EECS 16B Fall 2019 — Designing Information Devices and Systems II, UC Berkeley

Midterm 2

1. Stability (14pts)

Consider the complex plane below, which is broken into non-overlapping regions A through H. The circle drawn on the figure is the unit circle $|\lambda| = 1$.

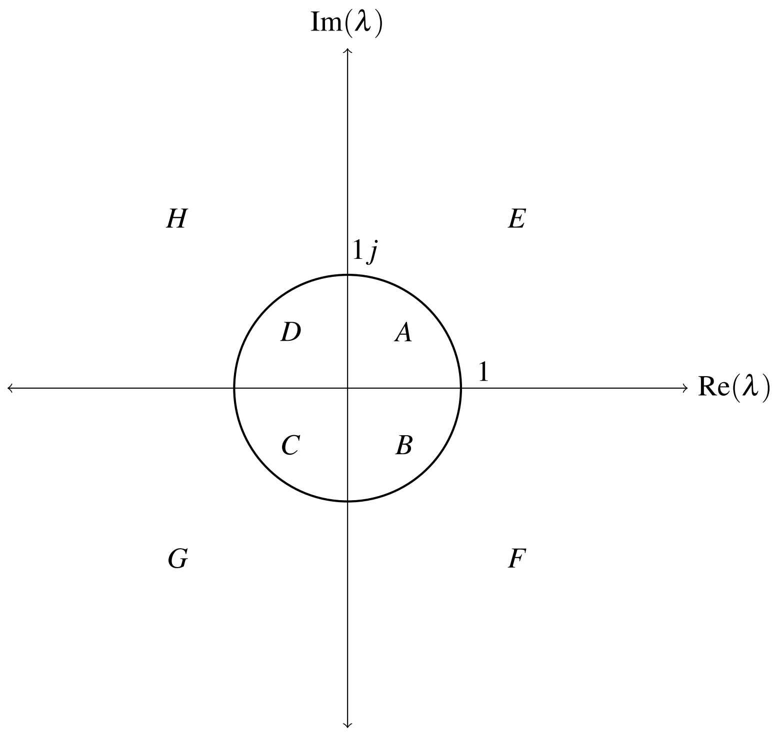

Figure 1: Complex plane divided into regions.

(a) (4pts) Consider the continuous-time system $\frac{d}{dt}x(t) = \lambda x(t) + v(t)$ and the discrete-time system $y(t+1) = \lambda y(t) + w(t)$.
In which regions can the eigenvalue λ be for a *stable* system? Fill out the table below to indicate *stable* regions. Assume that the eigenvalue λ does not fall directly on the boundary between two regions.

	A	**B**	**C**	**D**	**E**	**F**	**G**	**H**
Continuous Time System $x(t)$	◯	◯	◯	◯	◯	◯	◯	◯
Discrete Time System $y(t)$	◯	◯	◯	◯	◯	◯	◯	◯

(b) (10pts) Consider the continuous time system

$$\frac{d}{dt}x(t) = \lambda x(t) + u(t)$$

where λ is real and $\lambda < 0$.
Assume that $x(0) = 0$ and that $|u(t)| < \varepsilon$ for all $t \geq 0$.
Prove that the solution $x(t)$ will be bounded (i.e. $\exists k$ so that $|x(t)| \leq k\varepsilon$ for all time $t \geq 0$.).
(Hint: Recall that the solution to such a first-order scalar differential equation is:

$$x(t) = x_0 e^{\lambda t} + \int_0^t u(\tau) e^{\lambda(t-\tau)} d\tau$$

You may use this fact without proof.)

2. Computing the SVD (10pts)

Consider the matrix

$$A = \begin{bmatrix} 4 & -3 & 0 \\ 3 & 4 & 0 \end{bmatrix}.$$

Write out a singular value decomposition of the matrix A in the form $U\Sigma V^T$ where U is a 2×2 orthonormal matrix, Σ is a diagonal rectangular matrix, and V is a 3×3 orthonormal matrix.

