線形代数の整理帳

tomixy

2025 年 8 月 14 日

目次

第 1 章

ベクトルと座標

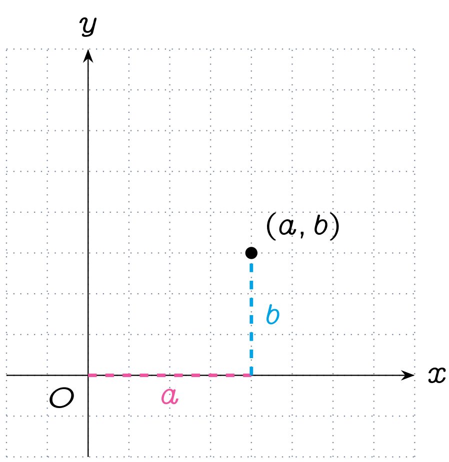
移動の表現としてのベクトル

平面上のある点の位置を表すのに、よく使われるのが直交座標系である。

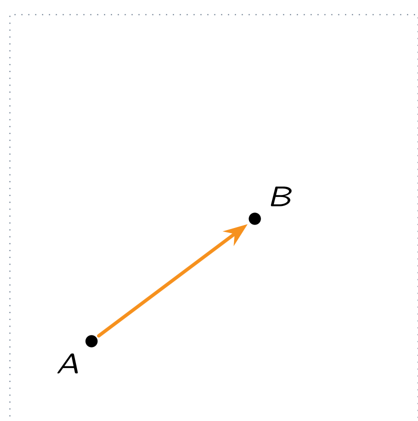
直交座標系では、x 軸と y 軸を垂直に張り、

- 原点 O からの x 軸方向の移動量（x 座標）
- 原点 O からの y 軸方向の移動量（y 座標）

という 2 つの数の組で点の位置を表す。



「位置の特定」という視点



「移動」という視点

座標とは、「x 軸方向の移動」と「y 軸方向の移動」という 2 回の移動を行った結果である。右にどれくらい、上にどれくらい、という考え方で平面上の「位置」を特定しているわけだが、単に「移動」を表したいだけなら、点から点へ向かう矢印で一気に表すこともできる。

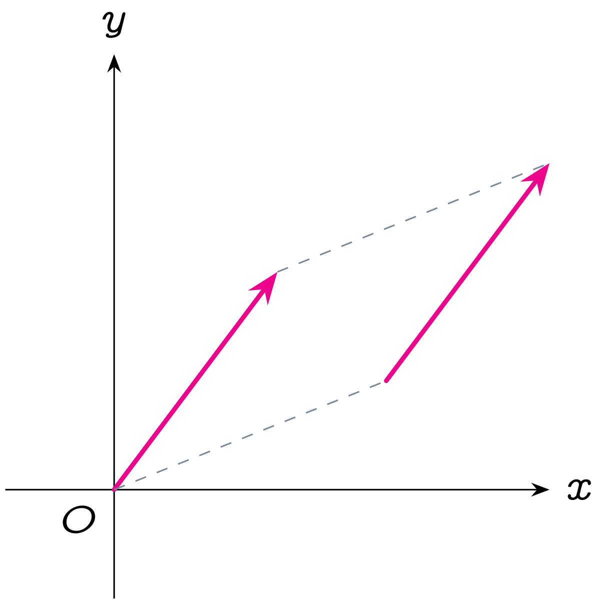
ある地点から別のある地点への「移動」を表す矢印を**ベクトル**という。

ベクトルが示す、ある地点からこのように移動すれば、この地点にたどり着く…といった「移動」の情報は、相対的な「位置関係」を表す上で役に立つ。

平行移動してもベクトルは同じ

座標は「位置」を表すものだが、ベクトルは「移動」を表すものにすぎない。
座標は「原点からの」移動量によって位置を表すが、ベクトルは始点の位置にはこだわらない。

たとえば、次の 2 つのベクトルは始点の位置は異なるが、同じ向きに同じだけ移動している矢印なので、同じベクトルとみなせる。

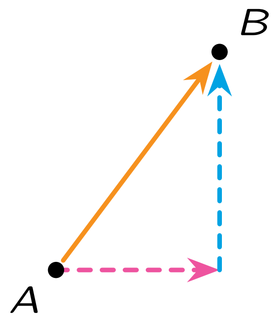


このような「同じ向きに同じだけ移動している矢印」は、平面内では平行な関係にある。
つまり、平行移動して重なる矢印は、同じベクトルとみなすことができる。

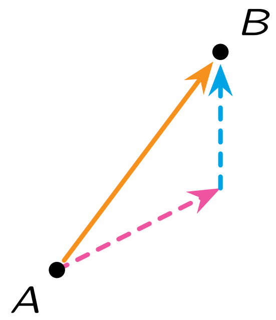
移動の合成とベクトルの分解

ベクトルは、各方向への移動の合成として考えることもできる。

純粋に「縦」と「横」に分解した場合は直交座標の考え方によく似ているが、必ずしも直交する方向のベクトルに分解する必要はない。



「縦」と「横」に分解

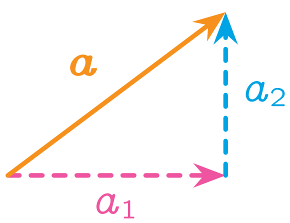


他の分解も考えられる

高次元への対応：数ベクトル

2 次元以上の空間内の「移動」を表すには、「縦」と「横」などといった 2 方向だけでなく、もっと多くの方向への移動量を組み合わせて考える必要がある。
また、4 次元を超えてしまうと、矢印の描き方すら想像がつかなくなってしまう。それは、方向となる軸が多すぎて、どの方向に進むかを表すのが難しくなるためだ。

そこで、一旦「向き」の情報を取り除くことで、高次元に立ち向かえないかと考える。
移動を表す矢印は「どの方向に進むか」と「どれくらい進むか」という向きと大きさの情報を持っているが、その「どれくらい進むか」だけを取り出して並べよう。



$$\boldsymbol{a} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$$

こうして単に「数を並べたもの」もベクトルと呼ぶことにし、このように定義したベクトルを数ベクトルという。

数を並べるとき、縦と横の 2 通りがある。それぞれ列ベクトル、行ベクトルとして定義する。

列ベクトル　数を縦に並べたものを列ベクトルという。

$$\boldsymbol{a} = (a_i) = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$$

行ベクトル　数を横に並べたものを行ベクトルという。

$$\boldsymbol{a} = (a_i) = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \end{pmatrix}$$

単に「ベクトル」と言った場合は、列ベクトルを指すことが多い。

行ベクトルは、列ベクトルを横倒しにしたもの（列ベクトルの転置）と捉えることもできる。

転置による行ベクトルの表現　行ベクトルは、列ベクトル $\boldsymbol{a}$ を転置したものとして表現できる。

$$\boldsymbol{a}^\top = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \end{pmatrix}$$

ベクトルの和

ベクトルによって数をまとめて扱えるようにするために、ベクトルどうしの演算を定義したい。

ベクトルどうしの足し算は、同じ位置にある数どうしの足し算として定義する。

ベクトルの和　2 つの n 次元ベクトル $\boldsymbol{a}$ と $\boldsymbol{b}$ の和を次のように定義する。

$$\boldsymbol{a}+\boldsymbol{b}=(a_i)+(b_i)=\begin{pmatrix} a_1+b_1 \\ a_2+b_2 \\ \vdots \\ a_n+b_n \end{pmatrix}$$

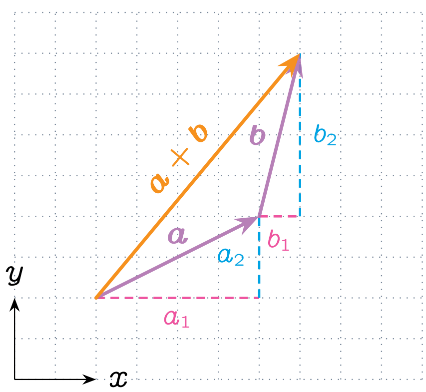
i 番目の数が $\boldsymbol{a}$ と $\boldsymbol{b}$ の両方に存在していなければ、その位置の数どうしの足し算を考えることはできない。

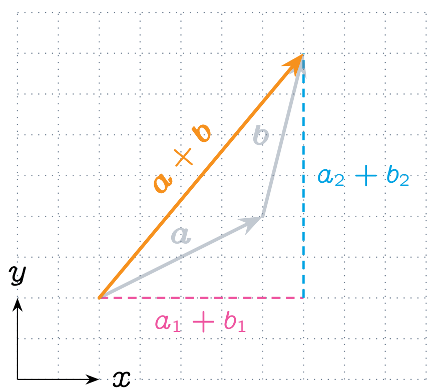
そのため、ベクトルの和が定義できるのは、同じ次元を持つ（並べた数の個数が同じ）ベクトルどうしに限られる。

移動の合成としてのイメージ

数ベクトルを「どれくらい進むか」を並べたものと捉えると、同じ位置にある数どうしを足し合わせるということは、同じ向きに進む量を足し合わせるということになる。

たとえば、x 軸方向に a_1、y 軸方向に a_2 進んだ場所から、さらに x 軸方向に b_1、y 軸方向に b_2 進む…というような「移動の合成」を表すのが、ベクトルの和である。





平行四辺形の法則

[Todo 1: 平行移動しても同じベクトルなので…]

ベクトルの差：逆向きにしてから足す

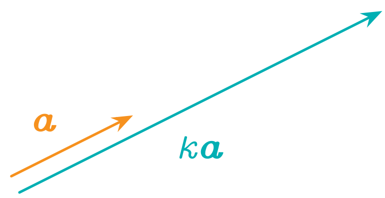
[Todo 2: irobutsu-linear-algebra 2.1.2 ベクトルの差]

矢に沿った移動で考える

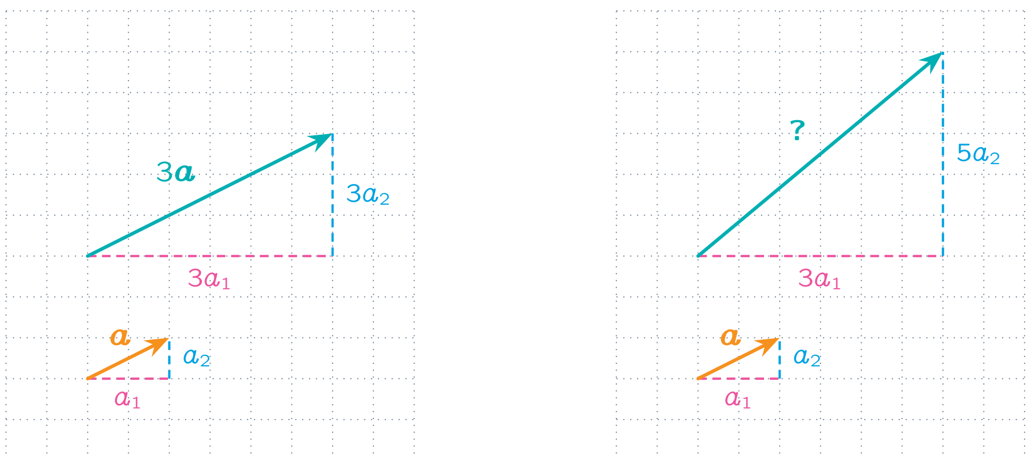
[Todo 3: 手持ちの画像を参考に、和と差の両方について書く]

ベクトルのスカラー倍

「どれくらい進むか」を表す数たち全員に同じ数をかけることで、向きを変えずにベクトルを「引き伸ばす」ことができる。



ここで向きごとにかける数を変えてしまうと、いずれかの方向に多く進むことになり、ベクトルの向きが変わってしまう。そのため、「同じ」数をかけることに意味がある。



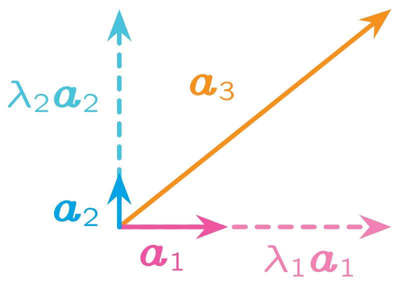
そこで、ベクトルの定数倍（スカラー倍）を次のように定義する。

ベクトルのスカラー倍　n 次元ベクトル $\boldsymbol{a}$ の k 倍を次のように定義する。

$$k\boldsymbol{a} = k(a_i) = \begin{pmatrix} ka_1 \\ ka_2 \\ \vdots \\ ka_n \end{pmatrix}$$

線形結合

ベクトルを「引き伸ばす」スカラー倍と、「つなぎ合わせる」足し算を組み合わせることで、あるベクトルを他のベクトルを使って表すことができる。



$$\boldsymbol{a}_3 = \lambda_1 \boldsymbol{a}_1 + \lambda_2 \boldsymbol{a}_2$$

このように、スカラー倍と和のみを使った形を一次結合もしくは線形結合という。

基底：座標を復元する

3 次元までのベクトルは、矢印によって「ある点を指し示すもの」として定義できる。

しかし、4 次元以上の世界に話を広げるため、ベクトルを単に「数を並べたもの」として再定義した。「数を並べたもの」としてのベクトルを、数ベクトルと呼んでいる。

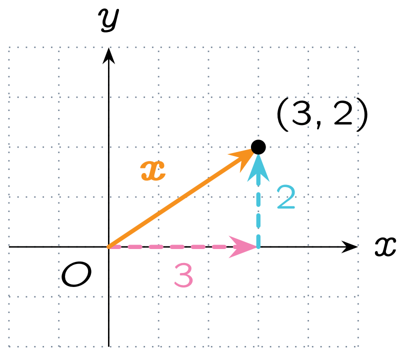
さて、2 次元平面や 3 次元空間で点を指し示すためのもう一つの概念として、座標がある。座標は、x 軸方向にこのくらい進み、y 軸方向にこのくらい進む…というように、「進む方向」と「進む長さ」によって表現される。

単なる数の並びである数ベクトルでは、「進む方向」については何も記述されていない。

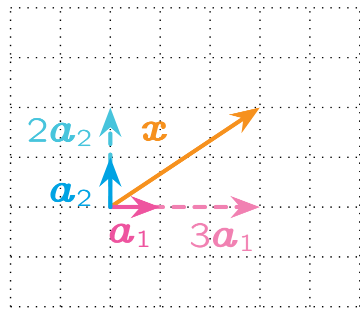
$$\begin{pmatrix} 3 \\ 2 \end{pmatrix}$$

しかし、「進む方向」を表すベクトル $\boldsymbol{a}_1$, $\boldsymbol{a}_2$ を新たに用意すれば、一次結合によって「進む方向」と「進む長さ」を持つベクトルを作ることができる。

$$\boldsymbol{x} = 3\boldsymbol{a}_1 + 2\boldsymbol{a}_2$$



$(3, 2)$



$\boldsymbol{x} = 3\boldsymbol{a}_1 + 2\boldsymbol{a}_2$

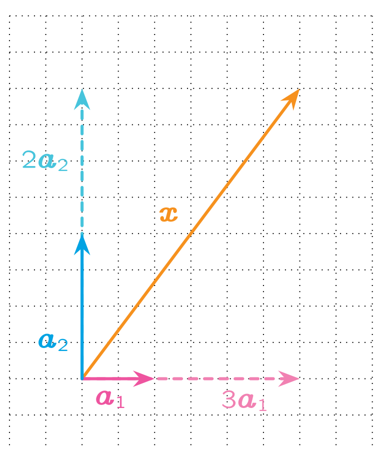
$\boldsymbol{a}_1$ と $\boldsymbol{a}_2$ のように、座標を復元するために向きの情報を付け加えるベクトルを、基底と呼ぶことにする。（厳密には「基底」と呼ぶための条件はいろいろあるが、それについては後々解説していく。）

基底が変われば座標が変わる

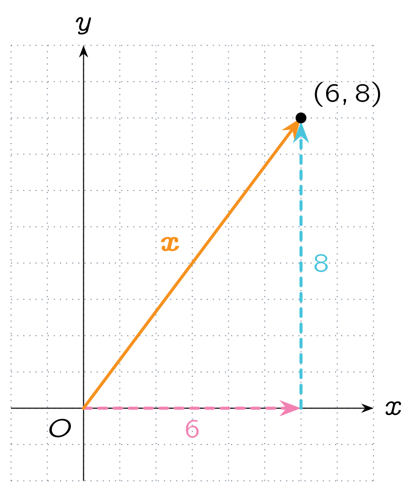
先ほどの例では、直交座標による点 $(3, 2)$ をベクトルの一次結合 $\boldsymbol{x} = 3\boldsymbol{a}_1 + 2\boldsymbol{a}_2$ で表現するために $\boldsymbol{a}_1$ と $\boldsymbol{a}_2$ を用意した。

$\boldsymbol{a}_1$ を x 軸方向の長さ 1 のベクトル、$\boldsymbol{a}_2$ を y 軸方向の長さ 1 のベクトルとすれば、$\boldsymbol{a}_1$ を 3 倍、$\boldsymbol{a}_2$ を 2 倍して足し合わせることで、点 $(5, 4)$ を指し示すベクトル $\boldsymbol{x}$ を作ることができる。

ここで、一次結合の式 $\boldsymbol{x} = 3\boldsymbol{a}_1 + 2\boldsymbol{a}_2$ は変えずに、$\boldsymbol{a}_1$ と $\boldsymbol{a}_2$ を変更すると、$\boldsymbol{x}$ が指し示す点も変わってしまう。



$\boldsymbol{x} = 3\boldsymbol{a}_1 + 2\boldsymbol{a}_2$



$(6, 8)$

このことから、

> 座標は使っている基底の情報とセットでないと意味をなさない

ものだといえる。

標準基底による直交座標系の構成

座標という数値の組は、使っている基底とセットでないと意味をなさないものである。

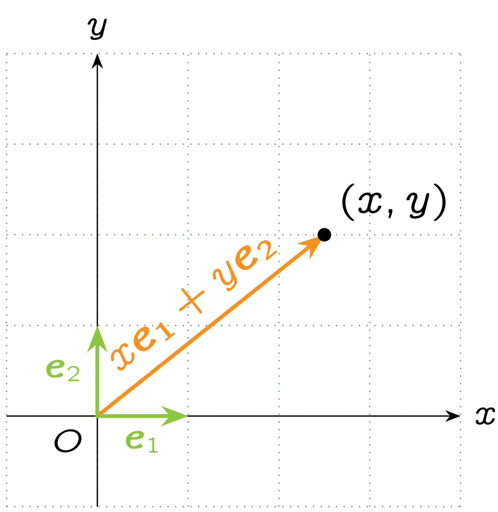
逆にいえば、

> 「こういう基底を使えば、このようなルールで座標を表現できる」

という考え方もできる。つまり、基底によって座標系を定義するということだ。

前の章で見た例を一般化して考えてみよう。

$\boldsymbol{e}_1$ を x 軸方向の長さ 1 のベクトル、$\boldsymbol{e}_2$ を y 軸方向の長さ 1 のベクトルとすれば、$\boldsymbol{e}_1$ を x 倍、$\boldsymbol{e}_2$ を y 倍して足し合わせたベクトル $x\boldsymbol{e}_1 + y\boldsymbol{e}_2$ で、2 次元直交座標系での点 (x, y) を指し示すことができる。



このとき、$\boldsymbol{e}_1$ と $\boldsymbol{e}_2$ は、各方向の 1 目盛に相当する。

これらをまとめて $\mathbb{R}^2$ 上の標準基底と呼び、$\{\boldsymbol{e}_1, \boldsymbol{e}_2\}$ と表す。（$\mathbb{R}^2$ とは、実数の集合である数直線 $\mathbb{R}$ を 2 本用意してつくった、2 次元平面を表す記号である。）

点 (x, y) を指し示す $x\boldsymbol{e_1} + y\boldsymbol{e_2}$ というベクトルは、直交座標による点の表現が「x 軸方向の移動」と「y 軸方向の移動」という 2 回の移動を行った結果であることをうまく表現している。

直交座標系をベクトルの言葉で言い換えると、

> 直交座標系は、標準基底である各ベクトル $\boldsymbol{e}_1$ と $\boldsymbol{e}_2$ を軸として、平面上の点の位置を標準基底の一次結合の係数 x と y の組で表す仕組み

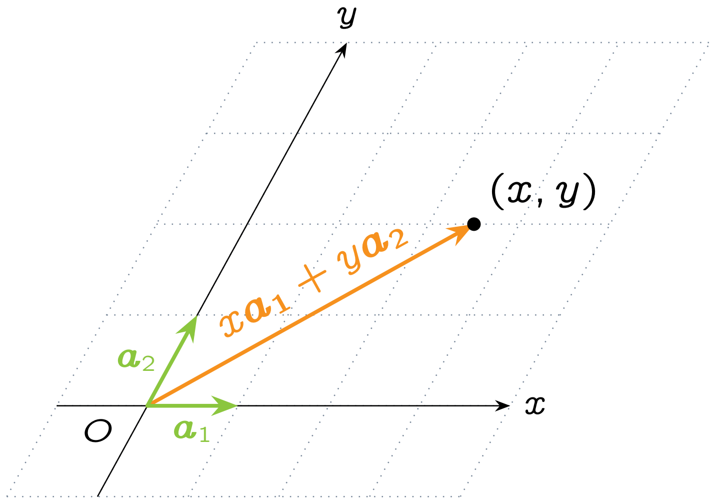
だといえる。

> 座標は点の位置を表す数の組のことで、座標系は点の位置を数の組で表すための仕組み（ルール）のことをいう。

基底を変えれば違う座標系を作れる

直交座標系は、標準基底という互いに直交するベクトルを基底に使っていたが、座標系を表現するにあたって必ずしも基底ベクトルが直交している必要はない。

座標系を基底ベクトルを使って捉え直しておくと、基底を取り替えることで、目的の計算に都合のいい座標系を作ることができる。
たとえば、次のように歪んだ空間を記述するための座標系を作ることも可能である。



ベクトルの次元

n 個の成分からなるベクトルは、n 次元ベクトルと呼ばれる。

ここで、「次元」とは何だろうか？

数ベクトルは、進む方向の数だけ「どれくらい進むか」を表す数値を並べたものとして導入した。

ここで、「進む方向」の情報は基底によって付け加えることができ、基底はいわば座標軸に対応する。

2 次元平面座標系は x 軸と y 軸という 2 つの座標軸で表されるように、次元とは座標軸の数、すなわち基底ベクトルの本数（基準となる方向がいくつあるか）に対応する。

次元とは、1 つの基底を構成するベクトルの本数

方向の数だけ移動量を並べた数ベクトルは、次元の数だけ成分を持つことになる。これが、n 個の成分からなるベクトルが n 次元ベクトルと呼ばれることに対する、ひとつの解釈である。

基底にできるベクトルの条件

座標系では、空間内（たとえば 2 次元空間であれば平面上）のあらゆる点を表すことができ、それらの点はベクトルで指し示す形でも表現できる。
基底が「座標系を設置するための土台」となるなら、基底とは、あらゆるベクトルを表すための材料とみなすことができる。

では、基底として使えるベクトルとは、どのようなベクトルだろうか？

基底とは過不足ない組み合わせ

まず、座標系とは、次の条件を当たり前に満たすものである。

i. あらゆる点を表すことができる
ii. 1 つの点の表し方は一通りである

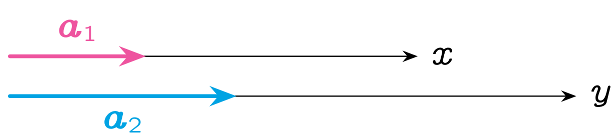
数ベクトルは基底が決まれば座標として使えるので、座標系の条件をベクトルの言葉に言い換えると、

i. 基底が決まれば、その線形結合であらゆるベクトルを表せる
ii. 基底が決まれば、1 つのベクトルの線形結合は一通りに定まる

つまり、空間内のすべてのベクトルを表現する上で、基底は不十分であってはいけないし、無駄があってもいけない。

不十分を考える（2 次元平面の例）

たとえば、2 次元座標系を表現するにあたって、必ずしも基底ベクトルが直交している必要はない。
しかし、平行なベクトルは明らかに基底（座標軸の土台）として使うことはできない。



x 軸と y 軸が平行だと、(x, y) の組で平面上の点を表すことはできない。

2 次元平面 $\mathbb{R}^2$ 上の点やベクトルは、2 つの方向を用意しないと表せないのだから、基底となるベクトルは互いに平行でない必要がある。

無駄を考える（2 次元平面の例）

平行な 2 つのベクトルは、互いに互いをスカラー倍で表現できてしまう。このようなベクトルの組は基底にはできない。

$$\boldsymbol{a}_2 = k\boldsymbol{a}_1$$

この平行な 2 つのベクトル $\{\boldsymbol{a}_1, \boldsymbol{a}_2\}$ に加えて、これらに平行でないもう 1 つのベクトル $\boldsymbol{a}_3$ を用意すれば、$\boldsymbol{a}_1$ と $\boldsymbol{a}_3$ の線型結合か、$\boldsymbol{a}_2$ と $\boldsymbol{a}_3$ の線型結合かのどちらかで、平面上の他のベクトルを表現できるようになる。

しかし、$\boldsymbol{a}_2$ は結局 $\boldsymbol{a}_1$ のスカラー倍（$\boldsymbol{a}_1$ と $\boldsymbol{a}_3$ の線型結合の特別な場合）で表現できてしまうのだから、「他のベクトルを表す材料」となるベクトルの組を考える上で、$\boldsymbol{a}_2$ は無駄なベクトルだといえる。

2 次元平面を表現するには 2 本の座標軸があれば十分なように、基底とは、「これさえあれば他のベクトルを表現できる」という、必要最低限のベクトルの組み合わせにしたい。
基底の候補の中に、互いに互いを表現できる複数のベクトルが含まれているなら、その中の 1 つを残せば十分である。

ここまでの考察から、あるベクトルの組を基底として使えるかどうかを考える上で、「互いに互いを表現できるか」という視点が重要になることがわかる。

- 互いに線型結合で表現できるベクトルだけでは不十分
- 互いに線型結合で表現できるベクトルが含まれていると無駄がある

ベクトルの組の「互いに互いを表現できるか」に着目した性質を表現する概念として、線型従属と線型独立がある。

- 線型従属：互いに互いを表現できるベクトルが含まれていること
- 線型独立：互いに互いを表現できない、独立したベクトルだけで構成されていること

線形従属

ベクトルの組を考え、どれか 1 つのベクトルがほかのベクトルの線形結合で表せるとき、それらのベクトルの組は線形従属であるという。

線形従属　k 個のベクトル $\{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_k\}$ が線形従属であるとは、少なくとも 1 つは 0 でない k 個の係数 $\{c_1, \ldots, c_k\}$ を用意すれば、それらを使った線形結合を零ベクトル $\boldsymbol{o}$ にできることをいう。

$$c_1\boldsymbol{a}_1 + c_2\boldsymbol{a}_2 + \cdots + c_k\boldsymbol{a}_k = \boldsymbol{o}$$

たとえば、c_1 が 0 でないとき、線形結合を零ベクトルにできるということは、次のような式変形ができることになる。

$$\boldsymbol{a}_1 = -\frac{c_2}{c_1}\boldsymbol{a}_2 - \frac{c_3}{c_1}\boldsymbol{a}_3 - \cdots - \frac{c_k}{c_1}\boldsymbol{a}_k$$

つまり、ベクトル $\boldsymbol{a}_1$ をほかのベクトルの線形結合で表せている。

「従属」という言葉を味わう

自分自身をほかのベクトルを使って表現できるということは、ほかのベクトルに依存している（従っている）ということになる。

たとえば、$\boldsymbol{a}_1$ と $\boldsymbol{a}_2$ の線形結合で表せるベクトル $\boldsymbol{a}_3$ は、 この 2 つのベクトル $\boldsymbol{a}_1, \boldsymbol{a}_2$ に従っているといえる。

$$\boldsymbol{a}_3 = 2\boldsymbol{a}_1 + \boldsymbol{a}_2$$

しかし、「$\boldsymbol{a}_3$ が $\boldsymbol{a}_1, \boldsymbol{a}_2$ に従っている」という一方的な主従関係になっているわけではない。その逆もまた然りである。
なぜなら、次のような式変形もできるからだ。

$$\boldsymbol{a}_2 = \boldsymbol{a}_3 - 2\boldsymbol{a}_1$$

この式で見れば、今度は $\boldsymbol{a}_2$ が $\boldsymbol{a}_1, \boldsymbol{a}_3$ に従っていることになる。

このように、線形従属とは、「どちらがどちらに従う」という主従関係ではなく、ベクトルの組の中での相互の依存関係を表すものである。

線形独立

線形従属は、いずれかのベクトルをほかのベクトルで表現できること、つまり基底の候補としては無駄が含まれている。そこで、その逆を考える。

互いに互いを表現できるような無駄なベクトルが含まれておらず、各々が独立している（無関係である）ベクトルの組は線形独立であるという。

線形独立　k 個のベクトル $\{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_k\}$ が線形独立であるとは、k 個の係数 $\{c_1, \ldots, c_k\}$ がすべて 0 であるときしか、それらを使った線形結合を零ベクトル $\boldsymbol{o}$ にできないことをいう。

$$c_1\boldsymbol{a}_1 + \cdots + c_k\boldsymbol{a}_k = \boldsymbol{o} \Longrightarrow c_1 = \cdots = c_k = 0$$

たとえば、係数 c_1 が 0 でないとすると、

$$\boldsymbol{a}_1 = -\frac{c_2}{c_1}\boldsymbol{a}_2 - \frac{c_3}{c_1}\boldsymbol{a}_3 - \cdots - \frac{c_k}{c_1}\boldsymbol{a}_k$$

のように、$\boldsymbol{a}_1$ をほかのベクトルで表現できてしまう。これでは線形従属である。
ほかの係数についても同様で、どれか 1 つでも係数が 0 でなければ、いずれかのベクトルをほかのベクトルで表現できてしまうのである。
このような式変形ができないようにするには、係数はすべて 0 でなければならない。

このように、線形独立には、互いに互いを表現できないようにする条件が課されているため、線形独立なベクトルの組は無駄なベクトルを含まず、基底の候補となり得る。

線形結合の一意性の言い換え

基底にできるベクトルの条件は、次のようなものだった。

i. 基底が決まれば、その線形結合であらゆるベクトルを表せる

ii. 基底が決まれば、1 つのベクトルの線形結合は一通りに定まる

このうち、(ii) の条件について、2 次元平面のイメージにとどまらず一般的に考察してみよう。

(ii) の条件は、次のように言い換えることができる。

ii. 同じ基底を用いた線形結合で表されるベクトルは、同じベクトルである

このことを数式で表すと、

$$x_1\boldsymbol{a}_1+\cdots+x_n\boldsymbol{a}_n=x_1'\boldsymbol{a}_1+\cdots+x_n'\boldsymbol{a}_n \Longrightarrow \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}=\begin{pmatrix} x_1' \\ \vdots \\ x_n' \end{pmatrix}$$

となるが、これは実は $\boldsymbol{a}_1,\ldots,\boldsymbol{a}_n$ が線型独立であることを意味している。

なぜなら、$\Longrightarrow$ の左側の等式を移項して、

$$(x_1-x_1')\boldsymbol{a}_1+\cdots+(x_n-x_n')\boldsymbol{a}_n=\boldsymbol{o}$$

ここで $u_i=x_i-x_i'$ とおくと、

$$u_1\boldsymbol{a}_1+\cdots+u_n\boldsymbol{a}_n=\boldsymbol{o}$$

また、$\Longrightarrow$ の右側の $x_i=x_i'$ という条件は、$u_i=0$ と書き換えられるので、

$$\begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}=\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

すなわち、

$$u_1=\cdots=u_n=0$$

まとめると、

$$u_1\boldsymbol{a}_1+\cdots+u_n\boldsymbol{a}_n=\boldsymbol{o} \Longrightarrow u_1=\cdots=u_n=0$$

となり、これは線型独立の定義式そのものである。

線型独立性における線形結合の一意性　線型独立性は、線形結合の一意性

$$\begin{aligned} c_1\boldsymbol{a}_1 + \cdots + c_k\boldsymbol{a}_k = c_1'\boldsymbol{a}_1 + \cdots + c_k'\boldsymbol{a}_k \\ \Longrightarrow c_1 = c_1', \ldots, c_k = c_k' \end{aligned}$$

と同値である。

この定理から、

線型独立性は、両辺の係数比較ができるという性質

であるとも理解できる。

また、基底となるベクトルの条件は、次のように言い換えられる。

基底となるベクトルは、

i. その線形結合であらゆるベクトルを表せる
ii. 線型独立である（線形結合が一意的である）

線形関係式

線形従属や線型独立の定義では、

$$\text{線形結合} = \boldsymbol{o}$$

という関係式を考えた。以降、この関係式を**線形関係式**と呼ぶことにする。

線形関係式　ベクトル $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_k$ に対する等式

$$c_1\boldsymbol{a}_1 + \cdots + c_k\boldsymbol{a}_k = \boldsymbol{o}$$

を、$\boldsymbol{a}_1, \ldots, \boldsymbol{a}_k$ の線形関係式という。

特に、$c_1 = \cdots = c_k = 0$ として得られる線形関係式を自明な線形関係式という。

これ以外の場合、つまり $c_i \neq 0$ となるような i が少なくとも 1 つあるならば、これは非自明な線形関係式である。

線型独立・線形従属の性質

[Note 1: 基底の存在の章に移動予定]

線形従属なベクトルでは、その中の 1 つのベクトルが、他のベクトルの線形結合で表される

線形結合によるベクトルの表現　$\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_m \in K^n$ を線型独立なベクトルとする

K^n のベクトル $\boldsymbol{a}$ と $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_m$ が一次従属であるとき、$\boldsymbol{a}$ は $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_m$ の線形結合で表される

すなわち、$c_1, c_2, \ldots, c_m \in K$ を用いて次のように書ける

$$\boldsymbol{a} = c_1\boldsymbol{a}_1 + c_2\boldsymbol{a}_2 + \cdots + c_m\boldsymbol{a}_m$$

証明

$\boldsymbol{a}, \boldsymbol{a}_1, \ldots, \boldsymbol{a}_m$ が一次従属であるので、少なくとも 1 つは 0 でない係数 $c, c_1, c_2, \ldots, c_m$ を用いて

$$c\boldsymbol{a} + c_1\boldsymbol{a}_1 + c_2\boldsymbol{a}_2 + \cdots + c_m\boldsymbol{a}_m = \mathbf{0}$$

が成り立つ

もし $c = 0$ だとすると、$c_1, c_2, \ldots, c_m$ のいずれかが 0 でないことになり、$\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_m$ が線型独立であることに矛盾する

よって、$c \neq 0$ である

そのため、上式を c で割ることができ、$\boldsymbol{a}$ は

$$\boldsymbol{a} = -\frac{c_1}{c}\boldsymbol{a}_1 - \frac{c_2}{c}\boldsymbol{a}_2 - \cdots - \frac{c_m}{c}\boldsymbol{a}_m$$

という $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_m$ の線形結合で表せる ■

非自明な線形関係式の存在と線形従属　ベクトルの集まりは、それらに対する非自明な線形関係式が存在するとき、そのときに限り線形従属である

証明

ベクトルの集まりが線型独立であることは、それらに対する線形関係式はすべて自明であるというのが定義である

それを否定すると、「自明でない線形関係式が存在する」となる ■

$k = 1$ の場合に、次の定理が成り立つ

単一ベクトルの線型独立性と零ベクトル

$$\boldsymbol{a}_1 \text{が線型独立} \iff \boldsymbol{a}_1 \neq \boldsymbol{0}$$

証明

$\Longrightarrow$

$\boldsymbol{a}_1$ が線型独立であるとする

すると、$\boldsymbol{a}_1$ に対する線形関係式

$$c_1\boldsymbol{a}_1 = \boldsymbol{0}$$

が成り立つのは、$c_1 = 0$ のときだけである

ここで、$\boldsymbol{a}_1 = \boldsymbol{0}$ と仮定すると、$c_1\boldsymbol{0} = \boldsymbol{0}$ が成り立つので、c_1 は任意の値をとることができる

これは、$\boldsymbol{a}_1$ に対する線形関係式が $c_1 = 0$ のときだけ成り立つという線型独立性の定義に反する

よって、$\boldsymbol{a}_1 \neq \boldsymbol{0}$ である ■

$\Longleftarrow$

$\boldsymbol{a}_1 \neq \boldsymbol{0}$ とする

このとき、もし $\boldsymbol{a}_1$ に対する線形関係式

$$c_1\boldsymbol{a}_1 = \boldsymbol{0}$$

が成り立つとしたら、$\boldsymbol{a}_1 \neq \boldsymbol{0}$ なので、c_1 は必ず 0 でなければならない

したがって、$\boldsymbol{a}_1$ に対する線形関係式は $c_1 = 0$ のときだけ成り立つ

これは、$\boldsymbol{a}_1$ が線型独立であることを意味する ■

第 2 章

線形写像と行列の演算

行列の導入

長方形に並んだ数の集まりを

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

などと書き、行列と呼ぶ

横の数字の並びを行、縦の数字の並びを列と呼ぶ

A は m 個の行と n 個の列をもつ行列である

第 i 行、第 j 列にある数字を a_{ij} と表し、これを (i,j) 成分と呼ぶ

行が m 個、列が n 個の行列は、m 行 n 列の行列、あるいは $m \times n$ 型の行列であるという

$n \times n$ 型の場合、行列は正方形なので n 次正方行列と呼ぶ

A の成分から第 j 列だけを取り出して $\mathbb{R}^m$ のベクトルとしたものが

$$\boldsymbol{a}_j = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix} \quad (1 \leq i \leq n)$$

であり、これを A の j 番目の列ベクトルという

A は、これらを横に並べたものという意味で

$$A = (\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n)$$

と書くことができる

行列とベクトルの積　$m \times n$ 型の行列 $A = (\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n)$ と $\boldsymbol{v} \in \mathbb{R}^n$ との積を

$$A\boldsymbol{v} = v_1\boldsymbol{a}_1 + v_2\boldsymbol{a}_2 + \cdots + v_n\boldsymbol{a}_n$$

により定める

ここで、v_i は $\boldsymbol{v}$ の第 i 成分である

$A\boldsymbol{v}$ を考えるとき、ほとんどの場合は、A が 1 つ与えられていて $\boldsymbol{v}$ がいろいろ動くという意識が強い

それは、行列 A のことを、ベクトルを与えて別なベクトルを作る

$$\text{入力ベクトル } \boldsymbol{v} \longrightarrow \text{出力ベクトル } A\boldsymbol{v}$$

という装置、すなわち写像だとみなすことである

行列のスカラー倍　A を行列、c をスカラーとするとき、A のすべての成分を c 倍して得られる行列を cA とする

行列とベクトルの積の性質　$A,\ B$ を $m \times n$ 型行列、$\boldsymbol{u},\ \boldsymbol{v} \in \mathbb{R}^n$、$c \in \mathbb{R}$ とするとき、次が成り立つ

i. $A(\boldsymbol{u} + \boldsymbol{v}) = A\boldsymbol{u} + A\boldsymbol{v}$

ii. $A(c\boldsymbol{v}) = cA\boldsymbol{v}$

証明

[Todo 4: book: 行列と行列式の基礎 p24（命題 1.4.3）]

線形写像の定義

線形写像と線形性　写像 $f\colon \mathbb{R}^n \to \mathbb{R}^m$ が線形写像であるとは、次の 2 つの条件が成立することである

i. $f(c\boldsymbol{v}) = cf(\boldsymbol{v})$ がすべての $c \in \mathbb{R},\ \boldsymbol{v} \in \mathbb{R}^n$ に対して成り立つ

ii. $f(\boldsymbol{u} + \boldsymbol{v}) = f(\boldsymbol{u}) + f(\boldsymbol{v})$ がすべての $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n$ に対して成り立つ

これらの性質を写像 f の線形性という

また、$m = n$ のとき、線形写像 $f\colon \mathbb{R}^n \to \mathbb{R}^n$ を $\mathbb{R}^n$ の線形変換と呼ぶ

線形変換は空間 $\mathbb{R}^n$ からそれ自身への写像なので、$\mathbb{R}^n$ 内において「ベクトルが変化している」（あるいは f が空間 $\mathbb{R}^n$ に作用している）ニュアンスとみることができる

$f\colon \mathbb{R}^n \to \mathbb{R}^m$ を線形写像とするとき、i より、

$$f(0 \cdot \boldsymbol{v}) = 0 \cdot f(\boldsymbol{v})$$

なので、

$$f(\mathbf{0}) = \mathbf{0}$$

が成り立つ

⚓ **零ベクトルの像**　零ベクトルは線形写像によって零ベクトルに写される

$m = n = 1$ のときは、線形写像 $f\colon \mathbb{R}^1 \longrightarrow \mathbb{R}^1$ は、通常の意味の関数である

このとき、i の性質から、

$$f(c) = f(c \cdot 1) = c \cdot f(1) \quad (c \in \mathbb{R} = \mathbb{R}^1)$$

が成り立つので、$a = f(1) \in \mathbb{R}$ とおくと、

$$f(x) = ax$$

と書ける

⚓ **一次元線形写像と比例関数の同一性**　線形写像 $f\colon \mathbb{R}^1 \longrightarrow \mathbb{R}^1$ は、a を比例定数とする比例関数である

線形写像の表現行列

$f\colon \mathbb{R}^n \longrightarrow \mathbb{R}^m$ を線形写像とするとき、各基本ベクトル $\boldsymbol{e}_j$ の f による像を

$$f(\boldsymbol{e}_j) = \boldsymbol{a}_j = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix}$$

と書くとする

これらを横に並べることによって、m 行 n 列の行列を作る

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix} = (\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n)$$

この行列 A を f の表現行列という

特に、$\mathbb{R}^n$ の線形変換の表現行列は n 次正方行列である

$\mathbb{R}^n$ の一般のベクトル $\boldsymbol{v}$ を、基本ベクトルの線型結合として

$$\boldsymbol{v} = \sum_{j=1}^{n} v_j \boldsymbol{e}_j$$

と書く

このとき、f の線形性より、

$$f(\boldsymbol{v}) = \sum_{j=1}^{n} v_j f(\boldsymbol{e}_j) = \sum_{j=1}^{n} v_j \boldsymbol{a}_j$$

となる

このベクトルの第 i 成分は

$$a_{i1}v_1 + a_{i2}v_2 + \cdots + a_{in}v_n$$

と書ける

これは $A\boldsymbol{v}$ の第 i 成分である

したがって、この記法を踏まえて、次のような表記ができる

⚓ 線形写像とその表現行列の関係

$$f(\boldsymbol{v}) = A\boldsymbol{v}$$

比例関数が比例定数 a だけで決まるのと同じように、線形写像は表現行列 A が与えられれば決まる

零写像と零行列　$f\colon \mathbb{R}^n \to \mathbb{R}^m$ を、すべての $\boldsymbol{v} \in \mathbb{R}^n$ に対して $f(\boldsymbol{v}) = \mathbf{0}$ と定めたものは明らかに線形写像であり、これを零写像と呼ぶ

その表現行列はすべての成分が 0 である行列である

この行列を零行列と呼び、O で表す

$m \times n$ 型であることを明示するために $O_{m,n}$ と書くこともある

また、n 次正方行列の場合は、O_n と書く

恒等写像と単位行列　$f\colon \mathbb{R}^n \to \mathbb{R}^n$ を、すべての $\boldsymbol{v} \in \mathbb{R}^n$ に対して $f(\boldsymbol{v}) = \boldsymbol{v}$ と定めたものは明らかに線形写像である

これを恒等写像と呼び、$f = \mathrm{id}_{\mathbb{R}^n}$ と書く

恒等写像の表現行列は、$f(\boldsymbol{e}_j) = \boldsymbol{e}_j \quad (1 \leq j \leq n)$ より

$$E = (\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n) = \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 1 \end{pmatrix}$$

であり、これを単位行列と呼ぶ

単位行列は正方行列であり、n 次であることを明示したいときは E_n と書く

線形写像 f から行列 A を作ったのとは逆に、任意の行列から線形写像を作ることができる

行列から線形写像を作る　$m \times n$ 型行列 A に対して、

$$f(\boldsymbol{v}) = A\boldsymbol{v} \quad (\boldsymbol{v} \in \mathbb{R}^n)$$

によって写像 $f\colon \mathbb{R}^n \to \mathbb{R}^m$ を定めれば、f は線形写像である

証明

行列とベクトルの積の性質より、f は線形写像である

また、f の定義から明らかに A は f の表現行列である ■

$\mathbb{R}^2$ の線形変換の例

[Todo 5: book: 行列と行列式の基礎 p51 - p56]

行列の積

線形写像の合成　$\mathbb{R}^n$ から $\mathbb{R}^m$ への線形写像 g と、$\mathbb{R}^m$ から $\mathbb{R}^l$ への線形写像 f が与えられているとき、これらを合成して得られる写像

$$f \circ g \colon \mathbb{R}^n \xrightarrow{g} \mathbb{R}^m \xrightarrow{f} \mathbb{R}^l$$

は、$\mathbb{R}^n$ から $\mathbb{R}^l$ への線形写像である

証明

[Todo 6: book: 行列と行列式の基礎 p56 （問 2.2）]

f と g の表現行列をそれぞれ $A = (a_{ij})$, $B = (b_{ij})$ とする

A は $l \times m$ 型、B は $m \times n$ 型の行列である

このとき、$f \circ g$ は $l \times n$ 型行列で表現される

それを C と書くことにして、その成分を計算しよう

そのためには、基本ベクトルの写り先を見ればよい

B を列ベクトルに分解して $B = (\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_n)$ と書くとき、

$$(f \circ g)(\boldsymbol{e}_j) = f(g(\boldsymbol{e}_j)) = f(\boldsymbol{b}_j) = A\boldsymbol{b}_j \quad (1 \leq j \leq n)$$

なので、

$$C = (A\boldsymbol{b}_1, A\boldsymbol{b}_2, \ldots, A\boldsymbol{b}_n)$$

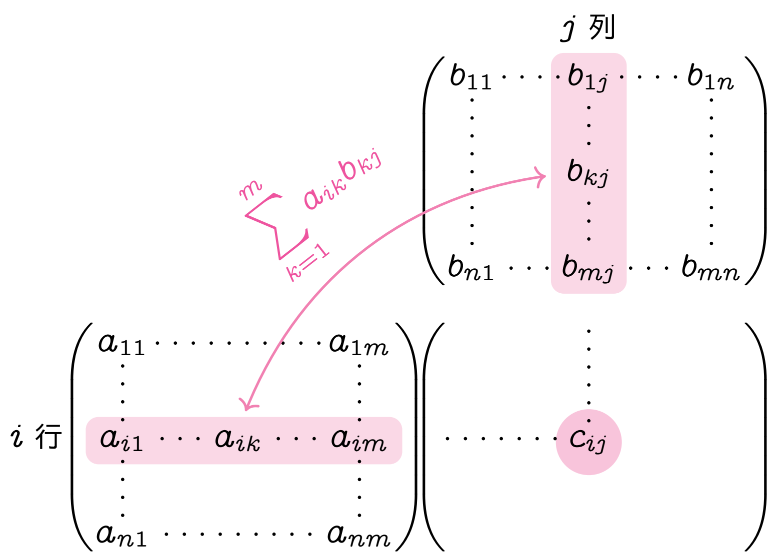
となる

C の (i, j) 成分は $A\boldsymbol{b}_j$ の第 i 成分なので、

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{im}b_{mj} = \sum_{k=1}^{m} a_{ik}b_{kj}$$

により与えられる

つまり、C の (i, j) 成分を計算するときは、A の第 i 行、B の第 j 列だけを見ればよい



このようにして得られた $l \times n$ 型行列 C を AB と書き、A と B の**積**と呼ぶ

⚓ **単位行列との積**　A を $m \times n$ 型とするとき、次が成り立つ

$$E_m A = A$$
$$A E_n = A$$

⚓ **零行列との積**　A を $m \times n$ 型とするとき、次が成り立つ

$$O_m A = A O_n = O_{m,n}$$

2 つの行列の積が順番に依らない場合、2 つの行列は**可換**であるという

一般には、2 つの行列は可換であるとは限らない

つまり、AB と BA は一般には異なる

[Todo 7: book: 行列と行列式の基礎 p58 （例 2.2.3, 2.2.4）]

行列の和とスカラー倍

A, B がともに $m \times n$ 型行列であるとき、それぞれの (i,j) 成分を足すことで行列の和 $A + B$ を定める

⚓ **分配法則**　積が定義できるとき、

$$A(B+C) = AB + AC$$
$$(B+C)A = BA + CA$$

⚓ **行列の積とスカラー倍の性質**　行列 A, B の積 AB が定義できるとき、つまり A の列の個数と B の行の個数が同じであるとき、$c \in \mathbb{R}$ に対して

$$(cA)B = A(cB) = c(AB)$$

が成り立つ

⚓ 線形写像の和　$f, g\colon \mathbb{R}^n \to \mathbb{R}^m$ を線形写像とし、

$$h(\boldsymbol{v}) = f(\boldsymbol{v}) + g(\boldsymbol{v}) \quad (\boldsymbol{v} \in \mathbb{R}^n)$$

により写像 $h\colon \mathbb{R}^n \to \mathbb{R}^m$ を定めるとき、h も線形写像である

また、f, g の表現行列を A, B とするとき、h の表現行列は $A + B$ である

なお、$h = f + g$ と書き、f, g の和と呼ぶ

証明

[Todo 8: book: 行列と行列式の基礎 p59 （問 2.5）]

行列の積の結合法則

⚓ 積の結合法則　積 AB, BC がともに定義できるとき、

$$(AB)C = A(BC)$$

写像による証明

A, B, C がそれぞれ $q \times m$, $m \times n$, $n \times p$ 型行列だとする

線形写像の合成

$$\mathbb{R}^p \xrightarrow{h} \mathbb{R}^n \xrightarrow{g} \mathbb{R}^m \xrightarrow{f} \mathbb{R}^q$$

を考え、f, g, h の表現行列をそれぞれ A, B, C とする

一般的な写像の合成の性質として、

$$(f \circ g) \circ h = f \circ (g \circ h)$$

が成り立つから、

$$(AB)C = A(BC)$$

がしたがう　■

積の計算規則による証明

AB の (i, l) 成分は、

$$(AB)_{il} = \sum_{k=1}^{m} a_{ik} b_{kl}$$

これを用いて、

$$\begin{aligned}((AB)C)_{ij} &= \sum_{l=1}^{n} (AB)_{il} c_{lj} \\ &= \sum_{l=1}^{n} \left(\sum_{k=1}^{m} a_{ik} b_{kl} \right) c_{lj}\end{aligned}$$

i, j はいま固定されているので、和には関係がない

動いているのは k, l だけ

ここで、次の書き換えができる

$$\begin{aligned}\sum_{l=1}^{n} \left(\sum_{k=1}^{m} a_{ik} b_{kl} \right) c_{lj} &= \sum_{l=1}^{n} \left(\sum_{k=1}^{m} a_{ik} b_{kl} c_{lj} \right) \\ &= \sum_{l=1}^{n} \sum_{k=1}^{m} a_{ik} b_{kl} c_{lj}\end{aligned}$$

$\displaystyle\sum_{l=1}^{n}$ の右にある式は l に関する和をとる前のものなので、l は止まっていると考えてよく、単純な分配法則を使っている

また、括弧がなくても、k に関する和を先にとって、その後で l に関する和をとっていると読むことができる

このとき、和の順番は交換してもよいので、

$$\begin{aligned}\sum_{l=1}^{n} \sum_{k=1}^{m} a_{ik} b_{kl} c_{lj} &= \sum_{k=1}^{m} \sum_{l=1}^{n} a_{ik} b_{kl} c_{lj} \\ &= \sum_{k=1}^{m} a_{ik} \left(\sum_{l=1}^{n} b_{kl} c_{lj} \right) \\ &= \sum_{k=1}^{m} a_{ik} (BC)_{kj}\end{aligned}$$

先ほどと同様に、$\displaystyle\sum_{k=1}^{m}$ の右では k は止まっていると考えている

そして、この結果は、$A(BC)$ の (i, j) である ■

結合法則が成り立つことが示されたので、$(AB)C$ または $A(BC)$ を表すとき、括弧を書かずに単に ABC と書いても問題ない

行列の個数が増えても同様である

また、A が正方行列の場合は、

$$A^2 = AA$$
$$A^3 = AAA$$

などのように書く

行列の区分け

行列を

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$$

のようなブロック型に区分けして計算することがよくある

A が $m \times n$ 型のとき、$m = m_1 + m_2$, $n = n_1 + n_2$ として、A_{ij} は $m_i \times n_j$ 型である

また、B が $n \times l$ 型で、$n = n_1 + n_2$, $l = l_1 + l_2$ と区分けして

$$B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}$$

とするとき、

$$\begin{aligned} AB &= \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} \\ &= \begin{pmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \end{pmatrix} \end{aligned}$$

のように A_{ij} などが行列の成分であるかのようにして（ただし積の順序は変えずに）積が計算できる

ここで、A の列の区分けと B の行の区分けの仕方が同じであることが必要である

3 つ以上のブロックに分ける場合も同様である

行列の転置

行列 $A = (a_{ij})$ に対し、その成分の行と列の位置を交換してできる行列を転置行列という

転置行列　$A = (a_{ij})$ を $m \times n$ 型行列とするとき、(i, j) 成分が a_{ji} である $n \times m$ 型行列を A の転置行列と呼び、${}^tA$ と表す

文字 t を左肩に書くのは、右肩に書くと t 乗に見えてしまうからである

t 乗と区別しつつ、右肩に書く流儀として、A^T と書く場合もある

ベクトルの転置

特別な場合として、n 次の数ベクトル $\boldsymbol{v}$ を $n \times 1$ 型行列とみて転置したもの ${}^t\boldsymbol{v}$ は $1 \times n$ 型行列となる

すなわち、数ベクトルの転置は横ベクトルになる

このことを利用して、たとえば

$$\begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$$

を ${}^t(v_1,\ v_2,\ \ldots,\ v_n)$ と表記することもある

転置の性質

転置は「行と列の入れ替え」であるので、明らかに次が成り立つ

⚓ 転置操作の反復不変性　${}^tA$ に対して、転置をもう一度して得られる行列は A と一致する

$$ {}^t({}^tA) = {}^{tt}A = A $$

⚓ 転置と行列積の順序反転性　行列 A, B の積 AB が定義できるとき、

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

証明

[Todo 9: book: 行列と行列式の基礎 p78 命題 2.5.3]

⚓ 行列の和に対する転置の分配性　A と B が同じ型の行列であるとき、

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

証明

[Todo 10:]

対称行列と交代行列

正方行列 A が「転置しても元と変わらない」としたら、A の成分は左上から右下にかけての対角線に関して**対称**（$a_{ij} = a_{ji}$）になっている

対称行列　正方行列 A が次を満たすとき、A を**対称行列**という

$$^{t}A = A$$

交代行列　正方行列 A が次を満たすとき、A を**交代行列**という

$$^{t}A = -A$$

正方行列のトレース

対角成分　正方行列 $A = (a_{ij})$ に対して、a_{ii} を**対角成分**と呼ぶ

トレース　正方行列 $A = (a_{ij})$ に対して、対角成分の和

$$\sum_{i=1}^{n} a_{ii}$$

を A の**トレース**と呼び、$\mathrm{tr}(A)$ と表す

トレースの性質

i. $\mathrm{tr}(A+B) = \mathrm{tr}(A) + \mathrm{tr}(B)$

ii. $\mathrm{tr}(cA) = c\,\mathrm{tr}(A)$

iii. $\mathrm{tr}(AB) = \mathrm{tr}(BA)$

証明

[Todo 11: book: 行列と行列式の基礎 p64 問 2.9]

行列と複素数

$$I = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

とおき、

$$aE + bI = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \quad (a, b \in \mathbb{R})$$

という形の行列を複素数と呼ぶことにより、複素数の定義ができる

この定義では、通常は $a + bi$ と書かれるものを行列として実現している

[Todo 12: book: 意味がわかる線形代数 p43〜49]

第 3 章

対角行列と三角行列

対角行列とスカラー行列

対角行列　対角成分以外の成分がすべて 0 である正方行列を対角行列と呼ぶ

$a_{ii} = c_i \quad (1 \leq i \leq n)$ である対角行列を次のように表す

$$\mathrm{diag}(c_1, c_2, \ldots, c_n) = \begin{pmatrix} c_1 & 0 & \cdots & 0 \\ 0 & c_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & c_n \end{pmatrix}$$

対角行列の特別な場合として、すべての対角成分が同じ値である行列はスカラー行列と呼ばれる

スカラー行列　c をスカラーとするとき、cE の形の行列をスカラー行列という

$$cE = \begin{pmatrix} c & 0 & \cdots & 0 \\ 0 & c & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & c \end{pmatrix}$$

対角行列とスカラー倍

行列 A にスカラー行列をかけることは、

$$(cE)A = A(cE) = cA$$

のように、スカラー c をかけるのと同じである

発展して、対角行列の場合には次のことがいえる

対角行列と列ベクトルのスカラー倍　右から対角行列をかけると、各列ベクトルがスカラー倍になる

すなわち、$A = (\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n)$ とすると、

$$A \cdot \mathrm{diag}(c_1, c_2, \ldots, c_n) = (c_1\boldsymbol{a}_1, c_2\boldsymbol{a}_2, \ldots, c_n\boldsymbol{a}_n)$$

が成り立つ

証明

[Todo 13: book: 行列と行列式の基礎 p63 （問 2.8）]

ブロック対角行列

対角行列の概念は、行列の各成分が数ではなく行列の場合にも拡張できる

ブロック対角行列　対角線上のブロックがすべて正方行列で、それ以外のブロックが零行列であるものをブロック対角行列という

$$\mathrm{diag}(A_1, A_2, \ldots, A_k) = \begin{pmatrix} A_1 & O & \cdots & O \\ O & A_2 & \cdots & O \\ \vdots & \vdots & \ddots & \vdots \\ O & O & \cdots & A_k \end{pmatrix}$$

ここで、対角成分に対応する行列 $A_1, A_2, \ldots, A_k$ を対角ブロックという

対角行列の嬉しさ：入出力の視点

ベクトルと行列を使うことで、入力 $\boldsymbol{x}$ と出力 $\boldsymbol{y}$ の関係を多次元の場合でも簡潔に表すことができる

$$\boldsymbol{y} = A\boldsymbol{x}$$

一般の行列による入出力

たとえば $\boldsymbol{x}$ と $\boldsymbol{y}$ をともに 3 次元ベクトルとすると、$\boldsymbol{y} = A\boldsymbol{x}$ は、

$$\begin{aligned} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} &= \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \\ &= \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 \end{pmatrix} \end{aligned}$$

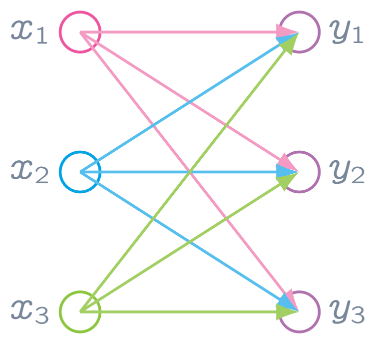
ここで、たとえば 2 行目に注目すると、

$$y_2 = a_{21}x_1 + a_{22}x_2 + a_{23}x_3$$

となり、y_2 の計算に $\boldsymbol{x}$ のすべての成分 x_1, x_2, x_3 が使われていることがわかる

各行に対応する出力 $\boldsymbol{y}_i$ は、入力 $\boldsymbol{x}$ のすべての成分に依存している

この依存関係を、次のようなダイアグラムで表すことにする

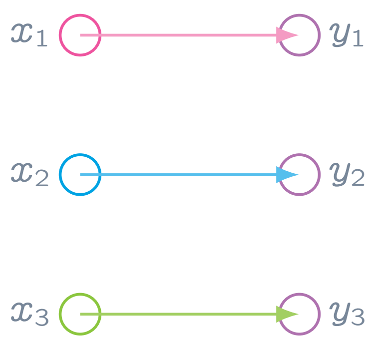


対角行列による入出力

A が対角行列の場合、$\boldsymbol{y} = A\boldsymbol{x}$ は次のような形になる

$$\begin{aligned}\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} &= \begin{pmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \\ &= \begin{pmatrix} a_{11}x_1 \\ a_{22}x_2 \\ a_{33}x_3 \end{pmatrix}\end{aligned}$$

ベクトルの各行に注目すると、各出力 $\boldsymbol{y}_i$ は、入力 $\boldsymbol{x}$ の対応する成分 $\boldsymbol{x}_i$ のみに依存していることがわかる



このように、A が対角行列の場合、$\boldsymbol{y} = A\boldsymbol{x}$ は独立な n 本のサブシステム

$$\begin{aligned} y_1 &= a_{11}x_1 \\ &\vdots \\ y_n &= a_{nn}x_n \end{aligned}$$

に分割されている

つまり、対角行列を使って関係を表現できれば、

> 見た目は n 次元問題でも、実質は 1 次元問題が n 本あるだけ

という状況になり、問題を大きく単純化できる

ブロック対角行列による入出力

ブロック対角行列は、

> 各ブロックごとに独立に変換される

という形の写像を表している

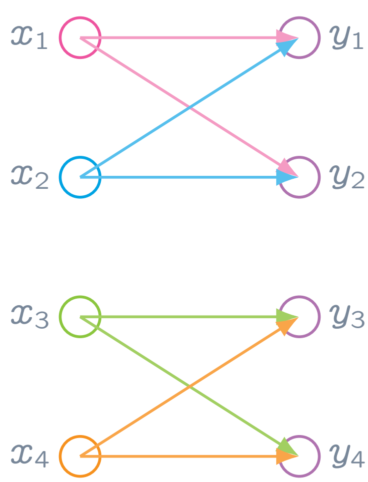
たとえば、

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & 0 & 0 \\ a_{21} & a_{22} & 0 & 0 \\ 0 & 0 & a_{33} & a_{34} \\ 0 & 0 & a_{43} & a_{44} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}$$

というブロック対角行列は、次のように分けて考えることができる

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$
$$\begin{pmatrix} y_3 \\ y_4 \end{pmatrix} = \begin{pmatrix} a_{33} & a_{34} \\ a_{43} & a_{44} \end{pmatrix} \begin{pmatrix} x_3 \\ x_4 \end{pmatrix}$$

ダイアグラムで表すと、2 つの独立なサブシステムに分解されている様子が見える



対角行列の嬉しさ：冪乗の計算

A が対角行列の場合、$\boldsymbol{y} = A\boldsymbol{x}$ は、行ごとのサブシステムとして各行を独立に計算できた

$$\begin{aligned} y_1 &= a_{11}x_1 \\ &\vdots \\ y_n &= a_{nn}x_n \end{aligned}$$

このように各行に分けて「1 次元問題が n 本あるだけ」と考えると、対角行列どうしの積や冪乗も、簡単に計算できることがわかる

$$\begin{pmatrix} a_1 & & \\ & \ddots & \\ & & a_n \end{pmatrix} \begin{pmatrix} b_1 & & \\ & \ddots & \\ & & b_n \end{pmatrix} = \begin{pmatrix} a_1 b_1 & & \\ & \ddots & \\ & & a_n b_n \end{pmatrix}$$

$$\begin{pmatrix} a_1 & & \\ & \ddots & \\ & & a_n \end{pmatrix}^k = \begin{pmatrix} a_1^k & & \\ & \ddots & \\ & & a_n^k \end{pmatrix}$$

三角行列

上三角行列の積　上三角行列どうしの積は上三角行列となる

証明

[Todo 14:]

第 4 章

連立一次方程式と掃き出し法

順問題と逆問題

世の中には、「ベクトル $\boldsymbol{x}$ を入力するとベクトル $\boldsymbol{y} = A\boldsymbol{x}$ が出力される」という形で表せる対象がたくさんある

この $\boldsymbol{y} = A\boldsymbol{x}$ という式は、

> 原因 $\boldsymbol{x}$ を知って結果 $\boldsymbol{y}$ を予測する

という場面（順問題）でそのまま使うことができる

一方、次のような

> 結果 $\boldsymbol{y}$ を知って原因 $\boldsymbol{x}$ を推定する

という問題（逆問題）を考えなければいけない場合もある

$\boldsymbol{y} = A\boldsymbol{x}$ という式（結果）から $\boldsymbol{x}$（原因）を求めるという問題は、連立一次方程式を解くということに他ならない

連立一次方程式の行列表記

未知数 $x_1, x_2, \ldots, x_n$ に関する連立方程式として

$$\begin{cases} a_{11}x_1 + a_{12}\,x_2 + \cdots + \;a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}\,x_2 + \cdots + \;a_{2n}x_n = b_2 \\ \quad\vdots \qquad\quad \vdots \qquad\qquad\qquad\quad \vdots \qquad\; \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases}$$

を考える

a_{ij} などは与えられた定数であり、係数と呼ばれる

i 番目の式の x_j の係数を a_{ij} と書いている

ここで、係数だけを集めて行列を作る

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

すると、先ほどの連立方程式は、ベクトル形で

$$x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + \cdots + x_n\boldsymbol{a}_n = \boldsymbol{b}$$

と書ける

また、n 個の未知数 $x_1, x_2, \ldots, x_n$ からベクトルを作る

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

すると、ベクトル形の方程式の左辺のベクトルを、行列 A とベクトル $\boldsymbol{x}$ の積と考えて、$A\boldsymbol{x}$ と表記できる

こうして、もとの連立一次方程式は、行列形の方程式

$$A\boldsymbol{x} = \boldsymbol{b}$$

に書き換えられる

行基本変形

連立一次方程式を行列によってとり扱うとき、1 つ 1 つの方程式は行列の行によって表されている

そこで、連立方程式の式変形に対応する操作として、行列の行に関する次のような操作（変形）を考える

行基本変形　行列への次の 3 種類の操作を**行基本変形**という

i. ある行の定数倍を他の行に加える
ii. ある行に 0 でない数をかける
iii. 2 つの行を交換する

拡大係数行列

次のような連立一次方程式を考える

$$\begin{cases} x + y = 3 \\ -x + z = -1 \\ -2x - y + 2z = -4 \end{cases}$$

0 や 1 の係数を省略せずに書くと、

$$\begin{cases} 1x + 1y + 0z = 3 \\ -1x + 0y + 1z = -1 \\ -2x - 1y + 2z = -4 \end{cases}$$

となるので、これを行列で表すと、

$$\begin{pmatrix} 1 & 1 & 0 \\ -1 & 0 & 1 \\ -2 & -1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 3 \\ -1 \\ -4 \end{pmatrix}$$

