Exam 3, Math 601
Fall 2018
University Of Wisconsin, Milwaukee

Registered as undegraduate non-degree student

Nasser M. Abbasi

December 13, 2018 Compiled on December 13, 2018 at 12:37am

Contents

List of Figures

1 Problem 1

1. Let $x(t)$ and $y(t)$ denote the population sizes of two biological species. If the two species are not competing for resources (occupy different biological niches) then a simple logistic model could be feasible to describe the dynamics of their coexistence.

$$
\begin{aligned}
x' &= a_1 x - b_1 x^2 \\
y' &= a_2 y - b_2 y^2.
\end{aligned}
$$

If however the two species are direct competitors, then their access to resources and their population growth rate could be reduced by a quantity that is proportional to the size of the competing species' population, leading to a *competition system* model

$$
\begin{aligned}
x' &= a_1 x - b_1 x^2 - c_1 xy \\
y' &= a_2 y - b_2 y^2 - c_2 xy.
\end{aligned}
$$

Assume that in a competition system (with appropriate units) the coefficients are given as

$$a_1 = 60, \quad a_2 = 42, \quad b_1 = 3, \quad b_2 = 3, \quad c_1 = 4, \quad c_2 = 2,$$

and determine all equilibria of the system as well as their corresponding stability properties. Give a short interpretation of your results in terms of the long term species dynamics (as $t \to \infty$).

Figure 1: Problem 1 Statement

<u>Solution</u>

$$
\begin{aligned}
x' &= a_1 x - b_1 x^2 - c_1 xy \\
y' &= a_2 y - b_2 y^2 - c_2 xy
\end{aligned}
$$

Using the values given in the problem, the above equations become

$$x' = 60x - 3x^2 - 4xy \tag{1A}$$

$$y' = 42y - 3y^2 - 2xy \tag{1B}$$

Or

$$
\begin{aligned}
x' &= f(x, y) \\
y' &= g(x, y)
\end{aligned}
$$

Equilibrium points are found by setting $f(x,y)=0$ and $g(x,y)$. This results in the following two equations to solve for x,y

$$60x-3x^2-4xy=0 \tag{1}$$
$$42y-3y^2-2xy=0 \tag{2}$$

The first equation (1A) becomes $x(60-3x-4y)=0$ which then gives one solution as

$$x=0 \tag{3}$$

And $60-3x-4y=0$ gives another solution as

$$x=\frac{60-4y}{3} \tag{4}$$

The second equation (1B) becomes $y(42-3y-2x)=0$ which gives one solution as

$$y=0 \tag{5}$$

And $42-3y-2x=0$ gives another solution as

$$y=\frac{42-2x}{3} \tag{6}$$

When $x=0$ then (6) results in $y=\frac{42}{3}=14$. When $x=\frac{60-4y}{3}$ then (6) results in $y=\frac{42-2\left(\frac{60-4y}{3}\right)}{3}=\frac{8}{9}y+\frac{2}{3}$, or $y=6$. Hence in this case $x=\frac{60-4(6)}{3}=12$.

Similarly, when $y=0$ then from (4) $x=\frac{60-4(0)}{3}=20$. The above shows that there are 4 equilibrium points. These are

$$\begin{aligned}x&=0,y=0\\x&=0,y=14\\x&=12,y=6\\x&=20,y=0\end{aligned}$$

To determine the type of stability of each equilibrium point, and since this is a nonlinear system, we must first linearize the system around each equilibrium point in order to determine the Jacobian matrix.

Once the system is linearized, then the eigenvalues of the Jacobian matrix are found in each case. From the values of eigenvalues we can then determine if the system is stable or not at each one of the above four equilibrium points.

The first step is then to linearize $f(x,y)$ and $g(x,y)$ around each of the equilibrium points. If we assume the equilibrium point is given by x_0,y_0 then expanding $f(x,y)$ in Taylor series around this point gives

$$f(x_0+\Delta x,y_0+\Delta y)=f(x_0,y_0)+\left.\frac{\partial f(x,y)}{\partial x}\right|_{x_0,y_0}(\Delta x+\Delta y)+\left.\frac{\partial f(x,y)}{\partial y}\right|_{x_0,y_0}(\Delta x+\Delta y)+\cdots$$

But $f(x_0, y_0) = 0$ since it is what defines an equilibrium point, the above becomes, after ignoring higher order terms since we are assuming small $\Delta x, \Delta y$

$$f(x_0 + \Delta x, y_0 + \Delta y) = \left.\frac{\partial f(x,y)}{\partial x}\right|_{x_0,y_0} (\Delta x + \Delta y) + \left.\frac{\partial f(x,y)}{\partial y}\right|_{x_0,y_0} (\Delta x + \Delta y)$$

Similarly for $g(x,y)$ we obtain the following

$$g(x_0 + \Delta x, y_0 + \Delta y) = \left.\frac{\partial g(x,y)}{\partial x}\right|_{x_0,y_0} (\Delta x + \Delta y) + \left.\frac{\partial g(x,y)}{\partial y}\right|_{x_0,y_0} (\Delta x + \Delta y)$$

Therefore a linearized f, g functions at the equilibrium point become

$$\begin{pmatrix} f(x_0 + \Delta x, y_0 + \Delta y) \\ g(x_0 + \Delta x, y_0 + \Delta y) \end{pmatrix} = \begin{pmatrix} \left.\frac{\partial f}{\partial x}\right|_{x_0,y_0} & \left.\frac{\partial f}{\partial y}\right|_{x_0,y_0} \\ \left.\frac{\partial g}{\partial x}\right|_{x_0,y_0} & \left.\frac{\partial g}{\partial y}\right|_{x_0,y_0} \end{pmatrix} \begin{pmatrix} \Delta x + \Delta y \\ \Delta x + \Delta y \end{pmatrix}$$

Replacing the original nonlinear $f(x,y), g(x,y)$ by the above linearized (approximation), the system can now be written as

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \\ \frac{\partial g}{\partial x} & \frac{\partial g}{\partial x} \end{pmatrix}_{\substack{x=x_0 \\ y=y_0}} \begin{pmatrix} x \\ y \end{pmatrix}$$

Where $\begin{pmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \\ \frac{\partial g}{\partial x} & \frac{\partial g}{\partial x} \end{pmatrix}$ is called the the Jacobian J matrix. Hence the system now can be written as

$$\vec{x}' = [J]\,\vec{x}$$

Now J is determined. From

$$\begin{aligned}
\frac{\partial f}{\partial x} &= \frac{\partial}{\partial x}\left(60x - 3x^2 - 4xy\right) = 60 - 6x - 4y \\
\frac{\partial f}{\partial y} &= \frac{\partial}{\partial y}\left(60x - 3x^2 - 4xy\right) = -4x \\
\frac{\partial g}{\partial x} &= \frac{\partial}{\partial x}\left(42y - 3y^2 - 2xy\right) = -2y \\
\frac{\partial g}{\partial y} &= \frac{\partial}{\partial y}\left(42y - 3y^2 - 2xy\right) = 42 - 6y - 2x
\end{aligned}$$

The Jacobian matrix becomes

$$J = \begin{pmatrix} 60 - 6x - 4y & -4x \\ -2y & 42 - 6y - 2x \end{pmatrix}_{x=x_0, y=y_0}$$

And the linearized system is

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} 60 - 6x - 4y & -4x \\ -2y & 42 - 6y - 2x \end{pmatrix}_{x=x_0, y=y_0} \begin{pmatrix} x \\ y \end{pmatrix}$$

Now each equilibrium point is examined using the above linearized system to determine the type of stability a that point.

case $x_0 = 0, y_0 = 0$

$$J = \begin{pmatrix} 60 - 6(0) - 4(0) & -4(0) \\ -2(0) & 42 - 6(0) - 2(0) \end{pmatrix} = \begin{pmatrix} 60 & 0 \\ 0 & 42 \end{pmatrix}$$

Hence the linearized system at this specific equilibrium point is

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} 60 & 0 \\ 0 & 42 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

Since J is a diagonal matrix, its eigenvalues are the values on the diagonal. Therefore $\lambda_1 = 60, \lambda_2 = 42$. Since the eigenvalues are positive, then this equilibrium point is not stable.

case $x_0 = 0, y_0 = 14$

$$J = \begin{pmatrix} 60 - 6(0) - 4(14) & -4(0) \\ -2(14) & 42 - 6(14) - 2(0) \end{pmatrix} = \begin{pmatrix} 4 & 0 \\ -28 & -42 \end{pmatrix}$$

Therefore linearized system at this specific equilibrium point is

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} 4 & 0 \\ -28 & -42 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

The eigenvalues can be found by solving $\begin{vmatrix} 4 - \lambda & 0 \\ -28 & -42 - \lambda \end{vmatrix} = 0$ to be $\lambda_1 = 4, \lambda_2 = -42$. Because one of the eigenvalues is positive, then this equilibrium point is not stable.

case $x_0 = 12, y_0 = 6$

$$J = \begin{pmatrix} 60 - 6(12) - 4(6) & -4(12) \\ -2(6) & 42 - 6(6) - 2(12) \end{pmatrix} = \begin{pmatrix} -36 & -48 \\ -12 & -18 \end{pmatrix}$$

The linearized system at this specific equilibrium point is

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} -36 & -48 \\ -12 & -18 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

The eigenvalues can be found to be $\lambda_1 = -52.632, \lambda_2 = -1.368$. Since both eigenvalues are now negative, then this equilibrium point is stable.

case $x_0 = 20, y_0 = 0$

$$J = \begin{pmatrix} 60 - 6(20) - 4(0) & -4(20) \\ -2(0) & 42 - 6(0) - 2(20) \end{pmatrix} = \begin{pmatrix} -60 & -80 \\ 0 & 2 \end{pmatrix}$$

The linearized system at this specific equilibrium point is

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} -60 & -80 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

The eigenvalues are $\lambda_1 = 2, \lambda_2 = -60$. Since one of the eigenvalues is positive, then this equilibrium point is not stable.

Summary of results obtained so far

equilibrium point	eigenvalues	type of stability
$x = 0, y = 0$	$\lambda_1 = 60, \lambda_2 = 42$	not stable (nodal source)
$x = 0, y = 14$	$\lambda_1 = 4, \lambda_2 = -42$	not stable (Saddle point)
$x = 12, y = 6$	$\lambda_1 = -52.632, \lambda_2 = -1.368$	stable (Nodal sink)
$x = 20, y = 0$	$\lambda_1 = 2, \lambda_2 = -60$	not stable (Saddle point)

To verify the above result, the phase plot for the original nonlinear system was plotted on the computer and the equilibrium points locations highlighted. The plot below agrees with the above result when looking at direction of arrows around each point. We see that the direction field arrows are all moving toward the stable point from any location near it. The stable equilibrium point was colored as green while the unstable ones colored in red.

```
ClearAll[x, y];
f1 = 60 x - 3 x^2 - 4 x y;
f2 = 42 y - 3 y^2 - 2 x y;
StreamPlot[{f1, f2}, {x, -5, 25}, {y, -5, 20},
 Epilog → {{Red, PointSize[0.03], Point[{{0, 0}, {0, 14}, {20, 0}}]},
   {Green, PointSize[.035], Point[{{12, 6}}]}},
 PlotLabel → Column[{"Phase plot for problem 1", "Showing location of all equilibrium points"},
   Alignment → Center],
 BaseStyle → 12,
 ImageSize → 400]
```

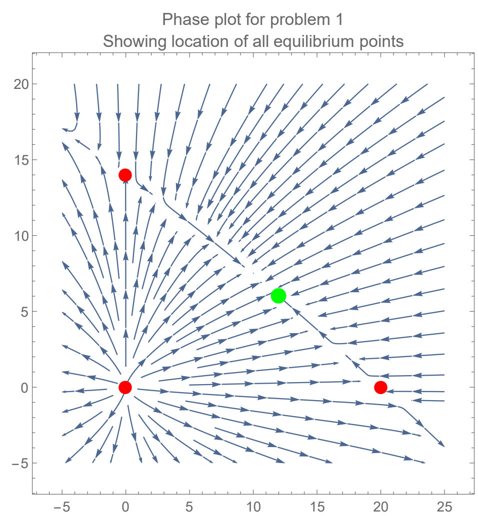


Figure 2: Phase plot for problem 1

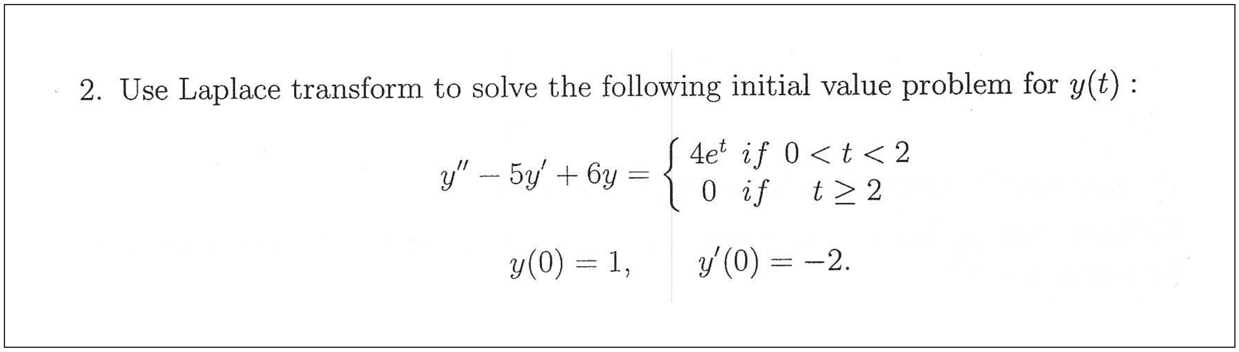
Interpretation of results Since the solution of the linearized system can be written as linear combination of solutions made up of terms that look like $c_i e^{\lambda_i t}$ where c_i are constants of integration and λ_i are the eigenvalues found above, then this implies when the real part of the eigenvalue is positive the solution will increase with time, moving away from the equilibrium point. Similarly, if the eigenvalue has a negative real part, it means it is a stable solution because solution will decay with time when perturbed slightly from the equilibrium.

Since this is second order system, there are two eigenvalues. Even if one eigenvalue is stable (i.e. negative), if the other eigenvalue is positive, then the system is unstable since one part of the solution will keep growing with time.

In terms of the dynamics of species, it means if the populations $x = 12$ and population $y = 6$, (this is the stable equilibrium) then these population will remain the same in long term even when one population becomes a little more or less than the other population. But for all other equilibrium populations sizes, such as $x = 20, y = 0$, then if the population y were to change slightly to become say $y = 1$ (may be by external influence) then this will cause both population to start changing, moving it away from $x = 20, y = 0$ as time increases, hence $x = 20, y = 0$ is not stable population size.

This seems to be sensitive to the parameters a_i, b_i, c_i given in the problem. It is not easy to give a more physical reasoning as why some population values is stable while other are not, other than to also note that all the unstable ones had at least one population at zero.

