Math 307: Problems for section 3.3

1. **Show that for $\mathbf{v}, \mathbf{w} \in \mathbb{C}^n$**

$$\|\mathbf{v}+\mathbf{w}\|^2 = \|\mathbf{v}\|^2 + \|\mathbf{w}\|^2 + 2\,\mathrm{Re}(\langle \mathbf{v}, \mathbf{w}\rangle)$$

and use this to prove the polarization identity

$$\langle \mathbf{v}, \mathbf{w}\rangle = \frac{1}{4}\Big(\|\mathbf{v}+\mathbf{w}\|^2 - \|\mathbf{v}-\mathbf{w}\|^2 + i\|\mathbf{v}-i\mathbf{w}\|^2 - i\|\mathbf{v}+i\mathbf{w}\|^2\Big)$$

$$\begin{aligned}
\|\mathbf{v}+\mathbf{w}\|^2 &= \langle \mathbf{v}+\mathbf{w}, \mathbf{v}+\mathbf{w}\rangle \\
&= \langle \mathbf{v}, \mathbf{v}\rangle + \langle \mathbf{v}, \mathbf{w}\rangle + \langle \mathbf{w}, \mathbf{v}\rangle + \langle \mathbf{w}, \mathbf{w}\rangle \\
&= \|\mathbf{v}\|^2 + \langle \mathbf{v}, \mathbf{w}\rangle + \overline{\langle \mathbf{v}, \mathbf{w}\rangle} + \|\mathbf{w}\|^2 \\
&= \|\mathbf{v}\|^2 + \|\mathbf{w}\|^2 + 2\,\mathrm{Re}(\langle \mathbf{v}, \mathbf{w}\rangle)
\end{aligned}$$

Thus

$$\begin{aligned}
\|\mathbf{v}+\mathbf{w}\|^2 - \|\mathbf{v}-\mathbf{w}\|^2 &= \|\mathbf{v}\|^2 + \|\mathbf{w}\|^2 + 2\,\mathrm{Re}(\langle \mathbf{v}, \mathbf{w}\rangle) - \big(\|\mathbf{v}\|^2 + \|\mathbf{w}\|^2 - 2\,\mathrm{Re}(\langle \mathbf{v}, \mathbf{w}\rangle)\big) \\
&= 4\,\mathrm{Re}(\langle \mathbf{v}, \mathbf{w}, \rangle)
\end{aligned}$$

and using this equality with $\mathbf{w}$ replaced by $-i\mathbf{w}$ yields

$$\begin{aligned}
\|\mathbf{v}-i\mathbf{w}\|^2 - \|\mathbf{v}+i\mathbf{w}\|^2 &= 4\,\mathrm{Re}(\langle \mathbf{v}, -i\mathbf{w}\rangle) \\
&= 4\,\mathrm{Re}(-i\langle \mathbf{v}, \mathbf{w}\rangle) \\
&= 4\,\mathrm{Im}(\langle \mathbf{v}, \mathbf{w}\rangle)
\end{aligned}$$

The last two equations give the result.

2. **Show that if $\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_n$ form a basis in $\mathbb{R}^n$, then they also form a basis when regarded as vectors in $\mathbb{C}^n$. In other words, show that 1.) if the only linear combination $c_1\mathbf{q}_1 + \cdots + c_n\mathbf{q}_n$ using real numbers $c_1, \ldots, c_n$ that equals zero has $c_1 = \cdots = c_n = 0$, then the same is true for complex numbers, and 2.) if every vector in $\mathbb{R}^n$ can be written as $c_1\mathbf{q}_1 + \cdots + c_n\mathbf{q}_n$ for some real numbers $c_1, \ldots, c_n$ then every vector in $\mathbb{C}^n$ can be written as a linear combination using complex numbers. If the basis $\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_n$ is orthonormal in $\mathbb{R}^n$ is is also orthonormal in $\mathbb{C}^n$?**

