



システム制御理論と統計的機械学習

第 12 章：ガウス過程回帰

加嶋 健司

October 10, 2025

京都大学 情報学研究科

本章の流れ

12.1 カーネル法

カーネル法とよばれる一連のノンパラメトリック手法の基本的な考え方

11.2 システム制御への応用

- ・ モデルフリーなパラメータチューニングなどに利用可能なベイズ最適化
- ・ 状態遷移関数をガウス過程回帰により同定するガウス過程回帰状態空間モデル

カーネル法

カーネルトリック (再掲)

集合 X , 特徴写像 $\phi: X \rightarrow \mathbb{R}^{n_f}$, $\{x_s\}_{s=1}^{\bar{s}}$, $x_s \in X$ および $p \sim \mathcal{N}(0, I)$

実数値関数 $f_p := p^\top \phi$ の代入値 $y_s = f_p(x_s) + w_s$, $w_s \sim \mathcal{N}(0, \sigma^2)$ を観測

$k(x, x') := \phi(x)^\top \phi(x')$, $(k(x_{1:\bar{s}}, \cdot))_i := k(x_i, \cdot) : X \rightarrow \mathbb{R}^{\bar{s}}$, $(K_\phi(x_{1:\bar{s}}))_{ij} := k(x_i, x_j) \in \mathbb{R}^{\bar{s} \times \bar{s}}$

データ $\{(x_s, y_s)\}_{s=1}^{\bar{s}}$ に対して f_p の MAP 推定関数は

$$\sum_s \|y_s - \rho^\top k(x_{1:\bar{s}}, x_s)\|^2 + \sigma^2 \rho^\top K_\phi(x_{1:\bar{s}}) \rho \quad (12.5)$$

を最小化する

$$\rho_* := (K_\phi(x_{1:\bar{s}}) + \sigma^2 I)^{-1} y_{1:\bar{s}} \quad (12.6)$$

を用いて,

$$\rho_*^\top k(x_{1:\bar{s}}, \cdot) \in \mathcal{F}_{k(x_{1:\bar{s}}, \cdot)} := \{\rho^\top k(x_{1:\bar{s}}, \cdot) : \rho \in \mathbb{R}^{\bar{s}}\}$$

で与えられる

半正定値カーネル (2) 特徴関数を用いない場合

$k(\cdot, \cdot)$ を $k(\mathbf{x}, \mathbf{x}') := \phi(\mathbf{x})^\top \phi(\mathbf{x}')$ の形に限定しない一般の二変数関数 k とする

- $k(\cdot, \mathbf{x}')$ が \mathbf{x}' にピークをもつ釣鐘形となる

$$k_{\text{SE}}(\mathbf{x}, \mathbf{x}') := \exp(-\|\mathbf{x} - \mathbf{x}'\|^2/c^2), \quad c > 0 \quad (12.7)$$

のような関数を選べば、 $\phi_s := k(\mathbf{x}_s, \cdot)$, $s = 1, \dots, \bar{s}$ の線形結合からなる MAP 推定関数は「 ρ の選択により、与えられたデータ付近の関数形を局所的に指定する」

- $K_\phi(\mathbf{x}_{1:\bar{s}})$ に対応する行列が半正定となることだけを仮定する。

半正定値カーネル (3) 定義

定義 12.1.1 – 半正定値カーネル

二変数関数 $k : X \times X \rightarrow \mathbb{R}$ が対称性 $k(x, x') = k(x', x)$ を持ち、任意の $N \in \mathbb{N}$ および $x_s, s = 1, \dots, N$ に対して行列

$$K(x_{1:N}) := \begin{bmatrix} k(x_1, x_1) & k(x_1, x_2) & \cdots & k(x_1, x_N) \\ k(x_2, x_1) & k(x_2, x_2) & \cdots & k(x_2, x_N) \\ \vdots & \vdots & \ddots & \vdots \\ k(x_N, x_1) & k(x_N, x_2) & \cdots & k(x_N, x_N) \end{bmatrix} \in \mathbb{R}^{N \times N} \quad (12.8)$$

