

AI 入門としての線形代数

TiroDuetto

2023 年 6 月 7 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。なお、このレジメでは章立てが a.b.c のようになっているが、を 1.a.b.c のようにした。^{*1}a を章、b を節、c を項と呼んでいる。

^{*1} documentclass を jsarticle から jsreport に変えたため。5/15 変更。

目次

第 1 部	線形代数	3
1.0	数学的記法における注意	3
1.0.1	集合	3
1.0.1.1	数の集合	3
1.0.1.2	直積	3
1.0.2	写像	3
1.1	線形代数とは何か?	4
1.1.1	計算機において	4
1.1.1.1	例	5
1.2	ベクトル	6
1.2.1	基本事項	6
1.2.1.1	演算	7
1.2.1.2	位置ベクトル	9
1.2.2	具体例	10
1.2.2.1	色の比較	10
1.2.2.2	Bag of Words	11
1.3	行列	11
1.3.1	基本事項	11
1.3.1.1	行列	11
1.3.1.2	和とスカラー倍	15
1.3.1.3	積	17
1.3.2	具体例	19
1.3.2.1	拡大・縮小	19
1.3.2.2	回転	20
1.3.3	連立 1 次方程式	21
1.3.3.1	基本変形	22
1.3.3.2	方程式を解く	24
1.3.3.3	正則行列・逆行列	27
1.3.3.4	正則性とその言い換え	28
1.3.4	行列式	29
1.3.4.1	余因子展開	29
1.3.4.2	外積	29

	1.3.4.3 Cayley-Hamilton の定理	29
1.4	ベクトル空間	29
	1.4.1 基本事項	29
	1.4.1.1 定義	29
	1.4.1.2 空間という言葉	29
	1.4.2 1 次結合と 1 次独立	29
	1.4.3 基底と次元	29
	1.4.3.1 基底の変換	29
1.5	線形写像	29
	1.5.1 写像	30
	1.5.2 基本事項	30
	1.5.3 表現行列	30
1.6	対角化	30
	1.6.1 固有値と固有ベクトル	30
	1.6.1.1 線形変換において	30
	1.6.1.2 一般の変換において	30
	1.6.2 対角化	30
付録 A	<u>おまけ: ランダム行列</u>	31
	A.1 基本の話	31
	A.1.1 積分測度	31
	A.1.2 分配関数	33
	A.2 Coulomb 気体の方法	34
付録 B	<u>おまけ: 微分形式</u>	35
参考文献		36

第 1 部

線形代数

1.0 数学的記法における注意

1.0.1 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

$$a \in A \quad (1.0.1)$$

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、

$$B \subset A \quad (1.0.2)$$

と書く。

1.0.1.1 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。

1.0.1.2 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。

1.0.2 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

$$f : V \rightarrow W \quad (1.0.3)$$

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

$$v \mapsto w = f(v) \quad (1.0.4)$$

と書く。 $v \mapsto f(v)$ とも書いてもよい。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

$$f: \mathbb{R} \rightarrow \mathbb{R}, \quad x \mapsto y = f(x) \quad (1.0.5)$$

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。

$$g: \mathbb{R}^2 \rightarrow \mathbb{C}, \quad (x, y) \mapsto z = x + iy \quad (1.0.6)$$

1.1 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。

1.1.1 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}

$$\vec{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \quad (1.1.1)$$

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。

$$\vec{v} = \begin{pmatrix} u_1 \\ \vdots \\ u_m \end{pmatrix} \quad (1.1.2)$$

^{*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。

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

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \quad (1.1.3)$$

これを用いて

$$\vec{u} = A\vec{v} \quad (1.1.4)$$

$$\Leftrightarrow \begin{pmatrix} u_1 \\ \vdots \\ u_m \end{pmatrix} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \quad (1.1.5)$$

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。

1.1.1.1 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 1.1 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 1.1 重さについては平均的な重さとしておく。(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

$$(\text{値段}) = 300 \times 1 + 100 \times 2 + 200 \times 5 = 1500 \quad (1.1.6)$$

$$(\text{重さ}) = 200 \times 1 + 30 \times 2 + 300 \times 5 = 1760 \quad (1.1.7)$$

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、

$$\begin{pmatrix} 1500 \\ 1760 \end{pmatrix} = \begin{pmatrix} 300 & 100 & 200 \\ 200 & 30 & 300 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 5 \end{pmatrix} \quad (1.1.8)$$

と表記すればよい。これも

$$\vec{x} = \begin{pmatrix} 1 \\ 2 \\ 5 \end{pmatrix}, \quad A = \begin{pmatrix} 300 & 100 & 200 \\ 200 & 30 & 300 \end{pmatrix}, \quad \vec{y} = \begin{pmatrix} 1500 \\ 1760 \end{pmatrix} \quad (1.1.9)$$

と書けば

$$\vec{y} = A\vec{x} \quad (1.1.10)$$

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。

1.2 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

$$\text{列ベクトル } \vec{a} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}, \quad \text{行ベクトル } \vec{b} = (v_1 \quad \cdots \quad v_n) \quad (1.2.1)$$

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。

1.2.1 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

$$\vec{OA} \quad (1.2.2)$$

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 \vec{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。

$$\vec{OB} = \vec{OA} + \vec{AB} \quad (1.2.3)$$

のように表せる。詳しくは後に述べる。

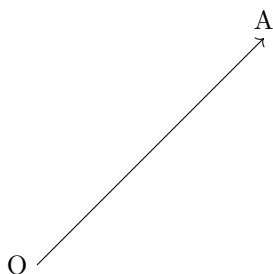


図 1.1 ベクトル \vec{OA}

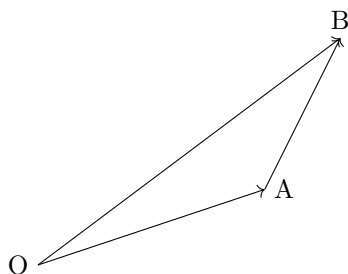


図 1.2 ベクトル $\vec{OB} = \vec{OA} + \vec{AB}$

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

$$\vec{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \quad (1.2.4)$$

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

$$|\vec{v}|, \quad \text{あるいは単に細字で } v \quad (1.2.5)$$

^{*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。

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

$$-\vec{v} \quad (1.2.6)$$

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

$$\vec{v} = \vec{w} \quad (1.2.7)$$

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。

$$\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \neq \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix} \quad (1.2.8)$$

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

$$e_{\vec{v}} = \frac{\vec{v}}{|\vec{v}|} \quad (1.2.9)$$

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。

$$\vec{0} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad (1.2.10)$$

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

$$(x, y, z) \leftrightarrow \begin{pmatrix} x \\ y \\ z \end{pmatrix} \quad (1.2.11)$$

で表すことができる。

1.2.1.1 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。

$$a\vec{v} = \begin{pmatrix} av_1 \\ \vdots \\ av_n \end{pmatrix} \quad (1.2.12)$$

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

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} + \begin{pmatrix} d \\ e \\ f \end{pmatrix} = \begin{pmatrix} a+d \\ b+e \\ c+f \end{pmatrix} \quad (1.2.13)$$

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。

$$-\vec{v} = (-1)\vec{v} \quad (1.2.14)$$

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

$$\vec{v} \cdot \vec{w} = v_1 w_1 + \cdots + v_n w_n = \sum_{i=1}^n v_i w_i \quad (1.2.15)$$

