Introduction to Discrepancy Theory

Aleksandar (Sasho) Nikolov

University of Toronto

Outline

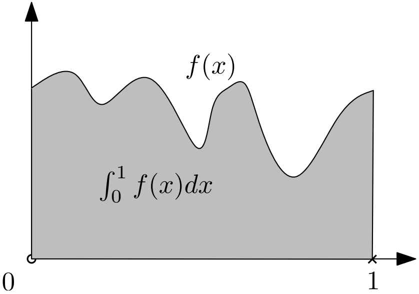
Discrepancy Theory

- How well can a *discrete* object approximation a *continuous* object?
- How well can a *small* object approximation a *big* object?

How to Compute an Integral?

A *fundamental* problem in sciences: *How to approximate the integral of a function?*

$$\int_0^1 f(x)dx = ?$$



The Issues

Didn't we learn this in calculus?

$$\int_0^1 xe^x dx = [xe^x]_0^1 - \int_0^1 e^x dx = [(x-1)e^x]_0^1 .$$

The Issues

Didn't we learn this in calculus?

$$\int_0^1 xe^x dx = [xe^x]_0^1 - \int_0^1 e^x dx = [(x-1)e^x]_0^1 .$$

- Integration is hard!
- Many interesting functions do not have a closed form integral at all.
- The function f may be *very* complicated!

The Issues Continue

Or we may not even know what f really is!
$f(x)$ may be:

- The speed of a particle at time x
- The price of a stock at time x
- The amount of energy released when we burn x grams of some chemical substance

The Issues Continue

Or we may not even know what f really is!
$f(x)$ may be:

- The speed of a particle at time x
- The price of a stock at time x
- The amount of energy released when we burn x grams of some chemical substance

Sometimes all we can do is treat f as a **black box**:

- we can *query* $f(x)$ at a specific x
- each query may require a new experiment: expensive!

The Issues Continue

Or we may not even know what f really is!
$f(x)$ may be:

- The speed of a particle at time x
- The price of a stock at time x
- The amount of energy released when we burn x grams of some chemical substance

Sometimes all we can do is treat f as a **black box**:

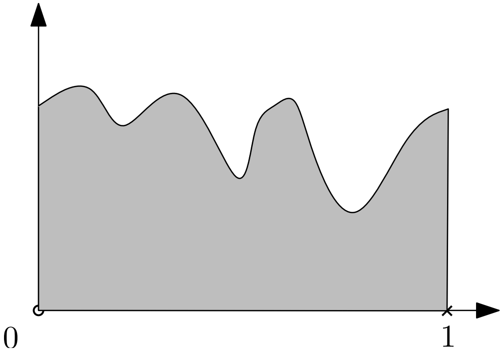
- we can *query* $f(x)$ at a specific x
- each query may require a new experiment: expensive!

Can we compute $\int_0^1 f(x)dx$ *with a* **black box** f*, under minimal assumptions, with few queries*?

The Monte Carlo Idea

During the Manhattan Project, von Neumann and Ulam had an idea (inspired by Ulam's uncle's gambling habbit):

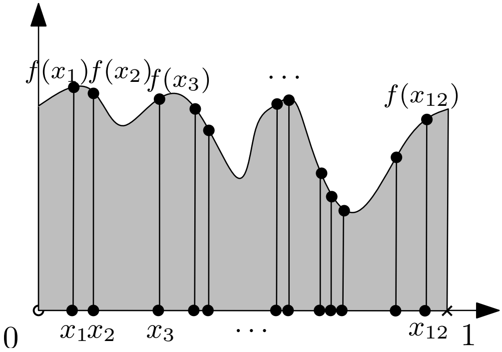
1. Pick n *random* points in $[0, 1]$
2. Estimate integral by average



The Monte Carlo Idea

During the Manhattan Project, von Neumann and Ulam had an idea (inspired by Ulam's uncle's gambling habbit):

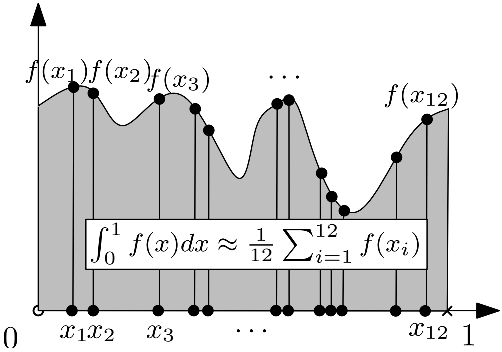
1. Pick n *random* points in $[0, 1]$
2. Estimate integral by average



