Bandit Convex Optimization e o algoritmo FKM

André Mazal Krauss - INF2982 Algoritmos e Incerteza

December 2, 2020

Outline

Bandit Convex Optimization

Resolução do Problema

- Tentativas
- Flaxman, Kalai, McMahan
  - OCO via Gradient Descent
  - Obtenção do Gradiente via One Point Estimation

BCO: Deseja-se escolher/jogar, para $t = 1..T$ rodadas, um ponto $x_t \in K$, $K \subset \mathbb{R}^d$, de maneira que minimize o regret dado por:

$$Regret_T = \sum_{t=1}^{T} f_t(x_t) - \min_{x \in K} \sum_{t=1}^{T} f_t(x)$$

Para escolher x_t na rodada t, temos acesso a:

- $x_1, x_2, ..., x_{t-1}$
- $f_1(x_1), f_2(x_2), ..., f_{t-1}(x_{t-1})$

Online Convex Optimization?

BCO difere do OCO visto em sala em que temos acesso somente a $f_t(x_t)$, e não a outros pontos da função f_t

Bandit Experts?

BCO pode ser entendido como uma extensão do Bandit Experts em que consideramos um espaço $K \subset \mathbb{R}^d$ em vez de um conjunto finito de escolhas.

FTRL?

Algorithm 1: Follow the Regularized Leader

x_1 é escolhido;
$R(x)$ é α - fortemente convexo;
for $t = 1..T$ **do**
  Jogue x_t, tome perda $f_t(x_t)$
  $x_{t+1} = \text{argmin}_{x \in K} \sum_{k=1}^{t} f_k(x) + \frac{1}{\eta} R(x)$
end

BCO via Experts?

No trabalho 2, construímos um algoritmo para resolução de OCO usando uma 'grid' de experts. Seria possível resolver BCO combinando essa ideia com o algoritmo UCB usado no Bandit Experts?

Zinkevich - 2003

Algorithm 2: OCO via Gradient Descent

Seja P_K um 'oráculo projetivo', ie:
$P_K(x) = \text{argmin}_{z \in K} \|x - z\|$
Seja η uma taxa de aprendizado
x_1 é escolhido
for $t = 1..T$ **do**
    Jogue x_t, tome perda $f_t(x_t)$
    $x_{t+1} = P_K(x_t - \eta \nabla f_t(x_t))$
end

Garantias

$$regret \leq DG\sqrt{T}$$

Algoritmo FKM

One point gradient estimator

2D

Seja:

$$\tilde{f}'(x) = \begin{cases} +\frac{1}{\delta} f(x+\delta) & \text{com prob.} \frac{1}{2} \\ -\frac{1}{\delta} f(x-\delta) & \text{com prob.} \frac{1}{2} \end{cases}$$

Temos portanto:

$$\lim_{\delta \to 0} \mathbb{E}[\tilde{f}'(x)] = \lim_{\delta \to 0} \frac{f(x+\delta) - f(x-\delta)}{2\delta} = f'(x)$$

Algoritmo FKM

One point gradient estimator

D-dimensional

$$\tilde{\nabla} f(x) \approx \mathbb{E}[(f(x+\delta u) - f(x))u]\frac{d}{\delta} = \mathbb{E}[f(x+\delta u)u]\frac{d}{\delta}$$

onde u é um vetor unitário aleatório, amostrado uniformemente

Algoritmo FKM

Gradient Descent via Gradient Estimator

Algorithm 3: Bandit Gradient Descent(FKM)

Seja $K_\alpha \subset K$ com α pequeno
Seja P_{K_α} um 'oráculo projetivo', ie:
$P_{K_\alpha}(x) = \text{argmin}_{z \in K_\alpha} \|x - z\|$
Seja η uma taxa de aprendizado
$x_1 = 0$
for $t = 1..T$ **do**
    amostre um vetor unitário aleatório u
    $y_t = x_t + \delta u$
    Jogue y_t, tome perda $f_t(y_t)$
    $x_{t+1} = P_{K_\alpha}(x_t - \eta f_t(y_t) u)$
end

Algoritmo FKM

Garantias

Garantia FKM

Supondo alguns valores para T, δ, η e α, podemos obter:

$$\mathbb{E}[\sum_{t=1}^{T} f_t(x_t)] - \min_{x \in K} \sum_{t=1}^{T} f_t(x) \leq 3CT^{\frac{5}{6}} \sqrt[3]{d\frac{R}{r}}$$

Se adicionalmente considerarmos que cada f_t é L-Lipschitz:

$$\mathbb{E}[\sum_{t=1}^{T} f_t(x_t)] - \min_{x \in K} \sum_{t=1}^{T} f_t(x) \leq 2T^{\frac{3}{4}} \sqrt{3RdC(L + \frac{C}{r})}$$

Referências I

Abraham D. Flaxman, Adam Tauman Kalai, H. Brendan McMahan
Online convex optimization in the bandit setting: gradient descent without a gradient.
2008.

Roi Livni
COS-511 at Princeton - Lecture 23: Bandit Convex Optimization.
2017.

Martin Zinkevich
Online Convex Programming and Generalized Infinitesimal Gradient Ascent
2003.