CONVEXITY CERTIFICATES FROM HESSIANS

Julien Klaus[1], Niklas Merk[1], Konstantin Wiedom[1], Sören Laue[2,3], Joachim Giesen[1]

[1]Friedrich Schiller University Jena, [2]Technical University of Kaiserslautern, [3]Data Assessment Solutions GmbH

Abstract

The Hessian of a differentiable convex function is positive semidefinite. Therefore, checking the Hessian of a given function is a natural approach to certify convexity. However, implementing this approach is not straightforward since it requires a representation of the Hessian that allows its analysis. Here, we implement this approach for a class of functions that is rich enough to support classical machine learning. For this class of functions it was recently shown how to compute computational graphs of their Hessians. We show how to check these graphs for positive semidefiniteness. We compare our implementation of the Hessian approach with the well-established disciplined convex programming (DCP) approach and prove that the Hessian approach is at least as powerful as the DCP approach for differentiable functions. Furthermore, we show for a state-of-the-art implementation of the DCP approach that, for differentiable functions, the Hessian approach is actually more powerful. That is, it can certify the convexity of a larger class of differentiable functions.

Acknowledgments

This work was supported by the Carl Zeiss Foundation within the project "Interactive Inference" and by the Ministry for Economics, Sciences and Digital Society of Thuringia.

Example

Is the function $f : \mathbb{R}^n \to \mathbb{R}$,

$$f(w) = \sum_{i=1}^{n} \log(\exp(w_i) + 1)$$

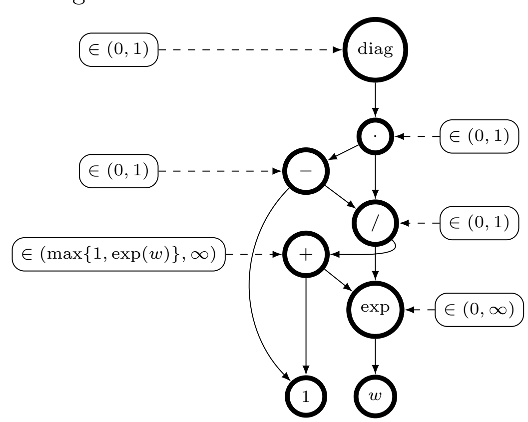
convex?

Solution

A function is convex, if its Hessian is positive semidefinite. We decide definiteness by traversing the Hessian and checking it for positivity, $\nabla_w^2 f(w)$

$$= \operatorname{diag}\left(\frac{\exp(w)}{\mathbb{1} + \exp(w)} \cdot \left(1 - \frac{\exp(w)}{\mathbb{1} + \exp(w)}\right)\right)$$

by traversing the directed acyclic graph representing this Hessian.



General Template

In general, we can not say something about the difference of two positive semidefinite matrices.

For example the **negative harmonic mean** $f(x) = -n / \sum_{i=1}^{n} x_i^{-1}$,

$$\nabla_x^2 f(x) = 2 \cdot f(x)^2 \Big(\operatorname{diag}\big(\text{vector}(1) \oslash (x \odot x \odot x)\big) + f(x) \cdot \ (\text{vector}(1) \oslash (x \odot x))(\text{vector}(1) \oslash (x \odot x))^\top\Big),$$

or ***p*-norms** $f(x) = \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p}$,

$$\nabla_x^2 f(x) = (p-1) \cdot \text{sum}((f(x))^{1/p-1} \cdot \operatorname{diag}(f(x) \oslash (x \odot x)) - (p-1) \cdot \text{sum}(f(x))^{1/p-2} \cdot (f(x) \oslash x)(f(x) \oslash x)^\top.$$

Interestingly both Hessians conform to the following positive semidefinite template that can be used for deciding convexity.