Algebra for the analysis of tensors and moment sequences

Bernard Mourrain
Inria of Univ. Côte d'Azur, Sophia Antipolis
Bernard.Mourrain@inria.fr

TENORS Learning week 1, 3-6 February 2025

Examples of problems

Guessing sequences

Given a sequence of values

$$\sigma_0, \sigma_1, \ldots, \sigma_s \in \mathbb{C},$$

find/guess the values of σ_n for all $n \in \mathbb{N}$.

☞ Find $r \in \mathbb{N}, \omega_i, \xi_i \in \mathbb{C}$ such that $\sigma_n = \sum_1^{\mathrm{r}} \omega_{\mathrm{i}} \, \xi_{\mathrm{i}}{}^n$, for all $n \in \mathbb{N}$.

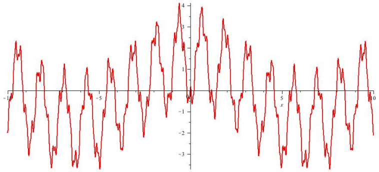
Example: 0, 1, 1, 2, 3, 5, 8, 13, ...

Solution:

- Find a recurrence relation valid for the first terms: $\sigma_{k+2} - \sigma_{k+1} - \sigma_k = 0$.
- Find the roots $\xi_1 = \frac{1+\sqrt{5}}{2}$, $\xi_2 = \frac{1-\sqrt{5}}{2}$ (golden numbers) of the characteristic polynomial: $x^2 - x - 1 = 0$.
- Deduce $\sigma_n = \frac{1}{\sqrt{5}}(\frac{1+\sqrt{5}}{2})^n - \frac{1}{\sqrt{5}}(\frac{1-\sqrt{5}}{2})^n$.

Reconstruction of signals

Given a function or signal $f(t)$:



decompose it as

$$f(t) = \sum_{i=1}^{r'} (a_i cos(\mu_i t) + b_i sin(\mu_i t)) e^{\nu_i t} = \sum_{i=1}^{r} \omega_i e^{\zeta_i t}$$

Blind identification

Observing $\boldsymbol{y}(t)$ with

$$\boldsymbol{y}(t) = H\ \boldsymbol{s}(t)$$

☞ **find** H **and** s(t)

- If the sources are statistically independent, using the high order statistics $\mathbb{E}(y_i\, y_j\, y_k \cdots)$ of the signal $\boldsymbol{y}(t)$, **decompose the symmetric tensor** $T = \sum_{i,j,k,\ldots} \mathbb{E}(y_i\, y_j\, y_k \cdots)\, x_i\, x_j\, x_k \cdots = \sum_{|\alpha|=d} \binom{d}{\alpha} \mathbb{E}(\boldsymbol{y}^\alpha)\boldsymbol{x}^\alpha$ as
$$T(\boldsymbol{x}) = \sum_{i=1}^{r} (H_i, \boldsymbol{x})^d$$
- Deduce the geometry of the sources $H = [H_1, \ldots, H_r]$ and $\boldsymbol{s}(t)$.

Tensor of matrix multiplication

Given two matrices $X = \begin{pmatrix} x_1 & x_2 \\ x_3 & x_4 \end{pmatrix}$, $Y = \begin{pmatrix} y_1 & y_2 \\ y_3 & y_4 \end{pmatrix} \in \mathbb{R}^{2\times 2} = \mathbb{R}^4$ the product $X\,Y$ is $Z = \begin{pmatrix} z_1' & z_2' \\ z_3' & z_4' \end{pmatrix}$ with

$$
\begin{aligned}
z_1' &= x_1 y_1 + x_2 y_3 \\
z_2' &= x_1 y_2 + x_2 y_4 \\
z_3' &= x_3 y_2 + x_4 y_4 \\
z_4' &= x_3 y_2 + x_4 y_4
\end{aligned}
$$

It defines the following tensor in $\mathbb{R}^4 \otimes \mathbb{R}^4 \otimes \mathbb{R}^4 = \mathbb{R}^{4\times 4\times 4}$:

$$T_{2,2} = (x_1 y_1 + x_2 y_3)\, z_1 + (x_1 y_2 + x_2 y_4)\, z_2 + (x_3 y_2 + x_4 y_4)\, z_3 + (x_3 y_2 + x_4 y_4)\, z_4$$

Decompositing minimally as

$$T_{2,2} = \sum_{i=1}^{r} l_i(x_1, \ldots, x_4) m_i(y_1, \ldots, y_4) n_i(z_1, \ldots, z_4)$$

where l_i, m_i, n_i are linear in x, y, z, allows to compute the matrix product **more efficiently**.

- Such a decomposition exists with $r = 7$ (< 8) for 2×2 matrices.
- For 3×3 matrices, $r = 22$ (< 27).
- These decompositions are useful to bound the **exponent of linear algebra**:

$$\text{Multiplication of } n \times n \text{ matrices} \in \mathcal{O}(n^\omega)$$

where $\omega < 2.38\ldots$ (Conjecture: $\omega = 2$!).

Multilinear, symmetric tensors and moment sequences over $\mathbb{K} = \mathbb{R}, \mathbb{C}, \ldots$

Multilinear tensors

- A multilinear tensor is of the form

$$T = [t_{i_1,\dots,i_l}] \in \mathbb{K}^{n_1 \times \cdots \times n_l}$$

as a **multi-dimensional array**

$$T(\mathrm{x}) = \sum_{i_1,\dots,i_l} t_{i_1,\dots,i_l} x_{1,i_1} \cdots x_{l,i_l}$$

as a **multilinear polynomial** in variables $x_{j,i}$

- Vector space denoted $\mathbb{K}^{n_1} \otimes \cdots \otimes \mathbb{K}^{n_l}$, dimension $= n_1 \times \cdots \times n_l$.

Symmetric tensors of order d

$T \in \mathbb{K}^{\overbrace{n \times \cdots n}^{d}}$ is **symmetric** iff $t_{i_{\tau(1)},\dots,i_{\tau(d)}} = t_{i_1,\dots,i_d}$ for any permutation τ of $[1,\dots,d]$.

- For any $i_1,\dots,i_d \in 1:n$ with $\#\{i_k = j\} = \alpha_j$, we have

$$t_{i_1,\dots,i_d} = t_{\overbrace{1,\dots,1}^{\alpha_1},\overbrace{2,\dots,2}^{\alpha_2},\dots,\overbrace{n,\dots,n}^{\alpha_n}} =: t_{\alpha_1,\dots,\alpha_n}$$

with $|\alpha| := \alpha_1 + \cdots + \alpha_n = d$.

- If T is symmetric and $\mathrm{x}_1 = \ldots = \mathrm{x}_d = \mathrm{x} = (x_1,\dots,x_n)$,

$$T(\mathrm{x}) = \sum_{i_1,\dots,i_l} t_{i_1,\dots,i_l} x_{i_1} \cdots x_{i_l} = \sum_{|\alpha|=d} t_\alpha \binom{d}{\alpha} \mathrm{x}^\alpha$$

where $\mathrm{x}^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$, $\binom{d}{\alpha} = \frac{d!}{\alpha_1! \cdots \alpha_n!}$.

= **homogeneous polynomial** of degree d in n variables $\mathrm{x} = (x_1,\dots,x_n)$

- Space denoted $\mathcal{S}_{n,d}$, of dimension $s_{n,d} := \binom{n-1+d}{d}$

Multilinear tensor

```
In [1]: T = [i+j+k for i in 0:2, j in 0:2,  k in 0:2 ]

Out[1]: 3×3×3 Array{Int64,3}:
        [:, :, 1] =
         0  1  2
         1  2  3
         2  3  4

        [:, :, 2] =
         1  2  3
         2  3  4
         3  4  5

        [:, :, 3] =
         2  3  4
         3  4  5
         4  5  6

In [2]: using DynamicPolynomials;
        X = @polyvar x0 x1 x2; Y = @polyvar y0 y1 y2; Z = @polyvar z0 z1 z2;

In [3]: Txyz = sum( T[i,j,k]*X[i+1]*Y[j+1]*Z[k+1] for i in 0:2, j in 0:2,  k in 0:2 )
```

Out[3]:

$$x_0y_0z1 + 2x_0y_0z_2 + x_0y_1z_0 + 2x_0y_1z_1 + 3x_0y_1z_2 + 2x_0y_2z_0 + 3x_0y_2z_1 + 4x_0y_2z_2 + x_1y_0z_0 + 2x_1y_0z_1 + 3x_1y_0z_2 + 2x_1y_1z_0 + 3x_1y_1z_1 + 4x_1y_1z_2 + 3x_1y_2z_0 + 4x_1y_2z_1 + 5x_1y_2z_2 + 2x_2y_0z_0 + 3x_2y_0z_1 + 4x_2y_0z_2 + 3x_2y_1z_0 + 4x_2y_1z_1 + 5x_2y_1z_2 + 4x_2y_2z_0 + 5x_2y_2z_1 + 6x_2y_2z_2$$

Symmetric tensor

```
In [4]: F = sum( T[i,j,k]*X[i+1]*X[j+1]*X[k+1] for i in 0:2, j in 0:2,  k in 0:2 )
```

Out[4]:

$$3x_0^2x_1 + 6x_0^2x_2 + 6x_0x_1^2 + 18x_0x_1x_2 + 12x_0x_2^2 + 3x_1^3 + 12x_1^2x_2 + 15x_1x_2^2 + 6x_2^3$$

Apolar product

Apolar product: For $F = \sum_{|\alpha|=d} f_\alpha \mathrm{x}^\alpha$, $F' = \sum_{|\alpha|=d} f'_\alpha \mathrm{x}^\alpha \in \mathcal{S}_{n,d}$,

$$\langle F, F' \rangle_d = \sum_{|\alpha|=d} \binom{d}{\alpha}^{-1} f_\alpha f'_\alpha.$$

Properties:

- $\langle F, (\boldsymbol{u} \cdot \mathrm{x})^d \rangle_d = F(\boldsymbol{u})$
- $\langle F, (\boldsymbol{v} \cdot x)(\boldsymbol{u} \cdot \boldsymbol{x})^{d-1} \rangle = \frac{1}{d} D_{\boldsymbol{v}} F(\boldsymbol{u})$
- $\langle F, (\boldsymbol{v}_1 \cdot x) \cdots (\boldsymbol{v}_k \cdot \boldsymbol{x})(\boldsymbol{u} \cdot \boldsymbol{x})^{d-k} \rangle = \frac{(d-k)!}{d!} D_{\boldsymbol{v}_1} \cdots D_{\boldsymbol{v}_k} F(\boldsymbol{u})$

Taylor expension: For an **orthonormal** basis $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n$ of $\mathbb{K}^n$:

$$\begin{aligned} F &= \sum_{|\alpha|=d} \frac{1}{\alpha_2! \cdots \alpha_n!} (\boldsymbol{u}_1 \cdot \boldsymbol{x})^{\alpha_1} \cdots (\boldsymbol{u}_n \cdot \boldsymbol{x})^{\alpha_n} D_{\boldsymbol{u}_2}^{\alpha_2} \cdots D_{\boldsymbol{u}_n}^{\alpha_n} F(\boldsymbol{u}_1) \\ &= \sum_{|\alpha|=d} \binom{d}{\alpha} (\boldsymbol{u}_1 \cdot \boldsymbol{x})^{\alpha_1} \cdots (\boldsymbol{u}_n \cdot \boldsymbol{x})^{\alpha_n} \langle F, (\boldsymbol{u}_1 \cdot \boldsymbol{x})^{\alpha_1} \cdots (\boldsymbol{u}_n \cdot \boldsymbol{x})^{\alpha_n} \rangle_d \end{aligned}$$

☞ $\left(\sqrt{\binom{d}{\alpha}} \prod_{i=1}^n (\boldsymbol{u}_i \cdot \boldsymbol{x})^{\alpha_i}\right)_{|\alpha|=d}$ **orthonormal basis of $\mathcal{S}_{n,d}$.**

Apolar duality

Definition

For $F \in \mathcal{S}_{n,d}$ and $h \in \mathcal{S}_{n,k}$,

- $F^* : p \in \mathcal{S}_{n,d} \mapsto \langle F, p \rangle \in \mathbb{K}$ is a linear functional $\in \mathcal{S}_{n,d}^*$
- $q \star F^* : p \in \mathcal{S}_{n,d-k} \mapsto \langle F, h\, p \rangle$

Theorem: **Weighted Sum of Evaluations (WSE)**

$$F(\boldsymbol{x}) = \sum_{i=1}^{r} \omega_i (\xi_i, \boldsymbol{x})^d \Leftrightarrow F^* = (\sum_{i=1}^{r} \omega_i \mathrm{e}_{\xi_i})^{[d]}.$$

where $\mathrm{e}_{\xi_i} : p \in S^d(\mathbb{K}^n) \mapsto p(\xi)$ is the **evaluation** ($=$ **Dirac** measure δ_ξ) at ξ_i, and $\Lambda^{[d]}$ is the restriction in degree d.

