Introductory Linear Algebra for AI

- Useful Matrix Properties

Notation Review

$\det(\mathbf{A})$	Determinant of $\mathbf{A}$
$\mathrm{Tr}(\mathbf{A})$	Trace of the matrix $\mathbf{A}$
$\mathrm{diag}(\mathbf{A})$	Diagonal matrix of the matrix $\mathbf{A}$, i.e. $(\mathrm{diag}(\mathbf{A}))_{ij} = \delta_{ij} A_{ij}$
$\mathrm{eig}(\mathbf{A})$	Eigenvalues of the matrix $\mathbf{A}$
$\mathrm{vec}(\mathbf{A})$	The vector-version of the matrix $\mathbf{A}$ (see Sec. 10.2.2)
$\sup$	Supremum of a set
$\|\|\mathbf{A}\|\|$	Matrix norm (subscript if any denotes what norm)
$\mathbf{A}^T$	Transposed matrix
$\mathbf{A}^{-T}$	The inverse of the transposed and vice versa, $\mathbf{A}^{-T} = (\mathbf{A}^{-1})^T = (\mathbf{A}^T)^{-1}$.

Inverse, Transpose

$$
\begin{aligned}
(\mathbf{AB})^{-1} &= \mathbf{B}^{-1}\mathbf{A}^{-1} \\
(\mathbf{ABC}...)^{-1} &= ...\mathbf{C}^{-1}\mathbf{B}^{-1}\mathbf{A}^{-1} \\
(\mathbf{A}^T)^{-1} &= (\mathbf{A}^{-1})^T \\
(\mathbf{A}+\mathbf{B})^T &= \mathbf{A}^T+\mathbf{B}^T \\
(\mathbf{AB})^T &= \mathbf{B}^T\mathbf{A}^T \\
(\mathbf{ABC}...)^T &= ...\mathbf{C}^T\mathbf{B}^T\mathbf{A}^T
\end{aligned}
$$

Inverse, Transpose

$$
\begin{aligned}
(\mathbf{AB})^{-1} &= \mathbf{B}^{-1}\mathbf{A}^{-1} \\
(\mathbf{ABC}...)^{-1} &= ...\mathbf{C}^{-1}\mathbf{B}^{-1}\mathbf{A}^{-1} \\
(\mathbf{A}^T)^{-1} &= (\mathbf{A}^{-1})^T \\
(\mathbf{A}+\mathbf{B})^T &= \mathbf{A}^T+\mathbf{B}^T \\
(\mathbf{AB})^T &= \mathbf{B}^T\mathbf{A}^T \\
(\mathbf{ABC}...)^T &= ...\mathbf{C}^T\mathbf{B}^T\mathbf{A}^T
\end{aligned}
$$

Introductory Linear Algebra for AI

- Practices

Eigenvalue and Eigenvector

$$A = \begin{bmatrix} -5 & 2 \\ 2 & -2 \end{bmatrix}$$

Linear Systems

$$A = \begin{bmatrix} -5 & 2 \\ 2 & -2 \end{bmatrix}$$

$$\frac{dx(t)}{dt} = Ax(t)$$

Linear Systems

General Solution:

$$x(t) = c_1 e^{\lambda_1 t} x_1(t) + c_2 e^{\lambda_2 t} x_2(t)$$

Linear Systems

What happens if eigenvalues are the same?

Linear Systems

Complex eigenvalues

Matrix Inverse: Gauss-Jordan Elimination

Inverse of $\begin{bmatrix} -1 & 1 & 2 \\ 3 & -1 & 1 \\ -1 & 3 & 4 \end{bmatrix}$

Matrix: Notations

Let A and B be 2×2 matrices.
(a) Prove that if $tr(A) = 0$, then A^2 is a scalar multiple of the identity matrix

Matrix: Notations

Let A and B be 2×2 matrices.
(b) Let [A,B] = AB-BA. Prove that the square of [A,B] commutes with every 2×2 matrix C.

Introductory Linear Algebra for AI

- Vector & Matrix
- Geometrical Meaning
- Application to AI

Fundamental Concepts of Linear Algebra

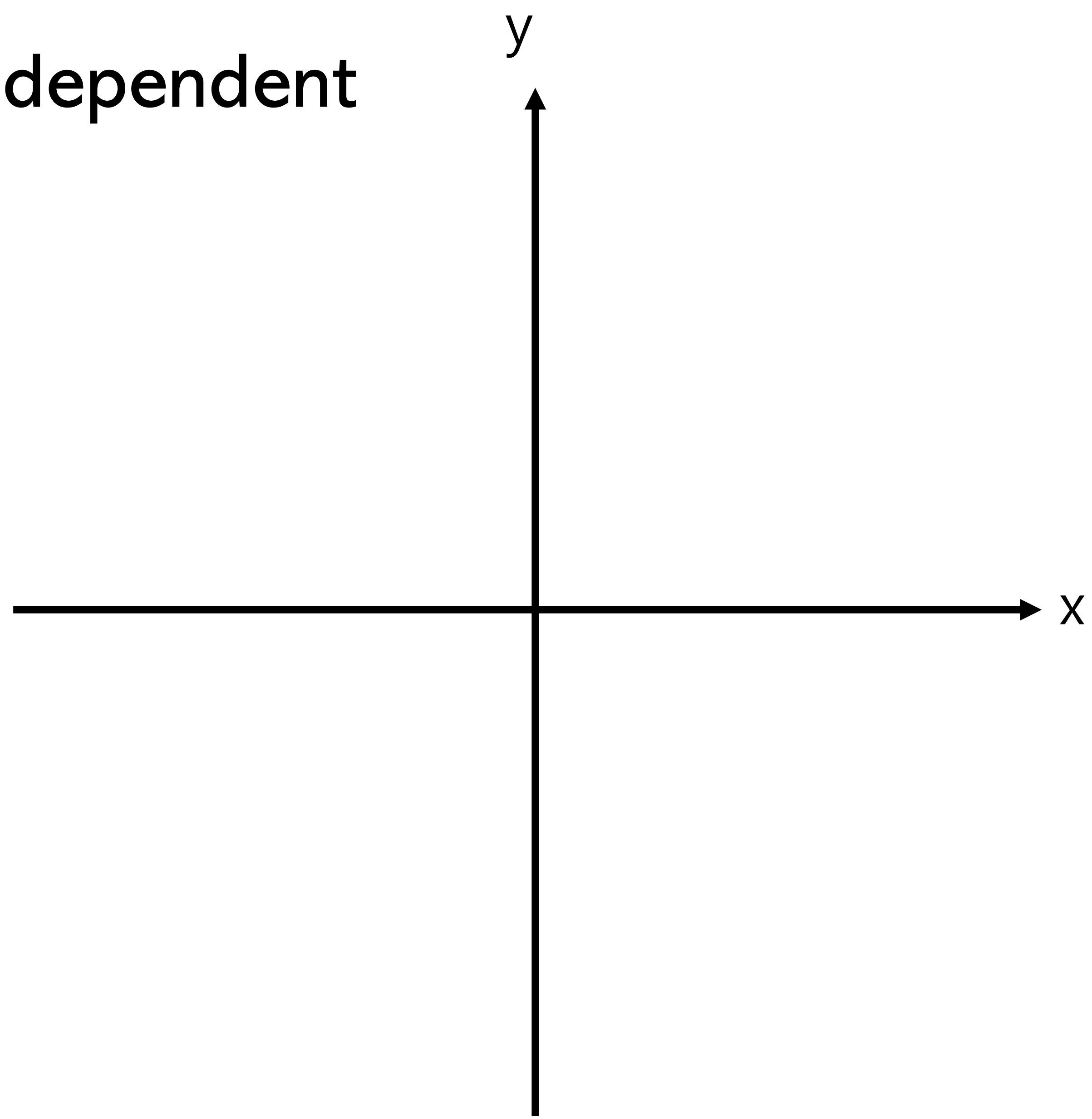
- Addition

- Scalar Multiplication (Scaling)