3. Outlier Removal (14pts)

Suppose we have a system where we believe that a 2-dimensional vector input $\vec{x}$ leads to scalar outputs in a linear way $\vec{p}^T\vec{x}$. However, the parameters $\vec{p}$ are unknown and must be learned from data. Our data collection process is imperfect and instead of directly seeing $\vec{p}^T\vec{x}$, we get observations $y = \vec{p}^T\vec{x} + w$ where the w is some kind of disturbance or noise that nature introduces.

To allow us to learn the parameters, we have 4 experimentally obtained data points: input-output pairs $(\vec{x}_i, y_i)$ where the index $i = 1, \ldots, 4$.

$$X = \begin{bmatrix} 1 & 2 \\ 3 & 2 \\ 5 & 1 \\ 4 & 0 \end{bmatrix} \text{ and } \vec{y} = \begin{bmatrix} 5 \\ 7 \\ 5000 \\ 300 \end{bmatrix}$$

Then we can express the approximate system of equations that we want to solve as $X\vec{p} \approx \vec{y}$.

(a) (6pts) Suppose we know that the third data point $\left(\begin{bmatrix} 5 \\ 1 \end{bmatrix}, 5000\right)$ may have been corrupted, and we wish to effectively remove it from the data set. We decide to do so by augmenting our approximate system of equations to be

$$[X, \vec{a}] \begin{bmatrix} \vec{p} \\ f \end{bmatrix} \approx \vec{y}. \tag{11}$$

Mark all of the following choices for $\vec{a}$ which have the effect of eliminating the third data point from the data set if we run least-squares on (11) to estimate $\vec{p}$. Fill in the circles corresponding to your selections.

○ $\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$, ○ $\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$, ○ $\begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$, ○ $\begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$, ○ $\begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$, ○ $\begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix}$,

○ $\begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}$, ○ $\begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}$, ○ $\begin{bmatrix} 3 \\ 0 \\ 0 \\ 0 \end{bmatrix}$, ○ $\begin{bmatrix} 0 \\ 3 \\ 0 \\ 0 \end{bmatrix}$, ○ $\begin{bmatrix} 0 \\ 0 \\ 3 \\ 0 \end{bmatrix}$, ○ $\begin{bmatrix} 0 \\ 0 \\ 0 \\ 3 \end{bmatrix}$

(b) (8pts) Now suppose that we know that both the third data point $(\begin{bmatrix}5\\1\end{bmatrix}, 5000)$ and the fourth data point $(\begin{bmatrix}4\\0\end{bmatrix}, 300)$ have been corrupted, and we wish to effectively remove both of them from the data set. We decide to do so by augmenting our approximate system of equations to be

$$[X, \vec{a}, \vec{b}] \begin{bmatrix} \vec{p} \\ f_a \\ f_b \end{bmatrix} \approx \vec{y}. \tag{12}$$

Mark all of the following choices for $\vec{a}, \vec{b}$ pairs which have the effect of eliminating the third and fourth data points from the data set if we run least-squares on (12) to estimate $\vec{p}$. Fill in the circles corresponding to your selections.

○ $\begin{bmatrix}0\\0\\0\\0\end{bmatrix}, \begin{bmatrix}0\\0\\1\\1\end{bmatrix}$; ○ $\begin{bmatrix}0\\0\\1\\0\end{bmatrix}, \begin{bmatrix}0\\0\\0\\1\end{bmatrix}$; ○ $\begin{bmatrix}0\\0\\0\\1\end{bmatrix}, \begin{bmatrix}0\\0\\1\\1\end{bmatrix}$;

○ $\begin{bmatrix}0\\1\\1\\1\end{bmatrix}, \begin{bmatrix}0\\1\\1\\0\end{bmatrix}$; ○ $\begin{bmatrix}3\\0\\0\\0\end{bmatrix}, \begin{bmatrix}0\\4\\0\\0\end{bmatrix}$; ○ $\begin{bmatrix}0\\0\\3\\4\end{bmatrix}, \begin{bmatrix}0\\0\\3\\4\end{bmatrix}$

4. Control (18 pts)

Suppose that we have a two-dimensional discrete-time system governed by:

$$\vec{x}(t+1) = \begin{bmatrix} -\frac{2}{3} & \frac{1}{3} \\ \frac{1}{3} & -\frac{7}{6} \end{bmatrix} \vec{x}(t) + \vec{w}(t).$$

(a) (2pts) **Is the system stable? Why or why not?**

Here, we give you that

$$A = \begin{bmatrix} -\frac{2}{3} & \frac{1}{3} \\ \frac{1}{3} & -\frac{7}{6} \end{bmatrix} = \begin{bmatrix} -\frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{bmatrix} \begin{bmatrix} -\frac{4}{3} & 0 \\ 0 & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} -\frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{bmatrix}$$

and that the characteristic polynomial $\det(\lambda I - A) = \lambda^2 + \frac{11}{6}\lambda + \frac{2}{3}$.

