EE263: Introduction to Linear Dynamical Systems
Final Exam

Name: ____________________________________________

SUID: ____________________________________________

Please circle the appropriate option for each of the following:

Grading option: Letter grade Credit/No credit

Date: August 14 – 15 (5pm) August 15 – 16 (10am)

August 15 – 16 (5pm) Other (please specify):

I acknowledge and accept the Honor Code.

(Signed) __________________________________________

(For EE263 staff only)

<table>
<thead>
<tr>
<th>Problem</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>/20</td>
</tr>
<tr>
<td>2</td>
<td>/20</td>
</tr>
<tr>
<td>3</td>
<td>/20</td>
</tr>
<tr>
<td>4</td>
<td>/20</td>
</tr>
<tr>
<td>5</td>
<td>/20</td>
</tr>
<tr>
<td>Total</td>
<td>/100</td>
</tr>
</tbody>
</table>
This is a twenty-four-hour take-home final exam. If you are taking the exam in person, please submit your exam in Bytes Cafe (in the Packard building) twenty-four hours after you pick it up; if you are an SCPD student taking the exam remotely, please email your solution to SCPD distribution (scpd-distribution@lists.stanford.edu) twenty-four hours after you receive the exam; if you are a non-SCPD student taking the exam remotely, please email your solution to ee263hw@gmail.com. (All of your work, including code, must be in one file.)

- You may use any books, notes, or computer programs (such as MATLAB), but you may not discuss the exam with others until Tuesday August 18, after everyone has taken the exam. The only exception is that you can ask the course staff for clarification by emailing the staff email address: ee263-sum1415-staff@lists.stanford.edu. Do not ask about the exam questions on Piazza! However, we’ve tried pretty hard to make the exam unambiguous and clear, so we’re unlikely to say much.

- Since you have twenty-four hours, we expect your solutions to be legible, neat, and clear. Do not hand in your rough notes, and please try to simplify your solutions as much as you can. We will deduct points from solutions that are technically correct, but much more complicated than they need to be.

- Please check your email a few times during the exam, just in case we need to send out a clarification or other announcement. It’s unlikely we’ll need to do this, but you never know.

- Attach the official exam cover page to your exam, and assemble your solutions to the problems in order: that is, problem 1, problem 2, and so on. Start each problem on a new page. We will not go hunting for your work if the problems are not arranged in order; make sure that the code for each problem appears with your discussion and solution. (In particular, do not put all the code at the end.)

- Please make a copy of your exam before handing it in. We have never lost one, but it might occur.

- When a problem involves some computation, we do not want just the final answers. We want a clear discussion and justification of exactly what you did, the MATLAB source code that produces the result, and the final numerical result.

- Some of the problems are described in a practical setting, such as energy consumption, population dynamics, or wireless communications. You do not need to understand anything about the application area to solve these problems. We’ve taken special care to make sure all the information and math needed to solve the problem are given in the problem description.

- Some of the problems require you to download and run a MATLAB file to generate problem data. These files can be found at the URL [http://www.stanford.edu/class/ee263/final2015s/Filename](http://www.stanford.edu/class/ee263/final2015s/Filename)
where you should substitute the particular filename (given in the problem) for $\text{FILENAME}$. 

*There are no links on the course web page pointing to these files, so you'll have to type in the whole URL yourself.*
1 Optimal operation of a two-state chemical reactor

Consider a chemical reactor containing $n$ compounds, labeled $1, \ldots, n$. Let $x_i(t)$ be the amount of compound $i$ in the reactor at time $t$. The chemical reactor has two modes of operation, labeled 1 and 2. (For example, the first mode may be operating the reactor at a low temperature, and the second mode may be operating the reactor at a high temperature.) For simplicity we assume that the mode of operation can be changed instantaneously. When we operate the reactor in mode $j$, the vector of compound amounts evolves according to the equation

$$\dot{x}(t) = A_j x(t).$$

We are given the vector $x(0) \in \mathbb{R}^n$ of initial compound amounts, and the dynamics matrices $A_1$ and $A_2$. Our objective is to maximize the amount of compound $k$ at time $T$, where $k \in \{1, \ldots, n\}$ and $T > 0$ are given.

