13.1350. Drone delivery. We want a drone to visit $m$ positions $p_1, \ldots, p_m \in \mathbb{R}^2$ at times $s_1, \ldots, s_m \in \mathbb{R}_+$. Define the matrix $P \in \mathbb{R}^{2 \times m}$ to have columns $p_j$, where $j = 1, \ldots, m$. The following Julia code (copy and paste works) defines these:

\[
P = \begin{bmatrix}
-1 & -0.5 & 0 & 0.5 & 1 & 0 \\
2 & 0 & 1 & 0 & 2 & 0
\end{bmatrix}; \ s = [3; 7; 10; 13; 17; 20];
\]

a) The dynamics of the drone are

\[
\ddot{q} = u,
\]

where $q : \mathbb{R} \to \mathbb{R}^2$ is the position of the drone as a function of time, and $u : \mathbb{R} \to \mathbb{R}^2$ is an input force as a function of time. Write this as a linear dynamical system of the form

\[
\dot{x} = Ax + Bu, \quad y = Cx,
\]

where $y : \mathbb{R} \to \mathbb{R}^4$ is the position and velocity of the drone. In particular, define $y$ to satisfy $y_1 = q_1, y_2 = q_2, y_3 = \dot{q}_1, y_4 = \dot{q}_2$. Is the system stable? Is the system controllable?

b) We will discretize this system with a sample interval $h > 0$. Assume the force is piecewise constant on sample intervals. Construct the exact discretization

\[
x_d(k + 1) = A_dx_d(k) + B_du_d(k), \quad y_d(k) = C_dx_d(k),
\]

where $x_d : \mathbb{Z}_+ \to \mathbb{R}^n$ satisfies $x_d(k) = x(kh)$, and likewise for $y_d$ and $u_d$. Give $A_d, B_d$ and $C_d$.

c) We operate the drone on the time interval $[0, T]$ where $T = s_m$. Define $n = T/h$. Assume the drone starts at the origin with velocity zero. Explain how to choose forces to steer the drone through the desired positions at the desired times, i.e. $y_{1,2}(s_i) = p_i$, while minimizing

\[
J_1 = \sum_{k=0}^{n-1} \|u_d(k)\|_2^2.
\]

The drone need not be stationary when passing through the points $p_1, \ldots, p_m$.

d) With $h = 0.1$, use your method to compute the optimal $u_d$. Report the optimal value of $J_1$ that you obtained. Plot the components of $u$ with respect to time. Plot the trajectory of the drone with axes $q_1$ and $q_2$, so that the plot shows the path followed by the drone. Mark on your plot the points $p_i$ where the drone has deliveries.

e) Suppose we want to find forces minimizing $J_1$ so that the drone is also stationary (velocity zero) when it is at position $i$, in order to make a clean drop. Explain how to do this. Report the optimal cost. Plot the trajectory and control inputs as before. (Even if you cannot solve this part, you may complete (f) without the condition that the drone is stationary.)
f) Suppose that, in addition to being stationary, we want to avoid the drone jerking too much, in order to avoid damaging its payload. We penalize the discrete jerk. Define
\[ J_2 = \sum_{k=1}^{n-1} \| u_d(k) - u_d(k-1) \|_2^2. \]

Explain how to find forces to steer the drone to be at the positions \( p_1, \ldots, p_m \) with velocity zero at the desired times \( s_1, \ldots, s_m \), while minimizing \( J_1 + \mu J_2 \), where \( \mu > 0 \) is given.

For \( \mu = 100 \), report \( J_1 \) and \( J_2 \). Plot the trajectory and controls as before.

Make a trade-off curve with \( J_2 \) on the horizontal axis, and \( J_1 \) on the vertical. Briefly interpret the endpoints.

