

拡張カルマンフィルター

1 はじめに

最も基本的な非線形フィルタリングである拡張カルマンフィルター（Extended Kalman filter, EKF）について説明する．はじめに，カルマンフィルター（Kalman filter, KF）の基礎となる線形最小分散推定（linear minimum mean square estimate, LMMSE）の理論について概説し，カルマンフィルターを誘導した後，拡張カルマンフィルターを誘導する．なお，ここでの説明や誘導方法は文献[1]に負うところが大きい．

2 線形最小分散推定（linear minimum mean square estimate, LMMSE）

カルマンフィルターの誘導にはいくつかの方法があるが（例えば，[2] [3] [4]），ここでは著者が分かりやすいと感じた線形最小分散推定に基づいた誘導方法を示す．よって，はじめにカルマンフィルターの基礎となる線形最小分散推定について簡単に説明する．

はじめに， x_t を時刻 t のある地点におけるある物理量の予測結果とし， y_t を時刻 t のある地点におけるある物理量の観測データと定義する．ここで，最適な推定値 x_t^{est} を2つのデータ x_t と y_t から推定する問題を考える．

x_t と y_t の真値を x_t^{true} とし， x_t と y_t の誤差をそれぞれ p_t , v_t とすると， x_t , y_t は以下のように表すことができる．

$$x_t = x_t^{\text{true}} + p_t \quad (1a)$$

$$y_t = x_t^{\text{true}} + v_t \quad (1b)$$

最適な x_t^{est} を推定するにあたり，誤差 p_t と v_t が，

$$E[p_t] = E[v_t] = 0 \quad (2)$$

$$E[p_t p_t] = P_t, \quad E[v_t v_t] = R_t \quad (3)$$

$$E[p_t v_t] = E[v_t p_t] = 0 \quad (4)$$

を満たし，かつ各誤差が正規分布（ガウス分布）に従うと仮定する．式(3)中の P_t , R_t は誤差分散を表す．

以上の準備から，最適解は，推定値 x_t^{est} を2つのデータ x_t と y_t の線形結合で表し，この推定値の誤差分散を最小にする重み α_x , α_y を求めることで得られる．

$$x_t^{\text{est}} = \alpha_x x_t + \alpha_y y_t \quad (5)$$

式(5)の期待値をとると、

$$E[x_t^{\text{est}}] = E[\alpha_x x_t] + E[\alpha_y y_t] = \alpha_x x_t^{\text{true}} + \alpha_y x_t^{\text{true}} \quad (6)$$

が得られ、

$$1 = \alpha_x + \alpha_y \quad (7)$$

のとき、 x_t^{est} は不偏推定値となる。式(7)を式(5)に代入すると、

$$x_t^{\text{est}} = \alpha_x x_t + (1 - \alpha_x) y_t \quad (8)$$

が得られ、 x_t^{est} の誤差分散 σ^2 は次式で表される。

$$\begin{aligned} \sigma^2 &= E[(x_t^{\text{est}} - x_t^{\text{true}})^2] \\ &= E[\{\alpha_x x_t + (1 - \alpha_x) y_t - x_t^{\text{true}}\}^2] \\ &= E[\{\alpha_x (x_t^{\text{true}} + p_t) + (1 - \alpha_x)(x_t^{\text{true}} + v_t) - x_t^{\text{true}}\}^2] \\ &= \alpha_x^2 E[p_t^2] + 2\alpha_x(1 - \alpha_x) E[p_t v_t] + (1 - \alpha_x)^2 E[v_t^2] \\ &= \alpha_x^2 P_t + (1 - \alpha_x)^2 R_t = (1 - \alpha_y)^2 P_t + \alpha_y^2 R_t \end{aligned} \quad (9)$$

