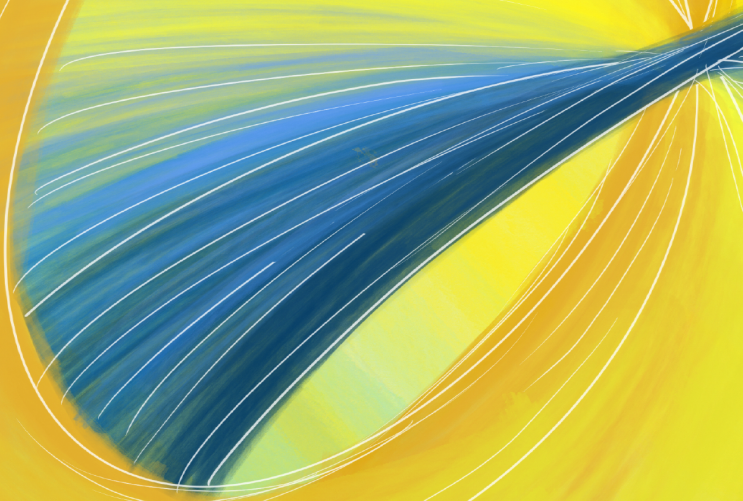
Scaling problems, algorithms and applications to Computer Science and Statistics



Rafael Oliveira
Akshay Ramachandran

impa

33º Colóquio Brasileiro de Matemática

Scaling problems, algorithms and applications to Computer Science and Statistics

Scaling problems, algorithms and applications to Computer Science and Statistics
Primeira impressão, julho de 2021

Publicado no Brasil / Published in Brazil.

ISBN 978-65-89124-18-4
MSC (2020) Primary: 14L24, Secondary: 62D99, 47N10, 90C26

Coordenação Geral Carolina Araujo

Produção Books in Bytes **Capa** Izabella Freitas & Jack Salvador

Realização da Editora do IMPA
IMPA
Estrada Dona Castorina, 110
Jardim Botânico
22460-320 Rio de Janeiro RJ

www.impa.br
editora@impa.br

Contents

1 Introduction

Several problems in mathematics, computer science, machine learning and statistics exhibit inherent symmetries which can be described by a group acting linearly on a vector space. Oftentimes, these symmetries are implicit or disguised in the "natural" description of the problems. Thus, many qualitative and quantitative properties inherent to these problems have laid dormant or unexplored until recent developments, which made crucial use of the group action structure, allowed for significant progress in such problems.
In this survey, we will give an overview of the phenomenon described above. Our emphasis will be on the geometric properties of such group actions and on the generalization of convexity that arises from natural optimization problems along group orbits, which we term by *geodesic convexity*.

1.1 Brief history

In the early days of invariant theory, the period known as classical invariant theory (late 1800s), the question of understanding geometric properties of plane curves which were invariant under changes of bases received a lot of attention. Notable mathematicians who worked on this question at the time include Aronhold, Clebsch, Gordan, Cayley, Sylvester and Hilbert. During this time, their focus was on

finding functions which associate a number to each curve that was independent of the choice of basis.

Mathematicians at the time came to realize that such a problem (invariance under change of basis) was about the action of a group on the ambient vector space, usually the action of the special linear group $\mathrm{SL}_n(\mathbb{C})$, and that the functions that they were studying were polynomial functions over the coefficients of the polynomials defining the curves being studied.

A simple example of the problem above, which is familiar to us all (but most likely not in this language), is the problem of deciding when a quadratic form in two variables, given by $ax^2+bxy+cy^2 \in \mathbb{C}[x,y]$, has a double root. As it turns out, the property of "having a double root" is independent of the choice of basis (that is, if we change basis $(x',y') = (x,y)A$, the quadratic will still have a double root) and it is characterized by the vanishing of the *discriminant* $\Delta := b^2-4ac$. Thus, the property of having a double root is completely captured by a polynomial function on the coefficients of the quadratic form (i.e. a,b,c).

The major research effort at the time was to determine the set of all polynomial invariants of "nice" group actions on certain vector spaces. Since the set of all invariant polynomials forms a $\mathbb{C}$-algebra, one of the main questions at the time, which was termed the first fundamental theorem of invariant theory, was to prove whether a group action had a finite set of generating invariants as a $\mathbb{C}$-algebra.

