Eluder Dimension and the Sample Complexity of Optimistic Exploration

Xuhui Liu

LAMDA, NJUAI

July 22, 2020

Outline

Table of Contents

Problem formulation

- A set of actions $\mathscr{A}$.
- A set of real-valued functions $\mathscr{F} = \{f_\rho : \mathscr{A} \mapsto \mathbb{R} | \rho \in \Theta\}$.
- At each time t, the agent is presented with a subset $\mathscr{A}_t \subset \mathscr{A}$.
- The agent selects an action $A_t \in \mathscr{A}_t$, and then recieve a reward R_t.
- H_t is the history $(\mathscr{A}_1, A_1, R_1, \cdots, \mathscr{A}_{t-1}, R_{t-1}, \mathscr{A}_t)$.
- The agent employs a policy $\pi = \{\pi_t | t \in \mathbb{N}\}$, $\pi_t(H_t)$ is a distribution over $\mathscr{A}$ with support $\mathscr{A}_t$.
- We assume that $\mathbb{E}[R_t | H_T, \theta, A_t] = f_\theta(A_t)$.

Problem formulation

- The T-proiod regret of a policy π is defined by

$$\text{Regret}(T, \pi, \theta) \sum_{t=1}^{T} \mathbb{E}[\max_{a \in \mathscr{A}_t} f_\theta(a) - f_\theta(A_t) | \theta].$$

- The T-period Bayesian regret is defined by $\mathbb{E}[\text{Regret}(T, \pi, \theta)]$, where the expectation is taken with respect to the prior distribution over θ. Hence,

$$\text{BayesRegret}(T, \pi) = \sum_{t=1}^{T} \mathbb{E}[\max_{a \in \mathscr{A}_t} f_\theta(a) - f_\theta(A_t)].$$

VC dimension

Given a sample $\mathcal{D}_n = \{(x_1, y_1), \ldots, (x_n, y_n)\}$, and define $S = \{x_1, \ldots, x_n\}$. Consider the set

$$\mathcal{H}_S = \mathcal{H}_{x_1,\ldots,x_n} = \{(h(x_1), \ldots, h(x_n) : h \in \mathcal{H}\}.$$

Definition (Growth Function). *The growth function is the maximum number of ways into which n points can be classified by the function class:*

$$G_{\mathcal{H}}(n) = \sup_{x_1,\ldots,x_n} |\mathcal{H}_S|.$$

Definition (VC Dimension). *The VC dimension of a class $\mathcal{H}$ is the largest $n = d_{VC}(\mathcal{H})$ such that*

$$G_{\mathcal{H}}(n) = 2^n.$$

In other words, VC dimension of a function class $\mathcal{H}$ is the cardinality of the largest set that it can shatters.

Example

- A finite binary-valued function class $\mathscr{F} = \{f_\rho : \mathscr{A} \mapsto \{0,1\} | \rho \in \{1,\ldots,n\}\}$.
- A finite action set $\mathscr{A} = \{1,\ldots,n\}$.
- Let $f_\rho(a) = \mathbf{1}(\rho = a)$.
- In time step t, $R_t = f_\rho(A_t)$.
- If ρ is uniformly distributed over $\{1,\ldots,n\}$, it's easyy to see that the Bayesian regret grows linearly with n.

Explanation

We fomulate this probelm as a supervised learning problem:

- At each time step, an action A_t is sampled uniformly from $\mathscr{A}$ and the reward $f_\theta(A_t)$ is observed.
- For large n, the time it takes to effectively learn to predict $f_\theta(A_t)$.

Eluder dimension

Definition 2. An action $a \in \mathcal{A}$ is *ϵ-dependent* on actions $\{a_1, \ldots, a_n\} \subseteq \mathcal{A}$ with respect to $\mathcal{F}$ if any pair of functions $f, \tilde{f} \in \mathcal{F}$, satisfying $\sqrt{\sum_{i=1}^{n}(f(a_i) - \tilde{f}(a_i))^2} \leq \epsilon$ also satisfies $f(a) - \tilde{f}(a) \leq \epsilon$. Further, a is ϵ-independent of $\{a_1, \ldots, a_n\}$ with respect to $\mathcal{F}$ if a is not ϵ-dependent on $\{a_1, \ldots, a_n\}$.

