PUBLICATIONS DE L'INSTITUT MATHÉMATIQUE
Nouvelle série tome 51 (65), 1992, 121–124

ONE CASE OF REDUCTION OF NONLINEAR REGRESSION TO LINEAR

Z. Ivković

Abstract. An effective procedure to find nonlinear regression X^* of $X = X(\xi_1, \dots, \xi_M)$ on $\eta_1, \dots, \eta_N$ in terms of the linear regression coefficients of ξ_k on $\eta_1, \dots, \eta_N$ is proposed. The variables $\xi_1, \dots, \xi_M, \eta_1, \dots, \eta_N$ are Gaussian. Some applications to N-ple Markov process are considered too.

Let $\xi_1, \dots, \xi_M, \eta_1, \dots, \eta_N$ be a finite set of real Gaussian variables centered at the expectations. Let $\mathcal{A}$ ($\mathcal{B}$) be the set of all random variables with finite variances and measurable with respect to $\xi_1, \dots, \xi_M$ $(\eta_1, \dots, \eta_N)$. Nonlinear regression X^* of X on $\eta_1, \dots, \eta_N$ is the conditional expectation $X^* = E(X \mid \eta_1, \dots, \eta_N)$. It is well-known that X^* is the projection of X onto $\mathcal{B}$.

The set of all polynomials in $\xi_1, \dots, \xi_M$ $(\eta_1, \dots, \eta_N)$ is complete (in the mean-square convergence) in $\mathcal{A}$ ($\mathcal{B}$). We have, using multidimensional Hermite polynomials H_p, the following representation

$$X = \sum_{p=1}^{\infty} \sum_{\alpha} A(\alpha) H_p(\xi(\alpha)), \qquad X^* = \sum_{p=1}^{\infty} \sum_{\beta} B(\beta) H_p(\eta(\beta)). \tag{1}$$

The letter α (β) represents a combination with repetitions of elements of $\{1, \dots, M\}$ ($\{1, \dots, N\}$), p at a time. So

$$\xi(\alpha) = \underbrace{\xi_1, \dots, \xi_1}_{k_1(\alpha)}, \dots, \underbrace{\xi_n, \dots, \xi_n}_{k_n(\alpha)}, \qquad 0 \le k_i(\alpha) \le p,\ k_1 + \dots + k_n = p.$$

In this paper we propose an effective procedure to find X^* in terms of the coefficients a_{ij}, $1 \le i \le M$, $1 \le j \le N$, of the linear (coinciding to nonlinear) regression $\xi_k^* = a_{k_1}\eta_1 + \dots + a_{k_N}\eta_N$ of ξ_k on $\eta_1, \dots, \eta_N$. The expression of a_{ij} in terms of the covariance matrix of $\xi_1, \dots, \xi_M, \eta_1, \dots, \eta_N$ is well-known. Multidimensional Hermite polynomials are the tool in the paper. Some of their properties are given in Appendix (see, for instance, [**4**] and [**2**]).

Applying the property (9) to (1), we have $X^* = \sum_{p=1}^{\infty} \sum_{\alpha} A(\alpha) H_p(\xi^*(\alpha))$.

AMS Subject Classification (1985): Primary 62 J 02, Secondary 60 J 25

Proposition.

$$H_p(\xi^*(\alpha)) = \sum\nolimits^{(M)} \frac{k_1(\alpha)}{l_{11}!\ldots l_{1N}!} a_{11}^{l(11)} \cdot \ldots \cdot a_{1N}^{l(1N)} \cdot \ldots \cdot \frac{k_n(\alpha)}{l_{M1}!\ldots l_{MN}!} a_{M1}^{l(M1)} \times$$
$$\times \ldots \cdot a_{MN}^{l(MN)} \cdot H_p(\underbrace{\eta_1,\ldots,\eta_1}_{l(11)+\cdots+l(M1)},\ldots,\underbrace{\eta_N,\ldots,\eta_N}_{l(1N)+\cdots+l(MN)}). \quad (2)$$

where $\sum^{(M)}$ *is the sum over all* $0 \le l_{1j} \le k_1(\alpha)$, $l_{11}+\cdots+l_{1N} = k_1(\alpha)$, $\ldots$, $0 \le l_{nj} \le k_n(\alpha)$, $l_{M1}+\cdots+l_{MN} = k_M(\alpha)$.

