Even Simpler Deterministic Matrix Sketching

Edo Liberty*

Abstract

This paper provides a one-line proof of Frequent Directions (FD) for sketching streams of matrices. It simplifies the main results in [1] and [2]. The simpler proof arises from sketching the *covariance* of the stream of matrices rather than the stream itself.

Introduction

Let $X_t \in \mathbb{R}^{d \times n_t}$ be a stream of matrices. Let $C = \sum_{t=1}^{T} X_t X_t^T \in \mathbb{R}^{d \times d}$ be their covariance matrix. Frequent Directions [1] maintains a rank deficient approximate covariance matrix $\tilde{C}_t \in \mathbb{R}^{d \times d}$ using Algorithm 1. Set $\tilde{C}_0 \in \mathbb{R}^{d \times d}$ to be the all zeros matrix. Then, at time $t = 1, \dots, T$ compute $\tilde{C}_t = \textsc{Update}(\tilde{C}_{t-1}, X_t, \ell)$.

Algorithm 1 Frequent Directions (FD) Update

1: **function** $\textsc{Update}(\tilde{C}_{t-1}, X_t, \ell)$
2: $\quad U_t \Lambda_t U_t^T = \tilde{C}_{t-1} + X_t X_t^T$
3: $\quad$ **return** $\tilde{C}_t = U \cdot \max(\Lambda - I \cdot \lambda_\ell^t, 0) \cdot U^T$
4: **end function**

Above, $U_t \Lambda_t U_t^T$ is the eigen-decomposition of $\tilde{C}_{t-1} + X_t X_t^T$ and λ_ℓ^t is the its ℓ'th largest eigenvalue. Note that the rank of $\tilde{C}_t$ is at most $\ell - 1$ for all t by construction. It can therefore be stored in $O(d\ell)$ space. Assuming $n_t < \ell$, the update operation itself also consumes at most $O(d\ell)$ space.

Lemma 1 (simplified from [2] and [1])**.** *Let $\tilde{C}$ denote the approximated covariance produced by FD and λ_i be the eigenvalues of the exact covariance C in descending order. For any ℓ and simultaneously for all $k < \ell$ we have*

$$\|C - \tilde{C}\| \le \frac{1}{\ell - k} \sum_{i=k+1}^{d} \lambda_i$$

*Pinecone

Short proof Lemma 1

Define $\Delta_t = X_t X_t^T - \tilde{C}_t + \tilde{C}_{t-1}$. Then $\sum_{t=1}^{T} \Delta_t = \sum_{t=1}^{T} X_t X_t^T - \sum_{t=1}^{T}(\tilde{C}_t - \tilde{C}_{t-1}) = C - \tilde{C}$ where $\tilde{C}$ stands for $\tilde{C}_T$, the final sketch.

Moreover, note that the top ℓ eigenvalues of Δ_t are all equal to one another because $\Delta_t = U_t \cdot \min(\Lambda_t, I \cdot \lambda_\ell^t) \cdot U_t^T$. As a result $\|\Delta_t\| < \frac{1}{\ell - k} \operatorname{tr}(\bar{P}_k \Delta_t \bar{P}_k)$ for any projection $\bar{P}_k$ having a null space of dimension at most k. Specifically, this holds for $\bar{P}_k$ whose null space contains the eigenvectors of C corresponding to its largest eigenvalues.

$$\begin{aligned}
\|C - \tilde{C}\| &= \|\sum_{t=1}^{T} \Delta_t\| \le \sum_{t=1}^{T} \|\Delta_t\| \\
&\le \frac{1}{\ell - k} \operatorname{tr}\left(\bar{P}_k \left(\sum_{t=1}^{T} \Delta_t\right) \bar{P}_k\right) \\
&\le \frac{1}{\ell - k} \operatorname{tr}\left(\bar{P}_k C \bar{P}_k\right) = \frac{1}{\ell - k} \sum_{i=k+1}^{d} \lambda_i
\end{aligned}$$

Here we used that $\operatorname{tr}(\bar{P}_k \tilde{C} \bar{P}_k) \ge 0$ because $\tilde{C}$ (and therefore $\bar{P}_k \tilde{C} \bar{P}_k$) is positive semidefinite. This completes the proof.

References

[1] Edo Liberty. Simple and deterministic matrix sketching. In *The 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD 2013, Chicago, IL, USA, August 11-14, 2013*, pages 581–588. ACM, 2013.

[2] Mina Ghashami and Jeff M. Phillips. Relative errors for deterministic low-rank matrix approximations. In Chandra Chekuri, editor, *Proceedings of the Twenty-Fifth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2014, Portland, Oregon, USA, January 5-7, 2014*, pages 707–717. SIAM, 2014.