Optimization-Aided Construction of Multivariate Chebyshev Polynomials
Supplementary Material

M. Dressler*, S. Foucart†, E. de Klerk‡, M. Joldes§, J. B. Lasserre¶, and Y. Xu‖

In this additional material, we discuss, not too rigorously, how our optimization procedure should be refined in order to deal with richer computational problems and thus open the way for a multivariate version of Basc [3]. Our first two points advocate the use of trigonometric moments instead of monomial moments, the latter being the core of GlobtiPoly 3. Our main premonition is that the resulting positive semidefinite matrices will inherit a Toeplitz structure instead of a Hankel structure, which seems to be much better suited numerically. The third and final point outlines a way to solve the best approximation problem in L_1-norm.

The use of trigonometric moments. Besides the aforementioned numerical stability, enabling one to attack problems of larger size, other advantages of trigonometric moments include fewer localization conditions (since measures are automatically restricted to the hypercube), the fact that polynomials are better represented via tensor products of univariate Chebyshev polynomials (at least in Chebfun, Chebfun2, and Chebfun3, [1, 8, 6]), and the fact that the semidefinite programs at stake could naturally be solved via the software CVX [4] instead of GlobtiPoly 3 [5].

We outline below a general strategy to approximate a monomial $m_{(k_1,\ldots,k_d)}$ with $k_1+\cdots+k_d = n$ by elements from $\mathcal{P}^d_{n-1}$ relatively to the L_∞-norm on a basic semialgebraic set $\Omega \subseteq [-1,1]^d$. Observe first that the question is equivalent to the approximation of $f(x) = T_{k_1}(x_1)\cdots T_{k_d}(x_d)$. We start again from the dual formulation

$$E_{\mathcal{P}^d_{n-1}}(f,\Omega) = \max_{\lambda \in C(\Omega)^*} \lambda(f) \qquad \text{subject to } \lambda_{|\mathcal{P}^d_{n-1}} = 0 \text{ and } \|\lambda\|_{C(\Omega)^*} = 1.$$

We then identify $\lambda \in C(\Omega)^*$ with a measure ν on $[-1,1]^d$ supported on Ω, which itself is identified with a measure μ on $[0,\pi]^d$ supported on some $\widetilde{\Omega}$ via

$$\lambda(h) = \int_{[-1,1]^d} h(x_1,\ldots,x_d)d\nu(x_1,\ldots,x_d) = \int_{[0,\pi]^d} h(\cos(\theta_1),\ldots,\cos(\theta_d))d\mu(\theta_1,\ldots,\theta_d), \quad h \in C(\Omega).$$

*School of Mathematics and Statistics, University of New South Wales, Australia, m.dressler@unsw.edu.au

†Department of Mathematics, Texas A&M University, United States, foucart@tamu.edu

‡Department of Econometrics and Operations Research, Tilburg University, The Netherlands, E.deKlerk@tilburguniversity.edu

§LAAS-CNRS, France, joldes@laas.fr

¶LAAS-CNRS and Toulouse School of Economics, France, lasserre@laas.fr

‖Department of Mathematics, University of Oregon, United States, yuan@uoregon.edu

Likewise, the nonnegative measures $\nu^\pm$ in the Jordan decomposition $\nu = \nu^+ - \nu^-$ are identified with nonnegative measures $\mu^\pm$ in the Jordan decomposition $\mu = \mu^+ - \mu^-$. Notice that $\lambda(f)$ is the kth trigonometric moment of μ, since

$$\lambda(f) = \int_{[-1,1]^d} T_{k_1}(x_1)\cdots T_{k_d}(x_d) d\nu(x_1,\dots,x_d) = \int_{[0,\pi]^d} \cos(k_1\theta_1)\cdots\cos(k_d\theta_d) d\mu(\theta_1,\dots,\theta_d) =: y_k.$$

