社団法人　電子情報通信学会
THE INSTITUTE OF ELECTRONICS,
INFORMATION AND COMMUNICATION ENGINEERS

信学技報
TECHNICAL REPORT OF IEICE.

[招待講演] スプライン関数の基礎と分位点回帰への応用

北原　大地†

† 立命館大学情報理工学部　〒 525–8577 滋賀県草津市野路東 1–1–1
E-mail: † d-kita@fc.ritsumei.ac.jp

あらまし　スプライン関数とは，微分も含んだ何らかの連続性条件を満たすように設計された，滑らかな区分的多項式のことである．特に，3 次の自然スプライン関数は，ある種の曲率を最小化する最も滑らかな関数となることから，基礎から応用まで幅広い分野で使用されている．本稿では，1 変数及び 2 変数スプライン関数の定義や重要な性質を，証明も交えて解説する．そして，変分問題の最適解となるスプライン関数の具体的かつ汎用的な構成方法を説明する．最後に，EUSIPCO 2020 で発表する予定である，分位点回帰への応用研究を簡単な数値実験とともに紹介する．
キーワード　スプライン関数，スプライン平滑化，分位点回帰，二次計画問題

[Invited Talk] Spline Basics and Its Application to Quantile Regression

Daichi KITAHARA†

† College of Information Science and Engineering, Ritsumeikan University
1–1–1 Nojihigashi, Kusatsu, Shiga, 525–8577 Japan
E-mail: † d-kita@fc.ritsumei.ac.jp

Abstract　A spline function is a piecewise polynomial which possesses certain-times continuous differentiability at the places where the polynomial pieces connect. In particular, natural cubic splines are widely used because they minimize a certain curvature. In this paper, we explain the definition and important properties of univariate and bivariate spline functions, including some proofs. We also introduce a specific and generalized construction method of spline functions as the optimal solutions to variational problems. Finally, we present our latest research on spline quantile regression, which we will present at EUSIPCO 2020, along with some simple numerical experiments.
Key words　Spline Function, Spline Smoothing, Quantile Regression, Quadratic Programming Problem

1. はじめに

スプライン関数は，何らかの連続性条件を満たすように設計された区分的多項式である [1–3]．連続な区分的多項式の中で，最も単純なものは折れ線グラフとなるが，「スプライン関数」と呼ぶ際には，2 階微分までの連続性を保証した滑らかな曲線を暗に意味する場合が多い．この理由は，現在最もよく使用されている「3 次の自然スプライン関数」が有する性質にある．3 次の自然スプライン関数は，与えられたデータ点を補間する際に，「2 階導関数の L_2 ノルム」という意味での曲率を，他のどんな C^2 級の関数よりも小さくすることができる．この滑らかさに関する最適性と区分的多項式というモデルの単純性及び柔軟性から，スプライン関数は補間・平滑化・回帰分析といった基本問題から，Computer-Aided Design (CAD)・Computer Graphics (CG) といった応用に至るまで幅広い分野で使われている．

筆者は，東京工業大学の学部 4 年次に恩師である山田功教授からスプライン関数の書籍 [1, 2] を拝借して以来，10 年近くに渡りスプライン関数の勉強と研究を少しずつ進めてきた．一見すると，既に解明し尽くされた関数のように思われるかもしれないが，筆者自身は，昨今の最適化アルゴリズムの発展を利用すれば，面白い研究がまだまだ実現可能であると考えている．スプライン関数に関しての最適化問題は，各区間における多項式の係数をパラメータとする有限次元空間の最適化問題に帰着できる．この際，区分的多項式というモデルのおかげで，通常では計算が困難となるような微分や積分の値を，閉形式として厳密に表現できる．したがって，帰着した有限次元空間の最適化問題を解くことができれば，近似計算を一切行わずに最適なスプライン関数を構成することが可能となるのである．

本稿の構成は以下の通りである．第 2 節では，主に 1 変数のスプライン関数が有する重要な性質を，証明も含めて解説する．まず記号を定義して，第 2.1 節で本稿におけるスプライン関数の定義を述べる．次に，幅広い分野で使われている「奇数次の自然スプライン関数」を紹介して，その一意性を示す (定理 1)．そして，3 次の自然スプライン関数が「曲率最小化の性質」を

有することを示す (定理 2)．第 2.2 節では与えられたデータ点に雑音が含まれている状況を想定して，データの補間ではなく平滑化を考える．そして，この場合でも 3 次の自然スプライン関数が C^2 級の関数全体の中で最適となることを示す (定理 3)．第 2.3 節では，最適解となるスプライン関数を具体的に求める方法を説明する．スプライン関数の構成方法には，いくつかのパターンが存在するが，筆者自身が最も汎用性が高いと考える方法を紹介する．第 2.4 節では，2 変数スプライン関数への 2 つの拡張方式を説明する．第 3 節では，筆者が EUSIPCO 2020 で発表する予定である分位点回帰への応用研究 [4] を，数値実験も踏まえて紹介する．最後に，第 4 節で本稿の内容を振り返る．

2. スプライン関数の基礎

実数全体と非負の整数全体の集合をそれぞれ $\mathbb{R}$ と $\mathbb{N}$ で表す．無限大を含む非負の整数 $\rho \in \mathbb{N} \cup \{\infty\}$ と，ある n 次元 ($n = 1$ または $n = 2$) 単連結領域 $\Omega \subseteq \mathbb{R}^n$ に対して，Ω 上で ρ 階連続微分可能な実数値関数全体の集合を $C^\rho(\Omega)$ で表す．微分可能な 1 変数実数値関数 f に対して，f'，f''，f'''，$f^{(l)}$ ($l \in \mathbb{N}$) でそれぞれ 1 階，2 階，3 階，l 階導関数を表す．有限次元のベクトルと行列をそれぞれボールド体の小文字と大文字で表す．

2.1 1 変数スプライン関数

スプライン関数は，「区分的に多項式として定義される連続関数」であり，多項式が切り替わる場所は節点と呼ばれる．節点 $-\infty =: \xi_{-1} < \xi_0 < \xi_1 < \cdots < \xi_b < \xi_{b+1} := \infty$ によって分割される区間を $I_0 := (-\infty, \xi_0]$，$I_i := [\xi_{i-1}, \xi_i]$ $(i = 1, 2, \ldots, b)$，$I_{b+1} := [\xi_b, \infty)$ と定義し，これら $b + 2$ 個の隣接区間を集合 $\sqcup := \{I_i\}_{i=0}^{b+1}$ として表現する．重なりのない隣接区間の集合 $\sqcup$ と 2 つの非負の整数 $\rho, d \in \mathbb{N}$ $(0 \le \rho < d)$ に対して，分割 $\sqcup$ 上の最大次数 d，滑らかさ ρ のスプライン関数全体の集合を

$$\mathcal{S}_d^\rho(\sqcup) := \{s \in C^\rho(I) \,|\, s = u_i \in \mathbb{P}_d^{\langle 1 \rangle} \text{ on } I_i \in \sqcup\} \qquad (1)$$