15.2200. Properties of symmetric matrices. In this problem \( P \) and \( Q \) are symmetric matrices. For each statement below, either give a proof or a specific counterexample.

a) If \( P \geq 0 \) then \( P + Q \geq Q \).
b) If \( P \geq Q \) then \( -P \leq -Q \).
c) If \( P > 0 \) then \( P^{-1} > 0 \).
d) If \( P \geq Q > 0 \) then \( P^{-1} \leq Q^{-1} \).
e) If \( P \geq Q \) then \( P^2 \geq Q^2 \).

Hint: you might find it useful for part (d) to prove \( Z \geq I \) implies \( Z^{-1} \leq I \).

15.2301. (Optional) Matrix norms and singular values.

a) Eigenvalues and singular values of a symmetric matrix. Suppose \( A \in \mathbb{R}^{n \times n} \) with \( A = A^\top \).

Let \( \lambda_1, \ldots, \lambda_n \) be the eigenvalues of \( A \), and assume that the eigenvalues are ordered such that \( |\lambda_1| \geq \cdots \geq |\lambda_n| \). Let \( \sigma_1, \ldots, \sigma_n \) be the singular values of \( A \); by definition the singular values are ordered such that \( \sigma_1 \geq \cdots \geq \sigma_n \geq 0 \). How are the eigenvalues and singular values of \( A \) related?

b) Suppose \( X \in \mathbb{R}^{n \times n} \). Is \( \sigma_{\max}(X) \geq \max_{1 \leq i \leq n} \sqrt{\sum_{1 \leq j \leq n} |X_{ij}|^2} \)? Prove or give a counterexample.

c) Suppose \( X \in \mathbb{R}^{n \times n} \). Is \( \sigma_{\min}(X) \geq \min_{1 \leq i \leq n} \sqrt{\sum_{1 \leq j \leq n} |X_{ij}|^2} \)? Prove or give a counterexample.

d) Suppose \( X \in \mathbb{R}^{n \times n} \). Is \( \sigma_{\max}(XY) \leq \sigma_{\max}(X)\sigma_{\max}(Y) \)? Prove or give a counterexample.

e) Suppose \( X, Y \in \mathbb{R}^{n \times n} \). Is \( \sigma_{\min}(XY) \geq \sigma_{\min}(X)\sigma_{\min}(Y) \)? Prove or give a counterexample.

f) Suppose \( X, Y \in \mathbb{R}^{n \times n} \). Is \( \sigma_{\min}(X + Y) \geq \sigma_{\min}(X) - \sigma_{\max}(Y) \)? Prove or give a counterexample.
g) Recall that the Frobenius norm of a matrix $A \in \mathbb{R}^{m \times n}$ is defined to be

$$\|A\|_F = \sqrt{\text{trace}(A^T A)}.$$ 

Show that

$$\|A\|_F = \left( \sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}|^2 \right)^{\frac{1}{2}}.$$ 

Thus, the Frobenius norm is simply the Euclidean norm of a matrix, when we think of the matrix as an element of $\mathbb{R}^{mn}$. Additionally, note that the Frobenius norm is much easier to compute than the spectral norm.

h) Show that if $U \in \mathbb{R}^{m \times m}$ and $V \in \mathbb{R}^{n \times n}$ are orthogonal, then

$$\|UA\|_F = \|AV\|_F = \|A\|_F.$$ 

Thus, multiplication by orthogonal matrices on the left or right does not change the Frobenius norm.

i) Show that

$$\|A\|_F = \sqrt{\sigma_1^2 + \cdots + \sigma_r^2},$$ 

where $\sigma_1, \ldots, \sigma_r$ are the nonzero singular values of $A$. Use this result to deduce that

$$\sigma_{\text{max}}(A) \leq \|A\|_F \leq \sqrt{r}\sigma_{\text{max}}(A).$$ 

In particular, we have that $\|Ax\| \leq \|A\|_F \|x\|$ for all $x \in \mathbb{R}^n$.

