atmarkIT

AI・機械学習のための数学超入門

～中学数学からのおさらい編～

羽山博［著］

01. AI・機械学習のための数学超入門 ―― 前提知識は四則演算だけ！

02.［AI・機械学習の数学］文字式を使いこなせば一気にレベルアップ

03.［AI・機械学習の数学］総和を表すΣは機械学習に必須の記号

04.［AI・機械学習の数学］番外編１　「0」の取り扱い

05.［AI・機械学習の数学］番外編２　指数と対数（指数編）

06.［AI・機械学習の数学］番外編３　指数と対数（対数編）

AI・機械学習のための数学超入門
――前提知識は四則演算だけ！

機械学習の数学は難しい!?　そう思っている人はここから学んでみよう。本連載では、小学校で習う「四則演算（足し算／引き算／掛け算／割り算）」を使って、機械学習の数学をできるだけ分かりやすく簡単に説明していく。だからサブタイトルは「――中学／高校数学のキホンから学べる」。今回は距離を求める中学数学をおさらいする。

羽山博，著（2020 年 03 月 02 日）

機械学習は、人間の日常とそう変わらない

　この連載を読もうとする方々は機械学習について興味をお持ちなので、回帰とか分類といった用語は既にどこかで見聞きしていることと思います。また、機械学習がどういうものかというイメージもそれぞれお持ちだと思います。しかし、多くの人が、機械学習を必要以上に難しく捉えすぎなのではないか、と最近つくづく感じています（私自身、そうでした）。そこで、もう少し身の丈に引き寄せて、機械学習を捉え直してみることにします。

　実際のところ、機械＝コンピューターがやっていることは、人間のやっていることとよく似ていて、意外に単純です（むしろ人間がやっていることの方がはるかに複雑です）。簡単な例で見てみましょう。皆さんは「数当てゲーム」をやったことがあると思います。あまりにも単純すぎて、オトナになってからはやらないかもしれませんが、プログラミングの練習問題としてやらされたことのある人もいるでしょう。

　試しにやってみましょう。今、私が考えている数字を当てて下さい。数字は **0** ～ **99** のどれかです。……といっても、読者の皆さんと私は直接対話しているわけではないので、皆さんの代表として誰か（A さんとします）に答えてもらうことにします。

A さん　それじゃ、**50** から

私　まあ、そうきますよね。もっと下です

A さん　じゃあ **25**

私　もうちょっと上です

A さん　**37** ぐらい？

私　今度はもうちょっと下です

A さん　**31**

私　もうちょっと下

A さん　**28**

私　近い。もうちょっとだけ上

A さん　**29**

私　正解です！

ここでAさんがやっていることは、機械学習ととてもよく似ています。Aさんは適当な初期値を決め、私からのヒントを基に値を調節していきます。そして、最適な解を見いだす、というわけです *1。ちなみに、Aさんは無意識のうちに「候補を半分ずつに絞っていく」という方法を採っています。これは（名前を覚える必要はないですが）**二分法**と呼ばれるアルゴリズムで、数字が **100 個**であれば **7 回**までの挑戦で確実に数字が当てられます（なぜ **7 回**なのかは、対数を使えば簡単に計算できますが、それは回を改めて説明します）。

*1　プログラミングの練習問題として作る数当てゲームでは、**人間**が数値を予測します（これは上記の「Aさん」の役割です）。**コンピューター**は、「その予測値が正解よりもどれくらい上か下か」というヒントを人間に提示します（これは上記の「私」の役割です）。そして、人間が正解に向けて数値を調整していきます。

この「数当てゲーム」の流れは、各役割を担う担当者が違うというだけで、機械学習でも同じです。

機械学習では、**コンピューター**が数値を予測します（＝「Aさん」の役割）。「その予測値が正解よりもどれくらい上か下か」というヒントは、（基本的に**人間**が作成した）正解情報から自動的に計算され、コンピューターに提示されます（＝「私」の役割）。そして、コンピューターが正解に向けて数値を調整していきます。

いや、そんなのは機械学習じゃないぞ、と言われる向きもあるかと思います。確かに二分法は単純なアルゴリズムなので、「機械学習」から見るとミジンコのようなものかもしれません（リアルなミジンコはとても複雑ですが）。しかし、出発点としてはこういうイメージで十分かと思います。つまり、**何らかの方法に従って少しずつ値を変えていき、最適な解を得る**ということですね。機械学習なんて数当てゲームの延長にすぎない……と考えれば、少しは身近に感じられるのではないでしょうか。学びを深めていけば、機械学習がどういうものであるかというイメージも少しずつ更新していけるでしょう。

機械学習の種類として登場する回帰や分類といった用語についても、上のような例から「実はカンタンなことなんだよ」というお話をしてもいいのですが、ダラダラと前置きを話していても退屈なだけなので、そろそろ本題に入っていくこととしましょう（必要に応じて、その都度、説明します。また、皆さんなりに、日常のこういうことと同じだよね、という例を探してもらうのもいいでしょう）。

→目次に戻る

機械学習で使われる数学は結局のところ四則演算だけ！

ところで、この連載のテーマは機械学習そのものではなく、「機械学習で使われる数学」です。これについても線形代数とか微分法とかが使われる、といった話はご存じでしょう。しかし、それらの用語を聞くだけで「難しい数学はちょっと……」と腰が引けてしまう人も多いようです。そこで諦めてしまうのはとても残念なので、この連載では、多くの人が理解できる中学／高校の数学レベルから機械学習に必要最小限の部分だけをおさらいして、難しそうな数式が書かれた機械学習の本を抵抗なく読めるようにしようというわけです ***2**。

***2**　こういう連載記事を書いているのだから、さぞかし著者は数学の得意な人だろう、と思われるかもしれません。しかし、私はドイツ文学をやりたくて文学部に入った人間です（なのに、第二外国語はフランス語を選択し、結局、哲学科心理学専攻に進みました……というのは余談ですが）。高校時代、数学の実力テストで **100 点**満点中たったの **17 点**しか取れなかったという輝かしい記録を持っているほど数学が苦手でした。それでも、東京大学やお茶の水女子大学で情報やアルゴリズムの授業を担当させてもらっているということは、純粋な数学は別として、道具としての数学はそれほど難しくないということなのでしょう。また、苦手だったからこそ、「どこでつまづくか」がよく分かっているつもりです。

実は、

> 線形代数にしても、微分法にしても、結局のところ**四則演算**をやっているだけ

なのです。ただ、何らかのパターンに従って計算するときに、いちいち「これらの値を足して、あの値で割って……」と、くどくど言うのが面倒なので、ちょっと変わった記号を使って簡潔に書くだけのことなのです。

そういうわけで、今回は四則演算を出発点として、「中学／高校で学ぶ数学の基本」からおさらいしていきましょう。既にお気付きだと思いますが、一歩ずつ皆さんに語りかけるように進めていくので、「お話のペースが遅い」と感じる人がいるかもしれません。しかし、**何かをマスターするための最短ルートは「ゆっくりやる」こと**です。音楽やスポーツを習ったことのある人は、ゆっくりと正確にやるのが最良の方法だと教わったことがあるでしょう。それは数学でも同じです。「自分が思っている自分のペース」は意外に速いものです。「回り道だ」とか「無駄だ」とあまり思わずに、ゆっくりとかみくだき、きちんとおなかの中に収めるように心がけるようにしましょう（食べやすくて、消化のよい説明を心がけます！）。

→目次に戻る

各項目には「目標」と「解説」と「練習問題」があります。「目標」を既に理解している人は、次の解説を読む必要はありません（でも、意外な発見があるかもしれないので、ぜひとも読んでみてください）。「練習問題」は必須ではありませんが、手を動かしてみたり、考えてみたりすることは、知識を確実に身につけるのに役立ちます。図解や式の変形が複雑になる場合には簡単な動画も用意してあります。動画では一歩ずつ手順を追いかけているので、すんなりと理解できない場合に参照してみてください。

四則演算のルールをざっとおさらい

目標

今回の目標は、四則演算のルールの確認です。「ハードルはできるだけ低く」をモットーにしてお話を進めるので、少しお付き合いください。

- **同じ数を何回か足す計算は、掛け算で表せる**
- **同じ数を何回か掛ける計算は、べき乗（詳細後述）で表せる**
- **計算には優先順位があり、かっこで囲んだ計算が優先される**

取りあえず、最低限の基本ということで、これだけを確認しておきましょう。これらのルールを確認した後、機械学習で使われる例として、「距離を求める」ための計算を見ていきます。

解説

目標に掲げたルールは、いくら数学が苦手な人でもなんとなく覚えているでしょうし、無意識に使っていることと思います。しかし、あらためて意識に上らせておくと、これからの学びが驚くほどスムーズになります。この連載のサブタイトルには「中学／高校数学のキホン」とありますが、まずは小学校の算数（四則演算：足し算／引き算／掛け算／割り算）から始めます。目標に掲げたルールを実例で確認しておきましょう。

ルール 1　同じ数を何回か足す計算は、掛け算で表せる

$$\text{例: } 2+2+2=2\times 3 \quad \ldots\ldots \quad 2 \text{ を } 3 \text{ 回足す} \rightarrow 2 \text{ 掛ける } 3$$

あまりに簡単だと、読むスピードが速くなりすぎてしまうので、穴埋め問題もやりながら着実に進めていきましょう。□の中には何が入るでしょうか。

$$3+3+3+3=3\times\square$$

はい、答えは 4 ですね。**3** を **4 回**足しているので、3 × 4 です。これらの式を見ると、掛け算は「足し算の延長」で、「決まったパターンの足し算を簡単に書く方法」と考えることもできます。上の例は簡単すぎるので、掛け算のありがたみが分かりませんが、例えば、**3** を **100 回**足すとしたらどうでしょう。

$$\underbrace{3+3+3+\cdots+3}_{100\text{個}}$$

このように書くのは、足し算が多すぎて面倒ですね（あまりにも面倒なので途中を省略してしまいました。数学的にちょっとイジワルなことを言うと「… 」の部分があいまいです)。それより、3 × 100 と表した方がはるかに簡単で、正確です。

さらに、**掛け算は面積を求めるためにも使われます。**図 1 のような長方形の面積は 3 × 4 = 12 となります。4 × 3 でも結果は同じです。

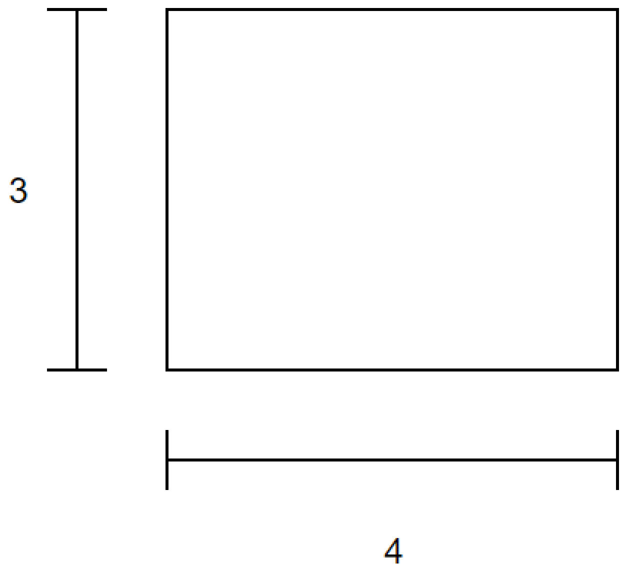


図 1　掛け算は面積の計算にも使われる

割り算も、「**0** で割っちゃダメ」というルールはありますが、「掛け算の延長」として捉えることができます。割り算は、**逆数** ***3** を掛けると考えればいいというだけのことです。

*3　例えば、元の数の逆数は、

$$\frac{1}{\text{元の数}}$$

という分数で表されます。

　例えば、3 の逆数は、

$$\frac{1}{3}$$

です。

　例えば、$6 \div 3$ という割り算は、どういう逆数の掛け算になるでしょうか？

$$6 \div 3 = 6 \times \frac{1}{3} = 2$$

ですね。

　さらに、**6** から **3** を何回引けるかということだと考えれば、割り算は「引き算の延長」にすぎないともいえます *4。その引き算は「負の数を足すこと」と考えれば、「足し算の延長」になります。数学的に細かな議論はさておき、**結局は全て「足し算の延長」**だと捉えることもできますね。

*4　$6 \div 3$ という割り算の場合、$6 - 3 = 3$、$3 - 3 = 0$ というように **6** から **3** が「**2** 回引ける」ので、やはり答えは 2 となります。

　全くの余談ですが、電卓を使って **0** で割り算をすると、当然のことながらエラーが表示されます。しかし、初期の電卓には、**0** で割るとパニックを起こしたかのように数字がめまぐるしく変わるものがありました。おそらく、「元の数」から「割る数」を何回引けるか、を数えていたのでしょう（**0** を引いても元の数のままなので、いつまでたっても計算が終わらず、回数だけがどんどん増えていくというわけです）。私が小学生の頃、それを偶然発見し、$1 \div 0$ などを計算させて、延々と数字が変わるのを飽きずに眺めていた記憶があります（残念ながら、その不思議体験は数学的才能の開花には結び付きませんでしたが）。

　ちなみに、先ほどは長方形の面積を計算しましたが、それ以外の「面積の計算」には**積分**という計算方法を使うケースもあります。**積分**では、足し算と掛け算を使います。例えば、（本連載で解説する）確率の計算（＝確率分布のグラフの面積を求める計算）などには積分が使われます。

→目次に戻る

また、（本連載で解説する）**微分**という計算方法では、引き算と割り算を使います。**微分**は「変化率を求める」ことと考えられます。

難しそうに見えますが、結局は四則演算なのです。

では、次のルールに進みましょう。今度は中学校にレベルアップします。

ルール 2　同じ数を何回か掛ける計算は、べき乗で表せる

$$\text{（例）}\quad 2\times2\times2=2^3\quad\ldots\ldots\quad 2\text{ を }3\text{ 回掛ける}\to 2\text{ の }3\text{ 乗}$$

念のための確認ですが、「2 の 3 乗」は「にのさんじょう」と読みますね。

これも、穴埋め問題で確認しておきましょう。□の中には何が入るでしょうか。

$$5\times5\times5\times5\times5\times5=5^{\square}$$

簡単ですね。答えは 6 です。**5** を **6 回**掛けているので、**5** の **6 乗**ということになります。**べき乗**は四則演算そのものには含まれませんが、「掛け算の延長」なので四則演算に毛が生えた程度のものと考えればいいですね。掛け算のことを「**乗算**」とも呼ぶのはご存じでしょう。そのことからも「**6 乗**」なら「**6 回**掛ける」ことだというのが分かります。

負の数を **2 乗**すると正の数になることも重要です。

$$\begin{aligned}4^2&=16\\(-4)^2&=16\end{aligned}$$

「何の何乗」における「何乗」に当たる数のことを**指数**（しすう）と呼びます。4^2 であれば、指数は **2** です。

また、今は覚える必要はありませんが「何の」に当たる数のことを**底**（てい）と呼びます。4^2 であれば、底は **4** です。

ルール 3　計算には優先順位があり、かっこで囲んだ計算が優先される

最後は、計算の順序に関するルールです。このルールに関しては、もう少し詳細に記しておきましょう。

- **足し算や引き算よりも、掛け算や割り算が優先する**
- **掛け算や割り算よりも、べき乗が優先する**
- **かっこで囲んだ計算が優先する**

足し算や掛け算などの記号は、一般に**演算子**と呼ばれます。演算子は優先順位が同じであれば、たいていは「左から右へ」と計算していきますが、べき乗は「右から左へ」（「上から下」と言った方がいいでしょうか）と計算します。優先順位の異なる演算子が式の中にある場合は、優先される計算の方を先に行います。優先順位を上げたいときにはかっこで囲みます。幾つかの例で確認しておきましょう。

$$3+2-1=4 \qquad \cdots \ (\ A\ )$$
$$1+2\times 3=7 \qquad \cdots \ (\ B\ )$$

（A）足し算と引き算は同じ優先順位なので、左から右へと計算する。よって、先に 3 + 2 (= 5) を計算して、次に 5 − 1 (= 4) を計算すればよい

（B）足し算よりも掛け算が優先するので、先に 2 × 3 (= 6) を計算する。次に 1 + 6 (= 7) を計算すればよい

上の例では、掛け算が足し算よりも優先されました。足し算の方を優先させるには () で囲みます。

$$(1+2)\times 3=9 \qquad \cdots \ (\ C\ )$$

（C）先に 1 + 2 (= 3) を計算する。次に 3 × 3 (= 9) を計算すればよい

コラム　計算の順序を工夫すると暗算が簡単になる

　ちょっとしたオマケの話ですが、$(1+2)\times 3$ の計算をするときに、**分配法則**を使うこともできます（図2）。$(1+2)$ を先に計算せずに、$\times 3$ を、文字通り 1 と 2 に分配して、$1\times 3+2\times 3$ という計算にしても同じ結果が得られます。

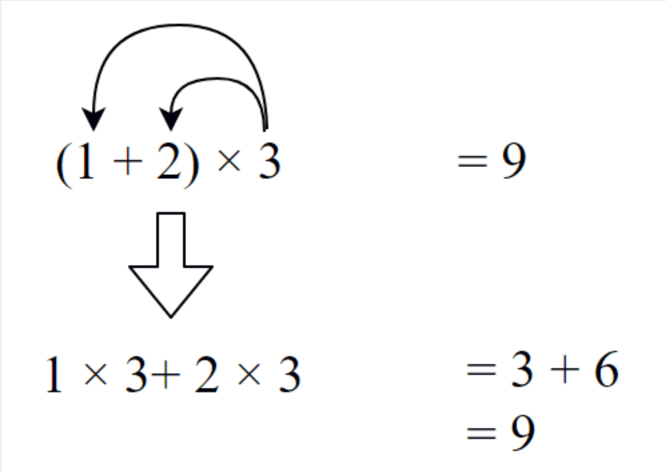


図 2　分配法則

足してから掛ける代わりに、個々の値に掛けてから足しても同じ。

　この法則を使うと、999×8 といった計算が暗算でも簡単にできるようになります。$999=1000-1$ なので、次のように計算できます。

$$
\begin{aligned}
999\times 8 &= (1000-1)\times 8\\
&= 1000\times 8-1\times 8\\
&= 8000-8\\
&= 7992
\end{aligned}
$$

　本筋に戻りましょう。基本的なルールを一通りおさらいしたので、念のため、練習問題にも取り組んでみましょう（※スキップしたい場合は、14 ページ目にジャンプしてください）。

練習問題

以下の$\square$に適切な数字や文字を入れましょう。

(1)　$7+7+7+7+7$は$7\times\square$と表せる

(2)　$8\times8\times8$は$8^{\square}$と表せる

(3)　$3+4\times5=\square$

(4)　$(3+4)\times5=\square$

(5)　$2^3-1=\square$

(6)　$2^{2^3}+1=2^{\square}+1=\square$

解答例

「四則演算ルール」練習問題の解答例は次の次ページに示します。14 ページ目からは「距離を求める」について説明します。

「四則演算ルール」練習問題に対する解答例です。

解答例

（1）　$7+7+7+7+7$ は、

$$7 \times \boxed{5}$$

と表せる

7 を **5 回**足しているので、掛け算で表すと 7×5 です。計算すると **35** となります。

（2）　$8 \times 8 \times 8$ は、

$$8^{\boxed{3}}$$

と表せる

計算すると **512** です。なお、8^4 は **4,096**、8^5 は **32,768**、8^6 は **262,144**、8^7 は **2,097,152** です。このように、**指数が少し大きくなるだけで、値が莫大なものになっていく**ことが分かります。

（3）

$$3+4 \times 5 = \boxed{23}$$

掛け算が優先されるので、先に $4 \times 5 = 20$ を計算します。

$$\begin{aligned} 3+4 \times 5 &= 3+20 \\ &= 23 \end{aligned}$$

（4）

$$(3+4) \times 5 = \boxed{35}$$

かっこ内を先に計算します。

$$\begin{aligned} (3+4) \times 5 &= 7 \times 5 \\ &= 35 \end{aligned}$$

（5）

$$2^3 - 1 = \boxed{7}$$

$2^3 = 2 \times 2 \times 2 = 8$ なので、以下のようになります。

$$\begin{aligned} 2^3 - 1 &= 8 - 1 \\ &= 7 \end{aligned}$$

（6）

$$2^{2^3} + 1 = 2^{\boxed{8}} + 1 = \boxed{257}$$

べき乗のべき乗は右上から計算します。したがって、 2^3 を先に計算します。

$$2^{2^3}+1=2^8+1=256+1=257$$

またまた余談ですが、Ruby や Python といった多くのプログラミング言語でも、べき乗の演算子の計算は右から左に行われます。ただし、Excel では、左から右に計算されることに注意が必要です。Excel で「=2^2^3」と入力すると、**256** ではなく、**64** という答えが表示されます。

