

Fitting absolutely monotonic functions

Juan Casanova

September 8, 2025

This is the abstract.

Contents

1	Introduction	1
1.1	Absolutely monotonic functions	2
1.2	Fitting exponential sums	4
1.3	Fitting monotonic and convex functions	5
1.4	Why absolutely monotonic functions	7
1.5	Summary	8
2	Method	9
2.1	Evaluation	9
3	Results	9
4	Conclusions	9

1 Introduction

This paper has two purposes. First, to present a method for fitting absolutely monotonic functions to data. Second, to discuss in this context three areas of mathematics research that, surprisingly, are rarely discussed together. We summarize the relevant similarities and connections of each of them with the aforementioned method, and consequently between them, as well as how this relates to potential applications of this work and existing literature.

1.1 Absolutely monotonic functions

A smooth function $f : [a, b) \rightarrow \mathbb{R}$ (where b can be $+\infty$) is *absolutely monotonic* if all of its derivatives are non-negative at every point in the interval. A simple example of an absolutely monotonic function is the exponential e^x , and in fact in a very significant sense, it can be shown that all absolutely monotonic functions are essentially sums of exponentials, a key idea supporting this work. More accurately, in his classic work [Bernstein, 1929], among other things, Bernstein characterized all absolutely monotonic functions as weighted averages of exponential functions with positive exponents.

Theorem 1.1 ([Bernstein, 1929]). *Let $f : [a, b) \rightarrow \mathbb{R}$ be an absolutely monotonic function. Then, there is a non-negative finite Borel measure μ such that*

$$f(x) = \int_0^\infty e^{tx} d\mu(t) \tag{1}$$

for every $x \in [a, b)$.

There is a direct connection between absolutely monotonic functions and *completely monotonic functions*, which are smooth functions with alternating sign derivatives. Specifically, a function $f(x)$ is absolutely monotonic in $[a, b)$ if and only if the function $f(-x)$ is completely monotonic in $(-b, -a]$. For example, e^{-x} is a completely monotonic function, and similarly, Bernstein's theorem, when applied to completely monotonic functions, characterizes them as weighted averages of exponential functions with negative exponents.

In informal terms, one can describe the function e^x to be absolutely monotonic *towards the right* and completely monotonic *towards the left*, and the absolutely monotonic or completely monotonic views of functions are more related with what side of the function we are focusing on, and the corresponding properties that they have on them.

For various significant reasons, there has been substantially more work on completely monotonic functions than on absolutely monotonic functions. While it is tempting to argue that they are simply two sides of the same coin, their relation is limited by the nature of the result we are focusing on. More precisely, a lot of the work on completely monotonic functions considers them in the half-line $[0, \infty)$, which when mirrored results in the interval $(-\infty, 0]$ for absolutely monotonic functions. However, the interesting monotonic behaviour of absolutely monotonic functions is most relevant in intervals on the right side of the real line. Similarly, completely monotonic

functions involve small absolute values and controlled decay, whereas absolutely monotonic functions involve very large absolute values and unfettered growth. This means that the types of interesting behaviours and challenges in each of these two sides of the coin can be quite different in practice.

Absolutely monotonic and completely monotonic functions are particular cases of exponential sums. When both positive, negative, and complex exponents are combined, more general types of functions appear, including periodic functions and variations on these.

Absolutely monotonic functions form a *convex cone*. In particular, weighted sums of absolutely monotonic functions are absolutely monotonic.

Lemma 1.1. *If f, g are absolutely monotonic functions and $\alpha, \beta \in [0, +\infty)$, then $\alpha f + \beta g$ is an absolutely monotonic function.*

The proof is trivial by looking at the definition of absolute monotonicity in terms of non-negative derivatives and linearity of derivation.

We do not concern ourselves too much with further theoretical properties of absolutely monotonic or completely monotonic functions in this work, but for the interested reader, [Koumandos, 2014] offers a modern summary of theoretical results and additional related families of functions, such as logarithmically completely monotonic functions, Bernstein functions, and Stieltjes functions.

