【タイトル仮】時系列データの汎化的予測モデル開発のための評価関数の推定

武藤 克弥 (Katsuya Mutoh)
u255018@st.pu-toyama.ac.jp

富山県立大学 大学院 電子・情報工学専攻 情報基盤工学部門

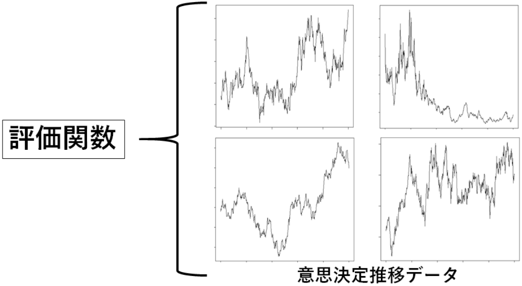
August 4, 2023

研究テーマについて

意思決定行動に関わる評価関数の推定

研究の背景 (仮説)

様々な意思決定活動はある評価関数で説明できるのではないか？
＝ 意思決定における行動は**ある評価関数をもとに決定される**



研究の目的

意思決定によって形成される時系列データ (金融・不動産・建設事業など) に対し，その結果の原因となった評価関数を推定するモデルを開発する．

研究テーマについて

評価関数推定のメリット

- 従来予測におけるデメリットの脱却
  - 従来の時系列予測にある「データの過学習」「パラメータ推定」などを行う手間がない
- 多分野への汎用性
  - 従来は少し分野が変わるだけで応用しにくいものもある
  - 高速道路の利用需要が上昇 ＝ 評価関数 V を最適化
    →新商品の売上予測

研究テーマについて

研究の新規性

意思決定モデルの拡張
→モデル (データ) に潜む評価関数を特定

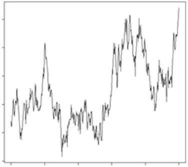
研究の流れ

1. 意思決定モデルの構築
2. 評価関数の推定手法の導入

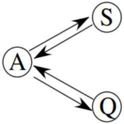
意思決定の枠組み

意思決定モデル [1]

- 時系列データに応じて，利益 (価値) の高くなる戦略を選択する



時系列データ
=世の中の情勢(株価など)



【仮定】
いずれかの戦略(アクティビティ)を選択

例) 建物の運用(3種のアクティビティ)
A: 運用継続
S: 運用休止
Q: 更地化(解体)

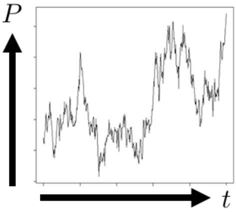
どの時間に何の戦略を取っているのが最適なのか？
→最適値関数で評価

[1]長江, 赤松., 2004.

アクティビティ価値の導出

時系列データの準備

- 確率微分方程式によって生成された時系列データを使用



時系列データ
=世の中の情勢(株価など)

P: 時刻 t の状態 (価格や需要量)

確率微分方程式 (幾何ブラウン運動)

$$dP(t) = \alpha(t, P)dt + \sigma(t, P)dZ(t)$$

- $t \in [0, T]$: 意思決定期間
- $\alpha(t, P) = \alpha P$
- $\sigma(t, P) = \sigma P$
- α: ドリフト係数 (期待収益率)
- σ: 拡散係数 (ボラティリティ)
- $Z(t)$: ノイズ (1 次元 Wiener 過程)

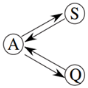
アクティビティ価値の導出

最適値関数: $V(t, P, n)$

- 各時間 $t \in [0, T]$ で全アクティビティの価値を計算
  →時間 t でどの戦略を選べば最適か分かる

$V(t, P, n)$: 時刻 t におけるアクティビティ n の将来価値

→3要素で構成 $\boxed{\text{(1) 利潤: } \pi_n(t, P), \quad \text{(2) 満期利潤}, \quad \text{(3) 推移コスト: } C_{n,m(t)}}$



- τ: アクティビティ切り替え時間

$$V(t,P,n) = \max_{\tau \in [t,T]} E\left[\underbrace{\int_t^{\tau} e^{-r(s-t)} \pi_n(s, P(s)) ds}_{\text{期間 } [t,\tau] \text{ での獲得利潤}} + \max_{m(\tau) \in O(n)} e^{-r(\tau - t)} \left\{ \underbrace{V(\tau, P(\tau), m(\tau))}_{\tau \text{ での満期利潤}} \underbrace{- C_{n,m(\tau)}}_{\tau \text{ で戦略 } n \to m \text{ への推移コスト}} \right\} \middle| P(t) = P \right]$$