が半正定値行列となるとき、 $k(\cdot, \cdot)$ を**半正定値カーネル** (positive semi-definitite kernel)、もしくは簡単に**カーネル**とよぶ。

半正定値カーネル (4) 例

例 12.1.2 – カーネルの例

- ・ ガウシアンカーネル

$$k_{\text{SE}}(\mathbf{x}, \mathbf{x}') := \exp(-\|\mathbf{x} - \mathbf{x}'\|^2 / c^2), \quad c > 0 \quad (12.8)$$

があり，SE カーネル（square exponential –），RBF カーネル（radially bounded function –）ともよばれる．

- ・ ラプラスカーネル

$$k(\mathbf{x}, \mathbf{x}') := \exp(-\alpha \sum_{i=1}^n |\mathbf{x}_i - \mathbf{x}'_i|) \quad (12.9)$$

- ・ ユークリッドカーネル $k(\mathbf{x}, \mathbf{x}') := \mathbf{x}^\top \mathbf{x}'$
- ・ 多項式カーネル $k(\mathbf{x}, \mathbf{x}') = (c + \mathbf{x}^\top \mathbf{x}')^d, \quad c > 0, \quad d \in \mathbb{N}$

半正定値カーネル (5) 特徴写像に対応するカーネル

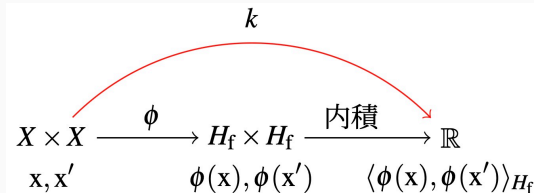
定理 12.1.3 – 特徴写像に対応するカーネル

X からある内積空間 H_f への関数 $\phi : X \rightarrow H_f$ に対して,

$$k(x, x') := \langle \phi(x), \phi(x') \rangle_{H_f} \quad (12.10)$$

により定義される k は半正定値カーネルである。

- 特徴写像の内積はカーネルとなる。
- カーネルは特徴写像の内積で書ける。
(注意 12.1.4 マーサー (Mercer) の定理)



ガウス過程回帰 (1) ガウス過程

定義 12.1.5 – ガウス過程

実数値関数 $\mu : X \rightarrow \mathbb{R}$, 半正定値カーネル $k : X \times X \rightarrow \mathbb{R}$, $f : X \rightarrow \text{rv}(\mathbb{R})$ をとる. このとき, 任意の $N \in \mathbb{N}$ および $\mathbf{x}_s \in X$, $s = 1, 2, \dots, N$ に対して

$$f(\mathbf{x}_{1:N}) \sim \mathcal{N}(\mu(\mathbf{x}_{1:N}), K(\mathbf{x}_{1:N})) \quad (12.12)$$

が成り立つとき, f は平均関数 μ , 共分散カーネル k をもつ**ガウス過程** (Gaussian process) であるといい, $f \sim \text{GP}(\mu, k)$ と表記する. ただし, $f(\mathbf{x}_{1:N})$ は

$$f(\mathbf{x}_{1:N}) = \begin{bmatrix} f(\mathbf{x}_1) & \dots & f(\mathbf{x}_N) \end{bmatrix}^\top \in \text{rv}(\mathbb{R}^N) \quad (12.13)$$

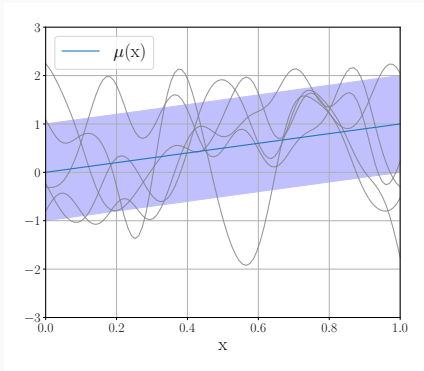
で与えられ, $\mu(\mathbf{x}_{1:N})$ も同様に定義する.

