Lecture 5

Math Foundations Team

BITS Pilani
Pilani | Dubai | Goa | Hyderabad

Introduction

- In the previous lecture, we discussed eigenvalues and eigenvectors of matrices
- In this lecture, we will look at two related methods for factorizing matrices into canonical forms.
- The first one is known as Eigenvalue decompostion. It uses the concepts of eigenvalues and eigenvectors to generate the decomposition
- The second method known as singular value decomposition or SVD is applicable to all matrices

Diagonal Matrices

- A diagonal matrix is a matrix that has value zero on all off diagonal elements.

$$\mathcal{D} = \begin{bmatrix} d_1 & & \\ & \ddots & \\ & & d_n \end{bmatrix}$$

- For a diagonal matrix $\mathcal{D}$, the determinant is the product of its diagonal entries.
- A matrix power $\mathcal{D}^k$ is given by each diagonal element raised to the power k.
- Inverse of a diagonal matrix is obtained by taking inverse of non-zero diagonal entry.

Diagonalizable Matrices

- A matrix $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ is diagonalizable if there exists an invertible matrix $\boldsymbol{P} \in \mathbb{R}^{n \times n}$ and a diagonal matrix $\mathcal{D}$ such that $\mathcal{D} = \boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$
- In the definition of diagonalization, it is required that $\boldsymbol{P}$ is an invertible matrix. Assume $p_1, p_2, \ldots.., p_n$ are the n columns of $\boldsymbol{P}$
- Rewriting we get $\boldsymbol{A}\boldsymbol{P} = \boldsymbol{P}\mathcal{D}$. By observing that $\mathcal{D}$ is a diagonal matrix, we can simplify as

$$\boldsymbol{A}p_i = \lambda_i p_i$$

where λ_i is the i^{th} diagonal entry in $\mathcal{D}$.

Diagonalizable Matrix

- Consider a square matrix

$$\boldsymbol{A} = \begin{bmatrix} 1 & 4 \\ 2 & 3 \end{bmatrix}$$

- Consider the invertible matrix

$$\boldsymbol{P} = \begin{bmatrix} -2 & 1 \\ 1 & 1 \end{bmatrix}$$

- Now consider the product $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ as follows

$$\begin{bmatrix} -2 & 1 \\ 1 & 1 \end{bmatrix}^{-1} \cdot \begin{bmatrix} 1 & 4 \\ 2 & 3 \end{bmatrix} \cdot \begin{bmatrix} -2 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & 5 \end{bmatrix}$$

Eigendecomposition of a matrix

- Recall the existence of eigenvalues and eigenvectors for square matrices
- Eigenvalues can be used to create a matrix decomposition known as Eigenvalue Decomposition
- A square matrix $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ can be factored into

$$\boldsymbol{A} = \boldsymbol{P}\boldsymbol{D}\boldsymbol{P}^{-1}$$

- where $\boldsymbol{P}$ is an invertible matrix of eigenvectors of $\boldsymbol{A}$ assuming we can find n eigenvectors that form a basis of $\mathbb{R}^n$
- and $\mathcal{D}$ is a diagonal matrix whose diagonal entries are the eigenvalues of $\boldsymbol{A}$

Example of Eigendecomposition

Let us compute the eigendecomposition of the matrix $\boldsymbol{A}$

$$\boldsymbol{A} = \begin{bmatrix} 2.5 & -1 \\ -1 & 2.5 \end{bmatrix}$$

- Step 1: Find the eigenvalues and eigenvectors

$$\boldsymbol{A} - \lambda \boldsymbol{I} = \begin{bmatrix} 2.5 - \lambda & -1 \\ -1 & 2.5 - \lambda \end{bmatrix}$$

- The characteristic equation is given by $\det(\boldsymbol{A} - \lambda \boldsymbol{I}) = 0$
- This leads to the equation $\lambda^2 - 5\lambda + \frac{21}{4} = 0$
- Solving the quadratic equation gives us $\lambda_1 = 3.5$ and $\lambda_2 = 1.5$

Example of Eigendecomposition

- The eigenvector corresponing to $\lambda_1 = 3.5$ is derived as
$p_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix}$
- The eigenvector corresponing to $\lambda_1 = 1.5$ is derived as
$p_2 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$
- Step 2 : Construct the matrix $\boldsymbol{P}$ to diagonalize $\boldsymbol{A}$

$$\boldsymbol{P} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

Example of Eigendecomposition

- The inverse of matrix $\boldsymbol{P}$ is given by

$$\boldsymbol{P}^{-1} = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

- The eigendecompostion of the matrix A is given by

$$\boldsymbol{A} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 3.5 & 0 \\ 0 & 1.5 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

- In summary we have obtained the required matrix factorization using eigenvalues and eigenvectors.