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、

$$|\vec{v}| = \sqrt{\vec{v} \cdot \vec{v}} \quad (1.2.16)$$

である。また、2つのベクトルのなす角 θ とすると

$$\vec{v} \cdot \vec{w} = |\vec{v}| |\vec{w}| \cos \theta \quad (1.2.17)$$

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

$$d: X \times X \rightarrow \mathbb{R} \quad (1.2.18)$$

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

$$d(\vec{v}, \vec{w}) = \sqrt{|v_1 - w_1|^2 + \cdots + |v_n - w_n|^2} \quad (1.2.19)$$

^{*4} 劣加法性とも言う。

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

$$\text{sim}(\vec{v}, \vec{w}) = \frac{\vec{v} \cdot \vec{w}}{|\vec{v}| |\vec{w}|} = \cos \theta \quad (1.2.20)$$

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。

1.2.1.2 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 1.3 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。

$$A(a, b) \leftrightarrow \vec{OA} = \begin{pmatrix} a \\ b \end{pmatrix} \quad (1.2.21)$$

「 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。

$$\vec{e}_x = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \vec{e}_y = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad (1.2.22)$$

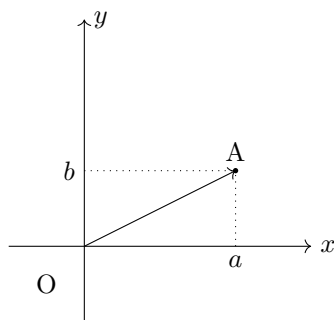


図 1.3 点 $A(a, b)$

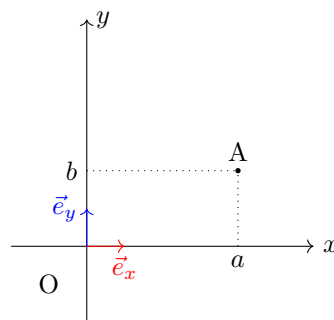


図 1.4 点 $A(a, b)$

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

^{*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。

ベクトルと呼ぶことのほうが多い。

$$\overrightarrow{OA} = \begin{pmatrix} a \\ b \end{pmatrix} = a \underbrace{\begin{pmatrix} 1 \\ 0 \end{pmatrix}}_{x \text{ 方向}} + b \underbrace{\begin{pmatrix} 0 \\ 1 \end{pmatrix}}_{y \text{ 方向}} \quad (1.2.23)$$

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

$$c_x \vec{e}_x + c_y \vec{e}_y = 0 \quad (1.2.24)$$

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。

1.2.2 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。

1.2.2.1 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。

$$\vec{o} = \begin{pmatrix} 255 \\ 165 \\ 0 \end{pmatrix}, \quad \vec{r} = \begin{pmatrix} 255 \\ 0 \\ 0 \end{pmatrix}, \quad \vec{y} = \begin{pmatrix} 255 \\ 255 \\ 0 \end{pmatrix}, \quad \vec{b} = \begin{pmatrix} 0 \\ 0 \\ 255 \end{pmatrix}, \quad (1.2.25)$$

まず距離を使って比較してみる。それぞれ (1.2.19) を用いて計算すると

$$d(\vec{o}, \vec{r}) = \sqrt{|255 - 255|^2 + |165 - 0|^2 + |0 - 0|^2} = 165 \quad (1.2.26)$$

$$d(\vec{o}, \vec{y}) = \sqrt{|255 - 255|^2 + |165 - 255|^2 + |0 - 0|^2} = 90 \quad (1.2.27)$$

$$d(\vec{o}, \vec{b}) = \sqrt{|255 - 0|^2 + |165 - 0|^2 + |0 - 255|^2} \approx 396.6 \quad (1.2.28)$$

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。(1.2.20) を用いてやれば

$$\text{sim}(\vec{o}, \vec{r}) = \frac{255 \times 255 + 165 \times 0 + 0 \times 0}{\sqrt{|255|^2 + |165|^2 + |0|^2} \sqrt{|255|^2 + |0|^2 + |0|^2}} \approx 0.84 \quad (1.2.29)$$

$$\text{sim}(\vec{o}, \vec{y}) = \frac{255 \times 255 + 165 \times 255 + 0 \times 0}{\sqrt{|255|^2 + |165|^2 + |0|^2} \sqrt{|255|^2 + |255|^2 + |0|^2}} \approx 0.98 \quad (1.2.30)$$

$$\text{sim}(\vec{o}, \vec{b}) = \frac{255 \times 0 + 165 \times 0 + 0 \times 255}{\sqrt{|255|^2 + |165|^2 + |0|^2} \sqrt{|0|^2 + |0|^2 + |255|^2}} = 0 \quad (1.2.31)$$

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。

1.2.2.2 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

$$\vec{A} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \quad \vec{B} = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \quad \vec{C} = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} \quad (1.2.32)$$

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。

1.3 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。

1.3.1 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。

1.3.1.1 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

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{pmatrix}, \quad \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix} \quad (1.3.1)$$

*6 やってないけど。

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。

$$A = \begin{pmatrix} a_{11} & \cdots & a_{12} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{i1} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & & \vdots & & \vdots \\ a_{m1} & \cdots & a_{mj} & \cdots & a_{mn} \end{pmatrix} \quad (1.3.2)$$

このうち

$$a_{ij}, \quad (a_{i1} \quad a_{i2} \quad \cdots \quad a_{in}), \quad \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix} \quad (1.3.3)$$

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

$$A = (a_{ij})_{m \times n} \quad (1.3.4)$$

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。

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、

$$A = B \quad (1.3.5)$$

である。等しくないときは

$$A \neq B \quad (1.3.6)$$

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

$$\begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \quad (1.3.7)$$

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

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad (1.3.8)$$

^{*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。

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

$$a_{ij} = 0 \quad (1.3.9)$$

である行列を**対角行列**という。

$$\begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix} \quad (1.3.10)$$

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

$$a_{ij} = \begin{cases} 0 & (i \neq j) \\ a & (i = j) \end{cases} \quad (1.3.11)$$

のとき**スカラー行列**という。

$$\begin{pmatrix} a & 0 & \cdots & 0 \\ 0 & a & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a \end{pmatrix} \quad (1.3.12)$$

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

$$a_{ij} = \begin{cases} 0 & (i \neq j) \\ 1 & (i = j) \end{cases} \quad (1.3.13)$$

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。

$$I_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} \quad (1.3.14)$$

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。

$$\delta_{ij} = \begin{cases} 0 & (i \neq j) \\ 1 & (i = j) \end{cases} \quad (1.3.15)$$

これを用いればスカラー行列や単位行列をそれぞれ

$$A = (a\delta_{ij})_{n \times n}, \quad I_n = (\delta_{ij})_{n \times n} \quad (1.3.16)$$

と表せる。

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

$$a_{ij} = 0 \quad (i > j) \quad (1.3.17)$$

のとき**上三角行列**といい、

$$a_{ij} = 0 \quad (i < j) \quad (1.3.18)$$

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、

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 0 \\ 2 & 4 & 0 \\ 3 & 5 & 6 \end{pmatrix} \quad (1.3.19)$$

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

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \xrightarrow{\text{転置!}} \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix} \quad (1.3.20)$$

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

$$(1) \quad \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} \quad \begin{pmatrix} 1 & 2 & 3 \\ 2 & 5 & 6 \\ 3 & 6 & 0 \end{pmatrix} \quad (1.3.21)$$

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

$$a_{ij} = a_{ji} \quad (1.3.22)$$

である。

以上で一気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。

$$\begin{pmatrix} 1 & 2 & 4 \\ 7 & 3 & 6 \end{pmatrix} \quad (1.3.23)$$

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。

$$\begin{pmatrix} 2 & 6 & 9 & 8 \\ 0 & 2 & 7 & 0 \\ 3 & 6 & 9 & 9 \\ 8 & 1 & 7 & 0 \end{pmatrix} \quad (1.3.24)$$

