

Geometric analysis of the condition of the convex feasibility problem

Dennis Amelunxen

November 5, 2009

Motivation

Condition of the convex feasibility problem

Introducing the Grassmann condition

Probabilistic analysis of the Grassmann condition

Estimation results

Motivation

Condition of the convex feasibility problem

Introducing the Grassmann condition

Probabilistic analysis of the Grassmann condition

Estimation results

Complexity of numerical algorithms

Decision problem:

$$f: \mathbb{R}^k \rightarrow \{0, 1\}, \quad A \mapsto f(A).$$

Assume we have a numerical algorithm that computes f .

How to analyze its running time?

Smale's 2-part scheme

1. Bound the running time $T(A)$ via

$$T(A) \leq \left(\text{size}(A) + (\log \text{ of } \text{condition}(A)) \right)^c.$$

2. analyze $\text{condition}(A)$ under random A

$$\text{Prob}(\text{condition}(A) \geq t) \leq \dots$$

2 forms of probabilistic analyses

Average-case analysis

- ▶ take 1 (global) distribution on the input space
- ▶ result depends heavily on the distribution
- ▶ usually too optimistic

Smoothed analysis

- ▶ take for every input \bar{a} a local distribution of dispersion σ around \bar{a} & consider the supremum over all \bar{a}
- ▶ $\sigma \rightarrow \text{diam}(\text{input space})$: average-case
 $\sigma \rightarrow 0$: “worst-case”
- ▶ result depends less on the distribution

Conic condition numbers

- ▶ input space = S^p (p -dimensional unit sphere)
- ▶ $\Sigma \subseteq S^p$ (ill-posed inputs)
- ▶ $\mathcal{C}(a) = \frac{1}{\sin d(a, \Sigma)}$, $d =$ geodesic distance on S^p (angle)

A general result

Theorem (Bürgisser, Cucker, Lotz)

$\Sigma \subseteq$ (zero set of a homog. polynom. of degree $\leq d$) $\neq S^p$,
 $\bar{a} \in S^p$, $\sigma \in (0, 1]$. Then

$$\mathbf{E}_{a \in B(\bar{a}, \sigma)} \ln \mathcal{C}(a) = \mathcal{O} \left(\ln p + \ln d + \ln \frac{1}{\sigma} \right).$$

Robustness (Cucker, Hauser, Lotz)

- ▶ extension to radially symmetric distributions
- ▶ even allowed: the density has a singularity in the center

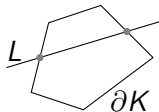
Motivating questions

- ▶ Can we get a general result for convex programming?
- ▶ What is the “degree” of convex programming?

The “degree” of the boundary of a convex body

$K \subset \mathbb{R}^m$ a convex body, L a *generic* line

$$|L \cap \partial K| = \begin{cases} 2, & \text{if } L \text{ hits } K \\ 0, & \text{else} \end{cases}$$



→ the “degree” of ∂K is 2

Motivation

Condition of the convex feasibility problem

Introducing the Grassmann condition

Probabilistic analysis of the Grassmann condition

Estimation results

$A \in \mathbb{R}^{m \times n}$ ($n > m$), $C \subseteq \mathbb{R}^n$ closed convex cone,

$$\check{C} := \{y \in \mathbb{R}^n \mid \forall x \in C : \langle y, x \rangle \geq 0\} \quad (\text{dual of } C).$$

Convex feasibility problem (CFP)

$$\exists x \in \mathbb{R}^n \setminus \{0\} : \begin{array}{l} Ax = 0 \\ x \in \check{C} \end{array} \quad (\text{P})$$

$$\exists y \in \mathbb{R}^m \setminus \{0\} : A^T y \in C \quad (\text{D})$$

$$\exists x \in \mathbb{R}^n \setminus \{0\} : \begin{array}{l} Ax = 0 \\ x \in \check{C} \end{array} \quad (P) \quad \Bigg| \quad \exists y \in \mathbb{R}^m \setminus \{0\} : A^T y \in C \quad (D)$$

