Sampling s-concave functions

Karthekeyan Chandrasekaran
Amit Deshpande
Santosh Vempala

Overview

- Setting
- Definition: s-concave functions
- Sampling algorithm - ball walk
- Earlier work
- Our results
- Isoperimetry

Setting

- Given: Oracle access to a function $f : \mathbb{R}^n \to \mathbb{R}_+$ with $supp(f) = K$, a convex body.

Setting

- Given: Oracle access to a function $f : \mathbb{R}^n \to \mathbb{R}_+$ with $supp(f) = K$, a convex body.
- Objective: Sample points distributed according to the distribution function π_f.

$$\pi_f(A) = \frac{\int_{x \in A} f(x)dx}{\int_{x \in K} f(x)dx} \qquad \forall A \subseteq K$$

Setting

- Given: Oracle access to a function $f : \mathbb{R}^n \to \mathbb{R}_+$ with $supp(f) = K$, a convex body.
- Objective: Sample points distributed according to the distribution function π_f.

$$\pi_f(A) = \frac{\int_{x \in A} f(x)dx}{\int_{x \in K} f(x)dx} \qquad \forall A \subseteq K$$

- Sample points from a distribution σ such that $d_{TV}(\sigma, \pi_f) \leq \epsilon$.

Setting

- Given: Oracle access to a function $f : \mathbb{R}^n \rightarrow \mathbb{R}_+$ with $supp(f) = K$, a convex body.
- Objective: Sample points distributed according to the distribution function π_f.

$$\pi_f(A) = \frac{\int_{x \in A} f(x)dx}{\int_{x \in K} f(x)dx} \qquad \forall A \subseteq K$$

- Sample points from a distribution σ such that $d_{TV}(\sigma, \pi_f) \leq \epsilon$.
  - **Total Variation Distance**: For probability distributions σ, π_f defined over K,

$$d_{TV}(\sigma, \pi_f) = \sup_{A \subseteq K} |\sigma(A) - \pi_f(A)|$$

Setting

- Given: Oracle access to a function $f : \mathbb{R}^n \rightarrow \mathbb{R}_+$ with $supp(f) = K$, a convex body.
- Objective: Sample points distributed according to the distribution function π_f.

$$\pi_f(A) = \frac{\int_{x \in A} f(x)dx}{\int_{x \in K} f(x)dx} \qquad \forall A \subseteq K$$

- Sample points from a distribution σ such that $d_{TV}(\sigma, \pi_f) \leq \epsilon$.
  - **Total Variation Distance**: For probability distributions σ, π_f defined over K,

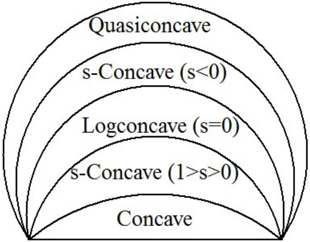
$$d_{TV}(\sigma, \pi_f) = \sup_{A \subseteq K} |\sigma(A) - \pi_f(A)|$$

- Sampling complexity = Number of function evaluations.

Definition: s-concave functions

- $f(x)^s$ is concave.
- $\forall \lambda \in [0,1], \forall x, y \in \mathbb{R}^n$

$$f(\lambda x + (1-\lambda)y) \geq (\lambda f(x)^s + (1-\lambda) f(y)^s)^{\frac{1}{s}}.$$



$f(x) = 1/\|x\|$

$f(x) = e^{-\|x\|^2}$

$f(x) = \|x\|^2$

$f(x) = \|x\|$

Ball Walk

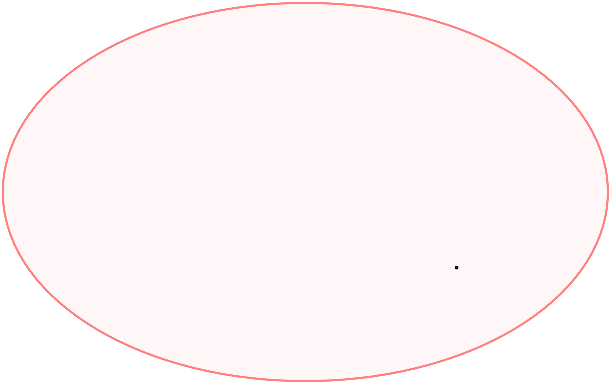
Ball Walk Step(x,r)

- Pick a uniform random point y from the ball of radius r centered at current point x.
- Move to y with probability $\min\{\frac{f(y)}{f(x)}, 1\}$; Stay at x with remaining probability.