2 Problem 2



2. Use Laplace transform to solve the following initial value problem for $y(t)$:

$$y'' - 5y' + 6y = \begin{cases} 4e^t \; if \; 0 < t < 2 \\ 0 \quad if \quad t \geq 2 \end{cases}$$

$$y(0) = 1, \qquad y'(0) = -2.$$

Figure 3: Problem 2 Statement

<u>Solution</u>

The ODE can be written as

$$\begin{aligned} y'' - 5y' + 6y &= 4e^t \left(U(t) - U(t-2)\right) \\ &= 4\left(e^t U(t) - e^t U(t-2)\right) \end{aligned}$$

Where $U(t)$ is the unit step function. In the following solutions, these Laplace transform relations obtained from table are used

$$\begin{aligned} U(t) &\Leftrightarrow \frac{1}{s} \\ U(t-\tau) &\Leftrightarrow \frac{1}{s}e^{-\tau s} \\ e^{-\alpha t}U(t) &\Leftrightarrow \frac{1}{s+\alpha} \\ \sin(\omega t) &\Leftrightarrow \frac{\omega}{s^2+\omega^2} \\ \cos(\omega t) &\Leftrightarrow \frac{s}{s^2+\omega^2} \end{aligned}$$

Assuming $\mathcal{L}[y(t)] = Y(s)$, and using the above relations of Laplace transform we find

$$\begin{aligned} \mathcal{L}\left[e^t U(t)\right] &= \frac{1}{s-1} \\ \mathcal{L}\left[e^t U(t-2)\right] &= \frac{e^{-2(s-1)}}{s-1} \end{aligned}$$

Now, taking the Laplace transform of the ODE results in

$$\left(s^2 Y(s) - sy(0) - y'(0)\right) - 5\left(sY(s) - y(0)\right) + 6Y(s) = 4\left(\frac{1}{s-1} - \frac{e^{-2(s-1)}}{s-1}\right)$$

Using $y(0)=1, y'(0)=-2$ the above simplifies to

$$
\begin{aligned}
\left(s^2Y(s)-s+2\right)-5\left(sY(s)-1\right)+6Y(s) &= \frac{4}{s-1}-\frac{4e^{-2(s-1)}}{s-1}\\
s^2Y(s)-s+2-5sY(s)+5+6Y(s) &= \frac{4}{s-1}-\frac{4e^{-2(s-1)}}{s-1}\\
Y(s)\left(s^2-5s+6\right)-s+7 &= \frac{4}{s-1}-\frac{4e^{-2(s-1)}}{s-1}\\
Y(s)\left(s^2-5s+6\right) &= \frac{4}{s-1}-\frac{4e^{-2(s-1)}}{s-1}+(s-7)
\end{aligned}
$$

But $\left(s^2-5s+6\right)=(s-3)(s-2)$ and the above becomes

$$
Y(s)=\frac{4}{(s-1)(s-3)(s-2)}-\frac{4e^2e^{-2s}}{(s-1)(s-3)(s-2)}+\frac{(s-7)}{(s-3)(s-2)} \tag{1}
$$

These are now simplified by partial fractions. The final result is only shown for brevity, since the process of performing partial fraction is a standard one.

$$
\frac{1}{(s-1)(s-3)(s-2)}=\frac{1}{2(s-3)}-\frac{1}{s-2}+\frac{1}{2(s-1)}
$$

And

$$
\frac{s-7}{(s-3)(s-2)}=\frac{-4}{s-3}+\frac{5}{s-2}
$$

Using the above result back in (1) results in

$$
\begin{aligned}
Y(s) &= \frac{2}{s-3}-\frac{4}{s-2}+\frac{2}{s-1}-e^2e^{-2s}\left(\frac{2}{(s-3)}-\frac{4}{s-2}+\frac{2}{(s-1)}\right)-\frac{4}{s-3}+\frac{5}{s-2}\\
&= \frac{-2}{s-3}+\frac{1}{s-2}+\frac{2}{s-1}-e^2\left(\frac{2e^{-2s}}{(s-3)}-\frac{4e^{-2s}}{s-2}+\frac{2e^{-2s}}{(s-1)}\right) \qquad (2)
\end{aligned}
$$

Now we apply the inverse Laplace transform. lookup table is also used for this purpose to obtain

$$
\begin{aligned}
-2\mathcal{L}^{-1}\left(\frac{1}{s-3}\right) &= -2e^{3t}\\
\mathcal{L}^{-1}\left(\frac{1}{s-2}\right) &= e^{2t}\\
2\mathcal{L}^{-1}\left(\frac{1}{s-1}\right) &= 2e^{t}
\end{aligned}
$$

And

$$
\begin{aligned}
2\mathcal{L}^{-1}\left(\frac{e^{-2s}}{s-3}\right) &= 2e^{3(t-2)}U(t-2)\\
4\mathcal{L}^{-1}\left(\frac{e^{-2s}}{s-2}\right) &= 4e^{2(t-2)}U(t-2)\\
2\mathcal{L}^{-1}\left(\frac{e^{-2s}}{s-1}\right) &= 2e^{(t-2)}U(t-2)
\end{aligned}
$$

Putting all these results back into (2) gives the response in time domain as

$$y(t) = -2e^{3t} + e^{2t} + 2e^{t} - e^{2}\left(2e^{3(t-2)} - 4e^{2(t-2)} + 2e^{(t-2)}\right)U(t-2)$$

The above can also be written as

$$y(t) = \begin{cases} -2e^{3t} + e^{2t} + 2e^{t} & 0 < t < 2 \\ -2e^{3t} + e^{2t} + 2e^{t} - e^{2}\left(2e^{3t} - 4e^{2t} + 2e^{t}\right) & t \geq 2 \end{cases}$$

Since the original ODE is not stable (due to damping term -5 negative in the given ODE, the solution will blow up with time). This is seen by the solution above, where the exponential are all positive, hence growing with time. The following is a plot of the above solution for up to $t = 2.2$

```
ClearAll[t];
mySol = -2 e^(3 t) + e^(2 t) + 2 e^t - e^2 (2 e^(3 (t-2)) - 4 e^(2 (t-2)) + 2 e^(t-2)) UnitStep[t - 2];
Plot[mySol, {t, 0, 2.2},
 Frame -> True,
 FrameLabel → {{"y(t)", None}, {"t (sec)", "Solution to problem 2"}},
 GridLines → Automatic, GridLinesStyle → LightGray, PlotStyle → Red]
```

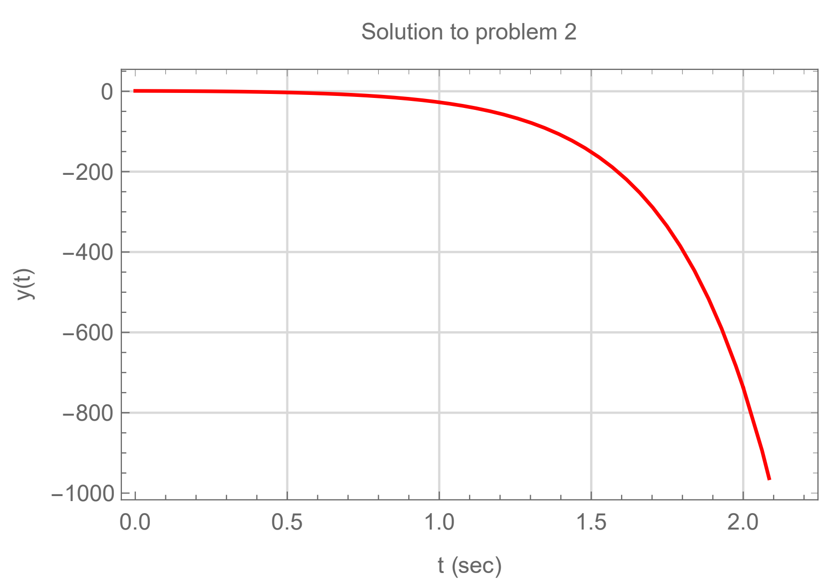


Figure 4: Plot of solution for problem 2

3 Problem 3

3. Solve the following linear system using Laplace transform

$$
\begin{aligned}
y_1'' + 5y_1 + y_2 &= 0\\
y_2'' - 2y_1 + 2y_2 &= 0\\
y_1(0) &= 3\\
y_1'(0) &= 0\\
y_2(0) &= 1\\
y_2'(0) &= 0
\end{aligned}
$$

Figure 5: Problem 3 Statement

<u>Solution</u>

Let $Y_1(s) = \mathcal{L}[y_1(t)]$ and let $Y_2(s) = \mathcal{L}[y_2(t)]$. Taking Laplace transform of the two ODE's gives

$$
\begin{aligned}
s^2Y_1(s) - sy_1(0) - y_1'(0) + 5Y_1(s) + Y_2(s) &= 0\\
s^2Y_2(s) - sy_2(0) - y_2'(0) - 2Y_1(s) + 2Y_2(s) &= 0
\end{aligned}
$$

Substituting the given initial conditions results in

$$s^2Y_1 - 3s + 5Y_1 + Y_2 = 0 \tag{1}$$

$$s^2Y_2 - s - 2Y_1 + 2Y_2 = 0 \tag{2}$$

The above two ODE's are now solved for $Y_1(s), Y_2(s)$

$$
\begin{aligned}
Y_1\left(s^2+5\right) + Y_2 &= 3s\\
Y_2\left(s^2+2\right) - 2Y_1 &= s
\end{aligned}
$$

or

$$\begin{pmatrix} s^2+5 & 1\\ -2 & s^2+2 \end{pmatrix}\begin{pmatrix} Y_1\\ Y_2 \end{pmatrix} = \begin{pmatrix} 3s\\ s \end{pmatrix}$$

Using Gaussian elimination: Adding $\left(\frac{2}{s^2+5}\right)$ times first row to second row gives

$$\begin{pmatrix} s^2+5 & 1 \\ 0 & s^2+2+\frac{2}{s^2+5} \end{pmatrix}\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} = \begin{pmatrix} 3s \\ s+\frac{6s}{s^2+5} \end{pmatrix}$$
$$\begin{pmatrix} s^2+5 & 1 \\ 0 & \frac{1}{s^2+5}\left(s^4+7s^2+12\right) \end{pmatrix}\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} = \begin{pmatrix} 3s \\ \frac{s}{s^2+5}\left(s^2+11\right) \end{pmatrix}$$

Back substitution: From last row

$$\begin{aligned} Y_2(s) &= \frac{\frac{s}{s^2+5}\left(s^2+11\right)}{\frac{1}{s^2+5}\left(s^4+7s^2+12\right)} \\ &= \frac{s\left(s^2+11\right)}{s^4+7s^2+12} \\ &= \frac{s\left(s^2+11\right)}{\left(s^2+4\right)\left(s^2+3\right)} \end{aligned} \tag{3}$$

First row gives

$$\left(s^2+5\right) Y_1 + Y_2 = 3s$$

Using $Y_2(s)$ found from (3), the above becomes

$$\begin{aligned} \left(s^2+5\right) Y_1 &= 3s - \frac{s\left(s^2+11\right)}{\left(s^2+4\right)\left(s^2+3\right)} \\ Y_1(s) &= \frac{3s}{\left(s^2+5\right)} - \frac{s\left(s^2+11\right)}{\left(s^2+5\right)\left(s^2+4\right)\left(s^2+3\right)} \end{aligned} \tag{4}$$

To obtain the time domain solution we need to inverse Laplace transform (3,4). Starting with (3), and applying partial fractions gives

$$Y_2(s) = \frac{s\left(s^2+11\right)}{\left(s^2+4\right)\left(s^2+3\right)} = \frac{8s}{3+s^2} - \frac{7s}{4+s^2} \tag{5}$$

From tables we see that

$$\begin{aligned} 8\mathcal{L}^{-1}\left[\frac{s}{3+s^2}\right] &= 8\cos\left(\sqrt{3}t\right) \\ 7\mathcal{L}^{-1}\left[\frac{s}{4+s^2}\right] &= 7\cos(2t) \end{aligned}$$

Hence (5) becomes in time domain as

$$y_2(t) = 8\cos\left(\sqrt{3}t\right) - 7\cos(2t)$$

Similarly for $Y_1(s)$, from (4) and applying partial fractions

$$\begin{aligned} Y_1(s) &= \frac{3s}{\left(s^2+5\right)} - \frac{s\left(s^2+11\right)}{\left(s^2+5\right)\left(s^2+4\right)\left(s^2+3\right)} \\ &= \frac{3s}{\left(s^2+5\right)} - \left(\frac{4s}{s^2+3} - \frac{7s}{s^2+4} + \frac{3s}{5+s^2}\right) \end{aligned} \tag{6}$$

From inverse Laplace transform table

$$
\begin{aligned}
3\mathcal{L}^{-1}\left[\frac{s}{(s^2+5)}\right] &= 3\cos\left(\sqrt{5}t\right)\\
4\mathcal{L}^{-1}\left[\frac{s}{s^2+3}\right] &= 4\cos\left(\sqrt{3}t\right)\\
7\mathcal{L}^{-1}\left[\frac{s}{s^2+4}\right] &= 7\cos(2t)\\
3\mathcal{L}^{-1}\left[\frac{s}{5+s^2}\right] &= 3\cos\left(\sqrt{5}t\right)
\end{aligned}
$$

Using these in (6), the solution $y_1(t)$ becomes

$$
\begin{aligned}
y_1(t) &= 3\cos\left(\sqrt{5}t\right) - \left(4\cos\left(\sqrt{3}t\right) - 7\cos(2t) + 3\cos\left(\sqrt{5}t\right)\right)\\
&= -4\cos\left(\sqrt{3}t\right) + 7\cos(2t)
\end{aligned}
$$

In summary

$$
\begin{aligned}
y_1(t) &= -4\cos\left(\sqrt{3}t\right) + 7\cos(2t)\\
y_2(t) &= 8\cos\left(\sqrt{3}t\right) - 7\cos(2t)
\end{aligned}
$$

The following is a plot of the solutions for 10 seconds.

```
ClearAll[y1, y2, t]
y1 = -4 Cos[√3 t] + 7 Cos[2 t];
y2 = 8 Cos[√3 t] - 7 Cos[2 t];
Plot[{y1, y2}, {t, 0, 10}, Frame → True,
 FrameLabel → {{"y1(t),y2(t)", None}, {"t (sec)", "Problem 3 solution"}},
 PlotLegends → {"y1(t)", "y2(t)"},
 BaseStyle → 14,
 GridLines → Automatic,
 GridLinesStyle → LightGray]
```

Out[ ]=

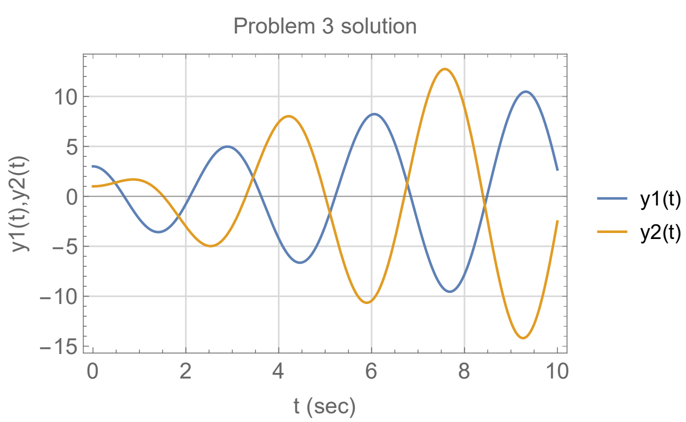


Figure 6: Plot of solution for problem 3

4 Problem 4

4. Consider an $n \times n$ real matrix A. Show that A can be uniquely written as a sum of a symmetric and a skew symmetric matrix

$$A = A_{sy} + A_{sk}.$$

We say that A is positive definite if $q(\mathbf{x}) > 0$ for $\mathbf{x} \neq \mathbf{0}$ where the quadratic form $q : \mathbb{R}^n \longrightarrow \mathbb{R}$ is defined as $q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$.

