Sampling from a Convex Body

Based on

1. Ravi Kannan's notes.
2. Jonathan Kelner's lecture notes.
3. "Techniques in Optimization and Sampling" by Yin Tat Lee and Santosh Vempala.
4. "Algorithmic Convex Geometry" by Santosh Vempala.
5. "Geometric Random Walks" by Santosh Vempala.

Sampling from a convex body K

Given a convex body $K \subset \mathbb{R}^n$, output a uniform random point from K.

Sampling from a convex body K

Given a convex body $K \subset \mathbb{R}^n$, output a uniform random point from K.

- Oracle access to K. Given $x \in \mathbb{R}^n$, oracle answers in $O(1)$ time whether $x \in K$.

Sampling from a convex body K

Given a convex body $K \subset \mathbb{R}^n$, output a uniform random point from K.

- Oracle access to K. Given $x \in \mathbb{R}^n$, oracle answers in $O(1)$ time whether $x \in K$.
- Also need some point in K as input. W.l.o.g., assume that $0 \in K$.
- Also given $r, R > 0$ such that $\mathcal{B}(0, r) \subseteq K \subseteq \mathcal{B}(0, R)$. W.l.o.g. assume that $r = 1$.

Sampling from a convex body K

Given a convex body $K \subset \mathbb{R}^n$, output a uniform random point from K.

- Oracle access to K. Given $x \in \mathbb{R}^n$, oracle answers in $O(1)$ time whether $x \in K$.
- Also need some point in K as input. W.l.o.g., assume that $0 \in K$.
- Also given $r, R > 0$ such that $\mathcal{B}(0, r) \subseteq K \subseteq \mathcal{B}(0, R)$. W.l.o.g. assume that $r = 1$.

Has applications in volume estimation, optimization, etc.

Sampling from a convex body K

Given a convex body $K \subset \mathbb{R}^n$, output a uniform random point from K.

- Oracle access to K. Given $x \in \mathbb{R}^n$, oracle answers in $O(1)$ time whether $x \in K$.
- Also need some point in K as input. W.l.o.g., assume that $0 \in K$.
- Also given $r, R > 0$ such that $\mathcal{B}(0, r) \subseteq K \subseteq \mathcal{B}(0, R)$. W.l.o.g. assume that $r = 1$.

Has applications in volume estimation, optimization, etc.
Estimating volumes is a classical problem. Closed form formulae known for rectangular solids, simplices, spheres, etc.

Computing volumes

Theorem (Informal statement)

There is no deterministic polynomial algorithm that, given a membership oracle for K, computes $\operatorname{vol}(K)$ *to within a polynomial factor.*

Computing volumes

Theorem (Informal statement)

There is no deterministic polynomial algorithm that, given a membership oracle for K, computes $\operatorname{vol}(K)$ *to within a polynomial factor.*

Proof idea.

Consider an oracle that answers "yes" to any point in the unit ball and "no" to any point outside. After m "yes" answers, the convex body K could be anything between the ball and the convex hull of the m query points. The ratio of these volumes in exponential in n. $\square$

The power of randomness!

Theorem (Informal statement)

There exists a randomized algorithm that, given a membership oracle for K and a parameter ϵ, runs in time polynomial in $n, 1/\epsilon$ and $\log R$ and outputs an estimate A such that w.h.p. we have $(1-\epsilon)\mathrm{vol}(K) \leq A \leq (1+\epsilon)\mathrm{vol}(K)$.

The power of randomness!

Theorem (Informal statement)

There exists a randomized algorithm that, given a membership oracle for K and a parameter ϵ, runs in time polynomial in $n, 1/\epsilon$ and $\log R$ and outputs an estimate A such that w.h.p. we have $(1-\epsilon)\mathrm{vol}\,(K) \leq A \leq (1+\epsilon)\mathrm{vol}\,(K)$.

- Main idea: algorithm for "approximately uniform" sampling from K.

The power of randomness!

Theorem (Informal statement)

There exists a randomized algorithm that, given a membership oracle for K and a parameter ϵ, runs in time polynomial in $n, 1/\epsilon$ and $\log R$ and outputs an estimate A such that w.h.p. we have $(1-\epsilon)\mathrm{vol}(K) \leq A \leq (1+\epsilon)\mathrm{vol}(K)$.

