MASSACHUSETTS INSTITUTE OF TECHNOLOGY
Department of Electrical Engineering and Computer Science

6.241: Dynamic Systems—Fall 2003

Homework 3 Solutions

Ex. 4.1 Note that for any $v \in C^m$, (show this!)

$$\|v\|_\infty \leq \|v\|_2 \leq \sqrt{m}\|v\|_\infty. \tag{1}$$

Therefore, for $A \in C^{m\times n}$ with $x \in \mathbb{C}^n$

$$\|Ax\|_2 \leq \sqrt{m}\|Ax\|_\infty \rightarrow \text{for } x \neq 0, \tfrac{\|Ax\|_2}{\|x\|_2} \leq \sqrt{m}\tfrac{\|Ax\|_\infty}{\|x\|_2}.$$

But, from equation (1), we also know that $\frac{1}{\|x\|_\infty} \geq \frac{1}{\|x\|_2}$. Thus,

$$\frac{\|Ax\|_2}{\|x\|_2} \leq \frac{\sqrt{m}\|Ax\|_\infty}{\|x\|_2} \leq \frac{\sqrt{m}\|Ax\|_\infty}{\|x\|_\infty} \leq \sqrt{m}\|A\|_\infty, \tag{2}$$

Equation (2) must hold for all $x \neq 0$, therefore

$$\max_{x\neq 0} \tfrac{\|Ax\|_2}{\|x\|_2} = \|A\|_2 \leq \sqrt{m}\|A\|_\infty.$$

To prove the lower bound $\frac{1}{\sqrt{n}}\|A\|_\infty \leq \|A\|_2$, reconsider equation (1):

$$\|Ax\|_\infty \leq \|Ax\|_2 \rightarrow \text{for } x \neq 0, \frac{\|Ax\|_\infty}{\|x\|_2} \leq \frac{\|Ax\|_2}{\|x\|_2} \leq \|A\|_2 \rightarrow \frac{\sqrt{n}\|Ax\|_\infty}{\|x\|_2} \leq \frac{\sqrt{n}\|Ax\|_2}{\|x\|_2} \leq \sqrt{n}\|A\|_2. \tag{3}$$

But, from equation (1) for $x \in C^n$, $\frac{\sqrt{n}}{\|x\|_2} \geq \frac{1}{\|x\|_\infty}$. So,

$$\frac{\|Ax\|_\infty}{\|x\|_\infty} \leq \frac{\sqrt{n}\|Ax\|_\infty}{\|x\|_2} \leq \sqrt{n}\|A\|_2$$

for all $x \neq 0$ including x that makes $\frac{\|Ax\|_\infty}{\|x\|_\infty}$ maximum, so,

$$\max_{x\neq 0} \tfrac{\|Ax\|_\infty}{\|x\|_\infty} = \|A\|_\infty \leq \sqrt{n}\|A\|_2,$$

or equivalently,

$$\frac{1}{\sqrt{n}}\|A\|_\infty \leq \|A\|_2.$$

Exercise 4.5 Any $m \times n$ matrix A, it can be expressed as

$$A = U \begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix} V',$$

where U and V are unitary matrices. The "Moore-Penrose inverse", or *pseudo-inverse* of A, denoted by A^+, is then defined as the $n \times m$ matrix

$$A^+ = V \begin{pmatrix} \Sigma^{-1} & 0 \\ 0 & 0 \end{pmatrix} U'.$$

a) Now we have to show that A^+A and AA^+ are symmetric, and that $AA^+A = A$ and $A^+AA^+ = A^+$. Suppose that Σ is a diagonal invertible matrix with the dimension of $r \times r$. Using the given definitions as well as the fact that for a unitary matrix U, $U'U = UU' = I$, we have

$$\begin{aligned} AA^+ &= U \begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix} V'V \begin{pmatrix} \Sigma^{-1} & 0 \\ 0 & 0 \end{pmatrix} U' \\ &= U \begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix} I \begin{pmatrix} \Sigma^{-1} & 0 \\ 0 & 0 \end{pmatrix} U' \\ &= U \begin{pmatrix} I_{r\times r} & 0 \\ 0 & 0 \end{pmatrix} U', \end{aligned}$$

