18.06 Problem Set 6 - Solutions

Due Wednesday, 24 October 2007 at 4 pm in 2-106.

Problem 1: (10=3+3+4) Do problem 4 from section 4.4 (P 228) in your book.

$\boxed{\text{Solution}}$ (a) Example: $Q = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}$, then it has orthonormal columns but

$$QQ^T = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \neq I.$$

Any such example must be an $m \times n$ matrices with $m > n$, columns are orthonormal. (If $m < n$, then there is no possible to find n orthonormal m-vectors; if $m = n$, then by definition we have $Q^T Q = I$, and thus must have $QQ^T = I$; if $m > n$, then QQ^T is an $m \times m$ matrix whose rank is no more than n, thus cannot be the identity matrix.)

(b) Example: $\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $\mathbf{v}_2 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$. Then they are orthogonal but linearly dependent.

(One of the vectors must be zero vector: two linearly dependent vectors will always lie in the same line; and if both are nonzero, their inner product cannot be zero.)

(c) Example: $\mathbf{v}_1 = \begin{pmatrix} 1/2 \\ 1/2 \\ 1/2 \\ 1/2 \end{pmatrix}$, $\mathbf{v}_2 = \begin{pmatrix} 1/2 \\ 1/2 \\ -1/2 \\ -1/2 \end{pmatrix}$, $\mathbf{v}_3 = \begin{pmatrix} 1/2 \\ -1/2 \\ 1/2 \\ -1/2 \end{pmatrix}$, $\mathbf{v}_4 = \begin{pmatrix} -1/2 \\ 1/2 \\ 1/2 \\ -1/2 \end{pmatrix}$. It is easy to check that they are an orthonormal basis for $\mathbb{R}^4$.

(Other exampls: add a negetive sign to fixed position(s) of each vector above, for example, $\mathbf{v}_1 = \begin{pmatrix} -1/2 \\ 1/2 \\ 1/2 \\ 1/2 \end{pmatrix}$, $\mathbf{v}_2 = \begin{pmatrix} -1/2 \\ 1/2 \\ -1/2 \\ -1/2 \end{pmatrix}$, $\mathbf{v}_3 = \begin{pmatrix} -1/2 \\ -1/2 \\ 1/2 \\ -1/2 \end{pmatrix}$, $\mathbf{v}_4 = \begin{pmatrix} 1/2 \\ 1/2 \\ 1/2 \\ -1/2 \end{pmatrix}$. In fact, all possible solutions comes from the above example by this way.)

Problem 2: (20=15+5) Apply the Gram-Schmidt algorithm to find an orthonormal basis for the subspace U of $\mathbb{R}^4$ spanned by the vectors:

$$\mathbf{v}_1 = (1,1,1,1), \mathbf{v}_2 = (1,1,2,4), \mathbf{v}_3 = (1,2,-4,-3).$$

Write down the QR decomposition of the matrix A whose columns are these three vectors.

$\boxed{\text{Solution}}$ We first choose $\mathbf{V}_1 = \mathbf{v}_1 = \begin{pmatrix}1\\1\\1\\1\end{pmatrix}$.

The second vector would be

$$\mathbf{V}_2 = \mathbf{v}_2 - \frac{\mathbf{V}_1^T\mathbf{v}_2}{\mathbf{V}_1^T\mathbf{V}_1}\mathbf{V}_1 = \begin{pmatrix}1\\1\\2\\4\end{pmatrix} - \frac{8}{4}\begin{pmatrix}1\\1\\1\\1\end{pmatrix} = \begin{pmatrix}-1\\-1\\0\\2\end{pmatrix}.$$

