Singular Value Decomposition (SVD)

Erisa Terolli

Talk @ Columbia University

15 February 2024

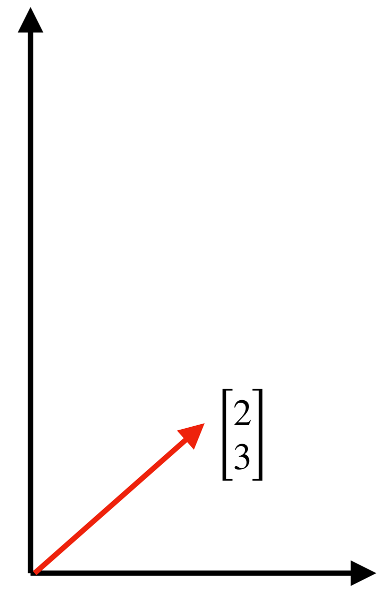
Applications in Data Science

Company/Industry	Application
Netflix	Personalized Content Recommendations
Google	PageRank Algorithm for Web Page Ranking
Facebook	Social Network Analysis
Amazon	Collaborative Filtering for Product Recommendations
Medical Imaging	MRI Reconstruction
Weather Forecasting	Climate Modeling
E-commerce	Customer Segmentation

2D Transformation Matrix

What happens when we multiply a 2D matrix with a vector?

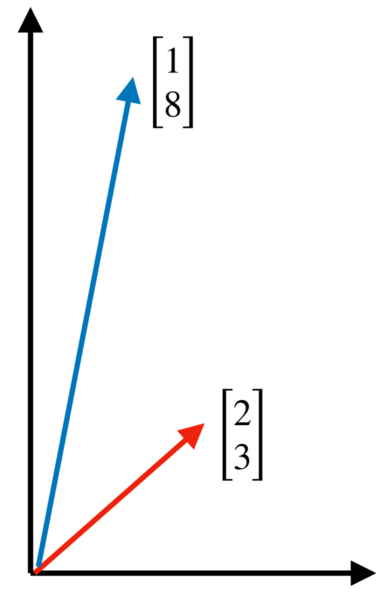
$$\begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \end{bmatrix} = ?$$



2D Transformation Matrix

What happens when we multiply a 2D matrix with a vector?

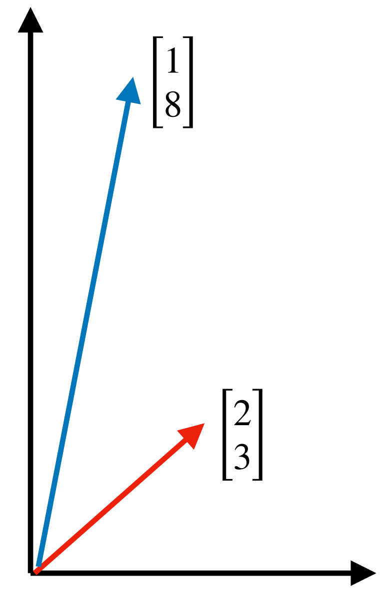
$$\begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 \\ 8 \end{bmatrix}$$



2D Transformation Matrix

What happens when we multiply a 2D matrix with a vector?

$$\begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 \\ 8 \end{bmatrix}$$



The matrix will transform the vector by rotating and stretching/shortening it

$$R = \begin{bmatrix} cos\theta & -sin\theta \\ sin\theta & cos\theta \end{bmatrix}$$

Rotation by θ

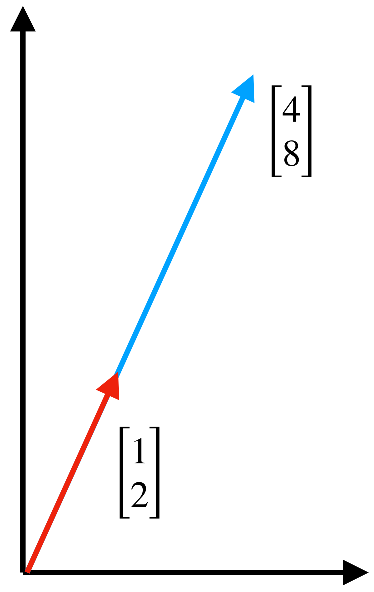
$$S = \begin{bmatrix} \alpha & 0 \\ 0 & \alpha \end{bmatrix}$$

Stretching by α

2D Transformation Matrix

What happens when we multiply a 2D matrix with a vector?

$$\begin{bmatrix} 2 & 1 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 4 \\ 8 \end{bmatrix}$$

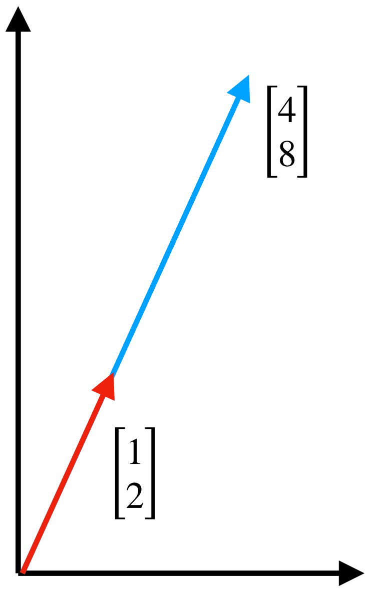


2D Transformation Matrix

What happens when we multiply a 2D matrix with a vector?

$$\begin{bmatrix} 2 & 1 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 4 \\ 8 \end{bmatrix}$$

$$4 \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 4 \\ 8 \end{bmatrix}$$

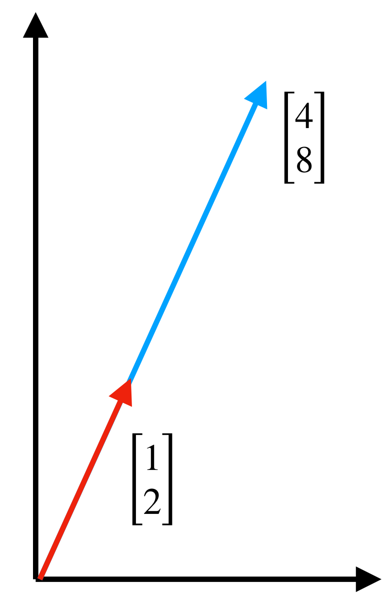


2D Transformation Matrix

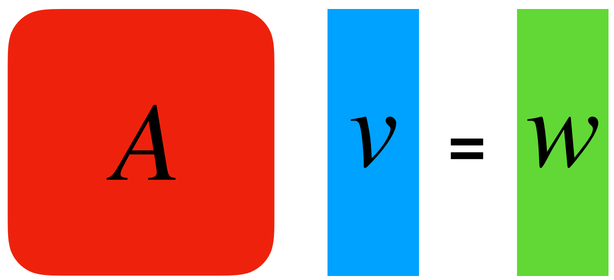
What happens when we multiply a 2D matrix with a vector?

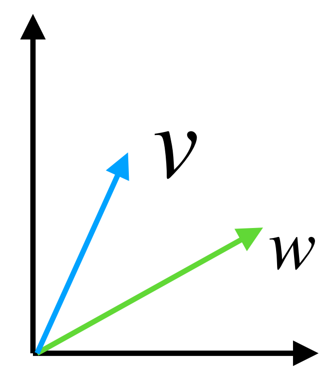
$$\begin{bmatrix} 2 & 1 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 4 \\ 8 \end{bmatrix}$$

$$\underbrace{4}_{\text{Eigen value}} \underbrace{\begin{bmatrix} 1 \\ 2 \end{bmatrix}}_{\text{Eigen vector}} = \begin{bmatrix} 4 \\ 8 \end{bmatrix}$$



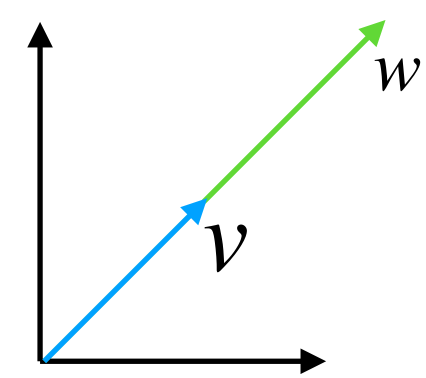
Eigenvectors & Eigenvalues





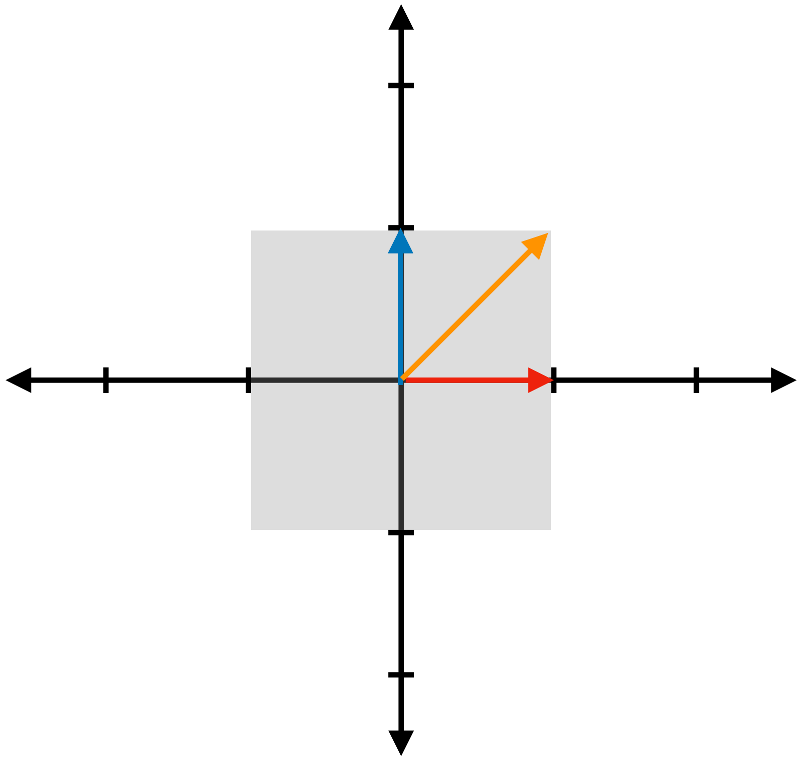
Transformation matrix A is applied to a vector v and outputs a vector w. If w points in the same direction as v (a.k.a. lies on the same 1 dimensional subspace), then v is an eigenvector of matrix A .

