Research Statement – Alex Dunbar

1. Overview

My research goal is to leverage algebraic and geometric structures to better understand optimization and machine learning problems. I am particularly interested in connections between optimization and real and tropical algebraic geometry.

Real algebraic geometry is the study of polynomial equations and inequalities defined over the real numbers. In recent years, a growing body of work has linked convex optimization and real algebraic geometry. A key observation which allows for this link is that a **sum of squares** certificate that a polynomial is globally nonnegative can be obtained via a **semidefinite programming** feasibility problem. **Tropical algebra**, on the other hand, studies the $(\max, +)$ semiring structure on $\mathbb{R} \cup \{-\infty\}$ and polynomial and rational functions over this semiring. The geometric objects associated to such functions are polyhedral. A recent line of work has connected the study of **ReLU neural networks** to tropical geometry, showing that functions which are representable by ReLU neural networks are also representable as tropical rational functions.

In my PhD work, I contribute to these growing connections. In the realm of real algebraic geometry, I study systems of quadratic equations and inequalities. In particular, in joint work with Greg Blekherman [BD24], we use tools from algebraic topology to connect certificates of emptiness and convexity properties of the solution set of a system of quadratic inequalities to properties of an associated curve. In ongoing joint work with Greg Blekherman and Rainer Sinn, we seek low-rank solutions to the semidefinite programming feasibility problems which certify that a ternary form is a sum of squares. In another direction, in ongoing joint work with Elizabeth Newman, we study semidefinite programming using the framework of tensor-tensor products for third-order tensors, connecting the choice of tensor-tensor product to the representation theory of finite groups. Moving to tropical algebra, in joint work with Lars Ruthotto [DR24], we provide a heuristic for regression problems over the class of tropical rational functions, motivated by the connection with ReLU networks. Prior to my PhD work, I worked on multiobjecitve integer programming. In joint work with Saumya Sinha and Andrew Schaefer [DSS23], we studied relaxations for such problems and developed analogs of Lagrangian and superadditive duality for the multiobjective setting.

Detailed summaries of my past and current projects, as well as potential directions for future research, are included below. Section 2 details work in real algebraic geometry and optimization, Section 3 details work on regression with tropical rational functions, and Section 4 describes work on multiobjective integer programming. Section 5 describes my plans for future research.

2. Real Algebraic Geometry and Optimization

Real algebraic geometry is concerned with the solution sets of polynomial equations and inequalities defined over the real numbers. Of particular importance are connections with convexity and convex optimization. This section details my contributions to the area, organized across three projects.

2.1. **Systems of Quadratics.** An important problem in real algebraic geometry is the certification that a variety has no real points. Semidefinite programming can be used for such certificates. If Q_1, Q_2, Q_3 are quadratic forms in the variables $x_0, x_1, \dots, x_n$ then $-(\sum_{i=0}^n x_i^2)$ is a sum of squares mod the ideal $\langle Q_1, Q_2, Q_3 \rangle$ if and only if there is a positive definite linear combination of the Q_i and these conditions certify the emptiness of $\mathcal{V}_{\mathbb{R}}(Q_1, Q_2, Q_3) \subseteq \mathbb{RP}^n$. However, if there is no positive definite linear combination of the Q_i, one needs to consider alternative certificates such as a hierarchy of SDPs.

In [BD24], we use algebraic topology to develop an alternate certificate. Our main tool is a spectral sequence developed by Agrachev and Lerario [AL12] which relates the homology of $\mathcal{V}_{\mathbb{R}}(Q_1, Q_2, Q_3)$ to the eigenvalues of linear combinations of the Q_i.

