The Cuckoo Problem: Finding the Globally Optimal Critical Value of a Polynomial Optimization Problem using the Macaulay Matrix

KU LEUVEN
FACULTY OF ENGINEERING SCIENCE

Hans van Rooij and Bart De Moor
KU Leuven - Department of Electrical Engineering (ESAT-STADIUS)

STADIUS
Center for Dynamical Systems, Signal Processing and Data Analytics

Abstract

Multivariate polynomial optimization arises in a variety of engineering areas such as system identification, model-order reduction, control theory, and robotics. For a multivariate polynomial objective function, the first-order necessary conditions for optimality can be represented as a system of multivariate polynomial equations. By identifying the common roots of these conditions, all critical points of the optimization problem can be determined and thus retrieving the optimal solutions that either maximize or minimize a specified cost function. In this poster, we will introduce the cuckoo problem, which involves evaluating a specific multivariate polynomial at the shared roots of a set of other multivariate polynomials. By using the Macaulay matrix, the critical values of the multivariate polynomial optimization problem can directly be computed by solving a generalized eigenvalue problem (GEP), without explicitly obtaining the corresponding critical points.

Multivariate (un)Constrained Polynomial Optimization

The first-order necessary conditions for optimality (KKT conditions) of a multivariate (un)constrained polynomial optimization problem lead to a system of multivariate polynomials:

$$\underset{\boldsymbol{x}}{\text{minimize}}\ J(\boldsymbol{x}),\ \ \boldsymbol{x} \in \mathbb{R}^n,$$
$$\text{subject to}\quad \mathrm{g}_i(\boldsymbol{x}, \boldsymbol{\mu}) = 0,\ \ i = 1, \ldots, n_g,\ \ \boldsymbol{\mu} \in \mathbb{R}^{n_h},$$

where $J(\boldsymbol{x}) \in \mathbb{R}[\boldsymbol{x}]$ is the objective function, the equality constraints are denoted by $\mathbf{g}(\boldsymbol{x}, \boldsymbol{\mu}) = [g_1(\boldsymbol{x}, \boldsymbol{\mu}), \ldots, g_{n_g}(\boldsymbol{x}, \boldsymbol{\mu})]^{\mathrm{T}}$ and $\boldsymbol{\mu}$ are slack variables.

By introducing Lagrange multipliers $\boldsymbol{\lambda} = [\lambda_1, \ldots, \lambda_{n_g}]^{\mathrm{T}} \in \mathbb{R}^{n_g}$, the Langrangian $\mathcal{L}(\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) : \mathbb{R}^{n+n_g+n_h} \to \mathbb{R}$ can be formulated as:

$$\mathcal{L}(\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = J(\boldsymbol{x}) + \boldsymbol{\lambda}^T \mathbf{g}(\boldsymbol{x}, \boldsymbol{\mu}).$$

The first order necessary conditions (FONC) for optimality are:

$$\begin{cases} \nabla_{\boldsymbol{x}} \mathcal{L}(\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) &= 0, \\ \nabla_{\boldsymbol{\lambda}} \mathcal{L}(\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) &= 0, \\ \nabla_{\boldsymbol{\mu}} \mathcal{L}(\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) &= 0. \end{cases} \tag{1}$$

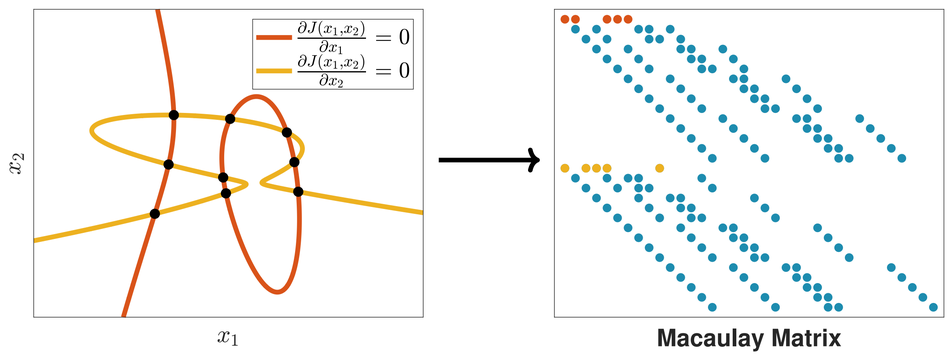
The Macaulay Approach

Consider the following unconstrained optimization problem (Himmelblau's function [1]):

