A Polynomial Number of Random Points does not Determine the Volume of a Convex Body

Ronen Eldan, Tel Aviv University

Cleveland, Aug 2, 2010

Volume Computing Algorithms

- A convex body in $\mathbb{R}^n$ is usually given by a certain **oracle**. An oracle is a "black box" which provides us information about the body. Several examples of such oracles:

Volume Computing Algorithms

- A convex body in $\mathbb{R}^n$ is usually given by a certain **oracle**. An oracle is a "black box" which provides us information about the body. Several examples of such oracles:
  - The **membership oracle**, which, for a point $x \in \mathbb{R}^n$, answers either $x \in K$ or $x \notin K$.

Volume Computing Algorithms

- A convex body in $\mathbb{R}^n$ is usually given by a certain **oracle**. An oracle is a "black box" which provides us information about the body. Several examples of such oracles:
  - The **membership oracle**, which, for a point $x \in \mathbb{R}^n$, answers either $x \in K$ or $x \notin K$.
  - The **separation oracle** (which answers $x \notin K$ and provides a separating hyperplane).

Volume Computing Algorithms

- A convex body in $\mathbb{R}^n$ is usually given by a certain **oracle**. An oracle is a "black box" which provides us information about the body. Several examples of such oracles:
  - The **membership oracle**, which, for a point $x \in \mathbb{R}^n$, answers either $x \in K$ or $x \notin K$.
  - The **separation oracle** (which answers $x \notin K$ and provides a separating hyperplane).
  - The **random point oracle** which just provides random points uniformly distributed in K.

Volume Computing Algorithms

- A convex body in $\mathbb{R}^n$ is usually given by a certain **oracle**. An oracle is a "black box" which provides us information about the body. Several examples of such oracles:
  - The **membership oracle**, which, for a point $x \in \mathbb{R}^n$, answers either $x \in K$ or $x \notin K$.
  - The **separation oracle** (which answers $x \notin K$ and provides a separating hyperplane).
  - The **random point oracle** which just provides random points uniformly distributed in K.
- The **complexity** or **running time** of the algorithm is defined as the number of questions posed to the oracle.

Volume Computing Algorithms

- A convex body in $\mathbb{R}^n$ is usually given by a certain **oracle**. An oracle is a "black box" which provides us information about the body. Several examples of such oracles:
  - The **membership oracle**, which, for a point $x \in \mathbb{R}^n$, answers either $x \in K$ or $x \notin K$.
  - The **separation oracle** (which answers $x \notin K$ and provides a separating hyperplane).
  - The **random point oracle** which just provides random points uniformly distributed in K.
- The **complexity** or **running time** of the algorithm is defined as the number of questions posed to the oracle.
- We are looking for polynomial time algorithms which approximate the volume up to a (multiplicative) constant C with probability ≈ 1.

Some History

- The first results in the subject appeared in the mid 80's.

Some History

- The first results in the subject appeared in the mid 80's.
  - G. Elekes realized that it is impossible to find a fast, **deterministic** algorithm which approximates the volume, in polynomial time. This follows from the fact that the convex hull of polynomially many points in the euclidean ball has a very small volume.

Some History

- The first results in the subject appeared in the mid 80's.
  - G. Elekes realized that it is impossible to find a fast, **deterministic** algorithm which approximates the volume, in polynomial time. This follows from the fact that the convex hull of polynomially many points in the euclidean ball has a very small volume.
  - Bárány and Z. Füredi later found to the best possible bound, which is: the volume is computable up to $(\frac{cn}{\log n})^n$.

Some History

- The first results in the subject appeared in the mid 80's.
  - G. Elekes realized that it is impossible to find a fast, **deterministic** algorithm which approximates the volume, in polynomial time. This follows from the fact that the convex hull of polynomially many points in the euclidean ball has a very small volume.
  - Bárány and Z. Füredi later found to the best possible bound, which is: the volume is computable up to $(\frac{cn}{\log n})^n$.
  - M. E. Dyer, A. M. Frieze, and R. Kannan were the first to find a polynomial time (non-deterministic) algorithm which approximates the volume up to some constant, using the **membership oracle**.

