Formalization and Extension of Statistical Linearization Techniques

N.N. Bakhtadze*, K.R. Chernyshov**, V.A. Lototsky***

V.A. Trapeznikov Institute of Control Sciences, 65 Profsoyuznaya, Moscow 117997, Russia
(e-mail: *bahfone@ipu.ru, **myau@ipu.ru, ***lotfone@ipu.ru)

Abstract: The paper deals with extensions of the correlation-based statistical linearization of the input-output mappings of nonlinear systems. The approach is based on considering the stochastic system's input and output as elements of a Hilbert space with an inner product of a general type. A proper choice of the inner product leads to using dispersion functions instead of correlation ones when deriving linearized models. The dispersion functions are much more complete measures of dependence between random processes than the correlation ones and provide eliminating the disadvantages of the correlation-based nonlinear system identification.

Keywords: Statistical linearization, Nonlinear systems, Input/output mappings, Measures of dependence, Dispersion functions.

1. INTRODUCTION


In turn, it is well known that applying the correlation functions frequently leads to negative results under study of nonlinear systems and systems having input random processes possessing a complex structure. This is the case since the correlation functions may vanish even under existence of a deterministic relationship between input and output processes of a considered system (Raißman, 1981, Rényi, 1959). More over, there exist examples when the cross correlation function does not reflect the actual dependence between two random variables even provided the regression of a variable to another one and vise versa is linear (Sarmanov and Bratoevo, 1967). To eliminate the negative phenomena in these cases, it is reasonable to use the technique provided by involving the dispersion functions (Durgaryan and Pashchenko, 1980, Pashchenko, 2006, Rajbman, 1981). The functions are more complete measures of dependence between the input and output processes of a stochastic system and, in particular, can completely handle a deterministic dependence between the processes.

2. PRELIMINARIES

A linearization-like approach is the dispersion linearization (Pashchenko, 2006). The dispersion version of statistical linearization involves the following functions characterizing the measure of nonlinear dependence between random processes:

- the proper cross dispersion function of the random processes \( y(t) \) and \( x(s) \)

\[
\theta_{yx}(t,s) = \mathbf{M} \left( \mathbf{M} \left( \frac{y(t)}{x(s)} \right) - \mathbf{M}y(t) \right)^2,
\]

- and auto dispersion function of the random process \( x(t) \)

\[
\theta_{xx}(t,s) = \mathbf{M} \left( \mathbf{M} \left( \frac{x(t)}{x(t)} \right) - \mathbf{M}x(t) \right)^2,
\]

- the generalized dispersion function of the random processes \( y(t), z(v) \) and \( x(s) \)

\[
\theta_{yxz}(t,v,s) = \mathbf{M} \left( \mathbf{M} \left( \frac{y(t)}{x(s)} \right) - \mathbf{M}y(t) \right) \times \\
\times \left( \mathbf{M} \left( \frac{z(v)}{x(s)} \right) - \mathbf{M}z(v) \right).
\]
Here $\mathbf{M}$ stands for the conditional mathematical expectation.

Within the dispersion linearization technique proposed, the input-output mapping of a studied system is approximated by a dependence of the type

$$y_M(t,u) = \varphi(t) + k\mathbf{M}\left(x(t)\right)/x(u)$$

(1)

or

$$y_M(t,u) = \varphi(t) + k\mathbf{M}\left(y(t)\right)/x(u).$$

(2)

Here $\varphi$ over a designation of a random value or process stands for centering, i.e. subtraction of the mathematical expectation. Centered random value or process means that of for centering, i.e. subtraction of the mathematical expectation.

Under the notations introduced, let us consider a problem of approximation of an unknown input/output mapping of a nonlinear system described generically by (unknown) joint distribution density of its output and input processes

$$\langle y(t), x(s), p(y,x,t,s) \rangle,$$

(3)

using realizations of the system’s output process $y(t)$ from the space $Y(t)$ and input process $x(s)$ from the space $X(s)$.

In the approximating input/output model of system described by density and processes (3),

$$y_M(t) = \varphi(t) + \int_T g(t,s)x(s)ds,$$

(4)

the non-random function $\varphi(t)$ and weight function $g(t,s)$ are subject to determination to meet the conditions

$$\mathbf{M}y_M(t) = \mathbf{M}y(t),$$

(5)

$$\int_T \frac{\varphi(t)}{g(t,s)} \rightarrow \inf_{g(t,s)}.$$  

(6)

From criterion (5), (6), the relations follows:

$$\varphi(t) = \mathbf{M}y(t) = m_y(t),$$

$$\int_T \int_{s=0}^{\infty} \left( y(t), x(s), p(y,x,t,s) \right)$$

(7)

which determine the desired approximation characteristics.