This research effort culminated in Hilbert's seminal works Hilbert (1890, 1893), where he proved such fundamental theorems as the Hilbert Basis Theorem, the Nullstellensatz, the Syzygy theorem, and the rationality of the Hilbert series. Hilbert's motivation to prove these theorems was to give a constructive proof that the ring of invariants was finitely generated, and to give a full description of the ring of invariants.

While the algebraic side of invariant theory has received much attention since the nineteenth century, it was only in the seminal works of Mumford and the striking developments by Kempf, Ness, Kostant and Kirwan, among others, that the geometric side of invariant theory really flourished. In geometric invariant theory,[1] given a group G acting on a vector space V, the goal is to understand the quotient space V/G given by the set of orbits of the group action on V.

In the development of geometric invariant theory by Mumford, a special optimization problem is central: the *null-cone problem*, which was already defined in the work of Hilbert (1893). We will study this problem in greater detail in Chap-

[1] The setting of geometric invariant theory is more general, and we have decided to remain with the setting of a group acting on a vector space for simplicity. For the more general treatment we refer the reader to Mumford, Fogarty, and Kirwan (1994) and Wallach (2017).

ter 2, but now through the lens of optimization over a Riemannian manifold.

1.2 Examples of scaling problems

In this section we describe some concrete examples of scaling problems which have seen important progress in recent years by the use of the optimization approach to geometric invariant theory. The beauty of these concrete examples, apart from being fundamental problems in their respective subareas of mathematics, is that we can state them even without the definitions from invariant theory, and we will do so in order to motivate the reader and to showcase how the inherent symmetries of a problem may be disguised in its statement.

1.2.1 Matrix Scaling

Given a non-negative $n \times n$ matrix $A \in \mathrm{Mat}_n(\mathbb{R})$, we say that A is *doubly-stochastic* if all row and column sums of A are equal to 1. An important problem, which appears in several disciplines ranging from economics, engineering, transportation theory and computer science, is the question of deciding when one can "transform" a non-negative matrix A (approximately) into a doubly-stochastic matrix B by multiplying the rows and columns of A by positive scalars. This problem motivates the following definition:

Definition 1 (Scaling of a matrix)**.** Given a non-negative matrix $A \in \mathrm{Mat}_n(\mathbb{R})$, we say that $\hat{A}$ is a *scaling* of A if it can be obtained by multiplying the rows and columns of A by positive scalars. In other words, $\hat{A}$ is a scaling of A if there exist positive diagonal matrices $R, C \in \mathrm{Mat}_n(\mathbb{R})$ such that $\hat{A} = RAC$.

As the reader can realize, the approximate version of the question is often needed, since the matrix $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ can be scaled arbitrarily close to a doubly-stochastic matrix (i.e. the identity) but it cannot be scaled exactly to a doubly-stochastic matrix (since the non-zero pattern of the matrix does not change by scaling). This motivates us to define a measure for how close a matrix is to being doubly-stochastic:

Definition 2 (Distance to doubly-stochastic)**.** Given a non-negative matrix $A \in \mathrm{Mat}_n(\mathbb{R})$, define its distance to doubly stochastic to be

$$\mathrm{ds}(A) = \sum_{i=1}^{n} (r_i - 1)^2 + \sum_{j=1}^{n} (c_j - 1)^2$$

where r_i and c_j denote the i^{th} row sum and the j^{th} column sum, respectively.

With the definition of distance as above, we can say that a non-negative matrix A is *approximately scalable* to doubly-stochastic if, and only if, for every $\varepsilon > 0$, there exists a scaling A_ε of A such that $\mathrm{ds}(A_\varepsilon) \leqslant \varepsilon$. We call such a scaling A_ε an ε-scaling of A.

Thus, given a non-negative matrix A, two natural questions arise: when is a matrix approximately scalable? If a matrix is scalable, can one efficiently find an ε-scaling, for a given parameter $\varepsilon > 0$?

We have arrived at the (computational version of the) matrix scaling problem:

Question 1.2.1 (Matrix Scaling)**.** Given a non-negative matrix $A \in \mathrm{Mat}_n(\mathbb{R})$ and an accuracy parameter $\varepsilon > 0$, is there a scaling B of A such that $\mathrm{ds}(B) \leqslant \varepsilon$? If there is such a scaling, output it.

