研究論文

# 改良型 SVR 感度分析法による 地域間死亡率格差の要因分析

田中草平[1]　松山直樹[2]

2023 年 1 月 31 日投稿
2023 年 2 月 23 日受理

**概要**

サポートベクター回帰（SVR）は、データの持つ複雑な非線形関係を表現できる反面で可読性に欠けるため、各説明変数に対して「感度」を導入し、感度の大小により説明変数の影響度の評価や説明変数の選択を可能にしたのが SVR 感度分析法である。SVR 感度分析法は説明変数空間が非常に大きく複雑な交絡関係が予想される問題に対しても有効な影響要因分析が可能であるが、感度に基づく変数削減と再学習のプロセスにおけるハイパーパラメータ探索における非効率性が問題となる。本研究ではこの SVR 感度分析法のハイパーパラメータ探索を容易にする新しいアルゴリズムを提案し、都道府県別の全死因・年齢調整死亡率の影響要因の分析において既存研究よりも優れた推定結果が得られることを示す。

キーワード：　サポートベクター回帰　感度分析　変数選択　全死因死亡率　影響要因

## 1. イントロダクション

ある事象への影響要因や影響の程度を調べる問題は、統計学的にはしばしば影響要因を説明変数、ある事象の変動の様子を目的変数の変化と考えた回帰問題として扱われる。そのような問題を扱う際に、説明変数の空間が大きい中で真に影響要因となる説明変数はそれほど多くないという仮定をスパース性といい、スパース性を仮定した際に真に影響する説明変数を選択するような技術はスパース推定と呼ばれる。ビックデータが蓄積されその解析を様々な分野で役立たせるようになった現代の社会では、スパース推定の技術の重要度は増しており、様々な手法が提案されている。本研究ではスパース推定の技術の中でもサポー

[1] 明治大学大学院先端数理科学研究科現象数理学専攻博士前期課程
[2] 明治大学総合数理学部

トベクター回帰 （Support Vector Regression: SVR）と呼ばれる手法に着目する。

影響要因分析を行う際に、古くから最も一般的に使われる手法は線形重回帰（Multiple Linear Regression: MLR）である。MLR は、可読性の高さや回帰係数の推定の容易さのため非常に扱いやすい手法である反面、データ数以上の説明変数を扱えないことや予測精度の低さ、多重共線性など問題点も多い。よって、説明変数が特に多い状況（場合によってはデータ数以上の状況）や、各要因同士が複雑に絡み合っており線形関係が認められないような状況においては、MLR では有効な分析ができないことも多い。そのような問題への対策として、機械学習の手法を用いることが考えられる。機械学習の手法の中には、非線形関係を学習できるものやデータ数以上の説明変数を扱うことができるものなど、MLR の問題点を解決したものが多くあり、MLR よりも有効な分析結果が得られる可能性がある。中でも、分類問題に多く使われる技術であるサポートベクターマシン（Support Vector Machine: SVM）を回帰問題に拡張した SVR は、MLR で発生する多くの問題点を改善できる手法であり、様々な問題への適用例が報告されている。SVR は、カーネル法による非線形化や $\varepsilon$ -不感損失関数の採用により、MLR に比べはるかに柔軟なデータへのフィッティングを実現することができる。田辺・鈴木(2020)は、SVR に「感度」という量を導入し各説明変数の影響の程度を数値化し、変数選択を可能にする方法を提案した。この感度分析法を同時に適用することで、説明変数空間が非常に大きく複雑な交絡関係が予想される問題に対しても有効な影響要因分析が可能となる。

MLR の様々な問題点を改善した手法である SVR の感度分析法だが、実装する中でハイパーパラメータ探索の非効率性の問題があることがわかった。SVR の感度分析法は、各説明変数に対して感度を計算し最も感度の絶対値が小さいものを消去した後モデルを再学習するということを繰り返すが、学習の際に一定の幅、範囲でハイパーパラメータの探索を行うと途中で最適とは言い難いモデルが生成されてしまい最適な変数の消去が行われないことがあるということが確認された。本研究では、この非効率性の解消のために、SVR において最も重要なハイパーパラメータである不感度$\varepsilon$の扱いを変え、ハイパーパラメータの表す意味が直感的にわかりやすいよう改善された拡張モデルであるニュー・サポートベクター回帰 $\nu$ -SVR（Scholkopf et al.2000）を使用した。$\nu$ -SVR を使用することで、ハイパーパラメータ $\nu$ の探索範囲を事前に絞り込むことができ変数消去の後の全てのタイミングで最適なモデルを得ることができるようになる。

本研究では、実際のデータを適用した分析を行い $\nu$ -SVR を使用した SVR の感度分析法の有効性を確認するために、地域間の死亡率格差の生まれる要因の分析を行った。具体的には、都道府県別の生活習慣・医療環境・労働環境などの指標を説明変数、全死因平均死亡率を目的変数とした回帰を行うことで健康への影響要因の特定を試みる。また、$\nu$ -SVR の感度分析法の有効性や、一般的に使われることの多い手法との違いを明らかにするために、通常の SVR や正則化を使用した MLR 等の他の様々な機械学習手法にも同じデータを適用し分析を行い、結果を比較した。

本論文の構成は以下のようになっている。2 節では、SVR の出発点とも言える技術である SVM に始まり SVR の感度分析法までを概説し、3 節で本研究のメインの技術である $\nu$ -SVR について述べ、4 節では実際のデータに対して分析を行った結果を述べる。5 節では考察を述べ、6 節で結論及び今後の発展可能性などについて述べ締めくくる。

## 2. 既存研究（SVR 感度分析法）

本節では、本研究で扱う SVR とその感度分析法について説明する。SVR は、元々分類問題の手法である SVM を回帰問題に拡張した手法である。よって本論文では、まず SVM について説明し、その拡張として SVR を導入する。

説明にあたり、各記号を以下のように定義し統一して扱うこととする。

$N$：データ数

$M$：説明変数の数

$\boldsymbol{x}_i \in \boldsymbol{R}^M$：$i$番目の説明変数ベクトル（$i = 1,2,\dots,N$）

$\boldsymbol{X} \in \boldsymbol{R}^{M\times N}$：説明変数ベクトルの行列（$= (\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_N)$）

$y_i$：$i$番目の目的変数

$\boldsymbol{y} \in \boldsymbol{R}^N$：目的変数のベクトル

$f$：予測に用いる回帰関数

$l$：損失関数

$\boldsymbol{w}$：回帰係数ベクトル

$\boldsymbol{a}^T$：$\boldsymbol{a}$の転置ベクトル

$||\quad||_1$：$L_1$ノルム（$\boldsymbol{a} = (a_1, a_2, \dots, a_n)$の時$||\boldsymbol{a}||_1 = |a_1| + |a_2| + \cdots + |a_n|$）

$||\quad||_2$：$L_2$ノルム（$\boldsymbol{a} = (a_1, a_2, \dots, a_n)$の時$||\boldsymbol{a}||_2 = \sqrt{{a_1}^2 + {a_2}^2 + \cdots + {a_n}^2}$）

### 2.1. サポートベクターマシン（Support Vector Machine: SVM）

SVM は、分類問題のために提案された機械学習手法であり、少ない計算コストで高精度な分類が可能となるノンパラメトリックな手法である。SVM は、最も基本的な 2 クラス分類をする線形関数から出発し、カーネル法（詳細は付録を参照）による非線形化・多クラス分類・異常検知・回帰など様々な拡張がされてきた。本論文では、回帰への拡張である SVR の説明に重点を置くので、最も基本的なアイディアである 2 クラス分類の場合に限って説明を行い、その後回帰問題への拡張として SVR を導入する。

まず初めに、全てのデータ点が線形分類関数で分類できる場合について導出し（ハードマ

ージン)、その後線形分離できないデータ点を含む場合について導出する（ソフトマージン)。線形分類境界を以下のように定義する：

$$f(\boldsymbol{x}) = \boldsymbol{x}^T\boldsymbol{w} + b = 0$$

分類問題なので、目的変数$y_i$は+1または−1のいずれかであり、$y_i = 1$であるデータ点$(y_i, \boldsymbol{x}_i)$に対し$f(\boldsymbol{x}_i) \geq 0$のとき正しく、$f(\boldsymbol{x}_i) < 0$のとき誤って（$y_i = -1$であるデータ点$(y_i, \boldsymbol{x}_i)$に対し$f(\boldsymbol{x}_i) \leq 0$のとき正しく、$f(\boldsymbol{x}_i) \geq 0$のとき誤って）分類しているとする。ここで、最適な分類関数$f$を求めるために以下の概念を導入する。

・サポートベクトル：分類境界$f(\boldsymbol{x}) = 0$と最も近い各クラスのデータ点

・マージン：サポートベクトルと分類境界$f(\boldsymbol{x}) = 0$との距離

SVM の基本的な考え方は、このマージンを選びマージンを最大化する分類境界を最適な判別関数とするということである。分類境界からあるデータ点$\boldsymbol{x}_i$までの距離$d_i$を以下の式で定義すると、マージンは以下の最小化問題として定式化される：

$$\text{距離：} d_i = \frac{|f(\boldsymbol{x}_i)|}{||\boldsymbol{w}||_2} = \frac{|\boldsymbol{x}_i\boldsymbol{w}+b|}{||\boldsymbol{w}||_2}$$

$$\text{マージン：} \min_i d_i \ \ ,\forall\ \ i \in \{1,2,\dots,N\}$$

よって、マージン最大化問題は以下のようになる：

$$\max_{\boldsymbol{w},b} \min_i d_i$$

$$s.t\ \ y_i\left(\boldsymbol{x}_i^T\boldsymbol{w} + b\right) \geq 0\ \ ,\forall i \in \{1,2,\dots,N\}$$

一般性を失うことなく、「サポートベクトルは、$f(\boldsymbol{x}) = +1$または$f(\boldsymbol{x}) = -1$の直線上にある」という仮定をする。この仮定を加えると、

$$\min_i d_i = \min_i \frac{\left|\boldsymbol{x}_i^T\boldsymbol{w} + b\right|}{||\boldsymbol{w}||_2} = \frac{1}{||\boldsymbol{w}||_2}$$