Theorem 1 ([BD24, Theorem 1.1]). *Suppose that $n \geq 4$ and that $g(\lambda) = \det(\lambda_1 Q_1 + \lambda_2 Q_2 + \lambda_3 Q_3)$ is smooth. Then, $\mathcal{V}_{\mathbb{R}}(Q_1, Q_2, Q_3) \subseteq \mathbb{RP}^n$ is empty if and only if g is hyperbolic and there is $\mu \in \mathbb{R}^3$ such that $\sum_{i=1}^3 \mu_i Q_i$ has n positive eigenvalues and one negative eigenvalue.*

The hyperbolicity of the polynomial g gives an interpretation in terms of convexity. A smooth **hyperbolic polynomial** is a homogeneous polynomial which defines a hypersurface that has maximally nested ovaloids, the innermost one bounding a convex set. Any hyperbolic plane curve possesses a definite determinantal representation by the Helton-Vinnikov theorem [HV07]. A consequence of Theorem 1 is that the only other determinantal representations of hyperbolic plane curves which define an empty subvariety of $\mathbb{RP}^n$ must have a combination which achieves n positive eigenvalues. In particular, there must be a hyperbolicity cone $\mathcal{P}$ of g which either has positive semidefinite matrices or matrices with exactly two negative eigenvalues.

The spectral sequence argument used to prove Theorem 1 also provides information about the solution sets of systems of quadratic *inequalities*. Let $S = \{x \in \mathbb{R}^n \mid Q_i(x, 1) \leq 0, i = 1, 2, 3\}$ be the affine set of solutions to the inequalities defined by the Q_i. Building on work in [DMS22; BDS24a], we seek a description of $\mathrm{conv}(S)$ in terms of **aggregations** (nonnegative linear combinations) of the defining inequalities. In [BD24, Theorem 1.2], we show that when the variety $\mathcal{V}_{\mathbb{R}}(Q_1, Q_2, Q_3)$ is empty, no nonzero aggregation lies in the hyperbolicity cone $\mathcal{P}$, and the set S has no points at infinity, then $\mathrm{conv}(S)$ can be obtained via finitely many aggregations. This expands the known cases where $\mathrm{conv}(S)$ can be obtained via aggregations beyond those known in [DMS22; BDS24a].

2.2. **Tensor-Tensor Products.** A line of work [New19; Kil+21; KKA15] has developed a family of tensor-tensor products for third order tensors which depends on an invertible matrix M. Formally, given $a, b \in \mathbb{C}^p$, the multiplication $a *_M b = M^{-1}\,\mathrm{diag}(Ma)Mb$ turns $\mathbb{C}^p$ into a commutative ring. Third order tensors in $\mathbb{C}^{m \times n \times p}$ are then viewed as matrices with entries in $\mathbb{C}^p$, refered to as **tubes**. Many matrix factorizations, such as the SVD, have analogs for third order tensors equipped with the $*_M$-product. In ongoing work with Elizabeth Newman, we offer two contributions to this area.

First, we investigate $*_M$-semidefinite tensors and semidefinite programming problems. For the remainder of this section, we assume that M is an orthogonal real matrix. A symmetric tensor $\mathcal{A} \in \mathbb{R}^{n \times n \times p}$ is M-PSD if $\langle \mathcal{X}, \mathcal{A} *_M \mathcal{X} \rangle \geq 0$ for all $\mathcal{X} \in \mathbb{R}^{n \times 1 \times p}$, where $\langle \mathcal{A}, \mathcal{B} \rangle = \sum_{i,j,k} a_{ijk} b_{ijk}$ and a tensor is symmetric if $a_{ijk} = a_{jik}$ for all $1 \leq i, j \leq n$. Such tensors share many properties analogous to those of PSD matrices. Moreover, given a symmetric tensor $\mathcal{A}$, membership in PSD_M^n can be determined facewise in the transform domain (i.e., by looking at the tensor $\hat{\mathcal{A}}$ with tubes $\hat{\mathcal{A}}_{i,j,:} = M \mathcal{A}_{i,j,:}$). Using the framework of M-PSD tensors, we study M-semidefinite programming problems, optimization problems of the form

$$(M\text{-SDP}) \qquad \max \langle \mathcal{C}, \mathcal{X} \rangle \ \text{ s.t. } \ \langle \mathcal{A}^{(i)}, \mathcal{X} \rangle = b^{(i)} \text{ for } i = 1, 2, \ldots, m, \ \text{ and } \mathcal{X} \in \mathrm{PSD}_M^n.$$

The problem (M-SDP) can be used to formulate a minimum nuclear norm tensor completion problem, similar to [KLL21].

