MASSACHUSETTS INSTITUTE OF TECHNOLOGY
Department of Electrical Engineering and Computer Science

6.241: Dynamic Systems—Fall 2003

Homework 5 Solutions

Exercise 5.1 Suppose that $A \in C^{m \times n}$ is perturbed by the matrix $E \in C^{m \times n}$.

1. Show that

$$|\sigma_{max}(A+E) - \sigma_{max}(A)| \leq \sigma_{max}(E).$$

Also find an E that achieves the upper bound.

Note that

$$A = A + E - E \rightarrow \|A\| = \|A + E - E\| \leq \|A + E\| + \|E\| \rightarrow \|A\| - \|A + E\| \leq \|E\|.$$

Also,

$$(A + E) = A + E \rightarrow \|A + E\| \leq \|A\| + \|E\| \rightarrow \|A + E\| - \|A\| \leq \|E\|.$$

Thus, putting the two inequalities above together, we get that

$$|\|A + E\| - \|A\|| \leq \|E\|.$$

Note that the norm can be any matrix norm, thus the above inequality holds for the 2-induced norms which gives us,

$$|\sigma_{max}(A+E) - \sigma_{max}(A)| \leq \sigma_{max}(E).$$

A matrix E that achieves the upper bound is

$$E = U \begin{pmatrix} -\sigma_1 & 0 & 0 & 0 \\ 0 & \ddots & 0 & 0 \\ \vdots & & -\sigma_r & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix} V' = -A,$$

where U and V form the SVD of A. Here, $A + E = 0$, thus $\sigma_{max}(A+E) = 0$, and

$$|0 + \sigma_{max}(A)| = \sigma_{max}(E)$$

is achieved.

2. Suppose that A has less than full column rank, *i.e.*, the rank$(A) < n$, but $A + E$ has full column rank. Show that

$$\sigma_{min}(A + E) \leq \sigma_{max}(E).$$

Since A does not have full column rank, there exists $x \neq 0$ such that

$$Ax = 0 \rightarrow (A+E)x = Ex \rightarrow \|(A+E)x\|_2 = \|Ex\|_2 \rightarrow \frac{\|(A+E)x\|_2}{\|x\|_2} = \frac{\|Ex\|_2}{\|x\|_2} \leq \|E\|_2 = \sigma_{max}(E).$$

But,

$$\sigma_{min}(A + E) \leq \frac{\|(A+E)x\|_2}{\|x\|_2},$$

as shown in chapter 4 (please refer to the proof in the lecture notes!). Thus

$$\sigma_{min}(A + E) \leq \sigma_{max}(E).$$

Finally, a matrix E that results in $A+E$ having full column rank and that achieves the upper bound is

$$E = U \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \ddots & 0 & 0 \\ \vdots & 0 & \sigma_{r+1} & \vdots \\ 0 & 0 & 0 & \sigma_{r+1} \\ & & 0 & \end{pmatrix} V',$$

for

$$A = U \begin{pmatrix} \sigma_1 & 0 & 0 & 0 \\ 0 & \ddots & 0 & 0 \\ \vdots & 0 & \sigma_r & \vdots \\ & & 0 & \end{pmatrix} V'.$$

Note that A has rank $r < n$, but that $A + E$ has rank n,

$$A + E = U \begin{pmatrix} \sigma_1 & 0 & 0 & 0 & 0 \\ 0 & \ddots & 0 & 0 & 0 \\ 0 & 0 & \sigma_r & 0 & 0 \\ 0 & 0 & 0 & \sigma_{r+1} & 0 \\ 0 & 0 & 0 & \dots & \sigma_{r+1} \\ & & & & \\ & & 0 & & \end{pmatrix} V'.$$

It is easy to see that $\sigma_{min}(A + E) = \sigma_{r+1}$, and that $\sigma_{max}(E) = \sigma_{r+1}$.

The result in part 2, and some extensions to it, give rise to the following procedure (which is widely used in practice) for estimating the rank of an unknown matrix A from a known matrix $A+E$, where $\|E\|_2$ is known as well. Essentially, the SVD of $A+E$ is computed, and the rank of A is then estimated to be the number of singular values of $A+E$ that are larger than $\|E\|_2$.

Exercise 5.2 Using SVD, A can be decomposed as

$$A = U \begin{pmatrix} \sigma_1 & & & & & \\ & \ddots & & & & \\ & & \sigma_{r+1} & & & \\ & & & \ddots & & \\ & & & & \sigma_k & \\ & & & & & 0 \end{pmatrix} V',$$

where U and V are unitary matrices. Following the given procedure, let's select the first r+1 columns of V : $\{v_1, v_2, \cdots, v_{r+1}\}$. From the definition of SVD those v_i's are orthonormal, and hence independent. Note that $\{v_1, v_2, \cdots, v_{r+1}, \cdots v_n\}$ span $\mathbb{R}^n$, and if $rank(E) = r$, then exactly r of the vectors, $\{v_1, v_2, \cdots, v_{r+1}, \cdots v_n\}$, span $\mathcal{R}(E') = \mathcal{N}^{\perp}(E)$. The remaining vectors span $\mathcal{N}(E)$. So, given any $r+1$ linearly independent vectors in $\mathbb{R}^n$, at least one must be in the nullspace of E. That is, $\exists \alpha_i$ for $i = 1, \cdots, r+1$ not all zero such that