推定値の誤差分散 σ^2 が最小となる重み α_y は、式(9)を微分することにより、以下のように求められる。

$$\frac{d(\sigma^2)}{d\alpha_y} = -2P_t + 2\alpha_y P_t + 2\alpha_y R_t = 0 \quad (10)$$

$$\alpha_y = \frac{P_t}{P_t + R_t} \quad (11)$$

以上より、最適な推定値 x_t^{est} とその誤差分散 σ^2 は次式で表される。

$$x_t^{\text{est}} = \frac{R_t}{P_t + R_t} x_t + \frac{P_t}{P_t + R_t} y_t = x_t + \frac{P_t}{P_t + R_t} (y_t - x_t) \quad (12)$$

$$\sigma^2 = (1 - \alpha_y)^2 P_t + \alpha_y^2 R_t$$

$$= \left(\frac{R_t}{P_t + R_t} \right)^2 P_t + \left(\frac{P_t}{P_t + R_t} \right)^2 R_t = \frac{R_t P_t}{P_t + R_t} \quad (13)$$

これまで示したように、推定値を予測結果 x_t と観測データ y_t の線形結合で表し、その誤差分散を最小にする推定法は線形最小分散推定と呼ばれる。

Fig.1 に、推定値 x_t^{est} と予測結果 x_t 、観測データ y_t それぞれの関係を示す。予測結果 x_t と観測データ y_t は真値からそれぞれ距離 p_t 、 v_t に位置する。 x_t と y_t は無相関であるため、真値 x_t^{true} から x_t 、 y_t へ向かうベクトルは直交関係にある。

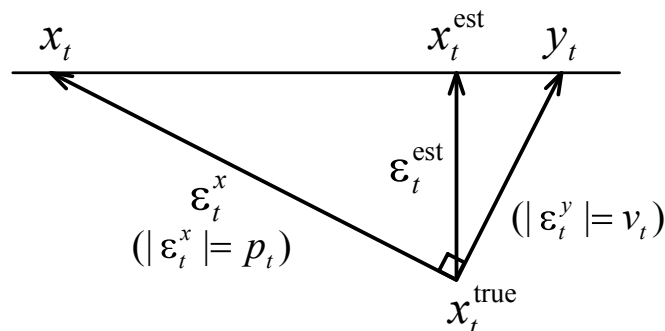


Fig.1 線形最小分散推定の幾何学的解釈[1]

最適な推定値は、予測結果 x_t 、観測データ y_t を通る直線上の、真値からの距離が最小となる点となる。また、**Fig.1** から式(13)を幾何学的に導くことも可能である。図より最適推定値の誤差は元の誤差の推定値の誤差よりも減少していることがわかる。

3 カルマンフィルタ（Kalman filter, KF）

前節の線形最小分散推定では、ある一時点のある一地点における予測結果と観測データから、最適な推定値を求める問題を考えた。ここでは時系列データに対する最適な状態推定法であるカルマンフィルタを、線形最小分散推定に基づき導出する。

対象とするシステム(問題)において、シミュレーションによって得られた予測結果を $x_1, x_2, x_3, \dots, x_T$ とする。 x_t は時刻 t における対象とするシステムの解析結果（状態変数）を表す多次元のベクトルを意味する。時刻 t の関数である状態変数 x の時間発展を与える線形作用素を $F_{t-1}(\cdot)$ とおくと、式(1a)は以下のように書き換えられる。

$$\begin{aligned} x_t &= F_{t-1}(x_{t-1}) + w_{t-1} \\ &= F_{t-1}(x_{t-1}^{\text{true}} + p_{t-1}) + w_{t-1} \end{aligned} \tag{14}$$

ここで w_{t-1} はシミュレーションによる時間発展に伴う誤差（システムノイズ、システム誤差）ベクトルである。式(14)のように状態の動的な性質を表すモデルは、一般的にシステムモデルと呼ばれる。

一方、対象とするシステムで観測されたデータを $y_1, y_2, y_3, \dots, y_T$ とし、 y_t は時刻 t において対象とするシステムの種々の地点で観測される観測データを表す多次元のベクトルとする。観測データ

に関しては以下の表現を用いる.

$$\mathbf{y}_t = \mathbf{H}_t(\mathbf{x}_t^{\text{true}}) + \mathbf{v}_t \quad (15)$$

ここで, \mathbf{H}_t は観測行列と呼ばれ, 全物理量 (状態変数) から一部のデータを取り出すためのマトリックスである. このように, システム方程式と観測方程式を連立させたシステムの表現 (モデル) は状態空間表現 (状態空間モデル, state space model) と呼ばれ, カルマンフィルタをはじめとする各種フィルタリング手法を説明する際に用いられる.