Show that A is positive definite if and only if A_{sy} is positive definite.

Figure 7: Problem 4 Statement

Solution

4.1 Part a

Note: In all the following it is assumed that x is a vector and that $x \neq 0$

Let

$$A = A_{sy} + A_{sk} \tag{1}$$

Where A_{sy} is a symmetric matrix, which means $A_{sy}^T = A_{sy}$ and A_{sk} is skew symmetric matrix which means $A_{sk}^T = -A_{sk}$. Taking the transpose of (1) gives

$$\begin{aligned} A^T &= \left(A_{sy} + A_{sk}\right)^T \\ &= A_{sy}^T + A_{sk}^T \\ &= A_{sy} - A_{sk} \end{aligned} \tag{2}$$

Adding (1)+(2) gives

$$\begin{aligned} A + A^T &= 2A_{sy} \\ A_{sy} &= \frac{A + A^T}{2} \end{aligned} \tag{3}$$

Subtracting (2)-(1) gives

$$\begin{aligned} A^T - A &= -2A_{sk} \\ A_{sk} &= \frac{A - A^T}{2} \end{aligned} \tag{4}$$

Therefore for any A,

$$A_{sy} = \frac{1}{2}\left(A + A^T\right) \tag{4A}$$

$$A_{sk} = \frac{1}{2}\left(A - A^T\right) \tag{4B}$$

To show that A_{sy} is indeed symmetric, this is done by construction :

$$\begin{aligned} A_{sy}^T &= \frac{1}{2}\left(A + A^T\right)^T \\ &= \frac{1}{2}\left(A^T + \left(A^T\right)^T\right) \end{aligned}$$

But $\left(A^T\right)^T = A$, and the above becomes

$$\begin{aligned} A_{sy}^T &= \frac{1}{2}\left(A^T + A\right) \\ &= A_{sy} \end{aligned}$$

Therefore A_{sy} is indeed symmetric.

To show that A_{sk} is skew symmetric matrix :

$$\begin{aligned} A_{sk}^T &= \frac{1}{2}\left(A - A^T\right)^T \\ &= \frac{1}{2}\left(A^T - \left(A^T\right)^T\right) \\ &= \frac{1}{2}\left(A^T - A\right) \\ &= -\frac{1}{2}\left(A - A^T\right) \\ &= -A_{sk} \end{aligned}$$

Hence A_{sk} is indeed skew symmetric.

Therefore any A matrix can be written as $A = A_{sy} + A_{sk}$ where A_{sy}, A_{sk} are given by (4A,4B).

Now we need to show that this is a unique was to write A. Proof is by contradictions. Let there be $\tilde{A}_{sy}$ matrix which is symmetric and $\tilde{A}_{sy} \neq A_{sy}$ and let there be $\tilde{A}_{sy}$ matrix which is skew matrix and $\tilde{A}_{sk} \neq A_{sk}$. And also let $A = \tilde{A}_{sy} + \tilde{A}_{sy}$ in addition to $A = A_{sy} + A_{sk}$. Then

$$\begin{aligned} A^T &= \left(\tilde{A}_{sy} + \tilde{A}_{sk}\right)^T \\ &= \tilde{A}_{sy}^T + \tilde{A}_{sk}^T \end{aligned}$$

Since $\tilde{A}_{sy}$ is assumed to be symmetric, then $\tilde{A}_{sy}^T = \tilde{A}_{sy}$ and since $\tilde{A}_{sk}$ is assumed to be skew symmetric, then $\tilde{A}_{sk}^T = -\tilde{A}_{sk}$ and the above becomes

$$A^T = \tilde{A}_{sy} - \tilde{A}_{sk}$$

Therefore

$$\begin{aligned}\frac{1}{2}\left(A+A^T\right) &= \frac{1}{2}\left(\tilde{A}_{sy}+\tilde{A}_{sk}+\tilde{A}_{sy}-\tilde{A}_{sk}\right)\\ &= \tilde{A}_{sy}\end{aligned}$$

But from (4A) above, we showed that $\frac{1}{2}\left(A+A^T\right) = A_{sy}$. Hence

$$A_{sy} = \tilde{A}_{sy}$$

Which is a contradiction to our assumption that $\tilde{A}_{sy} \neq A_{sy}$. Therefore A_{sy} is unique. The same is done for $\tilde{A}_{sk}$. From

$$\begin{aligned}\frac{1}{2}\left(A-A^T\right) &= \frac{1}{2}\left(\tilde{A}_{sy}+\tilde{A}_{sk}-\left(\tilde{A}_{sy}-\tilde{A}_{sk}\right)\right)\\ &= \tilde{A}_{sk}\end{aligned}$$

But from (4) above, we showed that $\frac{1}{2}\left(A-A^T\right) = A_{sk}$. Hence $A_{sk} = \tilde{A}_{sk}$ which is a contradiction. Therefore there is only way to write A as sum of symmetric and skew symmetric way, which is

$$A = \overbrace{\frac{A+A^T}{2}}^{A_{sy}} + \overbrace{\frac{1}{2}\left(A-A^T\right)}^{A_{sk}}$$

QED.

4.2 Part b

Starting with the forward direction. We need to show that given A is positive definite (p.d.) then this implies A_{sy} is also p.d.

From part (a) we found that A can be written as $A = A_{sy} + A_{sk}$. Since A is now assumed to be p.d. then this implies

$$\begin{aligned}x^T A x &> 0\\ x^T\left(A_{sy}+A_{sk}\right)x &> 0\\ x^T A_{sy} x + x^T A_{sk} x &> 0 \qquad (1)\end{aligned}$$

Now we will show that $x^T A_{sk} x = 0$ to finish the above proof. First we observe that

$$\begin{aligned}\left(x^T A_{sk} x\right)^T &= \left(A_{sk}x\right)^T x\\ &= x^T A_{sk}^T x\end{aligned}$$

But $A_{sk} = -A_{sk}^T$ by definition of skew symmetric matrix. Therefore the above becomes

$$\left(x^T A_{sk} x\right)^T = -\left(x^T A_{sk} x\right)$$

But $x^T A_{sk} x$ is a single number, say q. (To be precise, q is 1×1 matrix. but since it is 1×1 we can treat it as a number, since it is one element). But the transpose of a number (or 1×1 matrix) is itself. Hence the above relation says that

$$q^T = -q$$

For a number, this is the same as saying $q = -q$ and this only possible if $q = 0$ or in other words

$$x^T A_{sk} x = 0 \tag{2}$$

Using (2) in (1) shows immediately that

$$x^T A_{sy} x > 0$$

Therefore A_{sy} is positive definite.

Now we need to show the reverse direction. That is, we need to show that if A_{sy} is p.d. then this implies A is also p.d.

Since A_{sy} is now assumed to be p.d. then we can write

$$x^T A_{sy} x > 0$$

But $A = A_{sy} + A_{sk}$ therefore $A_{sy} = A - A_{sk}$ and the above becomes

$$x^T (A - A_{sk}) x > 0$$
$$x^T A x - x^T A_{sk} x > 0$$

But we showed in (2) that $x^T A_{sk} x = 0$. Therefore the above becomes

$$x^T A x > 0$$

Which implies that A is positive definite, which is what we are asked to show. QED

5 Problem 5

5. Consider

$$\mathbf{u_1} = [2 \quad 3 \quad 0 \quad 1], \quad \mathbf{u_2} = [-1 \quad 0 \quad 3 \quad 2]$$
$$\mathbf{u_3} = [2 \quad 2 \quad 1 \quad 4], \quad \mathbf{u_4} = [-6 \quad 4 \quad -2 \quad 0].$$

Show that three vectors $\mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3}$ out of the above four form an orthogonal set in $\mathbb{R}^4$. What are these vectors?
Find the best approximation $\mathbf{w} \in Span(\mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3})$ to the fourth vector using Gramm's theorem.
Find a basis $\mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3}, \mathbf{v_4}$ of orthogonal vectors for $\mathbb{R}^4$.

Figure 8: Problem 5 Statement

Solution

Two vectors $\vec{v}, \vec{u}$ are orthogonal if their dot product is zero. This is because $\vec{v} \cdot \vec{u} = \|\vec{v}\| \, \|\vec{u}\| \cos\theta$ where θ is the inner angle between the two vectors. Since the vectors are orthogonal, then $\cos 90^0 = 0$ and therefore $\vec{v} \cdot \vec{u} = 0$. To find which pairs are orthogonal to each others, we compute the inner product between all possible pairs :

$$\begin{aligned}
\vec{u}_1 \cdot \vec{u}_2 &= -2 + 0 + 0 + 2 = 0 \\
\vec{u}_1 \cdot \vec{u}_3 &= 4 + 6 + 0 + 4 = 14 \\
\vec{u}_1 \cdot \vec{u}_4 &= -12 + 12 + 0 + 0 = 0 \\
\vec{u}_2 \cdot \vec{u}_3 &= -2 + 0 + 3 + 8 = 9 \\
\vec{u}_2 \cdot \vec{u}_4 &= 6 + 0 - 6 + 0 = 0 \\
\vec{u}_3 \cdot \vec{u}_4 &= -12 + 8 - 2 + 0 = -2
\end{aligned}$$

We see from the above that $\vec{u}_1 \cdot \vec{u}_2 = 0, \vec{u}_1 \cdot \vec{u}_4 = 0, \vec{u}_2 \cdot \vec{u}_4 = 0$. Therefore

$$\begin{aligned}
S &= \{\vec{u}_1, \vec{u}_2, \vec{u}_4\} \\
&\equiv \{\vec{v}_1, \vec{v}_2, \vec{v}_3\}
\end{aligned}$$

Or

$$\begin{aligned}
\vec{v}_1 &= (2, 3, 0, 1) \\
\vec{v}_2 &= (-1, 0, 3, 2) \\
\vec{v}_3 &= (-6, 4, -2, 0)
\end{aligned} \tag{1B}$$

Form an orthogonal set in $\mathbb{R}^4$.

Now we need to find the best approximation of $\vec{w}=\vec{u}_3=(2,2,1,4)$ using the above orthogonal vectors $\vec{v}_1,\vec{v}_2,\vec{v}_3$. Using Gram's theorem, this approximation is

$$\vec{w}=c_1\vec{v}_1+c_2\vec{v}_2+c_3\vec{v}_3 \tag{1A}$$

Where the constants c_i are found from solving the system

$$\begin{pmatrix}\vec{v}_1\cdot\vec{v}_1 & \vec{v}_1\cdot\vec{v}_2 & \vec{v}_1\cdot\vec{v}_3\\ \vec{v}_2\cdot\vec{v}_1 & \vec{v}_2\cdot\vec{v}_2 & \vec{v}_2\cdot\vec{v}_3\\ \vec{v}_3\cdot\vec{v}_1 & \vec{v}_3\cdot\vec{v}_2 & \vec{v}_3\cdot\vec{v}_3\end{pmatrix}\begin{pmatrix}c_1\\c_2\\c_3\end{pmatrix}=\begin{pmatrix}\vec{v}_1\cdot\vec{w}\\ \vec{v}_2\cdot\vec{w}\\ \vec{v}_3\cdot\vec{w}\end{pmatrix}$$

But since $\vec{v}_i$ are all orthogonal to each others then $\vec{v}_i\cdot\vec{v}_j=0$ for $i\neq j$, and $\vec{v}_i\cdot\vec{v}_i=\|v_i\|^2$ and the above becomes

$$\begin{pmatrix}\|\vec{v}_1\|^2 & 0 & 0\\ 0 & \|\vec{v}_2\|^2 & 0\\ 0 & 0 & \|\vec{v}_3\|^2\end{pmatrix}\begin{pmatrix}c_1\\c_2\\c_3\end{pmatrix}=\begin{pmatrix}\vec{v}_1\cdot\vec{w}\\ \vec{v}_2\cdot\vec{w}\\ \vec{v}_3\cdot\vec{w}\end{pmatrix} \tag{1}$$

But

$$\begin{aligned}\vec{v}_1\cdot\vec{w}&=(2,3,0,1)\cdot(2,2,1,4)=4+6+4=14\\ \vec{v}_2\cdot\vec{w}&=(-1,0,3,2)\cdot(2,2,1,4)=-2+3+8=9\\ \vec{v}_3\cdot\vec{w}&=(-6,4,-2,0)\cdot(2,2,1,4)=-12+8-2=-6\end{aligned}$$

Hence (1) becomes

$$\begin{pmatrix}\|\vec{v}_1\|^2 & 0 & 0\\ 0 & \|\vec{v}_2\|^2 & 0\\ 0 & 0 & \|\vec{v}_3\|^2\end{pmatrix}\begin{pmatrix}c_1\\c_2\\c_3\end{pmatrix}=\begin{pmatrix}14\\9\\-6\end{pmatrix} \tag{2}$$

Since

$$\begin{aligned}\|\vec{v}_1\|^2&=\|(2,3,0,1)\|^2=4+9+1=14\\ \|\vec{v}_2\|^2&=\|(-1,0,3,2)\|^2=1+9+4=14\\ \|\vec{v}_3\|^2&=\|(-6,4,-2,0)\|^2=36+16+4=56\end{aligned}$$

Then (2) becomes

$$\begin{pmatrix}14 & 0 & 0\\ 0 & 14 & 0\\ 0 & 0 & 56\end{pmatrix}\begin{pmatrix}c_1\\c_2\\c_3\end{pmatrix}=\begin{pmatrix}14\\9\\-6\end{pmatrix} \tag{3}$$

From the above we see that

$$\begin{aligned}c_1&=1\\ c_2&=\frac{9}{14}\\ c_3&=\frac{-3}{28}\end{aligned}$$

Hence the best approximation using (1A) becomes

$$\begin{aligned}\vec{w} &= c_1\vec{v}_1 + c_2\vec{v}_2 + c_3\vec{v}_3\\ &= (2,3,0,1) + \frac{9}{14}(-1,0,3,2) - \frac{3}{28}(-6,4,-2,0)\\ &= \left(2, \frac{18}{7}, \frac{15}{7}, \frac{16}{7}\right)\end{aligned}$$

Therefore

$$\vec{w} = \frac{1}{7}(14, 18, 15, 16)$$

Now we need to find basis $\vec{v}_1, \vec{v}_2, \vec{v}_3, \vec{v}_4$ of orthogonal vectors in $\mathbb{R}^4$. We already found that from (1B) that $\vec{v}_1, \vec{v}_2, \vec{v}_3$ are three such vectors. So we just need to find another $\vec{v}_4 = [a_1, a_2, a_3, a_4]$ such that it is orthogonal to the other three, in other words we need to solve

$$\begin{aligned}\vec{v}_1 \cdot \vec{v}_4 &= 0\\ \vec{v}_2 \cdot \vec{v}_4 &= 0\\ \vec{v}_3 \cdot \vec{v}_4 &= 0\end{aligned}$$

This implies

$$\begin{aligned}[2,3,0,1] \cdot [a_1, a_2, a_3, a_4] &= 0\\ [-1,0,3,2] \cdot [a_1, a_2, a_3, a_4] &= 0\\ [-6,4,-2,0] \cdot [a_1, a_2, a_3, a_4] &= 0\end{aligned}$$

