

# サンプル値制御理論による確率密度関数の推定

永原正章 佐藤健治 山本裕 (京都大学)

## Density Estimation Based on the Sampled-Data Control Theory

M. Nagayara, \*K. Sato and Y. Yamamoto (Kyoto University)

**Abstract**— This paper is concerned with a nonparametric density estimation from a given histogram. The sampled-data  $H^\infty$  control theory is applied to this problem. Since the theory provides a powerful tool to reduce the worst-case estimation error taking inter-sample behavior into account, the estimated probability density function is reliable even if the number of samples are relatively small. Numerical examples are illustrated to show the effectiveness of the proposed method.

**Key Words:** probability density estimation, sampled-data control, signal processing

### 1 はじめに

不確かさのモデリングは、制御や信号処理において非常に重要な役割を果たす。不確かさには、限られた情報や知識に起因するもの（例えば非線形システムの線形近似や連続時間システムの離散化など）と、観測ノイズに起因するものの二つに大きく分けられる。制御工学においては、これらの不確かさは、有界なノルムを持つシステムまたは信号の集合としてモデリングされることが多い。特に  $H^\infty$  制御理論は、ノルムとして  $H^\infty$  ノルムを用い、この集合の中の最悪な要素に対して性能を最適化する手法（ミニマックス手法）を提供し、現在の制御工学の主流となっている。一方、信号処理では、不確かさを確率を用いて表現することが多い。確率を用いた不確かさのモデリングは、Bernoulli や Laplace の時代から現在まで非常に長い歴史を持つ。特に近年の計算機の高速化と記憶容量の巨大化により、これまで不可能であった複雑な計算も容易に実行できるようになり、適応信号処理や情報圧縮の技術が格段に進歩した。

これらの制御や信号処理を行うためには、その不確かさを見積る必要がある。 $H^\infty$  制御では、不確かさを表現するシステムの  $H^\infty$  ノルムの上界を知る必要があるが、これはシステム同定理論においてモデル集合同定と呼ばれ、盛んに研究されている（例えば、文献<sup>9)</sup>など）。一方、信号処理においては、不確かさを表現する確率分布（確率密度関数）を同定する必要がある。これを確率密度推定と呼び、特に学習理論の分野で盛んに研究されている問題である。いずれの問題においても、その同定精度が制御系や信号処理系の達成可能な性能に大きく影響することはいうまでもない。

しかし、確率密度推定においても不確かさの存在は不可避である。推定のための観測データは、一般に有限個であり、また観測ノイズも存在する。Bayes 推定では、これらに起因する不確かさもまた確率分布で表現し、この分布のパラメータ（超パラメータと呼ばれる）も同定する手法を提供する<sup>1)</sup>。さらにこれらを階層的に拡張した階層 Bayes モデルも提案されている。一方、本論文で提案する手法は、確率密度推定における不確かさを確定的な集合で表現する。すなわち、確率密度推定を  $H^\infty$  制御理論、特にサンプル値制御理論の枠組みで行うことを提案する。

サンプル値制御理論は文献<sup>12)</sup>によって近代的な方

法論が確立し、近年では、信号処理への適用<sup>3)</sup>を含めさまざまな応用が研究されている。得られる補償器は信号のアナログ特性を考慮した上で最適なものであるため、離散的なデータから連続特性を復元することができる。このような特徴は、密度推定問題に重要な示唆を与える。すなわち、有限回の試行によって連続値をもつ確率変数の密度関数を推定しなければならない状況下では、連続特性を最適に再現するサンプル値制御理論がひとつの解決策となりえる。

本研究の背景には、確率変数の値の離散化は信号の時間離散化に対応しているという Widrow ら<sup>10, 11)</sup>により提案されたアイデアがある。これは、エリアサンプルと呼ばれる操作により確率密度関数の定義域を離散化した関数は、ヒストグラムにおいて試行回数を無限大とした極限と一致するという事実によって正当化される。したがって、試行のヒストグラムをデータとして、実数値をとる確率変数の密度関数を復元する問題は、復元された関数が確率密度関数であるという制約のためにいくつか技術的な仮定が要求されるものの、従来のサンプル値制御理論による最適フィルタ設計問題<sup>3, 2, 5)</sup>と類似の性質を持つ。

