

AI 入門としての線形代数

TiroDuetto

2023 年 3 月 22 日

概要

表題には AI 入門と銘打ってあるが私が AI について詳しくないので、ほとんど線形代数の基本的な部分を記す。一応計算機による解析も勉強していくつもりであり、このノート内でもできる限り具体例を提示していきたい。内容はまだ書きかけなので随時更新していく。もし重大な誤植や論理の間違いがあれば指摘してほしい。なお、コンピュータとを計算機と呼んでいるのはそういう文化の人間が作った資料であるためであって深い意味はない。ゆえに表現が統一されていないところもあると思う。このノートでは「我々が扱う空間は 3 次元 Euclid 空間であり、これは内積の定義された計量ベクトル空間である」という文章が理解できるところまで進めたら幸いと思う。また、目次の 5 章と 6 章あたりの内容まで理解できれば計算機を用いた機械学習などにも応用できるだろうと期待している。

目次

| | | |
|-------|--------------|----|
| 0 | 数学的記法における注意 | 2 |
| 0.1 | 集合 | 2 |
| 0.1.1 | 数の集合 | 2 |
| 0.1.2 | 直積 | 3 |
| 0.2 | 写像 | 3 |
| 1 | 線形代数とは何か？ | 3 |
| 1.1 | 計算機において | 3 |
| 1.1.1 | 例 | 4 |
| 2 | ベクトル | 5 |
| 2.1 | 基本事項 | 5 |
| 2.1.1 | 演算 | 7 |
| 2.1.2 | 位置ベクトル | 8 |
| 2.2 | 具体例 | 9 |
| 2.2.1 | 色の比較 | 9 |
| 2.2.2 | Bag of Words | 10 |
| 3 | 行列 | 11 |
| 3.1 | 基本事項 | 11 |
| 3.1.1 | 行列 | 11 |
| 3.1.2 | 和とスカラー倍 | 14 |

| | | |
|-------|--------------|----|
| 3.1.3 | 積 | 16 |
| 3.2 | 連立 1 次方程式 | 16 |
| 3.2.1 | 基本変形 | 16 |
| 3.2.2 | 方程式を解く | 16 |
| 3.2.3 | 正則行列・逆行列 | 16 |
| 3.3 | 行列式 | 16 |
| 3.3.1 | 余因子展開 | 16 |
| 3.3.2 | 外積 | 16 |
| 4 | ベクトル空間 | 16 |
| 4.1 | 基本事項 | 17 |
| 4.1.1 | 定義 | 17 |
| 4.1.2 | 空間という言葉 | 17 |
| 4.2 | 1 次結合と 1 次独立 | 17 |
| 4.3 | 基底と次元 | 17 |
| 4.3.1 | 基底の変換 | 17 |
| 5 | 線形写像 | 17 |
| 5.1 | 写像 | 17 |
| 5.2 | 基本事項 | 17 |
| 5.3 | 表現行列 | 17 |
| 6 | 対角化 | 17 |
| 6.1 | 固有値と固有ベクトル | 17 |
| 6.1.1 | 線形変換において | 17 |
| 6.1.2 | 一般の変換において | 17 |
| 6.2 | 対角化 | 17 |

0 数学的記法における注意

0.1 集合

対象物の集まりを集合という。その集合を構成する対象物のことを元や要素という。集合 A に元 a が属することを

$$a \in A \quad (0.1)$$

と書く。またすべてが A の元からなる集合 B のことを、 A の部分集合といい、

$$B \subset A \quad (0.2)$$

と書く。

0.1.1 数の集合

ある数全体の集合には特別な記号が用意されている。複素数 \mathbb{C} , 実数 \mathbb{R} , 有理数 \mathbb{Q} , 整数 \mathbb{Z} , 自然数 \mathbb{N} など。なお、この書体は黒板太字と呼ばれ、通常出版物において単なる太字である字を黒板で書くときに使われるが、以

上の例のように特別な意味を持たせることもある。もちろん出版物においては通常の太字で書いても問題はない。例えば実数全体の集合を \mathbf{R} のように表している書籍もある。

0.1.2 直積

$c \in C, d \in D$ を用いた (c, d) 全体の集合のことを C と D の直積といい、 $C \times D$ と表す。例えば、座標平面上の点 (x, y) は $\mathbb{R} \times \mathbb{R}$ の元であると言える。なお、 $\mathbb{R} \times \mathbb{R} = \mathbb{R}^2$ とも書く。

0.2 写像

2つの集合があるとき、一方の集合の元に対して、もう一方の集合の元がただひとつに定まるような関係付けの規則を写像という。集合 V から集合 W への写像 f を

$$f : V \rightarrow W \quad (0.3)$$

のように書く。あるいは元どうしの関係で $v \in V, w \in W$ として

$$v \mapsto w = f(v) \quad (0.4)$$

と書く。 $v \mapsto f(v)$ と書いてもよい。大雑把に言えば v を入力したら f という規則によって w が出力されるということである。特に V や W が数の集合であれば、関数と呼ばれる。

