Computational Data Mining

Part 10: Matrix Decompositions III

Gram-Schmidt process

Instructor: Mahdi Vasighi

Department of Computer Science and Information Technology,
Institute for Advanced Studies in Basic Sciences, Zanjan, Iran

Outline

- Gram-Schmidt process
- QR decompostion

Orthogonality

Gram-Schmidt process

Due to the advantages of orthonormal bases which is the straightforwardness of coordinate representation, we will use GS for finding such a basis.

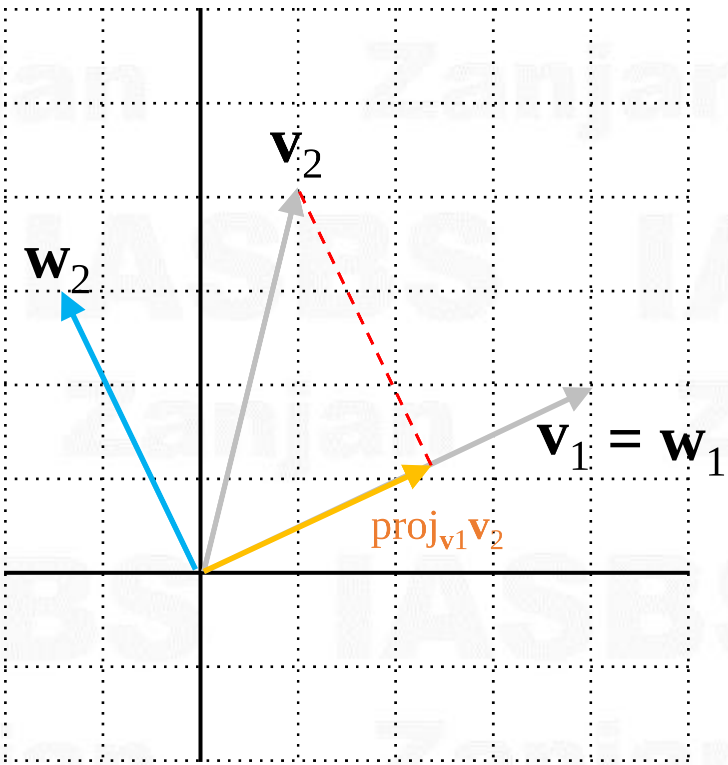
Non-normalized Classical Gram-Schmidt has three steps:

1. Begin with a basis for the inner product space.
2. Convert the basis to an orthogonal basis.
3. Normalize each vector in the orthogonal basis to form an orthonormal basis.

Orthogonality

Gram-Schmidt process

Let $\{\mathbf{v}_1, \mathbf{v}_2\}$ be a basis for R^2.
first choose one of the original vectors, say $\mathbf{v}_1$, and call it $\mathbf{w}_1$.



$$\mathbf{w}_1 = \mathbf{v}_1$$

Then we want to find a second vector orthogonal to $\mathbf{w}_1$.

$$\mathbf{w}_2 = \mathbf{v}_2 - \mathrm{proj}_{\mathbf{v}_1}\mathbf{v}_2$$

$$\mathbf{w}_2 = \mathbf{v}_2 - \frac{\mathbf{v}_2 \cdot \mathbf{w}_1}{\mathbf{w}_1 \cdot \mathbf{w}_1}\mathbf{w}_1$$

