

# アンサンブルカルマンフィルタ

中野 慎也

「データ科学: 理論から実用へ」

8 September 2023

# 逐次データ同化 (復習)

時刻  $t_{k-1}$  までの観測  $y_1, \dots, y_{k-1}$  を使った  $x_k$  の予測  $p(x_k | y_{1:k-1})$ :

$$p(x_k | y_{1:k-1}) = \int \overset{\text{Given}}{p(x_k | x_{k-1})} p(x_{k-1} | y_{1:k-1}) dx_{k-1}.$$

時刻  $t_k$  の観測  $y_k$  の情報を追加して  $x_k$  の予測を修正 (フィルタリング):

$$p(x_k | y_{1:k}) = \frac{\overset{\text{Given}}{p(y_k | x_k)} p(x_k | y_{1:k-1})}{\int \overset{\text{Given}}{p(y_k | x_k)} p(x_k | y_{1:k-1}) dx_k}.$$

- 予測とフィルタリングの計算を繰り返すことで、各時間ステップの推定を行う。

# カルマンフィルタのアルゴリズム (復習)

- 時間ステップ  $k = 1, \dots, K$  について, 以下を実行する.

- 一期先予測:

$$\bar{\mathbf{x}}_{k|k-1} = \mathbf{F}_k \bar{\mathbf{x}}_{k-1|k-1},$$

$$\mathbf{P}_{k|k-1} = \mathbf{F}_k \mathbf{P}_{k-1|k-1} \mathbf{F}_k^\top + \mathbf{Q}_k.$$

- フィルタリング:

$$\mathbf{K}_k = \mathbf{P}_{k|k-1} \mathbf{H}_k^\top (\mathbf{H}_k \mathbf{P}_{k|k-1} \mathbf{H}_k^\top + \mathbf{R}_k)^{-1},$$

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

$$\mathbf{P}_{k|k} = (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \mathbf{P}_{k|k-1}.$$

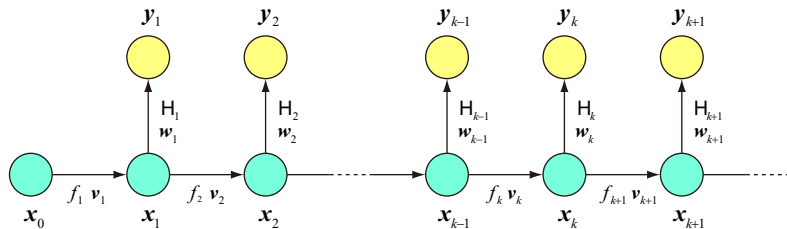
# 非線型システムの場合

実際にデータ同化で扱うのは、システムが非線型な場合が多いので、非線型システム

$$\mathbf{x}_k = \mathbf{f}_k(\mathbf{x}_{k-1}) + \mathbf{v}_k$$

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{w}_k$$

についても考える。



# 非線型システムの場合

線型システム  $x_k = F_k x_{k-1} + v_k$  では、 $x_{k-1}$  の分布  $p(x_{k-1} | y_{1:k-1})$  が  $\mathcal{N}(\bar{x}_{k-1|k-1}, P_{k-1|k-1})$  ならば、予測分布  $p(x_k | y_{1:k-1})$  は

- ガウス分布で、
- 平均ベクトルは  $F_k \bar{x}_{k-1|k-1}$ ,
- 分散共分散行列  $F_k P_{k-1|k-1} F_k^T + Q$ .

しかし、非線型システム  $x_k = f_k(x_{k-1}) + v_k$  では...

- 予測分布  $p(x_k | y_{1:k-1})$  がガウス分布になる保証はない。
- 平均ベクトルが  $f_k(x_{k-1})$  になるとは限らない。
- 分散共分散行列もどうなるかわからない。

# 拡張カルマンフィルタ

非線型システムであっても  $f_k$  の 1 次近似を考え、カルマンフィルタのアルゴリズムをそのまま適用するという手はある (拡張カルマンフィルタ)。

まず,  $f_k(x_{k-1})$  の 1 次のテイラー展開

$$f_k(x_{k-1}) \approx f_k(\bar{x}_{k-1|k-1}) + \tilde{\mathbf{F}}_{k, x_{k-1}|k-1} (x_{k-1} - \bar{x}_{k-1|k-1}) \quad (1)$$

を考える。但し,  $\tilde{\mathbf{F}}_{k, x_{k-1}|k-1}$  は  $f_k$  の  $x_{k-1} = \bar{x}_{k-1|k-1}$  におけるヤコビ行列, すなわち,

$$\tilde{\mathbf{F}}_{k, x_{k-1}|k-1} = \left. \frac{\partial f_k}{\partial \mathbf{x}_{k-1}^\top} \right|_{\mathbf{x}_{k-1} = \bar{\mathbf{x}}_{k-1|k-1}}.$$

$\|x_{k-1} - \bar{x}_{k-1|k-1}\|$  が小さければ

$$\begin{aligned} \mathbf{x}_k &= f_k(\mathbf{x}_{k-1}) + \mathbf{v}_k \\ &\approx f_k(\bar{\mathbf{x}}_{k-1|k-1}) + \tilde{\mathbf{F}}_{k, \mathbf{x}_{k-1}|k-1} (\mathbf{x}_{k-1} - \bar{\mathbf{x}}_{k-1|k-1}) + \mathbf{v}_k \end{aligned} \quad (2)$$

# 拡張カルマンフィルタ

$$\mathbf{x}_k \approx \mathbf{f}_k(\bar{\mathbf{x}}_{k-1|k-1}) + \tilde{\mathbf{F}}_{k, \mathbf{x}_{k-1|k-1}}(\mathbf{x}_{k-1} - \bar{\mathbf{x}}_{k-1|k-1}) + \mathbf{v}_k$$

のとき,

$$\begin{aligned} \bar{\mathbf{x}}_{k|k-1} &\approx \int \left[ \mathbf{f}_k(\bar{\mathbf{x}}_{k-1|k-1}) + \tilde{\mathbf{F}}_{k, \mathbf{x}_{k-1|k-1}}(\mathbf{x}_{k-1} - \bar{\mathbf{x}}_{k-1|k-1}) + \mathbf{v}_k \right] \\ &\quad \times p(\mathbf{x}_{k-1} | \mathbf{y}_{1:k-1}) p(\mathbf{v}_k) d\mathbf{x}_{k-1} d\mathbf{v}_k \\ &= \mathbf{f}_k(\bar{\mathbf{x}}_{k-1|k-1}). \end{aligned} \tag{3}$$

$$\therefore \int \mathbf{x}_{k-1} p(\mathbf{x}_{k-1} | \mathbf{y}_{1:k-1}) d\mathbf{x}_{k-1} = \bar{\mathbf{x}}_{k-1|k-1}, \quad \int \mathbf{v}_k p(\mathbf{v}_k) d\mathbf{v}_k = 0.$$

# 拡張カルマンフィルタ

$$\mathbf{x}_k \approx \mathbf{f}_k(\bar{\mathbf{x}}_{k-1|k-1}) + \tilde{\mathbf{F}}_{k, \mathbf{x}_{k-1|k-1}}(\mathbf{x}_{k-1} - \bar{\mathbf{x}}_{k-1|k-1}) + \mathbf{v}_k$$

のときの分散共分散行列は,

$$\mathbf{x}_k - \bar{\mathbf{x}}_{k|k-1} = \tilde{\mathbf{F}}_{k, \mathbf{x}_{k-1|k-1}}(\mathbf{x}_{k-1} - \bar{\mathbf{x}}_{k-1|k-1}) + \mathbf{v}_k$$

