

アンサンブル変換カルマンフィルタ

中野 慎也

「データ科学: 理論から実用へ」

14 September 2023

モンテカルロ近似

- モンテカルロ近似では、 $p(x)$ をガウス分布 $\mathcal{N}(0, \sigma^2 \mathbf{I})$ としたとき、平均値の推定誤差が σ/\sqrt{N} 程度ある。
- 誤差は抑えるには、ある程度、アンサンブルメンバー数が大きくする必要がある。
- しかし、アンサンブルカルマンフィルタでは、シミュレーション計算

$$\mathbf{x}_{k|k-1}^{(i)} = \mathbf{f}_k(\mathbf{x}_{k-1|k-1}^{(i)})$$

を N 回繰り返す必要がある。

- 大規模なシミュレーションモデルで、何回もシミュレーション計算を実行するのは不可能。現実には、計算機資源の制約でアンサンブルメンバー数 N が数十程度に抑えられる場合が多い。

ランダム性を回避した定式化

そこで、ランダムサンプリングをせずに、アンサンブルを構成してみる。
 状態変数の次元を m とする。平均 $\bar{x}_{k|k}$ ，分散共分散行列 $\mathbf{P}_{k|k}$ のガウス分布は，
 $\mathbf{P}_{k|k}$ のランクが $N - 1$ のとき，以下の式を満たす N 個の m 次元ベクトル
 $\{\mathbf{x}_{k|k}^{(1)}, \dots, \mathbf{x}_{k|k}^{(N)}\}$ で表現できる：

$$\bar{x}_{k|k} = \frac{1}{N} \sum_{i=1}^N \mathbf{x}_{k|k}^{(i)}, \quad \mathbf{P}_{k|k} = \frac{1}{N} \sum_{i=1}^N (\mathbf{x}_{k|k}^{(i)} - \bar{x}_{k|k})(\mathbf{x}_{k|k}^{(i)} - \bar{x}_{k|k})^{\top}. \quad (1)$$

ここでは，縦ベクトルの集合 $\{\mathbf{x}_{k|k}^{(1)}, \dots, \mathbf{x}_{k|k}^{(N)}\}$ をアンサンブルと呼び，アンサンブルを構成する各縦ベクトル $\mathbf{x}_{k|k}^{(1)}, \dots, \mathbf{x}_{k|k}^{(N)}$ をアンサンブルメンバー，あるいは粒子と呼ぶことにする。

$N = \text{rank} \mathbf{P}_{k|k} + 1 \ll \dim \mathbf{x}_k$ であれば，状態 \mathbf{x}_k の平均ベクトル $\bar{x}_{k|k}$ ，分散共分散行列 $\mathbf{P}_{k|k}$ を効率的に表現できることになる。

ランダム性を回避した定式化

ここで行列 $\mathbf{X}_{k|k}$ を

$$\mathbf{X}_{k|k} = \frac{1}{\sqrt{N}} \begin{pmatrix} \mathbf{x}_{k|k}^{(1)} - \bar{\mathbf{x}}_{k|k} & \mathbf{x}_{k|k}^{(2)} - \bar{\mathbf{x}}_{k|k} & \cdots & \mathbf{x}_{k|k}^{(N)} - \bar{\mathbf{x}}_{k|k} \end{pmatrix}, \quad (2)$$

と定義する．分散共分散行列 $\mathbf{P}_{k|k}$ は $\mathbf{P}_{k|k} = \mathbf{X}_{k|k} \mathbf{X}_{k|k}^T$ と書ける．この行列 $\mathbf{X}_{k|k}$ を用いると，平均ベクトル $\bar{\mathbf{x}}_{k|k}$ ，分散共分散行列 $\mathbf{P}_{k|k}$ のガウス分布に従う乱数が，以下の式から生成できる：

$$\mathbf{x}_k | \mathbf{y}_{1:k} \sim \bar{\mathbf{x}}_{k|k} + \mathbf{X}_{k|k} \mathbf{z}_k, \quad (\mathbf{z}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{I})). \quad (3)$$

$\mathbf{X}_{k|k}$ は縦長の行列で， \mathbf{x}_k よりも \mathbf{z}_k の方が圧倒的に次元が小さいのが普通なので，主成分分析と同様に，高次元のベクトルの不確かさを少数の潜在変数で表現していることになる．

アンサンブルの性質

アンサンブル平均を分布の平均と一致させるために,

$$\frac{1}{N} \sum_{i=1}^N (\mathbf{x}_{k|k}^{(i)} - \bar{\mathbf{x}}_{k|k}) = 0. \quad (4)$$

が成り立つ必要がある．そのため， $\mathbf{1} = (1 \ \cdots \ 1)^\top$ というベクトルを定義したとき，

$$\mathbf{X}_{k|k} \mathbf{1} = \frac{1}{\sqrt{N}} \begin{pmatrix} \mathbf{x}_{k|k}^{(1)} - \bar{\mathbf{x}}_{k|k} & \cdots & \mathbf{x}_{k|k}^{(N)} - \bar{\mathbf{x}}_{k|k} \end{pmatrix} \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = \mathbf{0} \quad (5)$$

となる．

アンサンブルの性質

$\mathbf{X}_{k|k} \mathbf{1} = \mathbf{0}$ が成り立つので、 $\mathbf{X}_{k|k}$ は必ず 0 固有値を (通常 1 個) 持つ。また、

$$\mathbf{X}_{k|k} = \mathbf{U}_x \mathbf{\Lambda} \mathbf{U}_z^T$$

$$= \begin{pmatrix} \mathbf{u}_x^{(1)} & \cdots & \mathbf{u}_x^{(m)} \end{pmatrix} \begin{pmatrix} \lambda_1 & & \mathbf{0} & 0 \\ & \ddots & & \vdots \\ \mathbf{0} & & \lambda_{N-1} & 0 \\ & & & \vdots \\ & \mathbf{0}_{(m-N+1) \times (N-1)} & & 0 \end{pmatrix} \begin{pmatrix} \mathbf{u}_z^{(1)T} \\ \vdots \\ \mathbf{u}_z^{(N-1)T} \\ \mathbf{u}_z^{(N)T} \end{pmatrix} \quad (6)$$

のように特異値分解したとき、 $\mathbf{u}_z^{(N)}$ を以下のようにできるはずである。

$$\mathbf{u}_z^{(N)} = \frac{1}{\sqrt{N}} \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}. \quad (7)$$

アンサンブルの性質

行列 $\mathbf{X}_{k|k}$ を

$$\mathbf{X}_{k|k} = \frac{1}{\sqrt{N}} \begin{pmatrix} \mathbf{x}_{k|k}^{(1)} - \bar{\mathbf{x}}_{k|k} & \cdots & \mathbf{x}_{k|k}^{(N)} - \bar{\mathbf{x}}_{k|k} \end{pmatrix},$$

と定義したことを思い出すと、先程の生成式

$$\mathbf{x}_k | \mathbf{y}_{1:k} \sim \bar{\mathbf{x}}_{k|k} + \mathbf{X}_{k|k} \mathbf{z}_k, \quad (\mathbf{z}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{I}))$$

において、各粒子 $\mathbf{x}_{k|k}^{(i)}$ は、 \mathbf{z}_k を i 番目の成分が \sqrt{N} で残りの成分が 0 となるようなベクトル $\mathbf{z}_k^{(i)}$ にしたとき、つまり

$$\begin{pmatrix} \mathbf{z}_k^{(1)} & \cdots & \mathbf{z}_k^{(N)} \end{pmatrix} = \begin{pmatrix} \sqrt{N} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \sqrt{N} \end{pmatrix} \quad (8)$$

としたときに得られる。

- 各 $\mathbf{z}_k^{(i)}$ に対する $p(\mathbf{z})$ の値 $p(\mathbf{z}_k^{(i)})$ は i によらず同じ。
- 正規分布 $\mathcal{N}(\bar{\mathbf{x}}_{k|k}, \mathbf{P}_{k|k})$ から、 $\mathbf{x}_{k|k}^{(i)}$ が得られる確率密度も i によらず同じ。

アンサンブルの性質

$\mathbf{X}_{k|k}\mathbf{1} = \mathbf{0}$ が成り立つため、先程の生成式

$$\mathbf{x}_k | \mathbf{y}_{1:k} \sim \bar{\mathbf{x}}_{k|k} + \mathbf{X}_{k|k} \mathbf{z}_k, \quad (\mathbf{z}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{I}))$$

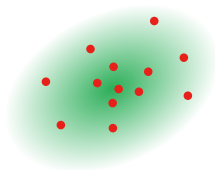
において、 \mathbf{z}_k が

$$\mathbf{z}_k = \mathbf{z}_k^{(i)} + \alpha \mathbf{1} \tag{9}$$

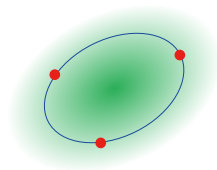
のような形であれば、 α によらず、 $\mathbf{x}_k = \mathbf{x}_k^{(i)}$ は成り立つ。

つまり、 $\mathbf{X}_{k|k}$ が 0 固有値を持つことに伴い、 $\mathbf{x}_k = \mathbf{x}_k^{(i)}$ を満たす \mathbf{z}_k は無限に存在するが、座標変換して 1 方向の成分を平均してやると、やはり、 $\mathbf{x}_k^{(i)}$ が生成される確率は i によらず同じ、ということになる。

アンサンブルの性質



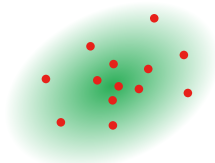
Monte Carlo representation



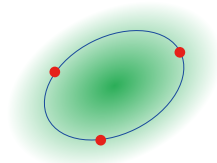
Simplex representation

- モンテカルロ法は、確率密度の高いところにたくさんの粒子を配して、確率密度分布を表現しているようなイメージ。
- 一方、非ランダムなアンサンブルでは、少ない状況では、確率密度が同じになるような場所、あるいは、分布の頂点から同じくらい離れた場所に粒子が配置されているような形になる。
- つまり、粒子の頂点を結ぶと $N - 1$ 次元の単体 (simplex) になるような状況。

アンサンブルの性質



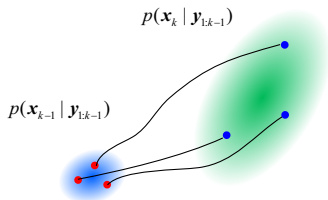
Monte Carlo representation



Simplex representation

- なお, $\text{rank} \mathbf{P}_{k|k}$ が大きく, $N < \text{rank} \mathbf{P}_{k|k}$ の場合には, モンテカルロ近似でも, $N - 1$ 次元の単体を成すような粒子配置になる.
- $\mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{I})$ からランダムにサンプル $z^{(i)}$ を生成すると, 得られるサンプルは, ほぼ長さ 1 で互いに直交ようになるため.

