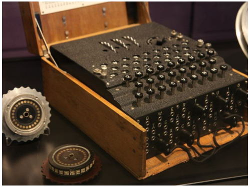
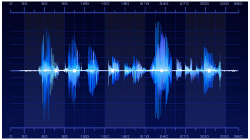
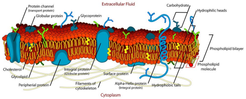



0 30 60 90 120 150 180 210 240 270 300 330 360


Extracellular Fluid
Protein channel
(transport protein)
Globular protein
Glycoprotein
Carbohydrate
Hydrophilic heads
Phospholipid bilayer
Phospholipid
molecule
Cholesterol
Glycolipid
Peripherial protein
Integral protein
(Globular protein)
Surface protein
Filaments of
cytoskeleton
Alpha-Helix protein
(Integral protein)
Hydrophobic tails
Cytoplasm

Skriti markovski modeli v finanÄŤnih ÄŤasovnih vrstah

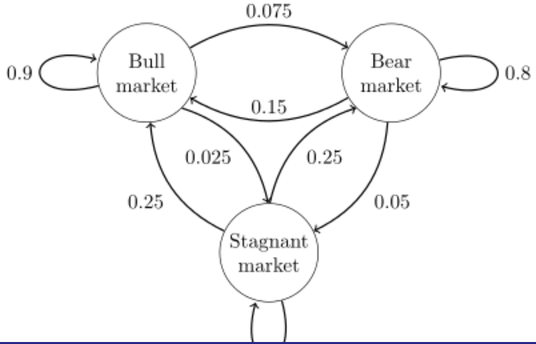
Fakulteta za matematiko in fiziko, Univerza v Ljubljani

14. april 2019

Martin PraÄŤek
Mentor: izr. prof. dr. Damjan Škulj

Markovski model

- Markovska lastnost je lastnost sluÄŤajnega procesa v diksretnem ÄŤasu, da je njegova vrednost v ÄŤasu t odvisna le od njegove vrednosti v ÄŤasu $t-1$.
- LoÄŤimo v celoti opazovan in delno opazovan ter
- Avtonomen in kotroliran sistem.



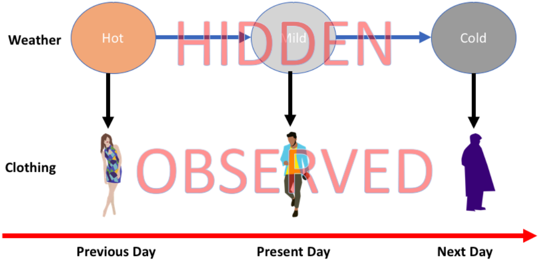
Delitev markovskih modelov

	V celoti opazovan	Le delno opazovan
Avtonomen	Markovska veriga	Skriti markovski model
Kontorliran	Markovski proces odločanja	Delno opazovalen proces odločanja

Skriti markovski model

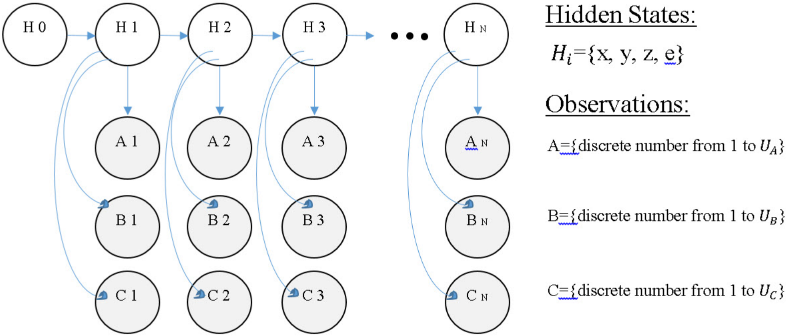
Skriti markovski model je statistiÄŤni markovski model, kjer predpostavljamo, da je modelirani sistem markovski proces z skritimi stanji.

Gre torej za tip modela, kjer lahko razberemo rezultat, ne moremo pa ugotoviti, kakĹˇna je bila funkcija, ki nam ga je dala.



Zahteve

- Markovska lastnost
- Enakomerno porazdeljeni ÄŤasi signalov O_t, ki jih poda resniÄŤni svet
- Sistem ima N stanj, vsako doloÄŤa sluÄŤajna spremenjivka S



- SluÄŤajnih spremenljivk skritih skoraj v nobenem ÄŤasu ne poznamo, poznamo pa sluÄŤajni proces Q, ki predstavlja signale
- Porazdelitveni zakon vsakega stanja i oznaÄŤimo z $b_i(x)$
- Vektor zaÄŤetnih stanj je π
- Prehodna matrika A, ki je neodvisna od ÄŤasa

Porazdelitveni zakon

- Gaussova meĹˇanica
- $b_i = \sum_{j=1}^{M} c_{ij} N(x; \mu_j, \sigma_j^2)$
- Ĺ tevilo porazdelitev M
- Matrika Γ, μ_{ij} predstavlja priÄŤakovano vrednost porazdelitve j v stanju i
- Matrika Σ, kjer σ_{ij} predstavlja varianco porazdelitve j v stanju i
- Matrika C, koeficienti c_{ij} iz Gaussove meĹˇanice

Osnovanje modela

Da bomo lahko delali z naĹˇim modelom, moramo najprej izvesti t.i. trening modela.

