Data Science for Actuaries (ACT6100)

Arthur Charpentier

Rappels # 3.2 (QR & SVD - Singular Value Decomposition)

automne 2020

https://github.com/freakonometrics/ACT6100/

QR Factorization

Let $\boldsymbol{A}$ be some $n \times k$ matrix, with k linarly independent vectors $\boldsymbol{a}_1, \cdots, \boldsymbol{a}_k$ in $\mathbb{R}^n$, $\boldsymbol{A} = [\boldsymbol{a}_1, \cdots, \boldsymbol{a}_k]$.
From Gram-Schmidt decomposition, there is $\boldsymbol{Q} = [\boldsymbol{q}_1, \cdots, \boldsymbol{q}_k]$ such that $\boldsymbol{Q}^\top \boldsymbol{Q} = \mathbb{I}$ and $\boldsymbol{R}$ upper triangular, such that $\boldsymbol{A} = \boldsymbol{Q}\boldsymbol{R}$.

Note: QR factorization is interesting to inverse matrices,

$$\boldsymbol{A}^{-1} = (\boldsymbol{Q}\boldsymbol{R})^{-1} = \boldsymbol{R}^{-1}\boldsymbol{Q}^\top$$

$\boldsymbol{R}^{-1}$ appears when solving linear systems

Spectral vs. Singular Value Decomposition

Example We've seen before the spectral decomposition

$$\boldsymbol{M}=\begin{bmatrix}1&2\\3&4\end{bmatrix}=\underbrace{\begin{bmatrix}-0.416&-0.825\\-0.909&0.566\end{bmatrix}}_{\boldsymbol{P}}\underbrace{\begin{bmatrix}5.372&0\\0&-0.372\end{bmatrix}}_{\boldsymbol{D}}\underbrace{\begin{bmatrix}-0.416&-0.825\\-0.909&0.566\end{bmatrix}^{-1}}_{\boldsymbol{P}^{-1}}$$

where

$$\underbrace{\begin{bmatrix}-0.416&-0.825\\-0.909&0.566\end{bmatrix}^{-1}}_{\boldsymbol{P}^{-1}}=\begin{bmatrix}-0.574&-0.837\\-0.923&0.422\end{bmatrix}$$

Alternatively, there are $\boldsymbol{U}$ and $\boldsymbol{V}$ two rotation matrices, and $\boldsymbol{\Delta}$ some diagonal matrix such that

$$\boldsymbol{M}=\boldsymbol{U}\boldsymbol{\Delta}\boldsymbol{V}^\top$$

Example

$$\boldsymbol{M}=\begin{bmatrix}1&3\\2&4\end{bmatrix}=\underbrace{\begin{bmatrix}-0.404&-0.914\\-0.914&0.404\end{bmatrix}}_{\boldsymbol{U}}\underbrace{\begin{bmatrix}5.465&0\\0&0.366\end{bmatrix}}_{\boldsymbol{\Delta}}\underbrace{\begin{bmatrix}-0.576&0.817\\-0.817&-0.576\end{bmatrix}}_{\boldsymbol{V}^\top}$$

Spectral vs. Singular Value Decomposition

```
> M=c(1,3,2,4)
> dim(M)=c(2,2)
> M
     [,1] [,2]
[1,]    1    2
[2,]    3    4
> eigen(M)
$values
[1]  5.3722813 -0.3722813

$vectors
           [,1]       [,2]
[1,] -0.4159736 -0.8245648
[2,] -0.9093767  0.5657675
```

```
> svd(M)
$d
[1] 5.4649857 0.3659662

$u
           [,1]       [,2]
[1,] -0.4045536 -0.9145143
[2,] -0.9145143  0.4045536

$v
           [,1]       [,2]
[1,] -0.5760484  0.8174156
[2,] -0.8174156 -0.5760484
>
```

$$\boldsymbol{D}=\begin{bmatrix}5.372 & 0\\ 0 & -0.372\end{bmatrix} \text{ and } \boldsymbol{\Delta}=\begin{bmatrix}5.465 & 0\\ 0 & 0.366\end{bmatrix}$$

$$\boldsymbol{M}^\top\boldsymbol{M}=\begin{bmatrix}10 & 14\\ 14 & 20\end{bmatrix}=\underbrace{\begin{bmatrix}0.576 & -0.817\\ 0.817 & 0.576\end{bmatrix}}_{\boldsymbol{P}'}\underbrace{\begin{bmatrix}29.866 & 0\\ 0 & 0.134\end{bmatrix}}_{\boldsymbol{D}'=\boldsymbol{\Delta}^2}\underbrace{\begin{bmatrix}0.576 & -0.817\\ 0.817 & 0.576\end{bmatrix}^{-1}}_{\boldsymbol{P}'^{-1}}$$