Or

$$\begin{aligned}2a_1 + 3a_2 + a_4 &= 0\\ -a_1 + 3a_3 + 2a_4 &= 0\\ -6a_1 + 4a_2 - 2a_3 &= 0\end{aligned}$$

Or

$$\begin{pmatrix}2 & 3 & 0 & 1\\ -1 & 0 & 3 & 2\\ -6 & 4 & -2 & 0\end{pmatrix}\begin{pmatrix}a_1\\ a_2\\ a_3\\ a_4\end{pmatrix} = \begin{pmatrix}0\\0\\0\end{pmatrix}$$

This system has three equations and 4 unknowns. Therefore it will have one free parameter giving an infinite number of solutions. Using Gaussian elimination:

$$\begin{pmatrix}2 & 3 & 0 & 1\\ -1 & 0 & 3 & 2\\ -6 & 4 & -2 & 0\end{pmatrix} \overset{R_2=R_2+\frac{1}{2}R_1}{\longrightarrow} \begin{pmatrix}2 & 3 & 0 & 1\\ 0 & \frac{3}{2} & 3 & \frac{5}{2}\\ -6 & 4 & -2 & 0\end{pmatrix} \overset{R_3=R_3+3R_1}{\longrightarrow}$$

$$\begin{pmatrix}2 & 3 & 0 & 1\\ 0 & \frac{3}{2} & 3 & \frac{5}{2}\\ 0 & 13 & -2 & 3\end{pmatrix} \overset{R_3=R_3-\frac{26}{3}R_2}{\longrightarrow} \begin{pmatrix}2 & 3 & 0 & 1\\ 0 & \frac{3}{2} & 3 & \frac{5}{2}\\ 0 & 0 & -28 & -\frac{56}{3}\end{pmatrix}$$

We stop the elimination here since no more elimination is possible. We have now this system

$$\begin{pmatrix} 2 & 3 & 0 & 1 \\ 0 & \frac{3}{2} & 3 & \frac{5}{2} \\ 0 & 0 & -28 & -\frac{56}{3} \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Back substitution: From last row we obtain the equation

$$\begin{aligned} -28a_3 - \frac{56}{3}a_4 &= 0 \\ a_4 &= -\frac{3}{2}a_3 \end{aligned}$$

The second row gives

$$\begin{aligned} \frac{3}{2}a_2 + 3a_3 + \frac{5}{2}a_4 &= 0 \\ a_2 &= \frac{2}{3}\left(-3a_3 - \frac{5}{2}a_4\right) \\ &= -2a_3 - \frac{5}{3}a_4 \end{aligned}$$

Since $a_4 = -\frac{3}{2}a_3$ the above becomes

$$\begin{aligned} a_2 &= -2a_3 - \frac{5}{3}\left(-\frac{3}{2}a_3\right) \\ &= \frac{1}{2}a_3 \end{aligned}$$

First row gives

$$\begin{aligned} 2a_1 + 3a_2 + a_4 &= 0 \\ a_1 &= \frac{1}{2}\left(-3a_2 - a_4\right) \end{aligned}$$

Since $a_2 = \frac{1}{2}a_3$ and $a_4 = -\frac{3}{2}a_3$ the above becomes

$$\begin{aligned} a_1 &= \frac{1}{2}\left(-3\left(\frac{1}{2}a_3\right) - \left(-\frac{3}{2}a_3\right)\right) \\ &= 0 \end{aligned}$$

Therefore the solution is

$$\begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{pmatrix} = \begin{pmatrix} 0 \\ \frac{1}{2} \\ 1 \\ -\frac{3}{2} \end{pmatrix} a_3$$

The above means that for any arbitrary a_3 value there is a solution. a_3 is just a scalar which only stretches or shrinks the vector but does not change its direction (orientation). Therefore the vector remains orthogonal to all others for any a_3. Let us pick $a_3 = 1$. Using this $\vec{v}_4$ becomes

$$\vec{v}_4 = \left[0, \frac{1}{2}, 1, -\frac{3}{2}\right]$$

To verify the result found, we will check that $\vec{v}_4$ is indeed orthogonal with the other three vectors :

$$\vec{v}_1 \cdot \vec{v}_4 = [2, 3, 0, 1] \cdot \left[0, \frac{1}{2}, 1, -\frac{3}{2}\right] = 0$$
$$\vec{v}_2 \cdot \vec{v}_4 = [-1, 0, 3, 2] \cdot \left[0, \frac{1}{2}, 1, -\frac{3}{2}\right] = 0$$
$$\vec{v}_3 \cdot \vec{v}_4 = [-6, 4, -2, 0] \cdot \left[0, \frac{1}{2}, 1, -\frac{3}{2}\right] = 0$$

QED.

6 Problem 6

6. For a general $n \times n$ nonsingular matrix A compute the total number of multiplications/divisions necessary to solve the linear system

$$A\mathbf{x} = \mathbf{b},$$

a) using Gauss elimination with back-substitution,
b) using Cramer's rule.

Figure 9: Problem 6 Statement

Solution

Let

$$A_{n\times n} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{pmatrix}, b = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_n \end{pmatrix}$$

6.1 Part (a)

The first step in Gaussian elimination is to reduce the above matrix to row echelon form :

$$A_{n\times n} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & a_{22} & a_{23} & \cdots & a_{2n} \\ 0 & 0 & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} b_1 \\ \tilde{b}_2 \\ \tilde{b}_3 \\ \vdots \\ \tilde{b}_n \end{pmatrix}$$

Row echelon form has zeros in its the lower left triangle. After this, back substitution starts by solving for x_n from the last row, then solving for x_{n-1} from the row above the last row and so on until we reach the first row.

Counting operations for forward pass

The first step is to zero out all entries in first column below a_{11} using a_{11} as pivot. Next is to zero out all entries in second column below the (updated) a_{22} value and so on.

To zero out an entry, for example a_{21}, we first need to do one division $\frac{a_{21}}{a_{11}} = \Delta$ and store this in memory, then do $a_{2i} = a_{2i} - \Delta a_{1i}$ for all entries in that row, which means for $i = 1 \cdots n$. (no need to count a_{21} since we know it will be zero). We have to remember that this is being applied to the b vector as well and not just for A matrix rows.

Hence we need one division to find Δ, and then $2n$ multiplication and addition/subtraction operations per row. The division is only needed once per row to find the pivot scaling Δ.

Since there are $n-1$ rows then there are $(n-1)$ divisions and $(2n)(n-1)$ multiplications/addition to zero out the first column. After this we have the following system reached

$$A_{n\times n} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & a_{22} & a_{23} & \cdots & a_{2n} \\ 0 & a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n2} & a_{n3} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} b_1 \\ \tilde{b}_2 \\ \tilde{b}_3 \\ \vdots \\ \tilde{b}_n \end{pmatrix}$$

The total cost now is therefore $(n-1) + (2n)(n-1)$.

We now switch to the second row and use the new value of a_{22} as <u>pivot</u> and repeat the same as above. The only difference now is that there are $n-2$ rows to process and $(n-2)$ divisions and therefore $2(n-1)(n-2)$ multiplications/addition to zero out the second column entries below the second row. After this we reach the following system

$$A_{n\times n} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & a_{22} & a_{23} & \cdots & a_{2n} \\ 0 & 0 & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & a_{n3} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} b_1 \\ \tilde{b}_2 \\ \tilde{b}_3 \\ \vdots \\ \tilde{b}_n \end{pmatrix}$$

The total cost of the above is therefore $(n-2) + 2(n-1)(n-2)$.

We now switch to the third row and use the new value of a_{33} as <u>pivot</u>. Now there are now $(n-3)$ divisions and $2(n-2)(n-3)$ multiplications/additions to obtain the following system

$$A_{n\times n} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & a_{22} & a_{23} & \cdots & a_{2n} \\ 0 & 0 & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} b_1 \\ \tilde{b}_2 \\ \tilde{b}_3 \\ \vdots \\ \tilde{b}_n \end{pmatrix}$$

The total cost of the above is therefore $(n-3) + 2(n-2)(n-3)$.

And so on until we reach the row before the last row, where there is only one row below it to process. The cost then is just one division and 2 additions and 2 multiplications. Therefore the number of total number of multiplication and additions operations for the forward pass is the sum of all the above operations, which can be written as

$$\overbrace{[(n-1)+(2n)(n-1)]}^{\text{row 2}}+\overbrace{[(n-2)+2(n-1)(n-2)]}^{\text{row 3}}+\cdots+\overbrace{[1+4]}^{\text{last row}}$$

Writing the above as $\sum_{k=1}^{n-1}(n-k)+2(n-k+1)(n-k)$ then we need to calculate this sum using known formulas for summations. Let this sum be Δ, hence

$$\begin{aligned}
\Delta &= \sum_{k=1}^{n-1} -3k+2k^2+3n-4kn+2n^2\\
&= -3\sum_{k=1}^{n-1}k+2\sum_{k=1}^{n-1}k^2+3\sum_{k=1}^{n-1}n-4\sum_{k=1}^{n-1}kn+2\sum_{k=1}^{n-1}n^2\\
&= -3\left(\frac{n(n-1)}{2}\right)+2\left(\frac{n^3}{3}-\frac{n^2}{2}+\frac{n}{6}\right)+3\left(n^2-n\right)-4\left(\frac{n^2(n-1)}{2}\right)+2\left(n^2(n-1)\right)\\
&= \frac{2}{3}n^3+\frac{1}{2}n^2-\frac{7}{6}n \qquad (1)
\end{aligned}$$

The above is the number of operations just for for the forward pass (elimination phase).

For example for matrix of size 3×3 the above gives 19 operations, and for matrix of size 4×4, it gives 46 operations and for 5×5 it gives 90 operations and so on.

Counting operations for backward pass In back substitution, we start from the end of the elimination phase above, which will be

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n}\\ 0 & a_{22} & a_{23} & \cdots & a_{2n}\\ 0 & 0 & a_{33} & \cdots & a_{3n}\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & 0 & \cdots & a_{nn}\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\\\vdots\\x_n\end{pmatrix}=\begin{pmatrix}b_1\\b_2\\b_3\\\vdots\\b_n\end{pmatrix}$$

First step is to solve for x_n by finding $x_n=\frac{b_n}{a_{nn}}$. This requires only one division. Next is to solve for x_{n-1} by finding $a_{n-1,n-1}x_{n-1}+a_{n-1,n}x_n=b_{n-1}$, or

$$x_{n-1}=\frac{b_{n-1}-\left(a_{n-1,n}\right)x_n}{\left(a_{n-1,n-1}\right)}$$

We see that this needs one subtraction, one multiplication and one division, or 3 operations. The next step is to solve for x_{n-2} from

$$\left(a_{n-2,n-2}\right)x_{n-2}+\left(a_{n-2,n-1}\right)x_{n-1}+\left(a_{n-2,n}\right)x_n=b_2$$

Hence

$$x_{n-2}=\frac{b_2-\left(a_{n-2,n-1}\right)x_{n-1}-\left(a_{n-2,n}\right)x_n}{\left(a_{n-2,n-2}\right)}$$

Therefore we need 2 subtractions, 2 multiplication and one division, or 5 operations. And so on until we reach the first row to solve for x_1. Therefore the total number of operations can be seen as

$$1+3+5+7+\cdots$$

The above can be written as the sum

$$\begin{aligned}\sum_{k=0}^{n-1}(2k+1) &= 2\sum_{k=0}^{n-1}k+\sum_{k=0}^{n-1}1\\ &= 2\left(\frac{n(n-1)}{2}\right)+n\\ &= n(n-1)+n\\ &= n^2-n+n\\ &= n^2 \qquad (2)\end{aligned}$$

We see that the cost of the elimination is much greater than the cost of back substitution. One is $O\left(n^3\right)$ while the other is $O\left(n^2\right)$.

From (1,2), the total number of operations for the complete Gaussian elimination process is

$$\begin{aligned}\Delta &= \frac{2}{3}n^3+\frac{1}{2}n^2-\frac{7}{6}n+n^2\\ &= \frac{2}{3}n^3+\frac{3}{2}n^2-\frac{7}{6}n\end{aligned}$$

For large n the above is $O\left(n^3\right)$.

6.2 Part (b)

Given a system of equations

$$\begin{pmatrix}a_{11}&a_{12}&a_{13}&\cdots&a_{1n}\\a_{21}&a_{22}&a_{23}&\cdots&a_{2n}\\a_{31}&a_{32}&a_{33}&\cdots&a_{3n}\\\vdots&\vdots&\vdots&\ddots&\vdots\\a_{n1}&a_{n2}&a_{n3}&\cdots&a_{nn}\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\\\vdots\\x_n\end{pmatrix}=\begin{pmatrix}b_1\\b_2\\b_3\\\vdots\\b_n\end{pmatrix}$$

Cramer method works as follows :

$$\begin{aligned}x_1 &= \frac{|A|}{|A_1|}\\ x_2 &= \frac{|A|}{|A_2|}\\ &\vdots\\ x_n &= \frac{|A|}{|A_n|}\end{aligned}$$

Where $|A|$ is the determinant of coefficient matrix $A_{n\times n}$ and $|A_i|$ is determinant of coefficient matrix but with the i^{th} column replaced by the column vector b.

An efficient way to find the determinant is to convert the matrix to row echelon form. In this form, the matrix is upper triangle. Hence the determinant is the product of all elements along the diagonal. This is more efficient than using the matrix cofactor expansion method.

In doing these row operations on the matrix to find $|A|$ the only difference from the elimination steps we did for part(a), is that we have to remember the following rules now during the elimination process

1. When adding multiple of one row to another row, the determinant is not affected.
2. When switching two rows, the determinant is multiplied by -1
3. When multiplying one row by some scalar, the determinant is also multiplied by the same scalar.

Given the above, let us assume that for each elimination step of a row, we do one multiplication to account for a possible multiplication by -1 or possible multiplication by a scalar. Since we do not know if this will happen every time as this clearly depends on the data in the matrix, then this will be the worst case counting.

This means there is an additional $(n-1)$ multiplications to add to the cost of doing the elimination step to reach row echelon form at the end.

Another small difference from part(a), is that now we do not have the b vector added during the forward step.

Therefore, as we did in part(a), the cost to reach this form

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & a_{22} & a_{23} & \cdots & a_{2n} \\ 0 & a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n2} & a_{n3} & \cdots & a_{nn} \end{pmatrix}$$

Is now $(n-1)+(2(n-1))(n-1)$. Recalling from part(a) the cost at this stage was $(n-1)+(2n)(n-1)$ here. So we changed $2n$ to $2(n-1)$, since there is no b vector, hence one less element. And as was done in part (a), the cost to reach

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & a_{22} & a_{23} & \cdots & a_{2n} \\ 0 & 0 & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & a_{n3} & \cdots & a_{nn} \end{pmatrix}$$

Now becomes $(n-2)+2(n-2)(n-2)$. Recalling from part(a) the cost at this stage was $(n-1)+2(n-1)(n-1)$. So we changed $2(n-1)$ to $2(n-2)$, since there is no b vector. This continues to the row before the last as in part (a). Therefore the number of total multiplication and additions operations for just the forward pass is

$$\overbrace{[(n-1)+(2(n-1))(n-1)]}^{\text{row 2}}+\overbrace{[(n-2)+2(n-2)(n-2)]}^{\text{row 3}}+\cdots+\overbrace{[1+2]}^{\text{last row}}$$

Hence the cost to put the matrix in row echelon form is

$$\begin{aligned}\sum_{k=1}^{n-1}(n-k)+2(n-k)^2 &= \sum_{k=1}^{n-1}n-\sum_{k=1}^{n-1}k+2\sum_{k=1}^{n-1}(n-k)^2\\ &= n(n-1)-\frac{n(n-1)}{2}+2\frac{n-3n^2+2n^3}{6}\\ &= \frac{2}{3}n^3-\frac{1}{2}n^2-\frac{1}{6}n \qquad (1)\end{aligned}$$

The above cost is very close to part(a) elimination phase as expected which was $\frac{2}{3}n^3+\frac{1}{2}n^2-\frac{7}{6}n$. Current cost is smaller because in part(a) we had the b vector there which added more operations, while here we just operated on A itself.

