1 Classical & Modified Gram-Schmidt

In this lecture, we'll talk about the classical and modified Gram-Schmidt algorithms.

1.1 Classical Algorithm

Given A, we want to find $\hat{Q}$ and $\hat{R}$ such that

$$A = \hat{Q}\hat{R}. \tag{1}$$

Note that we can rewrite (1) in the form

$$\begin{bmatrix} a_1 & a_2 & \dots & a_m \end{bmatrix} = \begin{bmatrix} q_1 & q_2 & \dots & q_m \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & \dots & r_{1m} \\ 0 & r_{22} & \dots & r_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & r_{mm} \end{bmatrix}.$$

The formula to finding these entries are

$$a_m = q_1 r_{1m} + q_2 r_{2m} + \ldots + q_m r_{mm}.$$

$$r_{ji} = \begin{cases} \langle a_i, q_j \rangle & j < i \\ \left|\left| a_i - \sum_{k=1}^{i-1} r_{ki} q_k \right|\right|_2 & j = i \\ 0 & j > i \end{cases}.$$

$$q_i = \frac{a_i - \sum_{j=1}^{i-1} r_{ji} q_j}{r_{ii}}.$$

a_i is a vector and r_{ij} is a scalar.

1.1.1 Worked Example

(Example.) Suppose we have

$$A = \begin{bmatrix} -1 & -1 & 1 \\ 1 & 3 & 3 \\ -1 & -1 & 5 \\ 1 & 3 & 7 \end{bmatrix}.$$

We can define

$$\vec{a_1} = \begin{bmatrix} -1 \\ 1 \\ -1 \\ 1 \end{bmatrix} \quad \vec{a_2} = \begin{bmatrix} -1 \\ 3 \\ -1 \\ 3 \end{bmatrix} \quad \vec{a_3} = \begin{bmatrix} 1 \\ 3 \\ 5 \\ 7 \end{bmatrix}.$$

In other words, our goal is to get something like

$$A = \begin{bmatrix} a_1 & a_2 & a_3 \end{bmatrix} = \begin{bmatrix} q_1 & q_2 & q_3 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ 0 & r_{22} & r_{23} \\ 0 & 0 & r_{33} \end{bmatrix}.$$

Then, we can find the elements of $\hat{Q}$ and $\hat{R}$.

$$a_1 = q_1 r_{11} \implies q_1 = \frac{a_1}{r_{11}}.$$

Since q_1 is a unit vector (remember that the q_i's are in an orthonormal set), it follows that $||q_1||_2 = 1$. Then,

$$r_{11} = ||a_1||_2 = \sqrt{(-1)^2 + 1^2 + (-1)^2 + 1^2} = 2.$$

Notice that

$$q_1 = \frac{a_1}{2} = \frac{1}{2}\begin{bmatrix} -1 \\ 1 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \end{bmatrix}.$$

$$a_2 = q_1 r_{12} + q_2 r_{22}.$$

Because q_1 and q_2 are orthonormal, we know that $\langle q_2, q_2 \rangle = 1$ and $\langle q_1, q_2 \rangle = 0$. So,

$$\langle a_2, q_1 \rangle = \langle q_1 r_{12} + q_2 r_{22}, q_1 \rangle = r_{12} \underbrace{\langle q_1, q_1 \rangle}_{1} + r_{22} \underbrace{\langle q_2, q_1 \rangle}_{0} = r_{12}$$

Then,

$$\begin{aligned} r_{12} = \langle a_2, q_1 \rangle &= \begin{bmatrix} -1 & 3 & -1 & 3 \end{bmatrix} \begin{bmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \end{bmatrix} \\ &= (-1)\left(-\frac{1}{2}\right) + (3)\left(\frac{1}{2}\right) + (-1)\left(-\frac{1}{2}\right) + (3)\left(\frac{1}{2}\right) \\ &= 4. \end{aligned}$$