Definition 3. The *ϵ-eluder dimension* $\dim_E(\mathcal{F}, \epsilon)$ is the length d of the longest sequence of elements in $\mathcal{A}$ such that, for some $\epsilon' \geq \epsilon$, every element is ϵ'-independent of its predecessors.

VC dimension

DEFINITION 4. An action a is *VC-independent* of $\tilde{\mathcal{A}} \subseteq \mathcal{A}$ if for any $f, \tilde{f} \in \mathcal{F}$, there exists some $\bar{f} \in \mathcal{F}$, which agrees with f on a and with $\tilde{f}$ on $\tilde{\mathcal{A}}$; that is, $\bar{f}(a) = f(a)$ and $\bar{f}(\tilde{a}) = \tilde{f}(\tilde{a})$ for all $\tilde{a} \in \tilde{\mathcal{A}}$. Otherwise, a is *VC-dependent* on $\tilde{\mathcal{A}}$.

DEFINITION 5. The VC dimension of a class of binary-valued functions with domain $\mathcal{A}$ is the largest cardinality of a set $\tilde{\mathcal{A}} \subseteq \mathcal{A}$ such that every $a \in \tilde{\mathcal{A}}$ is VC-independent of $\tilde{\mathcal{A}} \backslash \{a\}$.

Eluder dimension for common function class

- Finite action spaces.
  For all $\epsilon > 0$, the ϵ-eluder dimension of $\mathscr{A}$ is bounded by $|\mathscr{A}|$.
- Euclid space.
  The dimension of Euclid space for linear functions is 0-elider dimension of this space.

Linear models

- Reward functions are parameterized by a vector $\theta \in \Theta \subset \mathbb{R}^d$.
- Feature mapping $\phi : \mathscr{A} \mapsto \mathbb{R}^d$ such that $f_\theta(a) =< \phi(a), \theta >$.
- An ellipsoidal confidence set $\Theta_t = \{\rho \in \mathbb{R}^d : ||\rho - \hat{\theta}_t||_{V_t} \leq \sqrt{\beta_t}\}$, where $V_t := \sum_{k=1}^{t} \phi(A_t)\phi(A_t)^T + \lambda I$ for some $\lambda \in \mathbb{R}$ captures the amount of exploration carried put in each direction up to time t.

Algorithm 2 (Linear-Gaussian UCB)

1. **Update Statistics**:
   $\mu_t \leftarrow \mathbb{E}[\theta \mid H_t]$
   $\Sigma_t \leftarrow \mathbb{E}[(\theta - \mu_t)(\theta - \mu_t)^\top \mid H_t]$
2. **Select Action**:
   $\bar{A}_t \in \arg\max_{a \in \mathcal{A}} \{\langle \phi(a), \mu_t \rangle + \beta \log(t) \|\phi(a)\|_{\Sigma_t}\}$
3. **Increment** t **and Go to Step 1**.

Theorem

Assume there exist constants γ *and* S *such that for all* $a \in \mathscr{A}$ *and* $\rho \in \Theta$, $||\rho||_2 \leq S$ *and* $||\phi(a)||_2 \leq \gamma$. *Then* $dim_E(\mathscr{F}, \epsilon) \leq 3d(e/(e-1)) \log\{3 + 3((2S)/\epsilon)^2\} + 1$.

Linear models

Define

$$\omega_k := \sup\left\{(f_{\rho_1} - f_{\rho_2})(a_k) : \sqrt{\sum_{i=1}^{k-1}(f_{\rho_1} - f_{\rho_2})^2(a_i)} \leq \epsilon' \rho_1, \rho_2 \in \Theta\right\}.$$

Then the ϵ-eluder dimension is the longest sequence such that $\omega_k > \epsilon'$.

- Let $\phi_k = \phi(a_k)$,
- $\rho = \rho_1 - \rho_2$,
- and $\Phi_k = \sum_{i=1}^{k-1} \phi_i \phi_i^T$.
- In this case, $\sum_{i=1}^{k-1}(f_{\rho_1} - f_{\rho_2})^2(a_i) = \rho^T \Phi_k \rho$,
- and by the triangle inequality $||\rho||_2 \leq 2S$.

Step 1. If $\omega_k \geq \epsilon'$, then $\phi_k^T V_k^{-1} \phi_k \geq \frac{1}{2}$, where $V_k := \Phi_k + \lambda I$ and $\lambda = (\epsilon'/(2S))^2$.