Let us now add the $(n-1)$ multiplication we mentioned earlier to the result above. The cost now becomes

$$\begin{aligned}\Delta &= \frac{2}{3}n^3-\frac{1}{2}n^2-\frac{1}{6}n+(n-1)\\ &= \frac{2}{3}n^3-\frac{1}{2}n^2+\frac{5}{6}n-1\end{aligned}$$

We still need to calculate the product of the diagonal elements to find the determinant. For $n\times n$ matrix, this takes $n-1$ multiplications. Adding these to the above gives

$$\begin{aligned}\Delta &= \frac{2}{3}n^3-\frac{1}{2}n^2+\frac{5}{6}n-1+(n-1)\\ &= \frac{2}{3}n^3-\frac{1}{2}n^2+\frac{11}{6}n-2\\ &\approx \frac{2}{3}n^3-\frac{1}{2}n^2+\frac{11}{6}n\end{aligned}$$

We will use the above as the cost of finding the determinant.

How many times do we need to find determinants? We need to do it one time to find $|A|$ and then n more time for each $|A_i|$. Hence $(n+1)$ times. This is the main reason why Cramer method becomes much more costly compared to Gaussian elimination.

The number of operations now becomes

$$\begin{aligned}\Delta &= \left(\frac{2}{3}n^3-\frac{1}{2}n^2+\frac{11}{6}n\right)(n+1)\\ &= \frac{2}{3}n^4+\frac{1}{6}n^3+\frac{4}{3}n^2+\frac{11}{6}n\end{aligned}$$

We also need to add the cost of the final divisions $\frac{|A|}{|A_i|}$ to find each x_i. So we add n divisions to the above, giving the final cost as

$$\begin{aligned}\Delta &= \frac{2}{3}n^4 + \frac{1}{6}n^3 + \frac{4}{3}n^2 + \frac{11}{6}n + n \\ &= \frac{2}{3}n^4 + \frac{1}{6}n^3 + \frac{4}{3}n^2 + \frac{17}{6}n\end{aligned}$$

We see from above, that Cramer rule for large n is $O\left(n^4\right)$ while Gaussian elimination was $O\left(n^3\right)$. Hence Gaussian elimination is much more efficient for large n.

In summary

n	cost of Gaussian elimination $\frac{2}{3}n^3 + \frac{3}{2}n^2 - \frac{7}{6}n$	cost of Cramer $\frac{2}{3}n^4 + \frac{1}{6}n^3 + \frac{4}{3}n^2 + \frac{17}{6}n$
2	5	23
3	19	79
4	46	214
5	90	485
6	155	965
7	245	1743
8	364	2924
9	516	4629
10	705	6995

The following is a graphical illustration of the above

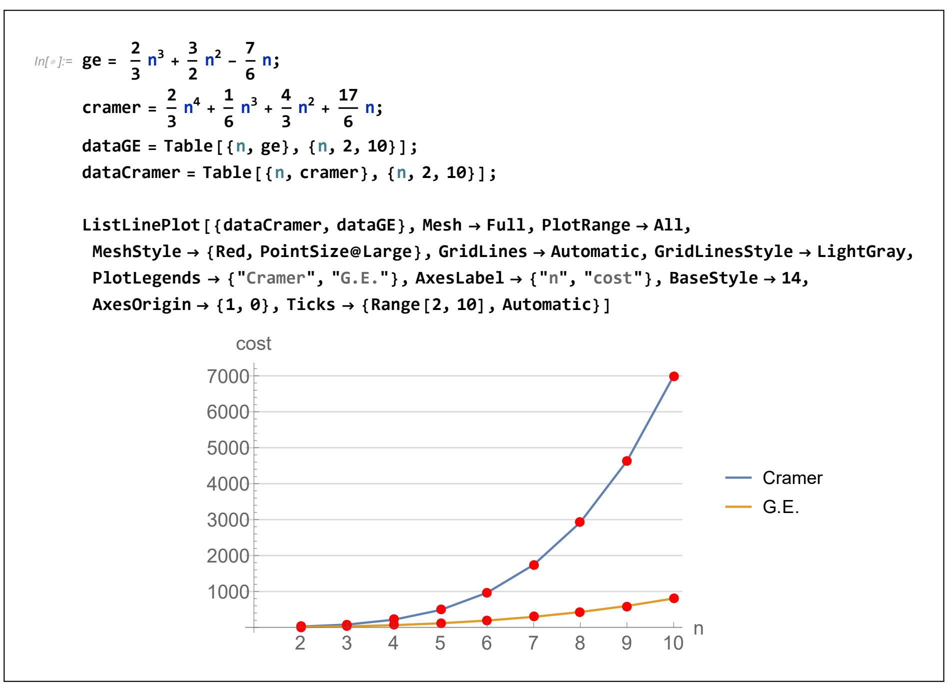


Figure 10: Cost of Gaussian elimination vs. Cramer method. Problem 6

7 Problem 7

7. Consider the matrix

$$A = \begin{bmatrix} \frac{1}{\sqrt{5}} & 0 & \frac{2}{\sqrt{5}} \\ \frac{-2}{\sqrt{5}} & 0 & \frac{1}{\sqrt{5}} \\ 0 & -1 & 0 \end{bmatrix}.$$

a) Show that its eigenvalues lie on the unit circle in $\mathbb{C}$.

b) Show that the linear transformation $T(\mathbf{x}) = A\mathbf{x}$ preserves the angles in $\mathbb{R}^3$, i.e. for any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^3$,

$$\angle(\mathbf{x}, \mathbf{y}) = \angle(A\mathbf{x}, A\mathbf{y}).$$

(Here $\angle(\mathbf{u}, \mathbf{v})$ denotes the angle between $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^3$. Note that $\angle(\mathbf{u}, \mathbf{v})$ can be expressed in terms of the norms and dot product of the vectors $\mathbf{u}$ and $\mathbf{v}$.)

Figure 11: Problem 7 Statement

Solution

7.1 Part a

The eigenvalues of A are found by solving $|A - \lambda I| = 0$ or

$$\begin{vmatrix} \frac{1}{\sqrt{5}} - \lambda & 0 & \frac{2}{\sqrt{5}} \\ \frac{-2}{\sqrt{5}} & -\lambda & \frac{1}{\sqrt{5}} \\ 0 & -1 & -\lambda \end{vmatrix} = 0$$

$$\left(\frac{1}{\sqrt{5}} - \lambda\right) \begin{vmatrix} -\lambda & \frac{1}{\sqrt{5}} \\ -1 & -\lambda \end{vmatrix} - 0 + \frac{2}{\sqrt{5}} \begin{vmatrix} \frac{-2}{\sqrt{5}} & -\lambda \\ 0 & -1 \end{vmatrix} = 0$$

$$\left(\frac{1}{\sqrt{5}} - \lambda\right)\left(\lambda^2 + \frac{1}{\sqrt{5}}\right) + \frac{2}{\sqrt{5}}\left(\frac{2}{\sqrt{5}}\right) = 0$$

$$\frac{1}{5}(\lambda - 1)\left(-5\lambda - 5\lambda^2 + \sqrt{5}\lambda - 5\right) = 0$$

$$\frac{1}{5}(\lambda - 1)\left(-5\lambda^2 + \lambda\left(\sqrt{5} - 5\right) - 5\right) = 0$$

$$(\lambda - 1)\left(-5\lambda^2 + \lambda\left(\sqrt{5} - 5\right) - 5\right) = 0$$

Hence $\lambda = 1$. The quadratic formula is used to solve $-5\lambda^2 + \lambda\left(\sqrt{5} - 5\right) - 5 = 0$. First it is normalized

$$5\lambda^2 - \lambda\left(\sqrt{5} - 5\right) + 5 = 0$$
$$\lambda^2 - \lambda\left(\frac{1}{5}\sqrt{5} - 1\right) + 1 = 0$$

Then $\lambda = \frac{-b}{2a} \pm \frac{1}{2a}\sqrt{b^2 - 4ac}$ where $b = -\left(\frac{1}{5}\sqrt{5} - 1\right)$, $c = 1$, $a = 1$ and the roots are

$$\begin{aligned}
\lambda &= \frac{\left(\frac{1}{5}\sqrt{5} - 1\right)}{2} \pm \frac{1}{2}\sqrt{\left(-\left(\frac{1}{5}\sqrt{5} - 1\right)\right)^2 - 4} \\
&= \left(\frac{\sqrt{5}}{10} - \frac{1}{2}\right) \pm \frac{1}{2}\sqrt{\left(\frac{1}{5}\sqrt{5} - 1\right)^2 - 4} \\
&= \left(\frac{\sqrt{5}}{10} - \frac{1}{2}\right) \pm \frac{1}{2}\sqrt{1 + \frac{5}{25} - \frac{2}{5}\sqrt{5} - 4} \\
&= \left(\frac{\sqrt{5}}{10} - \frac{1}{2}\right) \pm \frac{1}{2}\sqrt{1 + \frac{1}{5} - \frac{2}{5}\sqrt{5} - 4} \\
&= \left(\frac{\sqrt{5}}{10} - \frac{1}{2}\right) \pm \frac{1}{2}\sqrt{-\frac{14}{5} - \frac{2\sqrt{5}}{5}} \\
&= \left(\frac{\sqrt{5}}{10} - \frac{1}{2}\right) \pm \frac{1}{2}\sqrt{\frac{-14 - 2\sqrt{5}}{5}} \\
&= \left(\frac{\sqrt{5}}{10} - \frac{1}{2}\right) \pm \frac{1}{2}i\sqrt{\frac{14 + 2\sqrt{5}}{5}} \\
&= \left(\frac{\sqrt{5}}{10} - \frac{1}{2}\right) \pm \frac{1}{2}i\sqrt{\frac{5\left(14 + 2\sqrt{5}\right)}{(5)\,5}} \\
&= \left(\frac{\sqrt{5}}{10} - \frac{1}{2}\right) \pm \frac{1}{(2)\,(5)}i\sqrt{5\left(14 + 2\sqrt{5}\right)} \\
&= \left(\frac{\sqrt{5}}{10} - \frac{1}{2}\right) \pm \frac{1}{10}i\sqrt{10\sqrt{5} + 70}
\end{aligned}$$

Therefore the roots are

$$\begin{aligned}
\lambda_1 &= 1 \\
\lambda_2 &= \left(\frac{\sqrt{5}}{10} - \frac{1}{2}\right) + \frac{1}{10}i\sqrt{10\sqrt{5} + 70} \\
\lambda_3 &= \left(\frac{\sqrt{5}}{10} - \frac{1}{2}\right) - \frac{1}{10}i\sqrt{10\sqrt{5} + 70}
\end{aligned}$$

Numerically the above becomes

$$\begin{aligned}\lambda_1 &= 1\\ \lambda_2 &= -0.276 + 0.961i\\ \lambda_3 &= -0.276 - 0.961i\end{aligned}$$

The following plot shows the locations on the complex plane

```
eigs = {1, -0.276 + 0.961 I, -0.276 - 0.961 I};
eigs = ReIm@eigs;
img = Graphics[{
    {FaceForm[White], EdgeForm[Black], Disk[]},
    {Red, PointSize[.05], Point[eigs]}
   }, ImageSize → 300, Axes → True];
grid = Graphics[{}, GridLines → Automatic, PlotRangePadding → None,
   GridLinesStyle → LightGray, ImageSize → ImageDimensions@img];
Overlay[{img, grid}]
```

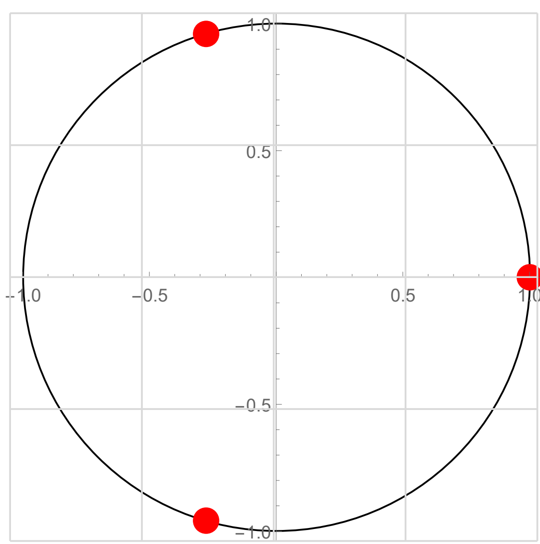


Figure 12: Graphical location of eigenvalues for problem 7

To show analytically that the eigenvalues lie on the unit circle means to show that the magnitude of each complex number is 1. Clearly λ_1 already satisfy this condition. We need to check now that $\|\lambda_2\| = 1$ and that $\|\lambda_3\| = 1$

$$\begin{aligned}
\|\lambda_2\| &= \sqrt{\operatorname{Re}(\lambda_2)^2 + \operatorname{Im}(\lambda_2)^2} \\
&= \sqrt{\left(\frac{\sqrt{5}}{10} - \frac{1}{2}\right)^2 + \left(\frac{1}{10}\sqrt{10\sqrt{5}+70}\right)^2} \\
&= \sqrt{\left(\frac{3}{10} - \frac{1}{10}\sqrt{5}\right) + \left(\frac{1}{10}\sqrt{5} + \frac{7}{10}\right)} \\
&= \sqrt{\frac{10}{10}} \\
&= 1
\end{aligned}$$

Similarly for λ_3 since it is the same except for the sign on the complex part (complex conjugate) which does not affect the norm. Therefore all the eigenvalues lie on unit circle in $\mathbb{C}$. QED.