$$Av = w$$
$$\lambda v = w$$
$$Av = \lambda v$$



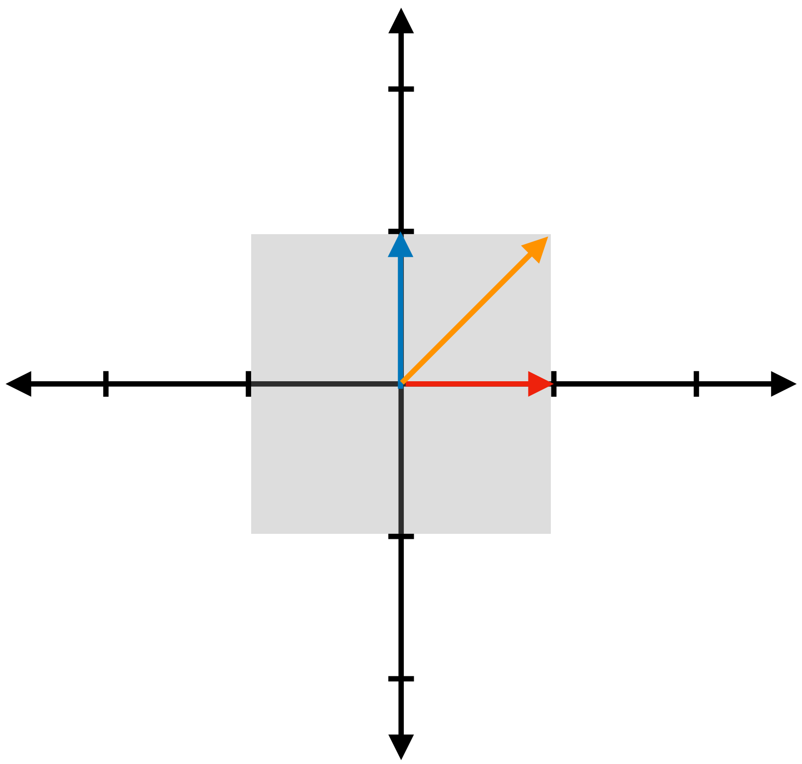
λ is an eigenvalue associated with eigenvector v of A .

