Answers to exam 1 — Math 4350/5350/7350 — Spring 02

Provide short, complete, precise solutions to **15** of the following problems. On the inside front cover of your exam book list the problem numbers you **omit**.

Insufficient justification earns 0 points. Read the problems carefully. Use complete sentences. Provide examples. *Above all be neat!*

1. What is the matrix representation of a linear transformation?

Let $T: V \to W$ be a linear transformation, and let E and B be bases in V and W, respectively. The matrix representation A of T with respect to E and B is defined by the equation $[T(x)]_E = A[x]_B$, where [ ] denotes the coordinates of a vector in the indicated basis.

2. Give an example of how to represent a vector in two different bases.

Consider a random matrix-vector multiplication:

$$\begin{bmatrix} 3 & 2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}.$$

From this equation we see that the vector on the right has coordinates 2 and 1 with respect to the standard basis, but has coordinates 1 and -1 with respect to the basis given by the columns of the matrix on the left.

3. Express matrix multiplication in terms of column vectors.

If $AB = C$ then the i-th column of C is a linear combination of the columns of A, using the entries in the i-th column of B as coefficients.

4. True or false: the eigenvalues of a unitary matrix have norm 1?

This is true. If A is unitary and $Ax = \lambda x$, with $x \neq 0$, then $A^*A = I$, and

$$\begin{aligned} \|x\|^2 &= x^*x = x^*A^*Ax = (Ax)^*(Ax) \\ &= (\lambda x)^*(\lambda x) = \overline{\lambda}\lambda x^*x = \|x\|^2 \, |\lambda|^2 . \end{aligned}$$

Since $\|x\| \neq 0$ we obtain that $|\lambda|^2 = 1$.

5. True or false: the eigenvalues of a hermitian matrix are real?

This is true. If A is hermitian and $Ax = \lambda x$, with $x \neq 0$, then $A^* = A$, and

$$\begin{aligned} \overline{\lambda} \|x\|^2 &= \overline{\lambda} x^*x = (\lambda x)^*x = (Ax)^*x = x^*A^*x \\ &= x^*Ax = x^*(\lambda x) = \lambda x^*x = \lambda \|x\|^2 . \end{aligned}$$

Since $\|x\| \neq 0$ we obtain that $\overline{\lambda} = \lambda$, which is to say that λ is real.

6. True or false: every rank-1 matrix is an outer product?

This is true. To say that A has rank 1 is to say that every column of A is a multiple of some column x (with at least one multiple being nonzero). If $A = \left[y_1 x \middle| y_2 x \middle| \cdots \middle| y_n x \right]$ then $A = xy^*$ where $y = \begin{bmatrix} \overline{y_1} \\ \vdots \\ \overline{y_n} \end{bmatrix}$. (See problem 3.)

7. True or false: the left singular vectors (U-vectors) of a matrix A are the eigenvectors of AA^*?

This is true. If $A = U\Sigma V^*$ then

$$AA^* = U\Sigma V^* V \Sigma^* U^* = U\Sigma^*\Sigma U^{-1}.$$

Now $\Sigma\Sigma^*$ is a diagonal matrix whose diagonal entries are the squares of the singular values σ_1^2, σ_2^2, ... (and possibly with a string of zeroes). This tells us that the columns of U — the left singular vectors for A — are eigenvectors for AA^*, and also that the squares of the singular values of A (and possibly some extra zeroes) are the eigenvalues of AA^*.

8. True or false: the singular values of a matrix A are the eigenvalues of A^*A?

This is false. For example consider the following diagonal matrix A:

$$\begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}^* .$$

We have displayed its singular value decomposition: its singular values are 3 and 2, and both its left and right singular vectors are the standard basis vectors. However,

$$A^*A = \begin{bmatrix} 9 & 0 \\ 0 & 4 \end{bmatrix},$$

whose eigenvalues are plainly 9 and 4. More generally, the *squares* of the singular values of a matrix A are the eigenvalues of A^*A, as one can show using calculation similar to the previous problem.