Now, we need to find

$$q_2 = \frac{a_2 - \sum_{j=1}^{1} r_{ji} q_j}{r_{22}}.$$

- Note that

$$\begin{aligned} r_{22} &= \left|\left| a_2 - \sum_{k=1}^{1} r_{k2} q_k \right|\right|_2 = ||a_2 - r_{12} q_1||_2 \\ &= \left|\left| \begin{bmatrix} -1 \\ 3 \\ -1 \\ 3 \end{bmatrix} - 4 \begin{bmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \end{bmatrix} \right|\right|_2 = \left|\left| \begin{bmatrix} -1 \\ 3 \\ -1 \\ 3 \end{bmatrix} - \begin{bmatrix} -2 \\ 2 \\ -2 \\ 2 \end{bmatrix} \right|\right|_2 \\ &= \left|\left| \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \right|\right|_2 = \sqrt{1^2 + 1^2 + 1^2 + 1^2} = \sqrt{4} = 2. \end{aligned}$$

- From this, it follows that

$$q_2 = \frac{a_2 - r_{12} q_1}{r_{22}} = \frac{1}{r_{22}}(a_2 - r_{12} q_1) = \frac{1}{2}\left(\begin{bmatrix} -1 \\ 3 \\ -1 \\ 3 \end{bmatrix} - 4 \begin{bmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \end{bmatrix} \right) = \frac{1}{2}\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{bmatrix}.$$

$$a_3 = r_{13}q_1 + r_{23}q_2 + r_{33}q_3.$$

At this point, we know that

$$q_1 = \begin{bmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \end{bmatrix} \quad q_2 = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{bmatrix} \quad a_3 = \begin{bmatrix} 1 \\ 3 \\ 5 \\ 7 \end{bmatrix}.$$

Additionally, $\langle q_1, q_3 \rangle = \langle q_2, q_3 \rangle = 0$ while $\langle q_3, q_3 \rangle = 1$. From there, we have

$$r_{13} = \langle a_3, q_1 \rangle = \begin{bmatrix} 1 & 3 & 5 & 7 \end{bmatrix} \begin{bmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \end{bmatrix} = 1\left(-\frac{1}{2}\right) + 3\left(\frac{1}{2}\right) + 5\left(-\frac{1}{2}\right) + 7\left(\frac{1}{2}\right) = 2$$

and

$$r_{23} = \langle a_3, q_2 \rangle = \begin{bmatrix} 1 & 3 & 5 & 7 \end{bmatrix} \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{bmatrix} = 8$$

and

$$\begin{aligned} r_{33} &= \left|\left| a_3 - \sum_{k=1}^{2} r_{k3}q_k \right|\right|_2 \\ &= ||a_3 - (r_{13}q_1 + r_{23}q_2)||_2 \\ &= \left|\left| \begin{bmatrix} 1 \\ 3 \\ 5 \\ 7 \end{bmatrix} - \left(2\begin{bmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \end{bmatrix} + 8\begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{bmatrix} \right) \right|\right|_2 \\ &= \left|\left| \begin{bmatrix} -2 \\ -2 \\ 2 \\ 2 \end{bmatrix} \right|\right|_2 \\ &= \sqrt{(-2)^2 + (-2)^2 + 2^2 + 2^2} \\ &= \sqrt{4+4+4+4} \\ &= 4. \end{aligned}$$

Finally,

$$q_3 = \frac{a_3 - \sum_{j=1}^{2} r_{j3}q_j}{r_{33}} = \frac{a_3 - (r_{13}q_1 + r_{23}q_2)}{r_{33}} = \frac{1}{r_{33}}(a_3 - (r_{13}q_1 + r_{23}q_2))$$
$$= \frac{1}{4}\begin{bmatrix} -2 \\ -2 \\ 2 \\ 2 \end{bmatrix} = \begin{bmatrix} -\frac{1}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{bmatrix}.$$