左辺の行列は、連立方程式の係数だけを取り出した行列になっているので、係数行列と呼ばれる

$$A = \begin{pmatrix} 1 & 1 & 0 \\ -1 & 0 & 1 \\ -2 & -1 & 2 \end{pmatrix}$$

また、右辺のベクトルは定数項をまとめたものになっているので、定数項ベクトルと呼ばれる

$$\boldsymbol{b} = \begin{pmatrix} 3 \\ -1 \\ -4 \end{pmatrix}$$

連立方程式の解を求める過程の式変形（行基本変形）によって、係数行列 A と定数項ベクトル $\boldsymbol{b}$ の成分が変化していく

そこで、変形の過程で変化する数（操作の対象）を、次のように 1 つの行列 $(A \mid \boldsymbol{b})$ としてまとめてしまおう

$$\left(\begin{array}{ccc|c} 1 & 1 & 0 & 3 \\ -1 & 0 & 1 & -1 \\ -2 & -1 & 2 & -4 \end{array}\right)$$

このように、係数行列と定数項ベクトルを 1 つの行列としてまとめたものを拡大係数行列という

拡大係数行列に対して行基本変形を繰り返し行うことで、連立一次方程式の解を求めることができる

拡大係数行列の変形と単位行列

連立方程式に対する式変形の結果として得られる形にはさまざまなパターンがあるが、まずは最も単純な場合を考える

たとえば、最終的に次のような形になれば、解 x, y, z が求まったことになる（★ はそれぞれ異なる数でよい）

$$\begin{cases} x & & = \bigstar \\ & y & = \bigstar \\ & & z = \bigstar \end{cases}$$

この理想形を拡大係数行列で表すと、次のようになる

$$\left(\begin{array}{ccc|c} 1 & 0 & 0 & \bigstar \\ 0 & 1 & 0 & \bigstar \\ 0 & 0 & 1 & \bigstar \end{array}\right)$$

つまり、$(A \mid \boldsymbol{b})$ という拡大係数行列に対して、$(E \mid \bigstar)$ という形を目指して行基本変形を施すことで、連立方程式を解くことができる

掃き出し法

具体的には、次のような手順によって、拡大係数行列を変形していく

ここで述べる手順は**掃き出し法**と呼ばれるものである

$$\begin{array}{c} \left(\begin{array}{ccc|c} 1 & 1 & 0 & 3 \\ -1 & 0 & 1 & -1 \\ -2 & -1 & 2 & -4 \end{array}\right) \end{array} \begin{array}{l} R_1 \\ R_2 \\ R_3 \end{array} \qquad \begin{cases} x + y & = 3 \\ -x \qquad + \ z & = -1 \\ -2x - y + 2z & = -4 \end{cases}$$

まず、$(1, 1)$ 成分より下の成分が 0 になるように基本変形を適用する

このことを、「$(1, 1)$ 成分を**要**にして第 1 列を掃き出す」と表現する

$$\left(\begin{array}{ccc|c} 1 & 1 & 0 & 3 \\ 0 & 1 & 1 & 2 \\ -2 & -1 & 2 & -4 \end{array}\right) R_2 + R_1 \qquad \begin{cases} x + y & = 3 \\ y + \ z & = 2 \\ -2x - y + 2z & = -4 \end{cases}$$

$$\left(\begin{array}{ccc|c} 1 & 1 & 0 & 3 \\ 0 & 1 & 1 & 2 \\ 0 & 1 & 2 & 2 \end{array}\right) R_3 + 2R_1 \qquad \begin{cases} x + y & = 3 \\ y + \ z & = 2 \\ y + 2z & = 2 \end{cases}$$

今度は、$(2,2)$ 成分を要にして第 2 列を掃き出す

$$\left(\begin{array}{ccc|c} 1 & 1 & 0 & 3 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 1 & 0 \end{array}\right) R_3 - R_2 \qquad \begin{cases} x + y \phantom{{}+z} = 3 \\ \phantom{x+{}} y + z = 2 \\ \phantom{x+y+{}} z = 0 \end{cases}$$

これで、対角成分がすべて 1 になった

最後に、対角成分以外の成分を 0 にするための行基本変形を施す

$$\left(\begin{array}{ccc|c} 1 & 1 & 0 & 3 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 0 \end{array}\right) R_2 - R_3 \qquad \begin{cases} x + y \phantom{{}+z} = 3 \\ \phantom{x+{}} y \phantom{{}+z} = 2 \\ \phantom{x+y+{}} z = 0 \end{cases}$$

$$\left(\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 0 \end{array}\right) R_1 - R_2 \qquad \begin{cases} x \phantom{{}+y+z} = 1 \\ \phantom{x+{}} y \phantom{{}+z} = 2 \\ \phantom{x+y+{}} z = 0 \end{cases}$$

拡大係数行列の変形と上三角形

ところで、先ほどの例では、対角成分以外の成分を 0 にしなくても、対角成分がすべて 1 になった時点で、解は十分に読み取れる形になっている

$$\left(\begin{array}{ccc|c} 1 & 1 & 0 & 3 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 1 & 0 \end{array}\right) \qquad \begin{cases} x + y \phantom{{}+z} = 3 \\ \phantom{x+{}} y + z = 2 \\ \phantom{x+y+{}} z = 0 \end{cases}$$

この時点でも、$z = 0$ を代入すれば $y = 2$ が得られ、さらに $y = 2$ を代入すれば $x = 1$ が得られることがすぐにわかる

このとき、係数行列は上三角行列になっているので、この形の方程式は上三角形と呼ばれる

上三角形を目指すことが、掃き出し法の基本方針である

しかし、いつでも上三角形に変形できるわけではない。

掃き出し法によって、係数行列を単位行列に変形できない場合もある。

- 解が一意に定まらない場合（解が無限個ある場合）
- 解が存在しない場合

解が無限個ある場合

まずは、解が一意に定まらない場合を見てみよう

$$\left(\begin{array}{ccc|c} 1 & -1 & -3 & 1 \\ 2 & 1 & 0 & 2 \\ 1 & 0 & -1 & 1 \end{array}\right) \begin{array}{l} R_1 \\ R_2 \\ R_3 \end{array} \qquad \begin{cases} x - y - 3z = 1 \\ 2x + y = 2 \\ x - z = 1 \end{cases}$$

$(1, 1)$ 成分を要にして第 1 列を掃き出す：

$$\left(\begin{array}{ccc|c} 1 & -1 & -3 & 1 \\ 0 & 3 & 6 & 0 \\ 0 & 1 & 2 & 0 \end{array}\right) \begin{array}{l} \\ R_2 - 2R_1 \\ R_3 - R_1 \end{array} \qquad \begin{cases} x - y - 3z = 1 \\ 3y + 6z = 0 \\ y + 2z = 0 \end{cases}$$

$(2, 2)$ 成分を 1 にする：

$$\left(\begin{array}{ccc|c} 1 & -1 & -3 & 1 \\ 0 & 1 & 2 & 0 \\ 0 & 1 & 2 & 0 \end{array}\right) \begin{array}{l} \\ \frac{1}{3}R_2 \\ \\ \end{array} \qquad \begin{cases} x - y - 3z = 1 \\ y + 2z = 0 \\ y + 2z = 0 \end{cases}$$

$(2, 2)$ 成分を要にして第 2 列を掃き出す：

$$\left(\begin{array}{ccc|c} 1 & -1 & -3 & 1 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right) \begin{array}{l} \\ \\ R_3 - R_2 \end{array} \qquad \begin{cases} x - y - 3z = 1 \\ y + 2z = 0 \\ 0 = 0 \end{cases}$$

1 を対角成分として持つ列の対角成分以外を 0 にする：

$$\left(\begin{array}{ccc|c} 1 & 0 & -1 & 1 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right) R_1 - R_2 \qquad \begin{cases} x \quad - \quad z \quad = 1 \\ \quad y + 2z \quad = 0 \\ \qquad\qquad 0 = 0 \end{cases}$$

この連立方程式は、実質的に 2 本の方程式しか持たないことがわかる

$$\begin{cases} x \quad - \quad z = 1 \\ \quad y + 2z = 0 \end{cases}$$

x, y について解くと、

$$\begin{cases} x \quad = z \quad + 1 \\ \quad y = -2z \end{cases}$$

となるので、z に任意の数 $z = \alpha$ を与えて解が得られる

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \alpha + 1 \\ -2\alpha \\ \alpha \end{pmatrix} = \alpha \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

解が存在しない場合

次のような連立一次方程式を考える

$$\left(\begin{array}{ccc|c} 1 & -1 & -3 & 1 \\ 2 & 1 & 0 & 1 \\ 1 & 0 & -1 & 1 \end{array}\right) \begin{array}{l} R_1 \\ R_2 \\ R_3 \end{array} \qquad \begin{cases} x - y - 3z = 1 \\ 2x + y \qquad = 1 \\ x \qquad - \quad z = 1 \end{cases}$$

$(1, 1)$ 成分を要にして第 1 列を掃き出す：

$$\left(\begin{array}{ccc|c} 1 & -1 & -3 & 1 \\ 0 & 3 & 6 & -1 \\ 0 & 1 & 2 & 0 \end{array}\right) \begin{array}{l} \\ R_2 - 2R_1 \\ R_3 - R_1 \end{array} \qquad \begin{cases} x - \quad y - 3z = 1 \\ \quad 3y + 6z = -1 \\ \quad y + 2z = 0 \end{cases}$$

$(2, 2)$ 成分を 1 にするため、第 2 行と第 3 行を入れ替える：

$$\begin{pmatrix} 1 & -1 & -3 & \bigm| & 1 \\ 0 & 1 & 2 & \bigm| & 0 \\ 0 & 3 & 6 & \bigm| & -1 \end{pmatrix} \begin{array}{l} \\ R_3 \\ R_2 \end{array} \qquad \begin{cases} x - y - 3z = 1 \\ y + 2z = 0 \\ 3y + 6z = -1 \end{cases}$$

$(2, 2)$ 成分を要にして第 2 列を掃き出す：

$$\begin{pmatrix} 1 & -1 & -3 & \bigm| & 1 \\ 0 & 1 & 2 & \bigm| & 0 \\ 0 & 0 & 0 & \bigm| & -1 \end{pmatrix} \begin{array}{l} \\ \\ R_3 - 3R_2 \end{array} \qquad \begin{cases} x - y - 3z = 1 \\ y + 2z = 0 \\ 0 = -1 \end{cases}$$

$0 = -1$ という式が現れたので、この連立方程式には解が存在しない

掃き出し法の段階ごとに得られる形

ここまで見てきた、掃き出し法による連立方程式の解法をまとめると、大まかには次のような手順を踏むことになる

1. 左の列から順に、対角成分を 1 にする
2. 対角成分が 1 となっている列の対角成分以外を 0 にする

手順 1 で得られる形を行階段行列と呼び、手順 2 で得られる形を既約行階段行列と呼ぶ

ただし、$0 = -1$ が現れたときのように、手順 1（行階段行列への変形）だけで解が存在するかはわかってしまう

解の存在以外にも、行階段行列に変形した時点で読み取れる情報はさまざまある

行階段行列

掃き出し法では、あるステップで下の成分がすべて 0 になって、

$$\begin{pmatrix} \spadesuit & * & * & * & * & * & * & * \\ 0 & \spadesuit & * & * & * & * & * & * \\ 0 & 0 & \spadesuit & * & * & * & * & * \\ 0 & 0 & 0 & \spadesuit & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

のような形になるのが典型例である。

ここで、0 でない成分を $\spadesuit$ で、任意の値をもつ成分を $*$ で表している。

零行

一般には、成分が 0 ばかりの行が下にくる。そのような行を零行という。

零行が現れない場合もあるし、複数現れる場合もある。

主成分

零行でない行に対して、一番左の 0 でない成分 $\spadesuit$ を主成分あるいは行に関する要と呼ぶ。

行階段行列の一般形

先ほど示した形では、行の主成分 $\spadesuit$ は左上から斜め右下 45° 方向にまっすぐ並んでいるが、一般にはそうできるとは限らない。

しかし、次のような形には必ずできる。

$$\begin{pmatrix} 0 & \spadesuit & * & * & * & * & * & * \\ 0 & 0 & 0 & \spadesuit & * & * & * & * \\ 0 & 0 & 0 & 0 & \spadesuit & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \spadesuit \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

行階段行列　次の条件を満たす行列を行階段行列という。

- 零行でない行の主成分が、下の行ほど 1 つ以上右にある

- 零行がある場合は、まとめてすべて下にある

どんな行列も、行基本変形の繰り返しで行階段行列にできる。

既約行階段行列

必要に応じて、行階段行列をさらに変形して次のような形にする

$$\begin{pmatrix} 0 & 1 & * & 0 & 0 & * & * & 0 \\ 0 & 0 & 0 & 1 & 0 & * & * & 0 \\ 0 & 0 & 0 & 0 & 1 & * & * & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

行の主成分はすべて 1 で、主成分のある列の主成分以外の成分はすべて 0 である

この形を簡約化された行階段行列あるいは既約行階段行列と呼ぶ

与えられた行列 A に対して、行基本変形の繰り返しで得られる行階段行列は一意的ではないが、既約行階段行列は一意的であることを後に議論する

そこで、既約行階段行列を $A_\circ$ と書くことにする

変形の過程を

$$\text{行列 } A \longrightarrow \text{行階段行列} \longrightarrow \text{簡約化された行階段行列 } A_\circ$$

と 2 段階にわけるのは、計算の効率以上の意味がある

行階段行列にするところまでで解決する問題（解の存在と一意性など）もあるからである

第 5 章

線形写像の像と核

線形写像と逆問題

$\boldsymbol{y} = A\boldsymbol{x}$ という形の式は、$\boldsymbol{x}$ と $\boldsymbol{y}$ の次元が同じならば、連立一次方程式として捉えることができた

$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

そして、このような形の連立方程式を解くことは、「$\boldsymbol{y}$ から $\boldsymbol{x}$ を推定する」という逆問題を解くことに相当する

一方、$\boldsymbol{y} = A\boldsymbol{x}$ という式は、線形写像を表す式とみることもできる
一般に、線形写像 $\boldsymbol{y} = A\boldsymbol{x}$ の表現行列 A は $m \times n$ 行列であり、$\boldsymbol{y}$ は m 次元ベクトル、$\boldsymbol{x}$ は n 次元ベクトルである

ここでは、$\boldsymbol{x}$ と $\boldsymbol{y}$ の次元が異なる場合の、「$\boldsymbol{y}$ から $\boldsymbol{x}$ を推定する」という逆問題を考えてみることにする

手がかりが足りない場合

手がかりとなる情報を $\boldsymbol{y}$ とし、知りたい情報を $\boldsymbol{x}$ とする

まずは、$\boldsymbol{y}$ の方が $\boldsymbol{x}$ より次元が小さい、すなわち $m < n$ の場合を考える
このとき、表現行列 A は横長の行列となる

$$\begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} = \begin{pmatrix} a_{11} & \cdots & \cdots & a_{1n} \\ \vdots & \ddots & \vdots & \vdots \\ a_{m1} & \cdots & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ \vdots \\ x_n \end{pmatrix}$$

$m < n$ の場合は、「知りたい量が n 個もあるのに、手がかりはたった m 個しかない」という状況になっている

見方を変えると、A を作用させたことによって「情報の一部が欠落してしまった」ともいえる

$m < n$ の場合の線形写像の写し方

$m < n$ のとき、A は、元より次元の低い空間に写す線形写像を表す
そのため、$\boldsymbol{x}$ はこの線形写像によって「潰される」ことになる

「潰される」とはどのようなことかというと、空間を張る $\boldsymbol{x}$ それぞれの居場所を、写す先では全員分用意することができないので、

> 複数の $\boldsymbol{x}$ を同じ $\boldsymbol{y}$ に写すしかない

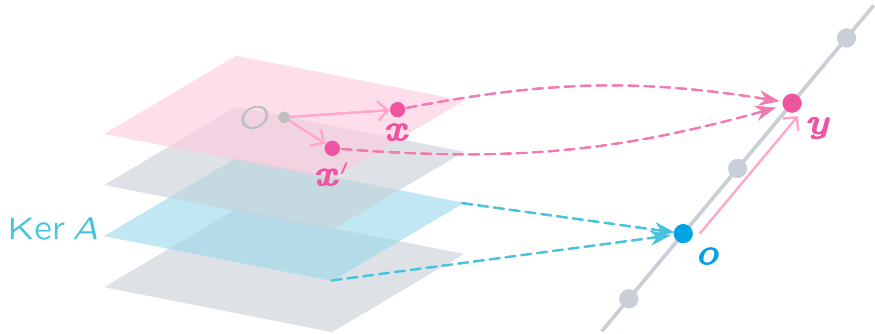
ということである

複数の $\boldsymbol{x}$ が同じ $\boldsymbol{y}$ に写ってしまうと、結果 $\boldsymbol{y}$ から元の $\boldsymbol{x}$ を特定することはできなくなる（つまり、情報が失われている）

線形写像の核

次の図は、1×3 行列 A による線形写像を表している

同じ平面上の点がすべて同じ点に写されることで、平面の集まりである立体（3 次元）が、点の集まりである直線（1 次元）へと「潰されている」ことがわかる



このとき、$A\boldsymbol{x} = \boldsymbol{o}$ に写ってくるような $\boldsymbol{x}$ の集合を、A の核あるいはカーネルといい、Ker A と表す

Ker A の次元

上の図では、零ベクトル $\boldsymbol{o}$（写り先の 1 次元空間の原点）に潰されている青い平面が Ker A に相当する

平面なので、この Ker A は 2 次元である

もしも $m < n$ でない場合、つまり潰れない写像の場合は、$A\boldsymbol{x} = \boldsymbol{o}$ に写る $\boldsymbol{x}$ は零ベクトル $\boldsymbol{o}$ だけなので、Ker A は 0 次元となる

Ker A に平行な方向の成分

[Todo 15: book: プログラミングのための線形代数 p114]

Ker f の定義

A が線形写像 f の表現行列であるとすると、Ker f を次のように定義できる

線形写像の核　線形写像 $f\colon V \to W$ に対して、f による $\{\boldsymbol{o}\}$ の逆像 $f^{-1}(\{\boldsymbol{o}\})$ を、線形写像 f の核や核空間、あるいはカーネルといい、$\mathrm{Ker}(f)$ と表記する

$$\mathrm{Ker}(f) = f^{-1}(\{\boldsymbol{o}\}) = \{\boldsymbol{v} \in V \mid f(\boldsymbol{v}) = \boldsymbol{o}\} \subset V$$

手がかりが多すぎる場合

今度は、$\boldsymbol{y}$ の方が $\boldsymbol{x}$ より次元が大きい、すなわち $m > n$ の場合を考える

このとき、表現行列 A は縦長の行列となる

$$\begin{pmatrix} y_1 \\ \vdots \\ \vdots \\ y_m \end{pmatrix} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

$m > n$ の場合は、「知りたい量はたった n 個しかないのに、手がかりが m 個もある」という状況になっている

この場合、手がかりどうしが矛盾することもある

$m > n$ の場合の線形写像の写し方

$m > n$ のとき、A は、元より次元の高い空間に写す線形写像を表す

そのため、写り先の空間すべてをカバーすることはできない

はみ出した $\boldsymbol{y}$ については、

> そこに写ってきてくれる $\boldsymbol{x}$ が存在しない

ことになる

現実の応用では、ノイズがのることで、はみ出した $\boldsymbol{y}$ が観測されることがある

そうなると、「手がかり $y_1, \ldots, y_m$ すべてに符号する $\boldsymbol{x}$ は存在しない」ということになってしまう

線形写像の像

与えられた A に対して、$\boldsymbol{x}$ をいろいろ動かしたときに A で写り得る $\boldsymbol{y} = A\boldsymbol{x}$ の集合を A の**像**といい、$\operatorname{Im} A$ で表す

別の言い方をすると、$\mathrm{Im}\, A$ は、元の空間全体を A で写した領域である

$\mathrm{Im}\, A$ 上にない $\boldsymbol{y}$ については、$\boldsymbol{y} = A\boldsymbol{x}$ となるような $\boldsymbol{x}$ は存在しない

$\mathrm{Im}\, f$ の定義

A が線形写像 f の表現行列であるとすると、$\mathrm{Im}\, f$ を次のように定義できる

線形写像の像　線形写像 $f\colon V \to W$ に対して、f による V の像 $f(V)$ を、線形写像 f の像や像空間といい、$\mathrm{Im}(f)$ と表記する

$$\mathrm{Im}(f) = f(V) = \{f(\boldsymbol{v}) \in W \mid \boldsymbol{v} \in V\} \subset W$$

単射と全射

ここまで、$\boldsymbol{y} = A\boldsymbol{x}$ という関係について、次の 2 つの観点で議論してきた

i. 同じ結果 $\boldsymbol{y}$ が出るような原因 $\boldsymbol{x}$ は唯一か
ii. どんな結果 $\boldsymbol{y}$ にも、それが出るような原因 $\boldsymbol{x}$ が存在するか

$\boldsymbol{y} = A\boldsymbol{x}$ を方程式と捉えると、(i) は解の一意性、(ii) は解の存在に対応する

単射

(i) は、次のようにも言い換えられる

i. 異なる原因 $\boldsymbol{x}, \boldsymbol{x}'$ が、A で同じ結果に写ることがないか

(i) の条件が成り立つとき、「線形写像 $\boldsymbol{y} = A\boldsymbol{x}$ は単射である」という

全射

(ii) は、次のようにも言い換えられる

ii. 元の空間全体（定義域）を A で写した領域 $\operatorname{Im} A$ が、行き先の空間全体（値域）に一致するか

(ii) の条件が成り立つとき、「線形写像 $\boldsymbol{y} = A\boldsymbol{x}$ は全射である」という

全単射

(i) と (ii) の両方が成り立つときは、「線形写像 $\boldsymbol{y} = A\boldsymbol{x}$ は全単射である」という

零ベクトルと単射性

零写像と射影を除けば、f によってベクトルが「つぶれない」という性質は、次のように表せる

$$\boldsymbol{v} \neq 0 \Longrightarrow f(\boldsymbol{v}) \neq \boldsymbol{o}$$

[Todo 16: book: 行列と行列式の基礎 p55 例 2.1.15]

この条件は、実は線形写像が単射であることを意味している

対偶をとって、次のように表現できる

零ベクトルへの写像による単射性の判定　線形写像 f が単射であることと次は同値である

$$f(\boldsymbol{v}) = \boldsymbol{o} \Longrightarrow \boldsymbol{v} = \boldsymbol{o}$$

証明

i. f が単射
ii. $f(\boldsymbol{v}) = \boldsymbol{o} \Longrightarrow \boldsymbol{v} = \boldsymbol{o}$

(i) $\Longrightarrow$ (ii)

零ベクトルの像は零ベクトルであることから、$f(\boldsymbol{v}) = \boldsymbol{o}$ は、

$$f(\boldsymbol{v}) = f(\boldsymbol{o})$$

と書き換えられる

f の単射性により、この式から、

$$\boldsymbol{v} = \boldsymbol{o}$$

がしたがう ■

(ii) $\Longrightarrow$ (i)

$f(\boldsymbol{v}_1) = f(\boldsymbol{v}_2)$ を満たす $\boldsymbol{v}_1, \boldsymbol{v}_2 \in \mathbb{R}^n$ を考える

このとき、f の線形性から、

$$f(\boldsymbol{v}_1 - \boldsymbol{v}_2) = f(\boldsymbol{v}_1) - f(\boldsymbol{v}_2)$$

となる

仮定 (ii) より、

$$f(\boldsymbol{v}_1 - \boldsymbol{v}_2) = \boldsymbol{o} \Longrightarrow \boldsymbol{v}_1 - \boldsymbol{v}_2 = \boldsymbol{o}$$

がいえるので、$\boldsymbol{v}_1 = \boldsymbol{v}_2$ が成り立つ

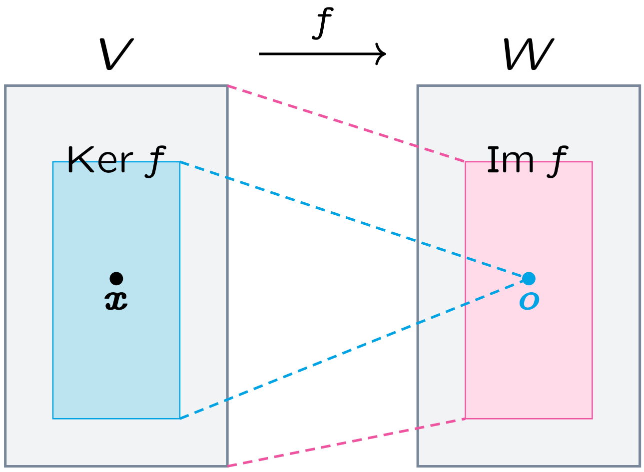
$f(\boldsymbol{v}_1) = f(\boldsymbol{v}_2)$ から $\boldsymbol{v}_1 = \boldsymbol{v}_2$ が導かれたことで、f は単射であることが示された ■

核・像と単射・全射

先ほどの定理で、線形写像 f によって「潰れない」という条件が、単射性と同値であることが示された

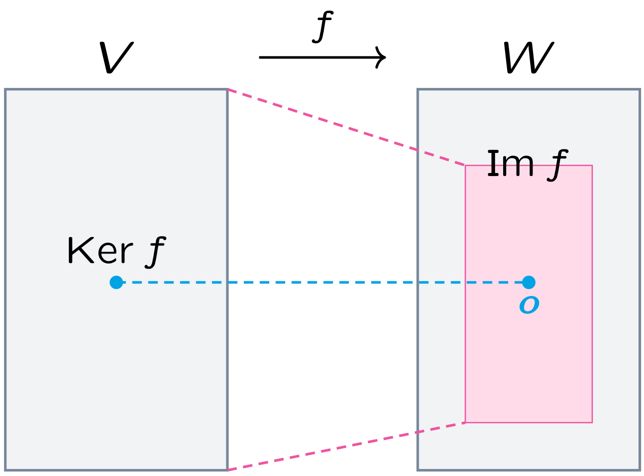
つまり、線形写像 f の核 $\mathrm{Ker}\, f$ が、f の単射性と関係しそうである

また、線形写像 f の像 $\mathrm{Im}\, f$ が値域と一致するかどうかが、f の全射性と関係する



単射となるときの核

線形写像 f が単射であるとは、「潰れない」ということなので、次のような状況である



つまり、$\mathrm{Ker}\, f$ が零ベクトル $\boldsymbol{o}$ のみを含む状態であればよい

線形写像の単射性と核の関係　f を線形写像とするとき、

$$f \text{ が単射} \iff \mathrm{Ker}\, f = \{\boldsymbol{o}\}$$

証明

$\mathrm{Ker}\, f$ の定義は

$$\mathrm{Ker}\, f = \{\boldsymbol{v} \in V \mid f(\boldsymbol{v}) = \boldsymbol{o}\}$$

これを踏まえて、次の 2 つが同値であることを示す

i. $f(\boldsymbol{v}) = \boldsymbol{o} \Longrightarrow \boldsymbol{v} = \boldsymbol{o}$

ii. $\operatorname{Ker} f = \{\boldsymbol{o}\}$

(i) $\Longrightarrow$ (ii)

このとき、$f(\boldsymbol{v}) = \boldsymbol{o}$ が $\boldsymbol{v} = \boldsymbol{o}$ を意味するので、$\operatorname{Ker} f$ の元は零ベクトルのみになる

よって、$\operatorname{Ker} f = \{\boldsymbol{o}\}$ が成り立つ ■

(ii) $\Longrightarrow$ (i)

$\operatorname{Ker} f = \{\boldsymbol{o}\}$ であれば、$\operatorname{Ker} f$ の元は零ベクトルのみである

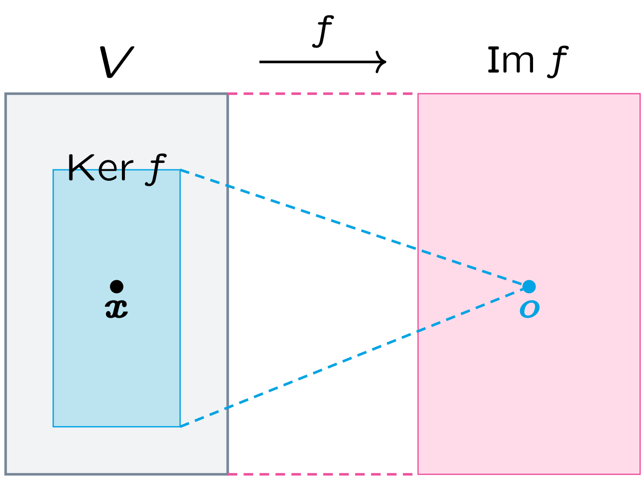
よって、$f(\boldsymbol{v}) = \boldsymbol{o}$ が成り立つとき、$\boldsymbol{v} = \boldsymbol{o}$ が成り立つことになる

すなわち、$f(\boldsymbol{v}) = \boldsymbol{o} \Longrightarrow \boldsymbol{v} = \boldsymbol{o}$ が成り立つ ■

全射となるときの像

線形写像 f が全射であるとは、$\operatorname{Im} f$ が行き先の空間全体を埋め尽くす状態である

このような状態であれば、たしかに $f(\boldsymbol{x})$ が $\operatorname{Im} f$ からはみ出してしまうことはない



この状況を式で表すと、線形写像 $f\colon V \to W$ が全射であるとは、

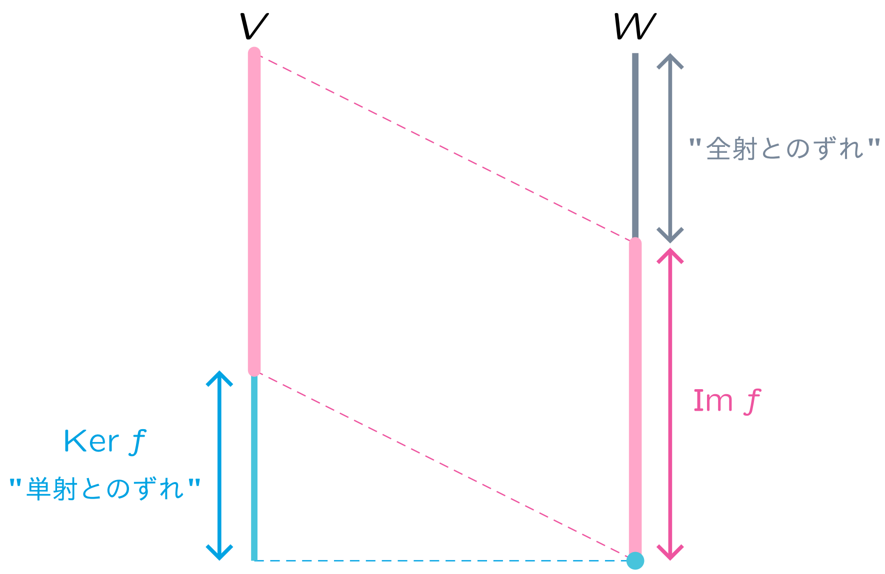
$$\operatorname{Im} f = W$$

という条件と同値である

単射・全射との離れ具合

Ker f が零ベクトルの集合に一致するなら f は単射であり、Im f が写り先全体に一致するなら f は全射である

このことから、Ker f と Im f は、それぞれ単射・全射と「どれくらいかけ離れているか」を測る尺度とも捉えられる



線形写像の像と線型独立性

⚓ 線形写像と線形独立性　$f\colon \mathbb{R}^n \to \mathbb{R}^m$ を線形写像、$\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n \in \mathbb{R}^n$ とする

ベクトル $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$ の f による像

$$f(\boldsymbol{v}_1), f(\boldsymbol{v}_2), \ldots, f(\boldsymbol{v}_n)$$

が線型独立であるとき、$\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ も線型独立である

証明

$\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$ の線形結合

$$c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \cdots + c_n\boldsymbol{v}_n = \boldsymbol{0}$$

を考える

この両辺を f で写すと、f の線形性と零ベクトルの像 $f(\boldsymbol{0}) = \boldsymbol{0}$ を使って

$$c_1 f(\boldsymbol{v}_1) + c_2 f(\boldsymbol{v}_2) + \cdots + c_n f(\boldsymbol{v}_n) = f(\boldsymbol{0}) = \boldsymbol{0}$$

仮定より $f(\boldsymbol{v}_1), f(\boldsymbol{v}_2), \ldots, f(\boldsymbol{v}_n)$ は線型独立なので、$c_1 = c_2 = \cdots = c_n = 0$ である

よって、

$$c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \cdots + c_n\boldsymbol{v}_n = \boldsymbol{0}$$

を満たす $c_1, c_2, \ldots, c_n$ は 0 しかないので、$\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ は線型独立である

■

次の定理は、平行なベクトルを線型写像で写した結果、平行でなくなったりはしないということを述べている

線形写像と線形従属性　$f\colon \mathbb{R}^n \to \mathbb{R}^m$ を線形写像、$\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n \in \mathbb{R}^n$ とする

$\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ が線形従属ならば、$\{f(\boldsymbol{v}_1), \ldots, f(\boldsymbol{v}_n)\}$ は線形従属である

証明

$\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ が線形従属であるとは、少なくとも 1 つは 0 でないある定数 $k_1, k_2, \ldots, k_n$ が存在して

$$k_1\boldsymbol{v}_1 + k_2\boldsymbol{v}_2 + \cdots + k_n\boldsymbol{v}_n = \boldsymbol{0}$$

が成り立つことを意味する

この両辺を f で写すと、線形性より

$$k_1 f(\boldsymbol{v}_1) + k_2 f(\boldsymbol{v}_2) + \cdots + k_n f(\boldsymbol{v}_n) = f(\mathbf{0}) = \mathbf{0}$$

が成り立つ

よって、$\{f(\boldsymbol{v}_1), f(\boldsymbol{v}_2), \ldots, f(\boldsymbol{v}_n)\}$ も線形従属である ■

たとえば平行四辺形の像が線分や 1 点になったりしないことなどは、「$\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ が線型独立ならば、$\{f(\boldsymbol{v}_1), f(\boldsymbol{v}_2), \ldots, f(\boldsymbol{v}_n)\}$ も線型独立である」と表現できる

単射な線型写像は線型独立性を保つ　線型写像 $f\colon \mathbb{R}^n \to \mathbb{R}^m$ が単射であるとき、$\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ が線型独立ならば、$\{f(\boldsymbol{v}_1), f(\boldsymbol{v}_2), \ldots, f(\boldsymbol{v}_n)\}$ も線型独立である

証明

$f(\boldsymbol{v}_1), f(\boldsymbol{v}_2), \ldots, f(\boldsymbol{v}_n)$ の線形結合

$$c_1 f(\boldsymbol{v}_1) + c_2 f(\boldsymbol{v}_2) + \cdots + c_n f(\boldsymbol{v}_n) = \mathbf{0}$$

を考える

f の線形性と零ベクトルの像 $f(\mathbf{0}) = \mathbf{0}$ より、次のように書き換えられる

$$f(c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \cdots + c_n\boldsymbol{v}_n) = \mathbf{0} = f(\mathbf{0})$$

f は単射だから、上式より

$$c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \cdots + c_n\boldsymbol{v}_n = \mathbf{0}$$

が成り立つ

ここで、$\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$ は線型独立なので、$c_1 = c_2 = \cdots = c_n = 0$ である

よって、$f(\boldsymbol{v}_1), f(\boldsymbol{v}_2), \ldots, f(\boldsymbol{v}_n)$ は線型独立である ■

第 6 章

正則な線形変換と逆行列

線形変換と逆問題

$\boldsymbol{y} = A\boldsymbol{x}$ という形の式は、$\boldsymbol{x}$ と $\boldsymbol{y}$ の次元が同じならば、連立一次方程式として捉えることができた。

$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

そして、このような形の連立方程式を解くことは、「$\boldsymbol{y}$ から $\boldsymbol{x}$ を推定する」という逆問題を解くことに相当する。

一方、$\boldsymbol{y} = A\boldsymbol{x}$ という式は、線形写像を表す式とみることもできる。
特に、$\mathbb{R}^n$ からそれ自身への線形写像 f を $\mathbb{R}^n$ の線形変換と呼ぶのだった。

言い換えると、表現行列 A で表される線形写像 $\boldsymbol{y} = A\boldsymbol{x}$ が線形変換と呼べるのは、$\boldsymbol{x}$ と $\boldsymbol{y}$ の次元が同じ場合である。
このように、線形変換と連立一次方程式を関連づけて考えることができる。

逆行列

「写り先 $\boldsymbol{y}$ から元の点 $\boldsymbol{x}$ を答える」という写像に対応する行列を逆行列といい、A^{-1} と表す。

この行列 A^{-1} は、

- どんな $\boldsymbol{x}$ を持ってきても、$A\boldsymbol{x} = \boldsymbol{y}$ ならば $A^{-1}\boldsymbol{y} = \boldsymbol{x}$
- どんな $\boldsymbol{y}$ を持ってきても、$A^{-1}\boldsymbol{y} = \boldsymbol{x}$ ならば $A\boldsymbol{x} = \boldsymbol{y}$

となるような行列である。

$$\begin{tikzcd} \boldsymbol{x} \arrow[r, bend left, "A"] & \boldsymbol{y} \arrow[l, bend left, "A^{-1}"] \end{tikzcd}$$

別の言い方をすると、

- A して A^{-1} したら元に戻る
- A^{-1} して A したら元に戻る

となるような行列 A^{-1} を逆行列として定義する。

逆行列　正方行列 A に対して、次式を満たす行列 X を A の逆行列といい、A^{-1} と表す。

$$AX = XE = E$$

正則性と全単射性

A の逆行列は、いつでも存在するとは限らない。

正則（行列の言葉で）　正方行列 A の逆行列が存在するとき、A は正則であるという。

A の逆行列が存在するには、A が表す写像が全単射である、つまり A によって「潰れない・はみ出さない」ことが必要である。

- 潰れてしまえば、元の $\boldsymbol{x}$ はわからない（単射でない場合）
- はみ出してしまえば、元の $\boldsymbol{x}$ は存在しない（全射でない場合）

正則（写像の言葉で）　線形変換 f が全単射であるとき、f は正則であるという。正方行列 A が正則な線形変換を与えるとき、A は正則行列であるという。

逆写像と逆行列の対応

一般に、写像 f が全単射であれば、逆写像 f^{-1} が存在する。

逆写像の線形性　f を $\mathbb{R}^n$ の正則な線形変換とするとき、逆写像 f^{-1} は線形である

証明

$\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$, $c \in \mathbb{R}$ とし、次の 2 つを示せばよい

i. $f^{-1}(\boldsymbol{x}+\boldsymbol{y}) = f^{-1}(\boldsymbol{x}) + f^{-1}(\boldsymbol{y})$
ii. $f^{-1}(c\boldsymbol{x}) = cf^{-1}(\boldsymbol{x})$

(i)

$f \circ f^{-1}$ は恒等写像であるから、

$$\begin{aligned}\boldsymbol{x} &= f \circ f^{-1}(\boldsymbol{x}) \\ \boldsymbol{y} &= f \circ f^{-1}(\boldsymbol{y}) \\ \boldsymbol{x}+\boldsymbol{y} &= f \circ f^{-1}(\boldsymbol{x}+\boldsymbol{y})\end{aligned}$$

また、f は線形写像であるから、

$$f \circ f^{-1}(\boldsymbol{x}+\boldsymbol{y}) = f(f^{-1}(\boldsymbol{x}) + f^{-1}(\boldsymbol{y}))$$

$f \circ f^{-1}(\boldsymbol{v})$ は、$f(f^{-1}(\boldsymbol{v}))$ を意味する記号なので、

$$f(f^{-1}(\boldsymbol{x}+\boldsymbol{y})) = f(f^{-1}(\boldsymbol{x}) + f^{-1}(\boldsymbol{y}))$$

両辺を f^{-1} で写すと、

$$f^{-1}(\boldsymbol{x}+\boldsymbol{y}) = f^{-1}(\boldsymbol{x}) + f^{-1}(\boldsymbol{y})$$

となり、(i) が示された ■

(ii)

$f \circ f^{-1}$ は恒等写像であるから、

$$\begin{aligned} \boldsymbol{x} &= f \circ f^{-1}(\boldsymbol{x}) = f(f^{-1}(\boldsymbol{x})) \\ c\boldsymbol{x} &= f \circ f^{-1}(c\boldsymbol{x}) = f(f^{-1}(c\boldsymbol{x})) \end{aligned}$$

$\boldsymbol{x} = f(f^{-1}(\boldsymbol{x}))$ の両辺に c をかけた、次も成り立つ

$$c\boldsymbol{x} = cf(f^{-1}(\boldsymbol{x}))$$

さらに、f は線形写像であるから、

$$cf(f^{-1}(\boldsymbol{x})) = f(cf^{-1}(\boldsymbol{x}))$$

ここまでの $c\boldsymbol{x}$ の複数の表現により、次式が成り立つ

$$f(f^{-1}(c\boldsymbol{x})) = f(cf^{-1}(\boldsymbol{x}))$$

両辺を f^{-1} で写すと、

$$f^{-1}(c\boldsymbol{x}) = cf^{-1}(\boldsymbol{x})$$

となり、(ii) が示された ■

n 次正則行列 A は、正則な線形変換 $f\colon \mathbb{R}^n \to \mathbb{R}^n$ と対応している。

逆写像 f^{-1} が存在し、線形であるから、ある n 次正方行列 B が対応するはずである。

$f \circ f^{-1} = f^{-1} \circ f = \mathrm{id}_{\mathbb{R}^n}$ であり、線形写像の合成は行列の積に対応するから、

$$AB = BA = E$$

が成り立つ。

このように、逆写像の性質から、逆行列の定義式を導くこともできる。

逆行列の一意性

逆行列は、存在するとしてもただ 1 つしか存在しない。

逆行列の一意性　正方行列 A に対して、A の逆行列が存在するならば、それは一意的である。

証明

A の逆行列が B_1 と B_2 の 2 つあるとする。

$$AB_1 = B_1A = E \quad \text{かつ} \quad AB_2 = B_2A = E$$

$AB_2 = E$ の両辺に B_1 をかけると、

$$B_1 = B_1AB_2 = (B_1A)B_2 = EB_2 = B_2$$

よって、$B_1 = B_2$ となり、逆行列は一意的である。 ■

逆行列による連立一次方程式の解

正則行列 A に対して、方程式 $A\boldsymbol{x} = \boldsymbol{b}$ のただ 1 つの解は次で与えられる。

$$\boldsymbol{x} = A^{-1}\boldsymbol{b}$$

これが「ただ 1 つ」の解といえるのは、係数行列 A が与えられれば、その逆行列 A^{-1} は一意的に定まるからである。

つまり、A が正則行列であり、その逆行列 A^{-1} が求まれば、行列のかけ算によって連立一次方程式の解が求められる。

行列の演算と逆行列

逆行列の逆行列

「A の取り消し」を取り消すには、A すればよい。

逆行列に対する逆行列　正則行列 A の逆行列 A^{-1} は正則であり、その逆行列は A である。

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

証明

A の逆行列が A^{-1} であることから、

$$AA^{-1} = A^{-1}A = E$$

この式は、A^{-1} が正則であり、その逆行列が A であることを示す式でもある。 ■

行列の積の逆行列

「B して A したもの」を元に戻すには、まず A を取り消してから B を取り消す必要がある。

正則行列の積に対する逆行列　正則行列 $A,\ B$ の積 AB は正則であり、その逆行列は次のようになる。

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

証明

$$\begin{aligned}(AB)(B^{-1}A^{-1}) &= A(BB^{-1})A^{-1} \\ &= AEA^{-1} \\ &= E\end{aligned}$$

であり、同様に

$$\begin{aligned}(B^{-1}A^{-1})(AB) &= B^{-1}(A^{-1}A)B \\ &= B^{-1}EB \\ &= E\end{aligned}$$

であるので、

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

が成り立つ。 ■

転置行列の正則性

正則行列の転置の正則性　正則行列 A に対して、その転置行列 ${}^tA$ も正則である。

証明

A が正則であることから、その逆行列 A^{-1} が存在し、

$$A^{-1}A = E$$

両辺の転置をとると、右辺の単位行列は転置しても単位行列であり、左辺には正則行列の積に対する逆行列の公式を用いて、

$${}^t(A^{-1}A) = {}^tA{}^t(A^{-1}) = E$$

この等式より、${}^tA$ の逆行列は ${}^t(A^{-1})$ であることがわかる。 ■

三角行列の正則性

上三角行列の正則性　対角成分がすべて 0 でない上三角行列は正則である。

証明

[Todo 17: book: 行列と行列式の基礎 p74 命題 2.4.9]

正則な上三角行列の逆行列　正則な上三角行列は、その逆行列も上三角行列である。

証明

[Todo 18:]

正則な上三角行列と関連して、次の事実が成り立つ。

行基本変形と対角行列　正則行列 A に対して、行のスカラー倍以外の行基本変形を繰り返し行って対角行列にできる。

証明

[Todo 19: book: 行列と行列式の基礎 p75 命題 2.4.12]

正則行列と対角行列

[Todo 20: book: プログラミングのための線形代数 p46〜47]

ブロック対角行列の正則性　次のようなブロック対角行列 M において、対角ブロック A, B が正則であれば、M も正則である。

$$\begin{array}{rcl} & \begin{array}{cc} \quad \ell \quad & \quad n-\ell \quad \end{array} & \\ M = & \left(\begin{array}{c|c} \quad A \quad & \quad O \quad \\ \hline \quad O \quad & \quad B \quad \end{array}\right) & \begin{array}{l} \ell \\ n-\ell \end{array} \end{array}$$

証明

A と B が正則であるから、逆行列 A^{-1} と B^{-1} が存在する。

それらを用いて、次のような積を考える。

$$\begin{aligned} \begin{pmatrix} A & O \\ O & B \end{pmatrix} \begin{pmatrix} A^{-1} & O \\ O & B^{-1} \end{pmatrix} &= \begin{pmatrix} AA^{-1} & O \\ O & BB^{-1} \end{pmatrix} \\ &= \begin{pmatrix} E_\ell & O \\ O & E_{n-\ell} \end{pmatrix} \\ &= E_n \end{aligned}$$

この等式は、M の逆行列の存在を示している。

$$M \begin{pmatrix} A^{-1} & O \\ O & B^{-1} \end{pmatrix} = E_n$$

つまり、対角ブロックがそれぞれ正則であれば、それらの逆行列を並べることで全体の逆行列が構成できる。

このようにして、M が正則であることがわかる。 ■

正則の判定

[Note 2: 「線型独立な列ベクトルと階数」の章に移動予定]

列ベクトルの線型独立性による正則の判定　n 次正方行列

$$A = (\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n)$$

に対して、次が成り立つ

$$A \text{ が正則行列} \iff \boldsymbol{a}_1, \ldots, \boldsymbol{a}_n \text{が線型独立}$$

証明

$\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n \in \mathbb{R}^n$ が線型独立であることは、

$$\mathrm{rank}(A) = n$$

と同値であることを以前示した

さらに、先ほど示した定理より、$\mathrm{rank}(A) = n$ は A が正則行列であることと同値である　■

逆行列の計算法

[Note 3: 「基本変形と基本行列」の章で扱う定理と被るので、削除予定]

正則行列 A の逆行列を計算するために、次の定理に注目しよう

逆行列の計算法の原理　正方行列 A に対して、$AB = E$ を満たす正方行列 B があるならば、A は正則であり、B は A の逆行列である

証明

[Todo 21: book: 行列と行列式の基礎 p72 命題 2.4.6]

上の定理の証明は、逆行列の計算法のヒントを含んでいる

A の逆行列 B を求めるには、n 個の線形方程式

$$A\boldsymbol{b}_i = \boldsymbol{e}_i \quad (1 \leq i \leq n)$$

を解けばよい

A は階数 n の n 次正方行列なので、行変形で A から E に到達することができる

$\boldsymbol{b}_i$ を求めるには、行変形により

$$(A \mid \boldsymbol{e}_i) \to \cdots \to (E \mid \boldsymbol{b}_i)$$

とすればよい

i ごとに掃き出し法を何度も実行しないといけないのかと思いきや、一度にまとめられる

$$\begin{gathered}(A \mid E) = (A \mid \boldsymbol{e}_1, \cdots, \boldsymbol{e}_n) \to \\ \cdots \to (E \mid \boldsymbol{b}_1, \cdots, \boldsymbol{b}_n) = (E \mid B)\end{gathered}$$

このようにすれば、行変形は 1 通りで十分である

第 7 章

基本変形と基本行列

行基本変形と基本行列

基本変形を行列のかけ算によって実現することができる

> **基本行列**　基本変形 α を単位行列 E に行った結果を E_α とするとき、E_α を α に対応する基本行列と呼ぶ

行基本変形とは、次の 3 種類の操作であった

i. 2 つの行を交換する
ii. ある行に 0 でない数をかける
iii. ある行の定数倍を他の行に加える

これらに対応して、行基本変形を表現する基本行列は、次の 3 種類がある

i. $F(i, j)$: E の i 行と j 行を交換したもの（$i \neq j$）
ii. $G(i; c)$: E の (i, j) 成分を 1 から c に置き換えたもの（$c \neq 0$）
iii. $H(i, j; c)$: E の (i, j) 成分を 0 から c に置き換えたもの（$i \neq j$）

$$F(i,j)=\begin{pmatrix}1 & & & & & & \\ & \ddots & & & & & \\ & & 0 & \dots & 1 & & \\ & & \vdots & & \vdots & & \\ & & 1 & \dots & 0 & & \\ & & & & & \ddots & \\ & & & & & & 1\end{pmatrix}$$

$$G(i;c)=\begin{pmatrix}1 & & & & & & \\ & \ddots & & & & & \\ & & c & & & & \\ & & & \ddots & & & \\ & & & & \ddots & & \\ & & & & & \ddots & \\ & & & & & & 1\end{pmatrix}$$

$$H(i,j;c)=\begin{pmatrix}1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & \dots & c & & \\ & & & \ddots & \vdots & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1\end{pmatrix}$$

行に関する基本変形は、基本行列を左からかけることに他ならない

基本行列による行基本変形の表現　行列 A に行基本変形 α を行って得られる行列を B とすると、

$$B=E_\alpha A$$

証明

$\boldsymbol{e}_k$ を k 列目が 1 で他が 0 の横ベクトルとし、A の k 行目の行ベクトルを $\boldsymbol{a}_k$ とする

行の交換

基本行列 $F(i,j)$ の k 行目は、

$$(F(i,j))_{k,*} = \begin{cases} \boldsymbol{e}_j & (k=i) \\ \boldsymbol{e}_i & (k=j) \\ \boldsymbol{e}_k & (k \neq i,j) \end{cases}$$

よって、$F(i,j)A$ の k 行目は、

$$(F(i,j)A)_{k,*} = \begin{cases} \boldsymbol{a}_j & (k=i) \\ \boldsymbol{a}_i & (k=j) \\ \boldsymbol{a}_k & (k \neq i,j) \end{cases}$$

となり、i 行目と j 行目が交換されていることがわかる ■

行の定数倍

基本行列 $G(i;c)$ の k 行目は、

$$(G(i;c))_{k,*} = \begin{cases} c\boldsymbol{e}_i & (k=i) \\ \boldsymbol{e}_k & (k \neq i) \end{cases}$$

よって、$G(i;c)A$ の k 行目は、

$$(G(i;c)A)_{k,*} = \begin{cases} c\boldsymbol{a}_i & (k=i) \\ \boldsymbol{a}_k & (k \neq i) \end{cases}$$

となり、i 行目が c 倍されていることがわかる ■

行の定数倍の加算

基本行列 $H(i,j;c)$ の k 行目は、

$$(H(i,j;c))_{k,*} = \begin{cases} \boldsymbol{e}_i + c\boldsymbol{e}_j & (k=i) \\ \boldsymbol{e}_j & (k=j) \\ \boldsymbol{e}_k & (k \neq i,j) \end{cases}$$

よって、$H(i,j;c)A$ の k 行目は、

$$(H(i,j;c)A)_{k,*} = \begin{cases} \boldsymbol{a}_i + c\boldsymbol{a}_j & (k=i) \\ \boldsymbol{a}_j & (k=j) \\ \boldsymbol{a}_k & (k \neq i,j) \end{cases}$$

となり、i 行目に j 行目の c 倍が加えられていることがわかる ■

列基本変形と基本行列

行基本変形と同様に、列に関する基本変形を考えることもできる

i. 2 つの列を交換する
ii. ある列に 0 でない数をかける
iii. ある列の定数倍を他の列に加える

列に関する基本変形は、基本行列を右からかけることで実現できる

基本行列による列基本変形の表現　行列 A に列基本変形 α を行って得られる行列を B とすると、

$$B = AE_\alpha$$

証明

転置すると A になるような行列 A' を考える

$$A' = {}^t(A)$$

転置すると行と列が入れ替わるので、A' に「行」基本変形を施した行列を転置すれば、A に同じ基本変形を列に関して施した行列が得られる

適用したい基本変形を α とし、これを列に関して施す基本行列が E_α なら、これを行に関して施す基本行列は ${}^t(E_\alpha)$ となる

よって、

$$B = {}^t({}^t(E_\alpha)A') = {}^t(A')^t({}^t(E_\alpha)) = AE_\alpha$$

というように、積の転置を取ると積の順序が入れ替わることから、行基本変形の場合とは積の順序が逆転することがいえる ■

基本行列の正則性

行基本変形も列基本変形も、基本行列によって定式化できる

この考えをさらに進めるため、基本行列の性質を述べる

基本行列の正則性 基本行列は正則である

証明

基本行列の表す変形を考えれば、

$$\begin{aligned} F(i,j)F(i,j) &= E \\ G(i;c)G(i;\frac{1}{c}) = G(i;\frac{1}{c})G(i;c) &= E \\ H(i,j;c)H(i;-c) = H(i,;-c)H(i,j;c) &= E \end{aligned}$$

が成り立つことがわかる

したがって、基本行列は逆行列を持つので正則である ■

つまり、各々の基本変形は可逆の変形、すなわち逆に戻ることのできる変形である

基本行列の積と逆行列

行基本変形が基本行列を左からかけることに対応することから、行基本変形とは線形写像であり、基本行列はその表現行列であるという見方もできる

そのため、行基本変形の合成は、基本行列の積として表現できる

このことから、行についての連続する複数の基本変形の繰り返しも可逆であることがいえる

基本行列の積による行変形の構成　行列の行変形 $A \to B$ に対し、$B = PA$ を満たす正則行列 P が存在する

このとき、P はいくつかの基本行列の積である

証明

行基本変形を $A \xrightarrow{\alpha_k} \cdots \xrightarrow{\alpha_1} B$ と合成して得られる行変形は、$E_{\alpha_1} \cdots E_{\alpha_k}$ を左からかけることで実現される

すなわち、

$$B = E_{\alpha_1} \cdots E_{\alpha_k} A$$

が成り立つ

個々の基本行列 $E_{\alpha_1}, \ldots, E_{\alpha_k}$ は正則であるので、これらの積 $P = E_{\alpha_1} \cdots E_{\alpha_k}$ も正則である ■

上の証明から、正則行列 P に対して、その逆行列を P^{-1} とすると、

$$P^{-1}B = P^{-1}E_{\alpha_1} \cdots E_{\alpha_k} A = P^{-1}PA = A$$

が成り立つことになる

ここで、$B = E$ の場合を考えると、$P^{-1}E = A$ となるので、次のことがいえる

単位行列への行変形による逆行列の構成　正方行列 A の単位行列への行変形 $A \to E$ に対応する基本変形の積は、A の逆行列を与える

つまり、任意の正方行列は行基本変形だけで単位行列に変形でき、その基本行列の積から逆行列を求めることができる

この章で得られた定理を組み合わせると、次の定理が得られる

基本行列の積による正則行列の表現　任意の正則行列はいくつかの基本行列の積である

証明

A を正則行列とすると、A の逆行列 A^{-1} は行変形 $A \to E$ に対応する基本変形の積によって与えられる

さらに、基本本行列の積による行変形の構成より、行変形 $A \to E$ に対し、

$$E = PA$$

を満たす正則行列 P が存在する

この等式より、$A^{-1} = P$ となり、P も基本行列の積であることがいえる　■

階数標準形

任意の行列 A は、行基本変形により、次のような既約行階段行列に変形できる

$$\begin{pmatrix} 0 & 1 & * & 0 & 0 & * & * & 0 \\ 0 & 0 & 0 & 1 & 0 & * & * & 0 \\ 0 & 0 & 0 & 0 & 1 & * & * & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

ここからさらに、列の交換によって、主成分のある列を左に集めることができる

$$\begin{array}{c c} \overleftrightarrow{\quad r \quad} \quad \overleftrightarrow{\quad n-r \quad} & \\ \left(\begin{array}{ccc|c} 1 & & & \\ & \ddots & & * \\ & & 1 & \\ \hline & O & & O \end{array}\right) & \begin{array}{l} \updownarrow r \\ \\ \updownarrow n-r \end{array} \end{array}$$

ここで、r は零行ではない行の個数、すなわち A の階数である

さらに、列の掃き出しで、左上のブロックの成分 $*$ をすべて 0 にできる

$$
\begin{array}{cc}
\begin{array}{cc} \quad r \quad & \quad n-r \quad \end{array} & \\
\left(\begin{array}{c|c} \begin{matrix} 1 & & \\ & \ddots & \\ & & 1 \end{matrix} & O \\ \hline O & O \end{array}\right) & \begin{array}{l} r \\ \\ n-r \end{array}
\end{array}
$$

この形を、A の階数標準形という

この形を得るまでの過程をまとめると、次のことがいえる

> **⚓ 基本変形による階数標準形の構成**　任意の行列は、行と列の基本変形を繰り返すことで、階数標準形に変形できる

ここで、P を行基本変形に対応する基本行列の積、Q を列基本変形に対応する基本行列の積とすると、A の階数標準形は PAQ で与えられる

基本行列の積は任意の正則行列を表すので、次のようにまとめられる

> **⚓ 正則行列による階数標準形の構成**　$m \times n$ 型行列 A に対して、行変形に対応する m 次正則行列 P、列変形に対応する n 次正則行列 Q が存在し、
>
> $$B = PAQ$$
>
> が階数標準形となる

第 8 章

行列の階数と解の性質

拡大係数行列と連立方程式の同値変形

A を m 行 n 列の行列、$\boldsymbol{b} \in \mathbb{R}^m$ とするとき、線形方程式

$$A\boldsymbol{x} = \boldsymbol{b}$$

は、n 個の文字 $x_1, \ldots, x_n$ に関する m 本の連立方程式である。

$\boldsymbol{x}$ は未知数 $x_1, x_2, \ldots, x_n$ を成分とするベクトルである。

このとき、A は方程式の係数行列と呼ばれ、A の右端に列ベクトル $\boldsymbol{b}$ を追加して得られる m 行 $(n+1)$ 列の行列

$$\tilde{A} = (A \mid \boldsymbol{b})$$

を拡大係数行列という。

掃き出し法による連立一次方程式の解法では、拡大係数行列 $\tilde{A}$ のうち、係数行列 A の部分を既約行階段形に変形することで、解を読み取ることを目指した。

$$\tilde{A}\text{の変形後} = \begin{array}{c} \begin{array}{ccccccc} j_1 & & j_2 & \cdots & j_r & & \end{array} \\ \left(\begin{array}{ccccccc|c} 1 & \star & 0 & \cdots & 0 & \star & \star & b_1 \\ 0 & 0 & 1 & \cdots & 0 & \star & \star & b_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & \star & \star & b_r \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 & b_{r+1} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 & b_m \end{array}\right) \\ n \end{array}$$

途中の行基本変形を表現する基本行列を順に $P_1, P_2, \ldots, P_k$ として、

$$P = P_k \cdots P_2 P_1$$

とおくと、行基本変形を施した後の行列は、

$$P\tilde{A} = P(A \mid \boldsymbol{b}) = (PA \mid P\boldsymbol{b})$$

すなわち、

$$PA = \begin{array}{c} \begin{array}{ccccccc} j_1 & & j_2 & \cdots & j_r & & \end{array} \\ \begin{pmatrix} 1 & \star & 0 & \cdots & 0 & \star & \star \\ 0 & 0 & 1 & \cdots & 0 & \star & \star \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & \star & \star \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 \end{pmatrix} \\ n \end{array}, \quad P\boldsymbol{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_r \\ b_{r+1} \\ \vdots \\ b_m \end{pmatrix}$$

と表せる。

つまり、

$$A\boldsymbol{x} = \boldsymbol{b} \Longrightarrow PA\boldsymbol{x} = P\boldsymbol{b}$$

であり、基本行列の積 P は正則であるから、両辺に左から P^{-1} をかけることで $A\boldsymbol{x} = \boldsymbol{b}$ に戻せるので、

$$A\boldsymbol{x} = \boldsymbol{b} \Longleftrightarrow PA\boldsymbol{x} = P\boldsymbol{b}$$

よって、与えられた連立方程式は、係数部分を既約行階段形に変形した形の連立方程式と同値である。

行基本変形によって得られる方程式の解は、
元の方程式の解と同じ

これが、掃き出し法による連立一次方程式の解法の根拠となる。
変形後の拡大係数行列から、解を読み取ればよい。

連立方程式の解のパターン

方程式を解くということは、次のような問題に答えることである。

A. **解の存在**：解は存在するか？
B. **解の一意性**：解が存在する場合、それはただ 1 つの解か？
C. **解の自由度**：解が複数存在する場合、どれくらい多く存在するのか？
D. **解のパラメータ表示**：解全体の集合をいかにしてわかりやすく表示できるか？

連立方程式において、解が 1 つに定まらない場合を不定、そもそも解が存在しない場合を不能と呼ぶ。
不定の場合は、問題 C と問題 D に答えることが方程式を「解く」ことになる。

不能かどうかを判断するのが問題 A、不定かどうかを判断するのが問題 B だが、これらは次のように変形された拡大係数行列から判断することができる。

$$\tilde{A}\text{の変形後} = \begin{array}{c} \begin{array}{ccccccc} j_1 & & j_2 & \cdots & j_r & & \end{array} \\ \left(\begin{array}{ccccccc|c} 1 & \star & 0 & \cdots & 0 & \star & \star & b_1 \\ 0 & 0 & 1 & \cdots & 0 & \star & \star & b_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & \star & \star & b_r \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 & b_{r+1} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 & b_m \end{array}\right) \\ \longleftrightarrow \\ \text{n} \end{array}$$

このような係数拡大行列において、方程式の解が存在するのは、

$$b_{r+1} = \cdots = b_m = 0$$

の場合に限る。

つまり、

$$\tilde{A}\text{の変形後} = \begin{array}{c} \begin{array}{ccccccc} j_1 & & j_2 & \cdots & j_r & & \end{array} \\ \left(\begin{array}{ccccccc|c} 1 & \star & 0 & \cdots & 0 & \star & \star & b_1 \\ 0 & 0 & 1 & \cdots & 0 & \star & \star & b_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & \star & \star & b_r \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \end{array}\right) \\ \longleftrightarrow \\ \text{n} \end{array}$$

というような形に変形された場合、$0 = 0$ という常に成り立つ等式の本数だけ解の不定性（解が定まらない変数）が残るが、解は存在する。

仮に $b_{r+1}, \ldots, b_m$ のうち、1 つでも 0 でないものがある場合、解は存在しない。

たとえば、

$$
\tilde{A}\text{の変形後} =
\begin{array}{c}
\begin{array}{ccccccc} \color{magenta}{j_1} & & \color{magenta}{j_2} & \color{magenta}{\cdots} & \color{magenta}{j_r} & & \end{array} \\
\left(\begin{array}{ccccccc|c}
1 & \star & 0 & \cdots & 0 & \star & \star & b_1 \\
0 & 0 & 1 & \cdots & 0 & \star & \star & b_2 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & \cdots & 1 & \star & \star & b_r \\
0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & \cdots & 0 & 0 & 0 & -1
\end{array}\right) \\
\xleftrightarrow{\hspace{6em}} \\
\color{cyan}{\text{n}}
\end{array}
$$

という形が得られた場合は、$0 = -1$ という常に成り立たない等式が含まれているので、連立されたすべての方程式を満たす解は存在しないことになる。

一般解のパラメータ表示

解が 1 つに定まらない場合は、解の全体像を知ることが方程式を「解く」ことになる。

無数個の解が存在する場合、解の集合が直線を成していたり、もっと高い次元の図形になっていることがある。

主変数と自由変数

係数行列 A の n 個の列が、n 個の変数に対応していることを思い出そう。

> **主変数と自由変数**　行列 A を行基本変形により行階段形にしたとき、主成分がある列に対応する変数を**主変数**と呼び、それ以外の変数を**自由変数**と呼ぶ。

たとえば、次のような既約行階段形に変形した拡大係数行列を考える。

$$
\tilde{A}_{\circ} = \begin{array}{c}
\begin{array}{ccccc|c} 1 & & 3 & 4 & & \end{array} \\
\left(\begin{array}{ccccc|c}
1 & 2 & 0 & 0 & -1 & -3 \\
0 & 0 & 1 & 0 & 2 & 1 \\
0 & 0 & 0 & 1 & 1 & 2 \\
0 & 0 & 0 & 0 & 0 & 0
\end{array}\right) \\
\longleftrightarrow \\
5
\end{array}
$$

変数を使って方程式の形に直すと、

$$
\begin{cases}
x_1 + 2x_2 + 0x_3 + 0x_4 - x_5 = -3 \\
0x_1 + 0x_2 + x_3 + 0x_4 + 2x_5 = 1 \\
0x_1 + 0x_2 + 0x_3 + x_4 + x_5 = 2
\end{cases}
$$

主成分がある列は 1, 3, 4 列なので、主変数は x_1, x_3, x_4 である。

それ以外の x_2, x_5 は自由変数となる。

自由変数とパラメータ

先ほどの方程式

$$
\begin{cases}
x_1 + 2x_2 \qquad\qquad - x_5 = -3 \\
\qquad\qquad x_3 \qquad + 2x_5 = 1 \\
\qquad\qquad\qquad x_4 + x_5 = 2
\end{cases}
$$

において、自由変数を含む項を左辺に移行すれば、

$$
\begin{cases}
x_1 = -2x_2 + x_5 - 3 \\
x_3 = -2x_5 + 1 \\
x_4 = -x_5 + 2
\end{cases}
$$

となる。

自由変数 x_2, x_5 に任意の値を代入したときの主変数 x_1, x_3, x_4 の値はすべてこの方程式の解になる。

自由変数の値は定まらないので、任意の値を取りうる文字として表すしかない。

そこで、

$$
x_2 = t_1, \qquad x_5 = t_2
$$

とおけば、

$$
\begin{cases}
x_1 & = -2t_1 + t_2 - 3 \\
x_3 & = \quad\ \ - 2t_2 + 1 \\
x_4 & = \quad\ \ - t_2 + 2
\end{cases}
$$

