18.700 Problem Set 5 solutions

1. (6 points) Suppose we are given

three distinct elements \(x_1, x_2, \text{ and } x_3\) in \(F\); \hspace{1cm} (1)

and

three arbitrary elements \(a, b, \text{ and } c\) in \(F\); \hspace{1cm} (2)

The problem is to find all polynomials

\[ p(x) = u_0 + u_1 x + u_2 x^2 + u_3 x^3 \] \hspace{1cm} (3)

of degree less than or equal to three satisfying the conditions

\[ p(x_1) = a, \quad p(x_2) = b, \quad p(x_3) = c. \] \hspace{1cm} (4)

a) The conditions (4) on \(p\) can be written as a system of three simultaneous linear equations in four unknowns. Write the augmented matrix of this system of equations.

The equations for the four unknowns \(u_0, u_1, u_2, \text{ and } u_3\) are

\[
\begin{align*}
&u_0 + u_1 x_1 + u_2 x_1^2 + u_3 x_1^3 = a \\
&u_0 + u_1 x_2 + u_2 x_2^2 + u_3 x_2^3 = b \\
&u_0 + u_1 x_3 + u_2 x_3^2 + u_3 x_3^3 = c 
\end{align*}
\]

for which the augmented matrix is

\[
\begin{pmatrix}
1 & x_1 & x_1^2 & x_1^3 & | & a \\
1 & x_2 & x_2^2 & x_2^3 & | & b - a \\
1 & x_3 & x_3^2 & x_3^3 & | & c - a 
\end{pmatrix}
\]

b) Perform elementary row operations to bring this augmented matrix to reduced row-echelon form. (This is a bit disconcerting, because some of the entries of the matrix are not “numbers” like 7, but symbols for numbers, like \(x_2\). You know from algebra how to add, subtract, and multiply such symbols. What requires care is dividing: before you divide by something like \(x_2\), you need to explain why it is not zero, or else worry separately about the case when it is zero. But you should be able to manage.

The first pivot is going to be in the upper left corner, where fortunately the entry is 1. So we begin by subtracting the first row from the second and from the third to clear the first column:

\[
\begin{pmatrix}
1 & x_1 & x_1^2 & x_1^3 & | & a \\
0 & x_2 - x_1 & x_2^2 - x_1^2 & x_2^3 - x_1^3 & | & b - a \\
0 & x_3 - x_1 & x_3^2 - x_1^2 & x_3^3 - x_1^3 & | & c - a 
\end{pmatrix}
\]

The second pivot will be in the second entry of the second row. We first divide that row by its leading entry \(x_2 - x_1\), which is legal since \(x_2 \neq x_1\) by (1):

\[
\begin{pmatrix}
1 & x_1 & x_1^2 & x_1^3 & | & a \\
0 & 1 & x_2 + x_1 & x_2^2 + x_1 x_2 + x_1^2 & | & \frac{b - a}{x_2 - x_1} \\
0 & x_3 - x_1 & x_3^2 - x_1^2 & x_3^3 - x_1^3 & | & c - a 
\end{pmatrix}
\]

Even though it’s not exactly what the algorithm in the notes calls for, the algebra will be simplified a little if at this point we divide the third row by \(x_3 - x_1\), which is nonzero by (1):

\[
\begin{pmatrix}
1 & x_1 & x_1^2 & x_1^3 & | & a \\
0 & 1 & x_2 + x_1 & x_2^2 + x_1 x_2 + x_1^2 & | & \frac{b - a}{x_2 - x_1} \\
0 & 1 & x_3 + x_1 & x_3^2 + x_1 x_3 + x_1^2 & | & \frac{b - a}{x_3 - x_1} 
\end{pmatrix}
\]
Now subtract \( x_1 \) times the second row from the first, and the second row from the third, to clear the other entries in the second column:

\[
\begin{pmatrix}
1 & 0 & -x_1x_2 & -x_1x_2^2 - x_1^2x_2 & \frac{ax_1 - bx_2}{x_2 - x_1} \\
0 & 1 & x_1 & x_1^2 + x_1x_2 + x_2^2 & \frac{b - a}{x_2 - x_1} \\
0 & 0 & x_3 - x_2 & (x_3 - x_2)(x_1 + x_2 + x_3) & \frac{(c-a)(x_2-x_3) - (b-a)(x_3-x_1)}{(x_2-x_1)(x_3-x_1)}
\end{pmatrix}
\]

