This homework is due April 25, 2016, at Noon.

1. Homework process and study group

(a) Who else did you work with on this homework? List names and student ID’s. (In case of hw party, you can also just describe the group.)

_Solution:_ I worked on this homework with...

(b) How long did you spend working on this homework? How did you approach it?

_Solution:_ I spent a total of X hours working on this assignment.
I first worked by myself for 2 hours, but got stuck on Problem 5 so I went to office hours on...
Then I went to homework party for a few hours, where I finished the homework.

_For self-grading, be sure to enter the number of hours alone in the comment field for this problem. This will help us better automatically collect this information._

2. Lecture Attendance

Did you attend live lecture this week? (the week you were working on this homework) What was your favorite part? Was anything unclear? Answer for each of the subparts below. If you only watched on YouTube, write that for partial credit.

(a) Monday lecture
(b) Wednesday lecture
(c) Friday lecture

_Solution:_ Full credit for attending live lecture and giving a comment (what you liked best, what was unclear) about that lecture. 8 points for attending live lecture but giving no comment. 5 points for watching on YouTube and giving a comment. 2 points for just watching on YouTube. 0 points for blank or not watching lecture at all.

3. Tracking a Desired Trajectory in Continuous Time

Many problems in 16AB so far have treated closed-loop control as being about holding a system steady at some desired operating operating point, which was designated as zero in a linear model. This control used the actual current state (and in principle can use an estimate of the state from an observer) to apply a control signal designed to bring the state to zero. Meanwhile, the idea of controllability itself was more general and allowed us to make an open-loop trajectory that went pretty much anywhere. This problem is about combining these two ideas together to make feedback control more practical — how can we get a system to more-or-less closely follow a desired trajectory, even though it might not start exactly where we wanted to start and in principle could be buffeted by small disturbances throughout.

The key conceptual idea is to realize that we can change coordinates in a time-varying way so that “zero” is the desired “open-loop” trajectory.
Consider a linear continuous-time system below ($\vec{x} \in \mathbb{R}^n$, $u \in \mathbb{R}$).

$$\dot{\vec{x}} = A\vec{x}(t) + Bu(t)$$

where $A \in \mathbb{R}^{n \times n}$, and $B \in \mathbb{R}^{n \times 1}$.

Here we use the physics-style Newton convention of denoting time derivatives by placing dots above variables.

(a) What is the condition for the system to be controllable?

**Solution:** As taught in lecture, by analogy to the condition discrete time system, the condition for the system to be controllable is to ensure that the controllability matrix, $C = [B \ AB \ \cdots \ A^{n-1}B]$, is full rank.

Now, consider the system

$$\dot{\vec{x}} = A\vec{x}(t) + Bu(t) = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix} \vec{x}(t) + \begin{bmatrix} 1 \\ 1 \end{bmatrix} u(t) \quad (1)$$

(b) Is the given system controllable?

**Solution:** By substituting the matrix $A$ and $B$ into the controllability matrix, we have:

$$C = [B \ AB] = \begin{bmatrix} 1 & 3 \\ 1 & 2 \end{bmatrix} \quad \text{full rank}$$

Since $C$ is full rank, the system is controllable, which means that we can navigate the system state to any desired state by choosing an input trajectory $u(t)$.

(c) Now, consider that we supposed that we started at $\vec{x}(0) = \vec{0}$ and had a nominal control signal $u_n(t)$ that would make the system follow the desired trajectory $\vec{x}_n(t)$ that satisfies (1) together with $u_n(t)$.

Change variables to $\vec{x}(t) = \vec{x}_n(t) + \vec{v}(t)$ and $u(t) = u_n(t) + u_v(t)$ and write out what (1) implies for the evolution of the trajectory deviation $\vec{v}(t)$ as a function of the control deviation $u_v(t)$.

**Solution:**

As indicated by the problem, we are provided with control signal $u_n(t)$ such that the system starting at $\vec{x}(0) = \vec{0}$ is able to follow a desired trajectory:

$$\dot{\vec{x}}_n(t) = A\vec{x}_n(t) + Bu_n(t) \quad (2)$$

To obtain the evolution of $\vec{v}(t)$, we substitute the change of variables into the original state equation:

$$\dot{\vec{x}}(t) = A\vec{x}(t) + Bu(t)$$
$$\dot{\vec{x}}_n(t) + \dot{\vec{v}}(t) = A\vec{x}_n(t) + A\vec{v}(t) + Bu_n(t) + Bu_v(t)$$
$$\dot{\vec{v}}(t) = A\vec{v}(t) + Bu_v(t)$$

where $A = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}$, $B = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, and the third equation is obtained by substituting (2). This describes the evolution of the trajectory deviation $\vec{v}(t)$ as a function of the control deviation $u_v(t)$.
(d) Now, add a bounded disturbance term \( \mathbf{w}(t) \) to the original state evolution in (1) and see if you can absorb that entirely within an evolution equation for \( \mathbf{v}(t) \) based on \( u_v(t) \). Write out the resulting equation for the dynamics as:

\[
\dot{\mathbf{v}} = A_v \mathbf{v} + B_v u_v + \mathbf{w}
\]

What are \( A_v \) and \( B_v \)?

