Linear Algebra Review

CS221: Introduction to Artificial Intelligence

Carlos Fernandez-Granda

10/11/2011

Index

What is a vector?

What is a vector ?

- An ordered tuple of numbers :

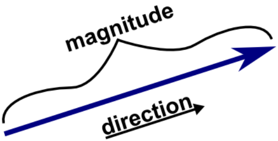
$$\mathbf{u} = \begin{pmatrix} u_1 \\ u_2 \\ \dots \\ u_n \end{pmatrix}$$

What is a vector ?

- An ordered tuple of numbers :

$$\mathbf{u} = \begin{pmatrix} u_1 \\ u_2 \\ \dots \\ u_n \end{pmatrix}$$

- A quantity that has a magnitude and a direction.



Vector spaces

- More formally a vector is an element of a **vector space**.

Vector spaces

- More formally a vector is an element of a **vector space**.
- A vector space $\mathcal{V}$ is a set that contains **all linear combinations** of its elements :

Vector spaces

- More formally a vector is an element of a **vector space**.
- A vector space $\mathcal{V}$ is a set that contains **all linear combinations** of its elements :
  - If vectors $\mathbf{u}$ and $\mathbf{v} \in \mathcal{V}$, then $\mathbf{u} + \mathbf{v} \in \mathcal{V}$.

Vector spaces

- More formally a vector is an element of a **vector space**.
- A vector space $\mathcal{V}$ is a set that contains **all linear combinations** of its elements :
  - If vectors $\mathbf{u}$ and $\mathbf{v} \in \mathcal{V}$, then $\mathbf{u} + \mathbf{v} \in \mathcal{V}$.
  - If vector $\mathbf{u} \in \mathcal{V}$, then $\alpha\, \mathbf{u} \in \mathcal{V}$ for any scalar α.

Vector spaces

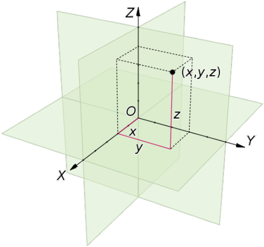
- More formally a vector is an element of a **vector space**.
- A vector space $\mathcal{V}$ is a set that contains **all linear combinations** of its elements :
  - If vectors $\mathbf{u}$ and $\mathbf{v} \in \mathcal{V}$, then $\mathbf{u} + \mathbf{v} \in \mathcal{V}$.
  - If vector $\mathbf{u} \in \mathcal{V}$, then $\alpha\, \mathbf{u} \in \mathcal{V}$ for any scalar α.
  - There exists $\mathbf{0} \in \mathcal{V}$ such that $\mathbf{u} + \mathbf{0} = \mathbf{u}$ for any $u \in \mathcal{V}$.

Vector spaces

- More formally a vector is an element of a **vector space**.
- A vector space $\mathcal{V}$ is a set that contains **all linear combinations** of its elements :
  - If vectors $\mathbf{u}$ and $\mathbf{v} \in \mathcal{V}$, then $\mathbf{u} + \mathbf{v} \in \mathcal{V}$.
  - If vector $\mathbf{u} \in \mathcal{V}$, then $\alpha\, \mathbf{u} \in \mathcal{V}$ for any scalar α.
  - There exists $\mathbf{0} \in \mathcal{V}$ such that $\mathbf{u} + \mathbf{0} = \mathbf{u}$ for any $u \in \mathcal{V}$.
- A **subspace** is a subset of a vector space that is **also** a vector space.

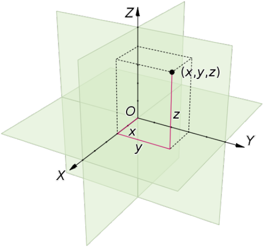
Examples

- Euclidean space $\mathbb{R}^n$.



Examples

- Euclidean space $\mathbb{R}^n$.

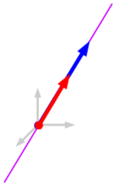


- The **span** of any set of vectors $\{\mathbf{u_1}, \mathbf{u_2}, \ldots, \mathbf{u_n}\}$, defined as :

$$\text{span}\left(\mathbf{u_1}, \mathbf{u_2}, \ldots, \mathbf{u_n}\right) = \{\alpha_1\mathbf{u_1} + \alpha_2\mathbf{u_2} + \cdots + \alpha_n\mathbf{u_n} \mid \alpha_i \in \mathbb{R}\}$$

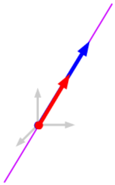
Subspaces

- A **line** through the origin in $\mathbb{R}^n$(span of a vector).

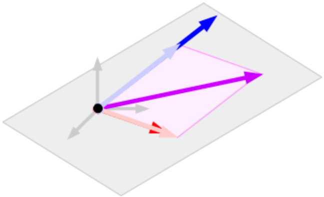


Subspaces

- A **line** through the origin in $\mathbb{R}^n$(span of a vector).



- A **plane** in $\mathbb{R}^n$(span of two vectors).



Linear Independence and Basis of Vector Spaces

- A vector $\mathbf{u}$ is linearly independent of a set of vectors $\{\mathbf{u_1}, \mathbf{u_2}, \ldots, \mathbf{u_n}\}$ if it does **not** lie in their **span**.

Linear Independence and Basis of Vector Spaces

- A vector $\mathbf{u}$ is linearly independent of a set of vectors $\{\mathbf{u_1}, \mathbf{u_2}, \ldots, \mathbf{u_n}\}$ if it does **not** lie in their **span**.
- A set of vectors is linearly independent if every vector is linearly independent of the rest.

Linear Independence and Basis of Vector Spaces

- A vector $\mathbf{u}$ is linearly independent of a set of vectors $\{\mathbf{u_1}, \mathbf{u_2}, \ldots, \mathbf{u_n}\}$ if it does **not** lie in their **span**.
- A set of vectors is linearly independent if every vector is linearly independent of the rest.
- A **basis** of a vector space $\mathcal{V}$ is a linearly independent set of vectors whose **span is equal to** $\mathcal{V}$.

Linear Independence and Basis of Vector Spaces

- A vector $\mathbf{u}$ is linearly independent of a set of vectors $\{\mathbf{u_1}, \mathbf{u_2}, \ldots, \mathbf{u_n}\}$ if it does **not** lie in their **span**.
- A set of vectors is linearly independent if every vector is linearly independent of the rest.
- A **basis** of a vector space $\mathcal{V}$ is a linearly independent set of vectors whose **span is equal to** $\mathcal{V}$.
- If a vector space has a basis with d vectors, its **dimension** is d.

Linear Independence and Basis of Vector Spaces

- A vector $\mathbf{u}$ is linearly independent of a set of vectors $\{\mathbf{u_1}, \mathbf{u_2}, \ldots, \mathbf{u_n}\}$ if it does **not** lie in their **span**.
- A set of vectors is linearly independent if every vector is linearly independent of the rest.
- A **basis** of a vector space $\mathcal{V}$ is a linearly independent set of vectors whose **span is equal to** $\mathcal{V}$.
- If a vector space has a basis with d vectors, its **dimension** is d.
- Any other basis of the same space will consist of **exactly d** vectors.

Linear Independence and Basis of Vector Spaces

- A vector $\mathbf{u}$ is linearly independent of a set of vectors $\{\mathbf{u_1}, \mathbf{u_2}, \ldots, \mathbf{u_n}\}$ if it does **not** lie in their **span**.
- A set of vectors is linearly independent if every vector is linearly independent of the rest.
- A **basis** of a vector space $\mathcal{V}$ is a linearly independent set of vectors whose **span is equal to** $\mathcal{V}$.
- If a vector space has a basis with d vectors, its **dimension** is d.
- Any other basis of the same space will consist of **exactly d** vectors.
- The dimension of a vector space can be infinite (function spaces).

Norm of a Vector

- A norm $||\mathbf{u}||$ measures the **magnitude** of a vector.

Norm of a Vector

- A norm $||\mathbf{u}||$ measures the **magnitude** of a vector.
- Mathematically, any function from the vector space to $\mathbb{R}$ that satisfies :

Norm of a Vector

- A norm $||\mathbf{u}||$ measures the **magnitude** of a vector.
- Mathematically, any function from the vector space to $\mathbb{R}$ that satisfies :
  - Homogeneity $||\alpha\ \mathbf{u}|| = |\alpha|\ ||\mathbf{u}||$

Norm of a Vector

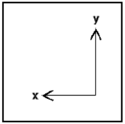
- A norm $||\mathbf{u}||$ measures the **magnitude** of a vector.
- Mathematically, any function from the vector space to $\mathbb{R}$ that satisfies :
  - Homogeneity $||\alpha\ \mathbf{u}|| = |\alpha|\ ||\mathbf{u}||$
  - Triangle inequality $||\mathbf{u} + \mathbf{v}|| \leq ||\mathbf{u}|| + ||\mathbf{v}||$

Norm of a Vector

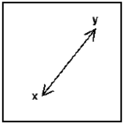
- A norm $||\mathbf{u}||$ measures the **magnitude** of a vector.
- Mathematically, any function from the vector space to $\mathbb{R}$ that satisfies :
  - Homogeneity $||\alpha\ \mathbf{u}|| = |\alpha|\ ||\mathbf{u}||$
  - Triangle inequality $||\mathbf{u} + \mathbf{v}|| \leq ||\mathbf{u}|| + ||\mathbf{v}||$
  - Point separation $||\mathbf{u}|| = 0$ if and only if $\mathbf{u} = 0$.

