Supplementary Material to "Efficient UCB-based Assignment Algorithm under Unknown Utility with Application in Mentor-Mentee Matching"

Table of Contents

A Proof of Lemmas **1**

A.1 Proof of Lemma 1 1

A.2 Proof of Lemma 2 3

B Proof of Theorems **4**

B.1 Proof of Theorem 1 4

B.2 Proof of Theorem 2 5

A Proof of Lemmas

A.1 Proof of Lemma 1

Proof. By the construction of $\widehat{\theta}^t$, we have

$$\begin{aligned}
\phi_{i,j}^{t\top}\widehat{\theta}^t - \phi_{i,j}^{t\top}\theta^* &= \phi_{i,j}^{t\top}M_{t-1}^{-1}r_{t-1} - \phi_{i,j}^{t\top}M_{t-1}^{-1}M_{t-1}\theta^* \\
&= \phi_{i,j}^{t\top}M_{t-1}^{-1}\sum_{\tau=1}^{t-1}\sum_{i=1}^{n_\tau}\phi_{i,\delta_\tau(i)}^{\tau}U_{i,\delta_\tau(i)}^{\tau} - \phi_{i,j}^{t\top}M_t^{-1}\left(\alpha I_d + \sum_{\tau=1}^{t-1}\sum_{i=1}^{n_t}\phi_{i,\delta_\tau(i)}^{\tau}\phi_{i,\delta_\tau(i)}^{\tau\top}\right)\theta^* \\
&= \phi_{i,j}^{t\top}M_{t-1}^{-1}\sum_{\tau=1}^{t-1}\sum_{i=1}^{n_\tau}\phi_{i,\delta_\tau(i)}^{\tau}\left(U_{i,\delta_\tau(i)}^{\tau} - \phi_{i,\delta_\tau(i)}^{\tau\top}\theta^*\right) - \alpha\phi_{i,j}^{t\top}M_{t-1}^{-1}\theta^*.
\end{aligned}$$

When $\|\theta^*\|_2 \leqslant B$, we have

$$\left|\phi_{i,j}^{t\top}\widehat{\theta}^t - \phi_{i,j}^{t\top}\theta^*\right| \leqslant \left|\phi_{i,j}^{t\top}M_{t-1}^{-1}\sum_{\tau=1}^{t-1}\sum_{i=1}^{n_\tau}\phi_{i,\delta_\tau(i)}^{\tau}\left(U_{i,\delta_\tau(i)}^{\tau} - \phi_{i,\delta_\tau(i)}^{\tau\top}\theta^*\right)\right| + \alpha B\|\phi_{i,j}^{t\top}M_{t-1}^{-1}\|_2. \tag{11}$$

Next, we bound the two terms on the right-hand side in (11) separately. To bound the first term, we first note that every $\left(U_{i,\delta_\tau(i)}^{\tau} - \phi_{i,\delta_\tau(i)}^{\tau\top}\theta^*\right)$ is σ-sub-Gaussian with mean 0, and mutually independent. Recall that we define $s_{i,j}^t = \sqrt{\phi_{i,j}^{t\top} M_{t-1}^{-1} \phi_{i,j}^t}$. Then, for any $\eta > 0$, we have

$$
\begin{aligned}
&P\left(\left|\phi_{i,j}^{t\top} M_{t-1}^{-1} \sum_{\tau=1}^{t-1} \sum_{i=1}^{n_\tau} \phi_{i,\delta_\tau(i)}^{\tau} \left(U_{i,\delta_\tau(i)}^{\tau} - \phi_{i,\delta_\tau(i)}^{\tau\top} \theta^*\right)\right| > \eta\right) \\
\leqslant & 2\exp\left(-\frac{\eta^2}{2\sigma^2 \sum_{\tau=1}^{t-1} \sum_{i=1}^{n_\tau} \left(\phi_{i,j}^{t\top} M_{t-1}^{-1} \phi_{i,\delta_\tau(i)}^{\tau}\right)^2}\right) \\
\leqslant & 2\exp\left(-\frac{\eta^2}{2\sigma^2 (s_{i,j}^t)^2}\right).
\end{aligned}
$$

