– Independent Work Report Spring 2019 –

$(\ln^s d, \ln^t \varepsilon^{-1})$-Weak Tractability

Gregory McCord
Advisor: Iasonas Petras

Abstract

Tractability studies the complexity of multivariate problems with respect to the accuracy ε *and the dimension* d *(number of variables) of the problem. Much research has gone into studying different tractability criteria, each of which uniquely defines a class of problems that satisfy the criterion. Two classic, prevalent notions are polynomial tractability and weak tractability. None of these criteria, however, reflect the complexity of problems with respect to both the number of bits in the dimension and the number of bits in the accuracy. Therefore, I propose the criterion* $(\ln^s d, \ln^t \varepsilon^{-1})$*-weak tractability for some constants* $s,t > 0$*. A problem is* $(\ln^s d, \ln^t \varepsilon^{-1})$*-weakly tractable if and only if its complexity is sub-exponential in* $\ln^s d$ *and* $\ln^t \varepsilon^{-1}$*. I analyze the general case for linear multivariate problems and the specific case of* $s = t = 1$ *for linear tensor product problems and provide theorems describing the necessary and sufficient conditions in each case.*

1. Introduction

Over the past decade, many theoretical computer scientists have analyzed tractability criteria, specifically those related to weak tractability. However, they have spent much of their time exploring variations of the tractability criteria categorizing continuous multidimensional problems in terms of their complexity. The complexity of a multidimensional problem typically depends on d and ε^{-1}. The tractability criteria impose restrictions on the scaling of the complexity of multivariate problems with respect to d and ε^{-1}. An interesting direction would be to study how imposing a more strict constraint on the scaling of the cost with respect to the dimension affects which problems survive the criterion. Specifically in this work, I analyze a criterion that studies multivariate problems with respect to the number of bits in the accuracy and the dimension.

Let $n(\varepsilon,d)$ be defined as the information complexity for a d dimensional continuous multivariate problem $S=\{S_d\}$, for $d\geq 1$. I will define S_d later in the cases of linear problems and linear tensor product problems. The goal is to compute an estimate of S within accuracy ε, and the complexity $n(\varepsilon,d)$ represents the minimum number of information operations required to estimate a solution to S_d within the error bound ε.

Before examining other criteria and exploring the motivation for my own, I define my criterion, $(\ln^s d, \ln^t \varepsilon^{-1})$-weak tractability, as follows:

$$\lim_{d+\varepsilon^{-1}\to\infty} \frac{\ln n(\varepsilon,d)}{\ln^s d+\ln^t \varepsilon^{-1}}=0 \tag{1}$$

In general terms, it suffices to say that problems that survive this bound have costs that are not exponential with respect to $\ln^s d$ and $\ln^t \varepsilon^{-1}$, for some $s,t>0$.

2. Background and Related Work

Tractability was first studied under the constraints of polynomial tractability and weak tractability. Polynomial tractability is defined as

$$n(\varepsilon,d)\leq C\cdot d^p\varepsilon^{-q},$$

for some $C,p,q\geq 0$ and has been studied in [5]. This criterion defines the class of problems solvable in time that is polynomial with respect to the dimension and accuracy. However, polynomial tractability is, in general, less relevant to my research and is included only for completeness.

Weak tractability on the other hand is a relevant and popular topic of research. Specifically, weak tractability is defined as

$$\lim_{d+\varepsilon^{-1}\to\infty} \frac{\ln n(\varepsilon,d)}{d+\varepsilon^{-1}}=0$$

and was analyzed in [5]. Weak tractability defines the class of problems that are solvable with cost that is sub-exponential in both ε^{-1} and d. One important point to note is that sub-exponential in d

includes the function $n(\varepsilon,d) = e^{\sqrt{d}}$. More specifically, to satisfy weak tractability, an algorithm must have complexity of the form $n(\varepsilon,d) = e^{o(d+\varepsilon^{-1})}$. You can also have a more general notion of weak tractability called (s,t)-weak tractability, which is defined as

$$\lim_{d+\varepsilon^{-1}\to\infty} \frac{\ln n(\varepsilon,d)}{d^s + \varepsilon^{-t}} = 0$$

for $s,t > 0$. This criterion was proposed and analyzed extensively in [8] and defines algorithms (similar to before) that satisfy $n(\varepsilon,d) = e^{d^{s_0}+\varepsilon^{-t_0}}$ for $s_0 < s, t_0 < t$.

Alternatively, one could consider altering weak tractability in other ways by tweaking the parameters themselves. As mentioned in Section-1, instead of considering the accuracy, one might want to impose a more stringent requirement and require the cost to be sub-exponential with respect to the number of bits in the accuracy $\ln\varepsilon^{-1}$. This is called EC-WT, or exponential convergence-weak tractability, and is defined as

$$\lim_{d+\varepsilon^{-1}\to\infty} \frac{\ln n(\varepsilon,d)}{d + \ln\varepsilon^{-1}} = 0,$$

which was studied extensively under [1–4]. While $\lg\varepsilon^{-1}$ is the true number of bits in the accuracy, since $\lg x$ and $\ln x$ are only a constant factor apart, the result extends from the base-e logarithm to the base-2 logarithm.

Further generalizations brought the research from EC-WT to $(\ln k)$-weak tractability, which was studied in [6], provides the criterion

$$\lim_{d+\varepsilon^{-1}\to\infty} \frac{\ln n(\varepsilon,d)}{d + \ln^k\varepsilon^{-1}} = 0$$

for $k > 0$. Now, we have the final iteration of weak tractability, which combines the principles of (s,t)-weak tractability and EC-WT for $(s,\ln k)$-weak tractability. Defined as

$$\lim_{d+\varepsilon^{-1}\to\infty} \frac{\ln n(\varepsilon,d)}{d^s + \ln^k\varepsilon^{-1}} = 0$$

this final criterion, studied in [7], shows the motivation for $(\ln^s d, \ln^t\varepsilon^{-1})$-weak tractability. In fact, the whole evolution of the tractability constraints demonstrate ways in which researchers have tweaked the criteria, making them more stringent and then generalizing them. These variations inspired my criterion and place it as the next evolution along the line of weak tractability research requiring the cost to be sub-exponential with respect to the number of bits in the dimension.

