COMPUTATIONAL SCIENCE AND TECHNIQUES
Volume 5, Number 1, 2017, 529-540

Publisher: Klaipėda University
http://journals.ku.lt/index.php/CST
Online ISSN: 2029-9966
DOI: 10.15181/csat.v5i1.1561

REKURENTINIS PASLĖPTŲJŲ MARKOVO MODELIŲ PARAMETRŲ VERTINIMO ALGORITMAS

Jūratė Vaičiulytė, Leonidas Sakalauskas
Vilniaus Universitetas, Matematikos ir informatikos institutas

Anotacija. Šiame darbe sudarytas rekurentinis paslėptųjų Markovo modelių parametrų vertinimo algoritmas. Paslėptieji Markovo modeliai modeliuojami Gauso skirstiniu, kurio parametrai pasiskirstę pagal daugiamatį normalųjį dėsnį su nežinomais vidurkių vektoriumi ir kovariacijų matrica. Nežinomų parametrų įverčiai gaunami didžiausio tikėtinumo metodu. Rekurentinis algoritmas sudarytas remiantis didžiausio tikėtinumo metodu išvestomis formulėmis ir klasikiniu EM algoritmu. Kadangi rekurentinio algoritmo vykdymo laikas yra proporcingas apdorojamų stebėjimų skaičiui, tai jis gali būti naudojamas modelio parametrų vertinimui realiu laiku. Realizuoto rekurentinio EM algoritmo savybės buvo ištirtos kompiuteriniu eksperimentu klasterizuojant duomenis. Jis taip pat gali būti taikomas duomenų klasifikavimo ir atpažinimo realiu laiku uždaviniams spręsti.

Pagrindiniai žodžiai: Paslėptieji Markovo Modeliai, tikėtinumo metodas, rekurentinis EM algoritmas

Įvadas

Gamtoje vykstantys realūs procesai yra charakterizuojami kaip signalai. Paslėptieji Markovo Modeliai yra dažnai taikomi stochastiniam signalų modeliavimui, kai charakterizuojamos tik statistinės signalų savybės (Rabiner, 1989). Šiame darbe realizuoti paslėptieji Markovo modeliai (PMM) su tolydžiais tikimybės tankiais, kai stebėjimai aprašomi Gauso skirstiniu. Realizuotas rekurentinis EM algoritmas modelio parametrų įvertinimui. Pavyzdžiui, pasinaudojant sukurtu algoritmu gali būti nagrinėjamas signalų klasifikavimas, klasterizavimas, atpažinimas realiu laiku.

1. Paslėptieji Markovo modeliai

Paslėptieji Markovo Modeliai – tai Markovo modeliai, kai stebėjimas yra atsitiktinė būsenos funkcija. Toks modelis (vadinamas paslėptu Markovo modeliu) yra tarsi dvigubas stochastinis procesas, kuriame paslėptas procesas gali būti stebimas tik per kitų stochastinių procesų sukurtą stebėjimų seką.

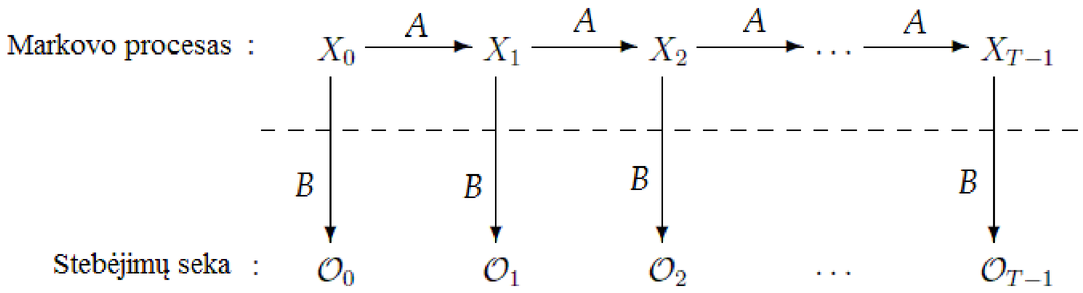
Markovo savybė: Paslėpto kintamojo $x(t)$ laiko momentu t sąlyginis tikimybinis skirstinys (angl. conditional probability distribution), kai duotos visais laiko momentais paslėpto kintamojo x reikšmės, priklauso tik nuo paslėpto kintamojo $x(t-1)$ reikšmės. Panašiai, stebimo kintamojo $O(t)$ reikšmė priklauso tik nuo paslėpto kintamojo $x(t)$ reikšmės.