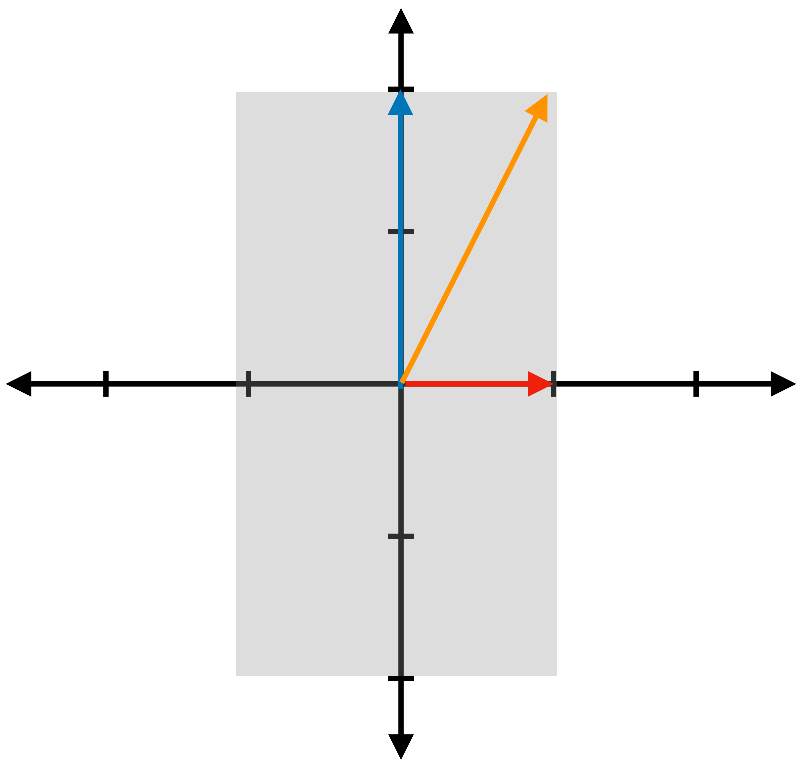
Eigenvectors & Eigenvalues



$$A = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$$

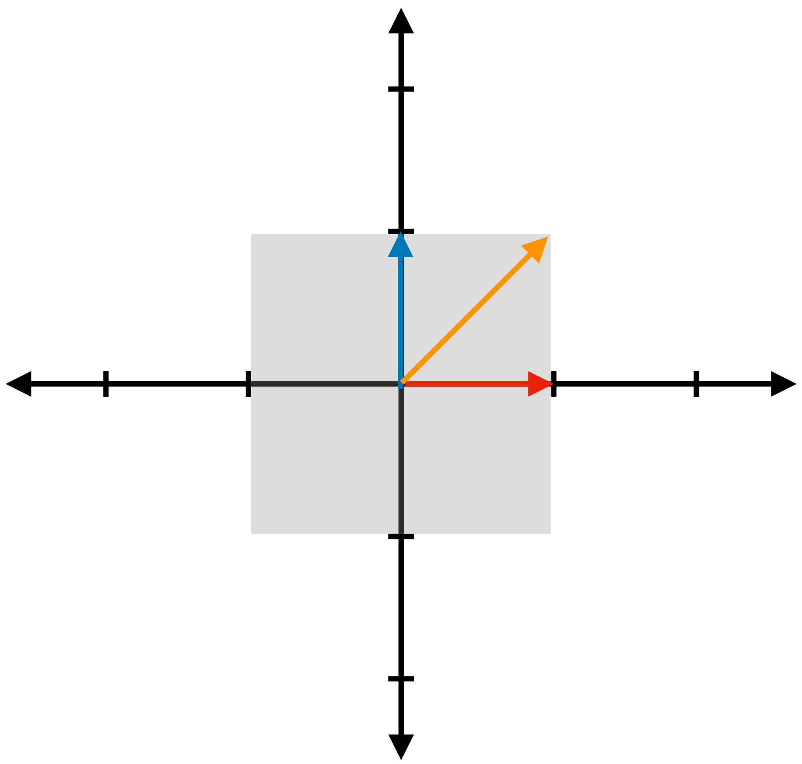
Eigenvectors & Eigenvalues

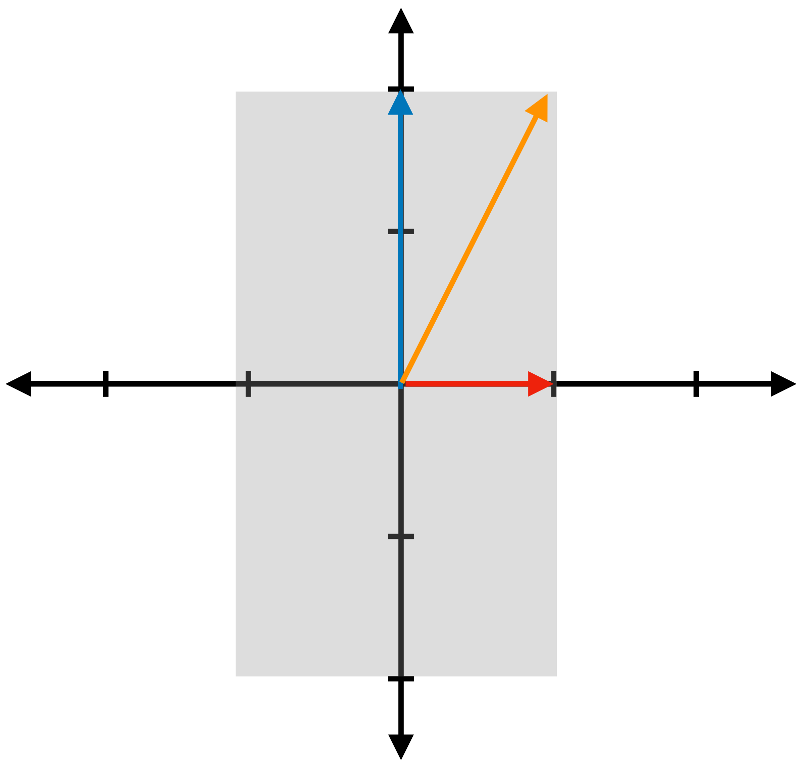




$$A = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$$

Eigenvectors & Eigenvalues

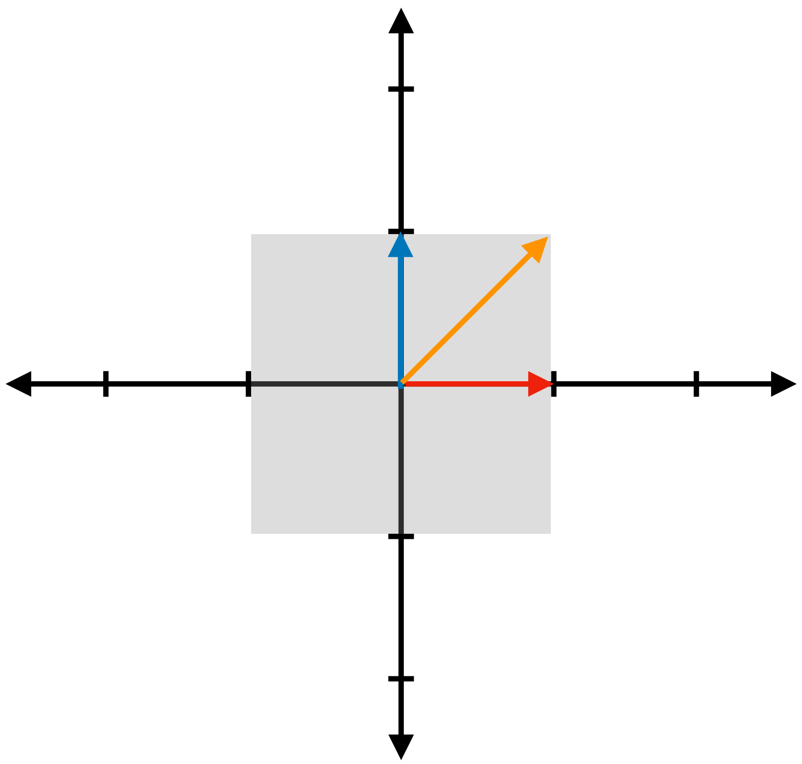


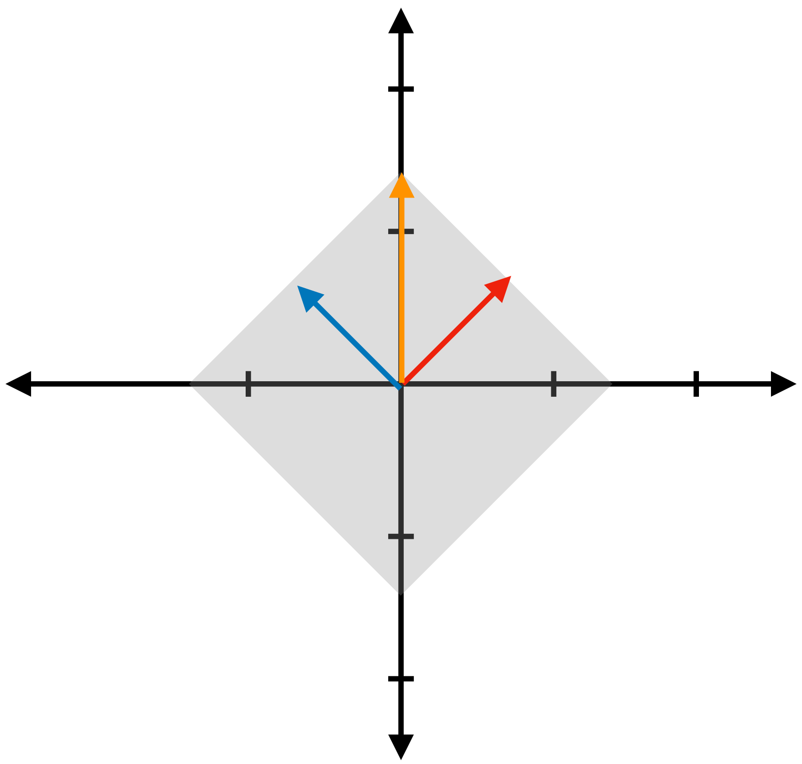


$$A = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$$

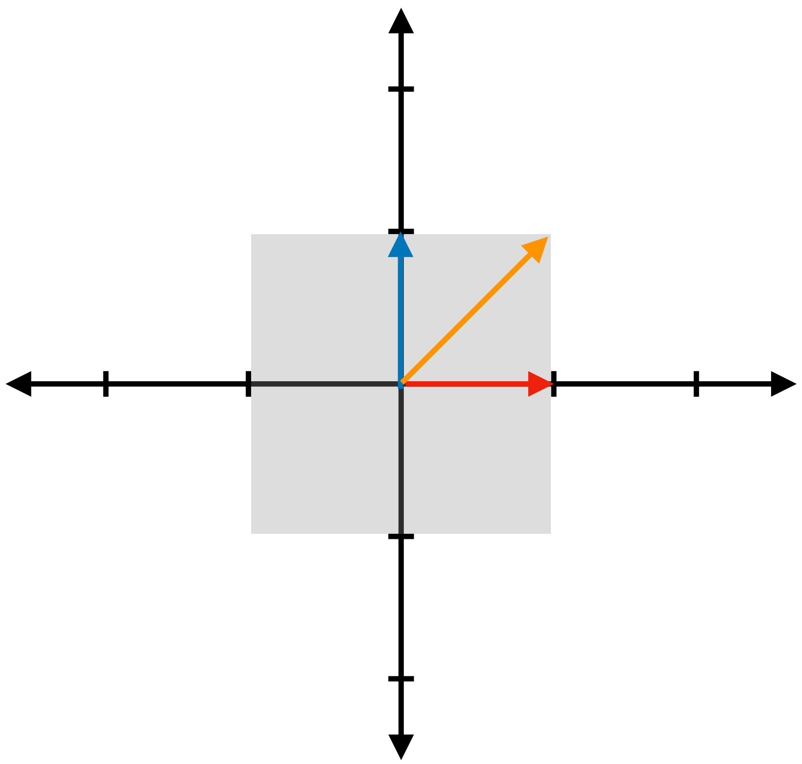
2 eigenvectors with values 1 and 2

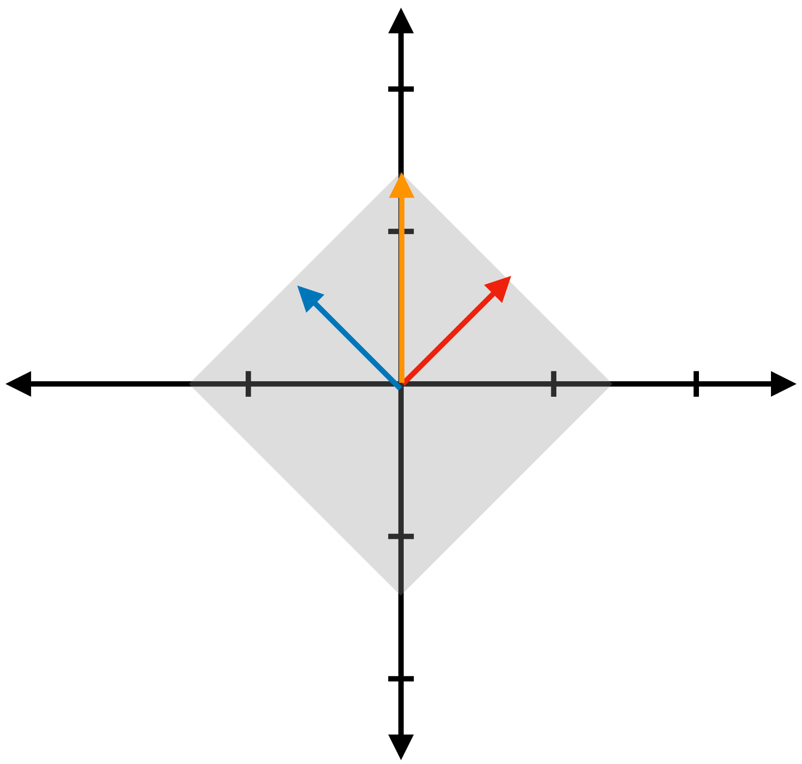
Eigenvectors & Eigenvalues





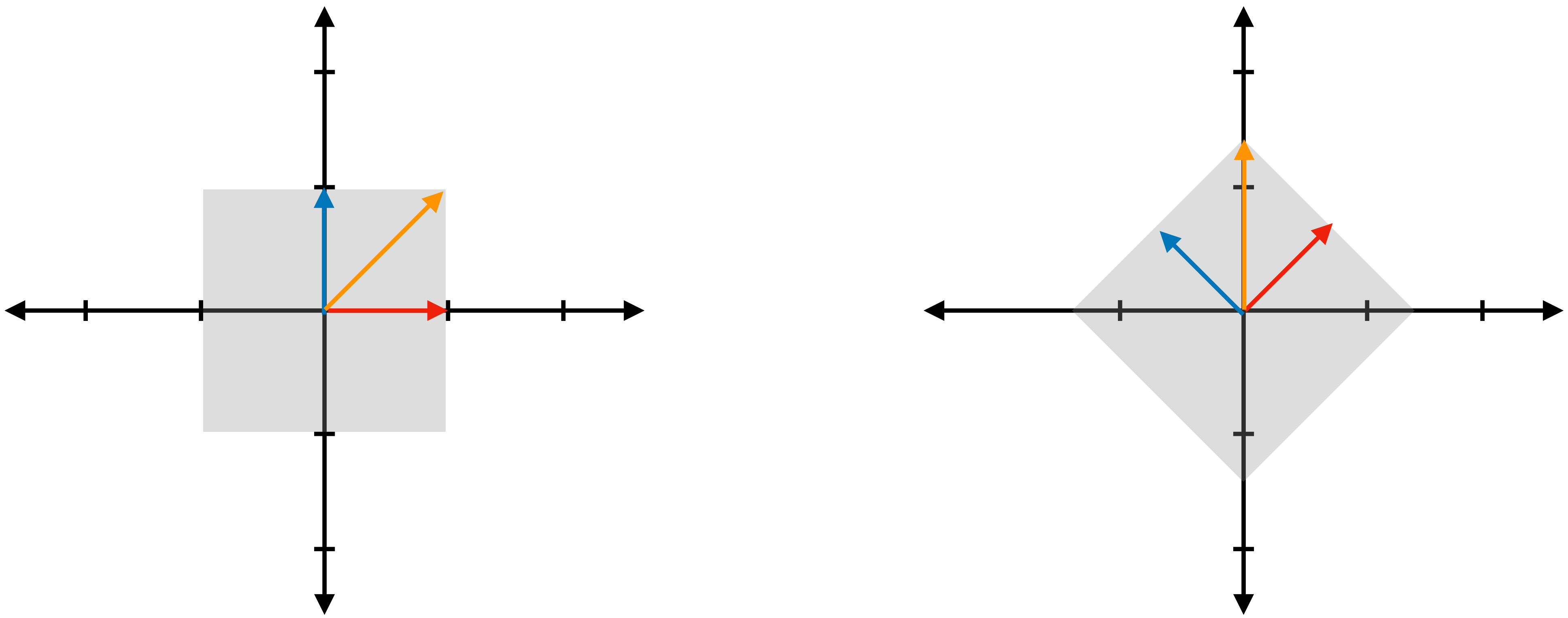
Eigenvectors & Eigenvalues





0 Eigenvectors

Eigenvectors & Eigenvalues



0 Eigenvectors

Check the vectors that lie on the same span after transformation and measure how much they magnitude changes

Finding Eigenvalues

$$Av = \lambda v$$

$$Av - \lambda v = 0$$

$$(A - \lambda)v = 0$$

Shift the matrix A by λ

$$(A - \lambda I)v = 0$$

It has nontrivial null space.
It must be singular.

$$A - \lambda I$$

$$det(A - \lambda I) = 0$$

The Determinant of singular matrices is 0.

Finding Eigenvectors

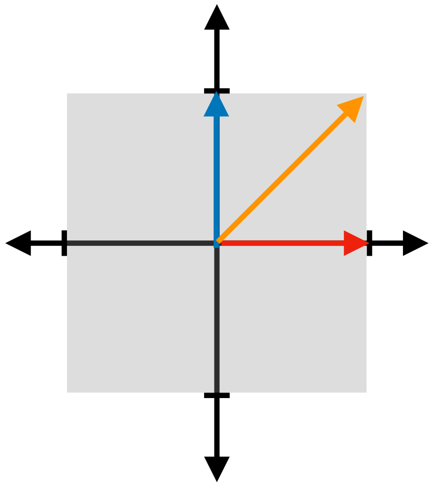
1. Find all eigenvalues λ by solving $det(A - \lambda I) = 0$.

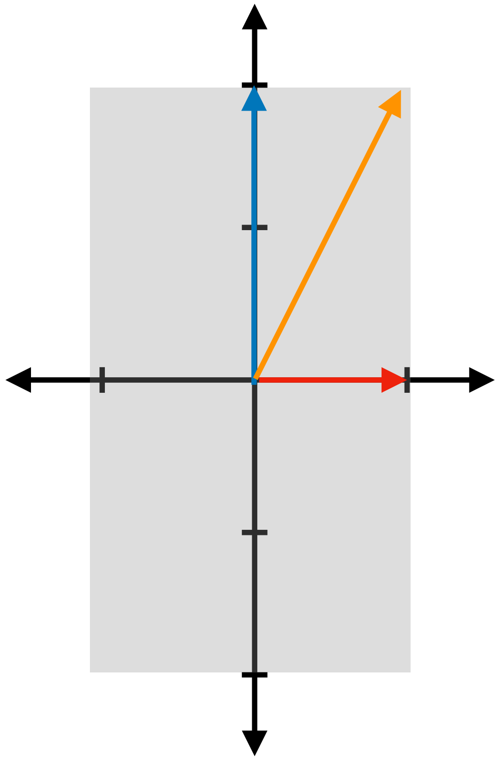
2. For each λ, find $v \in N(A - \lambda I)$.

   - Find a vector v that is in the null space of the matrix $(A - \lambda I)$.

EXAMPLE

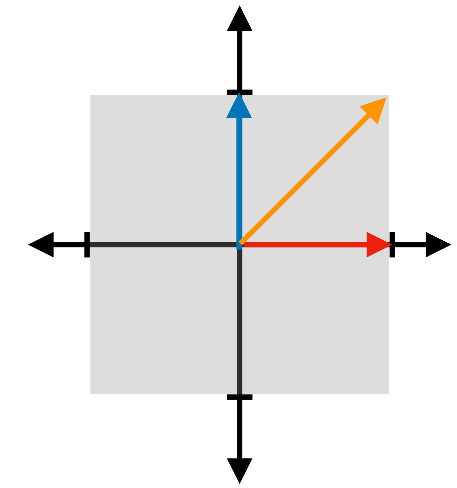