Suppose a complex linear combination of $\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_n$ equals zero. Writing the complex scalars as $c_k + id_k$ we have

$$(c_1 + id_1)\mathbf{q}_1 + \cdots + (c_n + id_n)\mathbf{q}_n = \mathbf{0}$$

Given the complex equation above, we can take the real and imaginary parts and get two equations involving real numbers. Thus

$$\mathrm{Re}\left((c_1 + id_1)\mathbf{q}_1 + \cdots + (c_n + id_n)\mathbf{q}_n\right) = \mathrm{Re}(\mathbf{0}) = \mathbf{0}$$

Since each $\mathbf{q}_k$ is a real vector, the real part of $(c_k + id_k)\mathbf{q}_k$ is $c_k\mathbf{q}_k$. Thus

$$c_1\mathbf{q}_1 + \cdots + c_n\mathbf{q}_n = \mathbf{0}.$$

But this implies $c_1 = \cdots = c_n = 0$ since we have assumed that $\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_n$ are linearly independent when using real numbers. Similarly (taking the imaginary part) we find $d_1 = \cdots = d_n = 0$ too. Thus each complex coefficient $c_k + id + k$ is zero. This proves linear independence using complex scalars.

To show that $\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_n$ span $\mathbb{C}^n$ when we use complex scalars, suppose that $\mathbf{w} \in \mathbb{C}^n$. We can write $\mathbf{w} = \mathbf{x} + i\mathbf{y}$ where $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. Then since the vectors in the basis span $\mathbb{R}^n$,

$$\mathbf{x} = c_1\mathbf{q}_1 + \cdots + c_n\mathbf{q}_n, \quad \mathbf{y} = d_1\mathbf{q}_1 + \cdots + d_n\mathbf{q}_n$$

for some choices of real numbers $c_1, \ldots, c_n$ and $d_1, \ldots, d_n$. This implies

$$\mathbf{w} = (c_1 + id_1)\mathbf{q}_1 + \cdots + (c_n + id_n)\mathbf{q}_n$$

Thus any $\mathbf{w} \in \mathbb{C}^n$ is a complex linear combination of the basis vectors.

Finally, since the vectors $\mathbf{q}_k$ are real the complex inner product and the real dot product agree. This implies that these vectors form an orthonormal set in $\mathbb{C}^n$ if they are orthonormal in $\mathbb{R}^n$.

3. **Show that any 2×2 orthogonal matrix is either a rotation matrix or a reflection matrix.**

   To obtain all possible 2×2 orthogonal matrices, we need to find all possible orthonormal bases to use as their columns. The first vector in an orthonormal basis is a vector of length one that can be written $\begin{bmatrix} \cos(\theta) \\ \sin(\theta) \end{bmatrix}$ for some $\theta \in [0, 2\pi)$. The next vector is one of the two vectors of unit length orthogonal to this one, either $\begin{bmatrix} -\sin(\theta) \\ \cos(\theta) \end{bmatrix}$ or $\begin{bmatrix} \sin(\theta) \\ -\cos(\theta) \end{bmatrix}$. This implies that every orthogonal matrix is either of the form $\begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}$ (a rotation matrix) or $\begin{bmatrix} \cos(\theta) & \sin(\theta) \\ \sin(\theta) & -\cos(\theta) \end{bmatrix}$ (a relection matrix).

4. **Let $Q = \begin{bmatrix} \mathbf{q}_1 | \mathbf{q}_2 | \cdots | \mathbf{q}_k \end{bmatrix}$ where $\mathbf{q}_1, \mathbf{q}_2, \ldots \mathbf{q}_k \in \mathbb{R}^n$ form an orthonormal set. (That is, they satisfy $\|\mathbf{q}_i\| = 1$ for $i = 1, \ldots, k$ and $\mathbf{q}_i \cdot \mathbf{q}_j = 0$ if $i \neq j$, but there might not be enough vectors to form a basis, i.e., possibly $k < n$). Identify the matrices Q^TQ and QQ^T. Show that the projection $\mathbf{p}$ of a vector $\mathbf{v}$ onto the subspace spanned by $\mathbf{q}_1, \mathbf{q}_2, \ldots \mathbf{q}_k$ can be written $\mathbf{p} = \sum_{i=1}^{k} \mathbf{q}_i\mathbf{q}_i^T\mathbf{v} = \sum_{i=1}^{k} (\mathbf{q}_i \cdot \mathbf{v})\mathbf{q}_i$.**

