Cálculo del Volumen en Dimensión Alta

Santosh Vempala

El Plan

1. Intro to high dimension and convexity
   - Volume distribution, logconcavity
   - Ellipsoids
   - Lower bounds
2. Algorithms
   - Rounding
   - Volume/Integration
   - Optimization
   - Sampling
3. Probability
   - Markov chains
   - Conductance
   - Mixing of ball walk
   - Mixing of hit-and-run
4. Geometry
   - Isoperimetry
   - Concentration
   - Localization and applications

- Open problems

Muestreo

Input: a convex set K with a membership oracle

Output: sample a uniform random point in K.

Reference	Complexity	New ingredient(s)
[Dyer et al. 1991]	n^{20}	Everything
[Lovász and Simonovits 1990]	n^{13}	Localization lemma
[Applegate and Kannan 1990]	n^{7}	Logconcave sampling
[Lovász 1990]	n^{7}	Ball walk
[Lovász and Simonovits 1993]	n^{5}	Many improvements
[Kannan et al. 1997]	n^{3}	Isotropy, speedy walk
[Lovász and Vempala 2003b]	n^{3}	hit-and-run

Conjectured Lower Bound: n^2.

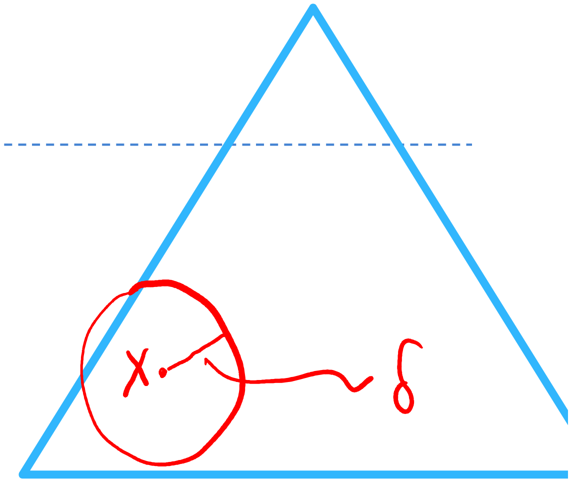
Parte 4: Geometría

- Isoperimetry
- Localization
- Concentration
- Other applications

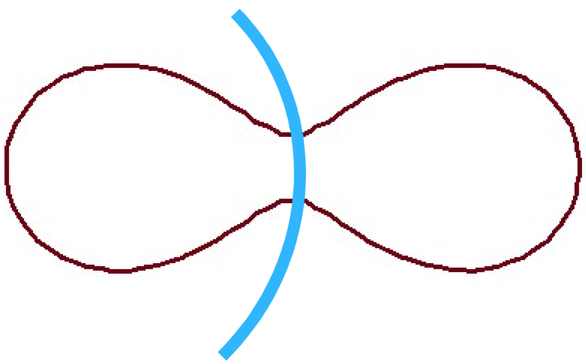
El Paseo de Bola

At x, pick random y from $x + \delta B_n$,

if y is in K, go to y.



This walk may get trapped on one side if the set is not convex.



Isoperimetria/Cheeger constant

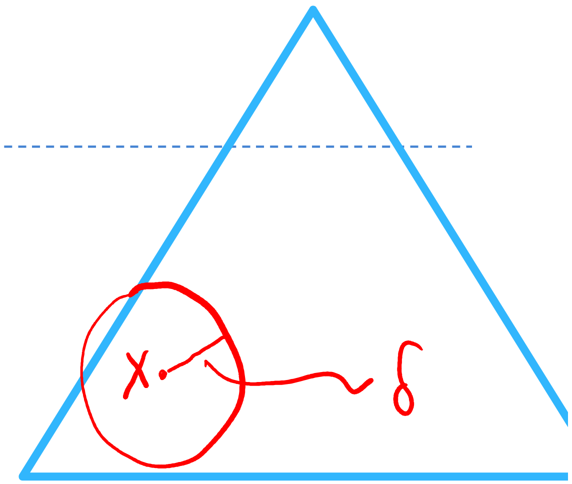
$$\frac{1}{\psi_K} = \inf_{S \subset K} \frac{vol(\partial S)}{\min(vol(S), vol(S^c))}$$

(∂S is the boundary of S)

El Paseo de Bola

At x, pick random y from $x + \delta B_n$,

if y is in K, go to y.



$$\frac{1}{\psi_K} = \min_S \frac{vol(\partial S)}{\min(vol(S), vol(S^c))}$$

<u>Theorem:</u>

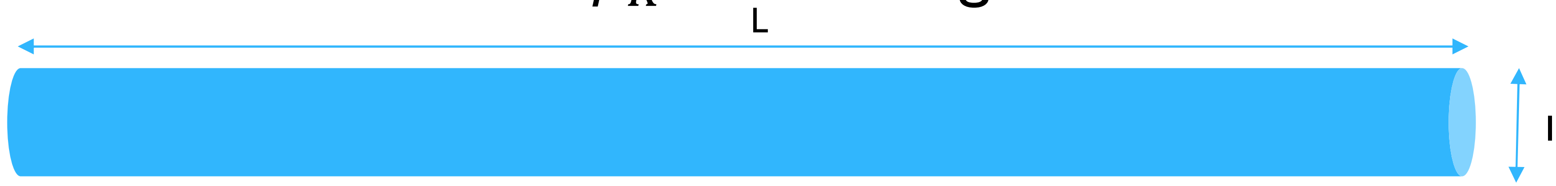
Mixing time of the ball walk is $O^*\left(n^2 \psi_{\mathrm{K}}^2\right)$.

