18.06 Spring 2018 Final - Solutions

1 (16 pts.)

The matrix A has a full SVD computed with Julia, where

$$A = \begin{pmatrix} 2 & 15 & 5 & 0 & 16 \\ 2 & 16 & 4 & 2 & 12 \\ 4 & 39 & 1 & 18 & -4 \end{pmatrix}.$$

The result is $A = U\Sigma V^T$ where

$$U = \begin{pmatrix} -0.341643 & -0.713606 & -0.611593 \\ -0.362087 & -0.500572 & 0.786334 \\ -0.867279 & 0.490095 & -0.0873704 \end{pmatrix}$$

$$\Sigma = \begin{pmatrix} 48.46518677202946 & 0 & 0 & 0 & 0 \\ 0 & 21.520354810093167 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

$$V = \begin{pmatrix} -0.10062 & -0.0217458 & 0.852961 & 0.292097 & -0.420167 \\ -0.923177 & 0.0186082 & 0.0386047 & -0.380465 & 0.0339893 \\ -0.0830254 & -0.236066 & 0.309187 & 0.29878 & 0.867475 \\ -0.33705 & 0.363403 & -0.343575 & 0.793231 & -0.0841173 \\ -0.130861 & -0.900773 & -0.239433 & 0.226805 & -0.25043 \end{pmatrix}$$

1.(a) (4 pts.) We wish to know which vectors in U and V correspond to the nullspace, the column space, the left nullspace, and the row space. Draw a box or circle around the vectors in each of these subspaces and label your box/circle.

1. (b) (2 pts.) What is the rank r of A?

1. (c) (5 pts.) Find the complete solution to $Ax = u_2$ where u_2 is the second column of U. (Okay to use symbols such as σ_1, u_1, or v_1. Be sure to have the complete solution, not just a solution.)

1. (d) (5 pts.) Ideally without precious time wasting computation, what is the determinant of the first three columns (reproduced below) of A? Justify your answer.

$$\begin{pmatrix} 2 & 15 & 5 \\ 2 & 16 & 4 \\ 4 & 39 & 1 \end{pmatrix}$$

Solution:

1. (a) There are two non-zero singular values and so this matrix has rank $r = 2$. The first two columns of U are therefore a basis for the column space, while the first two columns of V are a basis for the row space. The remaining one column of U is as basis for the left nullspace, while the remaining three columns of V are a basis for the nullspace.

1. (b) As stated, there are two non-zero singular values and so this matrix has rank $r = 2$.

1. (c) We can use the SVD to write A in rank-r form, i.e.

$$A = \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T$$

We can therefore identify that $x = v_2/\sigma_2$ is a particular solution, since

$$A\frac{v_2}{\sigma_2} = \sigma_1 u_1 v_1^T \left(\frac{v_2}{\sigma_2}\right) + \sigma_2 u_2 v_2^T \left(\frac{v_2}{\sigma_2}\right) = u_2 v_2^T v_2 = u_2$$

We can then add to this particular solution any multiple of independent vectors from the nullspace. The complete solution is therefore

$$\boxed{x = \frac{v_2}{\sigma_2} + c_1 v_3 + c_2 v_4 + c_3 v_5}$$

where c_1, c_2 and c_3 are arbitrary real constants.

1. (d) Since A is rank 2, any subset of its rows/columns can contain at most two linearly independent vectors. Therefore the three rows of A must be linearly dependent and so

$$\begin{vmatrix} 2 & 15 & 5 \\ 2 & 16 & 4 \\ 4 & 39 & 1 \end{vmatrix} = 0$$

2 (11 pts.) For 2(a) and 2(b) compute the gradient with respect to A (ideally without matrix elements or indices). It may be helpful to remember that trace(X)= trace(X^T)=trace($X^T I$) and trace (XY)=trace(YX).

2. (a) (3 pts.) $f(A) = \text{trace}(\frac{1}{2}A^T A)$, where the matrices are $n \times n$.

2. (b) (3 pts.) $f(A) = \text{trace}(A)$.

2. (c) (5 pts.) Compute $d(\exp.(A^3))$ in terms of A and dA. The dot indicates that we are taking the exponential e^x of every entry (not the matrix exponential). A^3 is the usual matrix multiplication of A times A times A.

