Journal of Uncertain Systems
Vol.8, No.x, pp.xx-xx, 2014
Online at: www.jus.org.uk

$\sqrt{x^2+\mu}$ is the Most Computationally Efficient Smooth Approximation to $|x|$: a Proof

Carlos Ramirez[1], Reinaldo Sanchez[1],
Vladik Kreinovich[1,2], and Miguel Argaez[1,3]
[1] *Computational Sciences Program*
[2] *Department of Computer Science*
[3] *Department of Mathematical Sciences*
University of Texas at El Paso
El Paso, TX 79968, USA
carlosrv19@gmail.com, vladik@utep.edu, margaez@utep.edu

Received 2 April 2013; Revised 20 June 2013

Abstract

In many practical situations, we need to minimize an expression of the type $\sum |c_i|$. The problem is that most efficient optimization techniques use the derivative of the objective function, but the function $|x|$ is not differentiable at 0. To make optimization efficient, it is therefore reasonable to approximate $|x|$ by a smooth function. We show that in some reasonable sense, the most computationally efficient smooth approximation to $|x|$ is the function $\sqrt{x^2+\mu}$, a function which has indeed been successfully used in such optimization.
©2014 World Academic Press, UK. All rights reserved.
Keywords: ℓ^1-norm, optimization, smooth approximation, efficient algorithms

1 Need to Approximate $|x|$ by Smooth Functions

Finding parameters of a model: general problem. In many practical situations, we need to determine the parameters of the model from the experimental data. In more precise terms:

- we know that the quantities $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_m)$ are related by a dependence $y = f(x, c)$ with some parameters $c = (c_1, \ldots, c_p)$;
- we measure the values $x^{(k)}$ and $y^{(k)}$ $(k = 1, \ldots, K)$ in several situations; and
- we want to find the values x for which $e^{(k)} \stackrel{\text{def}}{=} y^{(k)} - f(x^{(k)}, c) \approx 0$.

An important particular case of this class of problems is formed by *inverse problems*, when we need to find, e.g., the spatial density distribution c which leads to the observed values $y^{(k)}$ of the gravity field at different spatial locations $x^{(k)}$.

Finding parameters of the model: traditional approach. Usually, the main reason why the observed value $y^{(k)}$ is somewhat different from $f(x^{(k)}, c)$ is that measurements are never absolutely accurate. As a result, even if the model is exact, i.e., if $y = f(x, c)$ for the actual (unknown) values x and y, the measurement results $y^{(k)}$ and $x^{(k)}$ are slightly different from y and x, and thus, $y^{(k)} - f(x^{(k)}, c) \neq 0$.

Often, there are many different independent factors leading to measurement errors. It is known that, under certain reasonable conditions, the distribution of the joint effect of many independent factors is close to normal; this result is known as the Central Limit Theorem (see, e.g., [14]). Under these conditions, it is reasonable to conclude that the measurement errors $y^{(k)} - y$ are normally distributed, and thus, the differences $e^{(k)} = y^{(k)} - f(x^{(k)}, c)$ are normally distributed.

If the mean of the measurement error is different from 0, i.e., if the measuring instrument has a bias, then we can recalibrate this instrument and make this difference equal to 0. Thus, the differences $e^{(k)}$ are independent and normally distributed with mean 0. Usually, the mean square magnitude of the measurement

error does not change much within a reasonable range; so, we can safely assume that all the variables $e^{(k)}$ have the same standard deviation σ. Under this assumption, the probability density corresponding to each residual $e^{(k)}$ has the form

$$\frac{1}{\sqrt{2\pi}\cdot\sigma}\exp\left(-\frac{\left(e^{(k)}\right)^2}{2\sigma^2}\right).$$

Since measurement errors corresponding to different measurements are independent, the probability density corresponding to all the residual values $e^{(1)}$, ..., $e^{(K)}$ is equal to

$$L=\prod_{k=1}^{K}\frac{1}{\sqrt{2\pi}\cdot\sigma}\exp\left(-\frac{\left(e^{(k)}\right)^2}{2\sigma^2}\right).$$

It is reasonable to find the values of the parameters c for which this probability is the largest possible; this idea is known as the *Maximum Likelihood approach*. Maximizing L is equivalent to minimizing $-\ln(L)=\text{const}+\frac{1}{2\sigma^2}\cdot\sum_{k=1}^{K}\left(e^{(k)}\right)^2$, i.e., equivalently, to minimizing the sum $\sum_{k=1}^{K}\left(e^{(k)}\right)^2$.

