On barycentric subdivision

Persi Diaconis and Laurent Miclo

11 juin 2009

Plan of the talk

Motivation

The barycentric subdivision cuts a given triangle among its medians to produce six new triangles.
Choose uniformly one of them and iterate : we obtain a Markov chain $(\triangle(n))_{n\in\mathbb{N}}$.

Anti-motivation : this is not a good procedure to obtain nice triangularizations, because the triangles have a tendency to become flatter and flatter.

Very incomplete history

Blackwell was very interested in this problem and found, through a trial and error approach with the help of the computer, a martingale argument to show that the above triangle-valued chain becomes flatter and flatter exponentially fast.

Diaconis and McMullen and Hough also gave results in this direction, through dynamical system arguments.

Our goal : to propose another probabilistic approach and to go further in the description of the asymptotic behavior.

Normalization

Up to similitude, we can assume that the longest edge of the triangles is given by $[(0,0),(1,0)]$, that $(0,0)$ is also adjacent to shortest edge and that the triangles are included into the upper plane. We denote by $(X_n, Y_n)_{n\in\mathbb{N}}$ the coordinates of the other vertex.

If $Y_0 = 0$, the initial triangle is said to be flat. Then $Y_n = 0$ for all $n \in \mathbb{N}$, and we denote $(Z_n)_{n\in\mathbb{N}}$ the Markov chain of the abscissas in this situation.

Let also $(A_n)_{n\in\mathbb{N}}$ be the greatest angles.

Results (1)

We begin by recovering a result of Blackwell, confirming the tendency to flatness :

Theorem 1

Almost surely (a.s.) the stochastic sequence $(Y_n)_{n\in\mathbb{N}}$ converges to zero exponentially fast : there exists a constant $\chi > 0$ such that a.s. :

$$\limsup_{n\to\infty} \frac{1}{n} \ln(Y_n) \quad \leq \quad -\chi$$

We will numerically estimate that we can take $\chi \approx 0.07$.

Results (2)

Next we give a new proof of a result due to McMullen and Diaconis :

Theorem 2

The sequence $(A_n)_{n\in\mathbb{N}}$ is converging to π in probability.

Despite the convergence in probability, this result will be more difficult to deduce than Theorem 1, because the abscissa chain is not converging :

Theorem 3

Almost surely, the limit set of $(X_n)_{n\in\mathbb{N}}$ is $[0, 1/2]$.

The isoperimetric functional

The isoperimetric value of a (non-trivial) triangle $\triangle$ is

$$I(\triangle) \quad := \quad \frac{\mathcal{P}(\triangle)}{\sqrt{\mathcal{A}(\triangle)}} \quad := \quad \frac{\text{perimeter}}{\sqrt{\text{area}}}$$

It is related to the characteristic coordinates (x, y) of $\triangle$ through :

$$\sqrt{y}/3 \;\leq\; (I(\triangle))^{-1} \;\leq\; \sqrt{y}$$

Proposition 4

Almost surely, we have $\limsup_{n\to\infty} I(\triangle(n)) = +\infty$.

A simple geometrical idea

The proof is very simple : let L_1, L_2 and L_3 the lengths of the edges of a triangle $\triangle$ and l_1, l_2 and l_3 the lengths of its medians. Then $\frac{l_1+l_2+l_3}{L_1+L_2+L_3}$ is minimal for $\triangle$ equilateral. We deduce that

Lemma 5

For any $n \in \mathbb{N}$, we have, with $\alpha := \frac{\sqrt{3}+1-\sqrt{6}}{\sqrt{6}} > 0$,

$$\mathbb{E}[I(\triangle(n+1))|\triangle(n)] \quad \geq \quad (1+\alpha)\, I(\triangle(n))$$

This comes from the fact that the areas of the triangles obtained by the barycentric subdivision of $\triangle$ are the same and that the total sum of the perimeters is $L_1 + L_2 + L_3 + 2(l_1 + l_2 + l_3)$.

