Online Prediction

Logistical announcement:

HW 1 is on the website!
Due Friday after this one.

Until now (Stochastic Multi-armed Bandit)

K arms

Time t
Bandit (player) pick arm $i_t \in [K]$ to pull
Reward μ_{i_t} + noise r_t

Patient i $\longrightarrow$ Pick one of K treatments $\longrightarrow$ Moderna
$\longrightarrow$ AstraZeneca

(vaccine) → Pfizer
→ J&J
→ Do nothing (control)

Contextual bandit

$\mathbb{R}^d$

$\ni$

$x_t \longrightarrow$ pick arm $i(x_t, \dots, x_1, r_{t-1}, \dots, r_1)$ → Arm 1 (f_1^*)

Context

→ Arm 2

⋮

→ Arm K (f_K^*)

Reward of Arm i $= f_i^*(x_t) +$ noise

Class of functions

$$f_i^* \in \mathcal{F} \subseteq \{\mathbb{R}^d \to \mathbb{R}\}$$

$\forall f$ UNKNOWN

Two options for how to model $\{x_t\}$

Two options for how

1. Arbitrary fixed sequence $x_1 \dots x_T$ (could be adaptive also)

2. Stochastic context $x_1 \dots x_T \overset{iid}{\sim} \mathcal{D}_x$ on $\mathbb{R}^d$

Regret?

Oracle who knows $f_1^* \dots f_K^*$

$$i_t^* = \operatorname{argmax}_{i \in [K]} f_i^*(x_t)$$

$$\text{Regret} = \sum_{t=1}^{T} r_{\hat{i}_t} + \sum_{t=1}^{T} r_{i_t^*}$$

arm algorithm picks at round t

$$= \sum_{t=1}^{T} \left(r_{i_t^*} - r_{\hat{i}_t} \right)$$

Goal: minimize

$\mathbb{E}$ Regret

New Challenge in CB:

"Forecasting"

Generally speaking, CB algorithm looks like this.

NEW (learning) [1. Forecast $\hat{r}_1(x_t), \ldots \hat{r}_t(x_t)$

Like MAB?. [2. Given the forecasts, pick an arm $\hat{i}_t$

Challenge: how to learn online???

"Online prediction"

Online regression

[illegible]

Nature gives covariate x_t.

We predict $\hat{y}_t$ (based on $x_1 \ldots x_t$, $\hat{y}_1 \ldots \hat{y}_{t-1}$)

Nature shows true y_t (response)

$$\text{Regret} = \sum_{t=1}^{T} (\hat{y}_t - y_t)^2 - \text{ORACLE}$$

Performance

$$\left(\sum^{\text{more generally}} \ell(\hat{y}, y_t) \right)$$

$$f^* \in \mathcal{Y}$$

$$y_t = f^*(x_t) + \xi_t$$

ORACLE performance $\tilde{y}_t = f^*(x_t)$ Oracle

$$\text{Regret} = \sum_{t=1}^{T} (y_t - \hat{y}_t)^2 - \sum_{t=1}^{T} (y_t - \tilde{y}_t)^2$$

$$\sum_{t=1}^{T} \xi_t^2$$

Goal: Minimize

$\mathbb{E}$ Regret

$$\sum_{t=1}^{T} \mathbb{E}(y_t - \hat{y}_t)^2$$

$$= \mathbb{E} \sum_{t=1}^{T} \left(f^*(x_t) + \xi_t - \hat{y}_t\right)^2$$

$$= \mathbb{E} \sum_{t=1}^{T} \mathbb{E}\left[\left(f^*(x_t) + \xi_t - \hat{y}_t\right)^2 \,\middle|\, \begin{matrix} x_1, \ldots, x_t \\ y_1, \ldots, y_{t-1} \end{matrix}\right]$$

$$\left(f^*(x_t) - \hat{y}_t\right)^2 + \cancel{2\xi_t\left(f^*(x_t) - \hat{y}_t\right)}$$

$$+ \xi_t^2$$

$$= \mathbb{E} \sum_{t=1}^{T} \left(f^*(x_t) - \hat{y}_t\right)^2 + \mathbb{E} \sum_{t=1}^{T} \xi_t^2$$

non-negative ORACLE

$$\mathbb{E}\,\text{Regret} = \sum_{t=1}^{T} \mathbb{E}\left(f^*(x_t) - \hat{y}_t\right)^2$$

Let's focus on linear $f^*(x_t)$

$$= \langle w^*, x_t \rangle$$

$$\mathbb{E}\,\text{Regret} = \sum_{t=1}^{T} \mathbb{E}\underbrace{\left(\langle w^*, x_t \rangle - \hat{y}_t\right)^2}_{\langle w^* - \hat{w}_t,\, x_t \rangle^2}$$

$$\hat{y}_t = \langle \hat{w}_t, x_t \rangle$$

Chosen by algorithm based off of $x_1 \dots x_t$

$y_1 \dots y_{t-1}$

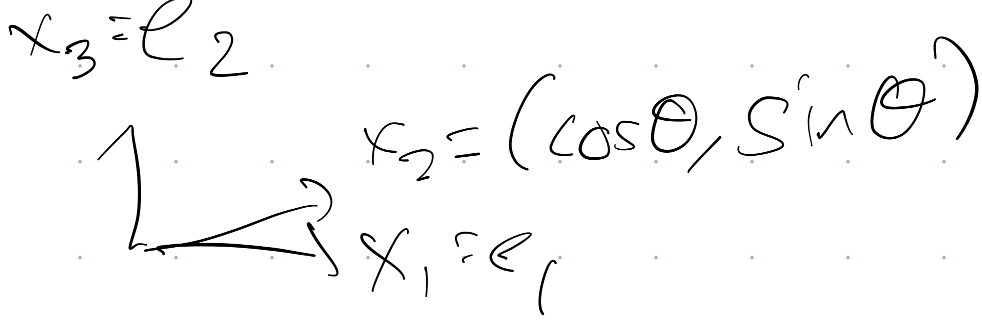
"Naive" strategy:

Follow the leader, FTL

$$\hat{w}_t = \arg\min_w \sum_{s=1}^{t-1} \left(y_t - \langle w, x_t \rangle\right)^2$$

Bad 2D example: $y_i \sim \mathcal{N}(0,1)$

θ small



Predict by FTL at round 3

$$\hat{w}_3 = \arg\min_w \left\| \begin{bmatrix} 1 & 0 \\ \cos\theta & \sin\theta \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} - \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \right\|_2^2$$

$$= \begin{bmatrix} 1 & 0 \\ \cos\theta & \sin\theta \end{bmatrix}^{-1} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

$$= \underbrace{\frac{1}{\sin\theta}\begin{bmatrix} \sin\theta & -\cos\theta \\ & 1 \end{bmatrix}}\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & \frac{-\cos\theta}{\sin\theta} \\ & \frac{1}{\sin\theta} \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

$$\hat{y}_3 = \frac{y_2}{\sin\theta} \rightarrow \text{either } +\infty \text{ or } -\infty$$

as $\theta \to 0$

What if we truncate?

e.g. assume $f^*(x_i) \in [-1, 1]$

Safe to truncate at $[\pm\sqrt{2\log T}]$

Truncation is not enough!

1D example, $y_i \overset{iid}{\sim} \mathcal{N}(0,1)$

$x_1 = 10^{-T}$, $x_2 = 10^{1-T}$, ..., $x_T = 1$

FTL does badly, even with

truncation