

Non-Parametric Signal Interpolation

Alexandr V. Dobrovidov

Russian Academy of Sciences, Moscow, Russia

Abstract: This paper considers the problem of interpolation (smoothing) of a partially observable Markov random sequence. For the dynamic observation models, an equation for the interpolation of the posterior probability density is derived. The main goal of this paper is to consider the smoothing problem for the case of unknown distributions of an unobservable component of a random Markov sequence. Successful results were obtained for the strongly stationary Markov processes with mixing and for the conditional density belonging to the exponential family of densities. The resulting method is based on the empirical Bayes approach and kernel nonparametric estimation. The equation for the optimal smoothing estimator is derived in the form independent of unknown distributions of an unobservable process. Such form of the equation allows to use the nonparametric estimates for some conditional functionals in the equation given a set of dependent observations. To compare the nonparametric estimators with optimal mean square smoothing estimators in Kalman scheme, simulation results are given.

Keywords: Markov Sequence, Interpolation, Non-Parametric Functional Estimates, Dependent Observations.

1 Introduction

This paper is devoted to synthesis of interpolation algorithms for an unobservable stationary sequence in conditional Markov scheme. The characteristic feature of the problem consists in absence of knowledge of the probability distribution family for the unobservable sequence. Nevertheless, some efforts are made to construct the sequence of estimators that have the asymptotic behavior (under a large volume of observations) close to the behavior of optimal interpolation estimators built under full statistical information on observable and unobservable signal components. For the case of full information, the transformation equations for the posterior probability density of unobservable sequences are known. They were obtained by R. Stratonovich for filtration and interpolation problems in 1966. The paper of Khazen (1978) represents the interpolation equation in the form of the normalizing product of filtration posterior probability densities in forward and backward time. This was done only for static observation models. The present paper shows that the similar equation is valid also for dynamic observation models (e.g., autoregressive models), however, in this case the equation is to be supplemented by another recursive equation that reflects the dynamic properties of observations.

Unfortunately, it is impossible to make use of this equations when the probability family for the unobservable sequence is unknown. However, for some conditional probability family of observations it is possible to transform these equations eliminating (explicit) dependence on statistic characteristics unknown *a priori*. The solution is found on the principles of the empirical Bayes approach and the theory of kernel non-parametric functional estimation (see Vasiliev, Dobrovidov, and Koshkin, 2004).

2 Interpolation Equation for Dynamic Observation Models

Interpolation (smoothing) of the partly observable Markov random sequence $(S_n, X_n)_{n \geq 1}$, $S_n \in \mathbb{R}^m$, $X_n \in \mathbb{R}^l$, is the problem of constructing estimators of the unobservable vector S_k or a known one-to-one function $Q(S_k)$ by observations $x_1^n = (x_1, \dots, x_n)^T$ of the sequence $(X_n)_{n \geq 1}$ for all $k \leq n$. It is well known, that the optimal mean-square smoothing estimator of $Q(S_k)$, $k \leq n$, equals to the conditional expectation

$$E(Q(S_k)|x_1^n) = \int_{\mathbb{R}^m} Q(s_k) \pi(s_k|x_1^n) ds_k, \quad (1)$$

where $\pi_k(s_k|x_1^n)$ is the conditional posterior probability density of S_k given all observable realizations x_1^n which is called to be the *interpolating* posterior density. There are some ways to calculate this conditional density. One of the interesting approaches examined below and referred to as the two-filter smoothing consists in recursive calculation of the *filtering* posterior probability density $w_k(s_k|x_1^k)$ in forward time and the *filtering* posterior density $\tilde{w}_k(s_k|x_k^n)$ in backward time. This algorithm is as follows (Khazen, 1978; Briers, Douset, and Maskell, 2003):

$$\pi_k(s_k|x_1^n) = \frac{f(x_1^k)f(x_k^n)}{f(x_1^n)} \cdot \frac{w_k(s_k|x_1^k)\tilde{w}_k(s_k|x_k^n)}{f(s_k, x_k)}, \quad (2)$$

