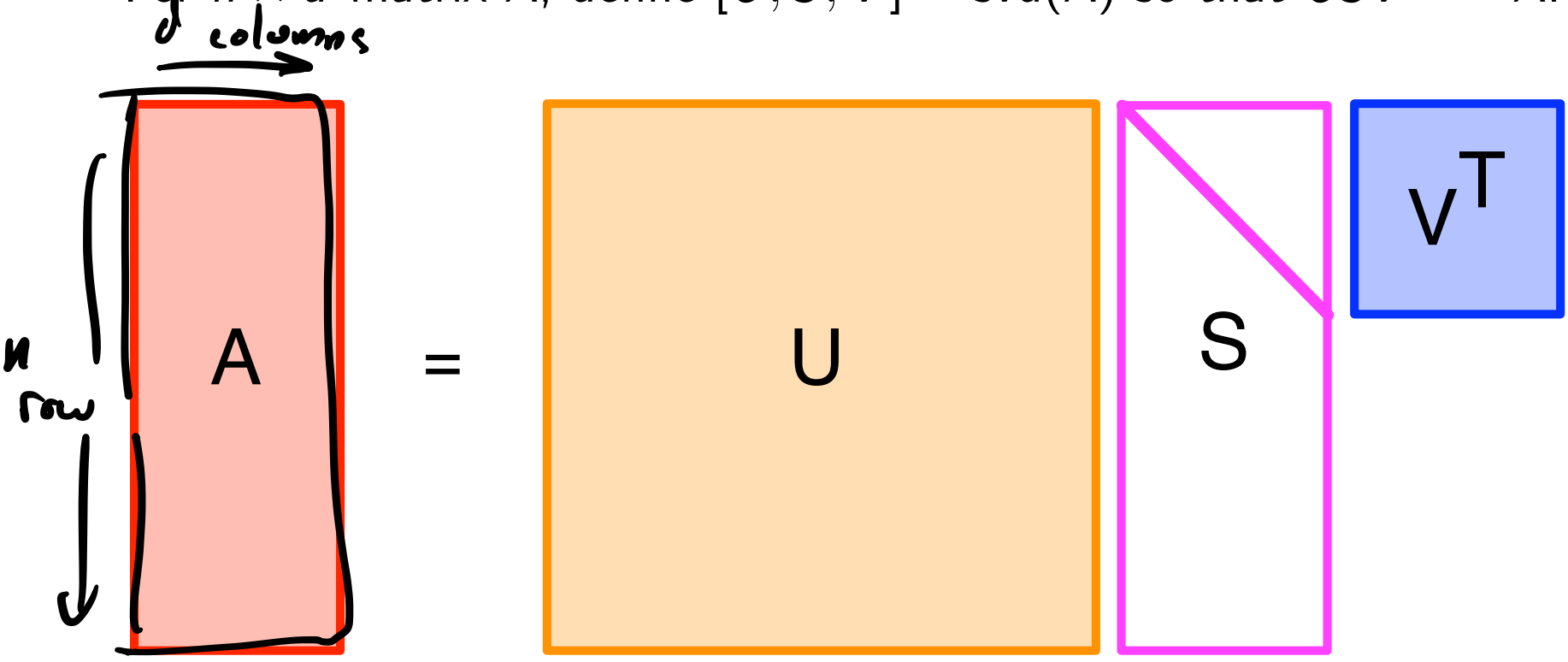
L16: Matrix Sketching

Jeff M. Phillips

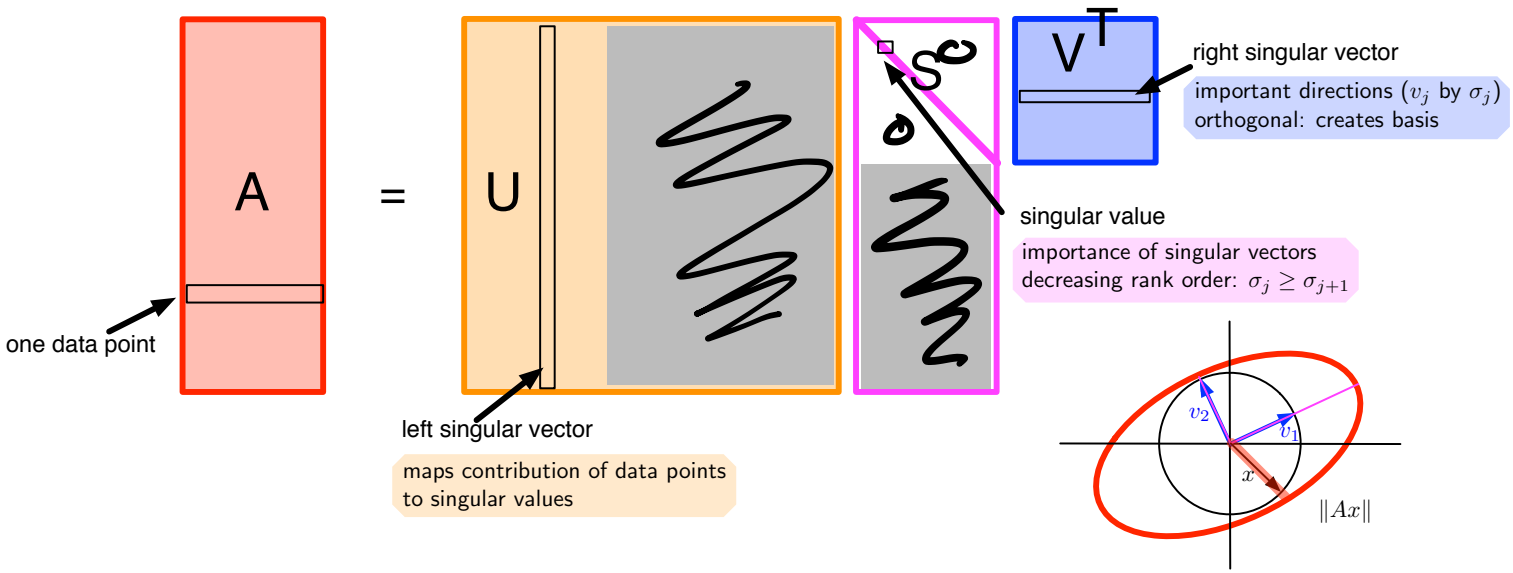
March 28, 2018

Singular Value Decomposition

For $n \times d$ matrix A, define $[U, S, V] = \text{svd}(A)$ so that $USV^T = A$.