すなわち、

$$
\begin{cases}
x_1 & = -2t_1 + t_2 - 3 \\
x_2 & = t_1 \\
x_3 & = \quad\ \ - 2t_2 + 1 \\
x_4 & = \quad\ \ - t_2 + 2 \\
x_5 & = \quad\ \ \ \ t_2
\end{cases}
$$

と書ける。

これをベクトル形に直すことで、一般的な解のパラメータ表示が得られる。

$$
\boldsymbol{x} = \begin{pmatrix} -3 \\ 0 \\ 1 \\ 2 \\ 0 \end{pmatrix} + t_1 \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + t_2 \begin{pmatrix} 1 \\ 0 \\ -2 \\ -1 \\ 1 \end{pmatrix}
$$

各行の方程式から得られる一般解

先ほどの具体例を一般化して考えてみよう。

次のように変形された係数拡大行列のうち、係数行列部分において主成分を含む列を $j_1, j_2, \ldots, j_r$ とする。

$$
\tilde{A}\text{の変形後} = \begin{array}{c}
\begin{array}{ccccccc|c}
j_1 & & j_2 & \cdots & j_r & & & \\
\end{array} \\
\left(\begin{array}{ccccccc|c}
1 & \star & 0 & \cdots & 0 & \star & \star & b_1 \\
0 & 0 & 1 & \cdots & 0 & \star & \star & b_2 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & \cdots & 1 & \star & \star & b_r \\
0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0
\end{array}\right) \\
\longleftarrow \quad n \quad \longrightarrow
\end{array}
$$

すると、各行の方程式は次のような形になる。

$$x_{j_i} + \sum_k \star x_k = b_i$$
$$(k > j_i, k \notin \{j_1, j_2, \ldots, j_r\})$$

ここで、x_k は j_i 列よりも右にある、$\star$ に対応する変数である。

既約行階段形では、主成分を含む列の主成分以外の要素はすべて 0 であるため、$\star$ に対応する自由変数のみが残る。
そのため、$k \notin \{j_1, j_2, \ldots, j_r\}$ という条件によって、$\star$ だけを残すようにしている。

移項して主変数 x_{j_i} について解いた形にすると、各行から得られる解は、

$$x_{j_i} = b_i - \sum_k \star x_k$$

となる。

この式から、$x_{j_1}, x_{j_2}, \ldots, x_{j_r}$ 以外の自由変数 x_k に勝手な数を与えるごとに、主変数 $x_{j_1}, x_{j_2}, \ldots, x_{j_r}$ が定まることがわかる。

このような自由変数は $n - r$ 個あるので、$A\boldsymbol{x} = \boldsymbol{b}$ の解は、$n - r$ 個のパラメータを用いて表すことができる。

基本解と特殊解

自由変数 $\star$ をパラメータ t_i とおき、各行の方程式から得られる解を縦に並べてベクトルの形にまとめると、

$$\boldsymbol{x} = \boldsymbol{b}' - \sum_{i=1}^{n-r} t_i \boldsymbol{u}_i$$

という形の一般解の表示が得られる。

ここで、パラメータ t_i をかけた列ベクトル $\boldsymbol{u}_i$ を連立方程式の基本解という。
また、パラメータをかけていない列ベクトル $\boldsymbol{b}'$ は、変形後の拡大係数行列 $(A' \mid \boldsymbol{b}')$ の定数項部分 $\boldsymbol{b}'$ から得られるベクトルであり、これを特殊解という。

行列の階数

ここで、主成分の個数 r に名前をつけておこう。

行階段行列において、主成分とは、零行でない行の中で一番左にある 0 でない成分のことを指す。
つまり、行階段行列の主成分の個数 r は、零行でない行の数と一致する。

行列の階数　行列 A を行階段行列に変形したとき、零行でない行の個数を A の**階数**あるいは**ランク**といい、$\mathrm{rank}\, A$ と書く。

零行でない行の個数は、既約行階段行列まで変形しなくても、行階段行列の時点で読み取れることに注意しよう。

変形の結果として得られる行階段行列は 1 通りとは限らないし、変形の途中の掃き出しの手順も 1 通りとは限らないが、

階数 $\mathrm{rank}\, A$ は A のみによって定まる値である

ことが後に証明できる。

階数のとりうる値の範囲

A が $m \times n$ 型ならば、行は m 個なので、$\mathrm{rank}\, A$ は 0 以上 m 以下の整数である。

また、階数は行階段行列に変形したときの主成分の個数でもあり、行基本行列の主成分は各列に高々 1 つなので、主成分の個数は列の個数 n を超えることはない。

したがって、次の重要な評価が成り立つ。

⚓ 行列の階数の範囲　$m \times n$ 型の行列 A の階数に対して、次の不等式が成り立つ。

$$0 \leq \mathrm{rank}(A) \leq \min(m, n)$$

解の自由度

連立方程式は、解が存在する場合、$n-r$ 個のパラメータを用いて一般解を表現できた。

パラメータの個数は、自由変数の個数でもあり、基本解の個数でもある。
パラメータの個数だけ、自由に値を決めることができる未知数が方程式に含まれているということである。

そこで、解を表すパラメータの個数を解の自由度と呼ぶ。

$$\begin{aligned} \text{解の自由度} &= \text{変数の個数} - \mathrm{rank}(A) \\ &= n - r \end{aligned}$$

解の自由度は、解全体のなす集合の大きさ、すなわち何次元の空間なのかを表している。

解の存在条件

連立一次方程式 $A\boldsymbol{x} = \boldsymbol{b}$ の解の存在条件について、階数を用いて議論することができる。

まず、行基本変形によって得られる方程式の解は、元の方程式の解と同じであった。

そこで、次のように変形した拡大係数行列をもとに考えると、

$$
\tilde{A}\text{の変形後} = \begin{array}{c}
\begin{array}{ccccccc} j_1 & & j_2 & \cdots & j_r & & \end{array} \\
\left(\begin{array}{ccccccc|c}
1 & \star & 0 & \cdots & 0 & \star & \star & b_1 \\
0 & 0 & 1 & \cdots & 0 & \star & \star & b_2 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & \cdots & 1 & \star & \star & b_r \\
0 & 0 & 0 & \cdots & 0 & 0 & 0 & b_{r+1} \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & \cdots & 0 & 0 & 0 & b_m
\end{array}\right) \\
\xleftrightarrow[n]{\qquad\qquad\qquad}
\end{array}
$$

この方程式の解が存在するのは、

$$
b_{r+1} = \cdots = b_m = 0
$$

の場合のみであることをすでに考察した。

ここで、拡大係数行列 $\tilde{A}$ は A の右端に 1 列追加して得られるので、零行でない行の個数、すなわち階数を考えると、$\operatorname{rank} \tilde{A}$ は $\operatorname{rank} A$ と等しいか、1 だけ増えるかのどちらかである。

$b_{r+1} = \cdots = b_m = 0$ の場合、$\operatorname{rank} \tilde{A}$ と $\operatorname{rank} A$ は一致する。

$$
\begin{array}{c}
\begin{array}{ccccccc} j_1 & & j_2 & \cdots & j_r & & \end{array} \\
\left(\begin{array}{ccccccc}
1 & \star & 0 & \cdots & 0 & \star & \star \\
0 & 0 & 1 & \cdots & 0 & \star & \star \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & \cdots & 1 & \star & \star \\
0 & 0 & 0 & \cdots & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & \cdots & 0 & 0 & 0
\end{array}\right) \\
\xleftrightarrow[n]{\qquad\qquad\qquad}
\end{array}
\qquad
\begin{array}{c}
\begin{array}{ccccccc} j_1 & & j_2 & \cdots & j_r & & \end{array} \\
\left(\begin{array}{ccccccc|c}
1 & \star & 0 & \cdots & 0 & \star & \star & b_1 \\
0 & 0 & 1 & \cdots & 0 & \star & \star & b_2 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & \cdots & 1 & \star & \star & b_r \\
0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0
\end{array}\right) \\
\xleftrightarrow[n]{\qquad\qquad\qquad}
\end{array}
$$

一方、$b_{r+1}, \ldots, b_m$ のうち、1 つでも 0 でないものがある場合は、拡大係数行列の右端の列に主成分が現れ、$\operatorname{rank} \tilde{A}$ と $\operatorname{rank} A$ は一致しない。

$$
\begin{array}{c}
\begin{array}{ccccccc}
j_1 & & j_2 & \cdots & j_r & & \\
\end{array} \\
\begin{pmatrix}
1 & \star & 0 & \cdots & 0 & \star & \star \\
0 & 0 & 1 & \cdots & 0 & \star & \star \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & \cdots & 1 & \star & \star \\
0 & 0 & 0 & \cdots & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & \cdots & 0 & 0 & 0
\end{pmatrix} \\
\xleftrightarrow[\text{n}]{\hspace{4cm}}
\end{array}
\qquad
\begin{array}{c}
\begin{array}{cccccccc}
j_1 & & j_2 & \cdots & j_r & & & \\
\end{array} \\
\left(\begin{array}{ccccccc|c}
1 & \star & 0 & \cdots & 0 & \star & \star & b_1 \\
0 & 0 & 1 & \cdots & 0 & \star & \star & b_2 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & \cdots & 1 & \star & \star & b_r \\
0 & 0 & 0 & \cdots & 0 & 0 & 0 & 1 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0
\end{array}\right) \\
\xleftrightarrow[\text{n}]{\hspace{4cm}}
\end{array}
$$

$b_{r+1}, \ldots, b_m$ のうち、0 でないものが 2 つ以上ある場合も、さらに行基本変形を行うことで、右上の拡大係数行列と同じ形にできるので、

$$\operatorname{rank} \tilde{A} = \operatorname{rank} A + 1$$

となる。

以上の考察から、連立一次方程式 $A\boldsymbol{x} = \boldsymbol{b}$ の解が存在する条件は、

係数行列と拡大係数行列の階数が等しい

ことだといえる。

⚓ 拡大係数行列と解の存在条件　A を $m \times n$ 型行列、$\boldsymbol{b} \in \mathbb{R}^m$ とする

$\tilde{A} = (A \mid \boldsymbol{b})$ とおくとき、

$$\operatorname{rank}(\tilde{A}) = \operatorname{rank}(A) \iff A\boldsymbol{x} = \boldsymbol{b} \text{ に解が存在する}$$

証明

[Todo 22: book: 行列と行列式の基礎 p31 （定理 1.5.1）]

⚓ 解の存在条件の系　A を $m \times n$ 型行列とするとき、

$$^{\forall}\boldsymbol{b} \in \mathbb{R}^m,\ A\boldsymbol{x} = \boldsymbol{b} \text{ の解が存在する} \iff \mathrm{rank}(A) = m$$

証明

[Todo 23: book: 行列と行列式の基礎 p32 （定理 1.5.2, 1.5.3）]

解の一意性

ここまでの議論から、次のことがいえる。

⚓ 解の一意性　$A\boldsymbol{x} = \boldsymbol{b}$ の解が存在するとき、

$$\text{解が一意的である} \iff \mathrm{rank}(A) = n$$

ここで、n は変数の個数である。

証明

$\Longleftarrow$

$\mathrm{rank}(A) = n$ であれば、解の自由度は $n - n = 0$、すなわち自由変数が存在しないことになる

自由変数がなければ「各変数＝定数」という式に変形できることになるので、解は明らかに一意的である ■

$\Longrightarrow$

対偶 $\mathrm{rank}(A) \neq n \Longrightarrow$ 解が一意的 を示す

$\mathrm{rank}(A) \leq n$ であるので、$\mathrm{rank}(A) \neq n$ は $\mathrm{rank}(A) < n$ を意味する

$\mathrm{rank}(A) < n$ であれば、自由変数が 1 つ以上存在するので解は無数にある

よって、解は一意的ではない ■

解のパラメータ表示の一意性

自由変数を $x_{j_1}, \ldots, x_{j_{n-r}}$ とするとき、一般解の表示

$$\boldsymbol{x} = \boldsymbol{x}_0 + t_1\boldsymbol{u}_1 + t_2\boldsymbol{u}_2 + \cdots + t_{n-r}\boldsymbol{u}_{n-r}$$

の j_k 番目の成分は等式

$$x_{j_k} = t_k$$

を意味するので、解が与えられたとき、パラメータの値は直接に読み取れる

このことから、

$$\boldsymbol{x} = \boldsymbol{x}_0 + t_1\boldsymbol{u}_1 + t_2\boldsymbol{u}_2 + \cdots + t_{n-r}\boldsymbol{u}_{n-r}$$

によって解を表示する際の $n-r$ 個のパラメータの値は一意的に定まることがわかる

この事実は、$\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_{n-r} \in \mathbb{R}^m$ が線形独立であると表現される

斉次形方程式の非自明解

$A\boldsymbol{x} = \boldsymbol{b}$ において、$\boldsymbol{b} = \boldsymbol{o}$ の場合、つまり、

$$A\boldsymbol{x} = \boldsymbol{o}$$

の形の線形連立方程式は斉次形であるという。

斉次形の場合は $\boldsymbol{x} = \boldsymbol{o}$ が明らかに解になっていて、これを自明解という。

したがって、斉次形の方程式では、自明解以外に解が存在するかどうかが基本的な問題となる。

斉次形方程式の非自明解の存在条件　斉次形の方程式 $A\boldsymbol{x} = \boldsymbol{o}$ において、

$$\text{自明解しか存在しない} \iff \mathrm{rank}(A) = n$$

ここで、n は変数の個数である。

証明

斉次形の場合は自明解が常に存在するので、解の一意性 $\mathrm{rank}(A) = n$ は、それ以外の解がないということを意味している。 ■

斉次形方程式と一般解

一般解のパラメータ表示の章で具体例として挙げた連立一次方程式

$$(A' \mid \boldsymbol{b}') = \begin{array}{c} \begin{array}{cccccc} 1 & & 3 & 4 & & \end{array} \\ \left(\begin{array}{ccccc|c} 1 & 2 & 0 & 0 & -1 & -3 \\ 0 & 0 & 1 & 0 & 2 & 1 \\ 0 & 0 & 0 & 1 & 1 & 2 \end{array}\right) \end{array}$$

すなわち、

$$\begin{cases} x_1 + 2x_2 \qquad\qquad - \ \ x_5 = -3 \\ \qquad\qquad x_3 \qquad + 2x_5 = 1 \\ \qquad\qquad\qquad x_4 + \ \ x_5 = 2 \end{cases}$$

の一般解は、次のようなパラメータ表示として得られた。

$$\boldsymbol{x} = \underset{\text{特殊解}}{\begin{pmatrix} -3 \\ 0 \\ 1 \\ 2 \\ 0 \end{pmatrix}} + t_1 \underset{\text{基本解}}{\begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}} + t_2 \underset{\text{基本解}}{\begin{pmatrix} 1 \\ 0 \\ -2 \\ -1 \\ 1 \end{pmatrix}}$$

この一般解を、ここでは別なアプローチで考察する。

特殊解は簡単に見つかる

自由変数 x_2, x_5 の値はなんでもよいので、これらを 0 とおいたもの

$$\begin{cases} x_1 & & & = -3 \\ & x_3 & & = 1 \\ & & x_4 & = 2 \end{cases}$$

すなわち、

$$\begin{cases} x_1 & & & & & = -3 \\ & x_2 & & & & = 0 \\ & & x_3 & & & = 1 \\ & & & x_4 & & = 2 \\ & & & & x_5 & = 0 \end{cases}$$

も方程式の解となる。

この解は、パラメータ表示された解 $\boldsymbol{x}$ の特殊解に一致しており、さらに、変形後の拡大係数行列 $(A' \mid \boldsymbol{b}')$ の定数項部分 $\boldsymbol{b}'$ からも直接読み取れることがわかる。

$$\boldsymbol{x}_0 = \underset{\text{特殊解}}{\begin{pmatrix} -3 \\ 0 \\ 1 \\ 2 \\ 0 \end{pmatrix}}$$

$$\begin{array}{ccccc} 1 & & 3 & 4 & \\ \end{array}$$

$$\left(\begin{array}{ccccc|c} 1 & 2 & 0 & 0 & -1 & -3 \\ 0 & 0 & 1 & 0 & 2 & 1 \\ 0 & 0 & 0 & 1 & 1 & 2 \end{array}\right)$$

解のパラメータ表示の再解釈

基本解と特殊解の章でも述べたように、特殊解を $\boldsymbol{x}_0$、基本解を $\boldsymbol{u}_i$ とすると、$A\boldsymbol{x} = \boldsymbol{b}$ の一般解は次のように表せる。

$$\boldsymbol{x} = \boldsymbol{x}_0 + t_1\boldsymbol{u}_1 + \cdots + t_{n-r}\boldsymbol{u}_{n-r}$$

ここで、実は、

$$A\boldsymbol{x}_0 = \boldsymbol{b}, \quad A\boldsymbol{u}_1 = \boldsymbol{o}, \quad \ldots \quad , A\boldsymbol{u}_{n-r} = \boldsymbol{o}$$

が成り立っている。

このことは次のように確かめられる。

$$\begin{aligned} A\boldsymbol{x} &= A(\boldsymbol{x}_0 + t_1\boldsymbol{u}_1 + \cdots + t_{n-r}\boldsymbol{u}_{n-r}) \\ &= A\boldsymbol{x}_0 + t_1A\boldsymbol{u}_1 + \cdots + t_{n-r}A\boldsymbol{u}_{n-r} \\ &= \boldsymbol{b} + t_1\boldsymbol{o} + \cdots + t_{n-r}\boldsymbol{o} \\ &= \boldsymbol{b} \end{aligned}$$

つまり、基本解 $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_{n-r}$ は、もとの方程式において $\boldsymbol{b} = \boldsymbol{o}$ とした斉次形方程式 $A\boldsymbol{x} = \boldsymbol{o}$ の解である。

1 つの解 $\boldsymbol{x}_0$ が見つかったら、あとは斉次形方程式 $A\boldsymbol{x} = \boldsymbol{o}$ の一般解を求めればよい

この方法で、$A\boldsymbol{x} = \boldsymbol{b}$ のすべての解が見つかることになる。

斉次形方程式の特殊解は自明解

特殊解を $\boldsymbol{x}_0$ とすると、次の関係が成り立っていた。

$$A\boldsymbol{x}_0 = \boldsymbol{b}$$

ここで、$\boldsymbol{b} = \boldsymbol{o}$ とした場合を考えると、

$$A\boldsymbol{x}_0 = \boldsymbol{o}$$

となるので、斉次形方程式 $A\boldsymbol{x} = \boldsymbol{o}$ の特殊解 $\boldsymbol{x}_0$ は、自明解 $\boldsymbol{x} = \boldsymbol{o}$ である。

つまり、斉次形方程式の一般解は、特殊解を除いた次のような形になる。

$$\boldsymbol{x} = t_1\boldsymbol{u}_1 + t_2\boldsymbol{u}_2 + \cdots + t_{n-r}\boldsymbol{u}_{n-r}$$

斉次形方程式の一般解を求めるということは、基本解を求めることに他ならない。

斉次形方程式の基本解も簡単に見つかる

斉次形方程式 $A\boldsymbol{x} = \boldsymbol{o}$ の基本解は、既約行階段形にした係数行列 $A_{\circ}$ の形を見てすぐに書き下せる。

たとえば、係数行列 A が次のように変形されたとする。

$$A_\circ = \begin{array}{c} \begin{array}{cccccc} 1 & & 3 & 4 & & \end{array} \\ \begin{pmatrix} 1 & 2 & 0 & 0 & 3 & -5 \\ 0 & 0 & 1 & 0 & 7 & 6 \\ 0 & 0 & 0 & 1 & -4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \end{array}$$

自由変数 x_2, x_5, x_6 をパラメータとして t_1, t_2, t_3 とおくと、基本解 $\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3$ を用いて、一般解 $\boldsymbol{x}$ は次のように表せる。

$$\boldsymbol{x} = t_1\boldsymbol{u}_1 + t_2\boldsymbol{u}_2 + t_3\boldsymbol{u}_3$$

ここで、$x_2 = t_1, x_5 = t_2, x_6 = t_3$ であるということは、$\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3$ が次のような形になっているはずである。

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{pmatrix} = t_1 \begin{pmatrix} \star \\ 1 \\ \star \\ \star \\ 0 \\ 0 \end{pmatrix} + t_2 \begin{pmatrix} \star \\ 0 \\ \star \\ \star \\ 1 \\ 0 \end{pmatrix} + t_3 \begin{pmatrix} \star \\ 0 \\ \star \\ \star \\ 0 \\ 1 \end{pmatrix}$$

この形であれば、各行を見ると、$x_2 = t_1, x_5 = t_2, x_6 = t_3$ が成り立つことがわかる。

$\star$ の部分については、変形後の係数行列から次のように読み取ればよい。

$$A_\circ = \begin{pmatrix} 1 & 2 & 0 & 0 & 3 & -5 \\ 0 & 0 & 1 & 0 & 7 & 6 \\ 0 & 0 & 0 & 1 & -4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\boldsymbol{x} = t_1 \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + t_2 \begin{pmatrix} -3 \\ 0 \\ -7 \\ 4 \\ 1 \\ 0 \end{pmatrix} + t_3 \begin{pmatrix} 5 \\ 0 \\ -6 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

このように係数行列から読み取った数の符号を変えたものを $\star$ の位置に並べるだけで解が得られる根拠は、次のように方程式 $A_\circ\boldsymbol{x} = \boldsymbol{o}$ の形に直すとわかる。

$$
\begin{cases}
x_1 + 2x_2 \qquad\qquad + 3x_5 \qquad - 5 = 0 \\
\qquad\qquad x_3 \qquad + 7x_5 + 6x_6 \qquad = 0 \\
\qquad\qquad\qquad x_4 - 4x_5 \qquad\qquad = 0
\end{cases}
$$

基本解 x_1, x_3, x_4 は、自由変数をパラメータに置き換えた上で移項することで求まるので、

$$
\begin{cases}
x_1 \qquad\qquad = -2t_1 - 3t_2 + 5t_3 \\
\qquad x_3 \qquad = \qquad - 7t_2 + 6t_3 \\
\qquad\qquad x_4 = \qquad 4t_2
\end{cases}
$$

この移項によって、変形後の係数行列に並んでいた数値から符号を変えたものが解として使われることになる。

斉次形方程式の一般解

先ほどの手順は、まず自由変数の位置をもとに単位ベクトルを考え、そこから主変数を構成する値を引いたものとして一般解を構成したと整理できる。

つまり、先ほど求めた一般解のパラメータ表示を

$$
\boldsymbol{x} = t_1\boldsymbol{u}_1 + t_2\boldsymbol{u}_2 + t_3\boldsymbol{u}_3
$$

とおくと、各基本解 $\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3$ は、次のように見ることもできる。

$$\boldsymbol{u}_1 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} - \begin{pmatrix} 2 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\boldsymbol{u}_2 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} - \begin{pmatrix} 3 \\ 0 \\ 7 \\ -4 \\ 0 \\ 0 \end{pmatrix}$$

$$\boldsymbol{u}_3 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} - \begin{pmatrix} -5 \\ 0 \\ 6 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

ここで、たとえば $\boldsymbol{u}_2$ は、さらに次のように展開できる。

$$\boldsymbol{u}_2 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} - \boldsymbol{e}_1 \begin{pmatrix} 3 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} - \boldsymbol{e}_3 \begin{pmatrix} 0 \\ 0 \\ 7 \\ 0 \\ 0 \\ 0 \end{pmatrix} - \boldsymbol{e}_4 \begin{pmatrix} 0 \\ 0 \\ 0 \\ -4 \\ 0 \\ 0 \end{pmatrix}$$

このことを一般化しておこう。

主変数の番号を $i_1, \ldots, i_r$、自由変数の番号を $j_1, \ldots, j_{n-r}$ とする。

$$A_\circ = \begin{array}{c} \begin{matrix} i_1 & j_1 & i_2 & \cdots & i_r & j_{n-r} \end{matrix} \\ \begin{pmatrix} 1 & b_{1,1} & 0 & \cdots & 0 & b_{1,n-r} \\ 0 & 0 & 1 & \cdots & 0 & b_{2,n-r} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & b_{r,n-r} \\ 0 & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix} \end{array}$$

$A_\circ$ の第 j_k 列を $(b_{ik})_{i=1}^m$ とするとき、基本解 $\boldsymbol{u}_k$ は、

$$\boldsymbol{u}_k = \boldsymbol{e}_{j_k} - \sum_{l=1}^{r} b_{lk} \boldsymbol{e}_{i_l} \quad (k = 1, \ldots, n-r)$$

と表すことができる。

線形写像の表現行列と連立方程式

線型写像の単射性・全射性は、その表現行列の階数によって連立一次方程式の解の性質と結びつく。

単射な線形写像と階数

線形写像 f が単射であることを、表現行列 A の性質として述べると、次のような言い換えができる。

⚓ **線形写像の単射性と表現行列**　線形写像 $f\colon \mathbb{R}^n \to \mathbb{R}^m$ の表現行列を A とするとき、次はすべて同値である。

i. f は単射
ii. $A\boldsymbol{x} = \boldsymbol{o}$ は自明な解しか持たない
iii. $\mathrm{rank}(A) = n$

証明

(i) $\Longleftrightarrow$ (ii)

線形写像 f は、表現行列 A を用いて次のように表せる

$$f(\boldsymbol{x}) = A\boldsymbol{x}$$

f が単射であることの言い換えは、

$$f(\boldsymbol{x}) = \boldsymbol{o} \Longrightarrow \boldsymbol{x} = \boldsymbol{o}$$

であり、$A\boldsymbol{x} = \boldsymbol{o}$ が自明解しか持たないことは、

$$A\boldsymbol{x} = \boldsymbol{o} \Longrightarrow \boldsymbol{x} = \boldsymbol{o}$$

が成り立つということである

$f(\boldsymbol{x}) = A\boldsymbol{x}$ であるから、これらの 2 つの条件は同値である ■

(ii) $\Longleftrightarrow$ (iii)

斉次形の非自明解の存在条件より、斉次形の方程式 $A\boldsymbol{x} = \boldsymbol{o}$ に自明解しか存在しないことと

$$\operatorname{rank}(A) = n$$

と同値である。 ■

全射な線形写像と階数

単射性と対比して、全射性についても表現行列の言葉で整理する。

線形写像の全射性と表現行列 線形写像 $f\colon \mathbb{R}^n \to \mathbb{R}^m$ の表現行列を A とするとき、次はすべて同値である。

i. f は全射
ii. 任意の $\boldsymbol{b} \in \mathbb{R}^m$ に対して、$A\boldsymbol{x} = \boldsymbol{b}$ には解が存在する

iii. $\mathrm{rank}(A) = m$

証明

(i) $\iff$ (ii)

線形写像 f は、表現行列 A を用いて次のように表せる

$$f(\boldsymbol{x}) = A\boldsymbol{x}$$

f が全射であることの言い換えは、

$$\forall \boldsymbol{b} \in \mathbb{R}^m, \exists \boldsymbol{x} \in \mathbb{R}^n, f(\boldsymbol{x}) = \boldsymbol{b}$$

であり、これは

$$\forall \boldsymbol{b} \in \mathbb{R}^m, A\boldsymbol{x} = \boldsymbol{b} \text{ に解が存在する}$$

と同値である

よって、これらの 2 つの条件は同値である ■

(ii) $\iff$ (iii)

連立方程式の解の存在条件より、$\mathrm{rank}(A) = m$ は、次の条件

$$^\forall \boldsymbol{b} \in \mathbb{R}^m, A\boldsymbol{x} = \boldsymbol{b} \text{ の解が存在する}$$

ことと同値である。 ■

全単射な線形変換と階数

一般の線形写像と対比して、線形変換の大きな特徴は次が成り立つことである。

単射と全射は、一般には一方から他方が導かれるわけではない 2 つの性質だが、$\mathbb{R}^n$ からそれ自身への線形写像（線形変換）の場合は同値になる。

線形代数における鳩の巣原理　f を $\mathbb{R}^n$ の線形変換とし、A を f の表現行列とするとき、次はすべて同値である。

i. f は単射
ii. f は全射
iii. f は全単射
iv. $\mathrm{rank}(A) = n$

証明

線形写像 $f\colon \mathbb{R}^n \to \mathbb{R}^m$ において、表現行列を A とすると、

$$
\begin{aligned}
f \text{ が単射} &\iff \mathrm{rank}(A) = n \\
f \text{ が全射} &\iff \mathrm{rank}(A) = m
\end{aligned}
$$

である。

線形変換は、線形写像 $f\colon \mathbb{R}^n \to \mathbb{R}^m$ の $m = n$ の場合であるので、f が単射であることも、全射であることも、

$$
\mathrm{rank}(A) = n
$$

という条件と同値になる。

つまり、線形変換は単射かつ全射であり、これは全単射であることも意味する。 ■

この定理は、いわば線形代数版「鳩の巣原理」である。

鳩の巣原理　有限集合 $X = \{1, 2, \ldots, n\}$ からそれ自身への写像 f に対して、単射と全射は同値である

鳩の巣原理は、歴史的には部屋割り論法とも呼ばれ、

n 個のものを m 個の箱に入れるとき、$n > m$ であれば、少なくとも 1 個の箱には 1 個より多いものが中にある

ことを指す。

ここで鳩の巣原理と呼んだのはこの命題そのものではないが、その変種と考えてよい。

階数による正則判定

線形代数における鳩の巣原理から、次のことがいえる。

正則の判定と階数　n 次正方行列 A に対して、

$$A \text{ が正則行列} \iff \operatorname{rank}(A) = n$$

この定理は、線形変換 f（もしくは正方行列 A）が正則かどうかについて、階数という 1 つの数値で判定できることを示している。

第 9 章

線型独立な列ベクトルと階数

非自明解の存在と有限従属性

斉次形方程式 $A\boldsymbol{x} = \boldsymbol{o}$ の非自明解の存在に対して、次の解釈もできる。

斉次形方程式の非自明解の存在と線形従属　$m \times n$ 型行列 A の列ベクトルを $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ とするとき、

$$A\boldsymbol{x} = \boldsymbol{o} \text{ に自明でない解がある} \iff \boldsymbol{a}_1, \ldots, \boldsymbol{a}_n \text{が線形従属}$$

証明

$A\boldsymbol{x} = \boldsymbol{o}$ は、ベクトルの等式

$$x_1\boldsymbol{a}_1 + \cdots + x_n\boldsymbol{a}_n = \boldsymbol{o}$$

と同じものである。

$\Longrightarrow$

もし自明でない解があるならば、$x_1, \ldots, x_n$ のうち少なくとも 1 つは 0 ではない。

$x_1\boldsymbol{a}_1 + \cdots + x_n\boldsymbol{a}_n = \boldsymbol{o}$ が成り立つもとで、0 でない係数が存在するということは、$\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ が線形従属であることを意味する。 ■

$\Longleftarrow$

対偶を示す。

$\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ が線形独立であれば、

$$x_1\boldsymbol{a}_1 + \cdots + x_n\boldsymbol{a}_n = \boldsymbol{o}$$

において、すべての係数 $x_1, \ldots, x_n$ は 0 でなければならない。

よって、0 以外の解（非自明解）は存在しないことになる。 ■

この命題の否定をとると、

$$A\boldsymbol{x} = \boldsymbol{o} \text{ には自明解しか存在しない} \iff \boldsymbol{a}_1, \ldots, \boldsymbol{a}_n \text{が線形独立}$$

となる。

ここで、斉次形方程式の非自明解の存在条件より、斉次形方程式 $A\boldsymbol{x} = \boldsymbol{o}$ において自明解しか存在しないことは、$\text{rank}(A) = n$、すなわち解の自由度が 0 であることと同値であった。

つまり、次が成り立つことがわかる。

⚓ 列ベクトルの線型独立性と階数 $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n \in \mathbb{R}^m$ に対して、$A = (\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n)$ とおくと、

$$\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n \text{が線型独立} \iff \text{rank}(A) = n$$

有限従属性（方程式の視点）

$\text{rank}(A) = n$ が成り立つ条件をさらに言い換えてみよう。

$\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n \in \mathbb{R}^m$ に対して、$A = (\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n)$ とおくと、A は $m \times n$ 型行列である。

階数のとりうる値の範囲より、

$$\operatorname{rank} A \leq \min(m, n)$$

であるから、もしも列の方が行よりも多い、つまり $n > m$ であれば、$\operatorname{rank} A \leq m < n$ となり、$\operatorname{rank} A = n$ が成り立つことはない。

よって、次の関係がいえる。

$$\begin{aligned} A\boldsymbol{x} = \boldsymbol{o} \text{ に自明でない解がある} &\iff \boldsymbol{a}_1, \ldots, \boldsymbol{a}_n \text{が線形従属} \\ &\iff \operatorname{rank} A \neq n \end{aligned}$$

ここで n は変数の個数、m は方程式の個数であるので、$n > m$ という状況を次のようにまとめることができる。

斉次形方程式における有限従属性　斉次線型方程式 $A\boldsymbol{x} = \boldsymbol{o}$ において、変数の個数が方程式の個数よりも多いときには、非自明な解が存在する。

有限従属性（ベクトルの集合における視点）

連立方程式の文脈に限定せず、より抽象的に言い換えたものが次の定理である。

有限従属性定理　$\mathbb{R}^m$ 内の m 個よりも多いベクトルからなる集合は線形従属である。

$n > m$ の場合、$\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n \in \mathbb{R}^m$ は線形従属となることを述べている。

この結論は、幾何的な直観からは自然だといえる。
たとえば、平面 $\mathbb{R}^2$ 内に 3 つ以上のベクトルがあれば、自動的に線形従属になる。

この事実は、次元の概念を議論する際の基礎となる。

列ベクトルの線型独立性と行基本変形

列ベクトルの線型独立性と階数では、列ベクトルの線型独立性を行列の階数で言い換えられることを示した。

行列の階数は、行基本変形を施した結果である行階段形からわかるものである。
もしも行基本変形によって列ベクトルの線型独立性が変化するとしたら、階数との関係も変わってしまうのではないだろうか。

この心配が杞憂であることは、次の定理からわかる。

行基本変形による線型独立性の不変性　行変形は列ベクトルの線形関係を保つ。

すなわち、行列 $A = (\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n)$ に行の変形を施して $B = (\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n)$ が得られたとするとき、

$$\sum_{i=1}^{n} c_i \boldsymbol{a}_i = \boldsymbol{o} \iff \sum_{i=1}^{n} c_i \boldsymbol{b}_i = \boldsymbol{o}$$

特に、

$$\{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n\} \text{ が線型独立} \iff \{\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n\} \text{ が線型独立}$$

証明

基本行列の積による行変形の構成より、P を基本行列の積（正則行列）とすると、$B = PA$ が成り立つ。

よって、$\boldsymbol{b}_i = P\boldsymbol{a}_i$ であり、線形関係式

$$\sum_{i=1}^{n} c_i \boldsymbol{a}_i = \boldsymbol{0}$$

に左から P をかけることで、

$$\sum_{i=1}^{n} c_i \boldsymbol{b}_i = \boldsymbol{0}$$

が得られる。

逆に、$\sum_{i=1}^{n} c_i \boldsymbol{b}_i = \mathbf{0}$ が成り立つとき、P^{-1} を左からかけることで、

$$\sum_{i=1}^{n} c_i \boldsymbol{a}_i = \mathbf{0}$$

が得られる。

したがって、

$$\sum_{i=1}^{n} c_i \boldsymbol{a}_i = \mathbf{0} \iff \sum_{i=1}^{n} c_i \boldsymbol{b}_i = \mathbf{0}$$

が成り立つ。 ■

主列ベクトルと階数の再解釈

行列の階数は、行階段形に変形したときの主成分の個数でもあった。

列ベクトルの線型独立性と階数の関係をさらに考察するために、「主成分のある列ベクトル」について考えてみよう。

主列ベクトル　行列 $A = (\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n)$ を行階段形にしたときに、主成分のある列番号を $i_1, \ldots, i_r$ とする。ここで、r は A の階数である。

このとき、$\boldsymbol{a}_{i_1}, \ldots, \boldsymbol{a}_{i_r}$ を主列ベクトルという。

掃き出し法は、主列ベクトルを選び出すためのアルゴリズムといえる。

行基本変形を行っても列ベクトルの線型独立性は変わらないことを根拠に、次の議論ができる。

主列ベクトルと線型独立性　行列の主列ベクトルの集合は線型独立である。

また、主列ベクトル以外の列ベクトルは、主列ベクトルの線形結合である。

証明

[Todo 24: book: 行列と行列式の基礎 p43 （命題 1.6.11）]

つまり、掃き出し法は、行列の列ベクトルの中から、$\mathrm{rank}\, A$ 個の線型独立な列ベクトルを選び出す方法を与えていることになる。

では、$\mathrm{rank}\, A$ 個よりも多くの列ベクトルを選ぶとどうなるのだろうか？

列ベクトルの線形従属性と階数　行列 A の列ベクトルから $\mathrm{rank}\, A$ 個よりも多いベクトルを選ぶと、線形従属になる。

証明

[Todo 25: book: 行列と行列式の基礎 p43 （命題 1.6.12）]

以上によって、行列の階数に関する次の理解が得られる。

階数と線型独立な列ベクトルの最大個数　行列 A の階数 $\mathrm{rank}\, A$ は、A の列ベクトルに含まれる線型独立なベクトルの最大個数と一致する。

証明

[Todo 26: book: 行列と行列式の基礎 p43 （定理 1.6.13）]

これで、「行変形を繰り返して行階段形にしたときの 0 でない段の数」として導入した階数という量の、より本質的な意味がわかった。

このように線型独立な列ベクトルの最大個数として階数を見直すことで、

行変形によって定めた階数が行変形の仕方によらない

こともわかる。

行基本変形による階数の不変性　行の基本変形で行列の階数は変化しない。

転置による階数の不変性

ここまで、行列の列ベクトルと階数の関係を考察してきたが、行ベクトルと階数の関係はどうだろうか？

正則行列による階数標準形の構成を用いて、次の重要な事実を証明することができる。

転置に関する階数の不変性　任意の行列 A に対して、

$$\mathrm{rank}\, A = \mathrm{rank}\, {}^tA$$

証明

A の階数標準形を B とすると、$B = PAQ$ となる正則行列 P, Q をとることができる。

両辺の転置をとると、

$${}^tB = {}^t(PAQ) = {}^tQ\,{}^tA\,{}^tP$$

となり、ここで、正則行列は転置をとっても正則なので、${}^tP, {}^tQ$ も正則行列である。

よって、${}^tA$ の階数標準形は ${}^tB$ である。

B は階数標準形であり、その形から明らかに

$$\mathrm{rank}\, B = \mathrm{rank}\, {}^tB$$

が成り立つので、変形前の行列 A についても

$$\operatorname{rank} A = \operatorname{rank} {}^t A$$

が成り立つ。 ■

行列 A の階数は A の線型独立な列ベクトルの最大個数であったが、この定理から次のこともいえるようになった。

階数と線型独立な行ベクトルの最大個数　行列 A の階数 $\operatorname{rank} A$ は、A の行ベクトルに含まれる線型独立なベクトルの最大個数と一致する。

この事実を連立方程式の視点で解釈すると、

係数行列 A の階数は、

独立な（本質的に意味を持つ）方程式の最大本数

を表しているといえる。

第 10 章

部分空間の基底と次元

線形部分空間

$m > n$ の場合、$m \times n$ 型行列 A は、写し先の空間をカバーしきれない写像を表していた。
つまり、写った結果が空間の一部、部分空間になるということである。

そこで、$\mathbb{R}^n$ の部分集合であって、ベクトル演算で閉じた集合について考える。これは、原点を含む直線や平面などを一般化した概念である。

線形部分空間　$\mathbb{R}^n$ のベクトルからなる空集合でない集合 V は、次が成り立つとき線形部分空間あるいは簡単に部分空間であるという。

i. すべての $\boldsymbol{u}, \boldsymbol{v} \in V$ に対して $\boldsymbol{u} + \boldsymbol{v} \in V$ が成り立つ
ii. すべての $c \in \mathbb{R}$, $\boldsymbol{u} \in V$ に対して $c\boldsymbol{u} \in V$ が成り立つ

入れものの空間 $\mathbb{R}^n$ のことはあまり意識せずに、集合 V とそのベクトル演算に着目して、ある $\mathbb{R}^n$ の線形部分空間のことを単に線形空間と呼ぶこともある。

$\mathbb{R}^n$ 自身も部分空間

たとえば、$\mathbb{R}^n$ 自身は明らかに $\mathbb{R}^n$ の部分空間である。

$n-1$ 次平面は部分空間

たとえば $\mathbb{R}^3$ において座標を (x, y, z) とするとき、xy 平面は $\mathbb{R}^3$ の部分空間である。

座標部分空間　$\{1, 2, \ldots, n\}$ の部分集合 I に対して、$x_i\ (i \in I)$ 以外の座標がすべて 0 である部分集合は $\mathbb{R}^n$ の部分集合である。

このようなものを座標部分空間といい、$\mathbb{R}^I$ と書く。

$$\mathbb{R}^I = \langle \boldsymbol{e}_i \mid i \in I \rangle$$

と表すこともできる。

零ベクトルだけからなる部分集合も部分空間

零ベクトル $\boldsymbol{o}$ だけからなる部分集合 $\{\boldsymbol{o}\}$ も部分空間である。

部分空間における零ベクトルの包含　部分空間は必ず零ベクトル $\boldsymbol{o}$ を含む。

証明

V は空集合でないので、ある $\boldsymbol{v} \in V$ をとるとき、線形部分空間の定義 (ii) より

$$0 \cdot \boldsymbol{v} = \boldsymbol{o} \in V$$

よって部分空間は必ず $\boldsymbol{o}$ を含む。　■

線形写像の像は部分空間

⚓ 線形写像の像は部分空間　線形写像 $f\colon V \to W$ の像 $\mathrm{Im}(f)$ は W の部分空間である。

証明

和について

$\boldsymbol{u}, \boldsymbol{v} \in \mathrm{Im}(f)$ とすると、$\boldsymbol{u} = f(\boldsymbol{v}_1)$, $\boldsymbol{v} = f(\boldsymbol{v}_2)$ とおける。

よって、f の線形性より、

$$\begin{aligned}\boldsymbol{u} + \boldsymbol{v} &= f(\boldsymbol{v}_1) + f(\boldsymbol{v}_2) \\ &= f(\boldsymbol{v}_1 + \boldsymbol{v}_2)\end{aligned}$$

となり、$\mathrm{Im}(f)$ は和について閉じている。　■

スカラー倍について

$\boldsymbol{u} \in \mathrm{Im}(f)$ と $c \in \mathbb{R}$ をとると、$\boldsymbol{u} = f(\boldsymbol{v})$ とおける。

よって、f の線形性より、

$$\begin{aligned}c\boldsymbol{u} &= cf(\boldsymbol{v}) \\ &= f(c\boldsymbol{v})\end{aligned}$$

となり、$\mathrm{Im}(f)$ はスカラー倍について閉じている。　■

線形写像の核は部分空間

⚓ 部分空間の零ベクトルと線形写像　部分空間 V, W の間の線形写像 $f\colon V \to W$ に対して、V の零ベクトルを $\boldsymbol{o}_V$、W の零ベクトルを $\boldsymbol{o}_W$ とすると、

$$f(\boldsymbol{o}_V) = \boldsymbol{o}_W$$

証明

任意の $\boldsymbol{v} \in V$, $\boldsymbol{w} \in W$ に対して、

$$
\begin{aligned}
0 \cdot \boldsymbol{v} &= \boldsymbol{o}_V \\
0 \cdot \boldsymbol{w} &= \boldsymbol{o}_W
\end{aligned}
$$

が成り立つ。

$f(\boldsymbol{o}_V)$ は、f の線形性により、次のように変形できる。

$$
f(\boldsymbol{o}_V) = f(0 \cdot \boldsymbol{v}) = 0 \cdot f(\boldsymbol{v})
$$

ここで、$f(\boldsymbol{v})$ は、f による $\boldsymbol{v} \in V$ の像であるので、W に属する。

そこで、$\boldsymbol{w} = f(\boldsymbol{v})$ とおくと、

$$
\begin{aligned}
f(\boldsymbol{o}_V) &= 0 \cdot f(\boldsymbol{v}) \\
&= 0 \cdot \boldsymbol{w} \\
&= \boldsymbol{o}_W
\end{aligned}
$$

となり、目標としていた式が示された。 ■

線形写像の核は部分空間　線形写像 $f : V \to W$ の核 $\mathrm{Ker}(f)$ は V の部分空間である。

証明

前述の定理の主張 $f(\boldsymbol{o}_V) = \boldsymbol{o}_W$ より、零ベクトルは核空間に属する。

$$
\boldsymbol{o} \in \mathrm{Ker}(f)
$$

和について

$\boldsymbol{u}, \boldsymbol{v} \in \mathrm{Ker}(f)$ とすると、$f(\boldsymbol{u}) = \boldsymbol{o}$ かつ $f(\boldsymbol{v}) = \boldsymbol{o}$ である。

よって、f の線形性より、

$$
\begin{aligned}
f(\boldsymbol{u} + \boldsymbol{v}) &= f(\boldsymbol{u}) + f(\boldsymbol{v}) \\
&= \boldsymbol{o} + \boldsymbol{o} = \boldsymbol{o}
\end{aligned}
$$

したがって、$\boldsymbol{u}+\boldsymbol{v} \in \mathrm{Ker}(f)$ である。 ■

スカラー倍について

$\boldsymbol{u} \in \mathrm{Ker}(f)$ と $c \in \mathbb{R}$ をとると、$f(\boldsymbol{u})=\boldsymbol{o}$ である。

よって、f の線形性より、

$$\begin{aligned} f(c\boldsymbol{u}) &= cf(\boldsymbol{u}) \\ &= c \cdot \boldsymbol{o} = \boldsymbol{o} \end{aligned}$$

したがって、$c\boldsymbol{u} \in \mathrm{Ker}(f)$ である。 ■

ベクトルが張る空間

$m \times n$ 型行列 A で写れる範囲を $\mathrm{Im}\,A$ として定義した。

$\boldsymbol{x}$ を n 次元ベクトルとすると、$\mathrm{Im}\,A$ は次のようなものといえる。

$\boldsymbol{x}$ をいろいろ動かしたときの、

$\boldsymbol{y}=A\boldsymbol{x}$ が動ける範囲が $\mathrm{Im}\,A$

ここで、A を列ベクトルを並べたもの $A=(\boldsymbol{a}_1,\dots,\boldsymbol{a}_n)$ として書き、$\boldsymbol{x}$ も成分 $x_1,\dots,x_n$ で書けば、

$$\boldsymbol{y}=\begin{pmatrix} \boldsymbol{a}_1 & \cdots & \boldsymbol{a}_n \end{pmatrix}\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = x_1\boldsymbol{a}_1+\cdots+x_n\boldsymbol{a}_n$$

つまり、

数 $x_1,\dots,x_n$ をいろいろ動かしたときの、

$x_1\boldsymbol{a}_1+\cdots+x_n\boldsymbol{a}_n$ が動ける範囲が $\mathrm{Im}\,A$

であり、この線形結合が動ける範囲を「ベクトル $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ の張る空間」という。

ベクトルが張る空間　k 個のベクトル $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_k \in \mathbb{R}^n$ を与えたとき、$\boldsymbol{a}_1, \ldots, \boldsymbol{a}_k$ の線形結合全体の集合を

$$\langle \boldsymbol{a}_1, \ldots, \boldsymbol{a}_k \rangle$$

あるいは

$$\operatorname{span}\{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_k\}$$

によって表し、これを $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_k$ が張る空間という。

ベクトルが張る空間は部分空間

ベクトルが張る空間は部分空間　$\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k \in \mathbb{R}^n$ が張る空間 $\langle \boldsymbol{v}_1, \ldots, \boldsymbol{v}_k \rangle$ は部分空間である

証明

[Todo 27: book: 行列と行列式の基礎 p94 命題 3.1.2]

部分空間の張る空間は部分空間　$V \subset \mathbb{R}^n$ を部分空間、$\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k \in V$ とすると、

$$\langle \boldsymbol{v}_1, \ldots, \boldsymbol{v}_k \rangle \subset V$$

証明

[Todo 28: book: 行列と行列式の基礎 p94 命題 3.1.4]

ベクトルが張る空間の幾何的解釈

ベクトル $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ の張る空間 $\langle \boldsymbol{a}_1, \ldots, \boldsymbol{a}_n \rangle$ は、$\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ で定まる平面の一般化といえる。（ここで、点は 0 次元平面、直線は 1 次元平面と考える。）

- $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ がすべて $\boldsymbol{o}$ なら、$\boldsymbol{o}$ ただ一点が $\langle \boldsymbol{a}_1, \ldots, \boldsymbol{a}_n \rangle$
- $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ がすべて一直線上にあれば、その直線が $\langle \boldsymbol{a}_1, \ldots, \boldsymbol{a}_n \rangle$
- $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ がすべて平面上にあれば、その平面が $\langle \boldsymbol{a}_1, \ldots, \boldsymbol{a}_n \rangle$

ベクトルが張る空間と有限従属性

⚓ 有限従属性定理の抽象版　$\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k \in \mathbb{R}^n$ とする

$\langle \boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k \rangle$ に含まれる k 個よりも多い個数のベクトルの集合は線形従属である

証明

[Todo 29: book: 行列と行列式の基礎 p41 （問 1.14）]

基底と次元

部分空間のパラメータ表示を与えるために基準として固定するベクトルの集合を定式化すると、基底という概念になる。

基底は、座標空間の「座標軸」に相当するものであり、部分空間を生成する独立なベクトルの集合として定義される。

基底　V を $\mathbb{R}^n$ の部分空間とする。ベクトルの集合 $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k\} \subset V$ は、次を満たすとき V の基底であるという。

i. $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k\}$ は線型独立である
ii. $V = \langle \boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k \rangle$

線形空間 V の基底 $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k\}$ を 1 つ見つけたら、ベクトルの個数を数えて、V の次元が k であるとする。

次元　線形空間 V の基底をなすベクトルの個数を V の次元といい、$\dim V$ と書く。
また、$\dim\{\boldsymbol{o}\} = 0$ と定義する。

基底の例：標準基底

たとえば、基本ベクトルの集合 $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n\}$ は $\mathbb{R}^n$ の基底であり、これを $\mathbb{R}^n$ の標準基底という。
標準基底 $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n\}$ は n 個のベクトルからなるため、$\mathbb{R}^n$ の次元は n である。

数ベクトル空間の標準基底　数ベクトル空間 K^n において、基本ベクトルの集合 $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n\}$ は K^n の基底である。

証明

部分空間を生成すること

任意のベクトル $\boldsymbol{v} \in K^n$ は、次のように表せる。

$$\boldsymbol{v} = v_1\boldsymbol{e}_1 + \cdots + v_n\boldsymbol{e}_n$$

したがって、K^n は $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ によって生成される。　■

線型独立であること

$\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ の線形関係式

$$c_1\boldsymbol{e}_1 + \cdots + c_n\boldsymbol{e}_n = \boldsymbol{o}$$

を考える。

このとき、左辺は

$$c_1\boldsymbol{e}_1 + \cdots + c_n\boldsymbol{e}_n = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$$

と書き換えられるので、これが零ベクトルになるためには、

$$c_1 = 0, \quad , \cdots , \quad c_n = 0$$

でなければならない。

よって、$\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ は線型独立である。 ■

基底と次元の定義の裏付け

このように基底と次元を定義するにあたって、次の保証が必要になる。

i. 任意の部分空間に、基底の定義を満たす有限個のベクトルが存在すること（基底の存在）
ii. 任意の部分空間に対して、基底をなすベクトルの個数が、基底の選び方によらず一定であること（次元の不変性）

基底の存在

線型独立なベクトルの延長

基底の構成と存在を示すために、次の補題を用いる。

⚓ 線型独立なベクトルの延長　V を K^n の $\{\boldsymbol{o}\}$ でない部分空間とする。
このとき、V の線型独立なベクトル $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_m$ と、V に入らないベクトル $\boldsymbol{a}$ は線型独立である。

証明

$\boldsymbol{a}, \boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_m$ が線型従属であるとする

すると、定理「線形結合によるベクトルの表現」より、$\boldsymbol{a}$ は $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_m$ の線形結合で表され、V に入り、矛盾する

よって、$\boldsymbol{a}, \boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_m$ は線型独立である　■

この定理は、ベクトルの集合が張る空間の記号を用いると、次のように簡潔にまとめられる。

⚓ 線型独立なベクトルの延長　$\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$ が線型独立であって、$\boldsymbol{v}_{k+1} \notin \langle \boldsymbol{v}_1, \ldots, \boldsymbol{v}_k \rangle$ ならば、$\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k, \boldsymbol{v}_{k+1}\}$ は線型独立である

線型独立なベクトルの基底への拡張

K^n の $\{\boldsymbol{o}\}$ でない部分空間 V の線型独立なベクトルは、V の基底に拡張できる。

⚓ 基底の存在　K^n の $\{\boldsymbol{o}\}$ でない部分空間 V には基底が存在する。

証明

$V \neq \{\boldsymbol{o}\}$ なので、V には少なくとも 1 つのベクトル $\boldsymbol{v}_1 \neq \boldsymbol{o}$ が存在する。

定理「単一ベクトルの線型独立性と零ベクトル」より、$\{\boldsymbol{v}_1\}$ は線型独立である

このとき、$\langle \boldsymbol{v}_1 \rangle \subset V$ であるが、もしも $\langle \boldsymbol{v}_1 \rangle = V$ ならば、$\{\boldsymbol{v}_1\}$ は V の基底である。

$\langle \boldsymbol{v}_1 \rangle \subsetneq V$ ならば、$\boldsymbol{v}_2 \not\subsetneq \langle \boldsymbol{v}_1 \rangle$ であるベクトルを V から選ぶことができる。
補題「線型独立なベクトルの延長」より、$\{\boldsymbol{v}_1, \boldsymbol{v}_2\}$ は線型独立である。

このとき、$\langle \boldsymbol{v}_1, \boldsymbol{v}_2 \rangle \subset V$ であるが、もしも $\langle \boldsymbol{v}_1, \boldsymbol{v}_2 \rangle = V$ ならば、$\{\boldsymbol{v}_1, \boldsymbol{v}_2\}$ は V の基底である。

$\langle \boldsymbol{v}_1, \boldsymbol{v}_2 \rangle \subsetneq V$ ならば、$\boldsymbol{v}_3 \not\subsetneq \langle \boldsymbol{v}_1, \boldsymbol{v}_2 \rangle$ であるベクトルを V から選ぶことができる。
補題「線型独立なベクトルの延長」より、$\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$ は線型独立である。

以下同様に続けると、$\langle \boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k \rangle = V$ となるまで、V に属するベクトルを選び続けることができる。

ここで線型独立なベクトルを繰り返し選ぶ操作が無限に続かないこと（有限値 k が存在すること）は、有限従属性定理により、K^n の中には n 個を超える線型独立なベクトルの集合は存在しないことから保証される。 ■

基底の延長

基底の存在証明で行った基底の構成をさらに続けることで、次の定理が得られる。

⚓ **基底の延長** V を n 次元の線形空間とし、線型独立なベクトル $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m \in V$ が与えられたとする。
このとき、$(n-m)$ 個のベクトル $\boldsymbol{v}_{m+1}, \ldots, \boldsymbol{v}_n \in V$ を追加して、$\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m, \boldsymbol{v}_{m+1}, \ldots, \boldsymbol{v}_n\}$ が V の基底になるようにできる。

証明

基底の存在の証明において、線型独立なベクトル $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m \in V$ が得られたところからスタートし、同様の手続きを繰り返せばよい。 ■

次元の不変性

⚓ 次元の不変性　K^n の部分空間 V の基底をなすベクトルの個数（次元）は一定である。

つまり、$\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$ と $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_l\}$ がともに V の基底ならば、$k = l$ である。

証明

$\boldsymbol{u}_1, \ldots, \boldsymbol{u}_l \in \langle \boldsymbol{v}_1, \ldots, \boldsymbol{v}_k \rangle$ であり、$\boldsymbol{u}_1, \ldots, \boldsymbol{u}_l$ は線型独立であるから、有限従属性定理の抽象版より、$l \leq k$ である。

同様にして $k \leq l$ も成り立つので、$k = l$ である。　■

線型独立なベクトルと次元

⚓ 線形独立なベクトルの最大個数と空間の次元　線形空間 V 中の線型独立なベクトルの最大個数は $\dim V$ と等しい。

証明

V の基底を $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$ とすると、V には k 個の線型独立なベクトルが存在する。

また、$V = \langle \boldsymbol{v}_1, \ldots, \boldsymbol{v}_k \rangle$ であるため、有限従属性定理の抽象版より、V 中の線型独立なベクトルの個数は k を超えることはない。

つまり、k は V に含まれる線型独立なベクトルの最大個数である。　■

⚓ **線形空間を生成するベクトルの最小個数と次元**　線形空間 V を張るベクトルの最小個数は $\dim V$ と等しい。

証明

[Todo 30: book: 行列と行列式の基礎 p100 問 3.3]

基底による線形写像の比較

一般に、$\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n\}$ を V の基底とするとき、線形写像 f に対して $f(\boldsymbol{u}_1), \ldots, f(\boldsymbol{u}_n)$ の値が測定できれば、f の形を特定できる。

$$
\begin{aligned}
f(\boldsymbol{v}) &= f(v_1\boldsymbol{u}_1 + \cdots + v_n\boldsymbol{u}_n) \\
&= \underbrace{v_1}_{\text{変数}} \overbrace{\boxed{f(\boldsymbol{u}_1)}}^{\text{定数}} + \cdots + \underbrace{v_n}_{\text{変数}} \overbrace{\boxed{f(\boldsymbol{u}_n)}}^{\text{定数}}
\end{aligned}
$$

上の式において、$v_1, \ldots, v_n$ は $\boldsymbol{v}$ によって決まる変数である。

$f(\boldsymbol{u}_1), \ldots, f(\boldsymbol{u}_n)$ の値さえ決まれば、どんな $\boldsymbol{v}$ を入れても f の値が定まる。これが「f の形が決まる」ということである。

このことを利用すると、基底に関する線形写像の値を比較することで、複数の線形写像を区別することができる。

⚓ **基底上の値による線型写像の同一性判定**　f, g をともに V から W への線型写像とする。f と g が等しいとは、V のある基底 $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ に対して、次が成り立つことと同値である。

$$
f(\boldsymbol{v}_i) = g(\boldsymbol{v}_i) \quad (i = 1, \ldots, n)
$$

証明

$f = g \Longrightarrow f(\boldsymbol{v}_i) = g(\boldsymbol{v}_i)$

$f = g$ ならば、任意の $\boldsymbol{v} \in V$ に対して $f(\boldsymbol{v}) = g(\boldsymbol{v})$ が成り立つ。

よって、基底ベクトル $\boldsymbol{v}_i$ に対しても $f(\boldsymbol{v}_i) = g(\boldsymbol{v}_i)$ が成り立つ。 ■

$f(\boldsymbol{v}_i) = g(\boldsymbol{v}_i) \Longrightarrow f = g$

V の任意のベクトル $\boldsymbol{v}$ は、基底ベクトルの線形結合として表される。

$$\boldsymbol{v} = c_1\boldsymbol{v}_1 + \cdots + c_n\boldsymbol{v}_n$$

このとき、線型写像の線形性から、

$$\begin{aligned} f(\boldsymbol{v}) &= f(c_1\boldsymbol{v}_1 + \cdots + c_n\boldsymbol{v}_n) \\ &= c_1 f(\boldsymbol{v}_1) + \cdots + c_n f(\boldsymbol{v}_n) \\ &= c_1 g(\boldsymbol{v}_1) + \cdots + c_n g(\boldsymbol{v}_n) \\ &= g(c_1\boldsymbol{v}_1 + \cdots + c_n\boldsymbol{v}_n) = g(\boldsymbol{v}) \end{aligned}$$

よって、任意の $\boldsymbol{v} \in V$ に対して $f(\boldsymbol{v}) = g(\boldsymbol{v})$ が成り立つので、$f = g$ である。 ■

第 11 章

線形写像の階数

線形写像の像と列空間

ベクトル $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ の張る空間の記号を用いると、ベクトルの張る空間と $\mathrm{Im}\, A$ に関する考察は次のようにまとめられる。

$$\mathrm{Im}\, A = \langle \boldsymbol{a}_1, \ldots, \boldsymbol{a}_n \rangle$$

つまり、A の列ベクトルが張る空間が $\mathrm{Im}\, A$ である。

このことから、$\mathrm{Im}\, A$ を A の列空間と呼ぶこともある。

線形写像の像と表現行列の列空間の一致　線形写像 $f\colon \mathbb{R}^n \to \mathbb{R}^m$ の像 $\mathrm{Im}\, f$ は、f の表現行列の列ベクトルが張る空間である。

証明

線形写像 $f\colon \mathbb{R}^n \to \mathbb{R}^m$ の表現行列を $A = (\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n)$ とするとき、$\boldsymbol{v} \in \mathbb{R}^n$ に対して、

$$f(\boldsymbol{v}) = A\boldsymbol{v} = v_1\boldsymbol{a}_1 + v_2\boldsymbol{a}_2 + \cdots + v_n\boldsymbol{a}_n$$

なので、

$$
\begin{aligned}
&\boldsymbol{u} \in \operatorname{Im} f \\
\iff& \exists \boldsymbol{v} \in \mathbb{R}^n \ s.t.\ \boldsymbol{u} = f(\boldsymbol{v}) \\
\iff& \exists v_1, \ldots, v_n \in \mathbb{R} \ s.t.\ \boldsymbol{u} = v_1 \boldsymbol{a}_1 + \cdots + v_n \boldsymbol{a}_n \\
\iff& \boldsymbol{u} \in \langle \boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n \rangle
\end{aligned}
$$

したがって、

$$\operatorname{Im} f = \operatorname{Im} A = \langle \boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n \rangle$$

が成り立つ。 ■

上述の証明の

$$\boldsymbol{u} \in \operatorname{Im} f \iff \exists \boldsymbol{v} \in \mathbb{R}^n \ s.t.\ \boldsymbol{u} = f(\boldsymbol{v})$$

という変形に着目すると、この定理は次のように線型方程式の文脈で言い換えられる。

⚓ **線形写像の像空間と方程式の解の存在**　$\boldsymbol{b} \in \mathbb{R}^m$ に対して

$$\boldsymbol{b} \in \operatorname{Im} A \iff \text{方程式 } A\boldsymbol{x} = \boldsymbol{b} \text{ が解を持つ}$$

$\boldsymbol{b} \in \mathbb{R}^m$ が $\operatorname{Im} A$ に属するかどうかを調べるためには階数による判定条件が使える。

線形写像の像空間の基底

線形写像の像空間は表現行列の列ベクトルによって張られるが、列ベクトルの集合は一般には線型独立ではない。
像空間の基底を得るためには、列ベクトルの部分集合、たとえば主列ベクトルを考えるのが自然である。

⚓ **主列ベクトルによる像空間の基底の構成**　行列 A の主列ベクトルの集合は $\operatorname{Im} A$ の基底である。

証明

[Todo 31: book: 行列と行列式の基礎 p97 定理 3.1.10]

線形写像の階数

次の定理は、行列の階数のさらに本質的な意味を明らかにし、行列の階数が行変形の仕方によらずに決まることを念押しするような定理である。

行列の階数と像空間の次元の一致　行列の階数は像空間の次元である。
すなわち、A を $m \times n$ 型行列とするとき、

$$\operatorname{rank} A = \dim \operatorname{Im} A$$

証明

主列ベクトルによる像空間の基底の構成で示したように、A の主列ベクトル $\boldsymbol{a}_{i_1}, \boldsymbol{a}_{i_2}, \ldots, \boldsymbol{a}_{i_r}$ は $\operatorname{Im} A$ の基底を成す。
よってその個数 $r = \operatorname{rank} A$ は $\operatorname{Im} A$ の次元である。 ■

この定理から、線形写像に対して、像空間の次元をその階数と定める。

線形写像の階数　$f\colon \mathbb{R}^n \to \mathbb{R}^m$ を線形写像とするとき、f の階数を

$$\operatorname{rank} f = \dim \operatorname{Im} f$$

と定義する。

核空間と斉次形方程式の解空間

線形写像 $f\colon \mathbb{R}^n \to \mathbb{R}^m$ の表現行列を A とするとき、

$$\operatorname{Ker} f = \{\boldsymbol{v} \in \mathbb{R}^n \mid A\boldsymbol{v} = \boldsymbol{o}\}$$

と定めると、$f(\boldsymbol{v}) = A\boldsymbol{v}$ という関係から、$\operatorname{Ker} f$ と $\operatorname{Ker} A$ は同じものを指す。

$\operatorname{Ker} A$ すなわち $\operatorname{Ker} f$ とは、f によって $\boldsymbol{o}$ に写ってしまうような、つまり $A\boldsymbol{x} = \boldsymbol{o}$ となるような $\boldsymbol{x}$ すべての集合である。

つまり、$\operatorname{Ker} A$ とは、斉次形の方程式 $A\boldsymbol{x} = \boldsymbol{o}$ の解空間そのものである。

核空間と一般解のパラメータ表示

解のパラメータ表示の再解釈で述べたように、$A\boldsymbol{x} = \boldsymbol{b}$ の解をすべて見つけるには、

1. 1 つの解（特殊解）$\boldsymbol{x}_0$ を見つける
2. $A\boldsymbol{x} = \boldsymbol{o}$ の一般解を求める
3. それらの和が $A\boldsymbol{x} = \boldsymbol{b}$ の一般解となる

という考え方を使うことができた。

このことを $\operatorname{Ker} A$ を用いて定式化できる。

特殊解と核の元による別解の構成　$\boldsymbol{x}_0$ が $A\boldsymbol{x} = \boldsymbol{b}$ の解であるとき、$\operatorname{Ker} A$ に属する任意のベクトル $\boldsymbol{u}$ を用いて、$\boldsymbol{x}_0 + \boldsymbol{u}$ もまた $A\boldsymbol{x} = \boldsymbol{b}$ の解となる。

証明

$\boldsymbol{x}_0$ が $A\boldsymbol{x} = \boldsymbol{b}$ の解であることから、

$$A\boldsymbol{x}_0 = \boldsymbol{b}$$

また、$\boldsymbol{u} \in \operatorname{Ker} A$ より、

$$A\boldsymbol{u} = \boldsymbol{o}$$

よって、

$$\begin{aligned} A(\boldsymbol{x}_0 + \boldsymbol{u}) &= A\boldsymbol{x}_0 + A\boldsymbol{u} \\ &= A\boldsymbol{x}_0 + \boldsymbol{o} \\ &= \boldsymbol{b} \end{aligned}$$

となり、$\boldsymbol{x}_0 + \boldsymbol{u}$ もまた $A\boldsymbol{x} = \boldsymbol{b}$ の解であることがわかる。 ■