Some History

- The first results in the subject appeared in the mid 80's.
  - G. Elekes realized that it is impossible to find a fast, **deterministic** algorithm which approximates the volume, in polynomial time. This follows from the fact that the convex hull of polynomially many points in the euclidean ball has a very small volume.
  - Bárány and Z. Füredi later found to the best possible bound, which is: the volume is computable up to $(\frac{cn}{\log n})^n$.
  - M. E. Dyer, A. M. Frieze, and R. Kannan were the first to find a polynomial time (non-deterministic) algorithm which approximates the volume up to some constant, using the **membership oracle**.
  - Since the early 90's many improvements were found, and the running time was improved from $O(n^{24})$ to $O(n^4)$, given by L. Lovász, S. Vempala in 2004. Additional works by B.Bollobás, N. Goyal, A.Kalai, L.Rademacher, M.Simonovits and other people whose names are also very hard to pronounce.

The Random Point Oracle

Theorem (E., 2009): There does not exist constants $C, p, \kappa > 0$ such that for every dimension n there exists an $O(n^\kappa)$ algorithm which estimates the volume of convex bodies up to C with probability at least p.

The Random Point Oracle

Theorem (E., 2009): There does not exist constants $C, p, \kappa > 0$ such that for every dimension n there exists an $O(n^\kappa)$ algorithm which estimates the volume of convex bodies up to C with probability at least p.

- The result does not involve any complexity arguments, it is of information-theoretical nature: The random points just do not contain enough information for the volume to be approximated.

The Random Point Oracle

Theorem (E., 2009): There does not exist constants $C, p, \kappa > 0$ such that for every dimension n there exists an $O(n^{\kappa})$ algorithm which estimates the volume of convex bodies up to C with probability at least p.

- The result does not involve any complexity arguments, it is of information-theoretical nature: The random points just do not contain enough information for the volume to be approximated.
- The volume cannot be approximated even if every point is given in an infinitely good accuracy.

The Random Point Oracle

Theorem (E., 2009): There does not exist constants $C, p, \kappa > 0$ such that for every dimension n there exists an $O(n^\kappa)$ algorithm which estimates the volume of convex bodies up to C with probability at least p.

- The result does not involve any complexity arguments, it is of information-theoretical nature: The random points just do not contain enough information for the volume to be approximated.
- The volume cannot be approximated even if every point is given in an infinitely good accuracy.
- A related result of N.Goyal and L.Rademacher shows that in order to learn a convex body one needs $2^{\Omega(\sqrt{n})}$ random points.

A very simple toy model

- Let us consider the n-dimensional unit-volume cube, and its obvious partition into 2^n cells.

A very simple toy model

- Let us consider the n-dimensional unit-volume cube, and its obvious partition into 2^n cells.
- We do the following random construction. Pick $0 < C < 1$. Create a subset of the cube by, for each one of the cells, removing it with probability $1 - C$.

A very simple toy model

- Let us consider the n-dimensional unit-volume cube, and its obvious partition into 2^n cells.
- We do the following random construction. Pick $0 < C < 1$. Create a subset of the cube by, for each one of the cells, removing it with probability $1 - C$.
- Generate n^{100} random points according the the uniform measure on the remaining body.

A very simple toy model

- Let us consider the n-dimensional unit-volume cube, and its obvious partition into 2^n cells.
- We do the following random construction. Pick $0 < C < 1$. Create a subset of the cube by, for each one of the cells, removing it with probability $1 - C$.
- Generate n^{100} random points according the the uniform measure on the remaining body.
- Claim: As long as C is not too close to 0, one cannot be guess this constant according to these n^{100} points.

A very simple toy model

- Let us consider the n-dimensional unit-volume cube, and its obvious partition into 2^n cells.
- We do the following random construction. Pick $0 < C < 1$. Create a subset of the cube by, for each one of the cells, removing it with probability $1 - C$.
- Generate n^{100} random points according the the uniform measure on the remaining body.
- Claim: As long as C is not too close to 0, one cannot be guess this constant according to these n^{100} points.
- The reason: Consider the event that two points are picked from the same cell. If we condition on this event, C does not change anything.

A very simple toy model

- Let us consider the n-dimensional unit-volume cube, and its obvious partition into 2^n cells.
- We do the following random construction. Pick $0 < C < 1$. Create a subset of the cube by, for each one of the cells, removing it with probability $1 - C$.
- Generate n^{100} random points according the the uniform measure on the remaining body.
- Claim: As long as C is not too close to 0, one cannot be guess this constant according to these n^{100} points.
- The reason: Consider the event that two points are picked from the same cell. If we condition on this event, C does not change anything.
- But since the number of cells is exponential, this event is negligible.

