EXERCISES ON SUMS-OF-SQUARES OVER FINITE SETS

CURRENT THEMES OF DISCRETE OPTIMIZATION:
BOOT-CAMP FOR EARLY CAREER RESEARCHERS

MAURICIO VELASCO

1. Exercises for lecture 1: The basic setup

(1) *(The coordinate ring of the hypercube).* Let n be a positive integer and $X := \{-1,1\}^n \subseteq \mathbb{R}^n$. Prove that $\mathbb{R}[X] = \mathbb{R}[x_1,\dots,x_n]/\left(x_i^2-1 : i=1,\dots,n\right)$ as follows:
  (a) Let $J(X)$ be the ideal generated by x_i^2-1 for $i=1,\dots,n$. Prove that $J(X) \subseteq I(X)$.
  (b) Prove that $\mathbb{R}[x_1,\dots,x_n]/J(X)$ is generated by the 2^n square-free monomials in $x_1,\dots,x_n$.
  (c) Conclude that $\mathbb{R}[X] = \mathbb{R}[x_1,\dots,x_n]/J(X)$.
  (d) Compute the Hilbert-samuel function $h_X(r)$ for $r = 0,1,2,\dots,n$ and verify that $\operatorname{reg}(X) = 1+n$.
  (e) Compute the coordinate ring of the hypercube $\{0,1\}^n \subseteq \mathbb{R}^n$.

(2) *(Every function on a finite X is a polynomial)* Let $X \subseteq \mathbb{R}^n$ be a finite set. Prove that every real-valued function on X is the restriction of some polynomial via the following steps:
  (a) Assume $\ell(x_1,\dots,x_n)$ is a linear form which distingusishes the points of X in the sense that $\ell(\alpha) = \ell(\beta)$ for $\alpha,\beta \in X$ implies $\alpha = \beta$. Prove that the polynomials

$$p_y(x_1,\dots,x_n) = \frac{\prod_{z\in X\setminus y}\left(\ell(x_1,\dots,x_n)-\ell(z)\right)}{\prod_{z\in X\setminus y}\left(\ell(y)-\ell(z)\right)}$$

  for $y \in X$ are Lagrange interpolation functions on X
  (b) Show that for every finite set X there exists a linear form which distinguishes the points of X.
  (c) Conclude that every real-valued function on X is the restriction of a polynomial and that the vector space dimension of $\mathbb{R}[X]$ equals $|X|$.

(3) *(Formulating $f_{(1)}$ as an SDP).* Let $X := \{-1,1\}^3 \subseteq \mathbb{R}^3$. In this exercise we want to minimize

$$f := 6 - 2(xy+yz+xz)$$

via semidefinite programming.
  (a) Write the optimization problem

$$f_{(1)} := \max\{\lambda : f-\lambda \in \Sigma_{\leq 1}\}$$

  in $\mathbb{R}[X]$ as an explicit semidefinite program.
  (b) Show that $f_{(1)} := \max\,[6-(A+E+H+J)]$ subject to

$$\begin{pmatrix} A & 0 & 0 & 0 \\ 0 & E & -1 & -1 \\ 0 & -1 & H & -1 \\ 0 & -1 & -1 & J \end{pmatrix} \succeq 0$$

  (c) Prove that $f_{(1)} = 0 = \min_{x\in X} f$.

2. Exercises for lecture 2: Sparsity and chordality

(1) *(Cliques of chordal graphs)* Prove the correctness of the algorithmic procedure explained Lecture 2 for finding all maximal cliques of a chordal graph given a perfect elimination order $v_1 < \cdots < v_n$ on its vertices.

(2) Use the Grone and Johnson Theorem to decide whether the following matrix $A \succeq 0$.

$$A = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{pmatrix} 2.2 & 1.1 & 0 & 1.1 \\ 1.1 & 2.2 & 1.1 & 0 \\ 0 & 1.1 & 2.2 & 1.1 \\ 1.1 & 0 & 1.1 & 2.2 \end{pmatrix} \end{array}$$

(3) *(The square-free monomial basis)* Let n be a positive integer and let $X := \{-1, 1\}^n$ be the hypercube. Prove the following statements:

(a) The square-free monomials $\{\chi_I : I \subseteq [n]\}$ form an orthonormal basis of $\mathbb{R}[X]$ for the inner product

$$\langle f, g \rangle := \frac{1}{|X|} \sum_{y \in X} f(x)g(x).$$

(b) Prove that the equality $\chi_I \cdot \chi_J = \chi_{I \triangle J}$ holds in $\mathbb{R}[X]$.