The third vector would be

$$\mathbf{V}_3 = \mathbf{v}_3 - \frac{\mathbf{V}_1^T\mathbf{v}_3}{\mathbf{V}_1^T\mathbf{V}_1}\mathbf{V}_1 - \frac{\mathbf{V}_2^T\mathbf{v}_3}{\mathbf{V}_2^T\mathbf{V}_2}\mathbf{V}_2 = \begin{pmatrix}1\\2\\-4\\-3\end{pmatrix} - \frac{-4}{4}\begin{pmatrix}1\\1\\1\\1\end{pmatrix} - \frac{-9}{6}\begin{pmatrix}-1\\-1\\0\\2\end{pmatrix} = \begin{pmatrix}1/2\\3/2\\-3\\1\end{pmatrix}.$$

Finally we normize them:

$$\mathbf{q}_1 = \begin{pmatrix}1/2\\1/2\\1/2\\1/2\end{pmatrix}, \mathbf{q}_2 = \begin{pmatrix}-1/\sqrt{6}\\-1/\sqrt{6}\\0\\2/\sqrt{6}\end{pmatrix}, \mathbf{q}_3 = \begin{pmatrix}\sqrt{2}/10\\3\sqrt{2}/10\\-6\sqrt{2}/10\\2\sqrt{2}/10\end{pmatrix}.$$

The QR decomposition is given by

$$A = QR = \begin{pmatrix}1/2 & -1/\sqrt{6} & \sqrt{2}/10\\1/2 & -1/\sqrt{6} & 3\sqrt{2}/10\\1/2 & 0 & -6\sqrt{2}/10\\1/2 & 2/\sqrt{6} & 2\sqrt{2}/10\end{pmatrix}\begin{pmatrix}2 & 4 & -2\\0 & \sqrt{6} & -3\sqrt{6}/2\\0 & 0 & 5\sqrt{2}/2\end{pmatrix}.$$

Problem 3: (25=5+5+8+7) In the Gram-Schmidt algorithm, at each step we subtract the projection of one vector onto the previous vectors, in order to make them orthogonal. The key operation is the inner product $\mathbf{x}^T\mathbf{y}$, sometimes denoted $\mathbf{x}\cdot\mathbf{y}$ or $\langle\mathbf{x},\mathbf{y}\rangle$. We can apply the same process to *any* vector space as long as we define a suitable "inner product" that obeys the same algebraic rules. The key rules that a inner product must obey (for real vector spaces) are:

$$\begin{aligned}
(a) &\quad \langle\mathbf{x}_1+\mathbf{x}_2,\mathbf{y}\rangle = \langle\mathbf{x}_1,\mathbf{y}\rangle + \langle\mathbf{x}_2,\mathbf{y}\rangle;\\
(b) &\quad \langle c\mathbf{x},\mathbf{y}\rangle = c\langle\mathbf{x},\mathbf{y}\rangle;\\
(c) &\quad \langle\mathbf{x},\mathbf{y}\rangle = \langle\mathbf{y},\mathbf{x}\rangle;\\
(d) &\quad \langle\mathbf{x},\mathbf{x}\rangle \ge 0 \text{ for all } \mathbf{x}, \text{ and } \langle\mathbf{x},\mathbf{x}\rangle = 0 \text{ if and only if } \mathbf{x}=0.
\end{aligned}$$

In particular, instead of the vector space $\mathbf{R}^m$ of column vectors, consider instead the vector space V of real-coefficient polynomial functions $f(x)$, $g(x)$, etc.

(1) Show that $\langle f,g\rangle := \int_0^1 f(x)g(x)\,dx$ defines an inner product on V, i.e., check that it satisfies the above four properties.

$\boxed{\text{Solution}}$ Property (a) and (b) comes from the linearity of integrals, property (c) is obvious since $f(x)g(x) = g(x)f(x)$, and the first part of property (d) is also obvious since $f(x)^2 \ge 0$ for any real function f. Finally we check the second part of property (d): Suppose $\langle f,f\rangle = 0$, where f is a polynomial. We can rewrite this as

$$\int_0^1 f(x)^2dx = 0.$$

Since f is a polynomial, in particular f is continuous, we must have $f(x) = 0$. This completes the proof.

(2) Show that the polynomials $f = 1+x$ and $g = 5-9x$ are orthogonal.

