【研究報告】

長瀬　永遠

富山県立大学　情報基盤工学講座

December 26, 2023

進捗報告 1

現在やっていること

年始あたりに数値実験ができるように本資料○ページ以降の内容をコーディング中.

今日の発表内容

数法則発見法の一つである RF 法（Rule extraction method from Facts）について，自身の理解度確認も含めて説明する.

進捗報告 2

RF 法の最終的な目標

複数の入力変数と一つの出力変数の間に成り立つ数法則を数値的アプローチによって求める.

RF 法の利点

- 求めたい多項式の項数や各項の指数を事前に指定しなくてもよい
- 各項の指数に小数を含めることができる
- 数値的アプローチなため，組合せ論的アプローチのような組合せ爆発が起こらない

進捗報告 3

RF5 で求めたい多変量多項式

$$y = f(\mathbf{x};\theta) + \epsilon, f(\mathbf{x};\theta) = w_0 + \sum_{j=1}^{J} w_j \prod_{k=1}^{K} x_k^{w_{jk}} \tag{1}$$

$\mathbf{x}$: 入力変数，y : 出力変数，w_0 : 定数項，w_j : 各項の係数，w_{jk} : 各項の各入力変数の指数，J : 項数，ϵ : 誤差項

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

$\mathbb{H}_1 = I$, $s = 1$ とする

- 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 $\nabla E(x;\theta_s)$ 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

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

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

次のページへ →

進捗報告 5

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．

- $g(0) = E$

\\ check 3 : ΔW_j, ΔW_{jk} の解釈

\\ check 4 : 計算のコーディングミス

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

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

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 6

$\boxed{1}$ $g'(0) > 0$ 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

進捗報告

先週までの内容

進捗報告 7

$\boxed{3}$ $\|\alpha\Delta\theta\| > 1.0$ のとき

$\alpha = 1.0 / \|\Delta\theta\|$ とする

また，$s = M$ とし，$\mathbb{H} = \mathbb{I}$ とする

$\boxed{4}$ $g(\alpha) \geq g(0)$ 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 $\boxed{3}$ までで得られた α の値

※ $g(\alpha) < g(0)$ になるまで $\boxed{4}$ を繰り返す

④ $\theta_{s+1} = \theta_s + \alpha_s \Delta\theta_s$ に修正する．

⑤ $s \equiv 0 \ (\mathrm{mod}\ M = KJ + J + 1)$ 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．

check 6：λ_s の解釈

（BFGS公式）

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

$$\mathbb{p} = \lambda_s \Delta\theta_s, \quad \mathbb{q} = \nabla E(\mathbb{x};\theta_{s+1}) - \nabla E(\mathbb{x};\theta_s)$$

α_s

進捗報告 8

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, … 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

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

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
先週までの内容

RF6.4 で求めたい多変量多項式

$$
\begin{aligned}
f(\mathbf{q}, \mathbf{x}; \boldsymbol{\Theta}) &= c_0 + \sum_{j=1}^{J} c_j m_j, \\
c_0 = \sum_{r=1}^{R} v_{0r}\sigma_r, \quad c_j &= \sum_{r=1}^{R} v_{jr}\sigma_r \\
\sigma_r = \sigma\left(\sum_{k=1}^{K_1}\sum_{l=1}^{L_k} v_{rkl}q_{kl}\right), \quad m_j &= \exp\left(\sum_{k=1}^{K_2} w_{jk}\ln x_k\right)
\end{aligned} \tag{4}
$$

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

進捗報告 10

今後のスケジュール

1. RF6.4 のコーディング
2. サンプリングで省かれたデータにおける所属確率復元の検討
3. アウトプットの見せ方検討（卒研時の転用）
4. 本論執筆＋数値実験

先週までの内容

これより後ろのスライドは先週までの内容になります.