[Rajba, 2011] is a particularly interesting piece of work partially bridging the gap between methods focused on convexity and absolute monotonicity. It presents an integral representation of *n-convex* functions, which are functions where a certain number of derivatives are positive. In other words, an absolutely monotonic function is a function that is *n-convex* for every *n*. The mathematics in this work and the representation it offers for *n-convex* functions are strongly reminiscent of the Bernstein theorem. In particular, there is a direct relationship between the terms in this work and the Taylor series of the exponential function, indicating that this integral representation is using “limited exponential functions” consisting of partial Taylor series of exponential functions, up to the derivative that is required. It is surprising, then, that the work does not mention the Bernstein theorem or the words “absolutely monotonic” or “completely monotonic” whatsoever, or cite any work in this subject.

While we do not use them in this paper, this work can be relevant in extending some of the notions here to less constrained cases like *n-convex*

functions, or to better understand the methods and their limitations.

1.2 Fitting exponential sums

There is a large body of work on fitting exponential sums to data. A lot of this work is related to signal analysis and therefore is concerned with general exponential sums, including positive, negative, and complex terms; though practical work is often limited to discrete sums.

One of the most significant families of methods for fitting exponential sums are the so-called Prony methods. Prony methods are spectral methods that use the relation between exponential sums, uniform sampling, and difference equations that enable a certain type of frequency analysis using linear algebra. [Keller and Plonka, 2021] offers a modern survey of Prony methods, focusing on common principles, such that they are linear methods on the space of operators on functions, and that they rely on sampling at specific distances to capture periodic behaviours in the signals. A lot of work on Prony methods is related to the optimization of the linear algebra aspects of the algorithms.

Prony methods are more general in its function space, present more constraints to the data, and focus more on frequency analysis of a signal than on good numerical fitting of a real function. Their goal is the analysis of signals, often assumed to be in the $[0, +\infty)$ half-line, rather than the fitting of observed real data on the absolutely monotonic side of the function. Therefore, they are not particularly well suited for fitting absolutely monotonic functions, even if they could theoretically be applied. It could be interesting to apply some of these methods to examples on absolutely monotonic functions in bounded intervals by mirroring them and using Prony methods, to compare with the results in this paper, but we have decided not to pursue this avenue further.

Other methods have been explored for fitting exponential sums over time. [Kammler, 1976] presents a theoretically focused and exploratory analysis of using Chebyshev polynomials to approximate completely monotonic functions. [Smith et al., 1976] discusses several potential methods, also with a focus on completely monotonic functions, with limited success, and with significant computational constraints that are not applicable in the current time. [Evans et al., 1980] uses a least squares algorithm to fit completely monotonic functions, but also suffers from being computationally obsolete. It does note, however, that with a least squares approach, numerical aspects

become critical, such as initialization of values or behaviour of the parameter space. We will discuss these aspects later in this work. In a slightly more modern survey, [Holmström and Petersson, 2002] discusses a wide array of methods for fitting exponential sums, including Prony methods, least squares approaches, and many others. None of these pay significant attention to absolutely monotonic functions.

There is definitely room for further exploring whether some of these methods may be applied successfully with a focus on absolutely monotonic functions, but work would be required to adapt them. We believe that between the focus on complex exponent signals, completely monotonic functions, the differences in error measurements between fitting a slow decay function and a fast growth function, and the outdated computational considerations of some of the older pieces of work, making any of these methods work successfully on absolutely monotonic functions would not be straightforward. Moreover, the method presented in §2 uses a least squares approach with a specific choice of representation of exponentials that improves numerical behaviour of the optimization algorithm, resulting in a very simple and relatively fast algorithm that gives significantly satisfying results. This suggests that there is no need to painfully adapt methods designed for other problems when a relatively simple one exists for this problem. It would, however, be interesting to verify this with an actual comparison.

1.3 Fitting monotonic and convex functions

The methods discussed in §1.2 are parametric or semi-parametric, using a predetermined family of functions and fitting using parameterizations. This ensures certain properties of the resulting fit by construction (complete monotonicity, periodic behaviour, etc). Another family of work focuses instead on fitting the data accurately while imposing the desired properties in less constructive ways. Specifically, we are talking about the wide range of work on fitting monotonic or convex functions in general.

