Online Learning with Kernel Losses

Aldo Pacchiano

UC Berkeley

Joint work with Niladri Chatterji and Peter Bartlett

Talk Overview

- Intro to Online Learning
- Linear Bandits
- **Kernel Bandits**

Online Learning

Online Learning

Learner

$$t = 1, \cdots, n$$

Adversary

Online Learning

Learner

$$t = 1, \cdots, n$$

Adversary

Learner chooses an action $a_t \in \mathcal{A}$

Online Learning

Learner

$$t = 1, \cdots, n$$

Adversary

Learner chooses an action $a_t \in \mathcal{A}$

Adversary reveals loss (or reward) $\ell_t \in \mathcal{W}$

Online Learning

Learner

$t = 1, \cdots, n$

Adversary

Learner chooses an action $a_t \in \mathcal{A}$

Adversary reveals loss (or reward) $\ell_t \in \mathcal{W}$

Can be i.i.d or adversarial

Online Learning

Learner

$$t = 1, \cdots, n$$

Adversary

Learner chooses an action $a_t \in \mathcal{A}$

Adversary reveals loss (or reward) $\ell_t \in \mathcal{W}$

Can be i.i.d or adversarial

$$\sum_{t=1}^{n} \ell_t(a_t)$$

Online Learning

Learner

$t = 1, \cdots, n$

Adversary

Learner chooses an action $a_t \in \mathcal{A}$

Adversary reveals loss (or reward) $\ell_t \in \mathcal{W}$

Can be i.i.d or adversarial

$$R(n) = \sum_{t=1}^{n} \ell_t(a_t) - \min_{a^* \in \mathcal{A}} \sum_{t=1}^{n} \ell_t(a_t)$$

Online Learning

Learner

$$t = 1, \cdots, n$$

Adversary

Learner chooses an action $a_t \in \mathcal{A}$

Adversary reveals loss (or reward) $\ell_t \in \mathcal{W}$

Can be i.i.d or adversarial

$$R(n) = \sum_{t=1}^{n} \ell_t(a_t) - \min_{a^* \in \mathcal{A}} \sum_{t=1}^{n} \ell_t(a_t)$$

The learner's objective is to minimize Regret

Full information vs Bandit feedback

Full information vs Bandit feedback

Full Information:

Learner gets to sees all of

$$\ell_t(\cdot)$$

Full information vs Bandit feedback

Full Information:

Learner gets to sees all of

$$\ell_t(\cdot)$$

Bandit Feedback:

Learner only sees the value

$$\ell_t(a_t)$$

Full information vs Bandit feedback

Full Information:

Learner gets to sees all of

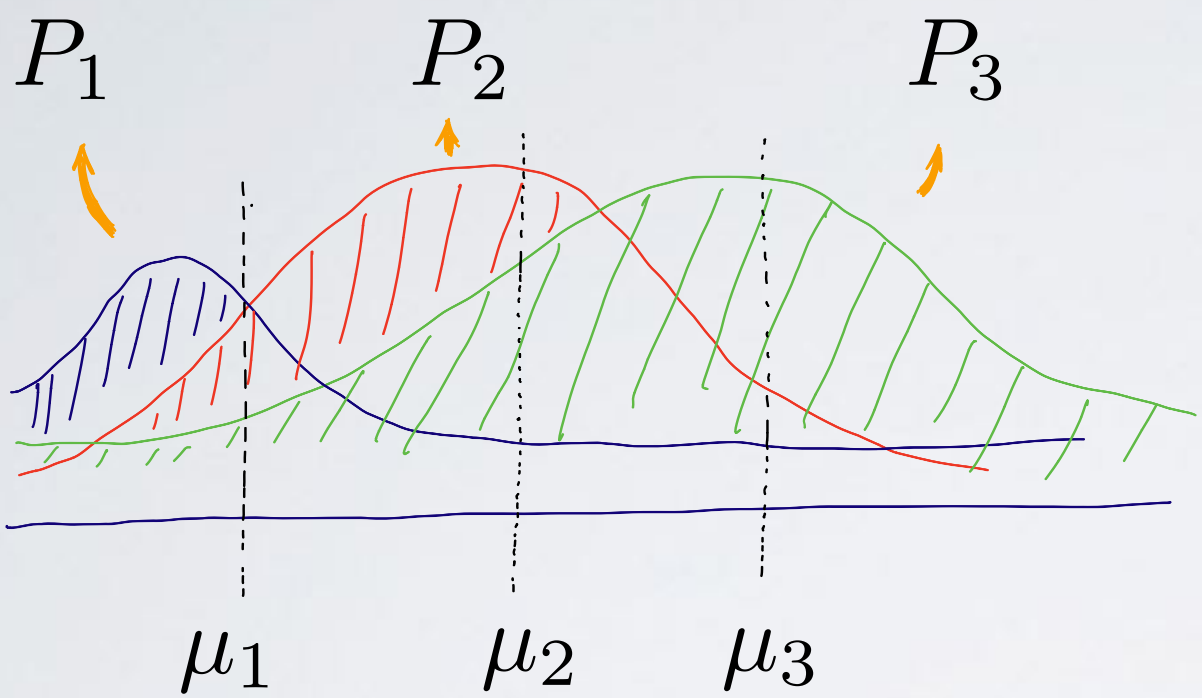
$$\ell_t(\cdot)$$

Bandit Feedback:

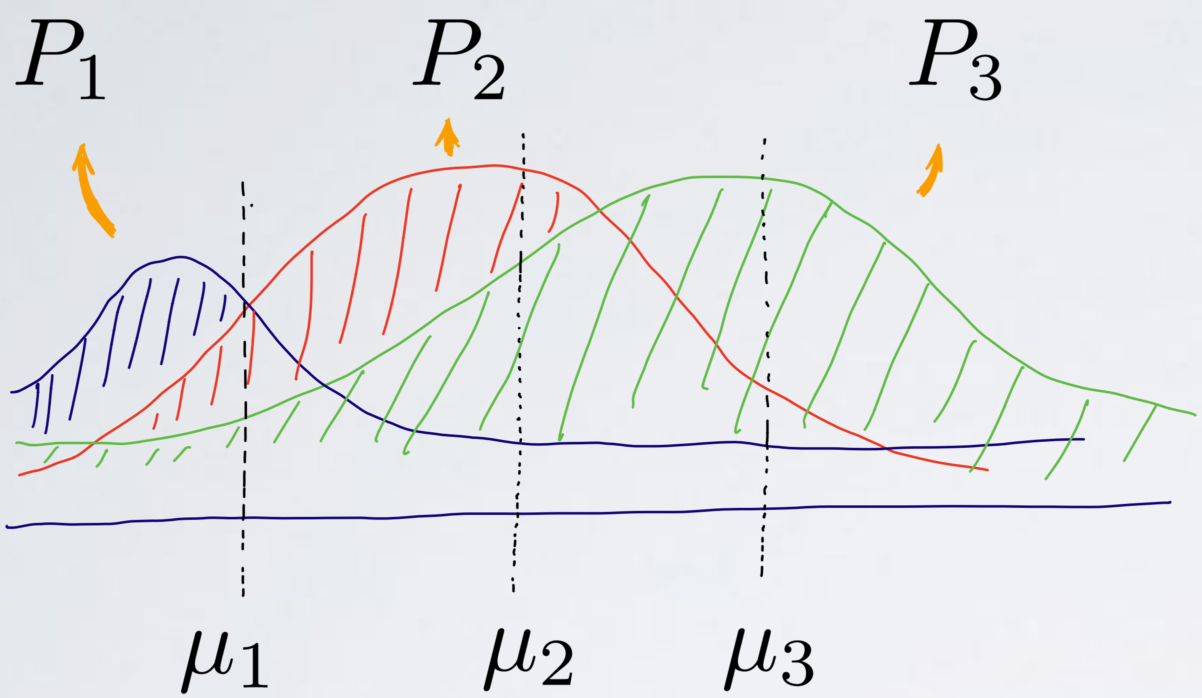
Learner only sees the value

$$\ell_t(a_t)$$

Multi Armed Bandits



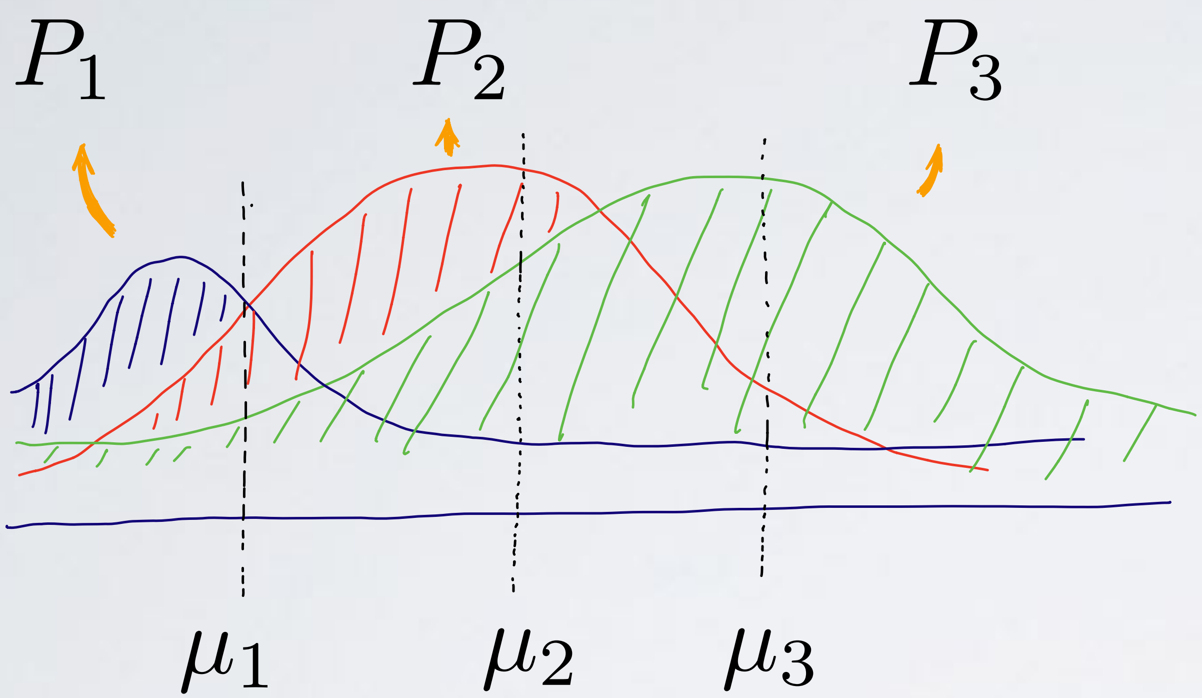
Multi Armed Bandits



Learner chooses

$$a_t \in \{1, \cdots, K\}$$

Multi Armed Bandits



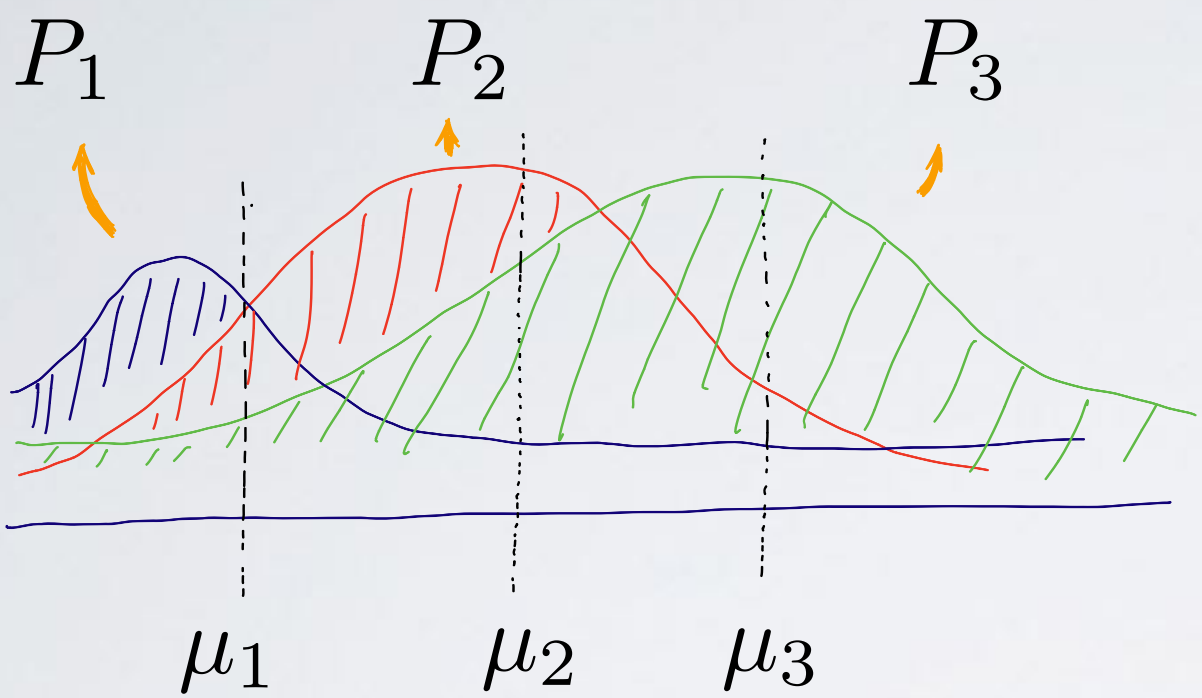
Learner chooses

$$a_t \in \{1, \cdots, K\}$$

Gets reward

$$X_{a_t} \sim P_{a_t}$$

Multi Armed Bandits



Learner chooses

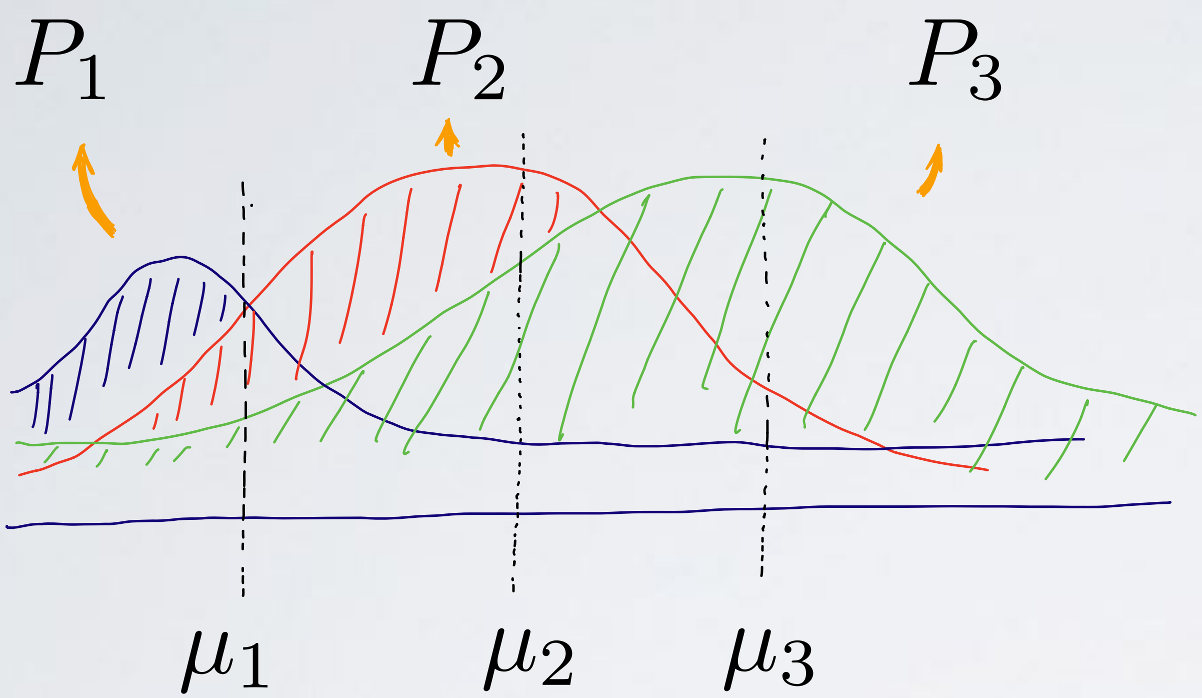
$$a_t \in \{1, \cdots, K\}$$

Gets reward

$$X_{a_t} \sim P_{a_t}$$

$$R(n) = \max_{a^* \in \{1, \cdots K\}} n\mu_{a^*} - \mathbb{E}\left[\sum_{t=1}^{n} X_{a_t}\right]$$

Multi Armed Bandits



Learner chooses

$$a_t \in \{1, \cdots, K\}$$

Gets reward

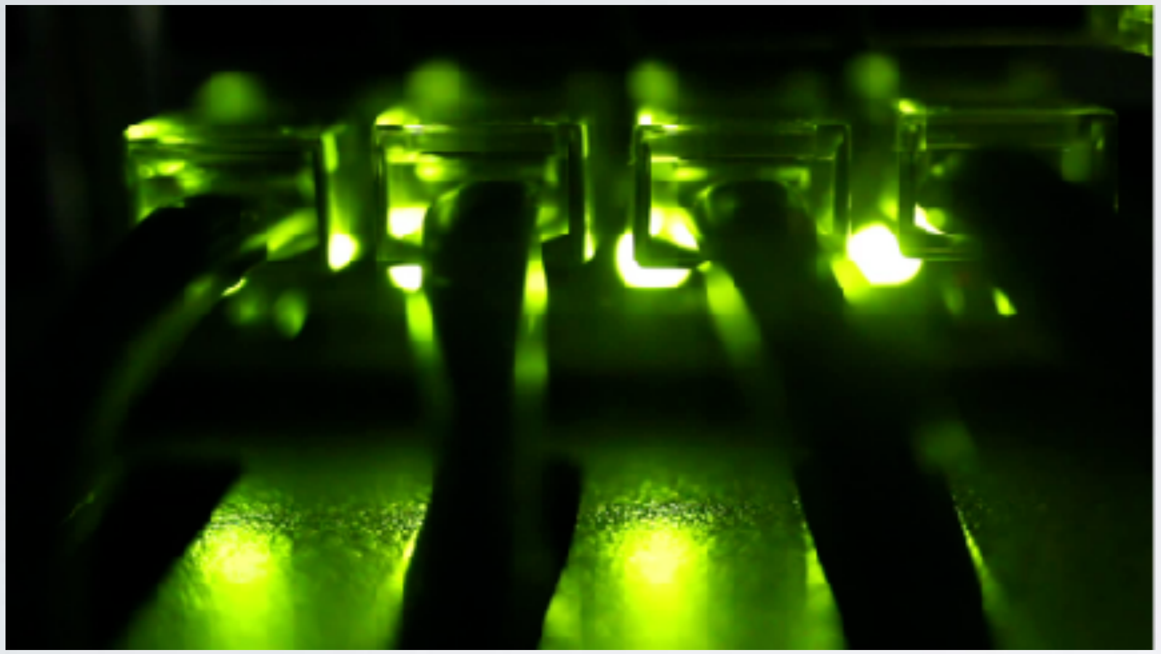
$$X_{a_t} \sim P_{a_t}$$

$$R(n) = \max_{a^* \in \{1, \cdots K\}} n\mu_{a^*} - \mathbb{E}\left[\sum_{t=1}^{n} X_{a_t}\right]$$