Dimension of Data

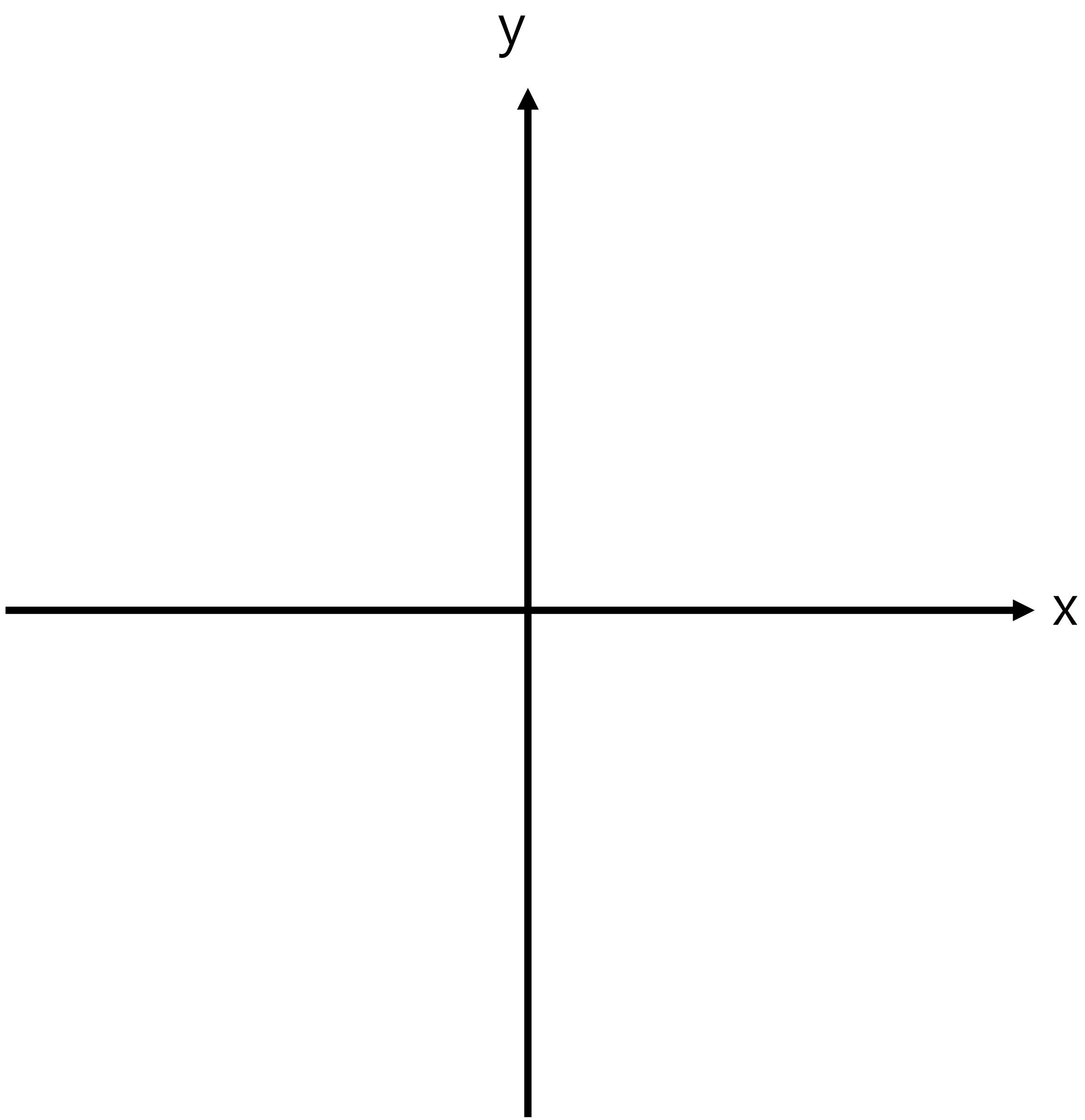
- Point
- Scalar
- Vector
- Matrix
- Tensor

