

くりくりすーがくれっすん
～AIのための数学第一歩～

文: いりす, 絵: 海牛

2019 年 8 月 16 日

目次

| | | |
|-----|-------------------------------|----|
| 1 | はじまり | 3 |
| 2 | くりくり♡せんけーだいすーにゆうもん♪ | 5 |
| 2.1 | ベクトル | 5 |
| | 平面ベクトルの定義と演算 | 5 |
| | n 次元ベクトル | 11 |
| | ティータイム：一般線形空間 | 13 |
| 3 | くりくり♡びぶんにゆうもん♪ | 17 |
| 3.1 | 微分のイントロダクション | 17 |

1 はじまり

梅雨も明け空気に茹るような暑気と夏の香を帯びた時候。午後を迎えた聖 MIDI 女学園中等部、一般の生徒は昼食を終えたあとの微睡みと奮闘しながら授業を受けている。そして、生徒会室では、ロージアがスマートフォンを片手に神妙な面持ちで深沈している。



「うむむ…」



「ロージアどうしたんですか？ 珍しく真面目な顔をして…」



「珍しくって何よ、あざと学ダンス主席で成績もと～っても優秀なロージアちゃんはいつも真面目です～」



「ろーじあはテストでいい点とってるし、とっても真面目なの～」



「やっぱりツキノはちゃんと私のこと見てくれてるね♪ いつもありがと♡ お兄ちゃんはロージアちゃんのこと見てくれる…？ (謎カメラ目線)」



「(誰にいつてるんですか…)」



「ロージアが真面目かはともかく、そんな顔でスマホみてるなんて珍しいやん？ どうしたん？」



「どうしたもこうしたも無いわよ、今度の台風見た？」



「台風 69 号でしたっけ…」



「そうそう、それね。その台風が今度のライブに来るかもしれないのよ」



「ほんまか！？ セっかくうちらも練習しとるし、見に来るファンのおにいやん達も現場のスタッフさんたちの為にも何事もなかったように開催してほしいなあ…」



「台風ってすぐにあぶないってわかったりしないの？」



「そうよね～最近はえーあい…？ が流行ってるし、なんとかなりそうだけどね。ねえねえホルミー、そういうの作れたりしないの？」



「『そういうの』って…天気予報ですら当てるのが難しいのによく言いますね…でも、ライブだけでなく電車を早めに止めたり的意思決定にも使えそうですし、作ってみる価値はあるかもしれませんね。私たちは寮だから登校はあまり関係なさそうですけど…」



