بسم الله الرحمن الرحیم

یادگیری بندیت

جلسه ۷ و ۸:

بندیت زمینه‌ای — فصل ۱۸ و ۱۹

ترم بهار ۱۳۹۹ - ۱۴۰۰

مرور بندیت دشمنانه

مسئله بندیت دشمنانه

$$R_n(\pi, x) = \max_{i \in [k]} \sum_{t=1}^{n} x_{ti} - \mathbb{E}\left[\sum_{t=1}^{n} x_{tA_t}\right]$$

$$R_n^*(\pi) = \sup_{x \in [0,1]^{n \times k}} R_n(\pi, x).$$

الگوریتم EXP3

Input: n, k, η

Set $\hat{S}_{0i} = 0$ for all i

for $t = 1, \ldots, n$ **do**

Calculate the sampling distribution P_t:

$$P_{ti} = \frac{\exp\left(\eta \hat{S}_{t-1,i}\right)}{\sum_{j=1}^{k} \exp\left(\eta \hat{S}_{t-1,j}\right)}$$

Sample $A_t \sim P_t$ and observe reward X_t

Calculate $\hat{S}_{ti}$:

$$\hat{S}_{ti} = \hat{S}_{t-1,i} + \underbrace{1 - \frac{\mathbb{I}\left\{A_t = i\right\}(1 - X_t)}{P_{ti}}}_{\widehat{X}_{t,i}}$$

end for

تحلیل الگوریتم EXP3

سود دسته i

سود ما

$$\hat{S}_{ni} - \hat{S}_n \leq \frac{\log(k)}{\eta} + \frac{\eta}{2}\sum_{t=1}^{n}\sum_{j=1}^{k} P_{tj}(\hat{X}_{tj} - 1)^2$$

تحلیل الگوریتم EXP3

سود دسته i

سود ما

$$\hat{S}_{ni} - \hat{S}_n \leq \frac{\log(k)}{\eta} + \frac{\eta}{2} \sum_{t=1}^{n} \sum_{j=1}^{k} P_{tj} (\hat{X}_{tj} - 1)^2$$

$$R_n(\pi, x) \leq 2\sqrt{nk \log(k)}.$$

بندیت زمینه‌ای

تعریف بندیت زمینه‌ای

دشمنانه

تعریف بندیت زمینه‌ای

- اطلاعات زمینه
  - مثال: پیشنهاد فیلم

تعریف بندیت زمینه‌ای

For rounds $t = 1, 2, \ldots, n$:

Learner observes context $c_t \in \mathcal{C}$ where $\mathcal{C}$ is an arbitrary fixed set of contexts.

Learner selects distribution $P_t \in \mathcal{P}_{k-1}$ and samples A_t from P_t.

Learner observes reward $X_t = x_{tA_t}$.

تعریف بندیت زمینه‌ای

Adversary secretly chooses rewards $(x_t)_{t=1}^n$ with $x_t \in [0,1]^k$

Adversary secretly chooses contexts $(c_t)_{t=1}^n$ with $c_t \in \mathcal{C}$

دشمنانه

For rounds $t = 1, 2, \ldots, n$:

Learner observes context $c_t \in \mathcal{C}$ where $\mathcal{C}$ is an arbitrary fixed set of contexts.

Learner selects distribution $P_t \in \mathcal{P}_{k-1}$ and samples A_t from P_t.

Learner observes reward $X_t = x_{tA_t}$.

تعریف بندیت زمینه‌ای

- بدون زمینه:

$$R_n(\pi, x) \leq 2\sqrt{nk\log(k)}.$$

تعریف بندیت زمینه‌ای

- بدون زمینه:

$$R_n(\pi, x) \leq 2\sqrt{nk\log(k)}.$$

زمینه‌ای چه فایده‌ای دارد؟

$$R_n(\pi, x, c) := \mathbb{E}\left[\sum_{c \in \mathcal{C}} \max_{i \in [k]} \sum_{t \in [n] : c_t = c} (x_{ti} - X_t)\right]$$

$$R_n^*(\pi) = \sup_{x \in [0,1]^{n \times k}, c \in \mathbb{C}} R_n(\pi, x, c)$$

روش ۱: یک EXP3 برای هر زمینه

هر زمینه،
$|C|/n$ بار

روش ۱: یک EXP3 برای هر زمینه

$$R_{nc} \leq 2\sqrt{k \sum_{t=1}^{n} \mathbb{I}\{c_t = c\} \log(k)},$$

هر زمینه،
$n/|C|$ بار

روش ۱: یک EXP3 برای هر زمینه

$$R_{nc} \leq 2\sqrt{k\sum_{t=1}^{n}\mathbb{I}\left\{c_t = c\right\}\log(k)},$$

$$R_n = \sum_{c\in\mathcal{C}} R_{nc} \leq 2\sum_{c\in\mathcal{C}}\sqrt{k\log(k)\sum_{t=1}^{n}\mathbb{I}\left\{c_t = c\right\}}.$$

هر زمینه،
$n/|C|$ بار

روش ۱: یک EXP3 برای هر زمینه

$$R_{nc} \leq 2\sqrt{k \sum_{t=1}^{n} \mathbb{I}\{c_t = c\} \log(k)},$$

$$R_n = \sum_{c \in \mathcal{C}} R_{nc} \leq 2 \sum_{c \in \mathcal{C}} \sqrt{k \log(k) \sum_{t=1}^{n} \mathbb{I}\{c_t = c\}}.$$

$$R_n \leq 2\sqrt{nk|\mathcal{C}| \log(k)}.$$

هر زمینه،

$n/|C|$ بار

روش ۱: یک EXP3 برای هر زمینه

$$R_{nc} \leq 2\sqrt{k \sum_{t=1}^{n} \mathbb{I}\{c_t = c\} \log(k)},$$

$$R_n = \sum_{c \in \mathcal{C}} R_{nc} \leq 2 \sum_{c \in \mathcal{C}} \sqrt{k \log(k) \sum_{t=1}^{n} \mathbb{I}\{c_t = c\}}.$$

$$R_n \leq 2\sqrt{nk|\mathcal{C}| \log(k)}.$$

هر زمینه،
$n/|C|$ بار

بندیت زمینه‌ای

توابع محدودتر

کمی واقعی‌تر!

$$R_n = \mathbb{E}\left[\sum_{c \in \mathcal{C}} \max_{i \in [k]} \sum_{t \in [n]: c_t = c} (x_{ti} - X_t)\right]$$

کمی واقعی‌تر!

$$R_n = \mathbb{E}\left[\sum_{c\in\mathcal{C}} \max_{i\in[k]} \sum_{t\in[n]:c_t=c} (x_{ti} - X_t)\right]$$

$$\max_{\phi:\mathbb{C}\to[k]} \sum_{c\in\mathbb{C}} \sum_{t:c_t=c} x_{t,\phi(c_t)} - X_t$$

کمی واقعی‌تر!

$$R_n = \mathbb{E}\left[\sum_{c\in\mathcal{C}} \max_{i\in[k]} \sum_{t\in[n]:c_t=c} (x_{ti} - X_t)\right]$$

$$\max_{\phi:\mathbb{C}\to[k]} \sum_{c\in\mathbb{C}} \sum_{t:c_t=c} x_{t,\phi(c_t)} - X_t$$

$$R_n = \mathbb{E}\left[\max_{\phi\in\Phi} \sum_{t=1}^{n} (x_{t\phi(c_t)} - X_t)\right]$$

مثال Φ

مثال Φ

- ۰) همه توابع

مثال Φ

- ۰) همه توابع
- ۱) افراز k_ بخشی روی C

$$\frac{1}{|\mathcal{C}|^2}\sum_{c,d\in\mathcal{C}}(1-s(c,d))\mathbb{I}\left\{\phi(c)\neq\phi(d)\right\}$$

مثال Φ

- ۰) همه توابع
- ۱) افراز k- بخشی روی C
- ۲) توابع با تغییرات کوچک

$$\frac{1}{|\mathcal{C}|^2} \sum_{c,d \in \mathcal{C}} (1 - s(c,d)) \mathbb{I}\{\phi(c) \neq \phi(d)\}$$

