Reducing Isotropy to KLS: An Almost Cubic Volume Algorithm

Santosh Vempala

with He Jia, Aditi Laddha and Yin Tat Lee

The Volume Problem

Given a measurable, compact set K in n-dimensional space, find a number A such that:

$$(1-\epsilon)\,\mathrm{vol}(K) \;\leq\; A \;\leq\; (1+\epsilon)\,\mathrm{vol}(K)$$

K is given by

- a point $x_0 \in K$, s.t. $x_0 + B_n \subseteq K \subseteq RB_n$
- a membership oracle: answers YES/NO to "$x \in K?$"

Volume: first attempt

- Divide and conquer:

- Difficulty: number of parts grows exponentially in n.

Volume: second attempt: Sandwiching

Thm (John). Maximum volume ellipsoid E contained in convex body K satisfies

$$E \subseteq K \subseteq nE.$$

K is in *isotropic position* if $E_K(x) = 0$ and $E_K(xx^T) = I$.

Thm (KLS95). For a convex body K in *isotropic position*,

$$\sqrt{\frac{n+1}{n}}B_n \subseteq K \subseteq \sqrt{n(n+1)}B_n$$

- Both are factor n sandwichings, with different ellipsoids.

Volume via Sandwiching

- The John ellipsoid can be approximated using the Ellipsoid algorithm, s.t.

$$E \subseteq K \subseteq n^{1.5} E$$

- The Inertial ellipsoid can be approximated to within any constant factor (we'll see how)
- Using either one,

$$E \subseteq K \subseteq n^{O(1)} E \ \Rightarrow vol(E) \leq vol(K) \leq n^{O(n)} \ vol(E).$$

- Polytime algorithm, $n^{O(n)}$ approximation
- Can we do better?

Complexity of Volume Estimation

Thm [E86, BF87]. For any deterministic algorithm that uses at most n^a membership calls to the oracle for a convex body K and computes two numbers A and B such that $\mathrm{A} \leq \mathrm{vol(K)} \leq \mathrm{B}$, there is some convex body for which the ratio B/A is at least

$$\left(\frac{cn}{a \log n}\right)^{\frac{\mathrm{n}}{2}}$$

where c is an absolute constant.

Thm [DF88]. Computing the volume of an explicit polytope $Ax \leq b$ is #P-hard, even for a totally unimodular matrix A and rational b.

Complexity of Volume Estimation

Thm [BF]. For deterministic algorithms:

# oracle calls	approximation lower bound
n^a	$\left(\frac{cn}{a \log n}\right)^{n/2}$
$\left(\frac{1}{\varepsilon}\right)^n$	$(1+\varepsilon)^n$

Thm [Dadush-V.13].

Matching upper bound of $(1+\epsilon)^n$ in time $\left(\frac{1}{\epsilon}\right)^{O(n)}$ poly(n).

Randomized Volume Computation

[Dyer-Frieze-Kannan1989]

Thm. There is a polynomial-time randomized algorithm that estimates the volume of a convex body to within relative error $(1+\epsilon)$ with probability at least $1-\delta$ in time poly$(n, \frac{1}{\epsilon}, \log\frac{R}{\delta})$.

Volume Computation: an ongoing adventure

	Exponent	New aspects
Dyer-Frieze-Kannan 89	23	everything
Lovász-Simonovits 90	16	localization
Applegate-K 90	10	logconcave integration
L 90	10	ball walk
DF 91	8	error analysis
LS 93	7	multiple improvements
KLS 97	5	speedy walk, isotropy
LV 03,04	4	annealing, wt. isoper.
LV 06	4	integration, local analysis
Cousins-V. 15 (well-rounded)	3	Gaussian cooling

Volume: third attempt: Sampling

- Pick random samples from ball/cube containing K.
- Compute fraction c of sample in K.
- Output c.vol(outer ball).

- Need too many samples!

Volume via Sampling [DFK89]

$$B \subseteq K \subseteq RB.$$

Let $K_i = K \cap 2^{i/n} B,\ \ i = 0, 1, \ldots, m = n \log R.$

$$vol(K) = vol(B). \frac{vol(K_1)}{vol(K_0)} \cdot \frac{vol(K_2)}{vol(K_1)} \ldots \frac{vol(K_m)}{vol(K_{m-1})}.$$

Estimate each ratio with random samples.

