1

Nonparametric estimators

1.1 Examples of nonparametric models and problems

1. Estimation of a probability density

Let $X_1, \ldots, X_n$ be identically distributed real valued random variables whose common distribution is absolutely continuous with respect to the Lebesgue measure on $\mathbf{R}$. The density of this distribution, denoted by p, is a function from $\mathbf{R}$ to $[0, +\infty)$ supposed to be unknown. The problem is to estimate p. An estimator of p is a function $x \mapsto p_n(x) = p_n(x, X_1, \ldots, X_n)$ measurable with respect to the observation $\mathbf{X} = (X_1, \ldots, X_n)$. If we know a priori that p belongs to a parametric family $\{g(x, \theta) : \theta \in \Theta\}$, where $g(\cdot, \cdot)$ is a given function, and Θ is a subset of $\mathbf{R}^k$ with a fixed dimension k independent of n, then estimation of p is equivalent to estimation of the finite-dimensional parameter θ. This is a *parametric* problem of estimation. On the contrary, if such a prior information about p is not available we deal with a *nonparametric* problem. In nonparametric estimation it is usually assumed that p belongs to some "massive" class $\mathcal{P}$ of densities. For example, $\mathcal{P}$ can be the set of all the continuous probability densities on $\mathbf{R}$ or the set of all the Lipschitz continuous probability densities on $\mathbf{R}$. Classes of such type will be called nonparametric classes of functions.

2. Nonparametric regression

Assume that we have n independent pairs of random variables $(X_1, Y_1), \ldots, (X_n, Y_n)$ such that

$$Y_i = f(X_i) + \xi_i, \quad X_i \in [0, 1], \tag{1.1}$$

where the random variables ξ_i satisfy $\mathbf{E}(\xi_i) = 0$ for all i and where the function f from $[0, 1]$ to $\mathbf{R}$ (called the regression function) is unknown. The problem of nonparametric regression is to estimate f given a priori that this function belongs to a nonparametric class of functions $\mathcal{F}$. For example, $\mathcal{F}$ can be the set of all the continuous functions on $[0, 1]$ or the set of

A. B. Tsybakov, *Introduction to Nonparametric Estimation*,
DOI 10.1007/978-0-387-79052-7_1,