Absolute monotonicity is a particular case of convexity, which is a particular case of monotonicity. Absolute monotonicity is a stronger constraint. Two important questions deserve to be explored. First, can methods to fit monotone or convex functions be extended to absolute monotonicity constraints? Second, since absolutely monotonic functions are convex, can methods constrained to absolutely monotonic functions be used when the target constraint is weaker (convexity)?

All or most of the work in this area expresses the fit as an optimiza-

tion problem (e.g. linear programming), that minimizes the error, and uses additional constraints on the optimization problem and/or steps in the algorithm to ensure the monotonicity/convexity of the result. They are all heavily focused on local properties of the functions and how to encode the constraints in this way, e.g. by looking at properties of derivatives/gradients/jacobians/hessian matrices.

We found two major families of research. First, work on imposing monotonicity constraints to splines. Second, fitting convex functions in a multivariate context.

A *spline* is a smooth piecewise polynomial. Spline interpolation is widely used because it is simple, very efficient, and produces smooth satisfying results for interpolation. However, while splines are smooth by design, imposing monotonicity in them is not easy. Work like [He and Shi, 1998, Zhang, 2004, Nagahara and Martin, 2013, Lu, 2015] considers a candidate spline interpolation, and defines a linear programming problem that minimizes its error with respect to the data. It then introduces the monotonicity constraint as additional constraints on the linear program. The results are satisfying but limited, because imposing monotonicity on splines can severely affect their ability to adequately approximate the data, or the constraints may not be strict and may result in functions that are not monotonic at every point. Moreover, extending this to absolutely monotonic functions seems extremely hard, since new ways of representing the constraints on all derivatives of absolutely monotonic functions in a linear program would need to be devised, without significantly affecting the computational and approximation properties of the resulting algorithm.

Work like [Hannah and Dunson, 2013] considers the more general problem of fitting convex multivariate functions to data. Particularly due to the multivariate nature, this problem is significantly more complicated. Various techniques are used to balance the convexity constraints with the goal of a good fit to the data. For example, [Lachand-Robert and Oudet, 2005] uses a projection step as part of a gradient descent algorithm to ensure the solutions are convex, while [Aguilera and Morin, 2008] focuses on finding local characterizations of convexity and discretization of the hessian matrix of the target function, to ensure the convexity of the solution. These solutions are significantly computationally expensive, and struggle with balancing difficult constraints in a large solution space, due to their multivariate nature. Moreover, while convex but not absolutely monotonic functions are more rare, the application space is still more limited than our target problem.

It might be worth comparing the performance, in computational cost and

goodness of fit, of these solutions to our proposed method for the problems in which they overlap (monotonic and convex functions). We expect these will be more cost for worse fit.

1.4 Why absolutely monotonic functions

Monotonic and convex functions are much more general than absolutely monotonic functions. Similarly, exponential sums appear in signal and frequency analysis, and completely monotonic functions can be relevant in economic modeling. These applications, and the lack of a comparatively as significant volume of applications for absolutely monotonic functions, explain the differences in the volume of work devoted to them. Why, then, do we want to fit absolutely monotonic functions? While this paper is focused on the method, we arrived at this problem from an applied setting. We wanted to use data to fit reward functions (for players in a video game). Absolute monotonicity is an extremely attractive constraint for a reward function, as it ensures not only that better results produce better rewards, but also that the returns increase, as it is often informally said, *exponentially*, or more accurately, absolutely monotonically. This means that the same comparative improvement in performance will always result in a larger improvement (and a faster increase in this improvement, and a faster increase on the speed of this improvement, etc.) on the reward the higher the performance already is. In other words, small differences become much more noticeable near the top, which is very attractive for reward functions, for example in a competitive setting. Another way to look at this is that an absolutely monotonic function is a way to re-scale a performance value into a reward space that is “exponential”, greatly accentuating differences at higher performance value while giving less relevance to differences at lower performance values.

Naturally, convex functions are often enough to represent this type of behaviour. While we have not thoroughly explored the economics literature in search for uses of convex functions, we conjecture that there might be some applications in economics in which convex functions are expected, that would benefit from using absolutely monotonic functions instead, by giving them stronger reward behaviours. Moreover, the method discussed in §2 is simple, fast, and successful enough that it might be used in place of other methods focused on convexity.