以上の準備より, 解析誤差共分散行列 \mathbf{P}_t を最小化する問題を考える. 推定値 $\mathbf{x}_t^{\text{est}}$ は, 線形最小分散推定における説明と同じように, 式(11)における $\mathbf{P}_t/(\mathbf{P}_t+\mathbf{R}_t)$ を重みマトリックス \mathbf{W}_t とおくと,

$$\begin{aligned} \mathbf{x}_t^{\text{est}} &= \mathbf{x}_t + \mathbf{W}_t(\mathbf{y}_t - \mathbf{H}_t\mathbf{x}_t) \\ &= \mathbf{F}_{t-1}(\mathbf{x}_{t-1}^{\text{true}} + \mathbf{p}_{t-1}) + \mathbf{w}_{t-1} + \mathbf{W}_t \{(\mathbf{H}_t\mathbf{x}_t^{\text{true}} + \mathbf{v}_t) - \mathbf{H}_t(\mathbf{F}_{t-1}(\mathbf{x}_{t-1}^{\text{true}} + \mathbf{p}_{t-1}) + \mathbf{w}_{t-1})\} \\ &= \mathbf{x}_t^{\text{true}} + \mathbf{F}_{t-1}\mathbf{p}_{t-1} + \mathbf{w}_{t-1} + \mathbf{W}_t(\mathbf{v}_t - \mathbf{H}_t\mathbf{F}_{t-1}\mathbf{p}_{t-1} - \mathbf{H}_t\mathbf{w}_{t-1}) \end{aligned} \quad (16)$$

と表されるため, 式(9)は以下のように書き換えられる.

$$\begin{aligned} \mathbf{P}_t &= E[(\mathbf{x}_t^{\text{est}} - \mathbf{x}_t^{\text{true}})^2] \\ &= E[(\mathbf{F}_{t-1}\mathbf{p}_{t-1} + \mathbf{w}_{t-1} + \mathbf{W}_t\mathbf{v}_t - \mathbf{W}_t\mathbf{H}_t\mathbf{F}_{t-1}\mathbf{p}_{t-1} - \mathbf{W}_t\mathbf{H}_t\mathbf{w}_{t-1})^2] \\ &= \mathbf{F}_{t-1}\mathbf{P}_{t-1}\mathbf{F}_{t-1}^T - 2\mathbf{F}_{t-1}\mathbf{P}_{t-1}\mathbf{F}_{t-1}^T\mathbf{H}_t^T\mathbf{W}_t^T + \mathbf{Q}_{t-1} - 2\mathbf{Q}_{t-1}\mathbf{H}_t^T\mathbf{W}_t^T \\ &\quad + \mathbf{W}_t\mathbf{R}_t\mathbf{W}_t^T + \mathbf{W}_t\mathbf{H}_t\mathbf{F}_{t-1}\mathbf{P}_{t-1}\mathbf{F}_{t-1}^T\mathbf{H}_t^T\mathbf{W}_t^T + \mathbf{W}_t\mathbf{H}_t\mathbf{Q}_{t-1}\mathbf{H}_t^T\mathbf{W}_t^T \end{aligned} \quad (17)$$

最適な重みマトリックスを求めるために, 行列の微分公式を用いて上式を \mathbf{W}_t で微分し, \mathbf{W}_t に関してまとめると次式が得られる.

