LECTURE 2

Convex sets and cones

A set $C \subset \mathbb{R}^n$ is <u>convex</u> whenever

$$\forall x, y \in C \quad \forall \lambda \in [0,1] \Rightarrow \lambda x + (1-\lambda) y \in C$$

Example: lines, linear spaces, disk, polyhedra...

We endow $\mathbb{R}^n$ with the Euclidean metric

$$\|x\| = \sqrt{\langle x, x \rangle} \quad \text{and} \quad d(x,y) = \|x - y\|$$

where $\langle x, y \rangle = \sum x_i y_i$ is the Euclidean (or standard) inner product.

Over the set of matrices $\mathbb{R}^{m \times m}$ we define

$$\mathbb{R}^{m \times m} \times \mathbb{R}^{m \times m} \longrightarrow \mathbb{R}$$

$$(M, N) \mapsto \mathrm{Trace}(MN^T)$$

and we denote it $\langle M, N \rangle := \mathrm{Trace}(MN^T)$ because it is an inner product (exercise).

Over $\mathrm{Sym}_m(\mathbb{R}) \subset \mathbb{R}^{m \times m}$ it reads

$$\langle M, N \rangle = \mathrm{Trace}(MN)$$

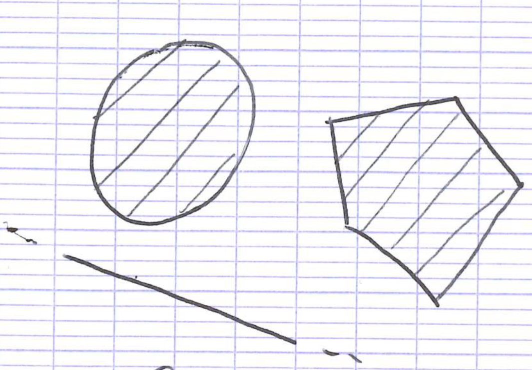
(because $N = N^T$)

Inner Product: $\langle \, , \rangle : V \times V \to \mathbb{R}$

- $\langle x, y \rangle = \langle y, x \rangle$ (symmetry)
- $\langle \alpha x_1 + \beta x_2, y \rangle = \alpha \langle x_1, y \rangle + \beta \langle x_2, y \rangle$ (linearity)
- $\langle x, x \rangle \geq 0$, $\langle x, x \rangle = 0$ iff $x = 0$. (positive-~~[illegible]~~ definiteness)

A set $C \subset \mathbb{R}^n$ is a convex cone if

$$\forall \alpha,\beta \in \mathbb{R}_{\geq 0} \ \forall x,y \in C \Rightarrow \alpha x + \beta y \in K$$



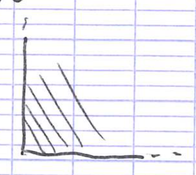
Convex Sets

Convex Cone

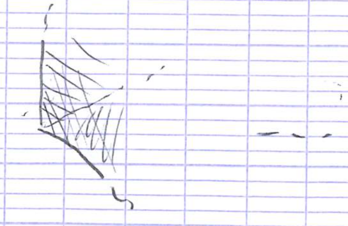
Examples of cones

(1) The positive orthant:

$$\mathbb{R}^n_{\geq 0} = \{ x \in \mathbb{R}^n : \ x_i \geq 0 \ \forall \ i = 1 \ldots n \}$$