(a) Suppose the reactor operates in mode 1 for $0 \leq t \leq T_0$, and mode 2 for $T_0 < t \leq T$. Explain how to choose the time $T_0$ in order to maximize the amount of compound $k$ at time $T$. Your answer only needs to be accurate to two decimal digits.

(b) Apply your method to the data given in chemical_reactor_data.m. Report the optimal value of $T_0$ and the corresponding amount of compound $k$ at time $T$; submit a plot showing all of the components of $x(t)$ as functions of time on a single set of axes.

(c) Suppose the reactor operates in mode 1 for $0 \leq t \leq T_1$ and $T_2 < t \leq T$, and mode 2 for $T_1 < t \leq T_2$. Explain how to choose the times $T_1$ and $T_2$ in order to maximize the amount of compound $k$ at time $T$. Your answers for $T_1$ and $T_2$ only need to be accurate to two decimal digits.

(d) Apply your method to the data given in chemical_reactor_data.m. Report the optimal values of $T_1$ and $T_2$ and the corresponding amount of compound $k$ at time $T$; submit a plot showing all of the components of $x(t)$ as functions of time on a single set of axes.
2 Ranking teams in a round-robin tournament

Suppose $n$ teams play a round-robin tournament (that is, each pair of teams plays exactly one game). For concreteness consider a round-robin tournament with four teams and the following outcomes:

- team 1 beats teams 2 and 3 by 10 points each,
- team 1 loses to team 4 by 1 point,
- team 2 beats teams 3 and 4 by 1 point each, and
- team 3 beats team 4 by 1 point.

After the tournament we assign each team an overall score that we hope reflects how good the team is. A simple method for assigning these scores is to let the score for each team be the fraction of games that the team won. We let the vector of these scores be $z(0) \in \mathbb{R}^n$. For the example given above, we have that

$$z_1(0) = \frac{2}{3}, \quad z_2(0) = \frac{2}{3}, \quad z_3(0) = \frac{1}{3}, \quad \text{and} \quad z_4(0) = \frac{1}{3}.$$ 

These scores do not seem entirely satisfactory for our example because

- teams 1 and 2 have the same ranking, even though team 1 seemed to dominate the tournament except for a close loss against team 4;
- teams 3 and 4 have the same ranking, even though team 4 managed to beat the dominant team 1, and only suffered slight defeats to teams 2 and 3.

We can try to improve our scores by accounting for the quality of the opponents that a team defeated, and the margin by which a team beat its opponents. Define the matrix $A \in \mathbb{R}^{n \times n}$ such that

$$A_{ij} = \begin{cases} 
\text{the margin by which team } i \text{ defeated team } j & \text{if team } i \text{ defeated team } j, \\
0 & \text{otherwise}.
\end{cases}$$

For the simple example given above, we have that

$$A = \begin{bmatrix}
1 & 2 & 3 & 4 \\
1 & 0 & 10 & 10 \\
2 & 0 & 0 & 1 \\
3 & 0 & 0 & 0 \\
4 & 1 & 0 & 0
\end{bmatrix}.$$

We refine our scores by letting $z(1) = Az(0)$; the refined ranking of team $i$ is

$$z_i(1) = \sum_{j=1}^{n} A_{ij} z_j(0).$$
This says that the refined ranking of team $i$ is a weighted combination of the margins of victory of team $i$ with weights given by our simple estimates of the strengths of the teams that team $i$ defeated. (Note that this is a very simple refinement; in particular, we have not accounted for team $i$’s losses.) Our refined estimates for the example above are

$$z_1(1) = 10, \quad z_2(1) = \frac{2}{3}, \quad z_3(1) = \frac{1}{3}, \quad \text{and} \quad z_4(1) = \frac{2}{3}. $$

