13

# Detection of Signals in Nonstationary Noise via Kalman Filter-Based Stationarization Approach

Hiroshi Ijima* & Akira Ohsumi**,***
*Wakayama University*
*** University of Miyazaki*
*Japan*



## 1. Introduction

Needless to say, the signal detection is one of the most important problems in the signal processing area for a long time, and a great deal of investigations has been done up to the present time. Most of the conventional approaches are based on the (binary) hypothesis-testing, and treat the corrupting (additive) noise as a stationary random process because stationary process is rather easy to handle and moreover its (invariant) statistical parameters can be readily calculated under the ergodic hypothesis. However, it will be no doubt that the actual random noise such as environmental noise is considered to be nonstationary because its statistical properties are not always unchanged but vary according to underlying physical circumstances.
Thus the problem of detecting signals in nonstationary random noise is the more important. For such problem, several interesting methods have been proposed. For example, Haykin (1996) and Haykin & Bhattacharya (1997) treat this problem and proposed a method named the modular learning strategy which incorporates such three fundamental blocks as time-frequency analysis, feature extraction and pattern classification. Also, Haykin & Thomson (1998) proposed an adaptive detector based on learning for the detection of the target signal buried in nonstationary background noises.
Philosophically different from their method, the authors have proposed an approach to the signal detection in nonstationary random noise, a new method of stationarization of the observation noise. The key of the approach is to convert the nonstationary random noise to a stationary one, and this procedure was named as stationarization of the observation data.
In Ijima, Okui & Ohsumi (2005) and Ijima, Ohsumi & Okui (2006), the signal detection is performed by testing the stationarized observation data whether there is some non-stationarized portion or not, based on the $KM_2O$-Langevin equation (which is the AR model with time-varying coefficients). If there exists such a portion in the data, the existence of a signal is decided. Related to the signal detection, the stationarization approach is also used in Ijima,

∗ Faculty of Education, Wakayama University, Sakaedani, Wakayama 640-8510, Japan;
e-mail: ijima@center.wakayama-u.ac.jp
∗∗ Graduate School of Engineering, University of Miyazaki, Kibana, Gakuen, Miyazaki, 889-2192, Japan
∗ ∗ ∗ Presently, Professor Emeritus of Kyoto Institute of Technology; e-mail: akiraspika@nifty.com

Ohsumi & Yamaguchi (2006) to estimate the time-delay of signals in nonstationary random noise, incorporated with the Wigner distribution-based maximum likelihood estimation.
In this paper the signal detection problem is investigated using the stationarization approach to nonstationary data. The model of the corrupting noise is given by an ARMA($p, q$) model with unknown time-varying coefficients. These coefficient parameters are estimated from the (original) observation data by the Kalman filter.

## 2. Problem Statement

Let $\{y(k)\}$ be the (scalar) observation data taken at sampling time instant $t_k$ ($k = 1, 2, \cdots$), and assume that it can be expressed as

$$y(k) = s(k) + n(k) \quad (k = 1, 2, \cdots), \tag{1}$$

where $s(\cdot)$ is a signal to be detected, whose form is surely known, and is assumed to exist in a brief interval if it exists; and $n(\cdot)$ is the nonstationary random noise. In consequence, the observation data $\{y(k)\}$ becomes nonstationary, but its trend time series is assumed to be removed by the process

$$y(k) = \Delta^d Y(k), \tag{2}$$

where $Y(k)$ is the original data received by the receiver; $\Delta Y(k) = Y(k) - Y(k-1)$; and $d$ indicates the order.
In this paper the random noise $n(k)$ is assumed to be given as the output of ARMA($p$, $q$) model with time-varying coefficient parameters:

$$n(k) + \sum_{i=1}^{p} \alpha_i(k) n(k-i) = \sum_{j=1}^{q} \beta_j(k) w(k-j) + w(k), \tag{3}$$

where $w(\cdot)$ is the white Gaussian noise with zero-mean and variance parameter $\sigma^2$; $\{\alpha_i(\cdot)\}$ and $\{\beta_j(\cdot)\}$ are slowly and smoothly varying parameters to be specified.
Then our purpose is to propose a method of detecting the signal $s(k)$ from the noisy observation data $\{y(k)\}$.
The procedure taken in this paper is as follows:
(i) First, based on the noise model (3), coefficient functions $\{\alpha_i(\cdot)\}$ and $\{\beta_j(\cdot)\}$ are estimated using Kalman filter from the observation data $\{y(k)\}$.
(ii) Using the estimates $\{\hat{\alpha}_i(\cdot)\}$ and $\{\hat{\beta}_j(\cdot)\}$ obtained in (i), the observation data $y(k)$ is modified to become stationary. This procedure is called the *stationarization of observation data*.
(iii) Using the stationarized observation data $\hat{y}(k)$, the signal detection is based on the model

$$\hat{y}(k) = \hat{s}(k) + w(k), \tag{4}$$

where $\hat{s}(k)$ is the modified signal. Equation (4) is familiar in the conventional signal detection problem where the noise is stationary.

