NYU

Linear Algebra

Lecture 14

Orthogonal Bases

Dr. Ralph Chikhany

Strang Sections 4.4 – Orthonormal Bases and Gram-Schmidt

Course notes adapted from N. Hammoud's NYU lecture notes.

Orthogonal Matrices

Orthogonal and Orthonormal Vectors

The vectors $\vec{q}_1, \dots, \vec{q}_n$ are orthogonal if

$$\vec{q}_i \cdot \vec{q}_j = \vec{q}_i^T \vec{q}_j = 0 \qquad (i \neq j)$$

$q_i \perp q_j \quad i \neq j$

The vectors $\vec{q}_1, \dots, \vec{q}_n$ are orthonormal if

$$\vec{q}_i^T \vec{q}_j = 0 \qquad (i \neq j)$$

"Orthogonal" means "Rotate"

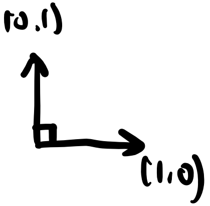
$$||q_i|| = 1$$

q_i is a unit vector

basis is not unique

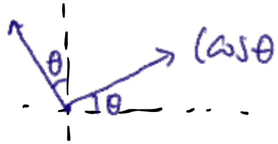
$\mathbb{R}^2$, $\{[0,1], [1,0]\}$

$\{[0,1], [1,1]\}$



all the orthogonal basis for $\mathbb{R}^2$

$q_2 = (-\sin\theta, \cos\theta)$



$(\cos\theta, \sin\theta) = q_1 \leftarrow q_1$ is a unit vector

$q_1 = (\cos\theta, \sin\theta)$

$q_2 = (-\sin\theta, \cos\theta)$

Matrices with Orthonormal Columns

A matrix that has orthonormal columns is denoted by Q, where

$$\underline{Q^T Q = I}$$ doesn't mean $QQ^T \neq I$

only case Q is a square matrix (Q is orthogonal Matrix)

$$Q = [\vec{q}_1 \ \vec{q}_2 \ldots \vec{q}_n] \quad \Longrightarrow \quad Q^T = \begin{bmatrix} \vec{q}_1^T \\ \vec{q}_2^T \\ \vdots \\ \vec{q}_n^T \end{bmatrix}$$

$$Q^T Q = \begin{bmatrix} \vec{q}_1^T \\ \vec{q}_2^T \\ \vdots \\ \vec{q}_n^T \end{bmatrix} [\vec{q}_1 \ \vec{q}_2 \ \cdots \ \vec{q}_n]$$

$$= \begin{bmatrix} \vec{q}_1 \cdot \vec{q}_1 & \vec{q}_1 \cdot \vec{q}_2 & \cdots & \vec{q}_1 \cdot \vec{q}_n \\ \vec{q}_2 \cdot \vec{q}_1 & \vec{q}_2 \cdot \vec{q}_2 & \cdots & \vec{q}_2 \cdot \vec{q}_n \\ \vdots & & & \\ \vec{q}_n \cdot \vec{q}_1 & \vec{q}_n \cdot \vec{q}_2 & \cdots & \vec{q}_n \cdot \vec{q}_n \end{bmatrix} = \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{bmatrix} = I_n$$

off diag : $q_i^T q_j \overset{=0}{} \quad i \neq j$

diag : $q_i^T q_i \overset{=1}{} \quad i = 1, \ldots, n.$

Orthogonal Matrices

If Q is a square matrix with orthonormal columns, then Q is called an orthogonal matrix. In this case $Q^TQ = I$ and $QQ^T = I$.

Q is invertible with $Q^{-1} = Q^T$

Orthogonal and Orthonormal Bases

Orthogonal Bases

A set of vectors $\{\vec{q}_1, \ldots, \vec{q}_n\}$ is called an orthogonal basis of a vector space V if $\vec{q}_1, \ldots, \vec{q}_n$ are orthogonal and they span V.

①

Theorem : $\{\vec{q}_1, \ldots, \vec{q}_n\}$ is an orthogonal set of nonzero vectors in $\mathbb{R}^m$, then $\vec{q}_1, \ldots, \vec{q}_n$ are linearly independent and they form a basis for the subspace $S = \text{span}\{\vec{q}_1, \ldots, \vec{q}_n\}$.