Sequences, series, duality (1D)

Sequences: $\sigma = (\sigma_k)_{k\in\mathbb{N}} \in \mathbb{K}^{\mathbb{N}}$ indexed by $k \in \mathbb{N}$.

Formal power series:

$$\sigma(y) = \sum_{k=0}^{\infty} \sigma_k\, y^k \in \mathbb{K}[[y]] \qquad\qquad \sigma(z) = \sum_{k=0}^{\infty} \sigma_k\, \frac{z^k}{k!} \in \mathbb{K}[[z]]$$

Linear functionals: $\mathbb{K}[x]^* = \{\Lambda : \mathbb{K}[x] \to \mathbb{K} \text{ linear}\}$.

$$\Lambda : p = \sum_{i=0}^{d} p_i x^i \in \mathbb{K}[x] \mapsto \langle \Lambda | p\rangle = \sum_{i\in\mathbb{N}} \Lambda_i p_i$$

Structure of $\mathbb{K}[x]$-module: $p \star \Lambda : q \mapsto \Lambda(p\, q)$.

$$\begin{array}{rclrcl} x \star \sigma(y) & = & \pi_+(y^{-1})\sigma(y)) & x \star \sigma(z) & = & \sum_{k=1}^{\infty} \sigma_k \frac{z^{k-1}}{(k-1)!} = \partial(\sigma(z)) \\ p(x) \star \sigma(y) & = & \pi_+(p(y^{-1})(\sigma(y))) & p(x) \star \sigma(z) & = & p(\partial)(\sigma(z)) \end{array}$$

Examples:

- $p \mapsto$ coefficient of x^i in $p = \frac{1}{i!}\partial^i(p)(0)$ represented as

$$y^i \quad \text{or} \quad \frac{1}{i!}z^i$$

  (y^k) **(resp.** $(\frac{z^k}{k!})$**)** is the dual basis of the monomial basis $(x^k)_{k\in\mathbb{N}}$.

  $x \star y^i = y^{i-1}$, $\quad x \star z^i = iz^{i-1}$ for $i > 0$ ($= 0$ for $i = 0$)

- $e_\zeta : p \mapsto p(\zeta)$ **evaluation** at ζ represented as

$$e_\zeta(y) = \sum_{k=0}^{\infty} \zeta^k y^k = \frac{1}{1-\zeta y} \in \mathbb{K}[[y]] \text{ or } e_\zeta(z) = \sum_{k=0}^{\infty} \zeta^k \frac{z^k}{k!} = e^{\zeta z} \in \mathbb{K}[[z]]$$

  $x \star e_\zeta = \zeta\, e_\zeta$

Sequences, series, duality (nD)

- **Multi-index sequences:** $\sigma = (\sigma_\alpha)_{\alpha\in\mathbb{N}^n} \in \mathbb{K}^{\mathbb{N}^n}$ indexed by $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$, called a **moment sequence**.

- **Formal power series:**

$$\sigma(\mathrm{y}) = \textstyle\sum_{\alpha\in\mathbb{N}^n} \sigma_\alpha \mathrm{y}^\alpha \in \mathbb{K}[[y_1, \dots, y_n]] \qquad \sigma(\mathrm{z}) = \sum_{\alpha\in\mathbb{N}^\mathrm{n}} \sigma_\alpha \frac{\mathrm{z}^\alpha}{\alpha!} \in \mathbb{K}[[\mathrm{z}_1, \dots, \mathrm{z}_\mathrm{n}]]$$

where $\alpha! = \prod \alpha_i!$ for $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}$.

- **Linear functionals:** $\Lambda \in R^* = \{\Lambda : R \to \mathbb{K}, \textbf{linear}\}$ where $R = \mathbb{K}[x_1, \dots, x_n]$

$$\Lambda : p = \sum_\alpha p_\alpha \mathrm{x}^\alpha \mapsto \langle\Lambda|p\rangle = \sum_\alpha \Lambda_\alpha p_\alpha$$

The coefficients $\langle\Lambda|\mathrm{x}^\alpha\rangle = \Lambda_\alpha \in \mathbb{K}$, $\alpha \in \mathbb{N}^n$ are called the **moments** of Λ.

$(\boldsymbol{y}^\alpha)_{\alpha\in\mathbb{N}^n}$ (resp. $(\frac{1}{\alpha!}\boldsymbol{z}^\alpha)_{\alpha\in\mathbb{N}^n}$) dual basis in R^* of the monomial basis $(\boldsymbol{x}^\alpha)_{\alpha\in\mathbb{N}^n}$

- **Structure of R-module:** $\forall p \in R, \sigma \in R^*$, $p \star \sigma : q \mapsto \langle\sigma|p\,q\rangle$:

$$p \star \sigma = \pi_+(p(y_1^{-1}, \dots, y_n^{-1})\sigma(\mathrm{y})) \qquad \mathrm{p} \star \sigma = \mathrm{p}(\partial_1, \dots, \partial_\mathrm{n})(\sigma)(\mathrm{z})$$

Truncated moment sequences

For $\boldsymbol{x} = (x_1, \ldots, x_n)$, $\mathcal{R}_n := \mathbb{K}[\boldsymbol{x}]$, $\mathcal{R}_{n,d} := \mathbb{K}[\boldsymbol{x}]_{\leq d}$ spanned by $\boldsymbol{x}^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$ with $|\alpha| = \alpha_1 + \cdots + \alpha_n \leq d$.

- **Truncated moment sequences:** $(\sigma_\alpha)_{|\alpha| \leq d}$
- **Truncated linear functionals:** $\mathcal{R}_{n,d}^* = \{\Lambda : \mathcal{R}_{n,d} \to \mathbb{K},\ \text{linear}\ \}$.
- **Truncated series:** For $\sigma(\boldsymbol{y}) = \sum_{\alpha \in \mathbb{N}^n} \sigma_\alpha \boldsymbol{y}^\alpha \in \mathbb{K}[[\boldsymbol{y}]]$,

  or $\sigma(\boldsymbol{z}) = \sum_{\alpha \in \mathbb{N}^n} \sigma_\alpha \frac{\boldsymbol{z}^\alpha}{\alpha!} \in \mathbb{K}[[\boldsymbol{z}]]_{\leq d}$,

$$\sigma(\boldsymbol{y})^{[d]} = \sum_{|\alpha| \leq d} \sigma_\alpha \boldsymbol{y}^\alpha \in \mathbb{K}[\boldsymbol{y}]_{\leq d} \text{ or } \sigma(\boldsymbol{z})^{[d]} = \sum_{|\alpha| \leq d} \sigma_\alpha \frac{\boldsymbol{z}^\alpha}{\alpha!} \in \mathbb{K}[\boldsymbol{z}]_{\leq d}$$

- **Coproduct:** For $p \in \mathcal{R}_{n,k}, \Lambda \in \mathcal{R}_{n,d}^*$, $p \star \Lambda \in \mathcal{R}_{n,d-k}^*$
- **From tensors to moment sequences:**

$F = \sum_{|\alpha|=d} F_\alpha x_0^{\alpha_0} x_1^{\alpha_1} \cdots x_n^{\alpha_n} \in S_{n+1,d} \Rightarrow \check{F} = \sum_{|\alpha|=d} F_\alpha \binom{d}{\alpha}^{-1} y_1^{\alpha_1} \cdots y_n^{\alpha_n} \in \mathcal{R}_{n,d}^*$ s.t.

$$F^* = \check{F} \circ \iota_0, \qquad \check{F} = F^* \circ h_{d,x_0}$$

where $\iota_0 : p(x_0, \ldots, x_n) \in \mathcal{S}_{n+1,d} \mapsto p(1, x_1, \ldots, x_n) \in \mathcal{R}_{n,d}$, $h_{d,x_0} : p \in \mathcal{R}_{n,d} \mapsto x_0^d p(\frac{\boldsymbol{x}}{x_0}) \in \mathcal{S}_{n+1,d}$.

Duality

We describe the relations between tensors, linear functionals, moment sequences and duality in action on effective examples.

See also https://github.com/tenors-network/TENORS-L1-Inria/blob/main/courses/Algebra-for-Analysis-of-Tensors/Duality.ipynb

[1]:
```
using DynamicPolynomials, MultivariateSeries
X = @polyvar x0 x1 x2
d = 3
F = x0^d + 2.0* (x0+x1-x2)^d
```

[1]:

$$-2.0x2^3+6.0x1x2^2-6.0x1^2x2+2.0x1^3+6.0x0x2^2-12.0x0x1x2+6.0x0x1^2-6.0x0^2x2+6.0x0^2x1+3.0x0^3$$

We compute the linear functional $F^* \in (\mathcal{S}_{3,d})^*$ by apolarity:

[2]:
```
Fstar = dual(F,d)
```

[2]:
```
-2.0dx2^3 + 2.0dx1*dx2^2 - 2.0dx1^2dx2 + 2.0dx1^3 + 2.0dx0*dx2^2 -
2.0dx0*dx1*dx2 + 2.0dx0*dx1^2 - 2.0dx0^2dx2 + 2.0dx0^2dx1 + 3.0dx0^3
```

The variables of the dual basis of the monomials basis are denoted dx_i.

We compute now the **affine** polynomial obtained by the substitution $x_0 \Rightarrow 1$:

[3]:
```
f = subs(F,x0=>1)
```

[3]:

$$3.0-6.0x2+6.0x1+6.0x2^2-12.0x1x2+6.0x1^2-2.0x2^3+6.0x1x2^2-6.0x1^2x2+2.0x1^3$$

Its dual $s = \check{f} \in (\mathcal{R}_{n,d})^*$ is:

[4]:
```
s = dual(f,d)
```

[4]:
```
3.0 - 2.0dx2 + 2.0dx1 + 2.0dx2^2 - 2.0dx1*dx2 + 2.0dx1^2 - 2.0dx2^3 +
2.0dx1*dx2^2 - 2.0dx1^2dx2 + 2.0dx1^3
```

We apply the linear functional s on x_1:

[5]:
```
dot(s,x1)
```

[5]:
```
2.0
```

We can notice that it coincides with $\langle \mathbf{e}_{0,0} + 2\,\mathbf{e}_{1,-1} | x_1 \rangle$ where $\mathbf{e}_{a,b}$ is the evaluation at the point (a, b).

We compute $x_2 \star s$, which coincides with $(-2\,\mathbf{e}_{1,-1})^{[2]}$ (i.e. $-2\,\mathbf{e}_{1,-1}$ truncated in degree ≤ 2)

[6]:
```
s2 = x2*s
```

[6]:
```
-2.0 + 2.0dx2 - 2.0dx1 - 2.0dx2^2 + 2.0dx1*dx2 - 2.0dx1^2
```

as we can check, when applying the linear functional $x_2 \star s$ on $x_1^2 + x_2^2$:

[7]:
```
dot(s2, x1^2+x2^2)
```

[7]:
```
-4.0
```

Decomposition of tensors and moment sequences

Matrices

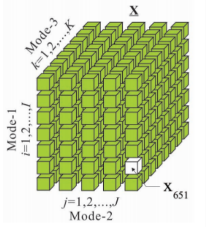
$M \in \mathbb{K}^{n_1 \times n_2} = \mathbb{K}^{n_1} \otimes \mathbb{K}^{n_2}$ is of **rank** r iff there exist $U \in \mathbb{K}^{n_1 \times n_1}$, $V \in \mathbb{K}^{n_2 \times n_2}$ invertible and Σ_r diagonal invertible s.t.

$$M = U \begin{pmatrix} \Sigma_r & 0 \\ 0 & 0 \end{pmatrix} V^t$$

- Σ_r not unique
- $\Sigma_r = I_r$ for some U, V.
- U, V unitary $\Rightarrow$ Singular Value Decomposition
- U, V are **eigenvectors** of $M\,M^t$ (resp. $M^t M$)
- Best low rank approximation from truncated SVD

Multilinear tensors of $\mathbb{K} = \mathbb{R}, \mathbb{C}, \ldots$

A tri-linear tensor $T \in \mathbb{K}^{n_1 \times n_2 \times n_3} = \mathbb{K}^{n_1} \otimes \mathbb{K}^{n_2} \otimes \mathbb{K}^{n_3}$



Decomposition of a trilinear tensor

$$T = \sum_{j=1}^{r} U_j \otimes V_j \otimes W_j \text{ with } U_j \in \mathbb{K}^{n_1}, V_j \in \mathbb{K}^{n_2}, W_j \in \mathbb{K}^{n_3}$$

with r minimal.

