Singular Vectors and Singular Value Decomposition

Last Lecture

- We defined singular values of an $m \times n$ matrix to behave like eigenvalues of a square matrix.

In This Lecture

- We will extend the similarity of singular values and eigenvalues by defining singular vectors and then use these to mimic orthogonal diagonalization for $m \times n$ matrices.

Singular Vectors and Singular Value Decomposition

Let A be an $m \times n$ matrix. Assume that $m \neq n$.

For an $n \times n$ matrix B we have $B\vec{v} = \lambda\vec{v}$, $\vec{v} \neq \vec{0}$. However, observe that we cannot have

$$A\vec{v} = \lambda\vec{v}$$

since $A\vec{v} \in \mathbb{R}^m$ and $\lambda\vec{v} \in \mathbb{R}^n$.

Hence, we see that we need to adjust this to

$$A\vec{v} = \sigma\vec{u}$$

where $\vec{v} \in \mathbb{R}^n$ and $\vec{u} \in \mathbb{R}^m$ with $\vec{v} \neq \vec{0}$, $\vec{u} \neq \vec{0}$.

By definition, for any non-zero singular value σ of A there is a vector $\vec{v} \neq \vec{0}$ such that

$$A^TA\vec{v} = \sigma^2\vec{v}$$

Thus, if we have $A\vec{v} = \sigma\vec{u}$, then we have

$$A^TA\vec{v} = A^T(\sigma\vec{u})$$

so

$$\sigma^2\vec{v} = \sigma A^T\vec{u}$$

Since σ is non-zero, we can divide by σ to get

$$A^T\vec{u} = \sigma\vec{v}$$

Moreover, by Theorem 10.6.1, if $\vec{v}$ is a unit eigenvector of A^TA, then we get that $\vec{u} = \frac{1}{\sigma}A\vec{v}$ is also a unit vector.

Therefore, for a non-zero singular value σ of A, we will want unit vectors $\vec{v}$ and $\vec{u}$ such that

$$A\vec{v} = \sigma\vec{u} \qquad \text{and} \qquad A^T\vec{u} = \sigma\vec{v}$$

However, our derivation does not work for $\sigma = 0$. In this case, we only need one of these conditions to be satisfied.

Singular Vectors and Singular Value Decomposition

Definition: Let A be an $m \times n$ matrix. If $\vec{v} \in \mathbb{R}^n$ and $\vec{u} \in \mathbb{R}^m$ are unit vectors and $\sigma \neq 0$ is a singular value of A such that

$$A\vec{v} = \sigma\vec{u} \qquad \text{and} \qquad A^T\vec{u} = \sigma\vec{v}$$

then we say that $\vec{u}$ is a left singular vector of A and $\vec{v}$ is a right singular vector of A.

Additionally, if $\vec{u}$ is a unit vector such that $A^T\vec{u} = \vec{0}$, then $\vec{u}$ is a left singular vector of A.
Similarly, if $\vec{v}$ is a unit vector such that $A\vec{v} = \vec{0}$, then $\vec{v}$ is a right singular vector of A.

Note: This definition of right singular vectors not only preserves our relationship of these vectors with the eigenvectors of A^TA, but we also get the corresponding result for the left singular vectors and the eigenvectors of AA^T.

Singular Vectors and Singular Value Decomposition

Theorem 10.6.5

Let A be an $m \times n$ matrix. A vector $\vec{v} \in \mathbb{R}^n$ is a right singular vector of A if and only if $\vec{v}$ is an eigenvector of A^TA. A vector $\vec{u} \in \mathbb{R}^m$ is a left singular vector of A if and only if $\vec{u}$ is an eigenvector of AA^T.

