Tik-61.231 Hahmontunnistuksen perusteet

Laskuharjoitus 3: 9.10.2000

1. Olkoon $x_1, x_2, \ldots, x_n$ riippumattomia ei-negatiivisia kokonaislukuja Poisson-jakaumasta jonka odotusarvo $E[x] = \lambda$. Tämä vastaa disreettiä jakaumaa $p(x|\lambda) = \lambda^x e^{-\lambda}/x!, x \geq 0$, kun $E[x] = \mathrm{Var}[x] = \lambda$. Etsi ML-estimaatti parametrille λ. Onko se harhaton?

2. Muotoile yksiulotteinen Parzen-estimaatti tiheysfunktiolle p_y annetuista näytteistä

$$y^{(i)} : 2.5, 2.8, 3.4, 4.2, 4.5, 4.7, 5.2, 5.6, 7.5$$

   Käytä suorakulmaista ikkunafunktiota.

3. Luokittele pisteet A, B ja C kuvaajassa käyttäen viiden lähimmän naapurin luokittelusääntöä ja Euklidista etäisyysfunktiota.

4. Kun käytetään lähimmän napurin luokittelusääntöä kahdelle luokalle (kuvaaja), saadaan päätösalueet ja raja näille luokille. Jos $\mathbf{y}^{(i)}$ ovat näytevektoreita luokista 1 ja $\mathbf{s}^{(i)}$ näytevektoreita luokasta 2, määritelmän perusteella mikä tahansa piste $\hat{\mathbf{y}}$ rajalla toteuttaa yhtälön

$$\min_i d(\hat{\mathbf{y}}, \mathbf{y}^{(i)}) = \min_i d(\hat{\mathbf{y}}, \mathbf{s}^{(i)})$$

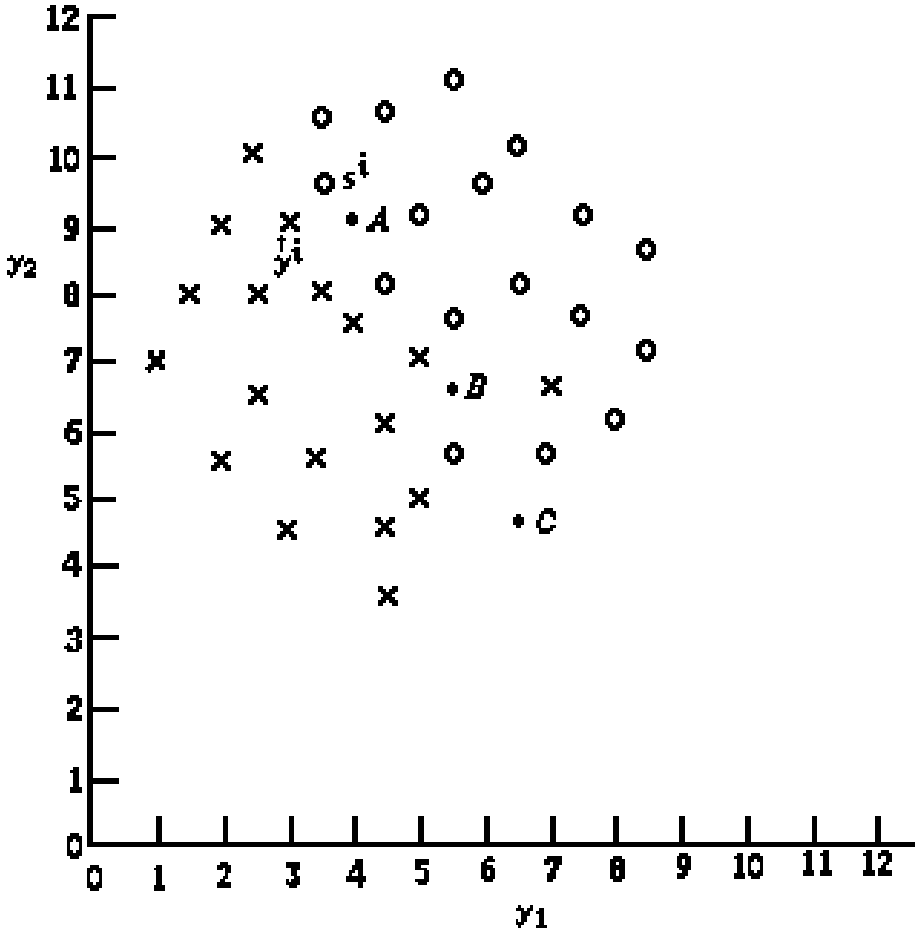
   Oletetaan etäisyysfunktion olevan Euklidinen.

   a) Mikä on päätösrajan yhtälö pienessä alueessa rajan lähellä, jos vakiovektori $\mathbf{y}^{(i)}$ on lähinnä rajaa luokasta 1 ja $\mathbf{s}^{(j)}$ on vakiovektori lähinnä rajaa luokasta 2?

   b) Piirrä täydellinen päätösraja luokille kuvaajassa.

   c) Etsi näytevektorit, joiden poistaminen ei muuta päätösrajaa.

5. Luokittimen virhettä pyritään usein estimoimaan testijoukolla (virheprosentti lasketaan testivektoreista). Joskus testijoukko on sama kuin millä luokitin rakennettiin. Osoite, että tämä ei ole hyvä tapa erityisesti jos kyseessä on 1-NN-luokitin.


12
11
10
9
8
7
6
5
4
3
2
1
0
0 1 2 3 4 5 6 7 8 9 10 11 12
y_2
y_1
s^i
y_1^i
A
B
C