となり、$||\boldsymbol{w}||_2 > 0$より、

$$\max_{\boldsymbol{w},b} \min_i d_i = \max_{\boldsymbol{w},b} \frac{1}{||\boldsymbol{w}||_2} \Leftrightarrow \min_{\boldsymbol{w},b} ||\boldsymbol{w}||_2 \Leftrightarrow \min_{\boldsymbol{w},b} \frac{1}{2}||\boldsymbol{w}||_2{}^2 \Leftrightarrow \min_{\boldsymbol{w},b} \frac{1}{2}\boldsymbol{w}^T\boldsymbol{w}$$

となる。以上より、簡単化したマージン最大化問題は以下のようになる。

$$\min_{\boldsymbol{w},b} \frac{1}{2}\boldsymbol{w}^T\boldsymbol{w}$$

$$s.t.\ \ y_i\left(\boldsymbol{x}_i^T\boldsymbol{w} + b\right) \geq 1\ \ ,\forall\ \ i \in \{1,2,\dots,N\}$$

ここで制約条件は、「サポートベクトルは、$f(\boldsymbol{x}) = +1$または$f(\boldsymbol{x}) = -1$の直線上にある（$-1 < f(\boldsymbol{x}) < 1$のの範囲にデータは存在しない)」という仮定をしているので、$y_i = 1$に対して$f(\boldsymbol{x}_i) \geq 0$、$y_i = -1$に対して$f(\boldsymbol{x}_i) < 0$と正しく分類する条件となる。

ここまでの議論は全てのデータを完全に分離する線形分類関数が存在する場合のみ有効であるが、そのような状況は現実的に考え難いので、線形分離できないデータに対して分類の失敗を許容するソフトマージンを導入する。分類境界から誤分類したデータ点までの距

離を誤差$\xi_i$とし、制約式を以下のように緩和する：

$$y_i(\boldsymbol{x}_i^T\boldsymbol{w}+b)+\xi_i \geq 1 \ ,\forall \ i \in \{1,2,\dots,N\}$$

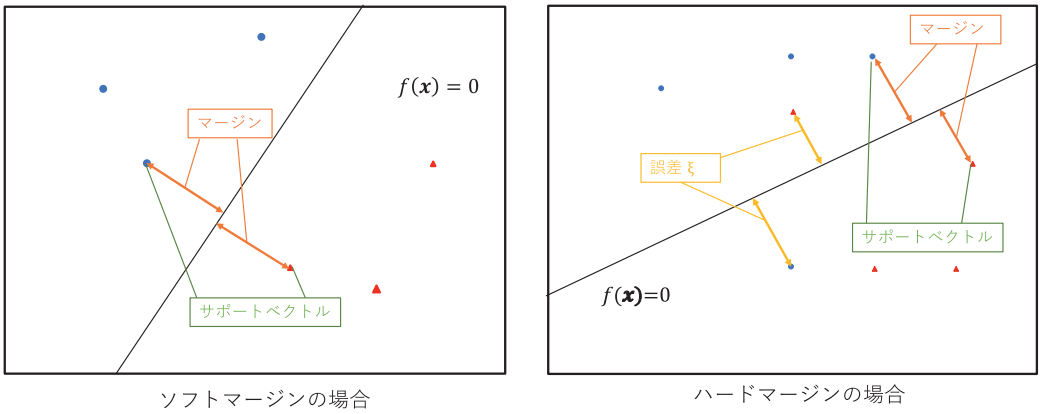
マージン最大化に加え、誤差を最小化するような分類境界を最良の分類関数とみなすと、ソフトマージン最大化問題は、以下のように書き換えることができる：

$$\min_{\boldsymbol{w},b,\xi}\left(\frac{1}{2}\boldsymbol{w}^T\boldsymbol{w}+C\sum_{i=1}^{N}\xi_i\right)$$

$$s.t \ y_i(\boldsymbol{x}_i^T\boldsymbol{w}+b)+\xi_i \geq 1 \ ,\xi_i \geq 0,\forall \ i \in \{1,2,\dots,N\}$$

ここで、$C$は誤差の重要度を調整するハイパーパラメータである。図 1 は、ハードマージンの場合とソフトマージンの場合それぞれにおけるサポートベクトル・マージン・誤差のイメージ図である。

この最適化問題は、ラグランジュの未定乗数法を用いて解くことができる。




ソフトマージンの場合　　ハードマージンの場合

図 1 ：2 次元の場合の SVM の最適化問題のイメージ図

2.2. サポートベクター回帰（Support Vector Regression; SVR）

SVR は、$\varepsilon$-不感損失関数という関数を損失関数に用い、カーネル法によって非線形化することで変数間の複雑な関係を表現することができる SVM の回帰分析への拡張である。

回帰関数を$f(\boldsymbol{x}_i)=\boldsymbol{x}_i\boldsymbol{w}+b$とすると、SVR の損失関数は以下の$\varepsilon$-不感損失関数$f_\varepsilon()$を用いて以下のように表せる：

$$\varepsilon\text{ -不感損失関数：}f_\varepsilon(y,f(\boldsymbol{x}))=\max\{0,|y-f(\boldsymbol{x})|-\varepsilon\}$$

$$\text{SVR の損失関数：}l(w,b)=\frac{1}{2}||\boldsymbol{w}||_2{}^2+C\sum_{i=n}^{N}f_\varepsilon(y_i,f(\boldsymbol{x}_i))$$

ここで、第 1 項は正則化項であり、$C$、$\varepsilon$ はそれぞれペナルティの強さ、誤差の許容範囲を

調整するハイパーパラメータである。損失関数の式を見るとε不感損失関数は絶対誤差を用いることで外れ値の影響を抑え、ε以下の誤差を無視することで過学習を防止するというアイディアからなる損失関数であるとわかる。また、SVM ではマージンが1以下のデータ点の損失を0と考えたが、SVR では、予測値と実測値の差がε以下であるデータ点の損失は0となり、SVM のアイディアを回帰問題に適用したものと解釈することができる。

ε-不感損失関数は、絶対値を場合分けにより外すと以下のように書くことができる。

$$l_\varepsilon\big(y, f(\boldsymbol{x})\big) = \max\{0, -(y - f(\boldsymbol{x})) - \varepsilon\} + \max\{0, y - f(\boldsymbol{x}) - \varepsilon\}$$

２つのスラック変数 $\xi_i^+$と$\xi_i^-$( $i \in \{1,2,\dots,N\}$)を導入すると、以下のように最適化問題を書き直すことができる:

$$\min_{\boldsymbol{w},b,\xi_i^+,\xi_i^-} \ \frac{1}{2}||\boldsymbol{w}||_2{}^2 + C\sum_{i=1}^{N}(\xi_i^+ + \xi_i^-)$$

$$s.t. \quad \xi_i^- \geq -\big(y - f(\boldsymbol{x})\big) - \varepsilon \ , \xi_i^- \geq 0 \ , \xi_i^+ \geq y - f(\boldsymbol{x}) - \varepsilon \ , \xi_i^+ \geq 0, \forall \ i \in \{1,2,\dots,N\}$$

この最適化問題は、SVM と同様にラグランジュの未定乗数法を用いて解くことができる。

SVR は、カーネル法の採用により以下の点で MLR よりも優れていると言うことができる:

- 非線形なカーネル関数により複雑な非線形関係を表現することができる（詳細は付録２.カーネル関数による非線形化を参照）
- 変数間の交絡関係をモデルに組み込むことができる
- 多重共線性を回避することができる

データ数を超える数の説明変数を用いた解析を行うことができる

MLR に比べ様々な点で優位と考えられる SVR だが、カーネル法の採用による多くの利点がある反面カーネル関数の複雑さにより可読性に欠けると言う欠点がある。そのため説明変数空間が大きく、それぞれの変数の目的変数への影響度を調べる問題への適用は難しいと考えられていた。そこで、SVR にある種の可読性を持たせようと考えられたのが本研究で注目する感度分析法である。

### 2.3. SVR 感度分析法

SVR は、複雑な非線形関係を表現できる反面可読性にかけるという欠点があることは前述の通りだが、各説明変数に対して「感度」と言う量を導入し、感度の大小により説明変数の影響度の評価や説明変数の選択を可能にしたのがこの感度分析法（田辺・鈴木(2020)）である。感度分析法は、以下の手順で各説明変数の感度を計算し、感度の値の小さいものを順次消去していくことで影響要因を特定する。

はじめに、以下の手順で各説明変数の感度を計算する。

感度の計算方法：

① 通常のデータ（データセット A とする）を用いてモデルを学習する

② 当該変数の値は実際のものを、それ以外の説明変数の値は全データの平均値に固定したデータセット（データセット B とする）を作る
③ ①で学習したモデルにデータセット B を入力し、出力値を求める
④ 当該変数の実測値を説明変数、上で求めた出力値を目的変数とする線形単回帰を行い、回帰直線の傾きを「感度」とする

以上の計算を全ての説明変数について行い、それぞれの変数に関して感度を求める。つまり感度とは、「ある説明変数の値の変化により、モデルの出力値がどの程度変化するか」を測る量と解釈することができる。感度の計算にあたり、説明変数のオーダーが異なると感度の値が大きく変わってきてしまい意味をなさないので、説明変数は全て同じスケールに正規化して分析を行う（本研究では最小 0、最大 1 に正規化して分析を行った）。感度を用いて、以下のアルゴリズムで変数を選択し最適なモデルを決定する。

変数選択のアルゴリズム：

① モデルを学習する
② モデルに説明変数を入力し、実測値と予測値を用いてモデルの精度指標（平均二乗誤差等）を計算する
③ それぞれの説明変数について感度を計算し、感度の絶対値が最小の説明変数を消去する
④ ①に戻る

以上のプロセスを繰り返しモデルの精度指標で最も良い結果となったモデルを採用し、その時の説明変数を影響要因とする。この操作の流れを図 2 に示す。感度分析法を採用することにより、入手可能な全ての説明変数を採用した非線形な分析をすることができる。

