Alternative Methods of Fractal Analysis

Jaromir Kukal

FNSPE, CTU in Prague

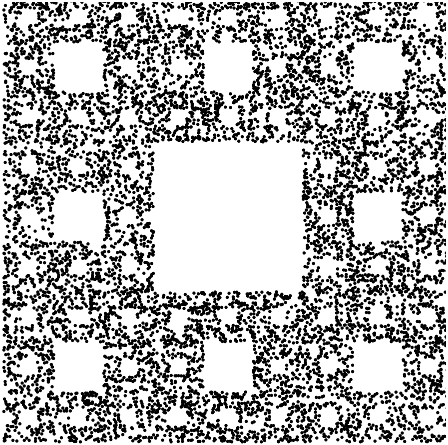
August 25, 2024

Motivation of Alternative Point Set Analysis

- Analysis of large 2D and 3D images
- Edge structure from edge detector
- Strucrure of image segmentation
- Structure of watershed
- Structure characterization by dimension value or values
- Correlation sum criticism:
  - quadratic time complexity
  - oscillations on log-log graph
- Box Counting criticism:
  - biased estimate of capacity dimension
  - sensitivity to grid translation and rotation

Uniform Sampling from Fractal Set

- Investigated structure $\mathcal{F} \subset \mathbb{R}^n$
- Uniform sampling $\mathbf{x} \sim \mathrm{U}(\mathcal{F})$
- Sample $\Phi = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_M\} \subset \mathcal{F}$



World of Entropies and Dimensions

- Renyi entropy decreases with grid spacing

$$H_\alpha = C_\alpha - D_\alpha \ln \epsilon$$

- H_0 is Hartley entropy
- H_1 is information entropy
- H_2 is collision entropy
- H_∞ is min-entropy
- D_0 is capacity dimension
- D_1 is information dimension
- D_2 is correlation dimension
- Fractal dimension order

$$D_\mathrm{T} \le D_\infty \le D_2 \le D_1 \le D_\mathrm{H} \le D_0 \le D_\mathrm{E} \le n$$

- Similarity dimension

$$D_\mathrm{SIM} = \frac{\log P}{-\log Q} = D_\mathrm{H} = H_\alpha, \forall \alpha \ge 0$$

New Method for Correlation Dimension Estimation

- Traditional correlation sum

$$\mathrm{C}(r) = \mathop{\mathrm{E}}_{\mathbf{x},\mathbf{y}\sim \mathrm{U}(\mathcal{F})} \mathrm{I}(||\mathbf{x}-\mathbf{y}|| \leq r)$$

- Estimation of D_2 using linearized regression

$$\ln \mathrm{C}(r) = A + D_2 \ln r$$

- Novel method is inspired by X-ray powder crystalography
- Debye and Scherrer performed power spectra averaging over SO(3) rotations of periodic structures
- We extend the point coordinates to $\mathbf{x} = (x_1, \ldots, x_n, 0, 0, \ldots) \in \mathcal{H}$ and perform power spectra averaging over SO(∞) rotations
- Average power spectrum oscillations are suppressed

Dlask, M., Kukal, J. Application of Rotational Spectrum for Correlation Dimension Estimation. Chaos, Solitons & Fractals, Elsevier, 99, 256-262, 2017.

Rotation of Power Spectrum in $\mathbb{R}^n$

- Fourier transform in $\mathbb{R}^n$ using angular frequency $\mathbf{k} \in \mathbb{R}^n$

$$\mathrm{F}(\mathbf{k}) = \int_{\mathbb{R}^n} \mathrm{f}(\mathbf{x}) \exp(-\mathrm{j}\mathbf{k} \cdot \mathbf{x}) \mathrm{d}\mathbf{x}$$

- Fourier transform of point set $\mathcal{F}$

$$\mathrm{F}(\mathbf{k}) = \mathop{\mathrm{E}}_{\mathbf{x} \sim \mathrm{U}(\mathcal{F})} \exp(-\mathrm{j}\mathbf{k} \cdot \mathbf{x})$$