A very simple toy model - continued

- A constant C induces a probability measure on the set of sequences of points, Ω. The **total variation** distance between two such measures is small.

$$d_{TV}(X_1, X_2) = \sup_{A \subset \Omega} |Prob(X_1 \in A) - Prob(X_2 \in A)|$$

A very simple toy model - continued

- A constant C induces a probability measure on the set of sequences of points, Ω. The **total variation** distance between two such measures is small.

$$d_{TV}(X_1, X_2) = \sup_{A \subset \Omega} |Prob(X_1 \in A) - Prob(X_2 \in A)|$$

- Of course, the geometry in this model was meaningless. If we took the convex hull after this construction, the same argument could doubtfully work (since some of points would be generated from "partial" cells, and we could learn their geometry).

A very simple toy model - continued

- A constant C induces a probability measure on the set of sequences of points, Ω. The **total variation** distance between two such measures is small.

$$d_{TV}(X_1, X_2) = \sup_{A \subset \Omega} |Prob(X_1 \in A) - Prob(X_2 \in A)|$$

- Of course, the geometry in this model was meaningless. If we took the convex hull after this construction, the same argument could doubtfully work (since some of points would be generated from "partial" cells, and we could learn their geometry).
- Maybe we can try some other random deletion process?

Outline of proof

- Since the algorithm does not make any decisions in the input phase, we may assume that it is **deterministic**.

Outline of proof

- Since the algorithm does not make any decisions in the input phase, we may assume that it is **deterministic**.
- Define $N = n^{\kappa}$ to be the number of random points and $\Omega = (\mathbb{R}^n)^N$. A deterministic algorithm is a function $F : \Omega \to \mathbb{R}$ which takes a sequence of points and returns the estimated volume of the body.

Outline of proof

- Since the algorithm does not make any decisions in the input phase, we may assume that it is **deterministic**.
- Define $N = n^\kappa$ to be the number of random points and $\Omega = (\mathbb{R}^n)^N$. A deterministic algorithm is a function $F : \Omega \to \mathbb{R}$ which takes a sequence of points and returns the estimated volume of the body.
- Suppose K_1 and K_2 convex-body-valued random variables with the following properties:

Outline of proof

- Since the algorithm does not make any decisions in the input phase, we may assume that it is **deterministic**.
- Define $N = n^{\kappa}$ to be the number of random points and $\Omega = (\mathbb{R}^n)^N$. A deterministic algorithm is a function $F : \Omega \to \mathbb{R}$ which takes a sequence of points and returns the estimated volume of the body.
- Suppose K_1 and K_2 convex-body-valued random variables with the following properties:
  - With high probability, $Vol(K_1) \approx C_1$ and $Vol(K_2) \approx C_2$, and $\frac{C_1}{C_2}$ is very large.

Outline of proof

- Since the algorithm does not make any decisions in the input phase, we may assume that it is **deterministic**.
- Define $N = n^{\kappa}$ to be the number of random points and $\Omega = (\mathbb{R}^n)^N$. A deterministic algorithm is a function $F : \Omega \to \mathbb{R}$ which takes a sequence of points and returns the estimated volume of the body.
- Suppose K_1 and K_2 convex-body-valued random variables with the following properties:
  - With high probability, $Vol(K_1) \approx C_1$ and $Vol(K_2) \approx C_2$, and $\frac{C_1}{C_2}$ is very large.
  - The random convex body K_i induces a probability measure on Ω in the following simple way: first generate a body K and then generate a sequence of uniformly distributed random points from K.

Outline of proof

- Since the algorithm does not make any decisions in the input phase, we may assume that it is **deterministic**.
- Define $N = n^{\kappa}$ to be the number of random points and $\Omega = (\mathbb{R}^n)^N$. A deterministic algorithm is a function $F : \Omega \to \mathbb{R}$ which takes a sequence of points and returns the estimated volume of the body.
- Suppose K_1 and K_2 convex-body-valued random variables with the following properties:
  - With high probability, $Vol(K_1) \approx C_1$ and $Vol(K_2) \approx C_2$, and $\frac{C_1}{C_2}$ is very large.
  - The random convex body K_i induces a probability measure on Ω in the following simple way: first generate a body K and then generate a sequence of uniformly distributed random points from K.
  - These corresponding measures have a small total variation distance.

