卒業研究発表会 150番

ガウス過程に基づく逐次抽出型 c-回帰モデルの検討

知能情報基礎研究室
学籍番号:1810370205
氏名:武川海斗
2022年1月24日

研究の概要

逐次抽出型
c-回帰モデル

＋

ガウス過程に
基づくc-回帰
モデル

＝ 提案手法

クラスタ数の自動推定　非線形な構造が得られる

逐次抽出型c-回帰モデル(SCRM)

目的関数 $J_{\mathrm{SCRM}}(\boldsymbol{U}, \boldsymbol{B}) = \sum_{k=1}^{n} u_{k1} \left(y_k - f(x_k; \beta_i)\right)^2 + \sum_{k=1}^{n} u_{k0} D$

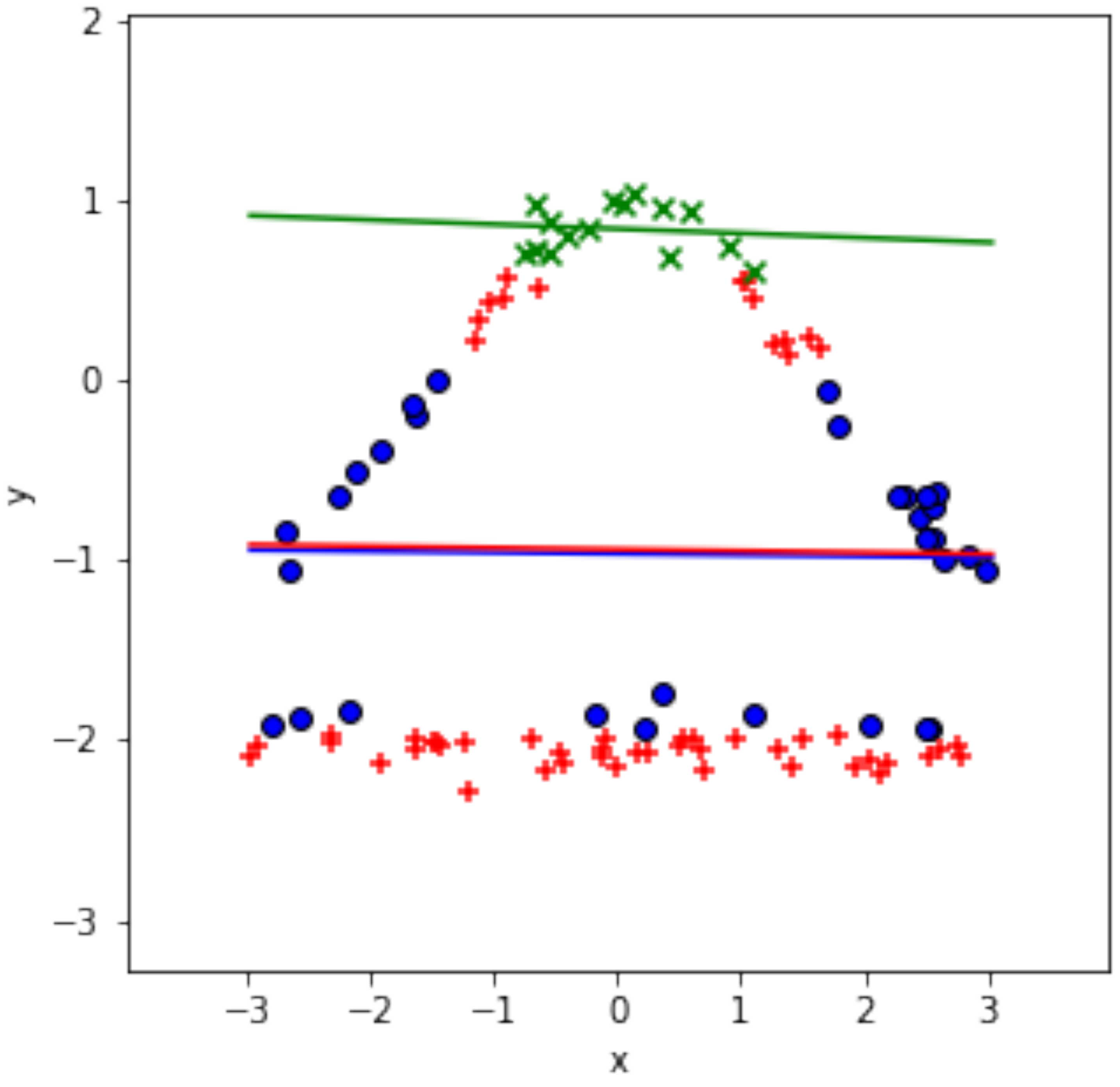
制約条件 $U_{\mathrm{SCRM}} = \left\{ \left(u_{ki}\right) : u_{ki} \in \{0,1\}, \sum_{j=1}^{c} u_{ki} = 1, {}^{\forall}k \right\}$

$$f(x_k; \beta_i) = \sum_{j=1}^{p} \beta_i^j x_k^j + \beta_i^{p+1}, x \in \mathfrak{R}^p, k = 1 \sim n, i = 1 \sim c$$

