Exploring the Critical Values of Multivariate Polynomial Optimization Problems in System Identification Applications

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

Abstract

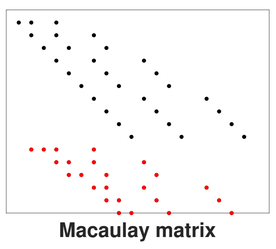
In optimization problems, the objective function is often expressed as a multivariate polynomial. The first-order necessary conditions for optimality can be represented as a system of multivariate polynomial equations. By identifying the common roots of these conditions, all critical points of the optimization problem can be determined. It is important to note that multiple critical points may have the same critical value. Interestingly, the set of critical values always has a Lebesgue measure of zero, as proven by Sard's theorem. In this poster, we show an elegant approach on how to compute the critical values of the objective function, thus finding the global minimum value, without explicitly calculating all the critical points. Once the smallest critical value has been extracted, a linear system can be solved to find its corresponding critical point, avoiding the computation of all other critical points.

The approach presented in this poster tackles this challenge by computing the critical value polynomial (CVP), which resides in the row space of the Macaulay matrix and is a univariate polynomial. The roots of the CVP correspond to the critical values of the objective function. Leveraging the critical value polynomial provides an intriguing way to identify the global optimal solutions in a range of system identification applications involving multivariate polynomial objective functions.

Multivariate polynomial optimization

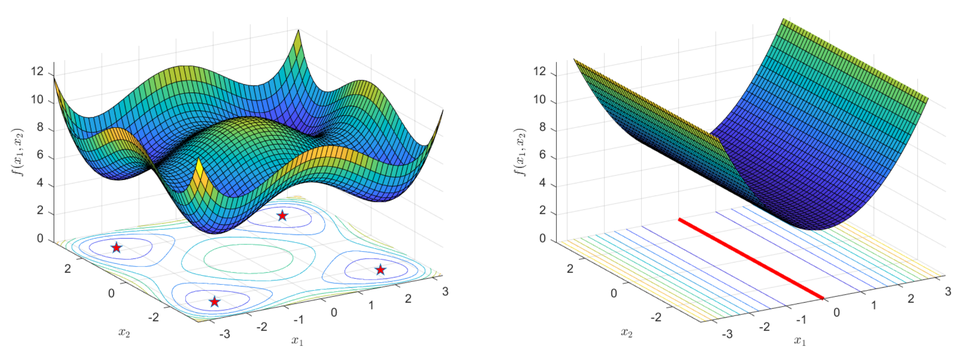
The first-order necessary conditions for optimality (KKT conditions) of a multivariate polynomial objective function $J(\mathbf{x})$ constitute a system of multivariate polynomial equations. We introduce an extra variable σ, which, upon solving Eq. (1), represents the critical values of the (un)constrained optimization problem. The coefficients of Eq. (1) can be encoded in a Macaulay matrix M_σ:

$$\begin{cases} \frac{\partial J(\mathbf{x})}{\partial x_1} &= 0 \\ \frac{\partial J(\mathbf{x})}{\partial x_2} &= 0 \\ \vdots & \\ \frac{\partial J(\mathbf{x})}{\partial x_n} &= 0 \\ \sigma - J(\mathbf{x}) &= 0 \end{cases} \quad (1) \longrightarrow$$



Sard's theorem

Multiple critical points can have the same corresponding critical value. This set of critical points can even be positive dimensional. However, due to Sard's theorem, the set of critical values always has a Lebesgue measure of zero.



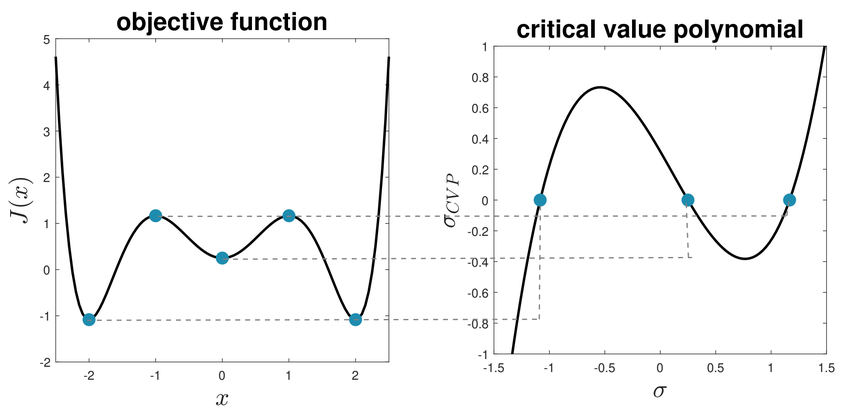
Critical value polynomial (CVP)

The critical value polynomial (CVP) is a univariate polynomial in σ having the critical values of $J(\mathbf{x})$ as its roots. Note that this polynomial is multiplicity-free, meaning that each critical value only appears once as a root of the CVP.

