AFRL-AFOSR-VA-TR-2024-0076

Structured Assignment: Geometric Optimization Algorithms for Large-Scale Matching

Solomon, Justin
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
77 MASSACHUSETTS AVE
CAMBRIDGE, MA, 02139
USA

12/20/2023
Final Technical Report

DISTRIBUTION A: Distribution approved for public release.

Air Force Research Laboratory
Air Force Office of Scientific Research
Arlington, Virginia 22203
Air Force Materiel Command

REPORT DOCUMENTATION PAGE

PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ORGANIZATION.

1. REPORT DATE	**2. REPORT TYPE**	**3. DATES COVERED**	
20231220	Final	**START DATE** 20190601	**END DATE** 20230531

4. TITLE AND SUBTITLE
Structured Assignment: Geometric Optimization Algorithms for Large-Scale Matching

5a. CONTRACT NUMBER	**5b. GRANT NUMBER**	**5c. PROGRAM ELEMENT NUMBER**
	FA9550-19-1-0319	61102F
5d. PROJECT NUMBER	**5e. TASK NUMBER**	**5f. WORK UNIT NUMBER**

6. AUTHOR(S)
Justin Solomon

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)	**8. PERFORMING ORGANIZATION REPORT NUMBER**
MASSACHUSETTS INSTITUTE OF TECHNOLOGY 77 MASSACHUSETTS AVE CAMBRIDGE, MA 02139 USA	

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	**10. SPONSOR/MONITOR'S ACRONYM(S)**	**11. SPONSOR/MONITOR'S REPORT NUMBER(S)**
Air Force Office of Scientific Research 875 N. Randolph St. Room 3112 Arlington, VA 22203	AFRL/AFOSR RTA2	AFRL-AFOSR-VA-TR-2024-0076

12. DISTRIBUTION/AVAILABILITY STATEMENT
A Distribution Unlimited: PB Public Release

13. SUPPLEMENTARY NOTES

14. ABSTRACT
This project aimed to develop a mathematical theory and accompanying numerical algorithms for geometrically structured instances of the assignment problem. We considered generalizations of assignment/transportation mathematical programs that included a variety of objective functions, with the commonality that they are parameterized by an underlying distance function along a geometric domain. This additional assumption—which arises naturally in applications including 3D shape analysis, logistical operations/planning, and machine learning—provides added structure that we leveraged to derive fine-grained understanding of the energy landscape and tractability.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**	UU	8
U	U	U		

19a. NAME OF RESPONSIBLE PERSON	**19b. PHONE NUMBER** *(Include area code)*
WARREN ADAMS	00000000

Standard Form 298 (Rev.5/2020)
Prescribed by ANSI Std. Z39.18

Final Project Report

Structured Assignment: Geometric Optimization Algorithms for Large-Scale Matching

PI: Justin Solomon, Massachusetts Institute of Technology

Accomplishments

Research objectives

This project aimed to develop a mathematical theory and accompanying numerical algorithms for geometrically structured instances of the assignment problem. We considered generalizations of assignment/transportation mathematical programs that included a variety of objective functions, with the commonality that they are parameterized by an underlying distance function along a geometric domain. This additional assumption—which arises naturally in applications including 3D shape analysis, logistical operations/planning, and machine learning—provides added structure that we leveraged to derive fine-grained understanding of the energy landscape and tractability.

Our original proposal divided the effort among a set of "tasks" related to this broad theme:

- *Task 1: Simultaneous Assignment and Distance Computation*
  Classical problems like linear assignment/transport consider the cost matrix as an input. But in modern settings, the cost matrix itself can be too large to store, or it can be expensive to compute, or both. In this part of the project, we considered alternatives when the cost function is structured in a way where it can be computed on the fly or implicitly.

- *Task 2: Non-Standard Objectives*
  Classical optimization theory divides assignment problems into linear and quadratic assignment cases; the former is a linear program while the latter is a potentially nonconvex quadratic program. In this part of the project, we consider non-non-standard objectives for matching, e.g. containing L^p norms or measures of distortion/elasticity derived from physics.

- *Task 3: Non-Standard Data*
  Most of optimal transport theory involves transport of scalar-valued quantities, like the density of a probability measure. Here, we consider transport applied to objects like vector fields and Riemannian metrics, and the associated nonlinear problems that come with data of this type

- *Task 4: Iterated assignment*
  In this task, we consider solving repeated instances of assignment problems with shared cost matrices or other shared structure. The goal is to make the amortized cost over all instances of the problem lower than solving each independently.