Notice that we're now done with the algorithm. In particular, we have

$$q_1 = \begin{bmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \end{bmatrix} \quad q_2 = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{bmatrix} \quad q_3 = \begin{bmatrix} -\frac{1}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{bmatrix},$$

and

$$r_{11} = 2 \quad r_{12} = 4 \quad r_{13} = 2 \quad r_{22} = 2 \quad r_{23} = 8 \quad r_{33} = 4.$$

This gives us the decomposition of

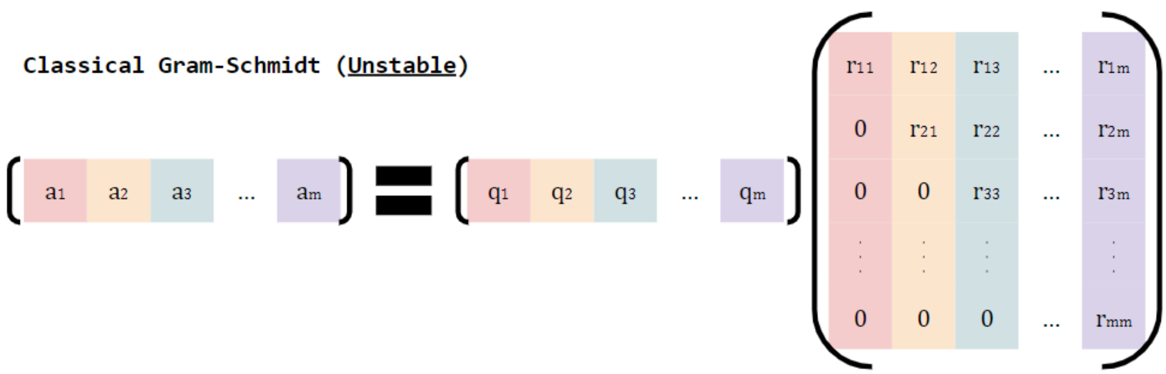
$$\underbrace{\begin{bmatrix} -1 & -1 & 1 \\ 1 & 3 & 3 \\ -1 & -1 & 5 \\ 1 & 3 & 7 \end{bmatrix}}_{A} = \underbrace{\begin{bmatrix} -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{bmatrix}}_{\hat{Q}} \underbrace{\begin{bmatrix} 2 & 4 & 2 \\ 0 & 2 & 8 \\ 0 & 0 & 4 \end{bmatrix}}_{\hat{R}}.$$

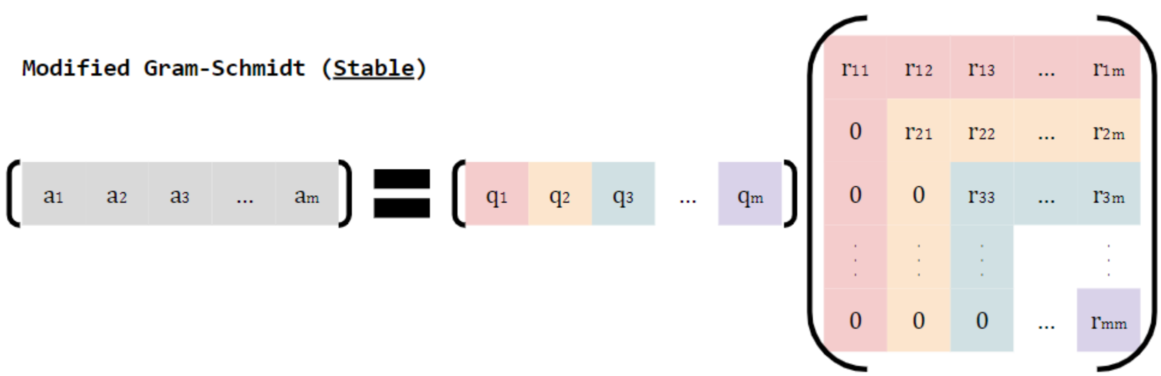
1.1.2 Summary

This is effectively how the **classical Gram-Schmidt** algorithm works. Notice how we went through each $\vec{a}_i$ entry (column by column) and found all the desired values of $\vec{q}_i$ and r_{ij}.