Nagrinėjamame PMM paslėptų kintamųjų būsenų erdvė yra diskreti, stebėjimai yra tolydieji dydžiai (Gauso skirstinio). PMM parametrai: būsenų perėjimo tikimybės, išvesties tikimybės. Būsenų perėjimo tikimybės kontroliuoja, kaip paslėpta būsena laiko momentu t yra parenkama, kai duota paslėpta būsena laiko momentu $t - 1$.

Paslėptų būsenų aibė turi vieną iš N galimų reikšmių, kurios sumodeliuotos pagal kategorinį skirstinį. Tai reiškia, kad kiekvienai iš galimų N būsenų, kurioje laiko momentu t gali būti paslėptas kintamasis, egzistuoja perėjimo tikimybė iš tos būsenos į kiekvieną iš paslėpto kintamojo N galimų būsenų laiko momentu $t + 1$, kai i viso yra N^2 perėjimo tikimybių ($N \times N$ perėjimo tikimybių matrica). Kadangi bet kuri perėjimo tikimybė gali būti nustatyta, kai yra žinomos kitos, iš viso yra $N(N - 1)$ perėjimo parametrų.

Kiekvienai i N galimų būsenų yra priskiriama išvesties tikimybių aibė, kuri valdo stebimo kintamojo skirstinį tam tikru laiko momentu, kai duota paslėpto kintamojo būsena tuo laiko momentu. Šios aibės dydis priklauso nuo stebimo kintamojo. Pavyzdžiui, jei stebimas kintamasis yra M-matis vektorius, pasiskirstęs pagal daugiamatį Gauso skirstinį, tai bus M parametrų, valdančių vidurkius ir $M(M+1)/2$ parametrų kontroliuojančių kovariacijų matricą, kai iš viso yra $O(NM^2)$ išvesties parametrų.

Markovo procesas (*1 pav. paslėptas po brūkšniuota linija*) yra nustatomas pagal esamą būseną ir perėjimo tikimybių matricą A. Mes galime stebėti tik O_i, kuris yra susijęs su matricos B paslėptomis Markovo proceso būsenomis (Stamp, 2015). Rodyklės diagramoje - rodo sąlygines priklausomybes.



1 pav. **Paslėptųjų Markovo modelių schema**

Taigi, norint aprašyti PMM, užtenka nusakyti rinkinį $\lambda=(A,B,\pi)$ (Rabiner, 1989). Apmokant PMM pagal stebėjimų duomenis O įvertinami modelio parametrai $\lambda=(A,B)$, t.y. perėjimo iš būsenos į būseną tikimybių matrica ir stebėjimų tikimybės tankio funkcijos kiekvienai modelio būsenai. Ji yra išreiškiama stebėjimų vektorių tikimybės tankio funkcija, nusakančia konkretaus stebėjimų vektoriaus tikimybę. Tikimybinis skirstinys gali būti įvairių formų, pvz.: tolydus skirstinys, normalusis ir t.t.

Kuriamas PMM matematinis modelis yra aprašomas nustatytais parametrais:

- būsenų skaičius N,
- būsenų perėjimų tikimybių matrica,

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1N} \\ \vdots & \cdots & \vdots \\ a_{N1} & \cdots & a_{NN} \end{pmatrix} \quad (1)$$

- pradinio buvimo būsenoje tikimybių matrica:

$$\pi = \begin{pmatrix} \pi_1 \\ \vdots \\ \pi_N \end{pmatrix} \qquad (2)$$

- tikimybės tankio funkcija

$$N(o, \mu_s, \sigma_s) = \frac{1}{\sqrt{(2\pi)^n |\sigma_s|}} \exp(-\frac{1}{2}(o-\mu_s)^T \sigma_s^{-1}(o-\mu_s)) \qquad (3)$$

čia stebėjimus aprašantys normalieji atsitiktiniai dydžiai su vidurkiais μ_s ir kovariacijų matricomis σ_s, $1 \leq S \leq N$

$$\mu_S = \begin{pmatrix} \mu_1 & \mu_2 & \dots & \mu_M \end{pmatrix}, \qquad \sigma_S = \begin{pmatrix} \sigma_{11} & \cdots & \sigma_{1M} \\ \vdots & \cdots & \vdots \\ \sigma_{M1} & \cdots & \sigma_{MM} \end{pmatrix}, \qquad (4)$$