3. Approach

The direction of this theoretical research has been fairly consistent since the beginning - to derive the theorems that accurately describe the constraints on the problems that will survive my proposed criterion. However, the nature of theoretical research in weak tractability is more streamlined than this somewhat open-ended question. Since the different weak tractability criteria that I mentioned in Section-2 share some similar characteristics, there were different things I could learn from the approach. The analysis progressed somewhat in the following fashion:

1. Understanding the original proof for weak tractability
2. Understanding how the proof changed when substituting $\ln \varepsilon^{-1}$ for ε^{-1} in EC-WT and $(\ln k)$-weak tractability
3. Deriving the specific case $s = t = 1$ for $(\ln^s d, \ln^t \varepsilon^{-1})$-weak tractability
4. Understanding how EC-WT further generalized when adding in exponents s, k for $(s, \ln k)$-weak tractability
5. Deriving the general case $s, t > 0$ for $(\ln^s d, \ln^t \varepsilon^{-1})$-weak tractability

My research began with learning about weak tractability and complexity theory in general before beginning to understand and manipulate the proofs used in more nuanced criteria. By understanding these techniques and how the criterion changed to satisfy a more stringent constraint on the accuracy of the problem, I was able to apply similar techniques in my analysis of a more stringent constraint on the dimension of the problem. This general flow continued until I successfully derived the desired theorems for $(\ln^s d, \ln^t \varepsilon^{-1})$-weak tractability for linear problems. This result can be found in Section-5.1. In the case of linear tensor product problems, I was able to derive the base case - $s = t = 1$. This result can be found in Section-5.2.

4. Setting

4.1. Information Based Complexity

Information Based Complexity (IBC) studies continuous multivariate problems and how they can be solved using partial information. It is concerned with providing estimates to the solutions of these continuous problems using only a subset of the provided d dimensional information. The goal is to generate an estimate of the solution within ε of the true value. The motivation behind this is that computers cannot truly process real valued functions. Rather, they sample functions and solve equations based on these samples - the partial information. The complexity of this analysis refers to the information complexity term - $n(\varepsilon, d)$ - referred to earlier. Additionally, IBC analysis can be done in the worst or average case. The analysis in this paper uses the worst case setting to find restrictions that are necessary and sufficient for the $(\ln^s d, \ln^t \varepsilon^{-1})$-weak tractability criterion to be satisfied. In other words, $n(\varepsilon, d)$ is the minimum amount of information necessary to produce a solution to the d dimensional problem with accuracy ε.

There are two different error definitions that are commonly used in the analysis of weak tractability criteria - absolute (ABS) and normalized (NOR) error criteria. I will consider both in this paper. In theorems and proofs, I will represent the general case (regardless of error criterion used) by using CRI_d. As defined by [5], $CRI_d = 1$ for the absolute error case and the error of the zero algorithm for the normalized error case (this will also be analyzed later). Therefore in this paper, I analyze my weak tractability criterion using the absolute and normalized error criteria in the worst case setting.

4.2. Linear Multivariate Problems in the Worst Case

Linear multivariate problems are d dimensional problems that map continuous functions between Hilbert spaces. These problems have been studied extensively in [5, Ch. 4.2, 4.4, 5.1]. I will now summarize key findings and properties about the operators.

Consider the multivariate problem $S = \{S_d\}$. S_d is defined as $S_d : H_d \to G_d$, where H_d and G_d are both d dimensional Hilbert spaces. Let S_d^* be defined as the problem that complements S_d via the mapping $S_d^* : G_d \to H_d$. Further, S_d^* is defined such that $\langle S_d(f), g\rangle_{G_d} = \langle f, S_d^*(g)\rangle_{H_d}$ for $f \in H_d, g \in G_d$ where $\langle \cdot, \cdot \rangle$ is the dot product operator. Then, let

$$W_d = S_d^* S_d : H_d \to H_d$$

be a compact, self-adjoint operator that is positive semi-definite with eigenvalues $\lambda_{d,1}, \lambda_{d,2}, \ldots$ and corresponding eigenvectors $e_{d,1}, e_{d,2}, \ldots$. Assume that the eigenvalues are sorted in non-decreasing order such that $\lambda_{d,1} \geq \lambda_{d,2} \geq \ldots \geq 0$. One point to note is that continuous linear operators such as W_d have an infinite number of eigenvalues.

Define a linear algorithm to be

$$A_{n,d}(f) = \sum_{i=1}^{n} \langle f, e_{d,i}\rangle S_d(e_{d,i})$$

According to [5, Cor. 4.12], $A_{n,d}(f)$ has minimum worst case error among all algorithms that estimate $S_d(f)$ using information from n continuous functionals, which begins to relate the error of the linear algorithm to the information complexity $n(\varepsilon, d)$. In fact, the corollary extends to say that the minimal error in the worst case is defined by

$$e^{wor}(n) = e^{wor}(A_{n,d}) = \sup_{f \in H_d, ||f||_{H_d} \leq 1} ||S_d(f) - A_{n,d}(f)|| = \sqrt{\lambda_{d,n+1}}.$$

Since the eigenvalues are non-decreasing, we can see that by increasing n, the number of continuous functionals, we decrease the error of the linear algorithm. Therefore, we define the information complexity as the minimum number of functionals to use in order to achieve error that is within our error threshold (i.e. less than ε). That is,

$$n(\varepsilon, d) = \min\left\{n : \sqrt{\lambda_{d,n+1}} \leq \varepsilon \cdot \sqrt{CRI_d}\right\} \tag{2}$$

or equivalently,

$$n(\varepsilon, d) = \left| \left\{ \lambda_{d,i} : \sqrt{\lambda_{d,i}} > \varepsilon \cdot \sqrt{CRI_d} \right\} \right| \quad (3)$$

If there were only a finite, constant number of eigenvalues, then computing sufficient eigenvalues would be achievable in constant time with respect to the size of the problem. As before, $CRI_d = 1$ for ABS while CRI_d under NOR is the error of the zero algorithm. Since, if $n = 0 \implies A_{0,d}(f) = 0$, we have that the error is just the largest eigenvalue, $\lambda_{d,1}$. That is, under NOR, $CRI_d = \sup_{f \in H_d, ||f||_{H_d} \leq 1} ||S_d(f)|| = \sqrt{\lambda_{d,1}}$. Therefore, for $CRI_d > 1$, estimating S_d within error ε is at least as easy as the absolute error case. The opposite holds for $CRI_d < 1$.

