Advanced Numerical Methods December 8, 2011

Lecture Notes 2.3: QR Decomposition

Lecturer: Che-Rung Lee *Scribe:Chang Yafang*

1 QR Decomposition

- Gram-Schmit process
- Householder reflector
- Givens rotation

1.1 Linear Algebra

- The meaning of a vector

$$\vec{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} = a_1 \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + a_2 \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \end{bmatrix} + \cdots + a_n \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$$

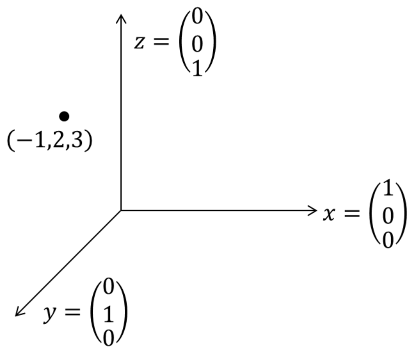


Figure 1: Example of a three dimensional vector $(-1, 2, 3)$.

- Example : If $\vec{x} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} = x_1\vec{a} + x_2\vec{b}$, shown as Figure 2(a) , find the values of x_1 and x_2.

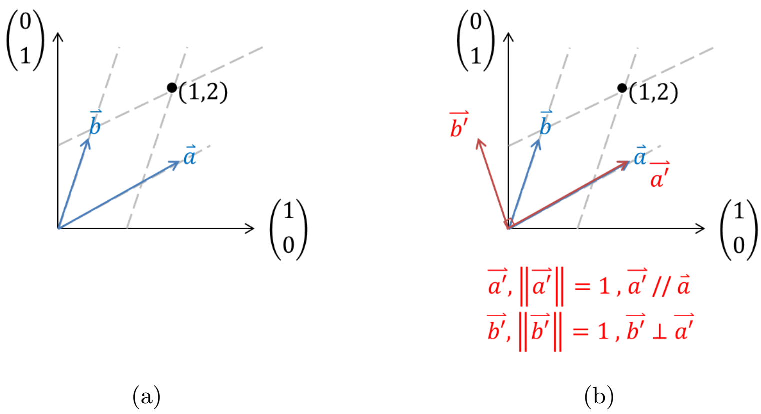


Figure 2: Plots of the example.

$$\vec{a} = \begin{bmatrix} 1 \\ 0.5 \end{bmatrix}, \vec{b} = \begin{bmatrix} 0.2 \\ 1 \end{bmatrix} \Rightarrow x_1\vec{a} + x_2\vec{b} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$
$$\Rightarrow \begin{bmatrix} \vec{a} & \vec{b} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$
$$\Rightarrow \begin{bmatrix} 1 & 0.2 \\ 0.5 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

  - We can use matrix $A = \begin{bmatrix} 1 & 0.2 \\ 0.5 & 1 \end{bmatrix} = \begin{bmatrix} \vec{a} & \vec{b} \end{bmatrix}$ to find x_1 and x_2. But it is not easy to solve, because A is not an orthogonal matrix.
  - Now we want to find a matrix $A' = \begin{bmatrix} \vec{a}' & \vec{b}' \end{bmatrix}$ (shown in Figure 4(b)) to make the computation become easier.
  -

$$A = \begin{bmatrix} \vec{a} & \vec{b} \end{bmatrix} = \underbrace{\begin{bmatrix} \vec{a}' & \vec{b}' \end{bmatrix}}_{Q} \underbrace{\begin{bmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{bmatrix}}_{R}$$

$\Rightarrow$

$$\vec{a} = \vec{a}'r_{11} + \vec{b}'r_{21} \quad (1)$$
$$\vec{b} = \vec{a}'r_{12} + \vec{b}'r_{22} \quad (2)$$

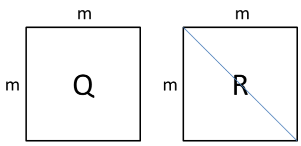
In (1) : $\vec{a} // \vec{a}'$ and $\vec{b}' \neq 0 \Rightarrow r_{21} = 0$, such that R will be upper triangular.

