MIT 18.06 Advanced Standing Exam Solutions

August 2016

Problem 1:

(a) $A - \mu I$ is singular if and only if μ is an eigenvalue, so μ_0 must be one of $1, 2, 3, \ldots, 10$.

(b) The exact solution is $x = \sum_{i=1}^{10} \frac{x_i x_i^T b}{(i-\mu) x_i^T x_i}$, i.e. we project the right-hand side b on to the basis of eigenvectors, using the fact that the eigenvectors *must be* orthogonal (but not necessarily normalized) for a real-symmetric A with distinct eigenvalues. For $\mu = 1.001$, we would typically expect (without prior knowledge of b) that it will be dominated by the term for the eigenvalue 1 (where $|\lambda - \mu|$ is smallest), and the best approximation for x in terms of a *single* eigenvector is therefore

$$x \approx -1000 \frac{x_1 x_1^T b}{x_1 x_1}.$$

Of course, in the event that b happened to be nearly orthogonal to x_1, then the best approximation would be by some other eigenvector (depending on which one b had the biggest projection onto), but without knowing anything more about b the above answer is the most reasonable guess. (For a random b, the above choice has orders of magnitude smaller average error than the projection onto any other eigenvector of A.)

Problem 2:

(a) Because x has four components, A must be an $m \times 4$ matrix, i.e. $\boxed{n = 4}$. Because there are two degrees of freedom in the solution, the nullspace $N(A)$ must have dimension 2; since this is $n - r$, it means that the rank is $\boxed{r = 2}$. All we can say about m is that $m \geq r$, and hence $\boxed{m \geq 2}$.

(b) Let's choose $m = 2$ since that is the smallest. It is also easiest to write down an A in reduced-row-echelon form $A = (I\, F)$, so that we can read off the nullspace vectors $\begin{pmatrix} -F \\ I \end{pmatrix}$. Noticing that the lowest two rows of the nullspace vectors $\begin{pmatrix} -2 \\ 1 \\ 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}$ from above indeed form a block I of the 2×2 identity matrix (if we swap their order), that means that the first two rows give $-F$. This gives a possible A of

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

(It is easy to check that the rows of A are orthogonal to the nullspace vectors above.) This choice of A gives $\vec{b} = A\vec{x} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$.

Of course, there are many other possible solution techniques. A very systematic approach (the sort of approach you might use on a computer if you had to deal with many problems of this sort), is to (i) orthogonalize the two nullspace vectors with Gram–Schmidt and then (ii) take two arbitrary vectors and project them onto the subspace orthogonal to $N(A)$ in order to form the row vectors of A. (The annoyance of doing Gram–Schmidt by hand is all the square roots in the normalizations, but those can be avoided in hand calculations by rescaling.) In this particular case, we would start by adding the column vectors (1,0,0,1) to (-2,1,1,0) to make them orthogonal, obtaining an orthonormal basis $(1, 0, 0, 1)^T/\sqrt{2}$ and $(-1, 1, 1, 1)^T/2$ for $N(A)$, and then project e.g. the vectors $(4, 0, 0, 0)$ and $(0, 4, 0, 0)$ (choosing the 4's here to avoid fractions) to be orthogonal to these two vectors, yielding rows (1,1,1,-1) and (1,3,-1,-1) of A. Hence in this case b would be $(0, 2)^T$.

Another systematic approach, given two nullspace vectors that form a 4×2 matrix $N = \begin{pmatrix} B \\ C \end{pmatrix}$, would be to notice that it forms the nullspace of the reduced-row echelon matrix $A = (I,\, -BC^{-1})$. Then you could use use Gaussian elimation (or the explicit 2×2 formula) to find C^{-1}, at which point you multiply it by B to get A.

Finally, for such a small matrix, you can just use trial-and-error. For $\begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}$ to be in the nullspace, the first and last columns of A must be equal and opposite, and picking them to be $(1, 0)$ and $(-1, 0)$ is a simple possibility. After that it just requires a little experimentation with the middle two columns to make the rows orthogonal to the second nullspace vector.