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 (1.3.24) の転置行列を書け。

例 3.7 実際に (1.3.21) 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

$$(a \quad b \quad c) \quad (1.3.25)$$

3 次の列ベクトルは

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} \quad (1.3.26)$$

などのようになる。

1.3.1.2 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

$$A = (a_{ij})_{n \times m}, \quad B = (b_{ij})_{n \times m} \quad (1.3.27)$$

に対して和 $A + B$ を

$$(A + B)_{n \times m} = (a_{ij} + b_{ij})_{n \times m} \quad (1.3.28)$$

と定義する。例えば 2×3 行列の和は

$$\begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix} + \begin{pmatrix} g & h & i \\ j & k & l \end{pmatrix} = \begin{pmatrix} a+g & b+h & c+i \\ d+j & e+k & f+l \end{pmatrix} \quad (1.3.29)$$

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

$$cA = (ca_{ij})_{n \times m} \quad (1.3.30)$$

である。例えば次の 2×3 行列を c 倍すると

$$c \begin{pmatrix} 1 & 2 & 4 \\ 7 & 3 & 6 \end{pmatrix} = \begin{pmatrix} c & 2c & 4c \\ 7c & 3c & 6c \end{pmatrix} \quad (1.3.31)$$

となる。例題として次の計算をせよ。

例 3.8

$$\begin{pmatrix} 1 & 2 & 4 \\ 7 & 3 & 6 \end{pmatrix} + \begin{pmatrix} 3 & 2 & 1 \\ 5 & 6 & 7 \end{pmatrix} \quad (1.3.32)$$

例 3.9

$$3 \begin{pmatrix} 2 & 6 & 9 \\ 0 & 2 & 7 \\ 3 & 6 & 9 \end{pmatrix} \quad (1.3.33)$$

例 3.10

$$\begin{pmatrix} 1 & 2 \\ 2 & -3 \end{pmatrix} + 2 \begin{pmatrix} 3 & 1 \\ -1 & 1 \end{pmatrix} \quad (1.3.34)$$

ここで (-1) 倍について述べておく。便利な記法として

$$(-1)A = -A \quad (1.3.35)$$

と書く。また、

$$A + (-B) = A - B \quad (1.3.36)$$

^{*9} 単に実数や複素数と思えばよい。

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 1.3.1 (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。この項の最後に、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 ${}^tA = -A$ 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

$$\begin{pmatrix} 0 & 2 \\ -2 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 & 2 \\ -1 & 0 & 3 \\ -2 & -3 & 0 \end{pmatrix} \quad (1.3.37)$$

などである。

定理 1.3.2. 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

$$a_{ij} = -a_{ji} \quad (i, j = 1, 2, \dots, n) \quad (1.3.38)$$

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。(1.3.38) は定義式 ${}^tA = -A$ 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、(1.3.38) と合わせて

$$a_{ii} = -a_{ii} \quad (1.3.39)$$

が成り立つ。よって

$$2a_{ii} = 0 \quad \therefore a_{ii} = 0 \quad (1.3.40)$$

である。

1.3.1.3 積

$l \times m$ 行列 $A = (a_{ij})_{l \times m}$ と $m \times n$ 行列 $B = (b_{jk})_{m \times n}$ の積 AB を次のように定義する。

$$AB = (c_{ik})_{l \times n}, \quad c_{ik} = \sum_{j=1}^m a_{ij}b_{jk} \quad (i = 1, 2, \dots, l, k = 1, 2, \dots, n) \quad (1.3.41)$$

この式の意味は、「行列 AB の (i, k) 成分 c_{ik} が、 A の行の成分と B の列の成分のそれぞれをかけて足したものとなるように定義する」ということである一般にこの積の順序は可換ではないことに注意。この節の始めの文章では太字にして強調してあるように、 A の列の個数と B の行の個数が等しいときにのみ積が定義されるので、これを入れ替えると定義できなくなるためである。

$$\begin{pmatrix} \cdots & \cdots & \cdots \\ \textcolor{red}{a_{i1}} & \cdots & \textcolor{blue}{a_{ij}} & \cdots & \textcolor{green}{a_{im}} \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{pmatrix} \begin{pmatrix} \textcolor{red}{b_{1k}} \\ \vdots \\ \textcolor{blue}{b_{jk}} \\ \vdots \\ \textcolor{green}{b_{mk}} \end{pmatrix} \quad (1.3.42)$$

図 1.5 この行列の積の第 (i, k) 成分は、ここで同じ色をつけたところをかけて足したものとなる。順にかけてすべて足すのである。

以下の計算を試してみよ。計算の結果として型がどうなるかをきちんと注意すること。

例 3.11 2 次正方行列どうしの積。これはすべて書いてしまっているが、きちんと計算があっているか確認してほしい。

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} e & f \\ g & h \end{pmatrix} = \begin{pmatrix} ae + bg & af + bh \\ ce + dg & cf + dh \end{pmatrix} \quad (1.3.43)$$

例 3.12(a) 2 次正方行列どうしの積。

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix} \quad (1.3.44)$$

例 3.12(b) 2 次正方行列どうしの積。

$$\begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \quad (1.3.45)$$

例 3.13 2×3 行列と 3×1 行列の積。

$$\begin{pmatrix} 0 & 1 & 2 \\ 1 & 2 & 0 \end{pmatrix} \begin{pmatrix} 3 \\ 4 \\ 5 \end{pmatrix} \quad (1.3.46)$$

例 3.14 3 次正方行列どうしの積。

$$\begin{pmatrix} 0 & 1 & 2 \\ 1 & 2 & 0 \\ 2 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (1.3.47)$$

例 3.15 2×1 行列と 1×2 行列の積。

$$\begin{pmatrix} 1 \\ 2 \end{pmatrix} \begin{pmatrix} 3 & 4 \end{pmatrix} \quad (1.3.48)$$

例 3.16 1×2 行列と 2×1 行列の積。

$$\begin{pmatrix} 1 & 2 \end{pmatrix} \begin{pmatrix} 3 \\ 4 \end{pmatrix} \quad (1.3.49)$$

例 3.12 は積の順序を入れ替えただけである。前述の通り、積の順序が可換で無いことがわかる。 n 次正方行列の A, B について、積 AB と BA はともに n 次正方行列となるが、これらが一致するとは限らないのである。 $AB = BA$ が成立するとき、この 2 つの行列 A と B は**可換**である、または**交換可能**であるという。なお、可換でないときは**非可換**であるという。例えば、任意の行列 A に対して、その行列と積の演算が可能な型の単位行列 I との積は可換である。例 3.14 のように行列の形は変わらない。

$$AI = IA = A \quad (1.3.50)$$

また、例 3.15, 3.16 は行ベクトルや列ベクトルの積と見れることに注目してほしい。例 3.15 は 2 次正方行列となるが、例 3.16 は 1 次正方行列、すなわち通常の数と見なせる。計算過程を細かく見れば、内積の計算 (1.2.15) と同じであることに気づく。よって $n \times 1$ 行列としての n 次列ベクトル \vec{v}, \vec{w} の**内積**が次のように表せる。行列の考えから、前側のベクトルは転置を取って $1 \times n$ 行列である行ベクトルにする。

$$\vec{v} \cdot \vec{w} = {}^t\vec{v}\vec{w} = (v_1 \quad \cdots \quad v_n) \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} = v_1 w_1 + \cdots + v_n w_n \quad (1.3.51)$$

これ以降、行列の和や積について考えるときは必ず演算が可能な型であるものとする。

定理 1.3.3. A, B, C を行列とすると、次の (A)~(D) が成り立つ。

- (A) $(AB)C = A(BC)$ (積の結合律)
- (B) $(A+B)C = AC + BC$ (分配律 I)
- (C) $A(B+C) = AB + AC$ (分配律 II)
- (D) スカラー c に対して $(cA)B = A(cB) = c(AB)$