Compute the eigenvectors and eigenvalues of $A = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$.





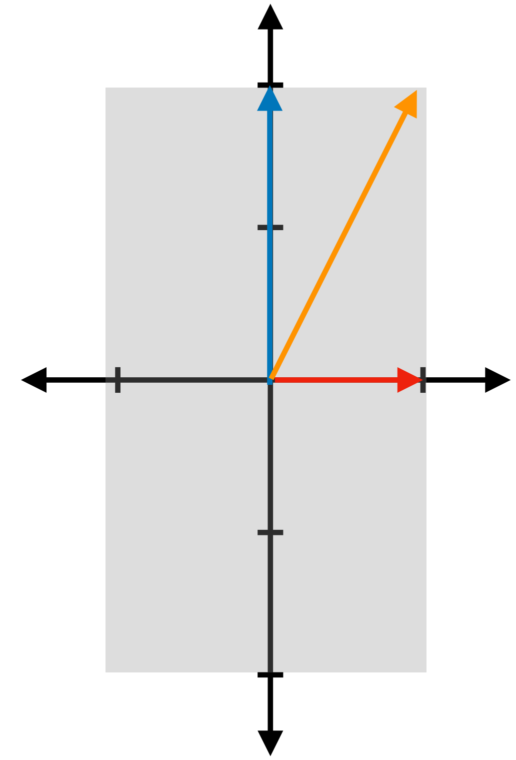
EXAMPLE

Compute the eigenvectors and eigenvalues of $A = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$.



$$\det(A - \lambda I) = 0$$

$$\begin{vmatrix} 1-\lambda & 0 \\ 0 & 2-\lambda \end{vmatrix} = (1-\lambda)(2-\lambda) = 0 \rightarrow \lambda_1 = 1,\ \lambda_2 = 2$$



- $\lambda_1 = 1,\ v_1 = \begin{bmatrix} x \\ y \end{bmatrix} \in N(A - 1I)$

$$\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},\ \begin{array}{c} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \\ \text{free col} \quad \text{pivot col} \end{array} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},\ \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} 1 \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$v_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \rightarrow$ Eigenvalue 1 with eigenvector $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$

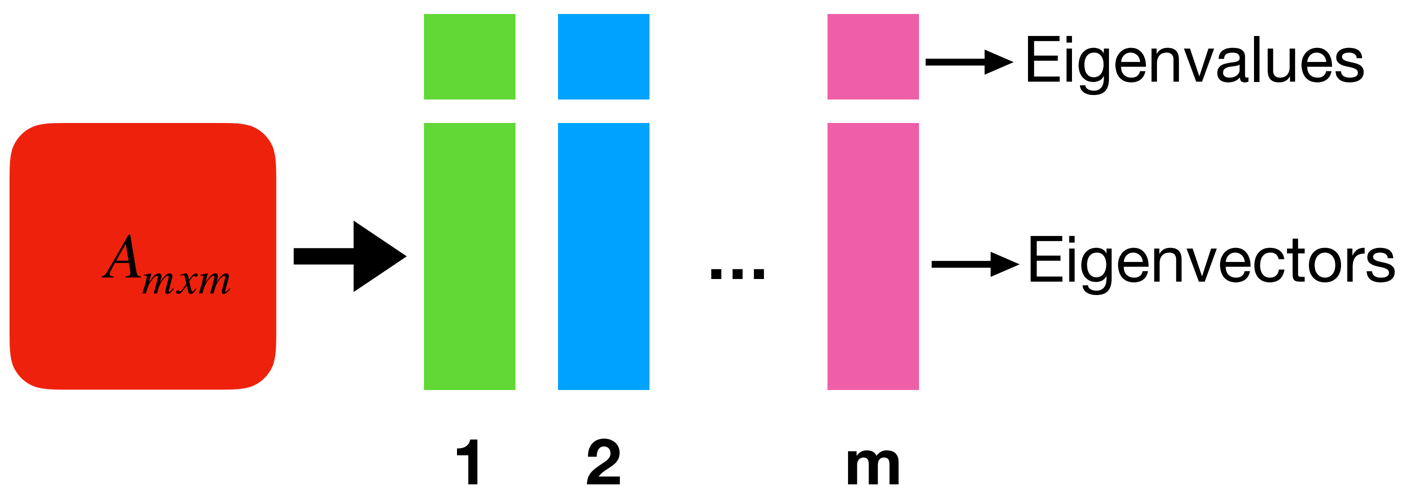
- $\lambda_2 = 2,\ v_2 = \begin{bmatrix} x \\ y \end{bmatrix} \in N(A - 2I)$

$$\begin{array}{c} \begin{bmatrix} -1 & 0 \\ 0 & 0 \end{bmatrix} \\ \text{pivot col} \quad \text{free col} \end{array} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},\ \begin{bmatrix} -1 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} x \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$v_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \rightarrow$ Eigenvalue 2 with eigenvector $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

