

Then by (12) there exists $T > 0$ such that $t > T$ implies

$$\mathbb{E} \inf_{t > T} \rho(\mathbf{y} - t\gamma' \mathbf{h}(\mathbf{x}, \alpha)) > 1 - \delta - 2\varepsilon. \quad (13)$$

Therefore (13) implies that for each $(\alpha, \gamma) \in A \times \Gamma$ there exist a neighborhood $U(\alpha, \gamma) \subset A \times \Gamma$ and $T(\alpha, \gamma) \in R_+$ such that

$$\mathbb{E} \inf_{(\alpha_1, \gamma_1) \in U(\alpha, \gamma)} \inf_{t > T(\alpha, \gamma)} \rho(\mathbf{y} - t\gamma_1' \mathbf{h}(\mathbf{x}, \alpha_1)) > 1 - \delta - 2\varepsilon = \lambda_0 + \frac{\xi}{2}. \quad (14)$$

The neighborhoods $\{U(\alpha, \gamma) : \alpha \in A, \gamma \in \Gamma\}$ are a covering of the compact set $A \times \Gamma$, and therefore there exists a finite subcovering thereof: $\{U_j = U(\alpha_j, \gamma_j)\}_{j=1}^N$. Let $T_0 = \max_j T(\alpha_j, \gamma_j)$.

We shall show that $\limsup_{n \rightarrow \infty} \|\hat{\beta}_n\| \leq T_0$ a.s. Put for brevity

$$\lambda_n(\alpha, \beta) = \frac{1}{n} \sum_{i=1}^n \rho(y_i - \beta' \mathbf{h}(x_i, \alpha)).$$

Then

$$\begin{aligned} \inf_{\|\beta\| > T_0} \inf_{\alpha \in A} \lambda_n(\alpha, \beta) &\geq \frac{1}{n} \sum_{i=1}^n \inf_{\alpha \in A, \gamma \in \Gamma} \inf_{t > T_0} \rho(y_i - t\gamma' \mathbf{h}(x_i, \alpha)) \\ &= \min_{j=1, \dots, N} \frac{1}{n} \sum_{i=1}^n \inf_{(\alpha, \gamma) \in U_j} \inf_{t > T_0} \rho(y_i - t\gamma' \mathbf{h}(x_i, \alpha)), \end{aligned}$$

and therefore (14) and the Law of Large Numbers imply

$$\lim_{n \rightarrow \infty} \inf_{\|\beta\| > T_0} \inf_{\alpha \in A} \lambda_n(\alpha, \beta) \geq \lambda_0 + \frac{\xi}{2} \text{ a.s.},$$

while

$$\lambda_n(\hat{\alpha}_n, \hat{\beta}_n) = \inf_{\beta \in B} \inf_{\alpha \in A} \lambda_n(\alpha, \beta) \leq \lambda_n(\alpha_0, \beta_0) \rightarrow \lambda_0 \text{ a.s.}$$

which shows that ultimately $\|\hat{\beta}_n\| \leq T_0$ with probability one.