* 余談　もう一つ、機械学習とは関係のない雑学ですが、上記の（5）番の問題のような「2 の何乗 -1」の形式の数を**メルセンヌ数**と呼び、(6) 番の問題のような「2 の (2 乗の何乗)+1」の形式の数を**フェルマー数**と呼びます。なお、問題例の **7** と **257** はいずれも**素数**（**= 1** より大きい自然数で、正の約数が **1** と**自分自身**のみである数のこと）になっていますが、メルセンヌ数とフェルマー数は必ずしも素数というわけではありません。

距離を求める　～　機械学習で使われる数学（1）

四則演算のルールを確認しただけでは、機械学習のどういう場面で役に立つのかが分からないでしょう。そこで簡単な利用例として、次は、距離の計算方法を見ておきましょう。今回は、次の章で終わりです。もう少しだけ頑張って数学の学習を進めましょう。

機械学習では、ほとんど全ての手法で距離を求める計算が使われます。例えば、最適な値を求めるために（仮に定めた最適な値と）各データとの距離を小さくしていく（これによって「仮に定めた最適な値」は徐々に「実際の最適な値」へ近づいていきます)、という場合に「距離の計算」が必要になります。最適な値と各データとの距離は「**誤差**（error)」と捉えることもできます。

直線上の距離を求める

図 3 のような数直線では、**2** という位置と **6** という位置の距離は引き算で求められます。つまり、6 − 2 という引き算になるので答えは 4 です。一方、2 − 6 という引き算もできます。そうすると答えは − 4 という負の値になります。

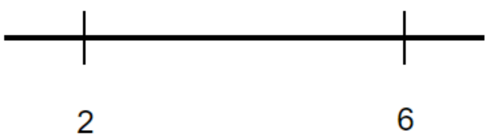


図 3　直線上の距離を求める

距離は正の値なので、どちらからどちらを引いても正になるようにしたいですね。そのような値を**絶対値**（+や−が付かない値）と呼び、| ～ | で囲んで表します。

→目次に戻る

なので、距離は **4** になります。

ただ、絶対値を取り扱うためには元の値が正の場合と負の場合に分けて考える必要があり、計算上の取り扱いが面倒です。そこで、ちょっとしたワザを使って計算を簡単にします。

9 ページ目でお話したように、負の値を **2 乗**すれば正の値になります。その後で、**正の平方根**（＝正の数の「平方根」)、つまり、√（ルート）を求めれば、絶対値になります。なお、「平方根」の意味については実際の計算を見ながら説明します。

まず、「6 − 2」と「2 − 6」という引き算の結果である「**4**」と「**-4**」を **2 乗**してみましょう。

$$
\begin{aligned}
(6-2)^2 &= 4^2 && \cdots \text{() の中を先に計算する} \\
&= 16 \\
(2-6)^2 &= (-4)^2 \\
&= (-4) \times (-4) \\
&= 16 && \cdots \ (\, D \,)
\end{aligned}
$$

- （D）**負の数** × **負の数**は正の数になる（＝負の数を **2 乗**すれば正の数になる）

となり、いずれも **16** という値になります。その後で、**正の平方根 *5** を求めます。

***5　平方根**（square root、ルート）とは、「**2 乗**する前の値」のことです。

例えば「**16**」は、何かを 2 乗した値です。何を **2 乗**したら **16** になるでしょうか？

答えは 4 と − 4 です。4^2 も $(-4)^2$ も 16 ですよね。つまり「**2 乗**する前の値」は **4** と **-4** になります。これらの平方根のうち、正の方を

$$\sqrt{16} = 4$$

のように√（ルート）という記号を使って書きます。負の方は、

$$-\sqrt{16} = -4$$

と表されます。

この平方根も中学校の数学ですが、ここではこれ以上詳しくは触れません。

→目次に戻る

16 の正の平方根は、

$$\sqrt{16} = 4$$

なので、ちゃんと **4** という距離が求められていますね。

ただし、いちいち√を計算せずに、距離の代わりに「距離の **2 乗**」をそのまま使うことがよくあります。文字式や関数を扱うようになると、2 乗のままの方が数式の取り扱いが便利ですし、必要になった時点で√の計算をすればいいので、取りあえずそのままにしておくというわけです（というわけで、平方根については詳しく触れませんでした）。

平面上の距離を求める

平面上の点は、横方向（x 方向）の値と、縦方向（y 方向）の値を組み合わせて、(2,4) のように表されます*6。横方向が **2**、縦方向が **4** の位置となります。

*6　このような (2,4) という値の組み合わせ（＝数値などの複数の要素を一列に並べてひとまとめにしたもの）は**ベクトル**と呼ばれ、やはり機械学習でよく使われます。横方向の値と縦方向の値を個別に取り扱うよりも、ひとまとめにして取り扱った方が便利ですからね。

図 4 には、(2,4) と (6,7) という 2 つの点があります。これらの距離を求めてみましょう。

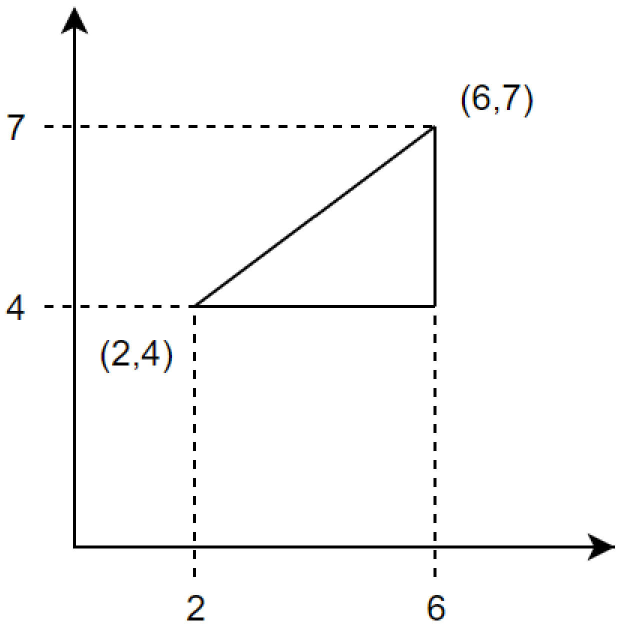


図 4　平面上の距離を求める

(2,4) と (6,7) の距離を求めるということは、図に示された直角三角形の斜辺の長さを求めることにほかなりません。直角三角形の斜辺の長さは、有名な**ピタゴラスの定理**（以下の式）で求められますね。

底辺も高さも引き算で求められます。つまり、以下のように計算できます。

$$\text{斜辺}^2 = \text{底辺}^2 + \text{高さ}^2$$

斜辺の **2 乗**が **25** なので、斜辺の長さは$\sqrt{25} = 5$です。しかし、これについても、いちいち$\sqrt{}$を求めずに、2乗したままの値を後の計算でそのまま使うことがよくあります。

$$\begin{aligned}\text{斜辺}^2 &= (6-2)^2 + (7-4)^2 \\ &= 4^2 + 3^2 \\ &= 16 + 9 \\ &= 25\end{aligned}$$

この計算過程を分かりやすく解説する動画（音声解説付き）も用意しました！

動画 1　「距離（＝斜辺）の計算」を分かりやすく解説 *7

*7　「基準となる点」の位置を **0**（ゼロ）としていますが、この点のことを原点と呼び、一般に **O**（オー）というアルファベットで表します。**O** は **Origin**（オリジン＝原点）の頭文字です。この例では、原点 **O** の位置は、x 座標の値も y 座標の値も **0** なので、正確に言うと (0,0) となります。

→目次に戻る

コラム　距離にはいろいろある

ここで「距離」と呼んでいるものは、**ユークリッド距離**と呼ばれるものです。実は、数学的には、距離は以下のような性質を満たすものとして定義されています。

1. 点 P と点 P の距離は、**0** である
2. 点 P と点 Q の距離は、点 Q と点 P の距離に等しい
3. 「点 P と点 Q」＋「点 Q と点 R」の距離は、点 P と点 R の距離以上である

最後の項目は、要するに点 P から点 R に行くときに点 Q に寄り道してもいい、ということです。「以上」ということなので、いわゆる最短距離でなくてもいいわけです。例えば、京都市内の中心部は道路が碁盤の目のようになっていますが、図 5 のように、その道筋を縦横に通った長さを「距離」としても上の定義に当てはまります。このような距離は**マンハッタン距離**と呼ばれます（日本なら平安京とか平城京の方がしっくりくるかもしれませんね）。

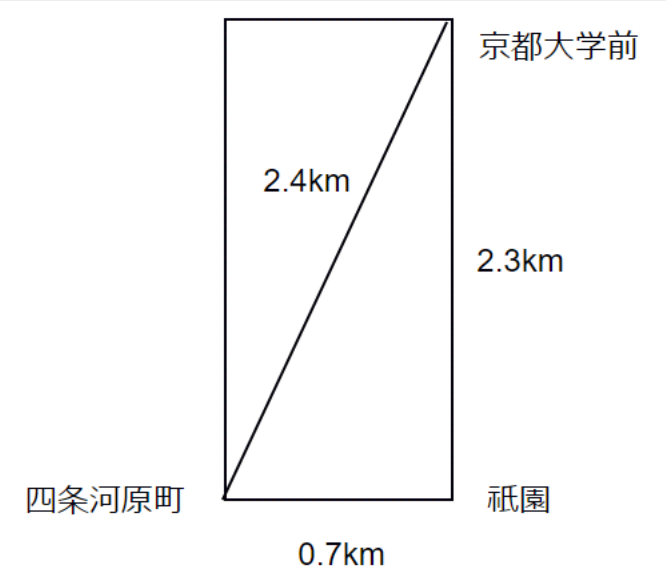


図 5　ユークリッド距離とマンハッタン距離

実際には、道路が少し曲がっているので、この計算通りにはなりませんが、四条河原町から京都大学前までの**マンハッタン距離**は、祇園を経由して歩くのと同じ距離なので、

$$0.7 + 2.3 = 3.0 \ \text{(km)}$$

となります。

一方、**ユークリッド距離**はピタゴラスの定理を使って、

$$\sqrt{0.7^2 + 2.3^2} = 2.4 \ \text{(km)}$$

となります。なお、この連載では「距離」は特に断りのない限り、**ユークリッド距離**を表すものとします。

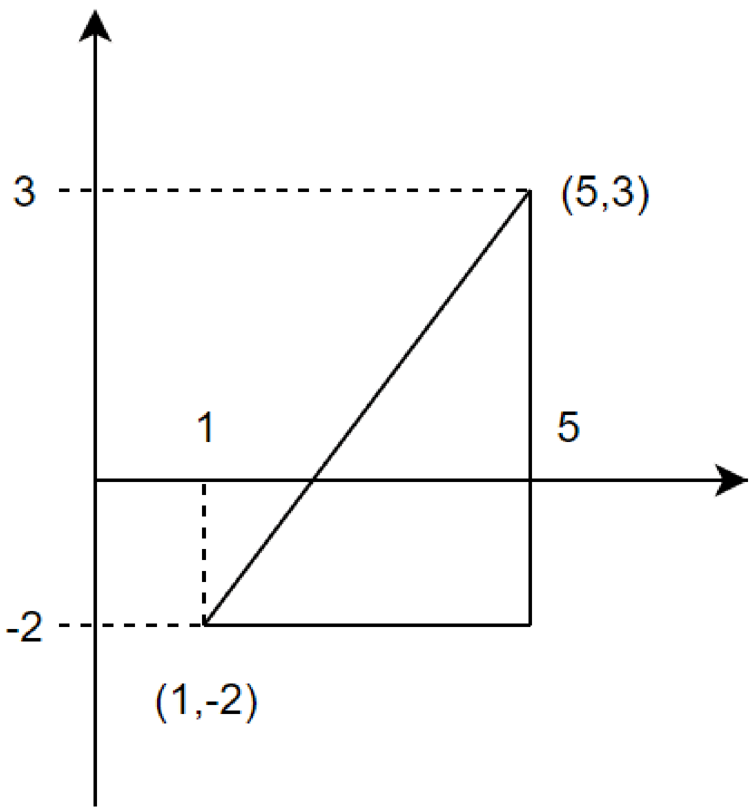
距離が計算できるようになったところ、念のため、練習問題にも取り組んでみましょう。

練習問題

以下の$\boxed{}$に適切な数字や文字を入れましょう。

（1）　数直線上にある4.5という値から2.3という値までの距離の2乗を求める式は、$(2.3-\boxed{})^{\boxed{}}$となる。かっこの中を先に計算するので、$(-\boxed{})^{\boxed{}}$の値を求めるとよい。答えは$\boxed{}$になる。

（2）　以下のような平面に$(1,-2)$という点と$(5,3)$という点がある。これらの点の距離を求めてみよう。図から、直角三角形の底辺の長さの2乗は$(5-\boxed{})^2=\boxed{}^2$であることが分かる。また、高さの2乗は$(3-(\boxed{}))^2=\boxed{}^2$となる。ピタゴラスの定理によって、斜辺^2$=\boxed{}^2+\boxed{}^2$となる。これを計算すると、斜辺^2$=41$。したがって、2点の距離は$\sqrt{41}\simeq 6.4$である（$\simeq$は「ほぼ等しい」という意味です。≒と書くこともあります）。



解答例

「距離を求める」練習問題の解答例は次ページに示します。また、次ページでは本稿の動画と次回の内容について紹介します。

→目次に戻る

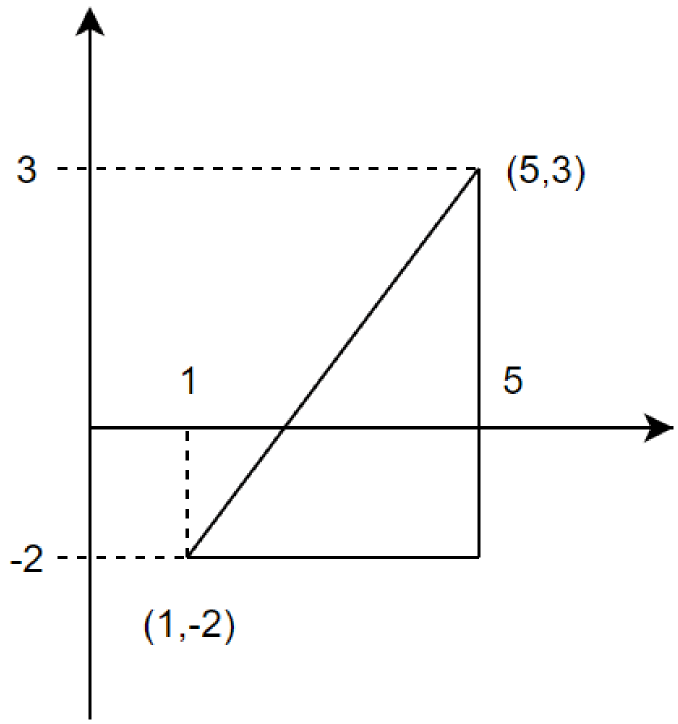
「距離を求める」練習問題に対する解答例です。

解答例

（1）　数直線上にある$\mathbf{4.5}$という値から$\mathbf{2.3}$という値までの距離の2乗を求める式は、$(\mathbf{2.3}-\boxed{\mathbf{4.5}})^{\boxed{2}}$となる。かっこの中を先に計算するので、$(-\boxed{\mathbf{2.2}})^{\boxed{2}}$の値を求めるとよい。答えは$\boxed{\mathbf{4.84}}$になる。

これは特に説明するまでもないでしょう。距離を求めるために 2 乗して正の平方根を求める、などということをやらなくても、人間であれば直感的に大きい値から小さい値を引いて 4.5 − 2.3 = 2.2 という距離が求められます。しかし、決まった手続きで計算を行いたい場合には、いちいちどちらが大きいかという判断をするよりも、単純に引き算をして 2 乗してしまった方が簡単です。

（2）　以下のような平面に$(\mathbf{1},-\mathbf{2})$という点と$(\mathbf{5},\mathbf{3})$という点がある。これらの点の距離を求めてみよう。図から、直角三角形の底辺の長さの2乗は$(\mathbf{5}-\boxed{\mathbf{1}})^2=\boxed{\mathbf{4}}^2$であることが分かる。また、高さの2乗は$(\mathbf{3}-(\boxed{-\mathbf{2}}))^2=\boxed{\mathbf{5}}^2$となる。ピタゴラスの定理によって、斜辺$^2=\boxed{\mathbf{4}}^2+\boxed{\mathbf{5}}^2$となる。これを計算すると、斜辺$^2=\mathbf{41}$。したがって、2点の距離は$\sqrt{\mathbf{41}}\simeq\mathbf{6.4}$である（$\simeq$は「ほぼ等しい」という意味です。≒と書くこともあります）。



平面上の点は (横方向の位置 , 縦方向の位置) で表されます。(1, − 2) であれば、横方向が **1**、縦方向が **-2** の点で、(5,3) は、横方向が **5**、縦方向が **3** の点です。なお、負の数を引くことは、足し算と同じです。3 − (− 2) = 3 + 2 = 5 です（証明は簡単ですが、まだ文字式のおさらいをしていないので今回は省略します）。

次回は……

今回は、基本の基本ということで四則演算をおさらいしました。簡単な計算でも、機械学習の中で使われる計算があることが確認できたでしょうか。

次回と次々回もまだおさらいです。取り上げるのは、次回は文字式で、次々回は総和。

総和とは「全て足す」ということで、機械学習に限らず、あらゆる場面で必要となる計算です。

また、文字式を使いこなせば、さまざまな計算を簡単に表せるので応用範囲も広がります。例えば、点と点だけでなく、ある点とある直線の距離を求める、といったことも簡単に表せます。

[AI・機械学習の数学] 文字式を使いこなせば一気にレベルアップ

数字（1、2、3、……）の式の次は、文字（a、b、x、y、……）を使った式をおさらいしよう。これまでに学んだ知識と平方完成を使って、距離の二乗和の最小値を求める計算も行う。

羽山博，著（2020 年 03 月 30 日）

前回は、「機械学習なんて人間が日常やっていることとそう変わらない」とか、「機械学習で使われる数学といっても結局のところ四則演算とその延長でしかない」といったお話をしました。簡単な例として、距離を求める方法について、具体的な数字を使った計算もやってみました。

今回は具体的な数字ではなく、

- 一般的な数を表したいときに便利な「文字式」（目標【その 1】）
- 目的に合わせて文字式を変形する方法（目標【その 2】）

をおさらいします。文字式を使いこなせると、数式で表現できる世界が一気に広がります。また、方程式や関数についてもさらっと（しれっと?）取りあげ、問題を簡単に解決するための基本を身に付けます。

目標【その 1】：文字式に慣れる

今回の目標の 1 つ目は文字式に慣れることです。具体的な数の代わりに、文字を使って式を表せば物事をより一般的に説明できるようになります。プログラミングでは何らかの値を表すのに count や sum といった単語を変数名として使うのが一般的ですが、数学では 1 つの数は原則として 1 文字で表します *1。どの文字を何のために使うかということについては、厳密なルールはありませんが、以下のような使われ方をします。

- a,b,c など …… 一般的な定数（何らかの決まった値）として使うことが多い
- x,y,z,t など …… 変数（さまざまな値を取る）として使うことが多い

必ずしも、上に記した目的の通りに使われるわけではありませんし、これ以外の使われ方もあります（必要に応じてその都度説明します）が、文字がどのような目的で使われているかを意識していれば、式を見たときにだいたいの意味が分かります。また、違和感があるときには、違った意味で使われているかもしれない、ということも分かります。

ともあれ、まずは文字式に慣れるために、簡単な例を使って見ていきましょう。

***1**　i,j,k などは、後述する**添字**（そえじ：「何番目の値か」ということを表すための文字）として使われることがよくあります。添字は a_i のように、文字の右下に小さく書くのが一般的です。この場合、厳密には 2 つの文字で 1 つの値を表していることになります。

解説：数の代わりに文字を使う

　今回もやはり小学校のレベルからスタートします。最初は、文字式を使わずに、具体的な数字を使って平均値（算術平均）を求めてみましょう。小学校のレベルとはいえ、平均値を求める計算は、機械学習に限らずあらゆるデータ処理に必須です。

　では、例題です。**56** と **78** と **91** という 3 つのデータの平均値を求める式を書いてみてください。計算結果は書かなくても構いません。

　平均値の計算方法は**数値を全て足して、個数で割る**です。答えは、

$$\frac{56+78+91}{3}$$

です。ちなみに計算結果は **75.0** です。カンタンでしたね。

　次は、中学校の数学です。中学に入ると、文字式を学ぶので、数字の代わりに a とか b といった文字を使うようになります。

　では、**56** とか **78** とか **91** といった数字の代わりに、文字を使ってみます。a と b と c の平均を求める式を書いてみましょう。

$$\frac{a + b + c}{3}$$

　はい、これもカンタンでした。a,b,c は、**どういう値か取りあえずは分からないけれど、何らかの決まった値**です。最初に言ったように、何らかの値の代わりに文字を使うというだけのことです ***2**。