$\boxed{\text{Solution}}$ We compute

$$\begin{aligned}
\langle f,g\rangle &= \int_0^1 (1+x)(5-9x)dx\\
&= \int_0^1 (5-4x-9x^2)dx\\
&= \left(5x-2x^2-3x^3\right)\Big|_0^1\\
&= 0.
\end{aligned}$$

This the polynomial f is orthogonal to the polynomial g.

(3) Apply the Gram-Schmidt algorithm to the set $\{1, x, x^2\}$ to obtain an orthonormal basis $\{f_0, f_1, f_2\}$ of all degree-2 polynomials.

$\boxed{\text{Solution}}$ Denote $g_0 = 1, g_1 = x$ and $g_2 = x^2$.

We begin by letting $G_0 = g_0 = 1$.

For G_1:

$$\begin{aligned}
G_1 &= g_1 - \frac{\langle G_0, g_1 \rangle}{\langle G_0, G_0 \rangle} G_0 \\
&= x - \frac{\int_0^1 x dx}{\int_0^1 dx} 1 \\
&= x - \frac{1}{2}.
\end{aligned}$$

Now G_2:

$$\begin{aligned}
G_2 &= g_2 - \frac{\langle G_0, g_2 \rangle}{\langle G_0, G_0 \rangle} G_0 - \frac{\langle G_1, g_2 \rangle}{\langle G_1, G_1 \rangle} G_1 \\
&= x^2 - \frac{\int_0^1 x^2 dx}{\int_0^1 dx} 1 - \frac{\int_0^1 (x - \frac{1}{2}) x^2 dx}{\int_0^1 (x - \frac{1}{2})^2 dx} (x - \frac{1}{2}) \\
&= x^2 - \frac{1}{3} - \frac{1/12}{1/12} (x - \frac{1}{2}) \\
&= x^2 - x + \frac{1}{6}.
\end{aligned}$$

Finally we normalize them and get

$$\begin{aligned}
f_0 &= \frac{G_0}{\|G_0\|} = \frac{1}{\sqrt{\int_0^1 1 dx}} = \frac{1}{1} \\
&= 1, \\
f_1 &= \frac{G_1}{\|G_1\|} = \frac{x - \frac{1}{2}}{\sqrt{\int_0^1 (x - \frac{1}{2})^2 dx}} = \frac{x - \frac{1}{2}}{1/2\sqrt{3}} \\
&= 2\sqrt{3}x - \sqrt{3}, \\
f_2 &= \frac{G_2}{\|G_2\|} = \frac{x^2 - x + \frac{1}{6}}{\sqrt{\int_0^1 (x^2 - x + \frac{1}{6})^2 dx}} = \frac{x^2 - x + \frac{1}{6}}{1/6\sqrt{5}} \\
&= 6\sqrt{5}x^2 - 6\sqrt{5}x + \sqrt{5}.
\end{aligned}$$

(4) Do the same thing *approximately* in Matlab, approximating a polynomial by its values over a set of 1000 discrete points, and the integral by a summation:

```
x = linspace(0,1,1000)';
A = [x.^0, x.^1, x.^2, x.^3, x.^4, x.^5];
```

That is, the columns of A are x^0, x^1, ..., x^5 (discretized). Matlab will do Gram-Schmidt for us via the function `qr` (passing zero as the second argument to `qr` will just do Gram-Schmidt of a non-square matrix rather than trying to construct a square orthogonal Q):

```
[Q,R] = qr(A, 0); Q = Q * sqrt(999);
```

The $\sqrt{999}$ factor is to change the normalization to match the approximate "integral" inner product rather than the ordinary dot product. (Why? Think about how the dot product compares to the approximate discretized integral.) Now plot the columns of Q versus x in order to see the orthogonal "polynomials" up to degree 5.

```
plot(x, Q)
```

Verify that they match your answers from part (3), up to degree 2 of course, within the numerical error, by plotting the curves on top of one another. (You can superimpose plots in Matlab by typing `hold on` at the Matlab prompt, which makes subsequent plot commands plot on top of one another; to stop superimposing, type `hold off`.)