4.3. Linear Tensor Product Problems in the Worst Case

Now, I will discuss a specific case of linear problems - linear tensor product problems. These are discussed extensively in [5, Ch. 5.2], but I will summarize the results and properties here. The d dimensional tensor product problem is a d-fold tensor product of a one-dimensional (univariate) linear problem. A tensor product is a generalization of the outer product that ultimately combines lower dimensional Hilbert Spaces into higher dimensional ones. Specifically, consider the linear problem in one-dimension $S_1 : H_1 \to G_1$ between Hilbert Spaces.

Consider S_1^* that is defined similarly as before. This leads to the compact, self-adjoint, positive semi-definite operator

$$W_1 = S_1^* S_1 : H_1 \to H_1.$$

Let the eigenvalues of this one-dimensional operator be defined as follows (let $d = 1$ be implied for simplicity of notation):

$$\lambda_1 \geq \lambda_2 \geq \ldots \geq 0$$

We define $H_d = \bigotimes_{i=1}^{d} H_1$ and $G_d = \bigotimes_{i=1}^{d} G_1$. Then, let

$$S_d = \bigotimes_{i=1}^{d} S_1 : H_d \to G_d$$

be the d dimensional linear tensor product problem of interest. Finally, we have $W_d = S_d^* S_d : H_d \to H_d$ defined as before. For general tensor products, we have that the eigenvalues of the resulting operator are the products of all possible combinations of eigenvalues of the one-dimensional problem. In other words, the eigenvalues and corresponding eigenvectors of the d dimensional problem are defined as follows:

$$\lambda_{d,j} = \prod_{i=1}^{d} \lambda_{j_i} \quad , \quad e_{d,j} = \bigotimes_{i=1}^{d} e_{j_i}$$

for all $j = [j_1, j_2, ..., j_d] \in \mathbb{N}^d$. I will make a slight addendum to the notation. We have that j is a vector corresponding to the indices of the eigenvalues of the one-dimensional problem that combine to form the d dimensional problem. Allow $\lambda_{d,j}$ to represent a specific eigenvalue of W_d made up of the j eigenvalues of W_1. Let $\lambda_{d,i}, \forall i \in \mathbb{N}$ be the i^{th} largest eigenvalue of W_d. Similarly, let $e_{d,i}, \forall i \in \mathbb{N}$ be the eigenvector corresponding to the i^{th} eigenvalue. Note that the largest eigenvalue of W_d is a product of the largest eigenvalue of W_1 d times. That is, $\lambda_{d,1} = \prod_{i=1}^{d} \lambda_1 = \lambda_1^d$.

Similar to Section-4.2, let the linear algorithm be defined as

$$A_{n,d}(f) = \sum_{i=1}^{n} \langle f, e_{d,i} \rangle S_d(e_{d,i}).$$

For the same principles as before, we have that $A_{n,d}(f)$ is optimal and therefore has minimal error for all estimates of $S_d(f)$ using information from up to n continuous linear functionals. The error in this case therefore follows the same equation as before and is defined by

$$e^{wor}(A_{n,d}) = \sup_{f \in H_d, ||f||_{H_d} \leq 1} ||S_d(f) - A_{n,d}(f)|| = \sqrt{\lambda_{d,n+1}}.$$

Therefore, the information complexity is defined as before as

$$n(\varepsilon, d) = \left| \left\{ \lambda_{d,i} : \sqrt{\lambda_{d,i}} > \varepsilon \cdot \sqrt{CRI_d} \right\} \right|.$$

However, due to the nature of $\lambda_{d,j}$, we can actually express this term in a more specific way for linear tensor product problems. That is,

$$n(\varepsilon,d) = \left|\left\{ j = [j_1, ..., j_d] \in \mathbb{N}^d : \sqrt{\lambda_{j_1} \cdots \lambda_{j_d}} > \varepsilon \cdot \sqrt{CRI_d} \right\}\right| \tag{4}$$

While seemingly more complicated, this is just the number of eigenvalues of W_d that are greater than the error (times the error criterion used). As before, $CRI_d = 1$ for ABS and the error of the zero algorithm for NOR. Since we derived the largest eigenvalue of W_d earlier, we now have that, under NOR,

$$CRI_d = \sup_{f \in H_d, ||f||_{H_d} \leq 1} ||S_d(f)|| = \sqrt{\lambda_{d,1}} = \lambda_1^{d/2}.$$

5. Results

5.1. Tractability of Linear Multivariate Problems

I study the case of general $s,t > 0$ for linear multivariate problems under the criterion $(\ln^s d, \ln^t \varepsilon^{-1})$-weak tractability in the worst case setting. Let a linear multivariate problem be as defined in Section-4.2. Note that I only consider d dimensional problems for $d \geq 2$. More research is necessary in order to understand the specific case of $d = 1$ under this criterion.