Volume via Sampling

$$K_i = K \cap 2^{i/n} B, \qquad i = 0, 1, \ldots, m = n \log R.$$

$$vol(K) = vol(B). \frac{vol(K_1)}{vol(K_0)} \frac{vol(K_2)}{vol(K_1)} \ldots \frac{vol(K_m)}{vol(K_{m-1})}.$$

Claim. $\mathrm{vol}(K_{i+1}) \le 2. \mathrm{vol}(K_i)$.

#samples for $(1 + \epsilon)$ approximation $= m. \frac{m}{\epsilon^2} = O^*\left(n^2\right)$.

But, how to sample?

Sampling

Input: function f: $R^n \rightarrow R_+$ specified by an oracle,

point x, error parameter ε.

Output: A point y from a distribution within distance ε of distribution with density proportional to f.

- Any logconcave density can be sampled in polytime.
- Many applications: Rounding, Optimization, Integration, Learning

Rounding via Sampling

1. Sample m random points from K
2. Compute sample mean and sample covariance matrix
   - $z = E(x) \qquad A = E((x-z)(x-z)^T).$
3. Output $B = A^{-\frac{1}{2}}$.

Then, $B(K-z)$ is nearly isotropic.

Thm. [Adamczak-Litvak-Pajor-TomczakJaegermann09]
For isotropic K, estimate $\hat{A}$ with $C(\epsilon) \cdot n$ random points satisfies

$$E(\|A - I\|_2) \leq \epsilon.$$

I.e., for any unit vector v, $\quad 1 - \epsilon \leq E\left(\left(v^T x\right)^2\right) \leq 1 + \epsilon.$

How to Sample?

Ball walk:

At x,

-pick random y from $x + \delta B_n$

-if y is in K, go to y

Hit-and-Run:

At x,

-pick a random chord L through x

-go to a random point y on L

Convergence depends on isoperimetry

- Technique: show that "conductance" of Markov chain is large.
  - (one-step overlap): Nearby points have overlapping one-step distributions
  - (isoperimetry) Large subsets have large boundaries

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

Isoperimetry and the KLS conjecture

$$\psi_K = \sup_{S:|S|\le\frac{|K|}{2}} \frac{vol_n(S)}{vol_{n-1}(\partial S)}$$

Thm. [KLS97]. Mixing rate of the ball walk in K from a warm start is $n^2\psi_K^2$.

$$A = E_K((x-\overline{x})(x-\bar{x})^T)$$

note: $\sum_i \lambda_i(A) = E_K\left(\|x-\bar{x}\|^2\right)$

		isotropic ($A = I$)	mixing time
Thm. [KLS95].	$\psi_K \le c\sqrt{\sum_i \lambda_i(A)}$	$O(\sqrt{n})$	n^3
Conj. [KLS95].	$\psi_K \le c\sqrt{\lambda_1(A)}$	$O(1)$	n^2
Thm. [Lee-V.17].	$\psi_K \le c\left(\sum_i {\lambda_i}^2(A)\right)^{1/4}$	$O(n^{1/4})$	$n^{2.5}$
Thm. [Chen20].	$\psi_K \le n^{o(1)} \cdot \sqrt{\lambda_1(A)}$	$O(n^{o(1)})$	$n^{2+o(1)}$

The Conjecture

For any logconcave density in any dimension, halfspaces minimize the isoperimetric ratio up to a universal constant.

Covariance of p: $A = \mathbb{E}_p((x-\bar{x})\left(x-\bar{x})^T\right)$.

$$\lambda_1(A) = \max_{\|v\|=1} v^T A v$$

Isoperimetry of halfspace cuts: project distribution to normal of halfspace, to get logconcave distribution with variance $\sigma_v^2 = v^T A v$. Then,

$$\psi(\text{halfspaces}) \lesssim \sigma_v \lesssim \sqrt{\lambda_1(A)}.$$

Sampling and KLS

Thm. [KLS] For any logconcave density q in $\boldsymbol{R}^n$ with KLS constant ψ_q, the mixing time of the ball walk from a warm start is $n^2\psi_q^2$.

- $n^{2+o(1)}$ for isotropic
- $n^2 \cdot Tr(A) = \tilde{O}(n^3)$ for a well-rounded body.

- KLS volume algorithm is $n \cdot n \cdot n^3 = O^*(n^5)$.

Simulated Annealing [LV03, Kalai-V.04]

To estimate $\int f$ consider a sequence $f_0, f_1, f_2, \ldots, f = f_m$ with $\int f_0$ being easy, e.g., constant function over ball.