のように定義する (図 1)．ここで，$I := \bigcup_{I_i \in \sqcup} I_i$ であり，この場合は $I = (-\infty, \infty) = \mathbb{R}$ となる．また，$\mathbb{P}_d^{\langle 1 \rangle}$ $(\subsetneq C^\infty(\mathbb{R}))$ は次数が高々 d の 1 変数実多項式全体の集合であり，具体的には $\mathbb{P}_d^{\langle 1 \rangle} := \{u : \mathbb{R} \to \mathbb{R} : x \mapsto \sum_{k=0}^d c_k x^k \,|\, c_k \in \mathbb{R}\}$ と表される．

実際には，「スプライン関数」という用語の定義にはいくつかの流派が存在し，$\mathcal{S}_d^{d-1}(\sqcup)$ に属する関数のみを意味するという厳格な定義や，1 つ以上の節点において不連続な場合も含めて C^ρ 級未満の連続性を許容するという区分的多項式とほぼ等価な定義などがある．また，区分的に 1 次多項式と指数関数の和として定義される「指数スプライン」や，指数関数の指数部がスプライン関数になっている「対数スプライン」等も研究されているため，論文を読む際には「スプライン関数」という用語が指し示す具体的な数式を注意深く確認する必要がある．

式 (1) のスプライン関数の定義に戻ろう．ある非負の整数 $k \in \mathbb{N}$ に対して，最大次数 $2k+1$，滑らかさ $2k$ のスプライン関数 $s \in \mathcal{S}_{2k+1}^{2k}(\sqcup)$ が，両端の区間 $I_0 = (-\infty, \xi_0]$ と $I_{b+1} = [\xi_b, \infty)$ では高々 k 次の多項式として表現されるとき，s は「$2k+1$ 次の自然スプライン関数」と呼ばれる．補間すべきデータ点と同じ場所を節点に用いれば，自然スプライン関数は一意に定まる．

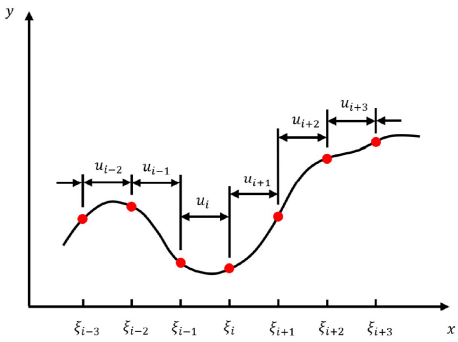

図 1　1 変数スプライン関数の節点 ξ_i と多項式 u_i の位置関係

［定理 1］$n + 1$ 個のデータ点 (x_i, y_i) $(x_0 < x_1 < \cdots < x_n)$ と n 以下の非負の整数 $k \in \mathbb{N}$ が与えられているものとする．スプライン関数の節点を $\xi_i := x_i$ $(i = 0, 1, \ldots, n$ かつ $b := n)$ と定義すると，$s(x_i) = y_i$ $(i = 0, 1, \ldots, n)$ を満たす $2k + 1$ 次の自然スプライン関数 $s \in \mathcal{S}_{2k+1}^{2k}(\sqcup)$ がただ一つ存在する．

［証明］$2k+1$ 次の自然スプライン関数 $s \in \mathcal{S}_{2k+1}^{2k}(\sqcup)$ は，両端の区間 I_0, I_{n+1} で k 次の多項式 $u_0, u_{n+1} \in \mathbb{P}_k^{\langle 1 \rangle}$ として，それ以外の区間 I_i $(i = 1, 2, \ldots, n)$ では $2k + 1$ 次の多項式 $u_i \in \mathbb{P}_{2k+1}^{\langle 1 \rangle}$ として表現される．これら $n + 2$ 個の多項式の各係数をスプライン関数の形状を決定するパラメータとして考えると，パラメータの数は $2(k+1)+n(2k+2) = (2k+2)(n+1)$ 個である．一方で各多項式が満たすべき条件を整理すると，$u_i(\xi_i) = y_i$ かつ $u_{i+1}(\xi_i) = y_i$ $(i = 0, 1, \ldots, n)$ で $2(n + 1)$ 個，$u_i^{(l)}(\xi_i) = u_{i+1}^{(l)}(\xi_i)$ $(i = 1, 2, \ldots, n-1;\, l = 1, 2, \ldots, 2k)$ で $2k(n-1)$ 個，$u_0^{(l)}(\xi_0) = u_1^{(l)}(\xi_0)$ かつ $u_n^{(l)}(\xi_n) = u_{n+1}^{(l)}(\xi_n)$ $(l = 1, 2, \ldots, k)$ で $2k$ 個，$u_1^{(l)}(\xi_0) = 0$ かつ $u_n^{(l)}(\xi_n) = 0$ $(l = k+1, k+2, \ldots, 2k)$ で $2k$ 個の条件が必要となる．これら，計 $(2k + 2)(n + 1)$ 個の条件は，パラメータに関する線形方程式として記述できるため，この線形方程式の解が一意に定まれば，データ点を補間する $2k + 1$ 次の自然スプライン関数 $s \in \mathcal{S}_{2k+1}^{2k}(\sqcup)$ も一意に定まる．解の一意性を証明するためには，$y_i = 0$ $(i = 0, 1, \ldots, n)$ である場合に自然スプライン関数 $s \in \mathcal{S}_{2k+1}^{2k}(\sqcup)$ が恒等的に 0 となる，つまり $s(x) \equiv 0$ となることを示せば十分である．このために，まず $s^{(k+1)}(x) \equiv 0$ であることを示す．$s^{(l)}(\xi_0) = s_n^{(l)}(\xi_n) = 0$ $(l = k+1, k+2, \ldots, 2k)$ と部分積分法を繰り返し用いると，

$$\begin{aligned}
&\int_{\mathbb{R}} \left|s^{(k+1)}(x)\right|^2 \mathrm{d}x = \int_{\xi_0}^{\xi_n} s^{(k+1)}(x) s^{(k+1)}(x)\,\mathrm{d}x \\
&= \left[s^{(k)}(x) s^{(k+1)}(x)\right]_{\xi_0}^{\xi_n} - \int_{\xi_0}^{\xi_n} s^{(k)}(x) s^{(k+2)}(x)\,\mathrm{d}x \\
&= 0 - \left[s^{(k-1)}(x) s^{(k+2)}(x)\right]_{\xi_0}^{\xi_n} + \int_{\xi_0}^{\xi_n} s^{(k-1)}(x) s^{(k+3)}(x)\,\mathrm{d}x \\
&= \cdots \\
&= (-1)^{k-1} \left(\left[s'(x) s^{(2k)}(x)\right]_{\xi_0}^{\xi_n} - \sum_{i=1}^{n} \int_{\xi_{i-1}}^{\xi_i} s'(x) s^{(2k+1)}(x)\,\mathrm{d}x \right) \\
&= (-1)^k \sum_{i=1}^{n} c_i \int_{\xi_{i-1}}^{\xi_i} s'(x)\,\mathrm{d}x = (-1)^k \sum_{i=1}^{n} c_i (y_i - y_{i-1}) = 0
\end{aligned}$$