Note that these refined scores reflect the dominance of team 1, and the fact that team 4 has a reasonable case for being ranked above team 3. Now that we have more refined estimates of the strength of each team, we repeat the refinement process, and let $z(2) = Az(1)$. We continue this process using the recursion $z(t + 1) = Az(t)$. In practice we only care about the relative scores of the teams, and the computation of $z(t)$ may be numerically unstable for large values of $t$. Therefore, we prefer to use the normalized scores

$$x(t) = \frac{1}{\|z(t)\|}z(t), \quad t = 0, 1, 2, \ldots.$$

Note that $x(t)$ satisfies the recursion

$$x(t + 1) = \frac{1}{\|Az(t)\|}Ax(t),$$

which is numerically stable, even for large values of $t$. Now return to the general case of a round-robin tournament. Assume that $A$ is diagonalizable, and that its eigenvalue decomposition is

$$A = \sum_{i=1}^{n} \lambda_i v_i w_i^T,$$

where the eigenvalues are ordered such that $|\lambda_1| \geq \cdots \geq |\lambda_n|$, and $v_1, \ldots, v_n$ are normalized to be unit vectors.

(a) Let $\bar{x} = \lim_{t \to \infty} x(t)$. According to our usual analysis, what is the value of $\bar{x}$? State the usual assumptions that we use to justify our analysis.

(b) Let $e(t) = x(t) - \bar{x}$ be the difference between $x(t)$ and its limiting value. The asymptotic convergence rate is defined to be

$$\rho = \lim_{t \to \infty} \frac{\|e(t + 1)\|}{\|e(t)\|}.$$

According to our usual analysis, what is the value of $\rho$? State the usual assumptions that we use to justify our analysis.

(c) Now consider the specific data given in `nba_ranking_data.m`. This file defines the following variables.

- $n$, the number of teams in the National Basketball Association (NBA)
- $T$, the number of iterations of refinement to perform
• **teams**, a cell array of length \( n \) whose \( i \)th entry is the name of the \( i \)th NBA team

• **records**, an \( n \times 2 \) matrix whose \( i \)th row is the number of wins (first column) and losses (second column) of the \( i \)th NBA team in the 2014 – 2015 regular season

• **scores**, an \( n \times n \) matrix whose \( (i,j) \) entry is the average margin of victory of team \( i \) over team \( j \) in the 2014 – 2015 NBA regular season (this is negative if team \( j \) outscored team \( i \))

Form the matrix \( A \) such that

\[
A_{ij} = \begin{cases} 
\text{scores}(i,j) & \text{scores}(i,j) > 0, \\
0 & \text{otherwise},
\end{cases}
\]

the vector \( z(0) \) such that

\[
z_j(0) = \text{fraction of its games that team } j \text{ won},
\]

and the vector \( x(0) = z(0)/\|z(0)\| \).

(i) Compute \( x(t) \) for \( t = 1, \ldots, T \). Submit a plot showing \( x_j(t) \) for \( j = 10, 20, 30 \) as functions of \( t \) on a single set of axes. Compute \( \bar{x} \), and indicate \( \bar{x}_j \) for \( j = 10, 20, 30 \) on your plot.

(ii) Compute \( e(t) \) for \( t = 0, \ldots, T \). Submit a plot showing \( \|e(t)\| \) as a function of \( t \); use a logarithmic scale for the vertical axis, and a linear scale for the horizontal scale. Report the value of \( \rho \). How is the slope of your plot related to \( \rho \)?

(iii) Are the assumptions that we usually use to justify our analysis of the asymptotic behavior of \( x(t) \) satisfied? Does our usual conclusion hold?

(iv) Are the assumptions that we usually use to justify our analysis of the asymptotic behavior of \( \|e(t+1)\|/\|e(t)\| \) satisfied? Does our usual conclusion hold?

(v) Which team is most overrated if we use win/loss record instead of our refined scoring system? (In other words, which value of \( i \) maximizes \( x_i(0) - \bar{x}_i \)?)