čia M - elementų skaičius

Būsenos parametrų skirstinys gali būti normalusis arba kelių Gauso dėsnių mišinys. Skirtingų būsenų X_i ir X_j parametrai μ ir σ yra skirtingi, kai $i \neq j$. Pastebėsime, kad prielaida, jog kiekvienos būsenos skirstinys yra normalusis, o ne kelių Gauso dydžių mišinys, nemažina bendrumo. Iš tikrųjų, jeigu kurioje nors PMM būsenoje yra stebimas Gauso skirstinio mišinys su svoriais c_1, c_2, ...c_l, tai tokią grandinę galima pakeisti jai ekvivalenčia, pakeitus minėtą mazgą l kitų mazgų. Jeigu grandinėje yra kelios būsenos, kuriose stebimi signalai yra su tais pačiais parametrais (μ ir σ), tuomet juos galima apjungti į vieną.

2. Klasifikavimas

Klasifikavimo metu kiekvienam gautam stebėjimui apskaičiuojama tikėtinumo funkcijos minimalios reikšmės argumentas ir pagal tai nustatoma, kuriai būsenai stebėjimas priklauso.

Sakykim, PMM sudarytas iš kelių būsenų su parametrais μ, σ, π. Tuomet logaritminė tikėtinumo funkciją $f(o, \mu, \sigma)$ galima užrašyti tokiu pavidalu:

$$f(o, \mu, \sigma) = \frac{(o-\mu)^T \cdot \sigma^{-1} \cdot (o-\mu)}{2} + \ln\left(\sqrt{|\sigma|}\right) \qquad (5)$$

Tuomet stebėjimas priskiriamas tai būsenai, kurioje apskaičiuota funkcijos reikšmė su atitinkamais parametrais yra: $arg\ min\ f(O, \mu_j, \sigma_j) - ln(p_j)$, kai $j \in [1;N]$ ir p_j - j-tosios būsenos tikimybė.

3. Kolmogorovo-Čepmeno lygtys

Vertinant parametrus reikia apskaičiuoti būsenos j tikimybę laiko momentu t, jei laiko momentu t-1 buvo būsenoje i. Tam pasinaudojama Kolmogorovo-Čepmeno lygtimis, buvimo būsenoje i tikimybę apskaičiuojant pagal formulę (Leveque, 2011):

$$\pi_t = A \cdot \pi_{t-1} \qquad (6)$$

4. PMM parametrų vertinimas: didžiausio tikėtinumo metodas

PMM parametrai gali būti įvertinami iteratyviu būdu. Modelio parametrų įvertinimui dažnai naudojamas didžiausio tikėtinumo metodas (Bishop, 2006; Vaseghi, 2000).

Didžiausio tikėtinumo metodą aptarsime tolydžiųjų atsitiktinų dydžių atveju. Tarkime, stebime atsitiktinį dydį *O*, kurio tankis $b(O)$ priklauso nuo nežinomų parametrų. Stebėjimo tikimybių tankis yra užrašomas tokiu pavidalu:

$$\sum_{j=1}^{N} \pi_j \cdot b_j(O) \tag{7}$$

čia π_j yra buvimo *j*-toje būsenoje tikimybė, o stebėjimo *j*-oje būsenoje tikimybė yra aprašoma daugiamačiu Gauso dėsniu:

$$b_j(O) = \frac{1}{\sqrt{(2\pi)^n \left|\sigma_j\right|}} e^{-\frac{1}{2}(o-\mu_j)^T \sigma_j{}^{-1}(o-\mu_j)} \tag{8}$$

Įveskime logaritminę tikėtinumo funkciją, aprašančią stebėjimą atskiroje būsenoje:

$$l(o,\mu,\sigma,\pi) = -\ln b(o) - \ln(\pi) = \frac{(o-\mu)^T \sigma^{-1}(o-\mu)}{2} + \ln\left(\frac{\sqrt{|\sigma|}}{\pi}\right) \tag{9}$$

į kurią įtraukta buvimo būsenoje tikimybė π.

