**Correlative sparsity structure of polynomial optimization problems in system identification**

Sándor Kolumbán, István Vajk

Budapest University of Technology and Economics  
Department of Automation and Applied Informatics  
{kolumban, vajk}@aut.bme.hu

**Abstract:** Every identification problem where we seek values of some parameters can be regarded as an optimization problem. In the case of linear time invariant systems identification this is a polynomial optimization problem (POP), meaning that the objective function as well as the constraints are given in terms of polynomials of the decision variables. It has been shown that the global solution to this type of problems can be approximated or even obtained by solving a sequence of semidefinite programs (SDP) with increasing size. The size of these SDPs is combinatorially increasing as we go along the sequence of SDPs. As the size grows so does the computational complexity of the obtaining the solutions to these SDPs. This motivated research to find ways to slow down the growth in size as much as possible. If the polynomials involved in the formulation of the POP have some specific properties then the increase in dimension along the sequence of SDPs can be reduced.

The aim of this paper is to examine the polynomial optimization problems rising in system identification in order to see how the sequence of SDPs related to identification behave. The SDP based solution of POPs is briefly introduced and the POP of the identification problems is presented in detail. The correlative sparsity graph of the POP is determined and it is shown that the graph is chordal. Based on these observations we formulate a sequence of SDPs which grows slower than the general solution. We also evaluate the performance of the method regarding its computational complexity and its ability to approximate the global solution of the optimization problem.

**Introduction:** We concentrate on identification of generalized Box-Jenkins models [1] which are given by the difference equation

|  |  |
| --- | --- |
|  |  |

where , , , and F are polynomials in and is the forward shift operator. The polynomials , , and F are monic. By rearranging (1) we get to the difference equation

|  |  |
| --- | --- |
|  |  |

We will represent the polynomials with their coefficient vectors, for example the polynomial will be represented as. As a slight abuse of notation whenever we write we will think of the coefficient vector of the polynomial . The degree of the polynomials , , , and is denoted with , , , and respectively. The notation will refer to the degree of the product . Using these notations the identification problem is given with the objective function to be minimized (3) and feasibility domain defined by the constraints (4)

|  |  |
| --- | --- |
|  |  |

|  |  |
| --- | --- |
|  |  |

denotes the number of samples, while and are column vectors built from the measurements , as .

**SDP relaxations for POPs:** This section summarizes the principles of SDP relaxations for polynomial optimization problems which are given by

|  |  |
| --- | --- |
|  |  |

where the feasibility domain is defined as

|  |  |
| --- | --- |
|  |  |

Lasserre showed that the global optimum of these problems can be approximated by a sequence of SDPs with growing dimensions [2].

The starting point of Lasserre’s paper is that we do not conduct our search for the optimal point in the space of the decision variables but instead, we search for a probability measure concentrated on that point. In terms of probability measures the optimization problem (5) can be written as follows

|  |  |
| --- | --- |
|  |  |

If we use decision variables then objective function becomes linear in these variables. What needs to be done is that the new optimization variables should correspond to a moment sequence of some probability measure and that measure should be concentrated on . These constraints can be expressed in terms of linear matrix inequalities with the moment matrix and localizing matrices which are defined as

|  |  |
| --- | --- |
|  |  |

and , with .

For any degree where the moment and localizing matrices are built only using moments where the SDP that approximates the global solution of (5) can be given as

|  |  |
| --- | --- |
|  |  |

In certain cases when the polynomials in the problem definition (5) and domain definition (6) have certain properties the matrices involved in are sparse. If this sparsity can be used to decrease the size of the SDPs then higher order problems can be solved. The specific requirements needed to be fulfilled by the polynomials can be found in [3] where the authors show how can the linear matrix inequalities in (9) be replaced with a number of smaller ones. The key to this decomposition is that the set can be covered by the union such that each constraint function is only concerned with variables indexed by one of these sets and this holds for the monomials of the objective function as well. Also, this cover should be found in a way that it has the so called running intersection property

|  |  |
| --- | --- |
|  |  |

The correlative sparsity graph defined in [3] can be described as follows. The nodes of the graph correspond to the variables in the original polynomial optimization problem. Two nodes are connected in the graph if they appear with nonzero exponent in at least one monomial of the objective function together or they appear together in one of the constraint functions . The index sets in the running intersection property can be calculated as the maximal cliques of the chordal extensions of the CSP graph.

**Structure of the identification POP:** The polynomial optimization problem is given with the objective function (3) and constraints (4). Using the notations of the previous sections , and where . The variables of the problem are the unknown elements of the vectors , denoted by and the elements of the unknown noise vector .