より,

$$\begin{aligned} & \mathbf{P}_{k|k-1} \\ & \approx \int \left[ \tilde{\mathbf{F}}_{k, \mathbf{x}_{k-1|k-1}}(\mathbf{x}_{k-1} - \bar{\mathbf{x}}_{k-1|k-1}) + \mathbf{v}_k \right] \\ & \quad \times \left[ \tilde{\mathbf{F}}_{k, \mathbf{x}_{k-1|k-1}}(\mathbf{x}_{k-1} - \bar{\mathbf{x}}_{k-1|k-1}) + \mathbf{v}_k \right]^\top p(\mathbf{x}_{k-1} | \mathbf{y}_{1:k-1}) p(\mathbf{v}_k) d\mathbf{x}_{k-1} d\mathbf{v}_k \\ & = \tilde{\mathbf{F}}_{k, \mathbf{x}_{k-1|k-1}} \mathbf{P}_{k-1|k-1} \tilde{\mathbf{F}}_{k, \mathbf{x}_{k-1|k-1}}^\top + \mathbf{Q}, \end{aligned} \tag{4}$$

$$\therefore \int (\mathbf{x}_{k-1} - \bar{\mathbf{x}}_{k-1|k-1}) \mathbf{v}_k^\top p(\mathbf{x}_{k-1} | \mathbf{y}_{1:k-1}) p(\mathbf{v}_k) d\mathbf{x}_{k-1} d\mathbf{v}_k = 0.$$



# 拡張カルマンフィルタ

- こうして、予測分布の平均ベクトル  $\bar{\mathbf{x}}_{k|k-1}$ ，分散共分散行列  $\mathbf{P}_{k|k-1}$  が，近似的に

$$\bar{\mathbf{x}}_{k|k-1} \approx \mathbf{f}_k(\bar{\mathbf{x}}_{k-1|k-1}), \quad (5a)$$

$$\mathbf{P}_{k|k-1} \approx \tilde{\mathbf{F}}_{k, \mathbf{x}_{k-1|k-1}} \mathbf{P}_{k-1|k-1} \tilde{\mathbf{F}}_{k, \mathbf{x}_{k-1|k-1}}^T + \mathbf{Q} \quad (5b)$$

となることが分かった．

- そこで、予測分布を平均ベクトル  $\bar{\mathbf{x}}_{k|k-1}$ ，分散共分散行列  $\mathbf{P}_{k|k-1}$  のガウス分布と考えて通常のカルマンフィルタと同じ手続きでフィルタリングを行うのが拡張カルマンフィルタ．

# 拡張カルマンフィルタのアルゴリズム (まとめ)

- 時間ステップ  $k = 1, \dots, K$  について, 以下を実行する.

- 一期先予測:

$$\begin{aligned}\bar{\mathbf{x}}_{k|k-1} &= \mathbf{f}_k(\bar{\mathbf{x}}_{k-1|k-1}), \\ \mathbf{P}_{k|k-1} &= \tilde{\mathbf{F}}_{k, \mathbf{x}_{k-1|k-1}} \mathbf{P}_{k-1|k-1} \tilde{\mathbf{F}}_{k, \mathbf{x}_{k-1|k-1}}^\top + \mathbf{Q}.\end{aligned}$$

- フィルタリング:

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

# 拡張カルマンフィルタの問題点

- $\tilde{\mathbf{F}}_{k, \mathbf{x}_{k-1|k-1}}$  をどうやって求めるかが問題になることが多い。
- さらに、状態変数  $\mathbf{x}$  が高次元の場合には、 $\tilde{\mathbf{F}}_{k, \mathbf{x}_{k-1|k-1}}$  を保持するために大きなメモリ容量も必要になる。
- 分散共分散行列  $\mathbf{P}_{k|k-1}$  を保持するのに要するメモリ容量も  $\tilde{\mathbf{F}}_{k, \mathbf{x}_{k-1|k-1}}$  と同じだけ必要である。

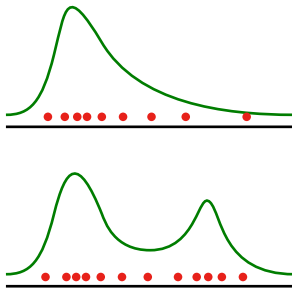
こうした問題を解決する方法として、モンテカルロ近似に基づくアンサンブルカルマンフィルタ (EnKF) がある。

# 高次元問題への適用

- データ同化で使われるシミュレーションモデルの状態変数の次元  $m$  は、1 万、100 万、1 億などかなり大規模になることが多い。
- カルマンフィルタの計算に使われる、分散共分散行列の要素数は  $m \times m$ 。1 億次元なら 1 京。何も考えずに計算するとペタバイト級のメモリが必要。
- しかし、実際のシミュレーションモデルは、高次元 (変数の数が多い) と言っても実質的な自由度はそこまで大きくないことが多い。(設定を色々変えてシミュレーションをする場合でも、変更するのは数個程度のパラメータであることが多い。)

# モンテカルロ近似

- 確率分布は、分布からランダムに抽出したサンプルでも表現できる。
- 個々のサンプルを粒子，アンサンブルメンバーとも呼ぶ。
- 粒子の集合をアンサンブルと呼ぶ。
- 粒子が密に分布していれば，確率密度が高いことを表す。
- ガウス分布に限らず，様々な分布の形状を表現できる。



# モンテカルロ近似

粒子の集合 (アンサンブル)  $\{\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(N)}\}$  による確率密度関数  $p(\mathbf{x})$  の近似は、デルタ関数の重ね合わせで以下のように表すこともある:

$$p(\mathbf{x}) \approx \frac{1}{N} \sum_{i=1}^N \delta(\mathbf{x} - \mathbf{x}^{(i)}). \quad (6)$$

こうすると、様々な連続関数との積の積分が

$$\int g(\mathbf{x}) p(\mathbf{x}) d\mathbf{x} \approx \frac{1}{N} \sum_{i=1}^N g(\mathbf{x}^{(i)}) \quad (7)$$

と表現できる.

# モンテカルロ近似

$$\int g(\mathbf{x}) p(\mathbf{x}) d\mathbf{x} \approx \frac{1}{N} \sum_{i=1}^N g(\mathbf{x}^{(i)})$$

を使うと，確率密度関数  $p(\mathbf{x})$  の平均，分散も

$$\int \mathbf{x} p(\mathbf{x}) d\mathbf{x} \approx \frac{1}{N} \sum_{i=1}^N \mathbf{x}^{(i)} \quad (8)$$

$$\int (\mathbf{x} - \boldsymbol{\mu})(\mathbf{x} - \boldsymbol{\mu})^{\top} p(\mathbf{x}) d\mathbf{x} \approx \frac{1}{N} \sum_{i=1}^N (\mathbf{x}^{(i)} - \boldsymbol{\mu})(\mathbf{x}^{(i)} - \boldsymbol{\mu})^{\top} \quad (9)$$

のようにアンサンブルの標本平均，標本分散で近似される．

必要なメモリ容量は， $N \times \dim \mathbf{x}$  のオーダー．

# 高次元の問題におけるモンテカルロ近似

$p(\mathbf{x})$  の平均, 分散は

$$\int \mathbf{x} p(\mathbf{x}) d\mathbf{x} \approx \frac{1}{N} \sum_{i=1}^N \mathbf{x}^{(i)}$$

$$\int (\mathbf{x} - \boldsymbol{\mu})(\mathbf{x} - \boldsymbol{\mu})^\top p(\mathbf{x}) d\mathbf{x} \approx \frac{1}{N} \sum_{i=1}^N (\mathbf{x}^{(i)} - \boldsymbol{\mu})(\mathbf{x}^{(i)} - \boldsymbol{\mu})^\top$$

のように近似したとき...

- $p(\mathbf{x})$  をガウス分布  $\mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{I})$  としたとき, 標本平均のばらつきは  $\sigma/\sqrt{N}$ , 標本分散のばらつきは  $\sqrt{2}\sigma^2/\sqrt{N}$ . 標本平均, 標本分散の近似精度は, 次元に依らず,  $N$  を増やせばよくなる.
- 但し, 標本分散共分散行列のランクは  $N - 1$  以下なので,  $N < \text{rank} + 1$  の場合, 分散共分散行列を完全に表すことはできない.