Proof. We consider first a particular α_1 where $k_1(\alpha_1) = p$. For an easier presentanion of the proof let us introduce the notations: $\xi_1 = \xi$, $\xi^* = a_1\eta_1 + \cdots + a_N\eta_N$. Consider the one-dimensional Hermite polynomial

$$H_p(\xi^*(\alpha_1)) = H_p(\underbrace{a_1\eta_1+\cdots+a_N\eta_N,\ldots,a_1\eta_1+\cdots+a_N\eta_N}_{p}). \quad (3)$$

The polynomial (3) is, according to the rule of addition (8), the sum of Hermite polynomials whose arguments are variations with repetitions of elements of $\{a_1\eta_1,\ldots,a_N\eta_N\}$, p at a time. Since an Hermite polynomial is a symmetric function of its arguments, this sum is

$$\sum\nolimits^{(1)} \frac{p!}{l_1!\ldots l_N!} H_p(\underbrace{a_1\eta_1,\ldots,a_1\eta_1}_{l(1)},\ldots,\underbrace{a_N\eta_N,\ldots,a_N\eta_N}_{l(N)}),$$

where $\sum^{(1)}$ is the sum over all combinations with repetitions of elements of $\{a_1\eta_1,\ldots,a_N\eta_N\}$, p at a time. By applying the property (7) we have

$$H_p(\xi^*(\alpha_1)) = \sum\nolimits^{(1)} \frac{p!}{l_1!\ldots l_N!} a_1^{l(1)} \cdot \ldots \cdot a_N^{l(N)} H_p(\underbrace{\eta_1,\ldots,\eta_1}_{l(1)},\ldots,\underbrace{\eta_N,\ldots,\eta_N}_{l(N)}).$$

In the general case, a variation with repetitions of elements $a_{ij}\eta_j$, $i = 1,\ldots,M$, $j = 1,\ldots,N$, p at a time, contains l_{11} of $a_{11}\eta_1$, $\ldots$, l_{M1} of $a_{M1}\eta_1$, l_{1N} of $a_{1N}\eta_N$, $\ldots$, l_{MN} of $a_{MN}\eta_N$. There are $k_1!/(l_{11}!\ldots l_{1N}!)\ \ldots\ k_n!/(l_{M1}!\ldots l_{MN}!)$ such summands, since the Hermite polynomial is a symmetric function. We see, by applying (7), that the argument η_1 appears $l_{11}+\cdots l_{M1}$ times, $\ldots$, the argument η_N appears $l_{1N}+\cdots+l_{MN}$ times. $\square$

The expression (2) for $H_p(\xi^*(\alpha))$ motivates us to introduce the following notation

$$H_p(\xi^*(\alpha)) = H_p[(a_{11}\eta_1+\cdots+a_{1N}\eta_N)^{k(1,\alpha)}\ \ldots\ (a_{M1}\eta_1+\cdots+a_{MN}\eta_N)^{k(M,\alpha)}]. \quad (4)$$

The final expression for nonlinear regresion X^* of X on $\eta_1,\ldots,\eta_N$ is

$$X^* = \sum_{p=1}^{\infty}\sum_{\alpha} A(\alpha)\cdot H_p[(a_{11}\eta_1+\cdots+a_{1N}\eta_N)^{k(1,\alpha)}$$
$$\ldots (a_{M1}\eta_1+\cdots+a_{MN}\eta_N)^{k(M,N)}]. \quad (5)$$

As an application consider the process $\{Y_n(t), t \geq 0\}$ defined by $Y_n(t) = \xi^n(t) - E\xi^n(t)$. The process $\{\xi(t), t \geq 0\}$ is a Gaussian N-ple Markov process with the linear prediction ξ^* of $\xi(t)$ by $\{\xi(u), u \leq s\}$, $s < t$, of the form

$$\xi^* = a_0(s,t)\xi(s) + a_1(s,t)\xi'(s) + \cdots + a_{N-1}(s,t)\xi^{(N-1)}(s).$$

The nonlinear prediction Y_n^* of $Y_n(t)$ by $\{Y_n(u), u \leq s\}$, was studied in [**2**] and [**3**]. We present here an explicit expresion for Y_n^* using (5). It is easy to expresss $Y_n((t))$ as a linear combination of one-dimensional Hermite polynomials $Y_n(t) = H_n(\xi(t)) + A_{n-2}(t)H_{n-2}(\xi(t)) + \cdots$. We have

$$\begin{aligned} Y_n^* &= H_n[(a_0\xi(s) + \cdots + a_{N-1}\xi^{(N-1)}(s))^n] \\ &\quad + A_{n-2}H_{n-2}[(a_0\xi(s) + \cdots + a_{N-1}\xi(s))^{n-2}] + \cdots . \end{aligned}$$

It remains only to express $\xi(s), \xi'(s), \ldots, \xi^{(N-1)}(s)$ by $Y(s), Y'(s), \ldots, Y^{(N-1)}(s)$.