なお、積の結合律より $(AB)C = A(BC) = ABC$ と括弧を省略して書いてよい。スカラー行列が cI と書けることから、(1.3.50) と合わせて考えればスカラー行列をかけることはスカラー倍と同じ意味があることがわかる。

積の結合律より、正方行列にはべき乗、すなわち同じ正方行列を何度もかけることが考えられる。正方行列 A に対して、これを n 回 ($n = 0, 1, 2, \dots$) かけたものを A^n と書いて、 A の n 乗という。ただし、 $n = 0$ のとき $A^0 = I$ とする。通常の数のように以下の**指数法則**が成り立つ。

$$A^n A^m = A^{n+m}, \quad (A^n)^m = A^{nm} \quad (n, m = 0, 1, 2, \dots) \quad (1.3.52)$$

次の計算をせよ。

例 3.17

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}^3 \quad (1.3.53)$$

例 3.18

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}^2 \quad (1.3.54)$$

ここで例 3.18 を見るとわかるように、通常の数の場合とは違って、ゼロ行列でなくとも何度かかけてゼロ行列になる行列がある。

べきゼロ行列 ある自然数 n に対して $A^n = O$ となるような行列を**べきゼロ行列**という。

この項の最後に、転置と行列の演算の関係について述べておく。

定理 1.3.4. A, B を行列とすると、次の (A)~(C) が成り立つ。

$$(A) \quad {}^t(A+B) = {}^tA + {}^tB$$

$$(B) \quad {}^t(AB) = {}^tB {}^tA$$

$$(C) \quad \text{スカラー } c \text{ に対して } {}^t(cA) = c {}^tA$$

積の転置を取ると、順序が入れ替わることに注意である。

1.3.2 具体例

以上の積の解説から (1.1.8) の例の気持ちが少しだけわかるかもしれない。しかし、行列の積の利用としてわかりやすいのは図形への応用である。例えば座標 (x, y) を何かの変換を施して (x', y') にすることは、変換行列 $A = (a_{ij})_{2 \times 2}$ を用いて、点を位置ベクトルで表し、

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x' \\ y' \end{pmatrix} \quad (1.3.55)$$

のように書ける。詳しくは後の章で述べるが、このような行列による変換を線形変換、または 1 次変換という。

恒等変換 単位行列をかけても形は変わらないので、この変換は**恒等変換**と呼ばれる。

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} \quad (1.3.56)$$

図形は点の集合であるため、変換をそれぞれの点に施すことで画像の変換処理を行うことができる。(1.3.55) は 1 つの点についてのみだが、複数の点について同時に変換することもできる。変換されるベクトルを位置ベクトル $\vec{r}_i = (x_i, y_i)$ を 2 つ列方向に並べて

$$X = (\vec{r}_1 \quad \vec{r}_2) = \begin{pmatrix} x_1 & x_2 \\ y_1 & y_2 \end{pmatrix} \quad (1.3.57)$$

と行列にすればよい。実際に AX がそれぞれの点 \vec{r}_i を変換していることを確かめよ。

1.3.2.1 拡大・縮小

まず、図形を 2 倍に拡大することを考える。つまり、点の位置を原点に対して 2 倍離れた場所に移動させればよい。例えば、 $(2, 1) \rightarrow (4, 2)$ としたい。

$$\begin{pmatrix} 2 \\ 1 \end{pmatrix} \rightarrow \begin{pmatrix} 4 \\ 2 \end{pmatrix} \quad (1.3.58)$$

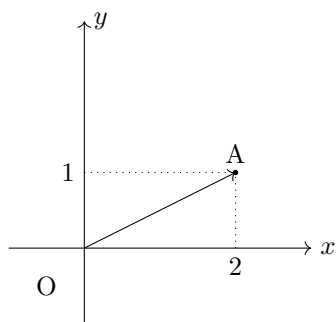


図 1.6 点 A(2, 1)

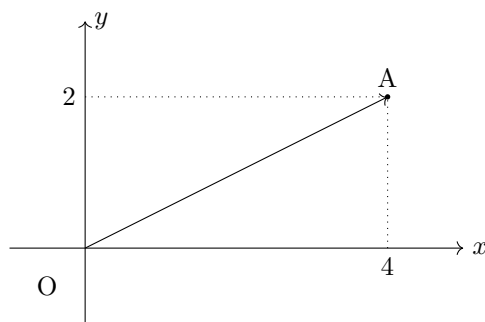


図 1.7 点 A'(4, 2)

このような変換は次の行列を用いて実現する。

$$\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \quad (1.3.59)$$

実際、次のようになることが確かめられる。

$$\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 2 \end{pmatrix} \quad (1.3.60)$$

もっと一般に x 方向に a 倍、 y 方向に b 倍したいときは

$$\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \quad (1.3.61)$$

を用いればよい。

1.3.2.2 回転

角度 θ の回転を表す行列 $R(\theta)$ は次で与えられる。

$$\begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \quad (1.3.62)$$

よって、点 (x, y) を原点に対して θ だけ回転させて点 (x', y') に移動したことを表すには

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \quad (1.3.63)$$

とかけばよい。例えば点 $(1, 0)$ を 60° 度回転させると、以下のように計算できる。

$$\begin{pmatrix} \cos 60^\circ & -\sin 60^\circ \\ \sin 60^\circ & \cos 60^\circ \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ \frac{\sqrt{3}}{2} \end{pmatrix} \quad (1.3.64)$$

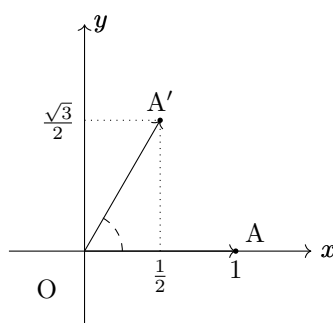


図 1.8 点 $A(1, 0)$ を回転させる。

1.3.3 連立 1 次方程式

ここでは行列を用いて連立 1 次方程式を解いていく。例として次の 2 元連立 1 次方程式を解く。

$$\begin{cases} 2x + 3y = 13 & (1) \\ x + 2y = 8 & (2) \end{cases} \quad (1.3.65)$$

「こんなもの中学生でも解ける」と思うかもしれない。

まず、中学生数学の知識を用いて解いてみる。一応確認しておくが、未知数が **2 つ** に対して与えられた方程式が **2 つ** なので、この式は解ける式である。(1) 式の x の係数が 2 なので、(2) 式を両辺 2 倍することで、 x の**係数を揃える**。

$$\begin{cases} 2x + 3y = 13 & (1) \\ 2x + 4y = 16 & (2) \times 2 \end{cases} \quad (1.3.66)$$

次に、(2) $\times 2$ と (1) を**入れ替える**。

$$\begin{cases} 2x + 4y = 16 & (2) \times 2 \\ 2x + 3y = 13 & (1) \end{cases} \quad (1.3.67)$$

(2) $\times 2$ に、(1) に -1 をかけたものを加える。すなわち、(2) $\times 2 - (1)$ より y だけの方程式にする。

$$y = 3 \quad (1.3.68)$$

最後にもとの (2) 式に $y = 3$ を代入することで $x = 2$ であることがわかる。ゆえに $x = 2, y = 3$ が解である。

さて、行列の知識を携えた上で考える。まず行列の積を用いれば (1.3.65) は次のように書き換えられることを確かめよ。

$$\begin{pmatrix} 2 & 3 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 13 \\ 8 \end{pmatrix} \quad (1.3.69)$$

