lin777@wisc.edu Yizhou Lin

1. a) We want to find SVD of $\mathcal{E}_r$,

$$\mathcal{E}_r = X - X_r = \sum_{i=1}^{n} \sigma_i u_i v_i^T - \sum_{i=1}^{r} \sigma_i u_i v_i^T$$

$$= \sum_{i=r}^{n-r} \sigma_i u_i v_i^T$$

b). Since X is full rank, $\operatorname{rank}\{\mathcal{E}_r\} = \operatorname{rank}\{X\} =$ number of σ_i, equals to ~~n~~. $\underline{p - r}$

c). $\|E_r\|_{op} = \left(\sum_{i=r}^{n-r} \sigma_i u_i v_i^T \right)^2$

d) From the questions above, X_r will be a good approximation to X iff the operator norm, i.e., $\|E_r\|_{op}$ is minimum.

2. a) Approximate rank of A is accurate.

The reason to plot log (singular value) is that s value is too large and log can efficiently solve this problem while at the same time, maintains the variability.

b). As the r increases, the approximation of zealey is ~~larger~~ better, since we got more singular vector to tech

c). The storage space for the original matrix, i.e. the one without the SVD process, is about, 10^2, 20^2, 50^2, 100x larger, in terms of storage space, than the ones with

$r \in \{10, 20, 50, 100\}$.

d. i) When $\sigma_g^2 = 10$, rank is approximately 80.

ii) When $\sigma_g^2 = 50$, rank is approximately 25.