Session 2: Linear transformations and matrices

Optimization and Computational Linear Algebra for Data Science

Léo Miolane

Contents

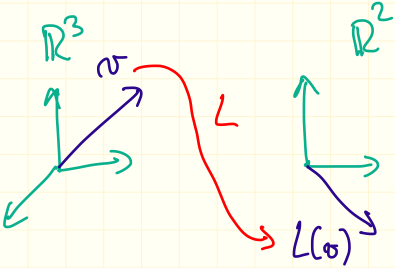
Questions?

Linear maps & matrices

Two sides of the same coin

Linear map | function / transformation

$$L : \mathbb{R}^m \to \mathbb{R}^n$$



$$\begin{cases} L(u+v) = L(u) + L(v) \\ L(\alpha u) = \alpha L(v) \end{cases}$$

for all $\alpha \in \mathbb{R}$, $u, v \in \mathbb{R}^m$

Matrix

$$\tilde{L} \in \mathbb{R}^{n\times m}$$

$$\tilde{L} = \begin{array}{cc} \overset{m}{\longleftrightarrow} & \\ \begin{pmatrix} l_{11} & \cdots & l_{1,m} \\ \vdots & & \vdots \\ l_{n,1} & \cdots & l_{n,m} \end{pmatrix} & \updownarrow n \end{array}$$

Two sides of the same coin

Linear map

$L : \mathbb{R}^m \to \mathbb{R}^n$

$\boxed{L(x) = \tilde{L}x}$ "matrix vector product"

for all $x \in \mathbb{R}^m$

(1)

Matrix

$\tilde{L} \in \mathbb{R}^{n\times m}$

$$\tilde{L} = \begin{array}{c} \overset{m}{\longleftrightarrow} \\ \begin{pmatrix} | & & | \\ L(e_1) & \cdots & L(e_m) \\ | & & | \end{pmatrix} \end{array} \updownarrow n$$

(2)

$L : \mathbb{R}^m \to \mathbb{R}^n$

$x \mapsto \tilde{L}x$

L is linear

$$\tilde{L} = \begin{pmatrix} l_{11} & \cdots & l_{1,m} \\ \vdots & & \vdots \\ l_{n,1} & \cdots & l_{n,m} \end{pmatrix}$$

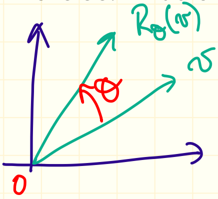
$$\begin{pmatrix} | & & | \\ L(e_1) & \cdots & L(e_m) \\ | & & | \end{pmatrix}$$

$$L(e_1) = \tilde{L}\, e_1 = \begin{pmatrix} l_{11} & \cdots & l_{1,m} \\ \vdots & & \vdots \\ l_{n,1} & \cdots & l_{n,m} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} l_{11} \\ \vdots \\ l_{n,1} \end{pmatrix}$$

Rotations in $\mathbb{R}^2$

Let $\theta \in \mathbb{R}$. The rotation $R_\theta : \mathbb{R}^2 \to \mathbb{R}^2$ of angle θ about the origin is linear.

Exercise: what is the canonical matrix of R_θ ?

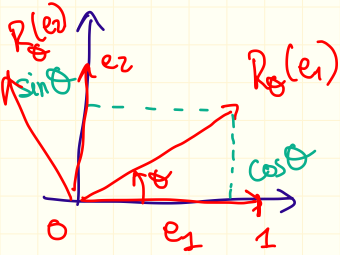


2

$$R_\theta = \begin{pmatrix} | & | \\ R_\theta(e_1) & R_\theta(e_2) \\ | & | \end{pmatrix}$$

2

$$R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$



$$\underline{R_\theta\left(\begin{pmatrix} x \\ y \end{pmatrix}\right)} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix}$$