$\left(y_k - f(x_k; \beta_i)\right)^2$：非類似度　x_k：説明変数

u_{k1}：抽出クラスタ　y_k：目的変数

u_{k0}：ノイズクラスタ　$D > 0$：ノイズパラメータ



非線形なクラスタ構造に対する
SCRMの出力図

問題点：SCRMでは非線形な分類が困難である

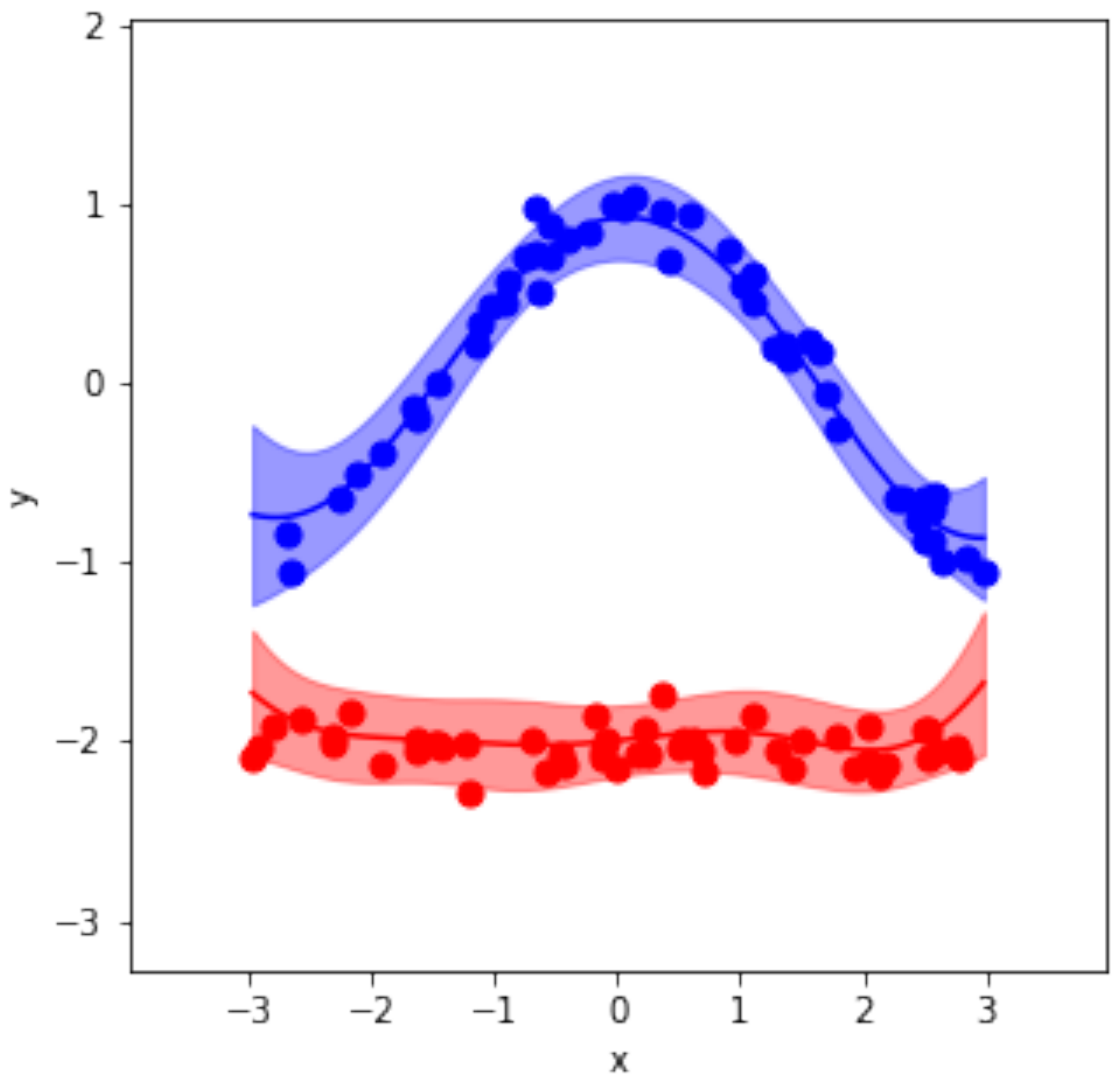
[1] S. Miyamoto and K. Arai, Different Sequential Clustering Algorithms and Sequential Regression Models, IEEE International Conference on Fuzzy Systems, pp. 1107–1112, 2009.

ガウス過程に基づくc-回帰モデル(GPCRM)

非線形な回帰線を求めることが可能

目的関数 $$J_{\mathrm{gpcrm}}(\boldsymbol{U},\boldsymbol{K}) = \sum_{k=1}^{n}\sum_{i=1}^{c} u_{ki}\left(y_k^{(i)} - \boxed{\boldsymbol{k}_*^{(i)T}\boldsymbol{K}^{(i)-1}\boldsymbol{Y}^{(i)}}\right)^2$$

制約条件 $$\sum_{i=1}^{c} u_{ki} = 1,(k = 1 \sim n)$$



GPCRMの出力図

$\left(y_k^{(i)} - \boldsymbol{k}_*^{(i)T}\boldsymbol{K}_i^{-1}\boldsymbol{Y}_i\right)^2$：非類似度

y_k：目的変数 $\boldsymbol{K}^{(i)}$：カーネル行列(入力点,入力点)

u_{ki}：帰属度 ただし，$\boldsymbol{K}^{(i)-1} = \left(\boldsymbol{K}^{(i)} + \lambda\boldsymbol{I}\right)^{-1}$

$\boldsymbol{k}_*^{(i)}$：カーネル関数(入力点, 予測したい点)

問題点：クラスタ数を事前に決定する必要がある

[2] 金月優斗, ガウス過程に基づくc-回帰モデル, 近畿大学大学院総合理工学研究科エレクトロニクス系工学専攻令和２年度修士論文

提案手法の定式化

SCRMとGPCRMを足し合わせた手法

SCRM

$$J_{\mathrm{SCRM}}(\boldsymbol{U},\boldsymbol{B}) = \sum_{k=1}^{n} u_{k1}(y_k - f(x_k;\beta_i))^2 + \sum_{k=1}^{n} u_{k0}D$$

+

GPCRM

$$J_{\mathrm{gpcrm}}(\boldsymbol{U},\boldsymbol{K}) = \sum_{k=1}^{n}\sum_{i=1}^{c} u_{ki}(y_k^{(i)} - \boldsymbol{k}_*^{(i)T}\boldsymbol{K}_i^{-1}\boldsymbol{Y}_i)^2$$

$$J_{\mathrm{gpscrm}}(\boldsymbol{U},\boldsymbol{K}) = \sum_{k=1}^{n} u_{k1}(y_k^{(1)} - \boldsymbol{k}_*^{(1)T}\boldsymbol{K}^{(1)-1}\boldsymbol{Y}^{(1)})^2 + \sum_{k=1}^{n} u_{k0}D$$

提案手法の目的関数

ガウス過程に基づく逐次抽出型c-回帰モデル(GPSCRM)