$n = 2$



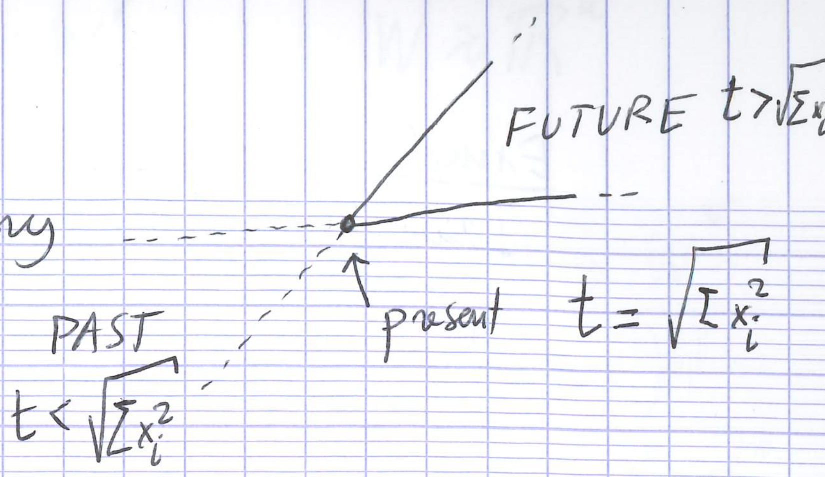
$n = 3$

(2) The Lorentz (or second-order) cone

$$\mathcal{L}^n = \left\{ \cancel{x \in \mathbb{R}^n} \ (x,t) \in \mathbb{R}^{n+1} : \ \sqrt{\sum_{i=1}^{n} x_i^2} \leq t \right\}$$

it defines a metric in the Einstein

space-time theory



(3) The cone of positive semidefinite matrices (or PSD cone):

$$Sym_m^+(\mathbb{R}) = \{M \in Sym_m(\mathbb{R}) : M \succeq 0\}$$

all eigenvalues are nonnegative

Conic and Semidefinite Programming

Let W be a f.d. real vector space. The dual

$$W^\vee = \{\ell : W \to \mathbb{R} : \ell \text{ linear}\}$$ is still a vector space.

Two There is an isomorphism $W \overset{\varphi}{\cong} W^\vee$.
(but we consider these spaces distinct because φ is non-canonical).

If $C \subset W$ is a convex cone, its dual cone is

$$C^* = \{\ell \in W^\vee : \ell(x) \geq 0 \ \forall x \in C\}$$

Examples

Identifying $(\mathbb{R}^n)^\vee \cong \mathbb{R}^n$ one has that $\mathbb{R}^n_{\geq 0}$ is self-dual that is $\mathbb{R}^n_{\geq 0} = \mathbb{R}^{n\,*}_{\geq}$.

Identifying $(\mathrm{Sym}_m(\mathbb{R}))^\vee \cong \mathrm{Sym}_m(\mathbb{R})$ one has

$$\mathrm{Sym}^+_m(\mathbb{R}) = \mathrm{Sym}^+_m(\mathbb{R})^*$$

This means that

$$M \succcurlyeq 0 \iff \underbrace{\langle M, N \rangle}_{=\,\mathrm{Trace}(MN)} \geq 0 \quad \forall\, N \succcurlyeq 0.$$

The Lorentz cone is self-dual, too.

A (primal) conic program is an opt. prob. of type

$$(P) \qquad p^* = \inf \langle c, x \rangle \quad \text{s.t.} \quad Ax = b, \quad x \in K$$

when K convex cone; and its dual conic program

$$(D) \qquad d^* = \sup \langle b, y \rangle \quad A^* y + s = c, \quad s \in K^*$$

$K \subset W$, $K^* \subset W^\vee$ $\qquad W \simeq \mathbb{R}^n$

$A : W \to \mathbb{R}^m$ linear function

A^* adjoint of A, $A^* : (\mathbb{R}^m)^\vee \longrightarrow W^\vee$

Strong duality: (under some conditions)

one could have $p^* = d^*$

Weak duality: it <u>always</u> holds $\quad p^* \geq d^*$

<u>Proof</u>

If one of the feasible sets is empty, this is trivial because one would have

$p^* = +\infty$ or $d^* = -\infty$.

If x, s are a primal-dual feasible couple:

$$\langle c, x \rangle - \langle b, y \rangle = \langle c, x \rangle - \langle Ax, y \rangle =$$

$$= \langle c, x \rangle - \langle x, A^* y \rangle = \langle x, s \rangle \geq 0$$

property of adjoint

$x \in K$
$s \in K^*$

Hence

$$\langle c, x \rangle \geq \langle b, y \rangle$$

$\forall\, x, y$ feasible $\to$ go to Inf & Sup $\square$

When both primal and dual feasible sets have an interior point, then strong duality holds (there are weaker criteria, too).