提案手法は Widrow ら<sup>10, 11)</sup>のアイデアを継承するものであるが、密度推定に用いられる手法が大きく異なり、推定対象となる確率密度関数のクラスは拡張される。文献<sup>10, 11)</sup>で提案された手法は Shannon のサンプリング定理<sup>7)</sup>に基づいたものであり、確率密度関数のフーリエ変換（特性関数）が有界なサポートを持つという仮定を排除できない。一方、一様分布のような有界なサポートを持つ密度関数の Fourier 変換は上記の仮定を満たさず、また実用上多くの分布（Erlang 分布など）も同様に適用できない。一方、提案手法は特性関数の減衰特性に先見的な情報を仮定するものの、特性関数のサポートの有界性のような強い仮定を必要としないため、応用上用いられる確率密度関数の多くが適用対象となる。

さらに、上記で述べた Bayes 推定では、確率分布を有限個のパラメータで表現し、そのパラメータを推定する手法がよく用いられる。例えば、確率分布を正規分布と仮定することにより、平均と分散の2つのパラメータに関する Bayes 更新則が得られる。一般には、共役事前分布を導入することにより、確率分布の関数形が保存される更新則が導出でき、次元の大きい確率

変数に対しても推定問題を効率的に解くことができる<sup>1)</sup>。これをパラメトリック推定と呼ぶ。しかし、分布の形状が単純なパラメータで表現しきれない状況も多く、このような場合はノンパラメトリックな手法を用いざるを得ない。そこで本研究では、分布の形状に仮定を置かないノンパラメトリック推定法を用いる。提案する手法はヒストグラム密度推定法の一つとみなせる。

以下、第2章で問題設定を行う。第3章でWidrowの手法について述べた後、第4章で提案手法を紹介する。第5章では、数値例によって提案手法の有効性を示す。

## 2 確率密度推定

本章では問題設定を行う。第1章でも述べた通り、量子化（縦軸の離散化）を密度関数のサンプリング（横軸の離散化）と対応付ける点が本質的である。

確率変数の試行によって得られたデータは、以下の一様量子化器によって有限個の値に制限されるとする。これは、計算機を用いた解析には不可避の想定であるし、少ないサンプルで意味のある密度推定を行うためにも必要な仮定である。

$$Q: \mathbb{R} \rightarrow \mathcal{A} := \{-M\Delta, (-M+1)\Delta, \dots, M\Delta\},$$

$$Qx := \begin{cases} k\Delta, & (k - \frac{1}{2})\Delta \leq x < (k + \frac{1}{2})\Delta, \\ & k = -M+1, \dots, M-1, \\ M\Delta, & x \geq (M - \frac{1}{2})\Delta, \\ -M\Delta, & x < (-M - \frac{1}{2})\Delta. \end{cases} \quad (1)$$

我々が扱う問題の基本形は以下の通りである。

問題 1. 未知の確率密度関数  $p(x)$  に従う確率変数の独立試行によって得られたサンプルを式 (1) で一様量子化したデータ  $\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_N$  が与えられているとする。このとき、 $p(x)$  を推定せよ。

試行により得られたデータの一様量子化器による量子化は、次式で定義されるエリアサンプリング<sup>10, 11)</sup>を用いた確率密度関数の横軸の離散化に対応している。

$$c[k] = \int_{(k-1/2)\Delta}^{(k+1/2)\Delta} p(x)dx =: (\tilde{S}_\Delta p)[k] \quad (2)$$

実は、大数の強法則により以下の補題が示される。

命題 2.  $A \subset \mathbb{R}$  を任意の Lebesgue 可測集合、 $\{X_n\}_{n=1}^\infty$  を確率密度関数  $p(x)$  に従う独立な確率変数の列であるとする。このとき、

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \chi_A(X_n) = \int_A p(x)dx \quad (3)$$

が、ほとんど確実に成り立つ。

この命題によって、量子化と密度関数離散化は以下のような関係にあることが分かる。まず、量子化幅  $\Delta$  で一様量子化してヒストグラムを生成する。このとき、試行の数を大きくするにしたがって、ある領域における頻度は式 (2) によって密度関数を離散化した値に近づいていく。

