15-859 Algorithms for Big Data Assignment 1

Fan Pu Zeng
fzeng@andrew.cmu.edu

1: Scratcy Scratch

$$T = \left\{ (x, \sin \frac{1}{x}) : x \in (0, 1] \right\} \cup \{(0, 0)\}$$

$$T = \{(x, \sin \frac{1}{x}) : x \in (0, 1]\} \cup \{(0, 0)\}$$
$$T = \left\{ \left(x, \sin \frac{1}{x} \right) : x \in (0, 1] \right\} \cup \{(0, 0)\}$$
$$T = \left\{ \left(x, \sin \frac{1}{x} \right) : x \in (0, 1] \right\} \cup \{(0, 0)\}$$

$$\left(\frac{1}{x} \right)$$
$$\left[\frac{1}{x} \right]$$
$$\left\{ \frac{1}{x} \right\}$$
$$\left\| \frac{1}{x} \right\|$$
$$|\frac{1}{x}|$$
$$\left\| \frac{1}{x} \right\|$$
$$\left\lceil \frac{1}{x} \right\rceil$$

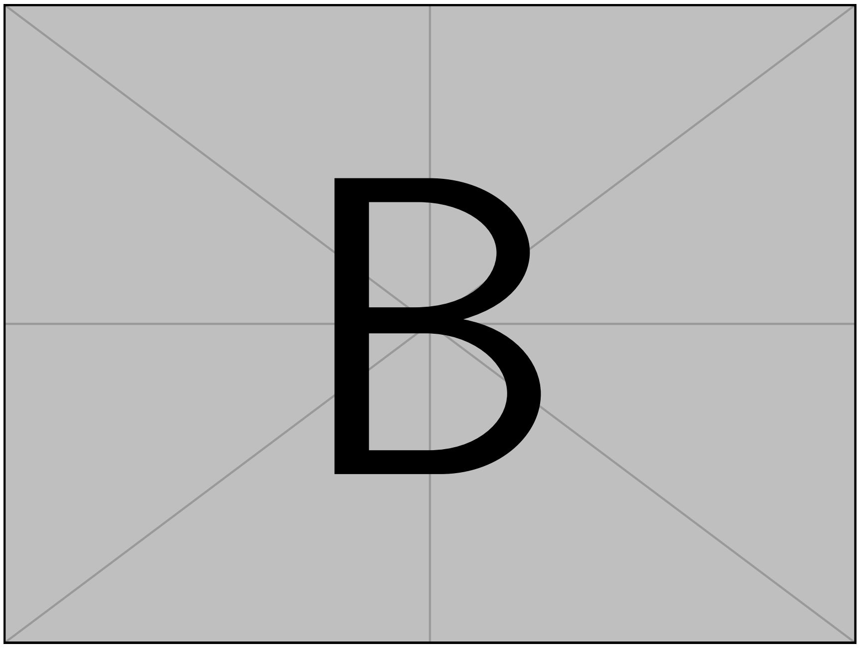

Figure 1: Landing/Login Page Example

We denote the historical trajectory as $\tau = (s_1, a_1, \ldots, a_{t-1}, s_t)$ and action-observation history (AOH) for player i as $\tau^i = \left(\Omega^i\left(s_1\right), a_1, \ldots, a_{t-1}, \Omega^i\left(s_t\right)\right)$, which encodes the trajectory from player i 's point of view.

We evaluated the policy periodically during training by testing it without exploration noise. Figure 2 shows the performance curve for a selection of environments. We also report results with components of our algorithm.

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

$$X = (X_1, ..., X_n) \qquad \text{(using "...")}$$
$$X = (X_1, \dots, X_n) \qquad \text{(using "\textbackslash dots")}$$

"Hello World!"

Problem: Show that if $(x_n)_n$ *converges to* x *in the usual sense, then* $\lim_{n\to\infty} x_n = \lim_{\mathcal{F}} x_n$.

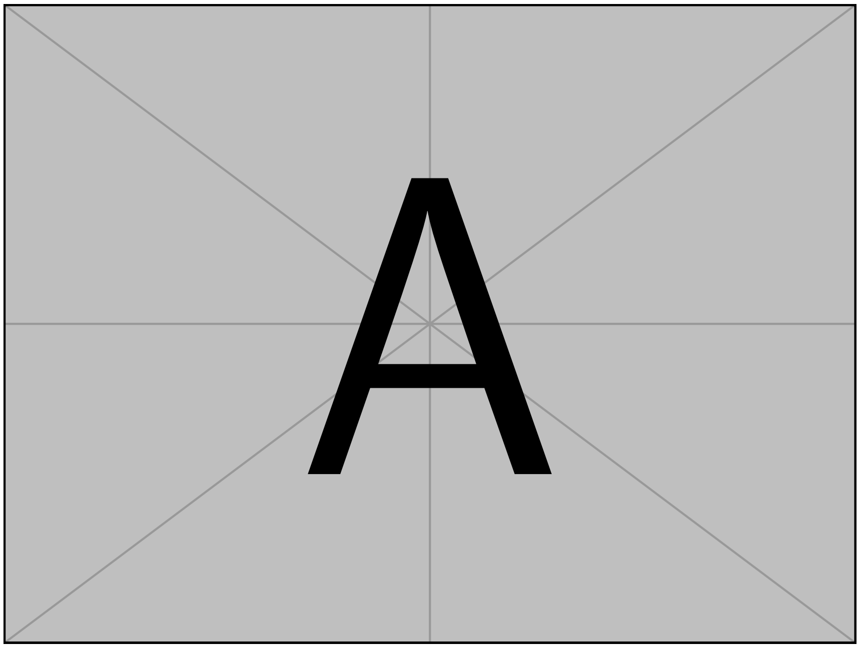

Figure 2: Landing/Login Page Example

Suppose that $(x_n)_n$ converges to x. We show that this x is also the $\mathcal{F}$-limit of $(x_n)_n$.

