EE464 Sparse Polynomials

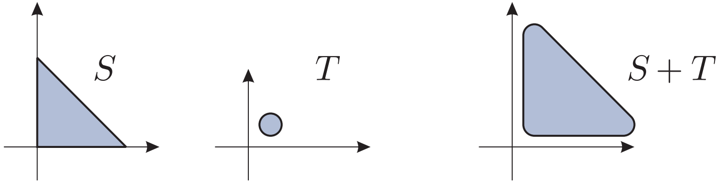
Minkowski sum

for subsets $S, T \subset \mathbb{R}^n$, the *Minkowski sum* is

$$S + T = \left\{\, x + y \mid x \in S,\ y \in T \,\right\}$$

also for $\lambda \in \mathbb{R}$, define

$$\lambda S = \left\{\, \lambda x \mid x \in S \,\right\}$$



convolution

for $S \in \mathbb{R}^N$ define the *indicator function* $I_S : \mathbb{R}^N \to \mathbb{R}$

$$I_S(x) = \begin{cases} 1 & \text{if } x \in S \\ 0 & \text{otherwise} \end{cases}$$

then the Minkowski sum corresponds to convolution

$$I_{S+T} = I_S * I_T$$

that is

$$I_{S+T}(x) = \int_y I_S(x-y) I_T(y)\, dy$$

properties

if S and T are convex, so is $S+T$

to see this, notice that the Cartesian product is convex

$$S \times T = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \;\middle|\; x \in S,\ y \in T \right\}$$

and the sum $S+T$ is image of the $S \times T$ under the linear map

$$\begin{bmatrix} x \\ y \end{bmatrix} \mapsto x + y$$

properties

in general $S + S \neq 2S$, for example

$$S = \{\, 0, 1 \,\} \qquad \text{and} \qquad S + S = \{\, 0, 1, 2 \,\}$$

if S is convex, then

$$(\lambda + \mu)S = \lambda S + \mu S$$

polyhedra

a set $S \subset \mathbb{R}^n$ is called a *polyhedron* if it is the intersection of a finite set of closed halfspaces

$$S = \left\{ x \in \mathbb{R}^n \mid Ax \leq b \right\}$$

- a bounded polyhedron is called a *polytope*
- the *dimension* of a polyhedron is the dimension of its affine hull

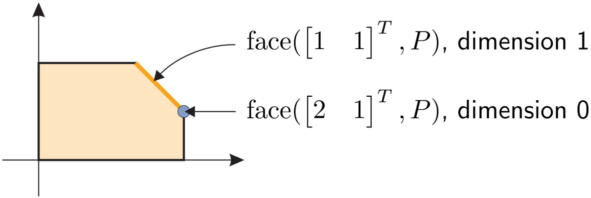
$$\text{affine}(S) = \left\{ \lambda x + \nu y \mid \lambda + \nu = 1,\ x, y \in S \right\}$$

- if $b = 0$ the polyhedron is a cone
- every polyhedron is convex

faces of polyhedra

given $a \in \mathbb{R}^n$, the corresponding *face* of polyhedron P is

$$\operatorname{face}(a, P) = \left\{ x \in P \mid a^T x \geq a^T y \text{ for all } y \in P \right\}$$



- faces of dimension 0 are called *vertices*

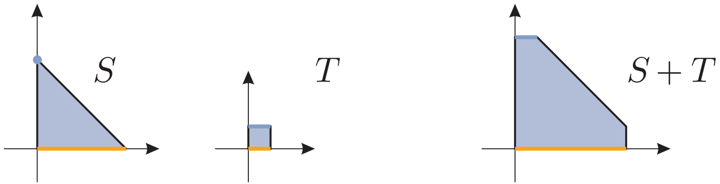
  1 *edges*

  $d-1$ *facets*, where $d = \dim(P)$
- facets are also said to have *codimension* 1

faces of polyhedra

- if F is a face of G, and G is a face of P then F is a face of P
  i.e., *is a face of* is transitive

- $\text{face}(a, S+T) = \text{face}(a, S) + \text{face}(a, T)$



- in particular, if x is a vertex of $S+T$, then

  $x = y + z$ for some y, a vertex of S and z, a vertex of T

  and the vertices y and z are unique

positive polynomials

suppose $f = c_d x^d + c_{d-1} x^{d-1} + \cdots + c_1 x + c_0$; then

$$f \text{ is PSD} \qquad \Longrightarrow \qquad d \text{ is even, } \ c_d > 0 \text{ and } c_0 \geq 0$$

what is the analogue in n variables?