$v = \begin{pmatrix} x \\ y \end{pmatrix}$ $\quad R_\theta(v)$

Operations on matrices

Addition and scalar multiplication

- Sum of two matrices of the **same** dimensions:

$$\begin{pmatrix} a_{1,1} & \cdots & a_{1,m} \\ \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & a_{n,m} \end{pmatrix} + \begin{pmatrix} b_{1,1} & \cdots & b_{1,m} \\ \vdots & \ddots & \vdots \\ b_{n,1} & \cdots & b_{n,m} \end{pmatrix} = \begin{pmatrix} a_{1,1}+b_{1,1} & \cdots & a_{1,m}+b_{1,m} \\ \vdots & \ddots & \vdots \\ a_{n,1}+b_{n,1} & \cdots & a_{n,m}+b_{n,m} \end{pmatrix}$$

- Multiplication by a scalar λ:

$$\lambda \begin{pmatrix} a_{1,1} & \cdots & a_{1,m} \\ \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & a_{n,m} \end{pmatrix} = \begin{pmatrix} \lambda a_{1,1} & \cdots & \lambda a_{1,m} \\ \vdots & \ddots & \vdots \\ \lambda a_{n,1} & \cdots & \lambda a_{n,m} \end{pmatrix}$$

A new vector space!

Proposition

- $\mathbb{R}^{n\times m}$ is a vector space.
- $\dim(\mathbb{R}^{n\times m}) =$ $n.m$

Proof. sketch

We can verify that the matrices

$$E_{i,j} = \begin{array}{cc} & \begin{array}{ccc} & & \downarrow j \end{array} \\ & \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & 1 & \vdots \\ 0 & \cdots & 0 \end{pmatrix} \leftarrow i \end{array}, \text{ for } \begin{array}{l} 1 \leq i \leq n \\ 1 \leq j \leq m \end{array}$$

form a basis of $\mathbb{R}^{n\times m}$, with $n.m$ elements.

□

Product of two matrices

Warning:

$$\begin{pmatrix} a_{1,1} & \cdots & a_{1,m} \\ \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & a_{n,m} \end{pmatrix} \times \begin{pmatrix} b_{1,1} & \cdots & b_{1,m} \\ \vdots & \ddots & \vdots \\ b_{n,1} & \cdots & b_{n,m} \end{pmatrix} \neq \begin{pmatrix} a_{1,1} \times b_{1,1} & \cdots & a_{1,m} \times b_{1,m} \\ \vdots & \ddots & \vdots \\ a_{n,1} \times b_{n,1} & \cdots & a_{n,m} \times b_{n,m} \end{pmatrix}$$

We will never consider such notion of product because it doesn't have any "geometrical meaning"

Matrix product

Let $L \in \mathbb{R}^{n\times m}$ and $M \in \mathbb{R}^{m\times k}$.

Definition (Matrix product)

The matrix product LM is the $n \times k$ matrix of the linear map $L \circ M$.

$$\begin{tikzcd}
\mathbb{R}^k \arrow[r, bend left, "M"] \arrow[rr, bend right, "L \circ M"'] & \mathbb{R}^m \arrow[r, bend left, "L"] & \mathbb{R}^n
\end{tikzcd}$$

$L \circ M(x) = L(M(x))$

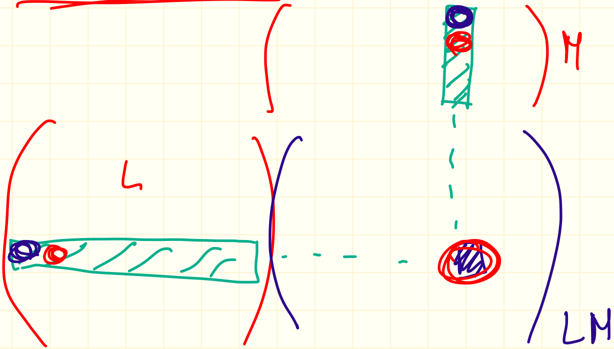
The composition $L \circ M$ is linear.

