Measuring the Likelihood of Numerical Constraints

Marco Console , **Matthias Hofer** and **Leonid Libkin**

School of Informatics, University of Edinburgh

{mconsole, mhofer, libkin}@inf.ed.ac.uk

Abstract

Our goal is to measure the likelihood of the satisfaction of numerical constraints in the absence of prior information. We study expressive constraints, involving arithmetic and complex numerical functions, and even quantification over numbers. Such problems arise in processing incomplete data, or analyzing conditions in programs without a priori bounds on variables. We show that for constraints on n variables, the proper way to define such a measure is as the limit of the part of the n-dimensional ball that consists of points satisfying the constraints, when the radius increases. We prove that the existence of such a limit is closely related to the notion of o-minimality from model theory. Thus, for constraints definable with the usual arithmetic and exponentiation, the likelihood is well defined, but adding trigonometric functions is problematic. We look at computing and approximating such likelihoods for order and linear constraints, and prove an impossibility result for approximating with multiplicative error. However, as the likelihood is a number between 0 and 1, an approximation scheme with additive error is acceptable, and we give it for arbitrary linear constraints.

1 Introduction

Our goal is to measure the likelihood of numerical constraints without having any additional knowledge such as a probability distribution. A constraint over n variables can be seen as a map $\varphi(\bar{x}) : \mathbb{R}^n \to \{0, 1\}$ or a subset $\varphi(\mathbb{R}^n) = \{\bar{a} \in \mathbb{R}^n \mid \varphi(\bar{a}) = 1\}$ of tuples in $\mathbb{R}^n$ satisfying the constraint. To express them, we can use standard arithmetic operations $+, -, \cdot, \div$, and even more complex ones such as $e^x, \ln x$. For example, $x^2 < y < e^x$ is a constraint, and we can even extend the language with quantifiers and write constraints such as $\exists z \, (e^z \cdot x = z \cdot y^2)$.

For a constraint $\varphi(\bar{x})$, we would like to measure how likely it is to be satisfied. If we are only interested in a bounded subset of $\mathbb{R}^n$, we can just compute the volume of the intersection of $\varphi(\mathbb{R}^n)$ with that subset. Without explicit bounds, if we have a probability distribution $p(\bar{x})$ on $\mathbb{R}^n$, this likelihood is just the expected value of φ, i.e., $\int_{\mathbb{R}^n} \varphi(\bar{x})p(\bar{x})d\bar{x}$.

But what happens when we have neither explicit bounds nor a probability distribution? Since there is no uniform distribution over the entire $\mathbb{R}^n$, we would like to, instead, estimate the "portion" that $\varphi(\mathbb{R}^n)$ cuts from $\mathbb{R}^n$, and view it as the likelihood of φ. But while it is intuitively clear that a tautological constraint like $1 > 0$ is likelier than $x \geq 1$ which in turn is likelier than $x = 0$, how do we formally explain this?

The motivation for these questions comes from several different areas. One is program analysis, whose relationship with constraint solving is discussed in [Gulwani *et al.*, 2008]. Likelihoods of a branch being taken, and more generally, of an execution path of a program, have been associated with volumes of certain sets in $\mathbb{R}^n$ ([Ma *et al.*, 2009]). This correspondence only works when there are explicit bounds on input variables (and thus volumes are finite). However, ad hoc bounds on variables can affect likelihoods drastically. To reason about execution paths without such bounds, we need to estimate likelihoods of arbitrary constraints over $\mathbb{R}^n$. Another motivating example comes from running queries on incomplete data, i.e., data with explicit markers saying that some values are unknown (e.g., nulls in databases). If we want to estimate the likelihood of answers to queries over such data, we need to estimate the likelihood of conditions specified in queries. This is useful for estimating how close an answer is to certainty, as was shown in [Libkin, 2018]. That work, however, only considered equality constraints, while database queries often use more complex conditions. How to estimate the likelihood of these conditions is yet to be defined. The third motivation comes from temporal constraint satisfaction ([Dechter *et al.*, 1991]). If we are not interested in mere satisfiability, but we want to estimate how likely the events described by a temporal constraint network are, we need to estimate the proportion of $\mathbb{R}^n$ that its constraints define. As a final example, a continuous analog of weighted model counting for propositional formulae ([Chavira and Darwiche, 2008]) was proposed in [Belle *et al.*, 2015]. The idea behind it is to compute volumes of sets of continuous parameters of models; but these volumes must be finite and hence parameters bounded. To lift this restriction, one would need to replace such volumes by measures of likelihoods of unbounded sets.

Thus, the question we are addressing is not only mathematically natural, but also relevant in several different scenarios. We now look at a few examples to indicate that the question is not trivial. Start with a simple constraint $x \geq y$ over $\mathbb{R}^2$. It
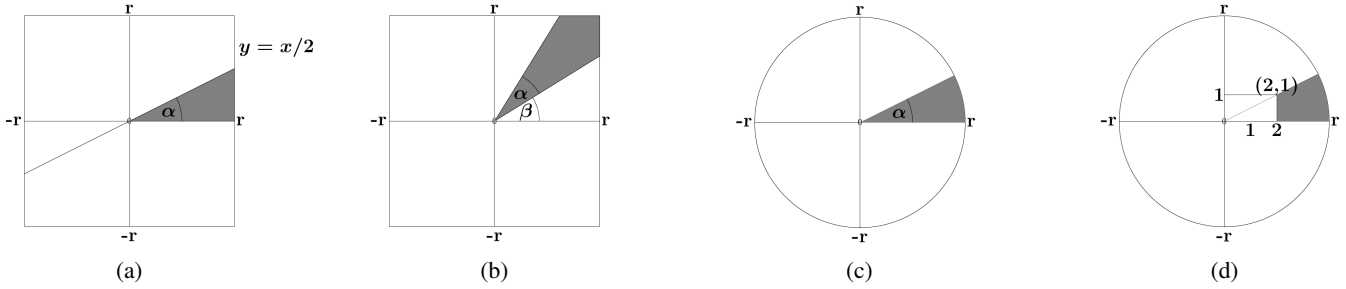
Figure 1: Illustration to the examples in the introduction

seems quite obvious that if we know nothing about x and y, its likelihood is $1/2$. But what about $x \geq 2y$? On a quick reflection one realizes that nothing changes: for $x \geq y$ we had symmetry along the $y = x$ line, and now we have symmetry along the $y = x/2$ line, so the likelihood is still $1/2$.

