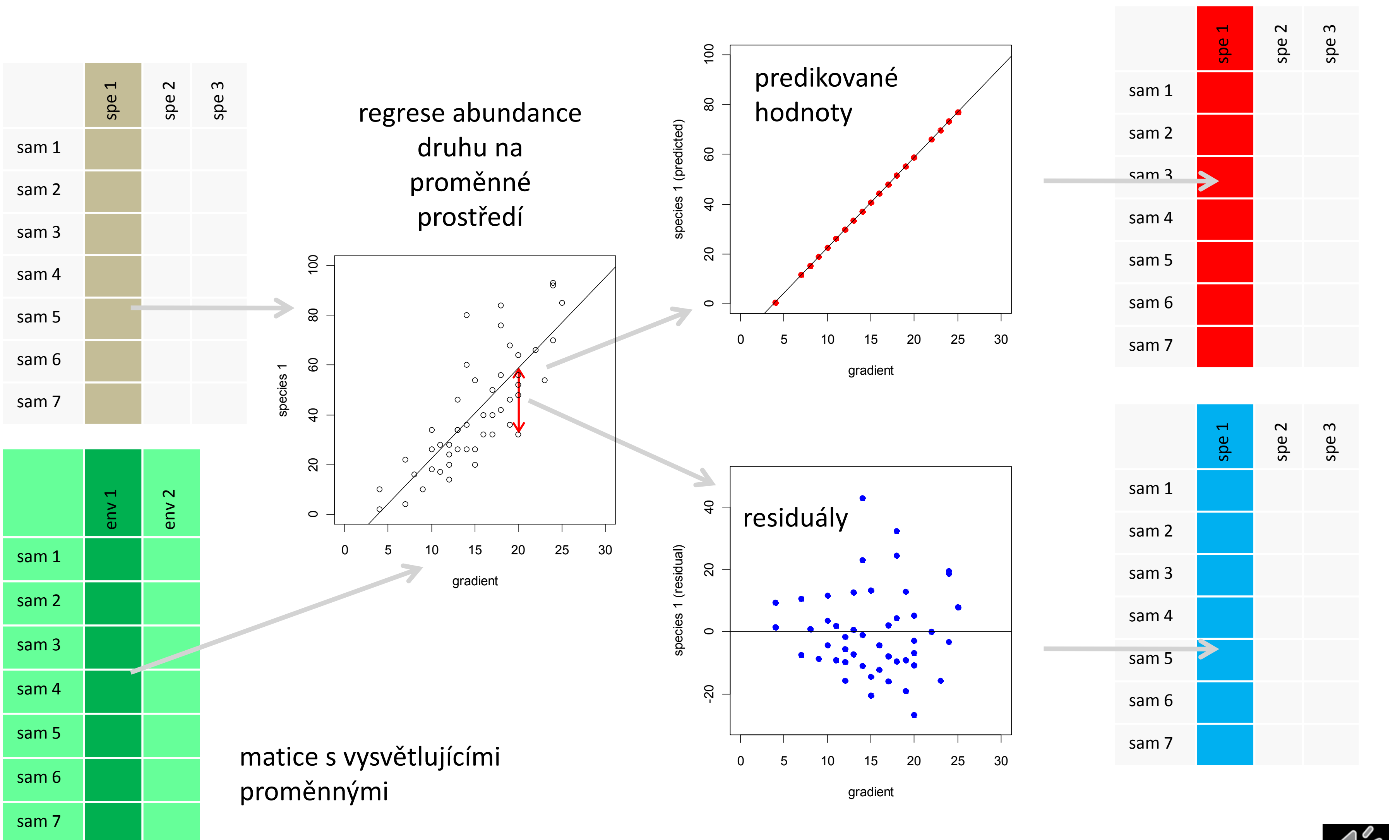
Přímá ordinační analýza

PŘEHLED METOD ORDINAČNÍ ANALÝZY

	raw-data-based (založené na primárních datech)		distance-based (založené na distanční matici)
	linear (lineární)	unimodal (unimodální)	
unconstrained (nepřímé)	PCA (analýza hlavních komponent)	CA, DCA (korespondenční a detrendovaná korespondenční analýza)	PCoA (analýza hlavních koordinát) NMDS (nemetrické mnohorozměrné škálování)
constrained (přímé)	RDA (redundanční analýza)	CCA (kanonická korespondenční analýza)	db-RDA (redundanční analýza založená na distanční matici)

PRINCIP PŘÍMÉ ORDINAČNÍ ANALÝZY (RDA)



OMEZENÉ ORDINAČNÍ OSY

- Skóre založené na fitovaných hodnotách z mnohonásobné. regrese druhů (RDA), nebo vzorků (CCA) na púroměnných prostředí
- Ordinační osy jsou lineárními kombinacemi prediktorů
- Reziduální variabilita je vyjádřena neomezenými osami, které jsou přítomné v každé přímé ordinaci.

