Image Compression and the Singular Value Decomposition

Brydon Eastman

REDEEMER UNIVERSITY COLLEGE

MAT-436 Linear Algebra

Redeemer University College, April 2, 2014

http://linalg.beastman.ca

http://cs.redeemer.ca/∼beastman/linalg

Intro

- How do computers store images?

Intro

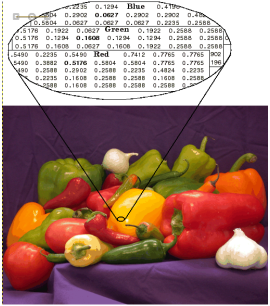
- How do computers store images?

A picture is really a matrix of colours

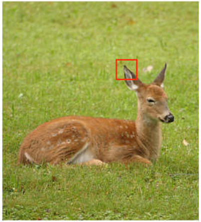
Intro

- How do computers store images?

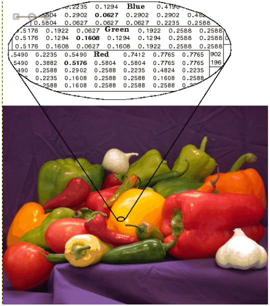
A picture is really a matrix of colours



- How do computers store images?



A picture is really a matrix of colours



The Problem

- A typical desktop background is 1920 pixels wide and 1080 pixels tall
- So we'd have to store 1920×1080 numbers for the red values, 1920×1080 numbers for the green values, and 1920×1080 numbers for the blue values.

The Problem

- A typical desktop background is 1920 pixels wide and 1080 pixels tall
- So we'd have to store 1920×1080 numbers for the red values, 1920×1080 numbers for the green values, and 1920×1080 numbers for the blue values.
- That amounts to storing $6,220,800$ different numbers

The Problem

- A typical desktop background is 1920 pixels wide and 1080 pixels tall
- So we'd have to store 1920×1080 numbers for the red values, 1920×1080 numbers for the green values, and 1920×1080 numbers for the blue values.
- That amounts to storing $6,220,800$ different numbers
- Computers are good at storing numbers like this, but that's just inefficient

The Problem

- A typical desktop background is 1920 pixels wide and 1080 pixels tall
- So we'd have to store 1920×1080 numbers for the red values, 1920×1080 numbers for the green values, and 1920×1080 numbers for the blue values.
- That amounts to storing $6,220,800$ different numbers
- Computers are good at storing numbers like this, but that's just inefficient

Why We Need Image Compression

The Problem

This is a picture I uploaded to Facebook last summer.
The camera I used stored the $35,426,304$ numbers ($\approx$ 65MB)
On Facebook it was only 12 MB. What happened?

The Math

Diagonalisable

A matrix A is diagonalisable if we can find an invertible matrix P such that $A = PDP^{-1}$ where D is a diagonal matrix.

Symmetric

A matrix A is symmetric if $A = A^{\intercal}$. Every symmetric matrix A can be written as $A = PDP^{\intercal}$ where P is an orthogonal matrix and D is a diagonal matrix.

The Math

Diagonalisable

A matrix A is diagonalisable if we can find an invertible matrix P such that $A = PDP^{-1}$ where D is a diagonal matrix.

Symmetric

A matrix A is symmetric if $A = A^\intercal$. Every symmetric matrix A can be written as $A = PDP^\intercal$ where P is an orthogonal matrix and D is a diagonal matrix.

- In both of these cases A had to be very special

The Math

Diagonalisable

A matrix A is diagonalisable if we can find an invertible matrix P such that $A = PDP^{-1}$ where D is a diagonal matrix.

Symmetric

A matrix A is symmetric if $A = A^{\intercal}$. Every symmetric matrix A can be written as $A = PDP^{\intercal}$ where P is an orthogonal matrix and D is a diagonal matrix.

- In both of these cases A had to be very special
- Today we will see that *every* matrix A can be written as $A = PDQ^{\intercal}$. Where P and Q are orthogonal.

The Math

Diagonalisable

A matrix A is diagonalisable if we can find an invertible matrix P such that $A = PDP^{-1}$ where D is a diagonal matrix.

Symmetric

A matrix A is symmetric if $A = A^\intercal$. Every symmetric matrix A can be written as $A = PDP^\intercal$ where P is an orthogonal matrix and D is a diagonal matrix.

- In both of these cases A had to be very special
- Today we will see that *every* matrix A can be written as $A = PDQ^\intercal$. Where P and Q are orthogonal.
- This is called **the singular value decomposition**

The Math

Diagonalisable

A matrix A is diagonalisable if we can find an invertible matrix P such that $A = PDP^{-1}$ where D is a diagonal matrix.

Symmetric

A matrix A is symmetric if $A = A^{\intercal}$. Every symmetric matrix A can be written as $A = PDP^{\intercal}$ where P is an orthogonal matrix and D is a diagonal matrix.

- In both of these cases A had to be very special
- Today we will see that *every* matrix A can be written as $A = PDQ^{\intercal}$. Where P and Q are orthogonal.
- This is called **the singular value decomposition**

$A^\intercal A$

- Recall that $A^\intercal A$ is square, symmetric, and has real eigenvalues (Theorem 5.18, p. 412).
- These eigenvalues are all non-negative

$A^{\intercal}A$

- Recall that $A^{\intercal}A$ is square, symmetric, and has real eigenvalues (Theorem 5.18, p. 412).
- These eigenvalues are all non-negative

$A^\intercal A$

- Recall that $A^\intercal A$ is square, symmetric, and has real eigenvalues (Theorem 5.18, p. 412).
- These eigenvalues are all non-negative

Suppose λ is an eigenvalue of $A^\intercal A$ with eigenvector $\vec{v}$. Then

- Recall that $A^\intercal A$ is square, symmetric, and has real eigenvalues (Theorem 5.18, p. 412).
- These eigenvalues are all non-negative

Suppose λ is an eigenvalue of $A^\intercal A$ with eigenvector $\vec{v}$. Then

$$0 \leq ||A\vec{v}||^2$$

- Recall that $A^{\intercal}A$ is square, symmetric, and has real eigenvalues (Theorem 5.18, p. 412).
- These eigenvalues are all non-negative

Suppose λ is an eigenvalue of $A^{\intercal}A$ with eigenvector $\vec{v}$. Then

$$0 \leq ||A\vec{v}||^2$$

$A^\intercal A$

- Recall that $A^\intercal A$ is square, symmetric, and has real eigenvalues (Theorem 5.18, p. 412).
- These eigenvalues are all non-negative