While---as with any mathematical research projects---individual conjectures and directions varied in success relative to our original proposal (this aspects make it impossible to assess detailed percentages of completion), overall we are happy report that this work was overwhelmingly successful and led to many scientific developments relevant to each articulated tasks; 100% of the tasks above led to meaningful research results and publications as articulated in this report and others associated with this proposal.

Accomplishments during reporting period

As this report is the final report for our project, it covers four years of dedicated research in the MIT Geometric Data Processing Group. While it is impossible to enumerate every single accomplishment of this large-scale project, below we summarize a few of the major outcomes; we refer to the annual reports for additional details.

Optimal transport. A major theme in the proposed research involved advancement of theory and algorithms related to optimal transport of probability measures. Many of the publications included in this report provide state-of-the-art algorithms and models for solving transport problems in a variety of dimensions and domains.

A major breakthrough in computational optimal transport fueled in part by research in our team is the use of neural networks to parameterize the unknown primal and/or dual variables in transport problems. At the cost of nonconvexity, this change in approach empirically leads to much stronger performance of optimal transport methods on a variety of datasets and problems. For example, our team developed the first optimal transport-based method for computing barycenters in Wasserstein space and proposed multiple transport-based algorithms for gradient flows in the space of probability measures. Several of our algorithms in this space are able to adapt to time-varying problems, relating to the ‘iterated’ ideas articulated in our original proposal.

Beyond the basic problem of evaluating transport distances/matchings, we also proposed numerous extensions of the problem that extend the applicability and robustness of optimal transport and inspire new challenges in optimization. For example, we showed how Wasserstein distances enable new models of robustness for machine learning models that have access to additional unlabeled data, and we used optimal transport to take problems in Bayesian inference even in the presence of symmetry.

Relevant to the fundamental aspects of the problem, we proposed an algorithm for approximating optimal transport distances from samples of probability measures, with rigorous proof of faster convergence rates in certain clustered regimes. We also showed how optimal transport distances provide a generalization of the notion of variance that can be used for certain statistical models, including lengthy analysis of our new “k-variance” measure. Several of our algorithms also are applied to sampling from distributions given their density functions, or approximating a distribution with a nearby empirical measure as measured by the Wasserstein distance.

Non-standard data. Our work on assignment and transport between collections of non-standard data led to algorithms that bring efficient algorithms to new fields.

One thrust of our proposed research involved matching between low-dimensional domains, with application to medical imaging, computer graphics, and other fields. In this space, we proposed state-of-the-art algorithms that optimize measures of elasticity drawn from physics and material science, with top performance in surface/volume matching tasks. One of our optimization algorithms, inspired by the popular ADMM algorithm but with three blocks and a nonconvex objective function, led us to make a detailed study of multi-block ADMM for a new class of nonconvex objectives; our publication on this work includes theoretical proof of convergence that to our knowledge is not an extension of known theory/approaches.

Our work on non-standard data also inspired effort related to inverse problems related to optimal transport. One recent paper in this space shows how to incorporate our neural network optimal transport models with Riemannian geometry to recover cost functions given pairs of nearby measures under an unknown transport metric. This work has application in understanding the evolution of population statistics over time.

Dataset matching. Our research into assignment problems led to several published works comparing datasets, learned embeddings, and related notions. For example, AFOSR-funded postdoc Dr. Tal Shnitzer proposed a new spectral distance between datasets that lower-bounds transport distances without needing a point-to-point matching; her work enables prediction of performance of machine learning models on new biased data by measuring its relationship to the original training set. A recent extension by Dr. Shnitzer considers a related problem of fusing embeddings learned from different modalities, e.g. images and text.

Optimization machinery. Our work in this project used and further developed machinery for a number of different optimization problems. For example, in a project inspired by his work on optimal transport, PhD student Lingxiao Li developed techniques for approximating the proximal operator associated to generic objective functions, with application to recovering a list of local optima associated to a given nonconvex optimization problems; his work is accompanied by thorough theoretical analysis predicting when his neural network representation is likely to succeed. Our work on ADMM and parameterization also includes new optimization techniques with careful convergence analysis, as mentioned above. On the convex optimization side, our studies of transport and matching between vector fields and between functions represented on graphs/meshes led to new approaches for PDE-constrained optimization that conserves convex structures after discretization.

Open problems and new collaborations. Our extended examination of assignment and transportation problems funded by this grant has led to a variety of open problems and new collaborations in optimization, mathematical theory, and beyond. For example, our new results in simultaneous optimization of assignment problems and computation of cost matrices has led to new theory in optimal transport---including a new formulation of the 2-Wasserstein distance combining aspects of static and fluid-dynamical versions of the problem---that has inspired a long-term collaboration with Prof. Nestor Guillen (Texas State University).

