ISSN: 1889-3066
© 2016 Universidad de Jaén
Web site: jja.ujaen.es

Jaen J. Approx. 8(1) (2016), 33–58

Jaen Journal on Approximation

On tensor decomposition, sparse interpolation and Padé approximation†

Annie Cuyt, Wen-shin Lee and Xianglan Yang

Communicated by F. Marcellán

Received July 20, 2015
Accepted May 10, 2016

Abstract

What is the connection of tensor decomposition in multilinear algebra with exponential analysis from signal processing, sparse interpolation from computer algebra, Gaussian quadrature from numerical analysis, and Padé approximation theory? These seemingly unrelated and diverse topics are nevertheless deeply intertwined, as we explain here. However, several of these connections have remained unexplored.

The various reformulations bring forth new ways to approach the problem of tensor decomposition (see Section 7) and suggestions for generalizations of existing methods (see Section 6). This may lead to important results since tensor decomposition has a number of grand applications [24], among others in chemometrics, neuroscience, computer vision, social network analysis, big data and the like.

In Section 1 we introduce the problem statement. Subsequently the connections are first established for two-dimensional tensors in the sections 2 and 3. Higher dimensional tensors are dealt with in the sections 4 and 5, with a discussion of the connections to the mentioned topics in the sections 6 and 7. We conclude in Section 8 with an illustration of the interrelationships and of our novel approach. Both the existing method presented in Section 5 and the new technique presented in Section 7 are shown.

Keywords: tensor decomposition, sparse interpolation, Gaussian integration, Padé approximation, exponential analysis.

MSC: Primary 41A55; Secondary 41A25, 42B35.

†This research is in part supported by an FWO project grant.

§1. Symmetric tensors and homogeneous polynomials

A tensor of order d is an element of the product space $\mathbb{C}^{n_1} \otimes \cdots \otimes \mathbb{C}^{n_d}$ and it is called cubical if its dimensions $n_i, i = 1, \ldots, d$ satisfy $n_1 = \cdots = n_d = n$. A cubical tensor of order d and dimension $n+1$ is represented by a multidimensional array $[t_{j_1,\ldots,j_d}]_{j_1,\ldots,j_d=0}^{n}$ and it is called symmetric if for any permutation π of $\{j_1, \ldots, j_d\}$ holds $t_{j_1,\ldots,j_d} = t_{\pi(j_1),\ldots,\pi(j_d)}$. With a symmetric tensor $[t_{j_1,\ldots,j_d}]_{j_1,\ldots,j_d=0}^{n}$ we can associate a homogeneous polynomial of degree d in $n+1$ variables:

$$p(x_0, \ldots, x_n) = \sum_{j_1,\ldots,j_d=0}^{n} t_{j_1,\ldots,j_d} x_{j_1} \cdots x_{j_d},$$

which can be written more compactly as

$$\sum_{j_1 \le \ldots \le j_d = 0}^{n} \left(\sum_{\pi} t_{\pi(j_1),\ldots,\pi(j_d)} \right) x_{j_1} \cdots x_{j_d},$$

or just

$$\sum_{|\kappa|=d} c_\kappa X^\kappa,$$

where $X = (x_0, \ldots, x_n)$, $\kappa = (k_0, \ldots, k_n)$, $|\kappa| = k_0 + \cdots + k_n$ and

$$X^\kappa = x_0^{k_0} x_1^{k_1} \cdots x_n^{k_n}.$$

For instance, the homogeneous polynomial associated with a symmetric tensor of order 3 and dimension 3 is

$$\begin{aligned} p(x_0, x_1, x_2) &= t_{000} x_0^3 + (t_{100} + t_{010} + t_{001}) x_0^2 x_1 + \ldots \\ &= c_{300} x_0^3 + c_{210} x_0^2 x_1 + \ldots \\ &= \sum_{|\kappa|=3} c_\kappa X^\kappa. \end{aligned}$$

For $[t_{j_1,j_2,j_3}]_{j_1,j_2,j_3=0}^{2}$ given by

$$\left(\begin{array}{ccc|ccc|ccc} t_{000} & t_{010} & t_{020} & t_{001} & t_{011} & t_{021} & t_{002} & t_{012} & t_{022} \\ t_{100} & t_{110} & t_{120} & t_{101} & t_{111} & t_{121} & t_{102} & t_{112} & t_{122} \\ t_{200} & t_{210} & t_{220} & t_{201} & t_{211} & t_{221} & t_{202} & t_{212} & t_{222} \end{array}\right) = \left(\begin{array}{ccc|ccc|ccc} 1 & 2 & 1 & 2 & 4 & 2 & 1 & 2 & 1 \\ 2 & 4 & 2 & 4 & 8 & 4 & 2 & 4 & 2 \\ 1 & 2 & 1 & 2 & 4 & 2 & 1 & 2 & 1 \end{array}\right),$$

$p(x_0, x_1, x_2)$ equals

$$x_0^3 + 6x_0^2x_1 + 3x_0^2x_2 + 12x_0x_1^2 + 12x_0x_1x_2 + 3x_0x_2^2 + 8x_1^3 + 12x_1^2x_2 + 6x_1x_2^2 + x_2^3. \tag{1.1}$$

The tensor decomposition or polynomial decomposition problem now consists in finding the least number r of linear forms

$$\ell_i(x_0, \ldots, x_n) = \sum_{k=0}^{n} \lambda_{ik} x_k$$

and weights w_i such that

$$p(x_0, \ldots, x_n) = \sum_{i=1}^{r} w_i \ell_i(x_0, \ldots, x_n)^d.$$

This minimal r is called the rank of the symmetric tensor $[t_{j_1,\ldots,j_d}]_{j_1,\ldots,j_d=0}^{n}$. For example, the polynomial in (1.1) can be decomposed as $(x_0 + 2x_1 + x_2)^3$ and hence $r = 1$.

Without loss of generality, we can set $x_0 = 1$ and assume that $\lambda_{i0} = 1$, because a change of variables can always guarantee this. Then the linear forms $\ell_i(x_0, \ldots, x_n)$ take the form

$$\ell_i(1, x_1, \ldots, x_n) = 1 + \sum_{k=1}^{n} \lambda_{ik} x_k, \qquad i = 1, \ldots, r.$$

§2. Two-dimensional tensors and Hankel systems

When $n = 1$, $x := x_1$, $k := k_1$ and $\lambda_i := \lambda_{i1}$, the polynomial $p(1, x)$ and its decomposition take the form

$$\begin{aligned} p(1, x) &= \sum_{k=0}^{d} c_k x^k \\ &= \sum_{i=1}^{r} w_i (1 + \lambda_i x)^d \\ &= \sum_{j=0}^{d} \sum_{i=1}^{r} w_i \binom{d}{j} \lambda_i^j x^j. \end{aligned} \tag{2.1}$$

Let us denote $\sigma_k := c_k / \binom{d}{k}$, $k = 0, \ldots, d$. Then for $i = 1, \ldots, r$ the w_i and λ_i must satisfy the polynomial system

$$\sum_{i=1}^{r} w_i \lambda_i^j = \sigma_j, \qquad j = 0, \ldots, d. \tag{2.2}$$

Now define

$$V(z) := \prod_{i=1}^{r}(z-\lambda_i) = z^r + \sum_{j=0}^{r-1} b_j z^j \tag{2.3}$$

and denote

$$H_r^{(k)} := \left(\sigma_{k+(i-1)+(j-1)}\right)_{i,j=1}^{r} = \begin{pmatrix} \sigma_k & \sigma_{k+1} & \cdots & \sigma_{k+r-1} \\ \sigma_{k+1} & \sigma_{k+2} & \cdots & \sigma_{k+r} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{k+r-1} & \sigma_{k+r} & \cdots & \sigma_{k+2r-2} \end{pmatrix}.$$