# モンテカルロ近似による予測

- 非線型システム  $\mathbf{x}_k = \mathbf{f}_k(\mathbf{x}_{k-1}) + \mathbf{v}_k$  を考える.
- $\mathbf{v}_k$  は正規分布  $\mathcal{N}(0, \mathbf{Q})$  に従うものとする.
- 時刻  $t_{k-1}$  でのフィルタ分布  $p(\mathbf{x}_{k-1} | \mathbf{y}_{1:k-1})$  が, 粒子  $\{\mathbf{x}_{k-1|k-1}^{(i)}\}_{i=1}^N$  によって,

$$p(\mathbf{x}_{k-1} | \mathbf{y}_{1:k-1}) \approx \frac{1}{N} \sum_{i=1}^N \delta(\mathbf{x}_{k-1} - \mathbf{x}_{k-1|k-1}^{(i)}) \quad (10)$$

のように近似されている状況で時刻  $t_k$  の状態の予測  $p(\mathbf{x}_k | \mathbf{y}_{1:k-1})$  を得たい.

# モンテカルロ近似による予測

$$p(\mathbf{x}_{k-1} | \mathbf{y}_{1:k-1}) \approx \frac{1}{N} \sum_{i=1}^N \delta(\mathbf{x}_{k-1} - \mathbf{x}_{k-1|k-1}^{(i)})$$

のように  $p(\mathbf{x}_{k-1} | \mathbf{y}_{1:k-1})$  が近似されているとする。

$\{\mathbf{x}_{k-1|k-1}^{(i)}\}$  と独立に  $p(\mathbf{v}_k) = \mathcal{N}(\mathbf{0}, \mathbf{Q})$  からのランダムサンプル  $\{\mathbf{v}_k^{(i)}\}$  を抽出し、

$$\mathbf{x}_{k|k-1}^{(i)} = \mathbf{f}_k(\mathbf{x}_{k-1|k-1}^{(i)}) + \mathbf{v}_k^{(i)} \quad (11)$$

のように  $\{\mathbf{x}_{k|k-1}^{(1)}, \dots, \mathbf{x}_{k|k-1}^{(N)}\}$  を生成すれば、これが  $p(\mathbf{x}_k | \mathbf{y}_{1:k-1})$  から抽出した  $\mathbf{x}_k$  のランダムサンプルになる。

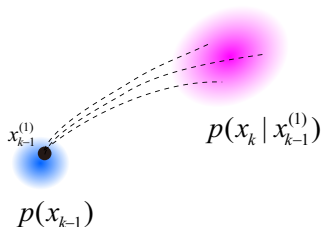
# モンテカルロ近似による予測

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

は、予測分布  $p(\mathbf{x}_k | \mathbf{y}_{1:k-1})$  を求める積分

$$p(\mathbf{x}_k | \mathbf{y}_{1:k-1}) = \int p(\mathbf{x}_k | \mathbf{x}_{k-1}) p(\mathbf{x}_{k-1} | \mathbf{y}_{1:k-1}) d\mathbf{x}_{k-1} \quad (12)$$

をそのままモンテカルロ法で計算したような形になっている。



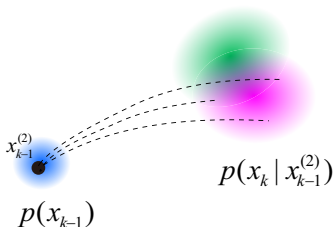
# モンテカルロ近似による予測

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

は、予測分布  $p(\mathbf{x}_k | \mathbf{y}_{1:k-1})$  を求める積分

$$p(\mathbf{x}_k | \mathbf{y}_{1:k-1}) = \int p(\mathbf{x}_k | \mathbf{x}_{k-1}) p(\mathbf{x}_{k-1} | \mathbf{y}_{1:k-1}) d\mathbf{x}_{k-1} \quad (12)$$

をそのままモンテカルロ法で計算したような形になっている。



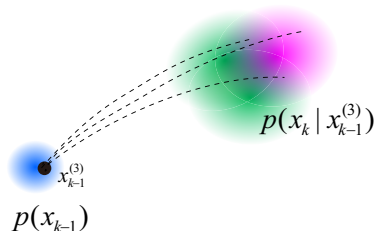
# モンテカルロ近似による予測

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

は、予測分布  $p(\mathbf{x}_k | \mathbf{y}_{1:k-1})$  を求める積分

$$p(\mathbf{x}_k | \mathbf{y}_{1:k-1}) = \int p(\mathbf{x}_k | \mathbf{x}_{k-1}) p(\mathbf{x}_{k-1} | \mathbf{y}_{1:k-1}) d\mathbf{x}_{k-1} \quad (12)$$

をそのままモンテカルロ法で計算したような形になっている。



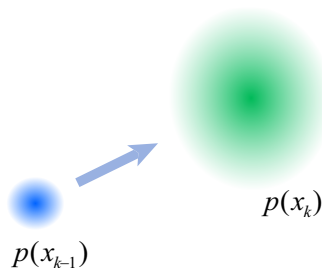
# モンテカルロ近似による予測

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

は、予測分布  $p(\mathbf{x}_k | \mathbf{y}_{1:k-1})$  を求める積分

$$p(\mathbf{x}_k | \mathbf{y}_{1:k-1}) = \int p(\mathbf{x}_k | \mathbf{x}_{k-1}) p(\mathbf{x}_{k-1} | \mathbf{y}_{1:k-1}) d\mathbf{x}_{k-1} \quad (12)$$

をそのままモンテカルロ法で計算したような形になっている。



# 平均ベクトル

$\mathbf{x}_{k|k-1}^{(i)} = \mathbf{f}_k(\mathbf{x}_{k-1|k-1}^{(i)}) + \mathbf{v}_k^{(i)}$  の操作によって、予測分布の平均が正しく得られることを確認しておく。

非線型システム  $\mathbf{x}_k = \mathbf{f}_k(\mathbf{x}_{k-1}) + \mathbf{v}_k$  において、 $p(\mathbf{x}_{k-1}|\mathbf{y}_{1:k-1})$  を与えた上で、時刻  $t_k$  の状態の予測  $\mathbf{x}_k = \mathbf{f}_k(\mathbf{x}_{k-1}) + \mathbf{v}_k$  の平均ベクトル  $\bar{\mathbf{x}}_{k|k-1}$  を計算してみると、

$$\begin{aligned}\bar{\mathbf{x}}_{k|k-1} &= \int (\mathbf{f}_k(\mathbf{x}_{k-1}) + \mathbf{v}_k) p(\mathbf{x}_{k-1}|\mathbf{y}_{1:k-1}) p(\mathbf{v}_k) d\mathbf{x}_{k-1} d\mathbf{v}_k \\ &= \int \mathbf{f}_k(\mathbf{x}_{k-1}) p(\mathbf{x}_{k-1}|\mathbf{y}_{1:k-1}) d\mathbf{x}_{k-1} + \int \mathbf{v}_k p(\mathbf{v}_k) d\mathbf{v}_k \\ &= \int \mathbf{f}_k(\mathbf{x}_{k-1}) p(\mathbf{x}_{k-1}|\mathbf{y}_{1:k-1}) d\mathbf{x}_{k-1}.\end{aligned}\quad (13)$$

# アンサンブルの標本平均

一方,  $\{\mathbf{x}_{k|k-1}^{(i)}\}$  の標本平均を計算すると,

$$\begin{aligned}\hat{\mathbf{x}}_{k|k-1} &= \frac{1}{N} \sum_{i=1}^N \mathbf{x}_{k|k-1}^{(i)} = \frac{1}{N} \sum_{i=1}^N \mathbf{f}(\mathbf{x}_{k-1|k-1}^{(i)}) + \frac{1}{N} \sum_{i=1}^N \mathbf{v}_k^{(i)} \\ &\approx \int \mathbf{f}_k(\mathbf{x}_{k-1}) p(\mathbf{x}_{k-1} | \mathbf{y}_{1:k-1}) d\mathbf{x}_{k-1} + \int \mathbf{v}_k p(\mathbf{v}_k) d\mathbf{v}_k \\ &= \int \mathbf{f}_k(\mathbf{x}_{k-1}) p(\mathbf{x}_{k-1} | \mathbf{y}_{1:k-1}) d\mathbf{x}_{k-1} = \bar{\mathbf{x}}_{k|k-1}.\end{aligned}\tag{14}$$

但し, 以下を使った:

$$p(\mathbf{x}) \approx \frac{1}{N} \sum_{i=1}^N \delta(\mathbf{x} - \mathbf{x}^{(i)}) \implies \int g(\mathbf{x}) p(\mathbf{x}) d\mathbf{x} \approx \frac{1}{N} \sum_{i=1}^N g(\mathbf{x}^{(i)}).\tag{15}$$