16.2980. (Optional) Smoothing. We have a discrete-time signal given by $x \in \mathbb{R}^n$. We get to measure $y \in \mathbb{R}^n$, given by

$$y_i = \sum_{k=-h}^{h} c_k x_{i+k} + w_i \quad \text{for } i = 1, \ldots, n$$

where $w_i$ is noise. Here we use the convention that $x_i = 0$ for $i < 1$ or $i > n$. That is, $y$ is $c$ convolved with $x$ plus noise. In applications, very often the effect of convolution with $c$ is to smooth or blur $x$, and we would like to undo this.

The file `regl_data.json` contains $c$, $w$ and $x$.

a) In Julia, construct the $n \times n$ matrix such that $y = Ax + w$. Plot the singular values $\sigma_k$ against $k$.

b) Plot the first 6 right singular vectors of $A$ (i.e. plot $V_{ij}$ against $i$ for $j = 1, \ldots, 6$.) Explain what you see.

c) Find and plot the least-squares estimate of $x$ given $y_{\text{meas}}$, computng $y_{\text{meas}}$ using $c$, $x$ and $w$ given in `regl_data.json`. Explain what happens.
d) Many of the singular values of $A$ are very small; this means that the measurement in the directions of the corresponding right singular vectors is being swamped by the noise. If we believe these components are small, we can remove them from our estimate of $x$ altogether by truncating the SVD of $A$ and using the truncated SVD to compute the estimate. This is called the \textit{truncated SVD regularization} of least-squares.

Suppose we decided only to keep the first $r$ components. Then truncate by letting $\tilde{V}$ and $\tilde{U}$ be the first $r$ columns of $V$ and $U$, and letting $\tilde{\Sigma}$ be the top-left $r \times r$ submatrix of $\Sigma$. Then we can construct an estimator that ignores the noise components by

$$A_{\text{est}} = \tilde{V}\tilde{\Sigma}^{-1}\tilde{U}^T$$

and set

$$x_{\text{est}} = A_{\text{est}}y_{\text{meas}}$$

For values of $r$ in $5, 10, 15, 30, 50$, compute and plot the corresponding estimates of $x$. Explain what you see.

e) For each $r$ between 1 and 35, compute the norm of the error

$$\|x - x_{\text{est}}\|$$

Plot this against $r$. Explain what you see.

f) Pick the ‘best’ $r$ and plot the corresponding estimate.

g) Another approach is to use Tychonov regularization. Find and plot the vector $x_{\text{reg}} \in \mathbb{R}^n$ that minimizes the function

$$\|Ax - y\|^2 + \mu\|x\|^2,$$

where $\mu > 0$ is the regularization parameter. Pick a value of $\mu$ that gives a good estimate, in your opinion.

h) The regularized solution is a linear function of $y$, so it can be expressed as $x_{\text{reg}} = By$ where $B \in \mathbb{R}^{n \times n}$. Express the SVD of $B$ in terms of the SVD of $A$. To be more specific, let

$$B = \sum_{i=1}^{n} \tilde{\sigma}_i \tilde{u}_i \tilde{v}_i^T$$

denote the SVD of $B$. Express $\tilde{\sigma}_i$, $\tilde{u}_i$, $\tilde{v}_i$ for $i = 1, \ldots, n$, in terms of $\sigma_i$, $u_i$, $v_i$, $i = 1, \ldots, n$ (and, possibly, $\mu$). Recall the convention that $\tilde{\sigma}_1 \geq \cdots \geq \tilde{\sigma}_n$.

i) Find the norm of $B$. Give your answer in terms of the SVD of $A$ (and $\mu$).

j) Find the worst-case relative inversion error, defined as

$$\max_{y \neq 0} \frac{\|ABy - y\|}{\|y\|}.$$

Give your answer in terms of the SVD of $A$ (and $\mu$).
18.2890. Linear dynamical systems for portfolio management. We consider a portfolio of $n$ financial assets (like stocks) and cash, which we manage over $T$ time steps of unit length (e.g. one month). We call $x_t \in \mathbb{R}^{n+1}$ for $t = 1, \ldots, T$ our state vector. The first $n$ elements are our positions in each of the assets, in dollars, and the last element is the dollar amount of cash we hold. Every element of $x$ can be either positive (for long positions) and negative (for short or borrowing). For $t = 1, \ldots, T - 1$, the transition from $x_t$ to $x_{t+1}$ is composed of two steps.

