SOCN course
"Low-rank approximation and its applications"

Ivan Markovsky
Vrije Universiteit Brussel

Spring 2014

Abstract

Established data modeling approaches are often derived in a stochastic setting. An alternative deterministic approximation approach, known in the systems and control literature as the behavioral approach, has been developed since the 80's by Jan C. Willems and co-workers. The behavioral approach differentiates between the abstract notion of a model and the concrete notion of a model representation. This distinction proves to be important for developing a coherent theory and effective algorithms for system identification, analysis, and control. The course presents a behavioral approach to system identification.

The highlight of the course is the low-rank approximation problem, which is a practical tool for modeling in the behavioral setting. A matrix constructed from the data being rank deficient implies that there is an exact low complexity linear model for that data. Moreover, the rank of the data matrix corresponds to the complexity of the model. In the generic case when an exact low-complexity model does not exist, the aim is to find a model that fits the data approximately. The corresponding computational problem is low-rank approximation. In the case of linear time-invariant dynamical models, the data matrix is, in addition, Hankel structured and the approximation should have the same structure.

Once the approximate system identification problem is formulated as a low-rank approximation problem, it is solved by generic methods. Except for a few special cases, however, low-rank approximation problems are nonconvex and a global solution is expensive to compute. In the course, we present methods based on local optimization, which lead to fast and effective algorithms. The cost function evaluation has the system theoretic interpretation of Kalman smoothing.

In addition to the theory and algorithms for exact and approximate system identification, the course presents examples from system theory (model reduction and distance to uncontrollability), computer algebra (approximate common divisor computation), and machine learning (recommender systems). Software implementation of the developed methods makes the theory applicable in practice. An essential part of the course are exercises, which give hands-on experience with the presented theory and methods.

Keywords: linear algebra, system theory, system identification, numerical algorithms

The course is based on the following references:

- I. Markovsky, Low rank approximation: Algorithms, implementation, applications, Springer, 2012.

- I. Markovsky, et al. Exact and approximate modeling of linear systems: A behavioral approach, SIAM, 2006.

- I. Markovsky and K. Usevich, Software for weighted structured low-rank approximation, preprint, 2012.

Overview

1. *Introduction:* The classical paradigm for data modeling invariably assumes that an input/output partitioning of the data is a priori given. For linear models, this paradigm leads to computational problems of solving approximately overdetermined systems of linear equations. Examples of most simple data fitting problems, however, suggest that the a priori fixed input/output partitioning of the data may be inadequate: 1) the fitting criteria often depend implicitly on the choice of the input and output variables, which may be arbitrary, and 2) the resulting computational problems may be ill-conditioned. An alternative paradigm for data modeling, sometimes refered to as the behavioral paradigm, does not assume a priori fixed input/output partitioning of the data. The corresponding computational problems involve approximation of a matrix constructed from the data by another matrix of lower rank. We will review applications in systems and control, signal processing, computer algebra, chemometrics, psychometrics, machine learning, and computer vision that lead to low-rank approximation problems. Finally, generic solution methods for solving low-rank approximation problems are outlined.

   (a) Classical and behavioral paradigms for data modeling
   (b) Motivating example for low-rank approximation
   (c) Overview of applications
   (d) Overview of algorithms

2. *The most powerful unfalsified model:* Three linear model representations—kernel, image, and input/output—are defined and the connections among them are shown. The problem of obtaining a state space representation of the system is known as system realization. The transition from impulse response to state space representation is a special exact identification problem and is at the heart of the subspace identification methods. Then the problem of deriving a representation of the system from a general exact trajectory of that system is considered. A key condition for existence of a unique solution is persistency of excitation of an input component of the data. The result leads to algorithms for exact identification. First, we present algorithms for passing from data to an impulse response representation. Then, we show how a state sequence of the model can be obtained directly from the data.

   (a) Representations of linear time-invariant systems
   (b) Exact system identification
   (c) The realization problem
   (d) Subspace identification algorithms

3. *Approximate linear time-invariant system identification:* Approximate data modeling is a trade-off between model complexity and model accuracy. One particular scalarization of this biobjective problem is the low-rank approximation problem. Fundamentally different ways of measuring the model accuracy are the misfit and latency. Misfit leads to errors-in-variables problems and latency leads to the prediction error problems. Independent of the particular problem formulation, there are three conceptually different approaches for solving the problem: convex relaxations, local optimization, and global optimization. We present a method based on local optimization. A bilinear structure of the problem is effectively used to eliminate part of the optimization variables. The computation of the cost function has system theoretic interpretation of Kalman smoothing. The remaining problem is a nonlinear least squares optimization over the model parameters.

   (a) Complexity-accuracy trade-off
   (b) Misfit vs latency
   (c) Misfit computation
   (d) Optimization algorithms

4. *Data modeling with missing data:* If time permits, we will present generalization of the local optimization approach, discussed in lecture 3, for approximation of missing data. Missing data are taken into account in the low-rank approximation setting by introducing a weighted cost function with zero weights associating to the missing values. The resulting misfit evaluation problem is a singular linear least norm problem. The obtained algorithms are illustrated on matrix completion, system identification, and data-driven simulation problems.

q 1. Generalized least-norm problems

- Exact and approximate matrix completion
- System identification with missing data
- Data-driven simulation and control

Short biography

Ivan Markovsky obtained MS degree in Control and Systems Engineering from the Technical University of Sofia in July 1998 and Ph.D. degree in Electrical Engineering from the Katholieke Universiteit Leuven in February 2005. From January 2007 to September 2012 he was a lecturer at the School of Electronics and Computer Science of the University of Southampton. Since October 2012 he is with the department ELEC of the Vrije Universiteit Brussel. His current research interests are structured low-rank approximation, system identification, and data-driven control.

For more information, visit his homepage.