(vi) Which team is most underrated if use win/loss record instead of our refined scoring system? (In other words, which value of \( i \) maximizes \( \bar{x}_i - x_i(0) \)?)
3 Minimum-sensitivity estimation of radiation levels

In the wake of a disaster at a nuclear power plant, you are interested in determining the radiation levels at the reactors. The power plant contained \( n \) reactors at known locations \((x_1, y_1), \ldots, (x_n, y_n) \in \mathbb{R}^2\); let \( r_j \in \mathbb{R} \) be the radiation level at the \( j \)th reactor. Due to the danger of radiation poisoning, you cannot take your measurements too close to the reactor sites. Let \((u_1, v_1), \ldots, (u_m, v_m) \in \mathbb{R}^2\) be candidate measurement locations on the border of the exclusion zone, where \( m > n \). The measurement at \((u_i, v_i)\) is

\[
w_i = \sum_{j=1}^{n} \frac{\alpha r_j}{\| (u_i, v_i) - (x_j, y_j) \|^2} + \delta w_i,
\]

where \( \alpha > 0 \) is a known proportionality constant, and \( \delta w_i \in \mathbb{R} \) is a measurement error. Intuitively, we assume that the radiation levels fall off according to an inverse-square law.

(a) You only have the resources to make \( n \) measurements. Let \( i_1, \ldots, i_n \) be the indices of the measurement locations that we choose. Let \( z = (w_{i_1}, \ldots, w_{i_n}) \) be our vector of measurements, and \( \delta z = (\delta w_{i_1}, \ldots, \delta w_{i_n}) \) be the corresponding vector of measurement errors. Estimating the radiation levels \( r_1, \ldots, r_n \) from these \( n \) measurements involves solving an \( n \times n \) system of linear equations. Let \( \delta r \) be the error in your estimate of \( r \). Explain how to choose \( n \) of the candidate measurement locations in order to minimize the sensitivity:

\[
\max_{r, z \in \mathbb{R}^n} \left\{ \frac{\| \delta r \|}{\| r \|} : \frac{\| \delta z \|}{\| z \|} \right\}
\]

of your estimate to the measurement errors.

(b) The file `nuclear_disaster_sensitivity_data.m` defines the following variables.

- \( n \) and \( m \), the numbers of reactor sites and candidate measurement locations, respectively
- `reactor_sites`, a \( 2 \times n \) matrix whose columns are the locations of the reactors
- `test_sites`, a \( 2 \times m \) matrix whose columns are the candidate measurement locations

Running `nuclear_disaster_sensitivity_data.m` also generates a plot showing the reactor sites as red \( \times \)s, and the candidate measurement locations as blue circles. Find the \( n \) measurement locations among the candidate measurement locations that minimize the sensitivity of your estimate of the radiation levels to the measurement errors. Report the indices of your chosen measurement locations (that is, which columns of `test_sites` you choose), and the corresponding sensitivity of your estimate to the measurement errors. Submit a plot indicating the chosen measurement locations with green squares on the plot generated by the data file.
4 Principal-components analysis of decathlon data

In the decathlon athletes compete in \( n = 10 \) different track-and-field events: 100 m run, long jump, shot put, high jump, 400 m run, 110 m hurdle, discus, pole vault, javelin, and 1500 m run. The file `decathlon_pca_data.m` contains the results for the \( m = 24 \) athletes who completed all of the events in the decathlon in the 2008 Olympics; in particular, it defines the following variables.

- \( m \) and \( n \), the numbers of athletes and events, respectively
- `names`, a cell array containing the last names and countries of the athletes
- `events`, a cell array containing the names of the events
- `scores`, a matrix containing the score of each athlete on each event; each row corresponds to an athlete, and each column corresponds to an event; the orderings of athletes and events are the same as in `names` and `events`, respectively

First, we standardize the data by forming the matrix \( X \in \mathbb{R}^{m \times n} \) such that

\[
X_{ij} = \frac{\text{scores}(i,j) - m_j}{s_j},
\]

where

\[
m_j = \frac{1}{m} \sum_{i=1}^{m} \text{scores}(i,j) \quad \text{and} \quad s_j = \sqrt{\frac{1}{m-1} \sum_{i=1}^{m} (\text{scores}(i,j) - m_j)^2}
\]

are, respectively, the mean and standard deviation of the scores for the \( j \)th event.

