Using the Macaulay Matrix to find the Globally Optimal Critical Value of a Multivariate Polynomial Optimization Problem

Hans van Rooij hans.vanrooij@kuleuven.be
Bart De Moor bart.demoor@kuleuven.be
KU Leuven, ESAT-STADIUS

March 26, 2024

Table of Contents

Table of Contents

Table of Contents

Multivariate Polynomial Optimization

Optimization Problem

$$\min_{\mathbf{x}} J(\mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^n,$$

$$\text{subject to} \quad \mathbf{g}\left(\mathbf{x}\right) = \left[g_1(\mathbf{x}), \ldots, g_{n_g}(\mathbf{x})\right]^T = \mathbf{0},$$

$$J(\mathbf{x}),\ g_i(\mathbf{x}) \in \mathbb{R}[\mathbf{x}]$$

$$\downarrow$$

$$\text{Langrangian: } \mathcal{L}(\mathbf{x}, \Lambda) = J(\mathbf{x}) + \Lambda^T \mathbf{g}(\mathbf{x}),$$

$$\Lambda = [\lambda_1, \ldots, \lambda_{n_g}]^T \in \mathbb{R}^n.$$

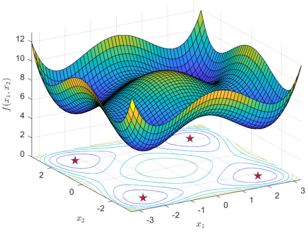
KKT conditions ⇒ System of Multivariate Polynomials

$$\begin{cases} \nabla_{\mathbf{x}} \mathcal{L}(\mathbf{x}, \Lambda) & = 0 = \nabla_{\mathbf{x}} J(\mathbf{x}) + \nabla_{\mathbf{x}} (\Lambda^T \mathbf{g}(\mathbf{x})), \\ \nabla_{\mathbf{\Lambda}} \mathcal{L}(\mathbf{x}, \Lambda) & = 0 = \mathbf{g}(\mathbf{x}). \end{cases}$$

Using the Critical Values

Set of critical points can be positive dimensional.

Sard's theorem [1]:
Set of critical values of a smooth function has a Lebesgue measure of zero.



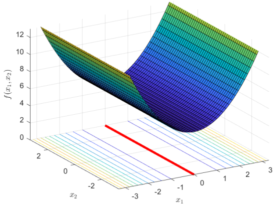


Table of Contents

Macaulay Matrix

Example (Bivariate Objective Function)

$$J(x_1, x_2) = x_1^4 + x_2^2 - 4x_1^2 + x_1 x_2 + 15.$$

$$\downarrow$$

$$\begin{cases} \partial_{x_1} J(x_1, x_2) &= 4x_1^3 - 8x_1 + x_2 &= 0, \\ \partial_{x_2} J(x_1, x_2) &= x_1 + 2x_2 &= 0. \end{cases}$$

$$\mathcal{M}(3) = \begin{array}{c} \\ \partial_{x_1} J(x_1,x_2) \\ \partial_{x_2} J(x_1,x_2) \\ x_1 \partial_{x_2} J(x_1,x_2) \\ x_2 \partial_{x_2} J(x_1,x_2) \\ x_1^2 \partial_{x_2} J(x_1,x_2) \\ x_1 x_2 \partial_{x_2} J(x_1,x_2) \\ x_2^2 \partial_{x_2} J(x_1,x_2) \end{array} \begin{array}{c} \begin{array}{cccccccccc} 1 & x_1 & x_2 & x_1^2 & x_1 x_2 & x_2^2 & x_1^3 & x_1^2 x_2 & x_1 x_2^2 & x_2^3 \end{array} \\ \begin{bmatrix} 0 & -8 & 1 & 0 & 0 & 0 & 4 & 0 & 0 & 0 \\ 0 & 1 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 2 \end{bmatrix} \end{array}.$$