## 3. Stationarization of Observation Data

Recalling the assumption that the duration of the signal $s(k)$ is short, neglect the signal in the observation data and consider the signal-free case, i.e., $y(k) = n(k)$, then the observation data $y(k)$ is expressed by (1) and (3) as follows:

$$y(k) = -\sum_{i=1}^{p} \alpha_i(k) y(k-i) + \sum_{j=1}^{q} \beta_j(k) w(k-j) + w(k). \tag{5}$$

In order to estimate the time-varying parameters $\{\alpha_i(k)\}$ and $\{\beta_j(k)\}$ in (5), suppose that they change from step $k-1$ to $k$ under random effects $\{e_\cdot(k)\}$. Define vectors

$$x(k) = \begin{bmatrix} -\alpha_1(k) \\ \vdots \\ -\alpha_p(k) \\ \beta_1(k) \\ \vdots \\ \beta_q(k) \end{bmatrix}, \quad v(k) = \begin{bmatrix} -e_1(k) \\ \vdots \\ -e_p(k) \\ e_{p+1}(k) \\ \vdots \\ e_{p+q}(k) \end{bmatrix}. \tag{6}$$

Then, $\{\alpha_i(k)\}$ and $\{\beta_j(k)\}$ are subject to the dynamics,

$$x(k+1) = x(k) + v(k), \tag{7}$$

where $\{e_\cdot(k)\}$ are assumed to be Gaussian with zero-means and variances $\tau_1^2, \cdots, \tau_{p+q}^2$. Then, Eq. (5) is expressed formally as

$$y(k) = H(k)x(k) + w(k) \tag{8}$$

in which $H(k)$ is given by

$$H(k) = [\, y(k-1), \cdots, y(k-p), w(k-1), \cdots, w(k-q)] \,. \tag{9}$$

At this stage it should be noted that the matrix $H(k)$ consists of the (unmeasurable) past noise sequence $\{w(\cdot)\}$. To remedy this inadequate situation, we resort to replace it by

$$\hat{H}(k) = [\, y(k-1), \cdots, y(k-p), \nu_m(k-1), \cdots, \nu_m(k-q)] \tag{10}$$

in which $\{\nu_m(\cdot)\}$ is the sequence modified from the innovation sequence $\nu(\cdot)$ as

$$\nu_m(\ell) = c(\ell)\, \nu(\ell) \quad (\ell = k-q, k-q+1, \cdots, k-1)\,, \tag{11}$$

where

$$\nu(\ell) = y(\ell) - \hat{H}(\ell)\hat{x}(\ell|\ell-1) \tag{12}$$

and

$$c(\ell) = \left[ 1 + \frac{1}{\sigma^2} \hat{H}(\ell) P(\ell|\ell-1) \hat{H}^T(\ell) \right]^{-\frac{1}{2}}. \tag{13}$$

Here, $\hat{x}(\ell|\ell-1)$ and $P(\ell|\ell-1)$ are the one-step prediction and its covariance matrix computed by Kalman filter for the past interval.

It is a simple exercise to show that the statistical properties of $\nu_m(\cdot)$ is the same as that of $w(\cdot)$, i.e., $E\{\nu_m(k)\} = 0$ and $E\{|\nu_m(k)|^2\} = \sigma^2$ (for proof, see Appendix). Then, instead of (8) we have the expression,

$$y(k) = \hat{H}(k)x(k) + w(k)\,. \tag{14}$$

The procedure for computing $\hat{H}(k)$ is stated as follows:
(i) *Preliminaries*: Assume for the past $k(<0)$ that $\{\nu_m(-1), \nu_m(-2), \cdots, \nu_m(-q)\}$ are set appropriately (may be set all zero), and preassign $\hat{x}(0|-1)$, $\hat{P}(0|-1)$ and $\hat{H}(0)$ as initial values. Then, at time $k$ $(k = 0, 1, 2, \cdots)$
(ii) *Computation of* $\nu(\ell)$ *and* $c(\ell)$: Compute the innovation $\nu(\ell)$ and coefficient $c(\ell)$ by (12) and (13) using $\hat{H}(\ell) = [\,y(\ell-1), \cdots, y(\ell-p), \nu_m(\ell-1), \cdots, \nu_m(\ell-q)]$.
(iii) *Computation of* $\nu_m(\ell)$: Compute $\nu_m(\ell)$ by (11) using $\nu(\ell)$ and $c(\ell)$ obtained in the previous step.
Repeat Steps (ii) and (iii) for $\ell = k-q, k-q+1, \cdots, k-1$ to obtain $\hat{H}(k)$. In computing (12) and (13), $\hat{x}(\ell|\ell-1)$ and $P(\ell|\ell-1)$ are computed by the Kalman filter (e.g., Jazwinski, 1970):

$$\hat{x}(\ell+1|\ell) = \hat{x}(\ell|\ell) \tag{15}$$

$$\hat{x}(\ell|\ell) = \hat{x}(\ell|\ell-1) + K(\ell)\nu(\ell), \tag{16}$$

$$K(\ell) = \frac{1}{\hat{H}(\ell)P(\ell|\ell-1)\hat{H}^T(\ell) + \sigma^2}\,P(\ell|\ell-1)\hat{H}^T(\ell) \tag{17}$$

$$P(\ell+1|\ell) = P(\ell|\ell) + Q \tag{18}$$

$$P(\ell|\ell) = P(\ell|\ell-1) - K(\ell)\hat{H}(\ell)P(\ell|\ell-1), \tag{19}$$

where $Q = \mathrm{diag}\,\{\tau_1^2, \cdots, \tau_{p+q}^2\}$.
Thus, the estimates of the coefficient parameters $\{\alpha_i(k)\}$ and $\{\beta_j(k)\}$ are obtained by the Kalman filter constructed for (7) and (14) (whose form is the same as (15)-(19) replacing $\ell$ by the present $k$). Under the basic assumption that the coefficient parameters vary slowly and smoothly, they can be treated like constants in an interval $I_k$ around the current time $k$. Write them as $\hat{\alpha}_{ik}$ and $\hat{\beta}_{jk}$ in $I_k$. Replacing the past $\{w(k-j)\}$ in (5) by the statistically equivalent sequence $\{\nu_m(k-j)\}$, define the sequence $\hat{y}(k)$ by

$$\hat{y}(k) := y(k) + \sum_{i=1}^{p} \hat{\alpha}_{ik} y(k-i) - \sum_{j=1}^{q} \hat{\beta}_{jk} \nu_m(k-j). \tag{20}$$

Then, we have the following adequate approximation for (5),

$$\hat{y}(k) = w(k) \tag{21}$$

which implies that the sequence $\{\hat{y}(k)\}$ is stationary because $w(k)$ is the stationary white noise.

