MATH 2121 - Lecture #13

* Get graded midterms in tutorials
* Next round of offline/online HW posted
  due next week

Outline : ① a little more about det
② abstract vector spaces

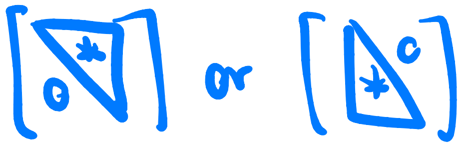
1 Last time: determinants

Let n be a positive integer.

Theorem. det is the unique function $\{ \ n \times n \text{ matrices } \} \to \mathbb{R}$ such that

1. $\det I_n = 1$ where $I_n = \begin{bmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{bmatrix}$ is the $n \times n$ identity matrix.
2. Switching two columns reverses the sign of the determinant. "alternating"
3. $\det A$ is linear as a function of a single column A if all other columns are fixed. "multilinear"

For 1×1 and 2×2 matrices, we have $\det \begin{bmatrix} a \end{bmatrix} = a$ and $\det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc$.



The *diagonal (positions)* of an $n \times n$ matrix are the positions $(1,1), (2,2), \ldots, (n,n)$.

The *diagonal entries* of a matrix are the entries in these positions.

A matrix A is *upper triangular* if $A_{ij} = 0$ whenever $i > j$.

A matrix A is *lower triangular* if $A_{ij} = 0$ whenever $i < j$.

A *triangular matrix* is a square matrix that is either upper or lower triangular.

A *diagonal matrix* is a matrix that is both upper and lower triangular.

Theorem. If A is triangular then $\det A$ is the product of the diagonal entries of A.

Theorem. A square matrix A is invertible if and only if $\det A \neq 0$.

Theorem. If A and B are $n \times n$ matrices then

$$\boxed{\det(AB) = (\det A)(\det B)} \quad \text{and} \quad \boxed{\det(A^T) = \det A.}$$

via RREF

Algorithm to compute $\det A$ for $n \times n$ matrix A.

1. Start by setting $\mathsf{denom} = 1$.
2. Row reduce A to an echelon form E, while doing the following:
   (a) When you switch two rows, multiply denom by -1.
   (b) When you rescale a row by a nonzero factor λ, multiply denom by λ.

Then $\boxed{\det A = \frac{\det E}{\mathsf{denom}}} = \frac{\text{the product of the diagonal entries of } E}{\mathsf{denom}}$.

(E is triangular matrix)

Theorem (One more method)**.** Consider an $n \times n$ matrix A.

Let $\boxed{A_{ij}}$ be entry in row i, column j. Form $\boxed{A^{(i,j)}}$ by deleting row i and column j.

Then $\boxed{\det A = A_{11} \det A^{(1,1)} - A_{12} \det A^{(1,2)} + A_{13} \det A^{(1,3)} - \cdots - (-1)^n A_{1n} \det A^{(1,n)}}$

when $n=3$:

$$\det \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} = a(ei-fh) - b(di-fg) + c(dh-eg)$$

2 Interpreting the determinant geometrically

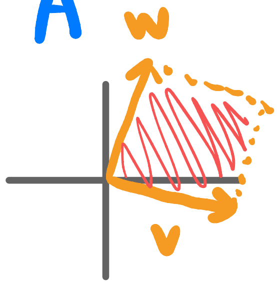
Let A be an $n \times n$ matrix. The determinant has a physical interpretation:

Proposition. The number $|\det A|$ is the volume of the *n-dimensional parallelogram*

$$P(A) = \{Av : v \in \mathbb{R}^n \text{ with } 0 \leq v_i \leq 1 \text{ for all } i = 1, 2, \ldots, n\}.$$