(b) (6pts) Suppose that there is no disturbance and we can now influence the system using a scalar input $u(t)$ to our system:

$$\vec{x}(t+1) = \begin{bmatrix} -\frac{2}{3} & \frac{1}{3} \\ \frac{1}{3} & -\frac{7}{6} \end{bmatrix} \vec{x}(t) + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u(t).$$

Is the system controllable?

(c) (10pts) We want to set the closed-loop eigenvalues of the system

$$\vec{x}(t+1) = \begin{bmatrix} -\frac{2}{3} & \frac{1}{3} \\ \frac{1}{3} & -\frac{7}{6} \end{bmatrix} \vec{x}(t) + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u(t)$$

to be $\lambda_1 = -\frac{5}{6}, \lambda_2 = \frac{5}{6}$ using state feedback

$$u(t) = -\begin{bmatrix} k_1 & k_2 \end{bmatrix} \vec{x}(t).$$

What specific numeric values of k_1 and k_2 should we use?

5. Upper Triangularization (12 pts)

In this problem, you need to upper-triangularize the matrix

$$A = \begin{bmatrix} 3 & -1 & 2 \\ 3 & -1 & 6 \\ -2 & 2 & -2 \end{bmatrix}$$

The eigenvalues of this matrix A are $\lambda_1 = \lambda_2 = 2$ and $\lambda_3 = -4$. We want to express A as

$$A = \begin{bmatrix} \vec{x} & \vec{y} & \vec{z} \end{bmatrix} \cdot \begin{bmatrix} \lambda_1 & a & b \\ 0 & \lambda_2 & c \\ 0 & 0 & \lambda_3 \end{bmatrix} \cdot \begin{bmatrix} \vec{x}^\top \\ \vec{y}^\top \\ \vec{z}^\top \end{bmatrix}$$

where the $\vec{x}, \vec{y}, \vec{z}$ are orthonormal. Your goal in this problem is to compute $\vec{x}, \vec{y}, \vec{z}$ so that they satisfy the above relationship for some constants a, b, c.

Here are some potentially useful facts that we have gathered to save you some computations, you'll have to grind out the rest yourself.

$$\begin{bmatrix} 3 & -1 & 2 \\ 3 & -1 & 6 \\ -2 & 2 & -2 \end{bmatrix} \begin{bmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \\ 0 \end{bmatrix} = \begin{bmatrix} \sqrt{2} \\ \sqrt{2} \\ 0 \end{bmatrix}.$$

We also know that

$$\left\{ \begin{bmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \\ 0 \end{bmatrix}, \begin{bmatrix} \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$$

is an orthonormal basis, and

$$\begin{bmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 3 & -1 & 2 \\ 3 & -1 & 6 \\ -2 & 2 & -2 \end{bmatrix} \cdot \begin{bmatrix} \frac{\sqrt{2}}{2} & 0 \\ -\frac{\sqrt{2}}{2} & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & -2\sqrt{2} \\ -2\sqrt{2} & -2 \end{bmatrix}.$$

We also know that $\begin{bmatrix} 0 & -2\sqrt{2} \\ -2\sqrt{2} & -2 \end{bmatrix}$ has eigenvalues 2 and -4. The normalized eigenvector corresponding to $\lambda = 2$ is $\begin{bmatrix} -\frac{\sqrt{6}}{3} \\ \frac{\sqrt{3}}{3} \end{bmatrix}$ and $\begin{bmatrix} \frac{\sqrt{3}}{3} \\ \frac{\sqrt{6}}{3} \end{bmatrix}$ is a vector that is orthogonal to that and also has norm 1.

