NLA sheet1 solutions for problems 8,9

Yuji Nakatsukasa

Problem 8 Let B be a square $n \times n$ matrix. Bound the ith singular values of AB using $\sigma_i(A)$ and $\sigma_i(B)$: Specifically, prove that for each i,

$$\sigma_i(A)\sigma_n(B) \le \sigma_i(AB) \le \sigma_i(A)\sigma_1(B).$$

(Solution:) By the Courant-Fisher theorem for singular values $\sigma_i(AB) = \max_{QQ^*=I_i} \min_{\|x\|=1} \|x^*QAB\|_2$. For any fixed vector $y^*(= x^*QA)$, we have $\|y^*B\|_2 \le \|y^*\|_2\|B\|_2$ (via C-F or directly via the SVD). Thus

$$\sigma_i(AB) = \max_{QQ^*=I_i} \min_{\|x\|=1} \|x^*QAB\|_2 \le \max_{QQ^*=I_i} \min_{\|x\|=1} \|x^*QA\|_2\|B\|_2 = \sigma_i(A)\sigma_1(B).$$

For the lower bound we use the fact $\sigma_{\min}(B) = \min_{\|x\|_2=1} \|x^*B\|_2$ (again via C-F or directly via the SVD), hence $\|x^*QAB\|_2 \ge \|x^*QA\|_2\sigma_{\min}(B)$ for any fixed x, to obtain

$$\sigma_i(AB) = \max_{QQ^*=I_i} \min_{\|x\|=1} \|x^*QAB\|_2 \ge \max_{QQ^*=I_i} \min_{\|x\|=1} \|x^*QA\|_2\sigma_{\min}(B) = \sigma_i(A)\sigma_n(B).$$

Problem 9 (optional; harder) Let $A \in \mathbb{R}^{m\times n}$, $m \ge n$ and $\sigma_i(A) \ge \sigma_i(A) \ge \cdots \ge \sigma_n(A) \ge 0$ be its singular values. Prove that for $k = 1, 2, \ldots, n$,

$$\sum_{i=1}^{k} \sigma_i(A) = \max_{Q^TQ=I_k, W^TW=I_k} \mathrm{trace}(Q^TAW).$$

($Q \in \mathbb{R}^{m\times k}, W \in \mathbb{R}^{n\times k}$ are orthonormal. Recall for an $k \times k$ matrix B, $\mathrm{trace}(B) = \sum_{i=1}^k B_{ii}$; a useful property is $\mathrm{trace}(CD) = \mathrm{trace}(DC)$ as long as CD is square.)

(solution:) Equality is seen to be attained when $Q = [u_1, u_2, \ldots, u_k], W = [v_1, \ldots, v_k]$, since then $\mathrm{trace}(Q^TAW) = \mathrm{trace}(\mathrm{diag}(\sigma_1(A), \ldots, \sigma_k(A))) = \sum_{i=1}^k \sigma_i(A)$. We need to prove this is an upper bound for $\mathrm{trace}(Q^TAW)$. First note that $\sigma_i(AB) \le \sigma_i(A)\|B\|$ holds for any A, B s.t. AB is defined (e.g. via Courant-Fisher). Now since Q, W are orthonormal, $\sigma_i(Q) = \sigma_i(W) = 1$ for all $i = 1, \ldots, k$. We thus have $\sigma_i(Q^TAW) \le \sigma_i(A)$ for all i. We are thus done if we prove $\mathrm{trace}(B) \le \sum_{i=1}^k \sigma_i(B)$ for any $k \times k$ matrix B. Let $B = U_B\Sigma_B V_B^T$ be the SVD. Then $\mathrm{trace}(B) = \mathrm{trace}(U_B\Sigma_B V_B^T) = \mathrm{trace}(\Sigma_B V_B^T U_B) = \sum_{i=1}^k \sigma_i(B)(V_B^TU_B)_{ii} \le \sum_{i=1}^k \sigma_i$, because $V_B^TU_B$ is orthogonal and so its entries are bounded by 1 in absolute value.