Tegul

$$L(o,\mu,\sigma,\pi) = \sum_{i=1}^{N} e^{-l(o,\mu_j,\sigma_j,\pi)} \tag{10}$$

Apskaičiuojamos *ln L* išvestinės pagal μ ir σ:

$$\left(L_{\mu_i}\right)' = \frac{e^{-l(o,\mu_i,\sigma_i,\pi_i)} \cdot (o-\mu_i) \cdot \sigma_i^{-1}}{\sum_i e^{-l(o\ ,\mu_i,\sigma_i,\pi)}} \tag{11}$$

$$\left(L_{\sigma_i}\right)' = \frac{e^{-l(o,\mu_i,\sigma_i,\pi_i)} (\sigma^{-1}(o-\mu)(o-\mu)^T \sigma^{-1} - \sigma^{-1})}{\sum_i e^{-l(o,\mu_i,\sigma_i,\pi_i)}} \tag{12}$$

Rastos išvestinės prilyginamos nuliui $(\ln L_\mu)' = 0$, $(\ln L_\sigma)' = 0$ ir gautos lygtys sprendžiamos μ ir σ atžvilgiu.

5. EM algoritmas

Kurtame PMM modelyje vidurkių ir dispersijų įverčiams maksimizuoti panaudotas EM algoritmas. EM algoritmas atliekamas šiais žingsniais (Jordan, 1999; Cappe, 2009):

- E-žingsnis: logaritminės tikėtinumo funkcijos sąlyginio vidurkio apskaičiavimas

$$L(\theta^i) = E[\log L(X^i|\theta)] \tag{13}$$

- M-žingsnis: tikėtinumo funkcijos sąlyginio vidurkio maksimizavimas

$$\theta^{i+1} \to \max_{\theta} L(\theta^i) \tag{14}$$

Pasinaudojus didžiausio tikėtinumo metodu išvestos formulės parametrų įvertinimui per sumas:

$$\bar{\mu}_j = \frac{\sum_{j=1}^{T} \gamma(j) \cdot O}{\sum_{j=1}^{T} \gamma(j)} \tag{15}$$

$$\bar{\sigma}_j = \frac{\sum_{j=1}^{T} \gamma(j) \cdot (O - \mu_j)(O - \mu_j)^T}{\sum_{j=1}^{T} \gamma(j)} \tag{16}$$

$$\gamma_j = e^{-l(o, \mu_i, \sigma_i, \pi_i)} \tag{17}$$

$a_{i,j}$ pakartotinio įvertinimo formulė:

$$\bar{a}_{ij} = \frac{skaicius\ perėjimu\ is\ busenos\ S_i\ i\ busena\ S_j}{skaicius\ perejimu\ is\ busenos\ S_i} \tag{18}$$

Bazinio EM algoritmo pseudokodas pateiktas ¿emiau.

Algoritmas Nr. 1: *Bazinis EM algoritmas: modelio parametrų vertinimas pagal turimus stebėjimo duomenis, kai skaičiavimai atliekami per sumos formules (15-18)*

1. *function Bazinis_EM_algoritmas (O, x, d, p, T, N, dim):*

   Įvestis: *O - stebėjimas, x - pradiniai vidurkių vektoriai, d - pradinės koviariacijų matricos, pi - pradinės buvimo būsenose tikimybės, p - būsenų perėjimo tikimybės, T - stebėjimų skaičius, N - būsenų skaičius, dim - vektorių dimensijų skaičius.*

   Išvestis: *parametrų reikšmės: tempM - vidurkių vektoriai, tempD - kovariacijų matricos, tempB - buvimo būsenoje tikimybės*

```
2. while(|x[iu]-x[iu-1]|>0.0001 and |d[iu]-d[iu-1]|>0.0001) do
      3. for i=1 to T
            4. pi[i+1] = p^T * pi[i]
      5. end for
         6. for i=1 to T
               7. sum_t_f = 0
               8. for j=1 to N
                     9. s[i][j] = exp(-l(O[i],x[j],d[j],p[i][j]))
```

```
10. sum_t_f[i] = sum_t_f[i] + s[i][j]
11. end for
12. for j=1 to N
13. s[i][j] = s[i][j] / sum_t_f[i]
14. end for
15. end for
16. for j=1 to N
17. v[j] = 0, vt = 0, vr = 0
18. for iu=1 to T
19. v[j] = v[j] + s[iu][j]
20. vt = vt + s[iu][j]*O[i]
21. vr = vr + (s[iu][j]*(O[i]-x[j])(O[i]-x[j])^Transp)
22. end for
23. x[j] = (1/v[j]) * vt
24. d[j] = (1/v[j]) * vr
25. end for
26. arg min l(O, x, d, pi)
26. end while
```

