WiLIで用いるなくしもの位置推定の理論について

品川風丸

2024年1月21日

1　はじめに

本ファイルではWiLIに用いるなくしもの位置推定の理論について述べる。またREADME.mdはですます調で書かれているが、本ファイルは数学に関する内容が主であることを踏まえだである調で書く。

なくしものの位置を推定するにあたり、まずなくしものの発生の原理について考える。経験則としてなくしものはある動作から別の動作へ移る際に手に持っているものを置く場合に発生しやすい。例えば筆者の実体験として、郵便受けから新聞紙を取ろうと手に持っていたスマホを床に置いてそのまま忘れる等がある。WiLIではこの経験則をもとに利用者の行動を確率論を用いてモデル化し、その下でなくしものの位置の確率分布を計算する。

2　理論

2.1　概要

利用者の行動について以下の2つの仮定を敷く。

- 仮定1：利用者は動作をノードとするマルコフ連鎖に従う。
- 仮定2：ある動作から別の動作へ遷移する際に確率的になくしものが発生する。またこの際の確率は遷移先の動作にのみ依存するとし、その確率を遷移失敗確率と呼ぶ。

この仮定の下では遷移を繰り返すといつかはなくしものが発生する。さらに各動作ごとになくしものが発生した際に遷移先がその動作であった確率を計算できる。この確率をなくしもの発生確率と呼ぶ。

ここで各動作ごとに利用者がその動作下でどこにいるかの傾向を表す位置分布を用意し、全動作に渡り利用者の位置分布をなくしもの発生確率を重みとして足し合わせる。こうして得られた分布はなくしものが発生した際に利

用者がいた位置の確率分布を表していると考えられる。そこでこの分布をなくしものの位置の確率分布とみなす。

2.2 定式化

表 2.1 のように記号を置く。

表 2.1: 記号

記号	概要
X	捜索範囲 $\mathbb{R}^2$ または $\mathbb{R}^3$ の部分集合
n	動作の数 （利用者が従うマルコフ連鎖のノードの数）
$M_i(i = 1, 2, \cdots, n)$	i 番目の動作
$m_t(t = 0, 1, 2, \cdots)$	t 回目の遷移を終えた直後の動作 ただし m_0 は初期動作とする
LI	なくしものが発生したことを表す
NLI	なくしものが発生しなかったことを表す
$r_t(t = 1, 2, \cdots)$	t 回目の遷移でなくしものが発生した場合は LI を、 発生しなかった場合は NLI を返す確率変数
τ	次を満たす数 任意の $t < \tau$ に対し $r_t = \mathrm{NLI}$ かつ $r_\tau = \mathrm{LI}$ （τ 回目までの遷移ではなくしものが発生せず、 τ 回目の遷移で発生する）
s_i	初期動作が M_i である確率 $\mathrm{P}(m_0 = M_i)$

これらの記号の下、仮定 1 と仮定 2 加えて利用者の位置分布を次のように定式化する。

条件 1 (仮定 1)**.** $\mathrm{P}(m_{t+1} = M_j \mid m_t = M_i)$ は i と j にのみ依存する。この確率は M_i から M_j への遷移確率と呼び、a_{ij} と書く。

条件 2 (仮定 2)**.** t に依らず次式が成り立つ。

$$\mathrm{P}(r_{t+1} = \mathrm{LI} \mid m_{t+1} = m_t) = 0 \tag{2.1}$$

条件 3 (仮定 2)**.** $\mathrm{P}(r_{t+1} = \mathrm{LI} \mid m_{t+1} = M_i, m_{t+1} \neq m_t)$ は i にのみ依存する。この確率を M_i への遷移失敗確率と呼び、θ_i と書く。

条件 4 (利用者の位置分布)**.** M_i に X 上の確率分布が付随しているとし、その密度関数を $b_i : X \to \mathbb{R}$ と書く。

また M_i でのなくしもの発生確率 $p_i (i = 1, 2, \cdots, n)$ を式 (2.2) で定める。

$$p_i = \mathrm{P}(\tau < \infty, m_\tau = M_i) \tag{2.2}$$

p_i は a_{ij}、s_i、θ_i を用いれば式 (2.3) のように表せる。導出は補足として 3 章に書く。

$$\boldsymbol{p} = \boldsymbol{L}(\boldsymbol{I} - \boldsymbol{K})^{-1}\boldsymbol{s} \tag{2.3}$$

ただし

$$\boldsymbol{p} = (p_1\ p_2\ \cdots\ p_n)^{\mathrm{T}}$$

$$\boldsymbol{L} = \begin{pmatrix} 0 & \theta_1 a_{12} & \cdots & \theta_1 a_{1n} \\ \theta_2 a_{21} & 0 & & \theta_2 a_{2n} \\ \vdots & & \ddots & \vdots \\ \theta_n a_{n1} & \theta_n a_{n2} & \cdots & 0 \end{pmatrix}$$