Need to go beyond normal distributions. Empirically, in about half of the cases, the distribution of measurement error is different from normal; see, e.g., [9, 10].

In many such situations, the results produced by the Least Squares method can be misleading; see, e.g., [7]. For example, in the simplest case when there is no input and $y=c$, we have a system of approximate equations $e^{(k)}=y^{(k)}-c\approx 0$. For this system, the Least Squares method leads to the arithmetic average $c=\frac{y^{(1)}+\ldots+y^{(K)}}{K}$. For non-normal distributions, we can have outliers (i.e., values for which $y^{(k)}\ll c$ or $y^{(k)}\gg c$) with a reasonable probability. In this case, if $c=0$ and almost all out of 100 observed values are within the interval $[-0.1,0.1]$, but one value $y^{(k)}=10^6$ is an outlier, the arithmetic average is equal to $\approx 10^4$.

In situations when we know the probability density functions, we can use the Maximum Likelihood method; however, in many practical situations, we do not know the probability distribution. To decrease the dependence on outliers in such situations, special *robust* statistical methods have been designed; see, e.g., [7]. Among the most empirically successful methods of this type are *ℓ^p-methods*, when we minimize $\sum\left|e^{(k)}\right|^p$ for some $p\le 2$. The parameter p has to be determined empirically; in many situations, $p\approx 1$ is the best estimate (see, e.g., [5]), so we arrive at the need to minimize the sum of absolute values $\sum\left|e^{(k)}\right|$.

Another important class of situations: sparse problems. In many practical situations, we know that only a few parameters c_i are different from 0, i.e., that the corresponding tuple c is *sparse*; see, e.g., [4, 6, 8]. In this case, it is reasonable to select, among all tuples c for which $e^{(k)}\approx 0$ for all k, a tuple with the smallest number of non-zero components.

It turns out that in general, finding such c is computationally intractable (NP-hard) [8], but under some reasonable assumptions, the corresponding optimization problem is equivalent to an easier-to-solve problem of minimizing the sum $\sum_{i=1}^{p}|c_i|$; see, e.g., [2, 3, 11, 12].

Summarizing: it is important to be able to minimize $\sum|c_i|$. We have considered two practical situations, in both of which there is a need to minimize an objective function of the type $\sum|c_i|$.

Need for a smooth approximation of $|x|$. It is known that the most efficient optimization techniques use derivatives of the objective function; see, e.g., [15]. Unfortunately, we cannot directly apply these techniques to an objective function of the type $\sum|c_i|$, since $|x|$ is not differentiable at $x=0$. To apply the corresponding efficient optimization techniques, it is therefore necessary to approximate $|x|$ by a smooth function.

A currently used approximation. In [1, 13, 12], a smooth approximation $|x|\approx\sqrt{x^2+\mu}$ was efficiently used.

A natural question and what we do in this paper. The empirical efficiency of the approximation $\sqrt{x^2+\mu}$ naturally prompts the following question: is this approximation the most efficient one, or there are other approximations which lead to even more efficient algorithms?

In this paper, we prove that, in some reasonable sense, the function $\sqrt{x^2+\mu}$ is the most computationally efficient smooth approximation to $|x|$.

2 Analysis of the Problem and the Main Result

Requirements for the desired smooth approximation. We want to select a smooth approximation $f(x) \approx |x|$. The main problem with the function $|x|$ is that it is not differentiable for $x = 0$; for large x, the function $|x|$ is perfectly differentiable, its derivative is equal to the sign $\text{sign}(x)$. It is therefore reasonable to require that the approximation be perfect for large x, i.e., that $\lim\limits_{x\to\pm\infty} \dfrac{f(x)}{|x|} = 1$ and that for the derivative $f'(x)$, we similarly have $\lim\limits_{x\to\pm\infty} \dfrac{f'(x)}{\text{sign}(x)} = 1$.

Definition 1. *We say that a function f from real numbers to real numbers is a* smooth approximation to $|x|$ *if this function is differentiable and satisfies the following two limit properties:* $\lim\limits_{x\to\pm\infty} \dfrac{f(x)}{|x|} = 1$ *and* $\lim\limits_{x\to\pm\infty} \dfrac{f'(x)}{\text{sign}(x)} = 1$.

