証拠に基づく政策立案のための潜在プロファイル分析と数法則発見法を用いた社会実情のモデル化と可視化

長瀬　永遠

富山県立大学　情報基盤工学講座

January 19, 2024

はじめに 1

本研究の背景

現在，EBPM の推進に向けて様々な取り組みがなされている．一方，いくつか課題も挙げられており，その中にデータの更新頻度が低いということがある．

本研究の目的

- 既存のデータから最新のデータを予測する手法を提案する
- 上記の手法の結果を用いて後述のシステムを改善する

はじめに2

先行研究

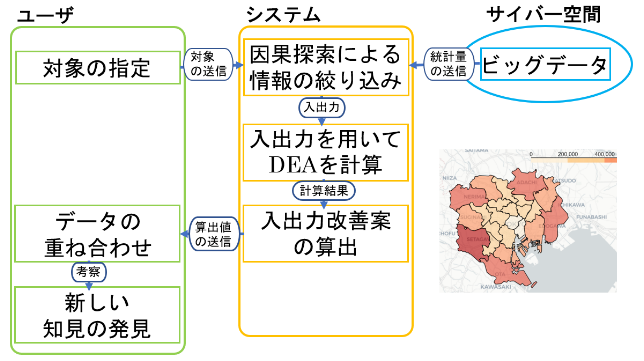

図 1: 先行研究

RF6.4 で求めたい多変量多項式

$$
\begin{aligned}
f(\mathbf{q}, \mathbf{x}; \mathbf{\Theta}) = c_0 + \sum_{j=1}^{J} c_j m_j, \\
c_0 = \sum_{r=1}^{R} v_{0r}\sigma_r, \quad c_j = \sum_{r=1}^{R} v_{jr}\sigma_r \\
\sigma_r = \sigma\left(\sum_{k=1}^{K_1}\sum_{l=1}^{L_k} v_{rkl} q_{kl}\right), \quad m_j = \exp\left(\sum_{k=1}^{K_2} w_{jk} \ln x_k\right)
\end{aligned} \tag{1}
$$

$\mathbf{q}$: 入力変数（質的）， $\mathbf{x}$: 入力変数（量的），
$\mathbf{\Theta}$: $\{\mathbf{v_0}, \mathbf{v_j}, \mathbf{v_r}, \mathbf{w_j}\}$

学習問題の定式化

$$f(\mathbf{x};\theta) = w_0 + \sum_{j=1}^{J} w_j m_j, \quad m_j = \exp\left(\sum_{k=1}^{K} w_{jk} \ln x_k\right) \tag{2}$$

$$E(\mathbf{x};\theta) = \frac{1}{2}\sum_{\mu=1}^{N} \left(f(\mathbf{x}^{\mu};\theta) - y^{\mu}\right)^2 \tag{3}$$

N : サンプル数

学習アルゴリズム

BPQ 法（Back Propagation based on Quasi-Newton）

パーセプトロンの学習：BPQ 法

① θ_1 の初期化

- パターン 1：$w_0 =$ y の平均値，$w_j = 0$，$w_{jk} = (-1, 1)$ の一様乱数
- パターン 2：全ての重みを $(-1, 1)$ の一様乱数

$\mathbb{H}_1 = \mathbb{I}$，$s = 1$ とする

- $\mathbb{I}$ は単位行列（対角成分が全て 1，その他が全て 0 の行列）

② $\Delta \theta_s = -\mathbb{H}_s \nabla E(x; \theta_s)$ を計算

停止条件を満たせば，反復を終了．

- 停止条件：勾配ベクトル $\nabla E(x; \theta_k)$ の全成分が 10^{-5} 未満 or 反復 20,000 回

（$\nabla E(x; \theta_s)$ の計算方法）

$$\nabla E(x; \theta_s) = \frac{\partial E}{\partial \theta} = \sum_{\mu=1}^{N} (f^{\mu} - y^{\mu}) \frac{\partial f^{\mu}}{\partial \theta}$$

\\ check1：ベクトル偏微分の解釈

$$= \sum_{\mu=1}^{N} (f^{\mu} - y^{\mu}) (1, \{m_j\}, \{w_j m_j \ln x_k\})$$

$$m_j = \exp\left(\sum_{k=1}^{K} w_{jk} \ln x_k\right)$$

次のページへ →

はじめに
進捗報告
先週までの内容

③ $E(x;\theta_s) + \alpha_s \Delta\theta_s)$ を最小にする α_s を求める

(用いる表記の解説)

- $E(x;\theta)$ は E, $f(x;\theta)$ は f と表記
- $E(x;\theta + \alpha\Delta\theta)$ は $g(\alpha)$ と表記

\\\\ check 2 : $g(0)$ の解釈

(α_s の求め方)

$g(\alpha) \approx g(0) + g'(0)\alpha + \frac{1}{2}g''(0)\alpha^2$ を最小化する α を求める.

\\\\ check 3 : ΔW_j, ΔW_{jk} の解釈

\\\\ check 4 : 計算のコーディングミス

- $g(0) = E$

- $g'(0) = \sum_{\mu=1}^{N} (f^{\mu} - y^{\mu}) f'^{\mu}$
- $f'^{\mu} = \Delta W_0 + \sum_{j=1}^{J} (\Delta W_j m_j^{\mu} + W_j m_j'^{\mu})$
- $m_j'^{\mu} = m_j^{\mu} \sum_{k=1}^{K} \Delta W_{jk} \ln x_k^{\mu}$.

