

カルマンフィルタによるカーネル回帰

柳川 実

Kernel regression based on the Kalman filtering

Minoru Yanagawa

Abstract

The theory of support vector machines (SVMs) with Gaussian kernels has been successfully applied to smooth curve fittings on a set of discrete data. In this paper, we develop an algorithm for obtaining kernel regression curves fitted to stochastic discrete data. We appeal to both Kalman filtering and SVM theories to construct our algorithm.

Key words : curve fitting, Kalman filter, nonlinear regression, support vector machine

1. 緒 言

先の論文¹⁾において、サポートベクターマシン (SVM) による離散データの曲線当てはめの有効性について述べた。本論文では、同一確率分布に従う、試行ごとに異なる離散データの集まりを考え、この集まりに対する回帰曲線を求めるアルゴリズムを SVM を利用して構成する。そのために、SVM に適合するカルマンフィルタのアルゴリズムを作成し、それを試行毎のデータに適用し回帰関数の係数を更新してゆく。

2. データの従う確率分布

x 軸上の各点 $x_i (i=0,1, \dots, 10)$ における確率変数 X_i は平均 μ_i 、分散 σ_i^2 の正規分布に従うと仮定する。表 1 で与えられる平均値と分散値を扱う。

図 1 は表 1 の平均値 μ の散布図である。

表 1 標本点の確率分布

x	0	1	2	3	4	5
μ	0	2	4	6	7	8
σ^2	0.04	0.25	0.49	1	1.69	2.25

x	6	7	8	9	10
μ	7	6	4	2	0
σ^2	1.69	1	0.49	0.25	0.04

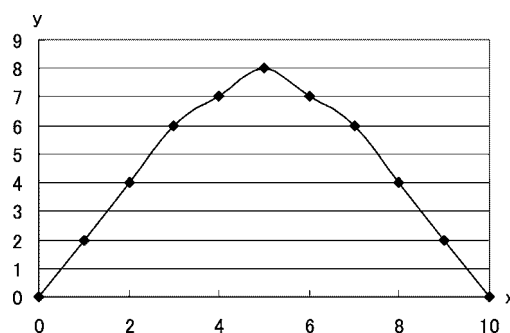


図 1 平均値の散布図

3. 回帰関数

ガウスカネルの組 $\{K(x, x_i)\}_{i=0}^m$ を用いる回帰関数を $F(x) = \sum_{i=0}^m a_i K(x, x_i)$ とする。ガウスカネルのグラム行列は \hat{K} とかく。係数の平均値が β , 共分散行列が M と与えられている。データ y と回帰値 Y に対して, 確率的に $y = Y + \nu$ の関係を仮定する。ただし, ν は測定ノイズである。与えられたデータ y に適する回帰係数 $\hat{\beta}$ は, R を ν の共分散行列とすると

$$\hat{\beta} = \beta + P\hat{K}^T R^{-1}(y - \hat{K}\beta)$$

$$P = (M^{-1} + \hat{K}^T R^{-1} \hat{K})^{-1}$$

を計算して得ることができる。

4. カルマンフィルタ

表1の確率分布から近似的に求めた回帰係数の共分散行列を P_0 とかく。求める際に用いる平均値は, 表1の平均 μ に対して行う SVM (サポートベクターマシン) 当てはめ¹⁾の係数 \hat{B}_0 を用いる。

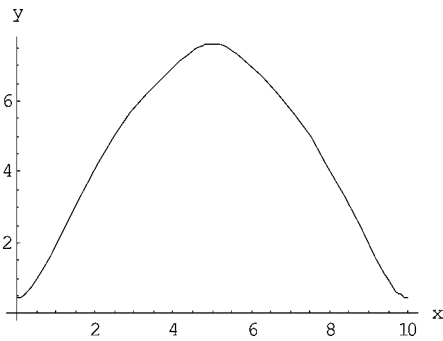


図2 平均 μ の SVM 当てはめ

図2は係数が \hat{B}_0 である回帰関数のグラフである。

第3節の結果に対してカルマンフィルタの理論を適用すると, 次のアルゴリズムが得られる。これをカーネルカルマンフィルタと呼ぶ。

$$\begin{cases} \hat{B}_{i+1} = \hat{B}_i + P_{i+1} \hat{K}^T R^{-1} (y_i - \hat{K} \hat{B}_i) \\ P_{i+1} = (P_i^{-1} + \hat{K}^T R^{-1} \hat{K})^{-1} \end{cases}$$