No martingale argument

From the above inequality we cannot deduce the a.s. divergence of $(I(\triangle(n)))_{n\in\mathbb{N}}$, but, taking into account the isoperimetric inequality, we get

$$\forall\, n,m \in \mathbb{N}, \qquad \mathbb{P}[I_{n+m} \geq (1+\alpha)^m 2\sqrt{\pi}|\triangle(n)] \;\geq\; \frac{1}{6^m}$$

Thus by stochastic comparison with an independent Bernoulli sequence of parameter $1/6^m$, we get the above proposition saying that the triangle chain always comes as close as we want to the set of flat triangles.
Another simple computation shows the following preliminary bound : there exist $0 < a < b < +\infty$ such that

$$\forall\, n \in \mathbb{N}, \qquad aI(\triangle(n)) \;\leq\; I(\triangle(n+1)) \;\leq\; bI(\triangle(n))$$

Kernels (1)

The Markov kernels Q and M of $(X_n, Y_n)_{n\in\mathbb{N}}$ and $(Z_n)_{n\in\mathbb{N}}$ can be naturally coupled and written under the following iterated random function form,

$$\begin{aligned} Q((x,y),\cdot\,) &= \frac{1}{6}\sum_{i\in[\![1,6]\!]}\delta_{(x_i,y_i)} \\ M(x,\cdot\,) &= \frac{1}{6}\sum_{i\in[\![1,6]\!]}\delta_{z_i(x)} \end{aligned}$$

As $y\to 0_+$, they are close in the sense that

Kernels (2)

Lemma 6

There exists a constant $K > 0$ such that for any characteristic point (x, y),

$$\forall\, i \in [\![1, 6]\!], \qquad \left\{ \begin{array}{rcl} |x_i(x,y) - z_i(x)| & \leq & Ky \\ |y_i(x,y)| & \leq & K\sqrt{y} \end{array} \right.$$

In the neighborhood of flat triangles (1)

The isporimetry value of flat triangles is $+\infty$, nevertheless, for $i \in [\![1,6]\!]$, the ratios $I(\triangle_i)/I(\triangle)$ admit a limit as $y \to 0_+$ "and x remains fixed", denoted $G(i,x)$, with

$$
\begin{aligned}
G(1,x) &= \sqrt{\tfrac{2}{3}}(1+x), & G(2,x) &= \sqrt{\tfrac{1}{6}}(2-x) \\
G(3,x) &= \sqrt{\tfrac{3}{2}}(1-x), & G(4,x) &= \sqrt{\tfrac{2}{3}}(2-x) \\
G(5,x) &= \sqrt{\tfrac{2}{3}}(2-x), & G(6,x) &= \sqrt{\tfrac{3}{2}}
\end{aligned}
$$

So to get $\mathbb{E}[\ln(I(\triangle(n+1)/I(\triangle(n))|\triangle(n)=\triangle] > 0$ for nearly flat triangles $\triangle$, it is sufficient to show that the mapping $[0,1/2] \ni x \mapsto \sum_{i\in[\![1,6]\!]} \ln(G(i,x))$ only takes positive values.

In the neighborhood of flat triangles (2)

Unfortunately, this is wrong and we need to iterate :

Proposition 7

There exist a constant $\gamma > 0$ and a neighborhood $\mathcal{N}$ of the set of the flat triangles, such that for any $n \in \mathbb{N}$,

$$\forall \triangle \in \mathcal{N}, \quad \mathbb{E}[\ln(l(\triangle(n+2))/l(\triangle(n)))|\triangle(n) = \triangle] \quad \geq \quad \gamma$$

Indeed, it is sufficient to check that the mapping

$$[0, 1/2] \ni x \quad \mapsto \quad \frac{1}{36} \sum_{i,j \in [\![1,6]\!]} \ln(G(j, z_i(x))G(i, x))$$

is positive. We used the computer to show that numerically. Since the exponential of this function is a polynomial mapping by pieces, this can be done algebraically via Sturm sequences.