上での操作「係数を揃える」「入れ替える」「かけて加える」は行列に対してどのように適用すれば

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \end{pmatrix} \quad (1.3.70)$$

を導けるだろうか。

ここでの問を一般化して整理すると次のようにまとめられる。定数 $a_{11}, a_{12}, \dots, a_{mn}$ と b_1, \dots, b_m をとり、 m 個の方程式からなる n 個の未知数 x_1, \dots, x_n についての連立 1 次方程式

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 & (1) \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 & (2) \\ \vdots & \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m & (m) \end{cases} \quad (1.3.71)$$

を考える。これは行列 $A = (a_{ij})_{m \times n}$ と n 次列ベクトル \vec{x} と m 次列ベクトル \vec{b} をそれぞれ

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}, \quad \vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad \vec{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} \quad (1.3.72)$$

と置くことで

$$A\vec{x} = \vec{b} \quad (1.3.73)$$

と書ける。この時 A を**係数行列**という。操作「係数を揃える」「入れ替える」「かけて加える」を整理すると

$$\begin{cases} (1) & 1 \text{ つの式に } 0 \text{ でない定数をかける。} \\ (2) & 2 \text{ つの式を入れ替える。} \\ (3) & 1 \text{ つの式に、他の式の } 0 \text{ でない定数をかけたものを加える。} \end{cases} \quad (1.3.74)$$

この節では、これらの操作を施して $\vec{x} = \dots$ の形で表す方法を提示する。

1.3.3.1 基本変形

(1.3.74) における操作では、式に定数をかけたとき両辺にかかることになる。したがって、(1.3.73) の係数行列だけでなく、右辺の \vec{b} の成分にも注目したほうがよい。ここで、係数行列 A に \vec{b} を並べた $m \times (n+1)$ 行列を $(A|\vec{b})$ と書き、**拡大係数行列**という。

$$(A|\vec{b}) = \left(\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_n \end{array} \right) \quad (1.3.75)$$

操作 (1.3.74) を、この拡大係数行列に対する変形と考えると、次の操作を行なうことに対応する。

$$\begin{cases} (1) & 1 \text{ つの行に } 0 \text{ でない定数をかける。} \\ (2) & 2 \text{ つの行を入れ替える。} \\ (3) & 1 \text{ つの行に、他の行の } 0 \text{ でない定数をかけたものを加える。} \end{cases} \quad (1.3.76)$$

この操作を**行基本変形**、あるいは**初等変形**という。行列に対して勝手にこんなことしてもよいのか?と思うかもしれない。しかし、この拡大係数行列はもともといくつもの等式を並べたものから作っていることに注意すれば、結局何も問題ないことがわかる。

それでは実際に (1.3.65) を例に取って、拡大係数行列を用いる方法で解いてみよう。

$$\begin{pmatrix} 2 & 3 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 13 \\ 8 \end{pmatrix} \Leftrightarrow \left(\begin{array}{cc|c} 2 & 3 & 13 \\ 1 & 2 & 8 \end{array} \right) \quad (1.3.77)$$

この拡大係数行列に基本変形を行っていく。ただし、第 i 行を r_i のように書いている。

$$\left(\begin{array}{cc|c} 2 & 3 & 13 \\ 1 & 2 & 8 \end{array} \right) \rightarrow_{r_1 \leftrightarrow r_2 \times 2} \left(\begin{array}{cc|c} 0 & -1 & -3 \\ 1 & 2 & 8 \end{array} \right) \quad (1.3.78)$$

$$\rightarrow_{r_1 \times (-1)} \left(\begin{array}{cc|c} 0 & 1 & 3 \\ 1 & 2 & 8 \end{array} \right) \quad (1.3.79)$$

$$\rightarrow_{r_2 - r_1 \times 2} \left(\begin{array}{cc|c} 0 & 1 & 3 \\ 1 & 0 & 2 \end{array} \right) \quad (1.3.80)$$

$$\rightarrow_{r_1 \leftrightarrow r_2} \left(\begin{array}{cc|c} 1 & 0 & 2 \\ 0 & 1 & 3 \end{array} \right) \quad (1.3.81)$$

$$(1.3.82)$$

これを再び方程式に戻す。

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \end{pmatrix} \quad (1.3.83)$$

よって $x = 2, y = 3$ である。” 中学数学の解き方” と一致している。きちんと比較して確認すること。このような解き方を**掃き出し法**という。このように、変形を繰り返して” 左側” の行列が単位行列になっていれば解ける。

操作 (1.3.76) を列において適用する変形を**列基本変形**という。この行と列についての変形をあわせて**基本変形**という。上では拡大係数行列について変形を行ったが、もちろん普通の行列でも行なうことができる。 $m \times n$ 行列 A について、基本変形を何回か行なうことで次のような形に変形できる。

$$A \rightarrow_{\text{基本変形}} \left(\begin{array}{c|c} I_r & O_{r,n-r} \\ \hline O_{m-r,r} & O_{m-r,n-r} \end{array} \right) = \left(\begin{array}{cccc|cccc} 1 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & 0 & 0 & \cdots & 0 \\ \hline 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \end{array} \right) \quad (1.3.84)$$

すなわち、変形を繰り返すことで $r (\leq m)$ 次の単位行列とそれ以外すべてゼロ行列という形にできる。ただしここの横線縦線は見やすさのために導入しているだけで具体的な意味はない。この形を A の**階数標準形**という。また、この r を A の**階数**といい、

$$r = \text{rank } A \quad (1.3.85)$$

と書く。なお、ゼロ行列については $\text{rank } O = 0$ とする。階数の定義より次の定理が導かれる。

定理 1.3.5. A を $m \times n$ 行列とすると

$$\text{rank } A \leq m, \quad \text{rank } A \leq n \quad (1.3.86)$$

ただし、階数を求めるだけであれば行基本変形のみで十分であり、階数標準形を求める必要はない。例えば行基本変形を繰り返して

$$\begin{pmatrix} 1 & * & * & * & * \\ 0 & 1 & * & * & * \\ 0 & 0 & 0 & 1 & * \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad (1.3.87)$$

のような形に変形できればよい。このとき階数は 3 である。この形の行列を**階段行列**という。

例 3.19 次の行列の階数標準形と階数を求めよ。

$$(1) \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} \quad (2) \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \quad (3) \begin{pmatrix} a & 1 & 1 \\ 1 & a & 1 \\ 1 & 1 & a \end{pmatrix} \quad (1.3.88)$$

例 3.20 2 次の正方行列に対する階数標準形と階段行列をすべて書け。ただし、* を用いて任意の数を表すこと。

例 3.21 次の連立方程式を解け。

$$\begin{cases} 4x - 3y = -9 \\ 3x - 7y = 17 \end{cases} \quad (1.3.89)$$

例 3.22 次の連立方程式を解け。

$$\begin{cases} x + 3y = -2 \\ x + 12y = 4 \end{cases} \quad (1.3.90)$$

例 3.23 次の連立方程式を解け。

$$\begin{cases} x + y + z = 9 \\ 2x + 3y - 2z = 5 \\ 3x - y + z = 7 \end{cases} \quad (1.3.91)$$