Suppose λ is an eigenvalue of $A^\intercal A$ with eigenvector $\vec{v}$. Then

$$\begin{aligned} 0 &\leq ||A\vec{v}||^2 \\ &= (A\vec{v}) \cdot (A\vec{v}) \end{aligned}$$

- Recall that $A^{\intercal}A$ is square, symmetric, and has real eigenvalues (Theorem 5.18, p. 412).
- These eigenvalues are all non-negative

Suppose λ is an eigenvalue of $A^{\intercal}A$ with eigenvector $\vec{v}$. Then

$$
\begin{aligned}
0 &\leq ||A\vec{v}||^2 \\
&= (A\vec{v}) \cdot (A\vec{v}) \\
&= (A\vec{v})^{\intercal} A\vec{v}
\end{aligned}
$$

- Recall that $A^\intercal A$ is square, symmetric, and has real eigenvalues (Theorem 5.18, p. 412).
- These eigenvalues are all non-negative

Suppose λ is an eigenvalue of $A^\intercal A$ with eigenvector $\vec{v}$. Then

$$\begin{aligned} 0 &\leq ||A\vec{v}||^2 \\ &= (A\vec{v}) \cdot (A\vec{v}) \\ &= (A\vec{v})^\intercal A\vec{v} \\ &= \vec{v}^\intercal A^\intercal A\vec{v} \end{aligned}$$

$A^{\intercal}A$

- Recall that $A^{\intercal}A$ is square, symmetric, and has real eigenvalues (Theorem 5.18, p. 412).
- These eigenvalues are all non-negative

Suppose λ is an eigenvalue of $A^{\intercal}A$ with eigenvector $\vec{v}$. Then

$$\begin{aligned} 0 &\leq ||A\vec{v}||^2 \\ &= (A\vec{v}) \cdot (A\vec{v}) \\ &= (A\vec{v})^{\intercal} A\vec{v} \\ &= \vec{v}^{\intercal} A^{\intercal} A\vec{v} \\ &= \vec{v}^{\intercal} \lambda \vec{v} \end{aligned}$$

$A^\intercal A$

- Recall that $A^\intercal A$ is square, symmetric, and has real eigenvalues (Theorem 5.18, p. 412).
- These eigenvalues are all non-negative

Suppose λ is an eigenvalue of $A^\intercal A$ with eigenvector $\vec{v}$. Then

$$\begin{aligned} 0 &\leq ||A\vec{v}||^2 \\ &= (A\vec{v}) \cdot (A\vec{v}) \\ &= (A\vec{v})^\intercal A\vec{v} \\ &= \vec{v}^\intercal A^\intercal A\vec{v} \\ &= \vec{v}^\intercal \lambda\vec{v} \\ &= \lambda||\vec{v}||^2 \end{aligned}$$

$A^{\intercal}A$

- Recall that $A^{\intercal}A$ is square, symmetric, and has real eigenvalues (Theorem 5.18, p. 412).
- These eigenvalues are all non-negative

Suppose λ is an eigenvalue of $A^{\intercal}A$ with eigenvector $\vec{v}$. Then

$$\begin{aligned} 0 &\leq ||A\vec{v}||^2 \\ &= (A\vec{v}) \cdot (A\vec{v}) \\ &= (A\vec{v})^{\intercal} A\vec{v} \\ &= \vec{v}^{\intercal} A^{\intercal} A\vec{v} \\ &= \vec{v}^{\intercal} \lambda \vec{v} \\ &= \lambda ||\vec{v}||^2 \end{aligned}$$

Singular Values

If A is an $m \times n$ matrix, the singular values of A are the square roots of the eigenvalues of $A^\intercal A$ and are denoted $\sigma_1, \sigma_2, \ldots, \sigma_n$. By convention $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n$.

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

Singular Values

Singular Values

If A is an $m \times n$ matrix, the singular values of A are the square roots of the eigenvalues of $A^{\intercal}A$ and are denoted $\sigma_1, \sigma_2, \ldots, \sigma_n$. By convention $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n$.

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

Singular Values

Singular Values

If A is an $m \times n$ matrix, the singular values of A are the square roots of the eigenvalues of $A^\intercal A$ and are denoted $\sigma_1, \sigma_2, \ldots, \sigma_n$. By convention $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n$.

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \therefore A^\intercal = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

Singular Values

Singular Values

If A is an $m \times n$ matrix, the singular values of A are the square roots of the eigenvalues of $A^\intercal A$ and are denoted $\sigma_1, \sigma_2, \ldots, \sigma_n$. By convention $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n$.

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \therefore A^\intercal = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

$$A^\intercal A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

Singular Values

Singular Values

If A is an $m \times n$ matrix, the singular values of A are the square roots of the eigenvalues of $A^\intercal A$ and are denoted $\sigma_1, \sigma_2, \ldots, \sigma_n$. By convention $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n$.

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \therefore A^\intercal = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

$$A^\intercal A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

$$\therefore c(x) = x^2 - 4x + 3$$

Singular Values

If A is an $m \times n$ matrix, the singular values of A are the square roots of the eigenvalues of $A^\intercal A$ and are denoted $\sigma_1, \sigma_2, \ldots, \sigma_n$. By convention $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n$.

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \therefore A^\intercal = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

$$A^\intercal A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

$$\therefore c(x) = x^2 - 4x + 3 \quad \therefore \lambda_1 = 3, \lambda_2 = 1$$

Singular Values

If A is an $m \times n$ matrix, the singular values of A are the square roots of the eigenvalues of $A^{\intercal}A$ and are denoted $\sigma_1, \sigma_2, \ldots, \sigma_n$. By convention $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n$.

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \therefore A^{\intercal} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

$$A^{\intercal}A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

$$\therefore c(x) = x^2 - 4x + 3 \quad \therefore \lambda_1 = 3, \lambda_2 = 1$$

Singular Values

If A is an $m \times n$ matrix, the singular values of A are the square roots of the eigenvalues of $A^\intercal A$ and are denoted $\sigma_1, \sigma_2, \ldots, \sigma_n$. By convention $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n$.

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \therefore A^\intercal = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

$$A^\intercal A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

$$\therefore c(x) = x^2 - 4x + 3 \quad \therefore \lambda_1 = 3, \lambda_2 = 1$$

$$\therefore \sigma_1 = \sqrt{3}, \sigma_2 = 1$$

Singular Values

Singular Values

If A is an $m \times n$ matrix, the singular values of A are the square roots of the eigenvalues of $A^\intercal A$ and are denoted $\sigma_1, \sigma_2, \ldots, \sigma_n$. By convention $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n$.