As a whole, we believe that any real-world modeling of a function that benefits from consistently accentuating differences at higher values, but fit

this function using existing data, such as reward functions, could benefit from imposing an absolute monotonicity constraint on the result and using the fitting method described in this paper.

1.5 Summary

In the field of mathematical analysis there is a significant theoretical understanding of absolutely monotonic functions, completely monotonic functions, exponential sums in general, and other related families of functions. At the same time, several methods of fitting some families of exponential sums to data exist, coming primarily from the signal analysis world. This applied work often does not use the existing theoretical results to their full extent, and sometimes fails to recognize the connection between the different families of exponential sums and their properties. At the same time, another significant research field studies algorithms to fit monotonic and convex functions, without acknowledging their relation to absolutely monotonic functions, and therefore exponential sums, and therefore not attempting (semi-)parameteric approaches to this problem, that could benefit from easier to fulfill constraints, and better computational performance and fitness.

In the remainder of this paper we discuss our proposed, simple method for fitting absolutely monotonic functions that sits in the middle of all of this, and that could help as an initial step towards bridging these gaps. Regardless of this new method, we believe it would be in the best interest of researchers in each of these fields to explore the results in each other's fields and consider how they might improve their approaches or reconsider some assumptions or choices that they make.

2 Method

2.1 Evaluation

3 Results

4 Conclusions

References

- [Aguilera and Morin, 2008] Aguilera, N. E. and Morin, P. (2008). Approximating optimization problems over convex functions. *Numerische Mathematik*, 111:1–34.
- [Bernstein, 1929] Bernstein, S. (1929). Sur les fonctions absolument monotones. *Acta Mathematica*, 52(1):1–66.
- [Evans et al., 1980] Evans, J. W., Gragg, W. B., and LeVeque, R. J. (1980). On least squares exponential sum approximation with positive coefficients. *Mathematics of Computation*, 34(149):203–211.
- [Hannah and Dunson, 2013] Hannah, L. A. and Dunson, D. B. (2013). Multivariate convex regression with adaptive partitioning. *J. Mach. Learn. Res.*, 14(1):3261–3294.
- [He and Shi, 1998] He, X. and Shi, P. (1998). Monotone b-spline smoothing. *Journal of the American Statistical Association*, 93(442):643–650.
- [Holmström and Petersson, 2002] Holmström, K. and Petersson, J. (2002). A review of the parameter estimation problem of fitting positive exponential sums to empirical data. *Appl. Math. Comput.*, 126(1):31–61.
- [Kammler, 1976] Kammler, D. W. (1976). Chebyshev approximation of completely monotonic functions by sums of exponentials. *SIAM Journal on Numerical Analysis*, 13(5):761–774.
- [Keller and Plonka, 2021] Keller, I. and Plonka, G. (2021). Modifications of prony’s method for the recovery and sparse approximation with generalized exponential sums. In Fasshauer, G. E., Neamtu, M., and Schumaker, L. L., editors, *Approximation Theory XVI*, pages 123–152, Cham. Springer International Publishing.

- [Koumandos, 2014] Koumandos, S. (2014). *On Completely Monotonic and Related Functions*, pages 285–321. Springer New York, New York, NY.
- [Lachand-Robert and Oudet, 2005] Lachand-Robert, T. and Oudet, E. (2005). Minimizing within convex bodies using a convex hull method. *SIAM Journal on Optimization*, 16(2):368–379.
- [Lu, 2015] Lu, M. (2015). Spline estimation of generalised monotonic regression. *Journal of Nonparametric Statistics*, 27(1):19–39.
- [Nagahara and Martin, 2013] Nagahara, M. and Martin, C. F. (2013). Monotone smoothing splines using general linear systems. *Asian Journal of Control*, 15(2):461–468.
- [Rajba, 2011] Rajba, T. (2011). New integral representations of nth order convex functions. *Journal of Mathematical Analysis and Applications*, 379(2):736–747.
- [Smith et al., 1976] Smith, M. R., Cohn-Sfetcu, S., and Buckmaster, H. A. (1976). Decomposition of multicomponent exponential decays by spectral analytic techniques. *Technometrics*, 18(4):467–482.
- [Zhang, 2004] Zhang, J.-T. (2004). A simple and efficient monotone smoother using smoothing splines. *Journal of Nonparametric Statistics*, 16(5):779–796.