(c) For each $z \in X$ let p_z be the Lagrange interpolation function centered at z. Use orthonormality of χ_I to find a formula expressing p_z in terms of the basis $\vec{\chi}$.

(d) Conclude that any nonnegative function $f \in \mathbb{R}[X]$ admits a representation $f = \vec{\chi}^t Q \vec{\chi}$ where $Q \succeq 0$ and the entries of Q are given by

$$Q_{\chi_I, \chi_J} := \langle f, \chi_{I \triangle J} \rangle.$$

(4) *(A general upper bound for the exactness degree)* Let n be an even integer and let $X := \{-1, 1\}^n$. Suppose d is even and $f \in \mathbb{R}[X]_{\leq d}$ is a nonnegative form. Prove that $f \in \Sigma_{\leq b}$ with $b = \frac{n+d}{2}$ via the following steps:

(a) Define the graph H whose vertices are all subsets of $[n]$ and connect two sets I, J with an edge if and only if the symmetric difference $I \triangle J$ has size at most d. Prove that $f := \vec{\chi}^t Q \vec{\chi}$ for some matrix $Q \succeq 0$ for which H is an admissible support graph.

(b) Define a graph $\hat{H}$ whose vertices are all subsets of $[n]$ and connect two sets I, J with an edge if and only if $||I| - |J|| \leq d$. Prove that $\hat{H}$ is a chordal cover of H with cliques $C_0, \dots, C_{n-d}$ where the vertices of C_i are all sets of sizes $i, i+1, \dots, i+d$.

(c) Note that multiplication $\chi_{[n]}$ exchanges the C_j's. Prove a Lemma explaining how.

(d) Prove that $f \in \Sigma_{\leq b}$ with $b = \frac{n+d}{2}$.

(e) What, if anything, changes when d or n are odd? Find the right statements and prove them.

3. Exercises for lecture 3: Symmetry in sums-of-squares.

(1) Let S_n act on $\mathbb{R}[\vec{x}] := \mathbb{R}[x_1, \dots, x_n]$ by permuting the coordinates.

(a) Prove that the set of homogeneous polynomials of any given degree d, denoted $\mathbb{R}[\vec{x}]_d$ is a stable subspace.

(b) Prove that the subspace $\langle \sum x_i \rangle \subseteq \mathbb{R}[\vec{x}]_1$ is a stable subspace.

(c) Prove that the orthogonal complement of $\sum x_i$ (with respect to the inner product $\langle x_i, x_j \rangle = \delta_{ij}$) is an *irreducible* stable subspace of $\mathbb{R}[\vec{x}]_1$

(d) Describe all the S_n-invariant sums-of-squares of linear forms in $\mathbb{R}[\vec{x}]$. Is every such form a sum of squares of invariant forms?

(2) *(A very useful isomorphism)* Let (V, ρ_1),(W, ρ_2) be representations of a finite group G. Prove the following statements:

(a) The pair (V^*, ρ_1^t) is a representation of G (here ρ_1^t denotes the transpose of ρ).

(b) The pair $(V \otimes W, \rho_1 \otimes \rho_2)$ is a representation of G.

(c) Prove that the vector space $\operatorname{Hom}(V, W)$ can be made into a representation of T by definining $\tau(g)(T) := \rho_2(g) \circ T \circ \rho_1(g^{-1})$.

(d) Prove that there is a canonical isomorphism *of representations*

$$V \otimes W \cong \operatorname{Hom}(V^*, W).$$

and in particular $(V \otimes W)^G$ is isomorphic $\operatorname{Hom}_G(V^*, W)$ defined as the space of morphisms of representations between V^* and W.

(e) Assume V, W are isotypical components of a representation Λ with summands isomorphic to irreducible representations T, S with $T^* \not\cong S$. Compute $(V \otimes W)^G$.

(3) *(Invariant sums-of-squares)* Let X be a finite set and let $G \subseteq \mathrm{Aut}(X)$ be a subgroup and suppose G acts on $\mathbb{R}[X]$ by the contragradient action ρ^*. Prove the following statements:

(a) If $W \subseteq \mathbb{R}[X]$ is a G-stable subspace then the multiplication map $\mu : W \otimes W \to \mathbb{R}[X]$ is a morphism of representations.