**Solution:**

The result can be obtained with a similar procedure as the previous part (c). This time, the system equation is given by:

\[
\dot{\mathbf{x}}(t) = A \mathbf{x}(t) + B u(t) = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix} \mathbf{x}(t) + \begin{bmatrix} 1 \\ 1 \end{bmatrix} u(t) + \mathbf{w}(t)
\]

(4)

By doing the changes of variables \( \mathbf{x}(t) = \mathbf{x}_n(t) + \mathbf{v}(t) \) and \( u(t) = u_n(t) + u_v(t) \) and (2) into the equation, we have:

\[
\dot{\mathbf{x}}(t) = A \mathbf{x}(t) + B u(t) + \mathbf{w}(t) \\
\dot{\mathbf{x}}_n(t) + \dot{\mathbf{v}}(t) = A \mathbf{x}_n(t) + A \mathbf{v}(t) + Bu_n(t) + Bu_v(t) + \mathbf{w}(t)
\]

Therefore, the resulting equation is given by

\[
\dot{\mathbf{v}}(t) = A_v \mathbf{v}(t) + B_v u_v(t) + \mathbf{w}(t).
\]

By comparison, we can see that \( A_v = A, B_v = B \).

Notice that the disturbance \( \mathbf{w}(t) \) is entirely something that must be dealt with in the \( \mathbf{v} \) dynamics. It doesn’t effect the nominal trajectory at all.

(e) Based on what you have found above, how will the system behave over time? If there is some small disturbance, will we end up following the intended trajectory \( \mathbf{x}_n(t) \) closely if we just apply the control \( u_n(t) \)?

**Solution:**

The key is to study the system equation for \( \mathbf{v} \), given by:

\[
\dot{\mathbf{v}}(t) = A_v \mathbf{v}(t) + \mathbf{w}(t)
\]

where \( u_v(t) = 0 \) as indicated in the problem, i.e., just apply the control \( u(t) = u_n(t) \).

If the state \( \mathbf{v}(t) \) fluctuates around \( \mathbf{0} \) under the small disturbance without any control, then we can end up following the intended trajectory \( \mathbf{x}_n(t) \).

Our task, therefore, is to study the stability of the system.

We know that if the eigenvalues are unstable (in the continuous-time case as here, this means that if any of them have a real part that is not strictly negative), then over time the state \( \mathbf{v} \) can grow without bound. This means that the system is getting further and further away from the original desired trajectory. Consequently, for stability reasons, we want the eigenvalues to have strictly negative real parts.

Since \( A_v \) is a upper-triangular matrix, its eigenvalues lie on the diagonal, namely, 2 and 2. In this case, since they clearly have real parts greater than zero, we can see that the system is vulnerable to any disturbances \( \mathbf{w}(t) \), and we will not end up following the intended trajectory \( \mathbf{x}_n(t) \).
Now, we want to apply state feedback control to the system to get it to more or less follow the desired trajectory.

(f) Just looking at the $\vec{v}(t), u_v(t)$ system, how would you apply state-feedback to choose $u_v(t)$ as a function of $\vec{v}(t)$ that would place both the eigenvalues of the closed-loop $\vec{v}(t)$ system at $-10$.

**Solution:**

We can assume that the input $u_v = [f_0 \quad f_1] \vec{v}$, which is a linear function of the state $\vec{v}$. With the new input, the system equation for $\vec{v}$ without any disturbance is given by:

$$\dot{\vec{v}}(t) = A_{cl} \vec{v}(t) + B [f_0 \quad f_1] \vec{v}$$

where we denote $A_{cl} = \begin{bmatrix} 2 + f_0 & 1 + f_1 \\ f_0 & 2 + f_1 \end{bmatrix}$ as the state matrix for the closed loop system. The characteristic polynomial for finding the eigenvalues of $A_{cl}$ is given by:

$$\det(\lambda I - A_{cl}) = \begin{vmatrix} \lambda - 2 - f_0 & -1 - f_1 \\ -f_0 & \lambda - 2 - f_1 \end{vmatrix} = \lambda^2 - (4 + f_0 + f_1)\lambda + f_0 + 2f_1 + 4$$

To set the eigenvalues to be where we want, we set this equal to $(\lambda + 10)(\lambda + 10) = \lambda^2 + 20\lambda + 100$.