Thus introduced, the space $Z_{t,s}$ is also a linear space $Z(t,s)$ with respect to the operations of addition and multiplication by a Borel function of the same arguments, which is bounded for any fixed argument values.

Let the inner product of a general type

$$\langle z_1(t,s), z_2(t,s) \rangle$$

is given in $Z(t,s)$, with the inner product satisfying the following conditions. The linear space $Z(t,s)$ is a complete metric space with respect to the metrics

$$\rho(\cdot, \cdot) = \| \cdot \|,$$

with

$$\| z(t,s) \| = \sqrt{\langle z(t,s), z(t,s) \rangle}$$

being the norm induced by the inner product. It is assumed with respect to $Z(t,s)$ that the norm of all its elements is finite for any fixed $t$ and $s$. Thus, $Z(t,s)$ is a Hilbert space.

Under the notations introduced, let us consider a problem of approximation of an unknown input/output mapping of a nonlinear system described generically by (unknown) joint distribution density of its output and input processes

$$\langle y(t), x(s), p(y,x,t,s) \rangle,$$

(3)

using realizations of the system's output process $y(t)$ from the space $Y(t)$ and input process $x(s)$ from the space $X(s)$.

In the approximating input/output model of system described by density and processes (3),

$$y_M(t) = \varphi(t) + \int_T g(t,s)x(s)ds,$$

(4)

the non-random function $\varphi(t)$ and weight function $g(t,s)$ are subject to determination to meet the conditions

$$\mathbf{M}y_M(t) = \mathbf{M}y(t),$$

(5)

$$\int_T \frac{\varphi(t)}{g(t,s)} \rightarrow \inf_{g(t,s)}.$$  

(6)

From criterion (5), (6), the relations follows:

$$\varphi(t) = \mathbf{M}y(t) = m_y(t),$$

$$\int_T \int_{s=0}^{\infty} \left( y(t), x(s), p(y,x,t,s) \right)$$

(7)

which determine the desired approximation characteristics.
Let now the inner product in \( Z(t,s) \) be given by the expression
\[
\langle z_1(t,s), z_2(t,s) \rangle = M(z_1(t,s)z_2(t,s)), \quad (8)
\]
and let \( \mathcal{A} \) be a self-adjoint positively defined bounded linear operator mapping the space \( Z(t,s) \) onto itself. Then the expression
\[
Mz_1(t,s), \mathcal{A}z_2(t,s) = \langle z_1(t,s), z_2(t,s) \rangle \mathcal{A}
\]
also determines an inner product in \( Z(t,s) \).

Consider the conditional mathematical expectation of the elements from \( Z(t,s) \) with respect to the section \( s_u \) of the random process \( x(s) \) in a point \( u \). At this case, the linear operator \( \mathcal{A} \) given by
\[
\mathcal{A}(t,s) = M\left\{ z(t,s) / x_u \right\} \quad (10)
\]
is a self-adjoint positively defined bounded operator in the sense of the inner product (8). In fact, let \( p(z_1,t,s) \), \( p(z_2,t,s) \), \( p(x,u) \), \( p(z_1,t,s,u) \), \( p(z_2,t,s,u) \) be the marginal and joint distribution densities of the processes \( z_1(t,s) \), \( z_2(t,s) \) and \( x(u) \). Then by virtue of (10),
\[
\langle \mathcal{A}_1(t,s), z_2(t,s) \rangle = \int \int \left\{ \int z_1 \frac{p(z_1,t,s,u)}{p(x,u)} dz_1 \right\} \frac{z_2 p(z_2,t,s,u) dx dz_2}{p(x,u)} =
\]
\[
= \int z_1 \int \int z_2 \frac{p(z_2,t,s,u)}{p(x,u)} dz_2 \frac{p(z_1,t,s,u) dx dz_1}{p(x,u)} =
\]
\[
= \langle z_1(t,s), \mathcal{A}_2(t,s) \rangle.
\]
The analogous statement is also valid for the multiple regression
\[
\mathcal{A}(t,s) = M\left\{ z(t,s) / x_u \right\}, \quad u \in U \quad (11)
\]
with \( U \) being a finite set of points at the positive semiaxis.

A remark is to be done here. Of course, the operator \( \mathcal{A} \) is positively definite only provided the corresponding conditional mean in (10) is nonzero almost surely for all elements in \( Z(t,s) \), i.e. \( x(t,s) \) and \( x_u \) are not independent. But the latter is determined by the general assumptions about the identifiability of the studied system. Generically, the identifiability of a system whose model is described by an input/output mapping may be considered as the following condition:

**the maximal correlation function** \( S_{xy}(t,s) \) **of the system’s input and output processes**

\[
S_{xy}(t,s) = \sup_{B,C} \frac{\mathbf{D}[y(t)] \mathbf{D}[x'(s)]}{\mathbf{D}[B y(t)] \mathbf{D}[C x'(s)]} \neq 0
\]
is not identically equal to zero.