Que es ψ_K?

$$\frac{1}{\psi_K} = \min_S \frac{\text{Area}(\partial S)}{\min(vol(S), vol(S^c))}$$

#steps = $O\left(n^2 \psi_K^2\right)$

Even for convex set K, ψ_K can be large.



$$\psi_K = L.$$

<u>Kannan-Lovász-Simonovits Conjecture:</u>

For any isotropic convex K, $\psi_K = O(1)$.

If true, the Ball Walk takes $O^*(n^2)$ for isotropic K.

KLS conjetura y conjeturas relacionados

Slicing Conjecture:

Any convex set K with $\mathrm{Cov}(K) = \lambda I$, we have $\lambda = O(\psi^2)$.

Thin-Shell Conjecture:

For isotropic convex K, $\mathbb{E}\left(\left(\|x\| - \sqrt{n}\right)^2\right) = O(\psi^2)$.

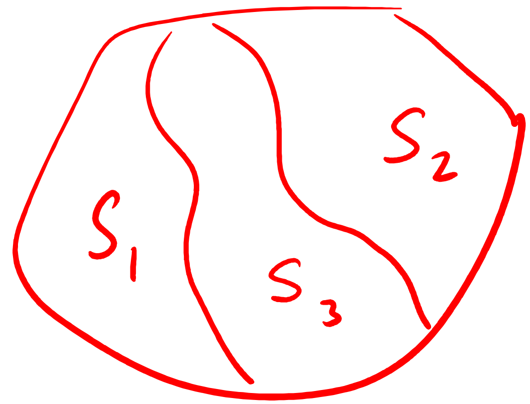
Generalized Levy concentration:

For logconcave distribution f, 1-Lipschitz f with $\mathbb{E}f = 0$,

$$\mathbb{P}(f(x) > t) = \exp(-\Omega(t/\psi)).$$

KLS conjecture implies all of them.

Isoperimetria



Thm. [LS90,DF91]

$$vol(S_3) \geq \frac{2d(S_1,S_2)}{D} \min vol(S_1), vol(S_2)$$

Extends to logconcave densities:

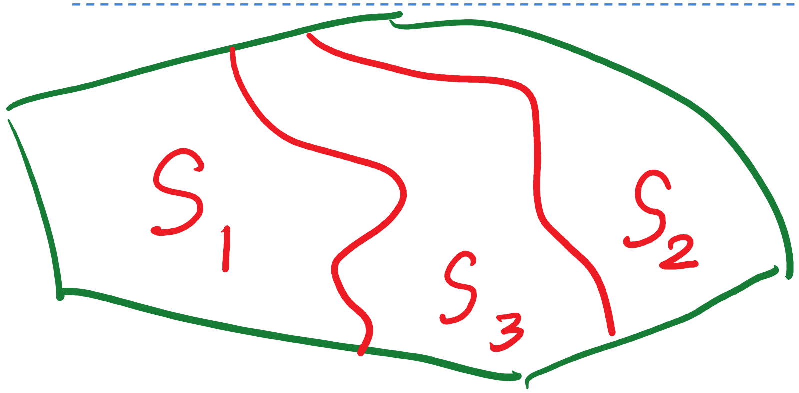
$$\pi_f(\partial S) \geq \frac{2}{D} \min \pi_f(S), \pi_f(\bar{S})$$

Equivalent to:

$$\psi_f \leq \frac{D}{2}$$

D is the diameter of the support.

Isoperimetria



$$\pi(S_3) \geq \frac{c}{D} d(S_1, S_2) \min \pi(S_1), \pi(S_2)$$

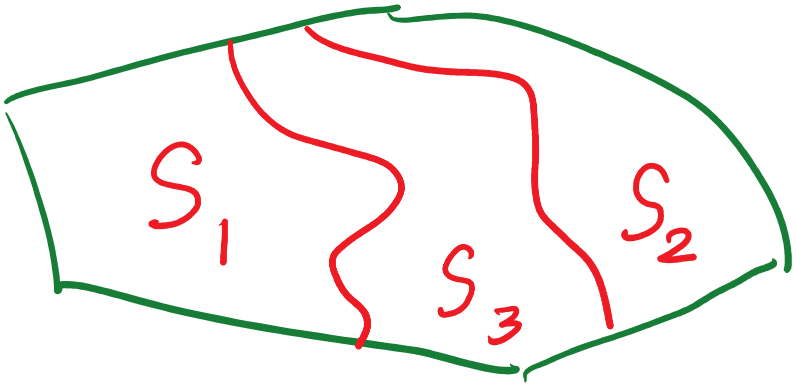
A = $E((x-\overline{x})(x-\bar{x})^T)$: covariance matrix of π

