A fun byproduct of my Master's thesis: symmetric tensor functions are dense in the space of permutation-invariant multivariate functions

Guillaume Wang
guillaume.wang@inf.ethz.ch

May 2021

During my Master's thesis, I encountered several interesting technical points that were not directly related to the thesis topic, so that I left them hanging. Here I will talk about one of them, encountered while working on Volterra series. (I chose to jump directly into my main point without giving any context on Volterra series, as it is not necessary; for a clean introduction to these objects, see chapter 4 of my Master's thesis report.)

1 Preliminaries

Notations and shorthands Fix some integer $n > 0$.

- For a point $\boldsymbol{t} = (t_1, ..., t_n) \in \mathbb{R}^n$ and a permutation $\sigma \in \mathfrak{S}_n$, $\boldsymbol{t}_\sigma$ denotes $(t_{\sigma(1)}, ..., t_{\sigma(n)})$.
- Call a multivariate function $g : \mathbb{R}^n \to \mathbb{R}$ *permutation-invariant* if for any permutation σ, it holds $g(\boldsymbol{t}) = g(\boldsymbol{t}_\sigma)$ for all $\boldsymbol{t} \in \mathbb{R}^n$.
- For any function $f : \mathbb{R} \to \mathbb{R}$, denote $f^{\otimes n} : [\mathbb{R}^n \to \mathbb{R}, \boldsymbol{t} \mapsto f(t_1)...f(t_n)]$. Call $f^{\otimes n}$ the associated *symmetric tensor function* [1] – tensor because it is a product of single-variable functions, and symmetric because all of those single-variable functions are the same.
- For any multivariate function $g : \mathbb{R}^n \to \mathbb{R}$, denote $\operatorname{Sym} g : \left[\mathbb{R}^n \to \mathbb{R}, \boldsymbol{t} \mapsto \frac{1}{n!} \sum_{\sigma \in \mathfrak{S}_n} g(\boldsymbol{t}_\sigma)\right]$.

We will sometimes write physicist-style $f(t)$ to mean a function $f : \mathbb{R} \to \mathbb{R}$, and similarly $g(\boldsymbol{t})$ instead of $g : \mathbb{R}^n \to \mathbb{R}$.

Some function spaces Fix $1 \leq p < \infty$ and q its conjugate exponent, i.e $1/p + 1/q = 1$.

- Let $L^p(\mathbb{R})$ be the Banach space of L^p-integrable functions over $\mathbb{R}$ (with the usual Lebesgue measure). Its dual space is $L^q(\mathbb{R})$.

[1] Disclaimer: the term "symmetric tensor function" may not be consistent with standard terminology, I haven't checked.

- Let $C_0(\mathbb{R})$ be the space of vanishing continuous functions over $\mathbb{R}$. Its dual space is $\mathcal{M}(\mathbb{R})$, the space of Radon measures. [2]
- Similarly define $L^p(\mathbb{R}^n)$, $L^q(\mathbb{R}^n)$, $C_0(\mathbb{R}^n)$ spaces of multivariate functions with n scalar variables.
- Denote $L^p_{\mathrm{Sym}}(\mathbb{R}^n)$, $L^q_{\mathrm{Sym}}(\mathbb{R}^n)$, $C_{0\,\mathrm{Sym}}(\mathbb{R}^n)$ the respective (closed) subspaces consisting of permutation-invariant functions.

Note that our shorthand Sym can be viewed as a projection operator from $L^p(\mathbb{R}^n)$ to $L^p_{\mathrm{Sym}}(\mathbb{R}^n)$, and from $C_0(\mathbb{R}^n)$ to $C_{0\,\mathrm{Sym}}(\mathbb{R}^n)$.

2 The result and why it looks surprising to me

Proposition 1. Let $1 \leq p < \infty$. The set $\{f^{\otimes n}(\boldsymbol{t});\ f \in L^p(\mathbb{R})\}$ has its linear span dense in $L^p_{\mathrm{Sym}}(\mathbb{R}^n)$.
The set $\{f^{\otimes n}(\boldsymbol{t});\ f \in C_0(\mathbb{R})\}$ has its linear span dense in $C_{0\,\mathrm{Sym}}(\mathbb{R}^n)$. [3]

More explicitly: *any* $g(\boldsymbol{t}) \in C_{0\,\mathrm{Sym}}(\mathbb{R}^n)$ *is arbitrarily-well uniformly approximated by finite sums of the form* $\sum_{i \leq m} f_i(t_1)...f_i(t_n)$ $(m < \infty,\ f_i \in C_0(\mathbb{R}))$.