- A small total variation distance between the random variables X_1, X_2 representing sequences of random points implies that for every algorithm $F : \Omega \to \mathbb{R}$, we would have a small total variation distance between $F(X_1)$ and $F(X_2)$.

- A small total variation distance between the random variables X_1, X_2 representing sequences of random points implies that for every algorithm $F : \Omega \to \mathbb{R}$, we would have a small total variation distance between $F(X_1)$ and $F(X_2)$.
- To summarize: we are looking for two different random constructions of convex bodies whose volumes are very different, yet whose output of random points is very similar, so that no algorithm can distiguish between them.

The random construction

Denote

$$T_0(\theta) = D_n \cap \{x; \langle x, \theta \rangle \leq n^{-\frac{1}{3}}\}.$$

The random construction

Denote

$$T_0(\theta) = D_n \cap \{x; \langle x, \theta \rangle \leq n^{-\frac{1}{3}}\}.$$

Note that it contains essentially all of the mass of the euclidean ball.

The random construction

Denote

$$T_0(\theta) = D_n \cap \{x; \langle x, \theta \rangle \leq n^{-\frac{1}{3}}\}.$$

Note that it contains essentially all of the mass of the euclidean ball. Next, Suppose that for every $\theta \in S^{n-1}$, $T(\theta)$ is a convex body satisfying

$$T_0(\theta) \subseteq T(\theta) \subseteq D_n$$

At this point, let us imagine that $T(\theta) = T_0(\theta)$.

The random construction

Denote

$$T_0(\theta) = D_n \cap \{x; \langle x, \theta \rangle \leq n^{-\frac{1}{3}}\}.$$

Note that it contains essentially all of the mass of the euclidean ball. Next, Suppose that for every $\theta \in S^{n-1}$, $T(\theta)$ is a convex body satisfying

$$T_0(\theta) \subseteq T(\theta) \subseteq D_n$$

At this point, let us imagine that $T(\theta) = T_0(\theta)$.
Let $\zeta = (\theta_1, ..., \theta_{m'})$ be a rotation invariant Poisson process on S^{n-1} with intensity m. Define

$$K = K_{T,m} = \bigcap_{1 \leq i \leq m'} T(\theta_i)$$

Weak correlation between generated points

Let $x_1, x_2 \in D_n$. Let us try to compare $\mathbb{P}(x_1 \in K \& x_2 \in K)$ with $\mathbb{P}(x_1 \in K)\mathbb{P}(x_2 \in K)$.

Weak correlation between generated points

Let $x_1, x_2 \in D_n$. Let us try to compare $\mathbb{P}(x_1 \in K \& x_2 \in K)$ with $\mathbb{P}(x_1 \in K)\mathbb{P}(x_2 \in K)$. We consider our deletion process as a superposition of deletion processes (divided into disjoint portions of S^{n-1}):

Weak correlation between generated points

Let $x_1, x_2 \in D_n$. Let us try to compare $\mathbb{P}(x_1 \in K \& x_2 \in K)$ with $\mathbb{P}(x_1 \in K)\mathbb{P}(x_2 \in K)$. We consider our deletion process as a superposition of deletion processes (divided into disjoint portions of S^{n-1}):

- A first process, S_1, which deletes x_1 and leaves x_2 intact.

Weak correlation between generated points

Let $x_1, x_2 \in D_n$. Let us try to compare $\mathbb{P}(x_1 \in K \& x_2 \in K)$ with $\mathbb{P}(x_1 \in K)\mathbb{P}(x_2 \in K)$. We consider our deletion process as a superposition of deletion processes (divided into disjoint portions of S^{n-1}):

- A first process, S_1, which deletes x_1 and leaves x_2 intact.
- A process S_2 which does the vise versa.

