MIT 18.06 Final Exam, Fall 2018: Solutions

Problem 1 (5+10 points):

The matrix A has the diagonalization $A = X\Lambda X^{-1}$ with

$$X = \begin{pmatrix} 1 & 1 & -1 & 0 \\ & 1 & 2 & 1 \\ & & 1 & 0 \\ & & & 1 \end{pmatrix}, \quad \Lambda = \begin{pmatrix} 1 & & & \\ & 2 & & \\ & & -2 & \\ & & & -1 \end{pmatrix}.$$

(a) Give a basis for the nullspace $N(M)$ of the matrix $M = A^4 - 2A^2 - 8I$. (Hint: not much calculation required!)

(b) Write down the solution $x(t)$ to the ODE $\frac{dx}{dt} = Ax$ for $x(0) = \begin{pmatrix} 2 \\ 3 \\ 1 \\ -1 \end{pmatrix}$.

Your final answer should contain no matrix exponentials or matrix inverses, just a sum of vectors (whose components you give explicitly as numbers) multiplied by given scalar coefficients (that may depend on t).

Solution:

(a) The eigenvalues of A are $\lambda = 1, 2, -2, -1$, from Λ. The eigenvalues of M are then $\lambda^4 - 2\lambda^2 - 8$, with the same corresponding eigenvectors. M therefore has two zero eigenvalues (which come from the ± 2 eigenvalues of A), and so the corresponding eigenvectors are a basis for $N(M)$, i.e.

$$N(M) = \text{span of } \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 2 \\ 1 \\ 0 \end{pmatrix}.$$

(b) We know that the general solution to the ODE may be written as

$$x(t) = c_1 e^t \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + c_2 e^{2t} \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + c_3 e^{-2t} \begin{pmatrix} -1 \\ 2 \\ 1 \\ 0 \end{pmatrix} + c_4 e^{-t} \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}$$

where c_1, c_2, c_3 and c_4 are constants that are determined by the initial condition: $Xc = x(0)$. Since X is upper triangular, we can solve for c by backsubstitution:

$$\begin{aligned} c_1 + c_2 - c_3 &= 2, \\ c_2 + 2c_3 + c_4 &= 3, \\ c_3 &= 1, \\ c_4 &= -1. \end{aligned}$$

Backsubstitution gives $c_1 = 1, c_2 = 2, c_3 = 1$ and $c_4 = -1$.

Problem 2 (3+3+3+3+3 points):

The real $m \times n$ matrix A has a QR factorization $A = QR$ of the form

$$Q = \begin{pmatrix} q_1 & q_2 & q_3 & q_4 & q_5 & q_6 \end{pmatrix}, \; R = \begin{pmatrix} 1 & -2 & 2 & 0 & 0 & 0 \\ & 2 & -3 & 0 & 0 & 0 \\ & & 1 & 0 & 0 & 0 \\ & & & 3 & 1 & -1 \\ & & & & 1 & 2 \\ & & & & & 1 \end{pmatrix}.$$

where $q_1, \ldots, q_6$ are six orthonormal vectors in $\mathbb{R}^m$.

(a) Give as much true information as possible about m, n, and the rank of A.

(b) If a_5 is the 5th column of A, write it in the basis $q_1, \ldots, q_6$, i.e. write it as $a_5 = c_1 q_1 + c_2 q_2 + \cdots + c_6 q_6$, by giving the numerical values of the coefficients $c_1, \cdots, c_6$.

(c) What is $\|a_5\|$?

(d) This pattern of zero entries in R means that columns .... of A must be .... to columns .... of A.

(e) If A is a square matrix, what is $|\det A|$ (the absolute value of the determinant)?

Solution:

(a) Since R is 6×6 , A must have 6 columns so that $n = 6$. The column space of A is spanned by 6 orthonormal vectors, so the rank of A is 6. The number of rows must be greater than or equal to the rank, so $m \geq 6$.

(b) From the QR factorization we see that $a_5 = q_4 + q_5$: this is just Q multiplied by the 5th column of R.

(c) $\|a_5\|^2 = a_5^T a_5 = (q_4 + q_5)^T (q_4 + q_5) = \|q_4\|^2 + \|q_5\|^2 = 2$ (from orthonormality of the q's) and so $\|a_5\| = \sqrt{2}$.