すなわち、密度関数のエリアサンプルは正確なヒストグラムと解釈することができて、問題 1 は（サンプル数無限大の極限において）以下の問題を考えることと同じである。

問題 3. 未知の確率密度関数  $p(x)$  をエリアサンプル (2) したデータが与えられている。このとき、 $p(x)$  を推定せよ。

もちろんこのような問題は、解の候補が無数に存在するため  $p(x)$  が属する関数空間に何らかの仮定を置く必要がある。Widrow らは Shannon の定理を適用するため完全帯域制限の仮定を用いた。我々は、安定で厳密に因果的な有限次元線形システムの出力が作る関数空間を採用する。

## 3 Widrow の量子化定理

この章では、Widrow らによる手法を簡単に説明する。式 (2) による密度関数の（横軸の）離散化は、

$$\begin{aligned} c[k] &= \int_{-\infty}^{\infty} \chi_{[-\Delta/2, \Delta/2)}(k\Delta - x)p(x)dx \\ &= (q_\Delta * p)(k\Delta) =: S_\Delta(q_\Delta * p)[k], \end{aligned}$$

$$q_\Delta(x) := \chi_{[-\Delta/2, \Delta/2)}(x)$$

と書き換えられる。ただし、 $\chi_I$  は  $I \subset \mathbb{R}$  上の定義関数であり、 $x \in I$  のとき  $\chi_I(x) = 1$ 、 $x \notin I$  のとき  $\chi_I(x) = 0$  と定義する。これにより、エリアサンプリングが  $q_\Delta$  と理想サンプリング  $S_\Delta$  によって表されることが分かる。Shannon の定理によれば、確率密度関数  $p$  と  $q_\Delta$  の合成積  $q_\Delta * p$  が完全帯域制限の仮定を満たすとき、問題 3 の完全な解が得られる。完全帯域制限された信号の空間を

$$BL^2 := \{f \in L^2 : (\mathcal{F}f)(\omega) = 0, |\omega| > \pi/\Delta\}$$

と定義する。ただし、 $\mathcal{F}$  は Fourier 変換を表す。Widrow らによる結果は定理として以下のようにまとめられる。

定理 4 (Widrow et al.<sup>10, 11)</sup>). 確率密度関数  $p(x)$  が  $BL^2$  に属するとする。このとき  $p(x)$  は、エリアサンプルにより得られたヒストグラム  $S_\Delta p$  により完全に再構成される。

## 4 $H^\infty$ 最適密度推定

### 4.1 関数空間 $FL^2$

$BL^2$  に属さない確率密度関数も応用上重要である。例えば、一様分布や Erlang 分布、Rayleigh 分布などは  $BL^2$  に属さないことが知られている<sup>6)</sup>。本研究では、以下の関数空間  $FL^2$  を採用する：

$$FL^2 := \{f \in L^2 : f = Fw, w \in L^2\}.$$

ただし、 $F$  は安定で厳密に因果的な有限次元線形システムである。この関数空間は完全帯域制限された関数を含まないが、上で述べたような確率密度関数を含んでいる<sup>6)</sup>。なお、システム  $F$  は密度推定の文献において平滑化パラメータと呼ばれているものに対応している<sup>1)</sup>。

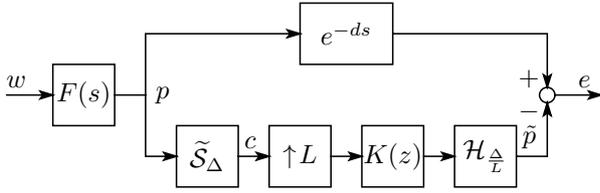


Fig. 1: Error system  $\mathcal{E}$

#### 4.2 サンプル値 $H^\infty$ 最適化による密度推定

考察する確率密度関数  $p(x)$  は  $FL^2$  に属すると仮定した上で, Fig. 1 の誤差系に関して信号復元問題を考察する.