Then, $\int f = \int f_0 . \frac{\int f_1}{\int f_0} . \frac{\int f_2}{\int f_1} \ldots \frac{\int f_m}{\int f_{m-1}}$.

Each ratio can be estimated by sampling:

1. Sample X with density proportional to f_i
2. Compute $Y = \frac{f_{i+1}(X)}{f_i(X)}$

Then, $E(Y) = \int \frac{f_{i+1}(X)}{f_i(X)} . \frac{f_i(X)}{\int f_i(X)} dX = \frac{\int f_{i+1}}{\int f_i}$.

Annealing [LV06]

- Define: $f_i(X) = e^{-a_i||X||}$
- $a_0 = 2R, \quad a_{i+1} = a_i / \left(1 + \frac{1}{\sqrt{n}}\right), \quad a_m = \frac{\epsilon}{2R}$
- $m \sim \sqrt{n} \log(2R/\epsilon)$ phases
- $f_0 . \frac{\int f_1}{\int f_0} . \frac{\int f_2}{\int f_1} \ldots \frac{\int f_m}{\int f_{m-1}}$.

Lemma. $VAR\left(Y = \frac{f_{i+1}(X)}{f_i(X)}\right) < 4\, E(Y)^2$.

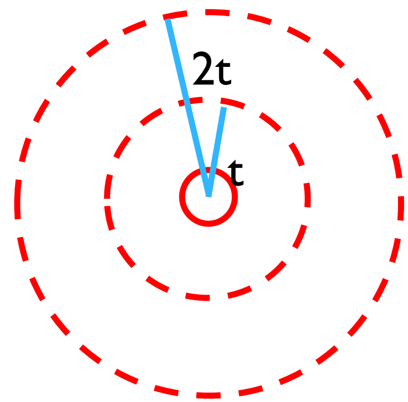
- Although expectation of Y can be large (exponential even), we need only a few samples to estimate it!

LoVe algorithm: $\sqrt{n} \times \sqrt{n} \times n^3 = n^4$

LV rounding algorithm

- Volume algorithm assumes body is near-isotropic
- But how to round? (make nearly isotropic)
- A sequence of doubling balls intersected with K:

$$K_i = K \cap 2^i B_n$$



Lemma. K_i isotropic $\Rightarrow K_{i+1}$ well-rounded, i.e., $\mathrm{E}\|x\|^2 = O(n)$.

- Repeat: use $O(n)$ samples to make K_{i+1} isotropic.
- Complexity: $\log n \cdot n \cdot n^2 \left(\sqrt{n}\right)^2 = \tilde{O}(n^4)$.

Gaussian Cooling [Cousins-V.2015]

- $f_i(X) = e^{-\frac{||X||^2}{2\sigma_i^2}}$
- $\sigma_0^2 = \frac{1}{n}, \sigma_m^2 = O(n).$

- Estimate $\frac{\int f_{i+1}}{\int f_i}$ using samples drawn according to f_i

Why Gaussian?

- KLS conjecture holds for Gaussian restricted to any convex body (Bobkov, Ledoux).

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

Thm. [Cousins-V. 13]. Ball walk applied to Gaussian $N(0, \sigma^2 I_n)$ restricted to convex body K containing the unit ball mixes in $n^2 \max\{1, \sigma^2\}$ steps from a warm start.

Gaussian Cooling

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

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

- Sampling time: $n^2 \max\{1, \sigma^2\} = n^2$
- #phases, #samples per phase: $\sqrt{n}$

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

Gaussian Cooling

- $f_i(X) = e^{-\frac{||X||^2}{2\sigma_i^2}}$
- For $\sigma_i^2 > 1$, we set $\sigma_i^2 = \sigma_{i-1}^2\left(1 + \frac{\sigma_{i-1}}{\sqrt{n}}\right)$
- For this accelerated schedule, we need K to be well-rounded

- Sampling time: $n^2 \max\{1, \sigma^2\} = \sigma^2 n^2$ (too much??)
- #phases to double σ is $\frac{\sqrt{n}}{\sigma}$, so #samples per phase is also $\frac{\sqrt{n}}{\sigma}$

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

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

Rounding and KLS?

- Can we round (isotropicize) faster than n^4 ?

