Frequent Items:

Keep k counters. If all are non-zero, subtract one from all of them. (for distinct elements).

[If elements exists increment counter; if not put element in empty space.]

Fact: each estimate $\le$ actual count and underestimate by at most n/k.

$\hat{f}_i \le f_i$ $\qquad$ $f_i - \frac{n}{k} \le \hat{f}_i \le f_i$.

BTW: suppose measure error by total loss. $\qquad (f_1 \ge f_2 \ge \cdots f_m)$

Optimal: underestimate $\boxed{OPT = \sum_{i=k+1}^{m} f_i}$

Suppose we maintain ℓ counters :- and then sps d decrements. then

$\sum$ counters $= \sum_{i=1}^{m} f_i - \ell(d)$. $= \sum_{i=1}^{k} \hat{f}_i \ge \sum_{i=1}^{k}(f_i - d) \Rightarrow d \le \frac{OPT}{\ell - k}$.

$\Rightarrow$ if we output the top k estimates then $\hat{f}_1 \ge \hat{f}_2 \cdots \ge \hat{f}_k$

$\Rightarrow$ total loss $= \sum_{i=1}^{m} f_i - (\sum_{i=1}^{k} \hat{f}_i) \le \sum_{i=1}^{m} f_i - (\sum_{i=1}^{k}(f_i - d))$

~~$= OPT(1 + kd)$~~ $\le OPT + \frac{k\, OPT}{\ell - k}$

$= OPT\left(\frac{\ell}{\ell - k}\right)$.

$\Rightarrow$ if $\ell = O(k/\epsilon)$ then loss $= (1 + O(\epsilon))\, OPT$.

So if we maintain ℓ different counters and outputs the best k, loss is small compared to OPT.

Similar to: suppose we find a matrix B that's a good approximation and $P_{B,k}$ be the projection onto the top k right singular vectors, then

want $\|A - AP_{B,k}\|_F^2 \le (1+\epsilon)\|A - A_k\|_F^2$. $\qquad$ {if $B = U\Sigma V^T$ then $P_{B,k} = V_k V_k^T$

Would be cool if we could get such a guarantee.

Liberty, et al:

Input: $A, K, \ell = K/\epsilon$. $A = \mathbb{R}^{n\times d}$

Start with $B \leftarrow 0^{\ell \times d}$. // invariant: B_ℓ is all zeros vector.

When a_t arrives $B_\ell \leftarrow a_t$.

let $B = U\Sigma V^T$ (assume $\Sigma \in \mathbb{R}^{d\times d}$. pad with zeros as necessary).

$C = \Sigma V^T$, $\delta_t \leftarrow \sigma_\ell^2$. (last singular value squared).

$B \leftarrow (\sqrt{\Sigma^2 - \delta I}) \cdot V^T$

Return B.

Let $B_{(t)}$ $C_{(t)}$, $\delta_{(t)}$ be these quantities at end of iteration t. $\Delta = \sum_t \delta_t$ is total "loss".

Property 1: $\forall x$ $0 \le \|Ax\|^2 - \|Bx\|^2 \le \Delta$. (each counter is off by at most d)

Property 2: $\|A\|_F^2 - \|B\|_F^2 \ge \Delta \ell$ (total mass lost $= d\ell$)

Thm 1.1: $\forall x$ $0 \le \|Ax\|^2 - \|Bx\|^2 \le \frac{\|A - A_k\|_F^2}{(\ell - k)}$

$\Leftrightarrow$ $0 \le x^T(A^TA - B^TB)x \le$ — $\Leftrightarrow$ $\|A^TA - B^TB\| \le \frac{\|A-A_k\|_F^2}{\ell - k}$.

Pf: $\Delta \ell \le \|A\|_F^2 - \|B\|_F^2$ (y_i = right singular vector of A) d-dimensional

$= \sum_{i=1}^{k} |Ay_i|^2 + \sum_{i=k+1}^{d} |Ay_i|^2 - \|B\|_F^2$

$= \sum_i |Ay_i|^2 + \|A - A_k\|_F^2 - \|B\|_F^2$

$\le \|A - A_k\|_F^2 + \sum_{i=1}^{k} (|Ay_i|^2 - |By_i|^2) \le \|A - A_k\|_F^2 + k\Delta$. (by Ppty #1.)

$\Rightarrow \Delta \le \frac{\|A-A_k\|_F^2}{(\ell - k)}$.

$\qquad\quad$ VI

$0 \le \|Ax\|^2 - \|Bx\|^2$

∎

Thm 1·2 : Let $B = U\Sigma V^T$ and V_k = top k singular values (right) of B.

$$P = V_k V_k^T$$

then $\|A - AP\|_F^2 \le \left(1 + \frac{k}{\ell - k}\right) \|A - A_k\|_F^2$.

$y_i \leftarrow$ right sing. of A

$v_i \leftarrow$ right sing of B.

Pf: $\|A - AP\|_F^2 = \|A\|_F^2 - \|AP\|_F^2$ (Pythagoras)

$$= \|A\|_F^2 - \sum_{i=1}^{k} |Av_i|^2$$

$$\le \|A\|_F^2 - \sum_{i=1}^{k} |Bv_i|^2 \quad \text{(Ppty 1)}.$$

$$\le \|A\|_F^2 - \sum_{i=1}^{k} |By_i|^2 \quad \text{(Eckert young)}$$

$$\le \|A\|_F^2 - \left(\sum_{i=1}^{k} |Ay_i|^2 - \Delta\right) \quad \text{(ppty 1)}$$

$$= \|A - A_k\|^2 + \Delta k \le \|A - A_k\|^2 \left(1 + \frac{k}{\ell - k}\right).$$

(by Thm 1.1.) ▨

(Recall)

$$\text{error} = \sum_i f_i - \sum_{i=1}^{k} \hat{f}_i \le \sum_i f_i - \left(\sum_{i=1}^{k} f_i - d\right) = OPT + dk. \le OPT\left(1 + \frac{k}{\ell - k}\right).$$