- $f(\mathbf{x}_{1:N})$ の統計的性質を規定することで, 間接的に f の性質を与える

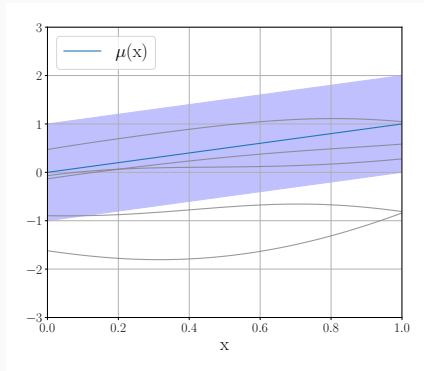
ガウス過程回帰 (2) ガウス過程の例

例 12.1.7 – ガウス過程

$\mu(x) = x$, $k_{SE}(x, x') := \exp(-\|x - x'\|^2/c^2)$ に対して $f \sim \text{GP}(\mu, k_{SE})$



$c = 0.1$



$c = 1$

ガウス過程回帰 (3) 事後分布

定理 12.1.8 – ガウス過程回帰

$f \sim \text{GP}(\mu, k)$, ある $\bar{s} \in \mathbb{N}$ および $\mathcal{D} := \{(\mathbf{x}_s, y_s)\}_{s=1:\bar{s}}$, $\mathbf{x}_s \in X$, $y_s \in \mathbb{R}$ に対して

$$y_s := f(\mathbf{x}_s) + w_s, w_s \sim \mathcal{N}(0, \sigma^2), s = 1, 2, \dots, \bar{s} \quad (12.15)$$

とし, f, w_1, \dots は独立であるとする. このとき, $y_{1:\bar{s}} = \mathbf{y}_{1:\bar{s}} \in \mathbb{R}^{\bar{s}}$ のもとでの $f \in \text{rv}(\mathcal{F})$ の事後分布は, 平均関数および共分散カーネル関数が

$$\mu(\bar{\mathbf{x}}|\mathcal{D}) := \mu(\bar{\mathbf{x}}) + k(\mathbf{x}_{1:\bar{s}}, \bar{\mathbf{x}})^\top (K(\mathbf{x}_{1:\bar{s}}) + \sigma^2 I)^{-1} (\mathbf{y}_{1:\bar{s}} - \mu(\mathbf{x}_{1:\bar{s}})) \quad (12.16)$$

$$k(\bar{\mathbf{x}}, \bar{\mathbf{x}}'|\mathcal{D}) := k(\bar{\mathbf{x}}, \bar{\mathbf{x}}') - k(\mathbf{x}_{1:\bar{s}}, \bar{\mathbf{x}})^\top (K(\mathbf{x}_{1:\bar{s}}) + \sigma^2 I)^{-1} k(\mathbf{x}_{1:\bar{s}}, \bar{\mathbf{x}}') \quad (12.17)$$

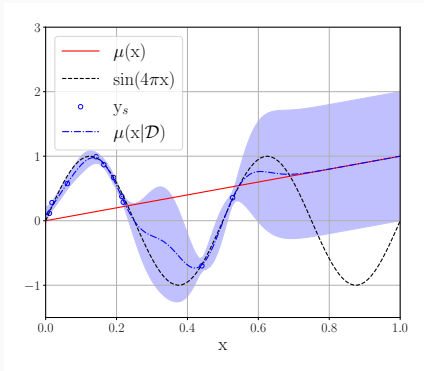
与えられるガウス過程 $f|_{\mathcal{D}} \sim \text{GP}(\mu, k|\mathcal{D})$ である.