Here the last inequality uses the fact that

$$
\begin{aligned}
\sum_{\tau=1}^{t-1} \sum_{i=1}^{n_\tau} \left(\phi_{i,j}^{t\top} M_{t-1}^{-1} \phi_{i,\delta_\tau(i)}^{\tau}\right)^2 &= \sum_{\tau=1}^{t-1} \sum_{i=1}^{n_\tau} \phi_{i,j}^{t\top} M_{t-1}^{-1} \phi_{i,\delta_\tau(i)}^{\tau} \phi_{i,\delta_\tau(i)}^{\tau\top} M_{t-1}^{-1} \phi_{i,j}^t \\
&\leqslant \phi_{i,j}^{t\top} M_{t-1}^{-1} \left(\alpha I_d + \sum_{\tau=1}^{t-1} \sum_{i=1}^{n_\tau} \phi_{i,\delta_\tau(i)}^{\tau} \phi_{i,\delta_\tau(i)}^{\tau\top}\right) M_{t-1}^{-1} \phi_{i,j}^t \\
&= \phi_{i,j}^{t\top} M_{t-1}^{-1} \phi_{i,j}^t = (s_{i,j}^t)^2.
\end{aligned}
$$

For given $\delta > 0$, letting $\eta = (\lambda - 1) s_{i,j}^t$ with our choice

$$
\lambda = BR + \sigma \sqrt{2 \log \frac{2\sum_{t=1}^T n_t^2}{\delta}},
$$

we have

$$
P\left(\left|\phi_{i,j}^{t\top} M_{t-1}^{-1} \sum_{\tau=1}^{t-1} \sum_{i=1}^{n} \phi_{i,\delta_\tau(i)}^{\tau} \left(U_{i,\delta_\tau(i)}^{\tau} - \phi_{i,\delta_\tau(i)}^{\tau\top} \theta^*\right)\right| > (\lambda - BR) s_{i,j}^t\right) \leqslant \frac{\delta}{\sum_{t=1}^T n_t^2}.
$$

Using a union bound argument, we have that

$$
P\left(\left|\phi_{i,j}^{t\top} M_{t-1}^{-1} \sum_{\tau=1}^{t-1} \sum_{i=1}^{n_\tau} \phi_{i,\delta_\tau(i)}^{\tau} \left(U_{i,\delta_\tau(i)}^{\tau} - \phi_{i,\delta_\tau(i)}^{\tau\top} \theta^*\right)\right| \leqslant (\lambda - BR) s_{i,j}^t,\ \forall t \in [T],\ i, j \in [n_t]\right) \geqslant 1 - \delta. \tag{12}
$$

We now bound the second term on the right-hand side of (11). Note that

$$
\|\phi_{i,j}^{t\top} M_{t-1}^{-1}\|_2^2 \leqslant \frac{1}{\alpha} \phi_{i,j}^{t\top} M_{t-1}^{-1} M_{t-1} M_{t-1}^{-1} \phi_{i,j}^t = \frac{1}{\alpha} (s_{i,j}^t)^2.
$$

Hence we have $\alpha B \|\phi_{i,j}^{t\top} M_{t-1}^{-1}\|_2 \leqslant \sqrt{\alpha} B s_{i,j}^t = BR s_{i,j}^t$. Combining this with (11) and (12), we conclude that

$$
P\left(\left|\phi_{i,j}^{t\top} \widehat{\theta}^t - \phi_{i,j}^{t\top} \theta^*\right| \leqslant \lambda s_{i,j}^t,\ \forall t \in [T],\ i, j \in [n_t]\right) \geqslant 1 - \delta.
$$

This completes the proof for Lemma 1.

□

A.2 Proof of Lemma 2

Proof. We first introduce the following lemma that will be used to bound $\|\phi_i^t\|^2_{M_{t-1}^{-1}}$.