But now change things a bit and consider $\varphi(x, y) = (x \geq 2y) \wedge (x, y \geq 0)$. If we consider an explicit bound $-r \leq x, y \leq r$, for some $r > 0$, then this constraint cuts $\frac{1}{16}$-th of the square $[-r, r]^2$, see Fig. 1 (a). More precisely, $\lambda(\varphi(\mathbb{R}^2) \cap [-r, r]^2)/\lambda([-r, r]^2) = \frac{1}{16}$, where $\lambda(\cdot)$ refers to the Lebesgue measure (area in 2d, volume in 3d).

This looks like a sensible way of defining the proportion we are interested in, until we apply a simple isometry transformation which should not change the value. In this case, rotate $\varphi(\mathbb{R}^2)$ anticlockwise β degrees where $\beta = \pi/4 - \arctan(1/2)/2$. For the new set C_β, shown in Fig. 1 (b), we have $\lambda(C_\beta \cap [-r, r]^2)/\lambda([-r, r]^2) = (1 - \tan(\beta))/4 \neq 1/16$. So, considering identical bounds $[-r, r]$ along both axes does not work well.

What if we use the radius r ball B_r^n instead? This looks like a better idea: in this case, $\lambda(\varphi(\mathbb{R}^2) \cap B_r^2)/\lambda(B_r^2) = \arctan(1/2)/2\pi$, see Fig. 1 (c). In fact every rotation or an isometry in general, applied to $\varphi(\mathbb{R}^2)$ (for instance, one resulting in C_β) will give us the same result.

Thus, using balls to measure the likelihood leads to more intuitive results. However, in the previous example we were lucky as the ratio did not depend on the radius r. This is not always so: consider a slight modification of φ, namely $\varphi'(x, y) = (x \geq 2y) \wedge (x, y \geq 1)$, see Fig. 1 (d). Then $\lambda(\varphi'(\mathbb{R}^2) \cap B_r^2)/\lambda(B_r^2) = \arctan(1/2)/2\pi - 1/\pi r^2$. How do we get rid of the dependence on the radius? The idea, similarly to what is done in the study of asymptotic behavior of properties defined by logical sentences ([Spencer, 2010]) is to consider the limit as r increases. In this case $\lim_{r \to \infty} \lambda(\varphi'(\mathbb{R}^2) \cap B_r^2)/\lambda(B_r^2) = \arctan(1/2)/2\pi$. Intuitively, the change from $x, y \geq 0$ to $x, y \geq 1$ affects a very small portion of the set defined by φ', and can be disregarded.

This is the proper way to define the likelihood that we shall follow. However, there is a caveat: do we know that the limit always exists? In fact it does not: consider a new constraint $\varphi''(x)$ saying that $2^{2k} \leq |x| \leq 2^{2k+1}$ for some $k \in \mathbb{N}$. In this case, as we shall formally prove later, the limit does not exist, and there seems to be no good way of defining the likelihood (in fact the likelihood will be oscillating in the $[\frac{1}{3} - \varepsilon, \frac{2}{3} - \varepsilon]$ interval, with decreasing ε as k increases).

These simple examples show that the question of defining the likelihood of numerical constraints is far from trivial and several questions need to be answered.

- How do we define it properly? The examples seem to indicate that the ball behaves better than the cube. Note that those were defining sets $\|\bar{a}\|_2 \leq r$ and $\|\bar{a}\|_\infty \leq r$ for the Euclidean norm $\|\cdot\|_2$ and the infinity norm $\|\cdot\|_\infty$. But there are many other norms one could consider.

- How do we know that the limit exists and the likelihood of a constraint is well defined?

- If the limit does exist and the likelihood is well defined, can we compute it? If it is not a rational number, like $\arctan(1/2)/2\pi$ in the example, can we approximate it, and how efficiently can we do so?

We answer these questions and establish the following results. First, we formally define the likelihood $\mu(\varphi)$ for numerical constraints φ and show that this can only be properly done when one considers the Euclidean norm $\|\cdot\|_2$.

We then address the question of existence of the limit. We prove that it exists for arbitrary constraints definable in the first-order theory of the structure $\langle \mathbb{R}, +, \cdot, e^x \rangle$. In particular such constraints can use $+, \cdot, -, \div, e^x, \ln x$. This relies on deep results in logic, related to the notion of o-minimal structures ([Van den Dries, 1998]). O-minimality states that subsets of $\mathbb{R}$ definable in first-order logic are finite unions of intervals. This condition is violated by adding, for instance, a predicate $\mathbb{N}(\cdot)$ for natural numbers, or the sine function. We prove that adding either a test for natural numbers, or any trigonometric function leads to a situation when the limit does not exist and the likelihood cannot be defined.

Finally, we look at computing and approximating such likelihoods. Even in the simplest case of order constraints, finding the exact likelihood is intractable. For linear constraints, the exact likelihoods, even for simplest constraints, are usually irrational numbers, and thus need to be approximated. We consider approximations of two standard kinds: with multiplicative error, and with additive error, see [Motwani and Raghavan, 1995]. We prove that the former does not exist under widely held complexity-theoretic assumptions. However, as we are approximating numbers between 0 and 1, approximation schemes with additive error are acceptable. We present such a schema for arbitrary linear constraints.

Organization. Basic notations are given in Section 2. The definition of the likelihood of constraints is given in Section 3.

Its existence is studied in Section 4. Complexity of computing and approximating likelihood is investigated in Section 5 for order constraints, and in Section 6 for linear constraints. Conclusions are in Section 7.