The Monte Carlo Idea

During the Manhattan Project, von Neumann and Ulam had an idea (inspired by Ulam's uncle's gambling habbit):

1. Pick n *random* points in $[0, 1]$
2. Estimate integral by average



Monte Carlo: Convergence

If we pick n random points $x_1, \ldots, x_n \in [0,1]$ then

$$\left| \int_0^1 f(x)dx - \frac{1}{n}\sum_{i=1}^{n} f(x_i) \right| \approx \frac{\mathcal{E}(f)}{\sqrt{n}},$$

where $\mathcal{E}(f)$ is a measure of the *energy* of f.

$$\mathcal{E}(f) = \left(\int_0^1 |f(x)|^2 dx \right)^{1/2}$$

Can we do better?

We have a sequence $\vec{x} = (x_1, x_2, x_3, \ldots)$ that we want to use to estimate integrals, using an average.

Can we do better?

We have a sequence $\vec{x} = (x_1, x_2, x_3, \ldots)$ that we want to use to estimate integrals, using an average.

When we want a more accurate estimate, we evaluate f at a few more points from the sequence.

Can we do better?

We have a sequence $\vec{x} = (x_1, x_2, x_3, \ldots)$ that we want to use to estimate integrals, using an average.

When we want a more accurate estimate, we evaluate f at a few more points from the sequence.

We want the error

$$\mathrm{Err}(f, \vec{x}, n) := \left| \int_0^1 f(x)dx - \frac{1}{n} \sum_{i=1}^{n} f(x_i) \right|$$

to be as small as possible.

- We know that if $\vec{x}$ is random, $\mathrm{Err}(f, \vec{x}, n) \ll 1/\sqrt{n}$? for f with constant energy.

Can we do better?

We have a sequence $\vec{x} = (x_1, x_2, x_3, \ldots)$ that we want to use to estimate integrals, using an average.

When we want a more accurate estimate, we evaluate f at a few more points from the sequence.

We want the error

$$\mathrm{Err}(f, \vec{x}, n) := \left| \int_0^1 f(x)dx - \frac{1}{n} \sum_{i=1}^{n} f(x_i) \right|$$

to be as small as possible.

- We know that if $\vec{x}$ is random, $\mathrm{Err}(f, \vec{x}, n) \ll 1/\sqrt{n}$? for f with constant energy.
- *Can we achieve* $\mathrm{Err}(f, \vec{x}, n) \ll 1/n$ *for all "nice"* f*?*

Outline

Intervals Are Enough

If a sequence $\vec{x}$ has small error for all *intervals*, then it has small error for all *smooth* functions.

$$\delta(\vec{x}, n) = \max_{a,b \in [0,1]} \left| |a - b| - \frac{1}{n} |\{i : a \le x_i \le b\}| \right|.$$

Intervals Are Enough

If a sequence $\vec{x}$ has small error for all *intervals*, then it has small error for all *smooth* functions.

$$\delta(\vec{x}, n) = \max_{a,b \in [0,1]} \left| |a - b| - \frac{1}{n} |\{i : a \le x_i \le b\}| \right|.$$

Koksma-Hlawka inequality:

$$\mathrm{Err}(f, \vec{x}, n) \le V(f)\delta(\vec{x}, n),$$

where $V(f)$ is a measure of the smoothness of f (*total variation*).

Intervals Are Enough

If a sequence $\vec{x}$ has small error for all *intervals*, then it has small error for all *smooth* functions.

$$\delta(\vec{x}, n) = \max_{a,b\in[0,1]} \left| |a-b| - \frac{1}{n}|\{i : a \le x_i \le b\}| \right|.$$

Koksma-Hlawka inequality:

