Kolmekaupa Markovi ahelate Viterbi raja lähendamine variatsiooniliste meetoditega

Jaan Erik Pihel

1. Kolmekaupa Markovi ahelate ja Viterbi raja probleem
2. Variatsioonilised Bayesi meetodid
3. Implementatsioon
4. Tulemused

1. **Kolmekaupa Markovi ahelate ja Viterbi raja probleem**
2. Variatsioonilised Bayesi meetodid
3. Implementatsioon
4. Tulemused

Kolmekaupa Markovi mudel

on mudel, millel on Markovi omadus

$$p(z_t|\mathbf{z}_{1:t-1}) = p(z_t|z_{-1})$$

ning mille komponente on võimalik kirjutada

$$z_t = (u_t, x_t, y_t),\ \ \mathbf{z} = (u_t, x_t, z_t)_{t=1}^T \in (\mathcal{U} \times \mathcal{X} \times \mathcal{Y})^T,$$

kusjuures fikseeritud vaatluste $\mathbf{y} = (y_t)_{t=1}^T$ puhul saame kirjutada

$$p(\mathbf{u}, \mathbf{x}) = \pi(u_1, x_1) \prod_{t=2}^{T} p_t(u_t, x_t|u_{t-1}, x_{t-1}) \,.$$

Viterbi rada

Diskreetsel mittehomogeensel paarikaupa Markovi mudelil

$$p(\mathbf{u}, \mathbf{x}) = \pi(u_1, x_1) \prod_{t=2}^{T} p_t(u_t, x_t | u_{t-1}, x_{t-1})$$

defineerime **Viterbi raja** kui $\mathbf{x}^* \in \{\mathbf{x} | \forall \mathbf{x}' \; p(\mathbf{x}) \geq p(\mathbf{x}')\}$,

kus $p(\mathbf{x}) = \sum_{\mathbf{u}} p(\mathbf{x}, \mathbf{u})$.

Viterbi raja leidmine on NP raske, seega on pikkade ahelate korral arvutuslikult võimatu. Viterbi raja lähendamiseks on kasutatud erinevaid meetodeid.

1. Kolmekaupa Markovi ahelate ja Viterbi raja probleem
2. **Variatsioonilised Bayesi meetodid**
3. Implementatsioon
4. Tulemused

Variatsioonilised Bayesi meetodid

Leiame *lihtsama* mõõdu *q*, mis oleks esialgsele mittehomogeensele paarikaupa

Markovi ahelale sarnane Kullback-Leibleri kauguse mõttes

$$\min_{q \in Q} D_{\mathrm{KL}}\left[q \| p\right],$$

kus

$$D_{\mathrm{KL}}\left[q \| p\right] = \sum_{\mathbf{u},\mathbf{x}} q(\mathbf{u},\mathbf{x}) \ln \frac{q(\mathbf{u},\mathbf{x})}{p(\mathbf{u},\mathbf{x})}.$$

.

Variatsioonilised Bayesi meetodid

Vaatleme juhte, kus

- $$Q_{\mathrm{BP}} = \{q \mid q(\mathbf{u}, \mathbf{x}) = q(\mathbf{u})q(\mathbf{x})\}$$

ja

- $$Q_{\mathrm{VMP}} = \{q \mid q(\mathbf{u}, \mathbf{x}) = \prod_{t=1}^{T} q_t(u_t, x_t)\},$$

ning KL kauguste minimiseerimine on võimalik iteratiivse protseduuri abil.

Esimene kitsendus annab meile *belief propagation* algoritmi ja teine *variational message passing* algoritmi.

1. Kolmekaupa Markovi ahelate ja Viterbi raja probleem
2. Variatsioonilised Bayesi meetodid
3. **Implementatsioon**
4. Tulemused

Belief propagation algoritm

Initsialisatsioon. Olgu $q_u^{(0)}$ Markovi mõõt ja kuulugu vaadeldavasse simpleksisse.