We use a similar notation for the trigonometric moments of $\mu^\pm$, namely

$$y_\ell^\pm := \int_{[0,\pi]^d} \cos(\ell_1\theta_1)\cdots\cos(\ell_d\theta_d) d\mu^\pm(\theta_1,\dots,\theta_d).$$

In this way, the constraint $\|\lambda\|_{C(\Omega)^*} = 1$ reads $y_0^+ + y_0^- = 1$, since

$$\|\lambda\|_{C(\Omega)^*} = \int_{[-1,1]^d} d|\nu| = \int_{[0,\pi]^d} d|\mu| = \int_{[0,\pi]^d} d(\mu^+ + \mu^-) = y_0^+ + y_0^-,$$

while the constraint $\lambda_{|\mathcal{P}_{n-1}^d} = 0$ reads $y_\ell^+ - y_\ell^- = 0$ whenever $|\ell| < n$, since

$$\lambda(T_{\ell_1}(x_1)\cdots T_{\ell_d}(x_d)) = \int_{[0,\pi]^d} \cos(\ell_1\theta_1)\cdots\cos(\ell_d\theta_d) d\mu(\theta_1,\dots,\theta_k) = y_\ell = y_\ell^+ - y_\ell^-.$$

All in all, replacing the measure μ by its sequence of moments $y \in \mathbb{R}^{\mathbb{N}^d}$, we arrive at

$$E_{\mathcal{P}_{n-1}^d}(f) = \sup_{y^\pm \in \mathbb{R}^{\mathbb{N}^d}} (y_k^+ - y_k^-) \quad \text{s.to} \;\; y_0^+ + y_0^- = 1,\; y_\ell^+ - y_\ell^- = 0 \text{ whenever } |\ell| < n,$$
$$\text{and } \; y^\pm \text{ represent nonnegative measures on } [0,\pi]^d \text{ supported on } \widetilde{\Omega}.$$

The latter constraint has two components (the second one being unnecessary if $\Omega = [-1,1]^d$):

1. by invoking the discrete multilinear trigonometric moment problem, the fact that $y^\pm$ represent nonnegative measures on $[0,\pi]^d$ translates into the positive semidefiniteness of the moment matrices $M(y^\pm) \in \mathbb{R}^{\mathbb{N}^d \times \mathbb{N}^d}$ with entries

$$M(y^\pm)_{\ell,\ell'} = \text{moment}_{|\ell-\ell'|}(\mu^\pm) = y^\pm_{|\ell-\ell'|}, \qquad \ell,\ell' \in \mathbb{N}^d.$$

2. the fact that $y^\pm$ represent measures that are supported on $\widetilde{\Omega}$ translates, say in the exemplary case of the shifted simplex

$$\Omega = \{(x_1,\dots,x_d) \in [-1,1]^d : x_1 + x_2 + \cdots + x_d \le 2 - d\},$$

into the fact that $(2 - d - x_1 - \cdots - x_d)\nu^\pm(x_1,\dots,x_d)$ are nonnegative measures on $[-1,1]^d$, or equivalently that

$$\widetilde{\mu}^\pm(\theta_1,\dots,\theta_d) := (2 - d - \cos(\theta_1) - \cdots - \cos(\theta_d))\mu^\pm(\theta_1,\dots,\theta_d)$$

are nonnegative measures on $[0,\pi]^d$. This is equivalent to the positive semidefiniteness of the matrices $N(y^\pm) \in \mathbb{R}^{\mathbb{N}^d \times \mathbb{R}^d}$ with entries

$$\begin{aligned}
&N(y^\pm)_{\ell,\ell'} = \mathrm{moment}_{|\ell-\ell'|}(\widetilde{\mu}^\pm) \\
&= \int_{[0,\pi]^d} \cos((\ell_1-\ell_1')\theta_1)\cdots\cos((\ell_d-\ell_d')\theta_d)(2-d-\cos(\theta_1)-\cdots-\cos(\theta_d))d\mu^\pm(\theta_1,\ldots,\theta_d) \\
&= (2-d)\,\mathrm{moment}_{|\ell-\ell'|}(\mu^\pm) \\
&- \sum_{i=1}^d \int_{[0,\pi]^d} \cos((\ell_1-\ell_1')\theta_1)\cdots\cos(\theta_i)\cos((\ell_i-\ell_i')\theta_i)\cdots\cos((\ell_d-\ell_d')\theta_d)d\mu^\pm(\theta_1,\ldots,\theta_d).
\end{aligned}$$