example

- suppose $f = x^3y^2 + xy + 1$

  substitute $x = t$ and $y = t$, i.e., evaluate f along the curve $x = y$,

  $$\hat{f} = t^5 + t^2 + 1$$

  so clearly f is not PSD

  this suggests that f is PSD implies f has even degree

- but for $f = x^3y^2 - xy^4 + x^2y^2 + 1$ the same substitution gives

  $$\hat{f} = t^4 + 1$$

the Newton polytope

suppose

$$f = \sum_{\alpha \in M} c_\alpha x^\alpha$$

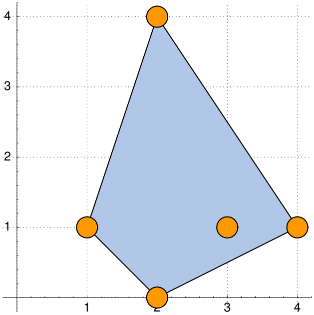
the set of monomials $M \subset \mathbb{N}^n$ is called the *frame* of f

the *Newton polytope* of f is its convex hull

$$\operatorname{new}(f) = \operatorname{co}\big(\operatorname{frame}(f)\big)$$

the example shows

$$f = 7\,x^4\,y + x^3\,y + x^2\,y^4 + x^2 + 3\,x\,y$$



necessary condition for nonnegativity

we'll evaluate the polynomial f along the curve

$$x_1 = z_1 t^{a_1}$$
$$\vdots$$
$$x_n = z_n t^{a_n}$$

for $f = \sum_{\alpha \in M} c_\alpha x^\alpha$ define

$$\hat{f} = \sum_{\alpha \in M} c_\alpha z^\alpha t^{a^T \alpha}$$

e.g., for $f = x^3 y + 2xy^7$ we have

$$\hat{f} = z_1^3 \, z_2 \, t^{3a_1 + a_2} \; + \; 2 \, z_1 \, z_2^7 \, t^{a_1 + 7a_2}$$

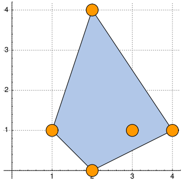
necessary condition for nonnegativity

if $f \in \mathbb{R}[x_1, \ldots, x_n]$ is PSD, then

> every vertex of $\mathrm{new}(f)$ has even coordinates, and a positive coefficient

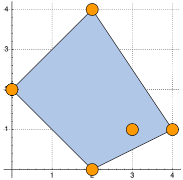
- $f = 7\,x^4\,y + x^3\,y + x^2\,y^4 + x^2 + 3\,x\,y$

  is not PSD, since term $3xy$ has coords $(1, 1)$



- $f = 7\,x^4\,y + x^3\,y - x^2\,y^4 + x^2 + 3\,y^2$

  is not PSD, since term $-x^2\,y^4$ has a negative coefficient



proof

if β is a vertex of $\text{new}(f)$, then there is some $a \in \mathbb{R}^n$ such that

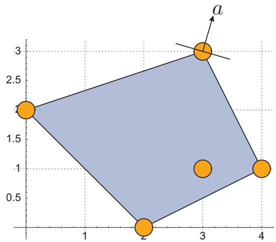
$$a^T\beta > a^T\alpha \text{ for all } \alpha \in M$$

evaluating $\hat{f}$ along the curve $x_i = z_i t^{a_i}$, gives

$$\hat{f} = c_\beta z^\beta t^{a^T\beta} + \text{ terms of lower degree in } t$$

as $t \to \infty$, the first terms dominates, so

$$c_\beta z^\beta \geq 0 \text{ for all } z \in \mathbb{R}^n$$



assume f is PSD, then

- picking $z = 1$ implies c_β must be positive
- picking $z_j = -1$ and $z_i = 1$ for $i \neq j$ implies β_i must be even

halfspaces containing the Newton polytope

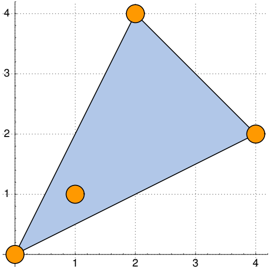
the Newton polytope of f is contained with the halfspace specified by a, b

$$\mathrm{new}(f) \subset \left\{ x \in \mathbb{R}^n \mid a^T x \leq b \right\}$$

if and only if

$$\lim_{t \to \infty} \left| t^{-b} \hat{f} \right| < \infty \qquad \text{for all } z \in \mathbb{R}^n$$

example



suppose $a = \begin{bmatrix}1 & 1\end{bmatrix}^T$ and $b = 6$

$$f = -x^4\, y^2 + x^2\, y^4 + x\, y + 1$$

then

$$\hat{f} = (-z_1^4 z_2^2 + z_1^2\, z_2^4)\, t^6 + z_1\, z_2\, t^2 + 1$$

we have

$$\operatorname{new}(f) \subset \left\{ x \in \mathbb{R}^n \mid a^T x \le b \right\} \quad \Longrightarrow \quad \lim_{t\to\infty} \left| t^{-b} \hat{f} \right| < \infty$$

the converse also holds, since by picking z arbitrarily we can arrange for the leading coefficient of $\hat{f}$ to be non-zero

