Journal of Computational and Applied Mathematics 256 (2014) 278–292

Contents lists available at ScienceDirect

Journal of Computational and Applied Mathematics

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

Software for weighted structured low-rank approximation

Ivan Markovsky *, Konstantin Usevich

Department ELEC, Vrije Universiteit Brussel, Belgium

ARTICLE INFO

Article history:
Received 28 June 2012
Received in revised form 26 February 2013

Keywords:
Mosaic Hankel matrix
Low-rank approximation
Total least squares
System identification
Deconvolution
Variable projection

ABSTRACT

A software package is presented that computes locally optimal solutions to low-rank approximation problems with the following features:

- *mosaic Hankel structure* constraint on the approximating matrix,
- *weighted* 2*-norm* approximation criterion,
- *fixed elements* in the approximating matrix,
- *missing elements* in the data matrix, and
- *linear constraints* on an approximating matrix's left kernel basis.

It implements a variable projection type algorithm and allows the user to choose standard local optimization methods for the solution of the parameter optimization problem. For an $m \times n$ data matrix, with $n > m$, the computational complexity of the cost function and derivative evaluation is $O(m^2 n)$. The package is suitable for applications with $n \gg m$. In statistical estimation and data modeling – the main application areas of the package – $n \gg m$ corresponds to modeling of large amount of data by a low-complexity model. Performance results on benchmark system identification problems from the database DAISY and approximate common divisor problems are presented.

1. Introduction

Structured low-rank approximation is defined as low-rank approximation

$$\text{minimize} \quad \text{over } \widehat{D} \ \|D - \widehat{D}\| \quad \text{subject to} \quad \operatorname{rank}(\widehat{D}) \le r$$

with the additional constraint that the approximating matrix $\widehat{D}$ has the same structure as the data matrix D. A typical example where a rank deficient structured matrix arises is when a sequence $p = (p_1, \ldots, p_{n_p})$ satisfies a difference equation with lag $\ell < \lceil n_p/2 \rceil$, *i.e.*,

$$R_0 p_t + R_1 p_{t+1} + \cdots + R_\ell p_{t+\ell} = 0, \quad \text{for } t = 1, \ldots, n_p - \ell. \tag{DE}$$

The system of Eqs. (DE) is linear in the vector of parameters $R := \begin{bmatrix} R_0 & R_1 & \cdots & R_\ell \end{bmatrix}$, so that it can be written as $R\mathscr{H}_{\ell+1,n_p-\ell}(p) = 0$, where $\mathscr{H}_{\ell+1,n_p-\ell}(p)$ is a Hankel matrix constructed from p. This shows that, for $R \neq 0$, the fact that p satisfies a difference equation (DE) is equivalent to rank deficiency of a Hankel matrix $\mathscr{H}_{\ell+1,n_p-\ell}(p)$.

Many problems in machine learning, system theory, signal processing, and computer algebra can be posed and solved as a structured low-rank approximation problem for different types of structures and different approximation criteria (see Section 5 and [1,2]). In identification and model reduction of linear time-invariant dynamical systems, the structure is

* Corresponding author.
E-mail addresses: imarkovs@vub.ac.be (I. Markovsky), kusevich@vub.ac.be (K. Usevich).

0377-0427/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.cam.2013.07.048

Table 1
Comparison of the old and new versions of the software.

Feature	Old version	New version
Matrix structure	Block-Hankel/Toeplitz	Matrix × mosaic-Hankel
Cost function	2-norm	Weighted 2-norm
Exact data	Whole blocks	Arbitrary elements
Missing data	Not allowed	Arbitrary elements
Constraints on the optimization variable	Unconstrained	Linear constraints
Interface	Matlab	Matlab, Octave, R

block-Hankel. In the computation of approximate greatest common divisor of two polynomials, the structure is Sylvester. In machine learning, the data matrix is often unstructured but the approximation criterion is a weighted 2-norm (or semi-norm, in the case of missing data).