- Adequate power spectrum

$$\mathrm{P}(\mathbf{k}) = |\mathrm{F}(\mathbf{k})|^2 = \mathop{\mathrm{E}}_{\mathbf{x},\mathbf{y} \sim \mathrm{U}(\mathcal{F})} \exp(-\mathrm{j}\mathbf{k} \cdot (\mathbf{x} - \mathbf{y}))$$

- Hypersphere $\mathcal{S}_{n-1} = \{\mathbf{x} \in \mathbb{R}^n : ||\mathbf{x}|| = 1\}$
- Support for SO(n) rotations

$$\mathbf{k} = \Omega \mathbf{u}, \quad \mathbf{x} - \mathbf{y} = ||\mathbf{x} - \mathbf{y}|| \mathbf{v}, \quad \mathbf{u}, \mathbf{v} \in \mathcal{S}_{n-1}$$

Limit Case of Power Spectrum Rotation

- Averaged power spectrum for $\Omega = ||\mathbf{k}||$

$$\mathrm{S}(\Omega) = \mathop{\mathrm{E}}_{\mathbf{x},\mathbf{y}\sim\mathrm{U}(\mathcal{F})} \mathop{\mathrm{E}}_{\mathbf{u},\mathbf{v}\sim\mathrm{U}(\mathcal{S}_{n-1})} \exp(-\mathrm{j}\Omega||\mathbf{x}-\mathbf{y}||\mathbf{u}\cdot\mathbf{v})$$

- Resulting formula in $\mathbb{R}^n$

$$\mathrm{S}(\Omega) = \mathop{\mathrm{E}}_{\mathbf{x},\mathbf{y}\sim\mathrm{U}(\mathcal{F})} \mathrm{H}_n(\Omega||\mathbf{x}-\mathbf{y}||)$$

$$\mathrm{H}_n(q) = \frac{2^{n/2-1}\Gamma(n/2)}{q^{n/2-1}} \mathrm{J}_{n/2-1}(q)$$

- Useful limit case

$$\lim_{n\to+\infty} \mathrm{H}_n(t\sqrt{n}) = \exp(-t^2/2)$$

Rotation of Power Spectrum in Hilbert Space $\mathcal{H}$

- Point coordinate extension

$$\mathbf{x} = (x_1, \ldots, x_n, 0, 0, \ldots) \in \mathcal{H}$$

- Resulting formula in $\mathcal{H}$

$$\mathrm{S}(\Omega) = \mathop{\mathrm{E}}_{\mathbf{x},\mathbf{y} \sim \mathrm{U}(\mathcal{F})} \exp(-\Omega^2 ||\mathbf{x} - \mathbf{y}||^2 / 2)$$

- Relationship to correlation dimension

$$\lim_{\Omega \to +\infty} \frac{\ln \mathrm{S}(\Omega)}{\ln \Omega} = -D_2$$

- Linearized regression for unbiased D_2 estimation

$$\ln \mathrm{S}(\Omega) = B - D_2 \ln \Omega$$

Resulting Monte Carlo I.

Perform *NMC*-times for given $\Omega > 0$:

- Generate different sample indices $i, j \sim \mathrm{U}(\{1, \ldots, M\})$
- Their distance is $d = ||\mathbf{x}_i - \mathbf{x}_j||$
- Evaluate pair contribution as $\exp(-\Omega^2 d^2/2)$

Estimate $\mathrm{S}(\Omega)$ as adequate mean value E

Using results for various Ω values, we employ linear regression and estimate D_2 as