## 4. Signal Detection

After obtained the estimates of coefficient parameters, the observation process (14) may be written using estimates as

$$y(k) = \hat{H}(k)\hat{x}(k|k) + w(k) \tag{22}$$

or

$$y(k) + \sum_{i=1}^{p} \hat{\alpha}_{ik} y(k-i) = \sum_{j=1}^{q} \hat{\beta}_{jk} \nu_m(k-j) + w(k). \tag{23}$$

Now, let us revive the signal $s(k)$ in the observation data. To do this, replace $\{y(k)\}$ formally by $\{y(k) - s(k)\}$ in (23) to obtain

$$y(k) + \sum_{i=1}^{p} \hat{\alpha}_{ik} y(k-i) = \left[ s(k) + \sum_{i=1}^{p} \hat{\alpha}_{ik} s(k-i) \right] + \sum_{j=1}^{q} \hat{\beta}_{jk} \nu_m(k-j) + w(k) \tag{24}$$

or

$$\hat{y}(k) = \hat{s}(k) + w(k), \tag*{$(4)_{\mathrm{bis}}$}$$

where $\hat{y}(k)$ has the same form as (20) and

$$\hat{s}(k) = s(k) + \sum_{i=1}^{p} \hat{\alpha}_{ik} s(k-i). \tag{25}$$

Note that $(4)_{\mathrm{bis}}$ is familiar to us as the mathematical model for the detection problem of signals in *stationary* noise (e.g., Van Trees, 1968).

Now, consider the binary hypotheses: $H^1$: $\hat{y}(k) = \hat{s}(k) + w(k)$, and $H^0$: $\hat{y}(k) = w(k)$, and let $\hat{Y}_k$ be the stationarized observation data taken up to $k$, $\hat{Y}_k = \{\hat{y}(\ell), \ell = 1, 2, \cdots, k\}$. Since the additive noise $w(k)$ is white Gaussian sequence with zero-mean and variance $\sigma^2$, the likelihood-ratio function $\Lambda(k) = p\{\hat{Y}_k | H^1\} / \hat{Y}_k | H^0\}$ is evaluated as follows:

$$\Lambda(k) = \frac{\displaystyle\prod_{\ell=1}^{k} (2\pi)^{-\frac{1}{2}} \exp\left\{ -\frac{\{\hat{y}(\ell) - \hat{s}(\ell)\}^2}{2\sigma^2} \right\}}{\displaystyle\prod_{\ell=1}^{k} (2\pi)^{-\frac{1}{2}} \exp\left\{ -\frac{\hat{y}^2(\ell)}{2\sigma^2} \right\}}. \tag{26}$$

We use rather its logarithmic form,

$$\begin{aligned} L(k) &:= \ln \Lambda(k) \\ &= \frac{1}{\sigma^2} \sum_{\ell=1}^{k} \hat{s}(\ell) \hat{y}(\ell) - \frac{1}{2\sigma^2} \sum_{\ell=1}^{k} \hat{s}^2(\ell) \end{aligned} \tag{27}$$

as the signal detector.

## 5. Simulation Studies

In this section, we provide a typical set of several simulation results to demonstrate the proposed method.

(i) *Experiment* 1.

The top of Fig.1 depicts a sample path of the observation data $\{Y(k)\}$ generated by calculating the output of the ARMA(4, 1)-model:

$$n(k) = -\sum_{i=1}^{4} \alpha_i(k) n(k-i) + \beta(k) w(k-1) + w(k)\,.$$

Time-varying coefficients $\{\alpha_i(k)\}$ and $\beta(k)$ are set as

$$\alpha_1(k) = -1.24\sin(0.002k - 0.95),\ \alpha_2(k) = 0.38 - 2\cos(0.004k - 1.89)$$
$$\alpha_3(k) = \alpha_1(k), \quad \alpha_4(k) = 1, \quad \beta(k) = 1.5.$$

The bottom of Fig.1 shows a signal embedded in the observation data around $k = 300$ given by

$$s(\ell) = 12\, e^{-2.78\ell^2} \sin(1.26\ell),$$

where $\ell = k - 300$. Figure 2 depicts trend-removed data and stationarized data $\hat{y}(k)$. The trend was removed by setting $d = 1$. For the Kalman filter (15)$\sim$(19), the parameters are set