Coefficent-wise: $T_{i_1,i_2,i_3} = \sum_{j=1}^{r} U_{i_1,j} V_{i_2,j} W_{i_3,j}$

Decomposition of a multilinear tensor

$$T = \sum_{j=1}^{r} U_j^1 \otimes \cdots \otimes U_j^l \text{ with } U_j^i \in \mathbb{K}^{n_i},$$

with r minimal.

Decomposition as polynomial:

$$T(\mathrm{x}) = \sum_{j=1}^{r} (U_j^1 \cdot \mathrm{x}_1) \cdots (U_j^l \cdot \mathrm{x}_l)$$

with $\mathrm{x}_k = (x_{k,1}, \ldots, x_{k,n_k})$ and $(U_j^k \cdot \mathrm{x}_k) = U_{1,j}^k x_{k,1} + \cdots + U_{n_k,j}^k x_{k,n_k}$.

Coefficent-wise: $T_{i_1,\ldots,i_l} = \sum_{j=1}^{r} U_{i_1,j}^1 \cdots U_{i_l,j}^l$

Symmetric tensor decomposition and Waring problem (1770)

Symmetric tensor decomposition problem:

Given a homogeneous polynomial F of degree d in the variables $\overline{\boldsymbol{x}} = (x_0, x_1, \dots, x_n)$ with coefficients $\in \mathbb{K}$:

$$F(\overline{\boldsymbol{x}}) = \sum_{|\alpha|=d} t_\alpha \, \overline{\boldsymbol{x}}^\alpha,$$

find a minimal decomposition of F of the form

$$F(\overline{\boldsymbol{x}}) = \sum_{i=1}^{r} \omega_i (\xi_{i,0} x_0 + \xi_{i,1} x_1 + \cdots + \xi_{i,n} x_n)^d$$

with $\xi_i = (\xi_{i,0}, \xi_{i,1}, \dots, \xi_{i,n}) \in \overline{\mathbb{K}}^{n+1}$ spanning disctint lines, $\omega_i \in \overline{\mathbb{K}}$.

The minimal r in such a decomposition is called the **rank** of T.

Moment sequence decomposition

Moment sequence decomposition

Given $\sigma = (\sigma_\alpha)_{|\alpha|\le d} \in \mathbb{N}^{s_{n,d}}$, find r minimal, $\omega_i \in \mathbb{K}$, $\xi_i \in \mathbb{K}^n$ for $i = 1, \ldots, r$ s.t.

$$\sigma(\boldsymbol{z}) := \sum_{|\alpha|\le d} \sigma_\alpha \frac{\boldsymbol{z}^\alpha}{\alpha!} = \Big(\sum_{i=1}^{r} \omega_i \, \mathrm{e}_{\xi_i}(\boldsymbol{z})\Big)^{[d]}$$

where $\mathrm{e}_\xi(\boldsymbol{z}) = \sum_{\alpha\in\mathbb{N}^n} \xi^\alpha \frac{\boldsymbol{z}^\alpha}{\alpha!} = \exp(\boldsymbol{z}, \xi)$.

☞ r is called the **rank** of σ.

Polynomial-exponential sequence decomposition

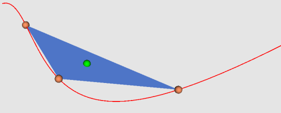
Given $\sigma = (\sigma_\alpha)_{|\alpha|\le d} \in \mathbb{N}^{s_{n,d}}$, find $r' \in \mathbb{N}, \omega_i(\boldsymbol{z}) \in \mathbb{K}[\boldsymbol{z}]$, $\xi_i \in \mathbb{K}^n$ s.t.

$$\sigma(\boldsymbol{z}) = \Big(\sum_{i=1}^{r'} \omega_i(\boldsymbol{z}) \, \mathrm{e}_{\xi_i}(\boldsymbol{z})\Big)^{[d]}$$

and $r = \sum_{i=1}^{r'} \dim\langle\langle \omega_i \rangle\rangle$ is minimal.

☞ r is called the **polynomial-exponential rank** of σ.

Geometric point of view



- $\mathcal{V}_{n+1,d} = \{\omega\,(\xi, \boldsymbol{x})^d, \omega \in \mathbb{K}, \xi \in \mathbb{K}^{n+1}\}$ **Veronese** variety (smooth except at 0)

$$F = \sum_{i=1}^{r} \omega_i (\xi_i, \boldsymbol{x})^d \text{ iff } F \in \overbrace{\mathcal{V}_{n+1,d} + \cdots + \mathcal{V}_{n+1,d}}^{r}$$

- $\mathrm{Sec}^r_{n+1,d} = \overline{\sum_{i=1}^r \mathcal{V}_{n+1,d}}$ r^{th}**-secant** variety of $\mathcal{V}_{n+1,d}$.
- For $S \subset \mathbb{R}^n$ (compact),

$$\begin{aligned} \mathcal{M}_+(S) &= \text{ **positive measures** supported on } S \\ &= \mathrm{Conv}(\delta_\xi \mid \xi \in S) \text{ where } \delta_\xi = \text{ **Dirac** at } \xi = \mathrm{e}_\xi \\ \mathcal{M}_+(S)^{[d]} &= \text{ **truncated moment sequences** of } \mu \in \mathcal{M}_+(S) \\ &= Conv((1 + (\xi, \boldsymbol{x}))^d \mid \xi \in S) \end{aligned}$$

- ☞ How to **decompose** a tensor or a moment sequence?
- ☞ Are there matrices **reaviling the rank**?
- ☞ How can we **deduce** the decomposition?

Flattening, matricisation

Multilinear tensors

For a multilinear tensor $T = [t_{i_1,\dots,i_l}] \in \mathbb{K}^{n_1\times\cdots\times n_l}$, **flattening** or **matricisation** in mode $(n_1 \times \cdots \times n_k, n_{k+1} \times \cdots \times n_l)$:

$$[t_{I,J}]_{I\in[n_1]\times\cdots\times[n_k],J\in[n_{k+1}]\times\cdots\times[n_l]}$$

$\Rightarrow$ matrix of size $M \times N$ with $M = n_1 \times \cdots \times n_k, N = n_{k+1} \times \cdots \times n_l$.

$$\begin{aligned} & \mathbb{K}^{n_1} \otimes \cdots \otimes \mathbb{K}^{n_k} \otimes \mathbb{K}^{n_{k+1}} \otimes \cdots \otimes \mathbb{K}^{n_l} \\ \sim \quad & \overbrace{(\mathbb{K}^{n_1} \otimes \cdots \otimes \mathbb{K}^{n_k})}^{E} \otimes \overbrace{(\mathbb{K}^{n_{k+1}} \otimes \cdots \otimes \mathbb{K}^{n_l})}^{F} \\ \sim \quad & E \otimes F \end{aligned}$$

Symmetric tensors

For $F = \sum_{|\gamma|=d} F_\gamma \, \mathrm{x}^\gamma \in S^d(\mathbb{K}^n)$, **matricisation** in degree $(k, d-k)$:

- $H_F^{k,d-k} = [\langle F, \mathrm{x}^{\alpha+\beta}\rangle_d]_{|\alpha|=k,|\beta|=d-k} = [\binom{d}{\alpha+\beta}^{-1} F_{\alpha+\beta}]_{|\alpha|=k,|\beta|=d-k}$

  also known as **flattening** or **Catalecticant** or **Hankel** matrix of F

  in degree $(k, d-k)$.

  $H_\mu^{k,d-k} = (\int \mathrm{x}^{\alpha+\beta} d\mu)_{|\alpha|=k,|\beta|=d-k}$ is a.k.a the **moment** matrix of $\mu \equiv \check{F}$.

  For $A \subset \mathcal{S}_{n,k}, B \subset \mathcal{S}_{n,d-k}$,

$$H_F^{A,B} = [\langle F, a\, b\rangle_d]_{a\in A, b\in B}$$

- **Catalecticant, Hankel** operator:

$$\begin{array}{rcl} H_F^{k,d-k} : \mathcal{S}_{n,d-k} & \rightarrow & S_{n,k}^* \\ b & \mapsto & b \star F^* \end{array}$$

Hankel matrix factorisation

Definition

For $\Xi = \{\xi_1, \ldots, \xi_r\}$ and $A = \{a_1, \ldots, a_s\} \subset S^k$

$$V_{A,\Xi} = \begin{bmatrix} a_1(\xi_1) & \cdots & a_1(\xi_r) \\ \vdots & & \vdots \\ a_s(\xi_1) & \cdots & a_s(\xi_r) \end{bmatrix}$$

is the **Vandermonde** matrix of A, Ξ.

Vandermonde factorization

If $F = \sum_{i=1}^{r} \omega_i (\xi \cdot \boldsymbol{x})^d$ and $\boldsymbol{m}_k = \{\boldsymbol{x}^\alpha\}_{|\alpha|=k}$, then

$$H_F^{k,d-k} = V_{\boldsymbol{m}_k,\Xi} \operatorname{diag}(\omega_1, \ldots, \omega_r)\, V^t_{\boldsymbol{m}_{d-k},\Xi}$$

Example with Fibonacci sequence $\sigma = (0, 1, 1, 2, 3, 5, 8, 13, \ldots)$, $d = 4$

- $F = \sum_{i=0}^{4} \sigma_i \binom{d}{i} x_0^{d-i} x_1^i = 4x_0^3 x_1 + 6x_0^2 x_1^2 + 8x_0 x_1^3 + 3x_1^4$
- $k = 2, d - k = 2$

$$H_F^{2,2} = (\langle F, x_0^{d-i-j} x_1^{i+j} \rangle_d)_{0 \leq i,j \leq 3} = (\sigma_{i+j})_{0 \leq i,j \leq 3} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 2 \\ 1 & 2 & 3 \end{pmatrix}$$

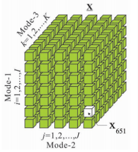
- rank $H_F^{2,2} = 2$
- $H_T^{2,2} = \begin{bmatrix} 1 & 1 \\ \xi_1 & \xi_2 \\ \xi_1^2 & \xi_2^2 \end{bmatrix} \mathrm{diag}(\omega_1, \omega_2) \begin{bmatrix} 1 & \xi_1 & \xi_1^2 \\ 1 & \xi_2 & \xi_2^2 \end{bmatrix}$

  with ξ_i roots of $X^2 - X - 1 = 0$ for $X = \frac{x_1}{x_0}$.

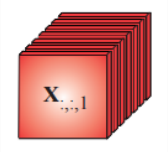
Decomposition via linear algebra

Multilinear tensors

$T \in \mathbb{K}^{n_1} \otimes \mathbb{K}^{n_2} \otimes \mathbb{K}^{n_3} \quad \equiv \quad [T_{[i]}]_{i=1}^{n_3}$ pencil of n_3 matrices of size $n_1 \times n_2$.



$\equiv$



For $T \in \mathbb{K}^{r} \otimes \mathbb{K}^{r} \otimes \mathbb{K}^{n_3}$,

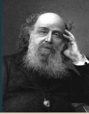
$$T = \sum_{j=1}^{r} U_j \otimes V_j \otimes W_j \text{ with } U, V \in \mathbb{K}^{r\times r}, W \in \mathbb{K}^{n_3 \times r}$$

$$\text{iff} \quad \mathrm{T}_{[i]} = U \operatorname{diag}(W_{i,1}, \ldots, W_{i,r}) V^t \quad i \in 1{:}n_3$$

If $T_{[1]}$ inv., $\quad U =$ matrix of **common eigenvectors** of $M_i = T_{[i]}\, T_{[1]}^{-1}$

$V^{-t} =$ matrix of **common eigenvectors** of $M_i' = T_{[1]}^{-1} T_{[i]}$.

