Quiz03

1. **Best Approximations**

We start by recalling the Eckart-Young Theorem: Consider any square matrix $C \in \mathbb{R}^{n\times n}$ and assume that we can write its full singular value decomposition as $C = U\Sigma V^\top$ where Σ is the $n \times n$ diagonal matrix with distinct diagonal entries $\sigma_1 > \sigma_2 > \cdots > \sigma_n > 0$. Then, for $0 \leq k \leq n$, the Eckart-Young Theorem for Frobenius norm states that

$$C_k = U_k \Sigma_k V_k^\top = \underset{\substack{B \in \mathbb{R}^{n\times n} \\ \text{rank}(B) \leq k}}{\text{argmin}} \|C - B\|_F$$

where $U_k \in \mathbb{R}^{n\times k}$ is the matrix consisting of the first k columns of U, $V_k \in \mathbb{R}^{n\times k}$ is the matrix consisting of the first k columns of V, and Σ_k is the $k \times k$ diagonal matrix with the top-k singular values $\sigma_1 > \cdots > \sigma_k$ as its diagonal entries.

(a) Now, consider any square matrix $C \in \mathbb{R}^{n\times n}$ and let $C_k \in \mathbb{R}^{n\times n}$ denote its best rank-k approximation in the Frobenius norm. Then, for any orthonormal matrix $W \in \mathbb{R}^{n\times n}$, show that

$$WC_kW^\top = \underset{\substack{B \in \mathbb{R}^{n\times n} \\ \text{rank}(B) \leq k}}{\text{argmin}} \|WCW^\top - B\|_F.$$

Solution: Using the fact that the Frobenius norm is invariant under orthogonal transformations, we have $\|B - WCW^\top\|_F = \|W^\top BW - C\|_F$. Therefore, we can write

$$\min_{\substack{B \in \mathbb{R}^{n\times n} \\ \text{rank}(B) \leq k}} \|B - WCW^\top\|_F = \min_{\substack{B \in \mathbb{R}^{n\times n} \\ \text{rank}(B) \leq k}} \|W^\top BW - C\|_F.$$

Next, we show that $\{W^\top BW : B \in \mathbb{R}^{n\times n}, \text{rank}(B) \leq k\} = \{Z \in \mathbb{R}^{n\times n} : \text{rank}(Z) \leq k\}$. For any $B \in \mathbb{R}^{n\times n}$ such that $\text{rank}(B) \leq k$, $Z = W^\top BW$ satisfies $\text{rank}(Z) \leq k$. On the other hand, for any $Z \in \mathbb{R}^{n\times n}$ such that $\text{rank}(Z) \leq k$, there exists $B = WZW^\top$ that satisfies $\text{rank}(B) \leq k$ and $W^\top BW = W^\top WZW^\top W = Z$ because W is an orthonormal matrix. Therefore,

$$\begin{aligned}\min_{\substack{B \in \mathbb{R}^{n\times n} \\ \text{rank}(B) \leq k}} \|W^\top BW - C\|_F &= \min_{\substack{Z \in \mathbb{R}^{n\times n} \\ \text{rank}(Z) \leq k}} \|Z - C\|_F \\ &= \|C_k - C\|_F.\end{aligned}$$

Lastly, noting that $\|WC_kW^\top - WCW^\top\|_F = \|W(C_k - C)W^\top\|_F = \|C_k - C\|_F$, we can write

$$\|WC_kW^\top - WCW^\top\|_F = \min_{\substack{B \in \mathbb{R}^{n\times n} \\ \text{rank}(B) \leq k}} \|B - WCW^\top\|_F.$$

Since $\text{rank}(WC_kW^\top) \leq k$, we conclude that $WC_kW^\top$ is a minimizer of the given optimization problem.

For the remainder of this problem, we consider a square matrix $A \in \mathbb{R}^{n\times n}$ and assume A has full rank. Using Gram-Schmidt Orthonormalization (GSO), we can write the matrix A as

$$A = QR \tag{1}$$

where $Q \in \mathbb{R}^{n\times n}$ is an orthonormal matrix and $R \in \mathbb{R}^{n\times n}$ is an upper triangular matrix.

(b) Recall that $A \in \mathbb{R}^{n \times n}$ is a square matrix with full rank, and we can write the matrix A as

$$A = QR \tag{2}$$

where $Q \in \mathbb{R}^{n \times n}$ is an orthonormal matrix and $R \in \mathbb{R}^{n \times n}$ is an upper triangular matrix.

Assume that $R = \text{diag}(r_1, r_2, \ldots, r_n) \in \mathbb{R}^{n \times n}$ is a diagonal matrix with

$$|r_1| > |r_2| > \cdots > |r_n|,$$

and all $r_1, r_2, \ldots, r_n$ are real numbers. Let $k < n$. Then, show that the best rank-k approximation to $AA^\top$ in the Frobenius norm is $QSQ^\top$, where S is a diagonal matrix defined as

$$S = \text{diag}(r_1^2, \ldots, r_k^2, 0, \ldots, 0) \in \mathbb{R}^{n \times n}.$$

Solution: We can write $AA^\top = QRR^\top Q^\top$. Note that Q is an orthonormal matrix and $RR^\top = \text{diag}(r_1^2, r_2^2, \ldots, r_n^2)$ is a diagonal matrix. Therefore, $AA^\top = QRR^\top Q^\top$ forms an SVD for $AA^\top$. Then, the result follows by the Eckart-Young Theorem.