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \therefore A^\intercal = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

$$A^\intercal A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

$$\therefore c(x) = x^2 - 4x + 3 \quad \therefore \lambda_1 = 3, \lambda_2 = 1$$

$$\therefore \sigma_1 = \sqrt{3}, \sigma_2 = 1$$

Some properties of Singular Values

- If A is an $m \times n$ matrix, then $A^{\intercal}$ is an $n \times m$ matrix.

Some properties of Singular Values

- If A is an $m \times n$ matrix, then $A^\intercal$ is an $n \times m$ matrix.
- Thus $A^\intercal A$ is an $n \times n$ matrix

Some properties of Singular Values

- If A is an $m \times n$ matrix, then $A^\intercal$ is an $n \times m$ matrix.
- Thus $A^\intercal A$ is an $n \times n$ matrix

Some properties of Singular Values

- If A is an $m \times n$ matrix, then $A^\intercal$ is an $n \times m$ matrix.
- Thus $A^\intercal A$ is an $n \times n$ matrix
- An $n \times n$ matrix has a degree n characteristic polynomial

Some properties of Singular Values

- If A is an $m \times n$ matrix, then $A^\intercal$ is an $n \times m$ matrix.
- Thus $A^\intercal A$ is an $n \times n$ matrix
- An $n \times n$ matrix has a degree n characteristic polynomial

Some properties of Singular Values

- If A is an $m \times n$ matrix, then $A^{\intercal}$ is an $n \times m$ matrix.
- Thus $A^{\intercal}A$ is an $n \times n$ matrix
- An $n \times n$ matrix has a degree n characteristic polynomial
- Thus (counting multiplicities) $A^{\intercal}A$ has n eigenvalues.

Some properties of Singular Values

- If A is an $m \times n$ matrix, then $A^\intercal$ is an $n \times m$ matrix.
- Thus $A^\intercal A$ is an $n \times n$ matrix
- An $n \times n$ matrix has a degree n characteristic polynomial
- Thus (counting multiplicities) $A^\intercal A$ has n eigenvalues.

Some properties of Singular Values

- If A is an $m \times n$ matrix, then $A^{\intercal}$ is an $n \times m$ matrix.
- Thus $A^{\intercal}A$ is an $n \times n$ matrix
- An $n \times n$ matrix has a degree n characteristic polynomial
- Thus (counting multiplicities) $A^{\intercal}A$ has n eigenvalues.
- Therefore A has n singular values (counting multiplicities).

Some properties of Singular Values

- If A is an $m \times n$ matrix, then $A^\intercal$ is an $n \times m$ matrix.
- Thus $A^\intercal A$ is an $n \times n$ matrix
- An $n \times n$ matrix has a degree n characteristic polynomial
- Thus (counting multiplicities) $A^\intercal A$ has n eigenvalues.
- Therefore A has n singular values (counting multiplicities).

Some properties of Singular Values

- If A is an $m \times n$ matrix, then $A^\intercal$ is an $n \times m$ matrix.
- Thus $A^\intercal A$ is an $n \times n$ matrix
- An $n \times n$ matrix has a degree n characteristic polynomial
- Thus (counting multiplicities) $A^\intercal A$ has n eigenvalues.
- Therefore A has n singular values (counting multiplicities).

Some properties of Singular Values

- If A is an $m \times n$ matrix, then $A^{\intercal}$ is an $n \times m$ matrix.
- Thus $A^{\intercal}A$ is an $n \times n$ matrix
- An $n \times n$ matrix has a degree n characteristic polynomial
- Thus (counting multiplicities) $A^{\intercal}A$ has n eigenvalues.
- Therefore A has n singular values (counting multiplicities).

Recall: If B is a $n \times n$ symmetric matrix, then there is an orthonormal basis for $\mathbb{R}^n$ that consists of eigenvectors of B. (Proof of the Spectral Theorem, p 414).

Some properties of Singular Values

- If A is an $m \times n$ matrix, then $A^{\intercal}$ is an $n \times m$ matrix.
- Thus $A^{\intercal}A$ is an $n \times n$ matrix
- An $n \times n$ matrix has a degree n characteristic polynomial
- Thus (counting multiplicities) $A^{\intercal}A$ has n eigenvalues.
- Therefore A has n singular values (counting multiplicities).

Recall: If B is a $n \times n$ symmetric matrix, then there is an orthonormal basis for $\mathbb{R}^n$ that consists of eigenvectors of B. (Proof of the Spectral Theorem, p 414).

Let $\{\vec{v}_1, \ldots, \vec{v}_n\}$ be such a basis containing eigenvectors of $A^{\intercal}A$ that correspond with eigenvalues λ_i.

Some properties of Singular Values

- If A is an $m \times n$ matrix, then $A^\intercal$ is an $n \times m$ matrix.
- Thus $A^\intercal A$ is an $n \times n$ matrix
- An $n \times n$ matrix has a degree n characteristic polynomial
- Thus (counting multiplicities) $A^\intercal A$ has n eigenvalues.
- Therefore A has n singular values (counting multiplicities).

Recall: If B is a $n \times n$ symmetric matrix, then there is an orthonormal basis for $\mathbb{R}^n$ that consists of eigenvectors of B. (Proof of the Spectral Theorem, p 414).

Let $\{\vec{v}_1, \ldots, \vec{v}_n\}$ be such a basis containing eigenvectors of $A^\intercal A$ that correspond with eigenvalues λ_i.

We already saw $||A\vec{v}_i||^2 = \lambda_i ||\vec{v}_i||^2$. But $\vec{v}_i$ is a unit vector, thus $||A\vec{v}_i|| = \sqrt{\lambda_i} = \sigma_i$.

Some properties of Singular Values

- If A is an $m \times n$ matrix, then $A^\intercal$ is an $n \times m$ matrix.
- Thus $A^\intercal A$ is an $n \times n$ matrix
- An $n \times n$ matrix has a degree n characteristic polynomial
- Thus (counting multiplicities) $A^\intercal A$ has n eigenvalues.
- Therefore A has n singular values (counting multiplicities).

Recall: If B is a $n \times n$ symmetric matrix, then there is an orthonormal basis for $\mathbb{R}^n$ that consists of eigenvectors of B. (Proof of the Spectral Theorem, p 414).

Let $\{\vec{v}_1, \dots, \vec{v}_n\}$ be such a basis containing eigenvectors of $A^\intercal A$ that correspond with eigenvalues λ_i.

We already saw $||A\vec{v}_i||^2 = \lambda_i ||\vec{v}_i||^2$. But $\vec{v}_i$ is a unit vector, thus $||A\vec{v}_i|| = \sqrt{\lambda_i} = \sigma_i$.

Singular Value Decomposition

- We are going to use these singular values to obtain the aforementioned decomposition.