Then we obtain from (2.3) that

$$\sum_{i=1}^{r} w_i \lambda_i^k V(\lambda_i) = 0$$

or

$$\sigma_{k+r} + \sum_{j=0}^{r-1} b_j \sigma_{k+j} = 0, \qquad k = 0,1,2,\ldots \tag{2.4}$$

The coefficients b_j of the polynomial $V(z)$ can therefore be obtained from the Hankel system

$$H_r^{(0)} \begin{pmatrix} b_0 \\ \vdots \\ b_{r-1} \end{pmatrix} = - \begin{pmatrix} \sigma_r \\ \vdots \\ \sigma_{2r-1} \end{pmatrix}, \tag{2.5}$$

under the condition that a sufficient number of moments σ_k is given. The regularity of $H_r^{(0)}$, and more generally $H_r^{(k)}$, is certified in (2.8). Since the $\lambda_i, i = 1, \ldots, r$, are distinct, the matrix factorizations

$$\begin{aligned} H_r^{(0)} &= VDV^T, \\ H_r^{(k)} &= VD \begin{pmatrix} \lambda_1^k & & \\ & \ddots & \\ & & \lambda_r^k \end{pmatrix} V^T, \qquad k = 1,2,\ldots \end{aligned} \tag{2.6}$$

where V and D respectively denote the Vandermonde matrix

$$V = \begin{pmatrix} 1 & 1 & \dots & 1 \\ \lambda_1 & \lambda_2 & \dots & \lambda_r \\ \vdots & \vdots & & \vdots \\ \lambda_1^{r-1} & \lambda_2^{r-1} & \dots & \lambda_r^{r-1} \end{pmatrix}$$

and the diagonal matrix

$$D = \begin{pmatrix} w_1 & & \\ & \ddots & \\ & & w_r \end{pmatrix},$$

hold and the roots λ_i of $V(z)$ can also be obtained from the generalized eigenvalue problem [20, p. 1226]

$$H_r^{(1)} u_i = \lambda_i H_r^{(0)} u_i, \qquad i = 1, \dots, r,$$

or the eigenvalue problem [5, p. 1859]

$$H_r^{(1)} \left(H_r^{(0)} \right)^{-1} \hat{u}_i = \lambda_i \hat{u}_i, \qquad i = 1, \dots, r.$$

Note that the (properly normalized) eigenvectors $\hat{u}_i, i = 1, \dots, r$, equal the columns of the matrix V.

Remains the problem of obtaining the rank r of the symmetric tensor. Here also sparse interpolation and Padé approximation, discussed in detail in Section 3, help us out. For $k \geq 0$ and denoting $|H_r^{(k)}| = \det H_r^{(k)}$, it is known [21, p. 603] that

$$|H_R^{(k)}| = 0, \qquad R > r, \tag{2.7}$$

and it is proved in [23] that

$$\begin{aligned} &|H_r^{(k)}| \neq 0, \\ &|H_s^{(k)}| = 0, \text{ only accidentally}, \qquad s < r. \end{aligned} \tag{2.8}$$

So the rank r can be deduced from the fact that the matrices $H_R^{(0)}$ with $R > r$ are singular. Missing data $\sigma_j, j > d$, in such a larger matrix, can be obtained from the fact that the matrices $G_k = H_r^{(k)} (H_r^{(0)})^{-1}$ and $G_\ell = H_r^{(\ell)} (H_r^{(0)})^{-1}$ with $k \neq \ell$ commute. Expressing

$$G_k G_\ell = G_\ell G_k, \qquad k \neq \ell,$$

delivers a set of equations from which the unknown moments can be computed.

§3. Connection with sparse interpolation and Padé approximation

When we re-express

$$\lambda_i = \Delta^{m_i}, \qquad \Delta = \exp(2\pi \mathrm{i}/M), \qquad M > 2 \max_{i=1,\ldots,r} |\Im(m_i)|,$$

then the computation of λ_i amounts to the solution of the sparse interpolation problem

$$\sum_{i=1}^{r} w_i z_j^{m_i} = \sigma_j, \qquad j = 0, \ldots, d, \qquad z_j = \Delta^j = \exp(2\pi \mathrm{i} j/M), \tag{3.1}$$

where the coefficients w_i and the complex exponents m_i need to be determined simultaneously in the polynomial (3.1). Since

$$\sum_{i=1}^{r} w_i z_j^{m_i} = \sum_{i=1}^{r} w_i \exp\left(2\pi \mathrm{i} j m_i/M\right),$$

statement (3.1) is also called an exponential analysis problem.

When introducing the linear functional

$$\gamma : z^j \to \sigma_j, \qquad j = 0, 1, 2, \ldots,$$

it follows from (2.4) that the polynomial $V(z)$ satisfies the formal orthogonality conditions

$$\gamma(z^k V(z)) = 0, \qquad k = 0, 1, 2, \ldots$$

The polynomial $V(z)$ is also called a Hadamard polynomial [21, p. 625] and can be written compactly as

$$\left|H_r^{(0)}\right| V(z) = \begin{vmatrix} \sigma_0 & \ldots & \sigma_{r-1} & \sigma_r \\ \sigma_1 & \ldots & \sigma_r & \sigma_{r+1} \\ \vdots & \ddots & \vdots & \vdots \\ \sigma_{r-1} & \ldots & \sigma_{2r-2} & \sigma_{2r-1} \\ 1 & \ldots & z^{r-1} & z^r \end{vmatrix}.$$

We now compute the interpolating polynomial of degree $r-1$ through the interpolation points $z = \lambda_i$, $i = 1, \ldots, r$, for the function $f(z) = (1 + zx)^d$ of the variable z, where x is treated as a parameter. Its Lagrange form equals

$$\sum_{i=1}^{r}(1+\lambda_i x)^d \frac{V(z)}{(z-\lambda_i)V'(\lambda_i)}. \tag{3.2}$$

Applying the linear functional γ to both $f(z)$ and (3.2) delivers on the one hand

$$\gamma\left((1+zx)^d\right) = \gamma\left(\sum_{j=0}^{d}\binom{d}{j} z^j x^j\right) = \sum_{j=0}^{d} c_j x^j = p(1,x)$$

and on the other hand

$$\sum_{i=1}^{r} A_i(1+\lambda_i x)^d, \qquad A_i = \gamma\left(\frac{V(z)}{(z-\lambda_i)V'(\lambda_i)}\right). \tag{3.3}$$

The two expressions equal each other if $A_i = w_i$, this means if the A_i satisfy (2.1). Remember that the λ_i are the zeroes of the formally orthogonal polynomial $V(z)$. Apparently the nodes λ_i provide jointly with the weights w_i an exact formula for the moments up to degree d, namely for the σ_j, $j = 0, \ldots, d$, as can be seen from (2.2). When $d = 2r-1$ as in (2.5), then formula (2.2) is a Gaussian integration rule. How the latter can be viewed as a Padé approximant of degree $r-1$ in the numerator and degree r in the denominator is explained in [6, pp. 62]. For completeness we mention that the denominator of this Padé approximant equals the reverse of the polynomial $V(z)$ given by (2.3); in other words

$$z^r V(1/z) = \sum_{j=0}^{r-1} b_j z^{r-j} + 1. \tag{3.4}$$

To see how this connection works in practice we take a simple example. Consider the two-dimensional tensor of order $d = 3$, associated with

$$p(x_0, x_1) = 4x_0^3 + 9x_0^2 x_1 + 18 x_0 x_1^2 + 17 x_1^3.$$

To decompose the polynomial $p(1,x)$ we construct

$$H_2^{(0)} = \begin{pmatrix} \sigma_0 & \sigma_1 \\ \sigma_1 & \sigma_2 \end{pmatrix} = \begin{pmatrix} 4 & 3 \\ 3 & 6 \end{pmatrix},$$

$$H_2^{(1)} = \begin{pmatrix} \sigma_1 & \sigma_2 \\ \sigma_2 & \sigma_3 \end{pmatrix} = \begin{pmatrix} 3 & 6 \\ 6 & 17 \end{pmatrix}.$$