「なるほどね、AI が止めれば簡単に休校になるってことでしょ♡ 面白そうじゃん、ホルミー、AI の作り方を教えて〜♡」



「えっ、私ですか…！？ AI を作るための数学なら教えられますけど…」



「えー数学ー…ま、全然大丈夫だけどね♡」



「つきの、すーがく苦手なの…」



「つきのはあまり数学とかは得意そうじゃないよね、ロージアちゃんが教えてあげるわ♡」



「ロージアが数学得意ってなんか珍しいやん」



「まあね〜これでもホルミー直々に鍛えてもらったからね〜*1」



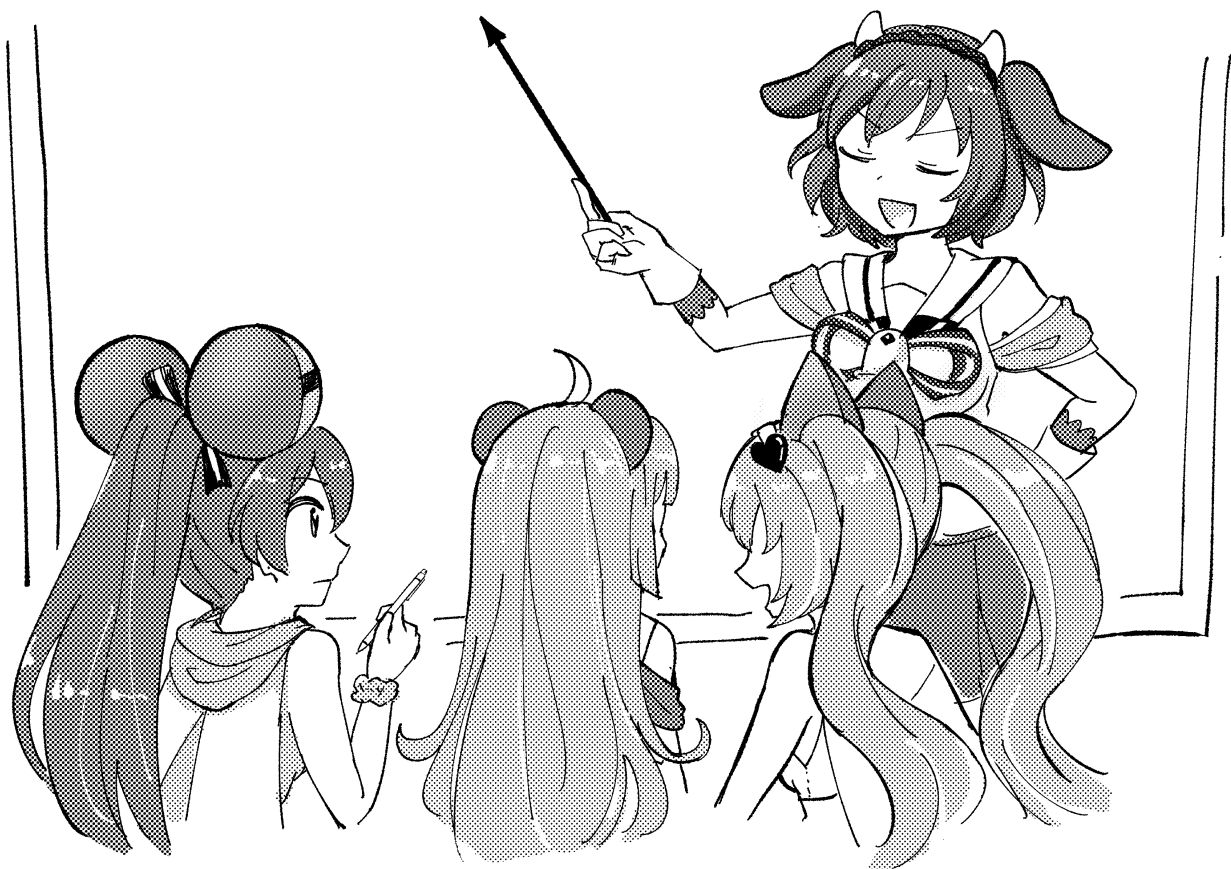
「ええな〜うちもホルミーに数学教えてもらいたいわあ〜」



「ええっと…はじめたいのですが…」

*1 以前の作品「ろーほるすーがくれっすん」のこと

<https://irisu-in-wonderland.tumblr.com/post/156533872627>



2 くりくり♡せんけーだいすーにゅうもん♪

2.1 ベクトル

平面ベクトルの定義と演算



「まずは、機械学習を議論する上で最も重要となるベクトルと行列を学びます」

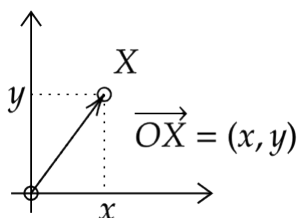


「ベクトル? 『チームのベクトル』みたいに使うし、向きって意味やっけ?」



「そうですね、ベクトルとは向きと大きさを表すもので、例えば二次元平面なら (x_1, x_2) ($x_1, x_2 \in \mathbb{R}$) という実数の組で記載します。

今回は簡単のため実数の集まり (\mathbb{R}) のみを考え、 x_1 が実数である、つまり実数集合の一つという意味で $x_1 \in \mathbb{R}$ と書きます。」



「ベクトルは例えば、原点 $(0,0)$ から (x_1, x_2) の点への矢印を表すものです。今は点とベクトルを区別するため、点は大文字アルファベット A, B, \dots で表し、ベクトルは $\mathbf{a}, \mathbf{b}, \dots$ と小文字アルファベット太文字で表します。特に原点は O , 点 A から点 B へのベクトルを \overrightarrow{AB} と書きます。最初はわかりやすく 2 次元平面で話していきますね。」



「ふーん、矢印ねえ…でも、矢印って平面上にいっぱいいる？ 同じものもあると思うんだけど、それって区別できるの??」



「いいところに気が付きましたね！ 違う地点からの矢印でも向きと長さ $((x_1, x_2))$ の組) が同じものは同じと見なして一つのものでします。さてここで、点 $O = (0,0)$ から点 $A = (1,1)$ へのベクトル \overrightarrow{OA} と点 $B = (1,0)$ から点 $C = (2,1)$ へのベクトル \overrightarrow{BC} を考えてみましょう。」



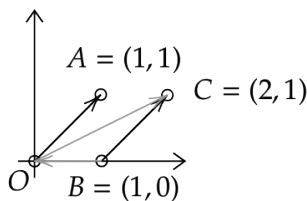
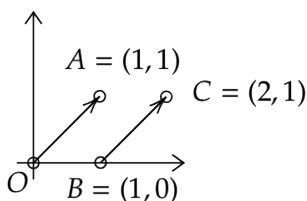
「どっちも $(1,1)$ じゃないの？」



「その通りです。ここで、 \overrightarrow{BC} を始端の点と終端の点と同じのベクトル同士をつなげることを $+$ で表して、考えると

$$(1,1) = \overrightarrow{BC} = \overrightarrow{BO} + \overrightarrow{OC} = (-1,0) + (2,1)$$

となります。 (\overrightarrow{BO}) は $\overrightarrow{OB} = (1,0)$ の逆向きなので $(-1,0)$ 」



「えーっと、これって右の足し算が $(-1,0) + (2,1)$ が $(-1+2, 0+1) = (1,1)$ になるの…？」



「はい！ このようにベクトル同士の足し算は次のように定義します

$$(x_1, x_2) + (y_1, y_2) := (x_1 + y_1, x_2 + y_2).$$

また、ベクトルは同じ矢印の向きと長さが同じものを一つに見なしてたので、 (y_1, y_2) の矢印を始端を (x_1, x_2) の終端に合わせることができ、すべてのベクトルに対して足し算を考えることができます。」



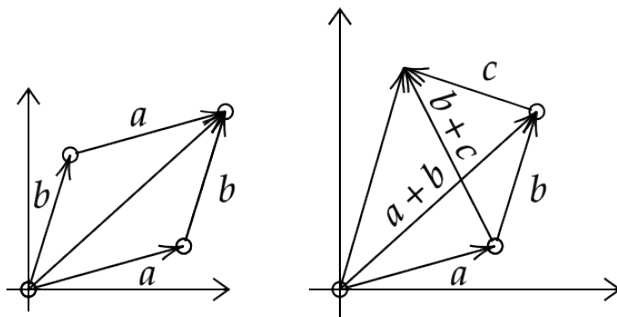
「ええっと、つまり、前の矢印に後ろの矢印を持ってきてあげるの…あれ、どっちも同じになるの」



「ツキノの言う通り、ベクトルの足し算には以下の性質が成り立ちます。」

命題 2.1. $\mathbf{o} = (0, 0)$ として、任意のベクトル $\mathbf{a}, \mathbf{b}, \mathbf{c}$ について

- (1) $\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}$ (交換律)
- (2) $(\mathbf{a} + \mathbf{b}) + \mathbf{c} = \mathbf{a} + (\mathbf{b} + \mathbf{c})$ (交換律)
- (3) $\mathbf{a} + \mathbf{o} = \mathbf{a}$



「さて、次にベクトルの引き算を考えてみます。 $\overrightarrow{OA} = (x_1, x_2)$ に対して、

$$\overrightarrow{AO} + (-\overrightarrow{AO}) = (0, 0) \Rightarrow -\overrightarrow{AO} = (-x_1, -x_2)$$

また、 $\overrightarrow{AO} + \overrightarrow{OA} = (0, 0)$ なので、 $-\overrightarrow{AO} = \overrightarrow{OA}$ となります。このマイナスの定義をよく見ると、元のベクトルの中身に -1 をかけてます。」



