MASSACHUSETTS INSTITUTE OF TECHNOLOGY
Department of Electrical Engineering and Computer Science

6.241: Dynamic Systems—Spring 2011

Homework 4 Solutions

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 $x = \alpha A x \to \frac{1}{\alpha} x = Ax$. 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

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

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:

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

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

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

Also,

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

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

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

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

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

where U and V are unitary matrices and $k \geq r+1$. Following the given procedure, let's select the first r+1 columns of $V : \{v_1, v_2, \cdots, v_{r+1}\}$. Since V is unitary, those v_i's are orthonormal and hence independent. Note that $\{v_1, v_2, \cdots, v_{r+1}, \cdots v_n\}$ span $\mathbf{R}^n$, 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 $\mathbf{R}^n$, at least one must be in the nullspace of E. That is there exists coefficients c_i for $i = 1, \cdots, r+1$, not all zero, such that

$$E(c_1 v_1 + c_2 v_2 + \cdots c_{r+1} v_{r+1}) = 0.$$

These coefficients can be normalized to obtain a nonzero vector z, $\|z\|_2 = 1$, given by

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

and 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 $\|z\|_2 = 1$. Hence

$$\|A-E\|_2 \geq \sigma_{r+1}$$

To show that the lower bound can be achieved, choose

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

E has rank r,

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

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

Exercise 6.1 The model is linear one needs to note that the integration operator is a linear operator. Formally one writes

$$\begin{aligned} S(\alpha u_1 + \beta u_2)(t) &= \int_0^\infty e^{-(t-s)}(\alpha u_1(s) + \beta u_2(s))ds \\ &= \alpha \int_0^\infty e^{-(t-s)} u_1(s) + \beta \int_0^\infty e^{-(t-s)} u_2(s) \\ &= \alpha(Su_1)(t) + \beta(Su_2)(t) \end{aligned}$$

It is non-causal since future inputs are needed in order to determine the current value of y. Formally one writes

$$(P_T S u)(t) = (P_T S P_T u)(t) + P_T \left(\int_T^\infty e^{-(t-s)} u(s) ds \right)$$

It is not memoryless since the current output depends on the integration of past inputs. It is also time varying since

$$(S\sigma_T u)(t) = (\sigma_T S u)(t) + \int_{-T}^0 e^{-(t-T-s)} u(s) ds$$

one can argue that if the only valid input signals are those where $u(t) = 0$ if $t < 0$ then the system is time invariant.

Exercise 6.4(i) linear , time varying , causal , not memoryless
(ii) nonlinear (affine, tranlated linear) time varying , causal , not memoryless
(iii) nonlinear, time invariant , causal, memoryless
(iv) linear, time varying , causal, not memoryless
(i),(ii) can be called time invariant under the additional requirement that $u(t) = 0$ for $t < 0$