　文字式ではさまざまな文字が値の代わりに使われますが、

$$x$$

は英語の筆記体とは違う書き方をすることがよくあります。これはギリシア文字の

$$\chi\ （カイ）$$

と混同するのを避けるためです（図 1）。

→目次に戻る

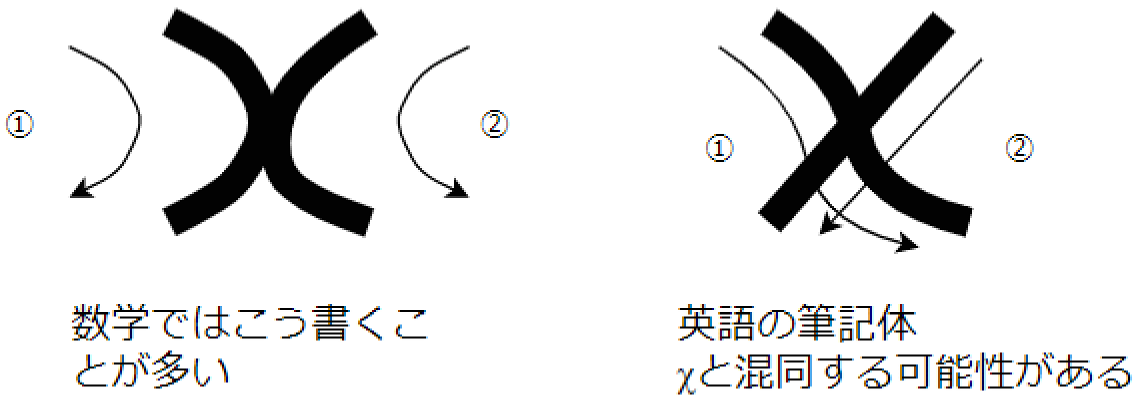


図１　x の書き方

*2　「数」の「代」わり、だから「代数」と呼ばれますね。ちなみに、文字式でも結局は四則演算をやっているだけです。今回の内容も四則演算ができれば全て理解できます。

* 余談　ちょっと余談です。数学に対する苦手意識を生みだす原因の一つに π , σ , θ といったギリシア文字があるかもしれませんね。

　ギリシア文字の小文字は、意味のある定数（＝決まった値）を表すのによく使われます。例えば、π（パイ）は**円周率**を表すのに使われ、σ（シグマ）は**母集団の標準偏差**を表すのに使われます。θ（シータ）は**角度**を表すのに使われます。また、θは、機械学習では変数の「重み」（影響の大きさ）を表すのに使われることもあります。

　πはご存じのように**3.141592**... という特定の値ですが、σやθは、**3.141592**... のような特定の値ではありません。基となるデータや状況によって異なるかもしれませんが、何らかの決まった値ということです。なお、ギリシア文字ではありませんが、**自然対数の底**と呼ばれる値は e と表します。この値は、**2.718281**... という特定の値です。

　ギリシア文字は確かになじみがないかもしれませんが、このような文字の用途や意味合いの違いを抜きにして、「この角度をθとすると」というように、呪文のような文字が数学の教科書にいきなり登場したことで、面食らった人も多いのではないかと思います（実際には、数学の授業で丁寧に説明されているのかもしれませんが、残念なことに私の記憶には残っていません）。

目標【その 2】：文字式の計算に便利な公式を利用する

計算の方法については、前回に見た四則演算の記号（＋ , − ,×,÷）やべき乗の表記がそのまま使えます。ただし、掛け算の記号（×）は省略するか、・で代用できます。

a× b は ab または、a・b と表せる

文字式の計算を簡単にするためにさまざまな公式が使えますが、ここでは必要最低限ということで、分配法則と二乗の展開公式を見ておきましょう。

- **分配法則：** $$a(x+y) = ax + ay$$
- **二乗の展開公式：** $$(a+b)^2 = a^2 + 2ab + b^2$$

この 2 つの公式について解説します。

解説：文字式の計算

上にも示したように、文字式の計算では、掛け算の記号が省略できます。例えば、

$$a \times b = ab$$

です。数字の場合、3 × 4 の × を省略して 34 と書いてしまうと「さんじゅうよん」と間違ってしまいます。しかし、ab なら a × b であることが明らかです。なお、3 × 4 を 3・4 のように省略して書くこともあります。数学者には、書く手間を少しでも省きたいという面倒くさがりな人が多いのかもしれませんね。

分配法則については、前回紹介しました。文字式の四則演算でも分配法則が使えます。

$$a(x+y) = ax + ay \qquad (1)$$

（※分配法則の計算過程を分かりやすく解説する音声解説付き動画も用意しました！　必要に応じてご視聴ください。）

動画１　分配法則の計算

念のため、

$$a(x+y) = ax + ay \qquad (1)$$

に数値を当てはめて確認しておきましょう。以下の式では、

- 3 が、（1）式の a に当たり
- 4 が、（1）式の x に当たり
- 5 が、（1）式の y に当たり

ます。慣れないうちは、どれがどれに対応しているかを「指さし確認」する習慣を付けるといいと思います。が、すぐに慣れるはずです。

（例）

$$\begin{aligned} \underline{3\times(4+5)} &= 3\times 9 \qquad \cdots \boxed{A} \\ &= 27 \\ \underline{3\times 4+3\times 5} &= 12+15 \qquad \cdots \boxed{B} \\ &= 27 \end{aligned}$$

［A］のアンダーラインを引いた部分は（1）式の左辺と同じ形です。一方、［B］のアンダーラインを引いた部分は（1）式の右辺と同じ形になっています *3。確かに結果は同じですね。

*3　等号（＝）や不等号（< , ≦ , ≧ , >）の左側の式を**左辺**、右側の式を**右辺**と呼びます。念のため。なお、「≦」と「≤」は同じ意味です。「≧」と「≥」も同じ意味です。

ただし、たまたまこの値だとうまくいったからかもしれないので、（1）式が成り立つことを示しておくと、以下の通りになります（※解説動画を用意しましたので、必要に応じてご視聴ください）。

動画 2　分配法則の証明

掛け算を足し算に戻して考えるといいですね。

$$
\begin{aligned}
a(x+y) &= \underbrace{(x+y)+(x+y)+\ldots+(x+y)}_{(x+y)\text{が}a\text{個}} && \cdots \boxed{A} \\
&= \underbrace{x+x+\ldots+x}_{x\text{が}a\text{個}} + \underbrace{y+y+\ldots+y}_{y\text{が}a\text{個}} && \cdots \boxed{B} \\
&= ax+ay && \cdots \boxed{C}
\end{aligned}
$$

- [A] …… a を掛けるということは、a 回足すということ
- [B] …… 足し算は順序を変えても同じなので、同類項で（x は x で、y は y で）まとめた
- [C] …… x を a 回足したものは、x 掛ける a なので、xa と表せる。掛け算は順序を変えても同じなので ax と書いてもよい（一般に、アルファベットの順に書きます）

この公式を利用すると、以下のような公式も導き出せます。いちいち証明はしませんが、特に下の公式（二乗の展開公式）は機械学習でよく使うので、ぜひ覚えておきましょう……と言いたいところですが、覚えておく必要はありません。ここに書いてある、ということだけを覚えておけば十分です。

$$(a+b)(x+y) = ax + ay + bx + by$$
$$(a+b)^2 = a^2 + 2ab + b^2$$

（なお、解説動画では、これらの公式が成り立つことも説明しています。）

動画 3　式の展開

これらの公式は、中学校で数学の基礎として学びますね。こういった公式を使って式を展開したり、因数分解したり……と無味乾燥な計算を延々とやらされて、こんなのがいったい何の役に立つんだろう、とうんざりしていた人も多いと思います。念のため確認しておくと、**式の展開**は以下の図 2 で左辺を右辺のように変形すること、因数分解は右辺を左辺のように変形することでした（**因数分解**とは何か、という定義を説明しても抽象的になりすぎて実感が湧かないでしょうから、あとの練習問題で少し具体例を見ます）。

→目次に戻る

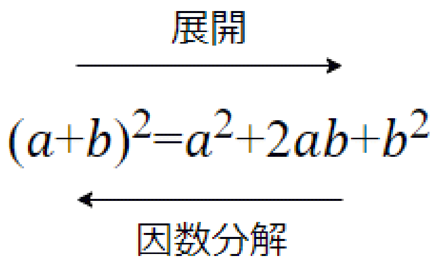


図 2　展開と因数分解

でも、機械学習のための数学ではこのような計算をよく使うのです。回帰式を求めるときに「あれ、なんだこれ二乗の展開公式使ってるじゃん」と気が付いたり、共分散を求めるときに「内積を求めるって、結局、積の公式じゃん」と気が付いたりします。「回帰式」や「共分散」についてはいずれ説明することとして、難しく感じられる手法も、結局はこういう基本的な計算（やはり四則演算です！）の積み重ねなのです。

とはいえ、ある程度の慣れは必要なので、軽く練習問題をこなしておきましょう（※スキップしたい場合は、32ページ目にジャンプしてください）。

練習問題

以下の$\square$に適切な文字や式を入れましょう。

(1)　$ab+ac+ad=\square(b+c+d)$

(2)　$(a+2)(b+3)=a\square+3a+2\square+6$

(3)
$$\begin{aligned}(x+3)^2 &= x^{\square}+\square\cdot 3x+3^{\square}\\ &= x^{\square}+\square x+\square\end{aligned}$$

(4)　$(ax+b)^2=(\square)^2+2\square\square x+b^{\square}$

解答例

「因数分解と式の展開」練習問題の解答例は次ページに示します。32ページ目からは「添え字の利用」と「文字式の応用例（距離の総和の最小値の求め方）」について説明します。

「因数分解と式の展開」練習問題に対する解答例です（解説動画も用意しました）。

動画 4　「因数分解と式の展開」の練習問題

解答例

（1）　$ab+ac+ad=\boxed{a}(b+c+d)$

（2）　$(a+2)(b+3)=a\boxed{b}+3a+2\boxed{b}+6$

（3）
$$\begin{aligned}(x+3)^2 &= x^{\boxed{2}}+\boxed{2}\cdot 3x+3^{\boxed{2}}\\ &= x^{\boxed{2}}+\boxed{6}x+\boxed{9}\end{aligned}$$

（4）　$(ax+b)^2=(\boxed{ax})^2+2\boxed{a}\boxed{b}x+b^{\boxed{2}}$

（1）は**分配法則**を逆に適用すればいいですね。この式の a のように、式を割り切ることのできる項のことを「因数」と呼び、**因数**との積の形に式を変形することを**因数分解**と呼びます。特に、全ての項に掛けられている項を因数として取り出すことを、俗に「くくる」とか「くくりだす」などと言います（この例なら「a でくくる」とか「a をくくりだす」といった感じです）。

→目次に戻る

（2）から（4）は公式を適用するだけでできますね。（3）と（4）にはちょっと注意を払ってください。単に公式を覚えて適用するだけでなく、公式の特徴をつかむようにしましょう。それが、後々ものすごく生きてきます。

（3）で見てみます。真ん中の 6x という項に注目してください。これは、前の x と後ろの 3 の積を 2 倍したものです。

（4）も同じですね。真ん中の 2abx は前の ax と後ろの b の積を 2 倍したものになっています。このことは、前が a^2、後ろが b^2 の形で、真ん中の項が 2ab の形になっていれば、あるいは、なんとかして **2ab の形に変形できれば**因数分解ができる、つまり（3）や（4）の逆の変形ができるということです。

* 余談　私自身、練習問題なんて面倒なだけだし、単調な計算ばかりするのは時間の無駄だよなぁ、とうんざりしていましたが、実は、次のステップに進むための「嗅覚」のようなものを養う意味があるのだな、と教える立場に立ってようやく分かってきました。例えば、上の例であれば、練習問題に取り組んで、その感覚を体得できれば、他の場面でも真ん中の項が 2ab の形式になっていることにすばやく気付くことができたり、その形式に持ち込めばうまくいくんじゃないか、という推理を働かせることができたりするようになります。そういう意味では、モチベーションがゼロの状態で機械的に課題をこなすだけというのは時間の無駄かもしれません。この連載では、練習問題は最小限にとどめますが、できるだけ、先につながるようなものにしようと思っています（うまくいけばいいのですが……）。

はい、これで中学校も卒業です。前回は中学校までで終わりましたが、今回は高校に進みます。

解説：添字の利用、文字式の応用例

具体的には、次のページで「添字の利用」と「文字式の応用例（距離の総和の最小値の求め方）」について解説します。

解説：添字の利用

高校に入ると扱うデータの数や種類も増えてきます。しかし、アルファベットは **26 文字**しかないので、**100 人**のデータの平均を取るといった場合には、a とか b といった英字だけではうまく書けません。そこで、x_1 や x_2 といった書き方で各データを表します。こう書けば、同じ種類の値が幾つかあって、それらを番号で区別したいときに便利です。学生 1、学生 2、学生 3 のように番号で区別するのと同じです。

例えば、x_1 と x_2 と x_3 の平均値を求めるなら、

$$\frac{x_1 + x_2 + x_3}{3}$$

となります。この小さな $_1$ とか $_2$ とか $_3$ のことを「**添字**（そえじ）」と呼びます ***4**。x_1 の読み方は「エックスいち」です。場合によっては「エックスのいち」とか「エックスのいちばん」ということもあるかもしれません。もちろん、小さな字だからといって小さな声で読むなどという必要はありません(そんなことをする人はいないでしょうが)。なお、個々のデータのことを「要素」ということもあります。

***4**　プログラミングでは、小さな文字の入力が面倒なので [] や () で囲んで表します。一般に、**1** からではなく **0** から始めた方が何かと都合がいいので、x の先頭の要素から **3 つ**足す場合、例えば、

$$x[0] + x[1] + x[2]$$

のような書き方になります。また、**添字**と呼ばずに**インデックス**と呼ぶことがよくあります。

添字を使えば、データがたくさんあっても 1 つの文字で複数の値が区別できますね。**100 人**のデータの平均を求めるなら、以下のように表せます。

$$\frac{x_1 + x_2 + \ldots + x_{100}}{100}$$

また、**n 人**のデータの平均を求めるなら、以下のように表せます。

$$\frac{x_1 + x_2 + \ldots + x_n}{n}$$

ここで、添字にも文字を使って表していることに注目してください。x_n なら、x の n 番目という意味になります。**n は一般にデータの個数を表すのに使う**ことが多いので、最後の値（x_n）ではなく「何番目か」の値を表したいときには i や j などを添字に使って表すのが一般的です。例えば、

$$x_i$$

のように書きます。i の値を変えれば好きな要素が指定できるというわけです。

ところで、**n 人**の平均値を求める式の分子の方を見て何か気が付かないでしょうか。式を見ているだけだと気にならないかもしれませんが、この式を書け、と言われるとどうでしょう。いちいち $x_1 + x_2 + \ldots + x_n$ と書くのはとても面倒ですね。**1 番**のデータから **n 番**のデータまでを合計するという決まり切った計算なので、もっと簡単に書きたいものです。

そこでいよいよ、Σ（シグマ）の登場です……と進みたいところなのですが、お話がかなり長くなってきたので、一応、予告だけということにして、詳しい説明は次回のお楽しみということにしましょう。

$$\sum_{i=1}^{n} x_i = x_1 + x_2 + \ldots + x_n$$

Σ、総和、つまり「全部足す」ということを表す記号ですが、機械学習のための計算でよく登場します。式の左辺は、Σが「全て足す」、下の $i = 1$ は「i の値は **1** から始める」、上の n は「i の値は **n** まで」、右の x_i は、「x_i を」ということになります。右辺のように長々と書く必要がなくなりますね。詳細については、次回ということで……。

文字式と文字式の計算の基本はここまでです。ホントに最低限ではありますが、これだけで十分先に進めます。このあとは少しだけ応用的な話題を取り上げ、文字式が問題解決に役立つことを見てみましょう。

各データとの距離の 2 乗の総和が最小になる値とは???

ここからは、文字式の計算に慣れるための練習をかねて、機械学習で使われる計算を少しだけ掘り下げて見ていきます（※これについては、計算過程を分かりやすく解説する音声解説付き動画も用意していますので、「難しい」と感じたら、ぜひ視聴してみてください）。

動画 5　距離の二乗和を求める

動画5　距離の二乗和を求める

問題：距離の二乗和の最小値を求める

最初に **56** と **78** と **91** という 3 つの値を例として使いました。ここでは、これらの値と数直線上での距離の 2 乗の総和（**二乗和**と略します）が最も小さくなる点を見つける計算をしてみましょう。つまり、以下の図 5 のような点を求めるというわけです。

x の上に ¯ が付いていますが、これはまあ飾りのようなものです。今は、あまり気にしないでください（読み方は「エックス・バー」です）。意味はあとで説明しますが、見慣れない記号にちょっと免疫を付けるためにお見せしたといった程度に考えておいてもらってけっこうです。

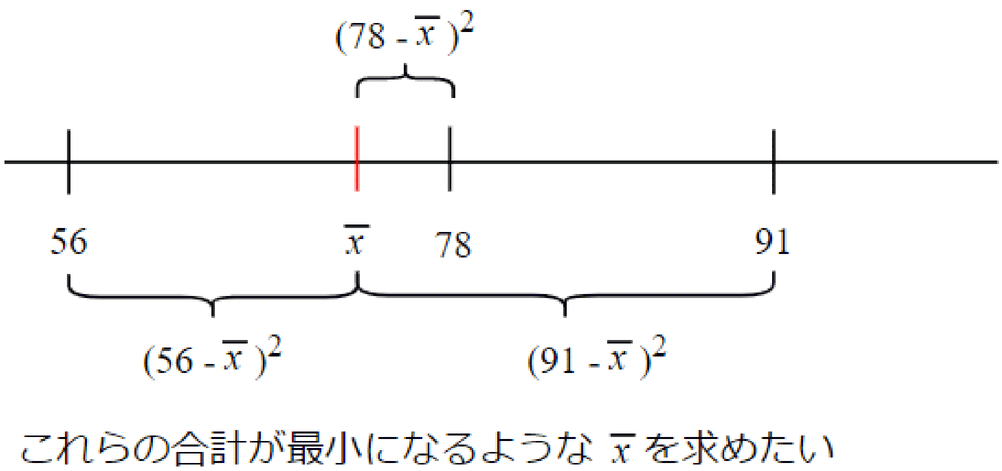


図 5　二乗和を最小にする点

仮にその求める値を

$$\bar{x}$$

としましょう。まず、距離の二乗和を求めてみます。

距離の二乗和を求める

前回見たように、引き算をして 2 乗するのでしたね。つまり、

$$\underline{(56-\bar{x})^2}+\underline{(78-\bar{x})^2}+\underline{(91-\bar{x})^2}$$

この式の下線の部分は、

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

の形になっています。＋と－の違いがありますが、b を引くのではなく、－ b を足すと考えれば公式が適用できます。

$$\begin{aligned}(a-b)^2 &= (a+(-b))^2 \\ &= a^2+2a(-b)+(-b)^2 && \cdots\boxed{A} \\ &= a^2-2ab+b^2 && \cdots\boxed{B}\end{aligned}$$

→目次に戻る

- [A]…… $-b$ を $+(-b)$ と考えて、公式を適用した
- [B]…… $(-b)^2 = b^2$ なので、こうなる。この公式を利用すればよい

では、これを使って、問題となる式を展開してみましょう。

$$
\begin{aligned}
&(56-\bar{x})^2+(78-\bar{x})^2+(91-\bar{x})^2 \\
&= 56^2 - 2\cdot 56\bar{x} + \bar{x}^2 \\
&\quad + 78^2 - 2\cdot 78\bar{x} + \bar{x}^2 \\
&\quad + 91^2 - 2\cdot 91\bar{x} + \bar{x}^2 \\
&= (56^2+78^2+91^2) - (2\cdot 56 + 2\cdot 78 + 2\cdot 91)\bar{x} + 3\bar{x}^2 \qquad \cdots \boxed{A}
\end{aligned}
$$