By comparing the coefficients, we have:

$$-(4 + f_0 + f_1) = 20$$
$$f_0 + 2f_1 + 4 = 100$$

Solving the above system of equations, we can find $f_0 = -144, f_1 = 120$; therefore, we can design the state-feedback $u_v(t) = [-144 \quad 120] \vec{v}$ which will place both the eigenvalues of the closed loop system at -10.

(g) Based on what you did in the previous parts, and given access to the desired trajectory $\vec{x}_n(t)$, the nominal controls $u_n(t)$, and the actual measurement of the state $\vec{x}(t)$, come up with a way to do feedback control that will keep the trajectory staying close to the desired trajectory no matter what the small bounded disturbance $\vec{w}(t)$ does.

**Solution:**

From the previous parts, we have succesfully found a feedback control law $u_v(t) = [f_0 \quad f_1] \vec{v}$ such that the closed-loop system for $\vec{v}$ can be kept around $\vec{0}$ as long as the disturbances are bounded. By changing variables $\vec{x}(t) = \vec{x}_n(t) + \vec{v}(t)$ and $u(t) = u_n(t) + u_v(t)$ that we performed in (c) and (d), the state $\vec{x}$, as a result, will stay close to the desired trajectory no matter what the small bounded disturbance $\vec{w}(t)$ does.

Explicitly $u(t) = u_n(t) + u_v(t) = u_n(t) + [-144 \quad 120] (\vec{x}(t) - \vec{x}_n(t))$ is the control law that we would invoke to achieve this.

4. Continuous-Time Analog Observer Design: Ship Autopilots
Modern ships use autopilots for steering. The main task of the autopilot is to maintain constant heading. A common system model used for ship steering controllers is the Nomoto first-order model. It is described using the following differential equation:

\[ T \ddot{\psi} + \dot{\psi} = K \delta, \]

where \( \psi \) is the ship heading, \( \delta \) is the rudder angle, and \( K \) and \( T \) are constants that are empirically estimated during sea trials. The “dot” notation used here is the physics convention (Newton’s notation) that is very convenient for problems where nothing more that a second derivative is needed.

The only sensor is a gyrocompass, which reports the ship’s current heading \( y(t) = \psi(t) \). We would also like to provide a good estimate of an additional important parameter, the rate of turn — the derivative of the ship’s current heading.

The input of the ship model is the rudder angle \( \delta \), and the output is the heading \( \psi \), as measured by the gyrocompass.

(Note for the curious: undoubtedly, some of you are wondering why we don’t just take the derivative of the measurement and be done with it. The reason is that although we are describing everything without any noise, in the real-world, all measurements are noisy. Taking the derivative of noise is a very bad idea because it is in the nature of noise to shake a lot and so the derivative gets swamped by the shaking of the noise.)

In this problem you’ll construct an analog continuous-time observer, and then analyze its behaviour.

(a) Choose your state variables so that you have a two-dimensional state.

**Solution:** This is a second-order differential equation in terms of \( \psi \). We can introduce a new variable \( r = \dot{\psi} \), the rate of turn. Together with \( \psi \), our state would be \[
\begin{bmatrix}
\psi \\
r
\end{bmatrix}
\].

(b) Write down the system as a state-space model with a two-dimensional state.

**Solution:** Replacing \( \ddot{\psi} \) with \( r \) and rearranging the terms, the equation becomes:

\[ T \dot{r} = K \delta - r. \]

And in matrix form:

\[
\frac{d}{dt} \begin{bmatrix}
\psi(t) \\
r(t)
\end{bmatrix} = \begin{bmatrix}
A & B \\
0 & -1/T
\end{bmatrix} \begin{bmatrix}
\psi(t) \\
r(t)
\end{bmatrix} + \begin{bmatrix}
0 \\
0
\end{bmatrix} \delta(t)
\]

\[ y(t) = \begin{bmatrix}
1 & 0
\end{bmatrix} \begin{bmatrix}
\psi(t) \\
r(t)
\end{bmatrix} \]

(c) Is the system observable?

**Solution:** Yes.

\[ CA = \begin{bmatrix}
1 & 0
\end{bmatrix} \begin{bmatrix}
0 & 1/T \\
0 & -1/T
\end{bmatrix} = \begin{bmatrix}
0 & 1
\end{bmatrix} \]

The observability matrix has two linearly independent vectors.