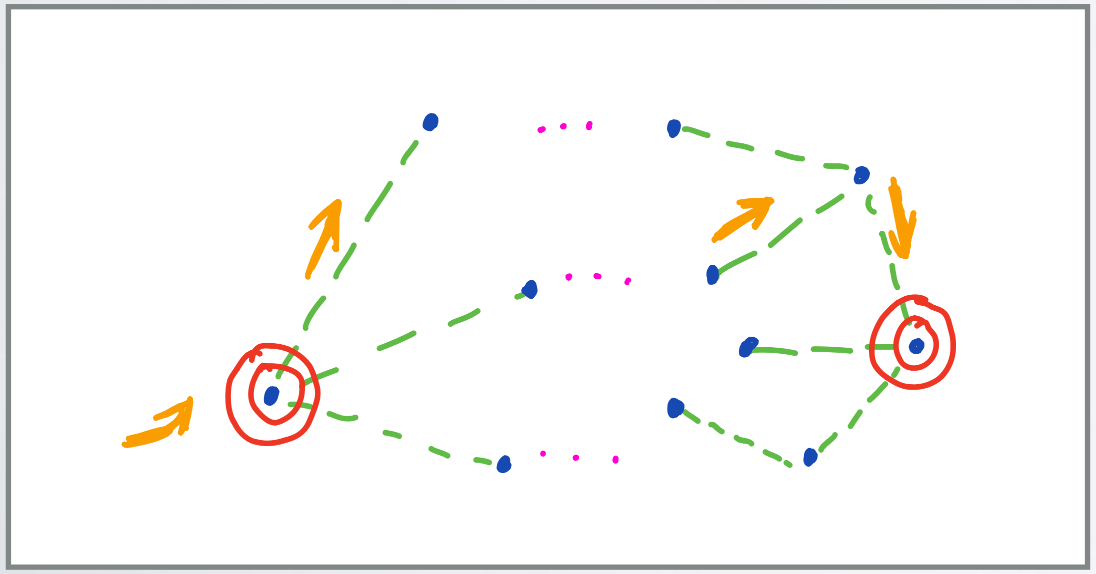
MAB regret $R(n) = \mathcal{O}(\sqrt{Kn\log(n)})$

[Auer et al. 2002]

Structured losses



Packet routing



Network (V, E)

Arms = Paths $a_t \in \mathcal{A} \subset \{0,1\}^E$

Loss = delay $w_t \in \mathcal{W} = [0,1]^E$

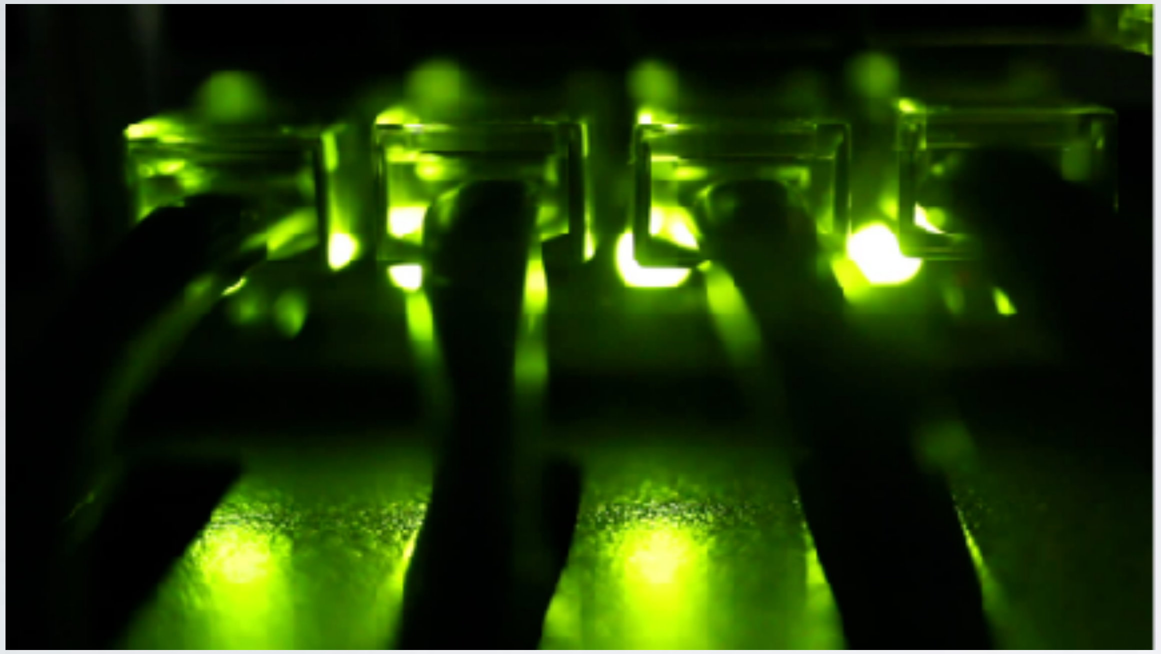
MAB regret

Exponential

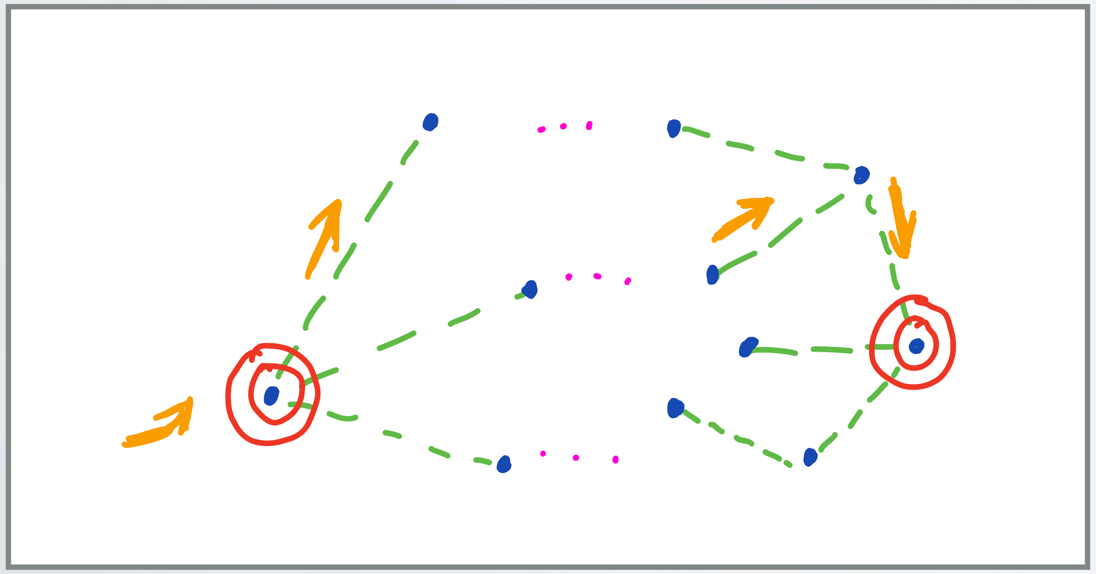
$$R(n) = \mathcal{O}(\sqrt{|\text{num paths}| \cdot n \log(n)})$$

Delay is linear $\langle a_t, w_t \rangle$

Structured losses



Packet routing



Network (V, E)

Arms = Paths $a_t \in \mathcal{A} \subset \{0,1\}^E$

Loss = delay $w_t \in \mathcal{W} = [0,1]^E$

MAB regret

Exponential

$$R(n) = \mathcal{O}(\sqrt{|\text{num paths}| \cdot n \log(n)})$$

Delay is linear $\langle a_t, w_t \rangle$

Linear Bandits

Linear Bandits

Learner chooses an action $a_t \in \mathcal{A} \subset \mathbb{R}^d$

Linear Bandits

Learner chooses an action $a_t \in \mathcal{A} \subset \mathbb{R}^d$

Adversary's loss $\ell_t(a) = \langle w_t, a \rangle$ for $w_t \in \mathcal{W} \subset \mathbb{R}^d$

Linear Bandits

Learner chooses an action $a_t \in \mathcal{A} \subset \mathbb{R}^d$

Adversary's loss $\ell_t(a) = \langle w_t, a \rangle$ for $w_t \in \mathcal{W} \subset \mathbb{R}^d$

Can be i.i.d or adversarial

Linear Bandits

Learner chooses an action $a_t \in \mathcal{A} \subset \mathbb{R}^d$

Adversary's loss $\ell_t(a) = \langle w_t, a \rangle$ for $w_t \in \mathcal{W} \subset \mathbb{R}^d$

Learner only experiences $\langle w_t, a_t \rangle$

Can be i.i.d or adversarial

Linear Bandits

Learner chooses an action $a_t \in \mathcal{A} \subset \mathbb{R}^d$

Adversary's loss $\ell_t(a) = \langle w_t, a \rangle$ for $w_t \in \mathcal{W} \subset \mathbb{R}^d$

Learner only experiences $\langle w_t, a_t \rangle$

Can be i.i.d or adversarial

Expected regret:

$$R(n) = \mathbb{E}\left[\sum_{t=1}^{n} \langle w_t, a_t \rangle - \inf_{a \in \mathcal{A}} \sum_{t=1}^{n} \langle w_t, a \rangle\right]$$

Linear Bandits

Learner chooses an action $a_t \in \mathcal{A} \subset \mathbb{R}^d$

Adversary's loss $\ell_t(a) = \langle w_t, a \rangle$ for $w_t \in \mathcal{W} \subset \mathbb{R}^d$

Learner only experiences $\langle w_t, a_t \rangle$

Can be i.i.d or adversarial

Expected regret:

$$R(n) = \mathbb{E}\left[\sum_{t=1}^{n} \langle w_t, a_t \rangle - \inf_{a \in \mathcal{A}} \sum_{t=1}^{n} \langle w_t, a \rangle\right]$$

MAB reduces to Linear Bandits

$$\mathcal{A} = \{e_1, \cdots, e_d\}, \qquad \mathcal{W} = [0, 1]^d$$

Exponential weights for adversarial linear bandits

Exponential weights for adversarial linear bandits

For $t = 1, \cdots, n$:

Sample mixture $a_t \sim p_t = \underbrace{(1-\gamma) q_t}_{\text{Exploitation}} + \underbrace{\gamma \nu}_{\text{Exploration}}$

Exponential weights for adversarial linear bandits

For $t = 1, \cdots, n$:

Sample mixture $a_t \sim p_t = \underbrace{(1-\gamma)q_t}_{\text{Exploitation}} + \underbrace{\gamma\nu}_{\text{Exploration}}$

Exponential weights for adversarial linear bandits

For $t = 1, \cdots, n$:

Sample mixture $a_t \sim p_t = \underbrace{(1-\gamma)q_t}_{\text{Exploitation}} + \underbrace{\gamma\nu}_{\text{Exploration}}$

Exponential weights for adversarial linear bandits

For $t = 1, \cdots, n$:

Sample mixture $a_t \sim p_t = \underbrace{(1-\gamma)q_t}_{\text{Exploitation}} + \underbrace{\gamma\nu}_{\text{Exploration}}$

See $\langle w_t, a_t \rangle$

Exponential weights for adversarial linear bandits

For $t = 1, \cdots, n$:

Sample mixture $a_t \sim p_t = \underbrace{(1-\gamma)q_t}_{\text{Exploitation}} + \underbrace{\gamma\nu}_{\text{Exploration}}$