**Transp – matricos transponavimas*

6. Rekurentinis EM algoritmas

Sudaryto bazinio EM algoritmo sudėtingumas yra antros eilės. Iš tikrųjų, nesunku matyti, kad esant fiksuotai imčiai didžiausio tikėtinumo įverčiams gauti reikės atlikti skaičiavimus, kurių apimtis yra proporcinga fiksuotam imties tūriui. Tokiu būdu, jeigu vertinimas atliekamas su kiekvienu stebėjimu, bendras reikalingų operacijų skaičius bus antros eilės. Stebint procesą reikalingų vertinimui operacijų skaičius labai augs ir vertinimas realiu laiku taps nebeįmanomas. Šiai problemai spręsti buvo išvestos rekurentinės formulės PMM parametrų vertinimui. Šiuo atveju PMM modelio parametrai yra atnaujinami su kiekvienu gautu nauju stebėjimu, neįsimenant ankstesnės mokymo aibės. Kiekvienoje rekurentinio EM algoritmo iteracijoje atliekamas baigtinis operacijų skaičius, priklausantis polinomiškai (antro laipsnio) nuo modelio parametrų skaiciaus, bet nepriklausantis nuo atliktų iteracijų skaičiaus. Tokiu atveju, sukurtojo algoritmo sudėtintgumas priklauso netiesiškai nuo atliktų iteracijų skaičiaus, t.y., $O(T)$. Klasikinis EM algoritmas reikalauja kiekvienoje iteracijoje apdoroti visą turimą stebėjimų imtį, todėl apdorojant visą imtį kiekvienoje iteracijoje algoritmo sudėtingumas tampa kvadratinis, t.y. $O(T^2)$.

Pasinaudojus gautomis parametrų vertinimo formulėmis per sumas (15-16), išvestos rekurentinės formulės vidurkių vektoriaus ir kovariacijų matricos įvertinimui:

$$\overline{\mu_t^i} = \mu_{t-1}^i + \frac{\left(O_t - \mu_{t-1}^i\right)}{t} \cdot \omega_t^i \tag{19}$$

$$\overline{\sigma_t^i} = \left(\frac{\beta_{t-1}^i \cdot (t-1)}{\beta_t^i \cdot t} \right) \cdot \left(\sigma_{t-1}^i + \frac{(O_t - \mu_{t-1}^i)(O_t - \mu_{t-1}^i)^T}{t} \cdot \omega_t^i \right) \quad (20)$$

Taip pat išvestos rekurentinės formulės būsenų perėjimo tikimybių skaičiavimui:

$$\beta_t^i = \beta_{t-1}^i + \frac{1}{t} \left(\frac{e^{-l(o_t, \mu_t^i, \sigma_t^i, \pi_i)}}{\gamma_t} - \beta_{t-1}^i \right) \quad (21)$$

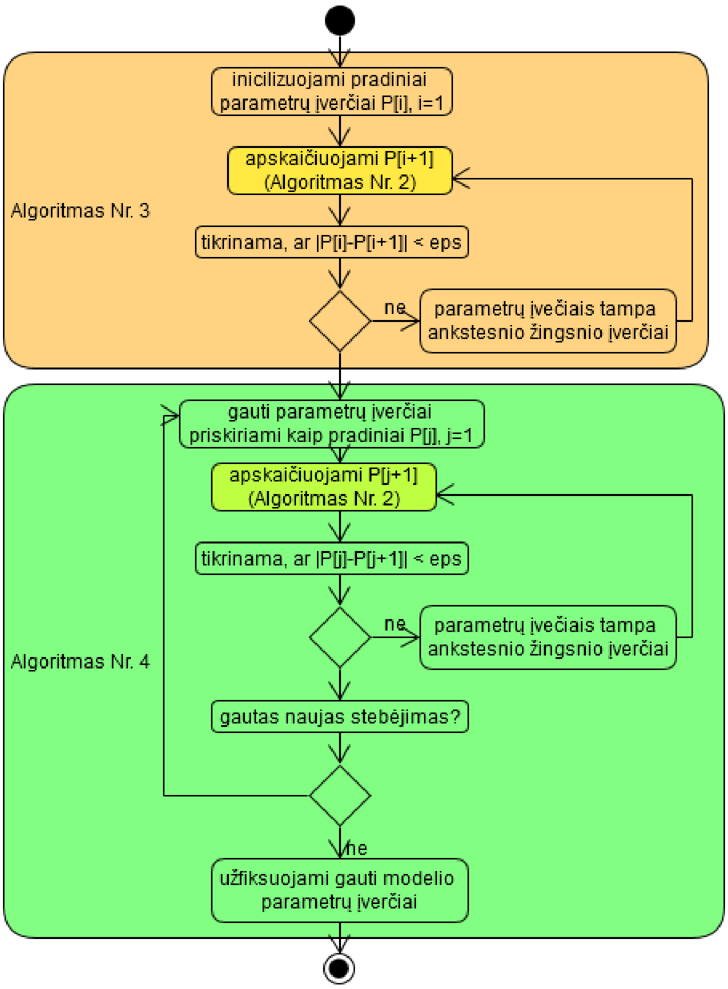
$$\beta_{t-1}^i = \frac{1}{t-1} \cdot \sum_{i=1}^{t} \frac{e^{-l(o_t, \mu_1^i, \sigma_1^i, \pi_1)}}{\gamma_t} \quad (22)$$

