The Annals of Mathematical Statistics
1970, Vol. 41, No. 3, 956–969

THE CONSISTENCY OF NONLINEAR REGRESSIONS[1]

By E. Malinvaud

1. Introduction. A sample of T observations on the variables x, z_1, $z_2, \cdots$, $z_m (j = 1, 2 \cdots m)$ has been generated according to the model:

$$x_t = g(z_{1t}, z_{2t} \cdots z_{mt}; \alpha_1, \alpha_2 \cdots \alpha_p) + \varepsilon_t, \tag{1}$$

in which x_t, z_{1t}, $z_{2t} \cdots z_{mt}$ designate the values taken by the variables in observation $t (t = 1, 2 \cdots T)$, $\alpha_1, \alpha_2 \cdots \alpha_p$ are p unknown parameters to be estimated ($k = 1, 2 \cdots p$), ε_t is an unobservable random variable with zero expected value and g is a known function of its $m+p$ arguments. By a regression on model (1) we mean the computation of estimates $\hat{\alpha}_1, \hat{\alpha}_2 \cdots \hat{\alpha}_p$ that minimize the mean square deviation of x from g:

$$T^{-1} L_T(\alpha) = T^{-1} \sum_{t=1}^{T} [x_t - g(z_{1t} \cdots z_{mt}; \alpha_1 \cdots \alpha_p)]^2. \tag{2}$$

The regression is said to be linear if the function g is linear in the vector α with components $\alpha_1, \alpha_2 \cdots \alpha_p$, a situation extensively explored in the literature.

Cases of model (1) with nonlinear functions g occur in various areas of applied statistics. They are frequent in econometrics, where the model often results from rather specific theories implying special forms for the dependence between the variables z_{jt} considered as exogenous and the variable x_t taken to be endogenous.

An important particular case of model (1) occurs when the function g is linear with respect to the variables z_{jt}:

$$x_t = \sum_{j=1}^{m} a_j(\alpha_1 \cdots \alpha_p) \cdot z_{jt} + \varepsilon_t \tag{3}$$

a_j being known functions of the parameters. In econometrics, distributed lag models are of this type, whereas any overidentified simultaneous equation system has a reduced form corresponding to the multivariate generalization of model (3), a generalization about which I shall make some comments in the concluding section.

The regression method for estimating the α_k is of widespread use and has some a priori appeal.[2] As is well known, important properties of linear regressions do not depend on the particular form assumed by the distribution of the random term ε_t, a feature that has definite advantages in most fields of applied statistics and that extends to asymptotic properties of nonlinear regressions.

Such being the case one may be surprised to realize how little developed is the statistical theory of nonlinear regressions. Research has been concentrated on the

---

Received August 23, 1967; revised June 16, 1969.

[1] This paper was written in the spring of 1967 during a visit at Berkeley and was supported by the Office of Naval Research under Contract Nonr-222(77) with the University of California.

[2] I probably should mention here, even though I shall not refer to it again, that the regression method gives maximum likelihood estimates if the ε_t are normally, independently and identically distributed with zero expected value.