$\boxed{\text{Solution}}$ The codes

```
>> x=linspace(0,1,1000)';
>> A=[x.^0, x.^1, x.^2, x.^3, x.^4, x.^5];
>> [Q,R]=qr(A,0);
>> Q=Q*sqrt(1000);
>> Q(:,1)=-Q(:,1);
>> Q(:,3)=-Q(:,3);
>> Q(:,5)=-Q(:,5);
>> plot(x,Q)
```

We changed the sign of columns 1, 3, 5 since the program choose the negative square root as the length. The picture is

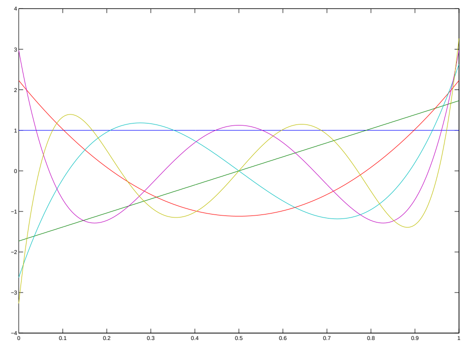

Figure 1: x-vs-Q

Now we plot the graphs of the polynomials we get, and then put the above graph together:

```
>> B=[x.^0]; C=[2*sqrt(3)*x.^1-sqrt(3)* x.^0];
>> D=[6*sqrt(5)*x.^2-6*sqrt(5)*x.^1+sqrt(5)*x.^0];
>> plot(x,B), hold on, plot(x,C), plot(x, D)
>> plot(x,Q)
```

The graphs are

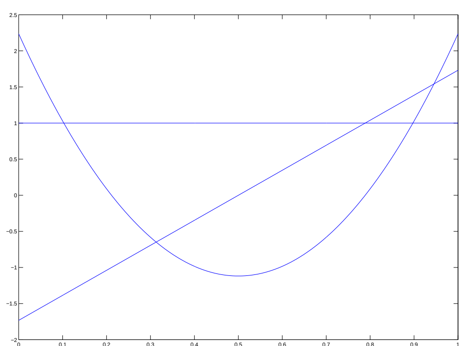

Figure 2: f_0, f_1, f_2

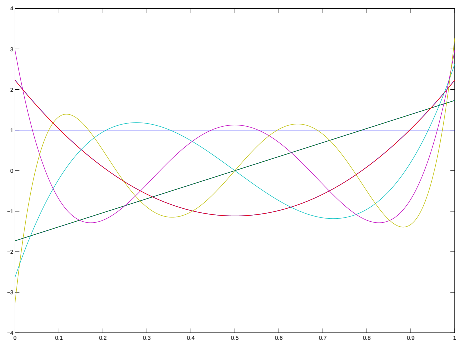

Figure 3: f_0, f_1, f_2 and x-vs-Q

The third graph coincides with the first one, which means that our functions in (3) coincides with those from (4).

Problem 4: (10=2+2+2+2+2) True or False: (Give reasons)
Suppose all matrices below are square matrices.

(1) $\det(-A) = -\det(A)$.

$\boxed{\text{Solution}}$ False.

Suppose A is $n \times n$ matrix, then we always have $\det(-A) = (-1)^n \det(A)$. Thus the above formula only holds for A of odd order, and fails for A of even order.

(2) $\det(A + B) = \det(A) + \det(B)$.

$\boxed{\text{Solution}}$ False.

Example: This is almost never true. As a simple example, we can take $A = B$ to be any nonsingular $n \times n$ matrix with $n > 1$, then

$$\det(A + B) = \det(2A) = 2^n \det(A) \neq 2\det(A) = \det(A) + \det(B).$$

(3) $\det(A - B) = \det(A) - \det(B)$.

$\boxed{\text{Solution}}$ False.

