

Rational Interpolation, the Loewner Framework and the Kolmogorov Superposition Theorem.

Athanasios C. Antoulas, Ion Victor Gosea, Charles Poussot-Vassal

Abstract

Problem #119, in [1], asks the question: *Are there actually functions of three variables?*

Stated differently: is it possible to use compositions of functions of two or fewer variables to express any function of three variables? This question is related to Hilbert's 13th problem: are there any genuine continuous multivariate functions. As a matter of fact, Hilbert conjectured the existence of a three-variable continuous function which cannot be expressed in terms of composition and addition of two-variable continuous functions. For a recent overview of this problem, see [2].

For continuous function, the Kolmogorov Superposition Theorem (KST) answers this question negatively. It shows namely that continued functions of several variables can be expressed as composition and superposition of functions of one variable. Thus, there are no *true* functions of three variables.

The present contribution presents connections between the Loewner framework for rational interpolation of multivariate functions and KST restricted to rational functions. The result is the formulation of KST for the special case of rational functions. As a byproduct *taming of the curse of dimensionality*, both in computational complexity, storage, and last but not least, numerical accuracy, is achieved.

Short summary of the Loewner framework. The Loewner framework is an interpolatory approach designed for approximating linear and nonlinear systems. Reference [3] extends this framework to linear parametric systems with an arbitrary number of parameters, in other words to multivariate functions of n variables. One main innovation established is the construction of data-based system realizations for any number of parameters. Equally importantly, [3] shows how to alleviate the computational burden, storage and numerical accuracy, by avoiding the explicit construction of higher dimensional Loewner matrices of size $N \times N$. Instead, the proposed methodology achieves decoupling of variables, leading to (i) a complexity reduction from $\mathcal{O}(N^3)$ to below than $\mathcal{O}(N^{1.5})$ when $N > 5$ and (ii) to memory storage bounded by the largest variable dimension rather than the product of all variable dimensions, thus taming the curse of dimensionality and making the solution scalable to large data sets.

After defining a new multivariate realization, we introduce the higher dimensional multivariate Loewner matrices and show that they can be computed by solving a coupled set of Sylvester equations. The *null space* of these Loewner matrices then allows the computation of the multivariate *barycentric weights* of the associated rational function. One of the main results of [3] is to show how the null space of N -dimensional Loewner matrices can be computed using a sequence of 1-dimensional Loewner matrices. This leads to a drastic computational burden reduction. This also leads to the formulation of KST for rational functions. Finally, two algorithms are proposed (one direct and one iterative) to construct, directly from data, multivariate (or parametric) realizations ensuring (approximate) interpolation. For details on the above material see [3].

The purpose of this contribution is to make contact of the above results with the Kolmogorov Superposition Theorem. For clarity of exposition we will illustrate the main features of our approach by means of a three-variable example.

Example. Consider the three-variable function $\mathbf{H}(s, t, x) = \frac{s^2 + xs + 1}{t + x + st + 2}$. Since the degrees in each variable are (2, 1, 1), we will need the integers $\nu_1 = 3$, $\nu_2 = 2$, and $\nu_3 = 2$. This implies that $N = \nu_1 \nu_2 \nu_3 = 12$. The right and left interpolation points are

$$\begin{aligned} s_1 = 1, s_2 = 2, s_3 = 3, \quad t_1 = 4, t_2 = 5, \quad x_1 = 6, x_2 = 7, \quad \text{and} \\ s_4 = 3/2, s_5 = 5/2, s_6 = 7/2, \quad t_3 = 9/5, t_4 = 11/5, \quad x_3 = 13/3, x_4 = 5, \quad \text{respectively.} \end{aligned}$$

Following the theory in [3], the right triples of interpolation points are $\mathbf{S} = [\mathbf{s}_1, \mathbf{s}_2, \mathbf{s}_3] \otimes \mathbb{I}_{1,2} \otimes \mathbb{I}_{1,2}$, $\mathbf{T} = \mathbb{I}_{1,3} \otimes [\mathbf{t}_1, \mathbf{t}_2] \otimes \mathbb{I}_{1,2}$, $\mathbf{X} = \mathbb{I}_{1,3} \otimes \mathbb{I}_{1,2} \otimes [\mathbf{x}_1, \mathbf{x}_2] \in \mathbf{C}^{1 \times N}$. Thus the resulting 3D-Loewner matrix has dimension $N \times N$ and the barycentric weights are

$$\mathbf{Bary} = \left[\frac{16}{29} \quad -\frac{17}{29} \quad -\frac{18}{29} \quad \frac{19}{29} \quad -\frac{40}{29} \quad \frac{42}{29} \quad \frac{46}{29} \quad -\frac{48}{29} \quad \frac{24}{29} \quad -\frac{25}{29} \quad -\frac{28}{29} \quad 1 \right]^T.$$

Again, the theory in [3], allows the demposition of this vector in a (pointwise) product of barycentric weights with respect to each variable, separately. Thus *decoupling* the problem is achieved, one of the important aspects of KST; in [3] we obtain:

$$\mathbf{Bary} = \mathbf{Bary}_x \odot \mathbf{Bary}_t \odot \mathbf{Bary}_s,$$

where \odot denotes the pointwise product. This is a special case of formula (5.5) in [3].

