Basic Mathematics, Fall 2020

Karen Keryan,
ASDS, YSU

October 22, 2020

If the image is all black, then all the entries of the corresponding matrix are euqal.

Example

Instead of sending $A = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix}$,

send $A = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \end{bmatrix}$.

Instead of sending 25 *numbers, send just 10 numbers.*

Low rank Images or Memory reduction

Example

Instead of sending $A = \begin{bmatrix} r & r & r & r & r & r \\ r & r & r & r & r & r \\ b & b & b & b & b & b \\ b & b & b & b & b & b \\ o & o & o & o & o & o \\ o & o & o & o & o & o \end{bmatrix}$,

send $A = \begin{bmatrix} r \\ r \\ b \\ b \\ o \\ o \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$.

Example

If $\mathrm{rk}(A) = 1$, *then it can be represented as* $A = \mathbf{u}_1 \mathbf{v}_1^T$.
If $\mathrm{rk}(A) = 2$, *then it can be represented as* $A = \mathbf{u}_1 \mathbf{v}_1^T + \mathbf{u}_2 \mathbf{v}_2^T$.
In general, if $A = U \Sigma V^T$ *and* $\mathrm{rk}(A) = r$, *then*

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

where now $\{\mathbf{u}_i\}, \{\mathbf{v}_j\}$ *are orthonormal sets of vectors.*
And we send intead of mn *numbers just* $r(m + m + 1)$ *numbers.*
To approximate A, **we keep larger** σ_i**'s, and discard smaller** σ_i**'s.**

Proposition

$\mathbf{u}_1, \ldots, \mathbf{u}_r$ *is an orthonormal basis for the* **column space**.
$\mathbf{u}_{r+1}, \ldots, \mathbf{u}_m$ *is an orthonormal basis for the* **left nullspace** $N(A^T)$.
$\mathbf{v}_1, \ldots, \mathbf{v}_r$ *is an orthonormal basis for the* **row space**.
$\mathbf{v}_{r+1}, \ldots, \mathbf{v}_n$ *is an orthonormal basis for the* **nullspace** $N(A)$.

Proposition

If the SVD of A is $A = U\Sigma V^T$ and $\mathrm{rk}(A) = r$, then

$$AV_r = U_r\Sigma_r,$$

equivalently

$$A\begin{bmatrix}\mathbf{v}_1 & \dots & \mathbf{v}_r\end{bmatrix} = \begin{bmatrix}\mathbf{u}_1 & \dots & \mathbf{u}_r\end{bmatrix}\begin{bmatrix}\sigma_1 & & \\ & \ddots & \\ & & \sigma_r\end{bmatrix},$$

called **reduced SVD.**

Example

If $A = \mathbf{x}\mathbf{y}^T$, with unit vectors $\mathbf{x}$ and $\mathbf{y}$. What is the SVD of A?

Example

Find SVD of the matrix $A = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 0 \end{bmatrix}$

Note that the matrices $A^T A$ *and* AA^T *are diagonal*

$$AA^T = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 9 \end{bmatrix}, \quad A^T A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 9 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Example

Find SVD of the matrix $A = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 0 \end{bmatrix}$

Note that the matrices $A^T A$ *and* AA^T *are diagonal*

$$AA^T = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 9 \end{bmatrix}, \quad A^T A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 9 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Remark

Here we get $A = U\Sigma V^T = 3\mathbf{u}_1\mathbf{v}_1^T + 2\mathbf{u}_2\mathbf{v}_2^T + 1\mathbf{u}_3\mathbf{v}_3^T$, *and* $\sigma_1\mathbf{u}_1\mathbf{v}_1^T$ *picks out the largest number* $A_{34} = 3$ *in the original matrix* A.

Remark

Removing the zero (last) row of A *just removes the last row of* Σ

Example

The flags of Sweden and Finland have rank 2.

$$A_{Finland} = A_{Sweden} = \begin{bmatrix} 1 & 2 & 1 & 1 \\ 2 & 2 & 2 & 2 \\ 1 & 2 & 1 & 1 \end{bmatrix}.$$

The singular values are 5.4016 0.9069 0.0000 and

$$\sigma_1 \mathbf{u}_1 \mathbf{v}_1^T = \begin{bmatrix} 1.17004\ldots & 1.63478\ldots & 1.17004\ldots & 1.17004\ldots \\ 1.77594\ldots & 2.48134\ldots & 1.77594\ldots & 1.77594\ldots \\ 1.17004\ldots & 1.63478\ldots & 1.17004\ldots & 1.17004\ldots \end{bmatrix}$$

is an approximation to the matrix $A_{Finland} = A_{Sweden}$.

