A Note on Bounding Regret of the C²UCB Contextual Combinatorial Bandit

Bastian Oetomo Malinga Perera Renata Borovica-Gajic Benjamin I. P. Rubinstein*

February 21, 2019

Abstract

We revisit the proof by [Qin et al., 2014] of bounded regret of the C²UCB contextual combinatorial bandit. We demonstrate an error in the proof of volumetric expansion of the moment matrix, used in upper bounding a function of context vector norms. We prove a relaxed inequality that yields the originally-stated regret bound.

1 Introduction

In deriving a regret bound on the C²UCB contextual combinatorial bandit, Qin et al. (2014) use the following equality within the proof of their Lemma 4.2.

Claim 1. Let \( k, m, n \) be natural numbers, \( V \) be a \( d \times d \) real and positive definite matrix, and \( S_t \subseteq [m] \) with \( |S_t| \leq k \leq m \) for \( t \in [n] \). Let \( x_t(i) \in \mathbb{R}^d \) be vectors for \( t \in [n], i \in [m] \), and define \( V_n = V + \sum_{i=1}^{n} \sum_{i \in S_t} x_t(i)x_t(i)^T \).

Then \( \det(V_n) = \det(V) \prod_{i=1}^{n} \left( 1 + \sum_{i \in S_t} \|x_t(i)\|^2_{V_{t-1}} \right) \), where we define \( \|a\|_M = \sqrt{a^TMa} \).

We present a counterexample to Claim 1 in Section 2, and then in Section 3 prove the relaxation given by,

Lemma 2. Under the same conditions as Claim 1, \( \det(V_n) \geq \det(V) \prod_{i=1}^{n} \left( 1 + \sum_{i \in S_t} \|x_t(i)\|^2_{V_{t-1}} \right) \).

In the setting of C²UCB, \( [n], [m] \) correspond to rounds and arms, \( S_t \) the (super arm of) played arms in round \( t \), \( x_t(i) \) the context vector for arm \( i \) at round \( t \), and \( V_t \) the covariance matrix from the played contexts added to \( V \) (taken to be a scaled identity, for achieving ridge regression reward estimates). We detail in Section 3 how Lemma 2 can be used within the remainder of the proof of (Qin et al., 2014) Lemma 4.2, ultimately yielding the C²UCB regret bound originally claimed. The regret analysis of C²UCB is based on previous analysis of contextual bandits (Auer, 2002; Dani et al., 2008; Chu et al., 2011). We demonstrate that the bound in Lemma 2 is sharp, by describing conditions for equality.

Notation. We denote by \( \lambda_j(A) \) the eigenvalues of the \( n \times n \) matrix \( A \), where, without loss of generality, \( \lambda_1(A) \leq \lambda_2(A) \leq \cdots \leq \lambda_n(A) \). We likewise order \( S_t = \{s_{(1,t)}, s_{(2,t)}, \ldots, s_{(|S_t|,t)}\} \), where \( s_{(1,t)} < s_{(2,t)} < \cdots < s_{(|S_t|,t)} \).

Generalised Matrix Determinant Lemma. We make use of the identity: Let \( A \) be an invertible \( n \times n \) matrix, and \( B, C \) be \( n \times m \) matrices, then \( \det(A + BC^T) = \det(I_n + C^TA^{-1}B) \det(A) \).

2 A Counterexample

Claim 1 derives from the assertion within the proof of (Qin et al., 2014) Lemma 4.2 that,

\[
\det(V_{n-1}) \det \left( I + \sum_{i \in S_n} (V_{n-1}^{-1/2} x_n(i))(V_{n-1}^{-1/2} x_n(i))^T \right) = \det(V_{n-1}) \det \left( I + \sum_{i \in S_n} \|x_n(i)\|^2_{V_{n-1}} \right).
\]

This appears to conflate outer and inner products, after basis transformation by \( V_{n-1}^{-1/2} \). The following counterexample to Claim 1 establishes that indeed it does not hold in general.