As mentioned in the beginning of this section, the matrix scaling problem has historically appeared independently in several scientific areas, and to solve the matrix scaling problem the following natural iterative algorithm has often been used: if the matrix is not *row-stochastic* (that is, the row sums are 1), make it row-stochastic by properly normalizing the rows. This may change the column sums. If the matrix is not *column-stochastic*, make it column-stochastic by normalizing the columns.

Input: a non-negative matrix $A \in \mathrm{Mat}_n(\mathbb{R})$, $\varepsilon > 0$.
Output: a scaling B of A such that $\mathrm{ds}(B) \leqslant \varepsilon$, if one exists. NO, otherwise.

- Set $B \leftarrow A$
- For T steps, while $\mathrm{ds}(B) > \varepsilon$:
  1. if B is not row-stochastic, multiply i^{th} row of B by $r_i(B)^{-1}$ for all $i \in [n]$
  2. if B is not column-stochastic, multiply j^{th} column by $c_j(B)^{-1}$ for all $j \in [n]$
- If at any point above $\mathrm{ds}(B) \leqslant \varepsilon$, return B, otherwise, after the T steps, return NO.

Algorithm 1: RAS algorithm

The algorithm above is a special case of a general optimization paradigm

known as *alternating minimization*, where to minimize a function one tries to alternately minimize simpler functions in an alternate fashion, where the idea is that the simpler functions are much easier to optimize (sometimes the optimum for the simpler functions can even be written in closed form, as is our case).

In Section 1.3, we will see an analysis of the algorithm shown above, as well as a striking application of using matrix scaling to obtain a deterministic approximation to the permanent of non-negative matrices, and the connection between matrix scaling and bipartite matchings.

For more background on the matrix scaling problem, we refer the reader to the surveys Garg and Oliveira (2018) and Idel (2016).

1.2.2 Optimal Transport Distances in Finite Distributions

Given two discrete probability measures $r, c \in \mathbb{R}_+^d$ over a finite set $[d] := \{1, 2, \ldots, d\}$, we define $U(r, c)$ to be the *transportation polytope* of r and c, which is given by

$$U(r, c) := \{P \in \mathrm{Mat}_d(\mathbb{R}_+) \mid P1_d = r,\ P^\dagger 1_d = c\}$$

where 1_d is the all ones vector of dimension d. An element of $U(r, c)$ is called a *transportation matrix* or joint distribution, as we will now see.

One can view $U(r, c)$ as the set of all *joint probability distributions* of two discrete random variables X, Y each taking values in $[d] := \{1, 2, \ldots, d\}$ where X has probability distribution r and Y has probability distribution c. In this case, each matrix $P \in U(r, c)$ is such that $P_{i,j} = \Pr[X = i, Y = j]$.

Given a cost matrix $M \in \mathrm{Mat}_d(\mathbb{R})$, the cost of mapping measure r to c using a transportation matrix P can be quantified by the Frobenius inner product $\langle M, P \rangle := \mathrm{Tr}[M^\dagger P]$. Thus, we have arrived at the *optimal transport* problem between r and c given cost M:

$$d_M(r, c) := \min_{P \in U(r,c)} \langle M, P \rangle.$$

Optimal transport of measures is a problem of great practical importance, having originated in the works of Monge (in 1871) and developed further by Kantorovich[2] (in 1942) in their studies on optimal allocation and transportation of

[2]Interestingly, Kantorovich is regarded as the father of Linear Programming.

resources. While the formulation above can be solved via standard convex optimization methods, or more specialized methods for linear programs, the complexity of solving the optimal transport problem above turns out to be $O(d^3 \log d)$ in practice, which turns out to be prohibitive for many applications.

In Cuturi (2013), the author proposed to add entropic constraints on the optimal transport problem to find optimal joint distributions which have *small mutual information*, as these solutions have applications to machine learning. Thus, Cuturi proposed to find solutions in the convex set

$$U_\alpha(r,c) := \{P \in U(r,c) \mid d_{KL}(P \mid\mid rc^\dagger) \leqslant \alpha\}$$

where $\alpha \geqslant 0$. Moreover, in the same work, Cuturi showed how one can use the matrix scaling algorithm from the previous section to solving the modified optimal transport problem! This yields a much simpler algorithm with a much better runtime in practice for computing such distances, and as showed in Cuturi (ibid.), these new distances have much better practical applications than the unconstrained original distances.