9. True or false: $\|A\|_1$ is the sum of the singular values of A?

This is false. If we look at the example in the previous problem we see that the sum of its singular values is 5. However, the 1-norm is the maximum column sum, which in this case is 3.

10. Show that if $x \in \mathbb{C}^m$ then $\|x\|_\infty \le \|x\|_2 \le \sqrt{m}\,\|x\|_\infty$.

Say that $\|x\|_\infty = |x_k|$. Then

$$|x_k|^2 \le |x_1|^2 + \cdots + |x_m|^2 \le |x_k|^2 + \cdots + |x_k|^2 = m\,|x_k|^2 .$$

Thats is $\|x\|_\infty^2 \le \|x\|_2^2 \le m\,\|x\|_\infty^2$.

11. Use problem 10 to show that if $A \in \mathbb{C}^{m\times n}$ then $\|A\|_2 \le \sqrt{m}\,\|A\|_\infty$.

Suppose that $x \in \mathbb{C}^n$ and $x \ne 0$. From problem 10 we have that

$$\frac{\|Ax\|_2}{\|x\|_2} \le \frac{\sqrt{m}\,\|Ax\|_\infty}{\|x\|_\infty} \le \sqrt{m}\,\|A\|_\infty .$$

Hence $\|A\|_2$, which is the supremum of the left-hand side of this inequality, cannot exceed $\sqrt{m}\,\|A\|_\infty$.

12. Use problem 10 to show that if $A \in \mathbb{C}^{m\times n}$ then $\|A\|_\infty \le \sqrt{n}\,\|A\|_2$.

Suppose that $x \in \mathbb{C}^n$ and $x \ne 0$. From problem 10 we have that

$$\frac{\|Ax\|_\infty}{\|x\|_\infty} \le \frac{\|Ax\|_2}{\|x\|_2 / \sqrt{n}} \le \sqrt{n}\,\|A\|_2 .$$

Hence $\|A\|_\infty$, which is the supremum of the left-hand side of this inequality, cannot exceed $\sqrt{n}\,\|A\|_2$.

13. Let $V = \operatorname{span}\{[2\ -3\ 0\ 1]^*,\ [1\ 0\ 1\ -1]^*\}$, $W = \{x \in \mathbb{C}^3 \mid x_1 = -x_2 = 2x_3\}$. Write down a matrix A whose range is V, whose nullspace is W, whose 2-norm is 7, and whose Frobenius norm is $\sqrt{50}$.

We first find orthonormal bases u_1, u_2 for V and v_1, v_2 for $W^\perp$, then use this to write down A using the outer product form of its singular value decomposition. From the description given, we deduce that the singular values of A are 7 and 1, whence $A = 7u_1v_1^* + u_2v_2^*$.[1]

14. Suppose P is a projector. Show that P is an orthogonal projector if and only if $\|P\|_2 = 1$.

By definition, $P^2 = P$, hence $\|P\|_2 = \left\|P^2\right\|_2 \le \|P\|_2^2$. If $P \ne 0$ then we find that $1 \le \|P\|_2$, for any projector. Now suppose that P is an orthogonal projector and consider its singular value decomposition of P, $U\Sigma V^*$. We know that $U = V$, since the decomposition can be put together from orthonormal bases of null(P) and range(P). Moreover, all of the singular values of P are either 1 (for those singular vectors in range(P)) or 0 (for those in null(P)). Hence $\|P\|_2 = 1$. Conversely, if P is a nonorthogonal projector, then there are nonorthogonal vectors $x \in \operatorname{range}(P)$ and $r \in \operatorname{null}(P)$. Note that $P(x + r) = x$ and that $\|x + r\|_2^2 = \|x\|_2^2 + \|r\|_2^2 + 2\operatorname{Real}(x^*r)$. If we scale x and r we may arrange that $\|r\|_2^2 + 2\operatorname{Real}(x^*r) < 0$, in which case $\|P(x)\|_2 > \|x\|_2$.