目的：クラスタ数の自動推定と非線形な回帰線を得ること

目的関数 $$J_{\mathrm{gpscrm}}(\boldsymbol{U},\boldsymbol{K}) = \sum_{k=1}^{n} u_{k1}(y_k^{(1)} - \boldsymbol{k}_*^{(1)T}\boldsymbol{K}^{(1)-1}\boldsymbol{Y}^{(1)})^2 + \sum_{k=1}^{n} u_{k0}D$$

制約条件 $$U_{gpscrm} = \left\{ (u_{ki}) : u_{ki} \in \{0,1\}, \sum_{i=0}^{1} u_{ki} = 1, {}^{\forall}k \right\}$$

$\left(y_k^{(i)} - \boldsymbol{k}_*^{(1)T}\boldsymbol{K}^{(i)-1}\boldsymbol{Y}^{(i)}\right)^2$：非類似度

$\boldsymbol{k}_*^{(1)}$：カーネル関数(入力点, 予測したい点)

y_k：目的変数

u_{k1}：抽出クラスタへの帰属度

$\boldsymbol{K}^{(1)}$：抽出クラスタに属する
カーネル行列(入力点,入力点)

u_{k0}：ノイズクラスタへの帰属度

$D > 0$：ノイズパラメータ

ただし，$\boldsymbol{K}^{(1)-1} = \left(\boldsymbol{K}^{(1)} + \lambda \boldsymbol{I}\right)^{-1}$