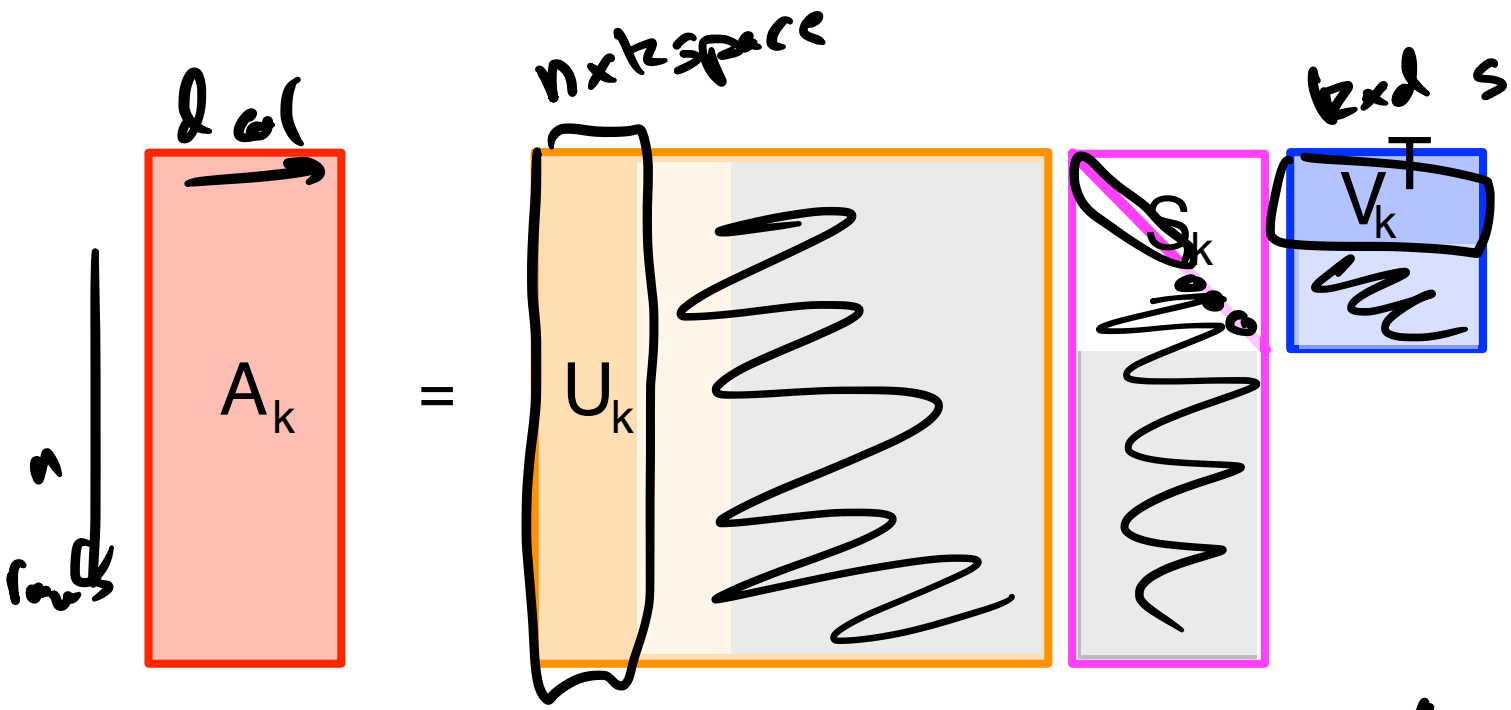


Singular Value Decomposition



Best Rank k-Approximation

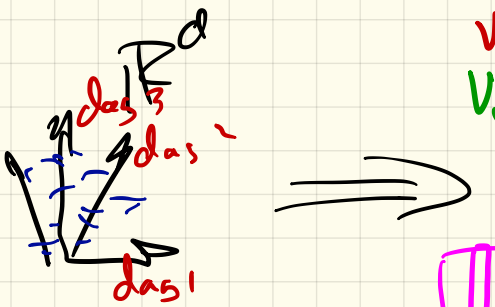
$A = (a_1, a_2 \dots a_n)$ a_i: row of A

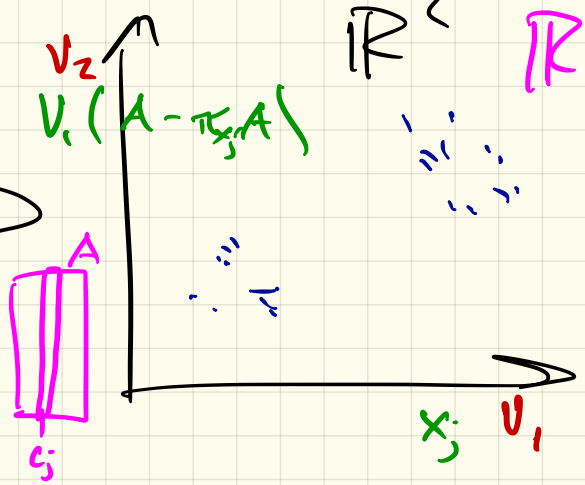


Take $O(nd^2)$ time (when $n > d$)

↳ Runtime $O(n \cdot d \cdot k)$ Approximate work in stream

If use SVD in PCA





$\mathbb{R}^2$ $\mathbb{R}^k$

using k original dimensions

$k = O\left(\frac{k'}{\varepsilon^2} \log \frac{1}{\delta}\right)$

columns A

$c_1, c_2, \ldots c_d$

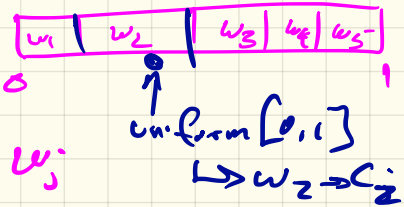
c_j weight $w_j = \|c_j\|^2 = \sum_i A_{ij}^2$

v_1 means $x_1 x_1 + x_2 x_2 \ldots x_d x_d$

$\alpha_j = \langle v_1, x_j \rangle$

select $C = \{c_1', c_2', \ldots c_{k'}'\}$

c_j' chosen at random iid proportionally to w_j



Reservior Sampling

Streams $A = \langle a_1, a_2, \ldots a_n \rangle$

$\langle w_1, w_2 \ldots w_n \rangle \quad [w_j = 1]$

Maintain a random sample.

r = random point from $A_i = \langle a_1, a_2, \ldots a_i \rangle$

step i:

- $w_{i+1} > \frac{1}{i+1}$ replace r w/ a_{i+1}
- otherwise keep $r = r$.

