Practice Problems

1. Remember that a matrix Q is *unitary* if $Q^H Q = I$. A matrix is *orthogonal* if it is real and unitary; that is, if it is real and $Q^T Q = I$.

a) Find the flaw in this argument:

   False Claim: all eigenvalues of an orthogonal matrix are ± 1. Indeed, if $Qx = \lambda x$,

$$\lambda^2 x^T x = (Qx)^T (Qx) = x^T (Q^T Q) x = x^T x,$$

   therefore $\lambda^2 = 1$, so $\lambda = \pm 1$. If you want, you can think about what happens for a rotation matrix

$$R = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

b) Correct the proof to show

   True Claim: all eigenvalues of a unitary matrix have magnitude 1 (e.g. $\lambda = e^{i\phi}$ for some ϕ).

c) Show that the eigenvectors for different eigenvalues of a unitary matrix are orthogonal.

d) Show that the determinant of any real unitary matrix (e.g., an orthogonal matrix) is ± 1 using eigenvalues. (Note: you already proved this on a previous pset in a different way.)

Solution. a) Let's think through this carefully. The equality

$$\lambda^2 x^T x = x^T x$$

from the argument above is true. This equality tells us that

$$\begin{aligned} \lambda^2 x^T x - x^T x &= 0 \\ x^T x(\lambda^2 - 1) &= 0. \end{aligned}$$

However, the second equality here tells us that either $\lambda^2 = 1$ *or* $x^T x = 0$. We are used to thinking about real vectors; if x is real, then $x^T x = ||x||^2 > 0$ since x cannot be the zero vector here (it is an eigenvector). However, here it's possible that x is complex, and if x is complex, it's easy for $x^T x = 0$ even if x is not the zero vector! For example, if

$$x = \begin{pmatrix} i \\ 1 \end{pmatrix}$$

then $x^T x = i^2 + 1 = 0$. So the flaw in the proof is the assumption that $x^T x = 0$.

b) To get a proof of the true claim, we switch every T in sight to an H (that is, we take the adjoint rather than the transpose.) So the proof of the true claim goes like this:

   If $Qx = \lambda x$, then

$$\lambda\bar{\lambda} x^H x = (Qx)^H (Qx) = x^H (Q^H Q) x = x^H x.$$

This means that

$$\lambda\overline{\lambda}x^H x - x^H x = 0$$
$$x^H x(\lambda\overline{\lambda} - 1) = 0.$$

Since $x^H x = ||x||^2 > 0$, we see that $\lambda\overline{\lambda} = |\lambda|^2 = 1$. This means that λ has magnitude 1, that is, $\lambda = e^{i\theta}$.

c) Say x_1, x_2 are eigenvectors with eigenvalues $\lambda_1 \neq \lambda_2$. We want to show that $x_1^H x_2 = 0$. We'll mimic the argument from b), but using x_1, x_2 rather than just a single eigenvector.

So we have

$$\overline{\lambda_1}\lambda_2 x_1^H x_2 = (Qx_1)^H(Qx_2) = x_1^H(Q^H Q)x_2 = x_1^H x_2.$$

This tells us that

$$x_1^H x_2(\overline{\lambda_1}\lambda_2 - 1) = 0.$$

So either $x_1^H x_2 = 0$ or $\overline{\lambda_1}\lambda_2 = 1$. The second situation is impossible: remember that since $\lambda_1 = e^{i\theta}$, its conjugate $\overline{\lambda_1} = e^{-i\theta} = 1/\lambda_1$. So if we were in the second situation, we would have $\overline{\lambda_1} = 1/\lambda_2$, which would imply $\lambda_1 = \lambda_2$. We assumed the two eigenvalues were *not* equal, so this is impossible, as claimed.

In summary, we must have $x_1^H x_2 = 0$, so the vectors are orthogonal.

d) Say Q is an $m \times m$ orthogonal matrix, with eigenvalues $\lambda_1, \dots, \lambda_m$. Since it's unitary, its eigenvalues must have the form $e^{i\theta}$. Since it's real, its complex eigenvalues must come in complex conjugate pairs. So the eigenvalues are p 1's, q (-1)'s and then r pairs of conjugate complex numbers. Remember that $\lambda\overline{\lambda} = |\lambda|^2$, so the product of the complex conjugate pairs is just 1.