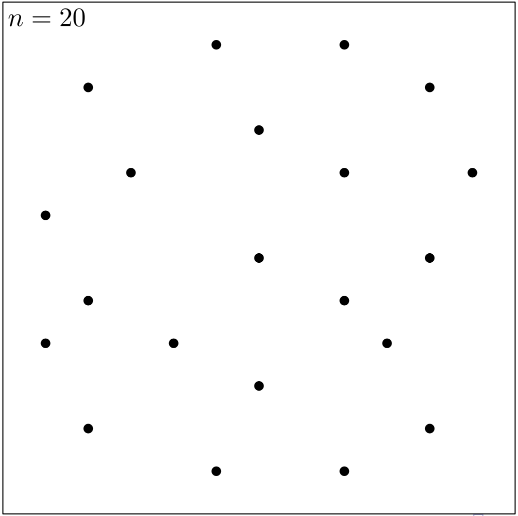
$$\mathrm{Err}(f, \vec{x}, n) \le V(f)\delta(\vec{x}, n),$$

where $V(f)$ is a measure of the smoothness of f (*total variation*).

van der Corput (1934): Can $\delta(\vec{x}, n) = O(1/n)$ for some sequence $\vec{x}$?

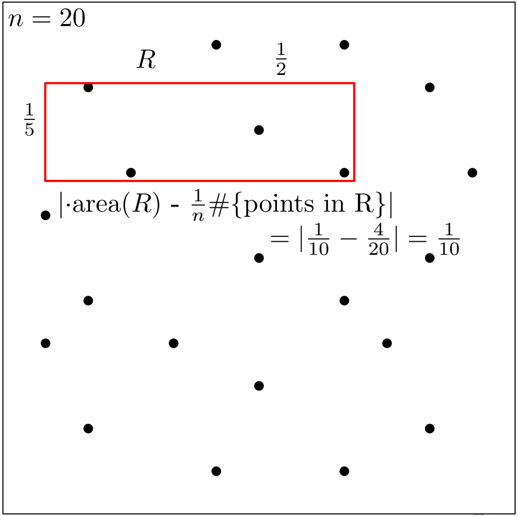
From Intervals to Rectangles

Roth showed that studying $\delta(\vec{x}, n)$ is equivalent to placing n points uniformly in a unit square. (Think of one dimension as the index.)



From Intervals to Rectangles

Roth showed that studying $\delta(\vec{x}, n)$ is equivalent to placing n points uniformly in a unit square. (Think of one dimension as the index.)



Discrepancy of Rectangles

For a set P of n points in $[0,1]^2$ and a rectangle $R = [a,b] \times [c,d]$

$$\begin{aligned} d(P,R) &= \left|\mathrm{area}(R) - \frac{|P \cap R|}{n}\right| \\ &= \left|(b-a)(d-c) - \frac{|P \cap R|}{n}\right| \\ d(P) &= \max_R D(P,R) \end{aligned}$$

Discrepancy of Rectangles

For a set P of n points in $[0,1]^2$ and a rectangle $R = [a,b] \times [c,d]$

$$d(P,R) = \left|\mathrm{area}(R) - \frac{|P \cap R|}{n}\right|$$
$$= \left|(b-a)(d-c) - \frac{|P \cap R|}{n}\right|$$
$$d(P) = \max_R D(P,R)$$

We can construct a sequence $\vec{x}$ with $\delta(\vec{x}, n) = O(f(n))$.

$$\Updownarrow$$

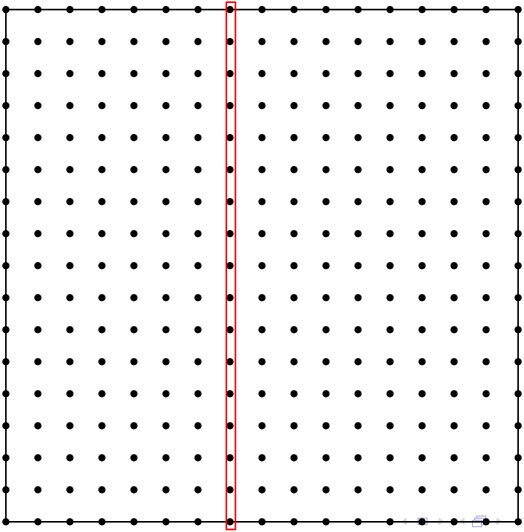
For any n, we can construct a set P of n points s.t. $d(P) = O(f(n))$.

Grid

Can we construct P s.t. $d(P) = O(1/n)$?

