

Theory of regression models with increasing number of unknown parameters.

Asaf Hajiyev

Department of Statistics,

Institute of Cybernetics,

Academy of Sciences,

F.Agayev str.,9, Baku,

370141, Azerbaijan

asaf@baku-az.net

It is considered a regression model $y_i = f(x_i, q) + e_i$, $i = 1, 2, \dots, N$; where

$f(x_i, q)$ is a some function relatively to unknown parameter q ;

e_i - is an error of observations, $q = (q_1, q_2, \dots, q_m)^T$ is the vector of unknown parameters.

It is assumed:

number of unknown parameters (m) depends on (N) and

$[m(N)/N] \rightarrow 0$, as $N \rightarrow \infty$ (1)

$E e_i = 0$, $E e_i^2 = s_i^2$ - are different, unknown,

but bounded ($s_i^2 \leq s_0^2$) (2)

Let $X(q)$ be a matrix of experiment and $l_1(q), l_2(q), \dots, l_m(q)$ are eigen values of the matrix $[X^T X/N]$, q^* is least square estimator of q .

Linear model. Let

$f(x_i, q) = q_1 j_1(x) + q_2 j_2(x) + \dots + q_m j_m(x)$, where $j_1(x), j_2(x), \dots, j_m(x)$ is a system of linearly independent functions.

Denote $C_N = \|c_{ij}\|$ covariance matrix of the vector $N^{1/2}(q - q^*)$,

$Y^* = Xq^*$, $c_{kl}^* = (y - y^*)^T I_{kl}(x) (y - y^*)$

For linear regression model if

$[m^{3/2}(N)l_1(q)]/N \rightarrow 0$, as $N \rightarrow \infty$

P

then $(q - q^*) \rightarrow 0$, $(c_{kl}^* - c_{kl}) \rightarrow 0$, as $N \rightarrow \infty$. [see, 1].

Similar problem with finite value of m investigated in [2].

The problem is to find least square estimator for nonlinear regression model and estimate the elements of covariance matrix.

Let $f(x_i, q)$ be nonlinear function of q and $F(q)$ is a matrix with elements

$$f_{ij}(q) = \frac{\partial f(x_j, q)}{\partial q_i}, \quad i=1, \dots, m; \quad j=1, \dots, N;$$

Then least square estimator is the solution of the following iterated process

$$q(s+1) = q(s) + [F^T(q(s)) F(q(s))]^{-1} F^T(q(s))(y - f(x, q(s))) \quad (3)$$

Let $B(r)$ is a sphere with radius $r > 0$ as a center at the point q .

Theorem 1. Under conditions (1), (2), if $q(0) \in B(r)$ and $m^{5/4} / [N(l_1(q))^3] \rightarrow 0$, as $N \rightarrow \infty$

then there exists the random variable $q_N(\infty)$ such that

$$(q_N(\infty) - q_N(s)) \rightarrow 0, \quad (\text{on probability}) \text{ as } s \rightarrow \infty, \quad r > 0.$$

In the capacity of least square estimator is considered $q_N(\infty)$.

Moreover, under conditions of theorem

$$(q_N(\infty) - q) \rightarrow 0, \quad (\text{on probability}) \text{ as } N \rightarrow \infty.$$

Further under of some conditions it is found consistent and asymptotically unbiased estimator of the elements of covariance matrix C_N .

These results allow to construct a confidence band for unknown function in nonlinear regression models.

References.

1. A. Hajiyev (1999) Proc. Inst. Mathem. Mech. v. X(XV111).
2. C. F. Wu (1986) Ann. Statist. v. 9, no. 3.

Resume.

For regression models with increasing number of unknown parameters constructed the iterated process for finding of least square estimator. Consistent and asymptotically unbiased estimator of the elements of covariance matrix is found. The results allow to construct a confidence band for unknown function in regression (linear and nonlinear) model with increasing number of unknown parameters and unknown variance of errors.