Sadly, the classical Gram-Schmidt is **unstable**[1]. For this reason, we'll introduce a *modified* Gram-Schmidt algorithm, which is **stable**. One notable difference is that

- The classical algorithm builds R one *column* at a time.
- The modified algorithm builds R one *row* at a time.





[1] A very small change of some entry in A can yield a significant difference in the resulting QR decomposition.

1.1.3 MATLAB Code

```
function [Q,R]=classicalGS(A)         % classical Gram-Schmidt

n = size(A,2);                        % number of columns; this formulation
                                      % does not need the number of rows
for i=1:n
Q(:,i) = A(:,i);                      % initialization

    for j=1:(i-1)
        R(j,i)=(A(:,i))'*Q(:,j);      % computing R(j,i) by going down the column
        Q(:,i)=Q(:,i)-R(j,i)*Q(:,j);  % updating Q(:,j)
    end

    R(i,i) = norm(Q(:,i));            % computing R(i,i)
    Q(:,i)=Q(:,i)/R(i,i);             % making Q(:,i) a unit vector
end
```

1.2 Modified Gram-Schmidt

As mentioned earlier, the modified Gram-Schmidt is a stable algorithm that builds R one row at a time.

1.2.1 MATLAB Code

```
function [Q,R]=modifiedGS(A)          % modified Gram-Schmidt

n = size(A,2);                        % number of columns; this formulation
                                      % does not need the number of rows
for i=1:n
    Q(:,i) = A(:,i);                  % initialization
end

for i=1:n
    R(i,i) = norm(Q(:,i));            % computing R(i,i)
    Q(:,i)=Q(:,i)/R(i,i);             % making Q(:,i) a unit vector

    for j=(i+1):n
        R(i,j)=(Q(:,i))'*Q(:,j);      % computing R(i,j) by going right on ith row
        Q(:,j)=Q(:,j)-R(i,j)*Q(:,i);  % updating Q(:,j)
    end

end
```

1.2.2 Worked Example

We'll base our algorithm on the MATLAB code above.

(Example.) We'll solve the same problem as in the previous example, *except* we'll use the modified algorithm instead. To reiterate, suppose we have

$$A = \begin{bmatrix} -1 & -1 & 1 \\ 1 & 3 & 3 \\ -1 & -1 & 5 \\ 1 & 3 & 7 \end{bmatrix}.$$

We can define

$$\vec{a_1} = \begin{bmatrix} -1 \\ 1 \\ -1 \\ 1 \end{bmatrix} \quad \vec{a_2} = \begin{bmatrix} -1 \\ 3 \\ -1 \\ 3 \end{bmatrix} \quad \vec{a_3} = \begin{bmatrix} 1 \\ 3 \\ 5 \\ 7 \end{bmatrix}.$$

In other words, our goal is to get something like

$$A = \begin{bmatrix} a_1 & a_2 & a_3 \end{bmatrix} = \begin{bmatrix} q_1 & q_2 & q_3 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ 0 & r_{22} & r_{23} \\ 0 & 0 & r_{33} \end{bmatrix}.$$

Keep this in mind, since we'll be assuming this. Then, we can find the elements of $\hat{Q}$ and $\hat{R}$. We'll run through the algorithm described in code above.