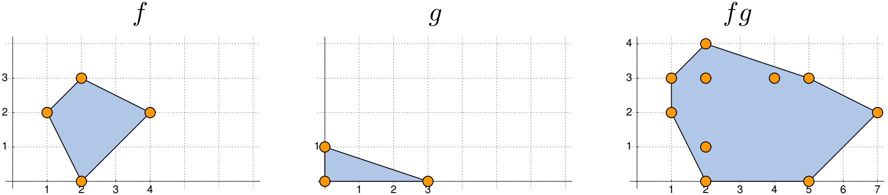
Newton polytopes of a product

$$\mathrm{new}(fg) = \mathrm{new}(f) + \mathrm{new}(g)$$

$$f = x^4\,y^2 + 2\,x^2\,y^3 - x^2 - x\,y^2$$

$$g = x^3 - y + 1$$

$$fg = x^7\,y^2 + 2\,x^5\,y^3 - x^5 - x^4\,y^3 - 2\,x^2\,y^4 + 2\,x^2\,y^3$$



Newton polytopes

we'd like to show $\mathrm{new}(fg) = \mathrm{new}(f) + \mathrm{new}(g)$

first, we'll show

$$\mathrm{new}(fg) \subset \mathrm{new}(f) + \mathrm{new}(g)$$

to see this, if $f = \sum_\alpha c_\alpha c^\alpha$ and $g = \sum_\beta d_\beta x^\beta$ then

$$fg = \sum_\alpha \sum_\beta c_\alpha d_\beta x^{\alpha+\beta}$$

so $\mathrm{frame}(fg) \subset \mathrm{frame}(f) + \mathrm{frame}(g)$

also we have $\mathrm{co}(S+T) \subset \mathrm{co}(S) + \mathrm{co}(T)$

Newton polytopes

it remains to show

$$\mathrm{new}(fg) \supset \mathrm{new}(f) + \mathrm{new}(g)$$

we'll show that if γ is a vertex of $\mathrm{new}(f) + \mathrm{new}(g)$ then $\gamma \in \mathrm{new}(fg)$

we know $\gamma = \alpha + \beta$ for unique $\alpha \in \mathrm{frame}(f)$ and $\beta \in \mathrm{frame}(g)$

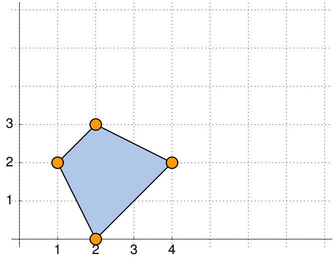
α and β are unique since γ is a vertex

the coefficient of x^γ in fg is $c_\alpha d_\beta$, which cannot be zero, so $\gamma \in \mathrm{new}(fg)$

Newton polytopes of squares

consequently we have

$$\operatorname{new}(f^n) = n \operatorname{new}(f)$$

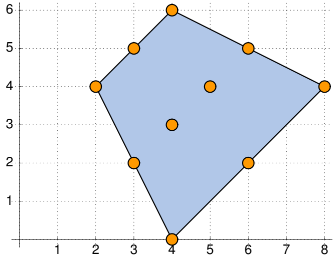


with

$$f = x^4\, y^2 + 2\, x^2\, y^3 - x^2 - x\, y^2$$

we have

$$f^2 = x^8\, y^4 + 4\, x^6\, y^5 - 2\, x^6\, y^2 - 2\, x^5\, y^4 + 4\, x^4\, y^6 - 4\, x^4\, y^3 + x^4 - 4\, x^3\, y^5 + 2\, x^3\, y^2 + x^2\, y^4$$