Most important cases

$$\text{LP} : C = \mathbb{R}_+^n = \mathbb{R}_+ \times \dots \times \mathbb{R}_+$$

$$\begin{aligned} \text{SOCP} : C &= C_1 \times \dots \times C_k, \\ C_i &= \{x \in \mathbb{R}^{n_i} \mid x_{n_i} \geq \|(x_1, \dots, x_{n_i-1})\|\} \end{aligned}$$

$$\text{SDP} : C = \{M \in \text{Sym}_\ell \mid M \text{ is pos. semidef.}\}$$

$$\text{Sym}_\ell := \{M \in \mathbb{R}^{\ell \times \ell} \mid M^T = M\}$$

Definition

$$\mathcal{F}_P := \{A \mid (P) \text{ is feasible}\},$$

$$\mathcal{F}_D := \{A \mid (D) \text{ is feasible}\},$$

$$\Sigma := \mathcal{F}_P \cap \mathcal{F}_D.$$

Definition (Renegar's condition number)

$$\mathcal{C}_R(A) := \frac{\|A\|}{d(A, \Sigma)}.$$

Motivation

Condition of the convex feasibility problem

Introducing the Grassmann condition

Probabilistic analysis of the Grassmann condition

Estimation results

Reformulation

$$\exists x \in \mathbb{R}^n \setminus \{0\} : \begin{array}{l} Ax = 0 \\ x \in \check{C} \end{array} \quad (\text{P})$$

$$\Leftrightarrow$$

$$\underbrace{\ker A}_{=: \mathcal{W}^\perp} \cap \check{C} \neq \{0\}$$

$$\exists y \in \mathbb{R}^m \setminus \{0\} : A^T y \in C \quad (\text{D})$$

$$\text{rk } A = m$$

$$\Leftrightarrow$$

$$\underbrace{\text{im } A^T}_{=: \mathcal{W}} \cap C \neq \{0\}$$

Coordinate-free CFP

$\mathcal{W} \in \text{Gr}_{n,m} := \{\mathcal{W} \subseteq \mathbb{R}^n \mid \mathcal{W} \text{ lin. subspace, } \dim \mathcal{W} = m\}$

$$\mathcal{W}^\perp \cap \check{C} \neq \{0\} \quad (\text{P})$$

$$\mathcal{W} \cap C \neq \{0\} \quad (\text{D})$$

Definition

$$\mathcal{F}_P := \{\mathcal{W} \in \text{Gr}_{n,m} \mid \mathcal{W}^\perp \cap \check{C} \neq \{0\}\} \quad (\text{primal feasible})$$

$$\mathcal{F}_D := \{\mathcal{W} \in \text{Gr}_{n,m} \mid \mathcal{W} \cap C \neq \{0\}\} \quad (\text{dual feasible})$$

$$\Sigma_m := \mathcal{F}_P \cap \mathcal{F}_D \quad (\text{ill-posed})$$

Definition (Grassmann condition)

$$\mathcal{C}_G(A) := \mathcal{C}(\mathcal{W}) := \frac{1}{\sin d(\mathcal{W}, \Sigma_m)},$$

if $\text{rk}(A) = m$, $\mathcal{W} := \text{im}(A^T)$, $d =$ geod. distance in $\text{Gr}_{n,m}$.

Some details: $d(\mathcal{W}, \Sigma_m)$

Let

$$C \subset \mathbb{R}^n$$

$$K := C \cap S^{n-1}$$

$$\check{C} = (\text{dual of } C)$$

$$\check{K} := \check{C} \cap S^{n-1}$$

$$\mathcal{W} \in \text{Gr}_{n,m}$$

$$W := \mathcal{W} \cap S^{n-1} .$$

Then

$$d(\mathcal{W}, \Sigma_m) = \begin{cases} d(W, K) & \text{if } \mathcal{W} \cap C = \{0\} \\ d(W^\perp, \check{K}) & \text{if } \mathcal{W}^\perp \cap \check{C} = \{0\} . \end{cases}$$

(dual feasible case: \rightarrow [Belloni, Freund, 2007])