(d) This pattern of zero entries in R means that columns 1–3 of A must be orthogonal to columns 4–6 of A. (Columns 1–3 of A are spanned by q_1, q_2, q_3, whereas columns 4–6 are spanned by q_4, q_5, q_6. Equivalently, to get that 3×3 block of zeros in the upper-right corner of R, a Gram–Schmidt process must have encountered dot products that were already zero when projecting columns 4–6 to be orthogonal to the first three columns.)

(e) By properties of the determinant, we know that $\det A = \det Q \det R$. Since Q is an orthogonal matrix, we know that $\det Q = \pm 1$. Since R is an upper triangular matrix, we know that $\det R$ is the product of the diagonal elements of R, i.e. $\det R = 6$. Hence, $|\det A| = 6$.

Problem 3 (8+3+4 points):

You are given the matrix

$$A = \begin{pmatrix} 1 & 2 & 1 & 0 \\ 2 & 5 & 1 & 1 \\ 0 & 1 & 1 & 3 \end{pmatrix}$$

(a) Give bases for the four fundamental subspaces of A.

(b) Which of $Ax = b$ or $A^T y = c$ will *not* have *unique* solutions for x or y (assuming a solution exists)?

(c) Which of $Ax = b$ or $A^T y = c$ may *not have a solution*? For that equation, **give a right-hand side** (b or c) for which a solution exists, and that has only *two nonzero* entries in the right-hand side.

Solution:

(a) We can use row operations to put A into reduced row echelon form (rref):

$$A \to \begin{pmatrix} 1 & 0 & 0 & -5 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 1 \end{pmatrix}.$$

From here we can see that A is a rank 3 matrix: The column space of A is three dimensional, but the column space of A is a subspace of $\mathbb{R}^3$. So any three, linearly independent vectors in $\mathbb{R}^3$ will form a basis for $C(A)$. For example, we can use the standard basis, or we could use the columns of A that correspond to the pivot columns:

$$\begin{aligned} C(A) &= \text{span of } \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \\ &= \text{span of } \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 5 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}. \end{aligned}$$

The row space is also three dimensional. Since the row space is preserved by row operations, we can use the rows of the rref matrix (this is useful for part c):

$$C(A^T) = \text{span of } \begin{pmatrix} 1 \\ 0 \\ 0 \\ -5 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}.$$

The nullspace has dimension $4-3=1$. So there is a single special solution to $Ax=0$:

$$N(A) = \text{span of } \begin{pmatrix} 5 \\ -2 \\ -1 \\ 1 \end{pmatrix}.$$

Finally, the left nullspace has dimension $3-3=0$, and so is a trivial vector space only containing the zero vector, i.e.

$$N(A^T) = \left\{\vec{0}\right\}$$

(b) Since $\dim(N(A)) > 0$, $Ax=b$ does not have a unique solution: we can add any nonzero vector from the nullspace to a particular solution of the system to get a different solution.

(c) Since $C(A^T) \neq \mathbb{R}^4$, $A^T y = c$ may not have a solution. In fact, it will only have a solution if $c \in C(A^T)$. To give an example of such a c with only two nonzero entries, we can use any of the three basis vectors we wrote down in part (a), e.g.

$$c = \begin{pmatrix} 1 \\ 0 \\ 0 \\ -5 \end{pmatrix}.$$

Alternatively, we can use the fact that $C(A^T) = N(A)^{\perp}$: we only have a solution if $c \perp \begin{pmatrix} 5 \\ -2 \\ -1 \\ 1 \end{pmatrix}$, where we used the $N(A)$ basis from (a). So, for example, if we look for a c of the form $\begin{pmatrix} c_1 \\ c_2 \\ 0 \\ 0 \end{pmatrix}$, we get the condition $5c_1 - 2c_2 = 0$. For example, we could use $c = \begin{pmatrix} 2 \\ 5 \\ 0 \\ 0 \end{pmatrix}$.

Problem 4 (5+10 points):

A vector $\hat{x}$ minimizes $\|Ax - b\|$ over all possible vectors x.

(a) If $\hat{x} = 0$, then b must be in which fundamental subspace of A?

(b) If $A = \begin{pmatrix} 1 & 2 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$ and $b = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$, what is the minimum possible $\|Ax - b\|$? Describe (with as much detail as you can) *all possible* $\hat{x}$ that give this minimum.

Solution:

(a) If $\hat{x}$ minimizes $\|Ax - b\|$ over all possible vectors x, then $A^T A\hat{x} = A^T b$. If $\hat{x} = 0$, then $A^T b = 0$, and so $b \in N(A^T)$. Another way to see this is that if $\hat{x} = 0$, then the projection of b onto the column space of A is the zero vector, which means that b must be orthogonal to $C(A)$, i.e. $b \in N(A^T)$.