$$E(\alpha_1 v_1 + \alpha_2 v_2 + \cdots \alpha_{r+1} v_{r+1}) = 0.$$

This implies that there exists a nonzero vector z such that

$$z = \sum_{i=1}^{r+1} \alpha_i v_i = \begin{pmatrix} v_1 & \cdots & v_{r+1} \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_{r+1} \end{pmatrix}$$

with $\|z\|_2 = 1$ such that $Ez = 0$. Thus,

$$(A-E)z = Az = U\Sigma \begin{pmatrix} - & {v_1}' & - \\ & \vdots & \\ - & {v_{r+1}}' & - \end{pmatrix} \left(\sum_{i=1}^{r+1} \alpha_i v_i \right) = U \begin{pmatrix} \sigma_1\alpha_1 \\ \sigma_2\alpha_2 \\ \vdots \\ \sigma_{r+1}\alpha_{r+1} \\ 0 \\ \vdots \\ 0 \end{pmatrix} \tag{1}$$

By taking 2-norm of both sides of the above equation,

$$\begin{aligned} \|(A-E)z\|_2 &= \|U \begin{pmatrix} \sigma_1\alpha_1 \\ \sigma_2\alpha_2 \\ \vdots \\ \sigma_{r+1}\alpha_{r+1} \\ 0 \\ \vdots \\ 0 \end{pmatrix} \|_2 = \| \begin{pmatrix} \sigma_1\alpha_1 \\ \sigma_2\alpha_2 \\ \vdots \\ \sigma_{r+1}\alpha_{r+1} \\ 0 \\ \vdots \\ 0 \end{pmatrix} \|_2 \quad \text{(since U is a unitary matrix)} \\ &= \left(\sum_{i=1}^{r+1} |\sigma_i\alpha_i|^2 \right)^{\frac{1}{2}} \geq \sigma_{r+1} \left(\sum_{i=1}^{r+1} |\alpha_i|^2 \right)^{\frac{1}{2}}. \end{aligned} \tag{2}$$

But, from our construction of z,

$$\|z\|_2^2 = 1 \rightarrow \|\begin{pmatrix} v_1 & \cdots & v_{r+1} \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_{r+1} \end{pmatrix} \|_2^2 = 1 \rightarrow \| \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_{r+1} \end{pmatrix} \|_2^2 = \sum_{i=1}^{r+1} |\alpha_i|^2 = 1.$$

Thus, equation(2) becomes

$$\|(A-E)z\|_2 \geq \sigma_{r+1}.$$

Finally, $\|(A-E)z\|_2 \leq \|A-E\|_2$ for all z such that $\|z\|_2 = 1$.

Now we have to achieve the lower bound. Choose

$$E = U \begin{pmatrix} \sigma_1 & & & \\ & \ddots & & \\ & & \sigma_r & \\ & & & 0 \end{pmatrix} V'.$$

E has rank r, and

$$A-E = U \begin{pmatrix} 0 & & & & & & \\ & \ddots & & & & & \\ & & 0 & & & & \\ & & & \sigma_{r+1} & & & \\ & & & & \ddots & & \\ & & & & & \sigma_k & \\ & & & & & & 0 \end{pmatrix} V'.$$

where $\|A-E\|_2 = \sigma_{r+1}$.

Exercise 5.8 (1) The statement is false, to construct a counterexample use $A = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$ and $B = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$.

(2) True, see exercise 5.2.

(3) True, first note that for any invertible matrix A we have $\sigma_{max}(A^{-1}) = \sigma_{min}(A)$, infact one can easily show that if $A \in \mathbf{C^{mxm}}$ then $\sigma_i(A) = \sigma_j(A^{-1})$ with $j = m - i + 1$ and $i \in \{1, ..., m\}$. Given that remark all we need to show is that $1 - \sigma_{max}(A) \leq \sigma_{min}(I - A)$. From the triangular inequality one has :

$$|\|Ix\|_2 - \|Ax\|_2| \leq \|(I - A)x\|_2$$

restricting our attention to unit vectors x, so $\|x\|_2 = 1$ and taking into account that $\sigma_{max}(A) < 1$ we obtain:

$$1 - \|Ax\|_2 \leq \|(I - A)x\|_2$$

but for any x on the unit sphere $1 - \sigma_{max}(A) \leq 1 - \|Ax\|_2$, thus

$$1 - \sigma_{max}(A) \leq \|(I - A)x\|_2$$

the above equation is still valid for any x on the unit sphere, ie $\|x\|_2 = 1$ so we can choose x such that $\|(I - A)x\|_2 = \sigma_{min}(I - A)$, thus

$$1 - \sigma_{max}(A) \leq \sigma_{min}(I - A)$$

completing the proof.

(4) Consider a matrix E with $rank(E) \leq i - 1$, but otherwise arbitrary. From the triangular inequality we have

$$\|A + B - E\|_2 \leq \|A - E\|_2 + \|B\|_2$$

from exercise 5.2 we know that $\sigma_i(A) \leq \|A - E\|_2$ if $rank(E) = i - 1$, thus

$$\sigma_i(A + B) \leq \|A - E\|_2 + \sigma_{max}(B)$$

now choose the particular E that achieves the bound, ie. $\|A - E\|_2 = \sigma_i(A)$ and we get

$$\sigma_i(A + B) \leq \sigma_i(A) + \sigma_{max}(B)$$

by letting $B = I$ we obtain the desired result.