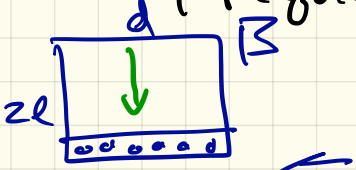
$$C = [c_1, c_2, \ldots, c_k]$$

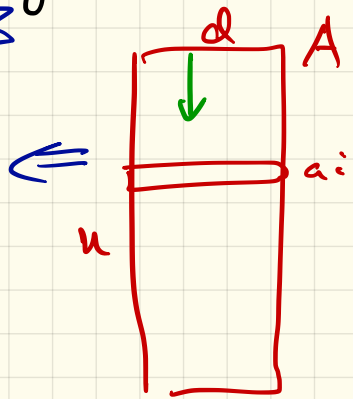
$$\Pi_C = C (C^T C)^{-1} C^T$$

$\Pi_C A$ $\qquad k = O\left(\frac{k}{\varepsilon^2}\right)$

$$\| A - \Pi_C A \|_F \leq \| A - A_k \|_F + \varepsilon \| A \|_F$$

Frequent Directions Alg (Matrix version MG alg)

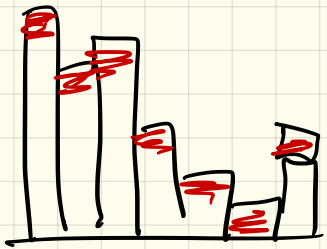




$A = [a_1, a_2 \dots$

a_i row A

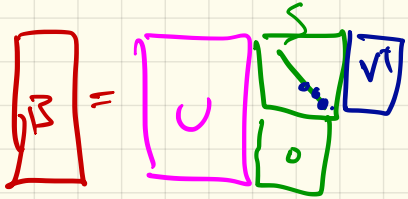
$a_i \in \mathbb{R}^d$



$\ell = k/\varepsilon$

Maintain Invariant

B has 1 all zeros row



0. Put first $2\ell-1$ rows of A in B

for $a_i \in A$ (in stream)

1. Put a_i in empty row of B
2. $[U, S, V] = \text{svd}(B)$ $O(d\ell^2)$
3. Set $\delta = \sigma_\ell^2$ ← diff sing. val B.
4. $S' = \text{diag}(\sqrt{\sigma_1^2 - \delta}, \sqrt{\sigma_2^2 - \delta}, \dots \sqrt{\sigma_{\ell-1}^2 - \delta}, 0, 0 \dots)$
5. $B = S'V^T$

only if no empty rows

FD $\ell = k/\varepsilon$ rows in B

① $\forall$ any unit vector $x \in \mathbb{R}^d$

$\sum_{j=k+1}^{d} \sigma_j^2$

$$0 \le \underbrace{\|Ax\|^2 - \|Bx\|^2}_{\|A^TA - B^TB\|_2} \le \frac{\|A - A_k\|_F^2}{\ell - k}$$

set $\ell = k + k/\varepsilon$

$$\le (1+\varepsilon)\|A - A_k\|_F^2$$

② $$\|A - A\pi_{B_k}\|_F^2 \le \frac{\ell}{\ell - k}\|A - A_k\|_F^2$$

$\ell = k + k/\varepsilon$

$$\le (1+\varepsilon)\|A - A_k\|_F^2$$

Priority Sampling

$A = \langle a_1, a_2 \ldots a_n \rangle$

$w_1 \; w_2 \ldots w_n$

$l_1 \; l_2 \; \ldots \; l_n$

without replacement

unbiased.

$$l_i = \frac{w_i}{v_i} \qquad v_i \in \text{Uniform}[0,1]$$

Keep top k priorities. $l_1', l_2' \ldots l_k'$

priority Queue

log k updates