Integr. equ. oper. theory 99 (9999) 0–0
0378-620X/99000-0 $ 1.50+0.20/0

Integral Equations and Operator Theory

Optimal Prediction of Generalized Stationary Processes

Vadim Olshevsky and Lev Sakhnovich

To Israel Gohberg on the occasion of his 75th anniversary with appreciation and friendship

Abstract. Methods for solving optimal filtering and prediction problems for the classical stationary processes are well-known since the late forties. Practice often gives rise to what are not classical stationary processes but generalized ones, e.g., to white noise and to many other examples. Hence it is of interest to carry over optimal prediction and filtering methods to them. For arbitrary generalized stochastic processes this could be a challenging problem. It was shown recently [OS04] that the generalized matched filtering problem can be efficiently solved for a rather general class of S_J*-generalized stationary* processes introduced in [S96]. Here it is observed that the optimal prediction problem admits an efficient solution for a slightly narrower class of T_J*-generalized stationary* processes. Examples indicate that the latter class is wide enough to include white noise, positive frequencies white noise, as well as generalized processes occurring when the smoothing effect gives rise the situation in which the distribution of probabilities may not exist at some time instances. One advantage of the suggested approach is that it connects solving the optimal prediction problem with inverting of the corresponding integral operators S_J. The methods for the latter, e.g., those using the Gohberg-Semençul formula, can be found in the extensive literature, and we include an illustrative example where a computationally efficient solution is feasible.

Mathematics Subject Classification (2000). Primary 60G20, 93E11 ; Secondary 60G25.

Keywords. Generalized Stationary Processes, Prediction, Filtering, Integral Equations, Gohberg-Semençul Formula.

1. Introduction

1.1. Optimal Prediction of Classical Stationary Processes. A complex-valued stochastic process $X(t)$ is called *stationary in the wide sense* (see, e.g., [D53]), if its expectation is a constant,

$$E[X(t)] = const, \qquad -\infty < t < \infty$$

and the correlation function depends only on the difference $(t-s)$, i.e.,

$$K_X(t,s) = E[X(t)\overline{X(s)}] = K_X(t-s).$$

We assume that $E[|X(t)|^2] < \infty$. Let us consider a system with the memory depth of ω that maps the input stochastic process $X(t)$ into the output stochastic process $Y(t)$ in accordance with the following rule:

$$Y(t) = \int_{t-\omega}^{t} X(s)g(t-s)ds, \qquad g(x) \in L(0,w). \tag{1.1}$$

In the optimal prediction problem one needs to find a filter $g(t)$ in

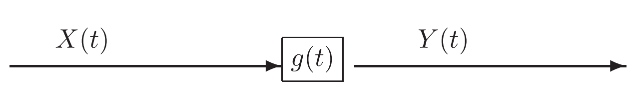

Figure I. Classical Optimal Filter.

This work was supported in part by the NSF contracts 0242518 and 0098222.

so that the output process $Y(t)$ is as close as possible to the true process $X(t+\tau)$ where $\tau > 0$ is a given constant. The measure of closeness is understood in the sense of minimizing the quantity

$$E[|X(t+\tau) - Y(t)|^2].$$

Wiener's seminal monograph [W49] solves the above problem for the case $\omega = \infty$ in (1.1). His results were extended to the case $\omega < \infty$ in [ZR50].

1.2. Generalized stationary processes. Motivation. The *white noise* $X(t)$ (having equal intensity at all frequencies within a broad band) is not a stochastic process in the *classical sense* as defined above. In fact, white noise can be thought of as the *derivative of a Brownian motion*, which is a continuous stationary stochastic process $W(t)$. It can be shown that $W(t)$ is nowhere differentiable, a fact explaining the highly irregular motions that Robert Brown observed. This means that white noise $\frac{dW(t)}{dt}$ does not exist in the ordinary sense. In fact, it is is a *generalized stochastic process* whose definition is stated in [VG61].

Generally, any receiving device has a certain "inertia" and hence instead of actually measuring the classical stochastic process $X(t)$ it measures its averaged value

$$\Phi(\varphi) = \int \varphi(t)X(t)dt, \tag{1.2}$$

where $\varphi(t)$ is a certain function characterizing the device. Small changes in φ yield small changes in $\Phi(\varphi)$ (small changes in the receiving devices yield closer measurements), hence Φ is a continuous linear functional (see (1.2)), i.e., a generalized stochastic process whose definition [VG61].