Comment. This is a rather weak definition of an approximation: e.g., according to this definition, for large N, a function $\sqrt{x^2+\mu}+N$ is a smooth approximation to $|x|$. A more realistic definition of an approximation should include other requirements as well. However, since our main result is applicable to all the functions satisfying Definition 1, it is thus applicable to all functions satisfying a stronger definition as well.

Which computations are efficient: a brief reminder. In a computer, arithmetic operations (addition, substraction, multiplication, and division) are hardware supported and therefore, very efficient. In contrast, computation of any other function, be it $\sqrt{x}$, $\sin(x)$, $\ln(x)$, $\exp(x)$, etc., consists of many consecutive arithmetic operations and is, therefore, order of magnitude slower than a single arithmetic operation.

So, if we want to make our computations most efficient, we must minimize the number of non-arithmetic functions used in computing $f(x)$ and $f'(x)$.

Observation: when computing $f(x)$, non-arithmetic operations are unavoidable. Let us first explain that a smooth approximation $f(x)$ to $|x|$ cannot be computed by only using arithmetic operations. By induction, we can prove that any function composed of arithmetic operations is a rational function $\dfrac{P(x)}{Q(x)}$, where $P(x) = p_0 + p_1 \cdot x + \ldots + p_\ell \cdot x^\ell$ and $Q(x) = q_0 + q_1 \cdot x + \ldots + q_j \cdot x^j$ are polynomials. One can check that for a rational function, when $x \to \pm\infty$, the ratio $\dfrac{P(x)}{Q(x)}$ is asymptotically equivalent to $\text{const} \cdot x^{\ell-j}$. This expression cannot be asymptotically equivalent to $|x|$ both for $x \to +\infty$ and for $x \to -\infty$; thus, a smooth approximation to $|x|$ cannot be computed by using only arithmetic operations – it requires at least one non-arithmetic operation for its computation.

Our idea. We have just shown that we cannot avoid using a (time-consuming) non-arithmetic operation when computing $f(x)$. At first glance, it may feel that a similar result holds for $f'(x)$: a rational function cannot be asymptotically equal to $\text{sign}(x)$ either. However, such a conclusion about $f'(x)$ would be somewhat misleading.

Indeed, in the optimization algorithms, we usually compute the derivative $f'(x)$ *after* we have computed $f(x)$. Because of this, when we compute the derivative $f'(x)$, we can use not only the input x, we can also use the already computed value $f(x)$.

Using time-consuming non-arithmetic operations is unavoidable when we compute $f(x)$. However, when we use the computed value $f(x)$ to compute the value of the derivative $f'(x)$, it would be nice not to use

non-arithmetic operations at all, and to use as few arithmetic operations as possible. Let us describe this idea in precise terms.

Definition 2.

- *We say that a function $F(x_1, \ldots, x_n)$ is* computable in zero arithmetic steps *if it coincides either with one of the inputs x_i or with a constant c.*
- *If a function $F_1(x_1, \ldots, x_n)$ is computable in s_1 arithmetic steps, a function $F_2(x_1, \ldots, x_n)$ is computable in s_2 arithmetic steps, and $\oplus$ is one of the four arithmetic operations, then we say that a function $F(x_1, \ldots, x_n) = F_1(x_1, \ldots, x_n) \oplus F_2(x_1, \ldots, x_n)$ is* computable in $s_1 + s_2 + 1$ arithmetic steps.

Example. A function $f(x) = (x - 1) \cdot (x + 1)$ is computable in three arithmetic steps:

- first, we compute $x - 1$;
- then, we compute $x + 1$;
- finally, we compute the product.

Let us describe this idea in terms of the above formal definition.

- On the first step, since x and 1 are computable in $s_1 = s_2 = 0$ arithmetic steps, we conclude that the difference $x - 1$ is computable in $s_1 + s_2 + 1 = 0 + 0 + 1 = 1$ arithmetic step.
- On the second step, since x and 1 are computable in $s_1 = s_2 = 0$ arithmetic steps, we conclude that the sum $x + 1$ is computable in $s_1 + s_2 + 1 = 0 + 0 + 1 = 1$ arithmetic step.
- Finally, since each of the two functions $F_1 = x - 1$ and $F_2 = x + 1$ can computed in $s_1 = s_2 = 1$ arithmetic step, we conclude that the product $(x-1)\cdot(x+1)$ is computable in $s_1+s_2+1 = 1+1+1 = 3$ arithmetic steps.