Alternative definition of $\mathcal{C}_G(A)$

Definition

$$\begin{aligned}\Sigma^\dagger &:= \{A \in \mathbb{R}^{m \times n} \mid \text{rk}(A) < m\} \\ &= \{A \in \mathbb{R}^{m \times n} \mid \kappa(A) = \infty\} \quad (\text{ill-posed})\end{aligned}$$

$$\Omega^\dagger := \{A \in \mathbb{R}^{m \times n} \mid \|A\| = \kappa(A) = 1\} \quad (\text{best-posed})$$

Proposition

$$A \in \mathbb{R}^{m \times n}, \text{rk}(A) = m$$

1. $\text{argmin}\{\|A - B\|_F \mid B \in \Omega^\dagger\} =: \tilde{A}$ is uniquely determined
2. $\ker A = \ker \tilde{A}$, $\text{im } A^T = \text{im } \tilde{A}^T$
3. $\mathcal{C}_G(A) = \mathcal{C}_R(\tilde{A})$

Some details: \tilde{A}

Let $A \in \mathbb{R}^{m \times n}$ have singular value decomposition

$$A = U \cdot \begin{pmatrix} \sigma_1 & & & \\ & \ddots & & \\ & & \sigma_m & \\ & & & 0 \end{pmatrix} \cdot V,$$

$U \in O(m)$, $V \in O(n)$. Then

$$\tilde{A} = U \cdot \begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & 0 \end{pmatrix} \cdot V.$$

Interpretation

- ▶ \tilde{A} corresponds to a preconditioning
- ▶ separation of “intrinsic” and “extrinsic” condition

Proposition

$A \in \mathbb{R}^{m \times n}$, $\text{rk}(A) = m$. Then

$$\mathcal{C}_G(A) \leq \mathcal{C}_R(A) \leq \kappa(A) \cdot \mathcal{C}_G(A).$$

(in the dual feasible case: [Belloni, Freund, 2007])

Motivation

Condition of the convex feasibility problem

Introducing the Grassmann condition

Probabilistic analysis of the Grassmann condition

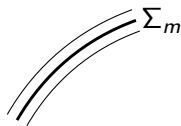
Estimation results

Main strategy

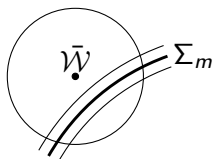
Tail estimates $\text{Prob}(\mathcal{C}(\mathcal{W}) \geq \varepsilon^{-1}) \stackrel{!}{\leq} \dots$ are volume estimates:

$$\text{Av. : } \text{Prob}(\mathcal{C}(\mathcal{W}) \geq \varepsilon^{-1}) = \frac{\text{vol}(\mathcal{T}(\Sigma_m, \alpha))}{\text{vol}(\text{Gr}_{n,m})},$$

($\varepsilon := \sin(\alpha)$)



$$\text{Sm. : } \text{Prob}(\mathcal{C}(\mathcal{W}) \geq \varepsilon^{-1}) = \frac{\text{vol}(B(\bar{\mathcal{W}}, \sigma) \cap \mathcal{T}(\Sigma_m, \alpha))}{\text{vol}(B(\bar{\mathcal{W}}, \sigma))}$$



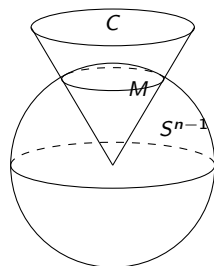
Properties of Σ_m

$M := \partial C \cap S^{n-1}$. Assume that

- ▶ M lies in an open halfsphere,
- ▶ M is a smooth hypersurface of S^{n-1} ,
- ▶ the curvature of M does not vanish.

Then

- ▶ $\Sigma_m \subset \text{Gr}_{n,m}$ orientable smooth hypersurface,
- ▶ isometry : $\Sigma_m \cong \text{Gr}(M, m-1)$ (Grassmann bundle).