Symmetric Matrices and Diagonalizability

- Recall that a matrix $\boldsymbol{A}$ is called symmetric matrix if $\boldsymbol{A} = \boldsymbol{A}^T$

$$\boldsymbol{A} = \begin{bmatrix} -2 & 1 \\ 1 & 1 \end{bmatrix}$$

- A Symmetric matrix $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ can always be diagonalized.
- This follows directly from the spectral theorem discussed in previous lecture
- Moreover the spectral theorem states that we can find an orthogonal matrix $\boldsymbol{P}$ of eigenvectors of $\boldsymbol{A}$.

Motivation for Singular Value Decomposition

- The singular value decomposition or (SVD) of a matrix is a central matrix decomposition method in linear algebra.
- The eigenvalue decomposition is applicable to square matrices only.
- The singular value decomposition exists for all rectangular matrices
- SVD involves writing a matrix as a product of three matrices $\boldsymbol{U}, \Sigma$ and $\boldsymbol{V}^T$.
- The three component matrices are derived by applying eigenvalue decomposition discussed previously

Singular Value Decomposition Theorem

- Let $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ be a rectangular matrix. Assume that $\boldsymbol{A}$ has rank r.
- The Singular value decomposition of $\mathcal{A}$ is defined as

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

- $\boldsymbol{U} \in \mathbb{R}^{m \times m}$ is an orthogonal matrix with column vectors u_i where $i = 1, \ldots, m$
- $\boldsymbol{V} \in \mathbb{R}^{n \times n}$ is an orthogonal matrix with column vectors v_j where $j = 1, \ldots, n$
- $\boldsymbol{\Sigma}$ is an m $\times$ n matrix with $\boldsymbol{\Sigma}_{ii} = \sigma_i > 0$
- The diagonal entries $\sigma_i, i = 1, \ldots, r$ of $\boldsymbol{\Sigma}$ are called the singular values.
- By convention, the singular values are ordered i.e $\boldsymbol{\Sigma}_{ii} > \boldsymbol{\Sigma}_{jj}$ where $i < j$.

Properties of Σ

- The singular value matrix $\boldsymbol{\Sigma}$ is unique.
- Observe that the $\boldsymbol{\Sigma} \in \mathbb{R}^{m \times n}$matrix is rectangular. In particular, $\boldsymbol{\Sigma}$ is of the same size as $\boldsymbol{A}$.
- This means that $\boldsymbol{\Sigma}$ has a diagonal submatrix that contains the singular values and needs additional zero padding.
- Specifically, if $m > n$, then the matrix $\boldsymbol{\Sigma}$ has diagonal structure up to row n and then consists of zero rows.
- If $m < n$, the matrix $\boldsymbol{\Sigma}$ has a diagonal structure up to column m and columns that consist of 0 from $m + 1$ to n.

Construction of V

- It can be observed that

$$\boldsymbol{A}^T\boldsymbol{A} = \boldsymbol{V}\Sigma^T\Sigma\boldsymbol{V}^T$$

- Since $\boldsymbol{A}^T\boldsymbol{A}$ has the following eigendecomposition

$$\boldsymbol{A}^T\boldsymbol{A} = \boldsymbol{P}\mathcal{D}\boldsymbol{P}^T$$

- Therefore, the eigenvectors of $\boldsymbol{A}^T\boldsymbol{A}$ that compose $\boldsymbol{P}$ are the right-singular vectors $\boldsymbol{V}$ of $\boldsymbol{A}$.
- The eigenvalues of $\boldsymbol{A}^T\boldsymbol{A}$ are the squared singular values of $\boldsymbol{\Sigma}$

Construction of U

- It can be observed that

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

- Since $\boldsymbol{A}\boldsymbol{A}^T$ has the following eigendecomposition

$$\boldsymbol{A}\boldsymbol{A}^T = \boldsymbol{S}\boldsymbol{D}\boldsymbol{S}^T$$

- Therefore, the eigenvectors of $\boldsymbol{A}\boldsymbol{A}^T$ that compose $\boldsymbol{S}$ are the left-singular vectors $\boldsymbol{U}$ of $\boldsymbol{A}$