■例

実数から実数 関数 $y = f(x)$ を以上の記法で表すと

$$f : \mathbb{R} \rightarrow \mathbb{R}, \quad x \mapsto y = f(x) \quad (0.5)$$

実数から複素数 虚数単位 $i = \sqrt{-1}$ を用いて、 $(x, y) \in \mathbb{R}^2$ から複素数 $z \in \mathbb{C}$ を $z = g(x, y) = x + iy$ のように定める。

$$g : \mathbb{R}^2 \rightarrow \mathbb{C}, \quad (x, y) \mapsto z = x + iy \quad (0.6)$$

1 線形代数とは何か？

世の中を大まかに把握したいとき、その現象を”真っ直ぐなもの”や”平らなもの”とみなしたり、単なる比例関係として扱うと便利だ。例えば球体である地上に暮らしていても自分の近傍を平面とみなしたり、アルバイトによる収入も「毎月この程度」と簡単に近似して計算するだろう。こういった考え方を「線形的に近似する」などと表現する。記憶にありそうな例で言えば、比例やその発展である1次関数である。つまり、名前の通り「線形的な」対称を研究するのが線形代数というわけである。また、代数という言葉は、簡単に言えば文字式による計算や方程式の研究する分野に用いられる。合わせて「1次関数を研究する」学問と思っていて貰えばよい。物理学においては本質的に非線形な現象であったとしても、そのうちの線形な部分を切り出して現象を理解したり、あるいは理論を構築していく。

1.1 計算機において

もともとはそういった動機から活かされていたのであるが、線形代数は奥が深く、応用範囲の広い分野である。その重要な対称がベクトルや行列である。今回の本題である計算機による解析では、いくつかのデータをベク

トルとして扱い、また新たなデータへの変換やその関係を行列で表す。 n 個のデータを n 成分ベクトルで次のように表すとしよう。^{*1}

$$\vec{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \quad (1.1)$$

このベクトルを m 個のデータを表すベクトル \vec{u} に変換すると考える。

$$\vec{v} = \begin{pmatrix} u_1 \\ \vdots \\ u_m \end{pmatrix} \quad (1.2)$$

単純計算でこの 2 つをつなげるためには nm 個の成分を持つ変換が必要であるとわかるだろう。この変換を A と書けば

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \quad (1.3)$$

これを用いて

$$\vec{u} = A \vec{v} \quad (1.4)$$

$$\Leftrightarrow \begin{pmatrix} u_1 \\ \vdots \\ u_m \end{pmatrix} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \quad (1.5)$$

このようにベクトル同士規則の変換を表すのが行列である。そのまま $m \times n$ 行列と呼ばれる。この m や n の順序については後ほど説明する。多数のデータを一度に扱いたい……それをきちんと定式化できたら……という願望を叶えてくれるのがこの計算である。すでに線形代数で研究されつくされているのだ。

1.1.1 例

さて、以上の話は一般論に寄りすぎているし、まだベクトルも行列の話もしていない中では何をしているのか？という疑問のほうが大きいと思われるので簡単な例示をしていく。なお [3] の例を少し改造しただけのものである。いま、スーパーマーケットにいる。桃、みかん、りんごの 1 個あたりの値段・重さの表 1 を知ったとき、果物を買った個数と合計の値段・重さの関係はどのように計算するのだろうか。

表 1 重さについては平均的な重さと思っておく。(例なのでそこまで重要でもないが)

| | 桃 | みかん | りんご |
|--------|-----|-----|-----|
| 値段 (円) | 300 | 100 | 200 |
| 重さ (g) | 200 | 30 | 300 |

それぞれの果物を 1 個、2 個、5 個買ったとする。もちろんこのときの計算は簡単で

$$(\text{値段}) = 300 \times 1 + 100 \times 2 + 200 \times 5 = 1500 \quad (1.6)$$

$$(\text{重さ}) = 200 \times 1 + 30 \times 2 + 300 \times 5 = 1760 \quad (1.7)$$

^{*1} わかりやすさのために矢印でベクトルを表しているが、 \vec{v} 以外にも太字の \mathbf{v} やあるいは先にベクトルであると宣言した上で通常の v のまま表すこともある。物理をやっている人はこのあたりにこだわりがあるらしく、3 次元は太字や矢印、4 次元以上は特に明示せず、量子力学の文脈では $|v\rangle$ などと書く。

とできる。これがベクトルと行列を用いれば一つの形式によって表せて便利だという話である。つまり、

$$\begin{pmatrix} 1500 \\ 1760 \end{pmatrix} = \begin{pmatrix} 300 & 100 & 200 \\ 200 & 30 & 300 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 5 \end{pmatrix} \quad (1.8)$$

と表記すればよい。これも

$$\vec{x} = \begin{pmatrix} 1 \\ 2 \\ 5 \end{pmatrix}, \quad A = \begin{pmatrix} 300 & 100 & 200 \\ 200 & 30 & 300 \end{pmatrix}, \quad \vec{y} = \begin{pmatrix} 1500 \\ 1760 \end{pmatrix} \quad (1.9)$$

と書けば

$$\vec{y} = A\vec{x} \quad (1.10)$$

の形で書ける。どう見ても1次関数そっくりである。もちろん、この話はやみくもに数字を並べただけでなく、きちんと線形代数の知識に則っている。勘のいい人は気づいたかもしれないが、結局のところ計算機はこのような計算を繰り返しているということである。