which is symmetric. Similarly,

$$\begin{aligned} A^+A &= V \begin{pmatrix} \Sigma^{-1} & 0 \\ 0 & 0 \end{pmatrix} U'U \begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix} V' \\ &= V \begin{pmatrix} \Sigma^{-1} & 0 \\ 0 & 0 \end{pmatrix} I \begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix} V' \\ &= V \begin{pmatrix} I_{r\times r} & 0 \\ 0 & 0 \end{pmatrix} V' \end{aligned}$$

which is again symmetric.
The facts derived above can be used to show the other two.

$$\begin{aligned} AA^+A &= (AA^+)A = U \begin{pmatrix} I_{r\times r} & 0 \\ 0 & 0 \end{pmatrix} U'A \\ &= U \begin{pmatrix} I_{r\times r} & 0 \\ 0 & 0 \end{pmatrix} U'U \begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix} V' \\ &= U \begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix} V' \\ &= A. \end{aligned}$$

Also,

$$
\begin{aligned}
A^+AA^+ &= (A^+A)A^+ = V\begin{pmatrix} I_{r\times r} & 0 \\ 0 & 0 \end{pmatrix} V'A^+ \\
&= V\begin{pmatrix} I_{r\times r} & 0 \\ 0 & 0 \end{pmatrix} V'V\begin{pmatrix} \Sigma^{-1} & 0 \\ 0 & 0 \end{pmatrix} U' \\
&= V\begin{pmatrix} \Sigma^{-1} & 0 \\ 0 & 0 \end{pmatrix} U' \\
&= A^+.
\end{aligned}
$$

b) We have to show that when A has full column rank then $A^+ = (A'A)^{-1}A'$, and that when A has full row rank then $A^+ = A'(AA')^{-1}$. If A has full column rank, then we know that $m \geq n$, $rank(A) = n$, and

$$
A = U\begin{pmatrix} \Sigma_{n\times n} \\ 0 \end{pmatrix} V'.
$$

Also, as shown in chapter 2, when A has full column rank, $(A'A)^{-1}$ exists. Hence

$$
\begin{aligned}
(A'A)^{-1}A' &= \left(V\begin{pmatrix} \Sigma' & 0 \end{pmatrix} U'U\begin{pmatrix} \Sigma \\ 0 \end{pmatrix} V'\right)^{-1} V\begin{pmatrix} \Sigma' & 0 \end{pmatrix} U' \\
&= \left(V\Sigma'\Sigma V'\right)^{-1} V\begin{pmatrix} \Sigma' & 0 \end{pmatrix} U' \\
&= V(\Sigma'\Sigma)^{-1}V'V\begin{pmatrix} \Sigma' & 0 \end{pmatrix} U' \\
&= V(\Sigma'\Sigma)^{-1}\begin{pmatrix} \Sigma' & 0 \end{pmatrix} U' \\
&= V\begin{pmatrix} \Sigma^{-1} & 0 \end{pmatrix} U' \\
&= A^+.
\end{aligned}
$$

Similarly, if A has full row rank, then $n \geq m$, $rank(A) = m$, and

$$
A = U\begin{pmatrix} \Sigma_{m\times m} & 0 \end{pmatrix} V'.
$$

It can be proved that when A has full row rank, $(A'A)^{-1}$ exists. Hence,

$$
\begin{aligned}
A'(AA')^{-1} &= V\begin{pmatrix} \Sigma' \\ 0 \end{pmatrix} U'\left(U\begin{pmatrix} \Sigma & 0 \end{pmatrix} V'V\begin{pmatrix} \Sigma' \\ 0 \end{pmatrix} U'\right)^{-1} \\
&= V\begin{pmatrix} \Sigma' \\ 0 \end{pmatrix} U'\left(U\Sigma\Sigma'U'\right)^{-1} \\
&= V\begin{pmatrix} \Sigma' \\ 0 \end{pmatrix} U'U(\Sigma\Sigma^{-1})U' \\
&= V\begin{pmatrix} \Sigma^{-1} \\ 0 \end{pmatrix} U' \\
&= A^+.
\end{aligned}
$$