が導かれるため，$s^{(k+1)}(x) \equiv 0$ が示される．ここで，最終行の

式変形では，各区間 I_i $(i=1,2,\dots,n)$ で s が $2k+1$ 次の多項式 $u_i \in \mathbb{P}^{\langle 1\rangle}_{2k+1}$ として表現されるため，$s^{(2k+1)}(x) \equiv u_i^{(2k+1)}(x)$ が恒等的にある定数 $c_i \in \mathbb{R}$ となることと，$s(\xi_i) = s(x_i) = y_i = 0$ $(i=0,1,\dots,n)$ であることを用いた．s が連続関数であることと $s^{(k+1)}(x) \equiv 0$ を満たすことから，s は高々 k 次の多項式となるが，$k+1$ 個以上の異なる位置の点 $(x_i, 0)$ $(i=0,1,\dots,n)$ を補間することから，$s(x) \equiv 0$ に限定されることが分かる．■

自然スプライン関数の中でも，特に 3 次の自然スプライン関数が様々な応用分野で用いられている [1, 2]．これは，3 次の自然スプライン関数によってデータを補間すると，「2 階導関数の L_2 ノルム」という意味での曲率を最小化できるためである．

［定理 2］$n+1$ 個のデータ点 (x_i, y_i) $(x_0 < x_1 < \cdots < x_n$ かつ $n \geq 1)$ が与えられているものとする．以下の変分問題(注1)

$$\underset{f\in C^2(\mathbb{R})}{\text{minimize}} \int_{\mathbb{R}} \left|f''(x)\right|^2 \mathrm{d}x \quad \text{subject to } \forall i\ f(x_i) = y_i \qquad (2)$$

はただ一つの最適解を有し，節点が $\xi_i = x_i$ $(i=0,1,\dots,n)$ である 3 次の自然スプライン関数 $s \in \mathcal{S}_3^2(\sqcup)$ が最適解となる．

［証明］定理 1 から，節点が $\xi_i = x_i$ $(i=0,1,\dots,n)$ であり，$s(x_i) = y_i$ $(i=0,1,\dots,n)$ を満たす 3 次の自然スプライン関数 $s \in \mathcal{S}_3^2(\sqcup)$ が常にただ一つ存在する．3 次の自然スプライン関数とは別に，ある関数 $f \in C^2(\mathbb{R})$ が $f(x_i) = y_i$ $(i=0,1,\dots,n)$ を満たしている状況を考える．このとき，$g(x) := f(x) - s(x)$ と定義すれば，$f''(x) = g''(x) + s''(x)$ が成り立つ．ここから，

$$\begin{aligned}\int_{\mathbb{R}} \left|f''(x)\right|^2 \mathrm{d}x &= \int_{\mathbb{R}} \left(\left|g''(x)\right|^2 + 2g''(x)s''(x) + \left|s''(x)\right|^2\right) \mathrm{d}x \\ &= \int_{\mathbb{R}} \left|g''(x)\right|^2 \mathrm{d}x + \int_{\mathbb{R}} \left|s''(x)\right|^2 \mathrm{d}x \geq \int_{\mathbb{R}} \left|s''(x)\right|^2 \mathrm{d}x\end{aligned}$$

が導かれ s の最適性が示される．1 行目から 2 行目の式変形は，

$$\begin{aligned}&\int_{\mathbb{R}} g''(x)s''(x)\,\mathrm{d}x = \int_{\xi_0}^{\xi_n} g''(x)s''(x)\,\mathrm{d}x \\ &= \left[g'(x)s''(x)\right]_{\xi_0}^{\xi_n} - \sum_{i=1}^{n} \int_{\xi_{i-1}}^{\xi_i} g'(x)s'''(x)\,\mathrm{d}x \\ &= -\sum_{i=1}^{n} c_i \int_{\xi_{i-1}}^{\xi_i} g'(x)\,\mathrm{d}x = -\sum_{i=1}^{n} c_i \left(g(\xi_i) - g(\xi_{i-1})\right) = 0\end{aligned}$$

から導かれている．ここで，各区間 I_i $(i=1,2,\dots,n)$ で s''' が定数 $c_i \in \mathbb{R}$ となることと，$g(\xi_i) = g(x_i) = f(x_i) - s(x_i) = y_i - y_i = 0$ $(i=0,1,\dots,n)$ となることを用いている．最後に，最適解の一意性を示す．s とは異なる関数 f も最適解であるならば，$\int_{\mathbb{R}} |f''(x)|^2\,\mathrm{d}x = \int_{\mathbb{R}} |s''(x)|^2\,\mathrm{d}x$ から $\int_{\mathbb{R}} |g''(x)|^2\,\mathrm{d}x = 0$ でなければならない．したがって，$g''(x) \equiv 0$ が成り立つ．g が連続関数であることと $g''(x) \equiv 0$ を満たすことから，g は高々 1 次の多項式となるが，2 個以上の異なる位置の点 $(x_i, 0)$ $(i=0,1,\dots,n)$ を補間することから，$g(x) \equiv 0$ に限定される．したがって $f(x) \equiv s(x)$ となり，f と s は完全に一致する．■

(注1)：探索変数が関数である最適化問題を「変分問題」と呼ぶ．2 階連続微分可能な関数の集合 $C^2(\mathbb{R})$ は凸集合ではあるものの閉集合ではないので，最小化する汎関数が凸であっても最適解が存在しない場合も多いので注意が必要である．

2.2 補間から平滑化への拡張

第 2.1 節で述べたように，3 次の自然スプライン関数を用いてデータ点を補間すれば，2 階導関数の L_2 ノルムの意味で最も滑らかな曲線を構成できる．しかしながら，真の関数値に雑音 $\epsilon_i \in \mathbb{R}$ が加わった観測値 $y_i = f^\star(x_i) + \epsilon_i$ から元の連続関数 $f^\star$ を推定する場合には，仮に $f^\star$ が非常に滑らかな関数であったとしても，式 (2) の変分問題は雑音の影響を強く受けてしまい，最適解でも激しく振動する関数となってしまう．元の滑らかな関数 $f^\star$ を高精度に推定するためには，各データ点との誤差をある程度許容しながら，最も滑らかな関数を探索すればよい．実は，このようにデータを平滑化する場合においても，3 次の自然スプライン関数が最適解となることを簡単に証明できる．