Norm of a Vector

- A norm $||\mathbf{u}||$ measures the **magnitude** of a vector.
- Mathematically, any function from the vector space to $\mathbb{R}$ that satisfies :
  - Homogeneity $||\alpha\ \mathbf{u}|| = |\alpha|\ ||\mathbf{u}||$
  - Triangle inequality $||\mathbf{u} + \mathbf{v}|| \leq ||\mathbf{u}|| + ||\mathbf{v}||$
  - Point separation $||\mathbf{u}|| = 0$ if and only if $\mathbf{u} = 0$.
- Examples :
  - Manhattan or ℓ_1 norm : $||\mathbf{u}||_1 = \sum_i |u_i|$.
  - Euclidean or ℓ_2 norm : $||\mathbf{u}||_2 = \sqrt{\sum_i u_i^2}$.



Manhattan



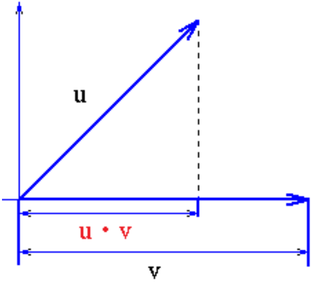
Euclidean

Inner Product

- Inner product between **u** and **v** : $\langle \mathbf{u}, \mathbf{v} \rangle = \sum_{i=1}^{n} u_i v_i$.

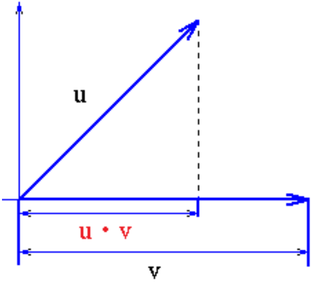
Inner Product

- Inner product between **u** and **v** : $\langle \mathbf{u}, \mathbf{v} \rangle = \sum_{i=1}^{n} u_i v_i$.
- It is the **projection** of one vector onto the other.



Inner Product

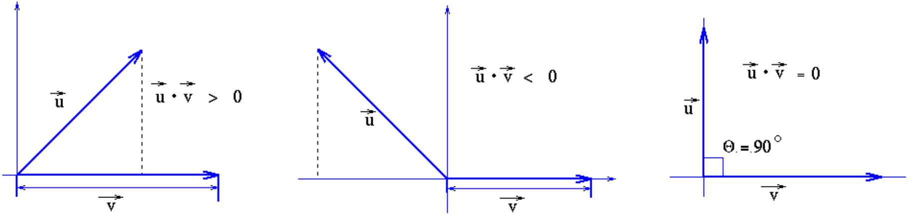
- Inner product between **u** and **v** : $\langle \mathbf{u}, \mathbf{v} \rangle = \sum_{i=1}^{n} u_i v_i$.
- It is the **projection** of one vector onto the other.



- Related to the Euclidean norm : $\langle \mathbf{u}, \mathbf{u} \rangle = ||\mathbf{u}||_2^2$.

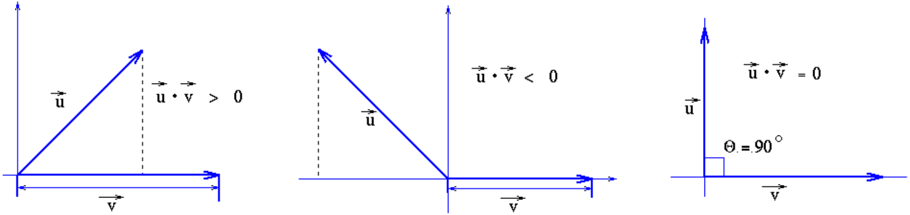
Inner Product

The inner product is a measure of **correlation** between two vectors, scaled by the norms of the vectors :



Inner Product

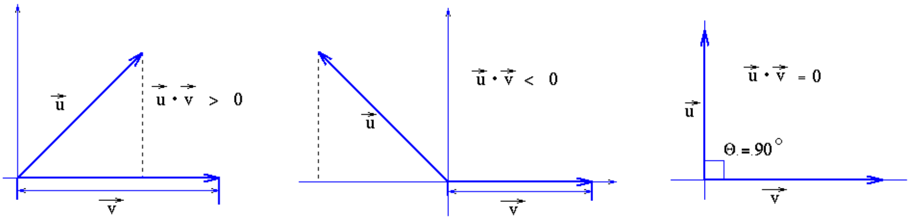
The inner product is a measure of **correlation** between two vectors, scaled by the norms of the vectors :



- If $\langle \mathbf{u}, \mathbf{v} \rangle > 0$, **u** and **v** are **aligned**.

Inner Product

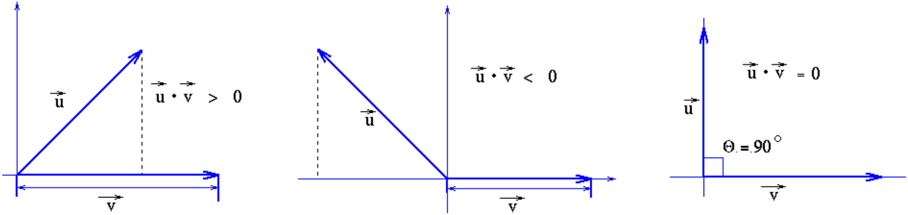
The inner product is a measure of **correlation** between two vectors, scaled by the norms of the vectors :



- If $\langle \mathbf{u}, \mathbf{v} \rangle > 0$, $\mathbf{u}$ and $\mathbf{v}$ are **aligned**.
- If $\langle \mathbf{u}, \mathbf{v} \rangle < 0$, $\mathbf{u}$ and $\mathbf{v}$ are **opposed**.

Inner Product

The inner product is a measure of **correlation** between two vectors, scaled by the norms of the vectors :



- If $\langle \mathbf{u}, \mathbf{v} \rangle > 0$, $\mathbf{u}$ and $\mathbf{v}$ are **aligned**.
- If $\langle \mathbf{u}, \mathbf{v} \rangle < 0$, $\mathbf{u}$ and $\mathbf{v}$ are **opposed**.
- If $\langle \mathbf{u}, \mathbf{v} \rangle = 0$, $\mathbf{u}$ and $\mathbf{v}$ are **orthogonal**.

Orthonormal Basis

- The vectors in an **orthonormal** basis have unit Euclidean norm and are orthogonal to each other.

Orthonormal Basis

- The vectors in an **orthonormal** basis have unit Euclidean norm and are orthogonal to each other.
- To express a vector $\mathbf{x}$ in an orthonormal basis, we can just take inner products.

Orthonormal Basis

- The vectors in an **orthonormal** basis have unit Euclidean norm and are orthogonal to each other.
- To express a vector $\mathbf{x}$ in an orthonormal basis, we can just take inner products.
- Example : $\mathbf{x} = \alpha_1 \mathbf{b_1} + \alpha_2 \mathbf{b_2}$.

Orthonormal Basis

- The vectors in an **orthonormal** basis have unit Euclidean norm and are orthogonal to each other.
- To express a vector $\mathbf{x}$ in an orthonormal basis, we can just take inner products.
- Example : $\mathbf{x} = \alpha_1 \mathbf{b_1} + \alpha_2 \mathbf{b_2}$.
- We compute $\langle \mathbf{x}, \mathbf{b_1} \rangle$:

$$\begin{aligned}
\langle \mathbf{x}, \mathbf{b_1} \rangle &= \langle \alpha_1 \mathbf{b_1} + \alpha_2 \mathbf{b_2}, \mathbf{b_1} \rangle \\
&= \alpha_1 \langle \mathbf{b_1}, \mathbf{b_1} \rangle + \alpha_2 \langle \mathbf{b_2}, \mathbf{b_1} \rangle \quad \text{By linearity.} \\
&= \alpha_1 + 0 \quad \text{By orthonormality.}
\end{aligned}$$

Orthonormal Basis

- The vectors in an **orthonormal** basis have unit Euclidean norm and are orthogonal to each other.
- To express a vector $\mathbf{x}$ in an orthonormal basis, we can just take inner products.
- Example : $\mathbf{x} = \alpha_1 \mathbf{b_1} + \alpha_2 \mathbf{b_2}$.
- We compute $\langle \mathbf{x}, \mathbf{b_1} \rangle$:

$$\begin{aligned} \langle \mathbf{x}, \mathbf{b_1} \rangle &= \langle \alpha_1 \mathbf{b_1} + \alpha_2 \mathbf{b_2}, \mathbf{b_1} \rangle \\ &= \alpha_1 \langle \mathbf{b_1}, \mathbf{b_1} \rangle + \alpha_2 \langle \mathbf{b_2}, \mathbf{b_1} \rangle \quad \text{By linearity.} \\ &= \alpha_1 + 0 \quad \text{By orthonormality.} \end{aligned}$$

- Likewise, $\alpha_2 = \langle \mathbf{x}, \mathbf{b_2} \rangle$.