少数粒子による予測



- モンテカルロ法では、たくさんの粒子を使えば、大数の法則で分布の平均、分散が本来欲しいものに一致するということであった。
- ここでは、システムノイズが 0 の場合、粒子数を $N = \text{rank} \mathbf{P}_{k|k} + 1$ とした単体による近似で、

$$\mathbf{x}_{k|k-1}^{(i)} = \mathbf{f}_k(\mathbf{x}_{k-1|k-1}^{(i)}) \quad (10)$$

を各粒子について計算すればそこそこのよい近似が得られることを、Wang et al. (2004) に概ね従って確認してみる。

平均ベクトル

予測分布の平均ベクトルは以下のように書ける:

$$\bar{\mathbf{x}}_{k|k-1} = \int \mathbf{f}_k(\mathbf{x}_{k-1}) \cdot p(\mathbf{x}_{k-1} | \mathbf{y}_{1:k-1}) d\mathbf{x}_{k-1}. \quad (11)$$

前のステップのフィルタ分布 $p(\mathbf{x}_{k-1} | \mathbf{y}_{1:k-1})$ に従う \mathbf{x}_{k-1} を

$$\mathbf{x}_{k-1} | \mathbf{y}_{1:k-1} \sim \bar{\mathbf{x}}_{k-1|k-1} + \mathbf{X}_{k-1|k-1} \mathbf{z}_{k-1}, \quad (\mathbf{z}_{k-1} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})), \quad (12)$$

のように書くと, 上の式は以下のように書ける:

$$\bar{\mathbf{x}}_{k|k-1} = \int \mathbf{f}_k(\bar{\mathbf{x}}_{k-1|k-1} + \mathbf{X}_{k-1|k-1} \mathbf{z}_k) \cdot p(\mathbf{z}_{k-1}) d\mathbf{z}_{k-1}. \quad (13)$$

テイラー展開

ここで、関数 $f(\mathbf{x})$ のテイラー展開を考える:

$$\begin{aligned}
 f(\mathbf{x}_0 + \Delta \mathbf{x}) &= f(\mathbf{x}_0) + \sum_{i=1}^m \Delta x_i \frac{\partial f}{\partial x_i}(\mathbf{x}_0) + \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^m \Delta x_i \Delta x_j \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}_0) + \cdots \\
 &= f(\mathbf{x}_0) + [\nabla^\top f(\mathbf{x}_0)] \Delta \mathbf{x} + \frac{1}{2} (\Delta \mathbf{x})^\top [\nabla^2 f(\mathbf{x}_0)] (\Delta \mathbf{x}) + \cdots
 \end{aligned} \tag{14}$$

但し、

$$\nabla^\top f = \left(\frac{\partial f}{\partial x_1} \quad \cdots \quad \frac{\partial f}{\partial x_m} \right), \quad \nabla^2 f = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_m \partial x_1} & \cdots & \frac{\partial^2 f}{\partial x_m \partial x_m} \end{pmatrix}.$$

(ハーヴィル: 統計のための行列代数 [上]などを参照)

平均ベクトル

したがって、 $f_k(\bar{\mathbf{x}}_{k-1|k-1} + \mathbf{X}_{k-1|k-1}\mathbf{z}_{k-1})$ のテイラー展開は、

$$\begin{aligned} & f_k(\bar{\mathbf{x}}_{k-1|k-1} + \mathbf{X}_{k-1|k-1}\mathbf{z}_{k-1}) \\ &= f_k(\bar{\mathbf{x}}_{k-1|k-1}) + \tilde{\mathbf{F}}_{\bar{\mathbf{x}}_{k-1|k-1}} \mathbf{X}_{k-1|k-1} \mathbf{z}_{k-1} \\ & \quad + \frac{1}{2} [\mathbf{X}_{k-1|k-1} \mathbf{z}_{k-1}]^\top [\nabla^2 f_k(\bar{\mathbf{x}}_{k-1|k-1})] [\mathbf{X}_{k-1|k-1} \mathbf{z}_{k-1}] + \cdots \end{aligned} \quad (15)$$

但し、 $\tilde{\mathbf{F}}_{\bar{\mathbf{x}}_{k-1|k-1}}$ は f_k のヤコビ行列．これを使って、式 (13) の積分

$$\bar{\mathbf{x}}_{k|k-1} = \int f_k(\bar{\mathbf{x}}_{k-1|k-1} + \mathbf{X}_{k-1|k-1}\mathbf{z}_{k-1}) \cdot p(\mathbf{z}_{k-1}) d\mathbf{z}_{k-1}$$

を $p(\mathbf{z}_{k-1}) = \mathcal{N}(\mathbf{0}, \mathbf{I})$ のもとで計算する．

平均ベクトル

$$\begin{aligned}
 & \mathbf{f}_k(\bar{\mathbf{x}}_{k-1|k-1} + \mathbf{X}_{k-1|k-1} \mathbf{z}_{k-1}) \\
 &= \mathbf{f}_k(\bar{\mathbf{x}}_{k-1|k-1}) + \tilde{\mathbf{F}}_{\bar{\mathbf{x}}_{k-1|k-1}} \mathbf{X}_{k-1|k-1} \mathbf{z}_{k-1} \\
 & \quad + \frac{1}{2} [\mathbf{X}_{k-1|k-1} \mathbf{z}_{k-1}]^\top \nabla^2 \mathbf{f}_k(\bar{\mathbf{x}}_{k-1|k-1}) [\mathbf{X}_{k-1|k-1} \mathbf{z}_{k-1}] + \cdots
 \end{aligned}$$

の右辺第 1 項は \mathbf{z}_{k-1} に依存しないので,

$$\int \mathbf{f}_k(\bar{\mathbf{x}}_{k-1|k-1}) \cdot p(\mathbf{z}_{k-1}) d\mathbf{z}_{k-1} = \mathbf{f}_k(\bar{\mathbf{x}}_{k-1|k-1}).$$

また、右辺第 2 項については $p(\mathbf{z}_{k-1}) = \mathcal{N}(\mathbf{0}, \mathbf{I})$ を掛けて積分すると 0 となる:

$$\int \tilde{\mathbf{F}}_{\bar{\mathbf{x}}_{k-1|k-1}} \mathbf{X}_{k-1|k-1} \mathbf{z}_{k-1} \cdot p(\mathbf{z}_{k-1}) d\mathbf{z}_{k-1} = \mathbf{0}.$$

平均ベクトル

右辺第 3 項 (2 次の項) については

$$\begin{aligned}
 & \frac{1}{2} \int [\mathbf{X}_{k-1|k-1} \mathbf{z}_{k-1}]^\top [\nabla^2 \mathbf{f}_k(\bar{\mathbf{x}}_{k-1|k-1})] [\mathbf{X}_{k-1|k-1} \mathbf{z}_{k-1}] p(\mathbf{z}_{k-1}) d\mathbf{z}_{k-1} \\
 &= \frac{1}{2} \int \mathbf{z}_{k-1}^\top \mathbf{X}_{k-1|k-1}^\top [\nabla^2 \mathbf{f}_k(\bar{\mathbf{x}}_{k-1|k-1})] \mathbf{X}_{k-1|k-1} \mathbf{z}_{k-1} \cdot p(\mathbf{z}_{k-1}) d\mathbf{z}_{k-1} \\
 &= \frac{1}{2} \text{tr} \left(\mathbf{X}_{k-1|k-1}^\top [\nabla^2 \mathbf{f}_k(\bar{\mathbf{x}}_{k-1|k-1})] \mathbf{X}_{k-1|k-1} \right).
 \end{aligned}$$

最後の式変形は、 $p(\mathbf{z}_{k-1}) = \mathcal{N}(\mathbf{0}, \mathbf{I})$ を使うと

$$\begin{aligned}
 \int \mathbf{z}^\top \mathbf{A} \mathbf{z} p(\mathbf{z}) d\mathbf{z} &= \int \sum_{i,j} (A^{ij} z^i z^j) p(z^1) \cdots p(z^N) dz^1 \cdots dz^N \\
 &= \sum_i A^{ii} = \text{tr} \mathbf{A}
 \end{aligned}$$

となることによる。

平均ベクトル

したがって、予測分布の平均 $\bar{x}_{k|k-1}$ は、テイラー展開の 2 次の項まで計算すると、

$$\begin{aligned}\bar{x}_{k|k-1} &= \int \mathbf{f}_k(\bar{x}_{k-1|k-1} + \mathbf{X}_{k-1|k-1} \mathbf{z}_{k-1}) \cdot p(\mathbf{z}_{k-1}) d\mathbf{z}_{k-1} \\ &= \mathbf{f}_k(\bar{x}_{k-1|k-1}) + \frac{1}{2} \text{tr} \left(\mathbf{X}_{k-1|k-1}^\top [\nabla^2 \mathbf{f}_k(\bar{x}_{k-1|k-1})] \mathbf{X}_{k-1|k-1} \right) + \cdots\end{aligned}\tag{16}$$

となる．

アンサンブル平均

一方, $x_{k|k-1}^{(i)} = f_k(x_{k-1|k-1}^{(i)})$ のように $x_{k|k-1}^{(i)}$ を各粒子について計算した場合の標本平均も計算する.

$$\begin{pmatrix} z_{k-1}^{(1)} & \cdots & z_{k-1}^{(N)} \end{pmatrix} = \begin{pmatrix} \sqrt{N} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \sqrt{N} \end{pmatrix}$$

とくと,

$$x_{k|k-1}^{(i)} = f_k(\bar{x}_{k-1|k-1} + \mathbf{X}_{k-1|k-1} z_{k-1}^{(i)})$$

と書けることを利用し, 先程のテイラー展開の式を使って変形していく.