We know

$$\det(Q) = \lambda_1\lambda_2\cdots\lambda_m = 1^p \cdot (-1)^q 1^r = \pm 1.$$

2. Here is a quick "proof" that the eigenvalues of **every** real matrix A are real:

$$\textbf{False Proof: } Ax = \lambda x \text{ gives } x^T Ax = \lambda x^T x, \quad \text{so } \lambda = \frac{x^T Ax}{x^T x} = \frac{\text{real}}{\text{real}}.$$

Find the flaw in this reasoning – a hidden assumption that is not justified. You can test those steps on the 90° rotation matrix

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \qquad \lambda = i,\ x = \begin{pmatrix} i \\ 1 \end{pmatrix}.$$

Solution. There are a couple of flaws in this argument. First, as we saw in the previous problem, it's possible for $x^T x = 0$ even if $x \neq 0$, so this division at the end might be dividing by zero. Even if we're not dividing by zero, $x^T Ax$ and $x^T x$ are not necessarily real, so the final equality is also not true.

3. a) If S is a positive definite matrix, show that S^{-1} is also positive definite.

b) If S and T are positive definite, show that their sum $S+T$ is also positive definite. If $S = A^H A$ and $T = B^H B$ for full-column-rank matrices A and B, then can you write down a full column-rank matrix C so that $S + T = C^T C$?

Solution. a) Remember that a matrix A is *positive definite* if it is Hermitian $A^H = A$ and its eigenvalues are all positive (or a couple of other equivalent conditions). So we know that $S^H = S$ and the eigenvalues of S are all positive. We need to check that the same properties hold for S^{-1}.

The eigenvalues of S^{-1} are just the reciprocals of eigenvalues of S, so they are all positive. We should also check that $(S^{-1})^H = S^{-1}$. We check this by starting with the true equation $S^{-1}S = I$ and taking the adjoint of both sides:

$$\begin{aligned}(S^{-1}S)^H &= I^H \\ S^H(S^{-1})^H &= I \\ S(S^{-1})^H &= I\end{aligned}$$

where we used the fact that $S^H = S$.

b) It's easiest to use a different characterization of positive definite matrices for this problem: $S^H = S$ and $x^H S x > 0$ for all vectors $x \neq 0$. We check these two properties for $S + T$:

$$(S+T)^H = S^H + T^H = S + T \qquad \text{and} \qquad x^H(S+T)x = x^H Sx + x^H Tx > 0.$$

The answer to the question is: you can write down the matrix C, but it's hard to relate it to A, B. In particular, from lecture, one way to get the matrix C (resp., A and B) is to look at the diagonalization of $S+T$ (resp, S and T). But the diagonalization of $S+T$ is hard to relate to the diagonalization of S and T separately; the eigenvectors and eigenvalues might be completely unrelated. (If anyone has a better answer to this question, I'd be happy to hear it!)

4. Say A is a 3×3 real matrix. The matrix $B = A + A^T$ has eigenvalues $\lambda_1 = 2, \lambda_2 = 0, \lambda_3 = 1$, with corresponding eigenvectors $x_1 = [1\ 2\ 1]$, $x_2 = [-2\ 1\ 0]$ and $x_3 = [1\ 2\ -5]$.

a) What is e^B? (It's fine to leave your answer as a product of several matrices, as long as each matrix is written down explicitly)

b) Let $C = (I - B)(I + B)^{-1}$. What are the eigenvalues and eigenvectors of C?

c) Give a good approximation for

$$y = C^{100}\begin{pmatrix}1\\1\\1\end{pmatrix}$$

in terms of a single eigenvector.

Solution. See `https://github.com/mitmath/1806/blob/fall18/exams/exam3sol.pdf`, Problem 3.

This is similar to problems we've done before in section. The only new part is to remember that B is a symmetric matrix, so the eigenbasis can be chosen to be orthogonal (you can also see by inspection that x_1, x_2, x_3 are orthogonal). This simplifies computing the diagonalization of B, and expanding y in terms of the eigenbasis.