MATHEMATICAL TOOLS - SOLUTION 7

Problem 1. Take $A = \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix}$. First, 1 and 2 are both eigenvalues, so A is diagonalizable. On the other hand:

$$AA^T = \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 2 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 5 & 1 \\ 1 & 1 \end{pmatrix}$$

the eigenvalues of AA^T are $3 \pm \sqrt{5}$, which means that the singular values of A are $\sqrt{3 \pm \sqrt{5}}$ which are different than 1 and 2. Now, $\|A\|_{op} = \sqrt{3+\sqrt{5}} \approx 2.2 > 2$.

Problem 2. We'll start with A. It holds:

$$A^T A = \begin{pmatrix} 4 & 0 \\ 0 & 1 \end{pmatrix}$$

and so $\sigma_1 = 2, \sigma_2 = 1$. Therefore $D = \begin{pmatrix} 2 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}$. Next, we want to find $\begin{pmatrix} x \\ y \end{pmatrix}$ s.t. $\left\|A\begin{pmatrix} x \\ y \end{pmatrix}\right\|_2^2 = \sigma_1^2 \left\|\begin{pmatrix} x \\ y \end{pmatrix}\right\|_2^2 = 4\left\|\begin{pmatrix} x \\ y \end{pmatrix}\right\|_2^2$. Well:

$$\left\|A\begin{pmatrix} x \\ y \end{pmatrix}\right\|_2^2 = \left(-\frac{24}{25}x + \frac{4}{5}y\right)^2 + \left(-\frac{32}{25}x - \frac{3}{5}y\right)^2 + \frac{36}{25}x^2 = 4x^2 + y^2$$

$$4\left\|\begin{pmatrix} x \\ y \end{pmatrix}\right\|_2^2 = 4x^2 + 4y^2$$

So we want to solve:

$$4x^2 + 4y^2 = 4x^2 + y^2 \iff y = 0$$

So we can take the normalized vector $v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, and its orthongonal $v_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. We then know:

$$Av_1 = \begin{pmatrix} -\frac{24}{25} \\ -\frac{32}{25} \\ -\frac{6}{5} \end{pmatrix}, Av_2 = \begin{pmatrix} \frac{4}{5} \\ -\frac{3}{5} \\ 0 \end{pmatrix}$$

So we may write $U = \begin{pmatrix} u_1 & u_2 & u_3 \end{pmatrix}$. Now, $Av_1 = UDVv_1 = UDe_1 = 2Ue_1 \implies u_1 = \frac{1}{2}Av_1 = \begin{pmatrix} -\frac{12}{25} \\ -\frac{16}{25} \\ -\frac{3}{5} \end{pmatrix}$. Similarly $Av_2 = UDVv_2 = UDe_2 = Ue_2 = u_2 = \begin{pmatrix} \frac{4}{5} \\ -\frac{3}{5} \\ 0 \end{pmatrix}$. Finally, we need u_3 to be orthogonal to u_1 and u_2. Take $u_3 = u_1 \times u_2 =$

$\begin{pmatrix} -\frac{9}{25} \\ -\frac{12}{25} \\ \frac{4}{5} \end{pmatrix}$. So:

$$\begin{pmatrix} -\frac{24}{25} & \frac{4}{5} & -\frac{9}{25} \\ -\frac{32}{25} & -\frac{3}{5} & -\frac{12}{25} \\ -\frac{6}{5} & 0 & \frac{4}{5} \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

We move on to B. First:

$$BB^T = \begin{pmatrix} 1 & 2 & 1 & 0 \\ 0 & 1 & 2 & 1 \\ 1 & 0 & 1 & 2 \\ 2 & 1 & 0 & 1 \end{pmatrix}$$

The characteristic polynomial is $x(x-2)^2(x-4)$ so the eigenvalues are $4, 2, 2, 0$ and the singular values are $2, \sqrt{2}, \sqrt{2}, 0$. Hence $D = \begin{pmatrix} 2 & & & \\ & \sqrt{2} & & \\ & & \sqrt{2} & \\ & & & 0 \end{pmatrix}$. Now, observe that $B \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \\ 2 \\ 2 \end{pmatrix}$ so we may take $v_1 = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}$. We have: $Bv_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = 2Ue_1 = 2u_1 \implies u_1 = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}$. Now, observe that $B \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \end{pmatrix}$, and $\left\| \begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \end{pmatrix} \right\|_2 = 2 = \sqrt{2}\sqrt{2} = \sqrt{2} \left\| \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix} \right\|_2$. So we'll take $v_2 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ -\frac{1}{\sqrt{2}} \\ 0 \end{pmatrix}$. Similarly, we'll take $v_3 = \begin{pmatrix} 0 \\ \frac{1}{\sqrt{2}} \\ 0 \\ -\frac{1}{\sqrt{2}} \end{pmatrix}$. We then have $u_2 = \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \end{pmatrix}, u_3 = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ -\frac{1}{2} \\ -\frac{1}{2} \end{pmatrix}$. Finally, we have $B \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix} = 0$, so we may take $v_4 = \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \\ -\frac{1}{2} \end{pmatrix}$ and $u_4 = \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \\ -\frac{1}{2} \end{pmatrix}$. Thus:

$$B = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} 2 & & & \\ & \sqrt{2} & & \\ & & \sqrt{2} & \\ & & & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} & 0 \\ 0 & \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \end{pmatrix}$$

Problem 3.

(1) By definition, the rank of a matrix is the size of the maximal set of linearly independent columns of the matrix. Consider any two columns in uv^T. They have the form: $v_i u, v_j u$, and they are clearly linearly dependent.