Singular Value Decomposition

- We are going to use these singular values to obtain the aforementioned decomposition.
- Let A be an $m \times n$ matrix with singuar values $\sigma_1 \geq \cdots \geq \sigma_r > 0$ and $\sigma_{r+1} = \sigma_{r+2} = \cdots = \sigma_n = 0$.

Singular Value Decomposition

- We are going to use these singular values to obtain the aforementioned decomposition.
- Let A be an $m \times n$ matrix with singuar values $\sigma_1 \geq \cdots \geq \sigma_r > 0$ and $\sigma_{r+1} = \sigma_{r+2} = \cdots = \sigma_n = 0$.
- U be an $m \times m$ orthogonal matrix, and V be an $n \times n$ orthogonal matrix.

Singular Value Decomposition

- We are going to use these singular values to obtain the aforementioned decomposition.
- Let A be an $m \times n$ matrix with singuar values $\sigma_1 \geq \cdots \geq \sigma_r > 0$ and $\sigma_{r+1} = \sigma_{r+2} = \cdots = \sigma_n = 0$.
- U be an $m \times m$ orthogonal matrix, and V be an $n \times n$ orthogonal matrix.

- We are going to use these singular values to obtain the aforementioned decomposition.
- Let A be an $m \times n$ matrix with singuar values $\sigma_1 \geq \cdots \geq \sigma_r > 0$ and $\sigma_{r+1} = \sigma_{r+2} = \cdots = \sigma_n = 0$.
- U be an $m \times m$ orthogonal matrix, and V be an $n \times n$ orthogonal matrix.

Singular Value Decomposition

- We are going to use these singular values to obtain the aforementioned decomposition.
- Let A be an $m \times n$ matrix with singuar values $\sigma_1 \geq \cdots \geq \sigma_r > 0$ and $\sigma_{r+1} = \sigma_{r+2} = \cdots = \sigma_n = 0$.
- U be an $m \times m$ orthogonal matrix, and V be an $n \times n$ orthogonal matrix.

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

Singular Value Decomposition

- We are going to use these singular values to obtain the aforementioned decomposition.
- Let A be an $m \times n$ matrix with singuar values $\sigma_1 \geq \cdots \geq \sigma_r > 0$ and $\sigma_{r+1} = \sigma_{r+2} = \cdots = \sigma_n = 0$.
- U be an $m \times m$ orthogonal matrix, and V be an $n \times n$ orthogonal matrix.

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

Where $\Sigma = \left[\begin{array}{c|c} D & O \\ \hline O & O \end{array}\right]$ and $D = \begin{bmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_r \end{bmatrix}$

Singular Value Decomposition

- We are going to use these singular values to obtain the aforementioned decomposition.
- Let A be an $m \times n$ matrix with singuar values $\sigma_1 \geq \cdots \geq \sigma_r > 0$ and $\sigma_{r+1} = \sigma_{r+2} = \cdots = \sigma_n = 0$.
- U be an $m \times m$ orthogonal matrix, and V be an $n \times n$ orthogonal matrix.

$$A = U\Sigma V^{\intercal}$$

$$\text{Where } \Sigma = \left[\begin{array}{c|c} D & O \\ \hline O & O \end{array}\right] \text{ and } D = \begin{bmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_r \end{bmatrix}$$

Examples of Σ

Some examples of Σ that may show up if $r = 2$:

$$\begin{bmatrix} 4 & 0 & 0 \\ 0 & 3 & 0 \end{bmatrix} \quad \begin{bmatrix} 2 & 0 \\ 0 & 2 \\ 0 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 5 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \begin{bmatrix} 9 & 0 \\ 0 & 6 \end{bmatrix}$$

How do we do the SVD?

- We already know how to find Σ

How do we do the SVD?

- We already know how to find Σ
- All we have to do is find U and V. We'll find V first.

How do we do the SVD?

- We already know how to find Σ
- All we have to do is find U and V. We'll find V first.

How do we do the SVD?

- We already know how to find Σ
- All we have to do is find U and V. We'll find V first.

How do we do the SVD?

- We already know how to find Σ
- All we have to do is find U and V. We'll find V first.

We know there are eigenvectors $\{\vec{v}_1, \ldots, \vec{v}_n\}$ of $A^\intercal A$ that form an orthonormal basis for $\mathbb{R}^n$.

How do we do the SVD?

- We already know how to find Σ
- All we have to do is find U and V. We'll find V first.

We know there are eigenvectors $\{\vec{v}_1, \ldots, \vec{v}_n\}$ of $A^\intercal A$ that form an orthonormal basis for $\mathbb{R}^n$.

Then $V = \begin{bmatrix} \vec{v}_1 & \cdots & \vec{v}_n \end{bmatrix}$ is an orthogonal $n \times n$ matrix.

How do we do the SVD?

- We already know how to find Σ
- All we have to do is find U and V. We'll find V first.

We know there are eigenvectors $\{\vec{v}_1, \ldots, \vec{v}_n\}$ of $A^\intercal A$ that form an orthonormal basis for $\mathbb{R}^n$.

Then $V = \begin{bmatrix} \vec{v}_1 & \cdots & \vec{v}_n \end{bmatrix}$ is an orthogonal $n \times n$ matrix.

Consider $\{A\vec{v}_1, \ldots, A\vec{v}_n\}$. I claim this is an orthogonal set of vectors. To see this suppose $\vec{v}_j$ is an eigenvector of $A^\intercal A$ with corresponding eigenvalue λ_j. For $i \neq j$ we have.

$$(A\vec{v}_i) \cdot (A\vec{v}_j) = (A\vec{v}_i)^\intercal (A\vec{v}_j)$$

How do we do the SVD?

- We already know how to find Σ
- All we have to do is find U and V. We'll find V first.

We know there are eigenvectors $\{\vec{v}_1, \ldots, \vec{v}_n\}$ of $A^\intercal A$ that form an orthonormal basis for $\mathbb{R}^n$.

Then $V = \begin{bmatrix} \vec{v}_1 & \cdots & \vec{v}_n \end{bmatrix}$ is an orthogonal $n \times n$ matrix.

Consider $\{A\vec{v}_1, \ldots, A\vec{v}_n\}$. I claim this is an orthogonal set of vectors. To see this suppose $\vec{v}_j$ is an eigenvector of $A^\intercal A$ with corresponding eigenvalue λ_j. For $i \neq j$ we have.

$$(A\vec{v}_i) \cdot (A\vec{v}_j) = (A\vec{v}_i)^\intercal (A\vec{v}_j)$$

How do we do the SVD?

- We already know how to find Σ
- All we have to do is find U and V. We'll find V first.

