ISSN: 2581-3064
sjrmcseditor@scischolars.com Online Publication Date: January 20, 2018 Volume 2, No. 2
SCHOLARS SCITECH RESEARCH ORGANIZATION
Scholars Journal of Research in Mathematics and Computer Science
www.scischolars.com

Iterated Extended Recursive Wiener Fixed-Point Smoothing and Filtering Algorithms in Discrete-Time Stochastic Systems with Nonlinear Observation Mechanism

Seiichi Nakamori

Department of Technology, Faculty of Education, Kagoshima University, 1-20-6, Korimoto, 890-0065 Japan.

k6165161@kadai.jp.

Abstract

This paper, by maximizing the conditional probability density function of the state vector given the observed values, presents discrete-time suboptimal maximum a posteriori (MAP) estimators for the nonlinear observation equation with additive white Gaussian observation noise. This paper, at first, proposes two kinds of iterated recursive Wiener extended algorithms for the fixed-point smoothing and filtering estimates. One is regarded as the Newton-Raphson algorithm, since it uses the first derivative in the Hessian with respect to the state vector, of the nonlinear observation function, and the other as the Newton's algorithm which uses up to the second derivative in the Hessian. Secondly, this paper proposes three kinds of extended recursive Wiener fixed-point smoothing and filtering algorithms in accordance with the medium-scale Quasi-Newton line search method, the Quasi-Newton (the limited memory Broyden-Fletcher-Goldfarb-Shanno (L-BFGS)) method and the Nelder-Mead simplex direct search method, in the relation with the MATLAB optimization toolbox respectively.

Keywords: Discrete-Time Stochastic Systems; Iterated Extended Recursive Wiener Estimators; Filter; Fixed-Point Smoother; Nonlinear Observation Mechanism.

1. Introduction

In nonlinear discrete-time stochastic systems, from the viewpoint of the maximum a posteriori (MAP) estimator, given the observed values, the conditional probability density function of the state vector is treated [1]. Maximization of the conditional probability density function with respect to the state vector is transformed into the minimization of the function (6) in [1]. Hence, the unconstrained optimization techniques can solve the minimization problem of the function. The simulation results in [1] show that the iterated extended Kalman filter (IEKF) is superior in estimation accuracy to the extended Kalman filter (EKF) and the unscented Kalman filter (UKF).

As for the unconstrained minimization of the objective function (6) in [1], there might be some useful programs such as "fminunc.m" and "fminsearch.m" in the optimization toolbox of MATLAB or "fmins.m" in Octave. In [2], it is indicated that the posterior mean and the mean by the unscented transformation agree to the third order. In [3], the unconstrained and constrained nonlinear optimization techniques are described.

In [4], the extended recursive Wiener fixed-point smoother and filter are proposed in discrete-time stochastic systems with nonlinear observation mechanism. The extended recursive Wiener estimators use the auto-covariance function of the state vector, expressed by the system matrix Φ, the observation vector C for the state vector $x(k)$, the variance $K(k,k)=K(0)$ of the state vector and the gradient of the nonlinear observation function with respect to the state vector. The recursive Wiener estimators do not use the information of the input matrix and the input noise variance in the state equation unlike the Kalman filter based on the state-space models. Hence, the recursive Wiener estimators are not influenced by the modelling errors of the input matrix and the input noise variance in the state equation. From the simulation results in [4], the extended recursive Wiener fixed-point smoother and filter are superior in estimation accuracy to the extended Kalman estimators. Theorem 1 shows the extended recursive Wiener fixed-point smoothing and filtering algorithms [4].

From the above viewpoints on the estimation accuracies of the iterated extended Kalman filter [1] and the extended recursive Wiener estimators [4], compared with the extended Kalman estimators, this paper, based on the minimization of(6) in [1], newly proposes discrete-time suboptimal MAP estimators for the nonlinear observation equation with additive white Gaussian observation noise. This paper, at first, proposes the iterated extended recursive Wiener algorithms for the fixed-point smoothing and filtering estimates. One is regarded as the Newton-Raphson algorithm, since it uses the first derivative in the Hessian with respect to the state vector, of the nonlinear observation function, and the other as the Newton's algorithm which uses up to the second derivative in the Hessian. Secondly, this paper proposes three kinds of iterated extended fixed-point smoothing and filtering algorithms in accordance with the medium-scale Quasi-Newton line search method, the Quasi-Newton (the limited memory Broyden-Fletcher-Goldfarb-Shanno (L-BFGS)) method and the Nelder-Mead simplex direct search method, in the relation with the MATLAB optimization toolbox, respectively.

The phase demodulation problem to estimate the signal, from the output that a phase was modulated by the signal, is important in the analog and digital communication systems [5]. In a numerical simulation example concerned with the phase demodulation, the following algorithms are compared.

1) Extended recursive Wiener fixed-point smoothing and filtering algorithms [4].
2) Iterated extended recursive Wiener fixed-point smoothing and filtering algorithms in the case of using the first derivative of the observation function with respect to the state vector in the Hessian.
3) Iterated extended recursive Wiener fixed-point smoothing and filtering algorithms in the case of using up to the second derivative of the observation function with respect to the state vector in the Hessian.
4) Iterated extended recursive Wiener fixed-point smoothing and filtering algorithms obtained based on the iterated extended Kalman filter in [6]. The algorithms are omitted in this paper.
5) Extended recursive Wiener fixed-point smoothing and filtering algorithms in accordance with the medium-scale Quasi- Newton line search method.
6) Extended recursive Wiener fixed-point smoothing and filtering algorithms in accordance with the Quasi-Newton (the L-BFGS) method.
7) Extended recursive Wiener fixed-point smoothing and filtering algorithms in accordance with the Nelder-Mead simplex direct search method.

2. Extended Recursive Wiener Estimation Algorithms for Nonlinear Observation Equation

Now, let a scalar nonlinear observation equation be given by

$$y(k) = f(z(k),k) + v(k), \quad z(k) = Cx(k), \tag{1}$$

where $z(k)$ represents the scalar signal, $x(k)$ the $n\times 1$ state vector and $v(k)$ the white Gaussian observation noise. Let the state equation for the state vector $x(k)$ be expressed by

$$x(k+1) = \Phi x(k) + w(k), \tag{2}$$

where Φ is the state-transition matrix and $w(k)$ white input noise. It is assumed that the signal and observation noise processes are mutually independent and zero-mean. Let the auto-covariance function of the observation noise $v(k)$ and the input noise $w(k)$ be expressed by

$$E[v(k)v(j)] = R\delta_K(k-j), \quad R>0 \tag{3}$$

$$E[w(k)w(j)] = Q\delta_K(k-j), \quad Q>0. \tag{4}$$

Here, $\delta_K(k-j)$ represents the Kronecker δ function.

Let $K(k,j) = K(k-j)$ represent the auto-covariance function of the state vector $x(k)$ and let $K(k,j)$ be expressed in the form of

$$K(k,j)=\begin{cases}A(k)B^T(j), 0\le j\le k\\ B(k)A^T(j), 0\le k\le j\end{cases} \quad (5)$$

in wide-sense stationary stochastic systems [7]. Here, $A(k)=\Phi^k$, $B^T(j)=\Phi^{-j}K(j,j)=\Phi^{-j}K(0)$. $K(0)$ represents the variance of the state vector $x(k)$.

Theorem 1 presents the extended recursive Wiener fixed-point smoothing and filtering algorithms [4] for the observation equation, including the nonlinear mechanism of the signal $z(k)$. It is noted that the proposed extended recursive Wiener estimators are sub-optimal because of the Taylor series approximation of the nonlinear observation function.

Theorem 1 [4]

Let the observation equation, which has the nonlinear mechanism of the signal $z(k)$, be given by (1) for the white Gaussian observation noise. Let the auto-covariance function of the state vector $x(k)$ be expressed by (5) and let the variance of white observation noise be R in wide-sense stationary stochastic systems. Then, the extended recursive Wiener fixed-point smoothing and filtering algorithms, using the covariance information of the signal and observation noise, consist of (6)-(16).