Sylvester approach (1851)



Theorem:

The binary form $T(x_0,x_1)=\sum_{i=0}^{d} t_i\binom{d}{i}x_0^{d-i}x_1^i$ can be decomposed as a sum of r distinct powers of linear forms

$$T=\sum_{k=1}^{r}\omega_k(\alpha_k x_0+\beta_k x_1)^d$$

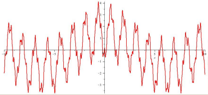
iff there exists a polynomial $p(x_0,x_1):=p_0x_0^r+p_1x_0^{r-1}x_1+\cdots+p_rx_1^r$ s.t.

$$\begin{bmatrix} t_0 & t_1 & \ldots & t_r \\ t_1 & & & t_{r+1} \\ \vdots & & & \vdots \\ t_{d-r} & \ldots & t_{d-1} & t_d \end{bmatrix}\begin{bmatrix} p_0 \\ p_1 \\ \vdots \\ p_r \end{bmatrix}=0$$

and of the form $p=c\prod_{k=1}^{r}(\beta_k x_0-\alpha_k x_1)$ with $(\alpha_k:\beta_k)$ distinct.

If $\alpha_k\neq 0$, $\xi_k=\frac{\beta_k}{\alpha_k}$ root of $p(x)=\sum_{i=0}^{r}p_ix^i$ (or generalized eigenvalues of (H_0,H_1)).

Prony's method (1795)



For the signal $f(t) = \sum_{i=1}^{\mathrm{r}} \omega_{\mathrm{i}} e^{\zeta_{\mathrm{i}} t}$, $(\omega_i, \zeta_i \in \mathbb{C})$,

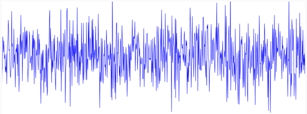
- Evaluate f at $2\,r$ regularly spaced points: $\sigma_0 := f(0), \sigma_1 := f(1), \ldots$
- Find $\mathrm{p} = [\mathrm{p}_0, \ldots, \mathrm{p}_\mathrm{r}]$ with $p_r = 1$ s.t. $\sigma_{k+r} + \sigma_{k+r_1} p_{r-1} + \cdots + \sigma_k p_0 = 0$ by solving:

$$\begin{bmatrix} \sigma_0 & \sigma_1 & \ldots & \sigma_r \\ \sigma_1 & & & \sigma_{r+1} \\ \vdots & & & \vdots \\ \sigma_{r-1} & \ldots & \sigma_{2r-1} & \sigma_{2r-1} \end{bmatrix} \begin{bmatrix} \mathrm{p}_0 \\ \mathrm{p}_1 \\ \vdots \\ \mathrm{p}_\mathrm{r} \end{bmatrix} = 0$$

- Compute the roots $\xi_1 = e^{\zeta_1}, \ldots, \xi_\mathrm{r} = e^{\zeta_\mathrm{r}}$ of $p(x) := \sum_{i=0}^{r} p_i x^i$ (or generalized eigenvalues of (H_0, H_1))
- Solve the system

$$\begin{bmatrix} 1 & \ldots & \ldots & 1 \\ \xi_1 & & & \xi_r \\ \vdots & & & \vdots \\ \xi_1^{r-1} & \ldots & \ldots & \xi_r^{r-1} \end{bmatrix} \begin{bmatrix} \omega_1 \\ \omega_2 \\ \vdots \\ \omega_\mathrm{r} \end{bmatrix} = \begin{bmatrix} \sigma_0 \\ \sigma_1 \\ \vdots \\ \sigma_{r-1} \end{bmatrix}$$

Decoding

$- - - - - -$  $- - - - - - - \rightarrow$

An algebraic code:

$$V = \{v(f) = [f(\xi_1), \ldots, f(\xi_m)] \mid f \in \mathbb{K}[x]; \deg(f) \leq d\}.$$

Encoding messages using the dual code:

$$C = V^{\perp} = \{\mathrm{c} \mid \mathrm{c} \cdot [f(\xi_1), \ldots, f(\xi_m)] = 0 \;\text{ for } f = x^k, 0 \leq k \leq d\}$$

Message received: $\mathrm{r} = \mathrm{m} + \mathrm{e}$ for $m \in C$ where $\mathrm{e} = [\omega_1, \ldots, \omega_m]$ is an error with $\omega_j \neq 0$ for $j = i_1, \ldots, i_r$ and $\omega_j = 0$ otherwise.

☞ **Find the error** e.

Berlekamp-Massey method (1969)

- Compute the syndrome $\sigma_k = v(x^k) \cdot \mathrm{r} = c(x^k) \cdot \mathrm{e} = \sum_{j=1}^{r} \omega_{i_j} \xi_{i_j}^k$.
- Compute the matrix

$$\begin{bmatrix} \sigma_0 & \sigma_1 & \dots & \sigma_r \\ \sigma_1 & & & \sigma_{r+1} \\ \vdots & & & \vdots \\ \sigma_{r-1} & \dots & \sigma_{2r-1} & \sigma_{2r-1} \end{bmatrix} \begin{bmatrix} p_0 \\ p_1 \\ \vdots \\ p_r \end{bmatrix} = 0$$

  and its kernel $p = [p_0, \dots, p_r]$.
- Compute the roots of the **error locator polynomial** $p(x) = \sum_{i=0}^{r} p_i \, x^i = p_r \prod_{j=1}^{r} (x - \xi_{i_j})$.
- Deduce the errors ω_{i_j} by solving a Vandermonde linear system.

Solving polynomial equations

How to solve $f = 0$ for $f = \sum_{i=0}^{d} f_i\, x^i \in \mathbb{K}[x]$?

The matrix of the **multiplication** by x modulo f, assuming $f_d = 1$:

$$\begin{aligned} \mathcal{M}_x : \mathbb{K}[x]/(f) &\to \mathbb{K}[x]/(f) \\ \overline{p} &\mapsto \overline{x\,p} \end{aligned}$$

In the mon. basis $1, x, \ldots, x^{d-1}$:

$$\begin{bmatrix} 0 & \ldots & 0 & -f_0 \\ 1 & \ddots & \vdots & \vdots \\ & \ddots & 0 & \vdots \\ 0 & & 1 & -f_{d-1} \end{bmatrix}$$

$\mathcal{M}_x(x^i) = x^{i+1}$ for $i \in 0{:}d\text{-}2$

$\mathcal{M}_x(x^{d-1}) = -\sum_{i=0}^{d-1} f_i x^i$

In the Lagrange basis (for simple roots):

$$\begin{bmatrix} \xi_1 & 0 & \ldots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & \xi_r \end{bmatrix}$$

$\mathcal{M}_x(\sum_{i=1}^{d} \lambda_i \boldsymbol{u}_i) = \sum_{i=1}^{d} \xi_i \lambda_i \boldsymbol{u}_i$

Theorem:

- The **eigenvalues** of $\mathcal{M}_x$ are the roots ξ_i of $f(x) = 0$.
- The **eigenvectors** of $\mathcal{M}_x$ are the Lagrange interpolation polynomials $\boldsymbol{u}_i$ (when the roots are simple).

- The dual space $\mathcal{A}^*$ of $\mathcal{A}$ is the set $\mathrm{Hom}_{\mathbb{K}}(\mathcal{A}, \mathbb{K})$ of linear forms $\Lambda : \mathcal{A} \to \mathbb{K}$

$$\mathcal{A}^* = (f)^\perp = \{\Lambda \in \mathbb{K}[x]^* \mid \forall p \in (f), \langle \Lambda | p \rangle = 0\}.$$

- The multiplication by x in $\mathcal{A}^*$ is the transposed of $\mathcal{M}_x$:

$$\begin{array}{rcl} \mathcal{M}_x^t : \mathcal{A}^* & \to & \mathcal{A}^* \\ \Lambda & \mapsto & x \star \Lambda \end{array}$$

where $x \star \Lambda : f \mapsto \langle \Lambda | x\, f \rangle = \langle \Lambda | \mathcal{M}_x(f) \rangle$.

- If the roots are simple, a basis of $\mathcal{A}^*$ is the set of **evaluations** $\mathrm{e}_{\xi_i} : f \mapsto f(\xi_i)$.

It is dual to the basis of **interpolation polynomials** $(\boldsymbol{u}_i)_{i=1:d}$.

Proposition

The evaluations e_{ξ_i} for $i \in 1\!:\!d$ are **eigenvectors** *of $\mathcal{M}_x^t$.*

Proof: $\mathcal{M}_x(\mathrm{e}_{\xi_i}) = x \star \mathrm{e}_{\xi_i} = (p \mapsto \xi_i p(\xi_i)) = \xi_i \mathrm{e}_{\xi_i}$.

The matrix of $(\mathrm{e}_{\xi_1}, \ldots, \mathrm{e}_{\xi_d})$ in the dual basis of $(1, x, \ldots, x^{d-1})$ is the **Vandermonde** matrix of $\xi_1, \ldots, \xi_d$:

$$V = \begin{pmatrix} 1 & 1 & \ldots & 1 \\ \xi_1 & \xi_2 & \ldots & \xi_d \\ \vdots & \vdots & & \vdots \\ \xi_1^{d-1} & \xi_2^{d-1} & \ldots & \xi_d^{d-1} \end{pmatrix}$$

If U is the coefficient matrix of Lagrange basis $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_d$ in the basis $(1, x, \ldots, x^{d-1})$ of $\mathcal{A} = \mathbb{K}[x]/(f)$, we have

$$V^t\, U = \mathrm{Id}$$

since $(1, \xi_i, \ldots, \xi_i^{d-1})U_j = \boldsymbol{u}_j(\xi_i) = \delta_{i,j}$

The roots by eigencomputation

Hypothesis: $\mathcal{V}_{\overline{\mathbb{K}}}(I) = \{\xi_1, \ldots, \xi_r\} \Leftrightarrow \mathcal{A} = \mathcal{R}/I$ Artinian (i.e. $\dim_{\mathbb{K}} \mathcal{A} = r < \infty$).

$$\begin{array}{rcl} \mathcal{M}_g : \mathcal{A} & \to & \mathcal{A} \\ a & \mapsto & g\, a \end{array} \qquad \begin{array}{rcl} \mathcal{M}_g^{\mathsf{t}} : \mathcal{A}^* & \to & \mathcal{A}^* \\ \Lambda & \mapsto & g \star \Lambda = \Lambda \circ \mathcal{M}_g \end{array}$$

Theorem:

- The eigenvalues of $\mathcal{M}_g$ are $\{g(\xi_1), \ldots, g(\xi_r)\}$.
- The common eigenvectors of all $(\mathcal{M}_g^{\mathsf{t}})_{g \in \mathcal{A}}$ are $\mathrm{e}_{\xi_i} : p \mapsto p(\xi_i)$ (up to a sc.).

Proposition

If the roots are **simple**,

- *the operators $\mathcal{M}_g$ are diagonalizable,*
- *their common eigenvectors are the* **interpolation polynomials** $\boldsymbol{u}_i$ *at the roots (up to a sc.), which form a basis of idempotents of $\mathcal{A}$.*

Structure of $\mathcal{A}$

Theorem:

If $I = Q_1 \cap \cdots \cap Q_{r'}$ with Q_i $\boldsymbol{m}_{\xi_i}$-primary, then

- $\mathcal{V}_{\overline{\mathbb{K}}}(I) = \{\xi_1, \ldots, \xi_{r'}\}$
- $\mathcal{A} = \mathcal{A}_1 \oplus \cdots \oplus \mathcal{A}_{r'}$ with $\mathcal{A}_i = \mathcal{R}/Q_i$
- $1 = \boldsymbol{u}_1 \oplus \cdots \oplus \boldsymbol{u}_{r'}$ with $\mathcal{A}_i = \boldsymbol{u}_i\,\mathcal{A}$, $\boldsymbol{u}_i^2 = \boldsymbol{u}_i$, $\boldsymbol{u}_i\boldsymbol{u}_j = 0$ if $i \neq j$.
  ($\boldsymbol{u}_i$ idempotents).

Theorem:

In a basis of $\mathcal{A}$, all the matrices M_g ($g \in \mathcal{A}$) are of the form

$$\mathtt{M}_g = \begin{bmatrix} \mathtt{M}_g^{\mathbf{1}} & & 0 \\ & \ddots & \\ 0 & & \mathtt{M}_g^{r'} \end{bmatrix} \text{ with } \mathtt{M}_g^{i} = \begin{bmatrix} g(\xi_i) & \star & \star \\ & \ddots & \star \\ 0 & & g(\xi_i) \end{bmatrix}$$

Corollary (Chow form)

$\Delta(\mathrm{u}) = \det(v_0 + v_1\,\mathtt{M}_{x_1} + \cdots + v_n\,\mathtt{M}_{x_n}) = \prod_{i=1}^{r}(v_0 + v_1\xi_{i,1} + \cdots + v_n\xi_{i,n})^{\mu_{\xi_i}}$ *where* $\mu_{\xi_i} = \dim \mathcal{A}_i$ *is the multiplicity of* ξ.