From solving the generalized eigenvalue problem we obtain the decomposition

$$p(x_0, x_1) = \frac{5}{8}(x_0 + 3x_1)^3 + \frac{27}{8}(x_0 + \frac{1}{3}x_1)^3.$$

The same result is obtained as follows. We form the Hadamard polynomial,

$$\begin{aligned} V(z) &= \begin{vmatrix} 4 & 3 & 6 \\ 3 & 6 & 17 \\ 1 & z & z^2 \end{vmatrix} / \left| H_2^{(0)} \right| \\ &= z^2 - \frac{10}{3}z + 1, \end{aligned}$$

which is the reverse of the denominator of the Padé approximant of degree 1 in the numerator and 2 in the denominator for the series

$$\sum_{j=0}^{\infty} \sigma_j z^j. \tag{3.5}$$

In [1] and [15] it is indicated that (3.5) is actually the Taylor series of a rational function with denominator (3.4).

The roots of $V(z)$ are $\lambda_1 = {}^1\!/_3$ and $\lambda_2 = 3$. Computing the A_i using the linear functional γ and formula (3.3), results in $A_1 = {}^{27}\!/_8$ and $A_2 = {}^5\!/_8$ and eventually in the same decomposition. For completeness, it is easy to verify that as stated in (2.7) and (2.8),

$$\left| H_R^{(0)} \right| = 0, \qquad R = 3 > r = 2,$$

with the missing σ_4 computed from the commutativity of $H_2^{(2)}(H_2^{(0)})^{-1}$ and $H_2^{(1)}(H_2^{(0)})^{-1}$ as follows. For $r = 2$, the products G_1G_2 and G_2G_1 are given by

$$\begin{aligned} G_1G_2 &= \begin{pmatrix} \frac{34-\sigma_4}{5} & \frac{-51+4\sigma_4}{15} \\ \frac{71-2\sigma_4}{3} & \frac{-132+8\sigma_4}{9} \end{pmatrix}, \\ G_2G_1 &= \begin{pmatrix} -\frac{10}{3} & \frac{91}{9} \\ \frac{51-4\sigma_4}{15} & \frac{-204+31\sigma_4}{45} \end{pmatrix}. \end{aligned}$$

From $G_1G_2 = G_2G_1$ we find $\sigma_4 = 152/3$ and eventually $|H_3^{(0)}| = 0$.

§4. Higher dimensional tensors and Hankel-like systems

By setting $x_0 = 1$ the homogeneous polynomial $p(x_0, \dots, x_n)$ takes the form

$$p(1, x_1, \dots, x_n) = \sum_{|\kappa| \le d} c_\kappa X^\kappa, \tag{4.1}$$

where we redefine $\kappa = (k_1, \dots, k_n)$ and $X = (x_1, \dots, x_n)$ since k_0 and x_0 are known by $k_0 = d - (k_1 + \cdots + k_n)$ and $x_0 = 1$. Again, the decomposition problem consists in finding a minimal number r such that the polynomial $p(1, x_1, \dots, x_n)$ of total degree d can be written as

$$p(1, x_1, \dots, x_n) = \sum_{i=1}^{r} w_i (1 + \lambda_{i1} x_1 + \cdots + \lambda_{in} x_n)^d; \tag{4.2}$$

in other words, such that

$$p(1, x_1, \dots, x_n) = \sum_{|\nu| \le d} \left(\sum_{i=1}^{r} w_i \binom{d}{\nu} \Lambda_i^\nu \right) X^\nu, \tag{4.3}$$

where $\Lambda_i = (\lambda_{i1}, \dots, \lambda_{in})$, $X = (x_1, \dots, x_n)$ and $\nu = (\nu_1, \dots, \nu_n)$ are all n-dimensional vectors. Similar to X^ν, the compact notation Λ_i^ν denotes the monomial

$$\Lambda_i^\nu = \lambda_{i1}^{\nu_1} \cdots \lambda_{in}^{\nu_n}.$$

For (4.3) to hold (see [5, p. 1856]), the Λ_i must satisfy

$$\sum_{i=1}^{r} w_i \Lambda_i^\nu = \sigma_\nu, \qquad \sigma_\nu = c_\nu \Big/ \binom{d}{\nu}, \qquad |\nu| \le d, \tag{4.4}$$

which is the n-variate generalization of (2.2), where as usual

$$\binom{d}{\nu} = \frac{d!}{\nu_1! \cdots \nu_n! (d - |\nu|)!}.$$

In what follows we use subsets $I \subset \mathbb{N}^n \cap \{\nu : |\nu| \le d\}$ of index vectors, and assume that their elements are ordered as $\nu(0), \nu(1), \nu(2), \dots$ with $\nu(0) = (0, \dots, 0)$. Each index vector $\nu(k)$ in the set I also needs to be connected to $(0, \dots, 0)$ ([5, p. 1860]), meaning that either the index vector is

$\nu(0) = (0, \ldots, 0)$ or for $\nu(k)$, $k > 0$, there exist $\nu(h)$ and i with $0 \le h < k$ and $1 \le i \le n$ such that $\nu(k) = \nu(h) + (\delta_{i1}, \ldots, \delta_{in})$, where $\delta_{ij} = 0$ if $i \neq j$ and $\delta_{ij} = 1$ if $i = j$. For $n = 1$ we have $\nu(j) = j$.

Then (4.4) can be rewritten as ([5, p. 1856])

$$\sum_{i=1}^{r} w_i \Lambda_i^{\nu(j)} = \sigma_{\nu(j)}, \qquad j = 0, \ldots, \binom{d+n}{n} - 1, \tag{4.5}$$

or

$$\begin{pmatrix} \Lambda_1^{\nu(0)} & \ldots & \Lambda_r^{\nu(0)} \\ \Lambda_1^{\nu(1)} & \ldots & \Lambda_r^{\nu(1)} \\ \vdots & & \vdots \\ \Lambda_1^{\nu(s)} & \ldots & \Lambda_r^{\nu(s)} \end{pmatrix} \begin{pmatrix} w_1 \\ \vdots \\ w_r \end{pmatrix} = \begin{pmatrix} \sigma_{\nu(0)} \\ \vdots \\ \sigma_{\nu(s)} \end{pmatrix}, \qquad s = \binom{d+n}{n} - 1.$$

The subsets of index vectors required in the sequel, are indexed as

$$I_0 = \{(0, \ldots, 0)\}, \qquad \# I_{k-1} = k,$$

and the index vectors in some selected I_{k-1}, which are all connected to $(0, \ldots, 0)$, are indicated as $\nu(0), \nu(1), \ldots, \nu(k-1)$. This enumeration of the index vectors may be different for different sets that contain k elements. We also define the boundary ∂I of an index set I by

$$\partial I = (\cup_{i=1}^{n} \{\nu + (\delta_{i1}, \ldots, \delta_{in}) : \nu \in I\}) \setminus I.$$

In the sequel we assume that the monomials $\Lambda_i^{\nu(j)}$ are such that the Vandermonde matrix (5.2), and hence the matrices $H_r^{(k)}$ in (5.1) for $k \ge 0$, are regular. We now indicate how one can solve for the vectors Λ_i, $i = 1, \ldots, r$, given the moments $\sigma_{\nu(j)}$ ([4, 26]). Afterwards it is easy to obtain the weights w_i, $i = 1, \ldots, r$ from the generalized Vandermonde system (4.5).