2 Preliminaries

We use $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, and $\mathbb{R}$ to denote natural, integer, rational, and real numbers, and superscript n to denote n-tuples of those, also referred to as (n-dimensional) points. For each $p \geq 1$, the p-norm $\|\bar{a}\|_p$ of $\bar{a} \in \mathbb{R}^n$ is defined as $(\sum_{i=1}^n |a_i|^p)^{1/p}$. Furthermore, $\|\bar{a}\|_\infty = \lim_{p\to\infty} \|\bar{a}\|_p = \max\{|a_1|, \ldots, |a_n|\}$. If p is omitted, we implicitly assume the Euclidean norm, i.e., $p = 2$ and $\|\bar{a}\| = \|\bar{a}\|_2$. For $\bar{c} \in \mathbb{R}^n$ and $r \in \mathbb{R}$, the n-dimensional p-ball of radius r and center $\bar{c}$ is $B_{p,r}^n(\bar{c}) = \{\bar{a} \in \mathbb{R}^n \mid \|\bar{a} - \bar{c}\|_p \leq r\}$. If p and $\bar{c}$ are omitted, we assume $p = 2$ and $\bar{c} = \bar{0}$, i.e., $B_r^n = B_{2,r}^n(\bar{0})$.

If $S \subseteq \mathbb{R}^n$, we denote by $\lambda^n(S)$ the n-dimensional Lebesgue measure of S. When n is clear from the context, we write $\lambda(S)$ for the n-dimensional Lebesgue measure of S. The Lebesgue measure is the unique extension of the standard notion of length, area and volume, in 1, 2 and 3 dimensions, to the n-dimensional case.

Constraint Languages

We consider first-order languages over structures on numbers, mainly the reals. Let Ω be the vocabulary containing function and predicate symbols. Function symbols will be standard numerical functions such as $+, \cdot, e^x, \ln x, \sin x$, etc, and, likewise, predicates are the standard predicates, such as $=, <$, etc. Constants are viewed as functions of arity zero. First-order formulae in the language of Ω are defined in the standard way. Assume a countably infinite set VAR of symbols, called set of variables. Each variable and each constant is a term. Then, if $t_1, \ldots, t_n$ are terms, and f is a function symbol of arity n, $f(t_1, \ldots, t_n)$ is a term, too. If P is a k-ary predicate symbol from Ω, and $t_1, \ldots, t_k$ are terms, $P(t_1, \ldots, t_k)$ is an atomic formulae. Furthermore, if φ and ψ are formulae, $\varphi \vee \psi, \varphi \wedge \psi, \neg\varphi$ are formulae, as are $\exists x\, \varphi$ and $\forall x\, \varphi$. The definitions of bound and free variables are standard. We write $t(\bar{x})$ and $\varphi(\bar{x})$ if the (distinct) free variables of t and φ are $\bar{x}$, and say that φ is of arity n (or n-ary) if it contains exactly n distinct free variables.

Semantics for terms and formulae in $\mathrm{FO}(\Omega)$ is defined with respect to a structure $\mathfrak{M} = \langle \mathbb{R}, \Omega \rangle$ and an assignment $\bar{a} \in \mathbb{R}^n$ for free variables $\bar{x} \in \mathrm{VAR}^n$. For the value of a term, we write $t(\bar{a})$. For example, given a term $t(x, y) = e^x + y$, the value $t(0, 1)$ is 2. We say that $\varphi(\bar{a})$ is true in $\mathfrak{M}$ when it is satisfied under the standard first-order semantics. For example, $\varphi(u, v, w) = \exists x\, (u \cdot x \cdot x + v \cdot x + w = 0)$ is formula in three free variables; $\varphi(a, b, c)$ is true iff $b^2 - 4ac \geq 0$.

For a formula $\varphi(\bar{x})$ with $\bar{x} \in \mathrm{VAR}^n$, let Ω_φ contain all the function and predicate symbols used in φ. Define

$$[\![\varphi]\!] = \{\bar{a} \in \mathbb{R}^n \mid \varphi(\bar{a}) \text{ is true in } \langle \mathbb{R}, \Omega_\varphi \rangle\}.$$

Since we assume the standard semantics of numerical functions and predicates, the above definition does not depend on the exact vocabulary Ω of a structure as long as $\Omega_\varphi \subseteq \Omega$. Indeed, then $\varphi(\bar{a})$ is true in $\langle \mathbb{R}, \Omega \rangle$ iff it is true in $\langle \mathbb{R}, \Omega_\varphi \rangle$.

3 Defining the Likelihood

To define the likelihood of a constraint given by a formula $\varphi(\bar{x})$ over $\langle \mathbb{R}, \Omega \rangle$, we need, as explained in the introduction, to estimate the portion of $\mathbb{R}^n$ occupied by $[\![\varphi]\!]$. We first look at the well-defined portion of $B_{p,r}^n = \{\bar{a} \in \mathbb{R}^n \mid \|\bar{a}\|_p \leq r\}$ occupied by $[\![\varphi]\!]$ in the radius r ball (with respect to the $\|\cdot\|_p$ norm), namely $\lambda([\![\varphi]\!] \cap B_{p,r}^n)/\lambda(B_{p,r}^n)$, and then consider its asymptotic behavior as the bound r increases. This gives us the following definition.

Definition 1. *Let $\varphi(\bar{x})$ be a formula with n-free variables over the vocabulary Ω. The p-likelihood of φ is defined as:*

$$\mu_p(\varphi) = \lim_{r\to\infty} \frac{\lambda([\![\varphi]\!] \cap B_{p,r}^n)}{\lambda(B_{p,r}^n)}$$

To answer the question which value p to choose, note that we do not expect the measure to change when simple volume-preserving transformations, such as rotations or reflections, are applied to $[\![\varphi]\!]$. These in general fall in the class of isometries, i.e., linear transformations given by orthogonal matrices M that satisfy $M^T = M^{-1}$ (i.e. the transpose of M equals its inverse). Note that, as long as Ω contains $+$ and $\cdot$, the set $M[\![\varphi]\!]$ is definable in the structure $\langle \mathbb{R}, \Omega \rangle$, i.e., there is a formula $M\varphi$ such that $[\![M\varphi]\!] = M[\![\varphi]\!]$.