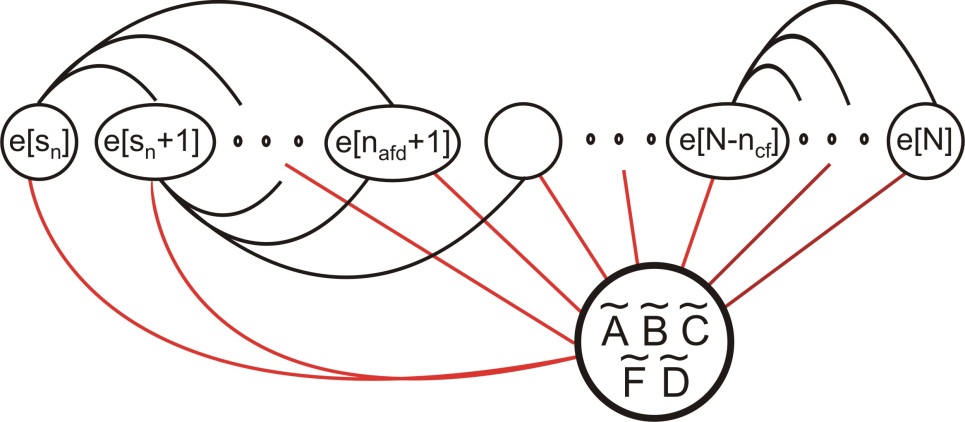


Fig 1. The structure of the correlative sparsity graph corresponding to the identification POP

The correlative sparsity (CSP) graph of the identification problem is given in Fig 1. We proved that this CSP graph is chordal. Since the maximal cliques are needed to form the reduced SRPs we also determined all maximal cliques of the graph. There are maximal cliques in the graph and these are

|  |  |
| --- | --- |
|  |  |

**Numerical results and performance analysis:** We tested the efficiency of the identification algorithm using the SDP sequences. Because the size of the SDPs depends on the sample count, this algorithm can only be used relatively small number of samples. In our case we used 20 samples on a second order system assuming output error noise model. The results were compared to the identification routine oe implemented in Matlab. In order to get a good picture of the results obtained from the two procedures we used 50 different noise realizations.

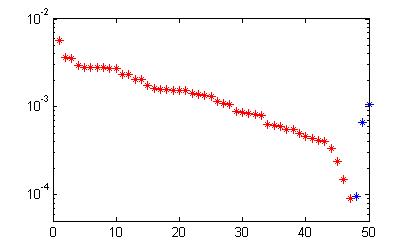
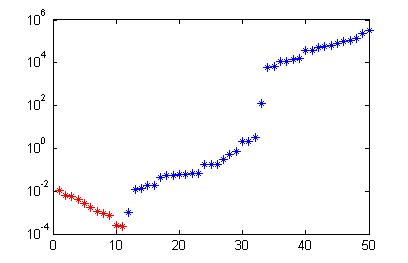


Fig. 2. The difference between the cost function values for the model obtained from the reduced Q2 and that from the Matlab oe routine with standard initialization

Fig. 2. Shows the difference between the objective functions values for models obtained by the Matlab oe routine and that for the solution extracted from the reduced SDPs for . The cases where the SDP based model resulted in a better fit are noted with blue. In cases when the Matlab oe routine performed better are visualized with red. The left picture shows the case when the Matlab routine was used with default initialization. In this case the majority of test cases resulted in oe stopping in a local minimum of the objective function. The right side of Fig.2. shows the case when oe was restarted from the points obtained by the SDP based model. It can be seen that the SDP based method is close to the global optimizer by higher values should be tried. This is not possible due to computational complexity.

**Conclusions:** We investigated the implications of the SDP based relaxation methods given by Lasserre [2] for solving POPs from system identification point of view. The structure of the identification POP was examined in detail. The method was tested on a small scale identification problem, this illustrated that fast convergence in occurs for POPs rising from identification. However, the computational complexity of solving the reduced relaxations of even for small values of is out of reach with the currently available computational power for problems with high degree of complexity or large sample count.

**Acknowledgement:** The work reported in the paper has been developed in the framework of the project „Talent care and cultivation in the scientific workshops of BME" project. This project is supported by the grant ***TÁMOP - 4.2.2.B-10/1--2010-0009***

**References:**

1. Lennart Ljung. System Identification - Theory for the User, 2nd edition, volume 2 of PTR Prentice Hall Information and System Sciences Series. 1999.
2. Jean Bernard Lasserre, Global optimization with polynomials and the problem of moments, SIAM Journal on Optimization, 11 (2001), 796-817.
3. Hayato Waki, Sunyoung Kim, Masakazu Kojima, and Masakazu Muramatsu, Sums of squares and semidenite program relaxations for polynomial optimization problems with structured sparsity, SIAM J. on Optimization, 17 (2006), pp. 218-242.