アンサンブル平均

第 2 項は

$$\sum_{i=1}^N \mathbf{X}_{k-1|k-1} \mathbf{z}_{k-1}^{(i)} = \sum_{i=1}^N (\mathbf{x}_{k-1|k-1}^{(i)} - \bar{\mathbf{x}}_{k-1|k-1}) = 0$$

となるので消える．第 3 項は以下の式を用いて変形する．

$$\frac{1}{N} \sum_{i=1}^N \mathbf{z}_{k-1}^{(i)\top} \mathbf{A} \mathbf{z}_{k-1}^{(i)} = \frac{1}{N} \sum_{i=1}^N N A^{ii} = \text{tr } \mathbf{A}$$

アンサンブル平均

最終的なアンサンブルの標本平均は、

$$\hat{x}_{k|k-1} = f_k(\bar{x}_{k-1|k-1}) + \frac{1}{2} \text{tr} \left(\mathbf{X}_{k-1|k-1}^T [\nabla^2 f_k(\bar{x}_{k-1|k-1})] \mathbf{X}_{k-1|k-1} \right) + \cdots \quad (17)$$

となり、解析的に求めた $\bar{x}_{k|k-1}$ の値と 2 次の項までは一致する。(計算しなかったが、3 次の項は一致しない。)

したがって、アンサンブルの平均 $\hat{x}_{k|k-1}$ により、システムモデルのテイラー展開の 2 次の項までを考慮した予測分布の平均の推定値が得られる。

分散共分散行列

予測分布の分散共分散行列については、以下を計算することになる:

$$\mathbf{P}_{k|k-1} = \int (\mathbf{f}_k(\mathbf{x}_{k-1}) - \bar{\mathbf{x}}_{k|k-1})(\mathbf{f}_k(\mathbf{x}_{k-1}) - \bar{\mathbf{x}}_{k|k-1})^\top p(\mathbf{x}_{k-1} | \mathbf{y}_{1:k-1}) d\mathbf{x}_{k-1} \quad (18)$$

$\mathbf{f}_k(\mathbf{x}_{k-1})$ のテイラー展開

$$\begin{aligned} \mathbf{f}_k(\mathbf{x}_{k-1}) &\approx \mathbf{f}_k(\bar{\mathbf{x}}_{k-1|k-1} + \mathbf{X}_{k-1|k-1} \mathbf{z}_{k-1}) \\ &= \mathbf{f}_k(\bar{\mathbf{x}}_{k-1|k-1}) + \tilde{\mathbf{F}}_{\bar{\mathbf{x}}_{k-1|k-1}} \mathbf{X}_{k-1|k-1} \mathbf{z}_{k-1} \\ &\quad + \frac{[\mathbf{X}_{k-1|k-1} \mathbf{z}_{k-1}]^\top \nabla^2 \mathbf{f}_k(\bar{\mathbf{x}}_{k-1|k-1}) [\mathbf{X}_{k-1|k-1} \mathbf{z}_{k-1}]}{2} + \dots \end{aligned}$$

と、平均

$$\bar{\mathbf{x}}_{k|k-1} = \mathbf{f}_k(\bar{\mathbf{x}}_{k-1|k-1}) + \frac{1}{2} \text{tr} \left(\mathbf{X}_{k-1|k-1}^\top [\nabla^2 \mathbf{f}_k(\bar{\mathbf{x}}_{k-1|k-1})] \mathbf{X}_{k-1|k-1} \right) + \dots$$

を考えると、

$$\mathbf{f}_k(\mathbf{x}_{k-1}) - \bar{\mathbf{x}}_{k|k-1} \approx \tilde{\mathbf{F}}_{\bar{\mathbf{x}}_{k-1|k-1}} \mathbf{X}_{k-1|k-1} \mathbf{z}_{k-1}.$$

分散共分散行列

したがって、分散共分散行列は、

$$\begin{aligned} \mathbf{P}_{k|k-1} &\approx \int \tilde{\mathbf{F}}_{\bar{\mathbf{x}}_{k-1|k-1}} \mathbf{X}_{k-1|k-1} \mathbf{z}_{k-1} \mathbf{z}_{k-1}^{\top} \mathbf{X}_{k-1|k-1}^{\top} \tilde{\mathbf{F}}_{\bar{\mathbf{x}}_{k-1|k-1}}^{\top} p(\mathbf{z}_{k-1}) d\mathbf{z}_{k-1} \\ &= \tilde{\mathbf{F}}_{\bar{\mathbf{x}}_{k-1|k-1}} \mathbf{X}_{k-1|k-1} \mathbf{X}_{k-1|k-1}^{\top} \tilde{\mathbf{F}}_{\bar{\mathbf{x}}_{k-1|k-1}}^{\top}. \end{aligned} \quad (19)$$

一方、アンサンブルの標本分散共分散行列は、

$$\begin{aligned} \hat{\mathbf{P}}_{k|k-1} &= \frac{1}{N} \sum_{i=1}^N \left[\mathbf{f}_k(\mathbf{x}_{k-1|k-1}^{(i)}) - \hat{\mathbf{x}}_{k|k-1} \right] \left[\mathbf{f}_k(\mathbf{x}_{k-1|k-1}^{(i)}) - \hat{\mathbf{x}}_{k|k-1} \right]^{\top} \\ &\approx \tilde{\mathbf{F}}_{\bar{\mathbf{x}}_{k-1|k-1}} \mathbf{X}_{k-1|k-1} \mathbf{X}_{k-1|k-1}^{\top} \tilde{\mathbf{F}}_{\bar{\mathbf{x}}_{k-1|k-1}}^{\top}. \end{aligned} \quad (20)$$

したがって、標本分散共分散行列 $\hat{\mathbf{P}}_{k|k-1}$ は、予測分布の分散共分散行列 $\mathbf{P}_{k|k-1}$ を、システムノイズの部分を除いて 2 次の項まで近似する。

問題はシステムノイズの寄与 \mathbf{Q} の扱いだが、これについては後で議論する。

線型の場合

線型システムの場合、 $\mathbf{x}_{k|k-1}^{(i)} = \mathbf{f}_k(\mathbf{x}_{k-1|k-1}^{(i)})$ の代わりに

$$\mathbf{x}_{k|k-1}^{(i)} = \mathbf{F}_k \mathbf{x}_{k-1|k-1}^{(i)} \quad (21)$$

で各粒子 $\mathbf{x}_{k|k-1}^{(i)}$ が得られる。

こうして得られた粒子 $\{\mathbf{x}_{k|k-1}^{(i)}\}$ のアンサンブル平均は、

$$\hat{\mathbf{x}}_{k|k-1} = \frac{1}{N} \sum_{i=1}^N \mathbf{F}_k \mathbf{x}_{k-1|k-1}^{(i)} = \mathbf{F}_k \bar{\mathbf{x}}_{k-1|k-1} = \bar{\mathbf{x}}_{k|k-1}. \quad (22)$$

線型の場合

$\mathbf{x}_{k|k-1}^{(i)} = \mathbf{F}_k \mathbf{x}_{k-1|k-1}^{(i)}$ で各粒子 $\mathbf{x}_{k|k-1}^{(i)}$ を得た場合、アンサンブルの分散共分散行列は、

$$\begin{aligned}
 \hat{\mathbf{P}}_{k|k-1} &= \frac{1}{N} \sum_{i=1}^N \left[\mathbf{F}_k \mathbf{x}_{k-1|k-1}^{(i)} - \bar{\mathbf{x}}_{k|k-1} \right] \left[\mathbf{F}_k \mathbf{x}_{k-1|k-1}^{(i)} - \bar{\mathbf{x}}_{k|k-1} \right]^\top \\
 &= \frac{1}{N} \sum_{i=1}^N \left[\mathbf{F}_k \mathbf{x}_{k-1|k-1}^{(i)} - \mathbf{F}_k \bar{\mathbf{x}}_{k-1|k-1} \right] \left[\mathbf{F}_k \mathbf{x}_{k-1|k-1}^{(i)} - \mathbf{F}_k \bar{\mathbf{x}}_{k-1|k-1} \right]^\top \\
 &= \frac{1}{N} \sum_{i=1}^N \mathbf{F}_k \left(\mathbf{x}_{k-1|k-1}^{(i)} - \bar{\mathbf{x}}_{k-1|k-1} \right) \left(\mathbf{x}_{k-1|k-1}^{(i)} - \bar{\mathbf{x}}_{k-1|k-1} \right)^\top \mathbf{F}_k^\top \\
 &= \mathbf{F}_k \mathbf{P}_{k-1|k-1} \mathbf{F}_k^\top = \mathbf{P}_{k|k-1}.
 \end{aligned} \tag{23}$$

したがって、アンサンブルの平均ベクトル、分散共分散行列は、分散共分散行列のシステムノイズの部分を除いて、カルマンフィルタの予測の式と一致する。

次元の圧縮

単体近似 ($N = \text{rank} \mathbf{P}_{k|k} + 1$) を使えば、状態 x_k の平均ベクトル $x_{k|k}$, 分散共分散行列 $\mathbf{P}_{k|k}$ が表現できるということだったが...

- 粒子の数 N だけ,

$$x_{k|k-1}^{(i)} = f_k(x_{k-1|k-1}^{(i)})$$

の計算が必要. つまり, N 回のシミュレーション計算を繰り返す必要がある.

- 大規模なシミュレーションモデルでは, そこまで何回もシミュレーション計算を実行するのが不可能な場合も多い.
- 現実には, システムの自由度 ($\text{rank} \mathbf{P}_{k|k}$) よりも, 計算機資源の制約で粒子数 N が決まることの方が多い.



この場合のアンサンブルは, 分散共分散行列の次元 (ランク) を落とした近似.

次元の圧縮

粒子数 N を，分散共分散行列 $\mathbf{P}_{k-1|k-1}$ の本当のランクよりも粒子数を少なく取った場合， $\mathbf{X}_{k-1|k-1} \mathbf{X}_{k-1|k-1}^\top$ は $\mathbf{P}_{k-1|k-1}$ と一致しなくなる．

別途， $\mathbf{X}_{k-1|k-1}^*$ を $\mathbf{P}_{k-1|k-1} = \mathbf{X}_{k-1|k-1}^* \mathbf{X}_{k-1|k-1}^{*\top}$ と $\mathbf{X}_{k-1|k-1}^* \mathbf{1} = \mathbf{0}$ が成り立つように定義すると，

$$\hat{x}_{k|k-1}^* = f_k(\bar{x}_{k-1|k-1}) + \frac{1}{2} \text{tr} \left(\mathbf{X}_{k-1|k-1}^{*\top} \left[\nabla^2 f_k(\bar{x}_{k-1|k-1}) \right] \mathbf{X}_{k-1|k-1}^* \right) + \dots$$

に対し，

$$\bar{x}_{k|k-1} = f_k(\bar{x}_{k-1|k-1}) + \frac{1}{2} \text{tr} \left(\mathbf{X}_{k-1|k-1}^\top \left[\nabla^2 f_k(\bar{x}_{k-1|k-1}) \right] \mathbf{X}_{k-1|k-1} \right) + \dots$$

なので，2 次の項も一部一致しなくなることに注意する．

少数の粒子によるアンサンブルカルマンフィルタ

- アンサンブルカルマンフィルタのアルゴリズムでは、観測ノイズの分布 $p(w_k)$ からのランダムサンプル $w_k^{(i)}$ を使うため、粒子数が少ない場合には、モンテカルロ誤差の影響が大きくなる可能性がある。
- 粒子数を少なく取った場合に適した方法として、乱数を使わずにフィルタリングを行うアンサンブル平方根フィルタ (Tippett et al., 2003; Livings et al., 2008) というカテゴリーに属する方法が幾つか存在する。
- 以下では、アンサンブル平方根フィルタの一種であるアンサンブル変換カルマンフィルタ (Bishop et al., 2001) を紹介する。

アンサンブル変換カルマンフィルタ (ETKF)

アンサンブル変換カルマンフィルタ (ensemble transform Kalman filter; ETKF) のおおまかな流れ:

- 1 予測分布を表現するアンサンブル $\{\mathbf{x}_{k|k-1}^{(1)}, \dots, \mathbf{x}_{k|k-1}^{(N)}\}$ から、平均ベクトル $\bar{\mathbf{x}}_{k|k-1}$ ，分散共分散行列の平方根 $\mathbf{X}_{k|k-1}$ を

$$\bar{\mathbf{x}}_{k|k-1} = \frac{1}{N} \sum_{i=1}^N \mathbf{x}_{k|k-1}^{(i)}, \quad (24)$$

$$\mathbf{X}_{k|k-1} = \frac{1}{\sqrt{N}} \begin{pmatrix} \mathbf{x}_{k|k-1}^{(1)} - \bar{\mathbf{x}}_{k|k-1} & \cdots & \mathbf{x}_{k|k-1}^{(N)} - \bar{\mathbf{x}}_{k|k-1} \end{pmatrix}, \quad (25)$$

のように求める。

- 2 $\bar{\mathbf{x}}_{k|k-1}$, $\mathbf{X}_{k|k-1}$ をもとに、フィルタ分布の平均ベクトル $\bar{\mathbf{x}}_{k|k}$ ，分散共分散行列の平方根 $\mathbf{X}_{k|k}$ を，すぐ後で述べる方法で求める。
- 3 求めた $\bar{\mathbf{x}}_{k|k}$, $\mathbf{X}_{k|k}$ から、フィルタ分布を表現するアンサンブル $\{\mathbf{x}_{k|k}^{(1)}, \dots, \mathbf{x}_{k|k}^{(N)}\}$ を生成する。

アンサンブル変換カルマンフィルタ (ETKF)

- フィルタ分布の平均は、カルマンフィルタのアルゴリズムに従って、

$$\bar{\mathbf{x}}_{k|k} = \bar{\mathbf{x}}_{k|k-1} + \mathbf{K}_k (\mathbf{y}_k - \mathbf{H}_k \bar{\mathbf{x}}_{k|k-1}). \quad (26)$$

- 一方、分散共分散行列の平方根 $\mathbf{X}_{k|k}$ は、行列 \mathbf{T}_k を

$$\mathbf{X}_{k|k} = \mathbf{X}_{k|k-1} \mathbf{T}_k \quad (27)$$

のように後ろから掛けることによって得る．

- 但し、 \mathbf{T}_k は、以下を 2 つの条件を満たすように設計する．

$$\begin{aligned} \text{条件 1: } \mathbf{P}_{k|k} & (= \mathbf{X}_{k|k} \mathbf{X}_{k|k}^\top = \mathbf{X}_{k|k-1} \mathbf{T}_k \mathbf{T}_k^\top \mathbf{X}_{k|k-1}^\top) \\ & = \mathbf{P}_{k|k-1} - \mathbf{P}_{k|k-1} \mathbf{H}_k^\top (\mathbf{H}_k \mathbf{P}_{k|k-1} \mathbf{H}_k^\top + \mathbf{R}_k)^{-1} \mathbf{H}_k \mathbf{P}_{k|k-1}, \end{aligned}$$

$$\text{条件 2: } \mathbf{X}_{k|k} \mathbf{1} = \mathbf{X}_{k|k-1} \mathbf{T}_k \mathbf{1} = \mathbf{0}.$$

アンサンブル変換カルマンフィルタ (ETKF)

- $\bar{\mathbf{x}}_{k|k}$ と $\mathbf{X}_{k|k}$ が求まったら, $\mathbf{X}_{k|k}$ を

$$\begin{pmatrix} \delta \mathbf{x}_{k|k}^{(1)} & \cdots & \delta \mathbf{x}_{k|k}^{(N)} \end{pmatrix} = \sqrt{N} \mathbf{X}_{k|k} \quad (28)$$

のように分解して,

$$\mathbf{x}_{k|k}^{(i)} = \bar{\mathbf{x}}_{k|k} + \delta \mathbf{x}_{k|k}^{(i)} \quad (29)$$

のように各粒子を生成すればよい.

- \mathbf{T}_k に課す 2 つの条件

$$\text{条件 1:} \quad \mathbf{P}_{k|k} = \mathbf{X}_{k|k} \mathbf{X}_{k|k}^{\top} = \mathbf{X}_{k|k-1} \mathbf{T}_k \mathbf{T}_k^{\top} \mathbf{X}_{k|k-1}^{\top}$$

$$\text{条件 2:} \quad \mathbf{X}_{k|k} \mathbf{1} = \mathbf{X}_{k|k-1} \mathbf{T}_k \mathbf{1} = \mathbf{0}.$$

のうち, 条件 1 は分散共分散行列をうまく表すための条件, 条件 2 は平均ベクトルが正しく $\bar{\mathbf{x}}_{k|k}$ になるための条件.

変換行列 \mathbf{T}_k の計算

カルマンフィルタの分散共分散行列の式を変形すると,

$$\begin{aligned}
 \mathbf{P}_{k|k} &= \mathbf{X}_{k|k} \mathbf{X}_{k|k}^\top = \mathbf{X}_{k|k-1} \mathbf{T}_k \mathbf{T}_k^\top \mathbf{X}_{k|k}^\top \\
 &= \mathbf{P}_{k|k-1} - \mathbf{P}_{k|k-1} \mathbf{H}_k^\top (\mathbf{H}_k \mathbf{P}_{k|k-1} \mathbf{H}_k^\top + \mathbf{R}_k)^{-1} \mathbf{H}_k \mathbf{P}_{k|k-1}, \\
 &= \mathbf{X}_{k|k-1} \mathbf{X}_{k|k-1}^\top \\
 &\quad - \mathbf{X}_{k|k-1} \mathbf{X}_{k|k-1}^\top \mathbf{H}_k^\top (\mathbf{H}_k \mathbf{X}_{k|k-1} \mathbf{X}_{k|k-1}^\top \mathbf{H}_k^\top + \mathbf{R}_k)^{-1} \mathbf{H}_k \mathbf{X}_{k|k-1} \mathbf{X}_{k|k-1}^\top, \\
 &= \mathbf{X}_{k|k-1} \left(\mathbf{I} - \mathbf{X}_{k|k-1}^\top \mathbf{H}_k^\top (\mathbf{H}_k \mathbf{X}_{k|k-1} \mathbf{X}_{k|k-1}^\top \mathbf{H}_k^\top + \mathbf{R}_k)^{-1} \mathbf{H}_k \mathbf{X}_{k|k-1} \right) \mathbf{X}_{k|k-1}^\top, \\
 &= \mathbf{X}_{k|k-1} \left(\mathbf{I} + \mathbf{X}_{k|k-1}^\top \mathbf{H}_k^\top \mathbf{R}_k^{-1} \mathbf{H}_k \mathbf{X}_{k|k-1} \right)^{-1} \mathbf{X}_{k|k-1}^\top.
 \end{aligned} \tag{30}$$

したがって, $\mathbf{T}_k \mathbf{T}_k^\top = \left(\mathbf{I} + \mathbf{X}_{k|k-1}^\top \mathbf{H}_k^\top \mathbf{R}_k^{-1} \mathbf{H}_k \mathbf{X}_{k|k-1} \right)^{-1}$ となれば, 条件 1 は満足される.

変換行列 \mathbf{T}_k の計算

条件 1 $\mathbf{T}_k \mathbf{T}_k^\top = \left(\mathbf{I} + \mathbf{X}_{k|k-1}^\top \mathbf{H}_k^\top \mathbf{R}_k^{-1} \mathbf{H}_k \mathbf{X}_{k|k-1} \right)^{-1}$ を満たす行列 \mathbf{T}_k を求めるために、以下の固有値分解を行う:

$$\mathbf{X}_{k|k-1}^\top \mathbf{H}_k^\top \mathbf{R}_k^{-1} \mathbf{H}_k \mathbf{X}_{k|k-1} = \mathbf{U} \mathbf{\Lambda} \mathbf{U}^\top. \quad (31)$$

すると、行列 \mathbf{T}_k は

$$\mathbf{T}_k \mathbf{T}_k^\top = (\mathbf{U} \mathbf{U}^\top + \mathbf{U} \mathbf{\Lambda} \mathbf{U}^\top)^{-1} = [\mathbf{U} (\mathbf{I} + \mathbf{\Lambda}) \mathbf{U}^\top]^{-1} = \mathbf{U} (\mathbf{I} + \mathbf{\Lambda})^{-1} \mathbf{U}^\top \quad (32)$$

を満たせばよいことになる。

これを満たす一番単純な行列は、 $\mathbf{U} (\mathbf{I} + \mathbf{\Lambda})^{-\frac{1}{2}}$ であるが、 Φ を任意の直交行列とし、 $\mathbf{U} (\mathbf{I} + \mathbf{\Lambda})^{-\frac{1}{2}} \Phi^\top$ としても成り立つ。

変換行列 \mathbf{T}_k の計算

$\mathbf{T}_k = \mathbf{U}(\mathbf{I} + \mathbf{\Lambda})^{-\frac{1}{2}} \mathbf{\Phi}^\top$ の形になっていれば、条件 1 は満足される。

↓

あとは、条件 2 の

$$\mathbf{X}_{k|k} \mathbf{1} = \mathbf{X}_{k|k-1} \mathbf{U}(\mathbf{I} + \mathbf{\Lambda})^{-\frac{1}{2}} \mathbf{\Phi}^\top \mathbf{1} = \mathbf{0}$$

を満たすように直交行列 $\mathbf{\Phi}$ を決めればよい。

変換行列 T_k の計算

$X_{k|k-1} U(I + \Lambda)^{-\frac{1}{2}} \Phi^T \mathbf{1} = 0$ という条件だけでは、一意に Φ は決まらないが、 $\Phi = U$ と置いて、

$$T_k = U(I + \Lambda)^{-\frac{1}{2}} U^T \quad (33)$$

とするのが一つの手段．

$\Phi = U$ でよいというのは、 U が固有値分解

$$X_{k|k-1}^T H_k^T R_k^{-1} H_k X_{k|k-1} = U \Lambda U^T$$

から出てきたことを思い出すとわかる．

変換行列 \mathbf{T}_k の計算

$\mathbf{X}_{k|k-1}\mathbf{1} = \mathbf{0}$ なので

$$\mathbf{X}_{k|k-1}^{\top} \mathbf{H}_k^{\top} \mathbf{R}_k^{-1} \mathbf{H}_k \mathbf{X}_{k|k-1} \mathbf{1} = \mathbf{0}. \quad (34)$$

したがって、 $\mathbf{X}_{k|k-1}^{\top} \mathbf{H}_k^{\top} \mathbf{R}_k^{-1} \mathbf{H}_k \mathbf{X}_{k|k-1}$ ベクトル $\mathbf{1}$ に平行なベクトルを固有ベクトルに持ち、それに対応する固有値は 0 になる。

$$\mathbf{X}_{k|k-1}^{\top} \mathbf{H}_k^{\top} \mathbf{R}_k^{-1} \mathbf{H}_k \mathbf{X}_{k|k-1} = \mathbf{U} \mathbf{\Lambda} \mathbf{U}^{\top}$$