Ball Walk

Ball Walk Step(x,r)

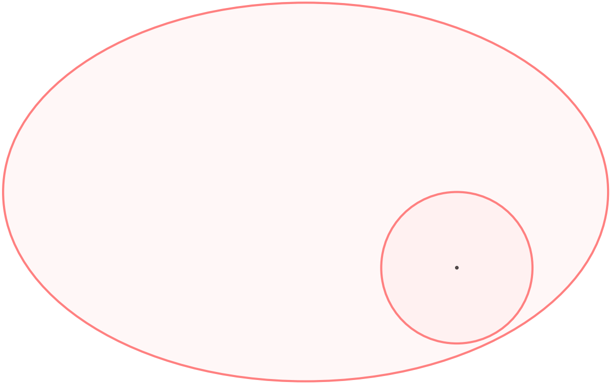
- Pick a uniform random point y from the ball of radius r centered at current point x.
- Move to y with probability $\min\{\frac{f(y)}{f(x)}, 1\}$; Stay at x with remaining probability.



Ball Walk

Ball Walk Step(x,r)

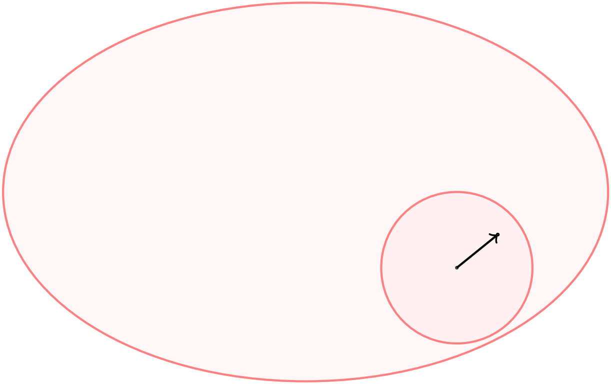
- Pick a uniform random point y from the ball of radius r centered at current point x.
- Move to y with probability $\min\{\frac{f(y)}{f(x)}, 1\}$; Stay at x with remaining probability.



Ball Walk

Ball Walk Step(x,r)

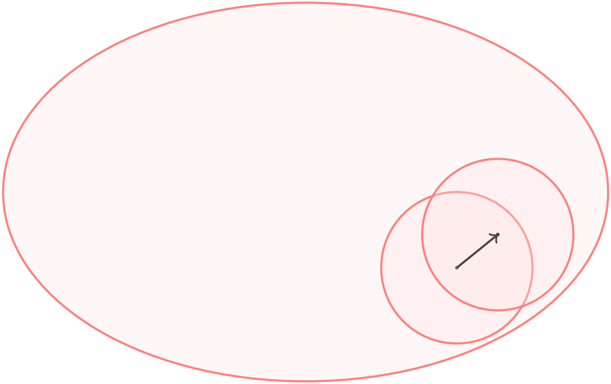
- Pick a uniform random point y from the ball of radius r centered at current point x.
- Move to y with probability $\min\{\frac{f(y)}{f(x)}, 1\}$; Stay at x with remaining probability.



Ball Walk

Ball Walk Step(x,r)

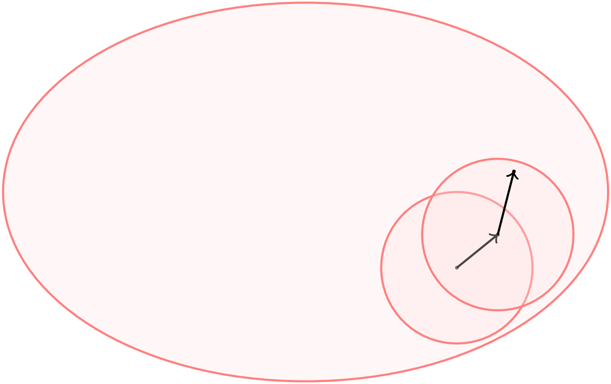
- Pick a uniform random point y from the ball of radius r centered at current point x.
- Move to y with probability $\min\{\frac{f(y)}{f(x)}, 1\}$; Stay at x with remaining probability.