abs val of $\det(A)$ = volume of a region of space associatied to A

$P(A) \subseteq Col(A)$

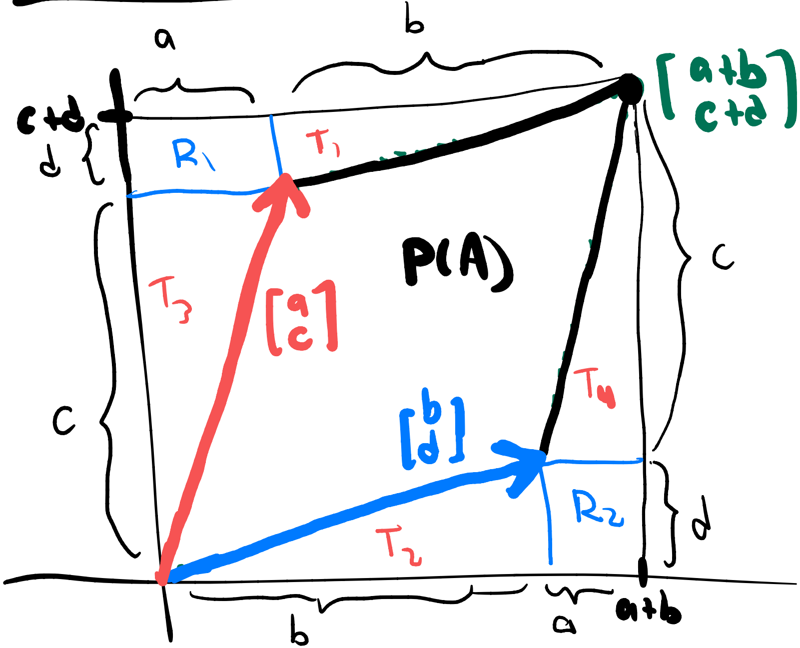
when $n=2$ and $A = [v, w]$ then $P(A) =$ 

$v, w \in \mathbb{R}^2$

Proof of this when $n=2$

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

when $n=2$, volume = area



$R_1 = ad$

$R_2 = ad$

R_0 = big rectangle

$T_1 = T_2 = \frac{1}{2}bd$

$T_3 = T_4 = \frac{1}{2}ac$

area $(a+b)(c+d)$

We can compute area of $P(A)$ by subtraction:

this is >0 → $-ad+bc$

$= |ad-bc|$

$=$

area of $P(A)$ $= R_0 - R_1 - R_2 - (T_1+T_2) - (T_3+T_4)$

$= (a+b)(c+d) - 2ad - bd - ac = -ad + bc = |\det A|$

we already saw this result when $S = \{ v \in \mathbb{R}^n \mid 0 \leq v_i \leq 1 \text{ for all } i \}$

has volume 1

Suppose $T : \mathbb{R}^n \to \mathbb{R}^n$ is linear with standard matrix A.

Corollary. If S is any subset of $\mathbb{R}^n$ with finite volume then

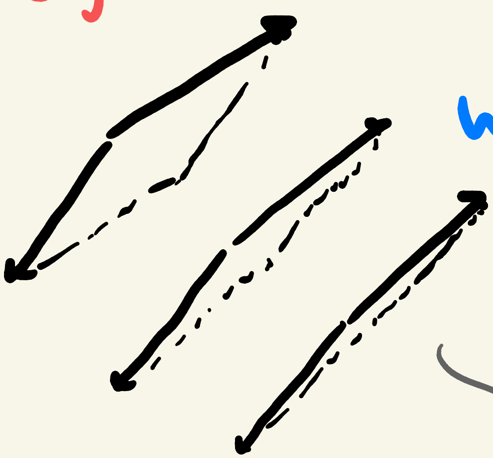
$$\text{volume of } T(S) = |\det(A)| \times (\text{volume of } S).$$

if you take any region of finite volume $S \subseteq \mathbb{R}^n$ then the new region $T(S) \overset{\text{def}}{=} \{ Av \mid v \in S \} \subseteq \mathbb{R}^n$ also has finite volume, given by $|\det A| \cdot \text{vol}(S)$