Newton polytopes and inequalities

if f and g are PSD polynomials then

$$f(x) \leq g(x) \text{ for all } x \in \mathbb{R}^n \qquad \Longrightarrow \qquad \text{new}(f) \subset \text{new}(g)$$

we'll show that any halfspace containing $\text{new}(g)$ also contains $\text{new}(f)$

if $\text{new}(g) \subset \{ x \mid a^T x \leq b \}$ then

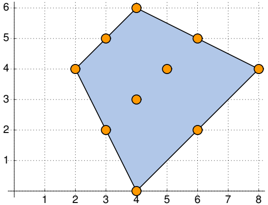
$$\lim_{t \to \infty} t^{-b} \hat{g} < \infty \qquad \text{for all } z$$

since $0 \leq f \leq g$ we therefore have the same holds for $\hat{f}$, and so

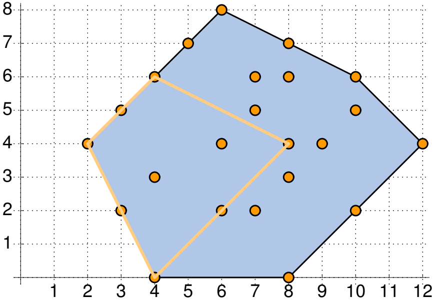
$$\text{new}(f) \subset \{ x \mid a^T x \leq b \}$$

example

$$f = x^4\, y^2 + 2\, x^2\, y^3 - x^2 - x\, y^2$$



$$\mathrm{new}(f^2)$$



$$\mathrm{new}\bigl(f^2(x^2y^2 + x^4 + 1)\bigr)$$

sparse SOS decomposition

this tells us which monomials we have in an SOS decomposition

$$f = \sum_{i=1}^{t} g_i^2 \qquad \Longrightarrow \qquad \operatorname{new}(g_i) \subset \frac{1}{2}\operatorname{new}(f)$$

because $0 \le g_i^2 \le f$ so

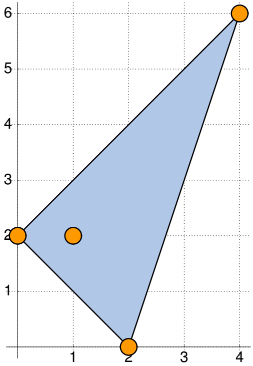
$$\begin{aligned} \operatorname{new}(f) &\supset \operatorname{new}(g_i^2) \\ &= 2\operatorname{new}(g_i) \end{aligned}$$

this holds for *every* SOS decomposition of f

example: sparse SOS decomposition

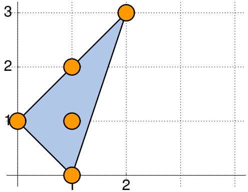
find an SOS representation for

$$f = 4\,x^4\,y^6 + x^2 - x\,y^2 + y^2$$



the squares in an SOS decomposition can only contain the monomials

$$\mathrm{new}(\frac{1}{2}f) \cap \mathbb{N}^n = \{x^2y^3, xy^2, xy, x, y\}$$



without using sparsity, we would include all 21 monomials of degree < 5 in the SDP

with sparsity, we only need 5 monomials

example continued

we find

$$f = 4\,x^4\,y^6 + x^2 - x\,y^2 + y^2$$

$$f = \begin{bmatrix} y \\ x \\ xy \\ xy^2 \\ x^2y^3 \end{bmatrix}^T \begin{bmatrix} 1 & 0 & -0.5 & 0 & -0.5 \\ 0 & 1 & 0 & -0.5 & 0 \\ -0.5 & 0 & 1 & 0 & 0 \\ 0 & -0.5 & 0 & 1 & 0 \\ -0.5 & 0 & 0 & 0 & 4 \end{bmatrix} \begin{bmatrix} y \\ x \\ xy \\ xy^2 \\ x^2y^3 \end{bmatrix}$$

and the matrix is PSD

homogeneous polynomials

polynomial f is called homogeneous if

$$f = \sum_{\alpha \in M} c_\alpha x^\alpha \qquad \text{with} \qquad \sum_{i=1}^{n} \alpha_i = d \text{ for all } \alpha \in M$$

if f is homogeneous, then for an SOS decomposition we need only look at monomials x^β such that

$$\sum_{i=1}^{n} \beta_i = \frac{d}{2}$$

for example

$$f = 4x^4 + 4x^3y - 7x^2y^2 - 2xy^3 + 10y^4$$

$$= \begin{bmatrix} x^2 \\ xy \\ y^2 \end{bmatrix}^T \begin{bmatrix} 4 & 2 & -\lambda \\ 2 & -7+2\lambda & -1 \\ -\lambda & -1 & 10 \end{bmatrix} \begin{bmatrix} x^2 \\ xy \\ y^2 \end{bmatrix}$$