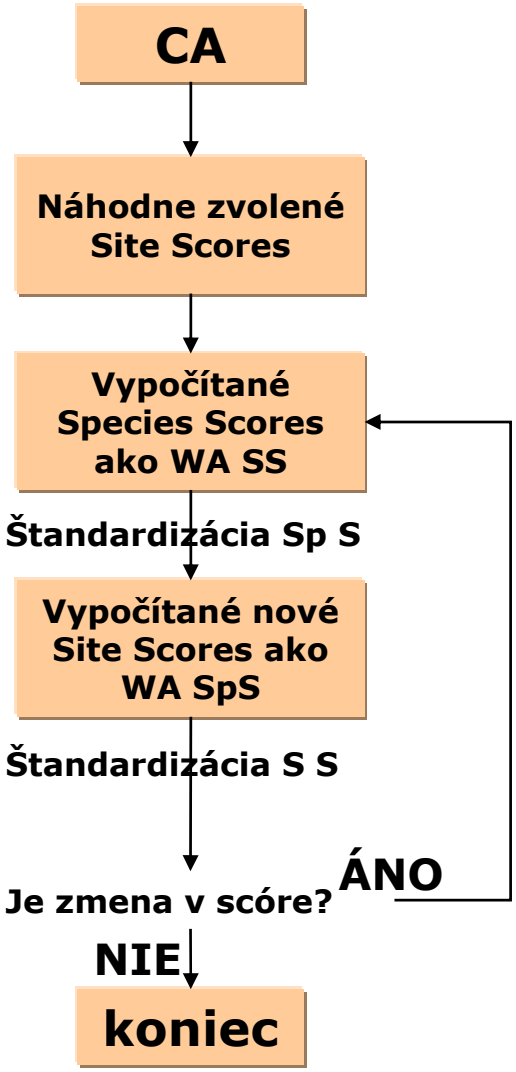
- Najjednoduchšou cestou ako odhadnúť optimum druhu pre unimodálny model je spočítať vážený priemer tých hodnôt charakteristík prostredia, pri ktorých sa druh vyskytuje.
- Ako váha sa pri výpočte používa početnosť ci iná dôležitostná hodnota druhu.
- Pri váženom priemerovaní je implicitne zahrnutá štandardizácia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

Korešpondenčná analýza v ekológii spoločenstiev

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
initial value	*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				



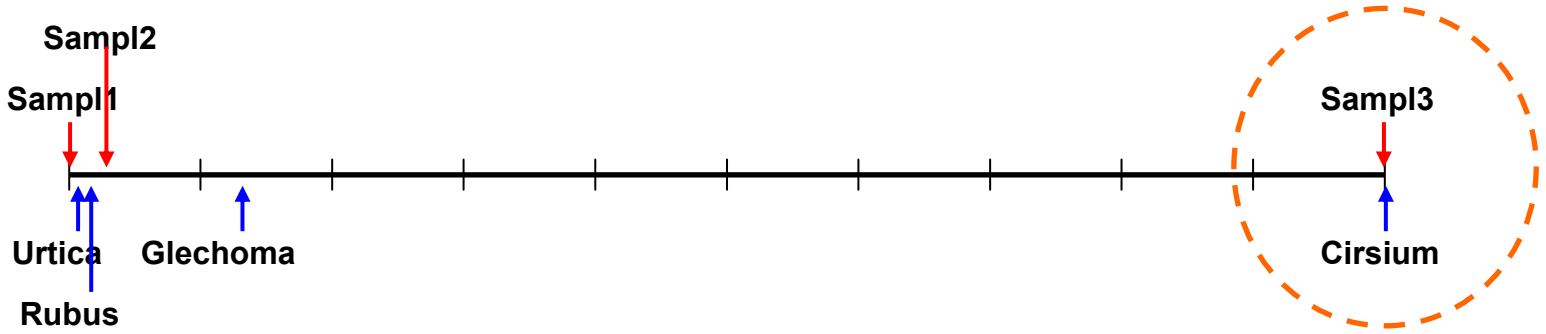
Korešpondenčná analýza v ekológii spoločenstiev

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)



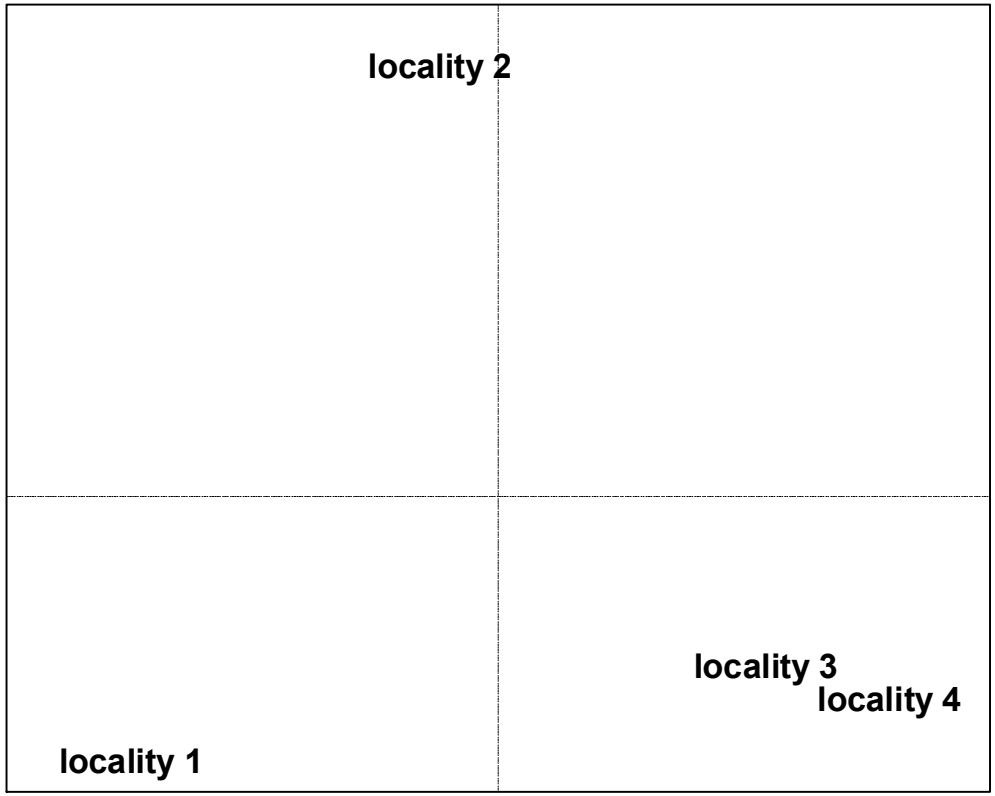
Korešpondenčná analýza: výsledky

- 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 v Statistica

Výstupy korešpondenčnej analýzy

STATISTICA - [Data: mark_pruzkum* (5v by 5c)]

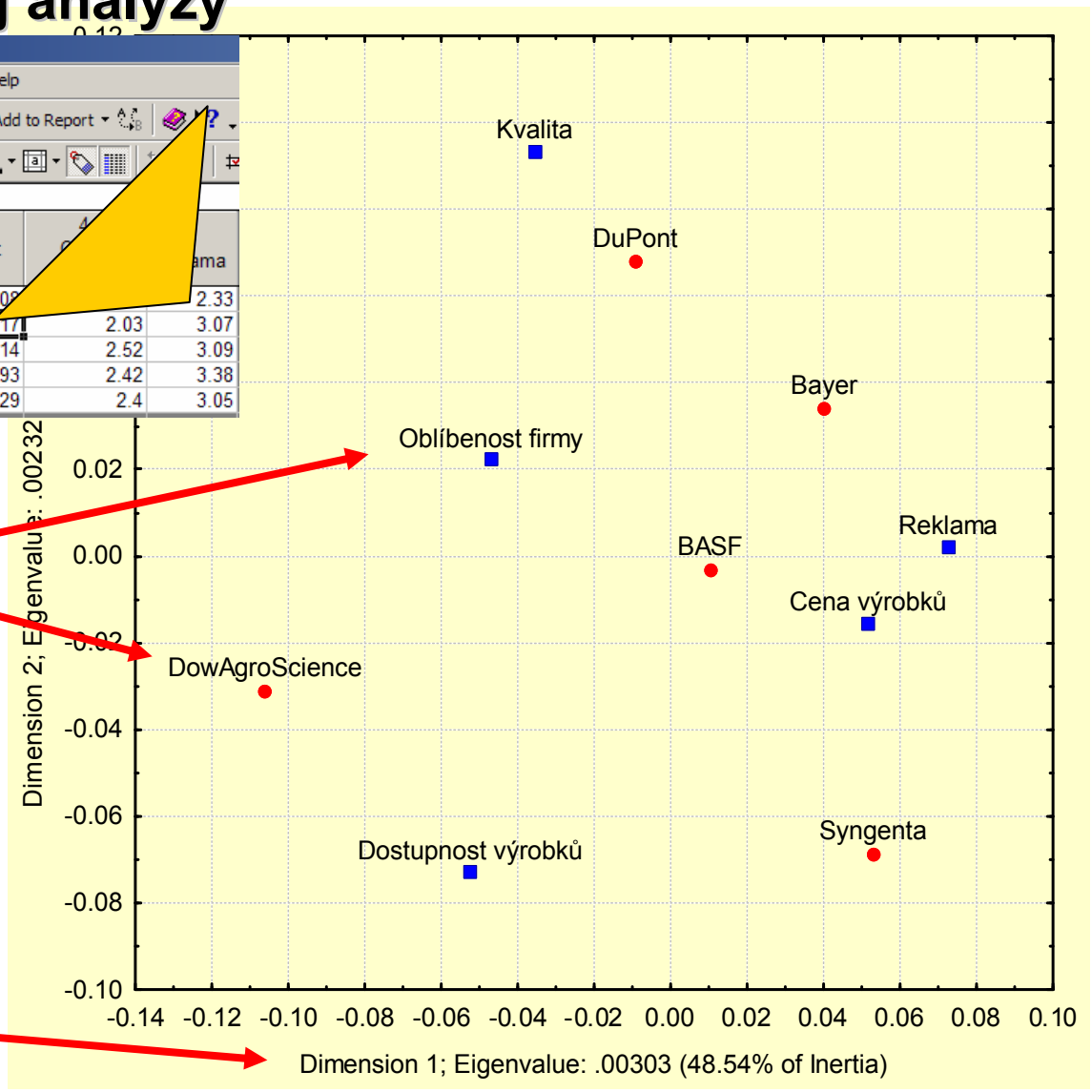
File Edit View Insert Format Statistics Graphs Tools Data Window Help

