Journal of Computational and Applied Mathematics 236 (2012) 2200–2222

Contents lists available at SciVerse ScienceDirect

Journal of Computational and Applied Mathematics

journal homepage: www.elsevier.com/locate/cam

On vector and matrix median computation

S. Setzer [a], G. Steidl [b,*], T. Teuber [b]

[a] Saarland University, Department of Mathematics and Computer Science, Campus E1.1, 66041 Saarbrücken, Germany

[b] University of Kaiserslautern, Department of Mathematics, Paul-Ehrlich-Str. 31, 67663 Kaiserslautern, Germany

ARTICLE INFO

Article history:
Received 16 September 2010

Keywords:
Vector and matrix median
Structure tensor
Proximation
ADMM
PPXA
Weiszfeld's algorithm

ABSTRACT

The aim of this paper is to gain more insight into vector and matrix medians and to investigate algorithms to compute them. We prove relations between vector and matrix means and medians, particularly regarding the classical structure tensor. Moreover, we examine matrix medians corresponding to different unitarily invariant matrix norms for the case of symmetric 2×2 matrices, which frequently arise in image processing. Our findings are explained and illustrated by numerical examples. To solve the corresponding minimization problems, we propose several algorithms. Existing approaches include Weiszfeld's algorithm for the computation of ℓ_2 vector medians and semi-definite programming, in particular, second order cone programming, which has been used for matrix median computation. In this paper, we adapt Weiszfeld's algorithm for our setting and show that also two splitting methods, namely the alternating direction method of multipliers and the parallel proximal algorithm, can be applied for generalized vector and matrix median computations. Besides, we compare the performance of these algorithms numerically and apply them within local median filters.

1. Introduction

While medians of one-dimensional data are well known in image processing, vector or matrix medians are not so common. The reason is that the generalization of the one-dimensional median to higher dimensions is not straightforward and, in contrast to the one-dimensional case, there exists in general no analytical expression. Instead, these medians have to be computed as solutions of certain minimization problems. The literature on this topic can, for example, be found in [1–4] and the references therein. Moreover, theoretical connections between vector median filters, morphology and PDEs are given in [5]. There exist various applications of multidimensional medians. Indeed, our interest in this topic comes from a dithering algorithm in [6] where a generalized vector median in $\mathbb{R}^2$ has to be computed in an intermediate step of the algorithm. We use the notation 'generalized' because, in contrast to the usual median, there appears an additional squared ℓ_2 term in the functional to be minimized.

In [7], a concept for matrix median computation was proposed and further extended in [8]. The authors suggest to apply semi-definite programming and second order cone programming (SOCP) to find the sought minimizers. Matrix medians of special rank-1 matrices are of interest in connection with the so-called structure tensor of Förstner and Gülch [9], which can be used to approximate image directions of constant gray values like at a straight edge. Recently, SL(2) invariant shape medians were considered in [10].

In this paper, we present a collection of theoretical results on vector and matrix medians. In particular, we investigate matrix medians for different unitarily invariant matrix norms for the case of symmetric 2×2 matrices and show relations

* Corresponding author.
E-mail addresses: setzer@mia.uni-saarland.de (S. Setzer), steidl@mathematik.uni-kl.de (G. Steidl), tteuber@mathematik.uni-kl.de (T. Teuber).

0377-0427/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.cam.2011.09.042

between certain vector and matrix problems. The findings are illustrated by numerical examples and compared to the results of the classical structure tensor.

Beyond that, we propose several algorithms to solve the involved minimization problems. As a first approach, we consider the alternating direction methods of multipliers (ADMM) for the generalized vector as well as matrix median computation. We introduce the algorithms for both problems systematically starting with the vector median computation and use for the matrix median computation a relation between the proximum with respect to a unitarily invariant matrix norm and the proximum with respect to its related gauge function. Next, we apply a relative of the ADMM algorithm, namely the parallel proximal algorithm (PPXA) from [11], which appears to be indeed slightly faster than ADMM. Besides, we briefly introduce second order cone programming (SOCP) and Weiszfeld's algorithm for our generalized ℓ_2 vector median problem. Next, we give a comparison of the computation times required by the different algorithms. Although ADMM and PPXA are slower than Weiszfeld's algorithm, it should be noted that ADMM and PPXA can be parallelized to a high degree so that we expect a significant speed-up for a parallel implementation, e.g., on a GPU.