$\boldsymbol{I}$ n 行 n 列の単位行列

$$\boldsymbol{K} = \begin{pmatrix} a_{11} & (1-\theta_1)a_{12} & \cdots & (1-\theta_1)a_{1n} \\ (1-\theta_2)a_{21} & a_{22} & & (1-\theta_2)a_{2n} \\ \vdots & & \ddots & \vdots \\ (1-\theta_n)a_{n1} & (1-\theta_n)a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

$$\boldsymbol{s} = (s_1\ s_2\ \cdots\ s_n)^{\mathrm{T}}$$

備考として $\boldsymbol{L}$ と $\boldsymbol{K}$ は a_{ij} に対し次式を満たす。

$$\boldsymbol{L} + \boldsymbol{K} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & & a_{2n} \\ \vdots & & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \tag{2.4}$$

なくしもの位置の確率分布の密度関数 $h : X \to \mathbb{R}$ は p_i と b_i を用いて式 (2.5) で求まる。

$$h = \sum_i p_i b_i \tag{2.5}$$

2.3 パラメータ学習方法

必要なパラメータは遷移確率 a_{ij}、初期確率 s_i、遷移失敗確率 θ_i、利用者の位置分布 b_i の 4 種である。

このうち a_{ij}、s_i、b_i について、これらはマルコフ連鎖の各ノードに確率分布が付随しているという点で隠れマルコフモデル（以下 HMM）に類似している。そのため利用者の位置分布を正規分布に限定すれば、利用者の位置推移から HMM の学習アルゴリズムによりパラメータの学習が可能であると考えられる。

残る θ_i はベイズ推定により学習する。$\boldsymbol{\theta} = (\theta_1\ \theta_2\ \cdots\ \theta_n)^{\mathrm{T}}$ が従う確率分布の密度関数 $f : [0,1]^n \to \mathbb{R}$ を真のなくしもの位置 $\boldsymbol{x}_{\mathrm{true}} \in X$ を用いて式 (2.6) により更新する。

$$f(\boldsymbol{\theta}|\boldsymbol{x}_{\mathrm{true}}) \propto h(\boldsymbol{x}_{\mathrm{true}}|\boldsymbol{\theta})f(\boldsymbol{\theta}) \tag{2.6}$$

2.4 なくしもの位置の推定方法

遷移失敗確率 θ_i の学習は点推定ではなく分布を推定している。そこでなくしもの位置の推定結果としてなくしもの位置の確率分布の $\boldsymbol{\theta}$ に関する平均値、すなわち式 (2.7) を密度関数としてもつ確率分布を採用する。

$$\begin{aligned} \mathrm{E}_{\boldsymbol{\theta}}[h]\ :\ & X \to \mathbb{R} \\ & \boldsymbol{x} \mapsto \int h(\boldsymbol{x}|\boldsymbol{\theta})f(\boldsymbol{\theta})\mathrm{d}\boldsymbol{\theta} \end{aligned} \tag{2.7}$$

3 補足：なくしもの発生確率の導出

本章では式 (2.3) の導出をする。

まず $p_{ti} = \mathrm{P}(\tau = t\ , m_\tau = M_i)$ とおく。これはなくしもの発生確率 p_i に対し次式を満たす。

$$p_i = \sum_{t=1}^{\infty} p_{ti} \tag{3.1}$$

また $\boldsymbol{p}_t = (p_{t1}\ p_{t2}\ \cdots\ p_{tn})^{\mathrm{T}}$ とおけば次のようにも書ける。