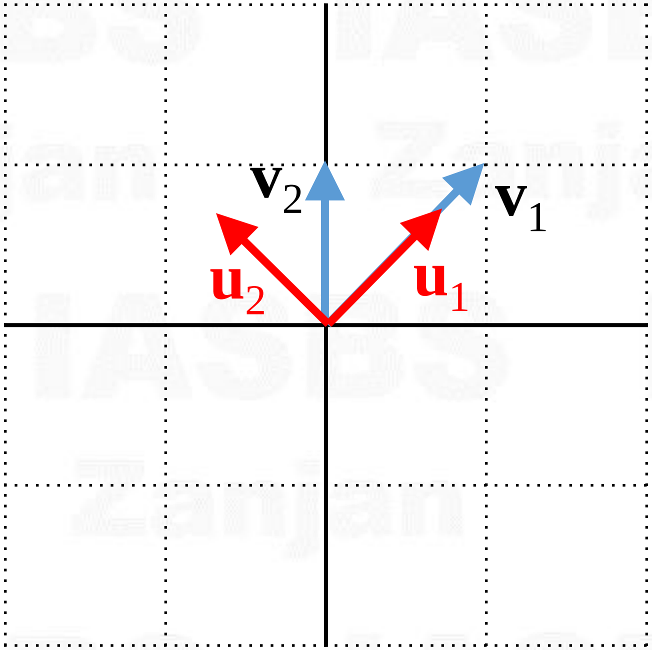
By normalizing $\mathbf{w}_1$ and $\mathbf{w}_2$, we obtain the orthonormal basis for R^2:

$$\{\mathbf{u}_1, \mathbf{u}_2\} = \left\{\frac{\mathbf{w}_1}{\|\mathbf{w}_1\|}, \frac{\mathbf{w}_2}{\|\mathbf{w}_2\|}\right\}$$

Orthogonality

Gram-Schmidt process

Example:



$$\mathbf{w}_1 = \mathbf{v}_1 = (1,1)$$

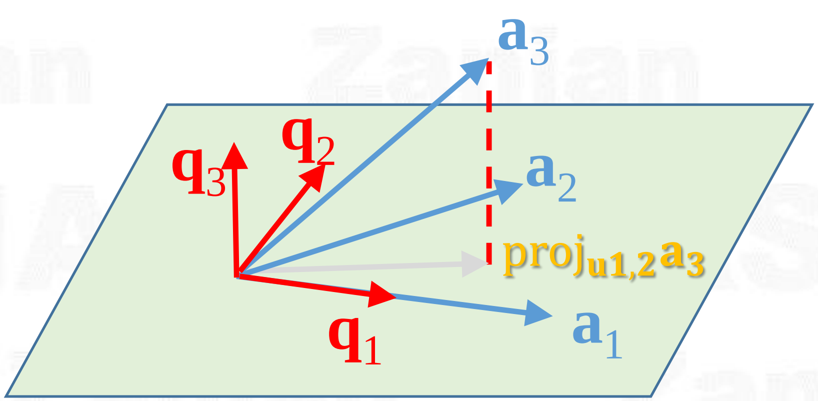
$$\mathbf{w}_2 = \mathbf{v}_2 - \frac{\mathbf{v}_2 \cdot \mathbf{w}_1}{\mathbf{w}_1 \cdot \mathbf{w}_1}\mathbf{w}_1 = (0,1) - \frac{1}{2}(1,1) = (-½, ½)$$

$$\mathbf{u}_1 = \frac{\mathbf{w}_1}{\|\mathbf{w}_1\|} = \frac{1}{\sqrt{2}}(1,1)$$

$$\mathbf{u}_2 = \frac{\mathbf{w}_2}{\|\mathbf{w}_2\|} = \sqrt{2}\left(-\frac{1}{2}, \frac{1}{2}\right) = \left(-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right)$$

Orthogonality

Gram-Schmidt process



$$\mathbf{A} = \{\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_n\}$$

$$\mathbf{u}_1 = \mathbf{a}_1 \,, \qquad \mathbf{q}_1 = \frac{\mathbf{u}_1}{\|\mathbf{u}_1\|}$$

$$\mathbf{u}_2 = \mathbf{a}_2 - (\mathbf{a}_2 \cdot \mathbf{u}_1)\mathbf{u}_1 \,, \qquad \mathbf{q}_2 = \frac{\mathbf{u}_2}{\|\mathbf{u}_2\|}$$

$$\mathbf{u}_3 = \mathbf{a}_3 - (\mathbf{a}_3 \cdot \mathbf{u}_1)\mathbf{u}_1 - (\mathbf{a}_3 \cdot \mathbf{u}_2)\mathbf{u}_2, \qquad \mathbf{q}_3 = \frac{\mathbf{u}_3}{\|\mathbf{u}_3\|}$$