Iteratsioon. Leiame mõõdu $q_x^{(0)}$, mis minimiseerib KL kaugust $D_{\mathrm{KL}}\left[q_u^{(0)} \times q_x \| p\right]$.

Seejärel leiame mõõdu $q_u^{(1)}$, mis minimiseerib KL kaugust $D_{\mathrm{KL}}\left[q_u \times q_x^{(0)} \| p\right]$.

Iteratsioonisammude järel saame mõõtude jada, kusjuures

- igal sammul KL kaugust minimiseeriv mõõt on samuti Markovi mõõt ning
- jada piirväärtuseks on alloleva minimiseerimisülesande lahend.

$$\min_{q \in Q} D_{\mathrm{KL}}\left[q \| p\right] \quad .$$

Belief propagation algoritm

Iteratsioonisamm on mõlemal juhul sarnane. Kirjutame selle välja q_x jaoks

$$\ln q_x^{(0)}(\mathbf{x}) = \frac{1}{Z_x} \sum_{\mathbf{u}} q_u^{(0)}(\mathbf{u}) \ln p(\mathbf{u}, \mathbf{x}),$$

kusjuures arvutuslikult ei ole sellise mõõdu leidmine raske edasi-tagasi algoritmi abil.

Kui mõõdud on koondunud, rakendame Viterbi raja lähendi leidmiseks Viterbi algoritmi mõõdule q_x.

.

Variational message passing algoritm

Initsialisatsioon. Olgu $q_{\setminus t}^{(0)}$ Markovi mõõt ja kuulugu vaadeldavasse simpleksisse.

Iteratsioon. Leiame mõõdu $q_t^{(0)}$, mis minimiseerib KL kaugust $D_{\mathrm{KL}}\left[q_t \times q_{\setminus t} \| p\right]$.

Kordame iga ajasammu $t = 1, \ldots, T$ jaoks.

Minimiseeriv mõõt on

$$\begin{aligned}\ln q_t(u_t, x_t) = \frac{1}{Z_t} \sum_{u_{t-1}, x_{t-1}} q_{t-1}^{(0)}(u_{t-1}, x_{t-1}) \ln p(u_t, x_t | u_{t-1}, x_{t-1}) \\ + \frac{1}{Z_t} \sum_{u_{t+1}, x_{t+1}} q_{t+1}^{(0)}(u_{t+1}, x_{t+1}) \ln p(u_{t+1}, x_{t+1} | u_t, x_t) \cdot\end{aligned}$$

1. Kolmekaupa Markovi ahelate ja Viterbi raja probleem
2. Variatsioonilised Bayesi meetodid
3. Implementatsioon
4. **Tulemused**

Vahelduva režiimiga mudel

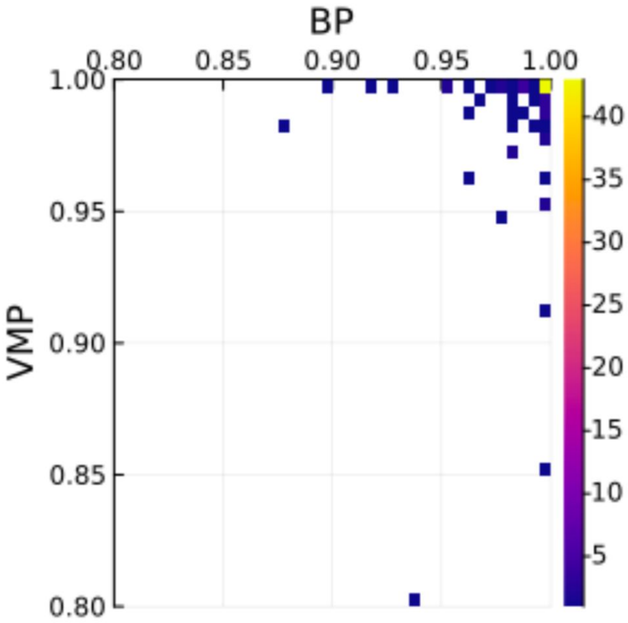
Olgu $\mathbf{U} = \{U_t\}_{t=1}^T$ Markovi ahel algtõenäosusega π_u ning üleminekumaatriksiga $\mathbf{r} = (r_{ij})_{i,j=1}^U$, mis on küllalt lähedane ühikmaatriksile. Seisundeid nim režiimideks.