Matrix product

Theorem

Let $L \in \mathbb{R}^{n\times m}$ and $M \in \mathbb{R}^{m\times k}$.
The entries matrix product LM are given by

$$\boxed{(LM)_{i,j} = \sum_{\ell=1}^{m} L_{i,\ell} M_{\ell,j},} \qquad \text{for } 1 \le i \le n \text{ and } 1 \le j \le k.$$



Proof

$$LM \overset{\text{def}}{=} \begin{pmatrix} | & & | \\ L\circ M(e_1) & \dots & L\circ M(e_n) \\ | & & | \end{pmatrix}$$

$$(LM)_{i,j} = \left(L\circ M(e_j)\right)_i \leftarrow i^{\text{th}} \text{ coefficient of } L\circ M(e_j)$$

$$= \left(L\left(\underline{M(e_j)}\right)\right)_i \qquad M(e_j) = \underbrace{\begin{pmatrix} M_{1,j} \\ \vdots \\ M_{m,j} \end{pmatrix}}_{=v}$$

$$= \left(L(v)\right)_i$$

$$= \left(Lv\right)_i$$

(matrix - vector product)

$$= \underline{\sum_{\ell=1}^{m} L_{i,\ell}\, v_\ell} = \boxed{\sum_{\ell=1}^{m} L_{i,\ell}\, M_{\ell,j}}$$

Rotations in $\mathbb{R}^2$

The R_a and R_b denote respectively the matrices of the rotations of angles a and b about the origin, in $\mathbb{R}^2$.

$$R_a = \begin{pmatrix} \cos a & -\sin a \\ \sin a & \cos a \end{pmatrix}$$

Exercise: Compute the product $R_a R_b$.

1st method

$$R_a R_b = \begin{pmatrix} \cos a & -\sin a \\ \sin a & \cos a \end{pmatrix} \cdot \begin{pmatrix} \cos b & -\sin b \\ \sin b & \cos b \end{pmatrix} = \begin{pmatrix} \cos a \cos b - \sin a \sin b & * \\ \cos b \sin a + \cos a \sin b & * \end{pmatrix}$$

2nd method : $R_a R_b$ is the composition of R_b and R_a which is a rotation of angle $a+b$:

$$R_a R_b = R_{a+b} = \begin{pmatrix} \cos(a+b) & -\sin(a+b) \\ \sin(a+b) & \cos(a+b) \end{pmatrix}$$

(As a by product : $\cos(a+b) = \cos a \cos b - \sin a \sin b$)

Matrix product properties

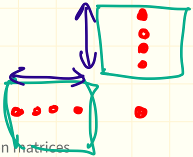
• Let $A, B \in \mathbb{R}^{n \times m}$ and $C, D \in \mathbb{R}^{m \times k}$

- $(A+B) \cdot C = A \cdot C + BC$
- $A(C+D) = AC + AD$

"1"

Recall $Id_n = \begin{pmatrix} 1 & & (0) \\ & \ddots & \\ (0) & & 1 \end{pmatrix}$: $\begin{cases} Id_n A = A \\ A \cdot Id_m = A \end{cases}$

Issue : $\boxed{AC \neq CA}$ if $k \neq n$



Can we divide two matrices ?

For instance if $\cancel{A}B = \cancel{A}C$

do we have $B = C$?

Take : $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ $\quad B = \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix}$ $\quad C = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$

- $AC = 0$ $\qquad \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix}$
- $AB = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$

$AB = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = AC$ but $C \neq B$.

Invertible matrices

Definition (Matrix inverse)

A **square** matrix $M \in \mathbb{R}^{n\times n}$ is called *invertible* if there exists a matrix $M^{-1} \in \mathbb{R}^{n\times n}$ such that

$$MM^{-1} = M^{-1}M = \mathrm{Id}_n.$$

Such matrix M^{-1} is unique and is called the *inverse* of M.

Exercise: Let $A, B \in \mathbb{R}^{n\times n}$. Show that if $AB = \mathrm{Id}_n$ then $BA = \mathrm{Id}_n$.

If $AB = AC$ and A is invertible

then $\underbrace{A^{-1}A}_{\mathrm{Id}_n} B = \underbrace{A^{-1}A}_{\mathrm{Id}_n} C$ hence $B = C$

Kernel and image

Definitions

Let $L : \mathbb{R}^m \to \mathbb{R}^n$ be a linear transformation.