Add to Workbook Add to Report

Arial 10 B I U

	1 Kvalita	2 Dostupnost výrobků	3 Oblíbenost firmy	4	ama
DowAgro Science	1.42	2.67	3.08		2.33
DuPont	1.76	2.34	3.17	2.03	3.07
Bayer	1.62	2.32	3.14	2.52	3.09
Syngenta	1.35	2.81	2.93	2.42	3.38
BASF	1.47	2.51	3.29	2.4	3.05

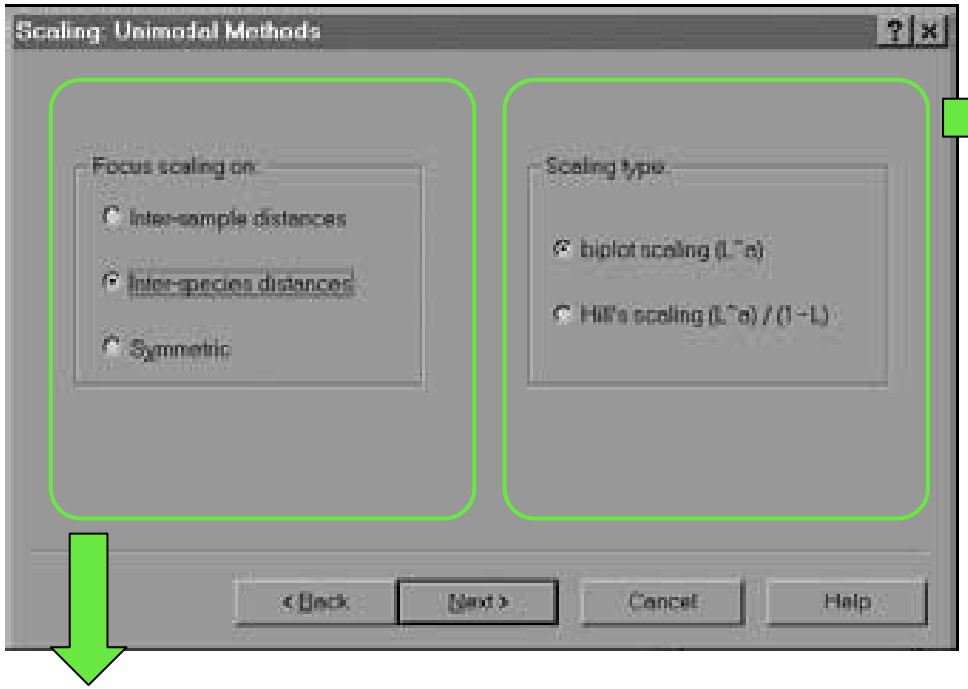
Vzájomná pozícia faktorov a skupín respondentov: vzájemnú pozíciu možno interpretovať