Linear Operator

- A linear operator $\mathcal{L} : \mathcal{U} \to \mathcal{V}$ is a map from a vector space $\mathcal{U}$ to another vector space $\mathcal{V}$ that satisfies :

Linear Operator

- A linear operator $\mathcal{L} : \mathcal{U} \to \mathcal{V}$ is a map from a vector space $\mathcal{U}$ to another vector space $\mathcal{V}$ that satisfies :
  - $\mathcal{L}(\mathbf{u}_1 + \mathbf{u}_2) = \mathcal{L}(\mathbf{u}_1) + \mathcal{L}(\mathbf{u}_2)$.

Linear Operator

- A linear operator $\mathcal{L} : \mathcal{U} \rightarrow \mathcal{V}$ is a map from a vector space $\mathcal{U}$ to another vector space $\mathcal{V}$ that satisfies :
  - $\mathcal{L}(\mathbf{u_1} + \mathbf{u_2}) = \mathcal{L}(\mathbf{u_1}) + \mathcal{L}(\mathbf{u_2})$.
  - $\mathcal{L}(\alpha\, \mathbf{u}) = \alpha\, \mathcal{L}(\mathbf{u})$ for any scalar α.

Linear Operator

- A linear operator $\mathcal{L} : \mathcal{U} \to \mathcal{V}$ is a map from a vector space $\mathcal{U}$ to another vector space $\mathcal{V}$ that satisfies :
  - $\mathcal{L}(\mathbf{u_1} + \mathbf{u_2}) = \mathcal{L}(\mathbf{u_1}) + \mathcal{L}(\mathbf{u_2})$.
  - $\mathcal{L}(\alpha\, \mathbf{u}) = \alpha\, \mathcal{L}(\mathbf{u})$ for any scalar α.
- If the dimension n of $\mathcal{U}$ and m of $\mathcal{V}$ are finite, $\mathcal{L}$ can be represented by an $m \times n$ matrix $\mathbf{A}$.

$$\mathbf{A} = \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ A_{21} & A_{22} & \dots & A_{2n} \\ & & \dots & \\ A_{m1} & A_{m2} & \dots & A_{mn} \end{pmatrix}$$

Matrix Vector Multiplication

$$\mathbf{Au} = \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ A_{21} & A_{22} & \dots & A_{2n} \\ & & \dots & \\ A_{m1} & A_{m2} & \dots & A_{mn} \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ \dots \\ u_n \end{pmatrix} = \begin{pmatrix} v_1 \\ v_2 \\ \dots \\ v_m \end{pmatrix}$$

Matrix Vector Multiplication

$$\mathbf{Au} = \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ A_{21} & A_{22} & \dots & A_{2n} \\ & & \dots & \\ A_{m1} & A_{m2} & \dots & A_{mn} \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ \dots \\ u_n \end{pmatrix} = \begin{pmatrix} v_1 \\ v_2 \\ \dots \\ v_m \end{pmatrix}$$

Matrix Vector Multiplication

$$\mathbf{Au} = \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ A_{21} & A_{22} & \dots & A_{2n} \\ & & \dots & \\ A_{m1} & A_{m2} & \dots & A_{mn} \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ \dots \\ u_n \end{pmatrix} = \begin{pmatrix} v_1 \\ v_2 \\ \dots \\ v_m \end{pmatrix}$$

Matrix Vector Multiplication

$$\mathbf{Au} = \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ A_{21} & A_{22} & \dots & A_{2n} \\ & & \dots & \\ A_{m1} & A_{m2} & \dots & A_{mn} \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ \dots \\ u_n \end{pmatrix} = \begin{pmatrix} v_1 \\ v_2 \\ \dots \\ v_m \end{pmatrix}$$

Matrix Product

Applying two linear operators **A** and **B** defines another linear operator, which can be represented by another matrix **AB**.

$$\mathbf{AB} = \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ A_{21} & A_{22} & \dots & A_{2n} \\ & & \dots & \\ A_{m1} & A_{m2} & \dots & A_{mn} \end{pmatrix} \begin{pmatrix} B_{11} & B_{12} & \dots & B_{1p} \\ B_{21} & B_{22} & \dots & B_{2p} \\ & & \dots & \\ B_{n1} & B_{n2} & \dots & B_{np} \end{pmatrix}$$

$$= \begin{pmatrix} AB_{11} & AB_{12} & \dots & AB_{1p} \\ AB_{21} & AB_{22} & \dots & AB_{2p} \\ & & \dots & \\ AB_{m1} & AB_{m2} & \dots & AB_{mp} \end{pmatrix}$$

Note that if **A** is $m \times n$ and **B** is $n \times p$, then **AB** is $m \times p$.

Matrix Product

Applying two linear operators **A** and **B** defines another linear operator, which can be represented by another matrix **AB**.

$$\mathbf{AB} = \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ A_{21} & A_{22} & \dots & A_{2n} \\ & & \dots & \\ A_{m1} & A_{m2} & \dots & A_{mn} \end{pmatrix} \begin{pmatrix} B_{11} & B_{12} & \dots & B_{1p} \\ B_{21} & B_{22} & \dots & B_{2p} \\ & & \dots & \\ B_{n1} & B_{n2} & \dots & B_{np} \end{pmatrix}$$

$$= \begin{pmatrix} AB_{11} & AB_{12} & \dots & AB_{1p} \\ AB_{21} & AB_{22} & \dots & AB_{2p} \\ & & \dots & \\ AB_{m1} & AB_{m2} & \dots & AB_{mp} \end{pmatrix}$$

Note that if **A** is $m \times n$ and **B** is $n \times p$, then **AB** is $m \times p$.

Matrix Product

Applying two linear operators **A** and **B** defines another linear operator, which can be represented by another matrix **AB**.

$$\mathbf{AB} = \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ A_{21} & A_{22} & \dots & A_{2n} \\ & & \dots & \\ A_{m1} & A_{m2} & \dots & A_{mn} \end{pmatrix} \begin{pmatrix} B_{11} & B_{12} & \dots & B_{1p} \\ B_{21} & B_{22} & \dots & B_{2p} \\ & & \dots & \\ B_{n1} & B_{n2} & \dots & B_{np} \end{pmatrix}$$

$$= \begin{pmatrix} AB_{11} & AB_{12} & \dots & AB_{1p} \\ AB_{21} & AB_{22} & \dots & AB_{2p} \\ & & \dots & \\ AB_{m1} & AB_{m2} & \dots & AB_{mp} \end{pmatrix}$$

Note that if **A** is $m \times n$ and **B** is $n \times p$, then **AB** is $m \times p$.

Matrix Product

Applying two linear operators **A** and **B** defines another linear operator, which can be represented by another matrix **AB**.

$$\mathbf{AB} = \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ A_{21} & A_{22} & \dots & A_{2n} \\ & & \dots & \\ A_{m1} & A_{m2} & \dots & A_{mn} \end{pmatrix} \begin{pmatrix} B_{11} & B_{12} & \dots & B_{1p} \\ B_{21} & B_{22} & \dots & B_{2p} \\ & & \dots & \\ B_{n1} & B_{n2} & \dots & B_{np} \end{pmatrix}$$

$$= \begin{pmatrix} AB_{11} & AB_{12} & \dots & AB_{1p} \\ AB_{21} & AB_{22} & \dots & AB_{2p} \\ & & \dots & \\ AB_{m1} & AB_{m2} & \dots & AB_{mp} \end{pmatrix}$$

Note that if **A** is $m \times n$ and **B** is $n \times p$, then **AB** is $m \times p$.

Matrix Product

Applying two linear operators **A** and **B** defines another linear operator, which can be represented by another matrix **AB**.

$$\mathbf{AB} = \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ A_{21} & A_{22} & \dots & A_{2n} \\ & & \dots & \\ A_{m1} & A_{m2} & \dots & A_{mn} \end{pmatrix} \begin{pmatrix} B_{11} & B_{12} & \dots & B_{1p} \\ B_{21} & B_{22} & \dots & B_{2p} \\ & & \dots & \\ B_{n1} & B_{n2} & \dots & B_{np} \end{pmatrix}$$

$$= \begin{pmatrix} AB_{11} & AB_{12} & \dots & AB_{1p} \\ AB_{21} & AB_{22} & \dots & AB_{2p} \\ & & \dots & \\ AB_{m1} & AB_{m2} & \dots & AB_{mp} \end{pmatrix}$$

Note that if **A** is $m \times n$ and **B** is $n \times p$, then **AB** is $m \times p$.

Matrix Product

Applying two linear operators **A** and **B** defines another linear operator, which can be represented by another matrix **AB**.

$$\mathbf{AB} = \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ A_{21} & A_{22} & \dots & A_{2n} \\ & & \dots & \\ A_{m1} & A_{m2} & \dots & A_{mn} \end{pmatrix} \begin{pmatrix} B_{11} & B_{12} & \dots & B_{1p} \\ B_{21} & B_{22} & \dots & B_{2p} \\ & & \dots & \\ B_{n1} & B_{n2} & \dots & B_{np} \end{pmatrix}$$

$$= \begin{pmatrix} AB_{11} & AB_{12} & \dots & AB_{1p} \\ AB_{21} & AB_{22} & \dots & AB_{2p} \\ & & \dots & \\ AB_{m1} & AB_{m2} & \dots & AB_{mp} \end{pmatrix}$$

Note that if **A** is $m \times n$ and **B** is $n \times p$, then **AB** is $m \times p$.