- $n \in \{A, S, Q\}$: アクティビティ
- r: 割引率
- $O(n)$: アクティビティ n の推移先集合

アクティビティ価値の導出

アクティビティ価値の導出 [2]

- 同時に全アクティビティ価値は求められない
  →グラフを分解して 1 つずつ求める

全アクティビティ導出手順 $V(t,P,n) = V_n(t,P)$

(1)終端アクティビティの価値を求める

これ以上遷移しないアクティビティ=④, ⑤

線形偏微分方程式

$$\boxed{\mathcal{L}_{n'} V_{n'}(t,P) + \pi_{n'}(t,P) = 0 \quad \forall t \in [0,T)}$$

終端条件: $V_{n'}(T,P(T)) = F_{n'}(P(T))$

偏微分作用素: $\mathcal{L}_n = \dfrac{\partial}{\partial t} + \dfrac{\partial}{\partial P}\alpha_n(t,P) + \dfrac{1}{2}\dfrac{\partial^2}{\partial P^2}\{\alpha_n(t,P)\}^2 - r$

(2)推移先が "終端アクティビティのみ" のアクティビティの価値を求める

→④, ⑤をもとに③の価値を求められる

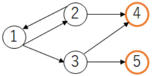
Fukushima型 merit関数による解の更新[2]

(3)サイクル構造内のアクティビティの価値を求める

→③, ④ (①⇔②の推移先)をもとに①, ②の価値を計算

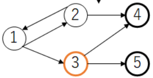
解の更新(2)と同様

(1)



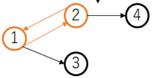
$V_4(t,P), V_5(t,P) \quad \text{for } 0\sim T$

(2)



$V_3(t,P) \leftarrow V_4(t,P), V_5(t,P) \quad \text{for } 0\sim T$

(3)



$\begin{matrix} V_1(t,P) \\ V_2(t,P) \end{matrix} \leftarrow V_3(t,P), V_4(t,P) \quad \text{for } 0\sim T$

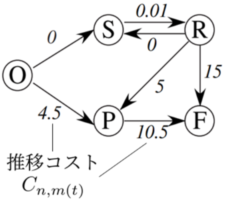
---

[2]Fukushima., 1992.

現在の取り組み

意思決定モデルの再現 [3]

- 有料道路の建設・運用事業における意思決定モデル



アクティビティ: $n \in \{O, S, R, P, F\}$

$V_n(t, P)$

(1) 利潤: $\pi_n(t, P)$,　~~(2) 満期利潤,~~　(3) 推移コスト: $C_{n,m(t)}$

O: 事前評価 $\pi_{\mathrm{O}}(t, P) = -M_{\mathrm{O}}$

S: 計画凍結 $\pi_{\mathrm{S}}(t, P) = 0$

R: 再評価 $\pi_{\mathrm{R}}(t, P) = -M_{\mathrm{R}}$

P: 1区間供用 $\pi_{\mathrm{P}}(t, P) = X_{\mathrm{P}}P - E_{\mathrm{P}}$

F: 2区間供用 $\pi_{\mathrm{F}}(t, P) = X_{\mathrm{F}}P - E_{\mathrm{F}}$

$$\boxed{\begin{array}{l} T = 20,\ r = 5\%,\ \alpha = 1\%,\ \sigma = 40\% \\ M_{\mathrm{O}} = 0.02,\ M_{\mathrm{R}} = 0.01 \\ X_{\mathrm{P}} = 0.5,\ E_{\mathrm{P}} = 0.6,\ X_{\mathrm{F}} = 1,\ E_{\mathrm{F}} = 1 \end{array}}$$

[3] 長江, 赤松., 2004.