［定理 3］$n+1$ 個の雑音を含んだデータ点 (x_i, y_i) $(x_0 < x_1 < \cdots < x_n$ かつ $n \geq 1)$ が与えられているものとする．変分問題

$$\underset{f\in C^2(\mathbb{R})}{\text{minimize}} \sum_{i=0}^{n} |y_i - f(x_i)|^2 + \lambda \int_{\mathbb{R}} \left|f''(x)\right|^2 \mathrm{d}x \qquad (3)$$

はただ一つの最適解を有し，節点が $\xi_i = x_i$ $(i=0,1,\dots,n)$ である 3 次の自然スプライン関数 $s \in \mathcal{S}_3^2(\sqcup)$ が最適解となる．ここで，$\lambda > 0$ は「平滑化パラメータ」と呼ばれる重みである．

［証明］背理法で証明する．式 (3) の変分問題の最適解 $f \in C^2(\mathbb{R})$ が 3 次の自然スプライン関数ではないと仮定する．定理 1 から，節点が $\xi_i = x_i$ $(i=0,1,\dots,n)$ であり，$s(x_i) = f(x_i)$ $(i=0,1,\dots,n)$ を満たす 3 次の自然スプライン関数 $s \in \mathcal{S}_3^2(\sqcup)$ が常にただ一つ存在する．このとき，式 (3) の第 1 項に関しては，$\sum_{i=0}^n |y_i - s(x_i)|^2 = \sum_{i=0}^n |y_i - f(x_i)|^2$ が成り立つ．さらに第 2 項に関しては，定理 2 から $\int_{\mathbb{R}} |s''(x)|^2\,\mathrm{d}x < \int_{\mathbb{R}} |f''(x)|^2\,\mathrm{d}x$ が成り立つ．結果として，f より s の方が式 (3) の汎関数の値を小さくでき，f が最適解であることと矛盾する．したがって，最適解は 3 次の自然スプライン関数となる．最後に，最適解の一意性を証明するために，異なる 3 次の自然スプライン関数 $s_1, s_2 \in \mathcal{S}_3^2(\sqcup)$ がどちらも最適解であると仮定する．$s_1 \neq s_2$ であることと定理 1 から，少なくとも 1 つ以上の節点 x_i において $s_1(x_i) \neq s_2(x_i)$ が成り立つ．ここで，任意の $\kappa \in (0,1)$ に対して $s_3 := \kappa s_1 + (1-\kappa)s_2$ と定義すれば，$s_3 \in \mathcal{S}_3^2(\sqcup)$ $(\subsetneq C^2(\mathbb{R}))$ となる．また，式 (3) の第 1 項は $(f(x_0), f(x_1), \dots, f(x_n))^{\mathrm{T}} \in \mathbb{R}^{n+1}$ に関して狭義凸関数であるため，$\sum_{i=0}^n |y_i - s_3(x_i)|^2 < \kappa \sum_{i=0}^n |y_i - s_1(x_i)|^2 + (1-\kappa) \sum_{i=0}^n |y_i - s_2(x_i)|^2$ が成立する．第 2 項は $f \in C^2(\mathbb{R})$ に関して凸関数であり，$\int_{\mathbb{R}} |s_3''(x)|^2\,\mathrm{d}x \leq \kappa \int_{\mathbb{R}} |s_1''(x)|^2\,\mathrm{d}x + (1-\kappa) \int_{\mathbb{R}} |s_2''(x)|^2\,\mathrm{d}x$ が成立する．すると，$\sum_{i=0}^n |y_i - s_3(x_i)|^2 + \lambda \int_{\mathbb{R}} |s_3''(x)|^2\,\mathrm{d}x < \sum_{i=0}^n |y_i - s_1(x_i)|^2 + \lambda \int_{\mathbb{R}} |s_1''(x)|^2\,\mathrm{d}x = \sum_{i=0}^n |y_i - s_2(x_i)|^2 + \lambda \int_{\mathbb{R}} |s_2''(x)|^2\,\mathrm{d}x$ が導かれ，s_3 の方が式 (3) の値を小さくでき，s_1 と s_2 が最適解であることと矛盾してしまうため，最適解はただ一つとなる．■

2.3 変分問題の最適解となるスプライン関数の導出

第 2.1 節及び第 2.2 節でスプライン関数の最適性を説明したので，次は最適解となるスプライン関数を実際に導出していく．スプライン関数の構成方法 [1–3] として，切断べき関数，差分商，バーンスタイン–ベジェ形式など，いくつかのパターンが存在し，この構成方法の多様さが初学者に対する高い壁として

立ちはだかる．しかしながら，これらの構成方法は証明の容易さ，変数の次元，数値的安定性，計算速度などが異なるだけで，表現可能なスプライン関数全体の集合は同じであるということを忘れてはならない．本稿では，筆者自身が最も汎用性が高く，理解もしやすいと考えるスプライン関数の構成方法を紹介する．

まず，両端の区間 $I_0=(-\infty,\xi_0]$ と $I_{b+1}=[\xi_b,\infty)$ に関しては，一般的にデータ点に対する「外挿」が行われることになる．何らかの強い理論的妥当性が存在しない限り，外挿結果を当てにしてはならない．したがって，これ以降は，両端の 2 区間を無視した b 個の隣接区間の集合 $\sqcup_b:=\{I_i\}_{i=1}^{b}$ のみを取り扱う．議論の対象となる区間も $\mathbb{R}$ ではなく，$I=\bigcup_{I_i\in\sqcup_b}I_i=[\xi_0,\xi_b]$ となる．さて，筆者が議論を展開しやすいと考えるスプライン関数 $s\in\mathcal{S}_d^\rho(\sqcup_b)$ $(2\leq\rho<d)$ の構成方法は，各区間 I_i の長さ $h_i:=\xi_i-\xi_{i-1}$ $(i=1,2,\dots,n)$ で正規化した以下の多項式表現

$$s(x):=u_i(x):=\sum_{k=0}^{d}c_k^{\langle i\rangle}\left(\frac{x-\xi_{i-1}}{h_i}\right)^k\quad\text{for } x\in I_i \tag{4}$$

である．ここで，$c_k^{\langle i\rangle}\in\mathbb{R}$ $(k=0,1,\dots,d)$ は高々 d 次の多項式 $u_i\in\mathbb{P}_d^{\langle1\rangle}$ $(i=1,2,\dots,b)$ の係数を表している．式 (4) の係数 $c_k^{\langle i\rangle}$ は全部で $b(d+1)$ 個あり，これらをまとめたベクトル $\boldsymbol{c}=(c_d^{\langle1\rangle},c_{d-1}^{\langle1\rangle},\dots,c_0^{\langle1\rangle},c_d^{\langle2\rangle},c_{d-1}^{\langle2\rangle},\dots,c_0^{\langle2\rangle},\dots,c_0^{\langle b\rangle})^{\mathrm{T}}\in\mathbb{R}^{b(d+1)}$ をスプライン関数 $s\in\mathcal{S}_d^\rho(\sqcup_b)$ の形状を定めるパラメータとする．ここで $x_i\in I$ $(i=0,1,\dots,n)$ が成立していれば，式 (4) から $(s(x_0),s(x_1),\dots,s(x_n))^{\mathrm{T}}=\boldsymbol{A}\boldsymbol{c}$ を満たすようなスパース行列 $\boldsymbol{A}\in\mathbb{R}^{(n+1)\times b(d+1)}$ が簡単に作成可能となる．

