

Predstavitev domačih nalog

Neža Kržan, Tom Rupnik

1 Diskriminantna analiza

1.1 Slide 1

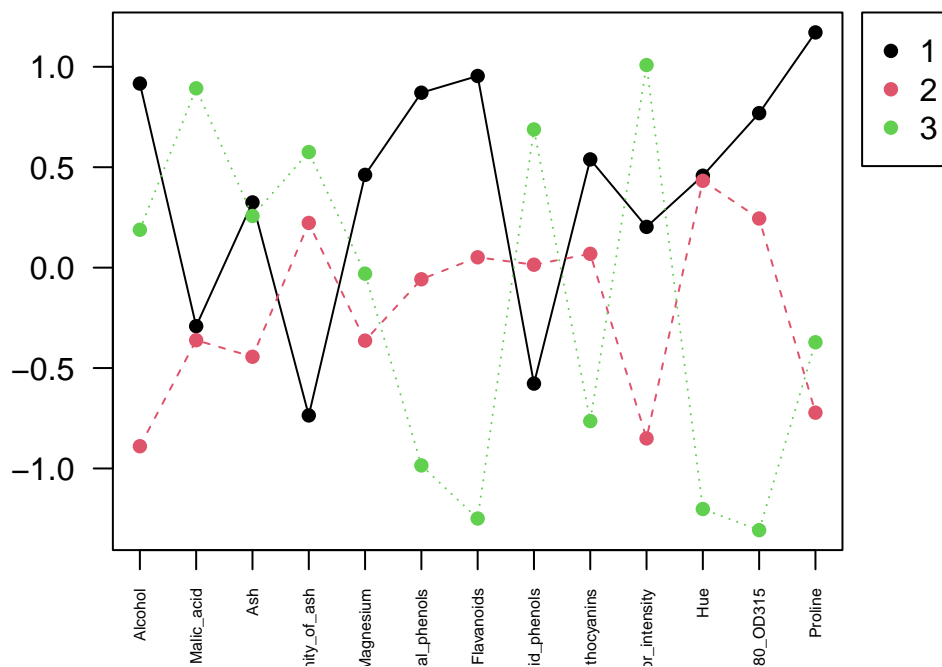
Podatkovni okvir z nalovom Wine - vsebuje 13 različnih lastnosti vina iz kemijske analize treh različnih sort. Iz teh komponent želimo določiti kateri sorti pripada vino.

Za skupine imamo sorte vina, označene so z številkami 1, 2 in 3 in posameznih vin je glede na sorto:

Tabela 1: Velikost skupin.

sorta vina	število vin	delež %
1	59	33.15
2	71	39.89
3	48	26.97

Poglejmo si še kako se sorte vina razlikujejo v povprečjih glede na lastnost na grafu, s katerega lahko vidimo, da večina spremenljivk dobro loči med sortami vina, je pa nekaj takih kjer to ne velja.