$$R^2 = E_\pi\left(\|x-\bar{x}\|^2\right) = Tr(A) = \sum_i \lambda_i(A)$$

Thm. [KLS95]. $\pi(S_3) \geq \frac{c}{R} d(S_1, S_2) \min \pi(S_1), \pi(S_2)$

$$\psi = O(R) = O\left(\sqrt{n}\right) \text{ for isotropic distributions}$$

Isoperimetria: Conjetura de KLS



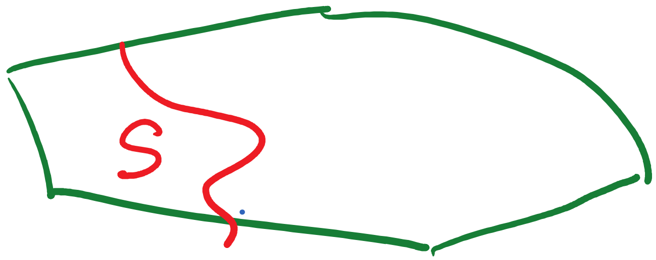
A = $E((x - \overline{x})(x - \bar{x})^T)$: covariance matrix of π

$$R^2 = E_\pi\left(\|x - \bar{x}\|^2\right) = Tr(A) = \sum_i \lambda_i(A)$$

Thm. [KLS95]. $\psi = O(R) = O\left(\sqrt{n}\right)$ for isotropic distributions

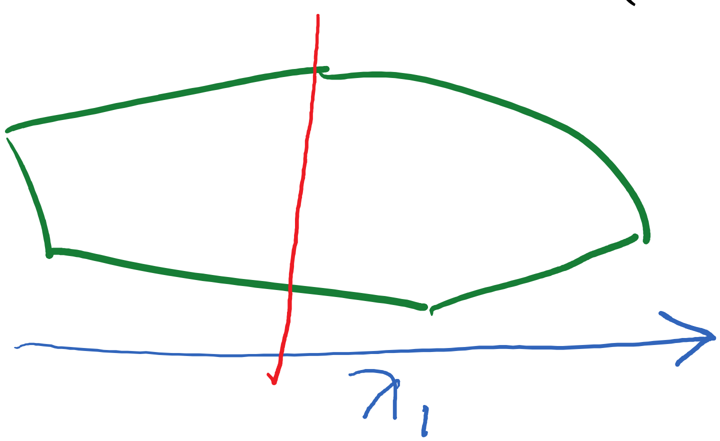
Conj. [KLS95]. $\psi = O\left(\sqrt{\lambda_1(A)}\right) = O(1)$ for isotropic dist.

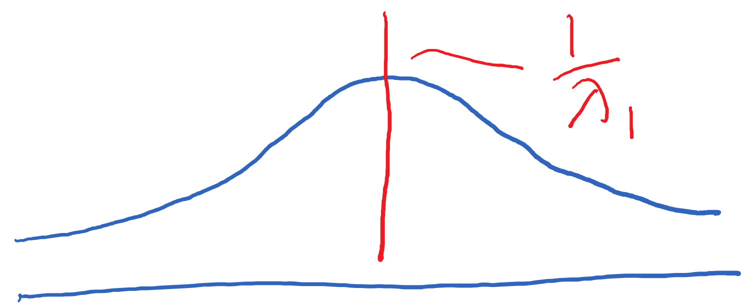
Conjetura de hiperplano KLS



$$A = E(xx^T)$$

Conj. [KLS95]. $\psi = O\left(\sqrt{\lambda_1(A)}\right)$



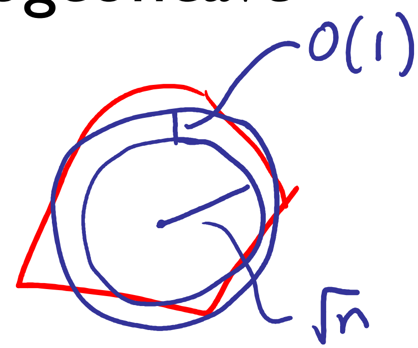


Conj. The minimum expansion/isoperimetry is given by a hyperplane cut up to a constant factor!

Conjetura de Cascaron Delgado

Conj. (Thin shell). For X from any isotropic logconcave distribution,

$$Var\left(\left|\left|X\right|\right|^2\right) = O(n)$$
$$E\left(\left(\left|\left|X\right|\right| - \sqrt{n}\right)^2\right) = O(1)$$



"Most of an isotropic logconcave distribution is contained in an annulus of constant thickness."

Current best bound [Guedon-E. Milman]: $n^{1/3}$

Conjetura de Slicing (o Hiperplano)

Conj. Any convex body of unit volume has a hyperplane section whose volume is at least a constant (independent of dimension).

Equivalently:

For any isotropic logconcave density $\max f^{1/n} = O(1)$.

Current best bound: $n^{1/4}$ [Bourgain86, Klartag]

KLS, Slicing, Thin-shell,…

	current bound	
slicing	$n^{1/4}$	[Bourgain; Klartag]
thin shell	$n^{1/3}$	[Guedon-Milman]
KLS	$n^{1/3}\sqrt{\log n}$	[Bobkov; Guedon-Milman+Eldan]
Poincare	$n^{1/3}\sqrt{\log n}$	
Lipschitz Conc.	$n^{1/3}\sqrt{\log n}$	

All are conjectured to be $O(1)$ for any logconcave density

Thm. [Cousins-V.13] True for any density proportional to the product of a logconcave function and a Gaussian.