また、得られた最終的に残った変数の感度の値を$\lambda_1, \lambda_2, \ldots, \lambda_m$としたとき、全体から見た相対的な変数の影響度として、$i$番目の感度の寄与率を以下の式で定義する：

$$C_i = \frac{|\lambda_i|}{\sum_{i=1}^{m} |\lambda_i|} \times 100(\%)$$

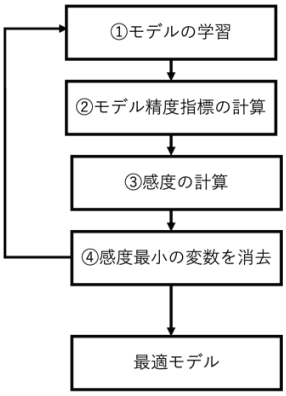



図 2 ：感度分析法による変数選択アルゴリズムの流れ

## ３．提案手法（$\nu$-SVR 感度分析法）

MLR の様々な問題点を改善した手法である SVR の感度分析法だが、実装する中で変数消去の都度行うハイパーパラメータ探索において後述するような非効率性の問題があることがわかった。この非効率性の解消のために、本研究では、SVR において最も重要なハイパーパラメータである不感度$\varepsilon$の扱いを変え、ハイパーパラメータの表す意味が直感的にわかりやすいよう改善された拡張モデルであるニュー・サポートベクター回帰$\nu$-SVR（Scholkopf et al.2000）を用いた感度分析法を提案する。

### 3.1. $\nu$-SVR の導入

通常の SVR では、誤差を無視する幅である不感度$\varepsilon$をハイパーパラメータとして扱ったが、これを変数（学習の際に最適化するパラメータ）として扱うのが$\nu$-SVR の基本的なアイデアである。不感度$\varepsilon$を変数として扱うために、ハイパーパラメータ$\nu$を導入し、$\nu$を介して不感度$\varepsilon$がモデルの複雑さ（正則化項$||\boldsymbol{w}||_2{}^2$）と学習データの損失とのトレードオフを制御するような以下の制約付き最適化問題を解くことで回帰関数$f()$を推定する。

$$\min_{\boldsymbol{w},\varepsilon,\xi_i^+,\xi_i^-} \ \frac{1}{2}||\boldsymbol{w}||_2{}^2 + C(\nu\varepsilon + \frac{1}{N}\sum_{i=1}^{N}(\xi_i^+ + \xi_i^-))$$

$$\begin{aligned} s.t. \quad & \xi_i^- \geq -\big(y - f(\boldsymbol{x})\big) - \varepsilon, \\ & \xi_i^- \geq 0, \\ & \xi_i^+ \geq y - f(\boldsymbol{x}) - \varepsilon, \\ & \xi_i^+ \geq 0, \\ & \varepsilon > 0 \end{aligned}$$

ハイパーパラメータ$\nu$は、サポートベクトル数の学習データ数に対する割合の下限となる。$\nu$-SVR のパラメータ推定のための最適化問題の定式化とカーネル法による非線形化については、付録を参照されたい。

### 3.2. 感度分析法に適用するメリット

SVR において、$\varepsilon$は SVM におけるマージンと解釈することができ、モデルの複雑さを決める重要なハイパーパラメータである。通常の SVR ではこの$\varepsilon$を直接調整するため、比較的大きな範囲と小さい変動幅でグリッドサーチにより最適な値を探索することとなる。一方$\nu$-SVR を用いる場合、$\varepsilon$ではなく「サポートベクトルとなるデータ数の割合の下限」である$\nu$を調整する。$\nu$はサポートベクトル数の割合なので変動範囲は$[0,1]$であり、現実的なモデルの複雑さを考慮すれば、調整すべき範囲を通常の SVR における$\varepsilon$に比べ狭めることができる。表 1 で通常の SVR と$\nu$-SVR それぞれの特徴を示す。

表 1：通常の SVR と $\nu$ -SVR の比較

| | SVR | $\nu$ -SVR |
|---|---|---|
| ハイパーパラメータ | $\varepsilon$ | $\nu$ |
| ハイパーパラメータの意味 | マージン<br>（$\varepsilon$以下の誤差は0とみなされる） | サポートベクトル数のデータ数に対する割合の下限 |
| ハイパーパラメータの変動範囲 | $\varepsilon \geq 0$<br>（目的変数の値のスケールに依存） | $\nu \in [0,1]$ |

このような特徴のため、ハイパーパラメータ探索の容易さにおいて両者には大きな違いが生ずる。実際、ハイパーパラメータの探索範囲をひとつに決めて、通常の SVR で感度分析法を行うと、変数消去過程の後半で最適なハイパーパラメータの値が見つからず、適切なモデルの学習が行われない場合があることがわかった。これは、変数消去の過程で探索範囲を適宜再設定する必要性を示唆しているが、計算プロセスの自動化の障害になり効率性が大きく損なわれる。一方で、$\nu$ -SVR ではこの問題は解消され、探索範囲を再設定しない自動化された計算プロセスでもハイパーパラメータ探索は順調に進み、結果的により自然な変数選択と高い予測精度が得られる。

このことを具体的に見るために、次節の設定の下で、変数消去数（横軸）に応じて、通常の SVR の感度分析法において探索範囲を固定(次節表 3 参照)したグリッドサーチの結果採用されたハイパーパラメータ $\varepsilon$ の値の推移（右）と、同じく $\nu$ -SVR の感度分析法において採用された $\nu$ の値の推移（左）を図 3 で示す。このグラフから、$\nu$ -SVR と違って、通常の SVR では変数消去過程の前半にくらべて後半のハイパーパラメータ $\varepsilon$ の変動域が著しく縮小して下限付近に張り付いており、スケールを見直して探索範囲を再設定すべきであることがわかる。

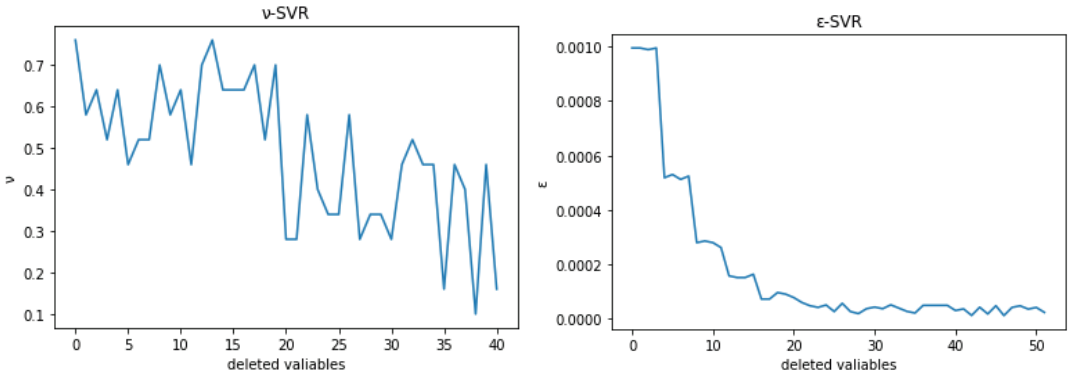



図 3：感度分析法でのグリッドサーチにおいて決定された $\varepsilon$ と $\nu$ の推移

## ４．データ分析

本節では、ν-SVR の感度分析法の使用例として、都道府県間の死亡率格差の要因分析を行う。

少子高齢化が深刻な問題となっている現在の日本だが、次なる問題として多死社会が危惧されている。それに伴い、医師不足、終末期医療など様々な社会問題の発生が予想され、国は地域自治体と協力して対策に取り組んでいる。近年、人の健康は遺伝子等の先天的な要因だけではなく、経済状況、労働環境などの後天的な要因の影響も受けるということが広く知られており、国や地方自治体が健康政策に取り組むに当たって、何が強く健康と結びついているのかを理解することの重要度は増していると言える。

本研究では、都道県別の社会的環境・生活習慣等の様々な指標を説明変数、全死因平均死亡率を目的変数とした回帰分析を行い、各説明変数の影響度指標の値の解釈及び予測の精度によりモデルの評価を行う。説明変数の影響度指標とは、例えば SVR の場合は感度の値と感度の最終的な寄与率、MLR の場合は回帰係数など説明変数の目的変数への影響度を表す指標のことを指す。このような分析を行うことにより、地域間で死亡率格差をうむ要因を特定するとともに政府や地域自治体の健康政策の手がかりとなるような情報の獲得を目指す。また、ν-SVR の比較対象として、通常の SVR を用いた感度分析法や他の様々な機械学習手法にも同じデータを適用した分析を行い、結果を考察することでν-SVR の感度分析法の有効性を検証する。

### 4.1. データ

本研究で使用するデータは先行研究を踏襲し以下の通りとした。観測年度に若干のズレは存在しているが環境要因は短期間には変動しにくいと考えられることから許容可能と判断した。

- 総務省統計局 都道府県・市区町村のすがた（社会・人口統計体系）2014 年度　2015 年度
- 厚生労働省　国民生活基礎調査 2013 年度　2016 年度
- 厚生労働省　国民健康・栄養調査 2012 年度

#### ① 目的変数

本研究では、都道府県ごとの健康の程度を表す目的変数として、全死因・年齢調整死亡率の男女平均を用いる。一般的に死亡率の水準は男女で異なるが、図 4 に示すように、男女の平均死亡率の都道府県別の傾向は概ね類似しており、政府や地域自治体が多死社会対策を検討するに当たっての参考情報を提供することが本データ分析の目的なので、簡潔さを重