Mimicking Orthogonal Diagonalization

Let A be an $m \times n$ matrix.
As was the case with singular vectors, if $m \neq n$, then P^TAP is not defined for a square matrix P.
Thus, we want to find an $m \times m$ orthogonal matrix U and an $n \times n$ orthogonal matrix V such that $U^TAV = \Sigma$, where Σ is an $m \times n$ "diagonal" matrix.
For orthogonal diagonalization, the columns of the matrix P are an orthonormal basis for $\mathbb{R}^n$ of eigenvectors of the matrix. In this case, we will need the columns of U to be an orthonormal basis for $\mathbb{R}^m$ of left singular vectors of A and the columns of V to be an orthonormal basis for $\mathbb{R}^n$ of right singular vectors of A.

We now prove that such orthonormal bases always exist.

Singular Vectors and Singular Value Decomposition

Lemma 10.6.6

Let A be an $m \times n$ matrix with $\text{rank}(A) = r$. If $\{\vec{v}_1, \dots, \vec{v}_n\}$ is an orthonormal basis for $\mathbb{R}^n$ consisting of the eigenvectors of A^TA arranged so that the corresponding eigenvalues $\lambda_1, \dots, \lambda_n$ are arranged from greatest to least and $\sigma_1, \dots, \sigma_n$ are the singular values of A, then $\{\frac{1}{\sigma_1}A\vec{v}_1, \dots, \frac{1}{\sigma_r}A\vec{v}_r\}$ is an orthonormal basis for $\operatorname{Col} A$.

Proof

By Corollary 10.6.4, we have that A has r non-zero singular values.
So, $\sigma_1, \dots, \sigma_r$ are all non-zero and $\sigma_{r+1} = \cdots = \sigma_n = 0$.
Now observe that for $1 \leq i, j \leq r$, $i \neq j$ we have

$$(A\vec{v}_i) \cdot (A\vec{v}_j) = (A\vec{v}_i)^T A\vec{v}_j = \vec{v}_i^T A^T A\vec{v}_j = \vec{v}_i^T(\lambda_j\vec{v}_j) = \lambda_j(\vec{v}_i^T\vec{v}_j) = \lambda_j(\vec{v}_i \cdot \vec{v}_j) = 0$$

since $\{\vec{v}_1, \dots, \vec{v}_n\}$ is an orthonormal set.
Hence, $\{A\vec{v}_1, \dots, A\vec{v}_r\}$ is orthogonal and so $\{\frac{1}{\sigma_1}A\vec{v}_1, \dots, \frac{1}{\sigma_r}A\vec{v}_r\}$ is orthonormal by Theorem 10.6.1.
Moreover, we know that $\dim \operatorname{Col}(A) = r$, so $\{\frac{1}{\sigma_1}A\vec{v}_1, \dots, \frac{1}{\sigma_r}A\vec{v}_r\}$ is an orthonormal basis for $\operatorname{Col} A$.

Singular Vectors and Singular Value Decomposition

Theorem 10.6.7

If A is an $m \times n$ matrix with rank r, then there exists an orthonormal basis $\{\vec{v}_1, \dots, \vec{v}_n\}$ for $\mathbb{R}^n$ of right singular vectors of A and an orthonormal basis $\{\vec{u}_1, \dots, \vec{u}_m\}$ for $\mathbb{R}^m$ of left singular vectors of A.

Proof

By Theorem 10.6.5 all eigenvectors of A^TA are right singular vectors of A.
Since A^TA is symmetric, we can find an orthonormal basis $\{\vec{v}_1, \dots, \vec{v}_n\}$ for $\mathbb{R}^n$ of right singular vectors of A.
By Lemma 10.6.6, the left singular vectors $\vec{u}_i = \frac{1}{\sigma_i}A\vec{v}_i$, $1 \leq i \leq r$, form an orthonormal basis for $\operatorname{Col}(A)$.
Also, by definition, a left singular vector $\vec{u}_j$ of A corresponding to $\sigma = 0$ lies in the nullspace of A^T.
Hence, by the Fundamental Theorem of Linear Algebra, if $\{\vec{u}_{r+1}, \dots, \vec{u}_m\}$ is an orthonormal basis for the nullspace of A^T, then $\{\vec{u}_1, \dots, \vec{u}_m\}$ is an orthonormal basis for $\mathbb{R}^m$ of left singular vectors of A.