SVD for Pseudoinverse

The SVD can be used for computing the pseudoinverse of a matrix. Indeed, the pseudoinverse of the matrix A with singular-value decomposition $A = U\Sigma V^T$ is

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

where Σ^+ is the pseudoinverse of Σ, which is formed by replacing every non-zero diagonal entry by its reciprocal and transposing the resulting matrix.

E.g. for $\Sigma = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$ the pseudoinverse Σ^+ is $\Sigma^+ = \begin{bmatrix} \frac{1}{4} & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$.

The pseudoinverse is one way to solve linear least squares problems.

Total least squares minimization

Recall that if the SVD of A is $A = U\Sigma V^T$ then $A\mathbf{v}_i = \sigma_i \mathbf{u}_i$ for $i = 1, \dots, r$, where $r = \mathrm{rk}(A)$.

A total least squares problem

Let A be an invertible matrix. Determine the vector $\mathbf{x}$ which solves the minimization problem

$$\min \|A\mathbf{x}\| \text{ subject to } \|\mathbf{x}\| = 1,$$

equivalently

$$\min_{\mathbf{x}} \frac{\|A\mathbf{x}\|}{\|\mathbf{x}\|}.$$

The solution turns out to be the right-singular vector $\mathbf{v}_r$ corresponding to the smallest singular value σ_r, i.e.

$$\sigma_r = \min_{\mathbf{x}} \frac{\|A\mathbf{x}\|}{\|\mathbf{x}\|} = \frac{\|A\mathbf{v}_r\|}{\|\mathbf{v}_r\|} = \frac{\|\sigma_r \mathbf{u}_r\|}{\|\mathbf{v}_r\|}$$

Proposition

If the SVD of a matrix $A \in \mathbb{R}^{n\times m}$ *is* $A = U\Sigma V^T$*, then*

$$\max_{\mathbf{x}} \frac{\|A\mathbf{x}\|}{\|\mathbf{x}\|} =$$

Proposition

If the SVD of a matrix $A \in \mathbb{R}^{n \times m}$ is $A = U\Sigma V^T$, then

$$\max_{\mathbf{x}} \frac{\|A\mathbf{x}\|}{\|\mathbf{x}\|} = \sigma_1$$

and maximum is attained for $\mathbf{x} =$

Proposition

If the SVD of a matrix $A \in \mathbb{R}^{n \times m}$ is $A = U\Sigma V^T$, then

$$\max_{\mathbf{x}} \frac{\|A\mathbf{x}\|}{\|\mathbf{x}\|} = \sigma_1$$

and maximum is attained for $\mathbf{x} = \mathbf{v}_1$.

$$\max_{\mathbf{x}\in\mathbb{R}^n} \frac{\|A\mathbf{x}\|}{\|\mathbf{x}\|} = \max_{\mathbf{x}\in\mathbb{R}^n} \frac{\|U\Sigma(V^T\mathbf{x})\|}{\|V^T\mathbf{x}\|} = \max_{\mathbf{y}\in\mathbb{R}^n} \frac{\|U(\Sigma\mathbf{y})\|}{\|\mathbf{y}\|} = \max_{\mathbf{y}\in\mathbb{R}^n} \frac{\|\Sigma\mathbf{y}\|}{\|\mathbf{y}\|}$$

$$= \max_{\mathbf{y}} \frac{\sqrt{\sigma_1^2 y_1^2 + \ldots + \sigma_r^2 y_r^2}}{y_1^2 + \ldots + y_n^2} = \sigma_1.$$