مثال Φ

- ۰) همه توابع
- ۱) افراز k_ بخشی روی C
- ۲) توابع با تغییرات کوچک

$$\frac{1}{|\mathcal{C}|^2}\sum_{c,d\in\mathcal{C}}(1-s(c,d))\mathbb{I}\{\phi(c)\neq\phi(d)\}$$

- ۳) توابع نامزد

$$\phi_1,\ldots,\phi_M:\mathcal{C}\to[k].$$

بندیت زمینه‌ای

راهنمایی متخصصین

$$\phi_1, \ldots, \phi_M : \mathcal{C} \to \mathbb{P}_k$$

$$R_n = \mathbb{E}\left[\max_{\phi \in \Phi} \sum_t \left(\sum_{i=1}^{k} \phi(c_t)_i \cdot x_{t,i} - X_t \right)\right]$$

$$R_n = \mathbb{E}\left[\max_{\phi\in\Phi}\sum_{t=1}^{n}(x_{t\phi(c_t)} - X_t)\right]$$

$$\phi_1, \ldots, \phi_M : \mathcal{C} \to \mathbb{P}_k$$

$$R_n = \mathbb{E}\left[\max_{\phi\in\Phi}\sum_{t}\left(\sum_{i=1}^{k}\phi(c_t)_i \cdot x_{t,i} - X_t\right)\right]$$

$$\phi_1, \ldots, \phi_M : \mathcal{C} \to [k].$$

$$R_n = \mathbb{E}\left[\max_{\phi \in \Phi} \sum_{t=1}^{n} (x_{t\phi(c_t)} - X_t)\right]$$

$$\phi_1, \ldots, \phi_M : \mathcal{C} \to \mathbb{P}_k$$

$$R_n = \mathbb{E}\left[\max_{\phi \in \Phi} \sum_{t} \left(\sum_{i=1}^{k} \phi(c_t)_i \cdot x_{t,i} - X_t \right)\right]$$

صورت مسئله «بندیت با راهنمایی متخصصین»

Adversary secretly chooses rewards $x \in [0,1]^{n \times k}$

Experts secretly choose predictions $E^{(1)}, \dots, E^{(n)}$

For rounds $t = 1, 2, \dots, n$:

Learner observes predictions of all experts, $E^{(t)} \in [0,1]^{M \times k}$.

Learner selects a distribution $P_t \in \mathcal{P}_{k-1}$.

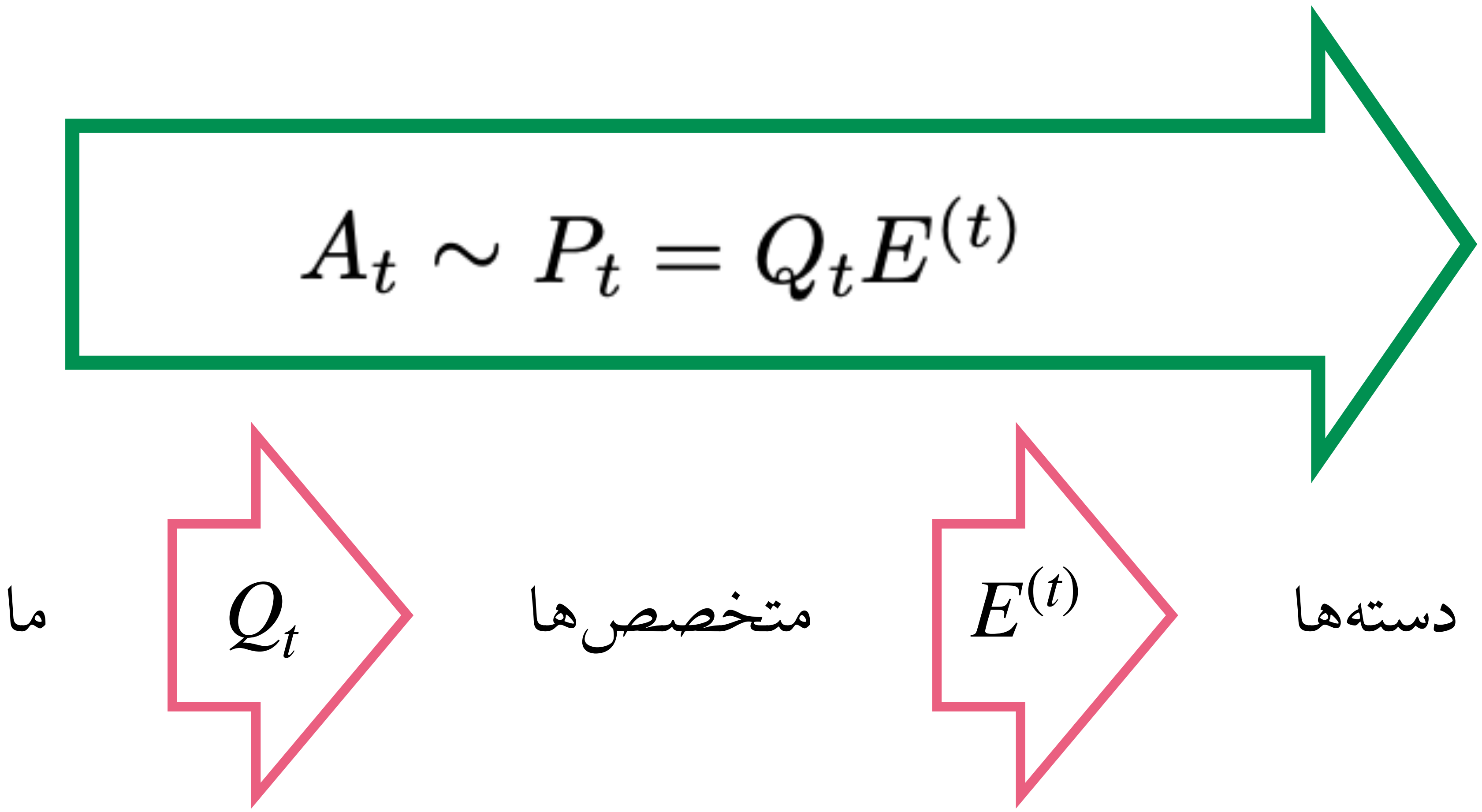
Action A_t is sampled from P_t and the reward is $X_t = x_{tA_t}$.

$$R_n = \mathbb{E}\left[\max_{m \in [M]} \sum_{t=1}^{n} E_m^{(t)} x_t - \sum_{t=1}^{n} X_t\right]$$

