

Cvičení 8

1. Načtěte data o rostlinných společenstvech v meziřadí vinohradů na Jižní Moravě – dataset vineyards. Zkuste je klasifikovat pomocí různých přístupů.
 - a. Proveďte Hellingerovu standardizaci a poté k-means a clusterovou analýzu s využitím Wardovy metody na matici eukleidovských vzdáleností. Konečné počty shluků určete pomocí šířky siluety.
 - b. Upravte data logaritmickou transformací a matici nepodobností na základě Bray-Curtisova indexu. Poté analyzujte pomocí PAM a clusterové analýzy s využitím single-linkage, complete linkage a UPGMA algoritmů. Konečné počty shluků určete pomocí šířky siluety.
 - c. Porovnejte výsledky získaných klasifikací mj. i pomocí šířky siluety.
 - d. Pro 1-2 klasifikace, které se vám zdají pěkné, a. popište výsledky pomocí analýzy diagnostických druhů a b. zobrazte výsledky v ordinační analýze (PCA, PcoA, nebo NMDS)
2. *Samostatně*: Zopakujte úlohu č. 1 s daty o pakomárech ze Svratky.

R funkce

Nehierarchický partitioning

kmeans k-means partitioning

cluster: pam partitioning around medoids

Klastrová analýza

hclust klastrová analýza

plot(hclust.object) Vykreslení dendrogramu

cutree stanovení skupin na základě oříznutí dendrogramu nutno specifikovat buď k (počet skupin) nebo h (výšku dendrogramu), kde se má provést řez

rect.hclust nakreslení boxů reprezentujících skupiny do dendrogramu

Diagnostika - šířka siluety

cluster: silhouette

silhouette.obj\$avg.width průměrná šířka siluety dané klasifikace

plot(silhouette object) nakreslí distribuci šířek siluety pro jednotlivá pozorování.

Popis shluků

indicspecies: multipatt provede popis shluků pomocí diagnostických druhů. Pro phi-koefficient je třeba zadat parametr `func="r.g"`. Pokud nechceme, aby se popisovaly i kombinace shluků, tak zadáme paramter `duleg = T`

summary(multipatt.object) Vypíše uživatelsky příjemnější popis shluků - diagnostické druhy pro jednotlivé shluky seřazené

seřazené od toho s nejvyšší vazbou k danému shluku.