Main new result:

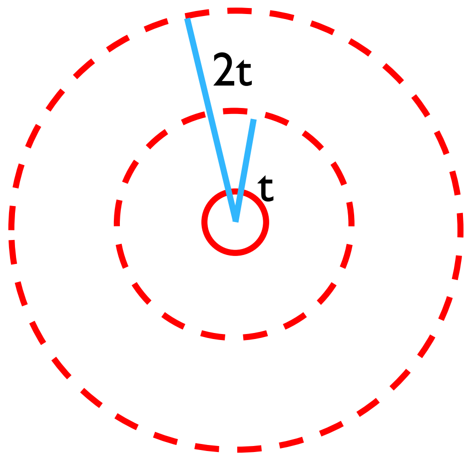
Thm. Any convex body can be brought into near-isotropic position using $\tilde{O}\left(n^3\psi_n^2\right)$ membership queries.

Cor. Volume of a convex body can be computed in $O^*\left(n^3\psi^2\right)$.

- $n^{3+o(1)}$ with current bound, n^3 if KLS holds.

Well-rounded → Isotropic

- Outer loop: For $i = 1$ to $\log R$, make $K_i = K \cap 2^i B_n$ isotropic.



- Inner loop: set $r = 1, k = \tilde{O}(r^2)$.
- While $r^2 \log n < n$,
  - Sample k points from K_i. Estimate covariance $\tilde{A}$.
  - Let V be subspace with "large" eigenvalues, $\lambda(\tilde{A}) > n$.
  - Scale up along all directions in $V^\perp$ by a factor of 2.

Well-rounded → Isotropic

Sample k points from K_i. Estimate covariance $\tilde{A}$.
Let V be subspace with "large" eigenvalues, $\lambda(\tilde{A}) > n$.
Scale up along all directions in $V^\perp$ by a factor of 2.

- Naïve algorithm: estimate covariance; scale up small eigenvalue directions.
- Difficulty: needs n samples, each takes n^3

Idea: use coarse estimates and gradually refine.

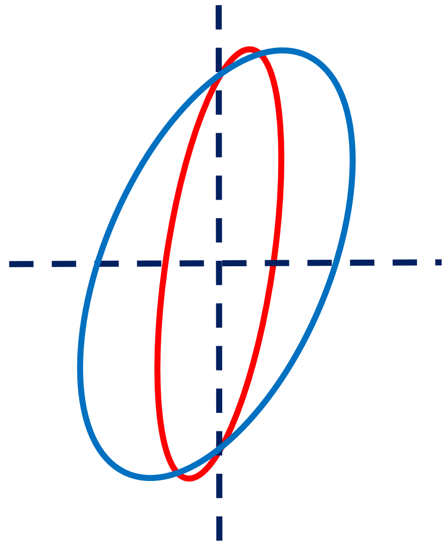
- Initially, estimate *very* large eigenvalues, scale up the rest
- Gradually estimate smaller eigenvalues.

Why is this better?

- When we get finer estimates, body is more isotropic and sampling is faster!

How to get round

- Find "big" directions using a few samples
- Scale up complementary subspace



- Repeat with more samples
- Threshold for "big" stays the same

Getting rounder, efficiently

- Sampling complexity is $n^2 \cdot \frac{\psi^2}{r^2} \le n^2 \cdot \frac{Tr(A_j)}{r^2}$

- But can be controlled better!

Thm. If $\psi_n = O(n^{1/p})$, then $\psi_q = \tilde{O}(\|A\|_p^{1/2})$.

- So sampling can be done in $n^2 \cdot \frac{\|A_j\|_p}{r_j^2}$

Getting rounder, efficiently

Lemma 1.

- $Tr(A_0) \leq Cn,\ Tr(A_j) \leq cr_j^2 \cdot Tr(A_0)$
- $\|A_j\|_p = \left(\sum_i \lambda_i^p(A_j)\right)^{1/p} \leq c \log n \cdot nr_j^{2/p}$

Lemma 2. Assume $n \geq 4r^2 \log n$. In each iteration, radius of inner ball, r, increases by a factor of $2\left(1 - \frac{1}{\log n}\right)$.

Lemma 3. In j'th phase, $k = \tilde{O}(r_j^2)$ samples suffice.

Therefore, complexity is

$$r_j^2 \cdot n^2 \cdot \frac{\|A_j\|_p}{r_j^2} = \tilde{O}\left(n^3 r_j^{2/p}\right) = \tilde{O}\left(n^{3+1/p}\right) = \tilde{O}\left(n^3 \psi^2\right)$$

To round, scale up skinny subspace

Lemma 1.

