Applications of MC methods

Marcin Chrząszcz
mchrzasz@cern.ch

University of Zurich[UZH]

Monte Carlo methods,
26 May, 2016

Optimization

Optimization problem:

We have a set $X \subset \mathbb{R}^m$ and a function $F: X \to \mathbb{R}$.
Task:
Find the optimum point:

$$x_{opt} \in X : \forall_{x \in X} \; F(x) \geqslant F(x_{opt})$$

$\Rightarrow$ This is completely different then normal function minimalization as we choose x_{opt} from a set X. This makes a big differences for numerical computations.

$\Rightarrow$ The MC algorithms for solving this problem:

- Hit and miss method - the simplest and the slowest.
- Sequence methods - MC interpretation of method of further approximations.
- Genetic methods, stat. optimization.

Optimization hit and miss

⇛ The algorithm acts as follows:

- We generate N points $x_1, ..., x_N \in X$ from a constant p.d.f. on X.
- We calculate the F function value in the points $x_1, ..., x_N$:

$$F_1 = F(x_1), \quad F_2 = F(x_2), \quad ..., \quad F_N$$

- We calculate $F^* = \min\{F_1, F_2, ..., F_N\}$.
- The solution is $x_j : F(x_j) = F^*$

⇛ Precision:

If $F^* = \min_{1 \leqslant j \leqslant N}\{F(x_j)\}$ where $x_j = 1, 2, ...N$ are random points from uniform p.d.f. on X. Then with the probability $1 - (1-\gamma)^N$ the volume of points x for which $F(x) < F^*$ is smaller then ϵ.

⇛ We can say that with probability $1 - (1-\gamma)^N$ the points x_{opt} was localized with the probability γ.

⇛ the smaller the volume the better the accuracy.

Optimization hit and miss

⇛ The algorithm acts as follows:

- We generate N points $x_1, ..., x_N \in X$ from a constant p.d.f. on X.
- We calculate the F function value in the points $x_1, ..., x_N$:

$$F_1 = F(x_1), \quad F_2 = F(x_2), \quad ..., \quad F_N$$

- We calculate $F^* = \min\{F_1, F_2, ..., F_N\}$.
- The solution is $x_j : F(x_j) = F^*$

⇛ Precision:

If $F^* = \min_{1 \leqslant j \leqslant N}\{F(x_j)\}$ where $x_j = 1, 2, ...N$ are random points from uniform p.d.f. on X. Then with the probability $1 - (1-\gamma)^N$ the volume of points x for which $F(x) < F^*$ is smaller then ϵ.

⇛ We can say that with probability $1 - (1-\gamma)^N$ the points x_{opt} was localized with the probability γ.
⇛ the smaller the volume the better the accuracy.

Optimization hit and miss

The smallest N that obeys: $1-(1-\gamma)^N \geqslant 1-\epsilon$:

γ	$1-\epsilon$					
	0.5	0.9	0.95	0.99	0.999	0.9999
0.5	1	4	5	7	10	14
0.1	7	22	29	44	66	88
0.05	14	45	59	90	135	180
0.01	69	230	299	459	688	917
0.001	693	2302	2995	4603	6905	9206
0.0001	6932	23025	29956	46050	69075	92099

$\Rrightarrow$ Example: $[0,1]^m$ and $F: X \rightarrow \mathbb{R}$.
How many points we need to generate to ahve 0.9 probability to be have half of the range of each direction in each precision:

- For $m=1$: $\gamma = 1/2 \rightarrowtail N=4$.
- For $m=2$: $\gamma = 1/4 \rightarrowtail N=9$.
- For $m=14$: $\gamma = 2^{-14} \rightarrowtail N > 23\,000$.

$\Rrightarrow$ Inefficient for multi-dimensions.

Optimization sequence

⇛ The algorithm:

- We choose the starting point $x_1 \in X$ from some p.d.f. on X set.
- After generating $x_1, x_2, ..., x_n$ check if some conditions are meet.
  - If YES then we stop and we put x_n as solution.
  - If NO then we generate x_{n+1} form a p.d.f. that depends on already generated points.

⇛ The basic sequence algorithm:

- Choose x_1.
- After we have $x_1, x_2, ..., x_n$ then we generate a temporary point ξ_n:

$$x_{n+1} = \begin{cases} x_n, & \text{if } F(x_n + \xi_n) \geqslant F(x_n) - \epsilon \\ x_n + \xi_n, & \text{if } F(x_n + \xi_n) < F(x_n) - \epsilon \end{cases}$$

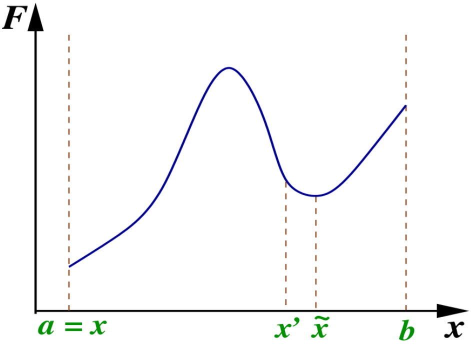
Optimization sequence

⇛ From the above algorithm we will get a sequence:

$$F(x_1) \geqslant F(x_2) \geqslant F(x_3) \geqslant ... \geqslant F(x_n) \geqslant F(x_{n+1})...$$

⇛ If the function is bounded from the bottom the the above sequence is converging.

⇛ How can we be sure it will converge to x_{opt}?



⇛ If we choose the correct the P_n every sequence starting from x' will converge to $\overline{x}$, where F has a local minimum.

Optimization sequence

$\Rightarrow$ There are two types of algorithms:

- If the algorithm can find the global minimum then we call it: global algorithm.
- If the algorithm can find only local minimum the we call it: local algorithm.

$\Rightarrow$ If in the sequence $\{x_n\}$ we find a point x' such that:

$$F(x_{opt}) < F(x') < F(x_{opt}) + \epsilon$$

then the above algorithm will converge only to x'.

$\Rightarrow$ Of course we can change the ϵ such that we escape the x'.

Q & A

Backup