Theorem 1. *Consider the non-zero linear multivariate problem $S = \{S_d\}$ for compact linear S_d defined over Hilbert spaces with $d \geq 2$. We study the problem S for the absolute or normalized error criterion in the worst case setting for the class of all linear functionals Λ^{all}. The problem is $(\ln^s d, \ln^t \varepsilon^{-1})$-weakly tractable, for $s,t > 0$ iff*

- *we have*

$$\lim_{j \to \infty} \frac{1}{\ln \frac{CRI_d}{\lambda_{d,j}}} (\ln j)^{1/t} = 0, \quad \text{and} \tag{5}$$

- *there exists a function* $f:(0,1/2]\to\mathbb{N}$, *such that*

$$M := \sup_{\beta\in(0,1/2]} \frac{1}{\beta^{1/t}} \sup_{\ln^s d\geq f(\beta)} \sup_{j\geq\lceil\exp(\sqrt{\beta}\ln^s d)\rceil+1} \frac{1}{\ln\frac{CRI_d}{\lambda_{d,j}}}(\ln j)^{1/t} < \infty \tag{6}$$

Proof of Theorem 1. I proceed similarly to the proof of [5, Thm. 5.3] and the proof of [6, Thm. 3] and adapt for s as in the proof of [7, Thm. 1]. First I will show that $(\ln^s d, \ln^t \varepsilon^{-1})$-weak tractability is necessary for, and therefore implies, equations (5) and (6) before proving that these equations are sufficient for $(\ln^s d, \ln^t \varepsilon^{-1})$-weak tractability. Without loss of generality, we may assume that $\varepsilon\in(0,1)$.

Let S be $(\ln^s d, \ln^t \varepsilon^{-1})$-weakly tractable. Then for every $\beta\in(0,1/2]$, $\exists N_\beta>0$ such that for all ε, d with $\ln^s d+\ln^t\varepsilon^{-1}\geq N_\beta$ (where also $\varepsilon\in(0,1)$ and $d\geq 2$), we have $\frac{\ln n(\varepsilon,d)}{\ln^s d+\ln^t\varepsilon^{-1}}\leq\beta$, which implies $n(\varepsilon,d)\leq\exp(\beta(\ln^s d+\ln^t\varepsilon^{-1}))$. We are interested in the eigenvalues for which $\lambda_{d,j+1}\leq\varepsilon^2\cdot CRI_d$. From the definition of $n(\varepsilon,d)$, we have

$$\frac{\lambda_{d,j}}{CRI_d}\leq\varepsilon^2, \text{ for } j=\lfloor\exp(\beta(\ln^s d+\ln^t\varepsilon^{-1}))\rfloor+1$$

Therefore, by extension of the definition of j, we have

$$\begin{aligned}
j &= \lfloor\exp(\beta(\ln^s d+\ln^t\varepsilon^{-1}))\rfloor+1\\
j-1 &\leq \exp(\beta(\ln^s d+\ln^t\varepsilon^{-1}))\\
\ln(j-1) &\leq \beta(\ln^s d+\ln^t\varepsilon^{-1})\\
\frac{\ln(j-1)}{\beta}-\ln^s d &\leq \ln^t\varepsilon^{-1}\\
\varepsilon^{-2} &\leq \exp\left(2\left(\frac{\ln(j-1)}{\beta}-\ln^s d\right)_+^{1/t}\right)\\
\varepsilon^2 &\leq \left(\frac{1}{\exp\left(2\left(\frac{\ln(j-1)}{\beta}-\ln^s d\right)_+^{1/t}\right)}\right)
\end{aligned}$$

Note that $_+$ is used since we only observe $\varepsilon \in (0,1)$, so its use appropriately limits the upper bound of ε. Because $\frac{\lambda_{d,j}}{CRI_d} \le \varepsilon^2$, we have

$$\frac{\lambda_{d,j}}{CRI_d} \le \left(\frac{1}{\exp\left(2\left(\frac{\ln(j-1)}{\beta} - \ln^s d\right)_+^{1/t}\right)} \right)$$

$$\ln\left(\frac{CRI_d}{\lambda_{d,j}}\right) \ge 2 \cdot \left(\frac{\ln(j-1)}{\beta} - \ln^s d\right)_+^{1/t}$$

$$\frac{1}{\ln\left(\frac{CRI_d}{\lambda_{d,j}}\right)} \le \frac{1}{2} \cdot \left(\frac{\beta}{\ln(j-1) - \beta \ln^s d}\right)_+^{1/t}$$

for $j = \lfloor \exp(\beta(\ln^s d + \ln^t \varepsilon^{-1})) \rfloor + 1$. For $\ln^s d + \ln^t \varepsilon^{-1} \ge N_\beta > 0$, we also get $j = \lfloor e^{\beta N_\beta} \rfloor + 1, \lfloor e^{\beta N_\beta} \rfloor + 2, \ldots$ by varying $\varepsilon \in (0,1)$.

Here we see why $d = 1$ does not satisfy the equation. For $d = 1 \implies \ln d = 0$. For small values of $\beta \ln^t \varepsilon^{-1}$, we have $j = 2$. However, for $j = 2, d = 1$, the expression $\ln(j-1) - \beta \ln^s d = 0$, resulting in an undefined upper bound for $1/\ln\left(\frac{CRI_d}{\lambda_{d,j}}\right)$.

For the first part of the theorem, let d be fixed and j be sufficiently large. In that case, $\frac{1}{\ln\left(\frac{CRI_d}{\lambda_{d,j}}\right)}$ is on the order of $\left(\frac{\beta}{\ln j}\right)^{1/t}$. For arbitrarily small β we have

$$\frac{1}{\ln\left(\frac{CRI_d}{\lambda_{d,j}}\right)} \le \frac{1}{2} \cdot \left(\frac{\beta}{\ln j}\right)^{1/t} \le \left(\frac{\beta}{\ln j}\right)^{1/t}$$

$$\frac{1}{\ln \frac{CRI_d}{\lambda_{d,j}}} (\ln j)^{1/t} \le \beta^{1/t}$$

$$\lim_{j \to \infty} \frac{1}{\ln \frac{CRI_d}{\lambda_{d,j}}} (\ln j)^{1/t} = 0, \quad \forall d \ge 2$$

which extends from the ε, δ definition of a limit. Therefore, we have demonstrated (5).