where the first factor is a normalizing constant depending only on observations. Here and further the function f without index denotes any probability density that may differ from another one even in the same formula, because its argument completely determines the object. The computation results for filtering densities in forward time and backward time are multiplied giving the interpolating posterior density. For static and dynamic observation models, the algorithms of such computation are different. For static models described by the conditional density $f(x_k|s_k)$, the joint probability density $f(s_k, x_k)$ in the denominator of (2) can be represented in the form of the product $p(s_k)f(x_k|s_k)$, where $p(s_k)$ is the prior density. For dynamic observation models described by the conditional density $f(x_k|x_{k-1}, s_k)$, such a product form can not be constructed, because the density $f(x_k|s_k)$ is unknown in this case. Nevertheless, the joint probability density $f(s_k, x_k)$ admits representation via $f(x_k|x_{k-1}, s_k)$ by means of the following recursive equation

$$f(s_k, x_k) = \int_{S_{k-1}} p(s_k|s_{k-1}) \int_{X_{k-1}} f(x_k|x_{k-1}, s_k) f(s_{k-1}, x_{k-1}) dx_{k-1} ds_{k-1} \quad (3)$$

with the initial condition $f(s_1, x_1)$ that is given by the prior density of the Markov process $(S_n, X_n)_{n \geq 1}$. This equation can be simply derived using the total probability formula. Consequently, for dynamic observation models, pair (2), (3) is the system of equations for interpolating posterior probability density.

The subsequent computation in (2), (3) can be carried out if all distributions of the composed Markov process $(S_n, X_n)_{n \geq 1}$ are known. But the main purpose of this paper consists in solving the interpolation problem with unknown distributions of the unobservable strongly stationary Markov sequence $(S_n)_{n \geq 1}$. The central idea to solve this

problem is based on the principle of the empirical Bayes approach and the theory of non-parametric functional estimation by weakly dependent observations (see Vasiliev et al., 2004). The empirical Bayes approach aims to construct the estimators that are explicitly independent of the probabilistic characteristics of the unobservable random variables. This can be achieved, for instance, by using the conditional densities of observations from the exponential density family (see Chentsov, 1972)

$$f(x_n|s_n) = \tilde{C}(s_n)h(x_n) \exp \{T^T(x_n)Q(s_n)\} , \quad (4)$$

where $T = (T_1, \dots, T_m)^T$, $Q = (Q^{[1]}, \dots, Q^{[m]})^T$, $h(\cdot)$, $Q^{[j]}(\cdot)$, and $T_j(\cdot)$, $j = 1, \dots, m$ are the given Borelian functions and $\tilde{C}(s_n)$ is the normalizing factor.

3 Optimal Interpolation Estimator under Unknown Probability Distribution of the Unobservable Signal

Some applied problems of extracting a signal from its mixture with a noise result in a situation when this useful signal is never observed in the pure form and consequently one can not gather any statistic about its distribution. Moreover, any information about possible parametric family which includes the probability distribution for the signal is usually absent. To solve this problem we propose to use the empirical Bayes approach to the interpolation problem. This means that one must make an attempt to find such a form of the estimation equation that does not depend explicitly on the unknown probability characteristics of the unobserved signal. Now there are no regular methods for constructing such equations for arbitrary observation models. The success can be attained for more narrow class of observation models described by conditional densities from the exponential family (4). At that, the success in constructing the equation is attained not for the signal posterior probability density (2), but directly for optimal mean square estimator (1) itself. To derive this equation let us use the Markovian property of the sequence $Z_k = (S_k, X_k)_{k \geq 1}$

$$f(s_1^n, x_1^n) = f(s_1, x_1) \prod_{k=2}^n g(s_k, x_k | s_{k-1}, x_{k-1}) , \quad (5)$$

where (s_k, x_k) is the value of the random variable (S_k, X_k) , and $f(s_1, x_1)$ and $g(\cdot|\cdot)$ are the prior density and the transition density of $(Z_n)_{n \geq 1}$. Taking into account (5), the interpolating posterior density $\pi_k(s_k|x_1^n)$ is representable in the form

$$\begin{aligned} \pi_k(s_k|x_1^n) &= \frac{1}{f(x_1^n)} f(s_k, x_1^n) = \frac{1}{f(x_1^n)} f(s_k, x_1^k, x_{k+1}^n) \\ &= \frac{1}{f(x_1^n)} \cdot f(x_1^k) w_k(s_k|x_1^k) \cdot f(x_{k+1}^n | s_k, x_k) . \end{aligned} \quad (6)$$