Problem 3:

(a) Defining $\vec{u}_k = \begin{pmatrix} f_{k+1} \\ f_k \end{pmatrix}$, re-express the above recurrence as $\vec{u}_{k+1} = A\vec{u}_k$, and give the matrix A.

(b) This recurrence is $\vec{u}_{k+1} = A\vec{u}_k$ for $A = \begin{pmatrix} 4 & 3 \\ 1 & 0 \end{pmatrix}$.

(c) Using the formula for the 2×2 characteristic polynomial, we get $\det(A - \lambda I) = \lambda^2 - \lambda \operatorname{trace} A + \det A = 0 = \lambda^2 - 4\lambda - 3$, which has solutions $\lambda_\pm = 2 \pm \sqrt{2^2 + 3} = 2 \pm \sqrt{7}$. The largest $|\lambda|$ is $2 + \sqrt{7}$, and hence we expect that this sequence (which corresponds to multiplying by A over and over again) will approach the corresponding eigenvector as $k \to \infty$. If we write this eigenvector as the column vector $x = (r, 1)^T$, then we see that $Ax = \lambda x$ implies that $r = \lambda$ (from the second row). Since r is exactly the ratio f_{k+1}/f_k , it follows that $\boxed{f_{k+1}/f_k \to 2 + \sqrt{7}}$.

A good check is that the first row of $Ax = \lambda_+ x$ is also satisfied: $4(2 + \sqrt{7}) + 3 = (2 + \sqrt{7})(2 + \sqrt{7})$, which is true (both sides give $11 + 4\sqrt{7}$).

(d) The sequence above starts with $f_0 = f_1 = 1$, and $|f_k|$ grows rapidly with k. Keep $f_0 = 1$, but give a *different* value of f_1 that will make the sequence (with the *same recurrence* $f_{k+2} = 4f_{k+1} + 3f_k$) approach *zero* ($f_k \to 0$) as $k \to \infty$.

(e) The other eigenvalue is $\lambda_- = 2 - \sqrt{7}$ has magnitude $|\lambda_-| < 1$ (you don't need a calculator to know this, since $2 = \sqrt{4} < \sqrt{7} < \sqrt{9} = 3$). Hence, if we start with something proportional to *that* eigenvector, $\begin{pmatrix} 2 - \sqrt{7} \\ 1 \end{pmatrix}$ from above, then the solutions will decay as $\lambda_-^k \to 0$. If $f_0 = 1$, this means $f_1 = 2 - \sqrt{7}$.

(Of course, if you try this on a computer, the solutions will eventually blow up anyway, because rounding errors will eventually pull in a nonzero component of the other eigenvector.)

Problem 4:

(a) **True**. [*Non-required* proof: $Q^TQ = I$, so $\det Q^TQ = 1 = (\det Q^T)(\det Q) = (\det Q)^2$, and the result follows. Alternatively, all the eigenvalues of a unitary matrix have magnitude 1, and the determinant is the product of these.]

(b) **False**. You get a steady state for *multiplying* repeatedly by Markov matrices, but here we have an ODE so we are *exponentiating* them. Counter-example: consider the 1×1 Markov matrix $A = 1$. $du/dt = Au = u$ has the solution $u = u(0)e^t$, which obviously blows up.

(c) **True**.

(d) **False**. Counter-example: Let S and T be two distinct lines through the origin, say the x axis and the y axis. Their union (the x and y axes) is not a subspace, because if you add a point $(1, 0, 0)$ on the x axis with a point $(0, 1, 0)$ on the y axis, then you get a point $(1, 1, 0)$ which is in neither, so the space is not closed under addition.

(e) **False.** Counter-example: Suppose A is the $n \times n$ identity matrix, and B is an $n \times n$ matrix of zeros. Then AB is also a matrix of zeros, whose column space is the zero vector, which does *not* contain the column space $C(A) = \mathbb{R}^n$.

(f) **True.** [*Non-required* proof: If $y \in C(AB)$, then $y = ABx$ for some x, hence $y = Az$ for $z = Bx$, hence $y \in C(A)$. Hence $C(AB) \subseteq C(A)$.]