c) Show that, of all x that minimize $\|y - Ax\|_2$, the one with the smallest length $\|x\|_2$ is given by $\hat{x} = A^+y$. If A has full row rank, we have shown in chapter 3 that the solution with the smallest length is given by

$$
\hat{x} = A'(AA')^{-1}y,
$$

and from part (b), $A'(AA')^{-1} = A^+$. Therefore

$$\hat{x} = A^+ y.$$

Similary, it can be shown that the pseudo inverse is the solution for the case when a matrix A has a full column rank (compare the results in chapter 2 with the expression you found in part (b) for A^+ when A has full column rank).

Now, let's consider the case when a matrix A is rank deficient, *i.e.*, $rank(A) = r < \min(m,n)$ where $A \in C^{m\times n}$ and is thus neither full row or column rank. Suppose we have a singular value decomposition of A as

$$A = U\Sigma V',$$

where U and V are unitary matrices. Then the norm we are minimizing is

$$\|Ax - y\| = \|U\Sigma V'x - y\| = \|U(\Sigma V'x - U'y)\| = \|\Sigma z - U'y\|,$$

where $z = V'x$, since $\|\cdot\|$ is unaltered by the orthogonal transformation, U. Thus, x minimizes $\|Ax - y\|$ if and only if z minimizes $\|\Sigma z - c\|$, where $c = U'y$. Since the rank of A is r, the matrix Σ has the nonzero singular values $\sigma_1, \sigma_2, \cdots, \sigma_r$ in its diagonal entries. Then we can rewrite $\|\Sigma z - c\|^2$ as follows:

$$\|\Sigma z - c\|^2 = \sum_{i=1}^{r}(\sigma_i z_i - c_i)^2 + \sum_{i=r+1}^{n} c_i^2.$$

It is clear that the minimum of the norm can be achived when $z_i = \frac{c_i}{\sigma_i}$ for $i = 1, 2, \cdots, r$ and the rest of the z_i's can be chosen arbitrarily. Thus, there are infinitely many solutions $\hat{z}$ and the solution with the minimum norm can be achieved when $z_i = 0$ for $i = r+1, r+2, \cdots, n$. Thus, we can write this $\hat{z}$ as

$$z = \Sigma_1 c,$$

where

$$\Sigma_1 = \begin{pmatrix} \Sigma_r^{-1} & 0 \\ 0 & 0 \end{pmatrix}$$

and Σ_r is a square matrix with nonzero singular values in its diagonal in decreasing order. This value of z also yields the value of x of minimal 2 norm since V is a unitary matrix.
Thus the solution to this problem is

$$\hat{x} = Vz = V\Sigma_1 c = V\Sigma_1 U'y = A^+ y.$$

It can be easily shown that this choice of A^+ satisfies all the conditions, or definitions, of pseudo inverse in a).

Exercise 4.7 Given a complex square matrix A, the definition of the *structured singular value function* is as follows.

$$\mu_{\underline{\Delta}}(A) = \frac{1}{\min_{\Delta \in \underline{\Delta}}\{\sigma_{max}(\Delta) |\ det(I - \Delta A) = 0\}}$$

where $\underline{\Delta}$ is some set of matrices.
a) If $\underline{\Delta} = \{\alpha I : \alpha \in \mathbf{C}\}$, then $det(I - \Delta A) = det(I - \alpha A)$. Here $det(I - \alpha A) = 0$ implies that there exists an $x \neq 0$ such that $(I - \alpha A)x = 0$. Expanding the left hand side of the equation yields $x = \alpha A x \rightarrow \frac{1}{\alpha} x = Ax$. Therefore $\frac{1}{\alpha}$ is an eigenvalue of A. Since $\sigma_{max}(\Delta) = |\alpha|$,

$$arg \min_{\delta \in \underline{\Delta}} \{\sigma_{max}(\Delta) | det(I - \Delta A) = 0\} = |\alpha| = |\frac{1}{\lambda_{max}(A)}|.$$