2 ベクトル

まず基本的な注意としてベクトルには列ベクトル・行ベクトルがある。同じ v_1, \dots, v_n を持つベクトルであったとしても

$$\text{列ベクトル } \vec{a} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}, \quad \text{行ベクトル } \vec{b} = (v_1 \quad \cdots \quad v_n) \quad (2.1)$$

の2つは区別される。この関係は $\vec{a} = \vec{b}^T$ と書ける。以下では全て列ベクトルで記すことにする。なお、高校数学ではベクトルを (v_1, v_2, \dots) のように書いていたがここでの行ベクトルにはカンマ, を入れないことにする。^{*2}

列ベクトルと行ベクトル違いについて、詳しくは行列を学ぶとわかる。

2.1 基本事項

ベクトルとは大きさと向きを持った量である。例えば、点 O から点 A へ向かうベクトルは

$$\overrightarrow{OA} \quad (2.2)$$

とでも書けばよい。決まった表し方があるわけではないのでわかりやすいように書くことをおすすめする。ではさらに点 A から点 B に向かうベクトル \overrightarrow{AB} を”加える”ことを考えてみよう。最終的には点 O を出発して点 B に向かうベクトルとなることがわかるだろう。これがベクトルの和である。

$$\overrightarrow{OB} = \overrightarrow{OA} + \overrightarrow{AB} \quad (2.3)$$

のように表せる。詳しくは後に述べる。

^{*2} ただし、「カンマは書いてはいけない」などと特に決まっているわけではない。人によっては普通に使っていたりする。

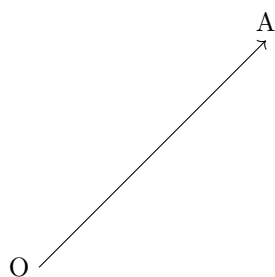


図1 ベクトル \vec{OA}

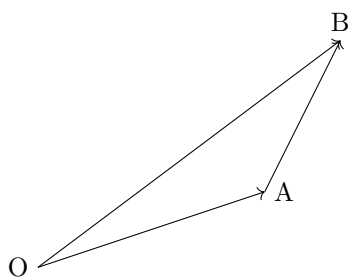


図2 ベクトル $\vec{OB} = \vec{OA} + \vec{AB}$

また、ここでは特に数を並べたものとも考えることも多い。この2つの表現は当然同じである。 n 成分をもつベクトル \vec{v} は

$$\vec{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \quad (2.4)$$

のように表される。前者の表現のほうが以下に用意する言葉のイメージは付きやすいかもしれない。向きを持たない量はスカラーという。ベクトル \vec{v} の大きさを絶対値あるいはノルムと呼び

$$|\vec{v}|, \text{ あるいは単に細字で } v \quad (2.5)$$

と書く。 $\|\vec{v}\|$ と2本線で囲むことも多い。数学書ではだいたいこう書かれる。大きさは向きを持たないのでスカラーである。大きさが同じで、向きが全く逆方向のベクトルを逆ベクトルと呼び

$$-\vec{v} \quad (2.6)$$

と負号を添えて書く。2つのベクトル \vec{v}, \vec{w} の大きさと向きがどちらも同じとき、これらのベクトルは等しいといって

$$\vec{v} = \vec{w} \quad (2.7)$$

と表す。これは両方のベクトルの各成分が等しいことを意味する。構成する数が同じであったとしても順番が違えば違うベクトルである。

$$\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \neq \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix} \quad (2.8)$$

大きさが1のベクトルを単位ベクトルと呼ぶ。例えば \vec{v} の方向の単位ベクトルは

$$e_{\vec{v}} = \frac{\vec{v}}{|\vec{v}|} \quad (2.9)$$

と表せる。このことは方向ベクトルとも呼ぶ。なお、方向を指定しないで単位ベクトルと言っても意味がないので注意。大きさが0のベクトル、すなわち全ての成分が0であるベクトルはゼロベクトルと呼ばれ、 $\vec{0}$ で表される。単に0と表すこともある。

$$\vec{0} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad (2.10)$$

また、空間中の点も数を並べて表現されていたので、この点を位置ベクトルという言葉で表せる。例えば 3 次元空間上の点について

$$(x, y, z) \leftrightarrow \begin{pmatrix} x \\ y \\ z \end{pmatrix} \quad (2.11)$$

で表すことができる。

2.1.1 演算

スカラー倍と 2 つのベクトルの和、内積を定義する。商は定義されない。成分の数が違うときは和や積を定義できないことに注意。なお、単に積ではなく内積という言葉を用いたのは外積と呼ばれる積もあるためである。外積はあとでやるかも。

スカラー倍 あるベクトル \vec{v} をスカラー a 倍したとき、各成分が a 倍される。

$$a\vec{v} = \begin{pmatrix} av_1 \\ \vdots \\ av_n \end{pmatrix} \quad (2.12)$$

図的な表現をすれば $a > 0$ ならば向きは同じで大きさが $|a|$ 倍され、 $a < 0$ ならば向きは反対で大きさが $|a|$ 倍されたベクトルとなる。