- 事後共分散カーネル $k(\cdot, \cdot|\mathcal{D})$ は \mathbf{y} (観測値) に依存しない

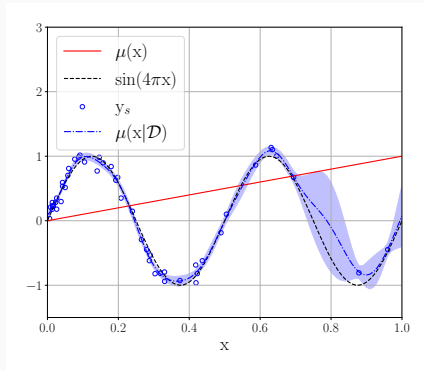
ガウス過程回帰 (4) ガウス過程回帰の例

例 12.1.9 – ガウス過程回帰

$\mu(x) = x$, $k_{SE}(x, x') := \exp(-\|x - x'\|^2/c^2)$ に対して $f \sim \mathbf{GP}(\mu, k_{SE})$



$\bar{s} = 10$



$\bar{s} = 50$

ガウス過程回帰 (5) 考察

事前平均 μ からの差 $f - \mu$ を新たに f とおくことで一般性を失わず $\mu = 0$ とする.

- 事後平均関数

$$\mu(\cdot|\mathcal{D}) = k(\mathbf{x}_{1:\bar{s}}, \cdot)(K(\mathbf{x}_{1:\bar{s}}) + \sigma^2 I)^{-1} \mathbf{y}_{1:\bar{s}} \quad (12.23)$$

はパラメトリック推定と完全に同じ

パラメトリック回帰と比較した場合のガウス過程回帰の利点は

- 事後分散関数により近似誤差が理論的に評価できる点
- $\mathcal{F}_{k(\mathbf{x}_{1:\bar{s}}, \cdot)} := \{\rho^\top k(\mathbf{x}_{1:\bar{s}}, \cdot) : \rho \in \mathbb{R}^{\bar{s}}\}$ の単純な関数形に限定する正当性を与えている点
- カーネルさえ選択すれば系統的かつ柔軟な手続きが与えられる点

計算量の観点からは、カーネル行列の逆行列の計算がデータ数 \bar{s} とともに増大する点が課題である。対策として、代表点の選定方法が提案されている。

再生核ヒルベルト空間 (1) 定義

確率によらない最小二乗法に対応するノンパラメトリックなアプローチ

定義 12.1.10 – 再生核ヒルベルト空間

\mathcal{H} を X 上の実数値関数を要素とし、内積 $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ を有するヒルベルト空間（ノルム $\|f\|_{\mathcal{H}} := \sqrt{\langle f, f \rangle_{\mathcal{H}}}$ に関して完備）とする．このとき、任意の $x \in X$ に対して、ある $C_x < \infty$ が存在し

$$|f(x)| \leq C_x \|f\|_{\mathcal{H}}, \quad \forall f \in \mathcal{H} \quad (12.24)$$

が成り立つとき、 \mathcal{H} は X -上の**再生核ヒルベルト空間** (reproducing kernel Hilbert space; RKHS) であるという．

- \mathcal{H} は \mathcal{F}_{ϕ} に対応する概念
- 「代入」という操作 $f \mapsto f(x)$ が有界という意味で、なめらかな関数からなる関数族

再生核ヒルベルト空間 (2) 反例としての L^2

関数を要素としてもつヒルベルト空間の代表例である L^2 -空間，すなわち内積 $\langle f, g \rangle_{L^2} := \int_{\mathbb{R}} f(x)g(x)dx$ を備えた二乗可積分な関数の集合はこの性質をもたない

例 12.1.11 – L^2 における代入操作の非有界性

任意の $c > 0$ に対して関数 $f_c(x) := \exp(-\|x\|^2/c^2)$ は L^2 の要素

- $x = 0$ を代入するという操作は $f_c(0) = 1, \forall c$ であるのに対して， $c \rightarrow 0$ のとき $\|f_c\|_{L^2} \rightarrow 0$ である．
- $g_\sigma(x) := \mathcal{N}(x|0, \sigma^2) \in L^2$ を考えると， $g_\sigma(0) = (\sqrt{2\pi}\sigma)^{-1}$ に対して， $\|g_\sigma\|_{L^2} = \sigma^{-1/2}\|g_1\|_{L^2}$ であるため， $x = 0$ の代入は $\sigma \rightarrow 0$ で有界とならない．
- 「いくらでも急峻に変動する関数」を含むような集合において「急峻な変動を加味しないノルム」を導入すると，代入操作は有界とならない．

