Symmetric Tensor Rank

Shiva Kaul
SELECT Lab 2013

homogeneous polynomials $\longleftrightarrow$ symmetric tensors

On $x \in \mathbb{R}^n$, a homogeneous polynomial with coefficients in $\mathbb{F}$, a field of characteristic 0 :

$$p(x) = 2x_1^2 + 9x_2^2 + 14x_1x_2$$

$$= 2x_1^2x_2^0 + 9x_1^0x_2^2 + 14x_1^1x_2^1$$

$$= 2x^{(2,0)} + 9x^{(0,2)} + 14x^{(1,1)}$$

$$= \sum_{|\alpha|=d} p_\alpha x^\alpha \qquad \alpha = (\alpha_1, \ldots, \alpha_n) \quad \alpha_i \in \{0,1,2,\ldots\} \qquad |\alpha| = \sum_{i=1}$$

$$= \left\langle [p_\alpha]_{|\alpha|=d} \, , \, [x^\alpha]_{|\alpha|=d} \right\rangle$$

multi-indexed vector $\qquad \mathbb{F}^N \leftarrow \binom{n+d-1}{d}$

Polynomial kernel

$$K(x,y)$$

$$= \langle x, y \rangle^d$$

$$= (x_1 y_1 + \cdots + x_n y_n) \cdots (x_1 y_1 + \cdots + x_n y_n)$$

$$= \sum_{\alpha} \binom{d}{\alpha} x^{\alpha} y^{\alpha} \qquad \binom{d}{\alpha} = \frac{d!}{\alpha_1! \, \alpha_2! \cdots \alpha_n!}$$

$$= \Big\langle \underbrace{\left[\binom{d}{\alpha}^{1/2} x^{\alpha} \right]_{|\alpha| = d}}_{\phi_d(x)}, \underbrace{\left[\binom{d}{\alpha}^{1/2} y^{\alpha} \right]_{|\alpha| = d}}_{\phi_d(y)} \Big\rangle \qquad : \mathbb{R}^n \to \mathbb{F}^N$$

Tensor products

$$\mathbb{F}^n \otimes \mathbb{F}^n = \mathbb{F}^{n^2}$$

$e_1, \ldots, e_n$ $\quad$ $e'_1, \ldots, e'_n$ $\quad$ $e_i \otimes e'_j = e_i e'^T_j$

A symmetric tensor $t \in \mathbb{F}^{n^d}$ satisfies

$$t_{j_1, j_2, \ldots, j_d} = t_{\pi(j_1), \pi(j_2), \ldots, \pi(j_d)}$$

for $j_i \in \{1, \ldots, n\}$ and permutation π

size of basis

When $d=2$, dimension is $\binom{n}{2} + n = \binom{n+1}{2}$

In general, it's $\binom{n+d-1}{d} = N$

Decomposition into rank-one parts

$$p = \sum_{j=1}^{r} \phi_d(z_j) \qquad\qquad p = \sum_{j=1}^{r} \underbrace{z_j \otimes \cdots \otimes z_j}_{d \text{ times}}$$

These are the same thing.

$$\langle \phi_2(x), \phi_2(y) \rangle = \langle x, y \rangle^2$$
$$= \left(\sum_{i=1}^{n} x_i y_i\right)\left(\sum_{j=1}^{n} x_i y_i\right)$$
$$= \sum_{i,j=1}^{n} x_i x_j y_i y_j = \langle x x^T, y y^T \rangle$$

so $\phi_2(x) \equiv x \otimes x$.

$\phi_4(x) = \phi_2(\phi_2(x)) \quad (\langle x, y\rangle^4 = \langle x x^T, y y^T \rangle^2)$ and so forth.

no coefficients: $c\langle x, y\rangle^d = \langle c^{1/d} x, y \rangle^d$

$S \subseteq \mathbb{F}^N$ is the coefficient-less span of $\{\phi_d(z) : z \in \cancel{\mathbb{R}^n\}}$

$\mathrm{srk}(p)$ = smallest r s.t. $p(x) = \sum_{j=1}^{r} \langle z_j, x \rangle^d$

Studied since 19th century.

For any n, d:

$\exists\, p \in S$ s.t. $\qquad\qquad \forall\, p \in S$,

$$\lceil N/n \rceil \leq \mathrm{srk}(p) \leq \lceil N/n \rceil$$

except for $(d,n) = (3,5)$, $(4,3)$, $(4,4)$, or $(4,5)$, which need one more term.

Known since at least '80.

Alexander - Hirschowitz '95

Very difficult and deep.

Want to understand these decompositions for ML.

- graphical model inference
- kernel approximations
- dimension reduction
- Kaul-Gordon TBA ☺