Duality:

$$\Sigma_m(C) \cong \Sigma_{n-m}(\check{C}).$$

Special cases:

$$\begin{aligned} \Sigma_1(C) &\cong M \\ \Sigma_{n-1}(C) &\cong \check{M}, \quad \check{M} := \partial \check{C} \cap S^{n-1}. \end{aligned}$$

Intermission: Intrinsic volumes

Euclidean Tube Formula

Theorem (Steiner, 1840)

$K \subset \mathbb{R}^2$ convex

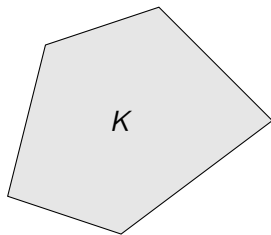
$$\text{vol}(\mathcal{I}(K, r)) = \text{vol}(K) + \text{vol}(\partial K) \cdot r + \pi \cdot r^2$$

Euclidean Tube Formula

Theorem (Steiner, 1840)

$K \subset \mathbb{R}^2$ convex

$$\text{vol}(\mathcal{I}(K, r)) = \text{vol}(K) + \text{vol}(\partial K) \cdot r + \pi \cdot r^2$$

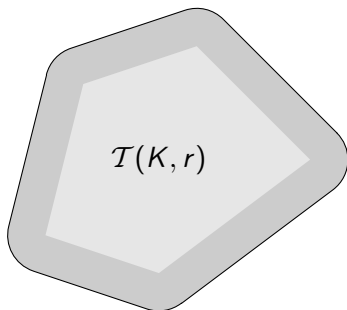


Euclidean Tube Formula

Theorem (Steiner, 1840)

$K \subset \mathbb{R}^2$ convex

$$\text{vol}(\mathcal{I}(K, r)) = \text{vol}(K) + \text{vol}(\partial K) \cdot r + \pi \cdot r^2$$

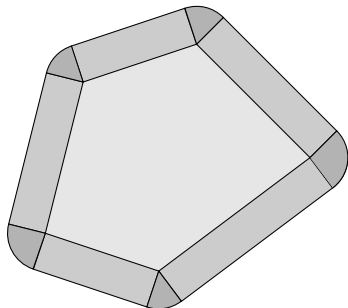


Euclidean Tube Formula

Theorem (Steiner, 1840)

$K \subset \mathbb{R}^2$ convex

$$\text{vol}(\mathcal{I}(K, r)) = \text{vol}(K) + \text{vol}(\partial K) \cdot r + \pi \cdot r^2$$

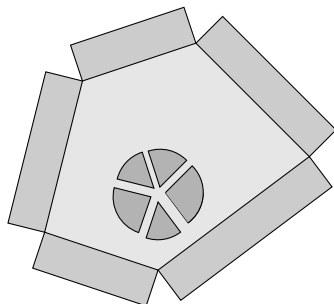


Euclidean Tube Formula

Theorem (Steiner, 1840)

$K \subset \mathbb{R}^2$ convex

$$\text{vol}(\mathcal{I}(K, r)) = \text{vol}(K) + \text{vol}(\partial K) \cdot r + \pi \cdot r^2$$



Euclidean Tube Formula

Theorem (Steiner, 1840)

$K \subset \mathbb{R}^2$ convex

$$\text{vol}(\mathcal{I}(K, r)) = \text{vol}(K) + \text{vol}(\partial K) \cdot r + \pi \cdot r^2$$

General case

$K \subset \mathbb{R}^n$ convex

$$\text{vol}(\mathcal{I}(K, r)) = \sum_{i=0}^n V_i(K) \cdot r^i$$

$$\mathcal{K}(\mathbb{R}^n) := \{K \subset \mathbb{R}^n \mid K \text{ compact \& convex}\},$$

$$K, K_1, K_2, K_1 \cup K_2 \in \mathcal{K}(\mathbb{R}^n).$$

$$(1) \quad V_i(K_1 \cup K_2) = V_i(K_1) + V_i(K_2) - V_i(K_1 \cap K_2)$$

$$(2) \quad V_i(\emptyset) = 0$$

$$(3) \quad V_i(Q \cdot K) = V_i(K) \quad \forall Q \in O(n)$$

$$(4) \quad V_i(a + K) = V_i(K) \quad \forall a \in \mathbb{R}^n$$

$$(5) \quad V_i \text{ is continuous .}$$

Theorem (Hadwiger, 1950s)

If $\mu: \mathcal{K}(\mathbb{R}^n) \rightarrow \mathbb{R}$ satisfies (1)-(5) then

$$\mu = \sum_{i=0}^n c_i \cdot V_i$$

for some uniquely determined $c_0, \dots, c_n \in \mathbb{R}$.