Solving a System of Polynomial Equations as an EVP

First-order necessary conditions

$$\begin{cases} \nabla_{\mathbf{x}}\mathcal{L}(\mathbf{x},\Lambda) & = & 0, \\ \nabla_{\mathbf{\Lambda}}\mathcal{L}(\mathbf{x},\Lambda) & = & 0. \end{cases} \quad (1)$$

$$\downarrow$$

$$\mathcal{M}\phi = 0.$$

Using (backward) shift-invariant null space [2, 3]

$$S_1\Phi D_\sigma = S_\sigma \Phi,$$

$$S_1 Z T D_\sigma = S_\sigma Z T,$$

$$T D_\sigma T^{-1} = (S_1 Z)^{-1} S_\sigma Z.$$

Example (Shift function σ)

- **First order monomial shift**: $\sigma = x_i \in \mathbf{x}, \quad i = 1, \ldots, n.$
- **Shift with objective function**: $\sigma = J(\mathbf{x}).$

Table of Contents

Using Critical Values

Using the Macaulay matrix to obtain the critical values in different ways.

- **Naive approach**: Obtain the critical points and explicitly substitute them in the objective function $J(\mathbf{x})$.

Using Critical Values

Using the Macaulay matrix to obtain the critical values in different ways.

- **Naive approach**: Obtain the critical points and explicitly substitute them in the objective function $J(\mathbf{x})$.
- **Cuckoo method**: Evaluate the objective function $J(\mathbf{x})$ in the set of critical points, without explicitly calculating these critical points. The obtained critical values result from an EVP.

Using Critical Values

Using the Macaulay matrix to obtain the critical values in different ways.

- **Naive approach**: Obtain the critical points and explicitly substitute them in the objective function $J(\mathbf{x})$.
- **Cuckoo method**: Evaluate the objective function $J(\mathbf{x})$ in the set of critical points, without explicitly calculating these critical points. The obtained critical values result from an EVP.
- **Critical Value Polynomial (CVP)**: Add extra variable σ and eliminate all other variables $\mathbf{x}$. The critical values are computed as the roots of a univariate polynomial.

Using Critical Values

Using the Macaulay matrix to obtain the critical values in different ways.

- **Naive approach**: Obtain the critical points and explicitly substitute them in the objective function $J(\mathbf{x})$.
- **Cuckoo method**: Evaluate the objective function $J(\mathbf{x})$ in the set of critical points, without explicitly calculating these critical points. The obtained critical values result from an EVP.
- **Critical Value Polynomial (CVP)**: Add extra variable σ and eliminate all other variables $\mathbf{x}$. The critical values are computed as the roots of a univariate polynomial.

Cuckoo: Objective Function as the Shift Polynomial

Compute multiplication matrix using $J(\mathbf{x})$ as a shift

$$TD_\sigma T^{-1} = (S_1 Z)^{-1} S_\sigma Z, \quad \sigma = J(\mathbf{x}).$$

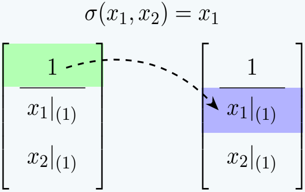
- Denote the degree of $J(\mathbf{x})$ as d_J.
- Furthermore, suppose the maximal degree of the set of standard monomials is d [3].
- Then a null space containing monomials up to degree $d + d_J$ is needed $\rightarrow$ computationally expensive.

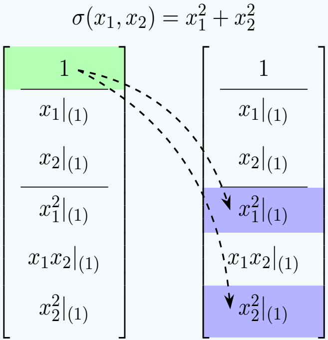
Table of Contents

Cuckoo: Objective Function as the Shift Polynomial

Example (High-Order Shift Operator)

$$J(\mathbf{x}) = x_1^2 + x_2^2 \quad \rightarrow \quad \begin{cases} \frac{\partial J(\mathbf{x})}{\partial x_1} = 0 = 2x_1, \\ \frac{\partial J(\mathbf{x})}{\partial x_2} = 0 = 2x_2. \end{cases}$$