We next show that the only norm, for which applying isometries does not change the likelihood μ_p, is the Euclidean norm. More precisely, we say that the likelihood μ_p is *well-behaved* if, for every structure $\langle \mathbb{R}, +, \cdot, \ldots \rangle$, every formula $\varphi(\bar{x})$, and every orthogonal matrix M, we have $\mu_p(\varphi) = \mu_p(M\varphi)$. Since μ_p is defined as the limit, the equality means that, if one side of the equation exists, then so is the other, and they are both equal.

Theorem 1. *The likelihood μ_p is well-behaved iff $p = 2$, i.e., only for the Euclidean norm.*

Proof idea. For all orthogonal matrices M and measurable sets $Z \subseteq \mathbb{R}^n$, we have $\lambda(Z) = \lambda(MZ)$. However $MB_{p,r}^n = B_{p,r}^n$ does not hold for all orthogonal matrices M unless $p = 2$. Hence, in general, also $M(S \cap B_{p,r}^n) \neq MS \cap B_{p,r}^n$, for $p \neq 2$. Therefore, $\lambda(S \cap B_{p,r}^n) = \lambda(M(S \cap B_{p,r}^n)) \neq \lambda(MS \cap B_{p,r}^n)$ for all orthogonal matrices M, unless $p = 2$. $\square$

In view of this, we now concentrate on $\mu_2(\varphi)$, and simply write $\mu(\varphi)$, omitting the subscript 2.

4 Existence of the Likelihood

Due to the way the likelihood is defined, a very natural question to ask is whether $\mu(\varphi)$ exists for all constraints φ definable in $\langle \mathbb{R}, \Omega \rangle$. We now show that the existence of the limit in Definition 1 is closely related to the subsets of $\mathbb{R}$ definable in $\langle \mathbb{R}, \Omega \rangle$. In fact, the existence of these limits is closely related to the property of the structure known as *o-minimality*, cf. [Van den Dries, 1998]. We use this to prove the existence results for very expressive constraints that can use functions such as $+, -, \cdot, \div, e^x, \ln x$ and others. We then show that the most common ways of violating o-minimality, such as adding a test for natural numbers of trigonometric functions, lead to situations where the limit does not exist.

Existence Results

The property of o-minimality states that for every formula in one free variable $\varphi(x)$, the set $\varphi(\mathbb{R}) \subseteq \mathbb{R}$ is a finite union of intervals. This is easily seen to be true for the real field $\langle \mathbb{R}, +, \cdot, 0, 1, < \rangle$. By quantifier-elimination (cf. [Caviness and Johnson, 1998]), every formula $\varphi(x)$ is equivalent to a Boolean combination of statements $p_1(x) > 0, \ldots, p_m(x) > 0$, where all the p_is are polynomials. Polynomials that are not identically zero have finitely many roots. Let $r_1 < r_2 < \ldots < r_k$ be all the roots of polynomials $p_1(x), \ldots, p_m(x)$; then in each interval (r_i, r_{i+1}), the truth of φ cannot change.

We now consider an expansion of this structure, the exponential field $\langle \mathbb{R}, +, \cdot, e^x \rangle$. Let us briefly list what is definable in $\mathrm{FO}(<, +, \cdot, e^x)$, to understand the power of these constraints. Constants 0 and 1 are definable as units of $+$ and $\cdot$; for example $\zeta_1(x) = \forall y\, (y \cdot x = y)$ defines 1. Every natural number is definable by a formula $\zeta_\ell(x) = (x = 1 + \ldots + 1)$, where the summation is taken ℓ times, and every rational number n/m is definable too by $\zeta_{n/m}(x) = \exists y \exists z\, (\zeta_n(y) \wedge \zeta_m(z) \wedge x \cdot z = y)$. Inverse operations $-, \div$, $\ln$ are definable (e.g., $x - y = z$ iff $x = y + z$) and for every rational s, the function s^x is definable since $s^x = e^{\ln s \cdot x}$ and both $\ln$ and arbitrary rational numbers are definable. Definability of a partial function means that its graph is definable by an FO formula.

A remarkable result by [Wilkie, 1996] showed that $\langle \mathbb{R}, +, \cdot, e^x \rangle$ is o-minimal. Thus, the set $\mathbb{N}$ is not definable in it, nor are trigonometric functions $\sin$ and $\cos$, as with them one can define sets that are not finite unions of intervals.

Theorem 2. *For every formula $\varphi \in \mathrm{FO}(<, +, \cdot, e^x)$, the likelihood $\mu(\varphi)$ exists.*

Proof sketch. Given a formula $\varphi(\bar{x})$ where $\bar{x} = (x_1, \ldots, x_n)$, consider a new formula $\psi(y, \bar{x}) = \varphi(\bar{x}) \wedge \|\bar{x}\| \leq y$ (the latter condition is definable with $+$ and $\cdot$). Then $\psi(r, \bar{a})$ holds iff $\bar{a} \in \llbracket \varphi \rrbracket \cap B_r^n$. For a fixed $\varepsilon > 0$, a formula $\beta_\varepsilon(y, z)$ is an ε-volume approximation for ψ if for each r, there is v such that $\beta_\varepsilon(r, v)$ holds, and moreover, whenever $\beta_\varepsilon(r, v)$ holds, it is the case that $|v - \lambda(\psi(r, \mathbb{R}^n))| < \varepsilon$, where $\psi(r, \mathbb{R}^n)$ is the set of all $\bar{a} \in \mathbb{R}^n$ such that $\psi(r, \bar{a})$ holds.

Using finiteness of VC dimension of definable families in o-minimal structures [Van den Dries, 1998], it was shown in [Karpinski and Macintyre, 1997; Koiran, 1995] that ε-approximations of volume exists for every ε. (In fact the results there even show that whenever $|v - \lambda(\psi(r, \mathbb{R}^n))| < \varepsilon/4$, then $\beta_\varepsilon(r, v)$ holds). Next we use the fact that o-minimal functions admit definable Skolemization, i.e., there is a definable function (a function with the definable graph) $f_\varepsilon : \mathbb{R} \to \mathbb{R}$ such that $\beta_\varepsilon(r, f_\varepsilon(r))$ holds for every r.