This is not Weierstrass with symmetrization As an obvious corollary, the proposition holds when $\mathbb{R}$ is replaced by a closed interval $I \subset \mathbb{R}$. In this case the result looks like a straightforward consequence of the Weierstrass approximation theorem, but it is not. Consider the following valid reasoning:

> Fix a continuous function $g(\boldsymbol{t})$ over the compact I^n and let $\varepsilon > 0$. By the Weierstrass approximation theorem, there exists a polynomial $P(\boldsymbol{t})$ such that $\|g - P\| := \sup_{I^n} |g - P| \leq \varepsilon$, and $P(\boldsymbol{t})$ can be written as $P(\boldsymbol{t}) = \sum_{\alpha \in \mathbb{N}^n} a_\alpha \boldsymbol{t}^\alpha$ (where there are only a finite number of nonzero coefficients a_α and the shorthand $\boldsymbol{t}^\alpha$ denotes $t_1^{\alpha_1}...t_n^{\alpha_n}$).
>
> If in addition $g(\boldsymbol{t})$ is permutation-invariant, then the lemma below shows that $\mathrm{Sym}\, P(\boldsymbol{t})$ is also an ε-approximation of $g(\boldsymbol{t})$, and it can be written as
>
> $$\mathrm{Sym}\, P(\boldsymbol{t}) = \sum_{\alpha \in \mathbb{N}^n} a_\alpha \,\mathrm{Sym}\, \boldsymbol{t}^\alpha = \sum_{\alpha \in \mathbb{N}^n} \sum_{\sigma \in \mathfrak{S}_n} \frac{a_\alpha}{n!} t_1^{\alpha_{\sigma(1)}} ... t_n^{\alpha_{\sigma(n)}}. \tag{1}$$

This does not show that $g(\boldsymbol{t})$ can be approximated by a finite combination of symmetric tensor functions, as the reasoning may yield approximators such as $t_1 t_2^3 + t_1^3 t_2$ (if $n = 2$), which are not of the required form.

Lemma 2 (The aforementioned lemma)**.** For any function $g(\boldsymbol{t})$ over $\mathbb{R}^n$ and any $1 \leq p \leq \infty$, it holds: $\|\mathrm{Sym}\, g\|_{L^p} \leq \|g\|_{L^p}$.
For any permutation-invariant $g(\boldsymbol{t})$ and any function $h(\boldsymbol{t})$, if $\|g - h\|_{L^p} \leq \varepsilon$, then $\|g - \mathrm{Sym}\, h\|_{L^p} \leq \varepsilon$.

[2] https://regularize.wordpress.com/2011/11/11/dual-spaces-of-continuous-functions/
[3] I'm pretty sure the same holds if C_0 is replaced by C_b i.e if we consider bounded continuous functions, instead of vanishing continuous.

Proof. Let any function $g(\boldsymbol{t})$ over $\mathbb{R}^n$ and any $1 \le p \le \infty$. By definition of the L^p norm,

$$\|\mathrm{Sym}\, g\|_{L^p} = \left\| \frac{1}{n!} \sum_{\sigma \in \mathfrak{S}_n} g(\boldsymbol{t}_\sigma) \right\|_{L^p} \le \frac{1}{n!} \sum_{\sigma \in \mathfrak{S}_n} \|g(\boldsymbol{t}_\sigma)\|_{L^p} = \|g\|_{L^p}\,. \tag{2}$$

Let any permutation-invariant function $g(\boldsymbol{t})$ and any function $h(\boldsymbol{t})$ such that $\|g-h\|_{L^p} \le \varepsilon$. Then

$$\|g - \mathrm{Sym}\, h\|_{L^p} = \|\mathrm{Sym}(g-h)\|_{L^p} \le \|g-h\|_{L^p} \le \varepsilon. \tag{3}$$

□

An example of surprise In fact the case of permutation-invariant polynomials over a compact set is already surprising to me... In fact just the following example is already surprising to me:

> Consider the function $g(\boldsymbol{t}) = t_1 + ... + t_n$ over $[0,1]^n$. According to the proposition, there exist $m < \infty$ and $f_i(t) \in C([0,1])$ $(i \le m)$ such that $g(\boldsymbol{t}) \approx \sum_{i \le m} f_i(t_1)...f_i(t_n)$, in the sense of uniform approximation over $[0,1]^n$.