For more background on optimal transport and its connections to matrix scaling and machine learning, we refer the reader to Cuturi (ibid.), where the connection presented above was first made, and where we drew this example from. For connections to image retrieval, see the seminal work of Rubner, Tomasi, and Guibas (2000). For a comprehensive treatment of optimal transport, see Villani (2008).

1.2.3 Paulsen Problem

The Paulsen problem is a central question in frame theory as discussed in Casazza and Kutyniok (2013).

Question 1.2.2. Let $U = \{u_1, ..., u_n\} \subseteq \mathbb{C}^d$ be a spanning set of vectors satisfying

$$\frac{1-\varepsilon}{d} I_d \preceq \sum_{j=1}^{n} u_j u_j^* \preceq \frac{1+\varepsilon}{d} I_d, \qquad \forall j \in [n] : \frac{1-\varepsilon}{n} \leqslant \|u_j\|_2^2 \leqslant \frac{1+\varepsilon}{n}. \tag{1.2.1}$$

What is the minimum distance $\sum_{j=1}^n \|v_j - u_j\|_2^2$ over all $V = \{v_1, ..., v_n\}$ satisfying Equation (1.2.1) exactly:

$$\sum_j v_j v_j^* = \frac{1}{d} I_d, \qquad \forall j \in [n] : \|v_j\|_2^2 = \frac{1}{n}.$$

Note that this is a different normalization, by a factor d, than normally given in the literature.

Vectors satisfying Equation (1.2.1) are known as ε-doubly balanced frames. The balance properties of doubly balanced frames, where $\varepsilon = 0$, are exploited to give strong results in coding theory and signal processing Casazza and Kutyniok (ibid.). Constructions of exactly doubly balanced frames are difficult and often rely on complicated algebraic structures. On the other hand, there are many simple algorithms to construct ε-doubly balanced frames. For example, a large enough set of random vectors will satisfy Equation (1.2.1) for some small ε with high probability. The Paulsen problem asks, for a given ε-doubly balanced frame, whether the conditions in Equation (1.2.1) can be corrected without moving too much. Since randomly generated frames are nearly doubly balanced, analyzing the distance bound in this case is of special importance.

Holmes and Paulsen (2004) studied frames from the perspective of coding theory, and showed that doubly balanced frames were optimally robust with respect to a single erasure. They also showed that Grassmannian frames, doubly balanced frames with large pairwise angles, were optimal for two erasures.

To address the difficulty of constructing these structured frames, the authors of Holmes and Paulsen (ibid.) suggested a simple numerical approach: first generate random frames, which approximately satisfy Equation (1.2.1), and then correct the conditions. Random frames are good candidates for both of these settings because they are approximately doubly balanced and have large pairwise angles with high probability. One goal of the Paulsen problem is then to validate this numerical algorithm as a simple method of constructing structured frames. The formalization below is from Cahill and Casazza (2013).

Conjecture 1.2.3 (Paulsen Problem)**.** Let $p(d, n, \varepsilon)$ be the smallest function such that for all ε-doubly balanced $U = \{u_1, ..., u_n\} \subseteq \mathbb{C}^d$, there exists a doubly balanced $V = \{v_1, ..., v_n\} \subseteq \mathbb{C}^d$ such that

$$\|V - U\|_F^2 = \sum_{j=1}^{n} \|v_j - u_j\|_2^2 \leqslant p(d, n, \varepsilon).$$

Then this distance function p can be taken independent of n.

The optimal function p has been unknown for almost twenty years, despite considerable attention in the frame theory literature. Prior to the work of Kwok, Lau, Lee, et al. (2017), the only known results on the function p were given

by Casazza, Fickus, and Mixon (2012) and Bodmann and Casazza (2010), and showed $p \leqslant poly(d, n, \varepsilon)$ when d, n are relatively prime and ε is small enough.

These results left open Conjecture 1.2.3, which was positively resolved in Kwok, Lau, Lee, et al. (2017).

Theorem 1.2.4 (Theorem 1.3.1 in Kwok, Lau, Lee, et al. (ibid.))**.** *The distance function can be bounded by* $p(d, n, \varepsilon) \lesssim d^{11/2}\varepsilon$. *In particular it can be taken independent of* n.