そして、どんな解もこの方法で作ることができる。

特殊解と核空間による一般解の構成　$A\boldsymbol{x} = \boldsymbol{b}$ を満たす 1 つの解 $\boldsymbol{x}_0$ が見つかれば、$A\boldsymbol{x} = \boldsymbol{b}$ の一般解は、$A\boldsymbol{x} = \boldsymbol{o}$ の一般解 $\boldsymbol{u}$ を用いて、$\boldsymbol{x}_0 + \boldsymbol{u}$ と表される。

証明

$A\boldsymbol{x} = \boldsymbol{b}$ の 1 つの解を $\boldsymbol{x}_0$、もう 1 つの解を $\boldsymbol{x}_1$ とおくと、

$$A\boldsymbol{x}_0 = \boldsymbol{b},\ A\boldsymbol{x}_1 = \boldsymbol{b}$$

が成り立つので、

$$\begin{aligned} A\boldsymbol{x}_1 - A\boldsymbol{x}_0 = \boldsymbol{b} - \boldsymbol{b} = \boldsymbol{o} \\ \therefore \quad A(\boldsymbol{x}_1 - \boldsymbol{x}_0) = \boldsymbol{o} \end{aligned}$$

となり、$\boldsymbol{x}_1 - \boldsymbol{x}_0$ は $A\boldsymbol{x} = \boldsymbol{o}$ の解である。

ここで、$A\boldsymbol{x} = \boldsymbol{o}$ の一般解 $\boldsymbol{u}$ が得られているなら、$\boldsymbol{x}_1 - \boldsymbol{x}_0$ も $\boldsymbol{u}$ で表すことができる。

したがって、$\boldsymbol{x}_0 + \boldsymbol{u}$ は $A\boldsymbol{x} = \boldsymbol{b}$ のすべての解を網羅する。 ■

解が 1 つ見つかれば、その解 $\boldsymbol{x}_0$ は固定して、$\mathrm{Ker}\,A$ に属するベクトル $\boldsymbol{u}$ をいろいろ変えることにより、$\boldsymbol{x}_0 + \boldsymbol{u}$ ですべての解が得られる。

核空間の基底と基本解

「$\boldsymbol{u}$ をいろいろ変えることにより」という部分をもう少し精密に述べよう。

いろいろ動かしてすべての解を網羅するには、解空間 $\mathrm{Ker}\, A$ の基底が必要である。すなわち、$\boldsymbol{u}$ は $\mathrm{Ker}\, A$ の基底 $\boldsymbol{u}_i$ を用いた次のような形で表される。

$$\boldsymbol{u} = c_1\boldsymbol{u}_1 + \cdots + c_d\boldsymbol{u}_d$$

ここで、$c_1, \ldots, c_d$ は任意であるので、この式は斉次形方程式 $A\boldsymbol{x} = \boldsymbol{o}$ の基本解のパラメータ表示そのものである。

⚓ **斉次形方程式の基本解と核空間の基底**　A を $m \times n$ 型行列とし、$\boldsymbol{u}_1, \ldots, \boldsymbol{u}_d$ を $A\boldsymbol{x} = \boldsymbol{o}$ の基本解とするとき、$\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_d\}$ は $\mathrm{Ker}\, A$ の基底である。

言い換えると、$\mathrm{Ker}\, A$ の元 $\boldsymbol{u}$ は、$A\boldsymbol{x} = \boldsymbol{o}$ の基本解 $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_d$ を使ってパラメータ表示できる。

パラメータの空間と座標部分空間

つまり、基本解 $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_d$ を基準として固定すれば、$\mathrm{Ker}\, A$ の元を 1 つ指定することは、パラメータの値の組

$$\begin{pmatrix} t_1 \\ \vdots \\ t_d \end{pmatrix} \in \mathbb{R}^d$$

を指定することと同じである。

斉次形方程式 $A\boldsymbol{x} = \boldsymbol{o}$ の主変数を $x_{i_1}, \ldots, x_{i_r}$、自由変数を $x_{j_1}, \ldots, x_{j_d}$ とすると、解のパラメータの空間は座標部分空間 $\mathbb{R}^{\{j_1, \ldots, j_d\}}$ である。

そして、そのパラメータ付けは、

$$\mathbb{R}^{\{j_1, \ldots, j_d\}} \ni \sum_{k=1}^{d} t_k \boldsymbol{e}_{j_k} \longmapsto \sum_{k=1}^{d} t_k \boldsymbol{u}_k \in \mathrm{Ker}\, A$$

によって与えられる。

核空間の次元と解の自由度

$\boldsymbol{b} = \boldsymbol{o}$ でない一般の連立方程式 $A\boldsymbol{x} = \boldsymbol{b}$ においても、基本解の個数 d は解の自由度であり、$\boldsymbol{u}_1, \ldots, \boldsymbol{u}_d$ は $\mathrm{Ker}\, A$ の基底をなすため、

$\mathrm{Ker}\, A$ の次元は、$A\boldsymbol{x} = \boldsymbol{b}$ の解の自由度と一致する

ということがいえる。

次元定理

連立方程式 $A\boldsymbol{x} = \boldsymbol{b}$ の解の自由度は、

$$\text{解の自由度} = (\text{変数の個数}) - \mathrm{rank}(A)$$

で表された

この関係は、$\boldsymbol{b} = \mathbf{0}$、すなわち斉次形の場合にも成り立つ

そこで、変数の個数を n とおくと、次のようにも書き換えられる

$$\mathrm{rank}(A) = n - (A\boldsymbol{x} = \mathbf{0} \text{ の解の自由度})$$

線型方程式と階数に関するこの関係を、線形写像と次元の言葉で言い換えたい

次のような線形写像

$$\begin{array}{cccc} f\colon & \mathbb{R}^n & \longrightarrow & \mathbb{R}^m \\ & \rotatebox{90}{$\in$} & & \rotatebox{90}{$\in$} \\ & \boldsymbol{x} & \longmapsto & A\boldsymbol{x} \end{array}$$

を考えると、

- 写像 f は、行列 A に対応する
- 変数の個数は、$\boldsymbol{x}$ の動く空間 $\mathbb{R}^n$ の次元 n に対応する
- $A\boldsymbol{x} = \mathbf{0}$ の解の自由度は、写像 f で $\mathbf{0}$ になってしまうものの次元に対応する

という関係が読み取れる

ここで、写像 f で $\mathbf{0}$ になってしまう縮退するものは、写像 f の核 $\mathrm{Ker}(f)$ である

このことを用いて関係式を表現し直すと、次のようになる

$$\mathrm{rank}(f) = n - \dim \mathrm{Ker}(f)$$

⚓ **線形写像の次元定理**　$f : \mathbb{R}^n \to \mathbb{R}^m$ を線形写像とすると、次が成り立つ

$$\mathrm{rank}(f) = n - \dim \mathrm{Ker}(f)$$

証明

A を f の表現行列とし、$\mathrm{rank}(f) = r$ とする

このとき、$\mathrm{Ker}(f)$ の次元は $A\boldsymbol{x} = \mathbf{0}$ の解空間の自由度 $n - r$ と一致するため、

$$\begin{aligned} \dim \mathrm{Ker}(f) &= n - r \\ &= n - \mathrm{rank}(f) \\ \therefore \quad \mathrm{rank}(f) &= n - \dim \mathrm{Ker}(f) \end{aligned}$$

となり、定理が成り立つ　■

核空間の次元による正則の判定

次元定理から、次のような正則判定法が得られる。

⚓ **核空間の次元による正則判定**　n 次正方行列 A に対して、

$$\begin{aligned} A \text{ が正則行列} &\iff \mathrm{Ker}\, A = \{\boldsymbol{o}\} \\ &\iff \dim \mathrm{Ker}\, A = 0 \end{aligned}$$

証明

階数による正則判定より、A が正則であることは、

$$\operatorname{rank} A = n$$

であることと同値である。

ここで、次元定理より、

$$\operatorname{rank} A + \dim \operatorname{Ker} A = n$$

$\operatorname{rank} A = n$ を代入し、整理すると、

$$\dim \operatorname{Ker} A = 0$$

が得られる。 ■

階数の性質

2 つの行列の階数の和 A, B を同じ型の行列とするとき、

$$\operatorname{rank}(A + B) \leq \operatorname{rank}(A) + \operatorname{rank}(B)$$

証明

[Todo 32: book: 行列と行列式の基礎 p44 問 1.15]

第 12 章

行列式

連立方程式の解の判別式としての行列式

[Todo 33:]

置換と互換

たとえば、$(1, 2, 3, 4)$ を並び替えた列 (i, j, k, l) があるとして、

$$
\begin{aligned}
1 &\longmapsto i \\
2 &\longmapsto j \\
3 &\longmapsto k \\
4 &\longmapsto l
\end{aligned}
$$

というように、番号を並び替える操作そのものを写像とみなし、置換と呼ぶ

置換　集合 $\{1, 2, \ldots, n\}$ からそれ自身への写像 σ が全単射であるとき、σ は n 次の置換であるという

たとえば、

$$\sigma(1) = 2, \quad \sigma(2) = 3, \quad \sigma(3) = 1$$

によって 3 次の置換を定めることができる

この置換を、

$$\sigma = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$$

と表記する

置換の積

写像とみる利点の 1 つは、積が定義できることである

もう 1 つの置換

$$\tau = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}$$

が与えられたとき、合成写像 $\sigma \circ \tau$ は、

$$\begin{aligned} 1 &\xmapsto{\tau} 1 \xmapsto{\sigma} 2 \\ 2 &\xmapsto{\tau} 3 \xmapsto{\sigma} 1 \\ 3 &\xmapsto{\tau} 2 \xmapsto{\sigma} 3 \end{aligned}$$

なので、

$$\sigma\tau = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$$

である

通常、合成の記号 $\circ$ を書かずに $\sigma\tau$ と表記する

なお、$\sigma\tau$ と $\tau\sigma$ は一般に異なる

写像の合成の結合法則から、置換の積でも結合法則が成り立つ

⚓ 置換の積の結合法則

$$(\sigma\tau)\rho = \sigma(\tau\rho)$$

恒等置換

恒等写像

$$\begin{gathered}\mathrm{id}\colon \{1, 2, \ldots, n\} \longmapsto \{1, 2, \ldots, n\} \\ \mathrm{id}(i) = i \quad (1 \leq i \leq n)\end{gathered}$$

は置換であるので、これを恒等置換と呼び、

$$e = \mathrm{id}$$

と書く

任意の置換 σ に対して、明らかに

$$\sigma e = e\sigma = \sigma$$

が成り立つ

また、次の性質はのちに行列式の性質を議論する際に重要になる

恒等置換の単調性による特徴づけ　$i \leq \sigma(i)$（あるいは $i \geq \sigma(i)$）を満たす置換 σ は恒等置換しか存在しない

証明

σ が恒等置換でないと仮定する

条件 $i \leq \sigma(i)$ より、「元の位置より後ろに移される」、すなわち「すべてが自分以上に移る」ことになる

たとえば、1 を 2 に、2 を 3 に、...、$n-1$ を n に写す置換を考える

しかし、集合 $\{1, 2, \ldots, n\}$ の要素は n 個しかないので、n を $n+1$ に写すことはできない

そこで、n を n に写すとすると、$n-1$ も n も n に写ることになり、これは置換が全単射であるという定義に反する

$i \geq \sigma(i)$ の場合も、「元の位置より前に移される」、すなわち「すべてが自分以下に移る」ことになると考えると、同様の矛盾が生じる

よって、σ は恒等置換でなければならない ■

逆置換

置換 σ は、定義より全単射であるので、逆写像 σ^{-1} が存在する

これを逆置換と呼ぶ

置換の集合

すべての n 次の置換からなる集合は群と呼ばれる構造を持っている

これを n 次対称群と呼び、記号 S_n で表す

互換

置換の中で最も基本的なのは、2 文字だけを交換する置換である

互換 $1 \leq i \neq j \leq n$ のとき、$\sigma(i) = j,\ \sigma(j) = i$ であって、k が $i,\ j$ 以外のとき $\sigma(k) = k$ とすることで得られる置換を

$$\sigma = (ij)$$

と書き、このような置換を互換という

たとえば、

$$(24) = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 4 & 3 & 2 & 5 \end{pmatrix}$$

互換の逆置換

互換は (ij) と書いても (ji) と書いても同じ操作を表す

i と j を交換してから j と i を交換すると元に戻るが、この (ij) と (ji) は互換としては同じなので、

> 互換の逆置換は自分自身

である

置換の一行表示

置換を表す 2 行の表示は、下の行だけで情報としては十分なので、たとえば

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 4 & 3 & 2 & 5 \end{pmatrix}$$

を $\sigma = 14325$ などと書いてしまうと便利である

これを σ の一行表示と呼ぶ

互換と置換の積

一行表示を用いた場合、互換と置換の積はたとえば次のように書ける

$\sigma = 14325$ とすると、

$$(12)\sigma = 24315, \quad \sigma(12) = 41325$$

$(12)\sigma$ は、$\sigma = 14325$ に互換 (12) を作用させて、24315 となる

$\sigma(12)$ は、12345 に互換 (12) を作用させて 21345 とし、さらに置換 σ を作用させることを意味する

置換 σ は、4 と 2 を入れ替える置換なので、21345 に対して σ を作用させると、41325 となる

この例の結果を一般的に述べると、次のようになる

⚓ **互換と置換の積** $\sigma \in S_n$ に対して、$\tau = (ij)$ を左からかけた $\tau\sigma$ の一行表示は、σ の数字 i と j を交換したものである

また、τ を右からかけた $\sigma\tau$ の一行表示は、σ の i 番目の数字と j 番目の数字を交換したものである

互換の積への分解

たとえば、$\sigma = 2413$ とすると、これは、

1. 1234 の 3 と 4 を交換して 1243
2. 1243 の 1 と 2 を交換して 2143
3. 2143 の 2 と 3 を交換して 2413

というように、互換に分解して考えることができる

数式でまとめると、

$$\sigma = (34)(12)(23)$$

互換の積への置換の分解　任意の置換 σ は、いくつかの互換の積として書ける

証明

n に対する帰納法を用いる

$n = 1$ のときは、互換の定義における i, j の条件を満たさず、i, j 以外の k について $\sigma(k) = k$ とすることで得られる置換に相当するので、1 つの互換とみなせる

$(n-1)$ 次以下の置換が互換の積で書けることを仮定する

σ を n 次の置換とし、$\sigma(n)$ の値を c とする

$c = n$ すなわち $\sigma(c) = c$ の場合、σ は c をまったく動かしていないため、実質的に $c-1$ までの数字だけを並び替えていることになる

そのため、σ は $c-1$ すなわち $(n-1)$ 次の置換とみなせるため、帰納法の仮定より、互換の積として書ける

$c \neq n$ の場合、$\sigma(c)$ を d とし、d と c を交換する互換 $\tau = (cd)$ を考える

このとき、$\tau\sigma$ は、σ の数字 c と d を交換したものであるので、

$$\tau\sigma = \begin{pmatrix} 1 & 2 & \cdots & c-1 & c & \cdots & n \\ 1 & 2 & \cdots & c-1 & \sigma(c) & \cdots & n \end{pmatrix}$$

c が n に一致しないという仮定をふまえると、

$$\tau\sigma(n) = n$$

であることが読み取れる

よって、$\tau\sigma$ は実質的に $(n-1)$ 次の置換とみなせるので、帰納法の仮定より、互換の積として書ける

$$\tau\sigma = \tau_1\tau_2\cdots\tau_m$$

ゆえに、

$$\sigma = \tau^{-1}\tau_1\tau_2\cdots\tau_m$$

であるが、互換の逆置換は自分自身であるので、

$$\sigma = \tau\tau_1\tau_2\cdots\tau_m$$

と書ける ■

置換の符号と偶奇

すべての置換は互換の積に分解できるが、その方法は一通りではない

しかし、互換の積の個数の偶奇性は、置換が与えられれば定まる

このことを証明するために、置換と多項式の関係を考察する

置換の多項式への作用

置換 $\sigma \in S_n$ と n 変数多項式 $f = f(x_1, x_2, \ldots, x_n)$ が与えられたとき、変数 x_i に $x_{\sigma(i)}$ を代入することにより、式 σf を

$$(\sigma f)(x_1, \ldots, x_n) = f(x_{\sigma(1)}, \ldots, x_{\sigma(n)})$$

と定める

⚓ 置換作用の結合法則　$f = f(x_1, \ldots, x_n)$ を n 変数の多項式とし、$\sigma, \tau \in S_n$ とするとき、

$$(\sigma\tau)f = \sigma(\tau f)$$

証明

式 τf は、

$$(\tau f)(x_1, \dots, x_n) = f(x_{\tau(1)}, \dots, x_{\tau(n)})$$

である

さらに σ を作用させると、$x_{\tau(i)}$ は $x_{\sigma(\tau(i))} = x_{(\sigma\tau)(i)}$ に置き換わるので、

$$\begin{aligned}(\sigma(\tau f)) &= f(x_{(\sigma\tau)(1)}, \dots, x_{(\sigma\tau)(n)}) \\ &= ((\sigma\tau)f)(x_1, \dots, x_n)\end{aligned}$$

が成り立つ ■

互換の差積への作用

次のような n 変数の多項式を差積と呼ぶ

$$\begin{array}{cccc} (x_1 - x_2) & (x_1 - x_3) & \cdots & (x_1 - x_n) \\ & (x_2 - x_3) & \cdots & (x_2 - x_n) \\ & & \ddots & \\ & & & (x_{n-1} - x_n) \end{array}$$

差積　次のような n 変数の多項式を差積と呼ぶ

$$\Delta_n = \prod_{1 \le i < j \le n} (x_i - x_j)$$

置換の符号を理解するために、差積を使うことができる

その第一歩となるのが、次の定理である

互換による差積の符号変化　τ を互換とするとき、

$$\tau \Delta_n = -\Delta_n$$

証明

$i < j$ として、$\tau = (ij)$ とすると、各因子 $x_s - x_t\,(1 \leq s < t \leq n)$ の変化は次のようになる

$x_i - x_j$ は $x_j - x_i$ になる

x_i と x_j を入れ替えることで、その差が逆転して符号が反転する

$$x_j - x_i = -(x_i - x_j)$$

よって、この項は -1 倍の効果をもたらす

$s < i < j$ のとき、$x_s - x_i$ と $x_s - x_j$ が入れ替わる

この場合、s は $i,\ j$ より前の添字である

- 互換前：$(x_s - x_i)(x_s - x_j)$
- 互換後：$(x_s - x_j)(x_s - x_i)$

2 つの項が交換されるだけなので、積の絶対値は変わらず、符号にも影響しない

$i < j < s$ のとき、$x_i - x_s$ と $x_j - x_s$ が入れ替わる

この場合、s は $i,\ j$ より後の添字である

- 互換前：$(x_i - x_s)(x_j - x_s)$
- 互換後：$(x_j - x_s)(x_i - x_s)$

この場合も、並び順だけが入れ替わり、符号には影響しない

$i < s < j$ のとき、$x_i - x_s$ と $x_s - x_j$ は…

この場合、s は i と j の間にある添字である

- 互換前：$(x_i - x_s)(x_s - x_j)$
- 互換後：$(x_j - x_s)(x_s - x_i)$

互換前の積を変形してみると、

$$\begin{aligned}(x_i - x_s)(x_s - x_j) &= -(x_i - x_s)(x_j - x_s)\\ &= (x_s - x_i)(x_j - x_s)\\ &= (x_j - x_s)(x_s - x_i)\end{aligned}$$

という形で、互換後の積が得られる

よって、この場合も積の符号は変わらない

以上をふまえると、符号が反転するのは $x_i - x_j$ の項だけである

よって、1 回の互換 (ij) によって、差積全体は (-1) 倍される ■

置換の符号

置換による差積の符号変化　置換 $\sigma \in S_n$ が s 個の互換の積として書けるならば、

$$\sigma\Delta_n = (-1)^s\Delta_n$$

が成り立つ

証明

置換 σ を s 個の互換の積 $\sigma = \tau_1 \cdots \tau_s$ と書いたとき、

$$\sigma\Delta_n = (\tau_1 \cdots \tau_s)\Delta_n$$

置換作用の結合法則を用いて、

$$\sigma\Delta_n = (\tau_1 \cdots \tau_{s-1})(\tau_s\Delta_n)$$

互換による差積の符号変化を繰り返し用いると、

$$\begin{aligned}\sigma\Delta_n &= (\tau_1 \cdots \tau_{s-1})(-\Delta_n)\\ &= (-1)(\tau_1 \cdots \tau_{s-1})\Delta_n\\ &= (-1)^s\Delta_n\end{aligned}$$

が最終的に得られる ■

この定理における $\sigma\Delta_n$ は、σ をどのような互換の積として表すかとは無関係に、σ が与えられれば決まる多項式である

そして、$(-1)^s$ という部分から、σ を互換の積で表したとき、その個数 s が偶数であれば符号は $+$ に、奇数であれば符号は $-$ になることがわかる

このようにして、次の定理が示されたことになる

置換の符号の存在　置換 σ を互換の積として書くとき、用いられる互換の個数の偶奇は σ のみによって決まる

そこで、置換の符号を次のように定義する

置換の符号　置換 $\sigma \in S_n$ を互換の積 $\sigma = \tau_1 \cdots \tau_i$ として書いたとき、σ の符号を

$$\mathrm{sgn}(\sigma) = (-1)^i$$

と定義する

そして、互換の個数の偶奇をそのまま、置換の偶奇として定める

偶置換と奇置換　置換 $\sigma \in S_n$ の符号 $\mathrm{sgn}(\sigma)$ が $+1$ であれば σ を偶置換と呼び、-1 であれば奇置換と呼ぶ

置換の性質

逆置換の符号

$$\mathrm{sgn}(\sigma^{-1}) = \mathrm{sgn}(\sigma)$$

証明

置換 σ を互換の積として書くと、逆置換はその互換の順序を逆にしたものになる

すなわち、$\sigma = \tau_1 \cdots \tau_s$ とすると、

$$\sigma^{-1} = \tau_s^{-1} \cdots \tau_1^{-1}$$

であるが、互換の逆置換は自分自身であるので、

$$\operatorname{sgn}(\sigma^{-1}) = (-1)^s = \operatorname{sgn}(\sigma)$$

が成り立つ ■

置換の符号の乗法性

$$\operatorname{sgn}(\sigma\tau) = \operatorname{sgn}(\sigma)\operatorname{sgn}(\tau)$$

証明

それぞれを互換の積 $\sigma = \tau_1 \cdots \tau_i$、$\tau = \rho_1 \cdots \rho_j$ と書くと、

$$\sigma\tau = \tau_1 \cdots \tau_i \rho_1 \cdots \rho_j$$

である

このとき、$\operatorname{sgn}(\sigma) = (-1)^i$, $\operatorname{sgn}(\tau) = (-1)^j$ なので、

$$\operatorname{sgn}(\sigma\tau) = (-1)^{i+j} = (-1)^i(-1)^j = \operatorname{sgn}(\sigma)\operatorname{sgn}(\tau)$$

が成り立つ ■

置換群の左右作用に対する和の不変性　f を S_n 上の関数とするとき、任意の $\tau \in S_n$ に対して、次が成り立つ

$$\sum_{\sigma \in S_n} f(\tau\sigma) = \sum_{\sigma \in S_n} f(\sigma) = \sum_{\sigma \in S_n} f(\sigma\tau)$$

証明

τ を固定して、σ をすべての置換（S_n の元）全体にわたって動かすとき、$\tau\sigma$ も S_n の全体を動く

言い換えると、写像 $S_n \to S_n$ を $\sigma \longmapsto \tau\sigma$ と定めると、これは全単射である

したがって、

$$\sum_{\sigma \in S_n} f(\sigma) = \sum_{\sigma \in S_n} f(\tau\sigma)$$

が成り立つ

同様に、写像 $S_n \to S_n$ を $\sigma \longmapsto \sigma\tau$ と定めると、これも全単射であるので、同様に、

$$\sum_{\sigma \in S_n} f(\sigma) = \sum_{\sigma \in S_n} f(\sigma\tau)$$

が成り立つことがわかる　■

行列式の定義

ある正方行列の行列式は、

1. 各列から 1 つずつ、行に重複がないように成分を選ぶ
2. それらをかけ合わせる
3. 符号をつけて足す

という手順で定まる値である

行列式　n 次正方行列 $A=(a_{ij})$ に対して、

$$\sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) \prod_{i=1}^{n} a_{i,\sigma(i)}$$

で定められる値を A の行列式と呼び、$|A|$ あるいは $\det(A)$ と表記する

三角行列の行列式

三角行列の場合、各列から 1 つずつ、0 でない成分を重複なく選び出す方法は、対角成分をすべて選ぶしかない

三角行列の行列式　三角行列の行列式は、対角成分の積である

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ & a_{22} & \dots & a_{2n} \\ & & \ddots & \vdots \\ \text{\huge 0} & & & a_{nn} \end{vmatrix} = a_{11}a_{22}\cdots a_{nn}$$

$$\begin{vmatrix} a_{11} & & & \text{\huge 0} \\ a_{21} & a_{22} & & \\ \vdots & & \ddots & \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} = a_{11}a_{22}\cdots a_{nn}$$

証明

行列式において、

$$a_{1,\sigma(1)}a_{2,\sigma(2)}\cdots a_{n,\sigma(n)} = 0$$

となる項は、和をとったときに消えてしまう

したがって、

$$a_{1,\sigma(1)}a_{2,\sigma(2)}\cdots a_{n,\sigma(n)} \neq 0$$

すなわち

$$a_{1,\sigma(1)} \neq 0, \ldots, a_{n,\sigma(n)} \neq 0$$

となるような選び方を考える

上三角行列の場合

上三角行列の定義より、$i > j$ ならば $a_{ij} = 0$ である

$a_{ij} \neq 0$ とするには、$i \leq j$ でなければならないので、$a_{i,\sigma(i)}$ においては、

$$i \leq \sigma(i)$$

である必要がある

そして、この条件を満たす置換は、恒等置換しか存在しないので、

$$\sigma(i) = i$$

より、a_{ii} の積によって行列式の値が構成される

また、恒等置換は 0（偶数）回の互換で構成されるので、各項の符号は正となる ■

下三角行列の場合

下三角行列の定義より、$i < j$ ならば $a_{ij} = 0$ である

$a_{ij} \neq 0$ とするには、$i \geq j$ でなければならないので、$a_{i,\sigma(i)}$ においては、

$$i \geq \sigma(i)$$

である必要がある

そして、この条件を満たす置換も、恒等置換しか存在しないので、上三角行列の場合と同様の結果が得られる ■

対角行列は、上三角行列でもあり下三角行列でもあるので、上の定理の特別な場合として次が成り立つ

対角行列の行列式　対角行列の行列式は、対角成分の積である

$$\begin{vmatrix} a_{11} & & & \\ & a_{22} & & \text{\huge 0} \\ & & \ddots & \\ & \text{\huge 0} & & \ddots & \\ & & & & a_{nn} \end{vmatrix} = a_{11}a_{22}\cdots a_{nn}$$

特に、対角成分がすべて 1 の場合が単位行列である

単位行列の行列式　単位行列の行列式は 1 である

$$|E| = 1$$

行列式の基本性質

次の性質により、以後議論する行列式の性質が列に対して成り立つなら、行に対しても成り立つといえるようになる

行列式の対称性

$$\det\left({}^tA\right) = \det(A)$$

証明

行列式の定義より、行列 ${}^tA$ の行列式は、行列 A の行列式に現れる $a_{i,\sigma(i)}$ の添字を入れ替えたもの $a_{\sigma(i),i}$ の積和になる

$$\det\left({}^tA\right) = \sum_{\sigma\in S_n} \mathrm{sgn}(\sigma) \prod_{i=1}^{n} a_{\sigma(i),i}$$

一方、$j = \sigma(i)$ とおくと、$i = \sigma^{-1}(j)$ となるので、添字の変数を変換して

$$\prod_{i=1}^{n} a_{\sigma(i),i} = \prod_{j=1}^{n} a_{j,\sigma^{-1}(j)}$$

よって、$\det\left({}^t A\right)$ の各項は、

$$\mathrm{sgn}(\sigma^{-1}) \prod_{j=1}^{n} a_{j,\sigma^{-1}(j)}$$

となるが、これは $\det(A)$ の定義式の σ^{-1} に対応する項と同じである

ここで、$\rho = \sigma^{-1}$ とおくと、$\sigma = \rho^{-1}$ であり、逆置換の符号から $\mathrm{sgn}(\sigma) = \mathrm{sgn}(\rho^{-1}) = \mathrm{sgn}(\rho)$ であるから、

$$\det\left({}^t A\right) = \sum_{\rho \in S_n} \mathrm{sgn}(\rho) \prod_{j=1}^{n} a_{j,\rho(j)} = \det(A)$$

よって、$\det\left({}^t A\right) = \det(A)$ が示された ■

行列式の列についての交代性　行列 $A = (\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n)$ において、2 つの列を入れ替えた行列を作ると、その行列の行列式の値は、元の行列 A の行列式の値の (-1) 倍になる

$$\begin{aligned} &\det(\boldsymbol{a}_1, \ldots, \boldsymbol{a}_i, \ldots, \boldsymbol{a}_j, \ldots, \boldsymbol{a}_n) \\ &\qquad = -\det(\boldsymbol{a}_1, \ldots, \boldsymbol{a}_j, \ldots, \boldsymbol{a}_i, \ldots, \boldsymbol{a}_n) \\ &\qquad\qquad (1 \leq i < j \leq n) \end{aligned}$$

証明

元々の行列 A の行列式の各項が、

$$f(\sigma) = \mathrm{sgn}(\sigma) a_{\sigma(1),1} \cdots a_{\sigma(i),i} \cdots a_{\sigma(j),j} \cdots a_{\sigma(n),n}$$

であるのに対し、第 i 列と j 列を入れ替えた行列の行列式の各項は、

$$\mathrm{sgn}(\sigma) a_{\sigma(1),1} \cdots a_{\sigma(i),j} \cdots a_{\sigma(j),i} \cdots a_{\sigma(n),n}$$

となる

ここで、i を j に、j を i に写す互換 $\sigma_0 = (ij)$ を考え、$\tau = \sigma\sigma_0$ とおくと、$\sigma(j) = \tau(i),\ \sigma(i) = \tau(j)$ となるので、

$$f(\tau) = \operatorname{sgn}(\tau) a_{\tau(1),1} \cdots a_{\tau(i),i} \cdots a_{\tau(j),j} \cdots a_{\tau(n),n}$$

このとき、置換群の左右作用に対する和の不変性より、

$$\sum_{\sigma \in S_n} f(\sigma) = \sum_{\sigma \in S_n} f(\sigma\sigma_0) = \sum_{\tau \in S_n} f(\tau)$$

すなわち、σ 全体の総和は τ 全体の総和に一致する

さらに、置換の符号の乗法性より、

$$\operatorname{sgn}(\tau) = \operatorname{sgn}(\sigma) \operatorname{sgn}(\sigma_0) = -\operatorname{sgn}(\sigma)$$

であるから、

$$f(\sigma) = -f(\tau)$$

よって、列の交換後、行列式全体が (-1) 倍される ■

行列式の列についての多重線形性　行列式を列の関数とみたとき、この関数は、どの列についても線形である

$$\begin{aligned} &\det(\boldsymbol{a}_1, \ldots, \alpha\boldsymbol{u} + \beta\boldsymbol{v}, \ldots, \boldsymbol{a}_n) \\ &\qquad = \alpha \det(\boldsymbol{a}_1, \ldots, \boldsymbol{u}, \ldots, \boldsymbol{a}_n) \\ &\qquad\qquad + \beta \det(\boldsymbol{a}_1, \ldots, \boldsymbol{v}, \ldots, \boldsymbol{a}_n) \end{aligned}$$

証明

$\sigma \in S_n$ に対応する各項について、

$$a_{\sigma(1),1} \cdots (\alpha u_{\sigma(i)} + \beta v_{\sigma(i)}) \cdots a_{\sigma(n),n}$$

$C = a_{\sigma(1),1} \cdots a_{\sigma(n),n}$ とし、$A = \alpha u_{\sigma(i)}$, $B = \beta v_{\sigma(i)}$ とおくと、

$$C(A+B) = CA + CB = \alpha C u_{\sigma(i)} + \beta C v_{\sigma(i)}$$

のように展開できる

よって、

$$\alpha(a_{\sigma(1),1} \cdots u_{\sigma(i)} \cdots a_{\sigma(n),n}) \\ + \beta(a_{\sigma(1),1} \cdots v_{\sigma(i)} \cdots a_{\sigma(n),n})$$

を用いれば、行列式の定義に基づいて定理が成り立つことがわかる ■

行列式の対称性より、次の定理も得られる

⚓ **行列式の行についての多重線形性と交代性**　行列式は行に関しても多重線形性と交代性をもつ

以降、列に対して成り立つ性質は行に対しても成り立つとし、列の場合のみを記載する

行列式の値が零になる条件

⚓ **列の重複による行列式の零化**　$A = (\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n)$ の n 個の列の中に、まったく同じものがあれば、

$$\det(A) = 0$$

となる

証明

行列 A の列ベクトルに、共通のベクトル $\boldsymbol{u}$ が含まれているとする

$$A = (\dots, \boldsymbol{u}, \dots, \boldsymbol{u}, \dots)$$

この 2 つの $\boldsymbol{u}$ の列を入れ替えると、

$$\det(\dots, \boldsymbol{u}, \dots, \boldsymbol{u}, \dots) = -\det(\dots, \boldsymbol{u}, \dots, \boldsymbol{u}, \dots)$$

ところが、入れ替えの前後で行列そのものは変化していない（まったく同じ列を入れ替えても行列は同じ）ので、行列式の値も変わらないはずである

すなわち、

$$\det A = -\det A$$

が成り立つ

ここで、両辺に $\det(A)$ を足すと、

$$2\det A = 0$$

より、$\det A = 0$ が成り立つ ■

列ベクトルの線形従属性による行列式の零化　$A = (\boldsymbol{a}_1, \dots, \boldsymbol{a}_n)$ の n 個の列ベクトルが線形従属であるとすれば、

$$\det(A) = 0$$

となる

証明

列ベクトルのうち 1 つ $\boldsymbol{a}_i$ が、残りのいくつかの線型結合で表されるとすると、

$$\det(\dots, \boldsymbol{a}_i, \dots) = \det\left(\dots, \sum_{j=1}^{k} c_j \boldsymbol{a}_j, \dots\right)$$

行列式の多重線形性より、

$$\det\left(\ldots, \sum_{j=1}^{k} c_j \boldsymbol{a}_j, \ldots\right) = \sum_{j=1}^{k} c_j \det(\ldots, \boldsymbol{a}_j, \ldots)$$

ここで、$\boldsymbol{a}_j$ は $\boldsymbol{a}_i$ 以外のいずれかの列ベクトルであるため、右辺の行列式では列ベクトルの重複が生じている

よって、行列式の値は 0 になる ■

この定理の対偶をとることにより、次の定理が得られる

⚓ 非零行列式による列ベクトルの線形独立性　$A = (\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n)$ の行列式の値が 0 でないならば、A の n 個の列ベクトル $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ は線形独立である

基本変形と行列式

行列式の性質から、行列の列や行に関する基本変形と行列式の関係が見えてくる

⚓ 基本変形と行列式の関係

i. 列（行）を交換すると行列式の符号が交換される
ii. 列（行）を定数倍すると、行列式の値も定数倍される
iii. 列（行）に他の列（行）の定数倍を加えても行列式の値は変化しない

(i) は行列式の交代性、(ii) は多重線形性であり、(iii) は次の定理によって示される

⚓ 列の掃き出しに関する不変性　$i \neq j$ のとき、

$$\begin{aligned}&\det(\ldots,\boldsymbol{a}_i+c\boldsymbol{a}_j,\ldots,\boldsymbol{a}_j\ldots)\\&\qquad=\det(\ldots,\boldsymbol{a}_i,\ldots,\boldsymbol{a}_j\ldots)\end{aligned}$$

証明

行列式の多重線形性より、

$$\begin{aligned}&\det(\ldots,\boldsymbol{a}_i+c\boldsymbol{a}_j,\ldots,\boldsymbol{a}_j\ldots)\\&\qquad=\det(\ldots,\boldsymbol{a}_i,\ldots,\boldsymbol{a}_j\ldots)+c\det(\ldots,\boldsymbol{a}_j,\ldots,\boldsymbol{a}_j\ldots)\end{aligned}$$

ここで、同じ列ベクトル $\boldsymbol{a}_j$ が 2 つ含まれている行列式の値は 0 になるので、

$$\det(\ldots,\boldsymbol{a}_i+c\boldsymbol{a}_j,\ldots,\boldsymbol{a}_j\ldots)=\det(\ldots,\boldsymbol{a}_i,\ldots,\boldsymbol{a}_j\ldots)$$

だけが残る ■

行列式の特徴づけ

n 個の与えられた n 次実ベクトル $\boldsymbol{a}_1,\ldots,\boldsymbol{a}_n$ に対して、ある実数が定まるとき、これを $F(\boldsymbol{a}_1,\ldots,\boldsymbol{a}_n)$ と表すことにする

多重線形性と交代性による行列式の特徴づけ　写像 $F\colon \mathbb{R}^n\times\cdots\times\mathbb{R}^n\to\mathbb{R}$ が多重線形性と交代性を満たすならば、

$$F(\boldsymbol{a}_1,\ldots,\boldsymbol{a}_n)=F(\boldsymbol{e}_1,\ldots,\boldsymbol{e}_n)\det(\boldsymbol{a}_1,\ldots,\boldsymbol{a}_n)$$

証明

多重線形性により、

$$\begin{aligned} F(\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n) &= F\left(\sum_{i=1}^{n} a_{i_1 1}\boldsymbol{e}_{i_1}, \ldots, \sum_{i=1}^{n} a_{i_n n}\boldsymbol{e}_{i_n}\right) \\ &= \sum_{i_1, \ldots, i_n} a_{i_1 1} \cdots a_{i_n n} F(\boldsymbol{e}_{i_1}, \ldots, \boldsymbol{e}_{i_n}) \end{aligned}$$

和において、各 i_k $(1 \leq k \leq n)$ は行番号なのでそれぞれ 1 から n まで動く

ここで、交代性から導かれる定理より、$(i_1, \ldots, i_n)$ に同じ添字が 2 つ以上ある場合には $F(\boldsymbol{e}_{i_1}, \ldots, \boldsymbol{e}_{i_n}) = 0$ である

したがって、この和は $(i_1, \ldots, i_n)$ がすべて異なる場合、すなわち $(i_1, \ldots, i_n)$ が $(1, \ldots, n)$ の置換である場合にのみ寄与する

よって、$(i_1, \ldots, i_n)$ にわたる和は、実際には n 次の置換

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ i_1 & i_2 & \cdots & i_n \end{pmatrix} \in S_n$$

にわたる和であるとみなせる

この対応により、$(i_1, \ldots, i_n)$ と $\sigma \in S_n$ を同一視すると、

$$F(\boldsymbol{e}_{i_1}, \ldots, \boldsymbol{e}_{i_n}) = F(\boldsymbol{e}_{\sigma(1)}, \ldots, \boldsymbol{e}_{\sigma(n)})$$

さらに、$(\boldsymbol{e}_{\sigma(1)}, \ldots, \boldsymbol{e}_{\sigma(n)})$ を $(\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n)$ に並び替えることを考える

すなわち、σ の逆置換 σ^{-1} を考えることになる

交代性によって、1 回の互換につき (-1) 倍されるが、全体の符号は互換の回数によって定まるので、$\mathrm{sgn}(\sigma^{-1}) = \mathrm{sgn}(\sigma)$ となる

$$F(\boldsymbol{e}_{\sigma(1)}, \ldots, \boldsymbol{e}_{\sigma(n)}) = \mathrm{sgn}(\sigma) F(\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n)$$

以上より、

$$\begin{aligned}
&F(\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n) \\
&= \sum_{\sigma \in S_n} a_{\sigma(1)1} \cdots a_{\sigma(n)n} F(\boldsymbol{e}_{\sigma(1)}, \ldots, \boldsymbol{e}_{\sigma(n)}) \\
&= \sum_{\sigma \in S_n} a_{\sigma(1)1} \cdots a_{\sigma(n)n} \operatorname{sgn}(\sigma) F(\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n) \\
&= \left(\sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{\sigma(1)1} \cdots a_{\sigma(n)n} \right) F(\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n) \\
&= \det(\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n) F(\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n)
\end{aligned}$$

となり、目的の等式が示された ■

ここで、$F(\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n) = 1$ であれば、

$$F(\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n) = \det(\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n)$$

と表せることになる

この $F(\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n) = 1$ を正規化の条件といい、行列式は

i. 双線形性
ii. 交代性
iii. 正規化の条件

によって特徴づけられる

すなわち、行列式は、この 3 つの条件を満たすような

> n 個の列ベクトル $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ で定まる関数

として定義することもできる

行列式の幾何学的意味

[Todo 34:]

行列の積と行列式

行列式の特徴づけから導ける性質として、次が重要である

行列式の乗法性　$A,\ B$ を同じ型の行列とするとき、

$$\det(AB) = \det(A)\det(B)$$

証明

B の列ベクトルを $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n$ とし、次の関数

$$F(\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n) = \det(A\boldsymbol{b}_1, \ldots, A\boldsymbol{b}_n)$$

を考える

ここで、det は列ベクトルに対して交代性をもつため、この関数 F も交代性をもつ

また、det の多重線形性に加え、A による作用は線形写像であるから、F も多重線形性を満たす

よって、多重線形性と交代性による行列式の特徴づけより、

$$F(\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n) = F(\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n)\det(B)$$

一方、F の引数を単位ベクトル $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ にしたもの

$$F(\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n) = \det(A\boldsymbol{e}_1, \ldots, A\boldsymbol{e}_n)$$

を考えると、

$$\begin{aligned} F(\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n) &= \det(A\boldsymbol{e}_1, \ldots, A\boldsymbol{e}_n) \\ &= \det(\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n) \\ &= \det(A) \end{aligned}$$

よって、

$$F(\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n) = \det(A)\det(B)$$

ここで、$F(\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n)$ の定義を思い出すと、

$$\det(A\boldsymbol{b}_1, \ldots, A\boldsymbol{b}_n) = \det(A)\det(B)$$

左辺の行列 $(A\boldsymbol{b}_1, \ldots, A\boldsymbol{b}_n)$ は、行列 B の各列ベクトルに対して A を左から作用させたものであり、行列 AB を意味している

したがって、

$$\det(AB) = \det(A)\det(B)$$

が成り立つ ■

行列式の乗法性を繰り返し適用することで、次の定理が得られる

⚓ **累乗行列の行列式と行列式の累乗** $A,\ B$ を正方行列とするとき、任意の自然数 n について、

$$\det(A^n) = \det(A)^n$$

行列式と正則性

行列式は、正則性の判定にも利用できる

⚓ **正則性と行列式の非零性**

$$A \text{ が正則行列} \iff \det(A) \neq 0$$

証明

$\Longrightarrow$

A が正則であることから、

$$AA^{-1} = E$$

両辺の行列式をとって、

$$\det\left(AA^{-1}\right) = \det(E)$$

左辺には行列式の乗法性を適用し、右辺は単位行列の行列式の値が 1 であることから、

$$\det(A)\det\left(A^{-1}\right) = 1$$

もし $\det(A) = 0$ だと仮定すると、$0 = 1$ という矛盾した式になる

よって、$\det(A) \neq 0$ でなければならない ■

$\Longleftarrow$

$\det(A) \neq 0$ であることから、行列 A の列ベクトルは線型独立である

そして、A の列ベクトルが線型独立であることと、A が正則であることは同値である ■

この定理の派生として、行列式を次の形で使うことが多い

消去法の原理 A を正方行列とするとき、

$$A\boldsymbol{x} = \boldsymbol{0} \text{ に非自明解が存在する} \iff \det(A) = 0$$

余因子展開

3 次正方行列において、第 1 列を次のようにとらえる

$$\begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \end{pmatrix} = a_{11} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + a_{21} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + a_{31} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

これをふまえて、3 次行列式を、第 1 列に関する線形性を用いて、次のような和に分解してみる

$$\begin{aligned} &\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \\ &= \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & a_{32} & a_{33} \end{vmatrix} + \begin{vmatrix} 0 & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ 0 & a_{32} & a_{33} \end{vmatrix} + \begin{vmatrix} 0 & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \\ &= a_{11} \begin{vmatrix} 1 & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & a_{32} & a_{33} \end{vmatrix} + a_{21} \begin{vmatrix} 0 & a_{12} & a_{13} \\ 1 & a_{22} & a_{23} \\ 0 & a_{32} & a_{33} \end{vmatrix} + a_{31} \begin{vmatrix} 0 & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 1 & a_{32} & a_{33} \end{vmatrix} \end{aligned}$$

ここで、たとえば、

$$\begin{vmatrix} 1 & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & a_{32} & a_{33} \end{vmatrix}$$

をどのように表せるかを考える

まず、$(1, 1)$ 成分を要にして第 1 行の掃き出しを行えば、

$$\begin{vmatrix} 1 & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & a_{32} & a_{33} \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 \\ 0 & a_{22} & a_{23} \\ 0 & a_{32} & a_{33} \end{vmatrix}$$

が得られる

そこで、

$$\boldsymbol{u}_1 = \begin{pmatrix} a_{22} \\ a_{32} \end{pmatrix}, \quad \boldsymbol{u}_2 = \begin{pmatrix} a_{23} \\ a_{33} \end{pmatrix}$$

とおき、

$$F(\boldsymbol{u}_1, \boldsymbol{u}_2) = \begin{vmatrix} 1 & 0 & 0 \\ 0 & a_{22} & a_{23} \\ 0 & a_{32} & a_{33} \end{vmatrix} = F(\boldsymbol{e}_1, \boldsymbol{e}_2) \cdot \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}$$

とみなす

ここで、

$$F(\boldsymbol{e}_1, \boldsymbol{e}_2) = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = 1$$

であるから、結局、

$$\begin{vmatrix} 1 & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & a_{32} & a_{33} \end{vmatrix} = \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}$$

が得られる

2 項めの行列式も同様に、掃き出し法によって、

$$\begin{vmatrix} 0 & a_{12} & a_{13} \\ 1 & a_{22} & a_{23} \\ 0 & a_{32} & a_{33} \end{vmatrix} = \begin{vmatrix} 0 & a_{12} & a_{13} \\ 1 & 0 & 0 \\ 0 & a_{32} & a_{33} \end{vmatrix}$$

これを、

$$\boldsymbol{u}_1 = \begin{pmatrix} a_{12} \\ a_{32} \end{pmatrix}, \quad \boldsymbol{u}_2 = \begin{pmatrix} a_{13} \\ a_{33} \end{pmatrix}$$

の関数 $F(\boldsymbol{u}_1, \boldsymbol{u}_2)$ とみなす

交代性より、

$$F(\boldsymbol{e}_1, \boldsymbol{e}_2) = \begin{vmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{vmatrix} = \det(\boldsymbol{e}_2, \boldsymbol{e}_1, \boldsymbol{e}_3)$$

$$= -\det(\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3) = -1$$

なので、

$$\begin{vmatrix} 0 & a_{12} & a_{13} \\ 1 & a_{22} & a_{23} \\ 0 & a_{32} & a_{33} \end{vmatrix} = -\begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix}$$

最後の項の行列式も同様にして、

$$\begin{vmatrix} 0 & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 1 & a_{32} & a_{33} \end{vmatrix} = \begin{vmatrix} 0 & a_{12} & a_{13} \\ 0 & 0 & 0 \\ 1 & a_{32} & a_{33} \end{vmatrix} = \begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix}$$

と表せる

以上より、3 次行列式は、次のような 2 次行列式の和に分解できる

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{21}\begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + a_{31}\begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix}$$

このような行列式の展開を一般化したものが、余因子展開である

余因子　n 次正方行列 $A = (a_{ij})$ から、第 i 行と第 j 列を取り除いて $(n-1)$ 次の正方行列 Δ_{ij} を作り、その行列式に符号 $(-1)^{i+j}$ をかけたものを、A の (i, j) 余因子と呼び、$\tilde{a}_{ij}$ と書く

$$\tilde{a}_{ij} = (-1)^{i+j} \det(\Delta_{ij})$$

余因子展開　$\det(A)$ は次のように余因子展開できる

第 j 列に関する展開

$$\det(A) = \tilde{a}_{1j}a_{1j} + \tilde{a}_{2j}a_{2j} + \cdots + \tilde{a}_{nj}a_{nj}$$

第 i 行に関する展開

$$\det(A) = \tilde{a}_{i1}a_{i1} + \tilde{a}_{i2}a_{i2} + \cdots + \tilde{a}_{in}a_{in}$$

証明

列に関する展開だけを示せば、行の方は行列式の対称性よりしたがう

行列 A を $A = (\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n)$ のように列ベクトル表示する

すると、

$$\boldsymbol{a}_j = a_{1j}\boldsymbol{e}_1 + \cdots + a_{nj}\boldsymbol{e}_n$$

なので、行列式の多重線形性を用いて、

$$\begin{aligned}\det(A) &= |\boldsymbol{a}_1, \ldots, \boldsymbol{a}_j, \ldots, \boldsymbol{a}_n| \\ &= \sum_{i=1}^{n} |\boldsymbol{a}_1, \ldots, a_{ij}\boldsymbol{e}_i, \ldots, \boldsymbol{a}_n| \\ &= \sum_{i=1}^{n} a_{ij}|\boldsymbol{a}_1, \ldots, \boldsymbol{e}_i, \ldots, \boldsymbol{a}_n|\end{aligned}$$

$|\boldsymbol{a}_1, \ldots, \boldsymbol{e}_i, \ldots, \boldsymbol{a}_n|$ に対して、(i, j) 成分を要にして第 i 行を掃き出す操作を行うと、

$$\begin{vmatrix} a_{11} & \cdots & 0 & \cdots & a_{1n} \\ \vdots & \ddots & 0 & \ddots & \vdots \\ a_{i1} & \cdots & 1 & \cdots & a_{in} \\ \vdots & \ddots & 0 & \ddots & \vdots \\ a_{n1} & \cdots & 0 & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & \cdots & 0 & \cdots & a_{1n} \\ \vdots & \ddots & 0 & \ddots & \vdots \\ 0 & \cdots & 1 & \cdots & 0 \\ \vdots & \ddots & 0 & \ddots & \vdots \\ a_{n1} & \cdots & 0 & \cdots & a_{nn} \end{vmatrix}$$

さらに、i 行目を 1 つ上の行と順に交換して 1 行目まで移動し、次に j 列目を 1 つ左の列と順に交換して 1 列目まで移動する

行や列の交換から生じる符号の変化は、$(i-1)+(j-1)$ の交換を行っているので、$(-1)^{i+j-2} = (-1)^2(-1)^{i+j} = (-1)^{i+j}$ となる

よって、次のような形が得られる

$$\begin{aligned}|\boldsymbol{a}_1, \ldots, \boldsymbol{e}_i, \ldots, \boldsymbol{a}_n| &= (-1)^{i+j}\begin{vmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & a_{11} & \cdots & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & a_{n1} & \cdots & \cdots & a_{nn} \end{vmatrix} \\ &= (-1)^{i+j}\begin{vmatrix} a_{11} & \cdots & \cdots & a_{1n} \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ a_{n1} & \cdots & \cdots & a_{nn} \end{vmatrix}\end{aligned}$$

ここで現れる行列式は、第 1 行・第 1 列に移動させた第 i 行・第 j 列を取り除いた $(n-1)$ 次正方行列の行列式である

よって、符号の部分も合わせて、余因子の定義より、次のように書ける

$$|\boldsymbol{a}_1, \ldots, \boldsymbol{e}_i, \ldots, \boldsymbol{a}_n| = \tilde{a}_{ij}$$

したがって、行列 A の行列式は、

$$\det(A) = \sum_{i=1}^{n} a_{ij}\tilde{a}_{ij}$$

と書けることが示された ■

余因子行列と逆行列の公式

[Todo 35:]

クラメルの公式

[Todo 36:]

第 13 章

線形同型

線形同型

線形同型は、部分空間が「同じ」であることを述べた概念である

線形同型写像　V, W を線形空間とし、線形写像 $f : V \to W$ が全単射であるとき、f は線形同型写像あるいは単に線形同型であるという

このとき、同型を表す記号 $\cong$ を用いて、

$$f : V \xrightarrow{\cong} W$$

と書くこともある

部分空間の線形同型　V と W の間に線形同型写像が存在するとき、V と W は線形同型であるとい、

$$V \cong W$$

と書く

線形同型の性質

ここでは、線形同型写像の恒等写像、逆写像、合成写像との関係を述べる

線形同型と恒等写像

恒等写像の線形同型性　恒等写像は線形同型写像である

証明

恒等写像は明らかに全単射であり、線形写像でもあるため、線形同型写像である ■

この事実は、部分空間の線形同型に関して次のように言い換えられる

部分空間の自己同型性　部分空間 V は V 自身と線形同型である

すなわち、

$$V \cong V$$

線形同型と逆写像

線形同型写像の逆写像　線形同型写像の逆写像は線形同型写像である

証明

[Todo 37: book: 図で整理！例題で納得！線形空間入門 p93～94]

この事実は、部分空間の線形同型に関して次のように言い換えられる

線形同型性の対称性　部分空間 V が部分空間 W と線形同型なら、W は V と線形同型である

すなわち、

$$V \cong W \Longrightarrow W \cong V$$

線形同型と合成写像

線形同型写像の合成　線形同型写像の合成は線形同型写像である

証明

[Todo 38: book: 図で整理！例題で納得！線形空間入門 p94]

この事実は、部分空間の線形同型に関して次のように言い換えられる

線形同型性の推移性　部分空間 V が部分空間 W と線形同型で、W が部分空間 U と線形同型ならば、V は U と線形同型である

すなわち、

$$V \cong W \wedge W \cong U \Longrightarrow V \cong U$$

ここまでで登場した、部分空間の線形同型に関する性質をまとめると、

線形同型の同値関係としての性質

i. $V \cong V$
ii. $V \cong W \Longrightarrow W \cong V$
iii. $V \cong W \wedge W \cong U \Longrightarrow V \cong U$

となり、これらは、

同型 $\cong$ が等号 $=$ と同じ性質をもつ

ことを意味している

線形同型写像と基底

線形同型写像による基底の保存　線形同型写像 f によって、部分空間の基底は基底に写る

証明

単射な線型写像は線型独立性を保つことから、f の単射性により、基底の線型独立性が保たれる

また、f の全射性により、基底の生成性も保たれる

よって、f によって基底は基底に写る　■

座標写像

座標写像　V を線形空間とし、$\mathcal{V} = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ を V の基底とする

このとき、K^n から V への線形写像 $\Phi_{\mathcal{V}} \colon K^n \to V$ を

$$\Phi_{\mathcal{V}}(\boldsymbol{x}) = \sum_{i=1}^{n} x_i \boldsymbol{v}_i \quad (\boldsymbol{x} \in (x_i)_{i=1}^{n} \in K^n)$$

を $\mathcal{V}$ で定まる座標写像と呼ぶ

このように定めた線形写像が座標写像と呼ばれる背景は、この座標写像が線形同型であることを示し、それがどんな意味を持つのかを考えることでわかる

線形空間の基底によって定まる線形同型写像　V を線形空間とし、$\mathcal{V} = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ を V の基底とする

このとき、K^n から V への線形写像 $\Phi_{\mathcal{V}} \colon K^n \to V$ を

$$\Phi_{\mathcal{V}}(\boldsymbol{x}) = \sum_{i=1}^{n} x_i \boldsymbol{v}_i \quad (\boldsymbol{x} \in (x_i)_{i=1}^{n} \in K^n)$$

と定めると、これは線形同型写像である

証明

線形写像 $\Phi_{\mathcal{V}}$ が全単射であることを示す

単射であること

基底 $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ の線型独立性は、

$$\sum_{i=1}^{n} x_i \boldsymbol{v}_i = \mathbf{0}$$

で表される線形結合が、$x_i = 0$ を満たすことを意味する

$\Phi_{\mathcal{V}}$ の定義をふまえると、この条件は、

$$\mathrm{Ker}(\Phi_{\mathcal{V}}) = \{\mathbf{0}\}$$

と書ける

よって、線形写像の単射性と核の関係より、$\Phi_{\mathcal{V}}$ は単射である　■

全射であること

基底 $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ が V を生成することは、

$$\begin{aligned}
\boldsymbol{u} \in V &\iff \boldsymbol{u} \in \langle \boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n \rangle \\
&\iff \exists (x_i)_{i=1}^n \in K^n \ s.t. \ \boldsymbol{u} = \sum_{i=1}^n x_i \boldsymbol{v}_i \\
&\iff \exists \boldsymbol{x} \in K^n \ s.t. \ \Phi_{\mathcal{V}}(\boldsymbol{x}) = \boldsymbol{u} \\
&\iff \boldsymbol{u} \in \mathrm{Im}(\Phi_{\mathcal{V}})
\end{aligned}$$

という言い換えにより、

$$V = \mathrm{Im}(\Phi_{\mathcal{V}})$$

を意味する

よって、像空間と全射性の関係により、$\Phi_{\mathcal{V}}$ は全射である ■

この定理を部分空間の線形同型に関して言い換えると、次のような主張になる

⚓ **有限次元部分空間と数ベクトル空間の線形同型性**　任意の部分空間は次元の等しい数ベクトル空間と線形同型である

つまり、

> 和とスカラー倍だけに着目すれば、
>
> どんな部分空間も数ベクトル空間と「同じ」

ということを意味する

この同型により、部分空間に座標を与えることができる

そしてその座標によって、ベクトルの成分表示が得られる

線形代数における鳩の巣原理

線形代数における鳩の巣原理の抽象版 V, W を同じ次元の線形空間とするとき、線形写像 $f\colon V \to W$ に関して、次はすべて同値である

i. f は単射
ii. f は全射
iii. f は線形同型
iv. $\operatorname{rank}(f) = \dim V = \dim W$

証明

$\mathcal{V}, \mathcal{W}$ をそれぞれ V, W の基底として、線形写像の合成

$$g\colon \mathbb{R}^n \xrightarrow{\Phi_{\mathcal{V}}} V \xrightarrow{f} W \xrightarrow{\Phi_{\mathcal{W}}^{-1}} \mathbb{R}^n$$

を考える

このとき、g は $\mathbb{R}^n$ の線形変換である

f が単射（全射）であると仮定すると、座標写像は全単射であるので、f との合成写像 g も単射（全射）となる

逆に、g が単射（全射）であると仮定した場合について考える

f は g を用いて次のように表現でき、

$$f = \Phi_{\mathcal{W}} \circ g \circ \Phi_{\mathcal{V}}^{-1}$$

座標写像は全単射であるので、g との合成写像 f も単射（全射）となる

以上より、f が単射（全射）であることと、g が単射（全射）であることは同値である

線形変換 g に対して、線形代数における鳩の巣原理より、

$$g \text{ が単射} \iff g \text{ が全射} \iff g \text{ が全単射}$$

が成り立つが、g の単射性・全射性は f についても成り立つことがわかったので、

$$f \text{ が単射} \iff f \text{ が全射} \iff f \text{ が線形同型}$$

がいえる

最後に、階数に関する条件を示す

像空間と全射性の関係により、f が全射であることは、$\mathrm{Im}(f) = W$ と同値であるから、

$$\dim \mathrm{Im}(f) = \dim W$$

より、

$$\mathrm{rank}(f) = \dim W = \dim V$$

が得られる ■

次元による部分空間の比較

次の事実は、数の一致で空間の一致が結論できる有用な結果である

次元の一致による部分空間の一致判定 2 つの線型空間について、$V \subset W$ ならば、

$$\dim V = \dim W \Longrightarrow V = W$$

証明

$\boldsymbol{v} \in V$ をそのまま W の元と考えることで得られる写像を $\iota\colon V \to W$ とする（包含写像）

この包含写像は、V の元 $\boldsymbol{v}$ を W の中にそのまま「埋め込む」操作を表しているため、$\iota(\boldsymbol{v})$ は $\boldsymbol{v}$ 自身である

$$\iota(\boldsymbol{v}) = \boldsymbol{v}$$

特に、$\iota(\boldsymbol{v}) = \mathbf{0}$ は $\boldsymbol{v} = \mathbf{0}$ そのものを意味する

$$\iota(\boldsymbol{v}) = \mathbf{0} \Longleftrightarrow \boldsymbol{v} = \mathbf{0}$$

したがって、零ベクトルへの写像による単射性の判定より、ι は単射である

また、ι が単射であることと、仮定 $\dim V = \dim W$ を合わせると、線形代数における鳩の巣原理の抽象版より、ι は全射であることがわかる

よって、全射の定義より、すべての $\boldsymbol{w} \in W$ に対して $\iota(\boldsymbol{v}) = \boldsymbol{w}$ となる $\boldsymbol{v}$ が存在する

すなわち、W の元はすべて V の元であり、$V \subset W$ もふまえると、これは $V = W$ を意味する ■

次元による部分空間の比較　K^n の部分空間 V, W について、$V \subseteq W$ ならば、

$$\dim V \leq \dim W$$

が成り立つ

等号が成立するのは、$V = W$ のときに限る

証明

$V \subseteq W$ であることから、基底の延長により、V の基底を延長して W の基底にできるので、

$$\dim V \leq \dim W$$

が成り立つ

等号が成立する場合については、前述の次元の一致による部分空間の一致判定を参照

■

核空間・像空間の次元

線形写像の単射性と核の次元　線形写像 $f\colon V \to W$ について、

$$f \text{ が単射} \iff \dim \operatorname{Ker}(f) = 0$$

証明

線形写像の単射性と核の関係より、f が単射であることは次と同値である

$$\operatorname{Ker}(f) = \{\mathbf{0}\}$$

次元の定義より、$\{\mathbf{0}\}$ の次元は 0 であるので、

$$\dim \operatorname{Ker}(f) = 0$$

が成り立つ　■

線形写像の全射性と像の次元　線形写像 $f\colon V \to W$ について、

$$f \text{ が全射} \iff \dim \operatorname{Im}(f) = \dim W$$

証明

線形代数における鳩の巣原理の抽象版の主張そのものである　■

第 14 章

表現行列と基底変換

基底に関する座標ベクトル

V を線形空間とし、$\mathcal{V}=\{\boldsymbol{v}_1,\boldsymbol{v}_2,\ldots,\boldsymbol{v}_n\}$ をその基底とする

V の任意のベクトル $\boldsymbol{v}$ は、

$$\boldsymbol{v}=\sum_{i=1}^{n}x_i\boldsymbol{v}_i$$

と一意的に書ける

ここで、$\Phi_{\mathcal{V}}$ を座標写像とすると、その定義から、

$$\Phi_{\mathcal{V}}^{-1}(\boldsymbol{v})=\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}\in\mathbb{R}^n$$

このベクトルを $\mathcal{V}$ に関する $\boldsymbol{v}$ の座標ベクトルあるいは成分表示と呼び、

$$\boldsymbol{v}=\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}_{\mathcal{V}}$$

と書くことにする

一般の基底に関する表現行列

V, W をそれぞれ次元が n, m の線形空間とし、f を V から W への線形写像とする

また、$\mathcal{V}, \mathcal{W}$ をそれぞれ V, W の基底とする

座標写像が線形同型写像であることは、任意の部分空間が数ベクトル空間と同型であることを意味していた

よって、V から W への線形写像 f は、数ベクトル空間との線形同型写像（座標写像）$\Phi_{\mathcal{V}}$, $\Phi_{\mathcal{W}}$ を合成すれば、

$$f' = \Phi_{\mathcal{W}}^{-1} \circ f \circ \Phi_{\mathcal{V}} : \mathbb{R}^n \to \mathbb{R}^m$$

として、数ベクトル空間の間の写像と考えることができる

この合成を図で整理して、次のように表す

$$\begin{tikzcd}
V \arrow[r, "f"] & W \\
\mathbb{R}^n \arrow[u, "\Phi_{\mathcal{V}}"] \arrow[r, "f'"] & \mathbb{R}^m \arrow[u, "\Phi_{\mathcal{W}}"]
\end{tikzcd}$$

このような図を**図式**という

下辺の矢印は、合成写像

$$\Phi_{\mathcal{W}}^{-1} \circ f \circ \Phi_{\mathcal{V}} : \mathbb{R}^n \to \mathbb{R}^m$$

を表していて、この写像は $\mathbb{R}^n$ から $\mathbb{R}^m$ への線形写像である

左下の $\mathbb{R}^n$ から右上の W への 2 通りの合成写像が一致するという意味で、この図式は**可換**であるという

数ベクトル空間の間の写像は、行列が定める線形写像であることから、この写像 f は $m \times n$ 型行列 A により表現される

$$\begin{tikzcd}
V \arrow[r, "f"] & W \\
\mathbb{R}^n \arrow[u, "\Phi_{\mathcal{V}}"] \arrow[r, "A"] & \mathbb{R}^m \arrow[u, "\Phi_{\mathcal{W}}"]
\end{tikzcd}$$

座標ベクトルの記法を用いると、写像 f は次で与えられる

$$f \colon \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}_{\mathcal{V}} \mapsto \left(A \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \right)_{\mathcal{W}}$$

このように、座標写像を用いることで、V から W への線形写像 f から、$m \times n$ 型行列が得られる

この行列 A を、基底 $\mathcal{V}, \mathcal{W}$ に関する f の表現行列という

つまり、

> 基底 $\mathcal{V}, \mathcal{W}$ を固定して考えるときは、f を A と同一視できる

ということになり、このとき、

> 表現行列は線形写像の「成分表示」

と解釈できる

表現行列の構成

数ベクトル空間の間の線形写像を定める行列は、各基本ベクトル $\boldsymbol{e}_j$ の f による像

$$f(\boldsymbol{e}_j) = \boldsymbol{a}_j = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix} \quad (1 \leq j \leq n)$$

を用いて、

$$(f(\boldsymbol{e}_1), \ldots, f(\boldsymbol{e}_n)) = (\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n) = A$$

のように構成された

この表現行列の構成を、部分空間 V, W の基底をそれぞれ $\mathcal{V}=\{\boldsymbol{v}_1,\ldots,\boldsymbol{v}_n\}$, $\mathcal{W}=\{\boldsymbol{w}_1,\ldots,\boldsymbol{w}_m\}$ として一般化する

このとき、$\boldsymbol{a}_j$ は座標写像 $\Phi_{\mathcal{W}}$ によって、

$$\Phi_{\mathcal{W}}(\boldsymbol{a}_j)=\sum_{i=1}^{m}a_{ij}\boldsymbol{w}_i \quad (1\leq j\leq n)$$

