ECE 586: Vector Space Methods
Lecture 20: Singular Value Decomposition

Henry D. Pfister
Duke University

8: Eigenvalue Decomposition

Definition

Let V be a vector space over F and let $T: V \to V$ be a linear operator. An eigenvalue of T is a scalar $\lambda \in F$ such that there exists a non-zero vector $\underline{v} \in V$ with $T\underline{v} = \lambda\underline{v}$. Any vector $\underline{v}$ such that $T\underline{v} = \lambda\underline{v}$ is called an eigenvector of T associated with the eigenvalue value λ.

Definition

The square matrix B is diagonalizable if there is an invertible matrix S (whose columns are eigenvectors) such that $S^{-1}BS = \Lambda$ is diagonal.

Theorem

Any Hermitian matrix B can be diagonalized by a unitary matrix U so that $U^H BU = \Lambda$ is a real-valued diagonal matrix.

Prove eigenval/vec properties for Hermitian matrices

Note: Matrices $A^H A$ and AA^H always Hermitian and positive semidefinite

Definition

The singular value decomposition (SVD) of a rank-r matrix $A \in \mathbb{C}^{m \times n}$ is

$$A = U \Sigma V^H = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_1^H \\ V_2^H \end{bmatrix} = U_1 \Sigma_1 V_1^H,$$

where (i) $U \in \mathbb{C}^{m \times m}$ and $V \in \mathbb{C}^{n \times n}$ are unitary and (ii) $U_1 \in \mathbb{C}^{m \times r}$, $U_2 \in \mathbb{C}^{m \times m-r}$, $V_1 \in \mathbb{C}^{n \times r}$, and $V_2 \in \mathbb{C}^{n \times n-r}$ have orthonormal columns. The diagonal matrix $\Sigma_1 \in \mathbb{R}^{r \times r}$ contains the non-zero singular values

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0.$$

The factorization $A = U \Sigma V^H$ is called the full SVD of the matrix A while the factorization $A = U_1 \Sigma_1 V_1$ is called the compact SVD of A.

The compact SVD of a rank-r matrix retains only the r columns of U, V associated with non-zero singular values.

Idea to find orthonormal changes of basis U, V so that $U^H AV$ is diagonal

- Let $\underline{v}_1, \ldots, \underline{v}_r$ be orthonormal eigenvectors of $A^H A$ with positive eigenvalues $\sigma_1^2, \ldots, \sigma_r^2$. Then,

$$\|A\underline{v}_i\|^2 = \underline{v}_i^H (A^H A\underline{v}_i) = \underline{v}_i^H (\sigma_i^2 \underline{v}_i) = \sigma_i^2$$

- This implies that $\|A\underline{v}_i\| = \sigma_i$. So $\underline{u}_i = \frac{1}{\sigma_i} A\underline{v}_i$ has $\|\underline{u}_i\| = 1$ and

$$AA^H \underline{u}_i = \frac{1}{\sigma_i} AA^H A\underline{v}_i = \frac{1}{\sigma_i} \sigma_i^2 A\underline{v}_i = \sigma_i^2 \underline{u}_i$$

$$\underline{u}_j^H \underline{u}_i = \left(\frac{1}{\sigma_j} A\underline{v}_j\right)^H \left(\frac{1}{\sigma_i} A\underline{v}_i\right) = \frac{1}{\sigma_i \sigma_j} \underline{v}_j^H (A^H A)\underline{v}_i = \delta_{i,j}$$

- For $U_1 = [\underline{u}_1, \ldots, \underline{u}_r]$ and $V_1 = [\underline{v}_1, \ldots, \underline{v}_r]$, this gives $AV_1 = U_1 \Sigma_1$ where Σ_1 is a $r \times r$ diagonal matrix with diagonal entries $\sigma_1, \ldots, \sigma_r$

- If cols of V_2 are an orthonormal basis for $\mathcal{N}(A)$, then $A[V_1 \; V_2] = U_1[\Sigma_1 \; 0]$. Thus, right multiplication by $V^H = [V_1^H \; V_2^H]$ gives the compact SVD

$$A = U_1 \Sigma_1 V_1^H,$$

where the columns of U_1, V_1 are orthonormal bases for $\mathcal{R}(A), \mathcal{R}(A^H)$