① $q_i^T q_j = 0$, $q_i^T q_i = 1$ $\Rightarrow$ ② $q_1 \cdots q_n$ are linear independent!

all solution of $c_1 q_1 + c_2 q_2 + \cdots + c_n q_n = \vec{0}$ is $c_1 = c_2 = \cdots = c_n = 0$

$$\vec{q}_i^T (c_1 \vec{q}_1 + c_2 \vec{q}_2 + \cdots + c_n \vec{q}_n) = \vec{q}_i^T \vec{0} = 0 \quad \Rightarrow c_i = 0 \text{ for all } i$$

$$\| $$

$$c_1 \underbrace{\vec{q}_i \cdot \vec{q}_1}_{=0} + c_2 \underbrace{\vec{q}_i \cdot \vec{q}_2}_{=0} + \cdots + c_i \underbrace{\vec{q}_i \cdot \vec{q}_i}_{=1} + \cdots + c_n \underbrace{\vec{q}_i \cdot \vec{q}_n}_{=0}$$

$$\|$$

$$0 \cdot c_1 + 0 \cdot c_2 + \cdots + 1 \cdot c_i + \cdots + 0 \cdot c_n$$

$$\| $$

$$c_i$$

Theorem of Coefficients

Let $\{\vec{q}_1, \ldots, \vec{q}_n\}$ be an orthogonal basis for a subspace $S \subset \mathbb{R}^m$. For each $\vec{v} \in S$,

if $\vec{v} \in \operatorname{span}\{\vec{q}_1 \cdots \vec{q}_n\}$

$$\vec{v} = c_1\vec{q}_1 + c_2\vec{q}_2 + \cdots + c_n\vec{q}_n$$

$\in$ L. C. of $q_1 \ldots q_n$

$\vec{q}_i^{\,T}\vec{q}_i = 1$

with $c_i = \cancel{\dfrac{\vec{q}_i^{\,T}\vec{v}}{\vec{q}_i^{\,T}\vec{q}_i}}$ for $1 \le i \le n$.

$c_i = \vec{q}_i^{\,T} \cdot \vec{v}$

$$\vec{q}_i^{\,T} \cdot \vec{v} = \vec{q}_i \cdot \left(c_1\vec{q}_1 + c_2\vec{q}_2 + \cdots + c_n\vec{q}_n \right)$$

$$= c_1 \vec{q}_i\vec{q}_1 + \cdots + c_i \vec{q}_i\vec{q}_i + \cdots + c_n \vec{q}_n^{\,T}\vec{q}_n$$

$$= c_i$$

Projection to a Orthogonal Basis.

$q_1 \cdots q_n$ are orthogonal,

$Q = [q_1 \cdots q_n] \Rightarrow Q^T Q = I$

Project vector b to Q

holds even $b \notin \text{span}\{q_1 \cdots q_n\}$

Coefficients : $\underbrace{(Q^T Q)}_{I}^{-1} Q^T b = Q^T b$

$$= \begin{bmatrix} \vec{q}_1^T \\ \vdots \\ \vec{q}_n^T \end{bmatrix} \vec{b} = \begin{bmatrix} \vec{q}_1 \cdot \vec{b} \\ \vec{q}_2 \cdot \vec{b} \\ \vdots \\ \vec{q}_n \cdot \vec{b} \end{bmatrix}$$

Vector : $Q \underbrace{(Q^T Q)}_{I}^{-1} Q^T b = \underbrace{Q Q^T}_{\text{the projection matrix}} b$

$$= Q \cdot \begin{bmatrix} \vec{q}_1 \cdot \vec{b} \\ \vdots \\ \vec{q}_n \cdot \vec{b} \end{bmatrix} = [\vec{q}_1 \cdots \vec{q}_n] \begin{bmatrix} \vec{q}_1 \cdot \vec{b} \\ \vdots \\ \vec{q}_n \cdot \vec{b} \end{bmatrix}$$

$$= \underbrace{(\vec{q}_1 \cdot \vec{b})}_{\text{coef}} \underbrace{\vec{q}_1}_{\text{vector}} + (\vec{q}_2 \cdot \vec{b}) \vec{q}_2 + \cdots + (\vec{q}_n \cdot \vec{b}) \vec{q}_n$$