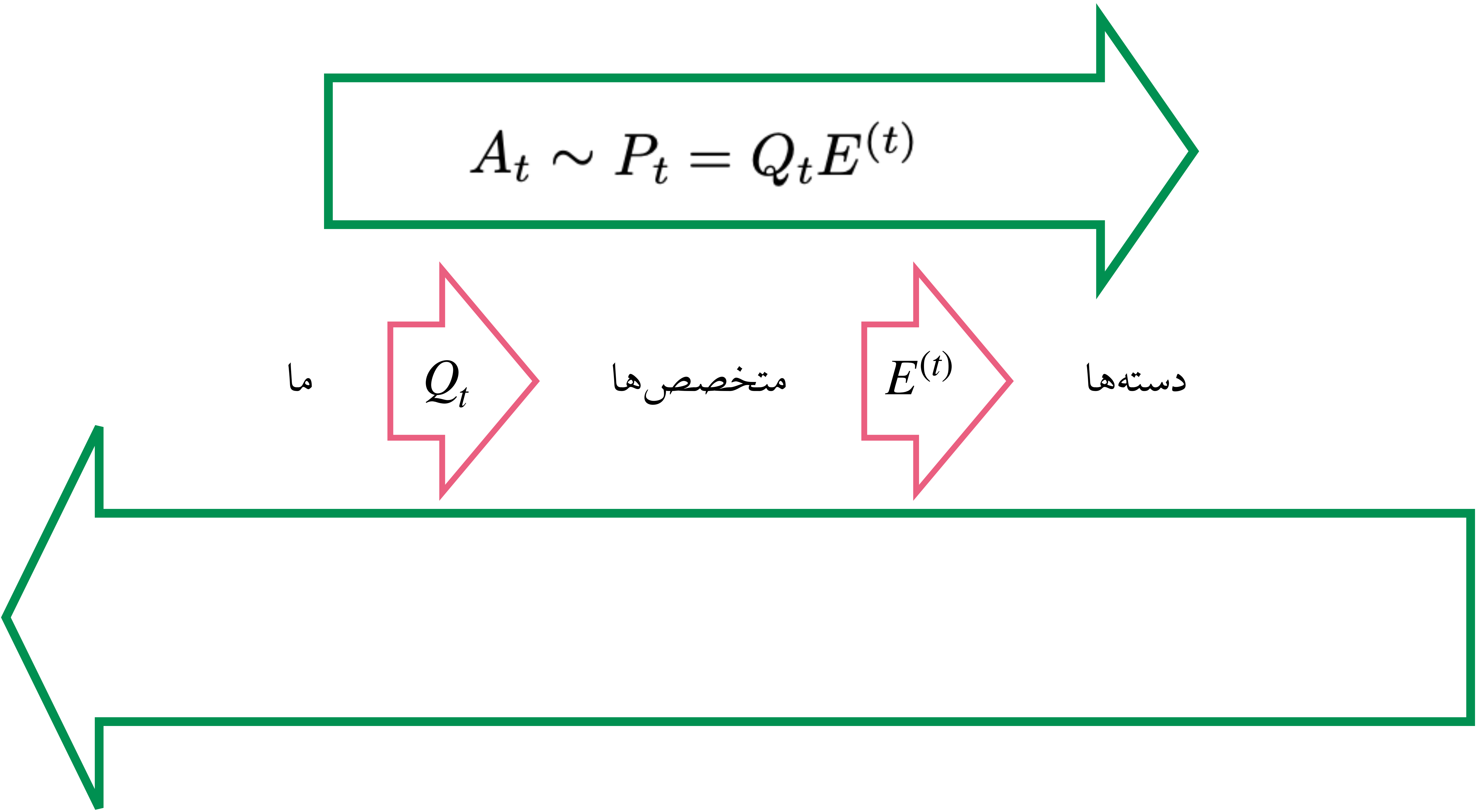
هدف

ایده

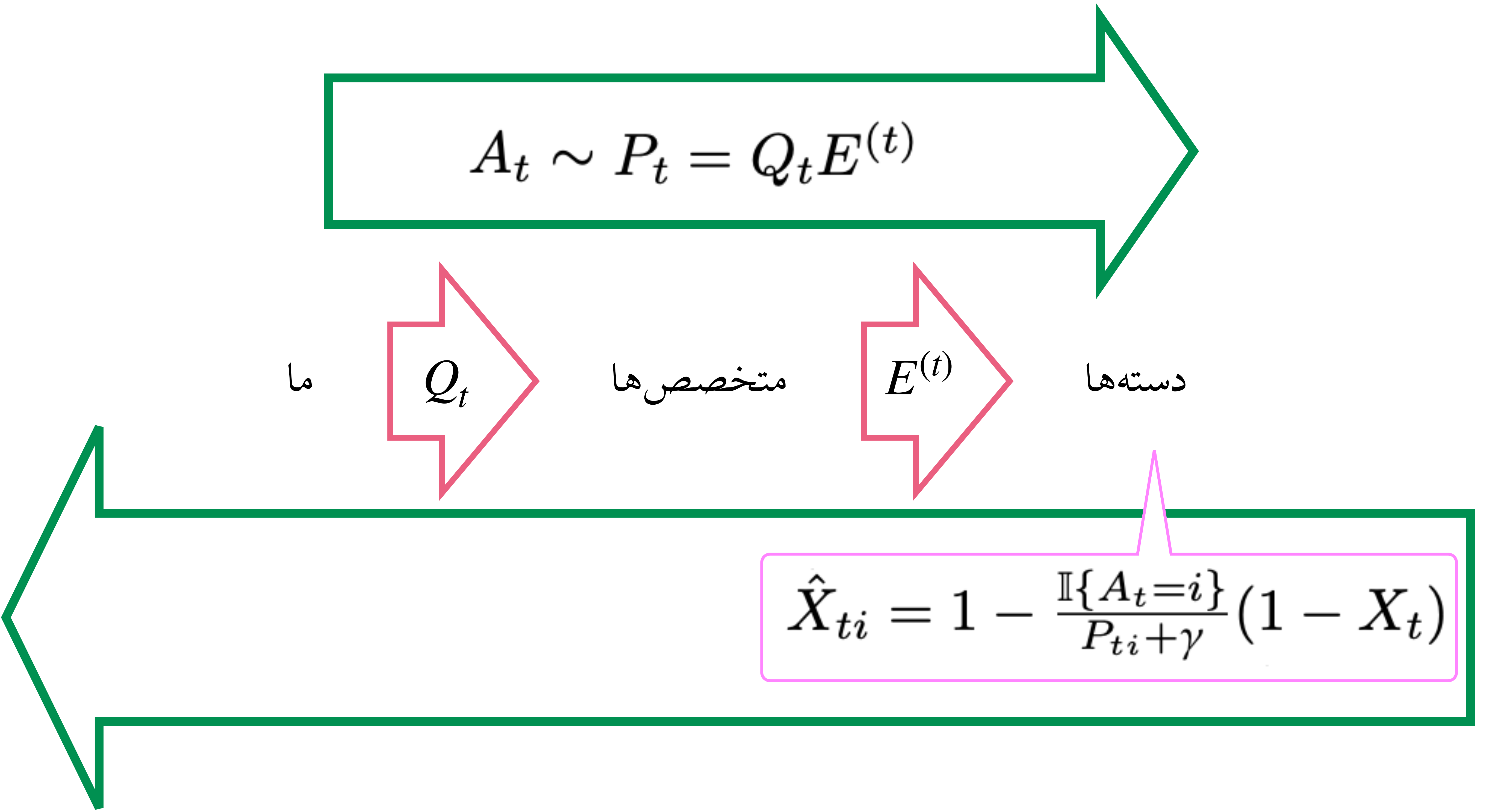
ایده



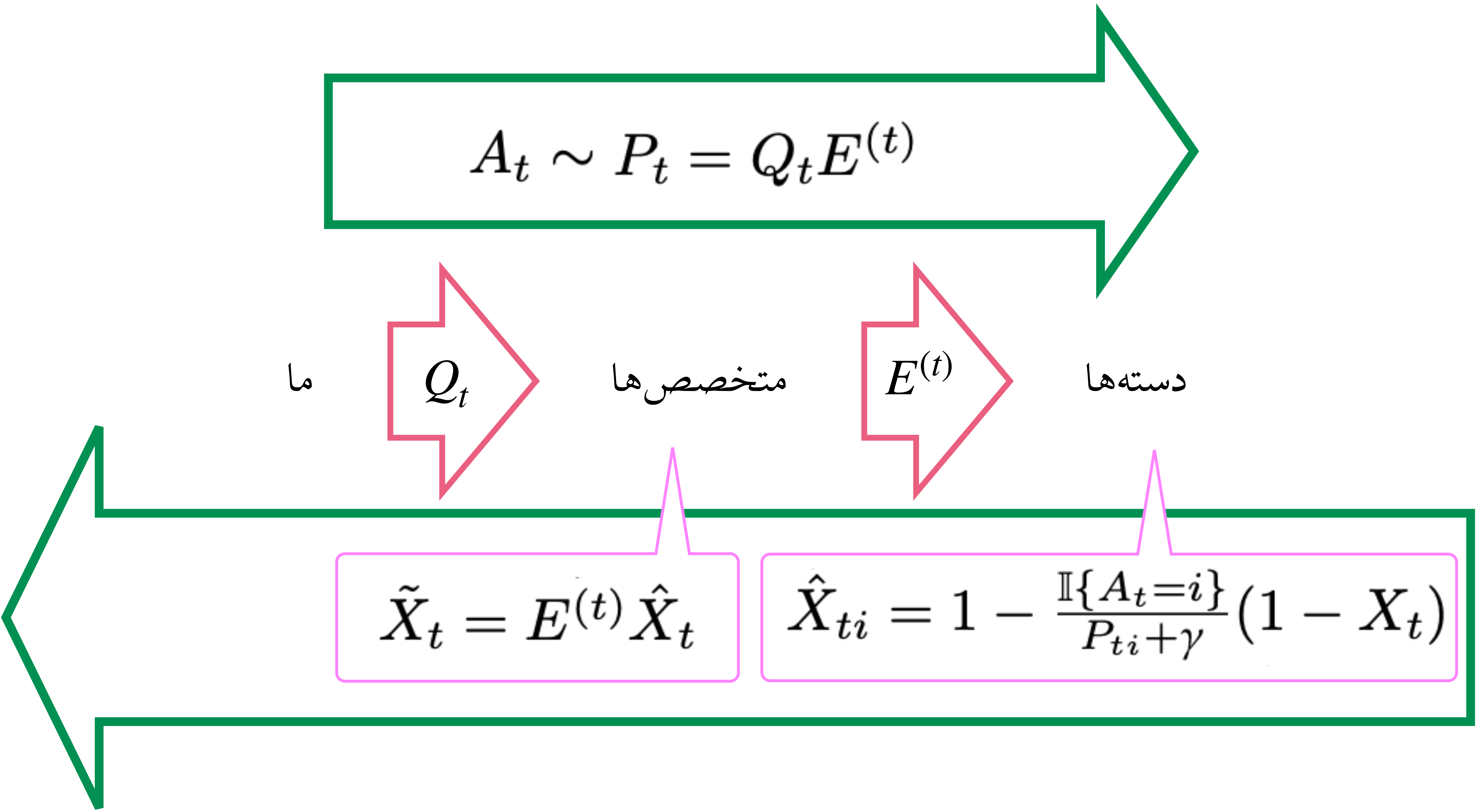
ایده



ایده


$A_t \sim P_t = Q_t E^{(t)}$
ما
Q_t
متخصص‌ها
$E^{(t)}$
دسته‌ها
$\hat{X}_{ti} = 1 - \frac{\mathbb{I}\{A_t = i\}}{P_{ti} + \gamma}(1 - X_t)$

ایده



$A_t \sim P_t = Q_t E^{(t)}$
ما
Q_t
متخصص‌ها
$E^{(t)}$
دسته‌ها
$\tilde{X}_t = E^{(t)} \hat{X}_t$
$\hat{X}_{ti} = 1 - \frac{\mathbb{I}\{A_t = i\}}{P_{ti} + \gamma}(1 - X_t)$

ایده

$$A_t \sim P_t = Q_t E^{(t)}$$

ما $\xrightarrow{Q_t}$ متخصص‌ها $\xrightarrow{E^{(t)}}$ دسته‌ها

$$\exp\left(\eta \widetilde{X}_{t,m}\right)$$

$$\tilde{X}_t = E^{(t)} \hat{X}_t$$

$$\hat{X}_{ti} = 1 - \frac{\mathbb{I}\{A_t = i\}}{P_{ti} + \gamma}(1 - X_t)$$

Lemma 18.2. *For any $m^* \in [M]$, it holds that*

$$\sum_{t=1}^{n} \tilde{X}_{tm^*} - \sum_{t=1}^{n} \sum_{m=1}^{M} Q_{tm} \tilde{X}_{tm} \leq \frac{\log(M)}{\eta} + \frac{\eta}{2} \sum_{t=1}^{n} \sum_{m=1}^{M} Q_{tm} (1 - \tilde{X}_{tm})^2.$$

Lemma 18.2. *For any $m^* \in [M]$, it holds that*

$$\sum_{t=1}^{n} \tilde{X}_{tm^*} - \sum_{t=1}^{n}\sum_{m=1}^{M} Q_{tm}\tilde{X}_{tm} \leq \frac{\log(M)}{\eta} + \frac{\eta}{2}\sum_{t=1}^{n}\sum_{m=1}^{M} Q_{tm}(1-\tilde{X}_{tm})^2.$$

$$R_n \leq \frac{\log(M)}{\eta} + \frac{\eta}{2}\sum_{t=1}^{n}\sum_{m=1}^{M} \mathbb{E}\left[Q_{tm}(1-\tilde{X}_{tm})^2\right].$$

ما $\xrightarrow{Q_t}$ متخصص‌ها $\xrightarrow{E^{(t)}}$ دسته‌ها

$$\exp\left(\eta \widetilde{X}_{t,m}\right)$$

$$\tilde{X}_t = E^{(t)}\hat{X}_t$$

$$\hat{X}_{ti} = 1 - \frac{\mathbb{I}\{A_t = i\}}{P_{ti} + \gamma}(1 - X_t)$$

$$\mathbb{E}\left[\sum_{m=1}^{M} Q_{tm}(1 - \tilde{X}_{tm})^2\right]$$