The new idea in this work was to use scaling algorithms like those studied recently in Garg, Gurvits, et al. (2016). To carry out this approach, Kwok, Lau, Lee, et al. (2017) defined a dynamical system which corrected approximately doubly balanced frames. This dynamical system could then be analyzed using tools from the operator scaling analysis of Garg, Gurvits, et al. (2016). The full proof of Kwok, Lau, Lee, et al. (2017) required a smoothed analysis approach coupled with an involved convergence analysis of the dynamical system.

Subsequently, in the aptly titled "Paulsen Problem made Simple", Hamilton and Moitra (2019) improved the distance bound to $p(d, n, \varepsilon) \lesssim d\varepsilon$, using a totally different and much shorter method. This almost matches the known lower bound, as there are simple examples showing $p \gtrsim \varepsilon$. Ramachandran (2021) revisits the dynamical system approach and closes this gap by using tools from geodesic convex optimization.

This dynamical system can also be analyzed to give a refined distance bound for the case of random frames, which answers the original motivation of the Paulsen problem.

Theorem 1.2.5 (Theorem 1.12 in Kwok, Lau, and Ramachandran (2019))**.** *For any* $n \geqslant poly(d)$ *large enough, if* $U = \{u_1, ..., u_n\} \subseteq \mathbb{R}^d$ *is generated such that each* u_j *is independent and uniformly distributed on* $\frac{1}{\sqrt{n}} S^{d-1}$*, then with high probability* U *is* ε*-doubly balanced for* $\varepsilon \leqslant \widetilde{O}(\sqrt{\frac{d}{n}})$*, and there exists doubly balanced* V *such that*

$$\|V - U\|_F^2 \lesssim \varepsilon^2.$$

This result validates the numerical approach suggested in Holmes and Paulsen (2004) to generate doubly balanced frames, and therefore gives a satisfactory answer to the original motivation for Question 1.2.2. It also gives the following corollary on Grassmannian frames.

Títulos Publicados — 33º Colóquio Brasileiro de Matemática

Geometria Lipschitz das singularidades – *Lev Birbrair e Edvalter Sena*

Combinatória – *Fábio Botler, Maurício Collares, Taísa Martins, Walner Mendonça, Rob Morris e Guilherme Mota*

Códigos Geométricos – *Gilberto Brito de Almeida Filho e Saeed Tafazolian*

Topologia e geometria de 3-variedades – *André Salles de Carvalho e Rafał Marian Siejakowski*

Ciência de Dados: Algoritmos e Aplicações – *Luerbio Faria, Fabiano de Souza Oliveira, Paulo Eustáquio Duarte Pinto e Jayme Luiz Szwarcfiter*

Discovering Euclidean Phenomena in Poncelet Families – *Ronaldo A. Garcia e Dan S. Reznik*

Introdução à geometria e topologia dos sistemas dinâmicos em superfícies e além – *Víctor León e Bruno Scárdua*

Equações diferenciais e modelos epidemiológicos – *Marlon M. López-Flores, Dan Marchesin, Vítor Matos e Stephen Schecter*

Differential Equation Models in Epidemiology – *Marlon M. López-Flores, Dan Marchesin, Vítor Matos e Stephen Schecter*

A friendly invitation to Fourier analysis on polytopes – *Sinai Robins*

PI-álgebras: uma introdução à PI-teoria – *Rafael Bezerra dos Santos e Ana Cristina Vieira*

First steps into Model Order Reduction – *Alessandro Alla*

The Einstein Constraint Equations – *Rodrigo Avalos e Jorge H. Lira*

Dynamics of Circle Mappings – *Edson de Faria e Pablo Guarino*

Statistical model selection for stochastic systems – *Antonio Galves, Florencia Leonardi e Guilherme Ost*

Transfer Operators in Hyperbolic Dynamics – *Mark F. Demers, Niloofar Kiamari e Carlangelo Liverani*

A Course in Hodge Theory Periods of Algebraic Cycles – *Hossein Movasati e Roberto Villaflor Loyola*

A dynamical system approach for Lane–Emden type problems – *Liliane Maia, Gabrielle Nornberg e Filomena Pacella*

Visualizing Thurston's Geometries – *Tiago Novello, Vinícius da Silva e Luiz Velho*

Scaling Problems, Algorithms and Applications to Computer Science and Statistics – *Rafael Oliveira e Akshay Ramachandran*

An Introduction to Characteristic Classes – *Jean-Paul Brasselet*

impa

Instituto de
Matemática
Pura e Aplicada

ISBN 978-65-89124-18-4

9 786589 124184