Construction of U continued

- $\boldsymbol{A} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^T$ can be rearranged to obtain a simple formulation for u_i
- By postmultiplying by $\boldsymbol{V}$ we get $\boldsymbol{A}\boldsymbol{V} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^T\boldsymbol{V}$
- By observing that $\boldsymbol{V}$ is orthogonal we obtain a simple form

$$\boldsymbol{A}\boldsymbol{V} = \boldsymbol{U}\boldsymbol{\Sigma}$$

- This is equivalent to the following

$$u_i = \frac{1}{\sigma_i}\boldsymbol{A}v_i \quad \forall i = 1, 2, ..., r$$

Computing Singular Value Decomposition 1

- We want to find SVD of the following rectangular matrix $\boldsymbol{A}$

$$\boldsymbol{A} = \begin{bmatrix} 1 & 0 & 1 \\ -2 & 1 & 0 \end{bmatrix}$$

- Let us consider the matrix $\boldsymbol{A}^T\boldsymbol{A}$ derived from $\boldsymbol{A}$ given by

$$\boldsymbol{A}^T\boldsymbol{A} = \begin{bmatrix} 5 & -2 & 1 \\ -2 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

- It is a symmetric matrix

Computing Singular Value Decomposition 2

- Derive the eigendecomposition of $\boldsymbol{A}^T\boldsymbol{A}$ in the form $\boldsymbol{P}\mathcal{D}\boldsymbol{P}^T$
- The matrix $\boldsymbol{P}$ is given by

$$\boldsymbol{P} = \begin{bmatrix} \frac{5}{\sqrt{30}} & 0 & \frac{-1}{\sqrt{6}} \\ \frac{-2}{\sqrt{30}} & \frac{1}{\sqrt{5}} & \frac{-2}{\sqrt{6}} \\ \frac{1}{\sqrt{30}} & \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{6}} \end{bmatrix}$$

- The matrix $\mathcal{D}$ is given by

$$\mathcal{D} = \begin{bmatrix} 6 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Computing Singular Value Decomposition 3

Now we construct the singular value matrix $\mathbf{\Sigma}$

- The matrix $\mathbf{\Sigma}$ has the dimension same as $\boldsymbol{A}$. In this case $\mathbf{\Sigma}$ is hence a 2×3 matrix.
- The diagonal entries of submatrix is obtained by taking square root of 6 and 1 respectively
- Singular-value matrix $\mathbf{\Sigma}$ is given by:

$$\mathbf{\Sigma} = \begin{bmatrix} \sqrt{6} & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

- The last column is a column of zeros only

Computing Singular Value Decomposition 4

Left singular vectors as the normalized image of the right singular vectors. Recall that $u_i = \frac{1}{\sigma_i}\boldsymbol{A}v_i$

- The first vector

$$u_1 = \frac{1}{\sigma_1}\boldsymbol{A}v_1 = \begin{bmatrix} \frac{1}{\sqrt{5}} \\ \frac{-2}{\sqrt{5}} \end{bmatrix}$$

- The second vector

$$u_2 = \frac{1}{\sigma_2}\boldsymbol{A}v_2 = \begin{bmatrix} \frac{2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \end{bmatrix}$$

Final Step : Combining U, Σ and V

We compile all the three matrices together to generate the SVD

- $$\boldsymbol{A} = \begin{bmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \\ \frac{-2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{bmatrix} \begin{bmatrix} \sqrt{6} & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \frac{5}{\sqrt{30}} & 0 & \frac{-1}{\sqrt{6}} \\ \frac{-2}{\sqrt{30}} & \frac{1}{\sqrt{5}} & \frac{-2}{\sqrt{6}} \\ \frac{1}{\sqrt{30}} & \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{6}} \end{bmatrix}^T$$
- The matrix $\boldsymbol{U}$ is an 2×2 matrix satifying orthogonality property.
- The matrix $\boldsymbol{V}$ is an 3×3 matrix satifying orthogonality property.

Comparing SVD and EVD

- The left-singular vectors of $\boldsymbol{A}$ are eigenvectors of $\boldsymbol{A}\boldsymbol{A}^T$
- The right-singular vectors of $\boldsymbol{A}$ are eigenvectors of $\boldsymbol{A}^T\boldsymbol{A}$
- The non-zero singular values of $\boldsymbol{A}$ are the square roots of the nonzero eigenvalues of $\boldsymbol{A}^T\boldsymbol{A}$.
- The SVD always exists for any matrix in $\mathbb{R}^{m\times n}$
- The eigendecomposition is only defined for square matrices in $\mathbb{R}^{n\times n}$ and only exists if we can find a basis of eigenvectors of $\mathbb{R}^n$

Comparing SVD and EVD

- The vectors in the eigendecomposition matrix $\boldsymbol{P}$ are not necessarily orthogonal but the matrices in decomposition are inverse of each other.
- On the other hand, the vectors in the matrices $\boldsymbol{U}$ and $\boldsymbol{V}$ in the SVD are orthonormal but $\boldsymbol{U}$ and $\boldsymbol{V}$ may not be inverse of each other.
- A key difference between the eigendecomposition and the SVD is that in the SVD, domain and codomain can be of different dimensions
- In the SVD, the entries in the diagonal matrix $\boldsymbol{\Sigma}$ are all real and nonnegative but in eigendecomposition diagonal matrix entries need not be real always.

