Chapter 60

Swarm Intelligence for Dimensionality Reduction: How to Improve the Non-Negative Matrix Factorization with Nature-Inspired Optimization Methods

Andreas Janecek
University of Vienna, Austria

Ying Tan
Peking University, China

ABSTRACT

Low-rank approximations allow for compact representations of data with reduced storage and runtime requirements and reduced redundancy and noise. The Non-Negative Matrix Factorization (NMF) is a special low-rank approximation that allows for additive parts-based, interpretable representation of the data. Various properties of NMF are similar to Swarm Intelligence (SI) methods: indeed, most NMF objective functions and most SI fitness functions are non-convex, discontinuous, and may possess many local minima. This chapter summarizes efforts on improving convergence, approximation quality, and classification accuracy of NMF using five different meta-heuristics based on SI and evolutionary computation. The authors present (1) new initialization strategies for NMF, and (2) an iterative update strategy for NMF. The applicability of the approach is illustrated on data sets coming from the areas of spam filtering and email classification. Experimental results show that both optimization strategies are able to improve NMF in terms of faster convergence, lower approximation error, and/or better classification accuracy.

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 and 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 et al. 2007) (Janecek and Gansterer 2010) (Lee and 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, etc.) 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 $f : \mathbb{R}^N \to \mathbb{R}$ is the nonlinear objective function of NMF, and $\Omega$ is the fea-
sible region (for NMF, $\Omega$ is restricted to non-negative values). $f$ is usually not convex, discontinuous
and may possess many local minima (Stadlthanner, Lutter et al. 2007). Since meta-heuristic optimization
algorithms are known to be able to deal well with such difficulties 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 optimi-
zation problems. They benefit from the fact that they are able to find acceptable results within a reason-
able 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 (Eberhart,
Shi et al. 2001): If a solution to a problem is good enough, fast enough and cheap enough, then it is suf-
ficient.

In this chapter 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 applica-
tions. 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 strategy aims at iteratively improving the approximation quality
of NMF during the first iterations.

The proposed optimization strategies can be considered successful if they are able to improve the
NMF in terms of either (i) faster convergence (i.e. better accuracy per runtime) (ii) lower final approxi-
mation error, (iii) or better classification accuracy. The optimization of different rows of $W$ and
different columns of $H$ can be split up into several partly independent sub-tasks and can thus be execut-
ed concurrently. Since this allows for a parallel and/or distributed computation of both update strategies,
Related Content

Soft-Computing-Based Optimization of Low Return Loss Multiband Microstrip Patch Antenna
www.igi-global.com/chapter/soft-computing-based-optimization-of-low-return-loss-multiband-microstrip-patch-antenna/153830?camid=4v1a

Assessment of Multi-Engine Machine Translation for English to Hindi Language (MEMTEHiL):
Using F&A and iBLEU Metrics
www.igi-global.com/article/assessment-of-multi-engine-machine-translation-for-english-to-hindi-language-memtehil/177183?camid=4v1a

Network-Driven Analysis Methods and their Application to Drug Discovery
www.igi-global.com/chapter/network-driven-analysis-methods-their/52321?camid=4v1a

Software-Based Validation of the Differentiation Scheme for Ethernet and Wireless LAN Access Network Types Using an End-to-End Approach