1. First, we define $q_1 = a_1$, $q_2 = a_2$, and $q_3 = a_3$.
2. Next, for we want to find r_{11}, r_{12}, r_{13} and then determine q_1. We can also update q_2 and q_3.
   - (Outer Loop: $i = 1$.) Note that

     $$r_{11} = ||q_1||_2 = ||a_1||_2 = \sqrt{(-1)^2 + 1^2 + (-1)^2 + 1^2} = 2.$$

     Here, $r_{11} = ||q_1||_2$ comes from the algorithm (since we *initially* set $q_1 = a_1$.)

     $$q_1 = \frac{q_1}{r_{11}} = \frac{q_1}{||a_1||_2} = \begin{bmatrix} -1/2 \\ 1/2 \\ -1/2 \\ 1/2 \end{bmatrix}.$$

     Here, we've updated the value of q_1.
   - In the *inner loop*, we do the following for $j = 2$ to 3:
     - (Inner Loop: $j = 2$.) Next, note that

       $$r_{12} = \langle q_1, q_2 \rangle = q_1^T q_2 = q_1^T a_2 = \left(-\frac{1}{2}\right)(-1) + \frac{1}{2}(3) + \left(-\frac{1}{2}\right)(-1) + \frac{1}{2}(3) = 4.$$

       From there, it follows that

       $$q_2 = q_2 - r_{12}q_1 = \begin{bmatrix} -1 \\ 3 \\ -1 \\ 3 \end{bmatrix} - 4 \begin{bmatrix} -1/2 \\ 1/2 \\ -1/2 \\ 1/2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}.$$

     - (Inner Loop: $j = 3$.) Next, we have

       $$r_{13} = \langle q_1, q_3 \rangle = q_1^T q_3 = q_1^T a_3 = \left(-\frac{1}{2}\right)(1) + \frac{1}{2}(3) + \left(-\frac{1}{2}\right)(5) + \frac{1}{2}7 = 2.$$

From there, it follows that

$$q_3 = q_3 - r_{13}q_1 = \begin{bmatrix}1\\3\\5\\7\end{bmatrix} - 2\begin{bmatrix}-1/2\\1/2\\-1/2\\1/2\end{bmatrix} = \begin{bmatrix}2\\2\\6\\6\end{bmatrix}.$$

3. After running through the first iteration of the outer loop discussed in the algorithm, we have

$$q_2 = \begin{bmatrix}1\\1\\1\\1\end{bmatrix} \quad q_3 = \begin{bmatrix}2\\2\\6\\6\end{bmatrix}.$$

   Now, we want to find r_{22}, r_{23}, and determine q_2. We also update q_3.

   - (Outer Loop: $i = 2$.) We have

$$r_{22} = ||q_2||_2 = \sqrt{1^2 + 1^2 + 1^2 + 1^2} = 2.$$

     So, updating q_2 gives us

$$q_2 = \frac{q_2}{r_{22}} = \begin{bmatrix}1/2\\1/2\\1/2\\1/2\end{bmatrix}.$$

   - In the *inner loop*, we do the following for $j = 3$ to 3:
     - (Inner Loop: $j = 3$.) We have

$$r_{23} = \langle q_2, q_3\rangle = q_2^T = q_3 = \frac{1}{2}(2) + \frac{1}{2}(2) + \frac{1}{2}(6) + \frac{1}{2}(6) = 8.$$

       From there,

$$q_3 = q_3 - r_{23}q_2 = \begin{bmatrix}2\\2\\6\\6\end{bmatrix} - 8\begin{bmatrix}1/2\\1/2\\1/2\\1/2\end{bmatrix} = \begin{bmatrix}-2\\-2\\2\\2\end{bmatrix}.$$

4. After running through the second iteration of the outer loop discussed in the algorithm, we have

$$q_3 = \begin{bmatrix}-2\\-2\\2\\2\end{bmatrix}.$$

   Now, we can find r_{33} and determine the value of q_3.

   - (Outer Loop: $i = 3$.) We have

$$r_{33} = ||q_3||_2 = \sqrt{(-2)^2 + (-2)^2 + 2^2 + 2^2} = 4.$$

     So,

$$q_3 = \frac{q_3}{r_{33}} = \begin{bmatrix}-1/2\\-1/2\\1/2\\1/2\end{bmatrix}.$$

- Notice that the inner loop is not executed since $j = 4$ is greater than 3.