Non-book problems for Assignment 14

NB 14.1. Let $A \in M_{2\times 2}(\mathbf{R})$ and let $f_A(t) = \det(A - tI)$, the characteristic polynomial of A with the variable named t. Show that

$$f_A(t) = t^2 - \mathrm{tr}(A)t + \det(A). \tag{1}$$

(*Note*: The simplicity of equation (1) is special to 2×2 matrices. In the next exercise, you will figure out what (1) generalizes to when $n > 2$.

NB 14.2. Let k, n be integers with $1 \le k \le n$. The k^{th} *elementary symmetric function* σ_k of variables $x_1, x_2, \dots, x_n$ is defined by

$$\sigma_k(x_1, \dots, x_n) = \sum_{i_1 < i_2 < \dots < i_k} x_{i_1} x_{i_2} \dots x_{i_k}.$$

(The sum is over all ordered k-tuples of integers $(i_1, i_2, \dots, i_n)$ satisfying $1 \le i_1 < i_2 < \dots < i_k \le n$.) For example, with $n = 3$,

$$\begin{aligned} \sigma_1(x_1, x_2, x_3) &= x_1 + x_2 + x_3 \ , \\ \sigma_2(x_1, x_2, x_3) &= x_1x_2 + x_1x_3 + x_2x_3 \ , \quad \text{and} \\ \sigma_3(x_1, x_2, x_3) &= x_1x_2x_3 \ . \end{aligned}$$

(The Greek letter σ is a lower-case sigma.)

(a) Let $\lambda_1, \lambda_2, \dots, \lambda_n \in \mathbf{R}$ and let D be the $n \times n$ diagonal matrix with $D_{ii} = \lambda_i, 1 \le i \le n$. Show that f_D, the characteristic polynomial of D, satisfies

$$\begin{aligned} (-1)^n f_D(t) &= (t - \lambda_1)(t - \lambda_2) \dots (t - \lambda_n) \\ &= t^n \ - \ \sigma_1(\lambda_1, \dots, \lambda_n)\, t^{n-1} \ + \ \sigma_2(\lambda_1, \dots, \lambda_n)\, t^{n-2} \\ &\quad - \ \dots \ + (-1)^n \sigma_n(\lambda_1, \lambda_2, \dots, \lambda_n). \end{aligned} \tag{2}$$

(b) Using part (a), show that if A is a diagonalizable $n \times n$ matrix with eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$ (not necessarily distinct; an eigenvalue of algebraic multiplicity m appears m times in this list) then the characteristic polynomial $f_A(t)$ is also given by the right-hand side of equation (2). In particular,

$$\begin{aligned} \det(A) \ = \ \det(A - tI)|_{t=0} \ &= \ f_A(0) \\ &= \ \text{constant term in } f_A(t) \\ &= \ \sigma_n(\lambda_1, \lambda_2, \dots, \lambda_n) \\ &= \ \lambda_1\lambda_2 \dots, \lambda_n \ . \end{aligned}$$

Thus *the determinant of A is the product of its eigenvalues* (counted with multiplicity):

$$\det(A) = \lambda_1\lambda_2 \dots \lambda_n \ . \tag{3}$$

Furthermore, the coefficient of t^{n-1} in $f_A(t)$ is always $(-1)^{n-1}\mathrm{tr}(A)$.

NB 14.3 (Explicit formula for the Fibonacci numbers, via "eigenstuff") The *Fibonacci numbers* are the terms of the sequence

$$1, 1, 2, 3, 5, 8, 13, \ldots, \tag{4}$$

in which the first two terms are 1, and every term after that is the sum of the two previous terms. In this problem you will use matrix algebra (specifically, "eigenstuff") to compute an explicit formula for the Fibonacci numbers and some related sequences.

For $n \geq 1$ let f_n be the n^{th} term of the Fibonacci sequence (4). Thus $f_1 = f_2 = 1$ and $f_{n+2} = f_{n+1} + f_n$ for $n \geq 1$. To simplify some formulas below, define $f_0 = 0$ (effectively, just inserting a 0 at the start of the sequence (4)), and observe that $f_2 = f_1 + f_0$, so that the recursive relation $f_{n+2} = f_{n+1} + f_n$ now holds for $n \geq 0$.