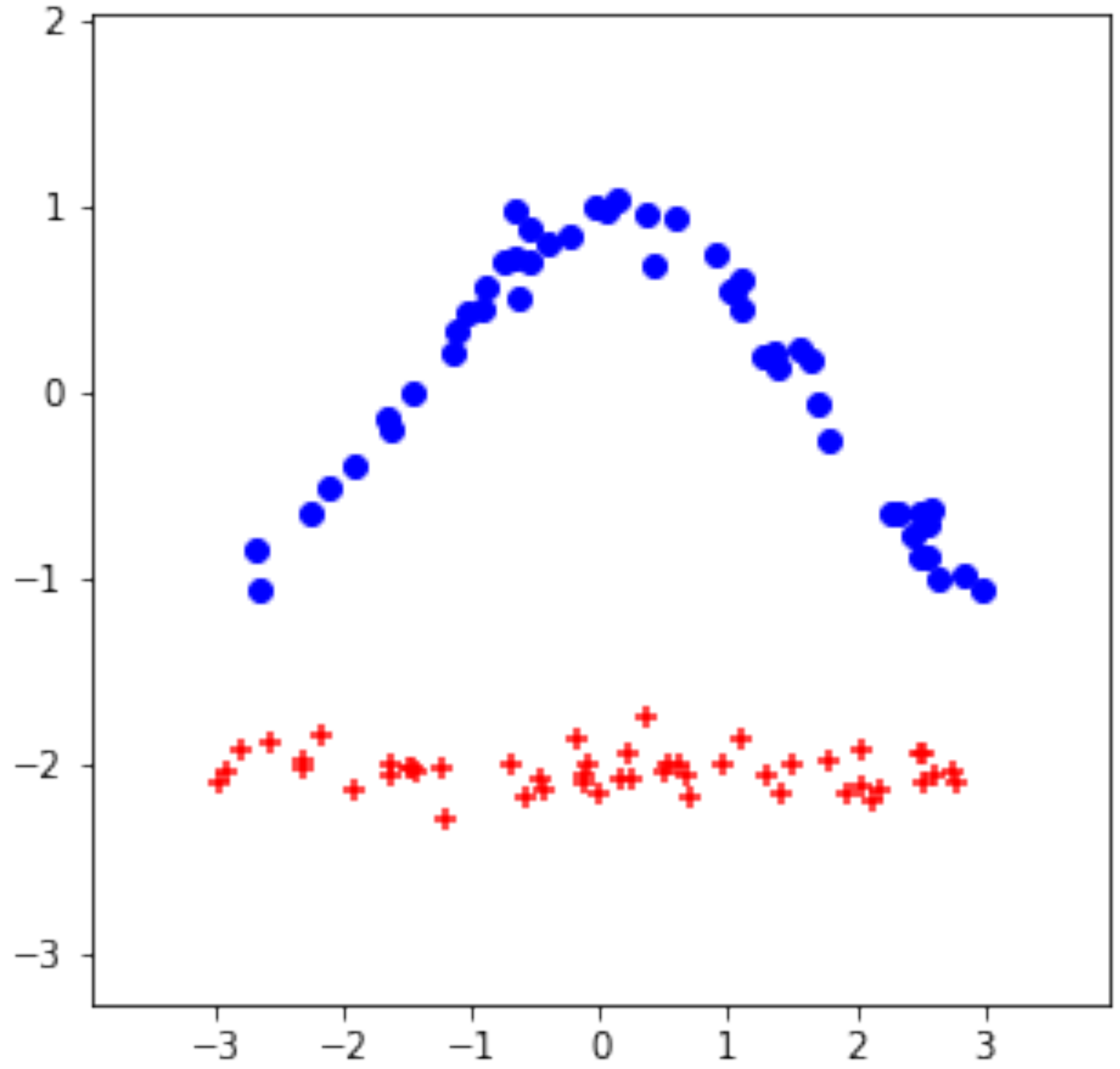
カーネル関数にはガウスカーネルを利用

実験条件

目的：提案手法が非線形な回帰線を得られるかどうかを調べること

- SCRM,　GPCRMと比較実験
- Adjusted Rand Index(ARI)[3]で評価
- 青,　赤,　緑の順番で出力する

データセット

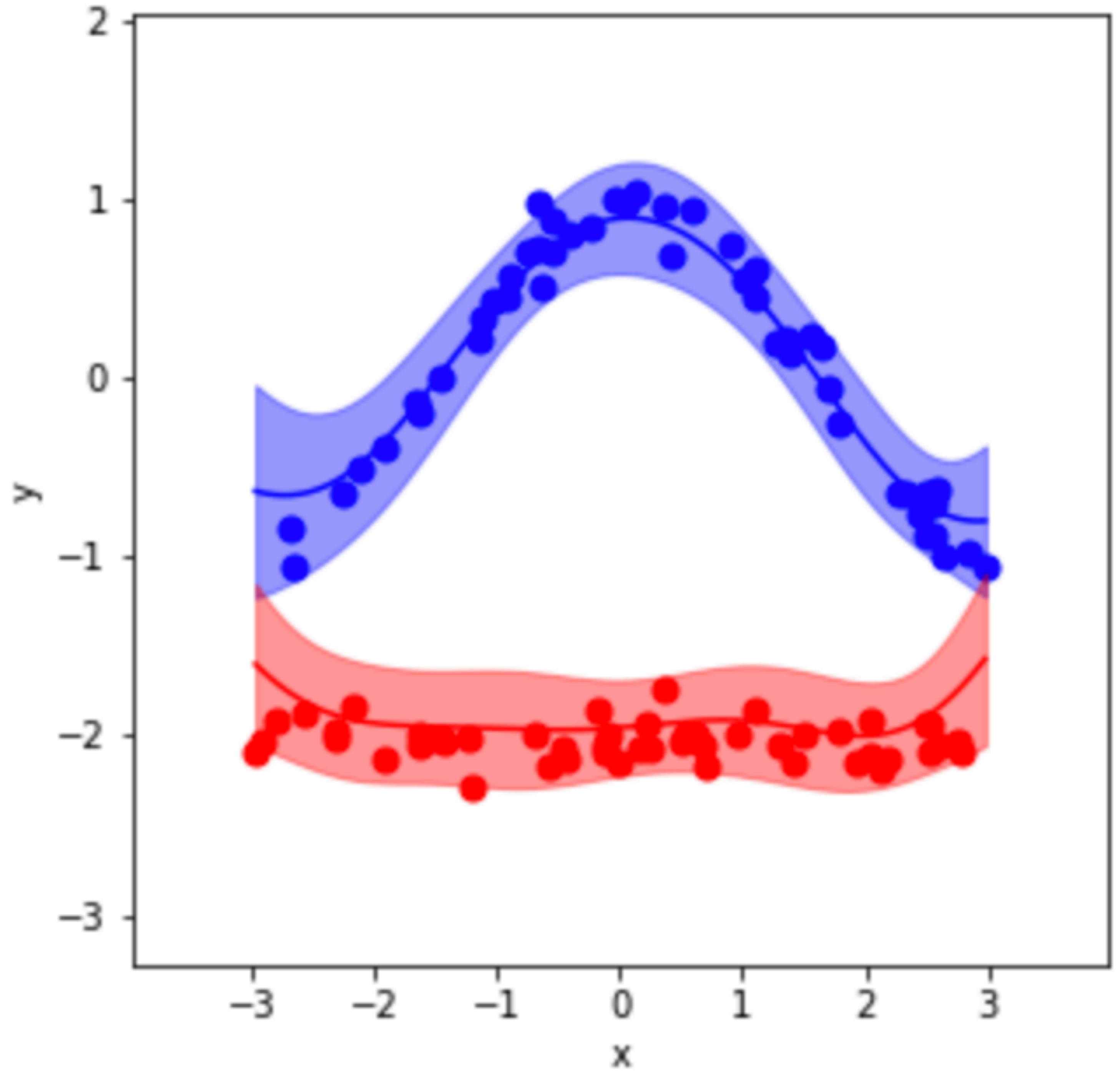


n=200, c=2

[3] L. Hubert and P. Arabie, Comparing Partitions, Journal of Classification, Vol.2, No.1, pp.193-218, 1985.

