Herrmann, Lukas; Opschoor, Joost A. A.; Schwab, Christoph
Constructive deep ReLU neural network approximation. (English) Zbl 07458301

Summary: We propose an efficient, deterministic algorithm for constructing exponentially convergent deep neural network (DNN) approximations of multivariate, analytic maps $f: [-1, 1]^K \to \mathbb{R}$. We address in particular networks with the rectified linear unit (ReLU) activation function. Similar results and proofs apply for many other popular activation functions. The algorithm is based on collocating particular networks with the rectified linear unit (ReLU) activation function. Similar results and proofs uniformly, with respect to the inputs. For analytic maps $f: [-1, 1]^K \to \mathbb{R}$, we prove exponential convergence of expression and generalization errors of the constructed ReLU DNNs. Specifically, for every target accuracy $\varepsilon \in (0, 1)$, there exists $N$ depending also on $f$ such that the error of the construction algorithm with $N$ evaluations of $\widehat{f}$ in input in the norm $L^\infty([-1, 1]^K; \mathbb{R})$ is smaller than $\varepsilon$ up to an additive data-corruption bound $\|f - \widehat{f}\|_{L^\infty([-1, 1]^K)}$ multiplied with a factor growing slowly with $\frac{1}{\varepsilon}$ and the number of non-zero DNN weights grows polylogarithmically with respect to $1/\varepsilon$. The algorithmic construction of the ReLU DNNs which will realize the approximations, is explicit and deterministic in terms of the function values of $\widehat{f}$ in tensorized Clenshaw-Curtis grids in $[-1, 1]^K$. We illustrate the proposed methodology by a constructive algorithm for (offline) computations of posterior expectations in Bayesian PDE inversion.

MSC:
41A50 Best approximation
41A10 Approximation by polynomials
65D05 Numerical interpolation
65D15 Algorithms for approximation of functions

Keywords:
deep ReLU neural networks; exponential convergence; neural network construction; generalization error

Full Text: DOI

References: