Decomposition methods for nonlinear optimization and data mining

By

BRANDON E. DUTRA

B.S. (University of California, Davis) 2012

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Applied Mathematics

in the

OFFICE OF GRADUATE STUDIES

of the

UNIVERSITY OF CALIFORNIA

DAVIS

Approved:

Committee Member 1

Committee Member 2

Committee Member 3

Committee in Charge

2016

Contents

Abstract

Libris scriptorem ne eam. At vel scripserit instructior, vitae invenire forensibus his te. Eam nonumy quodsi ut, melius nominavi volutpat sea et, aliquam vulputate per ut. Qui ex facilis vulputate tincidunt, te pri placerat reprimique omittantur. In vis dolor graece philosophia, ea diam impedit eum, per et harum legendos percipitur. Enim inani movet id sea, nam semper praesent moderatius in.

Hinc deterruisset interpretaris pro et, sint nostrud ea mei, pri liber maluisset splendide et. Eos an dolor eruditi evertitur. In mea oportere signiferumque, ius civibus definiebas constituam ex, dolores atomorum vis et. His ea nihil vitae, aliquam disputationi te usu, pri in bonorum legendos.

Acknowledgments

I'd like to thank the little people.

Libris scriptorem ne eam. At vel scripserit instructior, vitae invenire forensibus his te. Eam nonumy quodsi ut, melius nominavi volutpat sea et, aliquam vulputate per ut. Qui ex facilis vulputate tincidunt, te pri placerat reprimique omittantur. In vis dolor graece philosophia, ea diam impedit eum, per et harum legendos percipitur. Enim inani movet id sea, nam semper praesent moderatius in.

Hinc deterruisset interpretaris pro et, sint nostrud ea mei, pri liber maluisset splendide et. Eos an dolor eruditi evertitur. In mea oportere signiferumque, ius civibus definiebas constituam ex, dolores atomorum vis et. His ea nihil vitae, aliquam disputationi te usu, pri in bonorum legendos.

CHAPTER 1

Introduction

Allow me to introduce you to my dissertation.

CHAPTER 2

Background

In this chapter, we first review some polyhedral decompositions and how these relate to generating functions for the lattice points of a polyhedra.

2.1. Working with generating functions: an example

Let us start with an easy example. Consider the one dimensional polyhedra in $\mathbb{R}$ given by $P = [0, n]$. We encode the latticd points of $P \cap \mathbb{Z}$ by placing each integer point as the power of a monomial, there by obtaining the polynomial $S(P; z) := z^0 + z + z^2 + z^3 + \cdots + z^n$. The polynomial $S(P; z)$ is called the *generating function of* P. Notice that counting $P \cap \mathbb{Z}$ is equivalat to evaluating $S(P, 1)$.

In terms of the computational complexity, listing each monomial in the polynomial $S(P, z)$ results in a polynomial with exponential length in the bit length of n. However, we can rewrite the summation with one term:

$$S(P, z) = z^0 + z^1 + \cdots + z^n = \frac{1 - z^{n+1}}{1 - z}.$$

Counting the number of point in $|P \cap \mathbb{Z}|$ is no longer as simple as evaluating $\frac{1-z^{n+1}}{1-z}$ at $z = 1$ because this is a singularity. However, this singularity is removable. One could perform long-polynomial division, but this would result in a expoentially long polynomial in the bit length of n. Another option that yeilds a polynomail time algorithm would be to apply L'Hospital's rule:

$$\lim_{z \to 1} S(P, z) = \lim_{z \to 1} \frac{-(n+1)z^n}{1} = n + 1.$$

Notice that $S(P, z)$ can be written in two ways:

$$S(P, z) = \frac{1}{1 - z} - \frac{z^{n+1}}{1 - z} = \frac{1}{1 - z} + \frac{z^n}{1 - z^{-1}}.$$

The first two rational expressions have a nice description in terms of their series expansion:

$$1 + z + \cdots + z^n = (z^0 + z^1 + \cdots) - (z^{n+1} + z^{n+2} + \cdots).$$

For the secont two rational functions, we have to be careful about the domain of convergence when computing the series expansion. Notice that in the series expansion,

$$\frac{1}{1-z} = \begin{cases} z^0 + z^1 + z^2 \cdots & \text{if } |z| < 1 \\ -z^{-1} - z^{-2} - z^{-3} - \cdots & \text{if } |z| > 1 \end{cases}$$

$$\frac{z^n}{1-z^{-1}} = \begin{cases} -z^{n+1} - z^{n+2} - z^{n+3} - \cdots & \text{if } |z| < 1 \\ z^n + z^{n-1} + z^{n-2} + \cdots & \text{if } |z| > 1 \end{cases}$$

adding the terms when $|z| < 1$ or $|z| > 1$ results in the desired polynomial: $z^0 + z^1 + \cdots + z^n$. But we can also get the correct polynomial by adding the series that corresponds to different domains of convergence. However, to do this we must now add the series $\cdots + z^{-2} + z^{-1} + 0 + z + z^2 + \cdots$ which corresponds to the polyhedra that is the entire real line:

$$\begin{aligned} 1 + z + \cdots z^n = & (1 + z + z^2 + \cdots) \\ & + (z^n + z^{n-1} + \cdots) \\ & - (\cdots + z^{-2} + z^{-1} + 0 + z + z^2 + \cdots) \end{aligned}$$

and

$$\begin{aligned} 1 + z + \cdots z^n = & (-z^{-1} - z^{-2} - z^{-3} - \cdots) \\ & + (-z^{n+1} - z^{n+2} - z^{n+3} - \cdots) \\ & + (\cdots + z^{-2} + z^{-1} + 0 + z + z^2 + \cdots) \end{aligned}$$

Hence by including the series $\cdots + z^{-2} + z^{-1} + 0 + z + z^2 + \cdots$, we can perform the series expansion of $\frac{1}{1-z} + \frac{z^n}{1-z^{-1}}$ by computing the series expansion of each term on potentially different domains of convergence.

In the next sections, we will develop rigorous justification for adding the series $\cdots + z^{-2} + z^{-1} + 0 + z + z^2 + \cdots$.

2.2. Indicator functions

DEFINITION 2.2.1. The indicator function, $[A] : \mathbb{R}^d \to \mathbb{R}$, of a set $A \subseteq \mathbb{R}^d$ takes two values: $[A](x) = 1$ if $x \in A$ and $[A](x) = 0$ otherwise.

The set of indicator functions on $\mathbb{R}^d$ froms a vectorspace with pointwise additions and scalar multiplication. Notice that $[A] \cdot [B] = [A \cap B]$, and $[A] + [B] = [A \cup B] + [A \cap B]$.

DEFINITION 2.2.2. The *cone* of a set $A \subseteq \mathbb{R}^d$ is all conic combinations of the points from A:

$$\mathrm{Cone}(A) := \left\{ \sum_i \alpha_i a_i \mid a_i \in A, \alpha_i \in \mathbb{R}_{\geq 0} \right\}$$

DEFINITION 2.2.3. Let P be a polyhedron and $x \in P$. Then the *tangent cone*, of P at x is the polyhedral cone

$$\mathrm{TCone}(P, x) := x + \mathrm{Cone}(P - x)$$

Note that if x is a vertex of P, and P is given by an inequality description, then the tangent cone $\mathrm{TCone}(P, x)$ is the intersection of inequalities that are tight at x. Also, $\mathrm{TCone}(P, x)$ includes the affine hull of the face that x is in, so the tangent cone is pointed only if x is a vertex.

When F is a face of P, we will also use the notation $\mathrm{TCone}(P, F)$ to denote $\mathrm{TCone}(P, x)$ where x is any interior point of F.

THEOREM 2.2.4 ([**Bri37**], [**Gra71**]). *Let P be a polyhedron, then*

$$[P] = \sum_F (-1)^{\dim(F)} [\mathrm{TCone}(P, F)]$$

where the sum ranges over all faces F of P including $F = P$ but excluding $F = \emptyset$

This theorem is saying that if the generating function of a polytope is desired, it is sufficient to just find the generating function for every face of P. The next corollary takes this a step further and says it is sufficient to just construct the generating functions associated at each vertex. This is because, as we will see, the generating functions for non-pointed polyhedra can be ignored.

COROLLARY 2.2.5. *Let P be a polyhedron, then*

$$[P] \equiv \sum_{v\in V} [\mathrm{TCone}(P,v)] \quad (\text{mod } \textit{indicator functions of non-pointed polyhedra}),$$

where V is the vertex set of P.

2.3. Generating functions of simple cones

In this section, we quickly review the generating function for summation and integration when the polyhedra is a cone.

The next Proposition serves as a basis for all the summation algorithms we will discus.

PROPOSITION 2.3.1 (Theorem 13.8 in [**Bar08**]). *There exists a unique valuation $S(\,\cdot\,,\ell)$ which associates to every polyhedron $P \subset \mathbb{R}^d$ a meromorphic function in ℓ so that the following properties hold*

(1) *If $\ell \in \mathbb{R}^d$ such that $e^{\langle \ell,x\rangle}$ is summable over the lattice points of P, then*

$$S(P,\ell) = \sum_{P\cap\mathbb{Z}^d} e^{\langle \ell,x\rangle}.$$

(2) *For every point $s \in \mathbb{Z}^d$, one has*

$$S(s+P,\ell) = e^{\langle \ell,s\rangle} S(P,\ell).$$

(3) *If P contains a straight line, then $S(P,\ell)=0$.*

A consequence of the valuation property is the following fundamental theorem. It follows from the Brion–Lasserre–Lawrence–Varchenko decomposition theory of a polyhedron into the supporting cones at its vertices [**BHS09, Bri88, Bar08, Las83**].

LEMMA 2.3.2. *Let P be a polyhedron with set of vertices $V(P)$. For each vertex s, let $C_s(P)$ be the cone of feasible directions at vertex s. Then*

$$S(P,\ell) = \sum_{s\in V(P)} S(s + C_s(P),\ell).$$

This last lemma can be identified as the natural result of combining Corollary 2.2.5 and Proposition 2.3.1 part (3). A non-pointed polyhedra is another characterization of a polyhedra that contain a line.

Note that the cone $C_s(P)$ in Lemma 2.3.2 may not be simplicial, but for simplicial cones there are explicit rational function formulas. As we will see in Proposition 2.3.6, one can derive an explicit formula for the rational function $S(s+C_s(P),\ell)$ in terms of the geometry of the cones.

PROPOSITION 2.3.3. *For a simplicial full-dimensional pointed cone C generated by rays $u_1, u_2, \dots u_d$ (with vertex 0) and for any point s*

$$S(s+C,\ell) = \sum_{a\in(s+\Pi_C)\cap\mathbb{Z}^d} e^{\langle \ell,a\rangle} \prod_{i=1}^{d} \frac{1}{1-e^{\langle \ell,u_i\rangle}}$$

where $\Pi_c := \{\sum_{i=1}^d \alpha_i u_i \mid 0 \le \alpha_i < 1\}$ These identities holds as a meromorphic function of ℓ and pointwise for every ℓ such that $\langle \ell, u_i\rangle \neq 0$ for all u_i.

The continuous generating function for P almost mirrors the discrete case.

PROPOSITION 2.3.4 (Theorem 8.4 in [**Bar08**]). *There exists a unique valuation $I(\,\cdot\,,\ell)$ which associates to every polyhedron $P \subset \mathbb{R}^d$ a meromorphic function so that the following properties hold*

(1) *If ℓ is a linear form such that $e^{\langle \ell,x\rangle}$ is integrable over P with the standard Lebesgue measure on $\mathbb{R}^d$, then*
$$I(P,\ell) = \int_P e^{\langle \ell,x\rangle}\,\mathrm{d}x$$
(2) *For every point $s \in \mathbb{R}^d$, one has*
$$I(s+P,\ell) = e^{\langle \ell,s\rangle} I(P,\ell).$$
(3) *If P contains a line, then $I(P,\ell) = 0$.*

LEMMA 2.3.5. *Let P be a polyhedron with set of vertices $V(P)$. For each vertex s, let $C_s(P)$ be the cone of feasible directions at vertex s. Then*

$$I(P,\ell) = \sum_{s\in V(P)} I(s+C_s(P),\ell).$$

Again, this last lemma can be identified as the natural result of combining Corollary 2.2.5 and Proposition 2.3.4 part (3).

PROPOSITION 2.3.6. *For a simplicial full-dimensional pointed cone C generated by rays $u_1, u_2, \dots u_d$ (with vertex 0) and for any point s*

$$I(s + C, \ell) = vol(\Pi_C) e^{\langle \ell, s \rangle} \prod_{i=1}^{d} \frac{1}{-\langle \ell, u_i \rangle}.$$

These identities holds as a meromorphic function of ℓ and pointwise for every ℓ such that $\langle \ell, u_i \rangle \neq 0$ for all u_i.

2.4. Generating function for non-simple cones

In the last section, we reviewed results that reduced the problem of finding the generating function of $\sum_{x \in P \cap \mathbb{Z}^d} e^{\langle \ell, x \rangle}$ and $\int_P e^{\langle \ell, x \rangle} \, \mathrm{d}x$ of a general polytope P to finding the same generating functions of a simplicial cone C. This works perfectly if the poltope P is simplicial, meaning at every vertex v, the cone $\mathrm{TCone}(P, v)$ is simplicial. When $\mathrm{TCone}(P, v)$ is not simplicial, the solution is to triangulate it.

DEFINITION 2.4.1. A triangulation of a cone C is the set Γ of simplicial cones C_i of the same dimension as the affine hull of C such that

(1) the union of all the simplicial cones in Γ is C,
(2) the intersection of any pair of simplicial cones in Γ is a common face of both simplicial cones,
(3) and every ray of every simplicial cone is also a ray of C.

Let C be a full-dimensional pointed polyhedral cone, and $\Gamma = \{C_i \mid i \in I_1\}$ be a triangulation into simplicial cones C_i where I_1 is a finite index set. It is true that $C = \cup_{i \in I_1} C_i$, but $[C] = \sum_{i \in I_1} [C_i]$ is false as points on the boundary of two adjacent simplicial cones are counted multiple times. The correct approach is to use the inclusion-exclusion formula:

$$[C] = \sum_{\emptyset \neq J \subseteq I_1} (-1)^{|J|-1} [\cap_{j \in J} C_j]$$

Also, note that this still holds true when C (and the C_i) is shifted by a point s. When $|J| \geq 2$, $I(\cap_{j \in J} C_j, \ell) = 0$ as $\cap_{j \in J} C_j$ is not full-dimensional, and the integration is done with respect to the Lebesgue measure on $\mathbb{R}^d$. This leads us to the next lemma.

LEMMA 2.4.2. *For any triangulation D_s of the feasible cone C_s at each of the vertices s of the polytope P we have $I(P, \ell) = \sum_{s \in V(P)} \sum_{C \in D_s} I(s + C, \ell)$*

Lemma 2.4.2 states that we can triangulate a polytope's feasible cones and apply the integration formulas on each simplicial cone without worrying about shared boundaries among the cones. Note that there is no restriction on how the triangulation is performed.

More care is needed for the discrete case as $S(\cap_{j \in J} C_j, \ell) \neq 0$ when $|J| \geq 2$. We want to avoid using the inclusion-exclusion formula as it contains exponentially many terms (in size of $|I_1|$).

The discrete case has another complication. Looking at Proposition 2.3.3, we see that the sum

$$\sum_{a \in (s + \Pi_C) \cap \mathbb{Z}^d} e^{\langle \ell, a \rangle}$$

has to be enumerated. However, there could be an exponential number of points in $(s + \Pi_C) \cap \mathbb{Z}^d$ in terms of the bit length of the simplicial cone C.

We will illustrate one method for solving these problems called the *Dual Barvinok Algorithm.*

DEFINITION 2.4.3. Let $A \subseteq \mathbb{R}^d$, the *polar* of A is the set

$$A^\circ = \{x \in \mathbb{R}^d \mid \langle x, a \rangle \leq 1 \text{ for every } a \in A\}$$

LEMMA 2.4.4. *Cones enjoy many properties under the polar operation. Let C be a finitly generated cone in $\mathbb{R}^d$, then*

(1) $C^\circ = \{x \in \mathbb{R}^d \mid \langle x, c \rangle \leq 0, \forall c \in C\}$,
(2) C° *is also a cone,*
(3) $(C^\circ)^\circ = C$, *and*
(4) *if* $C = \{x \mid A^T x \leq 0\}$, *then* C° *is generated by the columns of* A.

The next lemma is core to Brion's "polarization trick" [**Bri88**] for dealing with the inclusion-exclusion terms.

Lemma 2.4.5 (Theorem 5.3 in [**Bar08**]). *Let $\mathcal{C}$ be the vector space spanned by the indicator functions of all closed convex sets in $\mathbb{R}^d$. Then there is a unique linear transformation $\mathcal{D}$ from $\mathcal{C}$ to itself such that $\mathcal{D}([A]) = [A^\circ]$ for all non-empty closed convex sets A.*

Instead of taking the non-simplicial cone C and triangulating it, we first compute C° and triangulate it to $\Gamma' = \{C_i^\circ \mid i \in I_2\}$. Then

$$[C^\circ] = \sum_{i\in I_2}[C_i^\circ] + \sum_{\emptyset\neq J\subseteq I_2, |J|>1} (-1)^{|J|-1}[\cap_{j\in J} C_j^\circ].$$

Applying the fact that $(C^\circ)C^\circ = C$ and the polar operator is linear, we get

$$[C] = \sum_{i\in I_2}[C_i] + \sum_{\emptyset\neq J\subseteq I_2, |J|>1} (-1)^{|J|-1}[(\cap_{j\in J} C_j^\circ)^\circ].$$

Notice that the polar of a full-dimensional pointed cone is another full-dimensional pointed cone. For each J with $|J| \geq 2$, $\cap_{j\in J} C_j^\circ$ is not a full-dimensional cone. The polar of a cone that is not full dimensional is a cone that contains a line. Hence $S((\cap_{j\in J} C_j^\circ)^\circ, \ell) = 0$.

For an extremely well written discussion on other practical algorithms to solve these problems using slightly different decompositions, see [**?**].

as improved by making use of see [6] and [4 , Remark 4.3]: The computations with rational generating fun ctions are in- variant with respect to the contribution of non-pointed con es (cones contain- ing a non-trivial linear subspace). The reason is that the ra tional generating function of every non-pointed cone is zero. By operating in t he dual space, i.e., by computing with the polars of all cones, lower-dimen sional cones can be safely discarded, because this is equivalent to discardi ng non-pointed cones in the primal space. Thus at each level of the decomposi tion, only at most d cones are created. This dual variant of Barvinoks algorithm has efficient implementations in LattE [8 , 9 , 10] and the library barvinok

K oppe [13] considered both irrational triangulations and irrational signed decompositions . He constructed a uniform irrational shifting vector v which ensures that (3) holds for all cones v + C i that are created during the course of the recursive Barvinok decomposition method. The implemen tation of this method in a version of LattE [14] was the first practically efficient variant of Barvinoks algorithm that works in the primal space. The benefits of a decomposition in the primal space are twofol d. First, it allows to effectively use the method of stopped decomposition [13], where the recursive decomposition of the cones is stopped before u nimodular cones are obtained. For certain classes of polyhedra, this techni que reduces the running time by several orders of magnitude. Second, for some classes of polyhedra such as the cross-poly topes, it is prohibitively expensive to compute triangulations of the v ertex cones in the dual space. An all-primal algorithm [13] that computes both triangulations and signed decompositions in the primal space is therefore a ble to handle problem instances that cannot be solved with a dual algorith m in reasonable time

the number of lattice points in $\Pi_C \cap \mathbb{Z}^d$ may be exponential in the encoding of C. A standard solution to these problems is to compute a triangulation for each cone C_s, and then decompose each simplicial cone into unimodular cones via Algorithm **??**.

(1) ex of a line via polynomial
(2) 2d tile case
(3) Brion, tangent cones
(4) Barvinok

Figure 2.1. A picture of a gull.

CHAPTER 3

Long Title of Second Chapter

Rain is wet. The conclusions are immediate and self-evident. We leave them as an exercise for the reader.

APPENDIX A

Long Title of Appendix A

Observations of non-wet rain have recently appeared in the literature. In this Appendix, we briefly consider the implications of these observations for the analysis offered in this dissertation.

Bibliography

[Bar08] A. I. Barvinok, *Integer points in polyhedra*, Zürich Lectures in Advanced Mathematics, European Mathematical Society (EMS), Zürich, Switzerland, 2008.

[BHS09] M. Beck, C. Haase, and F. Sottile, *(formulas of brion, lawrence, and varchenko on rational generating functions for cones)*, The Mathematical Intelligencer **31** (2009), no. 1, 9–17 (English).

[Bri37] C. Brianchon, *Théorème nouveau sur les polyèdres*, École Polytechnique **15** (1837), 317–319.

[Bri88] M. Brion, *Points entiers dans les polyèdres convexes*, Ann. Sci. École Norm. Sup. **21** (1988), no. 4, 653–663.

[Gra71] J. P. Gram., *Om rumvinklerne i et polyeder*, Tidsskrift for Math **3(4)** (1871), 161–163.

[Las83] J. B. Lasserre, *An analytical expression and an algorithm for the volume of a convex polyhedron in* $\mathbb{R}^n$, Journal of Optimization Theory and Applications **39** (1983), no. 3, 363–377 (English).