のように W に写される

また、$\boldsymbol{e}_j$ は座標写像 $\Phi_{\mathcal{V}}$ によって、

$$\Phi_{\mathcal{V}}(\boldsymbol{e}_j)=\sum_{i=1}^{n}e_{ij}\boldsymbol{v}_i \quad (1\leq j\leq n)$$

のように V に写されるが、これは $\boldsymbol{v}_j$ そのものである

たとえば、$j=1$ のときは、

$$\Phi_{\mathcal{V}}(\boldsymbol{e}_1)=\sum_{i=1}^{n}e_{i1}\boldsymbol{v}_i=\boldsymbol{v}_1$$

となる

よって、$\boldsymbol{e}_j\mapsto\boldsymbol{a}_j$ という写像は、

$$\boldsymbol{v}_j\mapsto\Phi_{\mathcal{W}}(\boldsymbol{a}_j)$$

という V から W への写像 f に対応する

（この対応は、可換図式からも明らか）

記号を書き換えると、

$$f(\boldsymbol{v}_j)=\Phi_{\mathcal{W}}(\boldsymbol{a}_j)=\sum_{i=1}^{m}a_{ij}\boldsymbol{w}_i$$

となり、右辺はさらに、

$$\sum_{i=1}^{m}a_{ij}\boldsymbol{w}_i=(\boldsymbol{w}_1,\ldots,\boldsymbol{w}_m)\begin{pmatrix}a_{1j}\\ \vdots\\ a_{mj}\end{pmatrix}$$

と変形できるので、まとめると、

$$(f(\boldsymbol{v}_1),\ldots,f(\boldsymbol{v}_n))=(\boldsymbol{w}_1,\ldots,\boldsymbol{w}_m)\,A$$

と表せる

線形変換の表現行列

V を n 次元の線形空間とし、f を V の線形変換、すなわち V から V 自身への線形写像とする

V の基底 $\mathcal{V}$ を選ぶとき、次の可換図式によって n 次正方行列 A が定められる

$$\begin{tikzcd} V \arrow[r, "f"] & V \\ \mathbb{R}^n \arrow[u, "\Phi_{\mathcal{V}}"] \arrow[r, "A\cdot"] & \mathbb{R}^n \arrow[u, "\Phi_{\mathcal{V}}"'] \end{tikzcd}$$

写像の定義される空間と、写す先の空間が同じなので、どちらに対しても同じ基底を用いることができる

もちろん、考える問題によっては別な基底を用いても構わないが、線形変換に対しては 1 つの基底を用いるのが自然である

数ベクトル空間の基底変換行列

$V = \mathbb{R}^n$ とし、標準基底 $\mathcal{E}$ によって行列 A で表現される線形変換を f とする

別な基底 $\mathcal{V}$ によって f を表現する行列を B とするとき、B をどうやって計算すればよいかを考える

B を定める原理は、表現行列の構成で議論したように、

$$(f(\boldsymbol{v}_1), \ldots, f(\boldsymbol{v}_n)) = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)B$$

ここで、$\boldsymbol{v}_i$ や $f(\boldsymbol{v}_i)$ は $\mathbb{R}^n$ の元なので、$(f(\boldsymbol{v}_1), \ldots, f(\boldsymbol{v}_n))$ や $(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$ は n 次の正方行列であるとみなせる

そこで、

$$P = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$$

とおくとき、次に示すように P は正則行列である

基底変換行列の正則性　基底の変換行列は正則行列である

証明

P の列ベクトルは基底であるため、線形独立である

列ベクトルの線型独立性による正則の判定で示したように、正則行列であることは、列ベクトルが線形独立であることと同値である ■

また、B を決める式

$$(f(\boldsymbol{v}_1), \ldots, f(\boldsymbol{v}_n)) = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)B$$

の左辺は、次のように書ける

$$\begin{aligned}(f(\boldsymbol{v}_1), \ldots, f(\boldsymbol{v}_n)) &= (A\boldsymbol{v}_1, \ldots, A\boldsymbol{v}_n) \\ &= A(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n) \\ &= AP\end{aligned}$$

よって、B を決める式は、

$$AP = PB$$

となり、P は正則である（逆行列が存在する）ので、両辺に左から P^{-1} をかけて、

$$B = P^{-1}AP$$

が得られる

行列 P は、標準基底 $\mathcal{E}$ から基底 $\mathcal{V}$ への基底変換行列と呼ばれる

線形空間の基底変換行列

V を線形空間とし、V の基底 $\mathcal{V} = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ を別な基底 $\mathcal{V}' = \{\boldsymbol{v}'_1, \ldots, \boldsymbol{v}'_n\}$ に取り替えることを考える

このとき、$\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ が V の基底であることから、V の元である $\boldsymbol{v}'_1, \ldots, \boldsymbol{v}'_n$ は、$\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ の線形結合で表される

そこで、

$$\boldsymbol{v}'_i = p_{1i}\boldsymbol{v}_1 + p_{2i}\boldsymbol{v}_2 + \cdots + p_{ni}\boldsymbol{v}_n$$

すなわち、

$$(\boldsymbol{v}'_1, \ldots, \boldsymbol{v}'_n) = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)(p_{ij})$$

とおく

このとき、写像 $f\colon V \to V$ を

$$\begin{cases} f(\boldsymbol{v}_1) & = & \boldsymbol{v}'_1 \\ f(\boldsymbol{v}_2) & = & \boldsymbol{v}'_2 \\ \vdots & & \vdots \\ f(\boldsymbol{v}_n) & = & \boldsymbol{v}'_n \end{cases}$$

を満たすものとして定義する

これはすなわち、基底 $\mathcal{V}$ を構成するそれぞれのベクトルを、基底 $\mathcal{V}'$ を構成するベクトルに順に写す線形変換であり、

$$(f(\boldsymbol{v}_1), \ldots, f(\boldsymbol{v}_n)) = (\boldsymbol{v}'_1, \ldots, \boldsymbol{v}'_n)$$

を満たすものである

すると、行列 $P = (p_{ij})$ を定める式は、

$$(f(\boldsymbol{v}_1), \ldots, f(\boldsymbol{v}_n)) = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)P$$

と書ける

よって、P は基底変換 $\mathcal{V} \rightharpoonup \mathcal{V}'$ を表す線形写像 f の表現行列である

この意味で、P を基底変換 $\mathcal{V} \rightharpoonup \mathcal{V}'$ の基底変換行列と呼ぶ

⚓ **線形空間の基底変換行列**　V を線形空間とし、$\mathcal{V} = \{\boldsymbol{v}_i\}_{i=1}^n$, $\mathcal{V}' = \{\boldsymbol{v}'_i\}_{i=1}^n$ を V の基底とするとき、基底変換 $\mathcal{V} \rightharpoonup \mathcal{V}'$ の変換行列 P は、

$$(\boldsymbol{v}'_1, \ldots, \boldsymbol{v}'_n) = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)P$$

により定まる

この行列 P は、座標写像を介して考えると、次の可換図式で定まるものである

$$
\begin{tikzcd}
\mathbb{R}^n \arrow[dd, "P\cdot"'] \arrow[dr, "\Phi_{\mathcal{V}'}"] & \\
& V \\
\mathbb{R}^n \arrow[ur, "\Phi_{\mathcal{V}}"] &
\end{tikzcd}
$$

つまり、P は、

$$\Phi_{\mathcal{V}}^{-1} \circ \Phi_{\mathcal{V}'} : \mathbb{R}^n \to \mathbb{R}^n$$

の標準基底に関する表現行列である

一方、この行列 P はベクトルの成分表示の変換に用いることもできる

座標ベクトルの変換則　基底変換 $\mathcal{V} \rightharpoonup \mathcal{V}'$ の変換行列を P とし、ベクトル $\boldsymbol{a} \in V$ の $\mathcal{V}$, $\mathcal{V}'$ に関する座標ベクトルをそれぞれ $\boldsymbol{x}$, $\boldsymbol{x}'$ とするとき、

$$\boldsymbol{x} = P\boldsymbol{x}'$$

が成り立つ

証明

ベクトル $\boldsymbol{a}$ の 2 種類の基底 $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ と $\boldsymbol{v}'_1, \ldots, \boldsymbol{v}'_n$ に関する成分

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad \begin{pmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_n \end{pmatrix}$$

を考えると、$\boldsymbol{a}$ の 2 通りの表現

$$\begin{aligned} \boldsymbol{a} &= x_1\boldsymbol{v}_1 + x_2\boldsymbol{v}_2 + \cdots + x_n\boldsymbol{v}_n \\ &= (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n) \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \end{aligned}$$

$$\begin{aligned}
\boldsymbol{a} &= x_1'\boldsymbol{v}_1' + x_2'\boldsymbol{v}_2' + \cdots + x_n'\boldsymbol{v}_n' \\
&= (\boldsymbol{v}_1', \ldots, \boldsymbol{v}_n') \begin{pmatrix} x_1' \\ x_2' \\ \vdots \\ x_n' \end{pmatrix} \\
&= (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n) P \begin{pmatrix} x_1' \\ x_2' \\ \vdots \\ x_n' \end{pmatrix}
\end{aligned}$$

が得られる

どちらも $(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$ との積の形、すなわち $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ の線形結合として表されている

ここで、基底 $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ の線型独立性から、その線形結合は一意的であるので、係数比較ができて、

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = P \begin{pmatrix} x_1' \\ x_2' \\ \vdots \\ x_n' \end{pmatrix}$$

が成り立つ ■

基底変換による表現行列の変化

$f \colon V \to W$ を線形写像とする

$$V \xrightarrow{f} W$$

V の基底 $\mathcal{V}$ と W の基底 $\mathcal{W}$ に関する f の表現行列を A とする

$$\begin{tikzcd}
V \arrow[r, "f"] & W \\
\mathbb{R}^n \arrow[u, "\Phi_{\mathcal{V}}"] \arrow[r, "A\cdot"] & \mathbb{R}^m \arrow[u, "\Phi_{\mathcal{W}}"']
\end{tikzcd}$$

また、V の基底 $\mathcal{V}'$ と W の基底 $\mathcal{W}'$ に基底を変えるとき、f の表現行列を B とする

$$
\begin{tikzcd}
\mathbb{R}^n \arrow[r, "B\cdot"] \arrow[d, "\Phi_{\mathcal{V}'}"] & \mathbb{R}^m \arrow[d, "\Phi_{\mathcal{W}'}"] \\
V \arrow[r, "f"] & W
\end{tikzcd}
$$

基底変換 $\mathcal{V} \rightharpoonup \mathcal{V}'$ の変換行列を P、$\mathcal{W} \rightharpoonup \mathcal{W}'$ の変換行列を Q とするとき、次の可換図式で整理できる

$$
\begin{tikzcd}
\mathbb{R}^n \arrow[rrr, "B\cdot"] \arrow[dd, "P\cdot"] \arrow[dr, "\Phi_{\mathcal{V}'}"] & & & \mathbb{R}^m \arrow[dd, "Q\cdot"] \arrow[dl, "\Phi_{\mathcal{W}'}"'] \\
& V \arrow[r, "f"] & W & \\
\mathbb{R}^n \arrow[rrr, "A\cdot"] \arrow[ur, "\Phi_{\mathcal{V}}"] & & & \mathbb{R}^m \arrow[ul, "\Phi_{\mathcal{W}}"']
\end{tikzcd}
$$

ここで、行列 A によって表現される写像を F_A、他の行列についても同様に表すと、

$$
\begin{tikzcd}
\mathbb{R}^n \arrow[rrr, "F_B"] \arrow[dd, "F_P"] \arrow[dr, "\Phi_{\mathcal{V}'}"] & & & \mathbb{R}^m \arrow[dd, "F_Q"] \arrow[dl, "\Phi_{\mathcal{W}'}"'] \\
& V \arrow[r, "f"] & W & \\
\mathbb{R}^n \arrow[rrr, "F_A"] \arrow[ur, "\Phi_{\mathcal{V}}"] & & & \mathbb{R}^m \arrow[ul, "\Phi_{\mathcal{W}}"']
\end{tikzcd}
$$

このとき、左上の $\mathbb{R}^n$ から右下の $\mathbb{R}^m$ への写像は、

$$F_A \circ F_P, \quad F_Q \circ F_B$$

という 2 通りの表現ができる

すなわち、

$$F_A \circ F_P = F_Q \circ F_B$$

合成写像は行列の積に対応するので、

$$AP = QB$$

ここで、Q は $\mathbb{R}^m \to \mathbb{R}^m$ の標準基底に関する表現行列であるから、正則行列である

そこで、左から Q^{-1} をかけて、

$$B = Q^{-1}AP$$

が得られる

基底変換に伴う表現行列の変換　線形写像 $f\colon V \to W$ の基底 $\mathcal{V}, \mathcal{W}$ に関する表現行列を A とし、同じ線形写像 f の別な基底 $\mathcal{V}', \mathcal{W}'$ に関する表現行列を B とするとき、基底変換 $\mathcal{V} \rightharpoonup \mathcal{V}'$ の変換行列を P、$\mathcal{W} \rightharpoonup \mathcal{W}'$ の変換行列を Q とすると、

$$B = Q^{-1}AP$$

が成り立つ

実用上は $V = W$ である場合が特に重要で、この場合には $P = Q$ とすることができるので、

$$B = P^{-1}AP$$

が成り立つ

基底変換に伴う表現行列の変換（線形変換の場合）　線形変換 $f\colon V \to V$ の基底 $\mathcal{V}$ に関する表現行列を A とし、同じ線形変換 f の別な基底 $\mathcal{V}'$ に関する表現行列を B とするとき、基底変換 $\mathcal{V} \rightharpoonup \mathcal{V}'$ の変換行列を P とすると、

$$B = P^{-1}AP$$

が成り立つ

相似変換

「ある種の操作を行ったら同一のものになるもの」を「互いに相似」と呼ぶ

たとえば、二つ以上の図形が「相似」であるとは、平行移動、回転、反転、拡大縮小などの操作を行うとそれら図形をぴったり重ねることができるという意味だった

行列に対する「相似」は、次のように定める

行列の相似　正方行列 A, B に対して、正則行列 P が存在して、

$$B = P^{-1}AP$$

が成り立つとき、A と B は相似であるという

このような変換を相似変換と呼ぶ

A と B が相似であるとき、A と B は 1 つの線形変換 f を異なる基底によって表現して得られた行列であるという関係にある

線形写像の階数標準形

線形写像に対して、うまく基底を選ぶと、表現行列を階数標準形にできる

線形写像の階数標準形　線形写像 $f\colon V \to W$ に対し、$r = \mathrm{rank}(f)$ とするとき、V, W のある基底に関する f の表現行列が次の形になる

$$\begin{pmatrix} E_r & O \\ O & O \end{pmatrix}$$

証明

V の基底を次のように分けて構成する

i. $\mathrm{Ker}(f)$ を張るベクトル（これは f によって零に写る）
ii. $\mathrm{Ker}(f)$ に属さないが、f によって像を生成するベクトル（これは f によって非零に写る）

V, W の次元をそれぞれ n, m とすると、線形写像の次元定理より、$\mathrm{Ker}(f)$ の次元は $n - r$ である

そこで、$\mathrm{Ker}(f) \subset V$ の基底を $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_{n-r}$ とする

さらに、$\boldsymbol{v}_1, \ldots, \boldsymbol{v}_r \in V$ を、$\boldsymbol{v}_1, \ldots, \boldsymbol{v}_r, \boldsymbol{u}_1, \ldots, \boldsymbol{u}_{n-r}$ が V の基底になるように選ぶ（基底の延長）

このとき、

$$\boldsymbol{w}_i = f(\boldsymbol{v}_i) \quad (i = 1, \ldots, r)$$

とおくと、$\boldsymbol{w}_1, \ldots, \boldsymbol{w}_r$ は線形独立である

実際、線形関係式

$$\sum_{i=1}^{r} c_i \boldsymbol{w}_i = \boldsymbol{0}$$

があるとすると、f は線形写像なので、

$$\sum_{i=1}^{r} c_i \boldsymbol{w}_i = \sum_{i=1}^{r} c_i f(\boldsymbol{v}_i) = f\left(\sum_{i=1}^{r} c_i \boldsymbol{v}_i\right) = \boldsymbol{0}$$

より、

$$\left(\sum_{i=1}^{r} c_i \boldsymbol{v}_i\right) \in \mathrm{Ker}(f)$$

この線形結合で表されるベクトルを $\boldsymbol{v}$ とする

$$\boldsymbol{v} = \sum_{i=1}^{r} c_i \boldsymbol{v}_i$$

すると、$\boldsymbol{v} \in \mathrm{Ker}(f)$ より、$\boldsymbol{v}$ は $\mathrm{Ker}(f)$ の基底 $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_{n-r}$ の線形結合でも表すことができる

$$\boldsymbol{v} = \sum_{j=1}^{n-r} d_j \boldsymbol{u}_j$$

したがって、$\boldsymbol{v}$ の 2 通りの表現から、次の等式が成り立つ

$$\sum_{i=1}^{r} c_i \boldsymbol{v}_i = \sum_{j=1}^{n-r} d_j \boldsymbol{u}_j$$

ここで、$\boldsymbol{v}_1, \ldots, \boldsymbol{v}_r, \boldsymbol{u}_1, \ldots, \boldsymbol{u}_{n-r}$ は V の基底なので、線型独立である

よって、等式

$$\sum_{i=1}^{r} c_i \boldsymbol{v}_i - \sum_{j=1}^{n-r} d_j \boldsymbol{u}_j = \boldsymbol{0}$$

が成り立つには、各係数が 0 でなければならない

$$\begin{aligned} c_i &= 0 \quad (i = 1, \ldots, r) \\ d_j &= 0 \quad (j = 1, \ldots, n-r) \end{aligned}$$

したがって、$\boldsymbol{w}_1, \ldots, \boldsymbol{w}_r$ は線形独立である

$\boldsymbol{w}_1, \ldots, \boldsymbol{w}_r$ はすべて $\mathrm{Im}(f)$ に属するので、これは $\mathrm{Im}(f) \subset W$ の基底となる

そこで、この基底を延長して、$\boldsymbol{w}_1, \ldots, \boldsymbol{w}_r, \boldsymbol{w}_{r+1}, \ldots, \boldsymbol{w}_m$ を W の基底とする

このように構成した V と W の基底に関する線形写像 f の表現行列を考える

$\mathrm{Ker}(f)$ を張るベクトル $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_{n-r}$ は f によって零に写ることと、$\boldsymbol{w}_1, \ldots, \boldsymbol{w}_r$ の定義より、

$$\begin{cases} f(\boldsymbol{v}_i) = \boldsymbol{w}_i & (i = 1, \ldots, r) \\ f(\boldsymbol{u}_j) = \boldsymbol{0} & (j = 1, \ldots, n-r) \end{cases}$$

よって、基底 $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_r, \boldsymbol{u}_1, \ldots, \boldsymbol{u}_{n-r}\}$ における f の表現行列は、

$$\begin{aligned} &(f(\boldsymbol{v}_1), \ldots, f(\boldsymbol{v}_r), f(\boldsymbol{u}_1), \ldots, f(\boldsymbol{u}_{n-r})) \\ &= (\boldsymbol{w}_1, \ldots, \boldsymbol{w}_r, \boldsymbol{0}, \ldots, \boldsymbol{0}) \\ &= (\boldsymbol{w}_1, \ldots, \boldsymbol{w}_r) \begin{pmatrix} E_r & O \\ O & O \end{pmatrix} \end{aligned}$$

として定まる ■

このように、線形空間 V, W の任意の基底変換を許すと、線形写像 f の表現行列をとても単純な形

$$\begin{pmatrix} E_r & O \\ O & O \end{pmatrix}$$

にできる

これを f の階数標準形という

基底変換に伴う表現行列の変換の原理を用いて、先ほどの定理を行列の言葉で表すことができる

線形写像 $f \colon V \to W$ の基底 $\mathcal{V}, \mathcal{W}$ に関する表現行列を A とし、同じ線形写像 f の別な基底 $\mathcal{V}', \mathcal{W}'$ に関する表現行列を階数標準形を B とする

このとき、正則行列による階数標準形の構成より、それぞれの基底変換行列 P, Q は行変形、列変形に対応する正則行列である

表現行列の階数標準形　$m \times n$ 型行列 A に対し、それぞれ n, m 次の正則行列 P, Q が存在して、

$$Q^{-1}AP = \begin{pmatrix} E_r & O \\ O & O \end{pmatrix}$$

となる

ここで、$r = \mathrm{rank}(A)$ である

第 15 章

直和分解と不変部分空間

部分空間の共通部分

与えられた部分空間から、新しく部分空間を作ることができる

線形部分空間の共通部分は部分空間　U, W を体 K 上の V の部分空間とするとき、共通部分 $U \cap W$ は V の部分空間である

証明

和について

$\boldsymbol{a}, \boldsymbol{b} \in U \cap W$ とすると、共通部分の定義より、$\boldsymbol{a}$ と $\boldsymbol{b}$ はどちらも U と W の両方に属していることになる

つまり、$\boldsymbol{a}, \boldsymbol{b} \in U$ かつ $\boldsymbol{a}, \boldsymbol{b} \in W$ である

U も W も部分空間なので、部分空間の定義より、

$$\boldsymbol{a} + \boldsymbol{b} \in U$$
$$\boldsymbol{a} + \boldsymbol{b} \in W$$

$\boldsymbol{a} + \boldsymbol{b}$ が U と W の両方に属していることから、$\boldsymbol{a} + \boldsymbol{b}$ は $U \cap W$ に属する

よって、$U \cap W$ は和について閉じている ■

スカラー倍について

$\boldsymbol{a} \in U \cap W$ と $c \in K$ をとる

共通部分の定義より、$\boldsymbol{a}$ は U と W の両方に属しているので、部分空間の定義より

$$c\boldsymbol{a} \in U$$
$$c\boldsymbol{a} \in W$$

よって、$c\boldsymbol{a}$ は $U \cap W$ に属するため、$U \cap W$ はスカラー倍について閉じている ■

部分空間の和

⚓ **線形部分空間の和は部分空間** U, W を体 K 上の V の部分空間とするとき、和空間

$$U + W := \{\boldsymbol{u} + \boldsymbol{w} \mid \boldsymbol{u} \in U, \boldsymbol{w} \in W\}$$

は V の部分空間である

証明

和について

$\boldsymbol{a}_1, \boldsymbol{a}_2 \in U$, $\boldsymbol{b}_1, \boldsymbol{b}_2 \in W$ とする

U と W は部分空間なので、部分空間の定義より

$$\boldsymbol{a}_1 + \boldsymbol{a}_2 \in U, \quad \boldsymbol{b}_1 + \boldsymbol{b}_2 \in W$$

一方、和空間の定義より、$\boldsymbol{a}_1 + \boldsymbol{b}_1$, $\boldsymbol{a}_2 + \boldsymbol{b}_2$ はそれぞれ $U + W$ の元である

これらの元の和をとったときに、その和も $U+W$ に属していれば、和空間は和について閉じているといえる

$$\begin{aligned}(\boldsymbol{a}_1+\boldsymbol{b}_1)+(\boldsymbol{a}_2+\boldsymbol{b}_2) &= (\boldsymbol{a}_1+\boldsymbol{a}_2)+(\boldsymbol{b}_1+\boldsymbol{b}_2) \\ &\in U+W\end{aligned}$$

上式で、和空間は和について閉じていることが示された ■

スカラー倍について

$\boldsymbol{a} \in U,\ \boldsymbol{b} \in W$ と $c \in K$ をとる

U と W は部分空間なので、部分空間の定義より

$$\begin{aligned} c\boldsymbol{a} &\in U \\ c\boldsymbol{b} &\in W \end{aligned}$$

一方、和空間の定義より、$\boldsymbol{a}+\boldsymbol{b}$ は $U+W$ の元である

この元をスカラー倍したときに、そのスカラー倍も $U+W$ に属していれば、和空間はスカラー倍について閉じているといえる

$$\begin{aligned} c(\boldsymbol{a}+\boldsymbol{b}) &= c\boldsymbol{a}+c\boldsymbol{b} \\ &\in U+W \end{aligned}$$

上式で、和空間はスカラー倍について閉じていることが示された ■

3 つ以上の部分空間の和も同様に考えて、一般に和空間は次のように定義される

和空間　線形空間 V と、その部分空間 $V_1, \ldots, V_k$ が与えられたときに、

$$\boldsymbol{v} = \boldsymbol{v}_1 + \cdots + \boldsymbol{v}_k \quad (\boldsymbol{v}_i \in V_i, i = 1, \ldots, k)$$

と表されるベクトル $\boldsymbol{v}$ 全体がなす集合を $V_1, \ldots, V_k$ の和空間といい、

$$\sum_{i=1}^{k} V_i$$

と書く

和空間を張るベクトル

部分空間を生成するベクトルを用いて、部分空間の和を表せる

部分空間の和と生成ベクトル　K^n の 2 つの部分空間 $U = \langle \boldsymbol{u}_1, \ldots, \boldsymbol{u}_m \rangle$ と $W = \langle \boldsymbol{w}_1, \ldots, \boldsymbol{w}_k \rangle$ に対して、和空間 $U + W$ は

$$U + W = \langle \boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_m, \boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_k \rangle$$

となる

証明

和空間 $U + W$ は

$$U + W = \{\boldsymbol{x} \in K^n \mid \boldsymbol{x} = \boldsymbol{u} + \boldsymbol{w},\ \boldsymbol{u} \in U,\ \boldsymbol{w} \in W\}$$

と定義される

また、$\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m, \boldsymbol{w}_1, \ldots, \boldsymbol{w}_k$ の張る部分空間は

$$H = \langle \boldsymbol{u}_1, \ldots, \boldsymbol{u}_m, \boldsymbol{w}_1, \ldots, \boldsymbol{w}_k \rangle$$

である

これらが等しいことを示せばよい

$U + W \subseteq H$

任意の $\boldsymbol{x} \in U + W$ に対し、$\boldsymbol{x} = \boldsymbol{u} + \boldsymbol{w}$（$\boldsymbol{u} \in U,\ \boldsymbol{w} \in W$）と書ける

すなわち、

$$\begin{aligned} \boldsymbol{u} &= a_1\boldsymbol{u}_1 + \cdots + a_m\boldsymbol{u}_m && (a_i \in K) \\ \boldsymbol{w} &= b_1\boldsymbol{w}_1 + \cdots + b_k\boldsymbol{w}_k && (b_j \in K) \end{aligned}$$

よって、

$$\boldsymbol{x} = \sum_{i=1}^{m} a_i\boldsymbol{u}_i + \sum_{j=1}^{k} b_j\boldsymbol{w}_j \in H$$

$H \subseteq U + W$

任意の $\boldsymbol{x} \in H$ は

$$\boldsymbol{x} = \sum_{i=1}^{m} a_i \boldsymbol{u}_i + \sum_{j=1}^{k} b_j \boldsymbol{w}_j$$

と書ける

ここで

$$\boldsymbol{u} = \sum_{i=1}^{m} a_i \boldsymbol{u}_i \in U$$

$$\boldsymbol{w} = \sum_{j=1}^{k} b_j \boldsymbol{w}_j \in W$$

とすれば、

$$\boldsymbol{x} = \boldsymbol{u} + \boldsymbol{w} \in U + W$$

以上より、$U + W \subseteq H$ と $H \subseteq U + W$ が成り立つので、$U + W = H$ が示された ■

和空間の包含関係

⚓ **和空間における部分空間の和集合の包含**　和空間は、和集合を部分集合として包含する

すなわち、U, W を V の部分空間とするとき、

$$U + W \supset U \cup W$$

が成り立つ

証明

部分空間はいずれも零ベクトルを含むので、たとえば、$U = \{\mathbf{0}\}$ の場合、

$$U + W \supset W$$

同様に、

$$U + W \supset U$$

よって、$U + W$ は U または W を包含することがわかる

すなわち、

$$U + W \supset U \cup W$$

が成り立つ ■

和空間の最小包含性 U, W を V の部分空間とする

和空間 $U + W$ は、U と W を含む部分空間のうち、最小のものである

証明

V の任意の部分空間のうち、U と W の両方を包含するもの V' を考える

このとき、部分空間は和に閉じているため、V' は $U + W$ も包含する

$$V' \supset U + W$$

よって、V' の任意性から、$U + W$ は U と W を含む部分空間のうち、最小のものとなる ■

このように、和空間 $U + W$ は、U や W を部分空間として含むが、U や W より真に大きい（U, W を真部分集合として含む）とは限らない

別の角度からいうと、

$$V = W_1 + W_2$$

という関係があるだけで、「V が W_1 と W_2 の和に分解された」というのは適当ではない

和空間が持つこの欠陥を補うために、和空間の概念をより精密化したものが、次に述べる直和である

直和分解

直和分解　線形空間 V の部分集合 W_1, W_2 に対して、任意の $\boldsymbol{v} \in V$ が $\boldsymbol{w}_1 \in W_1, \boldsymbol{w}_2 \in W_2$ によって

$$\boldsymbol{v} = \boldsymbol{w}_1 + \boldsymbol{w}_2$$

と一意的に表されるとき、V は W_1 と W_2 の直和である（直和に分解される）といい、

$$V = W_1 \oplus W_2$$

と書く

この定義は、次のように言い換えることができる

直和分解の同値条件　線形空間 V の部分集合 W_1, W_2 に対して $V = W_1 \oplus W_2$ が成り立つことと、

i. $V = W_1 + W_2$
ii. $W_1 \cap W_2 = \{\mathbf{0}\}$

の両方が成り立つことは同値である

証明

(i), (ii) $\Longrightarrow V = W_1 \oplus W_2$

$\boldsymbol{w}_1,\ \boldsymbol{w}'_1 \in W_1,\ \boldsymbol{w}_2,\ \boldsymbol{w}'_2 \in W_2$ とする

仮定 (i) と和空間の定義より、

$$\boldsymbol{v} = \boldsymbol{w}_1 + \boldsymbol{w}_2 = \boldsymbol{w}'_1 + \boldsymbol{w}'_2$$

この等式は、移項によって次のように変形できる

$$\boldsymbol{w}_1 - \boldsymbol{w}'_1 = \boldsymbol{w}'_2 - \boldsymbol{w}_2$$

部分空間は和に閉じているため、左辺は W_1 に、右辺は W_2 に属する

よって、このベクトルは $W_1 \cap W_2$ に属する

仮定 (ii) より、$W_1 \cap W_2$ の元は零ベクトルであるので、

$$\begin{aligned} \boldsymbol{w}_1 - \boldsymbol{w}'_1 &= \boldsymbol{0} \\ \boldsymbol{w}'_2 - \boldsymbol{w}_2 &= \boldsymbol{0} \end{aligned}$$

したがって、

$$\boldsymbol{w}_1 = \boldsymbol{w}'_1, \quad \boldsymbol{w}_2 = \boldsymbol{w}'_2$$

となり、$\boldsymbol{v}$ の表現の一意性が示された ■

$V = W_1 \oplus W_2 \Longrightarrow$ (i), (ii)

和空間の定義をふまえると、(i) は直和分解の定義に含まれる

(ii) を示すため、$\boldsymbol{v} \in W_1 \cap W_2$ とする

$\boldsymbol{v}$ は零ベクトルを用いて、

$$\boldsymbol{v} = \boldsymbol{v} + \boldsymbol{0} = \boldsymbol{0} + \boldsymbol{v}$$

と表せるが、直和分解の定義より、$\boldsymbol{v}$ の表現は一意的であるので、

$$\boldsymbol{v} = \boldsymbol{0}$$

を得る

よって、$W_1 \cap W_2 = \{\boldsymbol{0}\}$ が成り立つ ■

直和分解の一意性を表す条件

3 つ以上の部分空間による直和を考えるにあたって、直和分解の定義に含まれていた「一意性」を表す条件を定式化する

部分空間の和における表現の一意性　和空間 $\sum_{i=1}^{k} V_i$ の元 $\boldsymbol{v}$ を、部分空間 $V_1, \ldots, V_k$ の元 $\boldsymbol{v}_i \in V_i$ の和として

$$\boldsymbol{v} = \sum_{i=1}^{k} \boldsymbol{v}_i \quad (\boldsymbol{v}_i \in V_i)$$

と書くとする

このとき、次の条件が成り立てば、和に使われる $\boldsymbol{v}_i$ は $\boldsymbol{v}$ により一意的に定まる

$$\boldsymbol{v}_1 + \cdots + \boldsymbol{v}_k = \boldsymbol{0} \Longrightarrow \boldsymbol{v}_1 = \cdots = \boldsymbol{v}_k = \boldsymbol{0}$$

証明

仮に、$\boldsymbol{v}$ が 2 通りの和で表せるとする

$$\boldsymbol{v} = \sum_{i=1}^{k} \boldsymbol{v}_i = \sum_{i=1}^{k} \boldsymbol{v}'_i \quad (\boldsymbol{v}_i, \boldsymbol{v}'_i \in V_i)$$

このとき、

$$\sum_{i=1}^{k} (\boldsymbol{v}_i - \boldsymbol{v}'_i) = \boldsymbol{0}$$

となるが、ここで $\boldsymbol{v}_i - \boldsymbol{v}'_i$ は V_i に属する

そこで、$\boldsymbol{w}_i = \boldsymbol{v}_i - \boldsymbol{v}'_i \in V_i$ とおき、

$$\boldsymbol{w}_1 + \cdots + \boldsymbol{w}_k = \boldsymbol{0} \Longrightarrow \boldsymbol{w}_1 = \cdots = \boldsymbol{w}_k = \boldsymbol{0}$$

という条件を満たすとすると、$\boldsymbol{w}_i = \boldsymbol{0}$ より、

$$\boldsymbol{v}_i = \boldsymbol{v}'_i \quad (i = 1, \ldots, k)$$

が導かれる

したがって、$\boldsymbol{v}$ の和に使われる $\boldsymbol{v}_i$ は一意的に定まる　■

この一意性の条件を用いて、複数の部分空間による直和を次のように定義する

直和　線形空間 V と、その部分空間 $V_1, \ldots, V_k$ が与えられたとき、$\boldsymbol{v}_i \in V_i$, $\boldsymbol{v} \in \sum_{i=1}^{k} V_i$ に対して、

$$\boldsymbol{v}_1 + \cdots + \boldsymbol{v}_k = \mathbf{0} \implies \boldsymbol{v}_1 = \cdots = \boldsymbol{v}_k = \mathbf{0}$$

が成り立つとき、$\displaystyle\sum_{i=1}^{k} V_i$ は V の直和であるといい、

$$\bigoplus_{i=1}^{k} V_i$$

と書く

和空間と直和の次元

和空間の次元　K^n の部分空間 V, W に対して、

$$\dim(V + W) = \dim V + \dim W - \dim(V \cap W)$$

が成り立つ

証明

$\dim(V) = n$, $\dim(W) = m$ とする

$V \cap W$ の基底 $\mathcal{V} = \{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_d\}$ をとる

これを基底の延長の定理に基づいて、V の基底

$$\mathcal{V} \cup \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_{n-d}\}$$

に延長する

同様に、$\mathcal{V}$ を W の基底

$$\mathcal{V} \cup \{\boldsymbol{w}_1, \dots, \boldsymbol{w}_{m-d}\}$$

に延長する

このとき、$\boldsymbol{u}_1, \dots, \boldsymbol{u}_d, \boldsymbol{v}_1, \dots, \boldsymbol{v}_{n-d}, \boldsymbol{w}_1, \dots, \boldsymbol{w}_{m-d}$ が $V + W$ の基底になることを示す

$V + W$ を生成すること

$\boldsymbol{v} \in V,\ \boldsymbol{w} \in W$ とすると、それぞれ基底の線形結合で表すことができる

$$\boldsymbol{v} = \sum_{i=1}^{d} a_i \boldsymbol{u}_i + \sum_{j=1}^{n-d} b_j \boldsymbol{v}_j$$
$$\boldsymbol{w} = \sum_{i=1}^{d} c_i \boldsymbol{u}_i + \sum_{k=1}^{m-d} d_k \boldsymbol{w}_k$$

$V + W$ の任意の元は、$\boldsymbol{v} + \boldsymbol{w}$ と書けるので、

$$\boldsymbol{v} + \boldsymbol{w} = \sum_{i=1}^{d} (a_i + c_i) \boldsymbol{u}_i + \sum_{j=1}^{n-d} b_j \boldsymbol{v}_j + \sum_{k=1}^{m-d} d_k \boldsymbol{w}_k$$

となり、$\{\boldsymbol{u}_1, \dots, \boldsymbol{u}_d, \boldsymbol{v}_1, \dots, \boldsymbol{v}_{n-d}, \boldsymbol{w}_1, \dots, \boldsymbol{w}_{m-d}\}$ の線形結合で表せる ■

線型独立であること

$\boldsymbol{u}_1, \dots, \boldsymbol{u}_d, \boldsymbol{v}_1, \dots, \boldsymbol{v}_{n-d}, \boldsymbol{w}_1, \dots, \boldsymbol{w}_{m-d}$ が線型独立であることを示すために、次のような線形関係式を考える

$$\sum_{i=1}^{d} c_i \boldsymbol{u}_i + \sum_{j=1}^{n-d} c_{d+j} \boldsymbol{v}_j + \sum_{k=1}^{m-d} c_{d+n-d+k} \boldsymbol{w}_k = \boldsymbol{0}$$

ここで、$c_i \in K$ はスカラーである

この式を V と W の基底の線型結合として考えると、V の基底 $\boldsymbol{u}_i,\ \boldsymbol{v}_j$ に関する部分と W の基底 $\boldsymbol{u}_i,\ \boldsymbol{w}_k$ に関する部分がそれぞれ線形独立であるため、結局どの項においても $c_i = 0$ である必要がある

よって、$\boldsymbol{u}_1, \ldots, \boldsymbol{u}_d, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_{n-d}, \boldsymbol{w}_1, \ldots, \boldsymbol{w}_{m-d}$ は線型独立である

■

以上より、$\boldsymbol{u}_1, \ldots, \boldsymbol{u}_d, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_{n-d}, \boldsymbol{w}_1, \ldots, \boldsymbol{w}_{m-d}$ は $V + W$ の基底であることが示された

この基底をなすベクトルの個数（次元）について考えると、

$$\begin{aligned} \dim(V + W) &= d + (n - d) + (m - d) \\ &= n + m - d \end{aligned}$$

となる

ここで、$d = \dim(V \cap W)$ なので、

$$\dim(V + W) = \dim V + \dim W - \dim(V \cap W)$$

と書き換えられ、目的の式が得られた ■

直和分解に対して和空間の次元定理を適用すると、次のようにまとめられる

直和の次元　W_1, W_2 が V の部分空間であるとき、

$$\dim(W_1 \oplus W_2) = \dim W_1 + \dim W_2$$

が成り立つ

証明

直和分解の定義より $W_1 \cap W_2 = \{\mathbf{0}\}$ であるので、

$$\dim(W_1 \cap W_2) = 0$$

よって、和空間の次元の式から、$\dim(W_1 \cap W_2)$ が消えた形になる ■

また、和空間の次元定理の証明過程を、直和分解の場合で考えることで、次の定理が得られる

直和の基底　線形空間 V が部分空間 W_1, W_2 の直和に分解されることと、V の基底が W_1, W_2 の基底を合わせたものになることは同値である

証明

直和分解の場合、$W_1 \cap W_2 = \{\mathbf{0}\}$ であるため、和空間の次元定理の証明過程において、$W_1 \cap W_2$ の基底を考える必要がなくなる

よって、和空間の次元定理の証明と同様にして、W_1 の基底と W_2 の基底を合わせたものが V の基底になることが示される　■

不変部分空間

V 上の線形変換 f について、「変換 f で写しても変わらない」という性質を考える

f 不変　f を線形空間 V の線形変換とする

V の部分空間 W に対して、

$$f(W) \subset W$$

すなわち、

$$\forall \boldsymbol{w} \in W \Longrightarrow f(\boldsymbol{w}) \in W$$

が成り立つとき、W は f 不変な部分空間であるという

また、$V = \mathbb{R}^n$ で、f が正方行列 A によって定まっているときは、f 不変な部分空間 W を A 不変な部分空間ともいう

写像の制限と不変部分空間

不変部分空間による線形変換のブロック型行列表現　V を n 次元線形空間とし、線形変換 $f\colon V \to V$ を考える

このとき、V のある部分空間 W が f 不変ならば、V の適当な基底について、f は

$$\begin{pmatrix} * & * \\ O & * \end{pmatrix} \text{または} \begin{pmatrix} * & O \\ * & * \end{pmatrix}$$

という形の行列で表すことができる

証明

$\dim(W) = r$ とし、W の基底 $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_r$ を延長して V の基底 $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_r, \boldsymbol{v}_{r+1}, \ldots, \boldsymbol{v}_n$ をとる

このとき、表現行列の構成法より、

$$f(\boldsymbol{v}_j) = \sum_{i=1}^{r} a_{ij}\boldsymbol{v}_i + \sum_{i=r+1}^{n} a_{ij}\boldsymbol{v}_i \quad (1 \leq j \leq n)$$

とおける

ここで、W は f 不変であることは、$1 \leq j \leq r$ の範囲では $f(\boldsymbol{v}_j) \in W$ であることを意味する

W の元 $f(\boldsymbol{v}_j)$ は、W の基底だけを用いて表現できるので、

$$f(\boldsymbol{v}_j) = \sum_{i=1}^{r} a_{ij}\boldsymbol{v}_i \quad (1 \leq j \leq r)$$

すなわち、もともとの $f(\boldsymbol{v}_j)$ の式において、

$$\sum_{i=r+1}^{n} a_{ij}\boldsymbol{v}_i = \boldsymbol{0} \quad (1 \leq j \leq r)$$

となっている

$\boldsymbol{v}_i$ は基底なので線型独立であり、したがって、

$$a_{ij} = 0 \quad (1 \leq j \leq r,\ r+1 \leq i \leq n)$$

が成り立つ

この条件より、f の表現行列 (a_{ij}) は、

$$\begin{array}{c|c} \overset{r}{\longleftrightarrow} & \overset{n-r}{\longleftrightarrow} \end{array}$$

$$\left(\begin{array}{c|c} * & * \\ \hline O & * \end{array}\right)\begin{array}{l} \updownarrow r \\ \updownarrow n-r \end{array}$$

というような形になる

また、V の基底として、順序を変えた $\boldsymbol{v}_{r+1}, \ldots, \boldsymbol{v}_n, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_r$ を取ることもできる

この場合は、

$$f(\boldsymbol{v}_j) = \sum_{i=1}^{r} a_{ij}\boldsymbol{v}_i + \sum_{i=r+1}^{n} a_{ij}\boldsymbol{v}_i \quad (r+1 \leq j \leq n)$$

とおくと、$r+1 \leq j \leq n$ の範囲（V の基底の後半部分）で

$$\sum_{i=1}^{r} a_{ij}\boldsymbol{v}_i = \boldsymbol{0}$$

となるので、すなわち、

$$a_{ij} = 0 \quad (r+1 \leq j \leq n,\ 1 \leq i \leq r)$$

よって、f の表現行列 (a_{ij}) は、

$$\begin{array}{c|c} \overset{n-r}{\longleftrightarrow} & \overset{r}{\longleftrightarrow} \end{array}$$

$$\left(\begin{array}{c|c} * & O \\ \hline * & * \end{array}\right)\begin{array}{l} \updownarrow n-r \\ \updownarrow r \end{array}$$

という形になる

以上より、2 通りの f の表現行列の形が得られた ■

V の基底を $\boldsymbol{v}_1, \dots, \boldsymbol{v}_r, \boldsymbol{v}_{r+1}, \dots, \boldsymbol{v}_n$ ととった場合、$f(\boldsymbol{v}_j) \in W$ は

$$f(\boldsymbol{v}_j) = \sum_{i=1}^{r} a_{ij}\boldsymbol{v}_i \quad (1 \leq j \leq r)$$

だけで表現できた

この $1 \leq i \leq r,\ 1 \leq j \leq r$ の部分は、f の表現行列

$$A = (a_{ij}) = \begin{array}{cc} \begin{array}{cc} r & n-r \end{array} & \\ \left(\begin{array}{c|c} A_{11} & A_{12} \\ \hline O & A_{22} \end{array}\right) & \begin{array}{c} r \\ n-r \end{array} \end{array}$$

の、A_{11} の部分に対応する

つまり、この行列 A_{11} は、$\boldsymbol{v}_1, \dots, \boldsymbol{v}_r$ で張られる V の部分空間 W から W への線形写像 f' を、基底 $\boldsymbol{v}_1, \dots, \boldsymbol{v}_r$ について表現する行列になっている

f' は、f の定義域を W に制限したものになっているが、W の元に限定して考える限り、実質的には f と区別がないものである

この意味で、写像 f' を、写像 f の W への制限と呼び、$f|_W$ と表記する

写像の制限　写像 $f\colon X \to Y$ において、X のある部分集合 S が与えられたとき、定義域を S に限定したものを f の S に対する制限といい、

$$f|_S\colon S \to Y$$

と表す

同様に、V の基底を $\boldsymbol{v}_{r+1}, \dots, \boldsymbol{v}_n, \boldsymbol{v}_1, \dots, \boldsymbol{v}_r$ ととった場合、$f(\boldsymbol{v}_j) \in W$ は

$$\sum_{i=r+1}^{n} a_{ij}\boldsymbol{v}_i \quad (r+1 \leq j \leq n)$$

だけで表現できた

この $r+1 \leq i \leq n,\ r+1 \leq j \leq n$ の部分は、f の表現行列

$$A = (a_{ij}) = \begin{array}{cc} \begin{array}{cc} \quad n-r \quad & \quad r \quad \end{array} & \\ \left(\begin{array}{c|c} A_{11} & O \\ \hline A_{21} & A_{22} \end{array}\right) & \begin{array}{c} n-r \\ r \end{array} \end{array}$$

の、A_{22} の部分に対応する

つまり、この場合は、A_{22} が変換 f の W への制限 $f|_W$ を表現する行列になっている

不変部分空間への直和分解

不変部分空間による線形変換のブロック型行列表現の証明では、W の基底 $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_r$ を延長したものを V の基底 $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_r, \boldsymbol{v}_{r+1}, \ldots, \boldsymbol{v}_n$ とした

このとき、

$$W' = \langle \boldsymbol{v}_{r+1}, \ldots, \boldsymbol{v}_n \rangle$$

とおくと、V の基底が $W,\ W'$ の基底を合わせたものになっているため、直和の基底に関する定理より、

$$V = W \oplus W'$$

となる

ここで、もし W' も f 不変であれば、右上の A_{12} も零行列になって、表現行列は

$$A = (a_{ij}) = \begin{array}{cc} \begin{array}{cc} \quad r \quad & \quad n-r \quad \end{array} & \\ \left(\begin{array}{c|c} A_{11} & O \\ \hline O & A_{22} \end{array}\right) & \begin{array}{c} r \\ n-r \end{array} \end{array}$$

というブロック対角型になる

⚓ **不変部分空間への直和分解**　線形空間 V と、V 上の線形変換 f に対し、V が f 不変な部分空間 W_1 と W_2 の直和に分解することができれば、すなわち、

i. $V = W_1 \oplus W_2$

ii. W_1, W_2 は f 不変な V の部分空間

となる W_1, W_2 が存在すれば、適当な V の基底について、f は次のような形の行列で表せる

$$\begin{array}{cc} \begin{array}{cc} \dim(W_1) & \dim(W_2) \end{array} & \\ \begin{pmatrix} * & O \\ O & * \end{pmatrix} & \begin{array}{c} \dim(W_1) \\ \dim(W_2) \end{array} \end{array}$$

または

$$\begin{array}{cc} \begin{array}{cc} \dim(W_2) & \dim(W_1) \end{array} & \\ \begin{pmatrix} * & O \\ O & * \end{pmatrix} & \begin{array}{c} \dim(W_2) \\ \dim(W_1) \end{array} \end{array}$$

証明

W_1 の基底、W_2 の基底をこの順に並べるか、その反対の順に並べて、V の基底を構成することで、不変部分空間による線形変換のブロック型行列表現の証明と同様に示される ■

さらに、V をより細かい部分空間の直和に分解できる場合には、次のようになる

複数の不変部分空間への直和分解　線形空間 V と、V 上の線形変換 f について、

i. $V = W_1 \oplus W_2 \oplus \cdots \oplus W_k$

ii. 各部分空間 W_i は f 不変な V の部分空間

であるならば、適当な V の基底に対し、f は次のような形の行列で表せる

$$\begin{pmatrix} \boxed{*} & & & O \\ & \boxed{*} & & \\ & & \ddots & \\ O & & & \boxed{*} \end{pmatrix}$$

対角線上の各正方形の大きさは、各部分空間 W_i の次元に対応する

そこで、以上の議論を究極にまで押し進めると、次の定理になる

⚓ **一次元部分空間への直和分解**　n 次元部分空間 V が、n 個の f 不変の 1 次元部分空間の直和に分解できるとき、すなわち、

i. $V = W_1 \oplus W_2 \oplus \cdots \oplus W_n$

ii. $W_i \ (i = 1, 2, \ldots, n)$ は f 不変な 1 次元部分空間

となるときは、f は次のような対角行列で表せる

$$\begin{pmatrix} * & & & O \\ & * & & \\ & & \ddots & \\ O & & & * \end{pmatrix}$$

一次元不変部分空間

W を一次元部分空間とすると、これは基底 $\boldsymbol{w} \neq \boldsymbol{0}$ で張られる空間であるので、

$$W = \langle \boldsymbol{w} \rangle = \{\alpha \boldsymbol{w} \mid \alpha \in K\}$$

この一次元部分空間 W が f 不変であるとは、定義より、

$$\forall \boldsymbol{w} \in W \Longrightarrow f(\boldsymbol{w}) \in W$$

であり、これで W の元は $\alpha\boldsymbol{w}$ とも $f(\boldsymbol{w})$ とも表せることになるので、

$$f(\boldsymbol{w}) = \alpha\boldsymbol{w} \quad (\alpha \in K)$$

がいえる

以上をふまえて、一次元部分空間への直和分解という定理をより具体的に整理してみる

$\dim(V) = n$ とすると、V 上の線形変換 f の表現行列 A を構成する式は、

$$(f(\boldsymbol{w}_1), \ldots, f(\boldsymbol{w}_n)) = (\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n)A$$

となるが、ここで、

$$(f(\boldsymbol{w}_1), \ldots, f(\boldsymbol{w}_n)) = (\alpha_1\boldsymbol{w}_1, \ldots, \alpha_n\boldsymbol{w}_n)$$

であるので、

$$(f(\boldsymbol{w}_1), \ldots, f(\boldsymbol{w}_n)) = (\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n)\begin{pmatrix} \alpha_1 & 0 & \cdots & 0 \\ 0 & \alpha_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \alpha_n \end{pmatrix}$$

と書き換えられる

したがって、f は基底 $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n$ について、次の対角行列で表される

$$\begin{pmatrix} \alpha_1 & 0 & \cdots & 0 \\ 0 & \alpha_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \alpha_n \end{pmatrix}$$

ここで現れたスカラー α_i やベクトル $\boldsymbol{w}_i$ と、線形写像 f との関係が、固有値・固有ベクトルと行列の対角化という話題に発展する

第 16 章

行列の対角化

固有値と固有ベクトル

与えられた線形写像を表現する行列を単純化（対角化）する上で、一次元不変部分空間への直和分解が本質的な役割を果たす

一次元の f 不変部分空間 W の基底 $\boldsymbol{a}$ とは、

ある $\lambda \in K$ について $f(\boldsymbol{a}) = \lambda\boldsymbol{a}$

となるような $\mathbf{0}$ 以外のベクトルだった

固有値と固有ベクトル　体 K 上の線形空間 V 上の線形変換 $f\colon V \to V$ に対して、

$$f(\boldsymbol{a}) = \lambda\boldsymbol{a} \quad (\boldsymbol{a} \neq \mathbf{0})$$

となるベクトル $\boldsymbol{a} \in V$ が存在するとき、このようなスカラー $\lambda \in K$ を、線形変換 f の固有値という

また、このようなベクトル $\boldsymbol{a}$ を、f の固有値 λ に属する固有ベクトルという

線形変換 f の表現行列を A とすると、これは正方行列であり、$f(\boldsymbol{a}) = A\boldsymbol{a}$ と表せる

よって、固有値と固有ベクトルの定義は、次のようにも書ける

行列の固有値と固有ベクトル　正方行列 A に対して、

$$A\boldsymbol{a} = \lambda \boldsymbol{a} \quad (\boldsymbol{a} \neq \mathbf{0})$$

となるベクトル $\boldsymbol{a}$ とスカラー λ が存在するとき、このようなスカラー λ を行列 A の**固有値**という

また、このようなベクトル $\boldsymbol{a}$ を、行列 A の固有値 λ に属する**固有ベクトル**という

異なる固有値に属する固有ベクトル

異なる固有値に属する固有ベクトルの非一致性　異なる固有値 $\alpha_i,\ \alpha_j\ (\alpha_i \neq \alpha_j)$ に属する固有ベクトル $\boldsymbol{p}_i, \boldsymbol{p}_j$ は異なるベクトルである

証明

固有値と固有ベクトルの定義より、

$$\begin{cases} A\boldsymbol{p}_i &= \alpha_i \boldsymbol{p}_i \\ A\boldsymbol{p}_j &= \alpha_j \boldsymbol{p}_j \end{cases}$$

である

もし $\boldsymbol{p}_i = \boldsymbol{p}_j$ ならば、

$$\begin{aligned} \alpha_i \boldsymbol{p}_i &= \alpha_j \boldsymbol{p}_i \\ \therefore \quad (\alpha_i - \alpha_j)\boldsymbol{p}_i &= \mathbf{0} \end{aligned}$$

となるが、$\boldsymbol{p}_i$ は固有ベクトルであり $\mathbf{0}$ ではないので、$\alpha_i - \alpha_j = 0$ となる

すなわち、

$$\alpha_i = \alpha_j$$

が成立し、これは $\alpha_i \neq \alpha_j$ に反する

よって、$\boldsymbol{p}_i \neq \boldsymbol{p}_j$ でなければならない　■

この定理を発展させて、次のことがいえる

異なる固有値に属する固有ベクトルの線型独立性　$\alpha_1, \alpha_2, \ldots, \alpha_k$ が行列 A の相異なる固有値であるとすると、それぞれに属する固有ベクトル $\boldsymbol{p}_1, \boldsymbol{p}_2, \ldots, \boldsymbol{p}_k$ は線型独立である

証明

固有値の個数 k についての数学的帰納法によって証明する

$k=1$ のとき、$\boldsymbol{p}_1$ は固有ベクトルゆえ $\boldsymbol{0}$ ではないので、$\{\boldsymbol{p}_1\}$ は線型独立である

$k \geq 2$ として、$(k-1)$ 個以下の固有ベクトルについて定理の主張が成り立つと仮定する

このとき、線形関係式

$$c_1\boldsymbol{p}_1 + c_2\boldsymbol{p}_2 + \cdots + c_k\boldsymbol{p}_k = \boldsymbol{0}$$

を考える

両辺に A をかけると、$A\boldsymbol{p}_i = \alpha_i\boldsymbol{p}_i$ より、

$$c_1\alpha_1\boldsymbol{p}_1 + c_2\alpha_2\boldsymbol{p}_2 + \cdots + c_k\alpha_k\boldsymbol{p}_k = \boldsymbol{0}$$

この等式から、初めの線形関係式の α_k 倍を引いて

$$c_1(\alpha_1 - \alpha_k)\boldsymbol{p}_1 + \cdots + c_{k-1}(\alpha_{k-1} - \alpha_k)\boldsymbol{p}_{k-1} = \boldsymbol{0}$$

ここで、帰納法の仮定より、$\boldsymbol{p}_1, \boldsymbol{p}_2, \ldots, \boldsymbol{p}_{k-1}$ は線型独立であるため、係数はすべて 0 でなければならない

$$c_1(\alpha_1 - \alpha_k) = 0, \quad \ldots, \quad c_{k-1}(\alpha_{k-1} - \alpha_k) = 0$$

さらに、$\alpha_1, \alpha_2, \ldots, \alpha_k$ は相異なる固有値であるため、$\alpha_i - \alpha_k \neq 0$ $(i = 1, \ldots, k-1)$ である

よって、

$$c_1 = 0, \quad \ldots, \quad c_{k-1} = 0$$

が成り立つ

この結果を初めの線形関係式に代入すると、

$$c_k \boldsymbol{p}_k = \boldsymbol{0}$$

が残るが、$\boldsymbol{p}_k$ は固有ベクトルであり $\boldsymbol{0}$ ではないため、$c_k = 0$ も成り立つ

以上より、$\boldsymbol{p}_1, \boldsymbol{p}_2, \ldots, \boldsymbol{p}_k$ は線型独立である ■

固有ベクトルによる行列の対角化

一次元不変部分空間に関する議論で見たように、

$$f(\boldsymbol{a}_i) = \lambda_i \boldsymbol{a}_i \quad (i = 1, \ldots, n)$$

となるような $\boldsymbol{a}_i$ を基底として用いると、線形変換 f は次のような対角行列で表現できた

$$\begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

そして、このような $\boldsymbol{a}_i$ を固有ベクトル、λ_i を固有値として定義したため、

> 行列 A の固有ベクトルからなる基底が存在すれば、
>
> A は対角化できる

と言い換えられる

線形変換 f の表現行列を A とすると、A は正方行列である

たとえば基底を A の固有ベクトルに変換した際に、この線形変換 f の表現行列が $P^{-1}AP$ に変化するとして、この行列 $P^{-1}AP$ が対角行列となる場合が、A が対角化できるということである

対角化可能　与えられた正方行列 A が適当な正則行列 P により

$$P^{-1}AP = \begin{pmatrix} \lambda_1 & & & O \\ & \lambda_2 & & \\ & & \ddots & \\ O & & & \lambda_n \end{pmatrix}$$

と変形できるとき、A は対角化可能であるという

行列 A の固有ベクトルからなる基底が存在すれば、A は対角化可能である、ということを定式化しよう

対角化可能性と固有ベクトルの線型独立性　n 次元正方行列 A が対角化可能であるための必要十分条件は、線型独立な n 個の A の固有ベクトルが存在することである

証明

線型独立な A の固有ベクトルが存在 $\Longrightarrow$ A は対角化可能

A の固有ベクトルを $\boldsymbol{a}_i$ $(i = 1, \dots, n)$、それに対応する固有値を α_i とすると、固有値と固有ベクトルの定義より、次式が成り立つ

$$f(\boldsymbol{a}_i) = \alpha_i \boldsymbol{a}_i \quad (i = 1, \dots, n)$$

仮定より $\boldsymbol{a}_i$ は線型独立であり、一次元部分空間 $\{c\boldsymbol{a}_i \mid c \in K\}$ は $\boldsymbol{a}_i$ によって張られる空間である

よって、$\boldsymbol{a}_i$ を基底として用いることができるので、一次元不変部分空間に関する議論で見たように、A は対角行列で表現できる　■

A は対角化可能 $\Longrightarrow$ 線型独立な A の固有ベクトルが存在

A が対角化可能であることから、ある正則行列 P が存在して、

$$P^{-1}AP = \begin{pmatrix} \alpha_1 & & & O \\ & \alpha_2 & & \\ & & \ddots & \\ O & & & \alpha_n \end{pmatrix}$$

が成り立つので、両辺に P をかけて、

$$AP = P\begin{pmatrix} \alpha_1 & & & O \\ & \alpha_2 & & \\ & & \ddots & \\ O & & & \alpha_n \end{pmatrix}$$

が成り立つ

ここで、P を n 個の列ベクトル $\boldsymbol{p}_1, \boldsymbol{p}_2, \ldots, \boldsymbol{p}_n$ を横に並べたもの、すなわち、

$$P = (\boldsymbol{p}_1, \boldsymbol{p}_2, \ldots, \boldsymbol{p}_n)$$

とみなせば、上の等式は、

$$\begin{cases} A\boldsymbol{p}_1 = \alpha_1\boldsymbol{p}_1 \\ A\boldsymbol{p}_2 = \alpha_2\boldsymbol{p}_2 \\ \qquad \vdots \\ A\boldsymbol{p}_n = \alpha_n\boldsymbol{p}_n \end{cases}$$

という関係を意味する

これはすなわち、$\boldsymbol{p}_1, \boldsymbol{p}_2, \ldots, \boldsymbol{p}_n$ がそれぞれの固有値 $\alpha_1, \alpha_2, \ldots, \alpha_n$ に属する A の固有ベクトルであることを意味する

さらに、P は正則であるため、その列ベクトル $\boldsymbol{p}_1, \boldsymbol{p}_2, \ldots, \boldsymbol{p}_n$ は線型独立である ■

この定理と、異なる固有値に属する固有ベクトルの線型独立性から、次の定理が得られる

⚓ **固有値の相異性と対角化可能性**　n 次正方行列 A が異なる n 個の固有値 $\alpha_1, \ldots, \alpha_n$ をもつならば、A は対角化可能である

すなわち、ある n 次正則行列 P によって、

$$P^{-1}AP = \begin{pmatrix} \alpha_1 & & & O \\ & \alpha_2 & & \\ & & \ddots & \\ O & & & \alpha_n \end{pmatrix}$$

が成り立つ

証明

n 個の異なる固有値 $\alpha_1, \ldots, \alpha_n$ に属する固有ベクトル $\boldsymbol{p}_1, \ldots, \boldsymbol{p}_n$ は線型独立である

よって、固有ベクトルの線型独立性より、対角化可能性が導かれる ■

ただし、この定理の逆は成立しない

つまり、n 次正方行列 A が n 個の異なる固有値を持たなくても、対角化できることがある

実際、A がすでに対角行列になっているなら、最も単純な場合として $A = E$ をとると、A の固有値は 1 だけであるが、任意の正則行列 P に対して $P^{-1}EP$ は対角行列 E になる

よって、対角化のために本質的なのは、n 個の異なる固有値ではなく、

n 個の線型独立な固有ベクトル

であるといえる

特性多項式と特性方程式

λ が n 次正方行列 A の固有値であることは、

$$A\boldsymbol{x} = \lambda\boldsymbol{x} \quad (\boldsymbol{x} \neq \boldsymbol{0})$$

となるような $\boldsymbol{x} \in K^n$ が存在することである

ここで、$A\boldsymbol{x} = \lambda\boldsymbol{x}$ を次のように変形することができる

$$\begin{aligned} A\boldsymbol{x} - \lambda\boldsymbol{x} &= \boldsymbol{0} \\ A\boldsymbol{x} - \lambda E\boldsymbol{x} &= \boldsymbol{0} \\ (A - \lambda E)\boldsymbol{x} &= \boldsymbol{0} \end{aligned}$$

$\boldsymbol{x} \neq \boldsymbol{0}$ という条件により、$(A - \lambda E)\boldsymbol{x} = \boldsymbol{0}$ は非自明な解を持つ必要がある

固有ベクトルの斉次形方程式による定義　固有値 λ の固有ベクトルとは、斉次形方程式

$$(A - \lambda E)\boldsymbol{x} = \boldsymbol{0}$$

の非自明な解のことである

固有値を求める上で重要となるこの定理は、行列式を使って言い換えることができる

固有値の方程式による定義　行列 A の固有値 λ は、x についての n 次方程式

$$\det(A - xE) = 0$$

の K に含まれる解である

証明

λ が A の固有値であることは、斉次形方程式 $(A - \lambda E)\boldsymbol{x} = \boldsymbol{0}$ が非自明解を持つことと言い換えられる

そして、斉次形方程式が非自明解を持つことは、行列式が 0 になることと同値である

すなわち、

$$\det(A - \lambda E) = 0$$

が成り立ち、つまり $x = \lambda$ は方程式 $\det(A - xE) = 0$ の解である　■

$A = (a_{ij})$ とおいて、

$$\det(A - xE) = \begin{vmatrix} a_{11} - x & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - x & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - x \end{vmatrix}$$

を展開すると、x についての n 次式になる

特に、すべての列（あるいはすべての行）から、x を含む成分をとった場合の積は、

$$(a_{11} - x)(a_{22} - x) \cdots (a_{nn} - x)$$

であるので、これを展開して現れる項を中心に考察する

n 次の項

$(a_{11} - x)(a_{22} - x) \cdots (a_{nn} - x)$ の各因子から、$-x$ だけを選んでかけ合わせたものが

$$(-1)^n x^n$$

であり、これが最高次の項となる

$n - 1$ 次の項

$(a_{11} - x)(a_{22} - x) \cdots (a_{nn} - x)$ のうち、1 つだけ a_{ii} を選び、残りの因子からは $-x$ を選んでかけ合わせたものが

$$(-1)^{n-1}(a_{11} + a_{22} + \cdots + a_{nn})x^{n-1}$$

である

これは、トレースの定義より、

$$(-1)^{n-1} \operatorname{tr}(A) x^{n-1}$$

とも書き換えられる

$n - 2$ 次以下の項

行列式では、各列から 1 つずつ、行に重複がないように成分を選ぶ必要がある

そして、今取り上げている行列式では x を含む成分が対角線上にあるので、$n - 1$ 次の場合は、対角成分以外を選ぶことができなかった（対角成分以外から x でない数 a_{ij} を得ようとすると、同じ行もしくは列から 2 つ成分を選ぶことになってしまう）

しかし、$n-2$ 次以下の項では、x を含まない成分を 2 個以上選ぶことができるので、対角成分以外からも成分を選ぶことができる
そのため、$n-2$ 次以下の項は、上の展開式以外からも現れることになり、単純に計算はできない

定数項

定数項は、多項式において $x=0$ とおくことで得られるので、$\det(A-xE)$ に $x=0$ を代入した

$$\det(A)$$

が定数項となる

多項式の最高次の係数に $(-1)^n$ がつくのは面倒なので、$\det(A-xE)$ の代わりに、その $(-1)^n$ 倍である

$$\det(xE-A)$$

を考えることが多い

$(-1)^n$ は x に依存しない定数であり、方程式の解の集合を変えることはないので、どちらの行列式を使っても求まる固有値（0 になる x の値）は同じである

実際、$\det(xE-A)$ を展開すると、

$$\det(xE-A) = \begin{vmatrix} x-a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & x-a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{n1} & -a_{n2} & \cdots & x-a_{nn} \end{vmatrix}$$