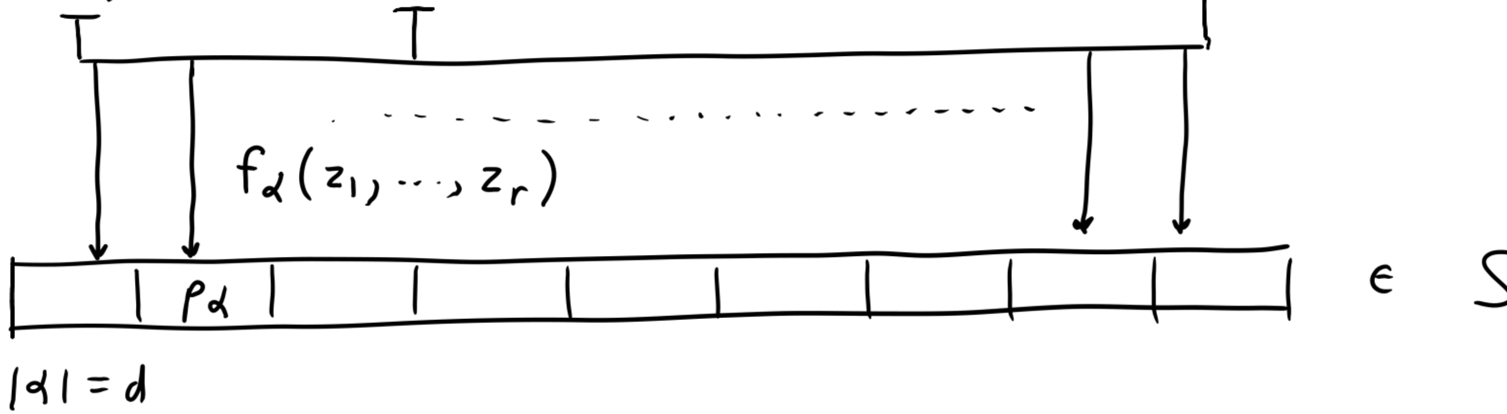
How? (Without becoming algebraic geometers....)

[AGHKT'12] focus on orthogonal z_j, i.e. $r \le n$
i.e. incredibly low rank. Then carry over some linear algebra.

Us: exploit the polynomial connection, but linearize in another way.

[HKL'12] use a part of this.

$$p(x) = \langle z_1, x\rangle^d + \langle z_2, x\rangle^d + \cdots + \langle z_r, x\rangle^d$$



$$F(z_1, \ldots, z_r) : \mathbb{F}^r \to \cancel{S}$$

$$= [f_\alpha(z_1, \ldots, z_r)]_{|\alpha| = d}$$

Want to show $S \subseteq \text{range}(F)$

Polynomials $f_1, \ldots, f_N$ are algebraically dependent if there is a non-zero "annihilating polynomial" A such that

$$A(f_1(z_1, \ldots, z_r), \ldots, f_N(z_1, \ldots, z_r)) = 0 \quad \forall z_1, \ldots, z_r$$

At most r algebraically independent polynomials on r variables.

$S \subseteq \mathrm{range}(F)$

$$\Updownarrow$$

$f_1, \ldots, f_N$ are algebraically independent

$\Rightarrow$ is easy: suppose A annihilates $f_1, \ldots, f_N$:

$$A(\underbrace{f_1(z_1, \ldots, z_r), \ldots, f_N(z_1, \ldots, z_r)}_{F(f_1, \ldots, f_N)}) = 0 \quad \forall z_1, \ldots, z_r$$

$$A(q) = 0 \quad \forall q \in \underset{\in S}{\mathrm{range}(F)}$$

so A is zero (on S)

$f_1, \ldots, f_N$ are algebraically independent

$\Big\Updownarrow$ First-order criterion for algebraic independence

$$J = \left[\frac{\partial f_\alpha}{d z_j} \right]_{\substack{|\alpha| = d \\ j \in [r]}} \text{ is full rank } \big(rk(J) = N \big)$$

$\Big\Updownarrow$ definition

$$\begin{bmatrix} g_1(z) & \cdots & g_N(z) \end{bmatrix}^T J = [0] \Rightarrow \text{all } g_\alpha = 0$$

$\Big\Updownarrow$ whp over v [Shwartz - Zippel]

$$rk\left(J|_{z=v} \right) = rk(J)$$

Suppose f_α are alg. dependent, annihilated by polynomial A of minimum degree: $A(f_1(z), \ldots, f_N(z)) = 0$

$$\left[\partial_{z_j} A(f_1(z), \ldots, f_N(z))\right]_{j \in [r]} = [0]_{j \in [r]}$$

$$\left[\frac{\partial A}{\partial z_j}\right] = \left[\frac{\partial A}{\partial f_\alpha}\right]^T \left[\frac{\partial f_\alpha}{\partial z_j}\right]$$

$$\left[\partial_{f_\alpha} A(f_1(z), \ldots, f_N(z))\right]^T_{|\alpha| = d} \underbrace{\left[\partial_{z_j} f_\alpha(z)\right]_{\substack{|\alpha| = d \\ j \in [r]}}}_{J} = [0]_{j \in [r]}$$

Each of these is zero because J is full rank

A's rate of change wrt. to every input is 0, so A is constant, and annihilating only if it's 0.

So full rank $\Rightarrow$ algebraic independence.

What's next?

Take $\left[\dfrac{\partial f_\alpha}{\partial z_j}\right]_{\substack{|\alpha|=d \\ j\in[r]}}$ (which we know)

and come up with a more computational criterion $\longrightarrow$ "Apolarity lemma" and tensor decomposition algorithms.

Thanks!

Polarization

$$D_c\left(\phi(z)(x)\right) = \sum_{i=1}^{n} c_i \frac{\partial}{\partial x_i} \langle z, x \rangle^d$$

$$= \sum_{i=1}^{n} c_i \, d \, \langle z, x \rangle^{d-1} x_i$$

$$= d \langle z, x \rangle^{d-1} \langle c, x \rangle$$

$$D_c\left(f(x)\right) = \sum_{i=1}^{n} c_i \frac{\partial f(x)}{\partial x_i}$$