視しこのデータを目的変数に採用した。

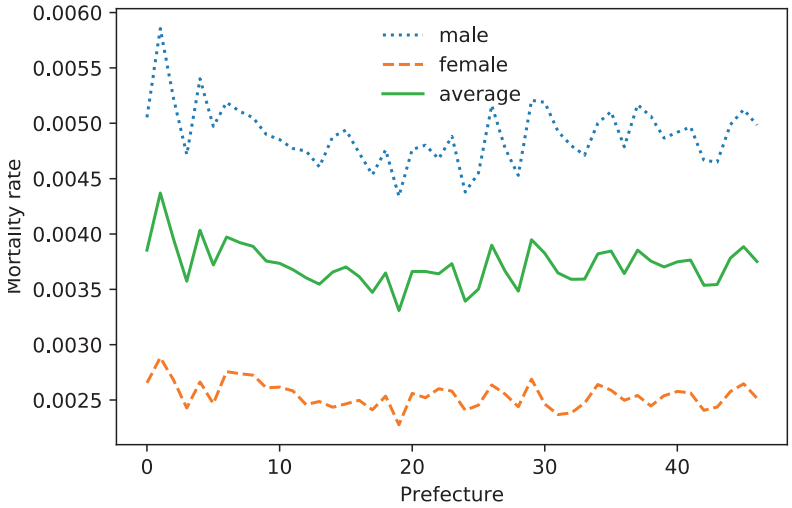



図４：都道府県別全死因年齢調整死亡率の男性・女性・男女平均のプロット

② 説明変数

本研究では、健康により強く影響する要因を特定し健康政策に役立てることが目的であるので、「健康に影響する可能性があり、かつ政策等により改善可能性があるもの」という基準のもとなるべく多くの指標を説明変数に採用した。健康に影響する指標としては、飲酒・喫煙など直接的に影響するものが考えられることが多いが、近年は社会・経済的要因も健康に強く影響すると考えられるようになってきているということから、一見健康との関連が薄いようにも思える指標の中にも新たな発見の可能性があるのではないかと考え積極的に取り入れた。以上の基準から、健康に影響する可能性のある指標として、(1)環境要因　(2)生活習慣要因　の２つに分類される合計４２種類の指標を説明変数として採用した。

データは、総務省統計局が実施している「社会人口統計体系　統計でみる市区町村のすがた」及び厚生労働省が実施している「国民生活基礎調査」、「国民健康栄養調査」、「特定健康診査」から取得した。全ての変数の定義、出典、調査年度及び目的変数との相関係数を表２に示す（注：表中の出典の略号は、社人…社会人口統計体系、国生…国民生活基礎調査、国健…国民健康・栄養調査、特健…特定健康診査）。

表 2：変数の定義・出典及び目的変数との相関係数

| | 分類 | 変数名 | 定義 | 相関係数 | 出典 | 調査年度 |
|---|---|---|---|---|---|---|
| 目的変数 | | 死亡率 | 全死因・年齢調整死亡率の男女平均 | - | 地域別 | 2015 |
| 説明変数 | 環境要因 | 核家族 | 核家族世帯割合 | -0.1571231 | 社人 | 2015 |
| | | 高齢単身 | 高齢単身世帯割合 | 0.32686275 | 社人 | 2015 |
| | | 婚姻 | 婚姻率 | -0.2529874 | 社人 | 2015 |
| | | 収入 | 世帯実収入平均 | -0.185917 | 社人 | 2015 |
| | | 収入格差 | 年間収入ジニ係数 | 0.18278988 | 社人 | 2014 |
| | | 世帯教育費 | 教育費割合（世帯） | -0.2519892 | 社人 | 2015 |
| | | 自治体教育費 | 教育費割合（自治体） | -0.3328989 | 社人 | 2015 |
| | | 人口集中 | 人口集中地区人口割合 | -0.1997187 | 社人 | 2015 |
| | | 労働災害 | 労働災害発生頻度 | 0.26471869 | 社人 | 2015 |
| | | 自殺 | 人口あたりの自殺件数 | 0.48863432 | 社人 | 2015 |
| | | 病床 | 人口あたりの一般病院病床 | 0.15315762 | 社人 | 2015 |
| | | 精神病床 | 人口あたりの精神病床 | 0.29127306 | 社人 | 2015 |
| | | 交通事故 | 人口あたりの交通事故死傷者数 | -0.1093242 | 社人 | 2015 |
| | | 公害 | 人口あたりの公害苦情件数 | -0.32377 | 社人 | 2015 |
| | | 保健師 | 人口あたりの保健師数 | -0.022624 | 社人 | 2014 |
| | | 社会福祉費 | 自治体の社会福祉費割合 | -0.158363 | 社人 | 2015 |
| | | 児童福祉施設 | 人口あたりの児童福祉施設数 | 0.06270742 | 社人 | 2015 |
| | | 高齢就業 | 高齢者就業割合 | -0.199635 | 社人 | 2015 |
| | | 養護老人ホーム | 高齢人口あたりの養護老人ホーム数 | -0.0237654 | 社人 | 2015 |
| | | 有料老人ホーム | 高齢人口あたりの有料老人ホーム数 | 0.23783781 | 社人 | 2015 |
| | | 介護療養型医療施設 | 高齢人口あたりの介護療養型医療施設 | -0.0193254 | 社人 | 2015 |
| | | 老人福祉センター | 高齢人口あたりの老人福祉センター数 | 0.07440065 | 社人 | 2015 |
| | | 介護老人福祉施設 | 高齢人口あたりの介護老人福祉施設数 | 0.19264803 | 社人 | 2015 |
| | | 高齢者学級 | 高齢人口あたりの高齢者学級・講座数 | 0.15059765 | 社人 | 2014 |
| | 生活習慣要因 | 喫煙 | 喫煙者の割合 | 0.4697738 | 国生 | 2013 |
| | | 飲酒 | 飲酒者の割合 | 0.20304369 | 国生 | 2013 |
| | | スポーツ | スポーツ行動者割合 | -0.5964439 | 社人 | 2011 |
| | | 趣味・娯楽 | 趣味・娯楽の平均時間 | -0.4338787 | 社人 | 2011 |
| | | 睡眠時間 | 一日の平均睡眠時間 | 0.45395314 | 国生 | 2013 |
| | | 歩数 | 一日の平均歩数 | -0.2881213 | 国健 | 2012 |
| | | 肥満 | BMIの平均値 | 0.33341286 | 国健 | 2012 |
| | | メタボ | メタボリックシンドローム該当者の割合 | 0.21173273 | 特健 | 2015 |
| | | 健診 | 人口あたりの健診受診率 | -0.2573411 | 国生 | 2013 |
| | | がん検診 | ５種のがん検診の平均受診率 | -0.1595202 | 国生 | 2015 |
| | | 健康無関心 | 日頃健康のために何もしていない人の割合 | 0.5296948 | 国生 | 2013 |
| | | 悩み・ストレス | 悩みやストレスのある人の割合 | -0.3900019 | 国生 | 2013 |
| | | 野菜摂取 | 一日の平均野菜摂取量 | -0.1374759 | 国健 | 2012 |
| | | 食塩摂取 | 一日の平均食塩摂取量 | -0.0032192 | 国健 | 2012 |
| | | 規則正しい食事 | 規則正しく食事している人の割合 | -0.2537301 | 国生 | 2013 |
| | | バランスのとれた食事 | バランスのとれた食事をしている人の割合 | -0.5421122 | 国生 | 2013 |
| | | 薄味の食事 | 薄味の食事をしている人の割合 | 0.02905257 | 国生 | 2013 |
| | | 食べ過ぎない意識 | 食べ過ぎないようにしている人の割合 | -0.1364806 | 国生 | 2013 |

### 4.2. ベンチマークとなる機械学習手法

本研究では、$\nu$-SVR の感度分析法の有効性を示すために、SVR、$\nu$-SVR 以外の機械学習手法にも同じデータを適用して分析を行い、結果から各手法の特性及び有効性の比較分析を行う。比較対象には、様々な形式の技術と網羅的に比較するために、古くから使われている伝統的な手法から比較的新しい手法まで分析することを考え、以下の３つを採用した：

- 線形重回帰（Multiple Linear Regression: MLR）
- 部分最小２乗回帰（Partial Least Squared Regression: PLSR）

● ランダムフォレスト（Random Forest: RF）

最も基本的な手法ある MLR だが、説明変数が多くなると多重共線性が生じることが知られており、本研究で扱うデータのように特に説明変数が多い場合全く解釈不能な結果しか得られないということも考えられるため、損失関数に正則化項を加えて分析を行った。正則化項は、一般に回帰係数の L1 ノルムの定数倍を加える場合（Lasso 回帰）と、L2 ノルムの２乗の定数倍を加える場合（Ridge 回帰）があるが、本研究ではその両方を採用し比較した。正則化を用いることで多重共線性のある程度の緩和は見込めるが、説明変数が多くなるとどうしてもいくつかの変数で解釈し難い結果が得られることが多い。そこで、多重共線性を回避するために、あらかじめ説明変数の中で相関が強いものを消去した状態で分析に臨むという方法が度々採用される。この方法は、本来目的変数に影響しているにも関わらず、他に自身と相関が強い変数があるという理由で消去されてしまう変数が出てきてしまう可能性があるため、結果への信頼度は全ての変数を採用する場合と比べて欠けるという問題点がある。しかしこの方法は、多重共線性の回避には繋がり予測精度は向上すると考えられるので、本研究では、MLR での分析に関して全ての説明変数を用いる場合と、説明変数同士の相関係数が 0.7 以上のものは目的変数との相関係数がより小さい方の変数を事前に消去したデータセットを用いる場合の２通りの分析を行った。

PLSR は、主成分分析のアイディアを用いて多重共線性の回避を目的として開発された手法である。通常の主成分分析によって得られる主成分に対して MLR を適用する主成分回帰（Principal Component Regression: PCR）に比べ、目的変数との相関を考慮した潜在変数空間で回帰を行うため、より高精度な分析が可能となる。

また、本研究では回帰木を用いたアンサンブル学習の手法であるランダムフォレスト（Random Forest: RF）も比較対象として採用した。回帰木は、回帰関数の形を定義しないノンパラメトリックな手法であり、MLR や SVR など関数の形を仮定する手法とどういった結果の違いを示すかはとても興味深く、発見が得られる可能性があるので比較対象として採用した。

これらの手法及び SVR、$\nu$-SVR の感度分析法で同じ都道府県別データを解析し、結果を分析する。SVR のカーネル関数は、一般的に使われる以下の４つのカーネル関数全てを変数消去前のデータに適用し、１０個抜き交差検証法による MSE（平均二乗誤差）の値が最も小さいものを採用した（その際のハイパーパラメータの値は SVM のパッケージ LIBSVM のデフォルトの値を使用した）：