Therefore, $\mu_{\underline{\Delta}}(A) = |\lambda_{max}(A)|$.

b) If $\underline{\Delta} = \{\Delta \in \mathbf{C}^{n \times n}\}$, then following a similar argument as in a), there exists an $x \neq 0$ such that $(I - \Delta A)x = 0$. That implies that

$$\begin{aligned} x = \Delta A x \quad & \rightarrow \quad \|x\|_2 = \|\Delta A x\|_2 \leq \|\Delta\|_2 \|Ax\|_2 \\ & \rightarrow \quad \frac{1}{\|\Delta\|_2} \leq \frac{\|Ax\|_2}{\|x\|_2} \leq \sigma_{max}(A) \\ & \rightarrow \quad \frac{1}{\sigma_{max}(A)} \leq \sigma_{max}(\Delta). \end{aligned}$$

Then, we show that the lower bound can be achieved. Since $\underline{\Delta} = \{\Delta \in \mathbf{C}^{n \times n}\}$, we can choose Δ such that

$$\Delta = V \begin{pmatrix} \frac{1}{\sigma_{max}(A)} & & & \\ & 0 & & \\ & & \ddots & \\ & & & 0 \end{pmatrix} U'.$$

where U and V are from the SVD of A, $A = U \Sigma V'$. Note that this choice results in

$$I - \Delta A = I - V \begin{pmatrix} 1 & & & \\ & 0 & & \\ & & \ddots & \\ & & & 0 \end{pmatrix} V' = V \begin{pmatrix} 0 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix} V$$

which is singular, as required. Also from the construction of Δ, $\sigma_{max}(\Delta) = \frac{1}{\sigma_{max}(A)}$. Therefore, $\mu_{\underline{\Delta}}(A) = \sigma_{max}(A)$.

c) If $\underline{\Delta} = \{diag(\alpha_1, \cdots, \alpha_n) | \alpha_i \in \mathbf{C}\}$ with $D \in \{diag(d_1, \cdots, d_n) | d_i > 0\}$, we first note that D^{-1} exists. Thus:

$$\begin{aligned} det(I - \Delta D^{-1} A D) &= det(I - D^{-1} \Delta A D) \\ &= det((D^{-1} - D^{-1} \Delta A) D) \\ &= det(D^{-1} - D^{-1} \Delta A) det(D) \\ &= det(D^{-1}(I - \Delta A)) det(D) \\ &= det(D^{-1}) det(I - \Delta A) det(D) \\ &= det(I - \Delta A). \end{aligned}$$

Where the first equality follows because Δ and D^{-1} are diagonal and the last equality holds because $det(D^{-1}) = 1/det(D)$. Thus, $\mu_{\underline{\Delta}}(A) = \mu_{\underline{\Delta}}(D^{-1}AD)$.

Now let's show the left side inequality first. Since $\underline{\Delta}_1 \subset \underline{\Delta}_2$, $\underline{\Delta}_1 = \{\alpha I | \alpha \in \mathbf{C}\}$ and $\underline{\Delta}_2 = \{diag(\alpha_1, \cdots, \alpha_n)\}$, we have that

$$\min_{\Delta \in \underline{\Delta}_1} \{\sigma_{max}(\Delta)|\ det(I - \Delta A) = 0\} \geq \min_{\Delta \in \underline{\Delta}_2} \{\sigma_{max}(\Delta)| det(I - \Delta A) = 0\},$$

which implies that

$$\mu_{\underline{\Delta}_1}(A) \leq \mu_{\underline{\Delta}_2}(A).$$

But from part (a), $\mu_{\underline{\Delta}_1}(A) = \rho(A)$, so,

$$\rho(A) \leq \mu_{\underline{\Delta}_2}(A).$$

Now we have to show the right side of inequality. Note that with $\underline{\Delta}_3 = \{\Delta \in \mathbf{C}\}$, we have $\underline{\Delta}_2 \subset \underline{\Delta}_3$. Thus by following a similar argument as above, we have

$$\min_{\Delta \in \underline{\Delta}_2} \{\sigma_{max}(\Delta)|\ det(I - \Delta A) = 0\} \geq \min_{\Delta \in \underline{\Delta}_3} \{\sigma_{max}(\Delta)|\ det(I - \Delta A) = 0\}.$$