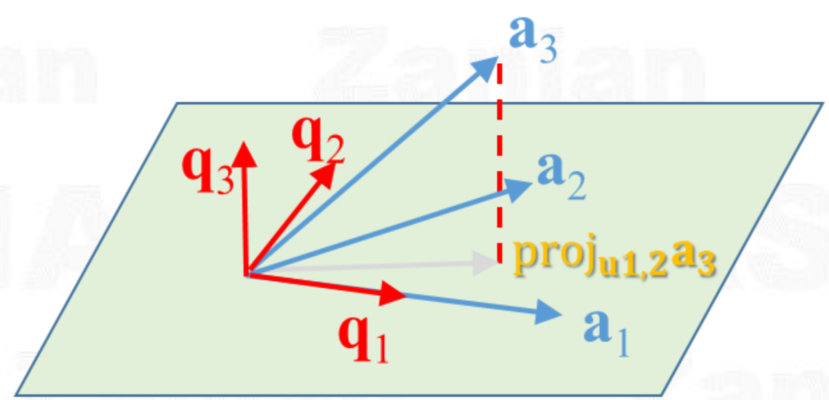
$$\mathbf{u}_{k+1} = \mathbf{a}_{k+1} - (\mathbf{a}_{k+1} \cdot \mathbf{u}_1)\mathbf{u}_1 - \cdots - (\mathbf{a}_{k+1} \cdot \mathbf{u}_k)\mathbf{u}_k,$$

$$\mathbf{q}_{k+1} = \frac{\mathbf{u}_{k+1}}{\|\mathbf{u}_{k+1}\|}$$

Orthogonality

Classical Gram-Schmidt

```
for j = 1 : n
   u_j = a_j
   for k = 1 : j - 1
      u_j = u_j - (q_j^T a_j) q_j
   end
   q_j = u_j / |u_j|
end
```



If an error is made in computing $\mathbf{q}_2$, so $\mathbf{q}_1^T\mathbf{q}_2=\delta$

$$\mathbf{u}_3 = \mathbf{a}_3 - (\mathbf{q}_1^T\mathbf{a}_3)\mathbf{q}_1 - (\mathbf{q}_2^T\mathbf{a}_3)\mathbf{q}_2$$

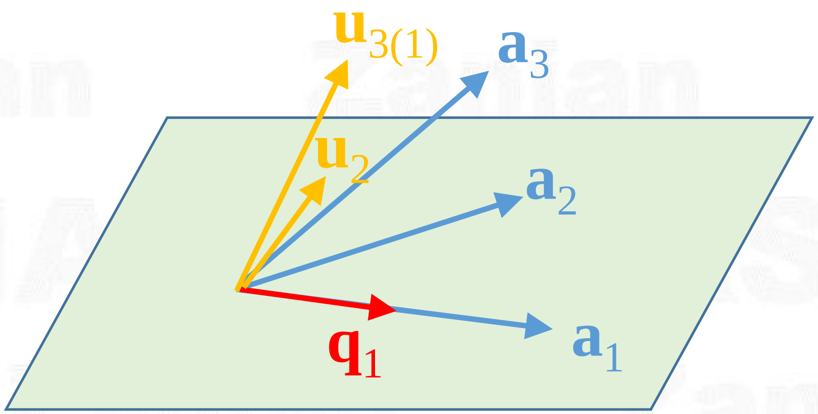
$$\mathbf{q}_2^T\mathbf{u}_3 = \mathbf{q}_2^T\mathbf{a}_3 - (\mathbf{q}_1^T\mathbf{a}_3)\delta - (\mathbf{q}_2^T\mathbf{a}_3) = -(\mathbf{q}_1^T\mathbf{a}_3)\delta$$

$$\mathbf{q}_1^T\mathbf{u}_3 = \mathbf{q}_1^T\mathbf{a}_3 - (\mathbf{q}_1^T\mathbf{a}_3) - (\mathbf{q}_2^T\mathbf{a}_3)\delta = -(\mathbf{q}_2^T\mathbf{a}_3)\delta$$