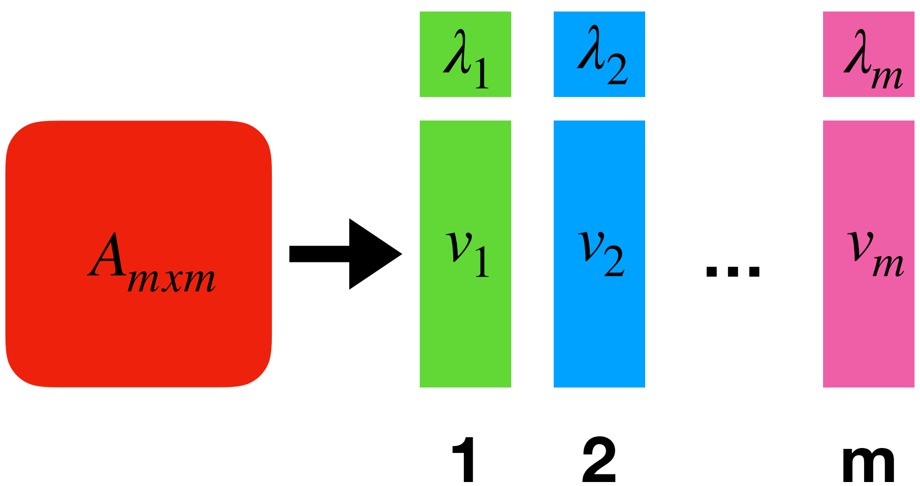
Eigenvalues	1	2
Eigenvectors	$\begin{bmatrix} 1 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 0 \\ 1 \end{bmatrix}$

Eigen Decomposition



Eigen Decomposition is the factorization of a square matrix into a canonical form, whereby the matrix is represented in terms of its eigenvalues and eigenvectors.

Diagonalization



$$Av_1 = \lambda_1 v_1$$
$$Av_2 = \lambda_2 v_2$$
$$\vdots$$
$$Av_m = \lambda_m v_m$$

$$\begin{aligned} Av_1 &= \lambda_1 v_1 \\ Av_2 &= \lambda_2 v_2 \\ Av_3 &= \lambda_3 v_3 \end{aligned} \longrightarrow \begin{bmatrix} v_{11} & v_{12} & v_{13} \\ v_{21} & v_{22} & v_{23} \\ v_{31} & v_{32} & v_{33} \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} = \begin{bmatrix} v_{11}\lambda_1 & v_{12}\lambda_2 & v_{13}\lambda_3 \\ v_{21}\lambda_1 & v_{22}\lambda_2 & v_{23}\lambda_3 \\ v_{31}\lambda_1 & v_{32}\lambda_2 & v_{33}\lambda_3 \end{bmatrix} \longrightarrow AV = V\Lambda$$

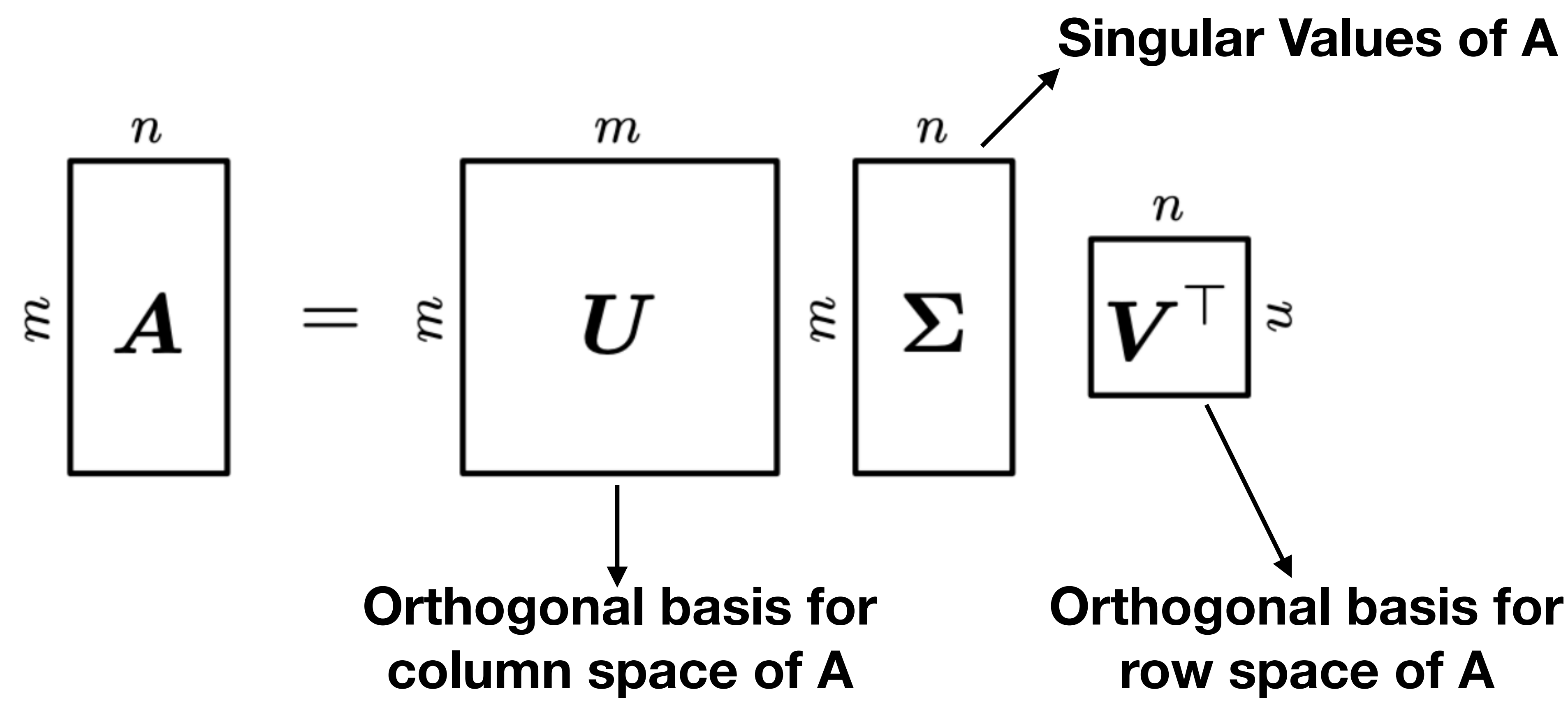
$$\boxed{\begin{aligned} A &= V\Lambda V^{-1} \\ V^{-1}AV &= \Lambda \end{aligned}}$$

Every symmetric matrix has the factorization $S = Q\Lambda Q^{-1}$ with real eigenvalues in Λ and orthonormal eigenvectors in the columns of $Q : S = Q\Lambda Q^{-1} = Q\Lambda Q^T$ with $Q^{-1} = Q^T$.