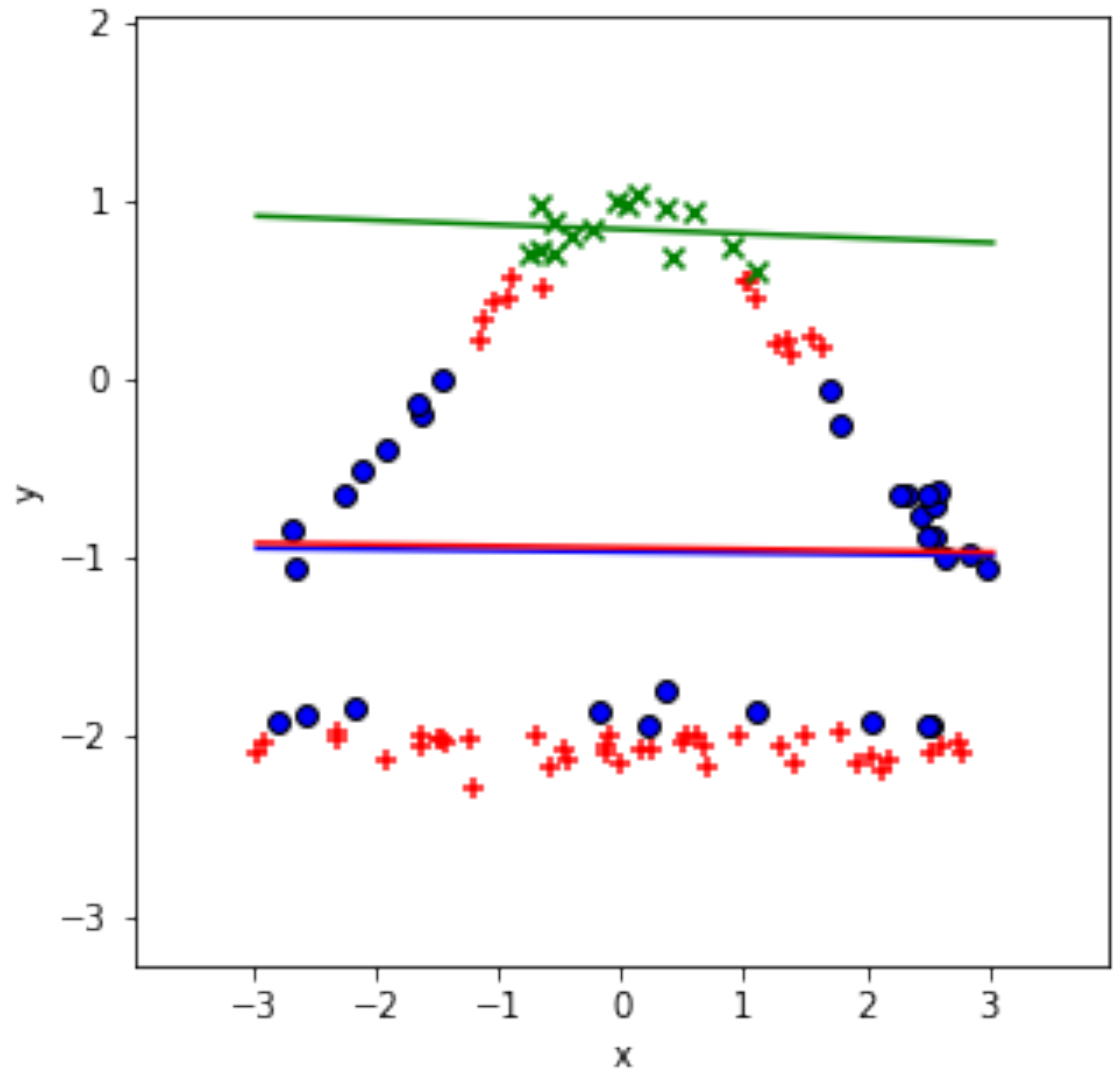
実験結果と考察

GPCRM



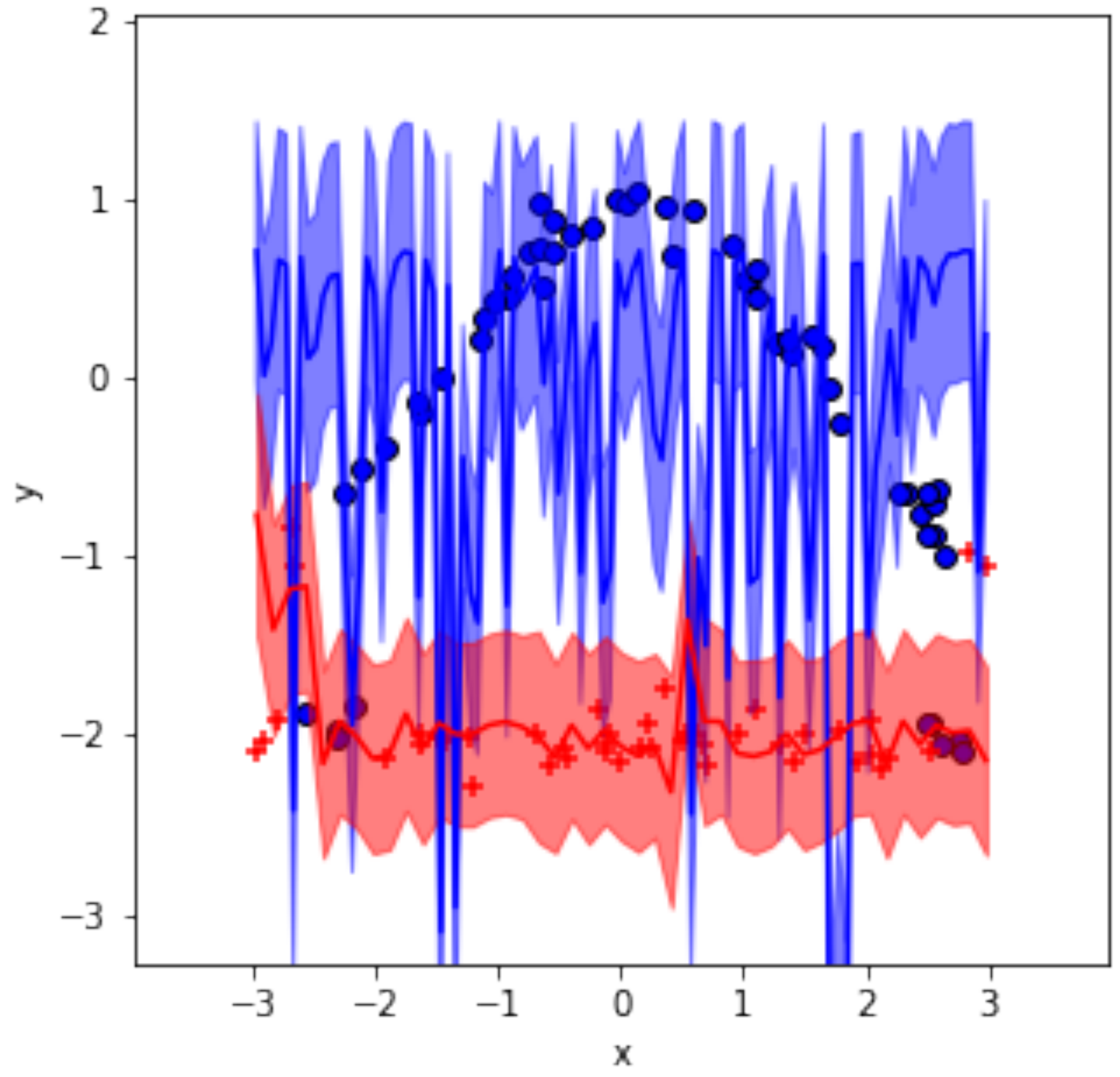
ARI=1.0, λ=0.5

SCRM



ARI=0.50, D=1.5

GPSCRM



ARI=0.80, D=1.5,

α=0.09, λ=0.5

回帰線が非線形な構造をうまく捉えられていない

まとめ

- 結果
  - 現状の提案手法では，良好な非線形な分割が行えなかった
  - 抽出ごとにハイパーパラメータを更新する必要がある

- 今後の課題
  - 抽出ごとにハイパーパラメータを自動推定する手法を考案する

付録

GPSCRMのアルゴリズム

1. ノイズパラメータD，正則化パラメータλ，ガウスパラメータαを設定する．

2. ランダムに帰属度u_{ki}を設定する．収束するまで繰り返す．

　2.1　パラメータ$\boldsymbol{K}^{(1)}$を更新する．

　2.2　非類似度d_{ki}を基に，帰属度u_{ki}を更新する．

$$u_{ki} = \begin{cases} i & (d_{ki} \leq D) \\ 1 - i & (otherwise) \end{cases} \quad (i = 0,1)$$

