SAMA Workshop on

Ensemble Methods in Meteorology and Oceanography

Institut Pierre Simon Laplace, Paris, May 15–16, 2008

Large Sample Asymptotics for the Ensemble Kalman Filter (EnKF)

François Le Gland

INRIA Rennes Bretagne–Atlantique

http://www.irisa.fr/aspi/

Valérie Monbet + Vu–Duc Tran

université de Bretagne Sud, Vannes

outline

- motivation : Kalman filter in large dimension
- EnKF as particle system with mean–field interaction
- identification of the limit
- large sample asymptotics of EnKF
- connection with particle filters
- conclusion and perspective

linear Gaussian state–space model

$$X_k = F_k \, X_{k-1} + W_k \qquad \text{with} \qquad W_k \sim \mathcal{N}(0, Q_k)$$

$$Y_k = H_k \, X_k + V_k \qquad \text{with} \qquad V_k \sim \mathcal{N}(0, R_k)$$

and Gaussian initial condition $X_0 \sim \mathcal{N}(m_0, \Sigma_0)$

observation noise covariance matrix R_k assumed invertible

conditional probability distribution

of hidden state X_k given past observations $Y_{0:k} = (Y_0, \cdots, Y_k)$

is Gaussian, with mean vector $\widehat{X}_k$ and covariance matrix P_k

Kalman filter equation

- prediction (forecast) step

$$\widehat{X}_k^- = F_k \, \widehat{X}_{k-1} \qquad \text{and} \qquad P_k^- = F_k \, P_{k-1} \, F_k^* + Q_k$$

- correction (analysis) step

$$\widehat{X}_k = \widehat{X}_k^- + K_k \, (Y_k - H_k \, \widehat{X}_k^-) \qquad \text{and} \qquad P_k = (I - K_k \, H_k) \, P_k^-$$

with the Kalman gain matrix defined by

$$K_k = P_k^- \, H_k^* \, (H_k \, P_k^- \, H_k^* + R_k)^{-1}$$

- initially $\widehat{X}_0^- = m_0$ and $P_0^- = \Sigma_0$.

if dimension m of hidden state is large, then computing and storing large $m \times m$ covariance matrices P_k^- and P_k is just impossible

matrix products in the prediction equation

$$P_k^- = F_k \, P_{k-1} \, F_k^* + Q_k$$

are even more problematic to work out

usually, dimension d of observation is much less, and matrix products in expression of Kalman gain matrix

$$K_k = P_k^- \, H_k^* \, (H_k \, P_k^- \, H_k^* + R_k)^{-1}$$

or in correction equation

$$P_k = (I - K_k \, H_k) \, P_k^- = P_k^- - P_k^- \, H_k^* \, (H_k \, P_k^- \, H_k^* + R_k)^{-1} \, H_k \, P_k^-$$

are much less problematic to work out

idea behind ensemble Kalman filter (EnKF) : use Monte Carlo samples and use empirical covariance matrix in place of prediction covariance matrix in practice, given an analysis ensemble $(X_{k-1}^{1,a}, \cdots, X_{k-1}^{N,a})$ of N elements, each ensemble element is propagated independently according to

$$X_k^{i,f} = F_k \, X_{k-1}^{i,a} + W_k^i \qquad \text{with} \qquad W_k^i \sim \mathcal{N}(0, Q_k)$$

notice that i.i.d. random vectors $(W_k^1, \cdots, W_k^N)$ are *simulated* here, with same statistics as additive Gaussian noise W_k in original state equation

initial ensemble $(X_0^{1,f}, \cdots, X_0^{N,f})$ is *simulated* as i.i.d. Gaussian random vectors with mean m_0 and covariance matrix Σ_0, i.e. with same statistics as initial condition X_0

empirical mean vector and covariance matrix of forecast elements $(X_k^{1,f}, \cdots, X_k^{N,f})$ are defined as

$$m_k^N = \frac{1}{N} \sum_{i=1}^N X_k^{i,f} \qquad \text{and} \qquad P_k^N = \frac{1}{N} \sum_{i=1}^N (X_k^{i,f} - m_k^N)\, (X_k^{i,f} - m_k^N)^*$$

respectively

this empirical covariance matrix is then used in correction step as follows

$$X_k^{i,a} = X_k^{i,f} + K_k^N \, (Y_k - H_k \, X_k^{i,f} - V_k^i) \qquad \text{with} \qquad V_k^i \sim \mathcal{N}(0, R_k)$$

with empirical Kalman gain matrix defined by

$$K_k^N = P_k^N \, H_k^* \, (H_k \, P_k^N \, H_k^* + R_k)^{-1}$$

notice that i.i.d. random vectors $(V_k^1, \cdots, V_k^N)$ are *simulated* here, with same statistics as additive Gaussian noise V_k in original observation equation in practice