In the future we must choose in equation (6) any factors depending on the observation x_k . For this purpose, the second factor of (6) is transformed as follows

$$f(x_1^k) w_k(s_k|x_1^k) = \int_{\mathbb{R}^m} f(s_1^{k-2}, x_1^{k-2}, s_{k-1}, x_{k-1}) g(s_k, x_k | s_{k-1}, x_{k-1}) ds_{k-1} .$$

But, since for dynamic models the transition density $g(s_k, x_k | s_{k-1}, x_{k-1}) = p(s_k | s_{k-1})f(x_k | x_{k-1}, s_k)$, we get

$$\begin{aligned} f(x_1^k)w_k(s_k | x_1^k) &= f(x_k | x_{k-1}, s_k) \int_{\mathbb{R}^m} p(s_k | s_{k-1})f(s_{k-1}, x_1^{k-1})ds_{k-1} \\ &= f(x_1^{k-1})f(x_k | x_{k-1}, s_k)f(s_k | x_1^{k-1}). \end{aligned} \quad (7)$$

Obviously, only the second factor of (7) depends on x_k .

Let us consider the third factor in (6). Then we get

$$\begin{aligned} f(x_{k+1}^n | s_k, x_k) &= \int_{\mathbb{R}^m} f(x_{k+2}^n | s_{k+1}, x_{k+1})g(s_{k+1}, x_{k+1} | s_k, x_k)ds_{k+1} \\ &= \int_{\mathbb{R}^m} f(x_{k+2}^n | s_{k+1}, x_{k+1})p(s_{k+1} | s_k)f(x_{k+1} | x_k, s_{k+1})ds_{k+1} \\ &= \int_{\mathbb{R}^m} f(x_{k+2}^n, s_{k+1}, x_{k+1}) \frac{p(s_{k+1} | s_k)}{p(s_{k+1})} \frac{f(x_{k+1} | x_k, s_{k+1})}{f(x_{k+1} | s_{k+1})} ds_{k+1}. \end{aligned} \quad (8)$$

Here only the last ratio $f(x_{k+1} | x_k, s_{k+1})/f(x_{k+1} | s_{k+1})$ depends on x_k . It is this dependence on x_k that does not yet permit us to construct a non-parametric version of the equation (2) for dynamic observation models. However, for static observation models of the form $X_n = \varphi(S_n, \eta_n)$, where φ is a measurable function and η_n is an independent noise, the conditional density of observations $f(x_{k+1} | x_k, s_{k+1}) = f(x_{k+1} | s_{k+1})$ does not depend on x_k , and this transforms the equation (6) to the form:

$$\begin{aligned} \pi_k(s_k | x_1^n) &= \frac{1}{f(x_1^n)} \cdot f(x_1^k)w_k(s_k | x_1^k) \cdot f(x_{k+1}^n | s_k, x_k) \\ &= \frac{1}{f(x_1^n)} f(x_k | s_k) f(s_k | x_1^{k-1}) \int_{\mathbb{R}^m} f(x_{k+2}^n, s_{k+1}, x_{k+1}) \frac{p(s_{k+1} | s_k)}{p(s_{k+1})} ds_{k+1} \\ &= \frac{\lambda_k(x_1^n \text{ without } x_k)}{f_k(x_k | x_1^n \text{ without } x_k)} f(x_k | s_k) f(s_k | x_1^{k-1}) f(s_k | x_{k+1}^n) p^{-1}(s_k), \end{aligned} \quad (9)$$

where only the second factor depends on x_k ; some new definitions for the normalizing constants are introduced here:

$$\lambda_k = \lambda_k(x_1^n \text{ without } x_k) \triangleq \frac{f(x_1^{k-1})f(x_{k+1}^n)}{f(x_1^{k-1}, x_{k+1}^n)}, \quad (10)$$

$$f_k(x_k | x_1^n \text{ without } x_k) \triangleq f(x_k | x_1^{k-1}, x_{k+1}^n). \quad (11)$$

Such a form of the interpolation equation allows us to obtain its counterpart independent of the statistical characteristics of the unobserved process $(S_n)_{n \geq 1}$.