Nõuame, et igas režiimis $u \in \mathcal{U} = \{A, B, C\}$ oleks juhuslik suurus $\mathbf{X} \in \{0, 1\}^T$ Markovi ahel.

$$p(x_t | x_{t-1}, u_{t-1} = u, u_{t-1} = u) = p_u(x_t | x_{t-1})$$

Ütleme, et juhuslikku suurust loetakse müraga, kui $y_t \sim \mathcal{N}(x_t, \sigma)$.

Saame genereerida mitu realisatsiooni ning rakendada kõigile kaht algoritmi.



Joonis 2: VMP ja BP algortimide protsentiilide kahedimensionaalne histogramm $N = 100$ erineva katse korral, kus $T = 15$ ja paarikaupa Markovi ahel on igal katsel sama. Graafikul kujutame eksperimendi $n \in \{1, \ldots, N\}$ korral VMP ja BP algoritmide Viterbi radade lähendite protsentiilid $r(\boldsymbol{x}^*_{\mathrm{VMP}}), r(\boldsymbol{x}^*_{\mathrm{BP}})$ histogrammina.

Tabel 3: VMP ja BP algortimide tulemused $N = 100$ erineva katse korral, kus $T = 15$ ja paarikaupa Markovi ahel on igal katsel sama. Leiame protsentiilid $r(\boldsymbol{x}^*_{\mathrm{VMP}}), r(\boldsymbol{x}^*_{\mathrm{BP}})$.
Järgmisena leiame KL kaugused marginaaljaotuse p_x ja lähendite vahel.
Leiame absoluutarvu radadest, mille marginaaltõenäosus oli suurem hinnangutest.
Leiame Hammingu kaugused erinevate algoritmide puhul ning viimasel neljal real kujutame 2., 3., 4. ja 5. kõige tõenäolisema raja Hammingu kaugusi Viterbi rajast.

	mean	min	max
$r(\boldsymbol{x}^*_{\mathrm{BP}})$	0.9403	0.005	1.0
$r(\boldsymbol{x}^*_{\mathrm{VMP}})$	0.9671	0.5037	1.0
$D_{\mathrm{KL}}\left[p_x \| q_x\right]$	11.5	0.7671	62.78
$D_{\mathrm{KL}}\left[p_x \| \sum_{\boldsymbol{u}} \prod_{t=1}^{T} q_t\right]$	5.145	0.5632	22.25
$D_{\mathrm{KL}}\left[q_x \| p_x\right]$	7.313	0.6023	25.4
$D_{\mathrm{KL}}\left[\sum_{\boldsymbol{u}} \prod_{t=1}^{T} q_t \| p_x\right]$	2.67	0.3981	7.225
$\sum_{\boldsymbol{x}} I_{p(\boldsymbol{x})>p(\boldsymbol{x}^*_{\mathrm{BP}})}(\boldsymbol{x})$	1955.0	0.0	32600.0
$\sum_{\boldsymbol{x}} I_{p(\boldsymbol{x})>p(\boldsymbol{x}^*_{\mathrm{VMP}})}(\boldsymbol{x})$			
H vmp	3.24	0.0	12.0
H bp	3.74	0.0	14.0
H orig	4.44	0.0	10.0
H naive	3.06	0.0	7.0
H_2	1.36	1.0	6.0
H_3	1.4	1.0	5.0
H_4	1.75	1.0	7.0
H_5	1.69	1.0	5.0