3. $\{x_k | u_{ki} = 1\}$の要素を$\boldsymbol{X}$から抽出する．

4. $\boldsymbol{X} = \emptyset$の場合終了．続行可能の場合，2に戻る．

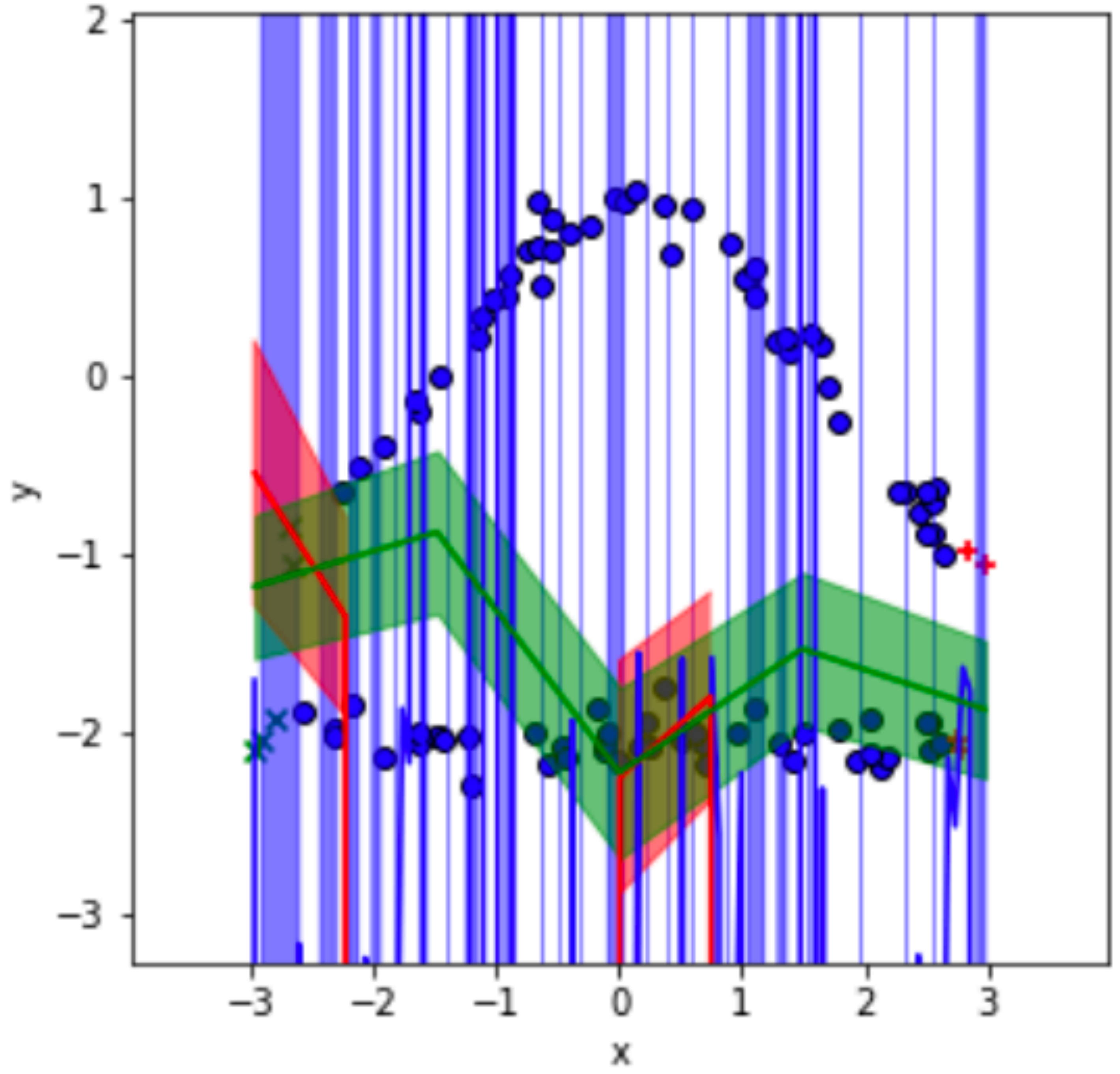
ハイパーパラメータについて

- カーネルパラメータ：α
  - 回帰線の形状を決定するパラメータ
  - 0.001 ~ 0.1の中から選ぶ
- ノイズパラメータ：D
  - 抽出クラスタの範囲を決定するパラメータ
  - D=1.5とし，抽出するたびにノイズパラメータDを更新する
- 正則化パラメータ：λ
  - データの過学習を防ぐためのパラメータ
  - 0.1 ~ 2.0の中から選ぶ