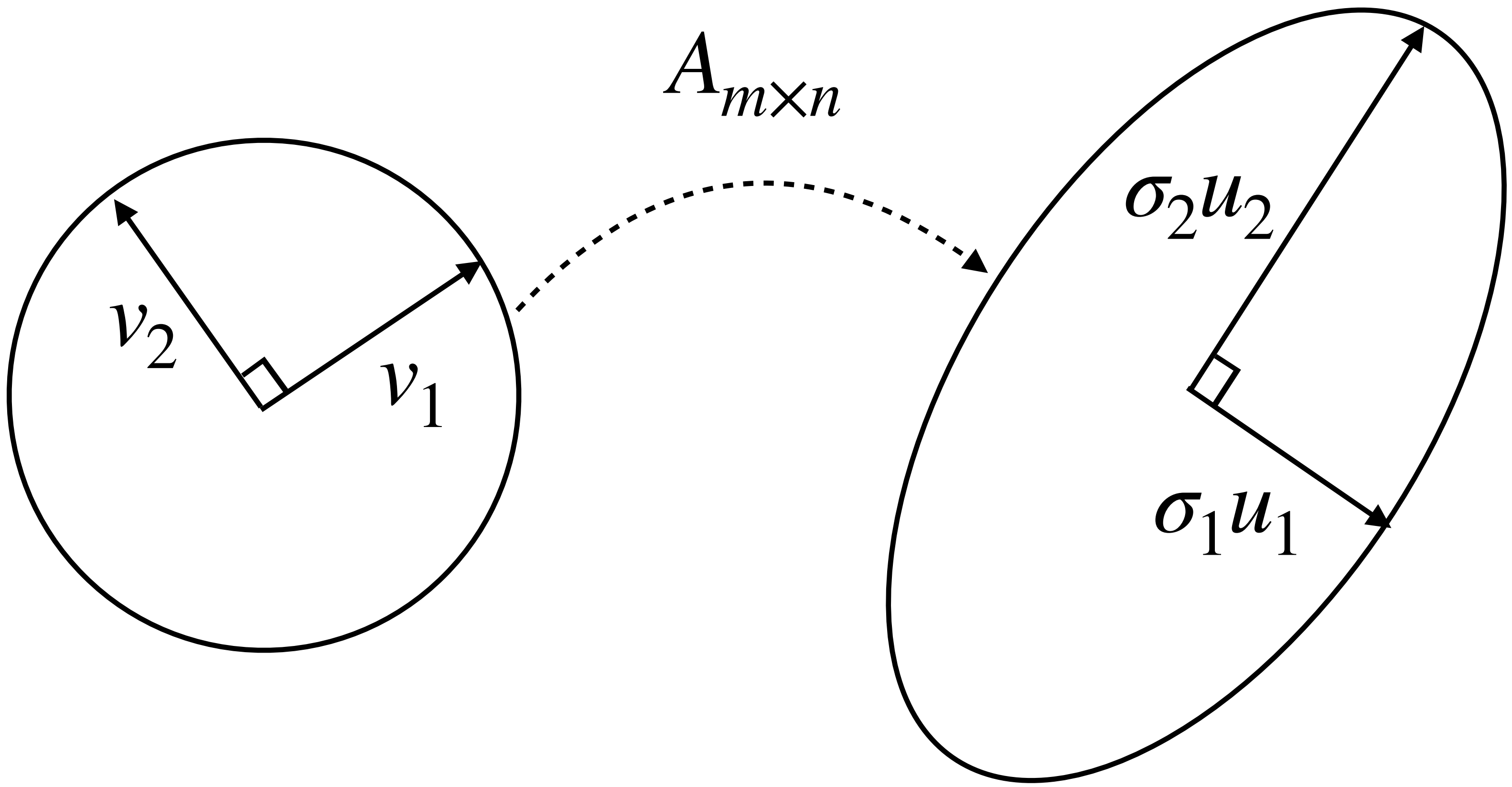
Singular Value Decomposition

SVD

The goal of SVD is to decompose a matrix A as the product of 3 other matrices $A = U\Sigma V^T$, where matrix V and U are orthogonal matrices and Σ is a diagonal matrix.



Formulation



$$\{\overrightarrow{v_1}, \overrightarrow{v_2}, \ldots, \overrightarrow{v_r}\} \rightarrow \{\overrightarrow{u_1}, \overrightarrow{u_2}, \ldots, \overrightarrow{u_r}\}$$

$$\overrightarrow{u_1}, \ldots, \overrightarrow{u_r} \ \ orthonormal$$

$$\sigma_1, \ldots, \sigma_r \ \ singular\ values$$

$$\sigma_1 \geq \sigma_2 \geq \sigma_3 \ldots \geq \sigma_r \geq 0$$

$$r = rank(A)$$

$$A\overrightarrow{v_1} = \sigma_1\overrightarrow{u_1},\ A\overrightarrow{v_2} = \sigma_2\overrightarrow{u_2} \rightarrow A\overrightarrow{v_j} = \sigma_j\overrightarrow{u_j},\ \forall j = 1,2,\ldots,r$$

$$\begin{bmatrix} & & \\ & A & \\ & & \end{bmatrix} \begin{bmatrix} | & | & & | \\ v_1 & v_2 & \ldots & v_r \\ | & | & & | \end{bmatrix} = \begin{bmatrix} | & | & & | \\ u_1 & u_2 & \ldots & u_r \\ | & | & & | \end{bmatrix} \begin{bmatrix} \sigma_1 & & & & \\ & \sigma_2 & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & \cdot \\ & & & & & \sigma_r \end{bmatrix}$$

$$AV = U\Sigma$$

$$A = U\Sigma V^{-1} = U\Sigma V^T$$

How to compute SVD?

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

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

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

$$A^TA = V\Sigma^T I\Sigma V^T$$

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

$$A^TA = V\Sigma^2 V^T$$

To find V compute eigendecomposition of A^TA where V will be the eigenvectors of A^TA and Σ^2 are the eigenvalues of A^TA.

$$AA^T = U\Sigma^2 U^T$$

To find U compute eigendecomposition of AA^T where U will be the eigenvectors of AA^T and Σ^2 are the eigenvalues of AA^T.

How to compute SVD?

- Columns of V are the eigenvectors of $A^T A$.
- Singular Values σ_i are the positive square root of the eigenvalues of $A^T A$.
- Columns of U are eigenvectors of AA^T.

$A=\begin{bmatrix}4 & 0\\ 3 & -5\end{bmatrix}$. Find U, Σ, V for A.

- Finding V and Σ

$$A^TA=\begin{bmatrix}4 & 3\\ 0 & -5\end{bmatrix}\begin{bmatrix}4 & 0\\ 3 & -5\end{bmatrix}=\begin{bmatrix}25 & -15\\ -15 & 25\end{bmatrix}$$

$$\det(A^TA-\lambda\underline{I})=\begin{vmatrix}25-\lambda & -15\\ -15 & 25-\lambda\end{vmatrix}=(25-\lambda)(25-\lambda)-225=0$$

$$\lambda^2-50\lambda+400=0 \rightarrow \lambda_1=40,\ \lambda_2=10$$

$$v_1\in N(A^TA-40\underline{I}) \rightarrow \begin{bmatrix}-15 & -15\\ -15 & -15\end{bmatrix}\begin{bmatrix}x\\ y\end{bmatrix}=\begin{bmatrix}0\\ 0\end{bmatrix} \rightarrow v_1=\begin{bmatrix}-1\\ 1\end{bmatrix},\ v_1=\begin{bmatrix}-1/\sqrt{2}\\ 1/\sqrt{2}\end{bmatrix}$$

$$v_2\in N(A^TA-10\underline{I}) \rightarrow \begin{bmatrix}15 & -15\\ -15 & 15\end{bmatrix}\begin{bmatrix}x\\ y\end{bmatrix}=\begin{bmatrix}0\\ 0\end{bmatrix} \rightarrow v_2=\begin{bmatrix}1\\ 1\end{bmatrix},\ v_2=\begin{bmatrix}1/\sqrt{2}\\ 1/\sqrt{2}\end{bmatrix}$$

$$V=\begin{bmatrix}| & |\\ v_1 & v_2\\ | & |\end{bmatrix}=\begin{bmatrix}-1/\sqrt{2} & 1/\sqrt{2}\\ 1/\sqrt{2} & 1/\sqrt{2}\end{bmatrix},\quad \Sigma=\begin{bmatrix}\sqrt{\lambda_1} & 0\\ 0 & \sqrt{\lambda_2}\end{bmatrix}=\begin{bmatrix}\sqrt{40} & 0\\ 0 & \sqrt{10}\end{bmatrix}$$

- Finding U.

$$Av_1 = \sigma_1 u_1 \rightarrow u_1 = \begin{bmatrix} -1/\sqrt{5} \\ -2/\sqrt{5} \end{bmatrix}$$

$$Av_2 = \sigma_2 u_2 \rightarrow u_2 = \begin{bmatrix} 2/\sqrt{5} \\ -1/\sqrt{5} \end{bmatrix}$$

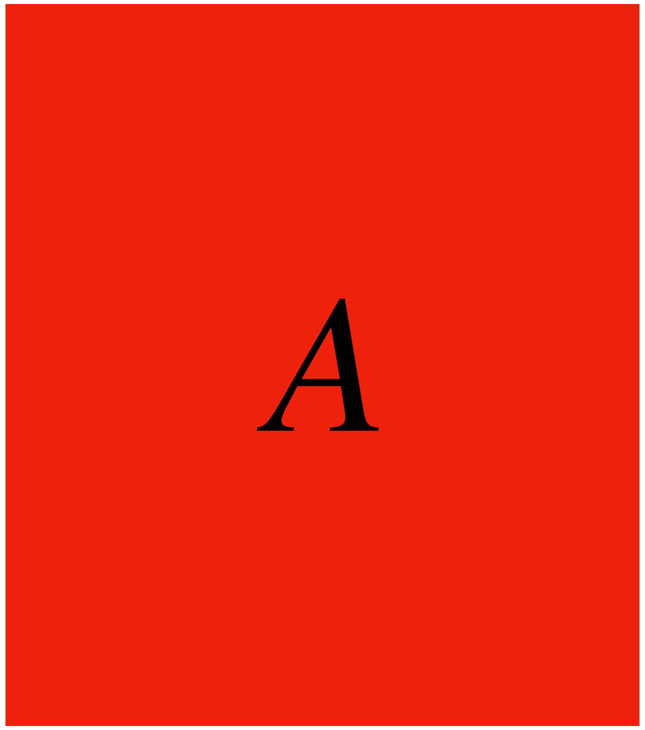
$$\longrightarrow U = \begin{bmatrix} | & | \\ u_1 & u_2 \\ | & | \end{bmatrix} = \begin{bmatrix} -1/\sqrt{5} & 2/\sqrt{5} \\ -2/\sqrt{5} & -1/\sqrt{5} \end{bmatrix}$$

$$U \Sigma V^T = \begin{bmatrix} -1/\sqrt{5} & 2/\sqrt{5} \\ -2/\sqrt{5} & -1/\sqrt{5} \end{bmatrix} \begin{bmatrix} \sqrt{40} & 0 \\ 0 & \sqrt{10} \end{bmatrix} \begin{bmatrix} -1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$$