\*School of Computing and Information Systems, University of Melbourne, Parkville, VIC 3010, Australia. {boetomo,wpperera}@student.unimelb.edu.au, {rborovica,brubinstein}@unimelb.edu.au
Example 3. Consider \( n = 1 \), the \( 2 \times 2 \) matrix \( V = 1.2I_2 \), \( S_t = \{1, 2, 3\} \) and let \( x_1(1) = \begin{bmatrix} 0.3 \\ 0.1 \end{bmatrix} \), \( x_1(3) = \begin{bmatrix} 0.1 \\ 0.5 \end{bmatrix} \). It follows that \( V_1 = \begin{bmatrix} 1.66 & 0.32 \\ 0.32 & 1.95 \end{bmatrix} \). Then we have

\[
\det(V) = \prod_{i=1}^{n} \left( 1 + \sum_{t \in S_t} ||x_t(i)||^2 \right)
\]

\[
= \det(V) \left( 1 + \sum_{t \in S_t} x_t(i)^T V^{-1} x_t(i) \right)
\]

\[
= \det(1.2I_2) \left( 1 + x_1(1)^T \left( \frac{1}{1.2} \right) x_1(1) + x_1(2)^T \left( \frac{1}{1.2} \right) x_1(2) + x_1(3)^T \left( \frac{1}{1.2} \right) x_1(3) \right)
\]

\[
= 1.2^2 \left( 1 + \frac{1}{1.2} (0.3^2 + 0.7^2) + \frac{1}{1.2} (0.6^2 + 0.1^2) + \frac{1}{1.2} (0.1^2 + 0.5^2) \right)
\]

\[
= 2.892 \neq 3.1346 = 1.66 \times 1.95 - 0.32 \times 0.32 = \det(V_1).
\]

3 Proof of Lemma [2]

Let \( X_n = [x_n(s_{1,n}) \ldots x_n(s_{|S_n|,n})] \). Then,

\[
\det(V_n) = \det \left( V + \sum_{t=1}^{n-1} \sum_{i \in S_t} x_t(i)x_t(i)^T + \sum_{i \in S_n} x_n(i)x_n(i)^T \right)
\]

\[
= \det(V_n - 1 + X_nX_n^T)
\]

\[
= \det(V_n - 1) \det \left( I_{|S_n|} + X_n^T V_{n-1}^{-1} X_n \right)
\]

\[
= \det(V_n - 1) \left[ \prod_{i=1}^{|S_n|} \lambda_i \left( I_{|S_n|} + X_n^T V_{n-1}^{-1} X_n \right) \right]
\]

where the fourth and final equalities follow from the Generalised Matrix Determinant Lemma and the fact that adding the identity to a square matrix increases eigenvalues by one. Now, the final line's product can be expanded as

\[
1 + \sum_{i=1}^{|S_n|} \lambda_i (X_n^T V_{n-1}^{-1} X_n) + \sum_{1 \leq t < i \leq |S_n|} \lambda_t \lambda_i (X_n^T V_{n-1}^{-1} X_n) (X_n^T V_{n-1}^{-1} X_n) + \cdots + \prod_{i=1}^{|S_n|} \lambda_i (X_n^T V_{n-1}^{-1} X_n).
\]

(1)

Since \( V \) is positive definite and \( x_t(i)x_t(i)^T \) is positive semi-definite (with one eigenvalue being \( x_t(i)^T x_t(i) \) and the remainder all zero) for all \( t \) and \( i \), we have that \( V_n - 1 = V + \sum_{t=1}^{n-1} \sum_{i \in S_t} x_t(i)x_t(i)^T \) is positive definite. Therefore, we conclude that \( V_{n-1}^{-1} \) is also positive definite, hence it has a symmetric square root matrix \( V_{n-1}^{-1/2} \). It also follows that \( X_n^T V_{n-1}^{-1} X_n \) is positive semi-definite. Therefore, the terms starting from the third term in the expansion (1) are all non-negative because they are products of the eigenvalues of \( X_n^T V_{n-1}^{-1} X_n \). Thus, we have,

\[
\det(V_n) = \det(V_n - 1) \left[ \prod_{i=1}^{|S_n|} \left( 1 + \lambda_i \left( X_n^T V_{n-1}^{-1} X_n \right) \right) \right]
\]

\[
\geq \det(V_n - 1) \left( 1 + \sum_{i=1}^{|S_n|} \lambda_i (X_n^T V_{n-1}^{-1} X_n) \right)
\]

\[
= \det(V_n - 1) \left( 1 + \text{tr}(X_n^T V_{n-1}^{-1} X_n) \right)
\]

\[
= \det(V_n - 1) \left( 1 + \sum_{i \in S_n} x_n(i) V_{n-1}^{-1} x_n(i) \right)
\]

2
= \det(V_{n-1}) \left( 1 + \sum_{i \in S_n} ||x_n(i)||^2_{V_{n-1}} \right),

where the third equality follows from expanding out the argument to the trace as

\[ X_n^T V_{n-1}^{-1} X_n = \begin{bmatrix}
    x_n(s_{(1,n)})^T V_{n-1}^{-1} x_n(s_{(1,n)}) & \cdots & x_n(s_{(1,n)})^T V_{n-1}^{-1} x_n(s_{(|S_n|,n)}) \\
    \vdots & \ddots & \vdots \\
    x_n(s_{(|S_n|,n)})^T V_{n-1}^{-1} x_n(s_{(|S_n|,n)}) & \cdots & x_n(s_{(|S_n|,n)})^T V_{n-1}^{-1} x_n(s_{(|S_n|,n)})
\end{bmatrix}. \]

Applying our recurrence relation on \( V_t \) for \( 1 \leq t \leq n \), we can telescope to arrive at the result.