This is the *key result* which allows the connection with KST and taming the curse of dimensionality. We have thus shown that the 3D multivariate function can be computed in terms of three 1D functions (one in each variable). These functions are denoted below by $\Phi(x)$, $\Psi(t)$ and $\Omega(s)$. Furthermore \mathbf{Lag}_x , \mathbf{Lag}_t and \mathbf{Lag}_s are the *Lagrange bases components* in each variable. Finally \mathbf{W} are the right interpolation values for the triples in $\mathbf{S} \times \mathbf{T} \times \mathbf{X}$. The ensuing numerical values are as follows:

$$\underbrace{\begin{bmatrix} -\frac{16}{17} \\ 1 \\ -\frac{18}{19} \\ 1 \\ -\frac{20}{21} \\ 1 \\ -\frac{23}{24} \\ 1 \\ -\frac{24}{25} \\ 1 \\ -\frac{28}{29} \\ 1 \end{bmatrix}}_{\mathbf{Bary}_x}, \underbrace{\begin{bmatrix} -\frac{17}{19} \\ -\frac{17}{19} \\ 1 \\ 1 \\ -\frac{7}{8} \\ -\frac{7}{8} \\ 1 \\ 1 \\ -\frac{25}{29} \\ -\frac{25}{29} \\ 1 \\ 1 \end{bmatrix}}_{\mathbf{Bary}_t}, \underbrace{\begin{bmatrix} \frac{19}{29} \\ \frac{19}{29} \\ \frac{19}{29} \\ \frac{19}{29} \\ -\frac{48}{29} \\ -\frac{48}{29} \\ -\frac{48}{29} \\ -\frac{48}{29} \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}}_{\mathbf{Bary}_s}, \underbrace{\begin{bmatrix} \frac{1}{x-6} \\ \frac{1}{x-7} \\ \frac{1}{x-6} \\ \frac{1}{x-7} \\ \frac{1}{x-6} \\ \frac{1}{x-7} \\ \frac{1}{x-6} \\ \frac{1}{x-7} \\ \frac{1}{x-6} \\ \frac{1}{x-7} \\ \frac{1}{x-6} \\ \frac{1}{x-7} \end{bmatrix}}_{\mathbf{Lag}_x}, \underbrace{\begin{bmatrix} \frac{1}{t-4} \\ \frac{1}{t-4} \\ \frac{1}{t-5} \\ \frac{1}{t-5} \\ \frac{1}{t-4} \\ \frac{1}{t-4} \\ \frac{1}{t-5} \\ \frac{1}{t-5} \\ \frac{1}{t-4} \\ \frac{1}{t-4} \\ \frac{1}{t-5} \\ \frac{1}{t-5} \end{bmatrix}}_{\mathbf{Lag}_t}, \underbrace{\begin{bmatrix} \frac{1}{s-1} \\ \frac{1}{s-1} \\ \frac{1}{s-1} \\ \frac{1}{s-1} \\ \frac{1}{s-2} \\ \frac{1}{s-2} \\ \frac{1}{s-2} \\ \frac{1}{s-2} \\ \frac{1}{s-3} \\ \frac{1}{s-3} \\ \frac{1}{s-3} \\ \frac{1}{s-3} \end{bmatrix}}_{\mathbf{Lag}_s}, \underbrace{\begin{bmatrix} \frac{1}{2} \\ \frac{9}{17} \\ \frac{4}{9} \\ \frac{9}{19} \\ \frac{17}{20} \\ \frac{19}{21} \\ \frac{17}{23} \\ \frac{19}{24} \\ \frac{7}{6} \\ \frac{31}{25} \\ 1 \\ \frac{31}{29} \end{bmatrix}}_{\mathbf{W}} \Rightarrow \begin{cases} \Phi(x) = \mathbf{Bary}_x \odot \mathbf{Lag}_x, \\ \Psi(t) = \mathbf{Bary}_t \odot \mathbf{Lag}_t, \\ \Omega(s) = \mathbf{Bary}_s \odot \mathbf{Lag}_s. \end{cases}$$