The third pivot will be the third entry in the third row, which is nonzero by (1). Dividing by it gives

\[
\begin{pmatrix}
1 & 0 & -x_1x_2 & -x_1x_2^2 - x_1^2x_2 & \frac{ax_1 - bx_2}{x_2 - x_1} \\
0 & 1 & x_1 & x_1^2 + x_1x_2 + x_2^2 & \frac{b - a}{x_2 - x_1} \\
0 & 0 & 1 & x_1 + x_2 + x_3 & \frac{(c-a)(x_2-x_3) - (b-a)(x_3-x_1)}{(x_2-x_1)(x_3-x_1)}
\end{pmatrix}
\]

Now add \( x_1x_2 \) times the third row to the first, and subtract \( x_1 + x_2 \) times the third row from the second, to clear the other entries in the third column:

\[
\begin{pmatrix}
1 & 0 & x_1x_2x_3 & \frac{a(x_2x_2^2 - x_2^2x_3) + b(x_2^2x_3 - x_1x_2^2) + c(x_1x_2^2 - x_1^2x_2)}{(x_1 - x_2)(x_2 - x_3)(x_3 - x_1)} \\
0 & 1 & -x_1x_2 - x_1x_3 - x_2x_3 & \frac{a(-x_1^2 + x_2^2) + b(-x_1^2 + x_2^2) + c(x_1^2 - x_2^2)}{(x_1 - x_2)(x_2 - x_3)(x_3 - x_1)} \\
0 & 0 & x_1 + x_2 + x_3 & \frac{a(x_3 - x_2) + b(x_1 - x_3) + c(x_2 - x_1)}{(x_1 - x_2)(x_2 - x_3)(x_3 - x_1)}
\end{pmatrix}
\]

c) Write all the polynomials of degree less than or equal to three satisfying the condition (4).

This means solving the system, which is equivalent to solving the system after row reduction. The free variable is \( u_3 \), which means that we can choose any value we like for \( u_3 \) (say \( z \)) and then solve equations like

\[
u_0 + x_1x_2x_3 \cdot z = \frac{a(x_2x_2^2 - x_2^2x_3) + b(x_2^2x_3 - x_1x_2^2) + c(x_1x_2^2 - x_1^2x_2)}{(x_1 - x_2)(x_2 - x_3)(x_3 - x_1)}
\]

Of course the solution is

\[
u_0 = -x_1x_2x_3 \cdot z + \frac{a(x_2x_2^2 - x_2^2x_3) + b(x_2^2x_3 - x_1x_2^2) + c(x_1x_2^2 - x_1^2x_2)}{(x_1 - x_2)(x_2 - x_3)(x_3 - x_1)}
\]

\[
u_1 = (x_1x_2 + x_1x_3 + x_2x_3) \cdot z + \frac{a(-x_1^2 + x_2^2) + b(-x_1^2 + x_2^2) + c(x_1^2 - x_2^2)}{(x_1 - x_2)(x_2 - x_3)(x_3 - x_1)}
\]

\[
u_2 = -(x_1 + x_2 + x_3) \cdot z + \frac{a(x_3 - x_2) + b(x_1 - x_3) + c(x_2 - x_1)}{(x_1 - x_2)(x_2 - x_3)(x_3 - x_1)}
\]

One way to simplify all the algebra is to write finally

\[
a \frac{(x - x_2)(x - x_3)}{(x_1 - x_2)(x_1 - x_3)} + b \frac{(x - x_1)(x - x_3)}{(x_2 - x_1)(x_2 - x_3)} + c \frac{(x - x_1)(x - x_2)}{(x_3 - x_1)(x_3 - x_2)} + z(x - x_1)(x - x_2)(x - x_3).
\]

Notice that the quadratic polynomial \( \frac{(x-x_2)(x-x_3)}{(x_1-x_2)(x_1-x_3)} \) vanishes at \( x_2 \) and \( x_3 \), and takes the value 1 at \( x_1 \). So the first three terms are the (unique) quadratic polynomial taking the desired values at \( x_1, x_2, \) and \( x_3 \). The last term represents any cubic polynomial that vanishes at \( x_1, x_2, \) and \( x_3 \).