Matrix Product

Applying two linear operators **A** and **B** defines another linear operator, which can be represented by another matrix **AB**.

$$\mathbf{AB} = \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ A_{21} & A_{22} & \dots & A_{2n} \\ & & \dots & \\ A_{m1} & A_{m2} & \dots & A_{mn} \end{pmatrix} \begin{pmatrix} B_{11} & B_{12} & \dots & B_{1p} \\ B_{21} & B_{22} & \dots & B_{2p} \\ & & \dots & \\ B_{n1} & B_{n2} & \dots & B_{np} \end{pmatrix}$$

$$= \begin{pmatrix} AB_{11} & AB_{12} & \dots & AB_{1p} \\ AB_{21} & AB_{22} & \dots & AB_{2p} \\ & & \dots & \\ AB_{m1} & AB_{m2} & \dots & AB_{mp} \end{pmatrix}$$

Note that if **A** is $m \times n$ and **B** is $n \times p$, then **AB** is $m \times p$.

Matrix Product

Applying two linear operators **A** and **B** defines another linear operator, which can be represented by another matrix **AB**.

$$\mathbf{AB} = \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ A_{21} & A_{22} & \dots & A_{2n} \\ & & \dots & \\ A_{m1} & A_{m2} & \dots & A_{mn} \end{pmatrix} \begin{pmatrix} B_{11} & B_{12} & \dots & B_{1p} \\ B_{21} & B_{22} & \dots & B_{2p} \\ & & \dots & \\ B_{n1} & B_{n2} & \dots & B_{np} \end{pmatrix}$$

$$= \begin{pmatrix} AB_{11} & AB_{12} & \dots & AB_{1p} \\ AB_{21} & AB_{22} & \dots & AB_{2p} \\ & & \dots & \\ AB_{m1} & AB_{m2} & \dots & AB_{mp} \end{pmatrix}$$

Note that if **A** is $m \times n$ and **B** is $n \times p$, then **AB** is $m \times p$.

Matrix Product

Applying two linear operators **A** and **B** defines another linear operator, which can be represented by another matrix **AB**.

$$\mathbf{AB} = \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ A_{21} & A_{22} & \dots & A_{2n} \\ & & \dots & \\ A_{m1} & A_{m2} & \dots & A_{mn} \end{pmatrix} \begin{pmatrix} B_{11} & B_{12} & \dots & B_{1p} \\ B_{21} & B_{22} & \dots & B_{2p} \\ & & \dots & \\ B_{n1} & B_{n2} & \dots & B_{np} \end{pmatrix}$$

$$= \begin{pmatrix} AB_{11} & AB_{12} & \dots & AB_{1p} \\ AB_{21} & AB_{22} & \dots & AB_{2p} \\ & & \dots & \\ AB_{m1} & AB_{m2} & \dots & AB_{mp} \end{pmatrix}$$

Note that if **A** is $m \times n$ and **B** is $n \times p$, then **AB** is $m \times p$.

Matrix Product

Applying two linear operators $\mathbf{A}$ and $\mathbf{B}$ defines another linear operator, which can be represented by another matrix $\mathbf{AB}$.

$$\mathbf{AB} = \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ A_{21} & A_{22} & \dots & A_{2n} \\ & & \dots & \\ A_{m1} & A_{m2} & \dots & A_{mn} \end{pmatrix} \begin{pmatrix} B_{11} & B_{12} & \dots & B_{1p} \\ B_{21} & B_{22} & \dots & B_{2p} \\ & & \dots & \\ B_{n1} & B_{n2} & \dots & B_{np} \end{pmatrix}$$

$$= \begin{pmatrix} AB_{11} & AB_{12} & \dots & AB_{1p} \\ AB_{21} & AB_{22} & \dots & AB_{2p} \\ & & \dots & \\ AB_{m1} & AB_{m2} & \dots & AB_{mp} \end{pmatrix}$$

Note that if $\mathbf{A}$ is $m \times n$ and $\mathbf{B}$ is $n \times p$, then $\mathbf{AB}$ is $m \times p$.

Transpose of a Matrix

- The transpose of a matrix is obtained by flipping the rows and columns.

$$\mathbf{A}^T = \begin{pmatrix} A_{11} & A_{21} & \dots & A_{n1} \\ A_{12} & A_{22} & \dots & A_{n2} \\ & & \dots & \\ A_{1m} & A_{2m} & \dots & A_{nm} \end{pmatrix}$$

Transpose of a Matrix

- The transpose of a matrix is obtained by flipping the rows and columns.

$$\mathbf{A}^T = \begin{pmatrix} A_{11} & A_{21} & \dots & A_{n1} \\ A_{12} & A_{22} & \dots & A_{n2} \\ & & \dots & \\ A_{1m} & A_{2m} & \dots & A_{nm} \end{pmatrix}$$

- Some simple properties :

Transpose of a Matrix

- The transpose of a matrix is obtained by flipping the rows and columns.

$$\mathbf{A}^T = \begin{pmatrix} A_{11} & A_{21} & \dots & A_{n1} \\ A_{12} & A_{22} & \dots & A_{n2} \\ & & \dots & \\ A_{1m} & A_{2m} & \dots & A_{nm} \end{pmatrix}$$

- Some simple properties :
  - $\left(\mathbf{A}^T\right)^T = \mathbf{A}$

Transpose of a Matrix

- The transpose of a matrix is obtained by flipping the rows and columns.

$$\mathbf{A}^T = \begin{pmatrix} A_{11} & A_{21} & \dots & A_{n1} \\ A_{12} & A_{22} & \dots & A_{n2} \\ & & \dots & \\ A_{1m} & A_{2m} & \dots & A_{nm} \end{pmatrix}$$

- Some simple properties :
  - $\left(\mathbf{A}^T\right)^T = \mathbf{A}$
  - $(\mathbf{A} + \mathbf{B})^T = \mathbf{A}^T + \mathbf{B}^T$

Transpose of a Matrix

- The transpose of a matrix is obtained by flipping the rows and columns.

$$\mathbf{A}^T = \begin{pmatrix} A_{11} & A_{21} & \dots & A_{n1} \\ A_{12} & A_{22} & \dots & A_{n2} \\ & & \dots & \\ A_{1m} & A_{2m} & \dots & A_{nm} \end{pmatrix}$$

- Some simple properties :
  - $\left(\mathbf{A}^T\right)^T = \mathbf{A}$
  - $(\mathbf{A}+\mathbf{B})^T = \mathbf{A}^T + \mathbf{B}^T$
  - $(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T$

Transpose of a Matrix

- The transpose of a matrix is obtained by flipping the rows and columns.

$$\mathbf{A}^T = \begin{pmatrix} A_{11} & A_{21} & \dots & A_{n1} \\ A_{12} & A_{22} & \dots & A_{n2} \\ & & \dots & \\ A_{1m} & A_{2m} & \dots & A_{nm} \end{pmatrix}$$

- Some simple properties :
  - $\left(\mathbf{A}^T\right)^T = \mathbf{A}$
  - $(\mathbf{A}+\mathbf{B})^T = \mathbf{A}^T + \mathbf{B}^T$
  - $(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T$
- The inner product for vectors can be represented as $\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u}^T\mathbf{v}$.

Identity Operator

$$\mathrm{I} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ & & \dots & \\ 0 & 0 & \dots & 1 \end{pmatrix}$$

Identity Operator

$$\mathtt{I} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ & & \dots & \\ 0 & 0 & \dots & 1 \end{pmatrix}$$

- For any matrix $\mathbf{A}$, $\mathbf{A}\mathtt{I} = \mathbf{A}$.

Identity Operator

$$\mathtt{I} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ & & \dots & \\ 0 & 0 & \dots & 1 \end{pmatrix}$$

- For any matrix $\mathbf{A}$, $\mathbf{A}\mathtt{I} = \mathbf{A}$.
- $\mathtt{I}$ is the **identity operator** for the matrix product.

Column Space, Row Space and Rank

- Let $\mathbf{A}$ be an $m \times n$ matrix.

Column Space, Row Space and Rank

- Let $\mathbf{A}$ be an $m \times n$ matrix.
- **Column space** :

Column Space, Row Space and Rank

- Let $\mathbf{A}$ be an $m \times n$ matrix.
- **Column space** :
  - Span of the columns of $\mathbf{A}$.

Column Space, Row Space and Rank

- Let $\mathbf{A}$ be an $m \times n$ matrix.
- **Column space** :
  - Span of the columns of $\mathbf{A}$.
  - Linear subspace of $\mathbb{R}^m$.

Column Space, Row Space and Rank

- Let $\mathbf{A}$ be an $m \times n$ matrix.
- **Column space** :
  - Span of the columns of $\mathbf{A}$.
  - Linear subspace of $\mathbb{R}^m$.
- **Row space** :

Column Space, Row Space and Rank

- Let **A** be an $m \times n$ matrix.
- **Column space** :
  - Span of the columns of **A**.
  - Linear subspace of $\mathbb{R}^m$.
- **Row space** :
  - Span of the rows of **A**.