Ball Walk

Ball Walk Step(x,r)

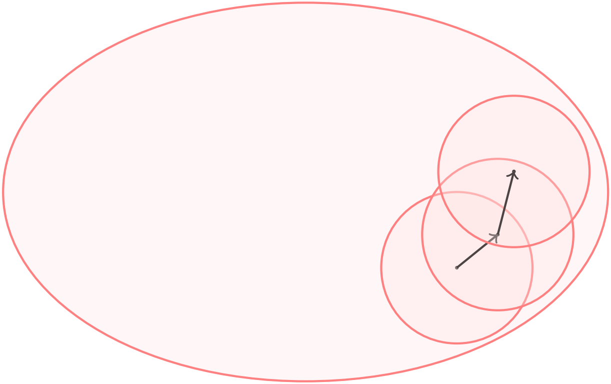
- Pick a uniform random point y from the ball of radius r centered at current point x.
- Move to y with probability $\min\{\frac{f(y)}{f(x)}, 1\}$; Stay at x with remaining probability.



Ball Walk

Ball Walk Step(x,r)

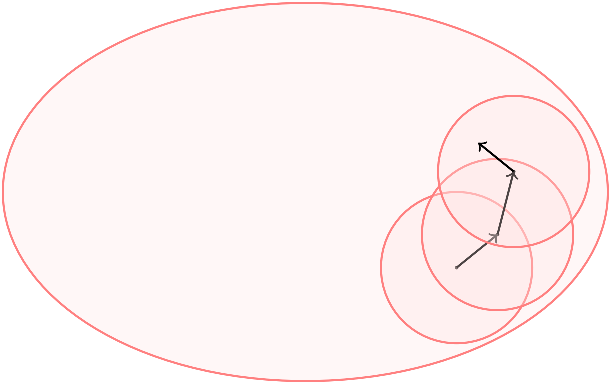
- Pick a uniform random point y from the ball of radius r centered at current point x.
- Move to y with probability $\min\{\frac{f(y)}{f(x)}, 1\}$; Stay at x with remaining probability.



Ball Walk

Ball Walk Step(x,r)

- Pick a uniform random point y from the ball of radius r centered at current point x.
- Move to y with probability $\min\{\frac{f(y)}{f(x)}, 1\}$; Stay at x with remaining probability.



Sampling Algorithm Overview

M-**warm distribution**: Distribution σ_0 on K such that

$$\forall A \subseteq K, \sigma_0(A) \leq M\pi_f(A)$$

INPUT:

- s-concave function oracle,
- x picked according to M-warm distribution and
- $\epsilon > 0$.

Sampling Algorithm Overview

M-**warm distribution**: Distribution σ_0 on K such that

$$\forall A \subseteq K, \sigma_0(A) \leq M\pi_f(A)$$

INPUT:

- s-concave function oracle,
- x picked according to M-warm distribution and
- $\epsilon > 0$.

OUTPUT: Point x from a distribution σ_m such that $d_{TV}(\sigma_m, \pi_f) \leq \epsilon$.

Sampling Algorithm Overview

M-**warm distribution**: Distribution σ_0 on K such that

$$\forall A \subseteq K, \sigma_0(A) \leq M\pi_f(A)$$

INPUT:

- s-concave function oracle,
- x picked according to M-warm distribution and
- $\epsilon > 0$.

OUTPUT: Point x from a distribution σ_m such that $d_{TV}(\sigma_m, \pi_f) \leq \epsilon$.
ALGORITHM: Run ball walk for m steps starting from x.

Sampling Algorithm Overview

M-**warm distribution**: Distribution σ_0 on K such that

$$\forall A \subseteq K, \sigma_0(A) \leq M\pi_f(A)$$

INPUT:

- s-concave function oracle,
- x picked according to M-warm distribution and
- $\epsilon > 0$.