となり、x の前に (-1) がつかずに済む

特性多項式　A を正方行列、x を変数として、

$$\Phi_A(x) = \det(xE-A)$$

とおく

これを特性多項式あるいは固有多項式と呼ぶ

特性多項式の構造　A を n 次正方行列とすると、特性多項式は、次のような n 次多項式である

$$\Phi_A(x) = x^n - \mathrm{tr}(A)x^{n-1} + \cdots + (-1)^n \det(A)$$

特性方程式　特性多項式 $\Phi_A(x)$ の根を求める方程式

$$\Phi_A(x) = 0$$

を、特性方程式あるいは固有方程式と呼ぶ

固有値の重複度

たとえば、次の方程式

$$(x-2)^3(x-1) = 0$$

の解は、$x = 2$ と $x = 1$ である

ここで、左辺を、

$$(x-2)(x-2)(x-2)(x-1) = 0$$

とみなすと、

$$\begin{aligned} x &= 2 \\ x &= 2 \\ x &= 2 \\ x &= 1 \end{aligned}$$

というように解が重複していることがわかる

このように、「何回同じ解が現れるか？」を数えたものを重複度という

方程式の解の重複度　多項式 $f(x)$ で表される方程式 $f(x)=0$ において、$f(x)$ が $(x-\alpha)^m$ で割り切れるが、$(x-\alpha)^{m+1}$ では割り切れないような定数 α と自然数 m が存在するとき、α はこの方程式の m 重解あるいは m 重根であるといい、m を α の重複度と呼ぶ

上の定義は難しく聞こえるが、「ちょうど m 回だけ $(x-\alpha)$ がかかっている」ということの言い換えにすぎない

たとえば、

$$(x-2)^3(x-1)$$

を $(x-2)^3$ で割ると、

$$(x-1)$$

として割り切れるが、$(x-2)^4$ で割ると、

$$\frac{(x-2)^3(x-1)}{(x-2)^4}=\frac{x-1}{x-2}=\frac{A}{x-1}+\frac{B}{x-2}$$

というように部分分数分解できるので、余りが出ていることがわかる（多項式の割り算における余りとは、$f(x)=g(x)q(x)+r(x)$ の $r(x)$ のことである）

つまり、$f(x)$ に因数 $(x-\alpha)$ が m 個含まれている場合、$f(x)$ は $(x-\alpha)^m$ で割り切れるが、m 個以上は含まれていないので、$(x-\alpha)^{m+1}$ で割ると余りが出てしまう

これはすなわち、「ちょうど m 回だけ $(x-\alpha)$ がかかっている」ということである

特性多項式が

$$\Phi_A(x)=(x-\alpha_1)^{k_1}\cdots(x-\alpha_s)^{k_s}$$

と因数分解できることは、代数学の基本定理によって保証されている

代数学の基本定理　定数でない任意の 1 変数多項式は、複素計数の 1 次式の積に分解できる

そこで、固有値の重複度を次のように定義する

固有値の重複度　特性多項式を因数分解して、

$$\Phi_A(x) = (x - \alpha_1)^{k_1} \cdots (x - \alpha_s)^{k_s}$$

とする

ここで、$\alpha_1, \ldots, \alpha_s$ は相異なるものとする

k_i は 1 以上の整数であり、これを固有値 α_i の**重複度**と呼ぶ

$\Phi_A(x)$ は n 次多項式であるから、

$$\sum_{i=1}^{s} k_i = n$$

が成り立つ

相似な行列の特性多項式

行列式の乗法性により、正方行列 A, B が、ある正則行列 P に対して

$$B = P^{-1}AP$$

となる（A と B が相似である）ならば、次のように A と B の特性多項式は一致する

$$\begin{aligned}
\det(xE - B) &= \det\left(xE - P^{-1}AP\right) \\
&= \det\left(xPP^{-1} - P^{-1}AP\right) \\
&= \det\left(P^{-1}P(x - A)\right) \\
&= \det\left(P^{-1}P(xE - A)\right) \\
&= \det(E(xE - A)) \\
&= \det(E)\det(xE - A) \\
&= \det(xE - A)
\end{aligned}$$

相似な行列の特性多項式　相似な行列の特性多項式は一致する

この事実は、すなわち次の事実を意味する

⚓ **相似な行列の固有値**　相似な行列の固有値は重複度も含めて一致する

n 次元正方行列 $A = (a_{ij})$ の特性多項式が、

$$\Phi_A(x) = x^n - \mathrm{tr}(A)x^{n-1} + \cdots + (-1)^n \det(A)$$

であることを思い出すと、次のことがいえる

⚓ **相似な行列のトレースと行列式**　A と B が相似ならば、

$$\begin{aligned} \mathrm{tr}(A) &= \mathrm{tr}(B) \\ \det(A) &= \det(B) \end{aligned}$$

さらに、A が対角化可能であるときには、

$$B = P^{-1}AP = \begin{pmatrix} \alpha_1 & 0 & \cdots & 0 \\ 0 & \alpha_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \alpha_n \end{pmatrix}$$

という行列 B と A が相似であるので、

$$\begin{aligned} \mathrm{tr}(A) &= \mathrm{tr}(B) = \alpha_1 + \alpha_2 + \cdots + \alpha_n \\ \det(A) &= \det(B) = \alpha_1\alpha_2\cdots\alpha_n \end{aligned}$$

であることがわかる

⚓ **対角化可能行列の固有値による不変量の表現**　行列 A が対角化可能であるとき、

- A のトレースは A の固有値の和
- A の行列式は A の固有値の積

さて、A と B が相似であるとき、A と B は 1 つの線形変換 f を異なる基底によって表現して得られた行列であるという関係にある

このとき、A と B の特性多項式が一致するということは、次のように言い換えられる

特性多項式の基底不変性　線形空間 V の線形変換 f に対して、V のある基底に関する表現行列 A の特性多項式 $\Phi_A(x)$ は、基底の選び方によらず f のみによって決まる

対角化可能な行列の特性多項式

A が P によって対角化されたとして、

$$P^{-1}AP = \mathrm{diag}(c_1, \ldots, c_n)$$

とすると、A の特性多項式は、

$$\begin{aligned}
\Phi_A(x) &= \Phi_{P^{-1}AP}(x) \\
&= \det\left(xE - P^{-1}AP\right) \\
&= \begin{vmatrix} x - c_1 & 0 & \cdots & 0 \\ 0 & x - c_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & x - c_n \end{vmatrix} \\
&= (x - c_1)(x - c_2)\cdots(x - c_n)
\end{aligned}$$

となる

一般の場合の特性多項式

$$(x - \alpha_1)^{k_1} \cdots (x - \alpha_s)^{k_s}$$

と見比べると、各 $1 \leq i \leq s$ に対して、$c_1, \ldots, c_n$ の中には α_i が k_i 個あることがわかる

つまり、対角成分として現れた数たち $c_1, \ldots, c_n$ は、重複度を含めて、特性多項式 $\Phi_A(x)$ の根と一致する

対角化と特性多項式の根　対角化可能な行列 A を対角化して得られる対角行列の対角成分たちは、重複度を含めて特性多項式の根と一致する

また、対角化可能性と固有ベクトルの線型独立性の証明過程を振り返ると、次のようにまとめられる

対角化行列の列ベクトルと固有ベクトルの対応　対角化可能な行列 A を対角化する正則行列 P の列ベクトルはすべて A の固有ベクトルであり、固有値 α_i のものが k_i 個ある
ここで、k_i は α_i の重複度である

固有空間

線形空間 V の中に、行列の固有ベクトルが「どれくらい」あるかを調べるため、各固有値 α に対して、α の固有ベクトルと $\mathbf{0}$ からなる V の部分集合を考える

α が A の固有値ならば、方程式

$$(\alpha E - A)\boldsymbol{x} = \mathbf{0}$$

の解空間、すなわち核空間 $\mathrm{Ker}(\alpha E - A)$ は、固有値 α を持つ A の固有ベクトルのすべてと $\mathbf{0}$ からなる

核空間は V の部分空間であり、これを固有値 α の固有空間と呼ぶ

固有空間　α が A の固有値であるとき、

$$W(\alpha) = \mathrm{Ker}(\alpha E - A)$$

を固有値 α の固有空間と呼ぶ

固有空間の次元

固有空間の次元と固有値の重複度　A の固有値 α_i の重複度 k_i と、固有空間 $W(\alpha_i)$ の次元 $\dim W(\alpha_i)$ に対し、次の不等式が成立する

$$\dim W(\alpha_i) \leq k_i \quad (1 \leq i \leq s)$$

証明

$W(\alpha_i)$ の基底 $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m$ をとる

$\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m, \boldsymbol{v}_{m+1}, \ldots, \boldsymbol{v}_n$ が K^n の基底となるように、$n-m$ 個のベクトル $\boldsymbol{v}_{m+1}, \ldots, \boldsymbol{v}_n$ を追加して基底を延長する

$P = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m, \boldsymbol{v}_{m+1}, \ldots, \boldsymbol{v}_n)$ とするとき、

$$AP = (A\boldsymbol{v}_1, \ldots, A\boldsymbol{v}_m, A\boldsymbol{v}_{m+1}, \ldots, A\boldsymbol{v}_n)$$

ここで、$W(\alpha_i)$ の基底 $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m$ は A の固有ベクトルであるので、固有値と固有ベクトルの定義より、

$$\begin{aligned}
AP &= (\alpha_i\boldsymbol{v}_1, \ldots, \alpha_i\boldsymbol{v}_m, A\boldsymbol{v}_{m+1}, \ldots, A\boldsymbol{v}_n) \\
&= (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n) \begin{array}{cc} \overset{m}{\longleftrightarrow} & \overset{n-m}{\longleftrightarrow} \\ \left(\begin{array}{ccc|c} \alpha_i & & & \\ & \ddots & & B \\ & & \alpha_i & \\ \hline & O & & C \end{array}\right) & \begin{array}{l} \updownarrow m \\ \updownarrow n-m \end{array} \end{array} \\
&= P \begin{pmatrix} \alpha_i E_m & B \\ O & C \end{pmatrix}
\end{aligned}$$

基底の線型独立性より、線型独立な列ベクトルを並べた行列 P は正則であるので、

$$P^{-1}AP = \begin{pmatrix} \alpha_i E_m & B \\ O & C \end{pmatrix}$$

この行列の特性多項式を考えると、

$$\begin{aligned}
\Phi_{P^{-1}AP}(x) &= \det\bigl(xE - P^{-1}AP\bigr) \\
&= \left| \begin{array}{ccc|c} x-\alpha_i & & & \\ & \ddots & & -B \\ & & x-\alpha_i & \\ \hline & O & & xE' - C \end{array} \right| \\
&= \begin{vmatrix} x-\alpha_i & & \\ & \ddots & \\ & & x-\alpha_i \end{vmatrix} \det(xE' - C) \\
&= (x-\alpha_i)^m \det(xE' - C)
\end{aligned}$$

より、固有値 α_i の重複度 k_i は m 以上となる

$$m \leq k_i$$

m は $W(\alpha)_i$ の基底を構成するベクトルの個数、すなわち $\dim W(\alpha_i)$ であるので、

$$\dim W(\alpha_i) \leq k_i$$

が成り立つ ■

この定理の証明過程で登場した特性方程式

$$\Phi_{P^{-1}AP}(x) = (x-\alpha_i)^m \det(xE' - C)$$

において、$\det(xE' - C)$ からも $(x - \alpha_i)$ が現れれば、α_i の重複度 k_i は m より大きくなることがわかる

$W(\alpha)$ の基底を構成するベクトルの個数は、固有値 α に属する線型独立な固有ベクトルの個数ともいえる

また、固有値 α の重複度が k であることは、特性方程式が $x=\alpha$ を k 重解にもつことを意味する

以上をふまえると、前述の定理は、特性方程式の視点で次のように言い換えられる

固有値の重複度と固有ベクトルの最大数　正方行列 A の特性方程式 $\Phi_A(x)=0$ が $x=\alpha$ を k 重解にもつとき、固有値 α に属する線型独立な固有ベクトルは、k 個以下しかとれない

対角化可能性

次の定理は、

> 各固有値の固有空間が「可能な限り大きい」

ときに限り、対角化可能であると述べている

固有空間次元と重複度の一致による対角化可能性　A の固有値を α_i、その重複度を k_i とする

A が対角化可能であることは、次と同値である

$$\dim W(\alpha_i)=k_i \quad (1\leq i\leq s)$$

証明

対角化可能 $\Longrightarrow$ 固有空間の次元と重複度が一致

A が対角化可能であるので、正則行列 P により $P^{-1}AP$ が対角行列になる

このとき、P の列ベクトルからなる A の固有ベクトルの集合には、固有値 α_i を持つものが k_i 個含まれる

各 i に対して、k_i 個の線型独立なベクトルが $W(\alpha_i)$ に含まれることになるため、

$$\dim W(\alpha_i) \geq k_i$$

がいえる

一方、固有空間の次元と固有値の重複度の不等式より、

$$\dim W(\alpha_i) \leq k_i$$

したがって、

$$\dim W(\alpha_i) = k_i$$

が成り立つ ■

固有空間の次元と重複度が一致 $\Longrightarrow$ 対角化可能

$\dim W(\alpha_i) = k_i$ が成り立つとし、$W(\alpha_i)$ の基底 $\mathcal{V}_i$ をとる

$\mathcal{V}_i$ は k_i 個の元からなり、これらは $W(\alpha_i)$ の基底であることから、線形独立な固有ベクトルである

さらに、異なる固有値に対応する固有ベクトルは線形独立であるから、$i \neq j$ とし $\mathcal{V}_i$ と $\mathcal{V}_j$ のベクトルは互いに線形独立である

そこで、すべての $\mathcal{V}_i$ を併せた集合

$$\mathcal{V} = \bigcup_{i=1}^{s} \mathcal{V}_i$$

を考えると、$\mathcal{V}$ のベクトルは線型独立である

このとき、$\mathcal{V}$ の元の個数は

$$\sum_{i=1}^{s} k_i = n$$

である

したがって、線型独立な n 個の固有ベクトルが存在するため、A は対角化可能である ■

次の補題をもとに、対角化可能性を特性方程式の言葉で述べることができる

特性方程式の単根性と固有空間の次元　特性方程式 $\Phi_A(x)$ において α_i が単根ならば、すなわち $k_i = 1$ ならば、

$$\dim W(\alpha_i) = 1$$

証明

α_i は固有値なので、$\alpha_i \neq \mathbf{0}$ より、$W(\alpha_i) \neq \{\mathbf{0}\}$ がいえる

これはつまり、

$$\dim W(\alpha_i) \geq 1$$

ということだが、固有空間の次元と固有値の重複度に関する不等式より、

$$\dim W(\alpha_i) \leq k_i = 1$$

も成り立つ

したがって、

$$\dim W(\alpha_i) = 1$$

である　■

特性方程式の単根性と対角化可能性　特性方程式 $\Phi_A(x)$ が重根を持たなければ、A は対角化可能である

証明

重根を持たないということは、各固有値の重複度 k_i は 1 である

よって、

$$\dim W(\alpha_i) = 1$$

となり、k_i と $\dim W(\alpha_i)$ が一致するので、A は対角化可能である ■

固有空間分解

[Todo 39:]

第 17 章

内積と計量空間

内積：ベクトルの「近さ」を返す関数

2 つのベクトルがどれくらい似ているかを議論するために、内積という尺度を導入しよう。

内積は、2 つのベクトルを引数にとり、その「近さ」を表すスカラー値を返す関数として定義する。

具体的な定義式を知る前に、「近さ」を測る道具として、どのような性質を持っていてほしいかを整理しておこう。
具体的な定義式は、その性質を満たすように「作る」ことにする。

内積の公理（$\mathbb{R}$ 上の線形空間） $\mathbb{R}$ 上の線形空間 V を考え、$\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in V$, $c \in \mathbb{R}$ とする。
2 つのベクトルを引数にとり、実数を返す関数 $(\cdot, \cdot) : V \times V \to \mathbb{R}$ として、次の性質を満たすものを内積という。

対称性

$$(\boldsymbol{u}, \boldsymbol{v}) = (\boldsymbol{v}, \boldsymbol{u})$$

双線形性 1. スカラー倍

$$(c\boldsymbol{u}, \boldsymbol{v}) = (\boldsymbol{u}, c\boldsymbol{v}) = c(\boldsymbol{u}, \boldsymbol{v})$$

双線形性 2. 和

$$(\boldsymbol{u} + \boldsymbol{w}, \boldsymbol{v}) = (\boldsymbol{u}, \boldsymbol{v}) + (\boldsymbol{w}, \boldsymbol{v})$$

$$(\boldsymbol{u}, \boldsymbol{v} + \boldsymbol{w}) = (\boldsymbol{u}, \boldsymbol{v}) + (\boldsymbol{u}, \boldsymbol{w})$$

正定値性

$$(\boldsymbol{u}, \boldsymbol{u}) \geq 0, (\boldsymbol{u}, \boldsymbol{u}) = 0 \iff \boldsymbol{u} = \boldsymbol{0}$$

内積が定められた線形空間を、計量線形空間、または単に計量空間という。（計量空間の厳密な定義は後に述べる。）

まずは、ここで示した内積の持つべき性質のそれぞれの意図を考えてみよう。

対称性

$\boldsymbol{u}$ が $\boldsymbol{v}$ にどれくらい近いか？という視点で測っても、$\boldsymbol{v}$ が $\boldsymbol{u}$ にどれくらい近いか？という視点で測っても、得られる「近さ」は同じであってほしい、という性質。

双線形性

どちらかのベクトルをスカラー倍してから「近さ」を測りたいとき、元のベクトルとの近さを測っておいて、それを定数倍することでも目的の「近さ」を求められる、という性質。

また、ほかのベクトルを足してから「近さ」を測りたいとき、足し合わせたいベクトルそれぞれについて近さを測っておいて、それを合計することでも目的の「近さ」を求められる、という性質。

これらは、近さを測るという「操作」と「演算」が入れ替え可能であるという、線形性と呼ばれる性質である。
2 つの引数 $\boldsymbol{u}, \boldsymbol{v}$ のどちらに関しても線形性があるということで、「双」がついている。

正定値性

ベクトルの「近さ」とは、向きがどれくらい近いか、という尺度でもある。
同じ方向なら正の数、逆の方向なら負の数をとるのが自然だと考えられる。

自分自身との「近さ」を測るとき、自分と自分は完全に同じ向きであるから、その「近さ」は正の数であるはずだ。
自分自身との「近さ」が 0 になるようなベクトルは、零ベクトル $\boldsymbol{o}$ だけである。

ノルム：ベクトルの「大きさ」

内積の正定値性より、自分自身との内積は次の性質を持つ。

- 常に正の数である
- 零ベクトルのときだけ 0 になる

零ベクトルのときだけ 0 になることから、自分自身との内積は、そのベクトル自身の大きさ（長さ）に応じてスケーリングする量なのではないか？と予想される。また、大きさを表す量は、当然正の数である必要がある。

このように、内積の正定値性は、自分自身との内積を使ってそのベクトルの「大きさ」を測ることができることを示唆する性質と考えることができる。

ベクトルの「大きさ」を表現する量をノルムという。

ベクトルの「大きさ」の測り方（ノルムの具体的な定義）はさまざま考えられ、用途によって使い分けられるが、ノルムを名乗るものはどれも次の性質を満たすように作る必要がある。

ノルムの公理（$\mathbb{R}$ 上の線形空間） $\mathbb{R}$ 上の線形空間 V を考え、$\boldsymbol{u}, \boldsymbol{v} \in V,\ c \in \mathbb{R}$ とする。
ベクトルを引数にとり、非負の実数を返す関数 $\|\cdot\| : V \to \mathbb{R}$ として、次の性質を満たすものをノルムという。

非負性

$$\|\boldsymbol{u}\| \geq 0, \|\boldsymbol{u}\| = 0 \iff \boldsymbol{u} = \boldsymbol{0}$$

斉次性

$$\|c\boldsymbol{u}\| = |c|\|\boldsymbol{u}\|$$

三角不等式

$$\|\boldsymbol{u} + \boldsymbol{v}\| \leq \|\boldsymbol{u}\| + \|\boldsymbol{v}\|$$

ノルムの非負性は、内積の正定値性そのものである。

斉次性についてはどうだろうか？

スカラー倍したベクトルの自身との内積を考えてみると、スカラー倍に対する内積の双線形性より、次のようになる。

$$(c\boldsymbol{u}, c\boldsymbol{u}) = c^2(\boldsymbol{u}, \boldsymbol{u})$$

これはノルムの斉次性を満たしていない。

自分自身との内積をそのままノルムとして使おうとすると、ベクトルを c 倍したらその長さは c^2 倍されるという、不自然な定義になってしまう。

そこで、二乗を消すために平方根をとることで、斉次性も満たす量を得ることができる。

$$\sqrt{(c\boldsymbol{u}, c\boldsymbol{u})} = \sqrt{c^2(\boldsymbol{u}, \boldsymbol{u})} = |c|\sqrt{(\boldsymbol{u}, \boldsymbol{u})}$$

このように自身との内積の平方根をとった形を、ベクトル $\boldsymbol{u}$ のノルムと定義することにする。

ベクトルのノルム　計量空間 V 上のベクトル $\boldsymbol{a}$ のノルム（長さ）を次のように定義する。

$$\|\boldsymbol{a}\| = \sqrt{(\boldsymbol{a}, \boldsymbol{a})}$$

コーシー・シュワルツの不等式

内積の公理だけを用いて、次の重要な不等式を導くことができる。

コーシー・シュワルツの不等式（ベクトルの内積） 計量空間 V 上のベクトル $\boldsymbol{u}, \boldsymbol{v}$ に対して、次が成り立つ。

$$(\boldsymbol{u}, \boldsymbol{v})^2 \leq (\boldsymbol{u}, \boldsymbol{u})(\boldsymbol{v}, \boldsymbol{v})$$

証明

内積の正定値性より、任意の $t \in \mathbb{R}$ に対して、

$$(\boldsymbol{u} - t\boldsymbol{v}, \boldsymbol{u} - t\boldsymbol{v}) \geq 0$$

が成り立つ。

ここで、内積の双線形性を用いて左辺を展開すると、

$$(\boldsymbol{u}, \boldsymbol{u}) - 2t(\boldsymbol{u}, \boldsymbol{v}) + t^2(\boldsymbol{v}, \boldsymbol{v}) \geq 0$$

これは t についての 2 次式であり、実数全体で非負ということは判別式が非正でなければならない。

$$\begin{aligned} (-2(\boldsymbol{u}, \boldsymbol{v}))^2 - 4(\boldsymbol{u}, \boldsymbol{u})(\boldsymbol{v}, \boldsymbol{v}) &\leq 0 \\ 4(\boldsymbol{u}, \boldsymbol{v})^2 &\leq 4(\boldsymbol{u}, \boldsymbol{u})(\boldsymbol{v}, \boldsymbol{v}) \end{aligned}$$

よって、両辺を 4 で割ると

$$(\boldsymbol{u}, \boldsymbol{v})^2 \leq (\boldsymbol{u}, \boldsymbol{u})(\boldsymbol{v}, \boldsymbol{v})$$

が得られる。 ■

この不等式は、$\boldsymbol{u}$ と $\boldsymbol{v}$ の近さを測って掛け合わせても、$\boldsymbol{u}$ 自身、$\boldsymbol{v}$ 自身との近さの積には勝てないことを表している。

$$(\boldsymbol{u}, \boldsymbol{v})(\boldsymbol{u}, \boldsymbol{v}) \leq (\boldsymbol{u}, \boldsymbol{u})(\boldsymbol{v}, \boldsymbol{v})$$

内積（の絶対値や 2 乗）は 2 つのベクトルが似ているほど大きくなり、2 つのベクトルが完全に一致する場合に最大となることを示唆しているようにも捉えられる。

このコーシー・シュワルツの不等式は、両辺の平方根をとることで、次のようにも書ける。

コーシー・シュワルツの不等式（ベクトルのノルム）　計量空間 V 上のベクトル $\boldsymbol{u}, \boldsymbol{v}$ に対して、次が成り立つ。

$$|(\boldsymbol{u}, \boldsymbol{v})| \leq \|\boldsymbol{u}\|\|\boldsymbol{v}\|$$

ベクトルのなす角

コーシー・シュワルツの不等式は、絶対値の性質から、次のように書き換えられる。

$$-\|\boldsymbol{u}\|\|\boldsymbol{v}\| \leq (\boldsymbol{u}, \boldsymbol{v}) \leq \|\boldsymbol{u}\|\|\boldsymbol{v}\|$$

$\boldsymbol{u}$, $\boldsymbol{v}$ が $\boldsymbol{o}$ でないときには、

$$-1 \leq \frac{(\boldsymbol{u}, \boldsymbol{v})}{\|\boldsymbol{u}\|\|\boldsymbol{v}\|} \leq 1$$

となるので、

$$\cos\theta = \frac{(\boldsymbol{u}, \boldsymbol{v})}{\|\boldsymbol{u}\|\|\boldsymbol{v}\|} \quad (0 \leq \theta \leq \pi)$$

を介して $\boldsymbol{u}$, $\boldsymbol{v}$ のなす角を定義できる。

ベクトルのなす角　計量空間 V 上のベクトル $\boldsymbol{u}, \boldsymbol{v}$ に対して、

$$\cos\theta = \frac{(\boldsymbol{u}, \boldsymbol{v})}{\|\boldsymbol{u}\|\|\boldsymbol{v}\|} \quad (0 \leq \theta \leq \pi)$$

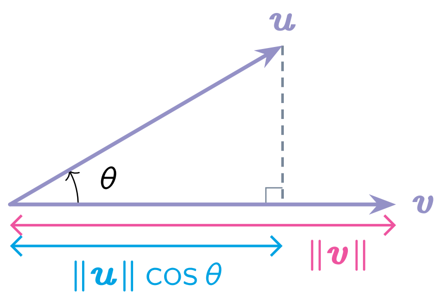
により定まる θ を $\boldsymbol{u}$, $\boldsymbol{v}$ のなす角という。

内積が表す「関係の強さ」

ベクトルのなす角の式から、内積の幾何的な解釈を捉えることができる。

$$(\boldsymbol{u}, \boldsymbol{v}) = \|\boldsymbol{u}\|\|\boldsymbol{v}\| \cos\theta$$

この式を、次のような図でイメージしてみよう。



この図から、内積は次のようにも解釈できる。

$\boldsymbol{u}$ の $\boldsymbol{v}$ との内積とは、

$\boldsymbol{v}$ 自身の長さと、$\boldsymbol{u}$ の $\boldsymbol{v}$ 方向の長さの積である

ここで、「$\boldsymbol{u}$ の $\boldsymbol{v}$ 方向の長さ」は、後に正射影という量として定義する。

自分自身との内積の再解釈

$\boldsymbol{v}$ と $\boldsymbol{u}$ が平行でまったく同じ方向を向いている場合、「$\boldsymbol{u}$ の $\boldsymbol{v}$ 方向の長さ」は、$\boldsymbol{u}$ の長さそのものである。

$$\|\boldsymbol{u}\| \cos\theta = \|\boldsymbol{u}\|$$

また、このとき、$\boldsymbol{v}$ と $\boldsymbol{u}$ は互いに正の数のスカラー倍で表すことができるので、

$$\|\boldsymbol{v}\| = k\|\boldsymbol{u}\| \quad (k > 0)$$

すると、内積は、

$$(\boldsymbol{u}, \boldsymbol{v}) = k\|\boldsymbol{u}\|^2$$

ここで、$\boldsymbol{v} = \boldsymbol{u}$ の場合は、$\boldsymbol{u}$ を特にスケーリングしなくても $\boldsymbol{v}$ に一致するので、$k = 1$ である。

$$(\boldsymbol{u}, \boldsymbol{u}) = \|\boldsymbol{u}\|^2$$

このように $\boldsymbol{u}$ の $\boldsymbol{u}$ 方向の長さは $\boldsymbol{u}$ 自身の長さであることから、自分自身との内積は 長さ^2 となる。
その平方根をとれば長さが得られるということで、ベクトルのノルムの定義

$$\|\boldsymbol{u}\| = \sqrt{(\boldsymbol{u}, \boldsymbol{u})}$$

を自然に解釈することができる。

平行の度合いと内積

同じ方向を向いているベクトルどうしは、平行に近ければ近いほど、これらは互いに似ていて「関係性の強い」ベクトルだといえる。

2 つのベクトルが同方向で完全に平行なとき、なす角 θ は 0 であるので、$\cos\theta$ の値は 1（$\cos\theta$ の最大値）となる。
つまり、同方向で平行に近い「似た」ベクトルほど、内積の値は最大値に近くなる。

逆方向と内積の符号

一方、2 つのベクトルが完全に平行で、逆の方向を向いているなら、片方のベクトルはもう片方のベクトルを負の数を使ってスカラー倍したものになる。

逆向きのベクトルどうしは、近い方向どころかむしろ「かけ離れた方向を向いている」といえる。
内積が「向きの似ている度合い」なら、「近い方向を向いている」度合いを正の数で、「かけ離れた方向を向いている」度合いを負の数で表すのが自然である。

実際、2 つのベクトルが逆向きで完全に平行なとき、$\cos\theta$ の値は -1（$\cos\theta$ の最小値）となる。
つまり、逆方向に近い「かけ離れた」ベクトルほど、内積の値は最小値に近くなる。

ベクトルの直交

「同じ向きに近い」場合と「逆向きに近い」場合が切り替わるのは、2 つのベクトルどうしが垂直なときである。
ならば、内積の正と負が切り替わる境界、すなわち内積が 0 になる場合とは、2 つのベクトルが直交する場合であるのが自然ではないだろうか。

別な考え方として、完全に垂直な 2 つのベクトルは、互いに全く影響を与えない方向を向いている。
2 つのベクトルが直交している場合、2 つのベクトルは互いに全く関係がないものとして、関係の強さを表す内積の値は 0 にしたい。

実際、内積の定義はこの解釈に沿うものになっている。
$\boldsymbol{u}$ と $\boldsymbol{v}$ のなす角が直角であるとき、$\cos\theta = 0$ となるので、内積も 0 になる。

幾何学的なイメージができない高次元の場合についても、内積が 0 になること、すなわち 2 つのベクトルが無関係であることを直交の定義としてしまおう。

ベクトルの直交　計量空間 V 上のベクトル $\boldsymbol{u}, \boldsymbol{v}$ に対して、

$$(\boldsymbol{u}, \boldsymbol{v}) = 0$$

が成り立つとき、$\boldsymbol{u}$ と $\boldsymbol{v}$ は直交するといい、

$$\boldsymbol{u} \perp \boldsymbol{v}$$

と表記する。

直交系と直交基底

直交系と直交基底　計量空間 V の $\mathbf{0}$ でないベクトル $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n$ がどの 2 つも互いに直交する、すなわち、

$$(\boldsymbol{a}_i, \boldsymbol{a}_j) = 0 \quad (i \neq j)$$

が成り立つとき、$\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n$ を直交系という

直交系が V の基底であるとき、直交基底と呼ばれる

直交系の線型独立性

直交系の線型独立性　計量空間の直交系 $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n$ は線型独立である

証明

係数 $c_1, c_2, \ldots, c_n \in K$ を用いた線形関係式

$$c_1\boldsymbol{a}_1 + c_2\boldsymbol{a}_2 + \cdots + c_n\boldsymbol{a}_n = \mathbf{0}$$

を考える

このとき、$\boldsymbol{a}_j \ (j = 1, 2, \ldots, n)$ との内積をとると、

$$(c_1\boldsymbol{a}_1 + c_2\boldsymbol{a}_2 + \cdots + c_n\boldsymbol{a}_n, \boldsymbol{a}_j) = 0$$

内積の双線形性より、

$$c_1(\boldsymbol{a}_1, \boldsymbol{a}_j) + c_2(\boldsymbol{a}_2, \boldsymbol{a}_j) + \cdots + c_n(\boldsymbol{a}_n, \boldsymbol{a}_j) = 0$$

$$\sum_{i=1}^{n} c_i(\boldsymbol{a}_i, \boldsymbol{a}_j) = 0$$

ここで、$\boldsymbol{a}_i$ は直交系であることから、$i \neq j$ の場合、

$$(\boldsymbol{a}_i, \boldsymbol{a}_j) = 0$$

よって、$i \neq j$ の項はすべて 0 になり、残るのは

$$c_j(\boldsymbol{a}_j, \boldsymbol{a}_j) = 0$$

ここで、直交系の定義より、$\boldsymbol{a}_j \neq \boldsymbol{0}$ なので、

$$(\boldsymbol{a}_j, \boldsymbol{a}_j) \neq 0$$

よって、$c_j = 0$ でなければならず、これは $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n$ が線型独立であることを意味する ■

正規直交系と正規直交基底

正規直交系と正規直交基底 計量空間 V の $\boldsymbol{0}$ でないベクトル $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n$ が直交系であり、さらに、どのベクトルもそのノルムが 1 に等しいとき、$\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n$ を正規直交系という

正規直交系が V の基底であるとき、正規直交基底と呼ばれる

正規直交基底と内積

正規直交基底どうしの内積は、クロネッカーのデルタ記号を用いて、簡潔に表現できる

クロネッカーのデルタ 次のように定義される δ_{ij} をクロネッカーのデルタという

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

正規直交基底同士の内積 計量空間 V の正規直交基底 $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n$ の内積に関して、次が成り立つ

$$(\boldsymbol{e}_i, \boldsymbol{e}_j) = \delta_{ij} \quad (i, j = 1, 2, \ldots, n)$$

証明

$\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n$ の直交性より、$i \neq j$ のときは、

$$(\boldsymbol{e}_i, \boldsymbol{e}_j) = 0$$

また、$\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n$ はすべてノルムが 1 であることから、$i = j$ のときは、

$$(\boldsymbol{e}_i, \boldsymbol{e}_i) = \|\boldsymbol{e}_i\|^2 = 1$$

この場合分けとそれぞれの結果は、クロネッカーのデルタ記号の定義と一致する ■

正規直交基底による表現行列の展開

$\mathbb{R}^n$ から $\mathbb{R}^m$ への線形写像は、ある $m \times n$ 型行列 A によって表現される

これを定める基本的な方法は、

1. 定義域 $\mathbb{R}^n$ に一つの正規直交基底（互いに直交する単位ベクトル）$\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n\}$ を定める
2. それぞれが写像されるべき m 次元ベクトル（像）$\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ を指定する

という手順であり、このとき、行列 A は

$$A = \begin{pmatrix} \boldsymbol{a}_1 \\ \vdots \\ \boldsymbol{a}_n \end{pmatrix} \begin{pmatrix} \boldsymbol{u}_1 & \cdots & \boldsymbol{u}_n \end{pmatrix} = \boldsymbol{a}_1 \boldsymbol{u}_1^\top + \cdots + \boldsymbol{a}_n \boldsymbol{u}_n^\top$$

と書くことができる（$\top$ は転置を表す）

正規直交基底による表現行列の展開　$\mathbb{R}^n$ から $\mathbb{R}^m$ への線形写像 f の表現行列 A は、$\mathbb{R}^n$ の正規直交基底 $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n\}$ を用いて、次のように表すことができる

$$A = \sum_{i=1}^{n} f(\boldsymbol{u}_i) \boldsymbol{u}_i^\top$$

実際、両辺に $\boldsymbol{u}_i$ をかけると、

$$A\boldsymbol{u}_i = \sum_{j=1}^{n} \boldsymbol{a}_j \boldsymbol{u}_j^\top \boldsymbol{u}_i = \sum_{j=1}^{n} \boldsymbol{a}_j \delta_{ij} = \boldsymbol{a}_i$$

より、

$$A\boldsymbol{u}_i = \boldsymbol{a}_i \quad (i = 1, \ldots, n)$$

が成り立つことがわかる

特に、$\mathbb{R}^n$ の正規直交基底として標準基底 $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ を選ぶと、行列 A は次のように表せる

$$\begin{aligned} A &= \sum_{i=1}^{n} \boldsymbol{a}_i \boldsymbol{e}_i^\top \\ &= \begin{pmatrix} \boldsymbol{a}_{11} \\ \vdots \\ \boldsymbol{a}_{m1} \end{pmatrix} \begin{pmatrix} 1 & \cdots & 0 \end{pmatrix} + \cdots + \begin{pmatrix} \boldsymbol{a}_{1n} \\ \vdots \\ \boldsymbol{a}_{mn} \end{pmatrix} \begin{pmatrix} 0 & \cdots & 1 \end{pmatrix} \\ &= \begin{pmatrix} \boldsymbol{a}_{11} & \cdots & \boldsymbol{a}_{1n} \\ \vdots & \ddots & \vdots \\ \boldsymbol{a}_{m1} & \cdots & \boldsymbol{a}_{mn} \end{pmatrix} \end{aligned}$$

すなわち、表現行列 A は、

像 $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ を列として順に並べた行列 $\begin{pmatrix} \boldsymbol{a}_1 & \cdots & \boldsymbol{a}_n \end{pmatrix}$

となる

$\mathbb{R}^n$ 上の内積

内積の公理と、正規直交基底の内積をもとに、$\mathbb{R}^n$ 上の内積を作ることができる。

まず、任意のベクトル $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^n$ を、正規直交基底の一次結合として表そう。

$$\boldsymbol{a} = a_1\boldsymbol{e}_1 + \cdots + a_n\boldsymbol{e}_n = \sum_{i=1}^{n} a_i\boldsymbol{e}_i$$

$$\boldsymbol{b} = b_1\boldsymbol{e}_1 + \cdots + b_n\boldsymbol{e}_n = \sum_{j=1}^{n} b_j\boldsymbol{e}_j$$

これらの内積を、双線形性を使って展開していく。

まず、和に関する双線形性より、「足してから内積を計算」と「内積を計算してから足す」は同じ結果になるので、シグマ記号 $\sum$ を内積の外に出すことができる。

また、スカラー倍に関する双線形性より、定数 a_i, b_j も内積の外に出すことができる。

$$\begin{aligned}(\boldsymbol{a}, \boldsymbol{b}) &= \left(\sum_{i=1}^{n} a_i\boldsymbol{e}_i, \sum_{j=1}^{n} b_j\boldsymbol{e}_j\right) \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} a_ib_j(\boldsymbol{e}_i, \boldsymbol{e}_j)\end{aligned}$$

正規直交基底の内積 $(\boldsymbol{e}_i, \boldsymbol{e}_j)$ はクロネッカーのデルタ δ_{ij} で表せるので、次のように書き換えられる。

$$\begin{aligned}(\boldsymbol{a}, \boldsymbol{b}) &= \sum_{i=1}^{n}\sum_{j=1}^{n} a_ib_j(\boldsymbol{e}_i, \boldsymbol{e}_j) \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} a_ib_j\delta_{ij}\end{aligned}$$

ここで、δ_{ij} は $i \neq j$ のとき 0 になるので、$i = j$ の項しか残らない。

$$\begin{aligned}(\boldsymbol{a}, \boldsymbol{b}) &= \sum_{i=1}^{n}\sum_{j=1}^{n} a_ib_j\delta_{ij} \\ &= \sum_{i=1}^{n} a_ib_i\delta_{ii}\end{aligned}$$

δ_{ii} は常に 1 なので、最終的に次のような式が得られる。

$$(\boldsymbol{a}, \boldsymbol{b}) = \sum_{i=1}^{n} a_ib_i$$

$\mathbb{R}^n$ 上の内積　$\boldsymbol{a} = (a_i)_{i=1}^n$, $\boldsymbol{b} = (b_i)_{i=1}^n \in \mathbb{R}^n$ に対して、

$$(\boldsymbol{a}, \boldsymbol{b}) = \sum_{i=1}^{n} a_i b_i$$

を $\mathbb{R}^n$ 上の内積と呼ぶ。

数ベクトルの同じ位置にある数どうしをかけ算して、それらを足し合わせる、という形になっている。

$\mathbb{R}^n$ 上の内積の性質

逆に、このように定義した $\mathbb{R}^n$ 上の内積が、内積の公理を満たしていることを確認してみよう。

$\mathbb{R}^n$ 上の内積の双線形性　$\boldsymbol{u}$, $\boldsymbol{v}$, $\boldsymbol{u}_1$, $\boldsymbol{u}_2$, $\boldsymbol{v}_1$, $\boldsymbol{v}_2 \in \mathbb{R}^n$, $c \in \mathbb{R}$ に対して、以下が成立する

i. $(\boldsymbol{u}_1 + \boldsymbol{u}_2, \boldsymbol{v}) = (\boldsymbol{u}_1, \boldsymbol{v}) + (\boldsymbol{u}_2, \boldsymbol{v})$
ii. $(c\boldsymbol{u}, \boldsymbol{v}) = c(\boldsymbol{u}, \boldsymbol{v})$
iii. $(\boldsymbol{u}, \boldsymbol{v}_1 + \boldsymbol{v}_2) = (\boldsymbol{u}, \boldsymbol{v}_1) + (\boldsymbol{u}, \boldsymbol{v}_2)$
iv. $(\boldsymbol{u}, c\boldsymbol{v}) = c(\boldsymbol{u}, \boldsymbol{v})$

証明

行列のかけ算と和に関する分配法則、行列のスカラー倍についての性質から従う　■

$\mathbb{R}^n$ 上の内積の対称性　$\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n$ に対して、次が成り立つ

$$(\boldsymbol{u}, \boldsymbol{v}) = (\boldsymbol{v}, \boldsymbol{u})$$

証明

実数同士の乗算は可換であることから、$\mathbb{R}^n$ 上の内積の定義により

$$(\boldsymbol{u}, \boldsymbol{v}) = \sum_{i=1}^{n} u_i v_i = \sum_{i=1}^{n} v_i u_i = (\boldsymbol{v}, \boldsymbol{u})$$

となり、明らかに成り立つ ■

$\mathbb{R}^n$ 上の内積の正値性　$\boldsymbol{u} \in \mathbb{R}^n$ に対して、

$$(\boldsymbol{u}, \boldsymbol{u}) \geq 0$$

であり、$\boldsymbol{u} = \boldsymbol{0}$ のときに限り、等号が成立する

証明

内積の定義より、

$$(\boldsymbol{u}, \boldsymbol{u}) = \sum_{i=1}^{n} u_i^2 \geq 0$$

である

ここで現れた u_i^2 は、u_i が 0 のときに限り 0 になるので、$\boldsymbol{u} = \boldsymbol{0}$ のときに限り、等号が成立する ■

$\mathbb{C}^n$ 上の内積

複素数 $z = a + bi$ に対して、

$$(a + bi)(a - bi) = a^2 + b^2 \geq 0$$

という式が成り立つ

このとき、$a - bi$ を z の**共役複素数**といい、$\overline{z}$ と表記する

また、$\sqrt{a^2 + b^2}$ は z の**絶対値**と呼ばれ、$|z|$ と表記する

すなわち、冒頭の不等式は、

$$|z|^2 = z\overline{z} \geq 0$$

と書き換えられる

このことを利用して、$\mathbb{C}^n$ 上の内積は、次のように定義すると $\mathbb{R}^n$ の場合の自然な拡張になる

$\mathbb{C}^n$ 上の内積（標準内積） $\boldsymbol{a} = (a_i)_{i=1}^n,\ \boldsymbol{b} = (b_i)_{i=1}^n \in \mathbb{C}^n$ に対して、

$$(\boldsymbol{a}, \boldsymbol{b}) = \sum_{i=1}^n a_i \overline{b_i}$$

を $\mathbb{C}^n$ 上の内積と定義する

この内積は標準内積、あるいは標準エルミート内積とも呼ばれる

このように定めることで、特に、

$$(\boldsymbol{a}, \boldsymbol{a}) = \sum_{i=1}^n a_i \overline{a_i} = \sum_{i=1}^n |a_i|^2 \geq 0$$

であるので、$\mathbb{R}^n$ の場合と同様に、ベクトルのノルムを定義できる

$\mathbb{R}^n$ 上の内積で成り立つ性質の多くは、$\mathbb{C}^n$ 上の内積でも成り立つが、対称性に関しては注意が必要である

標準内積の対称性 $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{C}^n$ に対して、次が成り立つ

$$(\boldsymbol{u}, \boldsymbol{v}) = \overline{(\boldsymbol{v}, \boldsymbol{u})}$$

証明

$\overline{\overline{z}} = z$ をふまえると、

$$
\begin{aligned}
\overline{(\boldsymbol{v}, \boldsymbol{u})} &= \overline{\sum_{i=1}^{n} v_i \overline{u_i}} \\
&= \sum_{i=1}^{n} \overline{v_i \overline{u_i}} \\
&= \sum_{i=1}^{n} \overline{v_i} u_i \\
&= \sum_{i=1}^{n} u_i \overline{v_i} \\
&= (\boldsymbol{u}, \boldsymbol{v})
\end{aligned}
$$

となり、目的の式が示された ■

複素数 $z = a + bi$ において、$b = 0$ の場合、z は実数である

このとき、$a + 0i = a - 0i = a$ であるから、z が実数の場合、

$$\overline{z} = z$$

が成り立つ

よって、$\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n$ であるなら、$\mathbb{C}^n$ 上の内積の対称性の式は

$$(\boldsymbol{u}, \boldsymbol{v}) = \overline{(\boldsymbol{v}, \boldsymbol{u})} = (\boldsymbol{v}, \boldsymbol{u})$$

と書き換えられ、これは $\mathbb{R}^n$ 上の内積の対称性そのものである

つまり、$\mathbb{C}^n$ 上の内積の対称性は、$\mathbb{R}^n$ 上の内積の対称性も含んだ表現になっている

転置による内積の表現

内積は、転置を用いて表現することもできる

⚓ 転置による内積の表現

$$(\boldsymbol{a}, \boldsymbol{b}) = {}^t\boldsymbol{a} \cdot \overline{\boldsymbol{b}} = (a_1,\ a_2,\ \ldots,\ a_n) \begin{pmatrix} \overline{b_1} \\ \overline{b_2} \\ \vdots \\ \overline{b_n} \end{pmatrix}$$

計量線形空間

内積の概念は、双線形性、対称性、正定値性を満たすものとして抽象化できる

計量線形空間　体 K 上の線形空間 V において、その任意の要素 $\boldsymbol{a},\ \boldsymbol{b} \in V$ に対し、次の性質

i. $(\boldsymbol{a}, \boldsymbol{b}_1 + \boldsymbol{b}_2) = (\boldsymbol{a}, \boldsymbol{b}_1) + (\boldsymbol{a}, \boldsymbol{b}_2)$
   $(\boldsymbol{a}_1 + \boldsymbol{a}_2, \boldsymbol{b}) = (\boldsymbol{a}_1, \boldsymbol{b}) + (\boldsymbol{a}_2, \boldsymbol{b})$
ii. $(c\boldsymbol{a}, \boldsymbol{b}) = c(\boldsymbol{a}, \boldsymbol{b})$
iii. $(\boldsymbol{a}, \boldsymbol{b}) = \overline{(\boldsymbol{b}, \boldsymbol{a})}$
iv. $(\boldsymbol{a}, \boldsymbol{a}) \geq 0, (\boldsymbol{a}, \boldsymbol{a}) = 0 \Longrightarrow \boldsymbol{a} = \boldsymbol{0}$

を満たす K の要素 $(\boldsymbol{a}, \boldsymbol{b})$ がただ一つ定まるとき、$(\boldsymbol{a}, \boldsymbol{b})$ を内積と呼び、V は計量線形空間、または単に計量空間であるという

内積の定義に

$$(\boldsymbol{a}, c\boldsymbol{b}) = c(\boldsymbol{a}, \boldsymbol{b})$$

が含まれていないことに注意しよう

この式は、(ii) と (iii) から導ける上、$K = \mathbb{C}^n$ の場合には成り立たない

$K = \mathbb{C}^n$ の場合を含め、一般に次が成り立つ

内積の共役線形性　計量空間 V の要素 $\boldsymbol{a}$, $\boldsymbol{b}$ の内積と $c \in K$ について、次の性質が成り立つ

$$(\boldsymbol{a}, c\boldsymbol{b}) = \overline{c}(\boldsymbol{a}, \boldsymbol{b}) \quad (c \in K)$$

証明

計量線形空間の定義の (ii) と (iii) を用いて、

$$\begin{aligned}(\boldsymbol{a}, c\boldsymbol{b}) &= \overline{(c\boldsymbol{b}, \boldsymbol{a})} \\ &= \overline{c}\overline{(\boldsymbol{b}, \boldsymbol{a})} \\ &= \overline{c}(\boldsymbol{a}, \boldsymbol{b})\end{aligned}$$

となる　■

第 18 章

正射影と直交化

観測装置としての内積

ここでは、内積 $(\boldsymbol{a}_1, \boldsymbol{a}_2)$ において、$\boldsymbol{a}_1$ と $\boldsymbol{a}_2$ の機能を分離して、

内積 $(\boldsymbol{a}_1, \boldsymbol{a}_2)$ とは、$\boldsymbol{a}_1$ で $\boldsymbol{a}_2$ を測って得られる値

という再解釈を行う。

縦ベクトルが基本

内積 $(\boldsymbol{a}_1, \boldsymbol{a}_2)$ は、次のように表すこともできた。

$$\boldsymbol{a}_1^\top \boldsymbol{a}_2$$

つまり、縦ベクトル $\boldsymbol{a}_2$ に、横ベクトル $\boldsymbol{a}_1^\top$ を作用させることで、内積が得られると読むことができる。

このように、$\boldsymbol{a}_2$ が測りたい対象で、$\boldsymbol{a}_1$ がその測定器であるという視点を持つことができる。

ブラとケット

この視点を表す上で有用なのが、ブラケット記法である。

- 横ベクトルに対応する記号として $\langle \boldsymbol{a}_1|$ を定義し、これをブラベクトルと呼ぶ
- 縦ベクトルに対応する記号として $|\boldsymbol{a}_2\rangle$ を定義し、これをケットベクトルと呼ぶ

これらの記号を用いると、内積は次のように表せる。

$$\langle \boldsymbol{a}_1|\boldsymbol{a}_2\rangle$$

このように、ブラとケットが組み合わされると、スカラー値 $\langle \boldsymbol{a}_1|\boldsymbol{a}_2\rangle$ が得られる。

ケットという「観測対象」に対して、結果としてスカラー値を返すような関数（内積）を考えたとき、そのための「観測装置」となるのがブラである。

観測装置であるブラ（横ベクトル）は、縦ベクトルからスカラー値を得るための写像 $\mathbb{R}^n \longrightarrow \mathbb{R}$ とみることもできる。
この捉え方を一般化すると、線形汎関数という概念に結びつく。

基底によるベクトルの展開

縦ベクトルをケット、横ベクトルをブラで表すことにする。

測定値からベクトルを特定する

基底 $|\boldsymbol{a}_1\rangle, |\boldsymbol{a}_2\rangle$ を用いて、ベクトル $|\boldsymbol{x}\rangle$ を次のような線形結合で表そう。

$$|\boldsymbol{x}\rangle = x_1 |\boldsymbol{a}_1\rangle + x_2 |\boldsymbol{a}_2\rangle$$

ここで、係数 x_1, x_2 を求めるために、左からブラ $\langle \boldsymbol{a}_1|, \langle \boldsymbol{a}_2|$ を作用させる。

$$\begin{aligned} \langle \boldsymbol{a}_1|\boldsymbol{x}\rangle &= x_1 \langle \boldsymbol{a}_1|\boldsymbol{a}_1\rangle + x_2 \langle \boldsymbol{a}_1|\boldsymbol{a}_2\rangle \\ \langle \boldsymbol{a}_2|\boldsymbol{x}\rangle &= x_1 \langle \boldsymbol{a}_2|\boldsymbol{a}_1\rangle + x_2 \langle \boldsymbol{a}_2|\boldsymbol{a}_2\rangle \end{aligned}$$

ブラとケットが組み合わされたブラケット $\langle \cdot|\cdot\rangle$ はスカラー値を表しているので、これは未知数 x_1, x_2 に関する連立方程式である。
この連立方程式を解けば、係数 x_1, x_2 が求まり、ベクトル $|\boldsymbol{x}\rangle$ を特定できる。

直交基底による展開

直交基底を用いると、線形結合の係数は連立方程式を解くことなく、内積を用いて直接計算できる。

直交基底 $|\boldsymbol{u}_1\rangle$, $|\boldsymbol{u}_2\rangle$ に対して、ベクトル $|\boldsymbol{x}\rangle$ が次のように表されるとする。

$$|\boldsymbol{x}\rangle = x_1 |\boldsymbol{u}_1\rangle + x_2 |\boldsymbol{u}_2\rangle$$

左からブラ $\langle\boldsymbol{u}_1|$, $\langle\boldsymbol{u}_2|$ を作用させると、

$$\begin{aligned}\langle\boldsymbol{u}_1|\boldsymbol{x}\rangle &= x_1 \langle\boldsymbol{u}_1|\boldsymbol{u}_1\rangle + x_2 \langle\boldsymbol{u}_1|\boldsymbol{u}_2\rangle \\ \langle\boldsymbol{u}_2|\boldsymbol{x}\rangle &= x_1 \langle\boldsymbol{u}_2|\boldsymbol{u}_1\rangle + x_2 \langle\boldsymbol{u}_2|\boldsymbol{u}_2\rangle\end{aligned}$$

ここで、直交していれば内積は 0 になるので、$\langle\boldsymbol{u}_1|\boldsymbol{u}_2\rangle$ や $\langle\boldsymbol{u}_2|\boldsymbol{u}_1\rangle$ は 0 となる。

$$\begin{aligned}\langle\boldsymbol{u}_1|\boldsymbol{x}\rangle &= x_1 \langle\boldsymbol{u}_1|\boldsymbol{u}_1\rangle \\ \langle\boldsymbol{u}_2|\boldsymbol{x}\rangle &= x_2 \langle\boldsymbol{u}_2|\boldsymbol{u}_2\rangle\end{aligned}$$

この式から、係数 x_1, x_2 は次のように求まる。

$$\begin{aligned}x_1 &= \frac{\langle\boldsymbol{u}_1|\boldsymbol{x}\rangle}{\langle\boldsymbol{u}_1|\boldsymbol{u}_1\rangle} \\ x_2 &= \frac{\langle\boldsymbol{u}_2|\boldsymbol{x}\rangle}{\langle\boldsymbol{u}_2|\boldsymbol{u}_2\rangle}\end{aligned}$$

⚓ **直交基底によるベクトルの展開**　計量空間 V の直交基底 $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n$ に対して、任意のベクトル $\boldsymbol{v} \in V$ は

$$\boldsymbol{v} = \sum_{i=1}^{n} \frac{(\boldsymbol{v}, \boldsymbol{u}_i)}{(\boldsymbol{u}_i, \boldsymbol{u}_i)} \boldsymbol{u}_i$$

と表すことができる。

証明

ベクトル $\boldsymbol{v}$ が次のような線形結合

$$\boldsymbol{v} = c_1 \boldsymbol{u}_1 + \cdots + c_n \boldsymbol{u}_n$$

で表されるとし、係数を求めることを目指す。

このとき、$\boldsymbol{u}_j \ (j = 1, 2, \ldots, n)$ との内積をとると、

$$\begin{aligned}(\boldsymbol{v}, \boldsymbol{u}_j) &= (c_1\boldsymbol{u}_1 + \cdots + c_n\boldsymbol{u}_n, \boldsymbol{u}_j) \\ &= c_1(\boldsymbol{u}_1, \boldsymbol{u}_j) + \cdots + c_n(\boldsymbol{u}_n, \boldsymbol{u}_j) \\ &= \sum_{i=1}^{n} c_i(\boldsymbol{u}_i, \boldsymbol{u}_j)\end{aligned}$$

となるが、$\boldsymbol{u}_i$ は直交系であるため、$i \neq j$ のとき $(\boldsymbol{u}_i, \boldsymbol{u}_j) = 0$ である。

よって、上の式において残るのは、$i = j$ の項だけとなり、

$$(\boldsymbol{v}, \boldsymbol{u}_j) = c_j(\boldsymbol{u}_j, \boldsymbol{u}_j)$$

ここで、直交系の定義より $\boldsymbol{u}_j \neq \boldsymbol{o}$ なので、$(\boldsymbol{u}_j, \boldsymbol{u}_j) \neq 0$ である。

そこで、両辺を $(\boldsymbol{u}_j, \boldsymbol{u}_j)$ で割ることができ、

$$c_j = \frac{(\boldsymbol{v}, \boldsymbol{u}_j)}{(\boldsymbol{u}_j, \boldsymbol{u}_j)}$$

が得られる。 ■

正規直交基底による展開

さらに、$|\boldsymbol{u}_1\rangle, |\boldsymbol{u}_2\rangle$ が正規直交基底であるなら、これらのノルムは 1 であるので、

$$\langle \boldsymbol{u}_1 | \boldsymbol{u}_1 \rangle = \|\boldsymbol{u}_1\|^2 = 1, \quad \langle \boldsymbol{u}_2 | \boldsymbol{u}_2 \rangle = \|\boldsymbol{u}_2\|^2 = 1$$

よって、係数 x_1, x_2 はさらに簡単な形で表すことができる。

$$\begin{aligned}x_1 &= \langle \boldsymbol{u}_1 | \boldsymbol{x} \rangle \\ x_2 &= \langle \boldsymbol{u}_2 | \boldsymbol{x} \rangle\end{aligned}$$

正規直交基底によるベクトルの展開　計量空間 V の正規直交基底 $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n$ に対して、任意のベクトル $\boldsymbol{v} \in V$ は

$$\boldsymbol{v} = \sum_{i=1}^{n} (\boldsymbol{v}, \boldsymbol{u}_i)\boldsymbol{u}_i$$

と表すことができる。

証明

正規直交基底の場合、

$$(\boldsymbol{u}_i, \boldsymbol{u}_i) = \|\boldsymbol{u}_i\|^2 = 1$$

であることを用いると、直交基底によるベクトルの展開式において分母が 1 となり、この形が得られる。 ■

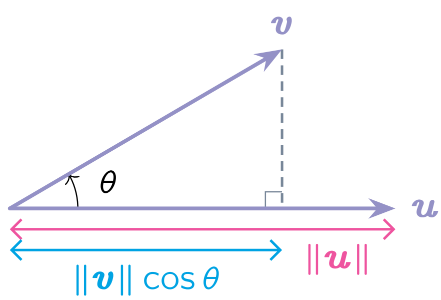
ベクトルの正射影

このように、直交基底や正規直交基底を用いることで計算が簡単になる場面が多くある。実は、任意の基底を直交基底や正規直交基底に作り変えることもできる。

正規直交系をつくるにあたって重要となる、正射影という概念を導入しよう。

ベクトルの「影」

内積の幾何学的解釈は、次のような図で内積の意味を捉えようとするものだった。



ベクトル $\boldsymbol{u}, \boldsymbol{v}$ のなす角 θ を用いると、$\boldsymbol{v}$ の $\boldsymbol{u}$ との内積は次のように表すことができる。

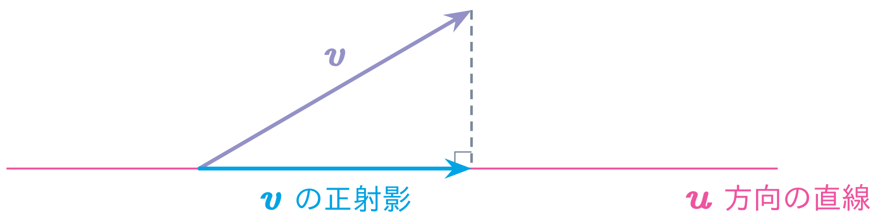
$$\langle \boldsymbol{u} | \boldsymbol{v} \rangle = \|\boldsymbol{u}\| \|\boldsymbol{v}\| \cos\theta$$

つまり、$\boldsymbol{u}$ と $\boldsymbol{v}$ の内積とは、$\boldsymbol{u}$ の長さ $\|\boldsymbol{u}\|$ と、$\boldsymbol{v}$ の $\boldsymbol{u}$ 方向の長さ $\|\boldsymbol{v}\| \cos\theta$ の積であるとも捉えられる。

厳密には、θ の値によっては $\cos\theta$ は負の数になり得るので、$\|\boldsymbol{v}\| \cos\theta$ をそのまま正射影の「長さ」と呼ぶのは適切ではないのだが…

ここで、$\boldsymbol{v}$ の $\boldsymbol{u}$ 方向の長さは、上から光を当てたときに $\boldsymbol{u}$ に投影される、「$\boldsymbol{v}$ の影」の長さのように見える。

そこで、$\boldsymbol{u}$ 方向の直線上に落とした $\boldsymbol{v}$ の影となるベクトルを、$\boldsymbol{v}$ の $\boldsymbol{u}$ への正射影と呼ぶ。

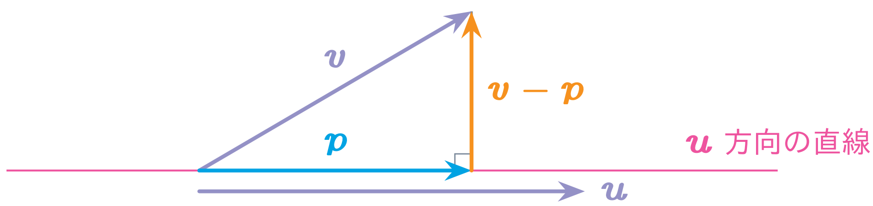


正射影の公式

なす角 θ を使わずに、$\boldsymbol{v}$ の $\boldsymbol{u}$ への正射影を表すことを考えてみよう。

$\boldsymbol{v}$ の $\boldsymbol{u}$ への正射影を $\boldsymbol{p}$ とおくと、次の関係が成り立っている。

- $\boldsymbol{p}$ は $\boldsymbol{u}$ と平行である
- $\boldsymbol{v}-\boldsymbol{p}$ は $\boldsymbol{u}$ と直交する



$\boldsymbol{p}$ は $\boldsymbol{u}$ と平行であることから、スカラー k を用いて次のように表せる。

$$\boldsymbol{p}=k\boldsymbol{u}$$

また、$\boldsymbol{v}-\boldsymbol{p}$ が $\boldsymbol{u}$ と直交することから、これらの内積は 0 になる。

$$(\boldsymbol{v}-k\boldsymbol{u},\boldsymbol{u})=0$$

内積の双線形性を用いて展開すると、

$$(\boldsymbol{v},\boldsymbol{u})-k(\boldsymbol{u},\boldsymbol{u})=0$$
$$\therefore\quad k\|\boldsymbol{u}\|^2=(\boldsymbol{v},\boldsymbol{u})$$

ここで、$\boldsymbol{u}\neq\boldsymbol{o}$ であれば、そのノルム $\|\boldsymbol{u}\|$ は 0 にはなり得ないので、$\|\boldsymbol{u}\|^2$ で両辺を割ることができる。

$$k=\frac{(\boldsymbol{v},\boldsymbol{u})}{\|\boldsymbol{u}\|^2}$$

よって、正射影ベクトル $\boldsymbol{p} = k\boldsymbol{u}$ は、次のように表すことができる。

$$\text{正射影ベクトル} = \frac{(\boldsymbol{v}, \boldsymbol{u})}{\|\boldsymbol{u}\|^2}\boldsymbol{u}$$

ここまでの話をまとめておこう。

正射影　$\boldsymbol{o}$ でないベクトル $\boldsymbol{u}$ が与えられているとき、ベクトル $\boldsymbol{v}$ に対し、

i. $\boldsymbol{p}$ が $\boldsymbol{u}$ と平行
ii. $\boldsymbol{v} - \boldsymbol{p}$ が $\boldsymbol{u}$ と直交

という条件を満たすベクトル $\boldsymbol{p}$ を、$\boldsymbol{v}$ の $\boldsymbol{u}$ への正射影という。

正射影の公式　ベクトル $\boldsymbol{v}$ の $\boldsymbol{u}$ への正射影 $\boldsymbol{p}$ は、次のように表される。

$$\boldsymbol{p} = \frac{(\boldsymbol{v}, \boldsymbol{u})}{\|\boldsymbol{u}\|^2}\boldsymbol{u}$$

正射影を測る観測装置

$\boldsymbol{v}$ の $\boldsymbol{u}$ への正射影の（符号付き）長さは、係数 k の部分で表される。

$$k = \frac{\langle \boldsymbol{u} | \boldsymbol{v} \rangle}{\|\boldsymbol{u}\|^2}$$

この長さの分だけ $\boldsymbol{u}$ をスケーリングしたものが、$\boldsymbol{v}$ の $\boldsymbol{u}$ への正射影 $\boldsymbol{p} = k\boldsymbol{u}$ となる。

ここで、$\boldsymbol{u}$ のノルムが 1 であれば、k は内積 $\langle \boldsymbol{u} | \boldsymbol{v} \rangle$ そのものになる。

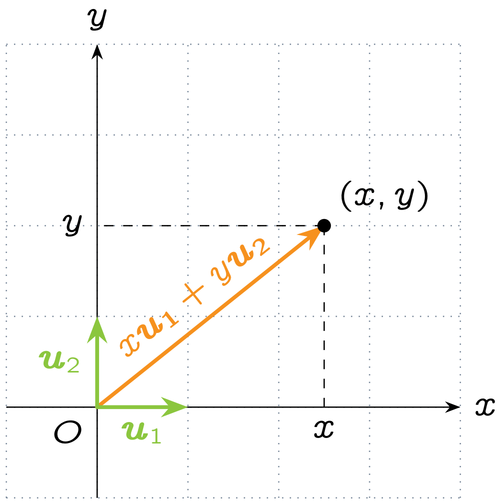
$$\langle \boldsymbol{u} | \boldsymbol{v} \rangle = \boldsymbol{u}^\top \boldsymbol{v} = k$$

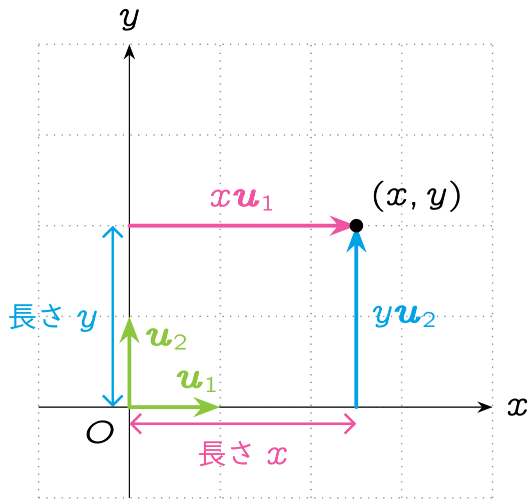
つまり、ノルムが 1 の横ベクトル $\langle \boldsymbol{u} | = \boldsymbol{u}^\top$ は、$\boldsymbol{v}$ の $\boldsymbol{u}$ への正射影の（符号つき）長さを測る観測装置として機能する。

単位ベクトル $\boldsymbol{u}$ との内積は、
「$\boldsymbol{u}$ 方向に正射影し、その長さを返す」操作になる

基底方向への正射影と座標

正規直交基底の場合、「基底方向への正射影の長さ」を測ることは、その方向の「成分（座標)」を得ることに相当する。





たとえば、ベクトル $\boldsymbol{v}$ が正規直交基底 $\boldsymbol{u}_1, \boldsymbol{u}_2$ に関して次のように表されるとする。

$$\boldsymbol{v} = \begin{pmatrix} x \\ y \end{pmatrix} = x\boldsymbol{u}_1 + y\boldsymbol{u}_2$$