KLS, Slicing, Thin-shell, Concentration

	bound	
slicing	$n^{1/4}$	[Bourgain; Klartag]
thin shell	$n^{1/3}$	[Guedon-E.Milman]
KLS	$n^{1/3}\sqrt{\log n}$	[Bobkov; Guedon-E.Milman+Eldan]
Poincaré	$n^{1/3}\sqrt{\log n}$	[Maz'ja; Cheeger]
Lipschitz Conc.	$n^{1/3}\sqrt{\log n}$	[Gromov-Milman]

All are conjectured to be $O(1)$ for any logconcave density

Thm. [Lee-V. December'16]. For isotropic logconcave distributions, $\psi = O(n^{1/4})$.

$$\psi \leq C\left(\sum_i \lambda_i(A)^2\right)^{\frac{1}{4}} = C\left(Tr(A^2)\right)^{\frac{1}{4}}$$

Implica que todas las constantes tienen el mismo limité!

Isoperimetria por localización

- $\pi_f(S_3) \geq \frac{2d(S_1,S_2)}{D} \min \pi_f(S_1),\ \pi_f(S_2)$

- Write as 2 inequalities:

$$\pi_f(S_1) \leq \pi_f(S_2), \qquad \pi_f(S_3) \geq \psi\, \pi_f(S_1)$$

Let $g(x) = f(x)(1_{S_2}(x) - 1_{S_1}(x)), h(x) = f(x)(\psi\, 1_{S_1}(x) - 1_{S_3}(x))$

- Then, need to show: $\int g \geq 0 \Rightarrow \int h \leq 0.$

- Suppose not, i.e., $\exists S_1, S_2, S_3$: $\int g \geq 0, \int h > 0.$

Idea:

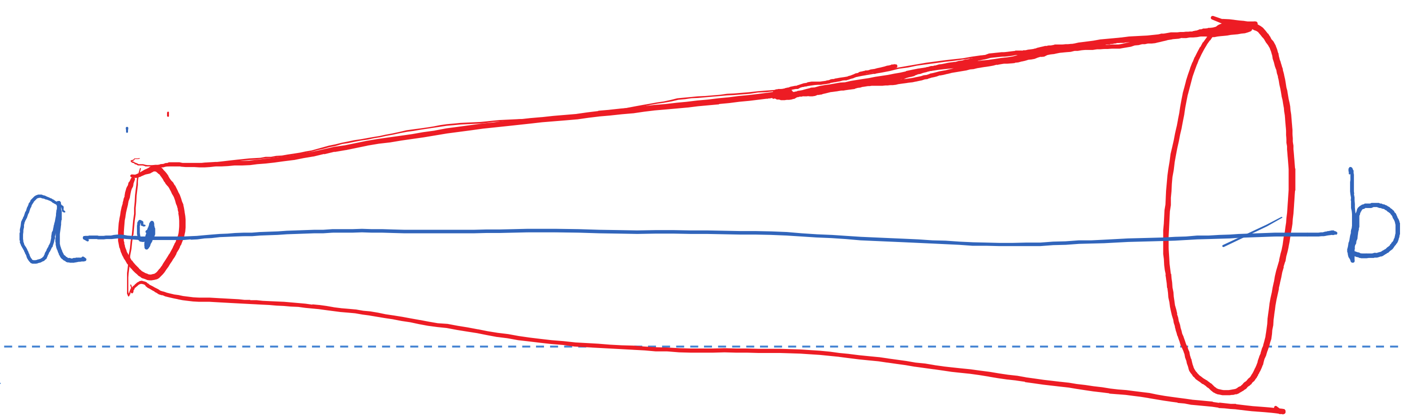
1. No such counterexample in one dimension
2. If such a counterexample exists in some dimension, then it also exists in 1 dimension.

Lemma de Localización [LS, KLS]

Lemma. Let $g, h: R^n \to R$ be integrable, lower semi-continuous functions. Suppose $\int g,\ \int h > 0$. Then, there exists an interval $[a, b] \subset R^n$ and a linear function $\ell: [0,1] \to R_+$ s.t.

$$\int_0^1 g\big((1-t)a + tb\big)\ell(t)^{n-1}\,dt > 0$$

$$\int_0^1 h\big((1-t)a + tb\big)\ell(t)^{n-1}\,dt > 0.$$



Isoperimetria por localización

$$vol(S_1) \leq vol(S_2) \Rightarrow vol(S_3) \geq \psi\, vol(S_1)$$

$$g(x) = 1_{S_2}(x) - 1_{S_1}(x), \qquad h(x) = \psi\, 1_{S_1}(x) - 1_{S_3}(x)$$

- Need to show: $\int g \geq 0 \Rightarrow \int h \leq 0$.
- Suppose not, i.e., for some partition, $\int g \geq 0, \int h > 0$.

Applying localization,

$$\int_0^1 g\big((1-t)a + tb\big)\ell(t)^{n-1}\,dt > 0, \int_0^1 h\big((1-t)a + tb\big)\ell(t)^{n-1}\,dt > 0.$$

Let $Z_i = \{t \in [0,1]: (1-t)a + tb \in S_i\},\ F(t) = f\big((1-t)a + tb\big)\ell(t)^{n-1}$.

