Math 510 — Fall 2022
Homework 6 — Drew Armstrong

Complete any 5 problems.

1. Principal Axes. Consider the following polynomial in two variables:

$$f(\mathbf{x}) = 1 + \begin{pmatrix} 1 & 0 \end{pmatrix} \mathbf{x} + \mathbf{x}^T \begin{pmatrix} 0 & 2 \\ 2 & -3 \end{pmatrix} \mathbf{x} = 1 + x + 4xy - 3y^2.$$

(a) Find a matrix Q satisfying $Q^TQ = I$ and $\det(Q) = 1$ and a vector $\mathbf{t} \in \mathbb{R}^2$ such that the change of variables $\mathbf{x} = Q\mathbf{u} + \mathbf{t}$ satisfies

$$f(\mathbf{x}) = f(Q\mathbf{u} + \mathbf{t}) = a + bu^2 + cv^2 \quad \text{for some } a, b, c \in \mathbb{R}^2.$$

(b) Graph the equation $f(x, y) = 0$ in the u, v-plane. [Hint: It's a hyperbola.]

(c) Graph the equation $f(x, y) = 0$ in the x, y-plane. [Hint: Your matrix Q is a rotation matrix R_θ. The trace of a rotation matrix is $\mathrm{tr}(R_\theta) = 2\cos\theta$.]

2. Gambler's Ruin. Two gamblers have \$3 between them. Every time a game is played, Player A gives \$1 to Player B with probability 2/3 and Player B gives \$1 to Player A with probability 1/3. The game continues until one of the players goes broke.

(a) Suppose that each player starts with a whole number of dollars. Then this game is a Markov chain with four states: $(0, 3), (1, 2), (2, 1), (3, 0)$. Write down the 4×4 transition matrix in partitioned form:

$$M = \left(\begin{array}{c|c} I & R \\ \hline O & Q \end{array}\right).$$

[Hint: Order the states as $(0, 3), (3, 0), (1, 2), (2, 1)$. The ij entry of M is the probability that state j leads to state i. Each column of M should sum to 1.]

(b) Compute the limit of M^n as $n \to \infty$. [Hint: Use a result from Homework 2.]

(c) If the game begins in state $(1, 2)$, what is the probability that A will eventually win? What if the game starts in state $(2, 1)$?

3. Diagonalization of the Transpose. Consider a square $n \times n$ matrix A.

(a) Show that A and A^T have the same eigenvalues. [Hint: For any scalar λ we have $(\lambda I - A)^T = \lambda I^T - A^T = \lambda I - A^T$.]

(b) Find a 2×2 example showing that A and A^T need not have the same eigen**vectors**.

(c) Suppose that A has a basis of eigenvectors $\mathbf{x}_1, \ldots, \mathbf{x}_n$ with $A\mathbf{x}_i = \lambda_i \mathbf{x}_i$. Let X be the invertible matrix with columns $\mathbf{x}_1, \ldots, \mathbf{x}_n$ and let $\mathbf{y}_1, \ldots, \mathbf{y}_n$ be the columns of $(X^{-1})^T$. In this case show that $\mathbf{y}_1, \ldots, \mathbf{y}_n$ is a basis of eigenvectors of A^T with $A^T\mathbf{y}_i = \lambda_i \mathbf{y}_i$. [Hint: Let Λ be the diagonal matrix with entries $\lambda_1, \ldots, \lambda_n$. The matrix equation $AX = X\Lambda$ encodes the n vector equations $A\mathbf{x}_i = \lambda_i \mathbf{x}_i$. What matrix equation encodes the vector equations $A\mathbf{y}_i = \lambda_i \mathbf{y}_i$?]

4. Equilibrium of a Markov Chain. Consider the following stochastic matrix:

$$M = \begin{pmatrix} 1/2 & 1/3 & 0 \\ 1/4 & 0 & 1/2 \\ 1/4 & 2/3 & 1/2 \end{pmatrix}.$$

(a) Show that $(1,1,1)$ is an eigenvector of M^T with eigenvalue 1. [Hint: This just expresses the fact that the entries in each column of M sum to 1.]

(b) Since M and M^T have the same eigenvalues, it follows from (a) that 1 is an eigenvalue of M. Find an eigenvector $M\mathbf{p} = \mathbf{p}$ whose entries sum to 1.