Denote

$$u_k(s_k) = f(s_k | x_1^{k-1})f(s_k | x_{k+1}^n)p^{-1}(s_k) \quad (12)$$

and remark once more that u_k is independent of x_k . Then, equation (9) can be represented in the form

$$\pi_k(s_k | x_1^n) = \frac{\lambda_k(x_1^n \text{ without } x_k)}{f_k(x_k | x_1^n \text{ without } x_k)} f(x_k | s_k) u_k(s_k). \quad (13)$$

Let us integrate this equation with respect to s_k and carry over the normalizing factor depending only of the observations to the left-hand side:

$$\frac{f_k(x_k|x_1^n \text{ without } x_k)}{\lambda_k(x_1^n \text{ without } x_k)} = \int_{\mathcal{S}_k} f(x_k|s_k)u_k(s_k)ds_k. \quad (14)$$

Assuming now that the density $f(x_k|s_k)$ belongs to the exponential family (4), differentiate (14) w.r.t. x_k . The possibility of differentiating under the sign of integral is justified by the assumption of existence of the second prior moment $E Q^T(S_k)Q(S_k)$, that is the natural restriction of signal power. The latter is sufficient for existence of the mean square risk. Differentiation w.r.t. x_k results in the equation

$$\frac{\nabla_{x_k} f_k(x_k|x_1^n \text{ without } x_k)}{\lambda_k(x_1^n \text{ without } x_k)} = \int_{\mathcal{S}_k} \nabla_{x_k} f(x_k|s_k)u_k(s_k)ds_k. \quad (15)$$

For the exponential conditional density $f(x_k|s_k)$,

$$\nabla_{x_k} f(x_k|s_k) = (\nabla_{x_k} \log h(x_k) + \nabla_{x_k} T^T(x_k)Q(s_k))f(x_k|s_k). \quad (16)$$

Substituting (16) to (15) and denoting by $Q(\hat{s}_k)$ the integral $\int Q(s_k)\pi_k(s_k|x_1^n)ds_k$, let us find the equation for the optimal mean square estimator $Q(\hat{s}_k)$

$$\mathcal{T}^T(x_k)Q(\hat{s}_k) = \nabla_{x_k} \log \frac{f_k(x_k|x_1^n \text{ without } x_k)}{h(x_k)}, \quad (17)$$

where \mathcal{T} is the Jacobi matrix with elements $\partial T_i / \partial x_k^{[j]}$, $i = 1, \dots, m$, $j = 1, \dots, r$.

The equation (17) is a simple linear vector equation w.r.t. $Q(\hat{s}_k)$, but it can be solved only for a certain density $f_k(x_k|x_1^n \text{ without } x_k)$. In the case where all probability distributions are known, this density can be computed, and the result coincides with (1) and (2). But in the case where $f_k(x_k|x_1^n \text{ without } x_k)$ can not be explicitly computed and even its parametric form is unknown, we restore it from the observations using the kernel non-parametric procedures.

4 Non-Parametric Interpolation Estimator

To solve the problem of interpolating on the basis of one realization x_1^n of a process $(X_k)_{1 \leq k \leq n}$, $X_k \in \mathbb{R}^l$, we can use the asymptotically ε -optimal interpolating procedure from Vasiliev et al. (2004), in which the truncated conditional density $\bar{f}(x_k|x_{k-\tau}^{k-1}, x_{k+1}^{k+\tau})$ is used instead of the conditional density $f(x_k|x_1^n \text{ without } x_k) \triangleq f(x_k|x_1^{k-1}, x_{k+1}^n)$, where the parameter τ is the order of the Markov process which approximates the non-Markovian weak dependent process $(X_n)_{n \geq 1}$. The criteria and methods of estimation of τ which were developed in Vasiliev et al. (2004) with regard to filtration, can be extended to the interpolation problems. In this way, the conditional density $\bar{f}(x_k|x_{k-\tau}^{k-1}, x_{k+1}^{k+\tau})$ can be written as the ratio

$$\bar{f}(x_k|x_{k-\tau}^{k-1}, x_{k+1}^{k+\tau}) = \frac{f(x_{k-\tau}^{k+\tau})}{f(x_{k-\tau}^{k-1}, x_{k+1}^{k+\tau})},$$

where the numerator is the marginal density of $l \times (2\tau + 1)$ -dimensional vector and the denominator is the marginal density of $l \times 2\tau$ -dimensional vector of observations. By substituting the multivariate non-parametric kernel estimators

$$f_N(x_1^n) = \frac{1}{N h_N^{lr}} \sum_{i=1}^N \prod_{k=1}^n \prod_{j=1}^l K \left(\frac{(x_k^{[j]} - X_k^{[j]}(i))}{h_N} \right) \quad (18)$$

for these densities, we get a non-parametric form of the equation (17) in the form

$$\mathcal{I}^T(x_k) Q(\tau \hat{s}_{k,N}) = \frac{\nabla_{x_k} f_N(x_{k-\tau}^{k+\tau})}{f_N(x_{k-\tau}^{k+\tau})} - \frac{\nabla_{x_k} h(x_k)}{h(x_k)}. \quad (19)$$

Interpretation of this equation is quite obvious. To construct the interpolation estimator at the point k , one uses the data before and later of k in a distance not exceeding τ .