Solution:

2. (a) We recall that the gradient ∇f of a scalar valued function of matrices may be found via the identity

$$df = \text{trace}((\nabla f)^T dA)$$

Applying the product rule to $f(A)$ we have

$$df = \text{trace}(\frac{1}{2}(dA)^T A) + \text{trace}(\frac{1}{2}A^T(dA)) = \text{trace}(\frac{1}{2}A^T dA) + \text{trace}(\frac{1}{2}A^T(dA)) = \text{trace}(A^T dA)$$

and so

$$\nabla f = A$$

2. (b) Using the same as the above, we have

$$df = \text{trace}(dA) = \text{trace}(I^T dA)$$

and so

$$\nabla f = I$$

2. (c) If $h(x)$ is a scalar valued function of scalars, and $h.A$ is this scalar function applied elementwise to the components of A, then

$$dh = h'. * dA,$$

where $.*$ denotes elementwise multiplication. In this, our scalar function $h(x) = \exp(x)$ is being applied elementwise to the components of A^3, and so

$$dh = \exp.(A^3). * d(A^3) = \exp.(A^3). * (A^2(dA) + A(dA)A + (dA)A^2)$$

3 (17 pts.)

Set up an idealized version of Bluebikes with bicycle stations in Allston, Boston, and Cambridge where on any given day, a bicycle has a 40% chance of remaining in the same city every night. There is a 30% chance of going to each of the other two cities. (Thus for example a bicycle starting in Cambridge has a 30% chance of ending up in Boston and a 30% chance of ending up in Allston.)

3. (a) (5 pts.) Write down the relevant Markov Matrix.

3. (b) (2 pts.) Is your matrix in 3(a) diagonalizible?

3. (c) (5 pts.) What are the eigenvalues of your matrix in 3(a)?

3. (d) (5 pts.) Suppose that on the first day 90% of the bicycles are in Cambridge, 1% in Allston, and 9% in Boston. What percentage of bicycles would you estimate would be in Cambridge once the bicycles have reached steady state?

Solution:

3. (a) The Markov matrix in this case is

$$M = \begin{pmatrix} 0.4 & 0.3 & 0.3 \\ 0.3 & 0.4 & 0.3 \\ 0.3 & 0.3 & 0.4 \end{pmatrix}$$

3. (b) M is symmetric and therefore it is necessarily diagonalizable.

3. (c) Since M is a positive Markov matrix, it has exactly one eigenvalue equal to 1. If we write down $M - \lambda I$, we have

$$M - \lambda I = \begin{pmatrix} 0.4-\lambda & 0.3 & 0.3 \\ 0.3 & 0.4-\lambda & 0.3 \\ 0.3 & 0.3 & 0.4-\lambda \end{pmatrix}$$

We immediately can see that choosing $\lambda = 0.1$ will produce a singular eigenvalue. Finally, the trace of M is 1.2, and so we can see that $\lambda = 0.1$ must be a repeated eigenvalue.

3. (d) The steady state vector is the vector corresponding to the eigenvalue $\lambda = 1$. Since M

is symmetric we can identify that

$$\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

is an eigenvector of M corresponding to $\lambda = 1$. However, the components of the steady state vector must sum to 1 because of conservation of probability. Hence in the long time limit, the probability vector will tend to

$$\begin{pmatrix} 1/3 \\ 1/3 \\ 1/3 \end{pmatrix}$$

and so $\sim 33\%$ of the bikes will be in Cambridge in the steady state limit.

4 (12 pts.)

Give a brief and convincing argument for each statement. (Not an example.)

4. (a) (3 pts.) The 2019^{th} power of a symmetric matrix is symmetric.

4. (b) (3 pts.) The 2019^{th} power of a positive definite matrix is positive definite (at least from the pure mathematics viewpoint).

4. (c) (3 pts.) The 2019^{th} power of a permutation matrix is a permutation matrix.

4. (d) (3 pts.) The 2019^{th} power of an orthogonal matrix is orthogonal.

Solution:

4. (a) If A is symmetric then $A = A^T$. Then

$$(A^{2019})^T = (AA...A)^T = A^T A^T ... A^T = AA...A = A^{2019}$$

and so A^{2019} is symmetric.

4. (b) If A is positive definite then all of its eigenvalues are strictly positive. If $\lambda > 0$ is an eigenvalue of A then $\lambda^{2019} > 0$ is an eigenvalue of A^{2019} and so all the eigenvalues of A^{2019} are positive and so A^{2019} is positive definite.

4. (c) A permutation matrix P acts to permute the elements of a vector v so that Pv has the same components of v but reordered. Applying P to the vector v 2019 times will continue to permute the elements of v and so P^{2019} will also be a permutation matrix.

4. (d) An orthogonal matrix obeys $Q^T Q = I$. Notice that

$$(Q^{2019})^T (Q^{2019}) = (Q^T Q^T ... Q^T)(QQ...Q) = I$$

and so Q^{2019} is also an orthogonal matrix.

5 (12 pts.)