PŘÍMÁ ORDINACE

INTERPRETACE VÝSLEDKŮ

RDA

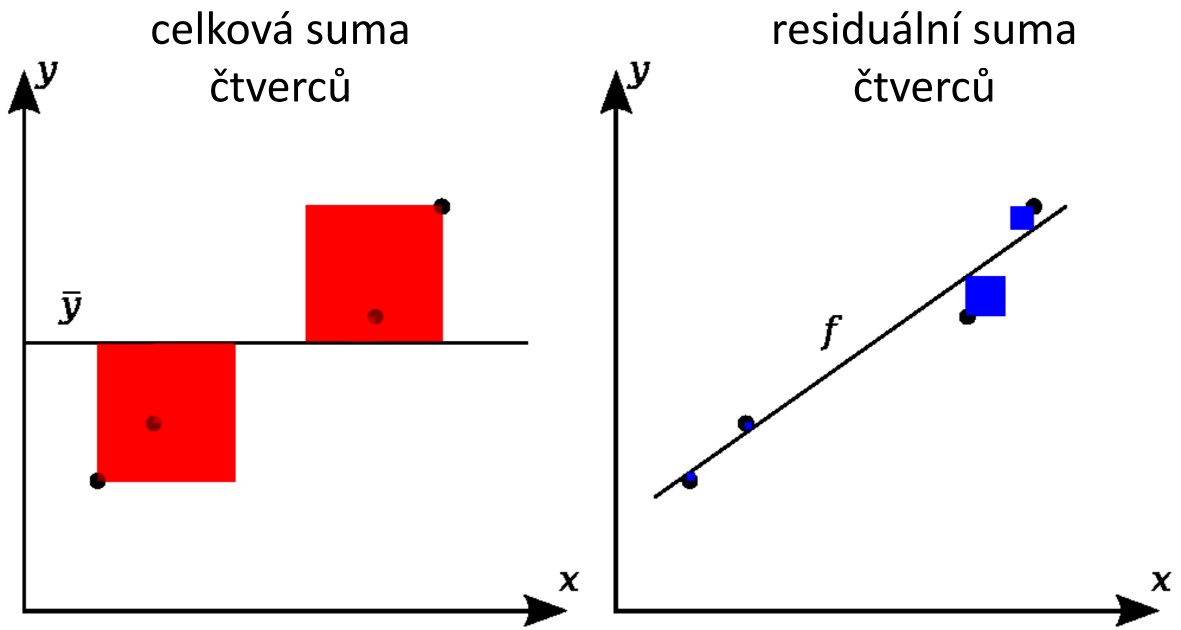
```
                Inertia Proportion Rank
Total            0.3783     1.0000
Constrained      0.1063     0.2808    2
Unconstrained    0.2721     0.7192   24
Inertia is variance

Eigenvalues for constrained axes:
   RDA1    RDA2
0.09240 0.01385

Eigenvalues for unconstrained axes:
    PC1     PC2     PC3     PC4     PC5     PC6     PC7     PC8
0.09543 0.03586 0.02688 0.02280 0.01706 0.01550 0.00999 0.00880
(Showed only 8 of all 24 unconstrained eigenvalues)
```