の \mathbf{U} の中にベクトル $\mathbf{1}$ に平行な固有ベクトルが含まれ、それに対応する固有値が 0 になるということなので、その固有ベクトルを

$$\mathbf{u}^{(N)} = \frac{1}{\sqrt{N}} \mathbf{1}$$

と書くことにする。

変換行列 \mathbf{T}_k の計算

$$\mathbf{X}_{k|k-1}^{\top} \mathbf{H}_k^{\top} \mathbf{R}_k^{-1} \mathbf{H}_k \mathbf{X}_{k|k-1} = \mathbf{U} \mathbf{\Lambda} \mathbf{U}^{\top}$$

の \mathbf{U} の中に含まれる $\mathbf{1}$ に平行な固有ベクトルを

$$\mathbf{u}^{(N)} = \frac{1}{\sqrt{N}} \mathbf{1}$$

として, $\mathbf{X}_{k|k-1}^{\top} \mathbf{H}_k^{\top} \mathbf{R}_k^{-1} \mathbf{H}_k \mathbf{X}_{k|k-1} = \mathbf{U} \mathbf{\Lambda} \mathbf{U}^{\top}$ を成分で書くと,

$$\begin{aligned} \mathbf{U} \mathbf{\Lambda} \mathbf{U}^{\top} &= \begin{pmatrix} \mathbf{u}^{(1)} & \dots & \mathbf{u}^{(N)} \end{pmatrix} \begin{pmatrix} \lambda_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \lambda_{N-1} \\ & & & 0 \end{pmatrix} \begin{pmatrix} \mathbf{u}^{(1)T} \\ \vdots \\ \mathbf{u}^{(N)T} \end{pmatrix} \quad (35) \\ &= \sum_{i=1}^{N-1} \lambda_i \mathbf{u}^{(i)} \mathbf{u}^{(i)T}. \end{aligned}$$

変換行列 \mathbf{T}_k の計算

$\mathbf{T}_k = \mathbf{U}(\mathbf{I} + \mathbf{\Lambda})^{-\frac{1}{2}} \mathbf{U}^\top$ を成分で書くと

$$\mathbf{T}_k = \mathbf{U}(\mathbf{I} + \mathbf{\Lambda})^{-\frac{1}{2}} \mathbf{U}^\top = \sum_{i=1}^N \frac{1}{\sqrt{1 + \lambda_i}} \mathbf{u}^{(i)} \mathbf{u}^{(i)T}. \quad (36)$$

$\mathbf{u}^{(N)} = \frac{1}{\sqrt{N}} \mathbf{1}$, $\lambda_N = 0$ を使うと,

$$\begin{aligned} \mathbf{T}_k \mathbf{1} &= \sqrt{N} \mathbf{U}(\mathbf{I} + \mathbf{\Lambda})^{-\frac{1}{2}} \mathbf{U}^\top \mathbf{u}^{(N)} = \sqrt{N} \left(\sum_{i=1}^N \frac{1}{\sqrt{1 + \lambda_i}} \mathbf{u}^{(i)} \mathbf{u}^{(i)T} \right) \mathbf{u}^{(N)} \\ &= \sqrt{N} \left(\frac{1}{\sqrt{1 + \lambda_N}} \mathbf{u}^{(N)} \right) = \sqrt{N} \mathbf{u}^{(N)} = \mathbf{1}. \end{aligned} \quad (37)$$

したがって、 $\mathbf{X}_{k|k-1} \mathbf{1} = \mathbf{0}$ が成り立っていれば

$$\mathbf{X}_{k|k} \mathbf{1} = \mathbf{X}_{k|k-1} \mathbf{T}_k \mathbf{1} = \mathbf{X}_{k|k-1} \mathbf{1} = \mathbf{0}. \quad (38)$$

カルマンゲイン

平均ベクトルの更新

$$\bar{\mathbf{x}}_{k|k} = \bar{\mathbf{x}}_{k|k-1} + \mathbf{K}_k (\mathbf{y}_k - \mathbf{H}_k \mathbf{x}_{k|k-1})$$

に用いるカルマンゲインは、以下のように変形しておくと便利である:

$$\begin{aligned}
 \mathbf{K}_k &= \mathbf{P}_{k|k-1} \mathbf{H}_k^\top (\mathbf{H}_k \mathbf{P}_{k|k-1} \mathbf{H}_k^\top + \mathbf{R}_k)^{-1} \\
 &= \mathbf{X}_{k|k-1} \mathbf{X}_{k|k-1}^\top \mathbf{H}_k^\top (\mathbf{H}_k \mathbf{X}_{k|k-1} \mathbf{X}_{k|k-1}^\top \mathbf{H}_k^\top + \mathbf{R}_k)^{-1} \\
 &= \mathbf{X}_{k|k-1} \left(\mathbf{I} + \mathbf{X}_{k|k-1}^\top \mathbf{H}_k^\top \mathbf{R}_k^{-1} \mathbf{H}_k \mathbf{X}_{k|k-1} \right)^{-1} \mathbf{X}_{k|k-1}^\top \mathbf{H}_k^\top \mathbf{R}_k^{-1} \\
 &= \mathbf{X}_{k|k-1} (\mathbf{I} + \mathbf{U} \mathbf{\Lambda} \mathbf{U}^\top)^{-1} \mathbf{X}_{k|k-1}^\top \mathbf{H}_k^\top \mathbf{R}_k^{-1} \\
 &= \mathbf{X}_{k|k-1} [\mathbf{U} (\mathbf{I} + \mathbf{\Lambda}) \mathbf{U}^\top]^{-1} \mathbf{X}_{k|k-1}^\top \mathbf{H}_k^\top \mathbf{R}_k^{-1} \\
 &= \mathbf{X}_{k|k-1} \mathbf{U} (\mathbf{I} + \mathbf{\Lambda})^{-1} \mathbf{U}^\top \mathbf{X}_{k|k-1}^\top \mathbf{H}_k^\top \mathbf{R}_k^{-1}.
 \end{aligned} \tag{39}$$

つまり、 \mathbf{T}_k を求めるのに使った固有値分解 $\mathbf{X}_{k|k-1}^\top \mathbf{H}_k^\top \mathbf{R}_k^{-1} \mathbf{H}_k \mathbf{X}_{k|k-1} = \mathbf{U} \mathbf{\Lambda} \mathbf{U}^\top$ がうまく活用できる。

アルゴリズム

- 1 各アンサンブルメンバーにシステムモデルを適用して予測分布を表現するアンサンブル $\{x_{k|k-1}^{(i)}\}$ を求める:

$$x_{k|k-1}^{(i)} = f_k(x_{k-1|k-1}^{(i)}) + v_k^{(i)}.$$

- 2 アンサンブル $\{x_{k|k-1}^{(i)}\}$ の平均ベクトル $\bar{x}_{k|k-1}$, 分散共分散行列の平方根 $X_{k|k-1}$ を求める.

- 3 固有値分解 $X_{k|k-1}^T H_k^T R_k^{-1} H_k X_{k|k-1} = U \Lambda U^T$.

- 4 カルマンゲイン K_k , 変換行列 T_k の計算:

$$K_k = X_{k|k-1} U (I + \Lambda)^{-1} U^T X_{k|k-1}^T H_k^T R_k^{-1}, \quad T_k = U (I + \Lambda)^{-\frac{1}{2}} U^T.$$

- 5 フィルタ分布の平均ベクトル $\bar{x}_{k|k}$, 分散共分散行列の平方根 $X_{k|k}$ を求める:

$$\bar{x}_{k|k} = \bar{x}_{k|k-1} + K_k (y_k - H_k \bar{x}_{k|k-1}), \quad X_{k|k} = X_{k|k-1} T_k.$$

- 6 求めた $\bar{x}_{k|k}$, $X_{k|k}$ $\left(= \left(\delta x_{k|k}^{(1)} \cdots \delta x_{k|k}^{(N)} \right) / \sqrt{N} \right)$ から, $x_{k|k}^{(i)} = \bar{x}_{k|k} + \delta x_{k|k}^{(i)}$ により, フィルタ分布を表現するアンサンブル $\{x_{k|k}^{(i)}\}$ を生成する.

実装上の利点

- アンサンブルメンバー数 N は、通常 x_k や y_k の次元よりずっと小さい。
- ETKF では、逆行列の計算を N 次元の空間で行うため、 $N \ll \dim x_k$ かつ $N \ll \dim y_k$ の場合には、かなり効率的である。

$$V_{k|k} \approx X_{k|k} X_{k|k}^T = X_{k|k-1} T_k T_k^T X_{k|k-1}^T$$

アンサンブル変換カルマンフィルタの性質

- アンサンブル変換カルマンフィルタ (ETKF) でフィルタ分布の平均ベクトル $\bar{x}_{k|k}$ は,

$$\begin{aligned}\bar{x}_{k|k} &= \bar{x}_{k|k-1} + \hat{\mathbf{K}}_k (\mathbf{y}_k - \mathbf{H}_k \bar{x}_{k|k-1}), \\ \hat{\mathbf{K}}_k &= \hat{\mathbf{X}}_{k|k-1} \mathbf{U} (\mathbf{I} + \mathbf{\Lambda})^{-1} \mathbf{U}^\top \hat{\mathbf{X}}_{k|k-1}^\top \mathbf{H}_k^\top \mathbf{R}_k^{-1}\end{aligned}$$

となり, 元のアンサンブル $\{x_{k|k-1}^{(i)}\}_{i=1}^N$ の線型結合で表せることが分かる.

- $\hat{\mathbf{X}}_{k|k} = \hat{\mathbf{X}}_{k|k-1} \mathbf{T}_k$ より, $\hat{\mathbf{X}}_{k|k}$ を構成する列ベクトルも $\{x_{k|k-1}^{(i)}\}_{i=1}^N$ の線型結合になっており,

$$\begin{pmatrix} \delta x_{k|k}^{(1)} & \cdots & \delta x_{k|k}^{(N)} \end{pmatrix} = \sqrt{N} \hat{\mathbf{X}}_{k|k}$$

から, $\{\delta x_{k|k}^{(i)}\}$ はすべて $\{x_{k|k-1}^{(i)}\}_{i=1}^N$ の線型結合であることが言える.