§5. Hankel-like generalized eigenvalue problems

We denote by $H_r^{(k)}$ the matrix

$$H_r^{(k)} := \left(\sigma_{\nu(k)+\nu(i-1)+\nu(j-1)}\right)_{i,j=1}^{r} = \begin{pmatrix} \sigma_{\nu(k)+\nu(0)+\nu(0)} & \cdots & \sigma_{\nu(k)+\nu(0)+\nu(r-1)} \\ \sigma_{\nu(k)+\nu(1)+\nu(0)} & \cdots & \sigma_{\nu(k)+\nu(1)+\nu(r-1)} \\ \vdots & & \vdots \\ \sigma_{\nu(k)+\nu(r-1)+\nu(0)} & \cdots & \sigma_{\nu(k)+\nu(r-1)+\nu(r-1)} \end{pmatrix}. \tag{5.1}$$

When $n = 1$ the matrix $H_r^{(k)}$ coincides with the one-dimensional definition and hence we do not introduce a different notation. How to obtain the proper value of r and the proper index set I_{r-1} of r index vectors connected to $(0, \dots, 0)$, to proceed with, is explained in [5, p. 1861].

Using the matrix factorizations

$$H_r^{(0)} = VDV^T,$$
$$H_r^{(k)} = VD \begin{pmatrix} \Lambda_1^{\nu(k)} & & \\ & \ddots & \\ & & \Lambda_r^{\nu(k)} \end{pmatrix} V^T,$$

where the matrices V and D respectively denote the Vandermonde-like matrix

$$V = \begin{pmatrix} \Lambda_1^{\nu(0)} & \dots & \Lambda_r^{\nu(0)} \\ \Lambda_1^{\nu(1)} & \dots & \Lambda_r^{\nu(1)} \\ \vdots & & \vdots \\ \Lambda_1^{\nu(r-1)} & \dots & \Lambda_r^{\nu(r-1)} \end{pmatrix} \tag{5.2}$$

and the diagonal matrix

$$D = \begin{pmatrix} w_1 & & \\ & \ddots & \\ & & w_r \end{pmatrix},$$

it is easy to see that ([4, p. 60], [5, p. 1859], [26, p. 38])

$$H_r^{(k)} u_i = \Lambda_i^{\nu(k)} H_r^{(0)} u_i, \qquad i = 1, \dots, r, \qquad k = 0, 1, 2, \dots \tag{5.3}$$

or

$$H_r^{(k)} \left(H_r^{(0)} \right)^{-1} \hat{u}_i = \Lambda_i^{\nu(k)} \hat{u}_i, \qquad i = 1, \dots, r, \qquad k = 0, 1, 2, \dots$$

The corresponding (properly normalized) generalized eigenvector $\hat{u}_i$ is given by

$$\hat{u}_i = \left(1, \Lambda_i^{\nu(1)}, \dots, \Lambda_i^{\nu(r-1)} \right)^T, \qquad i = 1, \dots, r,$$

which is independent of k.

With some sets of index vectors $\nu(j)$, the values $\lambda_{i1}, \ldots, \lambda_{in}, i = 1, \ldots, r$, can be read directly from the generalized eigenvalues or eigenvectors, or they can be obtained with a little extra effort. For instance, having

$$(0, \ldots, 0),\ (1, 0, \ldots, 0),\ \ldots,\ (0, \ldots, 0, 1)$$

as the first $n+1$ index vectors $\nu(j)$, $j = 0, \ldots, n$, hands us λ_{ij} for $j = 1, \ldots, n$ and $i = 1, \ldots, r$. Of course, this is not the only possibility for the $\nu(j)$ and it may not even be a suitable one.

If some of the moments in the matrices $H_r^{(k)}$ are unknown because their index lies outside the set $|\nu| \le d$, then the method can still be used. In that case one proceeds as follows. Using (5.3) it is easy to see that again the matrices $G_k := H_r^{(k)}(H_r^{(0)})^{-1}$ and $G_\ell := H_r^{(\ell)}(H_r^{(0)})^{-1}$ with $k \neq \ell$ commute. So by solving the equations

$$G_k G_\ell = G_\ell G_k, \qquad k \neq \ell,$$

for the unknown moments and substituting these in the Hankel-like matrices, one can continue the procedure.

§6. Connection with multivariate sparse interpolation and Padé approximation

The connection between tensor decomposition for $n = 1$, and sparse interpolation, exponential analysis, Gaussian quadrature or Padé approximation is through (2.2), which can be solved by Prony's method [22]. In case $n > 1$, several multivariate generalizations of these concepts are possible. Here we present a discussion of some connections with the tensor decomposition algorithm in Section 5. In the next section we present an entirely different approach built on a different multivariate generalization and its connections.

As a result of all these interrelationships, new algorithms for tensor decomposition can be expected in the future. For instance, the connections described here will make it possible (see [14]) to reduce the number of moments required in the tensor decomposition problem to precisely $(n+1)r$, which is the minimum because it equals the number of unkown parameters. A reduction from $O\left(r^n\right)$ when using gridded data, to $O\left(nr\right)$ in particular cases (see [18]), is yet possible via a method using 1-dimensional projections of the data on lines (see [25, 27, 28]). A further reduction, using the technique described in the forthcoming paper [14] on Prony's method in higher dimensions, is however possible. Here we only focus on the interrelationships between the different topics.

As in the 1-dimensional problem we re-express each λ_{ik} as

$$\lambda_{ik} = \Delta_k^{m_{ik}}, \qquad \Delta_k = \exp(2\pi \mathrm{i}/M_k), \qquad M_k > 2 \max_{i=1,\ldots,r} |\Im(m_{ik})|, \qquad k = 1, \ldots, n.$$

Then the tensor decomposition problem can be reformulated as the sparse interpolation problem with complex exponents,

$$\sum_{i=1}^{r} w_i Z_j^{\mu_i} = \sigma_{\nu(j)}, \qquad |\nu(j)| \le d, \qquad \mu_i = (m_{i1}, \ldots, m_{in}), \qquad Z_j = (\Delta_1^{\nu_{j1}}, \ldots, \Delta_n^{\nu_{jn}}), \tag{6.1}$$

where $\nu(j) = (\nu_{j1}, \ldots, \nu_{jn})$. Expression (6.1) is a multivariate version (not the most general though) of (3.1). This sparse interpolation problem is often called a multivariate exponential analysis problem since, as in (4.5),

$$\begin{aligned}\sum_{i=1}^{r} w_i Z_j^{\mu_i} &= \sum_{i=1}^{r} w_i \exp\left(2\pi\mathrm{i}\left(\frac{\nu_{j1} m_{i1}}{M_1} + \ldots + \frac{\nu_{jn} m_{in}}{M_n}\right)\right) \\ &= \sum_{i=1}^{r} w_i \lambda_{i1}^{\nu_{j1}} \cdots \lambda_{in}^{\nu_{jn}} = \sigma_{\nu(j)}.\end{aligned}$$