Hence it is very natural and important to solve the optimal filtering and prediction problems in the case of generalized stochastic processes.This correspondence is a sequel to [OS04] where we solved the problem of constructing the matched filtering problem for generalized stationary processes. Here, formulas for solving the optimal prediction problem are given.

1.3. The main result and the structure of the correspondence. In the next section 2 we recall the definition [VG61] of generalized stationary processes and describe the system action on it. Then in section 3 we introduce a class of T_J-generalized processes for which we will be solving the optimal prediction problem in section 4. Finally, in section 5 we consider a new model of colored noise. It is shown how our general solution to the optimal prediction problem can be a basis to provide a computationally efficient solution in this important example.

2. Generalized Stationary Processes. Auxiliary results

2.1. The Definition of [VG61]**.** Let $\mathcal{K}$ denotes the set of all infinitely differentiable finite functions. Let a stochastic functional Φ (i.e., a functional assigning to any $\varphi(t) \in \mathcal{K}$ a stochastic value $\Phi(\varphi)$) be linear, i.e.,

$$\Phi(\alpha\varphi + \beta\psi) = \alpha\Phi(\varphi) + \beta\Phi(\psi).$$

Let us further assume that all the stochastic values $\Phi(\varphi)$ have expectations given by

$$m(\varphi) = E[\Phi(\varphi)] = \int_{-\infty}^{\infty} x dF(x), \qquad \text{where} \qquad F(x) = P[\Phi(\varphi) \le x].$$

Notice that $m(\varphi)$ is a linear functional acting in the space $\mathcal{K}$ that depends continuously on φ. The bilinear functional

$$B(\varphi, \psi) = E[\Phi(\varphi)\overline{\Phi(\psi)}]$$

is a correlation functional of a stochastic process. It is supposed that that $B(\varphi, \psi)$ is continuously dependent on each of the arguments.

The stochastic process Φ is called *generalized stationary in the wide sense* [VG61], [S97] if for any functions $\varphi(t)$ and $\psi(t)$ from $\mathcal{K}$ and for any number h the equalities

$$m[\varphi(t)] = m[\varphi(t+h)], \tag{2.1}$$

$$B[\varphi(t), \psi(t)] = B[\varphi(t+h), \psi(t+h)] \tag{2.2}$$

hold true.

2.2. System Action on Generalized Stationary Processes. If the processes $X(t)$ and $Y(t)$ were classical, then it is standard to assume that the system action shown in figure 1 obeys

$$Y(t) = \int_0^T X(t-\tau)g(\tau)d\tau, \qquad \text{(where } T = w \text{ is the memory depth).} \tag{2.3}$$

Let us now consider a more general situation when the system shown in figure 2

$$\xrightarrow{\quad\Phi(t)\quad} \boxed{g(t)} \xrightarrow{\quad\Psi(t)\quad}$$

Figure 2. System action on generalized stationary processes

receives the generalized stationary signal Φ that we assume to be zero-mean. Then we define the system action as follows:

$$\Psi(\varphi) = \Phi[\int_0^T g(\tau)\varphi(t+\tau)d\tau], \tag{2.4}$$

The motivation for the latter definition is in that if $X(t)$ and $Y(t)$ were the classical stationary processes then the formula (2.4) for the corresponding functionals

$$\Phi(\varphi) = \int_{-\infty}^{\infty} X(t)\varphi(t)dt, \qquad \Psi(\varphi) = \int_{-\infty}^{\infty} Y(t)\varphi(t)dt \tag{2.5}$$

is equivalent to the classical relation (2.3), so that the former is a natural generalization of the latter.

3. S_J-Generalized and T_J-Generalized Stationary Processes

3.1. Definitions. Let us denote by $\mathcal{K}_J$ the set of the functions from $\mathcal{K}$ such that $\varphi(t) = 0$ when $t \notin J = [a, b]$. The correlation functional $B_J(\varphi, \psi)$ is called a *segment* of the correlation functional $B(\varphi, \psi)$ if

$$B_J(\varphi, \psi) = B(\varphi, \psi), \qquad \varphi, \psi \in \mathcal{K}_J. \tag{3.1}$$

Definition 3.1. Generalized stationary processes are called S_J*-generalized processes* if their segments satisfy

$$B_J(\varphi, \psi) = (S_J\varphi, \psi)_{L^2}, \tag{3.2}$$

where S_J is a bounded nonnegative operator acting in $L^2(a, b)$ having the form

$$S_J\varphi = \frac{d}{dt}\int_a^b \varphi(u)s(t-u)du. \tag{3.3}$$

Here $(\cdot, \cdot)_{L^2}$ is the inner product in the space $L^2(a, b)$.