# 分散共分散行列

$\mathbf{x}_k = \mathbf{f}_k(\mathbf{x}_{k-1}) + \mathbf{v}_k$  の分散共分散行列  $\mathbf{P}_{k|k-1}$  は,

$$\begin{aligned}\mathbf{P}_{k|k-1} &= \int (\mathbf{f}_k(\mathbf{x}_{k-1}) + \mathbf{v}_k - \bar{\mathbf{x}}_{k|k-1}) (\mathbf{f}_k(\mathbf{x}_{k-1}) + \mathbf{v}_k - \bar{\mathbf{x}}_{k|k-1})^\top \\ &\quad \times p(\mathbf{x}_{k-1} | \mathbf{y}_{1:k-1}) p(\mathbf{v}_k) d\mathbf{x}_{k-1} d\mathbf{v}_k \\ &= \int (\mathbf{f}_k(\mathbf{x}_{k-1}) - \bar{\mathbf{x}}_{k|k-1}) (\mathbf{f}_k(\mathbf{x}_{k-1}) - \bar{\mathbf{x}}_{k|k-1})^\top p(\mathbf{x}_{k-1} | \mathbf{y}_{1:k-1}) d\mathbf{x}_{k-1} \\ &\quad + \int \mathbf{v}_k \mathbf{v}_k^\top p(\mathbf{v}_k) d\mathbf{v}_k.\end{aligned}\tag{16}$$

但し、途中の式変形で、 $\mathbf{f}_k(\mathbf{x}_{k-1})$  と  $\mathbf{v}_k$  とが独立で、 $\mathbf{v}_k$  の平均が 0 と仮定し、

$$\int (\mathbf{f}_k(\mathbf{x}_{k-1}) - \bar{\mathbf{x}}_{k|k-1}) \mathbf{v}_k^\top d\mathbf{v}_k = 0\tag{17}$$

とした。

# 標本分散共分散行列

一方,  $\{\mathbf{x}_{k|k-1}^{(i)}\}$  の標本分散共分散行列を計算すると,

$$\begin{aligned}
 \hat{\mathbf{P}}_{k|k-1} &= \frac{1}{N} \sum_{i=1}^N (\mathbf{x}_{k|k-1}^{(i)} - \hat{\mathbf{x}}_{k|k-1})(\mathbf{x}_{k|k-1}^{(i)} - \hat{\mathbf{x}}_{k|k-1})^\top \\
 &= \frac{1}{N} \sum_{i=1}^N \left( \mathbf{f}(\mathbf{x}_{k|k-1}^{(i)}) - \hat{\mathbf{x}}_{k|k-1} \right) \left( \mathbf{f}(\mathbf{x}_{k|k-1}^{(i)}) - \hat{\mathbf{x}}_{k|k-1} \right)^\top + \frac{1}{N} \sum_{i=1}^N \mathbf{v}_k^{(i)} \mathbf{v}_k^{(i)\top} \\
 &\approx \int (\mathbf{f}_k(\mathbf{x}_{k-1}) - \hat{\mathbf{x}}_{k|k-1})(\mathbf{f}_k(\mathbf{x}_{k-1}) - \hat{\mathbf{x}}_{k|k-1})^\top p(\mathbf{x}_{k-1} | \mathbf{y}_{1:k-1}) d\mathbf{x}_{k-1} \\
 &\quad + \int \mathbf{v}_k \mathbf{v}_k^\top p(\mathbf{v}_k) d\mathbf{v}_k = \mathbf{P}_{k|k-1}.
 \end{aligned} \tag{18}$$

但し,  $\mathbf{f}(\mathbf{x}_{k|k-1}^{(i)})$  と  $\mathbf{v}_k$  とが独立ゆえに無相関とし, 以下を仮定した:

$$\frac{1}{N} \sum_{i=1}^N \left( \mathbf{f}(\mathbf{x}_{k|k-1}^{(i)}) - \hat{\mathbf{x}}_{k|k-1} \right) \mathbf{v}_k^{(i)\top} \approx 0. \tag{19}$$

# モンテカルロ近似による予測のまとめ

$p(\mathbf{x}_{k-1} | \mathbf{y}_{1:k-1})$  が

$$p(\mathbf{x}_{k-1} | \mathbf{y}_{1:k-1}) \approx \frac{1}{N} \sum_{i=1}^N \delta(\mathbf{x}_{k-1} - \mathbf{x}_{k-1|k-1}^{(i)})$$

のように近似されているとすると,  $\{\mathbf{x}_{k-1|k-1}^{(i)}\}$  と独立に  $p(\mathbf{v}_k) = \mathcal{N}(\mathbf{0}, \mathbf{Q})$  からのランダムサンプル  $\{\mathbf{v}_k^{(i)}\}$  を抽出し,

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

のようにすれば,  $p(\mathbf{x}_k | \mathbf{y}_{1:k-1})$  からのランダムサンプルによるアンサンブル  $\{\mathbf{x}_{k|k-1}^{(i)}\}$  が得られ, 平均  $\bar{\mathbf{x}}_{k|k-1}$ , 分散共分散行列  $\mathbf{P}_{k|k-1}$  をよく近似できる.

ヤコビ行列  $\tilde{\mathbf{F}}_{k, \mathbf{x}_{k-1|k-1}}$  が必要ないばかりか,  $\mathbf{f}_k$  を非線型のまま扱える.

# フィルタリング

一般には、システムが非線型であれば  $p(x_k | y_{1:k-1})$  はガウス分布にならないが、アンサンブルカルマンフィルタ (Evensen, 1994; 2003) ではガウス分布を事前分布として仮定したカルマンフィルタのアルゴリズムに基づいて、観測  $y_k$  の情報を取り入れた事後分布  $p(x_k | y_{1:k})$  を求める。

まず、予測分布を表現するアンサンブルの標本平均、標本分散共分散行列

$$\hat{x}_{k|k-1} = \frac{1}{N} \sum_{i=1}^N x_{k|k-1}^{(i)}, \quad (20)$$

$$\hat{P}_{k|k-1} = \frac{1}{N-1} \sum_{i=1}^N (x_{k|k-1}^{(i)} - \hat{x}_{k|k-1})(x_{k|k-1}^{(i)} - \hat{x}_{k|k-1})^T \quad (21)$$

を求める。

これでは、 $\dim x_k$  のサイズの標本分散共分散行列を保持するメモリが必要になってしまうのだが、実はうまい対処法がある。(後述)

# フィルタリング

標本分散からカルマンゲインを求める

$$\hat{\mathbf{K}}_k = \hat{\mathbf{P}}_{k|k-1} \mathbf{H}_k^T (\mathbf{H}_k \hat{\mathbf{P}}_{k|k-1} \mathbf{H}_k^T + \mathbf{R}_k)^{-1}. \quad (22)$$

その後は、カルマンフィルタと同様に

$$\begin{aligned} \hat{\mathbf{x}}_{k|k} &= \hat{\mathbf{x}}_{k|k-1} + \hat{\mathbf{K}}_k (\mathbf{y}_k - \mathbf{H}_k \hat{\mathbf{x}}_{k|k-1}), \\ \hat{\mathbf{P}}_{k|k} &= (\mathbf{I} - \hat{\mathbf{K}}_k \mathbf{H}_k) \hat{\mathbf{P}}_{k|k-1} \end{aligned}$$

を計算してもいいのだが、アンサンブルカルマンフィルタでは、各アンサンブルメンバー (粒子) に

$$\mathbf{x}_{k|k}^{(i)} = \mathbf{x}_{k|k-1}^{(i)} + \hat{\mathbf{K}}_k (\mathbf{y}_k - \mathbf{H}_k \mathbf{x}_{k|k-1}^{(i)} + \mathbf{w}_k^{(i)}) \quad (23)$$

を適用し、直接、事後分布  $p(\mathbf{x}_k | \mathbf{y}_{1:k})$  を表現するアンサンブル  $\{\mathbf{x}_{k|k}^{(i)}\}$  を求める。