- only samples are used
- empirical covariance matrix is never computed

indeed, to evaluate matrix–vector product $P_k^N \, u$ where u is a (column) vector of dimension m, only N scalar products need to be evaluated, since

$$P_k^N \, u = [\frac{1}{N} \, \sum_{i=1}^N (X_k^{i,f} - m_k^N) \, (X_k^{i,f} - m_k^N)^*] \, u = \frac{1}{N} \, \sum_{i=1}^N \lambda_k^i \, (X_k^{i,f} - m_k^N)$$

with $\lambda_k^i = (X_k^{i,f} - m_k^N)^* \, u$ for any $i = 1, \cdots, N$

in particular, H_k can be seen as a collection of d (row) vectors of dimension m, and to evaluate matrix products $P_k^N\ H_k^*$ and $H_k\ P_k^N\ H_k^*$, only $N \times d$ scalar products need to be evaluated, since

$$P_k^N\ H_k^* = [\frac{1}{N}\ \sum_{i=1}^{N} (X_k^{i,f} - m_k^N)\ (X_k^{i,f} - m_k^N)^*]\ H_k^* = \frac{1}{N}\ \sum_{i=1}^{N} (X_k^{i,f} - m_k^N)\ (h_k^i)^*$$

and

$$H_k\ P_k^N\ H_k^* = H_k\ [\frac{1}{N}\ \sum_{i=1}^{N} (X_k^{i,f} - m_k^N)\ (X_k^{i,f} - m_k^N)^*]\ H_k^* = \frac{1}{N}\ \sum_{i=1}^{N} h_k^i\ (h_k^i)^*$$

with $h_k^i = H_k\,(X_k^{i,f} - m_k^N)$ for any $i = 1, \cdots, N$

question : does the empirical mean of the ensemble elements converge to the Kalman filter, i.e. does

$$\frac{1}{N} \sum_{i=1}^{N} X_k^{i,f} \longrightarrow \widehat{X}_k^- \qquad \text{and} \qquad \frac{1}{N} \sum_{i=1}^{N} X_k^{i,a} \longrightarrow \widehat{X}_k$$

hold, as $N \uparrow \infty$?

answer is YES

outline

- motivation : Kalman filter in large dimension
- EnKF as particle system with mean–field interaction
- identification of the limit
- large sample asymptotics of EnKF
- connection with particle filters
- conclusion and perspective

ensemble Kalman filter idea has been extended to any system of the form

$$X_k = f_k(X_{k-1}) + W_k \qquad \text{with} \qquad W_k \sim \mathcal{N}(0, Q_k)$$

$$Y_k = H_k \, X_k + V_k \qquad \text{with} \qquad V_k \sim \mathcal{N}(0, R_k)$$

with non–necessarily Gaussian initial condition $X_0 \sim \mu_0$

in practice, given an analysis ensemble $(X_{k-1}^{1,a}, \cdots, X_{k-1}^{N,a})$ of N elements, each ensemble element is propagated independently according to the following set of decoupled equations

$$X_k^{i,f} = f_k(X_{k-1}^{i,a}) + W_k^i \qquad \text{with} \qquad W_k^i \sim \mathcal{N}(0, Q_k)$$

notice that i.i.d. random vectors $(W_k^1, \cdots, W_k^N)$ are *simulated* here, with same statistics as the additive Gaussian noise W_k in the original state equation

initial ensemble $(X_0^{1,f}, \cdots, X_0^{N,f})$ is *simulated* as i.i.d. random vectors with probability distribution μ_0, i.e. with the same statistics as the initial condition X_0

empirical mean vector and covariance matrix of forecast elements $(X_k^{1,f}, \cdots, X_k^{N,f})$ are defined as