For the second part of the theorem, take a function $f : (0, 1/2] \to \mathbb{N}$, such that $f(\beta) = N_\beta$. For $\ln^s d \ge f(\beta)$ and $j \ge \lceil e^{\beta \ln^s d} \rceil + 1 \ge \lceil e^{\sqrt{\beta} \ln^s d} \rceil + 1 \ge 3$ (since $\beta \in (0, 1/2], \ln d > 0$ for $d \ge 2$ and

$\sqrt{\beta} \geq \beta$) we have

$$j = \lfloor \exp(\beta(\ln^s d + \ln^t \varepsilon^{-1})) \rfloor + 1 \implies j \geq \lfloor \exp(\beta N_\beta) \rfloor + 1$$

which follows from $\ln d \geq f(\beta) = N_\beta$. Additionally, from above, we have

$$\begin{aligned} j &\geq \lceil e^{\sqrt{\beta} \ln^s d} \rceil + 1 \\ j - 1 &\geq \lceil e^{\sqrt{\beta} \ln^s d} \rceil \\ j - 1 &\geq e^{\sqrt{\beta} \ln^s d} \\ \ln(j-1) &\geq \sqrt{\beta} \ln^s d \end{aligned}$$

This leads to

$$\begin{aligned} \ln(j-1) &\geq \sqrt{\beta} \ln^s d \\ \sqrt{\beta} \ln(j-1) &\geq \beta \ln^s d \\ -\sqrt{\beta} \ln(j-1) &\leq -\beta \ln^s d \\ \ln(j-1) - \beta \ln^s d &\geq (1 - \sqrt{\beta}) \ln(j-1) > 0, \quad \text{for } j \geq 3, \beta \in (0, 1/2] \end{aligned} \tag{7}$$

Therefore, by using our result from before and some algebraic manipulation, we have

$$\frac{1}{\ln\left(\frac{CRI_d}{\lambda_{d,j}}\right)} \leq \frac{1}{2}\cdot\left(\frac{\beta}{\ln(j-1)-\beta\ln^s d}\right)_+^{1/t}$$

$$\frac{1}{\beta^{1/t}}\cdot\frac{1}{\ln\left(\frac{CRI_d}{\lambda_{d,j}}\right)}\cdot(\ln j)^{1/t} \leq \frac{1}{2}\cdot\left(\frac{\ln j}{\ln(j-1)-\beta\ln^s d}\right)_+^{1/t}$$

$$\leq \left(\frac{\ln j}{(1-\sqrt{\beta})\ln(j-1)}\right)_+^{1/t} \qquad \text{(from (7))}$$

$$= \left(\frac{\ln j}{\ln(j-1)}\cdot\frac{1}{1-\sqrt{\beta}}\right)_+^{1/t}$$

$$\leq \left(\frac{\ln 3}{\ln 2}\cdot(2+\sqrt{2})\right)^{1/t} < \infty \qquad \text{(inputs that maximize value)}$$

which demonstrates that the maximum possible value of the equation is finite based on the supremums of β and $j \geq \lceil e^{\sqrt{\beta}\ln^s d}\rceil + 1$ as well as the constraint $\ln^s d \geq f(\beta) = N_\beta > 0$. Note that the above would be undefined if $\ln d = 0 \implies j = 2$. This leads to the result in (6). That is,

$$M := \sup_{\beta\in(0,1/2]} \frac{1}{\beta^{1/t}} \sup_{\ln^s d \geq f(\beta)} \sup_{j \geq \lceil \exp(\sqrt{\beta}\ln^s d)\rceil + 1} \frac{1}{\ln\frac{CRI_d}{\lambda_{d,j}}}(\ln j)^{1/t} < \infty$$

Now we assume that the conditions (5), (6) hold. We demonstrate that they are sufficient for $(\ln^s d, \ln^t \varepsilon^{-1})$-weak tractability.

From condition (5), we have that for any $\beta \in (0,1/2]$, $\exists C_\beta \in \mathbb{Z}_{>2}$ such that $\forall j \geq C_\beta$ and $\ln^s d < f(\beta) \implies d = 2,3,...,\lfloor \exp\left(f(\beta)-1\right)^{1/s}\rfloor$ (since (5) holds for $d \geq 2$) the following holds

$$\frac{1}{\ln\frac{CRI_d}{\lambda_{d,j}}} \leq \frac{\beta}{(\ln j)^{1/t}}$$

$$\frac{\lambda_{d,j}}{CRI_d} \leq \exp^{-1}\left(\frac{(\ln j)^{1/t}}{\beta}\right)$$

As a result, we have $\lambda_{d,j} \leq \varepsilon^2 \cdot CRI_d$ for $j = \lceil e^{(2\beta\ln\varepsilon^{-1})^t}\rceil$. We can recover the previous equation through algebraic manipulation

$$
\begin{aligned}
j = \lceil e^{(2\beta \ln \varepsilon^{-1})^t} \rceil &\geq e^{(2\beta \ln \varepsilon^{-1})^t} \\
(\ln j)^{1/t} &\geq 2\beta \ln \varepsilon^{-1} \\
\frac{(\ln j)^{1/t}}{\beta} &\geq \ln \varepsilon^{-2} \\
\varepsilon^2 &\geq \exp^{-1}\left(\frac{(\ln j)^{1/t}}{\beta}\right) \\
\frac{\lambda_{d,j}}{CRI_d} &\leq \exp^{-1}\left(\frac{(\ln j)^{1/t}}{\beta}\right) \leq \varepsilon^2
\end{aligned}
$$