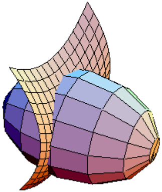
OUTPUT: Point x from a distribution σ_m such that $d_{TV}(\sigma_m, \pi_f) \leq \epsilon$.
ALGORITHM: Run ball walk for m steps starting from x.
Mixing time: m such that $d_{TV}(\sigma_m, \pi_f) \leq \epsilon$.

Conductance and Isoperimetry

- LS'93: Large conductance $\implies$ Fast mixing time.

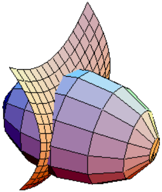
Conductance and Isoperimetry

- LS'93: Large conductance $\implies$ Fast mixing time.
- Isoperimetry:



Conductance and Isoperimetry

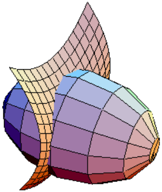
- LS'93: Large conductance $\implies$ Fast mixing time.
- Isoperimetry:



$$\pi_f(\delta S) \geq \phi_f \min\{\pi_f(S), \pi_f(K \backslash S)\}$$

Conductance and Isoperimetry

- LS'93: Large conductance $\implies$ Fast mixing time.
- Isoperimetry:

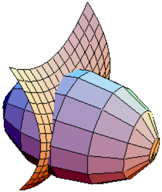


$$\pi_f(\delta S) \geq \phi_f \min\{\pi_f(S), \pi_f(K \backslash S)\}$$

- Good isoperimetry necessary for good conductance.

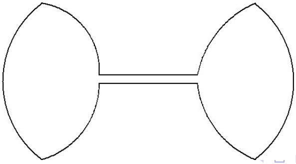
Conductance and Isoperimetry

- LS'93: Large conductance $\implies$ Fast mixing time.
- Isoperimetry:



$$\pi_f(\delta S) \geq \phi_f \min\{\pi_f(S), \pi_f(K \backslash S)\}$$

- Good isoperimetry necessary for good conductance.



Earlier work

- Uniform density over a convex body K (DFK'91 $\rightarrow$ LV'06).

$$m = O^*\left(\frac{n^2 D^2}{\epsilon^2}\right)$$

D: diameter of K.

Earlier work

- Uniform density over a convex body K (DFK'91 $\rightarrow$ LV'06).

$$m = O^*\left(\frac{n^2 D^2}{\epsilon^2}\right)$$

  D: diameter of K.

- Logconcave function $f : \mathbb{R}^n \rightarrow \mathbb{R}_+$ (AK'90 $\rightarrow$ LV'07).

$$m = O^*\left(\frac{n^2 Var(f)}{\epsilon^4}\right).$$

Earlier work

- Uniform density over a convex body K (DFK'91 $\rightarrow$ LV'06).

$$m = O^*\left(\frac{n^2 D^2}{\epsilon^2}\right)$$

D: diameter of K.

- Logconcave function $f : \mathbb{R}^n \rightarrow \mathbb{R}_+$ (AK'90 $\rightarrow$ LV'07).

$$m = O^*\left(\frac{n^2 Var(f)}{\epsilon^4}\right).$$

- For near isotropic convex bodies/logconcave-distributions, mixing time is $O^*(n^3)$ (LV'07).

Our Results 1/4

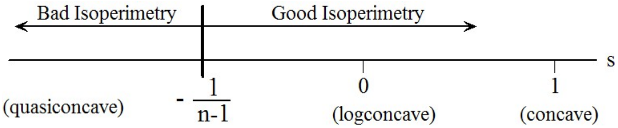
- Thm1 [ISOPERIMETRY]. Isoperimetry for s-concave functions for $s \geq -1/(n-1)$:
  For every partition S_1, S_2, S_3 of K,

$$\pi_f(S_3) \geq \frac{d(S_1, S_2)}{Diam} \min\left\{\pi_f(S_1), \pi_f(S_2)\right\}.$$