We know there are eigenvectors $\{\vec{v}_1, \ldots, \vec{v}_n\}$ of $A^\intercal A$ that form an orthonormal basis for $\mathbb{R}^n$.

Then $V = \begin{bmatrix} \vec{v}_1 & \cdots & \vec{v}_n \end{bmatrix}$ is an orthogonal $n \times n$ matrix.

Consider $\{A\vec{v}_1, \ldots, A\vec{v}_n\}$. I claim this is an orthogonal set of vectors. To see this suppose $\vec{v}_j$ is an eigenvector of $A^\intercal A$ with corresponding eigenvalue λ_j. For $i \neq j$ we have.

$$\begin{aligned} (A\vec{v}_i) \cdot (A\vec{v}_j) &= (A\vec{v}_i)^\intercal (A\vec{v}_j) \\ &= \vec{v}_i^\intercal A^\intercal A\vec{v}_j \end{aligned}$$

How do we do the SVD?

- We already know how to find Σ
- All we have to do is find U and V. We'll find V first.

We know there are eigenvectors $\{\vec{v}_1, \ldots, \vec{v}_n\}$ of $A^\intercal A$ that form an orthonormal basis for $\mathbb{R}^n$.

Then $V = \begin{bmatrix} \vec{v}_1 & \cdots & \vec{v}_n \end{bmatrix}$ is an orthogonal $n \times n$ matrix.

Consider $\{A\vec{v}_1, \ldots, A\vec{v}_n\}$. I claim this is an orthogonal set of vectors. To see this suppose $\vec{v}_j$ is an eigenvector of $A^\intercal A$ with corresponding eigenvalue λ_j. For $i \neq j$ we have.

$$\begin{aligned}(A\vec{v}_i) \cdot (A\vec{v}_j) &= (A\vec{v}_i)^\intercal (A\vec{v}_j) \\ &= \vec{v}_i^\intercal A^\intercal A \vec{v}_j \\ &= \vec{v}_i^\intercal \lambda_j \vec{v}_j\end{aligned}$$

How do we do the SVD?

- We already know how to find Σ
- All we have to do is find U and V. We'll find V first.

We know there are eigenvectors $\{\vec{v}_1, \ldots, \vec{v}_n\}$ of $A^\intercal A$ that form an orthonormal basis for $\mathbb{R}^n$.

Then $V = \left[\begin{array}{ccc} \vec{v}_1 & \cdots & \vec{v}_n \end{array}\right]$ is an orthogonal $n \times n$ matrix.

Consider $\{A\vec{v}_1, \ldots, A\vec{v}_n\}$. I claim this is an orthogonal set of vectors. To see this suppose $\vec{v}_j$ is an eigenvector of $A^\intercal A$ with corresponding eigenvalue λ_j. For $i \neq j$ we have.

$$\begin{aligned}(A\vec{v}_i) \cdot (A\vec{v}_j) &= (A\vec{v}_i)^\intercal (A\vec{v}_j) \\ &= \vec{v}_i^\intercal A^\intercal A\vec{v}_j \\ &= \vec{v}_i^\intercal \lambda_j \vec{v}_j \\ &= \lambda_j(\vec{v}_i \cdot \vec{v}_j) = 0\end{aligned}$$

- We already know how to find Σ
- All we have to do is find U and V. We'll find V first.

We know there are eigenvectors $\{\vec{v}_1, \ldots, \vec{v}_n\}$ of $A^\intercal A$ that form an orthonormal basis for $\mathbb{R}^n$.

Then $V = \begin{bmatrix} \vec{v}_1 & \cdots & \vec{v}_n \end{bmatrix}$ is an orthogonal $n \times n$ matrix.

Consider $\{A\vec{v}_1, \ldots, A\vec{v}_n\}$. I claim this is an orthogonal set of vectors. To see this suppose $\vec{v}_j$ is an eigenvector of $A^\intercal A$ with corresponding eigenvalue λ_j. For $i \neq j$ we have.

$$\begin{aligned} (A\vec{v}_i) \cdot (A\vec{v}_j) &= (A\vec{v}_i)^\intercal (A\vec{v}_j) \\ &= \vec{v}_i^\intercal A^\intercal A \vec{v}_j \\ &= \vec{v}_i^\intercal \lambda_j \vec{v}_j \\ &= \lambda_j (\vec{v}_i \cdot \vec{v}_j) = 0 \end{aligned}$$

How do we find U in the SVD?

Now recall that $\sigma_i = ||A\vec{v}_i||$ and r of these are nonzero. Thus

$$\left\{ \frac{1}{\sigma_1} A\vec{v}_1, \frac{1}{\sigma_2} A\vec{v}_2, \ldots, \frac{1}{\sigma_r} A\vec{v}_r \right\}$$

is an orthonormal set in $\mathbb{R}^m$.

How do we find U in the SVD?

Now recall that $\sigma_i = ||A\vec{v}_i||$ and r of these are nonzero. Thus

$$\left\{ \frac{1}{\sigma_1} A\vec{v}_1, \frac{1}{\sigma_2} A\vec{v}_2, \ldots, \frac{1}{\sigma_r} A\vec{v}_r \right\}$$

is an orthonormal set in $\mathbb{R}^m$.

We call $\vec{u}_i = \frac{1}{\sigma_i} A\vec{v}_i$.

How do we find U in the SVD?

Now recall that $\sigma_i = ||A\vec{v}_i||$ and r of these are nonzero. Thus

$$\left\{ \frac{1}{\sigma_1} A\vec{v}_1, \frac{1}{\sigma_2} A\vec{v}_2, \ldots, \frac{1}{\sigma_r} A\vec{v}_r \right\}$$

is an orthonormal set in $\mathbb{R}^m$.

We call $\vec{u}_i = \frac{1}{\sigma_i} A\vec{v}_i$.
Then let $\vec{u}_{r+1}, \ldots, \vec{u}_m$ be vectors such that

$$U = \begin{bmatrix} \vec{u}_1 & \cdots & \vec{u}_r & \vec{u}_{r+1} & \cdots & \vec{u}_m \end{bmatrix}$$

is an orthogonal matrix.

How do we find U in the SVD?

Now recall that $\sigma_i = ||A\vec{v}_i||$ and r of these are nonzero. Thus

$$\left\{ \frac{1}{\sigma_1}A\vec{v}_1, \frac{1}{\sigma_2}A\vec{v}_2, \ldots, \frac{1}{\sigma_r}A\vec{v}_r \right\}$$

is an orthonormal set in $\mathbb{R}^m$.