In view of the identity $\cos(a)\cos(b) = (\cos(a+b)+\cos(a-b))/2$, we derive

$$\begin{aligned}
&N(y^\pm)_{\ell,\ell'} = (2-d)\, y^\pm_{|\ell-\ell'|} \\
&- \sum_{i=1}^d \int_{[0,\pi]^d} \cos((\ell_1-\ell_1')\theta_1)\cdots\frac{\cos((\ell_i-\ell_i'+1)\theta_i)}{2}\cdots\cos((\ell_d-\ell_d')\theta_d)d\mu^\pm(\theta_1,\ldots,\theta_d) \\
&- \sum_{i=1}^d \int_{[0,\pi]^d} \cos((\ell_1-\ell_1')\theta_1)\cdots\frac{\cos((\ell_i-\ell_i'-1)\theta_i)}{2}\cdots\cos((\ell_d-\ell_d')\theta_d)d\mu^\pm(\theta_1,\ldots,\theta_d).
\end{aligned}$$

Altogether, this reads

$$N(y^\pm)_{\ell,\ell'} = (2-d)y^\pm_{|\ell-\ell'|} - \frac{1}{2}\sum_{i=1}^d \big(y^\pm_{|\ell-\ell'+\delta_i|} + y^\pm_{|\ell-\ell'-\delta_i|}\big).$$

In summary, the best approximation error $E_{\mathcal{P}^d_{n-1}}(f)$ can be computed by solving the semidefinite program (note that $M(y^\pm), N(y^\pm)$ depend linearly on $y^\pm$):

$$\underset{y^\pm\in\mathbb{R}^{\mathbb{N}^d}}{\text{maximize}}\ (y_n^+ - y_n^-) \qquad \text{s.to}\ \ y_0^+ + y_0^- = 1,\ y_\ell^+ - y_\ell^- = 0 \text{ whenever } |\ell| < n,$$
$$\text{and}\ \ M(y^\pm)\succeq 0,\ N(y^\pm)\succeq 0.$$

In practice, of course, the infinite vectors $y^\pm \in \mathbb{R}^{\mathbb{N}^d}$ are truncated to finite vectors $y^\pm \in \mathbb{R}^{N^d}$. This provides lower bounds for $E_{\mathcal{P}^d_{n-1}}(f)$, but it seems that these lower bounds agree with the genuine value even when the parameter N is moderate.

Atom extraction from Chebyshev moments. One of the advantages of GlobtiPoly 3 is that, if a measure is atomic, then it can extract the atoms from finitely many monomial moments. Actually, this is also valid for Chebyshev moments. The following procedure reveals how to do so, albeit in the univariate setting. It is an adaptation of Prony's method. Suppose that an atomic measure

$$\nu = \sum_{i=1}^s c_i \delta_{x^{(i)}}, \qquad x^{(1)},\ldots,x^{(s)} \in [-1,1], \quad c_1,\ldots,c_s \in \mathbb{R},$$

is available through its first $2s$ Chebyshev moments $y_0, \ldots, y_{2s-1}$, where

$$y_\ell = \int_{-1}^{1} T_\ell(x) d\nu(x) = \sum_{i=1}^{s} c_i T_\ell(x^{(i)}).$$