Column Space, Row Space and Rank

- Let $\mathbf{A}$ be an $m \times n$ matrix.
- **Column space** :
  - Span of the columns of $\mathbf{A}$.
  - Linear subspace of $\mathbb{R}^m$.
- **Row space** :
  - Span of the rows of $\mathbf{A}$.
  - Linear subspace of $\mathbb{R}^n$.

Column Space, Row Space and Rank

- Let $\mathbf{A}$ be an $m \times n$ matrix.
- **Column space** :
    - Span of the columns of $\mathbf{A}$.
    - Linear subspace of $\mathbb{R}^m$.
- **Row space** :
    - Span of the rows of $\mathbf{A}$.
    - Linear subspace of $\mathbb{R}^n$.
- **Important fact :**
  The column and row spaces of **any** matrix have **the same dimension**.

Column Space, Row Space and Rank

- Let $\mathbf{A}$ be an $m \times n$ matrix.
- **Column space** :
  - Span of the columns of $\mathbf{A}$.
  - Linear subspace of $\mathbb{R}^m$.
- **Row space** :
  - Span of the rows of $\mathbf{A}$.
  - Linear subspace of $\mathbb{R}^n$.
- **Important fact :**
  The column and row spaces of **any** matrix have **the same dimension**.
- This dimension is the **rank** of the matrix.

Range and Null space

- Let $\mathbf{A}$ be an $m \times n$ matrix.

Range and Null space

- Let **A** be an $m \times n$ matrix.
- **Range** :

Range and Null space

- Let $\mathbf{A}$ be an $m \times n$ matrix.
- **Range** :
    - Set of vectors equal to $\mathbf{Au}$ for some $\mathbf{u} \in \mathbb{R}^n$.

Range and Null space

- Let $\mathbf{A}$ be an $m \times n$ matrix.
- **Range** :
  - Set of vectors equal to $\mathbf{Au}$ for some $\mathbf{u} \in \mathbb{R}^n$.
  - Linear subspace of $\mathbb{R}^m$.

Range and Null space

- Let $\mathbf{A}$ be an $m \times n$ matrix.
- **Range** :
  - Set of vectors equal to $\mathbf{Au}$ for some $\mathbf{u} \in \mathbb{R}^n$.
  - Linear subspace of $\mathbb{R}^m$.
- **Null space** :

Range and Null space

- Let $\mathbf{A}$ be an $m \times n$ matrix.
- **Range** :
  - Set of vectors equal to $\mathbf{Au}$ for some $\mathbf{u} \in \mathbb{R}^n$.
  - Linear subspace of $\mathbb{R}^m$.
- **Null space** :
  - $\mathbf{u}$ belongs to the null space if $\mathbf{Au} = 0$.

Range and Null space

- Let $\mathbf{A}$ be an $m \times n$ matrix.
- **Range** :
  - Set of vectors equal to $\mathbf{Au}$ for some $\mathbf{u} \in \mathbb{R}^n$.
  - Linear subspace of $\mathbb{R}^m$.
- **Null space** :
  - $\mathbf{u}$ belongs to the null space if $\mathbf{Au} = 0$.
  - Subspace of $\mathbb{R}^n$.

Range and Null space

- Let $\mathbf{A}$ be an $m \times n$ matrix.
- **Range** :
  - Set of vectors equal to $\mathbf{Au}$ for some $\mathbf{u} \in \mathbb{R}^n$.
  - Linear subspace of $\mathbb{R}^m$.
- **Null space** :
  - $\mathbf{u}$ belongs to the null space if $\mathbf{Au} = 0$.
  - Subspace of $\mathbb{R}^n$.
  - Every vector in the null space is orthogonal to the rows of $\mathbf{A}$. The **null space** and **row space** of a matrix are **orthogonal**.

Range and Column Space

Another interpretation of the matrix vector product ($\mathbf{A_i}$ is column i of $\mathbf{A}$) :

$$\mathbf{Au} = \begin{pmatrix} \mathbf{A_1} & \mathbf{A_2} & \dots & \mathbf{A_n} \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ \dots \\ u_n \end{pmatrix}$$

$$= u_1 \mathbf{A_1} + u_2 \mathbf{A_2} + \cdots + u_n \mathbf{A_n}$$

Range and Column Space

Another interpretation of the matrix vector product ($\mathbf{A_i}$ is column i of $\mathbf{A}$) :

$$\mathbf{Au} = \begin{pmatrix} \mathbf{A_1} & \mathbf{A_2} & \dots & \mathbf{A_n} \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ \dots \\ u_n \end{pmatrix}$$

$$= u_1\mathbf{A_1} + u_2\mathbf{A_2} + \cdots + u_n\mathbf{A_n}$$

- The result is a **linear combination** of the columns of $\mathbf{A}$.

Range and Column Space

Another interpretation of the matrix vector product ($\mathbf{A_i}$ is column i of $\mathbf{A}$) :

$$\mathbf{Au} = \begin{pmatrix} \mathbf{A_1} & \mathbf{A_2} & \dots & \mathbf{A_n} \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ \dots \\ u_n \end{pmatrix}$$

$$= u_1\mathbf{A_1} + u_2\mathbf{A_2} + \cdots + u_n\mathbf{A_n}$$

- The result is a **linear combination** of the columns of $\mathbf{A}$.
- For any matrix, the **range** is **equal** to the **column space**.

Matrix Inverse

- For an $n \times n$ matrix $\mathbf{A}$: $\mathbf{rank} + \mathbf{dim(null\ space)} = \mathbf{n}$.

Matrix Inverse

- For an $n \times n$ matrix $\mathbf{A}$: **rank** $+$ **dim(null space)** $=$ **n**.
- If dim(null space)=0 then $\mathbf{A}$ is **full rank**.

Matrix Inverse

- For an $n \times n$ matrix $\mathbf{A}$: **rank** $+$ **dim(null space)** $=$ **n**.
- If dim(null space)=0 then $\mathbf{A}$ is **full rank**.
- In this case, the action of the matrix is invertible.

Matrix Inverse

- For an $n \times n$ matrix $\mathbf{A}$: $\mathbf{rank} + \mathbf{dim(null\ space)} = \mathbf{n}$.
- If dim(null space)=0 then $\mathbf{A}$ is **full rank**.
- In this case, the action of the matrix is invertible.
- The inversion is also linear and consequently can be represented by another matrix $\mathbf{A}^{-1}$.

Matrix Inverse

- For an $n \times n$ matrix $\mathbf{A}$: **rank** $+$ **dim(null space)** $=$ **n**.
- If dim(null space)=0 then $\mathbf{A}$ is **full rank**.
- In this case, the action of the matrix is invertible.
- The inversion is also linear and consequently can be represented by another matrix $\mathbf{A}^{-1}$.
- $\mathbf{A}^{-1}$ is the **only** matrix such that $\mathbf{A}^{-1}\mathbf{A} = \mathbf{A}\mathbf{A}^{-1} = \mathtt{I}$.

Orthogonal Matrices

- An **orthogonal** matrix $\mathbf{U}$ satisfies $\mathbf{U}^T\mathbf{U} = \mathtt{I}$.

Orthogonal Matrices

- An **orthogonal** matrix $\mathbf{U}$ satisfies $\mathbf{U}^T\mathbf{U} = \mathtt{I}$.
- Equivalently, $\mathbf{U}$ has **orthonormal columns**.

Orthogonal Matrices

- An **orthogonal** matrix $\mathbf{U}$ satisfies $\mathbf{U}^T\mathbf{U} = \mathtt{I}$.
- Equivalently, $\mathbf{U}$ has **orthonormal columns**.
- Applying an orthogonal matrix to two vectors **does not change** their inner product :

$$\begin{aligned} \langle \mathbf{Uu}, \mathbf{Uv} \rangle &= (\mathbf{Uu})^T \mathbf{Uv} \\ &= \mathbf{u}^T\mathbf{U}^T\mathbf{Uv} \\ &= \mathbf{u}^T\mathbf{v} \\ &= \langle \mathbf{u}, \mathbf{v} \rangle \end{aligned}$$

Orthogonal Matrices

- An **orthogonal** matrix $\mathbf{U}$ satisfies $\mathbf{U}^T\mathbf{U} = \mathtt{I}$.
- Equivalently, $\mathbf{U}$ has **orthonormal columns**.
- Applying an orthogonal matrix to two vectors **does not change** their inner product :

$$
\begin{aligned}
\langle \mathbf{Uu}, \mathbf{Uv} \rangle &= (\mathbf{Uu})^T \mathbf{Uv} \\
&= \mathbf{u}^T \mathbf{U}^T \mathbf{Uv} \\
&= \mathbf{u}^T \mathbf{v} \\
&= \langle \mathbf{u}, \mathbf{v} \rangle
\end{aligned}
$$

- Example : Matrices that represent **rotations** are orthogonal.

Trace and Determinant

- The **trace** is the sum of the diagonal elements of a square matrix.

Trace and Determinant

- The **trace** is the sum of the diagonal elements of a square matrix.
- The **determinant** of a square matrix $\mathbf{A}$ is denoted by $|\mathbf{A}|$.