# フィルタリング

$$\mathbf{x}_{k|k}^{(i)} = \mathbf{x}_{k|k-1}^{(i)} + \hat{\mathbf{K}}_k(\mathbf{y}_k - \mathbf{H}_k \mathbf{x}_{k|k-1}^{(i)} + \mathbf{w}_k^{(i)})$$

の  $\mathbf{w}_k^{(i)}$  は何かというと、観測ノイズの分布  $p(\mathbf{w}_k) = \mathcal{N}(\mathbf{0}, \mathbf{R})$  からランダムに生成したサンプル。これを加えておかないと、新しく作ったアンサンブル  $\{\mathbf{x}_{k|k}^{(i)}\}$  の分散共分散行列が過小になってしまう。(後述)

まずは、こうして作ったアンサンブル  $\{\mathbf{x}_{k|k}^{(i)}\}$  のアンサンブルの平均が、粒子数を無限大にすると、求めたい事後分布の平均ベクトルと一致することを確認する:

$$\begin{aligned} \hat{\mathbf{x}}_{k|k} &= \frac{1}{N} \sum_{i=1}^N \left[ \mathbf{x}_{k|k-1}^{(i)} + \hat{\mathbf{K}}_k(\mathbf{y}_k - \mathbf{H}_k \mathbf{x}_{k|k-1}^{(i)} + \mathbf{w}_k^{(i)}) \right] \\ &\rightarrow \bar{\mathbf{x}}_{k|k-1} + \mathbf{K}_k(\mathbf{y}_k - \mathbf{H}_k \bar{\mathbf{x}}_{k|k-1}) \quad (N \rightarrow \infty). \end{aligned} \tag{24}$$

但し、ここで  $\frac{1}{N} \sum_{i=1}^N \mathbf{x}_{k|k-1}^{(i)} \rightarrow \bar{\mathbf{x}}_{k|k-1}$ ,  $\frac{1}{N} \sum_{i=1}^N \mathbf{w}_k^{(i)} \rightarrow 0$ ,  $\hat{\mathbf{K}}_k \rightarrow \mathbf{K}_k$  ( $N \rightarrow \infty$ ) を使った。

# フィルタリング

次にアンサンブル  $\{\mathbf{x}_{k|k}^{(i)}\}$  の分散共分散行列を考える。

準備として  $\mathbf{x}_{k|k}^{(i)} - \hat{\mathbf{x}}_{k|k}$  を計算しておく。

$$\begin{aligned}\mathbf{x}_{k|k}^{(i)} &= \mathbf{x}_{k|k-1}^{(i)} + \hat{\mathbf{K}}_k(\mathbf{y}_k - \mathbf{H}_k\mathbf{x}_{k|k-1}^{(i)} + \mathbf{w}_k^{(i)}) \\ \hat{\mathbf{x}}_{k|k} &= \hat{\mathbf{x}}_{k|k-1} + \hat{\mathbf{K}}_k(\mathbf{y}_k - \mathbf{H}_k\hat{\mathbf{x}}_{k|k-1} + \hat{\mathbf{w}}_k).\end{aligned}$$

但し,  $\hat{\mathbf{w}}_k = \frac{1}{N} \sum_{i=1}^N \mathbf{w}_k^{(i)}$ . したがって,

$$\begin{aligned}\mathbf{x}_{k|k}^{(i)} - \hat{\mathbf{x}}_{k|k} &= (\mathbf{x}_{k|k-1}^{(i)} - \hat{\mathbf{x}}_{k|k-1}) - \hat{\mathbf{K}}_k\mathbf{H}_k(\mathbf{x}_{k|k-1}^{(i)} - \hat{\mathbf{x}}_{k|k-1}) + \hat{\mathbf{K}}_k(\mathbf{w}_k^{(i)} - \hat{\mathbf{w}}_k) \\ &= (\mathbf{I} - \hat{\mathbf{K}}_k\mathbf{H}_k)(\mathbf{x}_{k|k-1}^{(i)} - \hat{\mathbf{x}}_{k|k-1}) + \hat{\mathbf{K}}_k(\mathbf{w}_k^{(i)} - \hat{\mathbf{w}}_k).\end{aligned}\tag{25}$$

# フィルタリング

$\mathbf{x}_{k|k}^{(i)} - \hat{\mathbf{x}}_{k|k} = (\mathbf{I} - \hat{\mathbf{K}}_k \mathbf{H}_k)(\mathbf{x}_{k|k-1}^{(i)} - \hat{\mathbf{x}}_{k|k-1}) + \hat{\mathbf{K}}_k(\mathbf{w}_k^{(i)} - \hat{\mathbf{w}}_k)$  より,

$$\begin{aligned}\hat{\mathbf{P}}_{k|k} &= \frac{1}{N-1} \sum_{i=1}^N \left( \mathbf{x}_{k|k}^{(i)} - \hat{\mathbf{x}}_{k|k} \right) \left( \mathbf{x}_{k|k}^{(i)} - \hat{\mathbf{x}}_{k|k} \right)^\top \\ &= \frac{1}{N-1} \sum_{i=1}^N \left[ \left( \mathbf{I} - \hat{\mathbf{K}}_k \mathbf{H}_k \right) \left( \mathbf{x}_{k|k-1}^{(i)} - \hat{\mathbf{x}}_{k|k-1} \right) \left( \mathbf{x}_{k|k-1}^{(i)} - \hat{\mathbf{x}}_{k|k-1} \right)^\top \left( \mathbf{I} - \hat{\mathbf{K}}_k \mathbf{H}_k \right)^\top \right. \\ &\quad \left. + 2 \left( \mathbf{I} - \hat{\mathbf{K}}_k \mathbf{H}_k \right) \left( \mathbf{x}_{k|k}^{(i)} - \hat{\mathbf{x}}_{k|k} \right) (\mathbf{w}_k^{(i)} - \hat{\mathbf{w}}_k)^\top \hat{\mathbf{K}}_k^\top + \hat{\mathbf{K}}_k (\mathbf{w}_k^{(i)} - \hat{\mathbf{w}}_k) (\mathbf{w}_k^{(i)} - \hat{\mathbf{w}}_k)^\top \hat{\mathbf{K}}_k^\top \right] \\ &\rightarrow (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \mathbf{P}_{k|k-1} (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k)^\top + \mathbf{K}_k \mathbf{R} \mathbf{K}_k^\top. \quad (N \rightarrow \infty)\end{aligned}$$

(26)

但し,  $\mathbf{x}_{k|k}^{(i)}$  と  $\mathbf{w}_k^{(i)}$  を無相関と仮定し,  $N \rightarrow \infty$  に対して以下が成り立つことを使った:

$$\frac{1}{N-1} \sum_{i=1}^N \left( \mathbf{x}_{k|k}^{(i)} - \hat{\mathbf{x}}_{k|k} \right) (\mathbf{w}_k^{(i)} - \hat{\mathbf{w}}_k)^\top \rightarrow \mathbf{O}, \quad \frac{1}{N-1} \sum_{i=1}^N (\mathbf{w}_k^{(i)} - \hat{\mathbf{w}}_k) (\mathbf{w}_k^{(i)} - \hat{\mathbf{w}}_k)^\top \rightarrow \mathbf{R}.$$



# フィルタリング

$\hat{\mathbf{P}}_{k|k}$  をさらに計算すると,

$$\begin{aligned}
 \hat{\mathbf{P}}_{k|k} &\rightarrow (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \mathbf{P}_{k|k-1} (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k)^\top + \mathbf{K}_k \mathbf{R} \mathbf{K}_k^\top \\
 &= (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \mathbf{P}_{k|k-1} - (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \mathbf{P}_{k|k-1} \mathbf{H}_k^\top \mathbf{K}_k^\top + \mathbf{K}_k \mathbf{R} \mathbf{K}_k^\top \\
 &= (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \mathbf{P}_{k|k-1} - \mathbf{P}_{k|k-1} \mathbf{H}_k^\top \mathbf{K}_k^\top + \mathbf{K}_k \mathbf{H}_k \mathbf{P}_{k|k-1} \mathbf{H}_k^\top \mathbf{K}_k^\top + \mathbf{K}_k \mathbf{R} \mathbf{K}_k^\top \\
 &= (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \mathbf{P}_{k|k-1} - \mathbf{P}_{k|k-1} \mathbf{H}_k^\top \mathbf{K}_k^\top + \mathbf{K}_k (\mathbf{H}_k \mathbf{P}_{k|k-1} \mathbf{H}_k^\top + \mathbf{R}) \mathbf{K}_k^\top \\
 &= (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \mathbf{P}_{k|k-1} - \mathbf{P}_{k|k-1} \mathbf{H}_k^\top \mathbf{K}_k^\top + \mathbf{P}_{k|k-1} \mathbf{H}_k^\top \mathbf{K}_k^\top \\
 &= (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \mathbf{P}_{k|k-1}
 \end{aligned} \tag{27}$$

となり, 確かにカルマンフィルタの式から得られる分散共分散行列と一致する.