- Vector
  1. $\vec{a}$,$\vec{b}$ are orthogonal if $\vec{a}^T\vec{b} = 0$.
  2. $\vec{a}$,$\vec{b}$ are orthonormal if $\vec{a}$,$\vec{b}$ are orthogonal and $\|\vec{a}\| = \|\vec{b}\| = 1$.
- Matrix Q is orthogonal, and $Q = [\vec{q}_1, \vec{q}_2, ..., \vec{q}_n]$. (If it is in complex number domain, then Q is called unitary.)

$$\vec{q}_i^T \vec{q}_j = \begin{cases} 0, & \text{if } i \neq j \\ 1, & \text{if } i = j \end{cases}$$

  $\Rightarrow$

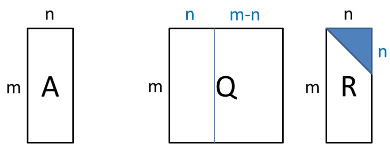
  1. $Q^{-1} = Q^T$. $\quad < proof > Q^T Q = I$
  2. $\forall \vec{v}$, $\|Q\vec{v}\| = \|\vec{v}\|$. $\quad < proof > \|Q\vec{v}\|^2 = (Q\vec{v})^T(Q\vec{v}) = \vec{v}^T Q^T Q \vec{v} = \vec{v}^T\vec{v} = \|\vec{v}\|^2$
- $A \in \mathbb{R}^{m \times n}$, $A = QR$.
  1. $m = n$,

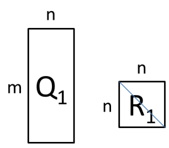


  2. $m > n$, there are two different sizes of decomposition.

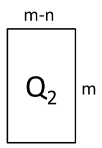
$$(2.1)\ A = QR \ \ and \ \ (2.2)\ A = Q_1R_1$$

  (2.1) equals (2.2), since $Q = [Q_1 \quad Q_2]$, $R = [R_1^T \quad 0]^T$
  $\Rightarrow QR = Q_1R_1 + Q_2 0 = Q_1R_1$.

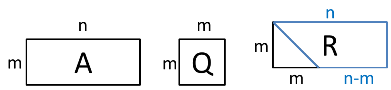




, where



  3. $m < n$



1.2 Gram-Schmit process

- $$A = [\vec{a}_1, \vec{a}_2, ..., \vec{a}_n], Q = [\vec{q}_1, \vec{q}_2, ..., \vec{q}_n], R = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ & r_{22} & \cdots & r_{2n} \\ & & \ddots & \ddots \\ & & & r_{nn} \end{bmatrix}$$

$$\begin{aligned} A &= QR \\ \Leftrightarrow \vec{a}_1 &= r_{11}\vec{q}_1 \\ \vec{a}_2 &= r_{12}\vec{q}_1 + r_{22}\vec{q}_2 \\ \vdots &= \vdots \\ \vec{a}_n &= r_{1n}\vec{q}_1 + r_{2n}\vec{q}_2 + \cdots + r_{nn}\vec{q}_n \end{aligned}$$

$$\begin{aligned} \Rightarrow \vec{q}_1 &= \frac{\vec{a}_1}{r_{11}} \, and \, \|\vec{q}_1\| = 1 \,\therefore r_{11} = \|\vec{a}_1\| \\ \|\vec{q}_1\|^2 &= \vec{q}_1^T\vec{q}_1 = \frac{\vec{a}_1^T\vec{a}_1}{\|\vec{a}_1\|^2} = 1 \\ \vec{q}_2 &= \frac{\vec{a}_2 - r_{12}\vec{q}_1}{r_{22}} \quad and \quad \vec{q}_1 \perp \vec{q}_2 \Leftrightarrow \vec{q}_1^T\vec{q}_2 = 0 \\ \vec{q}_1^T\vec{q}_2 &= \frac{\vec{q}_1^T\vec{a}_2 - r_{12}\vec{q}_1^T\vec{q}_1}{r_{22}} = 0 \Rightarrow r_{12} = \vec{q}_1^T\vec{a}_2 \\ \vec{q}_2 \; &// \; (\vec{a}_2 - r_{12}\vec{q}_1)\,,\ \|\vec{q}_2\| = 1\,,\ r_{22} = \|\vec{a}_2 - r_{12}\vec{q}_1\| \\ &\vdots \\ \vec{q}_3 &= \frac{\vec{a}_3 - r_{13}\vec{q}_1 - r_{23}\vec{q}_2}{r_{33}} \\ &\vdots \end{aligned}$$

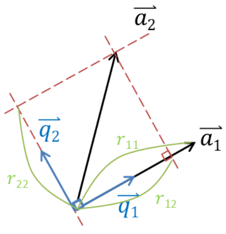
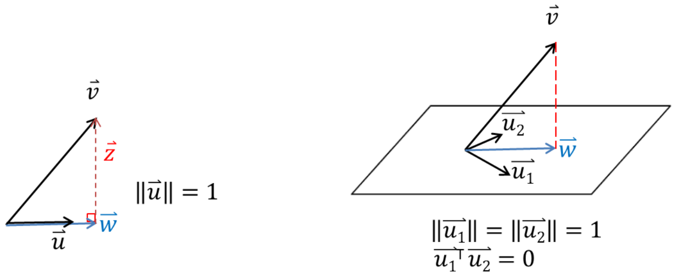


Figure 3: Part of schematic diagram of above derivation.