Singular Vectors and Singular Value Decomposition

Thus, for any $m \times n$ matrix A with rank r we have an orthonormal basis $\{\vec{v}_1, \dots, \vec{v}_n\}$ for $\mathbb{R}^n$ of right singular vectors of A corresponding to the n singular values $\sigma_1, \dots, \sigma_n$ of A and an orthonormal basis $\{\vec{u}_1, \dots, \vec{u}_m\}$ for $\mathbb{R}^m$ of left singular vectors of A such that

$$\begin{aligned} A\begin{bmatrix} \vec{v}_1 & \cdots & \vec{v}_n \end{bmatrix} &= \begin{bmatrix} A\vec{v}_1 & \cdots & A\vec{v}_n \end{bmatrix} \\ &= \begin{bmatrix} \sigma_1\vec{u}_1 & \cdots & \sigma_n\vec{u}_n \end{bmatrix} \\ &= \begin{bmatrix} \sigma_1\vec{u}_1 & \cdots & \sigma_r\vec{u}_r & \vec{0} & \cdots & \vec{0} \end{bmatrix} \\ &= \begin{bmatrix} \vec{u}_1 & \cdots & \vec{u}_m \end{bmatrix}\Sigma \end{aligned}$$

where Σ is the $m \times n$ matrix with $\Sigma_{ii} = \sigma_i$ for $1 \leq i \leq r$ and all other entries of Σ are 0.

Hence, we have that there exists an orthogonal matrix V and orthogonal matrix U such that $AV = U\Sigma$.

Instead of writing this as $U^T AV = \Sigma$, we typically write this as $A = U\Sigma V^T$ to get a matrix decomposition of A.

Singular Vectors and Singular Value Decomposition

Definition: A singular value decomposition of an $m \times n$ matrix A is a factorization of the form

$$A = U\Sigma V^T$$

where U is an orthogonal matrix containing left singular vectors of A, V is an orthogonal matrix containing right singular vectors of A, and Σ is the $m \times n$ matrix with $\Sigma_{ii} = \sigma_i$ for $1 \leq i \leq \operatorname{rank} A$ and all other entries of Σ are 0.

To find a singular value decomposition of a matrix A with $\operatorname{rank} A = r$, we follow what we did previously.

1. Find the eigenvalues $\lambda_1, \dots, \lambda_n$ of A^TA arranged from greatest to least and a corresponding set of orthonormal eigenvectors $\{\vec{v}_1, \dots, \vec{v}_n\}$.
2. Let $V = \begin{bmatrix} \vec{v}_1 & \cdots & \vec{v}_n \end{bmatrix}$ and let Σ be the $m \times n$ matrix whose first r diagonal entries are the non-zero singular values $\sigma_1, \dots, \sigma_r$ of A arranged from greatest to least.
3. Find left singular vectors of A by computing $\vec{u}_i = \frac{1}{\sigma_i} A\vec{v}_i$ for $1 \leq i \leq r$. Then extend the set $\{\vec{u}_1, \dots, \vec{u}_r\}$ to an orthonormal basis for $\mathbb{R}^m$ and let $U = \begin{bmatrix} \vec{u}_1 & \cdots & \vec{u}_m \end{bmatrix}$.

Then, $A = U\Sigma V^T$.

Singular Vectors and Singular Value Decomposition

Example

Find a singular value decomposition of $A = \begin{bmatrix} 1 & -1 \\ -2 & 2 \\ 2 & -2 \end{bmatrix}$.

Solution

We have $A^T A = \begin{bmatrix} 9 & -9 \\ -9 & 9 \end{bmatrix}$ which has eigenvalues $\lambda_1 = 18$ and $\lambda_2 = 0$.

We have that the corresponding orthonormal eigenvectors are $\vec{v}_1 = \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}$ and $\vec{v}_2 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$.

Then, $V = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$.

The singular values of A are $\sigma_1 = \sqrt{18}$, and $\sigma_2 = 0$, so we have $\Sigma = \begin{bmatrix} \sqrt{18} & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$.