$w_k^{(i)}$  を加えなかった場合には,

$$\hat{\mathbf{P}}_{k|k} \rightarrow (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \mathbf{P}_{k|k-1} (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k)^\top.$$

となるので, やや過小評価になる.

# アンサンブルカルマンフィルタ (EnKF) のまとめ

## ■ 予測:

- 各アンサンブルメンバー (粒子) にシステムモデルを適用.

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

## ■ フィルタリング:

- 予測分布を表現するアンサンブルの標本平均, 標本分散を求める.

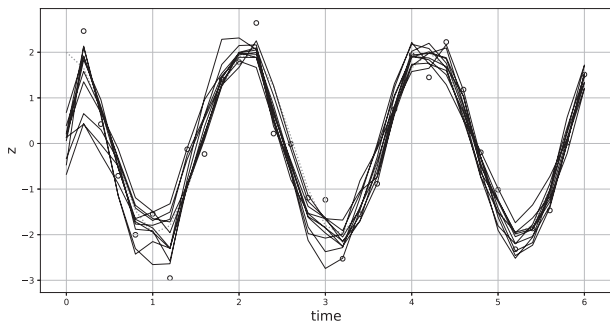
$$\hat{\mathbf{x}}_{k|k-1} = \frac{1}{N} \sum_{i=1}^N \mathbf{x}_{k|k-1}^{(i)},$$

$$\hat{\mathbf{P}}_{k|k-1} = \frac{1}{N-1} \sum_{i=1}^N (\mathbf{x}_{k|k-1}^{(i)} - \hat{\mathbf{x}}_{k|k-1})(\mathbf{x}_{k|k-1}^{(i)} - \hat{\mathbf{x}}_{k|k-1})^T$$

- カルマンゲイン  $\hat{\mathbf{K}}_k = \hat{\mathbf{P}}_{k|k-1} \mathbf{H}_k^T (\mathbf{H}_k \hat{\mathbf{P}}_{k|k-1} \mathbf{H}_k^T + \mathbf{R}_k)^{-1}$  を求めた上で, 各アンサンブルメンバー (粒子) に以下を適用し, 事後分布  $p(\mathbf{x}_k | \mathbf{y}_{1:k})$  を表現するアンサンブル  $\{\mathbf{x}_{k|k}^{(i)}\}_{i=1}^N$  を求める.

$$\mathbf{x}_{k|k}^{(i)} = \mathbf{x}_{k|k-1}^{(i)} + \hat{\mathbf{K}}_k (\mathbf{y}_k - \mathbf{H}_k \mathbf{x}_{k|k-1}^{(i)} + \mathbf{w}_k^{(i)}).$$

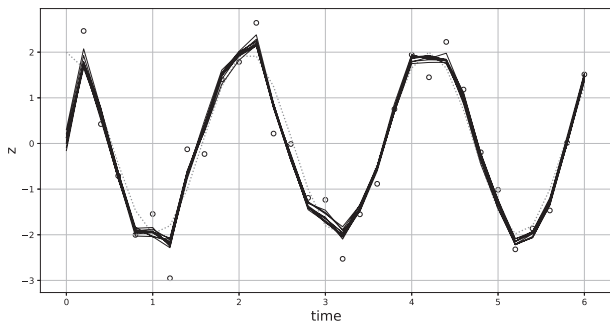
# 単振動の推定例



上の図は、単振動の問題にアンサンブルカルマンフィルタを適用した結果で、アンサンブルメンバー数は  $N = 5$  としている。

黒の実線は 10 個の異なる乱数の種を与えて推定した結果を示しているが、乱数の種によって結果がかなり異なっている。

# 単振動の推定例



上の図は、アンサンブルメンバー数を  $N = 50$  として、単振動の問題にアンサンブルカルマンフィルタを適用した結果である。

黒の実線は 10 個の異なる乱数の種を与えて推定した結果だが、 $N = 50$  にすると、どの実線も概ね同じ値になることが分かる。

# メモリの節約

- 今のアルゴリズムでは、標本分散共分散行列  $\hat{\mathbf{P}}_{k|k-1}$  を保持するために  $\dim \mathbf{x}_k \times \dim \mathbf{x}_k$  のオーダーのメモリが必要になってしまう。
- しかし、計算手順を工夫すれば、 $\hat{\mathbf{P}}_{k|k-1}$  をまじめに計算しなくても済む。
- 見通しをよくするために、

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

という行列を定義すると、標本分散共分散行列は  $\hat{\mathbf{P}}_{k|k-1} = \hat{\mathbf{X}}_{k|k-1} \hat{\mathbf{X}}_{k|k-1}^T$  のように分解できる。

- そうすると、カルマンゲインは以下のように変形できる。

$$\begin{aligned} \hat{\mathbf{K}}_k &= \hat{\mathbf{P}}_{k|k-1} \mathbf{H}_k^T (\mathbf{H}_k \hat{\mathbf{P}}_{k|k-1} \mathbf{H}_k^T + \mathbf{R}_k)^{-1} \\ &= \hat{\mathbf{X}}_{k|k-1} \hat{\mathbf{X}}_{k|k-1}^T \mathbf{H}_k^T (\mathbf{H}_k \hat{\mathbf{X}}_{k|k-1} \hat{\mathbf{X}}_{k|k-1}^T \mathbf{H}_k^T + \mathbf{R}_k)^{-1} \end{aligned} \quad (29)$$

# メモリの節約

- あとは,

$$\begin{aligned}\hat{\mathbf{K}}_k &= \hat{\mathbf{X}}_{k|k-1} \hat{\mathbf{X}}_{k|k-1}^\top \mathbf{H}_k^\top (\mathbf{H}_k \hat{\mathbf{X}}_{k|k-1} \hat{\mathbf{X}}_{k|k-1}^\top \mathbf{H}_k^\top + \mathbf{R}_k)^{-1} \\ &= \hat{\mathbf{X}}_{k|k-1} (\mathbf{H}_k \hat{\mathbf{X}}_{k|k-1})^\top \left( (\mathbf{H}_k \hat{\mathbf{X}}_{k|k-1}) (\mathbf{H}_k \hat{\mathbf{X}}_{k|k-1})^\top + \mathbf{R}_k \right)^{-1}\end{aligned}\quad (30)$$

のように, 先に  $\mathbf{H}_k \hat{\mathbf{X}}_{k|k-1}$  を計算してしまえば,  $\dim \mathbf{x}_k \times \dim \mathbf{x}_k$  のオーダーのメモリは必要なく,  $\dim \mathbf{y}_k \times N$  あるいは  $\dim \mathbf{y}_k \times \dim \mathbf{y}_k$  程度のメモリでカルマンゲインは計算できる.

- $N < \dim \mathbf{y}_k$  の場合は, さらに効率的に計算する方法がある.

