Week 15: Tuesday. 6.3: Singular Value Decomposition

Course Goals

After studying section 6.3: Singular Value Decomposition, you should

1. *Understand the singular value decomposition (SVD).*
2. *Understand the applicability of the SVD to other fields.*

Course Outcomes

To manifest that you have reached the above course goals, you should be able to

low rank approximation

1. *Be able to calculate the SVD of a matrix.*
2. *Understand how to relate the SVD to the pseudoinverse of a matrix.*
3. *Be able to identify bases/dimensions of the 4 fundamental subspaces from the SVD of a matrix.*

Week 15: Tuesday. 6.3: Singular Value Decomposition

Singular Value Decomposition

Any $m \times n$ matrix A can be factored into:

$$\underbrace{A}_{m\times n} = \underbrace{U}_{m\times m}\overbrace{\Sigma}^{m\times n}\underbrace{V^T}_{n\times n} = (\text{orthogonal})(\text{diagonal})(\text{orthogonal}).$$

- The columns of U $(m \times m)$ are the eigenvectors of AA^T.
- The columns of V $(n \times n)$ are the eigenvectors of $\cancel{AA^T}$. A^TA
- The r singular values on the diagonal of Σ $(m \times n)$ are the square-roots of the non-zero eigenvalues of both AA^T and A^TA.
- We order the singular values $\sigma_i \geq \sigma_{i+1}$.

Week 15: Tuesday. 6.3: Singular Value Decomposition

SVD and Fundamental Subspaces

U and V are orthonormal matrices and their columns are orthonormal bases for **all four fundamental subspaces**.

- First r columns of U are a basis for $C(A)$
- Last $(m - r)$ columns of U are a basis for $N(A^T)$
- First r columns of V are a basis for $C(A^T)$
- Last $(n - r)$ columns of V are a basis for $N(A)$.

$$A = U\Sigma V^T \implies \boxed{AV = U\Sigma}$$

Thus, when we map a column of V, $\vec{v}_j$ we end up with σ_j times the same column of U:

$$A\vec{v}_j = \sigma_j \vec{u}_j \qquad j = 1, \dots, r$$

Week 14: Thursday. 6.3: Singular Value Decomposition

- How do we know the singular value decomposition exists?
- How do we actually get the singular value decomposition for a given matrix?
- We will look at a brief **constructive proof** where we show how to build it.
- In practice, we use computers to calculate the SVD - because most of the cases we care about are cases where the matrix in question is quite large.

Week 15: Tuesday. 6.3: Singular Value Decomposition

SVD and Eigenvectors/Eigenvalues of A^TA and AA^T

- If $A = U\Sigma V^T$. Then

$$AA^T = (U\Sigma V^T)(U\Sigma V^T)^T = U\Sigma V^T V \Sigma^T U^T = U\boxed{\Sigma\Sigma^T}U^T$$

$$A^TA = (U\Sigma V^T)^T(U\Sigma V^T) = V\Sigma^T U^T U \Sigma V^T = V\boxed{\Sigma^T\Sigma}V^T$$

$\Sigma\Sigma^T$ and $\Sigma^T\Sigma$ are both diagonal matrices and let's just call them Σ^2

- But this means:

$$\underbrace{(AA^T)U = U\Sigma^2}_{\text{eigenvectors}} \text{ and } (A^TA)V = V\Sigma^2$$

- And this implies:

$$(AA^T)\vec{u}_i = \sigma_i^2\vec{u}_i \text{ and } (A^TA)\vec{v}_i = \sigma_i^2\vec{v}_i.$$

- Thus U and V are the eigenvectors of AA^T and A^TA respectively and the diagonal entries in Σ^2 are the eigenvalues.

Week 15: Tuesday. 6.3: Singular Value Decomposition

Positive Semi-Definite Matrix

Let A be an $m \times n$ matrix, then the matrix $A^T A$ is positive semi-definite.

- We will show that all eigenvalues are positive. Let λ and $\vec{x}$ be an eigenvalue and eigenvector of $A^T A$.

$$A^T A\vec{x} = \lambda\vec{x}.$$

- We know (because $A^T A$ is symmetric) that λ is real.
- Let's calculate $\|A\vec{x}\|^2$:

$$\|A\vec{x}\|^2 = (A\vec{x})^T(A\vec{x}) = \vec{x}^T A^T A\vec{x} = \lambda\vec{x}^T\vec{x} = \lambda\|\vec{x}\|^2.$$

Thus, $\boxed{\lambda = \|A\vec{x}\|^2/\|\vec{x}\| \geq 0} \implies \lambda \geq 0.$

- Thus $A^T A$ is positive semi-definite.