$y_{ti} = 1 - x_{ti}$ دسته‌ها $E^{(t)}$ متخصص‌ها Q_t ما

$$\hat{X}_{ti} = 1 - \frac{\mathbb{I}\{A_t = i\}}{P_{ti} + \gamma}(1 - X_t)$$

$$\tilde{X}_t = E^{(t)} \hat{X}_t$$

$$\exp\left(\eta \widetilde{X}_{t,m}\right)$$

$$\mathbb{E}\left[\sum_{m=1}^{M} Q_{tm}(1 - \tilde{X}_{tm})^2\right]$$

ما $\xrightarrow{Q_t}$ متخصص‌ها $\xrightarrow{E^{(t)}}$ دسته‌ها $\quad y_{ti} = 1 - x_{ti}$

$$\exp\left(\eta \widetilde{X}_{t,m}\right)$$

$$\tilde{X}_t = E^{(t)} \hat{X}_t$$

$$\hat{X}_{ti} = 1 - \frac{\mathbb{I}\{A_t = i\}}{P_{ti} + \gamma}(1 - X_t)$$

$$\hat{Y}_{ti} \doteq 1 - \hat{X}_{ti}$$

$$\hat{Y}_{ti} = \frac{A_{ti} y_{ti}}{P_{ti}}$$

$$\mathbb{E}\left[\sum_{m=1}^{M} Q_{tm}(1 - \tilde{X}_{tm})^2\right]$$

ما $\xleftarrow{Q_t}$ متخصص‌ها $\xleftarrow{E^{(t)}}$ دسته‌ها $\quad y_{ti} = 1 - x_{ti}$

$$\exp\left(\eta \widetilde{X}_{t,m}\right)$$

$$\tilde{X}_t = E^{(t)} \hat{X}_t$$

$$\hat{X}_{ti} = 1 - \frac{\mathbb{I}\{A_t = i\}}{P_{ti} + \gamma}(1 - X_t)$$

$$E^{(t)}(1 - \widehat{X}_t) = 1 - E^{(t)} \widehat{X}_t = 1 - \widetilde{X}_t$$

$$\tilde{Y}_{tm} \doteq 1 - \tilde{X}_{tm} \qquad \hat{Y}_{ti} \doteq 1 - \hat{X}_{ti}$$

$$\tilde{Y}_t = E^{(t)} \hat{Y}_t \qquad \hat{Y}_{ti} = \frac{A_{ti} y_{ti}}{P_{ti}}$$

$$\mathbb{E}\left[\sum_{m=1}^{M} Q_{tm}(1 - \tilde{X}_{tm})^2\right]$$

ما $\xrightarrow{Q_t}$ متخصص‌ها $\xrightarrow{E^{(t)}}$ دسته‌ها $\quad y_{ti} = 1 - x_{ti}$

$$\exp\left(\eta \widetilde{X}_{t,m}\right) \qquad \tilde{X}_t = E^{(t)} \hat{X}_t \qquad \hat{X}_{ti} = 1 - \frac{\mathbb{I}\{A_t = i\}}{P_{ti} + \gamma}(1 - X_t)$$

$$\begin{aligned} E^{(t)}(1 - \widehat{X}_t) \\ = 1 - E^{(t)} \widehat{X}_t \\ = 1 - \widetilde{X}_t \end{aligned}$$