# メモリの節約

- $N < \dim \mathbf{y}_k$  のときは、 $\hat{\mathbf{K}}_k$  を以下のように変形するとよい:

$$\begin{aligned}
 \hat{\mathbf{K}}_k &= \hat{\mathbf{P}}_{k|k-1} \mathbf{H}_k^\top (\mathbf{H}_k \hat{\mathbf{P}}_{k|k-1} \mathbf{H}_k^\top + \mathbf{R}_k)^{-1} \\
 &= \hat{\mathbf{X}}_{k|k-1} \hat{\mathbf{X}}_{k|k-1}^\top \mathbf{H}_k^\top (\mathbf{H}_k \hat{\mathbf{X}}_{k|k-1} \hat{\mathbf{X}}_{k|k-1}^\top \mathbf{H}_k^\top + \mathbf{R}_k)^{-1} \\
 &= \hat{\mathbf{X}}_{k|k-1} (\mathbf{I} + \hat{\mathbf{X}}_{k|k-1}^\top \mathbf{H}_k^\top \mathbf{R}_k^{-1} \mathbf{H}_k \hat{\mathbf{X}}_{k|k-1})^{-1} \hat{\mathbf{X}}_{k|k-1}^\top \mathbf{H}_k^\top \mathbf{R}_k^{-1}.
 \end{aligned} \tag{31}$$

但し、ここで逆行列補題の式

$$(\mathbf{P}^{-1} + \mathbf{H}^\top \mathbf{R}^{-1} \mathbf{H})^{-1} \mathbf{H}^\top \mathbf{R}^{-1} = \mathbf{P} \mathbf{H}^\top (\mathbf{H} \mathbf{P} \mathbf{H}^\top + \mathbf{R})^{-1}$$

を使った.

- $(\mathbf{I} + \hat{\mathbf{X}}_{k|k-1}^\top \mathbf{H}_k^\top \mathbf{R}_k^{-1} \mathbf{H}_k \hat{\mathbf{X}}_{k|k-1})$  は  $N \times N$  行列なので効率的に計算できる.

# アンサンブルカルマンフィルタ (改良版)

## ■ 予測:

- 各アンサンブルメンバー (粒子) にシステムモデルを適用.

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

## ■ フィルタリング:

- 以下を求める:

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

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

## ■ カルマンゲイン

$\hat{\mathbf{K}}_k = \hat{\mathbf{X}}_{k|k-1} (\mathbf{I} + \hat{\mathbf{X}}_{k|k-1}^T \mathbf{H}_k^T \mathbf{R}_k^{-1} \mathbf{H}_k \hat{\mathbf{X}}_{k|k-1})^{-1} \hat{\mathbf{X}}_{k|k-1}^T \mathbf{H}_k^T \mathbf{R}_k^{-1}$  を求めた上で、各アンサンブルメンバー (粒子) に以下を適用し、事後分布  $p(\mathbf{x}_k | \mathbf{y}_{1:k})$  を表現するアンサンブル  $\{\mathbf{x}_{k|k}^{(i)}\}_{i=1}^N$  を求める.

$$\mathbf{x}_{k|k}^{(i)} = \mathbf{x}_{k|k-1}^{(i)} + \hat{\mathbf{K}}_k (\mathbf{y}_k - \mathbf{H}_k \mathbf{x}_{k|k-1}^{(i)} + \mathbf{w}_k^{(i)}).$$



# 線型結合による実装

## アンサンブルカルマンフィルタの更新式

$$\begin{aligned}\mathbf{x}_{k|k}^{(i)} &= \mathbf{x}_{k|k-1}^{(i)} + \hat{\mathbf{K}}_k(\mathbf{y}_k - \mathbf{H}_k \mathbf{x}_{k|k-1}^{(i)} + \mathbf{w}_k^{(i)}) \\ &= \mathbf{x}_{k|k-1}^{(i)} + \hat{\mathbf{X}}_{k|k-1}(\mathbf{I} + \hat{\mathbf{X}}_{k|k-1}^\top \mathbf{H}_k^\top \mathbf{R}_k^{-1} \mathbf{H}_k \hat{\mathbf{X}}_{k|k-1})^{-1} \hat{\mathbf{X}}_{k|k-1}^\top \mathbf{H}_k^\top \mathbf{R}_k^{-1} \\ &\quad \times (\mathbf{y}_k - \mathbf{H}_k \mathbf{x}_{k|k-1}^{(i)} + \mathbf{w}_k^{(i)})\end{aligned}$$

で,

$$\boldsymbol{\beta}_k^{(i)} = (\mathbf{I} + \hat{\mathbf{X}}_{k|k-1}^\top \mathbf{H}_k^\top \mathbf{R}_k^{-1} \mathbf{H}_k \hat{\mathbf{X}}_{k|k-1})^{-1} \hat{\mathbf{X}}_{k|k-1}^\top \mathbf{H}_k^\top \mathbf{R}_k^{-1} (\mathbf{y}_k - \mathbf{H}_k \mathbf{x}_{k|k-1}^{(i)} + \mathbf{w}_k^{(i)})$$

とくと,

$$\mathbf{x}_{k|k}^{(i)} = \mathbf{x}_{k|k-1}^{(i)} + \hat{\mathbf{X}}_{k|k-1} \boldsymbol{\beta}_k^{(i)} \quad (32)$$

の形になる。

# 線型結合による実装

- $\mathbf{x}_{k|k}^{(i)} = \mathbf{x}_{k|k-1}^{(i)} + \hat{\mathbf{X}}_{k|k-1} \beta_k^{(i)}$  より、各  $\mathbf{x}_{k|k}^{(i)}$  は、元のアンサンブル  $\{\mathbf{x}_{k|k-1}^{(i)}\}_{i=1}^N$  の線型結合で表せる。但し、

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

を構成する列ベクトルが、それぞれ  $\{\mathbf{x}_{k|k-1}^{(i)}\}_{i=1}^N$  の線型結合になっていることに注意すること。

- 実際の  $\mathbf{x}_k$  には、様々な種類の変数が含まれている。気象シミュレーションの場合、圧力、温度、風速、など。
- しかし、 $\mathbf{x}_{k|k}^{(i)}$  が、元のアンサンブルの線型結合によって作られることから、圧力は圧力、風速は風速で別々に更新すればよいことが分かる。プログラムの実装もしやすい。

# 平滑化

- これまで、予測分布  $p(\mathbf{x}_k | \mathbf{y}_{1:k-1})$  とフィルタ分布  $p(\mathbf{x}_k | \mathbf{y}_{1:k})$  の計算方法を考えてきた。
- 将来を予測するためによい初期値を作るという目的であれば、フィルタ分布  $p(\mathbf{x}_k | \mathbf{y}_{1:k})$  でよいのだが、過去の状態を現在までのデータから推定したいという場合もあり得る。
- このようにデータよりも過去の状態を推定する操作を平滑化と呼ぶ。
- 過去を推定する（つまり平滑化を行う）には、状態ベクトル  $\mathbf{x}_k$  を過去の状態も含めた形で定義し直すとよい。すなわち、

$$\mathbf{x}_k^* = \left( \mathbf{x}_k^\top \mathbf{x}_{k-1}^\top \cdots \mathbf{x}_{k-L}^\top \right)^\top \quad (33)$$

のように、過去の状態も含む形で新たに状態ベクトル  $\mathbf{x}_k^*$  を定義する。

# 平滑化

状態ベクトルを

$$\mathbf{x}_k^* = (\mathbf{x}_k^\top \mathbf{x}_{k-1}^\top \cdots \mathbf{x}_{k-L}^\top)^\top$$

のように定義し直したのに伴い、状態空間モデルも以下のように修正する:

$$\mathbf{x}_k^* = \mathbf{f}_k^*(\mathbf{x}_{k-1}^*) + \mathbf{v}_k^*, \quad (34)$$

$$\mathbf{y}_k = \mathbf{H}_k^* \mathbf{x}_k^* + \mathbf{w}_k. \quad (35)$$

但し,  $\mathbf{f}_k^*$  は,

$$\mathbf{f}_k^*(\mathbf{x}_{k-1}^*) = \begin{pmatrix} \mathbf{f}_k(\mathbf{x}_{k-1}) \\ \mathbf{x}_{k-1} \\ \vdots \\ \mathbf{x}_{k-L} \end{pmatrix}, \quad \mathbf{v}_k^* = \begin{pmatrix} \mathbf{v}_k \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{pmatrix} \quad (36)$$

# 平滑化

$$\mathbf{f}_k^*(\mathbf{x}_{k-1}^*) = \begin{pmatrix} \mathbf{f}_k(\mathbf{x}_{k-1}) \\ \mathbf{x}_{k-1} \\ \vdots \\ \mathbf{x}_{k-L} \end{pmatrix}, \quad \mathbf{v}_k^* = \begin{pmatrix} \mathbf{v}_k \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{pmatrix}$$

