程序代写代做 CS编程辅导

CS861: Theoretical Foundations of Machine Learning — Lecture 27 - 06/11/2023

University of Wisconsin–Madison, Fall 2023

Lectures 2 … adient Descent, Contextual Bandits

Lecturer: Kirthe… — *Scribed by: Haoyue Bai & Deep Patel*

Disclaimer: *These n… jected to the usual scrutiny reserved for formal publications. They may be distributed outside this class only with the permission of the instructor.*

WeChat: cstutorcs

We have been looking at the framework of `Follow-The-Regularized Leader` (`FTRL`) that helps the learner choose actions w_t over rounds such that we have smaller cumulative regret w.r.t. the best action in hindsight. Specifically, we saw how choosing an appropriate regularizer can help obtain small regret when learner incurs linear losses. We then looked at a more general framework – `FTRL` with convex losses and strongly convex regularizers – which, in turn, led us to the following observation:

Assignment Project Exam Help

> "If we know/anticipate that $\{\nabla f_t\}_{t\geq 1}$ are small in some dual-norm $||\cdot||_*$, then it would be a good idea to run `FTRL` with a regularizer Λ which is strongly convex w.r.t. the corresponding norm[1] $(||\cdot||_*)_* = ||\cdot||$."

Email: tutorcs@163.com

We stated this formally as Theorem 1.1 (stated below) which we will prove in today's lecture. We will wrap up our discussion on the `FTRL` framework by applying this result in the context of at `Online Gradient Descent`. Finally, we will conclude this lecture and the course with a brief discussion of `Contextual Bandits` and a commonly-used algorithm for it – `EXP4 Algorithm`.

QQ: 749389476

1 FTRL and Online Gradient Descent

Theorem 1.1. *If f_t is convex, and $\Lambda(w) = \frac{1}{\eta}\lambda(w)$ where λ is 1-strongly convex in $||\cdot||$, then*

https://tutorcs.com

$$R_T(FTRL, \underline{f}) \leq \frac{1}{\eta}\left(\max_{w\in\Omega}\lambda(w) - \min_{w\in\Omega}\lambda(w)\right) + \eta\sum_{t=1}^{T}||g_t||_*^2$$

where $g_t \in \partial f_t(w_t)$ and $||\cdot||_*$ is the dual-norm of $||\cdot||$.

Proof In previous lecture, we proved the following for any $u \in \Omega$:

$$\sum_{t=1}^{T} f_t(w_t) - \sum_{t=1}^{T} f_t(u) \leq \sum_{t=1}^{T} f_t(w_t) - \sum_{t=1}^{T} f_t(w_{t+1}) + \left(\max_{w\in\Omega}\lambda(w) - \min_{w\in\Omega}\lambda(w)\right)$$

Using this, for a given policy π, we can say that:

$$\begin{aligned} R_T(\pi, \underline{f}) &= \sum_{t=1}^{T} f_t(w_t) - \sum_{t=1}^{T} f_t(w_*) \\ &\leq \frac{1}{\eta}\left(\max_{w\in\Omega}\lambda(w) - \min_{w\in\Omega}\lambda(w)\right) + \sum_{t=1}^{T}\left(f_t(w_t) - f_t(w_{t+1})\right) \end{aligned}$$

[1] dual of the dual-norm, which is the norm itself

程序代写代做 CS编程辅导

Therefore, it is sufficient to prove that the following holds on all rounds t:

$$f_t(w_t) - f_t(w_{t+1}) \le \eta ||g_t||_*^2 \tag{1}$$

By convexity and as g_t [illegible]

$$f_t(w_{t+1}) \ge f_t(w_t) + g_t^T(w_{t+1} - w_t)$$

By Hölder's inequality,

$$f_t(w_t) - f_t(w_{t+1}) \le g_t^T(w_t - w_{t+1})$$
$$\Rightarrow f_t(w_t) - f_t(w_{t+1}) \le ||g_t||_* ||w_t - w_{t+1}|| \tag{2}$$

Let's now denote $F_t \triangleq \sum_{s=1}^{t} f_t(w) + \frac{1}{\eta}\lambda(w)$. Since λ is 1-strongly convex in $||\cdot||$-norm by assumption, we have that F_t is $\frac{1}{\eta}$-strongly convex in $||\cdot||$-norm. Note that, by definition, w_{t+1} minimizes F_t and w_t minimizes F_{t-1}. Thus, as F_{t-1} and F_t are $\frac{1}{\eta}$-strongly convex, we can say that