線形カーネル：$K(\boldsymbol{x}_i, \boldsymbol{x}_j) = \boldsymbol{x}_i{}^T \boldsymbol{x}_j$

多項式カーネル：$K(\boldsymbol{x}_i, \boldsymbol{x}_j) = (\boldsymbol{x}_i{}^T \boldsymbol{x}_j + c)^d$

同型基底カーネル（RBF カーネル）：$K(\boldsymbol{x}_i, \boldsymbol{x}_j) = exp(-g||\boldsymbol{x}_i - \boldsymbol{x}_j||_2{}^2)$

シグモイドカーネル：$K(\boldsymbol{x}_i, \boldsymbol{x}_j) = \tanh(g\boldsymbol{x}_i{}^T \boldsymbol{x}_j + c), \tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$

※$c, d, g$はハイパーパラメータ

計算の結果、SVR と $\nu$ -SVR 両方において最も MSE の値が小さかった同型基底カーネルを採用した。各説明変数の影響度指標は、SVR の感度分析法の場合は感度、MLR 及び PLSR の場合は回帰係数、RF の場合は特徴量重要度を使用する。予測精度の評価には、本研究は都道府県別のデータを採用しておりデータ数が少ないことを考慮し、１個抜き交差検証法による MSE（平均二乗誤差）を用いた。モデルの実装には、Python における機械学習パッケージ（SVR 及び $\nu$ -SVR：LIBSVM-3.24　その他：scikit-learn 0.24.0）を用いた。モデルにおける全てのハイパーパラメータは、グリッドサーチにおいて１０分割交差検証法による MSE を最小化するものを採用した。表 3 は、グリッドサーチにより探索したハイパーパラメータと探索のグリッド幅、範囲及び得られた最適な値を示す。なお、SVR 感度分析法は変数消去のたびにハイパーパラメータを更新するため値の表記を省略したが、変数消去過程での $\varepsilon$ と $\nu$ の値の推移については前節の図 3 のグラフを参照されたい。

表 3：各ハイパーパラメータとグリッドサーチのグリッド幅・範囲及び得られた値

| 手法 | ハイパーパラメータ | グリッド幅・範囲 | 得られた値 |
|---|---|---|---|
| Ridge | 正則化項の係数$\lambda$ | $\{2^{-19}, 2^{-18}, \dots, 2^{-1}\}$ | $\lambda = 2^{-18}$ |
| Ridge 変数消去後 | 正則化項の係数$\lambda$ | $\{2^{-19}, 2^{-18}, \dots, 2^{-1}\}$ | $\lambda = 2^{-19}$ |
| Lasso | 正則化項の係数$\lambda$ | $\{2^{-19}, 2^{-18}, \dots, 2^{-1}\}$ | $\lambda = 2^{-3}$ |
| Lasso 変数消去後 | 正則化項の係数$\lambda$ | $\{2^{-19}, 2^{-18}, \dots, 2^{-1}\}$ | $\lambda = 2^{-2}$ |
| PSLR | 採用する主成分の数$M$ | $\{1,2, \dots, 42\}$ | $M = 5$ |
| RF | 木の数$n_estimators$ | $\{10,20, \dots, 100\}$ | $n_estimators = 30$ |
| | 分割の最適化の際の考慮する説明変数の割合$n_features$ | $\{0.1,0.2, \dots, 1.0\}$ | $max_features = 0.5$ |
| | ブートストラップ法において１つのデータ群のデータ割合$\mathrm{max}_samples$ | $\{0.1,0.2, \dots, 1.0\}$ | $max_samples = 0.6$ |
| SVR | カーネル関数RBFの係数$g$ | $\{2^{-19}, 2^{-18}, \dots, 2^{-6}\}$ | - |
| | 正則化項の係数$C$ | $\{2^{0}, 2^{0.2}, \dots, 2^{2.4}\}$ | - |
| | $\varepsilon$ -不感損失関数における不感度 $\varepsilon$ | $\{2^{-23}, 2^{-22}, \dots, 2^{-10}\}$ | - |
| $\nu$ -SVR | カーネル関数RBFの係数$g$ | $\{2^{-19}, 2^{-18}, \dots, 2^{-6}\}$ | - |
| | 正則化項の係数$C$ | $\{2^{0}, 2^{0.2}, \dots, 2^{2.4}\}$ | - |
| | サポートベクトルのデータ数に対しての率の下限$\nu$ | $\{0.1,0.16, \dots, 0.82\}$ | - |

### 4.3. 分析結果

分析結果をまとめると以下のようになる。表 3 は各手法により選ばれた変数名（変数選択機能がないものに関しては変数全て）とそれぞれの影響度指標の値を、表 4 は各モデルの MSE の値を示す。

## 表 4：各手法で選択された説明変数とその影響度指標の値

ν-SVR

| 変数 | 感度 | 感度の寄与率 |
|---|---|---|
| 健康無関心 | 0.0000724403 | 18.54% |
| 喫煙 | 0.0000432268 | 11.07% |
| 睡眠時間 | 0.0000237835 | 6.09% |
| 肥満 | 0.0000153395 | 3.93% |
| がん検診 | -0.0000094391 | 2.42% |
| 歩数 | -0.0000110632 | 2.83% |
| 婚姻 | -0.0000148872 | 3.81% |
| 収入 | -0.0000225245 | 5.77% |
| 病床 | -0.0000239422 | 6.13% |
| 趣味・娯楽 | -0.0000255150 | 6.53% |
| 精神病床 | -0.0000342187 | 8.76% |
| バランスのとれた食事 | -0.0000417481 | 10.69% |
| スポーツ | -0.0000525091 | 13.44% |

SVR

| 変数 | 感度 | 感度の寄与率 |
|---|---|---|
| 社会福祉費 | 0.000242355 | 14.83% |
| がん検診 | 0.000200803 | 12.29% |
| 健康無関心 | 0.000130650 | 8.00% |
| 歩数 | 0.000102463 | 6.27% |
| 睡眠時間 | 0.000093983 | 5.75% |
| 喫煙 | 0.000092079 | 5.64% |
| 婚姻 | 0.000050913 | 3.12% |
| 病床 | 0.000027796 | 1.70% |
| 精神病床 | 0.000016464 | 1.01% |
| 肥満 | 0.000006036 | 0.37% |
| 自治体教育費 | 0.000005386 | 0.33% |
| 公害 | -0.000057174 | 3.50% |
| スポーツ | -0.000147317 | 9.02% |
| バランスのとれた食事 | -0.000148733 | 9.10% |
| 趣味・娯楽 | -0.000155594 | 9.52% |
| 介護老人福祉施設 | -0.000156223 | 9.56% |

RF

| 変数 | 特徴量重要度 | 変数 | 特徴量重要度 |
|---|---|---|---|
| 自殺 | 18.13% | 収入 | 1.14% |
| スポーツ | 10.53% | 社会福祉費 | 1.12% |
| バランスのとれた食事 | 7.21% | 病床 | 1.08% |
| 高齢単身 | 6.47% | 核家族 | 1.05% |
| 喫煙 | 5.76% | 規則正しい食事 | 0.91% |
| 肥満 | 4.36% | 野菜摂取 | 0.87% |
| 健康無関心 | 4.21% | がん検診 | 0.80% |
| 労働災害 | 4.13% | 高齢就業 | 0.76% |
| 薄味の食事 | 3.81% | 食べ過ぎない意識 | 0.75% |
| 歩数 | 2.53% | 高齢者学級 | 0.74% |
| 自治体教育費 | 2.40% | メタボ | 0.70% |
| 健診 | 2.38% | 婚姻 | 0.66% |
| 人口集中 | 2.29% | 食塩摂取 | 0.60% |
| 収入格差 | 1.90% | 飲酒 | 0.60% |
| 交通事故 | 1.77% | 養護老人ホーム | 0.56% |
| 老人福祉センター | 1.53% | 精神病床 | 0.47% |
| 睡眠時間 | 1.40% | 介護老人福祉施設 | 0.43% |
| 悩み・ストレス | 1.36% | 児童福祉施設 | 0.27% |
| 公害 | 1.34% | 世帯教育費 | 0.21% |
| 有料老人ホーム | 1.24% | 保健師 | 0.19% |
| 介護療養型医療施設 | 1.22% | 趣味・娯楽 | 0.10% |

Ridge

| 変数 | 回帰係数 |
|---|---|
| 薄味の食事 | 0.00211524 |
| 喫煙 | 0.00181822 |
| 収入格差 | 0.00058323 |
| 健康無関心 | 0.00040309 |
| 労働災害 | 0.00008950 |
| がん検診 | 0.00008699 |
| 肥満 | 0.00008628 |
| 高齢単身 | 0.00006544 |
| 社会福祉費 | 0.00003483 |
| 高齢就業 | 0.00002797 |
| 睡眠時間 | 0.00001669 |
| スポーツ | 0.00001212 |
| 趣味・娯楽 | 0.00001024 |
| 自殺 | 0.00000829 |
| 核家族 | 0.00000599 |
| 介護老人福祉施設 | 0.00000205 |
| 収入 | 0.00000025 |
| 精神病床 | 0.00000020 |
| 高齢者学級 | 0.00000008 |
| 歩数 | 0.00000001 |
| 交通事故 | 0.00000001 |
| 病床 | -0.00000026 |
| 有料老人ホーム | -0.00000032 |
| 公害 | -0.00000157 |
| 野菜摂取 | -0.00000178 |
| 養護老人ホーム | -0.00000257 |
| 老人福祉センター | -0.00000272 |
| 人口集中 | -0.00000383 |
| 食塩摂取 | -0.00000776 |
| 介護療養型医療施設 | -0.00000844 |
| 自治体教育費 | -0.00000846 |
| 保健師 | -0.00001186 |
| 婚姻 | -0.00001770 |
| 世帯教育費 | -0.00002038 |
| 児童福祉施設 | -0.00002064 |
| 健診 | -0.00021900 |
| 悩み・ストレス | -0.00038642 |
| 食べ過ぎない意識 | -0.00041177 |
| 飲酒 | -0.00052955 |
| 規則正しい食事 | -0.00132717 |
| メタボ | -0.00165006 |
| バランスのとれた食事 | -0.00269072 |