- Projection Matrix



(a) (b)

Figure 4: Projection diagrams.

  - In Figure 4(a), $\vec{w} = (\vec{v}^T\vec{u})\vec{u} = [\|\vec{u}\|\ \|\vec{v}\|\cos\theta]\,\vec{u} = \vec{u}(\vec{u}^T\vec{v}) = (\vec{u}\vec{u}^T)\vec{v} = P\vec{v}$. P is projection matrix.
  - In Figure 4(b), if $\vec{w} = \alpha\vec{u}_1 + \beta\vec{u}_2$, find the values of α and β .

$$
\begin{aligned}
(\vec{v}-\vec{w})\perp\vec{u}_1 \quad & \textit{and} \quad (\vec{v}-\vec{w})\perp\vec{u}_2 \\
\vec{u}_1^T(\vec{v}-\vec{w}) \quad &= \quad \vec{u}_1^T\vec{v} - \vec{u}_1^T\vec{w} = \vec{u}_1^T\vec{v} - \vec{u}_1^T(\alpha\vec{u}_1) - \vec{u}_1^T(\beta\vec{u}_2) = 0 \\
\Rightarrow \alpha \quad &= \quad \vec{u}_1^T\vec{v} \\
\vec{u}_2^T(\vec{v}-\vec{w}) \quad &= \quad \vec{u}_2^T\vec{v} - \vec{u}_2^T(\alpha\vec{u}_1) - \vec{u}_2^T(\beta\vec{u}_2) = 0 \\
\Rightarrow \beta \quad &= \quad \vec{u}_2^T\vec{v} \\
\vec{w} \quad &= \quad \vec{u}_1\vec{u}_1^T\vec{v} + \vec{u}_2\vec{u}_2^T\vec{v} = \underbrace{\left(\vec{u}_1\vec{u}_1^T + \vec{u}_2\vec{u}_2^T\right)}_{P_{\vec{u}_1\vec{u}_2}:\ \textit{projection matrix}}\ \vec{v}
\end{aligned}
$$

    Let $U = (\vec{u}_1 \quad \vec{u}_2) \Rightarrow P_{\vec{u}_1\vec{u}_2} = UU^T = P_U$.
    Then if $U = [\vec{u}_1\ \vec{u}_2\ \vec{u}_3 \cdots \vec{u}_n]$, $P_U = UU^T$.

  - In Figure 4(a), $\vec{v} = \vec{w} + \vec{z}$, $\vec{w} \,//\, \vec{u}$, $\vec{z}\perp\vec{u}$

$$
\vec{z} = \vec{v} - \vec{w} = \vec{v} - UU^T\vec{v} = \underbrace{(I - UU^T)}_{P_\perp}\vec{v}
$$

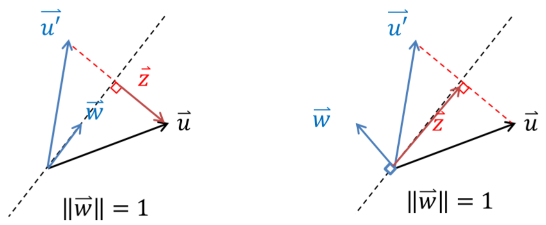
Algorithm 1 Gram-Schmit Process Algorithm

1. $Q_1 = \dfrac{\vec{a}_1}{\|\vec{a}_1\|}$, $r_1 = \|\vec{a}_1\|$
2. For $i = 0, 1, 2, \ldots, n$
3. $\quad \vec{v} = (I - Q_{i-1}Q_{i-1}^T)\vec{a}_i = \vec{a}_i - Q_{i-1}Q_{i-1}^T\vec{a}_i$
4. $\quad Q_i = \begin{bmatrix} Q_{i-1} & \dfrac{\vec{v}}{\|\vec{v}\|} \end{bmatrix}$, $r_{ii} = \|\vec{v}\|$

- Complexity

$$
\begin{aligned}
T(n) &= (\sum_{i=2}^{n} 4mi) + 3m \\
&= 4m\frac{(n+2)(n-1)}{2} + 3m \\
&= 2mn^2 + 2mn - m
\end{aligned}
$$

1.3 Householder reflector



(a) (b)

Figure 5: Reflector diagrams.