See $\langle w_t, a_t \rangle$

Build loss estimator $\hat{w}_t$

Exponential weights for adversarial linear bandits

For $t = 1, \cdots, n$:

Sample mixture $a_t \sim p_t = \underbrace{(1-\gamma)q_t}_{\text{Exploitation}} + \underbrace{\gamma\nu}_{\text{Exploration}}$

See $\langle w_t, a_t \rangle$

Build loss estimator $\hat{w}_t$

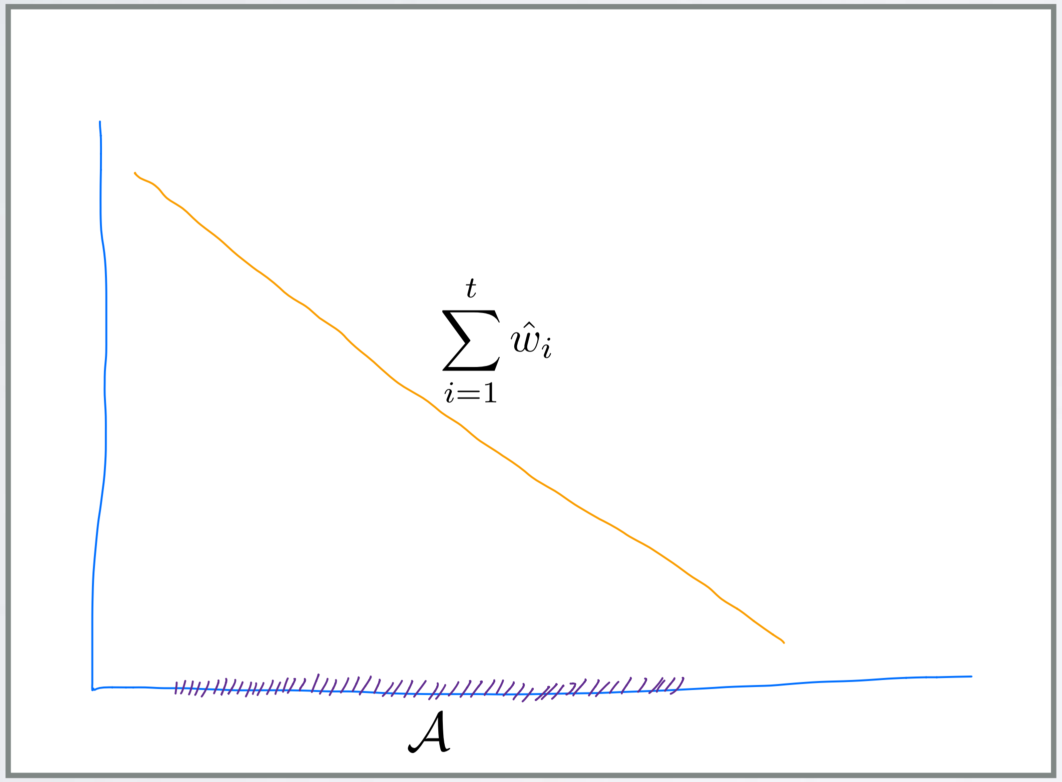
Update $\underbrace{q_t(a) \propto \exp(-\eta\langle \hat{w}_t, a\rangle) q_{t-1}(a)}_{\text{Exponential weights}}$

Exponential weights

$$q_t(a) \propto \exp(-\eta\langle\sum_{i=1}^{t}\hat{w}_i, a\rangle)$$

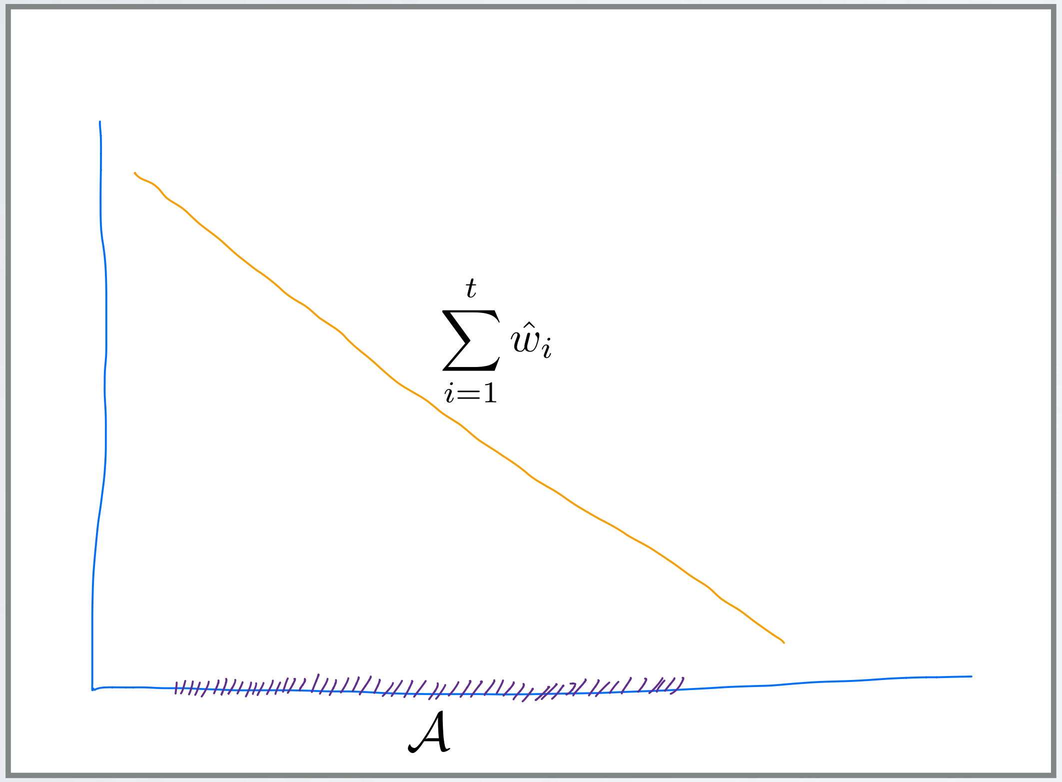
Exponential weights

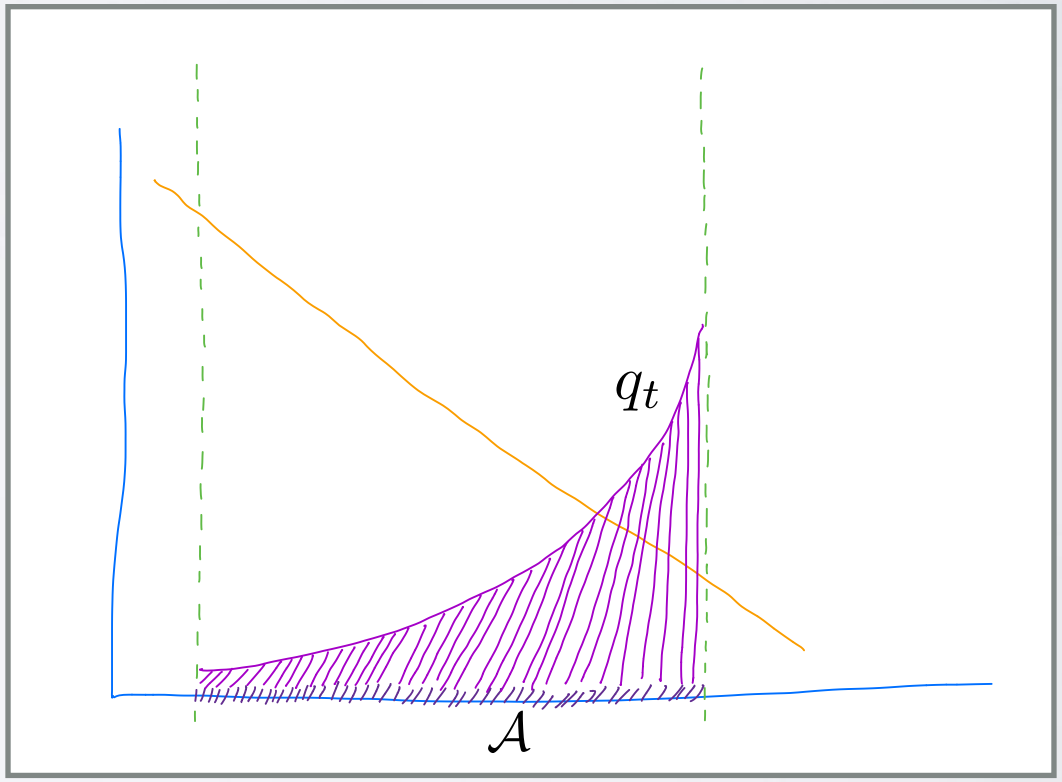
$$q_t(a) \propto \exp(-\eta\langle\sum_{i=1}^{t}\hat{w}_i, a\rangle)$$



Exponential weights

$$q_t(a) \propto \exp(-\eta \langle \sum_{i=1}^{t} \hat{w}_i, a \rangle)$$





Unbiased estimator of the loss

Let $\Sigma_t = \mathbb{E}_{a \sim p_t}\left[aa^\top\right]$ and set $\hat{w}_t = (\Sigma_t)^{-1} a_t \langle w_t, a_t \rangle$

Unbiased estimator of the loss

Let $\Sigma_t = \mathbb{E}_{a \sim p_t}\left[aa^\top\right]$ and set $\hat{w}_t = (\Sigma_t)^{-1} a_t \langle w_t, a_t \rangle$

Unbiased estimator of the loss

Let $\Sigma_t = \mathbb{E}_{a \sim p_t}\left[aa^\top\right]$ and set $\hat{w}_t = (\Sigma_t)^{-1} a_t \langle w_t, a_t \rangle$

$\hat{w}_t$ is an unbiased estimator of w_t :

Unbiased estimator of the loss

Let $\Sigma_t = \mathbb{E}_{a \sim p_t}\left[aa^\top\right]$ and set $\hat{w}_t = (\Sigma_t)^{-1} a_t \langle w_t, a_t \rangle$