Ridge 変数消去後

| 変数 | 回帰係数 |
|---|---|
| 喫煙 | 0.00156907 |
| 健康無関心 | 0.00037815 |
| 食べ過ぎない意識 | 0.00036463 |
| 収入格差 | 0.00019293 |
| 自殺 | 0.00007275 |
| 婚姻 | 0.00006216 |
| 高齢単身 | 0.00005397 |
| 児童福祉施設 | 0.00004222 |
| 肥満 | 0.00002421 |
| 世帯教育費 | 0.00002359 |
| 老人福祉センター | 0.00001905 |
| 公害 | 0.00000271 |
| 自治体教育費 | 0.00000186 |
| 収入 | 0.00000112 |
| 交通事故 | 0.00000008 |
| 歩数 | -0.00000005 |
| 高齢者学級 | -0.00000008 |
| 介護老人福祉施設 | -0.00000021 |
| 精神病床 | -0.00000028 |
| 野菜摂取 | -0.00000113 |
| 有料老人ホーム | -0.00000186 |
| 養護老人ホーム | -0.00000395 |
| 睡眠時間 | -0.00001642 |
| 高齢就業 | -0.00001731 |
| スポーツ | -0.00004712 |
| 労働災害 | -0.00007249 |
| 健診 | -0.00025986 |
| 薄味の食事 | -0.00026075 |
| 飲酒 | -0.00034947 |
| メタボ | -0.00038212 |
| 悩み・ストレス | -0.00081066 |
| 規則正しい食事 | -0.00097635 |

Lasso

| 変数 | 回帰係数 |
|---|---|
| 高齢単身 | 0.00009833 |
| 自殺 | 0.00004860 |
| 喫煙 | 0.00004367 |
| 老人福祉センター | 0.00002649 |
| 収入 | 0.00002178 |
| 肥満 | 0.00001810 |
| 高齢者学級 | 0.00000763 |
| 睡眠時間 | 0.00000668 |
| 収入格差 | 0.00000342 |
| 労働災害 | 0.00000083 |
| 有料老人ホーム | -0.00000075 |
| メタボ | -0.00000335 |
| 飲酒 | -0.00000859 |
| 自治体教育費 | -0.00000981 |
| 規則正しい食事 | -0.00001195 |
| 病床 | -0.00001901 |
| 悩み・ストレス | -0.00002386 |
| 健診 | -0.00002583 |
| 人口集中 | -0.00002710 |
| 養護老人ホーム | -0.00002953 |
| スポーツ | -0.00003245 |
| バランスのとれた食事 | -0.00004406 |
| 介護療養型医療施設 | -0.00004522 |
| 保健師 | -0.00008981 |

Lasso 変数消去後

| 変数 | 回帰係数 |
|---|---|
| スポーツ | 0.00008431 |
| 高齢単身 | 0.00003742 |
| 自殺 | 0.00003172 |
| 喫煙 | 0.00002165 |
| 悩み・ストレス | 0.00001726 |
| 保健師 | 0.00000324 |
| 肥満 | 0.00000321 |
| 健診 | -0.00000700 |
| 収入格差 | -0.00000813 |
| メタボ | -0.00001027 |
| 養護老人ホーム | -0.00001032 |
| 自治体教育費 | -0.00001091 |
| 老人福祉センター | -0.00001440 |
| 健康無関心 | -0.00002094 |
| 収入 | -0.00004303 |
| 飲酒 | -0.00006245 |
| バランスのとれた食事 | -0.00008546 |
| 人口集中 | -0.00008667 |

PLSR

| 変数 | 回帰係数 |
|---|---|
| 喫煙 | 0.000058855 |
| 高齢就業 | 0.000056707 |
| 野菜摂取 | 0.000032001 |
| 労働災害 | 0.000029185 |
| 健康無関心 | 0.000028645 |
| 食塩摂取 | 0.000027868 |
| 児童福祉施設 | 0.000023375 |
| バランスのとれた食事 | 0.000018282 |
| 病床 | 0.000018224 |
| 自治体教育費 | 0.000015856 |
| 公害 | 0.000015447 |
| 介護老人福祉施設 | 0.000015287 |
| 世帯教育費 | 0.000014275 |
| 社会福祉費 | 0.000013933 |
| 食べ過ぎない意識 | 0.000012733 |
| 薄味の食事 | 0.000011699 |
| 精神病床 | 0.000010129 |
| 介護療養型医療施設 | 0.000007109 |
| 有料老人ホーム | 0.000003684 |
| メタボ | 0.000001526 |
| 自殺 | -0.000000029 |
| がん検診 | -0.000003114 |
| 収入 | -0.000004557 |
| 歩数 | -0.000005996 |
| 趣味・娯楽 | -0.000006410 |
| スポーツ | -0.000006616 |
| 養護老人ホーム | -0.000007626 |
| 人口集中 | -0.000009141 |
| 収入格差 | -0.000009853 |
| 高齢者学級 | -0.000015649 |
| 悩み・ストレス | -0.000015686 |
| 保健師 | -0.000016339 |
| 飲酒 | -0.000021298 |
| 睡眠時間 | -0.000021548 |
| 高齢単身 | -0.000022651 |
| 肥満 | -0.000026395 |
| 規則正しい食事 | -0.000027021 |
| 健診 | -0.000027489 |
| 老人福祉センター | -0.000032035 |
| 交通事故 | -0.000040905 |
| 核家族 | -0.000054262 |
| 婚姻 | -0.000057833 |

※変数は影響度指標値の大きい順にソートされている

## 5. 変数選択結果の考察

### 5.1. 選択された変数についての考察

まず、変数選択の機能のある手法について、選ばれた変数の内容について各手法の特徴を考えてみる。変数選択機能がある手法は、SVR、$\nu$-SVR の感度分析法及び Lasso 正則化を

用いたMLRだが、それぞれの手法で選ばれた説明変数を見てみると、環境要因と生活習慣要因の比率に大きく違いがあることが見て取れる。Lassoで選ばれた２４個の変数のうち生活習慣要因は、喫煙・肥満・睡眠時間・メタボ・飲酒・規則正しい食事・悩みストレス・検診・スポーツ・バランスのとれた食事の１０個であり、Lasso（変数消去後）では選ばれた１８個の変数のうち生活習慣要因は、喫煙・肥満・メタボ・飲酒・悩みストレス・検診・スポーツ・バランスのとれた食事・健康無関心の９個であるのに対し、SVR では選ばれた１６個の変数のうち生活習慣要因は、健康無関心・スポーツ・喫煙・バランスのとれた食事・趣味娯楽・睡眠時間・肥満・歩数・がん検診の９個であり、ν-SVRでは選ばれた１３個の変数のうち生活習慣要因は、SVRで選ばれた変数と全く同じもの９個となり、SVR及びν-SVRはより高い割合で生活習慣要因の変数を選ぶ傾向があった。また、SVRとν-SVRを比較すると、生活習慣要因の変数に関しては全く同じものを選んでいるが、SVRの方がν-SVRよりも多く環境要因の変数を採用するという結果となった。

選ばれた変数の詳細を見てみると、Lasso では老人福祉センター・有料老人ホーム・養護老人ホーム・介護療養型医療施設が、Lasso（変数消去後）では養護老人ホーム・老人福祉センターが選ばれており老人施設を重要視しているのに対し、SVR では老人施設は介護老人福祉施設のみであり、ν-SVRでは高齢者福祉施設を１個も選んでいないという結果になった。全体としてみると、Lassoを使ったMLRでは、高齢者福祉施設等の外部的な要因を、SVR・ν-SVRでは日常生活の中での要因を重要視しているということがわかった。

5.2. 変数の影響度指標の値の符号と大きさについての考察

本研究において最も重要なのが、この影響度指標の符号と大きさに関する分析を行うことである。４２個の説明変数がある中で、各手法がどれだけ多重共線性を回避して変数の重要度を判定しているかを分析することは本研究の要点の１つである。

影響度指標を符号付きで比較することができるSVR、ν-SVR及びMLRを用いた５手法（Ridge・Ridge（変数消去後）・Lasso・Lasso（変数消去後）・PLSR）について選ばれた変数とその影響度指標の符号及び大きさについて変数の本来の意味と照らし合わせて考えてみると、ν-SVRが最も納得感を得られる結果となった。表５は、MLRを用いた手法において明らかに回帰係数の符号が逆だと考えられる変数とその値を示している（健康への影響として色々な説が考えられ判断が微妙なものは除いた）。表５の通り、MLRを用いた手法では現実世界と不整合な符号を示す変数はなくならず、この結果は、正則化や相関の高い変数の消去、無相関な潜在変数空間への写像などを用いても、線形の手法では多重共線性がぬぐいきれないというある種のMLRの限界を示していると考えることができる。

SVRでは、MLRを用いた手法ほどではないにしろ納得し難い結果となった。感度の値の符号については、「公害」の変数でマイナス（死亡率を低下させる働き）となったが、これは現実の問題を考えると逆であると考えられる。また、感度の値の大きさについては、SVRでは符号がプラスのものは値が大きい順に社会福祉費、がん検診、健康無関心、歩数、睡眠

時間、喫煙、…となっているが、直接的な健康への悪影響があることが知られている喫煙よりも歩数、睡眠などの変数が大きな値を示していることに違和感がある。

一方ν-SVR では、現実世界と明らかに不整合な符号を示す変数は見当たらず、全ての変数について感度の符号、値の両方についてある程度の納得感が得られる結果となっている。

表 5：MLR での結果において明らかに回帰係数の符号が逆な変数とその回帰係数の値

Ridge

| 変数 | 回帰係数 |
|---|---|
| 薄味の食事 | 0.00211524 |
| スポーツ | 0.00001212 |
| 趣味・娯楽 | 0.00001024 |
| 収入 | 0.00000025 |
| 公害 | -0.00000157 |
| 食塩摂取 | -0.00000776 |
| 悩み・ストレス | -0.00038642 |
| 飲酒 | -0.00052955 |
| メタボ | -0.00165006 |