Proof.

$$\begin{aligned} \omega_k &\leq \max\{\rho^T \phi_k : \rho^T \Phi_k \rho \leq (\epsilon')^2, \rho^T I \rho \leq (2S)^2\} \\ &\leq \max\{\rho^T \phi_k : \rho^T V_k \rho \leq 2(\epsilon')^2\} \\ &= \sqrt{2(\epsilon')^2} ||\phi_k||_{V_k^{-1}} \end{aligned}$$

□

Linear models

Step 2. If $\omega_i \geq \epsilon'$, for each $i < k$, then $\det V_k \geq \lambda^d (\frac{3}{2})^{k-1}$ and $\det V_k \leq ((\gamma^2(k-1))/d + \lambda)^d$.

Proof.

PROOF. Since $V_k = V_{k-1} + \phi_k \phi_k^T$, using the matrix determinant lemma,

$$\det V_k = \det V_{k-1}(1 + \phi_t^T V_k^{-1} \phi_t) \geq \det V_{k-1}\left(\frac{3}{2}\right) \geq \cdots \geq \det[\lambda I]\left(\frac{3}{2}\right)^{k-1} = \lambda^d \left(\frac{3}{2}\right)^{k-1}.$$

Recall that $\det V_k$ is the product of the eigenvalues of V_k, whereas $\mathrm{trace}[V_k]$ is the sum. As noted in Dani et al. [16], $\det V_k$ is maximized when all eigenvalues are equal. This implies $\det V_k \leq ((\mathrm{trace}[V_k])/d)^d \leq ((\gamma^2(t-1))/d + \lambda)^d$. □

□

Step 3. Complete proof.

Proof.

Step 2 shows that $(\frac{3}{2})^{(k-1)d} \leq \alpha_0[(k-1)/d] + 1$, where $\alpha_0 = \gamma^2/\lambda = (2S\gamma/\epsilon')^2$.
Let $B(x, \alpha) = \max\{B : (1+x)^B \leq \alpha B + 1\}$, then the number of times $\omega_k > \epsilon'$ can occur is bounded by $dB(1/2, \alpha_0) + 1$.

We now derive an explicit bound on $B(x, \alpha)$ for any $x \leq 1$. Note that any $B \geq 1$ must satisfy the inequality $\ln\{1+x\}B \leq \ln\{1+\alpha\} + \ln B$. Since $\ln\{1+x\} \geq x/(1+x)$, using the transformation of variables $y = B[x/(1+x)]$ gives

$$y \leq \ln\{1+\alpha\} + \ln\frac{1+x}{x} + \ln y \leq \ln\{1+\alpha\} + \ln\frac{1+x}{x} + \frac{y}{e} \quad \Longrightarrow \quad y \leq \frac{e}{e-1}\left(\ln\{1+\alpha\} + \ln\frac{1+x}{x}\right).$$

This implies $B(x, \alpha) \leq ((1+x)/x)(e/(e-1))(\ln\{1+\alpha\} + \ln((1+x)/x))$. The claim follows by plugging in $\alpha = \alpha_0$ and $x = 1/2$. □

□

Table of Contents

Posterior sampling

Algorithm 3 (Independent posterior sampling)

1. **Sample Model**:
   $\hat{\theta}_t \sim N(\mu_{t-1}, \Sigma_{t-1})$
2. **Select Action**:
   $A_t \in \arg\max_{a \in \mathcal{A}} \hat{\theta}_t(a)$
3. **Update Statistics**: For each a,
   $\mu_{ta} \leftarrow \mathbb{E}[\theta_a \mid H_t]$
   $\Sigma_{taa} \leftarrow \mathbb{E}[(\theta_a - \mu_{ta})^2 \mid H_t]$
4. **Increment t and Go to Step 1**.

Algorithm 4 (Linear posterior sampling)

1. **Sample Model**:
   $\hat{\theta}_t \sim N(\mu_{t-1}, \Sigma_{t-1})$
2. **Select Action**:
   $A_t \in \arg\max_{a \in \mathcal{A}} \langle \phi(a), \hat{\theta}_t \rangle$
3. **Update Statistics**:
   $\mu_t \leftarrow \mathbb{E}[\theta \mid H_t]$
   $\Sigma_t \leftarrow \mathbb{E}[(\theta - \mu_t)(\theta - \mu_t)^\top \mid H_t]$
4. **Increment t and Go to Step 1**.