(c) The Perron-Frobenius Theorem says that the 1-eigenspaces of M and M^T are one dimensional, and that the other eigenvalues of M and M^T satisfy $|\lambda| < 1$. Then from the Jordan Form we can find an invertible matrix $P = \left(\mathbf{p} \mid \mathbf{q} \mid \mathbf{r} \right)$ satisfying

$$M = P \left(\begin{array}{c|cc} 1 & 0 & 0 \\ \hline 0 & \multicolumn{2}{c}{\multirow{2}{*}{M'}} \\ 0 & & \end{array} \right) P^{-1},$$

where $(M')^n \to O$ as $n \to \infty$. Use this equation to show that

$$M^n \to \mathbf{p} \begin{pmatrix} 1 & 1 & 1 \end{pmatrix} = \left(\mathbf{p} \mid \mathbf{p} \mid \mathbf{p} \right) \quad \text{as } n \to \infty.$$

[Hint: Let $\mathbf{x}$ be the first column of $(P^{-1})^T$. Take transposes in the given equation to show that $M^T\mathbf{x} = 1\mathbf{x}$. Since the 1-eigenspace of M^T is one dimensional it follows from part (a) that $\mathbf{x}$ is a scalar multiple of $(1,1,1)$. Now use the fact that $\mathbf{x}^T\mathbf{p}$ is the top left entry of $P^{-1}P = I$ to show that $\mathbf{x} = (1,1,1)$.]

5. Dynamical Systems. Every real 2×2 matrix is similar to one of the following:

$$A = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}, \quad B = \begin{pmatrix} a & b \\ 0 & a \end{pmatrix}, \quad C = \begin{pmatrix} a & -b \\ b & a \end{pmatrix},$$

where a and b are real numbers.

(a) Compute the eigenvalues of A, B, C.

(b) Compute the exponential matrices $\exp(At)$, $\exp(Bt)$, $\exp(Ct)$. [Hint: Each of B and C can be written as a sum of commuting matrices:

$$B = \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix} + \begin{pmatrix} 0 & b \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad C = \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix} + \begin{pmatrix} 0 & -b \\ b & 0 \end{pmatrix}.]$$

(c) Let $\mathbf{x}(t) = (x(t), y(t))$. Use (a) to solve the dynamical systems $\mathbf{x}'(t) = A\mathbf{x}(t)$, $\mathbf{x}'(t) = B\mathbf{x}(t)$ and $\mathbf{x}'(t) = C\mathbf{x}(t)$, with initial condition $\mathbf{x}(0) = (x_0, y_0)$. Only use real numbers in your answers.

6. Jordan Blocks. Consider a $k \times k$ *Jordan block matrix* of the form

$$J = \begin{pmatrix} \lambda & 1 & & & \\ & \lambda & \ddots & & \\ & & \ddots & \ddots & \\ & & & \lambda & 1 \\ & & & & \lambda \end{pmatrix} \quad \text{for some scalar } \lambda,$$

where the blank entries are zero. We can also write $J = \lambda I + N$, where N is the $k \times k$ matrix with 1s on the superdiagonal and zeroes elsewhere:

$$N = \begin{pmatrix} 0 & 1 & & & \\ & 0 & \ddots & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ & & & & 0 \end{pmatrix}.$$

(a) Find an explicit formula for the nth powers N^n. Note that $N^n = O$ when $n \geq k$. [Hint: Work out the case $k = 4$ by hand. You will see a pattern.]

(b) Assuming $n \geq k$, use part (a) to show that

$$J^n = \begin{pmatrix} \lambda^n & \binom{n}{1}\lambda^{n-1} & \binom{n}{2}\lambda^{n-2} & \cdots & \binom{n}{k-1}\lambda^{n-k+1} \\ & \lambda^n & \binom{n}{1}\lambda^{n-1} & & \vdots \\ & & \ddots & \ddots & \vdots \\ & & & \lambda^n & \binom{n}{1}\lambda^{n-1} \\ & & & & \lambda^n \end{pmatrix}.$$

[Hint: Since λI and N commute, the binomial theorem gives $(\lambda I + N)^n = \sum_i \binom{n}{i}\lambda^{n-i}N^i$.]

(c) Use part (a) to show that

$$\exp(Jt) = \begin{pmatrix} e^{\lambda t} & te^{\lambda t} & \frac{t^2}{2}e^{\lambda t} & \cdots & \frac{t^{k-1}}{(k-1)!}e^{\lambda t} \\ & e^{\lambda t} & te^{\lambda t} & & \vdots \\ & & \ddots & \ddots & \vdots \\ & & & e^{\lambda t} & te^{\lambda t} \\ & & & & e^{\lambda t} \end{pmatrix}.$$

[Hint: Since λI and N commute, we have $\exp(Jt) = \exp(I\lambda t)\exp(Nt)$.]