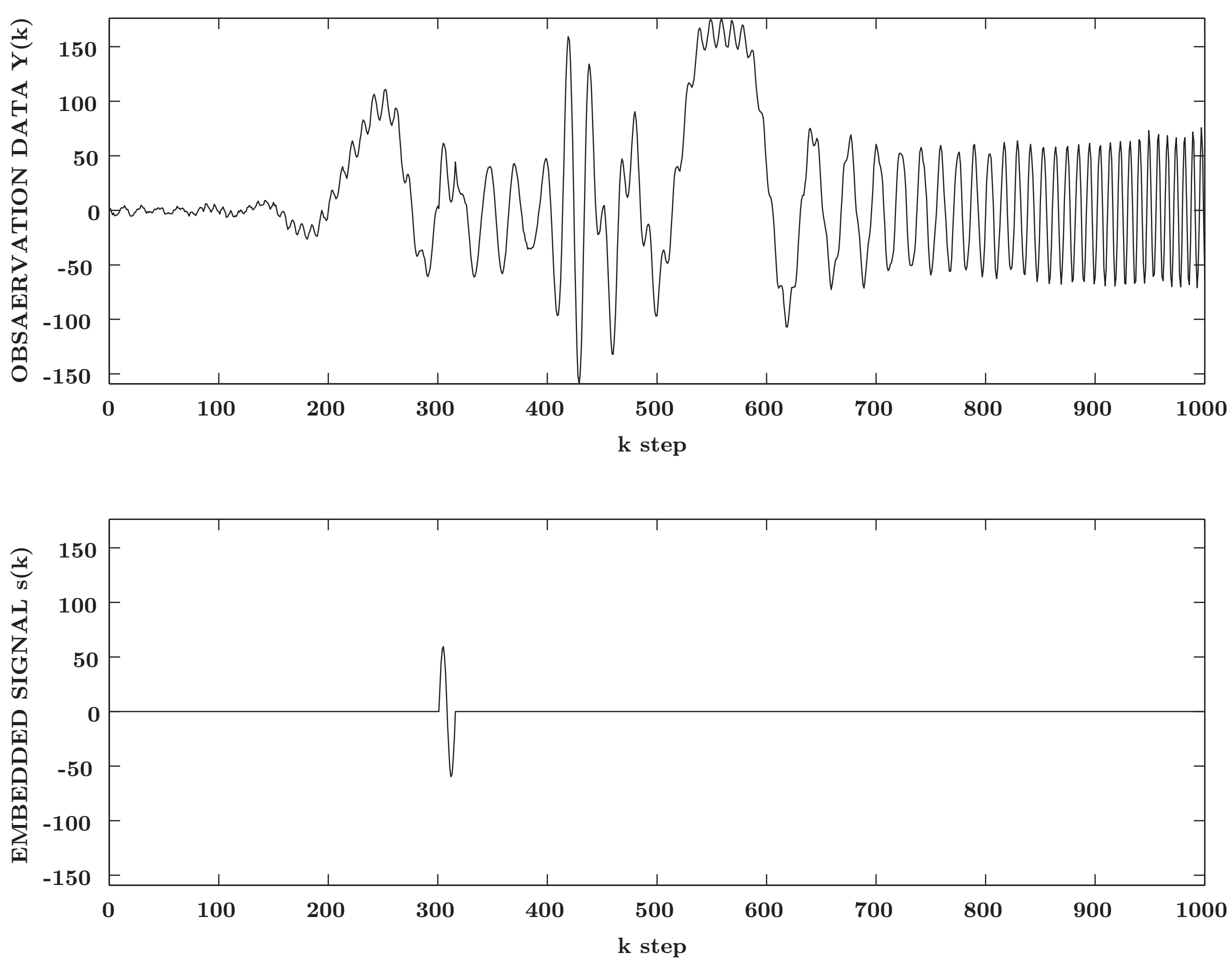


Fig. 1. A sample path of the observation data $Y(k)$ (top) and the embedded signal $s(k)$ (bottom).

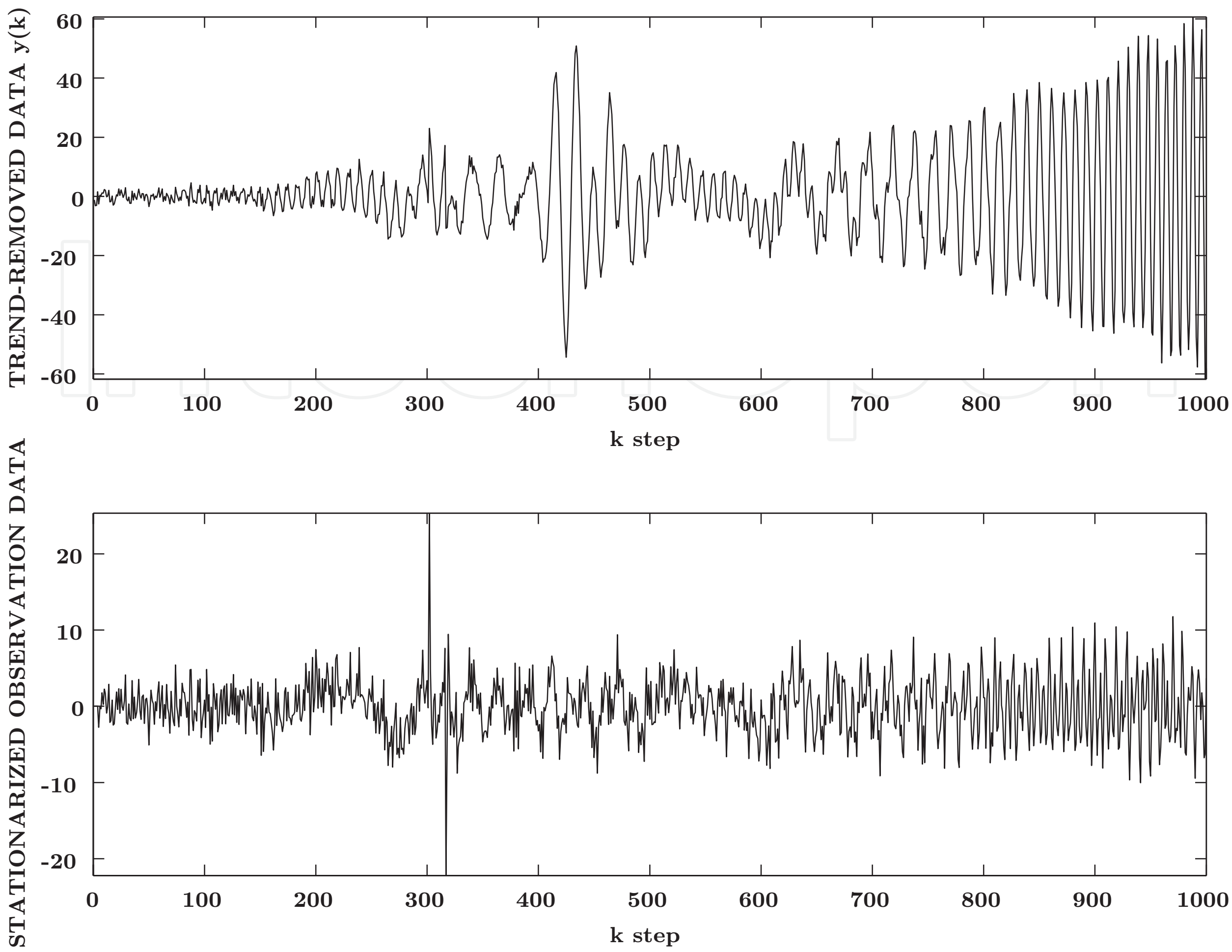


Fig. 2. The trend-removed data $y(k)$ (top) and the stationarized observation data $\hat{y}(k)$ (bottom).