Cuckoo: Objective Function as the Shift Polynomial

Example (Himmelblau's function)

$$\min_{x_1,x_2\in\mathbb{R}} J(x_1,x_2) = (x_1^2 + x_2 - 11)^2 + (x_1 + x_2^2 - 7)^2.$$

$$\downarrow$$

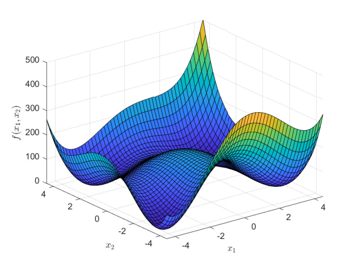
$$\begin{cases} \frac{\partial J(\mathbf{x})}{\partial x_1} = 0 = 4x_1^3 + 4x_1x_2 - 42x_1 + 2x_2^2 - 14, \\ \frac{\partial J(\mathbf{x})}{\partial x_2} = 0 = 2x_1^2 + 4x_1x_2 + 4x_2^3 - 26x_2 - 22. \end{cases}$$

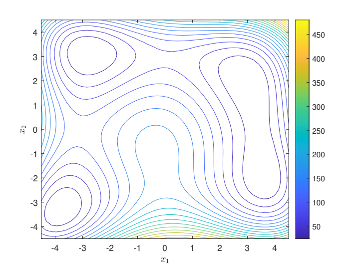
d	size $\mathcal{M}$	nullity $\mathcal{M}$	linearly independent monomials
3	2×10	8	$1,\ x_1,\ x_2,\ x_1^2,\ x_1x_2,\ x_2^2,\ x_1^2x_2,\ x_1x_2^2$
4	6×15	9	$1,\ x_1,\ x_2,\ x_1^2,\ x_1x_2,\ x_2^2,\ x_1^2x_2,\ x_1x_2^2,\ x_1^2x_2^2$
5	12×21	9	$1,\ x_1,\ x_2,\ x_1^2,\ x_1x_2,\ x_2^2,\ x_1^2x_2,\ x_1x_2^2,\ x_1^2x_2^2$
6	20×28	9	$1,\ x_1,\ x_2,\ x_1^2,\ x_1x_2,\ x_2^2,\ x_1^2x_2,\ x_1x_2^2,\ x_1^2x_2^2$
7	30×36	9	$1,\ x_1,\ x_2,\ x_1^2,\ x_1x_2,\ x_2^2,\ x_1^2x_2,\ x_1x_2^2,\ x_1^2x_2^2$
8	42×45	9	$1,\ x_1,\ x_2,\ x_1^2,\ x_1x_2,\ x_2^2,\ x_1^2x_2,\ x_1x_2^2,\ x_1^2x_2^2$

$\Rightarrow$ Compute eigenvalues of: $(S_1 Z(8))^{-1} S_{J(x_1,x_2)} Z(8)$.

Cuckoo: Objective Function as the Shift Polynomial

Example (Continued: Himmelblau's function)





critical value	multiplicity	classification
181.6165	1	local maximum
178.3372	1	saddle point
104.0152	1	saddle point
67.7192	1	saddle point
13.3119	1	saddle point
0	4	global minimum

Table of Contents

Cuckoo: Divide and Conquer (First Degree Shift Polynomials)

Compute multiplication matrices for all first order degree shifts

$$A_\sigma = T D_\sigma T^{-1} = (S_1 Z)^{-1} S_\sigma Z, \quad \sigma = x_i \in \mathbf{x}, \quad \forall i = 1, \ldots, n.$$

- The multiplication matrices commute.
- Substitute in objective function $J(\mathbf{x})$ to obtain a multiplication matrix objective function.
- The eigenvalues of this matrix correspond to the critical values of $J(\mathbf{x})$.
- No need to deal with high order monomial shifts!