The goal is to deduce the unknowns $x^{(1)}, \ldots, x^{(s)}$ and $c_1, \ldots, c_s$ from these moments. Consider the degree-s polynomial p vanishing exactly at the $x^{(i)}$'s and expand it in the Chebyshev basis as

$$p(x) := 2^{s-1}(x - x^{(1)}) \cdots (x - x^{(s)}) = \sum_{j=0}^{s} \gamma_j T_j(x), \qquad \text{where } \gamma_s = 1.$$

For any $\ell = 0, 1, \ldots, s-1$, observe that

$$\begin{aligned} 0 &= \sum_{i=1}^{s} c_i T_\ell(x^{(i)}) p(x^{(i)}) = \sum_{i=1}^{s} c_i T_\ell(x^{(i)}) \sum_{j=0}^{s} \gamma_j T_j(x^{(i)}) = \sum_{j=0}^{s} \gamma_j \sum_{i=1}^{s} c_i T_\ell(x^{(i)}) T_j(x^{(i)}) \\ &= \sum_{j=0}^{s} \gamma_j \sum_{i=1}^{s} c_i \frac{T_{\ell+j}(x^{(i)}) + T_{|\ell-j|}(x^{(i)})}{2} = \sum_{j=0}^{s} \gamma_j \frac{y_{\ell+j} + y_{|\ell-j|}}{2}. \end{aligned}$$

Using $\gamma_s = 1$, this identity rearranges into s linear equations in the s unknowns $\gamma_0, \ldots, \gamma_{s-1}$, namely (note that the coefficients of the system are known due to the availability of $y_0, \ldots, y_{2s-1}$)

$$\sum_{j=0}^{s-1} (y_{\ell+j} + y_{|\ell-j|}) \gamma_j = -(y_{\ell+s} + y_{s-\ell}), \qquad \ell = 0, 1, \ldots, s-1.$$

Solving this linear system provides the values of $\gamma_0, \ldots, \gamma_{s-1}$. Next, the unknown $x^{(1)}, \ldots, x^{(s)}$ are obtained as the roots of the polynomial $\sum_{j=0}^{s} \gamma_j T_j(x)$, computed as the eigenvalues of the colleague matrix C expressed (taking $\gamma_s = 1$ into account) as

$$C = \frac{1}{2} \begin{bmatrix} 0 & 2 & 0 & \cdots & 0 \\ 1 & 0 & 1 & \ddots & \vdots \\ 0 & 1 & 0 & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & 1 \\ 0 & \cdots & 0 & 1 & 0 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} 0 & \cdots & \cdots & 0 \\ \vdots & & & \vdots \\ \vdots & & & \vdots \\ 0 & \cdots & \cdots & 0 \\ \gamma_0 & \gamma_1 & \gamma_{s-2} & \gamma_{s-1} \end{bmatrix}.$$

Now that $x^{(1)}, \ldots, x^{(s)}$ are known, the values of $c_1, \ldots, c_s$ are deduced from the defining equations of $y_0, \ldots, y_{s-1}$, interpreted as s linear equations with unknowns $c_1, \ldots, c_s$.

We mention in passing that, in the multivariate setting, there is an analog of the relation between the roots of a polynomial expressed in the monomial basis and the eigenvalues of the companion matrix: this is Stickelberger eigenvalue theorem, see e.g. [7, Theorem 2.9]. For polynomials expressed in the Chebyshev basis and the colleague matrix, we are not aware of what the multivariate analog is.

