		doc_1		doc_2	decision	id
cases	authors		authors	• Keke Wu • Rui Du • Jingrun Chen • Xiang Zhou	DUPLICATES	43
	title	Understanding Loss Landscapes of Neural Network Models in Solving Partial Differential Equations	title	Understanding Loss Landscapes of Neural Network Models in Solving Partial Differential Equations		
	publication_date	2021-03-20 00:00:00	publication_date	2021-03-20 01:46:39+00:00		
	source	SupportedSources.PAPERS_WITH_CODE	source	SupportedSources.ARXIV		
	journal		journal	None		
	volume		volume			
	doi		doi			
	urls	• https://arxiv.org/pdf/2103.11069v1.pdf • https://github.com/wukekever/landscapes	urls	• http://arxiv.org/pdf/2103.11069v1 • http://arxiv.org/abs/2103.11069v1 • http://arxiv.org/pdf/2103.11069v1		
	id	id-8683175492600289561	id	id3424933303362828601		
	abstract	Solving partial differential equations (PDEs) by parametrizing its solution by neural networks (NNs) has been popular in the past a few years. However, different types of loss functions can be proposed for the same PDE. For the Poisson equation, the loss function can be based on the weak formulation of energy variation or the least squares method, which leads to the deep Ritz model and deep Galerkin model, respectively. But loss landscapes from these different models give arise to different practical performance of training the NN parameters. To investigate and understand such practical differences, we propose to compare the loss landscapes of these models, which are both high dimensional and highly non-convex. In such settings, the roughness is more important than the traditional eigenvalue analysis to describe the non-convexity. We contribute to the landscape comparisons by proposing a roughness index to scientifically and quantitatively describe the heuristic concept of "roughness" of landscape around minimizers. This index is based on random projections and the variance of (normalized) total variation for one dimensional projected functions, and it is efficient to compute. A large roughness index hints an oscillatory landscape profile as a severe challenge for the first order optimization method. We apply this index to the two models for the Poisson equation and our empirical results reveal a consistent general observation that the landscapes from the deep Galerkin method around its local minimizers are less rough than the deep Ritz method, which supports the observed gain in accuracy of the deep Galerkin method.	abstract	Solving partial differential equations (PDEs) by parametrizing its solution by neural networks (NNs) has been popular in the past a few years. However, different types of loss functions can be proposed for the same PDE. For the Poisson equation, the loss function can be based on the weak formulation of energy variation or the least squares method, which leads to the deep Ritz model and deep Galerkin model, respectively. But loss landscapes from these different models give arise to different practical performance of training the NN parameters. To investigate and understand such practical differences, we propose to compare the loss landscapes of these models, which are both high dimensional and highly non-convex. In such settings, the roughness is more important than the traditional eigenvalue analysis to describe the non-convexity. We contribute to the landscape comparisons by proposing a roughness index to scientifically and quantitatively describe the heuristic concept of "roughness" of landscape around minimizers. This index is based on random projections and the variance of (normalized) total variation for one dimensional projected functions, and it is efficient to compute. A large roughness index hints an oscillatory landscape profile as a severe challenge for the first order optimization method. We apply this index to the two models for the Poisson equation and our empirical results reveal a consistent general observation that the landscapes from the deep Galerkin method around its local minimizers are less rough than the deep Ritz method, which supports the observed gain in accuracy of the deep Galerkin method.		
	versions		versions			