Minimax Approach to Errors-in-Variables Linear Models

Yu. Golubev*

The talk deals with estimating unknown parameters $a, b \in \mathbb{R}$ in the simple Errors-in-Variables (EiV) linear model

$$Y_i = a + bX_i + \epsilon\xi_i,$$
$$Z_i = X_i + \sigma\zeta_i,$$

where $i = 1, \dots, n$, ξ_i, ζ_i are i.i.d. standard Gaussian random variables, $X_i \in \mathbb{R}$ are unknown nuisance variables, and ϵ, σ are noise levels which are assumed to be known. It is well known that the standard Maximum Likelihood (ML) estimates of a and b have unbounded moments. In order to improve the ML method and to construct good statistical estimates, we study the EiV model in Large Noise Regime assuming that as $n \to \infty$

$$\epsilon^2 = \epsilon_n^2 = (1 + o(1))\sqrt{n}\epsilon_\circ^2, \quad \sigma^2 = \sigma_n^2 = (1 + o(1))\sqrt{n}\sigma_\circ^2, \quad \text{with } \epsilon_\circ^2, \sigma_\circ^2 > 0.$$

Under these assumptions, we define the asymptotically minimax risk

$$r = \lim_{n\to\infty} \inf_{\tilde{b}} \sup_{X_i, a, b} \frac{n\mathbf{Var}_n^2(X)}{\sigma_n^2\epsilon_n^2}\mathbf{E}\big[\tilde{b} - b\big]^2,$$

where inf is taken over all estimate of b and

$$\mathbf{Var}_n(X) = \frac{1}{n}\sum_{i=1}^n (X_i - \bar{X})^2, \quad \bar{X} = \frac{1}{n}\sum_{i=1}^n X_i.$$

It is proved that

$$r = 1$$

and that the minimax estimate of b is a solution to a convex optimization problem. A fast algorithm for computing this estimate is also proposed.

*CNRS, Aix-Marseille Université, I2M, UMR 7373, 13453 Marseille, France.