Based on the above information, compute $\vec{x}, \vec{y}, \vec{z}$. Show your work.

You don't have to compute the constants a, b, c in the interests of time.

6. Minimum Norm Variants (52 pts)

In lecture and HW, you saw how to solve minimum norm problems in which we have a wide matrix A and solve $A\vec{x} = \vec{y}$ such that $\vec{x}$ is a minimum norm solution: $||\vec{x}|| \leq ||\vec{z}||$ for all $\vec{z}$ such that $A\vec{z} = \vec{y}$.

We also saw in the HW how we can solve some variants in which we were interested in minimizing the norm $||C\vec{x}||$ instead. You have solved the case where C is invertible and square or a tall matrix. This question asks you about the case when C is a wide matrix. The key issue is that wide matrices have nontrivial nullspaces — that means that there are "free" directions in which we can vary $\vec{x}$ while not having to pay anything. How do we best take advantage of these "free" directions?

Parts (a-b) are connected; parts (c-d) are another group that can be done independently of (a-b); and parts (e-g) are another group that can be started independently of either (a-b) or (c-d). If you can't do a part, move on. **During debugging, many TAs found it easier to start with parts (c-g), and coming back to (a-b) at the end.**

(a) (8pts) Given a wide matrix A (with m columns and n rows) and a wide matrix C (with m columns and r rows), we want to solve:

$$\min_{\vec{x} \text{ such that } A\vec{x}=\vec{y}} ||C\vec{x}|| \tag{13}$$

As mentioned above, the key new issue is to isolate the "free" directions in which we can vary $\vec{x}$ so that they might be properly exploited. Consider the full SVD of $C = U\Sigma_c V^\top = \sum_{i=1}^{\ell} \sigma_{c,i}\vec{u}_i\vec{v}_i^\top$. Here, we write $V = [V_c, V_f]$ so that $V_c = [\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_\ell]$ all correspond to singular values $\sigma_{c,i} > 0$ of C, and $V_f = [\vec{v}_{\ell+1}, \cdots, \vec{v}_m]$ form an orthonormal basis for the nullspace of C.

Change variables in the problem to be in terms of $\vec{\widetilde{x}} = \begin{bmatrix} \vec{\widetilde{x}}_c \\ \vec{\widetilde{x}}_f \end{bmatrix}$ where the ℓ-dimensional $\vec{\widetilde{x}}_c$ has components $\widetilde{x}_c[i] = \alpha_i \vec{v}_i^\top \vec{x}$, and the $(m-\ell)$-dimensional $\vec{\widetilde{x}}_f$ has components $\widetilde{x}_f[i] = \vec{v}_{\ell+i}^\top \vec{x}$. In vector/matrix form, $\vec{\widetilde{x}}_f = V_f^\top \vec{x}$ and $\vec{\widetilde{x}}_c = \begin{bmatrix} \alpha_1 & 0 & \cdots & 0 \\ 0 & \alpha_2 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \alpha_\ell \end{bmatrix} V_c^\top \vec{x}$. Or directly, $\vec{\widetilde{x}} = \begin{bmatrix} \vec{\widetilde{x}}_c \\ \vec{\widetilde{x}}_f \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} \alpha_1 & 0 & \cdots & 0 \\ 0 & \alpha_2 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \alpha_\ell \end{bmatrix} V_c^\top \\ V_f^\top \end{bmatrix} \vec{x}$.

Express $\vec{x}$ in terms of $\vec{\widetilde{x}}_f$ and $\vec{\widetilde{x}}_c$. Assume the $\alpha_i \neq 0$ so the relevant matrix is invertible.

What is $||C\vec{x}||$ in terms of $\vec{\widetilde{x}}_f$ and $\vec{\widetilde{x}}_c$? Simplify as much as you can for full credit.

(HINT: If you get stuck on how to express $\vec{x}$ in terms of the new variables, think about the special case when $\ell = 1$ and $\alpha_1 = \frac{1}{2}$. How is this different from when $\alpha_1 = 1$? The SVD of C might be useful when looking at $||C\vec{x}||$.)