和 2 つのベクトルの和は各成分ごとに和を取るものとする。例えば

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} + \begin{pmatrix} d \\ e \\ f \end{pmatrix} = \begin{pmatrix} a+d \\ b+e \\ c+f \end{pmatrix} \quad (2.13)$$

である。先に少し述べたような図の描像では、始点から途中の点を介して終点まで向かうように描かれ、最終的には単に始点から終点へ向かうベクトルとなる。なお、差については (-1) 倍して足したと考えるとよい。

$$-\vec{v} = (-1)\vec{v} \quad (2.14)$$

先に紹介した逆ベクトルとは足したらゼロベクトルになるベクトルとして定義される。ゼロベクトルは足しても引いてももとのベクトルは変わらない。^{*3}

内積 ベクトル特有の演算である。これは 2 つのベクトルからスカラーを作り出すものである。 \vec{v}, \vec{w} を n 成分のベクトルとすると

$$\vec{v} \cdot \vec{w} = v_1 w_1 + \cdots + v_n w_n = \sum_{i=1}^n v_i w_i \quad (2.15)$$

と定義される。内積を (\vec{v}, \vec{w}) や $\langle \vec{v}, \vec{w} \rangle$ で表すこともある。ベクトルのノルムはこの内積を用いて定義され、

$$|\vec{v}| = \sqrt{\vec{v} \cdot \vec{v}} \quad (2.16)$$

である。また、2 つのベクトルのなす角 θ とすると

$$\vec{v} \cdot \vec{w} = |\vec{v}| |\vec{w}| \cos \theta \quad (2.17)$$

と書ける。この書き方についてはきちんと図的な表し方から納得できることを確認してほしい。

また、以上に関係する諸概念についても紹介する。

■距離 ベクトルどうしの距離を定義したい。

まず距離というものを少し堅苦しい書き方ではあるが定義しておく。読み飛ばしても差し支えはない。集合 X について距離関数

$$d: X \times X \rightarrow \mathbb{R} \quad (2.18)$$

は任意の X の元 x, y, z について以下の条件を満たす。

- (A) $d(x, y) \geq 0$ (非負性)
- (B) $d(x, y) = d(y, x)$ (対称律)
- (C) $x = y$ のときのみ $d(x, y) = 0$
- (D) $d(x, y) \leq d(x, z) + d(z, y)$ (三角不等式^{*4})

ここでは、ベクトルどうしの距離の定義として **Euclid 距離**を用いる。すなわち、 \vec{v}, \vec{w} に対して

$$d(\vec{v}, \vec{w}) = \sqrt{|v_1 - w_1|^2 + \cdots + |v_n - w_n|^2} \quad (2.19)$$

位置ベクトルの考え方を思い出せば、これは単に 2 点間の距離を表しているにすぎないことがわかる。三平方の定理の n 次元への一般化である。このように定義した距離が上の条件 1. ~ 4. を満たすことを確認しておこう。

■コサイン類似度 2 つのベクトルがどれだけ”近い”かを図る尺度としてコサイン類似度を定義できる。ベクトル \vec{v}, \vec{w} に対してコサイン類似度 $\text{sim}(\vec{v}, \vec{w})$ を

$$\text{sim}(\vec{v}, \vec{w}) = \frac{\vec{v} \cdot \vec{w}}{|\vec{v}| |\vec{w}|} = \cos \theta \quad (2.20)$$

によって定める。これは $-1 \leq \text{sim}(\vec{v}, \vec{w}) \leq 1$ で”近さ”を判定するものである。例えば $\text{sim}(\vec{v}, \vec{w}) = 1$ であれば $\cos \theta = 1$ 、すなわち $\theta = 0$ なので全く同じ方向を向いていて、コサイン類似度の観点からは最も”近い”といえる。反対に、 $\text{sim}(\vec{v}, \vec{w}) = -1$ であれば $\cos \theta = -1$ 、すなわち $\theta = \pi$ であるから、全く逆方向に向いているベクトルである。このとき最も”遠い”と思えるだろう。なお、定義から分かる通り、コサイン類似度はベクトルの大きさは関係なく、角度で近さを測っている。

2.1.2 位置ベクトル

ベクトルと座標の関係をもう少し詳しく説明する。もともとベクトル自体は座標系によらない量^{*5}であるが、座標を指定すると計算において便利である。

xy 平面上で図 3 のように $A(a, b)$ という点があったとする。この数字の意味するところはもちろん、「 x 軸方向に a だけ進み、 y 軸方向に b だけ進んだところ」と解釈する。そして、ベクトル \vec{OA} を位置ベクトルとして、点 A を表すものとする。

$$A(a, b) \leftrightarrow \vec{OA} = \begin{pmatrix} a \\ b \end{pmatrix} \quad (2.21)$$

「 x 軸方向に a だけ進み、 y 軸方向に b だけ進んだところ」を上のように \vec{OA} で表せるが、この文言に従えば、もう少し違った表し方があるようにも思える。「 x 軸、 y 軸方向にそれぞれ基本となるベクトルがあって、それらを