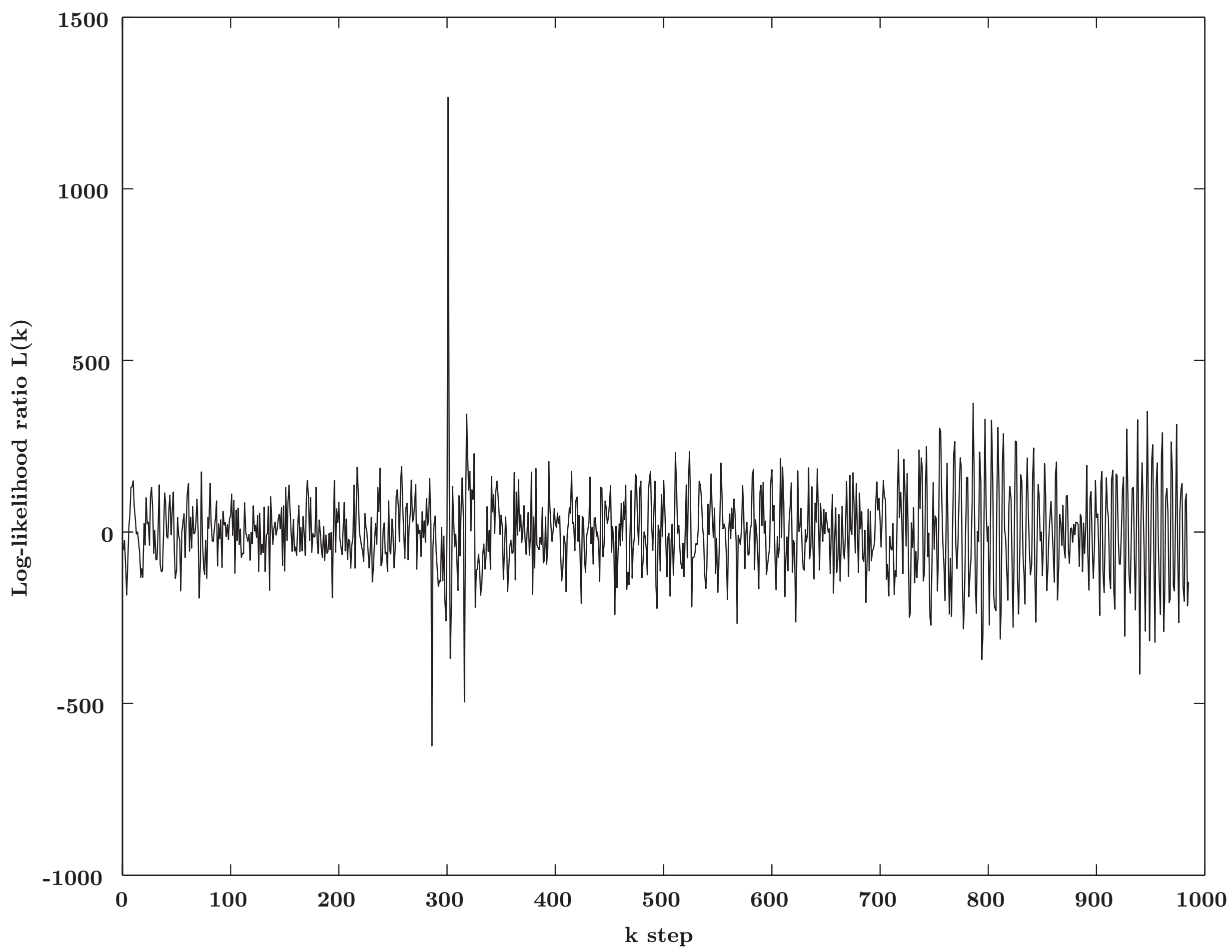


Fig. 3. Log-likelihood function $L(k)$.

as $Q = \text{diag}\,\{0.05,\ 0.05, 0.05,\ 0.05,\ 0.05\}$ and $\sigma^2 = 40$. It should be noted that from Fig. 2 the observation data is well stationarized and that even in this figure the signal emerges from the background noise.

Figure 3 shows the result of signal detection by the current log-likelihood ratio function $L(k)$. Clearly, it exhibits a salient peak around the true time instant $k = 300$ and this shows the existence of the signal.

(ii) *Experiment* 2.

Efficacy of the signal detector proposed in this paper is also tested for the pulse signal.

Figure 4 depicts observation data and embedded three pulses. Random noise $n(k)$ is generated by the same manner of previous simulation with same coefficients $\alpha_i(k)$ and $\beta(k)$. As a signals $s(k)$, a train of pulses with same magnitude is considered:

$$s(k) = \begin{cases} 20 & \text{for } D_i \le k < D_i + 5 \ \ (i = 1, 2, 3) \\ 0 & \text{otherwise,} \end{cases}$$

where $D_1 = 200$, $D_2 = 500$, $D_3 = 800$.