はじめに1

背景

近年，コンピューターの普及によって社会にはあらゆる分野において大量の情報が溢れるようになった．そのため，大量のデータから有用な情報を取り出す研究が数多くなされている．その中の一つに複数の入力データと単一の出力データからそれらの間に成り立つ数法則を発見する手法が挙げられる．

目的

社会事象について，データ間に成り立つ数法則を発見することによってそれらを表現する手法を提案する．また，それを行う上で課題となるデータの不完全さに考慮した手法を作成する．

はじめに2

新しい政策の効果検証がしたい…

理想

- 予算∞
- すぐに適用できる
- すぐに効果が出る

現実

- 予算簡単につかない
- 適用に数年
- 効果が出るのに数年

疑似的に効果検証できるようにすればよいのでは？

➡ そのためには社会を表現したモデルが必要

提案手法 1

RF5.0

複数の入力変数（量的）と一つの出力変数（量的）との間に成り立つ数法則を求めることができる 3 層パーセプトロンを用いた多変量多項式回帰法

数法則発見法を社会現象に適用する上での課題

行政のオープンデータは数年に一度しか更新されない
＝データを時間的な観点のみで収集するのは難しい
＝空間的な観点でのデータ収集が必要
→ 全ての自治体を同じ土俵で扱っていいのかという疑問が生まれる

提案手法 2

疑問

日本全国を見てみると人口や主要産業，地価など大きなばらつきがあるが，全てを同じ土俵で扱っていいのか？

仮説

- 全国の市区町村は潜在的にいくつかのタイプに分かれるのではないか
- タイプ分けすることで数法則発見の精度が向上するのではないか

アイデア

多数の入力変数（量的）を用いてクラスタリングを行い，発生したクラスを質的変数として入力変数に組み込むことで RF6.3 を適用する．

提案手法 3

計算量削減とインタラクティブな分析への工夫

潜在プロファイル分析（LPA）における計算量削減のために LPA には全データではなくサンプリングしたデータを用いる．サンプリングによって除外されたデータにおける各クラスへの所属確率は以下のファジィニューラルネットワークでモデルを作成することで求める．
また，所属確率を求めるためのモデルを作成しておくことで，ユーザが一部の入力を調整した場合の結果を見る際に LPA を行う必要がなくなる．

ファジィニューラルネットワーク

提案手法の概要

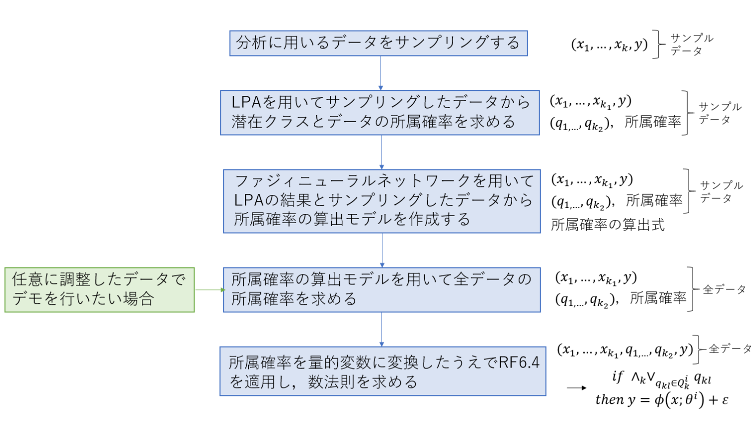

図 1: 提案手法のフロー

提案手法 5

所属確率から量的変数への変換

LPA によって求められた潜在クラスの数に基づいてカテゴリ数と閾値を設定し，所属確率を量的変数に変換する．

（潜在クラス数が 3 の場合）
各データの所属確率の合計は必ず 1 になることから以下のパターンが想定される．

1. 3 つの潜在クラスに属する → $q_1 = 0.33, q_2 = 0.33, q_3 = 0.33$
2. 2 つの潜在クラスに属する → $q_1 = 0.50, q_2 = 0.50, q_3 = 0.00$
3. 1 つの潜在クラスに属する → $q_1 = 1.00, q_2 = 0.00, q_3 = 0.00$

よって，以下のカテゴリによって量的変数を求める．
カテゴリ 1 $= 1.00 \sim 0.50$
カテゴリ 2 $= 0.50 \sim 0.33$
カテゴリ 3 $= 0.33 \sim 0.00$

今後の課題

- 提案手法のコーディング
- 数値実験の実施
- 修士論文の執筆