Sparse Low Rank Approximation in a Stream

Let $\mathbf{A} \in \mathbb{R}^{n \times d}$. Recall the usual low rank approximation problem:

$$\min_{\mathbf{X}} \|\mathbf{A} - \mathbf{X}\|_F^2. \tag{1}$$

where $\mathbf{X}$ ranges over rank k matrices. However, $\mathbf{X}$ returned by the above minimization problem can be dense, and it is often desirable to find a *sparse* $\mathbf{X}$. In the *sparse low rank approximation* problem, we seek to minimize Equation (1) over $\mathbf{X}$ which can be written as the sum of k rank 1 matrices, each of which is supported on a $s \times s$ submatrix. We write $\mathcal{S}_{s,k}$ for the set of such matrices.

It can be shown that solving the above problem exactly is NP hard! For polynomial time algorithms, we allow for the following relaxations:

- We only seek $(1+\varepsilon)$-approximations
- We allow for the matrix $\mathbf{X}$ to be supported on an $O(sk/\varepsilon) \times O(sk/\varepsilon)$ submatrix

That is, we seek a rank k matrix $\mathbf{Y}$ supported on a $O(sk/\varepsilon) \times O(sk/\varepsilon)$ submatrix such that

$$\|\mathbf{A} - \mathbf{Y}\|_F^2 \leq (1+\varepsilon) \min_{\mathbf{X} \in \mathcal{S}_{s,k}} \|\mathbf{A} - \mathbf{X}\|_F^2$$

Our approach is as follows:

- Identify $O(sk/\varepsilon)$ heavy rows and columns
- Obtain an entry-wise estimate $\hat{\mathbf{A}}$ to $\mathbf{A}$, and restrict it to the heavy rows and columns
- Output a rank k approximation of $\hat{\mathbf{A}}$ on the restriction

Let $\mathbf{B}$ an optimal matrix in $\mathcal{S}_{s,k}$, let W be an $sk \times sk$ support containing $\mathbf{B}$, and let X be the part of W lying in columns of squared norm at most τ.

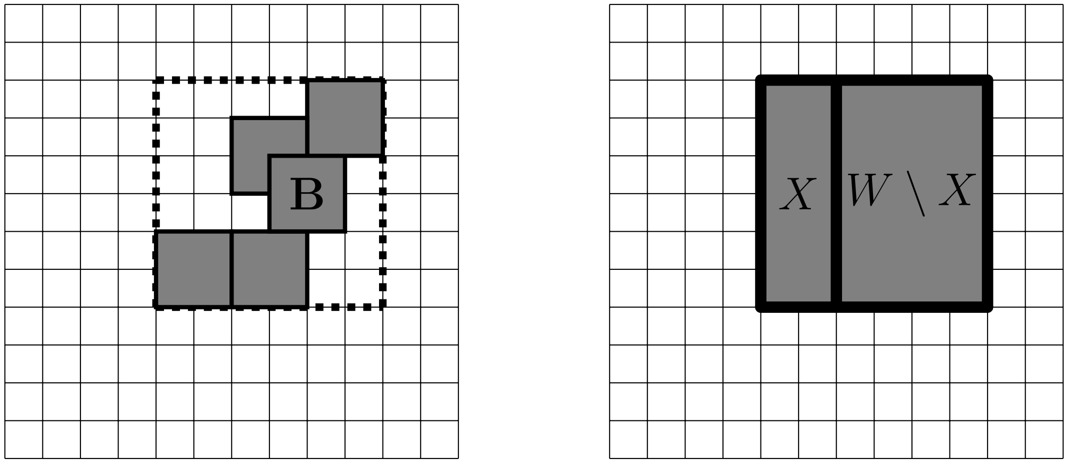

Figure 1: The supports $W, X, W \setminus X$ to be used in the proof.

Exercise. Show that

$$\|\mathbf{A} - (\mathbf{A}\mid_W)_k\|_F^2 \leq \|\mathbf{A} - \mathbf{B}\|_F^2.$$

Exercise. Show that

$$\|\mathbf{A} - (\mathbf{A}\mid_W)_k\|_F^2 = \left\|\mathbf{A} - (\mathbf{A}\mid_{W\setminus X})_k\right\|_F^2 - \|\mathbf{A}\mid_X\|_F^2 + \|\mathbf{A}\mid_W - (\mathbf{A}\mid_W)_k\|_F^2 - \left\|\mathbf{A}\mid_{W\setminus X} - (\mathbf{A}\mid_{W\setminus X})_k\right\|_F^2$$

Exercise. Show that

$$\left\| \mathbf{A} \mid_{W \setminus X} - (\mathbf{A} \mid_{W \setminus X})_k \right\|_F^2 \leq \| \mathbf{A} \mid_W - (\mathbf{A} \mid_W)_k \|_F^2$$

Exercise. Conclude that there exists a rank k matrix $\mathbf{C}$ supported on columns of squared norm at least τ such that

$$\|\mathbf{A}-\mathbf{C}\|_F^2 \leq \min_{\mathbf{X}\in\mathcal{S}_{s,k}} \|\mathbf{A}-\mathbf{X}\| + sk\tau$$

Fact. By a simple modification of the analysis of CountSketch, one can identify the rows $i \in [n]$ such that

$$\left\|\mathbf{e}_i^\top \mathbf{A}\right\|_2^2 \geq \alpha \left\|\mathbf{A}_{\overline{[1/\alpha]},*}\right\|_F^2$$

with $O(\alpha^{-1}\log^2 n)$ bits of space.

Exercise. Show how to identify a $O(sk/\varepsilon) \times O(sk/\varepsilon)$ submatrix $S \times T$ that contains a rank k matrix $\mathbf{C}$ such that

$$\|\mathbf{A} - \mathbf{C}\|_F^2 \leq (1+\varepsilon) \min_{\mathbf{X} \in \mathcal{S}_{s,k}} \|\mathbf{A} - \mathbf{X}\|_F^2$$

using $O(\varepsilon^{-1} sk \log^2 n)$ bits of space.

Exercise. Show how to recover a matrix $\hat{\mathbf{A}}$ such that

$$\left\|\hat{\mathbf{A}}\mid_{S\times T}-\mathbf{A}\mid_{S\times T}\right\|_\infty^2\le\frac{\varepsilon^4}{s^2k^2}\cdot\min_{\mathbf{X}\in\mathcal{S}_{s,k}}\|\mathbf{A}-\mathbf{X}\|_F^2$$

using $O(\varepsilon^{-4}s^2k^2\log n)$ bits of space. Conclude that

$$\left\|\hat{\mathbf{A}}\mid_{S\times T}-\mathbf{A}\mid_{S\times T}\right\|_F^2\le\varepsilon^2\cdot\min_{\mathbf{X}\in\mathcal{S}_{s,k}}\|\mathbf{A}-\mathbf{X}\|_F^2$$

Exercise. Show that

$$\left\| \mathbf{A} - (\hat{\mathbf{A}} \mid_{S \times T})_k \right\|_F^2 \leq (1+\varepsilon) \min_{\mathbf{X} \in \mathcal{S}_{s,k}} \|\mathbf{A} - \mathbf{X}\|_F^2$$