This is acturally the same as above. For example, we may take $A = -B$ to be any $n \times n$ ($n > 1$) nonsingular matrix.

(4) $\det(AB) = \det(A)\det(B)$.

$\boxed{\text{Solution}}$ True.

This is property 9 in book.

(5) $\det(A^{-1}) = \det(A)^{-1}$.

$\boxed{\text{Solution}}$ True. (We assume A is invertible, otherwise both sides are of no sense.)

We have $1 = \det(I) = \det(A)\det(A^{-1})$, which implies $\det(A^{-1}) = \det(A)^{-1}$.

Problem 5: (10) Calculate the determinant of the following 6×6 matrix:

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

$\boxed{\text{Solution}}$ We use basic properties in §5.1.

$$\det A = \det \begin{pmatrix} 1 & 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 & 1 \end{pmatrix} \overset{1}{=} \det \begin{pmatrix} 1 & 1 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 & 0 & -1 \\ -1 & 0 & 0 & 1 & 0 & 0 \\ 0 & -1 & 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 & 0 \end{pmatrix}$$

$$\overset{2}{=} \det \begin{pmatrix} 1 & 1 & 1 & 5 & 1 & 1 \\ 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \\ -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 \end{pmatrix} \overset{3}{=} \det \begin{pmatrix} 0 & 0 & 0 & 5 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \\ -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 \end{pmatrix}$$

$$\overset{4}{=} (-1)^3 \det \begin{pmatrix} 5 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \end{pmatrix} \overset{5}{=} 5,$$

where in step $\stackrel{1}{=}$ we minus the first row from the second through sixth row, in step $\stackrel{3}{=}$ we add columns 1, 2, 3,5, 6 to column 4, in step $\stackrel{3}{=}$ we minus $\frac{1}{5}$ times column 4 from columns 1, 2, 3, 5, 6, and in step $\stackrel{4}{=}$ we exchange column 1 with column 4, column 2 with column 5, column 3 with column 6. (You can use other orders of these operations or similar operations. However, both the big formula and the cofactor formula are not good for this problem – they encouter too much computations!)

Problem 6: (25=5+8+5+7) Construct a random symmetric 5×5 matrix A in Matlab:

```
A = rand(5,5); A = A' * A
```

Compute the QR decomposition of A, via:

```
[Q,R] = qr(A);
```

(a) Verify that $Q^T Q = I$ (that Q is orthogonal) and that $QR = A$, up to roundoff error (about 10^{-16}).

$\boxed{\text{Solution}}$ The codes

```
>> A=rand(5,5); A=A'*A;
>> [Q,R]=qr(A);
>> Q'*Q-eye(5)

ans =

   1.0e-15 *

         0   -0.1249    0.0590   -0.0902   -0.1388
   -0.1249    0.4441   -0.2290         0    0.2220
    0.0590   -0.2290    0.2220    0.0104   -0.0416
   -0.0902         0    0.0104   -0.3331    0.0278
   -0.1388    0.2220   -0.0416    0.0278   -0.1110
```

```
>> Q*R-A

ans =

   1.0e-15 *

    0.3331         0         0    0.2220    0.2220
         0   -0.4441         0   -0.2220    0.2220
   -0.1110   -0.2220   -0.2220         0   -0.2220
         0    0.2220    0.2220    0.4441         0
         0         0    0.2220         0    0.1110
```

(b) Compute

```
B = R * Q
```

Notice that B is symmetric. Why? (Hint: write R in terms of A.)

$\boxed{\text{Solution}}$ The codes

```
>> B=R*Q

B =

    6.2997    0.6777    0.0176    0.1072   -0.0004
    0.6777    0.7624    0.0781   -0.1105   -0.0012
    0.0176    0.0781    0.0552   -0.0200   -0.0001
    0.1072   -0.1105   -0.0200    0.1433   -0.0004
   -0.0004   -0.0012   -0.0001   -0.0004    0.0017
```

B is always symmetric: since

$$B = RQ = Q^T AQ,$$

and

$$A = A^T$$

is symmetric, we have

$$B^T = Q^T A^T Q = Q^T A Q = B$$

is symmetric.