$$\tilde{Y}_{tm} \doteq 1 - \tilde{X}_{tm} \qquad \hat{Y}_{ti} \doteq 1 - \hat{X}_{ti}$$

$$\tilde{Y}_t = E^{(t)} \hat{Y}_t \qquad \hat{Y}_{ti} = \frac{A_{ti} y_{ti}}{P_{ti}}$$

$$\mathbb{E}\left[\sum_{m=1}^{M} Q_{tm} (1 - \tilde{X}_{tm})^2\right]$$

$y_{ti} = 1 - x_{ti}$

دسته‌ها $\xrightarrow{E^{(t)}}$ متخصص‌ها $\xrightarrow{Q_t}$ ما

$$\exp\left(\eta \widetilde{X}_{t,m}\right)$$

$$\tilde{X}_t = E^{(t)} \hat{X}_t$$

$$\hat{X}_{ti} = 1 - \frac{\mathbb{I}\{A_t = i\}}{P_{ti} + \gamma}(1 - X_t)$$

$$E^{(t)}(1 - \widehat{X}_t) = 1 - E^{(t)} \widehat{X}_t = 1 - \widetilde{X}_t$$

$$\tilde{Y}_{tm} \doteq 1 - \tilde{X}_{tm}$$

$$\hat{Y}_{ti} \doteq 1 - \hat{X}_{ti}$$

$$\tilde{Y}_t = E^{(t)} \hat{Y}_t$$

$$\hat{Y}_{ti} = \frac{A_{ti} y_{ti}}{P_{ti}}$$

$$\mathbb{E}_t[\tilde{Y}_{tm}^2] = \mathbb{E}_t\left[\left(\frac{E^{(t)}_{mA_t} y_{tA_t}}{P_{tA_t}}\right)^2\right]$$

$$\mathbb{E}\left[\sum_{m=1}^{M} Q_{tm}(1 - \tilde{X}_{tm})^2\right]$$

ما $\Rightarrow Q_t \Rightarrow$ متخصص‌ها $\Rightarrow E^{(t)} \Rightarrow$ دسته‌ها $\quad y_{ti} = 1 - x_{ti}$

$$\exp\left(\eta \widetilde{X}_{t,m}\right) \qquad \tilde{X}_t = E^{(t)} \hat{X}_t \qquad \hat{X}_{ti} = 1 - \frac{\mathbb{I}\{A_t = i\}}{P_{ti} + \gamma}(1 - X_t)$$

$$E^{(t)}(1 - \widehat{X}_t) = 1 - E^{(t)} \widehat{X}_t = 1 - \widetilde{X}_t$$

$$\tilde{Y}_{tm} \doteq 1 - \tilde{X}_{tm} \qquad \hat{Y}_{ti} \doteq 1 - \hat{X}_{ti}$$

$$\tilde{Y}_t = E^{(t)} \hat{Y}_t \qquad \hat{Y}_{ti} = \frac{A_{ti} y_{ti}}{P_{ti}}$$

$$\mathbb{E}_t[\tilde{Y}_{tm}^2] = \mathbb{E}_t\left[\left(\frac{E_{mA_t}^{(t)} y_{tA_t}}{P_{tA_t}}\right)^2\right] = \sum_{i=1}^{k} \frac{\left(E_{mi}^{(t)} y_{ti}\right)^2}{P_{ti}}$$

$$\mathbb{E}\left[\sum_{m=1}^{M} Q_{tm}(1 - \tilde{X}_{tm})^2\right]$$

ما $\xrightarrow{Q_t}$ متخصص‌ها $\xrightarrow{E^{(t)}}$ دسته‌ها $\quad y_{ti} = 1 - x_{ti}$

$$\exp\left(\eta \widetilde{X}_{t,m}\right)$$

$$\tilde{X}_t = E^{(t)} \hat{X}_t$$

$$\hat{X}_{ti} = 1 - \frac{\mathbb{I}\{A_t = i\}}{P_{ti} + \gamma}(1 - X_t)$$

$$E^{(t)}(1 - \widehat{X}_t) = 1 - E^{(t)} \widehat{X}_t = 1 - \widetilde{X}_t$$

$$\tilde{Y}_{tm} \doteq 1 - \tilde{X}_{tm} \qquad \hat{Y}_{ti} \doteq 1 - \hat{X}_{ti}$$

$$\tilde{Y}_t = E^{(t)} \hat{Y}_t \qquad \hat{Y}_{ti} = \frac{A_{ti} y_{ti}}{P_{ti}}$$

$$\mathbb{E}_t[\tilde{Y}_{tm}^2] = \mathbb{E}_t\left[\left(\frac{E^{(t)}_{mA_t} y_{tA_t}}{P_{tA_t}}\right)^2\right] = \sum_{i=1}^{k} \frac{\left(E^{(t)}_{mi} y_{ti}\right)^2}{P_{ti}} \leq \sum_{i=1}^{k} \frac{E^{(t)}_{mi}}{P_{ti}}.$$

$$\mathbb{E}\left[\sum_{m=1}^{M} Q_{tm}(1 - \tilde{X}_{tm})^2\right]$$

ما $\xrightarrow{Q_t}$ متخصص‌ها $\xrightarrow{E^{(t)}}$ دسته‌ها $\quad y_{ti} = 1 - x_{ti}$

$$\exp\left(\eta \widetilde{X}_{t,m}\right) \qquad \tilde{X}_t = E^{(t)} \hat{X}_t \qquad \hat{X}_{ti} = 1 - \frac{\mathbb{I}\{A_t = i\}}{P_{ti} + \gamma}(1 - X_t)$$

$$E^{(t)}(1 - \widehat{X}_t) = 1 - E^{(t)} \widehat{X}_t = 1 - \widetilde{X}_t$$

$$\tilde{Y}_{tm} \doteq 1 - \tilde{X}_{tm} \qquad \hat{Y}_{ti} \doteq 1 - \hat{X}_{ti}$$

$$\tilde{Y}_t = E^{(t)} \hat{Y}_t \qquad \hat{Y}_{ti} = \frac{A_{ti} y_{ti}}{P_{ti}}$$

$$\mathbb{E}_t[\tilde{Y}_{tm}^2] = \mathbb{E}_t\left[\left(\frac{E^{(t)}_{mA_t} y_{tA_t}}{P_{tA_t}}\right)^2\right] = \sum_{i=1}^{k} \frac{\left(E^{(t)}_{mi} y_{ti}\right)^2}{P_{ti}} \leq \sum_{i=1}^{k} \frac{E^{(t)}_{mi}}{P_{ti}}.$$

$$\mathbb{E}\left[\sum_{m=1}^{M} Q_{tm}(1 - \tilde{X}_{tm})^2\right] \leq \mathbb{E}\left[\sum_{m=1}^{M} Q_{tm} \sum_{i=1}^{k} \frac{E^{(t)}_{mi}}{P_{ti}}\right]$$

ما $\xrightarrow{Q_t}$ متخصص‌ها $\xrightarrow{E^{(t)}}$ دسته‌ها $\quad y_{ti} = 1 - x_{ti}$

$$\exp\left(\eta \widetilde{X}_{t,m}\right) \qquad \tilde{X}_t = E^{(t)} \hat{X}_t \qquad \hat{X}_{ti} = 1 - \frac{\mathbb{I}\{A_t = i\}}{P_{ti} + \gamma}(1 - X_t)$$

$$\begin{aligned} &E^{(t)}(1 - \widehat{X}_t) \\ &= 1 - E^{(t)} \widehat{X}_t \\ &= 1 - \widetilde{X}_t \end{aligned}$$

$$\tilde{Y}_{tm} \doteq 1 - \tilde{X}_{tm} \qquad \hat{Y}_{ti} \doteq 1 - \hat{X}_{ti}$$

$$\tilde{Y}_t = E^{(t)} \hat{Y}_t \qquad \hat{Y}_{ti} = \frac{A_{ti} y_{ti}}{P_{ti}}$$

$$\mathbb{E}_t[\tilde{Y}_{tm}^2] = \mathbb{E}_t\left[\left(\frac{E_{mA_t}^{(t)} y_{tA_t}}{P_{tA_t}}\right)^2\right] = \sum_{i=1}^{k} \frac{\left(E_{mi}^{(t)} y_{ti}\right)^2}{P_{ti}} \leq \sum_{i=1}^{k} \frac{E_{mi}^{(t)}}{P_{ti}}.$$

$$\mathbb{E}\left[\sum_{m=1}^{M} Q_{tm}(1 - \tilde{X}_{tm})^2\right] \leq \mathbb{E}\left[\sum_{m=1}^{M} Q_{tm} \sum_{i=1}^{k} \frac{E_{mi}^{(t)}}{P_{ti}}\right] = \mathbb{E}\left[\sum_{i=1}^{k} \frac{\sum_{m=1}^{M} Q_{tm} E_{mi}^{(t)}}{P_{ti}}\right]$$