We call $\vec{u}_i = \frac{1}{\sigma_i}A\vec{v}_i$.
Then let $\vec{u}_{r+1}, \ldots, \vec{u}_m$ be vectors such that

$$U = \begin{bmatrix} \vec{u}_1 & \cdots & \vec{u}_r & \vec{u}_{r+1} & \cdots & \vec{u}_m \end{bmatrix}$$

is an orthogonal matrix.

Do these U and V work?

Now let's convince ourselves that $A = U\Sigma V^{\intercal}$.

Now since V is orthogonal $V^{\intercal}V = I$ (by definition) so it is enough to show

$$AV = U\Sigma$$

Do these U and V work?

Now let's convince ourselves that $A = U\Sigma V^\intercal$.

Now since V is orthogonal $V^\intercal V = I$ (by definition) so it is enough to show

$$AV = U\Sigma$$

(Aside: the columns of V are called $\vec{v}_i$ and the columns of U are called $\vec{u}_i$)

Do these U and V work?

Now let's convince ourselves that $A = U\Sigma V^{\intercal}$.

Now since V is orthogonal $V^{\intercal}V = I$ (by definition) so it is enough to show

$$AV = U\Sigma$$

(Aside: the columns of V are called $\vec{v}_i$ and the columns of U are called $\vec{u}_i$)

We constructed U such that $A\vec{v}_i = \sigma_i\vec{u}_i$ for $i = 1, \ldots, r$.

Do these U and V work?

Now let's convince ourselves that $A = U\Sigma V^\intercal$.

Now since V is orthogonal $V^\intercal V = I$ (by definition) so it is enough to show

$$AV = U\Sigma$$

(Aside: the columns of V are called $\vec{v}_i$ and the columns of U are called $\vec{u}_i$)

We constructed U such that $A\vec{v}_i = \sigma_i \vec{u}_i$ for $i = 1, \dots, r$.

And $||A\vec{v}_i|| = \sigma_i = 0$ for $i = r+1, \dots, n$.

Do these U and V work?

Now let's convince ourselves that $A = U\Sigma V^\intercal$.

Now since V is orthogonal $V^\intercal V = I$ (by definition) so it is enough to show

$$AV = U\Sigma$$

(Aside: the columns of V are called $\vec{v}_i$ and the columns of U are called $\vec{u}_i$)

We constructed U such that $A\vec{v}_i = \sigma_i \vec{u}_i$ for $i = 1, \ldots, r$.

And $||A\vec{v}_i|| = \sigma_i = 0$ for $i = r + 1, \ldots, n$.

Do these U and V work?

$$AV = A\left[\begin{array}{ccc} \vec{v}_1 & \cdots & \vec{v}_n \end{array}\right]$$

Do these U and V work?

$$
\begin{aligned}
AV &= A\begin{bmatrix} \vec{v}_1 & \cdots & \vec{v}_n \end{bmatrix} \\
&= \begin{bmatrix} A\vec{v}_1 & \cdots & A\vec{v}_n \end{bmatrix}
\end{aligned}
$$

Do these U and V work?

$$\begin{aligned}
AV &= A\left[\begin{array}{ccc} \vec{v}_1 & \cdots & \vec{v}_n \end{array}\right] \\
&= \left[\begin{array}{ccc} A\vec{v}_1 & \cdots & A\vec{v}_n \end{array}\right] \\
&= \left[\begin{array}{cccccc} \sigma_1\vec{u}_1 & \cdots & \sigma_r\vec{u}_r & 0 & \cdots & 0 \end{array}\right]
\end{aligned}$$

Do these U and V work?

$$
\begin{aligned}
AV &= A \begin{bmatrix} \vec{v}_1 & \cdots & \vec{v}_n \end{bmatrix} \\
&= \begin{bmatrix} A\vec{v}_1 & \cdots & A\vec{v}_n \end{bmatrix} \\
&= \begin{bmatrix} \sigma_1 \vec{u}_1 & \cdots & \sigma_r \vec{u}_r & 0 & \cdots & 0 \end{bmatrix} \\
&= \begin{bmatrix} \vec{u}_1 & \cdots & \vec{u}_m \end{bmatrix} \left[\begin{array}{ccc|c} \sigma_1 & \cdots & 0 & \\ \vdots & \ddots & \vdots & O \\ 0 & \cdots & \sigma_r & \\ \hline & O & & O \end{array} \right]
\end{aligned}
$$

Do these U and V work?

$$
\begin{aligned}
AV &= A \begin{bmatrix} \vec{v}_1 & \cdots & \vec{v}_n \end{bmatrix} \\
&= \begin{bmatrix} A\vec{v}_1 & \cdots & A\vec{v}_n \end{bmatrix} \\
&= \begin{bmatrix} \sigma_1 \vec{u}_1 & \cdots & \sigma_r \vec{u}_r & 0 & \cdots & 0 \end{bmatrix} \\
&= \begin{bmatrix} \vec{u}_1 & \cdots & \vec{u}_m \end{bmatrix} \left[\begin{array}{ccc|c} \sigma_1 & \cdots & 0 & \\ \vdots & \ddots & \vdots & O \\ 0 & \cdots & \sigma_r & \\ \hline & O & & O \end{array} \right] \\
&= U\Sigma
\end{aligned}
$$

Do these U and V work?

$$
\begin{aligned}
AV &= A \begin{bmatrix} \vec{v}_1 & \cdots & \vec{v}_n \end{bmatrix} \\
&= \begin{bmatrix} A\vec{v}_1 & \cdots & A\vec{v}_n \end{bmatrix} \\
&= \begin{bmatrix} \sigma_1 \vec{u}_1 & \cdots & \sigma_r \vec{u}_r & 0 & \cdots & 0 \end{bmatrix} \\
&= \begin{bmatrix} \vec{u}_1 & \cdots & \vec{u}_m \end{bmatrix} \left[\begin{array}{ccc|c} \sigma_1 & \cdots & 0 & \\ \vdots & \ddots & \vdots & O \\ 0 & \cdots & \sigma_r & \\ \hline & O & & O \end{array} \right] \\
&= U\Sigma
\end{aligned}
$$

- Now, other then exactly defining the "padding" we have to do with U, we have proven Theorem 1.1 on the handout

Do these U and V work?

$$
\begin{aligned}
AV &= A \begin{bmatrix} \vec{v}_1 & \cdots & \vec{v}_n \end{bmatrix} \\
&= \begin{bmatrix} A\vec{v}_1 & \cdots & A\vec{v}_n \end{bmatrix} \\
&= \begin{bmatrix} \sigma_1 \vec{u}_1 & \cdots & \sigma_r \vec{u}_r & 0 & \cdots & 0 \end{bmatrix} \\
&= \begin{bmatrix} \vec{u}_1 & \cdots & \vec{u}_m \end{bmatrix} \left[\begin{array}{ccc|c} \sigma_1 & \cdots & 0 & \\ \vdots & \ddots & \vdots & O \\ 0 & \cdots & \sigma_r & \\ \hline & O & & O \end{array} \right] \\
&= U\Sigma
\end{aligned}
$$

- Now, other then exactly defining the "padding" we have to do with U, we have proven Theorem 1.1 on the handout
- But how does this help us with image compression?