αとλの組み合わせから，ARIが良好な結果を選択する

カーネルパラメータαを変えた場合

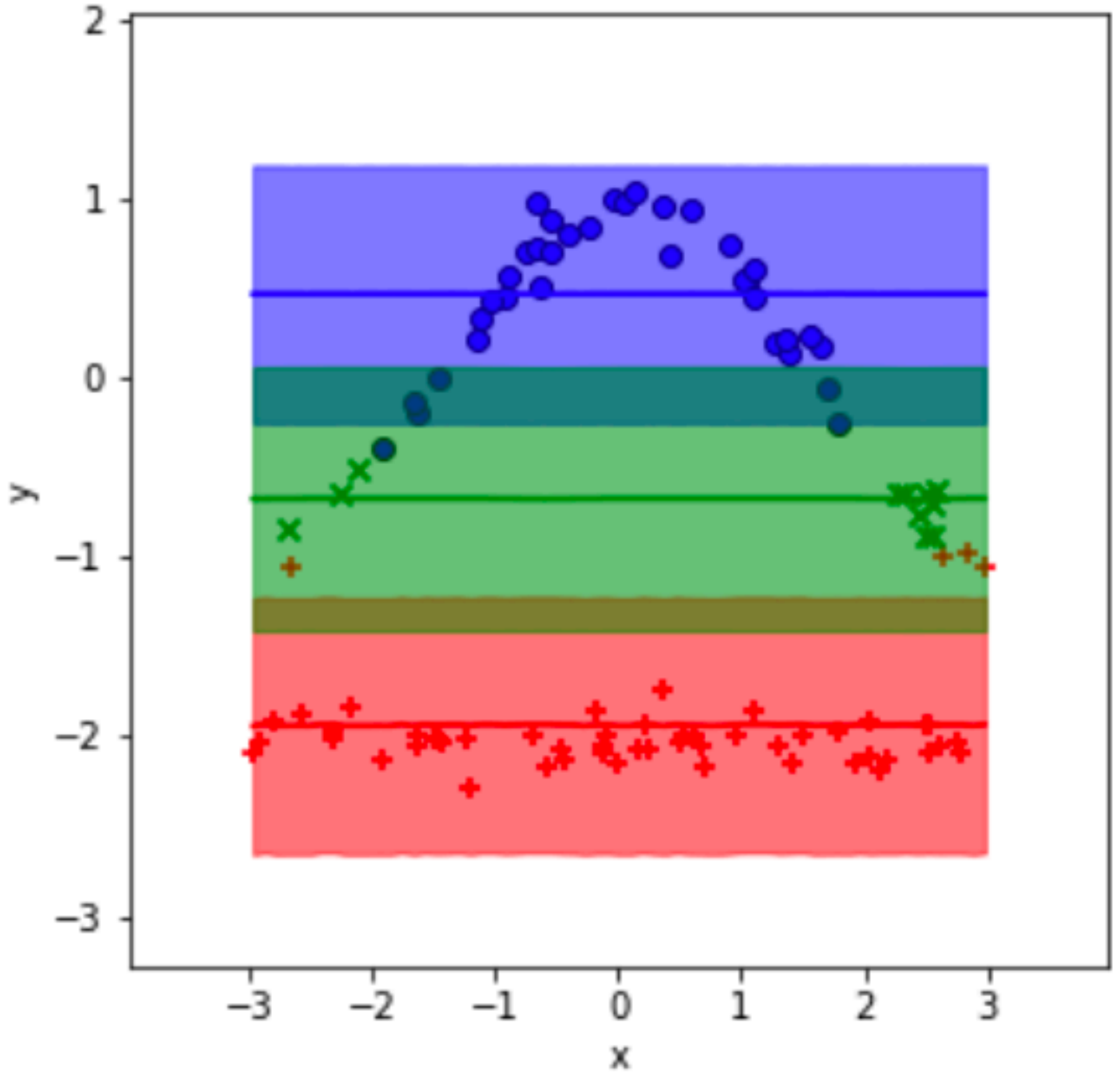
αを大きくした場合



GPSCRM による出力, D=1.5, α=5,

λ = 0.5, ARI = 0.027

αを小さくした場合

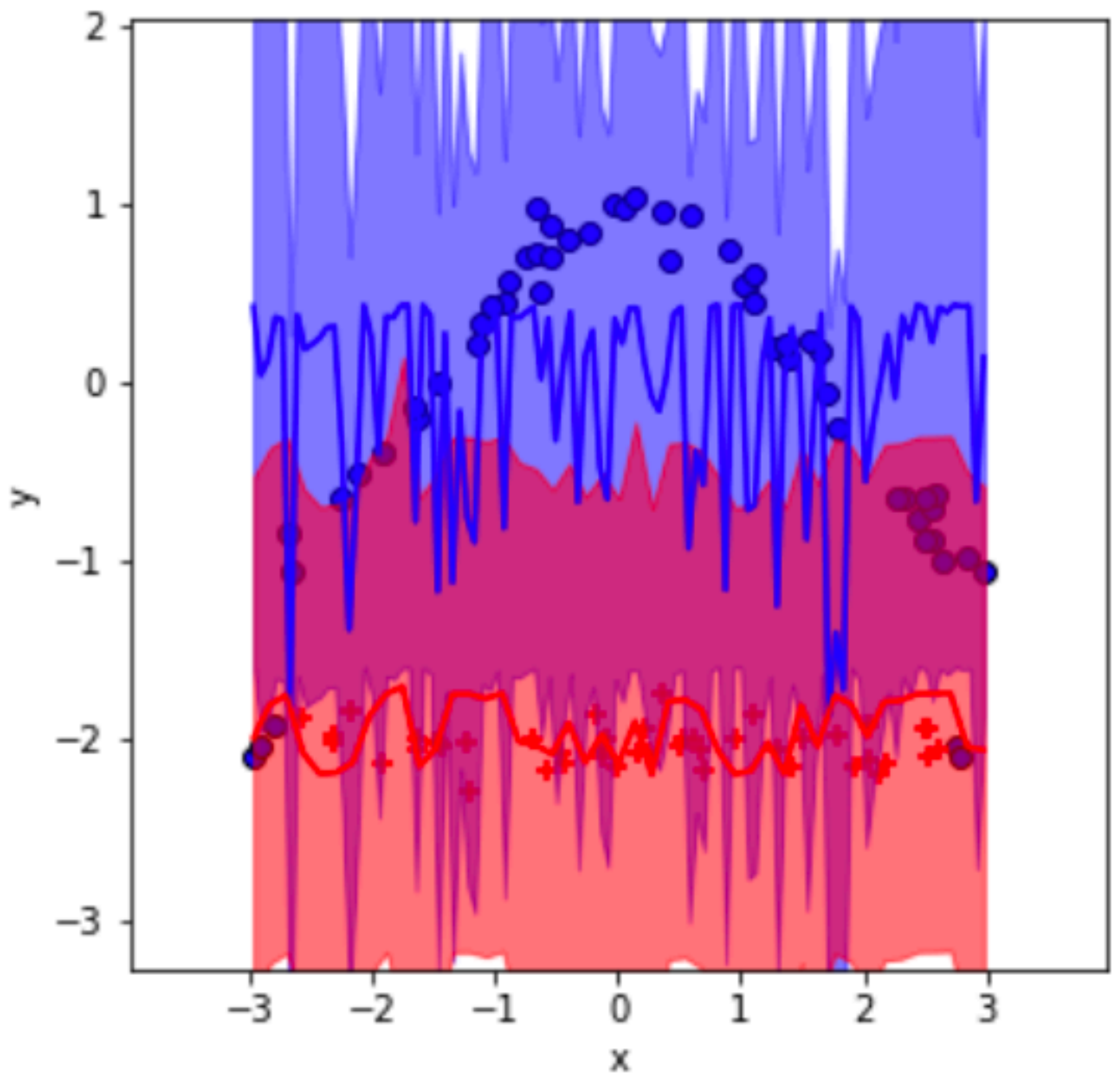


GPSCRM による出力, D=1.5, α=0.0001,

λ = 0.5, ARI = 0.97

正則化パラメータλを変えた場合

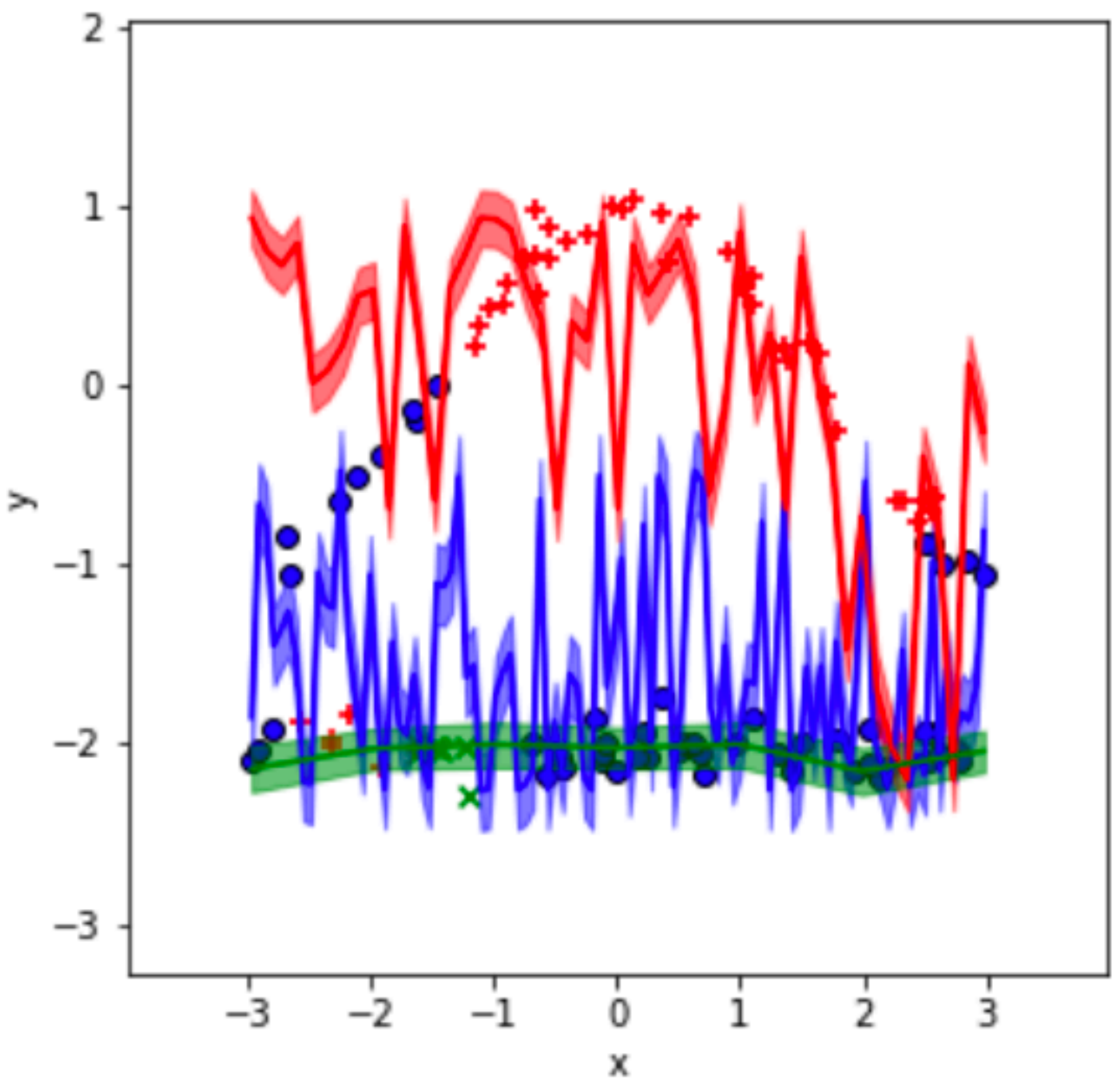
λを大きくした場合



GPSCRM による出力, D=1.5, α=0.09,

$\lambda = 4.0$, ARI = 0.80

λを小さくした場合



GPSCRM による出力, D=1.5, α=0.09,

$\lambda = 0.045$, ARI = 0.75

ガウス過程における期待値について

i番目のクラスタに属するデータに対するガウス過程の期待値：

$$\begin{pmatrix} \boldsymbol{y}^{(i)} \\ y_k^{(i)*} \end{pmatrix} \sim N\left(\mathbf{0}, \begin{pmatrix} \mathbf{K}_i & \mathbf{k}_k^{(i)} \\ \mathbf{k}_k^{(i)T} & k_{kk} \end{pmatrix}\right)$$