アンサンブル変換カルマンフィルタの性質

- ETKF では、フィルタ分布を表現するアンサンブルのメンバーを

$$\mathbf{x}_{k|k}^{(i)} = \bar{\mathbf{x}}_{k|k} + \delta \mathbf{x}_{k|k}^{(i)}$$

のように生成する．

- $\bar{\mathbf{x}}_{k|k}$, $\{\delta \mathbf{x}_{k|k}^{(i)}\}$ がすべて $\{\mathbf{x}_{k|k-1}^{(i)}\}_{i=1}^N$ の線型結合であることから, $\{\mathbf{x}_{k|k}^{(i)}\}$ もすべて $\{\mathbf{x}_{k|k-1}^{(i)}\}_{i=1}^N$ の線型結合となる．
- アンサンブルカルマンフィルタのところでも述べたように, このような性質は, 実装上は有利である．

初期アンサンブル

なお、上のアルゴリズムでは、初期値 x_0 の分布に関して、まず確率分布 $p(x_0)$ を与えた上で、アンサンブルメンバー $\{x^{(i)}\}_{i=1}^N$ を構成するという書き方にしている。

ただ、実用上は与えられた分布 $p(x_0)$ に応じてアンサンブルメンバーを用意するというよりも、先に何らかの方法で N 個のアンサンブルメンバーを用意し、その標本平均、標本分散共分散行列で $\bar{x}_{0|0}$, P を与えることが多い。例えば、

- シミュレーションモデルのパラメータ数個に摂動を与えて一定時間シミュレーションを実行した結果を N 個用意してアンサンブルを構成する、
- シミュレーションを長時間実行し、異なる N 個の時間ステップでの状態を保存して、アンサンブルを構成する、

などの方法が考えられる。

システムノイズについて

システムモデル $x_k = f_k(x_{k-1}) + v_k$ において、システムノイズを $v_k = 0$ とおいてしまうと、 N 個のアンサンブルメンバーが時間ステップごとに似た値を取るようになり、やがてほとんど同じ値になってしまうという現象が起きる。

- システムノイズ 0 の場合、フィルタリングの操作で分散共分散行列が縮小するのに対応して、アンサンブルメンバー間の距離が小さくなっていく。
- アンサンブルメンバー間の距離が小さくなり、すべてほぼ同じ値になってしまった場合、アンサンブルの分散共分散行列 $P_{k|k-1}$ は零行列に近くなる。
- このときカルマンゲインも零行列になり、観測の情報が推定に反映されなくなる。

システムノイズについて

- シミュレーションモデル f_k がシステムの時間発展を完璧に再現できない限り、推定される x_k の値は、現実の値から次第にずれて行く。
- 仮にシミュレーションモデル f_k が完璧だったとしても、有限のアンサンブルメンバーで表現される予測分布は近似誤差を含んでいる。この近似誤差もシステムノイズを 0 とした場合に推定が現実からずれる原因となり得る。

アンサンブルの分散共分散行列が縮小して推定が悪化するのを防ぐためには、システムノイズを考慮する必要がある。

但し、システムノイズを考慮した予測分布をうまく表現するようなアンサンブルを構成するのは容易ではない。

システムノイズの扱い

システムノイズの寄与 \mathbf{Q} を考慮するために、よく使われる方法:

- モンテカルロ近似のときと同様、 $p(\mathbf{v}_k) = \mathcal{N}(\mathbf{0}, \mathbf{Q})$ からのランダムサンプル $\{\mathbf{v}_k^{(i)}\}$ を抽出し、

$$\mathbf{x}_{k|k-1}^{(i)} = \mathbf{f}_k(\mathbf{x}_{k-1|k-1}^{(i)}) + \mathbf{v}_k^{(i)} \quad (40)$$

のように足す. (N が少ない場合、乱数による誤差の影響が問題となる.)

- 共分散膨張. 一旦システムノイズなしで時間発展を計算した上で、各アンサンブルメンバーを平均から離れる方向に動かす. 具体的には、フィルタアンサンブルの各メンバー $\mathbf{x}_{k-1|k-1}^{(i)}$ に対して、

$$\mathbf{x}_{k|k-1}^{\dagger(i)} = \mathbf{f}_k(\mathbf{x}_{k|k-1}^{(i)}) \quad (41)$$

をシミュレーションモデルで求め、その平均 $\bar{\mathbf{x}}_{k|k-1}$ を計算した上で

$$\mathbf{x}_{k|k-1}^{(i)} = \bar{\mathbf{x}}_{k|k-1} + (1 + \varepsilon) \left(\mathbf{x}_{k|k-1}^{\dagger(i)} - \bar{\mathbf{x}}_{k|k-1} \right) \quad (42)$$

のように動かす. ε は正の実数で $0.05 \sim 0.2$ 程度の値に設定することが多い. ($\mathbf{Q}_k = [(1 + \varepsilon)^2 - 1] \mathbf{P}_{k|k-1}^{\dagger}$ と仮定したことに相当する.)

Relaxation to prior

フィルタリングの操作を適用した後で、アンサンブルの拡がり具合をフィルタリング前の予測分布に近づける方法 (relaxation to prior) も提案されている。

- まず、システムノイズなしで予測アンサンブルを計算し、そのままアンサンブル変換カルマンフィルタを適用する。
- 次に、このとき得られる予測分布、フィルタ分布の分散共分散行列の平方根 $\mathbf{X}_{k|k-1}$, $\mathbf{X}_{k|k}$ を用いて、

$$\mathbf{X}_{k|k}^{\ddagger} = \alpha \mathbf{X}_{k|k-1} + (1 - \alpha) \mathbf{X}_{k|k} \quad (43)$$

のように $m \times N$ 行列 $\mathbf{X}_{k|k}^{\ddagger}$ を求める。

- 次の予測ステップでは、 $\mathbf{X}_{k|k}$ でなく $\mathbf{X}_{k|k}^{\ddagger}$ を使って生成したアンサンブルを用いる。すなわち、

$$\begin{pmatrix} \delta \mathbf{x}_{k|k}^{\ddagger(1)} & \cdots & \delta \mathbf{x}_{k|k}^{\ddagger(N)} \end{pmatrix} = \sqrt{N} \mathbf{X}_{k|k}^{\ddagger} \quad (44)$$

のように分解した後、

$$\mathbf{x}_{k|k}^{\ddagger(i)} = \bar{\mathbf{x}}_{k|k} + \delta \mathbf{x}_{k|k}^{\ddagger(i)} \quad (45)$$

カルマンフィルタの予測式の近似

カルマンフィルタにおける予測分布の分散共分散行列の式をうまく近似するようにアンサンブルを構成するという方法も考えられる．つまり，システムノイズなしでシミュレーションを実施して $\mathbf{x}_{k|k-1}^{\dagger(i)} = \mathbf{f}_k(\mathbf{x}_{k|k-1}^{(i)})$ を求め，その標本平均を $\bar{\mathbf{x}}_{k|k-1}$ とし，

$$\mathbf{X}_{k|k-1}^{\dagger} = \frac{1}{\sqrt{N}} \begin{pmatrix} \mathbf{x}_{k|k-1}^{\dagger(1)} - \bar{\mathbf{x}}_{k|k-1} & \cdots & \mathbf{x}_{k|k-1}^{\dagger(N)} - \bar{\mathbf{x}}_{k|k-1} \end{pmatrix}, \quad (46)$$

と定義する．予測分布の分散共分散行列の平方根 $\mathbf{X}_{k|k-1}$ は，

$$\mathbf{X}_{k|k-1} \mathbf{X}_{k|k-1}^{\top} \simeq \hat{\mathbf{P}}_{k|k-1} + \mathbf{Q} = \mathbf{X}_{k|k-1}^{\dagger} \mathbf{X}_{k|k-1}^{\dagger\top} + \mathbf{Q} \quad (47)$$

を満たすように求めればよい．

但し，式 (47) が $=$ ではなく \simeq で結ばれているのは， \mathbf{Q} の階数が N 以上の時，式 (47) を完全に満足する階数 $N - 1$ の行列 $\mathbf{X}_{k|k-1}$ が存在し得ないからである．

カルマンフィルタの予測式の近似

ここで,

$$\mathbf{Q} = \mathbf{Q}^{-\frac{1}{2}} \mathbf{Q}^{-\frac{1}{2}} \quad (48)$$

を満たす対称行列 $\mathbf{Q}^{-\frac{1}{2}}$ を定義する. $\mathbf{Q}^{-\frac{1}{2}}$ は, \mathbf{Q} の固有値分解から容易に得られる. 今の式

$$\mathbf{X}_{k|k-1} \mathbf{X}_{k|k-1}^T \simeq \mathbf{X}_{k|k-1}^\dagger \mathbf{X}_{k|k-1}^{\dagger T} + \mathbf{Q}$$

の左右から $\mathbf{Q}^{-\frac{1}{2}}$ を掛けて,

$$\mathbf{Q}^{-\frac{1}{2}} \mathbf{X}_{k|k-1} \mathbf{X}_{k|k-1}^T \mathbf{Q}^{-\frac{1}{2}} \simeq \mathbf{Q}^{-\frac{1}{2}} \mathbf{X}_{k|k-1}^\dagger \mathbf{X}_{k|k-1}^{\dagger T} \mathbf{Q}^{-\frac{1}{2}} + \mathbf{I}_m. \quad (49)$$

カルマンフィルタの予測式の近似

ここで、以下の特異値分解を考える

$$\mathbf{Q}^{-\frac{1}{2}} \mathbf{X}_{k|k-1}^\dagger = \check{\mathbf{U}}_x \check{\mathbf{\Gamma}} \check{\mathbf{U}}_z^\top. \quad (50)$$

$\check{\mathbf{U}}_x$, $\check{\mathbf{U}}_z$ は特異値 0 に対応する成分を除いたもので、それぞれ $m \times (N-1)$ 行列, $N \times (N-1)$ 行列になっており、したがって、 $\check{\mathbf{\Gamma}}$ は $(N-1)$ 次の対角行列である。式 (50) を式 (49) の右辺に代入すると、

$$\mathbf{Q}^{-\frac{1}{2}} \mathbf{X}_{k|k-1} \mathbf{X}_{k|k-1}^\top \mathbf{Q}^{-\frac{1}{2}} \simeq \check{\mathbf{U}}_x \check{\mathbf{\Gamma}}^2 \check{\mathbf{U}}_x^\top + \mathbf{I}_m. \quad (51)$$

そこで、右辺第 2 項を $\mathbf{I}_m \simeq \check{\mathbf{U}}_x \check{\mathbf{U}}_x^\top$ と近似し、

$$\mathbf{Q}^{-\frac{1}{2}} \mathbf{X}_{k|k-1} \mathbf{X}_{k|k-1}^\top \mathbf{Q}^{-\frac{1}{2}} = \check{\mathbf{U}}_x \check{\mathbf{\Gamma}}^2 \check{\mathbf{U}}_x^\top + \check{\mathbf{U}}_x \check{\mathbf{U}}_x^\top = \check{\mathbf{U}}_x (\check{\mathbf{\Gamma}}^2 + \mathbf{I}_{N-1}) \check{\mathbf{U}}_x^\top \quad (52)$$

が成り立つように $\mathbf{X}_{k|k-1}$ を求めることにする。式 (52) の両辺の左右から $\mathbf{Q}^{\frac{1}{2}}$ を掛けると、

$$\mathbf{X}_{k|k-1} \mathbf{X}_{k|k-1}^\top = \mathbf{Q}^{\frac{1}{2}} \check{\mathbf{U}}_x (\check{\mathbf{\Gamma}}^2 + \mathbf{I}_{N-1}) \check{\mathbf{U}}_x^\top \mathbf{Q}^{\frac{1}{2}}. \quad (53)$$