Hence,

$$\mu_{\underline{\Delta}_2}(A) = \mu_{\underline{\Delta}_2}(D^{-1}AD) \leq \mu_{\underline{\Delta}_3}(D^{-1}AD) = \sigma_{max}(D^{-1}AD).$$

Exercise 4.8 We are given a complex square matrix A with $rank(A) = 1$. According to the SVD of A we can write $A = uv'$ where u, v are complex vectors of dimension n. To simplify computations we are asked to minimize the Frobenius norm of Δ in the definition of $\mu_{\underline{\Delta}}(A)$. So

$$\mu_{\underline{\Delta}}(A) = \frac{1}{\min_{\Delta \in \underline{\Delta}} \{\ \|\Delta\|_F \ |\ det(I - \Delta A) = 0\}}$$

$\underline{\Delta}$ is the set of diagonal matrices with complex entries, $\underline{\Delta} = \{diag(\delta_1, \cdots, \delta_n) | \delta_i \in \mathbf{C}\}$. Introduce the column vector $\delta = (\delta_1, \cdots, \delta_n)^T$ and the row vector $B = (u_1 v_1^*, \cdots, u_n v_n^*)$, then the original problem can be reformulated after some algebraic manipulations as

$$\mu_{\underline{\Delta}}(A) = \frac{1}{\min_{\delta \in \mathbf{C^n}} \{\ \|\delta\|_2 \ |\ B\delta = 1\}}$$

To see this, we use the fact that $A = uv'$, and (from excercise 1.3(a))

$$\begin{aligned} det(I - \Delta A) &= det(I - \Delta uv') \\ &= det(1 - v'\Delta u) \\ &= 1 - v'\Delta u \end{aligned}$$

Thus $det(I - \Delta A) = 0$ implies that $1 - v'\Delta u = 0$. Then we have

$$\begin{aligned} 1 &= v'\Delta u \\ &= \begin{pmatrix} v_1^* & \cdots & v_n^* \end{pmatrix} \begin{pmatrix} \delta_1 & & \\ & \ddots & \\ & & \delta_n \end{pmatrix} \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} \\ &= \begin{pmatrix} v_1^* u_1 & \cdots & v_n^* u_n \end{pmatrix} \begin{pmatrix} \delta_1 \\ \vdots \\ \delta_n \end{pmatrix} \\ &= B\delta \end{aligned}$$

Hence, computing $\mu_{\underline{\Delta}}(A)$ reduces to a least square problem, i.e.,

$$\min_{\Delta \in \underline{\Delta}} \{ \|\Delta\|_F | \ det(I - \Delta A) = 0 \} \Leftrightarrow \min \|\underline{\delta}\|_2 \text{ s.t. } 1 = \text{B}\delta.$$

We are dealing with a underdetermined system of equations and we are seeking a minimum norm solution. Using the projection theorem, the optimal δ is given from $\delta^o = B'(BB')^{-1}$. Substituting in the expression of the structured singular value function we obtain:

$$\mu_{\underline{\Delta}}(A) = \sqrt{\sum_{i=1}^{n} |u_i v_i^*|^2}$$

In the second part of this exercise we define $\underline{\Delta}$ to be the set of diagonal matrices with real entries, $\underline{\Delta} = \{diag(\delta_1, \cdots, \delta_n) | \delta_i \in \mathbf{R}\}$. The idea remains the same, we just have to alter the constraint equation, namely $B\delta = 1 + 0j$. Equivalently one can write $D\delta = d$ where $D = \begin{pmatrix} Re(B) \\ Im(B) \end{pmatrix}$ and $d = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$. Again the optimal δ is obtained by use of the projection theorem and $\delta^o = D'(DD^T)^{-1}d$. Substituting in the expression of the structured singular value function we obtain:

$$\mu_{\underline{\Delta}}(A) = \frac{1}{\sqrt{d^T(DD^T)^{-1}d}}$$

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 $\|E\|_2$.