So far the connection with Prony's method. We now focus on the connection with concepts from approximation theory. Define $\mathcal{V}(Z)$ of the form

$$\mathcal{V}(Z) = Z^{\nu(r)} + \sum_{j=0}^{r-1} b_j Z^{\nu(j)}, \qquad \nu(r) \in \partial I_{r-1}, \qquad Z = (z_1, \ldots, z_n),$$

with $\mathcal{V}(\Lambda_i) = 0$. The coefficients $b_j, j = 0, \ldots, r-1$, in $\mathcal{V}(Z)$ are obtained from

$$\sigma_{\nu(k)+\nu(r)} + \sum_{j=0}^{r-1} b_j \sigma_{\nu(k)+\nu(j)} = 0, \qquad k = 0, 1, 2, \ldots, \tag{6.2}$$

or in matrix notation

$$H_r^{(0)} \begin{pmatrix} b_0 \\ \vdots \\ b_{r-1} \end{pmatrix} = - \begin{pmatrix} \sigma_{\nu(r)+\nu(0)} \\ \vdots \\ \sigma_{\nu(r)+\nu(r-1)} \end{pmatrix},$$

which follows from

$$\sum_{i=1}^{r} w_i \Lambda_i^{\nu(k)} \mathcal{V}(\Lambda_i) = 0, \qquad k = 0, 1, 2, \ldots,$$

in combination with (4.5). So a determinant representation for $\mathcal{V}(Z)$ is

$$\left|H_r^{(0)}\right| \mathcal{V}(Z) = \begin{vmatrix} \sigma_{\nu(0)+\nu(0)} & \cdots & \sigma_{\nu(0)+\nu(r-1)} & \sigma_{\nu(0)+\nu(r)} \\ \sigma_{\nu(1)+\nu(0)} & \cdots & \sigma_{\nu(1)+\nu(r-1)} & \sigma_{\nu(1)+\nu(r)} \\ \vdots & & \vdots & \\ \sigma_{\nu(r-1)+\nu(0)} & \cdots & \sigma_{\nu(r-1)+\nu(r-1)} & \sigma_{\nu(r-1)+\nu(r)} \\ & & & \\ Z^{\nu(0)} & \cdots & Z^{\nu(r-1)} & Z^{\nu(r)} \end{vmatrix}, \tag{6.3}$$

with $H_r^{(0)}$ as in Section 5.

This approach is still very much univariate in nature, as it can be seen from the fact that the indices in (6.2) and (6.3) take the form $\nu(k)+\nu(j)$ and not $\nu(k+j)$ as in the general multivariate Padé approximants in [11, 10]. It is not our goal to provide a detailed discussion of the connection to the latter approximants as well, but merely to provide a connection to this family of methods for the purpose of future developments in tensor decomposition. Another approach based on an underlying univariate principle is presented in the next section.

§7. A new approach via homogeneous multivariate Padé approximation

Let us switch from a cartesian to a spherical coordinate system. We write

$$X = (x_1, \ldots, x_n) = (\theta_1 x, \ldots, \theta_n x), \qquad ||(\theta_1, \ldots, \theta_n)||_2 = 1, \qquad x \in \mathbb{R},$$

where $\Theta = (\theta_1, \ldots, \theta_n)$ is a directional vector and x is the signed distance from the origin (for a similar approach we refer to [12]). Then the polynomial $p(1, x_1, \ldots, x_n)$ is rewritten as

$$\begin{aligned} p(\Theta; x) &= \sum_{j=0}^{d} \left(\sum_{|\nu|=j} c_\nu \Theta^\nu \right) x^j \\ &= \sum_{j=0}^{d} C_j(\Theta) x^j, \end{aligned}$$

where

$$C_j(\Theta) := \sum_{|\nu|=j} c_\nu \Theta^\nu, \qquad c_\nu \Theta^\nu = c_{\nu_1,\ldots,\nu_n} \theta_1^{\nu_1} \cdot \ldots \cdot \theta_n^{\nu_n}.$$

The decomposition problem takes the form

$$
\begin{aligned}
p(\Theta;x) &= \sum_{i=1}^{r} w_i \left(1+\Lambda_i(\Theta)x\right)^d \\
&= \sum_{j=0}^{d} \binom{d}{j} \left(\sum_{i=1}^{r} w_i \Lambda_i(\Theta)^j \right) x^j,
\end{aligned}
$$

where

$$\Lambda_i(\Theta) = \sum_{k=1}^{n} \lambda_{ik}\theta_k.$$

The counterpart of (2.2) is now

$$\sum_{i=1}^{r} w_i \Lambda_i(\Theta)^j = C_j(\Theta)/\binom{d}{j}, \qquad j=0\ldots,d.$$

Putting $S_j := C_j(\Theta)/\binom{d}{j}$ the system looks like

$$\sum_{i=1}^{r} w_i \Lambda_i(\Theta)^j = S_j(\Theta), \qquad j=0\ldots,d, \tag{7.1}$$

where the moments $S_j(\Theta)$ and the vectors $\Lambda_i(\Theta)$, $i=1,\ldots,r$, are now parameterized by the directional vector Θ. In the same way homogeneous multivariate Padé approximants were introduced in [7, 9].

Now define

$$V(\Theta;z) := \prod_{i=1}^{r} (z-\Lambda_i(\Theta)) = z^r + \sum_{j=0}^{r-1} B_j(\Theta) z^j, \tag{7.2}$$

where $B_j(\Theta)$ is a multivariate homogeneous polynomial in Θ of degree $r-j$, and denote

$$H_r^{(k)}(\Theta) := \left(S_{k+(i-1)+(j-1)}(\Theta)\right)_{i,j=1}^{r} = \begin{pmatrix} S_k(\Theta) & S_{k+1}(\Theta) & \ldots & S_{k+r-1}(\Theta) \\ S_{k+1}(\Theta) & S_{k+2}(\Theta) & \ldots & S_{k+r}(\Theta) \\ \vdots & \vdots & \ddots & \vdots \\ S_{k+r-1}(\Theta) & S_{k+r}(\Theta) & \ldots & S_{k+2r-2}(\Theta) \end{pmatrix}.$$

From (7.1) and (7.2) we obtain that

$$\sum_{i=1}^{r} w_i \Lambda_i^k(\Theta) V(\Theta; \Lambda_i(\Theta)) \equiv 0,$$

or

$$S_{k+r}(\Theta) + \sum_{j=0}^{r-1} B_j(\Theta) S_{k+j}(\Theta) \equiv 0, \qquad k = 0, 1, 2, \ldots \tag{7.3}$$

So the expressions $B_j(\Theta)$ that serve as parameterized coefficients in the polynomial $V(\Theta; z)$ in z, can be obtained from the parameterized Hankel system

$$H_r^{(0)}(\Theta) \begin{pmatrix} B_0(\Theta) \\ \vdots \\ B_{r-1}(\Theta) \end{pmatrix} = - \begin{pmatrix} S_r(\Theta) \\ \vdots \\ S_{2r-1}(\Theta) \end{pmatrix}, \tag{7.4}$$

under the condition that a sufficient number of moments $S_k(\Theta)$ is given. The existence of a nontrivial solution of (7.4) for general $n > 1$ is guaranteed in [8, pp. 60–62]. The same matrix factorizations of $H_r^{(k)}(\Theta)$ hold as for (2.5), but now with all entries parameterized by Θ. So the $\Lambda_i(\Theta)$ can also be obtained from the parameterized generalized eigenvalue problem

$$H_r^{(1)}(\Theta) u_i = \Lambda_i(\Theta) H_r^{(0)}(\Theta) u_i, \qquad i = 1, \ldots, r. \tag{7.5}$$