Cuckoo: Divide and Conquer (First Degree Shift Polynomials)

Example (Himmelblau's function revisited)

$$\min_{x_1,x_2\in\mathbb{R}} J(x_1,x_2) = (x_1^2+x_2-11)^2+(x_1+x_2^2-7)^2.$$

$$\downarrow$$

$$\begin{cases} \frac{\partial J(\mathbf{x})}{\partial x_1} & = & 0 & = 4x_1^3+4x_1x_2-42x_1+2x_2^2-14, \\ \frac{\partial J(\mathbf{x})}{\partial x_2} & = & 0 & = 2x_1^2+4x_1x_2+4x_2^3-26x_2-22. \end{cases}$$

- Divide:

$$A_{x_1} = (S_1 Z(5))^{-1} S_{x_1} Z(5),$$
$$A_{x_2} = (S_1 Z(5))^{-1} S_{x_2} Z(5).$$

- Conquer:

$$J(A_{x_1}, A_{x_2}) = (A_{x_1}^2 + A_{x_2} - 11I_9)^2 + (A_{x_1} + A_{x_2}^2 - 7I_9)^2.$$

Table of Contents

STADIUS
Center for Dynamical Systems, Signal Processing and Data Analytics

Critical Value Polynomial (CVP)

Optimization Problem

$$\min_{\mathbf{x}} J(\mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^n,$$

$$\text{subject to} \quad \mathbf{g}\left(\mathbf{x}\right) = \left[g_1(\mathbf{x}), \dots, g_{n_g}(\mathbf{x})\right]^T = \mathbf{0},$$

$$\downarrow$$

$$\text{Langrangian: } \mathcal{L}(\mathbf{x}, \Lambda) = J(\mathbf{x}) + \Lambda^T \mathbf{g}(\mathbf{x}),$$

$$\Lambda = [\lambda_1, \dots, \lambda_{n_g}]^T \in \mathbb{R}^n.$$

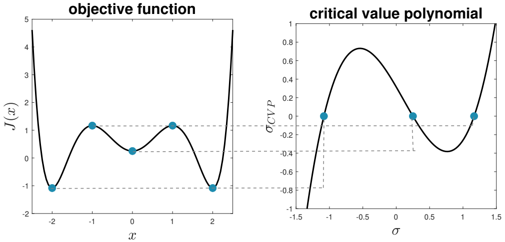
KKT conditions ⇒ System of Multivariate Polynomials

$$\begin{cases} \nabla_{\mathbf{x}} \mathcal{L}(\mathbf{x}, \Lambda) & = 0 = \nabla_{\mathbf{x}} J(\mathbf{x}) + \nabla_{\mathbf{x}} \mathbf{g}(\mathbf{x})^T \Lambda, \\ \nabla_{\mathbf{\Lambda}} \mathcal{L}(\mathbf{x}, \Lambda) & = 0 = \mathbf{g}(\mathbf{x}), \\ J(\mathbf{x}) - \sigma & = 0. \end{cases} \tag{2}$$

Critical Value Polynomial (CVP)

CVP: Univariate polynomial in σ having the critical values of $J(\mathbf{x})$ as its roots

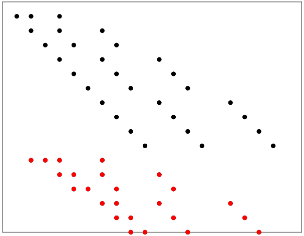
$$\sigma_{CVP} = \prod_{i=1}^{r} (\sigma - \sigma_i).$$



Calculating the CVP

Using Linear Algebra

- Use Macaulay Matrix $\mathcal{M}$ of 2 [4, 5].
- Find vector in row space of $\mathcal{M}$ with pre-specified zero pattern (e. g. using left or right singular vectors).



Macaulay matrix

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

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

Critical Value Polynomial (CVP)

Sufficient Condition to Find the CVP: Graßmann's dimension theorem

- 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 $\mathcal{I} \rightarrow E = I_n(\mathcal{I}, :)$.
- Find CVP in the intersection $\mathrm{row}(\mathcal{M}) \cap \mathrm{row}(E)$.
- Intersection if and only if $\dim(\mathrm{row}(\mathcal{M}) \cap \mathrm{row}(E)) > 0$.