- $Tr(A_0) \leq Cn, \;\; Tr\left(A_j\right) \leq c r_j^2 Tr(A_0)$
- $\left\|A_j\right\|_p^p = \sum_i \lambda_i^p(A_j) \leq c \log n \cdot n r_j^{2/p}$

$$\sum_i \lambda_i^p(A_{j+1}) \leq \sum_{i:\lambda_i \leq \lambda} (2\lambda)^p + \sum_{i:\lambda_i > \lambda} \lambda_i^p(A_j)$$

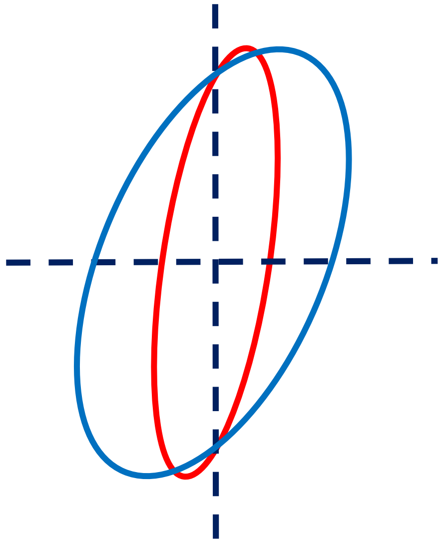
$$\leq \frac{Tr\left(A_j\right)}{\lambda} \cdot (2\lambda)^p + \sum_{i:\lambda_i > \lambda} \lambda_i^p(A_j)$$

Setting $\lambda = n$, in each iteration, increase in $\left\|A_j\right\|_p$ is at most

$$\left(\frac{r_j^2 n}{n} \cdot (2n)^p\right)^{1/p} = 2n r_j^{2/p}$$

Grow your inner ball

Lemma. While $\lambda \geq 4r^2 \log n$, r increases by a factor of at least $2\left(1 - \frac{1}{\log n}\right)$ in each iteration. (We use $\lambda = n$).



Proof sketch:

Scale up all directions with variance $< \lambda$.

V contains ellipsoid with minimum axis length λ

$V^\perp$ contains a ball of radius r that is scaled up by 2.

Then, new body contains a ball of radius nearly $2r$.

$x = \alpha y + (1-\alpha)$ where $\alpha \in [0,1]$, $y \in \partial B(2r) \cap V^\perp$, $z \in \partial B^n(\sqrt{\lambda}) \cap V$

Then,

$$\|x\|^2 = \alpha^2 4r^2 + (1-\alpha)^2 \lambda \geq \frac{4\lambda r^2}{\lambda + 4r^2} \geq 4 \cdot \frac{\log n}{\log n + 1} \cdot r^2$$

Matrix, Chernoff, and all that

Lemma 2. $k = \tilde{O}(r_j^2)$ samples suffice.

Lem. [Matrix-Chernoff].

A: covariance, $\hat{A}$: empirical covariance of k samples. Then,

$$-\varepsilon A - c \log^3 n \cdot \frac{Tr(A)}{\varepsilon k} I \preccurlyeq \hat{A} - A \preccurlyeq \varepsilon A + c \log^3 n \cdot \frac{Tr(A)}{\varepsilon k} I.$$

So with $\varepsilon = \frac{1}{2}$ and $k = 2c \cdot Tr(A) \cdot \frac{\log^3 n}{n}$,

$$\frac{1}{2} A - nI \preccurlyeq \hat{A} \preccurlyeq \frac{1}{2} A + nI$$

Suffices to detect eigenvalues larger than $4n$.

Localizing isoperimetry to the isotropic case

- Thm [LV16]. $\psi_q = O(\|A\|_F^{1/2})$ which is $O(n^{1/4})$ for isotropic.
- Thm[LV16]. Suppose $\psi_n = O(n^\beta)$, for some $0 < \beta \leq 1/2$. Then, for any logconcave density q in $\boldsymbol{R}^n$ with covariance A, we have

$$\psi_q \lesssim \frac{\sqrt{\log n}}{\beta} \|A\|_{1/2\beta}^{1/2}$$

- $\psi_n^2 = \Theta\left(n^{1/p}\right) \Rightarrow \psi_q^2 = O\left(\|A\|_p \log n\right)$
- Complexity is $n^{3+(1/p)} = n^3 \psi_n^2$.
- Thm [Chen20]. $\psi_n^2 = 2^{O\left(\sqrt{\log n \log\log n}\right)} = n^{o(1)}$.
- Thm. Complexity of Rounding/Volume is $\tilde{O}\left(n^{3+o(1)}\right)$.

Open question

- How true is the KLS conjecture?

Open question

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

$$Ax \leq b$$

Thank you!