Variabilita vyčerpaná danou faktorovou osou

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

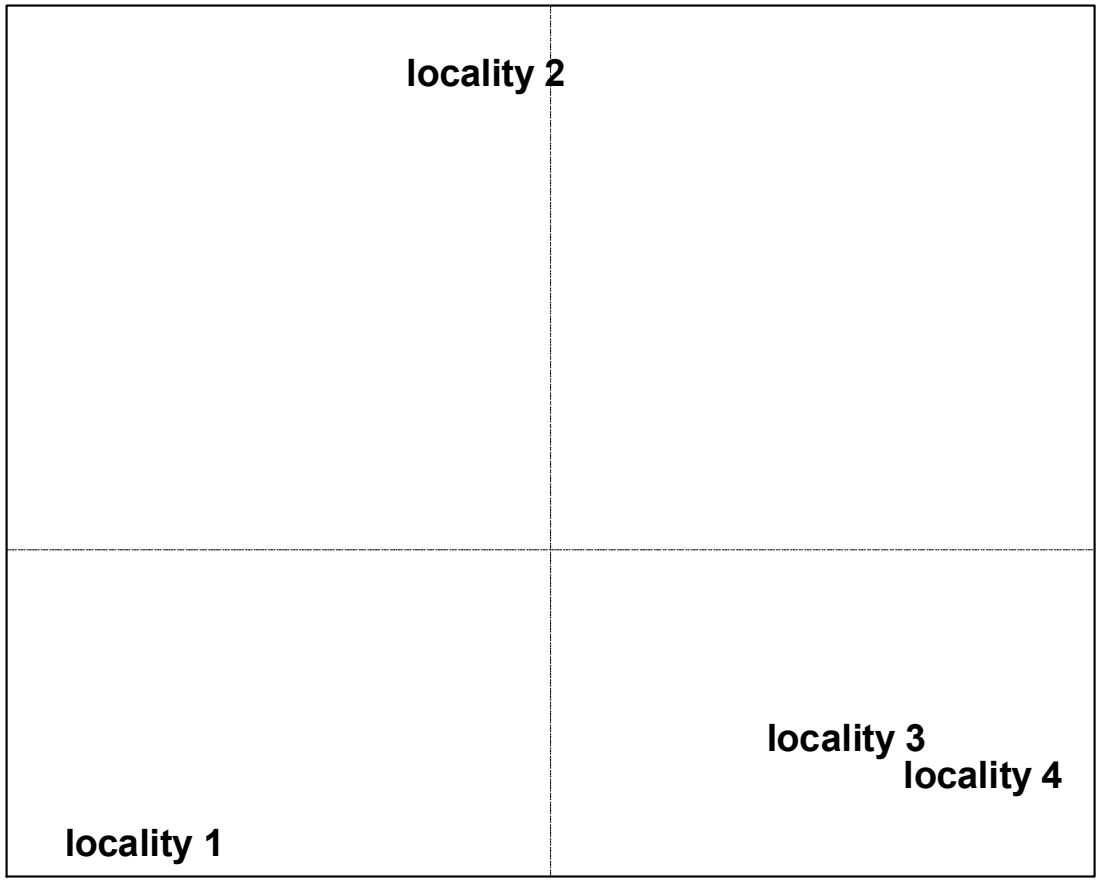
Correspondence analysis

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

REÁLNE DÁTA

- vtáčie druhy na 4 lokalitách
- dátová matica: 4 lokality x 38 dr. vtákov

hodnoty = priemerná abundancia



Korešpondenčná analýza: „arch effect“

Indirect gradient analysis

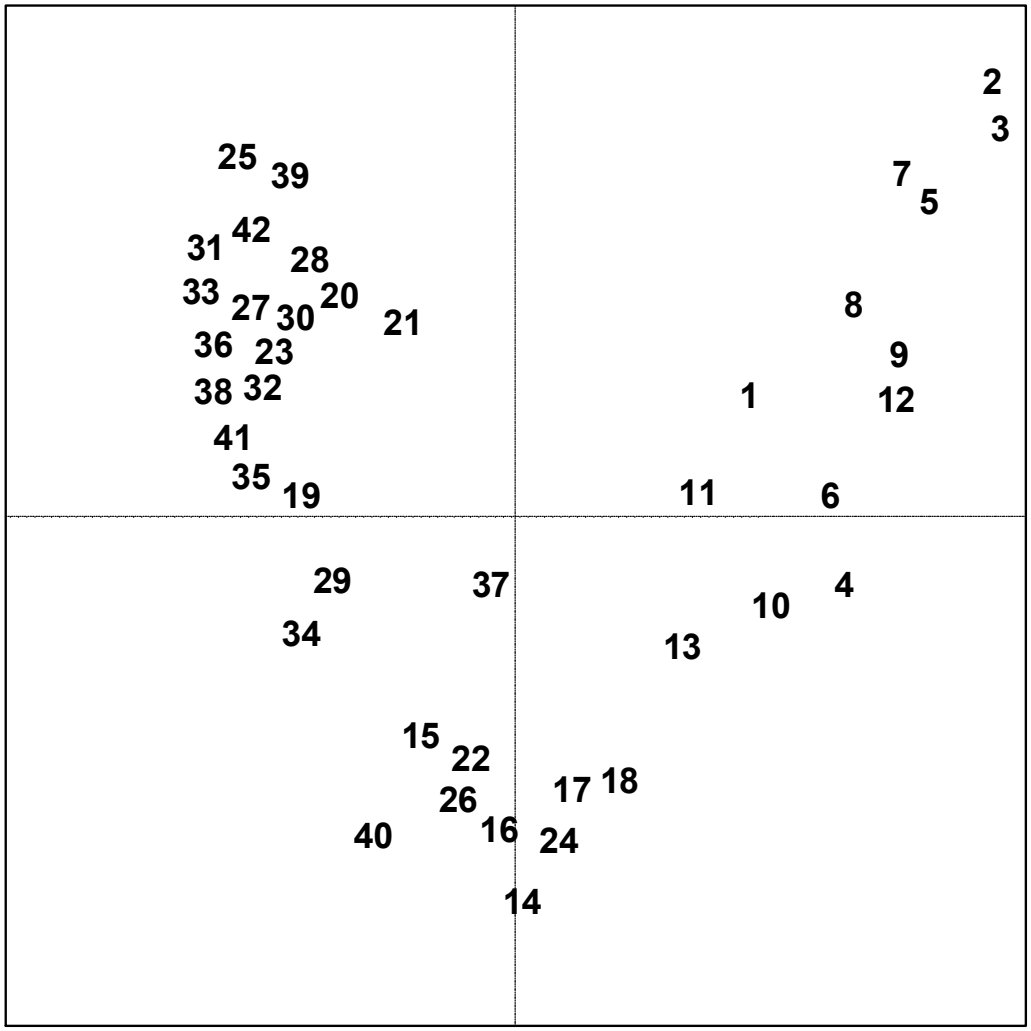
Correspondence analysis

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

REÁLNE DÁTA

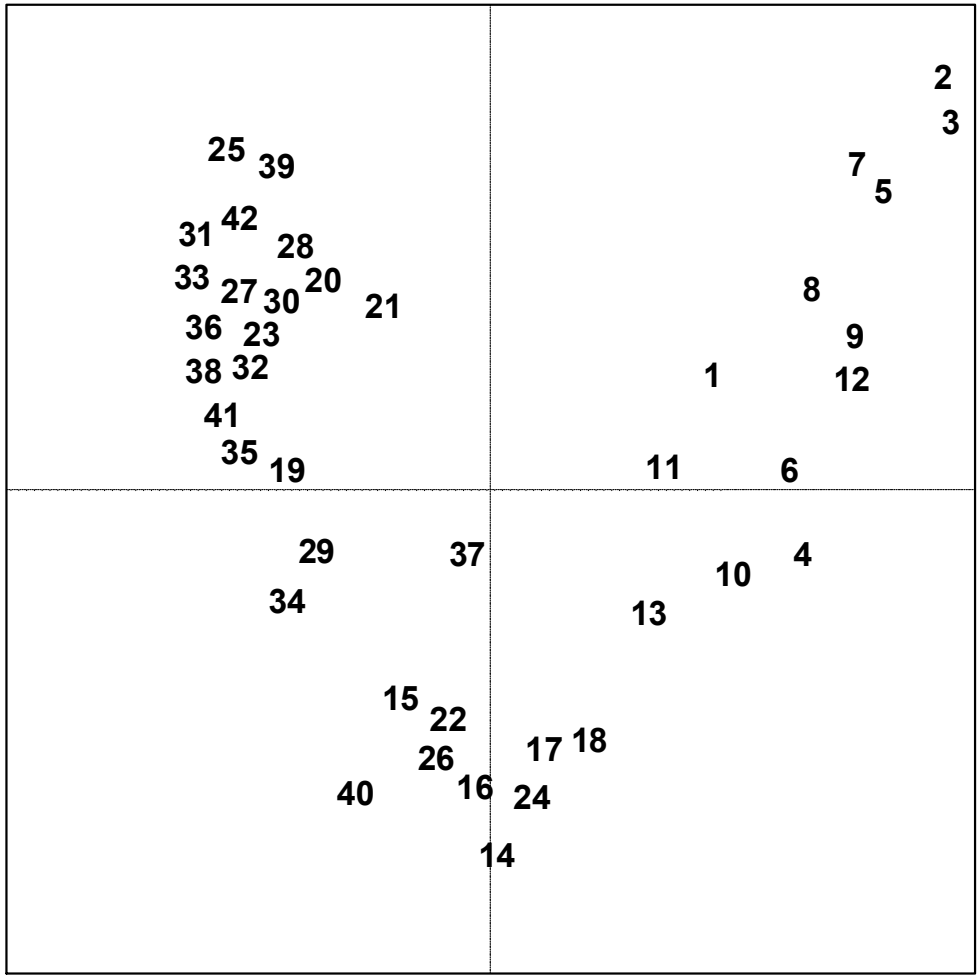
- suchozemské slimáky
- dátová matica: 42 lokalít x 33 dr. slimákov

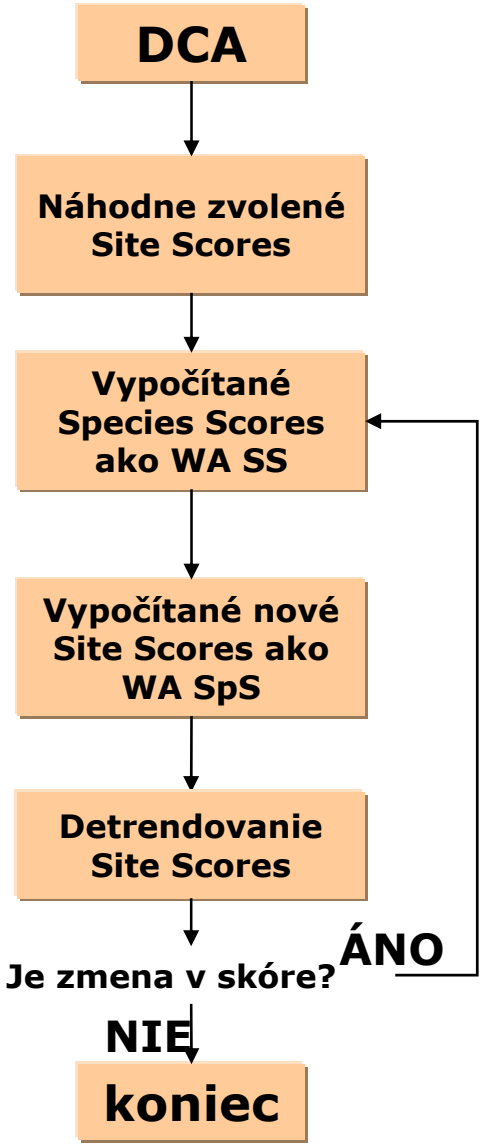
hodnoty = stupeň dominancie



Korešpondenčná analýza: „arch effect“

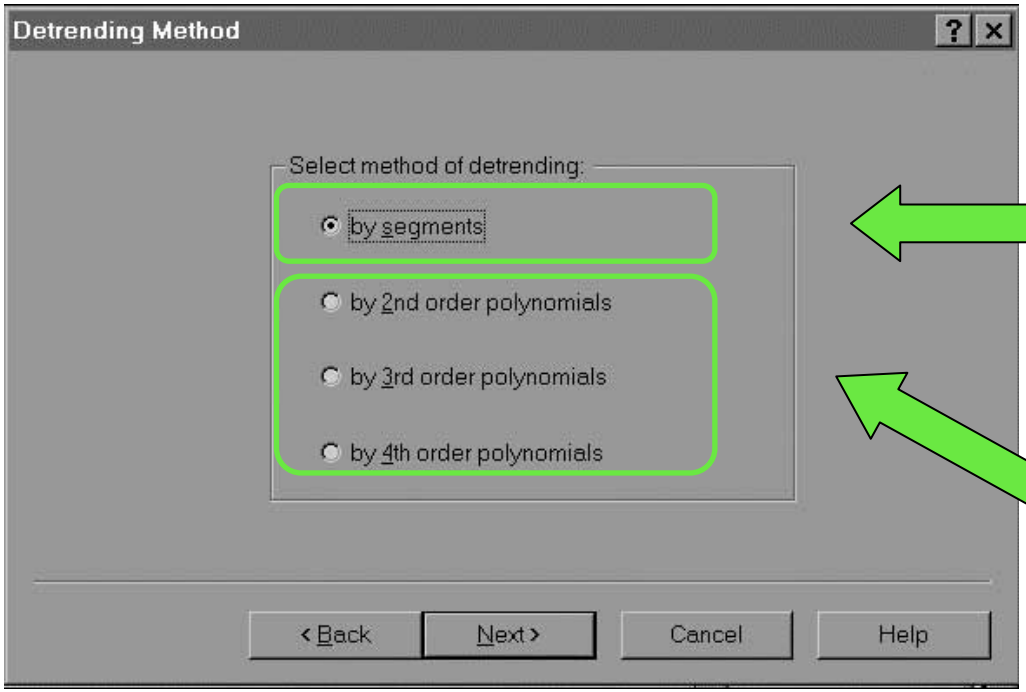
„arch effect“, „horse shoe effect“





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 súboru zvolených charakteristík prostredia.
- Doporučuje sa vylúčiť silne korelované premenné. Výber charakteristík prostredia, které sú mezi sebou korelované len minimálne, sa dá previesť postupnou selekciou charakteristík prostredia (*forward selection of environmental variables*).

