MODEL ANSWERS TO THE FIFTH HOMEWORK

1. The sequence g_n satisfies the recursion $g_n = g_{n-2} + 2g_{n-1}$. Thus the previous two elements determine the next element. Let

$$w_n = \begin{pmatrix} g_{n-1} \\ g_n \end{pmatrix}.$$

Then

$$w_{n+1} = \begin{pmatrix} g_n \\ g_{n+1} \end{pmatrix} = \begin{pmatrix} g_n \\ g_{n-1} + 2g_n \end{pmatrix}.$$

The last expression is Bw_n, where

$$B = \begin{pmatrix} 0 & 1 \\ 1 & 2 \end{pmatrix}.$$

In other words $w_n = B^{n-1}w_1$, where

$$w_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

To compute powers of B we diagonalise B. We first compute the eigenvalues of B, which are roots of the characteristic polynomial.

$$B - \lambda I_2 = \begin{pmatrix} -\lambda & 1 \\ 1 & 2-\lambda \end{pmatrix}.$$

The characteristic polynomial is

$$(-\lambda)(2-\lambda) - 1 = 0.$$

Rearranging gives

$$\lambda^2 - 2\lambda - 1 = 0.$$

Using the quadratic formula, we get

$$\lambda = \frac{2 \pm \sqrt{4+4}}{2} = 1 \pm \sqrt{2}.$$

If $\lambda = \lambda_1 = 1 + \sqrt{2}$ then

$$B - \lambda_1 I_2 = \begin{pmatrix} -1-\sqrt{2} & 1 \\ 1 & 1-\sqrt{2} \end{pmatrix}.$$

We compute the kernel. If we multiply the second row by $-1-\sqrt{2}$ then we get

$$\begin{pmatrix} -1-\sqrt{2} & 1 \\ -1-\sqrt{2} & -(1+\sqrt{2})(1-\sqrt{2}) = 1 \end{pmatrix}.$$

Therefore this matrix has rank one, as expected. The kernel is spanned by $v_1 = (1, 1+\sqrt{2})$ and this vector is an eigenvector with eigenvalue λ_1. Similarly $v_2 = (1, -1+\sqrt{2})$ is an eigenvector with eigenvalue $\lambda_2 = 1-\sqrt{2})$. It follows that $B = PDP^{-1}$, where

$$D = \begin{pmatrix} 1+\sqrt{2} & 0 \\ 0 & 1-\sqrt{2} \end{pmatrix},$$

and

$$P = \begin{pmatrix} 1 & 1 \\ 1+\sqrt{2} & 1-\sqrt{2} \end{pmatrix}.$$

It follows that

$$P^{-1} = \frac{1}{-2\sqrt{2}} \begin{pmatrix} 1-\sqrt{2} & -1 \\ -(1+\sqrt{2}) & 1 \end{pmatrix}.$$

One can check the equality $B = PDP^{-1}$. Now

$$\begin{aligned}
B^n w_1 &= PD^nP^{-1}w_1 \\
&= \frac{1}{-2\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1+\sqrt{2} & 1-\sqrt{2} \end{pmatrix} \begin{pmatrix} (1+\sqrt{2})^n & 0 \\ 0 & (1-\sqrt{2})^n \end{pmatrix} \begin{pmatrix} 1-\sqrt{2} & -1 \\ -(1+\sqrt{2}) & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\
&= \frac{1}{-2\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1+\sqrt{2} & 1-\sqrt{2} \end{pmatrix} \begin{pmatrix} (1+\sqrt{2})^n & 0 \\ 0 & (1-\sqrt{2})^n \end{pmatrix} \begin{pmatrix} -\sqrt{2} \\ -\sqrt{2} \end{pmatrix} \\
&= \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1+\sqrt{2} & 1-\sqrt{2} \end{pmatrix} \begin{pmatrix} (1+\sqrt{2})^n & 0 \\ 0 & (1-\sqrt{2})^n \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\
&= \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1+\sqrt{2} & 1-\sqrt{2} \end{pmatrix} \begin{pmatrix} (1+\sqrt{2})^n \\ (1-\sqrt{2})^n \end{pmatrix} \\
&= \frac{1}{2} \begin{pmatrix} 2g_n = (1+\sqrt{2})^n + (1-\sqrt{2})^n \\ 2g_{n+1} \end{pmatrix}.
\end{aligned}$$