Spectral vs. Singular Value Decomposition

```
> svd(M)
$d
[1] 5.4649857 0.3659662

$u
           [,1]       [,2]
[1,] -0.4045536 -0.9145143
[2,] -0.9145143  0.4045536

$v
           [,1]       [,2]
[1,] -0.5760484  0.8174156
[2,] -0.8174156 -0.5760484
> U = svd(M)$u
> V = svd(M)$v
```

```
> V%*%t(V)
     [,1] [,2]
[1,]    1    0
[2,]    0    1
> t(V)%*%V
     [,1] [,2]
[1,]    1    0
[2,]    0    1
> U%*%t(U)
          [,1]      [,2]
[1,] 1.000e+00 1.110e-16
[2,] 1.110e-16 1.000e+00
> round(U%*%t(U),10)
     [,1] [,2]
[1,]    1    0
[2,]    0    1
```

$\boldsymbol{U}$ and $\boldsymbol{V}$ are usually called rotation matrices but they are more precisely orthonormal matrices,

$$\boldsymbol{U}^\top\boldsymbol{U} = \boldsymbol{U}\boldsymbol{U}^\top = \mathbb{I} \text{ and } \boldsymbol{V}^\top\boldsymbol{V} = \boldsymbol{V}\boldsymbol{V}^\top = \mathbb{I}$$

Spectral vs. Singular Value Decomposition

$$\boldsymbol{M} = \boldsymbol{P}\boldsymbol{D}\boldsymbol{P}^{-1} \text{ or } \boldsymbol{U}\boldsymbol{\Delta}\boldsymbol{V}^\top \text{ ?}$$

with

- $\boldsymbol{P}$ is a non-singular matrix (i.e. $\boldsymbol{P}^{-1}$ exists)
- $\boldsymbol{U}$ and $\boldsymbol{V}$ are orthonormal, $\boldsymbol{U}\boldsymbol{U}^\top = \mathbb{I}$ and $\boldsymbol{V}\boldsymbol{V}^\top = \mathbb{I}$

λ is a singular value for $\boldsymbol{M}$ if

$$\exists \vec{\boldsymbol{u}}, \vec{\boldsymbol{v}} \text{ such that } \boldsymbol{M}\vec{\boldsymbol{v}} = \lambda\vec{\boldsymbol{u}} \text{ and } \boldsymbol{M}^\top\vec{\boldsymbol{u}} = \lambda\vec{\boldsymbol{v}}$$

$\vec{\boldsymbol{u}}$ and $\vec{\boldsymbol{v}}$ are left-singular and right-singular vectors.

Note that if $\boldsymbol{M} = \boldsymbol{U}\boldsymbol{\Delta}\boldsymbol{V}^\top$,

$$\boldsymbol{M}\boldsymbol{M}^\top = \boldsymbol{U}\boldsymbol{\Delta}\underbrace{\boldsymbol{V}^\top\boldsymbol{V}}\boldsymbol{\Delta}^\top\boldsymbol{U} = \boldsymbol{U}\underbrace{\boldsymbol{\Delta}\boldsymbol{\Delta}^\top}\boldsymbol{U}^\top$$

$$\boldsymbol{M}^\top\boldsymbol{M} = \boldsymbol{V}\boldsymbol{\Delta}^\top\underbrace{\boldsymbol{U}^\top\boldsymbol{U}}\boldsymbol{\Delta}\boldsymbol{V}^\top = \boldsymbol{V}\underbrace{\boldsymbol{\Delta}^\top\boldsymbol{\Delta}}\boldsymbol{V}^\top$$

Spectral vs. Singular Value Decomposition

$$M = PDP^{-1} \text{ or } U\Delta V^{\top} \text{ ?}$$

- eigenvalues $D_{i,i}$ can be negative
- entries $\Delta_{i,i}$ are all positive

```
> M=c(1,3,1,4)
> dim(M)=c(2,2)
> M
     [,1] [,2]
[1,]    1    1
[2,]    3    4
> eigen(M)
$values
[1] 4.7912878 0.2087122

$vectors
           [,1]       [,2]
[1,] -0.2550401 -0.7841904
[2,] -0.9669305  0.6205203
```

```
> svd(M)
$d
[1] 5.1925824 0.1925824

$u
           [,1]       [,2]
[1,] -0.2700013 -0.9628599
[2,] -0.9628599  0.2700013

$v
           [,1]       [,2]
[1,] -0.6082872 -0.7937170
[2,] -0.7937170  0.6082872
>
```

Spectral vs. Singular Value Decomposition

$$\boldsymbol{M} = \boldsymbol{P}\boldsymbol{D}\boldsymbol{P}^{-1} \text{ or } \boldsymbol{U}\boldsymbol{\Delta}\boldsymbol{V}^{\top} \text{ ?}$$

If $\boldsymbol{M}$ is a $n \times n$ symmetric matrix, the two coincide (up to signs)

```
> M=c(1,3,3,1)
> dim(M)=c(2,2)
> M
     [,1] [,2]
[1,]    1    3
[2,]    3    1
> eigen(M)
$values
[1]  4 -2

$vectors
          [,1]       [,2]
[1,] 0.7071068 -0.7071068
[2,] 0.7071068  0.7071068
```

```
> svd(M)
$d
[1] 4 2

$u
           [,1]       [,2]
[1,] -0.7071068 -0.7071068
[2,] -0.7071068  0.7071068

$v
           [,1]       [,2]
[1,] -0.7071068  0.7071068
[2,] -0.7071068 -0.7071068
>
```

