Week 8b

Thursday, October 29, 2020 10:24 AM

document

Chapter 22

Week 8b: Singular Value Decomposition (SVD)

Schedule

We previously met the Eigenvalue Decomposition (EVD), which we used on square matrices. There is no EVD for rectangular matrices, but there does exist a generalization known as the Singular Value Decomposition, which is one of the most useful matrix decompositions in applied linear algebra. In fact, we met the basic ingredients of the SVD in the previous class when we explored the user-movie rating data matrix. See the following webpage at the American Mathematical Society for a good geometric discussion of the SVD.

22.1 SVD - The Big Idea [40 minutes]

Rectangular matrices don't have eigenvalues and eigenvectors. However, they have a generalisation of these known as singular values and singular vectors.

- The singular values σ_i and singular vectors $\mathbf{u}_i, \mathbf{v}_i$ of an $n \times m$ rectangular matrix $\mathbf{A}$ satisfy the definition

$$\mathbf{A}\mathbf{v}_i = \sigma_i \mathbf{u}_i \tag{22.1}$$
$$\mathbf{A}^T\mathbf{u}_i = \sigma_i \mathbf{v}_i \tag{22.2}$$

  The singular vectors $\mathbf{v}_i$ are known as the **right singular vectors** and the singular vectors $\mathbf{u}_i$ are known as the **left singular vectors**.

- There are precisely $r = min(n, m)$ non-zero singular values. The singular vectors $\mathbf{v}_i$ are the eigenvectors of $\mathbf{A}^T\mathbf{A}$, and the singular vectors $\mathbf{u}_i$ are the eigenvectors of $\mathbf{A}\mathbf{A}^T$. The r non-zero eigenvalues of $\mathbf{A}^T\mathbf{A}$ and $\mathbf{A}\mathbf{A}^T$ are σ_i^2.

- The $n \times m$ matrix $\mathbf{A}$ has a *singular value decomposition* (SVD) of the form

$$\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T \tag{22.3}$$

  where $\mathbf{U}$ is an $n \times r$ orthogonal matrix whose columns are $\mathbf{u}_i$, $\mathbf{\Sigma}$ is an $r \times r$ diagonal matrix with r non-zero entries σ_i, and $\mathbf{V}$ is an $m \times r$ orthogonal matrix whose columns are $\mathbf{v}_i$. Please note that this version of the SVD is called the *reduced* or *economy* SVD - there is a more general form but this is the most useful in a practical setting.

Exercise 22.1

1. Read "The Big Idea" again!

2x3 1x2 1x3

2. Let's assume that $\mathbf{A}$ is a 3×2 matrix. What is the size of $\mathbf{A}^T$? What is the size of $\mathbf{v}_i$ and $\mathbf{u}_i$? What is the size of $\mathbf{A}^T\mathbf{A}$ and $\mathbf{A}\mathbf{A}^T$? How many eigenvalues will $\mathbf{A}^T\mathbf{A}$ have? How many eigenvalues will $\mathbf{A}\mathbf{A}^T$ have? What must be true about these eigenvalues according to "The Big Idea"?

size
$A^TA = 2\times2$
$AA^T = 3\times3$

3) $A v_i = \sigma_i u_i$ (22.1)
$A^TAv = \sigma_i A^T u_i$
$A^TAv = \sigma_i^2 v_i$
↑ e.g. of A^TA ← Times scalar

3. Show that σ_i^2 and $\mathbf{v}_i$ are the eigenvalues and eigenvectors of $\mathbf{A}^T\mathbf{A}$ by multiplying Equation (22.1) by $\mathbf{A}^T$ and then using Equation (22.2) to simplify.
4. Show that σ_i^2 and $\mathbf{u}_i$ are the eigenvalues and eigenvectors of $\mathbf{A}\mathbf{A}^T$ by multiplying Equation (22.2) by $\mathbf{A}$ and then using Equation (22.1) to simplify.

4) $A^T u_i = \sigma_i v_i$ (22.2)
$AA^T u_i = \sigma_i A v_i$
$AA^T u_i = \sigma_i^2 u_i$

5. Take the transpose of Equation (22.2) and justify the use of the term **left singular vector** for $\mathbf{u}_i$.

u^T is on the left of A

6. Why is it valid to write

Equivalent to 22.1

$$\mathbf{A}[\mathbf{v}_1 \ldots \mathbf{v}_r] = [\mathbf{u}_1 \ldots \mathbf{u}_r]\begin{bmatrix}\sigma_1 & & \\ & \cdot & \\ & & \sigma_r\end{bmatrix}$$

and why does this imply Equation (22.3)? multiply Both sides by V^T

5) $A^T u_i = \sigma_i v_i$ (22.2)
$u_i^T A = \sigma_i v_i$

7. Why does Equation (22.3) imply that

This is 22.3 expanded out

$$\mathbf{A} = \sigma_1\mathbf{u}_1\mathbf{v}_1^T + \sigma_2\mathbf{u}_2\mathbf{v}_2^T + \ldots + \sigma_r\mathbf{u}_r\mathbf{v}_r^T \tag{22.4}$$

8. An $n \times m$ matrix has nm data values. How many data values do you need to store σ_1, $\mathbf{u}_1$, and $\mathbf{v}_1$? What kind of compression ratio would you have if you only stored the first singular value and the first singular vectors? $1+m+n$ $\frac{1+m+n}{mn}$

