18.06 Problem Set 7 Solutions

Problem 1: Do problem 11 from section 8.3.

Solution A Markov matrix must conserve probability. Hence the columns must sum to 1:

$$A = \begin{pmatrix} .7 & .1 & .2 \\ .1 & .6 & .3 \\ .2 & .3 & .5 \end{pmatrix} \tag{1}$$

The steady state vector x satisfies $Ax = x$. In otherwords, x is an eigenvector of A with eigenvalue 1. Therefore x solves:

$$(A - I)x = \begin{pmatrix} -.3 & .1 & .2 \\ .1 & -.4 & .3 \\ .2 & .3 & -.5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 0 \tag{2}$$

$$\sim \begin{pmatrix} 0 & -.11 & .11 \\ .1 & -.4 & .3 \\ 0 & .11 & -.11 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 0 \tag{3}$$

$$\tag{4}$$

Take $x_3 = 1$, then $x_2 = 1$ and $.1x_1 = .4 - .3 = .1$ so $x_1 = 1$ and $x = (111)^T$.

In general, a symmetric Markov matrix $A^T = A$ has a steady solution $x = (111)^T$. This follows from:

- A is Markov $\rightarrow$ columns of A sum to 1.
- $A^T = A \rightarrow$ rows of A sum to 1.
- The vector $x = (111)^T$ sums the row elements. Hence, $Ax = x$.

Problem 2: Do problem 12 from section 8.3.

Solution

$$B = (A - I)x = \begin{pmatrix} -.2 & .3 \\ .2 & -.3 \end{pmatrix} \tag{5}$$

The eigenvalues satisfy the characteristic equation:

$$(-.2-\lambda)(-.3-\lambda)-(.2)(.3) = 0 \quad (6)$$
$$\lambda(\lambda+.5) = 0 \quad (7)$$

The eigenvalues are $\lambda = 0$ and $\lambda = -.5$.

In general, when A is Markov, $A - I$ will have a $\lambda = 0$ eigenvalue. Specifically, this follows from the (more general) fact that if μ is an eigenvalue of any matrix A, then $\mu - c$ is an eigenvalue of $A - cI$. Markov matrices have an eigenvalue of 1, hence $A - I$ must have an eigenvalue $1 - 1 = 0$.

The eigenvectors of B are, $\mathbf{x}_1 = (.3, .2)^T$ corresponding to $\lambda = 0$, and $\mathbf{x}_2 = (1, -1)^T$ corresponding to $\lambda = -0.5$. A general solution to the ODE $\frac{du}{dt} = (A - I)u$ has the form:

$$u = c_1 e^{0\cdot t}\mathbf{x}_1 + c_2 e^{-.5t}\mathbf{x}_2 \quad (8)$$

Here c_1 and c_2 are integration constants (determined by initial values). As $t \to \infty$, u becomes:

$$u \longrightarrow c_1\mathbf{x}_1 \quad (9)$$

Problem 3: Do problem 16 section 8.3.

Solution

$$A = \begin{pmatrix} .4 & .2 & .3 \\ .2 & .4 & .3 \\ .4 & .4 & .4 \end{pmatrix} \quad (10)$$

Since A is a Markov matrix, we know $\lambda = 1$ is an eigenvalue. In addition, $detA = 0$, so $\lambda = 0$ must also be an eigenvalue (ie $det(A - 0 \cdot I) = 0$). The third eigenvalue $\lambda = 0.2$ can be found by $i)$ inspection, $ii)$ using $MATLAB$, $iii)$ by direct computation. Using MATLAB, we can diagonalize $A = S\Lambda S^{-1}$, where the columns of S are eigenvectors of A:

$$S = \begin{pmatrix} .5145 & .7071 & -.4082 \\ .5145 & -.7071 & -.4082 \\ .686 & 0 & .8165 \end{pmatrix} \quad (11)$$

$$\Lambda = \begin{pmatrix} 1 & 0 & 0 \\ 0 & .2 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad (12)$$

Now $A^k = S\Lambda^k S^{-1}$, so:

$$\begin{aligned} A^k\mathbf{u}_0 &= (S\Lambda S^{-1})^k\mathbf{u}_0 && (13)\\ &= S\Lambda^k S^{-1}\mathbf{u}_0 && (14)\\ & && (15) \end{aligned}$$

Since

$$S^{-1} = \begin{pmatrix} .5831 & .5831 & .5831 \\ .7071 & -.7071 & 0 \\ -.4899 & -.4899 & .7348 \end{pmatrix} \tag{16}$$

then

$$S^{-1}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} .5831 \\ .7071 \\ -.4899 \end{pmatrix} \tag{17}$$