Fixed-point smoothing estimate of the signal $z(k)$ at the fixed point k: $\hat{z}(k|L)$

$$\hat{z}(k|L)=C\hat{x}(k|L) \quad (6)$$

Fixed-point smoothing estimate of the state vector $x(k)$ at the fixed point k: $\hat{x}(k|L)$

$$\hat{x}(k|L)=\hat{x}(k|L-1)+h(k,L,L)(y(L)-f(\hat{z}(L|L-1),L)) \quad (7)$$

Smoother gain: $h(k,L,L)$

$$\begin{aligned}&h(k,L,L)=(K(0)(\Phi^T)^{L-k}C^TH^T(L)-q(k|L-1)\Phi^TC^TH^T(L)\\&/(R+H(L)CK(L,L)C^TH^T(L)-H(L)C\Phi S(L-1)\Phi^TC^TH^T(L))\end{aligned} \quad (8)$$

$$\begin{aligned}&q(k|L)=q(k|L-1)\Phi^T+h(k,L,L)H(L)C(K(L,L)-\Phi S(L-1)\Phi^T),\\&q(k|k)=S(k)\end{aligned} \quad (9)$$

Filtering estimate of the signal $z(k)$: $\hat{z}(k|k)$

$$\hat{z}(k|k)=C\hat{x}(k,k) \quad (10)$$

Filtering estimate of the state vector $x(k)$: $\hat{x}(k|k)$

$$\hat{x}(k|k)=\Phi\hat{x}(k-1|k-1)+G(k)(y(k)-f(\hat{z}(k|k-1),k)), \hat{x}(0|0)=0 \quad (11)$$

One-step ahead prediction estimate of the signal $z(k)$: $\hat{z}(k|k-1)$

$$\hat{z}(k|k-1)=H(k)C\hat{x}(k,k-1) \quad (12)$$

One-step ahead prediction estimate of the state vector $x(k)$: $\hat{x}(k|k-1)$

$$\hat{x}(k|k-1)=\Phi\hat{x}(k-1|k-1) \quad (13)$$

Auto-variance function of $\hat{x}(k|k)$: $S(k)$

$$S(k)=\Phi S(k-1)\Phi^T+G(k)H(k)C(K(k,k)-\Phi S(k-1)\Phi^T), S(0)=0 \quad (14)$$

Filter gain: $G(k)$

$$\begin{aligned}&G(k)=(K(k,k)C^TH^T(k)-\Phi S(k-1)\Phi^TC^TH^T(k)/\\&(R+H(k)CK(k,k)C^TH^T(k)-H(k)C\Phi S(k-1)\Phi^TC^TH^T(k))\end{aligned} \quad (15)$$

Here, the function $H(k)$ is given by

$$H(k) = \frac{\partial f(Cx(k),k)}{\partial x(k)}\Big|_{x(k)=\hat{x}(k|k-1)}. \quad (16)$$

A necessary condition for the stability of the extended recursive Wiener estimators is given by $R + H(k)CK(k,k)C^T H^T(k) - H(k)C\Phi S(k-1)\Phi^T C^T H^T(k) > 0$.

Like the design of the extended Kalman filter, in the design of the extended recursive Wiener estimators, the function $H(k)$ is put as $H(k) = \frac{\partial f(Cx(k),k)}{\partial x(k)}\Big|_{x(k)=\hat{x}(k|k-1)}$ expanding the nonlinear observation function in a first-order Taylor series about $\hat{x}(k|k-1)$ [4]. Here, $\hat{z}(k|k-1) = C\Phi\hat{x}(k-1,k-1)$ represents the one-step ahead prediction estimate of the signal $z(k)$.

The difference of the extended recursive Wiener estimators from the extended Kalman estimators lies in the information used. The extended recursive Wiener estimators use the information of Φ, C, $K(0)$ and R. The extended Kalman estimators use the information of Φ, C and the variance of the white noise input, Q, in (2). Both estimators use the information of nonlinear observation function and the function $H(k)$. Since $S(k)$ is the auto-variance function of the filtering estimate $\hat{x}(k,k)$, the filtering error variance function $P(k|k)$ is given by $P(k|k) = K(0) - S(k)$. The variance function of the one-step ahead prediction estimate of x(k) is given by $\Phi S(k-1)\Phi^T$. One-step ahead prediction error variance function of x(k)is given by $P(k|k-1) = K(k,k) - \Phi S(k-1)\Phi^T$.

The simulation result in [4] shows that the extended recursive Wiener filter and fixed-point smoother are superior in estimation accuracy to the extended Kalman estimators respectively.

3. Iterated Extended Recursive Wiener Fixed-Point Smoothing and Filtering Algorithms

This section, likewise the derivation technique of the iterated extended Kalman filter [1], at first, proposes the iterated extended recursive Wiener fixed-point smoothing and filtering algorithms.

On a Bayesian analysis for the nonlinear Gaussian filtering problem, the probability density functions hold the following relationships [1].

$$p(x(k-1|y(k-1)) \approx N(x(k-1):\hat{x}(k-1|k-1), P(k-1|k-1)),$$

$$p(x(k|y(k-1)) \approx N(x(k):\hat{x}(k|k-1), P(k|k-1)),$$

$$p(x(k|y(k)) \approx N(x(k):\hat{x}(k|k), P(k|k)),$$

Where $p(x(k-1|y(k-1))$ is the initial prior density, $p(x(k|y(k-1))$ is the prediction density, $p(x(k|y(k))$ is the posterior density. Here, $Y_k = \{y(i)\}_{i=1}^{k}$ are the measurements. The conditional probability density function of $x(k+1)$ given Y_{k+1} is obtained, under the Gaussian assumption for all the related random variables, as follows.

$$p[x(k+1)|Y_{k+1}] = p[x(k+1)|y(k+1), Y_{k+1}]$$

$$= \frac{1}{b} p[y(k+1)|x(k+1)]p[x(k+1)|Y_k]$$

$$= \frac{1}{b} N[y(k+1): f(z(k+1),k+1), R]\cdot N[x(k+1):\hat{x}(k+1|k), P(k+1|k)],$$

Here, b is a constant value. Maximizing the conditional probability density function $p[x(k+1)|Y_{k+1}]$ with respect to $x(k+1)$ is equivalent to minimizing $J(x(k+1))$ with respect to $x(k+1)$ [1]. $J(x(k+1))$ is given by

$$J(x(k+1)) = \frac{1}{2}[y(k+1) - f(Cx(k+1),k+1)]^T R^{-1}[y(k+1) - f(Cx(k+1),k+1)]$$

$$+\frac{1}{2}[x(k+1) - \hat{x}(k+1|k)]^T P(k+1|k)^{-1}[x(k+1) - \hat{x}(k+1|k)] \quad (17)$$

$$= \frac{1}{2} r^T(k+1)r(k+1).$$

Here,

$$r(k+1)=\begin{bmatrix} R^{-\frac{1}{2}}(y(k+1)-f(Cx(k+1),k+1)) \\ P(k+1\mid k)^{-\frac{1}{2}}(x(k+1)-\hat{x}(k+1\mid k)) \end{bmatrix}.$$

In (17) $P(k+1|k)$ represents the prediction error variance function of $x(k+1)$. In the extended recursive Wiener filter [4], $S(k)$ represents the variance of the filtering estimate $\hat{x}(k,k)$ of the state vector $x(k)$. In terms of $S(k)$, the prediction error variance function $P(k+1|k)$ is given by $P(k+1|k)=K(k+1,k+1)-\Phi S(k)\Phi^T$, $K(k+1,k+1)=K(0)$, in wide-sense stationary stochastic systems.

In the Newton's method, the gradient vector and the Hessian matrix are required. It is necessary that the Hessian matrix $\nabla_x\nabla_x^T J[x(k+1)]|_{x(k+1)=\hat{x}^i(k+1|k+1)}$ is positive-definite since the $i+1$th filtering estimate $\hat{x}^{i+1}(k+1|k+1)$ is updated from the ith filtering estimate $\hat{x}^i(k+1|k+1)$ by

$$\hat{x}^{i+1}(k+1|k+1)=\hat{x}^i(k+1|k+1)+(\nabla_x\nabla_x^T J(x(k+1))|_{x(k+1)=\hat{x}^i(k+1|k+1)})^{-1}\nabla_x J(x(k+1))|_{x(k+1)=\hat{x}^i(k+1|k+1)}.$$

In the Newton's method, if the ith value $\hat{x}^i(k+1|k+1)$ is near to the true value $x(k+1)$, the error between the $i+1$th value $\hat{x}^{i+1}(k+1|k+1)$ and the true value is almost equal to the value, obtained by multiplying a constant to the square value of the error, provided that $\nabla_x\nabla_x^T J(x(k+1))|_{x(k+1)=\hat{x}^i(k+1|k+1)}>0$ [8].