ما $\xrightarrow{Q_t}$ متخصص‌ها $\xrightarrow{E^{(t)}}$ دسته‌ها $\quad y_{ti} = 1 - x_{ti}$

$$\exp\left(\eta \widetilde{X}_{t,m}\right) \qquad \tilde{X}_t = E^{(t)} \hat{X}_t \qquad \hat{X}_{ti} = 1 - \frac{\mathbb{I}\{A_t = i\}}{P_{ti} + \gamma}(1 - X_t)$$

$$\begin{aligned} &E^{(t)}(1 - \widehat{X}_t) \\ &= 1 - E^{(t)} \widehat{X}_t \\ &= 1 - \widetilde{X}_t \end{aligned}$$

$$\tilde{Y}_{tm} \doteq 1 - \tilde{X}_{tm} \qquad \hat{Y}_{ti} \doteq 1 - \hat{X}_{ti}$$

$$\tilde{Y}_t = E^{(t)} \hat{Y}_t \qquad \hat{Y}_{ti} = \frac{A_{ti} y_{ti}}{P_{ti}}$$

$$\mathbb{E}_t[\tilde{Y}_{tm}^2] = \mathbb{E}_t\left[\left(\frac{E_{mA_t}^{(t)} y_{tA_t}}{P_{tA_t}}\right)^2\right] = \sum_{i=1}^{k} \frac{\left(E_{mi}^{(t)} y_{ti}\right)^2}{P_{ti}} \le \sum_{i=1}^{k} \frac{E_{mi}^{(t)}}{P_{ti}}.$$

$$\mathbb{E}\left[\sum_{m=1}^{M} Q_{tm}(1 - \tilde{X}_{tm})^2\right] \le \mathbb{E}\left[\sum_{m=1}^{M} Q_{tm} \sum_{i=1}^{k} \frac{E_{mi}^{(t)}}{P_{ti}}\right] = \mathbb{E}\left[\sum_{i=1}^{k} \frac{\sum_{m=1}^{M} Q_{tm} E_{mi}^{(t)}}{P_{ti}}\right] = k.$$

$$R_n \leq \frac{\log(M)}{\eta} + \frac{\eta}{2} \sum_{t=1}^{n} \sum_{m=1}^{M} \mathbb{E}\left[Q_{tm}(1 - \tilde{X}_{tm})^2\right].$$

$$\mathbb{E}\left[\sum_{m=1}^{M} Q_{tm}(1 - \tilde{X}_{tm})^2\right] \leq k$$

$$R_n \leq \frac{\log(M)}{\eta} + \frac{\eta}{2} \sum_{t=1}^{n} \sum_{m=1}^{M} \mathbb{E}\left[Q_{tm}(1 - \tilde{X}_{tm})^2\right].$$

$$\mathbb{E}\left[\sum_{m=1}^{M} Q_{tm}(1 - \tilde{X}_{tm})^2\right] \leq k$$

$$\leq \frac{\log(M)}{\eta} + \frac{\eta n k}{2}$$

$$R_n \leq \frac{\log(M)}{\eta} + \frac{\eta}{2} \sum_{t=1}^{n} \sum_{m=1}^{M} \mathbb{E}\left[Q_{tm}(1 - \tilde{X}_{tm})^2\right].$$

$$\mathbb{E}\left[\sum_{m=1}^{M} Q_{tm}(1 - \tilde{X}_{tm})^2\right] \leq k$$

$$\leq \frac{\log(M)}{\eta} + \frac{\eta n k}{2} = \sqrt{2nk\log(M)}.$$

$$R_n \leq \frac{\log(M)}{\eta} + \frac{\eta}{2} \sum_{t=1}^{n} \sum_{m=1}^{M} \mathbb{E}\left[Q_{tm}(1 - \tilde{X}_{tm})^2\right].$$

$$\mathbb{E}\left[\sum_{m=1}^{M} Q_{tm}(1 - \tilde{X}_{tm})^2\right] \leq k$$

$$\leq \frac{\log(M)}{\eta} + \frac{\eta n k}{2} = \sqrt{2nk\log(M)}.$$

THEOREM 18.1. *Let $\gamma = 0$ and $\eta = \sqrt{2\log(M)/(nk)}$, and denote by R_n the expected regret of Exp4 defined in Algorithm 11 after n rounds. Then,*

$$R_n \leq \sqrt{2nk\log(M)}. \tag{18.7}$$

$$R_n \leq \sqrt{2nk\log(M)}.$$

$$R_n \leq \sqrt{2nk\log(M)}.$$

- مثال: Φ = همه توابع از C به [k]

$$R_n \leq \sqrt{2nk\log(M)}.$$

- مثال: Φ = همه توابع از C به [k]

$$M = k^C$$

$$R_n \leq \sqrt{2nk \log(M)}.$$

- مثال: Φ = همه توابع از C به [k]

$$M = k^{\mathcal{C}}$$

$$R_n \leq \sqrt{2nk|\mathcal{C}| \log(k)},$$

$$R_n \leq \sqrt{2nk \log(M)}.$$

- مثال: Φ = همه توابع از C به [k]

$$M = k^{\mathcal{C}}$$

$$R_n \leq \sqrt{2nk|\mathcal{C}| \log(k)},$$

- مثال: Φ = تعدادی تابع

$$R_n \leq \sqrt{2nk\log(M)}.$$

- مثال: Φ = همه توابع از C به [k]

$$M = k^{\mathcal{C}}$$

$$R_n \leq \sqrt{2nk|\mathcal{C}|\log(k)},$$

- مثال: Φ = تعدادی تابع

$$R_n \leq \sqrt{2nk\log(|\Phi|)}.$$

$$E_t^* = \sum_{s=1}^{t} \sum_{i=1}^{k} \max_{m \in [M]} E_{mi}^{(s)}$$

THEOREM 18.3. *Assume the same conditions as in Theorem 18.1, except let* $\eta_t = \sqrt{\log(M)/E_t^*}$. *Then there exists a universal constant* $C > 0$ *such that*

$$R_n \leq C\sqrt{E_n^* \log(M)}. \tag{18.14}$$

بندیت خطی تصادفی

تعامل با محیط در بندیت تصادفی

Adversary secretly chooses rewards $(x_t)_{t=1}^n$ with $x_t \in [0,1]^k$

Adversary secretly chooses contexts $(c_t)_{t=1}^n$ with $c_t \in \mathcal{C}$

یا تصادفی

For rounds $t = 1, 2, \ldots, n$:

Learner observes context $c_t \in \mathcal{C}$ where $\mathcal{C}$ is an arbitrary fixed set of contexts.

Learner selects distribution $P_t \in \mathcal{P}_{k-1}$ and samples A_t from P_t.

Learner observes reward $X_t = x_{tA_t}$.