That is, you could have written down an answer to (c) by pure thinking without all the nasty calculation in (b). The point of the problem was to understand row reduction a bit better.
2. (8 points) This problem is about the sequence of integers defined by
\[ a_0 = 0, \quad a_1 = 1, \quad a_{n+1} = a_n + 2a_{n-1} \quad (n \geq 1). \]
(This is like the formula for Fibonacci numbers, but more complicated.) Explicitly,
\[
\begin{align*}
a_0 &= 0, & a_1 &= 1, & a_2 &= 1 + 2 \cdot 0 = 1, \\
a_3 &= 1 + 2 \cdot 1 = 3, & a_4 &= 3 + 2 \cdot 1 = 5, & a_5 &= 5 + 2 \cdot 3 = 11, & \ldots
\end{align*}
\]
This sequence is related to the matrix \( A = \begin{pmatrix} 1 & 2 \\ 1 & 0 \end{pmatrix} \): the defining condition can be written
\[
\begin{pmatrix} a_{n+1} \\ a_n \end{pmatrix} = A \begin{pmatrix} a_n \\ a_{n-1} \end{pmatrix}.
\]

a) Find all eigenvalues and eigenvectors of \( A \).

The procedure from the text or class is to start with any nonzero vector \( v_0 \), and to apply \( A \) repeatedly until linear dependence appears. Starting with the standard basis vector \( v_0 = e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \) gives
\[
v_1 = e_1 + e_2, \quad v_2 = 3e_1 + e_2 = 2v_0 + v_1.
\]
This last dependence relation defines a polynomial
\[
p(x) = x^2 - x - 2 = (x - 2)(x + 1).
\]
The eigenvalues are the roots, which are 2 and -1. To find corresponding eigenvectors, we need nonzero solutions to the systems
\[
(A - 2I)u = 0, \quad (A + I)v = 0.
\]
You should know how to solve these: solutions are
\[
u = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \quad \text{(eigenvector for 2)},
\]
\[
v = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \quad \text{(eigenvector for -1)}.
\]

b) Write \( \begin{pmatrix} 1 \\ 0 \end{pmatrix} \) as a linear combination of eigenvectors of \( A \).

This system of two equations can be solved almost by inspection: \[
\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{1}{3} (u + v)
\]

c) Write a formula for \( A^n \begin{pmatrix} 1 \\ 0 \end{pmatrix} \) not using the matrix \( A \). (A good answer is a vector of formulas depending on \( n \): something like \( \begin{pmatrix} 2n + 1 \\ n^2 \end{pmatrix} \).)

Since \( A^n \) sends \( u \) to \( 2^n \cdot u \), and \( v \) to \( (-1)^n \cdot v \), we get
\[
A^n \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{1}{3} (A^n u + A^n v) = \begin{pmatrix} 2^{n+1} + (-1)^n \\ 2^n - (-1)^n \end{pmatrix}.
\]

d) Write a formula for \( a_n \) as a function of \( n \).

We know that \( \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} a_1 \\ a_0 \end{pmatrix} \), and that \( A \begin{pmatrix} a_n \\ a_{n-1} \end{pmatrix} = \begin{pmatrix} a_{n+1} \\ a_n \end{pmatrix} \). It follows by induction on \( n \) that
\[
A^n \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} a_{n+1} \\ a_n \end{pmatrix}.
\]
Part (c) therefore shows that \( a_n = \frac{2^n - (-1)^n}{3} \) (which is a perfect answer). It follows that \( a_n \) is the integer closest to \( 2^n/3 \).
3. (6 points) This problem is a generalization of $(5.30)$ in the text, and can be proved in a similar way.

