Scientific Computing HW 6

Ryan Chen

October 9, 2024

Note: For sums we assume the lower bound to be 1 unless otherwise specified, and we will not specify upper bounds since they are determined by the sizes of our matrices.

1. (a) Consider the following.

$$\sum_i \sum_j \sum_k a_{ik} b_{kj} c_{ji} = \sum_i \sum_j \sum_k b_{kj} c_{ji} a_{ik} = \sum_i \sum_j \sum_k c_{ji} a_{ik} b_{kj}$$

The first, second, and third expressions are, respectively, $\operatorname{tr}(ABC)$, $\operatorname{tr}(BCA)$, and $\operatorname{tr}(CAB)$.

(b) Compute

$$\begin{aligned}
\|A\|_F^2 &= \sum_i \sum_j a_{ij}^2 \\
&= \operatorname{tr}\big(A^T A\big) \\
&= \operatorname{tr}\big(V \Sigma U^T U \Sigma V^T\big) && A = U\Sigma V^T \\
&= \operatorname{tr}\big(V \Sigma^2 V^T\big) \\
&= \operatorname{tr}\big(\Sigma^2 V^T V\big) && \text{cyclic property of trace} \\
&= \operatorname{tr}\big(\Sigma^2\big) \\
&= \sum_i \sigma_i^2
\end{aligned}$$

(c) Compute

$$\begin{aligned}
\|A + B\|_F^2 &= \sum_i \sum_j (a_{ij} + b_{ij})^2 \\
&= \sum_i \sum_j a_{ij}^2 + \sum_i \sum_j b_{ij}^2 + 2 \sum_i \sum_j a_{ij} b_{ij} \\
&= \|A\|_F^2 + \|B\|_F^2 + 2 \langle A, B \rangle_F
\end{aligned}$$

2. Observe that

$$A_k = U_k \Sigma_k V_k^T = U \Sigma_k' V^T$$

where Σ_k' is obtained by adding zeros to Σ_k to make it the same size as Σ. Then

$$A - A_k = U(\Sigma - \Sigma_k')V^T$$

is an SVD of $A - A_k$ with the jth diagonal entry of $\Sigma - \Sigma_k'$ being 0 for $j \leq k$ and $\sigma_j(A)$ for $j > k$. Denoting the Ky Fan $p-$norm by $\|\cdot\|_{KF(p)}$,

$$\|A - A_k\|_{KF(p)}^p = \sum_j \sigma_j^p(A - A_k) = \sum_{j \geq k+1} \sigma_j^p(A)$$

Fix a matrix M with $\operatorname{rank} M \leq k$. By Lemma 1 in Section 4.3 of the lecture notes,

$$\sigma_{k+i}(A) \leq \sigma_i(A - M) + \sigma_{k+1}(M) = \sigma_i(A - M)$$

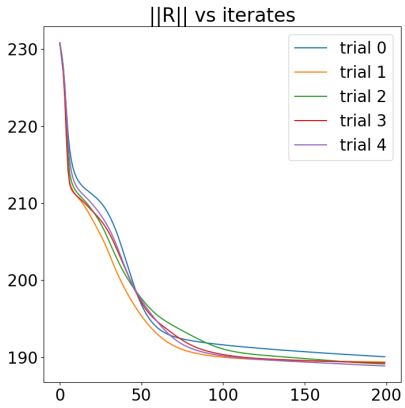
Then we have

$$\begin{aligned}
\|A - M\|_{KF(p)}^p &= \sum_i \sigma_i^p(A - M) \\
&\geq \sum_i \sigma_{k+i}^p(A) && \sigma_{k+i}(A) \leq \sigma_i(A - M) \\
&= \sum_{j \geq k+1} \sigma_j^p(A) && \text{change variables } j := k + i \\
&= \|A - A_k\|_{KF(p)}^p
\end{aligned}$$

Taking the pth root of both sides gives $\|A - M\|_{KF(p)} \geq \|A - A_k\|_{KF(p)}$.

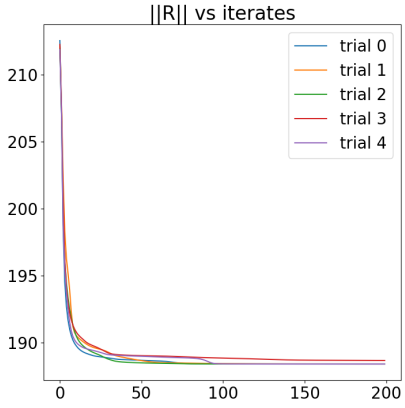
3. Code: `https://github.com/RokettoJanpu/scientific-computing-1-redux/blob/main/hw6.ipynb`

   (a) Here we run PGD five times.



   The residual norm settles at around 80 iterations. Its eventual value is about 190 in each trial. Examining entries of W greater than 0.2, we hypothesize that the documents are advertizements of events and services in various communities across the country.

   (b) Here we run HALS five times.



   The residual norm settles at around 20 iterations. Its eventual value is about 188 in each trial.

   (c) Using the SVD of A, we find $\|A - A_{10}\|_F \approx 187.72$, slightly below the residual norms of both PGD and HALS, as expected by the Eckart–Young–Mirsky theorem.