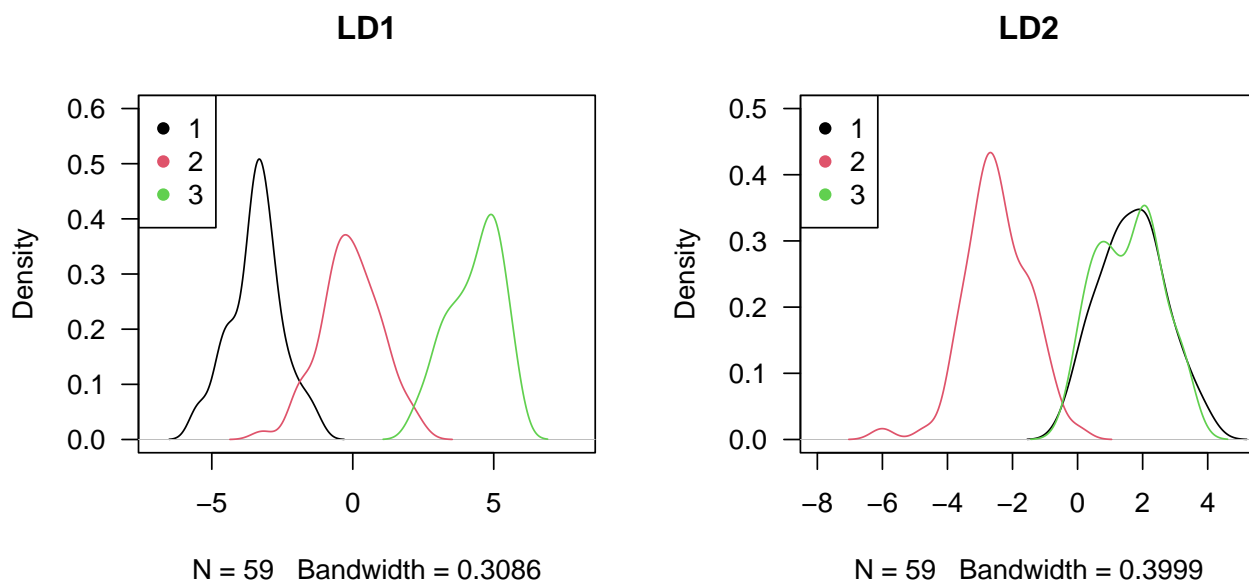


Slika 1: Povprečja neodvisnih spremenljivk po skupinah (standardizirane vrednosti).

1.2 Slide 2

Izvedli smo linearno diskriminantno analizo na standardiziranih podatkih in zanimalo nas je med katerimi sortami vina bolje ločuje posamezna diskriminantna funkcija.

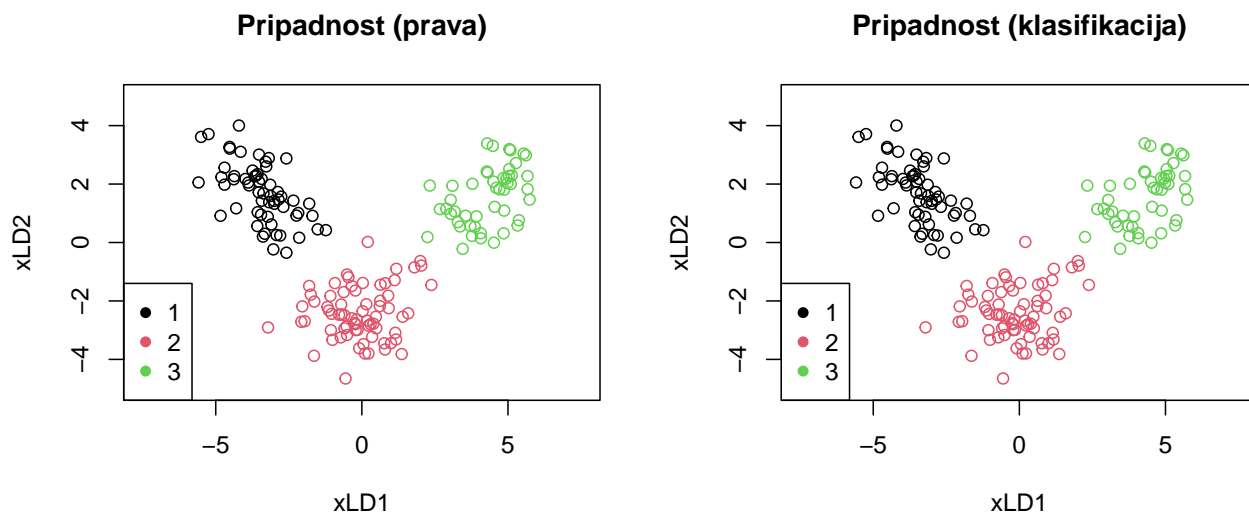
Prva linearna diskriminantna funkcija je najbolje ločevala med sorto vina 1 in sorto vina 3. Druga linearna diskriminantna funkcija pa najbolje ločuje med sorto vina 1 in 2 ter 3 in 2, kar se zelo dobro vidi na spodnjem grafu.



Slika 2: Grafični prikaz porazdelitev vrednosti po skupinah.

Pri prvi diskriminantni funkciji je imela intenziteta barve pomembno vlogo, pri drugi pa na primer ne.

1.3 Slide 3



Slika 3: Prikaz enot glede na pravo pripadnost sorti vina(levo) in pripadnost posamezni skupini glede na klasifikacijo po LDA(desno).