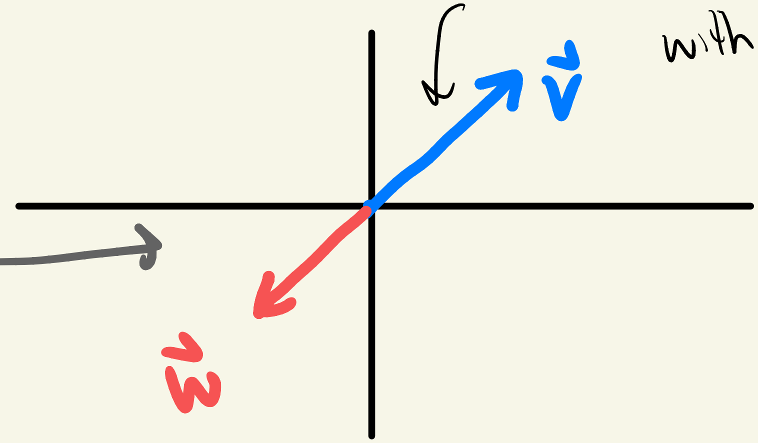
Suppose $n=2$. Why does $\det A = 0$ if A not invertible?

Reason: in this case both columns of A must be on same line (linearly dependent) so the parallelogram $P(A)$ is degenerate with zero volume

degeneration:



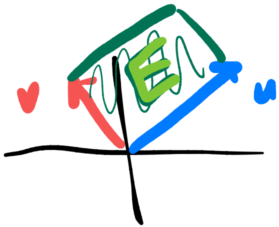
this is a parallelogram with area zero

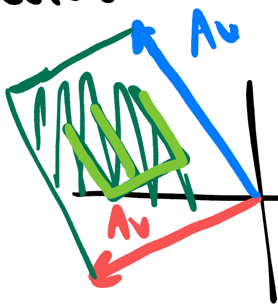


$A = [\vec{v}\ \vec{w}]$

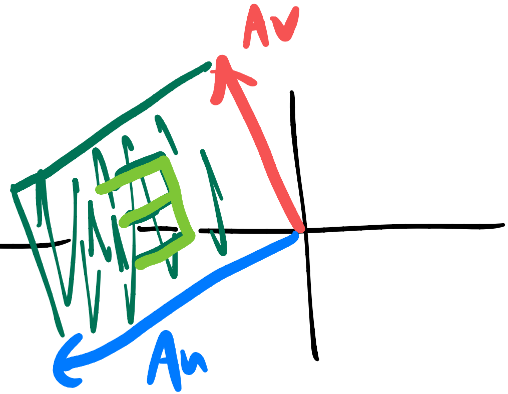
What is physical interpretation of sign of det A?

Suppose you have two vectors that look like this:





det A > 0



det A < 0

Let A be a 2x2 matrix

two options for Au, Av

Take any $v, u \in \mathbb{R}^2$

generalization / abstraction of our notion of subspaces of $\mathbb{R}^n$

3 Vector spaces

This course focuses on $\mathbb{R}^n$ and its subspaces.

These objects are examples of *(real) vector spaces*

There is also a notion of a *complex vector space* where our scalars can be complex numbers from $\mathbb{C}$ rather than just $\mathbb{R}$. Essentially all of the theory is the same, so for now we stick to real vector spaces which are more closely aligned with applications.

involves 3 things: ① a set V of what we will call vectors
② a way of adding vectors together
③ a way of doing scalar multiplication

The general definition of a vector space is given as follows: (subject to several conditions)

Definition. A *vector space* is a nonempty set V with two operations called *vector addition* and *scalar multiplication* satisfying several conditions.

We refer to the elements of V as $\boxed{\textit{vectors}.}$

required conditions

The vector addition operation for V must be a rule that takes two input vectors $u, v \in V$ and produces an output vector $u + v \in V$ such that

(a) $u + v = v + u$. commutative

(b) $(u + v) + w = u + (v + w)$. associative

(c) There exists a unique $\boxed{\textit{zero vector } 0 \in V}$ with $0 + v = v$ for all $v \in V$.

there is a zero vector

write scalar multiplication as cv or $c \cdot v$

$(c \in \mathbb{R},\ v \in V)$

more conditions

The scalar multiplication operation for V must be a rule that takes a scalar input $\boxed{c \in \mathbb{R}}$ and an input vector $\boxed{v \in V}$ and produces an output vector $\boxed{cv \in V}$ such that

(a) If $c = -1$ then $v + (-1)v = 0$ ← this 0 is the zero vector from prev slide

(b) $c(u + v) = cu + cv$.