Problem 5:

(a) (The proof here is very similar to the proof of orthogonality of the eigenvectors of a real-symmetric matrix.) By definition, $Ax = \lambda_x x$ and $A^T y = \lambda_y y$. Multiply both sides of the Ax equation by y^T on the left. Then $y^T Ax = \lambda_x y^T x = (A^T y)^T x = \lambda_y y^T x$, and hence $(\lambda_x - \lambda_y) y^T x = 0$. If $\lambda_x \neq \lambda_y$, then $\lambda_x - \lambda_y \neq 0$ and we must have $y^T x = 0$.

(b) We have $A^T = \begin{pmatrix} 1 & \epsilon \\ 1 & 1 \end{pmatrix}$. If we write $y = \begin{pmatrix} u \\ 1 \end{pmatrix}$ for an unknown u, then from the second row of $A^T y = \lambda y$ we get $u + 1 = \lambda$. Since $\lambda_\pm = 1 \pm \sqrt{\epsilon}$ (the same as A), we immediately get the corresponding eigenvectors to be:

$$y_\pm = \begin{pmatrix} \pm\sqrt{\epsilon} \\ 1 \end{pmatrix}.$$

We can then easily check that $y_-^T x_+ = -\sqrt{\epsilon} + \sqrt{\epsilon} = 0$

(c) For $\epsilon = 0$, then $x_+ = x_-$ and the matrix A is **defective** (no complete basis of eigenvectors). In this case, $y_+^T x_+ = 0$, as can easily be checked by direct calculation: $(0, 1) \cdot (1, 0)^T = 0$.

Problem 6:

(a) **Gaussian elimination,** because that corresponds to multiplying A on the *left* by an invertible matrix (the row operations) to get a new matrix (there, in upper-triangular form). In contrast, Gram–Schmidt corresponds to multiplying A on the *right* by an invertible matrix to get an orthogonal matrix $Q = AR^{-1}$ (the ortho-normalized columns). (Technically, QR factorization $A = QR$, via Gram–Schmidt or other means, could be interpreted as multiplying A on the *left* by the unitary matrix Q^* to get the upper-triangular matrix R; if you *explicitly* defended a choice of Gram–Schmidt in this way, then that *is* an acceptable answer.) Similar matrices correspond to multiplying A on *both* the left and right (by a matrix and its inverse) to get a matrix with the same eigenvalues. Diagonalization is a special case of similar matrices to get a diagonal matrix. Least-squares involves either computing A^TA for the normal equations, or doing QR factorization as discussed previously; also, typically least-squares is for cases where you can *not* solve $A\vec{x} = \vec{b}$ exactly.

(b) Multiplying A on the left by an invertible matrix, just as in Gaussian elimination, does not change the null space or the row space. Hence we can determine $N(A) = N(B)$ and $C(A^T) = C(B^T)$.

(c) For $Ax = b$ to be solvable for all b, we need A to have full row rank, i.e. if A is $m \times n$ then we need $C(A) = \mathbb{R}^m$, or $\text{rank}(A) = \text{rank}(B) = m$. Equivalently, B must have **full row rank**. Equivalently, $C(B) = \mathbb{R}^m$.

(d) Given $Ax = b$, we know $XAx = Bx = Xb$, hence

$$Xb = Bx = \begin{pmatrix} 1 & 2 & -1 \\ 4 & 9 & -6 \\ 1 & 4 & -5 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \\ -6 \end{pmatrix}.$$

(e) Since $N(A) = N(B)$, we should just take our solution $(1, 2, 3)^T$ and add any element of $N(B)$. The most straightforward way to find $N(B)$ is by Gaussian elimination to rref form:

$$\begin{pmatrix} 1 & 2 & -1 \\ 4 & 9 & -6 \\ 1 & 4 & -5 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 2 & -1 \\ 0 & 1 & -2 \\ 0 & 2 & -4 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 2 & -1 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{pmatrix},$$

at which point it is clear that B has rank 2 (2 pivots), and a basis for its nullspace is $(-3, 2, 1)^T$. (We can easily double-check this vector by multiplying B to get zero.) Hence the complete solution is

$$x = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + c \begin{pmatrix} -3 \\ 2 \\ 1 \end{pmatrix}$$

for any scalar c.