Let us introduce the linear functional

$$\Gamma : z^j \to S_j(\Theta), \qquad j = 0, 1, 2, \ldots$$

Then the polynomial $V(\Theta; z)$ satisfies the formal orthogonality conditions

$$\Gamma(z^k V(\Theta; z)) \equiv 0, \qquad k = 0, 1, 2, \ldots,$$

and equals the parameterized Hadamard polynomial $V(\Theta; z)$ given by

$$\left| H_r^{(0)}(\Theta) \right| V(\Theta; z) = \begin{vmatrix} S_0(\Theta) & \ldots & S_{r-1)}(\Theta) & S_r(\Theta) \\ S_1(\Theta) & \ldots & S_r(\Theta) & S_{r+1}(\Theta) \\ \vdots & \ddots & \vdots & \vdots \\ S_{r-1}(\Theta) & \ldots & S_{2r-2}(\Theta) & S_{2r-1}(\Theta) \\ 1 & \ldots & z^{r-1} & z^r \end{vmatrix}, \tag{7.6}$$

where $|H_r^{(0)}(\Theta)|$ denotes the determinant of $H_r^{(0)}(\Theta)$. From (7.6) we find that the expressions $B_j(\Theta)$ also are multivariate rational functions of Θ, of homogeneous degree $r^2 - j$ in the numerator and homogeneous degree $r^2 - r$ in the denominator ([29]). But if the tensor decomposition (7.1) exists, then we know from (7.3) that $|H_r^{(0)}(\Theta)|$ of homogeneous degree $r^2 - r$, is a common factor that can be cancelled in numerator and denominator of each of the $B_j(\Theta)$. In the same way as for (3.4), the parameterized polynomial $V(\Theta; z)$ is related to the denominator of the homogeneous Padé approximant of degree $r - 1$ in the numerator and r in the denominator by

$$z^r V(\Theta; 1/z) = 1 + \sum_{j=1}^{r} B_{r-j}(\Theta) z^j.$$

Some words on obtaining the value of r. This value need not be the same for all directional vectors Θ ([16]). Although we know that for $k \geq 0$,

$$\left| H_r^{(k)}(\Theta) \right| \not\equiv 0,$$

we cannot guarantee that $|H_r^{(0)}(\Theta)| \neq 0$ for all Θ. The expression $|H_r^{(0)}(\Theta)|$ is a polynomial in Θ and its zeroes constitute a set of Lebesgue measure zero in $\mathbb{C}^n$. However, we can guarantee that for $k \geq 0$,

$$\left| H_R^{(k)}(\Theta) \right| \equiv 0, \qquad R > r.$$

An illustration of this is given in the next section where we apply the methods of Section 5 and Section 7 on an example.

The relation with higher dimensional integration formulas, more precisely Gaussian cubature, is established subsequently, following the ideas in [2]. Recent results on this can be found in [3]. Proceeding as in (3.2), the Gaussian cubature formula can be viewed as a homogeneous multivariate Padé approximant of degree $r - 1$ in the numerator and r in the denominator ([2, 7]). So the different connections between tensor decomposition, multivariate Prony methods and sparse interpolation ([13]), multivariate orthogonality ([12]), Gaussian cubature ([2, 3]) and multivariate Padé approximation ([7]) exist clearly.

§8. Illustration

Consider the tensor of order 3 and dimension 3 associated with the homogeneous polynomial

$$p(x_0, x_1, x_2) = x_0 x_1 x_2 + x_1^3.$$

We first illustrate the technique described in Section 5 for the decomposition of this tensor. With

$$p(1, x_1, x_2) = x_1 x_2 + x_1^3,$$

we find that, except for

$$\sigma_{11} = {}^{1}\!/_{6}, \ \sigma_{30} = 1,$$

all σ_ν with $|\nu| \leq d$ are zero. Following [5, p. 1861] we find that $r = 4$ and

$$\nu(0) = (0,0), \ \nu(1) = (1,0), \ \nu(2) = (0,1), \ \nu(3) = (1,1), \tag{8.1}$$

$$I_3 = \{(0,0), (1,0), (0,1), (1,1)\}.$$

Construct the matrices

$$H_4^{(0)} = \begin{pmatrix} \sigma_{00} & \sigma_{10} & \sigma_{01} & \sigma_{11} \\ \sigma_{10} & \sigma_{20} & \sigma_{11} & \sigma_{21} \\ \sigma_{01} & \sigma_{11} & \sigma_{02} & \sigma_{12} \\ \sigma_{11} & \sigma_{21} & \sigma_{12} & \sigma_{22} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & \frac{1}{6} \\ 0 & 0 & \frac{1}{6} & 0 \\ 0 & \frac{1}{6} & 0 & 0 \\ \frac{1}{6} & 0 & 0 & \sigma_{22} \end{pmatrix},$$

$$H_4^{(1)} = \begin{pmatrix} \sigma_{10} & \sigma_{20} & \sigma_{11} & \sigma_{21} \\ \sigma_{20} & \sigma_{30} & \sigma_{21} & \sigma_{31} \\ \sigma_{11} & \sigma_{21} & \sigma_{12} & \sigma_{22} \\ \sigma_{21} & \sigma_{31} & \sigma_{22} & \sigma_{32} \end{pmatrix} = \begin{pmatrix} 0 & 0 & \frac{1}{6} & 0 \\ 0 & 1 & 0 & \sigma_{31} \\ \frac{1}{6} & 0 & 0 & \sigma_{22} \\ 0 & \sigma_{31} & \sigma_{22} & \sigma_{32} \end{pmatrix},$$

$$H_4^{(2)} = \begin{pmatrix} \sigma_{01} & \sigma_{11} & \sigma_{02} & \sigma_{12} \\ \sigma_{11} & \sigma_{21} & \sigma_{12} & \sigma_{22} \\ \sigma_{02} & \sigma_{12} & \sigma_{03} & \sigma_{13} \\ \sigma_{12} & \sigma_{22} & \sigma_{13} & \sigma_{23} \end{pmatrix} = \begin{pmatrix} 0 & \frac{1}{6} & 0 & 0 \\ \frac{1}{6} & 0 & 0 & \sigma_{22} \\ 0 & 0 & 0 & \sigma_{13} \\ 0 & \sigma_{22} & \sigma_{13} & \sigma_{23} \end{pmatrix}.$$

The missing σ_{ij} are determined from the commutativity of $H_4^{(1)}(H_4^{(0)})^{-1}$ and $H_4^{(2)}(H_4^{(0)})^{-1}$. They are found to equal

$$\sigma_{22} = 0, \ \sigma_{31} = {}^{5}\!/_{3}, \ \sigma_{32} = 1, \ \sigma_{13} = {}^{1}\!/_{6}, \ \sigma_{23} = 0.$$

Solving the generalized eigenvalue problem with $k = 1$ gives us the four eigenvectors

$$\hat{u}_1 = \begin{pmatrix} 1 \\ 4 \\ 1 \\ 4 \end{pmatrix}, \ \hat{u}_2 = \begin{pmatrix} 1 \\ -4 \\ 1 \\ -4 \end{pmatrix}, \ \hat{u}_3 = \begin{pmatrix} 1 \\ 2 \\ -1 \\ -2 \end{pmatrix}, \ \hat{u}_4 = \begin{pmatrix} 1 \\ -2 \\ -1 \\ 2 \end{pmatrix}.$$

Since

$$\hat{u}_i = (1, \lambda_{i1}, \lambda_{i2}, \lambda_{i1}\lambda_{i2})^T, \qquad i = 1, \ldots, 4,$$

we can deduce the Λ_i, $i = 1, \ldots, 4$, from these and subsequently solve the linear system (4.5) for the w_i. We finally obtain the decomposition

$$\begin{aligned} & p(x_0, x_1, x_2) \\ = \quad & \frac{1}{96}(x_0 + 4x_1 + x_2)^3 - \frac{1}{96}(x_0 - 4x_1 + x_2)^3 - \frac{1}{48}(x_0 + 2x_1 - x_2)^3 + \frac{1}{48}(x_0 - 2x_1 - x_2)^3. \end{aligned}$$

The boundary of I_3 contains 4 elements: $\partial I_3 = \{(2,0), (0,2), (2,1), (1,2)\}$. We compute the polynomial $\mathcal{V}(z_1, z_2)$ with each of these points in ∂I_3 as $\nu(4)$ and obtain:

$$\begin{aligned} \nu(4) &= (2,0), & \mathcal{V}(z_1, z_2) &= z_1^2 - 6z_2 - 10, \\ \nu(4) &= (0,2), & \mathcal{V}(z_1, z_2) &= z_2^2 - 1, \\ \nu(4) &= (2,1), & \mathcal{V}(z_1, z_2) &= z_1^2 z_2 - 10z_2 - 6, \\ \nu(4) &= (1,2), & \mathcal{V}(z_1, z_2) &= z_1(z_2^2 - 1). \end{aligned}$$