$$\sigma_{CVP} = \prod_{i=1}^{r} (\sigma - \sigma_i).$$

Example: $J(\mathbf{x}) = (2x^6 - 15x^4 + 24x^2)/12 + 1/4$

$$\sigma_{CVP} = (\sigma - 0.25)(\sigma + 1.08333)(\sigma - 1.16667)$$



This work was supported in part by the KU Leuven: Research Fund (C16/15/059, C3/19/053, C24/18/022, C3/20/117, C3I-21-00316, iBOF/23/064), Industrial Research Fund (13-0260, IOFm/16/004, IOFm/20/002), and LRD bilateral industrial projects; in part by Flemish Government agencies: FWO (S005319, I013218N, T001919N), EWI, and VLAIO (HBC.2019.2204, HBC.2021.0076); and in part by the European Commission (ERC Adv. Grant under grant 885682).

Computing the CVP

- The goal is to identify a vector ξ within the row space of M_σ that possesses a predetermined zero pattern: this vector must only exhibit non-zero entries corresponding to pure powers of σ in the multivariate Vandermonde monomial vector ϕ:

$$\phi^T = \begin{pmatrix} 1 & x_1 & x_2 & \cdots & x_n & \sigma & x_1^2 & \cdots & \sigma^2 & \cdots & x_1^d & \cdots & \sigma^d \end{pmatrix},$$

where d is the degree of M_σ.

Example: In case $\mathbf{x} = (x_1, x_2)^T \in \mathbb{R}^2$, the predetermined zero pattern to be found in the row space of M_σ of degree d looks as follows:

$$\begin{array}{c} \\ \xi^T = \end{array} \begin{array}{c} \begin{array}{ccccccccccccc} 1 & x_1 & x_2 & \sigma & x_1^2 & x_1x_2 & x_1\sigma & x_2^2 & x_2\sigma & \sigma^2 & x_1^3 & \cdots & \sigma^d \end{array} \\ \begin{pmatrix} \bullet & 0 & 0 & \bullet & 0 & 0 & 0 & 0 & 0 & \bullet & 0 & \cdots & \bullet \end{pmatrix}. \end{array}$$

where $\bullet$ stands for any arbitrary real number.

- The index set $\mathcal{I}$ contains the indices of the pure powers of sigma in ϕ.
- Define E as the identity matrix only containing the rows that are in $\mathcal{I}$.
- Find CVP in the intersection $\mathrm{row}(M_\sigma) \cap \mathrm{row}(E)$.

Using left singular vectors

The CVP can be found when the following condition holds:

$$\dim(\mathrm{row}(M_\sigma) \cap \mathrm{row}(E)) = \mathit{rank}(M_\sigma) - \mathit{rank}(M_\sigma(:, \mathcal{I}')) > 0$$

- Find vector η in left null space of $M_\sigma(:, \mathcal{I}') \Rightarrow \eta^T M_\sigma(:, \mathcal{I}') = 0$.
- Compute vector with predetermined zero pattern $\Rightarrow \xi^T = \eta^T M_\sigma \neq 0$.
- The CVP is then obtained as: $\sigma_{CVP} = \phi^T \xi$.

Using right singular vectors

The CVP can also be found if at least one of the principal angles between $\mathrm{row}(M_\sigma)$ and $\mathrm{row}(E)$ is zero.

- Orthonormal basis for $\mathrm{row}(M_\sigma)$ via SVD: $M_\sigma = U_{M_{\sigma,1}} \Sigma_{M_{\sigma,1}} V_{M_{\sigma,1}}^T$
- The rows of E form an orthonormal basis for $\mathrm{row}(E)$.
- The principal angles and directions can be extracted by calculating the SVD of $EV_{M_{\sigma,1}} = V_{M_{\sigma,1}}(\mathcal{I}, :) = PSQ^T$.
- The cosines of the principal angles between $\mathrm{row}(M_\sigma)$ and $\mathrm{row}(E)$ are the first $d+1$ singular values of $V_{M_{\sigma,1}}(\mathcal{I}, :)$.
- The principal directions $P_{dir,E}$ in $\mathrm{row}(E)$ are the first $d+1$ columns of $E^T P$.
- If there is an intersection between $\mathrm{row}(M_\sigma)$ and $\mathrm{row}(E)$, the CVP is obtained as: $\sigma_{CVP} = \phi^T P_{dir,E}(:, 1)$.

Computing global optimum

- Find smallest root of σ_{CVP}: σ^*.
- Solve homogeneous linear system of equations using σ^* to find corresponding global minimal critical point.

Applications and open problems

- The CVP finds applications in system identification, including tasks like $\mathcal{H}_2$-norm model order reduction.
- Currently only used on toy problems $\Rightarrow$ further exploit the structure of M_σ to increase performance.