čia koeficientas ω:

$$\omega_t^i = \frac{e^{-l(o_t, \mu_t^i, \sigma_t^i, \pi_i)}}{\gamma_t \cdot \beta_t} \quad (23)$$

$$\gamma_t = \sum_{i=1}^{N} e^{-l(o_t, \mu_t^i, \sigma_t^i, \pi_i)} \quad (24)$$

Rekurentinis EM algoritmas susideda iš dviejų pagrindinių dalių: pirmiausia vykdomas pradinis parametrų vertinimas, kai žinomi tik stebėjimų vektoriai, taikant rekurentinį EM algoritmą ir pasirinkus nedidelę fiksuoto dydžio stebėjimo imtį (15-18 formulės). Gavus pradinius įverčius toliau atpažinimas ir parametrų vertinimas atliekamas rekurentiniu būdu, stebint procesą realiu laiku (žr. 2 pav.). Pradiniai įverčiai rekurentiniam algoritmui yra reikalingi norint užtikrinti jo stabilumą ir išvengti konvergavimo į išsigimusius lokaliuosius tikėtinumo funkcijos ekstremumus.



2 pav. **Bendra rekurentinio EM algoritmo schema**

Žemiau pateiktas rekurentinio algoritmo pseudokodas.

Algoritmas Nr. 2: *Sekančio žingsnio parametrų reikšmių apskaičiavimas (19-24 formulės)*

1. function Parametrų-skaičiavimas (O, i, s, M, D, B):

Įvestis: *O - signalas, i - stebėjimo eilė, s - tikėtinumo f-jos rezultatas, M - vidurkių vektoriai, D - kovariacijų matricos, B - buvimo būsenoje tikimybės*
Išvestis: *sekančio žingsnio parametrų reikšmės: tempM - vidurkių vektoriai, tempD - kovariacijų matricos, tempB - buvimo būsenoje tikimybės*

2. tfsum = 0
3. ***for*** *j=1* ***to*** *N*
4. tfsum=tfsum+s[j]
5. ***end for***
6. ***for*** *j=1* ***to*** *N*

```
        7. sk[j]=s[j]/tfsum
    8. end for
    9. for j=1 to N
        10. tempB[j] = B[j] + (1/(i+1)) * (sk[j]-B[j])
        11. tempM[j] = M[j] + ((O-M[j])/(i+1)*(sk[j]/tempB[j])
        13. tempD[j] = ((B[j]*i)/(tempB[j]*(i+1)))*
                (D[j]+(O-M[j])(O-M[j])^Transp)*(sk[j]/tempB[j]*(i+1)))
    14. end for
    15. return (tempB, tempM, tempD)
```

Algoritmas Nr. 3: *Rekurentinis EM algoritmas : PMM parametrų vertinimas su fiksuota stebėjimų imtimi, kai gaunami pradiniai parametrų įverčiai*

1. *function Rekurentinis_EM (O, x, d, p, T, N, dim):*

Įvestis: *O – stebėjimų masyvas, x - pradiniai vidurkių vektoriai, d - pradinės koviariacijų matricos, p - būsenų perėjimo tikimybės, T - stebėjimų skaičius, N - būsenų skaičius, dim - vektorių dimensijų skaičius.*