TOYAMA Prefectural University

現在の取り組み

意思決定モデルの再現

- アルゴリズムをプログラム化し，同じ結果が得られるようにする

(1) 終端アクティビティ(F)価値の導出

$$\boxed{\mathcal{L}_{n'}V_{n'}(t,P)+\pi_{n'}(t,P)=0}$$

$\forall t \in [0,T)$

$\xrightarrow{\text{離散化}}$

$(t,P)\simeq(t^i,P^j)=(i\Delta T, P_0+j\Delta P)$ $(i=0,1,\cdots,I)$ $(j=0,1,\cdots,J)$

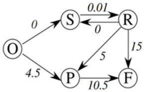
$$\boxed{\boldsymbol{V}_n^i=-(\boldsymbol{L}_n^i)(\boldsymbol{M}_n^i\boldsymbol{V}_n^{i+1}+\boldsymbol{\pi}_n^i)}$$

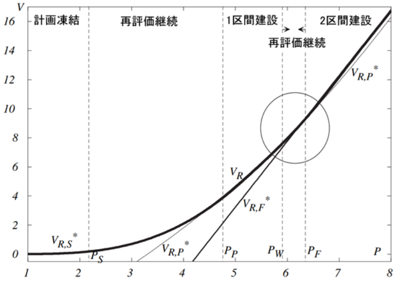
$I=J=400$

$\Delta t=\frac{20}{400},\quad \Delta P=\frac{5}{400}$

$\boldsymbol{L}_n^i,\ \boldsymbol{M}_n^i$ ：作用素行列

for all $n\in N$ **do** (終端条件)
$V_n^1 := \max.\left[F_n,\ \max_{m\in O(n)}.\left\{V_m^1-1C_{n,m}\right\}\right];$
end for





t=15におけるアクティビティRの価値

(2)(3) 残ったアクティビティ価値の導出

Fukushima型 merit関数[2]

[Alg-Merit]

Step 0 初期可能解 $X_c^{i(1)}\in\mathcal{R}_+$, $k:=1$.

Step 1 降下方向ベクトルの決定.

$$d^{(k)}:=H(X_c^{i(k)}).$$

Step 2 ステップ・サイズ α を，以下の一次元探索問題の解として求める.

$$\alpha^*=\arg.\min_{\alpha\in[0,1]}\Phi\left(X_c^{i(k)}+\alpha d^{(k)}\right).$$

Step 3 解の改訂. $X_c^{i(k+1)}:=X_c^{i(k)}+\alpha^* d^{(k)}$

Step 4 収束判定：収束していれば停止，そうでなければ $k:=k+1$ として Step 1 へ.

評価関数の推定

ロボットの起き上がり運動獲得のための
正規化ガウス関数ネットワーク (NGnet) を用いた Actor-critic 強化学習 [4]

研究概要

「目標出力 y(t) メートルまで頭が上がるように関節を動かす」ように学習させる

- ロボットは目標出力に達するように最適な「関節の動き」を探る
- 推定評価関数＝「この動きをすれば**目標出力**に達する」と考えた自己評価の集約
→この関数を最適化するような行動を行う

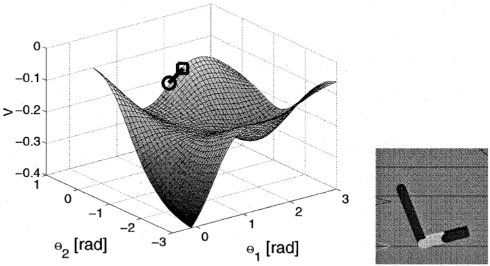
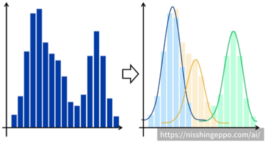


図 1: 推定評価関数と実際の行動

[4] 森本, 銅谷., 1999.

評価関数の推定

評価関数推定の概要



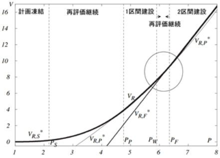
(1)関数の設定

利潤: $\pi_n(t, P)$, 推移コスト: $C_{n,m(t)}$

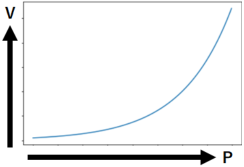
$\longrightarrow$ $b_{1,k}(\boldsymbol{x})$ $\quad$ $b_{2,k}(\boldsymbol{x})$

(2)アクティビティ価値導出アルゴリズム
(前々ページ)

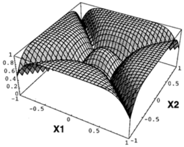
$$y(\mathbf{x}) = \sum_{k=1}^{K} w_k \underbrace{b_k(\mathbf{x})}_{\text{基底関数}}$$



2乗誤差e



V(価値)-P(状態)図の出力



(3)eが最小になるように w_k を更新

(4) (1)〜(3)を指定回数だけ繰り返す

(5)評価関数の出力

今後の予定

今後の予定

- アクティビティ価値導出のプログラムをまず完成させる