Now, let us expand $J[x(k+1)]$ in a Taylor series up to the second order about the ith iterated value of the estimate of $x(k+1)$.

$$\begin{aligned} J(x(k+1)) &= J[x^i(k+1)]+\nabla_x J(x(k+1))|_{x(k+1)=x^i(k+1)}(x(k+1)-x^i(k+1)) \\ &+\frac{1}{2}(x(k+1)-x^i(k+1))^T\nabla_x\nabla_x^T J(x(k+1))|_{x(k+1)=x^i(k+1)}(x(k+1)-x^i(k+1)) \end{aligned} \tag{18}$$

Here, $\nabla_x J(x(k+1))|_{x(k+1)=x^i(k+1)}$ and $\nabla_x\nabla_x^T J(x(k+1))|_{x(k+1)=x^i(k+1)}$ represent the gradient vector and the Hessian matrix, with respect to $x^i(k+1)$, respectively. By setting the gradient of (18), with respect to $x(k+1)$, to the zero vector, in accordance with the minimization of $J(x(k+1))$, we have an iterated equation of Newton's method for the $i+1$ th filtering estimate $\hat{x}^{i+1}(k+1|k+1)$ of the state vector $\mathrm{x}(\mathrm{k}+1)$ from the ith filtering estimate $\hat{x}^i(k+1|k+1)$ as follows.

$$\hat{x}^{i+1}(k+1|k+1)=\hat{x}^i(k+1|k+1)-[\nabla_x\nabla_x^T \mathrm{J}(\mathrm{x}(\mathrm{k}+1))|_{x(k+1)=\hat{x}^i(k+1|k+1)}]^{-1}\nabla_x \mathrm{J}[\mathrm{x}(\mathrm{k}+1)]|_{x(k+1)=\hat{x}^i(k+1|k+1)} \tag{19}$$

$$\begin{aligned} \nabla_x \mathrm{J}(\mathrm{x}(\mathrm{k}+1))|_{x(k+1)=\hat{x}^i(k+1|k+1)} &= -f_x\left(C\hat{x}^i(k+1|k+1),k+1\right)R^{-1} \\ &\times\left(y(k+1)-f\left(C\hat{x}^i(k+1|k+1),k+1\right)\right)+P(k+1|k)^{-1}\left(\hat{x}^i(k+1|k+1)-\hat{x}(\mathrm{k}+1|\mathrm{k})\right)R^{-1} \end{aligned} \tag{20}$$

$$\nabla_x\nabla_x^T J(x(k+1))|_{x(k+1)=\hat{x}^i(k+1|k+1)}\cong f_x(C\hat{x}^i(k+1|k+1),k+1)^T R^{-1} f_x(C\hat{x}^i(k+1|k+1),k+1)+P(k+1|k)^{-1} \tag{21}$$

Here, the Hessian is calculated approximately to the first derivative of $J(x(k+1))$ with respect to $x(k+1)$ and (19) is obtained by putting $x(k+1)$ and $x^i(k+1)$ to $\hat{x}^{i+1}(k+1|k+1)$ and $\hat{x}^i(k+1|k+1)$ respectively as in [1]. From retaining to the first derivative in (21), the minimization of (18) results in the Newton-Raphson algorithm, which yields an approximate MAP filtering estimate $\hat{x}(k+1|k+1)$ of $x(k+1)$. In [1], it is indicated that $(\nabla_x\nabla_x^T J[x(k+1)])^{-1}=P(k+1|k+1)$. Here, $P(k+1|k+1)$ represents the filtering error variance function of $x(k+1)$. Since $P(k+1|k+1)$ is a positive semi-definite matrix, the filtering estimate might not diverge. From the simulation result in [1], the estimation accuracy of the iterated extended Kalman filter is superior to the EKF and the UKF.

Now, in the followings, let us show the iterated recursive Wiener filtering and fixed-point smoothing algorithms in accordance with the Newton-Raphson method and the Newton's method.

Initial values at discrete time $k+1$:

(1) S(1)=0

(2) Filtering estimate of the state vector x(1) : $\hat{x}(1|1)=0$

(3) Filtering estimate of the signal $z(1)$: $\hat{z}(1|1)=0$

Iteration of discrete time k from 1 to last time:

One-step ahead prediction estimate of the state vector $x(k+1)$: $\hat{x}(k+1|k)$

$$\hat{x}(k+1|k)=\Phi\hat{x}(k|k)$$

One-step ahead prediction estimate of the signal $z(k+1)$: $\hat{z}(k+1|k)$

$$\hat{z}(k+1|k)=C\hat{x}(k+1|k)$$

One-step ahead prediction error variance function of $x(k+1)$: $P(k+1|k)$

$$\mathrm{P}(k+1|k)=\mathrm{K}(k+1,k+1)-\Phi\mathrm{S}(k)\Phi^T$$

$$\rho_0^1(k)=\hat{x}(k|k)$$

$$\hat{x}^1(k|k)=\hat{x}(k|k)$$

Iteration of j from 1 to I:

One-step ahead prediction estimate $\hat{x}^j(k+1|k)$ of $x(k+1)$ in the j th iteration: $\hat{x}^j(k+1|k)=\Phi\hat{x}^j(k|k)$

Differentiate the observation function with respect to $x(k)$ and put $x(k)=\hat{x}^j(k+1|k)$.

$$H^j(k+1)=\frac{\partial f(Cx(k),k)}{\partial x(k)}\Big|_{x(k)=\hat{x}^j(k+1|k)}$$

Observed value: $y(k+1)$

Filter gain in the jth iteration: $G^j(k+1)$

$$G^j(k+1)=\left(K(k+1,k+1)C^TH^j(k+1)-\Phi\mathrm{S}(k)\Phi^TC^TH^j(k+1)\right)$$
$$\times[\mathrm{R}+H^j(k+1)\left(\mathrm{CK}(k+1,k+1)C^T-\mathrm{C}\Phi\mathrm{S}(k)\Phi^TC^T\right)H^j(k+1)^T]^{-1}$$

(1) Newton-Raphson algorithm

$$\rho_1^{j+1}=\rho_0^j+[(H^j(k+1))^TR^{-1}H^j(k+1)+(K(k+1,k+1)-\Phi\mathrm{S}(k)\Phi^T)^{-1}]^{-1}[-H^j(k+1)^TR^{-1}(y(k+1)-f(C\rho_0^j,k+1))$$
$$+(\mathrm{K}(k+1,k+1)-\Phi\mathrm{S}(k)\ \Phi^T)^{-1}\left(\rho_0^j-\Phi\hat{x}(k|k)\right)]$$
$$\rho_0^{j+1}=\rho_1^{j+1}$$

(2) Newton's algorithm:

$$\rho_1^{(j+1)}=\rho_0^j+[(H^j(k+1))^TR^{-1}H^j(k+1)+(K(k+1,k+1)-\Phi S(k)\Phi^T)^{-1}-(y(k+1)-f(C\rho_0^j,k+1))R^{-1}f_{xx}(C\rho_0^j,k+1)]^{-1}$$
$$\times[-H^j(k+1)^TR^{-1}(y(k+1)-f(C\rho_0^j,k+1))+(K(k+1,k+1)-\Phi S(k)\Phi^T)^{-1}(\rho_0^j-\Phi\hat{x}(k|k))]$$

$$\rho_0^{j+1}=\rho_1^{j+1}$$

End of the iteration for j

Filtering estimate of $\mathrm{x}(k+1)$: $\hat{x}(k+1|k+1)$

$$\hat{x}(k+1|k+1)=\rho_1^{I+1}$$

Filtering estimate of the signal $z(k+1)$: $\hat{z}(k+1|k+1)$

$$\hat{z}(k+1|k+1) = C\hat{x}(k+1|k+1)$$

Time update for the variance of the filtering estimate $\hat{x}(k+1|k+1)$ of $x(k+1)$: $S(k+1)$

$$S(k+1) = \Phi S(k)\Phi^T + G^I(k+1)H^I(k+1)C\left(K(k+1,k+1) - \Phi S(k)\Phi^T\right)$$

Filtering error variance function of $x(k+1)$: $P(k+1|k+1)$

$$P(k+1|k+1) = K(k+1,k+1) - S(k+1)$$

Initial value of the fixed-point smoothing estimate: $\hat{x}(k+1|k+1)$

Initial value of $q(k+1)$: $q(k+1) = S(k+1)$

Iteration of i from 1 to Lag(=L):

Iteration of j from 1 to I:

One-step ahead prediction estimate $\hat{x}(k+i+1|k+i)$ of $x(k+i)$ in the jth iteration:

$$\hat{x}^j(k+i+1|k+i) = \Phi\hat{x}^j(k+i|k+i)$$

Differentiate the observation function with respect to $x(k)$ and put $x(k) = \hat{x}^j(k+i+1|k+i)$.