(d) Write down a model for the observer in matrix form using \( \tilde{\ell} \) to represent how you weigh the difference between the observed output \( y(t) \) and the estimated output \( \dot{\hat{y}}(t) \) coming from within your observer.
Solution:

\[
\begin{align*}
\frac{d}{dt} \begin{bmatrix} \hat{\psi}(t) \\ \hat{r}(t) \end{bmatrix} &= \begin{bmatrix} 0 & 1 \\ 0 & -\frac{1}{T} \end{bmatrix} \begin{bmatrix} \psi(t) \\ r(t) \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{K}{T} \end{bmatrix} \delta(t) - \hat{r}(\hat{y}(t) - y(t)) \\
\hat{y}(t) &= \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} \hat{\psi}(t) \\ \hat{r}(t) \end{bmatrix}
\end{align*}
\]

Note that the first row of the equation is

\[
\dot{\hat{\psi}} = \hat{r} - l_0 (\hat{\hat{y}} - y)
\]

We also know that \( \dot{\psi} = r. \) Does that mean we can cancel out \( \dot{\hat{\psi}} \) and \( \hat{r} \)?

The answer is no. \( \hat{r} \) does not necessarily equal \( \dot{\hat{\psi}}. \) These are just estimates of the original quantities, we intentionally add a new term \(-l_0 (\hat{\hat{y}} - y)\) to make sure the two are not always equal, but rather converge into the original \( \psi \) and \( r \) in time.

(e) Find \( \vec{l} = \begin{bmatrix} l_0 \\ l_1 \end{bmatrix} \) to place both the eigenvalues of the estimation error evolution at \(-2\).

**Solution:**

\[
A - \vec{l}C = \begin{bmatrix} 0 & 1 \\ 0 & -\frac{1}{T} \end{bmatrix} - \begin{bmatrix} l_0 \\ l_1 \end{bmatrix} \begin{bmatrix} 1 & 0 \end{bmatrix} = \begin{bmatrix} -l_0 & 1 \\ -l_1 & -\frac{1}{T} \end{bmatrix}
\]

The characteristic polynomial is:

\[
\lambda^2 + \left( l_0 + \frac{1}{T} \right) \lambda + \left( \frac{l_0}{T} + l_1 \right)
\]

To place the eigenvalues at \(-2\), the characteristic polynomial should be \((\lambda + 2)^2 = \lambda^2 + 4\lambda + 4. \) Therefore

\[
\begin{bmatrix} l_0 \\ l_1 \end{bmatrix} = \begin{bmatrix} 4 - \frac{1}{T} \\ (2 - \frac{1}{T})^2 \end{bmatrix}
\]

Now that we have designed the output-feedback and placed the eigenvalues of the estimation error. We’ll design a circuit implementing the observer.

We will represent the state variables as voltages. Each input, output, and state variable will be implemented as a node in our circuit. The output of the original systems (the gyrocompass) would be an input of this system, and so would the rudder angle.

Recall that in EE16A and previously in EE16B, you have seen how to implement the following operations using simpler circuit elements (mainly resistors, capacitors and op-amps): differentiation, integration, scaling, addition and negation. This will be enough to implement the observer.

(f) Design a circuit whose output is the integral of its input with respect to time.

**Solution:** We build an inverting integrator with an inverting amplifier connected in series:
This circuit’s output is
\[ v_{\text{out}} = \frac{R_3}{R_1 R_2 C} \int_0^t v_{\text{in}}(u) \, du. \]

We need to choose resistor and capacitor values so that \( R_1 R_2 C = R_3 \). For example \( R_1 = R_2 = 1 M\Omega, \ C = 1 nF, \ R_3 = 1 k\Omega \).

(g) Design a circuit whose output is a scaled version by a constant \( a_0 \) of its input.

**Solution:** We can just use an non-inverting amplifier

This circuit’s output is
\[ v_{\text{out}} = v_{\text{in}} \frac{R_1 + R_2}{R_1} . \]

Choose resistor values such that \( \frac{R_1 + R_2}{R_1} = a_0 \). If the constant is negative, we can always connect an inverting amplifier with a gain of -1.

(h) Design a circuit whose output is the negation of its input.

**Solution:** We can just use an inverting amplifier
This circuit’s output is
\[ v_{\text{out}} = -v_{\text{in}} \frac{R_2}{R_1}, \]

Choose resistor values to set the right gain so \( R_2 = R_1 \) which can be 100\( \text{k}\Omega \).

(i) Design a circuit whose output is the sum of its two inputs.

**Solution:** We can just use an inverting summing amplifier. In this example we have three inputs:

\[ v_{\text{out}} = R_4 \frac{R_2}{R_1} R_3 \left( v_{\text{in}1} + v_{\text{in}2} + v_{\text{in}3} \right) \]

We need to choose resistor values such that \( R_2 R_4 = R_1 R_3 \) which can be done by say, setting them all to 100\( \text{k}\Omega \).

