

Chapter 9

Swarm Intelligence for Non-Negative Matrix Factorization

Andreas Janecek

University of Vienna, Austria

Ying Tan

Peking University, China

ABSTRACT

The Non-negative Matrix Factorization (NMF) is a special low-rank approximation which allows for an additive parts-based and interpretable representation of the data. This article presents efforts to improve the convergence, approximation quality, and classification accuracy of NMF using five different meta-heuristics based on swarm intelligence. Several properties of the NMF objective function motivate the utilization of meta-heuristics: this function is non-convex, discontinuous, and may possess many local minima. The proposed optimization strategies are two-fold: On the one hand, a new initialization strategy for NMF is presented in order to initialize the NMF factors prior to the factorization; on the other hand, an iterative update strategy is proposed, which improves the accuracy per runtime for the multiplicative update NMF algorithm. The success of the proposed optimization strategies are shown by applying them on synthetic data and data sets coming from the areas of spam filtering/email classification, and evaluate them also in their application context. Experimental results show that both optimization strategies are able to improve NMF in terms of faster convergence, lower approximation error, and better classification accuracy. Especially the initialization strategy leads to significant reductions of the runtime per accuracy ratio for both, the NMF approximation as well as the classification results achieved with NMF.

DOI: 10.4018/978-1-4666-2479-5.ch009

1. INTRODUCTION

Low-rank approximations are utilized in several content based retrieval and data mining applications, such as text and multimedia mining, web search, etc. and achieve a more compact representation of the data with only limited loss in information. They reduce storage and runtime requirements, and also reduce redundancy and noise in the data representation while capturing the essential associations. The Non-negative Matrix Factorization (NMF) (Lee & Seung, 1999) leads to a low-rank approximation which satisfies non-negativity constraints. NMF approximates a data matrix A by $A \approx WH$, where W and H are the NMF factors. NMF requires all entries in A , W and H to be zero or positive. Contrary to other low-rank approximations such as the Singular Value Decomposition (SVD), these constraints force NMF to produce so-called “additive parts-based” representations. This is an impressive benefit of NMF, since it makes the interpretation of the NMF factors much easier than for factors containing positive and negative entries (Berry, Browne, Langville, Pauca, & Plemmons, 2007; Janecek & Gansterer, 2010; Lee & Seung, 1999).

The NMF is usually not unique if different initializations of the factors W and H are used. Moreover, there are several different NMF algorithms which all follow different strategies (e.g., mean squared error, least squares, gradient descent) and produce different results. Mathematically, the goal of NMF is to find a “good” (ideally the best) solution of an optimization problem with bound constraints in the form $\min_{x \in \Omega} f(x)$, where $f: \mathbb{R}^N \rightarrow \mathbb{R}$ is the nonlinear objective function of NMF, and Ω is the feasible region (for NMF, Ω is restricted to non-negative values). f is usually not convex, discontinuous and may possess many local minima (Stadlthanner, Lutter, Theis, Lang, Tome, Georgieva, & Puntonet, 2007). Since meta-heuristic optimization algorithms are known to be able to deal well with such difficul-

ties they seem to be a promising choice for improving the quality of NMF. Over the last decades nature-inspired meta-heuristics, including those based on swarm intelligence, have gained much popularity due to their applicability for various optimization problems. They benefit from the fact that they are able to find acceptable results within a reasonable amount of time for many complex, large and dynamic problems (Blackwell, 2007). Although they lack the ability to guarantee the optimal solution for a given problem (comparably to NMF), it has been shown that they are able to tackle various kinds of real-world optimization problems (Chiong, 2009). Meta-heuristics as well as the principles of NMF are in accordance with the *law of sufficiency* (Kennedy, Eberhart, & Shi, 2001): If a solution to a problem is good enough, fast enough and cheap enough, then it is sufficient.

In this article we present two different strategies for improving the NMF using five optimization algorithms based on swarm intelligence and evolutionary computing: Particle Swarm Optimization (PSO), Genetic Algorithms (GA), Fish School Search (FSS), Differential Evolution (DE), and Fireworks Algorithm (FWA). All algorithms are population based and can be categorized into the fields of swarm intelligence (PSO, FSS, FWA), evolutionary algorithms (GA), and a combination thereof (DE). The goal is to find a solution with smaller overall error at convergence, and/or to speed up convergence of NMF (i.e., smaller approximation error for a given number of NMF iterations) compared to identical NMF algorithms without applied optimization strategy. Another goal is to increase the classification accuracy in cases where NMF is used as dimensionality reduction method for machine learning applications. The concepts of the two optimization strategies are the following: In the first strategy, meta-heuristics are used to initialize the factors W and H in order to minimize the NMF objective function *prior* to the factorization. The second strat-

23 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:
www.igi-global.com/chapter/swarm-intelligence-non-negative-matrix/70646

Related Content

Modelling Dimensions for Agent Organizations

Luciano R. Coutinho, Jaime S. Sichman and Olivier Boissier (2009). *Handbook of Research on Multi-Agent Systems: Semantics and Dynamics of Organizational Models* (pp. 18-50).

www.irma-international.org/chapter/modelling-dimensions-agent-organizations/21094

An Agent-based Model for Portfolio Optimization Using Search Space Splitting

Yukiko Orito, Yasushi Kambayashi, Yasuhiro Tsujimura and Hisashi Yamamoto (2011). *Multi-Agent Applications with Evolutionary Computation and Biologically Inspired Technologies: Intelligent Techniques for Ubiquity and Optimization* (pp. 19-34).

www.irma-international.org/chapter/agent-based-model-portfolio-optimization/46197

The Science of Social Emergence

R. Keith Sawyer (2009). *Handbook of Research on Agent-Based Societies: Social and Cultural Interactions* (pp. 1-16).

www.irma-international.org/chapter/science-social-emergence/19614

Penguins Huddling Optimisation

Mohammad Majid al-Rifaie (2014). *International Journal of Agent Technologies and Systems* (pp. 1-29).

www.irma-international.org/article/penguins-huddling-optimisation/117203

From Text to Semantic Geodata Enrichment

Khaoula Mahmoudi and Sami Faiz (2014). *International Journal of Agent Technologies and Systems* (pp. 28-44).

www.irma-international.org/article/from-text-to-semantic-geodata-enrichment/109601