- [A]…… 同類項をまとめた

56^2 とか $2\cdot 56$ の計算は、何も暗算や筆算でやらなくても電卓を使えばいいですね。計算してまとめると、

$$17501 - 450\bar{x} + 3\bar{x}^2$$

となります（※ここまでの計算が難しいと感じたら、前掲の解説動画【距離の二乗和の最小値を求める】をご視聴ください）。

この式はどこかで見たことがある式ですね。そう、

$$y = ax^2 + bx + c$$

という二次関数の右辺です。項の順序を、

$$\bar{x}$$

の次数が高い方から（降べきの順に）整理して、$y =$ を付けると、

$$y = 3\bar{x}^2 - 450\bar{x} + 17501$$

となります。そこで、このグラフを描いてみましょう。

$$\bar{x}^2$$

の係数が正であれば、下に凸なグラフになりますね（図 6）。かなり縦長になるので、横幅を拡大して見やすくしてあります ***5**。

***5**　このグラフは macOS に付属している Grapher というアプリで描いたものです。数式を使ったグラフの描画には WolframAlpha や GeoGebra などのサイトやそれらのアプリも便利です。

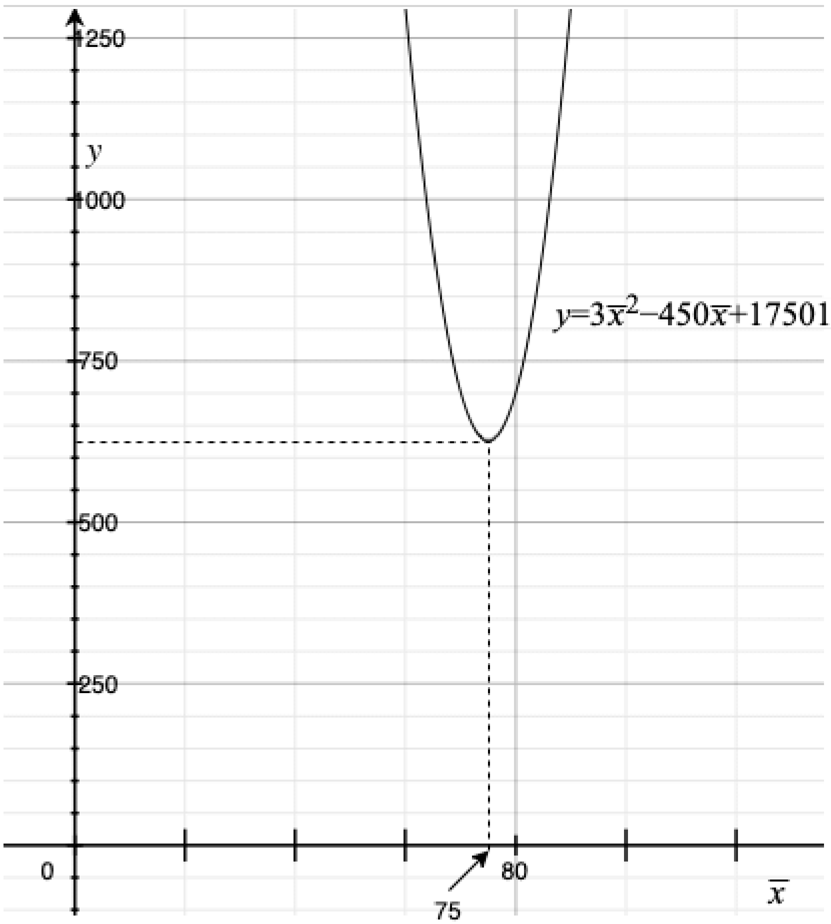


図 6　二次関数のグラフ

縦軸が y つまり、

$$3\bar{x}^2 - 450\bar{x} + 17501$$

の値です。グラフを見ると、この式の値が最小になる点があることが分かります。そのときの

$$\bar{x}$$

の値はいくらかというと **75** です。**75** というと……なんと、**56** と **78** と **91** の平均値ではありませんか！　私たちがグラフから見つけた値は、各データとの距離の二乗和が最小になる値でしたよね。つまり、**平均値とは各データとの距離の二乗和が最小になる値**だったのです。

最初、

$$\bar{x}$$

という表記のことをむにゃむにゃとごまかしていましたが、データ（x_i）から求めた平均値を、一般に

$$\bar{x}$$

と表します (最初にそれを言ってしまうとネタバレになるので、さらっと流したというわけです)。今後は、

$$\bar{x}$$

と書いてあれば、ああ、**平均値**なんだな、ということがひと目で分かるようになりましたね。

ところで、この **75** という値をグラフから読み取るのではなく、計算できちんと求めたいですよね。そんなもの、平均値なんだから全部足して個数で割ればいいじゃないか、と言われるかもしれませんが、それは結果としてそうなるだけのことです。そうではなく、各データとの距離の総和が最小になる値を計算で求めて、それが確かに平均値になっていることを確かめようというわけです。

コラム　μ と $\bar{x}$

平均値を表すのに、 μ （ミュー）という文字や $\bar{x}$ という文字が使われますが、これらの文字には違いがあります。 μ は母集団の平均値を表します。一方の $\bar{x}$ は母集団から取り出した一部のデータ（**サンプル**や**標本**と呼ばれます）を基に、母集団の平均値を推定した値です。「推定値である」「誤差を含む」といった意味合いを表すためには、 $\bar{x}$ などのように、文字の上に ¯ や ^ を付けることが一般的です（ ^ は、**キャレット**という文字ですが、数学では**ハット**と読みます）。

最小値を求めるために式を変形する（平方完成）

（※ここから本章の最後までに説明する「平方完成」の計算過程は動画でも分かりやすく説明しています。文章を読んで「難しい」と感じたら、ぜひ視してみてください。）

動画 6　二次式の最小値を求める（平方完成）

動画6　二次式の最小値を求める（平方完成）

高校の数学を少し覚えていれば、下に凸な二次関数

$$f(\bar{x})$$

の値を最小とする

$$\bar{x}$$

を求めるなら「微分して **0** と置く、だろ」と思われる方もいると思います。まさにその通りですし、その方が計算ははるかに簡単なのですが、文字式にもう少し慣れるために、ここはあえて微分を使わずにやってみましょう。実は中学校までの数学でできます ***6**。

→目次に戻る

***6**　ちなみに、

$$\bar{x}$$

で微分して **0** と置くと、

$$6\bar{x} - 450 = 0$$

となります。この方程式を解くと、

$$\bar{x} = 75$$

です。簡単すぎます！　瞬殺です！　……が、ぐっと我慢して微分はまたのお楽しみ、ということで。

　私たちがやりたいのは、

$$3\bar{x}^2 - 450\bar{x} + 17501 \qquad (1)$$

という式の値を最小にする

$$\bar{x}$$

の値を見つけるということでした。まず、この式の形に注目しましょう。この式は、

$$ax^2 + bx + c$$

と同じ形になっていますね。a、b、c を指さし確認してみましょう！

- （1）式の　3　が、a に当たり
- （1）式の　-450　が、b に当たり
- （1）式の　17502　が、c に当たります

　さて、最小値を求めるために、この式を以下のような形に変形してみたいと思います（この変形のことを**平方完成**と呼びます）。

$$a(x+p)^2 + q$$

　このとき、$(x+p)^2$ は、2 乗されているので必ず **0 以上**になることに注目してください。q の値は決まっていますから、a の値が**正**であれば（グラフが下に凸であれば）、全体が最小になるのは $(x+p)^2$ が **0** のとき、つまり $x = (-p)$ のときです（図 7）。

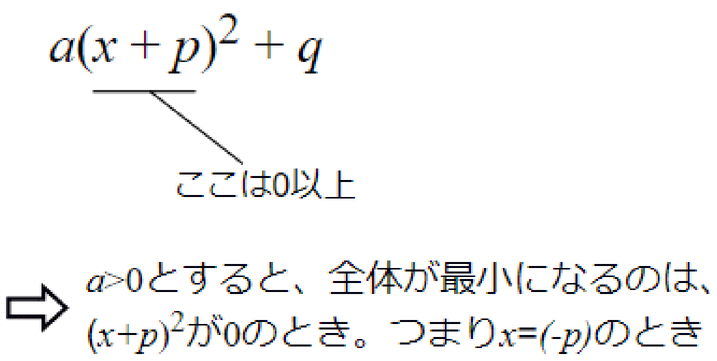


図 7　平方完成によって最小値を求める

数学の教科書や授業では、$ax^2 + bx + c$ を割ったり掛けたり足したり引いたりして $a(x + p)^2 + q$ の形に変形するのですが、ここでは思いっきり楽をしましょう。登山道で言えば、上級者向けコースを登るのではなく、初心者向けコースを降りてくる感じです。その分、ちょっと説明は長くなりますが、なだらかな道です（上級者向けの簡潔な変形も動画には含めてあります。興味がある方は後掲の動画【二次式から平方完成の式に変形する別の方法（上級者向け)】を視聴してください)。

というわけで、平方完成の逆のコースをたどり、$a(x + p)^2 + q$ を $ax^2 + bx + c$ の形にすればどうなるかを見ていくことにします。まずは、かっこの展開からです。

$$
\begin{aligned}
&a\underline{(x+p)^2}+q \\
&= a(\underline{x^2+2px+p^2})+q \cdots \boxed{A} \\
&= \underline{a}x^2+2\underline{a}px+\underline{a}p^2+q \cdots \boxed{B}
\end{aligned}
$$

- [A] …… $(a + b)^2 = a^2 + 2ab + b^2$ という公式を使った（二次式の展開)
- [B] …… $a(x + y) = ax + ay$ という公式を使った（分配法則）

これと、$ax^2 + bx + c$ を比較してみます。以下の図 8 のように、a はそのまま a に対応していますが、$b = 2ap$、$c = ap^2 + q$ です。

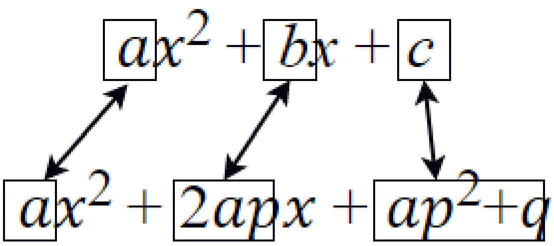


図 8　係数を比較する

ここで、以下の連立方程式が得られました。

$$
\begin{aligned}
b &= 2ap && (1) \\
c &= ap^2 + q && (2)
\end{aligned}
$$

平方完成した式の p について解く（連立方程式）

　おっと、いきなり連立方程式ですか！　とビビってしまう方もおられると思いますが、一歩ずつ丁寧に見ていけば怖れることはありません。まず、（1）の両辺を 2a で割っても左辺と右辺が等しいことに変わりはありませんね。従って、

$$\frac{b}{2a} = p$$

となります。つまり、

$$p = \frac{b}{2a} \qquad\qquad (3)$$

であることが分かりました（a>0 なので、a は **0** ではありません。従って割り算ができます）。実は、x の最小値を求めるだけなら、これですでに答が求められています。図 7 で示したように、x の値が－ p のとき、つまり、x の値が、

$$-\frac{b}{2a}$$

のときです。が、せっかくなので、これ以降の計算もついでにやっておきます（※計算はもううんざり、という人は「ゴール：最小値を求める」まで読み飛ばしてもらっても構いません）。

平方完成した式の q について解く（連立方程式）

　次に、（3）式を（2）式の p に代入しましょう。

$$p\text{と}\frac{b}{2a}\text{が等しい}$$

のだから、p の部分を、

$$\frac{b}{2a}$$

に書き替えても式は成り立ちますね。

$$c = a(\frac{b}{2a})^2 + q$$

となります。従って、

$$
\begin{aligned}
c - a(\frac{b}{2a})^2 &= q \cdots \boxed{A} \\
q &= c - a(\frac{b}{2a})^2 \cdots \boxed{B} \\
q &= c - a(\frac{b^2}{4a^2}) \cdots \boxed{C} \\
q &= c - \frac{b^2}{4a} \cdots \boxed{D}
\end{aligned}
$$

- $\boxed{A}$ …… $a(\frac{b}{2a})^2$を左辺に移項した（両辺から$a(\frac{b}{2a})^2$を引いたのと同じ）
- $\boxed{B}$ …… $q =$の形にしたいので、左辺と右辺を入れ替えた
- $\boxed{C}$ …… 2乗を展開した。$(\frac{n}{m})^2 = \frac{n^2}{m^2}$なので
- $\boxed{D}$ …… aで約分した。$a \cdot \frac{1}{a^2} = \frac{1}{a}$なので

これで、p と q が求められました。

$$a(x+p)^2 + q$$

に当てはめてみましょう。

$$a(x+p)^2 + q = a(\underline{x + \frac{b}{2a}})^2 + c - \frac{b^2}{4a}$$

ゴール：最小値を求める

ところで、私たちが関心を持っているのは、この式を最小にする x の値です。それは、アンダーラインを引いた部分から、

$$-\frac{b}{2a}$$

であることが分かります。そこで、

$$\text{最初の式}\quad 3x^2 - 450x + 17501$$

に戻ると、

$$
\begin{aligned}
a &= 3 \\
b &= -450
\end{aligned}
$$

なので、

$$-\frac{b}{2a} = -\frac{-450}{2 \cdot 3} = \frac{450}{6} = 75$$

　できました！　式の値、つまり、各データとの距離の二乗和が最小になる

$$\bar{x}$$

の値が計算で **75** と求められました。ちゃんと平均値になっていますね。

　平均値は、よく「極端に大きな値や極端に小さな値があると、それらの値に引きずられてしまう」と言われます。全部足して個数で割るという計算方法からは分かりませんが、そもそも元の値を二乗しているからなのです（例えば、大きな値を二乗すればより大きくなります）。

　さて、今回のお話はこのあたりでおしまいですが、次回（48 ページ目から）はΣを取りあげるので、これまでやってきた計算を一般化する（ここでは具体的な数値を使って確認しましたが、どんな値でも成り立つことを示す）、といったことにも挑戦してみたいと思います。

* 余談　なぜ、各データとの距離の総和を最小にする値ではなく、距離の **2 乗**の総和を最小にする値を求めるのか、と思われる方がいるかもしれませんね。前回は「絶対値は扱いが面倒だから」といった感じでさらっと流しましたが、距離を最小にする値が 1 つに決まらないことがあるというのが大きな理由です。具体例で見てみましょう。

例えば、2,3,5,7 というデータがあったとします。距離を最小にする値は、$|2-x|+|3-x|+|5-x|+|7-x|$ の最小値です。距離の 2 乗の総和を最小にする値は $(2-x)^2+(3-x)^2+(5-x)^2+(7-x)^2$ の最小値です。絶対値の計算はけっこう面倒なので、Grapher を使ってグラフを描いてみます。

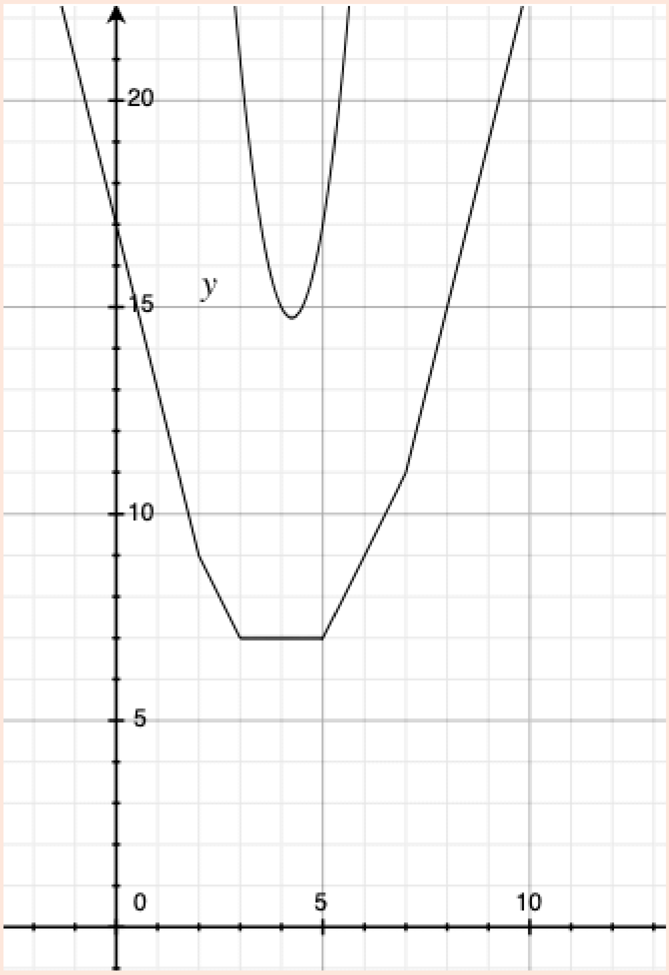


図 10　距離のグラフ

このグラフのカクカクした方が絶対値の総和のグラフです。このことからも分かるように、距離の総和の最小値は 1 つだけではありません。計算してみましょう。

→目次に戻る

$$
\begin{aligned}
x = 3\text{のとき、} & |2-x|+|3-x|+|5-x|+|7-x| \\
&= |2-3|+|3-3|+|5-3|+|7-3| \\
&= 1+0+2+4 \\
&= 7 \\
x = 4\text{のとき、} & |2-x|+|3-x|+|5-x|+|7-x| \\
&= |2-4|+|3-4|+|5-4|+|7-4| \\
&= 2+1+1+3 \\
&= 7
\end{aligned}
$$

となります。同様に、x = 5 のときも、最小値は **7** となります（3 < x ≦ 5 を満たす x であればいいので、無限にあります）。一方、2 乗を最小にする値は唯一に決まります。

蛇足ですが、三角形の各頂点からの距離の和を最小にする点を**フェルマー点**と呼びます（フェルマーさんはよく登場しますね）。また、三角形の各頂点からの距離の 2 乗の和を最小にする点は**重心**と呼ばれます。重心は三角形のように点が 3 つの場合だけでなく、点が幾つあっても 1 つに決まります。ちなみに、重心は **k-means 法**と呼ばれるクラスタリング（=グループ分け）のアルゴリズムなど、さまざまな手法で使われます。

（※次の動画は、前述した「平方完成の形への簡潔な変形方法（上級者向け）」を説明しています。興味がある方はご視聴ください。）

動画 7　二次式から平方完成の式に変形する別の方法（上級者向け）

次回は……

多くのデータを合計するというのは平均を求める場合だけでなく、さまざまな場面で登場する計算です。次回は、全てを合計する（＝総和を求める）のに使うΣという記号を取り上げます。Σを利用した計算ができるようになると、数式をより簡潔に、また正確に表せるようになり、複雑な問題をスッキリと解きほぐすのにも役立ちます。

→目次に戻る

[AI・機械学習の数学] 総和を表すΣは機械学習に必須の記号

「Σ」を理解して総和をマスターしよう。応用で、Σの公式を使って平均を求めてみる（最小二乗法につながる基礎知識）。さらに、平均を使って重心を求める計算も行う（クラスタリング「k-means 法」につながる基礎知識）。

羽山博，著（2020 年 04 月 16 日）

前回までで、中学までの数学のうち、最低限の知識についておさらいしました。演算のルールと文字式の取り扱いだけでしたが、それでも機械学習で使われる数学を垣間見ることができたと思います。添字は中学の数学には登場しなかったかもしれませんが、文字式の延長で理解できますね。

今回は、多くの数を足すときに便利な「**総和**」の記号Σの使い方を見ていきます。Σが使いこなせるようになると、複雑な計算も簡単に表せるようになります。

目標【その 1】：総和を求める

今回の最初の目標は、「総和」を求めるために使うΣの意味と書き方を理解することです。例えば、以下のような式に見覚えがあるでしょう（高校 2 年ぐらいの数学です）。

$$\sum_{i=1}^{n} x_i$$

Σは「総和を求める」つまり「全部足す」ということを表す記号で「シグマ」と読みます。機械学習のための計算でよく登場するので、書き方と意味を丁寧に見ていきましょう。

解説：Σの意味と書き方をマスターする

前回、平均値を求める計算のお話をしました。平均値を求めるには「全ての値を足して、個数で割る」という計算をしますが、その「全ての値を足す」という部分を前回までの知識で書くと、以下のようになります。

$$x_1 + x_2 + \ldots + x_n \tag{1}$$

上の式では、データの個数は **n 個**ということになります。合計を求めるだけならこの書き方だけで十分な気がしますが、+や ... という記号を幾つも書くのは面倒です。そこで、「こういうお約束に従って、全部足すんだよ」ということを簡単に表すための記号を使うことにしましょう。

それがΣです。この記号はギリシア文字のアルファベットの**シグマ**（大文字）で、英語で言えば、**合計（SUM）**の頭文字である **S** に当たる文字です。Σの筆順は以下の通りです。

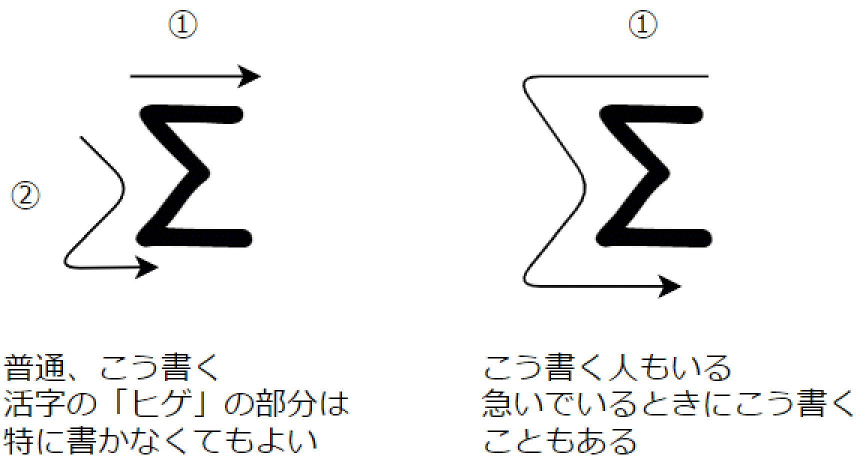


図１　Σの筆順

（1）式をもう一度見てください。x_i という書き方を使って各データを表すのであれば、i の値を **1,2,...n と**変えて全て足すことにすればいいでしょう。このとき「全て足す」「i の値は **1** から始める」「i の値は **n** まで」「x_i を」といちいち言うのが面倒なので、Σという記号を使って図２のように表します。