Semidefinite Programming

Fix $K = Sym_n^+(\mathbb{R})$, then the corresponding conic program is called Semidefinite programming ~~problem~~ (SDP).

We have $\quad W = Sym_n(\mathbb{R})$

$$A : Sym_n(\mathbb{R}) \longrightarrow \mathbb{R}^m$$
$$X \mapsto \left(\langle A_1, X \rangle, \ldots, \underset{\substack{\| \\ \mathrm{Trace}(A_i, X)}}{\langle A_m, X \rangle} \right)$$

In the case of SDP
the adjoint takes the form

$$A^* : (\mathbb{R}^m)^\vee \longrightarrow \left(Sym_n(\mathbb{R}) \right)^\vee$$
$$y \mapsto y_1 A_1 + y_2 A_2 + \cdots + y_m A_m$$

Hence we get

$$(P_{SDP}) \quad p^* = \operatorname{Inf} \langle C, X \rangle$$
$$\text{s.t.} \quad \langle A_i, X \rangle = b_i, \; i = 1..n$$
$$X \succcurlyeq 0 \quad (X \in K)$$

$$(D_{SDP}) \quad d^* = \operatorname{Sup} \langle b, y \rangle$$
$$\sum y_i A_i + S = C$$
$$S \succcurlyeq 0 \quad (S \in K^*)$$

Semidefinite Programming (SDP) can be "solved" efficiently in practice :

- good complexity (polynomial-time at fixed accuracy : Poly$(m, n, \log(1/\varepsilon))$),
- there exist solvers implemented f.e. in MATLAB (SeDuMi, ...) or in C++ (SDPA, ...) that have good performances

4h30

SOS Certificates

An expression of type

$$\boxed{f = f_1^2 + f_2^2 + \dots + f_r^2}$$

is called a SOS (positivity) certificate.

Q.: How can we compute such a certificate?

RECALL Method for univariate polynomials:

$$f = \prod_{\alpha \in R_{\mathbb{C}}} (x-\alpha)(x-\bar{\alpha}) \prod_{\beta \in R_{\mathbb{R}}} (x-\beta)^2 =$$

$$= h\bar{h} \cdot g^2 = (g \operatorname{Re}(h))^2 + (g \operatorname{Im}(h))^2.$$

But: for multivariate polynomials, we cannot ~~an~~ factor into linear forms.

Recap: for $d \in \mathbb{N}$

$$\Sigma_{n,2d} \subset \mathbb{R}[x]_{2d} \qquad x = (x_1 \cdots x_n)$$

Hilbert's 1888 theorem $\rightarrow \subsetneq \mathcal{P}_{2d}(\mathbb{R}^n) \subset$

$$A, B \in \mathrm{Sym}_m(\mathbb{R}) \rightarrow \langle A, B \rangle := \mathrm{Trace}(AB)$$

Inner product over $\mathrm{Sym}_m(\mathbb{R})$

For $\alpha \in \mathbb{N}^n \rightarrow x^{\alpha} := x_1^{\alpha_1} \cdots x_n^{\alpha_n}$

$$|\alpha| = \alpha_1 + \cdots + \alpha_n = \deg(x^{\alpha})$$

$$\deg\left(\sum_{\alpha \in I} c_{\alpha} x^{\alpha}\right) = \max_{\substack{c_\alpha \neq 0 \\ \alpha \in I}} |\alpha|$$

We define $\mathbb{N}^n_d = \{\alpha \in \mathbb{N}^n : |\alpha| \leq d\}$

$$|\mathbb{N}^n_d| = \dim \mathbb{R}[x]_d = \binom{n+d}{d} = \frac{(n+d)!}{n!\, d!}$$

<u>Definition</u> We fix the monomial basis of $\mathbb{R}[x]_d$:

$$v_d(x) = (1, x_1, x_2 \cdots x_n, x_1^2, x_1 x_2, \cdots x_n^2 \cdots x_n^d)^T$$