$\hat{w}_t$ is an unbiased estimator of w_t :

$$\begin{aligned}
\mathbb{E}_{a_t \sim p_t}\left[\hat{w}_t | \mathcal{F}_{t-1}\right] &= \left(\mathbb{E}_{a \sim p_t}\left[aa^T\right]\right)^{-1} \mathbb{E}_{a_t \sim p_t}\left[a_t \langle w_t, a_t \rangle | \mathcal{F}_{t-1}\right] \\
&= \left(\mathbb{E}_{a \sim p_t}\left[aa^T\right]\right)^{-1} \mathbb{E}_{a_t \sim p_t}\left[a_t a_t^\top | \mathcal{F}_{t-1}\right] w_t \\
&= w_t
\end{aligned}$$

Linear bandits regret

Theorem. *(Linear Bandits Regret).*

[See for example Bubeck '11]

$$R(n) \leq \gamma n + \frac{\log(|\mathcal{A}|)}{\eta} + \eta \sum_{t=1}^{n} \mathbb{E}\mathbb{E}_{a \sim p_t}(\langle \hat{w}_t, a \rangle)^2$$

Exploration over Barycentric Spanner, [Dani, Hayes, Kakade '08]

$$\mathcal{O}(d\sqrt{n \log(|\mathcal{A}|)}) = \mathcal{O}(d^{3/2}\sqrt{n})$$

Uniform over $\mathcal{A}$, [Cesa-Bianchi, Lugosi, '12]

$$\mathcal{O}(\sqrt{dn \log(|\mathcal{A}|)}) = \mathcal{O}(d\sqrt{n})$$

John's distribution [Bubeck, Cesa-Bianchi, Kakade '12]

$$\mathcal{O}(d\sqrt{n})$$

Linear bandits regret
Dimension dependence

Variance bound:

$$\mathbb{E}\left[\mathbb{E}_{a_t \sim p_t}\left[(\langle \hat{w}_t, a\rangle)^2\right]\right] \leq d$$

Linear bandits regret
Dimension dependence

Variance bound:

$$\mathbb{E}\left[\mathbb{E}_{a_t\sim p_t}\left[(\langle \hat{w}_t, a\rangle)^2\right]\right] \le d$$

Dimension dependence

Linear bandits regret
Dimension dependence

Variance bound:

$$\mathbb{E}\left[\mathbb{E}_{a_t \sim p_t}\left[(\langle \hat{w}_t, a\rangle)^2\right]\right] \leq d$$

Dimension dependence

$$R(n) \leq \gamma n + \frac{\log(|\mathcal{A}|)}{\eta} + \underbrace{\eta \sum_{t=1}^{n} \mathbb{E}\mathbb{E}_{a \sim p_t}(\langle \hat{w}_t, a\rangle)^2}_{\leq \eta d n}$$

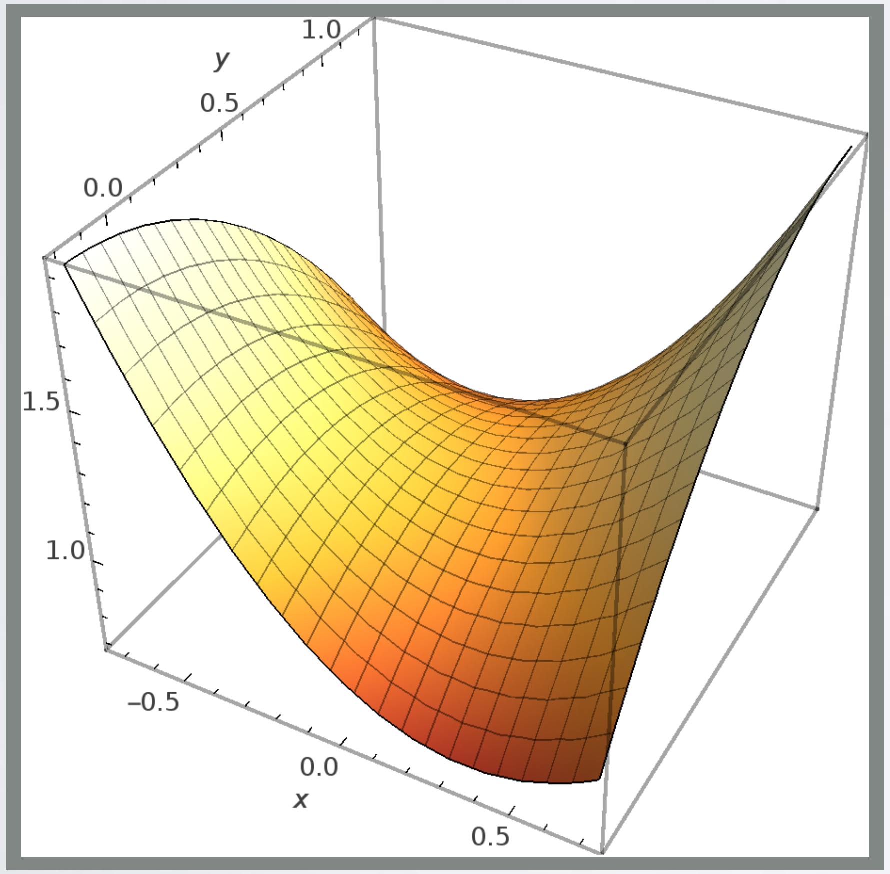
Recap

- Intro to Online Learning
- Linear Bandits
- **Kernel Bandits**

Online Quadratic losses

$$a_t \in \mathcal{A} = \{a \ \text{ s.t. } \ \|a\|_2 \leq 1\}$$

$$\ell_t(a) = \langle b_t, a \rangle + a^\top B_t a$$



$$z = x^2 - .5 * y^2 + x * y - .5 * x + .5y + 1$$

B_t Symmetric and possibly non convex

$$\min_{a \in \mathcal{A}} \ell_t(a)$$

Offline problem has polytime solution

Strong Duality

Peter Bartlett

Niladri Chatterji

Linearization of Quadratic losses

Quadratic losses are linear in the space of $\begin{pmatrix}\text{matrices}\\ \text{vector}\end{pmatrix}$

$$\ell(a) = \langle b_t, a\rangle + a^\top B_t a \quad \Rightarrow \quad \ell(a) = \left\langle \begin{pmatrix} B_t \\ b_t \end{pmatrix}, \begin{pmatrix} aa^\top \\ a \end{pmatrix} \right\rangle$$

We can use the linear bandits machinery

Exponential weights for quadratic bandits

Exponential weights for adversarial quadratic bandits

For $t = 1, \cdots, n$:

Sample mixture $a_t \sim p_t = \underbrace{(1-\gamma)q_t}_{\text{Exploitation}} + \underbrace{\gamma\nu}_{\text{Exploration}}$

See $\langle b_t, a_t \rangle + a_t^\top B_t a_t$

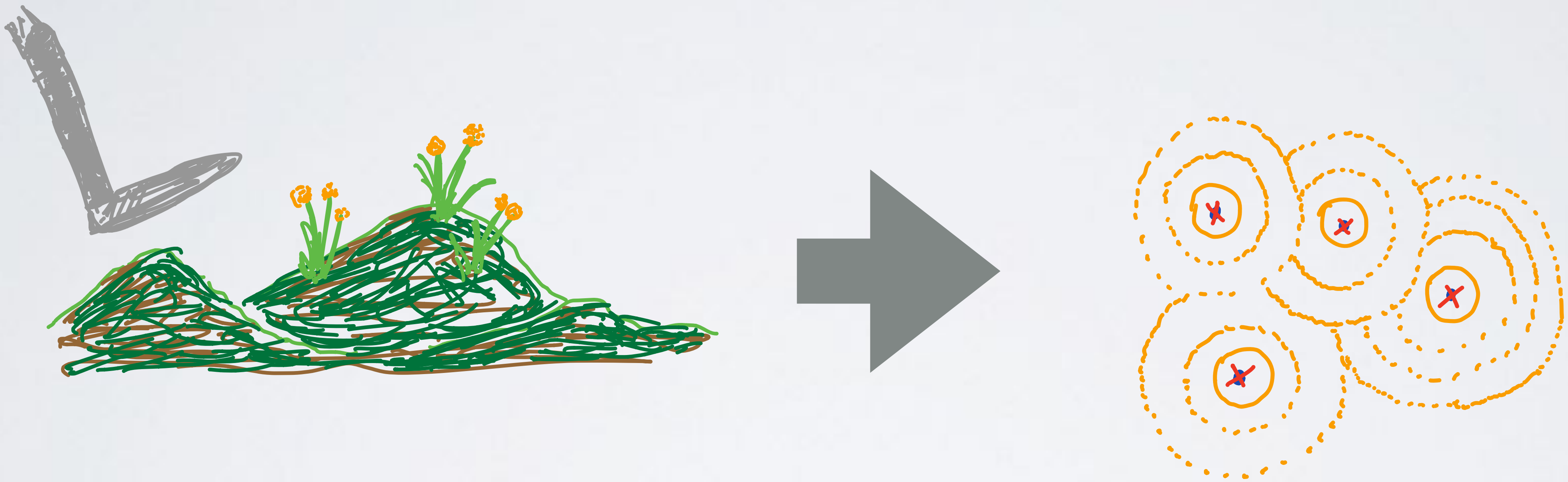
Build loss estimator $\begin{pmatrix} \hat{B}_t \\ \hat{b}_t \end{pmatrix}$

Update $q_t(a) \propto \underbrace{\exp(-\eta(\langle \hat{b}_t, a \rangle + a^\top \hat{B}_t a))q_{t-1}(a)}_{\text{Exponential weights}}$

Sampling is poly time

Beyond "Finite Dimensional" Losses

Evasion games:



Obstacle avoidance

$$\ell_t(a) = \exp(-\|a - w_t\|^2)$$

Gaussian kernel - Infinite dimensional

Space of Quadratics as a Reproducing Kernel Hilbert Space

Feature map
$x \to \Phi(x) \in \mathbb{R}^D$

Dot product
$\mathcal{K}(x, y) = \langle \Phi(x), \Phi(y) \rangle$

The Reproducing Kernel Hilbert Space $\mathcal{H}_\mathcal{K}$ of $\mathcal{K}$

Quadratics losses lie in an RKHS

$$\mathcal{K}(x, y) = \langle x, y \rangle + (\langle x, y \rangle)^2$$

Kernel Bandits

Kernel Bandits

If $\ell_t \in \mathcal{H}_{\mathcal{K}}$ can we leverage linearity?