Detrendovaná korešpondenčná analýza (DCA)

Indirect gradient analysis

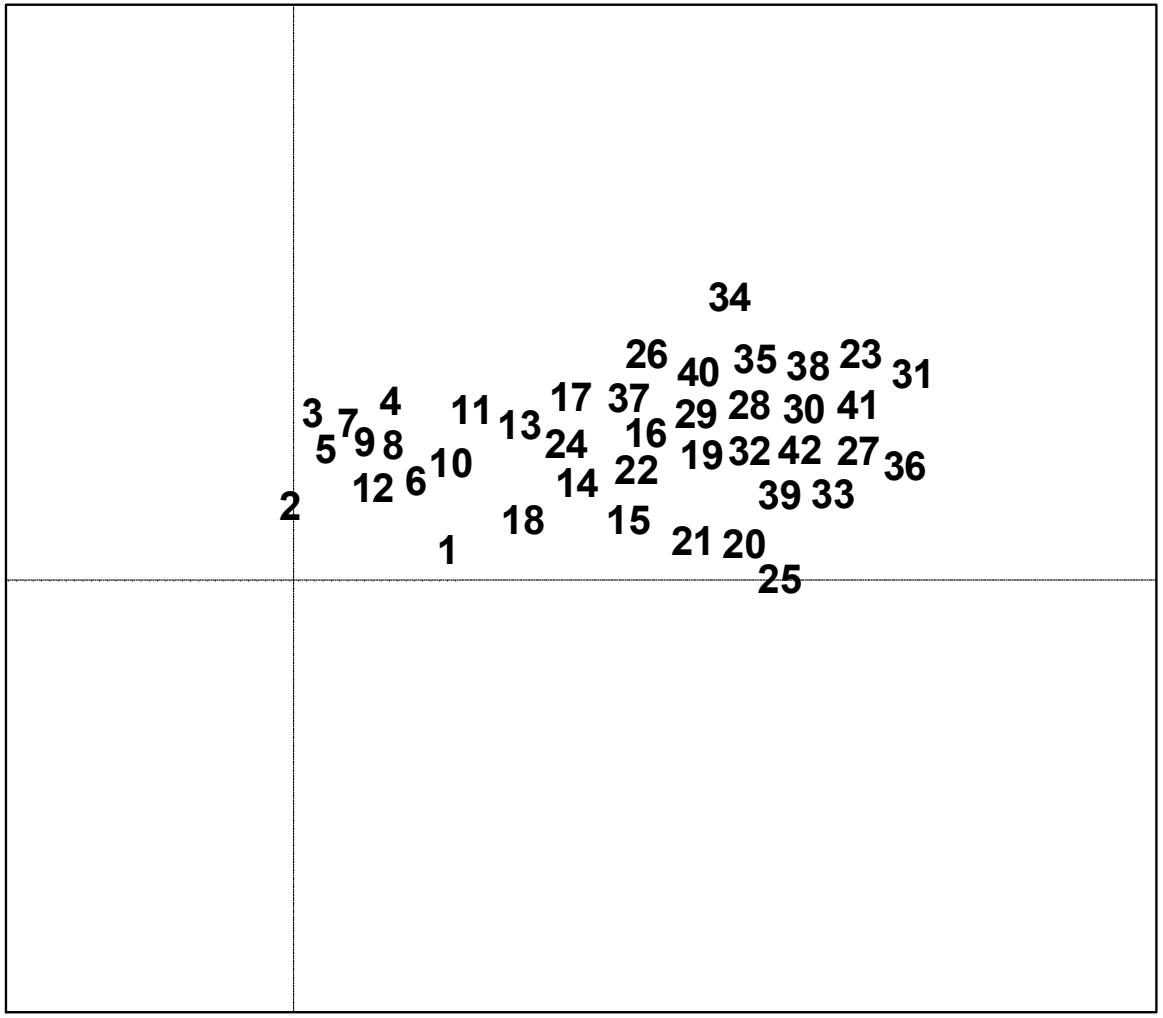
Detrended correspondence analysis

- DCA je postavená na unimodálnom modely
- 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 Haruštiaková

Podzim 2009

MU IBA Inštitút bioštatistiky a analýz, Masarykova univerzita

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.

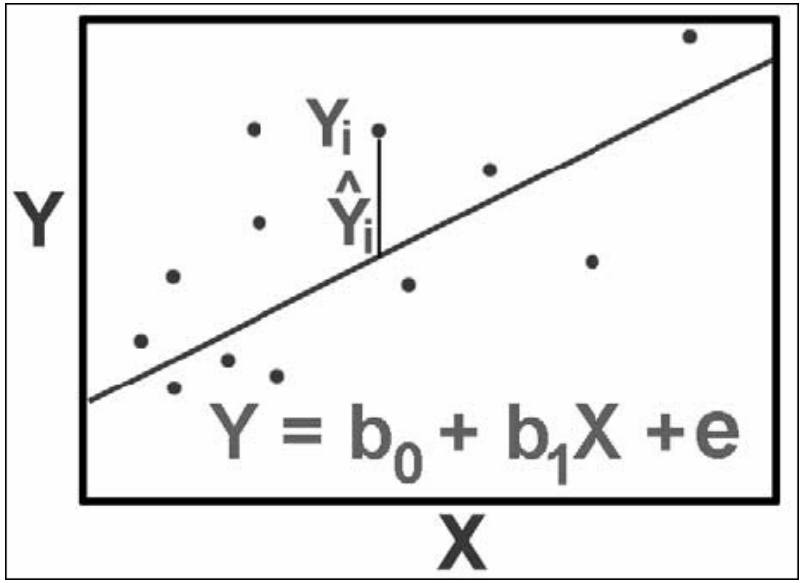
Ak je ich počet väčší než počet vzoriek zmenšený o jednu, tak sa ordinácia stáva nepriamou.

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 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 modely 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“.