Project to Orthogonal Matrix, $P = Q Q^T$

$$Q Q^T b = (\vec{q}_1 \cdot \vec{b}) \vec{q}_1 + (\vec{q}_2 \cdot \vec{b}) \vec{q}_2 + \cdots + (\vec{q}_n \cdot \vec{b}) \vec{q}_n$$

"orthogonal" is "rotate"

Thm. Q is an orthogonal Matrix (even if Q is not a square matrix)

$\|Qx\| = \|x\|$ $\quad$ angle between $\angle(Qx, Qy) = \angle(x, y)$

same proof

$$\|Qx\|^2 = (Qx)^T Qx$$

$$= x^T \underbrace{Q^T Q}_{I} x$$

$$= x^T x = \|x\|^2$$

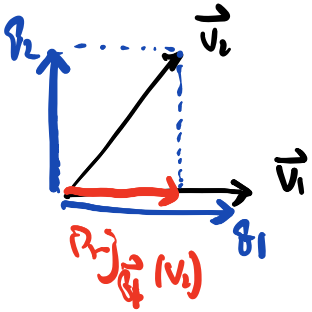
Gram-Schmidt

The Gram-Schmidt Process

Consider a vector space V with basis $\beta_V = \{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$

Gram-Schmidt (G-S) turns β_V into an orthogonal basis $\{\vec{q}_1, \vec{q}_2, \ldots, \vec{q}_n\}$ by using projections.

2-dimension.



G-S $\quad \{\vec{v}_1, \vec{v}_2\} \to \{\vec{q}_1, \vec{q}_2\}$

$$\vec{q}_2 = \vec{v}_2 - \text{Proj}_{\vec{q}_1}(\vec{v}_2) \Rightarrow \begin{cases} \vec{q}_1 = \vec{v}_1 + 0 \cdot \vec{v}_2 \\ \vec{q}_2 = \vec{v}_2 - c \cdot \vec{v}_1 \end{cases}$$

$$[\vec{q}_1, \vec{q}_2] = [\vec{v}_1 \ \vec{v}_2] \begin{bmatrix} 1 & -c \\ 0 & 1 \end{bmatrix} \qquad c = \frac{\vec{v}_2 \cdot \vec{v}_1}{\vec{v}_1 \cdot \vec{v}_1}$$

orthogonal matrix   upper traingular

$$[\vec{v}_1 \ \vec{v}_2] = \underbrace{[\vec{q}_1 \ \vec{q}_2]}_{Q} \underbrace{\begin{bmatrix} 1 & -c \\ 0 & 1 \end{bmatrix}^{-1}}_{\text{upper traingular matrix}}$$

The Gram-Schmidt Process

Consider a vector space V with basis $\beta_V = \{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$

Gram-Schmidt (G-S) turns β_V into an orthogonal basis $\{\vec{q}_1, \vec{q}_2, \ldots, \vec{q}_n\}$ by using projections.

$$
\begin{aligned}
\vec{q}_1 &= \vec{v}_1 \\
\vec{q}_2 &= \vec{v}_2 - \vec{p}_{21} = \vec{v}_2 - \frac{\vec{q}_1^{\,T}\vec{v}_2}{\vec{q}_1^{\,T}\vec{q}_1}\vec{q}_1 \\
\vec{q}_3 &= \vec{v}_3 - \vec{p}_{31} - \vec{p}_{32} \\
&= \vec{v}_3 - \frac{\vec{q}_1^{\,T}\vec{v}_3}{\vec{q}_1^{\,T}\vec{q}_1}\vec{q}_1 - \frac{\vec{q}_2^{\,T}\vec{v}_3}{\vec{q}_2^{\,T}\vec{q}_2}\vec{q}_2 \\
&\vdots \\
\vec{q}_n &= \vec{v}_n - \vec{p}_{n1} - \vec{p}_{n2} - \cdots - \vec{p}_{n(n-1)}
\end{aligned}
$$