Do these U and V work?

$$
\begin{aligned}
AV &= A \begin{bmatrix} \vec{v}_1 & \cdots & \vec{v}_n \end{bmatrix} \\
&= \begin{bmatrix} A\vec{v}_1 & \cdots & A\vec{v}_n \end{bmatrix} \\
&= \begin{bmatrix} \sigma_1 \vec{u}_1 & \cdots & \sigma_r \vec{u}_r & 0 & \cdots & 0 \end{bmatrix} \\
&= \begin{bmatrix} \vec{u}_1 & \cdots & \vec{u}_m \end{bmatrix} \left[\begin{array}{ccc|c} \sigma_1 & \cdots & 0 & \\ \vdots & \ddots & \vdots & O \\ 0 & \cdots & \sigma_r & \\ \hline & O & & O \end{array} \right] \\
&= U\Sigma
\end{aligned}
$$

- Now, other then exactly defining the "padding" we have to do with U, we have proven Theorem 1.1 on the handout
- But how does this help us with image compression?

The Outer Product Form of the SVD

- It turns out we have a different way of representing the SVD

The Outer Product Form of the SVD

- It turns out we have a different way of representing the SVD

$A = U\Sigma V^{\intercal}$

The Outer Product Form of the SVD

- It turns out we have a different way of representing the SVD

$A = U\Sigma V^{\intercal}$

The Outer Product Form of the SVD

- It turns out we have a different way of representing the SVD

$$A = U\Sigma V^{\intercal}$$

$$= \begin{bmatrix} \vec{u}_1 & \cdots & \vec{u}_m \end{bmatrix} \left[\begin{array}{ccc|c} \sigma_1 & \cdots & 0 & \\ \vdots & \ddots & \vdots & O \\ 0 & \cdots & \sigma_r & \\ \hline & O & & O \end{array}\right] \begin{bmatrix} \vec{v}_1^{\intercal} \\ \vdots \\ \vec{v}_n^{\intercal} \end{bmatrix}$$

The Outer Product Form of the SVD

- It turns out we have a different way of representing the SVD

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

$$= \begin{bmatrix} \vec{u}_1 & \cdots & \vec{u}_m \end{bmatrix} \left[\begin{array}{ccc|c} \sigma_1 & \cdots & 0 & \\ \vdots & \ddots & \vdots & O \\ 0 & \cdots & \sigma_r & \\ \hline & O & & O \end{array}\right] \begin{bmatrix} \vec{v}_1^\intercal \\ \vdots \\ \vec{v}_n^\intercal \end{bmatrix}$$

$$= \left[\begin{array}{ccc|ccc} \vec{u}_1 & \cdots & \vec{u}_r & \vec{u}_{r+1} & \cdots & \vec{u}_m \end{array}\right] \left[\begin{array}{ccc|c} \sigma_1 & \cdots & 0 & \\ \vdots & \ddots & \vdots & O \\ 0 & \cdots & \sigma_r & \\ \hline & O & & O \end{array}\right] \left[\begin{array}{c} \vec{v}_1^\intercal \\ \vdots \\ \vec{v}_r^\intercal \\ \hline \vec{v}_{r+1}^\intercal \\ \vdots \\ \vec{v}_n^\intercal \end{array}\right]$$

The Outer Product Form of the SVD

- It turns out we have a different way of representing the SVD

$$A = U\Sigma V^{\intercal}$$

$$= \begin{bmatrix} \vec{u}_1 & \cdots & \vec{u}_m \end{bmatrix} \left[\begin{array}{ccc|c} \sigma_1 & \cdots & 0 & \\ \vdots & \ddots & \vdots & O \\ 0 & \cdots & \sigma_r & \\ \hline & O & & O \end{array}\right] \begin{bmatrix} \vec{v}_1^{\intercal} \\ \vdots \\ \vec{v}_n^{\intercal} \end{bmatrix}$$

$$= \left[\begin{array}{ccc|ccc} \vec{u}_1 & \cdots & \vec{u}_r & \vec{u}_{r+1} & \cdots & \vec{u}_m \end{array}\right] \left[\begin{array}{ccc|c} \sigma_1 & \cdots & 0 & \\ \vdots & \ddots & \vdots & O \\ 0 & \cdots & \sigma_r & \\ \hline & O & & O \end{array}\right] \left[\begin{array}{c} \vec{v}_1^{\intercal} \\ \vdots \\ \vec{v}_r^{\intercal} \\ \hline \vec{v}_{r+1}^{\intercal} \\ \vdots \\ \vec{v}_n^{\intercal} \end{array}\right]$$

The Outer Product Form of the SVD

$$A = \begin{bmatrix} \vec{u}_1 & \cdots & \vec{u}_r \end{bmatrix} \begin{bmatrix} \sigma_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_r \end{bmatrix} \begin{bmatrix} \vec{v_1}^{\mathsf{T}} \\ \vdots \\ \vec{v}_r^{\mathsf{T}} \end{bmatrix} + \begin{bmatrix} \vec{u}_{r+1} & \cdots & \vec{u}_m \end{bmatrix} [O] \begin{bmatrix} \vec{v_{r+1}}^{\mathsf{T}} \\ \vdots \\ \vec{v}_n^{\mathsf{T}} \end{bmatrix}$$

The Outer Product Form of the SVD

$$A = \begin{bmatrix} \vec{u}_1 & \cdots & \vec{u}_r \end{bmatrix} \begin{bmatrix} \sigma_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_r \end{bmatrix} \begin{bmatrix} \vec{v_1}^\mathsf{T} \\ \vdots \\ \vec{v}_r^\mathsf{T} \end{bmatrix}$$

$$+ \begin{bmatrix} \vec{u}_{r+1} & \cdots & \vec{u}_m \end{bmatrix} [O] \begin{bmatrix} \vec{v_{r+1}}^\mathsf{T} \\ \vdots \\ \vec{v}_n^\mathsf{T} \end{bmatrix}$$

$$= \begin{bmatrix} \sigma_1 \vec{u}_1 & \cdots & \sigma_r \vec{u}_r \end{bmatrix} \begin{bmatrix} \vec{v_1}^\mathsf{T} \\ \vdots \\ \vec{v}_r^\mathsf{T} \end{bmatrix}$$

$$= \sigma_1 \vec{u}_1 \vec{v}_1^\mathsf{T} + \sigma_2 \vec{u}_2 \vec{v}_2^\mathsf{T} + \cdots + \sigma_r \vec{u}_r \vec{v}_r^\mathsf{T}$$