Formulas for *matched filters* for the above S_J-generalized processes have been recently derived in [OS04]. Here we consider a different problem of *optimal prediction* that is shown to have an efficient solution under additional assumption captured by the next definition.

Definition 3.2. An S_J-generalized process is referred to as a T_J-generalized process if in addition to (3.2) and (3.3) the kernel $s(t)$ of S_J in (3.3) has a continuous derivative (for $t \neq 0$) that we denote by $k(t)$, and moreover

$$s'(t) = k(t) \quad (t \neq 0), \qquad k(0) = \infty. \tag{3.4}$$

3.2. Examples. Before solving the optimal prediction problem we provide some illustrative examples.

Example 3.3. **White noise.** It is well-known that white noise W (which is the derivative of a nowhere differentiable Brownian motion) is not a continuous stochastic process. In fact, it is a generalized stationary process whose correlation functional is known [L68] to be

$$B'(\varphi, \psi) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \delta(t-s)\varphi(t)\overline{\psi(s)}ds$$

Thus, in this case we have $B'(\varphi, \psi) = (\varphi, \psi)_{L_2}$ and hence (3.2) implies that white noise Φ is a very special S_J-generalized stationary process with

$$S_J = I. \tag{3.5}$$

It means that the corresponding kernel function $s(t)$ has the form

$$s(t) = \begin{cases} \frac{1}{2} & t > 0 \\ -\frac{1}{2} & t < 0 \end{cases}$$

In accordance with (3.4) the white noise Φ is a T_J-generalized stationary process.

Example 3.4. **Positive frequencies white noise (PF-white noise).** Observe that the operator

$$S_J f = fD + \frac{j}{\pi} \fint_0^T \frac{f(t)}{x-t} dt, \quad f \in L_2[0,T], \tag{3.6}$$

(where $\fint_0^T$ is the Cauchy Principal Value integral, and $D \geq 1$) defines an S_J-generalized process. When $D > 1$ then S_J is invertible (see example 4.1), and if $D = 1$ then S_J is noninvertible. Notice that if $D = 1$ then the kernel of S_J is the Fourier transform of $f_{PW}(z) = 1$ having equal intensity at all *positive frequencies* and the zero intensity at the negative frequencies (hence the name *PF-white noise*). Observe that the process is, in fact, T_J-generalized. Indeed, it can be shown that the kernel function of S_J has the form

$$s(t) = \begin{cases} \frac{D}{2} + & \frac{i}{\pi} \ln t \quad t > 0 \\ -\frac{D}{2} + & \frac{i}{\pi} \ln |t| \quad t < 0 \end{cases}$$

which implies

$$k(t) = s'(t) = \frac{1}{t} \qquad (t \neq 0), \qquad\qquad k(0) = \infty.$$

In accordance with (3.4) the PF-white noise is a T_J-generalized stationary process.

4. Solution to the Optimal Prediction Problem

As is well known [L68], [M65], in the classical case the solution $g(t)$ of the optimal prediction problem can be found by solving

$$\int_0^w g(u) k_x(u-v) dv = k_x(u+\tau). \tag{4.1}$$

In the generalized case the solution g to the optimal prediction problem can be found via solving a generalization of (4.1),

$$S_J g = k_x(u+\tau) = s'(u+\tau). \tag{4.2}$$

If S_J is invertible then

$$g = S_J^{-1} s'(u+\tau). \tag{4.3}$$

Example 4.1. **PF-white noise revisited.** Here we return to the case considered in example 3.4. It can be shown that if $D > 1$ then S_J in (3.6) is positive definite and invertible with

$$S_J^{-1} f = f(x) D_1 - \beta \frac{j}{\pi} \fint_0^T \left(\frac{t}{T-y}\right)^{j\alpha} \left(\frac{x}{T-x}\right)^{-j\alpha} \frac{f(t)}{x-t} dt, \tag{4.4}$$

where

$$D_1 = \frac{D}{D^2-1}, \qquad \beta = \frac{1}{D^2-1}, \tag{4.5}$$

and the number α is obtained from

$$\cosh \alpha\pi = D \sinh \alpha\pi. \tag{4.6}$$

Clearly, (4.4) and (4.3) solve the optimal prediction problem in this case.