$$H^j(k+i+1) = \frac{\partial f(Cx(k),k)}{\partial x(k)}\Big|_{x(k)=\hat{x}^j(k+i+1|k+i)}$$

Observed value: $y(k+i+1)$

Filter gain in the jth iteration: $G^j(k+i+1)$

$$G^j(k+i+1) = \left(K(k+i+1,k+i+1)C^T H^j(k+i+1) - \Phi S(k+i)\Phi^T C^T H^j(k+i+1)\right)$$
$$\times[R + H^j(k+i+1)\left(CK(k+i+1,k+i+1)C^T - C\Phi S(k+i)\Phi^T C^T\right)H^j(k+i+1)^T]^{-1}$$

(1) Quasi-Raphson algorithm:

$$\rho_1^{(j+1)} = \rho_0^j + [(H^j(k+i+1))^T R^{-1} H^j(k+i+1) + (K(k+i+1,k+i+1) - \Phi S(k+i)\Phi^T)^{-1}]^{-1}$$
$$\times[-H^j(k+i+1)^T R^{-1}(y(k+i+1) - f(C\rho_0^j, k+i+1)) + (K(k+i+1,k+i+1) - \Phi S(k+i)\Phi^T)^{-1}(\rho_0^j - \Phi\hat{x}(k+i|k+i))],$$

$$\rho_0^{j+1} = \rho_1^{j+1}$$

(2) Newton's algorithm:

$$\rho_1^{(j+1)} = \rho_0^j + [(H^j(k+i+1))^T R^{-1} H^j(k+i+1) + (K(k+i+1,k+i+1) - \Phi S(k+i)\Phi^T)^{-1}$$
$$-(y(k+i+1) - f(C\rho_0^j, k+i+1))R^{-1} f_{xx}(C\rho_0^j, k+i+1)]^{-1}[-H^j(k+i+1)^T R^{-1}(y(k+i+1) - f(C\rho_0^j, k+i+1))$$
$$+(K(k+i+1,k+i+1) - \Phi S(k+i)\ \Phi^T)^{-1}\left(\rho_0^j - \Phi\hat{x}(k+i|k+i)\right)]$$

$$\rho_0^{j+1} = \rho_1^{j+1}$$

End of the iteration for j

Filtering estimate of $x(k+i+1)$: $\hat{x}(k+i+1|k+i+1)$

$$\hat{x}(k+i+1|k+i+1) = \rho_1^{I+1}$$

Filtering estimate of the signal $z(k+i+1)$: $\hat{z}(k+i+1|k+i+1)$

$\hat{z}(k+i+1 \mid k+i+1) = C\hat{x}(k+i+1 \mid k+i+1)$

Time update for the variance $S(k+i+1)$ of the filtering estimate $\hat{x}(k+i+1 \mid k+i+1)$ of $x(k+i+1)$: $S(k+i+1)$

$$S(k+i+1) = \Phi S(k+i)\Phi^T + G^I(k+i+1)H^I(k+i+1)C(K(k+i+1,k+i+1) \\ -\Phi S(k+i)\Phi^T)$$

Filtering error variance function of $x(k+i+1)$: $P(k+i+1|k+i+1)$

$$P(k+i+1|k+i+1) = K(k+i+1,k+i+1) - S(k+i+1)$$

One-step ahead prediction estimate of the state vector $x(k+1+i)$: $\hat{x}(k+1+i \mid k+i)$

$$\hat{x}(k+1+i \mid k+i) = \Phi\hat{x}(k+i \mid k+i)$$

One-step ahead prediction estimate of the signal $z(k+1+i)$: $\hat{z}(k+1+i \mid k+i)$

$$\hat{z}(k+1+i \mid k+i) = C\hat{x}(k+1+i \mid k+i)$$

Smoother gain: $h(k+1,k+1+i,k+1+i)$

$$h(k+1,k+i+1,k+i+1) = (K(k+i,k+i)(\Phi^T)^i C^T (H^I)^T (k+i+1) - q(k \mid k+i)\Phi^T C^T (H^I)^T (k+1+i)) \\ \cdot(R + H^I(k+i+1)CK(k+i,k+i)C^T(H^I)^T(k+i+1) - H^I(k+i+1)C\Phi S(k+i \mid k+i)\Phi^T C^T (H^I)^T (k+i+1))^{-1}$$

$$q(k+i+1) = q(k+i)\Phi^T + h(k+1,k+i+1,k+i+1)H^I(k+i+1)C\left(K(k+i,k+i) - S(k+i)\Phi^T\right)$$

Fixed-point smoothing estimate of $x(k+1)$ at the fixed point k: $\hat{x}(k+1 \mid k+i+1)$

$$\hat{x}(k+1|k+i+1) = \hat{x}(k+1|k+i) + h(k+1,k+i+1,k+i+1)(y(k+i+1) \\ - f(C\hat{x}(k+i+1 \mid k+i+1), k+i+1))$$

End of the iteration for i

End of the iteration for k

By the way, in accordance with the minimization of the function J(x(k+1)) of (17), let us show the extended recursive Wiener filtering and fixed-point smoothing algorithms based on the medium-scale Quasi-Newton line search method, the Quasi-Newton (the L-BFGS) method and the Nelder-Mead simplex direct search method.

Method 1: In accordance with the minimization of the function J(x(k+1)) of (17), the medium-scale Quasi-Newton line search method is used. The finite difference method gets the gradient of $f(Cx(k),k)$ in terms of the MATLAB program "fminunc.m" in the optimization toolbox.

Method 2: In accordance with the minimization of the function J(x(k+1)) of (17), by providing the gradient of the function $f(Cx(k),k)$, the Quasi-Newton (the L-BFGS) method is applied in terms of the MATLAB program "fminunc.m" in the optimization toolbox.

Method 3: In accordance with the minimization of the function J(x(k+1)) of (17), the Nelder-Mead simplex direct search method is applied in terms of the MATLAB program "fminsearch.m" in the optimization toolbox.

Initial values at discrete time $k=1$:

(1) $S(1) = 0$

(2) Filtering estimate of the state vector $x(1)$: $\hat{x}(1 \mid 1) = 0$

(3) Filtering estimate of the signal $z(1)$: $\hat{z}(1 \mid 1) = 0$

Iteration of discrete time k from 1 to last time:

One-step ahead prediction estimate of the state vector $x(k+1)$: $\hat{x}(k+1|k)$

$$\hat{x}(k+1|k) = \Phi\hat{x}(k|k)$$

One-step ahead prediction estimate of the signal $z(k+1)$: $\hat{z}(k+1|k)$

$$\hat{z}(k+1|k) = C\hat{x}(k+1|k)$$

One-step ahead prediction error variance function of $x(k+1)$: $P(k+1|k)$

$$P(k+1|k) = K(k+1,k+1) - \Phi S(k)\Phi^T$$

Observed value: $y(k+1)$

Let the optimal value, which minimizes the function J(x(k+1)) of

$$J(x(k+1)) = \frac{1}{2}[y(k+1) - f(Cx(k+1),k+1)]^T R^{-1}\left[y(k+1) - f(Cx(k+1),k+1)\right]$$
$$+\frac{1}{2}[x(k+1) - \hat{x}(k+1|k)]^T P(k+1|k)^{-1}\left[x(k+1) - \hat{x}(k+1|k)\right]$$

with respect to $x(k+1)$, be $xopt(k+1)^T$ through the following three methods respectively.

Method 1: In accordance with the minimization of the function J(x(k+1)), the medium-scale Quasi-Newton line search method is used. The finite difference method is used to get the gradient of the function $f(Cx(k+1),k+1)$ in the MATLAB program "fminunc.m".

Method 2: In accordance with the minimization of the function J(x(k+1)), by providing the gradient of the function $f(Cx(k+1),k+1)$, the Quasi-Newton (the L-BFGS) method is used in the MATLAB program "fminunc.m".

Method 3: In accordance with the minimization of the function J(x(k+1)), the Nelder-Mead simplex direct search method is used in the MATLAB program "fminsearch.m".