- Then this means that

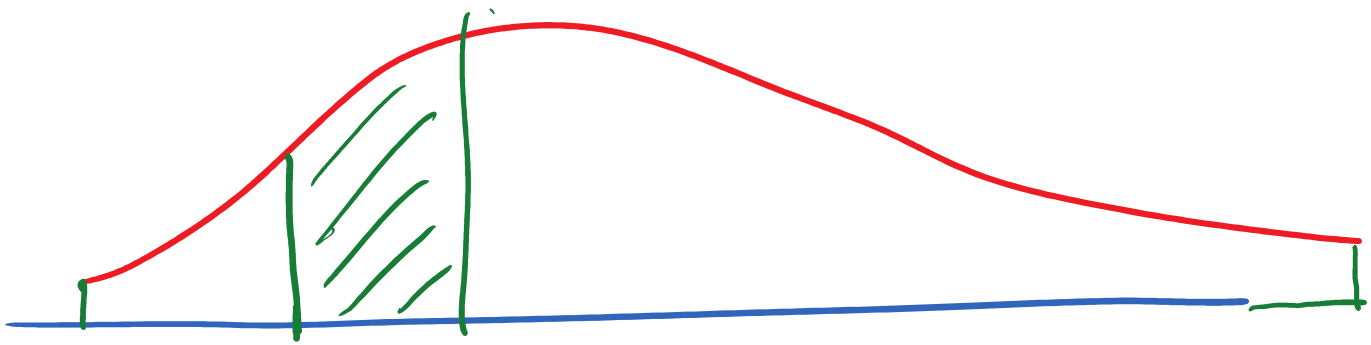
$\int_{Z_1} F \leq \int_{Z_2} F$, but $\int_{Z_3} F < \psi \int_{Z_1} F$, a 1-d counterexample must exist.

Isoperimetria en una dimensión

- For any logconcave function:

$$\int_{S_3} F \geq \frac{2d(S_1, S_2)}{D} \min \int_{S_1} F \, , \int_{S_2} F$$

- Suffices to show it for partition of into 3 intervals.



- Without factor of 2, follows from unimodality!
- Therefore, same isoperimetric ratio holds in R^n.

Localización: muchas aplicaciónes

- Carberry-Wright on concentration of polynomials
- Isoperimetry for cross-ratio distance, Dikin distance etc.
- For the analysis of algorithms

- Can be viewed as proof by contradiction
- Or as proof by induction [Chandrasekaran-Dadush-V., Eldan]

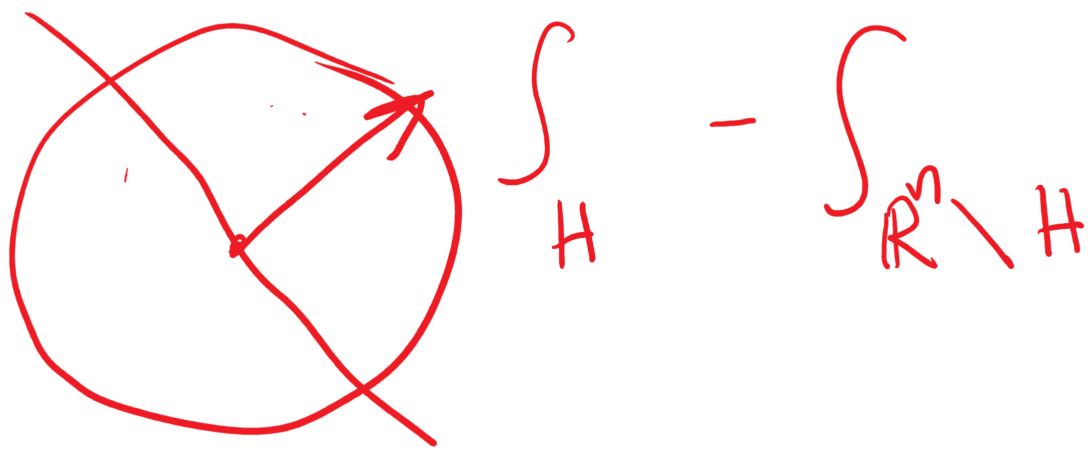
Prueba del lemma de localización

- $g, h: R^n \to R, \ \int g, \int h > 0.$

1. Find a bisecting halfspace for one function
2. Show support of limit of bisections is an interval or a point.
3. The limit function has a concave profile
4. Reduce to linear cross-sectional profile.

Bisección

- Given $\int g,\ \int h > 0$, find a halfspace H s.t. $\int_H g = \int_{R^n \setminus H} g$
- Claim: for any (n-2)-dim affine subspace A, there is a bisecting halfspace containing A in its bounding hyperplane.
- From this we get a halfspace H s.t. $\int_H g\ ,\ \int_H h > 0$
- Proof.



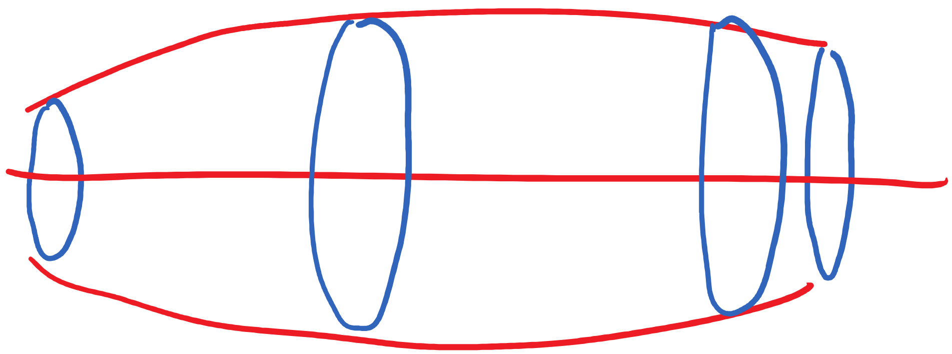
- Orthogonal to A, there is a 2-d space so the normal vector can be rotated in a complete circle. For a normal vector v, and corresponding halfspace H(v), consider $\int_{H(v)} g - \int_{R^n \setminus H(v)} g$. This is an odd, continuous function on the circle and therefore is zero somewhere.