$$m_k^{N,f} = \frac{1}{N} \sum_{i=1}^{N} X_k^{i,f} \qquad \text{and} \qquad P_k^{N,f} = \frac{1}{N} \sum_{i=1}^{N} (X_k^{i,f} - m_k^{N,f})\, (X_k^{i,f} - m_k^{N,f})^*$$

respectively

this empirical covariance matrix is then used in correction step to produce a new analysis ensemble $(X_k^{1,a}, \cdots, X_k^{N,a})$, according to set of equations with mean–field interaction

$$X_k^{i,a} = X_k^{i,f} + K_k(P_k^{N,f})\,(Y_k - H_k\, X_k^{i,f} - V_k^i) \qquad \text{with} \qquad V_k^i \sim \mathcal{N}(0, R_k)$$

with $m \times d$ gain matrix $K_k(P)$ defined by

$$K_k(P) = P\, H_k^*\, (H_k\, P\, H_k^* + R_k)^{-1}$$

for any $m \times m$ covariance matrix P

notice that i.i.d. random vectors $(V_k^1, \cdots, V_k^N)$ are *simulated* here, with same statistics as additive Gaussian noise V_k in original observation equation mean–field interaction : in view of

$$X_k^{i,a} = X_k^{i,f} + K_k(P_k^{N,f})\,(Y_k - H_k\, X_k^{i,f} - V_k^i)$$

each analysis element depends on whole forecast ensemble $(X_k^{1,f}, \cdots, X_k^{N,f})$ . . .

. . . but only through empirical probability distribution

$$\mu_k^{N,f} = \frac{1}{N} \sum_{i=1}^N \delta_{X_k^{i,f}}$$

of forecast elements, actually through empirical covariance matrix $P_k^{N,f}$

results in *dependent* analysis elements $(X_k^{1,a}, \cdots, X_k^{N,a})$

question : does the empirical probability distribution of the ensemble elements converge to the Bayesian filter, defined as

$$\mu_k^-(dx) = \mathbb{P}[X_k \in dx \mid Y_{0:k-1}] \qquad \text{and} \qquad \mu_k(dx) = \mathbb{P}[X_k \in dx \mid Y_{0:k}]$$

i.e. does

$$\mu_k^{N,f} = \frac{1}{N} \sum_{i=1}^N \delta_{X_k^{i,f}} \longrightarrow \mu_k^- \qquad \text{and} \qquad \mu_k^{N,a} = \frac{1}{N} \sum_{i=1}^N \delta_{X_k^{i,a}} \longrightarrow \mu_k$$

hold in some sense, as $N \uparrow \infty$?

answer is NO

propagation of chaos approach : to study asymptotic behaviour of empirical probability distributions

$$\mu_k^{N,f} = \frac{1}{N} \sum_{i=1}^{N} \delta_{X_k^{i,f}} \qquad \text{and} \qquad \mu_k^{N,a} = \frac{1}{N} \sum_{i=1}^{N} \delta_{X_k^{i,a}}$$

of forecast elements and analysis elements, respectively, approximating i.i.d. random vectors are introduced

in practice, these vectors are propagated independently according to following set of decoupled equations

$$\bar{X}_k^{i,f} = f_k(\bar{X}_{k-1}^{i,a}) + W_k^i \qquad \text{with} \qquad W_k^i \sim \mathcal{N}(0, Q_k)$$

and

$$\bar{X}_k^{i,a} = \bar{X}_k^{i,f} + K_k(\bar{P}_k^f)\,(Y_k - H_k\,\bar{X}_k^{i,f} - V_k^i) \qquad \text{with} \qquad V_k^i \sim \mathcal{N}(0, R_k)$$

where $\bar{P}_k^f$ denotes covariance matrix of any random vector $\bar{X}_k^{i,f}$

initially $\bar{X}_0^{i,f} = X_0^{i,f}$, i.e. initial set of i.i.d. random vectors coincides exactly with initial ensemble by definition

$$\bar{m}_k^f = \mathbb{E}[\bar{X}_k^{i,f}] \qquad \text{and} \qquad \bar{P}_k^f = \mathbb{E}[(\bar{X}_k^{i,f} - \bar{m}_k^f)\,(\bar{X}_k^{i,f} - \bar{m}_k^f)^*]$$

respectively

empirical mean vector and covariance matrix of i.i.d. random vectors $(\bar{X}_k^{1,f}, \cdots, \bar{X}_k^{N,f})$ are defined as

$$\bar{m}_k^{N,f} = \frac{1}{N} \sum_{i=1}^N \bar{X}_k^{i,f} \qquad \text{and} \qquad \bar{P}_k^{N,f} = \frac{1}{N} \sum_{i=1}^N (\bar{X}_k^{i,f} - \bar{m}_k^{N,f})\,(\bar{X}_k^{i,f} - \bar{m}_k^{N,f})^*$$

respectively

heuristics : these i.i.d. random vectors are close (contiguous) to elements in the ensemble Kalman filter, since they