確率密度関数  $p \in FL^2$  は  $L^2$  信号  $w$  を  $F(s)$  ( $F$  の伝達関数) でフィルタリングしたものであるとする. すなわち,  $F(s)$  は  $p$  の周波数モデルと考えられる. さらに,  $p$  をエリアサンプル  $\tilde{S}_\Delta$  により (正確な) ヒストグラム  $\{c[k]\}$  が得られる. 密度関数  $p$  の推定問題に対して, 本論文ではマルチレート信号処理<sup>8)</sup> で用いられる補間手法を採用する. まず, 離散データ  $\{c[k]\}$  を  $\uparrow L$  によってアップサンプルする. ただし,

$$((\uparrow L)c)[k] = \begin{cases} c[n], & k = Ln, n = 0, 1, \dots \\ 0, & \text{otherwise} \end{cases}$$

である. アップサンプルされた離散データ  $(\uparrow L)c$  を線形フィルタ  $K(z)$  により補間し, フィルタの出力は周期  $\Delta/L$  のゼロ次ホールド  $\mathcal{H}_{\Delta/L}$  で連続値に変換したものを密度関数の推定値  $\tilde{p}$  とする. 誤差評価は  $\tilde{p}(x)$  と  $p(x-d)$  の差に基づいて行われる. すなわち, Fig. 1 の誤差系を  $H^\infty$  ノルム

$$\mathcal{E} := \{e^{-ds} - \mathcal{H}_{\Delta/L}K(z)(\uparrow L)\tilde{S}_\Delta\}F(s). \quad (4)$$

とおき, その  $H^\infty$  ノルムを最小化する線形フィルタ  $K$  を求めることで, 最適な推定が得られる.

#### 4.3 確率密度関数の制約条件

以下では,  $K$  は FIR (finite impulse response) フィルタであると仮定する.

関数  $p(x)$  は確率密度であるから,

$$p(x) \geq 0, \quad (5)$$

$$\int_{-\infty}^{\infty} p(x)dx = 1. \quad (6)$$

を満たす. したがって, 密度関数の推定  $\tilde{p}(x)$  も同様の制約 (5), (6) を満たさなければならない. FIR フィルタ  $K(z)$  を

$$K(z) = \sum_{n=0}^N a_n z^{-n} \quad (7)$$

とする. 実は, 制約条件 (5), (6) は LMI (linear matrix inequality) と LME (linear matrix equality) で表現可能であることが示される<sup>4)</sup>.

補題 5. 推定された密度関数  $\tilde{p}(x)$  が任意の  $x \in \mathbb{R}$  に対して  $\tilde{p}(x) \geq 0$  を満たすとき, またそのときに限り,

$$a_n \geq 0, \quad n = 0, 1, \dots, N.$$

補題 6. エリアサンプルされた密度関数は正規化されているとする. すなわち,

$$\Delta \sum_{n=-\infty}^{\infty} c[n] = 1$$

が成り立つとする. このとき, 以下の条件 (i) と (ii) は同値である.

$$(i) \quad \int_{-\infty}^{\infty} \tilde{p}(x)dx = 1,$$

$$(ii) \quad \sum_{n=0}^N a_n = L.$$

したがって, 最適推定  $\tilde{p}$  が確率密度になる条件を FIR フィルタ  $K(z)$  の係数  $a_0, \dots, a_N$  に関する線形制約として記述することができる. このような制約は Widrow らの手法では考慮されていない.

#### 4.4 最適 FIR フィルタの設計

フィルタ  $K(z)$  を式 (7) で与えられる FIR フィルタと仮定すると, サンプル値系 (4) に対して, その  $H^\infty$  ノルムを近似する離散時間誤差系

$$E_d(z) = T_1(z) + W(\mathbf{a})T_2(z)$$

が得られる<sup>2)</sup>. ただし,

$$\mathbf{a} := [a_0, a_1, \dots, a_{N-1}],$$

$$W(\mathbf{a}) := \begin{bmatrix} a_0 I & a_1 I & \dots & a_{N-1} I \end{bmatrix}.$$

$E_d(z)$  は  $\mathbf{a}$ -アフィンであるから,  $E_d(z)$  の状態空間表現は次式で与えられる.