When we expand out $A^k\mathbf{u}_0$, we have:

$$A^k\mathbf{u}_0 = .5831(1)^k\begin{pmatrix} .5145 \\ .5145 \\ .686 \end{pmatrix} + .7071(0.2)^k\begin{pmatrix} .7071 \\ -.7071 \\ 0 \end{pmatrix} + (-.4899)(0)^k\begin{pmatrix} -.4082 \\ -.4082 \\ .8165 \end{pmatrix} \tag{18}$$

Hence, as $k \to \infty$, $A^k\mathbf{u}_0 \to .5831(.5145, .5145, .686)^T$. Similarly, for $\mathbf{u}_0 = (100, 0, 0)^T = 100(1, 0, 0)^T$, we can simply rescale the previous limit found for $\mathbf{u}_0 = (1, 0, 0)$ by 100:

$$A^k\mathbf{u}_0 = .5831 \cdot 100\begin{pmatrix} .5145 \\ .5145 \\ .686 \end{pmatrix} \tag{19}$$

Problem 4: Do problem 4 section 6.3.

Solution

To show $v + w$ is a constant, differentiate w.r.t. time:

$$\begin{aligned} \frac{d}{dt}(v + w) &= (w - v) + (v - w) && (20)\\ &= 0 && (21) \end{aligned}$$

We can cast the system into matrix form:

$$\frac{d}{dt}\begin{pmatrix} v \\ w \end{pmatrix} = \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} v \\ w \end{pmatrix} \tag{22}$$

The two eigenvalues are $\lambda = 0$ and $\lambda = 2$ (since the characteristic equation is $\lambda(\lambda + 2) = 0$). The corresponding eigenvectors are:

$$\mathbf{u}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \tag{23}$$

for $\lambda = 0$ and

$$\mathbf{u}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \tag{24}$$

for $\lambda = 2$.

To find v and w at $t = 1$ and $t \to \infty$, we solve the initial value problem. The general solution is:

$$\begin{pmatrix} v \\ w \end{pmatrix} = c_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + c_2 e^{-2t} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \tag{25}$$

The initial data $v(0) = 30$ and $w(0) = 10$ determine the constants c_1 and c_2:

$$\begin{pmatrix} 30 \\ 10 \end{pmatrix} = c_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + c_2 \begin{pmatrix} 1 \\ -1 \end{pmatrix} \tag{26}$$

Or $c_1 = 20$ and $c_2 = 10$. Hence $v(1) = 20 + 10e^{-2}$, $w(1) = 20 - 10e^{-2}$. Meanwhile $v(\infty) = 20$, $w(\infty) = 20$.

Problem 5: Do problem 11 in section 6.3.

Solution We have the ODE:

$$\frac{d}{dt} \begin{pmatrix} y \\ y' \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} y \\ y' \end{pmatrix} \tag{27}$$

The solution is:

$$\begin{pmatrix} y \\ y' \end{pmatrix} = e^{At} \begin{pmatrix} y(0) \\ y'(0) \end{pmatrix} \tag{28}$$

where

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \tag{29}$$

Note that

$$A^2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \tag{30}$$

and every other power $A^k = 0$ for $k > 1$. Therefore $e^{At} = I + At$:

$$\begin{pmatrix} y \\ y' \end{pmatrix} = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} \begin{pmatrix} y(0) \\ y'(0) \end{pmatrix} \tag{31}$$

so that we recover $y(t) = y(0) + ty'(0)$.

Problem 6: Do problem 6 in section 10.2.

Solution

- a) If A is a real matrix, then $A + \imath I$ is invertible. FALSE. Note that this statement is equivalent to asking "can a real matrix have an imaginary eigenvalue?". Take

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \tag{32}$$

  Then

$$A + \imath I = \begin{pmatrix} \imath & 1 \\ -1 & \imath \end{pmatrix} \tag{33}$$

  has determinant $\imath^2 + 1 = 0$ so that it is not invertible.

- b) If A is Hermitian then $A + \imath I$ is invertible. TRUE. If A is an $n \times n$ Hermitian matrix then A has n real eigenvalues. Therefore $A + \imath I$ has n complex eigenvalues none of which are zero. Therefore $A + \imath I$ is invertible.

- c) If U is unitary then $U + \imath I$ is invertible. FALSE. If U is unitary then every eigenvalue of U is of the form $e^{\imath\theta}$. We can construct a U which has an eigenvalue $-\imath$:

$$U = \begin{pmatrix} -\imath & 0 \\ 0 & -\imath \end{pmatrix} \tag{34}$$

  Note that $U^*U = I$, while $U + \imath I = 0$ which is not invertible.

Problem 7: Do problem 15 in section 10.2.

