Kalman Filter まとめ

竹田航太

2023 年 9 月 12 日

目次

概要

データ同化の基本的な概念である状態空間モデルと Kalman Filter についてまとめる.

1 状態空間モデル

Definition 1.1 (状態空間モデル)**.** 直接「見る」ことができない状態空間 $\mathbb{R}^N$ を仮定する， 観測空間 $\mathbb{R}^p$*:* 状態空間を観測したもの．状態空間の時間並進と観測は以下の方程式に従う．

$$x_i = M(x_{i-1}) + G_i(\xi_i) \tag{1.1}$$
$$y_i = H_i(x_i) + \eta_i \tag{1.2}$$

ただし,

- $M_i : \mathbb{R}^N \to \mathbb{R}^N$ は状態並進モデル.
- $H_i : \mathbb{R}^N \to \mathbb{R}^p$ は観測作用素.
- ξ_i はモデルノイズ. *Gaussian* とは限らない. 平均は $o \in \mathbb{R}^K$, 誤差共分散は正定値行列 $Q \in \mathbb{R}^{K \times K}$
- $G_i : \mathbb{R}^K \to \mathbb{R}^N$
- η_i は観測ノイズ. *Gaussian* を仮定. 平均は $o \in \mathbb{R}^p$, 誤差共分散は正定値行列 $R \in \mathbb{R}^{p \times p}$

Remark 1.2 (状態空間モデル構成のヒント)**.** 状態空間モデルを使うことで様々な現象を記述できる.

- 基本は状態空間はモデルを説明するための仮想的な空間なので一般に高次元であり, その一部が観測されて見えるようになるという考え方である.
- 自己回帰モデルや移動平均モデルは上記の状態空間モデルにより表現できる.
- $K = 1$ や $K = p$ となることが多い.

2　Kalman Filter のアルゴリズム

データ同化の気持ち:

データ同化はシミュレーションとデータを組み合わせてより良い推定をすることが目的. また根底には全ての推定, 観測には誤差があるという考え方があり, 誤差の推定も同時に扱う.

データ同化のステップ:

大きく forecast(予報) と analysis(解析) に分けられる.

forecast: モデルに従ってシミュレーションを行い, 観測データが得られている時刻の推定結果 x^f を返す.

analysis: 推定結果 x^f を観測データ y^o と誤差を元に補正した x^a を返す.

この 2 つを交互に行うことで真の状態を時系列で推定する. この時推定誤差 P^f, P^a も同時に時間発展していく.

2.1　Linear Kalman Filter(LKF)[1, Kalman1960]

最初に考案された基本的な Kalman Filter.

Assumption 2.1. 真のモデルは線形 *($M = \mathbf{M} \in \mathbb{R}^{N \times N}$)*, 観測も線形 *($H \in \mathbb{R}^{p \times N}$)* とする.

Initialization: (x_0, P_0)

推定値と推定誤差共分散の初期値を設定する. x_0 には適当な値を入れ, P_0 は大きな値にしておく.

Forecast: input $(x^a_{i-1}, P^a_{i-1}) \to$ output (x^f_i, P^f_i)

各時刻 i でモデル $\mathbf{M}$ に従って時間を進める．

$$x^f_i = \mathbf{M}x^a_{i-1} \tag{2.1}$$

誤差共分散行列も同時に発展する．

$$P^f_i = \mathbf{M}P^a_{i-1}\mathbf{M}^\top + Q \tag{2.2}$$

Analysis: input $(x^f_i, P^f_i, y^o) \to$ output (x^a_i, P^a_i)

予報誤差 P^f_i と観測誤差 R から最適な重みを決め，重み付き平均をとる．
推定と誤差共分散の発展は以下．

$$x^a_i = x^f_i + K_i(y^o - H_i x^f_i) \tag{2.3}$$

$$P^a_i = (i - K_i H_i)P^f_i \tag{2.4}$$

ただし，Kalman gain K_i(最適な重み) は以下で与えられる．

$$K_i = P^f_i H^\top_i (H_i P^f_i H^\top_i + R)^{-1} \tag{2.5}$$

2.2 Extended Kalman Filter(ExKF)

LKF で M を非線型の場合に拡張 (Extend) する．P^f を計算する際に M を線形化する必要がある．$\delta \ll 1$ に対して，$\mathbf{e}_j$ を $\mathbb{R}^N$ の j 方向の単位ベクトルとして，Taylor 近似より

$$M(x + \delta\mathbf{e}_j) = M(x) + \mathbf{M}_x \delta\mathbf{e}_j + O(\delta^2)$$