بندیت تصادفی زمینه‌ای

اولین
ساده‌سازی

$$r : \mathcal{C} \times [k] \rightarrow \mathbb{R}$$

تابع پاداش

$$X_t = r(C_t, A_t) + \eta_t,$$

نویز
۱ـ زیرگوسی

$$R_n = \mathbb{E}\left[\sum_{t=1}^{n} \max_{a \in [k]} r(C_t, a) - \sum_{t=1}^{n} X_t\right]$$

$$X_t = r(C_t, A_t) + \eta_t,$$

$$R_n = \mathbb{E}\left[\sum_{t=1}^{n} \max_{a \in [k]} r(C_t, a) - \sum_{t=1}^{n} X_t\right]$$

در بدترین حالت: $R_n \geq \Omega(\sqrt{nMk})$

$$X_t = r(C_t, A_t) + \eta_t,$$

$$R_n = \mathbb{E}\left[\sum_{t=1}^{n} \max_{a \in [k]} r(C_t, a) - \sum_{t=1}^{n} X_t\right]$$

😐 در بدترین حالت: $R_n \geq \Omega(\sqrt{nMk})$

$$X_t = r(C_t, A_t) + \eta_t,$$

$$R_n = \mathbb{E}\left[\sum_{t=1}^{n} \max_{a \in [k]} r(C_t, a) - \sum_{t=1}^{n} X_t\right]$$

در بدترین حالت: $R_n \geq \Omega(\sqrt{nMk})$

تابع r؟

دومین
ساده‌سازی

$$X_t = r(C_t, A_t) + \eta_t,$$

دومین
ساده‌سازی

$$X_t = r(C_t, A_t) + \eta_t,$$

$$r(c, a) = \langle \theta_*, \psi(c, a) \rangle,$$

دومین
ساده‌سازی

$$X_t = r(C_t, A_t) + \eta_t,$$

$$r(c, a) = \langle \theta_*, \psi(c, a) \rangle,$$

پارامتر
(ناشناخته)

دومین
ساده‌سازی

$$X_t = r(C_t, A_t) + \eta_t,$$

$\psi : \mathcal{C} \times [k] \to \mathbb{R}^d$

نگاشت ویژگی‌ها
(دانسته)

$$r(c, a) = \langle \theta_*, \psi(c, a) \rangle,$$

پارامتر
(ناشناخته)

بندیت خطی تصادفی

$$X_t = r(C_t, A_t) + \eta_t,$$

$$r(c, a) = \langle \theta_*, \psi(c, a) \rangle,$$

بندیت خطی تصادفی

$$X_t = r(C_t, A_t) + \eta_t,$$

$$r(c, a) = \langle \theta_*, \psi(c, a) \rangle,$$

$$X_t = \langle \theta_*, \psi(c_t, A_t) \rangle + \eta_t$$

بندیت خطی تصادفی

$$X_t = r(C_t, A_t) + \eta_t,$$

$$r(c, a) = \langle \theta_*, \psi(c, a) \rangle,$$

$$X_t = \langle \theta_*, \boxed{\psi(c_t, A_t)} \rangle + \eta_t$$

$$X_t = \langle \theta_*, A_t \rangle + \eta_t$$

بندیت خطی تصادفی

$$X_t = r(C_t, A_t) + \eta_t,$$

$$r(c, a) = \langle \theta_*, \psi(c, a) \rangle,$$

سومین
ساده‌سازی

$$X_t = \langle \theta_*, \boxed{\psi(c_t, A_t)} \rangle + \eta_t$$

$$X_t = \langle \theta_*, A_t \rangle + \eta_t$$

بندیت خطی تصادفی

$$X_t = \langle \theta_*, A_t \rangle + \eta_t$$

بندیت خطی تصادفی

$$X_t = \langle \theta_*, A_t \rangle + \eta_t$$

$$\hat{R}_n = \sum_{t=1}^{n} \max_{a \in \mathcal{A}_t} \langle \theta_*, a - A_t \rangle,$$

بندیت خطی تصادفی

$$X_t = \langle \theta_*, A_t \rangle + \eta_t$$

$$\hat{R}_n = \sum_{t=1}^{n} \max_{a \in \mathcal{A}_t} \langle \theta_*, a - A_t \rangle,$$

$$R_n = \mathbb{E}\left[\hat{R}_n\right] = \mathbb{E}\left[\sum_{t=1}^{n} \max_{a \in \mathcal{A}_t} \langle \theta_*, a \rangle - \sum_{t=1}^{n} X_t\right]$$

مثال

$$X_t = \langle \theta_*, A_t \rangle + \eta_t$$

مثال

$$X_t = \langle \theta_*, A_t \rangle + \eta_t$$

مثال

$$\mathcal{A}_t = \{e_1, \ldots, e_d\}$$

مثال

$$X_t = \langle \theta_*, A_t \rangle + \eta_t$$

مثال

$$\mathcal{A}_t = \{e_1, \ldots, e_d\}$$

بندیت تصادفی

(کلاسیک)

مثال

$$X_t = \langle \theta_*, A_t \rangle + \eta_t$$

مثال

$$\mathcal{A}_t = \{e_1, \ldots, e_d\}$$

بندیت تصادفی
(کلاسیک)

مثال

$$\mathcal{A}_t \subseteq \{0,1\}^d$$

مثال

$$X_t = \langle \theta_*, A_t \rangle + \eta_t$$

مثال

$$\mathcal{A}_t = \{e_1, \ldots, e_d\}$$

بندیت تصادفی
(کلاسیک)

مثال

$$\mathcal{A}_t \subseteq \{0,1\}^d$$

بندیت تصادفی
ترکیبیاتی!

UCB:

۰ـ میانگین تجربی برای هر دسته

۱ـ برای هر دسته یک بازه اطمینان

۲ـ خوش‌بینانه‌ترین تخمین برای هر دسته

۳ـ انتخاب دسته با بهترین خوش‌بینانه‌ترین تخمین

UCB:

۰ــ میانگین تجربی برای هر دسته

۱ــ برای هر دسته یک بازه اطمینان

۲ــ خوش‌بینانه‌ترین تخمین برای هر دسته

۳ــ انتخاب دسته با بهترین خوش‌بینانه‌ترین تخمین

$$X_t = \langle \theta_*, A_t \rangle + \eta_t$$

ایده UCB برای بندیت خطی تصادفی:

:UCB

۰ ـ میانگین تجربی برای هر دسته

۱ ـ برای هر دسته یک بازه اطمینان

۲ ـ خوش‌بینانه‌ترین تخمین برای هر دسته

۳ ـ انتخاب دسته با بهترین خوش‌بینانه‌ترین تخمین

$$X_t = \langle \theta_*, A_t \rangle + \eta_t$$

ایده UCB برای بندیت خطی تصادفی:

۰ ـ میانگین تجربی برای θ_*

اثر
متقابل
دسته‌ها

:UCB

۰ـ میانگین تجربی برای هر دسته

۱ـ برای هر دسته یک بازه اطمینان

۲ـ خوش‌بینانه‌ترین تخمین برای هر دسته

۳ـ انتخاب دسته با بهترین خوش‌بینانه‌ترین تخمین

$$X_t = \langle \theta_*, A_t \rangle + \eta_t$$

ایده UCB برای بندیت خطی تصادفی:

۰ـ میانگین تجربی برای θ_*

۱ـ یک بازه اطمینان برای θ_*

اثر متقابل دسته‌ها

UCB:

۰ـ میانگین تجربی برای هر دسته

۱ـ برای هر دسته یک بازه اطمینان

۲ـ خوش‌بینانه‌ترین تخمین برای هر دسته

۳ـ انتخاب دسته با بهترین خوش‌بینانه‌ترین تخمین

$$X_t = \langle \theta_*, A_t \rangle + \eta_t$$

ایده UCB برای بندیت خطی تصادفی:

۰ـ میانگین تجربی برای θ_*

۱ـ یک بازه اطمینان برای θ_*

۲ـ خوش‌بینانه‌ترین تخمین برای هر دسته

اثر متقابل دسته‌ها

UCB:

۰ـ میانگین تجربی برای هر دسته

۱ـ برای هر دسته یک بازه اطمینان

۲ـ خوش‌بینانه‌ترین تخمین برای هر دسته

۳ـ انتخاب دسته با بهترین خوش‌بینانه‌ترین تخمین

$$X_t = \langle \theta_*, A_t \rangle + \eta_t$$

ایده UCB برای بندیت خطی تصادفی:

۰ـ میانگین تجربی برای θ_*

۱ـ یک بازه اطمینان برای θ_*

۲ـ خوش‌بینانه‌ترین تخمین برای هر دسته

۳ـ انتخاب دسته با بهترین خوش‌بینانه‌ترین تخمین

اثر متقابل دسته‌ها

UCB:

۰ـ میانگین تجربی برای هر دسته

۱ـ برای هر دسته یک بازه اطمینان

۲ـ خوش‌بینانه‌ترین تخمین برای هر دسته

۳ـ انتخاب دسته با بهترین خوش‌بینانه‌ترین تخمین

$$X_t = \langle \theta_*, A_t \rangle + \eta_t$$

ایده UCB برای بندیت خطی تصادفی:

۰ـ میانگین تجربی برای θ_*

۱ـ یک بازه اطمینان برای θ_*

۲ـ خوش‌بینانه‌ترین تخمین برای هر دسته

۳ـ انتخاب دسته با بهترین خوش‌بینانه‌ترین تخمین

اثر متقابل دسته‌ها

ایده UCB برای بندیت خطی تصادفی

$$X_t = \langle \theta_*, A_t \rangle + \eta_t$$

- – میانگین تجربی برای θ_*

ایده UCB برای بندیت خطی تصادفی

$$X_t = \langle \theta_*, A_t \rangle + \eta_t$$

- میانگین تجربی برای θ_*

$$(A_1, X_1), (A_2, X_2), \ldots, (A_t, X_t) \Rightarrow \theta_t$$

ایده UCB برای بندیت خطی تصادفی

$$X_t = \langle \theta_*, A_t \rangle + \eta_t$$

◦ – میانگین تجربی برای θ_*

$$(A_1, X_1), (A_2, X_2), \ldots, (A_t, X_t) \Rightarrow \theta_t$$

$$\hat{\theta}_t = \mathrm{argmin}_{\theta \in \mathbb{R}^d} \left(\sum_{s=1}^{t} (X_s - \langle \theta, A_s \rangle)^2 + \lambda \|\theta\|_2^2 \right)$$

ایده UCB برای بندیت خطی تصادفی

ایده UCB برای بندیت خطی تصادفی

- – میانگین تجربی برای θ_*

ایده UCB برای بندیت خطی تصادفی

- میانگین تجربی برای θ_*

$$\hat{\theta}_t = \mathrm{argmin}_{\theta \in \mathbb{R}^d} \left(\sum_{s=1}^{t} (X_s - \langle \theta, A_s \rangle)^2 + \lambda \|\theta\|_2^2 \right)$$

-

ایده UCB برای بندیت خطی تصادفی

۰ــ میانگین تجربی برای θ_*

$$\hat{\theta}_t = \text{argmin}_{\theta \in \mathbb{R}^d} \left(\sum_{s=1}^{t} (X_s - \langle \theta, A_s \rangle)^2 + \lambda \|\theta\|_2^2 \right)$$

•

۱ــ یک بازه اطمینان برای θ_*: $\mathcal{C}_t$

ایده UCB برای بندیت خطی تصادفی

- ۰ــ میانگین تجربی برای θ_*

$$\hat{\theta}_t = \text{argmin}_{\theta \in \mathbb{R}^d} \left(\sum_{s=1}^{t} (X_s - \langle \theta, A_s \rangle)^2 + \lambda \|\theta\|_2^2 \right)$$

- ۱ــ یک بازه اطمینان برای θ_*: $\mathcal{C}_t$

$$\mathcal{C}_t \subseteq \mathcal{E}_t = \left\{ \theta \in \mathbb{R}^d : \|\theta - \hat{\theta}_{t-1}\|_{V_{t-1}}^2 \leq \beta_t \right\}$$

ایده UCB برای بندیت خطی تصادفی

۰ــ میانگین تجربی برای θ_*

- $$\hat{\theta}_t = \text{argmin}_{\theta \in \mathbb{R}^d} \left(\sum_{s=1}^{t} (X_s - \langle \theta, A_s \rangle)^2 + \lambda \|\theta\|_2^2 \right)$$

۱ــ یک بازه اطمینان برای θ_*: $\mathcal{C}_t$

- $$\mathcal{C}_t \subseteq \mathcal{E}_t = \left\{ \theta \in \mathbb{R}^d : \|\theta - \hat{\theta}_{t-1}\|_{V_{t-1}}^2 \leq \beta_t \right\}$$

۲ــ خوش‌بینانه‌ترین تخمین برای هر دسته

ایده UCB برای بندیت خطی تصادفی

۰ ــ میانگین تجربی برای θ_*

- $$\hat{\theta}_t = \text{argmin}_{\theta \in \mathbb{R}^d} \left(\sum_{s=1}^{t} (X_s - \langle \theta, A_s \rangle)^2 + \lambda \|\theta\|_2^2 \right)$$

۱ ــ یک بازه اطمینان برای θ_*: $\mathcal{C}_t$

- $$\mathcal{C}_t \subseteq \mathcal{E}_t = \left\{ \theta \in \mathbb{R}^d : \|\theta - \hat{\theta}_{t-1}\|_{V_{t-1}}^2 \leq \beta_t \right\}$$

۲ ــ خوش‌بینانه‌ترین تخمین برای هر دسته

- $$\text{UCB}_t(a) = \max_{\theta \in \mathcal{C}_t} \langle \theta, a \rangle$$

ایده UCB برای بندیت خطی تصادفی

۰ـ میانگین تجربی برای θ_*

- $$\hat{\theta}_t = \text{argmin}_{\theta \in \mathbb{R}^d} \left(\sum_{s=1}^{t} (X_s - \langle \theta, A_s \rangle)^2 + \lambda \|\theta\|_2^2 \right)$$

۱ـ یک بازه اطمینان برای θ_*: $\mathcal{C}_t$

- $$\mathcal{C}_t \subseteq \mathcal{E}_t = \left\{ \theta \in \mathbb{R}^d : \|\theta - \hat{\theta}_{t-1}\|_{V_{t-1}}^2 \leq \beta_t \right\}$$

۲ـ خوش‌بینانه‌ترین تخمین برای هر دسته

- $$\text{UCB}_t(a) = \max_{\theta \in \mathcal{C}_t} \langle \theta, a \rangle$$

۳ـ انتخاب دسته با بهترین خوش‌بینانه‌ترین تخمین

ایده UCB برای بندیت خطی تصادفی

۰ـ میانگین تجربی برای θ_*

- $$\hat{\theta}_t = \mathrm{argmin}_{\theta \in \mathbb{R}^d} \left(\sum_{s=1}^{t} (X_s - \langle \theta, A_s \rangle)^2 + \lambda \|\theta\|_2^2 \right)$$

۱ـ یک بازه اطمینان برای θ_*: $\mathcal{C}_t$

- $$\mathcal{C}_t \subseteq \mathcal{E}_t = \left\{ \theta \in \mathbb{R}^d : \|\theta - \hat{\theta}_{t-1}\|_{V_{t-1}}^2 \leq \beta_t \right\}$$

۲ـ خوش‌بینانه‌ترین تخمین برای هر دسته

- $$\mathrm{UCB}_t(a) = \max_{\theta \in \mathcal{C}_t} \langle \theta, a \rangle$$

۳ـ انتخاب دسته با بهترین خوش‌بینانه‌ترین تخمین

- $$A_t = \mathrm{argmax}_{a \in \mathcal{A}_t} \mathrm{UCB}_t(a).$$

ASSUMPTION 19.1. The following hold:

(a) $1 \leq \beta_1 \leq \beta_2 \leq \ldots \beta_n$.
(b) $\max_{t \in [n]} \sup_{a,b \in \mathcal{A}_t} \langle \theta_*, a - b \rangle \leq 1$.
(c) $\|a\|_2 \leq L$ for all $a \in \bigcup_{t=1}^n \mathcal{A}_t$.
(d) There exists a $\delta \in (0,1)$ such that with probability $1 - \delta$, for all $t \in [n]$, $\theta_* \in \mathcal{C}_t$ where $\mathcal{C}_t$ satisfies Eq. (19.7).

COROLLARY 19.3. *Under the conditions of Assumption 19.1, the expected regret of LinUCB with $\delta = 1/n$ is bounded by*

$$R_n \leq C d \sqrt{n} \log(nL),$$

where $C > 0$ is a suitably large universal constant.

پایان