Recall that $\lambda(B_r^n) = c_n \cdot r^n$ where $c_n = \pi^{n/2}/\Gamma(n/2+1)$. Consider the function $F_\varepsilon(r) = f_\varepsilon(r)/r^n$, still definable. Then $|F_\varepsilon(r) - c_n \cdot \lambda(\psi(r, \mathbb{R}^n))/\lambda(B_r^n)| \leq \varepsilon/r^n$. Since $\lambda(\psi(r, \mathbb{R}^n)) < \lambda(B_r^n)$, it tells us that F_ε is bounded when $r \to \infty$. From o-minimality, we know that for F_ε, as a definable function, there exists $r_0 \in \mathbb{R}$ such that F_ε is continuous and either monotonically increasing, or decreasing, for all $r > r_0$. Thus, $\lim_{r \to \infty} F_\varepsilon(r)$ exists, and therefore

by the above, $\lim_{r \to \infty} c \cdot \lambda(\psi(r, \mathbb{R}^n))/\lambda(B_r^n)$ exists. This means that $\mu(\varphi) = \lim_{r \to \infty} \lambda(\psi(r, \mathbb{R}^n))/\lambda(B_r^n)$ exists. $\quad\square$

The crucial ingredients of the proof are o-minimality and definability of volume approximations. The latter works in general for o-minimal expansions of the real field [Karpinski and Macintyre, 1997], of which $\langle \mathbb{R}, +, \cdot, e^x \rangle$ is the most prominent example.

Theorem 2 merely establishes the existence of the limit, but does not give a good approach to calculating it yet. Indeed, to approximate the limit, one would need to explicitly construct volume approximating formulae, and check their satisfiability. There are two obstacles to this approach in general. First, in the case of $\mathrm{FO}(<, +, \cdot, e^x)$ formulae, their decidability is not yet known: it was proved [Macintyre and Wilkie, 1996] under the assumption that Schanuel's conjecture – a major unsolved problem in number theory – holds. For the case when we have decidability of the real field $\langle \mathbb{R}, +, \cdot, < \rangle$, the bounds for ε-volume approximations can be huge: [Benedikt and Libkin, 2002] showed how to construct approximating formulae for very simple geometric shapes and $\varepsilon = 0.1$ that would involve over 10^{10} quantifiers. We shall propose a much better behaved approximation scheme, but only for the linear case, without multiplication.

Non-existence Results

O-minimality is the essential condition for proving the existence result. What happens without o-minimality? Of course there are many ways to produce structures that are not o-minimal, but it is known that any expansion of the real field with nice functions, that is not o-minimal, defines either the set of natural numbers or trigonometric functions. Specifically, the following was shown in [Miller, 2011]. If we look at the expansion of $\langle \mathbb{R}, +, \cdot \rangle$ with functions $f : \mathbb{R} \to \mathbb{R}^n$ that are solutions to $f'(x) = Mf(x)$ for some $n \times n$ matrix M, and the resulting expansion is not o-minimal, then one of the following is definable in this structure: either the set $\mathbb{N}$, or the logarithmic spiral $\{(e^x \cdot \cos(ax), e^x \cdot \sin(ax)) \mid x \in \mathbb{R}\}$. If the latter is added to $\mathrm{FO}(<, +, \cdot, e^x)$, then both $\sin$ and $\cos$ are definable.

In view of this, we consider non-o-minimal structures with a predicate for natural numbers, or a trigonometric function, and show that in them, the likelihood $\mu(\varphi)$ may not exist.

Theorem 3. *Let $\mathbb{N}(\cdot)$ be the unary predicate for natural numbers, and f any trigonometric function (*$\sin, \cos, \tan, \cot$*). Then in both $\mathrm{FO}(+, \cdot, \mathbb{N})$ and $\mathrm{FO}(+, \cdot, e^x, f)$ one can find formulae φ such that $\mu(\varphi)$ does not exist.*

Proof sketch. If unary predicate $\mathbb{N}$ is available, with $+$ and $\cdot$ one can define every computable predicate over $\mathbb{N}$. Hence one can define the following constraint: $\varphi(x) = \exists k\, \big((2^{2k} \leq |x|) \wedge (|x| \leq 2^{2k+1}) \wedge \mathbb{N}(k)\big)$. Note that in one dimension, $B_r^1 = [-r, r]$ and $\lambda(B_r^1) = 2r$. Then one can show that $\lambda(\llbracket \varphi \rrbracket \cap B_{2^{2k+i}}^1)/\lambda(B_{2^{2k+i}}^1) = (1+i)/3 + o(k^{-1})$ for $i = 0, 1$, proving that $\lim_{r \to \infty} \lambda(\llbracket \varphi \rrbracket \cap B_r^1)/\lambda(B_r^1)$ does not exist.

If we have any trigonometric function among $\cos, \tan, \cot$, then in $\mathrm{FO}(+, \cdot)$ we can define $\sin(x)$. Next consider

$\varphi(x, y) = (y > 0) \wedge (x > 0) \wedge (y \leq x \cdot \sin(\ln(x)))$. Then

$$\frac{\lambda(\llbracket \varphi \rrbracket \cap B_r^2)}{\lambda(B_r^2)} = \begin{cases} \frac{1}{5\pi} + o(\frac{1}{r}) & \text{if } r = e^{(2m+1)\pi}, \\ \frac{1}{5\pi e^{2\pi}} + o(\frac{1}{r}) & \text{if } r = e^{(2m+2)\pi}, \end{cases}$$

for $m \in \mathbb{N}$. Thus $\lim_{r \to \infty} \frac{\lambda(\llbracket \varphi \rrbracket \cap B_r^2)}{\lambda(B_r^2)}$ does not exist. $\qquad \square$

5 Order Constraints

We have seen that the likelihood $\mu(\varphi)$ exists for all the formulae of $\text{FO}(<, +, \cdot, e^x)$, although the existence proof does not lend itself to a natural algorithm for finding $\mu(\varphi)$. We thus now look at concrete algorithms for simpler constraints, starting with the basic case of $\text{FO}(<)$. Our results indicate hardness of the exact computation, thus motivating the need to look for approximations. To start with, we make the following easy observation.