次節で方程式の解き方を本格的に学ぶ。

1.3.3.2 方程式を解く

次のような 2 元連立 1 次方程式を考える。

$$\begin{cases} y = -x - 2 \\ x + y = 4 \end{cases} \quad (1.3.92)$$

解こうとすればすぐわかるように、この方程式は**解くことができない**。そもそも中学数学において初めに連立方程式が出てきたのは、2 つの関数の共有点を求めるときであった。つまり、 xy 平面上に図示すると、2 つの 1 次関数 $y = -x - 2$ と $x + y = 4$ は共有点を持たないことになる。また、次のような例ではどうだろうか。

$$\begin{cases} x - 2z = 0 \\ y + 4z = 0 \end{cases} \quad (1.3.93)$$

これらの式では未知数が 3 つに対して与えられた方程式が 2 つだけなので、解くことが出来ないものとなる。しかし上での例とは少し性質が違っていることがわかるだろうか。すなわち、任意の定数 c を用いて $z = c$ とおけば

$$x = 2c, \quad y = -4c, \quad z = c \quad (1.3.94)$$

と表すことも可能である。なお、この解をベクトルのように表して

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = c \begin{pmatrix} 2 \\ -4 \\ 1 \end{pmatrix} \quad (1.3.95)$$

のようにも書ける。^{*10}つまり、この連立方程式の**解は 1 つだけではない**^{*11}ことになる。ここで言いたいことは、与えられた連立方程式について、「解が存在しない」場合と「解が存在しても一意ではない」場合があるというこ

^{*10} のちに詳しく扱うが、このとき任意定数倍すべてが解になる。この解全体の集合を

$$W = \left\{ c \begin{pmatrix} 2 \\ -4 \\ 1 \end{pmatrix} \middle| c \in \mathbb{R} \right\} \quad (1.3.96)$$

のように書いて**解空間**という。

^{*11} **一意でない**と表現する。

とである。

以上のことを整理する。 $m \times n$ 行列 A 、 n 次列ベクトル \vec{x} 、 m 次列ベクトル \vec{b} として、連立方程式

$$A\vec{x} = \vec{b} \quad (1.3.97)$$

を考える。この方程式について

(A) 解は存在するのか、しないのか。

(B) 解が存在するとしたら、それは一意なのか、そうでないのか。

のことを問う必要がある。実はこれは前述した階数を用いて述べる事が可能である。

定理 1.3.6. (1.3.97) の解が存在する必要十分条件は拡大係数行列 $(A|\vec{b})$ と係数行列 A の階数が一致することである。すなわち、

$$\text{rank}(A|\vec{b}) = \text{rank } A \quad (1.3.98)$$

また、一意的に解が存在するための必要十分条件は、階数が係数行列の個数に一致することである。すなわち、

$$\text{rank}(A|\vec{b}) = \text{rank } A = n \quad (1.3.99)$$

なお、 $\text{rank}(A|\vec{b}) = \text{rank } A < n$ のとき、解は $(n - \text{rank } A)$ 個の任意定数を用いて表される。

証明はしないが、上で考えた例でこの事実を確認する。(1.3.92) は

$$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -2 \\ 4 \end{pmatrix} \quad (1.3.100)$$

と書けることに注意。係数行列は行基本変形によって

$$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \xrightarrow{r_2 - r_1} \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \quad (1.3.101)$$

より、階数 1 であることがわかる。しかし拡大係数行列は

$$\left(\begin{array}{cc|c} 1 & 1 & -2 \\ 1 & 1 & 4 \end{array} \right) \xrightarrow{r_2 - r_1} \left(\begin{array}{cc|c} 1 & 1 & -2 \\ 0 & 0 & 6 \end{array} \right) \quad (1.3.102)$$

より、階数 2 である。すなわち、定理 1.3.6 より解は存在しない。次に (1.3.93) においても同様に行基本変形を行なう。まず方程式が

$$\begin{pmatrix} 1 & 0 & -2 \\ 0 & 1 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad (1.3.103)$$

と書ける。なお、これはもう計算するまでもない。係数行列の階数は 2 で、拡大係数行列の階数も 2 である。すなわち、この方程式の解は存在する。しかし係数行列の列の個数は 3 である。したがって、解は一意的ではなく、 $3 - 2 = 1$ 個の任意定数を用いて解が表されるのであった。

それではいくつか解いてみよう。解が存在するか、存在するなら解は一意なのか、注意して計算すること。

例 3.24 次の連立方程式を解け。

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \quad (1.3.104)$$

例 3.25 次の連立方程式を解け。

$$\begin{pmatrix} 1 & 2 & 3 & 0 \\ 2 & 3 & 0 & 1 \\ 0 & 1 & 6 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \quad (1.3.105)$$

例 3.26 次の連立方程式を解け。

$$\begin{pmatrix} 1 & 2 & 3 & 0 \\ 2 & 3 & 0 & 1 \\ 0 & 1 & 6 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} \quad (1.3.106)$$

特別な例として (1.3.93) のように (1.3.97) において $\vec{b} = 0$ のような場合がある。つまり

$$A\vec{x} = 0 \quad (1.3.107)$$

と書かれる方程式を考える。この形の方程式を**同次**または**斉次**であるという。すぐに

$$\text{rank}(A|0) = \text{rank } A \quad (1.3.108)$$

であることはわかるので解は存在する。このとき、(1.3.107) において明らかに $\vec{x} = 0$ は解である。この解 $\vec{x} = 0$ を**自明な解**という。同次方程式について、定理 1.3.6 より、 $\text{rank } A = n$ のときは自明な解のみが存在する。 $m < n$ ならば自明でない解が存在する。

例 3.27 次の連立方程式を解け。

$$\begin{pmatrix} 1 & 2 & 3 & 0 \\ 2 & 3 & 0 & 1 \\ 0 & 1 & 6 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad (1.3.109)$$

ところで非同次な方程式

$$A\vec{x} = \vec{b} \quad (1.3.110)$$

の解を求めたい時、もちろんきちんと手順を追えば解けるのだが、方程式を満たす解が 1 つ見つかったとすると少し簡単になる。これを \vec{x}_0 とおく。この解を**特殊解**という。一方でこの方程式において $\vec{b} = 0$ とおいた同次形の方程式を考える。

$$A\vec{x} = 0 \quad (1.3.111)$$

これを満たす解を \vec{x}_1 とする。^{*12}もとの非同次方程式の任意の解 (**一般解**) は

$$\vec{x} = \vec{x}_0 + \vec{x}_1 \quad (1.3.112)$$

で与えられる。

連立方程式に関する練習問題を用意してある。

例 3.28 次の連立方程式において、係数行列と拡大係数行列の階数を求め、解を求めよ。

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \quad (1.3.113)$$

^{*12} 一般に同次形のほうが簡単に解ける。拡大係数行列は考えなくて良いため。

例 3.29 次の連立方程式において、係数行列と拡大係数行列の階数を求め、解を求めよ。

$$\begin{pmatrix} 1 & 1 & -2 \\ 1 & -2 & 1 \\ -2 & 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 3 \\ -3 \end{pmatrix} \quad (1.3.114)$$

例 3.30 次の連立方程式において、係数行列と拡大係数行列の階数を求め、解を求めよ。

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \\ 0 & 1 & 2 & 7 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \quad (1.3.115)$$

1.3.3.3 正則行列・逆行列

次の例題について考えてみよう。

例 3.31 以下の連立方程式

$$\begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 \\ -1 \end{pmatrix} \quad (1.3.116)$$

を解け。どう解いてもよい。次に行列

$$\begin{pmatrix} 3 & -2 \\ -2 & 1 \end{pmatrix} \quad (1.3.117)$$

を両辺の左側からかけて、結果を見よ。

普通の数について、**逆数**とは0ではない数 a に対してかけて1になる数 a^{-1} のことをいう。行列の世界で1に対応するものは単位行列 I である。したがって、ある行列 A にかけて単位行列 I になるような行列 B があればそれを逆行列とすることができるだろう。すなわち、 n 次正方行列 A に対して^{*13}、

$$AB = BA = I_n \quad (1.3.118)$$

となるような n 次行列 B があるとき、その行列 B を**逆行列**といい

$$B = A^{-1} \quad (1.3.119)$$

と表す。なお、普通の数と違い、もとの行列がゼロ行列でなくても逆行列が存在しないことがある。逆行列が存在する行列 A のことを**正則**であるといい、 A を**正則行列**という。ちなみに (1.3.118) では積の順序を入れ替えたどちらもが単位行列となるようにしたが、実は $AB = I_n$ または $BA = I_n$ の少なくとも一方が満たされるならばその B は A の逆行列である。