Rank of $A^T A$

Let A be an $m \times n$ matrix with rank r, then $A^T A$, which is an $n \times n$ matrix also has rank r.

- We know that $N(A) = n - \text{rank}(A)$ and $N(A^T A) = n - \text{rank}(A^T A)$.
- We will show that $N(A) = N(A^T A)$.
- $\vec{x} \in N(A) \implies \vec{x} \in N(A^T A)$.

$$\begin{aligned} A\vec{x} &= \vec{0} \\ A^T A\vec{x} &= A^T\vec{0} \\ A^T A\vec{x} &= \vec{0}. \end{aligned}$$

Thus, $\vec{x} \in N(A^T A)$.

- $\vec{x} \in N(A^T A) \implies \vec{x} \in N(A)$

$$\begin{aligned} A^T A\vec{x} &= \vec{0} \\ \vec{x}^T A^T A\vec{x} &= \vec{x}^T\vec{0} \\ \|A\vec{x}\|^2 &= 0 \end{aligned}$$

Since $\|A\vec{x}\| = 0 \implies A\vec{x} = 0$, we have $x \in N(A)$.

Non-Zero Eigenvalues and Eigenvectors A^TA

Let A be an $m \times n$ matrix with rank r, then A^TA, which we know an $n \times n$ positive semi-definite matrix with rank r has exactly r non-zero eigenvalues.

- If $\vec{x} \in N(A^TA)$ and $\vec{x} \neq \vec{0}$, then $\vec{x}$ is an eigenvector of eigenvalue 0.
$$A^TA\vec{x} = \vec{0} = 0\vec{x}.$$
- Since we know A^TA has an orthonormal set of eigenvectors (i.e., it's symmetric) we know that there will be $(n-r)$ eigenvectors corresponding to the eigenvalue of 0 with multiplicity $(n-r)$.
- Thus, the remaining r eigenvalues (with multiplicity) are positive and correspond to r eigenvectors.

A similar argument gives us the rank of A^T is equal to the rank of AA^T. And since the rank of A^T is equal to the rank of A so *the rank of all 4 matrices is the same!*

Proof of the Singular Value Decomposition

- Let A be an $m \times n$ matrix. Then $A^T A$ is a symmetric $n \times n$, matrix.
- Therefore it's eigenvalues are real and it has an orthonormal set of eigenvectors. Let's put these into the columns of an orthogonal matrix V.
- We list eigenvalues (and eigenvectors) in decreasing order:

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r > 0 = \lambda_{r+1} = \cdots = \lambda_n.$$

$r = \text{rank}$

$n-r$ 0 eigenvalues

That is, $\vec{v}_i$ corresponds to value λ_i.

- The **singular values** of A are given by:

$$\sigma_j = \sqrt{\lambda_j} \text{ for } j = 1, 2, \ldots n.$$

Then,

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0 = \sigma_{r+1} = \cdots = \sigma_n.$$

rank

Proof of the Singular Value Decomposition

- We will split the eigenvectors into two groups, the ones that correspond to $\sigma_i > 0$ and the ones where $\sigma_i = 0$.

$\lambda = 0$

$$V_1 = \underbrace{\begin{bmatrix} \vec{v}_1 & \vec{v}_2 & \dots & \vec{v}_r \end{bmatrix}} \text{ and } V_2 = \underbrace{\begin{bmatrix} \vec{v}_{r+1} & \vec{v}_{r+2} & \dots & \vec{v}_n \end{bmatrix}}.$$

- We will let Σ_1 be the $r \times r$ diagonal matrix of positive singular values:

$A = U\Sigma V^T$

$$\Sigma_1 = \begin{bmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_r \end{bmatrix}.$$

$r \times r$

- The full Σ is an $m \times n$ matrix padded by zero matrices: $\mathbf{O}$:

$$\Sigma = \begin{bmatrix} \Sigma_{1_{(r\times r)}} & \mathbf{O}_{r\times(n-r)} \\ \mathbf{O}_{(m-r)\times r} & \mathbf{O}_{(m-r)\times(n-r)} \end{bmatrix}.$$