$$\ln \mathrm{S}(\Omega) = B - D_2 \ln \Omega$$

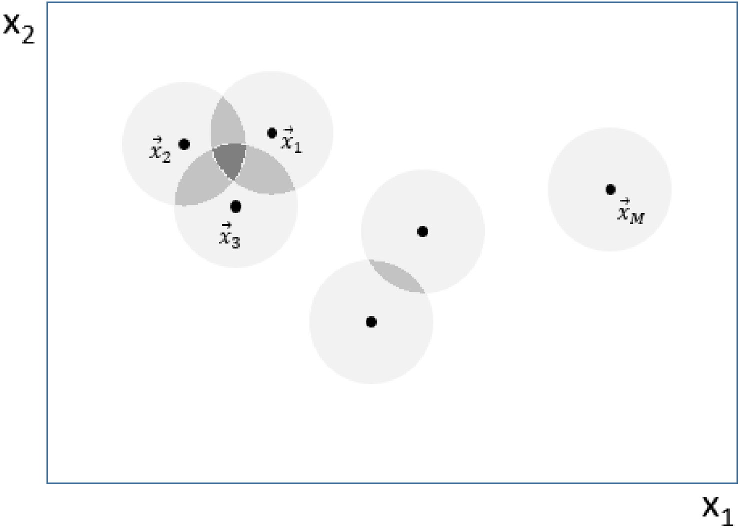
TSRM invariant estimation of correlation dimension D_2

New Method of Renyi Dimension Estimation

- Boxcounting offers biased estimate of capacity dimension D_0
- Particle counting for biased estimation of D_α
- Particle counting is sensitive to grid translation and rotation
- Investigation of finite sample from given uncountable set
- Basic tools: Parzen density estimate & Renyi entropy
- New term: Modified Renyi entropy H_α^*
- H_α^* is translation, rotation and mirroring invariant
- Monte Carlo estimation of H_α^* and D_α^*
- D_α^* estimate can be also scaling invariant
- Effective implementation of ϵ-querry using k-d tree
- Proven relationships: $D_0^* = D_0, \quad D_2^* = D_2$

Dlask, M., Kukal, J., Translation and Rotation Invariant Method of Renyi Dimension Estimation, Chaos, Solitons & Fractals, Elsevier, 114(C), 536-541, 2018.

TRM Invariant Parzen Estimate



Trivial Formulas First

- n-dimensional ball $\mathcal{B}(\mathbf{x},\epsilon)=\{\mathbf{y}\in\mathbb{R}^n : \|\mathbf{y}-\mathbf{x}\|_2\leq\epsilon\}$
- Ball volume $V_* = V_n\cdot\epsilon^n$
- Kernel density $\mathrm{f}_0(\mathbf{x},\epsilon)=\mathrm{I}(\|\mathbf{x}\|_2\leq\epsilon)/V_*$
- Parzen formula $\mathrm{f}(\mathbf{x},\Phi,\epsilon)=M^{-1}\sum_{k=1}^{M}\mathrm{f}_0(\mathbf{x}-\mathbf{x}_k,\epsilon)$
- Point degeneracy $\mathrm{G}(\mathbf{x},\Phi,\epsilon)=\sum_{k=1}^{M}\mathrm{I}(\|\mathbf{x}-\mathbf{x}_k\|_2\leq\epsilon)$
- Simplified PDF

$$\mathrm{f}(\mathbf{x},\Phi,\epsilon)=\frac{1}{M\cdot V_*}\sum_{k=1}^{M}\mathrm{I}(\|\mathbf{x}-\mathbf{x}_k\|_2\leq\epsilon)=\frac{\mathrm{G}(\mathbf{x},\Phi,\epsilon)}{M\cdot V_*}$$

- Minkowski sausage $\mathcal{S}\approx\mathcal{S}_M=\bigcup_{k=1}^{M}\mathcal{B}(\mathbf{x}_k,\epsilon)$
- Degeneracy range $\mathrm{G}(\mathbf{x},\Phi,\epsilon)\in\{1,\ldots,M\}$ for $\mathbf{x}\in\mathcal{S}_M$

Renyi Entropy

- Exponent range $\alpha \in [0,1) \cup (1,\infty)$
- Renyi entropy
  $H_\alpha = \frac{1}{1-\alpha} \ln \mathrm{E}\ p^{\alpha-1}$
- Hartley entropy $H_0 = \ln \mathrm{E}\ p^{-1}$
- Shannon entropy $H_1 = -\mathrm{E}\ \ln p$
- Collision entropy $H_2 = -\ln \mathrm{E}\ p$
- Min-entropy $H_\infty = -\ln \sup p$
- Event probabilities from grid
- Biased estimation