R^2 = 28.08 %

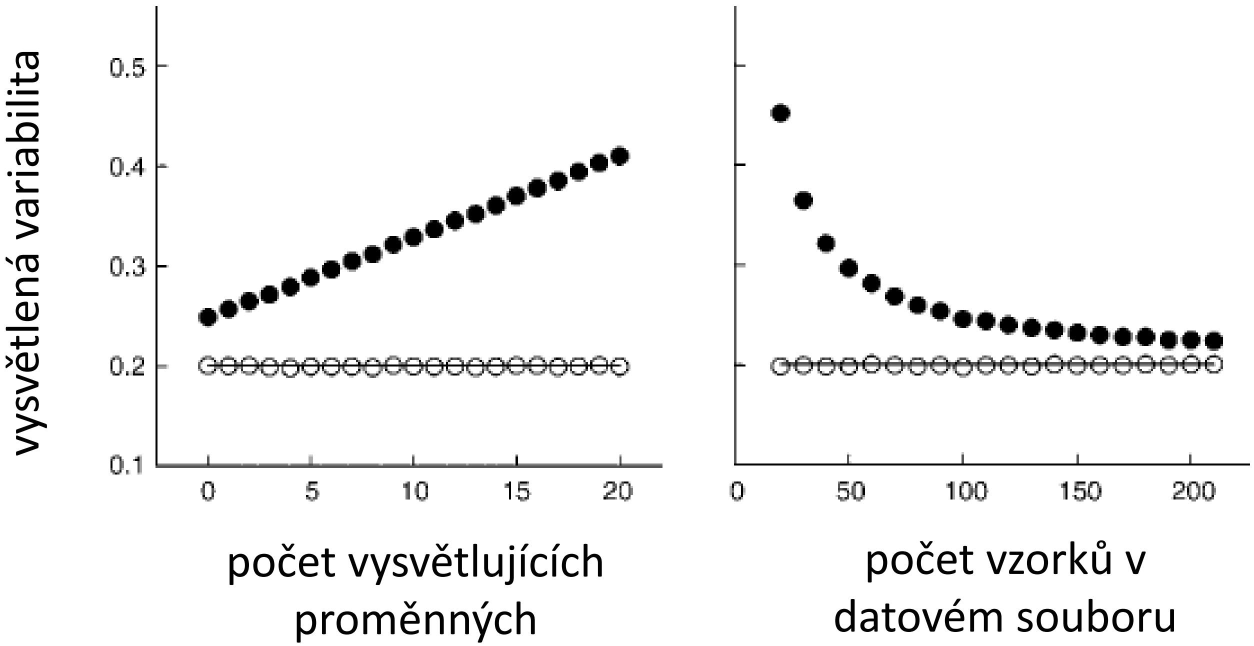
KOEFICIENT DETERMINACE V REGRESI



$$R^2 = 1 - \frac{SS_{\text{res}}}{SS_{\text{tot}}}$$

http://en.wikipedia.org/wiki/Coefficient_of_determination

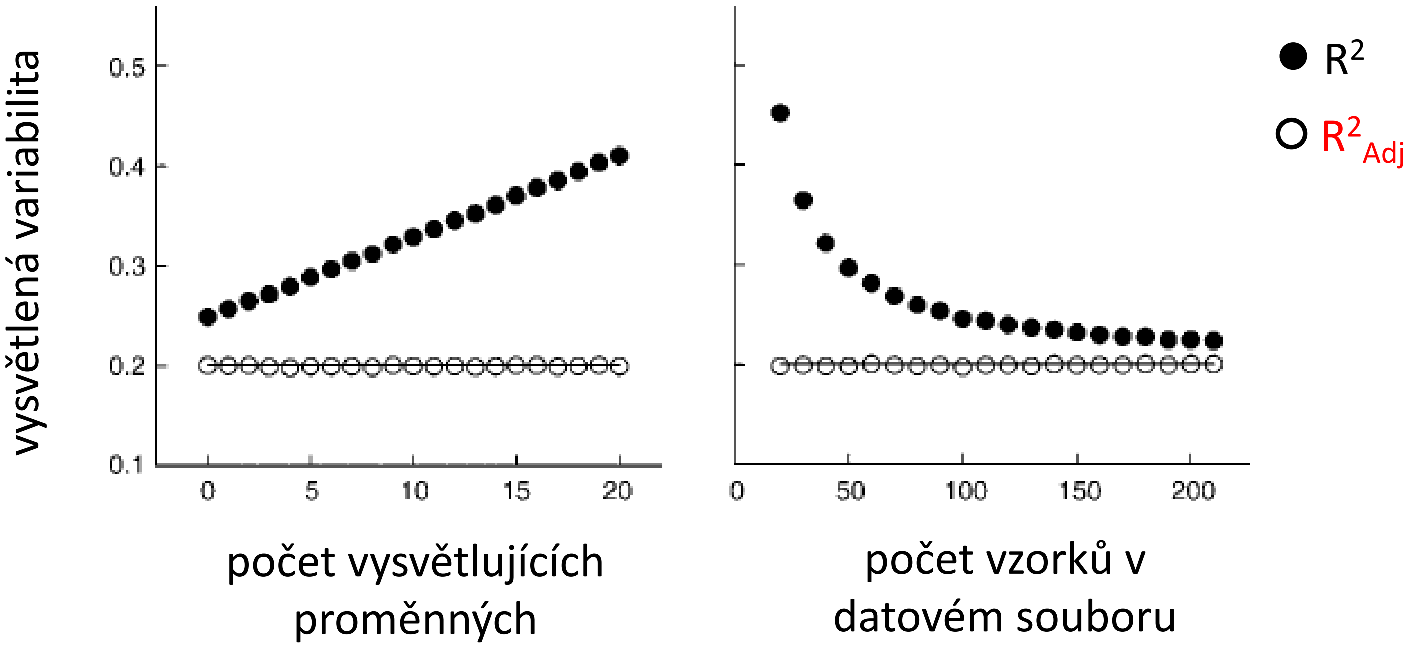
VYSVĚTLENÁ VARIABILITA (R^2)