$$\underset{x_1,\, x_2}{\text{minimize}}\ J(x_1, x_2) = (x_1^2 + x_2 - 11)^2 + (x_1 + x_2^2 - 7)^2.$$

The KKT conditions result in the following system of multivariate polynomials:

$$\begin{cases} \frac{\partial J(x_1,x_2)}{\partial x_1} &= 0 \ = 2x_1 + 4x_1(x_1^2 + x_2 - 11) + 2x_2^2 - 14, \\ \frac{\partial J(x_1,x_2)}{\partial x_2} &= 0 \ = 2x_2 + 4x_2(x_2^2 + x_1 - 7) + 2x_1^2 - 22. \end{cases} \tag{2}$$



The Macaulay matrix $\boldsymbol{M}$ is built recursively by performing forward multi-shift recursions. Thus, the system of multivariate polynomial equations, together with their respective monomial shifts, can be represented as follows [2]:

$$\boldsymbol{M}\boldsymbol{\phi} = 0,$$

where $\boldsymbol{\phi}$ represents a multivariate Vandermonde vector:

$$\boldsymbol{\phi} = \begin{bmatrix} 1 & x_1 & x_2 & x_1^2 & x_1x_2 & x_2^2 & x_1^3 & \cdots \end{bmatrix}^T.$$

Every solution of (2) represents one basis vector of the right null space $\boldsymbol{\Phi}$ of $\boldsymbol{M}$. The shift-invariant structure of $\boldsymbol{\Phi}$ can be used to extract all these solutions via a generalized eigenvalue problem (multidimensional realization theory) [3]:

$$\begin{bmatrix} 1 \\ x_1 \\ x_2 \end{bmatrix} x_1 = \begin{bmatrix} x_1 \\ x_1^2 \\ x_1x_2 \end{bmatrix} \ \rightarrow\ \boldsymbol{S}_1\boldsymbol{\Phi}\boldsymbol{D}_{\sigma(\boldsymbol{x})} = \boldsymbol{S}_{\sigma(\boldsymbol{x})}\boldsymbol{\Phi}.$$

We do not know $\boldsymbol{\Phi}$, so instead a numerical basis $\boldsymbol{Z}$ of the null space is computed via an SVD ($\boldsymbol{\Phi} = ZT$). The rank of $\boldsymbol{\Phi}$ (and thus $\boldsymbol{Z}$) stabilises in case of a zero-dimensional solution set.

$$\boldsymbol{S}_1\boldsymbol{Z}\boldsymbol{T}\boldsymbol{D}_{\sigma(\boldsymbol{x})} = \boldsymbol{S}_{\sigma(\boldsymbol{x})}\boldsymbol{Z}\boldsymbol{T} \iff \boldsymbol{A}_{\sigma(\boldsymbol{x})} = \boldsymbol{T}\boldsymbol{D}_{\sigma(\boldsymbol{x})}\boldsymbol{T}^{-1} = (\boldsymbol{S}_1\boldsymbol{Z})^{-1}\boldsymbol{S}_{\sigma(\boldsymbol{x})}\boldsymbol{Z}.$$

The Cuckoo Method

Evaluating a multivariate polynomial ($J(\boldsymbol{x})$) in the common roots of a system of multivariate polynomials (1), without explicitly computing these common roots.