Proposition 1. *If $\varphi \in \text{FO}(<)$, then $\mu(\varphi)$ is a computable rational number.*

Proof sketch. Consider $\varphi(x_1, \ldots, x_n)$. For each permutation π on $\{1, \ldots, n\}$, let $\chi_\pi(\bar{x})$ be $\bigwedge_{i \leq n-1}(x_{\pi(i)} < x_{\pi(i+1)})$. We can check, using decidability of $\langle \mathbb{R}, < \rangle$, whether χ_π implies φ. Then $\mu(\varphi)$ is the number of all π satisfying this condition, divided by $n!$ (since orderings of variables in which two are equal do not contribute to the n-dimensional measure). $\qquad \square$

We now address the exact complexity of computing $\mu(\varphi)$. Since it is a number, we must deal with counting complexity classes. Recall that #P is class of function problems that count the number of solutions of an NP problem, cf. [Arora and Barak, 2009]. Unless P = NP, problems which are hard for this class are intractable. Using this, we obtain the following lower bound for computing $\mu(\varphi)$ over $\text{FO}(<)$.

Theorem 4. *Computing $\mu(\varphi)$ for $\varphi \in \text{FO}(<)$ is #P-hard, even for quantifier-free formulae in disjunctive normal form.*

Proof sketch. We reduce from the #P-hard problem #DNF: for a propositional formula α in DNF over variables $\text{VAR}(\alpha)$, count the number #α of truth assignments that satisfy α. We construct an FO formula φ_α such that $\mu(\varphi_\alpha) = f(\#\alpha)$, and both φ_α and function f are polytime-computable. For this sketch, we assume that constant symbols are allowed in φ_α.

We assign to each $v \in \text{VAR}(\alpha)$ a distinct (numerical) variable $x \in \text{VAR}$. To define φ_α, for every clause $(l_1, \ldots, l_m)$ of α, assume that the variable of each literal l_i is v_i and that x_i is the (numerical) variable assigned to v_i. For each $(l_1, \ldots, l_m)$ of α, the formula φ_α contains a disjunct $(a_1 \wedge \cdots \wedge a_m)$ where each a_i is equal to $(x_i > 0)$ if l_i is a positive literal and a_i is equal to $(x_i < 0)$ if l_i is a negative literal. Note that φ_α is a DNF formula with free variables $x_1, \ldots, x_n$. One can show that $\mu(\varphi_\alpha)$ equals the number of orthants where φ_α is satisfied, divided by 2^n. The claim follows since $\mu(\varphi_\alpha) \cdot 2^n$ can be computed in polytime with respect to the size of α. $\qquad \square$

The decision version of the problem is intractable too.

Proposition 2. *Deciding whether $\mu(\varphi) > 0$ is NP-complete, and deciding whether $\mu(\varphi) = 1$ is coNP-complete for quantifier-free formulae $\varphi \in \text{FO}(<)$.*

Proof sketch. The construction in the previous proof shows that a DNF formula α is a tautology iff $\mu(\varphi_\alpha) = 1$, proving hardness. For membership, notice that for $\varphi \in \text{FO}(<)$, $\mu(\varphi) > 0$ (or equals one) iff some (every) strict linear ordering of the free variables of φ satisfies φ. This linear ordering can then be guessed and checked for satisfying φ. $\qquad \square$

6 Linear Constraints

We have seen that computing $\mu(\varphi)$ is hard for order constraints, which suggests that, as is common in such situations, one approximates it instead. When we move to more expressive linear constraints, i.e., $\text{FO}(<, +)$, there is another compelling reason for looking for approximations: numbers $\mu(\varphi)$ may not be rational. In fact, for proving #P-hardness of computing $\mu(\varphi)$ we needed a large number of variables, but in the linear case, irrational values appear already with very simple constraints over two variables.

Proposition 3. *Let $\varphi_a(x, y)$ be $(y \leq ax) \wedge (x \geq 0)$, where $a \in \mathbb{Q}$. Then $\mu(\varphi_a)$ is irrational unless $a \in \{0, \pm 1\}$.*

Proof sketch. We have $\mu(\varphi_a) = \arctan(a)/2\pi + 1/4$. Assume $\arctan(a) = \frac{m\pi}{n}$ for $m, n \in \mathbb{N}$, i.e., $\mu(\varphi_a) \in \mathbb{Q}$. Using Euler's formula we compute $\left(\cos(\frac{m\pi}{n}) \pm i\sin(\frac{m\pi}{n})\right)^{2n} = e^{\pm 2\pi mi} = 1$. Therefore $\cos(\frac{m\pi}{n}) \pm i\sin(\frac{m\pi}{n})$ are algebraic integers, i.e., roots of a polynomial with integer coefficients and leading coefficient 1. Using closure of algebraic integers under $+$ and $\cdot$ we see that $4\cos^2(\frac{m\pi}{n})$ is an algebraic integer, and by trigonometry it equals $4/(a^2 + 1)$. Since $a \in \mathbb{Q}$ and algebraic integers in $\mathbb{Q}$ are integers [Marcus, 1977], it means $4/(a^2 + 1) \in \mathbb{Z}$, which implies $a = 0$ or $a = \pm 1$. $\qquad \square$

For approximations, the standard notions of approximations are fully-polynomial randomized approximation schemes, FPRAS [Motwani and Raghavan, 1995]. For a problem of computing a function f, it is a randomized algorithm that takes an input I of f and $\varepsilon > 0$, and, in time polynomial in the size of I and $\frac{1}{\varepsilon}$, produces a random variable $A(I, \varepsilon)$ satisfying:

- $\mathbb{P}\big(|A(I, \varepsilon) - f(I)| \leq \varepsilon \cdot f(I)\big) \geq \frac{3}{4}$ (multiplicative error, in which case we refer to an FPRAS for the problem); or

- $\mathbb{P}\big(|A(I, \varepsilon) - f(I)| \leq \varepsilon\big) \geq \frac{3}{4}$ (additive error, in which case we refer to an AFPRAS for the problem).