- Main idea: algorithm for "approximately uniform" sampling from K.
- First proved by [Dyer, Frieze, Kannan - 89]. Many new developments since then leading to a rich theory of geometric random walks!

The power of randomness!

Theorem (Informal statement)

There exists a randomized algorithm that, given a membership oracle for K and a parameter ϵ, runs in time polynomial in $n, 1/\epsilon$ and $\log R$ and outputs an estimate A such that w.h.p. we have $(1-\epsilon)\mathrm{vol}(K) \leq A \leq (1+\epsilon)\mathrm{vol}(K)$.

- Main idea: algorithm for "approximately uniform" sampling from K.
- First proved by [Dyer, Frieze, Kannan - 89]. Many new developments since then leading to a rich theory of geometric random walks!

First attempt. Suppose $K \subseteq [-1/2, 1/2]^n$.

The power of randomness!

Theorem (Informal statement)

There exists a randomized algorithm that, given a membership oracle for K and a parameter ϵ, runs in time polynomial in $n, 1/\epsilon$ and $\log R$ and outputs an estimate A such that w.h.p. we have $(1-\epsilon)\text{vol}(K) \leq A \leq (1+\epsilon)\text{vol}(K)$.

- Main idea: algorithm for "approximately uniform" sampling from K.
- First proved by [Dyer, Frieze, Kannan - 89]. Many new developments since then leading to a rich theory of geometric random walks!

First attempt. Suppose $K \subseteq [-1/2, 1/2]^n$.

- Sample m uniform random points from $[-1/2, 1/2]^n$ and count how many belong to K. Let m' of them belong to K. Output m'/m.

The power of randomness!

Theorem (Informal statement)

There exists a randomized algorithm that, given a membership oracle for K and a parameter ϵ, runs in time polynomial in $n, 1/\epsilon$ and $\log R$ and outputs an estimate A such that w.h.p. we have $(1-\epsilon)\mathrm{vol}(K) \leq A \leq (1+\epsilon)\mathrm{vol}(K)$.

- Main idea: algorithm for "approximately uniform" sampling from K.
- First proved by [Dyer, Frieze, Kannan - 89]. Many new developments since then leading to a rich theory of geometric random walks!

First attempt. Suppose $K \subseteq [-1/2, 1/2]^n$.

- Sample m uniform random points from $[-1/2, 1/2]^n$ and count how many belong to K. Let m' of them belong to K. Output m'/m.

$$\mathbb{E}\left[\frac{m'}{m}\right] = \frac{\mathrm{vol}(K)}{\mathrm{vol}([-1/2, 1/2]^n)} = \mathrm{vol}(K).$$

How large does m need to be?

Consider the example of $K = \mathcal{B}(0, 1/2)$. Then $K \subseteq [-1/2, 1/2]^n$.

How large does m need to be?

Consider the example of $K = \mathcal{B}(0, 1/2)$. Then $K \subseteq [-1/2, 1/2]^n$.

- $n = 1$

$$\frac{\mathrm{vol}\,(K)}{\mathrm{vol}\,([-1/2, 1/2]^n)} = 1.$$

How large does m need to be?

Consider the example of $K = \mathcal{B}(0, 1/2)$. Then $K \subseteq [-1/2, 1/2]^n$.

- $n = 1$

$$\frac{\mathrm{vol}\,(K)}{\mathrm{vol}\,([-1/2, 1/2]^n)} = 1.$$

- $n = 2$

$$\frac{\mathrm{vol}\,(K)}{\mathrm{vol}\,([-1/2, 1/2]^n)} = \frac{\pi(1/2)^2}{1} = \frac{\pi}{4}.$$

How large does m need to be?

Consider the example of $K = \mathcal{B}(0, 1/2)$. Then $K \subseteq [-1/2, 1/2]^n$.

- $n = 1$

$$\frac{\operatorname{vol}(K)}{\operatorname{vol}([-1/2, 1/2]^n)} = 1.$$

- $n = 2$

$$\frac{\operatorname{vol}(K)}{\operatorname{vol}([-1/2, 1/2]^n)} = \frac{\pi(1/2)^2}{1} = \frac{\pi}{4}.$$

- $n = 3$

$$\frac{\operatorname{vol}(K)}{\operatorname{vol}([-1/2, 1/2]^n)} = \frac{4\pi/3(1/2)^3}{1} = \frac{\pi}{6}.$$

How large does m need to be?