Trace and Determinant

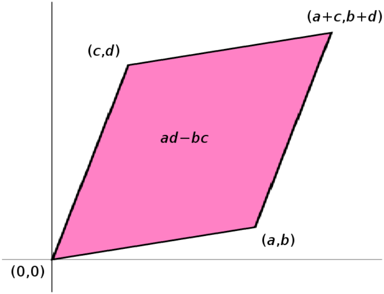
- The **trace** is the sum of the diagonal elements of a square matrix.
- The **determinant** of a square matrix $\mathbf{A}$ is denoted by $|\mathbf{A}|$.
- For a 2×2 matrix $\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$:

Trace and Determinant

- The **trace** is the sum of the diagonal elements of a square matrix.
- The **determinant** of a square matrix $\mathbf{A}$ is denoted by $|\mathbf{A}|$.
- For a 2×2 matrix $\mathbf{A} = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$:
  - $|\mathbf{A}| = ad - bc$

Trace and Determinant

- The **trace** is the sum of the diagonal elements of a square matrix.
- The **determinant** of a square matrix $\mathbf{A}$ is denoted by $|\mathbf{A}|$.
- For a 2×2 matrix $\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$:
    - $|\mathbf{A}| = ad - bc$
    - The absolute value of $|\mathbf{A}|$ is the area of the parallelogram given by the rows of $\mathbf{A}$.



Properties of the Determinant

- The definition can be generalized to larger matrices.

Properties of the Determinant

- The definition can be generalized to larger matrices.
- $|\mathbf{A}| = \left|\mathbf{A}^T\right|$.

Properties of the Determinant

- The definition can be generalized to larger matrices.
- $|\mathbf{A}| = \left|\mathbf{A}^T\right|$.
- $|\mathbf{AB}| = |\mathbf{A}|\,|\mathbf{B}|$

Properties of the Determinant

- The definition can be generalized to larger matrices.
- $|\mathbf{A}| = \left|\mathbf{A}^T\right|$.
- $|\mathbf{AB}| = |\mathbf{A}|\,|\mathbf{B}|$
- $|\mathbf{A}| = 0$ if and only if $\mathbf{A}$ is **not invertible**.

Properties of the Determinant

- The definition can be generalized to larger matrices.
- $|\mathbf{A}| = \left|\mathbf{A}^T\right|$.
- $|\mathbf{AB}| = |\mathbf{A}|\,|\mathbf{B}|$
- $|\mathbf{A}| = 0$ if and only if $\mathbf{A}$ is **not invertible**.
- If $\mathbf{A}$ is invertible, then $\left|\mathbf{A}^{-1}\right| = \frac{1}{|\mathbf{A}|}$.

Eigenvalues and Eigenvectors

- An **eigenvalue** λ of a square matrix $\mathbf{A}$ satisfies :

$$\mathbf{A}\mathbf{u} = \lambda\mathbf{u}$$

  for some vector $\mathbf{u}$, which we call an **eigenvector**.

Eigenvalues and Eigenvectors

- An **eigenvalue** λ of a square matrix **A** satisfies :

$$\mathbf{A}\mathbf{u} = \lambda\mathbf{u}$$

  for some vector **u**, which we call an **eigenvector**.
- Geometrically, the operator **A expands** $(\lambda > 1)$ or **contracts** $(\lambda < 1)$ eigenvectors but **does not rotate** them.

Eigenvalues and Eigenvectors

- An **eigenvalue** λ of a square matrix $\mathbf{A}$ satisfies :

$$\mathbf{Au} = \lambda\mathbf{u}$$

  for some vector $\mathbf{u}$, which we call an **eigenvector**.
- Geometrically, the operator $\mathbf{A}$ **expands** ($\lambda > 1$) or **contracts** ($\lambda < 1$) eigenvectors but **does not rotate** them.
- If $\mathbf{u}$ is an eigenvector of $\mathbf{A}$, it is in the null space of $\mathbf{A} - \lambda\mathtt{I}$, which is consequently **not invertible**.

Eigenvalues and Eigenvectors

- An **eigenvalue** λ of a square matrix $\mathbf{A}$ satisfies :

  $$\mathbf{A}\mathbf{u} = \lambda\mathbf{u}$$

  for some vector $\mathbf{u}$, which we call an **eigenvector**.
- Geometrically, the operator $\mathbf{A}$ **expands** ($\lambda > 1$) or **contracts** ($\lambda < 1$) eigenvectors but **does not rotate** them.
- If $\mathbf{u}$ is an eigenvector of $\mathbf{A}$, it is in the null space of $\mathbf{A} - \lambda\mathtt{I}$, which is consequently **not invertible**.
- The eigenvalues of $\mathbf{A}$ are the roots of the equation $|\mathbf{A} - \lambda\mathtt{I}| = 0$.

Eigenvalues and Eigenvectors

- An **eigenvalue** λ of a square matrix $\mathbf{A}$ satisfies :

$$\mathbf{A}\mathbf{u} = \lambda \mathbf{u}$$

  for some vector $\mathbf{u}$, which we call an **eigenvector**.
- Geometrically, the operator $\mathbf{A}$ **expands** ($\lambda > 1$) or **contracts** ($\lambda < 1$) eigenvectors but **does not rotate** them.
- If $\mathbf{u}$ is an eigenvector of $\mathbf{A}$, it is in the null space of $\mathbf{A} - \lambda \mathtt{I}$, which is consequently **not invertible**.
- The eigenvalues of $\mathbf{A}$ are the roots of the equation $|\mathbf{A} - \lambda \mathtt{I}| = 0$.
- **Not** the way eigenvalues are calculated in practice.

Eigenvalues and Eigenvectors

- An **eigenvalue** λ of a square matrix $\mathbf{A}$ satisfies :

$$\mathbf{A}\mathbf{u} = \lambda\mathbf{u}$$

  for some vector $\mathbf{u}$, which we call an **eigenvector**.
- Geometrically, the operator $\mathbf{A}$ **expands** ($\lambda > 1$) or **contracts** ($\lambda < 1$) eigenvectors but **does not rotate** them.
- If $\mathbf{u}$ is an eigenvector of $\mathbf{A}$, it is in the null space of $\mathbf{A} - \lambda\mathtt{I}$, which is consequently **not invertible**.
- The eigenvalues of $\mathbf{A}$ are the roots of the equation $|\mathbf{A} - \lambda\mathtt{I}| = 0$.
- **Not** the way eigenvalues are calculated in practice.
- Eigenvalues and eigenvectors can be **complex valued**, even if all the entries of $\mathbf{A}$ are real.

Eigendecomposition of a Matrix

- Let $\mathbf{A}$ be an $n \times n$ square matrix with n linearly independent eigenvectors $\mathbf{p}_1 \ldots \mathbf{p}_n$ and eigenvalues $\lambda_1 \ldots \lambda_n$.

Eigendecomposition of a Matrix

- Let $\mathbf{A}$ be an $n \times n$ square matrix with n linearly independent eigenvectors $\mathbf{p}_1 \dots \mathbf{p}_n$ and eigenvalues $\lambda_1 \dots \lambda_n$.
- We define the matrices :

$$\mathbf{P} = \begin{pmatrix} \mathbf{p}_1 & \dots & \mathbf{p}_n \end{pmatrix}$$

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ & & \dots & \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}$$

Eigendecomposition of a Matrix

- Let $\mathbf{A}$ be an $n \times n$ square matrix with n linearly independent eigenvectors $\mathbf{p}_1 \dots \mathbf{p}_n$ and eigenvalues $\lambda_1 \dots \lambda_n$.
- We define the matrices :

$$\mathbf{P} = \begin{pmatrix} \mathbf{p}_1 & \dots & \mathbf{p}_n \end{pmatrix}$$

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ & & \dots & \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}$$

- $\mathbf{A}$ satisfies $\mathbf{AP} = \mathbf{P}\Lambda$.

Eigendecomposition of a Matrix

- Let $\mathbf{A}$ be an $n \times n$ square matrix with n linearly independent eigenvectors $\mathbf{p}_1 \dots \mathbf{p}_n$ and eigenvalues $\lambda_1 \dots \lambda_n$.
- We define the matrices :

$$\mathbf{P} = \begin{pmatrix} \mathbf{p}_1 & \dots & \mathbf{p}_n \end{pmatrix}$$

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ & & \dots & \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}$$

- $\mathbf{A}$ satisfies $\mathbf{AP} = \mathbf{P}\Lambda$.
- $\mathbf{P}$ is full rank. Inverting it yields the **eigendecomposition** of $\mathbf{A}$:

$$\mathbf{A} = \mathbf{P}\Lambda\mathbf{P}^{-1}$$

Properties of the Eigendecomposition

- **Not all** matrices are **diagonalizable** (have an eigendecomposition). Example : $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$.

Properties of the Eigendecomposition

- **Not all** matrices are **diagonalizable** (have an eigendecomposition). Example : $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$.
- $\text{Trace}(\mathbf{A}) = \text{Trace}(\Lambda) = \sum_{i=1}^{n} \lambda_i$.

Properties of the Eigendecomposition

- **Not all** matrices are **diagonalizable** (have an eigendecomposition). Example : $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$.
- $\text{Trace}(\mathbf{A}) = \text{Trace}(\Lambda) = \sum_{i=1}^{n} \lambda_i$.
- $|\mathbf{A}| = |\Lambda| = \prod_{i=1}^{n} \lambda_i$.