Figure 5 depicts trend-removed data and stationarized data $\hat{y}(k)$. The trend was also removed by setting $d = 1$. The parameters of Kalman filter are set as the same of previous experiment. Figure 6 shows the result of signal detection. Clearly, log-likelihood ratio function $L(k)$ has large value around each time when each pulse exists. Thus the signal detection is well succeeded.

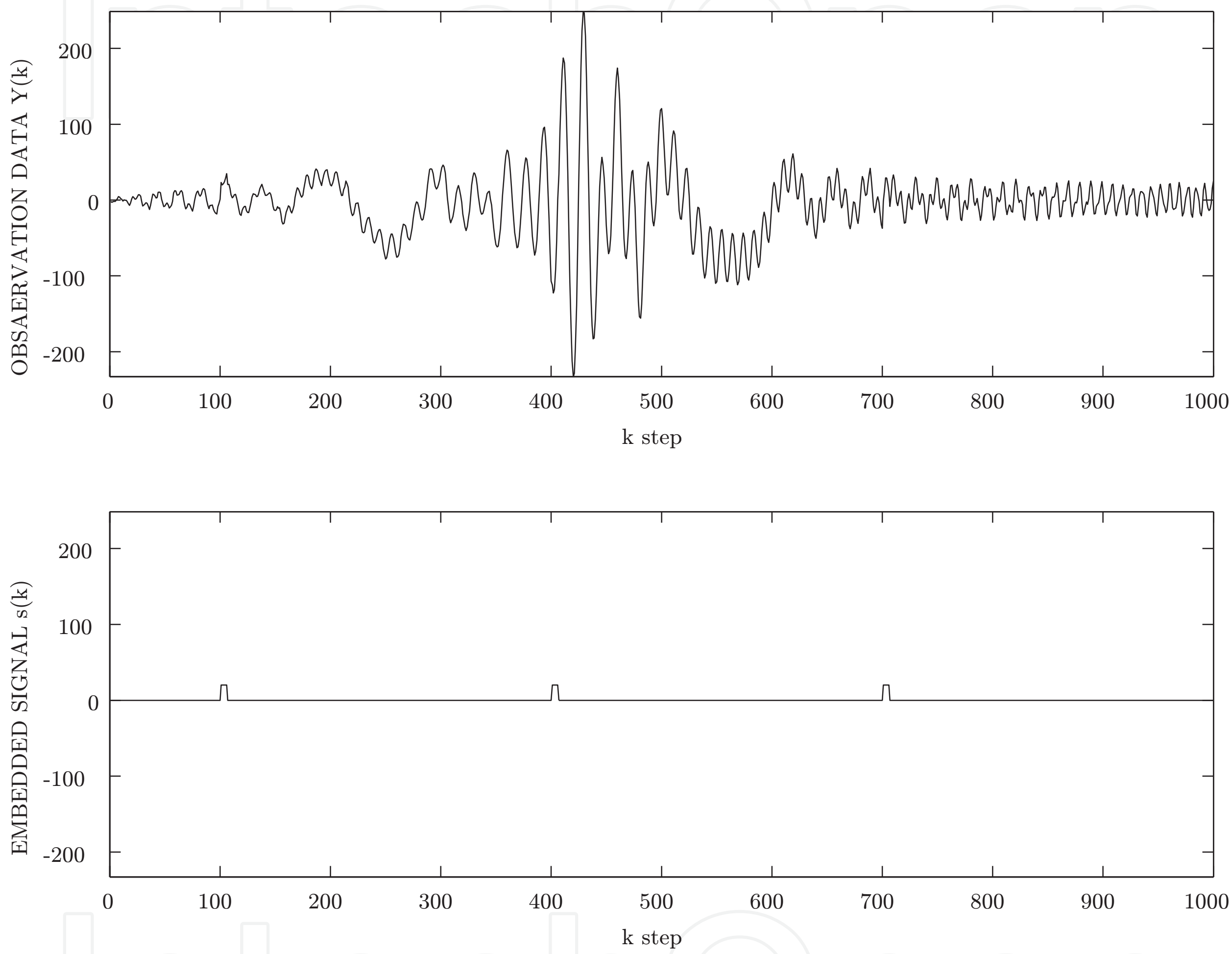


Fig. 4. A sample path of the observation data $Y(k)$ (top) and the pulse signal $s(k)$ (bottom).

