AI3602 Data Mining: Homework 2

Xiangyuan Xue (521030910387)

1. Let $\|\cdot\|_F$ denote the Frobenius norm of a matrix, $\sigma_k(\cdot)$ denote the k-th largest singular value of a matrix, and $\lambda_k(\cdot)$ denote the k-th largest eigenvalue of a matrix. For any matrix M, let M_k denote the rank-k approximation of M given by the method of singular value decomposition as is described. We have the following lemmas and corollaries.

Lemma 1. Frobenius norm of a matrix is orthogonally invariant, namely, for any matrix M and any orthogonal matrix O, it holds that

$$\|M\|_F = \|OM\|_F$$

Proof. Notice that $\|M\|_F^2 = \mathrm{tr}(M^T M)$, we have

$$\|OM\|_F^2 = \mathrm{tr}\left(M^T O^T O M\right) = \mathrm{tr}\left(M^T M\right) = \|M\|_F^2$$

which verifies the lemma.

Corollary 1. For any matrix M with rank r, it holds that

$$\|M\|_F^2 = \sum_{k=1}^{r} \sigma_k^2(M)$$

Proof. Let $M = U\Sigma V^T$ be the singular value decomposition of M, where U and V are orthogonal matrices and Σ is a diagonal matrix with singular values on the diagonal. Notice that orthogonal transformations preserve the Frobenius norm, we have

$$\|M\|_F^2 = \left\|U\Sigma V^T\right\|_F^2 = \|\Sigma\|_F^2 = \sum_{k=1}^{r} \sigma_k^2(M)$$

which verifies the corollary.

Lemma 2. For any matrix M, M' and M'' such that $M = M' + M''$, it holds that

$$\sigma_1(M') + \sigma_1(M'') \geq \sigma_1(M)$$

Proof. Notice that $\sigma_1(M) = \sqrt{\lambda_1(M^T M)} = \|M\|_2$, by the triangle inequality

$$\sigma_1(M') + \sigma_1(M'') = \|M'\|_2 + \|M''\|_2 \geq \|M\|_2 = \sigma_1(M)$$

which verifies the lemma.

Corollary 2. For any matrix M, M' and M'' such that $M = M' + M''$, it holds that

$$\sigma_i(M') + \sigma_j(M'') \geq \sigma_{i+j-1}(M)$$

where $1 \le i \le \text{rank}(M')$, $1 \le j \le \text{rank}(M'')$ and $1 \le i + j - 1 \le \text{rank}(M)$.

Proof. Notice that $\sigma_{k+1}(M) = \sigma_1(M - M_k)$, we have

$$\begin{aligned}
\sigma_i(M') + \sigma_j(M'') &= \sigma_1(M' - M'_{i-1}) + \sigma_1(M'' - M''_{j-1}) \\
&\ge \sigma_1(M - M'_{i-1} - M''_{j-1}) \\
&\ge \sigma_1(M - M_{i+j-2}) \\
&= \sigma_{i+j-1}(M)
\end{aligned}$$

which verifies the corollary.

Now we are ready to prove that A_k is the best rank-k approximation of A in terms of Frobenius norm error. Let u_i and v_i denote the i-th column of U and V respectively. The approximation error is given by

$$\begin{aligned}
\|A - A_k\|_F^2 = \left\| \sum_{i=k+1}^{n} \sigma_i u_i v_i^T \right\|_F^2 &= \text{tr}\left[\left(\sum_{i=k+1}^{n} \sigma_i v_i u_i^T \right) \left(\sum_{i=k+1}^{n} \sigma_i u_i v_i^T \right) \right] \\
&= \text{tr}\left(\sum_{i=k+1}^{n} \sigma_i^2 v_i u_i^T u_i v_i^T \right) = \sum_{i=k+1}^{n} \text{tr}\left(\sigma_i^2 v_i v_i^T \right) = \sum_{i=k+1}^{n} \sigma_i^2
\end{aligned}$$

Suppose P_k is another rank-k approximation of A, it holds that $A = (A - P_k) + P_k$ and $\sigma_{k+1}(P_k) = 0$. By the second corollary, we have

$$\sigma_i(A - P_k) = \sigma_i(A - P_k) + \sigma_{k+1}(P_k) \ge \sigma_{i+k}(A)$$

which holds for any $1 \le i \le n - k$. By the first corollary, we have

$$\|A - P_k\|_F^2 = \sum_{i=1}^{n} \sigma_i^2(A - P_k) \ge \sum_{i=k+1}^{n} \sigma_i^2(A) = \sum_{i=k+1}^{n} \sigma_i^2$$

which implies $\|A - P_k\|_F \ge \|A - A_k\|_F$ holds for any P_k. Therefore, we conclude that A_k is the best rank-k approximation of A in terms of Frobenius norm error.