Euclidean space \rightarrow Sphere

Transfer of integral geometric results from the euclidean space to the spherical setting was the subject of



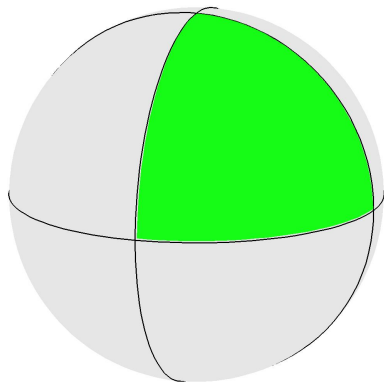
S. Glasauer.

Integralgeometrie konvexer Körper im sphärischen Raum.

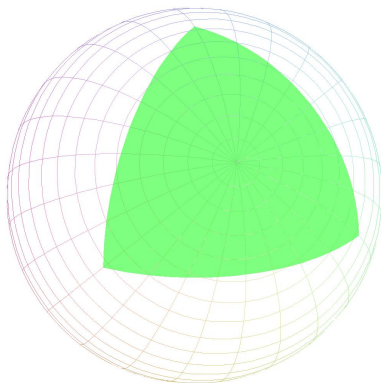
PhD Thesis, 1995.

(available at <http://www.hs-augsburg.de/~glasauer/publ/diss.pdf>)

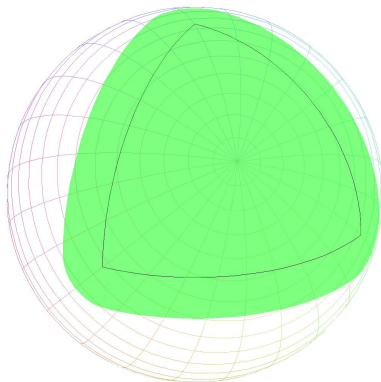
Example: Positive orthant



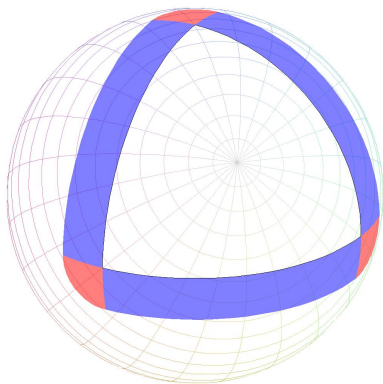
Example: Positive orthant



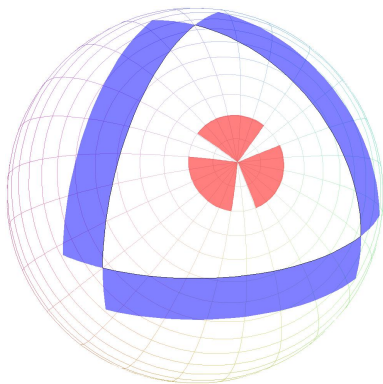
Example: Positive orthant



Example: Positive orthant



Example: Positive orthant



Spherical Steiner formula (Weyl)

Let

- ▶ $K \subset S^{n-1}$ spherically convex,
- ▶ $0 < \alpha < \frac{\pi}{2}$,
- ▶ $S^i \subset S^{n-1}$ subsphere.

Then

$$\text{vol}(\mathcal{I}(K, \alpha)) = \text{vol } K + \sum_{i=0}^{n-2} V_i(K) \cdot \text{vol}(\mathcal{I}(S^i, \alpha)).$$

Properties

- ▶ $V_i(S^k) = \delta_{ik}$
- ▶ If $K \subset S^{n-1}$ not a subsphere, then $V_i(K) \leq \frac{1}{2}$.
- ▶ $\sum_{i=0}^{n-2} V_i(K) = 1$

One first application

$A \in \mathbb{R}^{m \times n}$ Gaussian matrix

($\rightarrow \mathcal{W} := \text{im}(A^T) \in \text{Gr}_{n,m}$ uniformly at random).