2 次正方行列の例 2 次正方行列

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad (1.3.120)$$

の逆行列が存在する必要十分条件は $ad - bc \neq 0$ である。なお、このとき逆行列は

$$A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \quad (1.3.121)$$

である。

^{*13} 行列 A と B があるとき、積は AB と BA の2つが定義できるため、どちらに対しても単位行列となつてほしい。また、正方行列でない順序を入れかえた積を定義できない。

例 3.32 $AA^{-1} = I_n$ また $A^{-1}A = I_n$ となっていることを確認せよ。

"変換"としての意味 ある変換に対する逆の変換に対応する。つまり、変換前のデータを取り出す操作である。逆行列の存在しない行列による変換は、不可逆な変換に対応すると考えられるだろう。

なお、逆行列は存在すればそれは一意である^{*14}。これを示そう。 n 次正則行列 A の逆行列を B, C とする。どちらも逆行列であるのだから

$$AB = BA = I_n \quad (1.3.122)$$

$$AC = CA = I_n \quad (1.3.123)$$

を満たす。いま $B = C$ であれば逆行列が一意であることを示せたことになる。

$$B = BI_n = B(AC) \quad (\because 1.3.123) \quad (1.3.124)$$

$$= (BA)C \quad (\because \text{積の結合律}) \quad (1.3.125)$$

$$= I_n C = C \quad (\because 1.3.122) \quad (1.3.126)$$

以上より $B = C$ が言えたので、逆行列は存在すれば一意である。

1.3.3.4 正則性とその言い換え

逆行列が存在する行列のことを正則な行列であると述べたが、この正則性には同値な言い換えがいくつかある。

定理 1.3.7. n 次の正方行列 A について、以下の条件はすべて同値である。

- (A) A が正則。
- (B) $\text{rank } A = n$
- (C) A の階数標準形は単位行列である。((1.3.84) 参照)
- (D) 同次連立 1 次方程式 $A\vec{x} = 0$ の解は自明な解のみである。
- (E) 任意の n 次列ベクトル \vec{b} に対して、連立 1 次方程式 $A\vec{x} = \vec{b}$ の解が一意的に存在する。

これらの条件はそれぞれが互いに必要十分である。ここでは示さないで、藤岡 [1] などを参照のこと。この定理があるので任意の正則行列は基本変形を繰り返すことで逆行列を求められる。

$$(A|I_n) \rightarrow (I_n|A^{-1}) \quad (1.3.127)$$

となる。方法は拡大係数行列と同じであり、左側の部分で基本変形を行い、単位行列になったら、右側の部分が逆行列である。

例 3.33 基本変形を行うことで

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix} \quad (1.3.128)$$

の逆行列 A^{-1} が

$$A^{-1} = \begin{pmatrix} 0 & 1 & -1 \\ 1 & -1 & 1 \\ -1 & 1 & 0 \end{pmatrix} \quad (1.3.129)$$

であることを確認し、また、 $AA^{-1} = I_n$ であることを確認せよ。

^{*14} 一意である、すなわちこれが唯一のものであるということを示すには、2 つあることを仮定して議論を進め、結局それらが等しかったというふうにすればよい。

逆行列について基本的な性質を示す定理を述べておく。

定理 1.3.8. n 次正則行列 A, B について、以下が成り立つ。

$$(A) \quad (A^{-1})^{-1} = A$$

$$(B) \quad ({}^t A)^{-1} = {}^t(A^{-1})$$

$$(C) \quad (AB)^{-1} = B^{-1}A^{-1}$$

「かけて単位行列になる」という逆行列の意味を知ればこれらを示すことができるだろう。いくつか逆行列を求めてみよう。2 次の逆行列は (1.3.121) で与えられるので、3 次行列を考える。

例 3.34 次の行列の逆行列を求め、もとの行列との積が単位行列になることを確かめよ。

$$(A) \quad \begin{pmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{pmatrix}, \quad (B) \quad \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 2 \\ 1 & 2 & 1 \end{pmatrix} \quad (1.3.130)$$

1.3.4 行列式

行列式のイメージとその性質。Sarrus の方法と余因子展開で帰納的に定義する。

1.3.4.1 余因子展開

1.3.4.2 外積

1.3.4.3 Cayley-Hamilton の定理

行列の n 乗を計算する際、**かける回数を減らせる** 定理である。計算機を扱う場合には計算回数が少ない方が有利だと思われるので扱う。

1.4 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。

1.4.1 基本事項

1.4.1.1 定義

1.4.1.2 空間という言葉

1.4.2 1 次結合と 1 次独立

1.4.3 基底と次元

1.4.3.1 基底の変換

1.5 線形写像

こっちは AI のモデルでめちゃくちゃ使うらしい。いろんなデータの変換の基礎がこれ。

1.5.1 写像

1.5.2 基本事項

1.5.3 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。

1.6 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。

1.6.1 固有値と固有ベクトル

1.6.1.1 線形変換において

1.6.1.2 一般の変換において

1.6.2 対角化

ここまでできればその辺の理系学生より優秀。

付録 A

おまけ: ランダム行列

おまけとして、近年数学だけでなく物理学の観点からも活発に研究されている**ランダム行列**について述べる。ランダム行列は様々な分野と関係しており、量子情報や量子重力に関係する物理学だけでなく、金融工学、情報理論など、枚挙に暇がない。ここでは私の研究分野である、2次元の重力理論を記述する Jackiw-Teitelboim(JT) 重力理論や Sachdev-Ye-Kitaev(SYK) 模型に関わる話まで持っていきたい。勉強しながら備忘録代わりにメモを取っているので注意。

主に木村 [4] をもとにしている。さらっと読んだ感じ、かなり物理学への応用に重点を置いているようである。渡辺 et al. による入門書 [5] では、それよりも平易な言葉で書かれており、情報系や機械学習の理論についても触れているので、こちらを読んでもいいかもしれない。

A.1 基本の話

行列要素が乱数であるような行列を**ランダム行列**という。ただし、あまりに乱雑すぎても取り扱いが難しくなるため、様々な対称性を課することとなる。ここでは Hermite 共役性を持つ N 次の複素行列、すなわち**複素 Hermite ランダム行列**を考える。Hermite 行列は行列要素 $H_{ij} \in \mathbb{C}$ が

$$H_{ij} = H_{ji}^* \quad (\text{A.1.1})$$

であるような行列である。この定義から対角要素は $H_{ii} = H_{ii}^*$ となるため、実数である。

$$H_{ii} \in \mathbb{R} \quad (\text{A.1.2})$$

なお、ある行列の **Hermite 共役** (随伴) とは転置を取って複素共役を取ることである。これを \dagger を用いて表す。

$${}^t A^* = A^\dagger \quad (\text{A.1.3})$$

Hermite 行列とは $A = A^\dagger$ となる行列と言える。

A.1.1 積分測度

初めに扱うモデルは、Gauss 型ユニタリ集団 (gaussian Unitary Ensemble, GUE) である。行列 H が実現する確率分布を、Gauss 分布に従うとして

$$P(H)dH = \frac{1}{Z_N} e^{-\frac{1}{2}\text{Tr } H^2} dH \quad (\text{A.1.4})$$