(b) Prove that the averaging maps $\mathcal{A}' : W \otimes W \to (W \otimes W)^G$ and $\mathcal{A} : \mathbb{R}[X] \to \mathbb{R}[X]^G$ satisfy the commuting relation $\mathcal{A}' \circ \mu = \mu \circ \mathcal{A}$.

(c) Assume now that $\mathbb{R}[X]_{\leq r}$ is stable and that $\mathbb{R}[X]_{\leq r} := W_1 \oplus \cdots \oplus W_c$ is the G-isotypical decomposition. Assume this decomposition is real and furthermore that $W_j^* \cong W_j$ for every j (this always happens when the isotypical decomposition is real). Prove that:

(i) If $i \neq j$ then for every $w_i \in W_i$ and $w_j \in W_j$ the equality $\mathcal{A}(w_i w_j) = 0$ holds.

(ii) If $i = j$ then for every $w, w' \in W_i$ the average $\mathcal{A}(ww')$ lies in a subspace of dimension bounded above by the multiplicity of the isotypical component W_i.

(iii) Give a rigorous proof of the fact that

$$\Sigma_{\leq r}^G = \mathcal{A}(\Sigma_{W_1}) + \cdots + \mathcal{A}(\Sigma_{W_c})$$

4. Exercises for lecture 4: The polynomial kernel method.

(1) *(The reproducing or Christoffel-Darboux kernel)* Let X be a finite set. Endow $\mathbb{R}[X]$ with the inner product

$$\langle f, g \rangle := \frac{1}{|X|} \sum_{y \in X} f(y)g(y)$$

and let $W \subseteq \mathbb{R}[X]$ be any vector subspace. Prove the following statements

(a) If $f_1, \dots, f_L$ is a $\langle , \rangle$-orthonormal basis for W then the function

$$K(a,b) := \sum_{j=1}^{L} f_j(a) f_j(b)$$

called the reproducing or Christoffel-Darboux kernel has the following *reproducing property*

$$\frac{1}{|X|} \sum_{y \in X} K(x,y) f(y) = \begin{cases} f(x), \text{ if } f \in W \\ 0, \text{ if } f \in W^\perp \end{cases}$$

(b) Show that $K : X \times X \to \mathbb{R}$ is the unique function with the reproducing property on W. In particular show that $K(a,b)$ is independent of the chosen orthonormal basis for W.

(c) Suppose $h_1, \dots, h_L$ is a basis for W which is not necessarily orthonormal. Find a formula for computing the Christoffel-Darboux kernel $K(a,b)$ using the h's.

(2) *(The Funk-Hecke formula)* Suppose X is a finite metric space with distance function $d(x,y)$. Let $\mathbb{B}$ be the collection of automorphisms of the metric, that is $g : X \to X$ such that $d(gx, gy) = d(x,y)$. Assume $u(t)$ is a univariate polynomial and let $\Gamma : \mathbb{R}[X] \to \mathbb{R}[X]$ be defined by

$$\Gamma(f)(x) := \frac{1}{|X|} \sum_{y \in X} u(d(x,y)) f(y).$$

Prove the following statements:

(a) The map $\Gamma : \mathbb{R}[X] \to \mathbb{R}[X]$ is a morphism of $\mathbb{B}$-representations.

(b) Assume the isotypical decomposition of $\mathbb{R}[X]$ is of multiplicity one, that is

$$\mathbb{R}[X] := W_1 \oplus W_2 \oplus \cdots \oplus W_L$$

where the W_i are pairwise nonisomoprhic irreducible representations of $\mathbb{B}$. Prove that for each W_i there exists a constant λ_i such that Γ acts like $\lambda_i I$ on W_i.

(3) (*Generalized Krawtchouk polynomials*) We use the notation and assumptions from part (*b*) of the previous exercise. Fix $x_0 \in X$ and let $H := \mathrm{Stab}(x_0)$ that is, the collection of elements $g \in \mathbb{B}$ such that $gx_0 = x_0$. Assume that every $\mathbb{B}$-irreducible component contains a unique copy of the trivial H-representation and furthermore that $\mathbb{R}[X]^H$ is generated as an algebra by the function $\ell(y) := d(x_0, y)$ with $\mathbb{R}[X]^H = \mathbb{R}[\ell]/(q(\ell))$ where $q(\ell)$ is a univariate polynomial of degree $1 + \mathrm{diam}(X)$. Prove that

for each $j = 0, 1, \ldots, \operatorname{diam}(X)$ there is a unique univariate polynomial $u_j(t)$ with the property that $K(x, y) := u_j(d(x, y))$ is the reproducing kernel for W_j.