22.2 SVD - Example [20 minutes]

Consider the rectangular matrix

$$\mathbf{A} = \begin{bmatrix}1 & 2\\ 3 & 4\\ 5 & 6\\ 7 & 8\end{bmatrix}$$

Since this matrix is 4×2 we will form the 2×2 matrix $\mathbf{A}^T\mathbf{A}$,

$$\mathbf{A}^T\mathbf{A} = \begin{bmatrix}84 & 100\\ 100 & 120\end{bmatrix}$$

The eigenvalues of $\mathbf{A}^T\mathbf{A}$ are 203.6071 and 0.3929 respectively. The singular values are the square roots of these, namely $\sigma_1 = 14.2691$ and $\sigma_2 = 0.6268$ respectively. The associated eigenvectors are

$$\mathbf{v}_1 = \begin{bmatrix}0.6414\\ 0.7672\end{bmatrix}$$

and

$$\mathbf{v}_2 = \begin{bmatrix} -0.7672 \\ 0.6414 \end{bmatrix}$$

so that the matrix $\mathbf{\Sigma}$ is

$$\mathbf{\Sigma} = \begin{bmatrix} 14.2691 & 0 \\ 0 & 0.6268 \end{bmatrix}$$

and the matrix $\mathbf{V}$ is

$$\mathbf{V} = \begin{bmatrix} 0.6414 & -0.7672 \\ 0.7672 & 0.6414 \end{bmatrix}$$

Please note that each of the columns of $\mathbf{V}$ could be multiplied by -1 - the eigenvectors are only unique up to their direction (and the opposite direction).

To determine the $\mathbf{U}$ matrix, we recall that

$$\mathbf{A}\mathbf{v}_i = \sigma_i \mathbf{u}_i$$

which we can re-arrange and solve for $\mathbf{u}_i$

$$\mathbf{u}_i = \frac{1}{\sigma_i}\mathbf{A}\mathbf{v}_i$$

In this case $\mathbf{u}_1$ is given by

$$\mathbf{u}_1 = \frac{1}{14.2691}\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \\ 7 & 8 \end{bmatrix}\begin{bmatrix} 0.6414 \\ 0.7672 \end{bmatrix}$$

$$\mathbf{u}_1 = \begin{bmatrix} 0.1525 \\ 0.3499 \\ 0.5474 \\ 0.7448 \end{bmatrix}$$

and $\mathbf{u}_2$ is given by

$$\mathbf{u}_2 = \frac{1}{0.6268}\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \\ 7 & 8 \end{bmatrix}\begin{bmatrix} -0.7672 \\ 0.6414 \end{bmatrix}$$

$$\mathbf{u}_2 = \begin{bmatrix} 0.8227 \\ 0.4214 \\ 0.0201 \\ -0.3812 \end{bmatrix}$$

so that the $\mathbf{U}$ matrix is

$$\mathbf{U} = \begin{bmatrix} 0.1525 & 0.8227 \\ 0.3499 & 0.4214 \\ 0.5474 & 0.0201 \\ 0.7448 & -0.3812 \end{bmatrix}$$

The original matrix $\mathbf{A}$ therefore has the SVD

$$\mathbf{A} = \begin{bmatrix} 0.1525 & 0.8227 \\ 0.3499 & 0.4214 \\ 0.5474 & 0.0201 \\ 0.7448 & -0.3812 \end{bmatrix}\begin{bmatrix} 14.2691 & 0 \\ 0 & 0.6268 \end{bmatrix}\begin{bmatrix} 0.6414 & -0.7672 \\ 0.7672 & 0.6414 \end{bmatrix}^T$$

To check this we can define the matrix $\mathbf{A}$ in MATLAB and call the *svd* function with the "economy" option.

```
>> A = [1 2;3 4;5 6;7 8];
>> [U,Sigma,V] = svd(A,'econ')
```

Exercise 22.2

$0.3227 \to 0.8226$

all vals negated in U

1. Compare our result to the output from the use of the *svd* function and explain any discrepancies.
2. Confirm that the original matrix can be reconstituted by calculating

$$\sigma_1 \mathbf{u}_1 \mathbf{v}_1^T + \sigma_2 \mathbf{u}_2 \mathbf{v}_2^T$$

✓

3. How good is the reconstruction if we only keep the first part?

Not terrible, ±0.4

Exercise 22.3

1. Find the SVD of the following matrix by working through the steps outlined in this section. (You can use `eig` in MATLAB to get the relevant eigenvalues and eigenvectors.)

$$\begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{bmatrix}$$

2. Now use the `svd` command in MATLAB and check your work makes sense. You will need to use the "economy" option. ✓

22.3 SVD and User-Movie Data [10 mins]

22.4 Preview of the Homework and Project [10 minutes]

$$A^TA = \begin{bmatrix} 3 & 2 \\ 2 & 3 \\ 2 & -2 \end{bmatrix} \begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{bmatrix} = \begin{bmatrix} 13 & 12 & 2 \\ 12 & 13 & -2 \\ 2 & -2 & 8 \end{bmatrix}$$

$$v_1 = \begin{bmatrix} 0.2357 \\ -0.2357 \\ 0.9428 \end{bmatrix} \quad \sigma_1 = 3$$

$$v_2 = \begin{bmatrix} 0.7071 \\ 0.7071 \\ 0 \end{bmatrix} \quad \sigma_2 = 5$$

$$\begin{array}{cc} 2\times 3 & 2\times 1 \end{array}$$

$$u_1 = \frac{1}{\sigma_1} A v_1 = \frac{1}{3} \begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{bmatrix} \begin{bmatrix} 0.2357 \\ -0.2357 \\ 0.9428 \end{bmatrix} = \begin{bmatrix} .7071 \\ -.7071 \end{bmatrix}$$

$$u_2 = \begin{bmatrix} -.7071 \\ -.7071 \end{bmatrix}$$