   $Q^TQ = I_k$ (the $k \times k$ identity matrix) and $QQ^T = Q(Q^TQ)^{-1}Q^T$ is the projection onto the range of Q, that is, onto the subspace spanned by $\mathbf{q}_1, \mathbf{q}_2, \ldots \mathbf{q}_k$. The formula $\mathbf{p} = QQ^T\mathbf{v}$ is equivalent to $\mathbf{p} = \sum_{i=1}^{k} \mathbf{q}_i\mathbf{q}_i^T\mathbf{v} = \sum_{i=1}^{k} (\mathbf{q}_i \cdot \mathbf{v})\mathbf{q}_i$.

5. **For an $m \times n$ matrix A with linearly independent columns there is a factorization (called the QR factorization) $A = QR$ where Q is an $m \times n$ matrix whose columns form an orthonormal set, and R is an upper triangular matrix. For every $k = 1, 2, \ldots n$ the first k columns of Q spans the same subspace as the first k columns of A. In MATLAB/Octave the matrices Q and R in the QR decomposition of A are computed using `[Q R] = qr(A,0)`. (Without the second argument `0` a related but different decomposition is computed.)(For those of you who have learned about Gram-Schmidt: The columns of Q are the vectors obtained by applying the Gram-Schmidt procedure to the columns of A.**

Using MATLAB/Octave, compute and orthonormal basis $\mathbf{q}_1, \mathbf{q}_2$ for the plane in $\mathbb{R}^4$ spanned by $\mathbf{a}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$ and $\mathbf{a}_2 = \begin{bmatrix} -1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$ Compute the projection $\mathbf{p}$ of the vector $\mathbf{v} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ -1 \end{bmatrix}$ onto the plane. What are the coefficients of $\mathbf{p}$ when expanded in the basis $\mathbf{q}_1, \mathbf{q}_2$?

```
a1=[1;1;1;1];
a2=[-1;1;1;1];
A=[a1 a2];
[Q R] =qr(A,0)
Q =

  -0.50000   0.86603
  -0.50000  -0.28868
  -0.50000  -0.28868
  -0.50000  -0.28868

R =

  -2.00000  -1.00000
   0.00000  -1.73205

q1=Q(:,1)
 q1 =

  -0.50000
  -0.50000
  -0.50000
  -0.50000

q2=Q(:,2)
q2 =

   0.86603
  -0.28868
  -0.28868
  -0.28868

v=[1;1;1;-1];

p=Q*Q'*v
p =

   1.00000
   0.33333
   0.33333
   0.33333
```

The coefficients for the expansion of $\mathbf{p}$ in the basis $\mathbf{q}_1$, $\mathbf{q}_2$ are contained in the vector $Q^T\mathbf{v}$

```
Q'*v
ans =

  -1.00000
   0.57735
```

Check that the expansion with these coefficients gives $\mathbf{p}$.

```
(-1.00000)*q1 + (0.57735)*q2
ans =

   1.00000
   0.33333
   0.33333
   0.33333
```

6. **Using MATLAB/Octave and the discussion in the previous problem, find an orthonormal set of vectors $\mathbf{q}_1$, $\mathbf{q}_2$ and $\mathbf{q}_3$ with the same span as** $\begin{bmatrix}1\\1\\2\\0\\0\\0\end{bmatrix}$ $\begin{bmatrix}1\\0\\1\\0\\1\\0\end{bmatrix}$ **and** $\begin{bmatrix}0\\0\\1\\1\\1\\0\end{bmatrix}$**. Provide the commands that you used.**

```
a1=[1 1 2 0 0 0]';
a2=[1 0 1 0 1 0]';
a3=[0 0 1 1 1 0]';
A=[a1 a2 a3];
[Q R] = qr(A,0)

Q =

  -0.40825   0.40825   0.51640
  -0.40825  -0.40825  -0.00000
  -0.81650   0.00000  -0.25820
  -0.00000   0.00000  -0.77460
  -0.00000   0.81650  -0.25820
  -0.00000   0.00000   0.00000

R =

  -2.44949  -1.22474  -0.81650
   0.00000   1.22474   0.81650
   0.00000   0.00000  -1.29099
```

The three orthonormal vectors are the columns of Q.