Filtering estimate of $x(k+1)$: $\hat{x}(k+1|k+1)$

$$\hat{x}(k+1|k+1) = xopt(k+1)^T$$

Filtering estimate of the signal $z(k+1)$: $\hat{z}(k+1|k+1)$

$$\hat{z}(k+1|k+1) = C\hat{x}(k+1|k+1)$$

Differentiate the observation function with respect to $x(k)$ and put $x(k) = \hat{x}(k+1|k)$

$$H(k+1) = \frac{\partial f(Cx(k),k)}{\partial x(k)}\Big|_{x(k)=\hat{x}(k+1|k)}$$

Filter gain in the j th iteration: $G(k+1)$

$$G(k+1) = \left(K(k+1,k+1)C^T H(k+1) - \Phi S(k)\Phi^T C^T H(k+1)\right)$$
$$\times[R + H(k+1)\left(CK(k+1,k+1)C^T - C\Phi S(k)\Phi^T C^T\right)H(k+1)^T]^{-1}$$

Variance of the filtering estimate $\hat{x}(k+1|k+1)$ of $x(k+1)$: $S(k+1)$

$$S(k+1) = \Phi S(k)\Phi^T + G(k+1)H(k+1)C\left(K(k+1,k+1) - \Phi S(k)\Phi^T\right)$$

Filtering error variance function of $x(k+1)$: $P(k+1|k+1)$

$$P(k+1|k+1) = K(k+1,k+1) - S(k+1)$$

Initial value of the fixed-point smoothing estimate: $\hat{x}(k+1|k+1)$

Initial value of $q(k+1)$: $q(k+1) = S(k+1)$

Iteration of i from 1 to Lag(=L)

One-step ahead prediction estimate $\hat{x}(k+i+1|k+i)$ of $x(k+i)$: $\hat{x}(k+i+1|k+i)$

$$\hat{x}(k+i+1|k+i) = \Phi\hat{x}(k+i|k+i)$$

Observed value: $y(k+i+1)$

Let the optimal value, which minimizes the function $J(x(k+i+1))$ of

$$J(x(k+i+1)) = \frac{1}{2}[y(k+i+1) - f(Cx(k+i+1),k+i+1)]^T R^{-1}\left[y(k+i+1) - f(Cx(k+i+1),k+i+1)\right]$$
$$+\frac{1}{2}[x(k+i+1) - \hat{x}(k+i+1|k+i)]^T P(k+i+1|k+i)^{-1}\left[x(k+i+1) - \hat{x}(k+i+1|k+i)\right]$$

with respect to $x(k+i+1)$, be $xopt(k+i+1)^T$ through the following three methods respectively.

Method 1: In accordance with the minimization of the function $J(x(k+i+1))$, the medium-scale Quasi-Newton line search method is used. The finite difference method to get the gradient is used, without providing the gradient of the function $f(Cx(k+i+1),k+i+1)$, by the MATLAB program “fminunc.m”.

Method 2: In accordance with the minimization of the function $J(x(k+i+1))$, by providing the gradient of the function $f(Cx(k+i+1),k+i+1)$, the Quasi-Newton (the L-BFGS) method is used by the MATLAB program “fminunc.m”.

Method 3: In accordance with the minimization of the function $J(x(k+i+1))$, the Nelder-Mead simplex direct search method is used by the MATLAB program “fminsearch.m”.

Filtering estimate of $x(k+i+1)$: $\hat{x}(k+i+1|k+i+1)$

$$\hat{x}(k+i+1|k+i+1) = xopt(k+i+1)^T$$

Filtering estimate of the signal $z(k+i+1)$: $\hat{z}(k+i+1|k+i+1)$

$$\hat{z}(k+i+1|k+i+1) = C\hat{x}(k+i+1|k+i+1)$$

Differentiate the observation function with respect to $x(k)$ and put $x(k) = \hat{x}^j(k+i+1|k+i)$

$$H(k+i+1) = \frac{\partial f(Cx(k),k)}{\partial x(k)}\Big|_{x(k)=\hat{x}(k+i+1|k+i)}$$

Filter gain: $G(k+i+1)$

$$G(k+i+1) = \left(K(k+i+1,k+i+1)C^T H(k+i+1) - \Phi S(k+i)\Phi^T C^T H(k+i+1)\right)$$
$$\cdot[R + H(k+i+1)\left(CK(k+i+1,k+i+1)C^T - C\Phi S(k+i)\Phi^T C^T\right)H(k+i+1)^T]^{-1}$$

Variance of the filtering estimate $\hat{x}(k+i+1|k+i+1)$ of $x(k+i+1)$: $S(k+i+1)$

$$S(k+i+1)=\Phi S(k+i)\Phi^T+G(k+i+1)H(k+i+1)C\left(K(k+i+1,k+i+1)-\Phi S(k+i)\Phi^T\right)$$

Filtering error variance function of $x(k+i+1)$: $P(k+i+1|k+i+1)$

$$P(k+i+1|k+i+1)=K(k+i+1,k+i+1)-S(k+i+1)$$

Smoother gain: $h(k+1,k+i+1,k+i+1)$

$$h(k+1,k+i+1,k+i+1)=(K(k+i,k+i)(\Phi^T)^iC^TH^T(k+i+1)-q(k\,|\,k+i)\Phi^TC^TH^T(k+1+i))$$
$$\times(R+H(k+i+1)CK(k+i,k+i)C^TH^T(k+i+1)-H(k+i+1)C\Phi S(k+i\,|\,k+i)\Phi^TC^TH^T(k+i+1))^{-1}$$

$$q(k+i+1)=q(k+i)\Phi^T+h(k+1,k+i+1,k+i+1)H(k+i+1)C\left(K(k+i,k+i)-S(k+i)\Phi^T\right)$$

Fixed-point smoothing estimate of $x(k+1)$ at the fixed point k: $\hat{x}(k+1\,|\,k+i+1)$

$$\hat{x}(k+1|k+i+1)=\hat{x}(k+1|k+i)+h(k+1,k+i+1,k+i+1)(y(k+i+1)-f(C\hat{x}(k+i+1\,|\,k+i+1),k+i+1)))$$

End of the iteration for i

End of the iteration for k

Now, let us show a numerical simulation example in section 4.

4. A Numerical Simulation Example

Let a scalar observation equation with the nonlinear mechanism be given by

$$y(k)=f(z(k),k)+v(k),\ z(k)=Cx(k),$$

$$f(z(k),k)=\cos\left(2\pi f_ck\Delta+m_Az(k)\right),\ f_c=1{,}000\ \text{[Hz]},\ \Delta=0.000090703,\ m_A=1.2. \quad (22)$$

The nonlinear function in (22) shows the phase modulation in analogue communication systems [8]. Here, f_c, $\Delta$and m_A represent the carrier frequency, the sampling period of the signal $z(k)$ and the phase sensitivity respectively. Let $v(k)$ be the white Gaussian observation noise with the mean zero and the variance R, which is denoted by $N(0,R)$. The gradient of the nonlinear function $f(Cx(k),k)$ with respect to $x(k)$ is given by

$$H(k)=\frac{\partial f(x(k),k)}{\partial x(k)}|_{x(k)=\hat{x}(k|k-1)}=-m_AC\sin\left(2\pi f_ck\Delta+m_AC\hat{x}(k|k-1)\right),\hat{z}(k|k-1)=C\hat{x}(k|k-1), \quad (23)$$

where $x(k)$ is set to $\hat{x}(k\,|\,k-1)$.

Let the signal $z(k)$ be expressed by the state vector $x(k)$, which consists of the state variables as

$$x_1(k)=z(k),\ x_2(k)=z(k+1),\ \cdots,\ x_n(k)=z(k+n-1),$$

$$z(k)=Cx(k),\ x(k)=[x_1(k)\ \ x_2(k)\cdots\ \ x_n(k)]^T,\ z(k)=x_1(k),\ C=[1\quad 0\cdots\quad 0]. \quad (24)$$

Let us consider to estimate a vowel signal spoken by the author. Its phonetic symbol is written as "/i:/". The sampling frequency of the voice signal is 11.025 [kHz]. The auto-covariance data of the signal process are calculated in terms of the $N=350$ sampled data. Let the stochastic process of the vowel signal be modeled in terms of the autoregresive (AR) process of order $n=10$ as

$$z(k)=-a_1z(k-1)-a_2z(k-2)-\cdots a_nz(k-n)+e(k),\ E[e(k)e(s)]=\sigma^2\delta_K(k-s). \quad (25)$$

Let $K_z(i)$, $i=1,\cdots,n$, represent the auto-covariance data of the signal $z(k)$ in wide-sense stationary stochastic systems. The AR parameters a_i, $i=1,\cdots,n$, are calculated by the Yule-Walker equations.