Since membership in PSD_M^n can be determined facewise in the transform domain, solving problems of the form (M-SDP) amounts to solving block diagonal SDPs. We connect this block diagonalization to the invariant semidefinite programming problems studied by Gaterman and Parillo [GP04]. In doing so, we provide an interpretation of the $*_M$-product through the lens of representation theory, our second primary contribution.

Let $\rho : G \to GL_p(\mathbb{C})$ be a representation of a finite group and $\mathbb{C}^p \simeq \bigoplus_{i=1}^m V_i$ the decomposition into irreducibles. If M represents a change of basis from the standard basis on $\mathbb{C}^p$ to a basis compatible with the decomposition into irreducibles, then Schur's Lemma gives an explicit subspace W_ρ of tubes for which the multiplication map is ρ-equivariant. The linear equations defining W_ρ are described in terms of the dimensions of the V_i in the decomposition of $\mathbb{C}^p$ into irreducibles and the rows of M. If the affine constraints in (M-SDP) enforce membership in W_ρ, then the problem can be written as a standard SDP with matrix variable of size $np \times np$ which is invariant under the representation of G given by $\hat{\rho}(g) = I_m \otimes \rho(g)$.

2.3. **Pythagoras Numbers for Ternary Forms.** Sums of squares certificates for global nonnegativity of polynomials correspond to semidefinite programming feasibility problems. If $f \in \mathbb{R}[x_1, x_2, \ldots, x_n]_{2d}$ is a degree $2d$ homogeneous polynomial, then f is a sum of squares if and only if $f(x) = [x]_d^\top Q[x]_d$ for some positive semidefinite matrix Q, where $[x]_d$ is a vector containing all monomials of degree d in the variables $x_1, x_2, \ldots, x_n$. Such a matrix Q is called a Gram Matrix for f. The **Pythagoras number** $\mathrm{py}(n, 2d)$ is the minimum rank r such that every sum of squares $f \in \mathbb{R}[x_1, x_2, \ldots, x_n]_{2d}$ has a Gram matrix of rank at most r. One can computationally leverage low-rank structure in semidefinite programming problems through Burer-Monteiro methods, which replace the $\binom{n-1+d}{d} \times \binom{n-1+d}{d}$ matrix variable with a factorization into $BB^\top$, where B is $\binom{n-1+d}{d} \times r$, see e.g. [BM03].

One application of computing Pythagoras numbers in general is the problem of partially specified PSD matrix completion problems. This in turn has applications to rigidity theory and embedding discrete metric spaces into $\mathbb{R}^r$ [LV14b; LV14a].

In forthcoming work with Greg Blekherman and Rainer Sinn [BDS24b], we investigate the cases $n = 3$ and $8 \le 2d \le 12$. It is well-known that $d+1 \le \mathrm{py}(3, 2d) \le d+2$ (see e.g. [Ble+22; Sch17]). In our work, we relate the problem of computing $\mathrm{py}(3, 2d)$ to structure theorems for Artinian Gorenstien algebras [Die96]. Using this approach, we prove that $\mathrm{py}(3, 2d) = d + 1$ for $2d = 8, 10, 12$.

3. Regression with Tropical Rational Functions

A **tropical rational function** is a function of the form

$$f(x) = \max_{i=1,2,\ldots D}(\langle w^{(i)}, x\rangle + p_i) - \max_{i=1,2,\ldots,D}(\langle w^{(i)}, x\rangle + q_i).$$

Where $w_1, w_2, \ldots, w_n \in \mathbb{Z}^n$. Such functions are piecewise linear and known to be highly expressive–up to scaling of inputs, any function which can be written as a ReLU Neural Network can be written as a tropical rational function [ZNL18]. To this end, in [DR24], we propose a heuristic for ℓ_∞-regression over the space of tropical rational functions. Specifically, given a dataset $(x^{(1)}, y^{(1)}), (x^{(2)}, y^{(2)}), \ldots, (x^{(N)}, y^{(N)}) \in \mathbb{R}^n \times \mathbb{R}$ and a fixed set of exponents $w^{(1)}, w^{(2)}, \ldots, w^{(D)} \in \mathbb{Z}^n$, we want to solve

