Harp - DAAL SVD

Internship at IU

Anurag Sharma

Department of Mathematics and Statistics
IIT Kanpur

May 30, 2017

Outline

1. Introduction
   - Singular Value Decomposition
   - Details
2. Algorithm Details
   - Algorithm (Manually)
   - Algorithm Batch Mode
   - Algorithm Distributed Mode

Outline

Singular Value Decomposition

Definition

Factorizing a matrix in more mathematically intuitive and useful format

Mathematically:

$\mathcal{A} = \mathcal{U}^T \Sigma \mathcal{V}$

where $\mathcal{A}$ be any matrix; $\mathcal{U}$, $\mathcal{V}$ are orthogonal matrices and Σ is a diagonal matrix with positive real entries.

Outline

1. Introduction
   - Singular Value Decomposition
   - Details
2. Algorithm Details
   - Algorithm (Manually)
   - Algorithm Batch Mode
   - Algorithm Distributed Mode

Details

The SVD arises from finding an orthogonal basis for the row space that gets transformed into an orthogonal basis for the column space:

$$Av_i = \sigma_i u_i$$

v_i is a vector in row space and u_i is a vector in column space.
To find the orthogonal basis of row space we use Gram Schmidt Process.
The vectors in the nullspaces of A and A^T are taken care of by zeros on the diagonal of Σ

Outline

Method

- Calculate AA^T and A^TA
- Find eigenvectors of AA^T and A^TA
- Find eigenvalues of AA^T and A^TA

$$A^TA = \begin{bmatrix} 4 & -3 \\ 4 & 3 \end{bmatrix} \begin{bmatrix} 4 & 4 \\ -3 & 3 \end{bmatrix} = \begin{bmatrix} 25 & 7 \\ 7 & 25 \end{bmatrix}.$$

Two orthogonal eigenvectors of A^TA are $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$. To get an orthonormal basis, let $\mathbf{v}_1 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$ and $\mathbf{v}_2 = \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}$. These have eigenvalues $\sigma_1^2 = 32$ and $\sigma_2^2 = 18$. We now have:

$$\overset{A}{\begin{bmatrix} 4 & 4 \\ -3 & 3 \end{bmatrix}} = \overset{U}{\begin{bmatrix} & \\ & \end{bmatrix}} \overset{\Sigma}{\begin{bmatrix} 4\sqrt{2} & 0 \\ 0 & 3\sqrt{2} \end{bmatrix}} \overset{V^T}{\begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}}.$$

Similarly we calculate AA^T and find eigenvectors and normalize them.

$$
\begin{aligned}
AA^T &= \begin{bmatrix} 4 & 4 \\ -3 & 3 \end{bmatrix} \begin{bmatrix} 4 & -3 \\ 4 & 3 \end{bmatrix} \\
&= \begin{bmatrix} 32 & 0 \\ 0 & 18 \end{bmatrix}.
\end{aligned}
$$

Finally we get:

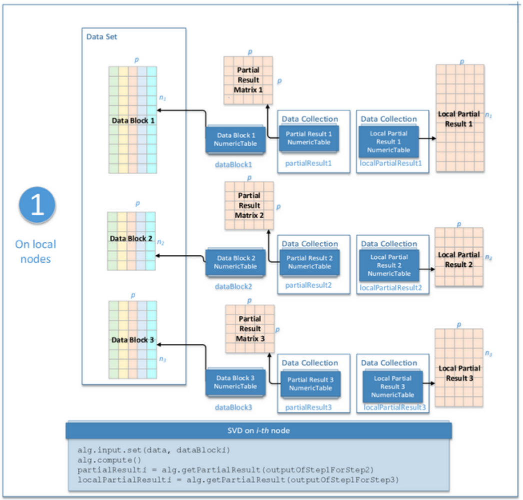
$$
\overset{A}{\begin{bmatrix} 4 & 4 \\ -3 & 3 \end{bmatrix}} = \overset{U}{\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}} \overset{\Sigma}{\begin{bmatrix} 4\sqrt{2} & 0 \\ 0 & 3\sqrt{2} \end{bmatrix}} \overset{V^T}{\begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}}.
$$

Outline

In batch mode we use modified QR decomposition to factorize and find eigenvectors.
For details : https://web.stanford.edu/class/cme335/lecture6.pdf

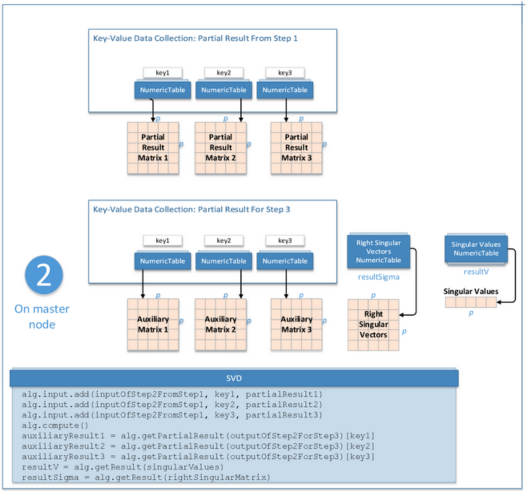
Outline

Step 1 - on Local Nodes



Algorithm
Distributed Mode

Step 2 - on Master Node



Algorithm
Distributed Mode