Structure of the dual $\mathcal{A}^*$

Definition (Polynomial-Exponential series)

$$\mathcal{P}ol\mathcal{E}xp = \left\{ \sigma(\boldsymbol{z}) = \sum_{i=1}^{r} \omega_i(\boldsymbol{z})\, \mathrm{e}_{\xi_i}(\boldsymbol{z}) \mid \omega_i(\boldsymbol{z}) \in \mathbb{K}[\boldsymbol{z}], \xi_i \in \mathbb{K}^n \right\}$$

where $\mathrm{e}_{\xi_i}(\boldsymbol{z}) = e^{z_1\xi_{i,1}+\cdots+z_n\xi_{i,n}} = \sum_\alpha \xi^\alpha \frac{\boldsymbol{z}^\alpha}{\alpha!}$ *is* the evaluation $\mathrm{e}_\xi : p \in \mathcal{R} \mapsto p(\xi)$.

Inverse system generated by $\omega_1(\boldsymbol{z}), \ldots, \omega_r(\boldsymbol{z}) \in \mathbb{K}[\boldsymbol{z}]$

$$\langle\langle \omega_1(\boldsymbol{z}), \ldots, \omega_r(\boldsymbol{z}) \rangle\rangle = \langle \partial_{\boldsymbol{z}}^\alpha(\omega_i), \alpha \in \mathbb{N}^n \rangle$$

Theorem:

For $\mathbb{K} = \overline{\mathbb{K}}$ algebraically closed and $\mathcal{A} = \mathcal{R}/I$ artinian with $I = Q_1 \cap \cdots \cap Q_{r'}$, Q_i $\boldsymbol{m}_{\xi_i}$-primary,

$$\mathcal{A}^* = I^\perp = \oplus_{i=1}^{r'} \mathcal{D}_i\, \mathrm{e}_{\xi_i}(\boldsymbol{z}) \quad \subset \quad \mathcal{P}ol\mathcal{E}xp$$

- $\mathcal{D}_i = Q_i^\perp = \langle\langle \omega_{i,1}(\boldsymbol{z}), \ldots, \omega_{i,l_i}(\boldsymbol{z}) \rangle\rangle$ with $\omega_{i,j}(\boldsymbol{z}) \in \mathbb{K}[\boldsymbol{z}]$.
- $\dim_{\mathbb{K}}(\mathcal{D}_i) = \mu_i$ multiplicity of ξ_i.

Example

$$\begin{cases} f_1 &= x_1^2 x_2 - x_1^2 \\ f_2 &= x_1 x_2 - x_2 \end{cases} \qquad I = (f_1, f_2) \subset \mathcal{R} = \mathbb{C}[x_1, x_2]$$

$$\mathcal{A} = \mathcal{R}/I \sim B = \langle 1, x_1, x_2 \rangle, \quad I = (x_1^2 - x_2, x_1 x_2 - x_2, x_2^2 - x_2)$$

$$M_1 = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix}, \quad M_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix} \qquad \begin{array}{c} \text{common} \\ \text{eig. vect.} \\ \text{of } M_1^t, M_2^t \end{array} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

$$I = Q_1 \cap Q_2 \quad \text{with primary comp.} \quad Q_1 = (x_1^2, x_2), \quad Q_2 = (x_1 - 1, x_2 - 1)$$

$$\mathcal{A} = \mathcal{A}_1 \oplus \mathcal{A}_2 \quad \text{with} \quad \mathcal{A}_1 = \langle \boldsymbol{u}_1, x_1 \boldsymbol{u}_1 \rangle \equiv \mathcal{R}/Q_1 \quad \mathcal{A}_2 = \langle \boldsymbol{u}_2 \rangle \equiv \mathcal{R}/Q_2$$

$$\mathcal{A}^* = I^\perp = Q_1^\perp \oplus Q_2^\perp = \quad \langle \mathrm{e}_{(0,0)}, z_1 \mathrm{e}_{(0,0)} \rangle \quad \oplus \quad \langle \mathrm{e}_{(1,1)} \rangle$$

where $\mathrm{e}_{(a,b)} : p \mapsto p(a,b)$, $z_1 \mathrm{e}_{(0,0)} : p \mapsto \partial_{x_1}(p)(0,0)$.

The algebraic structure from Hankel operators

Kronecker theorems

Univariate series:

Kronecker (1881)

The Hankel operator

$$\begin{array}{rcl} \mathcal{H}_\sigma : \mathbb{C}^{\mathbb{N},finite} & \to & \mathbb{C}^{\mathbb{N}} \\ (p_m) & \mapsto & (\sum_m \sigma_{m+n} p_m)_{n\in\mathbb{N}} \end{array}$$

is of **finite rank** r iff $\exists \omega_1, \dots, \omega_{r'} \in \mathbb{C}[z]$ and $\xi_1, \dots, \xi_{r'} \in \mathbb{C}$ distincts s.t.

$$\sigma(z) = \sum_{n\in\mathbb{N}} \sigma_n \frac{z^n}{n!} = \sum_{i=1}^{r'} \omega_i(z) \mathrm{e}_{\xi_i}(z)$$

with $\sum_{i=1}^{r'} (\deg(\omega_i) + 1) = r$.

Multivariate series:

Theorem: Generalized Kronecker Theorem [M'2018]

For $\sigma \in \mathcal{R}^*$, the Hankel operator

$$\begin{array}{rcl} \mathcal{H}_\sigma : \mathcal{R} & \to & \mathcal{R}^* \\ p & \mapsto & p \star \sigma \end{array}$$

is of rank r iff

$$\sigma(\boldsymbol{z}) = \sum_{i=1}^{r'} \omega_i(\boldsymbol{z})\, \mathrm{e}_{\xi_i}(\boldsymbol{z}) \quad \text{with } \omega_i(\boldsymbol{z}) \in \mathbb{K}[\boldsymbol{z}],$$

with $r = \sum_{i=1}^{r'} \dim\langle\langle \omega_i(\boldsymbol{z})\rangle\rangle = \sum_{i=1}^{r'} \dim\langle \partial_{\boldsymbol{z}}^{\gamma} \omega_i(\boldsymbol{z})\rangle$. In this case, we have

- $I_\sigma = \ker H_\sigma$ with $\mathcal{V}_{\mathbb{C}}(I_\sigma) = \{\xi_1, \ldots, \xi_{r'}\}$.
- $I_\sigma = Q_1 \cap \cdots \cap Q_{r'}$ with $Q_i^\perp = \langle\langle \omega_i \rangle\rangle\, \mathrm{e}_{\xi_i}(\boldsymbol{z})$.

☞ $\mathcal{A}_\sigma$ is **Gorenstein**[a]; $(a, b) \mapsto \langle \sigma | ab \rangle$ is non-degenerate in $\mathcal{A}_\sigma$.

☞ Can be generalized to $\sigma = (\sigma_1, \ldots, \sigma_m) \in (\mathcal{R}^*)^m$.

[a] $\mathcal{A}_\sigma^* = \mathcal{A}_\sigma \star \sigma$ is a free $\mathcal{A}_\sigma$-module of rank 1

For $F = \sum_{i=1}^{r} \omega_i (\xi_i, \boldsymbol{x})^d$ or $F^* = (\sum_{i=1}^{r} \omega_i \mathrm{e}_{\xi_i})^{[d]}$ with $\xi_{i,0} = 1$, let

- $I_\Xi := \{p \in \mathcal{R} \text{ s.t. } p(\xi_i) = 0\}$ be the defining ideal of $\Xi = \{\xi_1, \ldots, \xi_r\}$
- $\mathcal{A}_\Xi := \mathcal{R}/I_\Xi$ the quotient algebra by I_Ξ, of dimension r.

Theorem:

Let $A = \{a_1, \ldots, a_s\}, A' = \{a'_1, \ldots, a'_t\}$, and $H = H_F^{A,A'}$ be a **flattening** of F. Then

$$H_F^{A,A'} = V_{A,\Xi} \Delta_\omega V_{A',\Xi}^t$$

where $V_{A,\Xi} = \begin{bmatrix} a_1(\xi_1) & \cdots & a_1(\xi_r) \\ \vdots & & \vdots \\ a_s(\xi_1) & \cdots & a_s(\xi_r) \end{bmatrix}$ is the **Vandermonde** matrix A, Ξ.

If A and A' contain a basis of $\mathcal{A}_\Xi$, then

- $\ker H = I_\Xi \cap \langle A' \rangle$
- $\operatorname{im} H = (I_\Xi^\perp)_{|\langle A \rangle} = \operatorname{im} V_{A,\Xi}$ where

$$I_\Xi^\perp = \{\Lambda \in \mathbb{K}[\boldsymbol{x}]^* \mid \forall p \in I_\Xi, \langle \Lambda, p \rangle = 0\} = \mathcal{A}_\Xi^* = \langle \mathrm{e}_{\xi_1}, \ldots, \mathrm{e}_{\xi_r} \rangle.$$

Proof: A containts a basis of $\mathcal{A}_\Xi \Rightarrow V_{A,\Xi}$ of rank r.

Proposition

For $g \in \mathcal{R}$, $\mathcal{M}_g : a \in \mathcal{A}_\Xi \mapsto ga \in \mathcal{A}_\Xi$, $\Lambda \in \mathcal{R}^*$,
$\mathcal{H}_{g\star\Lambda} = \mathcal{H}_\Lambda \circ \mathcal{M}_g = \mathcal{M}_g^t \circ \mathcal{H}_\Lambda$

Proposition

The sets $B \subset A$, $B' \subset A'$ *of size* $r := \dim(\mathcal{A})$ *are bases of* $\mathcal{A}$ *iff* $H_0 = H_F^{B',B}$ *is invertible.*

Proposition

Assume that $x_i \cdot B \subset A$, *let* $H_i = H_F^{B',x_i B}$. *Then* $M_i = H_0^{-1} H_i$ *is the multiplication by* x_i *in* B *modulo* I_Ξ

Proposition

$\exists\, E, F$ *invertible such that*

$$\mathsf{H_i} = \mathsf{E}\,\mathrm{diag}(\xi_{1,i}, \ldots, \xi_{r,i})\,\mathsf{F} \Rightarrow \textbf{joint diagonalisation} \text{ of } H_0^{-1} H_i.$$

Example with Fibonacci sequence

$\sigma = (0,1,1,2,3,5,8,13,\ldots)$, $d=6$, $\psi = \sum_{i=0}^{d} \sigma_i \binom{d}{i} x_0^{d-i} x_1^i$.

- $H_\psi = \begin{pmatrix} 0 & 1 & 1 & 2 \\ 1 & 1 & 2 & 3 \\ 1 & 2 & 3 & 5 \\ 2 & 3 & 5 & 8 \end{pmatrix} H_0 = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}, H_1 = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$
- $M_x = H_0^{-1} H_1 = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$.
- Eigenvalues: $\xi_i = \frac{1 \pm \sqrt{5}}{2}$; Eigenvectors: $\mathbf{u}_i = \pm \frac{1}{\sqrt{5}}(x - \xi_i)$, $i=1,2$.
- Weights: $U^t H_0 U = \begin{pmatrix} \frac{1}{\sqrt{5}} & 0 \\ 0 & -\frac{1}{\sqrt{5}} \end{pmatrix}$.