- vysvětlená variabilita stoupá s počtem vysvětlujících proměnných (i když jsou náhodné) a klesá s počtem vzorků v datovém souboru
- platí pro přímou (kanonickou) ordinační analýzu i mnohonásobnou regresi

Peres-Neto et al. (2006) *Ecology*

VYSVĚTLENÁ VARIABILITA (R^2) A UPRAVENÝ (ADJUSTED) R^2



- upravený R^2 se nemění s počtem vysvětlujících proměnných a počtem vzorků v souboru

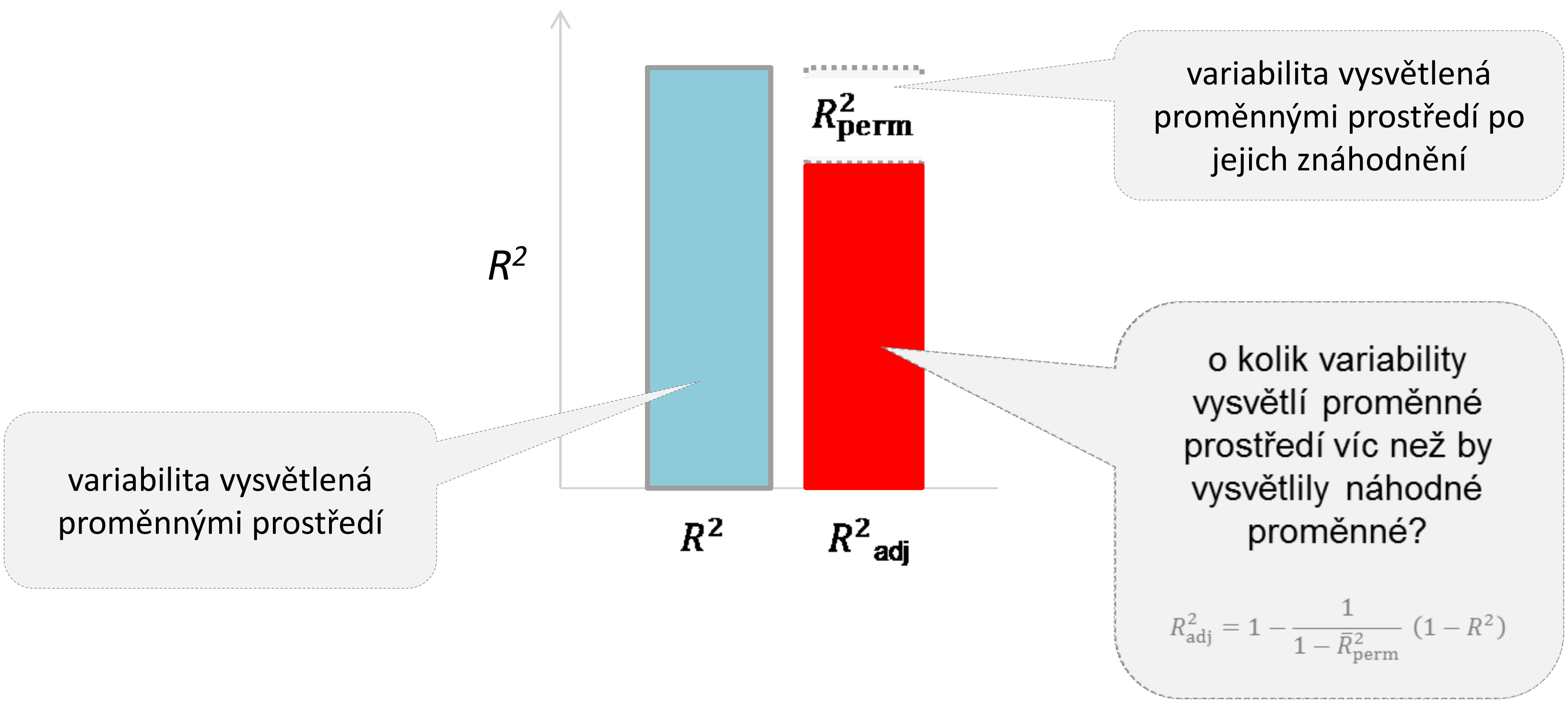
Peres-Neto et al. (2006) *Ecology*

Výpočet upraveného R^2 pomocí Ezekielovy formule (RDA)