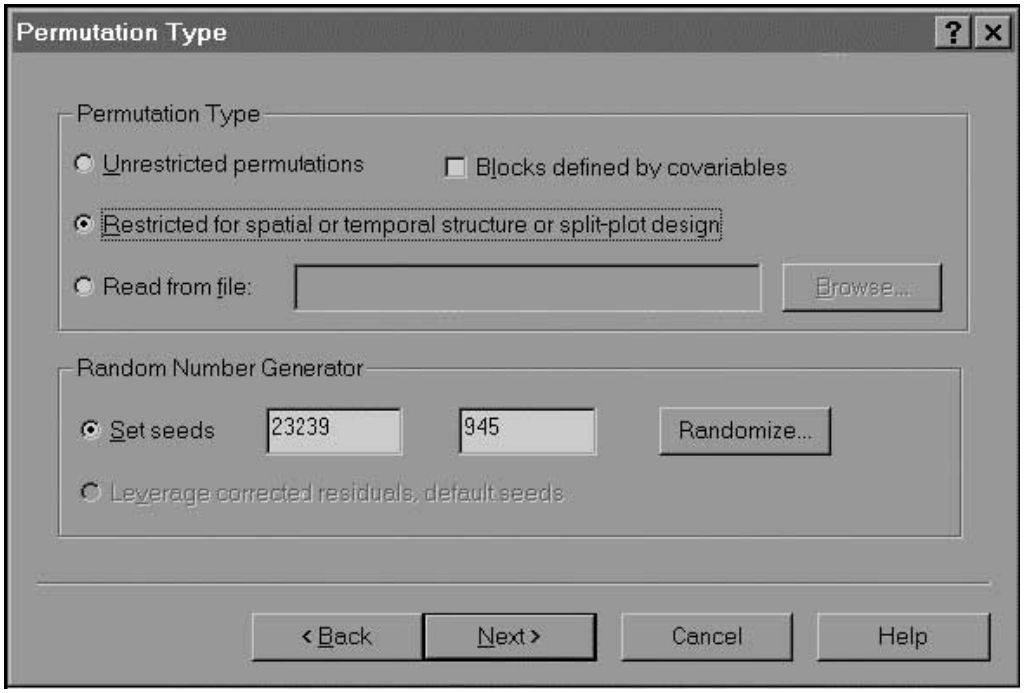
Priama gradientová analýza

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

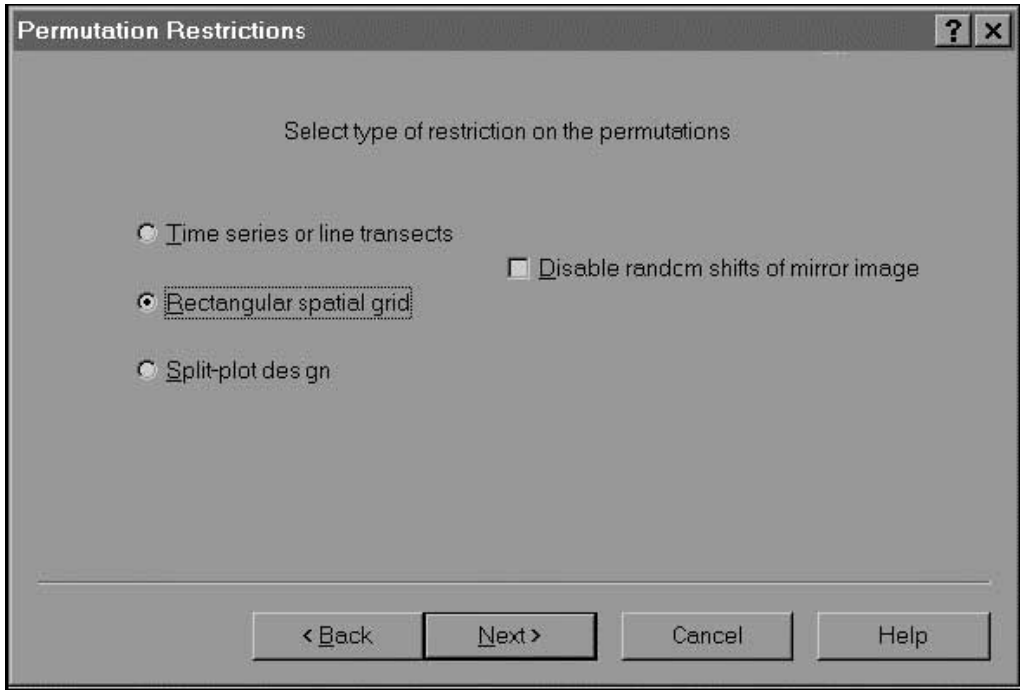
Výber štatisticky významných premenných

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

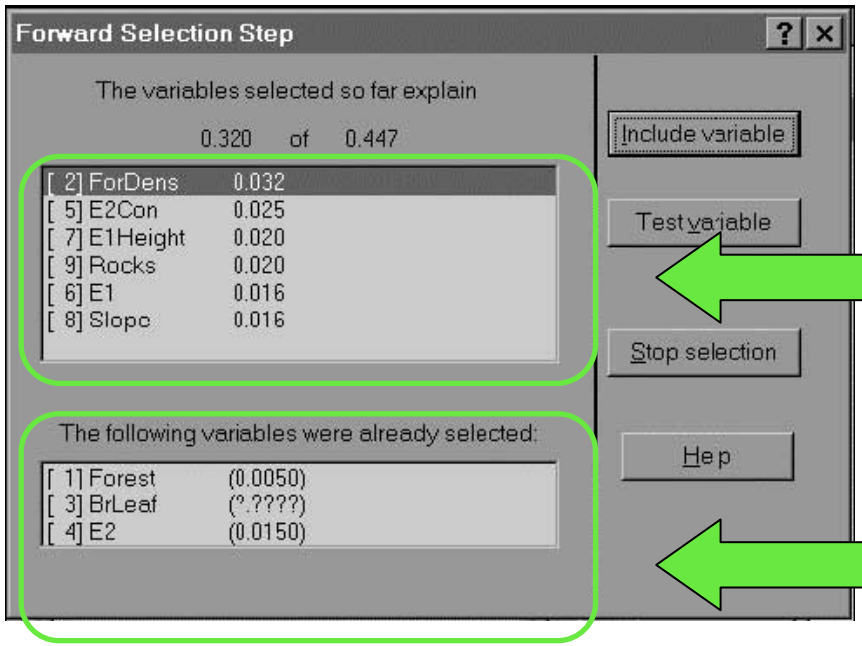


- vzorky pozdĺž časového alebo lineárneho transektu => permutácie „rotovaním“
- **split-plot design** permutácie v rámci bloku – ten je charakterizovaný niekoľkými nominálnymi premennými

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 Haruštiaková

Podzim 2009

Inštitút bioštatistiky a analýz, Masarykova univerzita

Redundačná analýza (RDA)

Direct gradient analysis — Redundancy analysis

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

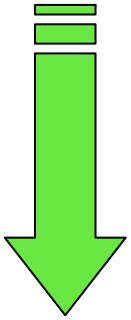
RDA je kanonická forma analýzy hlavných komponent (PCA)

- 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 Haruštiaková

Podzim 2009

Inštitút bioštatistiky a analýz, Masarykova univerzita

Kanonická korešpondenčná analýza (CCA)

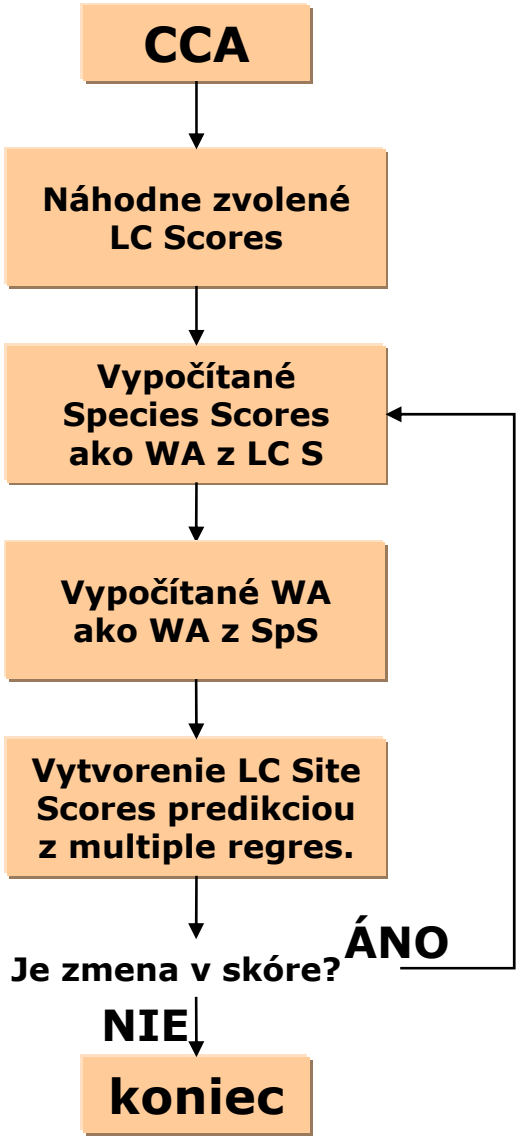
CCA je obmedzená ordinácia

- 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)