Lemma 3. *Let ϕ be a d-dimensional vector with $\|\phi\|_2 \leqslant R$. Let M be a $d \times d$ positive definite matrix with the minimum eigenvalue $\lambda_{\min}(M) \geqslant R^2$. Then we have*

$$\|\phi\|^2_{M^{-1}} \leqslant 2 \log \frac{\det(M + \phi\phi^\top)}{\det(M)}.$$

To prove Lemma 3, we note that

$$\begin{aligned} \det(M + \phi\phi^\top) &= \det(M) \det\left(I_d + M^{-1/2}\phi(M^{-1/2}\phi)^\top\right) \\ &= \det(M)\left(1 + \|\phi\|^2_{M^{-1}}\right) \end{aligned}$$

Since $\|\phi\|_2 \leqslant R$ and $\lambda_{\min}(M) \geqslant R^2$, we have $\|\phi\|^2_{M^{-1}} \leqslant 1$. Using the fact that $x \leqslant 2\log(1+x)$ for $x \in [0,1]$, we get

$$\|\phi\|^2_{M^{-1}} \leqslant 2\log(1 + \|\phi\|^2_{M^{-1}}) = 2 \log \frac{\det(M + \phi\phi^\top)}{\det(M)}.$$

The proof for Lemma 3 is thus completed.

We now continue to prove Lemma 2. Note that we assume $\|\phi_{i,j}^t\|_2 \leqslant R$ for all t, i, j. Also, by our construction of M_t's, we have $\lambda_{\min}(M_t) \geqslant R^2$ for all t. By defining $M_{t-1,i} = M_{t-1} + \phi^t_{i,\delta_t(i)}\phi^{t\top}_{i,\delta_t(i)}$ and using Lemma 3, we have

$$\begin{aligned} \sum_{t=1}^{T}\sum_{i=1}^{n_t} \frac{\|\phi^t_{i,\delta_t(i)}\|^2_{M_{t-1}^{-1}}}{n_t} &\leqslant \sum_{t=1}^{T}\sum_{i=1}^{n_t} 2\log\left[\frac{\det(M_{t-1,i})}{\det(M_{t-1})}\right]^{\frac{1}{n_t}} \\ &= 2\log\prod_{t=1}^{T}\frac{\left[\prod_{i=1}^{n_t}\det(M_{t-1,i})\right]^{\frac{1}{n_t}}}{\det(M_{t-1})} \\ &\leqslant 2\log\frac{\left[\prod_{i=1}^{n_T}\det(M_{T-1,i})\right]^{\frac{1}{n_T}}}{\det(M_0)}. \end{aligned} \tag{13}$$

Here the last inequality uses the fact that $\det(M_{t-1,i}) \leqslant \det(M_t)$ for every t and i, by our construction. Since $M_0 = R^2 I_d$ and $\|\phi^t_{i,j}\|_2 \leqslant R$, the maximum value $\det(M_{T-1,i})$ can take is $\left(R^2 + \frac{(\sum_{t=1}^{T-1} n_t)+1}{d}R^2\right)^d$, which can be further upper bounded by $\left(R^2 + \frac{\sum_{t=1}^{T} n_t}{d}R^2\right)^d$. Therefore, following (13), we have

$$\sum_{t=1}^{T}\sum_{i=1}^{n_t} \frac{\|\phi^t_{i,\delta_t(i)}\|^2_{M_{t-1}^{-1}}}{n_t} \leqslant 2d\log\left(R^2 + \frac{\sum_{t=1}^{T} n_t}{d}R^2\right).$$

This completes our proof for Lemma 2. □

B Proof of Theorems

B.1 Proof of Theorem 1

Proof. By the construction of $\widehat{\theta}^t$, we have

$$\begin{aligned}
\widehat{\theta}^t - \theta^* &= M_{t-1}^{-1} r_{t-1} - M_{t-1}^{-1} M_{t-1} \theta^* \\
&= M_{t-1}^{-1} \sum_{\tau=1}^{t-1} \sum_{i=1}^{n_\tau} \phi_{i,\delta_\tau(i)}^{\tau} U_{i,\delta_\tau(i)}^{\tau} - M_{t-1}^{-1} \left(\alpha I_d + \sum_{\tau=1}^{t-1} \sum_{i=1}^{n_t} \phi_{i,\delta_\tau(i)}^{\tau} \phi_{i,\delta_\tau(i)}^{\tau\top} \right) \theta^* \\
&= M_{t-1}^{-1} \sum_{\tau=1}^{t-1} \sum_{i=1}^{n_\tau} \phi_{i,\delta_\tau(i)}^{\tau} \left(U_{i,\delta_\tau(i)}^{\tau} - \phi_{i,\delta_\tau(i)}^{\tau\top} \theta^* \right) - \alpha M_{t-1}^{-1} \theta^*.
\end{aligned}$$