$$E_d(z) = \left[ \begin{array}{c|c} A_e & B_e \\ \hline C_e(\mathbf{a}) & D_e(\mathbf{a}) \end{array} \right].$$

ここで,  $A_e, B_e$  は定数行列であり,  $C_e(\mathbf{a}), D_e(\mathbf{a})$  は設計パラメータ  $\mathbf{a}$  にアフィンに依存する. KYP 補題により, 最適フィルタ設計問題は, 制約条件

$$\begin{aligned} P &= P^\top > 0, \\ \begin{bmatrix} A_e^\top P A_e - P & A_e^\top P B_e & C_e(\mathbf{a})^\top \\ B_e^\top P A_e & -\gamma I + B_e^\top P B_e & D_e(\mathbf{a})^\top \\ C_e(\mathbf{a}) & D_e(\mathbf{a}) & -\gamma I \end{bmatrix} &< 0, \\ V(\mathbf{a}) &\geq 0, \\ \text{tr} V(\mathbf{a}) &= L, \end{aligned}$$

の下で,  $\gamma > 0$  を最小化する FIR 係数  $\mathbf{a}$  を見つける問題となる. ただし,  $V(\mathbf{a}) := \text{diag}\{a_1, a_2, \dots, a_{N-1}\}$  とし,  $\text{tr} V(\mathbf{a})$  は行列  $V(\mathbf{a})$  のトレースを表す.

補題 5 により  $V(\mathbf{a}) \geq 0$  は制約条件 (5) に対応し,  $\text{tr} V(\mathbf{a}) = L$  は補題 6 から制約条件 (6) と等価な条件である. これらはすべて設計パラメータ  $\mathbf{a}$  に関して線形な制約であるから, MATLAB などの数値計算ソフトウェアを使って効率的に解くことができる.

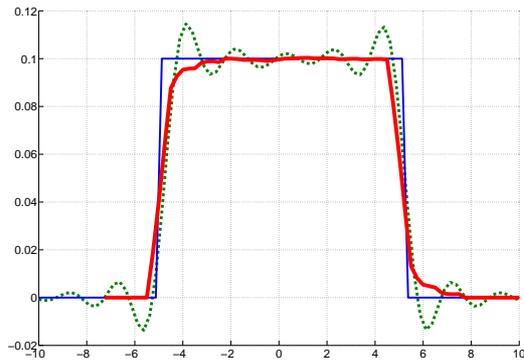


Fig. 2: 確率密度関数（一様分布）の推定：もとの分布（細実線），Widrow による推定（破線），提案法による推定（太実線）

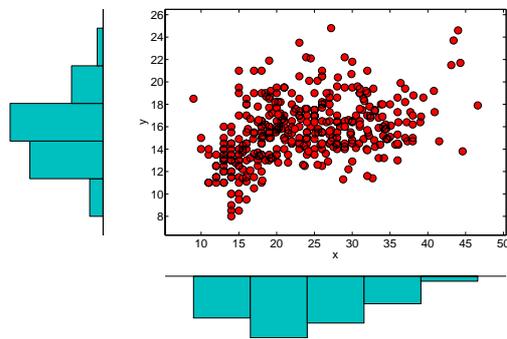


Fig. 3: 2次元データのヒストグラム

## 5 数値例

ここでは，1次元および2次元の確率密度推定の例題を示す．まず，区間  $[-5, 5]$  上の一様分布の推定を行う．この分布から  $10^5$  個のサンプルを取り， $\Delta = 1$  の幅のヒストグラムを求める．推定のためのフィルタ  $K(z)$  の設計パラメータは以下のとおりである．アップサンプリングの定数を  $L = 4$  とし，フィルタのタップ数を  $N = 16$  とする．平滑化パラメータである伝達関数  $F(s)$  を

$$F(s) = \frac{\omega_1 \omega_2}{(s + \omega_1)(s + \omega_2)}, \quad \omega_1 = \pi/10, \quad \omega_2 := \pi/20$$

とする．推定結果を Fig. 2 に示す．Widrow らの推定は，分布の境界付近で大きく波打っていることがわかる．また確率密度が負の値をとる部分も存在する．これは，一様分布が Widrow の定理の仮定を満たさないことが原因である．一方，我々の手法による推定はこのようなことが無く，良好な推定が得られる．