Now that we have the basic circuit elements. We’ll implement the observer as a circuit.

(j) Use the circuits you designed above to construct the observer as a circuit driven by the output of the gyrocompass.
**Solution:** We can rearrange the system into the following form:

\[
\frac{d}{dt} \begin{bmatrix} \dot{\psi} \\ \dot{r} \end{bmatrix} = \begin{bmatrix} f(\psi, \dot{r}, \delta, y) \\ g(\psi, \dot{r}, \delta, y) \end{bmatrix}
\]

\[
f(\psi, \dot{r}, \delta, y) = \dot{r} - \left(4 - \frac{1}{T}\right)(\psi - y)
\]

\[
g(\psi, \dot{r}, \delta, y) = -\frac{1}{T} \dot{r} + K T \delta - \left(2 - \frac{1}{T}\right)^2 (\psi - y)
\]

To implement \(f\) and \(g\) we’ll build sub-circuits whose inputs are \(\dot{\psi}, \dot{r}, \delta,\) and \(y\). We have all the basic components required to implement it:

We can build \(g\) similarly:

We now need to connect an integrator component to the output of \(f\) and connect the integrator’s output to the \(\dot{\psi}\) input of \(f\) and \(g\), (and similarly with the output of \(g\)). Now connect \(y\) and \(\delta\) to the corresponding inputs of \(f\) and \(g\).

**5. Lagrange interpolation by polynomials**

Given \(n\) distinct points and the corresponding evaluations/sampling of a function \(f(x), (x_i, f(x_i))\) for \(0 \leq i \leq n - 1\), the Lagrange interpolating polynomial is the polynomial of the least degree which passes through all the given points.

Given \(n\) distinct points and the corresponding evaluations, \((x_i, f(x_i))\) for \(0 \leq i \leq n - 1\), the Lagrange polynomial is

\[
P_n(x) = \sum_{i=0}^{i=n-1} f(x_i)L_i(x),
\]

where
\[ L_i(x) = \prod_{j=0; j \neq i}^{j=n-1} \frac{(x-x_j)}{(x_i-x_j)} = \frac{(x-x_0)}{(x_i-x_0)} \cdot \frac{(x-x_1)}{(x_i-x_1)} \cdot \ldots \cdot \frac{(x-x_{i-1})}{(x_i-x_{i-1})} \cdot \frac{(x-x_{i+1})}{(x_i-x_{i+1})} \cdot \ldots \cdot \frac{(x-x_{n-1})}{(x_i-x_{n-1})} \]

Here is an example: for two data points, \((x_0, f(x_0)) = (0, 4), (x_1, f(x_1)) = (-1, -3)\), we have

\[ L_0(x) = \frac{x-x_1}{x_0-x_1} = \frac{x-(-1)}{0-(-1)} = x + 1 \]

and

\[ L_1(x) = \frac{x-x_0}{x_1-x_0} = \frac{x-(0)}{(-1)-(0)} = -x \]

. Then

\[ P_2(x) = f(x_0)L_0(x) + f(x_1)L_1(x) = 4(x+1) + (-3)(-x) = 7x+4 \]

We can sketch those equations on the 2D plane as follows:

(a) Given three data points, \((2, 3), (0, -1)\) and \((-1, -6)\), find a polynomial \(f(x) = ax^2 + bx + c\) fitting the three points. Do this by solving a system of linear equations for the unknowns \(a, b, c\). Is this polynomial unique?

**Solution:** Plug in the three data points into \(f(x) = ax^2 + bx + c\), we get \(4a + 2b + c = 3, c = -1\) and \(a-b+c = -6\).

Solve these equations for the coefficients \(a, b\) and \(c\), and we get \(f(x) = -x^2 + 4x - 1\).

This polynomial is the unique degree 2 polynomial defined by these three distinct points.

(b) Like the monomial basis \(\{x, x^2, x^3, \ldots\}\), the set \(\{L_i(x)\}\) is a new basis for the subspace of degree \(n\) or lower polynomials. \(P_n(x)\) is the sum of the scaled basis polynomials. Find the \(L_i(x)\) corresponding to the three sample points in (a). Show your steps.