Note that $\left(U_{i,\delta_\tau(i)}^{\tau} - \phi_{i,\delta_\tau(i)}^{\tau\top} \theta^* \right)$ is assumed to be σ-sub-Gaussian with mean 0. Following the equation above, we have

$$\begin{aligned}
\|\widehat{\theta}^t - \theta^*\|_{M_{t-1}}^2 &= (\widehat{\theta}^t - \theta^*)^\top M_{t-1} (\widehat{\theta}^t - \theta^*) \\
&\leqslant 2 \left\| \sum_{\tau=1}^{t-1} \sum_{i=1}^{n_\tau} \phi_{i,\delta_\tau(i)}^{\tau} \left(U_{i,\delta_\tau(i)}^{\tau} - \phi_{i,\delta_\tau(i)}^{\tau\top} \theta^* \right) \right\|_{M_{t-1}^{-1}}^2 + 2\alpha^2 \|\theta^*\|_{M_{t-1}^{-1}}^2 .
\end{aligned} \tag{14}$$

Before we continue, we introduce a useful Lemma to help bound the first term on the right-hand side of (14).

Lemma 4 (Theorem 1 in Abbasi-Yadkori et al. (2011))**.** *Let $\{F_t\}_{t=0}^{\infty}$ be a filtration. Let $\{\epsilon_t\}_{t=1}^{\infty}$ be a real-valued stochastic process such that ϵ_t is F_{t-1}-measurable, and ϵ_t is conditionally σ-sub-Gaussian, i.e.,*

$$\mathbb{E}[\exp(\lambda \epsilon_t) | F_{t-1}] \leqslant \exp\left(\frac{\lambda^2 \sigma^2}{2} \right), \ \forall \lambda \in \mathbb{R}.$$

Let $\{X_t\}_{t=1}^{\infty}$ be an $\mathbb{R}^d$-valued stochastic process, such that X_t is F_{t-1}-measurable. Assume M is a $d \times d$ positive definite matrix. Define

$$\bar{M}_t = M + \sum_{s=1}^{t} X_s X_s^\top, \quad S_t = \sum_{s=1}^{t} \epsilon_s X_s.$$

Then for any $\delta > 0$, with probability at least $1 - \delta$,

$$\|S_t\|_{\bar{M}_t^{-1}}^2 \leqslant 2\sigma^2 \log\left(\frac{\det(\bar{M}_t)^{1/2} \det(M)^{-1/2}}{\delta} \right), \ \textit{for all } t \geqslant 0.$$

Using Lemma 4, with probability at least $1 - \delta$, we have

$$\begin{aligned}
\left\| \sum_{\tau=1}^{t-1} \sum_{i=1}^{n_\tau} \phi_{i,\delta_\tau(i)}^{\tau} \left(U_{i,\delta_\tau(i)}^{\tau} - \phi_{i,\delta_\tau(i)}^{\tau\top} \theta^* \right) \right\|_{M_{t-1}^{-1}}^2 &\leqslant 2\sigma^2 \log\left(\frac{\det(M_{t-1})^{1/2}}{\delta} \right) \\
&\leqslant \sigma^2 d \left[\log\left(R^2 + \frac{\sum_{\tau=1}^{t-1} n_\tau}{d} R^2 \right) + 2 \log\left(\frac{1}{\delta} \right) \right].
\end{aligned} \tag{15}$$

Here the last inequality uses the assumption that $\|\phi\|_2 \leqslant R$. Meanwhile, we also have

$$\|\theta^*\|^2_{M_{t-1}^{-1}} \leqslant \frac{1}{\alpha}\|\theta^*\|_2^2 \leqslant \frac{B^2}{\alpha}. \tag{16}$$

Plugging (15) and (16) into (14), and using $\alpha = R^2$, we finally obtain that

$$\|\widehat{\theta}^t - \theta^*\|^2_{M_{t-1}} \leqslant \sigma^2 d\left[\log\left(R^2 + \frac{\sum_{\tau=1}^{t-1} n_\tau}{d}R^2\right) + 2\log\left(\frac{1}{\delta}\right)\right] + 2B^2R^2.$$