$m \times n$

Proof of the Singular Value Decomposition

- We know that the rank of AA^T is the same, and is also positive semi-definite. Let's choose the eigenvectors for AA^T and split them the same way for U:

$N(AA^T)$

$$U_1 = \begin{bmatrix} \vec{u}_1 & \vec{u}_2 & \dots & \vec{u}_r \end{bmatrix} \text{ and } U_2 = \begin{bmatrix} \vec{u}_{r+1} & \dots & \vec{u}_m \end{bmatrix}$$

(The red is because we are not choosing a specific eigenvector for the columns in U_1 just yet.)

$A = U\Sigma V^T V$

$AV = U\Sigma$

- Then the SVD comes down to

$m \times n$

$$\begin{aligned}
AV &= U\Sigma \\
A\begin{bmatrix} V_1 & V_2 \end{bmatrix} &= \begin{bmatrix} U_{1_{m\times r}} & U_{2_{m\times(m-r)}} \end{bmatrix} \begin{bmatrix} \Sigma_{1_{(r\times r)}} & \mathbf{O}_{r\times(n-r)} \\ \mathbf{O}_{(m-r)\times r} & \mathbf{O}_{(m-r)\times(n-r)} \end{bmatrix} \\
\begin{bmatrix} AV_1 & AV_2 \end{bmatrix} &= \begin{bmatrix} (U_1\Sigma_1)_{m\times r} + (U_2\mathbf{O})_{m\times r} & (U_1\mathbf{O})_{m\times(n-r)} + (U_2\mathbf{O})_{m\times(n-r)} \end{bmatrix} \\
\begin{bmatrix} AV_1 & AV_2 \end{bmatrix} &= \begin{bmatrix} (U_1\Sigma_1)_{m\times r} & \mathbf{O}_{m\times(n-r)} \end{bmatrix}
\end{aligned}$$

- But V_2 has all the eigenvectors for eigenvalue 0, the condition $AV_2 = \mathbf{O}_{m\times(n-r)}$ is automatically met!

$$m\left[\underbrace{AV_1}_{r} \;\middle|\; \underbrace{AV_2}_{n-r}\right] = m\left[\underbrace{U_1\Sigma_1}_{r} \quad \underbrace{0}_{n-r}\right]$$

$\Rightarrow V_2 \Rightarrow$ matrix of eigenvectors of A^TA corresponding to 0

$AV_2 = 0$ ✓

$$\Rightarrow \boxed{AV_1 = U_1 \Sigma_1}$$

$n \times r \quad m \times r \quad r \times r$

$$\Rightarrow \boxed{A \approx U_1 \Sigma_1 V_1^T}$$

$V_1V_1^T = I$ (padding

$A = U \Sigma V^T \leftarrow 10\times 10$

$$= \underbrace{\left[\vec{u}_1 \; \dots\dots \; \vec{u}_m\right]}_{10} \begin{bmatrix} \sigma_1 & & \\ & \sigma_2 & \\ & & \ddots \end{bmatrix} \begin{bmatrix} \vec{v}_1 \\ \vec{v}_2 \\ \vec{v}_3 \\ \vdots \\ \vec{v}_n \end{bmatrix} 10$$

$\Leftarrow \underline{m \times 1} \; 1 \times 1 \; \underline{1 \times n}$

Proof of the Singular Value Decomposition

- Thus, for the eigenvalue 0 for U_2 and V_2 we can order our $(m-r)$ and $(n-r)$ eigenvectors in any order we want!
- We picked V_1 first, so let's use it the columns of U_1.

$$A\vec{v}_j = \sigma_j \vec{u}_j \Longrightarrow \boxed{\vec{u}_j = \frac{1}{\sigma_j} A\vec{v}_j.}$$

columns

- Then it automatically follows that:

$$AV_1 = U_1\Sigma_1.$$

- We have orthonormal $\vec{u}_j$. Let $1 \leq i \leq r$ and $1 \leq j \leq r$.