WeChat: cstutorcs

$$F_{t-1}(w_{t+1}) \ge F_{t-1}(w_t) + \frac{1}{2\eta}||w_{t+1} - w_t||^2$$

Assignment Project Exam Help

$$F_t(w_t) \ge F_t(w_{t+1}) + \frac{1}{2\eta}||w_t - w_{t+1}||^2$$

Summing both the sides above will give us

Email: tutorcs@163.com

$$f_t(w_t) - f_t(w_{t+1}) \ge \frac{1}{\eta}||w_t - w_{t+1}||^2 \tag{3}$$

Thus, Equations 2 and 3 imply that

QQ: 749389476

$$||w_t - w_{t+1}|| \le \eta ||g_t||_* \tag{4}$$

Now, combining Equation 2 and Equation 4 gives us the desired inequality as stated above in Equation 1:

https://tutorcs.com

$$f_t(w_t) - f_t(w_{t+1}) \le \eta ||g_t||_*^2$$

□

Example 2 (Online Gradient Descent). Let f_t be differentiable[2] and Ω be a compact, convex set. Choose $\lambda(w) = \frac{1}{2}||w||_2^2$. Let us say that we using the following `FTRL` framework to obtain w_t at the end of each round-t:

$$w_t \in \arg\min_{w \in \Omega} \sum_{s=1}^{t-1} f_s(w) + \frac{1}{2\eta}||w||_2^2$$

Although the actions w_t's obtained as above give us good regret rates, the problem with obtaining the w_t's this way is that, in general, the complexity of solving the aforementioned optimization problem grows with t – at the end of each round-t we have to compute a new gradient $\nabla f_t(w)$ which results in the computational cost growing linearly in t. We would like to keep the computational cost per round-t to be small, ideally not depending on t. So, we will start by rewriting

[2] We don't actually need this assumption. We are using it for simplicity in this class.

程序代写代做 CS编程辅导

the regret as follows:

$R_T(\pi, \{f_t\}$ [illegible] $\sum_{t=1}^{T} f_t(w)$

[illegible] $) - f_t(w)]\Big)$

[illegible] $w_t)(w_t - w)\Big)$

$$(\because f_t \text{ is convex } \iff f_t(w) \geq f_t(w_t) + (w - w_t)^T \nabla f_t(w_t)\ \forall w \in \Omega)$$

$$= \sum_{t=1}^{T} w_t^T \nabla f_t(w_t) - \min_{w \in \Omega} \sum_{t=1}^{T} w^T \nabla f_t(w_t)$$

WeChat: cstutorcs

$$= R_T\left(\pi,\ \underbrace{\{\nabla f_t(w_t)\}_{t=1}^{T}}_{\text{abuse of notation}^3}\right)$$

Assignment Project Exam Help

We will now apply FTRL on the linear losses $\tilde{f}_t(w) \triangleq w^T \nabla f_t(w_t)$ with $\lambda(w) = \frac{1}{2}||w||_2^2$ as shown below:

$$w_t = \arg\min_{w \in \Omega} \left(w^T \left(\sum_{s=1}^{t-1} \nabla f_s(w_s) \right) + \frac{1}{2\eta} ||w||_2^2 \right)$$

Email: tutorcs@163.com

$$= \arg\min_{w \in \Omega} ||w + \eta \sum_{s=1}^{t-1} \nabla f_s(w_s)||_2 \text{ (by completing the squares)}$$

QQ: 749389476

Hence, w_t will be the ℓ_2-projection of $-\eta \sum_{s=1}^{t-1} \nabla f_s(w_s)$ to Ω, which can be implemented in $\mathcal{O}(1)$-time[4] at each round-t as follows:

$$\begin{aligned} u_t &\longleftarrow u_{t-1} - \eta \nabla f_{t-1}(w_{t-1}) \\ w_t &\longleftarrow \arg\min_{w \in \Omega} ||w - u_t||_2 \end{aligned} \tag{5}$$

https://tutorcs.com

Now, we can show that

$$\begin{aligned} R_T(\pi, \{f_t\}_{t=1}^T) &\leq R_T(\pi, \{\nabla f_t(w_t)\}_{t=1}^T) \\ &\leq \frac{B}{\eta} + \eta T G^2 \text{ (By Theorem 1.1)} \\ &\in \mathcal{O}(G\sqrt{BT}) \left(\text{if } \eta = \sqrt{\frac{B}{TG^2}} \right) \end{aligned} \tag{6}$$

where $B = \max_{w \in \Omega} \lambda(w) - \min_{w \in \Omega} \lambda(w)$ and $||\nabla f_t(w_t)||_2 \leq G\ \forall t$.