Proof. Take any ε. Then we know that for some large enough N, if $n \geq N$, then $x_n \in B_\varepsilon(x)$. Since every non-principal ultrafilter on $\mathbb{N}$ contains $\mathcal{F}_\infty$, then $\mathcal{F}$ also contains $\{n : n \geq N\}$, since the complement is finite. Therefore since filters are closed upwards, any sequence items x_n with $n < N$ that happen to fall in the ball around x, i.e, $x_n \in B_\varepsilon(x)$ is also contained in some filter element, so $\{n \in \mathbb{N} : |x_n - x| < \varepsilon\} \in \mathcal{F}$, as desired.

□

Proof. Take any ε. Then we know that for some large enough N, if $n \geq N$, then $x_n \in B_\varepsilon(x)$. Since every non-principal ultrafilter on $\mathbb{N}$ contains $\mathcal{F}_\infty$, then $\mathcal{F}$ also contains $\{n : n \geq N\}$, since the complement is finite. Therefore since filters are closed upwards, any sequence items x_n with $n < N$ that happen to fall in the ball around x, i.e, $x_n \in B_\varepsilon(x)$ is also contained in some filter element, so $\{n \in \mathbb{N} : |x_n - x| < \varepsilon\} \in \mathcal{F}$, as desired. □

$$P(X) = \int P(X \mid z; \theta) P(z) dz.$$

$$P(X) = \int xyz dx.$$

1 A

Definition 1.1 (Using *)**.** A distribution on matrices $\boldsymbol{S} \in \mathbb{R}^{k \times n}$ has the $(\varepsilon, \delta, \ell)$-JL moment property if for all $x \in \mathbb{R}^n$ with $|x|_2 = 1$,

$$\mathop{\mathbf{E}}_{\boldsymbol{S}} \left| |\boldsymbol{S}x|_2^2 - 1 \right|^\ell \leq \varepsilon^\ell * \delta.$$

Fan Pu Zeng
fzeng@andrew.cmu.edu

Definition 1.2 (Using \cdot)**.** A distribution on matrices $\boldsymbol{S} \in \mathbb{R}^{k\times n}$ has the $(\varepsilon, \delta, \ell)$-JL moment property if for all $x \in \mathbb{R}^n$ with $|x|_2 = 1$,

$$\mathop{\mathbf{E}}_{\boldsymbol{S}} \left| |\boldsymbol{S}x|_2^2 - 1 \right|^{\ell} \leq \varepsilon^{\ell} \cdot \delta.$$

$\mathbf{Pr}\left[X \geq a\right] \leq \frac{\mathbf{E}[X]}{a}$

$x \,\mathbf{Pr}_X \left[X \geq a\right] \leq \frac{\mathbf{E}[X]}{a}$

$x\mathbf{Pr}_X \left[X \geq a\right] \leq \frac{\mathbf{E}[X]}{a}$

$\mathbf{E}_{X\sim} \left[X \geq a\right] \leq \frac{\mathbf{E}[X]}{a}$

$p(\boldsymbol{z}, \boldsymbol{x}) = p(\boldsymbol{z})p(\boldsymbol{x} \mid \boldsymbol{z})$

$p(\boldsymbol{z}, \boldsymbol{x}) = p(\boldsymbol{z})p(\boldsymbol{x}|\boldsymbol{z})$

$KL\left(P \parallel Q\right)$

$KL\left(P \parallel Q\right)$

We evaluated the policy periodically during training by testing it without exploration noise. Figure 2 shows the performance curve for a selection of environments.

$$\sum_{i=1}^{n}$$

$$\begin{aligned}
\nabla_\mu(\mathbb{E}_{x\sim q_\mu} f(x)) &= \nabla_\mu \int_x f(x) q_\mu(x) dx \\
&= \int_x f(x)(\nabla_\mu \log q_\mu(x)) q_\mu(x) dx \\
&= \mathbb{E}_{x\sim q_\mu} \left(f(x) \nabla_\mu \log q_\mu(x)\right)
\end{aligned}$$

$$\begin{aligned}
\nabla_\mu(\mathbb{E}_{x\sim q_\mu} f(x)) =& \nabla_\mu \int_x f(x) q_\mu(x) dx \\
=& \int_x f(x)(\nabla_\mu \log q_\mu(x)) q_\mu(x) dx \\
=& \mathbb{E}_{x\sim q_\mu} \left(f(x) \nabla_\mu \log q_\mu(x)\right)
\end{aligned}$$

$\mathbf{Pr}\left[X \geq a\right] \leq \frac{\mathbf{E}[X]}{a}$

$$\phi, \Phi, \varphi$$

,

Theorem 1.3 (Gibbs Variational Principle)**.** *Let $p(\mathbf{z}, \mathbf{x})$ be a joint distribution over latent variables and observables. Then,*

$$p(\mathbf{z} \mid \mathbf{x}) = \mathop{\arg\max}_{q(\mathbf{z}|\mathbf{x}):\ \textit{distribution over } \mathbf{z}}$$

Fan Pu Zeng
fzeng@andrew.cmu.edu

$$\min_x |\boldsymbol{A}x - b|_2^2$$

$$\boldsymbol{M}$$

$$|\langle \boldsymbol{u}, \boldsymbol{v} \rangle|^2 \leq \langle \boldsymbol{u}, \boldsymbol{u} \rangle \cdot \langle \boldsymbol{v}, \boldsymbol{v} \rangle$$