- start from same initial values exactly
- use same i.i.d. random vectors $(W_k^1, \cdots, W_k^N)$ and $(V_k^1, \cdots, V_k^N)$ exactly

already *simulated* and used in the ensemble Kalman filter

pros / cons

+ large sample asymptotics is simple to analyse, because of independance

− unknown covariance matrix $\bar{P}_k^f$ in general, hence unknown approximating i.i.d. random vectors

in contrast, elements in the ensemble Kalman filter are dependent, because they all contribute to / use empirical covariance matrix $P_k^{N,f}$ which results in mean–field interaction

but in counterpart this empirical covariance matrix is readily computable, and so are elements in the ensemble Kalman filter

outline

- motivation : Kalman filter in large dimension
- EnKF as particle system with mean–field interaction
- identification of the limit
- large sample asymptotics of EnKF
- connection with particle filters
- conclusion and perspective

limiting probability distributions $\bar{\mu}_k^f$ and $\bar{\mu}_k^a$ completely characterized by integrals of arbitrary bounded measurable functions

$$\int_{\mathbb{R}^m} \phi(x')\, \bar{\mu}_k^f(dx') = \int_{\mathbb{R}^m} \int_{\mathbb{R}^m} \phi(f_k(x) + w)\, \bar{\mu}_{k-1}^a(dx)\, p_k^W(dw)$$

where $p_k^W(dw)$ is Gaussian probability distribution with zero mean vector and covariance matrix Q_k, i.e. probability distribution of random vector W_k^i

sufficient conditions on drift function f_k can be given, under which $\bar{\mu}_k^f$ has finite second order moments, in which case covariance matrix $\bar{P}_k^f$ is finite, and

$$\int_{\mathbb{R}^m} \phi(x')\, \bar{\mu}_k^a(dx') = \int_{\mathbb{R}^m} \int_{\mathbb{R}^d} \phi(x + K_k(\bar{P}_k^f)\, (Y_k - H_k\, x + v))\, \bar{\mu}_k^f(dx)\, q_k^V(v)\, dv$$

where $q_k^V(v)$ is Gaussian density with zero mean vector and invertible covariance matrix R_k, i.e. probability density of random vector V_k^i

initially $\bar{X}_0^{i,f} \sim \mu_0$, which means $\bar{\mu}_0^f = \mu_0$

on the other hand, Bayesian filter satisfies

$$\int_{\mathbb{R}^m} \phi(x')\, \mu_k^-(dx') = \int_{\mathbb{R}^m} \int_{\mathbb{R}^m} \phi(f_k(x) + w)\, \mu_{k-1}(dx)\, p_k^W(dw)$$

and

$$\int_{\mathbb{R}^m} \phi(x')\, \mu_k(dx') = \frac{\displaystyle\int_{\mathbb{R}^m} \phi(x')\, q_k^V(Y_k - H_k\, x')\, \mu_k^-(dx')}{\displaystyle\int_{\mathbb{R}^m} q_k^V(Y_k - H_k\, x')\, \mu_k^-(dx')}$$

initially $\mu_0^- = \mu_0$

it follows from above discussion that $\bar{\mu}_0^f = \mu_0^-$, and if $\bar{\mu}_{k-1}^a = \mu_{k-1}$ then necessarily $\bar{\mu}_k^f = \mu_k^-$

but in general $\bar{\mu}_k^f = \mu_k^-$ does not necessarily imply $\bar{\mu}_k^a = \mu_k$, which means that in general limiting probability distributions $\bar{\mu}_k^f$ and $\bar{\mu}_k^a$ do not coincide with probability distributions μ_k^- and μ_k defining Bayesian filter