**Solution:** Assume \(x_0 = 2, x_1 = 0\) and \(x_2 = -1\), each corresponding \(L_i(x)\) is:

\[ L_0(x) = \frac{(x-x_1)}{(x_0-x_1)} \cdot \frac{(x-x_2)}{(x_0-x_2)} = \frac{(x-0)}{(2-0)} \cdot \frac{(x-(1))}{(2-(1))} = \frac{x^2+x}{6} \]

\[ L_1(x) = \frac{(x-x_0)}{(x_1-x_0)} \cdot \frac{(x-x_2)}{(x_1-x_2)} = \frac{(x-2)}{(0-2)} \cdot \frac{(x-(1))}{(0-(1))} = \frac{x^2-x-2}{-2} \]
\[ L_2(x) = \frac{(x-x_0)}{(x_2-x_0)} \frac{(x-x_1)}{(x_2-x_1)} = \frac{(x-2)}{(-1-2)} \frac{(x-0)}{(-1-0)} = \frac{x^2 - 2x}{3} \]

(c) Find the Lagrange polynomial \( P_n(x) \) for the three points in (a). Compare the result to the answer in (a). Are they different from each other? Why or why not?

Solution: \( P_n(x) = f(x_0)L_0(x) + f(x_1)L_1(x) + f(x_2)L_2(x) = 3 \times \frac{x^2 + x}{6} + (-1) \times \frac{x^2 - x - 2}{2} + (-6) \times \frac{x^2 - 2x}{3} \)

Therefore, \( P_n(x) = -x^2 + 4x - 1 \)

The answer is the same as the one we computed in (a). For these three points, we cannot have a polynomial of less than 2 degree fitting all of them. Lagrange interpolation must construct the least degree \((n = 2)\) polynomial passing the three points, which must be unique.

(d) Sketch \( P_n(x) \) and each \( f(x_i)L_i(x) \) on the 2D plane.

Solution:

(e) Show that the Lagrange interpolating polynomial must pass through all given points. In other words, show that \( P_n(x_i) = f(x_i) \) for all \( x_i \). Do this in general, not just for the example above.

Solution: For each \( x_i \), the corresponding \( L_i(x) \) is \( L_i(x) = \prod_{j=0, j \neq i}^{j=n-1} \frac{(x-x_j)}{(x_i-x_j)} \). Plug in \( x_i \), we will get

\[ L_i(x_i) = \prod_{j=0, j \neq i}^{j=n-1} \frac{(x_i-x_j)}{(x_i-x_i)} = 1. \]

For any other \( L_m(x) \), where \( m \neq i \), plug in \( x_i \), \( L_m(x_i) = \prod_{j=0, j \neq m}^{j=n-1} \frac{(x_i-x_j)}{(x_m-x_j)} \). Because \( m \neq i \), \( L_m(x_i) \) must have this term \( \frac{(x_i-x_i)}{(x_m-x_i)} = 0 \). Hence all \( L_m(x_i) \) must be zero.

Therefore, \( P_n(x_i) = \sum_{k=0}^{k=n-1} f(x_k)L_k(x_i) = f(x_i) \times 1 = f(x_i) \)
6. The vector space of polynomials

A polynomial of degree at most \( n \) on a single variable can be written as

\[
p(x) = p_0 + p_1 x + p_2 x^2 + \cdots + p_n x^n
\]

where we assume that the coefficients \( p_0, p_1, \ldots, p_n \) are real. Let \( P_n \) be the vector space of all polynomials of degree at most \( n \).

(a) Consider the representation of \( p \in P_n \) as the vector of its coefficients in \( \mathbb{R}^{n+1} \).

\[
\vec{p} = [p_0 \ p_1 \ \cdots \ p_n]^T
\]

Show that the set \( \mathcal{B}_n = \{1, x, x^2, \ldots, x^n\} \) forms a basis of \( P_n \), by showing the following.

- Every element of \( P_n \) can be expressed as a linear combination of elements in \( \mathcal{B}_n \).
- No element in \( \mathcal{B}_n \) can be expressed as a linear combination of the other elements of \( \mathcal{B}_n \).

(Hint: Use the aspect of the fundamental theorem of algebra which says that a nonzero polynomial of degree \( n \) has at most \( n \) roots, and use a proof by contradiction.)

Solution: We can write every polynomial of degree \( n \) on \( x \) as \( p(x) = p_0 + p_1 x + \cdots + p_n x^n \), which is a linear combination of the elements in \( \mathcal{B}_n \).

Let \( x^i \in \mathcal{B}_n \), for \( 0 \leq i \leq n \). We show that \( x^i \) cannot be written as a linear combination of the other elements of \( \mathcal{B}_n \). Suppose that there exist real constants \( c_j \) for \( j \neq i \), such that \( x^i = \sum_{j \neq i} c_j x^j \). This gives us an equation

\[
\sum_{0 \leq j \leq n, j \neq i} c_j x^j - x^i = 0.
\]

The left hand side is a nonzero polynomial of degree at most \( n \), which we know has at most \( n \) roots by the relevant aspect of the fundamental theorem of algebra. Thus the equality cannot hold for all \( x \), which is a contradiction.