I wonder what these f_i could look like. Since they are continuous over $[0,1]$, according to the Weierstrass approximation theorem we may assume without loss of generality that each f_i is polynomial. Then, developing the product $f_i(t_1)...f_i(t_n)$ would yield an a priori big polynomial, whereas directly using Weierstrass with symmetrization can yield simply $t_1 + ... + t_n$ itself.

3 Brief proof of the result

In this section we prove the L^p/L^q $(1 \le p < \infty)$ part of the proposition; the $C_0/\mathcal{M}$ part can be proved by the same arguments, with minor modifications.

I assume the reader is familiar with the basics of functional analysis and duality in Banach spaces. Recall the following density criterion, which is a consequence of the Hahn-Banach theorem (as are many things):

Lemma 3. Let E be a Banach space and A a subset.

$$\mathrm{span}(A) \text{ is dense in } E \qquad \iff \qquad \{X \in E';\ \forall a \in A, \langle a, X\rangle_E = 0\} = \{0_{E'}\}. \tag{4}$$

The set on the right is sometimes denoted $A^\perp$ and called the *annihilator* of A. Note that $A^\perp = (\mathrm{span}(A))^\perp$. In the case where E is a Euclidean space, $A^\perp$ is just (up to isometry) the orthogonal complement of $\mathrm{span}(A)$.

The proposition will be proved by applying the above density criterion. To do so we will need the following intuitively obvious lemma, characterizing the dual of $L^p_{\mathrm{Sym}}(\mathbb{R}^n)$. A formal and rather uninteresting proof can be found at the end of this document.

Lemma 4. The dual of $L^p_{\mathrm{Sym}}(\mathbb{R}^n)$ is isometrically isomorphic to $L^q_{\mathrm{Sym}}(\mathbb{R}^n)$.

We can now prove the proposition. The main argument is extracted from (Boyd Chua Desoer 1984, Theorem 2.5.2) in the context of Volterra series.

Proof of proposition. To apply the density criterion to $A = \{f^{\otimes n}(\boldsymbol{t});\ f \in L^p(\mathbb{R})\}$ and $E = L^p_{\mathrm{Sym}}(\mathbb{R}^n)$, let $h \in E' \simeq L^q_{\mathrm{Sym}}(\mathbb{R}^n)$ such that $\langle f^{\otimes n}, h\rangle_{L^p} = 0$ for all $f \in L^p(\mathbb{R})$. Let us show that $h = 0$, from which the proposition will follow.

Denote $\Phi_h : L^p(\mathbb{R}) \to \mathbb{R}$ the n-homogeneous map

$$\Phi_h[f] = \left\langle f^{\otimes n}, h\right\rangle_{L^p} = \int_{\mathbb{R}^n} d\boldsymbol{t}\ h(t_1, ..., t_n) f(t_1)...f(t_n) \tag{5}$$

and $\Psi_h : L^p(\mathbb{R})^n \to \mathbb{R}$ the associated n-linear system

$$\Psi_h\{f_1, ..., f_n\} = \langle f_1 \otimes ... \otimes f_n, h\rangle_{L^p} = \int_{\mathbb{R}^n} d\boldsymbol{t}\ h(t_1, ..., t_n) f_1(t_1)...f_n(t_n). \tag{6}$$

The n-linear system $\Psi_h\{\cdot, ..., \cdot\}$ is symmetric in its arguments, since h is permutation-invariant. So Ψ_h is completely determined by the n-homogeneous map $\Phi_h[\cdot]$ via the algebraic polarization identity [4]

$$n!\Psi_h\{f_1, ..., f_n\} = \left.\frac{\partial}{\partial\alpha_1...\partial\alpha_n}\right|_{\boldsymbol{\alpha}=0} \Phi_h\left[\sum_{i=1}^{n} \alpha_i f_i\right], \tag{7}$$

and the right-hand-side is the differential of an identically zero map. Consequently,

$$\forall f_1, ..., f_n \in L^p(\mathbb{R}),\ \ \Psi_h\{f_1, ..., f_n\} = 0. \tag{8}$$

Now evaluate this at $f_1(t) = \mathbb{1}_{t\in A_1}, ..., f_n(t) = \mathbb{1}_{t\in A_n}$ for intervals $A_i \subset \mathbb{R}$:

$$\Psi_h\{f_1, ..., f_n\} = \int_{\mathbb{R}^n} d\boldsymbol{t}\ h(t_1, ..., t_n)\mathbb{1}_{\boldsymbol{t}\in A_1\times...\times A_n} = 0. \tag{9}$$

Since this holds for all A_i, and hyperrectangles generate the Borel σ-algebra, then $h = 0$, as claimed. □

4 Is the result interesting/useful?

For the subjects that I'm currently leaning towards, the result presented in this document is actually pretty useless, as it only talks about approximability per se. It doesn't give any guarantees on the nature nor the number of functions f_i required to ε-approximate a given target function g.