span $\{q_1, q_2\}$ = span $\{v_1, v_2\}$

projection to span $\{q_1, q_2\}$

q_1, q_2 are orthogonal basis

So, it's easy to compute the projection

Example

Another Example

Find an orthogonal basis $\{u_1, u_2, u_3\}$ for $W = \mathrm{Span}\{v_1, v_2, v_3\} = \mathbf{R}^3$, where

$$v_1 = \begin{pmatrix}1\\1\\0\end{pmatrix} \qquad v_2 = \begin{pmatrix}1\\1\\1\end{pmatrix} \qquad v_3 = \begin{pmatrix}3\\1\\1\end{pmatrix}. \qquad \left\{ \begin{pmatrix}1\\1\\0\end{pmatrix}, \begin{pmatrix}0\\0\\1\end{pmatrix}, \begin{pmatrix}1\\-1\\0\end{pmatrix} \right\}$$

$$\vec{q}_1 = \vec{v}_1 = \begin{pmatrix}1\\1\\0\end{pmatrix}$$

They are not unit vector!

$$\vec{q}_2 = \vec{v}_2 - \mathrm{proj}_{q_1}(\vec{v}_2) = \begin{pmatrix}1\\1\\1\end{pmatrix} - \frac{\begin{pmatrix}1\\1\\1\end{pmatrix}\cdot\begin{pmatrix}1\\1\\0\end{pmatrix}}{\begin{pmatrix}1\\1\\0\end{pmatrix}\cdot\begin{pmatrix}1\\1\\0\end{pmatrix}} \underbrace{\begin{pmatrix}1\\1\\0\end{pmatrix}}_{q_1} = \begin{pmatrix}1\\1\\1\end{pmatrix} - \frac{2}{2}\begin{pmatrix}1\\1\\0\end{pmatrix} = \begin{pmatrix}0\\0\\1\end{pmatrix}$$

we normalize the legth of q_1

project to v_1

$$\vec{q}_3 = \vec{v}_3 - \mathrm{proj}_{q_1, q_2}(\vec{v}_3)$$

$$= \begin{pmatrix}3\\1\\1\end{pmatrix} - \frac{\overbrace{\begin{pmatrix}3\\1\\1\end{pmatrix}\cdot\begin{pmatrix}1\\1\\0\end{pmatrix}}^{4}}{\begin{pmatrix}1\\1\\0\end{pmatrix}\cdot\underbrace{\begin{pmatrix}1\\1\\0\end{pmatrix}}_{q_1}}\underbrace{\begin{pmatrix}1\\1\\0\end{pmatrix}}_{q_1} - \frac{\overbrace{\begin{pmatrix}3\\1\\1\end{pmatrix}\begin{pmatrix}0\\0\\1\end{pmatrix}}^{1}}{\begin{pmatrix}3\\1\\1\end{pmatrix}\underbrace{\begin{pmatrix}0\\0\\1\end{pmatrix}}_{q_2}}\underbrace{\begin{pmatrix}0\\0\\1\end{pmatrix}}_{q_2} = \begin{pmatrix}1\\-1\\0\end{pmatrix}$$

projection to span (q_1, q_2)

Example

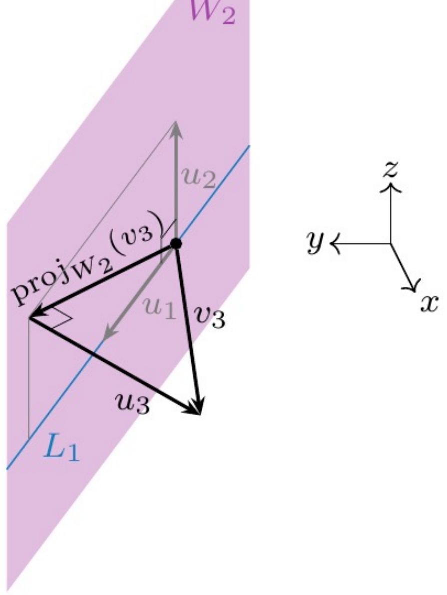
Find an orthogonal basis $\{u_1, u_2, u_3\}$ for $W = \mathrm{Span}\{v_1, v_2, v_3\} = \mathbf{R}^3$, where

$$v_1 = \begin{pmatrix}1\\1\\0\end{pmatrix},\ v_2 = \begin{pmatrix}1\\1\\1\end{pmatrix},\ v_3 = \begin{pmatrix}3\\1\\1\end{pmatrix} \xrightarrow{\text{G–S}} u_1 = \begin{pmatrix}1\\1\\0\end{pmatrix},\ u_2 = \begin{pmatrix}0\\0\\1\end{pmatrix},\ u_3 = \begin{pmatrix}1\\-1\\0\end{pmatrix}$$

They are not unit vectors!!

