

KLASIFIKACE NA ZÁKLADĚ HLOUBKY BODU

Ondřej Vencálek

Přírodovědecká fakulta

Univerzity Palackého v Olomouci

2.2.2010

Co na posteru je ...

Klasifikační úloha

Uvažujeme směs J různých pravděpodobnostních rozdělení P_1, \dots, P_J na \mathbb{R}^d .

Máme tzv. tréninkovou množinu:

pozorování č.	pozorované hodnoty	skupina
1	x_{11}, \dots, x_{1d}	g_1
2	x_{21}, \dots, x_{2d}	g_2
\vdots	\vdots	\vdots
N	x_{N1}, \dots, x_{Nd}	g_N

kde $g_i \in \{1, \dots, J\}$.

Náš cíl: přiřadit nové pozorování \mathbf{x} k některému z uvažovaných rozdělení.

Metody využití hloubky v klasifikační úloze

- ▶ Metoda maximální hloubky
- ▶ Metoda založená na DD-plotu
- ▶ Metoda založená na srovnání distribucí hloubek

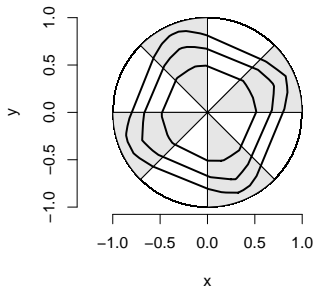
Optimalita

Tyto metody jsou asymptoticky optimální
(v Bayesovském smyslu) za předpokladu:

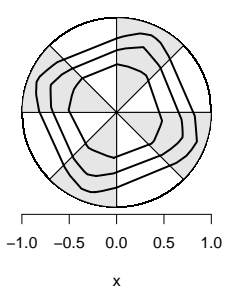
- ▶ elipticky symetrických rozdělání
- ▶ unimodálních
- ▶ lišících se jen parametrem polohy (stejná var. matice)
- ▶ stejných apriorních rozdělání

Co na posteru není ...

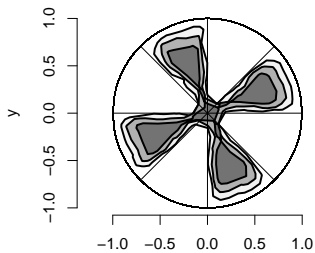
Distribution P1



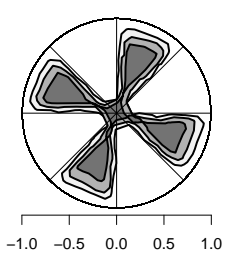
Distribution P2



Distribution P1



Distribution P2



Misclassification rates, N = 1000