Limité de bisecciónes

- Apply bisection as long as the support contains a 2-d plane with an interior rational point.
- Consider the sequence of supports $K \supset K_1 \supset K_2 \supset \cdots$
- The supports are convex
- The limiting support is 1-d: a point or an interval.

Perfil Concavo

Consider cross-sectional radius function.



- Replace each cross-section by ball of same volume; then by Brunn-Minkowski, the radius is a concave function
- Each concave function is bounded and the sequence converges to a concave limit.
- going to linear functions takes more work!
- Generalized to more inequalities [Fradelizi-Guedon].

Muestreo de Gaussiano

- KLS conjecture holds for Gaussian restricted to any convex body (via Brascamp-Lieb inequality).

Thm [Cousins-V.13].

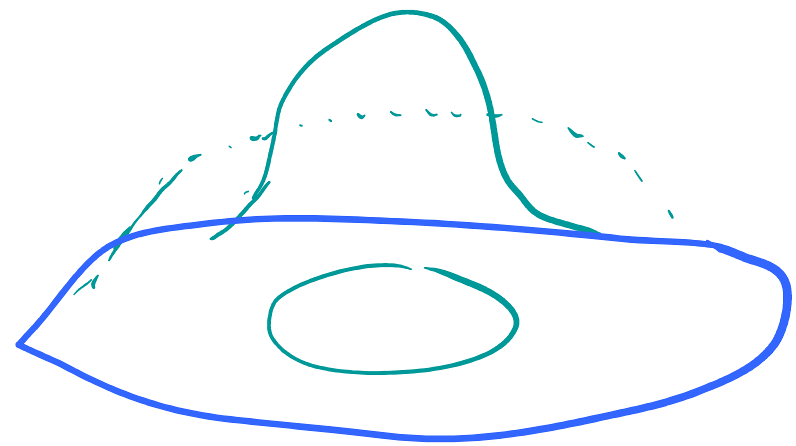
For any logconcave g, and $f(X) = g(X)e^{-\frac{||X||^2}{2\sigma^2}}$

$$\pi(S_3) \geq \frac{c}{\sigma} d(S_1, S_2) \min \pi(S_1), \pi(S_2)$$

- Leads to faster sampling:

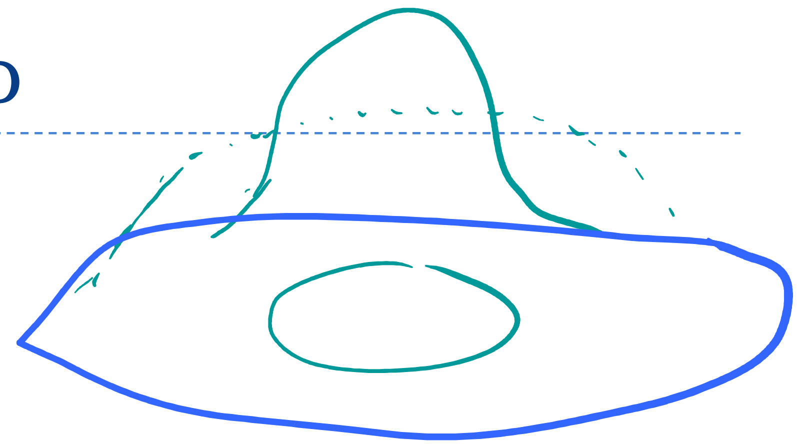
Thm. [CV13]. The ball walk applied to a Gaussian $N(0, \sigma^2 I_n)$ restricted to any convex body containing the unit ball mixes in $O^*\left(n^2 \max\ \{\sigma^2, 1\}\right)$ steps from a warm start.

Enfriamiento Gaussiano (Recocido adaptativo)



- $f_i(X) = e^{-\frac{\|X\|^2}{2\sigma_i^2}}$
- $\sigma_0^2 = \frac{1}{n}, \sigma_m^2 = O(n).$
- Estimate $\frac{\int f_{i+1}}{\int f_i}$ using samples drawn according to f_i
- For $\sigma_i^2 \leq 1$, set $\sigma_i^2 = \sigma_{i-1}^2\left(1 + \frac{1}{\sqrt{n}}\right)$
- For $\sigma_i^2 > 1$, set $\sigma_i^2 = \sigma_{i-1}^2\left(1 + \frac{\sigma_{i-1}}{\sqrt{n}}\right)$

Enfriamento Gaussiano



- $f_i(X) = e^{-\frac{||X||^2}{2\sigma_i^2}}$

For $\sigma_i^2 \leq 1$, we set $\sigma_i^2 = \sigma_{i-1}^2\left(1 + \frac{1}{\sqrt{n}}\right)$

- Sampling time: n^2,
- #phases, #samples per phase: $\sqrt{n}$

- So, total time = $n^2 \times \sqrt{n} \times \sqrt{n} = n^3$

Enfriamento Gaussiano

- $f_i(X) = e^{-\frac{||X||^2}{2\sigma_i^2}}$

For $\sigma_i^2 > 1$, we set $\sigma_i^2 = \sigma_{i-1}^2\left(1 + \frac{\sigma_{i-1}}{\sqrt{n}}\right)$

- Sampling time: $\sigma^2 n^2$ (too much?!)
- #phases to double σ is $\frac{\sqrt{n}}{\sigma}$
- #samples per phase is also $\frac{\sqrt{n}}{\sigma}$

- So, total time to double σ is $\frac{\sqrt{n}}{\sigma} \times \frac{\sqrt{n}}{\sigma} \times \sigma^2 n^2 = n^3$

Varianza del estimador del ratio

- Why can we set σ_i^2 as high as $\sigma_{i-1}^2\left(1+\frac{\sigma_{i-1}}{\sqrt{n}}\right)$?