と定義する。ここで規格化定数

$$Z_N = \int e^{-\frac{1}{2}\text{Tr } H^2} dH \quad (\text{A.1.5})$$

は**分配関数**と呼ばれる。これによって確率分布は規格化条件

$$\int P(H) dH = 1 \quad (\text{A.1.6})$$

を満たすようになる。連続分布としているので、行列 (演算子) \mathcal{O} の期待値は積分を用いて

$$\langle \mathcal{O} \rangle = \int dH P(H) \mathcal{O}(H) \quad (\text{A.1.7})$$

で定められる。ここで dH は Lebesgue 測度で、

$$dH = \prod_{i=1}^N dH_{ii} \prod_{i < j} dH_{ij}^{(0)} dH_{ij}^{(1)} \quad (\text{A.1.8})$$

のように書かれる。ただし $H_{ij}^{(0)} := \text{Re } H_{ij}$, $H_{ij}^{(1)} := \text{Im } H_{ij}$ である。

固有値を用いて書けるとうれしい。行列 H がユニタリ行列 $U \in U(N)$ で対角化されることを考える。すなわち、対角行列 $X = \text{diag}(x_1, \dots, x_N)$ を用いて、

$$H = UXU^\dagger \quad (\text{A.1.9})$$

と書ける。なお、対角化したときの対角成分はもとの行列の固有値であり、Hermite 性より実数であることに注意せよ。このときトレースの巡回性より

$$\text{Tr } H^2 = \text{Tr } (UXU^\dagger)(UXU^\dagger) = \text{Tr } X^2 = \sum_{i=1}^N x_i^2 \quad (\text{A.1.10})$$

である。測度について

$$dH = d(UXU^\dagger) = dUXU^\dagger + U dXU^\dagger + UX dU^\dagger \quad (\text{A.1.11})$$

また左右に D^\dagger, U をかけて

$$U^\dagger dHU = U^\dagger (dUXU^\dagger + U dXU^\dagger + UX dU^\dagger) U = dX + U^\dagger dUX + X dU^\dagger U \quad (\text{A.1.12})$$

ここで $U^\dagger U = I$ より $dU^\dagger U + U^\dagger dU = 0$ を用いると

$$U^\dagger dHU = dX + U^\dagger dUX - XU^\dagger dU \quad (\text{A.1.13})$$

これを成分で書きあらわすと

$$(U^\dagger dHU)_{ij} = dx_i \delta_{ij} + (x_j - x_i)(U^\dagger dU)_{ij} \quad (\text{A.1.14})$$

ただし

$$(U^\dagger dU)_{ik} X_{kj} - X_{ik} (U^\dagger dU)_{kj} = (U^\dagger dU)_{ik} x_k \delta_{kj} - x_i \delta_{ik} (U^\dagger dU)_{kj} \quad (\text{A.1.15})$$

であることを用いている。さらに、微分形式とウェッジ積を導入することを考える。 N^2 形式 ω_N を

$$\omega_N = \bigwedge_{i=1}^N dH_{ii} \bigwedge_{i < j, \alpha=0,1} dH_{ij}^{(\alpha)} \quad (\text{A.1.16})$$

と定めると行列積分を

$$\int dH = \int \omega_N \quad (\text{A.1.17})$$

と書ける。ここで (A.1.14) を外微分を施した時の偏微分たる係数と思い、また $U^\dagger dHU = dH$ であるから

$$\omega_N = \prod_{i < j}^N |x_j - x_i|^2 \bigwedge_{i=1}^N dx_i \bigwedge_{i < j, \alpha=0,1} (U^\dagger dU)_{ij}^{(\alpha)} \quad (\text{A.1.18})$$

この相似変換におけるヤコビアンは

$$|\Delta_N(X)|^2 := \prod_{i < j}^N |x_j - x_i|^2 \quad (\text{A.1.19})$$

で与えられ、Vandermonde 行列式と呼ばれる。

以上のように固有値と対角化された行列による表示が導かれたのでまとめると、確率分布は

$$P(H)dH = P(X)dXdU \quad (\text{A.1.20})$$

で与えられ、ここで

$$P(X) = \frac{1}{Z_N} e^{-\frac{1}{2} \sum_{i=1}^N x_i^2} \prod_{i < j}^N |x_j - x_i|^2 \quad (\text{A.1.21})$$

であり、測度は以下のように略記している。

$$dX := \bigwedge_{i=1}^N dx_i, \quad dU := \bigwedge_{i < j, \alpha=0,1} (U^\dagger dU)_{ij}^{(\alpha)} \quad (\text{A.1.22})$$

固有値による表示を観察してみる。固有値が近づくと $P(X)$ は小さくなり、特に縮退している場合はゼロとなる。固有値どうしが離れている方が寄与は大きくなるが、しかし共通の因子 $e^{-\frac{1}{2}x^2}$ によって原点付近へ束縛もされる。この競合によって、行列サイズを大きく取る極限で興味深い統計法則が生まれる。なお、ここで確率分布は $U(N)$ 変換で不変である。このような群不変測度のことを Haar 測度という。

A.1.2 分配関数

具体的に分配関数

$$Z_N = \int e^{-\frac{1}{2} \text{Tr } H^2} dH \quad (\text{A.1.23})$$

を求める。肩のトレースについて

$$\begin{aligned} \text{Tr } H^2 &= \sum_{i,j} \underbrace{H_{ij} H_{ji}}_{|H_{ij}|^2} = \sum_i H_{ii}^2 + \sum_{i \neq j} |H_{ij}|^2 \\ &= \sum_i H_{ii}^2 + 2 \sum_{i < j} |H_{ij}|^2 \\ &= \sum_i H_{ii}^2 + 2 \sum_{i < j, \alpha=0,1} \left(H_{ij}^{(\alpha)} \right)^2 \end{aligned} \quad (\text{A.1.24})$$

$$(\text{A.1.25})$$

であるから、これと測度

$$dH = \prod_{i=1}^N dH_{ii} \prod_{i < j, \alpha=0,1} dH_{ij}^{(\alpha)} \quad (\text{A.1.26})$$

をあわせてやると

$$Z_N = \int e^{-\frac{1}{2} \text{Tr } H^2} dH = \int \prod_{i=1}^N e^{-\frac{1}{2} H_{ii}^2} dH_{ii} \prod_{i < j, \alpha=0,1} e^{-\left(H_{ij}^{(\alpha)}\right)^2} dH_{ij}^{(\alpha)} \quad (\text{A.1.27})$$

となる。これは Gauss 積分

$$\int e^{-ax^2} dx = \sqrt{\frac{\pi}{a}} \quad (\text{A.1.28})$$

を用いて計算できる。よって

$$Z_N = \left(\int e^{-\frac{1}{2} H_{ii}^2} dH_{ii} \right)^N \left(\int e^{-\left(H_{ij}^{(\alpha)}\right)^2} dH_{ij}^{(\alpha)} \right)^{N(N-1)} = (2\pi)^{N/2} \pi^{N(N-1)/2} \quad (\text{A.1.29})$$

A.2 Coulomb 気体の方法

行列サイズ N を大きくした時の漸近的な振る舞いを調べるために、固有値を荷電粒子とみなした平衡状態としての固有値分布を議論する **Coulomb 気体の方法**を導入する。かなり物理的な取り扱いをするため、数学的な議論するには向かない。

付録 B

おまけ: 微分形式

自分用メモその 2

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]
- [4] 木村太郎, ランダム行列の数理, [森北出版,2021]
- [5] 渡辺澄夫, 永尾太郎, 樺島祥介, 田中利幸, 中島伸一, ランダム行列の数理と科学, [森北出版,2014]