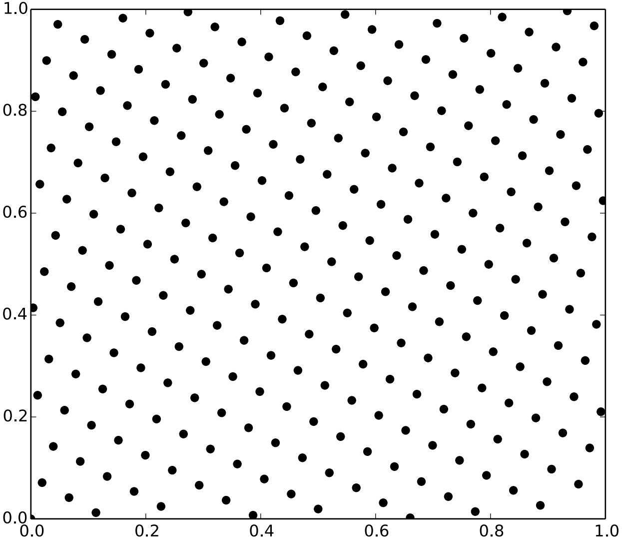
Grid

Can we construct P s.t. $d(P) = O(1/n)$?
Grid: $d(P) \approx \frac{1}{\sqrt{n}}$: area$(R) \approx 0$ and $|P \cap R| = \sqrt{n}$



Irrational Lattice

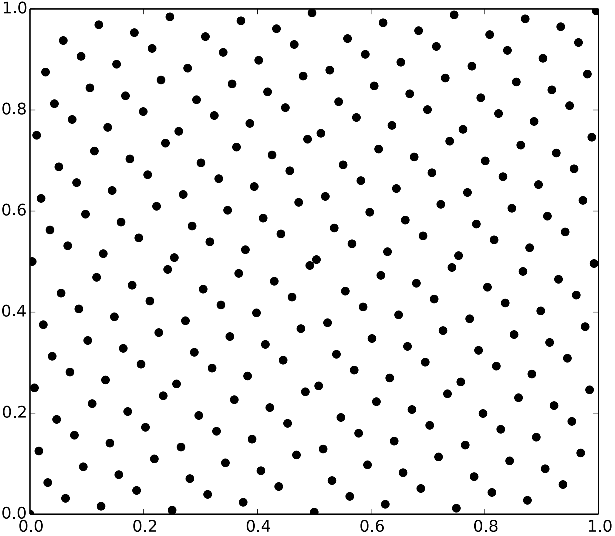
$P = \{(i/n, \{i \cdot \sqrt{2}\}) : i = 0, \ldots, n-1\}$: $d(P) = \Theta(\frac{\log n}{n})$
$\{x\}$ = fracional part of $x = x - \lfloor x \rfloor$



van der Corput set

$P = \{(i/n, \mathrm{rev}(i)) : i = 0, \ldots, n-1\}$: $D(P) = \Theta(\frac{\log n}{n})$

$\mathrm{rev}(b_k b_{k-1} \ldots b_1 b_0) = 0.b_1 b_2 \ldots b_k$



Roth's Lower Bound, and Questions

$D(P) = O(\frac{\log n}{n})$ is possible

$\Rightarrow$ we can estimate integrals with error $O\left(\frac{\log n}{n}\right)$

But what about $d(P) = O(1/n)$?

Roth's Lower Bound, and Questions

$D(P) = O(\frac{\log n}{n})$ is possible

$\Rightarrow$ we can estimate integrals with error $O\left(\frac{\log n}{n}\right)$

But what about $d(P) = O(1/n)$?

Theorem (Roth, 1954; Schmidt 1972)

For any n-point set P, $D(P) = \Omega(\frac{\log n}{n})$.

Roth's Lower Bound, and Questions

$D(P) = O(\frac{\log n}{n})$ is possible

$\Rightarrow$ we can estimate integrals with error $O\left(\frac{\log n}{n}\right)$

But what about $d(P) = O(1/n)$?

Theorem (Roth, 1954; Schmidt 1972)

For any n-point set P, $D(P) = \Omega(\frac{\log n}{n})$.

What about boxes in dimension 3? In dimension k?