Vector, Linear Independent

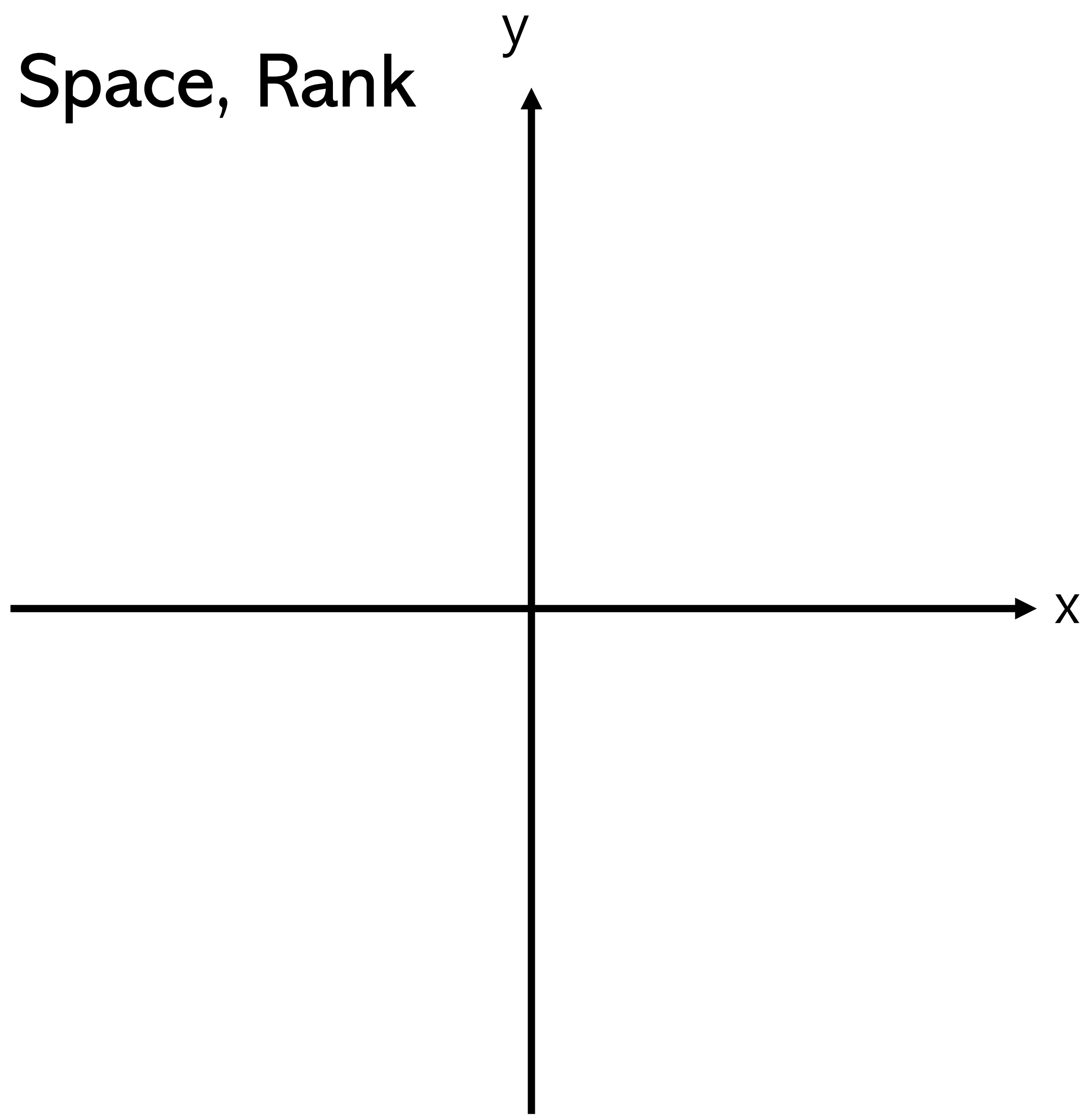


Basis

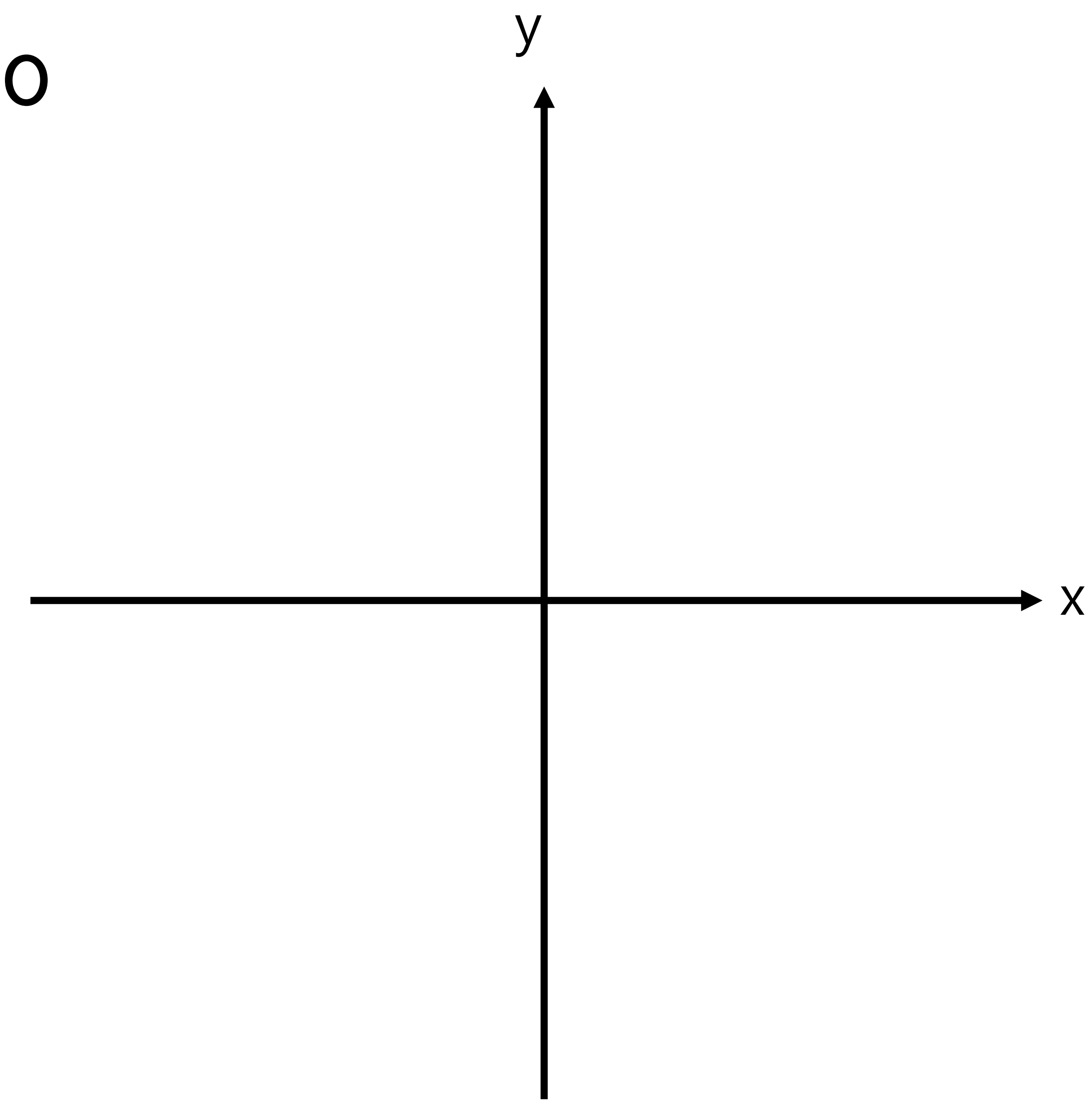
Span



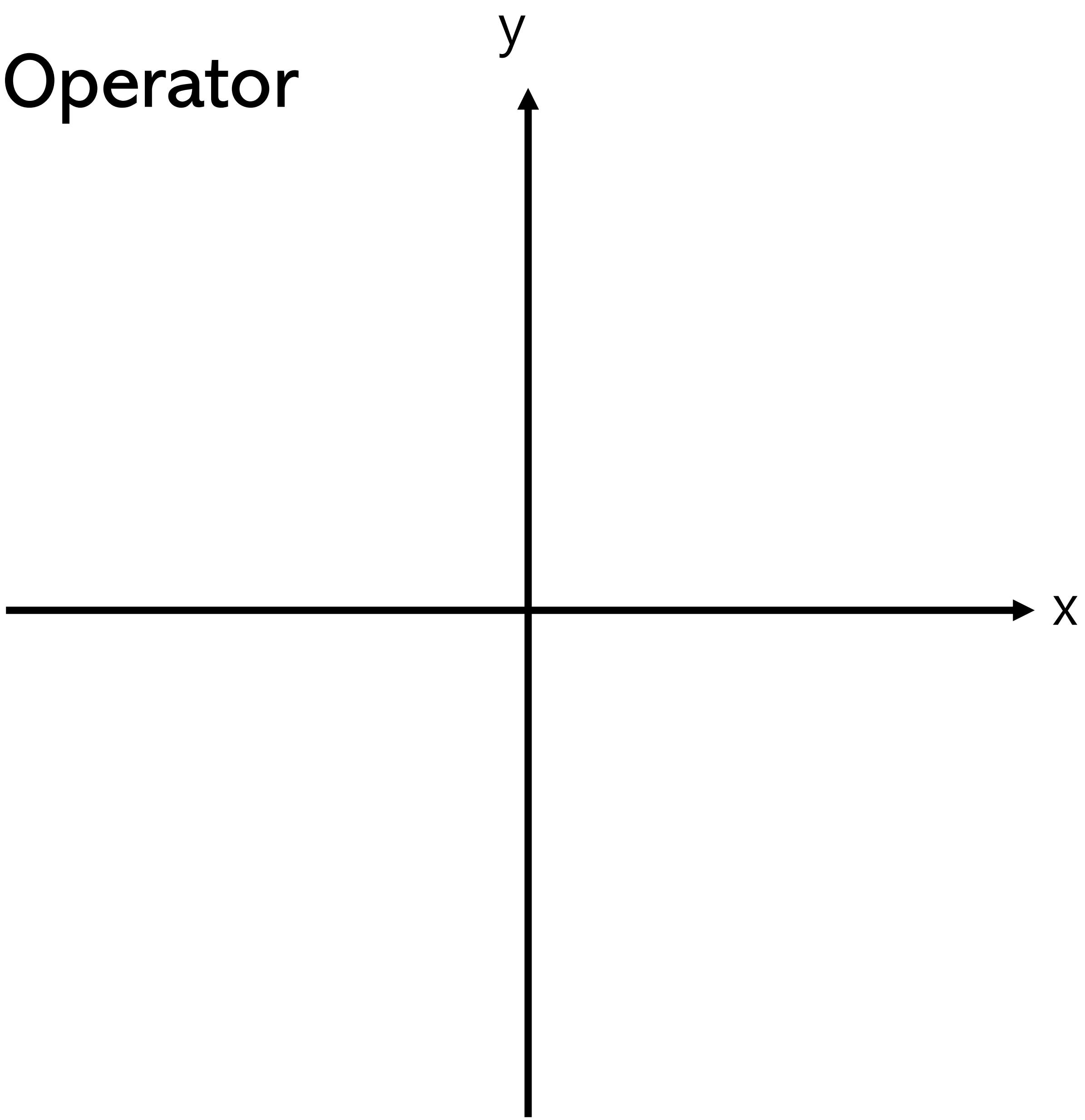
Matrix as Vector Space, Rank



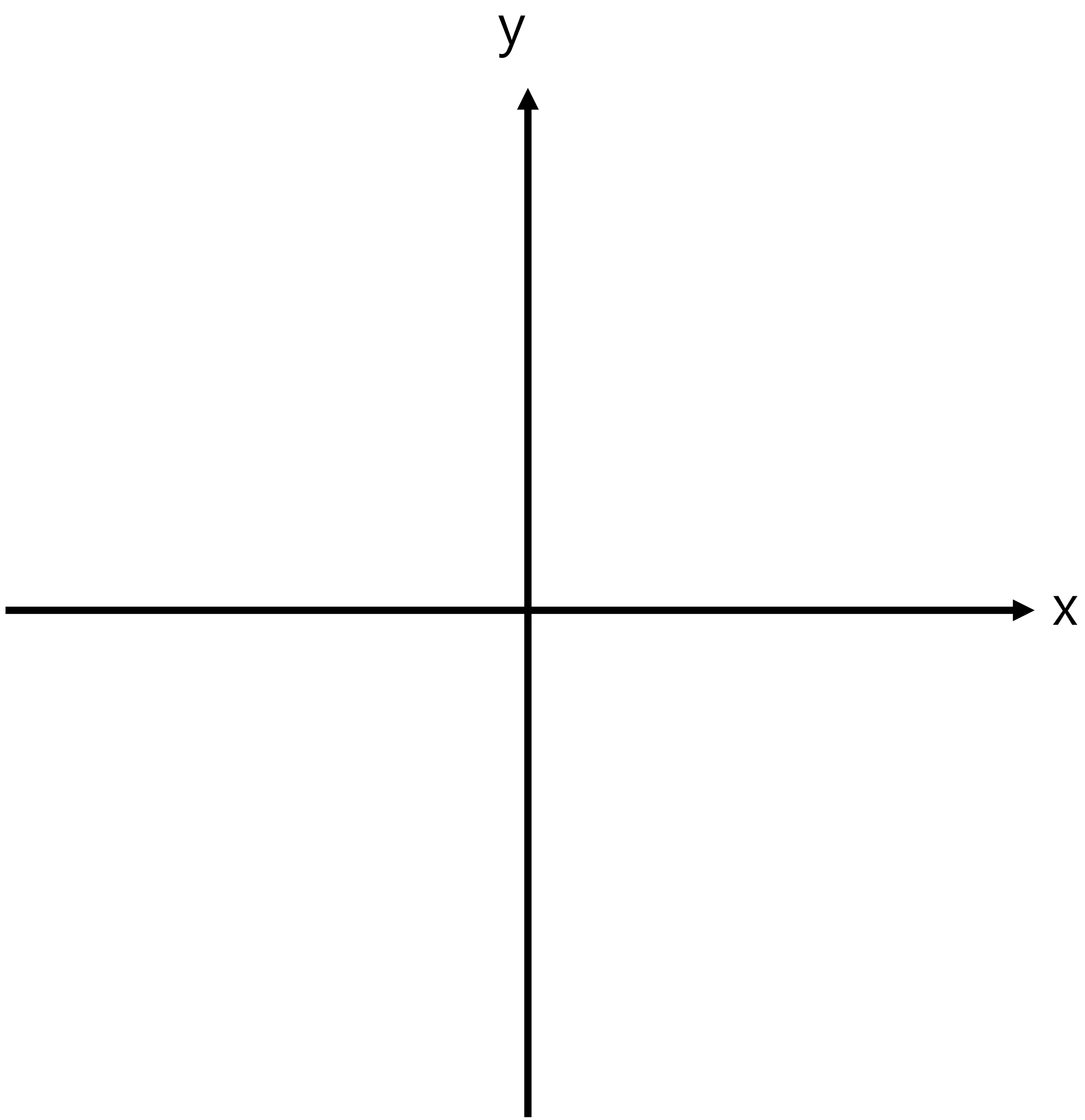
Null Space, Ax=0



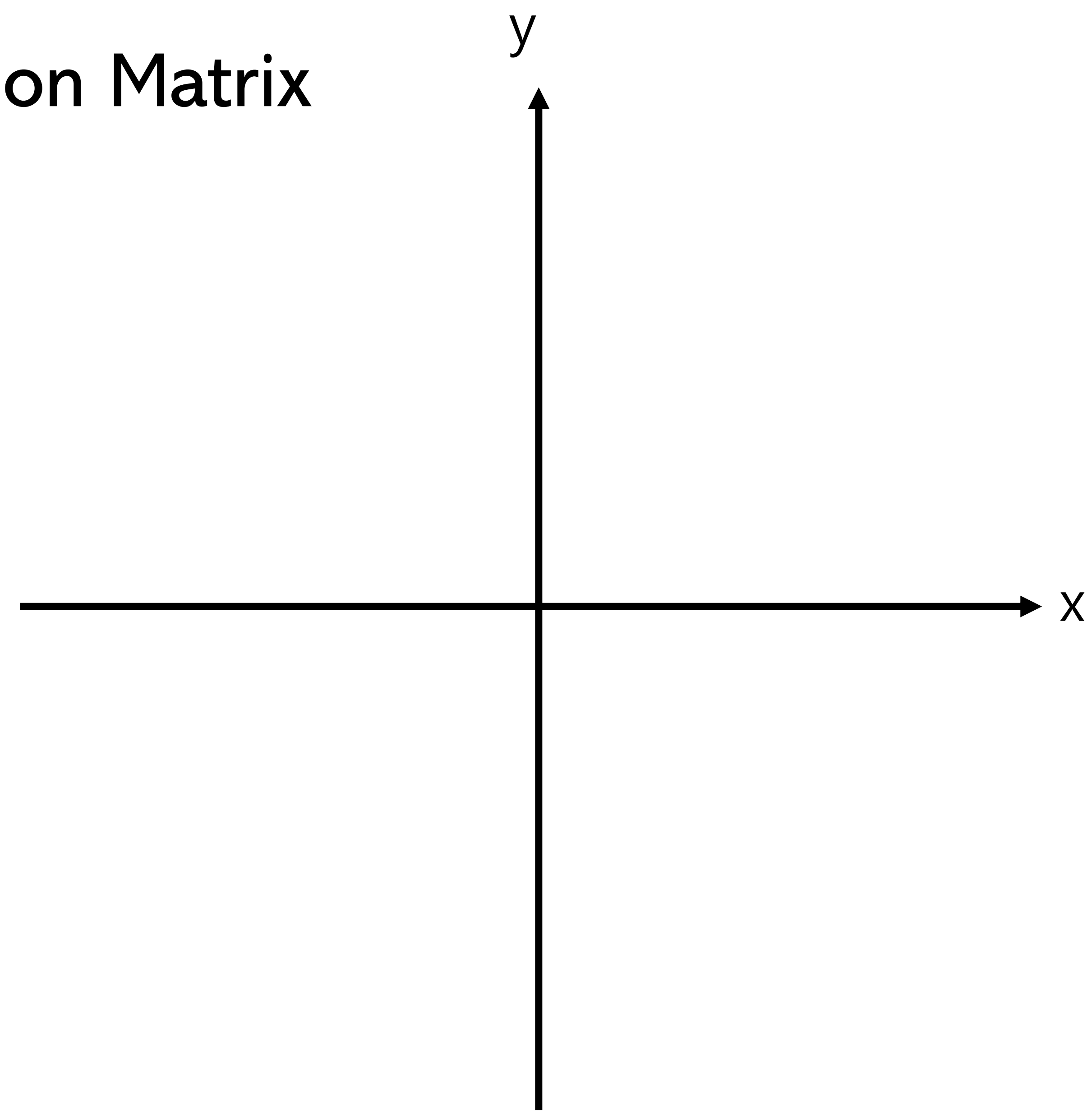
Matrix as Linear Operator



Vector to Vector



(Example) Rotation Matrix



Matrix Multiplication, AB – BA = ?

Matrix: Notations