「いわれてみればそうね」



「これはベクトルの長さ -1 倍していると考えることができます。これを一般化して実数 c についてベクトルを c 倍することを考えます。ベクトル (x_1, x_2) の c 倍は

$$c(x_1, x_2) := (cx_1, cx_2)$$

となり、これをスカラー倍といいます。スカラー倍には次の性質が成立します。」

命題 2.2. 任意のベクトル \mathbf{a}, \mathbf{b} , 実数 c, d について、

- (1) $c(\mathbf{a} + \mathbf{b}) = c\mathbf{a} + c\mathbf{b}$
- (2) $(c + d)\mathbf{a} = c\mathbf{a} + d\mathbf{a}$
- (3) $(cd)\mathbf{a} = c(d\mathbf{a})$
- (4) $1\mathbf{a} = \mathbf{a}$



「さて、続いて、次にベクトルの長さを定義します。ベクトル $\mathbf{x} = (x_1, x_2)$ の長さ、つまり、 $(0, 0)$ から (x_1, x_2) の距離になります。」



「あっ、これどこかで聞いたことある。^{*1}たしか三平方の定理から

$$\sqrt{(0 - x_1)^2 + (0 - x_2)^2} = \sqrt{x_1^2 + x_2^2}$$

じゃない？」



「その通りです！ ベクトルの $\mathbf{x} = (x_1, x_2)$ の長さをベクトルのノルムと言い、 $\|\mathbf{x}\|$ と表し

$$\|\mathbf{x}\| = \sqrt{x_1^2 + x_2^2}$$

となります。もちろんベクトルの引き算の定義から

$$\|\mathbf{x} - \mathbf{y}\| = \|(x_1 - y_1, x_2 - y_2)\| = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}$$

となり、 $\|\mathbf{x} - \mathbf{y}\|$ は (x_1, x_2) と (y_1, y_2) の距離となります。」



「まあ、確かにベクトルの引き算から距離っぽくなるよね。」

^{*1} ろーほるすーがくれっすん §3

<https://irisu-in-wonderland.tumblr.com/post/156533872627>



「続いて、ベクトル同士の角度を表すことができる**内積**を定義します。ベクトル \mathbf{x}, \mathbf{y} の内積を $\mathbf{x} \cdot \mathbf{y}$ と表し、

$$\mathbf{x} \cdot \mathbf{y} := x_1 \times y_1 + x_2 \times y_2$$

もちろん、定義から \mathbf{x} のノルムは内積を使って

$$\|\mathbf{x}\| = \sqrt{\mathbf{x} \cdot \mathbf{x}}$$

となります。」



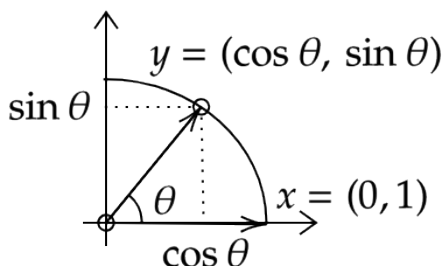
「うーん…これのどこが角度と関係するの？」



「簡単な例で考えてみましょう。

長さが1のベクトル \mathbf{x}, \mathbf{y} を考えます。 \mathbf{x}, \mathbf{y} は長さが1なので原点を中心とした半径1の円のどこかに位置します。

ここで、 $\mathbf{x} = (1, 0)$ として、円上のベクトル \mathbf{y} を考えてると、 \mathbf{x} と \mathbf{y} の角度 θ として、 $\mathbf{y} = (\cos \theta, \sin \theta)$ と書けます。」



「このとき

$$\mathbf{x} \cdot \mathbf{y} = \cos \theta$$

となります。このように内積は角度 θ についての関数 \cos と関係してます。

ただし、この θ は角度ですが、1回転が 360° で表記せず、1回転が $2\pi, 180^\circ$ が π と半径1の円の弧に対応する**弧度法**を使用します。」



「こどほう…？ 波動砲みたいやな??」



「ふーん、円の長さを角度とみるんだ、なんだか馴染みないわね。」



「実は、こうした方が都合が良いことがわかります*1。以下に度数法と弧度法の対応する表を載せておきます。

| | | | |
|-----|--------|-------|-----------------|
| 度数法 | 360° | 180° | 90° |
| 弧度法 | 2π | π | $\frac{\pi}{2}$ |

さて、一般的なベクトルに対して

$$\cos \theta = \frac{\mathbf{x} \cdot \mathbf{y}}{\|\mathbf{x}\| \|\mathbf{y}\|}$$

が成立します。」



「そう言われてみれば角度っぽいなあ…」



「ただこの説明で、何でもいから円上の点へのベクトル \mathbf{x}, \mathbf{y} を $\mathbf{x} = (1, 0)$ として、円上のベクトル \mathbf{y} で考えた部分は、内積がベクトルを t 回転させる関数 f_t について変わらないことを言わないといけません。つまり、

$$\mathbf{x} \cdot \mathbf{y} = f_t(\mathbf{x}) \cdot f_t(\mathbf{y})$$

が成立します。

次にどんなベクトル $\mathbf{x} = (x_1, x_2)$ もベクトル $\mathbf{e}_1 = (1, 0), \mathbf{e}_2 = (0, 1)$ として

$$\mathbf{x} = x_1 \mathbf{e}_1 + x_2 \mathbf{e}_2$$

となります。では、 $\mathbf{e}_1, \mathbf{e}_2$ だけでなく他の2つのベクトルで平面すべてを表すことを考えます。」



「他のベクトルで平面を表すってそんな簡単にできるの？」



「実は意外に簡単で、2つのベクトル \mathbf{a}, \mathbf{b} が平行でなければ良いです。すると、 \mathbf{a}, \mathbf{b} のベクトルで平行四辺形を作ることが出来、その平行四辺形を左右上下に張り付けていけば、全ての平面を覆うことが出来ると考えられます。ここで \mathbf{a}, \mathbf{b} が平行でないを式で表すと $\mathbf{a} = k\mathbf{b}$ となる実数 k が存在しないことであり、これは方程式

$$c_1 \mathbf{a} + c_2 \mathbf{b} = \mathbf{0}$$

の解 (c_1, c_2) が $(0, 0)$ のみであることが言えます。このような平行でないという性質を線形独立といいます。」

*1 微分を用いると $(\cos \theta)' = -\sin \theta$ となるが、今度数法で三角関数を定めると $\cos_a \theta = \cos \frac{\pi}{180} \theta$ とすると §2.2 で示す合成関数の微分より $(\cos_a \theta)' = \frac{\pi}{180} \sin \frac{\pi}{180} \theta$ となり、微分する度余計な定数が出てきて都合が悪い。