Then

$$\begin{aligned} \text{Prob}(A \text{ is dual feasible}) &= \text{Prob}(\mathcal{W} \cap C \neq \{0\}) \\ &= 2 \cdot \sum_{k=0}^{\lfloor \frac{m-1}{2} \rfloor} V_{n-m+2k}(K). \end{aligned}$$

(for LP: Wendel, 1962)

LP (Positive orthant) (polyhedral)

$$V_{n-m-1} = \frac{\binom{n}{m}}{2^n}$$

SOCP-1, (single Lorentz cone) (smooth)

$$V_{n-m-1} \approx \left(\frac{\binom{n}{m}}{2^n} \right)^{\frac{1}{2}}$$

SOCP (product of Lorentz cones) (stratified)

... (found closed formula)

SDP (Positive semidefinite cone) (stratified)

... (found closed formula)

End of intermission.

Spherical Tube Formula

$\mathcal{T}^\circ(K, \alpha) := \mathcal{T}(K, \alpha) \setminus K$ (outer tube)

$$\text{vol}(\mathcal{T}^\circ(K, \alpha)) = \sum_{i=0}^{n-2} V_i(K) \cdot \text{vol}(\mathcal{T}(S^i, \alpha))$$

Theorem (“Grassmannian Tube Formula”)

$$\text{vol}(\mathcal{T}^\circ(\Sigma_m, \alpha)) = f(n, m) \cdot \sum_{i=0}^{n-2} V_i(K) \cdot \sum_{k=0}^{n-2} c_{ik} \cdot \text{vol}(\mathcal{T}(S^k, \alpha))$$

$m \in \{1, n-1\}$: Spherical case,

$\text{vol } \Sigma_m$: [Firey, Schneider, Teufel, ...],

General case : [Glasauer]

→ closed formulas for the tail of \mathcal{C}_G

Summation pattern

$$\text{vol}(\mathcal{T}^\circ(\Sigma_m, \alpha)) = f(n, m) \cdot \sum_{i=0}^{n-2} V_i(K) \cdot \sum_{k=0}^{n-2} c_{ik} \cdot \text{vol}(\mathcal{T}(S^k, \alpha))$$

$[c_{ik}]$

$$\begin{array}{c}
 i \\
 \left[\begin{array}{cccccccc}
 * & & & & & & & \\
 & * & & & & & & \\
 & & * & & & & & \\
 & & & * & & & & \\
 & & & & * & & & \\
 & & & & & * & & \\
 & & & & & & * & \\
 & & & & & & & * \\
 & & & & & & & & *
 \end{array} \right] \\
 k
 \end{array}$$

$$n = 10, m = 1$$

Summation pattern

$$\text{vol}(\mathcal{I}^\circ(\Sigma_m, \alpha)) = f(n, m) \cdot \sum_{i=0}^{n-2} V_i(K) \cdot \sum_{k=0}^{n-2} c_{ik} \cdot \text{vol}(\mathcal{I}(S^k, \alpha))$$

$[c_{ik}]$

$$\begin{bmatrix} & & & * & & & \\ & & & * & * & & \\ & & * & * & * & & \\ * & * & * & * & * & * & \\ & * & * & * & * & * & * \\ & & * & * & * & * & \\ & & * & * & * & & \\ & & * & * & * & & \\ & & & * & * & & \\ & & & * & * & & \end{bmatrix}$$

$$n = 10, m = 5$$

Summation pattern

$$\text{vol}(\mathcal{T}^\circ(\Sigma_m, \alpha)) = f(n, m) \cdot \sum_{i=0}^{n-2} V_i(K) \cdot \sum_{k=0}^{n-2} c_{ik} \cdot \text{vol}(\mathcal{T}(S^k, \alpha))$$

$[c_{ik}]$

$$\begin{bmatrix} & & & & & * \\ & & & & * & * \\ & & & * & * & * \\ & & * & * & * & * \\ * & * & * & * & * & * \\ * & * & * & * & * & * \\ * & * & * & * & * & * \\ * & * & * & * & * & * \end{bmatrix}$$

$$n = 10, m = 6$$

Summation pattern

$$\text{vol}(\mathcal{T}^\circ(\Sigma_m, \alpha)) = f(n, m) \cdot \sum_{i=0}^{n-2} V_i(K) \cdot \sum_{k=0}^{n-2} c_{ik} \cdot \text{vol}(\mathcal{T}(S^k, \alpha))$$