Remark 1.1. *Some connections that we can make a note of:*

- *Suppose $f = f_t$ is a fixed function. This is similar to the standard Projected Gradient Descent (PGD) step:*

$$\begin{aligned} u_t &\longleftarrow w_{t-1} - \eta \nabla f(w_{t-1}) \\ w_t &\longleftarrow \arg\min_{w \in \Omega} ||w - u_t||_2 \end{aligned}$$

[3] We mean $w_t^T \nabla f_t(w_t)$ here

[4] We are not considering how this scales with the dimensionality, d, of $\Omega \subseteq \mathbb{R}^d$ at the moment

程序代写代做 CS编程辅导

Using update rule in Equation 5 on a fixed function leads to the following bound for convex optimization via Equation 6:

$$\frac{1}{T}\left(\sum_{t=1}^{T} f(w_t) - f(w_*)\right) \quad (\because \textit{min} \leq \ \textit{avg.})$$

$$\mathcal{O}\left(G\sqrt{\frac{B}{T}}\right)$$

Note that this ne... n optimal bound. We are simply showing an application of Theorem 1.1 to a ... oblem.

- *In machine learning, update rule defined in Equation 5 is similar to the Stochastic Gradient Descent (SGD) update where f_t is the loss for instance (x_t, y_t)*

WeChat: cstutorcs

2 Contextual Bandits

We will resume our discussion on `contextual bandits` now. Recall that, in the case of K-armed bandits, we had K-arms that can be pulled and we were competing against the single best arm/action in hindsight. However, in certain situations, the best arm/action depends on contextual information, which may be available to the learner. For e.g., K-armed bandits: advertising; contextual bandits: targeted advertising.

Assignment Project Exam Help

Definition 1 (The contextual bandit problem). *We will define the contextual bandit problem as follows:*

(i.) The environment picks a context $x_t \in \mathcal{X}$ and the learner observes x_t.

(ii.) Learner chooses action $A_t \in [K]$. Simultaneously, the environment picks a loss vector $\ell_t \in [0,1]^k$.

(iii.) Learner incurs the loss $\ell_t(A_t)$.

(iv.) Learner observes (only) $\ell_t(A_t)$.

Email: tutorcs@163.com

QQ: 749389476

https://tutorcs.com

Question: How do we define regret here?

- One option is to compete against the best action for the given context:

$$R_T(\pi, \underline{\ell}, \underline{x}) = \mathbb{E}\left[\sum_{t=1}^{T} \ell_t(A_t)\right] - \min_{e:\mathcal{X}\to[k]} \sum_{t=1}^{T} \ell_t(e(x_t))$$

  where $\mathbb{E}$ is w.r.t. the randomness of policy. Here, we are competing against the single best mapping from contexts to actions.

  - This is like running a separate bandit algorithm for different contexts.
  - And this is challenging if $|\mathcal{X}|$ is large, but also unnecessary if there are relationships between contexts. e.g., frying pan, non-stick skillet.

- Instead, we will consider a set of N experts, who map contexts to actions and we will now be competing against the single best expert in hindsight. Here, the experts could be, say, machine learning models trained on a variety of large datasets.

  - If the experts are $\{e_1, ..., e_N\}$, where $e_j : \mathcal{X} \to [K]\ \forall j \in [N]$, then

$$R_T(\pi, \underline{\ell}, \underline{x}) = \mathbb{E}\left[\sum_{t=1}^{T} \ell_t(A_t)\right] - \min_{j\in[k]} \sum_{t=1}^{T} \ell_t(e_j(x_t)).$$

程序代写代做 CS编程辅导

Question: Can we apply EXP3 algorithm here by treating the experts as arms?

- Yes. But the rge. $R_T \in O(\sqrt{TN\log(N)})$ which is fine as long as the number of ex ever, we are usually interested in cases where we have many more ex ons, $N \gg K$. We wish to avoid poly(N) dependence, but a poly $\log(N)$- ne.