- Matrix Transpose: $A^{\top}$

- Symmetric: $A^{\top} = A$
- Skew-Symmetric: $A^{\top} = -A$

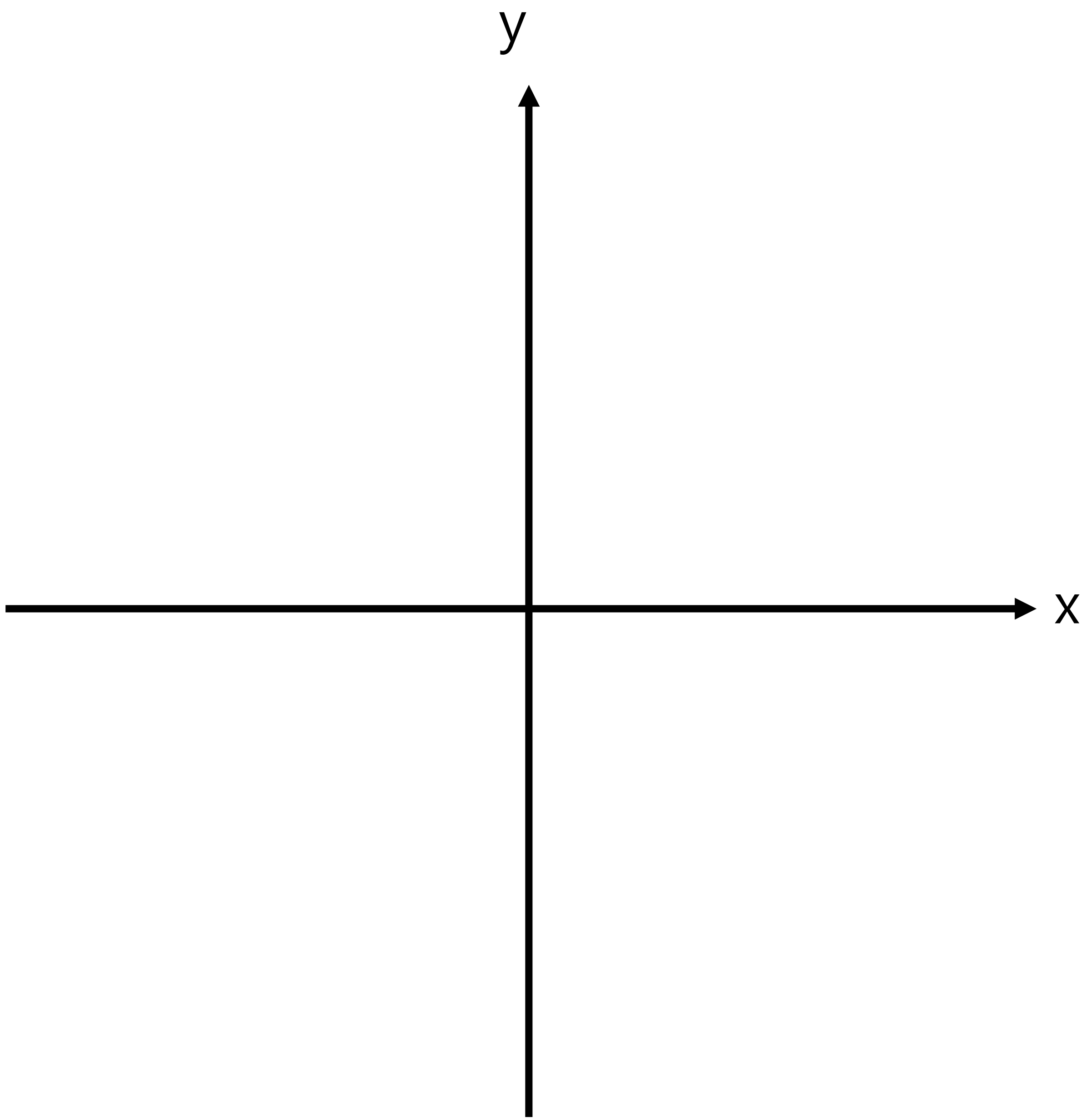
- Diagonal Matrix
- Triangular Matrix

(Advanced) Explaining Neural Network

(Example) Matrix and Vector

- Prove the following relation using the property of ‘matrix as linear operator’
- $\cos(\alpha+\beta) = \cos\alpha\cos\beta - \sin\alpha\sin\beta$
- $\sin(\alpha+\beta) = \sin\alpha\cos\beta - \cos\alpha\sin\beta$