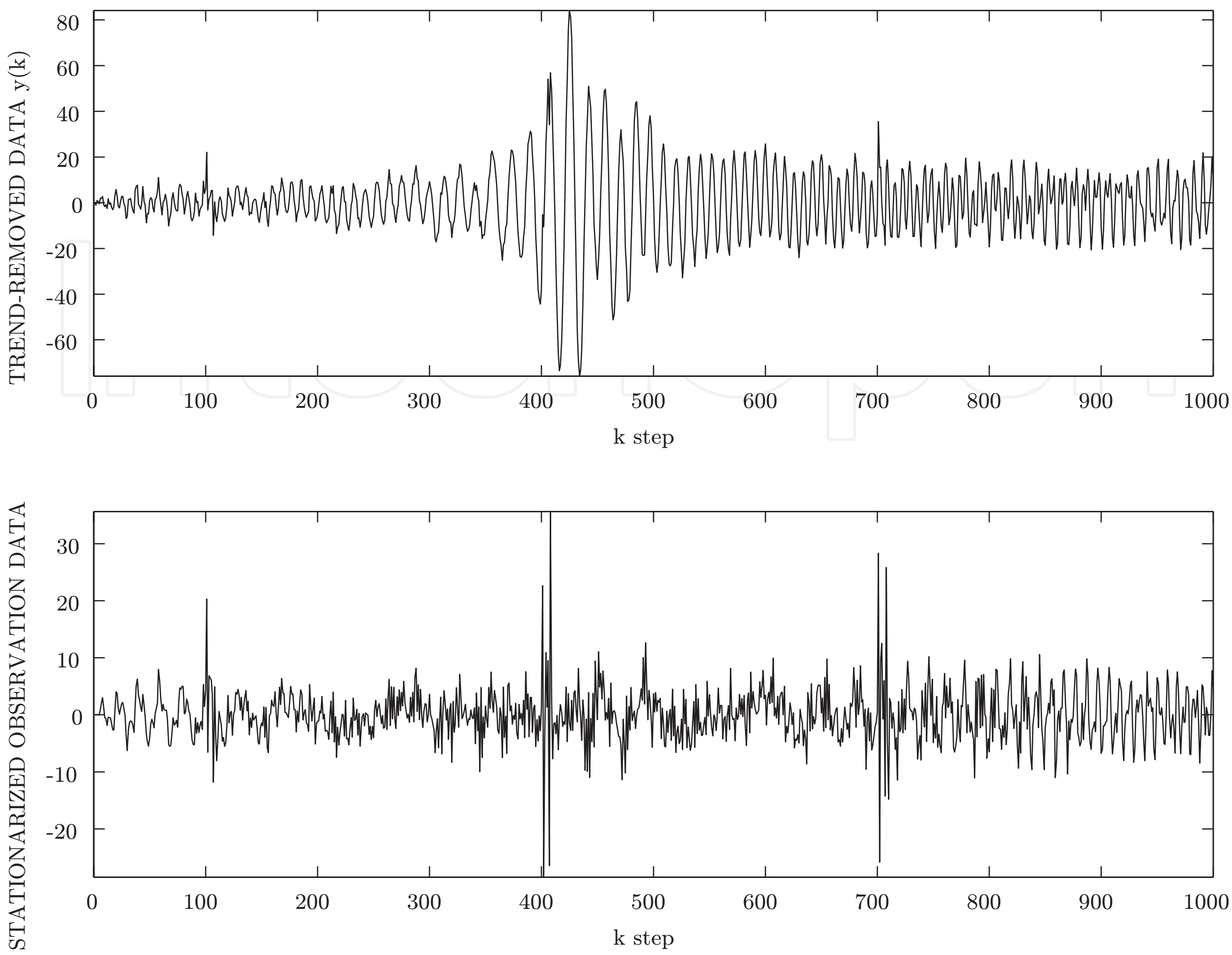


Fig. 5. The trend-removed data (top) and the stationarized observation data $\hat{y}(k)$ (bottom).

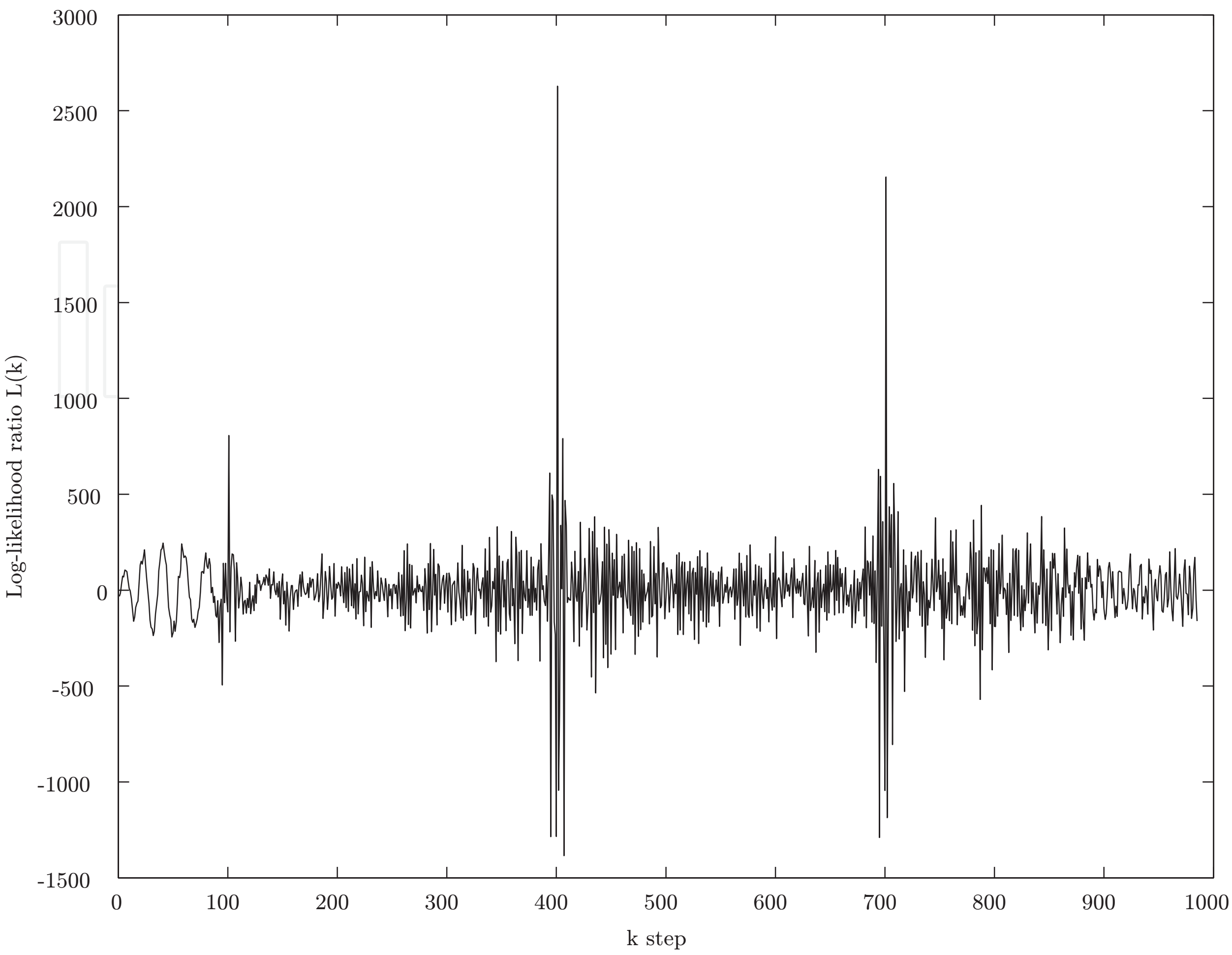


Fig. 6. Log-likelihood function $L(k)$.