この近似から $\mathbf{M}$ の j 列目は以下で与えられる．

$$\mathbf{M}_j = \mathbf{M}\mathbf{e}_j = \frac{(M(x + \delta\mathbf{e}_j) - M(x))}{\delta} \tag{2.6}$$

この線形化の計算に時間がかかる．

2.3 3 次元変分法 (3Dvar)

3 次元変分法．この「3 次元」というのは気象分野に由来する命名で空間 3 次元という意味．Kalman Filter とは導出方法は異なるが，ExKF で $P^f = P^a = P_0$ と推定誤差共分散に定数を入れたものと同値．計算が高速．

2.4 Ensemble Kalman Filter(EnKF)[2, Evensen1994]

ExKF での M の線形化の計算を回避するため．アンサンブルメンバーによって誤差を表現する．具体的には以下の m 個のメンバーを用いて推定を行う．

$$X = (x^{(1)}, \cdots, x^{(m)}) \tag{2.7}$$

Initialization: (X_0, P_0)

推定値と推定誤差共分散の初期値を設定する．$X_0 = (x_0^{(1)}, \cdots, x_0^{(m)})$ には適当な値を入れ，P_0 は大きな値にしておく．
例: Lorenz96 では X_0 はアトラクター上の任意の m 個の点，$P_0 = 25I$

Forecast: input $(x_{i-1}^{a(k)}, P_{i-1}^a) \to$ output $(x_i^{f(k)}, P_i^f)$

各時刻 i，各粒子でモデル $\mathbf{M}$ に従って時間を進める．

$$x_i^{f(k)} = \mathbf{M} x_{i-1}^{a(k)} \tag{2.8}$$

誤差共分散行列は統計的に計算する．$\delta X_i^f = (x_i^{f(1)} - \bar{x}_i^f, \cdots, x_i^{f(m)} - \bar{x}^f{}_i)$

$$P_i^f = \frac{1}{m-1} \delta X_i^f \delta X_i^{f\top} \in \mathbb{R}^{N \times N} \tag{2.9}$$

ただし，$\bar{x}_i^f = \frac{1}{m} \sum_{k=1}^m x_i^{f(k)}$

Analysis: input $(x_i^{f(k)}, P_i^f, y^o) \to$ output $(x_i^{a(k)}, P_i^a)$

予報誤差 P_i^f と観測誤差 R から最適な重みを決め，重み付き平均をとる．
推定は以下で行い，P^a は計算しない．

$$x_i^{a(k)} = x_i^{f(k)} + K_i(y^o - H_i x_i^{f(k)}) \tag{2.10}$$

ただし，Kalman gain K_i(最適な重み) は以下で与えられる．

$$K_i = P_i^f H_i^\top (H_i P_i^f H_i^\top + R)^{-1} \tag{2.11}$$

Remark 2.2 (実装上の注意)**.** 実用上 $m \ll N(\sim O(10^8))$ である場合が多いので，*EnKF* の実装ではメモリ節約のため $N \times N$ 行列は明示的に計算せず *Kalman gain* K_i を計算する．

Remark 2.3 (アルゴリズムの欠陥)**.** 上記のアルゴリズムの *Analysis step* で $X^a = (x^{a(1)}, \cdots, x^{a(m)})$ の統計的共分散を計算すると，$P^a = (I - KH)P^f(I - KH)^\top \neq (I - KH)P^f$ となって最適な共分散に一致しない．

2.5 Perturbed Obsevation method(PO 法)[3, Burgers1998]

EnKF では推定誤差共分散が最適な値に一致しないというアルゴリズムの欠陥がある．観測に摂動を加えることでこれを修正する，EnKF の Analysis step の (2.10) で

$$y^o \to y^o + \epsilon^{(k)} \tag{2.12}$$

と置き換える．ただし，$\epsilon^{(k)}$ は平均 $o \in \mathbb{R}^N$ 誤差共分散が R の正規分布に従う．

2.6 inflation と localization

Kalman Filter を安定的に動かす数値計算上のテクニックがある．

covariance inflation

数値計算上，誤差共分散行列は過小評価されることがある．その結果モデルを信用しすぎてモデルのカオス性から推定誤差が発達し非常に大きくなってしまう現象が起きる (filter divergence)．これを回避するためデータを同化する際に (Analysis step で) 誤差共分散を少し大きく補正する．このことを inflation という．方法はいくつかある．

(1) multiplicative inflation: 誤差共分散に定数 α をかける．一般的な方法.