The selection of $\sigma(\boldsymbol{x})$ determines the eigenvalues of $\boldsymbol{A}_{\sigma(\boldsymbol{x})}$. The critical values of $J(\boldsymbol{x})$ can thus be computed in two different ways:

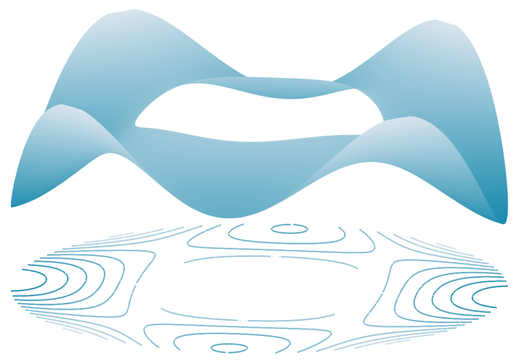
- Shifting with the objective function (CM1):
$$\boldsymbol{A}_{J(\boldsymbol{x})} = (\boldsymbol{S}_1\boldsymbol{Z})^{-1}\boldsymbol{S}_{J(\boldsymbol{x})}\boldsymbol{Z}.$$
- First-degree shift polynomials (CM2): First, all the multiplication matrices are computed:
$$\boldsymbol{A}_{x_i} = (\boldsymbol{S}_1\boldsymbol{Z})^{-1}\boldsymbol{S}_{x_i}\boldsymbol{Z}, \quad i = 1, \ldots, n.$$
Consequently, they are substituted into the objective function:
$$\boldsymbol{A}_{J(\boldsymbol{x})} = J(\boldsymbol{A}_{x_1}, \ldots, \boldsymbol{A}_{x_n}).$$
- The critical values of $J(\boldsymbol{x})$ are computed as the eigenvalues of $\boldsymbol{A}_{J(\boldsymbol{x})}$.

Numerical Example

The cuckoo method will be tested on the following constrained optimization problem:

$$\begin{aligned} \underset{x_1,\, x_2}{\text{minimize}}\ & J(x_1, x_2) = x_1^4 + x_2^4 - 4x_1^2 - 4x_2^2 + x_1x_2 + 15, \\ \text{subject to}\quad & x_1^2 + x_2^2 \leq 4 \ \rightarrow\ x_1^2 + x_2^2 - 4 + \mu_1^2 = 0, \\ & x_1^2 + x_2^2 \geq 1 \ \rightarrow\ x_1^2 + x_2^2 - 1 - \mu_2^2 = 0. \end{aligned}$$

The feasible region of $J(x_1, x_2)$ is:



- Significant time difference between the two cuckoo methods.
- Reduce computational complexity by tackling smaller subproblems.

Method	size $\boldsymbol{M}$	Time
CM1	99008×38760	1157.6 sec.
CM2	26026×12376	65.3 sec.

Critical Value	Multiplicity
4.875	8
5.000	4
8.875	8
9.000	4
11.000	4
11.125	16
11.999	4
12.125	8
15.000	4
15.125	8

References

[1] D. M. Himmelblau, *Applied Nonlinear Programming*. McGraw-Hill, 1972.

[2] P. Dreesen, "Back to the roots - polynomial system solving using linear algebra," Ph.D. dissertation, KU Leuven (Leuven, Belgium), 2013.

[3] C. Vermeersch, "The (block) macaulay matrix: Solving systems of multivariate polynomial equations and multiparameter eigenvalue problems," Ph.D. dissertation, KU Leuven (Leuven, Belgium), 2023.

This work was supported in part by the KU Leuven: Research Fund (projects iBOF/23/064, C3/20/117 and C31-21-00316) Industrial research fund and several Leuven Research Development bilateral industrial projects; in part by Flemish Government agencies: FWO (SBO project S005319, TBM project T001919N, PhD Grand 11K5623N), EWI (Flanders AI Research Program), and VLAIO (CSBO HBC.2021.0076, Baekeland PhD HBC.30192204); and in part by the European Commission: ERC Adv. Grant (grant 885682).