
Supplementary material to “A Gang of Bandits”

Nicolò Cesa-Bianchi
Università degli Studi di Milano, Italy
nicolo.cesa-bianchi@unimi.it

Claudio Gentile
University of Insubria, Italy
claudio.gentile@uninsubria.it

Giovanni Zappella
Università degli Studi di Milano, Italy
giovanni.zappella@unimi.it

A Appendix

This appendix contains the proof of Theorem 1.

Proof. Recall that

$$\tilde{U} = A_{\otimes}^{1/2} U \quad \text{where} \quad U = (\mathbf{u}_1^\top, \mathbf{u}_2^\top, \dots, \mathbf{u}_n^\top)^\top \in \mathbb{R}^{dn}.$$

Let then t be a fixed time step, and introduce the following shorthand notation:

$$\mathbf{x}_t^* = \operatorname{argmax}_{k=1, \dots, c_t} \mathbf{u}_{i_t}^\top \mathbf{x}_{t,k} \quad \text{and} \quad \tilde{\phi}_t^* = \operatorname{argmax}_{k=1, \dots, c_t} \tilde{U}^\top \tilde{\phi}_{t,k}.$$

Notice that, for any k we have

$$\tilde{U}^\top \tilde{\phi}_{t,k} = U^\top A_{\otimes}^{1/2} A_{\otimes}^{-1/2} \phi_{i_t}(\mathbf{x}_{t,k}) = U^\top \phi_{i_t}(\mathbf{x}_{t,k}) = \mathbf{u}_{i_t}^\top \mathbf{x}_{t,k}.$$

Hence we decompose the time- t regret r_t as follows:

$$\begin{aligned} r_t &= \mathbf{u}_{i_t}^\top \mathbf{x}_t^* - \mathbf{u}_{i_t}^\top \mathbf{x}_{t,k_t} \\ &= \tilde{U}^\top \tilde{\phi}_t^* - \tilde{U}^\top \tilde{\phi}_{t,k_t} \\ &= \tilde{U}^\top \tilde{\phi}_t^* - \mathbf{w}_{t-1}^\top \tilde{\phi}_t^* + \mathbf{w}_{t-1}^\top \tilde{\phi}_t^* + \text{CB}_t(\tilde{\phi}_t^*) - \text{CB}_t(\tilde{\phi}_t^*) - \tilde{U}^\top \tilde{\phi}_{t,k_t} \\ &\leq \tilde{U}^\top \tilde{\phi}_t^* - \mathbf{w}_{t-1}^\top \tilde{\phi}_t^* + \mathbf{w}_{t-1}^\top \tilde{\phi}_{t,k_t} + \text{CB}_t(\tilde{\phi}_{t,k_t}) - \text{CB}_t(\tilde{\phi}_t^*) - \tilde{U}^\top \tilde{\phi}_{t,k_t}, \end{aligned}$$

the inequality deriving from

$$\mathbf{w}_{t-1}^\top \tilde{\phi}_{t,k_t} + \text{CB}_t(\tilde{\phi}_{t,k_t}) \geq \mathbf{w}_{t-1}^\top \tilde{\phi}_{t,k} + \text{CB}_t(\tilde{\phi}_{t,k}), \quad k = 1, \dots, c_t.$$

At this point, we rely on [1] (Theorem 2 therein with $\lambda = 1$) to show that

$$|\tilde{U}^\top \tilde{\phi}_t^* - \mathbf{w}_{t-1}^\top \tilde{\phi}_t^*| \leq \text{CB}_t(\tilde{\phi}_t^*) \quad \text{and} \quad |\mathbf{w}_{t-1}^\top \tilde{\phi}_{t,k_t} - \tilde{U}^\top \tilde{\phi}_{t,k_t}| \leq \text{CB}_t(\tilde{\phi}_{t,k_t})$$

both hold simultaneously for all t with probability at least $1 - \delta$ over the noise sequence. Hence, with the same probability,

$$r_t \leq 2 \text{CB}_t(\tilde{\phi}_{t,k_t})$$

holds uniformly over t . Thus the cumulative regret $\sum_{t=1}^T r_t$ satisfies

$$\begin{aligned}
\sum_{t=1}^T r_t &\leq \sqrt{T \sum_{t=1}^T r_t^2} \\
&\leq 2 \sqrt{T \sum_{t=1}^T \left(\text{CB}_t(\tilde{\phi}_{t,k_t}) \right)^2} \\
&\leq 2 \sqrt{T \left(\sigma \sqrt{\ln \frac{|M_T|}{\delta}} + \|\tilde{U}\| \right)^2 \sum_{t=1}^T \tilde{\phi}_{t,k_t}^\top M_{t-1}^{-1} \tilde{\phi}_{t,k_t}} .
\end{aligned}$$

Now, using (see, e.g., [2])

$$\sum_{t=1}^T \tilde{\phi}_{t,k_t}^\top M_{t-1}^{-1} \tilde{\phi}_{t,k_t} \leq (1 + \max_{k=1,\dots,c_t} \|\tilde{\phi}_{t,k}\|^2) \ln |M_T| ,$$

with

$$\begin{aligned}
\max_{k=1,\dots,c_t} \|\tilde{\phi}_{t,k}\|^2 &= \max_{k=1,\dots,c_t} \phi_{i_t}(\mathbf{x}_{t,k}) A_\otimes^{-1} \phi_{i_t}(\mathbf{x}_{t,k}) \\
&\leq \max_{k=1,\dots,c_t} \|\phi_{i_t}(\mathbf{x}_{t,k})\|^2 \\
&= \max_{k=1,\dots,c_t} \|\mathbf{x}_{t,k}\|^2 \\
&\leq B^2 ,
\end{aligned}$$

along with $(a+b)^2 \leq 2a^2 + 2b^2$ applied with $a = \sigma \sqrt{\ln \frac{|M_T|}{\delta}}$ and $b = \|\tilde{U}\|$ yields

$$\sum_{t=1}^T r_t \leq 2 \sqrt{T \left(2\sigma^2 \ln \frac{|M_T|}{\delta} + 2\|\tilde{U}\|^2 \right) (1 + B^2) \ln |M_T|} .$$

Finally, observing that

$$\|\tilde{U}\|^2 = U^\top A_\otimes U = L(\mathbf{u}_1, \dots, \mathbf{u}_n)$$

gives the desired bound. □

References

- [1] Y. Abbasi-Yadkori, D. Pál, and C. Szepesvári. Improved algorithms for linear stochastic bandits. *Advances in Neural Information Processing Systems*, 2011.
- [2] O. Dekel, C. Gentile, and K. Sridharan. Robust selective sampling from single and multiple teachers. In *COLT*, pages 346–358, 2010.