(1)
$$\arg \min_{p_1,p_2,\ldots,p_D,q_1,q_2,\ldots,q_D} \left(\max_{j=1,2,\ldots,N} \left| \max_{i=1,2,\ldots D} (\langle w^{(i)}, x^{(j)} \rangle + p_i) - \max_{i=1,2,\ldots,D} (\langle w^{(i)}, x^{(j)} \rangle + q_i) - y_j \right| \right).$$

To do so, we leverage existing results on tropical polynomial regression to develop an alternating minimization heuristic for (1) which alternates between optimizing the variables $(p_1, p_2, \ldots, p_D)$ and $(q_1, q_2, \ldots, q_D)$. This builds on existing work in tropical polynomial regression (see e.g., [MT20]). Experimentally, our heuristic finds approximate solutions to (1); however, in the context of learning problems these solutions have large validation errors, suggesting a need for regularization. Theoretically, we show that the objective function $\mathcal{L}$ in the problem (1) is itself a tropical rational function of the variables $p_1, p_2, \ldots, p_D, q_1, q_2, \ldots, q_D$ and that the alternating minimization heuristic produces iterates which are contained in the nondifferentiability locus of $\mathcal{L}$, providing a geometric interpretation of the problem.

4. Multiobjective Integer Programming Duality

A **multiobjective integer (linear) program** is the problem of optimizing several, often competing, linear objective functions over the integral points of a polyhedron. Here, a feasible solution x is *efficient* to a problem with objective matrix C if $Cx - Cy \in \mathbb{R}^k_+$ (i.e. has nonnegative entries) for all feasible solutions y, and we seek the set of all efficient solutions. Such problems are typically written in the form

$$\text{Max } Cx \text{ s.t. } Ax \leqq b,\ x \in \mathbb{Z}^n_+, \tag{MOIP}$$

where $C \in \mathbb{Z}^{k\times n}$, $A \in \mathbb{R}^{m\times n}$, and $b \in \mathbb{R}^m$.

In [DSS23], we investigate **relaxations** and develop **dual programs** for (MOIP). A primary focus in our work is Lagrangian relaxation. Given a matrix Λ of nonnegative multipliers and a partition of the constraint matrix into two sets of rows $A^{(1)}$ and $A^{(2)}$, the Lagrangian relaxation of (MOIP) is the problem

$$\text{Max } Cx + \Lambda(b^{(1)} - A^{(1)}x) \text{ s.t. } A^{(2)}x \leqq b^{(2)},\ x \in \mathbb{Z}^n_+. \tag{LR(Λ)}$$

We show via example that there can be efficient solutions to (LR(Λ)) whose corresponding objective values provide strictly better bounds on the set of feasible objectives for (MOIP) than a convex hull relaxation. This contrasts with the single objective case, where the convex hull relaxation is the tightest possible relaxation. We introduce Lagrangian and superadditive dual problems, which generalize Lagrangian and superadditive dual problems for the single objective case and provide conditions for the dual to be strong at supported efficient solutions, i.e. those which are also efficient to both (MOIP) and its convex hull relaxation [DSS23, Theorems 7 and 8].

5. Future Directions

I have plans for several directions for future research, extending my current results and broadening into adjacent areas. A direct extension of the work in [BD24] is to consider sums of squares certificates for properties of systems of three quadratics.

Question 5.1. Let Q_1, Q_2, Q_3 be quadratic forms on $\mathbb{R}^{n+1}$. What is the minimum positive integer d such that $-(\sum_{i=0}^{n} x_i^2)^d$ is a sum of squares mod the ideal $\langle Q_1, Q_2, Q_3\rangle$ when $\mathcal{V}_{\mathbb{R}}(Q_1, Q_2, Q_3)$ is empty? What is the minimum nonnegative even integer $2d$ such that if Q_1 is positive on $\mathcal{V}_{\mathbb{R}}(Q_2, Q_3)$ then there is $f \in \Sigma_{n,2d}$ such that fQ_1 is SOS mod $\langle Q_2, Q_3\rangle$?