UCB regret decomposition

Consider a UCB algorithm with a UCB sequence $U = \{U_t | t \in \mathbb{N}\}$. Let $\bar{A}_t \in \arg\max_{a \in \mathscr{A}_t} U_t(a)$ and $A^* \in \arg\max_{a \in \mathscr{A}_t} f_\theta(a)$. We have the following regret decomposition:

$$
\begin{aligned}
f_\theta(A_t^*) - f_\theta(\bar{A}_t) &= f_\theta(A_t^*) - U_t(\bar{A}_t) + U_t(\bar{A}_t) - f_\theta(\bar{A}_t) \\
&\le [f_\theta(A_t^*) - U_t(A_t^*)] + [U_t(\bar{A}_t) - f_\theta(\bar{A}_t)]
\end{aligned}
$$

Bayesian regret:

$$
\text{BayesRegret}(T, \pi^U) \le \mathbb{E} \sum_{t=1}^{T} [U_t(\bar{A}_t) - f_\theta(\bar{A}_t)] + \mathbb{E} \sum_{t=1}^{T} [f_\theta(A_t^*) - U_t(A_t^*)]
$$

Posterior sampling regret decomposition

PROPOSITION 1. *For any UCB sequence* $\{U_t \mid t \in \mathbb{N}\}$,

$$\mathrm{BayesRegret}(T, \pi^{PS}) = \mathbb{E}\sum_{t=1}^{T}[U_t(A_t) - f_\theta(A_t)] + \mathbb{E}\sum_{t=1}^{T}[f_\theta(A_t^*) - U_t(A_t^*)]$$

for all $T \in \mathbb{N}$.

Posterior sampling regret decomposition

PROOF. Note that, conditioned on H_t, the optimal action A_t^* and the action A_t selected by posterior sampling are identically distributed, and U_t is deterministic. Hence $\mathbb{E}[U_t(A_t^*) \mid H_t] = \mathbb{E}[U_t(A_t) \mid H_t]$. Therefore

$$
\begin{aligned}
\mathbb{E}[f_\theta(A_t^*) - f_\theta(A_t)] &= \mathbb{E}[\mathbb{E}[f_\theta(A_t^*) - f_\theta(A_t) \mid H_t]] \\
&= \mathbb{E}[\mathbb{E}[U_t(A_t) - U_t(A_t^*) + f_\theta(A_t^*) - f_\theta(A_t) \mid H_t]] \\
&= \mathbb{E}[\mathbb{E}[U_t(A_t) - f_\theta(A_t) \mid H_t] + \mathbb{E}[f_\theta(A_t^*) - U_t(A_t^*) \mid H_t]] \\
&= \mathbb{E}[U_t(A_t) - f_\theta(A_t)] + \mathbb{E}[f_\theta(A_t^*) - U_t(A_t^*)].
\end{aligned}
$$

Summing over t gives the result. □

Bounds for general function classes

- Assumption 1. For all $f \in \mathscr{F}$ and $a \in \mathscr{A}$, $f(a) \in [0, C]$.
- Assumption 2. For all $t \in \mathbb{N}$, $R_t - f\theta(A_t)$ conditioned on (H_t, θ, A_t) is σ-sub-Gaussian.

Confidence Bounds

Let $L_{2,t}(f) = \sum_1^{t-1} (f(A_t) - R_t)^2$, and $\hat{f}_t^{LS} \in \arg\min_{f \in \mathscr{F}} L_{2,t}(f)$.

LEMMA 3. *For any $\delta > 0$ and $f: \mathscr{A} \mapsto \mathbb{R}$, with probability at least $1 - \delta$,*

$$L_{2,t}(f) \geq L_{2,t}(f_\theta) + \frac{1}{2}\|f - f_\theta\|_{2,E_t}^2 - 4\sigma^2 \log(1/\delta)$$

simultaneously for all $t \in \mathbb{N}$.