Our Results 2/4

- Thm2 [LOWER BOUND]. $\forall \epsilon > 0$, there exists a $\left(-\frac{1}{n-1-\epsilon}\right)$-concave function $f : \mathbb{R}^n \to \mathbb{R}_+$ and a partition of $supp(f)$ into S and T such that

$$\frac{\pi_f(\partial S)}{\min\{\pi_f(S), \pi_f(T)\}} \leq Cn(1+\epsilon)^{-\epsilon n}$$



Our Results 3/4

Definition: Function $f : \mathbb{R}^n \to \mathbb{R}_+$ is (α, δ)-**smooth** if $\max\left\{\frac{f(x)}{f(y)}, \frac{f(y)}{f(x)}\right\} \leq \alpha$ for all $x, y \in K$, $||x - y|| \leq \delta$.

Our Results 3/4

Definition: Function $f : \mathbb{R}^n \to \mathbb{R}_+$ is (α, δ)-**smooth** if $\max\left\{\frac{f(x)}{f(y)}, \frac{f(y)}{f(x)}\right\} \le \alpha$ for all $x, y \in K$, $||x - y|| \le \delta$.

- Thm3 [SAMPLING]: (α, δ)-smooth s-concave function f with $supp(f) = K$, a convex body and $s \ge -1/(n-1)$:

$$m = O\left(\left(\frac{nD^2}{\delta^2}\log\frac{2M}{\epsilon}\right)\cdot\max\left\{\frac{nM^2}{\epsilon^2}, \frac{(\alpha^{-s}-1)^2}{s^2}\right\}\right)$$

Our Results 4/4

Our Results 4/4

Cauchy density function:

$$f(x) \propto \frac{\det(A)^{-1}}{\left(1 + \|A(x-m)\|^2\right)^{(n+1)/2}}.$$

Our Results 4/4

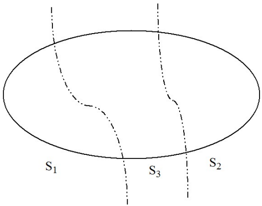
Cauchy density function:

$$f(x) \propto \frac{\det(A)^{-1}}{\left(1 + \|A(x-m)\|^2\right)^{(n+1)/2}}.$$

- Thm4 [SAMPLING CAUCHY]: Cauchy density function restricted to *any convex set* containing a ball of radius $||A^{-1}||_2$:

$$m = O\left(\frac{n^3 M^4}{\epsilon^4} \log \frac{2M}{\epsilon}\right)$$

Isoperimetry

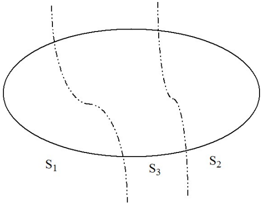


- Logconcave functions ($s \geq 0$) (DF'91):

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

where $d(S_1, S_2) = \min\{|u - v| : u \in S_1, v \in S_2\}$.

Isoperimetry



- Logconcave functions ($s \geq 0$) (DF'91):

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

where $d(S_1, S_2) = \min\{|u - v| : u \in S_1, v \in S_2\}$.

- s-concave functions ($s \geq -\frac{1}{n-1}$) (this talk):

$$\pi_f(S_3) \geq \frac{d(S_1, S_2)}{Diam} \min\{\pi_f(S_1), \pi_f(S_2)\}$$

Isoperimetry: Proof

Lemma 1: If $p : [0,1] \to \mathbb{R}_+$ is a unimodal function, then, for any partition of $[0,1]$ into Z_1, Z_2, Z_3,

$$\int_{Z_3} p(t)dt \geq d(Z_1, Z_2) \min \left\{ \int_{Z_1} p(t)dt, \int_{Z_2} p(t)dt \right\}.$$

Isoperimetry: Proof

Lemma 2. Suppose $f : \mathbb{R}^n \to \mathbb{R}_+$ is such that:

- For any affine line $L \subseteq \mathbb{R}^n$ and any linear function $l : K \cap L \to \mathbb{R}_+$, the function $h : K \cap L \to \mathbb{R}_+$,

$$h(x) := f(x) l(x)^{n-1}$$

  is unimodal.

