Greedy Algorithm for Neural Networks for Indefinite Elliptic Problems

Ziqian Li

joint work with Qingguo Hong (Missouri S&T),
Jiwei Jia (JLU) and Young Ju Lee (TSU)

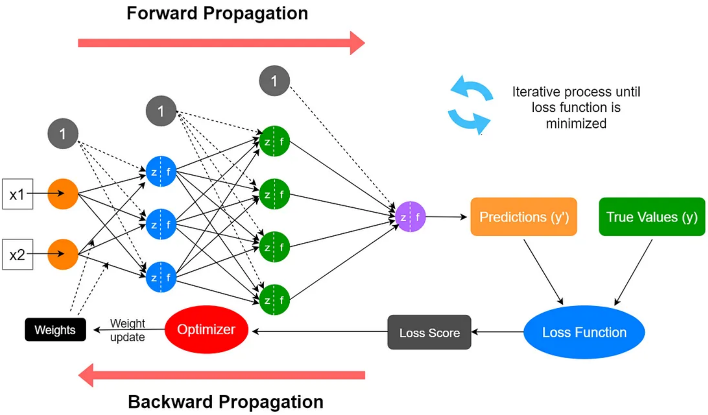
Erlangen, March 24, 2025

Friedrich-Alexander-Universität
DYNAMICS, CONTROL, MACHINE LEARNING AND NUMERICS

Contents

Optimization for neural networks

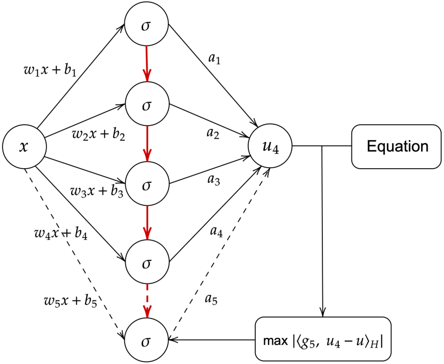
Classic way to train a neural network:



[0]image by Data science 365

Optimization for neural networks

Greedy way to train a neural network:



where

$$u_n = \sum_{i=1}^{n} a_i g_i, \quad g_i = \sigma(w_i x + b_i)$$

Preliminaries

For shallow neural networks with ReLU^k activation function $\sigma = \max\{0, x\}^k$, the dictionary $\mathbb{D}$ of d-dimension is taken as

$$\mathbb{D} = \mathbb{P}_k^d = \left\{ \sigma_k(\boldsymbol{\omega} \cdot \boldsymbol{x} + b) : \boldsymbol{\omega} \in S^{d-1}, \, b \in [c_1, c_2] \right\},$$

where $S^{d-1} = \{\boldsymbol{\omega} \in \mathbb{R}^d : |\boldsymbol{\omega}| = 1\}$ is the unit sphere.

Preliminaries

Consider the elliptic equation in $\mathbb{R}^d$:

$$\begin{cases} -\Delta u(x) + cu(x) = f(x), \quad \text{in } \Omega, \\ \dfrac{\partial u(x)}{\partial n} = 0, \quad \text{on } \partial\Omega, \end{cases}$$

the energy inner product of this equation is defined as

$$\langle u, v\rangle_H := a(u, v) = (\nabla u, \nabla v) + (cu, v), \quad \forall u, v \in \mathbb{R}^d.$$

When $c < 0$, the bilinear form $a(u, v)$ may not be coercive, this equation is then named indefinite elliptic equation (e.g. Helmholtz equation).

Step 1: Find n-th basis function g_n

In the n-th iternation, $n = 1, 2, 3, \ldots$, find g_n through

$$g_n = \arg\max_{g \in \mathbb{D}} |\langle g, u_{n-1} - u \rangle_H|, \tag{1}$$

where

$$\begin{aligned} \langle g, u_{n-1} - u \rangle_H &= (\nabla g, \nabla(u_{n-1} - u)) + (g, cu_{n-1} - cu) \\ &= (\nabla g, \nabla u_{n-1}) + (g, cu_{n-1}) - (\nabla g, \nabla u) - (g, cu) \\ &= (\nabla g, \nabla u_{n-1}) + (g, cu_{n-1}) - (f, g) \\ &= \int_\Omega (\nabla g \cdot \nabla u_{n-1} + cgu_{n-1} - fg) \mathrm{d}\boldsymbol{x}. \end{aligned}$$

Because u_{n-1}, c, f are known, we only need to choose $g \in \mathbb{D}$ to satisfy (1).

