18.06 Problem Set 9 - SOLUTIONS

Problem 1. Let $\sigma_{\max}(A)$ be the largest singular value of a matrix A. Show that $\sigma_{\max}(A^{-1})\sigma_{\max}(A) \geq 1$ for any square invertible matrix A.

Let A be an invertible $n \times n$ square matrix. Then the singular values of A are the square roots of the eigenvalues of AA^T or equivalently A^TA. Since $A^{-1}(A^{-1})^T = (A^TA)^{-1}$, the singular values of A^{-1} are the reciprocals of the singular values of A. Therefore, if σ_1 is the largest singular value of A and σ_n is the smallest singular value of A, $\sigma_{\max}(A^{-1})\sigma_{\max}(A) = \sigma_1/\sigma_n$. The known inequality $\sigma_1 \geq \sigma_n$ then yields the desired inequality $\sigma_{\max}(A^{-1})\sigma_{\max}(A) \geq 1$.

Problem 2. Suppose A has orthogonal columns $w_1, w_2, ..., w_n$ of lengths $\sigma_1, \sigma_2, ..., \sigma_n$. What are U, Σ, and V in the SVD?

The problem is slightly complicated (at least notationally) by allowing some lengths to be zero; in your solution you may assume that the σ_i are all nonzero (in that case, $k = n$ below). Suppose that the w_i are vectors in $\mathbb{R}^m$, so that A is a $m \times n$ matrix. Since A has pairwise orthogonal columns, A^TA is a $n \times n$ diagonal matrix with eigenvalues the squares of the lengths of the w_i, so the singular values of A are the nonzero σ_i's. Let k be the number of singular values. Let $f : \{1, ..., n\} \to \{1, ..., n\}$ be a permutation such that $\sigma_{f(1)} \geq ... \geq \sigma_{f(n)}$. Then Σ is a $m \times n$ matrix with $\Sigma_{ii} = \sigma_{f(i)}$ for $1 \leq i \leq k$ and 0 otherwise. V consists of the eigenvectors of A^TA, which are the standard basis vectors e_i, and we order the eigenvectors by the size of the corresponding eigenvalue. Therefore, V is a $n \times n$ permutation matrix with ith column equal to $e_{f(i)}$. We have $A = U\Sigma$, so U is the $m \times m$ matrix with first k columns $\frac{1}{\sigma_{f(1)}} w_{f(1)}, ..., \frac{1}{\sigma_{f(k)}} w_{f(k)}$, in that order, and the rest of the columns all 0.

If you do not follow the convention that the diagonal entries in Σ are listed in decreasing order, then we have a simpler SVD (which we also will accept as a valid solution). Namely, Σ is a $m \times n$ matrix with $\Sigma_{ii} = \sigma_i$, $V = I$ the $n \times n$ identity matrix, and U has columns $\frac{1}{\sigma_i} w_i$ (with the zero column if $\sigma_i = 0$).

Problem 3. If $A = QR$ with an orthogonal matrix Q, the SVD of A is almost the same as the SVD of R. Which of the three matrices U, Σ, V is changed because of Q?

We have $A^TA = (QR)^T(QR) = R^TR$ because $Q^TQ = I$. Thus the matrices Σ and V are unchanged. We have $AA^T = Q(RR^T)Q^T$, so if u is an eigenvector for AA^T, then Q^Tu is an eigenvector for RR^T. Thus the matrix U becomes Q^TU.

Problem 4. Let $n > 1$. Show that there is no n by n matrix A such that $AM = M^T$ for every n by n matrix M.

Let $M = I$. Then $AM = M^T$ forces $A = I$. However, we do not have $M = M^T$ for every matrix M, if $n > 1$. Therefore, no such A exists.

Problem 5. Let V be a vector space and $T : V \to V$ a linear transformation. Suppose that for every linear transformation $S : V \to V$, we have $S(T(v)) = T(S(v))$ for all vectors $v \in V$. Show that there exists a scalar c such that $T(v) = cv$ for all vectors $v \in V$.

Let $v_1, ..., v_n$ be a basis for V. For any vector $v \in V$, we may write $v = \sum_{i=1}^{n} a_i v_i$ for some $a_i \in \mathbb{R}$. Let's apply this to the vectors $T(v_i)$ to write $T(v_i) = a_{1i}v_1 + ... + a_{ni}v_n$.

Let $S_i : V \to V$ be the linear transformation defined by $S_i(v_i) = v_i$ and $S_i(v_j) = 0$ for $j \neq i$. Then for every i, the given equality $S_i(T(v_i)) = T(S_i(v_i))$ implies that $a_{ii}v_i = T(v_i) = a_{1i}v_1 + ... + a_{ni}v_n$, so $a_{ji} = 0$ for $j \neq i$. It remains to see that $a_{ii} = a_{jj}$ for all i, j. For this, let $S_{ij} : V \to V$ be the linear transformation defined by $S_{ij}(v_i) = v_j$, $S_{ij}(v_j) = v_i$, and $S_{ij}(v_k) = v_k$ for $k \neq i, j$ (actually, $S_{ij}(v_k)$ could be any vector for our purposes). Then $S_{ij}(T(v_i)) = a_{ii}v_j$ and $T(S_{ij}(v_i)) = a_{jj}v_j$, so $a_{ii} = a_{jj}$. Letting $c = a_{11}$, we conclude that $T(v) = cv$ for all vectors $v \in V$.