次に，滑らかさに関する正則化項 $\int_I|s''(x)|^2\,\mathrm{d}x$ の値を具体的に計算していく．$s\in\mathcal{S}_d^\rho(\sqcup_b)$ は区分的に定義されているため，

$$\int_I\left|s''(x)\right|^2\mathrm{d}x=\sum_{i=1}^{b}\int_{I_i}\left|s''(x)\right|^2\mathrm{d}x \tag{5}$$

のように積分区間を分割する．分割された各区間 I_i における積分値は，式 (4) の表現を用いれば，高校数学レベルの計算で

$$\int_{I_i}\left|s''(x)\right|^2\mathrm{d}x=\sum_{k=2}^{d}\sum_{l=2}^{d}\frac{k(k-1)l(l-1)}{h_i^3(k+l-3)}c_k^{\langle i\rangle}c_l^{\langle i\rangle} \tag{6}$$

のように求まる．式 (5) と式 (6) から $\int_I|s''(x)|^2\,\mathrm{d}x=\boldsymbol{c}^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{c}$ を満たす半正定値対称行列 $\boldsymbol{Q}\in\mathbb{R}^{b(d+1)\times b(d+1)}$ が作成される．

以上の結果から，式 (3) の汎関数の値は $\|\boldsymbol{y}-\boldsymbol{A}\boldsymbol{c}\|_2^2+\lambda\boldsymbol{c}^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{c}$ のように表現できる．ここで，$\boldsymbol{y}:=(y_0,y_1,\dots,y_n)^{\mathrm{T}}\in\mathbb{R}^{n+1}$ であり，$\|\cdot\|_2$ はベクトルの ℓ_2 ノルムを表す．パラメータ $\boldsymbol{c}$ は上記の関数値をただ最小化すれば良いという訳ではなく，s が I 上で ρ 階連続微分可能という条件を満たさなければならない．式 (4) から，隣接多項式 (u_i,u_{i+1}) $(i=1,2,\dots,b-1)$ の微分も含めた連続性条件 $u_i^{(l)}(\xi_i)=u_{i+1}^{(l)}(\xi_i)$ $(l=0,1,\dots,\rho)$ は，

$$\frac{1}{h_i^l}\sum_{k=l}^{d}\frac{k!}{(k-l)!}c_k^{\langle i\rangle}-\frac{l!}{h_{i+1}^l}c_l^{\langle i+1\rangle}=0\quad(l=0,1,\dots,\rho) \tag{7}$$

のように表される．式 (7) から $s\in C^\rho(I)\Leftrightarrow\boldsymbol{H}\boldsymbol{c}=\boldsymbol{0}$ を満たすようなスパース行列 $\boldsymbol{H}\in\mathbb{R}^{(b-1)(\rho+1)\times b(d+1)}$ を作成できる．

結果として，式 (3) において関数空間を $C^2(\mathbb{R})$ から $\mathcal{S}_d^\rho(\sqcup_b)$ に制限した変分問題は，以下の有限次元空間の最適化問題

$$\underset{\boldsymbol{c}\in\mathbb{R}^{b(d+1)}}{\text{minimize}}\ \|\boldsymbol{y}-\boldsymbol{A}\boldsymbol{c}\|_2^2+\lambda\boldsymbol{c}^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{c}\quad\text{subject to }\boldsymbol{H}\boldsymbol{c}=\boldsymbol{0} \tag{8}$$

に帰着される．式 (8) の問題は「二次計画問題」であるため，二次計画法のアルゴリズムで比較的高速に最適解が求まる．(注2) 最後に，節点を $\xi_i=x_i$ $(i=0,1,\dots,n$ かつ $b=n)$，次数を $d\geq3$，滑らかさを $\rho=2$ として式 (8) の問題を解けば，式 (3) の最適解となる 3 次の自然スプライン関数を導出できる．

2.4　2 変数スプライン関数への拡張

スプライン関数の多変数への拡張はそれほど簡単ではなく，安易に足を踏み入れると痛い目に遭う恐れがある．ここでは，2 変数スプライン関数への異なる拡張方式を簡単に紹介する．

まず，よく知られている 2 変数スプライン関数の定義を紹介する．これは厳密には，「双スプライン関数」という関数である．格子状に並んでいる節点 $-\infty<\xi_0<\xi_1<\cdots<\xi_{b_1}<\infty$ と $-\infty<\eta_0<\eta_1<\cdots<\eta_{b_2}<\infty$ で分割される長方形領域を $R_{i,j}:=[\xi_{i-1},\xi_i]\times[\eta_{j-1},\eta_j]$ $(i=1,2,\dots,b_1;\ j=1,2,\dots,b_2)$ と定義し，b_1b_2 個の隣接領域を集合 $\square_{b_1,b_2}:=\{R_{i,j}\}_{j=1,2,\dots,b_2}^{i=1,2,\dots,b_1}$ として表現する．重なりのない隣接長方形領域の集合 $\square_{b_1,b_2}$ と 2 つの非負の整数 $\rho,d\in\mathbb{N}$ $(0\leq\rho<d)$ に対して，分割 $\square_{b_1,b_2}$ 上の最大次数 d，滑らかさ ρ の双スプライン関数全体の集合を

$$\mathcal{S}_d^\rho(\square_{b_1,b_2}):=\{s\in C_{\rho,\rho}^{2\rho}(R)\,|\,s=u_{i,j}\in\mathbb{P}_{d,d}^{\langle2\rangle}\text{ on }R_{i,j}\in\square_{b_1,b_2}\} \tag{9}$$