☞ **Decomposition:**

$$\psi = \frac{1}{\sqrt{5}}(x_0 + \frac{1+\sqrt{5}}{2} x_1)^d - \frac{1}{\sqrt{5}}(x_0 + \frac{1-\sqrt{5}}{2} x_1)^d$$

Symmetric tensor decomposition

$$\begin{aligned}
\psi &= (x_0+3\,x_1-x_2)^4+(x_0+x_1+x_2)^4-3\,(x_0+2\,x_1+2\,x_2)^4\\
&= -x_0{}^4-24\,x_0{}^3x_2-8\,x_0{}^3x_1-60\,x_0{}^2x_2{}^2-168\,x_0{}^2x_1x_2-12\,x_0{}^2x_1{}^2\\
&\quad -96\,x_0x_2{}^3-240\,x_0x_1x_2^2-384\,x_0x_1^2x_2+16\,x_0x_1{}^3-46\,x_2{}^4-200\,x_1x_2^3\\
&\quad -228\,x_1{}^2x_2{}^2-296\,x_1^3x_2+34\,x_1{}^4\\
\langle\psi,p\rangle_4 &= \langle\psi^*|p\rangle \text{ where } \psi^*=e_{(3,-1)}+e_{(1,1)}-3\,e_{(2,2)} \quad (\textit{by apolarity})
\end{aligned}$$

$$H^{2,2}_{\psi^*} := \left[\begin{array}{cccccc}
-1 & -2 & -6 & -2 & -14 & -10\\
-2 & -2 & -14 & 4 & -32 & -20\\
-6 & -14 & -10 & -32 & -20 & -24\\
\hline
-2 & 4 & -32 & 34 & -74 & -38\\
-14 & -32 & -20 & -74 & -38 & -50\\
-10 & -20 & -24 & -38 & -50 & -46
\end{array}\right]$$

For $B'=\{x_{\mathbf{0}},x_{\mathbf{1}},x_{\mathbf{2}}\}, B=x_{\mathbf{0}}B'$

$$H_\psi^{B,x_{\mathbf{0}}B'} = \begin{bmatrix} -1 & -2 & -6\\ -2 & -2 & -14\\ -6 & -14 & -10\end{bmatrix}$$

$$H_\psi^{B,x_{\mathbf{1}}B'} = \begin{bmatrix} -2 & -2 & -14\\ -2 & 4 & -32\\ -14 & -32 & -20\end{bmatrix}$$

$$H_\psi^{B,x_{\mathbf{2}}B'} = \begin{bmatrix} -6 & -14 & -10\\ -14 & -32 & -20\\ -10 & -20 & -24\end{bmatrix}$$

- The matrix of multiplication by x_2 in $B' = \{x_0, x_1, x_2\}$ is

$$M_{\mathbf{2}} = (H_\psi^{B, x_{\mathbf{0}} B'})^{-\mathbf{1}} H_\psi^{B, x_{\mathbf{2}} B'} = \begin{bmatrix} 0 & -2 & -2 \\ 0 & \frac{1}{2} & \frac{3}{2} \\ 1 & \frac{5}{2} & \frac{3}{2} \end{bmatrix}.$$

- Its eigenvalues are $[-1, 1, 2]$ and the eigenvectors:

$$U := \begin{bmatrix} 0 & -2 & -1 \\ \frac{1}{2} & \frac{3}{4} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{4} & \frac{1}{2} \end{bmatrix}.$$

that is the polynomials

$$U(x) = \begin{bmatrix} \frac{1}{2}x_1 - \frac{1}{2}x_2 & -2 + \frac{3}{4}x_1 + \frac{1}{4}x_2 & -1 + \frac{1}{2}x_1 + \frac{1}{2}x_2 \end{bmatrix}.$$

- We deduce the weights and the frequencies:

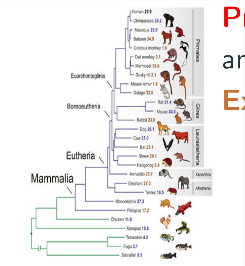
$$H_\psi^{[\mathbf{1}, x_{\mathbf{1}}, x_{\mathbf{2}}], U} = \begin{bmatrix} 1 & 1 & -3 \\ 1 \times 3 & 1 \times 1 & -3 \times 2 \\ 1 \times -1 & 1 \times 1 & -3 \times 2 \end{bmatrix}$$

Weights: $1, 1, -3$;
Frequencies: $(3, -1), (1, 1), (2, 2)$.

Decomposition:

$$\begin{aligned} \psi^* &= \mathrm{e}_{(3,-1)} + \mathrm{e}_{(1,1)} - 3\,\mathrm{e}_{(2,2)} \\ \psi(\mathrm{x}) &= (\mathrm{x}_0 + 3\,\mathrm{x}_1 - \mathrm{x}_2)^4 + (\mathrm{x}_0 + \mathrm{x}_1 + \mathrm{x}_2)^4 - 3\,(\mathrm{x}_0 + 2\,\mathrm{x}_1 + 2\,\mathrm{x}_2)^4 \end{aligned}$$

Phylogenetic tree multilinear tensors



Problem: study probability vectors for genes $[A, C, G, T]$ and the transitions described by Markov matrices M^i.

Example:

$$\begin{array}{rccc} \textit{Ancestor}: & & \mathcal{A} & \\ \textit{Transitions}: & M^1 & M^2 & M^3 \\ \textit{Species}: & \mathcal{S}_1 & \mathcal{S}_2 & \mathcal{S}_3 \end{array}$$

For $i_1, i_2, i_3 \in \{A, C, G, T\}$, the probability to observe i_1, i_2, i_3 is

$$p_{i_1,i_2,i_3} = \sum_{k=1}^{4} \omega_k \, M^1_{k,i_1} M^2_{k,i_2} M^3_{k,i_3} \Leftrightarrow \mathrm{p} = \sum_{k=1}^{4} \omega_k \, \mathrm{u_k} \otimes \mathrm{v_k} \otimes \mathrm{w_k}$$

where $\mathrm{u}_k = (M^1_{k,1}, \ldots, M^1_{k,4}), \mathrm{v}_k = (M^2_{k,1}, \ldots, M^2_{k,4}), \mathrm{w}_k = (M^3_{k,1}, \ldots, M^3_{k,4})$.

- ☞ p is a tensor $\in \mathbb{K}^4 \otimes \mathbb{K}^4 \otimes \mathbb{K}^4$ of rank ≤ 4.
- ☞ Its decomposition yields the M^i and the ancestor probability (ω_j).

1 Phylogenetic trees

We describe how the decomposition method works on a trilinear tensor of $\mathbb{R}^4 \otimes \mathbb{R}^4 \otimes \mathbb{R}^4$ of rank 4.

See also https://github.com/tenors-network/TENORS-L1-Inria/blob/main/courses/Algebra-for-Analysis-of-Tensors/Phylogenetic.ipynb

1.1 Problem

We are given a tensor T of the form

$$T = \sum_{i=1}^{4} \omega_i A[:,i] \otimes B[:,i] \otimes C[:,i]$$

where $A, B, C \in \mathbb{R}^{4\times 4}$ are Markov matrices and $\omega = (\omega_1, \ldots, \omega_4) \in (\mathbb{R}_+)^4$ are positive weights.

How to recover the weights ω and the factors A, B, C from the coefficients of T?

[1]:
```
using TensorDec, DynamicPolynomials, LinearAlgebra

# scale the columns of A, B, C and the weights w so that the sum of the colmuns is 1
normalize_markov! = function(w,A,B,C)
    for i in 1:size(A,2)
        l = sum(A[j,i] for j in 1:size(A,1) )
        A[:,i] /= l
        w[i] *= l
    end
    for i in 1:size(B,2)
        l = sum(B[j,i] for j in 1:size(B,1) )
        B[:,i] /=l
```

```
        w[i] *= l
    end
    for i in 1:size(C,2)
        l = sum(C[j,i] for j in 1:size(B,1) )
        C[:,i] /=l
        w[i] *= l
    end
    w, A,B,C
end;
# scale the columns of A so that the first row is [1, ..., 1]
normalize_affine = function(A0::AbstractMatrix) A = Matrix(A0); for i in 1:size(A,2) ␣
  ↪A[:,i] /= A[1,i] end; A end;
```

[2]:
```
A = rand(4,4); B = rand(4,4); C = rand(4,4); w = rand(4);
normalize_markov!(w,A,B,C);
```

We verify that the columns sum to `1.`:

[3]:
```
fill(1.,4)'*A, fill(1.,4)'*B, fill(1.,4)'*C
```

[3]:
```
([1.0 0.9999999999999999 1.0 0.9999999999999999], [1.0 1.0 1.0
0.9999999999999999], [1.0 1.0 0.9999999999999998 1.0])
```

[4]:
```
#w = fill(1.,4); A=[1 1 1 1; 0 1. 0 0; 0 0 2. 0; 0 0 0 3.]; B= A; C=A
```

[5]:
```
T = tensor(w,A,B,C)
```

```
[5]: 4×4×4 Array{Float64, 3}:
[:, :, 1] =
 0.63856   0.498562  0.609002  0.492611
 0.659005  0.576221  0.588409  0.480191
 0.331819  0.218192  0.31365   0.291366
 0.217415  0.142356  0.211895  0.192002

[:, :, 2] =
 0.459046  0.304443  0.516376  0.307711
 0.31125   0.259229  0.270975  0.217506
 0.323017  0.194411  0.324703  0.241424
 0.20303   0.116249  0.226682  0.158003

[:, :, 3] =
 0.147284   0.102441   0.122958   0.0951209
 0.0896452  0.0735928  0.0659168  0.0550932
 0.129666   0.0858343  0.0949834  0.0865085
 0.0728417  0.04477    0.0620078  0.0559637

[:, :, 4] =
 0.320664  0.218559   0.39174   0.197574
 0.240617  0.205661   0.21549   0.166941
 0.191538  0.117397   0.215483  0.130113
 0.124318  0.0727524  0.155088  0.0859419
```

The corresponding polynomial is:

[6]:
```
X = @polyvar x0 x1 x2 x3; Y = @polyvar y0 y1 y2 y3; Z = @polyvar z0 z1 z2 z3;
F= sum(T[i,j,k]*X[i]*Y[j]*Z[k] for i in 1:4 for j=1:4 for k=1:4)
```

[6]:

$$0.08594189851588788x3y3z3+0.05596366474953316x3y3z2+0.1580025087257289x3y3z1+0.19200166860542914x3y3z0+$$

[7]:
```
H = [T[:,:,i] for i in 1:4];
```

1.2 Matrices from the tensor

The flattening or Hankel matrix of the tensor indexed by monomials $A = (x_i)_{0\le i\le 3}$, $A' = (y_j\, z_k)_{0\le j,k\le 3}$ is:

[8]:
```
hcat(H...)
```

[8]:
```
4×16 Matrix{Float64}:
 0.63856   0.498562  0.609002  0.492611  …  0.218559   0.39174   0.197574
 0.659005  0.576221  0.588409  0.480191     0.205661   0.21549   0.166941
 0.331819  0.218192  0.31365   0.291366     0.117397   0.215483  0.130113
 0.217415  0.142356  0.211895  0.192002     0.0727524  0.155088  0.0859419
```

By scaling the factors and the weights of the decomposition, we assume that the factors U, V, W have their first coordinates equal to 1. We set $x_0 = 1$, $y_0 = 1$, $z_0 = 1$ and work on an affine chart of $\mathbb{P}^3 \times \mathbb{P}^3 \times \mathbb{P}^3$.

A basis of $\mathcal{A}$ is $B = \{1, y_1, y_2, y_3\}$. The operators M_i of multiplication by z_i in the basis B of $\mathcal{A}$ are:

[9]:
```
M = [inv(H[1])*H[i] for i in 2:4];
```

1.3 Joint diagonalisation

We take a random combination M_{rnd} of M_i and compute its eigenvectors:

[10]:
```
Mrnd = sum(M[i]*rand() for i in 1:3);
```

[11]:
```
E = eigen(Mrnd).vectors
```

[11]:
```
4×4 Matrix{Float64}:
 -0.677549   0.671611  -0.399447   0.818269
  0.69611   -0.69861    0.510161  -0.373124
  0.161483  -0.178923  -0.515211  -0.256734
  0.174016   0.169907   0.561013  -0.353981
```

We verify that the operators of multiplication M_i are diagonal (up to numerical error) in the basis of eigenvectors of M_{rnd}:

[12]:
```
D = [inv(E)*M[i]*E  for i in 1:3];
```

[13]:
```
D[1]
```

[13]:
```
4×4 Matrix{Float64}:
  0.384224      1.22402e-14  -4.45199e-14  -1.49325e-14
 -7.77156e-16   1.00375      -3.35287e-14  -1.22125e-14
 -6.73073e-15   3.69843e-15   3.18572       7.68829e-15
  9.4022e-15   -8.44463e-15   1.27676e-15   3.56338
```

[14]:
```
D[2]
```

[14]: 4×4 Matrix{Float64}:

```
  0.0804996    7.34135e-15  -3.58047e-15  9.50628e-15
 -6.43929e-15  0.389043      2.88658e-15  5.21805e-15
  3.08781e-16  9.09862e-16   0.434225     2.05391e-15
 -6.7793e-15   5.46785e-15  -1.44051e-14  2.95996
```

[15]:

```
D[3]
```

[15]: 4×4 Matrix{Float64}:

```
  0.330232      1.78468e-14  -1.29896e-14   6.10623e-15
 -8.60423e-15   0.472936     -1.9762e-14    4.21885e-15
  4.4964e-15   -6.50868e-15   2.82149      -6.18949e-15
  6.83481e-15  -3.38271e-15   3.11001e-14   1.79256
```

1.4 The factors from the eigenvectors

The corresponding terms on the diagonal give the factor W (with the first coordinate equal to 1):

[16]:

```
W = fill(1.,4,4); for i in 1:3 for j in 1:4 W[i+1,j] = D[i][j,j] end end; W
```