Consider the example of $K = \mathcal{B}(0, 1/2)$. Then $K \subseteq [-1/2, 1/2]^n$.

- $n = 1$
$$\frac{\operatorname{vol}(K)}{\operatorname{vol}([-1/2, 1/2]^n)} = 1.$$
- $n = 2$
$$\frac{\operatorname{vol}(K)}{\operatorname{vol}([-1/2, 1/2]^n)} = \frac{\pi(1/2)^2}{1} = \frac{\pi}{4}.$$
- $n = 3$
$$\frac{\operatorname{vol}(K)}{\operatorname{vol}([-1/2, 1/2]^n)} = \frac{4\pi/3(1/2)^3}{1} = \frac{\pi}{6}.$$
- For general n
$$\frac{\operatorname{vol}(K)}{\operatorname{vol}([-1/2, 1/2]^n)} \leq c^n$$
for some constant $c \in (0, 1)$.

Therefore, m will have to be exponential in n.

Second attempt

For any convex $S_1 \subset S_2 \ldots \subset S_L$ such that $S_1 \subseteq K \subseteq S_L$, we have

$$\mathrm{vol}\,(K) = \frac{\mathrm{vol}\,(K \cap S_L)}{\mathrm{vol}\,(K \cap S_{L-1})} \cdot \frac{\mathrm{vol}\,(K \cap S_{L-1})}{\mathrm{vol}\,(K \cap S_{L-2})} \cdots \frac{\mathrm{vol}\,(K \cap S_2)}{\mathrm{vol}\,(K \cap S_1)} \cdot \mathrm{vol}\,(S_1)\,.$$

Second attempt

For any convex $S_1 \subset S_2 \ldots \subset S_L$ such that $S_1 \subseteq K \subseteq S_L$, we have

$$\mathrm{vol}\left(K\right) = \frac{\mathrm{vol}\left(K \cap S_L\right)}{\mathrm{vol}\left(K \cap S_{L-1}\right)} \cdot \frac{\mathrm{vol}\left(K \cap S_{L-1}\right)}{\mathrm{vol}\left(K \cap S_{L-2}\right)} \cdots \frac{\mathrm{vol}\left(K \cap S_2\right)}{\mathrm{vol}\left(K \cap S_1\right)} \cdot \mathrm{vol}\left(S_1\right).$$

- Choose $\{S_i : i \in [L]\}$ such that $\mathrm{vol}\left(K \cap S_i\right)/\mathrm{vol}\left(K \cap S_{i-1}\right)$ can be estimated by sampling, i.e., sample points from $K \cap S_i$ and count how many of them belong to $K \cap S_{i-1}$.
- This can be done in polynomial time if $\mathrm{vol}\left(K \cap S_i\right)/\mathrm{vol}\left(K \cap S_{i-1}\right) \leq \mathrm{poly}\left(n\right)$ (H.W.).

Second attempt

For any convex $S_1 \subset S_2 \ldots \subset S_L$ such that $S_1 \subseteq K \subseteq S_L$, we have

$$\mathrm{vol}\left(K\right) = \frac{\mathrm{vol}\left(K \cap S_L\right)}{\mathrm{vol}\left(K \cap S_{L-1}\right)} \cdot \frac{\mathrm{vol}\left(K \cap S_{L-1}\right)}{\mathrm{vol}\left(K \cap S_{L-2}\right)} \cdots \frac{\mathrm{vol}\left(K \cap S_2\right)}{\mathrm{vol}\left(K \cap S_1\right)} \cdot \mathrm{vol}\left(S_1\right).$$

- Choose $\{S_i : i \in [L]\}$ such that $\mathrm{vol}\left(K \cap S_i\right)/\mathrm{vol}\left(K \cap S_{i-1}\right)$ can be estimated by sampling, i.e., sample points from $K \cap S_i$ and count how many of them belong to $K \cap S_{i-1}$.
- This can be done in polynomial time if $\mathrm{vol}\left(K \cap S_i\right)/\mathrm{vol}\left(K \cap S_{i-1}\right) \leq \mathrm{poly}\left(n\right)$ (H.W.).
- Choose $S_i = \mathcal{B}\left(0, R_i\right)$ where $R_1 = 1$ and $R_{i+1} = (1 + 1/n)R_i$.

