Simple proof for Frequent Directions

Edo Liberty*

Abstract

This paper provides a short proof of the main results in [1] and [2].

Proof of the main results

The main insight needed for producing the simplified proof is changing the item being sketched. We will consider sketching the *covariance* of a stream of matrices rather than the stream itself. The results turns out to be identical.

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 $\tilde{C}_t \in \mathbb{R}^{d \times d}$ as follows. Given $\tilde{C}_{t-1}$ and X_t compute $U_t \Lambda^t U_t^T = \tilde{C}_{t-1} + X_t X_t^T$. Let $\Delta_t = U_t \cdot \min(\Lambda^t, I \cdot \lambda_\ell^t) \cdot U_t^T$. Here, λ_ℓ^t is the ℓ'th largest eigenvalue of $\tilde{C}_{t-1} + X_t X_t^T$. Finally, set $\tilde{C}_0$ to be the all zeros matrix and let $\tilde{C}$ denote $\tilde{C}_T$. Set $\tilde{C}_t = \tilde{C}_{t-1} + X_t X_t^T - \Delta_t$. Note that $\tilde{C}_t = U_t \cdot \max(\Lambda^t - I \cdot \lambda_\ell^t, 0) \cdot U_t^T$ whose rank 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. Note that $\Delta_t = X_t X_t^T - \tilde{C}_t + \tilde{C}_{t-1}$ and so $\sum_{t=1}^{T} \Delta_t = \sum_{t=1}^{T} X_t X_t^T - \sum_{t=1}^{T} (\tilde{C}_t - \tilde{C}_{t-1}) = C - \tilde{C}$. Moreover, $\|\Delta_t\| \leq \frac{1}{\ell} \operatorname{tr}(\Delta_t)$ because the top ℓ eigenvalues of Δ_t are all equal to one another.

$$\|C - \tilde{C}\| = \|\sum_{t=1}^{T} \Delta_t\| \leq \sum_{t=1}^{T} \|\Delta_t\| \leq \frac{1}{\ell} \sum_{t=1}^{T} \operatorname{tr}(\Delta_t) = \frac{1}{\ell} \operatorname{tr}\left(\sum_{t=1}^{T} \Delta_t\right) = \frac{1}{\ell} \operatorname{tr}(C - \tilde{C}) \leq \frac{1}{\ell} \operatorname{tr}(C)$$

The last step is due to $\operatorname{tr}(\tilde{C_T}) \geq 0$ because $\tilde{C_T}$ is positive semidefinite. This completes the proof of claim 1 in [1]. For the main result in [2] note that $\|\Delta_t\| < \frac{1}{\ell - k} \operatorname{tr}(\bar{P}_k \Delta_t \bar{P}_k)$ for any projection $\bar{P}_k$ having a null space of dimension at most k. This is also because the top ℓ eigenvalues of Δ_t are all equal to one another.

$$\|C - \tilde{C}\| \leq \sum_{t=1}^{T} \|\Delta_t\| \leq \frac{1}{\ell - k} \operatorname{tr}\left(\bar{P}_k \left(\sum_{t=1}^{T} \Delta_t\right) \bar{P}_k\right) \leq \frac{1}{\ell - k} \operatorname{tr}\left(\bar{P}_k C \bar{P}_k\right)$$

Applying this using $\bar{P}_k$ which nullifying the top k eigenvalues of C_T gives $\|C - \tilde{C}\| < \frac{1}{\ell - k} \sum_{i=k+1}^{d} \lambda_i$ where λ_i are the eigenvalues of C sorted in descending order. This proves the main result in [2].

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.

*Pinecone