Math 115A
Homework 7 Comments

I graded 8 of the problems:
Section 5.2: 3b, 3f, 8, 9, 11
Section 6.1: 5, 9, 10

Each problem is worth 2 points. A grade of 0 indicates no solution or a substantially wrong solution. A grade of 2 indicates a correct or nearly correct solution. Otherwise the grade given is 1.

If you believe a problem was misgraded, or I made some addition or other error, please write a short note explaining the situation, attach it to your homework, and return it to me (either in person, in my mailbox, or under my office door). I'll take a look and afterwards leave your homework in a box outside my office.

The following are comments and occasionally solutions for the graded problems.

General Comments
The maximum number of points was 16. The high score was 14. I was disapointed no one got higher, but five people got this score and there were quite a few other high scores. The mean was 9.7. This is a little over 60%, and is the highest mean for a homework assignment yet.

5.2

3. I graded both parts (b) and (f) since they both seemed interesting and neither had the solution in the book. Each part was worth 2 points. The first task to to express T in terms of the standard basis. If you got this right (which most people did) but then for whatever reason said the matrix was not diagonalizable (which a lot of people did) I gave only a half point, since you avoided doing most of the work. I added the scores for both parts and rounded any resulting half point up. I was fairly harsh in grading these but I figure it's better to make your mistakes now and realize what they were than to make them on the exam.

   For part (b), choosing $\gamma = \{1, x, x^2\}$ as the standard basis of V you get

$$[T]_\gamma = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

   Choosing $\{x^2, x, 1\}$ as the order will give the same result. Some people ended up with $[T]_\gamma$ as the identity matrix, and I gave zero points for this because you should realize that that's just not possible; T is certainly not the identity transformation.

   The eigenvalues are $1, -1$ with $a(1) = 2$ and $a(-1) = 1$; determining the null space of each $T - \lambda I$ gives $g(1) = 2$ and $g(-1) = 1$, so T is diagonalizable and the change of basis matrix is

$$Q = [I]_\beta^\gamma = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & -1 \end{pmatrix}.$$

   Now the problem actually asked for a basis β of V, the original vector space $P_2(\mathbb{R})$, but I didn't take off points if you got at least far. The final answer is $\beta = \{1 + x^2, x, 1 - x^2\}$.

   For part (f), choosing $\gamma = \{E_{11}, E_{12}, E_{21}, E_{22}\}$ as the standard basis of V you get

$$[T]_\gamma = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

   Some people had trouble calculating the determinant, but if you do it right you get that the eigenvalues are again $1, -1$, but this time $a(1) = 3$ and $a(-1) = 1$. Determining the null space of each $T - \lambda I$ gives $g(1) = 3$ and $g(-1) = 1$, so again T is diagonalizable and now the change of basis matrix is

$$Q = [I]_\beta^\gamma = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

   Converting back to $M_{2\times 2}(\mathbb{R})$ we get $\beta = \{E_{11}, E_{12} + E_{21}, E_{22}, E_{12} - E_{21}\}$.

8. This was a test to see if you understand the definitions and theorems in the section; if you do the problem is very easy. The simplest solution is to note that we are given $\dim(E_{\lambda_1}) = n - 1$, and we know $\dim(E_{\lambda_2}) \geq 1$ since λ_2 is also an eigenvalue. So choose $n - 1$ linearly independent vectors from E_{λ_1} and a non-zero vector from E_{λ_2}. By Theorem 5.8 all of these vectors are linearly independent in V, and since V has dimenion n they must therefore form a basis, so A is diagonalizable.

   Another simple proof uses Theorem 5.7. In the notation used in lecture, Theorem 5.7 says that $1 \leq g(\lambda_1) \leq a(\lambda_1)$. Also since the characteristic polynomial has degree n we know $a(\lambda_1) + a(\lambda_2) \leq n$ It follows $a(\lambda_1) = n - 1$ or n. We also know $a(\lambda_2) \geq 1$ and so it must be the case that $a(\lambda_1) = n - 1$ and $a(\lambda_1) = 1$. Finally using Theorem 5.7 again we have $1 \leq g(\lambda_2) \leq a(\lambda_2) = 1$ which forces $g(\lambda_2) = 1$. It follows $g(\lambda_1) = a(\lambda_1)$ and $g(\lambda_2) = a(\lambda_2)$ and therefore A is diagonalizable.