7. The Rayleigh Quotient. Given a real $m \times n$ matrix A we want to find a nonzero vector $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$ that maximizes the *Rayleigh quotient* $r(\mathbf{x}) := \|A\mathbf{x}\|^2/\|\mathbf{x}\|^2$.

(a) Consider the symmetric $n \times n$ matrix $S = A^T A$ with ij entry s_{ij}. Show that

$$\|A\mathbf{x}\|^2 = \mathbf{x}^T S\mathbf{x} = \sum_{i,j} s_{ij}x_i x_j.$$

(b) Use part (a) to show that

$$\frac{\partial r}{\partial x_i} = \frac{2\|\mathbf{x}\|^2(S\mathbf{x})_i - 2\|A\mathbf{x}\|^2 x_i}{\|\mathbf{x}\|^4},$$

where $(S\mathbf{x})_i = \sum_j s_{ij}x_j$ is the ith entry of the vector $S\mathbf{x} \in \mathbb{R}^n$.

(c) Use part (b) to show that the gradient vector $\nabla r(\mathbf{x})$ is zero if and only if

$$A^T A\mathbf{x} = r(\mathbf{x})\mathbf{x}.$$

Since $A^T A$ has real non-negative eigenvalues, it follows that the maximum value of $r(\mathbf{x})$ is the largest eigenvalue of $A^T A$, and the vectors $\mathbf{x}$ that maximize $r(\mathbf{x})$ are the corresponding eigenvectors.

8. Total Least Squares. Consider the following three data points in the x, y-plane:

$$X = \left(\mathbf{x}_1 \mid \mathbf{x}_2 \mid \mathbf{x}_3 \right) = \left(\begin{array}{c|c|c} -1 & 0 & 1 \\ -4/3 & 5/3 & -1/3 \end{array} \right).$$

[Note: This data is centered at the origin: $\mathbf{x}_1 + \mathbf{x}_2 + \mathbf{x}_3 = \mathbf{0}$.] Given a vector $\mathbf{a} \in \mathbb{R}^2$, the orthogonal distance from data point $\mathbf{x}_i$ to the line $t\mathbf{a}$ is $\|(I - P_{\mathbf{a}})\mathbf{x}_i\|$, where $P_{\mathbf{a}} = \mathbf{a}\mathbf{a}^T/\mathbf{a}^T\mathbf{a}$ is the matrix that projects onto $t\mathbf{a}$. Our goal is to minimize the sum of the squared distances:

$$\sum \|(I - P_{\mathbf{a}})\mathbf{x}_i\|^2.$$

(a) Show that
$$\sum \|\mathbf{x}_i\|^2 = \sum \|P_{\mathbf{a}}\mathbf{x}_i\|^2 + \sum \|(I - P_{\mathbf{a}})\mathbf{x}_i\|^2.$$
[Hint: Show that $\mathbf{x}_i$ is the sum of orthogonal vectors $P_{\mathbf{a}}\mathbf{x}_i$ and $(I - P_{\mathbf{a}})\mathbf{x}_i$.] Since the sum on the left side is fixed by the data, we want to find $\mathbf{a}$ to maximize $\sum \|P_{\mathbf{a}}\mathbf{x}_i\|^2$.

(b) Show that
$$\sum_i \|P_{\mathbf{a}}\mathbf{x}_i\|^2 = \frac{\|X^T\mathbf{a}\|^2}{\|\mathbf{a}\|^2}.$$

(c) It follows from the problem on the Rayleigh quotient that $\sum_i \|P_{\mathbf{a}}\mathbf{x}_i\|^2$ is maximized when $\mathbf{a}$ is an eigenvector for the largest eigenvalue of XX^T. Use this to find $\mathbf{a}$. Sketch the data points and the best fit line.

9. Euler's Rotation Theorem. Let A be a real 3×3 matrix satisfying $A^TA = I$ and $\det(A) = 1$. We will show that A is a rotation.

(a) Show that 1 is an eigenvalue of A. [Hint: The characteristic polynomial factors as $(x - \lambda_1)(x - \lambda_2)(x - \lambda_3)$ for some $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{C}$, not necessarily distinct. Since A is real, the non-real eigenvalues come in conjugate pairs. Since $A^TA = I$ we have $|\lambda_j| = 1$ for all j. And since $\det(A) = 1$ we have $\lambda_1\lambda_2\lambda_3 = 1$.]

