## Matrix Analysis and Fast Solvers for Neural Network Computations

## Jianlin Xia

## Abstract

Neural networks provide a powerful tool for machine learning and other data science techniques. They can also serve as new ways for mathematical and numerical tasks such as function approximations and PDE solutions. Although there have been significant developments in neural network methods, the analysis of relevant matrices and the design of relevant fast and stable matrix computation techniques are typically overlooked.

In fact, neural network methods provide highly interesting new opportunities to perform matrix analysis and design matrix algorithms that can benefit modern data analysis and machine learning. Examples of scenarios where large and challenging matrices arise include the following.

- In neural network least-squares approximations of functions, large mass matrices and Hessian matrices may be constructed from activation functions such as ReLU functions as basis functions.
- Sparse structured matrices have been often used in the design of effective neural networks and efficient training algorithms (and a simple example is the use of sparse Toeplitz matrices as weight matrices in some neural networks).
- In optimization and training algorithms such as ADAM and BFGS, the underlying matrices are often closely related to certain preconditioners.

In this talk, we aim to bridge the gap between some neural network methods and fast and reliable matrix computations. We present rigorous analysis for some of these matrices and show two aspects.

- Why some of these matrices pose significant challenges (say, in the conditioning, spectrum distribution, and frequency modes) for matrix computations.
- Why it is feasible to design new fast and reliable solvers for these problems based on certain underlying structures.

In particular, consider the approximation of a function  $u: \Omega(\subset \mathbb{R}^d) \to \mathbb{R}$  by

$$v = \sum_{i=1}^{n} c_i \sigma(\mathbf{w}_i^T \mathbf{x} + b_i), \quad \mathbf{x} \in \Omega,$$

where  $\mathbf{w}_i$ 's are weight vectors,  $b_i$ 's are biases,  $c_i$ 's are scalar coefficients, and  $\sigma(t) = \max\{t, 0\}$  is the ReLU function. Let  $W = (\mathbf{w}_1, \ldots, \mathbf{w}_n)$ ,  $\mathbf{b} = (b_1, \ldots, b_n)^T$ , and  $\mathbf{c} = (c_1, \ldots, c_n)$ . The least-squares approximation of u by v solves the following optimization problem:

$$\min_{W,\mathbf{b},\mathbf{c}} \mathcal{J} \quad \text{with} \quad \mathcal{J} := \langle v - u, v - u \rangle = \frac{1}{2} \int_{\Omega} (v - u)^2 d\mathbf{x}.$$

By viewing **c** as a set of linear parameters and  $\{W, \mathbf{b}\}$  as nonlinear parameters, we can look at the gradient of  $\mathcal{J}$  with respect to one of the two sets of parameters with the other set fixed [1]. Setting

the gradients to be 0 leads to a linear system for the linear parameters and a nonlinear system for the nonlinear parameters. The former system has a mass matrix  $\tilde{A}$  as the coefficient matrix. The solution of the latter system with Newton or Gauss-Newton methods lead to linear systems involving a Hessian or Gauss-Newton matrix  $\tilde{H}$ . In a highly simplified setting,  $\tilde{A}$  and  $\tilde{H}$  may be related to the following matrix forms, respectively:

$$A = (A_{ij})_{n \times n}, \quad A_{ij} = \left\langle \sigma(\mathbf{w}_i^T \mathbf{x} + b_i), \sigma(\mathbf{w}_j^T \mathbf{x} + b_j \right\rangle, \\ H = (H_{ij})_{n \times n}, \quad H_{ij} = \left\langle h(\mathbf{w}_i^T \mathbf{x} + b_i), h(\mathbf{w}_j^T \mathbf{x} + b_j \right\rangle,$$

where  $h(t) = \sigma'(t) = \begin{cases} 1, & t > 0, \\ 0, & t < 0. \end{cases}$ 

Some interesting matrix analysis may be performed for A and H. For example, we can show the following aspects.

- A and H are positive definite (with modest assumptions) but are highly ill conditioned due to the fast decay of the eigenvalues. For instance, even in the 1D case with uniform breakpoints that define the ReLU basis functions, A has 2-norm condition number proportional to  $\frac{1}{n^4}$ .
- The behaviors of the low and high frequency modes further make them challenging for iterative solvers.
- Some preconditioning strategies may be designed based on basis function modifications, but the effectiveness is limited.
- On the other hand, the matrices have nice inherent structures. In particular, relevant offdiagonal blocks of the matrices are low rank (the 1D case) or numerically low rank (with appropriate conditions). These structures make it feasible to design fast and stable direct solvers for the relevant linear systems.

These problems thus give a nice opportunity to apply structured matrix methods. This also shows how advanced matrix analysis may benefit modern neural network methods. The talk includes joint with Z. Cai, T. Ding, M. Liu, and X. Liu.

## References

[1] Z. CAI, T. DING, M. LIU, X. LIU, AND J. XIA, A structure-guided Gauss-Newton method for shallow ReLU neural network, arXiv:2404.05064, submitted to SIAM J. Sci. Comput., (2024).