$$R^2_{(\mathbf{Y}|\mathbf{X})\mathrm{adj}} = 1 - \frac{n-1}{n-p-1}\left(1 - R^2_{\mathbf{Y}|\mathbf{X}}\right)$$

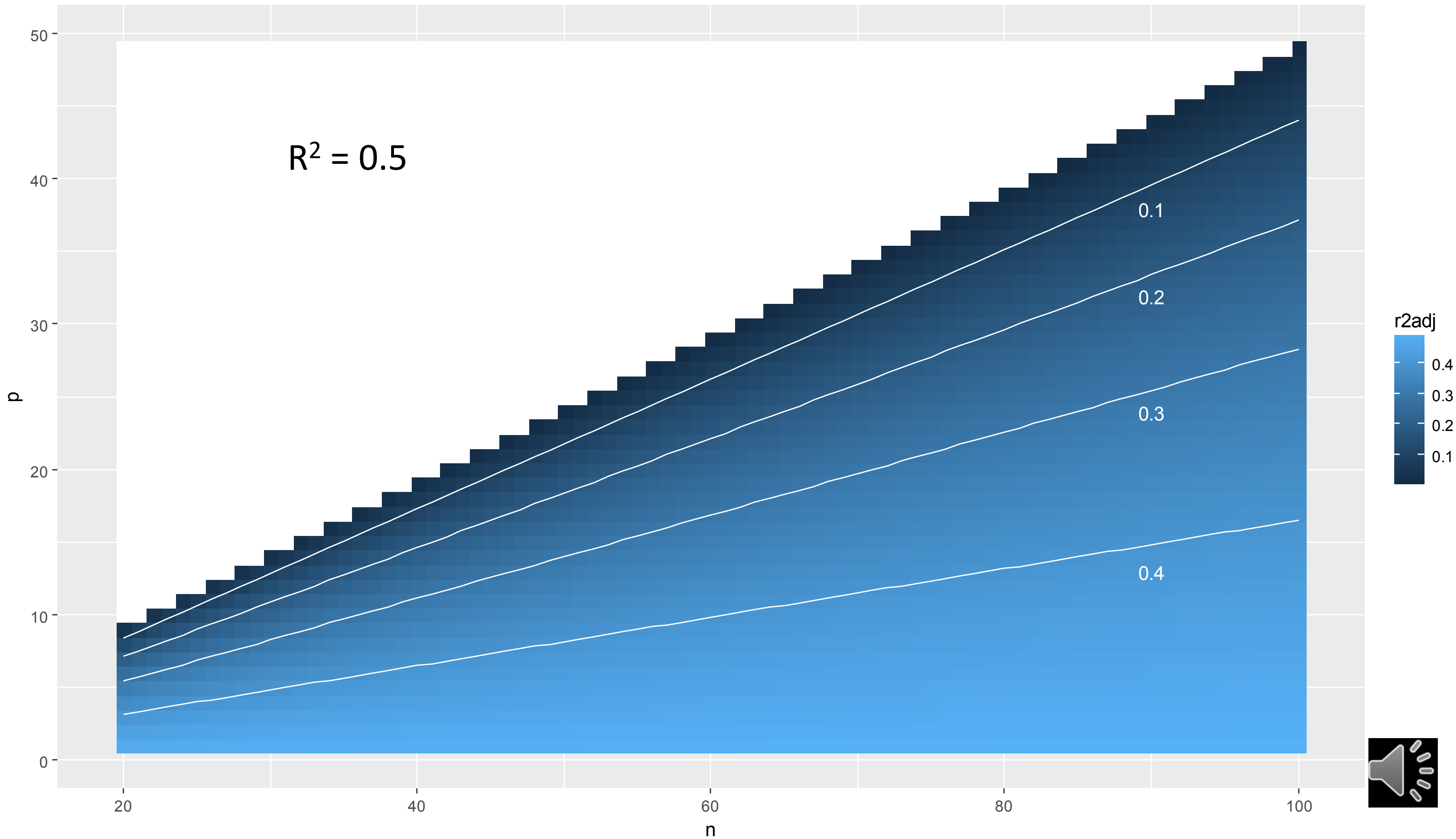
n ... počet vzorků
p ... počet vysvětlujících proměnných
$R^2_{Y|X}$ ... vysvětlená variabilita bez adjustace