The paper is organized as follows. In Section 2, we recall special proximation problems with vector norms and unitarily invariant matrix norms, which we need for our median computations. Then, Section 3 deals with vector median computations. After collecting a number of theoretical results, we propose the above mentioned algorithms for the ℓ_p vector median computation, namely ADMM and PPXA. Furthermore, SOCP and Weiszfeld's algorithm are applied for the ℓ_2 vector norm. The proof of the convergence of Weiszfeld's algorithm for our slightly more general setting is given in the Appendix. In Section 4, we are interested in matrix median computations with respect to different unitarily invariant matrix norms. In particular, we deduce the ADMM algorithm and PPXA for matrix median computations in Section 4.1. Then, in Section 4.2, we prove several relations for the matrix mean/medians of 2×2 rank-1 matrices $P_i = p_i p_i^{\mathrm{T}}$ and show connections to special vector mean/medians appearing from the vectors p_i. Numerical experiments are reported in Section 5. In Section 5.1, ADMM, PPXA and Weiszfeld's algorithm are compared with respect to their computation time. It appears that SOCP implemented in the optimization toolbox MOSEK cannot compare to these algorithms for our settings. Section 5.2 illustrates the behavior of the matrix/vector mean and medians within local filters and illuminates the results obtained in Section 4.2. The paper ends with conclusions in Section 6.

2. Proximation with vector and matrix norms

In this section, we recall some special proximation problems, which we need for our median computations. First, we are interested in

$$\hat{x} = \operatorname*{argmin}_{x \in \mathbb{R}^d} \left\{ \frac{1}{2} \|f - x\|_2^2 + \lambda \|x\|_p \right\}, \quad 1 \le p \le \infty \tag{1}$$

for a given data vector $f \in \mathbb{R}^d$. The Fenchel conjugates of the ℓ_p-norms in $\mathbb{R}^d$ are given by

$$\|x\|_p^* := \sup_{y \in \mathbb{R}^d} \{\langle y, x \rangle - \|y\|_p\} = \begin{cases} \infty & \text{if } \|x\|_q > 1, \\ 0 & \text{if } \|x\|_q \le 1, \end{cases} \tag{2}$$

where $\frac{1}{p} + \frac{1}{q} = 1$ and as usual $p = 1$ corresponds to $q = \infty$ and conversely. Now the minimizer can be found by $\hat{x} = f - \hat{v}$, where $\hat{v}$ is the solution of the dual problem

$$\hat{v} = \operatorname*{argmin}_{v \in \mathbb{R}^d} \left\{ \frac{1}{2} \|f - v\|_2^2 + \lambda \|v/\lambda\|_p^* \right\}.$$

By (2), this can be rewritten as the constrained problem

$$\|f - v\|_2 \to \min_v \quad \text{subject to } \|v\|_q \le \lambda.$$

Hence, $\hat{v} = \Pi_{B_{q,\lambda}}(f)$ is the orthogonal projection of f onto the ℓ_q-ball $B_{q,\lambda}$ with radius λ and center 0 and

$$\hat{x} = f - \Pi_{B_{q,\lambda}}(f).$$

For $p = 1, 2, \infty$, the orthogonal projections onto $B_{q,\lambda}$ are given by

$$\mathbf{p = 1}: \quad \Pi_{B_{\infty,\lambda}}(f) = (P_\lambda(f_i))_{i=1}^d \quad \text{and} \quad \hat{x} = (S_\lambda(f_i))_{i=1}^d, \quad \text{where}$$

$$P_\lambda(f_i) := \begin{cases} f_i & \text{if } |f_i| \le \lambda, \\ \lambda \operatorname{sgn}(f_i) & \text{if } |f_i| > \lambda, \end{cases} \quad \text{and} \quad S_\lambda(f_i) = \begin{cases} 0 & \text{if } |f_i| \le \lambda, \\ f_i - \lambda \operatorname{sgn}(f_i) & \text{if } |f_i| > \lambda. \end{cases}$$

The function S_λ is known as *soft-shrinkage*; see [12].

$$\mathbf{p = 2}: \quad \Pi_{B_{2,\lambda}}(f) = \begin{cases} f & \text{if } \|f\|_2 \le \lambda, \\ \lambda \dfrac{f}{\|f\|_2} & \text{if } \|f\|_2 > \lambda \end{cases} \quad \text{and} \quad \hat{x} = \begin{cases} 0 & \text{if } \|f\|_2 \le \lambda, \\ f\left(1 - \dfrac{\lambda}{\|f\|_2}\right) & \text{if } \|f\|_2 > \lambda. \end{cases}$$