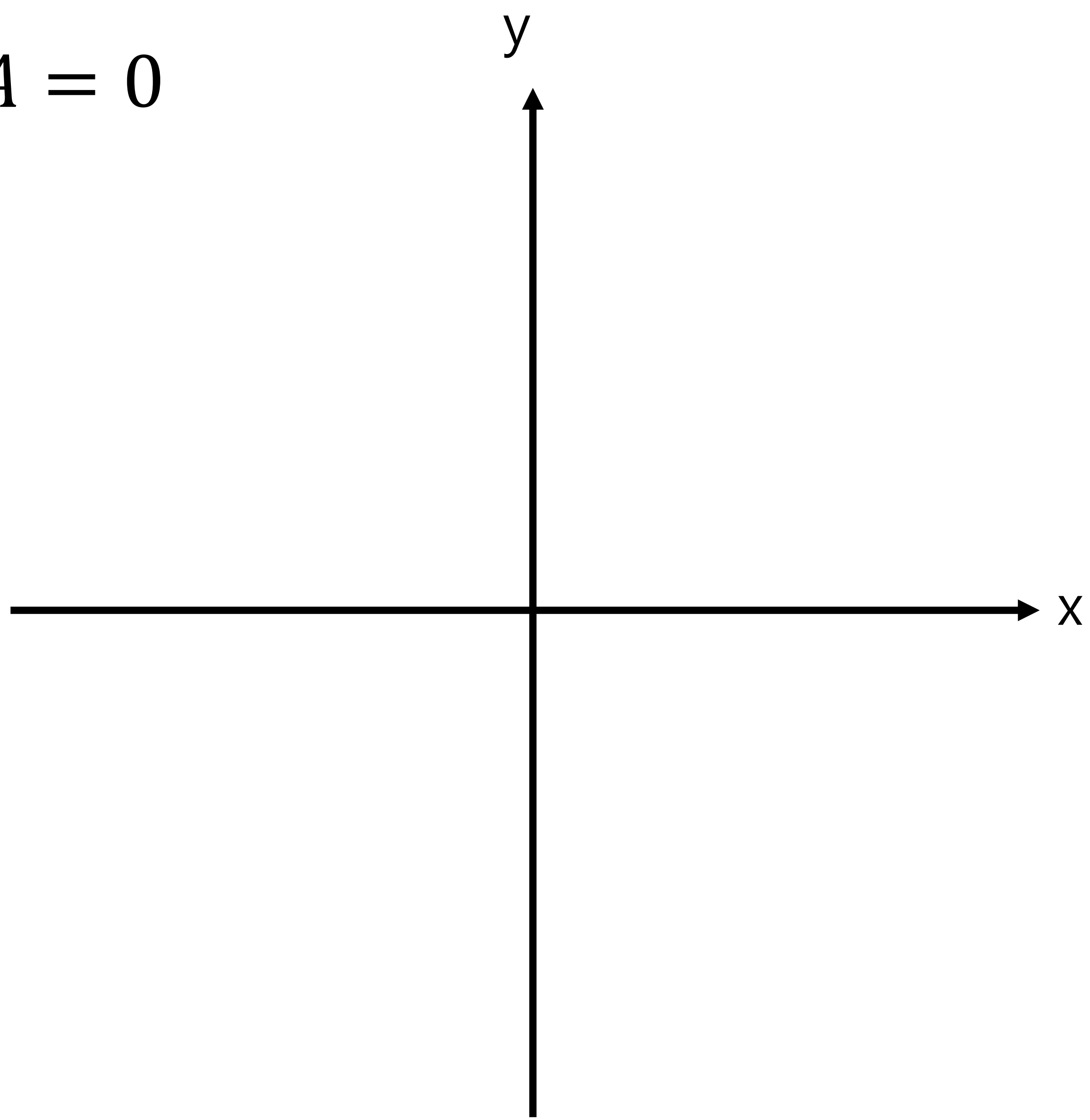
Determinant = Transformation of *Volume*