Novelty: Modified Renyi Entropy

- Main integral

$$\mathrm{J}(\Phi,\alpha,\epsilon)=\int_{\mathbf{x}\in\mathbb{R}^n}\mathrm{f}^{\alpha}(\mathbf{x},\Phi,\epsilon)\mathrm{d}\mathbf{x}$$

- Referential integral

$$\mathrm{J}_0(\alpha,\epsilon)=\int_{\mathbf{x}\in\mathbb{R}^n}\mathrm{f}_0^{\alpha}(\mathbf{x},\epsilon)\mathrm{d}\mathbf{x}=V_*^{1-\alpha}$$

- Modified Renyi entropy

$$\mathrm{H}_{\alpha}^{*}(\Phi,\epsilon)=\frac{\ln \mathrm{J}(\Phi,\alpha,\epsilon)-\ln \mathrm{J}_0(\alpha,\epsilon)}{1-\alpha}$$

- Relationship to degeneracy

$$H_{\alpha}^{*}=\ln M+\frac{\ln \mathrm{E}\ G^{\alpha-1}}{1-\alpha}$$

Particular Cases and Modified Renyi Dimension

- Modified Hartley entropy
  $H_0^* = \ln M + \ln \mathrm{E}\ G^{-1}$
- Modified Shannon entropy
  $H_1^* = \lim_{\alpha \to 1} H_\alpha^* = \ln M - \mathrm{E}\ \ln G$
- Modified collision entropy
  $H_2^* = \ln M - \ln \mathrm{E}\ G$
- Modified min-entropy
  $H_\infty^* = \lim_{\alpha \to \infty} H_\alpha^* = \ln M - \ln \max G$
- Modified Renyi dimension
  $D_\alpha^* = \lim_{\epsilon \to 0^+} \frac{\mathrm{H}_\alpha^*(\epsilon)}{-\ln \epsilon}$

Resulting Procedure

Perform *NMC*-times for given α and fixed $\epsilon > 0$:

- Generate sample index $j \sim \mathrm{U}(\{1, \ldots, M\})$
- Generate point $\mathbf{x} \sim \mathrm{U}(\mathcal{B}(\mathbf{x}_j, \epsilon))$
- Calculate degeneracy $G = \sum_{k=1}^{M} \mathrm{I}(\|\mathbf{x} - \mathbf{x}_k\|_2 \leq \epsilon)$

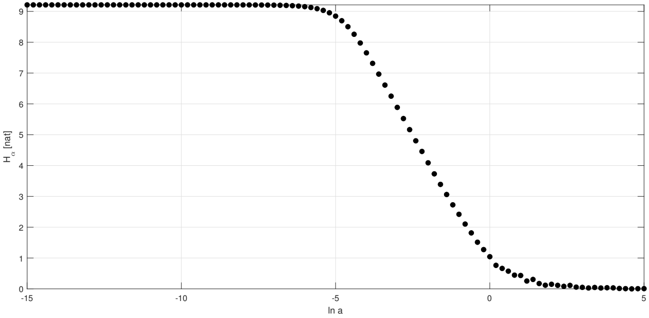
Estimate adequate mean value E
Estimate modified Renyi entropy D_α^*
Using results for various ϵ values, we employ linear regression and estimate D_α^* as

$$\mathrm{H}_\alpha^*(\epsilon) = C_\alpha - D_\alpha^* \ln \epsilon$$

TSRM invariant estimation of Renyi dimension D_α^*

How to Suppress Bias of D^*_α and Save Its MSE?

- Range of $H^*_\alpha \in [0, \ln M]$
- Fitting near inflex point
- Suggested setting $N \geq 10^7$, $NMC \geq 10^6$



How to Suppress Time Complexity?

- Effective implementation of ϵ-querry
- Batch construction of k-d tree with $\mathrm{T}(N) = \mathrm{O}(N \log^2 N)$
- Single querry complexity $\mathrm{T}(N, G) = \mathrm{O}(\log N) + G$
- Use only small values of ϵ
- Degeneracy constrain $G < N^\beta$ with $\beta \in (1/2, 3/4)$

Thank you for your attention