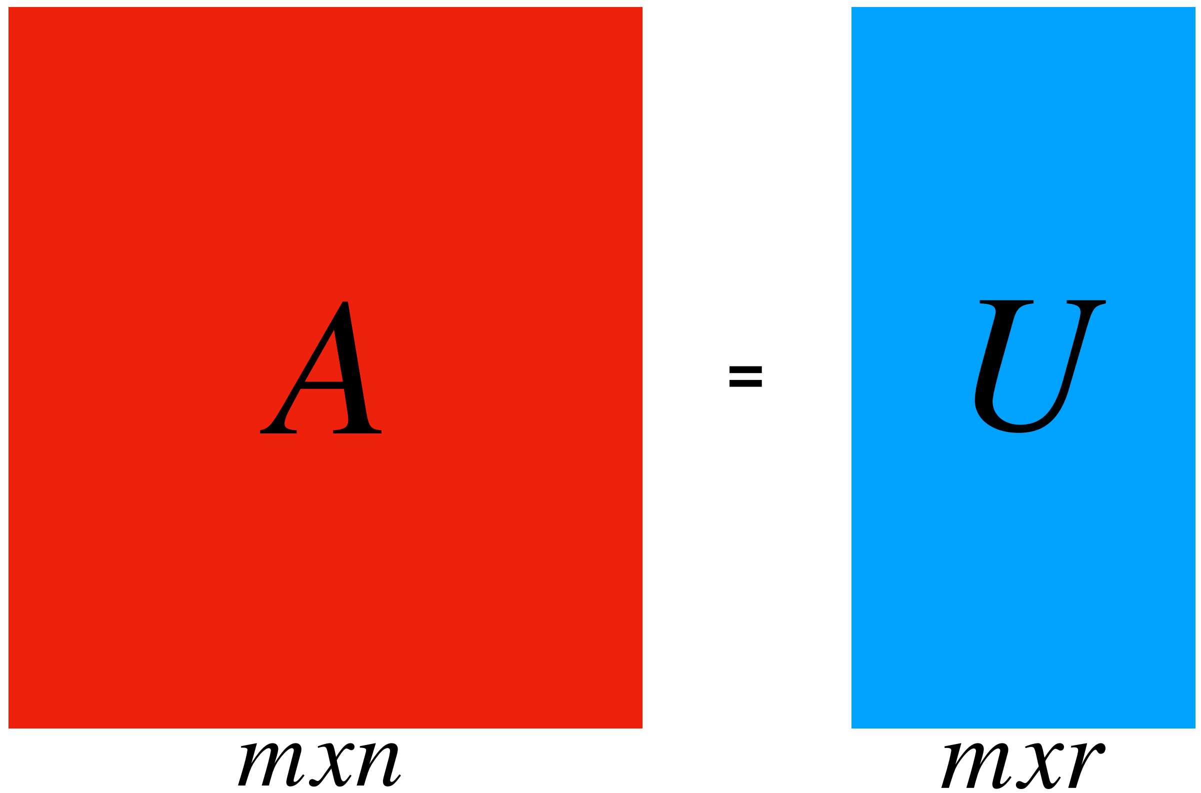
$$= \begin{bmatrix} -1/\sqrt{5} & 2/\sqrt{5} \\ -2/\sqrt{5} & -1/\sqrt{5} \end{bmatrix} \begin{bmatrix} -2\sqrt{5} & 2\sqrt{5} \\ \sqrt{5} & \sqrt{5} \end{bmatrix}$$

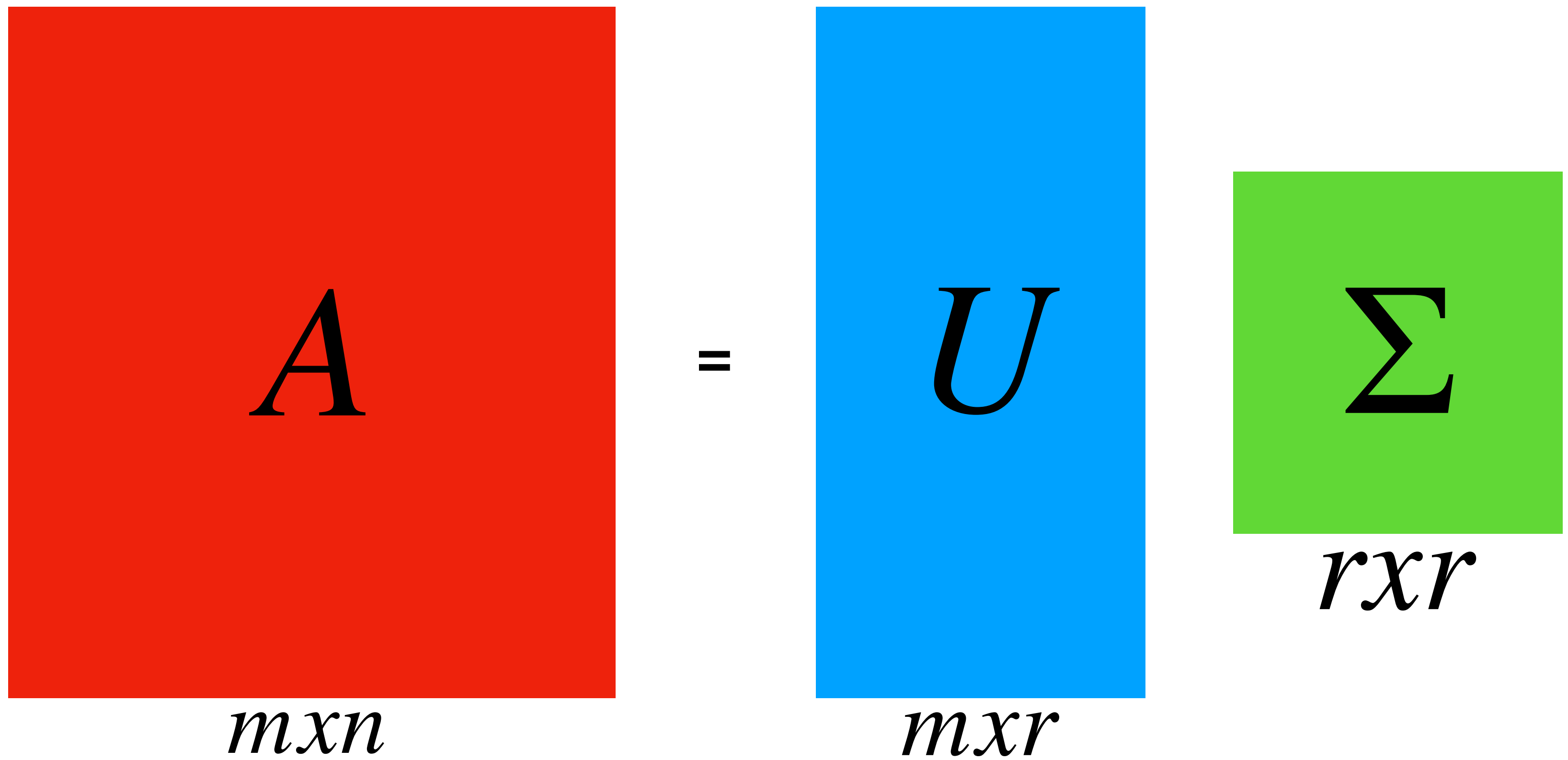
$$= \begin{bmatrix} 4 & 0 \\ 3 & -5 \end{bmatrix} = A \checkmark$$