The 5×5 matrix A has a QR decomposition where R is the diagonal matrix with $1, 2, 3, 4, 5$ on the diagonal, and 0 off the diagonal.

5. (a) (3 pts.) If a_5 is the fifth column of A, what is $\|a_5\|$?

5. (b) (3 pts.) If a_4 is the fourth column of A, what is the inner product of a_4 and a_5. (Reminder this means compute $a_4^T a_5$) Explain your answer.

5. (c) (3 pts.) The linear transformation that takes x to Ax takes the five dimensional unit cube to a parallelopiped. What is the unsigned volume of the image of the five dimensional unit cube? (Remember this means the absolute value of the volume of the parallelopiped determined by the columns of A.)

5. (d) (1 pt.) Pick the best one (without explanation) of {must be, might be, can't be}: The matrix A ________________ orthogonal.

5. (e) (1 pt.) Pick the best one (without explanation) of {must be, might be, can't be}: The matrix A ________________ (symmetric) positive definite.

5. (f) (1 pt.) Pick the best one (without explanation) of {must be, might be, can't be}: The columns of Q ________________ semi-axes of the ellpsoid that is the image of the unit sphere under A.

Solution:

5. (a) Since R is a diagonal matrix, the column of A are parallel to the columns of Q, but multiplied by the entries on the diagonal of R. We can therefore deduce that

$$a_5 = 5q_5 \implies \|a_5\| = 5\|q_5\|.$$

Since Q is an orthogonal matrix, all of its columns are normalized so that $\|q_i\| = 1$, and so

$$\boxed{\|a_5\| = 5}$$

5. (b) We know that $a_4 = 4q_4$ and $a_5 = 5q_5$, and so

$$a_4^T a_5 = 20 q_4^T q_5.$$

Since Q is orthogonal, its columns are mutually orthogonal and so $q_4^T q_5 = 0$, meaning that

$$\boxed{a_4^T a_5 = 0}$$

5.(c) We know that the volume of the unit cube is 1. The image of the unit cube under the transformation described by A will have unsigned volume $|\det A|$. However, we know from the QR factorization $A = QR$ that

$$\det A = \det Q \det R = \pm(5 \times 4 \times 3 \times 2 \times 1) = \pm 120,$$

and so the unsigned volume of the unit cube under the transformation is 120.

5. (d) The matrix A **can't** be orthogonal

5. (e) The matrix A **might** be positive definite

5. (f) The columns of Q **must be** semi-axes of the ellipsoid that is the image of the unit sphere under A.

6 (10 pts.)

6. (a) (6 pts.) Set up a matrix least squares problem if we are interested in taking n data points (x_i, y_i) for $i = 1, \ldots, n$ and we wish to find the best function $f(x) = c_1 \sin(x) + c_2 \cos(x) + c_3 \tan(x)$ through the data points. Note: Setting up a matrix least squares problem means setting up a matrix A and a vector b in a least squares equation such that $x = \hat{x}$ minimizes $\|Ax - b\|$.

Also write the solution to the least squares problem in terms of the compact SVD of your matrix.

6. (b) (4 pts.) Set up a matrix least squares problem if we are interested in taking n data points (x_i, y_i, z_i) in R^3, and we wish to fit a function $f(x, y) = c_1 e^{x+y} + c_2 \sin(x - y)$.

Solution

6. (a) (6 pts.) The least squares problem in this case is to minimize

$$\|Ax - b\|,$$

where

$$A = \begin{pmatrix} \sin x_1 & \cos x_1 & \tan x_1 \\ \sin x_2 & \cos x_2 & \tan x_2 \\ \vdots & \vdots & \vdots \\ \sin x_n & \cos x_n & \tan x_n \end{pmatrix}, \quad x = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix}, \quad b = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

The least squares solution is

$$\hat{x} = V \Sigma_r^{-1} U^T b$$

6.(b) The least squares problem in this case is to minimize

$$\|Ax - b\|,$$

where

$$A = \begin{pmatrix} e^{x_1+y_1} & \sin x_1 - y_1 \\ e^{x_2+y_2} & \sin x_2 - y_2 \\ \vdots & \vdots \\ e^{x_n+y_n} & \sin x_n - y_n \end{pmatrix}, \quad x = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}, \quad b = \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{pmatrix}$$

7 (4 pts.) The 4x4 symmetric matrix A satisfies $A^2 = A$ and has all four diagonal elements equal to 1/2.

The four roots to the equation $\det(A - \lambda I) = 0$ are not distinct. Allowing for multiple eigenvalues, they are ________, ________, ________, and ________. (Provide a brief explanation.)