Calculation of Determinant: Cramer's Rule

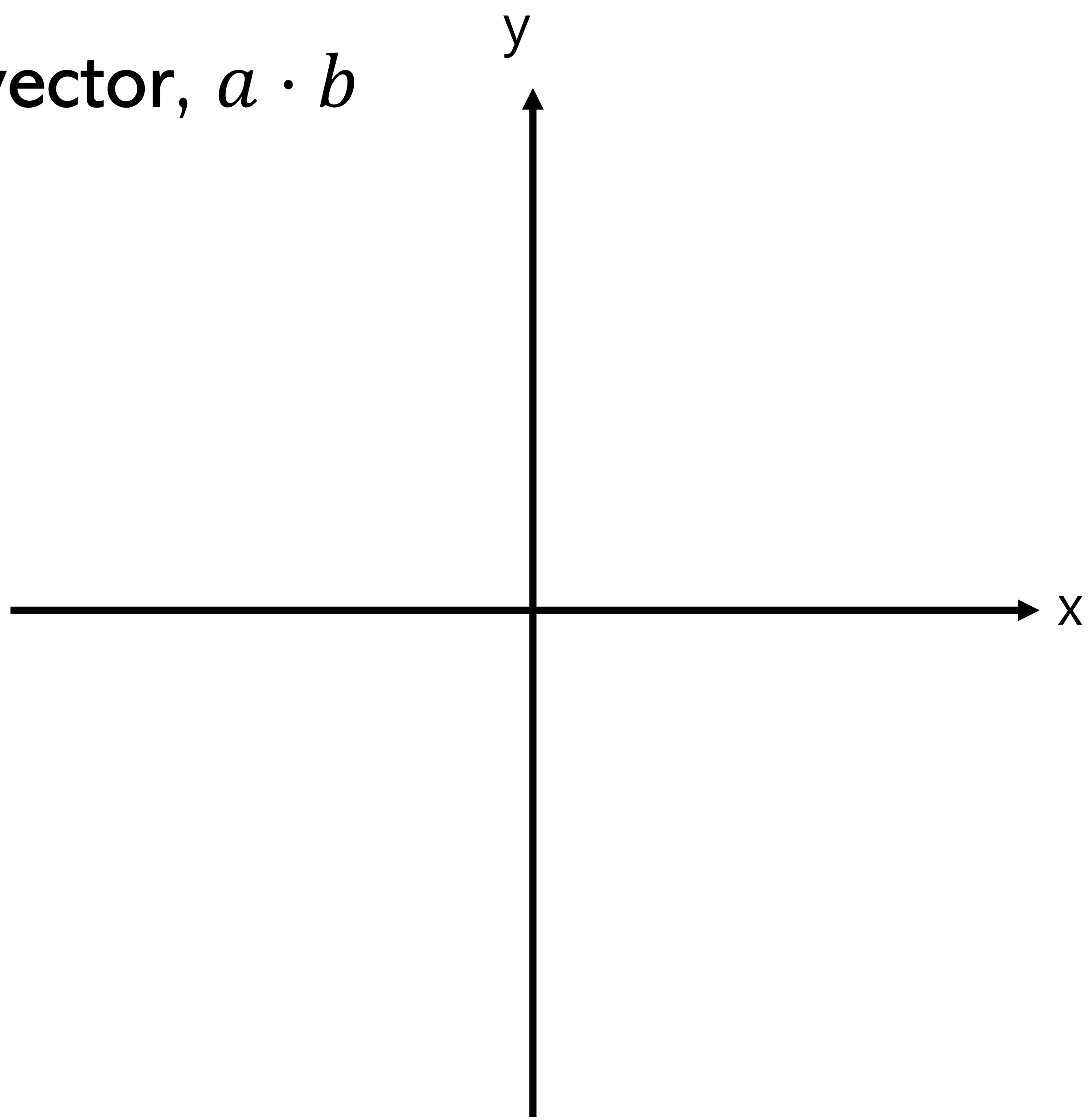
(example) 3x3 matrix

Meaning of $\det A = 0$

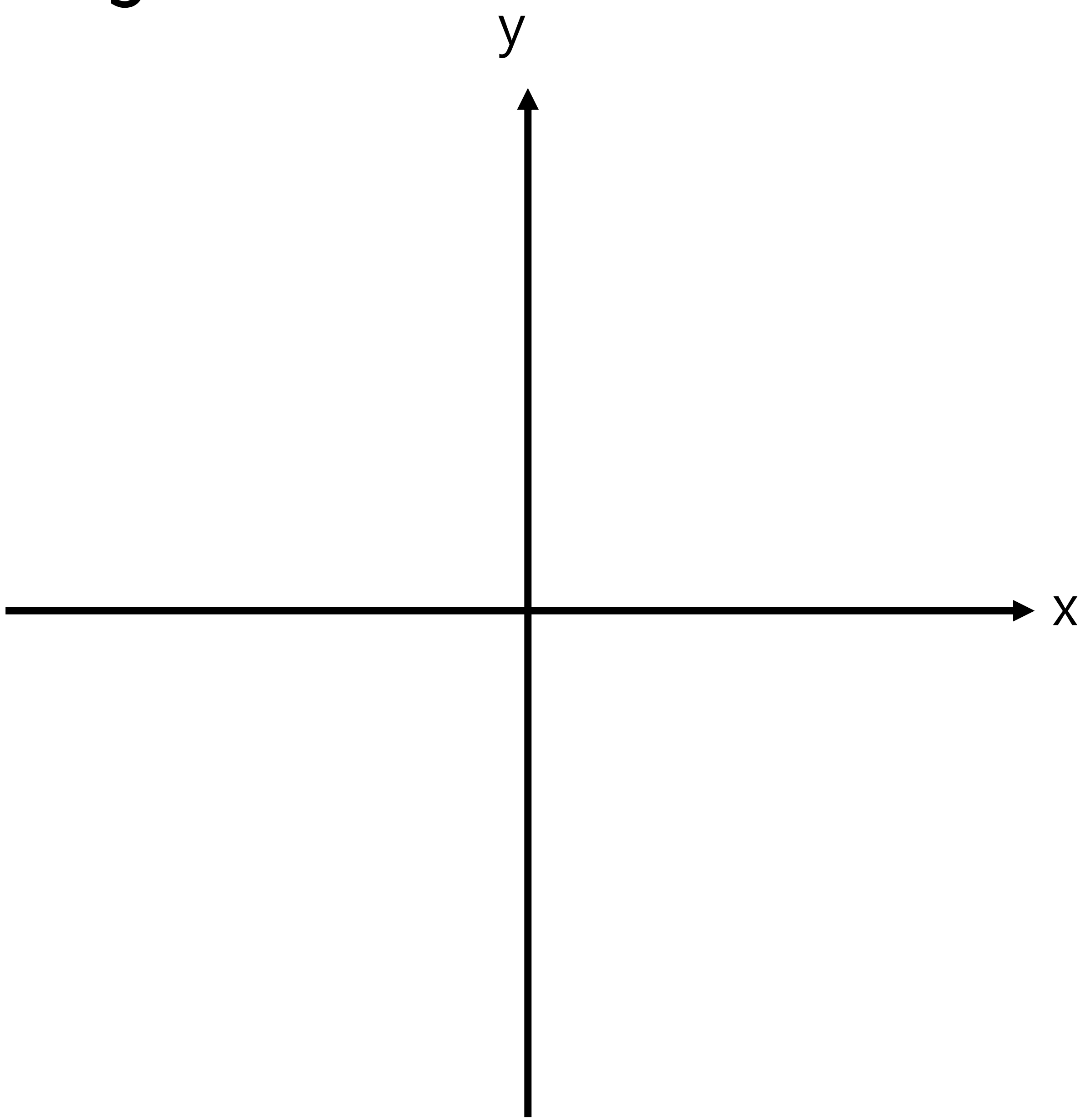


Matrix Inverse, $AA^{-1} = I$

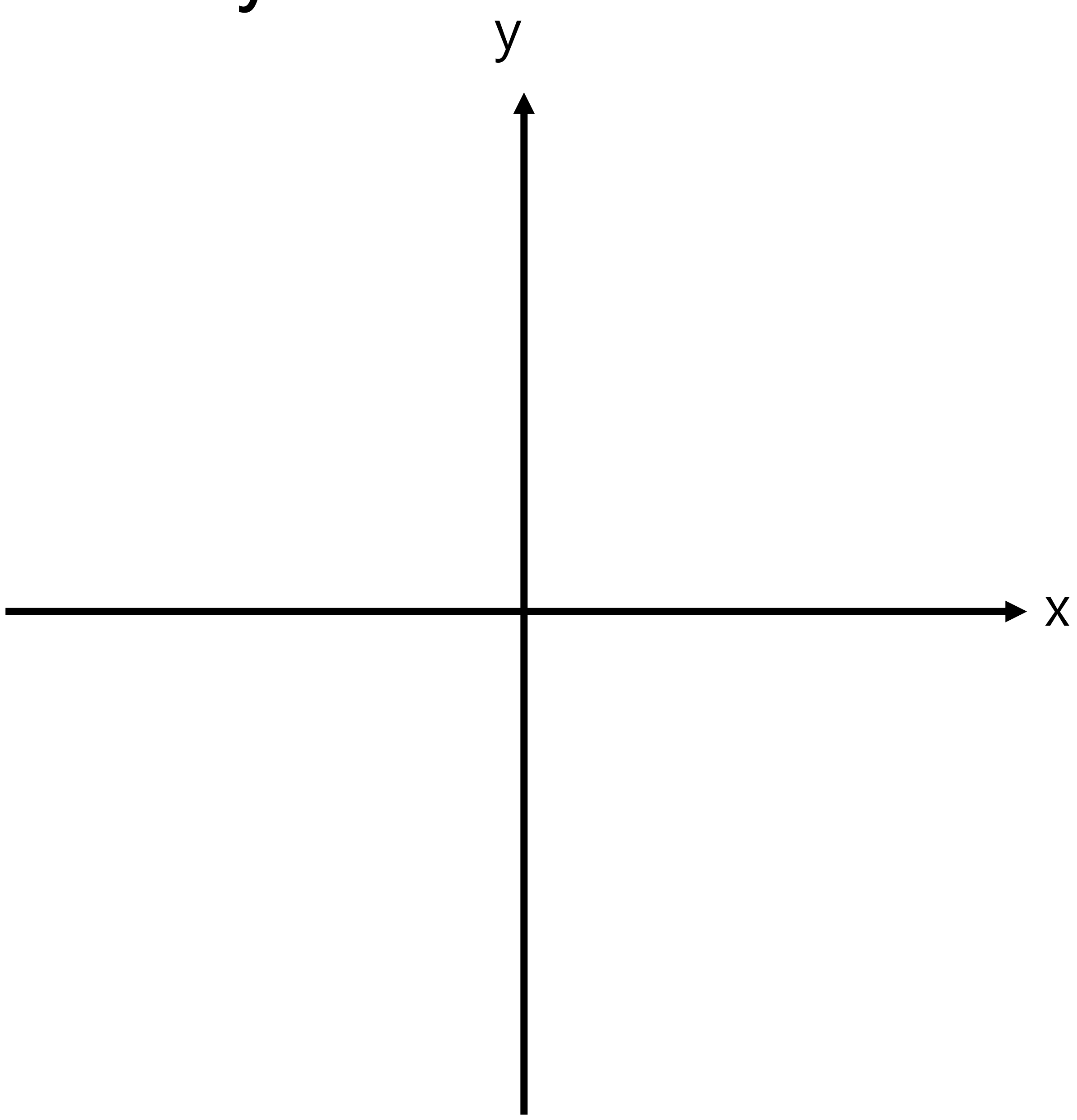
Dot product of vector, $a \cdot b$



Geometrical Meaning of Dot Product

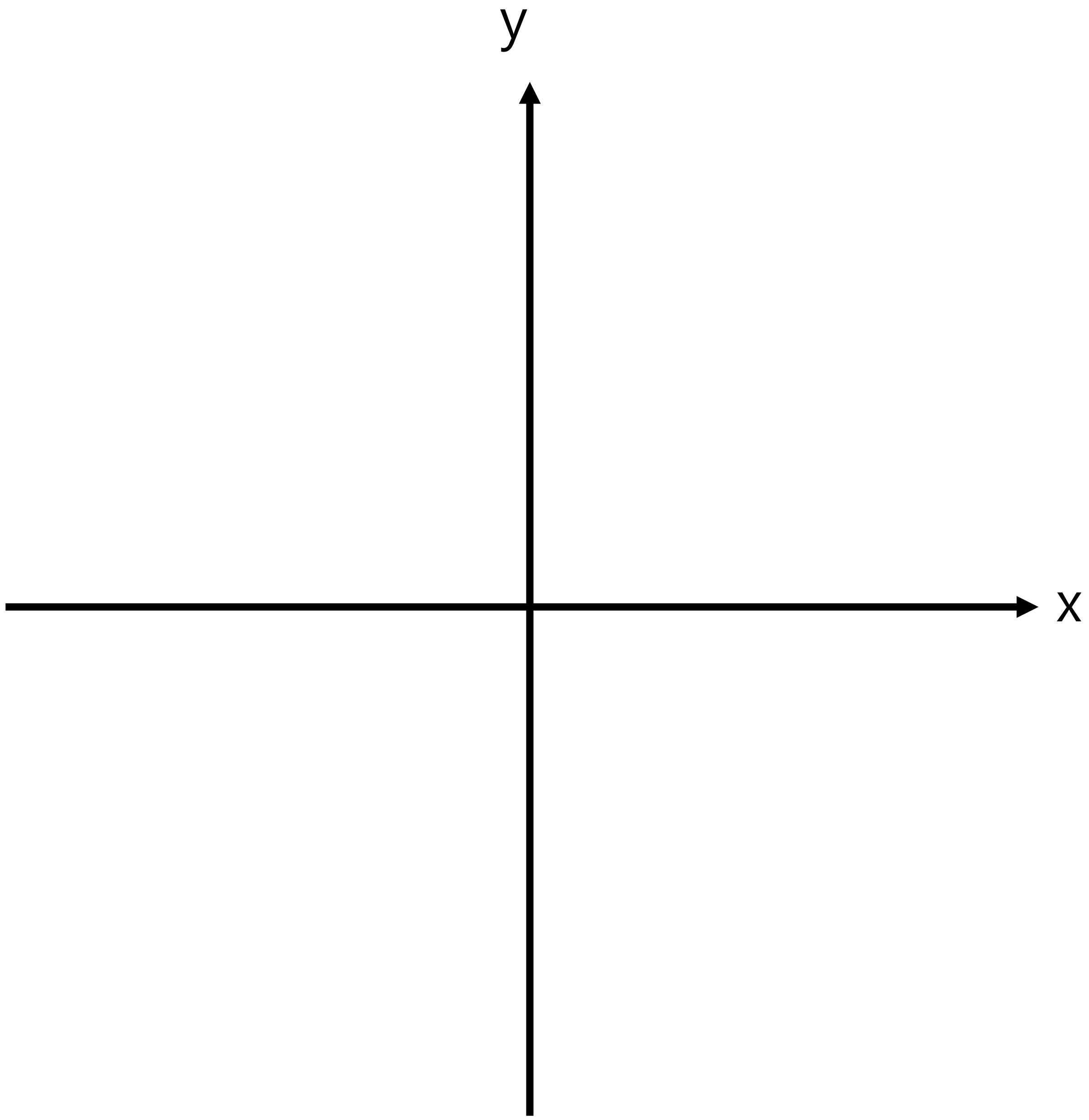


Dot Product as Similarity Index

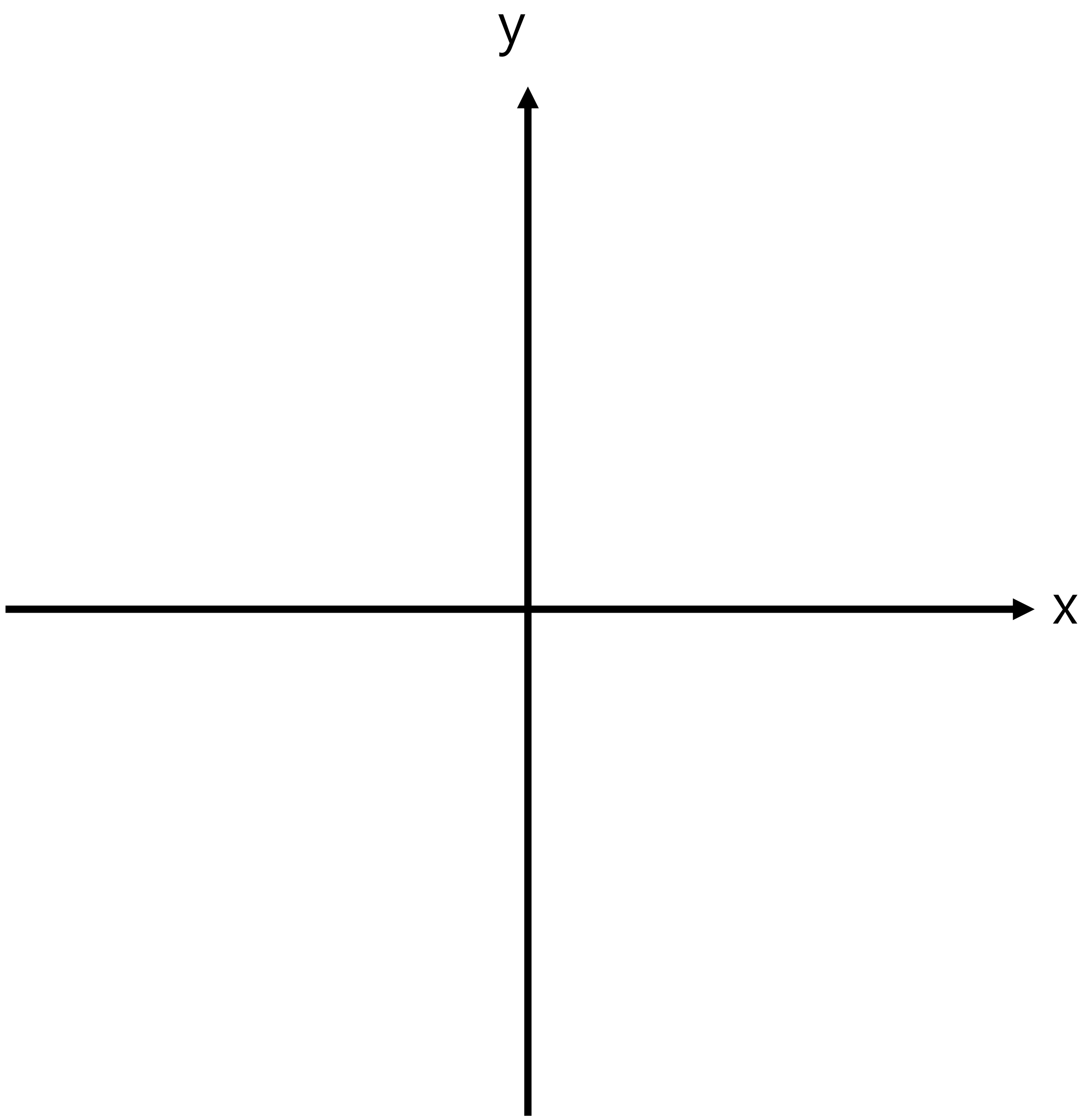


(Advanced) Dot Product Attention

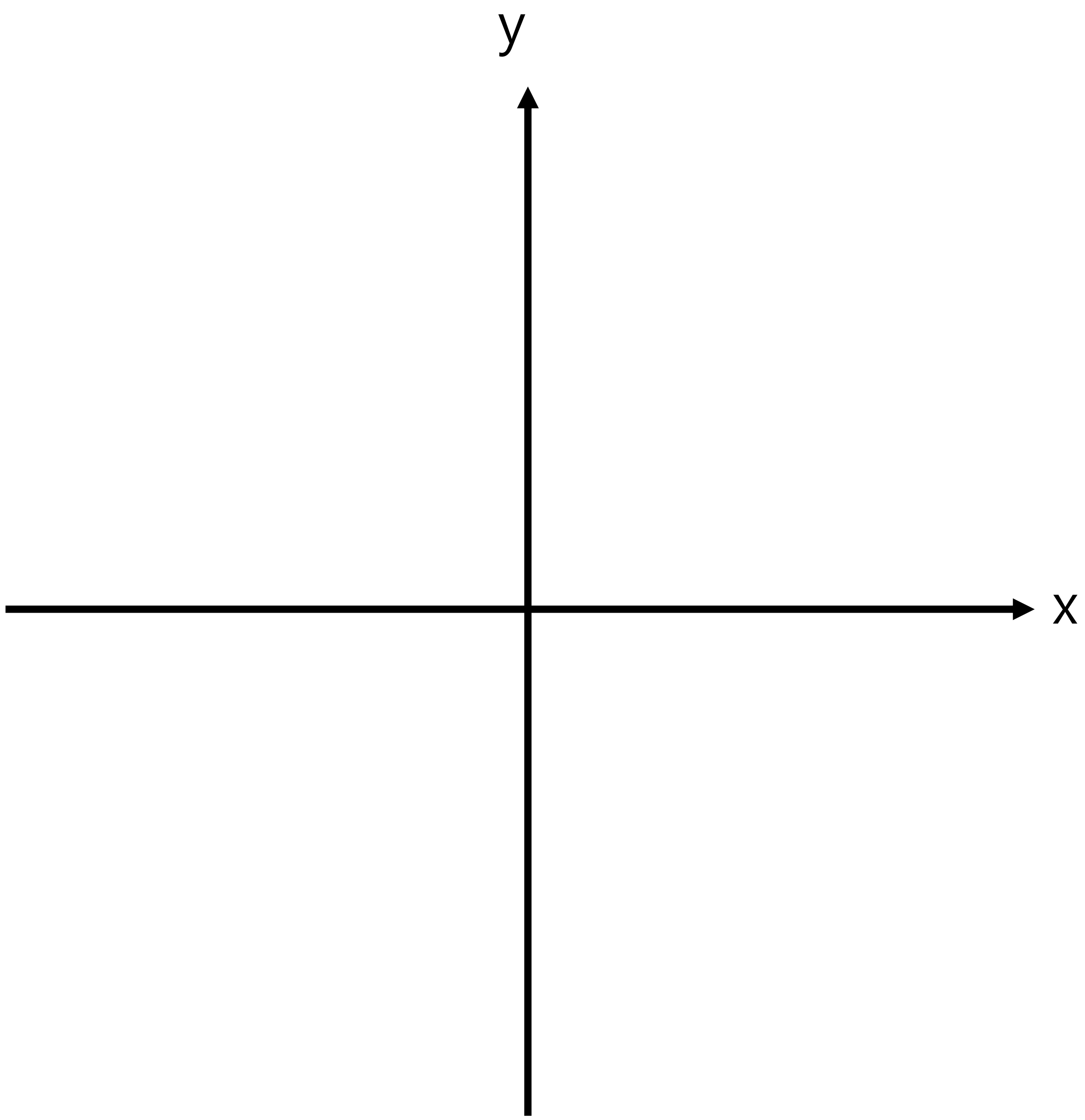
Rank, Null-Space, Determinant, Inverse



Cross product, $a \times b$



Cross product vs Determinant



Eigenvalue & Eigenvector, $Ax = \lambda x$

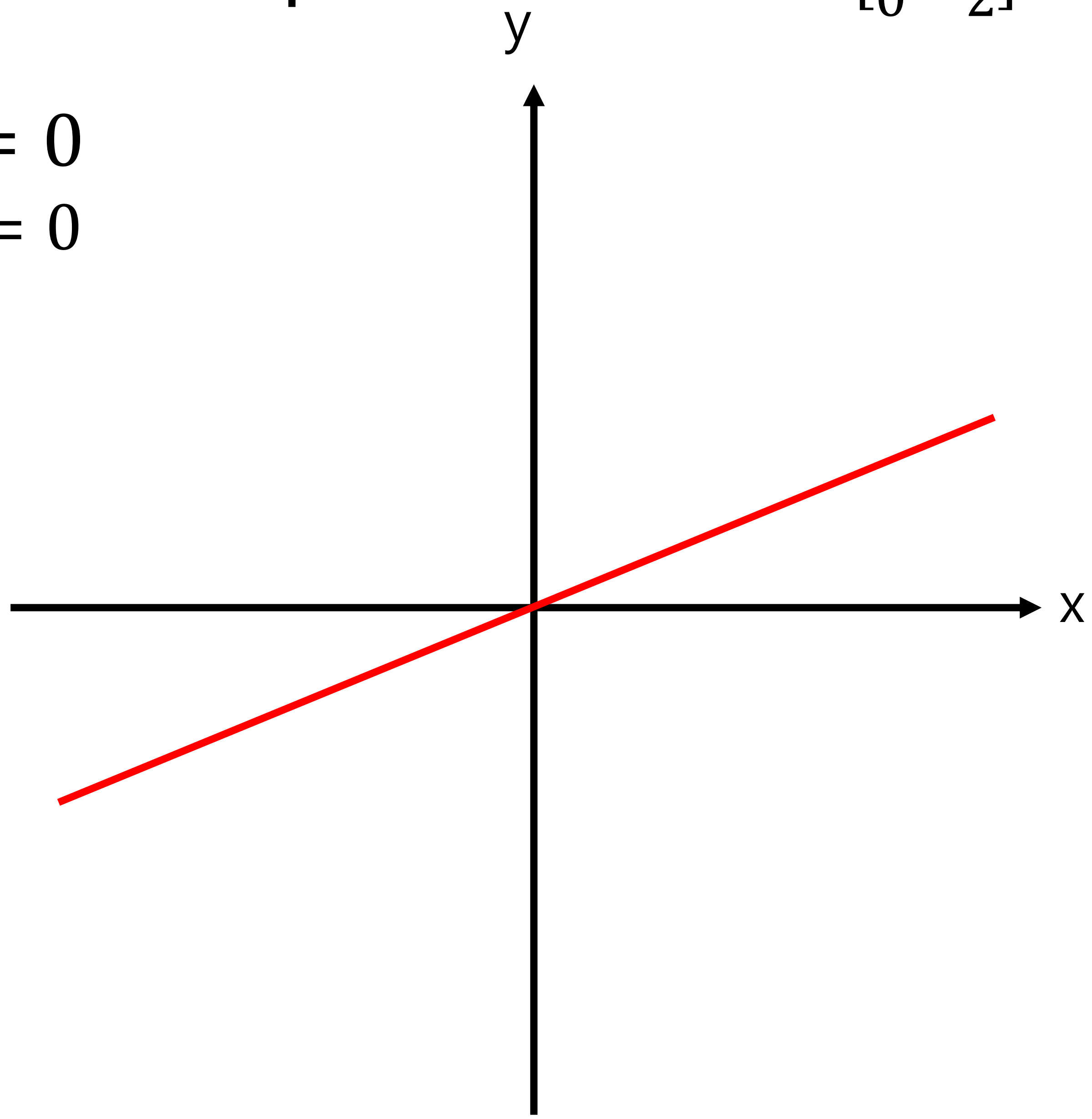
Finding Eigenvalue & Eigenvector

$$A = \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix}$$

Eigenvalue: Span & Off-Span