- First, the portfolio positions change value because of market returns. Let $\mu \in \mathbb{R}^n_{++}$ be the vector of returns, where $\mathbb{R}^n_{++}$ is the set of all vectors of length $n$ with strictly positive entries. We define the post-return portfolio $\tilde{x}_t$ to be

$$\tilde{x}_t = \begin{cases} 
\mu_i (x_t)_i & i = 1, \ldots, n, \\
(x_t)_i & i = n + 1.
\end{cases}$$

(Intuitively, cash is unchanged, and the asset positions are multiplied by the corresponding element of the vector of returns.) For simplicity we assume that the vector of returns does not change in time.

- Then we trade. We can exchange any amount of cash for the corresponding amount of any of the assets. Note that the only valid trades are cash for asset. If you wish to trade some amount of an asset with the same amount of another asset, you have to perform two trades: trade the first asset with cash, and then trade cash with the second asset. (Think carefully about this definition of trade when you formulate the transaction costs.) For example, if we buy $c > 0$ dollars of the first asset and sell $d > 0$ dollars of the second asset the state evolves as

$$x_{t+1} = \tilde{x}_t + \begin{bmatrix} c \\
-d \\
0 \\
\vdots \\
0 \\
-(c-d)
\end{bmatrix}.$$ 

Finally, we define the portfolio value $v_t \in \mathbb{R}$ for $t = 1, \ldots, T$ to be

$$v_t = 1^T x_t.$$ 

a) Formulate the problem as a linear dynamical system of the form

$$x_{t+1} = Ax_t + Bu_t, \quad t = 1, \ldots, T - 1.$$ 

The control vector $u_t$ should have dimension $n$.

b) Assume that our trades incur quadratic transaction costs with parameter $\rho > 0$. For example, if at time $t$ we buy $c > 0$ dollars of the first asset, and we sell $d > 0$ dollars of the second asset (the example above), then the transaction costs for the transition $x_t$ to $x_{t+1}$ are

$$\rho(c^2 + d^2).$$
(Be careful, they are not $\rho(c + d)^2$.) Explain how to solve the problem of maximizing the final value of the portfolio $v_T$ minus the total transaction costs. (The sequence of controls $u_1, \ldots, u_{T-1}$ that achieves the maximum should be a function of $A$, $B$, and $\rho$). Use methods from EE263.

c) Apply your method to the following data.

\[
\begin{align*}
T &= 12; \\
x_1 &= [1000, 1000, 0, 1000, 0, 0]; \\
\mu &= [1.001, 1.003, 1.004, 1.006, 1.007]; \\
\rho &= 0.0001;
\end{align*}
\]

What is the final value $v_T$? What are the total transaction costs? Plot the trajectories of the portfolio positions $x_t$. (On the same plot you should draw $n + 1$ lines, one for each of the assets and cash, with time on the $x$-axis.)

d) Now assume that we aim to liquidate an initial portfolio, which means that at time $T$ we want to have zero positions in any of the $n$ assets and only hold cash. We thus impose the constraint $(x_T)_i = 0$, for $i = 1, \ldots, n$. Explain how to solve the problem of maximizing the final portfolio value (in this case, all cash) minus the transaction costs with this additional constraint. Use methods from EE263.

e) Apply your method to the data given above. What is the final value $v_T$? What are the total transaction costs? Plot the trajectories of the portfolio positions $x_t$. (On the same plot you should draw $n + 1$ lines, one for each of the assets and cash, with time on the $x$-axis.)