(c) $(c + d)v = cv + dv$ for $c, d \in \mathbb{R}$.

(d) $c(dv) = (cd)v$ for $c, d \in \mathbb{R}$.

(e) If $c = 1$ then $1v = v$.

distributivity

Notation: If $v \in V$ then we define $-v = (-1)v$ and $u - v = u + (-v)$.

$$u - v \overset{\text{def}}{=} u + (-1)\cdot v$$

$$-v \overset{\text{def}}{=} (-1)\cdot v$$

* most important example of vector spaces:

① **Example.** $\mathbb{R}^n$ and any subspace of $\mathbb{R}^n$ are vector spaces, with the usual operations of vector addition and scalar multiplication.

* another example: (not a subspace of $\mathbb{R}^n$)

② **Example.** The set $\mathbb{R}^{m\times n}$ of $m \times n$ matrices is a vector space. Its zero vector is the zero $m \times n$ matrix.

what is zero vector? $0 = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}$ ($m\times n$ zero matrix)

third example:

③ **Example.** The set of linear functions $f : \mathbb{R}^n \to \mathbb{R}^m$ is a vector space. Its zero vector is the function f with the formula $f(x) = 0$ for all $x \in \mathbb{R}^n$.

what is zero vector? $0 =$ (the function $\mathbb{R}^n \to \mathbb{R}^m$ with formula $x \mapsto \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$)

Claim: ② and ③ are different vector spaces, but are somehow "the same"

idea: taking standard matrix is an invertible map between ② and ③ compatible w/ + and ·

Less familiar example: ① set of vectors = $\left(\begin{array}{c} x \in \mathbb{R} \\ x > 0 \end{array}\right)$

Example. The set of positive real numbers $\mathbb{R}^+ = \{x \in \mathbb{R} : x > 0\}$ is a vector space, but *not for the usual addition and multiplication operations.*

Instead, define a new version of addition $\oplus$ by $x \oplus y = xy$ for $x, y \in \mathbb{R}^+$.

Then define a new version of multiplication $\odot$ by $c \odot x = x^c$ for $c \in \mathbb{R}$ and $x \in \mathbb{R}^+$.

* The zero vector for this vector space $(\mathbb{R}^+, \oplus, \odot)$ is the number 1.

We also have $\ominus x = 1/x$ and $x \ominus y = x/y$.

② new addition: $x \oplus y$ " = " x times " y " $= xy$

$2 \oplus 6 = 12$

$\ominus x = -1 \odot x = x^{-1}$

③ new scalar mult: $c \odot x = x^c$

$2 \odot 6 = 6^2 = 36$

$1 \odot x = x^1 = x$

It is rarely necessary to check the axioms of a vector space in detail, and there is not much need to memorize the abstract definition. If we have a set with operations that look like vector addition and scalar multiplication for $\mathbb{R}^n$, then we usually have a vector space. Moreover, it's typically easy to identify every vector space we encounter as a special case of a few general constructions like the following:

function space

Example. Let X be any set. Let $\mathsf{Functions}(X, \mathbb{R})$ to be the set of functions

$$f : X \to \mathbb{R}.$$

Given $f, g \in \mathsf{Functions}(X, \mathbb{R})$ define $f + g$ to be the function with the formula

$$(f + g)(x) = f(x) + g(x) \qquad \text{for } x \in X.$$

Given $c \in \mathbb{R}$ and $f \in \mathsf{Functions}(X, \mathbb{R})$, define cf to be the function with the formula

$$(cf)(x) = cf(x) \qquad \text{for } x \in X.$$

point-wise addition and scalar mult. operations

The set $\mathsf{Functions}(X, \mathbb{R})$ is a vector space relative to these operations.

Zero vector in $\mathsf{Functions}(X, \mathbb{R})$ is function with formula $f(x) = 0$ for all $x \in X$.

$X = \{1,2,3,\ldots,n\}$

* The vector space $\mathbb{R}^n$ "is the same as" $\mathsf{Functions}(\{1,2,3,\ldots,n\},\mathbb{R})$.