再生核ヒルベルト空間 (2) カーネル関数

再生核ヒルベルト空間はとても自由度の高い集合に見えるが、実際には対応する唯一のカーネル関数により特徴づけられる。

定理 12.1.12 – 再生核ヒルベルト空間とカーネルの対応

任意の X -上の再生核ヒルベルト空間 \mathcal{H} に対して、二つの条件

- 任意の $x \in X$ に対して、 $k(\cdot, x) \in \mathcal{H}$
- (再生性) 任意の $f \in \mathcal{H}$, $x \in X$ に対して、

$$\langle f(\cdot), k(\cdot, x) \rangle_{\mathcal{H}} = f(x) \quad (12.25)$$

を満たす唯一の X -上の半正定値カーネル k が存在し、これを \mathcal{H} の再生核とよぶ。

逆に、任意の半正定値カーネル k に対して、再生核 k をもつ再生核ヒルベルト空間 \mathcal{H}_k が唯一つ存在する。

再生核ヒルベルト空間 (3) 例

例 12.1.13 – ReLU 関数と RKHS

\mathbb{R}_+ 上で定義された絶対連続かつ $f(0) = 0$ を満たす関数からなる集合に内積

$$\langle f, g \rangle_{\mathcal{H}} := \int_0^{\infty} \nabla f(x) \cdot \nabla g(x) dx \quad (12.28)$$

を導入したヒルベルト空間 \mathcal{H} を考える．したがって，

$$\|f\|_{\mathcal{H}}^2 := \int_0^{\infty} |\nabla f(x)|^2 dx \quad (12.29)$$

である．このとき，この \mathcal{H} は再生核

$$k(x, x') := \min(x, x') = -\text{ReLU}(-(x - x')) + x' \quad (12.30)$$

をもつ再生核ヒルベルト空間となる．

再生核ヒルベルト空間 (4) リッジ回帰の一般化

定理 12.1.14 – リプレゼンター定理

再生核 k をもつ X 上の再生核ヒルベルト空間 \mathcal{H}_k , $\mathcal{D} := \{(x_s, y_s)\}_{s=1}^{\bar{s}}$ および $L : (X \times \mathbb{R}^2)^{\bar{s}} \rightarrow \mathbb{R} \cup \{\infty\}$, 厳密に単調増加する関数 $r : [0, \infty) \rightarrow \mathbb{R}$ が与えられたとき,

$$L((x_1, y_1, f(x_1)), \dots, (x_{\bar{s}}, y_{\bar{s}}, f(x_{\bar{s}}))) + r(\|f\|_{\mathcal{H}_k}) \quad (12.31)$$

を最小化する $f \in \mathcal{H}_k$ は $\mathcal{F}_{k(\cdot, x_{1:\bar{s}})} := \{\rho^\top k(\cdot, x_{1:\bar{s}}) : \rho \in \mathbb{R}^{\bar{s}}\}$ の要素である.

系 12.1.15 – ガウス過程回帰の事後平均との対応

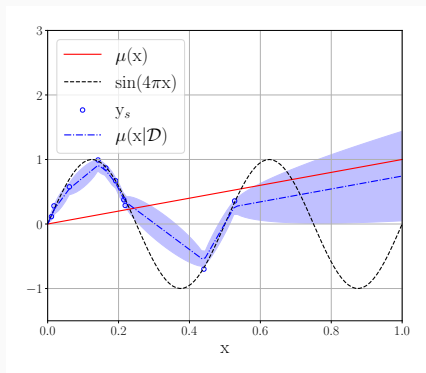
ガウス過程回帰による MAP 推定関数 $\mu(\cdot | \mathcal{D}) = k(\cdot, x_{1:\bar{s}})(K(x_{1:\bar{s}}) + \sigma^2 I)^{-1} y_{1:\bar{s}}$ は

$$\sum_{s=1}^{\bar{s}} |f(x_s) - y_s|^2 + \sigma^2 \|f\|_{\mathcal{H}_k}^2 \quad (12.32)$$