   We didn't go over this problem in section but it appears several people must have discussed the problem with Prof. Krishna in office hours. Some people took his hints and came up with a good solution, while others just copied down what he apparently wrote on the board without exhibiting any understanding.

9. The key to this problem is actually problem 12 from Section 5.1, which was not assigned but is discussed in the text (the top of page 249) and I'm sure was discussed in lecture as well. The point is that similar matrices have the same characteristic polynomial, and this is what allows us to define the characteristic polynomial of a linear operator—the fact that it doesn't depend on what basis you choose when you represent the operator as a matrix. I wanted people to reference this fact in their solutions, but only a few people did. I was generous in grading, however.

   In problem 9 of Section 5.1 in last week's homework you proved that the characteristic polynomial of $[T]_\beta$ is a product of linear factors (see my solution posted on the web), so part (a) follows from this, and most people got this.

   People had more trouble formulating an analog for matrices. Some people said let's assume A is upper triangular. This is too weak. Many people just rewrote the statement using L_A in place of T. This is okay, but not quite in the spirit of what was asked for. The best formulation would go something like this: Let A be a matrix that is similar to some upper triangular matrix B; then the characteristic polynomial of A splits. Again the solution relies on the fact that the characteristic polynomial of A is the same as that for B.

11. There were two key points to the solution of this problem. The first is problem 9 of Section 5.1 again: The eigenvalues of an upper triangular matrix its diagonal entries. In fact to be more careful you should note that the solution to this problem also shows that each eignvalue appears on the diagonal with the correct multiplicity. This gives the trace result immediately for any upper triangular matrix; for determinant you also use property 4 on page 235.

    In the problem however we are only given that A is similar to some upper triangular matrix B. So the second key point is that $\mathrm{tr}(A) = \mathrm{tr}(B)$ and $\det(A) = \det(B)$. The first follows from problem 10 of Section 2.5 (on old homework), and the second from property 9 of determinants on page 236.

    Many people missed at least one of the key points. Several people assumed A is diagonalizable which may not be true and in any case is not necessary. I gave one point for a reasonable effort.

6.1

5. I graded just properties (a) and (c), which were worth one point apiece. Property (b) is just like (a), and (d) I did in section although I made a rather substantial error so I'd like to correct that here.

   For property (a) the slickest solution is to note that $\langle x + z, y\rangle = (x + z)Ay^* = xAy^* + zAy^* = \langle x, y\rangle + \langle z, y\rangle$ by the properties of matrix multiplication. Property (b) follows the same way: $\langle cx, y\rangle = (cx)Ay^* = c(xAy^*) = c\langle x, y\rangle$.

   Property (c) requires a little more work: First note that since xAy^* is a 1×1 matrix, it is equal to its transpose. So $xAy^* = (xAy^*)^t = \bar{y}A^tx^t$. Now we have $\overline{\langle x, y\rangle} = \overline{xAy^*} = \overline{\bar{y}A^tx^t} = yA^*x^* = yAx^* = \langle y, x\rangle$ since A is self-adjoint.

   Several people had these slick solutions with no explanation, and I assume they came from Prof. Krishna but I gave full credit. Quite a few other people solved these by multiplying everything out and comparing, and I gave full credit here if it looked like you did enough work since I didn't want to check the details.

For property (d) I solved this in section by multiplying everything out. Letting $x = (x_1 x_2)$ we have $xAx^* = |x_1|^2 + ix_1\overline{x_2} - i\overline{x_1}x_2 + 2|x_2|^2$. Now in section I said the middle two terms cancel. However this is not true; actually $ix_1\overline{x_2} - i\overline{x_1}x_2 = 2\Im(\overline{x_1}x_2)$ where $\Im$ denotes taking the imaginary part. This shows that the product is real at least but we still have to do more work to show at it is greater than zero.