It is interesting to compare mean-square error $d_n^2 = E(Y_n(t) - Y^*)^2$ of the nonlinear prediction and the mean-squre error $d_1^2 = E(\xi(t) - \xi^*)^2$ of the linear prediction. Since $EH_p^2(\xi) = \|H_p(\xi)\|^2 = p!\,b^{2p}$, $b^2 = \|\xi\|^2$, it follows that ($b_1^2(s,t) = \|\xi^*\|^2$)

$$\begin{aligned} d_n^2 &= \|Y_n(t) - Y^*\|^2 = \|Y_n(t)\|^2 - \|Y^*\|^2 \\ &= \left(\|H_n(\xi(t))\|^2 - \|H_n(\xi^*)\|^2\right) + A_{n-2}^2\left(\|H_{n-2}(\xi(t))\|^2 - \|H_{n-2}(\xi^*)\|^2\right) \\ &= n!\,(b^{2n} - b_1^{2n}) + A_{n-2}^2(n-2)!\,(b^{2(n-2)} - b_1^{2(n-2)}) \\ &= (b^2 - b_1^2)\left(n!\sum_{k=0}^{n-1} b^{2k}b_1^{2(n-1-k)} + A_{n-2}^2(n-2)!\sum_{k=0}^{2(n-3-k)} b^{2k}b_1^{2(n-3-k)}\right). \end{aligned}$$

Hence, $d_n^2 = B_n(s,t)d_1^2$, $B_n(s,t) > 0$.

As an example, let $\xi(t) = \int_0^t (t-u)\,dW(u)$ be the proper canonical representation [**1**] of $\{\xi(t)\}$, where $\{W(t), t \geq 0\}$ is the Wiener process. Then

$$\begin{aligned} \xi^* &= \int_0^s (t-u)\,dW(u) = \xi(s) + (t-s)\xi'(s), \\ b^2(t) &= t^3/3, \qquad b_1^2(s,t) = t^2 s - ts^2 + s^3/3. \end{aligned}$$

Let $n = 4$. From $H_4(\xi(t)) = \xi^4(t) - 6b^2\xi(t) + 3b^4$, we have $Y_4(t) = H_4(\xi(t)) + 6b^2H_2(\xi(t))$ and $d_4^2 = 24(b^2 + b_1^2)(4b^2 + b_1^2)d_1^2$.

Appendix. The explicit expression for $H_p(\xi_1, \ldots, \xi_p)$ is

$$\xi_1 \ldots \xi_p - \sum_{k=i,j} b_{i(1)j(1)}\xi_{k(1)} \ldots \xi_{k(p-2)} + \sum_{k\neq i,j} b_{i(1)j(1)}b_{i(2)j(2)}\xi_{k(1)} \ldots \xi_{k(p-4)} - \cdots,$$
$$b_{ij} = \operatorname{cov}(\xi_i, \xi_j), \qquad (6)$$

where the first sum is over all combinations (i,j) of elements of the set $\{1, \ldots, p\}$ and the second sum is over all disjoint pairs of (i_1, j_1), (i_2, j_2) and so on. For example, $H_2(\xi_1, \xi_2) = \xi_1\xi_2 - b_{12}$, $H_3(\xi_1, \xi_2, \xi_3) = \xi_1\xi_2\xi_3 - b_{23}\xi_1 - b_{13}\xi_2 - b_{12}\xi_3$.

For a real a we have

$$H_p(a\xi_1, \xi_2, \dots, \xi_p) = aH_p(\xi_1, \dots, \xi_p). \tag{7}$$

It follows immediately from (6) that $H_p(\eta_1 + \eta_2, \xi_2, \dots, \xi_p) = H_p(\xi_1, \eta_2, \dots, \eta_p) + H_p(\eta_2, \xi_2, \dots, \xi_p)$. Generalizing this property of addition, we have

$$H_p\Big(\sum\nolimits_1^I \xi_i, \sum\nolimits_1^J \eta_j, \dots, \sum\nolimits_1^K \zeta_k\Big) = \sum H_p(\xi(i), \eta(j), \dots, \zeta(k)), \tag{8}$$

where the sum is over all arrangements ($I \cdot J \cdot \ldots \cdot K$ in numbers) of indices $i, j, \dots, k$. Finally, let $\{\xi(t), t \in T\}$ be a set of Gaussian variables. Denote by ξS the σ-algebra generated by $\{\xi(t), t \in S\}$, $S \subset T$. We have [**2**]:

$$E(H_p(\xi(t_1), \dots, \xi(t_p)) \mid \xi S) = H_p(E(\xi(t_1) \mid \xi S),\ \dots,\ E(\xi(t_p) \mid \xi S)). \tag{9}$$

REFERENCES

[**1**] T. Hida and G. Kallianpur, *The square of a Gaussian Markov process and non-linear prediction*, J. Multivariate Anal. **5** (1975), 451–461.

[**2**] Z. Ivković and Z. Lozanov, *Hermite polynomiaqls of Gaussian process*, Rev. Res. Math. Ser., Fac. of Sci., Novi Sad, **12** (1982), 105–115.

[**3**] Z. Ivković, *Non-linear prediction of the degreen of a gaussian N-ple Markov process*, J. Multivariate Anal. **19** (1986), 327–329.

[**4**] Ю. А. Розанов, *Гауссовские бесконечномерные распределения*, Труды Мат. ин.-та им. В. А. Стеклова **108** (1968).

Matematički fakultet
Studentski trg
11001 Beograd, p,p. 550
Jugoslavija

(Received 15 02 1991)