Report on "Exact full-RSB SAT/UNSAT transition in infinitely wide two-layer neural networks"

1 Summary

This paper conducts a detailed analysis of the phase diagram of the tree committee machine, a model of two-layer neural networks which includes the perceptron as a special case. The paper considers the case where the weights of the first layer are on the sphere and the weights of the second layer are fixed, and derives via the replica method the (free) entropy of the set of solutions ('interpolators') when there are $P = \alpha N$ random Gaussian patterns in dimension N in the large system limit under a general full RSB ansatz, and provides an accurate estimate of the SAT/UNSAT transition for various popular activation functions.

The paper then considers the special case of the negative perceptron. The main contribution of the paper is to unveil the existence of a Gardner phase bordering the satisfiability threshold for a large unbounded range of negative values of the margin κ. In this regime the two-replica overlap distribution exhibits an atom at its left-end followed by a gap followed by a continuous part, a possibility which was not considered in previous literature where only the continuous part was established/assumed to exist.

The existence of an 'overlap gap' in this Gardner phase has an important algorithmic consequence. The incremental approximate message passing algorithm (IAMP) of ref. [26] is known to succeed at returning a solution when there is no gap in the overlap distribution. The authors show that the 'no overlap gap' (FRSB) phase is a bounded region bordering the SAT/UNSAT transition for relatively moderate values of the negative margin κ, followed by the Gardner phase where this algorithm must fails. In particular the IAMP algorithm does not succeed up to satisfiability for all $\kappa < 0$.

The authors also conduct a numerical experiment with gradient descent with two loss functions (the quadratic hinge loss and the cross entropy) and show that this algorithm fails much below the SAT/UNSAT threshold.

2 Comments

This work unveils a new phenomenon in the negative perceptron with important consequences regarding the computational tractability of finding solutions in toy models of neural networks. I have two comments:

1. I find the shape of the boundary between the fRSB phase and the Gardner phase (the dotted blue line $\alpha^{1+\text{fRSB}}$ in Figure 5) particularly intriguing for the following reason: let's say that a point in the (κ, α) plane is *solvable* if there exists an efficient algorithm for returning a solution to the perceptron problem at those parameters. Then any point to the bottom-left of a solvable point is also solvable since one can add constraints and/or make the margin tighter, solve the harder problem then return the solution found. (This is a simple argument justifying the fact that the SAT/UNSAT transition line is decreasing.)

Now, if the boundary $\kappa \mapsto \alpha^{1+\text{fRSB}}(\kappa)$ is increasing as depicted in Figure 5, then there must exist solvable points in the Gardner and 1RSB phases of the model. It is not clear that this is impossible but it would be interesting since this would imply an absence of (worst case) overlap gap in a 1RSB phase. It would be instructive if the authors could provide more details on the estimation of this line.

2. My second comment is about the gradient descent experiments. The success probability should transition sharply from 1 to 0 at the some threshold α_{GD} for large system size. The transition does appear to be sharp in Figure 7 but not so much in Figure 6. One run appears to find solutions beyond the critical threshold α_c, which means finite size effects are still present. Perhaps the authors could run larger instances and average more runs so all curves are nearly vertical and collapse on top of each other. Moreover, ref. [36] has a rigorous analysis of a linear programming algorithm and experimental results on some version of gradient descent. Perhaps depicting their transition line in the same phase diagram would also be instructive for the sake of comparison.

I support acceptance of the paper conditional on a revision addressing the previous points.