^{*4} 劣加法性とも言う。

^{*5} 直交座標を使おうが極座標を使おうがその矢印そのものは全く変わらずそこに”在る”ため。

何倍かして足し合わせたもの」と表したらどうだろうか。

$$\vec{e}_x = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \vec{e}_y = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad (2.22)$$

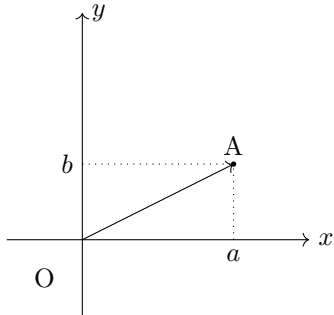


図3 点 $A(a, b)$

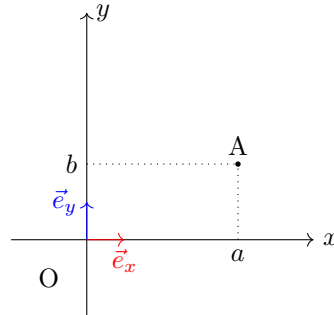


図4 点 $A(a, b)$

この基本となるベクトルのことをそのまま**基本ベクトル**、または**基底ベクトル**と呼ぶ。線形代数の文脈では基底ベクトルと呼ぶことのほうが多い。

$$\overrightarrow{OA} = \begin{pmatrix} a \\ b \end{pmatrix} = \underbrace{a \begin{pmatrix} 1 \\ 0 \end{pmatrix}}_{x \text{ 方向}} + \underbrace{b \begin{pmatrix} 0 \\ 1 \end{pmatrix}}_{y \text{ 方向}} \quad (2.23)$$

基底ベクトルはよく \vec{e} で表される。詳しいことは後の項で記すが、この基底ベクトルのように

$$c_x \vec{e}_x + c_y \vec{e}_y = 0 \quad (2.24)$$

を満たすような係数 c が $c_x = c_y = 0$ のみであるとき、**1 次独立**であるという。また、このようないくつかのベクトルをスカラー倍して和を取ったものを**1 次結合**という。この例から、2次元平面上では1次独立なベクトルが2つあればそれらの1次結合によってどんなベクトルも表すことができるということが理解できるとよい。

2.2 具体例

ここまでの内容を用いてできる具体例を記す。ここでは、距離とコサイン類似度を用いた類似性の判定についてである。

2.2.1 色の比較

もっとも簡単に色を表す方法は **RGB** だろう。これは赤、緑、青の三原色を用いて一つの色を表す方法である。よって、色は **3 成分**のベクトルで表されるものとなる。上での例では成分は2つだったが、3つに増えても同じである。ここでは各成分が0から255までの値で表されるものとする。

ここでの問題設定は「オレンジ色は赤色、黄色、青色のどの色に最も近いか」としよう。オレンジ色、赤色、黄色、青色のベクトルをそれぞれ $\vec{o}, \vec{r}, \vec{y}, \vec{b}$ とおくと、次のように表せる。

$$\vec{o} = \begin{pmatrix} 255 \\ 165 \\ 0 \end{pmatrix}, \quad \vec{r} = \begin{pmatrix} 255 \\ 0 \\ 0 \end{pmatrix}, \quad \vec{y} = \begin{pmatrix} 255 \\ 255 \\ 0 \end{pmatrix}, \quad \vec{b} = \begin{pmatrix} 0 \\ 0 \\ 255 \end{pmatrix}, \quad (2.25)$$

まず距離を使って比較してみる。それぞれ (2.19) を用いて計算すると

$$d(\vec{o}, \vec{r}) = \sqrt{|255 - 255|^2 + |165 - 0|^2 + |0 - 0|^2} = 165 \quad (2.26)$$

$$d(\vec{o}, \vec{y}) = \sqrt{|255 - 255|^2 + |165 - 255|^2 + |0 - 0|^2} = 90 \quad (2.27)$$

$$d(\vec{o}, \vec{b}) = \sqrt{|255 - 0|^2 + |165 - 0|^2 + |0 - 255|^2} \approx 396.6 \quad (2.28)$$

となる。これはそのままどれだけ離れているかを表しているの、オレンジ色には黄色、赤色、青色の順で似ているといえる。

次にコサイン類似度を用いてみる。角度で近さを考えるので、ベクトルのなす角が 0、すなわち値が 1 に近いほど似ていて、なす角が π 、すなわち値が -1 に近いほど似ていないということになる。(2.20) を用いてやれば

$$\text{sim}(\vec{o}, \vec{r}) = \frac{255 \times 255 + 165 \times 0 + 0 \times 0}{\sqrt{|255|^2 + |165|^2 + |0|^2} \sqrt{|255|^2 + |0|^2 + |0|^2}} \approx 0.84 \quad (2.29)$$

$$\text{sim}(\vec{o}, \vec{y}) = \frac{255 \times 255 + 165 \times 255 + 0 \times 0}{\sqrt{|255|^2 + |165|^2 + |0|^2} \sqrt{|255|^2 + |255|^2 + |0|^2}} \approx 0.98 \quad (2.30)$$

$$\text{sim}(\vec{o}, \vec{b}) = \frac{255 \times 0 + 165 \times 0 + 0 \times 255}{\sqrt{|255|^2 + |165|^2 + |0|^2} \sqrt{|0|^2 + |0|^2 + |255|^2}} = 0 \quad (2.31)$$