Matrix Approximation

- We considered the SVD as a way to factorize $\boldsymbol{A} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^T$ into the product of three matrices, where $\boldsymbol{U}$ and $\boldsymbol{V}$ are orthogonal and $\boldsymbol{\Sigma}$ contains the singular values on its main diagonal.
- Instead of doing the full SVD factorization, we will now investigate how the SVD allows us to represent a matrix $\boldsymbol{A}$ as a sum of simpler matrices $\boldsymbol{A}_i$
- This representation which lends itself to a matrix approximation scheme that is cheaper to compute than the full SVD.

Matrix Approximation continued

- A matrix $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ of rank r can be written as a sum of rank-1 matrices so that $\boldsymbol{A} = \sum_{i=1}^{r} \sigma_i u_i v_i^T$
- The diagonal structure of the singular value matrix $\boldsymbol{\Sigma}$ multiplies only matching left and right singular vectors $u_i v_i^T$ and scales them by the corresponding singular value σ_i.
- All terms $\sigma_i u_i v_i^T$ vanish for $i \neq j$ because $\boldsymbol{\Sigma}$ is a diagonal matrix.
- Any term for $i > r$ would vanish because the corresponding singular value is 0.

Rank k Approximation

- We summed up the r individual rank-1 matrices to obtain a rank r matrix $\boldsymbol{A}$.
- If the sum does not run over all matrices $\boldsymbol{A}_i \;\; i = 1, ..., r$ but only up to an intermediate value k we obtain a rank-k approximation
- The approximation represented by $\hat{\boldsymbol{A}}(k)$ is defined as follows

$$\hat{\boldsymbol{A}}(k) = \sum_{i=1}^{k} \sigma_i u_i v_i^T$$

- To measure the difference between $\boldsymbol{A}$ and its rank-k approximation we need the notion of a norm which is introduced next

Spectral Norm of a matrix

- We introduce the notation of a subscript in the matrix norm
- Spectral Norm of a Matrix. For $\boldsymbol{x} \in \mathbb{R}^n$, $x \neq \mathbf{0}$, the spectral norm norm of a matrix $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ is defined as

$$\|\boldsymbol{A}\|_2 = \max_{\boldsymbol{x}} \frac{\|\boldsymbol{A}\boldsymbol{x}\|_2}{\|\boldsymbol{x}\|_2}$$

  where $\|\boldsymbol{y}\|_2$ is the euclidean norm of $\boldsymbol{y}$
- Theorem : The spectral norm of a matrix $\boldsymbol{A}$ is its largest singular value

Example : Spectral Norm of a matrix

- Example : Consider the following matrix $\boldsymbol{A}$

$$\boldsymbol{A} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$$

- Singular value decomposition of this matrix will provide the matrix $\boldsymbol{\Sigma}$ as follows

$$\boldsymbol{\Sigma} = \begin{bmatrix} 5.465 & 0 \\ 0 & 0.366 \end{bmatrix}$$

- The 2 singular values are 5.4650 and 0.366.
- By definition the spectral norm is the largest singular value.
- Hence, the spectral norm is 5.4650

Application

- Consider the individual saturation levels of each pixel in the original black and white image as the numerical entries in a matrix.
- Compute the SVD of that matrix and remove the singular values (from smallest to largest), converting the modified matrices (with removed values) back into a series of images.
- This process of decomposition can reduce the image storage size without losing the quality needed to fully represent the image

Application

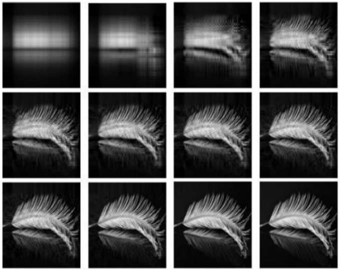

Figure 2: Number of Singular Values: {1, 2, 5, 10}{15, 18, 24, 30}{35, 60, 120, 680}

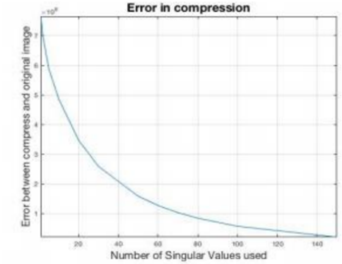

Figure 3: Error in compression when applied to the grayscale feather image.

The original image has approximately 680 singular values, but there is a close resemblance to the original image using only 120 singular values .
(https://www.lagrange.edu/academics/undergraduate/undergraduate-research/citations/18-Citations2020.Compton.pdf)