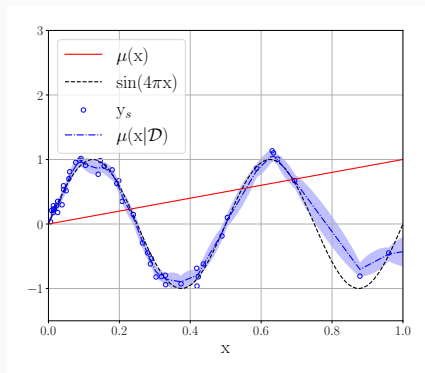
を最小化する \mathcal{H}_k の要素である.

再生核ヒルベルト空間 (5) リッジ回帰の一般化の例

$$\|f\|_{\mathcal{H}}^2 := \int_0^{\infty} |\nabla f(x)|^2 dx$$



$\bar{s} = 10$



$\bar{s} = 50$

システム制御への応用

ベイズ最適化 (1) ブラックボックス最適化

ブラックボックス最適化：

具体的な関数形は未知の関数 $f : X \rightarrow \mathbb{R}$ に対して、指定した $x_s \in X$ に対して

$$y_s := f(x_s) + w_s \quad (12.33)$$

のみが観測できる状況で、 $x^* := \arg \max_{x \in X} f(x)$ を求める最適化問題

回帰同様に f に関して何らかの滑らかさの仮定をおき、代入回数を削減したい。

Example

制御系設計において、たとえば x を PID 制御ゲインの組みとし、対応する制御性能 $f(x)$ を最大化する x^* を探したいとしよう。このとき、関数 f を求めることが現実的でなくても、与えられた \bar{x} のもとで実験を実施して $f(\bar{x})$ (の近似値) の値を測定することはできる状況であれば、ブラックボックス最適化とみなすことができ、実験回数を削減することは有用であると考えられる。

ベイズ最適化 (2) 獲得関数

ベイズ最適化 (Bayesian optimization) :

未知関数をガウス過程 $f \sim \text{GP}(\mu, k)$ とし、事後分布を活用してつぎの評価点を選択する
ブラックボックス最適化手法

与えられた $\mathcal{D} := \{(x_i, y_i)\}_{i=1:s}$ に対して、**獲得関数 (acquisition function)** を最大化

- s 番目までの最適値 $\bar{f}_s := \arg \max_{i \in [s]} y_i$ に対して最大値の改善する確率 (probability of improvement) $\text{PI}(x) := \mathbb{P}[f(x|\mathcal{D}) > \bar{f}_s]$
- 改善量の期待値 (expected improvement) $\text{EI}(x) := \text{ReLU}(f(x|\mathcal{D}) - \bar{f}_s)$
- 重み定数 $\beta > 0$ を用いて平均と分散の和 (upper confidence bound; UCB)
 $\text{UCB}(x) := \mu(x|\mathcal{D}) + \beta k(x, x|\mathcal{D})$

ガウス過程回帰状態空間モデル (1) 遷移関数の近似

前章までは、雑音 v と（確定的な）関数 $f: \mathbb{X} \times \mathbb{U} \times \mathbb{V} \rightarrow \mathbb{X}$ を用いて $x_{k+1} = f(x_k, u_k, v_k)$ により確率的な状態遷移を表現