このとき、各係数 x, y は、次のような正射影の長さとして得られる。

- $x\boldsymbol{u}_1$ を $\boldsymbol{u}_1$ 方向へ正射影したものの長さが x
- $y\boldsymbol{u}_2$ を $\boldsymbol{u}_2$ 方向へ正射影したものの長さが y

ここで、「正射影を測る観測装置」で述べた、

単位ベクトルとの内積は、その方向に正射影したときの長さを与える

という解釈と合わせると、次のように各座標（成分）を得ることができる。

$$x = \langle \boldsymbol{u}_1 | \boldsymbol{v} \rangle, \quad y = \langle \boldsymbol{u}_2 | \boldsymbol{v} \rangle$$

直交化と正規化

ベクトル $\boldsymbol{a}$ が、互いに直交するベクトル $\boldsymbol{u}_1, \boldsymbol{u}_2, \widehat{\boldsymbol{u}}_3$ を用いて、次のように書けるとする。

$$\boldsymbol{a} = x_1\boldsymbol{u}_1 + x_2\boldsymbol{u}_2 + \widehat{\boldsymbol{u}}_3$$

ここで、$\boldsymbol{u}_1$ と $\boldsymbol{u}_2$ のノルムは 1 だが、$\widehat{\boldsymbol{u}}_3$ はそうではない。

$\widehat{\boldsymbol{u}}_3$ は、ノルムが 1 のベクトル $\boldsymbol{u}_3$ を用いると、$\widehat{\boldsymbol{u}}_3 = x_3\boldsymbol{u}_3$ となるようなベクトルであるとする。

このとき、$\boldsymbol{u}_1, \boldsymbol{u}_2$ は単位ベクトルであるから、$\boldsymbol{a}$ との内積でそれぞれの係数を得ることができる。

$$\boldsymbol{a} = \langle \boldsymbol{u}_1|\boldsymbol{a}\rangle\, \boldsymbol{u}_1 + \langle \boldsymbol{u}_2|\boldsymbol{a}\rangle\, \boldsymbol{u}_2 + \widehat{\boldsymbol{u}}_3$$

この式を変形すると、$\widehat{\boldsymbol{u}}_3$ は次のように表せる。

$$\widehat{\boldsymbol{u}}_3 = \boldsymbol{a} - \langle \boldsymbol{u}_1|\boldsymbol{a}\rangle\, \boldsymbol{u}_1 - \langle \boldsymbol{u}_2|\boldsymbol{a}\rangle\, \boldsymbol{u}_2$$

ベクトル $\boldsymbol{a}$ を構成するときに、$\boldsymbol{u}_1, \boldsymbol{u}_2$ に直交するような成分が $\widehat{\boldsymbol{u}}_3$ である。

これを求めるためには、$\boldsymbol{a}$ から $\widehat{\boldsymbol{u}}_3$ 以外の方向 $\boldsymbol{u}_1, \boldsymbol{u}_2$ への正射影を引けばよい。

互いに直交するベクトルは、正射影を利用することで連鎖的に得ることができる。

実際、$\boldsymbol{u}_1$ と $\boldsymbol{u}_2$ を直交化しておけば、これらへの正射影を引くという形で、さらにこれらに直交するベクトル $\widehat{\boldsymbol{u}}_3$ をつくることができる。

なお、ベクトルのノルムを 1 にすることを正規化という。

$\boldsymbol{u}_1$ と $\boldsymbol{u}_2$ が正規化されていれば、これらへの正射影は内積だけで簡単に求まる。

グラム・シュミットの直交化法

正規化と直交化を連鎖させることで正規直交系をつくろうとするのが、グラム・シュミットの直交化法である。

具体的には、次の手順を繰り返すことで、計量空間 V の線型独立なベクトル $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ から正規直交系 $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ をつくることができる。

1. ノルムが 1 のベクトルをつくる

2. ノルムが 1 のベクトルとの内積で、正射影をつくる
3. 元のベクトルから正射影を引くことで、直交ベクトルをつくる

正規化

まずは、$\boldsymbol{a}_1$ から、ノルムが 1 であるベクトルをつくる。

ベクトルのノルムを 1 にすることを正規化という。

$\boldsymbol{a}_1$ を正規化したベクトルは、次のように求められる。

$$\boldsymbol{e}_1 = \frac{\boldsymbol{a}_1}{\|\boldsymbol{a}_1\|}$$

ここで、$\boldsymbol{e}_1$ は $\boldsymbol{a}_1$ をスカラー $\dfrac{1}{\|\boldsymbol{a}_1\|}$ 倍しただけなので、$\boldsymbol{e}_1$ と $\boldsymbol{a}_1$ は平行である。

正射影による直交化

次に、$\boldsymbol{e}_1$ と直交するような $\boldsymbol{u}_2$ をつくる。

そのために、$\boldsymbol{a}_2$ から $\boldsymbol{a}_2$ の $\boldsymbol{e}_1$ への正射影を引いたものは、$\boldsymbol{e}_1$ と直交することを利用する。

$\boldsymbol{a}_2$ の $\boldsymbol{e}_1$ への正射影は、次のように計算できる。

$$(\boldsymbol{a}_2, \boldsymbol{e}_1)\boldsymbol{e}_1$$

そこで、

$$\boldsymbol{u}_2 = \boldsymbol{a}_2 - (\boldsymbol{a}_2, \boldsymbol{e}_1)\boldsymbol{e}_1$$

とおくと、$\boldsymbol{u}_2$ は $\boldsymbol{e}_1$ と直交する。

再び正規化

ここで、もし $\boldsymbol{u}_2 = \boldsymbol{o}$ ならば、$\boldsymbol{a}_2$ は $\boldsymbol{a}_1$ の線形結合で表されることになり、$\boldsymbol{a}_1, \boldsymbol{a}_2$ は線型従属になってしまう。

$$\boldsymbol{a}_2 = (\boldsymbol{a}_2, \boldsymbol{e}_1)\boldsymbol{e}_1 = (\boldsymbol{a}_2, \boldsymbol{e}_1)\frac{1}{\|\boldsymbol{a}_1\|}\boldsymbol{a}_1$$

ここでは、$\boldsymbol{a}_1, \boldsymbol{a}_2$ は線型独立なベクトルと仮定しているので、$\boldsymbol{u}_2 \neq \boldsymbol{o}$ である。

よって、$\boldsymbol{u}_2$ を次のように正規化することができる。

$$\boldsymbol{e}_2 = \frac{\boldsymbol{u}_2}{\|\boldsymbol{u}_2\|}$$

$\boldsymbol{u}_2$ は $\boldsymbol{e}_1$ と直交するベクトルであるので、$\boldsymbol{e}_2$ も $\boldsymbol{e}_1$ と直交する。

さらに、$\boldsymbol{e}_2$ はノルムが 1 のベクトルになっている。

そこで、前工程「正射影への直交化」に戻って、今度は $\boldsymbol{e}_2$ と直交するようなベクトル $\boldsymbol{u}_3$ をつくることができる。

正規化と直交化の連鎖

以上の手順を繰り返すことで、線型独立なベクトル $\boldsymbol{a}_1, \dots, \boldsymbol{a}_n$ から、正規直交系 $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n$ を得ることができる。

このような方法をグラム・シュミットの直交化法という。

⚓ **グラム・シュミットの直交化法**　計量空間 V の線型独立なベクトル $\boldsymbol{a}_1, \dots, \boldsymbol{a}_n$ から、正規直交系 $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n$ を次のように構成できる。

$$\boldsymbol{u}_k = \boldsymbol{a}_k - \sum_{j=1}^{k-1} (\boldsymbol{a}_k, \boldsymbol{e}_j)\boldsymbol{e}_j$$

$$\boldsymbol{e}_k = \frac{\boldsymbol{u}_k}{\|\boldsymbol{u}_k\|}$$

ここで、$k = 1, 2, \dots, n$ である。

正規直交基底の構成

次の定理により、グラム・シュミットの直交化法は、線型独立なベクトルから正規直交系を得るだけでなく、任意の基底から正規直交基底を得る手法としても利用できる。

⚓ **グラム・シュミットの直交化と生成空間の不変性**　計量空間 V の線型独立なベクトル $\boldsymbol{a}_1, \dots, \boldsymbol{a}_n$ から、グラム・シュミットの直交化法を用いて得られた正規直交系を $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n$ とする。

このとき、$\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ が張る空間と $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ が張る空間は一致する。

$$\langle \boldsymbol{a}_1, \ldots, \boldsymbol{a}_n \rangle = \langle \boldsymbol{e}_1, \ldots, \boldsymbol{e}_n \rangle$$

証明

グラム・シュミットの直交化法では、各ステップ k において、まず $\boldsymbol{a}_k$ からその前に得られた直交ベクトル $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_{k-1}$ への射影を引くことで、$\boldsymbol{a}_k$ に直交するベクトルを構成する。

すなわち、

$$\boldsymbol{u}_k = \boldsymbol{a}_k - \sum_{j=1}^{k-1} (\boldsymbol{a}_k, \boldsymbol{e}_j)\boldsymbol{e}_j$$

と定義し、その後これを正規化して $\boldsymbol{e}_k$ とする。

ここで、$\boldsymbol{u}_k$ は右辺の形から明らかなように、$\boldsymbol{a}_k$ と $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_{k-1}$ の線型結合である。

そしてさらに各 $\boldsymbol{e}_j$（$j < k$）は、それ以前の $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_j$ の線型結合であることから、$\boldsymbol{u}_k$ は結局 $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_k$ の線型結合として書ける。

したがって、$\boldsymbol{e}_k$ も $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_k$ の線型結合となり、$\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ はすべて $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ の線型結合である。

よって、すべての $\boldsymbol{e}_k$ は $\langle \boldsymbol{a}_1, \ldots, \boldsymbol{a}_n \rangle$ に属することになり、

$$\langle \boldsymbol{e}_1, \ldots, \boldsymbol{e}_n \rangle \subset \langle \boldsymbol{a}_1, \ldots, \boldsymbol{a}_n \rangle$$

が成り立つ。

両辺の部分空間の次元を考えると、$\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ が線型独立であるため、$\langle \boldsymbol{a}_1, \ldots, \boldsymbol{a}_n \rangle$ の次元は n である。

一方、$\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ も直交系であることから線型独立であるため、$\langle \boldsymbol{e}_1, \ldots, \boldsymbol{e}_n \rangle$ の次元も n である。

よって、部分空間の次元が等しいことから、両者は一致する。 ■

正規直交基底の存在

このように、グラム・シュミットの直交化法によって、任意の基底から正規直交基底をつくることができる。

つまり、グラム・シュミットの直交化法は、内積が定められている空間（計量空間）には正規直交基底が必ず存在することを示している。

⚓ **正規直交基底の存在**　$\{\boldsymbol{o}\}$ でない任意の計量空間は正規直交基底を持つ。

線形従属なベクトルの正規直交化

与えられたベクトルが線型独立でない場合にグラム・シュミットの直交化法を適用すると、いずれ射影を引いたベクトルが $\boldsymbol{o}$ になる。

ある $\boldsymbol{a}_k$ が、前のベクトルたち $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_{k-1}$ の線形結合として表される、すなわち線形従属であるとする。

ここで、$\boldsymbol{a}_1, \ldots, \boldsymbol{a}_{k-1}$ から得られた正規直交系を $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_{k-1}$ とすると、これらの張る空間は一致する。

$$\langle \boldsymbol{a}_1, \ldots, \boldsymbol{a}_{k-1} \rangle = \langle \boldsymbol{e}_1, \ldots, \boldsymbol{e}_{k-1} \rangle$$

よって、$\boldsymbol{a}_k$ は、すでに得られた正規直交系 $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_{k-1}$ の線形結合として表すこともできる。

$$\boldsymbol{a}_k = \sum_{i=1}^{k-1} (\boldsymbol{a}_k, \boldsymbol{e}_i) \boldsymbol{e}_i$$

そのため、射影をすべて引くと、次のように残りが $\boldsymbol{o}$ になってしまう。

$$\boldsymbol{u}_k = \boldsymbol{a}_k - \sum_{i=1}^{k-1} (\boldsymbol{a}_k, \boldsymbol{e}_i) \boldsymbol{e}_i = \boldsymbol{o}$$

このように、グラム・シュミットの直交化法における射影を引く操作は、すでにある正規直交基底に重なっている成分（従属部分）を消し去ってしまう。

この性質により、グラム・シュミットの直交化法は線形従属な場合でも破綻せずに使える。

しかし、結果として新しい成分がゼロになる（つまり新しい情報がない）ため、得られる直交系は完全な基底にはならない。

計量同型

[Todo 40:]

第 19 章

複素行列と対角化

転置行列と随伴行列

複素正方行列 A の転置行列において、各成分をその共役複素数に置き換えた行列を随伴行列という

随伴行列　複素正方行列 $A = (a_{ij})$ に対し、$\overline{a_{ji}}$ を (i, j) 成分にもつ行列 ${}^t\overline{A}$ を A の随伴行列といい、A^* と表す

実数 x の複素共役は $\overline{x} = x$ であるので、A が実行列のときは、

$$A^* = {}^tA$$

すなわち、

実行列の世界では、随伴行列は転置行列

にすぎない

転置と似た性質

転置を二回行うと元に戻ることと同様に、次が成り立つ

⚓ **随伴行列の自己反転性**　複素正方行列 A に対し、随伴行列を二回とると元に戻る

$$(A^*)^* = A$$

証明

随伴行列の定義より、

$$(A^*)^* = {}^t\overline{A^*} = {}^t\overline{{}^t\overline{A}}$$

$A = (a_{ij})$ とすると、A の各成分を共役複素数にした行列は、

$$\overline{A} = (\overline{a_{ij}})$$

これを転置すると、

$${}^t\overline{A} = (\overline{a_{ji}})$$

さらに、もう一度各成分の複素共役をとると、

$$\overline{{}^t\overline{A}} = (\overline{\overline{a_{ji}}}) = (a_{ji})$$

したがって、

$$(A^*)^* = {}^t\overline{{}^t\overline{A}} = (a_{ij}) = A$$

が成り立つ　■

⚓ **積に対するエルミート共役の順序反転性**　複素行列 A B の積 AB が定義できるとき、

$$(AB)^* = B^*A^*$$

証明

[Todo 41:]

随伴による内積の表現

標準内積は、随伴を用いて表現することもできる

⚓ 随伴による標準内積の表現

$$(\boldsymbol{a}, \boldsymbol{b}) = \boldsymbol{b}^* \cdot \boldsymbol{a}$$

証明

標準内積の対称性より、

$$(\boldsymbol{a}, \boldsymbol{b}) = \overline{(\boldsymbol{b}, \boldsymbol{a})}$$

ここで、右辺の内積を転置を用いて表すと、

$$\overline{(\boldsymbol{b}, \boldsymbol{a})} = \overline{\boldsymbol{b}^\top \cdot \overline{\boldsymbol{a}}} = \overline{\boldsymbol{b}^\top} \cdot \overline{\overline{\boldsymbol{a}}} = \boldsymbol{b}^* \cdot \boldsymbol{a}$$

よって、

$$(\boldsymbol{a}, \boldsymbol{b}) = \overline{(\boldsymbol{b}, \boldsymbol{a})} = \boldsymbol{b}^* \cdot \boldsymbol{a}$$

が成り立つ ■

随伴公式

随伴行列と標準内積は、次のような関係で結ばれる

⚓ 随伴公式　複素行列 A と計量空間上のベクトル $\boldsymbol{u}$, $\boldsymbol{v}$ に対し、

$$(A\boldsymbol{u}, \boldsymbol{v}) = (\boldsymbol{u}, A^*\boldsymbol{v})$$

証明

転置を用いて内積を表すと、

$$(A\boldsymbol{u}, \boldsymbol{v}) = {}^t(A\boldsymbol{u})\overline{\boldsymbol{v}}$$

転置と行列積の順序反転性より、${}^t(A\boldsymbol{u}) = {}^t\boldsymbol{u}{}^tA$ なので、

$$(A\boldsymbol{u}, \boldsymbol{v}) = ({}^t\boldsymbol{u}{}^tA)\overline{\boldsymbol{v}}$$

行列の積の結合法則を用いて、

$$(A\boldsymbol{u}, \boldsymbol{v}) = {}^t\boldsymbol{u}({}^tA\overline{\boldsymbol{v}})$$

ここで、$\overline{{}^t\overline{A}}$ は、$A = (a_{ij})$ とすると、

1. $\overline{A} = (\overline{a_{ij}})$
2. ${}^t\overline{A} = (\overline{a_{ji}})$
3. $\overline{{}^t\overline{A}} = (\overline{\overline{a_{ji}}}) = (a_{ji}) = {}^tA$

となり、${}^tA$ と一致する

これを用いて書き換えると、

$$(A\boldsymbol{u}, \boldsymbol{v}) = {}^t\boldsymbol{u}(\overline{{}^t\overline{A}}\overline{\boldsymbol{v}})$$

複素共役の積の性質 $\overline{z_1} \cdot \overline{z_2} = \overline{z_1 z_2}$ を用いて、

$$(A\boldsymbol{u}, \boldsymbol{v}) = {}^t\boldsymbol{u}\overline{{}^t\overline{A}\boldsymbol{v}}$$

この時点で、右辺を内積として書き直すと、$A\boldsymbol{v}$ の複素共役がなくなることに注意して、

$$(A\boldsymbol{u}, \boldsymbol{v}) = (\boldsymbol{u}, {}^t\overline{A}\boldsymbol{v})$$

随伴行列の定義 $A^* = {}^t\overline{A}$ より、

$$(A\boldsymbol{u}, \boldsymbol{v}) = (\boldsymbol{u}, A^*\boldsymbol{v})$$

となり、目的の等式が得られた ■

ユニタリ行列と直交行列

ユニタリ行列　複素正方行列 A が次を満たすとき、A をユニタリ行列という

$$A^* = A^{-1}$$

ユニタリ行列と内積

2 つのベクトルそれぞれにユニタリ行列を左からかけても、それらの内積は変わらない

ユニタリ行列の特徴づけとしての内積不変性　n 次複素行列 A がユニタリ行列であることと、任意の $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{C}^n$ に対し、

$$(A\boldsymbol{u}, A\boldsymbol{v}) = (\boldsymbol{u}, \boldsymbol{v})$$

が成り立つことは同値である

証明

ユニタリ行列ならば内積を保つ

随伴公式より、

$$(A\boldsymbol{u}, A\boldsymbol{v}) = (\boldsymbol{u}, A^*A\boldsymbol{v})$$

ここで、A がユニタリ行列であることは、

$$A^*A = E$$

と言い換えられるので、これを用いると、

$$(A\boldsymbol{u}, A\boldsymbol{v}) = (\boldsymbol{u}, \boldsymbol{v})$$

が成り立つ　■

内積を保つならばユニタリ行列

転置を用いて内積を表すと、

$$(A\boldsymbol{u}, A\boldsymbol{v}) = {}^t(A\boldsymbol{u})\overline{(A\boldsymbol{v})}$$
$$(\boldsymbol{u}, \boldsymbol{v}) = {}^t\boldsymbol{u}\overline{\boldsymbol{v}}$$

これらが一致するというのが仮定なので、

$$ {}^t(A\boldsymbol{u})\overline{(A\boldsymbol{v})} = {}^t\boldsymbol{u}\overline{\boldsymbol{v}}$$

この関係を用いて、行列 ${}^tA\overline{A}$ の (i, j) 成分を考えると、

$$\begin{aligned} {}^t(A\boldsymbol{e}_i)\overline{(A\boldsymbol{e}_j)} &= {}^t\boldsymbol{e}_i\overline{\boldsymbol{e}_j} \\ &= \delta_{ij} \end{aligned}$$

となり、これはすなわち、

$$ {}^tA\overline{A} = E$$

よって、両辺の複素共役をとることで、

$$A^*A = E$$

を得る

したがって、A はユニタリ行列である ■

この定理において、$\boldsymbol{u} = \boldsymbol{v}$ の場合を考えると、ユニタリ行列とノルムに関する性質が導かれる

ユニタリ行列とノルム

ユニタリ行列を左からかけても、ベクトルのノルムは変わらない

⚓ **ユニタリ行列の特徴づけとしてのノルム不変性** n 次複素行列 A がユニタリ行列であることと、任意の $\boldsymbol{v} \in \mathbb{C}^n$ に対し、

$$\|A\boldsymbol{v}\| = \|\boldsymbol{v}\|$$

が成り立つことは同値である

証明

A がユニタリ行列であることと、任意の $\boldsymbol{u}$, $\boldsymbol{v} \in \mathbb{C}^n$ に対し、

$$(A\boldsymbol{u}, A\boldsymbol{v}) = (\boldsymbol{u}, \boldsymbol{v})$$

が成り立つことは同値であった

ここで、$\boldsymbol{u} = \boldsymbol{v}$ とすると、

$$(A\boldsymbol{v}, A\boldsymbol{v}) = (\boldsymbol{v}, \boldsymbol{v})$$

が成り立つことになり、ノルムの定義より、

$$\|A\boldsymbol{v}\|^2 = \|\boldsymbol{v}\|^2$$

すなわち、

$$\|A\boldsymbol{v}\| = \|\boldsymbol{v}\|$$

がしたがう ■

ユニタリ行列と直交性

A が実正方行列のときは、

$$A \text{ がユニタリ行列} \iff {}^tA = A^{-1}$$

となり、このような A は直交行列と呼ばれる

直交行列　実正方行列 A が次を満たすとき、A を直交行列という

$${}^tA = A^{-1}$$

直交行列という名前の由来は、次のように考えられる

A を n 個の列ベクトルを横一列に並べたものとみなし、

$$A = (\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n)$$

とおくと、${}^tA = A^{-1}$、すなわち ${}^tAA = E$ は、

$$\begin{pmatrix} {}^t\boldsymbol{a}_1 \\ {}^t\boldsymbol{a}_2 \\ \vdots \\ {}^t\boldsymbol{a}_n \end{pmatrix} (\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_n) = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

と表される

これは、ベクトル $\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_n$ が、次の性質

$${}^t\boldsymbol{a}_i\boldsymbol{a}_j = (\boldsymbol{a}_i, \boldsymbol{a}_j) = \delta_{ij}$$

を満たすことを意味する

すなわち、直交行列 A の列ベクトル $\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_n$ は、互いに直交する単位ベクトルである

この事実は、複素行列に対しても成立する

⚓ **ユニタリ行列の列ベクトルの直交正規性**　複素正方行列 A を $A = (\boldsymbol{a}_1, \dots, \boldsymbol{a}_n)$ と列ベクトル分解するとき、

$$A \text{ がユニタリ行列} \iff (\boldsymbol{a}_i, \boldsymbol{a}_j) = \delta_{ij}$$

すなわち、ユニタリ行列の列ベクトルは、互いに直交する単位ベクトルである

証明

A がユニタリ行列であることは、任意の $\boldsymbol{u},\ \boldsymbol{v} \in \mathbb{C}^n$ に対し、

$$(A\boldsymbol{u}, A\boldsymbol{v}) = (\boldsymbol{u}, \boldsymbol{v})$$

が成り立つことと同値であった

ここで、$\boldsymbol{u} = \boldsymbol{e}_i,\ \boldsymbol{v} = \boldsymbol{e}_j$ とすると、

$$(A\boldsymbol{e}_i, A\boldsymbol{e}_j) = (\boldsymbol{e}_i, \boldsymbol{e}_j)$$

が成り立つことになる

左辺の $A\boldsymbol{e}_i$ について考えると、

$$
\begin{aligned}
A\boldsymbol{e}_i &= (\boldsymbol{a}_1, \ldots, \boldsymbol{a}_i, \ldots, \boldsymbol{a}_n)\begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \begin{matrix} \\ \\ \leftarrow i \\ \\ \\ \end{matrix} \\
&= \boldsymbol{a}_i
\end{aligned}
$$

$A\boldsymbol{e}_j$ についても同様なので、

$$
\begin{aligned}
(A\boldsymbol{e}_i, A\boldsymbol{e}_j) = (\boldsymbol{a}_i, \boldsymbol{a}_j) = (\boldsymbol{e}_i, \boldsymbol{e}_j) = \delta_{ij} \\
\therefore \quad (\boldsymbol{a}_i, \boldsymbol{a}_j) = \delta_{ij}
\end{aligned}
$$

となり、A がユニタリ行列であることは、$(\boldsymbol{a}_i, \boldsymbol{a}_j) = \delta_{ij}$ へと同値変形できる ■

ユニタリ行列と随伴・転置

⚓ **ユニタリ行列の随伴不変性**　ユニタリ行列 U の随伴行列 U^* もユニタリ行列である

証明

随伴行列を二回とると元に戻るので、

$$
(U^*)^* = U
$$

また、ユニタリ行列の定義より、

$$
U^* = U^{-1}
$$

したがって、

$$
\begin{aligned}
(U^*)^* &= U \\
U &= (U^*)^{-1}
\end{aligned}
$$

すなわち、

$$
U^* = (U^*)^{-1}
$$

となるので、U^* もユニタリ行列である ■

上の定理は、実行列の世界では、次の定理に対応する

直交行列の転置不変性　直交行列 Q の転置行列 ${}^tQ$ も直交行列である

エルミート行列と対称行列

エルミート行列　複素正方行列 A が次を満たすとき、A をエルミート行列という

$$A^* = A$$

A が実正方行列のときは、

$$A\ \text{がエルミート行列} \iff {}^tA = A$$

となり、このような A は対称行列、あるいは実対称行列と呼ばれる

エルミート行列の固有値

行列の成分が実数であっても、特性方程式の根は一般には実数とは限らない

つまり、固有値は一般には複素数であるが、エルミート行列については次が成り立つ

エルミート行列の固有値の実数性　エルミート行列の固有値はすべて実数である

証明

エルミート行列 A の固有ベクトルを $\boldsymbol{v}$ とし、その固有値を $\alpha \in \mathbb{C}^n$ とすると、

$$A\boldsymbol{v} = \alpha\boldsymbol{v}$$

より、次が成り立つ

$$\begin{aligned}(A\boldsymbol{v}, \boldsymbol{v}) &= (\alpha\boldsymbol{v}, \boldsymbol{v}) \\ &= \alpha(\boldsymbol{v}, \boldsymbol{v})\end{aligned}$$

一方、随伴公式から、次のようにも書ける

$$(A\boldsymbol{v}, \boldsymbol{v}) = (\boldsymbol{v}, A^*\boldsymbol{v})$$

A がエルミート行列であることから、$A^* = A$ なので、

$$\begin{aligned}(A\boldsymbol{v}, \boldsymbol{v}) &= (\boldsymbol{v}, A\boldsymbol{v}) \\ &= (\boldsymbol{v}, \alpha\boldsymbol{v})\end{aligned}$$

内積の共役線形性に注意して、

$$(A\boldsymbol{v}, \boldsymbol{v}) = \overline{\alpha}(\boldsymbol{v}, \boldsymbol{v})$$

ここまでで得られた $(A\boldsymbol{v}, \boldsymbol{v})$ の 2 通りの表現をまとめると、

$$\alpha(\boldsymbol{v}, \boldsymbol{v}) = \overline{\alpha}(\boldsymbol{v}, \boldsymbol{v})$$

移項して、

$$(\alpha - \overline{\alpha})(\boldsymbol{v}, \boldsymbol{v}) = 0$$

ここで、$\boldsymbol{v}$ は固有ベクトルなので、$\boldsymbol{v} \neq \boldsymbol{0}$ である

よって、$(\boldsymbol{v}, \boldsymbol{v}) \neq 0$ で両辺を割ることができ、次を得る

$$\alpha = \overline{\alpha}$$

すなわち、α は実数である ■

エルミート行列では、固有値が実数であることがうまく活きて、次の性質も成り立つ

エルミート行列の固有値の直交性　エルミート行列の相異なる固有値を持つ固有ベクトルは直交する

すなわち、エルミート行列 A の固有ベクトル $\boldsymbol{u}$, $\boldsymbol{v}$ がそれぞれ固有値 α, $\beta \in \mathbb{R}^n$ を持つとし、$\alpha \neq \beta$ ならば、

$$(\boldsymbol{u}, \boldsymbol{v}) = 0$$

が成り立つ

証明

固有値と固有ベクトルの定義より、

$$A\boldsymbol{u} = \alpha\boldsymbol{u}$$
$$A\boldsymbol{v} = \beta\boldsymbol{v}$$

が成り立つ

一方、随伴公式より、

$$(A\boldsymbol{u}, \boldsymbol{v}) = (\boldsymbol{u}, A^*\boldsymbol{v})$$

であるが、A はエルミート行列なので、$A^* = A$ が成り立つ

$$(A\boldsymbol{u}, \boldsymbol{v}) = (\boldsymbol{u}, A\boldsymbol{v})$$

先ほどの固有値と固有ベクトルの関係を代入して、

$$(\alpha\boldsymbol{u}, \boldsymbol{v}) = (\boldsymbol{u}, \beta\boldsymbol{v})$$

ここで、α, β は実数なので、内積の共役線形性を考慮しても、

$$\alpha(\boldsymbol{u}, \boldsymbol{v}) = \beta(\boldsymbol{u}, \boldsymbol{v})$$

として、スカラーをそのまま外に出すことができる

よって、

$$(\alpha - \beta)(\boldsymbol{u}, \boldsymbol{v}) = 0$$

であるが、$\alpha \neq \beta$ なので、$(\alpha - \beta) \neq 0$ で両辺を割ることができ、

$$(\boldsymbol{u}, \boldsymbol{v}) = 0$$

を得る ■

エルミート行列の対角化に向けた考察

H を n 次エルミート行列とすると、その固有値は n 個の実数として $\alpha_1, \dots, \alpha_n$ とおける

そして、α_i に属する固有ベクトル $\boldsymbol{v}_i$ をとると、$\boldsymbol{v}_1, \dots, \boldsymbol{v}_n$ はどの 2 つも互いに直交する

そこで、それぞれを次のように正規化する

$$\boldsymbol{u}_i = \frac{\boldsymbol{v}_i}{\|\boldsymbol{v}_i\|} \quad (i = 1, \dots, n)$$

すると、$\boldsymbol{u}_1, \dots, \boldsymbol{u}_n$ は互いに直交する単位ベクトルであるので、

$$U = (\boldsymbol{u}_1, \dots, \boldsymbol{u}_n)$$

とおけば、U はユニタリ行列となる

$\boldsymbol{u}_i$ は H の各固有ベクトル $\boldsymbol{v}_i$ をスカラー倍したものなので、

$$H\boldsymbol{u}_i = \alpha_i \boldsymbol{u}_i$$

という関係が成り立つ

つまり、U の列ベクトル $\boldsymbol{u}_1, \dots, \boldsymbol{u}_n$ はそれぞれ H の固有値 $\alpha_1, \dots, \alpha_n$ に属する固有ベクトルである

さらに、ユニタリ行列はその定義から明らかに正則行列であるので、対角化行列の列ベクトルと固有ベクトルの対応を振り返ると、

> エルミート行列はユニタリ行列を用いて対角化できる

という「予感」がしてくる

まだ「予感」としかいえないのは、エルミート行列の固有値 $\alpha_1, \dots, \alpha_n$ が重複している可能性があるからである

正規行列

エルミート行列の対角化について議論するために、エルミート行列・ユニタリ行列を含むより包括的な概念として正規行列を導入する

正規行列　複素正方行列 A が次を満たすとき、A を正規行列という

$$AA^* = A^*A$$

正規行列の例

A をエルミート行列とすると、$A^* = A$ なので、

$$AA^* = A^2$$
$$A^*A = A^2$$

となり、正規行列の定義を満たす

エルミート行列の正規行列性　エルミート行列は正規行列である

また、A をユニタリ行列とすると、$A^* = A^{-1}$ なので、

$$AA^* = AA^{-1} = E$$
$$A^*A = A^{-1}A = E$$

となり、こちらも正規行列の定義を満たす

ユニタリ行列の正規行列性　ユニタリ行列は正規行列である

正規行列の性質

正規行列と随伴によるノルム保存性　複素正方行列 A が正規行列であることは、任意の $\boldsymbol{v} \in \mathbb{C}^n$ に対し、

$$\|A\boldsymbol{v}\| = \|A^*\boldsymbol{v}\|$$

が成り立つことと同値である

証明

[Todo 42: book: 行列と行列式の基礎 p262 問 6.9 (1)]

正規行列における固有ベクトルの随伴対応　A を正規行列とするとき、$\boldsymbol{v}$ が A の固有値 α の固有ベクトルならば、$\boldsymbol{v}$ は A^* の固有値 $\overline{\alpha}$ の固有ベクトルである

すなわち、

$$A\boldsymbol{v} = \alpha\boldsymbol{v} \Longrightarrow A^*\boldsymbol{v} = \overline{\alpha}\boldsymbol{v}$$

証明

[Todo 43: book: 行列と行列式の基礎 p262 問 6.9 (2)]

正規行列の対角化

A の固有値 α に属する線型独立な固有ベクトルがちょうど k 個存在することは、

$$\dim\{\boldsymbol{x} \mid A\boldsymbol{x} = \alpha\boldsymbol{x}\} = k$$

と表せる

これは、固有値 α の固有空間の次元が k であること、噛み砕くと、固有値 α の固有ベクトル $\boldsymbol{x}$ の集合が部分空間であり、k 個の固有ベクトルがこの部分空間の基底を成す（線型独立である）ことを意味する

固有空間は核空間 $\mathrm{Ker}(A-\alpha E)$ と定義されるため、この次元が k であることは、次のようにも書ける

$$\dim \mathrm{Ker}(A-\alpha E)=k$$

正規行列について、一般に次が成り立つ

正規行列における固有空間の次元と固有値の重複度の一致　n 次複素正方行列 A が正規行列であるとき、$\Phi_A(x)$ における固有値 α の重複度 k について、次の等式が成り立つ

$$k=n-\mathrm{rank}(A-\alpha E)$$

次元定理を用いて言い換えると、α の固有空間 $W(\alpha)$ について、

$$\dim W(\alpha)=k$$

が成り立つ

証明

$l=n-\mathrm{rank}(A-\alpha E)$ とおく（l が重複度 k に等しいことを示すことが目標）

すなわち、

$$\mathrm{rank}(A-\alpha E)=n-l$$

であると仮定する

また、固有値 α の固有ベクトルは、斉次形方程式

$$(A-\alpha E)\boldsymbol{x}=\boldsymbol{0}$$

の非自明解である

この方程式の解空間は $\mathrm{Ker}(A-\alpha E)$ であるが、次元定理より、

$$\dim \mathrm{Ker}(A-\alpha E)=n-\mathrm{rank}(A-\alpha E)=l$$

であるので、$\mathrm{Ker}(A - \alpha E)$ は次元 l の部分空間である

すなわち、方程式 $(A - \alpha E)\boldsymbol{x} = \boldsymbol{0}$ を満たす l 個の線型独立なベクトルが存在する

これらを $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_l$ とすると、これらはすべて固有値 α の固有ベクトルである

これらが正規直交系でない場合は、グラム・シュミットの直交化法を用いて正規直交系に変換し、それを改めて $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_l$ とする

次に、これら l 個のベクトルを補う形で、正規直交基底 $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_l, \boldsymbol{v}_{l+1}, \ldots, \boldsymbol{v}_n$ を作る

これらを用いて、行列 U を

$$U = (\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_l, \boldsymbol{v}_{l+1}, \ldots, \boldsymbol{v}_n)$$

とおくと、U はユニタリ行列である

さらに、$\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_l$ は A の固有値 α に属する固有ベクトルであることから、

$$U^{-1}AU = \begin{array}{cc} \begin{array}{cc} \qquad l \qquad & \qquad n-l \qquad \end{array} & \\ \left(\begin{array}{ccc|c} \alpha & & & \\ & \ddots & & B \\ & & \alpha & \\ \hline & O & & C \end{array}\right) & \begin{array}{l} l \\ \\ n-l \end{array} \end{array}$$

ユニタリ行列 U の定義より、$U^{-1} = U^*$ が成り立つので、

$$U^*AU = \begin{array}{cc} \begin{array}{cc} \qquad l \qquad & \qquad n-l \qquad \end{array} & \\ \left(\begin{array}{ccc|c} \alpha & & & \\ & \ddots & & B \\ & & \alpha & \\ \hline & O & & C \end{array}\right) & \begin{array}{l} l \\ \\ n-l \end{array} \end{array}$$

ここで、両辺の随伴行列をつくることを考える

左辺は、積の随伴行列をつくると積の順序が入れ替わることに注意して、

$$(U^*AU)^* = U^*A^*(U^*)^* = U^*A^*U$$

右辺は、転置してから各成分を共役複素数に置き換えればよいので、

$$U^*A^*U = \begin{array}{cc} \begin{array}{cc} \overset{l}{\longleftrightarrow} & \overset{n-l}{\longleftrightarrow} \end{array} & \\ \left(\begin{array}{ccc|c} \overline{\alpha} & & & \\ & \ddots & & O \\ & & \overline{\alpha} & \\ \hline & B^* & & C^* \end{array}\right) & \begin{array}{l} \updownarrow l \\ \updownarrow n-l \end{array} \end{array}$$

一方、A が正規行列であることから、$\boldsymbol{v}_1, \ldots, \boldsymbol{v}_l$ は、A^* の固有値 $\overline{\alpha}$ に属する固有ベクトルでもあるので、

$$U^{-1}A^*U = U^*A^*U = \begin{array}{cc} \begin{array}{cc} \overset{l}{\longleftrightarrow} & \overset{n-l}{\longleftrightarrow} \end{array} & \\ \left(\begin{array}{ccc|c} \overline{\alpha} & & & \\ & \ddots & & B' \\ & & \overline{\alpha} & \\ \hline & O & & C' \end{array}\right) & \begin{array}{l} \updownarrow l \\ \updownarrow n-l \end{array} \end{array}$$

とも表せる

ここで、B と C は $l \times (n-l)$ 型行列、B' と C' は $(n-l) \times l$ 型行列であり、型が一致するので成分を比較できる

よって、

$$B^* = O, \quad C^* = C'$$

0 の複素共役は 0 であることから、$B^* = O$ より、

$$B = O$$

がしたがう

このことをふまえて、あらためて U^*AU を表すと、

$$U^*AU = \begin{array}{cc} \begin{array}{cc} \overset{l}{\longleftrightarrow} & \overset{n-l}{\longleftrightarrow} \end{array} & \\ \left(\begin{array}{ccc|c} \alpha & & & \\ & \ddots & & O \\ & & \alpha & \\ \hline & O & & C \end{array}\right) & \begin{array}{l} \updownarrow l \\ \updownarrow n-l \end{array} \end{array}$$

となる

ここで、A と U^*AU の特性多項式は一致するので、実際に計算すると、

$$
\begin{aligned}
\det(xE - A) &= \det(xE - U^*AU) \\
&= \left|\begin{array}{ccc|c} x-\alpha & & & \\ & \ddots & & O \\ & & x-\alpha & \\ \hline & O & & xE_{n-l} - C \end{array}\right| \\
&= \left|\begin{array}{ccc} x-\alpha & & \\ & \ddots & \\ & & x-\alpha \end{array}\right| \det(xE_{n-l} - C) \\
&= (x-\alpha)^l \det(xE_{n-l} - C)
\end{aligned}
$$

また、$\alpha E - U^*AU$ を考えると、

$$
\alpha E - U^*AU = \begin{array}{c} \overset{l}{\longleftrightarrow} \quad \overset{n-l}{\longleftrightarrow} \\ \left(\begin{array}{ccc|c} 0 & & & \\ & \ddots & & O \\ & & 0 & \\ \hline & O & & \alpha E_{n-l} - C \end{array}\right) \end{array} \begin{array}{l} \updownarrow l \\ \updownarrow n-l \end{array}
$$

より、

$$
\mathrm{rank}(\alpha E - U^*AU) = \mathrm{rank}(\alpha E_{n-l} - C)
$$

ここで、A と U^*AU は相似な行列であり、相似な行列の固有値（特性方程式の根）は重複度も含めて一致するので、

$$
\mathrm{rank}(\alpha E - U^*AU) = \mathrm{rank}(\alpha E - A) = n - l
$$

よって、

$$
\mathrm{rank}(\alpha E_{n-l} - C) = n - l
$$

つまり、$\alpha E_{n-l} - C$ は行列の階数が次数 $n-l$ に等しいので、正則行列である

ゆえにその行列式は、

$$
\det(\alpha E_{n-l} - C) \neq 0
$$

となることから、$x = \alpha$ は方程式 $\det(xE_{n-l} - C) = 0$ の解ではないことがわかる

よって、$\det(xE - A) = 0$ の解 $x = \alpha$ は、$(x-\alpha)^l$ の部分から現れることになるため、$x = \alpha$ は l 重解である

したがって、α の重複度 k は l に等しいことが示された ■

固有空間の次元と重複度が一致すれば対角化可能であることから、正規行列は対角化可能である

さらに、上の定理の証明過程から、正規行列はユニタリ行列によって対角化できることもわかる

正規行列のユニタリ対角化　複素正方行列 A について、A が正規行列であることと、A がユニタリ行列を用いて対角化できることは同値である

証明

正規行列 $\Longrightarrow$ ユニタリ行列を用いて対角化可能

正規行列における固有空間の次元と固有値の重複度の一致の定理の証明過程より明らか ■

ユニタリ行列を用いて対角化可能 $\Longrightarrow$ 正規行列

A がユニタリ行列 U を用いて、次のように対角化されたとする

$$U^*AU = \begin{pmatrix} \alpha_1 & & O \\ & \ddots & \\ O & & \alpha_n \end{pmatrix}$$

このとき、両辺に左から U をかけ、右から U^* をかけると、ユニタリ行列の定義より $U^*U = UU^* = E$ であることから、

$$A = U \begin{pmatrix} \alpha_1 & & O \\ & \ddots & \\ O & & \alpha_n \end{pmatrix} U^*$$

と変形できる

よって、A^* は、積の随伴行列をつくると積の順序が入れ替わることに注意して、

$$
\begin{aligned}
A^* &= (U^*)^* \begin{pmatrix} \overline{\alpha_1} & & O \\ & \ddots & \\ O & & \overline{\alpha_n} \end{pmatrix} U^* \\
&= U \begin{pmatrix} \overline{\alpha_1} & & O \\ & \ddots & \\ O & & \overline{\alpha_n} \end{pmatrix} U^*
\end{aligned}
$$

以上をふまえて、AA^* と A^*A をそれぞれ計算すると、

$$
\begin{aligned}
AA^* &= U \begin{pmatrix} \alpha_1 & & O \\ & \ddots & \\ O & & \alpha_n \end{pmatrix} U^* U \begin{pmatrix} \overline{\alpha_1} & & O \\ & \ddots & \\ O & & \overline{\alpha_n} \end{pmatrix} U^* \\
&= U \begin{pmatrix} \alpha_1\overline{\alpha_1} & & O \\ & \ddots & \\ O & & \alpha_n\overline{\alpha_n} \end{pmatrix} U^*
\end{aligned}
$$

$$
\begin{aligned}
A^*A &= U \begin{pmatrix} \overline{\alpha_1} & & O \\ & \ddots & \\ O & & \overline{\alpha_n} \end{pmatrix} U^* U \begin{pmatrix} \alpha_1 & & O \\ & \ddots & \\ O & & \alpha_n \end{pmatrix} U^* \\
&= U \begin{pmatrix} \overline{\alpha_1}\alpha_1 & & O \\ & \ddots & \\ O & & \overline{\alpha_n}\alpha_n \end{pmatrix} U^*
\end{aligned}
$$

となり、$\alpha_i\overline{\alpha_i} = \overline{\alpha_i}\alpha_i$ なので、たしかに、

$$
AA^* = A^*A
$$

が成り立つ

これは、A が正規行列であることを意味する ■

実対称行列の対角化

エルミート行列は正規行列なので、次のことがいえる

⚓ **エルミート行列のユニタリ対角化**　エルミート行列はユニタリ行列を用いて対角化できる

この定理を実行列の世界にもってくると、次のようになる

⚓ **実対称行列の直交対角化**　実対称行列は直交行列を用いて対角化できる

第 20 章

計量空間上の変換

ユニタリ変換

体 $\mathbb{C}$ 上の計量空間において、内積を保つ線形変換をユニタリ変換という

ユニタリ変換　体 $\mathbb{C}$ 上の計量空間 V における線形変換 f がユニタリ変換であるとは、任意の $\boldsymbol{u},\ \boldsymbol{v} \in V$ に対し、

$$(f(\boldsymbol{u}), f(\boldsymbol{v})) = (\boldsymbol{u}, \boldsymbol{v})$$

が成り立つことである

体 $\mathbb{R}$ 上のユニタリ変換は、直交変換と呼ばれる

ユニタリ変換の表現行列

ユニタリ行列の性質である内積不変性

$$(A\boldsymbol{u}, A\boldsymbol{v}) = (\boldsymbol{u}, \boldsymbol{v})$$

から、ユニタリ変換の表現行列はユニタリ行列であることがわかる

⚓ **ユニタリ変換とユニタリ行列表現**　計量空間上の線形変換 f がユニタリ変換であることと、f の表現行列 A がユニタリ行列であることは同値である

このことから、ユニタリ行列の性質は、ユニタリ変換の性質として言い換えることができる

ユニタリ変換とノルム

ユニタリ行列のノルム不変性から、

> ユニタリ変換はベクトルの長さを変えない変換

でもあることがわかる

⚓ **ユニタリ変換とノルム保存性**　計量空間 V における線形変換を f がユニタリ変換であることと、任意の $\boldsymbol{v} \in V$ に対し

$$\|f(\boldsymbol{v})\| = \|\boldsymbol{v}\|$$

が成り立つことは同値である

エルミート変換

エルミート変換　体 $\mathbb{C}$ 上の計量空間 V における線形空間 f がエルミート変換であるとは、任意の $\boldsymbol{u}, \boldsymbol{v} \in V$ に対し、

$$(f(\boldsymbol{u}), \boldsymbol{v}) = (\boldsymbol{u}, f(\boldsymbol{v}))$$

が成り立つことである

体 $\mathbb{R}$ 上のエルミート変換は、対称変換と呼ばれる

随伴写像

[Todo 44:]

随伴変換

[Todo 45:]

正規変換

[Todo 46:]

第 21 章

三角化と行列多項式

行列の三角化

対角化の次善の策として、三角化という方法がある

三角化定理　A を n 次複素正方行列とするとき、ある正則行列 P が存在して、$P^{-1}AP$ が上三角行列になる

その対角成分は重複度を含めて A の固有値と一致する

証明

三角化できること

n に関する帰納法を用いる

$n = 1$ のとき、A は 1×1 型行列なので、上三角行列である

$n \geq 2$ のとき、$\boldsymbol{v}_1$ を A の固有ベクトルとし、その固有値を α_1 とする

$\boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$ を追加して、$\mathbb{C}^n$ の基底に延長する

$P_1 = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$ とおくと、

$$P_1^{-1}AP_1 = \begin{pmatrix} \alpha_1 & * \\ \mathbf{0} & A_1 \end{pmatrix}$$

ここで、A_1 は $(n-1)$ 次正方行列である

帰納法の仮定より、$(n-1)$ 次の正則行列 P_2 を選べば、$P_2^{-1}A_1P_2$ は上三角行列になる

そこで、

$$P = P_1 \begin{pmatrix} 1 & {}^t\mathbf{0} \\ \mathbf{0} & P_2 \end{pmatrix}$$

とおくと、P_2 が正則であることから、P は正則である

P の逆行列は、

$$P^{-1} = \begin{pmatrix} 1 & {}^t\mathbf{0} \\ \mathbf{0} & P_2^{-1} \end{pmatrix} P_1^{-1}$$

であるので、

$$\begin{aligned} P^{-1}AP &= \begin{pmatrix} 1 & {}^t\mathbf{0} \\ \mathbf{0} & P_2^{-1} \end{pmatrix} P_1^{-1}AP_1 \begin{pmatrix} 1 & {}^t\mathbf{0} \\ \mathbf{0} & P_2 \end{pmatrix} \\ &= \begin{pmatrix} 1 & {}^t\mathbf{0} \\ \mathbf{0} & P_2^{-1} \end{pmatrix} \begin{pmatrix} \alpha_1 & * \\ \mathbf{0} & A_1 \end{pmatrix} \begin{pmatrix} 1 & {}^t\mathbf{0} \\ \mathbf{0} & P_2 \end{pmatrix} \\ &= \begin{pmatrix} \alpha_1 & * \\ \mathbf{0} & P_2^{-1}A_1P_2 \end{pmatrix} \end{aligned}$$

$P_2^{-1}A_1P_2$ は上三角行列であるから、$P^{-1}AP$ も上三角行列となる ■

対角成分が固有値と一致すること

一般に、三角行列の行列式は対角成分の積になる

このことから、n 次上三角行列 $B = (b_{ij})$ に対して、

$$\Phi_B(x) = \det(xE - B) = (x - b_{11}) \cdots (x - b_{nn})$$

が成り立つため、B の固有値は、特性方程式

$$(x - b_{11}) \cdots (x - b_{nn}) = 0$$

の解 $b_{11}, \ldots, b_{nn}$ となる

さて、$P^{-1}AP$ と A は相似な行列であるので、その特性多項式は一致する

$$\Phi_A(x) = \Phi_{P^{-1}AP}(x) = \det\bigl(xE - P^{-1}AP\bigr)$$

よって、$P^{-1}AP$ が上三角行列ならば、$A = (a_{ij})$ とおくと、

$$\Phi_A(x) = (x - a_{11}) \cdots (x - a_{nn})$$

が成り立ち、A の固有値は A の対角成分 $a_{11}, \ldots, a_{nn}$ となる ■

QR 分解

任意の正則行列は、ユニタリ行列（直交行列）と上三角行列の積に分解することができる

この分解は **QR 分解**と呼ばれ、行列の数値計算等で利用されている

正則行列に対する QR 分解の存在　任意の n 次複素正則行列 A に対して、$A = QR$ となるユニタリ行列 Q と上三角行列 R が存在する

証明

A は正則行列であるので、その列ベクトルは線型独立である

そこで、$A = (\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n)$ とおくと、$\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ からグラム・シュミットの直交化法を用いて、正規直交基底 $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n$ を作ることができる

このとき、次のように正規直交基底を用いた線形結合で $\boldsymbol{a}_j$ を表すことができる

$$\boldsymbol{a}_j = \sum_{i=1}^{j} (\boldsymbol{a}_j, \boldsymbol{u}_i)\boldsymbol{u}_i \quad (j = 1, \ldots, n)$$

この等式は、次のように書き換えられる

$$\boldsymbol{a}_j = (\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n) \begin{pmatrix} (\boldsymbol{a}_j, \boldsymbol{u}_1) \\ \vdots \\ (\boldsymbol{a}_j, \boldsymbol{u}_j) \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

そこで、

$$r_{jk} = \begin{cases} (\boldsymbol{a}_j, \boldsymbol{u}_k) & (1 \leq k \leq j) \\ 0 & (j < k \leq n) \end{cases}$$

とおくと、行列 $R = (r_{jk})$ は上三角行列である

また、$\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n$ は正規直交系であるので、これらを列ベクトルとした行列 $Q = (\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n)$ はユニタリ行列である

したがって、$A = QR$ となる ■

この定理の証明において、$\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ から正規直交基底を構成するところを工夫すると、正則行列だけでなく、任意の複素正方行列もユニタリ行列と上三角行列の積で表せることがわかる

⚓ **QR 分解の存在**　任意の n 次複素正方行列 A に対して、$A = QR$ となるユニタリ行列 Q と上三角行列 R が存在する

証明

A の列ベクトルは線型独立であるとは限らないが、グラム・シュミットの直交化法は線形従属なベクトルに対しても適用できる

ただし、線形従属なベクトルにグラム・シュミットの直交化法を適用すると、零ベクトルが得られることがある

零ベクトルは基底として使うことはできないため除外し、残った正規直交ベクトルを $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k$ とする

ここに $n-k$ 個のベクトルを補う形で、正規直交基底 $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k, \boldsymbol{u}_{k+1}, \ldots, \boldsymbol{u}_n$ を構成する

この場合も、正規直交基底を用いた線形結合で任意のベクトル $\boldsymbol{a}_j$ を表すことができるので、

$$\boldsymbol{a}_j = (\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n) \begin{pmatrix} (\boldsymbol{a}_j, \boldsymbol{u}_1) \\ \vdots \\ (\boldsymbol{a}_j, \boldsymbol{u}_k) \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

となり、以降は A が正則な場合と同様に示すことができる ■

ユニタリ行列による三角化

ユニタリ行列によって対角化できる行列は正規行列であった

したがって、正規行列以外の行列は、ユニタリ行列によって対角化することはできないが、ユニタリ行列によって三角化することはできる

シューア分解の存在　n 次複素正方行列 A に対して、適当なユニタリ行列 U により、$U^{-1}AU$ を上三角行列（A の**シューア形**）にすることができる

証明

任意の正則行列 P は QR 分解でき、ユニタリ行列 U と上三角行列 T を用いて、$P = UT$ と表せる

このとき、P の列ベクトルを $\boldsymbol{p}_1, \ldots, \boldsymbol{p}_n$、$U$ の列ベクトルを $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n$ とすると、

$$(\boldsymbol{p}_1, \ldots, \boldsymbol{p}_n) = (\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n)T$$

となり、P の列ベクトルは P の正則性より線型独立、U の列ベクトルは正規直交性から線型独立であるので、T は $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n$ から $\boldsymbol{p}_1, \ldots, \boldsymbol{p}_n$ への基底変換行列とみなせる

よって、基底変換行列 T は正則であるので、その逆行列を用いて、

$$U = PT^{-1}$$

と書くことができる

また、U はユニタリ行列であるので、その定義から正則である

U の逆行列は、正則行列の積に対する逆行列の公式と、逆行列の逆行列をとると元に戻ることを用いて、

$$U^{-1} = (PT^{-1})^{-1} = (T^{-1})^{-1}P^{-1} = TP^{-1}$$

と計算できる

これらを用いると、

$$\begin{aligned} U^{-1}AU &= (TP^{-1})APT^{-1} \\ &= T(P^{-1}AP)T^{-1} \end{aligned}$$

となり、ここで、

- T は上三角行列
- $P^{-1}AP$ は三角化定理より、上三角行列
- T^{-1} は上三角行列の逆行列より、上三角行列

であることと、上三角行列の積は上三角行列になることから、$U^{-1}AU$ も上三角行列になる ■

行列多項式

[Todo 47:]

フロベニウスの定理

[Todo 48:]

ケイリー・ハミルトンの定理

[Todo 49:]

最小多項式

[Todo 50:]

第 22 章

直交補空間と射影行列

直交補空間

内積を導入したことで、ベクトルの長さや直交性が利用できるようになった。

直交性は、ベクトルだけでなく、部分空間に対しても拡張できる。

計量空間の部分空間に直交するベクトルの集合を、直交補空間という。

直交補空間　計量空間 V の部分空間 W に対し、W の直交補空間 $W^{\perp}$ を次のように定義する。

$$W^{\perp} := \{\boldsymbol{v} \in V \mid \forall \boldsymbol{w} \in W, (\boldsymbol{v}, \boldsymbol{w}) = 0\}$$

直交補空間は V の部分空間

直交補空間もまた、計量空間の部分空間になっている。

直交補空間の部分空間性　計量空間 V の部分空間 W の直交補空間 $W^{\perp}$ は、計量空間 V の部分空間である。

証明

和について

$\boldsymbol{a}_1, \boldsymbol{a}_2 \in W^\perp$ とすると、任意の $\boldsymbol{b} \in W$ に対して、

$$(\boldsymbol{a}_1 + \boldsymbol{a}_2, \boldsymbol{b}) = (\boldsymbol{a}_1, \boldsymbol{b}) + (\boldsymbol{a}_2, \boldsymbol{b}) = 0 + 0 = 0$$

となるので、$\boldsymbol{a}_1 + \boldsymbol{a}_2 \in W^\perp$ である。 ■

スカラー倍について

$\boldsymbol{a} \in W^\perp$ とすると、任意のスカラー $c \in K$ と任意の $\boldsymbol{b} \in W$ に対して、

$$(c\boldsymbol{a}, \boldsymbol{b}) = c(\boldsymbol{a}, \boldsymbol{b}) = c \cdot 0 = 0$$

となるので、$c\boldsymbol{a} \in W^\perp$ である。 ■

直交補空間による直和分解

「直交補空間」という名前は、「補集合」と同様に、何らかの集合を補う集合であることを想起させる。

実際、直交補空間 $W^\perp$ は、もとの集合 W を補い、V 全体を構成するような性質を持つ。

直交補空間を用いた計量空間の分解　計量空間 V の部分空間 W に対して、

$$V = W + W^\perp$$

証明

$W = \{\boldsymbol{o}\}$ の場合は、任意の $\boldsymbol{v} \in V$ に対して $\boldsymbol{o}$ との内積は 0 になることから、$W^\perp$ は V 全体となる。

$$W^\perp = \{\boldsymbol{v} \in V \mid (\boldsymbol{v}, \boldsymbol{o}) = 0\} = V$$

よって、

$$V = W + W^{\perp} = \{\boldsymbol{o}\} + V = V$$

が成り立つ。

以降、$W \neq \{\boldsymbol{o}\}$ とする。

W の基底 $\{\boldsymbol{w}'_1, \ldots, \boldsymbol{w}'_k\}$ を 1 つとり、これに対してグラム・シュミットの直交化法を適用して、正規直交基底 $\{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_k\}$ を得る。

任意の $\boldsymbol{v} \in V$ をとり、次のようにおく。

$$\boldsymbol{u} = \boldsymbol{v} - \sum_{i=1}^{k} (\boldsymbol{v}, \boldsymbol{w}_i)\boldsymbol{w}_i$$

$\boldsymbol{u}$ と $\boldsymbol{w}_i$ の内積を計算すると、

$$\begin{aligned}
(\boldsymbol{u}, \boldsymbol{w}_i) &= \left(\boldsymbol{v} - \sum_{j=1}^{k} (\boldsymbol{v}, \boldsymbol{w}_j)\boldsymbol{w}_j, \boldsymbol{w}_i\right) \\
&= (\boldsymbol{v}, \boldsymbol{w}_i) - \sum_{j=1}^{k} (\boldsymbol{v}, \boldsymbol{w}_j)(\boldsymbol{w}_j, \boldsymbol{w}_i) \\
&= (\boldsymbol{v}, \boldsymbol{w}_i) - \sum_{j=1}^{k} (\boldsymbol{v}, \boldsymbol{w}_j)\delta_{ij} \\
&= (\boldsymbol{v}, \boldsymbol{w}_i) - (\boldsymbol{v}, \boldsymbol{w}_i) \\
&= 0
\end{aligned}$$

このように、任意の $i = 1, \ldots, k$ に対して、$\boldsymbol{u}$ と $\boldsymbol{w}_i$ の内積が 0 になることから、$\boldsymbol{u} \in W^{\perp}$ である。

一方、$\boldsymbol{u}$ の定義式を $\boldsymbol{v}$ を表す式として整理すると、

$$\boldsymbol{v} = \boldsymbol{u} + \sum_{i=1}^{k} (\boldsymbol{v}, \boldsymbol{w}_i)\boldsymbol{w}_i$$

となるが、$\boldsymbol{w}_i$ が W の正規直交基底であることから、

$$\sum_{i=1}^{k} (\boldsymbol{v}, \boldsymbol{w}_i)\boldsymbol{w}_i$$

の部分は、W の任意の元を表す。

よって、V の任意の元 $\boldsymbol{v}$ は、W の元と $W^{\perp}$ の元 $\boldsymbol{u}$ の和として表されるので、

$$V = W + W^{\perp}$$

が成り立つ。■

さらに、次の定理が成り立つことで、単なる空間の和ではなく、直和として分解できることがわかる。

直交補空間との交わり　計量空間 V の部分空間 W に対して、

$$W \cap W^{\perp} = \{\boldsymbol{o}\}$$

証明

$\boldsymbol{a} \in W \cap W^{\perp}$ とすると、$\boldsymbol{a} \in W$ かつ $\boldsymbol{a} \in W^{\perp}$ である。

$\boldsymbol{a} \in W^{\perp}$ より、$\boldsymbol{a} \in W$ に対しても内積が 0 になるので、

$$(\boldsymbol{a}, \boldsymbol{a}) = 0$$

ここで、内積の性質より、

$$(\boldsymbol{a}, \boldsymbol{a}) = \|\boldsymbol{a}\|^2 \geq 0$$

であり、等号が成立するのは、$\boldsymbol{a} = \boldsymbol{o}$ のときのみである。

よって、$\boldsymbol{a} = \boldsymbol{o}$ である。

零ベクトルは任意のベクトルと直交し（内積が 0 になり）、また任意の部分空間に属するので、明らかに $\boldsymbol{o} \in W \cap W^{\perp}$ である。

$\boldsymbol{a}$ は $W \cap W^{\perp}$ の任意の元であり、$\boldsymbol{a} = \boldsymbol{o} \in W \cap W^{\perp}$ であることがわかったので、

$$W \cap W^{\perp} = \{\boldsymbol{o}\}$$

がいえる。■

こうして、次の両方が成り立つことから、

i. $V = W + W^{\perp}$

ii. $W \cap W^{\perp} = \{\boldsymbol{o}\}$

計量空間 V は部分空間 W とその直交補空間 $W^{\perp}$ の直和として分解できる。

よって、直和の次元公式より、次の定理が従う。

直交補空間と次元　計量空間 V の部分空間 W に対して、

$$\dim V = \dim W + \dim W^{\perp}$$

直交補空間の性質

直交補空間の直交補空間はもとの空間

部分空間の双直交補と元空間の一致　計量空間 V の部分空間 W に対して、次が成り立つ。

$$(W^{\perp})^{\perp} = W$$

証明

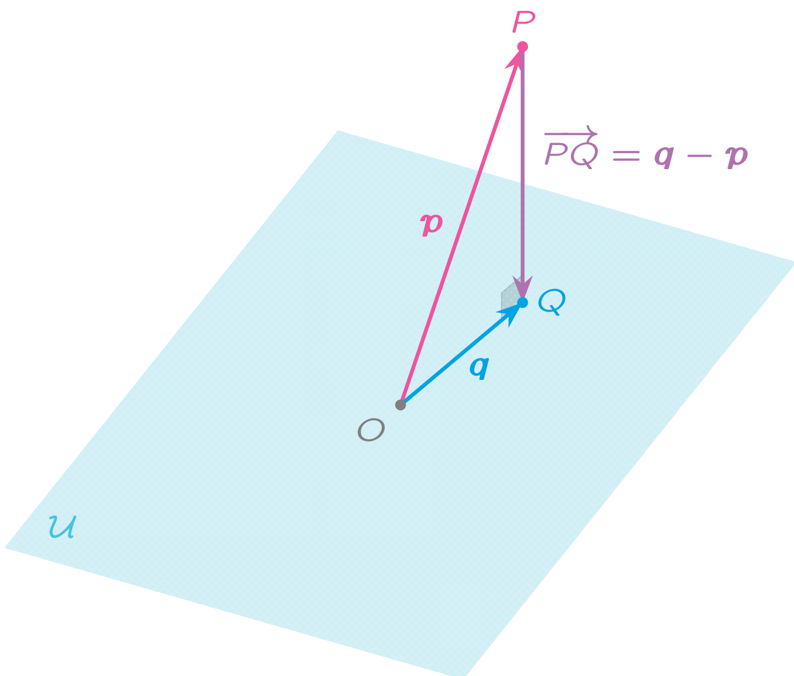
[Todo 51:]

直交射影と反射影

$\mathbb{R}^n$ 上の点 P に対して、部分空間 $\mathcal{U}$ 上の点 $Q \in \mathcal{U}$ のうち、$\overrightarrow{PQ}$ が $\mathcal{U}$ に直交するような点 Q を、点 P の $\mathcal{U}$ への**直交射影**あるいは**正射影**という

また、$\overrightarrow{QP}$ を点 Q の $\mathcal{U}$ からの**反射影**という

射影前のベクトルを $\boldsymbol{p}$、射影後のベクトルを $\boldsymbol{q}$ とすると、直交射影とは、$\boldsymbol{q}$ と $\boldsymbol{q}-\boldsymbol{p}$ が直交するように射影することである



このとき、次のような関係が成り立っている

$$\overrightarrow{OP} = \overrightarrow{OQ} + \overrightarrow{QP}$$
$$\overrightarrow{OQ} \in \mathcal{U}, \quad \overrightarrow{QP} \in \mathcal{U}^{\perp}$$

ここで、$\mathcal{U}^{\perp}$ は部分空間 $\mathcal{U}$ に直交するベクトルの全体であり、$\mathcal{U}$ の**直交補空間**と呼ばれる

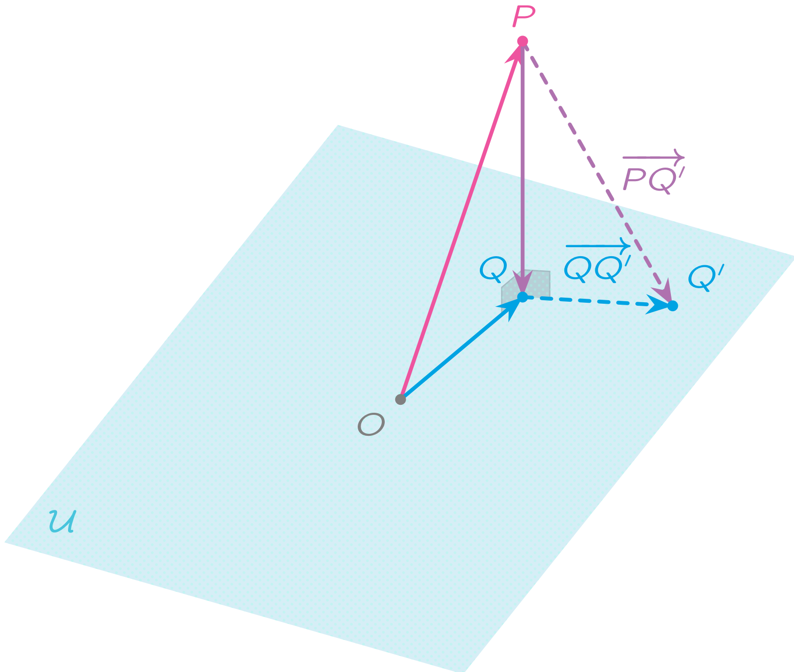
$\mathbb{R}^n$ の部分空間 $\mathcal{U}$ の直交補空間 $\mathcal{U}^{\perp}$ も、$\mathbb{R}^n$ の部分空間となる

$\overrightarrow{OP}$ は $\mathbb{R}^n$ の任意のベクトルを表すことから、$\mathbb{R}^n$ のベクトルは、$\mathcal{U}$ への射影 $\overrightarrow{OQ}$ と、$\mathcal{U}$ からの反射影 $\overrightarrow{QP}$ の和として表されることがわかる