Výpočet adjustovaného R^2 permutačním modelem (RDA, CCA)



UPREAVENÉ R^2 PRO KOMBINACE N A P

$$R^2_{(\mathbf{Y}|\mathbf{X})\mathrm{adj}} = 1 - \frac{n-1}{n-p-1}\left(1 - R^2_{\mathbf{Y}|\mathbf{X}}\right)$$



PŘÍMÁ ORDINAČNÍ ANALÝZA

MONTE-CARLO PERMUTAČNÍ TEST

- testuje nulovou hypotézu, že druhové složení je nezávislé na jedné nebo více vysvětlujících proměnných pomocí randomizace

ošetřeno	plocha	pokr.
1	1	8
1	2	7
1	3	7
1	7	6
1	8	5
1	9	6
0	4	5
0	5	5
0	6	6
0	10	1
0	11	2
0	12	2
	F =	0.413

RANDOMIZACE: Každé ploše se přiřadí náhodně jeden vegetační zápis (v našem případě jedna hodnota pokryvnosti). Tím se zruší jakákoliv závislost mezi proměnnou prostředí (tj. ošetřením) a hodnotami pokryvnosti.

Z hodnot pokryvnosti spočtu množství variability vysvětlené proměnnými prostředí (F-statistika). Pro hodnoty:

výchozí randomizované

ran.1	ran.2	ran.3	ran.4
7	2	5	7
6	5	8	2
5	6	1	5
1	8	6	7
6	1	7	6
6	5	6	2
8	5	6	5
5	6	7	1
5	7	2	6
2	6	2	5
7	2	5	6
2	7	5	8
0.175	0.005	0.005	0.076

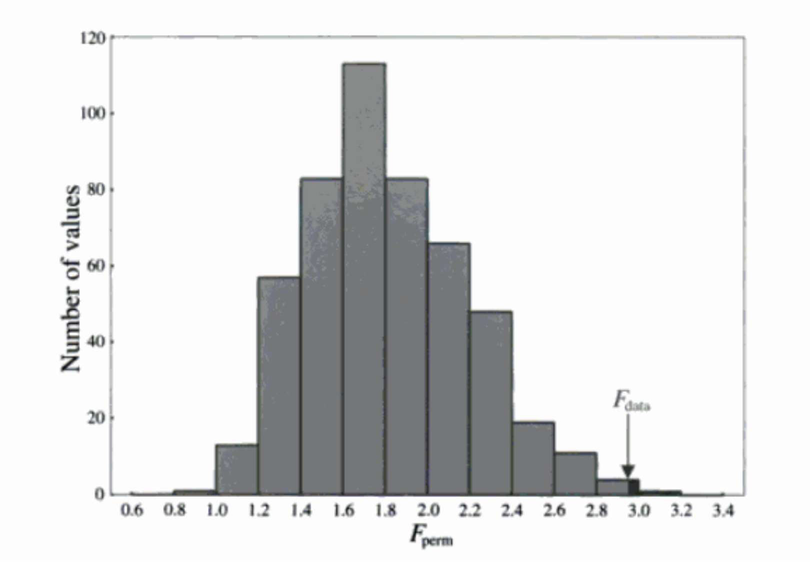
A tak dále. Až po dosažení kýženého počtu randomizací.

Herben & Münzbergová (2001)

PŘÍMÁ ORDINAČNÍ ANALÝZA

MONTE-CARLO PERMUTAČNÍ TEST

- Randomizace produkuje rozdělení hodnot, které by mělo testové kritérium za platnosti nulové hypotézy
- testuje nulovou hypotézu, že druhové složení je nezávislé na jedné nebo více vysvětlujících proměnných
- test první kanonické osy – vliv jen jedné kvantitativní proměnné
- test všech kanonických os – vliv všech proměnných, nebo vliv jedné kategoriální proměnné s více kategoriemi (počet os = počet kategorií – 1)
- testová statistika – **F_{data} (pseudo-F)**