---

[1] Note: If you had said this much I would have given you full credit, since I want to know how much you learned, not how well you do arithmetic. Nevertheless, for your amusement, let's finish the computation.

To find these bases we start with bases for V and $W^\perp$ and use orthogonal projection — or one step of Gram-Schmidt, if you prefer — to produce orthogonal bases. To wit, let x_1 and x_2 denote the given basis for V, let $y_1 = [1\ 1\ 0]^*$ and $y_2 = [1\ 0\ -2]^*$, which is a basis for $W^\perp$. Now $x_1^*x_1 = 14$, $x_1^*x_2 = 1$, $y_1^*y_1 = 2$, and finally $y_1^*y_2 = 1$. Thus the computations

$$u_1 = \frac{1}{\sqrt{14}}x_1 = \begin{bmatrix} 2/\sqrt{14} \\ -3/\sqrt{14} \\ 0 \\ 1/\sqrt{14} \end{bmatrix}$$

$$u_2 = \frac{x_2 - \frac{1}{14}x_1}{\left\|x_2 - \frac{1}{14}x_1\right\|} = \frac{14}{\sqrt{12^2 + 3^2 + 14^2 + 15^2}} \begin{bmatrix} 12/14 \\ 3/14 \\ 1 \\ -15/14 \end{bmatrix} = \begin{bmatrix} 12/\sqrt{574} \\ 3/\sqrt{574} \\ 14/\sqrt{574} \\ -15/\sqrt{574} \end{bmatrix}$$

$$v_1 = \frac{1}{\sqrt{2}}y_1 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix}$$

$$v_2 = \frac{y_2 - \frac{1}{2}y_1}{\left\|y_2 - \frac{1}{2}y_1\right\|} = \frac{2}{\sqrt{1^2 + 1^2 + 4^2}} \begin{bmatrix} 1/2 \\ -1/2 \\ -2 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{17} \\ -1/\sqrt{17} \\ -4/\sqrt{17} \end{bmatrix},$$

yield the required orthonormal bases. So, finally

$$A = \begin{bmatrix} \frac{14}{\sqrt{28}} + \frac{12}{\sqrt{9758}} & \frac{14}{\sqrt{28}} - \frac{12}{\sqrt{9758}} & -\frac{48}{\sqrt{9758}} \\ -\frac{21}{\sqrt{28}} + \frac{3}{\sqrt{9758}} & -\frac{21}{\sqrt{28}} - \frac{3}{\sqrt{9758}} & -\frac{12}{\sqrt{9758}} \\ -\frac{14}{\sqrt{9758}} & \frac{14}{\sqrt{9758}} & \frac{14}{\sqrt{9758}} \\ \frac{7}{\sqrt{28}} - \frac{15}{\sqrt{9758}} & \frac{7}{\sqrt{28}} + \frac{15}{\sqrt{9758}} & \frac{60}{\sqrt{9758}} \end{bmatrix}$$

15. Suppose P is an orthogonal projector. Show that $I - 2P$ is unitary. Describe the linear transformation $I - 2P$ geometrically.

We know that if P is an orthogonal projector then $P^* = P$ — this was proved in our text, using the singular value decomposition, as in the previous problem — whence $(I - 2P)^* = I^* - 2P^* = I - 2P$. Therefore

$$\begin{aligned}(I-2P)^*(I-2P) &= (I-2P)(I-2P)^* = (I-2P)^2 \\ &= I - 4P + 4P^2 = I - 4P + 4P = I.\end{aligned}$$

So, $I - 2P$ is both its own adjoint and its own inverse. In particular it is unitary. In fact it is a reflection thru null(P). This is because if $x = y + r$, where $y \in \text{range}(P)$ and $r \in \text{null}(P)$, then $Px = -y + r$.