Confidence bounds

$$\beta_t^*(\mathcal{F}, \delta, \alpha) := 8\sigma^2 \log(N(\mathcal{F}, \alpha, \|\cdot\|_\infty)/\delta) + 2\alpha t\left(8C + \sqrt{8\sigma^2 \ln(4t^2/\delta)}\right).$$

PROPOSITION 6. *For all $\delta > 0$ and $\alpha > 0$, if*

$$\mathcal{F}_t = \{f \in \mathcal{F} : \|f - \hat{f}_t^{LS}\|_{2, E_t} \leq \sqrt{\beta_t^*(\mathcal{F}, \delta, \alpha)}\}$$

for all $t \in \mathbb{N}$, then

$$\mathbb{P}\left(f_\theta \in \bigcap_{t=1}^{\infty} \mathcal{F}_t\right) \geq 1 - 2\delta.$$

Confidence bounds

DEFINITION 1. The *Kolmogorov dimension* of a function class $\mathscr{F}$ is given by

$$\dim_K(\mathscr{F}) = \limsup_{\alpha \downarrow 0} \frac{\log N(\mathscr{F}, \alpha, \|\cdot\|_\infty)}{\log(1/\alpha)}.$$

In particular, we have the following result.

PROPOSITION 7. *For any fixed class of functions* $\mathscr{F}$,

$$\beta_t^*(\mathscr{F}, 1/t^2, 1/t^2) = 16(1 + o(1) + \dim_K(\mathscr{F})) \log t.$$

Bayesian regret bounds

LEMMA 4. *For all* $T \in \mathbb{N}$, *if* $\inf_{\rho \in \mathcal{F}_\tau} f_\rho(a) \leq f_\theta(a) \leq \sup_{\rho \in \mathcal{F}_\tau} f_\rho(a)$ *for all* $\tau \in \mathbb{N}$ *and* $a \in \mathcal{A}$ *with probability at least* $1 - 1/T$, *then*

$$\text{BayesRegret}(T, \pi^{PS}) \leq C + \mathbb{E} \sum_{t=1}^{T} w_{\mathcal{F}_t}(A_t).$$

Bayesian regret bounds

PROPOSITION 8. *If* $(\beta_t \geq 0 \mid t \in \mathbb{N})$ *is a nondecreasing sequence and* $\mathcal{F}_t := \{f \in \mathcal{F}: \|f - \hat{f}_t^{LS}\|_{2, E_t} \leq \sqrt{\beta_t}\}$, *then*

$$\sum_{t=1}^{T} \mathbf{1}(w_{\mathcal{F}_t}(A_t) > \epsilon) \leq \left(\frac{4\beta_T}{\epsilon^2} + 1\right) \dim_E(\mathcal{F}, \epsilon)$$

for all $T \in \mathbb{N}$ *and* $\epsilon > 0$.

Bayesian regret bounds

LEMMA 5. *If* $(\beta_t \geq 0 \mid t \in \mathbb{N})$ *is a nondecreasing sequence and* $\mathcal{F}_t := \{f \in \mathcal{F} \colon \|f - \hat{f}_t^{LS}\|_{2, E_t} \leq \sqrt{\beta_t}\}$*, then*

$$\sum_{t=1}^{T} w_{\mathcal{F}_t}(A_t) \leq 1 + \dim_E(\mathcal{F}, T^{-1})C + 4\sqrt{\dim_E(\mathcal{F}, T^{-1})\beta_T T}$$

for all $T \in \mathbb{N}$.

PROPOSITION 10. *For any fixed class of functions* $\mathcal{F}$,

$$\text{BayesRegret}(T, \pi^{PS}) \leq 1 + [\dim_E(\mathcal{F}, T^{-1}) + 1]C + 16\sigma\sqrt{\dim_E(\mathcal{F}, T^{-1})(1 + o(1) + \dim_K(\mathcal{F}))\log(T)T}$$

Table of Contents

Proposition 8. *If* $(\beta_t \geq 0 \mid t \in \mathbb{N})$ *is a nondecreasing sequence and* $\mathcal{F}_t := \{f \in \mathcal{F} : \|f - \hat{f}_t^{LS}\|_{2,E_t} \leq \sqrt{\beta_t}\}$, *then*

$$\sum_{t=1}^{T} \mathbf{1}(w_{\mathcal{F}_t}(A_t) > \epsilon) \leq \left(\frac{4\beta_T}{\epsilon^2} + 1\right) \dim_E(\mathcal{F}, \epsilon)$$

for all $T \in \mathbb{N}$ *and* $\epsilon > 0$.