「平行じゃないだけでそんな性質あるんやなあ…」

n 次元ベクトル



「さて、これまで 2 次元のベクトルについて学んできましたが、これを n 次元に拡張してみます。機械学習に使う数学では n 次元を扱うので、 n 次元ベクトルを定義します。」



「 n 次元…全然想像できないの…」



「まあ、私たちは 3 次元しか見えてないのでイメージしにくいと思います…しかし、定義は非常に明快です。」



「もしかして、2 次元のときは 2 つの数の組だったんやから、 n 次元の場合は n 個の数があればえんとちゃう？」



「ジャクリン、正解です！ n 次元のベクトル \mathbf{x} は以下のように定義します。

$$\mathbf{x} = (x_1, x_2, \dots, x_n) \quad (x_1, x_2, \dots, x_n \in \mathbb{R})$$

そして、 n 次元ベクトル $\mathbf{x} = (x_1, x_2, \dots, x_n), \mathbf{y} = (y_1, y_2, \dots, y_n), c \in \mathbb{R}$ について、足し算とスカラー倍を

$$\mathbf{x} + \mathbf{y} = (x_1 + y_1, x_2 + y_2, \dots, x_n + y_n)$$

$$c\mathbf{x} = (cx_1, cx_2, \dots, cx_n)$$

と定めます。 \mathbb{R} の n 次元ベクトルの集合を \mathbb{R}^n と書きます。

n 次元の場合についても 2 次元の場合と同じように次の命題が成立します。」

命題 2.3. $\mathbf{o} = (0, 0)$ として、任意のベクトル $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{R}^n$ について

- (1) $\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}$ (交換律)
- (2) $(\mathbf{a} + \mathbf{b}) + \mathbf{c} = \mathbf{a} + (\mathbf{b} + \mathbf{c})$ (結合律)
- (3) $\mathbf{a} + \mathbf{o} = \mathbf{a}$

証明 $\mathbf{a} = (\dots, a_i, \dots), \mathbf{b} = (\dots, b_i, \dots), \mathbf{c} = (\dots, c_i, \dots)$ とかく。すると
交換律:

$$\mathbf{a} + \mathbf{b} = (\dots, a_i + b_i, \dots) = (\dots, b_i + a_i, \dots) = \mathbf{b} + \mathbf{a},$$

結合率:

$$(\mathbf{a} + \mathbf{b}) + \mathbf{c} = (\dots, a_i + b_i, \dots) + \mathbf{c} = (\dots, (a_i + b_i) + c_i, \dots) = (\dots, a_i + (b_i + c_i), \dots) = \mathbf{a} + (\mathbf{b} + \mathbf{c})$$

最後は明らか。

□



「また、 n 次元ベクトルのノルムと内積を次のように定めます。」

$$\|x\| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2} = \sqrt{\sum_{i=1}^n x_i^2}$$

$$x \cdot y = x_1y_1 + x_2y_2 + \cdots + x_ny_n = \sum_{i=1}^n x_iy_i$$



「変なギザギザが出てきたの…」



「これはシグマやな～ x_1 から x_n を足すときにこう書くんやで～」



『シ熊』…ツキノのお仲間さんなの？ なんだか仲良くなれそうなの～！」



「違うと思いますが…」



「なんだか、 n 次元ベクトルってどこかで見たことあるのよね～」



「この前、 $\sqrt{2}$ について教えたときに n 次元座標は教えたじゃないですか…」



「この前って言ってもずいぶん前のように思えるのは気のせいかしら？」



「それは言わないお約束です…

さて、2次元の時と同様に線形独立が定義されます。ただし、 n 次元の場合は $2 \leq k \leq n$ の k 個のベクトル x_1, \dots, x_k について

$$\sum_{i=1}^k c_i x_i = 0$$

となる (c_1, \dots, c_k) が 0 のみであることを線形独立と言います。」



「ところでさ、ベクトルと点って何が違うの？ 平面だと (x, y) という 2 組の座標で表して、立体だと (x, y, z) って 3 組の座標で表すけど、これってベクトルも同じじゃない？」



「鋭いところに気が付きましたね！ 点とベクトルの違いは、ベクトルには足し算とスカラー倍といった演算があり、点については演算を言及してないところが大きな違いだと思います。^{*2} ベクトルで n 次元の点を表すことが出来るので、今後は点という言葉を使わずにベクトルのみで考えていきます。」



「ふーん、まあどっちでもいいってことよね。」

ティータイム：一般線形空間



「ふうー疲れた。お茶にしようよ。」



「良いですね、お茶のついでに数学の話でも…」



「えっ…お茶の時間って休むものなんやないの…？」



「ま、お茶のついでに数学の話をする、そういうものよ。」



「実は、実数の組で表した矢印でなく、曲線や関数などの一般的な対象にも、足し算とスカラー倍が定められ命題 2.1, 2.2 と逆ベクトルが唯一つ存在することが言えればベクトルとみなすことができます。」



「んんん？？？ どういうこと？」



「ええっと、ベクトルが矢印だけだと似たような性質をもったものに应用できません。実数に留まらず、他にも使えるようにベクトルで大事な性質をまとめて、何らかの集まり V に足し算 $+: V \times V \rightarrow V$ (これは V と V の組に対して V への関数という意味) とスカラー体 K とスカラー倍 $\cdot: K \times V \rightarrow V$ があって、あるルールを満たすものをベクトルとしようとしています。」

^{*2} 点は何らかの幾何構造を有した空間での要素とベクトルはベクトル空間の要素なので、この説明は語弊があるかもだけど見逃してください…

また、点がベクトル表示可能という主張はユークリッドやヒルベルトの公理的幾何学がデカルト座標で表示されることを示さないといけないが、これは『R. ハーツホーン 幾何学 I (丸善出版)』を参照 (代数幾何学じゃないよ！)