のように定義する．ここで，$R:=\bigcup_{R_{i,j}\in\square_{b_1,b_2}}R_{i,j}=[\xi_0,\xi_{b_1}]\times[\eta_0,\eta_{b_2}]$ である．$\mathbb{P}_{d,d}^{\langle2\rangle}$ $(\subsetneq C^\infty(\mathbb{R}^2))$ は，各変数に関して次数が高々 d の 2 変数実多項式全体の集合であり，具体的には $\mathbb{P}_{d,d}^{\langle2\rangle}:=\{u:\mathbb{R}^2\to\mathbb{R}:(x,y)\mapsto\sum_{k=0}^{d}\sum_{l=0}^{d}c_{k,l}x^ky^l\,|\,c_{k,l}\in\mathbb{R}\}$ となる．$C_{\rho,\rho}^{2\rho}(R)$ $(\subsetneq C^\rho(R))$ は，R 上で $\frac{\partial^{i+j}f}{\partial x^i\partial y^j}$ $(i=0,1,\dots,\rho;\ j=0,1,\dots,\rho)$ がどれも連続となる 2 変数関数全体の集合である．双 3 次自然スプライン(注3)で格子状のデータ $(x_i,y_j,z_{i,j})$ $(x_0<x_1<\cdots<x_{n_1};\ y_0<y_1<\cdots<y_{n_2})$ を補間すると，曲率 $\iint_R|\frac{\partial^4f}{\partial x^2\partial y^2}|^2\,\mathrm{d}x\mathrm{d}y$ が最小化されるが，これは一種の「4 階導関数の L_2 ノルム」であり，定理 2 の完全な一般化(注4)ではない．双スプライン関数は非格子状のデータに対して使い勝手が悪く，$y=x$ という直線上では多項式の次数が 2 倍となる問題もある．

次に，より柔軟な 2 変数スプライン関数の定義を紹介する．まず非格子状のデータにも対応するために，対象領域を 2 次元の多角形領域(注5)Ω $(\subsetneq\mathbb{R}^2)$ とする．2 次元多角形領域 Ω を最も細かく分割するには，三角形分割を用いればよい．3 つの頂点 $\boldsymbol{v}_k:=(x_k,y_k)\in\Omega$ $(k=1,2,3)$ によって定まる三角形領域を $T:=\{t_1\boldsymbol{v}_1+t_2\boldsymbol{v}_2+t_3\boldsymbol{v}_3\,|\,t_1,t_2,t_3\in[0,1]\text{ and }t_1+t_2+t_3=1\}$

(注2)：式 (4) の構成方法でも，次数 d が大きくなると数値的不安定性が生じてしまう．筆者の経験では，$d=7$ 程度までであれば安定して最適解を求められる．

(注3)：双 3 次自然スプライン s の節点は $(\xi_i,\eta_j)=(x_i,y_j)$ で，$x=\xi_0,\xi_{n_1}$ のとき $\forall y\ \frac{\partial^2s}{\partial x^2}(x,y)=0$，$y=\eta_0,\eta_{n_2}$ のとき $\forall x\ \frac{\partial^2s}{\partial y^2}(x,y)=0$ となる．

(注4)：実は，任意の位置の有限個のデータ点を全て補間する C^2 級の 2 変数関数の中で，「2 階導関数の L_2 ノルム」$\iint_{\mathbb{R}^2}(|\frac{\partial^2f}{\partial x^2}|^2+2|\frac{\partial^2f}{\partial x\partial y}|^2+|\frac{\partial^2f}{\partial y^2}|^2)\,\mathrm{d}x\mathrm{d}y$ を最小にする関数は 2 変数スプライン関数にならず，区分的な関数にもならない．

(注5)：長方形領域 R も多角形領域であるため，Ω の具体例として取り扱える．

と定義する．重なりのない隣接三角形領域の集合 $\Delta_b := \{T_i\}_{i=1}^b$ が $\bigcup_{T_i\in\Delta_b} T_i = \Omega$ を満たし，$T_i \cap T_j$ $(i \neq j)$ は空集合，共通の辺，共通の頂点のいずれかであるとする．このとき，分割 Δ_b 上の最大次数 d，滑らかさ ρ の 2 変数スプライン関数全体の集合を

$$\mathcal{S}_d^\rho(\Delta_b) := \{s \in C^\rho(\Omega) \,|\, s = u_i \in \mathbb{P}_d^{\langle 2\rangle} \text{ on } T_i \in \Delta_b\} \quad (10)$$

のように定義する．ここで，$\mathbb{P}_d^{\langle 2\rangle}$ $(\subsetneq \mathbb{P}_{d,d}^{\langle 2\rangle})$ は次数が高々 d の 2 変数実多項式全体の集合であり，具体的には $\mathbb{P}_d^{\langle 2\rangle} := \{u : \mathbb{R}^2 \to \mathbb{R} : (x,y) \mapsto \sum_{k=0}^d \sum_{l=0}^{d-k} c_{k,l} x^k y^l \,|\, c_{k,l} \in \mathbb{R}\}$ となる．式 (10) の 2 変数スプライン関数は，式 (9) の双スプライン関数より柔軟な表現が可能となるが，バーンスタイン–ベジェ形式 [3] を理解できない限り，構成することが難しい．また，分割 Δ_b が全ての $R_{i,j}$ を分割できていれば，$\mathcal{S}_d^\rho(\square_{b_1,b_2}) \subsetneq \mathcal{S}_{2d}^\rho(\Delta_b)$ が成立する．

3. 分位点回帰への応用

ここでは，筆者が EUSIPCO 2020 で発表する予定である，スプライン関数による分位点回帰の研究 [4] を簡単に紹介する．

3.1 平均値回帰・中央値回帰

2 つの確率変数の組 (X,Y) の $n+1$ 個の実現値 $(x_i, y_i) \in \mathbb{R}^2$ $(i = 0, 1, \ldots, n)$ が与えられているものとする．ここで，(X,Y) の同時確率密度関数 $f_{X,Y}$ は，ある長方形領域 $\Omega := (x_{\inf}, x_{\sup}) \times (y_{\inf}, y_{\sup}) \subseteq \mathbb{R}^2$ 上で正の値をとり，$\iint_\Omega f_{X,Y}(x,y)\,\mathrm{d}x\mathrm{d}y = 1$ を満たすとする．この場合には，「X が与えられたときの Y の条件付き確率密度関数」が $f_{Y|X}(y|x) := f_{X,Y}(x,y)/f_X(x) := f_{X,Y}(x,y)/\int_{y_{\inf}}^{y_{\sup}} f_{X,Y}(x,y)\,\mathrm{d}y > 0$ のように定義される．

観測された 2 次元データ $\{(x_i, y_i)\}_{i=0}^n$ から，確率変数の組 (X,Y) の傾向を探ろうとする (2 次元データの縮約という) 際に，傾向の連続的な変化を捉えるために，以下の最小二乗回帰

$$\underset{\boldsymbol{\theta}}{\text{minimize}} \sum_{i=0}^n w_i |y_i - r_{\boldsymbol{\theta}}(x_i)|^2 \quad (11)$$