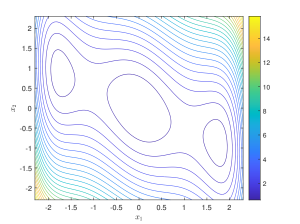
Critical Value Polynomial (CVP)

Example (Three-hump camel function)

$$\min_{x_1,x_2\in\mathbb{R}} J(x_1,x_2) = 2x_1^2 - 1.05x_1^4 + \frac{x_1^6}{6} + x_1x_2 + x_2^2.$$

$$\downarrow$$

$$\begin{cases} \frac{\partial J(\mathbf{x})}{\partial x_1} & = & 0 & = x_1^5 - \frac{21}{5}x_1^3 + 4x_1 + x_2, \\ \frac{\partial J(\mathbf{x})}{\partial x_2} & = & 0 & = x_1 + 2x_2, \\ \sigma - J(x_1,x_2) & = & 0 & = \sigma - (2x_1^2 - 1.05x_1^4 + \frac{x_1^6}{6} + x_1x_2 + x_2^2). \end{cases}$$



STADIUS

Critical Value Polynomial (CVP)

Example (Continued Three-hump camel function)

d	size $\mathcal{M}$	nullity $\mathcal{M}$
6	61×84	24
7	98×120	27
8	150×165	29
9	220×220	30
10	311×286	30
11	426×364	30

d	$\dim\left(\text{row}(\mathcal{M}) \cap \text{row}(E)\right)$
9	0
10	0
11	1
12	2
13	3
14	4

σ_{CVP}	roots (critical values)
$\sigma^3 - 1.17600\sigma^2 + 0.26201\sigma$	{0, 0.29863, 0.87736}

Possible Problems using the CVP

- An additional variable σ is needed.
- A small perturbation in the coefficients can drastically influence the roots of the CVP!

Example (Wilkinson's Polynomial)

$$\prod_{i=1}^{20}(\sigma - i) = x^{20} - 210x^{19} + ... + 2432902008176640000.$$

If -210 is perturbed by 2^{-23} to -210.0000001192, then the root at $\sigma = 20$ changes to $\sigma = 20.8$! The roots $\sigma = 18$ and $\sigma = 19$ collide and create a double root at $\sigma \approx 18.62$!

Extract critical point from corresponding critical value

- Find smallest critical value: σ^*.
- Solve homogeneous linear system of equations using σ^* to find corresponding global minimal critical point.

- Homogeneous linear system: $((S_1 Z)^{-1} S_\sigma Z - \sigma^* I)t = 0$.
- Eigenvector containing critical point: $k = Zt$.

Corresponding critical points

$$k = \begin{bmatrix} 1 & x_1^* & x_2^* & \cdots & x_n^* & \sigma^* \cdots & x_1^{*d} & x_1^{*d-1} x_2^* & \cdots & \sigma^{*d} \end{bmatrix}^T$$

Table of Contents

References I

- J. M. Lee, "Sard's theorem," *Introduction to smooth manifolds*, pp. 125–149, 2012.
- K. Batselier, P. Dreesen, and B. De Moor, "On the null spaces of the macaulay matrix," *Linear Algebra and its Applications*, vol. 460, pp. 259–289, 2014.
- P. Dreesen, "Back to the roots - polynomial system solving using linear algebra," Ph.D. dissertation, ESAT-STADIUS, KU Leuven (Leuven, Belgium), 2013.
- F. S. MacAulay, "Some Formulæ in Elimination," *Proceedings of the London Mathematical Society*, vol. s1-35, no. 1, pp. 3–27, 05 1902. [Online]. Available: https://doi.org/10.1112/plms/s1-35.1.3

References II

F. S. Macaulay, *The algebraic theory of modular systems*. Cambridge University Press, 1994, vol. 19.