(a) Let \( X = \sum_{i=1}^{m} \sigma_i u_i v_i^T \) be the singular-value decomposition of \( X \). Submit a plot showing \( \sigma_j \) versus \( j \). The fraction of the total power that is contained in the first \( j \) dyads of the singular-value decomposition is defined to be

\[
p_j = \frac{\sum_{j=1}^{j} \sigma_j^2}{\sum_{j=1}^{n} \sigma_j^2}.
\]

Submit a plot of \( p_j \) versus \( j \), and report \( p_2 \), the fraction of the total power that is contained in the first two dyads.

(b) Use the command `text` to create a plot in which the point \((v_1)_j, (v_2)_j\) is labeled with the name of the \( j \)th event. Do similar events appear to be close to each other on your plot?

(c) Let \( t_i = \sum_{j=1}^{n} X_{ij} \) be the total standardized score of the \( i \)th athlete. Then, we have that

\[
r_j = \sum_{i=1}^{m} t_i X_{ij}
\]

is proportional to the correlation between an athlete’s total standardized score and his standardized score on the \( j \)th event. Download the function `spatial_plot` from the course website. Let \( v_1 \) and \( v_2 \) be the first two right singular vectors of \( X \). Submit a copy of the plot generated using the following command.
Which right singular vector seems to represent $r$?

(d) Let $u_1$ and $u_2$ be the first two left singular vectors of $X$. Define the vector $\delta \in \mathbb{R}^m$

$$\delta_i = (X_{i1} + X_{i5} + X_{i6}) - (X_{i3} + X_{i7} + X_{i9}),$$

and let $t = (t_1, \ldots, t_m)$ be the vector of the total standardized scores of the athletes. Submit copies of the plots generated using the following commands.

```
spatial_plot(u1 , u2 , t , length(names));
spatial_plot(u1 , u2 , delta , length(names));
```

Give intuitive interpretations of the first two left singular vectors of $X$. 
5 Stock-market prediction using maximum-correlation linear functions

Let \( x(1), \ldots, x(T) \in \mathbb{R}^n \) and \( y(1), \ldots, y(T) \in \mathbb{R}^m \) be two discrete-time vector signals. Given constant vectors \( \alpha \in \mathbb{R}^n \) and \( \beta \in \mathbb{R}^m \), define the discrete-time scalar signals

\[
    u(t) = \alpha^T x(t) \quad \text{and} \quad v(t) = \beta^T y(t).
\]

The correlation of \( u(t) \) and \( v(t) \) is defined to be

\[
    r(u, v) = \frac{1}{T} \sum_{t=1}^{T} (u(t) - \bar{u})(v(t) - \bar{v}),
\]

where \( \bar{u} = \frac{1}{T} \sum_{t=1}^{T} u(t) \) and \( \bar{v} = \frac{1}{T} \sum_{t=1}^{T} v(t) \) are the means of \( u(t) \) and \( v(t) \), respectively. Intuitively, we think of \( r(u, v) \) as a measure of the strength of the linear relationship between \( u(t) \) and \( v(t) \) – we have that \( |r(u, v)| \) is close to one when there is a strong (linear) relationship, and close to zero when the (linear) relationship between \( u(t) \) and \( v(t) \) is weak. Since \( r(u, v) \) does not depend on the scalings of \( \alpha \) and \( \beta \), we standardize the problem by taking \( \|\alpha\| = \|\beta\| = 1 \).

(a) Define the matrices \( \tilde{X} \in \mathbb{R}^{n \times T} \) and \( \tilde{Y} \in \mathbb{R}^{m \times T} \) such that

\[
    \tilde{X} = [x(1) \cdots x(T)] \left( I - \frac{1}{T} 1 1^T \right) \quad \text{and} \quad \tilde{Y} = [y(1) \cdots y(T)] \left( I - \frac{1}{T} 1 1^T \right).
\]

Intuitively, \( \tilde{X} \) is the matrix whose columns are \( x(t) - \bar{x} \), where \( \bar{x} = \frac{1}{T} \sum_{t=1}^{T} x(t) \) is the sample mean of \( x(t) \); \( \tilde{Y} \) has a similar interpretation. Show that

\[
    r(u, v) = \frac{\alpha^T \tilde{X} \tilde{Y}^T \beta}{\| \tilde{X} \tilde{Y}^T \| \| \tilde{Y} \|}.
\]