(b) The minimum possible $\|Ax - b\|$ occurs when $p = Ax$ is the projection of b onto the column space of A. The column space of A is one dimensional, and is spanned by the vector $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, and so we can compute this projection easily using the projection formula:

$$p = \frac{\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}^T \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}}{\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}^T \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

So the minimum value of $\|Ax - b\|$ is given by $\left\| \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\| = \sqrt{2}$.

To find a particular $\hat{x}$ that gives this minimum, we firstly find an $\hat{x}$ that satisfies $A\hat{x} = p$, for instance $\hat{x} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$. We can then add to this any multiple of a vector in $N(A)$, which is spanned by $\begin{pmatrix} 2 \\ -1 \end{pmatrix}$. The most general $\hat{x}$ then takes the form:

$$\hat{x} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} + c \begin{pmatrix} 2 \\ -1 \end{pmatrix},$$

where c is some constant.

Problem 5 (5+10 points):

Suppose that $A = A^T$ is a real-symmetric 10×10 matrix whose eigenvalues are $11, 10, 9, 8, 7, 6, 5, 4, 3, 2$. Corresponding eigenvectors of A, each normalized to have length $\|x\| = 1$, are (to three significant digits) the columns of the matrix:

$$X = \begin{pmatrix} -0.179 & 0.415 & -0.173 & 0.115 & -0.125 & 0.079 & -0.676 & -0.066 & 0.3 & -0.423 \\ -0.239 & -0.168 & -0.53 & -0.08 & 0.232 & -0.281 & -0.386 & -0.223 & -0.353 & 0.413 \\ -0.35 & 0.116 & -0.201 & -0.41 & 0.547 & 0.197 & 0.182 & 0.373 & 0.38 & 0.029 \\ -0.399 & -0.238 & -0.079 & -0.46 & -0.279 & -0.139 & 0.259 & -0.427 & -0.028 & -0.468 \\ -0.378 & -0.132 & 0.013 & 0.28 & -0.374 & -0.332 & 0.096 & 0.056 & 0.582 & 0.4 \\ -0.227 & 0.352 & 0.195 & 0.336 & 0.388 & 0.137 & 0.242 & -0.657 & 0.042 & 0.109 \\ -0.404 & 0.327 & 0.566 & -0.143 & 0.024 & -0.399 & -0.123 & 0.277 & -0.371 & 0.03 \\ -0.001 & 0.628 & -0.369 & -0.162 & -0.46 & 0.172 & 0.3 & 0.05 & -0.202 & 0.261 \\ -0.303 & -0.07 & -0.331 & 0.6 & 0.07 & -0.052 & 0.288 & 0.329 & -0.29 & -0.389 \\ -0.425 & -0.285 & 0.195 & 0.047 & -0.213 & 0.732 & -0.194 & 0.051 & -0.194 & 0.195 \end{pmatrix}.$$

(That is, the first column of X is an eigenvector for $\lambda = 11$, and so on.)

Consider the recurrence relation $x_{n+1} - x_n = -\alpha A x_n$ on vectors $x_n \in \mathbb{R}^{10}$, where $\alpha > 0$ is some positive real number.

(a) For what values of α (if any) can the recurrence have solutions that *diverge* in magnitude as $n \to \infty$?

(b) For $\alpha = 1$, give a good approximation for the vector x_{100} given $x_0 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$. You can leave your answer in the form of some vector times some coefficient(s) without carrying out the multiplications, but give all the numbers in your coefficients and vectors to 3 significant digits.

Solution:

(a) We can rewrite the recurrence relation as $x_{n+1} = (I - \alpha A)x_n$. The solutions x_n can diverge provided $I - \alpha A$ has at least one eigenvalue with *absolute value* greater than 1. The eigenvalues of $I - \alpha A$ are $1 - \alpha\lambda$, where λ is an eignvalue of A. Since $\alpha > 0$, and all the eigenvalues are positive, divergence can first occur when $1 - 11\alpha < -1$, i.e. when $\alpha > \frac{2}{11}$.