It was shown by Barvinok in [Bar93] that the emptiness of a real variety defined by m quadratic forms can be determined in time polynomial in n. Question 5.1 is additionally related to existing results in the real algebraic geometry literature, where it is known that any nonnegative quadratic form on a variety of minimal degree is a sum of squares of linear forms [BSV16]. A variety X defined by the complete intersection of two quadratics is of *almost minimal degree*, that is, $\operatorname{codim}(X) = \deg(X) + 2$ and therefore a natural candidate for future work. The problem of minimizing a quadratic function subject to a finite number of quadratic constraints has also been studied in the optimization literature [Bie16], where it is known that a polynomial time algorithm can solve the problem. We propose to answer these questions by building on the results in [BD24]. Theorem 1 shows that there are limited possibilities for the sturcture of the degree 2 part of the ideal $I = \langle Q_1, Q_2, Q_3\rangle$. We intend to leverage this dichotomy to understand conditions for existence of a linear functional ℓ which separates $\Sigma_{n,2d}$ and I_{2d}. Moreover, due to the limited structure of I_2, we conjecture that the integer d in Question 5.1 is independent of the number of variables n.

This is similar to many problems in convex algebraic geometry, where it is necessary to understand linear functionals on the vector space of polynomials $\ell \in \mathbb{R}[x_0, \dots, x_n]^*_{2d}$. In many cases, one wishes to decompose ℓ into the sum of point evaluations $\ell = \sum_{i=1}^{k} ev_{z_i}$, with an upper bound on the number of points k. By identifying $\ell = \langle F, \bullet\rangle$ through the apolar inner product, this becomes a question about **symmetric tensor decomposition**. Moreover, one often needs the points z_i to satisfy some reality conditions, for example coming in complex conjugate pairs or being almost real, meaning that all points are real except possibly one complex conjugate pair. The study of almost-real rank of forms $F \in \mathbb{R}[x_0, x_1]_d$ was initiated in [BCJ24], and we propose extending these results with a focus on the underlying geometry.

Question 5.2. What are typical and maximal almost real ranks of forms $F \in \mathbb{R}[x_0, \dots, x_n]_d$? What are the geometric properties of schemes $\Gamma \subseteq \mathbb{P}^n$ which certify the almost real rank of F? More generally, what are the answers to the analogous questions if we are concerned with less restrictive reality requirements on the points?

Symmetric tensor ranks have also recently appeared in the study of neural networks with monomial activation functions, as important components in descriptions of the *neurovariety*, the Zariski closure of the set of functions representable with a given architecture [Koh24]. I plan to investigate analogous questions for networks with ReLU activation functions from a tropical perspective.

Question 5.3. Is the set of functions representable by a ReLU neural network of fixed architecture given by $\operatorname{trop}(X)$ for some semialgebraic set X defined over $\mathbb{R}\{\{t\}\}$? How do factorizations of tropical polynomials affect representability?

Answering Question 5.3 would help to strengthen the connection between ReLU network architecture and the geometry of tropical functions. Ultimately, this could establish a use for applying the alternating heuristic for the solution of (1) to the training of ReLU networks. Additionally, determining if (1) has a solution with objective less than some fixed $\delta > 0$ is a tropical linear programming feasiblity problem, inviting the use of tropical convexity to understand bounds on approximation error for ReLU networks of specified architecture.

References

[AL12] A. Agrachev and A. Lerario. "Systems of quadratic inequalities". In: *Proc. Lond. Math. Soc. (3)* 105.3 (2012), pp. 622–660.

[Bar93] Alexander I. Barvinok. "Feasibility testing for systems of real quadratic equations". In: *Discrete Comput. Geom.* 10.1 (1993), pp. 1–13.

[BCJ24] Grigoriy Blekherman, Justin Chen, and Jaewoo Jung. "Sums of squares, Hankel index and almost real rank". In: *Forum of Mathematics, Sigma* 12 (2024), e70.

[BD24] Grigoriy Blekherman and Alex Dunbar. "A Topological Approach to Simple Descriptions of Convex Hulls of Sets Defined by Three Quadrics". In: *arXiv preprint arXiv:2405.18282* (2024).

[BDS24a] Grigoriy Blekherman, Santanu S. Dey, and Shengding Sun. "Aggregations of Quadratic Inequalities and Hidden Hyperplane Convexity". In: *SIAM J. Optim.* 34.1 (2024), pp. 98–126.

