

Optymalizacja globalna

Adam Przemysław Chojecki

Mini PW

09.03.2023

Najłatwiej zacząć cytatem

Zgadywanie jest tanie.
Błędne zgadywanie jest
kosztowne

stare chińskie przysłowie

No Free Lunch

Learning Theory: Można wnioskować o zachowaniu całej funkcji na podstawie próbki.

Matematyk: Nie można :)

Learning Theory: No ok, ale można zamodelować rozkład, który często będzie dawał niewielkie błędy.

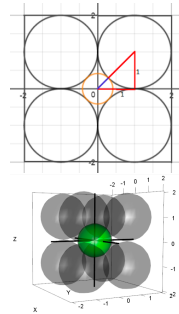
Matematyk: No Free Lunch: Średnio wyjdzie nic XDDD

Learning Theory: Nie chcemy średniej po wszystkich, tylko po tych, które mnie interesują 😎

Na podstawie [4] 5.2.1

Przekleństwo wymiarowości (Ang. Curse of Dimensionality)

- Ogromna liczba wymiarów jest typowym setupem. Analiza obrazów. Badania genetyczne itp.
- Ciekawie o Curse of Dimensionality: [8] oraz [4] (5.11.1)
- Środkowa kula w wielu wymiarach wypełnia coraz większą przestrzeń. Już w 10D środkowa kula wychodzi poza pudełko

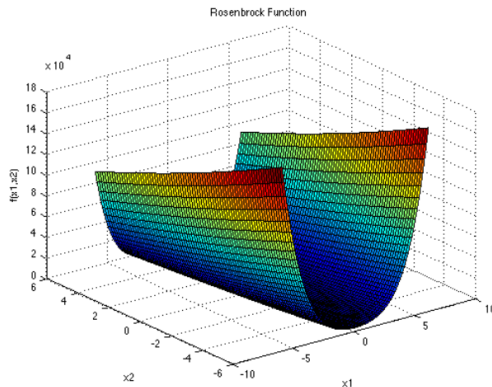


Rysunek: Obrazki wzięte z [6]

Błogosławieństwo wymiarowości

Jeśli mamy jakiś dobry pomysł na interpretowanie tych nowych wymiarów, to działa to na naszą korzyść (podobnie jak więcej obserwacji).

Większość obserwowanych minimów lokalnych jest podobnej jakości.

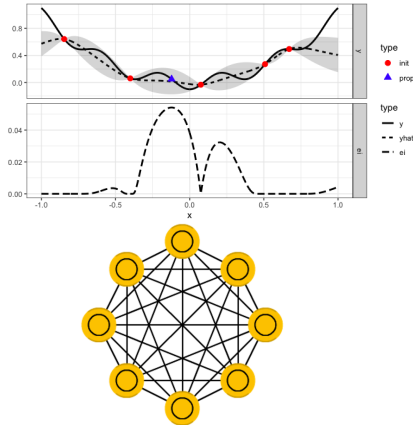


$$f(\mathbf{x}) = \sum_{i=1}^{d-1} [100(x_{i+1} - x_i^2)^2 + (x_i - 1)^2]$$

Rysunek: Obrazek wzięty z [0]

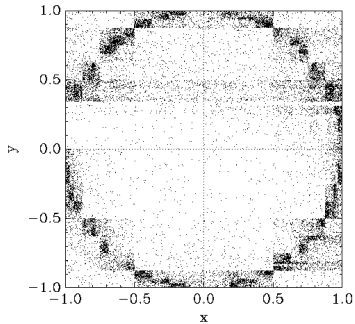
Przykładowe setupy optymalizacyjne

- Algorytmy Ewolucyjne
- Algorytmy Metaheurystyczne
- Monte Carlo Markov Chain
- Model-Based; Optymalizacja Bayesowska
- Problem SAT
- Wiele innych
- Sieci Hopfielda

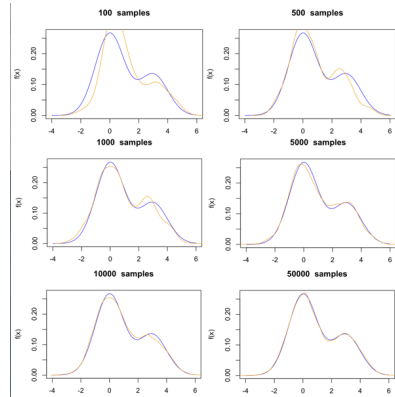


Rysunek: Obrazki wzięte z [2], [1]

Monte Carlo



Rysunek: Obrazek wzięty z [10]



Rysunek: Obrazek wzięty z [9]

Tuning hiperparametrów

The virgin Grid / Random Search

Chodzi całkiem losowo,
albo całkiem
przewidywalnie



Potrzebuje dużo czasu,
aby cokolwiek znaleźć

Brak weny i nudny

Nie interesują go wartości
funkcji jaką optymalizuje
WTF?