Here \( \mathbf{D}[\cdot] \) stands for the variance. The supremum is taken all over the nonlinear transformations such that theirs variances exist and are nonzero ones. The optimal transformations exist provided the condition
\[
\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{p^2(y,x,t,s)}{p(x,s)p(y,t)} < \infty \quad \text{holds}
\]
(Rényi, 1959, Sarmanov, 1963a, 1963b). Under conditions (9), (10) by virtue of the properties of the dispersion functions, equation (7) takes the form
\[
\theta_{yxx}(t,s,u) = \int g(t,v) \theta_{yxx}(s,v,u) dv, \quad (12)
\]
with \( \theta_{yxx}(t,s,u) \) being the generalized cross dispersion function,
\[
\theta_{yxx}(t,s,u) = \mathbf{M}\left[ \mathbf{M}\left[ \frac{y(t)}{x(u)} \right] - \mathbf{M}y(t) \right] \times
\]
\[
\times \left[ \mathbf{M}\left[ \frac{x(s)}{x(u)} \right] - \mathbf{M}x(s) \right],
\]
and \( \theta_{xxx}(s,v,u) \) being the generalized dispersion function,
\[
\theta_{xxx}(s,v,u) = \mathbf{M}\left[ \mathbf{M}\left[ \frac{x(s)}{x(u)} \right] - \mathbf{M}x(s) \right] \times
\]
\[
\times \left[ \mathbf{M}\left[ \frac{x(v)}{x(u)} \right] - \mathbf{M}x(v) \right].
\]
The functions are determined by the common expression for the generalized dispersion function \( \theta_{yxx}(t,v,s) \) presented in Section 2.

Under conditions (9), (10) the dispersion functions in expression (12) are replaced with the corresponding multiple dispersion functions \( \theta_{yxx}(t,s,u) \) and \( \theta_{xxt}(s,v,u) \).

When the inner product in \( Z(t,s) \) is given by expressions (9), (10) or (9), (11), linearization criterion (6) takes the form
\[
\theta_{ext}(t,u) \to \inf_{g(t,s)} \epsilon(t), \quad \text{or} \quad \theta_{ext}(t,u) \to \inf_{g(t,s)} \epsilon(t), \quad (13)
\]
with \( \epsilon(t) = y(t) - y_M(t) \) being the linearization error. Expression (13) is a form of the dispersion identification criterion (Durgaryan and Pashchenko, 1974, Pashchenko, 2006). Following this, the approach proposed is natural to be referred as the statistical linearization with the dispersion criterion.
4. MULTI INPUT SYSTEMS

The approach derived is naturally expanded to the statistical linearization of multi input systems with input processes \( x_1(s), \ldots, x_n(s) \) which are considered as elements of the space \( \mathcal{X}(s) \). Following the setting, to approximate an input-output relationship of a nonlinear system, described, as above, by (unknown) joint distribution density of the output and input processes,

\[
\langle y(t), x_1(s), \ldots, x_n(s), p(y, x_1, \ldots, x_n, t, s) \rangle, \tag{14}
\]

the model

\[
y_M(t) = \varphi(t) + \sum_{i=1}^{n} g_i(t, s) x_i(s) ds
\]

is applied. In the model, the functions \( \varphi(t) \) and \( g_i(t, s), \ i=1, \ldots, n \), are subject to determination from the conditions

\[
\mathbf{M}_{yM}(t) = \mathbf{M}_y(t), \tag{16}
\]

\[
\inf_{s \in (t, \infty]} \| y_M(t) - y(t) \|_{g_i(t, s), t=1, \ldots, n}. \tag{17}
\]

From criterion (16), (17), the relationships follow:

\[
\varphi(t) = \mathbf{M}_y(t) = m_y(t), \tag{18}
\]

determining the required approximation characteristics.