(b) Suppose that the coefficients \( p_0, \ldots, p_n \) of \( p \) are unknown. To determine the coefficients, we evaluate \( p \) on \( n + 1 \) points, \( x_0, \ldots, x_n \). Suppose that \( p(x_i) = y_i \) for \( 0 \leq i \leq n \). Find a matrix \( V \) in terms of the \( x_i \), such that

\[
V \begin{pmatrix} p_0 \\ p_1 \\ \vdots \\ p_n \end{pmatrix} = \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_n \end{pmatrix}.
\]

Solution: We can write each equation \( p(x_i) = y_i \) as

\[
p_0 + p_1 x_i + p_2 x_i^2 + \cdots + p_n x_i^n = \begin{pmatrix} 1 & x_i & x_i^2 & \cdots & x_i^n \end{pmatrix} \begin{pmatrix} p_0 \\ p_1 \\ \vdots \\ p_n \end{pmatrix} = y_i.
\]

If we arrange these \( n + 1 \) equations into a matrix, we get

\[
\begin{pmatrix} 1 & x_0 & x_0^2 & \cdots & x_0^n \\ 1 & x_1 & x_1^2 & \cdots & x_1^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^n \end{pmatrix} \begin{pmatrix} p_0 \\ p_1 \\ \vdots \\ p_n \end{pmatrix} = \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_n \end{pmatrix}.
\]

The matrix \( V \) above with rows \( \begin{pmatrix} 1 & x_i & x_i^2 & \cdots & x_i^n \end{pmatrix} \) is also known as a Vandermonde matrix.
(c) For the case where \( n = 2 \), compute the determinant of \( V \) and show that it is equal to

\[
\det(V) = \prod_{0 \leq i < j \leq n} (x_j - x_i).
\]

Conclude that if \( x_0, \ldots, x_n \) are distinct, then we can uniquely recover the coefficients \( p_0, \ldots, p_n \) of \( p \).

This holds for \( n > 2 \) in general, but consider only the case where \( n = 2 \) for now.

**Solution:** If \( n = 2 \), we can write \( V \) as

\[
V = \begin{pmatrix}
1 & x_0 & x_0^2 \\
1 & x_1 & x_1^2 \\
1 & x_2 & x_2^2
\end{pmatrix}
\]

We can compute the determinant of \( V \) by using row reduction. This gives us the matrix

\[
\begin{pmatrix}
1 & x_0 & x_0^2 \\
0 & x_1 - x_0 & x_1^2 - x_0^2 \\
0 & x_1^2 - x_1x_2 - x_0x_2 + x_0x_1 & x_2^2 - x_1x_2 - x_0x_2 + x_0x_1
\end{pmatrix}
\]

Since this is upper-triangular, we can compute the determinant by taking the product of the diagonal entries.

\[
\det(V) = x_1x_2 - x_1^2x_2 - x_0x_2^2 + x_0^2x_2 + x_0x_1^2 - x_0^2x_1.
\]

We can check that this is equal to

\[
(x_2 - x_1)(x_2 - x_0)(x_1 - x_0) = x_1x_2^2 - x_1^2x_2 - x_0x_2^2 + x_0^2x_2 + x_0x_1^2 - x_0^2x_1
\]

Note that \( \det(V) \) is nonzero, or \( V \) is invertible, if and only if all \( x_0, x_1, x_2 \) are distinct. If this is the case, then there is a unique solution for \( p_0, p_1, p_2 \).

(d) (optional) Argue using Lagrange interpolation that indeed such matrices \( V \) above must always be invertible if the \( x_i \) are distinct.

**Solution:** If the \( x_i \) are distinct, we can construct \( n+1 \) Lagrange polynomials \( L_i(x) \) such that \( L_i(x_i) = 1 \) and \( L_i(x_j) = 0 \) if \( j \neq i \). Let \( \vec{I}_i \) be the vector of the coefficients of \( L_i(x) \) in the basis \( \mathcal{B}_n \). From part (b), we know that \( V\vec{I}_i \) is the vector with \( i \)-th entry 1 and 0 elsewhere. Let \( L \) be the matrix with columns \( \vec{I}_0, \ldots, \vec{I}_n \). \( L \) is \( (n+1) \times (n+1) \) since each \( \vec{I}_i \) has \( n+1 \) entries, and \( V \) has the same dimensions. Thus \( VL \) is the identity matrix. This implies that \( L = V^{-1} \), so \( V \) is invertible.