(It is worth mentioning that if we rewrite the above expression in the equivalent form $x \cdot x - 1$, we can conclude that the function $(x - 1) \cdot (x + 1)$ can be also computed in two arithmetic steps.)

Proposition. *Among all smooth approximations to $|x|$, the function for which it takes the smallest number of arithmetic steps to compute $f'(x)$ from x and $f(x)$ in the function $f(x) = \sqrt{x^2 + \mu}$.*

Discussion.

- In this sense, the function $\sqrt{x^2 + \mu}$ is indeed the most computationally efficient smooth approximation to $|x|$.
- This result uses a weak definition of a smooth approximation (Definition 1). It is therefore applicable to any stronger definition of a smooth approximation – as long as:
  - every smooth approximation in the sense of this new definition satisfies Definition 1, and
  - $f(x) = \sqrt{x^2 + \mu}$ is a smooth definition in the new sense as well.

Proof.

1°. Let us first show that it is not possible to have a smooth approximation to $|x|$ for which $f'(x)$ can be computed from x and $f(x)$ in zero arithmetic steps.

Indeed, in this case, we would have $f'(x) = x$, $f'(x) = f(x)$, or $f'(x) = c$. In all these three cases, we do not get the correct asymptotic for $f'(x)$ when $x \to \pm\infty$:

- in the first case, $f'(x) = x \to \infty$, while, according to Definition 1, we should have $f'(x) \to \pm 1$;
- in the second case $f'(x) = f(x)$, while, according to Definition 1, we should have $f'(x) \sim \text{sign}(x) \neq |x| \sim f(x)$;

- in the third case, $f'(x) = c$ has the same limit c when $x \to +\infty$ and when $x \to -\infty$, while, according to Definition 1, we should have two different limits $+1$ and -1.

2°. Because of Part 1 of this proof, we need at least one arithmetic step to compute $f'(x)$ from x and $f(x)$, i.e., we need to apply at least one arithmetic operation $\oplus$ to the values x, $f(x)$, and c.

Let us consider all possible situations when exactly one arithmetic operation $\oplus$ is applied.

2.1°. When the operation $\oplus$ is addition, we get three possible cases: $f'(x) = f(x) + x$, $f'(x) = f(x) + c$, and $f'(x) = x + c$. In all these cases, the corresponding equality is inconsistent with the asymptotics described in Definition 1.

2.2°. When the operation $\oplus$ is subtraction, we get six possible cases: $f'(x) = f(x) - x$, $f'(x) = x - f(x)$, $f'(x) = f(x) - c$, $f'(x) = c - f(x)$, $f'(x) = x - c$, and $f'(x) = c - x$. One can check that in all these six cases, the corresponding equality is inconsistent with the asymptotics described in Definition 1.

2.3°. When the operation $\oplus$ is multiplication, we get three possible cases: $f'(x) = f(x) \cdot x$, $f'(x) = f(x) \cdot c$, and $f'(x) = x \cdot c$. In all these cases, the corresponding equality is inconsistent with the asymptotics described in Definition 1.

2.4°. Finally, when the operation $\oplus$ is division, we get six possible cases: $f'(x) = f(x)/x$, $f'(x) = x/f(x)$, $f'(x) = f(x)/c$, $f'(x) = c/f(x)$, $f'(x) = x/c$, and $f'(x) = c/x$. One can check that in the last four cases, the corresponding equality is inconsistent with the asymptotics described in Definition 1. Thus, the only remaining cases are $f'(x) = f(x)/x$ and $f'(x) = x/f$. Let us consider these two cases one by one.

2.4.1°. When $f'(x) = \dfrac{df}{dx} = \dfrac{f}{x}$, we can move all the terms related to f to one side and all the terms related to x to another side and get $\dfrac{df}{f} = \dfrac{dx}{x}$. Integrating this equality, we conclude that $\ln(f) = \ln(x) + \text{const}$. Applying $\exp(z)$ to both sides of this equality, we conclude that $f(x) = \text{const} \cdot x$. One can easily check that this linear function does not satisfy asymptotics required by Definition 1.