7.2 Part b

Let two vectors in the domain of A be $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ and $\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$. And let the two vector after the mapping, which now lie in the range of A be $\tilde{\mathbf{x}} = \begin{pmatrix} \tilde{x}_1 \\ \tilde{x}_2 \\ \tilde{x}_3 \end{pmatrix}$ and $\tilde{\mathbf{y}} = \begin{pmatrix} \tilde{y}_1 \\ \tilde{y}_2 \\ \tilde{y}_3 \end{pmatrix}$. Since $\mathbf{x} \cdot \mathbf{y} = \|\mathbf{x}\| \, \|\mathbf{y}\| \cos\theta$ where θ is the inner angle between the vectors, and since $\tilde{\mathbf{x}} \cdot \tilde{\mathbf{y}} = \|\tilde{\mathbf{x}}\| \, \|\tilde{\mathbf{y}}\| \cos\tilde{\theta}$, then we need to show that

$$\theta = \tilde{\theta}$$

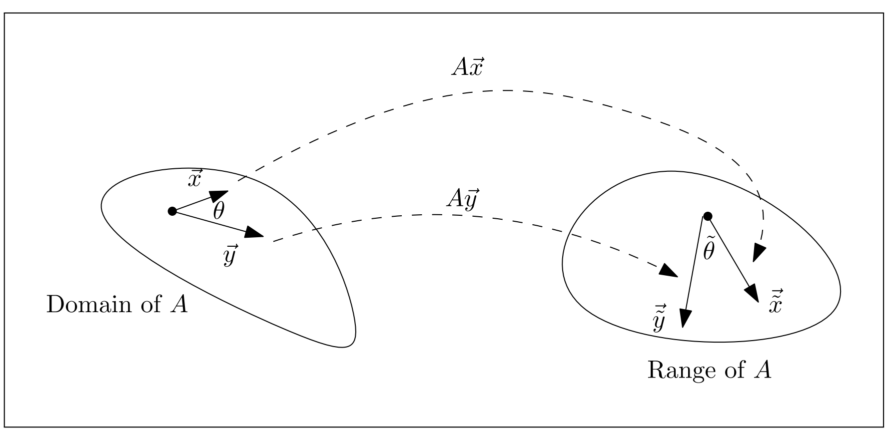


Figure 13: Linear transformation Ax preserves angles. Problem 7

Let

$$A\mathbf{x} = \tilde{\mathbf{x}}$$
$$A\mathbf{y} = \tilde{\mathbf{y}}$$

Using the A given, then

$$\begin{pmatrix} \frac{1}{\sqrt{5}} & 0 & \frac{2}{\sqrt{5}} \\ \frac{-2}{\sqrt{5}} & 0 & \frac{1}{\sqrt{5}} \\ 0 & -1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{5}} x_1 + \frac{2}{\sqrt{5}} x_3 \\ \frac{-2}{\sqrt{5}} x_1 + \frac{1}{\sqrt{5}} x_3 \\ -x_2 \end{pmatrix}$$

Hence

$$\tilde{\mathbf{x}} = \begin{pmatrix} \frac{1}{\sqrt{5}} x_1 + \frac{2}{\sqrt{5}} x_3 \\ \frac{-2}{\sqrt{5}} x_1 + \frac{1}{\sqrt{5}} x_3 \\ -x_2 \end{pmatrix}$$

We see from the above that

$$\tilde{x}_1 = \frac{1}{\sqrt{5}} x_1 + \frac{2}{\sqrt{5}} x_3$$
$$\tilde{x}_2 = \frac{-2}{\sqrt{5}} x_1 + \frac{1}{\sqrt{5}} x_3$$
$$\tilde{x}_3 = -x_2$$

Similarly

$$\tilde{\mathbf{y}} = \begin{pmatrix} \frac{1}{\sqrt{5}} y_1 + \frac{2}{\sqrt{5}} y_3 \\ \frac{1}{\sqrt{5}} y_1 - \frac{2}{\sqrt{5}} y_3 \\ -y_2 \end{pmatrix}$$

We see from the above that

$$\tilde{y}_1 = \frac{1}{\sqrt{5}} y_1 + \frac{2}{\sqrt{5}} y_3$$
$$\tilde{y}_2 = \frac{-2}{\sqrt{5}} y_1 + \frac{1}{\sqrt{5}} y_3$$
$$\tilde{y}_3 = -y_2$$

We now need to determine θ and $\tilde{\theta}$ and show they are the same. From the definition above

$$\theta = \arccos\left(\frac{\mathbf{x} \cdot \mathbf{y}}{\|\mathbf{x}\| \, \|\mathbf{y}\|}\right)$$

But $\mathbf{x} \cdot \mathbf{y} = x_1 y_1 + x_2 y_2 + x_3 y_3$ and $\|\mathbf{x}\| = \sqrt{x_1^2 + x_2^2 + x_3^2}$ and $\|\mathbf{y}\| = \sqrt{y_1^2 + y_2^2 + y_3^2}$, therefore the above becomes

$$\theta = \arccos\left(\frac{x_1 y_1 + x_2 y_2 + x_3 y_3}{\sqrt{x_1^2 + x_2^2 + x_3^2}\sqrt{y_1^2 + y_2^2 + y_3^2}}\right) \tag{1}$$

Similarly, $\tilde{\mathbf{x}} \cdot \tilde{\mathbf{y}} = \tilde{x}_1\tilde{y}_1 + \tilde{x}_2\tilde{y}_2 + \tilde{x}_3\tilde{y}_3$. Using the values of $\tilde{x}_i, \tilde{y}_i$ found above the dot product becomes

$$
\begin{aligned}
\tilde{\mathbf{x}} \cdot \tilde{\mathbf{y}} &= \left(\frac{1}{\sqrt{5}}x_1 + \frac{2}{\sqrt{5}}x_3\right)\left(\frac{1}{\sqrt{5}}y_1 + \frac{2}{\sqrt{5}}y_3\right) + \left(\frac{-2}{\sqrt{5}}x_1 + \frac{1}{\sqrt{5}}x_3\right)\left(\frac{-2}{\sqrt{5}}y_1 + \frac{1}{\sqrt{5}}y_3\right) + (-x_2)(-y_2) \\
&= \frac{1}{5}(x_1 + 2x_3)(y_1 + 2y_3) + \frac{1}{5}(x_3 - 2x_1)(y_3 - 2y_1) + x_2y_2 \\
&= \frac{1}{5}x_1y_1 + \frac{2}{5}x_1y_3 + \frac{2}{5}x_3y_1 + \frac{4}{5}x_3y_3 + \frac{4}{5}x_1y_1 - \frac{2}{5}x_1y_3 - \frac{2}{5}x_3y_1 + \frac{1}{5}x_3y_3 + x_2y_2
\end{aligned}
$$

Which simplifies to

$$\tilde{\mathbf{x}} \cdot \tilde{\mathbf{y}} = x_1y_1 + x_2y_2 + x_3y_3$$

And $\|\tilde{\mathbf{x}}\| = \sqrt{\tilde{x}_1^2 + \tilde{x}_2^2 + \tilde{x}_3^2}$. Using the values of $\tilde{x}_i$ found above, this becomes

$$
\begin{aligned}
\|\tilde{\mathbf{x}}\| &= \sqrt{\left(\frac{1}{\sqrt{5}}x_1 + \frac{2}{\sqrt{5}}x_3\right)^2 + \left(\frac{-2}{\sqrt{5}}x_1 + \frac{1}{\sqrt{5}}x_3\right)^2 + (-x_2)^2} \\
&= \sqrt{\frac{1}{5}x_1^2 + \frac{4}{5}x_1x_3 + \frac{4}{5}x_3^2 + \frac{4}{5}x_1^2 - \frac{4}{5}x_1x_3 + \frac{1}{5}x_3^2 + x_2^2}
\end{aligned}
$$

Which simplifies to

$$\|\tilde{\mathbf{x}}\| = \sqrt{x_1^2 + x_2^2 + x_3^2}$$

Similarly, $\|\tilde{\mathbf{y}}\| = \sqrt{\tilde{y}_1^2 + \tilde{y}_2^2 + \tilde{y}_3^2}$ and using the values of $\tilde{y}_i$ found above, then this becomes

$$
\begin{aligned}
\|\tilde{\mathbf{y}}\| &= \sqrt{\left(\frac{1}{\sqrt{5}}y_1 + \frac{2}{\sqrt{5}}y_3\right)^2 + \left(\frac{-2}{\sqrt{5}}y_1 + \frac{1}{\sqrt{5}}y_3\right)^2 + (-y_2)^2} \\
&= \sqrt{\frac{1}{5}y_1^2 + \frac{4}{5}y_1y_3 + \frac{4}{5}y_3^2 + \frac{4}{5}y_1^2 - \frac{4}{5}y_1y_3 + \frac{1}{5}y_3^2 + y_2^2}
\end{aligned}
$$

Which simplifies to

$$\|\tilde{\mathbf{y}}\| = \sqrt{y_1^2 + y_2^2 + y_3^2}$$

Therefore

$$
\begin{aligned}
\cos\tilde{\theta} &= \frac{\tilde{\mathbf{x}} \cdot \tilde{\mathbf{y}}}{\|\tilde{\mathbf{x}}\|\,\|\tilde{\mathbf{y}}\|} \\
\tilde{\theta} &= \arccos\left(\frac{\sqrt{x_1^2 + x_2^2 + x_3^2}}{\sqrt{x_1^2 + x_2^2 + x_3^2}\sqrt{y_1^2 + y_2^2 + y_3^2}}\right) \qquad (2)
\end{aligned}
$$

Comparing (1) and (2) shows they are the same. Therefore $\theta = \tilde{\theta}$. QED.

8 Problem 8

8. Write the expressions

$$\cos(nx - mx) - \cos(nx + mx), \cos(nx - mx) - \cos(nx + mx),$$

$$\sin(nx - mx) - \sin(nx + mx), \sin(nx - mx) + \sin(nx + mx)$$

with integers n and m in terms of

$$\cos nx, \sin nx, \cos mx, \sin mx.$$

Consider the Hilbert space of real square integrable functions $L^2[-\pi\ \pi]$ on the $[-\pi, \pi]$ interval, equipped with the inner product $< f, g >= \int_{-\pi}^{\pi} fg\, dx$. Show that the set of functions

$$S = \left\{ \frac{1}{\sqrt{2\pi}}, \frac{\cos x}{\sqrt{\pi}}, \frac{\sin x}{\sqrt{\pi}}, \frac{\cos 2x}{\sqrt{\pi}}, \frac{\sin 2x}{\sqrt{\pi}}, \ldots, \frac{\cos 5x}{\sqrt{\pi}}, \frac{\sin 5x}{\sqrt{\pi}} \right\} \subset L^2[-\pi, \pi]$$

is an orthogonal but not an orthonormal set in $L^2[-\pi, \pi]$. How would you change S to make it orthonormal? Use Gram's theorem to compute the best approximating function $H_5(x)$ from the subspace of S-linear combinations *span*(S) of $L^2[-\pi, \pi]$ to the function

$$H(x) = \pi - \mid x \mid \quad \in L^2[-\pi, \pi].$$

Figure 14: Problem 8 Statement

Correction: The set S shown above should be

$$S = \left\{ \frac{1}{2\pi}, \frac{\cos x}{\pi}, \frac{\sin x}{\pi}, \frac{\cos 2x}{\pi}, \frac{\sin 2x}{\pi}, \cdots, \frac{\cos 5x}{\pi}, \frac{\sin x}{\pi} \right\}$$

<u>Solution</u>

Two functions f, g are orthogonal on $[-\pi, \pi]$ if $\int_{-\pi}^{\pi} fgdx = 0$. To show this for the set of functions given, we pick $f = \frac{1}{2\pi}$ and then for g we pick $\frac{\cos mx}{\pi}$ and then $\frac{\sin mx}{\pi}$. i.e.

$$I_1 = \int_{-\pi}^{\pi} \frac{1}{2\pi} \frac{\cos mx}{\pi} dx$$

$$I_2 = \int_{-\pi}^{\pi} \frac{1}{2\pi} \frac{\sin mx}{\pi} dx$$

For the rest, we have to determine the following 3 cases

$$I_3 = \int_{-\pi}^{\pi} \frac{\cos mx}{\pi} \frac{\cos nx}{\pi} dx$$
$$I_4 = \int_{-\pi}^{\pi} \frac{\cos mx}{\pi} \frac{\sin nx}{\pi} dx$$
$$I_5 = \int_{-\pi}^{\pi} \frac{\sin mx}{\pi} \frac{\sin nx}{\pi} dx$$

These will take care of all possible combination of any two function in the set S. We could always replace m, n by a number from $1 \cdots 5$ after evaluating the integrals in order to obtain a specific case. Starting with I_1

$$I_1 = \frac{1}{2\pi^2} \int_{-\pi}^{\pi} \cos mx dx$$

But cos function has period 2π and therefore the integral above is zero $\underline{I_1 = 0}$. This shows that $f = \frac{1}{2\pi}$ is orthogonal with all $\frac{\cos mx}{\pi}$ functions in the set.

$$I_2 = \frac{1}{2\pi^2} \int_{-\pi}^{\pi} \sin mx dx$$

As above, sin function has period of 2π and therefore the integral above is zero $\underline{I_2 = 0}$. This shows that $f = \frac{1}{2\pi}$ is orthogonal with all $\frac{\sin mx}{\pi}$ functions in the set.

$$I_3 = \frac{1}{\pi^2} \int_{-\pi}^{\pi} \cos mx \cos nx dx$$

From tables, using $\underline{\cos A \cos B = \frac{1}{2} (\cos (A - B) + \cos (A + B))}$, then

$$\cos mx \cos nx = \frac{1}{2} (\cos ((m - n) x) + \cos ((m + n) x))$$

And I_3 now becomes

$$\begin{aligned} I_3 &= \frac{1}{2\pi^2} \int_{-\pi}^{\pi} \cos ((m - n) x) + \cos ((m + n) x) dx \\ &= \frac{1}{2\pi^2} \left(\int_{-\pi}^{\pi} \cos (m - n) x dx + \int_{-\pi}^{\pi} \cos ((m + n) x) dx \right) \end{aligned}$$

Since the problem is asking us to show orthogonality of different functions in the set, then we assume $m \neq n$, otherwise the integral will have to be handled as special case when $m = n$ due to the division.

$$I_3 = \frac{1}{2\pi^2} \left(\frac{1}{m - n} [\sin (m - n) x]_{-\pi}^{\pi} + \frac{1}{m + n} [\sin (m + n) x]_{-\pi}^{\pi} \right)$$

But since n, m are integers, then both terms above are zero since $\sin (N\pi) = 0$ for integer N. Hence $\underline{I_3 = 0}$. This shows that $\frac{\cos mx}{\pi}$ is orthogonal with $\frac{\cos nx}{\pi}$ when $m \neq n$.

$$I_4 = \frac{1}{\pi^2} \int_{-\pi}^{\pi} \sin nx \cos mx dx$$

Using $\underline{\sin A \cos B = \frac{1}{2}(\sin(A-B) + \sin(A+B))}$, the above becomes

$$
\begin{aligned}
I_4 &= \frac{1}{2\pi^2} \int_{-\pi}^{\pi} (\sin(n-m)x + \sin(n+m)x)\,dx \\
&= \frac{1}{2\pi^2} \left(\int_{-\pi}^{\pi} \sin(n-m)x dx + \int_{-\pi}^{\pi} \sin(n+m)x dx \right)
\end{aligned}
$$

Again, since $n \neq m$, then the above becomes

$$
\begin{aligned}
I_4 &= \frac{1}{2\pi^2} \left(\frac{-1}{n-m} [\cos(n-m)x]_{-\pi}^{\pi} + \frac{-1}{n+m} [\cos(m+n)x]_{-\pi}^{\pi} \right) \\
&= \frac{1}{2\pi^2} \left(\frac{-1}{n-m} [\cos((n-m)\pi) - \cos((n-m)(-\pi))] + \frac{-1}{n+m} [\cos((n+m)\pi) - \cos((n+m)(-\pi))] \right)
\end{aligned}
$$

But $\cos(-x) = \cos(x)$ and the above becomes

$$
\begin{aligned}
I_4 &= \frac{1}{2\pi^2} \left(\frac{-1}{n-m} [\cos((n-m)\pi) - \cos((n-m)\pi)] + \frac{-1}{n+m} [\cos((n+m)\pi) - \cos((n+m)\pi)] \right) \\
&= \frac{1}{2\pi^2} \left(\frac{-1}{n-m} [0] + \frac{-1}{n+m} [0] \right) \\
&= 0
\end{aligned}
$$

Hence $\underline{I_4 = 0}$. This shows that $\frac{\sin mx}{\pi}$ is orthogonal with $\frac{\sin nx}{\pi}$ when $m \neq n$.