The constant 3/4 can be replaced by any number in $(1/2, 1)$. From the results presented in Proposition 2, we get, under the widely held complexity-theoretic assumption that RP is different from NP:

Corollary 1. *There is no FPRAS for computing $\mu(\varphi)$ even for order constraints from $\text{FO}(<)$, assuming RP $\neq$ NP.*

In general, FPRAS is viewed as a preferred type of approximation scheme, since there are no a priori bounds on the values $f(I)$. But in our case $\mu(\varphi) \in [0, 1]$, and in such cases AFPRAS provide a good notion of approximation. We now show that such approximations exist for $\text{FO}(<, +)$. Note that every formula over $\text{FO}(<, +, -, 0, 1)$ can be effectively transformed, using Fourier-Motzkin elimination, into a quantifier-free formula; such a transformation is polynomial in the for-

mula and exponential in the number of variables (assuming a fixed quantifier alternation depth) [Weispfenning, 1988].

We thus assume that our linear constraints are *quantifier-free formulae in disjunctive normal form with rational coefficients*. This subclass of $\mathsf{FO}(<,+)$ will be denoted by $\mathsf{DNF}(<,+)$. Its formulae express unions of systems of linear equalities an inequalities. For the input representation, we assume that a formula is represented as the set of its disjuncts. In turn, each system of linear inequalities will be represented as a pair of a matrix and a vector in the usual way. Note that computing $\mu(\varphi)$ remains #P-hard for formulae $\mathsf{DNF}(<,+)$, as follows from Theorem 4. Thus, the assumption does not simplify the task of finding an approximation scheme but instead allows us to focus on the intrinsic source of complexity of the problem. Our main result then states:

Theorem 5. *There is an* AFPRAS *for computing $\mu(\varphi)$ for $\varphi \in \mathsf{DNF}(<,+)$.*

In the rest of the section we explain how our AFPRAS works. We first need to set the computational model for the algorithm. For problems of a geometric nature, it is standard to use the RAM model of computation. In it, all the basic arithmetic operations are performed in constant time. As it is customary, we will also assume that the machine can uniformly at random select a real number in the range $[0,1]$ in constant time (see, e.g., [Bringmann and Friedrich, 2010]).

Intuitively, to compute the likelihood of a formula $\varphi \in \mathsf{DNF}(<,+)$ we need to examine the asymptotic behavior of its disjuncts. A system of linear inequalities in n variables (i.e., an n-dimensional polyhedron) can be represented as the Minkowski sum of two sets: the convex hull $conv(E)$ of a finite set $E \subset \mathbb{R}^n$, and the conic hull $cone(C)$ of a finite set $C \subset \mathbb{R}^n$. This is due to the Minkowski-Weyl Theorem ([Solodovnikov, 1980]). Of these two sets, the only one that matters for the likelihood of $\varphi \in \mathsf{DNF}(<,+)$ is $cone(C)$. This is due to the fact that, at the limit, the finite diameter of $conv(E)$ will be eventually marginalized. Using this observation, we can prove the following statement. In what follows, $\tilde{\varphi}$ denotes the formula obtained from $\varphi \in \mathsf{DNF}(<,+)$ by replacing each disjunct $\varphi_i \equiv (A\bar{x} \leq \bar{b})$ with its homogeneous version $\tilde{\varphi}_i \equiv (A\bar{x} \leq \bar{0})$.

Lemma 1. *For $\varphi \in \mathsf{DNF}(<,+)$ we have $\mu(\varphi) = \mu(\tilde{\varphi})$.*

Proof idea. To obtain $\tilde{\varphi}_i$ from φ_i, we shift each individual hyperplane of φ_i. These shifts preserve the slope of the hyperplanes, and thus the angles defining $cone(C)$. $\square$

In light of Lemma 1, to compute the likelihood of $\varphi \in \mathsf{DNF}(<,+)$ we focus directly on $\tilde{\varphi}$. This observation makes the problem much easier, as the following lemma proves.

Lemma 2. *For an n-ary formula $\varphi \in \mathsf{DNF}(<,+)$, we have $\mu(\varphi) = \lambda(\llbracket \tilde{\varphi} \rrbracket \cap B_1^n)/\lambda(B_1^n)$.*

Proof idea. For an homogeneous system, $cone(C)$ is centered in the origin. The ratio $\lambda(\llbracket \tilde{\varphi} \rrbracket \cap B_r^n)/\lambda(B_r^n)$, then, does not depend on the value of r. $\square$

Using Lemma 2, to estimate the value of $\mu(\varphi)$ we can proceed as follows. Given any $\varepsilon > 0$, our goal is to produce a value $A(\varphi, \varepsilon)$ such that $\mathbb{P}(|A(\varphi, \varepsilon) - \mu(\varphi)| \leq \varepsilon) \geq \frac{3}{4}$. To

---

Algorithm 1 apx-mes

Input: n-ary $\varphi \in \mathsf{DNF}(<,+)$ and $\varepsilon > 0$
Output: $\xi \in \mathbb{R}$ such that $\mathbb{P}(|\xi - \mu(\varphi)| \leq \varepsilon) \geq 3/4$

1: $count := 0; s = \lceil \varepsilon^{-2} \rceil$;
2: **for** i = 1, \ldots, s **do**
3:     sample a point from $\mathcal{B}_1^n$.
4:     **if** $\tilde{\varphi}(\bar{p})$ is true **then** $count := count + 1$;
5: **end for**
6: **return** $\left(\frac{count}{s} \right)$

---

this end, we model $A(\varphi, \varepsilon)$ as a (scaled) binomial random variable A_s as follows. Let $(X_i)_{1 \leq i \leq s}$ be a sequence of independent and identically distributed Bernoulli random variables over the measure space $(B_1^n, \mathcal{F}(B_1^n))$, where $\mathcal{F}(B_1^n)$ is the standard Lebesgue σ-algebra over B_1^n. More precisely, we require that $X_i(\bar{v}) = 1$ if and only if $\varphi(\bar{v})$ is true. Defining $A_s = (\sum_{i=1}^{s} X_i)/s$, we can prove the following.