Suppose $F$ is any field, and suppose $n = p + q$ with $p$ and $q$ positive integers. Suppose that $A$ is a $p \times p$ matrix, $B$ is a $p \times q$ matrix, and $D$ is a $q \times q$ matrix. The $n \times n$ matrix

$$T = \begin{pmatrix} A & B \\ 0 & D \end{pmatrix}$$

is called block upper triangular; here $0$ means the $q \times p$ matrix of zeros.

**a) Show that $T$ is invertible if and only if both $A$ and $D$ are invertible.**

Suppose first that $T$ is invertible. This means that $T$ is surjective, and therefore that the set of columns of $T$ is a basis of $F^n$. In particular, the first $p$ columns must be linearly independent. These columns are

$$\begin{pmatrix} a_1 \\ 0_q \\ \vdots \\ a_p \\ 0_q \end{pmatrix},$$

with $a_j$ the $j$th $(p \times 1)$ column of $A$, and $0_q$ the $(q \times 1)$ zero vector. It’s clear by inspection that these columns are linearly independent in $F^n$ if and only if $(a_1, \ldots, a_p)$ is linearly independent in $F^p$. The conclusion is that the columns of $A$ are linearly independent, and therefore that $A$ is invertible. Exactly the same argument (applied to the last $q$ rows of $T$) shows that $D$ must also be invertible.

Now suppose that $A$ and $D$ are invertible. We will show that the $n$ columns of $T$ form a linearly independent list. Suppose not; that is, that some column of $T$ is a linear combination of earlier columns. It cannot be one of the first $p$ columns, because we already saw that these are linearly independent (since $A$ is invertible). So it must be one of the last $q$ columns

$$\begin{pmatrix} b_1 \\ d_1 \\ \vdots \\ b_q \\ d_q \end{pmatrix}.$$ 

This means that we have an equation (for some $k$ between $1$ and $q$)

$$\begin{pmatrix} \cdot \\ + \\ \cdot \end{pmatrix} = x_1 \begin{pmatrix} a_1 \\ 0_q \end{pmatrix} + \cdots + x_p \begin{pmatrix} a_p \\ 0_q \end{pmatrix} + y_1 \begin{pmatrix} b_1 \\ d_1 \end{pmatrix} + \cdots + y_{k-1} \begin{pmatrix} b_{k-1} \\ d_{k-1} \end{pmatrix}.$$

The last $q$ coordinates of this equation are $d_k = y_1d_1 + \cdots + y_{k-1}d_{k-1}$, but this contradicts the linear independence of the columns of $D$. The contradiction shows that the columns of $T$ are linearly independent, as we wished to show.

**b) Show that $\lambda$ is an eigenvalue of $T$ if and only if either $\lambda$ is an eigenvalue of $A$ or $\lambda$ is an eigenvalue of $D$ (or both).**

We know that $\lambda$ is an eigenvalue of $T$ if and only if $T - \lambda I_n$ fails to be invertible. The matrix $T - \lambda I_n$ is still block upper-triangular, with diagonal blocks $A - \lambda I_p$ and $D - \lambda I_q$. According to part (a), $T - \lambda I_n$ fails to be invertible if and only if at least one of $A - \lambda I_p$ and $D - \lambda I_q$ fails to be invertible; that is, if and only if $\lambda$ is an eigenvalue of $A$ or of $D$ (or both).

**c) Give an example of a $4 \times 4$ real matrix $T$ so that $T$ has no (real) eigenvalues.**

In class I discussed the $2 \times 2$ real (rotation) matrix $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, proving that it has no real eigenvalues. (The complex eigenvalues are $+i$ and $-i$.) According to (b), we can take for $T$ any $4 \times 4$ block upper-triangular matrix with two diagonal blocks equal to $A$, such as

$$T = \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$ 

(The complex eigenvalues of $T$ are $\pm i$, each with a two-dimensional eigenspace.)