$$P = \frac{n_x + 1}{N + 1}$$

P – hladina signifikance

n_x – počet permutací, kde F_{perm} >= F_{data}

N – celkový počet permutací

POSTUPNÝ VÝBĚR VYSVĚTLUJÍCÍCH PROMĚNNÝCH

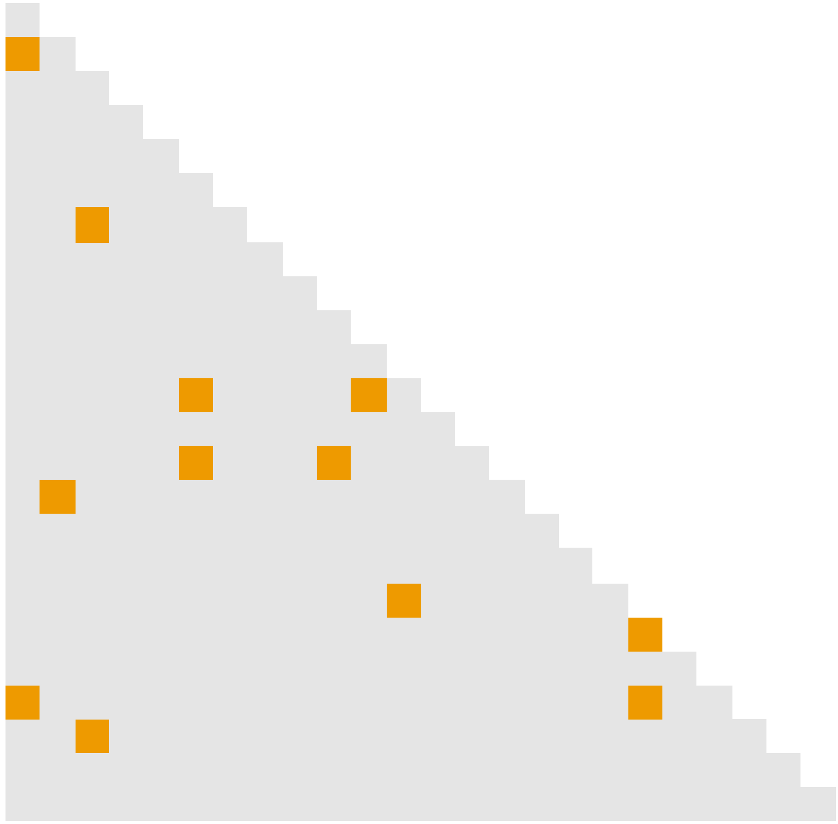
FORWARD SELECTION

- ze souboru vysvětlujících proměnných umožňuje vybrat jen ty, které mají průkazný vliv
- v každém kroku testuje zvlášť vliv jednotlivých proměnných (Monte-Carlo permutační test)
- vybere tu proměnnou, která vysvětlí nejvíce variability a zároveň je signifikantní; tuto proměnnou pak do modelu zahrne jako kovariátu
- v dalším kroku znovu testuje vliv jednotlivých proměnných na druhová data (s odstraněním vlivu kovariát) a opakuje předchozí kroky
- testy signifikance jsou zatíženy mnohonásobným porovnáním, a jsou proto poměrně liberální (počet signifikantních proměnných je často nerealisticky vysoký a vyžaduje např. Bonferroniho korekci)

Problém mnohonásobného porovnání
Forward selection příliš liberální

Simulace:

- 25 náhodně vygenerovaných proměnných
- otestování průkaznosti korelace každé proměnné s každou (čtvercová matice)
- průkazné korelace ($p < 0.05$) jsou označeny červeně
- dohromady 300 analýz, z nich je 16 průkazných



ŘEŠENÍ PŘÍLIŠ LIBERÁLNÍ FORWARD SELECTION

- globální test
- pokud významný, pak forward selection
- končí, pokud:
  - R^2_{Adj} má překročit globální R^2_{Adj}
  - by přidávaná proměnná nebyla významná