[16]: 4×4 Matrix{Float64}:

```
 1.0        1.0       1.0       1.0
 0.384224   1.00375   3.18572   3.56338
 0.0804996  0.389043  0.434225  2.95996
 0.330232   0.472936  2.82149   1.79256
```

The eigenvectors `E` are (up to a scalar) the interpolation polynomials at the points in the basis B. Therefore, `inv(E)'` is (up to a scaling of the columns) the **Vandermonde** matrix of the points in $B = \{1, y_1, y_2, y_3\}$:

$$V = \begin{pmatrix} 1 & 1 & 1 & 1 \\ v_{1,1} & v_{2,1} & v_{3,1} & v_{4,1} \\ v_{1,2} & v_{2,2} & v_{3,2} & v_{4,2} \\ v_{1,3} & v_{2,3} & v_{3,3} & v_{4,3} \end{pmatrix}$$

This gives the factor `V`

[17]:
```
V = inv(E)'
```

[17]:
```
4×4 adjoint(::Matrix{Float64}) with eltype Float64:
 4.54037  2.71845  -0.717092  2.40036
 4.10939  1.28546  -0.264018  2.21873
 4.05693  2.53245  -1.75912   0.421948
 3.2216   3.09232  -0.103501  0.0789611
```

We remove the scaling factors, by normalizing the columns so that the first coordinate of the points is 1:

[18]:
```
V = normalize_affine(V)
```

[18]:
```
4×4 Matrix{Float64}:
 1.0       1.0       1.0       1.0
 0.905077  0.472864  0.368179  0.924333
 0.893524  0.93158   2.45313   0.175785
 0.709546  1.13753   0.144335  0.0328955
```

Since $T = \sum_i \omega_i U[:,i] \otimes V[:,i] \otimes W[:,i]$ and $W[1,i] = 1$, we have $H_1 = T[:,:,1] = \sum_i \omega_i U[:,i] \otimes V[:,i]$ and the columns of U coincides up to a scaling with the columns $H_1(V')^{-1} = H_1 E$:

[19]:
```
U = normalize_affine(H[1]*E)
```

[19]:
```
4×4 Matrix{Float64}:
 1.0       1.0       1.0        1.0
 1.35261   0.29567   0.0795717  0.555197
 0.288573  1.15391   0.507942   1.11787
 0.197175  0.746456  0.411169   0.420357
```

1.5 The weights from the factors

We solve a Vandermonde system

$$\begin{pmatrix} \vdots & & \vdots \\ u_{1,i}v_{1,j}w_{1,k} & & u_{4,i}v_{4,j}w_{4,k} \\ \vdots & & \vdots \end{pmatrix} \begin{pmatrix} \omega_1 \\ \vdots \\ \omega_4 \end{pmatrix} = \begin{pmatrix} \vdots \\ T[i,j,k] \\ \vdots \end{pmatrix}$$

selecting only the rows where $k = 1$, s.t. $w_{l,1} = 1$ just to simplify the matrix constructions.

[20]:
```
Vdm = hcat([[U[i,k]*V[j,k] for i in 1:4 for j in 1:4] for k in 1:4]...)
theta = [H[1][i,j] for i in 1:4 for j in 1:4];

w1 = Vdm\theta
```

[20]:
```
4-element Vector{Float64}:
 0.4470649786874177
 0.15032225861369847
 0.027340579174909594
 0.013831789444849434
```

[21]:

```
normalize_markov!(w1,U,V,W); T1 = tensor(w1,U,V,W); norm(T-T1)
```

[21]: 5.750349421984536e-14

[ ]:

Sparse interpolation

$$f(\mathrm{x}) = \sum_{i=1}^{r} \omega_i\, \mathrm{x}^{\alpha_i} \quad \Rightarrow \quad \sigma = \sum_{\gamma} f(\varphi^{\gamma})\, \frac{\mathrm{y}^{\gamma}}{\gamma!} = \sum_{i=1}^{r} \omega_i\, \mathrm{e}_{\varphi^{\alpha_i}}(\mathrm{y})$$

Example: $f(x_1, x_2) = x_1^{33} x_2^{12} - 5\, x_1 x_2^{45} + 101.$

- Compute $\sigma_\alpha = f(\varphi_1^{\alpha_1}, \varphi_2^{\alpha_2})$ for $\alpha_1 + \alpha_2 \leq 3$ and $\varphi_1 = \varphi_2 = e^{\frac{2\mathrm{i}\pi}{50}}$.
- Compute the Hankel matrix $H_\sigma^{1,2}$:

$$\begin{bmatrix} 97.00000 & 97.01771 + 3.93695\,\mathrm{i} & 95.50360 - 1.47099\,\mathrm{i} & 98.46280 + 4.88062\,\mathrm{i} & 97.42748 + 1.820 \\ 97.01771 + 3.93695\,\mathrm{i} & 98.46280 + 4.88062\,\mathrm{i} & 97.42748 + 1.82098\,\mathrm{i} & 102.35770 + 3.77300\,\mathrm{i} & 99.50853 + 5.294 \\ 95.50360 - 1.47099\,\mathrm{i} & 97.42748 + 1.82098\,\mathrm{i} & 95.73130 - .33862\,\mathrm{i} & 99.50853 + 5.29465\,\mathrm{i} & 95.42134 + 1.472 \end{bmatrix}$$

- Deduce the decomposition of $\sigma = \sum_{i=1}^{3} \omega_i \mathrm{e}_{\xi_i}$:

$$\Xi = \begin{bmatrix} 0.99211 + 0.12533\mathrm{i} & 0.80902 - 0.58779\mathrm{i} \\ 1.00000 + 4.86234e^{-11}\mathrm{i} & 1.00000 - 6.91726e^{-10}\mathrm{i} \\ -0.53583 - 0.84433\mathrm{i} & 0.06279 + 0.99803\mathrm{i} \end{bmatrix} \omega = \begin{bmatrix} -5.00000 - 4.43772e^{-7}\mathrm{i} \\ 101.00000 + 4.65640e^{-7}\mathrm{i} \\ 1.00000 - 1.92279e^{-8}\mathrm{i} \end{bmatrix}$$

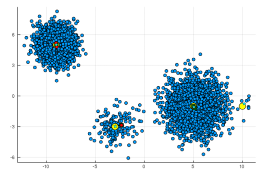
- and the exponents $\frac{50\,\Xi}{2\pi i}$ mod 50 of the terms of f:

$$\begin{bmatrix} 1.00000 - 0.414119e^{-7}\,\mathrm{i} & -5.00000 + 0.270858e^{-6}\,\mathrm{i}, \\ 0.386933e^{-9} + 0.137963e^{-8}\,\mathrm{i} & -0.550458e^{-8} - 0.38761e^{-8}\,\mathrm{i} \\ -17.00000 - 0.100085e^{-6}\,\mathrm{i} & 12.00000 + 0.700984e^{-6}\,\mathrm{i} \end{bmatrix}$$

A mixture of (spherical) Gaussian distributions

$$g(\boldsymbol{y}) = \sum_{k=1}^{r} \omega_k f(\boldsymbol{y}, \mu_k, \sigma_k)$$

where



- $f(\boldsymbol{y}, \mu_k, \sigma_k)$ is the normal distribution of mean $\mu_k \in \mathbb{R}^n$ and covariance $\Sigma_k = \mathrm{diag}(\sigma_k^2) \in \mathbb{R}^{n\times n}$,
- ω_k is the proportion of mixture of the k^{th} normal distribution $f(\boldsymbol{y}, \mu_k, \sigma_k)$.

Theorem:

For $\overline{\sigma}$ the smallest eigenvalue of $\mathbb{E}[\boldsymbol{y} \otimes \boldsymbol{y}] - \mathbb{E}[\boldsymbol{y}] \otimes \mathbb{E}[\boldsymbol{y}]$ and $\boldsymbol{v}$ its unit eigenvector,

- $M_1(\boldsymbol{x}) := \mathbb{E}[\langle \boldsymbol{v}, \boldsymbol{y} - \mathbb{E}[\boldsymbol{y}]\rangle^2 (\boldsymbol{y} \cdot \boldsymbol{x})] = \sum_k \omega_k \, \sigma_k^2 \, (\mu_k \cdot \boldsymbol{x})$
- $M_2(\boldsymbol{x}) := \mathbb{E}[(\boldsymbol{y} \cdot \boldsymbol{x})^2] - \overline{\sigma} \, \|\boldsymbol{x}\|^2 = \sum_k \omega_k \, (\mu_k \cdot \boldsymbol{x})^2$
- $M_3(\boldsymbol{x}) := \mathbb{E}[(\boldsymbol{y} \cdot \boldsymbol{x})^3] - 3M_1(\boldsymbol{x}) \, \|\boldsymbol{x}\|^2 = \sum_k \omega_k \, (\mu_k \cdot \boldsymbol{x})^3$

Gaussian Mixtures

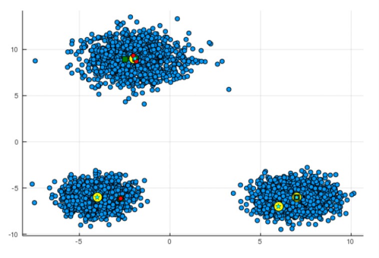
$g(\boldsymbol{x}) = \sum_{k=1}^{r} \omega_k f(\mathrm{x}, \mu_k, \sigma_k)$ where $f(\mathrm{x}, \mu_k, \sigma_k)$ of means μ_k, covariance $\sigma_k^2\, Id$.

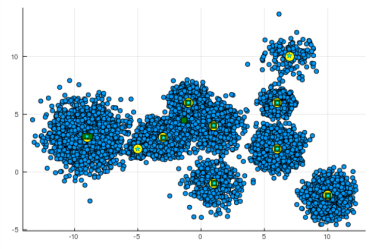
Expectation Maximisation (EM):

$$\max \sum_{i=1}^{p} \log\left(\sum_{k=1}^{r} \omega_k f(\mathrm{x}_i, \mu_k, \sigma_k)\right)$$

by alternate iterative optimization from an initial start.

Comparaison with k-means, split and tensor decomposition:





Examples with $n = 6, r = 4$; $n = 30, r = 10$

See [Khouja-Mattei-M'22]

How the tensor rank differs from the matrix rank?

- $\mathrm{rank}(F) \geq \mathrm{rank}$ of any matricization of F.
- $\mathrm{rank}(F)$ can be **strictly bigger** than the rank of its matricization.
- The rank for a random/generic multilinear (resp. symmetric) tensor is $\lceil \frac{\prod n_i}{\sum n_i} \rceil$ (resp. $r_g = \lceil \frac{\binom{d+n-1}{d}}{n} \rceil$) except in few exceptional cases (symmetric tensor exceptions: $d = 2$ and $(d, n) \in \{(3,5), (4,3), (4,4), (4,5)\}$, Alexander-Hirschovitz'95).
- The rank of a tensor can be **strictly bigger** than r_g (not for matrices).
- The rank decomposition of a tensor F is generically **unique** if $\mathrm{rank}(F) < r_g$ (not for matrices).
  There are **infinitely** many rank decompositions if $\mathrm{rank}(F) > r_g$ (not possible for matrices).

Low rank Hankel matrices and Generalized Additive Decomposition

Low rank decomposition of Hankel matrices

Rank 1 Hankel matrices: $H_\xi = [\xi^{\alpha+\beta}]_{\alpha\in A,\beta\in B}$ for some $\xi \in \mathbb{K}^n$ or $\mathbb{P}^n$.