$\lambda_j = \sigma_j^2$

$$\begin{aligned} \vec{u}_i^T \vec{u}_j &= \left(\frac{1}{\sigma_i}\vec{v}_i^T A^T\right)\left(\frac{1}{\sigma_j}A\vec{v}_j\right) = \frac{1}{\sigma_i\sigma_j}\vec{v}_i^T A^T A\vec{v}_j \\ &= \frac{\sigma_j^2}{\sigma_i\sigma_j}\vec{v}_i^T\vec{v}_j = \boxed{\frac{\sigma_j}{\sigma_i}\vec{v}_i^T\vec{v}_j} = \delta_{i,j}. \end{aligned}$$

- Also, by our earlier argument, U_1 consists of eigenvectors of AA^T. Thus, we have proven the SVD and how to build it!

Example

Find the singular value decomposition of:

$$A = \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix} \text{ and } B = \begin{bmatrix} 2 & -2 \\ 1 & 2 \end{bmatrix}.$$

- Find eigenvalues for $A^T A$ and eigenvectors. (This gives us V_1 and V_2).
- Find eigenvectors for the eigenvalue 0 for AA^T (i.e., a basis for the nullspace of AA^T).
- Enforce $\boxed{A\vec{v}_j = \sigma_j \vec{u}_j}$ for $j = 1, 2, \dots r$.

Week 15: Tuesday. 6.3: Singular Value Decomposition

$$A = \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix}, A^T A = \begin{bmatrix} 5 & 5 \\ 5 & 5 \end{bmatrix}, AA^T = \begin{bmatrix} 2 & 4 \\ 4 & 8 \end{bmatrix}.$$

- Find eigenvalues for $A^T A$:

$$0 = \det(A^T A - \lambda I) = \lambda^2 - 10\lambda = \lambda(\lambda - 10) \implies \lambda_1 = 10, \lambda_2 = 0.$$

- Find eigenvectors for $A^T A$:

$$\vec{0} = (A^T A - 10I)\vec{x}_1 \implies \left[\begin{array}{cc|c} -5 & 5 & 0 \\ 5 & -5 & 0 \end{array}\right] \implies \left[\begin{array}{cc|c} -5 & 5 & 0 \\ 0 & 0 & 0 \end{array}\right].$$

$V = [V_1 \quad V_2]$

$$\vec{x}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \implies \vec{v}_1 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}.$$

Week 15: Tuesday. 6.3: Singular Value Decomposition

$$A = \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix}, A^T A = \begin{bmatrix} 5 & 5 \\ 5 & 5 \end{bmatrix}, AA^T = \begin{bmatrix} 2 & 4 \\ 4 & 8 \end{bmatrix}.$$

- Find eigenvectors for $A^T A$:

$$\underline{\vec{0} = (A^T A - 0I)\vec{x}_2} \implies \left[\begin{array}{cc|c} 5 & 5 & 0 \\ 5 & 5 & 0 \end{array}\right] \implies \left[\begin{array}{cc|c} 5 & 5 & 0 \\ 0 & 0 & 0 \end{array}\right].$$

$$\vec{x}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix} \implies \vec{v}_2 = \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}. \checkmark$$

- Build V_1 and V_2:

r $\quad n-r$

$$V = \begin{bmatrix} V_1 & V_2 \end{bmatrix} = \begin{bmatrix} \vec{v}_1 & \vec{v}_2 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}. \checkmark$$

Week 15: Tuesday. 6.3: Singular Value Decomposition

$$A = \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix}, A^TA = \begin{bmatrix} 5 & 5 \\ 5 & 5 \end{bmatrix}, AA^T = \begin{bmatrix} 2 & 4 \\ 4 & 8 \end{bmatrix}.$$

- Find eigenvectors for eigenvalue 0 for AA^T:

$$\vec{0} = AA^T\vec{x} \implies \left[\begin{array}{cc|c} 2 & 4 & 0 \\ 4 & 8 & 0 \end{array}\right] = \left[\begin{array}{cc|c} 2 & 4 & 0 \\ 0 & 0 & 0 \end{array}\right].$$

$\|\vec{x}_2\| = \sqrt{2^2+1^2} = \sqrt{5}$

$$\vec{x}_2 = \begin{bmatrix} -2 & 1 \end{bmatrix} \implies \vec{u}_2 = \begin{bmatrix} -2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix},$$

- Enforce: $A\vec{v}_i = \sigma_i \vec{u}_i$ (Recall: $\sigma_i = \sqrt{\lambda_i}$)

$$\frac{1}{\sqrt{10}} A\vec{v}_1 = \vec{u}_1 \implies \frac{1}{\sqrt{10}} \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} = \frac{1}{\sqrt{10}} \begin{bmatrix} \sqrt{2} \\ 2\sqrt{2} \end{bmatrix} = \begin{bmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix}$$