Išvestis: *parametrų reikšmės: tempM - vidurkių vektoriai, tempD - kovariacijų matricos, tempB - buvimo būsenoje tikimybės*

```
  3. for j=1 to N
      4. tempB[j][0]=exp(-f(O[0],x[j],d[j],p[0][j]))
  5.  sumB = sumB + tempB[j][0]
6. end for
7. for j=1 to N
        8.  x[j][0] = O[0]
        9.  d[j][0] = empty matrix
        10. b[j][0] = tempB[j][0] / sumB
11. end for

12. iu = 1
13. while(|x[iu]-x[iu-1]|>0.0001 and |d[iu]-d[iu-1]|>0.0001) do
        14. for i=1 to T
                15. for j=1 to N
                16. s[i][j] = exp(-l(O[i],x[iu-1][j],d[iu-1][j],p[i][j]))
                17. end for
                18. x[i],d[i],p[i],tempM,tempD,tempB <-
                Parametrų-skaičiavimas(O,i,s[i],x[i-1],d[i-1],p[i-1])
        19. end for
        20. iu = iu + 1
21. end while
22. return (tempB, tempM, tempD)
```

Algoritmas Nr. 4: *Rekurentinis EM algoritmas: atpažinimas ir parametrų įvertinimas atliekamas rekurentiniu būdu, stebint procesą realiu laiku, kai pradiniai parametrų įverčiai gauti Algoritmu Nr. 3*

1. *function Rekurentinis_EM_atpazinimas (O, x, d, p, T, N, dim):*

   Įvestis: *O - stebėjimas, x - pradiniai vidurkių vektoriai, d - pradinės koviariacijų matricos, p - būsenų perėjimo tikimybės, T - stebėjimų skaičius, N - būsenų skaičius, dim - vektorių dimensijų skaičius.*

   Išvestis: *parametrų reikšmės: tempM - vidurkių vektoriai, tempD - kovariacijų matricos, tempB - buvimo būsenoje tikimybės*

   x, d, p <- Rekurentinis_EM (O, x, d, p, T, N, dim)

2. ***for*** *i=1* ***to*** *T*
   3. *iu = 1*
   4. ***while***(|*s[iu]-s[iu-1]*|*>0.0001)* ***do***
      5. ***for*** *j=1* ***to*** *N*
         6. *s[i][j] = exp(-l(O[i],x[iu-1][j],d[iu-1][j],p[i][j]))*
      7. ***end for***
      8. *tempM,tempD,tempB <- Parametrų-skaičiavimas(O,i,s[i],x[i-1],d[i-1],p[i-1])*
      9. *iu = iu + 1*
   10. ***end while***
11. ***end for***
12. ***return*** *(tempB, tempM, tempD)*

7. Eksperimento eiga. PMM parametrų vertinimas

Sudarytam algoritmui patikrinti buvo atliktas kompiuterinis eksperimentas: stebėjimo vektorių klasterizavimas. Duoti dvimačiai Gauso stebėjimų vektoriai ($T=8192$ ir $T=2048$), kurie sudaro du klasterius ($k=2$). Šių klasterių tikrieji centrai: $\begin{pmatrix}500\\500\end{pmatrix}$ ir $\begin{pmatrix}600\\600\end{pmatrix}$.

Parenkami atsitiktiniai pradiniai duomenys:

– vidurkių vektoriai μ : $\begin{pmatrix}483\\444\end{pmatrix}$, $\begin{pmatrix}616\\589\end{pmatrix}$

– kovariacijų matricos σ :

$$\begin{pmatrix}11520 & 2643\\2643 & 11530\end{pmatrix}, \quad \begin{pmatrix}8159 & 2458\\2458 & 7987\end{pmatrix}$$

– perėjimų iš būsenos į būseną tikimybių matrica A, kur visos tikimybės yra vienodos: $\begin{pmatrix}0.5 & 0.5\\0.5 & 0.5\end{pmatrix}$

- pradinio buvimo būsenoje tikimybių matrica π, kur visos tikimybės yra vienodos: $\begin{pmatrix}0.5\\0.5\end{pmatrix}$.

PMM modelio parametrų vertinimas buvo nutraukiamas, kai skirtumas tarp dviejų gretimų modelio parametrų įverčių tapdavo mažesnis už pasirinktą reikšmę $\varepsilon=0.0001$.

Buvo stebimas algoritmo vykdymo laikas. Rekurentinis EM algoritmas parametrų vertinimą atliko keturis kartus greičiau nei bazinis EM algoritmas. Gautos parametrų reikšmės ir algoritmų vykdymo laikas pateikti 2 lentelėje.