(b) From part (a) there exists a real unit vector $\mathbf{u} \in \mathbb{R}^3$ satisfying $A\mathbf{u} = \mathbf{u}$. Let $\mathbf{v}, \mathbf{w} \in \mathbb{R}^3$ be any orthonormal basis for the plane $\mathbf{u}^\perp$, so that $\mathbf{u}, \mathbf{v}, \mathbf{w}$ is an orthonormal basis for $\mathbb{R}^3$. Show that $A\mathbf{v}$ and $A\mathbf{w}$ are also in the plane $\mathbf{u}^\perp$. [Hint: Show that $(A\mathbf{x})^T(A\mathbf{y}) = \mathbf{x}^T\mathbf{y}$ for any vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^3$. Now put $\mathbf{x} = \mathbf{u}$ and $\mathbf{y} = \mathbf{v}, \mathbf{w}$.]

(c) It follows from (b) that
$$A = Q \left(\begin{array}{c|cc} 1 & 0 & 0 \\ \hline 0 & \multicolumn{2}{c}{\multirow{2}{*}{A'}} \\ 0 & & \end{array} \right) Q^{-1}.$$
Where $Q^TQ = I$ is the real orthogonal matrix with columns $\mathbf{u}, \mathbf{v}, \mathbf{w}$, and A' is some real 2×2 matrix. Show that $(A')^TA' = I$ and $\det(A') = 1$. [Hint: Show that $Q^{-1}AQ$ has these properties.]

(d) Finally, show that
$$A' = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$
for some angle θ. [Hint: From (c) the columns of A' are an orthonormal basis for $\mathbb{R}^2$.] It follows that A is a rotation around the vector $\mathbf{u}$ by angle θ. In retrospect, we could have found the angle right away because $\mathrm{tr}(A) = 1 + 2\cos\theta$.

(e) The following matrix satisfies $A^TA = I$ and $\det(A) = 1$, hence it is a rotation:[1]
$$A = \frac{1}{7}\begin{pmatrix} 3 & 2 & 6 \\ -6 & 3 & 2 \\ -2 & -6 & 3 \end{pmatrix}.$$
Find the axis $\mathbf{u}$ and the angle of rotation θ.

10. The Moore-Penrose Pseudoinverse. Two $m \times n$ matrices A and B are called *pseudoinverses* when the following properties hold:

[1]Where did I find an orthogonal matrix with rational entries? If $S^T = -S$ then one can check that $A = (I - S)^{-1}(I + S)$ satisfies $A^TA = I$ and $\det(A) = 1$. Furthermore, if S has rational entries then A has rational entries. This construction is called the *Cayley transform*.

(1) $ABA = A$ and $BAB = B$,
(2) $AB = B^*A^*$ and $BA = A^*B^*$.

(a) Suppose that A, B are pseudoinverses, and A, C are pseudoinverses. In this case, prove that $B = C$, hence A has a unique pseudoinverse, which we call A^+. [Hint: Use properties (1) and (2) to show that B and C are both equal to CAB. The proofs begin with $B = BAB = A^*B^*B = \cdots$ and $C = CAC = CC^*A^* \cdots$.]
(b) If A has independent columns, show that $A^+ = (A^*A)^{-1}A^*$. If A has independent rows, show that $A^+ = A^*(AA^*)^{-1}$. If A is square and invertible, show that $A^+ = A^{-1}$.
(c) Let A be an $m \times n$ matrix of rank r, and consider a singular value decomposition $A = U\Sigma V^*$. This means that:
  - The matrices U, Σ, V have shapes $m \times r$, $r \times r$, $n \times r$.
  - The matrices U and V satisfy $U^*U = V^*V = I_r$.
  - The matrix Σ is diagonal and invertible.

  In this case, prove that
$$A^+ = V\Sigma^{-1}U^*.$$

More ideas...

Diagonalization of Companion Matrices. Application to differential equations.

Vandermonde Matrices. Inspired by companion matrices.

The Fourier Matrix. Inspired by Vandermonde matrices. I still feel bad that we didn't have time for discrete Fourier transforms.

Real and Imaginary Parts of a Normal Operator.

Simultaneous Diagonalization. Commuting diagonalizable operators are simultaneously diagonalizable.