これもまた、黄色、赤色、青色の順で似ているという結果になった。

これを応用すれば、画像データから色を抽出してベクトルにして画像同士の似ている・似ていないを定量的に計算できるだろう*6。

2.2.2 Bag of Words

コサイン類似度は文章の近さについても使える。文章をベクトルにするとき **Bag of Words** という手法を使う。例えば以下の 3 つの文章を考える。

(A) : 「私はラーメンと餃子が好きです」

(B) : 「私は餃子が嫌いです」

(C) : 「私はラーメンが好きです」

上の文章に出てくる自立語は「私」「ラーメン」「餃子」「好き」「嫌い」である。つまり、これらの自立語の有無によって文章を特徴づける。この”特徴づけ”をベクトルにするため、5 成分のベクトルとして、順にその言葉が出たら 1 を、出なければ 0 をその値とする。文章のベクトルをそれぞれ $\vec{A}, \vec{B}, \vec{C}$ で表せば

$$\vec{A} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \quad \vec{B} = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \quad \vec{C} = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} \quad (2.32)$$

と書ける。このように用意すれば sim を用いて近さを測ることができる。この方法でベクトルにしたとき、距離の使用は類似性の判定に向かない。なぜならば文章が長いほど多くの成分が 1 になり、また短ければ多くの成分が 0 となるためである。すなわち、長い文章と短い文章を比較したときには内容にかかわらず距離が大きくなってしまふのである。

*6 やってないけど。

3 行列

新たに出てくる言葉や概念が多いので注意深く見ていく。

3.1 基本事項

行列について、ゼロ行列、正方行列、単位行列、対角行列、スカラー行列、上三角行列、下三角行列、転置、対称行列、反対称行列 (交代行列)。行ベクトル、列ベクトル。添字。

3.1.1 行列

横と縦に長方形に数を並べたものを行列という。並べられた数をそれぞれ成分といい、数は $()$ や $[]$ で囲む。横に並んだ成分を行、縦に並んだ成分を列という。例えば 3 行 4 列の行列であるとすれば成分を a_{ij} ($i = 1, 2, 3, j = 1, 2, 3, 4$) のように表して

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{pmatrix}, \quad \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix} \quad (3.1)$$

である。一般に m 行 n 列である行列は $m \times n$ 行列や (m, n) 型の行列などという。この行と列の組 (m, n) を行列の型やサイズという。

$$A = \begin{pmatrix} a_{11} & \cdots & a_{12} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{i1} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & & \vdots & & \vdots \\ a_{m1} & \cdots & a_{mj} & \cdots & a_{mn} \end{pmatrix} \quad (3.2)$$

このうち

$$a_{ij}, \quad (a_{i1} \quad a_{i2} \quad \cdots \quad a_{in}), \quad \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix} \quad (3.3)$$

の部分それぞれ行列 A の (i, j) 成分、第 i 行、第 j 列という。上の表し方で a_{ij} の ij を添字という。^{*7}この添字は必ず行・列の順になっていることに注意。 $m \times n$ 行列 A の成分を a_{ij} と書いたとき

$$A = (a_{ij})_{m \times n} \quad (3.4)$$

と略記することもある。なお、行列の型が明らかなきは $m \times n$ を省略することもある。以下にいくつか行列に関する言葉を用意する。多いので少しずつ覚えてほしい。また、添字の記法を多用しているのでこれも慣れていくことをおすすめする。

1×1 行列 (a) のように表されるがこの場合普通の数 a と同一視して単に a と表す。

^{*7} 単に変数や文字を簡潔に表す記法である。しかしこの添字記法は計算を便利にし、Einstein 規約によって様々な計算が一気に簡単になる。なお、下付きの添字以外に上付き添字を使うこともあるので添字の記法を見たときは十分に注意すること。

行列の相等 2つの行列 A と B が等しいとは、それぞれの型が同じで、対応する成分がそれぞれすべて等しいことを意味する。すなわち、 $A = (a_{ij})_{m \times n}$ と $B = (b_{kl})_{p \times q}$ において $m = p$ かつ $n = q$ で、任意の i, j において $a_{ij} = b_{ij}$ が成り立つとき、

$$A = B \quad (3.5)$$

である。等しくないときは

$$A \neq B \quad (3.6)$$

と書かれる。もちろんベクトルと同じようにどれか一つの成分でも違えば $A \neq B$ である。

ゼロ行列 すべての成分が 0 である $m \times n$ 行列をゼロ行列といい、 $O_{m,n}$ や O と書く。例えば 3×2 のゼロ行列は

$$\begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \quad (3.7)$$

のようになる。

正方行列 行列 A の行と列の数が同じとき、すなわち $n \times n$ 行列のとき A を n 次の正方行列、または n 次行列という。 $A = (a_{ij})_{n \times n}$ のうち成分 $a_{11}, a_{22}, \dots, a_{nn}$ を A の対角成分という。例えば 2 次正方行列は