Dissemination

The main mechanism for disseminating results relevant to this work is academic publication. A total of 56 papers published in peer-reviewed journals and conferences are included in this report, which spans roughly 4 years of research. PDF-format versions of all these papers are freely available on Prof. Solomon's web page; he can also send copies directly upon request.

Beyond publications, the team released software packages with many of the published research papers. Each can be found with an open source license (typically, the MIT license) on the GitHub website, facilitating reproduction of our scientific results and future research on related topics.

Finally, Prof. Solomon and his team regularly present their research at external events to spread word about their work. More comprehensive lists of relevant talks and related engagements can be found in annual reports associated with this proposal.

Note: The list of participants included with this report covers only those participants who salaries were derived from AFOSR funding, plus the project PI. Many other members of Prof. Solomon's team contributed to this work, since it is a central aspect of the team's research focus, as can be observed via the MIT-internal coauthors on almost all publications accompanying this report.

Impacts

Development of the principal disciplines of the project

This project led to several advances related to practical solution of assignment and transport problems in both low dimensions (with application e.g. to computer vision/graphics) and high dimensions (with application e.g. to machine learning).

In low dimensions, we have developed state-of-the-art algorithms for optimal transport along curved surfaces, as well as computation of low-distortion maps between surfaces and between volumes. These algorithms combine insight from differential geometry with efficient optimization techniques. Our extensions to non-standard data also allowed new applications e.g. to diffusion MRI signals, wherein the matched quantities are no longer scalar.

In high dimensions, our ideas and others from colleagues in the field have driven advancements in optimal transport for AI applications. Optimal transport ideas now are commonplace in the construction of generative AI models, and the basic techniques we and others developed enable links between these probabilistic algorithms and mathematical theory. Our work has pushed forward the feasibility of computing gradient flows, matchings, and variational approximations for high-dimensional data.

Other disciplines

Our work on distances between datasets and recovery of Riemannian metrics from optimal transport has been applied to biological data. Dr. Tal Shnitzer, the first author of the work on dataset distances, is now a full-time researcher at the Broad Institute, which studies problems in this space. Our work on surface matching also led to a long-term collaboration on BOLD MRI imaging of the human placenta, wherein our algorithms formed the basis for statistical shape analysis methods needed in that application.

Development of human resources

Locally, support from this grant was provided at a critical time for Prof. Solomon's research group, which has grown and become more established during the period of funding. Thanks in large part to the research funded from this grant, Prof. Solomon recently was awarded tenure in MIT's Department of Electrical Engineering and Computer Science, and he received the MIT Edgerton Award in recognition of his top performance among junior faculty.

More broadly, Solomon's team has a strong focus on outreach and broader impacts of their work. Several members of the project team have been involved in teaching and service, both internally to MIT and as part of conference tutorial programs. More prominently, during the period of support, Solomon also launched the Summer Geometry Initiative (SGI), an annual summer outreach program providing six weeks of research training to students drawn from underrepresented and underserved groups in geometry research; this program has had a measurable effect on the composition of the cohort of PhD applicants in geometry and has become an established part of MIT's broader plans in Diversity, Equity, and Inclusion (DEI).

Teaching and educational experiences

Educational materials related to work funded in this project have been included in Prof. Solomon's graduate course on Shape Analysis, whose contents are freely available online; some have also appeared in the Symposium on Geometry Processing (SGP) Graduate School program. Prof. Solomon has also used this research to refine the contents of the second edition of his textbook, *Numerical Algorithms*.

Infrastructure

Nothing to report.

Society

From a technical perspective, research in this proposal related to optimal transport in the space of graph partitions has been applied to the analysis of districting plans for compliance with civil rights law and other challenges.

Our SGI program (see above) has also had a concrete effect on the composition of incoming classes of PhD students in our field.

Changes

Changes in approach

Although the details of research projects change as mathematical and technical ideas become refined, our team largely followed the research agenda outlined in the original proposal. We do note a slight shift in approach inspired by broader trends in statistics and machine learning toward use of neural network representations for solution of certain assignment and transportation problems; we view this adjustment as a positive change that has led to increased performance of our models in a variety of settings.

Problems or delays

The project was extended one year due to delays associated with the Covid pandemic.

Expenditure

Nothing to report.

Human subjects, vertebrates, or biohazards

Nothing to report.

Place of performance

Nothing to report.

Technical Updates

Please see discussion above and accompanying published research papers. If there are particular research papers or projects on which it would be valuable to provide detailed technical updates beyond what we have included in our reports, please contact Prof. Solomon.