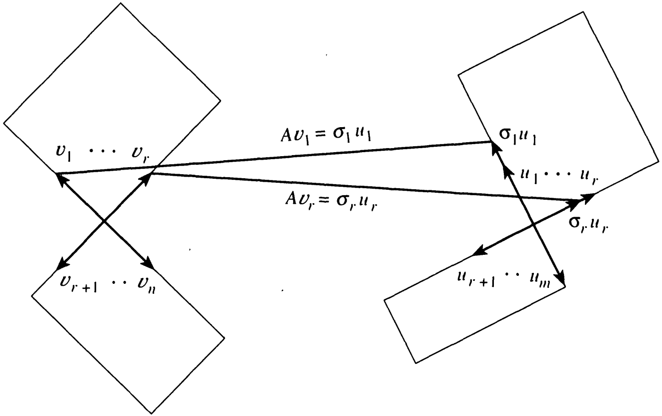

Figure 3. Orthonormal bases that diagonalize A.

For $V = [\underline{v}_1, \ldots, \underline{v}_n]$ and $U = [\underline{u}_1, \ldots, \underline{u}_m]$, $AV = U\Sigma$ where $\Sigma \in \mathbb{R}^{m \times n}$ has diagonal $\sigma_1, \ldots, \sigma_r$. Thus, $A = U\Sigma V^H$.

Figure from "The Fundamental Theorem of Linear Algebra" by Gilbert Strang, The American Mathematical Monthly, Nov. 1993

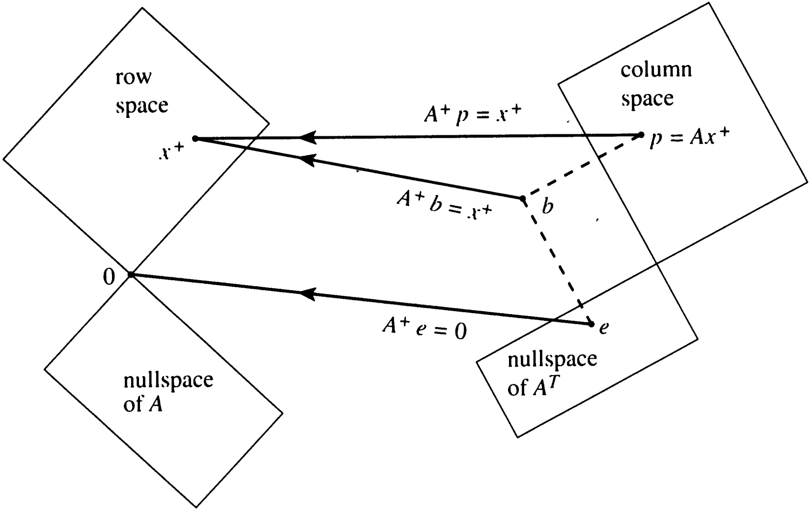

Figure 4. The inverse of A (where possible) is the pseudoinverse A^+.

9.2: Singular Value Decomposition Example

Consider the matrix

$$A = \begin{bmatrix} 1 & 1 \\ 5 & -1 \\ -1 & 5 \end{bmatrix}.$$

An eigenvalue decomposition of $A^H A$ is given by

$$A^H A = \begin{bmatrix} 27 & -9 \\ -9 & 27 \end{bmatrix} = V \Lambda V^H = \left(\frac{1}{\sqrt{2}} \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix} \right) \begin{bmatrix} 36 & 0 \\ 0 & 18 \end{bmatrix} \left(\frac{1}{\sqrt{2}} \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix} \right)$$

This implies $\Sigma_1 = \Lambda^{1/2}$ and $V_1 = V$. Thus, we find $U_1 = A V_1 \Sigma_1^{-1}$ with

$$U_1 = \begin{bmatrix} 1 & 1 \\ 5 & -1 \\ -1 & 5 \end{bmatrix} \left(\frac{1}{\sqrt{2}} \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix} \right) \begin{bmatrix} \frac{1}{\sqrt{36}} & 0 \\ 0 & \frac{1}{\sqrt{18}} \end{bmatrix} = \begin{bmatrix} 0 & \frac{1}{3} \\ \frac{1}{\sqrt{2}} & \frac{2}{3} \\ -\frac{1}{\sqrt{2}} & \frac{2}{3} \end{bmatrix}$$

Putting this all together, we have the compressed SVD

$$A = U_1 \Sigma_1 V_1^H = \begin{bmatrix} 0 & \frac{1}{3} \\ \frac{1}{\sqrt{2}} & \frac{2}{3} \\ -\frac{1}{\sqrt{2}} & \frac{2}{3} \end{bmatrix} \begin{bmatrix} \sqrt{36} & 0 \\ 0 & \sqrt{18} \end{bmatrix} \left(\frac{1}{\sqrt{2}} \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix} \right).$$