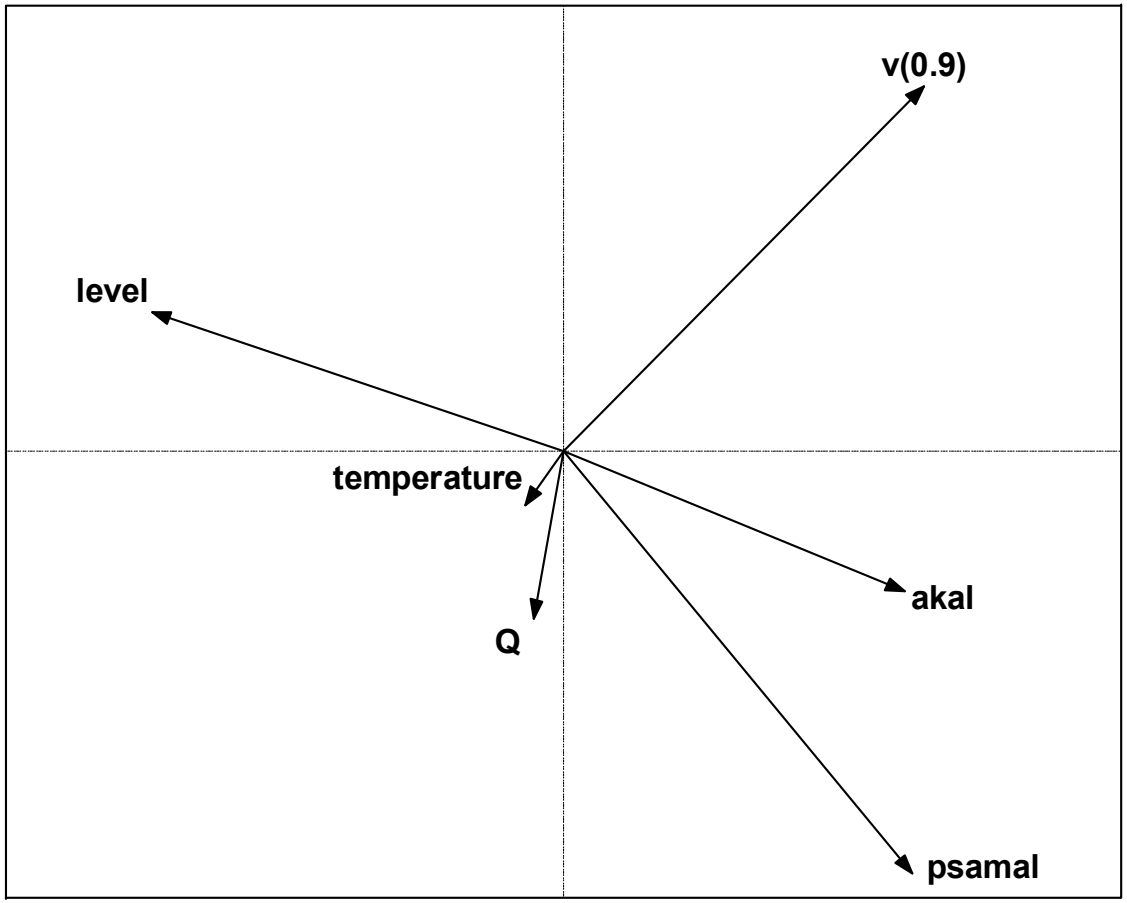
Direct gradient analysis

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 Haruštiaková

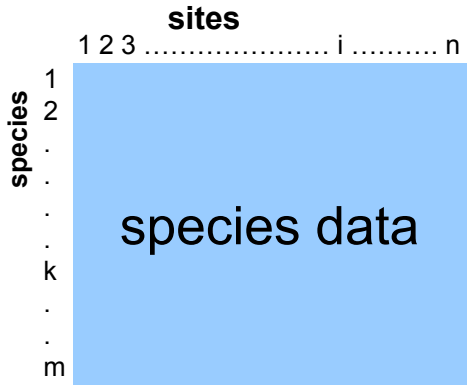
Podzim 2009

Inštitút bioštatistiky a analýz, Masarykova univerzita

Parciálna ordinácia

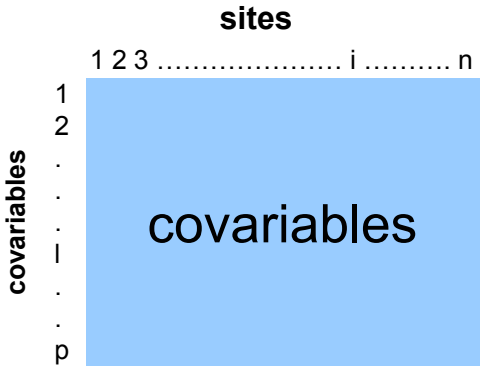
Indirect gradient analysis

Druhové dáta



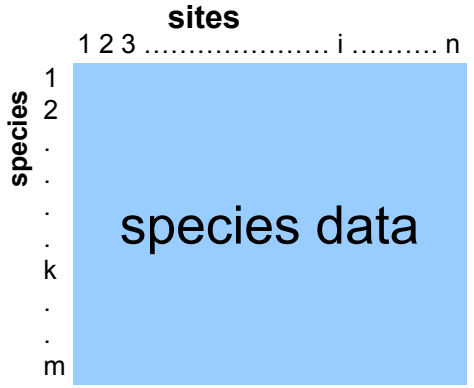
+

Kovariáty



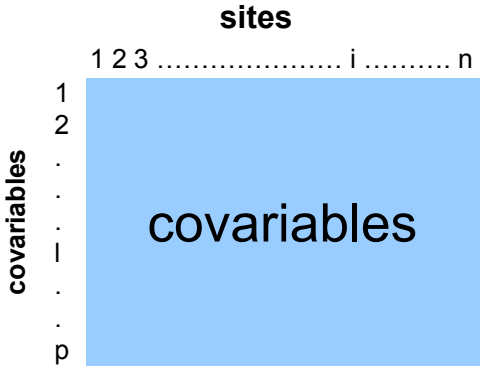
Direct gradient analysis

Druhové dáta



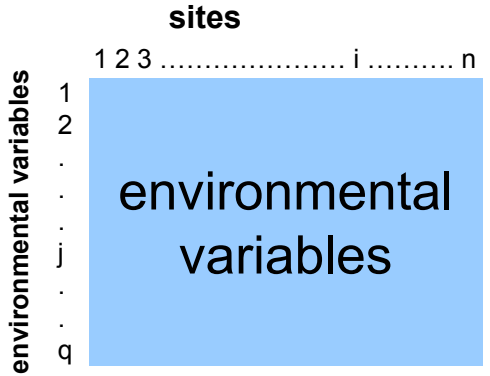
+

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)

→ druhové dáta + kovariáty

- Redundancy analysis (RDA)
- Canonical correspondence analysis (CCA)

→ druhové dáta + charakteristiky prostredia + kovariáty

Priame vs. nepriame ordinačné metódy

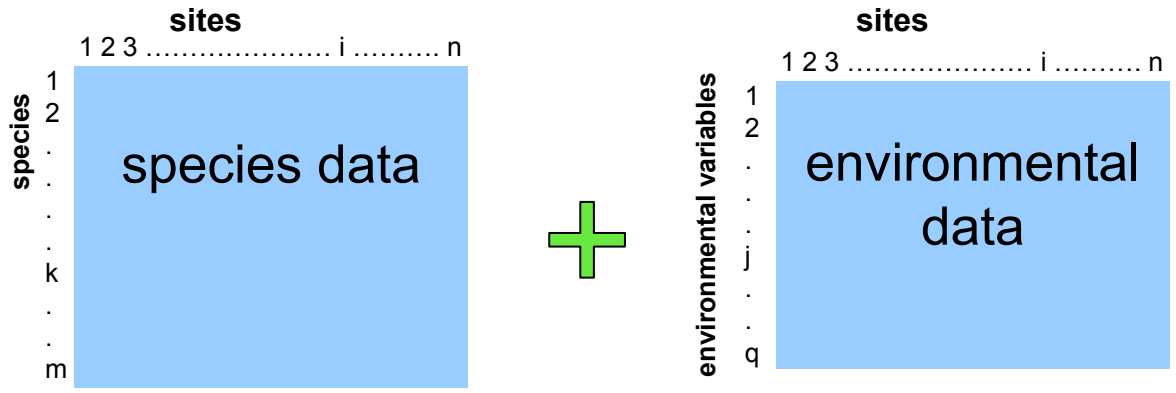
Danka Haruštiaková

Podzim 2009

Inštitút bioštatistiky a analýz, Masarykova univerzita

Priama či nepriama gradientová analýza?

Máme **druhové dáta** aj **charakteristiky prostredia**.



Môžeme použiť oba prístupy: priamu aj nepriamu ordináciu.

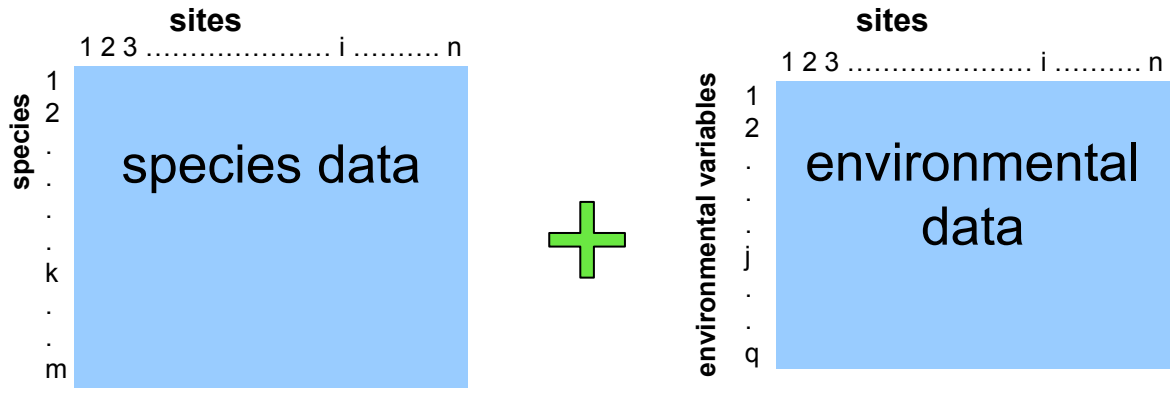
1. Spočítame najprv nepriamu ordináciu s následnou regresiou ordinačných osí na merané charakteristiky prostredia (tj. premietnutie týchto charakteristík do ordinačného diagramu)
2. Spočítame priamu (obmedzenú) ordináciu.

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.

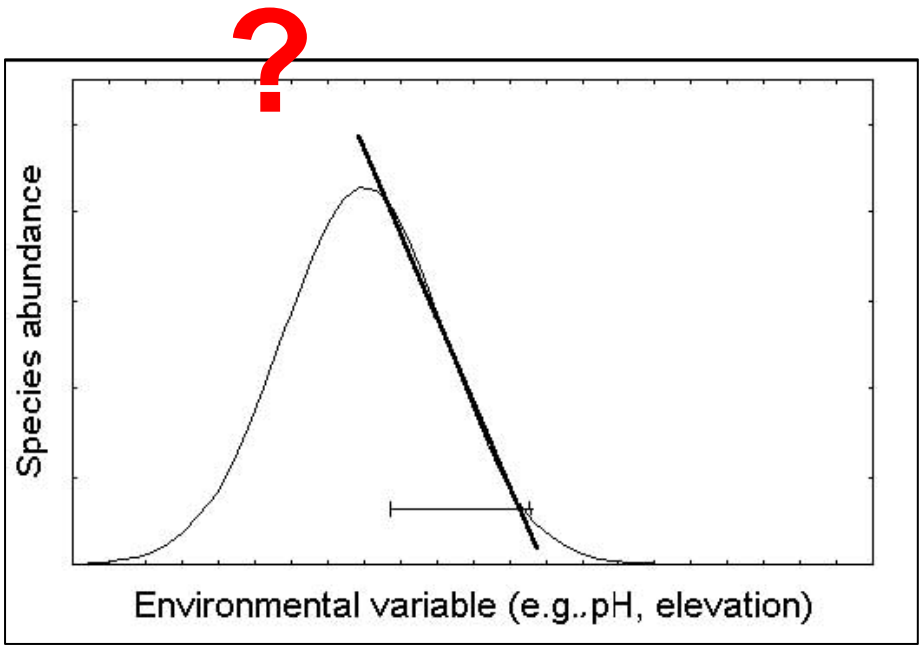
Lineárny alebo unimodálny model?