5. Some practical consequences. A connection to the Gohberg-Semençul formula

The main focus of the sections 2 - 4 had mostly a theoretical nature. In this section we indicate that the formula (4.3) offers a novel technique allowing one to work out practical problems. Specifically:

- Filtering problems for *classical* stationary processes typically lead to non-invertible operators S_J, and to find the solution $g(t)$ to the optimal prediction problem one needs to solve (4.1). In the case of *generalized* stationary processes the operator S_J is often invertible and hence there is a better formula (4.3).
- Secondly, the operator S_J in (3.2) and (3.3) can be seen as a new way of modeling colored noise. This model is useful since the existing integral equations literature already describes many particular examples on inverting S_J, either explicitly or numerically. Hence (4.3) solves the corresponding optimal prediction problem.

Before providing one such examples let us rewrite the operator S_J of (3.3),

$$S_J f = \frac{d}{dx} \int_a^b f(t) s(x-t) dt, \qquad f(x) \in L^2(a,b)$$

in a more familiar form.

Proposition 5.1. *Let*

$$a = 0, \qquad b = T, \qquad s(0) - s(-0) = 1, \qquad s'(t) = k(t) \quad (t \neq 0).$$

With these settings S_J can be rewritten as

$$S_J f = f(x) + \int_0^T f(t)k(x-t)dt. \tag{5.1}$$

If $k(t)$ is continuous for $t \neq 0$ then the corresponding process is T_J-generalized.

The integral equations literature (see, e.g., [GS72], [S96]) contains results on the inversion of the operator S_J of the form (5.1). The following theorem is well-known.

Theorem 5.2. **(Gohberg-Semençul)** *Let the operator S_J has the form (5.1) with $k(x) \in L(-w, w)$. If there are two functions $\gamma_\pm(x) \in L(0, w)$ such that*

$$S_J\gamma_+(x) = k(x), \qquad S_J\gamma_-(x) = k(x-w) \tag{5.2}$$

then $S_{[0,w]}$ is invertible in $L^p(0, w)$ ($p \geq 1$) and

$$S_J^{-1} f = f(x) + \int_0^w f(t)\gamma(x,t)dt, \tag{5.3}$$

where $\gamma(x, t)$ is given by (5.4).

$$\gamma(x,t) = \begin{cases} -\gamma_+(x-t) - \displaystyle\int_t^{w+t-x} [\gamma_-(w-s)\gamma_+(s+x-t) - \gamma_+(w-s)\gamma_-(s+x-t)]ds & x > t, \\ -\gamma_-(x-t) - \displaystyle\int_t^{w} [\gamma_-(w-s)\gamma_+(s+x-t) - \gamma_+(w-s)\gamma_-(s+x-t)]ds & x < t \end{cases} \tag{5.4}$$

The latter result leads to a number of interesting special cases when the operator S_J can be explicitly inverted and hence the formula (4.3) solves the optimal prediction problem in these cases. For example, the processes corresponding to $k(x) = |x|^{-h}$, with $0 < h < 1$, or to $k(x) = -\log|x - t|$ are of interest. We elaborate the details for another example next.

Example 5.3. **Colored noise approximated by rational functions. The exponential kernel.** Let us consider *colored noise* approximated by a combination of rational functions with the fixed poles $\{\pm i\alpha_m\}$,

$$f(t) = \sum_{m=1}^N \gamma_m \frac{1}{t^2 + \alpha_m^2}, \qquad \alpha_m > 0, \qquad \gamma_m > 0.$$

Let us use its Fourier transform

$$k(x) = \sum_{m=1}^N \beta_m e^{-\alpha_m|x|}, \qquad \beta_j = \frac{\pi}{\alpha_m}\gamma_m \tag{5.5}$$

to define the operator S_J via (5.1).

Solution to the filtering problem. The situation is exactly the one captured by theorem 5.2 where the operator (5.1) has the special kernel (5.5). A procedure to solve (5.2), and hence to find the inverse of S_J is obtained next.

Theorem 5.4. **(Computational Procedure)** *Let S_J be given by (5.1) and its kernel $k(x)$ has the special form (5.5). Then S_J^{-1} is given by the formula (5.3), (5.4), (5.2), where*

$$\gamma_+(x) = -\gamma(x, 0), \qquad \gamma_-(x) = -\gamma(w - x, 0). \tag{5.6}$$