Rank r Hankel matrices are not necessarily the sum of r rank one matrices:

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \neq \lambda_1 \begin{bmatrix} 1 & \xi_1 & \xi_1^2 \\ \xi_1 & \xi_1^2 & \xi_1^3 \\ \xi_1^2 & \xi_1^3 & \xi_1^4 \end{bmatrix} + \lambda_2 \begin{bmatrix} 1 & \xi_2 & \xi_2^2 \\ \xi_2 & \xi_2^2 & \xi_2^3 \\ \xi_2^2 & \xi_2^3 & \xi_2^4 \end{bmatrix}$$

but

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \lim_{\epsilon\to 0} \frac{1}{2\,\epsilon} \begin{bmatrix} 1 & \epsilon & \epsilon^2 \\ \epsilon & \epsilon^2 & \epsilon^3 \\ \epsilon^2 & \epsilon^3 & \epsilon^4 \end{bmatrix} - \frac{1}{2\,\epsilon} \begin{bmatrix} 1 & -\epsilon & \epsilon^2 \\ -\epsilon & \epsilon^2 & -\epsilon^3 \\ \epsilon^2 & -\epsilon^3 & \epsilon^4 \end{bmatrix}$$

Generalized Additive Decomposition

Recall: For an orthogonal basis $\boldsymbol{u} = (\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n)$ of $\mathbb{K}^n$ and $P \in \mathcal{S}_{n,d}$,

$$\langle (\boldsymbol{u}_1, \boldsymbol{x})^{d-k} (\boldsymbol{u}_2, \boldsymbol{x})^{\beta_2} \cdots (\boldsymbol{u}_n, \boldsymbol{x})^{\beta_n}, P \rangle_d = \frac{1}{d \cdots (d-k+\beta_1+1)} D_{\boldsymbol{u}_2}^{\beta_2} \cdots D_{\boldsymbol{u}_n}^{\beta_n}(P)(\boldsymbol{u}_1)$$

Theorem: Weighted Sum of Dirac Differentials (WSD)

$$F(\boldsymbol{x}) = \textstyle\sum_{i=1}^{r'} w_i(\boldsymbol{x}) (\xi_i, \boldsymbol{x})^{d-k_i}$$

with $w_i(\boldsymbol{x}) = \sum_{|\beta|=k_i} \omega_{i,\beta} \, (\xi_i \cdot \boldsymbol{x})^{\beta_1} \prod_{j=2}^{n} (\zeta_{i,j} \cdot \boldsymbol{x})^{\beta_j}$, $(\xi_i, \zeta_{i,2}, \ldots, \zeta_{i,n})$ ortho. b.

$$\Leftrightarrow$$

$$F^* = \Big(\sum_{i=1}^{r'} \sum_{|\beta|=k_i} \omega_{i,\beta} \frac{(d-k_i+\beta_1)!}{d!} \mathrm{e}_\xi \circ \prod_{j=2}^{n} D_{\zeta_{i,j}}^{\beta_j} \Big)^{[d]}$$

If $\xi_{i,1} = 1$, then

$$\begin{aligned} \check{F} &= \Big(\textstyle\sum_{i=1}^{r} \sum_{|\beta|=k_i} \omega_{i,\beta} \frac{(d-k_i+\beta_1)!}{d!} (\zeta_{i,2}, \boldsymbol{z})^{\beta_2} \cdots (\zeta_{i,n}, \boldsymbol{z})^{\beta_n} \, \mathrm{e}_\xi(\boldsymbol{z}) \Big)^{[d]} \\ &= \Big(\textstyle\sum_{i=1}^{r'} \check{\omega}_i(\boldsymbol{z}) \, \mathrm{e}_{\xi_i}(\boldsymbol{z}) \Big)^{[d]} \in \mathcal{R}_{n-1,d}^* \end{aligned}$$

Definition (Generalized Additive Decomposition)

find r', $w_i(\boldsymbol{x}) \in \mathcal{S}_{n,k_i}$ and $\Xi = [\xi_1, \dots, \xi_r] \in \mathbb{K}^{n \times r'}$ such that

$$F = \sum_{i=1}^{r'} w_i(\boldsymbol{x})\,(\xi_i, \boldsymbol{x})^{d-k_i}$$

with $\sum_{i=1}^{r'} \dim\langle\langle \breve{\omega}_i \rangle\rangle$ **minimal**.

Example: For $d > 5$, $F = x_0^{d-1} x_1 + (x_0 + x_1 + 2\,x_2)^{d-2}(x_0 - x_1)^2$ is a GAD of

$$\begin{aligned} \mathrm{rank}_{gad}(F) &= \dim\langle\langle z_1 \rangle\rangle + \dim\langle\langle (z_1 - 1)^2 \rangle\rangle \\ &= \dim\langle 1, z_1 \rangle + \dim\langle 1, z_1 - 1, (z_1 - 1)^2 \rangle = 5 \end{aligned}$$

with $\xi_1 = [1, 0, 0]$, $\xi_2 = [1, 1, 2]$.

Geometric point of view

- $\mathcal{V}_{n,d} = \{\omega\,(\xi, \boldsymbol{x})^d, \omega \in \mathbb{K}, \xi \in \mathbb{K}^n \xi \neq 0\}$ **Veronese** variety
- $\mathcal{T}_{n,d} = \{\omega(\boldsymbol{x})\,(\xi, \boldsymbol{x})^{d-1}, \omega(\boldsymbol{x}) \in \mathcal{S}_{n,1}, \xi \in \mathbb{K}^n \xi \neq 0\}$ **tangential** variety (= points on tangents to $\mathcal{V}_{n,d}$).
- $\mathcal{O}^k_{n,d} = \{\omega(\boldsymbol{x})\,(\xi, \boldsymbol{x})^{d-k}, \omega(\boldsymbol{x}) \in \mathcal{S}_{n,k}, \xi \in \mathbb{K}^n \xi \neq 0\}$ **osculating** variety (= points on oscullating linear spaces to $\mathcal{V}_{n,d}$).

Proposition

The singular locus of $\mathcal{O}^k_{n,d}$ is $\mathcal{O}^{k-1}_{n,d}$.

$$F = \sum_{i=1}^{r'} \omega_i(\boldsymbol{x})(\xi_i, \boldsymbol{x})^{d-k_i} \quad \text{iff} \quad F \in \sum_{i=1}^{r'} \mathcal{O}^{k_i}_{d,n}$$

For $F = \sum_{i=1}^{r'} \omega_i(\boldsymbol{x})(1 + (\xi_i, \boldsymbol{x}))^{d-k_i}$, let $A, A' \subset \mathcal{R}_n$ and

- $I_F := (\ker H_{\check{F}}^{A',A}) \subset \mathcal{R}_n (= \mathcal{R})$.
- $\mathcal{A}_F := \mathcal{R}/I_F$ the quotient algebra by I_F.

Theorem:

Let $B \subset A, B' \subset A'$ s.t. $B^+ = B \cup x_1 B \cup \cdots \cup x_n B \subset A$, $B'^+ \subset A'$ and $|B| = |B'| = r$. Assume that $\mathrm{rank}_{\check{F}}^{B',B} = \mathrm{rank} H_{\check{F}}^{A',A} = r$.

- $\mathrm{rank}_{gad}(F) = r$
- $\mathcal{A}_F = \mathcal{A}_1 \oplus \cdots \oplus \mathcal{A}_{r'}$ with $\mathcal{A}_i^* = \langle\langle \check{\omega}_i(\boldsymbol{z}) \rangle\rangle \mathrm{e}_{\xi_i}(\boldsymbol{z})$
- $\mathrm{rank}_{gad}(F) = \mathrm{rank}(H_{\check{F}}^{A',A}) = r = \mu_1 + \cdots \mu_{r'}$ where $\mu_i = \dim\langle\langle \check{\omega}_i(\boldsymbol{z}) \rangle\rangle$
- $H_0 = H_{\check{F}}^{B',B}$ is **invertible**.
- For $H_i = H_{\check{F}}^{B',x_i B}$, $M_i = H_0^{-1} H_i =$ **multiplication** by x_i in the basis B of $\mathcal{A}_F$.

☞ GAD via joint triangularization (e.g. joint Schur factorization) of the M_i.

Different notions of rank of a tensor $F \in \mathcal{S}_{n,d}$:

- **rank(F)** = minimal r such that $F = \sum_{i=1}^{r} \omega_i \, (\xi_i, \boldsymbol{x})^d$ or $F^* = (\sum_{i=1}^{r} \omega_i \, \mathrm{e}_{\xi_i})^{[d]}$ or $F^* \in I_{\Xi}^{\perp}$
- $\mathbf{rank}_{gad}$**(F)** = minimal r such that $F = \sum_{i=1}^{r'} \omega_i(\boldsymbol{x}) \, (\xi_i, \boldsymbol{x})^d$ or $\check{F} = (\sum_{i=1}^{r'} \breve{\omega}_i(\boldsymbol{z}) \, \mathrm{e}_{\xi_i}(\boldsymbol{z}))^{[d]}$ with $\sum_i \dim\langle\langle \breve{\omega}_i(\boldsymbol{z}) \rangle\rangle = r$
- $\mathbf{rank}_{\mathrm{border}}$**(F)** = minimal r s.t. $F = \lim_{n \to \infty} F_n$ with $\mathrm{rank}(F_n) = r$
- $\mathbf{rank}_{cactus}$**(F)** = minimal r s.t. $F^* \in I^{\perp}$ and $\dim(\mathcal{S}/I)_{[l]} = r$ for $l \gg 0$.

Many open questions

- How to compute the rank of specific tensors (e.g. matrix multiplication tensors) ?
- How to compute $\mathrm{rank}(F)$, $\mathrm{rank}_{gad}(T)$ ...when it is hight but not too hight (less than the generic rank)?
- How to characterize algebraically varieties of tensors of given rank_{border}, (resp. rank_{gad}), (resp. rank_{cactus})?
- How to find close low rank tensors ? the best low rank tensor approximation ?
- ...

J. Alexander and A. Hirschowitz.
La méthode d'Horace éclatée : application à l'interpolation en degré quatre.
Inventh. math., 107:585–602, 1992.

Gaspard Riche Baron de Prony.
Essai Expérimental et Analytique: Sur Les Lois de La Dilatabilité de Fluides Élastiques et Sur Celles de La Force Expansive de La Vapeur de l'alcool, á Différentes Températures.
J. Ecole Polyt., 1:24–76, 1795.

Alessandra Bernardi, Jérôme Brachat, Pierre Comon, and Bernard Mourrain.
General tensor decomposition, moment matrices and applications.
Journal of Symbolic Computation, 52:51–71, 2013.

Alessandra Bernardi, Jérôme Brachat, and Bernard Mourrain.
A comparison of different notions of ranks of symmetric tensors.
Linear Algebra and its Applications, 460:205-230, 2014.

Alessandra Bernardi, Alessandro Oneto, and Daniele Taufer.
On schemes evinced by generalized additive decompositions and their regularity.
Journal de Mathématiques Pures et Appliquées, 188:446-469, 2024.

Jérôme Brachat, Pierre Comon, Bernard Mourrain, and Elias P. P. Tsigaridas.
Symmetric tensor decomposition.
Linear Algebra and its Applications, 433(11-12):1851–1872, 2010.

Luca Chiantini, Giorgio Ottaviani, and Nick Vannieuwenhoven.
On generic identifiability of symmetric tensors of subgeneric rank.
Transactions of the American Mathematical Society, 369(6):4021-4042, 2016.

Pierre Comon, Gene Golub, Lek-Heng Lim, and Bernard Mourrain.
Symmetric tensors and symmetric tensor rank.
SIAM Journal on Matrix Analysis and Applications, 30(3):1254–1279, September 2008.

D. Cox, J. Little, and D. O'Shea.
Ideals, Varieties, and Algorithms.
Undergraduate Texts in Mathematics. Springer-Verlag, New York, 2nd edition, 1997.

M. Elkadi and B. Mourrain.
***Introduction á la résolution des systémes polynomiaux*, volume 59 of *Mathématiques et Applications*.**
Springer, 2007.

A. Iarrobino and V. Kanev.
***Power sums, Gorenstein algebras, and determinantal loci*, volume 1721 of *Lecture Notes in Computer Science*.**
Springer-Verlag, Berlin, 1999.

Cédric Josz, Jean-Bernard Lasserre, and Bernard Mourrain.
Sparse polynomial interpolation: sparse recovery, super resolution, or Prony?
Advances in Computational Mathematics, 45(3):1401–1437, 2019.

Rima Khouja, Pierre-Alexandre Mattei, and Bernard Mourrain.
Tensor decomposition for learning Gaussian mixtures from moments.
Journal of Symbolic Computation, 113:193–210, 2022.

J. M. Landsberg.
Tensors: Geometry and applications.
Graduate Studies in Mathematics 128. Providence, RI: American Mathematical Society (AMS). 439 p. , 2012.

Monique Laurent and Bernard Mourrain.
A generalized flat extension theorem for moment matrices.
Archiv der Mathematik, 93(1):87–98, 2009.

Bernard Mourrain.
Polynomial-Exponential Decomposition from Moments.
Foundations of Computational Mathematics, 18(6):1435–1492, 2018.

Bernard Mourrain.
From moments to sparse representations, an algebraic, geometric and algorithmic viewpoint.
Lecture, 2019.

James Joseph Sylvester.
***Essay on Canonical Form*.**
The collected mathematical papers of J. J. Sylvester, Vol. I, Paper 34, Cambridge University Press. 1909 (XV und 688). G. Bell, 1851.