(c) Repeat this process. Compute:

```
[Q,R] = qr(B); B = R * Q
```

over and over again. You should find that, after repeating this a number of times, the result stops changing (if you ignore numbers smaller than 10^{-15}. Compute

```
eig(A)
```

which prints the eigenvalues of A (something you learned in 18.03, and we'll study in more detail soon). Compare this to the B resulting from the above process. How are they related?

$\boxed{\text{Solution}}$ Codes

```
>> [Q,R]=qr(B);B=R*Q

B =

    6.3820    0.0727   -0.0002    0.0020    0.0000
    0.0727    0.7166    0.0089   -0.0204    0.0000
   -0.0002    0.0089    0.0487   -0.0142   -0.0000
    0.0020   -0.0204   -0.0142    0.1134    0.0000
    0.0000    0.0000   -0.0000    0.0000    0.0016

>> [Q,R]=qr(B);B=R*Q

B =

    6.3829    0.0081   -0.0000    0.0000   -0.0000
    0.0081    0.7165    0.0019   -0.0030   -0.0000
   -0.0000    0.0019    0.0609   -0.0289    0.0000
    0.0000   -0.0030   -0.0289    0.1004   -0.0000
   -0.0000   -0.0000    0.0000   -0.0000    0.0016
```

```
>> [Q,R]=qr(B);B=R*Q;
>> [Q,R]=qr(B);B=R*Q;
>> [Q,R]=qr(B);B=R*Q;
>> [Q,R]=qr(B);B=R*Q;
>> [Q,R]=qr(B);B=R*Q;
>> [Q,R]=qr(B);B=R*Q;
>> [Q,R]=qr(B);B=R*Q;
>> [Q,R]=qr(B);B=R*Q;
>> [Q,R]=qr(B);B=R*Q;
>> [Q,R]=qr(B);B=R*Q;
>> [Q,R]=qr(B);B=R*Q;
>> [Q,R]=qr(B);B=R*Q;
>> [Q,R]=qr(B);B=R*Q;
>> [Q,R]=qr(B);B=R*Q;
>> [Q,R]=qr(B);B=R*Q;
>> [Q,R]=qr(B);B=R*Q;
>> [Q,R]=qr(B);B=R*Q;
>> [Q,R]=qr(B);B=R*Q;
>> [Q,R]=qr(B);B=R*Q;
>> [Q,R]=qr(B);B=R*Q

B =

    6.3830   -0.0000    0.0000    0.0000   -0.0000
    0.0000    0.7165    0.0000   -0.0000    0.0000
   -0.0000    0.0000    0.1156   -0.0000   -0.0000
    0.0000   -0.0000   -0.0000    0.0457   -0.0000
   -0.0000   -0.0000    0.0000    0.0000    0.0016

>> [Q,R]=qr(B);B=R*Q

B =

    6.3830   -0.0000    0.0000    0.0000    0.0000
    0.0000    0.7165    0.0000   -0.0000   -0.0000
   -0.0000    0.0000    0.1156   -0.0000    0.0000
    0.0000   -0.0000   -0.0000    0.0457    0.0000
    0.0000    0.0000   -0.0000   -0.0000    0.0016
```

```
>> eig(A)

ans =

    0.0016
    0.0457
    0.1156
    0.7165
    6.3830
```

B will finally be an diagonal matrix, with diagonal entries the eigenvalues of A.

(d) Explain why, if B ever becomes a completely diagonal matrix, it should not change at all when we do the step in part (c).

$\boxed{\text{Solution}}$ If B is a completely diagonal matrix, then its column are already orthogonal and are a scaled version of the identity matrix. Therefore Gram-Schmidt will just give you $Q = I$. Thus $R = Q^T B = B$ and $RQ = BI = B$: the matrix doesn't change.