## 6. Conclusion

The efficacy of the proposed signal detection method based on the stationarization of nonstationary observation data has been confirmed by simulation studies. The key to use the Kalman filter to estimate the coefficient parameters of the ARMA noise model is laid on the replacement of the unobservable past noise sequence by the equivalent (modified) innovation sequence which is observation data-measurable. The stationarization of a nonstationary data as introduced in this paper will have potential ability to treat the nonstationary noise or observation data in the signal processing.

**Appendix.** *Proof of Statistical Equivalence Between* $\{w(k)\}$ *and* $\{\nu_m(k)\}$

The mean of the modified innovation sequence $\nu_m(k)$ is clearly zero. Indeed,

$$\begin{aligned}\mathcal{E}\{\nu_m(k)\} &= c(k)\mathcal{E}\{\nu(k)\} \\ &= c(k)\mathcal{E}\{y(k) - \hat{H}(k)\hat{x}(k|k-1)\}.\end{aligned}$$

Here, recalling that $y(k)$ is given by the form (14), we have

$$\begin{aligned} &= c(k)[\hat{H}(k)\mathcal{E}\{x(k) - \hat{x}(k|k-1)\} + \mathcal{E}\{w(k)\}] \\ &= c(k)\hat{H}(k)\mathcal{E}\{\mathcal{E}\{x(k) - \hat{x}(k|k-1)|Y_{k-1}\}\} \\ &= c(k)\hat{H}(k)\mathcal{E}\{\mathcal{E}\{x(k)|Y_{k-1}\} - \hat{x}(k|k-1)\} \\ &= 0, \end{aligned}$$

where $Y_{k-1} = \{y(\ell), 0 \le \ell \le k-1\}$.
Next, the variance of $\nu_m(k)$ is evaluated as follows:

$$\begin{aligned} \mathcal{E}\{\nu_m^2(k)\} &= c^2(k)\mathcal{E}\{\nu^2(k)\} \\ &= c^2(k)\mathcal{E}\{[\hat{H}(k)[x(k) - \hat{x}(k|k-1)] + w(k)]^2\} \\ &= c^2(k)[\hat{H}(k)\mathcal{E}\{[x(k) - \hat{x}(k|k-1)][x(k) - \hat{x}(k|k-1)]^T\}\hat{H}^T(k) + \mathcal{E}\{w^2(k)\}] \\ &= c^2(k)[\hat{H}(k)P(k|k-1)\hat{H}^T(k) + \sigma^2]. \end{aligned}$$

If we select $c(k)$ as (13), the variance of $\nu_m(k)$-sequence becomes $\sigma^2$ which is just the variance of $\{w(k)\}$.
(Q.E.D.)

## 7. References

Haykin, S. (1996). Neural networks expand SP's horizons. *IEEE Signal Processing Mag.*, Vol.13, No.2, pp.24-29

Haykin, S. & Bhattacharya, T. K. (1997). Modular learning strategy for signal detection in a nonstationary environment. *IEEE Trans. Signal Processing*, Vol.45, No.6, pp.1619-1637

Haykin, S. & Thomson, D. J. (1998). Signal detection in a nonstationary environment reformulated as an adaptive pattern classification problem. *Proc. of the IEEE*, Vol.86, No.11, pp.2325-2344

Ijima, H., Ohsumi, A. & Okui, R. (2006). A method of detection of signals corrupted by nonstationary random noise via stationarization of the data, *Trans. IEICE, Fundamentals of Electronics, Communications and Computer Sciences*, Vol. J89-A, No.6, pp.535-543 (in Japanese)

Ijima, H., Ohsumi, A. & Yamaguchi, S. (2006). Nonlinear parametric estimation for signals in nonstationary random noise via stationarization and Wigner distribution, *Proc. 2006 Int. Symp. Nonlinear Theory and its Applic.* (*NOLTA 2006*), Bologna, Italy, pp.851-854

Ijima, H., Okui, R. & Ohsumi, A. (2005). Detection of signals is nonstationary random noise via staionarization and stationary test, *Proc. IEEE Workshop on Statistical Signal Processing* (*SSP'05*), Bordeaux, France, Paper ID 68

Jazwinski A. H. (1970). *Stochastic Processes and Filtering Theory*, Academic Press, New York

Van Trees, H. L. (1968). *Detection, Estimation, and Modulation Theory,* Part I, John Wiley


**Signal Processing**
Edited by Sebastian Miron

ISBN 978-953-7619-91-6
Hard cover, 528 pages
**Publisher** InTech
**Published online** 01, March, 2010
**Published in print edition** March, 2010

This book intends to provide highlights of the current research in signal processing area and to offer a snapshot of the recent advances in this field. This work is mainly destined to researchers in the signal processing related areas but it is also accessible to anyone with a scientific background desiring to have an up-to-date overview of this domain. The twenty-five chapters present methodological advances and recent applications of signal processing algorithms in various domains as telecommunications, array processing, biology, cryptography, image and speech processing. The methodologies illustrated in this book, such as sparse signal recovery, are hot topics in the signal processing community at this moment. The editor would like to thank all the authors for their excellent contributions in different areas of signal processing and hopes that this book will be of valuable help to the readers.