Kernel Bandits

If $\ell_t \in \mathcal{H}_{\mathcal{K}}$ can we leverage linearity?

Main challenge:

Dimension of $\mathcal{H}_{\mathcal{K}}$ might be infinite.

Naive Linear regret $\mathcal{O}(\sqrt{dn \log(|\mathcal{A}|)})$

Kernel Bandits

If $\ell_t \in \mathcal{H}_{\mathcal{K}}$ can we leverage linearity?

Main challenge:

Dimension of $\mathcal{H}_{\mathcal{K}}$ might be infinite.

Naive Linear regret $\mathcal{O}(\sqrt{dn \log(|\mathcal{A}|)})$

Encouraging facts:

Kernel spaces are "small"

Towards an Algorithm

Algorithm Strategy:

Towards an Algorithm

Algorithm Strategy:

1) Construct an $m < \infty$ dimensional proxy kernel that *uniformly* approximates the original kernel over $\mathcal{A} \times \mathcal{A}$.

$$\mathcal{K}_m(x, y) \approx \mathcal{K}(x, y)$$

Towards an Algorithm

Algorithm Strategy:

I) Construct an $m < \infty$ dimensional proxy kernel that *uniformly* approximates the original kernel over $\mathcal{A} \times \mathcal{A}$.

$$\mathcal{K}_m(x, y) \approx \mathcal{K}(x, y)$$

II) Exponential weights using the proxy kernel. Control the bias.

Towards an Algorithm

Algorithm Strategy:

I) Construct an $m < \infty$ dimensional proxy kernel that *uniformly* approximates the original kernel over $\mathcal{A} \times \mathcal{A}$.

$$\mathcal{K}_m(x, y) \approx \mathcal{K}(x, y)$$

II) Exponential weights using the proxy kernel. Control the bias.

Receive
$\mathcal{K}(w_t, a_t)$

Towards an Algorithm

Algorithm Strategy:

I) Construct an $m < \infty$ dimensional proxy kernel that *uniformly* approximates the original kernel over $\mathcal{A} \times \mathcal{A}$.

$$\mathcal{K}_m(x, y) \approx \mathcal{K}(x, y)$$

II) Exponential weights using the proxy kernel. Control the bias.

Receive
$\mathcal{K}(w_t, a_t)$ →

Towards an Algorithm

Algorithm Strategy:

I) Construct an $m < \infty$ dimensional proxy kernel that *uniformly* approximates the original kernel over $\mathcal{A} \times \mathcal{A}$.

$$\mathcal{K}_m(x, y) \approx \mathcal{K}(x, y)$$

II) Exponential weights using the proxy kernel. Control the bias.

Receive $\mathcal{K}(w_t, a_t)$ → Pretend it was $\mathcal{K}_m(w_t, a_t)$

Kernel functions are "small"

Kernel function evaluations can be uniformly approximated by a small number of basis functions.

Kernel functions are "small"

Kernel function evaluations can be uniformly approximated by a small number of basis functions.

Mercer's Theorem

There exist functions $\{\phi_i\}_{i=1}^{\infty}$ and nonnegative values $\{\mu_i\}_{i=1}^{\infty}$ such that:

$$\mathcal{K}(x, y) = \sum_{i=1}^{\infty} \mu_i \phi_i(x) \phi_i(y) \qquad \forall x, y \in \mathcal{A} \times \mathcal{A}$$

Eigendecay

$$\mathcal{K}(x, y) = \sum_{i=1}^{\infty} \mu_i \phi_i(x) \phi_i(y)$$

Eigendecay

$$\mathcal{K}(x, y) = \sum_{i=1}^{\infty} \mu_i \phi_i(x) \phi_i(y)$$

Gaussian Kernel

$\mathcal{K}(x, y) = \exp(-\|x - y\|^2)$

Sobolev Kernel

$\mathcal{K}(x, y) = \min(x, y)$

Eigendecay

$$\mathcal{K}(x,y) = \sum_{i=1}^{\infty} \mu_i \phi_i(x) \phi_i(y)$$

Gaussian Kernel

$\mathcal{K}(x,y) = \exp(-\|x-y\|^2)$

Exponential decay

$\mu_j \leq Ce^{-\beta j}$

$\{\phi_j(x)\} = \{\sin(j\pi x), \cos(j\pi x)\}$

$\mu_j \approx e^{-cj\log(j)}$

Sobolev Kernel

$\mathcal{K}(x,y) = \min(x,y)$

Eigendecay

$$\mathcal{K}(x, y) = \sum_{i=1}^{\infty} \mu_i \phi_i(x) \phi_i(y)$$

Gaussian Kernel

$\mathcal{K}(x, y) = \exp(-\|x - y\|^2)$

Exponential decay
$\mu_j \leq C e^{-\beta j}$

$\{\phi_j(x)\} = \{\sin(j\pi x), \cos(j\pi x)\}$

$\mu_j \approx e^{-cj \log(j)}$

Sobolev Kernel

$\mathcal{K}(x, y) = \min(x, y)$

Polynomial decay
$\mu_j \leq C j^{-\beta}$

$\phi_j(x) \approx \sin\left(\frac{2j\pi x}{2}\right)$

$\mu_j \approx \frac{1}{j^2}$

Eigendecay

$$\mathcal{K}(x,y) = \sum_{i=1}^{\infty} \mu_i \phi_i(x)\phi_i(y)$$

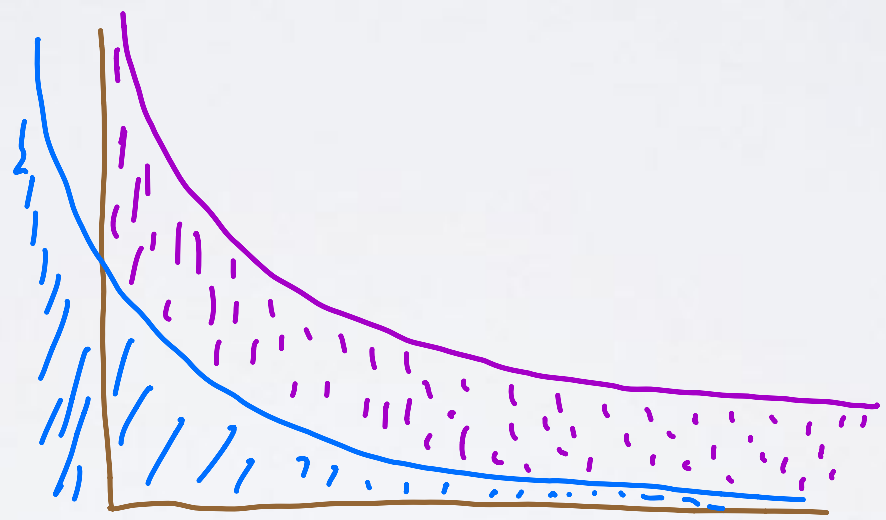
<u>*Gaussian Kernel*</u>

$\mathcal{K}(x,y) = \exp(-\|x-y\|^2)$

Exponential decay
$\mu_j \leq Ce^{-\beta j}$

$\{\phi_j(x)\} = \{\sin(j\pi x), \cos(j\pi x)\}$

$\mu_j \approx e^{-cj\log(j)}$



<u>*Sobolev Kernel*</u>

$\mathcal{K}(x,y) = \min(x,y)$

Polynomial decay
$\mu_j \leq Cj^{-\beta}$

$\phi_j(x) \approx \sin\left(\frac{2j\pi x}{2}\right)$

$\mu_j \approx \frac{1}{j^2}$

Construction of a finite dimensional Proxy Kernel

Eigenfunctions $\{\phi_i\}_{i=1}^{\infty}$ with eigenvalues $\{\mu_i\}_{i=1}^{\infty}$

$$\mathcal{K}(x,y) = \sum_{i=1}^{\infty} \mu_i \phi_i(x) \phi_i(y)$$

Construction of a finite dimensional Proxy Kernel

Eigenfunctions $\{\phi_i\}_{i=1}^{\infty}$ with eigenvalues $\{\mu_i\}_{i=1}^{\infty}$

$$\mathcal{K}(x, y) = \sum_{i=1}^{\infty} \mu_i \phi_i(x) \phi_i(y)$$

Truncate at $i = m$

Construction of a finite dimensional Proxy Kernel

Eigenfunctions $\{\phi_i\}_{i=1}^{\infty}$ with eigenvalues $\{\mu_i\}_{i=1}^{\infty}$

$$\mathcal{K}(x,y) = \sum_{i=1}^{\infty} \mu_i \phi_i(x) \phi_i(y)$$

Truncate at $i = m$

$$\mathcal{K}^o(x,y) = \sum_{i=1}^{m} \mu_i \phi_i(x) \phi_i(y)$$

Construction of a finite dimensional Proxy Kernel

Eigenfunctions $\{\phi_i\}_{i=1}^{\infty}$ with eigenvalues $\{\mu_i\}_{i=1}^{\infty}$

$$\mathcal{K}(x,y) = \sum_{i=1}^{\infty} \mu_i \phi_i(x) \phi_i(y)$$

Truncate at $i = m$

Deterministic Proxy Kernel

$$\mathcal{K}^o(x,y) = \sum_{i=1}^{m} \mu_i \phi_i(x) \phi_i(y)$$

Construction of a finite dimensional Proxy Kernel

Eigenfunctions $\{\phi_i\}_{i=1}^{\infty}$ with eigenvalues $\{\mu_i\}_{i=1}^{\infty}$

$$\mathcal{K}(x,y) = \sum_{i=1}^{\infty} \mu_i \phi_i(x) \phi_i(y)$$

Truncate at $i = m$

Deterministic Proxy Kernel

$$\mathcal{K}^o(x,y) = \sum_{i=1}^{m} \mu_i \phi_i(x) \phi_i(y)$$

Building proxy Kernel with samples
Kernel PCA

Exponential weights for adversarial kernel bandits

Exponential weights for adversarial kernel bandits

Build $\hat{\mathcal{K}}_m(x, y) = \langle \Phi_m(x), \Phi_m(y) \rangle$ from $\mathbb{P}$

Exponential weights for adversarial kernel bandits

Build $\hat{\mathcal{K}}_m(x,y) = \langle \Phi_m(x), \Phi_m(y) \rangle$ from $\mathbb{P}$