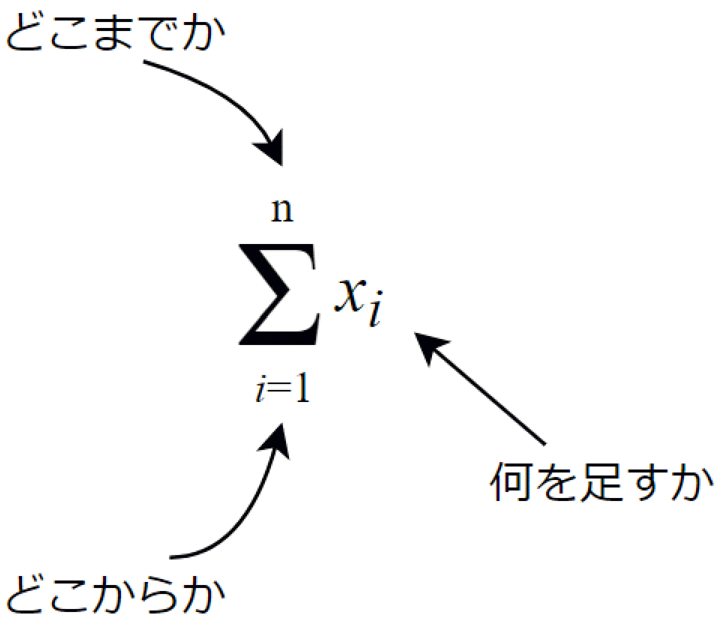


図２　Σの書き方と意味

書き方を確認しておきましょう。Σの下には「**どこ**からか」を、上に「**どこ**までか」を書き、右に「何を足すのか」を書きます。上の式の意味を改めて日本語で書き下すと「i が **1** から」「**n** まで」「x_i を足す」ということになります。人によって読み方が違うかもしれませんが、この式は「シグマ、アイ、イコール、いちから、エヌまで、エックスアイ（をたす）」のように読みます。() 内に書いたように、「シグマ」と意味は重なりますが、最後に「をたす」と付けることもあります。従って、

$$\sum_{i=1}^{n} x_i = x_1 + x_2 + \ldots + x_n$$

です。Σという記号はちょっとイカツイ面構えをしていますが、左辺を見ると、式を書く手間もスペースも省けることが分かります。しかも、足し算で表していたときの、途中の「...」が要らなくなったので、あいまいさがなくなり、どのような計算をするのかが正確に表せるようになりました。

なお、文章の中で数式を書く場合は、**下限**と**上限**を右下と右上に書いて、

$$\textstyle\sum_{i=1}^{n} x_i$$

のように表すこともあります。せっかくなので、ちょっと読んでみてください。

$\displaystyle\sum_{i=1}^{n} x_i$ はシグマ、アイ、イコール、いちから、エヌまで、エックスアイ（をたす）と読みます

また、データが幾つかあって、それらを全部足すといった場合など、「**どこ**からか」「**どこ**までか」が文脈から明らかな場合は、**下限**と**上限**を省略して、以下のように書くことがあります。

$$\sum x_i$$

では、ちょっとしたクイズです。x_1 から x_n までの平均値を求める式を、Σを使って書いてみてください。ここでは、「**どこ**から」「**どこ**まで」を省略せずに書いてみましょう。

$$\frac{\boxed{}}{n}$$

答えは以下の通り。真ん中にΣを書いて、下に添字の下限、上に上限、右に足し算をする値（文字）を書くだけです。既に見た通りでしたね。

$$\frac{\displaystyle\sum_{i=1}^{n} x_i}{n}$$

Σとは要するに**足し算**です。**何番**から**何番**まで何を足す、と言う代わりに記号を使って簡潔に表しているだけのことです。別に謎の呪文を使ってイジワルをしようというわけではなく、カンタンに表したいだけのことなのです。

ここで、Σに関する言い回しで重要な補足をしておきます。これまで、書き方の説明として「Σの右に x_i を書く」といった言い方をしてきましたが、意味としては、この x_i はΣを使って合計する範囲の中にあるものなので、今後は「Σ**の中の** x_i」のような表し方にします。以下の例で確認しておいてください。

$$\sum(a_i + b) \quad \cdots\cdots \boxed{A}$$
$$\sum a_i + b \quad \cdots\cdots \boxed{B}$$

[A] …… Σの中の式は $(a_i + b)$

[B] …… Σの中の式は a_i だけ。b はΣの外

* 余談　ちょっと余談です。上限や下限には無限大（∞）や無限小（-∞）を書くこともできます。例えば、

$$\sum_{n=0}^{\infty} \frac{1}{n!}$$

のような書き方です。ちなみに、この値は前の余談でお話しした自然対数の底 e = 2.718281... になります。なお、n! は「**n** の階乗（かいじょう）」と読み、**n** から **1** ずつ減らしていき **1** になるまでの値を全て掛けたもの、つまり、**n×(n-1)×...2×1** を表します。なお、0 != 1 なので、上の式を展開すると以下のようになります。

$$\begin{aligned}&\sum_{n=0}^{\infty} \frac{1}{n!}\\ &= \frac{1}{0!} + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \ldots\\ &= \frac{1}{1} + \frac{1}{1} + \frac{1}{2 \cdot 1} + \frac{1}{3 \cdot 2 \cdot 1} + \ldots\\ &= 2.718281\ldots\end{aligned}$$

今回はもう少しお話がありますが、Σの書き方と意味が一通り分かったので、いったん練習問題に取り組んでみましょう。

練習問題

（1）　以下の$\square$に適切な文字や式を入れましょう。

$$a_1 + a_2 + \ldots + a_n = \sum_{\square}^{\square} \square$$

（2）　$ax_1 + ax_2 + \ldots + ax_n$を$\sum$を使って表してみましょう。

（3）　上の式で、aが3、x_iの値が順に$\{5, 7, 2\}$のとき、合計の値を求めてみましょう。

（4）　$\sum_{i=1}^{n}(ax_i + b)$を足し算で表してみましょう。

解答例

「総和」練習問題の解答例は次ページに示します。53 ページ目からは「Σの公式を知る」について説明します。

前のページで示した「総和」練習問題に対する解答例です。

解答例

$$(1)\quad a_1 + a_2 + \ldots + a_n = \sum_{\boxed{i=1}}^{\boxed{n}} \boxed{a_i}$$

$$(2)\quad \sum_{i=1}^{n} a x_i$$

$$(3)\quad \begin{aligned} & 3 \times 5 + 3 \times 7 + 3 \times 2 \\ & = 15 + 21 + 6 \\ & = 42 \end{aligned}$$

$$(4)\quad a(x_1 + x_2 + \ldots + x_n) + nb$$

（1）と（2）はΣの書き方を確認するための問題です。なお、（2）については、

$$a \sum_{i=1}^{n} x_i$$

に変形できることを後で学びます。

（3）はΣで計算がどのように行われるかの確認です。順番に足していくだけなので、計算ミスをしなければ大丈夫でしょう。

（4）は、素直に足し算にすると、以下のようになります。

$$(ax_1 + b) + (ax_2 + b) + \ldots + (ax_n + b)$$

足し算はカッコを外しても、順序を変えても結果は同じなので、同じ文字（同類項）はできるだけまとめましょう。つまり、以下のようになります。

$$\begin{aligned} & (ax_1 + b) + (ax_2 + b) + \ldots + (ax_n + b) \\ & = ax_1 + b + ax_2 + b + \ldots + ax_n + b \quad \cdots \qquad \cdots \boxed{A} \\ & = ax_1 + ax_2 + \ldots + ax_n + \underbrace{b + b + \ldots + b}_{n\text{個}} \quad \cdots \qquad \cdots \boxed{B} \end{aligned}$$

[A] …… カッコを外した

[B] …… 同類項がまとまるように順序を変えた

ここで、b をよく見てみると、同じ値を単に **n 回**足しているだけであることが分かります。ということは、掛け算にできますね。

$$上の式 = ax_1 + ax_2 + \ldots + ax_n + nb$$

さらに、a に注目してみると、以下のように書けます（a でくくりました）。

$$上の式 = a(x_1 + x_2 + \ldots + x_n) + nb$$

これで式がだいぶ見やすくなりました。意味もよく分かりますね。式を見やすくしておくことも理解を助けます。ちょっとひと息ついてから次に進みましょう。

目標【その 2】：Σの公式を知る

次は、Σを応用するために使える公式を最小限のものにしぼって見ていきます。

$$\sum (x_i + y_i) = \sum x_i + \sum y_i$$

$$\sum ax_i = a \sum x_i$$

解説：いったん足し算に戻してみれば当たり前のこと！

上の公式は、Σを素直に足し算で表せば簡単に導き出せます（厳密な証明ではありませんが）。え、簡単じゃないよぉ、と思われる方はぜひ、動画 1 を見て式の変形方法を確認しておいてください。

動画 1　Σの公式が成り立つことの確認

→目次に戻る

$$
\begin{aligned}
&\sum (x_i + y_i) \\
&= (x_1 + y_1) + (x_2 + y_2) + \ldots + (x_n + y_n) && \cdots \boxed{A} \\
&= \underbrace{x_1 + x_2 + \ldots + x_n}_{n\text{個}} + \underbrace{y_1 + y_2 + \ldots + y_n}_{n\text{個}} && \cdots \boxed{B} \\
&= \sum x_i + \sum y_i && \cdots \boxed{C}
\end{aligned}
$$

[A] …… 足し算で表してみた

[B] …… 同類項をまとめた

[C] …… Σを使って書いた

$$
\begin{aligned}
&\sum a x_i \\
&= a x_1 + a x_2 + \ldots + a x_n && \cdots \boxed{A} \\
&= a(x_1 + x_2 + \ldots + x_n) && \cdots \boxed{B} \\
&= a \sum x_i && \cdots \boxed{C}
\end{aligned}
$$

[A] …… 足し算で表してみた

[B] …… 分配法則を逆に使った（a でくくった。因数分解した、とも言いましたね）

[C] …… () の中をΣを使って書いた

ここまでの計算は、前掲の動画 1 でも説明しています。

ここまでは比較的素直に進んできましたが、たまには、ひっかけ問題を出してみましょう。

$$\sum_{i=1}^{n} \bar{x}$$

これを足し算で表し、それをさらに掛け算で表してみてみましょう。x の上に ‾ が付いていますが、ここでは特に意味はありません。ちょっとした飾りです *1。

*1　前回も見たように、

$$\bar{x}$$

は、x の平均値という意味でよく使われます。‾ は棒線（バー）なので、

$$\bar{x}$$

は「エックス、バー」と読みます。

→目次に戻る

あれ、おや、これはどうすればいいんだろう、と悩みモードに入ってしまった人もいるかもしれませんが、

$$\bar{x}\text{に添字が付いていないこと}$$

に注意してください。要するに、

$$\text{同じ}\bar{x}\text{という値を}n\text{回足しているだけ}$$

のことです。

$$\begin{aligned}\sum_{i=1}^{n}\bar{x} &= \underbrace{\bar{x}+\bar{x}+\bar{x}+\ldots+\bar{x}}_{n\text{個}} \\ &= n\bar{x}\end{aligned}$$

となります。Σの中の式が x_i なら、i の値によって、異なる値を足すことになります。しかし、添字が付いていないので、

$$\text{同じ}\bar{x}\text{を}n\text{回足す}$$

ということになります。だから、

$$n\bar{x}$$

と等しいのですね。

Σの意味と書き方、公式についての基本はここまでです。ホントに最低限ではありますが、これだけで十分先に進めます。この後は少し応用的な話題に入ります。

応用：Σを使って平均値を求める式を導き出す

前回、具体的な数値を使って、各データとの距離の二乗和が最小になるときの

$$\bar{x}$$

の値こそが平均値である、というお話をしました。また平均値の話か、と思われた方もおられるかもしれませんが、目標【その 2】の公式を使ったΣの計算に慣れるのに恰好（かっこう）のネタなので、少しお付き合いください。また、新たな気づきも得られるはずです。

さて、今回は、数値の代わりに文字を使い、総和をΣで表してみます。そうすれば、平均値を求める式を一般的に表すことができそうですね。そこで、各データを $\{x_1,x_2, \ldots , x_n\}$ とし、各データとの距離の二乗和が最小になるときの値を

$$\bar{x}$$

として、

$$\bar{x}\text{の値（つまり平均値）}$$

を求める式を導き出してみましょう。

図で表すと、以下の

$$(x_i - \bar{x})の2乗の総和$$

が

$$最小になるときの\bar{x}の値$$

を求めるということです（図 3）。なお、2 乗の総和のことを「二乗和（じじょうわ）」と略して呼びます。ここからは動画でも分かりやすく解説しているので、難しく感じた人は次の動画 2 を視聴してみてください。

動画 2　距離の二乗和を最小にする「$\bar{x}$」を求める（総和Σ活用編）

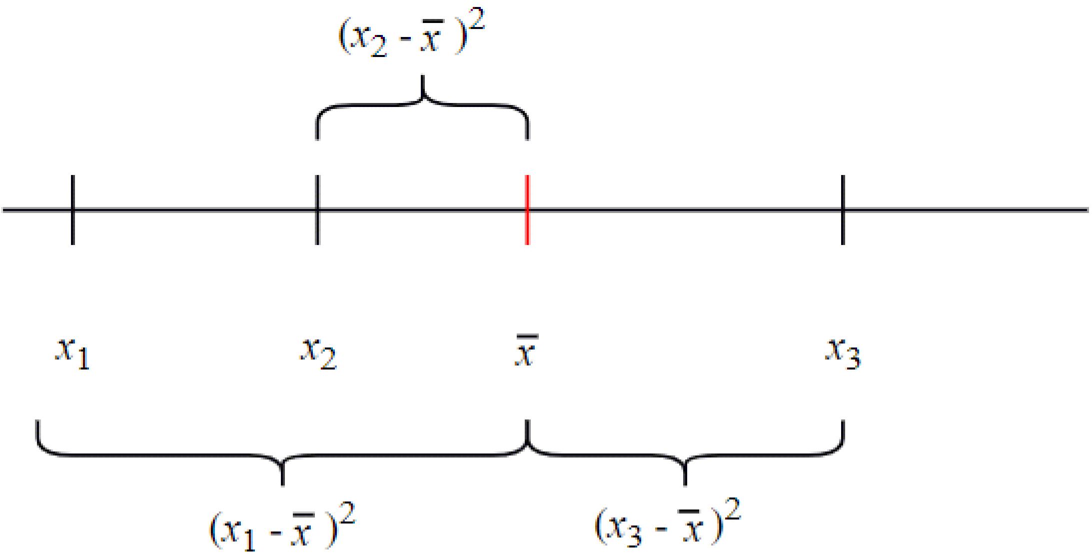


図 3　各データからの距離の二乗和が最も小さくなる点を求める

では、前回と同じようにやってみましょう。まず、

$$各データから\bar{x}を引いて、2乗$$

します（＝距離の 2 乗を求める)。そして、それらを全て足します（＝総和を求める)。

$$(x_1-\bar{x})^2+(x_2-\bar{x})^2+\ldots+(x_n-\bar{x})^2$$

です。ここで、Σを使ってこの式を書いてみてください。またまた穴埋めです。「**どこ**から」「**どこ**まで」は自明なので省略しちゃいましょう。

$$\sum(\boxed{\phantom{x_i-\bar{x}}})^2$$

答えは以下の通りです。

$$\sum(\boxed{x_i-\bar{x}})^2$$

要するに、Σの中の式では、$x_1,x_2,\ldots,x_n$ を一般的に x_i と表せばいい、ということですね。

続いて、Σの中の式を展開していきましょう。展開？　え？　何言ってるか分からない、という人もいるかもしれませんが、これまでに学んだ知識を思い出してください。

$$(x_i-\bar{x})^2$$

の部分を見てみると、$(a-b)^2=a^2-2ab+b^2$ の公式の左辺と同じ形になっていることに気がつきます。

$$aの代わりにx_iを使い、$$
$$bの代わりに\bar{x}を使っているだけ$$

です。従って、以下のように展開できます。

$$\sum\underline{(x_i-\bar{x})^2}=\sum\underline{(x_i^2-2x_i\bar{x}+\bar{x}^2)}$$

次に、目標【その 2】で見たΣ $(x_i+y_i)=\Sigma x_i+\Sigma y_i$ という公式を使ってみます。

$$\begin{aligned}\sum(x_i-\bar{x})^2&=\sum(x_i^2-2x_i\bar{x}+\bar{x}^2)\\&=\sum{x_i}^2\underbrace{-2(\sum x_i)\bar{x}}_{\boxed{A}}+\underbrace{n\bar{x}^2}_{\boxed{B}}\end{aligned}$$

［A］の項を見てください。

$$2と\bar{x}$$

には添字が付いていないので定数として扱えます。

$$2\text{は}\sum\text{の前に}$$
$$\bar{x}\text{は}\sum\text{の後ろに}$$

書かれていますが、単なる掛け算です。これは、

$$\sum ax_i = a\sum x_i$$

という公式を使っただけです。

［B］の項は単に

$$\bar{x}^2\text{を}n\text{回足している}$$

だけなので、

$$n\bar{x}^2$$

となりますね。

$$\bar{x}\text{の次数が高い方から項を並べ替え}$$

てみると、以下のようになります。

$$n\bar{x}^2 - 2(\sum x_i)\bar{x} + \sum {x_i}^2 \tag{2}$$

　これが、

$$\text{各データと}\bar{x}\text{との距離の2乗の総和（二乗和）}$$

ですね。よくよく見ると、これは

$$\bar{x}\text{の二次式}$$

です。ということは前回と同じように、平方完成すれば、

$$\text{この式の値を最小にする}\bar{x}$$

が求められます。が、わざわざ計算しなくても、係数だけを見れば

$$\text{式の値を最小にする}\bar{x}$$

が求められることが分かっていました。つまり、

$$ax^2 + bx + c$$

という式で a が正であれば、

$$-\frac{b}{2a}$$

を求めるだけでいい、というわけです（※この内容が分からない場合は、前回の平方完成の説明をもう一度確認してみてください）。そこで、式を見比べてみましょう。枠内には（2）式の係数を当てはめてみてください。

$$a = \boxed{}$$
$$b = -2\boxed{}$$

答えは、

$$a = \boxed{n}$$
$$b = -2\boxed{\sum x_i}$$

ですね。従って、

$$\begin{aligned} -\frac{b}{2a} &= -\frac{-2\boxed{\sum x_i}}{2\boxed{n}} \\ &= \frac{\sum x_i}{n} \end{aligned}$$

となります。これで、（2）式全体を最小にする値が

$$\frac{\sum x_i}{n}$$

だということが分かりました。ところで、

$$\frac{\sum x_i}{n}$$

って何でしょう。Σは全ての値を足すこと、n は値の個数なので、日本語で言うと「全ての値を足して個数で割ったもの」となります。これは平均値を求めるための式そのものです！

ということで、各データとの距離の 2 乗の総和（二乗和）を最小にする値は、全ての値を足して個数で割れば求められること、そして、それが**平均値（算術平均）**である、ということが一般的に表せた、というわけです。

ここでは、Σの公式に慣れるために上のような計算をしてみましたが、実はそれ以上の意味合いがあります。それは、このような式の変形は、たくさんの項を+でつないで総和を求める書き方だと簡単には表せないということです。Σという記号を使ったからこそ、このような計算ができたのだ、ということが重要です。

数学で、さまざまな記号を使うのには、複雑な計算が簡単に表せるというメリットがあるからです。繰り返しになりますが、謎の呪文を使ってイジワルをしようというわけではないのです。

最後に、「重心を求める」について解説します。これまでに学んだ知識を総動員しましょう。復習も兼ねています。

→目次に戻る

応用：重心を求める

もう少し応用的なお話もしておきましょう。機械学習の一つに**クラスタリング**と呼ばれる方法があります。これは、データを幾つかのグループに分けるのに使われる手法です。そのときに、各グループのデータの「重心」を求める必要があり、そのためにも平均値が使われます。