$0 \text{ "=" } \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$
notation

Definition (Kernel)

The kernel $\operatorname{Ker}(L)$ (or nullspace) of L is defined as the set of all vectors $v \in \mathbb{R}^m$ such that $L(v) = 0$, i.e.

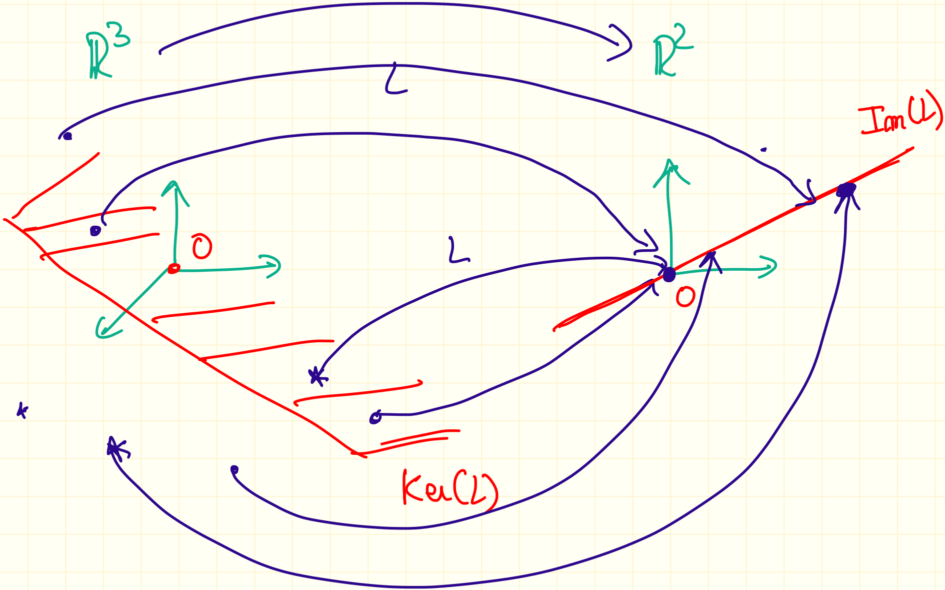
$$\operatorname{Ker}(L) \stackrel{\text{def}}{=} \{v \in \mathbb{R}^m \mid \overbrace{L(v)}^{\in \mathbb{R}^n} = 0\}.$$

zero vector of $\mathbb{R}^n$

Definition (Image)

The image $\operatorname{Im}(L)$ (or column space) of L is defined as the set of all vectors $u \in \mathbb{R}^n$ such that there exists $v \in \mathbb{R}^m$ such that $L(v) = u$.

Picture



Remarks

Let $L : \mathbb{R}^m \to \mathbb{R}^n$ be a linear transformation.

Proposition

- $\mathrm{Ker}(L)$ is a subspace of $\mathbb{R}^m$.
- $\mathrm{Im}(L)$ is a subspace of $\mathbb{R}^n$.

Proved in HW2.

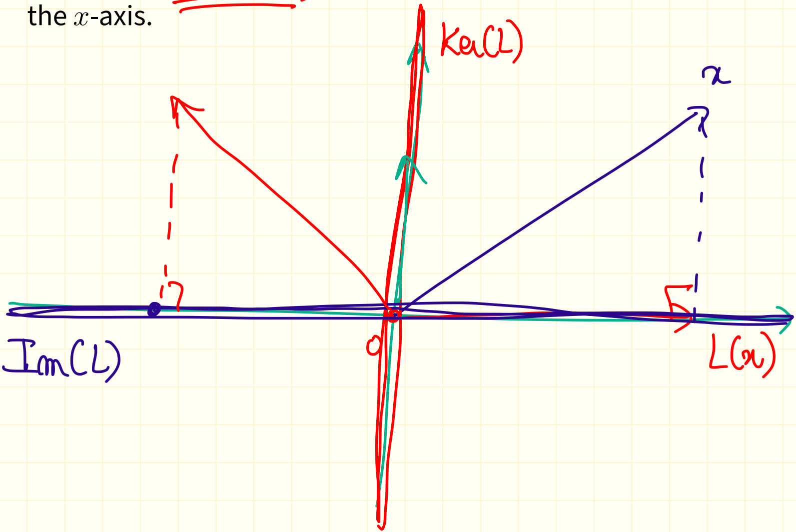
Remark: $\mathrm{Im}(L)$ is also the Span of the columns of the matrix representation of L.