カルマンフィルタの予測式の近似

式 (50) から $\mathbf{Q}^{\frac{1}{2}} \check{\mathbf{U}}_x = \mathbf{X}_{k|k-1}^\dagger \check{\mathbf{U}}_z \check{\mathbf{\Gamma}}^{-1}$ が成り立つので,

$$\mathbf{X}_{k|k-1} \mathbf{X}_{k|k-1}^\top = \mathbf{X}_{k|k-1}^\dagger \check{\mathbf{U}}_z \check{\mathbf{\Gamma}}^{-1} (\check{\mathbf{\Gamma}}^2 + \mathbf{I}_{N-1}) \check{\mathbf{\Gamma}}^{-1} \check{\mathbf{U}}_z^\top \mathbf{X}_{k|k-1}^\dagger. \quad (54)$$

これを満たし、かつ $\mathbf{X}_{k|k-1} \mathbf{1}_N = \mathbf{0}$ を満たすには,

$$\mathbf{X}_{k|k-1} = \mathbf{X}_{k|k-1}^\dagger \check{\mathbf{U}}_z \check{\mathbf{\Gamma}}^{-1} (\check{\mathbf{\Gamma}}^2 + \mathbf{I}_{N-1}) \check{\mathbf{U}}_z^\top \quad (55)$$

とおけばよい. ($\mathbf{X}_{k|k-1}^\dagger \mathbf{1}_N = \mathbf{0}$ なので, 行列 $\mathbf{Q}^{-\frac{1}{2}} \mathbf{X}_{k|k-1}^\dagger$ の特異値分解で 0 に対応する成分を除いて得られる $\check{\mathbf{U}}_z$ は, $\check{\mathbf{U}}_z^\top \mathbf{1}_N = \mathbf{0}$ を満たす.)

$\mathbf{X}_{k|k-1}$ が得られたら, これを予測アンサンブルの分散共分散行列の平方根として使い, アンサンブル変換カルマンフィルタのアルゴリズムを適用すればよい.

カルマンフィルタの予測式の近似

なお、 $\mathbf{X}_{k|k-1}$ を得るのに特異値分解

$$\mathbf{Q}^{-\frac{1}{2}} \mathbf{X}_{k|k-1}^\dagger = \check{\mathbf{U}}_x \check{\mathbf{\Gamma}} \check{\mathbf{U}}_z^\top$$

を行ったが、

$$\mathbf{X}_{k|k-1} = \mathbf{X}_{k|k-1}^\dagger \check{\mathbf{U}}_z \check{\mathbf{\Gamma}}^{-1} (\check{\mathbf{\Gamma}}^2 + \mathbf{I}_{N-1}) \check{\mathbf{U}}_z^\top$$

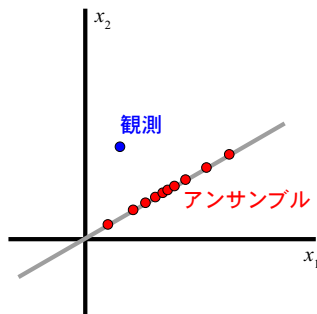
を得るのに必要な行列 $\check{\mathbf{U}}_z$, $\check{\mathbf{\Gamma}}$ は、以下の固有値分解を利用して求めることもできる:

$$\mathbf{X}_{k|k-1}^\top \mathbf{Q}^{-1} \mathbf{X}_{k|k-1} = \mathbf{U}_z \mathbf{\Gamma}^2 \mathbf{U}_z^\top. \quad (56)$$

このとき、 N 次対角行列 $\mathbf{\Gamma}$ から対角要素が 0 の部分を除いた $N-1$ 次対角行列が $\check{\mathbf{\Gamma}}$, \mathbf{U}_z から 0 固有値に対応する固有ベクトルを除いた行列が $\check{\mathbf{U}}_z$ になる。

少数粒子による推定の制約

- アンサンブルカルマンフィルタや ETKF で得られるアンサンブルメンバー $\{x_{k|k}^{(i)}\}$ はすべて $\{x_{k|k-1}^{(i)}\}_{i=1}^N$ の線型結合で得られる。
- そのため、フィルタ分布は $\{x_{k|k-1}^{(i)}\}_{i=1}^N$ で張られる $N - 1$ 次元の部分空間に制限される。
- このことは、 $\hat{P}_{k|k-1} = \hat{X}_{k|k-1} \hat{X}_{k|k-1}^T$ が $N - 1$ 次元空間に制限されることから分かる。
- N がシステムの自由度と同程度あればよいが、 N をそこまで大きくできない場合も少なくない。その場合、以下のような問題が起こり得る：
 - データが十分にあっても、推定結果をデータに合わせきれない、
 - 離れた場所に偽の相関が現れることもある。



局所化

- 地球科学分野のデータ同化では、このような問題への対処策として、シミュレーションの各グリッド点の近傍で得られる観測データのみを考慮するような処理が行われる。このような処理を局所化 (localization) と呼ぶ。
- 局所化には、大きく分けて 2 種類の手法がある。
 - 1 アンサンブルの標本分散共分散行列で表現される予測分布の分散共分散行列から、離れた場所同士の (偽の) 相関を取り除いた上でフィルタ分布を計算する B 局所化。
 - 2 シミュレーションのグリッド点ごとに近傍の観測データだけを参照してフィルタ推定値を求める R 局所化。

なお、アンサンブル変換カルマンフィルタに適用できるのは、R 局所化のみとなる。

B 局所化

B 局所化は、アンサンブルカルマンフィルタで得られる予測アンサンブルの標本分散共分散行列 $\hat{\mathbf{P}}_{k|k-1}$ を以下のように調整する:

$$\hat{\mathbf{P}}_{k|k-1}^{\ell} = \mathbf{L} \circ \hat{\mathbf{P}}_{k|k-1}. \quad (57)$$

\mathbf{L} は空間的な距離に応じて重み付けをする局所化行列であり、 \circ は行列の要素ごとの積 (シュア積 Schur product, あるいはアダマール積 Hadamard product と呼ぶ) を表し,

$$\mathbf{L} \circ \mathbf{A} = \begin{pmatrix} L_{11}A_{11} & L_{12}A_{12} & \cdots & L_{1m}A_{1m} \\ L_{21}A_{21} & L_{22}A_{22} & & L_{2m}A_{2m} \\ \vdots & & \ddots & \vdots \\ L_{m1}A_{m1} & L_{m2}A_{m2} & \cdots & L_{mm}A_{mm} \end{pmatrix} \quad (58)$$

である.

かつて予測分布の分散共分散行列 $\mathbf{P}_{k|k-1}$ をデータ同化の分野では \mathbf{B} と表記することが多かったため、式 (57) の操作を B 局所化と呼んでいる.

B 局所化

B 局所化において、局所化行列 \mathbf{L} の各要素は、空間上の 2 点 $\mathbf{r}_{j_1}, \mathbf{r}_{j_2}$ が与えられたときに、2 点間の距離に依存する関数で与えることが多い。よく用いられるのは、Gaspari–Cohn 関数と呼ばれるもので、以下で定義される:

$$\rho(r) = 1 - \frac{5}{3}r^2 + \frac{5}{8}r^3 + \frac{1}{2}r^4 - \frac{1}{4}r^5 \quad (0 \leq r < 1), \quad (59a)$$

$$\rho(r) = 4 - 5r + \frac{5}{3}r^2 + \frac{5}{8}r^3 - \frac{1}{2}r^4 + \frac{1}{12}r^5 - \frac{2}{3r} \quad (1 \leq r < 2), \quad (59b)$$

$$\rho(r) = 0 \quad (r \geq 2), \quad (59c)$$

但し、 $r = \|\mathbf{r}_{j_1} - \mathbf{r}_{j_2}\|/\lambda_B$.

この関数は、Gauss 関数を近似し、かつ $r \geq 2$ で 0 となるように設計されている。 r は \mathbf{r}_{j_1} と \mathbf{r}_{j_2} との距離を λ_B でスケーリングしたもの。 λ_B のことを局所化半径と言う。

B 局所化

局所化行列 \mathbf{L} を Gaspari–Cohn 関数のような 2 点間の距離に依存する関数で与える場合、状態変数 x_k の j 番目の要素 $x_{k,j}$ に対応するシミュレーションのグリッド点の位置を r_j として、

$$\mathbf{L} = \begin{pmatrix} \rho\left(\frac{\|r_1 - r_1\|}{\lambda_B}\right) & \cdots & \rho\left(\frac{\|r_1 - r_m\|}{\lambda_B}\right) \\ \vdots & \ddots & \vdots \\ \rho\left(\frac{\|r_m - r_1\|}{\lambda_B}\right) & \cdots & \rho\left(\frac{\|r_m - r_m\|}{\lambda_B}\right) \end{pmatrix} \quad (60)$$

とすればよい。

ρ が Gaspari–Cohn 関数ならば、距離が $2\lambda_B$ 以上離れた 2 点間の相関は 0 となるから、各グリッド点の推定に離れた場所の観測データが影響することはなくなる。

B 局所化

$\hat{\mathbf{P}}_{k|k-1}^\ell = \mathbf{L} \circ \hat{\mathbf{P}}_{k|k-1}$ とおくというだけでは、 $m \times m$ の大きなサイズの行列になるため、 m が大きい場合に計算機で扱うのが難しくなる。

しかし、観測 y_k の各要素が、状態変数 x_k の要素のどれかを直接観測したものになっていれば、 y_k の各要素は対応する x_k の要素のグリッド点と対応付けられる。そこで、 y_k の j 番目の要素 $y_{k,j}$ に対応するグリッド点の位置を \mathbf{r}'_j として、

$$\mathbf{L}_{yy} = \begin{pmatrix} \rho\left(\frac{\|\mathbf{r}'_1 - \mathbf{r}'_1\|}{\lambda_B}\right) & \cdots & \rho\left(\frac{\|\mathbf{r}'_1 - \mathbf{r}'_n\|}{\lambda_B}\right) \\ \vdots & \ddots & \vdots \\ \rho\left(\frac{\|\mathbf{r}'_n - \mathbf{r}'_1\|}{\lambda_B}\right) & \cdots & \rho\left(\frac{\|\mathbf{r}'_n - \mathbf{r}'_n\|}{\lambda_B}\right) \end{pmatrix} \quad (61)$$

$$\mathbf{L}_{xy} = \begin{pmatrix} \rho\left(\frac{\|\mathbf{r}_1 - \mathbf{r}'_1\|}{\lambda_B}\right) & \cdots & \rho\left(\frac{\|\mathbf{r}_1 - \mathbf{r}'_n\|}{\lambda_B}\right) \\ \vdots & \ddots & \vdots \\ \rho\left(\frac{\|\mathbf{r}_m - \mathbf{r}'_1\|}{\lambda_B}\right) & \cdots & \rho\left(\frac{\|\mathbf{r}_m - \mathbf{r}'_n\|}{\lambda_B}\right) \end{pmatrix} \quad (62)$$

とおく。

B 局所化

行列 \mathbf{L}_{yy} , \mathbf{L}_{xy} を導入すると,

$$\mathbf{H}_k(\mathbf{L} \circ \mathbf{P}_{k|k-1})\mathbf{H}_k^\top = \mathbf{L}_{yy} \circ (\mathbf{H}_k\mathbf{P}_{k|k-1}\mathbf{H}_k^\top) = \mathbf{L}_{yy} \circ (\hat{\mathbf{Y}}_{k|k-1}\hat{\mathbf{Y}}_{k|k-1}^\top), \quad (63)$$

$$(\mathbf{L} \circ \mathbf{P}_{k|k-1})\mathbf{H}_k^\top = \mathbf{L}_{xy} \circ (\mathbf{P}_{k|k-1}\mathbf{H}_k^\top) = \mathbf{L}_{xy} \circ (\hat{\mathbf{X}}_{k|k-1}\hat{\mathbf{Y}}_{k|k-1}^\top) \quad (64)$$

が成り立つ.