A martingale argument (1)

Let $\gamma > 0$ and $A > 0$ be two given constants. Assume that for any R large enough, we are given a stochastic chain $(V_n^{(R)})_{n\in\mathbb{N}}$ and a martingale $(N_n^{(R)})_{n\in\mathbb{N}}$, satisfying $V_0^{(R)} = R$, $N_0^{(R)} = 0$ and such that for any time $n \in \mathbb{N}$,

$$\begin{aligned} \left|N_{n+1}^{(R)} - N_n^{(R)}\right| &\leq A \\ V_{n+1}^{(R)} - V_n^{(R)} &\geq \gamma + N_{n+1}^{(R)} - N_n^{(R)} \end{aligned}$$

Lemma 8

There exists a constant $A' > 0$ only depending on A and γ :

$$\mathbb{P}[\exists\ n \in \mathbb{N} : V_n^{(R)} < R/2] \leq \exp(-\gamma R/A') \frac{1}{1 - \exp(-\gamma^2/(4A'))}$$

A martingale argument (2)

Furthermore, we have a.s.

$$\liminf_{n\to\infty} \frac{V_n^{(R)}}{n} \geq \gamma$$

These results follow from standard considerations, via exponential martingale bounds and the iterated logarithm law.

The latter lemma is applied with $V^{(R)}$ a sequence of the kind $(\ln(I(\triangle(2n))))_{n\in\mathbb{N}}$, appropriately started and stopped. Choosing conveniently R (corresponding to a neighborhood of the set of flat triangles), we get from our preliminary bounds :

A.s. convergence

Proposition 9

Let $\mathcal{N}' := \{\triangle\ :\ \ln(I(\triangle)) > R_1\}$. *There exists a large enough constant* $R_2 \geq 2R_1$ *such that for any finite stopping time* T *satisfying* $\ln(I(\triangle(T))) \geq R_2$, *we have*

$$\mathbb{P}[\exists\ n \in \mathbb{N}\ :\ \triangle(T+n) \notin \mathcal{N}' | \mathcal{T}_T] \quad < \quad 1/2$$

Furthermore on the event $\{\forall\ n \in \mathbb{N}\ :\ \triangle(T+n) \in \mathcal{N}'\}$, *we have a.s.*

$$\liminf_{n\to\infty} \frac{\ln(I_n)}{n} \quad \geq \quad \gamma/2$$

Then Theorem 1 with $\chi = \gamma/2$ follows from Proposition 4.

Ergodicity

To go further, we need to show that Z is ergodic : M admits a unique attracting invariant probability μ, namely satisfying $\mu M = \mu$ and for any probability ν on $[0, 1/2]$, $\lim_{n\to\infty} \nu M^n = \mu$.
We also would like an estimation of the speed of convergence in Wasserstein distance (because in the end we will need to couple X and Z).
Barnsley and Elton (1988) gave such results under the assumption that there exists $r < 0$ such that

$$\forall\, x \neq y \in [0, 1/2], \qquad \sum_{i\in[\![1,6]\!]} \frac{1}{6} \ln\left(\frac{|z_i(y) - z_i(x)|}{|y - x|}\right) \quad \leq \quad r$$

Iterated random functions

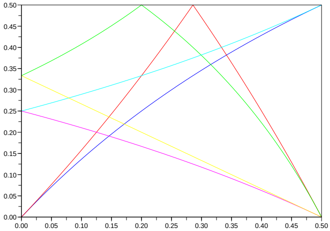

FIG.: Graphs of the z_i, for $i \in [\![1, 6]\!]$

On Barnsley and Elton's criterion

This criterion implies that

$$\sup_{x\in[0,1/2]} \sum_{i\in[\![1,n]\!]} \frac{1}{6}\ln(|z_i'|(x)) \quad \leq \quad r$$

but in general this is not a sufficient condition for ergodicity. Nevertheless, taking into account that our random functions are piecewise homographical mappings, we can go back. Except that we have to apply these considerations to the iterated mappings $z_i \circ z_j$, for $i, j \in [\![1, 6]\!]$. Curiously, we end up with a computation already encountered, since

$$\forall\, i \in [\![1,6]\!],\ \forall\, x \in [0,1/2], \qquad \left|z_i'\right|(x) \quad = \quad \frac{1}{G^2(i,x)}$$

On the invariant probability μ

We need two other facts :

Proposition 10

The probability μ contains no atom, in particular $\mu(\{0\}) = 0$. The support of μ is the whole segment $[0, 1/2]$.