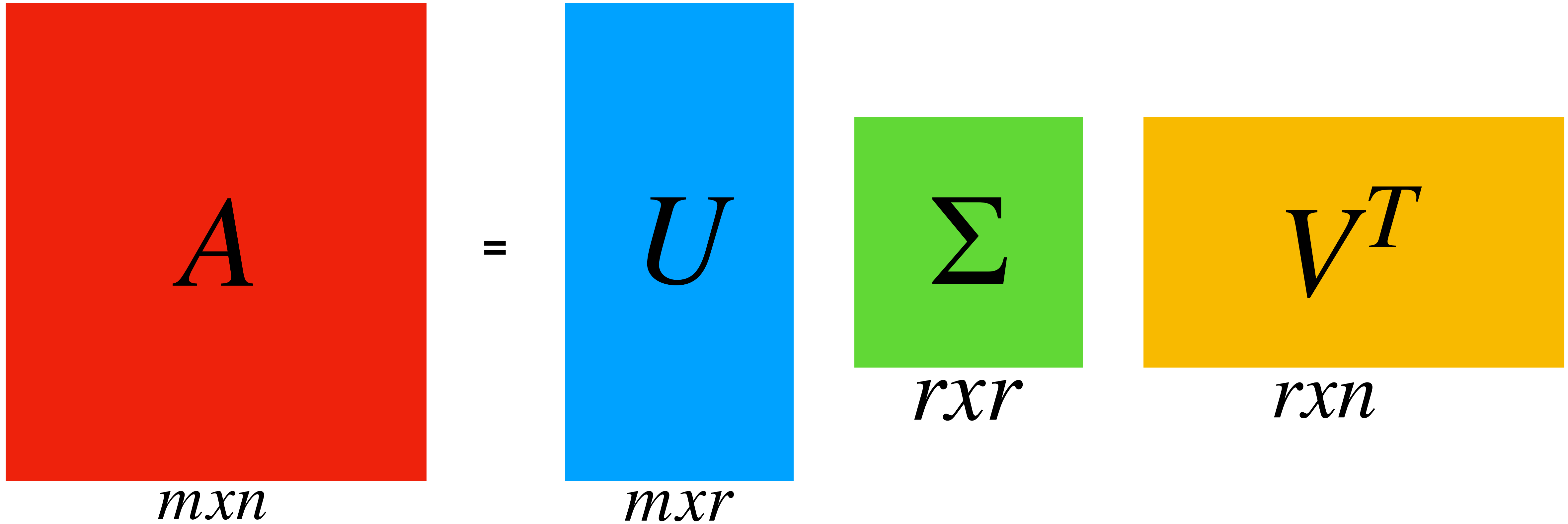
Matrix Approximation with SVD



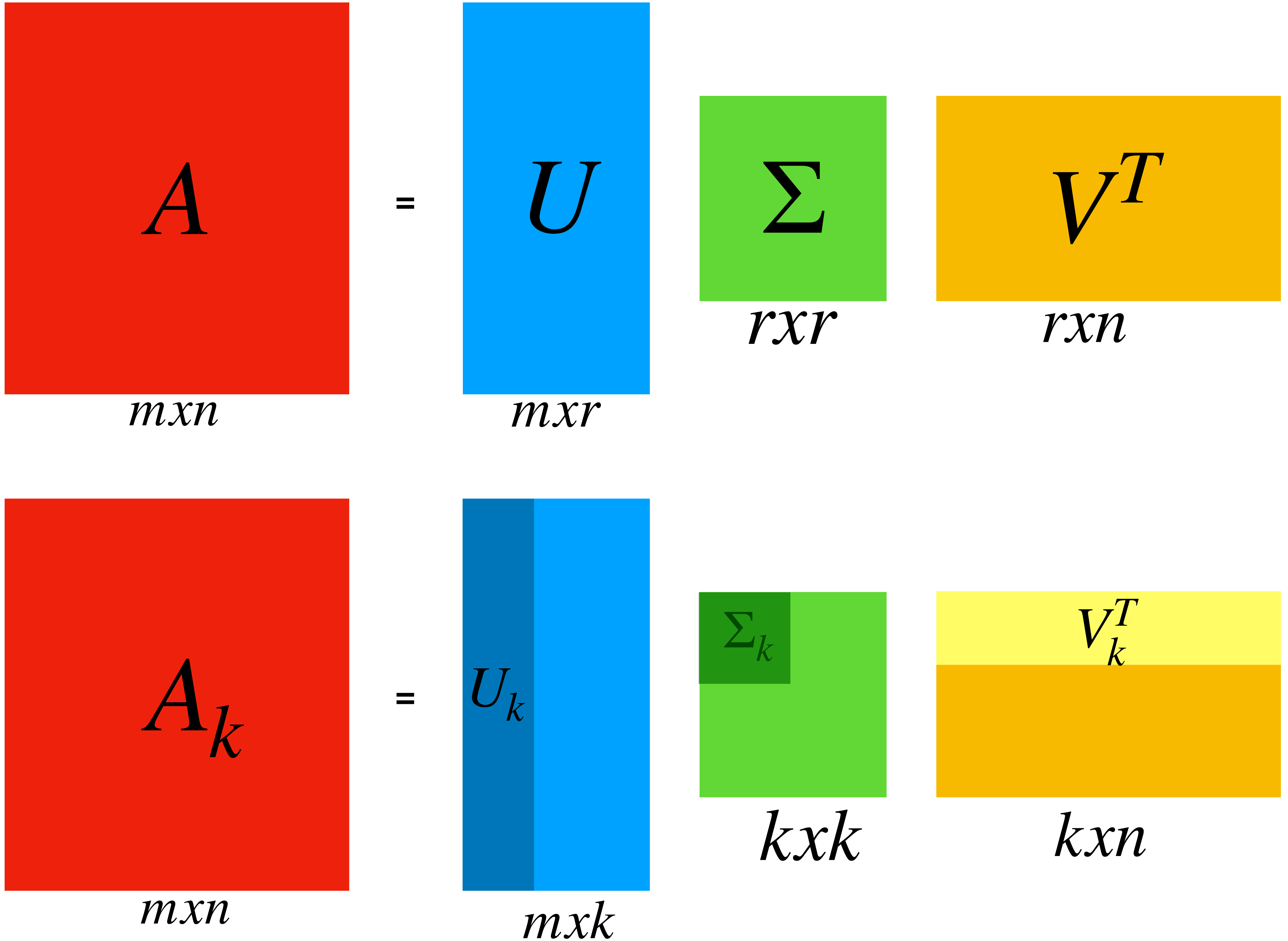
mxn


A
=
U
mxn
mxr

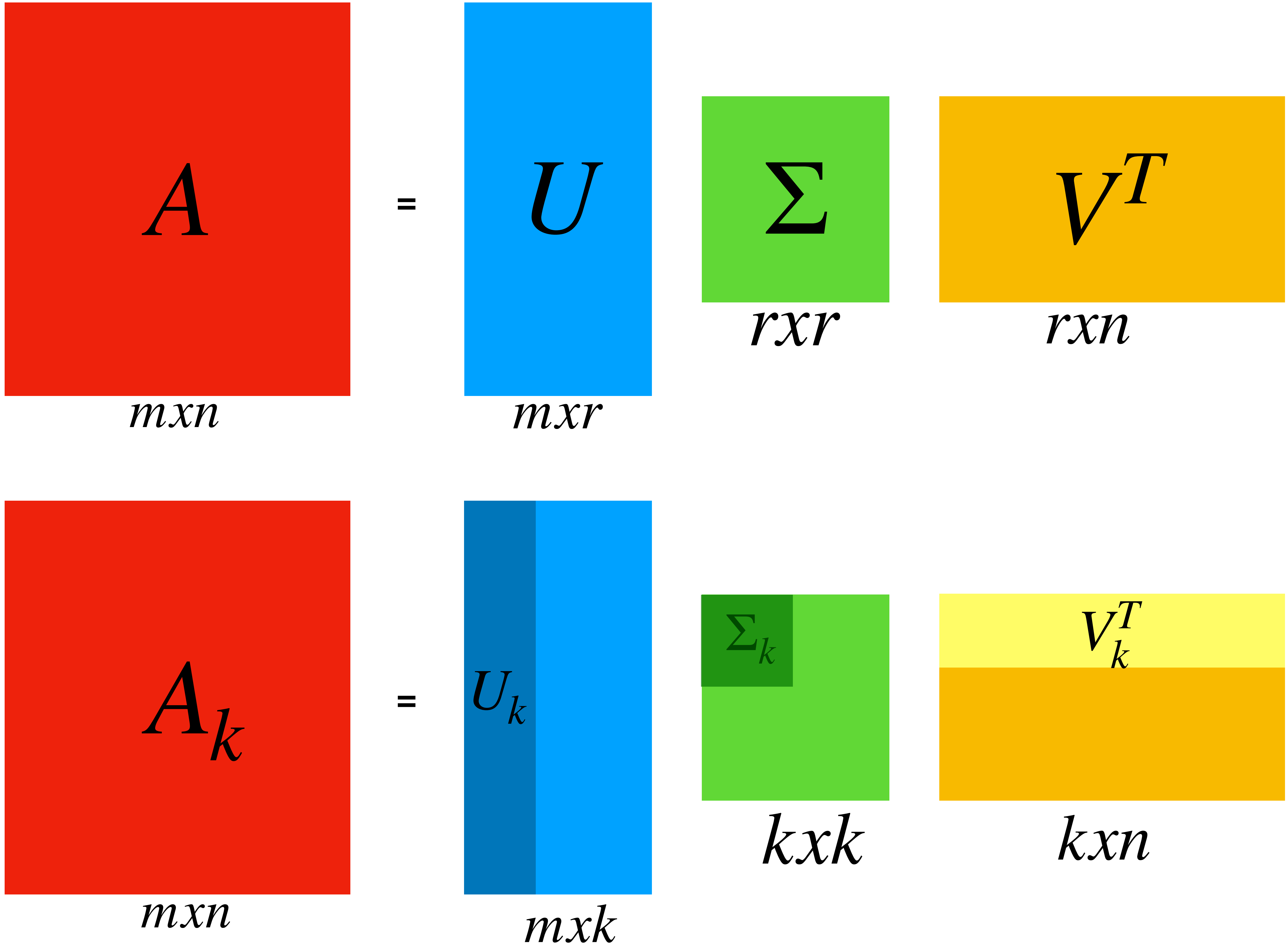

A
=
U
Σ
mxn
mxr
rxr



A
=
U
Σ
V^T
rxr
rxn
mxn
mxr



A
mxn
=
U
mxr
Σ
rxr
V^T
rxn
A_k
mxn
=
U_k
mxk
Σ_k
kxk
V_k^T
kxn



Eckart - Young Theorem 1936

The reconstruction matrix $A_k = U_k \Sigma_k V_k^T$ is the closest rank-k matrix to A.

Lab: Image Compression

https://github.com/terollierisa/svd

Conclusion

- Matrix Factorization: SVD breaks down a matrix into three distinct matrices $U\Sigma V^T$.
- Universality of SVD: SVD is applicable to any matrix.
- Diverse Applications: SVD is extensively utilized in image compression, recommender systems, dimensionality reduction, signal processing, and various other fields.

References

- Strang, Gilbert. *Introduction to linear algebra*. Wellesley-Cambridge Press, 2022.
- Brunton, Steven L., and J. Nathan Kutz. Data-driven science and engineering: Machine learning, dynamical systems, and control. Cambridge University Press, 2022.
- Eckart, Carl, and Gale Young. "The approximation of one matrix by another of lower rank." *Psychometrika* 1.3 (1936): 211-218.

Thank you!

Erisa Terolli
Stevens Institute of Technology
erisa.terolli90@gmail.com
Slides + Code: https://github.com/terollierisa/svd