重心とは、各データからの距離の二乗和が最小になる点のことなので、やはり平均値に他なりません。小学校の頃から慣れ親しんできた平均値がここでも大活躍です。

平面上の二点の距離を求める方法については、この連載の第１回で見ました。そのときには具体的な数値を使って計算しましたが、文字式を使って表してみましょう。

$$\text{点}(x_1, y_1)\text{と点}(\bar{x}, \bar{y})\text{の距離}$$

は、図４の斜辺の長さに当たります。

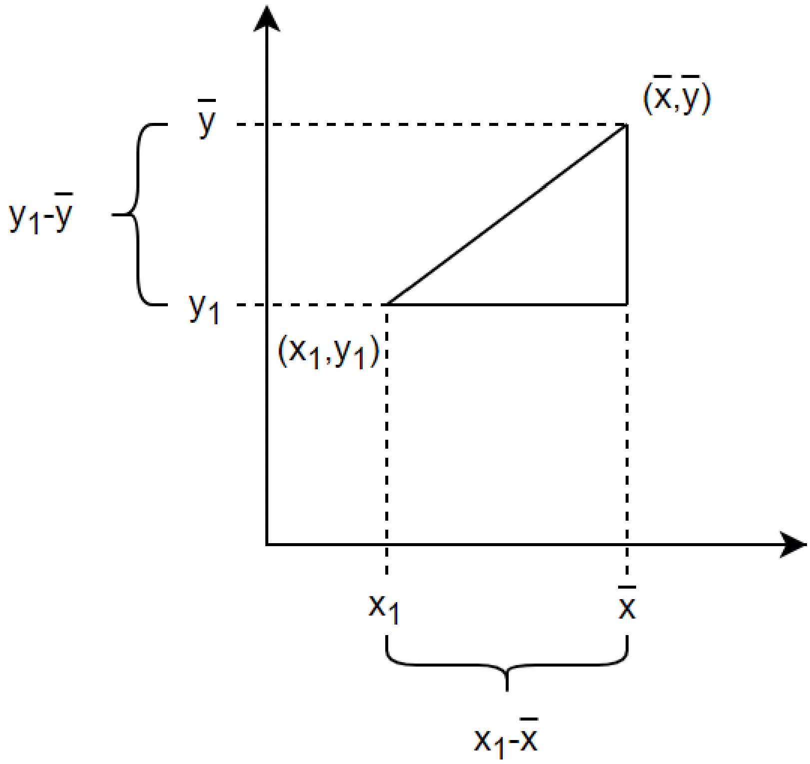


図４　２点の距離

ピタゴラスの定理を使って計算すると、

$$\text{点}(x_1, y_1)\text{と点}(\bar{x}, \bar{y})\text{の距離}$$

は、

$$\sqrt{(x_1 - \bar{x})^2 + (y_1 - \bar{y})^2}$$

ですね。いちいち√を求めるのも面倒なので、2 乗のままにしておきましょう。つまり、

$$点(x_1,y_1)と点(\bar{x},\bar{y})の距離の2乗$$

は、

$$(x_1-\bar{x})^2+(y_1-\bar{y})^2$$

ということになります。

次に、複数の点の重心を求めてみましょう。重心の座標を

$$(\bar{x},\bar{y})$$

とします。

以下のような図5の各点と、
$(\bar{x},\bar{y})$との距離（図5の矢印の長さ）
の二乗和

が

$$最小になるような(\bar{x},\bar{y})$$

を求めればいいですね。

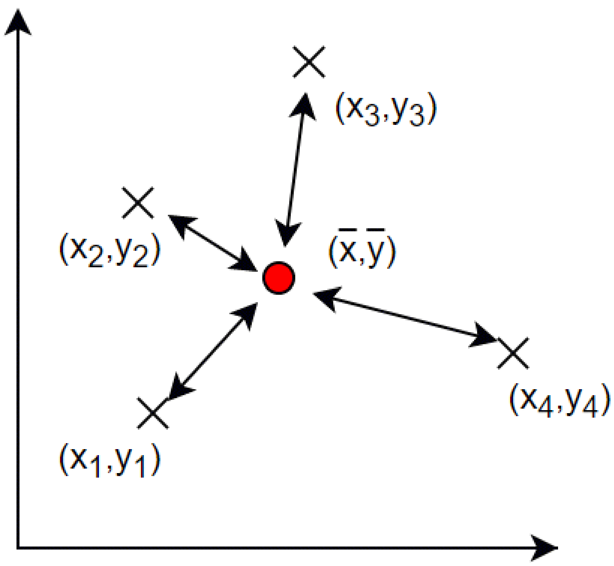


図 5　重心

上の図では、点は 4 つしか描かれていませんが、点が $(x_1,y_1),(x_2,y_2)\dots(x_n,y_n)$ のようにたくさんあるなら、

$$(x_1,y_1),(x_2,y_2)\dots(x_n,y_n)の各点と点(\bar{x},\bar{y})の距離の二乗和$$

は、

$$\begin{aligned}&(x_1-\bar{x})^2+(y_1-\bar{y})^2\\&+(x_2-\bar{x})^2+(y_2-\bar{y})^2\\&+\dots\\&+(x_n-\bar{x})^2+(y_n-\bar{y})^2\end{aligned}$$

となりますね。これをΣを使って表せば、

$$\begin{aligned}&\sum((x_i-\bar{x})^2+(y_i-\bar{y})^2)\\=&\sum(x_i-\bar{x})^2+\sum(y_i-\bar{y})^2\end{aligned}$$

となります。この式の値を最小にする

$$(\bar{x},\bar{y})$$

の値は、それぞれの平均値なので、

$$\left(\frac{\sum x_i}{n},\frac{\sum y_i}{n}\right)$$

です。これが重心の座標です *2。

*2

式の値を最小にする$\bar{x}$や$\bar{y}$の値

を求めるためには、微分法（偏微分）を使うのが便利です。いずれお話することになると思います。お楽しみに。

軸（変数）がもう一つ増えて、三次元になっても同様に計算できます（図 6）。考え方は同じですが、ちょっと複雑に見えるので、動画での解説も用意してあります（動画 3）。

動画 3　重心の座標を求める（三次元編）

→目次に戻る

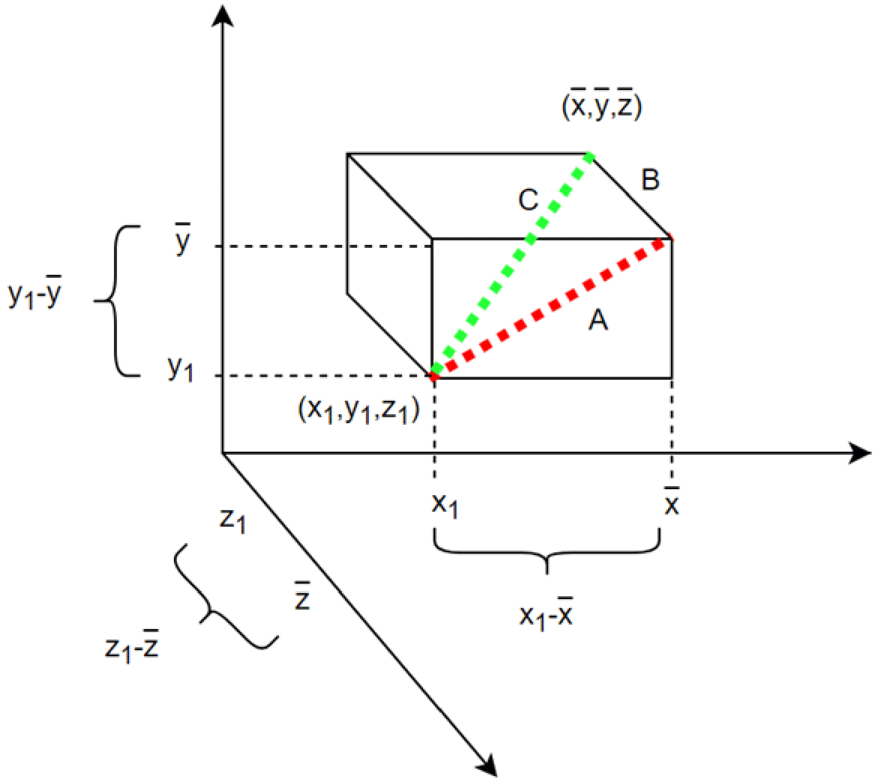


図 6　2 点の距離（三次元）

求めたい距離は、

$$(x_1, y_1, z_1)\text{から}(\bar{x}, \bar{y}, \bar{z})\text{までの距離}$$

で、図 6 の C の長さに当たります。ここでは、三角形 ABC が直角三角形になっていることに注目してください。直角三角形ということは、やはりピタゴラスの定理が使えますね。つまり、

$$C^2 = A^2 + B^2$$

です。

A^2 は上で求めた通り、

$$(x_1 - \bar{x})^2 + (y_1 - \bar{y})^2$$

です。B^2 は

$$(z_1 - \bar{z})^2$$

です。よって、C^2 は、

$$C^2 = (x_1 - \bar{x})^2 + (y_1 - \bar{y})^2 + (z_1 - \bar{z})^2$$

であることが分かります。

重心も同様に求められます。点が $(x_1,y_1,z_1),(x_2,y_2,z_2)\ldots(x_n,y_n,z_n)$ のようにたくさんあるなら、

$$\text{それらの点と}\bar{x}, \bar{y}, \bar{z}\text{との距離の二乗和}$$

は、二次元の場合と同様、

$$\sum(x_i-\bar{x})^2+\sum(y_i-\bar{y})^2+\sum(z_i-\bar{z})^2$$

となり、この式の値が最小になるときの

$$\bar{x},\bar{y},\bar{z}$$

の値は、それぞれの平均値なので、重心の座標は、

$$\left(\frac{\sum x_i}{n},\frac{\sum y_i}{n},\frac{\sum z_i}{n}\right)$$

となります。さらに変数が増えても考え方は同じです ***3**。

***3** 軸の数が増えてくると x,y,z では変数を表しにくくなるので、$x_1,x_2,x_3,\ldots$ のように添字を使って変数を区別することがあります。この場合の添字は個々のデータを区別するものではなく、あくまで軸（変数）を区別するためのものです。従って、x_1 と x_2 は別の変数です（x という変数の 1 番目と 2 番目ではありません！）。x_1 という変数の個々のデータを区別するためには、もう一つ添字を使って、

$$x_{11},x_{12},\ldots,x_{1n}$$

のように表します。これは、x_1 という変数の **1 番目**の値、x_1 という変数の **2 番目**の値、....、x_1 という変数の **n 番目**の値という意味になります。

ちなみに、**m 次元**の **n 個**の点の重心の座標は、

$$\left(\frac{\sum x_{1i}}{n},\frac{\sum x_{2i}}{n},\ldots,\frac{\sum x_{mi}}{n}\right)$$

と表されます。

ここではクラスタリングなどの詳細については触れませんが、例えば **k-means** 法では、以下のような手順でグループを分けます（これまでに学んだ知識だけでできます）。

1. 最初はランダムに幾つかのグループの重心の座標を決める
2. 各点を、重心との距離が一番近いグループに入れる
3. 各グループの重心の座標を計算し直す
4. 2 に戻る（重心が動かなくなるまで繰り返す）

*余談　これまで、機械学習で使われる数学を理解するために x,y,z といった変数を使って数式の意味や使い方を解説してきました。それでは実感が湧かない、現実の仕事にどう使われているのか分からないという人も多いでしょう。

　今回のお話とは異なる例ですが、重回帰分析では「面積」「駅歩」から「家賃」を予測するというようなことを行います。この場合は「面積」を x、「駅歩」を y、「家賃」を z で表す、と考えれば実感が湧くのではないかと思います。さらに「築年数」も考慮すると、変数が多くなって、文字が足りなくなるので、「面積」を x_1、「駅歩」を x_2、「築年数」を x_3 という変数で表し、「家賃」を y という変数で表すといった方法を使います。

　念のために繰り返しておきますが、この場合の添字は各データを区別するためのものではなく、変数を区別するためのものです。

次回は……

　これまでのお話では二次式までしか扱ってきませんでした。機械学習では、三次、四次といった高次の関数や指数関数、対数関数などを扱うことも考えられます。また変数をたくさん使い、それらの連立方程式を解く、といったことも必要になってきます。さらに、値の変化を見て最適な解を求めることもあります。そういった問題を解決するための強力な助っ人が「微分」です。

　本書「中学数学からのおさらい編」の本編は以上です。今後の電子書籍化として検討している「微分／偏微分編」では、微分法を取り上げ、さまざまな関数の解析ができるようにするための基礎を学びます。計算方法は平方完成よりもはるかに簡単なので、ご心配なく。

次のページ以降では番外編（Appendix）として、「0」の取り扱い、指数、対数について簡単に説明しています。

[AI・機械学習の数学]番外編1　「0」の取り扱い

連載の通常の流れとは別の番外編。「0」の取り扱いについて 3 つのポイントを解説。0 で割ること、0 乗、0 の階乗について説明する。

羽山博，著（2020 年 08 月 03 日）

今回は連載の通常の流れとは別枠の「番外編その 1」です。コラムを読むような気持ちで気楽に読んでみてください。通常とは異なり、例題や練習問題はありません。ポイント解説のみの短い記事です。

数学史では **0** の発見などについての興味深いさまざまな話がありますが、ここでは、そういった話ではなく、実用面に絞って **0** の取り扱い方を整理しておきます。今回は中学／高校数学の基礎のおさらいとなっており、基本的な約束事からべき乗や階乗での **0** の取り扱いについて見ていきます。

ポイント 1　「0」で割ってはいけない

0 で割り算（**0** 除算）することはできない、というのはもはや常識といってもいいでしょう。一方、**0** を何かで割ることはできます。その場合の答えは **0** です。

$$10 \div 0 \cdots \text{この計算はできない}$$
$$0 \div 10 = 0$$

プログラミング言語では、多くの場合、リスト 1.1 のように **0** で割り算をするとエラーになります（中には「無限大（∞）」と表示されたりするものもあります）。

```
>>> a=10
>>> b=0
>>> print(a/b)
Traceback (most recent call last):
  File "<stdin>", line 1, in <module>
ZeroDivisionError: division by zero  …… エラーが表示された
>>>
```

リスト 1.1　Python で 0 除算をした例

具体的な数字を使っている場合には間違えることはないと思いますが、文字式では、割る数（つまり分母）が **0** になっていないことを確認しておく必要があります。また、プログラミングでは **0** での割り算が起こったときに備えての対応が必要です。なお、「**0** に限りなく近づく」という場合は **0** ではありません。以下の式は本編の第 4 回で登場する導関数の定義です。この場合、h の値は限りなく **0** に近づきますが **0** ではありません。

$$\lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$

ポイント 2　何かを 0 乗すると 1 になる

どのような値であっても **0 乗**すると **1** になる、というのを一般的に表すと以下のようになります。

$$x^0 = 1$$

このことを具体的な値で確かめておきましょう。べき乗というのは、同じ数を何回か掛けることでしたね。例えば、**2** の **3 乗**は **2** を **3 回**掛けることを表します。

$$\begin{aligned} 2^3 &= 2 \times 2 \times 2 \\ &= 8 \end{aligned}$$

答えは **8** です。では、**2** の **3 乗**を **2** で割るとどうなるでしょうか。

$$\begin{aligned} 2^3 \div 2 &= 2 \times 2 \times 2 \div 2 \\ &= \frac{2 \times 2 \times 2}{2} \qquad \cdots \boxed{A} \\ &= 2 \times 2 \\ &= 2^2 \end{aligned}$$

[A] …… 2 で約分できる

2 の **3 乗**を **2** で割ると、**2** の **2 乗**になりますね。答えは **4** です。このことから、同じ値で割り算をするなら、指数を 1 つ減らせばいいということが分かります。では、**2** の **3 乗**を **2** の **2 乗**で割ってみるとどうなるでしょうか。

$$\begin{aligned} 2^3 \div 2^2 &= 2 \times 2 \times 2 \div (2 \times 2) \\ &= \frac{2 \times 2 \times 2}{2 \times 2} \qquad \cdots \boxed{B} \\ &= 2 \\ &(= 2^1) \end{aligned}$$

[B] …… 2×2 で約分できる

2 の **3 乗**を **2** の **2 乗**で割ると、答えは **2** になります。これは **2** の **1 乗**です。やはり、以下のように指数を **2** 減らせばいいということですね。

$$2^3 \div 2^2 = 2^{(3-2)}$$

では、**2 の 3 乗**を **2 の 3 乗**で割ってみましょう。ある数を同じ数で割っているので、答えは **1** になるはず。これまでと同じように、掛け算と割り算でも表してみます。

$$\begin{aligned} 2^3 \div 2^3 &= 2 \times 2 \times 2 \div (2 \times 2 \times 2) \\ &= \frac{2 \times 2 \times 2}{2 \times 2 \times 2} \\ &= 1 \end{aligned}$$

これを、指数を加減した式で表すと、以下のようになります。

$$\begin{aligned} 2^3 \div 2^3 &= 2^{(3-3)} \\ &= 2^0 \end{aligned}$$

というわけで、$2^0 = 1$ ということが分かりました。このことは、どんな値についても言えることなので、一般的に表すと最初に示したような以下の式になります。

$$x^0 = 1$$

さらに、**2 の 3 乗**を **2 の 4 乗**で割ってみましょう。

$$\begin{aligned} 2^3 \div 2^4 &= 2 \times 2 \times 2 \div (2 \times 2 \times 2 \times 2) \\ &= \frac{2 \times 2 \times 2}{2 \times 2 \times 2 \times 2} \qquad \cdots \boxed{C} \\ &= \frac{1}{2} \end{aligned}$$

[C] …… 2×2×2 で約分できる

指数を加減した式で表すと、以下のようになります。

$$\begin{aligned} 2^3 \div 2^4 &= 2^{(3-4)} \\ &= 2^{-1} \end{aligned}$$

つまり、

$$2^{-1} = \frac{1}{2}$$

となります。これも一般的に表すと以下のようになります。

$$x^{-n} = \frac{1}{x^n}$$

0 を通り過ぎて、マイナスの話になってしまいましたが、指数を減らして「**0 乗**」になる場面は、微分や偏微分でも登場します。

ポイント 3　「0」の階乗は 1

階乗とは、値を **1** つずつ減らしていきながら **1** まで掛けるという計算で、値の後に「!」を付けて表します。例えば、**4** の階乗であれば、以下のようになります。

$$\begin{aligned} 4! &= 4 \times 3 \times 2 \times 1 \\ &= 24 \end{aligned}$$

0 の階乗、つまり 0! の答えはいくらになるでしょうか。直感的には **0** のような気がしますが、実はこの答えは **1** です。

$$0! = 1$$

これは「そう取り決めておくと都合がいい」からです。

せっかくなので、階乗についても簡単に復習しておきましょう。階乗は順列や組み合わせの計算、確率の計算でよく使われます。例えば、A、B、C の **3 人**を順に並べる方法が何通りあるかを求めるには、最初の人が **3 人**のうちのいずれかなので **3 通り**、そのそれぞれに対して、次が **2 人**のうちのいずれかなので **2 通り**、最後は **1 人**なので **1 通り**です。従って、

$$\begin{aligned} 3! &= 3 \times 2 \times 1 \\ &= 6 \text{（通り）} \end{aligned}$$

となるわけです。また、A、B、C、D、E の **5 人**から **3 人**を選んで順に並べるなら、以下のようになります。

$$\begin{aligned} & 5 \times 4 \times 3 \\ &= 60 \text{（通り）} \end{aligned}$$

階乗を使ってこれを表すなら、以下のようにできますね。

$$\begin{aligned} \frac{5!}{2!} &= \frac{5 \times 4 \times 3 \times 2 \times 1}{2 \times 1} \\ &= 60 \text{（通り）} \end{aligned}$$

一般に、n 個のうちから r 個を選んで順に並べる場合の数（順列）は、以下の式で表されます。

$$n \times (n-1) \times (n-2) \times \cdots \times (n-r+1) = \frac{n!}{(n-r)!}$$

A、B、C、D、E の **5 人**から **3 人**を選んで順に並べる場合なら、n が **5**、r が **3** になります。上の式にそれらの値を当てはめてみると納得できるはずです。ところで、n = r の場合はどうなるでしょう。つまり n 人の中から n 人を選んで並べる場合です。

例えば、最初の例は **3 人**のうちから **3 人**を順に並べることになるので、それに当てはまりますね。上の式（の右辺）で $n = 3$、$r = 3$ とすると、以下のようになります。

$$\frac{3!}{(3-3)!} = \frac{3!}{0!}$$

答えは最初に計算した通り、$3!$ でした。ということは、$0! = 1$ としないとつじつまが合わないですね。そういった理由もあって、$0! = 1$ と定義されている、ということです。

今回の番外編は **0** の取り扱いについての留意点を見ましたが、次回は、指数と対数の取り扱いや「自然対数の底 e」のお話など、横断的なテーマを取り上げることとします（ちょっと先取りになりますが、指数関数や対数関数の微分法についても、公式と例を紹介しておきます）。

