19 March 2025 ·

Qeios

Peer Review

Review of: "Scalable Thermodynamic Second-order Optimization"

Filippo M. Miatto[1]

1. Xanadu Quantum Technologies (Canada), Toronto, Canada

The paper describes an application of thermodynamic computing. Specifically, K-FAC optimization approximates the local curvature of the cost function to second order, and therefore it needs matrix inversion / linear system solving as part of the optimization procedure. The claim is that thermodynamic computing is a cheaper way to perform such operations and therefore it can speed up this kind of optimization.

The paper is generally correct, though I found it a bit sloppily written. Also, the main contribution (the ability of thermodynamic computing to solve linear systems) has already been presented in other works by the same team, which makes this mostly a presentation of a specific application. There are so many algorithms that require linear system solving or matrix inversion that one might question the need to address each of them individually.

The first part of the paper summarizes the K-FAC algorithm from the original K-FAC paper. The original paper specifies that the cost function should be the negative log probability of a cross entropy or a least squares objective function, but this is missing in this work. Additionally, the definition of the cost function (\cal{L}) in eq. 1 is a function of the parameters \theta, but then from eq. 2 onwards, it becomes a function of y and f(x). A better way to write the derivative would be to write \partial_{\theta}(\cal L) and then evaluate this where needed. Instead, it's confusing to leave y and f(x) in the expression because \cal L is not a function of x and y (having summed over them). Finally (and perhaps this is more of a personal preference), a lot of space is dedicated to index gymnastics with tensor products and vectorizations, which could be simplified by specifying the tensor indices directly.

A question that I had while reading the paper was about the complexity of sampling the output of the thermodynamic device, which I expected would be included in the calculations. Clearly, the thermodynamic device reaches a state such that some of its degrees of freedom are distributed like the

doi.org/10.32388/KTPXU1

quantity that is sought (e.g., the solution of the linear system) in linear time, but I am also interested in understanding the complexity of extracting such values. For a fair comparison with non-thermodynamic methods, I would like to see the relationship between the sampling/time complexity and the error in estimating the solution added to the global complexity calculation.

Below is a list of typos:

Page 4, line 2: missing “been”

Before eq. 18, the expression for ds should have \ell

In the line after eq. 21, it should read “fewer” not “less” sums

In the third line of sec 4.2, there’s an extra “to”

In the third line after eq 22, there’s an extra s in “corresponds”

The last line of page 10 is missing a “the”

Above figure 2, it should read “reasonably wide layers” with a y

In the middle of page 13, there’s a missing “information” after Fisher

In the second line of page 14, there’s a missing verb between “we” and “strategies”

In the fourth line of sec 4.5.1, it should read “...have a catastrophic...”

Declarations

Potential competing interests: No potential competing interests to declare.

doi.org/10.32388/KTPXU1