- 1. Given $\vec{u}$ and $\vec{w}$, $\|\vec{w}\| = 1$, shown as Figure 5(a), compute $\vec{u'}$.

$$
\begin{aligned}
\vec{z} &= (I - \vec{w}\vec{w}^T)\vec{u} \\
\vec{u} &= \vec{z} + \vec{w}\vec{w}^T\vec{u}
\end{aligned}
$$

$$
\begin{aligned}
\Rightarrow \vec{u'} &= \vec{w}\vec{w}^T\vec{u} - \vec{z} \\
&= \vec{w}\vec{w}^T\vec{u} - \vec{u} + \vec{w}\vec{w}^T\vec{u} \\
&= -\underbrace{(I - 2\vec{w}\vec{w}^T)}_{reflector}\vec{u}
\end{aligned}
$$

NOTE : We can choose $\vec{w}$ as another direction, shown as Figure 5(b). Then we can find

$$\vec{u'} = (I - \vec{w}\vec{w}^T)\vec{u} - \vec{w}\vec{w}^T\vec{u} = \underbrace{(I - 2\vec{w}\vec{w}^T)}_{reflector}\vec{u}$$

2. Given $\vec{u}$ and $\vec{u'}$, shown as Figure 5(a), compute $\vec{w}$.

$$\vec{v} = \frac{\vec{u} + \vec{u'}}{2}, \vec{w} = \frac{\vec{v}}{\|\vec{v}\|}$$

NOTE : We can choose $\vec{w}$ as another direction, shown as Figure 5(b). Then we can find

$$\vec{v} = \frac{\vec{u'} - \vec{u}}{2}, \vec{w} = \frac{\vec{v}}{\|\vec{v}\|}$$

- Householder reflector features :
  $H = I - 2\vec{w}\vec{w}^T$, $\|\vec{w}\| = 1$

  1. H is symmetric.
  $$H^T = I - 2(\vec{w}^T)^T\vec{w}^T = I - 2\vec{w}\vec{w}^T = H$$
  2. H is orthogonal.
  $$\begin{aligned} H^T H &= (I - 2\vec{w}\vec{w}^T)(I - 2\vec{w}\vec{w}^T) \\ &= I - 2\vec{w}\vec{w}^T - 2\vec{w}\vec{w}^T + 4\vec{w}\vec{w}^T\vec{w}\vec{w}^T \\ &= I - 4\vec{w}\vec{w} + 4\vec{w}\vec{w} \\ &= I \\ &\Rightarrow H^{-1} = H^T = H \end{aligned}$$

- Example :

$$\begin{aligned} \vec{v}_1 &= \vec{a}_1 - \|\vec{a}_1\|\vec{e}_1 \ , \ \vec{w}_1 = \frac{\vec{v}_1}{\|\vec{v}_1\|} \ , \ A = [\vec{a}_1, \vec{a}_2, \cdots, \vec{a}_n] \\ H_1 &= I - 2\vec{w}_1\vec{w}_1^T \\ H_1 A &= \begin{bmatrix} \|\vec{a}_1\| & \triangle & \cdots & \triangle \\ 0 & & & \\ \vdots & & \mathbf{A^{(1)}} & \\ 0 & & & \end{bmatrix} , \ A_{(n\times n)} \ , \ A^{(1)}_{(n-1\times n-1)} \\ H_2 A^{(1)} &= \begin{bmatrix} \|\vec{a}_1^{(1)}\| & \triangle & \cdots & \triangle \\ 0 & & & \\ \vdots & & \mathbf{A^{(2)}} & \\ 0 & & & \end{bmatrix} , \ A^{(2)}_{(n-2\times n-2)} \\ H_2 &= I - 2\vec{w}_2\vec{w}_2^T \end{aligned}$$

Question : $A = QR$, how can we find Q? $\Rightarrow$ Using $\hat{H}_2 = \begin{bmatrix} 0 & \vec{0}^T \\ \vec{0} & H_2 \end{bmatrix}$.