However I still find the result technically interesting and surprising. I never heard about it before but I'm certain it must be somewhere out there already – I would be glad to know where and in what context.

A Proofs

Proof of lemma that $L^p_{\mathrm{Sym}}(\mathbb{R}^n)' \simeq L^q_{\mathrm{Sym}}(\mathbb{R}^n)$. View the shorthand $\mathrm{Sym}_p : g \mapsto \frac{1}{n!}\sum_\sigma g(\boldsymbol{t}_\sigma)$ as an operator from $L^p(\mathbb{R}^n)$ to $L^p_{\mathrm{Sym}}(\mathbb{R}^n)$, and denote $\mathrm{Sym}_p^* : L^p_{\mathrm{Sym}}(\mathbb{R}^n)' \to L^p(\mathbb{R}^n)' \simeq L^q(\mathbb{R}^n)$ its adjoint operator, i.e such that

$$\forall \nu \in L^p_{\mathrm{Sym}}(\mathbb{R}^n)', \forall g \in L^p(\mathbb{R}^n),\ \ \left\langle g, \mathrm{Sym}_p^* \nu\right\rangle_{L^p} = \left\langle \mathrm{Sym}_p g, \nu\right\rangle_{L^p_{\mathrm{Sym}}}. \tag{10}$$

[4]See https://en.wikipedia.org/wiki/Polarization_of_an_algebraic_form.

Since Sym_p is surjective, then Sym_p^* is injective, and it is not hard to check that Sym_p^* is isometric. It remains to check that the image space of Sym_p^* is $L^q_{\mathrm{Sym}}(\mathbb{R}^n)$.

Indeed: for any $\nu \in L^p_{\mathrm{Sym}}(\mathbb{R}^n)'$ and $h = \mathrm{Sym}_p^* \nu \in L^q(\mathbb{R}^n)$,

$$\forall g \in L^p(\mathbb{R}^n),\ \ \langle g, h \rangle_{L^p} = \left\langle \mathrm{Sym}_p g, \nu \right\rangle_{L^p} = \left\langle \mathrm{Sym}_p \mathrm{Sym}_p g, \nu \right\rangle_{L^p} = \left\langle \mathrm{Sym}_p g, h \right\rangle_{L^p} \tag{11}$$

since $\mathrm{Sym} \circ \mathrm{Sym} = \mathrm{Sym}$, and

$$\forall g \in L^p(\mathbb{R}^n),\ \ \left\langle \mathrm{Sym}_p g, h \right\rangle_{L^p} = \frac{1}{n!} \sum_{\sigma \in \mathfrak{S}_n} \langle g(\boldsymbol{t}_\sigma), h(\boldsymbol{t}) \rangle = \frac{1}{n!} \sum_{\sigma \in \mathfrak{S}_n} \langle g(\boldsymbol{t}), h(\boldsymbol{t}_\sigma) \rangle = \left\langle g, \mathrm{Sym}_q h \right\rangle_{L^p}. \tag{12}$$

Thus $h = \mathrm{Sym}_q h$ (since they coincide as elements of $L^p(\mathbb{R}^n)'$), so h is permutation-invariant. Hence $\mathrm{Im}(\mathrm{Sym}_p^*) \subset L^q_{\mathrm{Sym}}(\mathbb{R}^n)$.

Conversely: for any $h \in L^q_{\mathrm{Sym}}(\mathbb{R}^n) \subset L^q(\mathbb{R}^n) \simeq L^p(\mathbb{R}^n)'$, denote $\nu \in L^p_{\mathrm{Sym}}(\mathbb{R}^n)'$ its restriction to $L^p_{\mathrm{Sym}}(\mathbb{R}^n)$. Then one can check that $h = \mathrm{Sym}_p^* \nu$. $\square$