More generally, if V is any vector space then the set of functions

$$\mathsf{Functions}(X,V) = \{f : X \to V\}$$

is a vector space for similar definitions of vector addition and scalar multiplication.

an element of $\mathbb{R}^n$ is a list of numbers $v = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$

a function $f : \{1,2,\ldots,n\} \to \mathbb{R}$ is also encoded as a list of numbers: $\begin{bmatrix} f(1) \\ f(2) \\ \vdots \\ f(n) \end{bmatrix}$

if we write functions $f : \{1,2,\ldots,n\} \to \mathbb{R}$ as column vectors in this way, then $f+g$ and $c \cdot f$ correspond to our usual vector ops

① any scalar times zero vector = zero vector

② scalar $0 \in \mathbb{R}$ times any vector = zero vector

As an example of how one can use the axioms to prove properties of a general vector space, consider the following identities which are obvious for subspaces of $\mathbb{R}^n$.

Proposition. If V is a vector space then $0v = 0$ and $c0 = 0$ for all $c \in \mathbb{R}$ and $v \in V$.

Proof. We have $0v = (0+0)v = 0v + 0v$ so

$$0 = 0v - 0v = (0v + 0v) - 0v = 0v + (0v - 0v) = 0v + 0 = 0v.$$

Similarly, $c0 = c(0+0) = c0 + c0$ so

$$0 = c0 - c0 = (c0 + c0) - c0 = c0 + (c0 - c0) = c0 + 0 = c0.$$

□

We will not focus very much in this course on the art of coming up with these sorts of algebraic derivations. Mostly, we can just rely on our intuition from subspaces of $\mathbb{R}^n$ when working with more general spaces.

4 Subspaces, bases, and dimension

Notions of subspaces, bases, and dimension for vector spaces are the same as for $\mathbb{R}^n$.

Same defn as for subspaces of ~~$\mathbb{R}^n$~~ V

Definition. A $\boxed{\textit{subspace}}$ of a vector space V is a subset H containing the zero vector of V, such that if $u, v \in H$ and $c \in \mathbb{R}$ then $u + v \in H$ and $cv \in H$.

If $H \subset V$ is a subspace then H is itself a vector space with the same operations of scalar multiplication and vector addition.

Example. V is a subspace of itself and $\{0\} \subset V$ is a subspace.

Ex $\{$upper-Δ $n \times n$ matrices$\}$ is subspace of $\mathbb{R}^{n \times n}$

(all $n \times n$ matrices)

(upper or lower)

Ex. $\{$all triangular $n \times n$ matrices$\}$ is NOT a subspace of $\mathbb{R}^{n \times n}$

why: $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ = not triangular

Example. $\mathbb{R}^2$ is technically not a subspace of $\mathbb{R}^3$ since $\mathbb{R}^2$ is not a subset of $\mathbb{R}^3$.

If you want a subspace of $\mathbb{R}^3$ that "looks like" $\mathbb{R}^2$, three candidates are

$$\left\{\begin{bmatrix} x \\ y \\ 0 \end{bmatrix} : x, y \in \mathbb{R}\right\}, \quad \left\{\begin{bmatrix} x \\ 0 \\ y \end{bmatrix} : x, y \in \mathbb{R}\right\}, \quad \text{and} \quad \left\{\begin{bmatrix} 0 \\ x \\ y \end{bmatrix} : x, y \in \mathbb{R}\right\}.$$

There is nothing intrinsic that makes one of these more natural than the rest.

function subspace

Example. Let X be any set. Let $Y \subset X$ be a subset.

Define H as the subset of $\mathsf{Functions}(X, \mathbb{R})$ consists of the functions $f : X \to \mathbb{R}$ with $f(y) = 0$ for all $y \in Y$. Then H is a subspace. (needs checking)

Example. The set of all functions $\mathsf{Functions}(\mathbb{R}^n, \mathbb{R}^m)$ is a vector space since $\mathbb{R}^m$ is a vector space. The subset of linear functions $f : \mathbb{R}^n \to \mathbb{R}^m$ is a subspace.

Let V be a vector space.

same def as in $\mathbb{R}^n$

Definition. A *linear combination* of $v_1, v_2, \ldots, v_k \in V$ is a vector of the form

$$c_1 v_1 + c_2 v_2 + \cdots + c_k v_k \in V$$

for some scalars $c_1, c_2, \ldots, c_k \in \mathbb{R}$.