Consider a linear self-adjoint positively defined operator \( \mathcal{A} \) mapping \( Z(t, s) \) onto itself and given by the relationship

\[
\mathcal{A}(t, s) = \mathbf{M} \left( \frac{z(t, s)}{x_{lu}, \ldots, x_{nu}} \right), \tag{19}
\]

with \( x_{lu} \) being the section of the random process \( x_j(s) \) in the point \( u, \ i=1, \ldots, n \). Then under conditions (8), (9), (19) by virtue of the properties of the dispersion functions, equations (18) take correspondingly the form

\[
\theta_{yxt}(t, s, u) = \sum_{j=1}^{n} g_j(t, s) \theta_{x_jxt}(s, v, u) dv, \ i=1, \ldots, n,
\]

with \( \theta_{xxt}(\cdot, \cdot, \cdot) \) being the corresponding multiple dispersion functions which are determined by the multiple regression

\[
\mathbf{M} \left( \frac{\mathbf{x}_{lu}, \ldots, x_{nu}}{x_{lu}, \ldots, x_{nu}} \right).
\]

5. CRITERION OF COINCIDENCE OF THE NORMS OF OUTPUTS

Another criterion to select the functions \( \varphi(t) \) and \( g(t, s) \) in model (3) is the condition of coincidence of the mathematical expectations and norms of the system’s and model’s outputs, i.e.

\[
\mathbf{M}_{yM}(t) = \mathbf{M}_y(t), \tag{20}
\]

From criterion (20), the relationships follow

\[
\varphi(t) = m_y(t), \tag{21}
\]

\[
\| y(t) \|^2 = \int_{T} g(t, s) g(t, v) \theta_{xxt}(s, v, u) dv ds dv.
\]

which determine the required approximation characteristics.

Under conditions (8)-(10), equation (21) is written in the form

\[
\theta_{yxt}(t, u) = \int_{T} g(t, s) g(t, v) \theta_{xxt}(s, v, u) dv ds dv. \tag{22}
\]

6. STATIONARY SYSTEMS AND ESTIMATION OF THE DISPERSION FUNCTIONS

Let now \( y(t) \) from the space \( \mathcal{Y}(t) \) and \( x(s) \) from the space \( \mathcal{X}(s) \) be stationary and joint stationary in the strict sense random processes. For the case, the linearized model is searched in the form

\[
y_M(t) = \varphi + \int_{0}^{\infty} g(v) x(t-\tau) d\tau,
\]

with \( \tau = t-s \), and the weight function \( g(\tau) \) vanishes as \( \tau < 0 \). Then, for the stationary model, it finally follows

\[
\varphi = M_y(t) = m_y, \tag{23}
\]

\[
\theta_{yxt}(t-u, v-u) = \int_{0}^{\infty} g(\tau) \theta_{xxt}(t-u-\tau, v-u) d\tau, \tag{24}
\]

\[
0 \leq t < \infty.
\]
where the arguments of the time are presented in the form providing an evidence of representation, since

\[
\theta_{yxz}(t-u,v-u) = M \left( M^{o} \frac{y(t)}{x(u)} \right) \times \left( M^{o} \frac{x(v)}{x(u)} \right),
\]

\[
\theta_{xxz}(t-u,v-u) = M \left( M^{o} \frac{x(t-r)}{x(u)} \right) \times \left( M^{o} \frac{x(v)}{x(u)} \right).
\]

Within the section worth while to note also the most general dispersion function, the R-function (Durgaryan and Pashchenko, 1980, Pashchenko, 2006) of the stationary and joint stationary in the strict sense ergodic centered random processes \( y(t+s+v), z(t+s), x(t+s), w(t) \). It is, generically, a function in three arguments (what, formally, was not pointed out in literature):

\[
R_{yzzw}(v,\sigma,\tau) = M \left( M^{o} \frac{y(t+s+v)}{x(t+s)} \right) \times \left( M^{o} \frac{z(t+\sigma)}{w(t)} \right).
\]

where the argument \( v \) is formed as \( (t+s+v)-(t+s) \), the argument \( \sigma \) is formed as \( (t+\sigma)-(t) \), and the argument \( \tau \) is formed as \( (t+s)-(t) \).

7. CONCLUSIONS

Within the approach derived in the paper, the statistical linearization of nonlinear dynamic systems is treated as a search of optimal in the sense of given criteria linear transformations in the Hilbert space of random functions \( Z(t,s) \). A specific selection of the type of the inner product in \( Z(t,s) \) gives rise to various variants of models within a unique representation of the linearized models, with selection of the inner product as the mixed moment leading to the ordinary statistical linearization based on using the correlation functions of the input and output processes of the system. Since the correlation functions may vanish even under existence of deterministic dependence between random values/processes, applying these measures of dependence may correspondingly lead to unacceptable results.

In the paper, selecting the inner product in \( Z(t,s) \) which provides to derive required characteristics of linearized models, based on the dispersion relationships involving dispersion functions, i.e. the covariances of conditional means, has been proposed. The dispersion functions provide handling the nonlinear structure of a system and its input and output processes, and hence deriving adequate models linearized.

REFERENCES