$f\left(\sigma_i^2, x\right) = e^{-\frac{||x||^2}{2\sigma^2}}$ for $x \in K$

$F\left(\sigma_i^2\right) = \int f\left(\sigma_i^2, x\right) dx$

Lemma. $Y = \frac{f(\sigma^2, X)}{f\left(\frac{\sigma^2}{1+\alpha}, X\right)}$ $\quad E(Y) = \frac{F(\sigma^2)}{F\left(\frac{\sigma^2}{1+\alpha}\right)}$

$$\frac{E(Y^2)}{E(Y)^2} = \frac{F\left(\frac{\sigma^2}{1+\alpha}\right) F\left(\frac{\sigma^2}{1-\alpha}\right)}{F(\sigma^2)^2} = e^{O\left(\frac{\alpha^2 n}{\sigma^2}\right)} = O(1)$$

for $\alpha = O\left(\frac{\sigma}{\sqrt{n}}\right)$.

Varianza del estimador del ratio

$$\frac{E(Y^2)}{E(Y)^2} = \frac{F\left(\frac{\sigma^2}{1+\alpha}\right)F\left(\frac{\sigma^2}{1-\alpha}\right)}{F(\sigma^2)^2} = e^{O\left(\frac{\alpha^2 n}{\sigma^2}\right)}$$

First use localization to reduce to 1-d inequality,

for a restricted family of logconcave functions:

For $K \in R \cdot B_n$ and $-R \leq l \leq u \leq R$

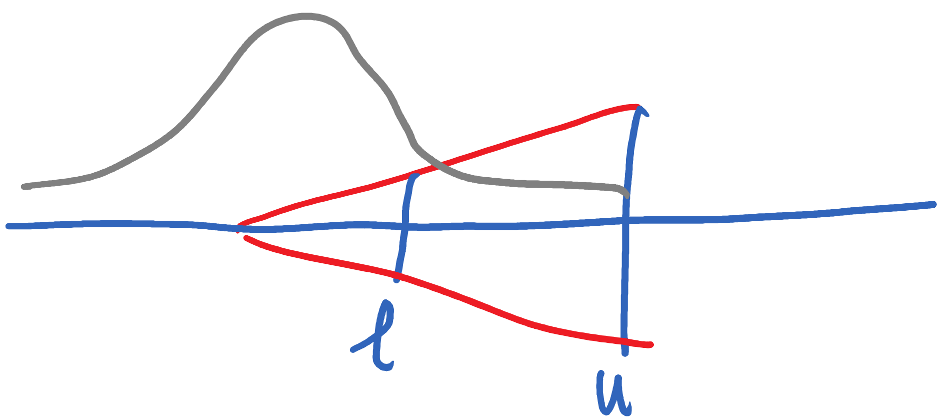
$$G(\sigma^2) = \int_l^u (t+b)^{n-1}\, e^{-\frac{\mathrm{t}^2}{2\sigma^2}} \mathrm{dt}$$

$$\frac{G\left(\frac{\sigma^2}{1+\alpha}\right)G\left(\frac{\sigma^2}{1-\alpha}\right)}{G(\sigma^2)^2} = e^{O\left(\frac{\alpha^2 R^2}{\sigma^2}\right)}$$

Varianza del estimador del ratio

$$\frac{G\left(\frac{\sigma^2}{1+\alpha}\right)G\left(\frac{\sigma^2}{1-\alpha}\right)}{G(\sigma^2)^2} = e^{O\left(\frac{\alpha^2 R^2}{\sigma^2}\right)} \quad \Leftrightarrow$$

$$E\left(t^2\right) = \frac{\int_l^u t^2 (t+b)^{n-1} e^{-\frac{t^2}{2\sigma^2}} dt}{\int_l^u (t+b)^{n-1} e^{-\frac{t^2}{2\sigma^2}} dt}, \qquad \frac{dE\left(t^2\right)}{\mathrm{d}\sigma^2} \leq c$$



Enfriamento Gaussiano [CV15]

- Adaptive annealing
- $1 + \frac{\sigma}{\sqrt{n}}$ is best possible rate

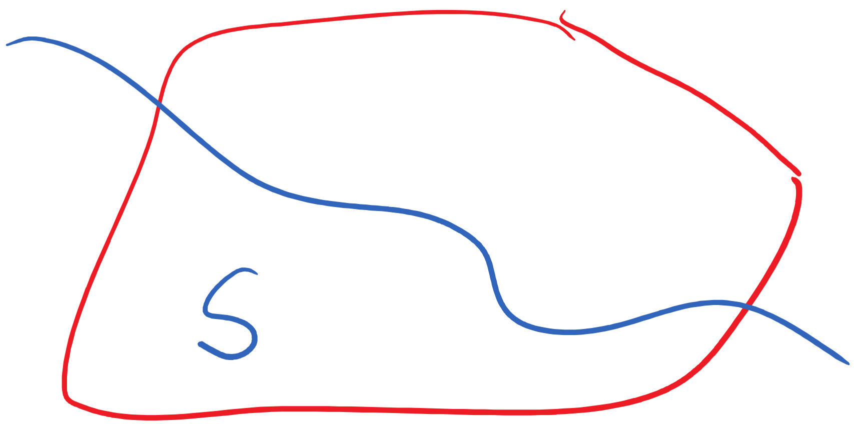
- Thm. The volume of a well-rounded convex body K can be estimated using $O^*(n^3)$ membership queries.