Weak correlation between generated points

Let $x_1, x_2 \in D_n$. Let us try to compare $\mathbb{P}(x_1 \in K \& x_2 \in K)$ with $\mathbb{P}(x_1 \in K)\mathbb{P}(x_2 \in K)$. We consider our deletion process as a superposition of deletion processes (divided into disjoint portions of S^{n-1}):

- A first process, S_1, which deletes x_1 and leaves x_2 intact.
- A process S_2 which does the vise versa.
- A process S_b which removed both together (in the same "cut").

Weak correlation between generated points

Let $x_1, x_2 \in D_n$. Let us try to compare $\mathbb{P}(x_1 \in K \& x_2 \in K)$ with $\mathbb{P}(x_1 \in K)\mathbb{P}(x_2 \in K)$. We consider our deletion process as a superposition of deletion processes (divided into disjoint portions of S^{n-1}):

- A first process, S_1, which deletes x_1 and leaves x_2 intact.
- A process S_2 which does the vise versa.
- A process S_b which removed both together (in the same "cut").

$$\frac{\mathbb{P}(x_1 \in K \& x_2 \in K)}{\mathbb{P}(x_1 \in K)\mathbb{P}(x_2 \in K)} = \frac{e^{-(m(S_1)+m(S_2)+m(S_b))}}{e^{-(m(S_1)+m(S_b)+m(S_2)+m(S_b))}} = e^{m(S_b)}$$

Weak correlation between generated points - continued

- It follows from concentration properties of the euclidean ball that typical points x_1, x_2 are almost orthogonal.

Weak correlation between generated points - continued

- It follows from concentration properties of the euclidean ball that typical points x_1, x_2 are almost orthogonal.
- For such points, one has,

$$m(S_b) \ll m(S_1)$$

Weak correlation between generated points - continued

- It follows from concentration properties of the euclidean ball that typical points x_1, x_2 are almost orthogonal.
- For such points, one has,

$$m(S_b) \ll m(S_1)$$

- Therefore, even if the chances of a single point to be removed are rather high, the chances for both to be removed in the same deletion are very small.

Weak correlation between generated points - continued

- It follows from concentration properties of the euclidean ball that typical points x_1, x_2 are almost orthogonal.
- For such points, one has,

$$m(S_b) \ll m(S_1)$$

- Therefore, even if the chances of a single point to be removed are rather high, the chances for both to be removed in the same deletion are very small.
- This means that we can cut a large portion of mass, still having

$$\mathbb{P}(x_1 \in K \& x_2 \in K)(1 + O(e^{-n^{\epsilon}})) = \mathbb{P}(x_1 \in K)\mathbb{P}(x_2 \in K)$$

for some $\epsilon > 0$.

Concentration of the volume

- By Fubini,

$$\mathbb{E}(Vol(K)) = \int_{D_n} P(x \in K)dx$$

Concentration of the volume

- By Fubini,

$$\mathbb{E}(Vol(K)) = \int_{D_n} P(x \in K)dx$$

- The above "almost independence" implies,

$$(\mathbb{E}Vol(K))^2 = \int_{D_n \times D_n} P(x_1 \in K)P(x_2 \in K)dx1dx2 \approx$$

$$\int_{D_n \times D_n} P(x_1, x_2 \in K)dx1dx2 = \mathbb{E}(Vol(K)^2)$$

Concentration of the volume

- By Fubini,

$$\mathbb{E}(Vol(K)) = \int_{D_n} P(x \in K)dx$$

- The above "almost independence" implies,

$$(\mathbb{E}Vol(K))^2 = \int_{D_n \times D_n} P(x_1 \in K)P(x_2 \in K)dx1dx2 \approx$$

$$\int_{D_n \times D_n} P(x_1, x_2 \in K)dx1dx2 = \mathbb{E}(Vol(K)^2)$$

- Hence, the volumes are very concentrated around their expectation.

How does the weak correlation help us?

- For $A \subset \Omega$,

$$P(A) = \mathbb{E}\frac{Vol(A \cap K^N)}{Vol(K^N)}$$

How does the weak correlation help us?

- For $A \subset \Omega$,

$$P(A) = \mathbb{E}\frac{Vol(A \cap K^N)}{Vol(K^N)}$$

- Since the volume is concentrated, and since

$$\mathbb{E} Vol(K^N \cap A) = \int_A \mathbb{P}(\forall i, x_i \in K),$$

How does the weak correlation help us?