Using the inequality $\lceil x \rceil \leq 2x$ for $x \geq 1$, we have $\lambda_{d,j} \leq \varepsilon^2 \cdot CRI_d$ for $j = \lceil e^{(2\beta \ln \varepsilon^{-1})^t} \rceil \leq 2e^{(2\beta \ln \varepsilon^{-1})^t}$. Therefore, we get

$$
n(\varepsilon, d) \leq \max\left(C_\beta, 2e^{(2\beta \ln \varepsilon^{-1})^t}\right) \tag{8}
$$

We now consider the case where $\ln^s d \geq f(\beta)$. Condition (6) implies that for every $\beta \in (0, 1/2]$ and $j \geq \lceil \exp(\sqrt{\beta} \ln^s d) \rceil + 1$ we have

$$
\begin{aligned}
\frac{1}{\ln \frac{CRI_d}{\lambda_{d,j}}} &\leq M\left(\frac{\beta}{\ln j}\right)^{1/t} < \infty \\
\frac{\lambda_{d,j}}{CRI_d} &\leq \exp^{-1}\left(M^{-1}\left(\frac{\ln j}{\beta}\right)^{1/t}\right)
\end{aligned}
$$

As a result, we have $\lambda_{d,j} \leq \varepsilon^2 \cdot CRI_d$ for $j = \lceil e^{\beta(2M \ln \varepsilon^{-1})^t} \rceil$. Therefore we have $j = \max(\lceil \exp(\sqrt{\beta} \ln^s d) \rceil + 1, \lceil e^{\beta(2M \ln \varepsilon^{-1})^t} \rceil)$. From the inequality $\lceil x \rceil \leq 2x$ for $x \geq 1$, we get that the information complexity is

$$
n(\varepsilon, d) \leq \max(2e^{\sqrt{\beta} \ln^s d} + 1, 2e^{\beta(2M \ln \varepsilon^{-1})^t}) \tag{9}
$$

From (8), (9) we get

$$\frac{\ln n(\varepsilon,d)}{\ln^s d+\ln^t \varepsilon^{-1}} \le \max\left(\frac{\ln C_\beta}{\ln^s d+\ln^t \varepsilon^{-1}}, \frac{(2\beta\ln\varepsilon^{-1})^t}{\ln^s d+\ln^t \varepsilon^{-1}}+\frac{\ln 2}{\ln^s d+\ln^t \varepsilon^{-1}}, \right.$$
$$\left. \frac{\beta(2M\ln\varepsilon^{-1})^t}{\ln^s d+\ln^t \varepsilon^{-1}}+\frac{\ln 2}{\ln^s d+\ln^t \varepsilon^{-1}}, \frac{\sqrt{\beta}\ln^s d}{\ln^s d+\ln^t \varepsilon^{-1}}+\frac{\ln 2}{\ln^s d+\ln^t \varepsilon^{-1}}\right)$$

For sufficiently large $\ln^s d+\ln^t \varepsilon^{-1}$, the terms involving the constants (C_β and $\ln 2$) in the numerator run to 0. Additionally, the maximum value for $\frac{\ln^s d}{\ln^s d+\ln^t \varepsilon^{-1}}$ and $\frac{\ln^t \varepsilon^{-1}}{\ln^s d+\ln^t \varepsilon^{-1}}$ is ≤ 1. Therefore we get

$$\frac{\ln n(\varepsilon,d)}{\ln^s d+\ln^t \varepsilon^{-1}} \le \max\left((2\beta)^t, \beta(2M)^t, \sqrt{\beta}\right)$$

which implies

$$\lim_{d+\varepsilon^{-1}\to\infty} \frac{\ln n(\varepsilon,d)}{\ln^s d+\ln^t \varepsilon^{-1}} = 0$$

via the ε,δ definition of a limit since β can be arbitrarily small which in turn implies that S is $(\ln^s d, \ln^t \varepsilon^{-1})$-weakly tractable. □

5.2. Tractability of Linear Tensor Product Problems

Moving on from the tractability of linear problems, I study the tractability of linear tensor product problems for the absolute error criterion for the class of all linear functionals Λ^{all} in the worst case. Let a linear tensor product problem be as defined in Section-4.3. Consider the case of $\lambda_2 = 0$. We have a trivial case since, for any $\lambda_1 \ge 0$, $\lambda_{d,1} = \lambda_1^d$ and $\lambda_{d,i} = 0$ for $i > 1 \implies n(\varepsilon,d) \le 1$. Additionally, consider the cases $\lambda_1 = \lambda_2 = 1$ and $\lambda_1 > 1, \lambda_2 > 0$. In fact, according to Theorem 5.5 in [5], both of these cases are intractable for less stringent definitions of weak tractability. Therefore, for $(\ln^s d, \ln^t \varepsilon^{-1})$-weak tractability, they are also intractable. We now consider the case of $\lambda_1 \le 1$ and $\lambda_2 \in (0,1)$.

Theorem 2. *Consider the linear tensor product problem $S = \{S_d\}$ in the worst case setting. For the* *absolute error criterion and for the class of all linear functionals Λ^{all}, where $\lambda_1 \leq 1$ and $\lambda_2 \in (0,1)$ we have*

- *S is not $(\ln^1 d, \ln^1 \varepsilon^{-1})$-weakly tractable.*

This theorem states that, for $s = t = 1$, there are no linear tensor product problems such that $(\ln^s d, \ln^t \varepsilon^{-1})$-weak tractability holds when the first two eigenvalues of the one-dimensional problem S_1 satisfy $\lambda_1 \leq 1$ and $\lambda_2 \in (0,1)$. At this time, it is unknown what the precise bounds are for general $s > 1$ or $t > 1$ (or both) under this criterion.