CV algorithm: $\frac{\sqrt{n}}{\sigma} \times \frac{\sqrt{n}}{\sigma} \times \sigma^2 n^2 \times \log n = O^*(n^3)$

- For a large class of discrete problems, adaptive annealing improves the reduction from counting to sampling [Stefankovic-V.-Vigoda].

Preguntas Abiertas: Geometría

- How true is the KLS conjecture?



$$\inf_{|S|\le\frac{|K|}{2}} \frac{|\partial S|}{|S|} > \frac{1}{\sqrt{\lambda_1(A)}}$$

- KLS Thm: $\psi = O\left(\sqrt{\sum_i \lambda_i\,(A)}\right) = O\left(\sqrt{Tr(A)}\right)$.

- A ~~conjecture~~ Theorem [Lee-V.16]

$$\psi \le C\left(\sum_i \lambda_i(A)^2\right)^{1/4} = C\left(Tr\left(A^2\right)\right)^{1/4}.$$

Preguntas Abiertas: Probabilidad

Q1. Does ball walk mix rapidly starting at a single nice point, e.g., the centroid?

Q2. When to stop? How to check convergence to stationarity on the fly? Does it suffice to check that the measures of all halfspaces have converged?

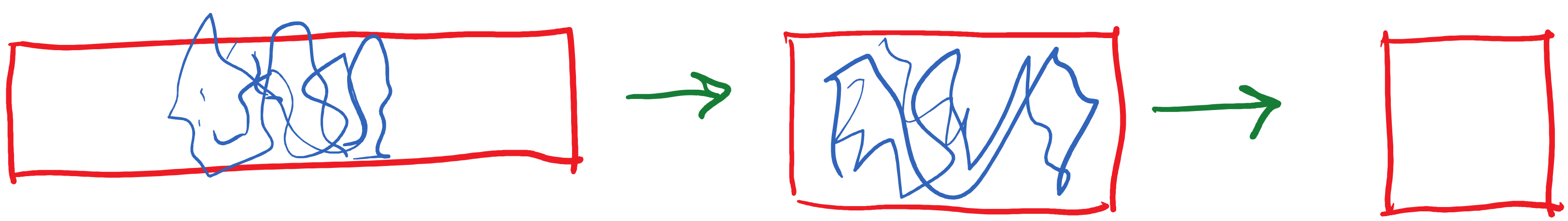
(Note: poly(n) sample can estimate all halfspace measures approximately)

Preguntas Abiertas: Probabilidad/Algoritmos

- Faster isotropy/rounding?
  - How to get information before reaching stationary?

To make isotropic:

run for N steps; transform using covariance; repeat.



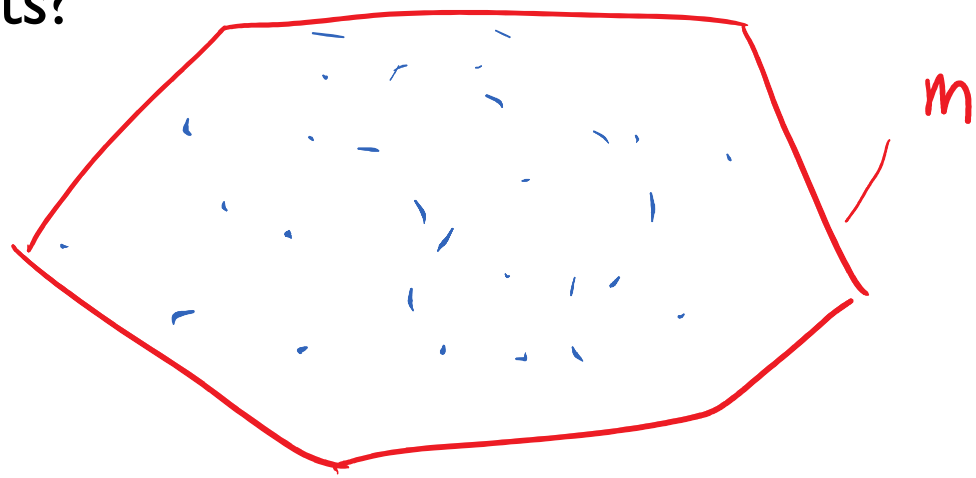
Preguntas Abiertas: Algoritmos

- Analyze Coordinate Hit-and-Run: saves a factor of n in implementing each step compared to ball walk or hit-and-run.

- Other walks?

Preguntas Abiertas: Algoritmos

- How efficiently can we learn a polytope P given only random points?



- With O(mn) points, cannot "see" structure, but enough information to estimate the polytope! Algorithms?
- For convex bodies:
  - [KOS][GR] need $2^{\Omega(\sqrt{n})}$ points to learn P
  - [Eldan] need 2^{n^c} even to estimate the volume of P

Preguntas Abiertas: Algoritmos

- Can we estimate the volume of an explicit polytope in *deterministic* polynomial time?

$$Ax \leq b$$