「ここで、スカラー体 K (一般に体) とは実数のように足し算 $+$ と掛け算 \times が定義されていて、任意の $a \in K$ について $a + o = o$ となる o に値するゼロ元、 $a \times e = a$ となる e に値する単位元などが存在し、 $a + (-a) = o$ となる $(-a)$ がたった一つ存在したり、 $a \neq o$ のときに $a \times a^{-1} = e$ となる逆数にあたる a^{-1} が存在するといった実数の足し算掛け算の性質をもったものです。

現代の数学では、ベクトルの集まりはなんでも良く (それこそ猫や椅子でも OK) 演算とルールが合えばベクトルで良いとしております。このような集まり $(V, K, +, \cdot)$ をベクトル空間 (線形空間) といいます。」



「そういうものなんやねえ～ということはロージアも足し算とスカラー倍が命題 2.1, 2.2 を満たせばベクトル空間になるってことなん？」



「??? ま、まあそういうことですね…」

定義 2.4. 集合 V が K 上のベクトル空間 (または線形空間) であるとは、ゼロベクトル o 、任意のベクトル $x, y, z \in V$ 、任意のスカラー $a, b \in K$ について足し算とスカラー倍が定義されていて、以下が成立することを言う。

- (1) $x + y = y + x$ (交換律)
- (2) $(x + y) + z = x + (y + z)$ (結合律)
- (3) $x + o = x$
- (4) $x + (-x) = o$ となる $-x$ が唯一存在する
- (5) $a \cdot (x + y) = a \cdot x + a \cdot y$ (分配律)
- (6) $(a + b) \cdot x = a \cdot x + b \cdot x$
- (7) $(ab) \cdot x = a \cdot (b \cdot x)$
- (8) $1 \cdot x = x$



「ちょっと定義に隠れて胡麻化そうとしてるけど、私、ベクトルじゃないんだけど」



「ロージアはベクトルになってもかわいいよ？」



「そういう問題じゃないと思いますが…」

えーっと…話を戻しますね…それまで私たちは n 次元と言ってましたが、 n 次元はどのように定められるかを考えます。実は線形独立が重要なカギとなっていて、線形独立なベクトルの集まり $A = \{x_1, x_2, \dots\} \subset V^{*1}$ について

$$\left\{ \sum c_i x_i; c_i \in K \right\} = V$$

となっているとき A を基底といいます。」



「つまり、今まで考えてた \mathbb{R}^n での $(1, 0, \dots, 0), (0, 1, \dots, 0), \dots, (0, 0, \dots, 1)$ は基底ってことよね。」



「その通り！ これは特に標準基底といい e_1, \dots, e_n と書きます。

そして、一般の線形空間において基底 A の要素の個数を次元といいます。基底の取り方は色々ありますが、実はその個数は唯一つであることが分かります*2。

とくに V の中に線形独立のベクトルが無限に存在する場合、次元は無限になります。」



「無限次元…想像つかないの…」



「そうですね、一般の線形空間の例を見てみましょう。有理数 \mathbb{Q} と実数 \mathbb{R} は実数 \mathbb{R} に含まれています ($\mathbb{Q} \subset \mathbb{R}, \mathbb{R} \subset \mathbb{R}$) ここで \mathbb{R} はその足し算 + 掛け算 \cdot について、 \mathbb{Q} 上の線形空間、 \mathbb{R} 上の線形空間となっています。」



「ん？ 実数が有理数上の線形空間ってことは実数が有理数のベクトルとして見る事ができるってこと??」



「はい、そうです。では、実数を線形空間として見たとき、その次元はどうなるでしょうか？ \mathbb{R} 上の線形空間としては 1 を選んでやれば、どんな実数 x も $x = x \cdot 1$ となり、基底となっているので 1 次元です。では、有理数上の実数の線形空間の次元はどうでしょうか？ 実は無限次元になります。なぜならば、素数の平方根

$$\sqrt{2}, \sqrt{3}, \sqrt{5}, \dots, \sqrt{p_n}, \dots$$

が有理数上で線形独立になっているからです。まとめると次の事が言えます。」

命題 2.5. \mathbb{R} は \mathbb{Q} 上線形空間で、無限次元。

*1 $A \subset B$ は集合 A が B に含まれるという意味で、任意の $x \in A$ に対して $x \in B$ が成立することを言う。

*2 斉藤毅 線形代数の世界 定理 1.5.4、現代数学概説 I の第 6 章 §1 定理 2 を参照。



「あれ、これどこかで見たことある…*3」



「以前のスタインハウスの問題 n 次元を解く点ですね。実はこの問題は素数の平方根の集まりが線形独立であることを本質的に使ってます。また、 \mathbb{Q} 上の線形空間を考えることで 5 次以上の方程式には解の公式が存在しないことや、定規とコンパスで何が作図できるかを考えられます。*4」



「へー！ 無味乾燥な定義だと思ったけど、色々なところに繋がってるのね！」



「そうなんです！ そこが数学の面白さなんですよ！」

*3 「ろーほるすーがくれっすん」の §3 のスタインハウスの問題を n 次元に拡張したもの

<https://irisu-in-wonderland.tumblr.com/post/156533872627>

*4 E.Artin, Galois Theory に載っているような現代的なガロア理論のこと。群はこの研究が初出と言われており、 \mathbb{Q} の方程式が解ける体 K を考え、その線形空間と群が対応されるので、群について考えればよいという理論である。