Lemma 3. *Let $\varphi \in \mathsf{DNF}(<,+)$ and $\varepsilon \in (0,1]$. Then, for $s \geq \varepsilon^{-2}$ the following inequality holds:*

$$\mathbb{P}(|A_s - \mu(\varphi))| \leq \varepsilon) \geq 3/4$$

Using Lemma 3, we can obtain an AFPRAS for $\mu(\varphi)$ as follows. First, observe that under the assumption of the RAM model we can generate uniform samples from the n-dimensional unit ball in time polynomial w.r.t. n (see, e.g., [Blum *et al.*, 2016]). To estimate the value of μ within an additive error ε then, we can sample $s = \lceil \varepsilon^{-2} \rceil$ such points and compute the ratio of these points falling inside $\llbracket \varphi \rrbracket$. Algorithm 1 formalizes this intuition.

Lemma 4. *Algorithm 1 runs in time polynomial in the size of φ and $1/\varepsilon$.*

Lemma 3 and Lemma 4 together prove that Algorithm 1 is an AFPRAS for the computation of the likelihood of formulae in $\mathsf{DNF}(<,+)$, thus concluding the proof of Theorem 5.

7 Conclusions

We provided a framework for estimating likelihoods of numerical constraints defining unbounded sets in $\mathbb{R}^n$, showed when such estimates exist, and explained how to approximate them for linear constraints. There are several extensions of the framework we would like to pursue. One is the extension of the approximation scheme to polynomial constraints. Another direction is related to application of these estimate in different areas, such as program analysis, querying incomplete data, and temporal constraint networks, outlined in the introduction. Finally we would like to consider constraints over $\mathbb{Z}^n$. We expect that our results will help estimate the likelihood of constraints over integers, since for many bodies (e.g., balls), the number of integer lattice points can serve as an approximation of the volume.

Acknowledgments

References

[Arora and Barak, 2009] Sanjeev Arora and Boaz Barak. *Computational Complexity - A Modern Approach*. Cambridge University Press, 2009.

[Belle *et al.*, 2015] Vaishak Belle, Andrea Passerini, and Guy Van den Broeck. Probabilistic inference in hybrid domains by weighted model integration. In *IJCAI*, pages 2770–2776, 2015.

[Benedikt and Libkin, 2002] Michael Benedikt and Leonid Libkin. Aggregate operators in constraint query languages. *J. Comput. Syst. Sci.*, 64(3):628–654, 2002.

[Blum *et al.*, 2016] Avrim Blum, John Hopcroft, and Ravindran Kannan. Foundations of data science. *Vorabversion eines Lehrbuchs*, 2016.

[Bringmann and Friedrich, 2010] Karl Bringmann and Tobias Friedrich. Approximating the volume of unions and intersections of high-dimensional geometric objects. *Comput. Geom.*, 43(6-7):601–610, 2010.

[Caviness and Johnson, 1998] Bob Caviness and Jeremy Johnson. *Quantifier Elimination and Cylindrical Algebraic Decomposition*. Springer, 1998.

[Chavira and Darwiche, 2008] Mark Chavira and Adnan Darwiche. On probabilistic inference by weighted model counting. *Artif. Intell.*, 172(6-7):772–799, 2008.

[Dechter *et al.*, 1991] Rina Dechter, Itay Meiri, and Judea Pearl. Temporal constraint networks. *Artif. Intell.*, 49(1-3):61–95, 1991.

[Gulwani *et al.*, 2008] Sumit Gulwani, Saurabh Srivastava, and Ramarathnam Venkatesan. Program analysis as constraint solving. In *PLDI*, pages 281–292. ACM, 2008.

[Karpinski and Macintyre, 1997] Marek Karpinski and Angus Macintyre. Approximating the volume of general Pfaffian bodies. In *Structures in Logic and Computer Science, A Selection of Essays in Honor of Andrzej Ehrenfeucht*, pages 162–173, 1997.

[Koiran, 1995] Pascal Koiran. Approximating the volume of definable sets. In *FOCS*, pages 134–141, 1995.

[Libkin, 2018] Leonid Libkin. Certain answers meet zero-one laws. In *PODS*, pages 195–207, 2018.

[Ma *et al.*, 2009] Feifei Ma, Sheng Liu, and Jian Zhang. Volume computation for boolean combination of linear arithmetic constraints. In *CADE*, pages 453–468, 2009.

[Macintyre and Wilkie, 1996] Angus Macintyre and Alex Wilkie. On the decidability of the real exponential field. In *Kreiseliana*, pages 441–467. A. K. Peters, 1996.

[Marcus, 1977] Daniel Marcus. *Number Fields*. Springer, 1977.

[Miller, 2011] Chris Miller. Expansions of o-minimal structures on the real field by trajectories of linear vector fields. *Proceedings of the American Mathematical Society*, 139(1):319–330, 2011.

[Motwani and Raghavan, 1995] Rajeev Motwani and Prabhakar Raghavan. *Randomized Algorithms*. Cambridge University Press, 1995.

[Solodovnikov, 1980] Aleksandr Solodovnikov. *Systems of Linear Inequalities*. University of Chicago Press, 1980.

[Spencer, 2010] Joel Spencer. *The Strange Logic of Random Graphs*. Springer, 2010.

[Van den Dries, 1998] Lou Van den Dries. *Tame Topology and o-minimal Structures*, volume 248. Cambridge university press, 1998.

[Weispfenning, 1988] Volker Weispfenning. The complexity of linear problems in fields. *J. Symb. Comput.*, 5(1/2):3–27, 1988.

[Wilkie, 1996] Alex J Wilkie. Model completeness results for expansions of the ordered field of real numbers by restricted Pfaffian functions and the exponential function. *Journal of the American Mathematical Society*, 9(4):1051–1094, 1996.