[AI・機械学習の数学]番外編2　指数と対数（指数編）

連載の通常の流れとは別の番外編。AI や機械学習でよく使う「指数」を解説。指数関数の性質や指数関数の微分法についても簡単に紹介する。

羽山博，著（2021 年 05 月 13 日）

この連載の本編では、微分やベクトル、行列などのいわば数学の「縦糸」に当たるテーマを取り上げていますが、番外編では、さまざまなテーマにまたがる、いわば「横糸」に当たるテーマを取り上げます。今回は、AI や機械学習に登場する数式の中でよく使われる「指数」に焦点を当てることにします。なお、指数と切っても切り離せない「対数」については、番外編 3 で取り扱うことにします。

指数は、同じ数を何回か掛けることを簡単に表すのに使います。その「何回か」は自然数（**1 以上**の整数）でなくても構いません。ここでは、有理数（負の数を含めて、分数で表せる範囲の値）について、指数の取り扱いを見ていきます。また、指数関数の性質や指数関数の微分法についても簡単に紹介します。

ポイント 1　指数関数の基本公式

指数の計算方法については、本編の第 1 回や番外編 1 でも取り扱ったので、具体的な計算の例はそちらに譲ることとして、まずは基本公式をまとめておきましょう。ただ、公式だけだと実感が湧かないでしょうから、簡単な例も併せて示すこととします。

掛け算と割り算の公式

$$\begin{aligned} x^m \cdot x^n &= x^{m+n} \\ \frac{x^m}{x^n} &= x^{m-n} \end{aligned}$$

これらの式で、x に当たる値のことを「**底**（てい）」と呼びます。m や n は「**指数**」ですね。以下の例では、いずれも底は **2** です。

（例 1）

$$\begin{aligned} 2^3 \times 2^4 &= 2^{3+4} \\ &= 2^7 \\ &= 128 \end{aligned}$$

（例 2）

$$\begin{aligned} \frac{2^7}{2^3} &= 2^{7-3} \\ &= 2^4 \\ &= 16 \end{aligned}$$

指数の部分が掛け算になっている場合は、以下のようになります。

$$x^{m \cdot n} = (x^m)^n \cdots \boxed{A}$$

指数の部分がさらに指数になっている場合は、以下のように右から左へと計算します。

$$x^{m^n} = x^{(m^n)} \cdots \boxed{B}$$

14 ページ目でも説明しましたが、Excel でべき乗を表す演算子「^」を使って「=2^3^2」と入力すると、［A］のような計算方法になり、結果は、

$$(2^3)^2 = 8^2 = 64$$

となります。一方、Python でべき乗を表す演算子「**」を使って「=2**3**2」と入力すると［B］のような計算方法になり、結果は、

$$2^{(3^2)} = 2^9 = 512$$

となります。

何を 0 乗しても 1 になる

$$x^0 = 1 \quad ただし、x \neq 0$$

これは、割り算の公式で m = n である場合に当たります。なお、0^0 は定義されません（このような計算は許されません）が、0^n は **0** です。

マイナス乗の場合

$$x^{-n} = \frac{1}{x^n}$$

これは、割り算の公式で m = 0 である場合に当たります。

（例 3）

$$\begin{aligned} 2^{-3} &= \frac{1}{2^3} \\ &= \frac{1}{8} \end{aligned}$$

コラム　n 進数のしくみと指数

私たちが普段使っている十進数は、各桁の値と **10** を底とした指数の積を全て足した値と考えられます。例えば、

$$1234 = \underbrace{1 \times 10^3}_{10^3=1000\text{の位の個数}} + \underbrace{2 \times 10^2}_{10^2=100\text{の位の個数}} + \underbrace{3 \times 10^1}_{10^1=10\text{の位の個数}} + \underbrace{4 \times 10^0}_{10^0=1\text{の位の個数}}$$

となります。右端が 1 の位（10^0 の位）です。この例だと、1 の位が **4 個**あるということですね。10 の位（10^1 の位）が **3 個**、100 の位（10^2 の位）が **2 個**、1000 の位（10^3 の位）が **1 個**あるというわけです。指数に注目すると、左の桁に行けば指数が **1** 増えることが分かります。逆に言うと、右の桁に行けば指数が **1** 減るわけですね。そして、1 の位の指数は **0** となります。

ここで、小数点以下を考えてみましょう。右に行けば指数が **1** 減るので、1 の位の右、つまり小数第一位の指数は -1 になり、さらにその右の小数第二位の指数は -2 になると考えられます。例えば、

$$3.14 = 3 \times 10^0 + \underbrace{1 \times 10^{-1}}_{10^{-1}=\frac{1}{10}\text{の位の個数}} + \underbrace{4 \times 10^{-2}}_{10^{-2}=\frac{1}{10^2}\text{の位の個数}}$$

と表せます。$10^{-1} = 1/10 = 0.1$ となり、$10^{-2} = 1/100 = 0.01$ となるので、**3.14** は、1 の位（10^0 の位）が **3 個**、0.1 の位（10^{-1} の位）が **1 個**、0.01 の位（10^{-2} の位）が **4 個**集まった数であることが分かります。ちゃんと計算が合っていますね。

なお、Python の round 関数や Excel の ROUND 関数で四捨五入の桁位置を指定する場合には、上記のように表現したときの「**- 指数**」の値を指定します。例えば、Python で小数点以下第 2 位まで求めるときには 10 の -2 乗の位まで求めるので、**-(-2)** つまり、**2** を桁位置として指定し、round(1.234,2) と書きます。結果は **1.23** ですね。round(4649,-1) なら 10 の 1 乗の位まで求めるので、結果は **4650** となります。

上のような表し方は何進数でも同じです。例えば二進数であれば、底が **2** になるだけのことです。例えば、**110.101** という二進数であれば、

$$110.101 = \underbrace{1 \times 2^2 + 1 \times 2^1 + 0 \times 2^0}_{\text{整数部}} + \underbrace{1 \times 2^{-1} + 0 \times 2^{-2} + 1 \times 2^{-3}}_{\text{小数部}}$$

と表せます。整数部は **6** ですね。小数部は $2^{-1} = 1/2^1 = 0.5$、$2^{-3} = 1/2^3 = 0.125$ なので、**0.625** になります。つまり、二進数の **110.101** は十進数の **6.625** であることが分かります。

ポイント 2　指数が分数の場合

指数が 1/n である場合は、**n 乗根**の**主要根**になります *1。まず、公式を記しておきます。

$$x^{\frac{1}{n}} = \sqrt[n]{x}$$
$$x^{\frac{m}{n}} = \sqrt[n]{x^m}$$

上の式の

$$\sqrt[n]{x}$$

とは、**n 乗**すると x になるような値のことです。例えば、

$$\sqrt[3]{8}$$

は、**3 乗**すると **8** になる値です。$2^3 = 8$ なので、この値は **2** となります。なお、

$$\sqrt[2]{x} \quad \text{は単に} \quad \sqrt{x}$$

と書きます。

下の式は、指数の部分が m と 1/n の掛け算になっているものと考えられます。すでに見た［A］式、つまり、

$$x^{m \cdot n} = (x^m)^n$$

で、n のところに 1/n を書けば、

$$x^{m \cdot \frac{1}{n}} = (x^m)^{\frac{1}{n}} = \sqrt[n]{x^m}$$

となりますね。

*1　n 乗根は**冪根**（べきこん）や**累乗根**（るいじょうこん）とも呼ばれ、n = 2 の場合を**平方根**、n = 3 の場合を**立方根**と呼びます。n 乗根は 1 つとは限りません。例えば、**16** の 4 乗根は「4 乗すれば **16** になる値」のことなので、実数の範囲であれば、**2** と **-2** の 2 つになります。このうち、正の値の方を**主要根**と呼びます。

つまり、**16** の 4 乗根は、

$$\sqrt[4]{16} = 2$$
$$-\sqrt[4]{16} = -2$$

の 2 つですが、

$$16^{\frac{1}{4}} = \sqrt[4]{16} = 2$$

です。

なお、複素数にまで範囲を広げると、**16** の 4 乗根は、ほかにも 2i と－ 2i（i は虚数単位）があります。

指数が分数の場合の例はこれまで出てこなかったので、少し見ておきましょう。

(例 4)

$$\begin{aligned}125^{\frac{1}{3}} &= \sqrt[3]{125}\\ &= \sqrt[3]{5\times5\times5}\\ &= 5 \qquad\qquad \cdots \text{3乗すると125になる値は5}\end{aligned}$$

(例 5)

$$\begin{aligned}256^{\frac{3}{4}} &= \sqrt[4]{256^3}\\ &= \sqrt[4]{16777216}\\ &= \sqrt[4]{64\times64\times64\times64}\\ &= 64 \qquad\qquad \cdots \text{4乗すると16777216になる値は64}\end{aligned}$$

実際の計算は Python などのプログラミング言語や Excel などにまかせてしまえばいいのですが、

$$\sqrt[n]{x^n} = x$$

なので、根号（√）の中を素因数分解し、素因数の n 乗の形にして根号の外に出すと、式を簡単にできます。

(例 6)

$$\begin{aligned}\sqrt[3]{24} &= \sqrt[3]{2\cdot2\cdot2\cdot3}\\ &= \sqrt[3]{2^3\cdot3}\\ &= 2\sqrt[3]{3}\end{aligned}$$

コラム　Python で底が負の場合は√と 1/2 乗の結果が異なる

n の正の平方根は√ n ですが、これは、

$$n^{\frac{1}{2}}$$

と表すこともできます。ただし、n が負の場合は虚数（複素数）となります。Python では、リスト 1 のようにべき乗で表した場合は複素数が返されます。

```
>>> (-4)**(1/2)
(1.2246467991473532e-16+2j)
```

リスト 1　Python の基本機能で負の数の 1/2 乗を計算した結果

-4 の 1/2 乗は **2j**（**j** は虚数単位）となります。前の **1.2246467991473532e-16** は本来は **0** ですが、わずかな誤差が表示されています。

一方、リスト 2 のように math モジュールの sqrt 関数を使って、引数に負の値を指定するとエラーになります。

```
>>> from math import sqrt
>>> sqrt(-4)
Traceback (most recent call last):
  File "<stdin>", line 1, in <module>
ValueError: math domain error
```

リスト 2　Python の math モジュールで負の数の平方根を計算した結果

ポイント 3　指数関数の値とグラフ

指数関数は、$f(x) = a^x$ のように、変数 x が指数になっているものです。$f(x) = 2^x$ であれば、以下のような値になります。二次関数 $f(x) = x^2$ との違いも見比べておきましょう。

x	0	1	2	3	4	5	6	7	8	9	10
二次関数 x^2	0	1	4	9	16	25	36	49	64	81	100
指数関数 2^x	1	2	4	8	16	32	64	128	256	512	1024

表 1　二次関数と指数関数の値の変化

これだけの値を見ても分かるように、指数関数では x の値が大きくなると急激に f(x) の値が大きくなります。グラフも描いておきましょう。

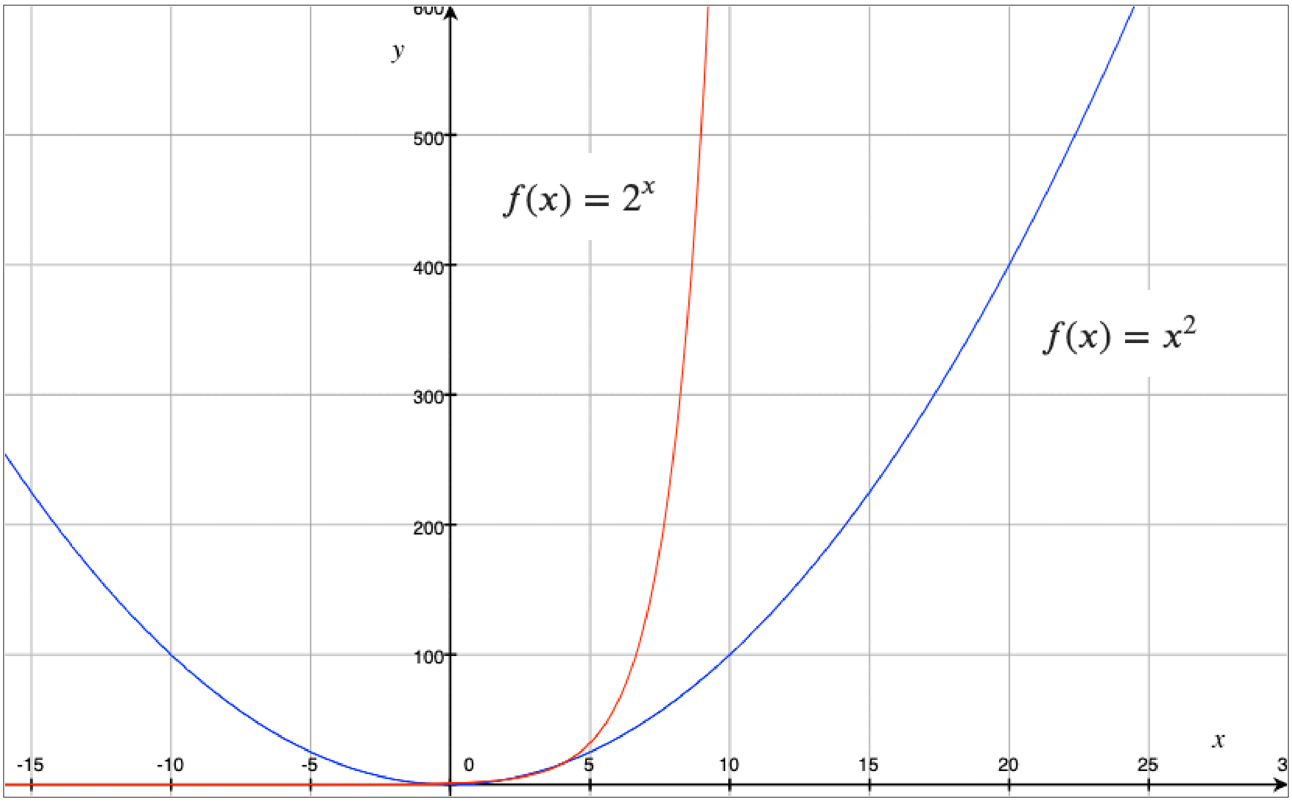


図 1　二次関数と指数関数のグラフ

$x = 4$ を超えたあたりから二次関数と指数関数の差が急激に大きくなる。例えば、$x = 20$ のとき、$20^2 = 400$ だが、$2^{20} = 1048576$ となり、図をはみ出してしまう。

ポイント４　指数関数の微分法

指数関数の微分法については、公式と簡単な例を記すにとどめます。

$$f(x)=a^x$$

のとき、

$$f'(x)=a^x\log_e a\cdots\boxed{C}$$

です。log は**対数**を表すので、対数の話をしないと上の式は使えませんね。というわけで、対数のお話をしましょう……といいたいところですが、それは番外編３で見ることとして、もう一つ重要な公式を記しておきます。

$$(e^x)'=e^x\cdots\boxed{D}$$

e は**自然対数**の底と呼ばれる値で、実際の値は **2.7182...** です。e^x を微分しても e^x になるというところが重要なポイントです。e は［C］式にもさりげなく登場していますね。

対数のお話をしないと指数関数の微分はできないのですが、雰囲気だけでもつかめるように、単に公式を適用しただけのものですが、簡単な具体例を記しておきます。

（例７）

$$(2^x)'=2^x\log_e 2$$

（例８）

$$\begin{aligned}(e^x)'&=e^x\log_e e\\&=e^x\cdot 1\\&=e^x\end{aligned}$$

例８も公式を適用しただけですが、［D］が成り立つことが分かります。次回の番外編３で解説しますが、$\log_e e=1$ という公式も適用しました。

コラム　正規分布やシグモイド関数などで使われる自然対数の底 e

自然対数の底 e は、機械学習や統計学などのさまざまな場面で登場します。本編の第 14 回で紹介した正規分布の確率密度関数でも使われています。そこで見たように、e^x は exp(x) と表すことがあります。x の部分の式が複雑になるような場合に便利な表し方です。

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

ほかにも、e はニューラルネットワークの活性化関数の一つである**シグモイド関数**の中でも使われます。シグモイド関数は以下の式で表されます。

$$f(x) = \frac{1}{1+e^{-x}}$$

シグモイド関数をグラフ化すると以下のようになります。

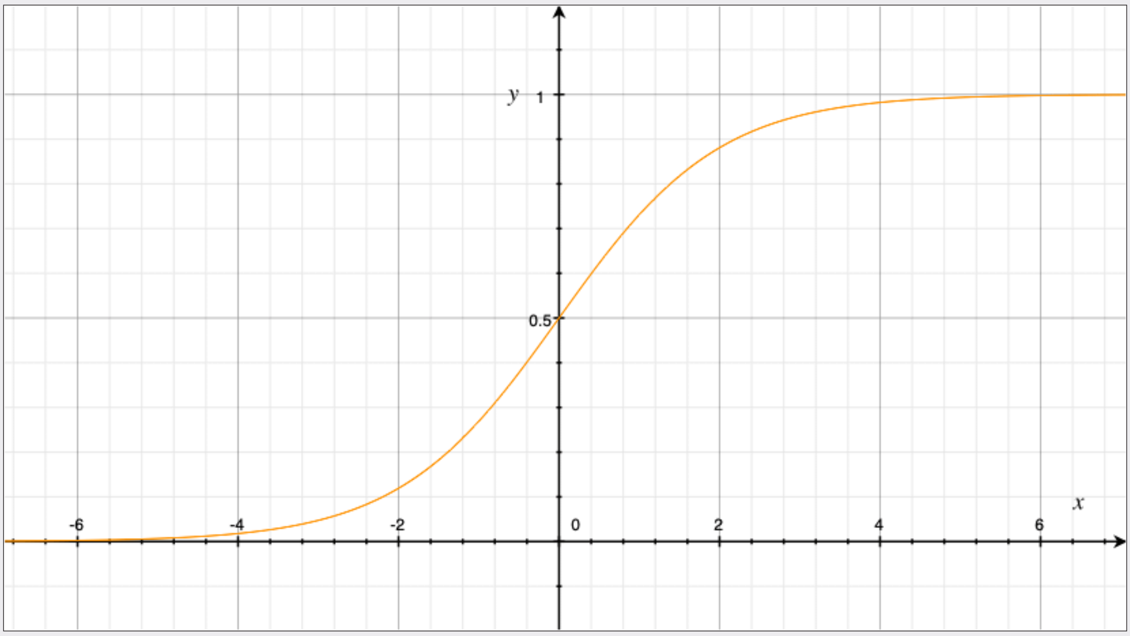


図 2　シグモイド関数のグラフ

シグモイド関数は、下限が **-1**、上限が **1** で、x = 0 のとき、f(x) = 0.5 となる。

活性化関数やシグモイド関数の詳細については、AI・機械学習の用語事典などをぜひ参照してください。

ここでは 2 つだけ例を挙げましたが、自然対数の底 e は数学や物理などのさまざまな分野の数式に、頻繁に登場します。詳細については説明しませんが、世界一美しい数式として有名なオイラーの公式（$e^{i\pi}+1=0$）にも登場します。

今回は、指数の計算方法と指数関数、指数関数の微分法について簡単に整理しました。最後に、指数関数の微分に「対数」が必要であるということに少し触れましたが、対数については、次の番外編 3 で詳しく見ていくことにします。

[AI・機械学習の数学]番外編3　指数と対数(対数編)

連載の通常の流れとは別の番外編。指数と切っても切れ離せない「対数」を解説。対数関数の性質や対数関数の微分法についても簡単に紹介する。

羽山博，著（2021 年 05 月 17 日）

この連載の本編では、微分やベクトル・行列などのいわば数学の「縦糸」にあたるテーマを取り上げていますが、番外編では、さまざまなテーマにまたがる、いわば「横糸」にあたるテーマを取り上げます。前回は「指数」を取り上げましたが、今回は指数と切っても切り離せない「対数」について見ていくこととします。

対数とは、端的にいえばある数の指数部分を取り出す計算です。対数を使えば、掛け算になっている式を足し算で表せるように変形できます。また、指数関数の値の対数を取れば、グラフが直線で表せます。まずは、対数の表し方から見ていきましょう。

ポイント 1　対数の表し方

対数を理解するには、指数をきちんと理解しておく必要があります。おさらいをかねて、具体例で見てみましょう。$2^3 = 8$ の意味をあらためて確認しておきます。図 1 の上半分を見てください。**2** は「**底**（てい）」でしたね。**3** はもちろん「指数」です。**2** の 3 乗が **8** になるというわけです。