Define a sequence of vectors $\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \ldots$ in $\mathbf{R}^2$ by

$$\mathbf{x}_n = \begin{pmatrix} f_n \\ f_{n+1} \end{pmatrix}. \tag{5}$$

(a) Show that this sequence of vectors $(\mathbf{x}_n)_{n=0}^{\infty}$ satisfies

$$\mathbf{x}_{n+1} = A\mathbf{x}_n \quad \text{for all } n \geq 0, \tag{6}$$

where

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}. \tag{7}$$

Then use (6) to deduce that

$$\mathbf{x}_n = A^n \mathbf{x}_0, \quad \text{for all } n \geq 1. \tag{8}$$

(b) (i) Find the eigenvalues of A. You should find that these are two distinct, real numbers λ_1 and λ_2.

By problem NB 14.1, $\lambda_1\lambda_2 = \det(A) = -1$, so $|\lambda_1|\,|\lambda_2| = 1$. From your formula for the eigenvalues, you should *easily* (without a calculator) be able to see that one of the eigenvalues has absolute value greater than 1, so the other must have absolute value less than 1 (which can also be seen easily, but slightly less easily). Let λ_1 be the eigenvalue with $|\lambda_1| > 1$ and let λ_2 be the other eigenvalue. Below, let D be the diagonal matrix $\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$.

(ii) Find eigenvectors $\mathbf{v}_1, \mathbf{v}_2$ corresponding to λ_1, λ_2 respectively. Since $\lambda_1 \neq \lambda_2$, the set $\{\mathbf{v}_1, \mathbf{v}_2\}$ is linearly independent set in the two-dimensional vector space $\mathbf{R}^2$, hence is a basis of $\mathbf{R}^2$ (an A-eigenbasis).

(c) Express the vector $\mathbf{x}_0 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ in terms of $\mathbf{v}_1$ and $\mathbf{v}_2$. I.e. find $c_1, c_2 \in \mathbf{R}$ such that $\mathbf{x}_0 = c_1\mathbf{v}_1 + c_2\mathbf{v}_2$. Then show that

$$A^n\mathbf{x}_0 = c_1\,{\lambda_1}^n\,\mathbf{v}_1 + c_2\,{\lambda_2}^n\,\mathbf{v}_2\ , \tag{9}$$

and use equation (9) (together with equation (8) and the definition of $\mathbf{x}_n$) to write down an explicit formula for the n^{th} Fibonacci number f_n.

If you've done everything correctly, in your formula for f_n you'll see the irrational number $\sqrt{5}$ appearing in two fractions that are raised to higher and higher powers. Yet the Fibonacci numbers are integers! Remarkably, not only do all the $\sqrt{5}$'s cancel, allowing your formula to work out to a *rational* number for each n, the fractions "conspire" with each other to produce an *integer*.

(d) For each $n \geq 1$, compute D^n explicitly in terms of the eigenvalues of D. (Recall that for a diagonal matrix, the eigenvalues are precisely the diagonal entries.) Relate this to FIS exercise 5.1/16b.

(e) Use the information found in part (b) to construct an invertible matrix $C \in M_{2\times 2}(\mathbf{R})$ such that $D = C^{-1}AC$.

Since $D = C^{-1}AC$, we also have $A = CDC^{-1}$ (why?). From non-book problem NB 11.3, we then have

$$A^n = CD^nC^{-1} \quad \text{for any } n \geq 1. \tag{10}$$

Since $\mathbf{x}_n = A^n\mathbf{x}_0$, we now have two more ways of computing $\mathbf{x}_n$:

(i) Compute A^n explicitly, and then multiply $\mathbf{x}_0$ by the result.

(ii) Use the associativity of matrix multiplication to compute $A^n\mathbf{x}_0 = CD^nC^{-1}\mathbf{x}_0$ without ever computing A^n itself, by doing the matrix computation in "right-to-left" order (as indicated by the parentheses below):

$$\underbrace{C\Big(\overbrace{D^n(\ \underbrace{C^{-1}\mathbf{x}_0}_{\text{compute first}}\ )}^{\text{compute second}}\Big)}_{\text{compute third}}. \tag{11}$$

Note that methods (i), (ii), and the method in part (c) should all give the same answer! After you've done the computation all three ways, compare the methods and see how and where the same information is packaged differently.