Eckart-Young Theorem

Let the SVD of A be given by $A = \sum_{i=1}^{r} \sigma_i \mathbf{u}_i \mathbf{v}_i^T$. If $k < r = \operatorname{rank}(A)$ and $A_k = \sum_{i=1}^{k} \sigma_i \mathbf{u}_i \mathbf{v}_i^T$, then

$$\min_{\operatorname{rank}(B)=k} ||A - B||_2 = \|A - A_k\|_2 = \sigma_{k+1}$$

Example

The best rank 2 approximation of the matrix (of rank 3) $A = \begin{bmatrix} 4 & 2 & 1 \\ 6 & 3 & 2 \\ 2 & 2 & 3 \end{bmatrix}$ *is*

$$A_2 = \begin{bmatrix} 4.0265\ldots & 1.9287\ldots & 1.0294\ldots \\ 5.9809\ldots & 3.0510\ldots & 1.9788\ldots \\ 2.0034\ldots & 1.9907\ldots & 3.0038\ldots \end{bmatrix}$$

In many applications (such as PCA), we only require **a few eigenvectors**(with the largest eigenvalues). It would be wasteful to compute the full decomposition, and then discard all eigenvectors with eigenvalues that are beyond the first few. Iterative processes, which directly optimize only the first few eigenvectors, are computationally more efficient than a full eigendecomposition (or SVD). In the extreme case of only needing the first eigenvector, a simple method called the **power iteration** is very efficient. Power iteration chooses a random vector $\mathbf{x}_0 \notin \mathrm{null}(A)$ and follows the iteration

$$\mathbf{x}_{k+1} = \frac{A\mathbf{x}_k}{\|A\mathbf{x}_k\|}, k = 0, 1, ...$$

We always have $\|\mathbf{x}_k\| = 1$. This sequence of vectors converges to the eigenvector associated with the largest eigenvalue of A. The original Google PageRank algorithm uses such an algorithm for ranking web pages based on their hyperlinks.

Image compression

The popular "Lena" image (512 x 512, gray scale) is tested for the compression scheme. The figure below shows the results of the compression with different ranks used for the re-constructed images.

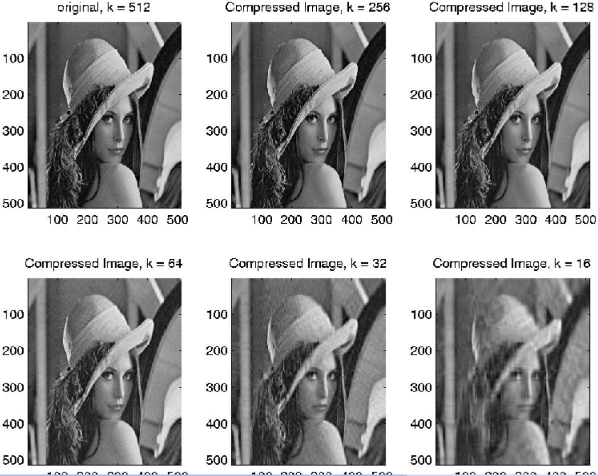

Image compression

Original Image:	16.4MB
Rank 100:	1.4MB
Rank 75:	1.0MB
Rank 50	0.7MB

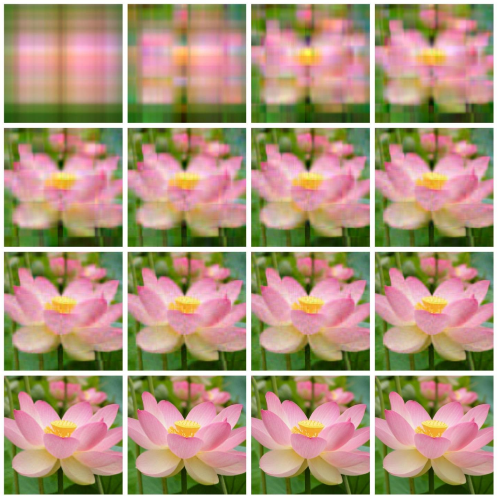

Image size 2800x2052 and its approximations with ranks $\{\{1,2,4,6\},\{8,10,12,14\},\{16,18,20,25\},\{50,75,100,\text{original image}\}\}$
The Matlab codes can be found at `http://www.math.utah.edu/~goller/F15_M2270/BradyMathews_SVDImage.pdf`

Some more fascinating applications of SVD can be found following the link
https://people.maths.ox.ac.uk/porterm/papers/s4.pdf