$$\frac{\partial \text{tr}(\mathbf{P}_t)}{\partial \mathbf{W}_t} = -2\mathbf{F}_{t-1}\mathbf{P}_{t-1}\mathbf{F}_{t-1}^T\mathbf{H}_t^T - 2\mathbf{Q}_{t-1}\mathbf{H}_t^T + 2\mathbf{W}_t\mathbf{R}_t + 2\mathbf{W}_t\mathbf{H}_t\mathbf{F}_{t-1}\mathbf{P}_{t-1}\mathbf{F}_{t-1}^T\mathbf{H}_t^T + 2\mathbf{W}_t\mathbf{H}_t\mathbf{Q}_{t-1}\mathbf{H}_t^T \quad (18)$$

$$\mathbf{W}_t = (\mathbf{F}_{t-1}\mathbf{P}_{t-1}\mathbf{F}_{t-1}^T + \mathbf{Q}_{t-1})\mathbf{H}_t^T \{ \mathbf{R}_t + \mathbf{H}_t(\mathbf{F}_{t-1}\mathbf{P}_{t-1}\mathbf{F}_{t-1}^T + \mathbf{Q}_{t-1})\mathbf{H}_t^T \}^{-1} \quad (19)$$

式(19)は式(12)をマトリックス表現したものであり, 最適な重み \mathbf{W}_t がシステムノイズと観測ノイズの和に対するシステムノイズの比を表している. カルマンフィルタでは行列 \mathbf{W}_t をとくにカルマンゲイン \mathbf{K}_t と呼び, 多くの教科書では式(19)の代わりに次式が用いられている.

$$K_t = P_t H_t^T (R_t + H_t P_t H_t^T)^{-1} \tag{20}$$

$$P_t = F_{t-1} P_{t-1} F_{t-1}^T + Q_{t-1} \tag{21}$$

ここで、 P_t は時刻 t における解析誤差共分散行列である。式(21)から明らかなように、 P_t は時刻 $t-1$ における諸量から計算することができ、観測データが得られていない状態での解析誤差共分散行列の時間発展を表している。このような解析誤差分散の時間発展を表す式はリヤプノフ方程式 (Lyapunov equation) と呼ばれている。

これまでの説明を整理するために、時刻 $t = 0 \sim T$ でシステムの状態推定を行う場合のカルマンフィルターのフローを Fig.2 に示す。

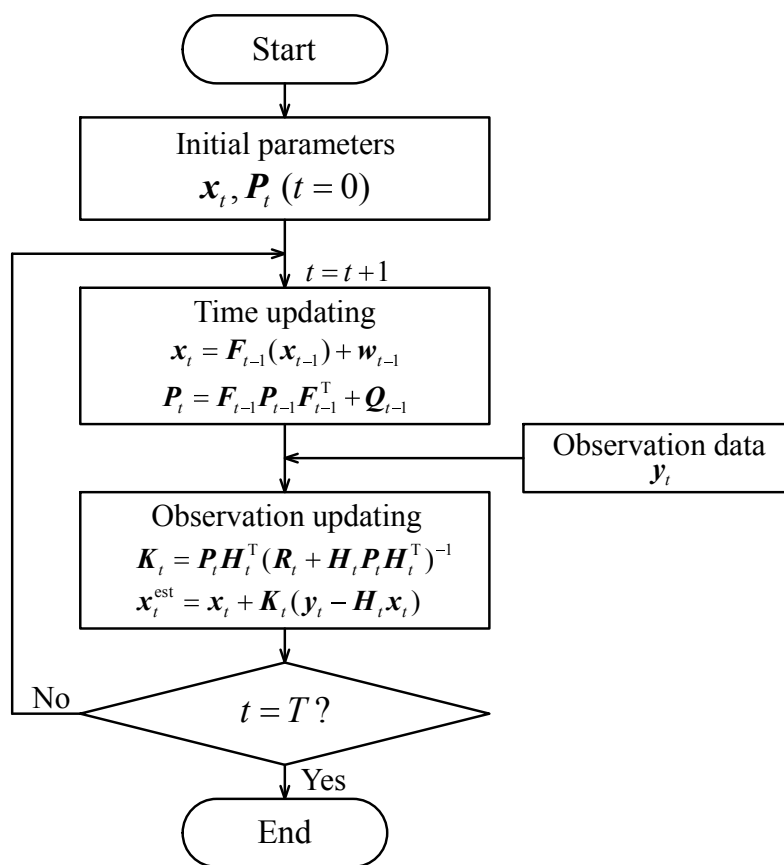


Fig.2 カルマンフィルターのフロー

4 拡張カルマンフィルター (extended Kalman filter, EKF)