With the above notation we can express \mathbf{H} as the quotient of two rational functions:

$$\left. \begin{aligned} \hat{\mathbf{n}}(s, t, x) &= \sum_{\text{rows}} [\mathbf{W} \odot \Phi(\mathbf{x}) \odot \Psi(\mathbf{t}) \odot \Omega(\mathbf{s})] \\ \hat{\mathbf{d}}(s, t, x) &= \sum_{\text{rows}} [\Phi(\mathbf{x}) \odot \Psi(\mathbf{t}) \odot \Omega(\mathbf{s})] \end{aligned} \right\} \Rightarrow \frac{\hat{\mathbf{n}}(s, t, x)}{\hat{\mathbf{d}}(s, t, x)} = \mathbf{H}(s, t, x).$$

Consequently, KST for rational functions, as *composition and superposition* of one-variable functions, takes the form:

$$\left. \begin{aligned} \hat{\mathbf{n}}(s, t, x) &= \sum_{\text{rows}} \exp [\log \mathbf{W} + \log \Phi(\mathbf{x}) + \log \Psi(\mathbf{t}) + \log \Omega(\mathbf{s})] \\ \hat{\mathbf{d}}(s, t, x) &= \sum_{\text{rows}} \exp [\log \Phi(\mathbf{x}) + \log \Psi(\mathbf{t}) + \log \Omega(\mathbf{s})] \end{aligned} \right\} \quad (*)$$

Some details of the above computation. To compute (i) \mathbf{Bary}_x , computation of the nullspace of six 1D Loewner matrices of size 2×2 is needed, (ii) \mathbf{Bary}_t , computation of three 1D Loewner matrices of size 2×2 is needed, and (iii) \mathbf{Bary}_s , computation of one 1D Loewner matrix of size 3×3 is needed. The resulting total computation using 1D Loewner matrices is $\nu_1^3\nu_2\nu_3 + \nu_2^3\nu_3 + \nu_3^3 = 99$ flops as opposed to $(\nu_1\nu_2\nu_3)^3 = 1728$ flops, when working with 3D Loewner matrices. For details on the computational complexity, storage and numerical accuracy, we refer to [3]. Note also that the n D Loewner matrix is of dimension 12×12 while in the 1D case, a maximum of 3×3 matrix is needed.

Comparison of KST and (*). A number of researchers have contributed in sharpening Kolmogorov's original result, so currently it is often referred to as the *Kolmogorov, Arnol'd, Kahane, Lorenz and Sprecher* Theorem (see [2], theorem 2.1). For simplicity we will follow [2] and state this result for $n = 3$, so that we can compare it with (*).

Theorem. Given a continuous function $f : [0, 1]^3 \rightarrow \mathbb{R}$ of three variables, there exist real numbers λ_i , $i = 1, 2$, and single-variable continuous functions $\phi_k : [0, 1] \rightarrow \mathbb{R}$, $k = 1, \dots, 7$, and a single-variable function $g : \mathbb{R} \rightarrow \mathbb{R}$, such that

$$f(x_1, x_2) = \sum_{k=1}^7 g(\phi_k(x_1) + \lambda_1\phi_k(x_2) + \lambda_2\phi_k(x_3)), \quad \forall (x_1, x_2, x_3) \in [0, 1]^3.$$

In the above result, λ_i and ϕ_k do not depend on f . Thus for $n = 3$, eight functions are needed together with two real scalars λ_i .

Similarities and differences between KST and (*).

1. While KST refers to continuous functions defined on $[0, 1]^n$, (*) is concerned with rational functions defined on \mathbb{C}^n .
2. In its present form (*) is valid in a particular basis, namely the *Lagrange* basis. Multiplication of functions in (*), is defined with respect to this basis.
3. The composition and superposition property holds for the numerator and denominator. Notice that in KST no explicit denominators are considered. This is important in our case because (*) preserves interpolation conditions.
4. The parameters needed are $n = 3$ Lagrange bases (one in each variable) and the barycentric coefficients of numerator and denominator.
5. Both KST and (*) accomplish the goal of replacing the computation of multivariate functions, by means of a series of computations involving single-variable functions, KST for general continuous functions and (*) for rational functions. Notice also that (*) provides a different formulation of the problem than KST.

References

- [1] G. Pólya and G. Szegő, *Problems and Theorems in Analysis*, Volume 1, Springer Verlag, 1972.
- [2] Sidney A. Morris, *Hilbert 13: Are there any genuine continuous multivariate functions?*, Bulletin (New Series) of the AMS, **58**: 107-118 (2021).
- [3] Athanasios C. Antoulas, Ion Victor Gosea, Charles Poussot-Vassal, *The Loewner framework for parametric systems: Taming the curse of dimensionality*, <https://arxiv.org/abs/2405.00495>. Submitted to SIREV, May 2024.