$$\begin{bmatrix} K_z(0) & K_z(1) & \cdots & K_z(\text{n}-2) & K_z(\text{n}-1) \\ K_z(1) & K_z(0) & \cdots & K_z(\text{n}-3) & K_z(\text{n}-2) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ K_z(\text{n}-2) & K_z(\text{n}-3) & \cdots & K_z(0) & K_z(1) \\ K_z(\text{n}-1) & K_z(\text{n}-2) & \cdots & K_z(1) & K_z(0) \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_{n-1} \\ a_n \end{bmatrix} = \begin{bmatrix} -K_z(1) \\ -K_z(2) \\ \vdots \\ -K_z(n-1) \\ -K_z(n) \end{bmatrix} \tag{26}$$

By referring to [10], [11], the $1\times n$ observation vector C, the auto-variance function $K(0)$ of the state vector $x(k)$ and the system matrix Φ are obtained in terms of the auto-covariance data of the signal as follows:

$$C = [1 \quad 0 \quad \cdots \quad 0], \tag{27}$$

$$K(0) = \begin{bmatrix} K_z(0) & K_z(1) & \cdots & K_z(\text{n}-2) & K_z(\text{n}-1) \\ K_z(1) & K_z(0) & \cdots & K_z(\text{n}-3) & K_z(\text{n}-2) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ K_z(\text{n}-2) & K_z(\text{n}-3) & \cdots & K_z(0) & K_z(1) \\ K_z(\text{n}-1) & K_z(\text{n}-2) & \cdots & K_z(1) & K_z(0) \end{bmatrix}, \tag{28}$$

$$\Phi = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ -a_n & -a_{n-1} & -a_{n-2} & \cdots & -a_2 & -a_1 \end{bmatrix} \tag{29}$$

$K(0)$ is also called the Hankel matrix. As indicated in [11],[12], a finite dimensional realization for $z(k)$ exists if and only if the rank of the Hankel matrix is n.

Figure 1 illustrates the signal $z(k)$, the filtering estimate $\hat{z}(k|k)$ and the fixed-point smoothing estimate $\hat{z}(k|k+5)$, by the proposed iterated extended recursive Wiener fixed-point smoother and filter in the case of using the first derivative of $f(Cx(k),k)$ with respect to $x(k)$ in the Hessian, vs. k for the signal to noise ratio (SNR) 20 [dB]. From Figure1, it is seen that the fixed-point smoothing estimate $\hat{z}(k|k+5)$ is superior in estimation accuracy to the filtering estimate. Here, the iteration number I in the calculation of the filtering and fixed-point smoothing estimate is 20. Table 1 shows the mean of 20 trials for the mean-square values (MSVs) of the estimation errors in terms of $10log_{10}\frac{\sum_{k=1}^{550}(z(k)-\hat{z}(k|k))^2/550}{Var(z(k))}$ [dB] and $10log_{10}\frac{\sum_{k=1}^{550}(z(k)-\hat{z}(k|k+Lag))^2/550}{Var(z(k))}$ [dB], $Lag = 1, 2, \ldots, 5$, $Var(z(k)) = 0.9377$, by the extended recursive Wiener filter and fixed-point smoother [4]. From Table 1, in each signal-to-noise ratio (SNR) [dB], the estimation accuracy for the fixed-point smoother is superior to that of the filter. Table 2 shows the mean of 20 trials for the MSVs of the estimation errors by the iterated extended recursive Wiener fixed-point smoother and filter in the case of using the first derivative of $f(x(k),k)$ with respect to $x(k)$ in the Hessian. From Table 2, in each SNR [dB], the estimation accuracy for the fixed-point smoother is superior to that of the filter. Here, the iteration number I is 20.Table 3 shows the mean of 20 trials of the MSVs of the estimation errors by the iterated extended recursive Wiener fixed-point smoother and filter in the case of using up to the second derivative of $f(x(k),k)$ with respect to $x(k)$ in the Hessian. From Table 3, in each SNR [dB], the estimation accuracy for the fixed-point smoother is superior to that of the filter. Here, the iteration number I is 20.Table 4 shows the mean of 20 trials for the MSVs of the estimation errors by the iterated extended recursive Wiener fixed-point smoother and filter, which are obtained based on the iterated extended Kalman filter in [6].From Table 1 to Table 4,the extended recursive Wiener estimators [4], the extended recursive Wiener estimators in the case of using the first derivative of $f(x(k),k)$ with respect to $x(k)$ in the Hessian and the iterated extended recursive Wiener estimators in the case of using up to the second derivative of $f(x(k),k)$ with respect to $x(k)$ in the Hessian have almost the same

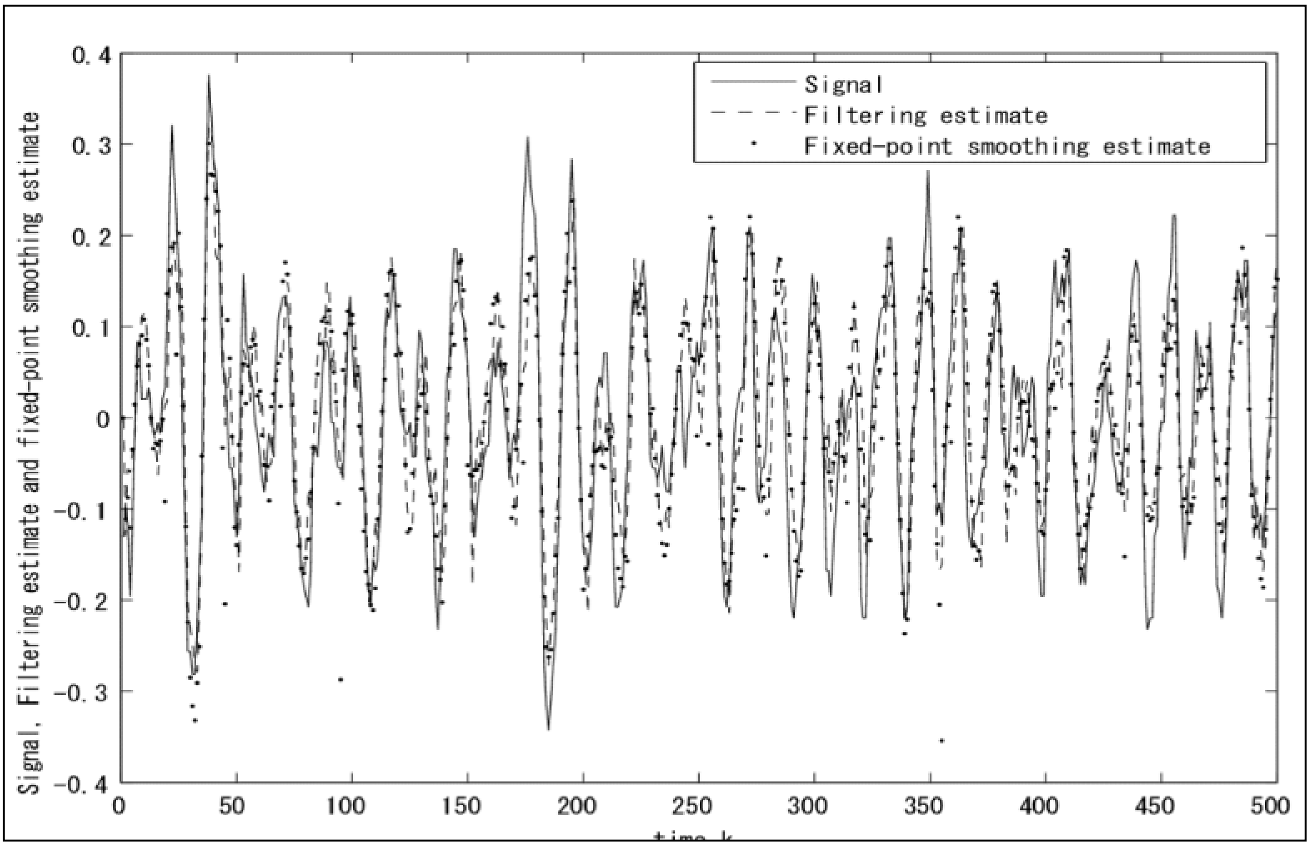

Figure 1 Signal $z(k)$, the filtering estimate $\hat{z}(k|k)$ and the fixed-point smoothing estimate $\hat{z}(k|k+5)$, by the proposed iterated extended recursive Wiener fixed-point smoother and filter in the case of using the first derivative of $f(x(k),k)$ with respect to $x(k)$ in the Hessian, vs. k for the signal to noise ratio (SNR) 20 [dB].