Note that $-1 < 1-\sqrt{2} < 0$. So for large n, $(1-\sqrt{2})^n$ is negative and very small. Therefore g_n is the closest integer to $(1+\sqrt{2})^n/2$.
It is fun to check that this works for various values of n. For $n = 3$ we get 7.03553, which is very close to the real answer, 7. For $n = 5$ we get 41.00609, which even closer to the real answer, 41.
2. Let

$$B = A - \lambda I_2 = \begin{pmatrix} 1-\lambda & 1 \\ 0 & 1-\lambda \end{pmatrix}.$$

The characteristic polynomial is

$$(1-\lambda)(1-\lambda) = 0.$$

This has $\lambda = 1$ as a repeated root.

$$B - I_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

The kernel is spanned by $(0, 1)$. This is an eigenvector with eigenvalue 1.

(ii) If A were diagonalisable it would have a basis of eigenvectors. 1 is the only eigenvalue and the corresponding eigenspace is one dimensional. Therefore A does not have a basis of eigenvectors and so A is not diagonalisable.

3. (i) This matrix represents rotation through Θ radians. If A is diagonalisable then it has a basis of eigenvectors. An eigenvector would span a line that is fixed under rotation through Θ radians. There are only two possibilities, $\Theta = 0$ and $\Theta = \pi$. In the first case $A = I_2$, which is surely diagonalisable. In the second every vector is an eigenvector, with eigenvalue -1. So A certainly has a basis of eigenvectors and A is diagonalisable.

(ii) We just have to show that A has two distinct complex eigenvalues. The characteristic polynomial is a quadratic polynomial. If $\Theta = 0$ or $\Theta = \pi$ then A has two real eigenvectors with real eigenvalues. Otherwise the eigenvalues are not real. But the roots to a real quadratic polynomial come in complex conjugate pairs. So if one root is not real the other one is the complex conjugate. In particular the two roots are different. For each root we can find an eigenvector and so we have a basis of eigenvectors, and since the dimension is two, we have a basis of eigenvectors.

(iii) Now we have to do some work. Let

$$B = A - \lambda I_2 = \begin{pmatrix} \cos\theta - \lambda & -\sin\theta \\ \sin\theta & \cos\theta - \lambda \end{pmatrix}.$$

The characteristic polynomial is

$$(\cos\theta - \lambda)(\cos\theta - \lambda) + \sin\theta \sin\theta = 0.$$

Rearranging gives,

$$\cos^2\theta + \sin^2\theta - 2\lambda\cos\theta + \lambda^2 = 0.$$

Using a well-known identity we finally get

$$\lambda^2 - 2\cos\theta\lambda + 1 = 0.$$

Using the quadratic formula gives,

$$\begin{aligned}\lambda &= \frac{2\cos\theta \pm \sqrt{4\cos^2\theta - 4}}{2}\\ &= \cos\theta \pm \sqrt{\cos^2\theta - 1}\\ &= \cos\theta \pm \sqrt{-\sin^2\theta}\\ &= \cos\theta \pm i\sin\theta\\ &= e^{\pm i\theta}.\end{aligned}$$

As expected the two roots are complex conjugate. Suppose we plug in $\lambda_1 = e^{i\theta}$. We get

$$B = \begin{pmatrix} -i\sin\theta & -\sin\theta \\ \sin\theta & -i\sin\theta \end{pmatrix}.$$

The second row is i times the first row. Solving in the usual way to find the kernel gives $v_1 = (1, -i)$. The other eigenvector is $v_2 = (1, i)$ (what else but the complex conjugate of v_1). Therefore $A = PDP^{-1}$, where

$$D = \begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix},$$

and

$$P = \begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix}.$$

Note that

$$P^{-1} = \frac{1}{2i}\begin{pmatrix} i & -1 \\ i & 1 \end{pmatrix}.$$

4. Suppose that there are scalars $r_1, r_2, \dots, r_n$ such that

$$0 = r_1v_1 + r_2v_2 + \dots + r_nv_n.$$

If we apply T to both sides we get

$$0 = T(r_1v_1 + r_2v_2 + \dots + r_nv_n).$$

By linearity the RHS expands to

$$r_1T(v_1) + r_2T(v_2) + \dots + r_nT(v_n).$$

Thus

$$0 = r_1T(v_1) + r_2T(v_2) + \dots + r_nT(v_n).$$

By independence of the vectors $T(v_1)$, $T(v_2)$, $\dots$, $T(v_n)$, we have $r_1 = r_2 = \dots = r_n = 0$. But then $v_1, v_2, \dots, v_n$ are independent.