Singular Value Decomposition $2\times 2 \ \to \times 2\times 3$

More generally

$$\begin{bmatrix}1 & 2 & 3\\ 4 & 5 & 6\end{bmatrix} = \underbrace{\begin{bmatrix}-0.386 & -0.922\\ -0.922 & 0.386\end{bmatrix}}_{\boldsymbol{U}} \underbrace{\begin{bmatrix}9.508 & 0\\ 0 & 0.773\end{bmatrix}}_{\boldsymbol{\Delta}} \underbrace{\begin{bmatrix}-0.429 & -0.566 & -0.704\\ 0.806 & 0.112 & -0.581\end{bmatrix}}_{\boldsymbol{V}^\top}$$

$$\begin{bmatrix}1 & 2 & 3\\ 4 & 5 & 6\end{bmatrix} = \underbrace{\begin{bmatrix}-0.39 & -0.92\\ -0.92 & 0.39\end{bmatrix}}_{\boldsymbol{U}} \underbrace{\begin{bmatrix}9.51 & 0 & 0\\ 0 & 0.77 & 0\end{bmatrix}}_{\boldsymbol{\Delta}} \underbrace{\begin{bmatrix}-0.42 & -0.57 & -0.70\\ 0.81 & 0.11 & -0.58\\ \star & \star & \star\end{bmatrix}}_{\boldsymbol{V}^\top}$$

$$\begin{bmatrix}1 & 2 & 3\\ 4 & 5 & 6\end{bmatrix} = \underbrace{\begin{bmatrix}-0.39 & -0.92\\ -0.92 & 0.39\end{bmatrix}}_{\boldsymbol{U}} \underbrace{\begin{bmatrix}9.51 & 0 & 0\\ 0 & 0.77 & 0\end{bmatrix}}_{\boldsymbol{\Delta}} \underbrace{\begin{bmatrix}-0.42 & -0.57 & -0.70\\ 0.81 & 0.11 & -0.58\\ 0.41 & -0.82 & 0.41\end{bmatrix}}_{\boldsymbol{V}^\top}$$

where

- $\boldsymbol{U}$ and $\boldsymbol{V}$ are (square) rotation matrices (orthogonal)
- $\boldsymbol{\Delta}$ is a "diagonal" matrix

Singular Value Decomposition

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} = \underbrace{\begin{bmatrix} -0.39 & -0.92 \\ -0.92 & 0.39 \end{bmatrix}}_{\boldsymbol{U}} \underbrace{\begin{bmatrix} 9.51 & 0 & 0 \\ 0 & 0.77 & 0 \end{bmatrix}}_{\boldsymbol{\Delta}} \underbrace{\begin{bmatrix} -0.42 & -0.57 & -0.70 \\ 0.81 & 0.11 & -0.58 \\ 0.41 & -0.82 & 0.41 \end{bmatrix}}_{\boldsymbol{V}^\top}$$

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} = \underbrace{\begin{bmatrix} -3.67 & -0.71 & 0 \\ -8.8 & 0.30 & 0 \end{bmatrix}}_{\boldsymbol{U\Delta}} \underbrace{\begin{bmatrix} -0.42 & -0.57 & -0.70 \\ 0.81 & 0.11 & -0.58 \\ 0.41 & -0.82 & 0.41 \end{bmatrix}}_{\boldsymbol{V}^\top}$$

Spectral vs. Singular Value Decomposition

Let $\boldsymbol{M}$ be a $m \times n$ matrix with $\text{rank}(\boldsymbol{M}) = r$, then

$$\boldsymbol{M} = \boldsymbol{U}\boldsymbol{\Delta}\boldsymbol{V}^\top$$

where $\boldsymbol{U}$ and $\boldsymbol{V}$ are $m \times m$ and $n \times n$ orthogonal matrices

$$\boldsymbol{U}^\top\boldsymbol{U} = \mathbb{I}_m \text{ and } \boldsymbol{V}^\top\boldsymbol{V} = \mathbb{I}_n$$

and $\boldsymbol{\Delta}$ is a $m \times n$ matrix with 0's everywhere, except on the m main diagonal, with entries $\boldsymbol{\Delta}_{ii}$ such that

$$\boldsymbol{\Delta}_{11} \geq \boldsymbol{\Delta}_{22} \geq \boldsymbol{\Delta}_{rr} > 0 \text{ and } \boldsymbol{\Delta}_{ii} = 0,\ \forall i > r$$

. Hence

$$\boldsymbol{M} = \sum_{i=1}^{r} \boldsymbol{\Delta}_{ii}\boldsymbol{U}_{\cdot i}\boldsymbol{V}_{\cdot i}^\top$$

Note: $\boldsymbol{M}'_k = \sum_{i=1}^{k} \boldsymbol{\Delta}_{ii}\boldsymbol{U}_{\cdot i}\boldsymbol{V}_{\cdot i}^\top$ is an approximation of $\boldsymbol{M}$ of rank k. It is the best rank-k approximation (Schmidt–Mirsky–Eckart–Young).