のように定義すると、 $\mathbf{x}_k^*$  に対するアンサンブルカルマンフィルタの予測の手続き

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

は、以下のような操作に対応する:

$$\mathbf{x}_{k|k-1}^{*(i)} = \begin{pmatrix} \mathbf{x}_{k|k-1}^{(i)} \\ \mathbf{x}_{k-1|k-1}^{(i)} \\ \vdots \\ \mathbf{x}_{k-L|k-1}^{(i)} \end{pmatrix} = \begin{pmatrix} \mathbf{f}_k(\mathbf{x}_{k-1|k-1}^{(i)}) + \mathbf{v}_k^{(i)} \\ \mathbf{x}_{k-1|k-1}^{(i)} \\ \vdots \\ \mathbf{x}_{k-L|k-1}^{(i)} \end{pmatrix}. \quad (37)$$

つまり、 $\mathbf{x}_{k|k-1}^{(i)}$  を得る操作は変更なし、それ以外も自明。

# 平滑化

フィルタリングの操作は、先程述べたように

$$\mathbf{x}_{k|k}^{*(i)} = \mathbf{x}_{k|k-1}^{*(i)} + \hat{\mathbf{X}}_{k|k-1}^* \boldsymbol{\beta}_k^{(i)}$$

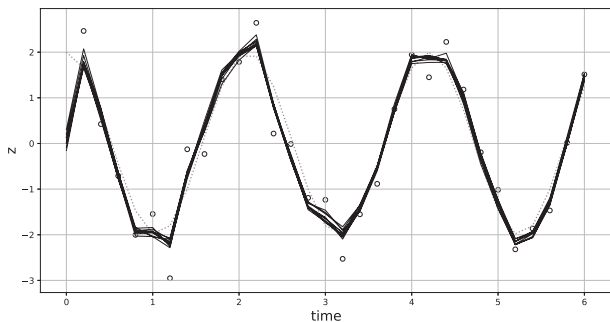
と書ける。この式をばらすと、

$$\begin{aligned}\mathbf{x}_{k|k}^{(i)} &= \mathbf{x}_{k|k-1}^{(i)} + \hat{\mathbf{X}}_{k|k-1} \boldsymbol{\beta}_k^{(i)} \\ \mathbf{x}_{k-1|k}^{(i)} &= \mathbf{x}_{k-1|k-1}^{(i)} + \hat{\mathbf{X}}_{k-1|k-1} \boldsymbol{\beta}_k^{(i)} \\ &\vdots \\ \mathbf{x}_{k-L|k}^{(i)} &= \mathbf{x}_{k-L|k-1}^{(i)} + \hat{\mathbf{X}}_{k-L|k-1} \boldsymbol{\beta}_k^{(i)}\end{aligned}$$

となるので、 $\mathbf{x}_k$  だけでなく、 $\mathbf{x}_{k-1}, \dots, \mathbf{x}_{k-L}$  についても同じ重み  $\boldsymbol{\beta}_k^{(i)}$  で線型結合を取ればよい。但し、

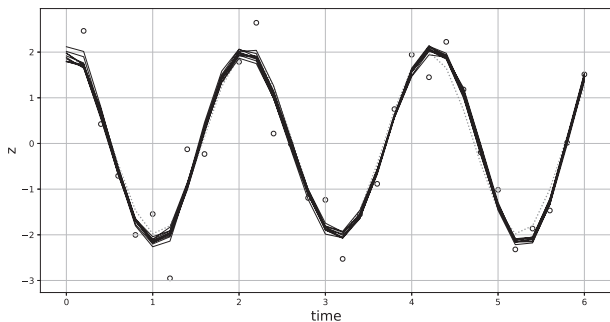
$$\hat{\mathbf{X}}_{k-l|k-1} = \frac{1}{\sqrt{N-1}} \begin{pmatrix} \mathbf{x}_{k-l|k-1}^{(1)} - \hat{\mathbf{x}}_{k-l|k-1} & \cdots & \mathbf{x}_{k-l|k-1}^{(N)} - \hat{\mathbf{x}}_{k-l|k-1} \end{pmatrix}, \quad (l = 0, \dots, L)$$

# 単振動の推定例



上の図は、アンサンブルメンバー数を  $N = 50$  として、単振動の問題にアンサンブルカルマンフィルタを適用した結果である。

# 単振動の推定例



同じアンサンブルメンバー数 ( $N = 50$ ) で、アンサンブルカルマンスムーザを適用すると上のようになる。



# 平滑化

- 後で述べるアンサンブル変換カルマンフィルタ、粒子フィルタでも、同様にして、過去の状態も含む新たな状態ベクトル  $x_k^*$  を定義し直すことで、平滑化が実現できる。
- 線型・ガウスの場合のカルマンフィルタでも、同様。
- 但し、線型・ガウスの場合、固定区間平滑化 (Rauch-Tung-Striebel スムーズ) と呼ばれる方法が使われることも多い。

## 課題 2

- 先程，カルマンフィルタを適用したのと同じ気象時系列データで同じ 1 変数状態空間モデル

$$x_k = x_{k-1} + v_k,$$

$$y_k = x_k + w_k$$

のもと，アンサンブルカルマンフィルタを適用し，各時点の  $x_k$  を推定せよ． $Q_k, R_k$  はデータの特徴に応じて適宜設定すること．（ $N$  は数十程度でよいと思う．）

## 課題 3

- アンサンブルカルマンフィルタを実際に試してみるためには、適当なモデルから、擬似的なデータを作り、アンサンブルカルマンフィルタで、同じモデルを使って推定できることを確認するとよい。
- ここでは、Lorenz 方程式 (Lorenz, 1963) を例として取り上げる:

$$\begin{aligned}\frac{dx}{dt} &= -s(x - y) \\ \frac{dy}{dt} &= rx - y - xz \\ \frac{dz}{dt} &= xy - bz,\end{aligned}$$

但し、パラメータは、元の論文に従って  $s = 10$ ,  $r = 28$ ,  $b = 8/3$  と設定する。

- 以下では、この Lorenz 方程式を

$$\frac{dx}{dt} = g(x)$$

と表す。

## 課題 3

手順:

1 まず, Lorenz 方程式

$$\frac{dx}{dt} = g(x)$$

を 4 次の Runge-Kutta 法

$$\kappa_1 = g(x_{k-1})$$

$$\kappa_2 = g\left(x_{k-1} + \frac{\Delta t}{2}\kappa_1\right)$$

$$\kappa_3 = g\left(x_{k-1} + \frac{\Delta t}{2}\kappa_2\right)$$

$$\kappa_4 = g(x_{k-1} + \Delta t\kappa_3)$$

$$x_k = x_{k-1} + \frac{\Delta t}{6}(\kappa_1 + 2\kappa_2 + 2\kappa_3 + \kappa_4)$$

などで積分し, 「真のシナリオ」を作る ( $\Delta t = 0.01$  程度).

## 課題 3

- 2 観測可能な時間間隔を設定し，真のシナリオにガウス乱数 (例えば  $\mathcal{N}(0, \mathbf{I})$ ) を加えて，擬似観測データを生成する． (例えば  $t$  が 0.1 増えるごとに観測が得られると仮定すればよい．)
- 3 元の 4 次の Runge-Kutta 法による時間発展のモデルを  $f_k$  として使い，アンサンブルカルマンフィルタで  $x_k$  を推定する．  $\mathbf{R} = \mathbf{I}$ ,  $\mathbf{Q} = 0.01\mathbf{I}$  程度に設定する．

# まとめ

- アンサンブルカルマンフィルタでは、モンテカルロ近似を用いることにより、非線型システムにおいても予測分布の計算を比較的容易に扱うことができる。
- アンサンブルカルマンフィルタにおけるフィルタの計算は、粒子数を無限大に取ったときにカルマンフィルタの結果と一致するように行われる。但し、予測の修正は各粒子ごとに行われる。
- 粒子数を少なく設定した場合には、計算手順を工夫することでカルマンフィルタで見られたメモリの問題も解決される。