5. The matrices

$$\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \quad \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \quad \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

are diagonalisable as they are diagonal to start with. The matrices

$$\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \quad \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} \quad \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix} \quad \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix},$$

all have distinct eigenvalues (namely 0 and 1) and so they are diagonalisable as well. The matrices

$$\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \quad \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix},$$

have characteristic polynomial

$$(1+\lambda)\lambda + 1.$$

Neither 0 nor 1 is a root of this polynomial and so these matrices don't have any eigenvectors. Therefore these matrices cannot be diagonalised.

This leaves

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \quad \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix},$$

which have characteristic polynomial λ^2 and

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

which have characteristic polynomial $(\lambda + 1)^2$. In all six cases one can compute that the corresponding eigenspace is one dimensional. Therefore these matrices cannot be diagonalised.

6. (a) Consider the set of all functions W^X from a set X to a vector space W. This is naturally a vector space, where we define addition and scalar multiplication pointwise. We just need to check that the set of linear maps is a vector subspace.

The function which sends every vector to zero is easily seen to be linear. Therefore the subset of linear functions is non-empty. Given two linear transformations $\phi\colon F^n \longrightarrow F^m$ and $\psi\colon F^n \longrightarrow F^m$, let

$$\phi + \psi\colon F^n \longrightarrow F^m,$$

be the function

$$(\phi + \psi)(v) = \phi(v) + \psi(v).$$

We check that this is a linear transformation. Given two vectors v and w, we have

$$\begin{aligned}(\phi+\psi)(v+w) &= \phi(v+w)+\psi(v+w)\\ &= \phi(v)+\phi(w)+\psi(v)+\psi(w)\\ &= \phi(v)+\psi(v)+\phi(w)+\psi(w)\\ &= (\phi+\psi)(v)+(\phi+\psi)(w).\end{aligned}$$

Thus $\phi+\psi$ respects addition. Now suppose that v is a vector and r is a scalar. We have

$$\begin{aligned}(\phi+\psi)(rv) &= \phi(rv)+\psi(rv)\\ &= r\phi(v)+r\psi(v)\\ &= r(\phi(v)+\psi(v))\\ &= r(\phi+\psi)(v).\end{aligned}$$

Thus $\phi+\psi$ respects scalar multiplication. It follows that $\phi+\psi$ is a linear transformation. Thus the set of linear transformations is closed under addition.

Now suppose we are given $\phi\colon F^n \longrightarrow F^m$ and a scalar λ. Define a function

$$\lambda\phi\colon F^n \longrightarrow F^m,$$

by the rule $(\lambda\phi)(v) = \lambda\phi(v)$. Suppose that v and w are vectors. Then

$$\begin{aligned}(\lambda\phi)(v+w) &= \lambda(\phi(v+w))\\ &= \lambda(\phi(v)+\phi(w))\\ &= \lambda\phi(v)+\lambda\phi(w)\\ &= (\lambda\phi)(v)+(\lambda\phi)(w).\end{aligned}$$

Therefore $\lambda\phi$ respects addition. Now suppose that v is a vector and r is a scalar. Then

$$\begin{aligned}(\lambda\phi)(rv) &= \lambda\phi(rv)\\ &= \lambda r\phi(v)\\ &= r\lambda\phi(v)\\ &= r(\lambda\phi)(v).\end{aligned}$$

Therefore $\lambda\phi$ respects scalar multiplication. It follows that $\lambda\phi$ is a linear transformation. Hence the set of linear transformations is closed under scalar multiplication. But then the set of linear transformations is a subvector space of the set of all functions $(F^m)^{F^n}$.

(b) Let ϕ and ψ be two linear functions and let A and B be the associated matrix.

$$(\phi + \psi)(e_i) = \phi(e_i) + \psi(e_i) = Ae_i + Be_i = (A + B)e_i.$$

Therefore $A + B$ is the matrix associated to $\phi + \psi$. Thus the function which assigns to a linear transformation the corresponding matrix respects addition. Now suppose that λ is a scalar. Then

$$(\lambda\phi)(e_i) = \lambda(\phi(e_i)) = \lambda(Ae_i) = (\lambda A)e_i.$$

Thus the matrix associated to $\lambda\phi$ is λA. Thus the function which assigns to a linear transformation the corresponding matrix respects scalar multiplication. Thus the function which assigns to a linear transformation the corresponding matrix is linear.