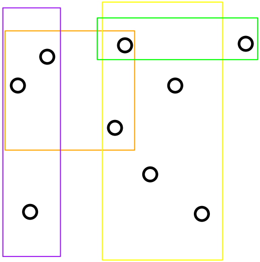
$$\frac{(\log n)^{(k-1)/2+\eta_k}}{n} \lesssim d(P) \lesssim \frac{(\log n)^{k-1}}{n}$$

for $\eta_k \to 0$ as $k \to \infty$.

Outline

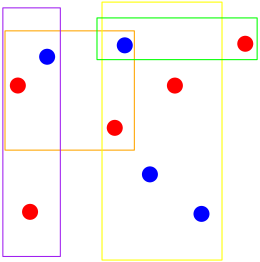
Tusnády's Problem

Given: Set Q of n points in the unit square
Goal: Color each point $p \in Q$ red or blue so that each rectangle R is as *balanced* as possible.



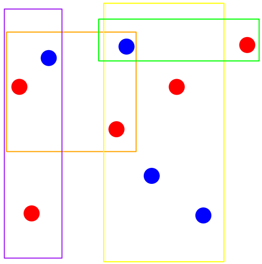
Tusnády's Problem

Given: Set Q of n points in the unit square
Goal: Color each point $p \in Q$ red or blue so that each rectangle R is as *balanced* as possible.



Tusnády's Problem

Given: Set Q of n points in the unit square
Goal: Color each point $p \in Q$ red or blue so that each rectangle R is as *balanced* as possible.



$$\mathrm{disc}(Q) := \min_{\chi} \max_{R} \Big| \sum_{p \in R \cap P} \chi(p) \Big|,$$

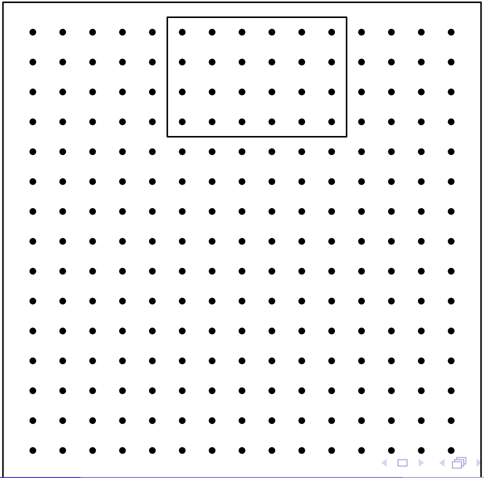
where $\chi \colon P \to \{-1, 1\}$ is a coloring.

From Combinatorial to Geometric Discrepancy

For any n there exists an n-point set P s.t.

$$d(P) \lesssim \frac{1}{n} \max_{Q} \mathrm{disc}(Q),$$

where Q ranges over n-point sets in the unit square.

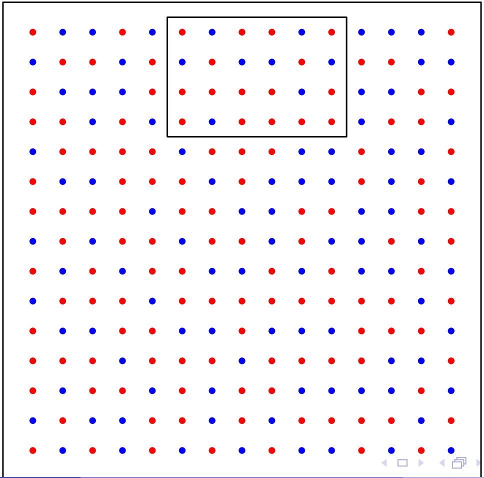


From Combinatorial to Geometric Discrepancy

For any n there exists an n-point set P s.t.

$$d(P) \lesssim \frac{1}{n} \max_{Q} \operatorname{disc}(Q),$$

where Q ranges over n-point sets in the unit square.