The final integral is

$$
I_5 = \frac{1}{\pi^2} \int_{-\pi}^{\pi} \sin mx \sin nx dx
$$

Using $\sin A \sin B = \frac{1}{2}(\cos(A-B) - \cos(A+B))$ the above becomes

$$
\begin{aligned}
I_5 &= \frac{1}{2\pi^2} \int_{-\pi}^{\pi} \cos((m-n)x) - \cos((m+n)x)\,dx \\
&= \frac{1}{2\pi^2} \left(\int_{-\pi}^{\pi} \cos((m-n)x)\,dx - \int_{-\pi}^{\pi} \cos((m+n)x)\,dx \right)
\end{aligned}
$$

<u>Case $n = m$</u>

$$
\begin{aligned}
I_5 &= \frac{1}{2\pi^2} \left(\int_{-\pi}^{\pi} dx - \int_{-\pi}^{\pi} dx \right) \\
&= 0
\end{aligned}
$$

This shows that $\frac{\sin mx}{\pi}$ is orthogonal with all $\frac{\cos nx}{\pi}$ when $m = n$.

<u>Case $n \neq m$</u>

$$
I_5 = \frac{1}{2\pi^2} \left(\frac{1}{m-n} [\sin((m-n)x)]_{-\pi}^{\pi} - \frac{1}{m+n} [\sin((m+n)x)]_{-\pi}^{\pi} \right)
$$

But since n, m are integers, then both terms above are zero since $\sin(N\pi) = 0$ for integer N. Hence $\underline{I_5 = 0}$. This shows that $\frac{\sin mx}{\pi}$ is orthogonal $\frac{\cos nx}{\pi}$.

The above shows that all the functions in S are pairwise orthogonal.

To make the set S orthonormal, we need to find weight k such that $\|kf(x)\| = 1$ or for functions, this is the same as

$$\sqrt{\int_{-\pi}^{\pi} (kf(x))^2\,dx} = 1$$

For $f = \frac{1}{2\pi}$, this becomes

$$\begin{aligned}
\sqrt{\int_{-\pi}^{\pi} \left(k\frac{1}{2\pi}\right)^2 dx} &= 1\\
\frac{k}{2\pi}\sqrt{\int_{-\pi}^{\pi} dx} &= 1\\
\frac{k}{2\pi}\sqrt{2\pi} &= 1\\
k &= \sqrt{2\pi}
\end{aligned}$$

For $f = \frac{\cos mx}{\pi}$

$$\begin{aligned}
\sqrt{\int_{-\pi}^{\pi} \left(k\frac{\cos mx}{\pi}\right)^2 dx} &= 1\\
\frac{k}{\pi}\sqrt{\int_{-\pi}^{\pi} \cos^2 mx dx} &= 1\\
\frac{k}{\pi}\sqrt{\int_{-\pi}^{\pi} \frac{1}{2} + \frac{1}{2}\cos 2mx dx} &= 1\\
\frac{k}{\pi}\sqrt{\left(\int_{-\pi}^{\pi} \frac{1}{2}dx + \frac{1}{2}\int_{-\pi}^{\pi} \cos 2mx dx\right)} &= 1\\
\frac{k}{\pi}\sqrt{\left(\pi + \frac{1}{2}\overbrace{\left[\frac{\sin(2mx)}{2m}\right]_{-\pi}^{\pi}}^{0}\right)} &= 1\\
\frac{k}{\pi}\sqrt{\pi} &= 1\\
k &= \sqrt{\pi}
\end{aligned}$$

For $f = \frac{\sin mx}{\pi}$

$$
\begin{aligned}
\sqrt{\int_{-\pi}^{\pi}\left(k\frac{\sin mx}{\pi}\right)^{2}dx} &= 1\\
\frac{k}{\pi}\sqrt{\int_{-\pi}^{\pi}\sin^{2}mxdx} &= 1\\
\frac{k}{\pi}\sqrt{\int_{-\pi}^{\pi}\frac{1}{2}-\frac{1}{2}\cos 2mxdx} &= 1\\
\frac{k}{\pi}\sqrt{\left(\int_{-\pi}^{\pi}\frac{1}{2}dx-\frac{1}{2}\int_{-\pi}^{\pi}\cos 2mxdx\right)} &= 1\\
\frac{k}{\pi}\sqrt{\left(\pi-\frac{1}{2}\overbrace{\left[\frac{\sin(2mx)}{2m}\right]_{-\pi}^{\pi}}^{0}\right)} &= 1\\
\frac{k}{\pi}\sqrt{\pi} &= 1\\
k &= \sqrt{\pi}
\end{aligned}
$$

Therefore the orthonormal set now becomes, after using the weights found above as

$$
\begin{aligned}
\tilde{S} &= \left\{\sqrt{2\pi}\frac{1}{2\pi},\sqrt{\pi}\frac{\cos x}{\pi},\sqrt{\pi}\frac{\sin x}{\pi},\sqrt{\pi}\frac{\cos 2x}{\pi},\sqrt{\pi}\frac{\sin 2x}{\pi},\cdots,\sqrt{\pi}\frac{\cos 5x}{\pi},\sqrt{\pi}\frac{\sin x}{\pi}\right\}\\
&= \left\{\frac{1}{\sqrt{2\pi}},\frac{\cos x}{\sqrt{\pi}},\frac{\sin x}{\sqrt{\pi}},\frac{\cos 2x}{\sqrt{\pi}},\frac{\sin 2x}{\sqrt{\pi}},\cdots,\frac{\cos 5x}{\sqrt{\pi}},\frac{\sin x}{\sqrt{\pi}}\right\}
\end{aligned}
$$

We now need to approximate $H(x) = \pi - |x|$ using $\tilde{S}$. The following is a plot of $H(x)$ over $[-\pi, \pi]$

```
f = Pi - Abs[x];
Plot[f, {x, -Pi, Pi}, Frame -> True, GridLines → Automatic,
 GridLinesStyle → LightGray, PlotStyle → Red,
 FrameLabel → {{"H(x)", None}, {"x", "Function to approximate"}},
 BaseStyle → 12,
 FrameTicks → {{Automatic, None}, {{-Pi, -Pi / 2, 0, Pi / 2, Pi}, None}}]
```

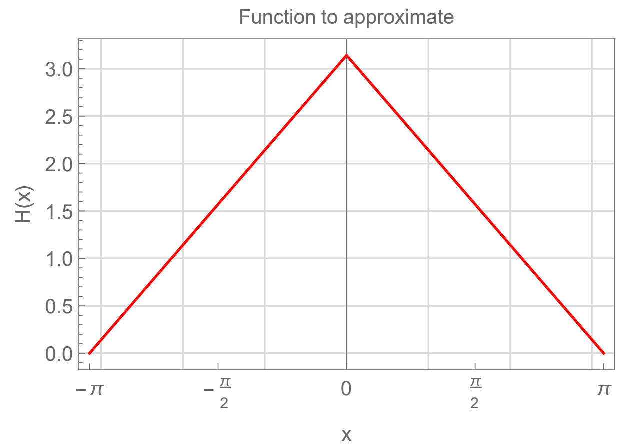


Figure 15: Function $H(x)$ to approximate. Problem 8

Counting the number of functions in $\tilde{S}$, there are 11 of them. Using Gram's theorem, this approximation is

$$H(x) \approx c_1 S_1 + c_2 S_2 + c_3 S_3 + \cdots \tag{1}$$

Where $S_1 = \frac{1}{\sqrt{2\pi}}, S_2 = \frac{\cos x}{\sqrt{\pi}}, \cdots, S_{10} = \frac{\cos 5x}{\sqrt{\pi}}, S_{11} = \frac{\sin x}{\sqrt{\pi}}$. Hence

$$H(x) \approx c_1 \frac{1}{\sqrt{2\pi}} + c_2 \frac{\cos x}{\sqrt{\pi}} + c_3 \frac{\sin x}{\sqrt{\pi}} + c_4 \frac{\cos 2x}{\sqrt{\pi}} + c_5 \frac{\sin 2x}{\sqrt{\pi}} + \cdots + c_{10} \frac{\cos 5x}{\sqrt{\pi}} + c_{11} \frac{\sin 5x}{\sqrt{\pi}}$$

where the constants c_i are found from solving

$$\begin{pmatrix} \langle S_1, S_1 \rangle & \langle S_1, S_2 \rangle & \langle S_1, S_3 \rangle & \cdots & \langle S_1, S_{11} \rangle \\ \langle S_2, S_1 \rangle & \langle S_2, S_2 \rangle & \langle S_2, S_3 \rangle & \cdots & \langle S_2, S_{11} \rangle \\ \langle S_3, S_1 \rangle & \langle S_3, S_2 \rangle & \langle S_3, S_3 \rangle & \cdots & \langle S_3, S_{11} \rangle \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \langle S_{11}, S_1 \rangle & \langle S_{11}, S_2 \rangle & \langle S_{11}, S_3 \rangle & \cdots & \langle S_{11}, S_{11} \rangle \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ \vdots \\ c_{11} \end{pmatrix} = \begin{pmatrix} \langle S_1, H(x) \rangle \\ \langle S_2, H(x) \rangle \\ \langle S_3, H(x) \rangle \\ \vdots \\ \langle S_{11}, H(x) \rangle \end{pmatrix}$$

But since $\left\langle S_i, S_j \right\rangle = 0$ for $i \neq j$, because we showed above they are orthogonal to each others, and since S_i are all normalized now, then $\langle S_i, S_i \rangle = \|S_i\|^2 = 1$. Hence the above reduces to

$$\begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ \vdots \\ c_{11} \end{pmatrix} = \begin{pmatrix} \langle S_1, H(x) \rangle \\ \langle S_2, H(x) \rangle \\ \langle S_3, H(x) \rangle \\ \vdots \\ \langle S_{11}, H(x) \rangle \end{pmatrix} \tag{2}$$

So we just need to evaluate $\langle S_i, H(x) \rangle$. But we need to do this only for three cases. These are $\left\langle \frac{1}{\sqrt{2\pi}}, H(x) \right\rangle, \left\langle \frac{\cos mx}{\sqrt{\pi}}, H(x) \right\rangle, \left\langle \frac{\sin mx}{\sqrt{\pi}}, H(x) \right\rangle$ and then set $m = 1 \cdots 5$.

$$\begin{aligned}
\left\langle \frac{1}{\sqrt{2\pi}}, H(x) \right\rangle &= \int_{-\pi}^{\pi} \frac{1}{\sqrt{2\pi}} H(x)\, dx \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} H(x)\, dx \\
&= \frac{1}{\sqrt{2\pi}} \left(\int_{-\pi}^{0} (\pi + x)\, dx + \int_{0}^{\pi} (\pi - x)\, dx \right) \\
&= \frac{1}{\sqrt{2\pi}} \left(\left[\pi x + \frac{x^2}{2} \right]_{-\pi}^{0} + \left[\pi x - \frac{x^2}{2} \right]_{0}^{\pi} \right) \\
&= \frac{1}{\sqrt{2\pi}} \left(\left[0 - \left(-\pi^2 + \frac{\pi^2}{2} \right) \right] + \left[\pi^2 - \frac{\pi^2}{2} \right] \right) \\
&= \frac{1}{\sqrt{2\pi}} \left(\left[\frac{\pi^2}{2} \right] + \left[\pi^2 - \frac{\pi^2}{2} \right] \right) \\
&= \frac{1}{\sqrt{2\pi}} \pi^2 \\
&= \frac{\pi^{\frac{3}{2}}}{\sqrt{2}}
\end{aligned}$$

And

$$\begin{aligned}
\left\langle \frac{\cos mx}{\sqrt{\pi}}, H(x) \right\rangle &= \int_{-\pi}^{\pi} \frac{\cos mx}{\sqrt{\pi}} H(x)\, dx \\
&= \frac{1}{\sqrt{\pi}} \left(\int_{-\pi}^{0} (\pi + x) \cos mx dx + \int_{0}^{\pi} (\pi - x) \cos mx dx \right) \\
&= \frac{1}{\sqrt{\pi}} \left(\int_{-\pi}^{0} \pi \cos mx + \int_{-\pi}^{0} x \cos mx dx + \int_{0}^{\pi} \pi \cos mx dx - \int_{-\pi}^{0} x \cos mx dx \right) \\
&= \frac{1}{\sqrt{\pi}} \left(\int_{-\pi}^{\pi} \pi \cos mx + \int_{-\pi}^{0} x \cos mx dx - \int_{-\pi}^{0} x \cos mx dx \right) \qquad (3)
\end{aligned}$$

$\int x \cos mx dx$ can be evaluated by integration by parts. Let $u = x, dv = \cos mx \rightarrow du = 1, v =$

$\frac{\sin mx}{m}$ hence

$$\begin{aligned}\int_{-\pi}^{0} x\cos mx dx &= \left[x\frac{\sin mx}{m}\right]_{-\pi}^{0} - \int_{-\pi}^{0}\frac{\sin mx}{m}dx \\ &= 0 - \frac{1}{m}\int_{-\pi}^{0}\sin mx dx \\ &= -\frac{1}{m}\left(-\frac{\cos mx}{m}\right)_{-\pi}^{0} \\ &= \frac{1}{m^2}\left(1-\cos m\pi\right) \qquad (4)\end{aligned}$$

And

$$\begin{aligned}\int_{0}^{\pi} x\cos mx dx &= \left[x\frac{\sin mx}{m}\right]_{0}^{\pi} - \int_{0}^{\pi}\frac{\sin mx}{m}dx \\ &= 0 - \frac{1}{m}\int_{0}^{\pi}\sin mx dx \\ &= -\frac{1}{m}\left(-\frac{\cos mx}{m}\right)_{0}^{\pi} \\ &= \frac{1}{m^2}\left(\cos m\pi - 1\right) \qquad (5)\end{aligned}$$

And $\int_{-\pi}^{\pi}\pi\cos mx = \pi\int_{-\pi}^{\pi}\pi\cos mx = 0$. Using (4,5) in (3), then

$$\begin{aligned}\left\langle\frac{\cos mx}{\sqrt{\pi}}, H(x)\right\rangle &= \frac{1}{\sqrt{\pi}}\left(\frac{1}{m^2}\left(1-\cos m\pi\right) - \frac{1}{m^2}\left(\cos m\pi - 1\right)\right) \\ &= \frac{1}{m^2\sqrt{\pi}}\left(1-\cos m\pi - \cos m\pi + 1\right) \\ &= \frac{2\left(1-\cos m\pi\right)}{m^2\sqrt{\pi}}\end{aligned}$$

Hence

$$\begin{aligned}\left\langle\frac{\cos mx}{\sqrt{\pi}}, H(x)\right\rangle_{m=1} &= \left\langle\frac{\cos x}{\sqrt{\pi}}, H(x)\right\rangle = \frac{2\left(1-\cos\pi\right)}{\sqrt{\pi}} = \frac{2\left(1+1\right)}{\sqrt{\pi}} = \frac{4}{\sqrt{\pi}} \\ \left\langle\frac{\cos mx}{\sqrt{\pi}}, H(x)\right\rangle_{m=2} &= \left\langle\frac{\cos 2x}{\sqrt{\pi}}, H(x)\right\rangle = \frac{2\left(1-\cos 2\pi\right)}{4\sqrt{\pi}} = 0 \\ \left\langle\frac{\cos mx}{\sqrt{\pi}}, H(x)\right\rangle_{m=3} &= \left\langle\frac{\cos 3x}{\sqrt{\pi}}, H(x)\right\rangle = \frac{2\left(1-\cos 3\pi\right)}{9\sqrt{\pi}} = \frac{2\left(1+1\right)}{9\sqrt{\pi}} = \frac{4}{9\sqrt{\pi}} \\ \left\langle\frac{\cos mx}{\sqrt{\pi}}, H(x)\right\rangle_{m=4} &= \left\langle\frac{\cos 4x}{\sqrt{\pi}}, H(x)\right\rangle = \frac{2\left(1-\cos 4\pi\right)}{16\sqrt{\pi}} = 0 \\ \left\langle\frac{\cos mx}{\sqrt{\pi}}, H(x)\right\rangle_{m=5} &= \left\langle\frac{\cos 5x}{\sqrt{\pi}}, H(x)\right\rangle = \frac{2\left(1-\cos 5\pi\right)}{25\sqrt{\pi}} = \frac{4}{25\sqrt{\pi}}\end{aligned}$$

Similarly (we expect all the following integrals to be zero, this is because we see from above that $H(x)$ is an even function and sin is odd, hence the product is an odd function and the integral is over the period). This is the same as when in doing Fourier series expansion (which is what we are doing here essentially but using Gram's theorem instead), all the b_n terms will be zero when the function being approximated is even and all the a_n terms will be zero when the function being approximation is odd.