Ridge 変数消去後

| 変数 | 回帰係数 |
|---|---|
| 食べ過ぎない意識 | 0.00036463 |
| 労働災害 | -0.00007249 |
| 飲酒 | -0.00034947 |
| メタボ | -0.00038212 |
| 悩み・ストレス | -0.00081066 |

Lasso

| 変数 | 回帰係数 |
|---|---|
| メタボ | -0.00000335 |
| 悩み・ストレス | -0.00002386 |

Lasso 変数消去後

| 変数 | 回帰係数 |
|---|---|
| スポーツ | 0.00008431 |
| メタボ | -0.00001027 |
| 健康無関心 | -0.00002094 |
| 飲酒 | -0.00006245 |

PLSR

| 変数 | 回帰係数 |
|---|---|
| 野菜摂取 | 0.000032001 |
| バランスのとれた食事 | 0.000018282 |
| 食べ過ぎない意識 | 0.000012733 |
| 薄味の食事 | 0.000011699 |
| 自殺 | -0.000000029 |
| 悩み・ストレス | -0.000015686 |
| 飲酒 | -0.000021298 |
| 肥満 | -0.000026395 |
| 交通事故 | -0.000040905 |

### 5.3. RF での分析結果に関する考察

RF では、特徴量重要度の値の大小を用いて説明変数の目的変数への影響の程度を見ることができる。各説明変数の特徴量重要度の大きさを考えて見ると、RF も SVR と同様に、生活習慣要因が多く上位に来ており、特徴量重要度の大きさと変数の関係は MLR を用いた場合よりも納得感のある結果が得られた。

しかし RF における特徴量重要度の値は「ある変数のモデルの予測精度への影響度」を表す指標であり、MLR における回帰係数のように符号で目的変数に与える影響の方向の判断ができず、想定通りの影響の仕方をしているのかという疑問が残る。また RF は、ブートストラップ法の採用や木ごとのランダムな説明変数選択など乱数を用いた処理が多いため、用いる回帰木の個数が少ない場合、実行のたびに結果が大きく変わることが確認された。

この乱数処理に関する問題は、木の個数を増やし、例えば 1000 以上で固定することによって解消される可能性がある。しかしながら、特にグリッドサーチにおいて筆者の計算環境では負荷が重くなりすぎることと、特徴量重要度の符号が不明となる問題は残ることから、今回は対応を見送った。

### 5.4. 予測誤差に関する考察

表 6 は、各手法の学習済みモデルに対する 1 個抜き交差検証法を用いた MSE の値を示す。予測精度としては、精度が高い方からν-SVR、Lasso（変数消去後）、SVR、Lasso、PLSR、RF、Ridge（変数消去後）、Ridge という順になった。予測精度の観点からも、ν-SVR が最も誤差が小さく有効であるという結果となった。

表６：各手法のモデルの MSE

| 手法 | $\nu$-SVR | SVR | Ridge | Ridge 変数消去後 | Lasso | Lasso 変数消去後 | PLSR | RF |
|---|---|---|---|---|---|---|---|---|
| MSE | $9.870\times10^{-9}$ | $1.324\times10^{-8}$ | $4.510\times10^{-8}$ | $3.116\times10^{-8}$ | $1.306\times10^{-8}$ | $1.279\times10^{-8}$ | $1.768\times10^{-8}$ | $2.532\times10^{-8}$ |

## 6. まとめ

本研究では、SVR の感度分析法を行う際のハイパーパラメータ探索の非効率性解消を目指して $\nu$-SVR を導入し、その手法を都道県別の平均死亡率と環境・生活習慣要因の指標のデータに適用し、他の様々な機械学習手法と比較し結果の分析を行うことで有効性を検証した。分析結果としては、本研究で採用したデータに対しては選択された変数に対する納得感、予測精度の両面において提案手法が最も有効であるという結論を得た。

ただし、本研究で比較した各手法において、ハイパーパラメータの探索方法によって、どの程度ハイパーパラメータが変化し、どの程度予測値に影響を与えるかは重要な論点であるが、計算負荷も大きいため、今回は検討できていない。しかしながら、図３を見る限り、グリッドサーチの設定によって $\nu$-SVR は影響を受けにくく、通常の SVR は影響を受けやすい可能性が示唆されている。これらの検証は今後の課題である。

今後の発展の方向性としては、様々な種類のデータに対する適用が考えられる。本研究では SVR の感度分析法の概説と $\nu$-SVR による拡張手法の提案が主な目的だったため、データ分析ではオープンなデータでかつ直感的な理解が得られやすい問題を扱った。本研究で扱ったテーマは説明変数よりもデータ数が多い状況であった（データ数４７に対して説明変数４２）が、SVR はデータ数以上の説明変数を使うことができるというスパース推定において重要な性質を持っている。よって、従来の手法では扱いが難しいとされてきた説明変数の数がさらに膨大な問題に対しても有効な分析結果が得られる可能性がある。テーマの例としては、特定の病気・腫瘍の発生と特定の遺伝子の保有の関係を調べるなどの Bio Informatics（生命情報科学）の分野や、損害保険のクレームデータに対しクレーム額・件数と各クレームの特徴の関係を調べるなどの損害保険数理の分野等が挙げられる。そういった目的変数と説明変数の関係が直感的にわかりづらい分野の問題への適用が考えられる。

謝辞
本論文の作成にあたり有益なアドバイスを与えてくれた匿名査読者に感謝する。

## 参考文献

Bernhard Scholkopf, Alex J. Smola, Robert C. Williamson, Peter L. Bartlett [2000] “New Support Vector Algorithms”, *Neural Computation* 12, pp.1207-1245
Leo Breiman [2001] “Random Forests” *Machine Learning*, 45, pp.5-32
LIBSVM 3.14 —A Library for Support Vector Machine
(https://www.csie.ntu.edu.tw/~cjlin/libsvm/)
scikit-learn Machine Learning in Python
(https://scikit-learn.org/stable/index.html)
国民健康・栄養調査(https://www.mhlw.go.jp/bunya/kenkou/kenkou_eiyou_chousa.html)
国民生活基礎調査(https://www.mhlw.go.jp/toukei/list/20-21.html)
社会人口統計体系　統計でみる都道府県・市区町村のすがた
(https://www.stat.go.jp/data/ssds/index.html)
田辺和俊,鈴木孝弘[2020]「サポートベクター回帰における感度分析法による変数選択の有効性の検証―都道府県別全死因死亡率の影響要因の分析―」,*統計数理* 第 68 巻,第 1 号, pp.175-192
特定健康診査(https://www.mhlw.go.jp/stf/seisakunitsuite/bunya/0000161103.html)

## 付録　$\nu$-SVR の最適化アルゴリズムと非線形化

ここでは、３節で述べた$\nu$-SVR の最適化問題の解法とカーネル関数を用いた非線形化について概説する。

### １．$\nu$-SVR の最適化問題の解法

３節で記載の通り、最適化問題は以下のようになる：

$$\min_{\boldsymbol{w},\varepsilon,\xi_i^+,\xi_i^-} \frac{1}{2}||\boldsymbol{w}||_2{}^2 + C(\nu\varepsilon + \frac{1}{N}\sum_{i=1}^{N}(\xi_i^+ + \xi_i^-))$$

$$s.t. \quad \xi_i^- \geq -(y - f(\boldsymbol{x})) - \varepsilon \ldots (1)$$

$$\xi_i^- \geq 0 \quad \ldots (2)$$

$$\xi_i^+ \geq y - f(\boldsymbol{x}) - \varepsilon \quad \ldots (3)$$

$$\xi_i^+ \geq 0 \ldots (4)$$

$$\varepsilon > 0 \ldots (5)$$

ここで、$\xi_i^-, \xi_i^+$はそれぞれ$y - f(\boldsymbol{x}) \leq 0, y - f(\boldsymbol{x}) > 0$の場合の誤差を表し、制約条件(1), (3)は、$\xi_i^-, \xi_i^+$が$\varepsilon$以上の誤差を持つデータ（非サポートベクトル）の誤差であることに対応する。未定乗数を$\beta, \alpha_i^-, \alpha_i^+, \eta_i^-, \eta_i^+$として、ラグランジュ関数は以下のようになる：

$$L(\lambda, \gamma, \boldsymbol{w}, b, \boldsymbol{\xi}) = \frac{1}{2}||\boldsymbol{w}||_2{}^2 + C\nu\varepsilon + \frac{C}{N}\sum_{i=1}^{N}(\xi_i^+ + \xi_i^-)) - \beta\varepsilon - \sum_{i=1}^{N}(\eta_i^-\xi_i^- + \eta_i^+\xi_i^+)$$
$$- \sum_{i=1}^{N}\alpha_i^-\{\xi_i^- + y_i - (\boldsymbol{x}_i\boldsymbol{w} + b) + \varepsilon\} - \sum_{i=1}^{N}\alpha_i^+\{\xi_i^+ + y_i - (\boldsymbol{x}_i\boldsymbol{w} + b) + \varepsilon\}$$

$$\beta, \alpha_i^-, \alpha_i^+, \eta_i^-, \eta_i^+ \geq 0 \ldots (6)$$

$L$を$\boldsymbol{w}, b, \varepsilon, \xi_i^+, \xi_i^-$について最小化し、$\beta, \alpha_i^-, \alpha_i^+, \eta_i^-, \eta_i^+$について最大化することでパラメータを求める。最適化問題は以下のように書き換えられる：

$$\max_{\beta, \alpha_i^-, \alpha_i^+, \eta_i^-, \eta_i^+} \min_{\boldsymbol{w}, b, \varepsilon, \xi_i^+, \xi_i^-} L$$