$$P_i\left(y_k^{(i)*} \mid x_k^{(i)*}, D_i\right) = N\left(\boxed{\mathbf{k}_k^{(i)T}\mathbf{K}_i^{-1}y^i}, \mathbf{k}_k^{**} - \mathbf{k}_{ki}^{*T}\mathbf{K}_i^{-1}\mathbf{k}_{ki}^{*}\right)$$

ガウス過程回帰による期待値を表す部分

$\boldsymbol{K}_i$：カーネル関数の行列

$\boldsymbol{k}_k^{(i)}$：カーネル関数のベクトル

k_{kk}：カーネル関数

正則化項λについて

- カーネル関数によって表現力の高さから過学習を引き起こす

$$P_i\left(y_k^{(i)*} \mid x_k^{(i)*}, D_i\right) = N\left(\mathbf{k}_k^{(i)T}\mathbf{K}_i^{-1}y^i, \mathbf{k}_k^{**} - \mathbf{k}_{ki}^{*T}\mathbf{K}_i^{-1}\mathbf{k}_{ki}^{*}\right)$$

カーネル行列に対して，正則化項を追加する

$$\boxed{\mathbf{K}_i^{-1} = (\boldsymbol{K}_i + \lambda\mathbf{I}_{C_i})^{-1}}$$