$[c_{ik}]$

$$\begin{bmatrix} & & & & & & & * \\ & & & & & * & * & * \\ & & & & * & * & * & * \\ & & & * & * & * & * & * \\ & & * & * & * & * & * & * \\ & * & * & * & * & * & * & * \\ * & * & * & * & * & * & * & * \\ * & * & * & * & * & * & * & * \end{bmatrix}$$

$$n = 10, m = 7$$

Summation pattern

$$\text{vol}(\mathcal{T}^\circ(\Sigma_m, \alpha)) = f(n, m) \cdot \sum_{i=0}^{n-2} V_i(K) \cdot \sum_{k=0}^{n-2} c_{ik} \cdot \text{vol}(\mathcal{T}(S^k, \alpha))$$

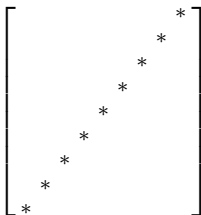
$$[c_{ik}]$$

$$\begin{bmatrix} & & & & & & & * \\ & & & & & & * & * \\ & & & & & * & * & * \\ & & & & * & * & * & * \\ & & * & * & * & * & * & * \\ & * & * & * & * & * & * & * \\ * & * & * & * & * & * & * & * \end{bmatrix}$$

$$n = 10, m = 8$$

Summation pattern

$$\text{vol}(\mathcal{T}^o(\Sigma_m, \alpha)) = f(n, m) \cdot \sum_{i=0}^{n-2} V_i(K) \cdot \sum_{k=0}^{n-2} c_{ik} \cdot \text{vol}(\mathcal{T}(S^k, \alpha))$$

 $[c_{ik}]$


$$n = 10, m = 9$$

Motivation

Condition of the convex feasibility problem

Introducing the Grassmann condition

Probabilistic analysis of the Grassmann condition

Estimation results

Main Theorem

Let

- ▶ $A \in \mathbb{R}^{m \times n}$ Gaussian matrix,
- ▶ $\varepsilon := \sin(\alpha) < n^{-\frac{3}{2}}$.

Then for any convex cone $C \subset \mathbb{R}^n$ we have

$$\text{Prob} \left(\mathcal{C}_G(A) > \frac{1}{\varepsilon} \right) < 6 \cdot n \cdot \varepsilon .$$

In particular,

$$\mathbf{E}(\ln \mathcal{C}_G(A)) < 2.5 \cdot \ln(n) + 2.8 .$$

Using result by Chen, Dongarra (2005)

$$\mathbf{E}(\ln \kappa(A)) < \ln \left(\frac{m+1}{2} \right) + 2.26$$

we get:

Corollary

$A \in \mathbb{R}^{m \times n}$ Gaussian matrix. Then

$$\mathbf{E}(\ln \mathcal{C}_R(A)) < 3.5 \cdot \ln(n) + 5 .$$

Special cases

$$\text{arbitrary cone} : \mathbf{E}(\ln \mathcal{C}_G(A)) < 2.5 \cdot \ln(n) + 2.8$$

Using the corresponding formulas for $V_i(K)$, we get

$$\text{LP} : \mathbf{E}(\ln \mathcal{C}_G(A)) < 1.5 \cdot \ln(m) + 6$$

$$\text{SOCP-1} : \mathbf{E}(\ln \mathcal{C}_G(A)) < \ln(m) + 8 .$$

From average analysis to smoothed analysis

- ▶ 1st approach: Uniform distribution on geodesic balls in $\text{Gr}_{n,m}$:

$$\frac{\text{vol}(B(\bar{W}, \sigma) \cap \mathcal{T}(\Sigma_m, \alpha))}{\text{vol}(B(\bar{W}, \sigma))} \stackrel{!}{<} \dots$$

- ▶ Grassmannian Tube formula relies on “local intrinsic volumes” (curvature measures).
 - Proof techniques from [Bürgisser, Cucker, Lotz] should carry over.