Journal of Computational Mathematics
Vol.38, No.6, 2020, 937–956.

http://www.global-sci.org/jcm
doi:10.4208/jcm.1907-m2017-0224

SOLUTION OF OPTIMAL TRANSPORTATION PROBLEMS USING A MULTIGRID LINEAR PROGRAMMING APPROACH*

Adam M. Oberman
Department of Mathematics, McGill University, Montreal, Canada
Email: adam.oberman@mcgill.ca
Yuanlong Ruan
Department of Mathematics, Beihang University, Beijing, China
Email: ruanylpku@gmail.com

Abstract

We compute and visualize solutions to the Optimal Transportation (OT) problem for a wide class of cost functions. The standard linear programming (LP) discretization of the continuous problem becomes intractable for moderate grid sizes. A grid refinement method results in a linear cost algorithm. Weak convergence of solutions is established and barycentric projection of transference plans is used to improve the accuracy of solutions. Optimal maps between nonconvex domains, partial OT free boundaries, and high accuracy barycenters are presented.

Mathematics subject classification: 49M99, 65K15, 90C05.
Key words: Optimal Transportation, Linear Programming, Monge-Kantorovich, Barycenter.

1. Introduction

There has recently been a great deal of interest in Computational Optimal Transportation (OT) with a number of new, effective numerical methods being proposed, including Entropic Regularization [14,15], extensions of the classical Benamou-Brenier methodology [4] and semi-discrete methods [19,20]. Computational OT is being applied to Machine Learning [10], Inverse Problems [21, 30], Data Assimilation [27]. Extension of the numerical methods have lead to methods for Wasserstein Gradient Flows [6] and Mean Field Games [8]. These methods have advanced far beyond the early methods which include the classical Benamou-Brenier method [3], Haker and Tannembaum's method [18] and more recent PDE based methods [9]. See also the book [22].

The applications of Optimal Transportation are varied, and the numerical methodology has advanced considerably. In this article, we focus on geometric applications of the OT problem, and visualization of classical problems. We achieve accurate results, using variable reduction for the linear programming method, but without the rigorous verification achieved by Schmitzer [24]. However our method is much simpler to implement: it can be implemented using vector operations in Matlab. Moreover, at least in the quadratic cost case, our heuristic is very close to the rigorous block data checking of Schmitzer. We implement a practical Linear Programming (LP) method for solving Optimal Transportation problems. The computational cost is linear in terms of solution time and memory requirements (see Table 4.1). We compute solutions

* Received September 19, 2017 / Revised version received April 26, 2019 / Accepted July 9, 2019 / Published online September 13, 2019 /

with problem sizes going up to half a million variables on a laptop computer using academic or commercial Linear Programming optimization software. Available parallel LP solvers allow the method to scale to even larger problem sizes.

Linear Programming solutions have very low accuracy compared to PDE methods. This is because the Kantorovich solutions are much weaker: the approximate solutions are plans (see Figure 1.1). This reduces the accuracy of solutions, and makes visualization difficult. Barycentric projection is used to recover the approximate map and improves the accuracy of solutions, leading to accuracy comparable to the accuracy of the PDE method.

A weak convergence proof is established using available stability results. Barycentric projection commutes with weak convergence, in the case of convex costs, so the convergence proof remains valid.

We use a heuristic grid refinement method, where the support of the solution on a coarse grid is used to infer the support of the solution on a refined grid. A rigorous approach to grid refinement is available: the work of Schmitzer [23, 24] ensures that at each refinement, the sparse solution of the linear program is optimal for the full problem. However this method introduces significant complexity to the code.

The efficiency and accuracy of the method reveals solution features not otherwise available, including optimal maps between nonconvex sets, and for non-quadratic cost functions. The method is generalized to other problems, allowing for the computation of accurate free boundaries in partial optimal transportation, and high resolution barycenters for shapes.

1.1. Related numerical work

Entropic regularization methods modify the optimization problem by adding a small multiple of an entropy term [14, 15] to the objective (or cost) function. The regularized problem can be solved by a Bregman iterative solution method [5]. The recent article [26] applied the method to large scale geometric problems.

Multiscale solvers have previously been used to improve solver performance. Merigot solved a sequence of OT problems where the target is a sum of Diracs [20]. Schmitzer [23, 24] used a grid refinement procedure which was applied to both LP and combinatorial optimization solvers. A one step grid refinement was used in [13] to find the support of the barcyenter. A grid refinement procedure was used in [7].

The early Benamou-Brenier formulation leads to a fluid mechanics solver, by adding a synthetic time variable to the problem [3], adding one dimension to the problem. The Monge-Ampère Partial Differential formulation was recently used to solve the OT problem using a convergent finite difference method [9]. This method is for quadratic cost, and places regularity requirements on the densities, one of which must have convex support. Our method has comparable accuracy to the PDE approach, see §4.3.

1.2. Background on the Optimal Transportation problem

The Monge formulation of the Optimal Transportation problem which seeks optimal *maps* while the Kantorovich formulation seeks optimal *transference plans*. Transference plans are a weaker notion of solution which can be computed using linear programming.

Given two probability measures, μ, ν with bounded supports $X, Y \subset \mathbb{R}^d$, respectively and the cost function $c(x, y) : X \times Y \to \mathbb{R}$. The goal is to rearrange one measure into the other,