- $g''(0) = \sum_{\mu=1}^{N} ((f'^{\mu})^2 + (f^{\mu} - y^{\mu}) f''^{\mu})$
- $f''^{\mu} = \sum_{j=1}^{J} (2\Delta W_j m_j'^{\mu} + W_j m_j''^{\mu})$
- $m_j''^{\mu} = m_j^{\mu} \left(\sum_{k=1}^{K} \Delta W_{jk} \ln x_k^{\mu}\right)^2$

ただし, ΔW_0, ΔW_j, ΔW_{jk} は ② で計算される ΔW_0, ΔW_j, ΔW_{jk} の変化量

進捗報告 5

はじめに
進捗報告
先週までの内容

$\boxed{1}$ $g'(0) > 0$ のとき　　\\ check 5 : $\boxed{1}$ の解釈（$\boxed{1}$ の後, $\boxed{2}$ の処理を行うか）

② の結果を $\Delta\theta = -\nabla E$, $\mathbb{H} = \mathbb{I}$ に置き換えて $\boxed{2\text{-}1}$, $\boxed{2\text{-}2}$ のどちらかへ進む

$\boxed{2\text{-}1}$ $g'(0) < 0$ かつ $g''(0) > 0$ のとき

$$\alpha = -\frac{g'(0)}{g''(0)}$$

$\boxed{2\text{-}2}$ $g'(0) < 0$ かつ $g''(0) \leq 0$ のとき

$$\alpha = -\frac{g'(0)}{\sum_{\mu=1}^{N} (f'^{\mu})^2}$$

また, $\alpha > 0$ のとき, $S = M$ とし, $\mathbb{H} = \mathbb{I}$ とする.

└ パラメータ θ の次元 : $(KJ + J + 1)$

次のページへ →

はじめに
進捗報告
先週までの内容

[3] $\|\alpha \Delta \boldsymbol{\theta}\| > 1.0$ のとき
$\alpha = 1.0 / \|\Delta \boldsymbol{\theta}\|$ とする
また，$s = M$ とし，$\mathbb{H} = \mathbb{I}$ とする

[4] $g(\alpha) \geq g(0)$ のとき

$$\alpha = -\frac{g'(0)\tilde{\alpha}^2}{2(g(\tilde{\alpha}) - g(0) - g'(0)\tilde{\alpha})}$$

，ただし，$\tilde{\alpha}$ は [3] までで得られた α の値

※ $g(\alpha) < g(0)$ になるまで [4] を繰り返す

④ $\boldsymbol{\theta}_{s+1} = \boldsymbol{\theta}_s + \alpha_s \Delta \boldsymbol{\theta}_s$ に修正する．

⑤ $s \equiv 0 \pmod{M = KJ + J + 1}$ ならば $\mathbb{H}_{s+1} = \mathbb{I}$ とし，
それ以外の場合，BFGS公式を用いて $\mathbb{H}_{s+1}$ を更新する．
$s \leftarrow s + 1$ として ② に戻る．

\\ check6: λ_s の解釈

(BFGS公式)

$$\mathbb{H}_{s+1} = \mathbb{H}_s + \left(1 + \frac{\boldsymbol{q}^T \mathbb{H}_s \boldsymbol{q}}{\boldsymbol{p}\boldsymbol{q}}\right)\frac{\boldsymbol{p}\boldsymbol{p}^T}{\boldsymbol{p}^T\boldsymbol{q}} - \frac{\boldsymbol{p}\boldsymbol{q}^T\mathbb{H}_s + \mathbb{H}_s\boldsymbol{q}\boldsymbol{p}^T}{\boldsymbol{p}^T\boldsymbol{q}},$$

$$\boldsymbol{p} = \underbrace{\lambda_s}_{\alpha_s} \Delta \boldsymbol{\theta}_s, \quad \boldsymbol{q} = \nabla E(\boldsymbol{x}; \boldsymbol{\theta}_{s+1}) - \nabla E(\boldsymbol{x}; \boldsymbol{\theta}_s)$$

はじめに

進捗報告

先週までの内容

最適中間ユニット数 J の決定

このモデルの場合，最適中間ユニット数 J と多変数多項式の項数は同じになるので，BPQ 法での学習は事前に J を指定し，$J = 1, 2, 3, \cdots$ と変化させながら行う．最適なモデルの選択（項数の選択）は各 J の値において以下の BIC（ベイズ情報量基準）を算出し，その値が最も小さいものを選択する．

モデル選択：BIC

Jの値を $1, 2, \cdots$ と変化させ，以下の$BIC(J)$が最小となるJを最適なモデルに選択する

$$BIC(J) = \frac{N}{2} \log\left(\frac{1}{N}\sum_{n=1}^{N}\left(f(x^n; \hat{\theta}_J) - y^n\right)^2\right) + \frac{M}{2}\log N$$

ただし，N：総サンプル数，M：総パラメータ数，
$\hat{\theta}_J$：各Jにおける BPQ法の結果

進捗報告 8

やること

1. コードのデバッグ
2. 数値実験
3. 卒論時システムと合わせる
4. 本論執筆