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 simple 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 its covariance matrix. Frequent Directions [1] maintains a rank deficient approximate covariance matrix \( \tilde{C} \in \mathbb{R}^{d \times d} \) using Algorithm 1. Set \( \tilde{C}_0 \in \mathbb{R}^{d \times d} \) to be the all zeros matrix. Then, at time \( t = 1, \ldots, T \) compute \( \tilde{C}_t = \text{UPDATE}(\tilde{C}_{t-1}, X_t, \ell) \).

Algorithm 1 Frequent Directions (FD) Update

1: function UPDATE(\( \tilde{C}_{t-1}, X_t, \ell \))
2: \( U_t \Lambda_t U_t^T = \tilde{C}_{t-1} + X_t X_t^T \)
3: return \( \tilde{C}_t = U \cdot \max(\Lambda - I \cdot \lambda^T, 0) \cdot U^T \)
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 \( \lambda_i \) is the \( i \)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 \( \lambda_i \) be the eigenvalues of the exact covariance \( C \) in descending order. For any \( \ell \) and simultaneously for all \( k < \ell \) we have

\[
\| C - \tilde{C} \| \leq \frac{1}{\ell - k} \sum_{i=k+1}^{d} \lambda_i
\]

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

\[
\| C - \tilde{C} \| = \| \sum_{t=1}^{T} \Delta_t \| \leq \sum_{t=1}^{T} \| \Delta_t \| \\
\leq \frac{1}{\ell - k} \text{tr} (\tilde{P}_k \left( \sum_{t=1}^{T} \Delta_t \right) \tilde{P}_k) \\
\leq \frac{1}{\ell - k} \text{tr} (\tilde{P}_k C \tilde{P}_k) = \frac{1}{\ell - k} \sum_{i=k+1}^{d} \lambda_i
\]

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

References