Why does this work?

- Once we have u_1 and u_2, then we're sad because v_3 is not orthogonal to u_1 and u_2.
- Fix: let $W_2 = \mathrm{Span}\{u_1, u_2\}$, and let $u_3 = (v_3)_{W_2^\perp} = v_3 - \mathrm{proj}_{W_3}(u_3)$.
- By construction, $u_1 \cdot u_3 = 0 = u_2 \cdot u_3$ because $W_2 \perp u_3$.



Check:

$u_1 \cdot u_2 = 0$ ✓

$u_1 \cdot u_3 = 0$ ✓

$u_2 \cdot u_3 = 0$ ✓

QR Factorization

The QR Factorization

Given an $m \times n$ matrix $A = [\vec{v}_1 \ \vec{v}_2 \ \dots \ \vec{v}_n]$, such that $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$ are linearly independent. Then, we can factorize A as

$R = Q^T A$

$$A = QR \rightarrow \text{upper traingular Matrix}$$

orthogonal matrix

Existence : Show in (✱) Slide of G-S

Q from G-S process.

$$Q^T A = \underbrace{Q^T Q}_{= I} R = R$$

The QR Factorization

Given an $m \times n$ matrix $A = [\vec{v}_1 \; \vec{v}_2 \; \dots \; \vec{v}_n]$, such that $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$ are linearly independent. Then, we can factorize A as

$$A = QR$$

Finding Q: Let $Q = [\vec{q}_1 \; \vec{q}_2 \; \dots \; \vec{q}_n]$

To find $\vec{q}_1, \vec{q}_2, \dots, \vec{q}_n$, we use G-S on the vectors $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$, then make them orthonormal by dividing each $\vec{q}_i$ by its magnitude.

Finding R: $A = QR \implies$ multiply both sides by Q^T

$$\implies Q^T A = Q^T Q R \implies R = Q^T A$$

Example

$$v_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \quad v_2 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \quad v_3 = \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix} \xrightarrow{G-S} q_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \quad q_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad q_3 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$$

are not unit vector!

change each vector to unit vector

$$Q = \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \end{bmatrix}$$

Q is an orthogonal matrix $\quad Q^T Q = I$

$$R = Q^T V = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \end{bmatrix} \begin{bmatrix} v_1 & v_2 & v_3 \end{bmatrix}$$

check by your self

it's an upper Triangular Matrix

Why QR $\rightarrow$ least square or Projection Much easier.

$Ax = b$ may have no Solution

find Solution $\hat{x}$ such that

$\| A\hat{x} - b \|_2^2$ by solve $\min_x \| Ax - b \|_2^2$

$A = QR$

$\text{Col}(Q) = \text{Col}(A)$

Projection of b to the Col(A) $\Rightarrow$ $Q Q^T b$

Claim. $\hat{x}$ can be solve easily by $Rx = Q^T b$ (not require)

The size of adapt to the dimension of the space!

$QRx = QQ^Tb \Rightarrow Rx = Q^T b$

$A \in \mathbb{R}^{3\times 2} \xrightarrow{G-S} Q \in \mathbb{R}^{3\times 2}$ 2 vectors in 3-dimension

$R = \underbrace{Q^T}_{2\times 3} \underbrace{A}_{3\times 2} \rightarrow 2\times 2$ square matrix