For $t = 1, \cdots, n$:

Sample mixture $a_t \sim p_t = \underbrace{(1-\gamma) q_t}_{\text{Exploitation}} + \underbrace{\gamma \nu}_{\text{Exploration}}$

Exponential weights for adversarial kernel bandits

Build $\hat{\mathcal{K}}_m(x, y) = \langle \Phi_m(x), \Phi_m(y) \rangle$ from $\mathbb{P}$

For $t = 1, \cdots, n$:

Sample mixture $a_t \sim p_t = \underbrace{(1-\gamma)q_t}_{\text{Exploitation}} + \underbrace{\gamma\nu}_{\text{Exploration}}$

See $\mathcal{K}(w_t, a_t)$

Exponential weights for adversarial kernel bandits

Build $\hat{\mathcal{K}}_m(x, y) = \langle \Phi_m(x), \Phi_m(y) \rangle$ from $\mathbb{P}$

For $t = 1, \cdots, n$:

Sample mixture $a_t \sim p_t = \underbrace{(1-\gamma)q_t}_{\text{Exploitation}} + \underbrace{\gamma\nu}_{\text{Exploration}}$

See $\mathcal{K}(w_t, a_t)$

Build loss estimator $\hat{w}_t$

Exponential weights for adversarial kernel bandits

Build $\hat{\mathcal{K}}_m(x, y) = \langle \Phi_m(x), \Phi_m(y) \rangle$ from $\mathbb{P}$

For $t = 1, \cdots, n$:

Sample mixture $a_t \sim p_t = \underbrace{(1-\gamma)q_t}_{\text{Exploitation}} + \underbrace{\gamma\nu}_{\text{Exploration}}$

See $\mathcal{K}(w_t, a_t)$

Build loss estimator $\hat{w}_t$

Update $q_t(a) \propto \underbrace{\exp(-\eta\langle \hat{w}_t, \Phi_m(a)\rangle)q_{t-1}(a)}_{\text{Exponential weights}}$

Exponential weights for adversarial kernel bandits

Build $\hat{\mathcal{K}}_m(x, y) = \langle \Phi_m(x), \Phi_m(y) \rangle$ from $\mathbb{P}$

For $t = 1, \cdots, n$:

Sample mixture $a_t \sim p_t = \underbrace{(1-\gamma)q_t}_{\text{Exploitation}} + \underbrace{\gamma\nu}_{\text{Exploration}}$

See $\mathcal{K}(w_t, a_t)$

Build loss estimator $\hat{w}_t$

Update $q_t(a) \propto \underbrace{\exp(-\eta\langle \hat{w}_t, \Phi_m(a)\rangle) q_{t-1}(a)}_{\text{Exponential weights}}$

Sampling might not be poly time

Biased loss estimates

Let $\Sigma_m^{(t)} = \mathbb{E}_{a\sim p_t}\left[\Phi_m(a)\Phi_m(a)^\top\right]$ and set $\hat{w}_t := \mathcal{K}(a_t, y_t)\left((\Sigma_m^{(t)})^{-1}\Phi_m(a_t)\right)$

$\hat{w}_t$ is biased:

$$\begin{aligned}\mathbb{E}_{a_t\sim p_t}\left[\hat{w}_t|\mathcal{F}_{t-1}\right] &= \mathbb{E}\left[\mathcal{K}(a_t, y_t)\left((\Sigma_m^{(t)})^{-1}\Phi_m(a_t)\right)\Big|\mathcal{F}_{t-1}\right] \\ &= \Phi_m(y_t) + \mathbb{E}\left[\underbrace{\left(\mathcal{K}(a_t, y_t) - \hat{\mathcal{K}}_m(a_t, y_t)\right)\left((\Sigma_m^{(t)})^{-1}\Phi_m(a_t)\right)}_{=:\xi_t,\text{ the bias}}\Big|\mathcal{F}_{t-1}\right]\end{aligned}$$

Kernel Bandits Regret

Expected regret

$$R(n) = \mathbb{E}\left[\sum_{t=1}^{n} \mathcal{K}(w_t, a_t) - \inf_{a \in \mathcal{A}} \sum_{t=1}^{n} \mathcal{K}(w_t, a^*)\right]$$

Theorem.

$$R(n) \leq 2\gamma n + \underbrace{\eta m n + \frac{2\epsilon n}{\eta}}_{Bias\ \ variance} + 2\epsilon n + \frac{1}{\eta}\log(|\mathcal{A}|).$$

ξ_t bias

$\mathcal{K}_m$ game vs $\mathcal{K}$ game

Kernel Bandits Regret

Expected regret

$$R(n) = \mathbb{E}\left[\sum_{t=1}^{n} \mathcal{K}(w_t, a_t) - \inf_{a \in \mathcal{A}} \sum_{t=1}^{n} \mathcal{K}(w_t, a^*)\right]$$

Theorem.

$$R(n) \leq 2\gamma n + \underbrace{\eta m n + \frac{2\epsilon n}{\eta}}_{Bias\ variance} + 2\epsilon n + \frac{1}{\eta}\log(|\mathcal{A}|).$$

ξ_t bias

$\mathcal{K}_m$ game vs $\mathcal{K}$ game

Kernel Bandits Regret

Expected regret

$$R(n) = \mathbb{E}\left[\sum_{t=1}^{n} \mathcal{K}(w_t, a_t) - \inf_{a \in \mathcal{A}} \sum_{t=1}^{n} \mathcal{K}(w_t, a^*)\right]$$

Theorem.

$$R(n) \leq 2\gamma n + \underbrace{\eta m n + \frac{2\epsilon n}{\eta}}_{\textit{Bias variance}} + 2\epsilon n + \frac{1}{\eta}\log(|\mathcal{A}|).$$

ξ_t bias

$\mathcal{K}_m$ game vs $\mathcal{K}$ game

Kernel Bandits Regret

Expected regret

$$R(n) = \mathbb{E}\left[\sum_{t=1}^{n} \mathcal{K}(w_t, a_t) - \inf_{a \in \mathcal{A}} \sum_{t=1}^{n} \mathcal{K}(w_t, a^*)\right]$$

Theorem.

$$R(n) \leq \underbrace{2\gamma n + \eta m n + \frac{2\epsilon n}{\eta}}_{Bias\ variance} + 2\epsilon n + \frac{1}{\eta}\log(|\mathcal{A}|).$$

ξ_t bias

Bias variance

$\mathcal{K}_m$ game vs $\mathcal{K}$ game

Kernel Bandits Regret

Corollary.

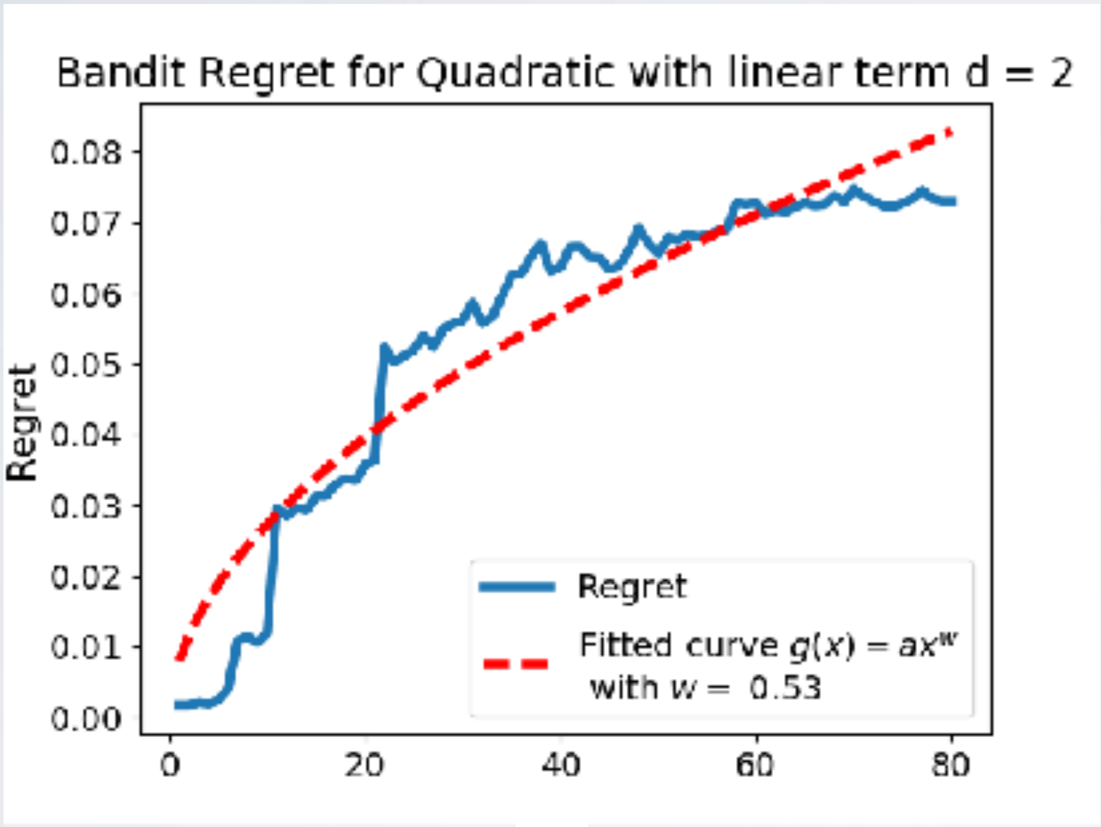
1 Polynomial Decay. $\mu_j \leq C j^{-\beta}$ *and* $\beta > 2$*:*