各変数に関する微分を計算し、

$$\begin{cases} \dfrac{\partial L}{\partial \boldsymbol{w}} = 0 \\ \dfrac{\partial L}{\partial b} = 0 \\ \dfrac{\partial L}{\partial \varepsilon} = 0 \\ \dfrac{\partial L}{\partial \xi^-} = 0 \\ \dfrac{\partial L}{\partial \xi^+} = 0 \end{cases} \Leftrightarrow \begin{cases} \boldsymbol{w} = \displaystyle\sum_{i=1}^{N}(\alpha_i^+ - \alpha_i^-)\boldsymbol{x}_i \ldots ① \\ C\nu - \beta + \displaystyle\sum_{i=1}^{N}(\alpha_i^+ + \alpha_i^-) = 0 \ldots ② \\ \displaystyle\sum_{i=1}^{N}(\alpha_i^+ - \alpha_i^-) = 0 \ldots ③ \\ \dfrac{C}{N} - \alpha_i^- - \eta_i^- = 0 \ldots ④ \\ \dfrac{C}{N} - \alpha_i^+ - \eta_i^+ = 0 \ldots ⑤ \end{cases}$$

$L$に代入すると、

$$L(\lambda, \gamma, \boldsymbol{w}, b, \boldsymbol{\xi}) = \frac{1}{2}||\boldsymbol{w}||_2{}^2 + C\nu\varepsilon + \frac{C}{N}\sum_{i=1}^{N}(\xi_i^+ + \xi_i^-)) - \beta\varepsilon - \sum_{i=1}^{N}(\eta_i^-\xi_i^- + \eta_i^+\xi_i^+) + \sum_{i=1}^{N}(\alpha_i^+ - \alpha_i^-)y_i$$
$$- \sum_{i=1}^{N}\alpha_i^-\{\xi_i^- + y_i - (\boldsymbol{x}_i\boldsymbol{w} + b) + \varepsilon\} - \sum_{i=1}^{N}\alpha_i^+\{\xi_i^+ + y_i - (\boldsymbol{x}_i\boldsymbol{w} + b) + \varepsilon\}$$
$$= \frac{1}{2}\left|\left|\boldsymbol{w}\right|\right|_2{}^2 + \sum_{i=1}^{N}(\alpha_i^+ - \alpha_i^-)y_i + (C\nu - \beta)\varepsilon + \frac{C}{N}\sum_{i=1}^{N}(\xi_i^+ + \xi_i^-) - \sum_{i=1}^{N}(\eta_i\xi_i + \eta_i^*\xi_i^*)$$
$$- \sum_{i=1}^{N}(\alpha_i^- - \alpha_i^+)\boldsymbol{x}_i\boldsymbol{w} - \sum_{i=1}^{N}(\alpha_i^- - \alpha_i^+)b - \sum_{i=1}^{N}\alpha_i^-(\xi_i^- + \varepsilon) - \sum_{i=1}^{N}\alpha_i^+(\xi_i^+ + \varepsilon)$$
$$= -\frac{1}{2}\sum_{i,j}(\alpha_i^+ - \alpha_i^-)(\alpha_j^+ - \alpha_j^-)\boldsymbol{x}_i^T\boldsymbol{x}_j + \sum_{i=1}^{N}(\alpha_i^+ - \alpha_i^-)y_i - \sum_{i=1}^{N}\varepsilon(\alpha_i^- + \alpha_i^+)$$

ここで、制約条件(6)と式④⑤より、

$$0 \leq \alpha_i^- \leq \frac{C}{N}, 0 \leq \alpha_i^+ \leq \frac{C}{N}$$

また、制約条件(6)と式②より、

$$\frac{\sum_{i=1}^{N}(\alpha_i^+ + \alpha_i^-)}{C} \leq \nu$$

よって、双対問題は以下のようになる：

$$\min_{\boldsymbol{\alpha}^-,\boldsymbol{\alpha}^+} \ -\frac{1}{2}\sum_{i,j}(\alpha_i^+ - \alpha_i^-)(\alpha_j^+ - \alpha_j^-)\boldsymbol{x}_i^T\boldsymbol{x}_j + \sum_{i=1}^{N}(\alpha_i^+ - \alpha_i^-)y_i - \sum_{i=1}^{N}\varepsilon(\alpha_i^- + \alpha_i^+)$$

$$s.t. 0 \leq \alpha_i^- \leq \frac{C}{N} \ldots (7),$$

$$0 \leq \alpha_i^+ \leq \frac{C}{N} \ldots (8),$$

$$\frac{\sum_{i=1}^{N}(\alpha_i^+ + \alpha_i^-)}{C} \leq \nu \ldots (9),$$

$$\sum_{i=1}^{N}(\alpha_i^+ - \alpha_i^-) = 0$$

ここで、KKT 条件より、$\alpha_i^-\{\xi_i^- + y_i - (\boldsymbol{x}_i\boldsymbol{w} + b) + \varepsilon\} = 0, \alpha_i^+\{\xi_i^+ + y_i - (\boldsymbol{x}_i\boldsymbol{w} + b) + \varepsilon\} = 0$なので、$\alpha_i^- \neq 0$の時$\xi_i^- + y_i - (\boldsymbol{x}_i\boldsymbol{w} + b) + \varepsilon = 0$となり、これは$\boldsymbol{x}_i$がサポートベクトルであることに対応する。また、制約条件(7)~(9)より、$\nu$はサポートベクトル数の学習データ数に対しての率の下限となる。

双対問題を導出することができたので、これ以降は SVM の場合と同じ手順で解くことができる。また関係式を用いると、回帰直線は以下のように表せる。

$$f(\boldsymbol{x}) = (\sum_{i=1}^{N}(\alpha_i^+ - \alpha_i^-)\boldsymbol{x}_i)^T\boldsymbol{x} + b$$

## ２．カーネル関数による非線形化

ここでは、SVM や SVR を非線形化する方法であるカーネル法について、回帰の場合に限定して説明する。

説明変数ベクトル

$\boldsymbol{x}$を特徴空間$\boldsymbol{F}$に写像する関数$\boldsymbol{\phi}: \boldsymbol{R}^M \to \boldsymbol{F}$を考える。この$\boldsymbol{\phi}(\boldsymbol{x})$を新たな説明変数ベクトルだと考えると、双対問題は以下のように書き換えられる：

$$\max_{\boldsymbol{\alpha}^+,\boldsymbol{\alpha}^-} -\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}(\alpha_i^+ - \alpha_i^-)(\alpha_j^+ - \alpha_j^-)\,\boldsymbol{\phi}(\boldsymbol{x}_i)^T\boldsymbol{\phi}(\boldsymbol{x}_j) + \sum_{i=1}^{N}(\alpha_i^+ - \alpha_i^-)y_i - \sum_{i=1}^{N}(\alpha_i^+ - \alpha_i^-)\varepsilon$$

$$s.t.\sum_{i=1}^{N}(\alpha_i^+ - \alpha_i^-) = 0 \ \ ,0 \le \alpha_i^+ \le C \ \ ,0 \le \alpha_i^- \le C, \forall \ \ i \in \{1,2,\dots,N\}$$

写像$\boldsymbol{\phi}$による変換が非線形であれば、回帰関数$f(\boldsymbol{x})$はもとの$\boldsymbol{x}$の空間では非線形になり得る。一方、$f(\boldsymbol{x})$は$\boldsymbol{\phi}(\boldsymbol{x})$に関して１次なので$\boldsymbol{\phi}(\boldsymbol{x})$の空間$\boldsymbol{F}$上において$f(\boldsymbol{x})$は線形な回帰関数となっている。そのため、変換後の$\boldsymbol{\phi}(\boldsymbol{x})$を新たな説明変数ベクトルとみなして$\boldsymbol{x}$と置き換えれば、これまでの導出はそのまま成立する。

ここで、SVR の双対問題において$\boldsymbol{\phi}(\boldsymbol{x})$は内積の形$\boldsymbol{\phi}(\boldsymbol{x}_i)^T\boldsymbol{\phi}(\boldsymbol{x}_j)$でのみ現れているので、$\boldsymbol{\phi}(\boldsymbol{x})$を直接計算する必要はなく、内積$\boldsymbol{\phi}(\boldsymbol{x}_i)^T\boldsymbol{\phi}(\boldsymbol{x}_j)$さえ計算できればよい。この内積$\boldsymbol{\phi}(\boldsymbol{x}_i)^T\boldsymbol{\phi}(\boldsymbol{x}_j)$をカーネル関数として以下のように定義する：

$$K(\boldsymbol{x}_i, \boldsymbol{x}_j) = \phi(\boldsymbol{x}_i)^T\phi(\boldsymbol{x}_j)$$

ある特定の性質を満たす関数を用いると、$\boldsymbol{\phi}(\boldsymbol{x})$の計算を行うことなく直接$K(\boldsymbol{x}_i, \boldsymbol{x}_j)$の計算ができることが知られている。カーネル関数はこれまで様々な種類のものが提案されており、文字列のデータに適したものやグラフ構造のあるデータに対して適したものなど用途によって最も適したものを使用することができる。内積${\boldsymbol{x}_i}^T\boldsymbol{x}_j$をカーネル関数$K(\boldsymbol{x}_i, \boldsymbol{x}_j)$で置き換えると回帰関数は以下のようになる：

$$f(\boldsymbol{x}) = \sum_{i=1}^{N}(\alpha_i^+ - \alpha_i^-)\,K(\boldsymbol{x}_i, \boldsymbol{x}) + b$$

Factor Analysis of Regional Mortality Disparities Using Improved SVR Sensitivity Analysis Method

Sohei Tanaka　Naoki Matsuyama

Abstract

While Support Vector Regression (SVR) can express complex nonlinear relationships in data, it lacks readability. The SVR sensitivity analysis method introduces a "sensitivity" for each explanatory variable, which allows the influence of explanatory variables to be evaluated and explanatory variables to be selected according to the magnitude of the sensitivity. However, the hyperparameter search in the process of sensitivity-based variable reduction and re-learning is unstable. In this study, we propose a new algorithm that facilitates the hyperparameter search for this SVR sensitivity analysis method, and show that it provides better estimation results than existing studies in the analysis of the influence factors of age-adjusted mortality rates by prefecture.