Then $\mathbb{R}[x]_d = \langle v_d(x) \rangle_{\mathbb{R}}$

<u>Example</u> $n=2$, $d=2$

$$v_2\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right) = \begin{pmatrix} 1 \\ x_1 \\ x_2 \\ x_1^2 \\ x_1 x_2 \\ x_2^2 \end{pmatrix} \qquad \dim \mathbb{R}[x_1, x_2]_2 = 6 \text{ over } \mathbb{R}.$$

<u>Proposition</u> (Gram representation)

A polynomial $g \in \mathbb{R}[x]_{2d}$ is a sum of squares (that is, $g \in \Sigma_{n,2d}$) if and only if there exists $Q \in \mathrm{Sym}^+_{\binom{n+d}{d}}(\mathbb{R})$ such that

$$\forall x \in \mathbb{R}^n: \; g(x) = v_d(x)^T Q \, v_d(x).$$

<u>Example</u> (quadratic forms)

$g \in \mathbb{R}[x]_2 \quad (d=1) \to \binom{n+1}{1} = n+1$

$$\hookrightarrow Q =: \begin{pmatrix} 1 & q^T \\ q & \tilde{Q} \end{pmatrix}$$

$$v_1(x) = \begin{pmatrix} 1 \\ x_1 \\ \vdots \\ x_n \end{pmatrix} \qquad g = (1 \; x_1 \dots x_n) \, Q \begin{pmatrix} 1 \\ x_1 \\ \vdots \\ x_n \end{pmatrix} =$$

$$= q_0 + q^T x + x^T \tilde{Q} x$$

$\to$ In degree 2 the matrix Q is unique. If $Q \succcurlyeq 0$ then $g \in P_2(\mathbb{R}^n)$

$\Longrightarrow$ hence g is a sum of squares!

__Proof__ (of Gram representation)

Suppose there is $Q \in Sym^{+}_{\binom{n+d}{d}}(\mathbb{R})$ st

$$g = v_d(x)^T Q \, v_d(x) \overset{\text{then}}{=} ?$$

$Q = HH^T$ for some $H \in \mathbb{R}^{\binom{n+d}{d} \times r}$

where $r = \operatorname{rank} Q$. (proof by exercise).

Hence

$$g = v_d(x)^T H H^T v_d(x) = \left(H^T v_d(x)\right)^T H^T v_d(x) =$$

$$= \| H^T v_d(x) \|^2$$

$\lceil$ $w^T w = \|w\|^2$ Euclidean norm $\rfloor$

$$H^T v_d(x) = \begin{pmatrix} g_1 \\ \vdots \\ g_r \end{pmatrix} \to g = g_1^2 + \dots g_r^2$$

$g_i \in \mathbb{R}[x]_d$ $\qquad r = \operatorname{rank} Q$.

Conversely, supp. $g \in \Sigma_{n,2d}$

$$g = g_1^2 + \dots + g_r^2 \to \deg g_i \le d \quad \text{(prove it)}$$

$$\to g_i = h_i^T v_d(x) \qquad h_i \in \mathbb{R}^{\binom{n+d}{d}}$$

$\lceil$ example

$g_i = 1 - 3x_1 + 2x_1^2 + 5x_2$

$\to h_i = (1, -3, 5, 2, 0, 0)$

$v_2(x) = (1, x_1, x_2, x_1^2, x_1 x_2, x_2^2)$ $\rfloor$

$$\to \quad g = \sum_{i=1}^{r} (h_i^T v_d(x))^2 = \sum_{i=1}^{r} (h_i^T \overset{\overset{\mathbb{R}[x]}{\downarrow}}{v_d(x)})(h_i^T v_d(x)) =$$

$$= \sum_{i=1}^{r} v_d^T(x) \underbrace{h_i h_i^T}_{\substack{\downarrow \\ \text{rank 1 psd matrix}}} v_d(x) = v_d^T(x) \underbrace{\left(\sum h_i h_i^T\right)}_{\substack{Q \succeq 0 \\ \text{rank } r \\ \text{psd matrix}}} v_d(x)$$

$\square$

<u>Exercise</u> (Cholesky decomposition)

$\underset{\in \mathrm{Sym}_m^+(\mathbb{R})}{Q \succeq 0}$ of rank r iff $\exists\, H \in \mathbb{R}^{m \times r}$ s.t. $Q = HH^T$.