Proof of Theorem 2. I proceed similarly to the proof of [6, Thm. 4]. First, we will consider the case of $\lambda_1 = 1$ and $\lambda_2 \in (0,1)$. Take $j \in \mathbb{N}^d$ such that

$$\lambda_{d,j} = \prod_{i=1}^{d} \lambda_{j_i} > \varepsilon^2$$

Let k be the number of indices j_i that are greater than 1. We have

$$\lambda_2^k = \lambda_1^{d-k} \lambda_2^k \geq \lambda_{d,j} > \varepsilon^2$$

since $\lambda_1 = 1$ by assumption for the first case and by definition of $n(\varepsilon, d)$ for linear tensor product problems. Since the eigenvalues of the one-dimensional problem are sorted in non-increasing order, we know that $\lambda_i \geq \lambda_j, i < j$. Therefore,

$$\lambda_{d,j} = \lambda_{j_i}^{d-k} \prod_{i:j_i>1} \lambda_{j_i} = \prod_{i:j_i>1} \lambda_{j_i}$$

which is upper bounded by λ_2^k. There are $\binom{d}{k}$ ways to select the d values of j_i (where $j_i \in \{1,2\}, \forall i \in [d]$). Therefore, since we have at least $\binom{d}{k}$ eigenvalues, all of which are greater than ε^2, it extends that

$$n(\varepsilon, d) \geq \binom{d}{k}$$

Consider error ε_d such that

$$\lambda_2^{(\lfloor d/2 \rfloor+1)/2} < \varepsilon_d < \lambda_2^{\lfloor d/2 \rfloor/2}$$

This is equivalent to $\varepsilon_d^2 < \lambda_2^{\lfloor d/2 \rfloor}$, which by parallel implies $k = \lfloor d/2 \rfloor$. Therefore,

$$n(\varepsilon_d, d) \geq \binom{d}{\lfloor d/2 \rfloor} \geq 2^{\lfloor d/2 \rfloor}$$

where the final inequality comes from

$$\begin{aligned}
\binom{d}{\lfloor d/2 \rfloor} &= \frac{d!}{(d-\lfloor d/2 \rfloor)!\lfloor d/2 \rfloor!} \\
&\approx \frac{\sqrt{2\pi d}\left(\frac{d}{e}\right)^d}{\sqrt{2\pi(d-\lfloor d/2 \rfloor)}\left(\frac{(d-\lfloor d/2 \rfloor)}{e}\right)^{d-\lfloor d/2 \rfloor}\sqrt{2\pi(\lfloor d/2 \rfloor)}\left(\frac{(\lfloor d/2 \rfloor)}{e}\right)^{\lfloor d/2 \rfloor}}
\end{aligned}$$

via Stirling's Approximation. For large d, this becomes

$$\begin{aligned}
\binom{d}{\lfloor d/2 \rfloor} &\approx \frac{\sqrt{2\pi d}\left(\frac{d}{e}\right)^d}{\left(\sqrt{\pi d}\left(\frac{d}{2e}\right)^{d/2}\right)^2} && \text{(large } d\text{)} \\
&= \frac{\sqrt{2}\left(\frac{d}{e}\right)^d}{\sqrt{\pi d}\left(\frac{d}{2e}\right)^d} \\
&= \sqrt{\frac{2}{\pi d}} \cdot 2^d \\
&\approx 2^d && \text{(large } d\text{)} \\
&\geq 2^{\lfloor d/2 \rfloor}
\end{aligned}$$

Now we can see that in this case $(\ln^1 d, \ln^1 \varepsilon^{-1})$-weak tractability does not hold. Let $\varepsilon = \varepsilon_d$ (ε is fixed) with $d \to \infty$. Note the following two inequalities: $\lfloor x \rfloor > x - 1, \forall x$ and $\lfloor x/2 \rfloor + 1 \leq x, x \geq 1$.

Then we get

$$\begin{aligned}
\lim_{\substack{d\to\infty,\\ \lambda_2^{\frac{\lfloor d/2\rfloor+1}{2}}<\varepsilon_d<\lambda_2^{\frac{\lfloor d/2\rfloor}{2}}}} \frac{\ln n(\varepsilon,d)}{\ln d+\ln\varepsilon^{-1}} &> \lim_{d\to\infty}\frac{\ln n(\varepsilon,d)}{\ln d+\frac{\lfloor d/2\rfloor+1}{2}\ln\lambda_2^{-1}} \\
&\geq \lim_{d\to\infty}\frac{\lfloor d/2\rfloor\ln 2}{\ln d+\frac{\lfloor d/2\rfloor+1}{2}\ln\lambda_2^{-1}} \\
&\geq \lim_{d\to\infty}\frac{\left(\frac{d}{2}-1\right)\ln 2}{\ln d+\frac{d}{2}\ln\lambda_2^{-1}} \\
&= \lim_{d\to\infty}\frac{\left(\frac{1}{2}-\frac{1}{d}\right)\ln 2}{\frac{\ln d}{d}+\frac{1}{2}\ln\lambda_2^{-1}} \\
&= \frac{\ln 2}{\ln\lambda_2^{-1}} > 0
\end{aligned}$$

which implies that S is not $(\ln^1 d, \ln^1 \varepsilon^{-1})$-weakly tractable. Next, we consider the case $\lambda_1 < 1$ and $\lambda_2 > 0$. Some results extend from the previous part of the proof. Namely, for $k = \lfloor d/2\rfloor$ we have

$$n(\varepsilon_d,d) \geq \binom{d}{\lfloor d/2\rfloor} \geq 2^{\lfloor d/2\rfloor}$$

Similar to above we let k be the number of indices $j_i > 1$. We therefore have

$$\lambda_1^{d-k}\lambda_2^k \geq \lambda_{d,j} > \varepsilon^2$$

Since $\lambda_1 < 1$, we will now choose error ε_d such that

$$\lambda_1^{(d-\lfloor d/2\rfloor-1)/2}\lambda_2^{(\lfloor d/2\rfloor+1)/2} < \varepsilon_d < \lambda_1^{(d-\lfloor d/2\rfloor)/2}\lambda_2^{\lfloor d/2\rfloor/2}$$