The interpolation estimator in the equation (19) is obviously consistent, but it depends on the logarithmic gradient of the conditional probability density which is an unstable functional that takes infinite values, when the denominator is equal to zero. For a more strong convergence, one should construct a piecewise smooth approximation (see Vasiliev et al., 2004) providing the mean-square convergence under some additional regularity conditions.

5 Comparison of Non-Parametric Interpolation Estimator with the Optimal Estimator in the Kalman Scheme

To illustrate the proposed estimator and its performance, consider an example with univariate state and observation models ($m = l = 1$):

$$S_{n+1} = aS_n + b\xi_{n+1}, \quad b^2 = \sigma^2(1 - a^2), \quad (20)$$

$$X_n = AS_n + B\eta_n, \quad S_n, X_n \in \mathbb{R}. \quad (21)$$

Here, S_1 , ξ_n and η_n are the mutually independent random variables with the Gaussian distributions $N(0, \sigma^2)$ for S_1 and $N(0, 1)$ for ξ_n and η_n , $n \geq 1$. The coefficients a , b , A , B are known, $|a| < 1$. With suitable initial conditions, such equations generate a strongly stationary process. For the model (20), (21), the conditional probability density of observations is Gaussian and, therefore, the Kalman filter and the optimal forward and backward recursive linear interpolation equations associated with it can be obtained using the results of Liptser and Shiryaev (1977).

The Kalman filter:

$$\begin{aligned} \hat{S}_{k+1} &= a\hat{S}_k + \frac{Ab^2 + a^2A\gamma_k}{B^2 + A^2b^2 + A^2a^2\gamma_k} [x_{k+1} - Aa\hat{S}_k], \\ \gamma_{k+1} &= \frac{B^2(a^2\gamma_k + b^2)}{A^2(a^2\gamma_k + b^2) + B^2} \end{aligned} \quad (22)$$

with the initial conditions

$$\widehat{S}_1 = \frac{A\sigma^2}{A^2\sigma^2 + B^2}x_1, \quad \gamma_1 = \frac{B^2\sigma^2}{A^2\sigma^2 + B^2},$$

where $\widehat{S}_k = E[S_k|x_1^k]$, $\gamma_k = E[(S_k - \widehat{S}_k)^2|x_1^k]$.

Forward interpolation:

$$\begin{aligned} D_k &= A^2(a^2\gamma_k + \sigma^2(1 - a^2)) + B^2 \\ \widetilde{S}_k &= \widehat{S}_k + Aa\gamma_k(X_{k+1} - Aa\widehat{S}_k)/D_k \\ \widetilde{\gamma}_k &= A^2\sigma^2(1 - a^2) + B^2\gamma_k/D_k, \quad k = 2, \dots, n, \end{aligned} \quad (23)$$

where $\widehat{S}_k = E[S_k|x_1^{k+1}]$, $\widetilde{\gamma}_k = E[(S_k - \widetilde{S}_k)^2|x_1^{k+1}]$.

Backward interpolation:

$$\begin{aligned} \widetilde{\widetilde{S}}_k &= \widetilde{S}_k + \widetilde{\gamma}_k a \sigma^2 (1 - a^2) (\widetilde{\widetilde{S}}_{k+1} - \widetilde{S}_k) / d_k \gamma_{k+1} \\ \widetilde{\widetilde{\gamma}}_k &= \widetilde{\gamma}_k + \widetilde{\gamma}_k^2 (\sigma^2 (1 - a^2))^2 \widetilde{\widetilde{\gamma}}_{k+1} / D_k^2 \gamma_{k+1}, \quad k = 2, \dots, n - 1, \end{aligned} \quad (24)$$

where $\widetilde{\widetilde{S}}_k = E[S_k|x_1^n]$, $\widetilde{\widetilde{\gamma}}_k = E[(S_k - \widetilde{\widetilde{S}}_k)^2|x_1^n]$.