Let $x_1 = a + bi$ and $x_2 = c + di$, where $a, b, c, d \in \mathbb{R}$. Then

$$\begin{aligned} xAx^* &= (a+bi)(a-bi) + i(a+bi)(c-di) - i(a-bi)(c+di) + 2(c+di)(c-di) \\ &= a^2 + b^2 + i[(ac+bd) + (bc-ad)i] - i[(ac+bd) + (ad-bc)i] + 2c^2 + 2d^2 \\ &= a^2 + b^2 + 2(ad-bc) + 2c^2 + 2d^2 \\ &= (a^2 + 2ad + d^2) + (b^2 - 2bc + c^2) + c^2 + d^2 \\ &= (a+d)^2 + (b-c)^2 + c^2 + d^2 \end{aligned}$$

You can check that this is always positive as long as at least one of a, b, c, d is not zero.

Although I didn't grade this part I did look through solutions to see if anyone caught my error and corrected it. I'll bet a lot of people saw the error, but probably no one wanted to deal with fixing it. Only one person seemed to solve this correctly, although I didn't check the computation.

The reason that A gives rise to an inner product this way is that not only is it self-adjoint, but it is also positive definite, which is a stronger condition. Some equivalent conditions for a matrix (or linear operator) to be positive definite are given in problem 17 of Section 6.4. Note that condition (b) is exactly the one used by this problem, so that a matrix gives rise to an inner product in this way if and only if it is positive definite.

I didn't grade the computation at the end of the problem although looking through solutions people came up with all sorts of answers. I believe the correct answer is $6 + 21i$.

9. Each part was worth one point. I originally misread this problem and thought that $\langle x, z\rangle = 0$ for all $z \in V$, the vector space, instead of just all $z \in \beta$, the basis. If the former had been true then we'd just let $z = x$ which means $\langle x, x\rangle = 0$ and so $x = 0$ by Theorem 6.1(d). However this misreading still provides a method to proceed: Let $x = a_1z_1 + \cdots + a_nz_n$ be the representation of x in the basis. Then $\langle x, x\rangle = \langle x, a_1z_1\rangle + \cdots + \langle x, a_nz_n\rangle = \overline{a_1}\langle x, z_1\rangle + \cdots + \overline{a_n}\langle x, z_n\rangle$ by conjugate linearity. However by hypothesis each $\langle x, z_i\rangle = 0$ and so $\langle x, x\rangle = 0$. We can now use Theorem 6.1(d) to conclude $x = 0$.

   Some people tried to use Theorem 6.1(c) to solve this, but it doesn't apply at all.

   For part (b), the easiest way to proceed is to just use part (a). If $\langle x, z\rangle = \langle y, z\rangle$ for all $z \in \beta$, then by linearity $\langle x - y, z\rangle = 0$ for all $z \in \beta$, and so by part (a) it follows $x - y = 0$ and thus $x = y$.

10. This was a giveaway: Since x and y are orthogonal, by definition $\langle x, y\rangle = 0$ and therefore also $\langle y, x\rangle = \overline{\langle x, y\rangle} = 0$. Thus $\|x + y\|^2 = \langle x + y, x + y\rangle = \langle x, x\rangle + \langle x, y\rangle + \langle y, x\rangle + \langle y, y\rangle = \|x\|^2 + \|y\|^2$. Most people got this.

    Quite a few people invoked $\langle x, y\rangle = \|x\|\,\|y\|\cos\theta$ and the definition of orthogonality from their lower division course, but this only applies to $\mathbb{R}^n$ and not to more abstract vector spaces. So you should make sure to understand the new definition which is in fact simpler! And in the abstract setting we do things backwards—we define the angle between two vectors in an arbitrary inner product space to be $\theta = \cos^{-1}\frac{\langle x,y\rangle}{\|x\|\|y\|}$.

    I'm not sure what was desired when they ask to deduce the Pythagorean theorem in $\mathbb{R}^2$. I guess it was enough just to draw a picture of a right triangle and how how the vectors correspond to the sides, which is what many people did.