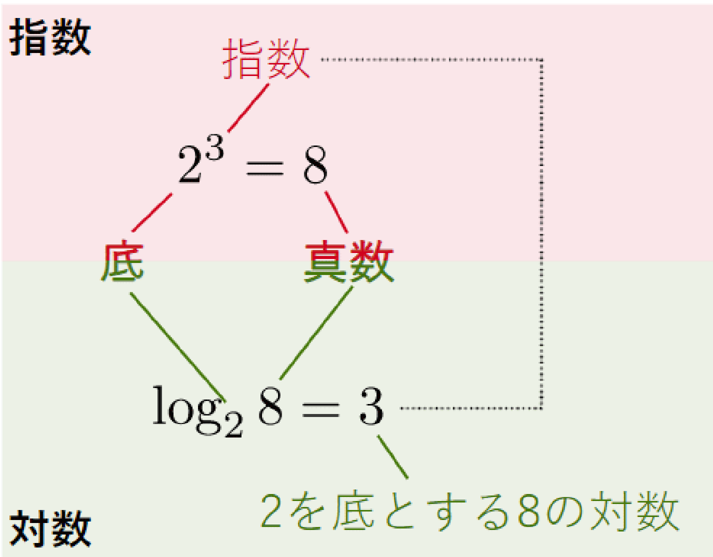


図 1　指数と対数の関係

対数とは、ある数の指数を求めることにほかならない。

では、図 1 の下半分に移ります。書き方を見ながら対数の意味を明らかにしていきましょう。対数を表すには、最初に log（ログ）と書き、log の右下に小さく底を書きます。この場合であれば、底は **2** です。そして、その隣に「真数（しんすう）」と呼ばれる値を書きます。**真数**とは、底を何乗かした値のことです。この例だと、**8** が真数です。つまり、**2** を何乗かしたら **8** になるということです。

- $\log_2 8$ …… **2** を何乗かしたら **8** になる、その「何乗か」を求める

答えはもちろん **3** ですが、これは、その「何乗か」という値のことですね。それが対数です。つまり、対数とは、底を決めたときに、真数を得るための指数はいくらであるかということです。図 **1** を見れば、指数と対数は表裏一体の関係にあることが分かります。少し例を見ておきましょう。

(例 1)

$$\log_3 81 = 4$$

3 を何乗したら **81** になるか……$3^4 = 81$ なので、答えは **4**。

(例 2)

$$\log_2 10 = 3.3219\ldots$$

2 を何乗したら **10** になるか……$2^x = 10$ を解けば $x =$ **3.3219...** となります（この方程式は簡単には解けませんが $2^3 = 8, 2^4 = 16$ であることから、**3** と **4** の間の値であることは分かりますね）。

なお、底が **10** の対数（例えば $\log_{10}81$ など）のことを**常用対数**と呼び、底が e の対数（例えば $\log_e 81$ など）のことを**自然対数**と呼びます。特に底が指定されていない場合（例えば log 81 など）は e が底であるものと見なされます。ただし、底が e でない場合でも、文脈から底の値が明らかであるときには、底を省略して書くこともあります。

ポイント 2 対数の性質や公式のまとめ

対数の意味が分かれば、以下の式の意味も簡単に分かると思います。

$$\begin{aligned}\log_a a^x &= x \cdots a\text{を何乗したら}a^x\text{になるか（答え：}x\text{乗）}\\ \log_a a &= 1 \cdots a\text{を何乗したら}a\text{になるか（答え：1乗）}\\ \log_a 1 &= 0 \cdots a\text{を何乗したら1になるか（答え：0乗）}\end{aligned}$$

続いて、基本的な公式についてもまとめておきます。以下の公式は底が **1 以外**の正の実数であれば成り立つので、底を省略しています。

$$\begin{aligned}\log xy &= \log x + \log y & \cdots \boxed{A}\\ \log \frac{x}{y} &= \log x - \log y & \cdots \boxed{B}\\ \log x^y &= y \log x & \cdots \boxed{C}\\ \log_x y &= \frac{\log_a y}{\log_a x} & \cdots \boxed{D}\end{aligned}$$

［A］…… 掛け算は足し算になる（ただし x,y は正の実数）

［B］…… 割り算は引き算になる（ただし x,y は正の実数）

［C］…… べき乗は log の外に出せる（ただし x は正の実数、y は実数）

［D］…… 底を a に変換する（x,y,a は正の数、x ≠ 1,a ≠ 1）

［B］式で x の値が **1** であれば、log x = 0 になるので、以下の式が成り立つことも分かります。

$$\begin{aligned}\log\frac{1}{y} &= \log 1 - \log y\\ &= 0 - \log y\\ &= -\log y\end{aligned}$$

それぞれ、例を見ておきましょう。

（例 3）

$$\begin{aligned}\log_2(4\times 8) &= \log_2 4 + \log_2 8\\ &= 2+3=5\end{aligned}$$

左辺は **2** を何乗したら **32**(= 4×8) になるかということです。$2^5 = 32$ なので答えは **5** です。右辺の計算と合っていますね。これは、指数に関する以下の公式に対応するものです。指数の部分だけを見ると同じ形になっていますね（a = 2,m = 2,n = 3 となります）。

$$a^m \times a^n = a^{m+n}$$

（例 4）

$$\begin{aligned}\log_{10}\frac{10}{1000} &= \log_{10}10 - \log_{10}1000\\ &= 1-3\\ &= -2\end{aligned}$$

（例 5）

$$\begin{aligned}\log_2 10^3 &= 3\log_2 10\\ &= 3\times 3.3219\\ &= 9.9657\end{aligned}$$

例 5 では答えを小数点以下 4 桁まで求めています。既に例 2 で $\log_2 10 = 3.3219$ であることは見たので、その値を使いました。結果が整数にならない場合は手計算では難しいので、Python や Excel などの関数を使って $\log_2 10^3$ の値と $3\log_2 10$ を比べておきます。Python なら以下の通りです。

→目次に戻る

```
>>> import math
>>> math.log(10**3,2)
9.965784284662087
>>> 3*math.log(10,2)
9.965784284662089
```

リスト 1　Python で（例 5）の検算をしてみた

最下位の値が異なるのは浮動小数点数の計算時に起こる誤差のため。数学的には等しい。

Excel なら、「=LOG(10^3,2)」と「=3*LOG(10,2)」を任意のセルに入力して結果を見比べるといいでしょう。

（例 6）

$$
\begin{aligned}
\log_5 10 &= \frac{\log_2 10}{\log_2 5} \\
&= \frac{3.32}{2.32} \\
&= 1.43
\end{aligned}
$$

ここでは、小数点以下第二位まで計算しました。$\log_2 10 = 3.32...$、$\log_2 5 = 2.32...$ です。このように、底を変換すると、そのままでは計算しづらい場合でも、既知の値を使った計算にできることがあります。

対数の公式についてはこれぐらいにして、次に対数関数について見ていきます。が、その前にコラムを 1 つ挟んでおきましょう。以下のコラムは、これまでに見た公式を活用するので、対数の計算方法を確認するのにも役立ちます。

コラム　対数と情報量の表し方

情報量とは、平たくいうと「情報の価値」です。ある情報が得られたとき、それがまれにしか起こらないものだと、その情報の価値は高いと考えられます（いわば「スクープ」に当たりますね）。ということは、その事象が起こる確率 p の小さい方が情報量は大きくなると考えられます。そこで、情報量の尺度として、以下のように確率 p の逆数を取ることにします。p が小さくなると 1/p は大きくなるからです。

$$\frac{1}{p}$$

この対数を取ると以下のようになります。なぜ対数を取るのかは後でお話することとして、これが情報量の定義です。底は **2** です。また、情報量の単位はビットです。

$$\log \frac{1}{p} = -\log p \cdots \text{情報量の定義}$$

例えば、さまざまな身分のメンバー 8 人の中から学生を選びたいとき、学生が 3 人と分かった場合の情報量を求めてみましょう。無作為に選んだ人が学生である確率は 3/8 です。$\log_2 3 = 1.58$ とすると、情報量は以下のように **1.42 ビット**となります。

$$
\begin{aligned}
-\log\frac{3}{8} &= -(\log 3 - \log 8) \\
&= -(1.58 - 3) \\
&= 1.42
\end{aligned}
$$

なお、各々の情報量に確率を掛けた期待値のことを**平均情報量（エントロピー）**と呼びます。例えば、上の例で、学生である確率は 3/8、学生でない確率は 5/8 なので、平均情報量は、log 3/8 = 1.42、log 5/8 = 0.68 を利用して、

$$
\begin{aligned}
\frac{3}{8}(-\log\frac{3}{8}) + \frac{5}{8}(-\log\frac{5}{8}) &= \frac{3\cdot 1.42}{8} + \frac{5\cdot 0.68}{8} \\
&= 0.96
\end{aligned}
$$

となります。全てが等確率であるときには平均情報量は最大（1）になり、情報に偏りがある場合には平均情報量は小さくなります。

さて、対数を取る理由ですが、それはズバリ、掛け算を足し算にして表したいからです。具体例で見てみます。例えば、1 ～ 10 までの数字が書かれた赤玉と白玉があるものとしましょう（全部で **20 個**です）。この中で、どれが当たりであるか、ヒントを教えてもらえるものとしましょう。

- 情報 A：当たりは赤　→ 確率は 10/20 = 1/2、その逆数は **2**
- 情報 B：当たりは 5 番　→　確率は 2/20 = 1/10、その逆数は **10**（こちらの価値が高い）

さらに以下のようにヒントが得られたものとします。

- 情報 C：当たりは赤の 5 番　→　確率は 1/20、その逆数は **20**

これらの情報の価値には以下のような関係が成り立っていることが分かります（2×10 = 20 ですね）。

$$A \times B = C$$

情報量を取り扱うときには、「情報 A ＋情報 B ＝情報 C」のように、掛け算ではなく足し算として表した方がスッキリするので、各項の対数を取ります。すると、

$$\log A + \log B = \log C$$

となります。このように、対数を使えば掛け算を足し算に変えられるというわけです。左辺と右辺をそれぞれ小数点以下第二位まで計算してみましょう。log 10 = 3.32、log 5 = 2.32 です。

$$
\begin{aligned}
\text{左辺:}\quad & \log 2 + \log 10 = 1 + 3.32 \\
& \qquad\qquad\qquad = 4.32 \\[1em]
\text{右辺:}\quad & \log 20 = \log 4 \cdot 5 \\
& \qquad\quad = \log 4 + \log 5 \\
& \qquad\quad = 2 + 2.32 \\
& \qquad\quad = 4.32
\end{aligned}
$$

ちゃんと合っていますね。このことを情報量の加法性と呼びます。

ポイント 3　対数関数の値とグラフ

対数関数は、$f(x) = \log_a x$ のように、対数を使って表された関数です。$f(x) = \log_2 x$ であれば、以下のような値を取ります。

x	1	2	3	4	5	6	7	8	9	10
対数関数 $\log_2 x$	0	1	1.58	2	2.32	2.58	2.81	3	3.17	3.32

表 1　対数関数の値の変化

グラフも描いておきましょう。

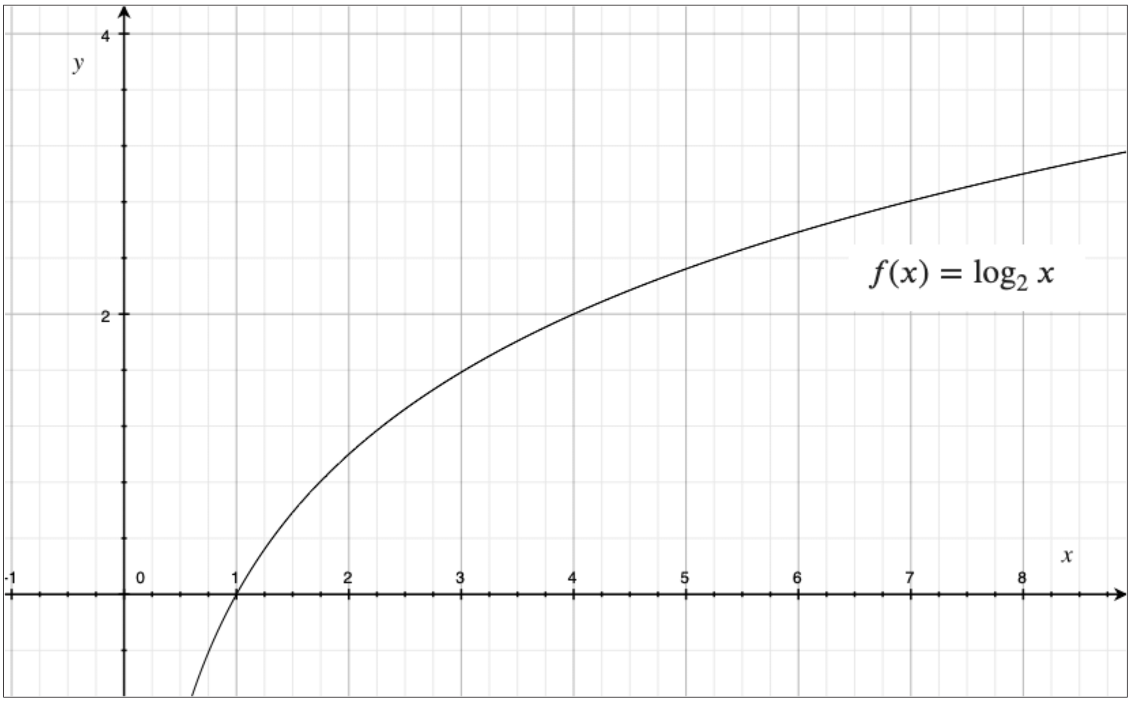


図 2　対数関数のグラフ

対数関数のグラフは指数関数のグラフを y = x に対して対称にした形になる。

→目次に戻る

番外編２でも見たように、指数関数は急激に値が増えていきます。そのため、x が少し大きくなるだけですぐにグラフが上にはみ出してしまいます。そのような場合には対数を取ってグラフ化すると便利です。回帰式が指数関数になるようなデータの場合も、対数を取ると近似曲線が直線になるので、扱いやすくなります。以下に値の変化の様子と、グラフを示しておきます。

x	1	2	3	4	5	6	7	8	9	10
指数関数 2^x	2	4	8	16	32	64	128	256	512	1024
対数を取る	1	2	3	4	5	6	7	8	9	10

表２　指数関数の対数を取る

対数を取るということは、2^x の x を求めるということなので、当然のことながら直線的な関係になる。

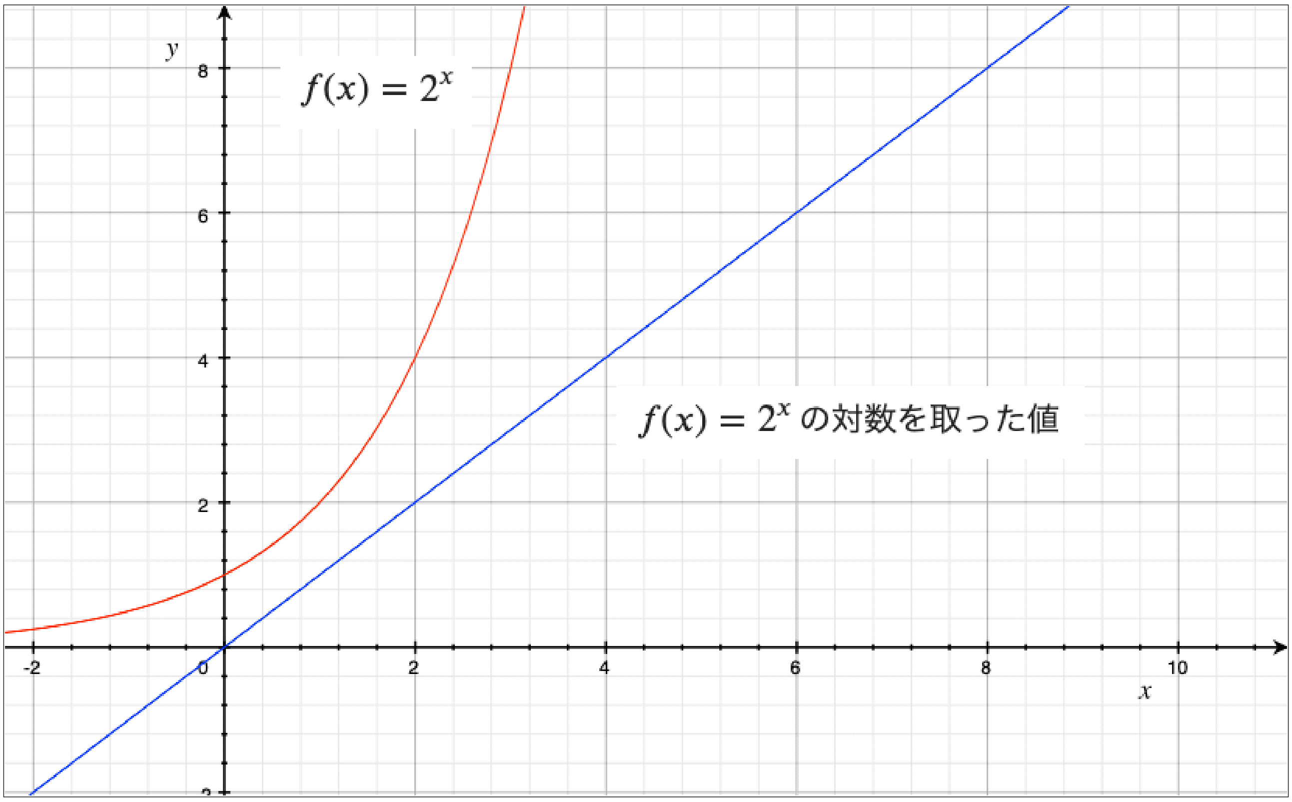


図３　対数関数とその対数を取ったグラフ

指数関数のグラフは急激に値が大きくなって図からはみ出してしまうが、対数を取れば直線になってグラフが見やすくなる。

ポイント 4　対数関数の微分

対数関数の微分法については、公式と計算例を記すにとどめます。計算例には、簡単な例と、公式をフル活用するちょっと難しめのものを含めておきます。

$$f(x) = \log_a x$$

のとき、

$$f'(x) = \frac{1}{x \log_e a} \cdots \boxed{E}$$

となります。

f(x) の底が e のときは、

$$\begin{aligned} f'(x) &= \frac{1}{x \log_e e} \\ &= \frac{1}{x} \cdots \boxed{F} \end{aligned}$$

となります。

(例 7)

$$(\log_2 x)' = \frac{1}{x \log_e 2}$$

これは、公式［E］を適用しただけで求められました。

最後に、多少複雑なものも紹介しておきましょう。本編の第 8 回で見た合成関数の微分を使います。ちょっと難しいので、すぐに理解できなくても問題はありませんが、少しずつでも慣れるようにするといいでしょう。

(例 8)

$$\begin{aligned} (\log_2 (3x+1))' &= \left(\frac{\log_e (3x+1)}{\log_e 2} \right)' \cdots \boxed{1} \\ &= \frac{1}{\log_e 2} \underbrace{(\log(3x+1))'}_{(a)} \cdots \boxed{2} \end{aligned}$$

［1］では、底を e にそろえました。これは公式［D］を使った変形です（下に再掲）。分母の $\log_e 2$ は定数なので、くくりだしたのが［2］の式です。

$$\log_x y = \frac{\log_a y}{\log_a x} \cdots \boxed{D}$$

次に、［2］の（a）の部分で、3x + 1 = u と置き、連鎖律を使うと、

$$(\log_e u)' = \underbrace{\frac{d\log_e u}{du}}_{(b)} \cdot \underbrace{\frac{du}{dx}}_{(c)}$$

となります。（b）の部分は、公式［F］を使って微分できますね。

$$\begin{aligned}\frac{d\log_e u}{du} &= (log_e u)' \\ &= \frac{1}{u} \\ &= \frac{1}{3x+1}\end{aligned}$$

（c）の部分は、(3x + 1)' = 3 となるので、結局、［2］式は、

$$\frac{1}{\log_e 2} \cdot \frac{3}{3x+1} = \frac{3}{(3x+1)\log_e 2} \cdots \boxed{3}$$

となります。このままでも構わないのですが

$$3 = \log_2 8 = \frac{\log_e 8}{\log_e 2} \cdots \boxed{4}$$

であることを利用して、以下のように変形できます。

$$\begin{aligned}\frac{1}{\log_e 2} \cdot \frac{3}{3x+1} &= \frac{3}{(3x+1)\log_e 2} \\ &= \frac{3}{(x \cdot \frac{\log_e 8}{\log_e 2} + 1) \cdot \log_e 2} \\ &= \frac{3}{x\log_e 8 + \log_e 2} \cdots \boxed{5}\end{aligned}$$

［5］では、分配法則を使って、$\log_e 2$ を x・$\log_e 8$ / $\log_e 2$ と **1** に掛けました。スッキリした形になりましたね。

今回は、対数の計算方法と対数関数、指数関数の微分法について整理しました。最後に、これまでに登場した公式の活用や本編で見た合成関数の微分などの計算方法を使いました。

編集：@IT 編集部
発行：アイティメディア株式会社