3 The EXP4 al

WeChat: cstutorcs

Just like we built on `Hedge` to arrive at the `EXP3` algorithm, we are going to build on the `EXP3` algorithm to arrive at the `EXP4` algorithm here. We will treat the experts as arms here and run `EXP3` on them, but what we will do differently here is the following: We will use the fact that when we observe feedback, we will discount all the experts who would have chosen the action. Based on this, we can express the pseudocode of `EXP4` as shown below in Algorithm 1:

Assignment Project Exam Help

Email: tutorcs@163.com

QQ: 749389476

https://tutorcs.com

Algorithm 1: The EXP-4 algorithm

Input: Time horizon T, learning rate η
Let $\tilde{L}_0 \leftarrow \mathbf{0}_N$
for $t = 1, ..., T$ **do**
  Observe x_t
  Compute $\tilde{p}_t(i) = \frac{e^{-\eta \tilde{L}_{t-1}(i)}}{\sum_{j=1}^{N} e^{-\eta \tilde{L}_{t-1}(j)}}, \forall i \in [N]$
  Let $p_t(a) = \sum_{j=1}^{N} \tilde{p}_t(j)\mathbb{I}(e_j(x_t) = a)$
  Sample $A_t \sim p_t$
  Observe $\ell_t(A_t)$
  $\hat{\ell}_t(a) \leftarrow \frac{\ell_t(a)}{p_t(a)}\mathbb{I}(A_t = a)\ \forall a \in [k]$
  $\tilde{\ell}_t(j) \leftarrow \hat{\ell}_t(e_j(x_t))\ \forall j \in [N]$
  $\tilde{L}_t(j) \leftarrow \tilde{L}_{t-1}(j) + \tilde{\ell}_t(j)\ \forall j \in [N]$
end

Remark 3.1. *We can note the following about the EXP4 algorithm:*

- *Lines (ii.), (iii.), and (iv.) can be implemented as* `Expert`$\sim \tilde{p}_t$ *and* $A_t =$ `Expert`(x_t).
- *We can write the loss update as*

$$\tilde{L}_t(j) \longleftarrow \tilde{L}_{t-1}(j) + \mathbb{I}\{e_j(x_t) = A_t\} \cdot \frac{\ell_t(A_t)}{p_t(A_t)} \tag{7}$$

  ▲ *Note that we are not discounting only one expert here. That is, we are NOT utilizing the following update rule:*

$$\tilde{L}_t(j) \longleftarrow \tilde{L}_{t-1}(j) + \mathbb{I}\{E_t = j\} \cdot \frac{\ell_t(A_t)}{\tilde{p}_t(E_t)} \tag{8}$$

  See Remark 3.2 below for more on this.

Theorem 3.1 (Regret bound for `EXP4`)**.** *Assume that the loss vectors on each round satisfy $\ell_t \in [0,1]^K$ and let $x_t \in \mathcal{X}$ be drawn arbitrarily. Then, $\forall\ \underline{\ell}\ (= (\ell_1, \ldots, \ell_T))$, $\forall\ \underline{x}\ (= (x_1, \ldots, x_T))$, EXP4 satisfies:*

$$R_T(\pi, \underline{\ell}, \underline{x}) \leq \frac{\log N}{\eta} + \eta K T$$

程序代写代做 CS编程辅导

where N is the number of experts. With $\eta = \sqrt{\frac{\log N}{KT}}$, we see that the regret bound becomes

$$\ell, x) \leq 2\sqrt{KT \log N}$$

Proof First, we will $\mathbb{E}[\tilde{\ell}_t^2(j)|\tilde{p}_t]$.

$$_j(x_t))) \times 0 + p_t(e_j(x_t)) \cdot \frac{\ell_t(e_j(x_t))}{p_t(e_j(x_t))} \tag{9}$$

Similarly,

$$\begin{aligned}\mathbb{E}[\tilde{\ell}_t^2(j)|\tilde{p}_t] &= (1 - p_t(e_j(x_t))) \times 0 + p_t(e_j(x_t)) \cdot \frac{\ell_t^2(e_j(x_t))}{p_t^2(e_j(x_t))} \\ &= \frac{\ell_t^2(e_j(x_t))}{p_t(e_j(x_t))}\end{aligned} \tag{10}$$