however, in linear Gaussian case, (probability distributions defining) Bayesian filter coincide with (Gaussian distributions associated with) Kalman filter, i.e. probability distribution μ_k^- is Gaussian, with mean vector $\widehat{X}_k^-$ and covariance matrix P_k^-, and the probability distribution μ_k is Gaussian, with mean vector and covariance matrix

$$\widehat{X}_k = \widehat{X}_k^- + K_k \, (Y_k - H_k \, \widehat{X}_k^-) \qquad \text{and} \qquad P_k = (I - K_k \, H_k) \, P_k^-$$

respectively

if $\bar{\mu}^a_{k-1} = \mu_{k-1}$, then it follows from general case that $\bar{\mu}^f_k = \mu^-_k$, and in particular $\bar{m}^f_k = \widehat{X}^-_k$ and $\bar{P}^f_k = P^-_k$

if $\bar{\mu}^f_k = \mu^-_k$, then in particular $\bar{P}^f_k = P^-_k$ hence $K_k(\bar{P}^f_k) = K_k$, and by definition $\bar{X}^{i,f}_k$ is Gaussian random vector with mean $\widehat{X}^-_k$ and covariance matrix P^-_k

since V^i_k is another independent Gaussian random vector with zero mean and covariance matrix R_k, then

$$\bar{X}^{i,a}_k = \bar{X}^{i,f}_k + K_k(\bar{P}^f_k)\,(Y_k - H_k\,\bar{X}^{i,f}_k - V^i_k)$$

is Gaussian random vector with mean

$$\bar{m}^a_k = \widehat{X}^-_k + K_k\,(Y_k - H_k\,\widehat{X}^-_k) = \widehat{X}_k$$

and covariance matrix

$$\bar{P}^a_k = (I - K_k\,H_k)\,P^-_k\,(I - K_k\,H_k)^* + K_k\,R_k\,K^*_k = (I - K_k\,H_k)\,P^-_k = P_k$$

which means that probability distribution of $\bar{X}^{i,a}_k$ is μ_k, or in other words $\bar{\mu}^a_k = \mu_k$

Assumption A globally Lipschitz continuous drift function, i.e.

$$|f_k(x) - f_k(x')| \leq L\, |x - x'|$$

for any $x, x' \in \mathbb{R}^m$

Assumption B locally Lipschitz continuous drift function, with at most polynomial growth at infinity, i.e.

$$|f_k(x) - f_k(x')| \leq L\, |x - x'|\, (1 + |x|^s + |x'|^s\,)$$

for any $x, x' \in \mathbb{R}^m$ and for some $s \geq 0$

under Assumption A, drift function has at most linear growth at infinity, i.e.

$$|f_k(x)| \leq M\, (1 + |x|\,)$$

under Assumption B, drift function has at most polynomial growth at infinity, i.e.

$$|f_k(x)| \leq M\, (1 + |x|^{s+1}\,)$$

for any $x \in \mathbb{R}^m$

a priori estimates (existence of moments)

Proposition if Assumption A holds, and if random vector X_0 has finite moment of order p, for some $p \geq 2$, then random vectors $\bar{X}_k^{i,f}$ and $\bar{X}_k^{i,a}$ have finite moments of the same order p, and in particular covariance matrix $\bar{P}_k^f$ is finite

if Assumption B holds, and if random vector X_0 has finite moments of any order, then random vectors $\bar{X}_k^{i,f}$ and $\bar{X}_k^{i,a}$ have finite moments of any order, and in particular covariance matrix $\bar{P}_k^f$ is finite

outline

- motivation : Kalman filter in large dimension
- EnKF as particle system with mean–field interaction
- identification of the limit
- large sample asymptotics of EnKF
- connection with particle filters
- conclusion and perspective

empirical covariance matrix $P_k^{N,f}$ of EnKF forecast elements
vs. covariance matrix $\bar{P}_k^f$ of limiting i.i.d. sequence

contiguity of empirical covariance matrices

$$\| P_k^{N,f} - \bar{P}_k^{N,f} \| \leq 2\, |\Delta_k^{N,2,f}|^2 + C\, \Delta_k^{N,2,f}$$

where $C > 0$ depends on (existing) finite moments of limiting sequence, and

$$\Delta_k^{N,2,f} = (\frac{1}{N} \sum_{i=1}^{N} |X_k^{i,f} - \bar{X}_k^{i,f}|^2)^{1/2}$$