Second attempt

For any convex $S_1 \subset S_2 \ldots \subset S_L$ such that $S_1 \subseteq K \subseteq S_L$, we have

$$\mathrm{vol}\,(K) = \frac{\mathrm{vol}\,(K \cap S_L)}{\mathrm{vol}\,(K \cap S_{L-1})} \cdot \frac{\mathrm{vol}\,(K \cap S_{L-1})}{\mathrm{vol}\,(K \cap S_{L-2})} \cdots \frac{\mathrm{vol}\,(K \cap S_2)}{\mathrm{vol}\,(K \cap S_1)} \cdot \mathrm{vol}\,(S_1)\,.$$

- Choose $\{S_i : i \in [L]\}$ such that $\mathrm{vol}\,(K \cap S_i)\,/\mathrm{vol}\,(K \cap S_{i-1})$ can be estimated by sampling, i.e., sample points from $K \cap S_i$ and count how many of them belong to $K \cap S_{i-1}$.
- This can be done in polynomial time if $\mathrm{vol}\,(K \cap S_i)\,/\mathrm{vol}\,(K \cap S_{i-1}) \leq \mathrm{poly}\,(n)$ (H.W.).
- Choose $S_i = \mathcal{B}\,(0, R_i)$ where $R_1 = 1$ and $R_{i+1} = (1 + 1/n)R_i$.
- Then $R_i = (1 + 1/n)^{i-1} R_1 = (1 + 1/n)^{i-1}$.

Since we want $K \subseteq S_L$, it suffices to have $R_L \geq R$. Therefore,

$$R \leq R_L = \left(1 + \frac{1}{n}\right)^{L-1} \leq e^{(L-1)/n} \implies L - 1 \geq n \log R.$$

Since we want $K \subseteq S_L$, it suffices to have $R_L \geq R$. Therefore,

$$R \leq R_L = \left(1 + \frac{1}{n}\right)^{L-1} \leq e^{(L-1)/n} \implies L - 1 \geq n \log R.$$

Fact

To estimate the product of L quatities to relative error ϵ, it suffices to estimate each quantity to a relative error of $O(\epsilon/L)$.

Since we want $K \subseteq S_L$, it suffices to have $R_L \geq R$. Therefore,

$$R \leq R_L = \left(1 + \frac{1}{n}\right)^{L-1} \leq e^{(L-1)/n} \implies L - 1 \geq n \log R.$$

Fact

To estimate the product of L quatities to relative error ϵ, it suffices to estimate each quantity to a relative error of $O(\epsilon/L)$.

- Since $S_i \cap K \subseteq (1 + 1/n)\,(S_{i-1} \cap K)$, we have $\frac{\mathrm{vol}(S_i \cap K)}{\mathrm{vol}(S_{i-1} \cap K)} \leq e$ (why?).

Since we want $K \subseteq S_L$, it suffices to have $R_L \geq R$. Therefore,

$$R \leq R_L = \left(1 + \frac{1}{n}\right)^{L-1} \leq e^{(L-1)/n} \implies L - 1 \geq n \log R.$$

Fact

To estimate the product of L quatities to relative error ϵ, it suffices to estimate each quantity to a relative error of $O(\epsilon/L)$.

- Since $S_i \cap K \subseteq (1 + 1/n)(S_{i-1} \cap K)$, we have $\frac{\mathrm{vol}(S_i \cap K)}{\mathrm{vol}(S_{i-1} \cap K)} \leq e$ (why?).
- Therefore, $\mathrm{vol}(S_i \cap K) / \mathrm{vol}(S_{i-1} \cap K)$ can be estimated using a polynomial number of uniform random samples from $S_i \cap K$.

Since we want $K \subseteq S_L$, it suffices to have $R_L \geq R$. Therefore,

$$R \leq R_L = \left(1 + \frac{1}{n}\right)^{L-1} \leq e^{(L-1)/n} \implies L - 1 \geq n \log R.$$

Fact

To estimate the product of L quatities to relative error ϵ, it suffices to estimate each quantity to a relative error of $O(\epsilon/L)$.