Spectral theorem (of linear algebra):

$$Q = \sum_{i=1}^{m} \lambda_i u_i u_i^T \quad \text{for } Q \text{ real symmetric}$$

If $Q \succeq 0$ of rank r: $\quad Q = \sum_{i=1}^{r} \lambda_i u_i u_i^T$ and $\lambda_i > 0$

then $\quad Q = UDU^T \qquad D = \begin{pmatrix} \lambda_1 & & & & \\ & \ddots & & & \\ & & \lambda_r & & \\ & & & 0 & \\ & & & & \ddots \\ & & & & & 0 \end{pmatrix}$

Let

$$E = \begin{pmatrix} \sqrt{\lambda_1} & & & & \\ & \ddots & & & \\ & & \sqrt{\lambda_r} & & \\ & & & 0 & \\ & & & & \ddots \\ & & & & & 0 \end{pmatrix} \to E^2 = D \qquad E = E^T$$

$$\to \quad Q = UE^2U^T = UEEU^T = UE^TEU^T =$$

$$= (\underbrace{EU^T}_{H^T})^T EU^T = HH^T$$

$\blacksquare$

Notation

$$v_d(x)\, v_d(x)^T = \binom{n+d}{d} \times \binom{n+d}{d} \text{ matrix} = \sum_{\alpha \in \mathbb{N}^n_{2d}} B_\alpha x^\alpha$$

with polynomial entries of degree $2d$

Example

$$v_1(x) = \begin{pmatrix} 1 \\ x_1 \\ x_2 \end{pmatrix} \qquad n=2,\ d=1 \qquad \binom{n+d}{d} = \binom{2+1}{1} = 3$$

$$\begin{pmatrix} 1 \\ x_1 \\ x_2 \end{pmatrix} (1\ x_1\ x_2) = \begin{pmatrix} 1 & x_1 & x_2 \\ x_1 & x_1^2 & x_1x_2 \\ x_2 & x_1x_2 & x_2^2 \end{pmatrix} = \sum_{\alpha \in \mathbb{N}^2_2} B_\alpha x^\alpha =$$

$$= \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}}_{= B_{\binom{0}{0}}} + \underbrace{\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}}_{= B_{\binom{1}{0}}} x_1 + \underbrace{\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}}_{= B_{\binom{0}{1}}} x_2 +$$

$$+ \underbrace{\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}}_{= B_{\binom{2}{0}}} x_1^2 + \underbrace{\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}}_{= B_{\binom{1}{1}}} x_1x_2 + \underbrace{\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{= B_{\binom{0}{2}}} x_2^2$$

So B_α are indexed in $\mathbb{N}^n_{2d}$ and with entries in $\{0,1\}$.

<u>Proposition</u> Checking whether $g \in \Sigma_{n,2d}$ (and computing the decomposition $g = g_1^2 + \dots + g_z^2$) is equivalent to solving one semidefinite (feasibility) programming problem.

<u>Proof</u>

$$g = \sum_{\alpha \in \mathbb{N}^n_{2d}} g_\alpha x^\alpha = v_d(x)^T Q\, v_d(x) \quad \text{with } Q \succeq 0$$

iff $\langle Q, B_\alpha \rangle = g_\alpha$. Indeed:

$$g = v_d(x)^T Q\, v_d(x) = \operatorname{Trace}\left(v_d(x)^T Q\, v_d(x)\right) =$$

$$= \operatorname{Trace}\left(Q\, v_d(x)\, v_d(x)^T\right) =$$

$$= \operatorname{Trace}\left(Q \cdot \sum_{\alpha \in \mathbb{N}^n_{2d}} B_\alpha x^\alpha\right) =$$