A linear combination of an infinite set is a linear combination of some finite subset.

A linear combination by definition **only involves finitely many vectors**.

Definition. The *span* of a set of vectors is the set of all linear combinations that can be formed from the vectors. It is important to note that each such linear combination can **only involve finitely many vectors**.

The span of a set of vectors in V is a subspace of V.

new

Note: $\mathbb{R}^n = \mathbb{R}\text{-span}\{e_1, e_2, \ldots, e_n\} = \mathbb{R}\text{-span}\{v \in \mathbb{R}^n\}$

old

real valued functions

Example. Let $V = \mathsf{Functions}(\mathbb{R}, \mathbb{R})$. The span of the infinite set of functions

$$\text{monomials: } \quad 1, x, x^2, x^3, \cdots \in V$$

is the subspace of $\boxed{\textit{polynomial functions}}$. = finite lin. comb. of monomials

Each polynomial function is a linear combination of a finite number of monomials

$$c_n x^n + c_{n-1} x^{n-1} + \cdots + c_1 x + c_0.$$

An infinite sum like $\underline{1 + x + x^2 + \ldots \text{ is } \textbf{not} \text{ a polynomial}}$.

$1 + x + x^2 + \cdots = \frac{1}{1-x}$ not a polynomial

not exactly rigorous

Definition. A finite set of vectors $v_1, v_2, \ldots, v_k \in V$ is *linearly independent* if it is impossible to express

$$0 = c_1 v_1 + c_2 v_2 + \cdots + c_k v_k$$

for some $c_1, c_2, \ldots, c_k \in \mathbb{R}$ except when $c_1 = c_2 = \cdots = c_k = 0$.

An infinite set is linearly independent if every finite subset is linearly independent.

Definition. A *basis* of a vector space V is a subset of **linearly independent** vectors whose **span is** V.

Saying that $b_1, b_2, b_3, \ldots$ is a basis for V is the same thing as saying that each $v \in V$ can be expressed as a uniquely linear combination of basis elements.

Theorem. Let V be a vector space.

1. V has at least one basis.
2. Every basis of V has the same size.
3. If A is a subset of linearly independent vectors in V then V has a basis B with
$$A \subset B.$$
4. If C is a subset of vectors in V whose span is V then V has a basis B with
$$B \subset C.$$

Definition. As for subspaces of $\mathbb{R}^n$, we define the $\boxed{\textit{dimension}}$ of a vector space V to be the common number of elements in any of its bases.

Denote the dimension of V by $\dim V$.

Corollary. If $H \subset V$ is a subspace then $\dim H \leq \dim V$.

Moreover, if $H \subset V$ is a subspace with $\dim H = \dim V$ then $H = V$.

Example. If X is a finite set then $\dim \mathsf{Functions}(X, \mathbb{R}) = |X|$ is the size of X.

A basis is given by the functions $\delta_y : X \to \mathbb{R}$ for $y \in X$, defined by the formulas

$$\delta_y(x) = \begin{cases} 1 & \text{if } x = y \\ 0 & \text{if } x \neq y \end{cases} \qquad \text{for } x \in X.$$

Suppose U and V are vector spaces.

Definition. A function $f : U \to V$ is *linear* if

$$f(u+v) = f(u) + f(v) \qquad \text{and} \qquad f(cv) = cf(v) \qquad \text{for all } c \in \mathbb{R} \text{ and } u, v \in U.$$

Define $\mathsf{range}(f) = \{f(x) : x \in U\}$ and $\mathsf{kernel}(f) = \{x \in U : f(x) = 0\}$.

Proposition. If $f : U \to V$ is linear then $\mathsf{range}(f)$ and $\mathsf{kernel}(f)$ are subspaces.

These subspaces are generalizations of the column space and null space of a matrix.

Proposition. If U, V, W are vector spaces and $f : V \to W$ and $g : U \to V$ are linear functions then $f \circ g : U \to V \to W$ is linear, where $f \circ g(x) = f(g(x))$.