Orthogonality

Modified Gram-Schmidt

for $j = 1 : n$
  $\mathbf{u}_j = \mathbf{a}_j$
end
for $j = 1 : n$
  $\mathbf{q}_j = \mathbf{u}_j \;/\; |\mathbf{u}_j|$
  for $k = j+1 : n$
    $\mathbf{u}_k = \mathbf{u}_k - (\mathbf{q}_j^T \mathbf{u}_k)\, \mathbf{q}_j$
  end
end



Orthogonality

Modified Gram-Schmidt

for $j = 1 : n$

 $\mathbf{u}_j = \mathbf{a}_j$

end

for $j = 1 : n$

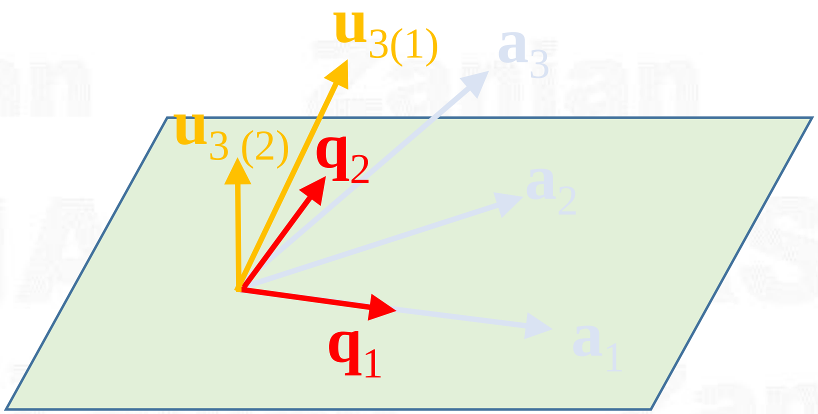
 $\mathbf{q}_j = \mathbf{u}_j \,/\, |\mathbf{u}_j|$

 for $k = j+1 : n$

  $\mathbf{u}_k = \mathbf{u}_k - (\mathbf{q}_j^T \mathbf{u}_k)\, \mathbf{q}_j$

 end

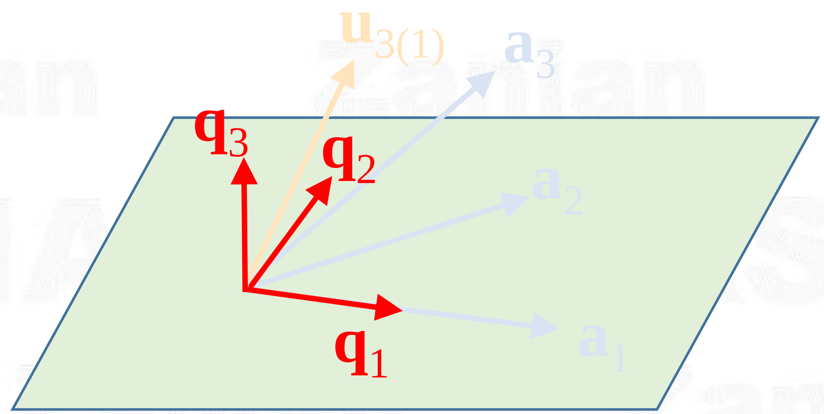
end



Orthogonality

Modified Gram-Schmidt

for $j = 1 : n$
  $\mathbf{u}_j = \mathbf{a}_j$
end
for $j = 1 : n$
  $\mathbf{q}_j = \mathbf{u}_j \,/\, |\mathbf{u}_j|$
  for $k = j+1 : n$
    $\mathbf{u}_k = \mathbf{u}_k - (\mathbf{q}_j^T \mathbf{u}_k)\, \mathbf{q}_j$
  end
end