Blanchet et al. 2008

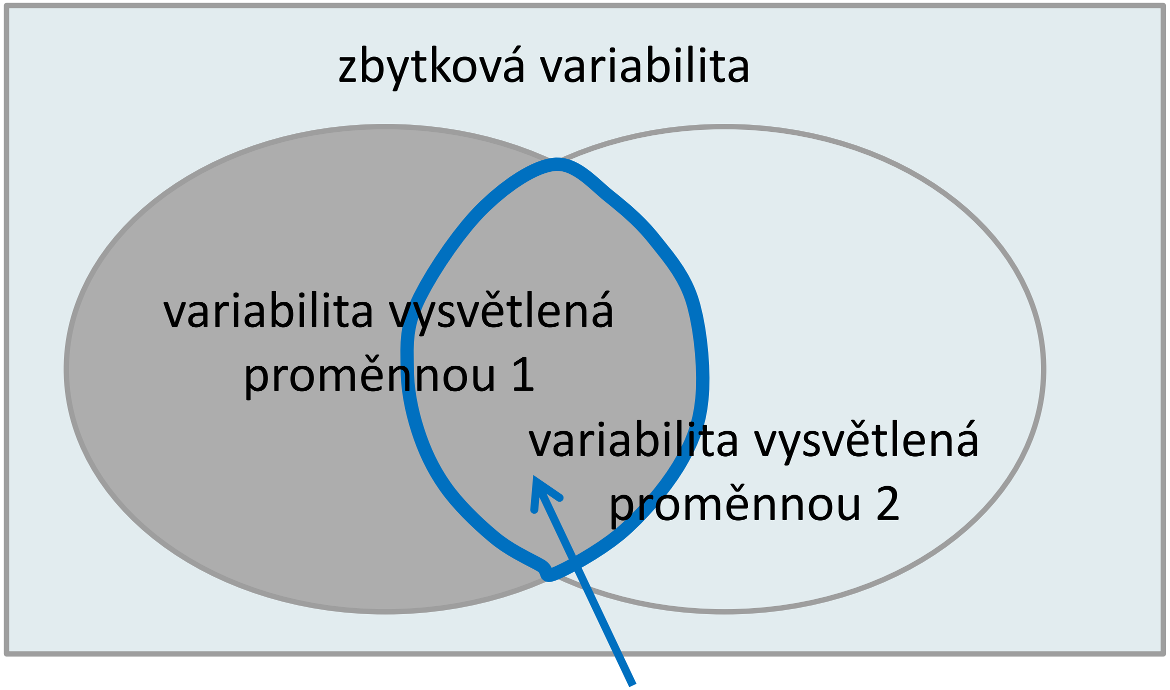
Parciální ordinace

PARTIAL ORDINATION

- odstraňuje část variability vysvětlené proměnnými, jejichž vliv chceme odečíst, ale ne přímo kvantifikovat (například vliv umístění ploch do bloků)
- následně se přímou nebo nepřímou ordinací analyzuje zbytková variabilita
- „odečítané“ proměnné se definují jako **kovariáty**
- pokud následuje přímá ordinace – ordinační osy představují čistý vliv ostatních vysvětlujících proměnných bez vlivu kovariát
- pokud následuje nepřímá ordinace – ordinační osy zachycují zbytkovou variabilitu v druhových datech po odstranění vlivu kovariát
- Pomocí kovariát můžeme dále testovat i čistý efekty více prediktorů, pokud jsou tyto navzájem korelované

ROZKLAD VARIANCE

VARIANCE PARTITIONING



vysvětlená variabilita sdílená proměnnou 1 a proměnnou 2 kvůli jejich vzájemné korelaci

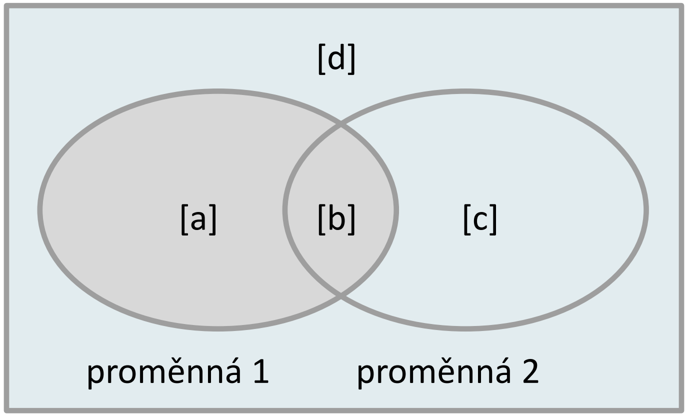
Borcard et al. 1992, *Ecology* 73: 1045–1055

ROZKLAD VARIANCE

VARIANCE PARTITIONING

Tři analýzy:

vysvětlující proměnná	kovariáta	vysvětlená variabilita
1 a 2	není	[a]+[b]+[c]
1	2	[a]
2	1	[c]



sdílená variabilita [b] = ([a]+[b]+[c]) – [a] – [c]
nevysvětlená variabilita [d] = *Total inertia* – ([a]+[b]+[c])

[a]+[b] – celkový (*marginal*) vliv proměnné 1
[a] – čistý (*partial, conditional*) vliv proměnné 1 (bez vlivu prom. 2)

Borcard et al. 1992, *Ecology* 73: 1045–1055