(e) We can define an inner product on \( P_n \) by setting

\[
\langle p, q \rangle = \int_{-1}^{1} p(x)q(x) \, dx.
\]

Show that this satisfies the following properties of a real inner product. (We would have to put in a complex conjugate on \( p \) if we wanted a complex inner product.)

- \( \langle p, p \rangle \geq 0 \), with equality if and only if \( p = 0 \).
- For all \( a \in \mathbb{R} \), \( \langle ap, q \rangle = a \langle p, q \rangle \).
- \( \langle p, q \rangle = \langle q, p \rangle \).

**Solution:**

- \( \langle p, p \rangle = \int_{-1}^{1} p^2(x) \, dx \geq 0 \) since \( p^2(x) \geq 0 \) for all \( x \). The integral is 0 if and only if \( p = 0 \) for all \( x \), or \( p \) is the zero polynomial.
- \( \langle ap, q \rangle = \int_{-1}^{1} (ap(x))(q(x)) \, dx = a \int_{-1}^{1} p(x)q(x) \, dx = a \langle p, q \rangle \).
(f) Now that we have an inner product on \( P_n \), we can consider orthonormality. If \( \mathcal{B} = \{ b_0, b_1, \ldots, b_n \} \) is a basis for \( P_n \), we say that it is an orthonormal basis if

- \( \langle b_i, b_j \rangle = 0 \) if \( i \neq j \).
- \( \langle b_i, b_i \rangle = 1 \).

We can also compute projections. For any \( p, u \in P_n, u \neq 0 \), the projection of \( p \) onto \( u \) is

\[
\text{proj}_u p = \frac{\langle p, u \rangle}{\langle u, u \rangle} u.
\]

Consider the case where \( n = 2 \). From part (a), we have the basis \( \{1, x, x^2\} \) for \( P_2 \). Convert this into an orthonormal basis using the Gram-Schmidt process.

**Solution:** Using the Gram-Schmidt process, the first basis element is \( b_0 = \frac{1}{\sqrt{\langle 1, 1 \rangle}} \). We can compute the denominator as

\[
\langle 1, 1 \rangle = \int_{-1}^{1} dx = 2
\]

so we have \( b_0 = \frac{1}{\sqrt{2}} \). To compute the next basis element, we first compute

\[
x - \langle x, b_0 \rangle b_0 = x - \frac{1}{\sqrt{2}} \int_{-1}^{1} \frac{1}{\sqrt{2}} x dx = x.
\]

We compute \( b_1 \) by normalizing this. We first compute

\[
\langle x, x \rangle = \int_{-1}^{1} x^2 dx = \frac{2}{3}
\]

So we have \( b_1 = \sqrt{\frac{3}{2}} x \). To compute \( b_2 \), we first compute

\[
x^2 - \langle x^2, b_0 \rangle b_0 - \langle x^2, b_1 \rangle b_1 = x^2 - \left( \frac{1}{\sqrt{2}} \int_{-1}^{1} \frac{1}{\sqrt{2}} x^2 dx \right) - \left( \sqrt{\frac{3}{2}} \int_{-1}^{1} \sqrt{\frac{3}{2}} x^3 dx \right) = x^2 - \frac{1}{3}
\]

We compute \( b_2 \) by normalizing this.

\[
b_2 = \frac{x^2 - \frac{1}{3}}{\sqrt{\langle x^2 - \frac{1}{3}, x^2 - \frac{1}{3} \rangle}} = \sqrt{\frac{45}{8}} \left( x^2 - \frac{1}{3} \right).
\]

(g) (optional) An alternative inner-product could be placed upon real polynomials if we simply represented them by a sequence of their evaluations at 0, 1, \ldots, \( n \) and adopted the standard Euclidean inner product on sequences of real numbers. Can you give an example of an orthonormal basis with this alternative inner product?

**Solution:** For \( p, q \in P_n \), this alternative inner product is given by

\[
\langle p, q \rangle = \sum_{i=0}^{n} p(i)q(i).
\]

An example of an orthonormal basis with this inner product would be the Lagrange polynomials \( L_i(x) \) constructed on the points 0, 1, \ldots, \( n \). Since \( L_i(i) = 1 \) and \( L_i(j) = 0 \) for \( j \neq i, 0 \leq i, j \leq n \), thus for \( i \neq j \) we clearly have \( \langle L_i, L_j \rangle = 0 \) and we can also check that \( \langle L_i, L_i \rangle = 1 \).
7. Your Own Problem

Write your own problem related to this week’s material and solve it. You may still work in groups to brainstorm problems, but each student must submit a unique problem. What is the problem? How to formulate it? How to solve it? What is the solution?

Contributors:

- Ming Jin.
- Baruch Sterin.