$$L = \begin{pmatrix} | & & | \\ c_1 & \cdots & c_m \\ | & & | \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} = x_1 c_1 + \cdots + x_m c_m$$

$$\mathrm{Im}(L) = \mathrm{Span}(c_1, \ldots, c_m)$$

Example: orthogonal projection

Consider $L : \mathbb{R}^2 \to \mathbb{R}^2$ to be the orthogonal projection onto the x-axis.



Why do we care about this ?

Linear systems

Assume that we given a dataset:

$$a_i = (a_{i,1}, \dots, a_{i,m}) \in \mathbb{R}^m, \quad y_i \in \mathbb{R} \qquad \text{for} \quad i = 1, \dots, n.$$

We would like to find $x \in \mathbb{R}^m$ such that

$$x_1 a_{i,1} + \cdots + x_m a_{i,m} = y_i \qquad \text{for all} \quad i \in \{1, \dots, n\}.$$

We need to solve :

$$\begin{cases} x_1 a_{11} + \cdots + x_m a_{1,m} = y_1 \\ \vdots \\ x_1 a_{n,1} + \cdots + x_m a_{n,m} = y_n \end{cases}$$

Matrix notation

Let's write

$$A = \underbrace{\begin{pmatrix} a_{1,1} & \cdots & a_{1,m} \\ \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & a_{n,m} \end{pmatrix}}\in \mathbb{R}^{n\times m} \quad \text{and} \quad y = \underbrace{\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}} \in \mathbb{R}^n.$$

$\begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}$

Then the system of equations becomes $\boxed{A \underline{\underline{x}} = y}$

- Case 1 : $y \notin \mathrm{Im}(A)$: my system does not have any solution.
- Case 2 : $y \in \mathrm{Im}(A)$, then there exists $x_0 \in \mathbb{R}^m$ such that $Ax_0 = y$.

Let's find all solutions !

Is x_0 the only solution to $Ax = y$?

x is solution $\iff Ax = y$ $\quad Ax_0 = y$

$\iff Ax - Ax_0 = y - y$

$\iff A(x - x_0) = 0.$

$\iff x - x_0 \in \mathrm{Ker}(A)$

$\iff$ there exists $v \in \mathrm{Ker}(A)$

such that : $x = x_0 + v.$

The set of all solution is

$S = \{ x_0 + v \mid v \in \mathrm{Ker}(A) \}$

Conclusion: 3 possible cases

1. $y \notin \mathrm{Im}(A)$: there is no solution to $Ax = y$.

2. $y \in \mathrm{Im}(A)$, then there exists $x_0 \in \mathbb{R}^m$ such that $Ax_0 = y$. The set of solutions in then

$$S = \left\{ x_0 + v \,\middle|\, v \in \mathrm{Ker}(A) \right\}.$$

- If $\mathrm{Ker}(A) = \{0\}$, then $S = \{x_0\}$: x_0 is the unique solution.
- If $\mathrm{Ker}(A) \neq \{0\}$, then $\mathrm{Ker}(A)$ contains infinitely many vectors: there are infinitely many solutions.