- Since $S_i \cap K \subseteq (1 + 1/n)(S_{i-1} \cap K)$, we have $\frac{\mathrm{vol}(S_i \cap K)}{\mathrm{vol}(S_{i-1} \cap K)} \leq e$ (why?).
- Therefore, $\mathrm{vol}(S_i \cap K) / \mathrm{vol}(S_{i-1} \cap K)$ can be estimated using a polynomial number of uniform random samples from $S_i \cap K$.
- $S_i \cap K$ is convex. We use random walks for sampling "approximately uniform" random points from $S_i \cap K$. This will also suffice for estimating volume.

Grid Walk(δ)

- Impose a grid of side length δ on K.
- The neighbours of point x on the grid are $\{x \pm \delta e_i : i \in [n]\} \cap K$.

Grid Walk(δ)

- Impose a grid of side length δ on K.
- The neighbours of point x on the grid are $\{x \pm \delta e_i : i \in [n]\} \cap K$.
- At point x,
  1. Probability of moving to each neighbor of x is $1/(4n)$.
  2. Walk stays at x with remaining probability.

Grid Walk(δ)

- Impose a grid of side length δ on K.
- The neighbours of point x on the grid are $\{x \pm \delta e_i : i \in [n]\} \cap K$.
- At point x,
  1. Probability of moving to each neighbor of x is $1/(4n)$.
  2. Walk stays at x with remaining probability.
- Walk can be performed using membership oracle for K.

Grid Walk(δ)

- Impose a grid of side length δ on K.
- The neighbours of point x on the grid are $\{x \pm \delta e_i : i \in [n]\} \cap K$.
- At point x,
  1. Probability of moving to each neighbor of x is $1/(4n)$.
  2. Walk stays at x with remaining probability.
- Walk can be performed using membership oracle for K.
- Number of vertices can be exponential in n, for e.g., $K = [0, 1]^n$.

Grid Walk(δ)

- Impose a grid of side length δ on K.
- The neighbours of point x on the grid are $\{x \pm \delta e_i : i \in [n]\} \cap K$.
- At point x,
  1. Probability of moving to each neighbor of x is $1/(4n)$.
  2. Walk stays at x with remaining probability.
- Walk can be performed using membership oracle for K.
- Number of vertices can be exponential in n, for e.g., $K = [0,1]^n$.
- Stationary distribution π is the uniform distribution on the vertices.

Grid Walk(δ)

- Impose a grid of side length δ on K.
- The neighbours of point x on the grid are $\{x \pm \delta e_i : i \in [n]\} \cap K$.
- At point x,
  1. Probability of moving to each neighbor of x is $1/(4n)$.
  2. Walk stays at x with remaining probability.
- Walk can be performed using membership oracle for K.
- Number of vertices can be exponential in n, for e.g., $K = [0, 1]^n$.
- Stationary distribution π is the uniform distribution on the vertices.

Theorem (Informal)

Mixing time of the Grid walk is polynomial.

Grid Walk

Final point output by walk will not be random from all of K, but only from a subset of K.

Grid Walk

Final point output by walk will not be random from all of K, but only from a subset of K.

- If x is the final point in the walk, output a random point in $\prod_{i \in [n]} [x_i - \delta/2, x_i + \delta/2]$.

Grid Walk

Final point output by walk will not be random from all of K, but only from a subset of K.

- If x is the final point in the walk, output a random point in $\Pi_{i\in[n]}[x_i - \delta/2, x_i + \delta/2]$.

This may give some points not in K, and may miss some points in K.

Grid Walk

Final point output by walk will not be random from all of K, but only from a subset of K.

- If x is the final point in the walk, output a random point in $\Pi_{i \in [n]}[x_i - \delta/2, x_i + \delta/2]$.

This may give some points not in K, and may miss some points in K.

- Sample points from the cube till we get a point from K?

Grid Walk

Final point output by walk will not be random from all of K, but only from a subset of K.

- If x is the final point in the walk, output a random point in $\Pi_{i\in[n]}[x_i - \delta/2, x_i + \delta/2]$.

This may give some points not in K, and may miss some points in K.

- Sample points from the cube till we get a point from K?

This may not give uniform distribution on points from K. Each cube needs to be chosen with probability proportional to its intersection with K.