[BDS24b] Grigoriy Blekherman, Alex Dunbar, and Rainer Sinn. "Pythagoras Numbers of Ternary Forms". In Preparation. 2024.

[Bie16] Daniel Bienstock. "A note on polynomial solvability of the CDT problem". In: *SIAM J. Optim.* 26.1 (2016), pp. 488–498.

[Ble+22] Grigoriy Blekherman et al. "Sums of squares and quadratic persistence on real projective varieties". In: *J. Eur. Math. Soc. (JEMS)* 24.3 (2022), pp. 925–965.

[BM03] Samuel Burer and Renato D. C. Monteiro. "A nonlinear programming algorithm for solving semidefinite programs via low-rank factorization". In: vol. 95. 2. Computational semidefinite and second order cone programming: the state of the art. 2003, pp. 329–357.

[BSV16] Grigoriy Blekherman, Gregory G. Smith, and Mauricio Velasco. "Sums of squares and varieties of minimal degree". In: *J. Amer. Math. Soc.* 29.3 (2016), pp. 893–913.

[Die96] Susan J. Diesel. "Irreducibility and dimension theorems for families of height 3 Gorenstein algebras". In: *Pacific J. Math.* 172.2 (1996), pp. 365–397.

[DMS22] Santanu S. Dey, Gonzalo Muñoz, and Felipe Serrano. "On obtaining the convex hull of quadratic inequalities via aggregations". In: *SIAM J. Optim.* 32.2 (2022), pp. 659–686.

[DR24] Alex Dunbar and Lars Ruthotto. "Alternating minimization for regression with tropical rational functions". In: *Algebr. Stat.* 15.1 (2024), pp. 85–111.

[DSS23] Alex Dunbar, Saumya Sinha, and Andrew J Schaefer. "Relaxations and duality for multiobjective integer programming". In: *Mathematical Programming* (2023), pp. 1–40.

[GP04] Karin Gatermann and Pablo A. Parrilo. "Symmetry groups, semidefinite programs, and sums of squares". In: *J. Pure Appl. Algebra* 192.1-3 (2004), pp. 95–128.

[HV07] J. William Helton and Victor Vinnikov. "Linear matrix inequality representation of sets". In: *Communications on Pure and Applied Mathematics* 60.5 (2007), pp. 654–674.

[Kil+21] Misha E. Kilmer et al. "Tensor-tensor algebra for optimal representation and compression of multiway data". In: *Proc. Natl. Acad. Sci. USA* 118.28 (2021), Paper No. e2015851118, 12.

[KKA15] Eric Kernfeld, Misha Kilmer, and Shuchin Aeron. "Tensor-tensor products with invertible linear transforms". In: *Linear Algebra Appl.* 485 (2015), pp. 545–570.

[KLL21] Hao Kong, Canyi Lu, and Zhouchen Lin. "Tensor Q-rank: new data dependent definition of tensor rank". In: *Mach. Learn.* 110.7 (2021), pp. 1867–1900.

[Koh24] Kathlen Kohn. "The geometry of the neuromanifold". In: *SIAM News* 57.6 (2024), pp. 2, 4.

[LV14a] M. Laurent and A. Varvitsiotis. "Positive semidefinite matrix completion, universal rigidity and the strong Arnold property". In: *Linear Algebra Appl.* 452 (2014), pp. 292–317.

[LV14b] Monique Laurent and Antonios Varvitsiotis. "A new graph parameter related to bounded rank positive semidefinite matrix completions". In: *Math. Program.* 145.1-2 (2014), pp. 291–325.

[MT20] Petros Maragos and Emmanouil Theodosis. "Multivariate tropical regression and piecewise-linear surface fitting". In: *ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE. 2020, pp. 3822–3826.

[New19] Elizabeth Newman. "A Step in the Right Dimension: Tensor Algebra and Applications". PhD thesis. Tufts University, May 2019.

[Sch17] Claus Scheiderer. "Sum of squares length of real forms". In: *Math. Z.* 286.1-2 (2017), pp. 559–570.

[ZNL18] Liwen Zhang, Gregory Naitzat, and Lek-Heng Lim. "Tropical geometry of deep neural networks". In: *International Conference on Machine Learning*. PMLR. 2018, pp. 5824–5832.