(b) (12pts) Continuing the previous part, **give appropriate values for the α_i so that the problem (13) becomes**

$$\min_{\vec{\tilde{x}}=\begin{bmatrix}\vec{\tilde{x}}_c\\ \vec{\tilde{x}}_f\end{bmatrix} \text{ such that } [A_c, A_f]\begin{bmatrix}\vec{\tilde{x}}_c\\ \vec{\tilde{x}}_f\end{bmatrix}=\vec{y}} \|\vec{\tilde{x}}_c\| \tag{15}$$

Give explicit expressions for A_c and A_f in terms of the original A and terms arising from the SVD of C. Because you have picked values for the α_i, there should be no α_i in your final expressions for full credit.

(HINT: How do the singular values $\sigma_{c,i}$ interact with the α_i? Then apply the appropriate substitution to (13) to get (15).)

(c) (5pts) Let us focus on a simple case. *(You can do this even if you didn't get the previous parts.)* Suppose that $A = [A_c, A_f]$ where the columns of A_f are orthonormal, as well as orthogonal to the columns of A_c. The columns of A together span the entire n-dimensional space. We directly write $\vec{x} = \begin{bmatrix} \vec{x}_c \\ \vec{x}_f \end{bmatrix}$ so that $A\vec{x} = A_c\vec{x}_c + A_f\vec{x}_f$. Now suppose that we want to solve $A\vec{x} = \vec{y}$ and only care about minimizing $\|\vec{x}_c\|$. We don't care about the length of $\vec{x}_f$ — it can be as big or small as necessary.

In other words, we want to:

$$\min_{\vec{x}=\begin{bmatrix} \vec{x}_c \\ \vec{x}_f \end{bmatrix} \text{ such that } [A_c, A_f]\begin{bmatrix} \vec{x}_c \\ \vec{x}_f \end{bmatrix} = \vec{y}} \|\vec{x}_c\| \tag{16}$$

Show that the optimal solution has $\vec{x}_f = A_f^\top \vec{y}$.

(HINT: Multiplying both sides of something by $A_f^\top$ might be helpful.)

(d) (8pts) Continuing the previous part, **compute the optimal $\vec{x}_c$.** Show your work.

(HINT: What is the work that $\vec{x}_c$ needs to do? $\vec{y} - A_f A_f^\top \vec{y}$ might play a useful role, as will the SVD of $A_c = \sum_i \sigma_i \vec{t}_i \vec{w}_i^\top$.)

(e) (5pts) Now suppose that A_c did not necessarily have its columns orthogonal to A_f. Continue to assume that A_f has orthonormal columns. *(You can do this part even if you didn't get any of the previous parts.)* Write the matrix $A_c = A_{c\perp} + A_{cf}$ where the columns of A_{cf} are all in the column span of A_f and the columns of $A_{c\perp}$ are all orthogonal to the columns of A_f. **Give an expression for A_{cf} in terms of A_c and A_f.**

(HINT: What does this have to do with projection and least squares?)

(f) (8pts) Continuing the previous part, **compute the optimal** $\vec{x}_c$ that solves (16): (copied below)

$$\min_{\vec{x}=\begin{bmatrix}\vec{x}_c\\ \vec{x}_f\end{bmatrix} \text{ such that } [A_c, A_f]\begin{bmatrix}\vec{x}_c\\ \vec{x}_f\end{bmatrix}=\vec{y}} \|\vec{x}_c\|$$

Show your work. Feel free to call the SVD as a black box as a part of your computation.
(HINT: What is the work that $\vec{x}_c$ needs to do? The SVD of $A_{c\perp}$ might be useful.)

(g) (6pts) Continuing the previous part, **compute the optimal $\vec{x}_f$.** Show your work.

You can use the optimal $\vec{x}_c$ in your expression just assuming that you did the previous part correctly, even if you didn't. You can also assume a decomposition $A_c = A_{c\perp} + A_{cf}$ from further above in part (e) without having to write what these are, just assume that you did them correctly, even if you didn't do them at all.

(HINT: What is the work that $\vec{x}_f$ needs to do? How is A_{cf} relevant here?

Contributors:

- Sidney Buchbinder.
- Pavan Bhargava.
- Anant Sahai.
- Regina Eckert.
- Kuan-Yun Lee.