- For $A \subset \Omega$,

$$P(A) = \mathbb{E}\frac{Vol(A \cap K^N)}{Vol(K^N)}$$

- Since the volume is concentrated, and since

$$\mathbb{E}Vol(K^N \cap A) = \int_A \mathbb{P}(\forall i, x_i \in K),$$

we get

$$P(A) \approx \frac{1}{\mathbb{E}(Vol(K^N))} \int_A \prod_i \mathbb{P}(\forall i, x_i \in K)$$

So...?

- The above means that the following processes are almost equivalent:
  - Generate the body K and then generate N random points, uniformly in K.

So...?

- The above means that the following processes are almost equivalent:
  - Generate the body K and then generate N random points, uniformly in K.
  - Generate the body K and generate one random point. Repeat N times.

So...?

- The above means that the following processes are almost equivalent:
  - Generate the body K and then generate N random points, uniformly in K.
  - Generate the body K and generate one random point. Repeat N times.
- Since the construction was centrally symmetric, all we have to do in order for the total variation distance between two induced measures to be small is to make sure that the distribution of $|x_1|$ is approximately the same in the two families.

So...?

- The above means that the following processes are almost equivalent:
  - Generate the body K and then generate N random points, uniformly in K.
  - Generate the body K and generate one random point. Repeat N times.
- Since the construction was centrally symmetric, all we have to do in order for the total variation distance between two induced measures to be small is to make sure that the distribution of $|x_1|$ is approximately the same in the two families.
- To do this, we use the liberty of choosing the shape of the cut, and the intensity of the poisson processes.

Possible further research

In fact, our proof gives a slightly stronger result:

Theorem

There exists $\varepsilon > 0$ *and a number* $n_0 \in \mathbb{N}$ *such that for all* $n > n_0$*, there does not exist an algorithm which takes* $e^{n^{\varepsilon}}$ *points generated randomly according to the uniform measure in a convex body* $K \subset \mathbb{R}^n$*, which determines* $Vol(K)$ *up to* $e^{n^{\varepsilon}}$ *with probability more than* $e^{-n^{\varepsilon}}$ *to be correct.*

Possible further research

In fact, our proof gives a slightly stronger result:

Theorem

There exists $\varepsilon > 0$ and a number $n_0 \in \mathbb{N}$ such that for all $n > n_0$, there does not exist an algorithm which takes e^{n^ε} points generated randomly according to the uniform measure in a convex body $K \subset \mathbb{R}^n$, which determines $Vol(K)$ up to e^{n^ε} with probability more than e^{-n^ε} to be correct.

- The volume radius of a convex body $K \subset \mathbb{R}^n$ is defined as $Volrad(K) = Vol(K)^{\frac{1}{n}}$. Clearly, it is much easier to estimate the volume radius, but is it possible (say, up to 1.001)?

Possible further research

In fact, our proof gives a slightly stronger result:

Theorem

There exists $\varepsilon > 0$ and a number $n_0 \in \mathbb{N}$ such that for all $n > n_0$, there does not exist an algorithm which takes e^{n^ε} points generated randomly according to the uniform measure in a convex body $K \subset \mathbb{R}^n$, which determines $Vol(K)$ up to e^{n^ε} with probability more than e^{-n^ε} to be correct.

- The volume radius of a convex body $K \subset \mathbb{R}^n$ is defined as $Volrad(K) = Vol(K)^{\frac{1}{n}}$. Clearly, it is much easier to estimate the volume radius, but is it possible (say, up to 1.001)?
- The above question is equivalent to the question whether or not the isotropic constant of a body can be estimated using a polynomial number of points.

What about a general convex body?

- This result gives us two families of convex bodies that cannot be distinguished if we only have a polynomial number of random points. What can be said about the distribution of n^{100} random points in a general convex body?

What about a general convex body?

- This result gives us two families of convex bodies that cannot be distinguished if we only have a polynomial number of random points. What can be said about the distribution of n^{100} random points in a general convex body?
- If we first apply a random rotation to the body, then the marginal distribution of a single point is spherically symmetric and therefore all of the information is contained in the Gramm matrix.
- The recent **thin shell** results imply that the main diagonal entries are all concentrated around n.
- What about the off diagonal entries?
  - Is their distribution gaussian?
  - For a small enough number of points, the off diagonal entries may be distributed in the same way for all convex bodies.

Thank you!