Step 1: If $\omega_t(A_t) > \epsilon$, then A_t is ϵ-dependent on fewer than $4\beta_T/\epsilon^2$ disjooint subsequences of $(A_1, \ldots, A_{t-1})$ for $T > t$.

$$\begin{aligned}
\omega_t(A_t) > \epsilon &\Rightarrow \exists \bar{f}, \underline{f} \in \mathscr{F}_t, \bar{f}(A_t) - \underline{f}(A_t) > \epsilon \\
&\Rightarrow \text{If } A_t \text{ is } \epsilon\text{-dependent on a subsequence of } (A_1, \ldots, A_{t-1}), \\
&\quad\ \text{then} \sum_{j=1}^{k} (\bar{f}(A_j) - \underline{f}(A_j))^2 > \epsilon^2 \\
&\Rightarrow \text{If } A_t \text{ is } \epsilon\text{-dependent on } K \text{ disjoint subsequences} \\
&\quad\ \text{then} ||\bar{f} - \underline{f}||_{2,E_t}^2 > K\epsilon^2
\end{aligned}$$

$$||\bar{f} - \underline{f}||_{2,E_t} \leq ||\bar{f} - \hat{f}_t^{LS}||_{2,E_t} + ||\underline{f} - \hat{f}_t^{LS}||_{2,E_t} \leq 2\sqrt{\beta_t} \leq 2\sqrt{\beta_T}$$

It's follows that $K < 4\beta_T/\epsilon^2$.

Step 2: For any action sequence $(a_1, \ldots, a_\tau)$, there is some element a_j that is ϵ-dependent on at least $\tau/d - 1$ disjoint subsequences, where $d := \dim_E(\mathscr{F}, \epsilon)$.

Step 3:
Consider taking $(a_1, \ldots, a_\tau)$ to be the subsequence of $(A_1, \ldots, A_T)$, in which the elements satisfies $\omega_t(A_t) > \epsilon$.

$$\Rightarrow \ \tau/d - 1 \leq 4\beta_T/\epsilon^2.$$

LEMMA 5. *If* $(\beta_t \geq 0 \mid t \in \mathbb{N})$ *is a nondecreasing sequence and* $\mathcal{F}_t := \{f \in \mathcal{F} : \|f - \hat{f}_t^{LS}\|_{2, E_t} \leq \sqrt{\beta_t}\}$, *then*

$$\sum_{t=1}^{T} w_{\mathcal{F}_t}(A_t) \leq 1 + \dim_E(\mathcal{F}, T^{-1})C + 4\sqrt{\dim_E(\mathcal{F}, T^{-1})\beta_T T}$$

for all $T \in \mathbb{N}$.

Step 1:

PROOF. To reduce notation, write $d = \dim_E(\mathcal{F}, T^{-1})$ and $w_t = w_t(A_t)$. Reorder the sequence $(w_1, \ldots, w_T) \to (w_{i_1}, \ldots, w_{i_T})$, where $w_{i_1} \geq w_{i_2} \geq \cdots \geq w_{i_T}$. We have

$$\sum_{t=1}^{T} w_{\mathcal{F}_t}(A_t) = \sum_{t=1}^{T} w_{i_t} = \sum_{t=1}^{T} w_{i_t} \mathbf{1}\{w_{i_t} \leq T^{-1}\} + \sum_{t=1}^{T} w_{i_t} \mathbf{1}\{w_{i_t} > T^{-1}\} \leq 1 + \sum_{t=1}^{T} w_{i_t} \mathbf{1}\{w_{i_t} \geq T^{-1}\}.$$

Step 2:

- $d = \dim_E(\mathscr{F}, T^{-1}) \geq \dim_E(\mathscr{F}, \epsilon)$, for all $\epsilon > T^{-1}$.
- $\sum_{t=1}^{T} \mathbf{1}(\omega_t > \epsilon) < ((4\beta_T)/\epsilon^2 + 1)d$.
- $\omega_t > \epsilon \Rightarrow \sum_{k=1}^{T} \mathbf{1}(\omega_k > \epsilon) \geq t$.
- $\epsilon < \sqrt{(4\beta_T d)/(t-d)}$.
- $w_t \leq \min\{C, \sqrt{(4\beta_T d)/(t-d)}\}$.

Step 3:

$$\sum_{t=1}^{T} w_{i_t} \mathbf{1}\{w_{i_t} > T^{-1}\} \leq dC + \sum_{t=d+1}^{T} \sqrt{\frac{4d\beta_T}{t-d}} \leq dC + 2\sqrt{d\beta_T} \int_{t=0}^{T} \frac{1}{\sqrt{t}}\, dt = dC + 4\sqrt{d\beta_T T}.$$

Thanks!