(2) Let $u_1^T, \dots, u_k^T \in \mathbb{R}$ be linearly independent rows of A. Then for every row a_i^T of A there exist $v_i \in \mathbb{R}^k$ s.t. $a_i^T = v_i^T \begin{pmatrix} u_1^T \\ \dots \\ u_k^T \end{pmatrix}$. So $A = \begin{pmatrix} v_1^T \\ \dots \\ v_m^T \end{pmatrix} \begin{pmatrix} u_1^T \\ \dots \\ u_k^T \end{pmatrix}$. Let $b_1, \dots, b_k$ be the columns of $\begin{pmatrix} v_1^T \\ \dots \\ v_m^T \end{pmatrix}$. Then $A = \begin{pmatrix} b_1 & \dots & b_k \end{pmatrix} \begin{pmatrix} u_1^T \\ \dots \\ u_k^T \end{pmatrix} = \sum_{i=1}^k b_i u_i^T$. By an argument similar to the previous part we know that $u_i b_i^T$ is of rank 1 or 0. Now $u_i \neq 0$ because $u_1, \dots, u_k$ are linearly independent. Also, for every i, $b_i \neq 0$ because for some j, $a_j^T = u_j^T$, which means that $(b_i)_j = 1$ (since $u_j^T = a_j^T = v_j^T \begin{pmatrix} u_1^T \\ \dots \\ u_k^T \end{pmatrix}$ which by linear independence yields $v_j^T = e_i^T$). Thus $b_i u_i^T \neq 0$ so it is of rank 1.

(3) Let UDV^T be a SVD of A. Take v_i^T to be the ith row of V^T and u_i to be the ith column of U.

Problem 4. Let UDV^T bs a SVD of A. Let $x \in \mathbb{R}^m$. Let v_i^T be the rows of V and let u_i be the columns of U. Then, applying Pythagoras' theorem and the fact that $\{v_i\}, \{u_i\}$ are orthonormal:

$$Ax = \sum \sigma_i u_i \langle x, v_i \rangle \implies \|Ax\|_2^2 = \sum \sigma_i^2 \langle x, v_i \rangle^2 \geq \sigma_n^2 \sum \langle x, v_i \rangle^2 = \sigma_n^2 \|Ax\|_2^2$$

which implies the result.

Problem 5.

(1) Basic calculus - non-empty closed and bounded subsets of $\mathbb{R}$ attain a maximum. S_A is trivially non-empty, so it's enough to show that it's bounded and closed.

(2) Recall all norms on $\mathbb{R}^n$ are equivalent, so there exists some $C > 0$ s.t. for all $x \in \mathbb{R}^n$, $\|x\|_1 \leq C N_n(x)$. Thus, for all $x \in \mathbb{R}^n$ s.t. $N_n(x) = 1$, $\|x\|_1 \leq C$. So the set $B_1 = \{x \in \mathbb{R}^n : N_n(x) = 1\}$ is ℓ_1-bounded. Now, let $x = \sum_{i=1}^n x_i e_i \in B_1$. It holds:

$$N_m(Ax) \leq \sum_{i=1}^n |x_i| N_m(Ae_i) \leq \left(\max_{i \in [n]} N_m(Ae_i)\right) \|x\|_1 \leq \left(\max_{i \in [n]} N_m(Ae_i)\right) C$$

Thus S_A is bounded above by $\left(\max_{i \in [n]} N_m(Ae_i)\right) C$ and below by 0.

Now, let $0 \neq x \in \mathbb{R}^n$. Then $\frac{x}{N_n(x)} \in B_1$. $N_m(Ax) = N_n(x) N_m\left(A\left(\frac{x}{N_n(x)}\right)\right) \leq N_n(x) \left(\max_{i \in [n]} N_m(Ae_i)\right) C$.

(3) Let $\{s_k\}_{k=1}^\infty \subseteq S_A$ be a sequence that converges to $s \in \mathbb{R}$. For every k, let $x_k \in B_1$ be s.t. $N_m(Ax_k) = s_k$. The sequence $\{x_k\}_{k=1}^\infty \subseteq B_1 \subseteq \mathbb{R}^n$ is ℓ_1-bounded by what we showed in the previous section. Therefore, by the Bolzano-Weierstrass theorem it has an ℓ_1-convergent subsequence

$\{x_{k_\ell}\}_{\ell=1}^{\infty}$. Say x_{k_ℓ} converges to $x \in \mathbb{R}^n$. We'll show that $x \in B_1$ and that $N_m(Ax) = s$. Then $\lim_{\ell\to\infty} \|x_{k_\ell} - x\|_1 = 0$. But there exists some $D > 0$ s.t. for all $y \in \mathbb{R}^n$, $N_n(y) \le D\|y\|_1$. Therefore, $\lim_{\ell\to\infty} N_n(x_{k_\ell} - x) \le \lim_{\ell\to\infty} D\|x_{k_\ell} - x\|_1 = 0$. Using the triangle inequality:

$$|N_n(x) - 1| = |N_n(x) - N_n(x_{k_\ell})| \le N_n(x - x_{k_\ell}) \to 0 \implies N_n(x) = 1$$

Therefore $x \in B_1$. Furthermore, using the triangle inequality and linearity of A:

$$|N_m(Ax) - s_{k_\ell}| = |N_m(Ax) - N_m(Ax_{k_\ell})| \le N_m(A(x - x_{k_\ell})) \le N_n(x - x_{k_\ell})\left(\max_{i\in[n]} N_m(Ae_i)\right) C \to 0$$

So

$$N_m(Ax) = \lim_{\ell\to\infty} s_{k_\ell} = s$$

as desired.