$$= \operatorname{Trace}\left(\sum_{\alpha \in \mathbb{N}^n_{2d}} Q B_\alpha x^\alpha\right) =$$

$$= \sum_{\alpha \in \mathbb{N}^n_{2d}} \operatorname{Trace}\left(Q B_\alpha \overset{\text{scalar}}{x^\alpha}\right) =$$

$$= \sum_{\alpha \in \mathbb{N}^n_{2d}} \operatorname{Trace}(Q B_\alpha) \cdot x^\alpha = \sum_{\alpha \in \mathbb{N}^n_{2d}} \langle Q, B_\alpha \rangle x^\alpha$$

Hence

$$g = \sum_\alpha g_\alpha x^\alpha = \sum_\alpha \langle B, Q_\alpha \rangle x^\alpha$$

by the Identity Principle of Polynomials

$$g_\alpha = \langle Q, B_\alpha \rangle .$$

So the sos-decomposition can be checked via the SDP (feasibility) problem

$$\boxed{\begin{array}{lll} \text{Find} & Q \in \mathrm{Sym}_{\binom{n+d}{d}}(\mathbb{R}) & \\ \text{s.t.} & Q \succeq 0 & \left(Q \in K = \mathrm{Sym}^+_{\binom{n+d}{d}}(\mathbb{R}) \right) \\ & \langle Q, B_\alpha \rangle = g_\alpha . & \end{array}}$$

$\square$

<u>Example</u> (Lasserre's book Ex. 2.2)

$n = 2$, $d = 4$ homogeneous

$$f(x) = 2x_1^4 + 2x_1^3x_2 - x_1^2x_2^2 + 5x_2^4$$

$$f = \begin{pmatrix} x_1^2 & x_2^2 & x_1x_2 \end{pmatrix} \begin{pmatrix} q_{11} & q_{12} & q_{13} \\ q_{12} & q_{22} & q_{23} \\ q_{13} & q_{23} & q_{33} \end{pmatrix} \begin{pmatrix} x_1^2 \\ x_2^2 \\ x_1x_2 \end{pmatrix} =$$

$$= q_{11}x_1^4 + (q_{33} + 2q_{12})x_1^2x_2^2 + q_{22}x_2^4 + 2q_{13}x_1^3x_2 +$$

$+ 2q_{23}x_1x_2^3$. Equeting coeffcients

$q_{11}=2$, $q_{22}=5$, $q_{33}+2q_{12}=-1$, $2q_{13}=2$, $q_{23}=0$

$$Q=\begin{pmatrix} 2 & q_{12} & 1 \\ q_{12} & 5 & 0 \\ 1 & 0 & -1-2q_{12} \end{pmatrix} \Longleftarrow$$

Linear matrix and $Q(q_{12}) \succeq 0$ is called LMI (Linear Matrix Inequality, or feasible problem of SDP).

we find that for ~~some~~ $q_{12} = -3$, $Q(-3) \succeq 0$

$$Q(-3)=\begin{pmatrix} 2 & -3 & 1 \\ -3 & 5 & 0 \\ 1 & 0 & 5 \end{pmatrix} = \frac{1}{2}\underbrace{\begin{pmatrix} 2 & 0 \\ -3 & 1 \\ 1 & 3 \end{pmatrix}}_{=H}\underbrace{\begin{pmatrix} 2 & -3 & 1 \\ 0 & 1 & 3 \end{pmatrix}}_{=H^T}$$

$\boxed{6h00}$ $\longrightarrow$ $f = \frac{1}{2}(2x_1^2-3x_2^2+x_1x_2)^2+\frac{1}{2}(x_2^2+3x_1x_2)^2$ ☑

LECTURE 3

<u>Definition</u> A polynomial $f \in \mathbb{R}[x]$ is called <u>homogeneous</u> (of degree d) if every monomial in the expansion of f has degree d.

<u>Example</u>

* $x^2y+3z^3-t^3+xtz$ <u>is</u> homogeneous of deg 3
* x^2+3z^5y is <u>not</u> homogeneous