Properties of the Eigendecomposition

- **Not all** matrices are **diagonalizable** (have an eigendecomposition). Example : $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$.
- $\text{Trace}(\mathbf{A}) = \text{Trace}(\Lambda) = \sum_{i=1}^{n} \lambda_i$.
- $|\mathbf{A}| = |\Lambda| = \prod_{i=1}^{n} \lambda_i$.
- The **rank** of **A** is equal to the **number of nonzero eigenvalues**.

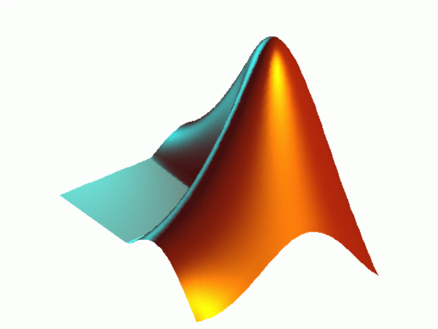
Properties of the Eigendecomposition

- **Not all** matrices are **diagonalizable** (have an eigendecomposition). Example : $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$.
- $\text{Trace}(\mathbf{A}) = \text{Trace}(\Lambda) = \sum_{i=1}^{n} \lambda_i$.
- $|\mathbf{A}| = |\Lambda| = \prod_{i=1}^{n} \lambda_i$.
- The **rank** of **A** is equal to the **number of nonzero eigenvalues**.
- If λ is a nonzero eigenvalue of **A**, $\frac{1}{\lambda}$ is an eigenvalue of $\mathbf{A}^{-1}$ with the **same eigenvector**.

Properties of the Eigendecomposition

- **Not all** matrices are **diagonalizable** (have an eigendecomposition). Example : $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$.
- $\text{Trace}(\mathbf{A}) = \text{Trace}(\Lambda) = \sum_{i=1}^{n} \lambda_i$.
- $|\mathbf{A}| = |\Lambda| = \prod_{i=1}^{n} \lambda_i$.
- The **rank** of **A** is equal to the **number of nonzero eigenvalues**.
- If λ is a nonzero eigenvalue of $\mathbf{A}$, $\frac{1}{\lambda}$ is an eigenvalue of $\mathbf{A}^{-1}$ with the **same eigenvector**.
- The eigendecomposition makes it possible to compute **matrix powers** efficiently :
  $\mathbf{A}^m = \left(\mathbf{P}\Lambda\mathbf{P}^{-1}\right)^m = \mathbf{P}\Lambda\mathbf{P}^{-1}\ \mathbf{P}\Lambda\mathbf{P}^{-1}\ \ldots\mathbf{P}\Lambda\mathbf{P}^{-1} = \mathbf{P}\Lambda^m\mathbf{P}^{-1}$.

Matlab Example



Eigendecomposition of a Symmetric Matrix

- If $\mathbf{A} = \mathbf{A}^T$ then $\mathbf{A}$ is **symmetric**.

Eigendecomposition of a Symmetric Matrix

- If $\mathbf{A} = \mathbf{A}^T$ then $\mathbf{A}$ is **symmetric**.
- The eigenvalues of symmetric matrices are **real**.

Eigendecomposition of a Symmetric Matrix

- If $\mathbf{A} = \mathbf{A}^T$ then $\mathbf{A}$ is **symmetric**.
- The eigenvalues of symmetric matrices are **real**.
- The eigenvectors of symmetric matrices are **orthonormal**.

Eigendecomposition of a Symmetric Matrix

- If $\mathbf{A} = \mathbf{A}^T$ then $\mathbf{A}$ is **symmetric**.
- The eigenvalues of symmetric matrices are **real**.
- The eigenvectors of symmetric matrices are **orthonormal**.
- Consequently, the eigendecomposition becomes :
  $\mathbf{A} = \mathbf{U}\Lambda\mathbf{U}^T$ for Λ real and $\mathbf{U}$ orthogonal.

Eigendecomposition of a Symmetric Matrix

- If $\mathbf{A} = \mathbf{A}^T$ then $\mathbf{A}$ is **symmetric**.
- The eigenvalues of symmetric matrices are **real**.
- The eigenvectors of symmetric matrices are **orthonormal**.
- Consequently, the eigendecomposition becomes : $\mathbf{A} = \mathbf{U}\Lambda\mathbf{U}^T$ for Λ real and $\mathbf{U}$ orthogonal.
- The eigenvectors of $\mathbf{A}$ are an orthonormal **basis** for the **column space** (and the row space, since they are equal).

Eigendecomposition of a Symmetric Matrix

- The action of a symmetric matrix on a vector $\mathbf{Au} = \mathbf{U}\Lambda\mathbf{U}^T\mathbf{u}$ can be decomposed into :

Eigendecomposition of a Symmetric Matrix

- The action of a symmetric matrix on a vector $\mathbf{Au} = \mathbf{U}\Lambda\mathbf{U}^T\mathbf{u}$ can be decomposed into :
  1. Projection of $\mathbf{u}$ onto the column space of $\mathbf{A}$ (multiplication by $\mathbf{U}^T$).

Eigendecomposition of a Symmetric Matrix

- The action of a symmetric matrix on a vector $\mathbf{Au} = \mathbf{U}\Lambda\mathbf{U}^T\mathbf{u}$ can be decomposed into :
  1. Projection of $\mathbf{u}$ onto the column space of $\mathbf{A}$ (multiplication by $\mathbf{U}^T$).
  2. Scaling of each coefficient $\langle \mathbf{U}_i, \mathbf{u} \rangle$ by the corresponding eigenvalue (multiplication by Λ).

Eigendecomposition of a Symmetric Matrix

- The action of a symmetric matrix on a vector $\mathbf{Au} = \mathbf{U}\Lambda\mathbf{U}^T\mathbf{u}$ can be decomposed into :
  1. Projection of $\mathbf{u}$ onto the column space of $\mathbf{A}$ (multiplication by $\mathbf{U}^T$).
  2. Scaling of each coefficient $\langle \mathbf{U}_i, \mathbf{u} \rangle$ by the corresponding eigenvalue (multiplication by Λ).
  3. Linear combination of the eigenvectors scaled by the resulting coefficient (multiplication by $\mathbf{U}$).

Eigendecomposition of a Symmetric Matrix

- The action of a symmetric matrix on a vector $\mathbf{Au} = \mathbf{U}\Lambda\mathbf{U}^T\mathbf{u}$ can be decomposed into :
  1. Projection of $\mathbf{u}$ onto the column space of $\mathbf{A}$ (multiplication by $\mathbf{U}^T$).
  2. Scaling of each coefficient $\langle \mathbf{U}_i, \mathbf{u} \rangle$ by the corresponding eigenvalue (multiplication by Λ).
  3. Linear combination of the eigenvectors scaled by the resulting coefficient (multiplication by $\mathbf{U}$).
- Summarizing :

$$\mathbf{Au} = \sum_{i=1}^{n} \lambda_i \langle \mathbf{U}_i, \mathbf{u} \rangle \mathbf{U}_i$$

Eigendecomposition of a Symmetric Matrix

- The action of a symmetric matrix on a vector $\mathbf{Au} = \mathbf{U}\Lambda\mathbf{U}^T\mathbf{u}$ can be decomposed into :
  1. Projection of $\mathbf{u}$ onto the column space of $\mathbf{A}$ (multiplication by $\mathbf{U}^T$).
  2. Scaling of each coefficient $\langle \mathbf{U}_i, \mathbf{u} \rangle$ by the corresponding eigenvalue (multiplication by Λ).
  3. Linear combination of the eigenvectors scaled by the resulting coefficient (multiplication by $\mathbf{U}$).
- Summarizing :

$$\mathbf{Au} = \sum_{i=1}^{n} \lambda_i \langle \mathbf{U}_i, \mathbf{u} \rangle \mathbf{U}_i$$

- It would be great to generalize this to all matrices...

Singular Value Decomposition

- **Every** matrix has a **singular value decomposition** :

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

Singular Value Decomposition

- **Every** matrix has a **singular value decomposition** :

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

- The columns of $\mathbf{U}$ are an orthonormal basis for the **column space** of $\mathbf{A}$.

Singular Value Decomposition

- **Every** matrix has a **singular value decomposition** :

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

- The columns of **U** are an orthonormal basis for the **column space** of **A**.
- The columns of **V** are an orthonormal basis for the **row space** of **A**.

Singular Value Decomposition

- **Every** matrix has a **singular value decomposition** :

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

- The columns of $\mathbf{U}$ are an orthonormal basis for the **column space** of $\mathbf{A}$.
- The columns of $\mathbf{V}$ are an orthonormal basis for the **row space** of $\mathbf{A}$.
- Σ is diagonal. The entries on its diagonal $\sigma_i = \Sigma_{ii}$ are positive real numbers, called the **singular values** of $\mathbf{A}$.