3 くりくり♡びぶんにゆうもん♪

3.1 微分のイントロダクション



「それでは微分について学んでいきます。」



「微分ってあれでしょ、ドラマでよく頭良さそうな人がやってる如何にも理系～な感じのやつでしょ？」



「黒板に難しい数式がビッシリ書いてあるやつやな～」



「とっても難しそうな…」

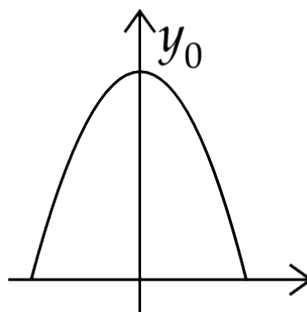


「そうですね…多分全然馴染みのない概念だと思いますので、まず簡単な具体例を考えてみましょう。

物体を地点 y_0 から落下させることを考えます。そのときの地上からの高さは時間 t に対して

$$y(t) = y_0 - \frac{1}{2}gt^2 \quad (g \text{ は定数})$$

となることが実験的に分かっています。グラフを描くと以下のようになります。」



「理科の実験とかであるやつやな～」



「この y について、時間 t_1 での速度を求めてみましょう。ただし $y(t_1) > 0$ です。速度を求めるために次のように考えてみます。小学校の計算では

$$(\text{距離}) = (\text{速度}) \times (\text{時間})$$

だったんで速度 \bar{v} は t_1 秒から $t_1 + h$ 秒後の変化に、どれくらい移動するかの割合（変化率）として考えてみます。」



「つまり、式で書くと

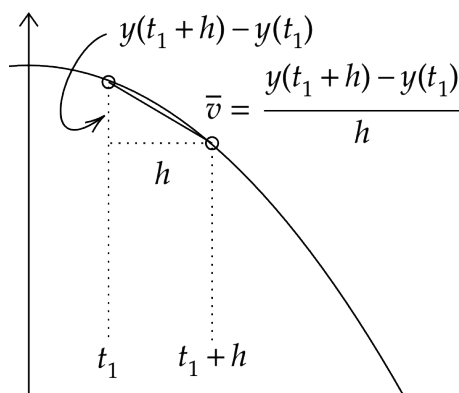
$$\begin{aligned}\bar{v}h &= y(t_1 + h) - y(t_1) = -(y_0 - y_0) - \frac{1}{2}g(t_1 + h)^2 + \frac{1}{2}gt_1^2 \\ &= -\frac{1}{2}g(t_1^2 + 2t_1h + h^2 - t_1^2) = -\frac{1}{2}g(2t_1h + h^2)\end{aligned}$$

よって

$$\bar{v} = -g(t_1 + \frac{1}{2}h)$$

と考えます。これを t_1 秒から $t_1 + h$ 秒の平均速度といいます。

これは、 $y(t_1)$ と $y(t_1 + h)$ のグラフの点を結んだ直線の傾きになってます。」



「たしかにこれで t_1 から $t_1 + h$ の地面からの位置は求められそうだけど、 t_1 から $t_1 + h$ からの間は速度は変わり続けているからズレが生じるんじゃない？」

距離 = 速度 × 時間の式みたいに単純に距離が求められないから速度としては不十分じゃないかしら？」



「その通りです、平均速度はあくまである区間での平均した速度となっているため、区間内の変化を捉えることができません。

そこで、 h を非常に小さくした瞬間速度を考えます。今、 t_1 から $t_1 + h$ の平均速度 \bar{v} が

$$\bar{v} = -g(t_1 + \frac{1}{2}h) = -gt_1 - \frac{1}{2}gh$$

でしたが、 h が非常に小さいので、第2項の $\frac{1}{2}gh$ は0に近づくので、瞬間速度 v は

$$v = -gt_1$$

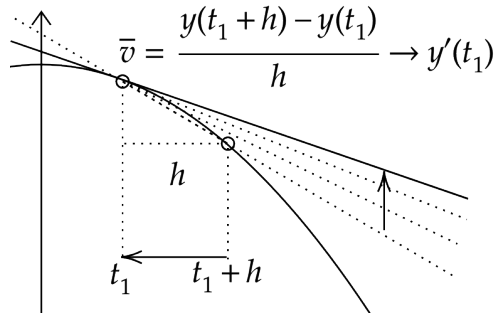
となります。ここで、 h を0に近付けたときに \bar{v} が $-gt_1$ に近づくということを

$$\lim_{h \rightarrow 0} \bar{v} = -gt_1$$

または

$$\bar{v} \rightarrow -gt_1 \quad (h \rightarrow 0)$$

と書き、 \bar{v} の極限といいます。」



「言われてみるとそうなりそうやなあ」



「ここで、平均速度 t_1 から $t_1 + h$ の点を結ぶ直線の傾きとして見れましたが瞬間速度は t_1 の点における接線の傾きになっています。

さて、この瞬間速度を求めた方法をどんな関数 $f(x)$ にも適用することを考えます。関数 $f(x)$ の $x = a$ から $x = a + h$ での平均変化率を求めると

$$\frac{f(a + h) - f(a)}{h}$$

となります。」



「ここで、 h を限りなく 0 に近付けた

$$\lim_{h \rightarrow 0} \frac{f(a+h) - f(a)}{h}$$

の値が存在すれば、それを $f(x)$ の $x = a$ における瞬間変化率、導値、微分係数と言ひ、この値を求めることを微分するといいます。 f の $x = a$ における微分係数を $f'(a)$, $\frac{d}{dx}f(a)$ など書きます。微分係数は瞬間速度と同じように、その点での接線の傾きとなっています。

f に対して微分係数が定義できる点から微分係数を対応させた関数として見ることができ、これを導関数といいます。導関数の値は各点での傾きとなるので、0 となることで平らとなり、周りの値よりか高いか低いかが分かります。これが微分と最大最小との関係を示唆する意味となっております。」



「坂を上ってるときに、傾きが緩くなったらもしかしたら頂上かもって錯覚する感じがやな！」



「限りなく 0 に近い数…想像つかないの…」



「そうですね、ツキノの言う通り、限りなく 0 に近づくという操作が非常に曖昧であり、数学では極限という概念は 18 世紀に誕生したのですが、19 世紀ではその定義を巡って物議をかもしました。

そして、この単純な極限の定義（無限小式）は見直されて厳密な定義（ $\epsilon\delta$ 式）が与えられました。しかし、20 世紀になり、数学が発展すると旧来の定義も $\epsilon\delta$ 式も同じであることが証明されました。^{*1}」



「ふーん、微分って昔からあるのね。」



「そうですね、200 年前からありますが、微分は数学の歴史上大きな影響を与えたのは確かです。定義だとわかりづらいので、単純な定義で、具体例を計算してみましょう。」