- RazvrĹˇÄŤanje v skupine (*k*-means clustering)
- Akaikov informacijski kriterij

IzraÄŤun $P(O|\lambda)$

Z λ oznaÄŤimo vse ostale parametre $\lambda = (\pi, A, C, \Gamma, \Sigma)$. Za uÄŤinkovit izraÄŤun le teh si pomagamo z iterativnim postopkom, podobnim sistemu EM, *naprej* – *nazaj*.

$$\alpha_1(i) = \pi_i b_i(O_1) = \pi_i b_i = \sum_{k=1}^{M} c_{ik}$$

$$\Downarrow$$

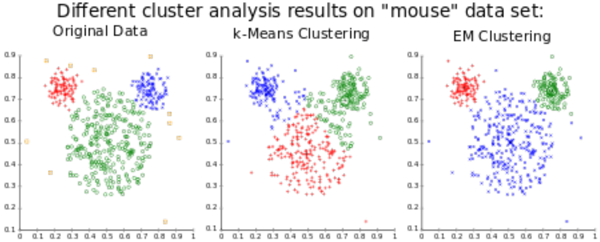
$$P(O|\lambda) = \sum_{i=1}^{N} \alpha_T(i)$$

Trening modela

Da bomo lahko $P(O|\lambda)$ maksimizirali, potrebujemo zaÄŤetne ocene parametrov. Za to ne poznamo analitiÄŤnega postopka, lahko pa lokalno maksimiziramo, z naprimer, Baum-Welchovim algoritmom. Pri tem nas ne skrbijo ocene za A ter π, kjer moramo paziti le na neniÄŤelnost le teh. VeÄŤ problemov nam povzroÄŤajo C, Σ in Γ.

C, Σ in Γ

Za dobro zaÄŤetno oceno si lahko pomagamo z razvrĹˇÄŤanjem v skupine. Za C to pomeni, da bo element $c_{ij} = 1/M$ za vsak par ij, kjer je M Ĺˇtevilo skupin. PriÄŤakovane vrednosti in variance nato pridobimo iz vrednosti teh skupin.



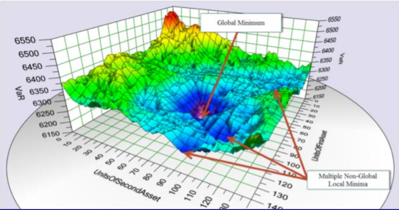
Lokalno maksimiziranje

Za potrebe lokalne maksimizacije doloÄŤimo dodatne funkcije:

$$\xi_t(i,j) = \frac{\alpha_t(i) a_{ij} b_j(O_t) \beta_{t+1}(j)}{P(O|\lambda)}$$

$$\gamma_t(i) = \frac{\alpha_t(i)\beta_t(i)}{P(O|\lambda)}$$

$$\gamma_t(j,k) = \gamma_t(j) \frac{c_{jk} N(x; \mu_{jk}, \sigma^2_{jk})}{\sum_{m=1}^{M} c_{jm} N(x; \mu_{jm}, \sigma^2_{jm})}$$



Lokalno maksimiziranje

Prek dodatnih funkcij definiramo iteracijske postopke za naĹˇe spremenljivke:

$$\overline{\pi_i} = \gamma_1(i)$$

$$\overline{a_{ij}} = \frac{\sum_{t=1}^{T-1} \xi_t}{\sum_{t=1}^{T-1} \gamma_t(i)}$$

$$\overline{c_{jk}} = \frac{\sum_{t=1}^{T} \gamma_t(j,k)}{\sum_{t=1}^{T} \sum_{m=1}^{M} \gamma_t(j,m)}$$

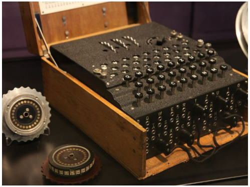
$$\overline{\mu_{jk}} = \frac{\sum_{t=1}^{T} \gamma_t(j,k) O_t}{\sum_{t=1}^{T} \gamma_t(j,k)}$$

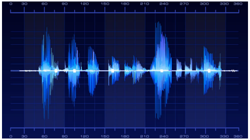
Tako definiramo veĂŤfazni iterativni proces, pri katerem popravljamo vrednosti parametrov do konvergence.

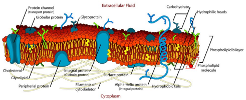
ZaÄŤetno stanje in Viterbijev algoritem

Za delo s tem algoritmom moramo definirati $\delta_t(i)$, ki za vsako stanje i vrne najveÄŤjo verjetnost vzdolĹľ poti v ÄŤasu t. Prek $\delta_t(i)$ nato induktivno izvedemo algoritem.
Viterbijev algoritem nam vrne $p*$, ki je najveÄŤja verjetnost in q_T*, ki nam pove stanje v ÄŤasu T, ki nam to verjetnost vrne.

Uporaba




0 30 60 90 120 150 180 210 240 270 300 330 360
0 30 60 90 120 150 180 210 240 270 300 330 360


Extracellular Fluid
Protein channel
(transport protein)
Globular protein
Glycoprotein
Carbohydrate
Hydrophilic heads
Phospholipid bilayer
Phospholipid
molecule
Cholesterol
Glycolipid
Peripherial protein
Integral protein
(Globular protein)
Surface protein
Filaments of
cytoskeleton
Alpha-Helix protein
(Integral protein)
Hydrophobic tails
Cytoplasm

V dolgi predstavitvi se bom bolj posvetil sami finanÄŤni analizi, ki sem jo tokrat zaenkrat pustil pri miru. V prihodnje bom tudi sam poizkusil doloÄŤiti skriti markovski model na svojem setu podatkov.

Viri

1. I. MacDonald, W. Zucchini. Hidden Markov and Other Models for Discrete-valued Time Series. Chapman & Hall/CRC Monographs on Statistics & Applied Probability. Taylor & Francis. (1997)
2. D.Roman, G. Mitra,N. Spagnolo. Hidden Markov models for financial optimization problems.IMA Journal of Management Mathematics. 21 (2010).