Singular Value Decomposition

- **Every** matrix has a **singular value decomposition** :

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

- The columns of $\mathbf{U}$ are an orthonormal basis for the **column space** of $\mathbf{A}$.
- The columns of $\mathbf{V}$ are an orthonormal basis for the **row space** of $\mathbf{A}$.
- Σ is diagonal. The entries on its diagonal $\sigma_i = \Sigma_{ii}$ are positive real numbers, called the **singular values** of $\mathbf{A}$.
- The action of $\mathbf{A}$ on a vector $\mathbf{u}$ can be decomposed into :

$$\mathbf{Au} = \sum_{i=1}^{n} \sigma_i \langle \mathbf{V}_i, \mathbf{u} \rangle \mathbf{U}_i$$

Properties of the Singular Value Decomposition

- The eigenvalues of the symmetric matrix $\mathbf{A}^T\mathbf{A}$ are equal to the **square** of the singular values of $\mathbf{A}$:

$$\mathbf{A}^T\mathbf{A} = \mathbf{V}\Sigma\mathbf{U}^T\mathbf{U}\Sigma\mathbf{V}^T = \mathbf{V}\Sigma^2\mathbf{V}^T$$

Properties of the Singular Value Decomposition

- The eigenvalues of the symmetric matrix $\mathbf{A}^T\mathbf{A}$ are equal to the **square** of the singular values of $\mathbf{A}$:

$$\mathbf{A}^T\mathbf{A} = \mathbf{V}\Sigma\mathbf{U}^T\mathbf{U}\Sigma\mathbf{V}^T = \mathbf{V}\Sigma^2\mathbf{V}^T$$

- The **rank** of a matrix is equal to the number of nonzero singular values.

Properties of the Singular Value Decomposition

- The eigenvalues of the symmetric matrix $\mathbf{A}^T\mathbf{A}$ are equal to the **square** of the singular values of $\mathbf{A}$:

$$\mathbf{A}^T\mathbf{A} = \mathbf{V}\Sigma\mathbf{U}^T\mathbf{U}\Sigma\mathbf{V}^T = \mathbf{V}\Sigma^2\mathbf{V}^T$$

- The **rank** of a matrix is equal to the number of nonzero singular values.
- The largest singular value σ_1 is the solution to the optimization problem :

$$\sigma_1 = \max_{\mathbf{x}\neq 0} \frac{||\mathbf{A}\mathbf{x}||_2}{||\mathbf{x}||_2}$$

Properties of the Singular Value Decomposition

- The eigenvalues of the symmetric matrix $\mathbf{A}^T\mathbf{A}$ are equal to the **square** of the singular values of $\mathbf{A}$:

$$\mathbf{A}^T\mathbf{A} = \mathbf{V}\Sigma\mathbf{U}^T\mathbf{U}\Sigma\mathbf{V}^T = \mathbf{V}\Sigma^2\mathbf{V}^T$$

- The **rank** of a matrix is equal to the number of nonzero singular values.
- The largest singular value σ_1 is the solution to the optimization problem :

$$\sigma_1 = \max_{\mathbf{x}\neq 0} \frac{||\mathbf{Ax}||_2}{||\mathbf{x}||_2}$$

- It can be verified that the largest singular value satisfies the properties of a norm, it is called the **spectral norm** of the matrix.

Properties of the Singular Value Decomposition

- The eigenvalues of the symmetric matrix $\mathbf{A}^T\mathbf{A}$ are equal to the **square** of the singular values of $\mathbf{A}$:

$$\mathbf{A}^T\mathbf{A} = \mathbf{V}\Sigma\mathbf{U}^T\mathbf{U}\Sigma\mathbf{V}^T = \mathbf{V}\Sigma^2\mathbf{V}^T$$

- The **rank** of a matrix is equal to the number of nonzero singular values.
- The largest singular value σ_1 is the solution to the optimization problem :

$$\sigma_1 = \max_{\mathbf{x} \neq 0} \frac{||\mathbf{A}\mathbf{x}||_2}{||\mathbf{x}||_2}$$

- It can be verified that the largest singular value satisfies the properties of a norm, it is called the **spectral norm** of the matrix.
- In statistics analyzing data with the singular value decomposition is called **Principal Component Analysis**.

Compression using the Singular Value Decomposition

- The singular value decomposition can be viewed as a sum of rank 1 matrices :

$$\mathbf{A} = \sigma_1 \mathbf{u}_1 \mathbf{v}_1^T + \sigma_2 \mathbf{u}_2 \mathbf{v}_2^T \cdots + \sigma_R \mathbf{u}_R \mathbf{v}_R^T$$

Compression using the Singular Value Decomposition

- The singular value decomposition can be viewed as a sum of rank 1 matrices :

$$\mathbf{A} = \sigma_1 \mathbf{u}_1 \mathbf{v}_1^T + \sigma_2 \mathbf{u}_2 \mathbf{v}_2^T \cdots + \sigma_R \mathbf{u}_R \mathbf{v}_R^T$$

- If some of the singular values are very small, we can **discard** them.

Compression using the Singular Value Decomposition

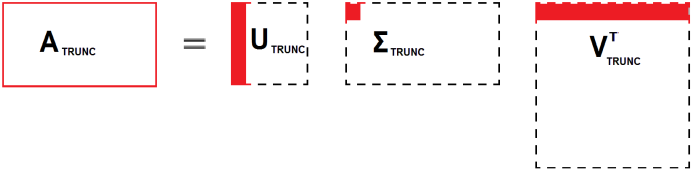
- The singular value decomposition can be viewed as a sum of rank 1 matrices :

$$\mathbf{A} = \sigma_1 \mathbf{u}_1 \mathbf{v}_1^T + \sigma_2 \mathbf{u}_2 \mathbf{v}_2^T \cdots + \sigma_R \mathbf{u}_R \mathbf{v}_R^T$$

- If some of the singular values are very small, we can **discard** them.
- This is a form of **lossy compression**.

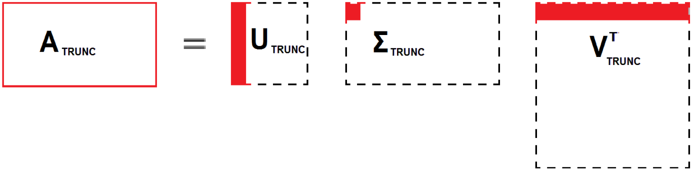
Compression using the Singular Value Decomposition

- Truncating the singular value decomposition allows us to represent the matrix with **less parameters**.



Compression using the Singular Value Decomposition

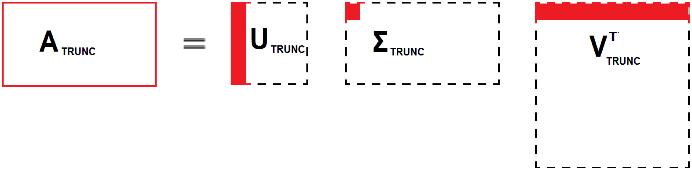
- Truncating the singular value decomposition allows us to represent the matrix with **less parameters**.



- For a 512×512 matrix :

Compression using the Singular Value Decomposition

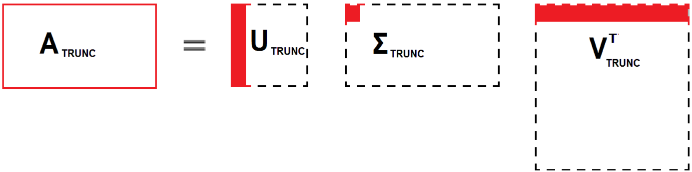
- Truncating the singular value decomposition allows us to represent the matrix with **less parameters**.



- For a 512×512 matrix :
  - Full representation : $512 \times 512 =$**262 144**

Compression using the Singular Value Decomposition

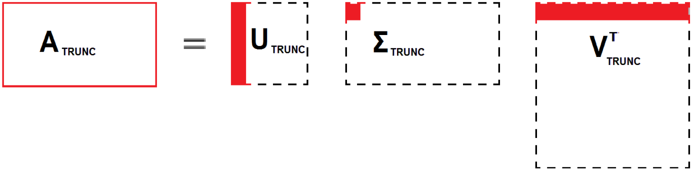
- Truncating the singular value decomposition allows us to represent the matrix with **less parameters**.



- For a 512×512 matrix :
  - Full representation : $512 \times 512 =$**262 144**
  - Rank 10 approximation : $512 \times 10 + 10 + 512 \times 10 =$**10 250**

Compression using the Singular Value Decomposition

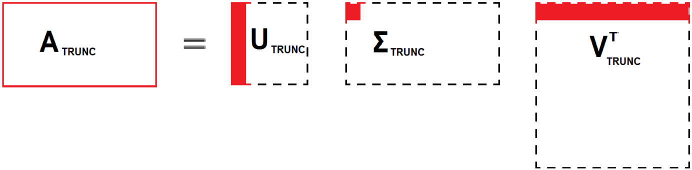
- Truncating the singular value decomposition allows us to represent the matrix with **less parameters**.



- For a 512×512 matrix :
  - Full representation : $512 \times 512 =$**262 144**
  - Rank 10 approximation : $512 \times 10 + 10 + 512 \times 10 =$**10 250**
  - Rank 40 approximation : $512 \times 40 + 40 + 512 \times 40 =$**41 000**

Compression using the Singular Value Decomposition

- Truncating the singular value decomposition allows us to represent the matrix with **less parameters**.



- For a 512×512 matrix :
  - Full representation : $512 \times 512 =$**262 144**
  - Rank 10 approximation : $512 \times 10 + 10 + 512 \times 10 =$**10 250**
  - Rank 40 approximation : $512 \times 40 + 40 + 512 \times 40 =$**41 000**
  - Rank 80 approximation : $512 \times 80 + 80 + 512 \times 80 =$**82 000**