Solution Diagonalize:

$$K = \begin{pmatrix} 0 & -1+\imath \\ 1+\imath & \imath \end{pmatrix} \tag{35}$$

The characteristic equation is:

$$\begin{aligned} det(K - \lambda I) &= -\lambda(\imath - \lambda) - (1+\imath)(-1+\imath) && (36) \\ &= -\lambda(\imath - \lambda) - (1+\imath)(-1+\imath) && (37) \\ &= \lambda^2 - \imath\lambda - 2\imath^2 && (38) \\ &= (\lambda - 2\imath)(\lambda + \imath) && (39) \end{aligned}$$

For the $\lambda = 2\imath$ eigenvalue:

$$(K - 2\imath I)\mathbf{x} = \begin{pmatrix} -2\imath & -1+\imath \\ 1+\imath & -\imath \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \tag{40}$$

Note that the second row of the matrix is just a multiple $(\frac{-2\imath}{1+\imath})$ of the first row. If we set $x_1 = 1$, then $-2\imath + (-1+\imath)x_2 = 0$ or $x_2 = 1 - \imath$. The first eigenvector is:

$$\mathbf{x} = \begin{pmatrix} 1 \\ 1-\imath \end{pmatrix} \tag{41}$$

For the eigenvalue $\lambda = -\imath$:

$$(K + \imath I)\mathbf{y} = \begin{pmatrix} \imath & -1+\imath \\ 1+\imath & 2\imath \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \tag{42}$$

We can take $y_2 = 1$, $y_1 = -1 - \imath$:

$$\mathbf{y} = \begin{pmatrix} -1-\imath \\ 1 \end{pmatrix} \tag{43}$$

Note that all the eigenvalues are purely imaginary. This follows from $K^* = -K$. If we write $H = \imath K$ then $H^* = -\imath K^* = H$ is Hermitian and therefore has real eigenvalues.

To diagonalize K in the form requested $K = U\Lambda U^*$, we must ensure that U is a unitary matrix (ie. $U^* = \bar{U^T}$, where $\bar{A}$ means to conjugate every element of matrix A.). We can construct a unitary U out of the eigenvectors of K provided they are normalized to 1. To normalize each eigenvector we multiply them by $1/\sqrt{3}$. K is then diagonailzed as:

$$K = \begin{pmatrix} 1/\sqrt{3} & -(1+\imath)/\sqrt{3} \\ (1-\imath)/\sqrt{3} & 1/\sqrt{3} \end{pmatrix} \begin{pmatrix} 2\imath & 0 \\ 0 & -\imath \end{pmatrix} \begin{pmatrix} 1/\sqrt{3} & (1+\imath)/\sqrt{3} \\ (-1+\imath)/\sqrt{3} & 1/\sqrt{3} \end{pmatrix} \quad (44)$$

Problem 8: Do problem 16 in section 10.2.

Solution

Diagonalize:

$$Q = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \quad (45)$$

The characteristic equation is:

$$det(Q - \lambda I) = (\cos\theta - \lambda)^2 + \sin^2\theta \quad (46)$$

$$(47)$$

which yields:

$$\lambda = \cos\theta \pm \imath \sin\theta \quad (48)$$

$$= e^{\pm\imath\theta} \quad (49)$$

For the $\lambda = \cos\theta + \imath\sin\theta$ eigenvalue:

$$(Q - (\cos\theta + \imath\sin\theta)I)\mathbf{x} = \begin{pmatrix} -\imath\sin\theta & -\sin\theta \\ \sin\theta & -\imath\sin\theta \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \quad (50)$$

Note that the second row of the matrix is just a multiple ($\imath$) of the first row. If we set $x_1 = 1$, then $x_2 = -\imath$. The first normalized eigenvector is:

$$\mathbf{x} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -\imath \end{pmatrix} \quad (51)$$

Note that since Q is a real matrix, both the eigenvalues and eigenvectors must come in complex conjugate pairs (Why must this be true?). Therefore the second eigenvector, (for $\lambda = \cos\theta - \imath\sin\theta$) is:

$$\mathbf{y} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ \imath \end{pmatrix} \quad (52)$$

Note that all the eigenvalues lie on the complex unit circle. This is always true for an orthogonal (or unitary) matrix. For instance, if we take an eigenvector x with eigenvalue λ, we can write $Ux = \lambda x$ and $x^* U^* = \bar{\lambda} x^*$. Mulitplying these row and column vectors we have: $x^* U^* U x = \lambda \bar{\lambda} x^* x$. Since $U^* U = I$, and x is nonzero, we have $|\lambda|^2 = 1$ or that the magnitude of the eigenvalue is 1.

To diagonalize Q in the form requested $Q = U \Lambda U^*$:

$$K = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -\imath\sqrt{2} & \imath/\sqrt{2} \end{pmatrix} \begin{pmatrix} e^{\imath\theta} & 0 \\ 0 & e^{-\imath\theta} \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} & \imath/\sqrt{2} \\ 1/\sqrt{2} & -\imath/\sqrt{2} \end{pmatrix} \tag{53}$$

Problem 9: Do problem 15 in section 10.3.