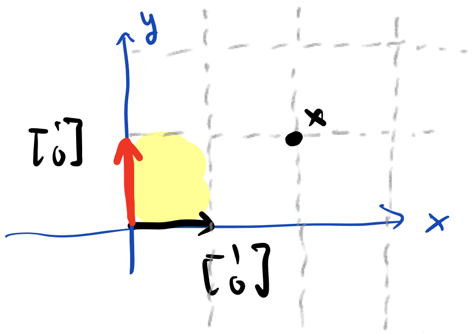
$A = [\,|\ \ |\,]$

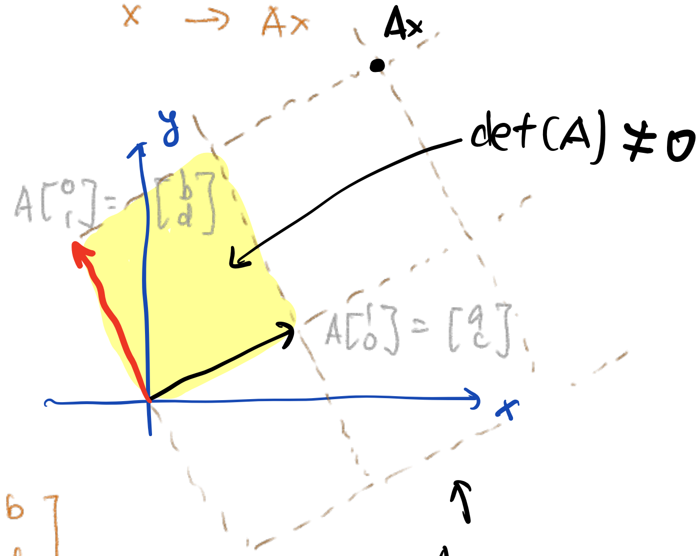
$A \in \mathbb{R}^{2\times 2}$ $\qquad$ $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$

$\mathbb{R}^{2\times 2}$ matrix is a linear transform from $\mathbb{R}^2 \to \mathbb{R}^2$

$x \to Ax$

$A\begin{bmatrix}1\\0\end{bmatrix} = \begin{bmatrix}a\\c\end{bmatrix}$ $\qquad$ $A\begin{bmatrix}0\\1\end{bmatrix} = \begin{bmatrix}b\\d\end{bmatrix}$



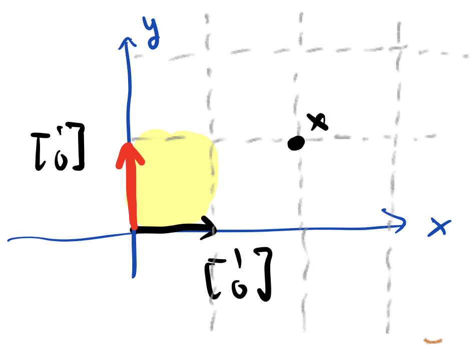


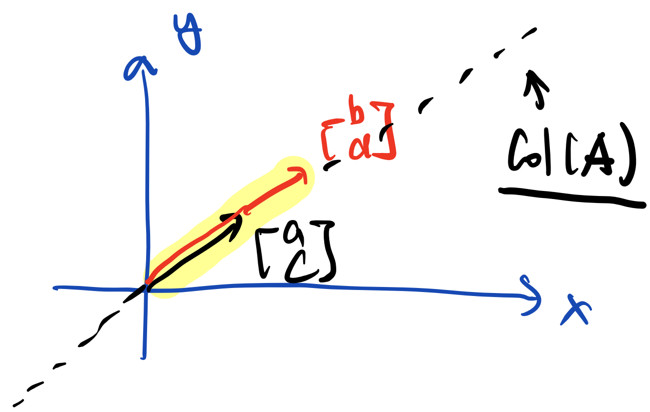
A is invertible

$A\begin{bmatrix}c_1\\c_2\end{bmatrix} = c_1\begin{bmatrix}a\\c\end{bmatrix} + c_1\begin{bmatrix}b\\d\end{bmatrix}$

$A\left(c_1\begin{bmatrix}1\\0\end{bmatrix} + c_2\begin{bmatrix}0\\1\end{bmatrix}\right) \to c_1\left(A\begin{bmatrix}1\\0\end{bmatrix}\right) + c_2\left(A\begin{bmatrix}0\\1\end{bmatrix}\right)$

preserve the linear Relation! ← linear transformation.





det(A) = 0 $\qquad$ A is not invertible.

$-\det(\otimes) = 0$