Here

$$\gamma(x, 0) = G(x) \begin{bmatrix} F_1 \\ F_2 \end{bmatrix}^{-1} B, \tag{5.7}$$

where a $1 \times 2N$ row $G(x)$, $N \times 2N$ matrices F_1, F_2 and a $2N \times 1$ column B are defined by

$$G(x) = \begin{bmatrix} e^{\nu_1 x} & e^{\nu_2 x} & \cdots & e^{\nu_{2N} x} \end{bmatrix}, \qquad F_1 = \begin{bmatrix} \frac{1}{\alpha_i + \nu_k} \end{bmatrix}_{1 \leq i \leq N, 1 \leq k \leq 2N},$$

$$F_2 = \begin{bmatrix} \frac{-e^{\nu_k w}}{\alpha_i - \nu_k} \end{bmatrix}_{1 \leq i \leq N, 1 \leq k \leq 2N}, \qquad B = [\underbrace{1 \ \cdots \ 1}_{N} \ \underbrace{0 \ \cdots \ 0}_{N}].$$

The numbers $\{\nu_k\}$ *are the roots (we assume them to be pairwise different) of the polynomial*

$$Q(z) = P(z) - 2\sum_{m=1}^{N} \delta_m \sum_{s=1}^{m} z^{2(m-s)} \sum_{k=1}^{N} \alpha_k^{2s-1}\beta_k, \tag{5.8}$$

where the numbers $\{\delta_k\}$ *are the coefficients of the polynomial*

$$P(z) = \prod_{m=1}^{N}(z^2 - \alpha_m^2) = \sum_{m=1}^{N} \delta_m z^{2m}. \tag{5.9}$$

Hence, in this important case, the formula (4.3) allows us to find the explicit solution $g(t)$ to the optimal prediction problem via plugging in it the formulas (5.3), (5.4) together with (5.6) (5.7). Notice that there are fast and superfast algorithms to solve the Cauchy-like linear system in (5.7), see, e.g., [O03] and the references therein.

Remark, that the example 5.3 can also be solved by using the Kalman filtering method, see, e.g., [K61, KSH00].

We would like to conclude this paper with an interesting problem of extending the Kalman filtering method to generalized stationary processes. Currently it was done for the simplest case of classical stationary processes with added white noise, such as the case considered in the example 5.3.

We would like to thank the anonymous referee for the very careful reading of the manuscript, and for a number of helpful suggestions.

References

[D53] J.L. Doob, Stochastic processes, Wiley, 1953

[GF74] I. Gohberg and I. A. Feldman, *Convolution equations and projection methods for their solution,* Transl. Math. Monographs , Providence, Rhode Island, 1974.

[GS72] I. Gohberg and A.Semençul, *On the inversion of finite Toeplitz matrices and their continual analogues,* Math. Issled. 7:2, 201-223. (In Russian.)

[H00] S.Haykin, *Communications Systems*, Wiley, 4th edition, 2000.

[K61] T.Kailath, *Lectures on Wiener and kalman filtering*, Springer Verlag, New York, 1961.

[KSH00] T.Kailath, A.S.Sayed, and B.Hassibi , *Linear Estimation*, Prentice Hall, 2000.

[L68] B.R.Levin, *Theoretical Foundations of Statistical Radio Engineering*, Moskow, 1968.

[LM04] Nadav Levanon, Eli Mozeson, *Radar Signals*, Wiley, 2004.

[M65] D. Middleton, *Topics in Communication Theory,* McGraw-Hill, 1965.

[M77] N.I. Muskhelishvili, *Singular Integral Equations,* Aspen Publishers Inc, 1977.

[O03] V. Olshevsky, *Pivoting for structured matrices and rational tangential interpolation*, in Fast Algorithms for Structured Matrices: Theory and Applications, CONM/323, p. 1 - 75, AMS publications, May 2003.

[OS04] V.Olshevsky and L.Sakhnovich, *Matched filtering for generalized Stationary Processes,* 2004, submitted.

[S96] L.A.Sakhnovich, *Integral Equations with Difference Kernels on Finite Intervals,* Operator Theory Series, v.84, Birkhauser Verlag, 1996.

[S97] L.A.Sakhnovich, *Interpolation Theory and its Applications,* Kluwer Academic Publications, 1997.

[VG61] N.Y.Vilenkin and I.M.Gelfand, *Some Applications of Harmonic Analysis*, 1961.

[W49] N.Wiener, *Extrapolation, Interpolation and Smoothing of Stationary Time Series*, Cambridge, 1949.

[ZR50] L.A.Zadeh and J.R.Ragazzini, *An extension of Wiener's Theory of Prediction*, J.Appl.Physics, 1950, 21:7, 645-655.

Vadim Olshevsky
Department of Mathematics
University of Connecticut
Storrs, CT 06269
E-mail: olshevsky@math.uconn.edu

Lev Sakhnovich
Department of Mathematics
University of Connecticut
Storrs, CT 06269
E-mail: lev.sakhnovich@verizon.net