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad (3.8)$$

のように書いて、対角成分は a, d である。文字通り対角線の成分のことである。

対角行列 正方行列 A が対角成分以外を持たない、すなわち $i \neq j$ のとき

$$a_{ij} = 0 \quad (3.9)$$

である行列を対角行列という。

$$\begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix} \quad (3.10)$$

スカラー行列 対角行列のうち、すべての対角成分が等しい行列、すなわち

$$a_{ij} = \begin{cases} 0 & (i \neq j) \\ a & (i = j) \end{cases} \quad (3.11)$$

のときスカラー行列という。

$$\begin{pmatrix} a & 0 & \cdots & 0 \\ 0 & a & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a \end{pmatrix} \quad (3.12)$$

単位行列 スカラー行列のうちすべての対角成分が 1 である、すなわち

$$a_{ij} = \begin{cases} 0 & (i \neq j) \\ 1 & (i = j) \end{cases} \quad (3.13)$$

のとき単位行列といい、 $n \times n$ 行列であれば n 次単位行列といい I_n や E_n 、あるいは単に I や E と書く。

$$I_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} \quad (3.14)$$

なお、以下のように $i = j$ のときは 1、それ以外は 0 となる記号として **Kronecker** のデルタを定義すると便利である。

$$\delta_{ij} = \begin{cases} 0 & (i \neq j) \\ 1 & (i = j) \end{cases} \quad (3.15)$$

これを用いればスカラー行列や単位行列をそれぞれ

$$A = (a\delta_{ij})_{n \times n}, \quad I_n = (\delta_{ij})_{n \times n} \quad (3.16)$$

と表せる。

三角行列 正方行列 $A = (a_{ij})_{n \times n}$ について

$$a_{ij} = 0 \quad (i > j) \quad (3.17)$$

のとき上三角行列といい、

$$a_{ij} = 0 \quad (i < j) \quad (3.18)$$

のとき下三角行列という。3 次行列でそれぞれの例を書けば、

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 0 \\ 2 & 4 & 0 \\ 3 & 5 & 6 \end{pmatrix} \quad (3.19)$$

転置行列 $m \times n$ 行列 A の行と列を入れ替えて得られる $n \times m$ 行列を A の転置行列といい、 tA とかく。^{*8}

例えば

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \xrightarrow{\text{転置!}} \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix} \quad (3.20)$$

転置行列を求めることを、「転置を取る」と表現する。

対称行列 もとの行列と転置行列が一致するとき、その行列を対称行列という。例えば

$$(1) \quad \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} \quad \begin{pmatrix} 1 & 2 & 3 \\ 2 & 5 & 6 \\ 3 & 6 & 0 \end{pmatrix} \quad (3.21)$$

なお、 n 次正方行列 $A = (a_{ij})$ が対称行列であるときの必要十分条件は

$$a_{ij} = a_{ji} \quad (3.22)$$

である。

以上で一氣に言葉を定義したので、おそらくすぐに慣れることは難しいだろう。例題を少し用意したので少し手を動かしてやってみてほしい。

例 3.1 次の行列の型、 $(1, 2)$ 成分、第 2 行、第 3 列をそれぞれ答えよ。

$$\begin{pmatrix} 1 & 2 & 4 \\ 7 & 3 & 6 \end{pmatrix} \quad (3.23)$$

例 3.2 次の正方行列の型を答えよ。また、対角成分をすべて書き出し、足しあわせよ。

$$\begin{pmatrix} 2 & 6 & 9 & 8 \\ 0 & 2 & 7 & 0 \\ 3 & 6 & 9 & 9 \\ 8 & 1 & 7 & 0 \end{pmatrix} \quad (3.24)$$

例 3.3 2 次単位行列を具体的に書いてみよ。

例 3.4 (i, j) 成分が $2\delta_{ij}$ で表される 3 次行列を具体的に書き表せ。

例 3.5 (i, j) 成分が $2\delta_{i,j+1}$ で表される 3 次行列を具体的に書き表せ。

例 3.6 (3.24) の転置行列を書け。

例 3.7 実際に (3.21) の転置を取って元の行列と一致するか確かめよ。

なお、 $1 \times n$ 行列を n 次の行ベクトル、 $m \times 1$ 行列を m 次の列ベクトルといい、行ベクトルと列ベクトルをあわせて数ベクトルという。例えば 3 次の行ベクトルは

$$(a \quad b \quad c) \quad (3.25)$$

3 次の列ベクトルは

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} \quad (3.26)$$

などのようになる。

3.1.2 和とスカラー倍

さて、行列に関する演算を定義していく。まず和について考える。ともに $n \times m$ 行列である A と B に対して和を、各成分同士を足すことと定義する。型が違う行列については演算が定義できないことに注意せよ。添字を用いて書き表わせば

$$A = (a_{ij})_{n \times m}, \quad B = (b_{ij})_{n \times m} \quad (3.27)$$

に対して和 $A + B$ を

$$(A + B)_{n \times m} = (a_{ij} + b_{ij})_{n \times m} \quad (3.28)$$

と定義する。例えば 2×3 行列の和は

$$\begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix} + \begin{pmatrix} g & h & i \\ j & k & l \end{pmatrix} = \begin{pmatrix} a+g & b+h & c+i \\ d+j & e+k & f+l \end{pmatrix} \quad (3.29)$$