2 lentelė. Eksperimento rezultatai

	Bazinis EM algoritmas	Rekurentinis EM algoritmas
T=2048	μ : $\begin{pmatrix}499.02\\497.45\end{pmatrix}$, $\begin{pmatrix}598.47\\599.59\end{pmatrix}$ σ: $\begin{pmatrix}1638 & -143.12\\-143.12 & 1615\end{pmatrix}$, $\begin{pmatrix}1722 & 62.87\\62.87 & 1531\end{pmatrix}$ π: $\begin{pmatrix}0.5\\0.5\end{pmatrix}$ **Vykdymo laikas:** 8.127 s	μ : $\begin{pmatrix}499.01\\497.44\end{pmatrix}$, $\begin{pmatrix}598.46\\599.57\end{pmatrix}$ σ: $\begin{pmatrix}1638 & -143.48\\-143.48 & 1615\end{pmatrix}$, $\begin{pmatrix}1723 & 63.64\\63.64 & 1532\end{pmatrix}$ π: $\begin{pmatrix}0.49\\0.50\end{pmatrix}$ **Vykdymo laikas:** 2.075 s
T=8192	μ : $\begin{pmatrix}502.35\\504.36\end{pmatrix}$, $\begin{pmatrix}595.14\\592.67\end{pmatrix}$ σ: $\begin{pmatrix}7785 & 442.76\\442.76 & 8298\end{pmatrix}$, $\begin{pmatrix}7591 & 562.39\\562.39 & 7323\end{pmatrix}$ π: $\begin{pmatrix}0.5\\0.5\end{pmatrix}$ **Vykdymo laikas:** 192.879 s	μ : $\begin{pmatrix}502.34\\504.38\end{pmatrix}$, $\begin{pmatrix}595.15\\592.65\end{pmatrix}$ σ: $\begin{pmatrix}7784 & 443.52\\443.52 & 8301\end{pmatrix}$, $\begin{pmatrix}7590 & 563.19\\563.19 & 7325\end{pmatrix}$ π: $\begin{pmatrix}0.5\\0.5\end{pmatrix}$ **Vykdymo laikas:** 26.476 s

Išvados

Darbe sudarytas paslėptųjų Markovo modelių parametrų vertinimo algoritmas. Atlikus skaičiavimus, gaunami nežinomų parametrų įverčiai. Klasikinio EM algoritmo sudėtingumas yra antros eilės ir jis netinka parametrų vertinimui realiu laiku. Todėl buvo išvestos rekurentinės formulės PMM parametrų vertinimui, kai jie yra atnaujinami su kiekvienu gautu nauju stebėjimu, neįsimenant ankstesnės mokymo aibės. Iš atlikto kompiuterinio eksperimento matyti, kad modelio parametrų vertinimas realizuotu rekurentiniu EM algoritmu vykdomas kelis kartus greičiau nei klasikinis EM algoritmas. Sudarytas rekurentinis algoritmas gali būti taikomas stebėjimų klasterizavimui ir atpažinimui realiu laiku.

Literatūra

Bishop, C. (2006). *Pattern recognition and machine learning.* New York [u.a.]: Springer.

Cappe, O., Moulines, E. and Ryden, T. (2009). *Inference in Hidden Markov Models.* [ebook] Available at: https://www.ime.usp.br/ebp/ebp13/mainbras.pdf [Accessed 12 Sep. 2017].

Leveque, O. (2011). *Lecture notes on Markov chains*. [ebook] Available at: http://www.hamilton.ie/ollie/Downloads/Mar1.pdf [Accessed 12 Sep. 2017].
Jordan, M. (1999). *Learning in graphical models*. Cambridge, Mass.: MIT Press. 355 – 368.
Rabiner, L. (1989). A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the IEEE*, 77(2), pp.257-286.
Stamp, M. (2015). *A Revealing Introduction to Hidden Markov Models*. [ebook] Available at: http://www.cs.sjsu.edu/faculty/stamp/RUA/HMM.pdf [Accessed 12 Sep. 2017].
Vaseghi, S. (2000). *Advanced digital signal processing and noise reduction*. Chichester, U.K.: J. Wiley & Sons.

RECURRENT ESTIMATION OF HIDDEN MARKOV MODEL PARAMETERS

Jūratė Vaičiulytė, Leonidas Sakalauskas

Summary

This work contains recurrent algorithm for Hidden Markov Model parameter estimation. We are modelling Hidden Markov Models with Gaussian distribution, which parameters are distributed by multivariate normal with unknown parameters – mean vector and covariance matrix. The estimation of unknown parameters are calculated by the maximum likelihood parameters estimation. Formulas for recurrent parameter estimation by EM algorithm was derived.

Key words: Hidden Markov Models, likelihood method, recurrent EM algorithm