(b) Suppose $\alpha = 1$. We know that $x_{100} = (I - A)^{100} x_0$. We can write any x_0 in terms of the eigenvector basis as $x_0 = \sum_{i=1}^{10} c_i v_i$, where v_i is the

ith column of X. The eigenvalue of $(I - A)$ with largest absolute value is $1 - 11 = -10$, with eigenvector v_1, and so $x_{100} \approx c_1(-10)^{100}v_1$. Since the eigenvectors are orthogonal (A is real-symmetric with distinct λ's), we can compute c_1 using projection:

$$c_1 = \frac{v_1^T x_0}{v_1^T v_1} = -0.179,$$

(where $v_1^T v_1 = 1$ since you were told that the columns are normalized) and so:

$$x_{100} \approx -0.179 \times (-10)^{100} v_1 = 0.179 \times 10^{100} \begin{pmatrix} 0.179 \\ 0.239 \\ 0.35 \\ 0.399 \\ 0.378 \\ 0.227 \\ 0.404 \\ 0.001 \\ 0.303 \\ 0.425 \end{pmatrix}.$$

(Indeed, we could write down the "exact" x_{100} by including all of the other eigenvectors. But you actually *gain no accuracy* by doing so: you were only given the eigenvectors to three significant digits, but 9^{100} is more than 10^4 times smaller than 10^{100}, so all of the "corrections" from the other eigenvectors are too small to be distinguished from the unknown digits of our coefficients.)

Problem 6 (5+10 points):

You are trying to fit of a sequence of four data points $(t_k, y_k) = (1, 6), (4, 7), (9, 12), (16, 14)$ to a function of the form $y(t) = \alpha + \beta\sqrt{t}$ for unknown coefficients α and β .

(a) Write down a minimization problem of the form "minimize .... over all possible α, β" that we could solve for a "best fit" curve using 18.06 techniques. (Don't solve it.)

(b) Write down a 2×2 system of equations of the form

$$\begin{pmatrix} \text{some} \\ \text{matrix} \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} \text{some} \\ \text{vector} \end{pmatrix}$$

whose solution gives the "best" fit coefficients α and β for your minimization problem from part (a). You don't need to solve it! You can leave the matrix and vector as a product of other matrices/vectors/transposes, but give the numerical values of each of the terms (no matrix inverses allowed).

Solution:

(a) The thing we know how to do in 18.06 is to find the least-squares fit of our data. For a function of the form $y(t) = \alpha + \beta\sqrt{t}$, we therefore want to minimize

$$\sum_{k=1}^{4}(y_k - \alpha - \beta\sqrt{t_k})^2,$$

over all (real) $\alpha$and β.

(b) We can write the minimization problem in matrix form as

$$\min_x \|Ax - b\|^2,$$

where

$$A = \begin{pmatrix} 1 & \sqrt{1} \\ 1 & \sqrt{4} \\ 1 & \sqrt{9} \\ 1 & \sqrt{16} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{pmatrix}, \quad x = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}, \quad b = \begin{pmatrix} 6 \\ 7 \\ 12 \\ 14 \end{pmatrix}.$$

The solution to this minimization problem is then given by $\hat{x}$, where $\hat{x}$ satisfies the normal equations:

$$A^T A\hat{x} = A^T b.$$

Problem 7 (10 points):

Suppose that you want to multiply each of the three matrices

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \; B = \begin{pmatrix} 1 & 1 \\ -1 & 0 \\ 0 & 1 \end{pmatrix}, \; C = \begin{pmatrix} 4 & 5 \end{pmatrix},$$

by each of the two vectors $x, y \in \mathbb{R}^2$. That is, you want to compute the *six* vectors $x_A = Ax$, $x_B = Bx$, $x_C = Cx$, $y_A = Ay$, $y_B = By$, and $y_C = Cy$. Write *all* of these products in the form of a *single* matrix–matrix multiplication:

$$\begin{pmatrix} \text{some matrix in terms of} \\ x_A, x_B, x_C, y_A, y_B, y_C \end{pmatrix} = \begin{pmatrix} \text{some matrix in terms of} \\ A, B, C \end{pmatrix} \begin{pmatrix} x & y \end{pmatrix}.$$

Note that the third matrix here is the 2×2 matrix whose columns are x and y. That is, **give the sizes of the other two matrices and how their contents are arranged.**

[For example, a possible but wrong(!) answer for the second matrix would be the 2×6 matrix $\begin{pmatrix} A & B^T & C^T \end{pmatrix}$.]

Solution:

We can write all of these products as the following single matrix-matrix multiplication:

$$\begin{pmatrix} x_A & y_A \\ x_B & y_B \\ x_C & y_C \end{pmatrix} = \begin{pmatrix} A \\ B \\ C \end{pmatrix} \begin{pmatrix} x & y \end{pmatrix}.$$

The first and second matrices are both 6×2 matrices, while the third is a 2×2 matrix.