ガウス過程回帰状態空間モデル：

（確率的に定まる関数である）ガウス過程 f を用いて、

$$x_{k+1} = f(x_k, u_k) \quad (12.37)$$

- 実数値をとるガウス過程しか考えていないため、 f の i 番目の要素 $f_i: \mathbb{R}^{n_x+n_u} \rightarrow \text{rv}(\mathbb{R})$ が平均関数 μ_i 、共分散カーネル k_i をもつ $\mathbb{X} \times \mathbb{U}$ -上のガウス過程
- f_i が各要素および各時刻で独立であるとする、

$$x_{k+1} = \mu(x_k, u_k) + \text{diag}\left\{\left(\sqrt{k_i((x_k, u_k), (x_k, u_k))}\right)_{i=1}^{n_x}\right\}v_k, \quad v_k \sim \mathcal{N}(0, I) \quad (12.38)$$

$$\mu := \begin{bmatrix} \mu_1 & \cdots & \mu_{n_x} \end{bmatrix}^\top \quad (12.39)$$

- 事前の粗いモデルを事前平均 μ に用い、モデル化誤差をデータから回帰

ガウス過程回帰状態空間モデル (2) 正規分布のガウス過程への代入

\mathbb{R}^d 上のガウス過程 $f \sim \text{GP}(\mu, k)$ と $x \in \mathbb{R}^d$ に対して $f(x) \in \text{rv}(\mathbb{R})$

f と独立な $x \sim \mathcal{N}(\mu_x, \Sigma_x) \in \text{rv}(\mathbb{R}^d)$ に対して $f(x) \in \text{rv}(\mathbb{R})$

$\varrho \in \mathbb{R}^{\bar{s}}$, 正定値行列 $M \in \mathbb{R}^{\bar{s} \times \bar{s}}$ および $k_{\text{SE}} := \exp(-\|x - x'\|^2 / c^2)$, $c > 0$ を用いて

$$\mu(\cdot) := k_{\text{SE}}(x_{1:\bar{s}}, \cdot)^\top \varrho, \quad (12.40)$$

$$k(x, x') := k_{\text{SE}}(x, x') - k_{\text{SE}}(x_{1:\bar{s}}, x)^\top M^{-1} k_{\text{SE}}(x_{1:\bar{s}}, x') \quad (12.41)$$

ならば, 塔特性より

$$\mathbb{E}[f(x)] = \int_{\mathbb{R}^d} \varphi(x) \mathbb{E}[f(x)] dx = \int_{\mathbb{R}^d} \varphi(x) k_{\text{SE}}(x_{1:\bar{s}}, x)^\top dx \varrho \quad (12.42)$$

は具体的にかき下せる.

$\mathbb{E}[(f(x))^2]$, $\mathbb{E}[f_1(x)f_2(x)]$ も同様

ガウス過程回帰状態空間モデル (3) 予測分布の近似計算 (PILCO)

- ・ ガウス過程回帰状態空間モデル $x_{k+1} = f(x_k, u_k)$
- ・ 線形制御則 $u_k = K_k x_k + \beta_k$ や決定論的ガウス過程方策

$$u_k = \pi(x_k) := k_{\text{SE}}(\bar{x}_{1:N}, x_k)^\top p, \quad \bar{x}_i \in \mathbb{X}, \quad i = 1, 2, \dots, N, \quad p \in \mathbb{R}^N \quad (12.47)$$

- ・ ステージコスト

$$\ell_k(x) := 1 - \exp\left(-\frac{1}{2\sigma_\ell^2} \|x_* - x\|^2\right) \quad (12.45)$$

拡張カルマンフィルタの場合のように $\begin{bmatrix} x_k^\top & u_k^\top \end{bmatrix}^\top$ の分布を正規分布により近似する。

コスト $\mathbb{E} \left[\sum_k \ell_k(x_k) \right]$ の K_k や \bar{x}_i , p に関する偏微分も解析的な近似計算が可能であり、勾配法により最適化できる。

第 12 章

1. カーネル法

半正定値カーネル

ガウス過程回帰

再生核ヒルベルト空間

2. システム制御への応用

ベイズ最適化

ガウス過程回帰状態空間モデル