(f) (i) Compute $\lim_{n\to\infty} \frac{f_{n+1}}{f_n}$ and relate the answer to eigenvalue(s) of A (equivalently, eigenvalues of D). Then look back at equation (9) and see how you could have predicted the value of this limit.

(ii) Consider a modified Fibonacci sequence $(a_n)_{n=0}^{\infty}$ in which the first two terms are replaced by arbitrary numbers a_0 and a_1, not both 0, but the recursive rule $a_n = a_{n-1} + a_{n-2}$ is still used for $n \geq 2$. Using (9), show that $\lim_{n\to\infty} \frac{a_{n+1}}{a_n}$ always exists, and is "almost" independent of the initial terms a_0 and a_1, in the following sense: the limit always exists, and is always either λ_1 or λ_2, but is λ_1 unless the vector $\begin{pmatrix} a_0 \\ a_1 \end{pmatrix}$ lies in the λ_2-eigenspace of A. (So, if you threw a dart at the xy plane to select the vector $\begin{pmatrix} a_0 \\ a_1 \end{pmatrix}$, the above limit would virtually always be λ_1; the chance of your hitting the line that represents the λ_2-eigenspace would be virtually zero.)

NB 14.4 In the setting of non-book problem NB 12.1, a function $f : V \times V \to W$ is called *symmetric* if $f(v, w) = f(w, v)$ (no sign-change).

For any vector spaces V and W, the concepts of *bilinearity* and *symmetry* of a function $f : V \times V \to W$ make sense. A concept that makes sense *only for the case* $W = \mathbf{R}$ (regardless of V) is *positive-definiteness*: a symmetric, bilinear function $f : V \times V \to \mathbf{R}$ is called *positive-definite* if $f(v, v) > 0$ for all $v \neq 0_V$.

Show that a function $f : V \times V \to \mathbf{R}$ is an inner product on V if and only if f is symmetric, bilinear, and positive-definite.

(As seen in the definition on pp. 327–328 of FIS, for such a function, it is common to use notation such as "$\langle v, w\rangle$" for the output of such a function when the inputs are v and w, rather than to require choosing a name, e.g. f, for the function, and writing "$f(v, w)$" for the output.)

> FYI: A common *definition* of " inner product" on a real vector space V is "symmetric, bilinear, positive-definite function from $V \times V$ to $\mathbf{R}$." When we want to define inner products for *complex* vector spaces, as in FIS, we cannot avail ourselves of this simple wording.

NB 14.5 Let V, W be vector spaces, let $\langle\cdot, \cdot\rangle$ be an inner product on W, and let $T : V \to W$ be a linear trasnformation. Define a function $\langle\cdot, \cdot\rangle' : V \times V \to \mathbf{R}$ by

$$\langle v_1, v_2\rangle' = \langle\, T(v_1), T(v_2)\,\rangle.$$

Show that $\langle\cdot, \cdot\rangle'$ is an inner product on V if and only if T is one-to-one.

(When working this out, you should find that "one-to-one-ness" is relevant to only one of the criteria in the definition of "inner product".)

NB 14.6 Let $m, n \geq 1$, let $\langle\cdot, \cdot\rangle$ and $\langle\cdot, \cdot\rangle'$ be the standard inner products on $\mathbf{R}^m$ and $\mathbf{R}^n$ respectively, and let $A \in M_{m\times n}(\mathbf{R})$. Let $\vec{a} = \begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix} \in \mathbf{R}^m, \vec{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \in \mathbf{R}^n$.

Show that

$$\langle \vec{a}, A\vec{b} \rangle \ = \ \sum_{i,j} a_i A_{ij} b_j \ = \ \vec{a}^{\,t} A \,\vec{b} \ = \ \langle A^t \vec{a}, \vec{b} \rangle'. \tag{12}$$

(Here "$\sum_{i,j}$" is short-hand for the double-sum "$\sum_{i=1}^{m} \sum_{j=1}^{n}$", and we are treating a 1×1 matrix as a real number [the matrix's sole entry]; the 1×1 matrix $\vec{a}^{\,t} A \,\vec{b}$ is a product of the three matrices $\vec{a}^{\,t} \in M_{1\times m}(\mathbf{R})$, $A \in M_{m\times n}(\mathbf{R})$, and $\vec{b} \in M_{n\times 1}(\mathbf{R})$.)