We compute

$$\vec{u}_1 = \frac{1}{\sigma_1} A\vec{v}_1 = \begin{bmatrix} 1/3 \\ -2/3 \\ 2/3 \end{bmatrix}$$

and from our work above $\{\vec{u}_1\}$ is an orthonormal basis for $\operatorname{Col}(A)$.

To find $\vec{u}_2$ and $\vec{u}_3$, we find an orthonormal basis for $\operatorname{Null}(A^T)$.

Singular Vectors and Singular Value Decomposition

Example

Find a singular value decomposition of $A = \begin{bmatrix} 1 & -1 \\ -2 & 2 \\ 2 & -2 \end{bmatrix}$.

Solution

Row reducing A^T, we get

$$\begin{bmatrix} 1 & -2 & 2 \\ -1 & 2 & -2 \end{bmatrix} \sim \begin{bmatrix} 1 & -2 & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

Hence, a basis for $\operatorname{Null}(A^T)$ is $\left\{ \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix} \right\}$.

But, we need an orthonormal basis for $\operatorname{Null}(A^T)$, so we apply the Gram-Schmidt procedure. We get

$$\vec{u}_2 = \begin{bmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \\ 0 \end{bmatrix}, \text{ and } \vec{u}_3 = \begin{bmatrix} -2/\sqrt{45} \\ 4/\sqrt{45} \\ 5/\sqrt{45} \end{bmatrix}$$

Thus, we take $U = \begin{bmatrix} \vec{u}_1 & \vec{u}_2 & \vec{u}_3 \end{bmatrix}$ and we have singular value decomposition $A = U\Sigma V^T$.

Singular Vectors and Singular Value Decomposition

Example

Find a singular value decomposition of $B = \begin{bmatrix} 2 & -4 \\ 2 & 2 \\ -4 & 0 \\ 1 & 4 \end{bmatrix}$.

Solution

We have $B^TB = \begin{bmatrix} 25 & 0 \\ 0 & 36 \end{bmatrix}$. Thus, the eigenvalues are $\lambda_1 = 36$ and $\lambda_2 = 25$ with corresponding unit eigenvectors $\vec{v}_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ and $\vec{v}_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

So, we take $V = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. Now, the singular values of B are $\sigma_1 = 6$ and $\sigma_2 = 5$ so, we get $\Sigma = \begin{bmatrix} 6 & 0 \\ 0 & 5 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$.

We compute

$$\vec{u}_1 = \frac{1}{\sigma_1} B\vec{v}_1 = \begin{bmatrix} -2/3 \\ 1/3 \\ 0 \\ 2/3 \end{bmatrix}, \qquad \vec{u}_2 = \frac{1}{\sigma_2} B\vec{v}_2 = \begin{bmatrix} 2/5 \\ 2/5 \\ -4/5 \\ 1/5 \end{bmatrix}$$

Next, we extend $\{\vec{u}_1, \vec{u}_2\}$ to an orthonormal basis for $\mathbb{R}^4$ by adding on a basis for the nullspace of B^T.

Singular Vectors and Singular Value Decomposition

Example

Find a singular value decomposition of $B = \begin{bmatrix} 2 & -4 \\ 2 & 2 \\ -4 & 0 \\ 1 & 4 \end{bmatrix}$.

Solution

Applying the Gram-Schmidt procedure to a basis for $\mathrm{Null}(B^T)$ we get vectors

$$\vec{u}_3 = \begin{bmatrix} 1/3 \\ -2/3 \\ 0 \\ 2/3 \end{bmatrix}, \quad \vec{u}_4 = \begin{bmatrix} 8/15 \\ 8/15 \\ 3/5 \\ 4/15 \end{bmatrix}$$

We let $U = \begin{bmatrix} \vec{u}_1 & \vec{u}_2 & \vec{u}_3 & \vec{u}_4 \end{bmatrix}$ and then we have singular value decomposition $U\Sigma V^T$ of B.