7. **Do the following computational experiment. First start with a random symmetric 10×10 matrix A (for example `B=rand(10,10); A=B'*B;` will produce such a matrix) and compute its QR factorization. Call the factors Q_1 and R_1. Now multiply Q_1 and R_1 in the "wrong" order to obtain $A_2 = R_1Q_1$ and compute the QR factorization of the resulting matrix A_2.**

Repeat this step to obtain a sequence of matrices Q_k, R_k and A_k. Do these sequences converge? If so can you identify the limit? (Hint: `eig(C)` computes the eigenvalues of `C`).

We can automate the iterative procedure. Here I'll do it ten times, check the value of R and compare to the eigenvalues of A

```
A=rand(10,10);
A=A'*A;
[Q R] = qr(A);
for k=[1:10]
  [Q R] = qr(R*Q,0);
end
R

R =

 Columns 1 through 7:

  -28.26041   -0.00000   -0.00000    0.00000   -0.00000   -0.00000   -0.00000
    0.00000   -2.20459   -0.48470   -0.00121    0.00004   -0.00000    0.00000
    0.00000    0.00000   -2.23187    0.00364   -0.00022    0.00000   -0.00000
    0.00000    0.00000    0.00000   -1.15933    0.05953   -0.00037    0.00003
    0.00000    0.00000    0.00000    0.00000   -0.76320   -0.00532    0.00035
    0.00000    0.00000    0.00000    0.00000    0.00000   -0.45709    0.01440
    0.00000    0.00000    0.00000    0.00000    0.00000    0.00000   -0.29274
    0.00000    0.00000    0.00000    0.00000    0.00000    0.00000    0.00000
    0.00000    0.00000    0.00000    0.00000    0.00000    0.00000    0.00000
    0.00000    0.00000    0.00000    0.00000    0.00000    0.00000    0.00000

 Columns 8 through 10:

   -0.00000   -0.00000   -0.00000
   -0.00000    0.00000   -0.00000
    0.00000   -0.00000    0.00000
   -0.00000   -0.00000    0.00000
   -0.00000    0.00000   -0.00000
    0.00000   -0.00000    0.00000
    0.00001   -0.00000    0.00000
   -0.11023    0.00000   -0.00000
    0.00000   -0.04297    0.00005
    0.00000    0.00000    0.01267

eig(A)

ans =

    0.012671
    0.042971
    0.110228
    0.292491
    0.457455
    0.761467
```

```
    1.162016
    1.988698
    2.474165
   28.260409
```

Okay, lets do it 100 more times.

```
for k=[1:100]
  [Q R] = qr(R*Q);
end
R

R =

 Columns 1 through 7:

  -28.26041   -0.00000   -0.00000   -0.00000    0.00000   -0.00000   -0.00000
    0.00000   -2.47417   -0.00000    0.00000   -0.00000    0.00000   -0.00000
    0.00000    0.00000   -1.98870   -0.00000    0.00000   -0.00000   -0.00000
    0.00000    0.00000    0.00000   -1.16202    0.00000   -0.00000   -0.00000
    0.00000    0.00000    0.00000    0.00000   -0.76147    0.00000   -0.00000
    0.00000    0.00000    0.00000    0.00000    0.00000   -0.45745    0.00000
    0.00000    0.00000    0.00000    0.00000    0.00000    0.00000   -0.29249
    0.00000    0.00000    0.00000    0.00000    0.00000    0.00000    0.00000
    0.00000    0.00000    0.00000    0.00000    0.00000    0.00000    0.00000
    0.00000    0.00000    0.00000    0.00000    0.00000    0.00000    0.00000

 Columns 8 through 10:

   -0.00000   -0.00000   -0.00000
    0.00000    0.00000   -0.00000
   -0.00000   -0.00000    0.00000
    0.00000   -0.00000    0.00000
   -0.00000   -0.00000   -0.00000
    0.00000    0.00000    0.00000
    0.00000   -0.00000    0.00000
   -0.11023   -0.00000    0.00000
    0.00000   -0.04297    0.00000
    0.00000    0.00000    0.01267
```