5. カーネル回帰曲線

表1の確率分布に従うデータを基にして, 回帰の様子を見る。

各点 x_i において, 正規乱数の発生により得られた標本を図3に与える。

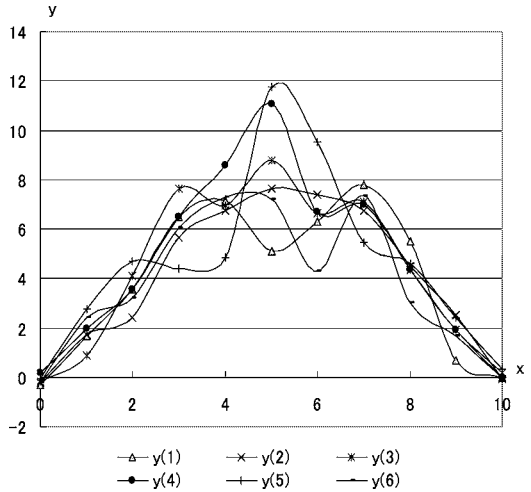


図3 標本 $y(1) \cdots y(6)$

最初に y_1 に対してアルゴリズムを適用すると, 図4に与えられる回帰曲線を導くことが出来る。

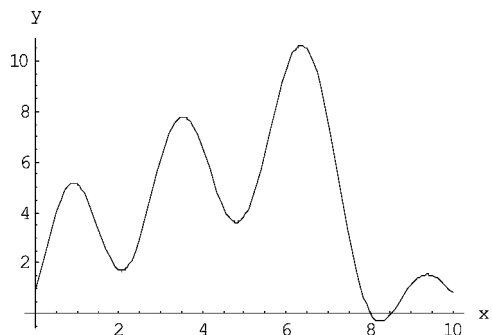


図4 係数 \hat{B}_1 の回帰曲線

カーネルカルマンフィルタのアルゴリズムを逐次 y_6 まで更新し続けてゆくと, 図5の如く, 図2のグラフにより近いカーネル回帰関数のグラフが得られる。

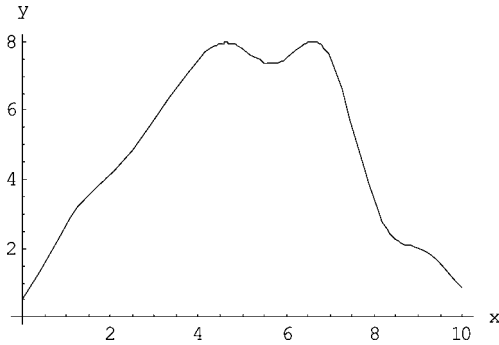


図5 係数 \hat{B}_6 の回帰曲線

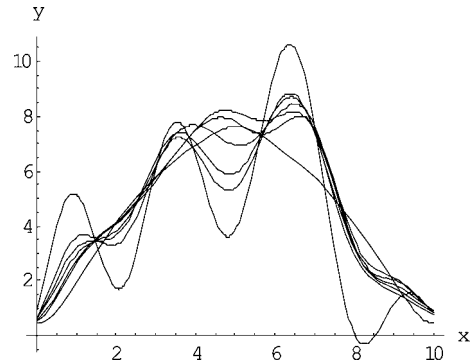


図6 回帰関数列の収束

6. 結 言

アルゴリズムを実行する際の最初のデータとして、第5節のデータのどれをとってもよい。例えば、 y_1 の代わりに y_2 を初期データとしても結果は同一である。もちろん、逐次実行するアルゴリズムの初期の段階においては幾分かの差はあるが、十分回数を重ねることにより一定の回帰曲線に収束してゆく。

図6に同一座標軸上に図2のグラフに併せて y_1 から y_6 の定める回帰関数のグラフが描かれ

ている。回帰関数列の収束の様子がグラフの束から見て取れ、真の回帰曲線が十分に推測できる。

文 献

- 1) 柳川実 (2008) 曲率の小さい曲線当てはめに関する一考察. 日大歯紀 35, 1-4
- 2) 片山徹 (1983) 応用カルマンフィルタ. 朝倉書店, 東京
- 3) N. クリスティアニーニ, J.S. テイラー (大北剛訳) (2005) サポートベクターマシン入門. 共立出版, 東京