

UNIVERZA V LJUBLJANI
FAKULTETA ZA MATEMATIKO IN FIZIKO

Finančna matematika – 1. stopnja

Martin Praček

Skriti markovski modeli v časovnih vrstah

Delo diplomskega seminarja

Mentor: prof. dr./doc. dr. Damjan Škulj

Ljubljana, 2019

KAZALO

1. Uvod	4
Slovar strokovnih izrazov	4
2. Uporaba	4
2.1. Procesiranje govora	4
2.2. Uporaba v biologiji in biokemiji	4
Literatura	4

Skriti markovski modeli v časovnih vrstah

POVZETEK

V povzetku na kratko opiši vsebinske rezultate dela. Sem ne sodi razlaga organizacije dela – v katerem poglavju/razdelku je kaj, pač pa le opis vsebine.

Hidden Markov Models in Time Series

ABSTRACT

Prevod zgornjega povzetka v angleščino.

Math. Subj. Class. (2010): navedi vsaj eno klasifikacijsko oznako – dostopne so na www.ams.org/mathscinet/msc/msc2010.html

Ključne besede: skriti markovski modeli, časovne vrste, slučajni proces navedi nekaj ključnih pojmov, ki nastopajo v delu

Keywords: hidden markov models, time series, angleški prevod ključnih besed

1. UVOD

SLOVAR STROKOVNIH IZRAZOV

2. UPORABA

Skriti markovski modeli so zelo široko uporabno matematično orodje za modeliranje. Načini uporabe se zelo razlikujejo in gredo od zelo bioloških do ekonomskih. Ekonomskim, torej tistim, ki jih delujejo kot napovedovalci cen vrednostnih papirjev v prihodnosti, se bom najgloblje posvetil v tem delu, zato sedaj raje pogledimo ostale načine uporabe.

2.1. Procesiranje govora. Ena najbolj široko uporabljenih načinov uporabe pa je procesiranje govora za posamične glasovne enote. Gre za sistem, kjer želimo prepoznati posamične izgovorjene besede, ne moremo pa ga uporabiti za splošen govor.

Da bomo procesiranje lahko zmodelirali moramo najprej določiti slovar glasov, iz katerega vemo, da bo glas prišel. Tu uporabljam besedo glas, ker ni nujno, da bo prepoznan glas dejansko beseda; eden izmed glasov bi lahko bil tudi zgolj črka A. Ko je slovar določen, bo slučajni proces St , kjer t označuje realizacijo v končnem času t , določen z dvema mehanizmoma:

- (1) homogena Markova veriga Ct , $t \in 1, \dots, m$, ki določa položaj glasilk v vsakem času,
- (2) ter verjetnostne porazdelitve, po eno za vsako stanje.

Te pogoje natančno določimo s parametri δ , ki predstavlja $C1$, Π , kjer je to matrika velikosti $n \times m$ in pa prehodno matriko Γ za markovsko verigo.

Ko imamo vse to določeno s pomočjo Baum-Welchovega algoritma določimo rezultat.

2.2. Uporaba v biologiji in biokemiji. V zadnjem času se skrite markovske modele vedno več uporablja tudi za modeliranje bioloških in biokemijskih procesov, med katerimi je najbolj znan primer modeliranja proteinov v celični membrani, ki ga bom malo opisal, poleg tega pa med drugim še za **napovedovanje genov?**

Čeprav lahko nekatere proteine in njihovo delovanje znotraj celične membrane enostavno predstavimo, to ne velja za vse. Tako imenovani β -**barrel** membranski proteini zahtevajo več dela. Da bi ugotovili njihovo delovanje in ga primerjali z delovanjem v vodi topnih proteinov so Bagos, Liakoupos et al. razvili model ki to naredi. Gre za model, ki temelji na skritih markovskih modelih, njegov trening pa poteka na diskretni bazi podatkov in ni namenjen maksimiziranju **likelihood**, temveč je namenjen maksimiziranju verjetnosti pravih napovedi.

Z delom potem nadaljujemo podobno kot pri procesiranju govora, le da je naš začetni slovar tu dolg 20 znakov, toliko kolikor je aminokislin.

LITERATURA