次に2次元データにもとづく確率密度の推定を行う．フィルタは上記と同じものを用い，画像処理と同様の手法<sup>13)</sup>により信号処理を2次元に拡張して推定を行う．2次元データは MATLAB の Statistics Toolbox に付属の `carbigr` を用いた．Fig. 3 にデータの分布図およびヒストグラムを示す．得られた推定を Fig. 4 に示す．非常に粗いヒストグラムから推定できていることがわかる．

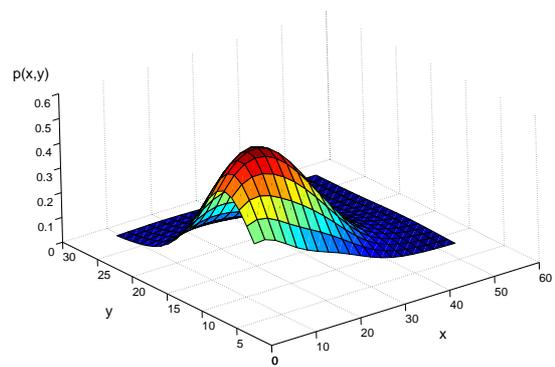


Fig. 4: 確率密度関数の推定

## 6 おわりに

本論文では，確率密度関数の推定問題をサンプル値  $H^\infty$  制御問題として定式化し，不確かさを考慮した推定法を提案した．Shannon の定理にもとづく Widrow らの推定法に比べ，提案手法はより広い分布に対して適用できる．また，画像信号処理（2次元信号処理）の手法を使えば，2次元の分布の推定も可能である．今後の課題は，より高次元の分布の推定法を確立することである．

## 参考文献

- 1) C. M. Bishop: *Pattern Recognition and Machine Learning*, Springer, 2006.
- 2) 石井, 山本: マルチレート D/A 変換器のサンプル値  $H^\infty$ ,  $H^2/H^\infty$  設計, システム制御情報学会論文誌, Vol. 11, No. 10, pp. 585–592, 1998.
- 3) P. P. Khargonekar and Y. Yamamoto: Delayed signal reconstruction using sampled-data control, *Proc. of 35th IEEE Conf. on Decision and Control*, pp. 1259–1263, 1996.
- 4) M. Nagahara, K. Sato and Y. Yamamoto:  $H^\infty$  optimal nonparametric density estimation from quantized samples, *Proc. of ISICIE SSS*, 2008 (submitted).
- 5) M. Nagahara and Y. Yamamoto: A new design for sample-rate converters, *Proc. of 39th IEEE Conf. on Decision and Control*, pp. 4296–4301, 2000.
- 6) A. Papoulis and S. U. Pillai: *Probability, Random Variables and Stochastic Processes*, Fourth Ed., Mc Graw Hill, 2002.
- 7) M. Unser: Sampling — 50 years after Shannon, *Proc. IEEE*, Vol. 88, No. 4, pp. 569–587, 2000.
- 8) P. P. Vaidyanathan: *Multirate Systems and Filter Banks*, Prentice Hall, Englewood Cliffs, 1993.
- 9) D. K. De Vries and P. M. J. Van den Hof: Quantification of uncertainty in transfer function estimation: a mixed probabilistic – worst-case approach, *Automatica*, Vol. 31, No. 4, pp. 543–557, 1995.
- 10) B. Widrow: A study of rough amplitude quantization by means of nyquist sampling theory, *IRE Trans. on Circuit Theory*, Vol. 3, No. 4, pp. 266 – 276, 1956.
- 11) B. Widrow, I. Kollar, and M.-C. Liu: Statistical theory of quantization, *IEEE Trans. on Instrumentation and Measurement*, Vol. 45, No. 2, pp. 353 – 361, 1996.
- 12) Y. Yamamoto: A function space approach to sampled-data control systems and tracking problems, *IEEE Trans. Autom. Control*, Vol. 39, No. 4, pp. 703 – 712, 1994.
- 13) 山本, 永原: デジタル制御理論による信号処理, 映像情報メディア学会誌, Vol. 61, No. 12, pp. 1710–1715, 2007.