Week 15: Tuesday. 6.3: Singular Value Decomposition

- Write down the V, U and Σ:

$\vec{u}_1$ $\quad$ $\vec{u}_2$

$$V = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}, U = \begin{bmatrix} 1/\sqrt{5} & -2/\sqrt{5} \\ 2/\sqrt{5} & 1/\sqrt{5} \end{bmatrix}, \Sigma = \begin{bmatrix} \sqrt{10} & 0 \\ 0 & 0 \end{bmatrix}.$$

$r \quad (n-r) \qquad r \quad (m-r)$

Week 15: Tuesday. 6.3: Singular Value Decomposition

$$B = \begin{bmatrix} 2 & -2 \\ 1 & 2 \end{bmatrix}, B^T B = \begin{bmatrix} 5 & -2 \\ -2 & 8 \end{bmatrix}, BB^T = \begin{bmatrix} 8 & -2 \\ -2 & 5 \end{bmatrix}$$

- Find eigenvalues of $B^T B$:

  $0 = \det(B^T B - \lambda I) = \lambda^2 - 13\lambda + 36 = (\lambda - 9)(\lambda - 4) \implies \lambda_1 = 9, \lambda_2 = 4.$

  Note: We have order λ_i from largest to smallest!
- Find eigenvectors of $B^T B$:

$$\left[\begin{array}{cc|c} 5-9 & -2 & 0 \\ -2 & 8-9 & 0 \end{array}\right] = \left[\begin{array}{cc|c} -4 & -2 & 0 \\ -2 & -1 & 0 \end{array}\right] = \left[\begin{array}{cc|c} -2 & -1 & 0 \\ 0 & 0 & 0 \end{array}\right]$$

$$\vec{x}_1 = \begin{bmatrix} 1 \\ -2 \end{bmatrix} \implies \vec{v}_1 = \begin{bmatrix} 1/\sqrt{5} \\ -2/\sqrt{5} \end{bmatrix}.$$

Week 15: Tuesday. 6.3: Singular Value Decomposition

$$B = \begin{bmatrix} 2 & -2 \\ 1 & 2 \end{bmatrix}, B^T B = \begin{bmatrix} 5 & -2 \\ -2 & 8 \end{bmatrix}, BB^T = \begin{bmatrix} 8 & -2 \\ -2 & 5 \end{bmatrix}$$

- Find eigenvectors of $B^T B$:

$$\left[\begin{array}{cc|c} 5-4 & -2 & 0 \\ -2 & 8-4 & 0 \end{array}\right] = \left[\begin{array}{cc|c} 1 & -2 & 0 \\ -2 & 4 & 0 \end{array}\right] = \left[\begin{array}{cc|c} 1 & -2 & 0 \\ 0 & 0 & 0 \end{array}\right]$$

$$\vec{x}_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \implies \vec{v}_2 = \begin{bmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix}.$$

- We do not have 0 as an eigenvalue, so we skip directly to:

$$B\vec{v}_j = \sigma_j \vec{u}_j$$

- $\sigma_1 = 3$:

$$\vec{u}_1 = \frac{1}{3}B\vec{v}_1 = \frac{1}{3}\begin{bmatrix} 2 & -2 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 1/\sqrt{5} \\ -2/\sqrt{5} \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 6/\sqrt{5} \\ -3/\sqrt{5} \end{bmatrix} = \begin{bmatrix} 2/\sqrt{5} \\ -1/\sqrt{5} \end{bmatrix}$$

- $\sigma_2 = 2$

$$\vec{u}_2 = \frac{1}{2}B\vec{v}_2\begin{bmatrix} 2 & -2 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 2/\sqrt{5} \\ 4/\sqrt{5} \end{bmatrix} = \begin{bmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix}.$$

- Complete the matrices:

$$U = \begin{bmatrix} 2/\sqrt{5} & 1/\sqrt{5} \\ -1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix}, \Sigma = \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix} V = \begin{bmatrix} 1/\sqrt{5} & 2/\sqrt{5} \\ -2/\sqrt{5} & 1/\sqrt{5} \end{bmatrix}.$$