^{*1} 17 世紀、フェルマーやデカルトによって関数の接線を求める研究がされ、ニュートン、ライプニッツによって微分という手法は定式化された。そして極限の取り扱い（ $\epsilon\delta$ 論法）、実数の性質を見直し厳密に定めたのはコーシー、ワイエルシュトラス、デデントらの 19 世紀の数学者であった。そして、20 世紀、ロビンソンによって数学基礎論を用いて、それまでの $\epsilon\delta$ 論法と旧来の無限小解析が両立する超準解析が構成された。

参考: ブルバキ, 数学史

■例: 微分を計算してみよう



「では、次の関数を微分してみましょう。」

(1) $f(x) = C$ (C :定数) について $f'(x_0)$

(2) $f(x) = x^n$ について $f'(x_0)$

(3) $f(x) = |x|$ について $f'(0)$

(1)



「まずは (1) ですが、これは簡単なので、計算してみてください。」

「やってみるで～ええと、微分の定義の通り計算すると…



$$f'(x_0) = \lim_{h \rightarrow 0} \frac{f(x_0 + h) - f(x_0)}{h} = \lim_{h \rightarrow 0} \frac{C - C}{h} = 0$$

やな！」



「0 になっちゃったわね。」



「そうですね、この関数を見てみると、とくに値が変化せずに一定の値なので、変化率は 0 となります。よってどんな点の微分係数も 0 となります。」

(2)



「続いて (2) を解きますが、これを解く前に次の命題を示しましょう」

命題 3.1. 任意の実数 a, b について $n \geq 2$ について

$$(a + b)^n = a^n + na^{n-1}b + R^{(n-2)}(a, b)$$

ただし、 $R^{(n-2)}(a, b)$ は

$$R^{(n-2)}(a, b) = \sum_{i=0}^{n-2} c_i^{(n-2)} g(b) a^i$$

と書け、 $b^2 | R^{(n-2)}(a, b)$ となる。^{*1}



「なんか変な記号でたの、 $b^2 | R^{(n-2)}(a, b)$ ってなんなのなの？」



「これは割り切るといって $p|q$ は $q = pr$ となるような r が存在することを言います。^{*2}」

証明 $n = 2$ のとき、

$$(a + b)^2 = a^2 + 2ab + b^2$$

となり、 $R^{(0)}(a, b) = b^2$ とすれば主張は正しい。

$n - 1$ のときに成立するとして、 n のとき

$$\begin{aligned} (a + b)^n &= (a + b)^{n-1}(a + b) = (a^{n-1} + (n-1)a^{n-2}b + R^{(n-3)}(a, b))(a + b) \\ &= (a^n + (n-1)a^{n-1}b + aR^{(n-3)}(a, b)) + (a^{n-1}b + (n-1)a^{n-2}b^2 + bR^{(n-3)}(a, b)) \\ &= a^n + na^{n-1}b + (aR^{(n-3)}(a, b) + (n-1)a^{n-2}b^2 + bR^{(n-3)}(a, b)) \end{aligned}$$

となる。ここで $R^{(n-2)}(a, b) = aR^{(n-3)}(a, b) + (n-1)a^{n-2}b^2 + bR^{(n-3)}(a, b)$ とすれば、 $aR^{(n-3)}(a, b) = \sum_{i=0}^{n-3} c_i^{(n-3)} g(b)a^{i+1}$ であり、 a のべき乗の数は $n - 2$ を超えない。また、 $b^2|aR^{(n-3)}(a, b), b^2|(n-1)a^{n-2}b^2, b^2|bR^{(n-3)}(a, b)$ なので $b^2|R^{(n-2)}(a, b)$ となる。よって、示された。□



「なんか不思議な証明なの。」



「このように n までの式で一番最初（今回の場合は $n = 2$ ）を示し、 $n - 1$ を示したあとに n を示す証明法を数学的帰納法といいます。このように証明すればどんな自然数の n にも証明することができます。」



「ええっと、前に教えてもらったっけ？」

^{*1} 一般に $\sum_{i=0}^n c_i X^i, c_i \in \mathbb{R}$ となる式を n 次の X について多項式といい、その集合を $\mathbb{R}[X]$ とかく、また各 $c_i X^i$ を単項式という。とくに $R^{(n-2)}(a, b)$ は 2 変数多項式で $\in \mathbb{R}[a, b]$ である。

^{*2} 前作「ろーほろすーがくれっすん」(<https://irisu-in-wonderland.tumblr.com/post/156533872627>) にも同様の記号が登場するが、前作は整数 \mathbb{Z} の範囲だった。

今回の $b^2|R^{(n-2)}(a, b)$ は $\mathbb{R}[a, b]$ で考えてる。このように整数と多項式の間で足し算と掛け算のみに着目すれば同じことを考えられる。掛け算と足し算が定められ、全ての元に逆数が存在するとは限らない構造を環という。



「そうですね…私もうろ覚えですが教えた気がします…*3

では、準備がそろいましたので、 x^n の微分をしてみましょう。まず $n = 1$ の場合です。
これは簡単で

$$f'(x_0) = \lim_{h \rightarrow 0} \frac{f(x_0 + h) - f(x_0)}{h} = \lim_{h \rightarrow 0} \frac{x_0 + h - x_0}{h} = 1$$

となります。続いて $n \geq 2$ の場合はどうなるかというと…

$$f'(x_0) = \lim_{h \rightarrow 0} \frac{f(x_0 + h) - f(x_0)}{h} = \lim_{h \rightarrow 0} \frac{(x_0 + h)^n - x_0^n}{h}$$

ここで $(x_0 + h)^n$ に 3.1 を用いて

$$\begin{aligned} \lim_{h \rightarrow 0} \frac{(x_0 + h)^n - x_0^n}{h} &= \lim_{h \rightarrow 0} \frac{(x_0^n + nx_0^{n-1}h + R^{(n-2)}(x_0, h)) - x_0^n}{h} \\ &= \lim_{h \rightarrow 0} \frac{nx_0^{n-1}h + R^{(n-2)}(x_0, h)}{h} \\ &= \lim_{h \rightarrow 0} \left(nx_0^{n-1} + \frac{R^{(n-2)}(x_0, h)}{h} \right) \end{aligned}$$

となります。今、 $R^{(n-2)}(x_0, h) = h^2 Q(x_0, h)$ と表せるので、

$$\lim_{h \rightarrow 0} \frac{R^{(n-2)}(x_0, h)}{h} = \lim_{h \rightarrow 0} h Q(x_0, h) = 0$$