We see that the matrix R is a diagonal matrix with ± 1 times the eigenvalues on the diagonal. If the MATLAB/Octave `qr` function returned the factorization we carried out in class (where the diagonal entries of R are always positive), this calculation would return the eigenvalues on the diagonal. In this situation, Q has ± 1 on the diagonal (as you can check), and the sign will indicate whether we need to flip the entry in R to get the eigenvalue. A more coherent way of saying this is that QR will be diagonal with the eigenvalues on the diagonal. Let's check

```
Q*R

ans =
```

```
Columns 1 through 7:

  28.26041    0.00000    0.00000    0.00000   -0.00000    0.00000    0.00000
   0.00000    2.47417    0.00000   -0.00000    0.00000   -0.00000    0.00000
   0.00000    0.00000    1.98870    0.00000   -0.00000    0.00000    0.00000
  -0.00000    0.00000   -0.00000    1.16202   -0.00000    0.00000    0.00000
   0.00000   -0.00000    0.00000   -0.00000    0.76147   -0.00000    0.00000
  -0.00000    0.00000   -0.00000    0.00000    0.00000    0.45745   -0.00000
   0.00000   -0.00000    0.00000   -0.00000   -0.00000   -0.00000    0.29249
   0.00000    0.00000   -0.00000    0.00000    0.00000   -0.00000   -0.00000
  -0.00000   -0.00000    0.00000    0.00000   -0.00000    0.00000    0.00000
   0.00000    0.00000   -0.00000   -0.00000    0.00000   -0.00000   -0.00000

Columns 8 through 10:

   0.00000    0.00000    0.00000
  -0.00000   -0.00000    0.00000
   0.00000    0.00000   -0.00000
  -0.00000    0.00000   -0.00000
   0.00000    0.00000    0.00000
  -0.00000   -0.00000   -0.00000
  -0.00000    0.00000   -0.00000
   0.11023    0.00000   -0.00000
  -0.00000    0.04297   -0.00000
   0.00000   -0.00000    0.01267
```

8. **If U_1 and U_2 are unitary matrices, is $U_1 U_2$ a unitary matrix too?**

   Yes. We only need to verify that $U_1 U_2$ preserves the lengths of vectors. But $\|U_1 U_2 \mathbf{w}\| = \|U_2 \mathbf{w}\|$ for every $\mathbf{w}$ (because U_1 is unitary) and $\|U_2 \mathbf{w}\| = \|\mathbf{w}\|$ for every $\mathbf{w}$ (because U_2 is unitary). Thus $\|U_1 U_2 \mathbf{w}\| = \|\mathbf{w}\|$ for every $\mathbf{w}$.

9. **If $\mathbf{q}_1, \ldots, \mathbf{q}_n$ is an orthonormal basis for $\mathbb{C}^n$ do the complex conjugated vectors $\overline{\mathbf{q}}_1, \ldots, \overline{\mathbf{q}}_n$ form an orthonormal basis as well? Give a reason.** Yes. We have

   $$\langle \overline{\mathbf{q}}_i, \overline{\mathbf{q}}_j \rangle = \mathbf{q}_i^T \overline{\mathbf{q}}_j = \overline{\overline{\mathbf{q}}_i^T \mathbf{q}_j} = \overline{\langle \mathbf{q}_i, \mathbf{q}_j \rangle}$$

   Thus, if $\mathbf{q}_1, \ldots, \mathbf{q}_n$ is an orthonormal basis then

   $$\langle \overline{\mathbf{q}}_i, \overline{\mathbf{q}}_j \rangle = \begin{cases} \overline{1} = 1 & i = j \\ \overline{0} = 0 & i \neq j \end{cases}$$

   This shows that the vectors $\overline{\mathbf{q}}_1, \ldots, \overline{\mathbf{q}}_n$ form an orthonormal basis too.