\section{Implication of Lemma 2}

By rearranging the inequality, we know that

\[ \prod_{t=1}^{n} \left( 1 + \sum_{i \in S_t} ||x_t(i)||^2_{V_{t-1}} \right) \leq \frac{\det(V_n)}{\det(V)}, \]

provided that \( \det(V) > 0 \), which is guaranteed for our positive definite \( V \). The next steps of (Qin et al., 2014, Lemma 4.2)'s proof follow the original pattern now with the second inequality in what follows (due to our Lemma 2 and monotonicity), rather than the original equality:

\[ \sum_{t=1}^{n} \sum_{i \in S_t} ||x_t(i)||^2_{V_{t-1}} \leq 2 \sum_{t=1}^{n} \log \left( 1 + \sum_{i \in S_t} ||x_t(i)||^2_{V_{t-1}} \right) = 2 \log \left( \prod_{t=1}^{n} \left( 1 + \sum_{i \in S_t} ||x_t(i)||^2_{V_{t-1}} \right) \right) \]

\[ \leq 2 \log \left( \frac{\det(V_n)}{\det(V)} \right) = 2 \log(\det(V_n)) - 2 \log(\det(V)), \]

which yields the regret bound as presented by Qin et al. (2014), without further modification to the proof of their Lemma 4.2.

\section{Discussion}

The proof of Lemma 2 offers intuition as to when the inequality holds with equality. Namely, it is true when the matrix \( X_t^T V_{t-1}^{-1} X_t \) has at most one non-zero eigenvalue \( i.e. \), be either a rank-1 or rank-0 matrix for all \( 1 \leq t \leq n \). This is because the terms that we dropped in calculating the determinant of \( I_{|S_t|} + X_t^T V_{t-1}^{-1} X_t \) are then identically 0. This agrees with the result of the non-generalised matrix determinant lemma.

This occurs when intra-round, played context vectors are co-linear to each other: if the context vector of arm \( i \) can be written as \( x_t(i) = a_i u_i \), then we can write \( X_t = u_i a_i^T \), where \( a_i \) is a column vector with \( a_{ij} \) as its components. The matrix we are interested in becomes \( X_t^T V_{t-1}^{-1} X_t = (u_i a_i^T)^T V_{t-1}^{-1} (u_i a_i^T) = ||u_i||^2_{V_{t-1}} a_i a_i^T \), which is a rank-1 matrix. Thus, it also follows that the trace of this matrix is \( ||u_i||^2_{V_{t-1}} ||a_i||^2 \). One interesting thing to notice here is that the context vectors need not to be co-linear across rounds.

A special case of the co-linearity scenario is the non-combinatorial bandit. In this scenario, \( |S_t| = 1 \) for all \( t \). This means that given a particular round \( t \), there is only one context vector available. In particular, \( \det \left( I_{|S_t|} + X_t^T V_{t-1}^{-1} X_t \right) = \det \left( I_t + x_t^T V_{t-1}^{-1} x_t \right) = 1 + x_t^T V_{t-1}^{-1} x_t \), which is the bound that we had for calculating \( \det \left( I_{|S_t|} + X_t^T V_{t-1}^{-1} X_t \right), \) were \( |S_t| = 1 \).

\section{References}

Peter Auer. Using confidence bounds for exploitation-exploration trade-offs. \textit{Journal of Machine Learning Research}, 3(Nov):397–422, 2002.

Wei Chu, Lihong Li, Lev Reyzin, and Robert Schapire. Contextual bandits with linear payoff functions. In \textit{Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics}, pages 208–214, 2011.

Varsha Dani, Thomas P. Hayes, and Sham M. Kakade. Stochastic linear optimization under bandit feedback. In \textit{Proceedings of the 21st Annual Conference on Learning Theory}, pages 355–366, 2008.

Lijing Qin, Shouyuan Chen, and Xiaoyan Zhu. Contextual combinatorial bandit and its application on diversified online recommendation. In \textit{Proceedings of the 2014 SIAM International Conference on Data Mining}, pages 461–469, 2014.

\footnote{Here as in the original proof, we leverage assumptions: \( \lambda_1(V) \geq k \) and the context vectors are of bounded norm \( ||x_t(i)||_2 \leq 1 \).}