となり、よって

$$f'(x_0) = nx_0^{n-1}$$

が証明されました。」



「とっても盛沢山ね…」



「せやなあ～計算複雑やったわ～」

「実は命題 3.1 は厳密には

$$(a + b)^n = {}_nC_0 a^n + {}_nC_1 a^{n-1}b + {}_nC_2 a^{n-2}b^2 \cdots + {}_nC_{n-1} a b^{n-1} + {}_nC_n b^n$$



とかけます。ここで ${}_nC_r$ は

$${}_nC_0 = 1, {}_nC_n = 1, {}_nC_r = \frac{n \cdot (n-1) \cdots (n-r+1)}{r \cdot (r-1) \cdots 1}$$

です。」

*3 「ろーほるすーがくれっすん」 (<https://irisu-in-wonderland.tumblr.com/post/156533872627>) の付録 A 参照



「あ、なんか教科書でみたことあるかもしれんなあ〜」

(3)



「あと、もう少しの辛抱ですよ！ 最後の問題を解いてみましょう。」



「 $f(x) = |x|$ の $f'(0)$ を求めるんだよね？ でも、 $f(x) = |x|$ って $x > 0$ のとき $f(x) = x$ だから、(2) から $f'(0) = 1$ なんじゃないの？」



「たしかにそう考えられます。しかし、次のようにも考えられます。 $x < 0$ のとき $f(x) = -x$ だから $x < 0$ から近付けたらその微分係数は $f'(0) = -1$ になりませんか？」



「言われてみればそうね…どっちが正しいんだろ…」



「このように、極限が近づく向きによって答えが変わるケースがあります。特に近づく方向を明示するために x が a に $a < x$ から近づく場合は

$$\lim_{x \rightarrow a+0} f(x), \lim_{x \rightarrow a, x > a} f(x)$$

などと書きます。最初にロージアが考えたのがこちらですね。

また、 x が a に $a < x$ から近づく場合は

$$\lim_{x \rightarrow a-0} f(x), \lim_{x \rightarrow a, x < a} f(x)$$

と書きます。」



「結構ややこしいの…これって結局どっちが正しいの？」



「実は、どっちも正しいです。 $x \rightarrow a$ は $x \rightarrow a+0$ と $x \rightarrow a-0$ の値が等しい時のみ定義されます。つまり

$$\lim_{x \rightarrow a} f(x) = \lim_{x \rightarrow a, x < a} f(x) = \lim_{x \rightarrow a, x > a} f(x)$$

となります。逆に異なる場合は極限を定義できないことになります。

つまり、この問題は

$$\lim_{h \rightarrow 0, h > 0} \frac{f(h) - f(0)}{h} = 1, \lim_{h \rightarrow 0, h < 0} \frac{f(h) - f(0)}{h} = -1$$

となり、極限は存在しないことになります。」



「へ～極限が存在しないんやな～そんなんあるんやな！」



「解答が存在しないってなんか反則じゃない？」



「いえ、数学の世界では存在しないことを証明することはよくありますよ？」



「せやで～目潰しも金的もありのバリトウードの世界なんやで？」



「やってやるなの～」



「バリトウードで…」

■例：微分で関数の大まかな形を見てみよう



「先ほど、微分を使えば関数がどこで平らになるか、どこで傾いてるかがわかるといいました。関数を微分することで関数の大まかな形（概形）をみることができます。たとえば $y = (x - a)(x - b)^2$ を考えてみましょう。」



「これってたしか $x = a, x = b$ で 0 になるような二次関数よね。」



「そうですね。では、導関数を求めることでその形になることが大まかにわかることを見ていきます。 $y = (x - a)(x - b)^2$ を展開すると

$$(x - a)(x - b)^2 = x^2 - (a + b)x + ab$$

となります。実は、関数の微分は各項の微分となることが分かってます*1。つまり、

$$(x^2 - (a + b)x + ab)' = (x^2)' - ((a + b)x)' + (ab)'$$

よって、先ほどの例から

$$(x^2 - (a + b)x + ab)' = 2x - (a + b)$$

となります。」



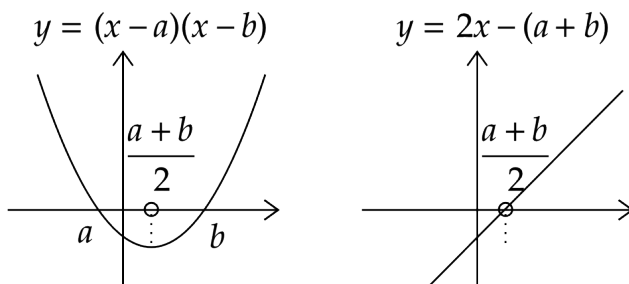
「これが $y = (x - a)(x - b)$ のかたちになってるの…？」



「一見わかりませんが、導関数を見ていきましょう。導関数 $y' = 2x - (a + b)$ は $x = \frac{a+b}{2}$ で 0 になり、その前後で - から + に代わってます。

| | | | | | |
|------|-----------|---------------------|---------------------|---------------------|----------|
| x | $-\infty$ | $x < \frac{a+b}{2}$ | $x = \frac{a+b}{2}$ | $\frac{a+b}{2} < x$ | ∞ |
| y' | $-\infty$ | - | 0 | + | ∞ |

つまり、傾きが $x = \frac{a+b}{2}$ より左では左斜め下で、 $x = \frac{a+b}{2}$ で平らになり、より右にいくと右上に上がるようになっています。さらに + に行けば行くほど y' の値は大きくなり - に行けば行くほど小さくなるので $|x|$ の値が大きくなるほど傾きは急になっていきます。」



「こういう感じで微分を使えば関数の形が分かるんやね～」



「そうですね、二次関数だけでなく、三次関数、四次関数、さらに一般の関数についてもその大まかな形がわかります。」



「例を見てわかる通り、微分を計算するのに極限について、厳密に考える必要があります。極限とは何なのか、さらには実数とは何なのかについて考えていきましょう！」

*1 関数 f, g について $(f(x) + g(x))' = f(x)' + g(x)'$ となる。これは §2.2.1 で示す。