The intersection of the 4 zero curves delivers the points $(z_1, z_2) = (\lambda_{i1}, \lambda_{i2})$, $i = 1, \ldots, 4$, given by

$$(4,1),\ (-4,1),\ (2,-1),\ (-2,-1). \tag{8.2}$$

Let us now follow the parameterized approach described in Section 7, which is fundamentally different from the above. We write

$$\begin{aligned} p(1, x_1, x_2) &= \theta_1\theta_2 x^2 + \theta_1^3 x^3 & \qquad x_1 = \theta_1 x,\ x_2 = \theta_2 x, \\ &= C_2(\Theta)x^2 + C_3(\Theta)x^3 & \Theta = (\theta_1, \theta_2). \\ &= p(\Theta; x), \end{aligned}$$

Then the decomposition problem takes the form

$$p(\Theta; x) = \sum_{j=0}^{3} \left(\sum_{i=1}^{r} w_i \Lambda_i(\Theta)^j \right) \binom{d}{j} x^j, \qquad \Lambda_i(\Theta) = \lambda_{i1}\theta_1 + \lambda_{i2}\theta_2.$$

We have

$$S_0(\Theta) = 0,\ S_1(\Theta) = 0,\ S_2(\Theta) = \frac{1}{3}\theta_1\theta_2,\ S_3(\Theta) = \theta_1^3.$$

Construct

$$H_4^{(0)}(\Theta) = \begin{pmatrix} 0 & 0 & \frac{1}{3}\theta_1\theta_2 & \theta_1^3 \\ 0 & \frac{1}{3}\theta_1\theta_2 & \theta_1^3 & S_4(\Theta) \\ \frac{1}{3}\theta_1\theta_2 & \theta_1^3 & S_4(\Theta) & S_5(\Theta) \\ \theta_1^3 & S_4(\Theta) & S_5(\Theta) & S_6(\Theta) \end{pmatrix},$$

$$H_4^{(1)}(\Theta) = \begin{pmatrix} 0 & \frac{1}{3}\theta_1\theta_2 & \theta_1^3 & S_4(\Theta) \\ \frac{1}{3}\theta_1\theta_2 & \theta_1^3 & S_4(\Theta) & S_5(\Theta) \\ \theta_1^3 & S_4(\Theta) & S_5(\Theta) & S_6(\Theta) \\ S_4(\Theta) & S_5(\Theta) & S_6(\Theta) & S_7(\Theta) \end{pmatrix}.$$

The generalized eigenvalues $\Lambda_i(\Theta)$ are obtained from the generalized eigenvalue problem (7.5), after the missing moments are determined from the commutativity of the matrix products $H_4^{(2)}(\Theta)(H_4^{(0)}(\Theta))^{-1}$ and $H_4^{(1)}(\Theta)(H_4^{(0)}(\Theta))^{-1}$. The latter are given by

$$\begin{aligned} S_4(\Theta) &= \frac{20}{3}\theta_1^3\theta_2 + \frac{2}{3}\theta_1\theta_2^3, \\ S_5(\Theta) &= 20\theta_1^5 + 10\theta_1^3\theta_2^2, \\ S_6(\Theta) &= \frac{100}{3}\theta_1^3\theta_2^3 + \theta_1\theta_2^5 + 136\theta_1^5\theta_2, \\ S_7(\Theta) &= 336\theta_1^7 + 420\theta_1^5\theta_2^2 + 35\theta_1^3\theta_2^4. \end{aligned}$$

We remark that for this computation also $H_2^{(2)}(\Theta)$ needs to be constructed. The parameterized generalized eigenvalues equal

$$\begin{aligned} \Lambda_1(\Theta) &= 4\theta_1 + \theta_2, \\ \Lambda_2(\Theta) &= -4\theta_1 + \theta_2, \\ \Lambda_3(\Theta) &= 2\theta_1 - \theta_2, \\ \Lambda_4(\Theta) &= -2\theta_1 - \theta_2. \end{aligned}$$

These can also be found from the factorization of $V(\Theta; z)$ which is given by

$$V(\Theta; z) = (4\theta_1 + \theta_2 - z)(-4\theta_1 + \theta_2 - z)(2\theta_1 - \theta_2 - z)(-2\theta_1 - \theta_2 - z).$$

For the coefficients in the decomposition we can either solve (7.1) for the unknown w_i or apply the functional Γ to $V(\Theta; z)/\left[(z - \Lambda_i(\Theta))\, V'(\Theta; \Lambda_i(\Theta))\right]$ and obtain the $A_i(\Theta)$. The latter method gives

$$w_1 = A_1(\Theta) = \frac{1}{96}, \qquad w_2 = A_2(\Theta) = \frac{-1}{96}, \qquad w_3 = A_3(\Theta) = \frac{-1}{48}, \qquad w_4 = A_4(\Theta) = \frac{1}{48}.$$

As indicated in [12], the $A_i(\Theta)$ are independent of Θ in this case. Combining the parameterized zeroes $\Lambda_i(\Theta)$ of $V(\Theta;z)$ with the $A_i(\Theta)$ eventually gives the same decomposition

$$\begin{aligned} p(\Theta;z) \quad &= \quad \frac{1}{96}\left(x_0+(4\theta_1+\theta_2)x\right)^3-\frac{1}{96}\left(x_0+(-4\theta_1+\theta_2)x\right)^3 \\ &\quad -\frac{1}{48}\left(x_0+(2\theta_1-\theta_2)x\right)^3+\frac{1}{48}\left(x_0+(-2\theta_1-\theta_2)x\right)^3, \qquad x_1=\theta_1 x,\ x_2=\theta_2 x. \end{aligned} \tag{8.3}$$

For completeness, we mention that

$$\left|H_5^{(0)}(\Theta)\right| \equiv 0$$

independent of Θ, thereby indicating that $r=4$ indeed.

The question remains whether $r=4$ for all directional vectors. Let's take a look at the polynomial $|H_4^{(0)}(\Theta)|$ of degree 12,

$$\left|H_4^{(0)}(\Theta)\right| = \frac{1}{81}\theta_1^4(\theta_1+\theta_2)^2(3\theta_1+\theta_2)^2(\theta_1-\theta_2)^2(3\theta_1-\theta_2)^2.$$

When either $\theta_2=\pm\theta_1$ or $\theta_2=\pm 3\theta_1$, one of the four terms in the decomposition (8.3) reduces to one of the other three terms and so after simplification only 3 terms remain, reducing r from 4 to 3 for these four directional vectors.

To conclude we reassure the reader that the symbolic expressions in this new method need not be computed using symbolic methods. So the methods in the sections 5 and 7 do not need a different computational environment. All multivariate polynomials involved, $|H_r^{(k)}(\Theta)|$, $|H_R^{(0)}(\Theta)|$ as well as missing $S_j(\Theta)$ and the $\Lambda_i(\Theta)$ for $1\le i\le r$, can easily be computed using numeric samples of these polynomials in combination with Prony's method. As an example we compute $|H_4^{(0)}(\Theta)|$ in this way ([13, 19]).