Moore–Penrose Pseudo Inverse

For a matrix $A \in \mathbb{C}^{m\times n}$, the matrix $A^+ \in \mathbb{C}^{n\times m}$ is the pseudo-inverse iff:

1. $AA^+A = A$ (implies AA^+ is idempotent)
2. $A^+AA^+ = A^+$ (implies A^+A is idempotent)
3. $(AA^+)^H = AA^+$ (implies AA^+ is Hermitian)
4. $(A^+A)^H = A^+A$ (implies A^+A is Hermitian)

Lemma

From the compact SVD $A = U_1\Sigma_1 V_1^H$, one finds that $A^+ = V_1\Sigma_1^{-1}U_1^H$.

Proof.

1. $AA^+A = U_1(\Sigma_1 V_1^H V_1 \Sigma_1^{-1} U_1^H U_1)\Sigma_1 V_1^H = A$
2. $A^+AA^+ = V_1(\Sigma_1^{-1} U_1^H U_1 \Sigma_1 V_1^H V_1)\Sigma_1^{-1} U_1^H = A^+$
3. $(AA^+)^H = \left(U_1(\Sigma_1 V_1^H V_1 \Sigma_1^{-1})U_1^H\right)^H = U_1 U_1^H = AA^+$
4. $(A^+A)^H = \left(V_1(\Sigma_1^{-1} U_1^H U_1 \Sigma_1)V_1^H\right)^H = V_1 V_1^H = A^+A$ □

Thus, AA^+ and A^+A are projection matrices onto $\mathcal{R}(A)$ and $\mathcal{R}(A^H)$.

Approximation Property of the SVD

Definition

Let $A \in \mathbb{C}^{m \times n}$ have SVD $A = U \Sigma V^H$ where $\underline{u}_1, \ldots, \underline{u}_m$ and $\underline{v}_1, \ldots, \underline{v}_n$ are the columns of U and V. Then, the *k*-truncated SVD expansion of *A* is

$$\mathcal{T}_k(A) \triangleq \sum_{i=1}^{k} \sigma_i \, \underline{u}_i \, \underline{v}_i^H$$

Theorem

In terms of the Frobenius norm $\|A\|_F^2 \triangleq \sum_{ij} |a_{ij}|^2$, *the best rank-k approximation of* $A \in \mathbb{C}^{m \times n}$ *is given by the k-truncated SVD expansion:*

$$\min_{B \in \mathbb{C}^{m \times n} :\, rank(B) = k} \|A - B\|_F = \|A - \mathcal{T}_k(A)\|_F$$

The Frobenius norm is induced by inner product $\langle A, B \rangle \triangleq \mathrm{Tr}(B^H A)$ on $\mathbb{C}^{m \times n}$

Principal Component Analysis (PCA)

Problem: For a given set of N data points $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_N \in \mathbb{R}^n$, what p-dimensional affine subspace $W \subset \mathbb{R}^n$ minimizes the approximation error

$$\sum_{i=1}^{N} \|\underline{x}_i - P_W(\underline{x}_i)\|^2.$$

Solution: Using $\underline{w}_0 = \frac{1}{N}\sum_{i=1}^{N} \underline{x}_i$, we define the mean-corrected data matrix

$$A = [\underline{x}_1 - \underline{w}_0, \ldots, \underline{x}_N - \underline{w}_0].$$

Then, the problem can solved using the SVD $A = U\Sigma V^T$. In particular, one can define $U_p \triangleq [\underline{u}_1, \ldots, \underline{u}_p]$ and choose $W = \mathrm{span}(U_p) + \underline{w}_0$ so that

$$P_W(\underline{x}) = U_p U_p^T(\underline{x} - \underline{w}_0) + \underline{w}_0.$$

For dimension reduction, one stores $\underline{y}_i = U_p^T \underline{x}_i \in \mathbb{R}^p$ instead of $\underline{x}_i \in \mathbb{R}^n$.

Note: This solution essentially replaces A by the p-truncated SVD $\mathcal{T}_p(A)$

Next Steps

- To continue studying after this video –
  - Try the required reading from website:

    The Fundamental Theorem of Linear Algebra by Gilbert Strang
  - Or the recommended reading: Course Notes EF 6.5 - 6.6.1, 9.1 - 9.3
  - Also, look at the problems in Assignments 7 and 8