Step 2: Orthogonal projection

Apply orthogonal projection on g_i, $i = 1, 2, \ldots, n$ to update all the coefficients a_i, $i = 1, 2, \ldots, n$. Due to the weak form of the equation

$$(\nabla u, \nabla v) + (cu, v) = (f, v), \quad \forall u, v \in \mathbb{R}^d,$$

we obtain the following linear system with a_i as the solution:

$$\int_\Omega \nabla \left(\sum_{i=1}^{n} a_i g_i \right) \cdot \nabla g_j \mathrm{d}\boldsymbol{x} + \int_\Omega c \left(\sum_{i=1}^{n} a_i g_i \right) g_j \mathrm{d}\boldsymbol{x} = \int_\Omega f g_j \mathrm{d}\boldsymbol{x},$$

$j = 1, \cdots, n$. Then a_i, $i = 1, 2, \ldots, n$ are obtained.

Helmholtz equation

We consider the 2D Helmholtz equation given as follows:

$$\begin{cases} -\Delta u - k^2 u = f, \quad x \in (0,1)^2, \\ \dfrac{\partial u}{\partial n} = 0, \quad x \in \partial(0,1)^2, \end{cases}$$

where $k > 1$ is the wavenumber. The source term $f = k^2 \cos(kx)\cos(ky) - k^2$ and the analytical solution is $u(x,y) = \cos(kx)\cos(ky) + 1$.

Helmholtz equation - $k = 2\pi$

Performance of greedy algorithm with ReLU^3 activation function:

n	dof	$\|u - u_h\|_{L^2}$	order	$\|u - u_h\|_{H^1}$	order
16	48	2.801e-02	-	1.382e-01	-
32	96	2.229e-03	3.65	2.600e-02	2.41
64	192	2.952e-04	2.92	6.435e-03	2.01
128	384	3.887e-05	2.93	1.649e-03	1.96
256	768	1.176e-05	1.73	4.430e-04	1.90

Performance of FEM with P_2 quadratic elements:

h	dof	$\|u - u_h\|_{L^2}$	order	$\|u - u_h\|_{H^1}$	order
$1/16$	1023	7.957e-02	-	1.267e-01	-
$1/32$	4095	4.084e-02	0.96	6.469e-02	0.97
$1/64$	16383	2.069e-02	0.98	3.273e-02	0.98
$1/128$	65535	1.041e-02	0.99	1.647e-02	0.99
$1/256$	262143	5.224e-03	0.99	8.260e-03	1.00

Helmholtz equation - $k = 10\pi$

Performance of greedy algorithm with ReLU^3 activation function:

n	dof	$\|u - u_h\|_{L^2}$	order	$\|u - u_h\|_{H^1}$	order
16	48	2.150e+00	-	3.399e+00	-
32	96	7.305e-01	1.56	1.193e+00	1.51
64	192	6.410e-02	3.51	1.719e-01	2.80
128	384	5.674e-03	3.50	4.195e-02	2.03
256	768	8.196e-04	2.79	1.134e-02	1.89

Performance of FEM with P_2 quadratic elements:

h	dof	$\|u - u_h\|_{L^2}$	order	$\|u - u_h\|_{H^1}$	order
1/16	1023	1.419e+00	-	2.261e+00	-
1/32	4095	1.906e-02	-	9.205e-02	-
1/64	16383	2.073e-02	-	4.017e-02	-
1/128	65535	1.041e-02	0.99	1.749e-02	1.20
1/256	262143	5.224e-03	0.99	8.391e-03	1.06

Thank you!

[0]Jinchao Xu, et al. Greedy training algorithms for neural networks and applications to PDEs[J]. Journal of Computational Physics, 2023, 484: 112084.

[0]Qingguo Hong, Jiwei Jia, Young Ju Lee, Ziqian Li. Greedy Algorithm for Neural Networks for Indefinite Elliptic Problems[J]. arXiv preprint arXiv:2410.19122, 2024.