(2) addtive inflation: 誤差共分散行列に定数行列を足す．足す行列の形は単位行列の定数倍などいろいろ考えられる．

(3) adaptive inflation: （これは他の 2 つと並列すべきではないが）各 step で適切な inflation 大きさを推定する．

Obsevation error inflation

観測誤差共分散を大きく補正することもあるらしい．

Covariance localization

EnKF において，予測誤差共分散行列の rank を改善するために用いられる．アンサンブルサイズが小さいときに必要となる．

Obsevation localization

時間発展する場を同化する問題に適用される．同化する位置と観測の位置が一致しない場合に距離に応じて観測の信頼度を下げる方法.

2.7 Ensemble Square Root Filter(EnSRF)[4, Whitaker2002]

PO とは別のアプローチで EnKF の欠陥を修正し最適な誤差共分散を達成する．計算時に平方根をとる操作が入るので Square Root Filter と呼ばれる．

2.8 Ensemble Transform Kalman Filter(ETKF)[5, Bishop2001]

Analysis step で平均からの Perturbation δX^f を適切な線形変換 (Transform) により update する．具体的には $P^a = \frac{1}{m-1}\delta X^a \delta X^{a\top} = (I - KH)P^f$ を満たすように

$$\delta X^a = \delta X^f T$$

という線形変換 T を定める．

$$T = [I - \delta Y^\top (\delta Y \delta Y^\top + R)^{-1} \delta Y]^{1/2} \in \mathbb{R}^{m\times m} \tag{2.13}$$

と線形変換 T を定めると良い．ただし，$\delta X^f = X^f - \bar{x^f}, \delta Y = H\delta X^f$.

注意: 平方根 T は一意には定まらない．直交行列 U を用いて $\tilde{T} = TU$ としても，$\tilde{T}\tilde{T}^\top = TU(TU)^\top = TUU^\top T^\top = TT^\top$.

2.9 Local Ensemble Transform Kalman Filter(LETKF)[6, Hunt2007]

遠くの観測の影響を同化の際に小さくする localization という手法を ETKF に適用する．計算量が増える．

3 全体のまとめ

3.1 線形

LKF

$$x^f = \mathbf{M}x^a \tag{3.1}$$

$$P^f = \mathbf{M}P^a\mathbf{M}^T + Q \tag{3.2}$$

$$x^a = x^f + K(y^o - Hx^f) \tag{3.3}$$

$$P^a = (I - KH)P^f \tag{3.4}$$

$$K = P^f H^T (HP^f H^T + R)^{-1} \tag{3.5}$$

- M が非線型の場合に対応していない．

3.2 非線型

ExKF

$$x^f = M(x^a) \tag{3.6}$$
$$P^f = \mathbf{M}P^a\mathbf{M}^T + Q \tag{3.7}$$
$$x^a = x^f + K(y^o - Hx^f) \tag{3.8}$$
$$P^a = (I - KH)P^f \tag{3.9}$$
$$K = P^f H^T (HP^f H^T + R)^{-1} \tag{3.10}$$

- $\mathbf{M}$ の計算が大変.

3Dvar

$P^f = P^a = P_0$

- 高速.

3.3 EnKF

EnKF

$X = (x^{(1)}, \cdots, x^{(m)}) \in \mathbb{R}^{N\times m}, \delta X^f = X^f - \bar{x^f}$

$$x^{f(k)} = M(x^{a(k)}) \qquad k = 1, \cdots, m \tag{3.11}$$
$$P^f = \frac{1}{m-1}\delta X^f \delta X^{f\,T} \tag{3.12}$$
$$x^{a(k)} = x^{f(k)} + K(y^o - Hx^f) \qquad k = 1, \cdots, m \tag{3.13}$$
$$K = P^f H^T (HP^f H^T + R)^{-1} \tag{3.14}$$

- $\mathbf{M}$ の計算をしなくてすむ. - $P^a \neq (I - KH)P^f$ となり欠陥.

3.3.1 PO

PO 法

$X = (x^{(1)}, \cdots, x^{(m)}) \in \mathbb{R}^{N\times m}, \delta X^f = X^f - \bar{x^f}$

$$x^{a(k)} = x^{f(k)} + K(y^o + \epsilon^{o(k)} - Hx^f) \qquad k = 1, \cdots, m \tag{3.15}$$

- 新たなノイズを導入. - 大きいアンサンブルメンバー数が必要.

3.3.2 EnSRF

ETKF

$X = (x^{(1)}, \cdots, x^{(m)}) \in \mathbb{R}^{N\times m}, \delta X^f = X^f - \bar{x^f}, \delta Y = \frac{1}{\sqrt{m-1}} H\delta X^f$

$$\bar{x^a} = \bar{x^f} - K(H\bar{x^f} - y) \tag{3.16}$$

$$T = [I - \delta Y^T(\delta Y\delta Y^T + R)^{-1}\delta Y]^{1/2} \tag{3.17}$$

$$X^a = \bar{x^a} + \delta X^f T \tag{3.18}$$

- 少ないメンバー数で動く.

LETKF

$X = (x^{(1)}, \cdots, x^{(m)}) \in \mathbb{R}^{N\times m}, \delta X^f = X^f - \bar{x^f}, \delta Y = H\delta X^f$ とする．各 $i = 1, \ldots, N$ で $\boldsymbol{\rho}_i = [\rho(d(i,j)/l)]_{j=1}^N$ とおく，ただし，$\rho : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ は localization 関数，$l > 0$ は localization 半径.

$$\tilde{P^a}_i = \left[\frac{m-1}{\alpha} I + \delta Y^T \boldsymbol{\rho}_i \circ R^{-1}\delta Y\right]^{-1} \tag{3.19}$$

$$T_i = \tilde{P^a}\delta Y \boldsymbol{\rho}_i \circ R^{-1}(y^o - H\bar{x^f}) + [(m-1)\tilde{P^a}]^{1/2} \tag{3.20}$$

$$X^a[i] = \bar{x^f}[i] + \delta X^f T_i[i] \tag{3.21}$$

$\boldsymbol{a} \in \mathbb{R}^N$ と $A \in \mathbb{R}^{N\times N}$ に対して，$\boldsymbol{a}A = (a_j A_{i,j})_{i,j=1}^N \in \mathbb{R}^{N\times N}$ と定める.

- 計算量が多い. - 並列化可能.

3.4 その他

- Uncented KF

参考文献

[1] R. E. Kalman. A New Approach to Linear Filtering and Prediction Problems. *Journal of Basic Engineering*, 82(1):35–45, 03 1960.

[2] Geir Evensen. Sequential data assimilation with a nonlinear quasi-geostrophic model using monte carlo methods to forecast error statistics. *Journal of Geophysical Research: Oceans*, 99(C5):10143–10162, 1994.

[3] Gerrit Burgers, Peter Jan van Leeuwen, and Geir Evensen. Analysis Scheme in the Ensemble Kalman Filter. *Monthly Weather Review*, 126(6):1719–1724, 06 1998.

[4] Jeffrey S. Whitaker and Thomas M. Hamill. Ensemble Data Assimilation without Perturbed Observations. *Monthly Weather Review*, 130(7):1913–1924, 07 2002.

[5] Craig Bishop, Brian Etherton, and Sharanya Majumdar. Adaptive sampling with the ensemble transform kalman filter. part i: Theoretical aspects. *Monthly Weather Review - MON WEATHER REV*, 129, 03 2001.

[6] Brian R. Hunt, Eric J. Kostelich, and Istvan Szunyogh. Efficient data assimilation for spatiotemporal chaos: A local ensemble transform kalman filter. *Physica D: Nonlinear Phenomena*, 230(1):112 – 126, 2007. Data Assimilation.