がよく用いられる．ここで，$r_{\boldsymbol{\theta}}$ はある回帰モデルを表し，$\boldsymbol{\theta}$ はそのパラメータ(注6)である．第 i データ (x_i, y_i) の重み $w_i > 0$ は，Y の条件付き分散 $V[Y|X = x_i]$ や，X の確率密度関数値 $f_X(x_i)$ から設計される場合が多い．式 (11) でデータ数が無限大と思えるほどたくさんある理想的な状況，つまり $n \to \infty$ を考える．実はこの理想状況では，式 (11) の最適解 $r_{\boldsymbol{\theta}^*}$ は「X が与えられたときの Y の条件付き平均値」に近づいていく(注7)：

$$r_{\boldsymbol{\theta}^*}(x) \to \mu_Y(x) := E[Y|X = x] = \int_{y_{\inf}}^{y_{\sup}} y f_{Y|X}(y|x)\,\mathrm{d}y$$

ただし，$r_{\boldsymbol{\theta}}$ が μ_Y を表現可能でなければならない．

式 (11) で用いている二乗誤差は外れ値に敏感であるため，裾の重い分布に対しては最適解 $r_{\boldsymbol{\theta}^*}$ の形状が大きく変動しやすい．外れ値に対して頑健な手法として，以下の最小絶対値回帰

$$\underset{\boldsymbol{\theta}}{\text{minimize}} \sum_{i=0}^n w_i |y_i - r_{\boldsymbol{\theta}}(x_i)| \quad (12)$$

が有名である．式 (12) の絶対値誤差は，二乗誤差と違い外れ値を過剰に評価しないため，裾の重い分布に対しても最適解 $r_{\boldsymbol{\theta}^*}$ が変動しづらくなる．また，$n \to \infty$ であるとき，最適解 $r_{\boldsymbol{\theta}^*}$ は「X が与えられたときの Y の条件付き中央値」に近づいていく：

$$r_{\boldsymbol{\theta}^*}(x) \to m_Y(x) \text{ satisfying } \int_{y_{\inf}}^{m_Y(x)} f_{Y|X}(y|x)\,\mathrm{d}y = 0.5$$

ただし，$r_{\boldsymbol{\theta}}$ が m_Y を表現可能でなければならない．

(注6)：$r_{\boldsymbol{\theta}}$ を多項式とした場合，$\boldsymbol{\theta}$ は多項式の係数をまとめたベクトルとなる．
(注7)：ここでは議論を分かりやすくするために「収束」を意味する記号 '$\to$' を用いているが，実際には確率変数の収束であるため，注意深い議論が必要となる．

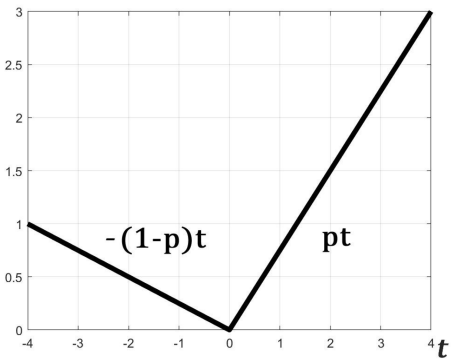

図 2　非対称絶対値関数 J_p $(p = 0.75)$

3.2 分位点回帰

第 3.1 節の中央値回帰の結果を一般化することで，任意の分位点，つまり任意のパーセンタイル曲線を求めることができる．まず，「X が与えられたときの Y の条件付き累積分布関数」を

$$F_{Y|x}(y) := \int_{y_{\inf}}^{y} f_{Y|X}(t|x)\,\mathrm{d}t \quad \text{for } y \in (y_{\inf}, y_{\sup})$$

のように定義する．全ての $(x,y) \in \Omega$ に対して $f_{Y|X}(y|x) > 0$ が成り立つことから，$F_{Y|x}$ は狭義単調増加関数となり，その逆関数 $F_{Y|x}^{-1} : (0,1) \ni p \mapsto F_{Y|x}^{-1}(p) \in (y_{\inf}, y_{\sup})$ が一意に定義される．この逆関数は，「X が与えられたときの Y の条件付き分位関数」$Q_{Y|x}$ と呼ばれる：

$$Q_{Y|x}(p) := F_{Y|x}^{-1}(p) \quad \text{for } p \in (0,1)$$

ここで，$q_{p,Y}(x) := Q_{Y|x}(p)$ $(x \in (x_{\inf}, x_{\sup}))$ と定義すると，$q_{p,Y}$ は「X が与えられたときの Y の条件付き第 p 分位点」と呼ばれる．$p = 0.5$ のとき，$q_{p,Y}$ は条件付き中央値 m_Y となる．

さらに $p \in (0,1)$ に対して，図 2 のような非対称絶対値関数

$$J_p(t) := \begin{cases} pt & \text{if } t \geq 0, \\ -(1-p)t & \text{if } t < 0, \end{cases}$$

を定義して，以下の最適化問題

$$\underset{\boldsymbol{\theta}}{\text{minimize}} \sum_{i=0}^n w_i J_p(y_i - r_{\boldsymbol{\theta}}(x_i)) \quad (13)$$

を考える [5]．すると $n \to \infty$ であるとき，式 (13) の最適解 $r_{\boldsymbol{\theta}^*}$ は Y の条件付き第 p 分位点 $q_{p,Y}$ に近づいていく：

$$r_{\boldsymbol{\theta}^*}(x) \to q_{p,Y}(x) \text{ satisfying } \int_{y_{\inf}}^{q_{p,Y}(x)} f_{Y|X}(y|x)\,\mathrm{d}y = p$$

ただし，$r_{\boldsymbol{\theta}}$ が $q_{p,Y}$ を表現可能でなければならない．最適解 $r_{\boldsymbol{\theta}^*}$ が $q_{p,Y}$ に近づいていく理由は，最適解 $r_{\boldsymbol{\theta}^*}$ に対して下側の誤差 $e_i = y_i - r_{\boldsymbol{\theta}^*}(x_i) < 0$ の総量と上側の誤差 $e_i = y_i - r_{\boldsymbol{\theta}^*}(x_i) > 0$ の総量が，$p : 1-p$ の比率で釣り合っているためである．

