

Rješenje zadatka 5.2 predmeta Strojno učenje

Siniša Biđin

5. veljače 2013.

- (a) Algoritam k-srednjih vrijednosti koristi se prilikom tzv. “čvrstog” grupiranja gdje svaki primjer pripada samo jednoj grupi. Kod algoritma maksimizacije očekivanja (EM-algoritam) grupiranje je “meko” i probabilističko te primjer stoga može pripadati u više grupa.

(b)

$$p(\mathbf{x}) = \sum_{k=1}^K \pi_k p(\mathbf{x}|\boldsymbol{\theta}_k)$$
$$\ln \mathcal{L}(\boldsymbol{\theta}|\mathcal{D}) = \sum_{i=1}^N \ln \sum_{k=1}^K \pi_k p(\mathbf{x}^{(i)}|\boldsymbol{\theta}_k)$$

- (i) Svaka komponenta mješavine $p(\mathbf{x}|\boldsymbol{\theta}_k)$ ima svoje parametre $\boldsymbol{\theta}_k$ (na primjer, u slučaju Gaussove gustoće vjerojatnosti, to su parametri $\boldsymbol{\mu}_k$ i $\boldsymbol{\Sigma}_k$). Uz njih, parametri su još i svi koeficijenti mješavine π_k .
- (ii) Ne, jer zbog neoznačenih primjera to više ne možemo činiti zasebno za svaku klasu (ne znamo koji primjer pripada kojoj komponenti). Nije nam dostupno rješenje u zatvorenoj formi, stoga parametre optimiramo iterativnim metodama.
- (iii) Da, mogli bismo primijeniti gradijentni spust. (*Todo izraz.*)
- (c) (*Todo izvod.*)
- (d) EM-algoritam degenerira u algoritam k-srednjih vrijednosti kada komponente mješavine imaju dijeljenu i izotropnu kovarijacijsku matricu $\boldsymbol{\Sigma} = \sigma^2 \mathbf{I}$ te kada se vrijednosti $h_k^{(i)}$ zaokružuju na 1 i 0. (*Todo pokaži.*)