$$\boldsymbol{p} = \sum_{t=1}^{\infty} \boldsymbol{p}_{ti} \tag{3.2}$$

そこで式 (2.3) を導くために p_{ti} の t に関する一般項を求める。

p_{ti} の一般項を求めるにあたって $q_{ti} = \mathrm{P}(\tau > t\ , m_t = M_i)$ の一般項を求め

る。q_{ti} は次の漸化式を満たす。

$$
\begin{aligned}
q_{t+1\ i} &= \sum_k \mathrm{P}(\tau > t+1\ , m_{t+1} = M_i\ , m_t = M_k) \\
&= \sum_k \mathrm{P}(\tau \neq t+1, m_{t+1} = M_i, m_t = M_k | \tau > t)\mathrm{P}(\tau > t) \\
&= \sum_k \left(\begin{array}{l} \mathrm{P}(\tau \neq t+1 | \tau > t, m_{t+1} = M_i, m_t = M_k) \\ \times \mathrm{P}(m_{t+1} = M_i, m_t = M_k | \tau > t)\mathrm{P}(\tau > t) \end{array} \right) \\
&= \sum_k \left(\begin{array}{l} \mathrm{P}(r_t = \mathrm{NLI} | \forall t' \leq t \text{ s.t. } r_{t'} = \mathrm{NLI}, m_{t+1} = M_i, m_t = M_k) \\ \times \mathrm{P}(m_{t+1} = M_i, m_t = M_k)\mathrm{P}(\tau > t) \end{array} \right) \\
&= \sum_k \left(\begin{array}{l} \mathrm{P}(r_t = \mathrm{NLI} | m_{t+1} = M_i, m_t = M_k) \\ \times \mathrm{P}(m_{t+1} = M_i | m_t = M_k)\mathrm{P}(m_t = M_k)\mathrm{P}(\tau > t) \end{array} \right) \\
&= \sum_k \left(\begin{array}{l} \mathrm{P}(r_t = \mathrm{NLI} | m_{t+1} = M_i, m_t = M_k) \\ \times \mathrm{P}(m_{t+1} = M_i | m_t = M_k)\mathrm{P}(\tau > t, m_t = M_k) \end{array} \right) \\
&= \sum_{k \neq i} (1-\theta_i) a_{ki} q_{tk} + a_{ii} q_{ti}
\end{aligned}
\tag{3.3}
$$

この式 (3.3) は $\boldsymbol{q}_t = (q_{t1}\ q_{t2}\ \cdots\ q_{tn})^{\mathrm{T}}$ とおけば次式のようにも書ける。

$$
\boldsymbol{q}_{t+1} = \begin{pmatrix} a_{11} & (1-\theta_1)a_{21} & \cdots & (1-\theta_1)a_{n1} \\ (1-\theta_2)a_{12} & a_{22} & \cdots & (1-\theta_2)a_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ (1-\theta_n)a_{1n} & (1-\theta_n)a_{2n} & \cdots & a_{nn} \end{pmatrix} \boldsymbol{q}_t \tag{3.4}
$$

式 (3.4) の係数行列を $\boldsymbol{K}$ とおく。一方初項 q_{1i} についても $\mathrm{P}(\tau > 0) = 1$ であることに注意すれば式 (3.3) と同様にして次式を得る。

$$
\begin{aligned}
q_{1i} &= \sum_k \left(\begin{array}{l} \mathrm{P}(r_1 = \mathrm{NLI} \mid m_1 = M_i, m_0 = M_k) \\ \times \mathrm{P}(m_1 = M_i \mid m_0 = M_k)\mathrm{P}(m_0 = M_k) \end{array} \right) \\
&= \sum_{k \neq i} (1-\theta_i) a_{ki} s_k + a_{ii} s_i
\end{aligned}
\tag{3.5}
$$

この式 (3.5) は $\boldsymbol{s} = (s_1\ s_2\ \cdots\ s_n)^{\mathrm{T}}$ と $\boldsymbol{K}$ を使えば次式のようになる。

$$
\boldsymbol{q}_1 = \boldsymbol{K}\boldsymbol{s} \tag{3.6}
$$

よって式 (3.4)、(3.6) から $\boldsymbol{q}_t$ の一般項は次式となる。

$$
\boldsymbol{q}_t = \boldsymbol{K}^t \boldsymbol{s} \tag{3.7}
$$

次に p_{ti} の一般項について、式 (3.3) と同様にして次式を得る。

$$
\begin{aligned}
p_{t+1\ i} &= \sum_k \mathrm{P}(\tau = t+1, \sigma_{t+1} = S_i, \sigma_t = S_k) \\
&= \sum_k \left(\begin{array}{l} \mathrm{P}(r_t = \mathrm{LI} \mid m_{t+1} = M_i, m_t = M_k) \\ \times \mathrm{P}(m_{t+1} = M_i \mid m_t = M_k) \mathrm{P}(\tau > t, m_t = M_k) \end{array} \right) \\
&= \sum_{k \neq i} \theta_i a_{ki} q_{tk}
\end{aligned} \tag{3.8}
$$

この式 (3.8) は $\boldsymbol{p}_t$ を使えば次のようになる。

$$
\boldsymbol{p}_{t+1} = \begin{pmatrix} 0 & \theta_1 a_{21} & \cdots & \theta_1 a_{n1} \\ \theta_2 a_{12} & 0 & \cdots & \theta_2 a_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ \theta_n a_{1n} & \theta_n a_{2n} & \cdots & 0 \end{pmatrix} \boldsymbol{q}_t \tag{3.9}
$$

