STROJNO UČENJE

2. Domaća Zadaća

Krešimir Špes
0036419866
ak. god. 2011. / 2012.

a) prvo ćemo procijeniti parametar μ_{ML} i $\sum_{ML}$ po sljedećim formulama:

$$\hat{\boldsymbol{\mu}}_{\mathrm{ML}} = \frac{1}{N}\sum_{i=1}^{N}\mathbf{x}^{(i)}. \qquad \hat{\boldsymbol{\Sigma}}_{\mathrm{ML}} = \frac{1}{N}\sum_{i=1}^{N}(\mathbf{x}^{(i)} - \hat{\boldsymbol{\mu}}_{\mathrm{ML}})(\mathbf{x}^{(i)} - \hat{\boldsymbol{\mu}}_{\mathrm{ML}})^{\mathrm{T}}$$

μ_{ML} = [-0.5320 -0.6550 -0.5560 -1.7550]

Radi korekcije pristranosti $\sum_{ML}$ ćemo korigirati sa N/(1 – N) zbog malog broja primjera

$$\sum_{ML} = \begin{bmatrix} 2.4418 & 0.6105 & 1.2935 & 0.0102 \\ 0.6105 & 0.5002 & 0.2914 & 0.1758 \\ 1.2935 & 0.2914 & 1.9438 & -0.0264 \\ 0.0102 & 0.1758 & -0.0264 & 0.4008 \end{bmatrix}$$

vrijednosti su izračunate pomoću matlaba, kod je dan u zad1_a.m datodeci.

b)

Mahalanobisova udaljenost: **4.0301**

Ekulidska udaljenost: **2.8284**

Nije jednako ekulidskoj udaljenosti jer je Mahalanobisova udaljenost neosjetljiva na razlike u varijanci izmedu pojedinih dimenzija i na korelacije izmedu varijabli.

Rekao bih da je to dobro jer nam u statističkom pogledu Mahalanobisova udaljenost daje iskoristljiviju mjeru udaljenosti. To je pogotovo korisno u više od 3 dimenzije jer euklidske udaljenosti ionako nemožemo vizualizirati za više od 3 dimenzije.

c)

Varijable su statistički nezavisne ako vrijedi P(X, Y) = P(X) * P(Y)

Kako bi dokazali (ne)zavisnost varijable, izračunamo kovarijaciju dviju varijabla. Tako je kovarijacija za x1 i x2 = -0.252568 što znači da nisu nezavisne.

Ne možemo definitivno odgovoriti da li su varijable nezavisno jer iako možemo dokazati da su varijable nekorelirane, to ne znači da varijable nisu nelinearno korelirane (koeficijent korelacije mjeri samo linearnu zavisnost varijabli)

U nastavku je dan python kod za izračun kovarijacije varijabli x1 i x2:

```python
x1 = [-1.28,-1.83,3.24,-2.47,-1.07,-1.58,-1.00,-0.53,0.50,0.70]
x2 = [-1.51,-0.99,0.11,-1.95,0.15,0.06,-0.95,-0.67,-0.91,0.11]

E1 = 0
for i in x1:
	E1 += i
E1 /= len(x1)

E2 = 0
for i in x2:
	E2 += i
E2 /= len(x2)

E12 = 0
for i in range(0,len(x1)):
	E12 += x1[i]*x2[i]
E12 /= len(x1)*len(x1)

cov = E12 - E1*E2

print "cov:", cov
```