Grid Walk

Final point output by walk will not be random from all of K, but only from a subset of K.

- If x is the final point in the walk, output a random point in $\Pi_{i \in [n]}[x_i - \delta/2, x_i + \delta/2]$.

This may give some points not in K, and may miss some points in K.

- Sample points from the cube till we get a point from K?

This may not give uniform distribution on points from K. Each cube needs to be chosen with probability proportional to its intersection with K.

- If the final point does not belong to K, then start again from beginning.

Grid Walk

Final point output by walk will not be random from all of K, but only from a subset of K.

- If x is the final point in the walk, output a random point in $\Pi_{i \in [n]}[x_i - \delta/2, x_i + \delta/2]$.

This may give some points not in K, and may miss some points in K.

- Sample points from the cube till we get a point from K?

This may not give uniform distribution on points from K. Each cube needs to be chosen with probability proportional to its intersection with K.

- If the final point does not belong to K, then start again from beginning.

The space of grid points in K may not be connected.

- Work with $K' := (1+\alpha)K$.

Set $\alpha := \delta\sqrt{n}$.

Set $\alpha := \delta\sqrt{n}$.

- Any cube that intersects K will be fully contained in K' (H.W.).

Set $\alpha := \delta\sqrt{n}$.

- Any cube that intersects K will be fully contained in K' (H.W.).
- Any two grid points in K are connected in K'.

Set $\alpha := \delta\sqrt{n}$.

- Any cube that intersects K will be fully contained in K' (H.W.).
- Any two grid points in K are connected in K'.

Algorithm: Perform grid walk on K'. If x is the final point in the walk, output a random point in $\Pi_{i\in[n]}[x_i - \delta/2, x_i + \delta/2]$. If the final point does not belong to K, then start again from beginning.

Set $\alpha := \delta\sqrt{n}$.

- Any cube that intersects K will be fully contained in K' (H.W.).
- Any two grid points in K are connected in K'.

Algorithm: Perform grid walk on K'. If x is the final point in the walk, output a random point in $\Pi_{i\in[n]}[x_i - \delta/2, x_i + \delta/2]$. If the final point does not belong to K, then start again from beginning.

- Probability of success is $\mathrm{vol}(K)$ divided by the volume of all δ-cubes whose centers are in K'. By the argument above, every such cube is contained in $(1+\alpha)K'$. Therefore, probability of success is at least

$$\frac{\mathrm{vol}(K)}{\mathrm{vol}((1+\alpha)K')} = \frac{\mathrm{vol}(K)}{\mathrm{vol}((1+\alpha)^2 K)} = \frac{1}{(1+\alpha)^{2n}} = (1+\delta\sqrt{n})^{-2n}.$$

Set $\alpha := \delta\sqrt{n}$.

- Any cube that intersects K will be fully contained in K' (H.W.).
- Any two grid points in K are connected in K'.

Algorithm: Perform grid walk on K'. If x is the final point in the walk, output a random point in $\Pi_{i\in[n]}[x_i - \delta/2, x_i + \delta/2]$. If the final point does not belong to K, then start again from beginning.

- Probability of success is $\mathrm{vol}\,(K)$ divided by the volume of all δ-cubes whose centers are in K'. By the argument above, every such cube is contained in $(1+\alpha)K'$. Therefore, probability of success is at least

$$\frac{\mathrm{vol}\,(K)}{\mathrm{vol}\,((1+\alpha)K')} = \frac{\mathrm{vol}\,(K)}{\mathrm{vol}\,((1+\alpha)^2 K)} = \frac{1}{(1+\alpha)^{2n}} = (1+\delta\sqrt{n})^{-2n}.$$

Therefore, if we choose $\delta = 1/(cn^{1.5})$, then probability of success is at least a constant.

Some other random walks

Definition (Ball Walk(δ))

1. Pick a uniform random point y from the ball of radius δ centered at the current point x.
2. If y is in K, then go to y; else stay at x.

Some other random walks

Definition (Ball Walk(δ))

1. Pick a uniform random point y from the ball of radius δ centered at the current point x.
2. If y is in K, then go to y; else stay at x.

Definition (Hit-and-run)

1. Pick a uniform random line ℓ through the current point x.
2. Go to a uniform random point on the chord $\ell \cap K$.