$$A = \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix}$$

$$Ax - \lambda x = 0$$
$$(A - \lambda I)x = 0$$



Eigenvalue: Geometrical Meaning

$$Ax - \lambda x = 0$$
$$(A - \lambda I)x = 0$$

Eigenvalue: Imaginary Number

(Example) Rotation Matrix

Useful Formulas

$$tr(A) = \sum \lambda$$

$$\det(A) = \prod \lambda$$

Matrix Diagonalization, $A = PDP^{-1}$

$$Av_1 = \lambda_1 v_1$$
$$Av_2 = \lambda_2 v_2$$
$$\dots$$

Principal Component Analysis (PCA)

- Method of ‘dimension reduction’

- Data compression, Noise elimination
- Inappropriate for highly nonlinear dataset

Principal Component Analysis (PCA)

- Step 1: Normalization

Subtract the mean, then divide by the standard deviation

Principal Component Analysis (PCA)

- Step 2: Covariance Matrix

$$V = \frac{1}{n-1}\,(X - \bar{X})^{\top}(X - \bar{X})$$

Principal Component Analysis (PCA)

- Step 3: Eigenvalue Decomposition

Calculate eigenvalue and eigenvector of the covariance matrix

Principal Component Analysis (PCA)

- Step 4: Select Principal Components

(ex) Select top m eigenvectors

Principal Component Analysis (PCA)

- Step 5: Use the result to reduce the dimension

Matrix Exponentials

$$A = \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix}$$

$$e^{A} = ?$$

Matrix Exponentials

(Definition)

$$e^{A} = A + \frac{1}{2!}A^2 + \frac{1}{3!}A^3 + \frac{1}{4!}A^4 + \cdots$$

Matrix Exponentials

(Properties)

$$e^{0} = I$$

$$e^{aA+bB} = e^{aA}e^{bB}$$

$$\frac{d}{dt}e^{At} = Ae^{tA}$$

$$e^{A} = Pe^{D}P^{-1}$$

Solution of Linear Systems

$$\frac{d}{dt}x(t) = Ax(t)$$

Pseudo-Inverse and Least-Squares

$$Ax = b$$
$$x = A^{-1}b$$

What happens if A is rectangular?

Pseudo-Inverse and Least-Squares

$$Ax = b$$
$$A^\top Ax = A^\top b$$
$$x = (A^\top A)^{-1}A^\top b$$