But we will go ahead and do the integrals to show that this is indeed the case.

$$\begin{aligned}
\left\langle \frac{\sin mx}{\sqrt{\pi}}, H(x) \right\rangle &= \int_{-\pi}^{\pi} \frac{\sin mx}{\sqrt{\pi}} H(x)\, dx \\
&= \frac{1}{\sqrt{\pi}} \left(\int_{-\pi}^{0} (\pi + x) \sin mx dx + \int_{0}^{\pi} (\pi - x) \sin mx dx \right) \\
&= \frac{1}{\sqrt{\pi}} \left(\int_{-\pi}^{0} \pi \sin mx + \int_{-\pi}^{0} x \sin mx dx + \int_{0}^{\pi} \pi \sin mx dx - \int_{-\pi}^{0} x \sin mx dx \right) \\
&= \frac{1}{\sqrt{\pi}} \left(\int_{-\pi}^{\pi} \pi \sin mx + \int_{-\pi}^{0} x \sin mx dx - \int_{-\pi}^{0} x \sin mx dx \right) \qquad (6)
\end{aligned}$$

$\int x \sin mx dx$ is evaluated by integration by parts. Let $u = x, dv = \sin mx \longrightarrow du = 1, v = \frac{-\cos mx}{m}$ hence

$$\begin{aligned}
\int_{-\pi}^{0} x \sin mx dx &= -\frac{1}{m} [x \cos mx]_{-\pi}^{0} - \int_{-\pi}^{0} \frac{-\cos mx}{m} dx \\
&= -\frac{1}{m} [0 - (-\pi \cos m\pi)] + \frac{1}{m} \int_{-\pi}^{0} \cos mx dx \\
&= -\frac{\pi}{m} [\cos m\pi] + \frac{1}{m} \left(\frac{\sin mx}{m} \right)_{-\pi}^{0} \\
&= -\frac{\pi}{m} [\cos m\pi] \qquad (7)
\end{aligned}$$

And

$$\begin{aligned}
\int_{0}^{\pi} x \sin mx dx &= -\frac{1}{m} [x \cos mx]_{0}^{\pi} - \int_{0}^{\pi} \frac{-\cos mx}{m} dx \\
&= -\frac{1}{m} [\pi \cos m\pi] + \frac{1}{m} \int_{0}^{\pi} \cos mx dx \\
&= -\frac{\pi}{m} [\cos m\pi] \qquad (8)
\end{aligned}$$

And $\int_{-\pi}^{\pi} \pi \sin mx = 0$. Using (7,8) in (6), then

$$\begin{aligned}
\left\langle \frac{\sin mx}{\sqrt{\pi}}, H(x) \right\rangle &= \frac{1}{\sqrt{\pi}} \left(-\frac{\pi}{m} [\cos m\pi] + \frac{\pi}{m} [\cos m\pi] \right) \\
&= 0
\end{aligned}$$

Hence as expected all the inner products now are zero

$$\left\langle \frac{\sin mx}{\sqrt{\pi}}, H(x) \right\rangle_m = 0 \qquad m = 1, 2, 3, 4, 5$$

Using all the above results in (2) gives

$$\begin{pmatrix} 1&0&0&0&0&0&0&0&0&0&0\\ 0&1&0&0&0&0&0&0&0&0&0\\ 0&0&1&0&0&0&0&0&0&0&0\\ 0&0&0&1&0&0&0&0&0&0&0\\ 0&0&0&0&1&0&0&0&0&0&0\\ 0&0&0&0&0&1&0&0&0&0&0\\ 0&0&0&0&0&0&1&0&0&0&0\\ 0&0&0&0&0&0&0&1&0&0&0\\ 0&0&0&0&0&0&0&0&1&0&0\\ 0&0&0&0&0&0&0&0&0&1&0\\ 0&0&0&0&0&0&0&0&0&0&1 \end{pmatrix} \begin{pmatrix} c_1\\ c_2\\ c_3\\ c_4\\ c_5\\ c_6\\ c_7\\ c_8\\ c_9\\ c_{10}\\ c_{11} \end{pmatrix} = \begin{pmatrix} \left\langle \frac{1}{\sqrt{2\pi}}, H(x) \right\rangle\\ \left\langle \frac{\cos x}{\sqrt{\pi}}, H(x) \right\rangle\\ \left\langle \frac{\sin x}{\sqrt{\pi}}, H(x) \right\rangle\\ \left\langle \frac{\cos 2x}{\sqrt{\pi}}, H(x) \right\rangle\\ \left\langle \frac{\sin 2x}{\sqrt{\pi}}, H(x) \right\rangle\\ \left\langle \frac{\cos 3x}{\sqrt{\pi}}, H(x) \right\rangle\\ \left\langle \frac{\sin 3x}{\sqrt{\pi}}, H(x) \right\rangle\\ \left\langle \frac{\cos 4x}{\sqrt{\pi}}, H(x) \right\rangle\\ \left\langle \frac{\sin 4x}{\sqrt{\pi}}, H(x) \right\rangle\\ \left\langle \frac{\cos 5x}{\sqrt{\pi}}, H(x) \right\rangle\\ \left\langle \frac{\sin 5x}{\sqrt{\pi}}, H(x) \right\rangle \end{pmatrix}$$

Using the results found above, the above becomes

$$\begin{pmatrix} 1&0&0&0&0&0&0&0&0&0&0\\ 0&1&0&0&0&0&0&0&0&0&0\\ 0&0&1&0&0&0&0&0&0&0&0\\ 0&0&0&1&0&0&0&0&0&0&0\\ 0&0&0&0&1&0&0&0&0&0&0\\ 0&0&0&0&0&1&0&0&0&0&0\\ 0&0&0&0&0&0&1&0&0&0&0\\ 0&0&0&0&0&0&0&1&0&0&0\\ 0&0&0&0&0&0&0&0&1&0&0\\ 0&0&0&0&0&0&0&0&0&1&0\\ 0&0&0&0&0&0&0&0&0&0&1 \end{pmatrix} \begin{pmatrix} c_1\\ c_2\\ c_3\\ c_4\\ c_5\\ c_6\\ c_7\\ c_8\\ c_9\\ c_{10}\\ c_{11} \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}}\pi^{\frac{3}{2}}\\ \frac{4}{\sqrt{\pi}}\\ 0\\ 0\\ 0\\ \frac{4}{9\sqrt{\pi}}\\ 0\\ 0\\ 0\\ \frac{4}{25\sqrt{\pi}}\\ 0 \end{pmatrix}$$

Therefore we see that

$$\begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \\ c_5 \\ c_6 \\ c_7 \\ c_8 \\ c_9 \\ c_{10} \\ c_{11} \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}}\pi^{\frac{3}{2}} \\ \frac{4}{\sqrt{\pi}} \\ 0 \\ 0 \\ 0 \\ \frac{4}{9\sqrt{\pi}} \\ 0 \\ 0 \\ 0 \\ \frac{4}{25\sqrt{\pi}} \\ 0 \end{pmatrix}$$

The above shows that $c_1 = \frac{1}{\sqrt{2}}\pi^{\frac{3}{2}}, c_2 = \frac{4}{\sqrt{\pi}}, c_6 = \frac{4}{9\sqrt{\pi}}, c_{10} = \frac{4}{25\sqrt{\pi}}$ and all other $c's$ are zero. Therefore the best approximation is

$$\begin{aligned} H(x) &\approx c_1\frac{1}{\sqrt{2\pi}} + c_2\frac{\cos x}{\sqrt{\pi}} + c_3\frac{\sin x}{\sqrt{\pi}} + c_4\frac{\cos 2x}{\sqrt{\pi}} + c_5\frac{\sin 2x}{\sqrt{\pi}} + \cdots + c_{10}\frac{\cos 5x}{\sqrt{\pi}} + c_{11}\frac{\sin 5x}{\sqrt{\pi}} \\ &= \frac{1}{\sqrt{2}}\pi^{\frac{3}{2}}\frac{1}{\sqrt{2\pi}} + \frac{4}{\sqrt{\pi}}\frac{\cos x}{\sqrt{\pi}} + \frac{4}{9\sqrt{\pi}}\frac{\cos 3x}{\sqrt{\pi}} + \frac{4}{25\sqrt{\pi}}\frac{\cos 5x}{\sqrt{\pi}} \\ &= \frac{1}{2}\pi + \frac{4}{\pi}\cos x + \frac{4}{9\pi}\cos 3x + \frac{4}{25\pi}\cos 5x \end{aligned}$$

Or

$$\boxed{H(x) \approx \frac{1}{2}\pi + \frac{4}{\pi}\cos x + \frac{4}{9\pi}\cos 3x + \frac{4}{25\pi}\cos 5x}$$

To verify the approximation, the above was plotted against the original $H(x)$, first using one term $H_1(x) \approx \frac{1}{2}\pi$ then using 2 terms $H_2(x) \approx \frac{1}{2}\pi + \frac{4}{\pi}\cos x$ then using 3 terms $H_3(x) \approx \frac{1}{2}\pi + \frac{4}{\pi}\cos x + \frac{4}{9\pi}\cos 3x$ and then using all terms $H_4(x) \approx \frac{1}{2}\pi + \frac{4}{\pi}\cos x + \frac{4}{9\pi}\cos 3x + \frac{4}{25\pi}\cos 5x$. The plot below shows that the approximation improved as more terms added giving the best approximation when all terms are added as expected.

```
ClearAll[x, n];
f = Pi - Abs[x];
approx = {1/2 π, 4/π Cos[x], 4/(9 π) Cos[3 x], 4/(25 π) Cos[5 x]};
data = Table[
   Plot[{f, Total[approx[[1 ;; n]]]}, {x, -Pi, Pi}, Frame -> True,
    GridLines → Automatic, GridLinesStyle → LightGray,
    PlotStyle → {Red, Blue}, FrameLabel → {{"H(x)", None}, {"x", Total[approx[[1 ;; n]]]}},
    BaseStyle → 12,
    FrameTicks → {{Automatic, None}, {{-Pi, -Pi / 2, 0, Pi / 2, Pi}, None}},
    ImageSize → 300],
   {n, 1, 4}
  ];
Grid[Partition[data, 2]]
```

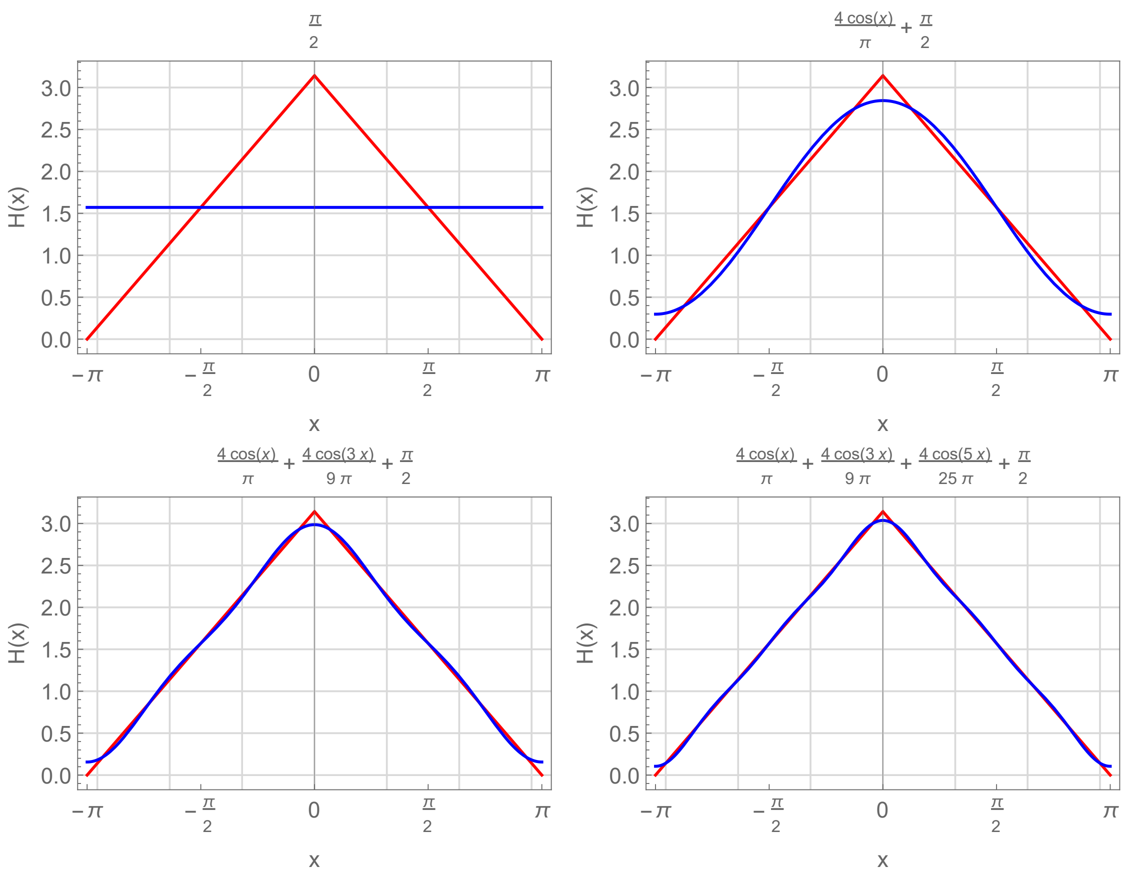


Figure 16: $H(x)$ approximation final resul. Problem 8t