カルマンフィルターでは、線形なシステム方程式 (シミュレーションモデル) に対して最適な推定値を求めたが、実際に取り扱う問題の多くは非線形である。ここでは非線形システムに適用

可能な拡張カルマンフィルタについて簡単に説明する。

以下に示す非線形状態空間モデルに対して、カルマンフィルタを適用することを考える。

$$\mathbf{x}_t = \mathbf{f}_t(\mathbf{x}_{t-1}) + \mathbf{w}_{t-1} \quad (22a)$$

$$\mathbf{y}_t = \mathbf{H}_t(\mathbf{x}_t^{\text{true}}) + \mathbf{v}_t \quad (22b)$$

ここに、 $\mathbf{f}_t(\cdot)$ は非線形作用素を表す。カルマンフィルタを非線形モデルに適用するために、非線形作用素 \mathbf{f}_t を状態ベクトル \mathbf{x}_{t-1} の推定値 $\mathbf{x}_{t-1}^{\text{est}}$ の周りで線形化（微分）し、式(14)と同様な線形作用素 $\hat{\mathbf{F}}_t$ として扱うことを考える。

$$\hat{\mathbf{F}}_t = \left. \frac{\partial \mathbf{f}_t(\mathbf{x})}{\partial \mathbf{x}} \right|_{\mathbf{x}=\mathbf{x}_{t-1}^{\text{est}}} \quad (23)$$

$\hat{\mathbf{F}}_t$ はいわゆるヤコビ行列である。以上の準備から、カルマンフィルタ説明する際に用いた式(13)と(20)の代わりに

$$\mathbf{x}_t = \mathbf{f}_t(\mathbf{x}_{t-1}^{\text{est}}) + \mathbf{w}_{t-1} \quad (14)'$$

$$\mathbf{P}_t = \hat{\mathbf{F}}_{t-1} \mathbf{P}_{t-1} \hat{\mathbf{F}}_{t-1}^T + \mathbf{Q}_{t-1} \quad (21)'$$

を用いることで、カルマンフィルタのアルゴリズムが非線形システムの最適推定問題に適用できる。これを拡張カルマンフィルタという。なお、非線形観測モデル \mathbf{h}_t を扱う場合でも、システムモデルと同様に線形化してカルマンフィルタの更新式を使うことができる。

最後に、拡張カルマンフィルタの問題点について触れる。拡張カルマンフィルタでは各時点で微分による線形化を行うため、解析的に微分できない場合には数値微分を行う必要がある。また、高次項を無視しているため、非線形性の強いシステムに対しては計算が不安定になることが知られている。このような問題に対しては、アンセンテッドカルマンフィルタ（unscented Kalman filter, UKF）やアンサンブルカルマンフィルタ（ensemble Kalman filter, EnKF）が有効となる。

参考文献

- [1] 淡路敏之，蒲池政文，池田元美，石川洋一：データ同化 観測・実験とモデルを融合するイノベーション，京都大学出版会，pp.15-19，2009.
- [2] 樋口知之，上野玄太，中野慎也，中村和幸，吉田 亮：データ同化入門—次世代のシミュレーション技術—，朝倉書店，pp.47-77，2011.
- [3] 片山 徹：新版 応用カルマンフィルタ，朝倉書店，pp.83-107，2000.
- [4] 谷萩隆嗣：カルマンフィルタと適応信号処理，コロナ社，pp.1-47，2005.