Now we will follow a similar derivation as before. Let $\varepsilon = \varepsilon_d$ with $d \to \infty$. Then we get

$$\lim_{\substack{d\to\infty,\\ \lambda_1^{\frac{d-\lfloor d/2\rfloor-1}{2}}\lambda_2^{\frac{\lfloor d/2\rfloor+1}{2}}<\varepsilon_d<\lambda_1^{\frac{d-\lfloor d/2\rfloor}{2}}\lambda_2^{\frac{\lfloor d/2\rfloor}{2}}}} \frac{\ln n(\varepsilon,d)}{\ln d+\ln\varepsilon^{-1}} > \lim_{d\to\infty}\frac{\lfloor d/2\rfloor \ln 2}{\ln d+\frac{d-\lfloor d/2\rfloor-1}{2}\ln\lambda_1^{-1}+\frac{\lfloor d/2\rfloor+1}{2}\ln\lambda_2^{-1}}$$

$$\geq \lim_{d\to\infty}\frac{\left(\frac{d}{2}-1\right)\ln 2}{\ln d+\frac{d-\lfloor d/2\rfloor-1}{2}\ln\lambda_1^{-1}+\frac{d}{2}\ln\lambda_2^{-1}}$$

$$\geq \lim_{d\to\infty}\frac{\left(\frac{d}{2}-1\right)\ln 2}{\ln d+\frac{d}{2}\ln\lambda_1^{-1}+\frac{d}{2}\ln\lambda_2^{-1}}$$

$$= \lim_{d\to\infty}\frac{\left(\frac{1}{2}-\frac{1}{d}\right)\ln 2}{\frac{\ln d}{d}+\frac{1}{2}\ln(\lambda_1^{-1}\lambda_2^{-1})}$$

$$= \frac{\ln 2}{\ln(\lambda_1\lambda_2)^{-1}} > 0$$

which implies that S is not $(\ln^1 d, \ln^1 \varepsilon^{-1})$-weakly tractable. $\square$

6. Future Work

There are several cases that I've mentioned that will require additional work. Specifically, there remains the case of $d = 1$ for linear multivariate problems. As mentioned before, when you take $\ln d$ for $d = 1$, you get 0. Normally this isn't a problem, however there are some cases where $\ln\varepsilon^{-1}$ is small such that division might result in an undefined result. More analysis is necessary to see what case these types of problems degenerate into. That is, for $d = 1$, what form does the information complexity take and how does this react with my criterion.

Additionally, the analysis of linear tensor product problems in the general case for $s > 1$ or $t > 1$ (or both) remains open. In particular, it would be interesting to find necessary and sufficient conditions on the rate of decay of the eigenvalues of the linear tensor product problems that satisfy the criterion.

Finally, in this paper, I have done analysis using the absolute and normalized error criteria in the worst case setting for linear and linear tensor product problems. However, information based complexity analysis can also be used to study the average case of these problems under different criteria. Future research could consider this criterion in the average case setting or introduce other error criteria (apart from ABS and NOR). Further, researchers could also apply this criterion to a wider group of problems - not just linear and linear tensor product problems.

7. Conclusion

Over the years, there has been much research into different weak tractability criteria. From the original criterion to EC-WT to $(\ln^s d, \ln^t \varepsilon^{-1})$-weak tractability, each new evolution altered the parameters slightly to either generalize or impose more stringent constraints on a previous iteration of the criterion. With each newly proposed criterion came the goal of understanding how these changes affected the class of problems that survived the bound. Finding a class of problems that survive a tractability criterion is equivalent to stating necessary and sufficient conditions for the problems to satisfy the criterion. Theorem 1 details the bounds for the general case of linear multivariate problems while theorem 2 states that the criterion cannot be satisfied for the case of $s = t = 1$ for linear tensor product problems. More analysis is necessary to derive the conditions for the general case.

8. Acknowledgements

I'd like to thank Dr. Petras for his help and advice during meetings. From helping me choose a new topic in a new field after our initial one fell through to teaching me the basics of complexity theory and tractability to checking over my proofs to giving me recommendations and notes on my paper and methodologies, he has been there throughout the process.

I'd also like to thank my brother, Chris, for always being there to talk through things or offer advice.

Lastly I'd like to thank my parents, Mike and Kathy, for supporting and believing in me. They have always been there for me.

9. Ethics

I hereby declare that I am the sole author of this paper.

Honor Code: I pledge my honor that this paper represents my own work in accordance with university regulations.

-Gregory M. McCord

References

[1] J. Dick *et al.*, "Approximation of analytic functions in korobov spaces," *Journal of Complexity*, vol. 30, no. 2, pp. 2–28, 2014.

[2] J. Dick *et al.*, "Exponential convergence and tractability of multivariate integration for korobov spaces," *Mathematics of Computation*, vol. 80, no. 274, pp. 905–930, 2011.

[3] C. Irrgeher *et al.*, "Integration in hermite spaces of analytic functions," *Journal of Complexity*, vol. 31, no. 3, pp. 380–404, 2015.

[4] P. Kritzer, F. Pillichshammer, and H. Woźniakowski, "Multivariate integration of infinitely many times differentiable functions in weighted korobov spaces," *Mathematics of Computation*, vol. 83, no. 287, pp. 1189–1206, 2014.

[5] E. Novak and H. Wozniakowski, "Tractability of multivariate problems, volume i: Linear information, european math," *Soc., Zürich*, vol. 3, 2008.

[6] A. Papageorgiou and I. Petras, "A new criterion for tractability of multivariate problems," *Journal of Complexity*, vol. 30, no. 5, pp. 604–619, 2014.

[7] A. Papageorgiou, I. Petras, and H. Woźniakowski, "(s, $\ln\kappa$)-weak tractability of linear problems," *Journal of Complexity*, vol. 40, pp. 1–16, 2017.

[8] P. Siedlecki and M. Weimar, "Notes on (s, t)-weak tractability: A refined classification of problems with (sub) exponential information complexity," *Journal of Approximation Theory*, vol. 200, pp. 227–258, 2015.