16. Compute the orthogonal projector onto span$\{[1\ 2\ 0]^*,\ [1\ {-1}\ {-1}]^*\}$.

There are at least two ways to do this. If A is the matrix whose columns are the given vectors, then $A(A^*A)^{-1}A^*$ — often denoted AA^+ — is the orthogonal projection onto range(A):

$$\begin{aligned}A(A^*A)^{-1}A^* &= \begin{bmatrix}1 & 1\\ 2 & -1\\ 0 & -1\end{bmatrix}\begin{bmatrix}3/14 & 1/14\\ 1/14 & 5/14\end{bmatrix}\begin{bmatrix}1 & 2 & 0\\ 1 & -1 & -1\end{bmatrix} \\ &= \begin{bmatrix}10/14 & 2/14 & -6/14\\ 2/14 & 13/14 & 3/14\\ -6/14 & 3/14 & 5/14\end{bmatrix}\end{aligned}$$

Alternatively, we can compute an orthonormal basis from the given vectors, using vector projection, form a matrix Q from these, and then the product QQ^* will also yield the orthogonal projection. That they come out to the same thing can be seen by using the QR factorization of A.

17. Explain how to use the QR factorization of a matrix to solve a system of equations. Give an example.

If it is known that $A = QR$, where Q is unitary and R is upper triangular, then any equation $Ax = b$ can be solved fairly easily. First, $Rx = Q^*b$, and the latter is easy to compute. Second, the equation $Rx = y$ is relatively easy to solve for x, by the method of back substitution. For example, if

$$Q = \begin{bmatrix}1/\sqrt{2} & -1/\sqrt{2}\\ 1/\sqrt{2} & 1/\sqrt{2}\end{bmatrix},\quad R = \begin{bmatrix}2 & -1\\ 0 & 3\end{bmatrix},\quad b = \begin{bmatrix}1\\ 1\end{bmatrix}$$

then

$$Q^*b = \begin{bmatrix}\sqrt{2}\\ 0\end{bmatrix},\quad x_2 = \frac{0}{3} = 0,\ \ x_1 = \frac{\sqrt{2}+0}{2} = 1/\sqrt{2}.$$

18. Describe in detail how both the classical and modified Gram-Schmidt algorithms work. What is the advantage of the modified algorithm over the classical algorithm?

Both algorithms start with vectors $a_1, \ldots, a_n$ and proceed by a sequence of vector projections to produce an orthonormal set $q_1, \ldots, q_n$ with the property that span$\{a_1, \ldots, a_k\}$ = span$\{q_1, \ldots, q_k\}$ for all k. The classical algorithm projects each x_k so that it is orthogonal to the vectors already produced, $q_1, \ldots, q_{k-1}$, and then normalizes to length 1. The modified algorithm, which is more numerically stable, first normalizes x_k to length 1 and then projects all of the *remaining* vectors $x_{k+1}, \ldots, x_n$ so that they are orthogonal to it.

19. Why is the Gram-Schmidt procedure described as "triangular orthogonalization"? Why is the Householder algorithm described as "orthogonal triangularization"? What do these algorithms produce?

Both algorithms produce the QR factorization of a matrix A, altho the Householder algorithm does not yield Q explicitly. It proceeds by a sequence of orthogonal operations which culminate in a triangularization of A: $Q_n \cdots Q_1 A = R$ and $Q = Q_1 \cdots Q_n$. (The Q_i are hermitian.) By contrast, the (classical) Gram-Schmidt algorithm proceeds by a sequence of triangular operations which culminate in an orthogonalization of A: $AR_1 \cdot R_n = Q$ and $R = R_n^{-1} \cdots R_1^{-1}$. (The inverses and this product are trivial to compute.)

20. Suppose $A \in \mathbb{C}^{m \times n}$. Show that A has rank n if and only if A^*A is invertible.

Since A^*A is square, it is invertible if and only if its nullity is 0. Now if $x \in \text{null}(A)$ then $A^*Ax = 0$, so certainly if A has rank *less than* n then A^*A is *not* invertible. Conversely, if A has rank n then its nullity is 0. Suppose $x \in \text{null}(A^*A)$. We have $\|Ax\|^2 = x^*A^*Ax = 0$, whence $Ax = 0$. By hypothesis this means that $x = 0$. That is, the nullity of A^*A is 0, as claimed.