(b) Suppose \( \tilde{X} \) and \( \tilde{Y} \) are both fat and full rank. Define \( a \in \mathbb{R}^n \) and \( b \in \mathbb{R}^m \) such that

\[
    a = (\tilde{X} \tilde{X}^T)^{-\frac{1}{2}} \alpha \quad \text{and} \quad b = (\tilde{Y} \tilde{Y}^T)^{-\frac{1}{2}} \beta.
\]

Then, we have that

\[
    r(u, v) = \frac{a^T P b}{\|a\| \|b\|},
\]

where

\[
    P = (\tilde{X} \tilde{X}^T)^{-\frac{1}{2}} \tilde{X} \tilde{Y}^T (\tilde{Y} \tilde{Y}^T)^{-\frac{1}{2}}.
\]

Explain how to choose \( a \) and \( b \) in order to maximize \( r(u, v) \). Then, explain how to find unit vectors \( \alpha \) and \( \beta \) that maximize \( r(u, v) \).

(c) Consider a market with \( n \) assets at times \( t = 1, \ldots, T + 1 \). Let \( z_i(t) \) be the change in the price of asset \( i \) in time period \( t \) for \( i = 1, \ldots, n \) and \( t = 1, \ldots, T + 1 \). Define

\[
    x(t) = z(t) \quad \text{and} \quad y(t) = z(t+1), \quad t = 1, \ldots, T.
\]
If we choose $\alpha$ and $\beta$ such that $u(t) = \alpha^T x(t)$ and $v(t) = \beta^T y(t)$ are positively correlated, then we have a method for trading assets in the market. At time $t$ we compute $u(t) = \alpha^T x(t) = \alpha^T z(t)$. If $u(t) > 0$, then we expect $v(t) = \beta^T y(t) = \beta^T z(t + 1)$ to be positive, so we buy a portfolio with weights given by $\beta$. (Note that a negative weight indicates a short position.) Similarly, if $u(t) < 0$, then we expect $v(t)$ to be negative, so we sell a portfolio with weights given by $\beta$. We make money if $\text{sgn}(u(t)) = \text{sgn}(v(t))$, and lose money otherwise. The file `stock_prediction_maximum_correlation_data.m` defines the following variables.

- $n$, the number of assets in the market
- $T$, the length of the data set
- $Z$, a matrix whose columns are $z(1), \ldots, z(T + 1)$

Compute the vectors $\alpha$ and $\beta$ that maximize $r(u, v)$. Report your values of $\alpha$ and $\beta$, and the corresponding value of $r(u, v)$.

(d) The file `stock_prediction_maximum_correlation_data.m` also defines the variables

- $T_{\text{test}}$, the length of the test data set, and
- $Z_{\text{test}}$, a matrix whose columns are $z_{\text{test}}(1), \ldots, z_{\text{test}}(T_{\text{test}} + 1)$.

This is a different set of data from the same market. Define $x_{\text{test}}(t) = z_{\text{test}}(t)$ and $y_{\text{test}}(t) = z_{\text{test}}(t + 1)$ for $t = 1, \ldots, T_{\text{test}}$. Using the values of $\alpha$ and $\beta$ that you found above, compute the fraction of the time that $\text{sgn}(u_{\text{test}}(t)) = \text{sgn}(v_{\text{test}}(t))$, where $u_{\text{test}}(t) = \alpha^T x_{\text{test}}(t)$ and $v_{\text{test}}(t) = \beta^T y_{\text{test}}(t)$. This is the fraction of the time that we make money on the test set using the trading strategy based on $\alpha$ and $\beta$.

Consider a simple index fund that places the same weight on each asset. Suppose we buy the index fund if its price increases in the current time period, and sell (or short) the index fund otherwise. We make money using this strategy in time period $t + 1$ if $\text{sgn}(1^T x_{\text{test}}(t)) = \text{sgn}(1^T y_{\text{test}}(t))$, and lose money otherwise. Compute the fraction of the time that the simple trading strategy makes money.