Solution A symmetric matrix has real eigenvalues. If $A^2 = A$, then the only possible eigenvalues are 0 and 1. If all the diagonal elements are equal to 1/2 then $Trace(A) = 2$. The only possible sum of 0 and 1 giving a trace of 2 is $\lambda_1 = \lambda_2 = 0$ and $\lambda_3 = \lambda_4 = 1$.

8 (10 pts.)

There are six permutation matrices in R^3. Let P_1, P_2, P_3 be the three with determinant -1 (labelled arbitrarily.) The other three are I and P_4 and P_5.

8 (a) (4 pts.) Give an example of one permutation matrix with determinant -1, and one that is not the identity but has determinant +1.

8. (b) (6 pts.) Consider the matrix $1000P_1 + 800P_3 + 6P_5$. Give an eigenvalue, eigenvector pair for this matrix (ideally by not writing down the matrix).

Solution 8. (a) The 3×3 permutation matrices with determinant -1 are

$$P_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad P_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad P_3 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

The remaining 3×3 permutation matrices with determinant 1 are

$$P_4 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}, \quad P_5 = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

8. (b) All permutation matrices have an eigenvector $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ with eigenvalue 1. Therefore $M = 1000P_1 + 800P_3 + 6P_5$ has an eigenvector $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ with eigenvalue $1000 + 800 + 6 = 1806$.

9 (8 pts.)

Given n numbers $h_1, \dots, h_n$, a Hankel matrix is defined as a matrix H such that $H_{ij} = h_{|i-j|+1}$. They take the form

$$H = \begin{pmatrix} h_1 & h_2 & h_3 & \dots & h_{n-1} & h_n \\ h_2 & h_1 & h_2 & \dots & h_{n-2} & h_{n-1} \\ h_3 & h_2 & h_1 & \dots & h_{n-3} & h_{n-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \\ h_{n-1} & h_{n-2} & h_{n-3} & \dots & h_1 & h_2 \\ h_n & h_{n-1} & h_{n-2} & \dots & h_2 & h_1 \end{pmatrix}.$$

(a) (4 pts.) Are the set of Hankel matrices a vector subspace of symmetric $n \times n$ matrices? (Explain)

(b) (4 pts.) Find a basis for the vector space of Hankel matrices. What is the dimension?

Solution 9. (a) Consider two Hankel matrices

$$H = \begin{pmatrix} h_1 & h_2 & h_3 & \dots & h_{n-1} & h_n \\ h_2 & h_1 & h_2 & \dots & h_{n-2} & h_{n-1} \\ h_3 & h_2 & h_1 & \dots & h_{n-3} & h_{n-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \\ h_{n-1} & h_{n-2} & h_{n-3} & \dots & h_1 & h_2 \\ h_n & h_{n-1} & h_{n-2} & \dots & h_2 & h_1 \end{pmatrix}, \quad G = \begin{pmatrix} g_1 & g_2 & g_3 & \dots & g_{n-1} & g_n \\ g_2 & g_1 & g_2 & \dots & g_{n-2} & g_{n-1} \\ g_3 & g_2 & g_1 & \dots & g_{n-3} & g_{n-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \\ g_{n-1} & g_{n-2} & g_{n-3} & \dots & g_1 & g_2 \\ g_n & g_{n-1} & g_{n-2} & \dots & g_2 & g_1 \end{pmatrix}.$$

If we take an arbitrary linear combination $aH + bG$ where $a, b \in \mathbb{R}$ then we obtain

$$aH+bG = \begin{pmatrix} ah_1+bg_1 & ah_2+bg_2 & ah_3+bg_3 & \dots & ah_{n-1}+bg_{n-1} & ah_n+bg_n \\ ah_2+bg_2 & ah_1+bg_1 & ah_2+bg_2 & \dots & ah_{n-2}+bg_{n-2} & ah_{n-1}+bg_{n-1} \\ ah_3+bg_3 & ah_2+bh_2 & ah_1+bh_1 & \dots & ah_{n-3}+bg_{n-3} & ah_{n-2}+bg_{n-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \\ ah_{n-1}+bg_{n-1} & ah_{n-2}+bg_{n-2} & ah_{n-3}+bh_{n-3} & \dots & ah_1+bg_1 & ah_2+bg_2 \\ ah_n+bg_n & ah_{n-1}+bg_{n-1} & ah_{n-2}+bg_{n-2} & \dots & ah_2+bg_2 & ah_1+bg_1 \end{pmatrix}$$

which is a symmetric matrix in Hankel form. Therefore these form a vector subspace.

9. (b) A possible basis for this vector space is the set of matrices H_m, where the $h_i = 0$ except for $i = m$ where $h_m = 1$. This vector space has dimension n.