Function values are almost optimal for (deterministic) L_2-approximation

Mario Ullrich

Let $F \subset L_2$ be a class of complex-valued functions on a set D, such that, for all $x \in D$, point evaluation $f \mapsto f(x)$ is a continuous linear functional. We study the L_2-approximation of functions from F and want to compare the power of function values with the power of arbitrary linear information.

To be precise, the *sampling width* $e_n(F)$ is the minimal worst-case error that can be achieved with n function values, whereas the *linear width* $a_n(F)$ is the minimal worst-case error that can be achieved with n pieces of arbitrary linear information (like derivative values or Fourier coefficients), using linear algorithms. We show, under mild assumptions on F, that

$$e_{cn}(F) \;\le\; c_p \sqrt{\log n} \left(\frac{1}{n} \sum_{k \ge n} a_k(F)^p \right)^{1/p}, \qquad n \ge 3,$$

for all $0 < p < 2$ and constants $c, c_p > 0$ that only depend on p.
In particular, the linear widths and the sampling widths have the same polynomial order of convergence: If we assume that $a_n(F) \lesssim n^{-\alpha} (\log n)^{\beta}$ for some $\alpha > 1/2$ and $\beta \in \mathbb{R}$, then we obtain

$$e_n(F) \;\lesssim\; n^{-\alpha} (\log n)^{\beta + 1/2}.$$

Despite its generality, this result is sharp enough to improve upon several existing bound in specific settings. Moreover, our proof reveals the fascinating fact that i.i.d. random sampling points, together with a suitable (weighted) least squares method, are with high probability as good as all known, sophisticated constructions.