The second assertion is an immediate consequence of a result due to Dubins and Freedman (1966), saying that the support of μ is the whole state space if it can be covered by the images of the functions z_i which are strict contractions.
The first assertion would be a consequence of the same paper, if the iterated random functions were one-to-one. A more careful investigation is necessary in our case.

Couplings

The criterion of Barnsley and Elton applied in our situation implies that there exist $q \in (0, 1]$ and $\rho \in (0, 1)$ such that for any $n \in \mathbb{N}$, we can construct a coupling such that

$$\mathbb{E}[|Z_{n+2} - Z'_{n+2}|^q | Z_n, Z'_n] \quad \leq \quad \rho |Z_n - Z'_n|^q$$

where Z and Z' are Markov chains with M as kernel.
A similar coupling gives, for a constant $K' > 0$,

$$\mathbb{E}[|Z_{n+2} - X_{n+2}|^q | Z_n, X_n, Y_n] \quad \leq \quad \rho |Z_n - X_n|^q + K' Y_n^{q/2}$$

It follows easily that the random variable $|X_n - Z_n|$ converges in probability to zero for large time $n \in \mathbb{N}$.

Convergence in probability

Due to our renormalization and to Theorem 1, the convergence in probability of the angle A_n is equivalent to the convergence of Y_n/X_n toward zero. Let $\epsilon, \eta > 0$ be given and write for any $n \in \mathbb{N}$,

$$\begin{aligned} \mathbb{P}[Y_n/X_n \geq \epsilon] &\leq \mathbb{P}[Y_n \geq 2\epsilon\eta] + \mathbb{P}[X_n \leq 2\eta] \\ &\leq \mathbb{P}[Y_n \geq 2\epsilon\eta] + \mathbb{P}[|X_n - Z_n| \leq \eta] + \mathbb{P}[Z_n \leq \eta] \end{aligned}$$

Taking into account that the attractive probability μ of Z is continuous, we get by letting n going to infinity

$$\limsup_{n\to\infty} \mathbb{P}[Y_n/X_n \geq \epsilon] \leq \mu([0,\eta])$$

and the wanted result follows by letting η going to zero.

Limit set of X

The bounds obtained on the ergodicity of Z implies that for any open set included in $[0,1/2]$, we can find $\eta > 0$ and $N \in \mathbb{N}^*$ such that

$$\inf_{z\in[0,1/2]} \mathbb{P}_z[Z_N \in \mathcal{O}] \quad \geq \quad \eta$$

Using what we have already seen, this can be translated into the fact that there exist $\eta' > 0$ and $N' \in \mathbb{N}^*$ such that

$$\inf_{(x,y)\in\mathcal{D}} \mathbb{P}_{(x,y)}[X_{N'} \in \mathcal{O}'] \quad \geq \quad \eta' \tag{1}$$

where $\mathcal{D}$ is the set of characteristic points.

Simulation of μ

Here is an approximation of μ using the strong law of large numbers with $(Z_n)_{0\leq n\leq 100000}$. The following histogram uses 100 bars.

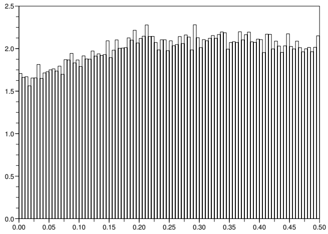

FIG.: An approximation of μ

Evaluation of χ

Instead of iterating twice the kernel M, we can iterate it N times and in the limit as N goes to infinity, we get an evaluation of the constant χ of Theorem 1 :

$$\begin{aligned}\chi &\approx \lim_{N\to\infty} \frac{1}{N} \sum_{n\in[\![0,N-1]\!]} \mathbb{E}_x\left[\ln(G(I_{n+1}, Z_n))\right] \\ &= \frac{1}{6} \sum_{i\in[\![1,6]\!]} \int \ln(G(i,x))\,\mu(dx)\end{aligned}$$

Using the above simulation of μ, we get $\chi \approx 0.07$.