This completes our proof for Theorem 1. □

B.2 Proof of Theorem 2

Proof. For notational simplicity, we first define

$$V_t = \frac{1}{n}\sum_{i=1}^{n}\phi(\boldsymbol{x}_i^t, \boldsymbol{z}^t_{\delta_t(i)})^\top\theta^*,$$
$$V_t^* = \frac{1}{n}\sum_{i=1}^{n}\phi(\boldsymbol{x}_i^t, \boldsymbol{z}^t_{\delta_t^*(i)})^\top\theta^*,$$

where δ_t and δ_t^* denotes the assignment decided by the algorithm and the oracle assignment at t, respectively. Then we have

$$R_T = \sum_{t=1}^{T}(V_t^* - V_t).$$

We first upper bound the performance gap $(V_t^* - V_t)$ for each single round t, and then upper bound their summation as the total regret.

Under the events $\mathcal{E}_i^t$'s, we can first bound the term $(V_t^* - V_t)$ as below:

$$\begin{aligned}
V_t^* - V_t &= \frac{1}{n_t}\sum_{i=1}^{n_t}(\phi^t_{i,\delta_t^*(i)} - \phi^t_{i,\delta(i)})^\top\theta^* \\
&\leqslant \frac{1}{n_t}\sum_{i=1}^{n_t}\left(\phi^{t\top}_{i,\delta_t^*(i)}\widehat{\theta}^t + \lambda s^t_{i,\delta_t^*(i)}\right) - \sum_{i=1}^{n_t}\phi^{t\top}_{i,\delta_t(i)}\theta^* \\
&\leqslant \frac{1}{n_t}\sum_{i=1}^{n_t}\left(\phi^{t\top}_{i,\delta_t(i)}\widehat{\theta}^t + \lambda s^t_{i,\delta_t(i)}\right) - \sum_{i=1}^{n_t}\phi^{t\top}_{i,\delta_t(i)}\theta^* \\
&\leqslant \frac{1}{n_t}\sum_{i=1}^{n_t}2\lambda s^t_{i,\delta_t(i)}.
\end{aligned} \tag{17}$$

Here the second inequality is by the construction of our assignment δ_t that maximizes the total upper confidence bound, and the first and third inequality uses the events $\mathcal{E}^t_{i,j}$'s.

From (17), we have

$$R_T \leqslant 2\lambda\sum_{t=1}^{T}\sum_{i=1}^{n_t}\frac{s^t_{i,\delta_t(i)}}{n_t}. \tag{18}$$

To further bound the right-hand side, we use Lemma 2 to obtain that

$$\begin{aligned}
\sum_{t=1}^{T}\sum_{i=1}^{n_t}\frac{s^t_{i,\delta_t(i)}}{n_t} &\leqslant \sqrt{T\sum_{t=1}^{T}\left(\sum_{i=1}^{n_t}\frac{s^t_{i,\delta_t(i)}}{n_t}\right)^2} \\
&\leqslant \sqrt{T\sum_{t=1}^{T}\sum_{i=1}^{n_t}\frac{\left(s^t_{i,\delta_t(i)}\right)^2}{n_t}} \\
&= \sqrt{T\sum_{t=1}^{T}\sum_{i=1}^{n_t}\frac{\|\phi^t_{i,\delta_t(i)}\|^2_{M^{-1}_{t-1}}}{n_t}} \\
&\leqslant \sqrt{2dT\log\left(R^2+\frac{\sum_{t=1}^{T}n_t}{d}R^2\right)}. \qquad (19)
\end{aligned}$$

Combining (19) with our choice of λ in (7), we obtain that

$$\begin{aligned}
R_T \leqslant & 2\lambda\sum_{t=1}^{T}\sum_{i=1}^{n_t}\frac{s^t_{i,\delta_t(i)}}{n_t} \\
\leqslant & 4\sigma\sqrt{dT\log\frac{2\sum_{t=1}^{T}n_t^2}{\delta}\log\left(R^2+\frac{\sum_{t=1}^{T}n_t}{d}R^2\right)} \\
& + 2BR\sqrt{2dT\log\left(R^2+\frac{\sum_{t=1}^{T}n_t}{d}R^2\right)},
\end{aligned}$$

which completes our proof for Theorem 2. □