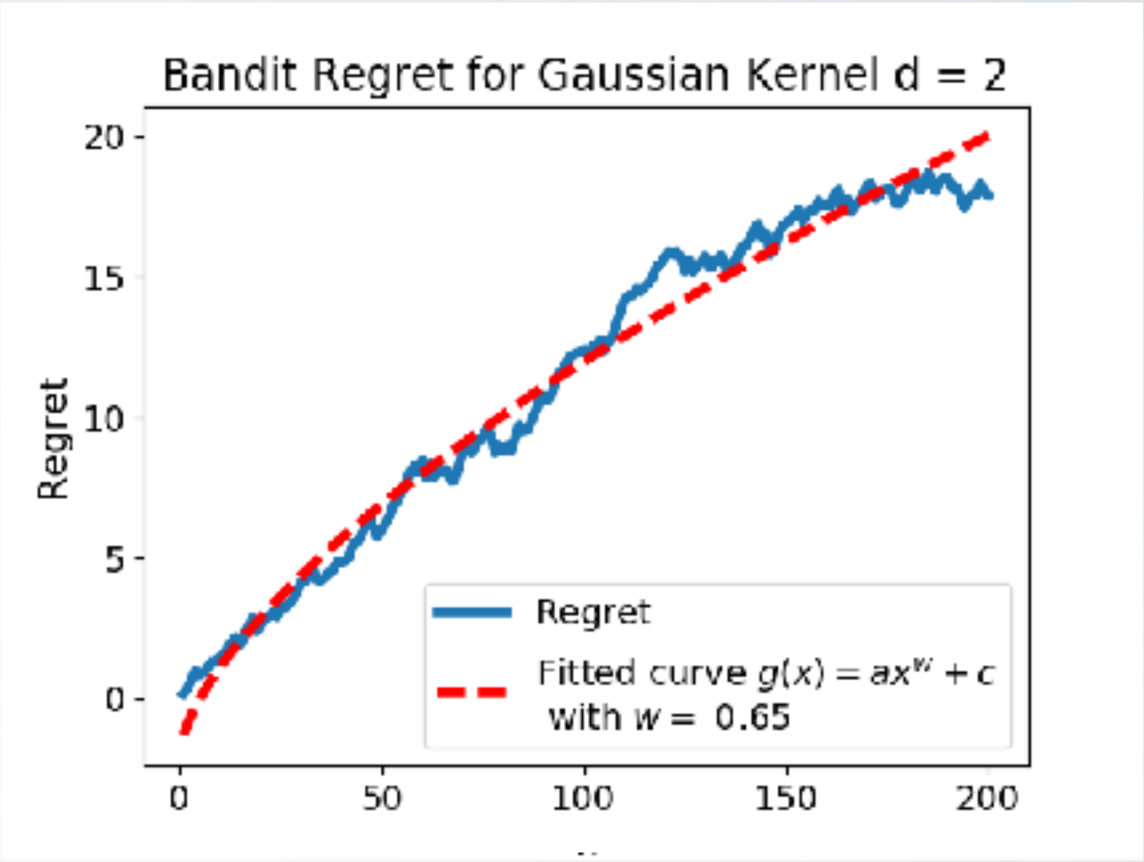
$$R(n) \leq \mathcal{O}\left(\log(|\mathcal{A}|)^{\frac{\beta-2}{2(\beta-1)}} n^{\frac{\beta}{2(\beta-1)}}\right)$$

2 Exponential Decay. $\mu_j \leq C e^{-\beta j}$*:*

$$R(n) \leq \mathcal{O}\left(\sqrt{\log(|\mathcal{A}|)\log(n)n}\right)$$

Experiments





Lower Bound

Lower Bound

Polynomial decay $\mu_j \leq C j^{-\beta}$:

Lower Bound

Polynomial decay $\mu_j \leq Cj^{-\beta}$:

$$\mathcal{R}_n \geq \Omega\left(n^{\frac{\beta+1}{2\beta}}\right)$$

Lower Bound

Polynomial decay $\mu_j \leq Cj^{-\beta}$:

$$\mathcal{R}_n \geq \Omega\left(n^{\frac{\beta+1}{2\beta}}\right)$$

Exponential decay $\mu_j \leq C\exp(-\beta j)$:

Lower Bound

Polynomial decay $\mu_j \leq Cj^{-\beta}$:

$$\mathcal{R}_n \geq \Omega\left(n^{\frac{\beta+1}{2\beta}}\right)$$

Exponential decay $\mu_j \leq C\exp(-\beta j)$:

$$\mathcal{R}_n \geq \Omega\left(n^{1/2}\right)$$

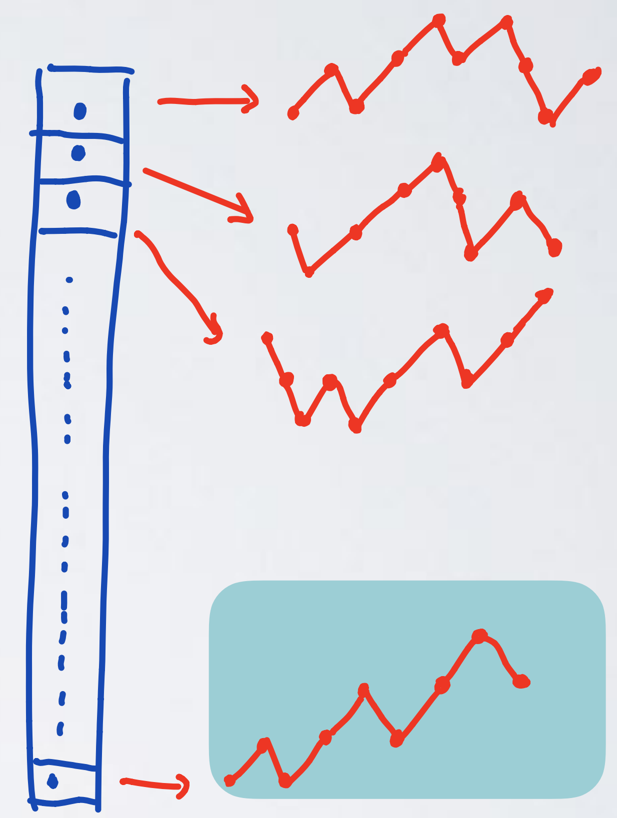
Lower Bound

Polynomial decay $\mu_j \leq Cj^{-\beta}$:

$$\mathcal{R}_n \geq \Omega\left(n^{\frac{\beta+1}{2\beta}}\right)$$

Exponential decay $\mu_j \leq C\exp(-\beta j)$:

$$\mathcal{R}_n \geq \Omega\left(n^{1/2}\right)$$



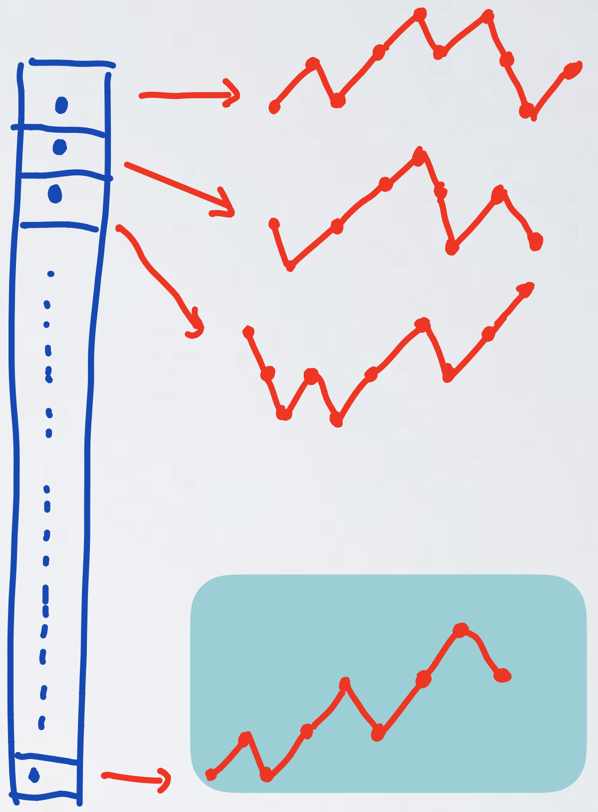
Lower Bound

Polynomial decay $\mu_j \leq C j^{-\beta}$:

$$\mathcal{R}_n \geq \Omega\left(n^{\frac{\beta+1}{2\beta}}\right)$$

Exponential decay $\mu_j \leq C \exp(-\beta j)$:

$$\mathcal{R}_n \geq \Omega\left(n^{1/2}\right)$$



$$\mathcal{A} = \{(A_j)_{j=1}^{\infty} \text{ s.t. } |A_j| = 1 \ \forall j\}$$

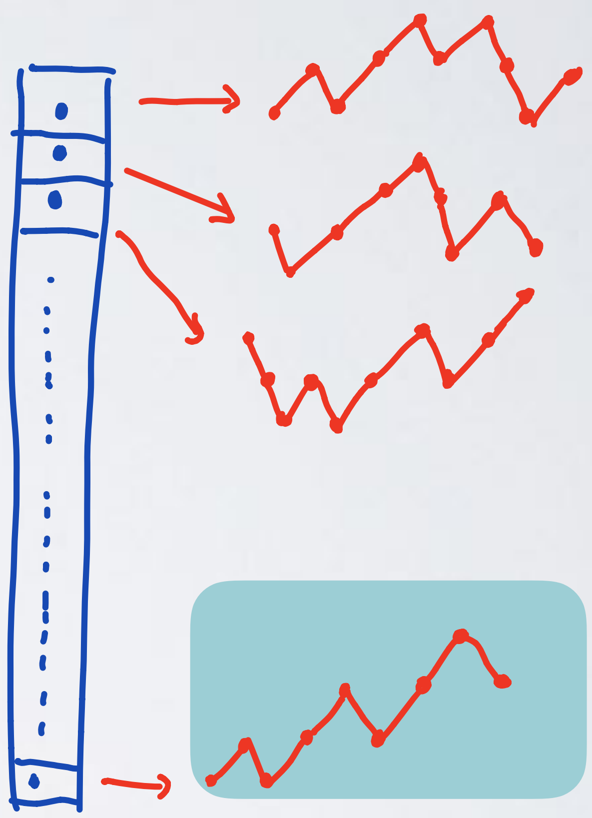
Lower Bound

Polynomial decay $\mu_j \leq Cj^{-\beta}$:

$$\mathcal{R}_n \geq \Omega\left(n^{\frac{\beta+1}{2\beta}}\right)$$

Exponential decay $\mu_j \leq C\exp(-\beta j)$:

$$\mathcal{R}_n \geq \Omega\left(n^{1/2}\right)$$



$$\mathcal{A} = \{(A_j)_{j=1}^{\infty} \text{ s.t. } |A_j| = 1\ \forall j\}$$
$$\mathcal{W} = \{(w_j)_{j=1}^{\infty} \text{ s.t. } |w_j| = \mu_j\ \forall j\}$$

Final thoughts

Thanks for your attention