Inspecting $|H_4^{(0)}(\Theta)|$ tells us that a priori bounds on the partial degrees of this polynomial determinant in θ_1 and θ_2, are given by 16 and 12. As mutually coprime bounds, 17 and 13 can be taken. Evaluating $|H_4^{(0)}(\Theta)|$ at $\theta_1=\exp(2\pi \mathrm{i}(10/17))$ and $\theta_2=\exp(2\pi\mathrm{i}(3/13))$ and following [13, 19] we find that $|H_4^{(0)}(\Theta)|$ only contains 5 terms, namely

$$|H_4^{(0)}(\Theta)| = d_1\theta_1^8\theta_2^4+d_2\theta_1^{10}\theta_2^2+d_3\theta_1^4\theta_2^8+d_4\theta_1^6\theta_2^6+d_5\theta_1^{12}.$$

Numerical values for the real-valued coefficients d_i, $i=1,\ldots,5$, have imaginary parts of the order of 10^{-10}:

$$d_1 = +1.45679012358018 - 3.38031824831369\times 10^{-11}\mathrm{i},$$

$$d_2 = -2.22222222139352 - 2.63621714615090 \times 10^{-10}\mathrm{i},$$
$$d_3 = +0.01234567926619 - 5.75165916323563 \times 10^{-10}\mathrm{i},$$
$$d_4 = -0.24691358095272 + 7.01071048519302 \times 10^{-10}\mathrm{i},$$
$$d_5 = +0.99999999949986 + 1.71519431770272 \times 10^{-10}\mathrm{i}.$$

Computing the regular continued fraction representation ([17, p. 175])

$$\Re(d_i) = \delta_0^{(i)} + \cfrac{1}{\delta_1^{(i)} + \cfrac{1}{\delta_2^{(i)} + \cdots}}$$

of the real parts of the coefficients d_i, delivers for rather small m_i a partial denominator $\delta_{m_i}^{(i)}$ that is very large, indicating that the real part of d_i is close to a rational number:

$$d_1 = 1 + \cfrac{1}{2 + \cfrac{1}{5 + \cfrac{1}{3 + \cfrac{1}{2 + \cfrac{1}{1235237 + \ldots}}}}},$$

$$d_2 = -2 - \cfrac{1}{4 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{14897602 + \ldots}}}},$$

$$d_3 = \cfrac{1}{80 + \cfrac{1}{1 + \cfrac{1}{600429 + \ldots}}},$$

$$d_4 = -\cfrac{1}{4 + \cfrac{1}{20 + \cfrac{1}{215945 + \ldots}}},$$

$$d_5 = \cfrac{1}{1 + \cfrac{1}{1999439551 + \dots}}.$$

Within truncation error bounds of less than 8.3×10^{-10} for the real parts of the d_i, we find

$$|H_4^{(0)}(\Theta)| = \frac{118}{81}\theta_1^8\theta_2^4 - \frac{20}{9}\theta_1^{10}\theta_2^2 + \frac{1}{81}\theta_1^4\theta_2^8 - \frac{20}{81}\theta_1^6\theta_2^6 + \theta_1^{12}.$$

References

[1] Bajzer Z., A. C. Myers, S. S. Sedarous and F. G. Prendergast (1989)
Padé-Laplace method for analysis of fluorescence intensity decay, Biophys J. 56(1), 79–93.

[2] Benouahmane B. and A. Cuyt (2000)
Multivariate orthogonal polynomials, homogeneous Padé approximants and Gaussian cubature, Numer. Algorithms 24, 1–15.

[3] Benouahmane B., A. Cuyt and I. Yaman (2014)
Near-minimal cubature formulae on the disk, Technical report, Universiteit Antwerpen.

[4] Bernardi A., J. Brachat, P. Comon and B. Mourrain (2013)
General tensor decomposition, moment matrices and applications, J. Symbolic Comput. 52, 51–71.

[5] Brachat J., P. Comon, B. Mourrain and E. P. Tsigaridas (2010)
Symmetric tensor decomposition, Linear Algebra Appl. 433(11-12), 1851–1872.

[6] Brezinski C. (1980)
Padé type Approximation and General Orthogonal Polynomials, International Series of Numerical Mathematics, 50. Birkhäuser Verlag, Basel-Boston, Mass.

[7] Cuyt A. (1983)
Multivariate Padé approximants, J. Math. Anal. Appl. 96, 283–293.

[8] Cuyt A. (1984)
Padé Approximants for Operators: Theory and Applications, Lecture Notes in Mathematics, 1065. Springer-Verlag, Berlin.

[9] Cuyt A. (1985)
A review of multivariate Padé approximation theory, J. Comput. Appl. Math. 12-13, 221–232.

[10] Cuyt A. (1986)
Multivariate Padé approximants revisited, BIT 26, 71–79.

[11] Cuyt A. (1992)
Padé approximation in one and more variables, in: Approximation theory, spline functions and applications (Maratea, 1991), pp. 41-68, NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci., 356, Kluwer Acad. Publ., Dordrecht.

[12] Cuyt A., B. Benouahmane, Hamsapriye and I. Yaman (2011)
Symbolic-numeric Gaussian cubature rules, Appl. Numer. Math. 61, 929–945.

[13] Cuyt A. and W.-s. Lee (2008)
A new algorithm for sparse interpolation of multivariate polynomials, Theoret. Comput. Sci. 409, 180–185.

[14] Cuyt A. and W.-s. Lee (2015)
How many samples are required in higher dimensional exponential analysis?, Technical report, University of Antwerp.

[15] Cuyt A. and W.-s. Lee (2016)
Sparse interpolation and Rational approximation, Contemp. Math. 661, 229–242.

[16] Cuyt A. and D. S. Lubinsky (1997)
A de Montessus theorem for multivariate homogeneous Padé approximants. The heritage of P. L. Chebyshev: a Festschrift in honor of the 70th birthday of T. J. Rivlin, Ann. Numer. Math. 4, 217–228.

[17] Cuyt A., V. B. Petersen, B. Verdonk, H. Waadeland and W. B. Jones (2008)
Handbook of Continued Fractions for Special Functions, Springer, New York.

[18] Diederichs B. and A. Iske (2015)
Parameter Estimation for Bivariate Exponential Sums, in: Proceedings of the 2015 International Conference Sampling Theory and Applications, SampTA 2015, pp.493–497, IEEE, Washington, DC, USA.

[19] Giesbrecht M., G. Labahn and W.-s. Lee (2006)
Symbolic-numeric sparse interpolation of multivariate polynomials, ISSAC 2006, 116–123.

[20] Golub G. H., P. Milanfar and J. Varah (1999)
A stable numerical method for inverting shape from moments, SIAM J. Sci. Comput. 21, 1222–1243.

[21] Henrici P. (1974)
Applied and Computational Complex Analysis. Volume 1: Power seriesintegrationconformal mappinglocation of zeros. Pure and Applied Mathematics, Wiley-Interscience [John Wiley & Sons], New York-London-Sydney.

[22] Hildebrand F. B. (1956)
Introduction to Numerical Analysis, Mc Graw Hill Book Company, Inc., New York-Toronto-London.

[23] Kaltofen E., W.-s. Lee and A. Lobo (2000)
Early termination in Ben-Or/Tiwari sparse interpolation and a hybrid of Zippel's algorithm, in: Proceedings of the 2000 International Symposium on Symbolic and Algebraic Computation (St. Andrews), pp. 192–201, ACM, New York.

[24] Kolda T. G. and B. W. Bader (2009)
Tensor decompositions and applications, SIAM Rev. 51(3), 455–500.

[25] Kunis S., P. Thomas, T. Römer and U. von der Ohe (2016)
A multivariate generalization of Prony's method, Linear Algebra Appl. 490, 31–47.

[26] Lopez Q. M. (2011)
Symmetric tensor decomposition and algorithms, Master's thesis, Universidad Complutense de Madrid, Spain.

[27] Potts D. and M. Tasche (2013)
Parameter estimation for multivariate exponential sums, Electron. Trans. Numer. Anal. 40, 204–224.

[28] Tomas S. (2016)
Prony's method in several variables, Technical report, Universität Passau.

[29] Werner H. (1986)
Multivariate Padé approximation, Numer. Math. 48, 429–440.

Annie Cuyt, Wen-shin Lee and Xianglan Yang,
Dept Mathematics and Computer Science,
University of Antwerp,
Middelheimlaan 1,
B-2020 Antwerpen, Belgium.
annie.cuyt@uantwerpen.be
wen-shin.lee@uantwerpen.be
xianglan.yang@uantwerpen.be