この係数行列を $\boldsymbol{L}$ とおき式 (3.7) を代入すれば次式を得る。

$$
\boldsymbol{p}_t = \boldsymbol{L}\boldsymbol{K}^{t-1}\boldsymbol{s} \tag{3.10}
$$

$\boldsymbol{p}_t$ の一般項が求まったので次に $\boldsymbol{p}$ について求める。式 (3.10) を式 (3.2) に代入すると次のようになる。

$$
\begin{aligned}
\boldsymbol{p} &= \sum_{t=1}^{\infty} \boldsymbol{L}\boldsymbol{K}^{t-1}\boldsymbol{s} \\
&= \boldsymbol{L} \left(\sum_{t=0}^{\infty} \boldsymbol{K}^t \right) \boldsymbol{s}
\end{aligned} \tag{3.11}
$$

ここで $\sum \boldsymbol{K}^t$ が収束性について、ゲルシュゴリンの定理 [1] より $\boldsymbol{K}^{\mathrm{T}}$ の任意の固有値 λ はいずれかの i に対し次式を満たす。

$$
a_{ii} - \sum_{j \neq i} (1 - \theta_j) a_{ij} \leq \lambda \leq a_{ii} + \sum_{j \neq i} (1 - \theta_j) a_{ij} \tag{3.12}
$$

式 (3.12) の右辺は次のように評価できる。

$$
\begin{aligned}
\lambda &\leq a_{ii} + \sum_{j \neq i} (1 - \theta_j) a_{ij} \\
&\leq a_{ij} + \sum_{j \neq i} a_{ij} \\
&= 1
\end{aligned}
$$

等号が成り立つのは任意の j に対し $a_{ij} > 0 \Rightarrow \theta_j = 0$ が成り立つときに限

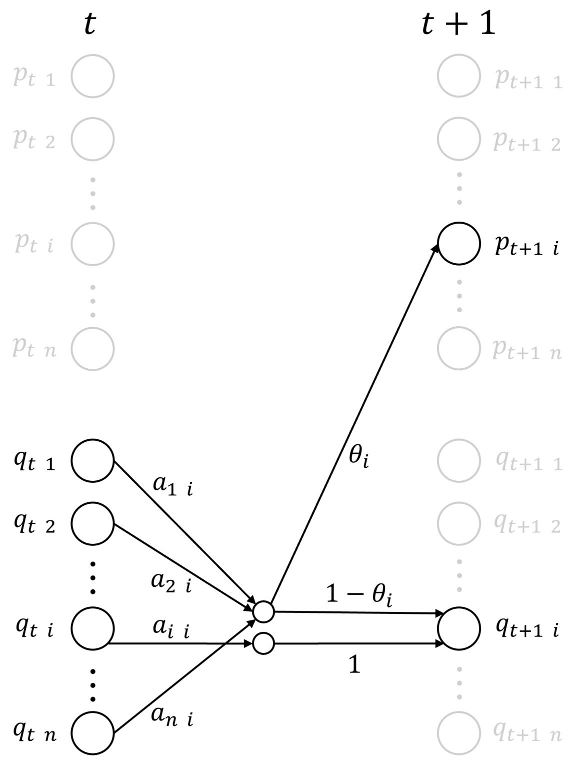

図 3.1: p_{ti} と q_{ti} の漸化式のイメージ

る。また式 (3.12) の左辺についても次のように評価できる。

$$\begin{aligned}\lambda &\geq a_{ii} - \sum_{j\neq i}(1-\theta_j)a_{ij} \\ &\geq a_{ii} \\ &\geq 1\end{aligned}$$

等号が成り立つのは $a_{ii}=1$ のときに限る。よって任意の i に対し $a_{ii}<1$ かつ、任意の i に対し適当な j が存在して $a_{ij}>0$ かつ $\theta_j>0$ となるとき $\boldsymbol{K}^{\mathrm{T}}$ の固有値はいずれも絶対値が 1 未満であり、また $\boldsymbol{K}$ と $\boldsymbol{K}^{\mathrm{T}}$ の固有値は等しいので $\boldsymbol{K}$ の固有値も絶対値が 1 未満となる。このとき $\sum \boldsymbol{K}^t$ は収束し収束先は行列等比級数の公式 [2] により次のように求まる。

$$\sum_{t=0}^{\infty}\boldsymbol{K}^t = (\boldsymbol{I}-\boldsymbol{K})^{-1} \tag{3.13}$$

ただし $\boldsymbol{I}$ は n 行 n 列の単位行列である。

最後に式 (3.13) を式 (3.11) に代入すれば式 (2.3) を得る。

参考文献

[1] 齊藤：数值解析学入門．pp.108–109，東京大学出版会，2012.

[2] 齋藤：基礎数学 1　線形代数入門．pp.212，東京大学出版会，1966.