と書かれる。

次にスカラー倍を定義する。スカラー、すなわち大きさのみを持つ数^{*9} c に対して、 $n \times m$ 行列 A の c によるスカラー倍とは、すべての成分を c 倍するものと定義する。ここでも添字を用いて表せば

$$cA = (ca_{ij})_{n \times m} \quad (3.30)$$

^{*9} 単に実数や複素数と思えばよい。

である。例えば次の 2×3 行列を c 倍すると

$$c \begin{pmatrix} 1 & 2 & 4 \\ 7 & 3 & 6 \end{pmatrix} = \begin{pmatrix} c & 2c & 4c \\ 7c & 3c & 6c \end{pmatrix} \quad (3.31)$$

となる。例題として次の計算をせよ。

例 3.8

$$\begin{pmatrix} 1 & 2 & 4 \\ 7 & 3 & 6 \end{pmatrix} + \begin{pmatrix} 3 & 2 & 1 \\ 5 & 6 & 7 \end{pmatrix} \quad (3.32)$$

例 3.9

$$3 \begin{pmatrix} 2 & 6 & 9 \\ 0 & 2 & 7 \\ 3 & 6 & 9 \end{pmatrix} \quad (3.33)$$

例 3.10

$$\begin{pmatrix} 1 & 2 \\ 2 & -3 \end{pmatrix} + 2 \begin{pmatrix} 3 & 1 \\ -1 & 1 \end{pmatrix} \quad (3.34)$$

ここで (-1) 倍について述べておく。便利な記法として

$$(-1)A = -A \quad (3.35)$$

と書く。また、

$$A + (-B) = A - B \quad (3.36)$$

と書くことに注意せよ。和とスカラー倍について以下の定理が成り立つ。これは後に扱うベクトル空間が満たすべき基本的性質である。

定理 3.1 (和とスカラー倍). $m \times n$ 行列 A, B, C またゼロ行列 $O_{m,n}$ とスカラー c, d に対して、以下の (A) ~ (H) が成り立つ。

- (A) $A + B = B + A$ (和の交換律)
- (B) $(A + B) + C = A + (B + C)$ (和の結合律)
- (C) $A + O_{m,n} = O_{m,n} + A = A$
- (D) $c(dA) = (cd)A$ (スカラー倍の結合律)
- (E) $(c + d)A = cA + dA$ (分配律 I)
- (F) $c(A + B) = cA + cB$ (分配律 II)
- (G) $1A = A$
- (H) $0A = O_{m,n}$

なお和の結合律により $(A + B) + C = A + (B + C) = A + B + C$ と書いてよい。ここで対称行列と対になる概念として交代行列を定義する。

交代行列 ${}^t A = -A$ となる正方行列 A を交代行列、あるいは反対称行列という。

例えば

$$\begin{pmatrix} 0 & 2 \\ -2 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 & 2 \\ -1 & 0 & 3 \\ -2 & -3 & 0 \end{pmatrix} \quad (3.37)$$

などである。

定理 3.2. n 次正方行列 $A = (a_{ij})_{n \times n}$ が交代行列であるための必要十分条件は

$$a_{ij} = -a_{ji} \quad (i, j = 1, 2, \dots, n) \quad (3.38)$$

であり、交代行列の対角成分はすべて 0 である。

この主張を示していく。(3.38) は定義式 ${}^t A = -A$ を成分で表したものであるため、必要十分条件である。次に対角成分がすべて 0 であることを示す。対角成分は $i = j$ であるため、(3.38) と合わせて

$$a_{ii} = -a_{ii} \quad (3.39)$$

が成り立つ。よって

$$2a_{ii} = 0 \quad \therefore a_{ii} = 0 \quad (3.40)$$

である。

3.1.3 積

3.2 連立 1 次方程式

3.2.1 基本変形

3.2.2 方程式を解く

3.2.3 正則行列・逆行列

3.3 行列式

3.3.1 余因子展開

3.3.2 外積

4 ベクトル空間

計算機で使うかは知らないけど、線形代数および数学を学ぶ以上このあたりの考えは避けて通れないと思う。このあたりから抽象的な言葉が増えてしんどくなってくる。

4.1 基本事項

4.1.1 定義

4.1.2 空間という言葉

4.2 1 次結合と 1 次独立

4.3 基底と次元

4.3.1 基底の変換

5 線形写像

こっちは AI のモデルでめっちゃくちゃ使うらしい。

5.1 写像

5.2 基本事項

5.3 表現行列

線形写像はベクトル空間の基底を選べば表現行列で表せる。

6 対角化

対角化すると行列のべき乗や指数関数の計算が簡単になる。

6.1 固有値と固有ベクトル

6.1.1 線形変換において

6.1.2 一般の変換において

6.2 対角化

ここまでできればその辺の理系学生より優秀。

参考文献

- [1] 藤岡敦, 手を動かして学ぶ線形代数,[裳華房,2017(第 2 版)]
- [2] 石村園子, 改訂版 すぐわかる線形代数,[東京図書,2012(改訂版)]
- [3] 中西崇文,Python ハンズオンによるはじめての線形代数,[森北出版,2021]