1. $\mathbf{u}_3^{(1)} = \mathbf{a}_3 - (\mathbf{q}_1^T \mathbf{a}_3)\mathbf{q}_1$
2. $\mathbf{u}_3 = \mathbf{u}_3^{(1)} - \left(\mathbf{q}_2^T \mathbf{u}_3^{(1)}\right)\mathbf{q}_2$
3. $\mathbf{q}_2^T \mathbf{u}_3 = \mathbf{q}_2^T \mathbf{u}_3^{(1)} - \left(\mathbf{q}_2^T \mathbf{u}_3^{(1)}\right) = 0$
4. $\mathbf{q}_1^T \mathbf{u}_3 = \mathbf{q}_1^T \mathbf{u}_3^{(1)} - \left(\mathbf{q}_2^T \mathbf{u}_3^{(1)}\right)\delta$
5. $\mathbf{q}_1^T \mathbf{u}_3^{(1)} = \mathbf{q}_1^T \mathbf{a}_3 - (\mathbf{q}_1^T \mathbf{a}_3) = 0$
6. $\mathbf{q}_2^T \mathbf{u}_3^{(1)} = \mathbf{q}_2^T \mathbf{a}_3 - (\mathbf{q}_1^T \mathbf{a}_3)\delta$
7. $\mathbf{q}_1^T \mathbf{u}_3 = -(\mathbf{q}_2^T \mathbf{a}_3 - (\mathbf{q}_1^T \mathbf{a}_3)\delta)\delta$
8. $\mathbf{q}_1^T \mathbf{u}_3 = -\mathbf{q}_2^T \mathbf{a}_3\,\delta + (\mathbf{q}_1^T \mathbf{a}_3)\delta^2$

Orthogonality

QR decomposition

Similar to LU decomposition by finding an orthogonal coordinate system in which our transformation becomes upper triangular.

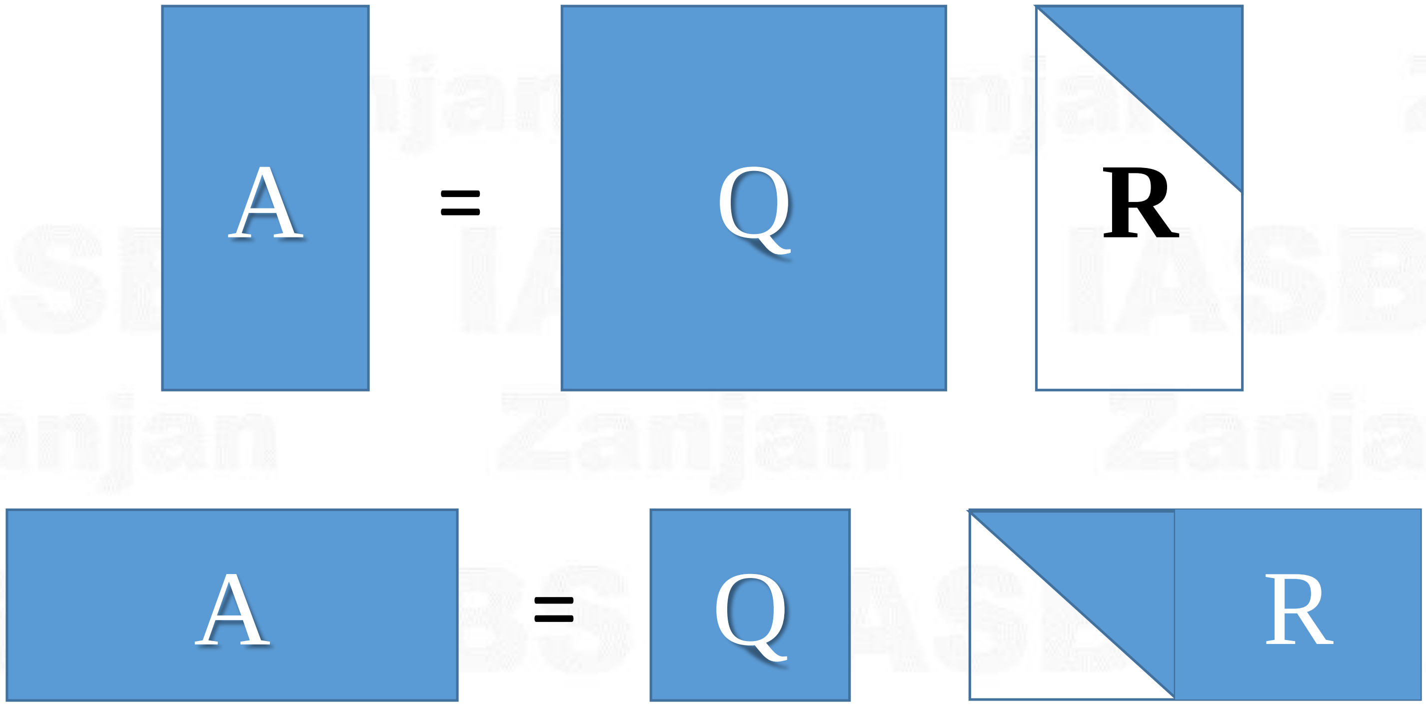
Given any n×n matrix **A**, if **A** is invertible, then there is an orthogonal matrix **Q** and an upper triangular matrix **R** with diagonal entries such that