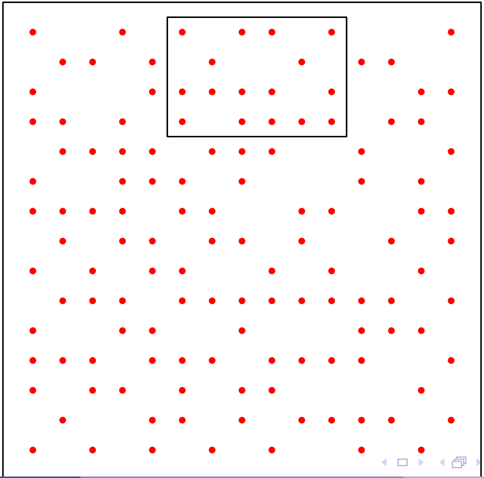


From Combinatorial to Geometric Discrepancy

For any n there exists an n-point set P s.t.

$$d(P) \lesssim \frac{1}{n} \max_Q \operatorname{disc}(Q),$$

where Q ranges over n-point sets in the unit square.

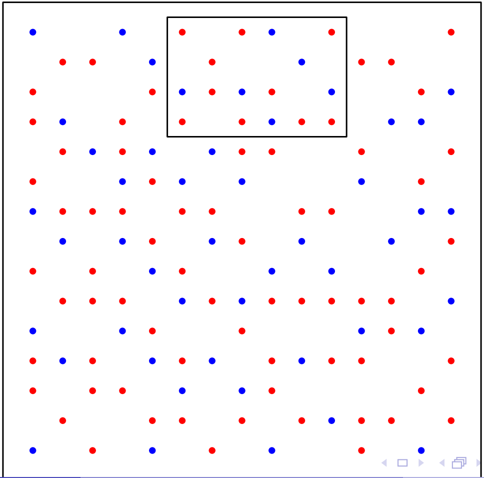


From Combinatorial to Geometric Discrepancy

For any n there exists an n-point set P s.t.

$$d(P) \lesssim \frac{1}{n} \max_Q \mathrm{disc}(Q),$$

where Q ranges over n-point sets in the unit square.

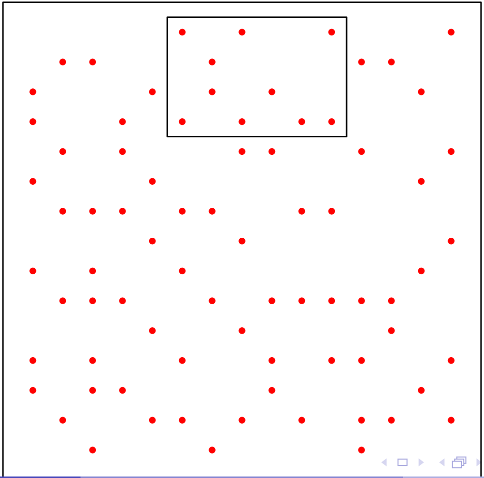


From Combinatorial to Geometric Discrepancy

For any n there exists an n-point set P s.t.

$$d(P) \lesssim \frac{1}{n} \max_{Q} \mathrm{disc}(Q),$$

where Q ranges over n-point sets in the unit square.



Bounds for Tusnády

Theorem (Nikolov, Matoušek, Talwar, 2014)

$$\log(n)^{d-1} \lesssim \max_{Q} \mathrm{disc}(Q) \lesssim \log(n)^{d+1/2}$$

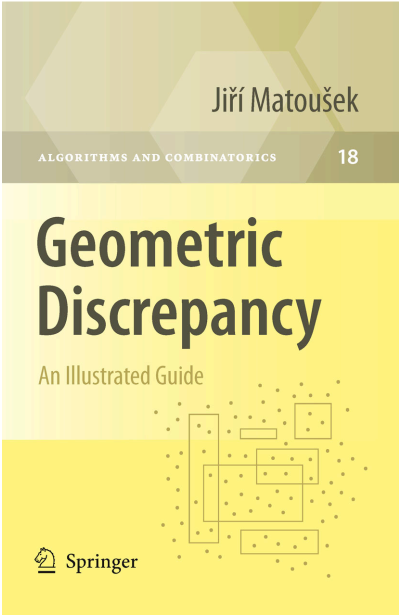
The proof uses (the analysis of) an algorithm to estimate discrepancy.

Computational Questions

- How can we *efficiently* (i.e. fast) find balanced colorings?
- Can we compute $\mathrm{disc}(Q)$?

This kind of balanced colorings problem has many other applications:

- computational geometry
- data structures
- approximation algorithms
- private data analysis.


Jiří Matoušek
ALGORITHMS AND COMBINATORICS
18
Geometric Discrepancy
An Illustrated Guide
Springer