Gaussian elimination

$$A = \begin{pmatrix} 1 & -1 & 0 & 1 \\ 2 & 0 & 1 & -1 \\ -1 & 5 & 2 & 0 \end{pmatrix} \in \mathbb{R}^{n\times m} \qquad \text{and} \qquad y = \begin{pmatrix} 1 \\ 1 \\ 4 \end{pmatrix} \in \mathbb{R}^n.$$

$$\begin{cases} x_1 - x_2 + x_4 = 1 & (R_1) \\ 2x_1 + x_3 - x_4 = 1 & (R_2) \\ -x_1 + 5x_2 + 2x_3 = 4 & (R_3) \end{cases}$$

"leading coefficient"

$$\begin{cases} x_1 - x_2 + x_4 = 1 & (R_1) \\ 0 + 2x_2 + x_3 - 3x_4 = -1 & (R_2) - 2(R_1) \\ 0 + 4x_2 + 2x_3 + x_4 = 5 & (R_3) + (R_1) \end{cases}$$

leading coefficient

Gaussian elimination

$$\begin{cases} x_1 - x_2 + x_4 = 1 & (R_1) \\ 2x_2 + x_3 - 3x_4 = -1 & (R_2) \\ 0 + 0 + 7x_4 = 7 & (R_3) - 2(R_2) \end{cases}$$

$$\begin{cases} x_1 - x_2 + \underline{x_4} = 1 \\ 2x_2 - \underline{3x_4} = -1 - x_3 \\ \underline{x_4} = \underline{1} \end{cases}$$

x_3 will play a role of "parameter" of the set of solutions

$$\begin{cases} x_1 = 1 - x_3/2 \\ x_2 = 1 - \frac{x_3}{2} \\ x_3 = x_3 \\ x_4 = 1 \end{cases}$$

Gaussian elimination

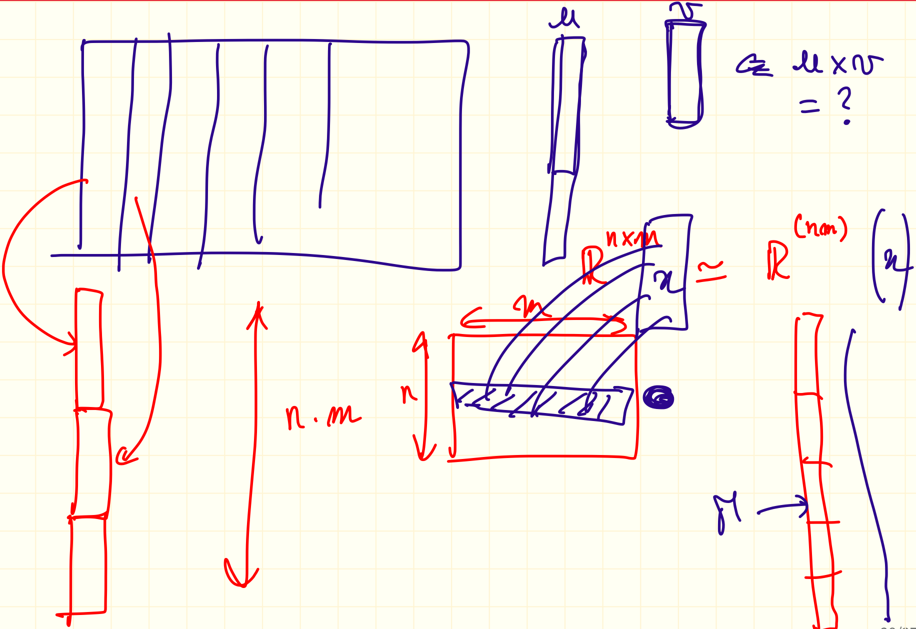
The set of solution

$$S = \left\{ \begin{pmatrix} 1 - x_3/2 \\ 1 - x_3/2 \\ x_3 \\ 1 \end{pmatrix} \;\middle|\; \underline{\underline{x_3 \in \mathbb{R}}} \right\}$$

$$= \left\{ \underbrace{\begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \end{pmatrix}}_{x_0} + v \;\middle|\; v \in \underbrace{\mathrm{Span}\left(\begin{pmatrix} -1/2 \\ -1/2 \\ 1 \\ 0 \end{pmatrix}\right)}_{\mathrm{Ker}(A)} \right\}$$

Gaussian elimination

Questions?



u
v
$\approx u \times v$
$= ?$
$n \cdot m$
$\mathbb{R}^{n \times m}$
$\simeq \mathbb{R}^{(nm)}$
m
n

Questions?

Questions?