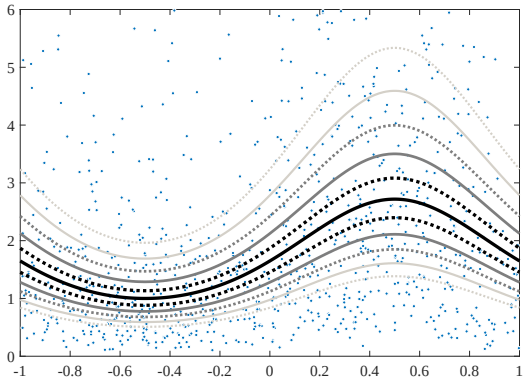

図 3　第 0.25/0.3/0.35/0.4/0.45/0.5/0.55/0.6/0.65/0.7/0.75 分位点

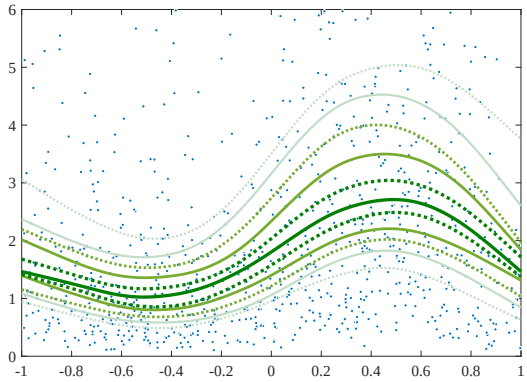

図 4　式 (14) による各分位点の推定結果

3.3　スプライン関数を用いた分位点回帰

分位点回帰の性能を最大限発揮するためには，様々な形状の分位点 $q_{p,Y}$ を表現できる，柔軟な回帰モデル $r_{\boldsymbol{\theta}}$ を用いるべきである．本研究では，節点を細かく設定したスプライン関数 $s \in \mathcal{S}_d^{\rho}(\sqcup_b)$ を回帰モデル $r_{\boldsymbol{\theta}}$ として用いることを考える．節点を細かくしているためスプライン関数は高い表現能力を持つが，一方で標本サイズ n が比較的小さい場合にデータに過適合してしまうという問題が生じる．過適合を回避するため，真の分位点 $q_{p,Y}$ がある程度滑らかな曲線であることを仮定し，式 (3) と同様の滑らかさに関する正則化項を加えた以下の変分問題

$$\underset{s\in\mathcal{S}_d^{\rho}(\sqcup_b)}{\text{minimize}} \sum_{i=0}^{n} w_i J_p(y_i - s(x_i)) + \lambda \int_I \left|s''(x)\right|^2 \mathrm{d}x \quad (14)$$

を考える．そして，式 (14) の最適解を分位点回帰の結果とする．式 (4) の表現を用いれば式 (14) の問題は，各区間における多項式の係数をまとめたベクトル $\boldsymbol{c} \in \mathbb{R}^{b(d+1)}$ に関する最適化問題

$$\underset{\boldsymbol{c}\in\mathbb{R}^{b(d+1)}}{\text{minimize}} \sum_{i=0}^{n} w_i J_p(y_i - \boldsymbol{a}_i^{\mathrm{T}}\boldsymbol{c}) + \lambda \boldsymbol{c}^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{c} \quad \text{subject to } \boldsymbol{H}\boldsymbol{c} = \boldsymbol{0}$$

へと帰着される．そして，この有限次元空間の凸最適化問題を何らかのアルゴリズムで解けば，式 (14) の最適解を構成できる．

最後に，式 (14) を用いて実際に分位点回帰を行う．実験対象とした Y の条件付き確率密度関数は，以下の対数正規分布

$$f_{Y|X}(y|x) := \frac{1}{\sqrt{2\pi}y} e^{-\frac{(\log y - \hat{\mu}(x))^2}{2}} \quad \text{for } y \in (0,\infty) \quad (15)$$

である．ここで，$\hat{\mu}(x) := 0.5\sin(\pi x) + 0.5$ とした．ランダムに生成した 1,000 個の実現値 $\{(x_i, y_i)\}_{i=0}^{999}$ から，11 通りの $p \in \{0.25, 0.3, 0.35, 0.4, 0.45, 0.5, 0.55, 0.6, 0.65, 0.7, 0.75\}$ に関して図 3 に表示されている第 p 分位点 $q_{p,Y}$ を推定する．各実現値 (x_i, y_i) は，まず $x_i \in [-1, 1] =: I$ を以下の混合正規分布

$$f_X(x) := \frac{0.3}{\sqrt{2\pi}\sigma_1} e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}} + \frac{0.7}{\sqrt{2\pi}\sigma_2} e^{-\frac{(x-\mu_2)^2}{2\sigma_2^2}}$$

から生成し，その後 y_i を式 (15) の対数正規分布から生成した．ここで，$(\mu_1, \mu_2, \sigma_1, \sigma_2) = (-0.7, 0.4, 0.3, 0.45)$ とした．1 変数スプライン関数の設定は，$d = 5$，$\rho = 2$，$b = 40$，$h_i = \frac{2}{b} = \frac{1}{20}$ $(i = 1, 2, \ldots, 40)$ とした．図 4 は，式 (14) による各分位点 $q_{p,Y}$ の推定結果を表している．ここで，青い点は各実現値 (x_i, y_i) の位置を表している．図 4 を見てみると，各分位点の推定は概ね成功しているものの，隣接分位点のそれぞれの推定結果が互いに近づきすぎたり，離れすぎたりする箇所があることが分かる．これは，全ての分位点 $q_{p,Y}$ $(p \in (0,1))$ はただ一つの条件付き確率密度関数 $f_{Y|X}$ から定義されているにも関わらず，式 (14) では各分位点を独立に推定していることが原因と考えられる．

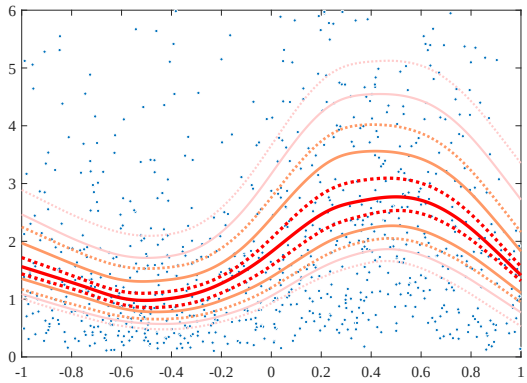

図 5　EUSIPCO 2020 で発表予定の同時分位点回帰による推定結果

図 5 は，EUSIPCO 2020 で発表予定の同時分位点回帰 [4] の結果を表している．提案法では隣接分位点の類似性も考慮して各分位点を同時に推定するため，高精度な回帰結果が得られる．

4.　おわりに

本稿では，1 変数スプライン関数の定義及び最適性を説明し，最適解としてのスプライン関数の具体的な構成方法を紹介した．2 変数スプライン関数の異なる定義に触れた後に，分位点回帰への応用研究も紹介した．駆け足ではあったが，1 人でも多くの読者がスプライン関数に興味を抱いて頂ければ幸いである．

文　献

[1] 市田 浩三, 吉本 富士市, スプライン関数とその応用. 東京: 教育出版, 1979.
[2] 桜井 明, スプライン関数入門. 東京: 東京電機大学出版局, 1981.
[3] C. K. Chui, 桜井 明 (翻訳), 新井 勉 (翻訳), マルチスプライン. 東京: 東京電機大学出版局, 1991.
[4] D. Kitahara, K. Leng, Y. Tezuka, and A. Hirabayashi, "Simultaneous spline quantile regression under shape constraints," in *Proc. EUSIPCO 2020*, Aug. 2020, to appear.
[5] R. Koenker, *Quantile Regression*. New York, NY: Cambridge University Press, 2005.