Table 1 Mean of 20 trials for the MSVs of the estimation errors by the extended recursive Wiener fixed-point smoother and filter [4].

Signal to noise ratio (SNR) [dB]	MSV in [dB] for $\hat{z}(k\|k)$	MSV in [dB] for $\hat{z}(k\|k+1)$	MSV in [dB] for $\hat{z}(k\|k+2)$	MSV in [dB] for $\hat{z}(k\|k+3)$	MSV in [dB] for $\hat{z}(k\|k+4)$	MSV in [dB] for $\hat{z}(k\|k+5)$
5 [dB]	-18.5811	-18.6989	-18.7773	-18.6728	-18.6098	-18.6792
10 [dB]	-19.4003	-19.6863	-19.7576	-19.9175	-19.8346	-19.8226
20 [dB]	-23.1121	-24.1060	-24.2740	-24.4684	-23.2599	-24.5408

estimation accuracy and are superior in estimation accuracy to the iterated extended recursive Wiener estimators obtained based on the iterated extended Kalman filter in [6].Table 5 shows the mean of 20 trials for the MSVs of the estimation errors by the extended recursive Wiener fixed-point smoother and filter in accordance with the medium-scale Quasi-Newton line search method in minimizing the function $J(x(k))$ .The finite difference method to get the gradient is used by the MATLAB program "fminunc.m" in the optimization toolbox. Table 6 shows the mean of 20 trials for the MSVs of the estimation errors by the extended recursive Wiener fixed-point smoother and filter in accordance with the Quasi Newton (the L-BFGS) method in minimizing the function $J(x(k))$. In the minimization, the MATLAB program "fminunc.m" in the optimization tool box is used under the condition that the gradient of $f(x(k),k)$ is provided. Table 7 shows the mean of 20 trials for the MSVs of the estimation errors by the extended recursive Wiener fixed-point smoother and filter in accordance with the Nelder-Mead simplex direct search method. In the minimization of the function $J(x(k))$,

Table 2 Mean of 20 trials for the MSVs of the estimation errors by the iterated extended recursive Wiener fixed-point smoother and filter in the case of using the first derivative of $f(x(k),k)$ with respect to $x(k)$ in the Hessian.

Signal to noise ratio (SNR) [dB]	MSV in [dB] for $\hat{z}(k\|k)$	MSV in [dB] for $\hat{z}(k\|k+1)$	MSV in [dB] for $\hat{z}(k\|k+2)$	MSV in [dB] for $\hat{z}(k\|k+3)$	MSV in [dB] for $\hat{z}(k\|k+4)$	MSV in [dB] for $\hat{z}(k\|k+5)$
5 [dB]	-18.5812	-18.6576	-18.5958	-18.5315	-18.8368	-18.7306
10 [dB]	-19.4395	-19.8349	-19.7868	-19.8576	-19.9862	-19.8912
20 [dB]	-23.1917	-24.2057	-24.2218	-24.1370	-23.9037	-23.7592

Table 3 Mean of 20 trials for the MSVs of the estimation errors by the iterated extended recursive Wiener fixed-point smoother and filter in the case of using up to the second derivative of $f(x(k),k)$ with respect to $x(k)$ in the Hessian.

Signal to noise ratio (SNR) [dB]	MSV in [dB] for $\hat{z}(k\|k)$	MSV in [dB] for $\hat{z}(k\|k+1)$	MSV in [dB] for $\hat{z}(k\|k+2)$	MSV in [dB] for $\hat{z}(k\|k+3)$	MSV in [dB] for $\hat{z}(k\|k+4)$	MSV in [dB] for $\hat{z}(k\|k+5)$
5 [dB]	-18.5725	-18.6896	-18.7188	-18.8099	-18.7423	-18.5943
10 [dB]	-19.4225	-19.6977	-19.9219	-19.9660	-19.7660	-19.9009
20 [dB]	-23.0858	-24.1181	-24.1798	-24.2277	-23.8132	-23.6467

Table 4 Mean of 20 trials for the MSVs of the estimation errors by the iterated extended recursive Wiener fixed-point smoother and filter obtained based on the iterated extended Kalman filter in [6].

Signal to noise ratio (SNR) [dB]	MSV in [dB] for $\hat{z}(k\|k)$	MSV in [dB] for $\hat{z}(k\|k+1)$	MSV in [dB] for $\hat{z}(k\|k+2)$	MSV in [dB] for $\hat{z}(k\|k+3)$	MSV in [dB] for $\hat{z}(k\|k+4)$	MSV in [dB] for $\hat{z}(k\|k+5)$
5 [dB]	-16.4859	-16.6537	-16.7070	-16.7405	-16.7597	-16.7504
10 [dB]	-16.9881	-17.4304	-17.5647	-17.6469	-17.9209	-17.9468
20 [dB]	-20.6227	-22.4695	-22.8444	-22.7787	-22.8733	-22.9444

Table 5 Mean of 20 trials for the MSVs of the estimation errors by the extended recursive Wiener fixed-point smoother and filter in accordance with the medium-scale Quasi-Newton line search method.

Signal to noise ratio (SNR) [dB]	MSV in [dB] for $\hat{z}(k\|k)$	MSV in [dB] for $\hat{z}(k\|k+1)$	MSV in [dB] for $\hat{z}(k\|k+2)$	MSV in [dB] for $\hat{z}(k\|k+3)$	MSV in [dB] for $\hat{z}(k\|k+4)$	MSV in [dB] for $\hat{z}(k\|k+5)$
5 [dB]	-10.6724	-11.2141	-11.4688	-11.7531	-11.9821	-12.1963
10 [dB]	-13.9068	-14.4664	-14.7144	-15.0748	-15.4149	-15.5331
20 [dB]	-20.0936	-20.4134	-20.4698	-21.0171	-21.3275	-21.5210

Table 6 Mean of 20 trials for the MSVs of the estimation errors by the extended recursive Wiener fixed-point smoother and filter in accordance with the Quasi-Newton (the L-BFGS) method.

Signal to noise ratio (SNR) [dB]	MSV in [dB] for $\hat{\boldsymbol{z}}(\boldsymbol{k}\,|\,\boldsymbol{k})$	MSV in [dB] for $\hat{\boldsymbol{z}}(\boldsymbol{k}\,|\,\boldsymbol{k}+1)$	MSV in [dB] for $\hat{\boldsymbol{z}}(\boldsymbol{k}\,|\,\boldsymbol{k}+2)$	MSV in [dB] for $\hat{\boldsymbol{z}}(\boldsymbol{k}\,|\,\boldsymbol{k}+3)$	MSV in [dB] for $\hat{\boldsymbol{z}}(\boldsymbol{k}\,|\,\boldsymbol{k}+4)$	MSV in [dB] for $\hat{\boldsymbol{z}}(\boldsymbol{k}\,|\,\boldsymbol{k}+5)$
5 [dB]	-10.6414	-11.1847	-11.4303	-11.6960	-11.8523	-12.0151
[dB]	-13.4928	-14.1294	-14.6363	-14.9573	-15.1699	-15.4152
20 [dB]	-19.9816	-20.3124	-20.3429	-20.8418	-21.0235	-21.2226

Table 7 Mean of 20 trials for the MSVs of the estimation errors by the extended recursive Wiener fixed-point smoother and filter in accordance with the Nelder-Mead simplex direct search method.