by definition

consistency of empirical covariance matrice for limiting sequence

since $(\bar{X}_k^{1,f}, \ldots, \bar{X}_k^{N,f})$ are independent random variables, then

$$\| \bar{P}_k^{N,f} - \bar{P}_k^f \| \longrightarrow 0$$

almost surely, as $N \uparrow \infty$ by the law of large numbers

almost sure contiguity of ensemble elements

introduce

$$\Delta_k^{N,p,\bullet} = (\frac{1}{N} \sum_{i=1}^{N} |X_k^{i,\bullet} - \bar{X}_k^{i,\bullet}|^p)^{1/p}$$

Proposition if Assumption A holds, and if the random vector X_0 has finite moment of order p for some $p \geq 2$, then

$$\Delta_k^{N,p,\bullet} \longrightarrow 0$$

for the same order p, almost surely as $N \uparrow \infty$

if Assumption B holds, and if the random vector X_0 has finite moments of any order, then

$$\Delta_k^{N,p,\bullet} \longrightarrow 0$$

for any order p, almost surely as $N \uparrow \infty$

$\mathbb{L}^p$–contiguity of ensemble elements

recall

$$\Delta_k^{N,p,\bullet} = \Big(\frac{1}{N} \sum_{i=1}^{N} |X_k^{i,\bullet} - \bar{X}_k^{i,\bullet}|^p \Big)^{1/p}$$

and introduce

$$D_k^{N,p,\bullet} = \big(\mathbb{E}|X_k^{i,\bullet} - \bar{X}_k^{i,\bullet}|^p \big)^{1/p}$$

Lemma

$$D_k^{N,p \wedge q,\bullet} \leq \big(\mathbb{E}|\Delta_k^{N,p,\bullet}|^q \big)^{1/q} \leq D_k^{N,p \vee q,\bullet}$$

Proposition if Assumption B holds, and if random vector X_0 has finite moments of any order, then

$$\sup_{N \geq 1} \sqrt{N}\, D_k^{N,p,\bullet} < \infty$$

for any order p

almost sure convergence

Theorem let ϕ be a locally Lipschitz continuous function, with at most polynomial growth at infinity, i.e.

$$|\phi(x) - \phi(x')| \leq L\, |x - x'|\, (1 + |x|^\sigma + |x'|^\sigma)$$

for any $x, x' \in \mathbb{R}^m$ and for some $\sigma \geq 0$

if Assumption A holds, and if the random vector X_0 has finite moment of order p for some $p \geq 2$, then

$$\frac{1}{N} \sum_{i=1}^N \phi(X_k^{i,\bullet}) \longrightarrow \int_{\mathbb{R}^m} \phi(x)\, \bar\mu_k^\bullet(dx)$$

for the same order p, almost surely as $N \uparrow \infty$

if Assumption B holds, and if the random vector X_0 has finite moments of any order, then

$$\frac{1}{N} \sum_{i=1}^N \phi(X_k^{i,\bullet}) \longrightarrow \int_{\mathbb{R}^m} \phi(x)\, \bar\mu_k^\bullet(dx)$$

for any order p, almost surely as $N \uparrow \infty$

$\mathbb{L}^p$–convergence and rate of convergence

Theorem let ϕ be a locally Lipschitz continuous function, with at most polynomial growth at infinity, i.e.

$$|\phi(x) - \phi(x')| \leq L\ |x - x'|\ (1 + |x|^\sigma + |x'|^\sigma)$$

for any $x, x' \in \mathbb{R}^m$ and for some $\sigma \geq 0$

if Assumption B holds, and if the random vector X_0 has finite moments of any order, then

$$\sup_{N \geq 1} \sqrt{N}\ (\,\mathbb{E}|\,\frac{1}{N}\sum_{i=1}^N \phi(X_k^{i,\bullet}) - \int_{\mathbb{R}^m} \phi(x)\ \bar{\mu}_k^\bullet(dx)\,|^p\,)^{1/p} < \infty$$

for any order p

outline

- motivation : Kalman filter in large dimension
- EnKF as particle system with mean–field interaction
- identification of the limit
- large sample asymptotics of EnKF
- connection with particle filters
- conclusion and perspective