$$\mathbf{A} = \mathbf{QR}$$

Orthogonality

QR decomposition

QR can also be applied to non-square matrices, resulting in factors that look like this:



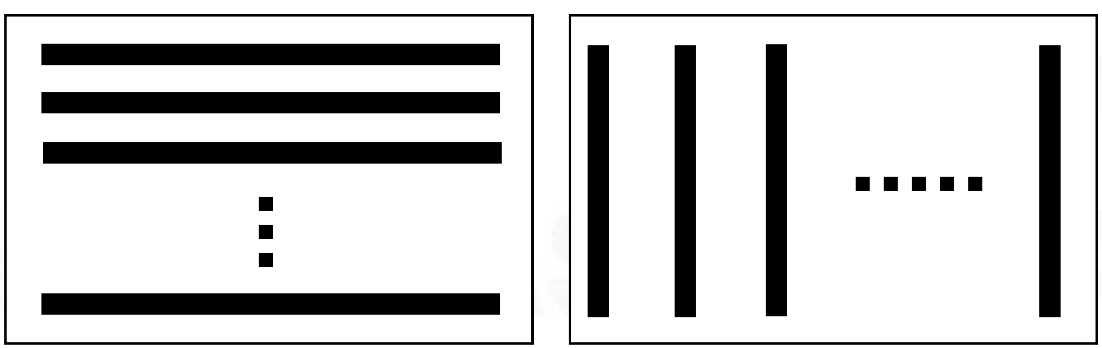
Orthogonality

QR decomposition

There are several methods for computing the QR decomposition. One of them is the Gram-Schmidt process.

$$[\mathbf{a}_1 \quad \mathbf{a}_2 \quad \dots \quad \mathbf{a}_3] = [\mathbf{q}_1 \quad \mathbf{q}_2 \quad \dots \quad \mathbf{q}_3] \begin{bmatrix} \mathbf{a}_1.\mathbf{q}_1 & \mathbf{a}_2.\mathbf{q}_1 & \cdots & \mathbf{a}_n.\mathbf{q}_1 \\ 0 & \mathbf{a}_2.\mathbf{q}_2 & \cdots & \mathbf{a}_n.\mathbf{q}_2 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{a}_n.\mathbf{q}_n \end{bmatrix}$$

$\mathbf{Q}^{\mathrm{T}}$ $\mathbf{A}$



$$= \begin{bmatrix} \mathbf{a}_1.\mathbf{q}_1 & \mathbf{a}_2.\mathbf{q}_1 & \cdots & \mathbf{a}_n.\mathbf{q}_1 \\ 0 & \mathbf{a}_2.\mathbf{q}_2 & \cdots & \mathbf{a}_n.\mathbf{q}_2 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{a}_n.\mathbf{q}_n \end{bmatrix}$$

Householder Transformation and relation to **QR** decomposition

Orthogonality

Any Question?