2.4.2°. The only remaining case is when $f'(x) = \dfrac{df}{dx} = \dfrac{x}{f}$. In this case, we can also move all the terms related to f to one side and all the terms related to x to another side and get $f \cdot df = x \cdot dx$. Integrating this equality, we conclude that $\dfrac{f^2}{2} = \dfrac{x^2}{2} + C$, for some integration constant C. Thus, $f^2 = x^2 + \mu$, where we denoted $\mu \stackrel{\text{def}}{=} 2 \cdot C$. Hence, we get $f(x) = \sqrt{x^2 + \mu}$.

3°. Summarizing:

- there is no smooth approximation $f(x)$ to $|x|$ for which $f'(x)$ can be computed in zero arithmetic steps, and
- the only smooth approximation for which $f'(x)$ can be computed in one arithmetic step is $f(x) = \sqrt{x^2 + \mu}$.

Thus, among all smooth approximations to $|x|$, the function $f(x) = \sqrt{x^2 + \mu}$ indeed has the property that the value $f'(x)$ can be computed from the known values x and $f(x)$ in the smallest number of arithmetic steps. The proposition is proven.

Acknowledgments.

This work was supported in part by the National Science Foundation grants HRD-0734825 and HRD-1242122 (Cyber-ShARE Center of Excellence) and DUE-0926721, by Grants 1 T36 GM078000-01 and 1R43TR000173-01 from the National Institutes of Health, and by a grant N62909-12-1-7039 from the Office of Naval Research.

References

[1] M. Argáez, C. Ramirez, and R. Sanchez, "An ℓ_1 algorithm for underdetermined systems and applications", *IEEE Proceedings of the 2011 Annual Conference on North American Fuzzy Information Processing Society NAFIPS'2011*, El Paso, Texas, March 18–20, 2011, pp. 1–6.

[2] E. J. Candès, J. Romberg and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements", *Comm. Pure Appl. Math.*, 2006, Vol. 59, pp. 1207–1223.

[3] E. J. Candès, and T. Tao, "Decoding by linear programming", *IEEE Transactions on Information Theory*, 2005, Vol. 51, No. 12, pp. 4203–4215.

[4] D. L. Donoho, "Compressed sensing", *IEEE Transactions on Information Theory*, 2006, Vol. 52, No. 4, pp. 1289–1306.

[5] D. I. Doser, K. D. Crain, M. R. Baker, V. Kreinovich, and M. C. Gerstenberger "Estimating uncertainties for geophysical tomography", *Reliable Computing*, 1998, Vol. 4, No. 3, pp. 241–268.

[6] M. Elad, *Sparse and Redundant Representations*, Springer Verlag, 2010.

[7] P. J. Huber, *Robust Statistics*, Wiley, Hoboken, New Jersey, 2004.

[8] B. K. Natarajan, "Sparse approximate solutions to linear systems", *SIAM Journal on Computing*, 1995, Vol. 24, pp. 227–234.

[9] P. V. Novitskii and I. A. Zograph, *Estimating the Measurement Errors*, Energoatomizdat, Leningrad, 1991 (in Russian).

[10] A. I. Orlov, "How often are the observations normal?", *Industrial Laboratory*, 1991, Vol. 57, No. 7, pp. 770–772.

[11] C. Ramirez, V. Kreinovich, and M. Argaez, "Why ℓ_1 Is a Good Approximation to ℓ_0: A Geometric Explanation", *Journal of Uncertain Systems*, 2013, Vol. 7, to appear.

[12] C. Ramirez and M. Argaez, "An ℓ_1 minimization algorithm for non-smooth regularization in image processing", *Signal, Image and Video Processing*, to appear, DOI:10.1007/s11760-013-0454-1

[13] R. Sanchez, M. Argaez, and P. Guillen, "Sparse Representation via l^1-minimization for Underdetermined Systems in Classification of Tumors with Gene Expression Data", *Proceedings of the IEEE 33rd Annual International Conference of the Engineering in Medicine and Biology Society EMBC'2011 "Integrating Technology and Medicine for a Healthier Tomorrow"*, Boston, Massachusetts, August 30 – September 3, 2011, pp. 3362–3366.

[14] D. J. Sheskin, *Handbook of Parametric and Nonparametric Statistical Procedures*, Chapman & Hall/CRC, Boca Raton, Florida, 2007.

[15] S. A. Vavasis, *Nonlinear Optimization: Complexity Issues*, Oxford University Press, New York, 1991.