return to any system of the form

$$X_k = f_k(X_{k-1}) + W_k \qquad \text{with} \qquad W_k \sim \mathcal{N}(0, Q_k)$$

$$Y_k = H_k \, X_k + V_k \qquad \text{with} \qquad V_k \sim \mathcal{N}(0, R_k)$$

with non–necessarily Gaussian initial condition $X_0 \sim \mu_0$

approximation of the Bayesian filter

$$\mu_k(dx) = \mathbb{P}[X_k \in dx \mid Y_{0:k}]$$

in the form of a weighted empirical probability distribution

$$\mu_k \approx \mu_k^N = \sum_{i=1}^N w_k^i \, \delta_{\xi_k^i} \qquad \text{with} \qquad \sum_{i=1}^N w_k^i = 1$$

associateds with a system of N particles, characterized by their

- positions $(\xi_k^1, \cdots, \xi_k^N)$
- and weights $(w_k^1, \cdots, w_k^N)$

a particle approximation with optimal importance distribution

- mutation step : independently for any $i = 1, \ldots, N$

$$\xi_k^{i,-} = f_k(\xi_{k-1}^i) + W_k^i \qquad \text{with} \qquad W_k^i \sim \mathcal{N}(0, Q_k)$$

and

$$\xi_k^i = \xi_k^{i,-} + \Gamma_k \, (Y_k - H_k \, \xi_k^{i,-} - V_k^i) \qquad \text{with} \qquad V_k^i \sim \mathcal{N}(0, R_k)$$

with gain matrix

$$\Gamma_k = K_k(Q_k) = Q_k \, H_k^* \, (H_k \, Q_k \, H_k^* + R_k)^{-1}$$

so that, conditionnally w.r.t. ξ_{k-1}^i

$$\xi_k^i = (I - \Gamma_k \, H_k) \, (f_k(\xi_{k-1}^i) + W_k^i) + \Gamma_k \, (Y_k - V_k^i)$$

is Gaussian with mean vector

$$(I - \Gamma_k) \, H_k) \, f_k(\xi_{k-1}^i) + \Gamma_k \, Y_k = f_k(\xi_{k-1}^i) + \Gamma_k \, (Y_k - H_k \, f_k(\xi_{k-1}^i))$$

and covariance matrix

$$(I - \Gamma_k \, H_k) \, Q_k \, (I - \Gamma_k \, H_k)^* + \Gamma_k \, R_k \, \Gamma_k^* = (I - \Gamma_k \, H_k) \, Q_k$$

▶ weighting step : independently for any $i = 1, \ldots, N$

$$w_k^i \propto q(Y_k - H_k\, f_k(\xi_{k-1}^i), H_k\, Q_k\, H_k^* + R_k)$$

where $q(x, \Sigma)$ is Gaussian density with zero mean vector and invertible covariance matrix Σ

▶ selection step : discard / multiply particles according to their (relative) weights (many variants) many convergence results hold as population size N goes to infinity, with Bayesian filter μ_k as the limit, for this particular and for many other particle filters

Theorem convergence in $\mathbb{L}^p$–mean

$$(\mathbb{E}| \sum_{i=1}^N w_k^i \, \phi(\xi_k^i) - \int_{\mathbb{R}^m} \phi(x) \, \mu_k(dx) \, |^p \,)^{1/p} \longrightarrow 0$$

for any order p, as $N \uparrow \infty$

Theorem central limit theorem

$$\sqrt{N} \, (\sum_{i=1}^N w_k^i \, \phi(\xi_k^i) - \int_{\mathbb{R}^m} \phi(x) \, \mu_k(dx) \,) \Longrightarrow \mathcal{N}(0, v(\phi))$$

in distribution as $N \uparrow \infty$, with (more or less explicit) expression for the asymptotic variance $v(\phi)$

outline

- motivation : Kalman filter in large dimension
- EnKF as particle system with mean–field interaction
- identification of the limit
- large sample asymptotics of EnKF
- connection with particle filters
- conclusion and perspective

in his PhD thesis, Nicolas Papadakis has proposed a weighted version of the EnKF, where the evolution of the ensemble elements is seen as a special case of a mutation step, and appropriate weights are introduced

issues

- convergence and CLT for WEnKF, as ensemble size goes to infinity
- comparison with particle filters, on the basis of asymptotic variances

workprogramme of national project submitted to the ANR by Olivier Talagrand