WeChat: cstutorcs

We will now apply result (i.) from the `Hedge` Theorem (Theorem 1, Lecture 23). As we have N experts, the loss $\tilde{\ell}_t$ is non-negative and $\tilde{p}_t^T \tilde{\ell}_t \leq 1\ \forall t$, we have, for any expert i:

Assignment Project Exam Help

$$\sum_{t=1}^{T} \tilde{p}_t^T \tilde{\ell}_t - \sum_{t=1}^{T} \tilde{\ell}_t(i) \leq \frac{\log N}{\eta} + \eta \sum_{t=1}^{T} \tilde{p}_t^T \tilde{\ell}_t^2 \tag{11}$$

We will apply Equation 11 by setting $i = i^*$ (i.e., best expert in hindsight),

Email: tutorcs@163.com

$$i^* = \arg\min_{i \in [N]} \sum_{t=1}^{T} \ell_t(e_j(x_t))$$

QQ: 749389476

From Equation 9, we get

$$\mathbb{E}[\tilde{\ell}_t(i^*)] = \ell_t(e_{i*}(x_t))$$

From Equation 9 again, we get

https://tutorcs.com

$$\begin{aligned}\mathbb{E}[\tilde{p}_t^T \tilde{\ell}_t|\tilde{p}_t] &= \tilde{p}_t^T \mathbb{E}[\tilde{\ell}_t|\tilde{p}_t] \\ &= \sum_{j=1}^{N} \tilde{p}_t(j)\ell_t(e_j(x_t)) \ \ \text{(By Equation 9)} \\ &= \sum_{j=1}^{N} \tilde{p}_t(j) \left(\sum_{a=1}^{K} \ell_t(a)\mathbb{I}\{e_j(x_t) = a\} \right) \\ &= \sum_{a=1}^{K} \ell_t(a) \underbrace{\sum_{j=1}^{N} \tilde{p}_t(j)\mathbb{I}\{e_j(x_t) = a\}}_{p_t(a)} \\ &= \sum_{a=1}^{K} p_t(a)\ell_t(a) = p_t^T \ell_t \\ &= \mathbb{E}\left[\ell_t(A_t)|p_t\right]\end{aligned}$$

程序代写代做 CS编程辅导

Next, using Equation 10, we can say

$$\frac{\ell_t^2(e_j(x_t))}{p_t(e_j(x_t))} \quad \text{(By Equation 10)}$$

$$\sum_{a=1}^{K} \frac{\ell_t^2(a)}{p_t(a)} \mathbb{I}\{e_j(x_t) = a\}$$

$$\underbrace{\sum_{j=1}^{N} \tilde{p}_t(j) \mathbb{I}\{e_j(x_t) = a\}}_{p_t(a)}$$

$$= \sum_{a=1}^{K} \ell_t^2(a) \leq K \quad (\because \ell_t(a) \in [0,1] \, \forall a \in [K], \, \forall t)$$

WeChat: cstutorcs

Remark 3.2. *We discount all experts that predict a at round-t instead of just one expert. If we were to discount only one expert, we would get a dependence on N as shown below which we clearly do not want in case where we have large N.*

Assignment Project Exam Help

$$\sum_{j=1}^{N} \tilde{p}_t(j) \frac{\ell_t^2(e_j(x_t))}{\tilde{p}_t(j)} \leq N$$

Email: tutorcs@163.com

Intuitively, since $p_t(A_t) \geq \tilde{p}_t(E_t)$, we see that the update rule in Equation 7 is better as it reduces variance in observed loss values compared to the case where we use the update rule in Equation 8.

QQ: 749389476

Thus, we can finally see that

https://tutorcs.com

$$\begin{aligned}
\mathbb{E}[\text{LHS of Equation 11}] &= \mathbb{E}\left[\sum_{t=1}^{T} \ell_t(A_t)\right] - \sum_{t=1}^{T} \ell_t(e_{i^*}(x_t)) \\
&= R_T(\pi, \underline{\ell}, \underline{x}) \\
\mathbb{E}[\text{RHS of Equation 11}] &\leq \frac{\log N}{\eta} + \eta K T \\
&\in \mathcal{O}(\sqrt{KT \log N}) \quad \left(\text{if } \eta = \sqrt{\frac{\log N}{KT}}\right)
\end{aligned}$$

$\square$