The Outer Product Form of the SVD

$$A = \begin{bmatrix} \vec{u}_1 & \cdots & \vec{u}_r \end{bmatrix} \begin{bmatrix} \sigma_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_r \end{bmatrix} \begin{bmatrix} \vec{v_1}^\intercal \\ \vdots \\ \vec{v_r}^\intercal \end{bmatrix}$$

$$+ \begin{bmatrix} \vec{u}_{r+1} & \cdots & \vec{u}_m \end{bmatrix} [O] \begin{bmatrix} \vec{v_{r+1}}^\intercal \\ \vdots \\ \vec{v_n}^\intercal \end{bmatrix}$$

$$= \begin{bmatrix} \sigma_1 \vec{u}_1 & \cdots & \sigma_r \vec{u}_r \end{bmatrix} \begin{bmatrix} \vec{v_1}^\intercal \\ \vdots \\ \vec{v_r}^\intercal \end{bmatrix}$$

$$= \sigma_1 \vec{u}_1 \vec{v}_1^\intercal + \sigma_2 \vec{u}_2 \vec{v}_2^\intercal + \cdots + \sigma_r \vec{u}_r \vec{v}_r^\intercal$$

The Outer Product Form of the SVD

$$A = \begin{bmatrix} \vec{u}_1 & \cdots & \vec{u}_r \end{bmatrix} \begin{bmatrix} \sigma_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_r \end{bmatrix} \begin{bmatrix} \vec{v_1}^T \\ \vdots \\ \vec{v}_r^T \end{bmatrix}$$

$$+ \begin{bmatrix} \vec{u}_{r+1} & \cdots & \vec{u}_m \end{bmatrix} [O] \begin{bmatrix} \vec{v_{r+1}}^T \\ \vdots \\ \vec{v}_n^T \end{bmatrix}$$

$$= \begin{bmatrix} \sigma_1 \vec{u}_1 & \cdots & \sigma_r \vec{u}_r \end{bmatrix} \begin{bmatrix} \vec{v_1}^T \\ \vdots \\ \vec{v}_r^T \end{bmatrix}$$

$$= \sigma_1 \vec{u}_1 \vec{v}_1^T + \sigma_2 \vec{u}_2 \vec{v}_2^T + \cdots + \sigma_r \vec{u}_r \vec{v}_r^T$$

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

The Outer Product Form of the SVD

$$
\begin{aligned}
A &= \begin{bmatrix} \vec{u}_1 & \cdots & \vec{u}_r \end{bmatrix} \begin{bmatrix} \sigma_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_r \end{bmatrix} \begin{bmatrix} \vec{v_1}^{\mathsf{T}} \\ \vdots \\ \vec{v}_r^{\mathsf{T}} \end{bmatrix} \\
&\qquad + \begin{bmatrix} \vec{u}_{r+1} & \cdots & \vec{u}_m \end{bmatrix} [O] \begin{bmatrix} \vec{v_{r+1}}^{\mathsf{T}} \\ \vdots \\ \vec{v}_n^{\mathsf{T}} \end{bmatrix} \\
&= \begin{bmatrix} \sigma_1 \vec{u}_1 & \cdots & \sigma_r \vec{u}_r \end{bmatrix} \begin{bmatrix} \vec{v_1}^{\mathsf{T}} \\ \vdots \\ \vec{v}_r^{\mathsf{T}} \end{bmatrix} \\
&= \sigma_1 \vec{u}_1 \vec{v}_1^{\mathsf{T}} + \sigma_2 \vec{u}_2 \vec{v}_2^{\mathsf{T}} + \cdots + \sigma_r \vec{u}_r \vec{v}_r^{\mathsf{T}}
\end{aligned}
$$

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

The heart of SVD Compression

$$A = \sigma_1 \vec{u}_1 \vec{v}_1^\intercal + \sigma_2 \vec{u}_2 \vec{v}_2^\intercal + \cdots + \sigma_r \vec{u}_r \vec{v}_r^\intercal$$

We can talk about an approximation A_k for some $k \leq r$.

The heart of SVD Compression

$$A = \sigma_1 \vec{u}_1 \vec{v}_1^\intercal + \sigma_2 \vec{u}_2 \vec{v}_2^\intercal + \cdots + \sigma_r \vec{u}_r \vec{v}_r^\intercal$$

We can talk about an approximation A_k for some $k \leq r$.

$$A_k = \sigma_1 \vec{u}_1 \vec{v}_1^\intercal + \sigma_2 \vec{u}_2 \vec{v}_2^\intercal + \cdots + \sigma_r \vec{u}_k \vec{v}_k^\intercal$$

$$A_r = A$$

The heart of SVD Compression

$$A = \sigma_1 \vec{u}_1 \vec{v}_1^{\mathsf{T}} + \sigma_2 \vec{u}_2 \vec{v}_2^{\mathsf{T}} + \cdots + \sigma_r \vec{u}_r \vec{v}_r^{\mathsf{T}}$$

We can talk about an approximation A_k for some $k \leq r$.

$$A_k = \sigma_1 \vec{u}_1 \vec{v}_1^{\mathsf{T}} + \sigma_2 \vec{u}_2 \vec{v}_2^{\mathsf{T}} + \cdots + \sigma_r \vec{u}_k \vec{v}_k^{\mathsf{T}}$$

$$A_r = A$$

$$\sigma_i \in \mathbb{R}, \ \ \vec{u}_i \in \mathbb{R}^m, \ \ \vec{v}_i \in \mathbb{R}^n$$

So for A_k we're really only storing $k(1 + n + m)$ numbers.

The heart of SVD Compression

$$A = \sigma_1 \vec{u}_1 \vec{v}_1^\intercal + \sigma_2 \vec{u}_2 \vec{v}_2^\intercal + \cdots + \sigma_r \vec{u}_r \vec{v}_r^\intercal$$

We can talk about an approximation A_k for some $k \leq r$.

$$A_k = \sigma_1 \vec{u}_1 \vec{v}_1^\intercal + \sigma_2 \vec{u}_2 \vec{v}_2^\intercal + \cdots + \sigma_r \vec{u}_k \vec{v}_k^\intercal$$

$$A_r = A$$

$$\sigma_i \in \mathbb{R}, \ \ \vec{u}_i \in \mathbb{R}^m, \ \ \vec{v}_i \in \mathbb{R}^n$$

So for A_k we're really only storing $k(1 + n + m)$ numbers.