Signal to noise ratio (SNR) [dB]	MSV in [dB] for $\hat{\boldsymbol{z}}(\boldsymbol{k}\,|\,\boldsymbol{k})$	MSV in [dB] for $\hat{\boldsymbol{z}}(\boldsymbol{k}\,|\,\boldsymbol{k}+1)$	MSV in [dB] for $\hat{\boldsymbol{z}}(\boldsymbol{k}\,|\,\boldsymbol{k}+2)$	MSV in [dB] for $\hat{\boldsymbol{z}}(\boldsymbol{k}\,|\,\boldsymbol{k}+3)$	MSV in [dB] for $\hat{\boldsymbol{z}}(\boldsymbol{k}\,|\,\boldsymbol{k}+4)$	MSV in [dB] for $\hat{\boldsymbol{z}}(\boldsymbol{k}\,|\,\boldsymbol{k}+5)$
5 [dB]	-10.1890	-10.8414	-10.7074	-11.5343	-11.7529	-12.1399
10 [dB]	-13.1325	-13.9186	-14.3813	-14.7564	-14.9933	-15.1571
20 [dB]	-19.1967	-19.7477	-19.8762	-20.4420	-20.8427	-21.0225

the MATLAB program "fminsearch.m" in the optimization toolbox is used. From Table 5, Table 6 and Table 7, in each SNR [dB], the estimation accuracy for the fixed-point smoother is superior to that of the filter and, as the Lag increases, there is a tendency that the estimation accuracy of the fixed-point smoother is improved. Also, the estimation accuracies for the extended recursive Wiener fixed-point smoother and the filter, in accordance with the medium-scale Quasi-Newton line search method, the Quasi-Newton (L-BFGS) method and the Nelder-Mead simplex direct search method, are almost same. The extended recursive Wiener estimators [4] and the three kinds of iterated extended recursive Wiener estimators are superior in estimation accuracy to the extended recursive Wiener estimators in accordance with the medium-scale Quasi-Newton line search method, the Quasi-Newton (L-BFGS) method and the Nelder-Mead simplex direct search method. Table 8 compares the computation times of the filtering estimate $\hat{z}(k\,|\,k)$ and the fixed-point smoothing estimate $\hat{z}(k\,|\,k+1)$, $1 \le \mathrm{k} \le 550$ per trial for seven estimation techniques. Among the seven estimation algorithms, the extended recursive Wiener fixed-point smoother and filter [4] show the shortest computational time. It takes long computational times in the extended recursive Wiener fixed-point smoothing and filtering algorithms in accordance with the medium-scale Quasi-Newton line search method, the Quasi-Newton (the L-BFGS) method and the Nelder-Mead simplex direct search method. Here, in the computations of the estimates, MATLAB R2013a is used on DELL PC Optiplex 3020 (Intel® Celeron® CPU G1820 @ 2.70 GHz, OS: Windows 10, RAM: 4 GB).

For the smooth nonlinear modulation function whose deviations from the reference (nominal) trajectory are small enough to allow linear perturbation techniques [13], the extended recursive Wiener estimators are feasible in estimation accuracy. Likewise, from the above simulation, the iterated extended recursive Wiener estimators, in the both cases of using the first derivative of $\mathrm{f}\left(\mathrm{x}(k),k\right)$ and up to the second derivative of $\mathrm{f}\left(\mathrm{x}(k),k\right)$ with respect to $\mathrm{x}\left(k\right)$ in the Hessian, have shown the feasible estimation performance.

Table 8 Comparison of computation times of the filtering estimate $\hat{z}(k|k)$ and the fixed-point smoothing estimate $\hat{z}(k|k+1)$, $1 \le k \le 550$ per trial for seven estimation techniques.

Nonlinear estimation techniques (Table No.)	Computation time (numbered in ascending order of computation time of the seven methods)
1. Extended recursive Wiener fixed-point smoother and filter [4]. (Table 1)	0.642188 [s](1)
2. Iterated extended recursive Wiener fixed-point smoother and filter in the case of using the first derivative of $f(x(k),k)$ with respect to $x(k)$ in the Hessian. (Table 2)	7.882021 [s] for the iteration I = 20(3)
3. Iterated extended recursive Wiener fixed-point smoother and filter in the case of using up to the second derivative of $f(x(k),k)$ with respect to $x(k)$ in the Hessian. (Table 3)	8.120086 [s] for the iteration I = 20(4)
4. Iterated extended recursive Wiener fixed-point smoother and filter obtained based on the extended Kalman filter in [6]. (Table 4)	1.711135 [s] for the iteration I = 20(2)
5. Extended recursive Wiener fixed-point smoother and filter in accordance with the medium-scale Quasi-Newton line search method. (Table 5)	96.268837 [s](7)
6. Extended recursive Wiener fixed-point smoother and filter algorithms in accordance with the Quasi-Newton (L-BFGS) method. (Table 6)	61.558414 [s](5)
7. Extended recursive Wiener fixed-point smoother and filter in accordance with the Nelder-Mead simplex direct search method. (Table 7)	83.857940 [s](6)

5. Conclusions

This paper has proposed approximate MAP extended recursive Wiener fixed-point smoother and filter for the white Gaussian observation noise. This paper, at first, has devised two kinds of iterated extended recursive Wiener extended algorithms for the fixed-point smoothing and filtering estimates. One is regarded as the Newton-Raphson algorithm, since it use the first derivative in the Hessian with respect to the state vector, of the nonlinear observation function, and the other as the Newton's algorithm which uses up to the second derivative in the Hessian. Secondly, this paper has proposed three kinds of extended recursive Wiener fixed-point smoothing and filtering algorithms in accordance with the medium-scale Quasi-Newton line search method, the Quasi-Newton (the L-BFGS) method and the Nelder-Mead simplex direct search method.

Among the seven estimation algorithms, the extended recursive Wiener fixed-point smoother and filter [4] have shown the shortest computational time. It takes long computational times in the extended recursive Wiener fixed-point smoother and filter in accordance with the medium-scale Quasi-Newton line search method, the Quasi-Newton (the L-BFGS) method and the Nelder-Mead simplex direct search method. From the aspect of the estimation accuracy, these three estimators are inferior to the extended recursive Wiener estimators [4] and the three kinds of iterated extended recursive Wiener estimators. The extended recursive Wiener estimators [4], the iterated extended recursive Wiener estimators in the both cases of using the first derivative of $f(x(k),k)$ and up to the second derivative of $f(x(k),k)$ with respect to $x(k)$ in the Hessian have almost the same estimation accuracy and are superior in estimation accuracy to the iterated extended recursive Wiener estimators, obtained based on the iterated extended Kalman filter in [6].

References

[1] Liang-qun, Li, Hong-bing, Ji and Jun-hui, Luo. (2005). The iterated extended Kalman particle filter. International Symposium on Communications and Information Technologies 2005 (ISCIT 2005), 1172-1175.

[2] Wan, E. A. and van der Merwe, R. (2001). The Unscented Kalman Filter. In *Kalman Filtering and Neural Networks***.** Chapter 7(62 pages), Haykin, S. (ed). New York: Wiley Publishing, Inc.

[3] Himmelblau, D. M. (2009). *Applied Nonlinear Programming*. New York: McGraw-Hill Book Co.

[4] Nakamori, S. (2007). Design of extended recursive Wiener fixed-point smoother and filter in discrete-time stochastic systems. Digital Signal Processing, 17, 360-370.

[5] Blahut, R. E. (1990). *Digital Transmission of Information*.MA: Addison-Wesley Publishing Co.

[6] Simon, D. (2006).*Optimal State Estimation Kalman,* H_∞ *, and Nonlinear Approaches*.New York: Wiley Publishing, Inc.

[7] Sage, A.P. and Melsa, J.L. (1971).*Estimation Theory with Applications to Communications and Control*. New York: McGraw-Hill Book Co.

[8] Press, W. H., Teukolsky, S. A., Vetterling, W. T. and Flannery, B. P. (2002). *Numerical Recipes in C, The Art of Scientific Computing*, Second Edition. Cambridge: Cambridge University Press.

[9] Van Trees, H. L. (1968).*Detection, Estimation, and Modulation Theory, Part 1*.New York: Wiley Publishing, Inc.

[10] Nakamori, S. (1995). Recursive estimation technique of signal from output measurement data in linear discrete-time systems. IEICE Trans. Fundamentals of Electronics, Communications and Computer Sciences, E82-A, 600-607.

[11] Nakamori, S. (1997). Reduced-order estimation technique using covariance data of observed value in linear discrete-time systems. Digital Signal Processing, 7, 55-62.

[12] Akaike, H. (1974). Stochastic theory of minimal realization. IEEE Trans. Automatic Control, 19, 667-674.

[13] Maybeck, P. S. (1982). *Stochastic Models, Estimation, and Control*. In Mathematics in Science and Engineering, A Series of Monographs and Textbooks, 2 (141-2), Bellman, R. (ed). New York: Academic Press.

Author Biography

Seiichi Nakamori is a Specially Appointed Professor in the Faculty of Education, Kagoshima University, from April, 2017 to date. He was appointed Professor Emeritus, Kagoshima University, in April 2016. He was a Professor, Department of Technology, Graduate School of Education, Kagoshima University from April, 1994 to March 2016, Associate Professor, in the Department from April, 1987 to March, 1994 and a lecturer, Interdisciplinary Chair (Applied Mathematics), Faculty of Engineering, Oita University, from April, 1985 to March, 1987.

He received the B.E. degree in Electronic Engineering from Kagoshima University in 1974 and the Dr. Eng. Degree in Applied Mathematics and Physics from Kyoto University in 1982. He is mainly interested in stochastic signal estimation and image restoration.