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)



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.

Diskriminačná analýza (CVA, DFA)

Samostatný PPT

Inštitút bioštatistiky a analýz, Masarykova univerzita

Neparametrická ordinácia (NMDS)

Danka Haruštiaková

Podzim 2009

Inštitút bioštatistiky a analýz, Masarykova univerzita

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 hocijakú 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 zhlukovania K-means (v prípadoch keď nie je možné použiť euklidovské vzdialenosti)

počet dimenzií musí byť určený vopred

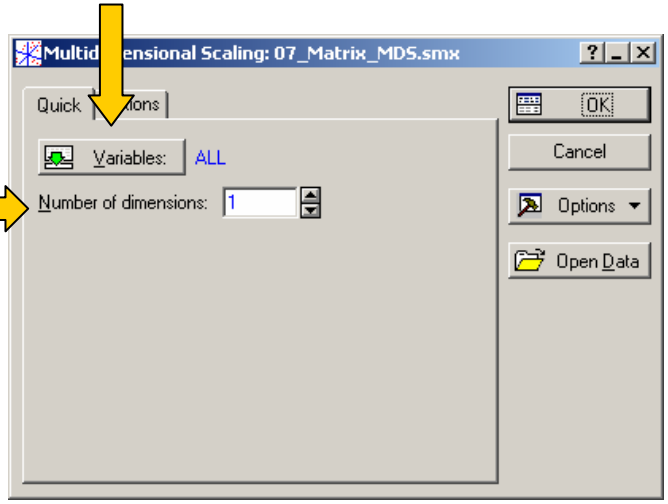
ťažko interpretovateľné výsledky

Mnohonásobné škálovanie v Statistica

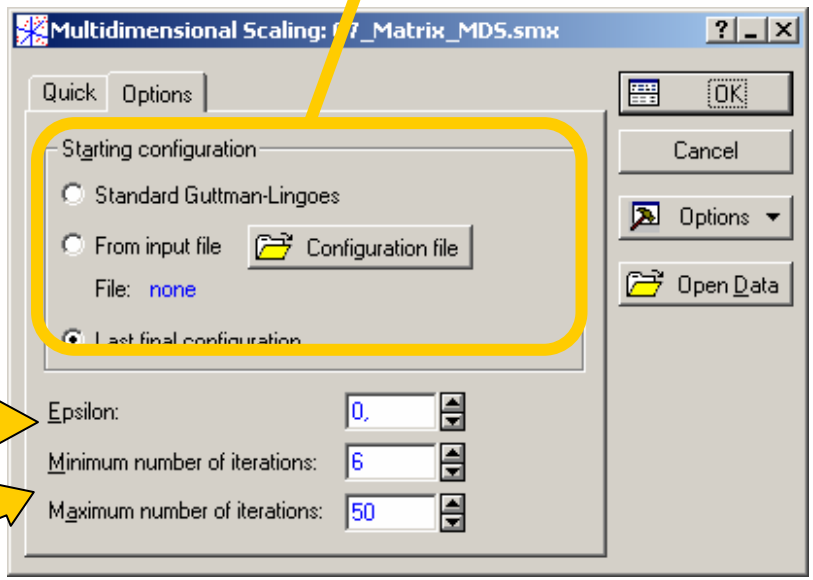
Multidimensional scaling dokáže na základě asociační matice s libovolnou metrikou vytvořit její Euklidovskou reprezentaci (příklad: na základě tabulky vzdáleností měst vytvoří mapu).

Výběr parametrů (vstupní soubor musí mít formát asociační matice)

Počet dimenzí k extrakci



Počáteční konfigurace



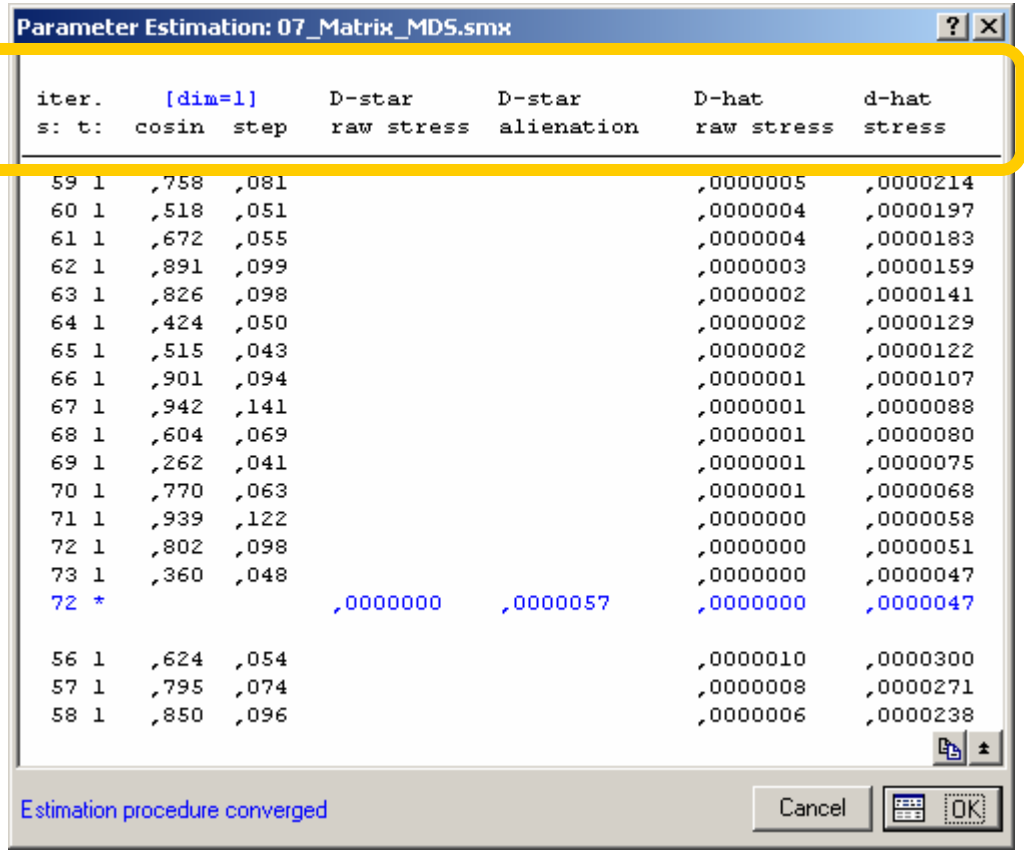
Vzdálenosti menší než jsou považovány za 0

Počty iterací

Mnohonásobné škálovanie v Statistica

Výpočet

Multidimensional scaling může sloužit pro přípravu podkladů pro k-means clustering pokud nemůžeme na naše data použít Euklidovskou vzdálenost. Metoda je výpočetně velmi náročná.