Another question : is $\hat{H}_2$ orthogonal?

$$\hat{H}_2 = \begin{bmatrix} 0 & \vec{0}^T \\ \vec{0} & H_2 \end{bmatrix} \Rightarrow \hat{H}_2{}^T\hat{H}_2 = \begin{bmatrix} 0 & \vec{0}^T \\ \vec{0} & H_2^T \end{bmatrix}\begin{bmatrix} 0 & \vec{0}^T \\ \vec{0} & H_2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & I \end{bmatrix}$$

$\Rightarrow$ $\hat{H}_2$ is not orthogonal. So we can do some modification of $\hat{H}_2$.

$$\hat{H}_2 = \begin{bmatrix} 1 & \vec{0}^T \\ \vec{0} & H_2 \end{bmatrix} \Rightarrow \hat{H}_2{}^T\hat{H}_2 = \begin{bmatrix} 1 & \vec{0}^T \\ \vec{0} & H_2^T \end{bmatrix}\begin{bmatrix} 1 & \vec{0}^T \\ \vec{0} & H_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & I \end{bmatrix} = I$$

$\Rightarrow$: $\hat{H}_2$ is orthogonal right now.

NOTE : If A and B are orthogonal matrices, the AB is also orthogonal.
$< proof >$ $(AB)^T AB = B^T A^T AB = B^T B = I$

Algorithm 2 QR decomposition using Householder reflector

1. Let $R^{(0)} = A$
2. For $i = 1, 2, \ldots, n-1$
3. $\quad \hat{H}_i = \begin{bmatrix} I & 0 \\ 0 & H_i \end{bmatrix}$ where $H_i R^{(i-1)}(i, i:n) = \|R^{(i-1)}(i, i:n)\| e_1$
4. $\quad R^{(i)} = \hat{H}_i R^{(i-1)}$

   End for
5. $\quad R = R^{(n-1)}, Q = \hat{H}_1^T \hat{H}_2^T \cdots \hat{H}_{(n-1)}^T$

- Complexity

  Let size of H_i being $k \times k$, where $k = n - i + 1$.

  - Step 3, $H_i = I - 2\vec{w}_i\vec{w}_i^T$, $\vec{w}_i = \dfrac{\vec{a}_1 - \|\vec{a}_1\|\vec{e}_1}{\|\vec{a}_1 - \|\vec{a}_1\|\vec{e}_1\|}$
    1. $\|\vec{a}_1\|$: $2k$.
    2. $\vec{a}_1 - \|\vec{a}_1\|\vec{e}_1$: $O(1)$.
    3. Normalize $\vec{w}_i$: $O(1)$.
  - Step 4,

    $\hat{H}_i R^{(i-1)} = H_i \underbrace{R^{(i-1)}(i:n, i:n)}_{T_i} = (I - 2\vec{w}_i\vec{w}_i^T)T_i = T_i - 2\vec{w}_i(\vec{w}_i^T T_i)$: $4k^2$.

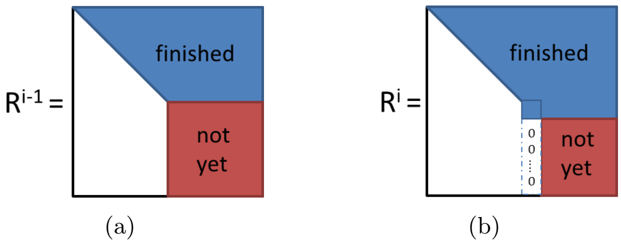


Figure 6: Schematic diagrams of Algorithm 2.

  -

$$
\begin{aligned}
T(n) &= T(n-1) + 2n + 4n^2 \\
&= \sum_{k=1}^{n} [4k^2 + 2k + c] \\
&= 4 \cdot \frac{1}{6} n(n+1)(2n+1) + 2 \cdot \frac{1}{2} n(n+1) + c \cdot n \\
&\approx \frac{4}{3} n^3
\end{aligned}
$$

  - Step 5,

    $I - 2\vec{w}\vec{w}^T$: $2n^2$.

$$\hat{H}_i = \begin{bmatrix} I & 0 \\ 0 & H_i \end{bmatrix}, \ B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}, \ \hat{H}_i B = \begin{bmatrix} B_{11} & B_{12} \\ H_i B_{21} & H_i B_{22} \end{bmatrix} : \ 4kn.$$

$$\begin{aligned} T(n) &= T(n-1) + 2n^2 + \sum_{k=2}^{n-1} 4kn \\ &= 2n^2 + 4n\frac{(n+1)(n-1)}{2} \\ &\approx 2n^3 \end{aligned}$$