THE CHAD MODEL BASED OPTIMIZATON

Każdy krok
dogłębnie
przemyślany i
optymalny



Wystarczy kilka obserwacji i
już masz super hiperparametry

Niesamowicie ciekawy

Jedyne co go interesuje
to wartości funkcji jaką
optymalizuje

Rysunek: Na podstawie dokumentacji mlrMBO [2]

Problem SAT

Są gotowe biblioteki sprawdzające spełnialność formuł logicznych.
Każdy problem NP można sprowadzić do problemu SAT.

Po co powstała sieć Hopfielda

Oryginalne zastosowanie sieci Hopfielda pięknie opisane w tym filmie na YT z okazji SoME2: [3]. Okazuje się, że sieć Hopfielda minimalizuje pewną formę kwadratową.

$$E = \sum_{i=1}^n \sum_{j=1}^n t_{ij} v_i v_j - \sum_{i=1}^n l_i v_i$$

Użyteczna w optymalizacji

Możemy znaleźć taką formę, która odpowiada naszemu problemowi i zbudować sieć ją minimalizującą.

Np. dla problemu komiwojażera możemy wybrać $E = E_1 + E_2$, gdzie:

$$E_1 = \sum_{x=1}^n \sum_{\substack{i,j=1 \\ i \neq j}}^n v_{xi} v_{xj} + \sum_{x=1}^n \sum_{\substack{i,j=1 \\ i \neq j}}^n v_{ix} v_{jx} + \left(\left(\sum_{x=1}^n \sum_{i,j=1}^n v_{ij} \right) - n \right)^2$$

$$E_2 = \sum_{x=1}^n \sum_{\substack{i,j=1 \\ i \neq j}}^n d_{ij} v_{ix} v_{j,x+1}$$

Na podstawie [5] (4.12 i 4.16)

Bibliografia I

- [1] *Artykuł o sieciach hopfielda na stronie primo.ai.* [http://primo.ai/index.php?title=Hopfield_Network_\(HN\)](http://primo.ai/index.php?title=Hopfield_Network_(HN)). [Online; 09-03-2023].
- [2] *Dokumentacja rozszerzenia MBO do mlr.* <https://mlrmbo.mlr-org.com>. [Online; 09-03-2023].
- [3] *Film an YT z okazji SoME2 opowiadający o modelu Hopfielda.* <https://youtu.be/piF6D6CQxUw>. [Online; 09-03-2023].
- [4] Ian Goodfellow, Yoshua Bengio i Aaron Courville. *Deep Learning*. <http://www.deeplearningbook.org>. MIT Press, 2016.

Bibliografia II

- [5] Jacek Mańdziuk. *Sieci neuronowe typu Hopfielda*.
https://pages.mini.pw.edu.pl/~mandziukj/PRACE/Sieci_Hopfielda-EXIT.pdf. [Online; 09-03-2023].
- [0] *Opis funkcji Rosenbrocka w Virtual Library of Simulation Experiments*.
<https://www.sfu.ca/~ssurjano/rosen.html>. [Online; 09-03-2023].
- [6] psychometroskar. *Wpis na blogu o kulach wpisanych w wielowymiarową kostkę*.
<https://psychometroskar.com/imagining-multiple-dimensions-is-hard/>. [Online; 09-03-2023].

Bibliografia III

- [8] Michel Verleysen i Damien François. „The Curse of Dimensionality in Data Mining and Time Series Prediction”. W: *Computational Intelligence and Bioinspired Systems*. Red. Joan Cabestany, Alberto Prieto i Francisco Sandoval. Berlin, Heidelberg: Springer Berlin Heidelberg, 2005, s. 758–770. ISBN: 978-3-540-32106-4.
- [9] *Wikipedia o Algorytmie Metropolisa-Hastingsa*. https://pl.wikipedia.org/wiki/Algorytm_Metropolisa-Hastingsa. [Online; 09-03-2023].
- [10] *Wikipedia o całkowaniu Monte Carlo*. https://en.wikipedia.org/wiki/Monte_Carlo_integration. [Online; 09-03-2023].

Dziękuję za uwagę

Najbardziej polecam przeczytać [2], [3], [4].