Despite the academic popularity and numerous applications of the structured low-rank approximation problem, the only efficient publicly available software package for structured low-rank approximation is the one of [3]. The package presented in this paper is a significantly extended version of the software in [3]. The main extensions are summarized in Table 1 and are described in more details in Section 2.

The paper is organized as follows. Section 3 defines the considered weighted structured low-rank approximation and presents Matlab/Octave and R interfaces for calling the underlying C++ solver. Section 4 gives details about the solution method for solving the resulting parameter optimization problem. Implementation and software design issues are extracted in Appendix A. Section 5 lists applications of the package and describes in more details an application for solving scalar autonomous linear time-invariant identification and approximate common divisor problems. Appendix B lists extra options of the software for choosing the optimization method.

2. Main features of the software

1. *Matrix structure specification*
   The software package supports matrix structures of the form

$$\mathscr{S}(p) := \Phi \mathscr{H}_{\mathbf{m},\mathbf{n}}(p), \tag{$\mathscr{S}$}$$

where Φ is a full row rank matrix and $\mathscr{H}_{\mathbf{m},\mathbf{n}}$ is a *mosaic Hankel* structure [4], *i.e.*, a $q \times N$ block matrix

$$\mathscr{H}_{[m_1 \ \cdots \ m_q],[n_1 \ \cdots \ n_N]}(p) = \begin{bmatrix} \mathscr{H}_{m_1,n_1}(p^{(11)}) & \cdots & \mathscr{H}_{m_1,n_N}(p^{(1N)}) \\ \vdots & & \vdots \\ \mathscr{H}_{m_q,n_1}(p^{(q1)}) & \cdots & \mathscr{H}_{m_q,n_N}(p^{(qN)}) \end{bmatrix}, \tag{$\mathscr{H}_{\mathbf{m},\mathbf{n}}$}$$

with scalar *Hankel* blocks

$$\mathscr{H}_{m,n}(p) := \begin{bmatrix} p_1 & p_2 & p_3 & \cdots & p_n \\ p_2 & p_3 & \iddots & & p_{n+1} \\ p_3 & \iddots & & & \vdots \\ \vdots & & & & \\ p_m & p_{m+1} & \cdots & & p_{m+n-1} \end{bmatrix} \in \mathbb{R}^{m\times n}. \tag{$\mathscr{H}_{m,n}$}$$

($\mathscr{H}_{\mathbf{m},\mathbf{n}}$) is more general than the block-Hankel and "flexible structure specification", used in the old version of the software (see Section 3 in [3]). In fact, the "flexible structure specification" is equivalent to $\mathscr{H}_{\mathbf{m},\mathbf{n}}$ with equal n_i's. Mosaic Hankel matrices with blocks of different column dimension allow us to solve, for example, system identification problems with multiple trajectories of different lengths [5].

The matrix Φ further extends the class of mosaic Hankel matrices to (mosaic) Hankel-like matrices. A trivial example is the Toeplitz structure achieved by

$$\Phi = J_m := \begin{bmatrix} & & 1 \\ & \iddots & \\ 1 & & \end{bmatrix} \in \mathbb{R}^{m\times m} \quad \text{and} \quad \mathscr{H}_{\mathbf{m},\mathbf{n}}(p) = \mathscr{H}_{m,n}(p).$$

(Empty spaces in a matrix denote zeros.) A more interesting example is the *Toeplitz-plus-Hankel* structure, achieved by

$$\Phi = \begin{bmatrix} I_m & J_m \end{bmatrix} = \begin{bmatrix} 1 & & & & & 1 \\ & \ddots & & & \iddots & \\ & & 1 & 1 & & \end{bmatrix} \in \mathbb{R}^{m\times 2m} \quad \text{and} \quad \mathscr{H}_{\mathbf{m},\mathbf{n}}(p) = \begin{bmatrix} \mathscr{H}_{m,n}(p^{(1)}) \\ \mathscr{H}_{m,n}(p^{(2)}) \end{bmatrix}. \tag{$\mathscr{T}+\mathscr{H}$}$$