Dealing with the L_1-norm. We conclude by outlining a procedure to compute the error of best approximation from $\mathcal{P}^d_{n-1}$ to a monomial m_k—or any polynomial $f \in \mathcal{P}^d_N$ with $N \geq n$—relative to the L_1-norm on a basic semialgebraic set Ω. In other words, we want to solve the problem

$$\underset{p \in \mathcal{P}^d_{n-1}}{\text{minimize}} \, \|f - p\|_{L_1(\Omega)}, \qquad \text{where} \quad \|f - p\|_{L_1(\Omega)} = \int_\Omega |(f-p)(x)| dx.$$

Fixing $p \in \mathcal{P}^d_{n-1}$ for now, with the Jordan decomposition of the signed measure $(f-p)(x)dx$ in the back of our mind, we observe that

$$\|f - p\|_{L_1(\Omega)} = \inf_{\mu^\pm} \int d(\mu^+ + \mu^-)(x) \qquad \text{subject to} \quad d(\mu^+ - \mu^-)(x) = (f-p)(x)dx,$$

with infimum taken over nonnegative Borel measures μ^+ and μ^- supported on Ω. Then, viewing the measures $\mu^\pm$ in terms of their sequences $y^\pm \in \mathbb{R}^{\mathbb{N}^d}$ of moments—monomial or trigonometric—we notice that the objective function is simply $y^+_0 + y^-_0$ while the constraint reads $y^+_\ell - y^-_\ell = z_\ell(p)$, $\ell \in \mathbb{N}^d$, for some sequence $z(p) \in \mathbb{R}^{\mathbb{N}^d}$ depending affinely on p. Thus, adding moment and localization constraints stating that some infinite matrices $M(y^\pm)$ and $N(y^\pm)$ depending linearly on $y^\pm$ are positive semidefinite, we arrive at

$$\|f - p\|_{L_1(\Omega)} = \inf_{y^\pm \in \mathbb{R}^{\mathbb{N}^d}} \; y^+_0 + y^-_0 \qquad \text{subject to } y^+ - y^- = z(p), \; M(y^\pm) \succeq 0, \; N(y^\pm) \succeq 0.$$

At this point, unfixing $p \in \mathcal{P}^d_{n-1}$, we can see that the above expression can further be minimized over this approximating polynomial, leading to an infinite-dimensional semidefinite program with variables $p \in \mathcal{P}^d_{n-1}$ and $y^\pm \in \mathbb{R}^{\mathbb{N}^d}$. In practice, of course, we truncate the latter to obtain a finite-dimensional semidefinite program whose minimizers can be shown to converge to the sought-after minimizer when the truncation parameter grows. We even expect the availability of error esimates, albeit *a posteriori* ones, similar to the ones derived carefully in [2] for a univariate setting.

References

[1] T. A. Driscoll, N. Hale, and L. N. Trefethen (eds.). Chebfun Guide. Pafnuty Publications, Oxford, 2014.

[2] S. Foucart and J. B. Lasserre. *Computation of Chebyshev polynomials for union of intervals.* Computational Methods and Function Theory, 19/4, 625–641, 2019.

[3] S. Foucart and V. Powers. Basc*: constrained approximation by semidefinite programming.* IMA Journal of Numerical Analysis, 37/2, 1066–1085, 2017.

[4] M. Grant and S. Boyd. CVX: Matlab software for disciplined convex programming, version 2.0 beta. `https://cvxr.com/cvx`, September 2013.

[5] D. Henrion, J. B. Lasserre, and J. Loefberg. *GloptiPoly 3: moments, optimization and semidefinite programming.* Optimization Methods and Software, 24/4-5, 761–779, 2009. `https://homepages.laas.fr/henrion/software/gloptipoly3/`

[6] B. Hashemi and L. N. Trefethen. *Chebfun in three dimensions.* SIAM Journal on Scientific Computing, 39/5, C341–C363, 2017.

[7] M. Laurent. Sums of Squares, Moment Matrices and Optimization over Polynomials. In: Emerging Applications of Algebraic Geometry, Vol. 149 of IMA Volumes in Mathematics and its Applications, M. Putinar and S. Sullivant (eds.), Springer, pp. 157–270, 2009.

[8] A. Townsend and L. N. Trefethen. *An extension of Chebfun to two dimensions.* SIAM Journal on Scientific Computing, 35/6, C495–C518, 2013.