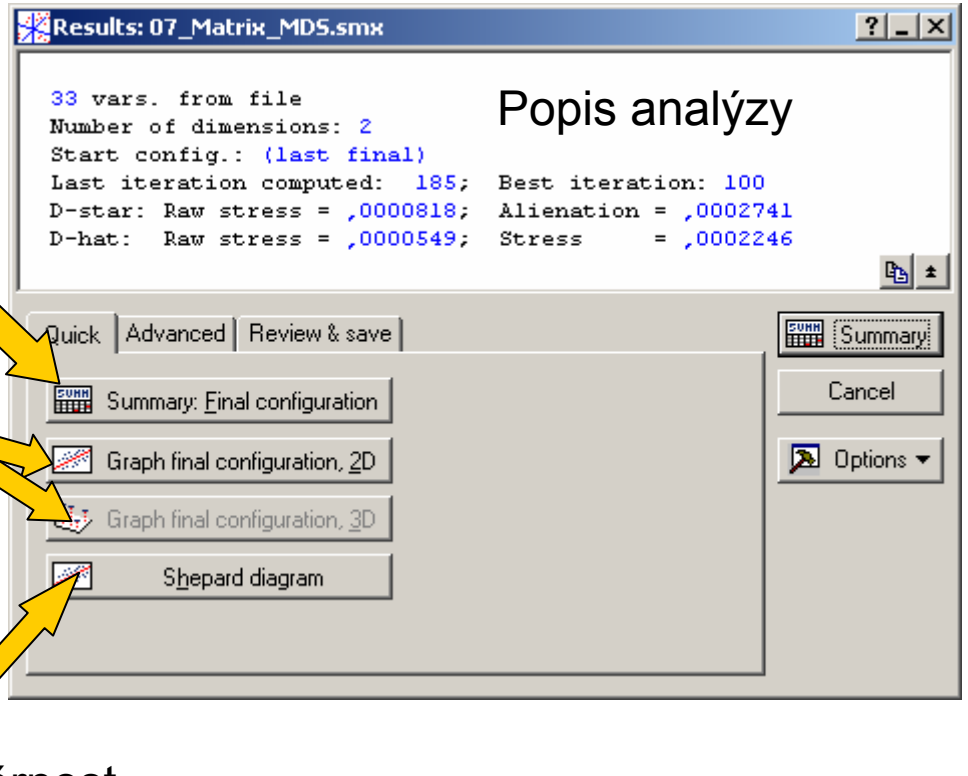
Parametry měnící se při přepočtech

Mnohonásobné škálovanie v Statistica

Výsledky Quick

Výstup nových dimenzí + charakteristiky

Výstupní 2D a D graf



Shephard diagram ~ věrnost reprezentace

Mnohonásobné škálovanie v Statistica

Výsledky tabulky

objekty

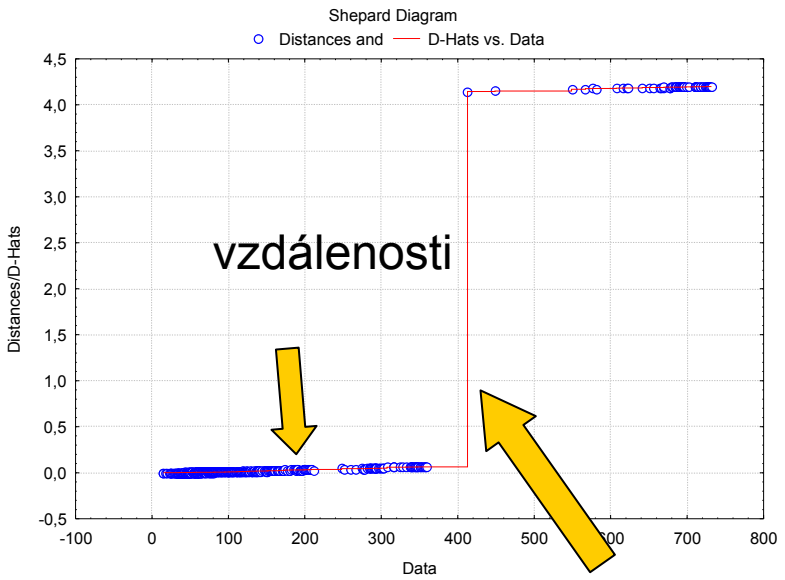
Final Configuration (07_Matrix_MDS.smx)
D-star: Raw stress = ,0000818; Alienation = ,0002741
D-hat: Raw stress = ,0000549; Stress = ,0002246

	DIM. 1	DIM. 2
HEL	-0,254837	0,00
HVE	-0,254777	,000982
MEL	-0,254272	0,000491
ROH	-0,255098	-0,002506

Nové dimenze

Stress – měřítko reprezentace, čím nižší, tím lepší reprezentace
Alienation – cizost, čím nižší, tím lepší reprezentace

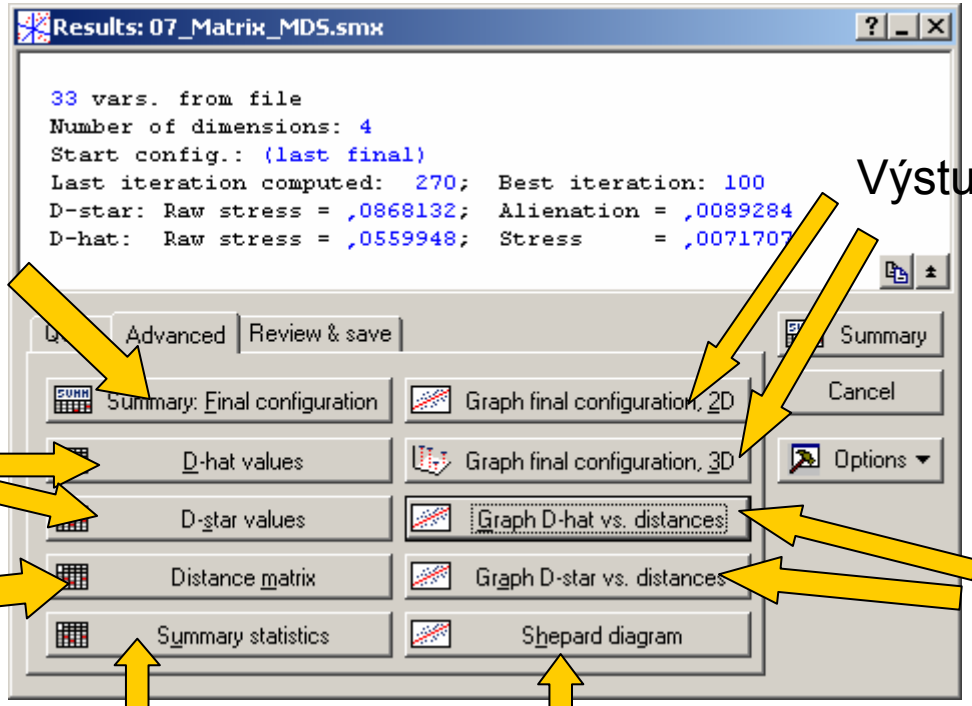
Shepard diagram



D-hat ~ průběh vzdáleností při dobré reprezentaci

Mnohonásobné škálovanie v Statistica

Výsledky Advanced



Výstup nových dimenzí + charakteristiky

Výstupní 2D a 3D graf

D-hat, D-star

Matice vzdáleností (reprodukovaná)

D-hat, D-star versus reprodukovaná vzdálenost ~ věrnost reprodukce

Sumární hodnoty (reprodukovaná vzdálenost, D-hat, D-star)

Shepard diagram

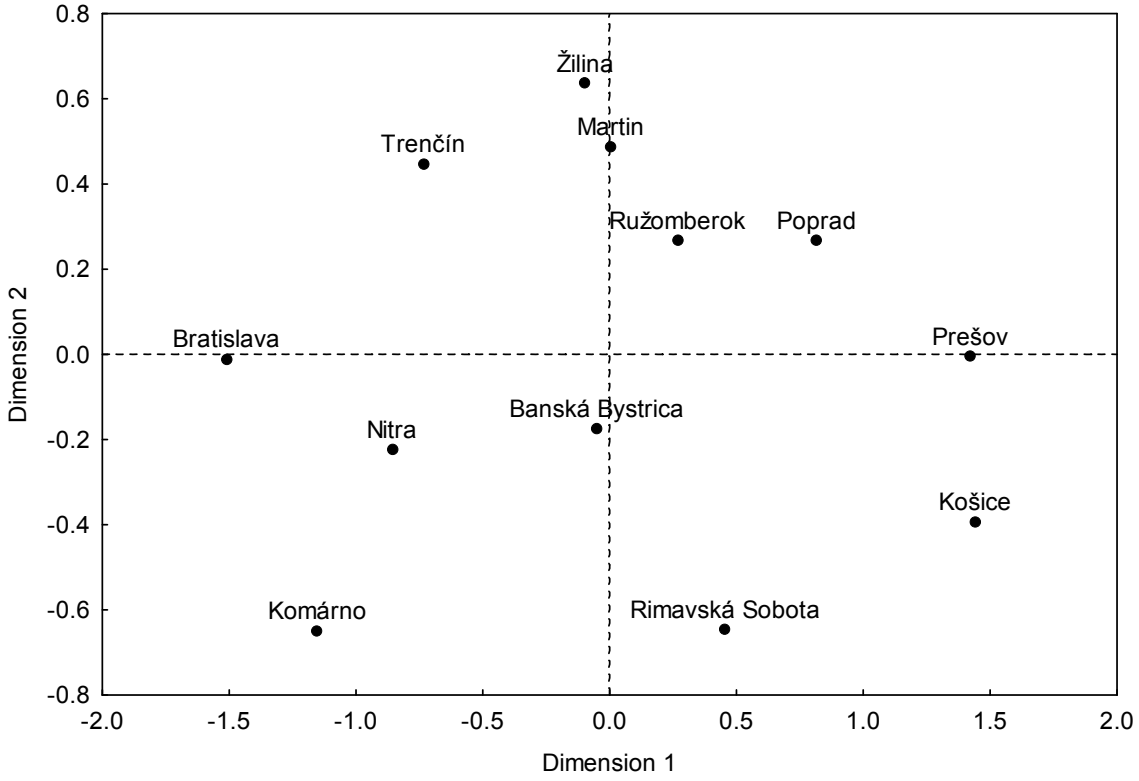
Mnohonásobné škálovanie – príklad

- máme k dispozícii maticu vzdialeností miest Slovenska z mapy
- cieľ: zreprodukovať vzdialenosti medzi mestami v dvojrozmernom priestore

	Banská Bystrica	Bratislava	Komárno	Košice	Martin	Nitra	Poprad	Prešov	Rimavská Sobota	Ružomberok	Trenčín	Žilina
B. Bystrica	0	204	188	214	92	119	124	208	105	53	139	117
Bratislava	204	0	100	402	227	85	328	412	273	257	124	202
Komárno	188	100	0	342	214	69	312	396	213	241	160	238
Košice	214	402	342	0	234	317	120	36	129	195	337	259
Martin	92	227	214	234	0	145	114	198	171	39	103	25
Nitra	119	85	69	317	145	0	243	327	188	172	91	169
Poprad	124	328	312	120	114	243	0	84	133	75	217	139
Prešov	208	412	396	36	198	327	84	0	165	159	301	223
R. Sobota	105	273	213	129	171	188	133	165	0	140	208	196
Ružomberok	53	257	241	195	39	172	75	159	140	0	142	64
Trenčín	139	124	160	337	103	91	217	301	208	142	0	78
Žilina	117	202	238	259	25	169	139	223	196	64	78	0

Mnohonásobné škálovanie – príklad

- Výsledok mnohonásobného škálovania



Mnohonásobné škálovanie – príklad

- Ukážka Shepardovho diagramu (príklad miest Slovenska)