Isoperimetry: Proof

Lemma 2. Suppose $f : \mathbb{R}^n \to \mathbb{R}_+$ is such that:

- For any affine line $L \subseteq \mathbb{R}^n$ and any linear function $l : K \cap L \to \mathbb{R}_+$, the function $h : K \cap L \to \mathbb{R}_+$,

$$h(x) := f(x)l(x)^{n-1}$$

  is unimodal.
- Then for every partition S_1, S_2, S_3 of K,

$$\pi_f(S_3) \geq \frac{d(S_1, S_2)}{Diam} \min \{\pi_f(S_1), \pi_f(S_2)\}.$$

Isoperimetry: Proof

Lemma 2. Suppose $f : \mathbb{R}^n \to \mathbb{R}_+$ is such that:

- For any affine line $L \subseteq \mathbb{R}^n$ and any linear function $l : K \cap L \to \mathbb{R}_+$, the function $h : K \cap L \to \mathbb{R}_+$,

$$h(x) := f(x) l(x)^{n-1}$$

  is unimodal.
- Then for every partition S_1, S_2, S_3 of K,

$$\pi_f(S_3) \geq \frac{d(S_1, S_2)}{Diam} \min \{\pi_f(S_1), \pi_f(S_2)\}.$$

Proof. By Localization Lemma [LS'93, KLS'95].

Isoperimetry: Proof

Lemma 2. Suppose $f : \mathbb{R}^n \to \mathbb{R}_+$ is such that:

- For any affine line $L \subseteq \mathbb{R}^n$ and any linear function $l : K \cap L \to \mathbb{R}_+$, the function $h : K \cap L \to \mathbb{R}_+$,

$$h(x) := f(x)l(x)^{n-1}$$

  is unimodal.

- Then for every partition S_1, S_2, S_3 of K,

$$\pi_f(S_3) \geq \frac{d(S_1, S_2)}{Diam} \min\left\{\pi_f(S_1), \pi_f(S_2)\right\}.$$

Proof. By Localization Lemma [LS'93, KLS'95].

- Localization Lemma reduces proving target inequality to proving similar isoperimetric inequality for h.

Isoperimetry: Proof

Lemma 2. Suppose $f : \mathbb{R}^n \to \mathbb{R}_+$ is such that:

- For any affine line $L \subseteq \mathbb{R}^n$ and any linear function $l : K \cap L \to \mathbb{R}_+$, the function $h : K \cap L \to \mathbb{R}_+$,

$$h(x) := f(x) l(x)^{n-1}$$

  is unimodal.
- Then for every partition S_1, S_2, S_3 of K,

$$\pi_f(S_3) \geq \frac{d(S_1, S_2)}{Diam} \min \left\{ \pi_f(S_1), \pi_f(S_2) \right\}.$$

Proof. By Localization Lemma [LS'93, KLS'95].

- Localization Lemma reduces proving target inequality to proving similar isoperimetric inequality for h.
- By Lemma 1, the required isoperimetric inequality holds for h.

Isoperimetry: Proof

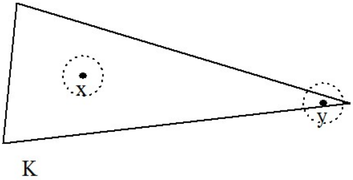
Lemma 3. If f is a $\left(-\frac{1}{n-1}\right)$-concave function, for any affine line $L \subseteq \mathbb{R}^n$ and any linear function $l : K \cap L \to \mathbb{R}_+$, the function $h : K \cap L \to \mathbb{R}_+$

$$h(x) := f(x)l(x)^{n-1}$$

is unimodal.

Sampling: Proof Issues

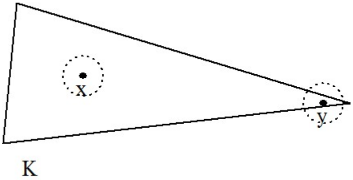
- Points of good local conductance.



  - Pick radius appropriately.

Sampling: Proof Issues

- Points of good local conductance.



  - Pick radius appropriately.
- Coupling: One-step distributions of geometrically close points overlap.

Further Work

- Remove the dependence on the smoothness parameter.
- How to perform warm start?
- Any better characterisation for isoperimetry?

Thank you!
Questions?