Solution

This question concerns counting multiplications for computing convolutions via FFT. The answer is, each of the two FFT operations require $1/2 n \log n$ multiplications. Meanwhile, the convolution (when performed in Fourier space), require n pointwise multiplications.

The idea behind using FFT for computing fast computations is based on the formula $F(x * y) = F(x) \cdot F(y)$, where $F(x)$ and $F(y)$ are the Fourier transformed vectors of x and y, and $F(x) \cdot F(y)$ implies pointwise multiplication (ie. multiplying $(1, 2, 3)$ with $(4, 5, 6)$ pointwise yields $(4, 10, 18)$). The textbook writes this in matrix notation as $x * y = F(E(F^{-1} x))$, where E denotes the pointwise multiplication by $F(y)$ (note the textbook reverses the definition engineers and physics use for a Fourier transform. This is why F and F^{-1} have been swapped in the last formula.)

To see why the FFT requires $1/2 n \log n$ operations, consider the case when n is a power of 2, let $n = 2^m$. If x is a vector of length 2^n, the FFT relies on the recursive formula for the kth component $F(x)_k = F(x_{even})_k + \omega^k F(x_{odd})_k$. Here x_{even} is a vector of length 2^{n-1} composed of the elements at even locations and x_{odd} has length 2^{n-1} with elements at the odd locations, while $\omega = e^{\imath 2\pi/n}$.

Ie. let $x = (x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8)$. Then

$$\begin{aligned} F(x)_k &= x_1 + x_2\omega^k + x_3\omega^{2k} + x_4\omega^{3k} + x_5\omega^{4k} && (54) \\ &+ x_6\omega^{5k} + x_7\omega^{6k} + x_8\omega^{7k} && (55) \end{aligned}$$

which we regroup as:

$$\begin{aligned} F(x)_k &= (x_1 + x_3\omega^{2k} + x_5\omega^{4k} + x_7\omega^{6k}) && (56) \\ &+ (x_2 + x_4\omega^{2k} + x_6\omega^{4k} + x_8\omega^{6k})\omega^k && (57) \end{aligned}$$

and again to:

$$F(x)_k = [(x_1 + x_5\omega^{4k}) + (x_3 + x_7\omega^{4k})\omega^{2k}] \quad (58)$$

$$+ [(x_2 + x_6\omega^{4k}) + (x_4 + x_8\omega^{4k})\omega^{2k}]\omega^k \quad (59)$$

Note that in the last equation, the term inside each set of rounded brackets is a 2 element Fourier transform. ie. we need to calculate the term inside the rounded brackets only when $k = 0$ and $k = 1$ since $k = 2$ corresponds to $\omega^8 = 1$. Meanwhile, the terms inside the rectangular brackets are Fourier transforms of length 4.

Now to count things up, introduce the function G_m as the number of multiplications required to compute $F(x)$ where x has length $n = 2^m$. Then G_m satisfies the recursion formula:

$$G_m = 2G_{m-1} + 2^m/2 \quad (60)$$

That is to say, each Fourier transform requires *i*) two Fourier transforms of half the length, *ii*) plus $2^m/2$ additional multiplications required to reconstruct the total transform. There are $2^m/2$ additional multiplications and not 2^m multiplications because $\omega^{n/2} = -1$. Therefore terms such as $[(x_2 + x_6\omega^{4k}) + (x_4 + x_8\omega^{4k})\omega^{2k}]\omega^k$, for $k = 0 \ldots 3$ are just negatives of $k = 4 \ldots 7$.

To solve for G_m, note that $G_0 = 0$ and $G_1 = 1$:

$$G_m = 2(2G_{m-2} + 2^{m-2}) + 2^{m-1} \quad (61)$$

$$= 4G_{m-2} + 2^{m-1} + 2^{m-1} \quad (62)$$

$$\ldots \quad (63)$$

$$= 2^m G_0 + m2^{m-1} \quad (64)$$

$$= m2^{m-1} \quad (65)$$

$$= \frac{n}{2}\log_2 n \quad (66)$$

Alternatively, we can use induction. If for example we conjecture $G_m = m2^{m-1}$, then certainly $G_0 = 0$, $G_1 = 1$. Then for $G_{m+1} = 2G_m + 2^m$. Using the induction hypothesis:

$$G_{m+1} = 2m2^{m-1} + 2^m \quad (67)$$

$$= 2^m(m+1) \quad (68)$$

$$= 2^{(m+1)-1}(m+1) \quad (69)$$

Hence, $G_m = m2^{m-1}$ holds for all integers m.