A non-parametric interpolation can be constructed using only one observation equation (21) and the sample of data of size n . In this univariate case, non-parametric interpolation equation (17) is

$$\widehat{S}_k^\tau = \frac{B^2}{A} \frac{\partial/\partial x_k f(x_{k-\tau}^{k+\tau})}{f(x_{k-\tau}^{k+\tau})} + \frac{x_k}{A}. \quad (25)$$

It should be noted that this equation does not involve the parameters a, b of state-space equation (20). The non-parametric kernel estimator for the denominator $f(x_{k-\tau}^{k+\tau})$ in (25) is defined by formula (18). The numerator in (25) contains a partial derivative of the multivariate density $f(x_{k-\tau}^{k+\tau})$ at the point x_k . Usually, the estimator of the density derivative is selected as the derivative of the density estimator. In this case, the bandwidths and kernels in both estimators are coincide. But it is not always so. For instance, if we want to have identical rate of convergence for the derivative estimator and the density estimator in the ratio (25), we need to choose not only different bandwidths, but also different kernels.

The non-parametric estimator for the ratio of the density derivative to the density itself is described by the following expression:

$$\frac{(\partial/\partial x_k) f_N(x_{k-\tau}^{k+\tau})}{f_N(x_{k-\tau}^{k+\tau})} = h_N^{-2} \frac{\sum_{i=\tau+1}^{n-\tau} (x_i - x_k) \prod_{l=\tau}^{\tau} \exp\left(-\frac{(x_{k+l} - x_{i+l})^2}{2h_N^2}\right)}{\sum_{i=\tau+1}^{n-\tau} \prod_{l=\tau}^{\tau} \exp\left(-\frac{(x_{k+l} - x_{i+l})^2}{2h_N^2}\right)}, \quad (26)$$

where $N = n - 2\tau$ and h_N is the bandwidth parameter; its optimal value depends on unknown functions. In the experiment, these parameters are modified to get a “good value” of the risk

$$\widehat{R} = \frac{1}{M} \sum_{j=1}^M \left(\frac{1}{N} \sum_{i=1}^N (S_i(j) - \widehat{S}_i(j))^2 \right)^{1/2}, \quad (27)$$

where M is the number of repeated experiments. This performance measure is the mean square deviation of the process estimator $(\hat{S}_i)_{1 \leq i \leq N}$ from the process of the signal $(S_i)_{1 \leq i \leq N}$ averaged by M experiments. We consider three estimators: Kalman estimator \hat{S}_k from the equation (22) with the risk R_K , optimal backward interpolation \tilde{S}_k from the equation (24) with the risk R_{OI} , and non-parametric interpolation from the equation (25) with the risk R_{NI} . Because

$$R_{OI} \leq R_K \quad \text{and} \quad R_{OI} \leq R_{NI},$$

for convenience of comparison, let us introduce the relative errors in percentage as

$$\varepsilon_K = \frac{R_K - R_{OI}}{R_{OI}} \cdot 100, \quad \varepsilon_{NI} = \frac{R_{NI} - R_{OI}}{R_{OI}} \cdot 100.$$

It shows how much one estimator is better or worse than another one. The simulation results are presented in Figure 1 for $n = 1000$, $\sigma^2 = 2$, $a = 0.7$, $b = 1$, $A = B = 1$, and $\tau = 1$.