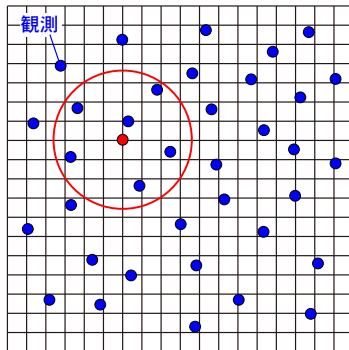
これを用いると, 局所化を適用したカルマンゲイン

$$\hat{\mathbf{K}}_k = \mathbf{L}_{xy} \circ (\hat{\mathbf{X}}_{k|k-1}\hat{\mathbf{Y}}_{k|k-1}^\top) \left(\mathbf{L}_{yy} \circ (\hat{\mathbf{Y}}_{k|k-1}\hat{\mathbf{Y}}_{k|k-1}^\top) + \mathbf{R}_k \right)^{-1} \quad (65)$$

が得られる. あとは通常のアンサンブルカルマンフィルタのアルゴリズムにしたがって, 各アンサンブルメンバーを修正すればよい.

R 局所化

- もう一つの方法として、シミュレーションのグリッド点ごとに、近くのデータだけを参照して推定を行うという方法がある (Ott et al., 2004).
- このような方法を R 局所化と言う。
- 典型的には、該当するグリッド点から一定の距離 λ 以下の点で得られた観測データを参照する。このときの λ は、R 局所化における局所化半径となる。



R 局所化

R 局所化を適用して x_k の j 番目の要素 $x_{k,j}$ を推定するには、以下を定義する:

- $\bar{x}_{k|k-1}$ の j 番目の要素 $\bar{x}_{k|k-1,j}$,
- 予測アンサンブルのメンバー $x_{k|k-1}^{(i)}$ の j 番目の要素を $x_{k|k-1,j}^{(i)}$, j 番目の要素の平均を $\bar{x}_{k|k-1,j}$ として得られる横ベクトル:

$$\hat{e}_{k|k-1,j}^\top = \frac{1}{\sqrt{N-1}} \left(x_{k|k-1,j}^{(1)} - \bar{x}_{k|k-1,j} \cdots x_{k|k-1,j}^{(N)} - \bar{x}_{k|k-1,j} \right),$$

- $x_{k,j}$ の割り当てられているグリッド点の推定に用いる観測データをまとめたベクトル $y_{k,j}$,
- H_k から $y_{k,j}$ に関する行だけを取り出した $\check{H}_{k,j}$.
- R_k から $y_{k,j}$ に関する行だけを取り出した $\check{R}_{k,j}$.

R 局所化

R 局所化では, $N < \dim \mathbf{y}_k$ の場合のカルマンゲイン計算式

$$\mathbf{K}_k = \hat{\mathbf{X}}_{k|k-1} (\mathbf{I} + \hat{\mathbf{X}}_{k|k-1}^T \mathbf{H}_k^T \mathbf{R}_k^{-1} \mathbf{H}_k \hat{\mathbf{X}}_{k|k-1})^{-1} \hat{\mathbf{X}}_{k|k-1}^T \mathbf{H}_k^T \mathbf{R}_k^{-1}. \quad (66)$$

に基づき, 局所化されたカルマンゲイン

$$\mathbf{k}_{k,j}^T = \hat{\mathbf{e}}_{k|k-1,j}^T (\mathbf{I} + \hat{\mathbf{X}}_{k|k-1}^T \check{\mathbf{H}}_{k,j}^T \check{\mathbf{R}}_{k,j}^{-1} \check{\mathbf{H}}_{k,j} \hat{\mathbf{X}}_{k|k-1})^{-1} \hat{\mathbf{X}}_{k|k-1}^T \check{\mathbf{H}}_{k,j}^T \check{\mathbf{R}}_{k,j}^{-1} \quad (67)$$

を求め, $x_{k|k}^{(i)}$ の j 番目の要素 $x_{k|k,j}^{(i)}$ を以下のように得る.

$$x_{k|k,j}^{(i)} = x_{k|k-1,j}^{(i)} + \mathbf{k}_{k,j}^T (\mathbf{y}_{k,j} - \check{\mathbf{H}}_{k,j} \bar{\mathbf{x}}_{k|k-1} + \mathbf{w}_{k,j}^{(i)}). \quad (68)$$

但し, $\mathbf{w}_{k,j}^{(i)}$ は $\mathcal{N}(\mathbf{0}, \check{\mathbf{R}}_{k,j})$ に従う乱数.

局所アンサンブル変換カルマンフィルタ

R 局所化を適用したアンサンブル変換カルマンフィルタが局所アンサンブル変換カルマンフィルタ (local ensemble transform Kalman filter; LETKF) (Hunt et al., 2007)。

局所アンサンブル変換カルマンフィルタでは、 x_k の j 番目の要素 $x_{k,j}$ を推定するために以下の固有値分解を行う：

$$\mathbf{Y}_{k|k-1,j}^T \check{\mathbf{R}}_{k,j}^{-1} \mathbf{Y}_{k|k-1,j} = \mathbf{U}_j \mathbf{\Lambda}_j \mathbf{U}_j^T. \quad (69)$$

予測分布の平均 $\bar{x}_{k|k-1}$ の j 番目の要素を $\bar{x}_{k|k-1,j}$ ，予測アンサンブルのメンバー $x_{k|k-1}^{(i)}$ の j 番目の要素を $x_{k|k-1,j}^{(i)}$ とし，

$$\mathbf{e}_{k|k-1,j}^T = \frac{1}{\sqrt{N}} \left(x_{k|k-1,j}^{(1)} - \bar{x}_{k|k-1,j} \quad \cdots \quad x_{k|k-1,j}^{(N)} - \bar{x}_{k|k-1,j} \right) \quad (70)$$

と定義すると，フィルタ分布の平均 $\bar{x}_{k|k}$ の j 番目の要素 $\bar{x}_{k|k,j}$ は，

$$\bar{x}_{k|k,j} = \bar{x}_{k|k-1,j} + \mathbf{e}_{k|k-1,j}^T \mathbf{U}_j (\mathbf{I} + \mathbf{\Lambda}_j)^{-1} \mathbf{U}_j^T \mathbf{X}_{k|k-1}^T \check{\mathbf{H}}_{k,j} \check{\mathbf{R}}_{k,j}^{-1} (\mathbf{y}_{k,j} - \check{\mathbf{H}}_{k,j} \bar{\mathbf{x}}_{k|k-1}). \quad (71)$$

局所アンサンブル変換カルマンフィルタ

フィルタ分布を表現するアンサンブルを得るには、まず、各 j に対して

$$\mathbf{T}_{k,j} = \mathbf{U}_j(\mathbf{I} + \mathbf{\Lambda}_j)^{-\frac{1}{2}} \mathbf{U}_j^{\top}. \quad (72)$$

のように $N \times N$ 行列 $\mathbf{T}_{k,j}$ を求める。さらに

$$\mathbf{e}_{k|k,j}^{\top} = \mathbf{e}_{k|k-1,j}^{\top} \mathbf{T}_{k,j} \quad (73)$$

から得られるベクトル $\mathbf{e}_{k|k,1}^{\top}, \dots, \mathbf{e}_{k|k,m}^{\top}$ を N 個まとめた行列

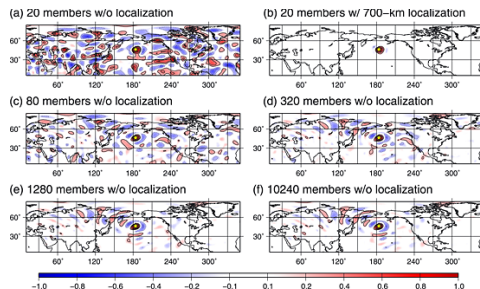
$$\mathbf{X}_{k|k} = \begin{pmatrix} \mathbf{e}_{k|k,1}^{\top} \\ \vdots \\ \mathbf{e}_{k|k,m}^{\top} \end{pmatrix} \quad (74)$$

を構成する。 $\sqrt{N} \mathbf{X}_{k|k}$ を $\begin{pmatrix} \delta \mathbf{x}_{k|k}^{(1)} & \dots & \delta \mathbf{x}_{k|k}^{(N)} \end{pmatrix} = \sqrt{N} \mathbf{X}_{k|k}$ のように列ベクトルに分解すれば、 $i = 1, \dots, N$ に対して以下を適用することでフィルタ分布を表現するアンサンブルが得られる:

$$\mathbf{x}_{k|k}^{(i)} = \bar{\mathbf{x}}_{k|k} + \delta \mathbf{x}_{k|k}^{(i)}. \quad (75)$$

局所化の効果

- N が小さく、局所化も施さない場合、離れた場所に見せかけの相関が現れる。
- 局所化を施すことで、見せかけの相関はなくなる。しかし、少し離れた場所の意味のある相関の情報も落ちてしまう。
- 意味のある相関を考慮するには、 N を十分大きく取る必要がある。



Effect of localization (Miyoshi et al., Geophys. Res. Lett., 2014)

まとめ

- アンサンブル変換カルマンフィルタは、アンサンブルカルマンフィルタのように乱数を使う必要がなく、少数の粒子で推定ができるため、計算コストを抑えることができる。
- 但し、高次元の状態変数の確率分布を少数の粒子で表現するため、データがたくさんあっても推定結果をデータに合わせきれない、離れた場所に偽の相関が現れるといった問題が起こり得る。
- こうした問題の解決策として、各グリッド点ごとに近傍のデータだけを参照して推定を行う局所化という手法が提案されている。