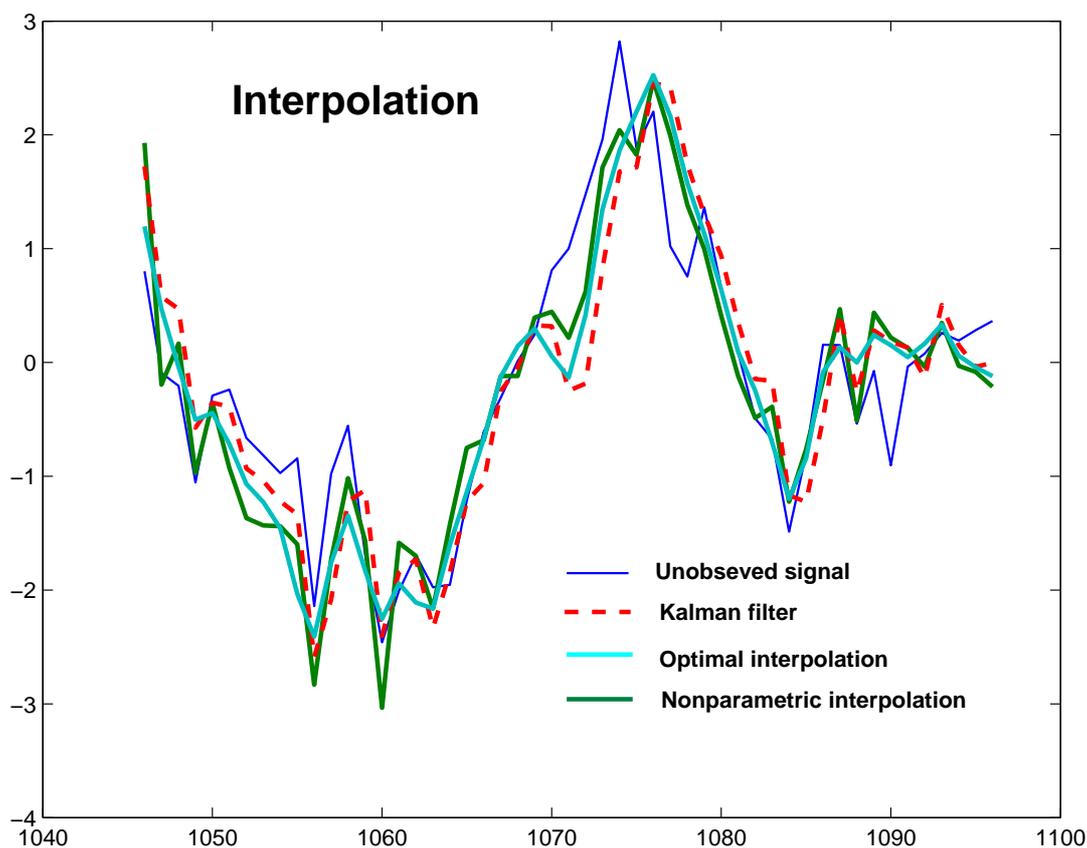


Figure 1: Comparison of the non-parametric smoothing estimator with the optimal estimators

The relative errors ε_K and ε_{NI} are given in Table 1. The simulation shows that non-parametric estimators can be superior to the optimal Kalman filtering estimators by performance, but it is always inferior w.r.t. the optimal backward interpolation. In spite of

Table 1: Relative excess of the empirical risk over the optimal smoothing risk

M	Optimal	Kalman ε_K	Non-par ε_{NI}
50	0%	7.48%	4.42%

existence of some methods for bandwidth parameter selection in non-parametric density estimators (e.g., plug-in or cross-validation), there are no corresponding methods for non-parametric estimation of the density derivative. Therefore, the way for completely automatic bandwidth selection still remains open.

6 Conclusion

The paper presents three interpolating algorithms for estimation of an unobservable signal corrupted by noise on the fixed time interval. By the example of a linear model, the simulation experiments illustrate the performance of the proposed non-parametric interpolation estimator in comparison with the optimal Kalman filtering estimator and with the optimal backward interpolation.

Acknowledgements

The paper was supported by RFFR Grant 06-08-00072-a

References

- Briers, M., Douset, A., and Maskell, S. (2003). Smoothing algorithms for state-space models. *IEEE Transactions on signal processing*, 55, 1-21.
- Chentsov, N. (1972). *Statistical Decision Rules and Optimal Inference* (2nd ed.). Moscow: Nauka (in Russian).
- Khazen, E. (1978). Restoration of the component of a multivariate markov process under the observations of other its components. *Probl. Control Inform. Theory*, 7 (4), 263-275.
- Liptser, R., and Shiryaev, A. (1977). *Statistics of Random Processes I: General Theory (1977). II: Applications (1978)* (2nd ed.). New York: Springer-Verlag.
- Vasiliev, V., Dobrovidov, A., and Koshkin, G. (2004). *Nonparametric Estimation of Functionals of Stationary Sequences Distributions* (2nd ed.). Moscow: Nauka (in Russian).

Author's Address:

Alexandr V. Dobrovidov
 Trapeznikov Institute of Control Sciences
 Russian Academy of Sciences
 6-3-174 Sumskaya, Moscow, 113587, Russia
 E-mail: dobrovid@ipu.rssi.ru