このような表し方は一意的であり、$\overrightarrow{OP}$ の $\mathcal{U}$ と $\mathcal{U}^{\perp}$ への**直和分解**という

直交射影と最短距離

点 Q を $\mathcal{U}$ 上の別の点 Q' に移動した場合を考える



このとき、三平方の定理より、

$$\|\overrightarrow{PQ'}\|^2 = \|\overrightarrow{PQ}\|^2 + \|\overrightarrow{QQ'}\|^2 > \|\overrightarrow{PQ}\|^2$$

となるから、

直交射影した点 Q は、

点 P から最短となる $\mathcal{U}$ 上の点

であることがわかる

射影行列

ベクトルの射影の概念は、射影行列を用いて表現できる。

任意のベクトル $\boldsymbol{x} \in \mathbb{R}^n$ は、$\boldsymbol{u} \in \mathcal{U}$、$\boldsymbol{u}^\perp \in \mathcal{U}^\perp$ を用いて

$$\boldsymbol{x} = \boldsymbol{u} + \boldsymbol{u}^\perp$$

と一意的に分解できる（直和分解）。

ここで、$\boldsymbol{x}$ の $\mathcal{U}$ への射影を表すのは、$\boldsymbol{u}$ である。

つまり、$\mathcal{U}$ への射影とは $\boldsymbol{x}$ のうち、$\mathcal{U}$ に含まれる成分 $\boldsymbol{u}$ だけを取り出す操作といえる。

そこで、部分空間 $\mathcal{U}$ へ射影する写像を $P_{\mathcal{U}}$ とすると、

$$P_{\mathcal{U}}\boldsymbol{x} = \boldsymbol{u}$$

という、「$\boldsymbol{x}$ に $P_{\mathcal{U}}$ を作用させると $\boldsymbol{u}$ だけが残る」という形で書ける。

部分空間への射影

このとき、$\boldsymbol{x}$ がもともと $\mathcal{U}$ の元である場合は、$\boldsymbol{u}^{\perp} = \boldsymbol{o}$ の場合と考えて、

$$\boldsymbol{x} = \boldsymbol{u} + \boldsymbol{o} = \boldsymbol{u}$$

つまり、射影しても変わらないので、

$$P_{\mathcal{U}}\boldsymbol{x} = \boldsymbol{u} = \boldsymbol{x} \quad (\boldsymbol{x} \in \mathcal{U})$$

である。

一方、$\boldsymbol{x}$ が $\mathcal{U}$ の直交補空間 $\mathcal{U}^{\perp}$ の元の場合は、$\boldsymbol{u} = \boldsymbol{o}$ の場合と考えて、

$$P_{\mathcal{U}}\boldsymbol{x} = \boldsymbol{u} = \boldsymbol{o} \quad (\boldsymbol{x} \in \mathcal{U}^{\perp})$$

となる。

以上をまとめると、次のように書ける。

$$P_{\mathcal{U}}\boldsymbol{x} = \begin{cases} \boldsymbol{x} & (\boldsymbol{x} \in \mathcal{U}) \\ \boldsymbol{o} & (\boldsymbol{x} \in \mathcal{U}^{\perp}) \end{cases}$$

同様に、直交補空間 $\mathcal{U}^{\perp}$ へ射影する写像を $P_{\mathcal{U}^{\perp}}$ とした場合、

$$P_{\mathcal{U}^{\perp}}\boldsymbol{x} = \begin{cases} \boldsymbol{o} & (\boldsymbol{x} \in \mathcal{U}) \\ \boldsymbol{x} & (\boldsymbol{x} \in \mathcal{U}^{\perp}) \end{cases}$$

とまとめられる。

射影行列の展開

$\mathbb{R}^n$ が $\mathcal{U}$ と $\mathcal{U}^\perp$ の直和に分解されることから、$\mathbb{R}^n$ の基底は $\mathcal{U}$ の基底と $\mathcal{U}^\perp$ の基底を合わせたものになる。

そこで、部分空間 $\mathcal{U}$ の正規直交基底 $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_r\}$ を選ぶと、これを $\mathbb{R}^n$ の正規直交基底 $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_r, \boldsymbol{u}_{r+1}, \ldots, \boldsymbol{u}_n\}$ に拡張できる。
ここで、$\{\boldsymbol{u}_{r+1}, \ldots, \boldsymbol{u}_n\}$ は $\mathcal{U}^\perp$ の正規直交基底になる。

このとき、

$$P_{\mathcal{U}}\boldsymbol{x} = \begin{cases} \boldsymbol{x} & (\boldsymbol{x} \in \mathcal{U}) \\ \boldsymbol{o} & (\boldsymbol{x} \in \mathcal{U}^\perp) \end{cases}$$

という式は、$P_{\mathcal{U}}$ が $\mathbb{R}^n$ の正規直交基底

$$\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_r, \boldsymbol{u}_{r+1}, \ldots, \boldsymbol{u}_n\}$$

を、それぞれ次のように写像することを意味する。

$$\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_r, \boldsymbol{o}, \ldots, \boldsymbol{o}\}$$

同様に、

$$P_{\mathcal{U}^\perp}\boldsymbol{x} = \begin{cases} \boldsymbol{o} & (\boldsymbol{x} \in \mathcal{U}) \\ \boldsymbol{x} & (\boldsymbol{x} \in \mathcal{U}^\perp) \end{cases}$$

という式は、$P_{\mathcal{U}^\perp}$ が $\mathbb{R}^n$ の正規直交基底

$$\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_r, \boldsymbol{u}_{r+1}, \ldots, \boldsymbol{u}_n\}$$

を、それぞれ次のように写像することを意味する。

$$\{\boldsymbol{o}, \ldots, \boldsymbol{o}, \boldsymbol{u}_{r+1}, \ldots, \boldsymbol{u}_n\}$$

ゆえに、正規直交基底による表現行列の展開より、$P_{\mathcal{U}}$ と $P_{\mathcal{U}^\perp}$ は次のように表現できる。

$$\begin{aligned} P_{\mathcal{U}} &= \boldsymbol{u}_1\boldsymbol{u}_1^\top + \cdots + \boldsymbol{u}_r\boldsymbol{u}_r^\top \\ P_{\mathcal{U}^\perp} &= \boldsymbol{u}_{r+1}\boldsymbol{u}_{r+1}^\top + \cdots + \boldsymbol{u}_n\boldsymbol{u}_n^\top \end{aligned}$$

$P_{\mathcal{U}}$ と $P_{\mathcal{U}^\perp}$ をそれぞれ、部分空間 $\mathcal{U}$、およびその直交補空間 $\mathcal{U}^\perp$ への射影行列と呼ぶ。

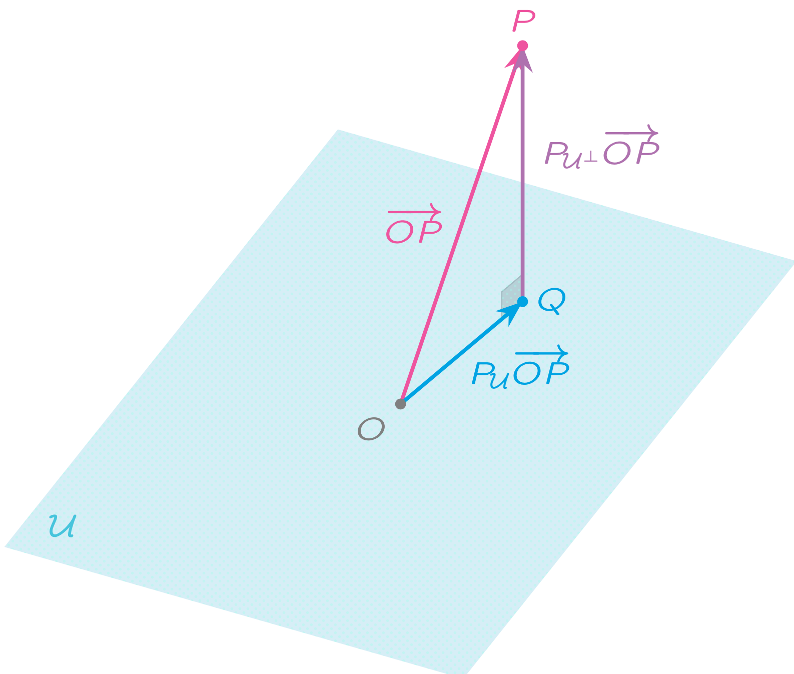
単位行列の射影行列への分解

直交射影と反射影の章で示した、

$$\overrightarrow{OP} = \overrightarrow{OQ} + \overrightarrow{QP}$$
$$\overrightarrow{OQ} \in \mathcal{U}, \quad \overrightarrow{QP} \in \mathcal{U}^{\perp}$$

という関係は、射影行列を用いて、次のようにも表せる

$$\begin{aligned}\overrightarrow{OP} &= P_{\mathcal{U}}\overrightarrow{OP} + P_{\mathcal{U}^{\perp}}\overrightarrow{OP} \\ &= (P_{\mathcal{U}} + P_{\mathcal{U}^{\perp}})\overrightarrow{OP}\end{aligned}$$



$\mathbb{R}^n$ 内のすべての点 P に対して、$\overrightarrow{OP} = (P_{\mathcal{U}} + P_{\mathcal{U}^{\perp}})\overrightarrow{OP}$ が成り立つことから、

$$P_{\mathcal{U}} + P_{\mathcal{U}^{\perp}} = E$$

が成り立っている

これはすなわち、単位行列 E が、部分空間 $\mathcal{U}$ その直交補空間 $\mathcal{U}^{\perp}$ への射影行列の和に分解できることを意味する

$$E = \underbrace{\boldsymbol{u}_1\boldsymbol{u}_1^{\top} + \cdots + \boldsymbol{u}_r\boldsymbol{u}_r^{\top}}_{P_{\mathcal{U}}} + \underbrace{\boldsymbol{u}_{r+1}\boldsymbol{u}_{r+1}^{\top} + \cdots + \boldsymbol{u}_n\boldsymbol{u}_n^{\top}}_{P_{\mathcal{U}^{\perp}}}$$

この式により、単位行列 E 自体を、空間全体 $\mathbb{R}^n$ への射影行列と考えることもできる

射影行列とノルム

$P_{\mathcal{U}}\overrightarrow{OP}$ と $P_{\mathcal{U}^\perp}\overrightarrow{OP}$ は直交するから、三平方の定理より、

$$\|\overrightarrow{OP}\|^2 = \|P_{\mathcal{U}}\overrightarrow{OP}\|^2 + \|P_{\mathcal{U}^\perp}\overrightarrow{OP}\|^2$$

がいえる

ゆえに、任意のベクトル $\boldsymbol{x} = \overrightarrow{OP}$ に対して、

$$\|\boldsymbol{x}\|^2 = \|P_{\mathcal{U}}\boldsymbol{x}\|^2 + \|P_{\mathcal{U}^\perp}\boldsymbol{x}\|^2$$

が成り立つ

射影行列の冪等性と対称性

「一度射影した点をもう一度射影しても変化しない」という性質は、次のような数式として表現できる

射影行列の冪等性　射影行列 $P_{\mathcal{U}}$ は冪等である

$$P_{\mathcal{U}}^2 = P_{\mathcal{U}}$$

証明

$$\begin{aligned}
P_{\mathcal{U}}^2 &= \left(\sum_{i=1}^{r} \boldsymbol{u}_i\boldsymbol{u}_i^\top\right)\left(\sum_{j=1}^{r} \boldsymbol{u}_j\boldsymbol{u}_j^\top\right) \\
&= \sum_{i=1}^{r}\sum_{j=1}^{r} \boldsymbol{u}_i\boldsymbol{u}_i^\top\boldsymbol{u}_j\boldsymbol{u}_j^\top \\
&= \sum_{i=1}^{r}\sum_{j=1}^{r} \boldsymbol{u}_i(\boldsymbol{u}_i^\top\boldsymbol{u}_j)\boldsymbol{u}_j^\top \\
&= \sum_{i=1}^{r}\sum_{j=1}^{r} \delta_{ij}\boldsymbol{u}_i\boldsymbol{u}_j^\top
\end{aligned}$$

ここで、δ_{ij} を含むことから、$i=j$ の場合のみ項が残り、

$$P_{\mathcal{U}}^2 = \sum_{i=1}^{r} \boldsymbol{u}_i \boldsymbol{u}_i^\top = P_{\mathcal{U}}$$

が得られる ■

この $P_{\mathcal{U}}^2 = P_{\mathcal{U}}$ という式は、一般の（必ずしも直交射影でない）射影の定義として用いられる

次の性質は、射影が直交射影であることを示すものである

射影行列の対称性　射影行列 $P_{\mathcal{U}}$ は、対称行列である

$$P_{\mathcal{U}}^\top = P_{\mathcal{U}}$$

証明

$$P_{\mathcal{U}}^\top = \left(\sum_{i=1}^{r} \boldsymbol{u}_i \boldsymbol{u}_i^\top \right)^\top$$

和の転置は転置の和であることを用いて、

$$P_{\mathcal{U}}^\top = \sum_{i=1}^{r} (\boldsymbol{u}_i \boldsymbol{u}_i^\top)^\top$$

また、積の転置は転置の積だが、積の順序が入れ替わることに注意して、

$$P_{\mathcal{U}}^\top = \sum_{i=1}^{r} (\boldsymbol{u}_i^\top)^\top \boldsymbol{u}_i^\top$$

転置の転置をとると元に戻るので、

$$P_{\mathcal{U}}^\top = \sum_{i=1}^{r} \boldsymbol{u}_i \boldsymbol{u}_i^\top = P_{\mathcal{U}}$$

が得られる ■

射影行列は冪等かつ対称であるが、その逆も成り立つ

対称性と冪等性による射影行列の特徴づけ　対称かつ冪等な行列は、ある部分空間への射影行列となる

証明

n 次対称行列 P は、n 個の実数の固有値 $\lambda_1, \ldots, \lambda_n$ を持つ

これらに属する固有ベクトルを $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n$ とすると、$\boldsymbol{u}_i$ は互いに直交する

固有値と固有ベクトルの関係から、

$$P\boldsymbol{u}_i = \lambda_i \boldsymbol{u}_i$$

両辺に左から P をかけると、

$$P^2\boldsymbol{u}_i = \lambda_i P\boldsymbol{u}_i = \lambda_i \cdot \lambda_i \boldsymbol{u}_i = \lambda_i^2 \boldsymbol{u}_i$$
$$\therefore \quad P^2\boldsymbol{u}_i = \lambda_i^2 \boldsymbol{u}_i$$

ここで、P は冪等なので、$P^2 = P$ が成り立つ

$$P^2\boldsymbol{u}_i = P\boldsymbol{u}_i = \lambda_i \boldsymbol{u}_i$$

これを用いると、

$$\lambda_i \boldsymbol{u}_i = \lambda_i^2 \boldsymbol{u}_i$$

固有ベクトル $\boldsymbol{u}_i$ は零ベクトルではないので、

$$\lambda_i = \lambda_i^2$$

よって、

$$\begin{aligned} \lambda_i^2 - \lambda_i &= 0 \\ \lambda_i(\lambda_i - 1) &= 0 \\ \lambda_i &= 0, 1 \end{aligned}$$

すなわち、固有値は 0 か 1 のいずれかである

そこで、

$$\lambda_1 = \cdots = \lambda_r = 1$$
$$\lambda_{r+1} = \cdots = \lambda_n = 0$$

とおくと、

$$P\boldsymbol{u}_i = \begin{cases} \boldsymbol{u}_i & (i = 1, \ldots, r) \\ 0 & (i = r+1, \ldots, n) \end{cases}$$

よって、P は $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_r\}$ の張る部分空間への射影行列である ■

第 23 章

アフィン空間と超平面

超平面

[Todo 52:]

射影長

[Todo 53:]

第 24 章

抽象線形空間

線形空間の公理

線形代数の理論は線型独立性や線形写像を基礎にしている

これらは線形結合、すなわちベクトルの和とスカラー倍を用いて定義された

任意のベクトルは線形結合で表され、線形写像は線形結合を保つ写像として定義される

そこで、和とスカラー倍が定義された一般の集合に対しても、線型空間の理論を適用できないか？と考える

和とスカラー倍が定義された一般の集合を、改めて線形空間として定義する

そして、その集合の元をベクトルと呼ぶことにする

和とスカラー倍が定義されていれば、線形結合によりその元を表すことができるからだ

線形空間　集合 V の任意の元 $\boldsymbol{a}, \boldsymbol{b}$ と体 K の任意の元 k に対して、V の元 $\boldsymbol{a}+\boldsymbol{b}$（和）が定まり、$V$ の元 $k\boldsymbol{a}$（スカラー倍）が定まるとする

これらの演算が次の条件を満たすとき、V を K 上の線形空間、あるいは K 線型空間と呼び、線型空間の元をベクトルと呼ぶ

i. 交換法則：$\boldsymbol{a}+\boldsymbol{b}=\boldsymbol{b}+\boldsymbol{a}$

ii. 結合法則：$(\boldsymbol{a}+\boldsymbol{b})+\boldsymbol{c}=\boldsymbol{a}+(\boldsymbol{b}+\boldsymbol{c})$、$k(l\boldsymbol{a})=(kl)\boldsymbol{a}$

iii. 分配法則：$k(\boldsymbol{a}+\boldsymbol{b})=k\boldsymbol{a}+k\boldsymbol{b}$、$(k+l)\boldsymbol{a}=k\boldsymbol{a}+l\boldsymbol{a}$

iv. $1\boldsymbol{a}=\boldsymbol{a}$（1 は体 K の乗法に関する単位元）

v. 零元の存在：$\mathbf{0}$ と書かれる特別な元が存在し、任意の $\boldsymbol{a} \in V$ に対して $\boldsymbol{a}+\mathbf{0}=\boldsymbol{a}$

vi. 和に関する逆元の存在：任意の $\boldsymbol{a} \in V$ に対して $-\boldsymbol{a}$ と書かれる特別な元が存在し、$\boldsymbol{a}+(-\boldsymbol{a})=(-\boldsymbol{a})+\boldsymbol{a}=\mathbf{0}$

線形写像の空間

$V,\ W$ をともに有限次元 K 上の線形空間とする

V から W への線形写像全体の集合を $\mathrm{Hom}(V,W)$ と書く

特に、V の線形変換全体の集合は $\mathrm{End}(V)$ と書く

このとき、$\mathrm{Hom}(V,W)$ に線型空間の構造（和とスカラー倍）を次のように導入する

線形写像の和とスカラー倍　線形写像 $f,g \in \mathrm{Hom}(V,W)$ と $c \in K$ に対して、和とスカラー倍を次のように定義する

$$(f+g)(v) := f(v)+g(v)$$
$$(cf)(v) := c \cdot f(v)$$

これらの演算は、再び $V \to W$ の線形写像を定めることが確認できる

線形写像全体による線形空間　線形写像全体の集合 $\mathrm{Hom}(V,W)$ は K 上の線形空間である

証明

加法が線形性を満たす

f, g をともに線形写像とし、任意の $v_1, v_2 \in V$ と $a, b \in K$ に対して、

$$\begin{aligned} &(f+g)(av_1+bv_2) \\ &= f(av_1+bv_2)+g(av_1+bv_2) \\ &= af(v_1)+bf(v_2)+ag(v_1)+bg(v_2) \\ &= a(f(v_1)+g(v_1))+b(f(v_2)+g(v_2)) \\ &= a(f+g)(v_1)+b(f+g)(v_2) \end{aligned}$$

よって、$f+g$ は線形写像である ■

スカラー倍が線形性を満たす

f を線形写像とし、任意の $v_1, v_2 \in V$ と $a, b, c \in K$ に対して、

$$\begin{aligned} (cf)(av_1+bv_2) &= cf(av_1+bv_2) \\ &= c \cdot (f(av_1+bv_2)) \\ &= c \cdot (f(av_1)+f(bv_2)) \\ &= c \cdot (af(v_1)+bf(v_2)) \\ &= a(cf)(v_1)+b(cf)(v_2) \end{aligned}$$

よって、cf は線形写像である ■

線型空間の公理をすべて満たすことも、容易に確認できる ■

第 25 章

広義固有空間

第 26 章

双対空間

内積から線形汎関数へ

横ベクトル（$1 \times n$ 型行列）を縦ベクトル（$n \times 1$ 型行列）にかけると、1×1 のスカラー値が得られる。

$$\begin{pmatrix} a_1 & \cdots & a_n \end{pmatrix} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = a_1 v_1 + \cdots + a_n v_n$$

上の式は、数ベクトル空間の内積そのものである。

$$\langle \boldsymbol{a} | \boldsymbol{v} \rangle = \boldsymbol{a}^\top \boldsymbol{v} = a_1 v_1 + \cdots + a_n v_n$$

さて、「観測装置としての内積」の章で述べたように、

> 内積 $\langle \boldsymbol{a} | \boldsymbol{v} \rangle$ は、観測装置 $\langle \boldsymbol{a} |$ によるベクトル $| \boldsymbol{v} \rangle$ の測定結果

という捉え方もできる。

ここで、観測装置である横ベクトル $\langle \boldsymbol{a} |$ を、縦ベクトル $| \boldsymbol{v} \rangle$ から内積を返す関数 $\phi_{\boldsymbol{a}}$ とみることにしよう。

$$\phi_{\boldsymbol{a}}(\boldsymbol{v}) = \langle \boldsymbol{a} | \boldsymbol{v} \rangle = a_1 v_1 + \cdots + a_n v_n$$

$\phi_{\boldsymbol{a}}$ は、縦ベクトル $\boldsymbol{v}$ を入力とし、スカラー値 $\langle \boldsymbol{a} | \boldsymbol{v} \rangle$ を返す、$\mathbb{R}^n$ から $\mathbb{R}$ への写像である。

さらに、内積の双線形性から、$\phi_{\boldsymbol{a}}$ は線形写像であることがわかる。

$$\begin{aligned} \phi_{\boldsymbol{a}}(c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2) &= (\boldsymbol{a}, c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2) \\ &= c_1(\boldsymbol{a}, \boldsymbol{v}_1) + c_2(\boldsymbol{a}, \boldsymbol{v}_2) \\ &= c_1\phi_{\boldsymbol{a}}(\boldsymbol{v}_1) + c_2\phi_{\boldsymbol{a}}(\boldsymbol{v}_2) \end{aligned}$$

この関数 $\phi_{\boldsymbol{a}}$ は、線形汎関数と呼ばれる写像の一例である。

$\mathbb{R}^n$ 上の線形汎関数　$\mathbb{R}^n$ 上の関数 $\phi\colon \mathbb{R}^n \to \mathbb{R}$ が線形写像であるとき、ϕ を $\mathbb{R}^n$ 上の線形汎関数あるいは線形形式という。

線形汎関数のベクトル表示

$\mathbb{R}^n$ 上の線形汎関数は、すべて内積から定めることができる。

$\mathbb{R}^n$ 上の線形汎関数の内積による表現　$\mathbb{R}^n$ 上の任意の線形汎関数 $\psi\colon \mathbb{R}^n \to \mathbb{R}$ に対し、ある $\boldsymbol{a} \in \mathbb{R}^n$ がただ一つ存在して、次を満たす。

$$\psi = \phi_{\boldsymbol{a}} = \langle \boldsymbol{a} | \cdot \rangle$$

証明

$\mathbb{R}^n$ の標準基底を $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ とする。

このとき、任意のベクトル $\boldsymbol{v} \in \mathbb{R}^n$ は、次のように表される。

$$\boldsymbol{v} = v_1\boldsymbol{e}_1 + \cdots + v_n\boldsymbol{e}_n$$

これに ψ を作用させると、線形汎関数 ψ は線形性をもつので、

$$\begin{aligned}\psi(\boldsymbol{v}) &= \psi(v_1\boldsymbol{e}_1 + \cdots + v_n\boldsymbol{e}_n) \\ &= v_1\psi(\boldsymbol{e}_1) + \cdots + v_n\psi(\boldsymbol{e}_n) \\ &= \begin{pmatrix}\psi(\boldsymbol{e}_1) & \cdots & \psi(\boldsymbol{e}_n)\end{pmatrix}\begin{pmatrix}v_1 \\ \vdots \\ v_n\end{pmatrix}\end{aligned}$$

ここで、

$$\boldsymbol{a} = \begin{pmatrix}\psi(\boldsymbol{e}_1) & \cdots & \psi(\boldsymbol{e}_n)\end{pmatrix}$$

とおけば、次が成り立つ。

$$\psi(\boldsymbol{v}) = \langle \boldsymbol{a} | \boldsymbol{v} \rangle = \phi_{\boldsymbol{a}}(\boldsymbol{v})$$

$\boldsymbol{v}$ は任意のベクトルなので、

$$\psi = \phi_{\boldsymbol{a}} = \langle \boldsymbol{a} | \cdot \rangle$$

となるような $\boldsymbol{a} \in \mathbb{R}^n$ の存在が示された。

さらに、次式を振り返ると、ψ が決まれば $\boldsymbol{a}$ が一意に定まることがわかる。

$$\boldsymbol{a} = \begin{pmatrix}\psi(\boldsymbol{e}_1) & \cdots & \psi(\boldsymbol{e}_n)\end{pmatrix}$$

よって、ψ に対して $\boldsymbol{a}$ はただ一つ存在する。 ■

上の定理の証明で現れた次の式は、2 通りの読み方ができる。

$$\boldsymbol{a} = \begin{pmatrix}\psi(\boldsymbol{e}_1) & \cdots & \psi(\boldsymbol{e}_n)\end{pmatrix}$$

ψ が決まれば、$\psi(\boldsymbol{e}_1), \ldots, \psi(\boldsymbol{e}_n)$ の値が決まるので、$\boldsymbol{a}$ がただ一つ定まる。

逆に、基底 $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ に対する ψ の値が決まれば ψ の形が決まるので、上の式のように $\boldsymbol{a}$ を定めれば、$\boldsymbol{a}$ に対応して ψ の形がただ一つに定まることになる。

まとめると、

- すべてのベクトル $\boldsymbol{a}$ は線形汎関数 ψ をひとつ定める
- すべての線形汎関数 ψ はベクトル $\boldsymbol{a}$ をひとつ定める

$\boldsymbol{a}$ から ψ への対応は一対一であり、ψ から $\boldsymbol{a}$ への対応も一対一である。

すなわち、$\mathbb{R}^n$ のベクトルと $\mathbb{R}^n$ 上の線形汎関数の間には、全単射が存在する。

全単射な対応は、本来同じものに「異なる表現を与えている」と捉えることができる。

縦ベクトルと横ベクトルによる線形汎関数の表現

次の式も、先ほどの定理の証明で現れたものである。

$$\psi(\boldsymbol{v}) = \begin{pmatrix} \psi(\boldsymbol{e}_1) & \cdots & \psi(\boldsymbol{e}_n) \end{pmatrix} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$$

この式もまた、2 通りの読み方ができる。

$\boldsymbol{a}$ を横ベクトルとみるなら、

$$\psi(\boldsymbol{v}) = \underbrace{\begin{pmatrix} \psi(\boldsymbol{e}_1) & \cdots & \psi(\boldsymbol{e}_n) \end{pmatrix}}_{\boldsymbol{a}} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \boldsymbol{a}\boldsymbol{v}$$

この見方では、線形汎関数は横ベクトル $\boldsymbol{a}$ との「行列としての積」である。

線形汎関数を行列の積として定義すれば、「横」ベクトル $\boldsymbol{a}$ が線形汎関数の表現行列に相当すると捉えられる。

一方、$\boldsymbol{a}$ を縦ベクトルとみるなら、

$$\psi(\boldsymbol{v}) = \underbrace{\begin{pmatrix} \psi(\boldsymbol{e}_1) & \cdots & \psi(\boldsymbol{e}_n) \end{pmatrix}}_{\boldsymbol{a}^\top} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \boldsymbol{a}^\top \boldsymbol{v} = (\boldsymbol{a}, \boldsymbol{v})$$

この見方では、線形汎関数は縦ベクトル $\boldsymbol{a}$ との「内積」である。

線形汎関数を内積として定義すれば、「縦」ベクトル $\boldsymbol{a}$ が線形汎関数の表現行列に相当すると捉えられる。

このように、線形汎関数という同じものに対して、横ベクトルと縦ベクトルは「異なる表現を与えている」とも解釈できる。

横ベクトルと縦ベクトルが転置という関係で結ばれていることで、この 2 通りの見方が可能になる。

線形汎関数の空間

内積の双線形性は、任意のベクトル $\boldsymbol{v}$ に対して、

$$(c_1\boldsymbol{a}_1 + c_2\boldsymbol{a}_2, \boldsymbol{v}) = c_1(\boldsymbol{a}_1, \boldsymbol{v}) + c_2(\boldsymbol{a}_2, \boldsymbol{v})$$

が成り立つというものだった。

これは、$\mathbb{R}^n$ 上の線形汎関数が満たす関係式と読み替えることができる。

$$\phi_{c_1\boldsymbol{a}_1+c_2\boldsymbol{a}_2}(\boldsymbol{v}) = c_1\phi_{\boldsymbol{a}_1}(\boldsymbol{v}) + c_2\phi_{\boldsymbol{a}_2}(\boldsymbol{v})$$

この関係式は、$\mathbb{R}^n$ 上の線形汎関数の集合に、線形空間としての構造をもたらす。

$\mathbb{R}^n$ 上の線形汎関数の集合を $(\mathbb{R}^n)^*$ と書くことにしよう。
この集合 $(\mathbb{R}^n)^*$ に和とスカラー倍の演算を導入することで、$(\mathbb{R}^n)^*$ を線形空間とみなすことができる。

線形汎関数の空間の基底

$\mathbb{R}^n$ の基底を $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n\}$ とするとき、任意のベクトル $\boldsymbol{v} \in \mathbb{R}^n$ は、

$$\boldsymbol{v} = v_1\boldsymbol{u}_1 + \cdots + v_n\boldsymbol{u}_n = \begin{pmatrix} \boldsymbol{u}_1 & \cdots & \boldsymbol{u}_n \end{pmatrix} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$$

という線形結合で表すことができる。
ここで、$v_1, \ldots, v_n$ は、基底 $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n\}$ に関する $\boldsymbol{v}$ の成分あるいは座標と呼ばれる。

このうち第 j 座標 v_j を取得する関数を ϕ_j と定めよう。

$$\phi_j(\boldsymbol{v}) = v_j$$

このような関数を座標関数と呼ぶことにする。

また、ϕ_j は線形であるため、$\mathbb{R}^n$ 上の線形汎関数である。

補足：ϕ_j の線形性

任意の $\boldsymbol{v}, \boldsymbol{w} \in \mathbb{R}^n$ が基底 $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n\}$ に関して次のように表せるとする。

$$\boldsymbol{v} = \sum_{i=1}^{n} v_i \boldsymbol{u}_i, \quad \boldsymbol{w} = \sum_{j=1}^{n} w_j \boldsymbol{u}_j$$

このとき、ϕ_j は次のように定義される。

$$\phi_j(\boldsymbol{v}) = v_j, \quad \phi_j(\boldsymbol{w}) = w_j$$

ベクトルの和を考えると、

$$\boldsymbol{v} + \boldsymbol{w} = \sum_{i=1}^{n} (v_i + w_i) \boldsymbol{u}_i$$

より、第 j 座標は $v_j + w_j$ となるので、

$$\phi_j(\boldsymbol{v} + \boldsymbol{w}) = v_j + w_j = \phi_j(\boldsymbol{v}) + \phi_j(\boldsymbol{w})$$

ベクトルのスカラー倍を考えると、

$$\alpha \boldsymbol{v} = \sum_{i=1}^{n} (\alpha v_i) \boldsymbol{u}_i$$

より、第 j 座標は αv_j となるので、

$$\phi_j(\alpha \boldsymbol{v}) = \alpha v_j = \alpha \phi_j(\boldsymbol{v})$$

以上より、$\phi_j \colon \mathbb{R}^n \to \mathbb{R}$ は線形写像であることが示された。

ϕ_j を用いると、$\boldsymbol{v}$ を表す線形結合は次のように書ける。

$$\boldsymbol{v} = \phi_1(\boldsymbol{v}) \boldsymbol{u}_1 + \cdots + \phi_n(\boldsymbol{v}) \boldsymbol{u}_n$$

ここで、たとえば $\boldsymbol{v}$ を $\boldsymbol{u}_1$ に置き換えた式を考える。

$$\boldsymbol{u}_1 = \phi_1(\boldsymbol{u}_1) \boldsymbol{u}_1 + \cdots + \phi_n(\boldsymbol{u}_1) \boldsymbol{u}_n$$

この等式が成り立つには、

- $\phi_1(\boldsymbol{u}_1) = 1$
- $\phi_2(\boldsymbol{u}_1) = 0, \ldots, \phi_n(\boldsymbol{u}_1) = 0$

でなければならない。

右辺の $\boldsymbol{u}_1$ だけが残り、他の項が消えることで、$\boldsymbol{u}_1 = \boldsymbol{u}_1$ という等式が成り立つ。

同様に考えると、$\boldsymbol{v}$ を $\boldsymbol{u}_i$ に置き換えた式

$$\boldsymbol{u}_i = \phi_1(\boldsymbol{u}_i)\boldsymbol{u}_1 + \cdots + \phi_n(\boldsymbol{u}_i)\boldsymbol{u}_n$$

が成り立つには、$\boldsymbol{u}_i$ だけが残り、他の項が消えなければならないので、

$$\phi_j(\boldsymbol{u}_i) = \delta_{ij} = \begin{cases} 1 & (i = j) \\ 0 & (i \neq j) \end{cases}$$

と定める必要がある。

この式により、$\mathbb{R}^n$ の基底 $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n$ を選べば、それらに対応する線形汎関数 $\phi_1, \ldots, \phi_n$ が定まることがわかる。

そしてこのとき、$\phi_1, \ldots, \phi_n$ は $(\mathbb{R}^n)^*$ の基底となっている。

$\mathbb{R}^n$ における基底に対応する線形汎関数の構成　$\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n\}$ を $\mathbb{R}^n$ の基底とするとき、$\phi_j \in (\mathbb{R}^n)^*$ を次のように定める。

$$\phi_j(\boldsymbol{u}_i) = \delta_{ij}$$

このような $\phi_1, \ldots, \phi_n$ は $(\mathbb{R}^n)^*$ の基底をなす。

証明

$\phi_1, \ldots, \phi_n$ が線型独立であること

次のような $\phi_1, \ldots, \phi_n$ の線形関係式を考える。

$$c_1\phi_1 + \cdots + c_n\phi_n = 0$$

このとき、任意の j に対して、

$$
\begin{aligned}
(c_1\phi_1 + \cdots + c_n\phi_n)(\boldsymbol{u}_j) &= c_1\phi_1(\boldsymbol{u}_j) + \cdots + c_n\phi_n(\boldsymbol{u}_j) \\
&= \sum_{i=1}^{n} c_i\phi_i(\boldsymbol{u}_j) = \sum_{i=1}^{n} c_i\delta_{ij} \\
&= c_j = 0
\end{aligned}
$$

が成り立たなければならない。

これは $\phi_1, \ldots, \phi_n$ が線型独立であることを示している。 ■

$\phi_1, \ldots, \phi_n$ が $(\mathbb{R}^n)^*$ を張ること

$\psi \in (\mathbb{R}^n)^*$ を任意にとると、$\boldsymbol{u}_j$ に対する値 $\alpha_j = \psi(\boldsymbol{u}_j)$ が定まる。

このとき、α_j を係数とする $\phi_1, \ldots, \phi_n$ の線形結合を作ると、

$$
\begin{aligned}
(\alpha_1\phi_1 + \cdots + \alpha_n\phi_n)(\boldsymbol{u}_j) &= \alpha_1\phi_1(\boldsymbol{u}_j) + \cdots + \alpha_n\phi_n(\boldsymbol{u}_j) \\
&= \sum_{i=1}^{n} \alpha_i\phi_i(\boldsymbol{u}_j) = \sum_{i=1}^{n} \alpha_i\delta_{ij} = \alpha_j \\
&= \psi(\boldsymbol{u}_j)
\end{aligned}
$$

ϕ_j, ψ はともに $\mathbb{R}^n$ から $\mathbb{R}$ への線形写像であり、ϕ_j の線形結合もまた $(\mathbb{R}^n)^*$ の元なので $\mathbb{R}^n$ から $\mathbb{R}$ への線形写像である。

よって、$\mathbb{R}^n$ の基底 $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n\}$ に対して同じ値をとることから、

$$
\psi = \alpha_1\phi_1 + \cdots + \alpha_n\phi_n
$$

がいえる。

したがって、任意の $\psi \in (\mathbb{R}^n)^*$ は $\phi_1, \ldots, \phi_n$ の線形結合として表すことができるため、

$$
(\mathbb{R}^n)^* = \langle \phi_1, \ldots, \phi_n \rangle
$$

が示された。 ■

線形汎関数の空間の次元

$\mathbb{R}^n$ の基底 $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n\}$ と、それに対応する $(\mathbb{R}^n)^*$ の基底 $\{\phi_1, \ldots, \phi_n\}$ は、どちらも n 個のベクトルの組になっている。

> ここでいう「ベクトル」とは、「線形空間の元」という意味である。$(\mathbb{R}^n)^*$ も線形空間であるので、その元である線形汎関数も「ベクトル」と呼んでいる。

基底をなすベクトルの個数は、その空間の次元として定義されるので、次のことがいえる。

⚓ $\mathbb{R}^n$ とその線形汎関数の空間の次元の一致　$\mathbb{R}^n$ 上の線形汎関数の空間 $(\mathbb{R}^n)^*$ の次元は、$\mathbb{R}^n$ の次元と等しい。

$$\dim \mathbb{R}^n = \dim(\mathbb{R}^n)^* = n$$

また、次元が等しいことから、$\mathbb{R}^n$ と $(\mathbb{R}^n)^*$ は線形同型である。
すなわち、$\mathbb{R}^n$ の元（縦ベクトル）と $(\mathbb{R}^n)^*$ の元（$\mathbb{R}^n$ 上の線形汎関数）の間には、全単射が存在する。

基底を決めれば、縦ベクトルと線形汎関数を同一視する（同じものの「異なる表現」と捉える）ことができる。

横ベクトルと座標関数

$n \times 1$ 型行列（n 次の縦ベクトル）全体の集合は $\mathbb{R}^n$ と表された。
$1 \times n$ 型行列（n 次の横ベクトル）全体の集合を ${}^t\mathbb{R}^n$ と表すことにする。

${}^t\mathbb{R}^n$ の元は $1 \times n$ 型行列なので、$\mathbb{R}^n$ から $\mathbb{R}$ への線形写像（すなわち $\mathbb{R}^n$ 上の線形汎関数）を表現している行列だと考えることができる。

座標関数の表現行列

基本ベクトルを転置したもの ${}^t\boldsymbol{e}_j \in {}^t\mathbb{R}^n$ を縦ベクトル $\boldsymbol{v} \in \mathbb{R}^n$ にかけると、$\boldsymbol{v}$ の j 番目の成分が得られる。

たとえば、$n = 3,\ j = 2$ の場合、

$$ {}^t\boldsymbol{e}_2 \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = v_2 $$

といった具合に、2 番目の成分 v_2 が得られる。

このように、ベクトル $\boldsymbol{v} \in \mathbb{R}^n$ に対して、その j 番目の成分を返す**座標関数**を x_j と表記することにしよう。

$$ \begin{array}{cccc} x_j\colon & \mathbb{R}^n & \longrightarrow & \mathbb{R} \\ & \Cup & & \Cup \\ & \boldsymbol{v} & \longmapsto & v_j = {}^t\boldsymbol{e}_j\boldsymbol{v} \end{array} $$

このとき、$x_j\colon \mathbb{R}^n \to \mathbb{R}$ は $\mathbb{R}^n$ 上の**線形汎関数**である。

${}^t\boldsymbol{e}_j\boldsymbol{v}$ を行列の積として見ると、横基本ベクトル ${}^t\boldsymbol{e}_j \in {}^t\mathbb{R}^n$ は線形汎関数 x_j の表現行列だと捉えることができる。

[Todo 54: 「基底方向への正射影」という観点についても述べる？]

横ベクトルと線形汎関数の同一視

任意の縦ベクトルは、基本ベクトル（標準基底）の線形結合として一意的に表現できる。

$$ |\boldsymbol{v}\rangle = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = v_1\boldsymbol{e}_1 + \cdots + v_n\boldsymbol{e}_n $$

同様に、任意の横ベクトルは、横基本ベクトルの線形結合として一意的に表現できる。

$$ \langle\boldsymbol{a}| = \begin{pmatrix} a_1 & \cdots & a_n \end{pmatrix} = a_1{}^t\boldsymbol{e}_1 + \cdots + a_n{}^t\boldsymbol{e}_n $$

ここで、横ベクトル $\langle \boldsymbol{a}|$ は観測装置という視点に戻って、縦ベクトルを入力したら $\boldsymbol{a}$ との内積を返す線形汎関数を ϕ とおくと、

$$\begin{aligned}\phi(\boldsymbol{v}) = \boldsymbol{a}^\top \boldsymbol{v} &= \begin{pmatrix} a_1 & \cdots & a_n \end{pmatrix} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \\ &= a_1 v_1 + \cdots + a_n v_n \\ &= a_1 {}^t\boldsymbol{e}_1 \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} + \cdots + a_n {}^t\boldsymbol{e}_n \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \\ &= a_1 x_1(\boldsymbol{v}) + \cdots + a_n x_n(\boldsymbol{v})\end{aligned}$$

よって、任意の線形汎関数 $\phi \in (\mathbb{R}^n)^*$ は、座標関数 $x_1, \ldots, x_n$ の線型結合として表すことができる。

$$\phi = a_1 x_1 + \cdots + a_n x_n$$

また、x_i の表現行列が ${}^t\boldsymbol{e}_i$ であることを思い出すと、

$$\phi = a_1 {}^t\boldsymbol{e}_1 + \cdots + a_n {}^t\boldsymbol{e}_n = \langle \boldsymbol{a}|$$

というように、線形汎関数 ϕ は横ベクトル $\langle \boldsymbol{a}|$ と同一視することができる。

$\{{}^t\boldsymbol{e}_1, \ldots, {}^t\boldsymbol{e}_n\}$ を基底としてどんな横ベクトルも表現できることは、$\{x_1, \ldots, x_n\}$ を基底としてどんな線形汎関数も表現できることに対応する。
これより、横ベクトルの空間 ${}^t\mathbb{R}^n$ と、線形汎関数の空間 $(\mathbb{R}^n)^*$ は、同じ空間とみなすことができる。

縦ベクトルと横ベクトルの双対性

$\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n\}$ を $\mathbb{R}^n$ の基底とするとき、任意の縦ベクトル $\boldsymbol{v} \in \mathbb{R}^n$ は、

$$\boldsymbol{v} = v_1 \boldsymbol{u}_1 + \cdots + v_n \boldsymbol{u}_n$$

という線形結合で表すことができる。

ここで、$v_1, \ldots, v_n$ は基底 $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n\}$ に関する $\boldsymbol{v}$ の座標である。

このうち、j 番目の座標 v_j を取得する関数を $\phi_j\colon \mathbb{R}^n \to \mathbb{R}$ と定めると、ϕ_j は、

$$\phi_j(\boldsymbol{u}_i) = \delta_{ij}$$

を満たし、$\{\phi_1, \ldots, \phi_n\}$ が $(\mathbb{R}^n)^*$ の基底となる。

このとき、$(\mathbb{R}^n)^*$ の元（線形汎関数）を横ベクトルと同一視すると、任意の横ベクトル $\phi \in {}^t\mathbb{R}^n$ は、

$$\phi = c_1\phi_1 + \cdots + c_n\phi_n$$

という線形結合で表すことができる。

ここで、$c_1, \ldots, c_n$ は基底 $\{\phi_1, \ldots, \phi_n\}$ に関する ϕ の座標である。
このうち、j 番目の座標 c_j を取得する関数を $\psi_j\colon {}^t\mathbb{R}^n \to \mathbb{R}$ と定めると、ψ_j は、

$$\psi_j(\phi_i) = \delta_{ij}$$

を満たし、$\{\psi_1, \ldots, \psi_n\}$ が $({}^t\mathbb{R}^n)^*$ の基底となる。

さて、基底を変えれば座標も変わってしまうので、ψ_j はあくまでも基底が $\{\phi_1, \ldots, \phi_n\}$ のときの横ベクトルの座標を返す関数である。
さらに、ϕ_j は $\mathbb{R}^n$ の基底が $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n\}$ のときの縦ベクトルの座標を返す関数である。

つまり、ψ_j は $\mathbb{R}^n$ の基底 $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n\}$ に依存しているので、$\boldsymbol{u}_j \in \mathbb{R}^n$ を入力として ψ_j を定める関数 ι を考えてみる。

$$\begin{array}{cccc} \iota\colon & \mathbb{R}^n & \longrightarrow & ({}^t\mathbb{R}^n)^* \\ & \Psi & & \Psi \\ & \boldsymbol{u}_j & \longmapsto & \psi_j \end{array}$$

ι を用いると、次のように書ける。

$$\iota(\boldsymbol{u}_j) = \psi_j$$

このとき、基底に対して座標は一意的であり、基底が変わると座標が変わることから、

i. 基底 $\{\boldsymbol{u}_j\}_{j=1}^n$ を固定すれば、$\iota(\boldsymbol{u}_j) = \psi_j$ を満たす座標 $\{\psi_j\}_{j=1}^n$ は一意に定まる
ii. 座標 $\{\psi_j\}_{j=1}^n$ を固定すれば、$\iota(\boldsymbol{u}_j) = \psi_j$ を満たす基底 $\{\boldsymbol{u}_j\}_{j=1}^n$ は一意に定まる

という 2 通りの見方ができる。

このように、$\boldsymbol{u}_j \in \mathbb{R}^n$ と $\psi_j \in ({}^t\mathbb{R}^n)^*$ には、「互いに測り、測られる」という対称性がある。このような対称性を双対性という。

この性質を意識し、${}^t\mathbb{R}^n$ を $\mathbb{R}^n$ の双対空間という。

$$\begin{tikzcd} \mathbb{R}^n \arrow[r, leftrightarrow, "\text{同一視}"] \arrow[dr, "\text{裏返し}"'] & ({}^t\mathbb{R}^n)^* \\ & {}^t\mathbb{R}^n \arrow[u, "\text{裏返し}"'] \end{tikzcd}$$

双対とは、「裏返しにした関係」と解釈できる。

${}^t\mathbb{R}^n$ が $\mathbb{R}^n$ の双対空間であるとは、「横ベクトルの空間 ${}^t\mathbb{R}^n$ を裏返しにしたもの $({}^t\mathbb{R}^n)^*$ は、縦ベクトルの空間 $\mathbb{R}^n$ と同一視できる」ということである。

逆に、$\mathbb{R}^n$ は ${}^t\mathbb{R}^n$ の双対空間である。「縦ベクトルの空間 $\mathbb{R}^n$ を裏返しにしたもの $(\mathbb{R}^n)^*$ は、横ベクトルの空間 ${}^t\mathbb{R}^n$ と同一視できる」ということでもある。
すなわち、線形汎関数の空間 $(\mathbb{R}^n)^*$ を横ベクトルの空間 ${}^t\mathbb{R}^n$ と同一視できる。

そこで、${}^t\mathbb{R}^n$ を $(\mathbb{R}^n)^*$ に書き換えると、

$$\begin{tikzcd} \mathbb{R}^n \arrow[r, leftrightarrow, "\text{同一視}"] \arrow[dr, "\text{裏返し}"'] & ((\mathbb{R}^n)^*)^* \\ & (\mathbb{R}^n)^* \arrow[u, "\text{裏返し}"'] \end{tikzcd}$$

という関係が見えてくる。$(\mathbb{R}^n)^*$ を $\mathbb{R}^n$ の双対空間という。
表 $\mathbb{R}^n$ の裏は $(\mathbb{R}^n)^*$ であり、裏の裏 $((\mathbb{R}^n)^*)^*$ は表 $\mathbb{R}^n$ になる。

双対空間と双対基底

ここまでの話を、一般の線形空間 V に拡張しよう。

まず、V 上の線形汎関数を次のように定義する。

線形汎関数　V を $\mathbb{R}$ 上の線形空間とする。V から $\mathbb{R}$ への線形写像 $\phi\colon V \to \mathbb{R}^n$ を V 上の線形汎関数あるいは線形形式という。

V から $\mathbb{R}$ への線形写像、すなわち V 上の線形汎関数全体の集合を考える。

双対空間　V 上の線形汎関数全体の集合を V の双対空間といい、V^* と表す。

$$V^* := \mathrm{Hom}(V, \mathbb{R}) = \{\phi\colon V \to \mathbb{R} \mid \phi\text{は線形写像}\}$$

線形空間 V が有限次元の場合は、選んでおいた V の基底に対して、双対基底（dual basis）という双対空間 V^* の基底を考えることができる。

双対基底の構成　V を n 次元の線形空間とし、$\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ を V の基底とする。このとき、$\phi_j \in V^*$ を次のように定める。

$$\phi_j(\boldsymbol{v}_i) = \delta_{ij}$$

このような $\phi_1, \ldots, \phi_n$ は V^* の基底をなす。

この定理は $V = \mathbb{R}^n$ の場合と同様に示すことができる。

また、この定理から次が成り立つ。

双対空間の次元　n 次元線形空間 V の双対空間 V^* の次元は、V の次元と等しい。

$$\dim V = \dim V^* = n$$

これより、V と V^* は線形同型であることがいえるが、この同型は基底に依存していることに注意しよう。

一旦ここまでの話をまとめると、次のような関係が成り立っている。

$$\begin{tikzcd}
V \arrow[r, leftrightarrow, "?"] \arrow[dr, "\text{基底による同型}"'] & (V^*)^* \\
& V^* \arrow[u, "\text{基底による同型}"']
\end{tikzcd}$$

再双対空間による自然同型

線形空間 V の双対空間 V^* もまた線形空間になるので、さらにその双対空間 $(V^*)^*$ を考えることができる。

$(V^*)^*$ を V の再双対空間あるいは第 2 双対空間といい、V^{**} と書くこともできる。

実は $(V^*)^*$ と V は線形同型であり、この同型は V の基底に依存しないことが示される。

$$\begin{tikzcd} V \arrow[r, leftrightarrow, "\text{自然な同型}"] \arrow[dr, "\text{基底による同型}"'] & (V^*)^* \\ & V^* \arrow[u, "\text{基底による同型}"'] \end{tikzcd}$$

再双対空間への写像

線形汎関数 $\phi \in V^*$ に $\boldsymbol{v} \in V$ を入力して得られるスカラー値を次のように書くことにする。

$$\langle \phi, \boldsymbol{v} \rangle := \phi(\boldsymbol{v})$$

$\boldsymbol{v} \in V$ を固定したとき、任意の線形汎関数（V^* の元）に $\boldsymbol{v}$ を入力したもの $\langle -, \boldsymbol{v} \rangle$ を考えることができる。

> $-$ はプレースホルダーであり、（線形汎関数なら）なんでも入れられることを意味する。具体的な線形汎関数が決まっていないときは、$-(\boldsymbol{v})$ と書くよりも、$\langle -, \boldsymbol{v} \rangle$ と書いた方がわかりやすい。

ここで、具体的な $\phi \in V^*$ を与えれば、スカラー値 $\langle \phi, \boldsymbol{v} \rangle$ が確定する。

$$\begin{array}{cccc} \Phi_{\boldsymbol{v}} \colon & V^* & \longrightarrow & \mathbb{R} \\ & \rotatebox{90}{$\in$} & & \rotatebox{90}{$\in$} \\ & \phi & \longmapsto & \langle \phi, \boldsymbol{v} \rangle \end{array}$$

この写像 $\phi \mapsto \langle \phi, \boldsymbol{v} \rangle$ を $\Phi_{\boldsymbol{v}}$ と書くことにしよう。

$$\Phi_{\boldsymbol{v}}(\phi) = \langle \phi, \boldsymbol{v} \rangle = \phi(\boldsymbol{v})$$

このように定めた $\Phi_{\boldsymbol{v}} \colon V^* \to \mathbb{R}$ は線形写像であるので、$(V^*)^*$ 上の線形汎関数である。

補足：$\Phi_{\boldsymbol{v}}$ の線形性

$\phi_1, \phi_2 \in V^*$、$c_1, c_2 \in \mathbb{R}$ とする。

ϕ_1, ϕ_2 は線形写像であるので、線形写像の和とスカラー倍の定義より、

$$\begin{aligned}\Phi_{\boldsymbol{v}}(c_1\phi_1 + c_2\phi_2) &= (c_1\phi_1 + c_2\phi_2)(\boldsymbol{v}) \\ &= c_1\phi_1(\boldsymbol{v}) + c_2\phi_2(\boldsymbol{v}) \\ &= c_1\Phi_{\boldsymbol{v}}(\phi_1) + c_2\Phi_{\boldsymbol{v}}(\phi_2)\end{aligned}$$

となるので、$\Phi_{\boldsymbol{v}}$ は線形写像である。

余談だが、上の式変形は次のように書くこともできる。

$$\begin{aligned}\Phi_{\boldsymbol{v}}(c_1\phi_1 + c_2\phi_2) &= \langle c_1\phi_1 + c_2\phi_2, \boldsymbol{v}\rangle \\ &= c_1\langle\phi_1, \boldsymbol{v}\rangle + c_2\langle\phi_2, \boldsymbol{v}\rangle \\ &= c_1\Phi_{\boldsymbol{v}}(\phi_1) + c_2\Phi_{\boldsymbol{v}}(\phi_2)\end{aligned}$$

この見方に慣れておくと、後の議論に対して戸惑いが少なくなる。

また、$\Phi_{\boldsymbol{v}}$ は $\boldsymbol{v}$ に依存しているので、各 $\boldsymbol{v} \in V$ に $\Phi_{\boldsymbol{v}} \in (V^*)^*$ を対応させる写像 ι を考えることができる。

$$\begin{array}{cccc}\iota\colon & V & \longrightarrow & (V^*)^* \\ & \rotatebox{90}{$\in$} & & \rotatebox{90}{$\in$} \\ & \boldsymbol{v} & \longmapsto & \Phi_{\boldsymbol{v}}\end{array}$$

このように定めた $\iota\colon V \to (V^*)^*$ は線形写像である。

補足：ι の線形性

$\boldsymbol{v}_1, \boldsymbol{v}_2 \in V$、$c_1, c_2 \in \mathbb{R}$ とすると、

$$\begin{aligned}\iota(c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2) &= \Phi_{c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2} \\ &= \langle -, c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2\rangle \\ &= c_1\langle -, \boldsymbol{v}_1\rangle + c_2\langle -, \boldsymbol{v}_2\rangle \\ &= c_1\Phi_{\boldsymbol{v}_1} + c_2\Phi_{\boldsymbol{v}_2} \\ &= c_1\iota(\boldsymbol{v}_1) + c_2\iota(\boldsymbol{v}_2)\end{aligned}$$

となるので、ι は線形写像である。

$\iota\colon V \to (V^*)^*$ は線形写像であるので、ι が線形同型写像であることを示せば、V と $(V^*)^*$ の同型が導かれる。

そのためには、ι の全単射性を証明できればよい。

双対空間の分離性

特に ι が単射であることを示すために、次の定理を用いる。

⚓ **双対空間の分離性**　有限次元線形空間 V において、任意の $\boldsymbol{v} \in V$ で $\boldsymbol{v} \neq \boldsymbol{o}$ ならば、$\phi(\boldsymbol{v}) \neq 0$ となるような線形汎関数 $\phi \in V^*$ が存在する。

証明

$\boldsymbol{v} \neq \boldsymbol{o}$ より、$\boldsymbol{v}$ は線型独立である。

よって、基底の延長により、$\boldsymbol{v}$ を含む V の基底 $\{\boldsymbol{v}, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ を選ぶことができる。

この基底に対応する双対基底 $\phi_1, \phi_2, \ldots, \phi_n \subset V^*$ を考えると、それぞれの ϕ_i は、次の性質をもつ。

$$\phi_i(\boldsymbol{v}_j) = \delta_{ij} \quad (i, j = 1, 2, \ldots, n)$$

このとき $\phi_1(\boldsymbol{v}) = 1$ であるので、$\phi = \phi_1$ をとれば、任意の $\boldsymbol{v} \neq \boldsymbol{o}$ に対して $\phi(\boldsymbol{v}) = 1$ となる。 ■

再双対空間との同型

⚓ **再双対空間との自然な同型**　V が有限次元ならば、$\iota\colon V \to (V^*)^*$ は線形同型である。

証明

写像 ι は単射

$\iota(\boldsymbol{v}) = 0$ すなわち、任意の $\phi \in V^*$ に対して

$$\iota(\boldsymbol{v})(\phi) = \phi(\boldsymbol{v}) = 0$$

であると仮定する。

この仮定は、すべての線形汎関数が $\boldsymbol{v}$ を 0 に写すことを意味する。

ここで、$\boldsymbol{v} \neq \boldsymbol{o}$ とすると、双対空間の分離性より、$\phi(\boldsymbol{v}) \neq 0$ となるような線形汎関数 ϕ が存在する。

これは $\iota(\boldsymbol{v}) = 0$ という仮定と矛盾するので、$\iota(\boldsymbol{v}) = 0$ のもとでは、$\boldsymbol{v} = \boldsymbol{o}$ でなければならない。

したがって、

$$\iota(\boldsymbol{v}) = 0 \Longrightarrow \boldsymbol{v} = \boldsymbol{o}$$

となり、これは線形写像 ι が単射であることを示している。 ■

写像 ι は全射

双対空間の次元を考えると、

$$\dim(V^*)^* = \dim V^* = \dim V$$

[Note 4: 次元定理と全射性との関係を加筆したら、その記載箇所へのリンクを貼る]

ι が単射であることから $\mathrm{Ker}(\iota) = \{\boldsymbol{o}\}$ なので、線形写像の次元定理より、$\dim(V^*)^* = \dim V$ は $\iota \colon V \to (V^*)^*$ が全射であることを示している。

■

双対ペアリング

V と $(V^*)^*$ の間には、線形同型写像 $\iota\colon V \to (V^*)^*$ が存在する。

このことから、V と $(V^*)^*$ は線形同型であることがいえる。

このように、V が有限次元の場合は、V と $(V^*)^*$ を自然に（基底によらずに）同一視することができる。

ここで、再双対空間への写像を考える際に登場した次の式を再解釈してみよう。

$$\Phi_{\boldsymbol{v}}(\phi) = \phi(\boldsymbol{v})$$

V と $(V^*)^*$ の同型により、$\boldsymbol{v} \in V$ と $\Phi_{\boldsymbol{v}} \in (V^*)^*$ も同一視することができる。

そこで、$\Phi_{\boldsymbol{v}}$ を単に $\boldsymbol{v}$ と書くことにすると、次の関係が得られる。

$$\boldsymbol{v}(\phi) = \phi(\boldsymbol{v})$$

これは、$\boldsymbol{v} \in V$ と $\phi \in V^*$ に対し、

値 $\phi(\boldsymbol{v})$ をとることは、$\boldsymbol{v}$ から見ても ϕ から見ても対等

であることを表している。

この平等さを表すために、次のような記法を使うこともある。

$$\langle \phi, \boldsymbol{v} \rangle = \langle \boldsymbol{v}, \phi \rangle = \phi(\boldsymbol{v})$$

この記号 $\langle \cdot, \cdot \rangle$ を、双対を表すペアリングと呼ぶ。

双対写像

[Placeholder 1: 改編予定]

線形空間の間の線形写像が与えられると、双対空間の間の線形写像を定めることができる。

線形空間 V, W の間の線形写像 $f\colon V \to W$ が与えられたとする。

W 上の線形汎関数を $\varphi \in W^*$ とすると、次のような関係になっている。

$$
\begin{tikzcd}
V \arrow[r, "f"] & W \arrow[d, "\varphi"] \\
 & \mathbb{R}
\end{tikzcd}
$$

このとき、合成写像 $\varphi \circ f\colon V \to \mathbb{R}$ を考えることができる。

$$
\begin{tikzcd}
V \arrow[r, "f"] \arrow[dr, "\varphi \circ f"'] & W \arrow[d, "\varphi"] \\
 & \mathbb{R}
\end{tikzcd}
$$

線形写像の合成もまた線形写像になるので、$\varphi \circ f$ は V 上の線形汎関数である。これを $f^*(\varphi) \in V^*$ と書くことにする。

$$
\begin{tikzcd}
V \arrow[r, "f"] \arrow[dr, "f^*(\varphi)"'] & W \arrow[d, "\varphi"] \\
 & \mathbb{R}
\end{tikzcd}
$$

$f^*(\varphi)$ は V 上の線形汎関数なので、$\boldsymbol{v} \in V$ を入力するとスカラーを返す。

合成写像の記法より、$(\varphi \circ f)(\boldsymbol{v})$ は、$\varphi(f(\boldsymbol{v}))$ とも書けるので、

$$
f^*(\varphi)(\boldsymbol{v}) = \varphi(f(\boldsymbol{v})) = \langle \varphi, f(\boldsymbol{v}) \rangle
$$

と整理できる。

また、f^* は W^* 上の線形汎関数 φ を入力として、V^* 上の線形汎関数 $f^*(\varphi)$ を返す線形写像である。

$$
\begin{tikzcd}
V \arrow[r, "f"] \arrow[dr, "f^*(\varphi)"' name=L] & W \arrow[d, "\varphi" name=P] \\
 & \mathbb{R}
\arrow[from=P, to=L, bend left=90, "f^*"]
\end{tikzcd}
$$

補足：なぜ f^* は線形写像といえるのか

$W^* = \mathrm{Hom}(W, \mathbb{R})$ における和とスカラー倍の定義から示すことができる。

$\varphi_1, \varphi_2 \in W^*$ に対して、

$$\begin{aligned} f^*(\varphi_1 + \varphi_2)(\boldsymbol{v}) &= (\varphi_1 + \varphi_2)(f(\boldsymbol{v})) \\ &= \varphi_1(f(\boldsymbol{v})) + \varphi_2(f(\boldsymbol{v})) \\ &= f^*(\varphi_1)(\boldsymbol{v}) + f^*(\varphi_2)(\boldsymbol{v}) \\ &= (f^*(\varphi_1) + f^*(\varphi_2))(\boldsymbol{v}) \end{aligned}$$

また、$c \in \mathbb{R}$ に対して、

$$\begin{aligned} f^*(c\varphi)(\boldsymbol{v}) &= (c\varphi)(f(\boldsymbol{v})) \\ &= c \cdot \varphi(f(\boldsymbol{v})) \\ &= c \cdot f^*(\varphi)(\boldsymbol{v}) \\ &= (cf^*(\varphi))(\boldsymbol{v}) \end{aligned}$$

以上より、

$$\begin{aligned} f^*(\varphi_1 + \varphi_2) &= f^*(\varphi_1) + f^*(\varphi_2) \\ cf^*(\varphi) &= f^*(c\varphi) \end{aligned}$$

となり、写像 f^* は線形性を満たすことがわかる。 ■

このように定まる線形写像 $f^* \colon W^* \to V^*$ を、f の双対写像という。

$$\begin{array}{cccc} f^* \colon & W^* & \longrightarrow & V^* \\ & \rotatebox{90}{$\in$} & & \rotatebox{90}{$\in$} \\ & \varphi & \longmapsto & f^*(\varphi) \end{array}$$

双対写像　V, W を線形空間とし、$f \colon V \to W$ を線形写像とするとき、f の双対写像 $f^* \colon W^* \to V^*$ を次のように定義する。

$$f^*(\varphi) \coloneqq \varphi \circ f \quad (\varphi \in W^*)$$

双対を表すペアリングの記法を用いると、次のように整理できる。

$$f^*(\varphi)(\boldsymbol{v}) = \langle f^*(\varphi), \boldsymbol{v} \rangle = \langle \varphi, f(\boldsymbol{v}) \rangle = \varphi(f(\boldsymbol{v})) \quad (\boldsymbol{v} \in V)$$

双対写像の表現行列

[Placeholder 2: 改編予定]

線形写像の双対写像の表現行列は、元の線形写像の表現行列の転置になる。

このことから、双対写像は**転置写像**とも呼ばれる。

双対写像の行列表現　V, W を有限次元の線形空間とし、$f\colon V \to W$ を線型写像とする。また、$\dim V = n, \dim W = m$ とする。

V の基底 $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$、$W$ の基底 $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_m$ を選び、これらの双対基底をそれぞれ $\phi_1, \ldots, \phi_n$、$\psi_1, \ldots, \psi_m$ とする。

このとき、$\{\boldsymbol{v}_i\}$、$\{\boldsymbol{w}_j\}$ に関する f の表現行列を A とすると、$\{\psi_j\}, \{\phi_i\}$ に関する f^* の表現行列は $A^\top$ によって与えられる。

証明

$f\colon V \to W$ に対して、その双対写像 $f^*\colon W^* \to V^*$ は次で定義される。

$$\begin{tikzcd} V \arrow[r, "f"] \arrow[dr, "\psi \circ f"'] & W \arrow[d, "\psi"] \\ & \mathbb{R} \end{tikzcd}$$

$$f^*(\psi) = \psi \circ f \quad (\psi \in W^*)$$

$f^*(\psi)$ は V 上の線形汎関数なので、任意の $\boldsymbol{v} \in V$ を入力するとスカラーを返す。

合成写像の記法より、$(\psi \circ f)(\boldsymbol{v})$ は、$\psi(f(\boldsymbol{v}))$ とも書けるので、

$$f^*(\psi)(\boldsymbol{v}) = \psi(f(\boldsymbol{v}))$$

が成り立つ。

一方、f の表現行列 A は次のように構成される。

$$f(\boldsymbol{v}_i) = \sum_{j=1}^{m} a_{ji}\boldsymbol{w}_j \quad (1 \leq i \leq n)$$

したがって、任意の i に対し、

$$\psi_k(f(\boldsymbol{v}_i)) = \psi_k\left(\sum_{j=1}^{m} a_{ji}\boldsymbol{w}_j\right) = \sum_{j=1}^{m} a_{ji}\psi_k(\boldsymbol{w}_j)$$

ここで、$\{\psi_k\}$ は $\{\boldsymbol{w}_j\}$ の双対基底なので、$\psi_k(\boldsymbol{w}_j) = \delta_{kj}$ より、

$$\psi_k(f(\boldsymbol{v}_i)) = a_{ki}$$

また、$f^*(\psi_k) \in V^*$ は V 上の線形汎関数なので、V の双対基底 $\{\phi_i\}$ の線形結合として表せる。

$$f^*(\psi_k) = \sum_{i=1}^{n} b_{ik}\phi_i$$

この係数 b_{ik} を並べた行列を B とすると、B は f^* の表現行列である。

このとき、

$$f^*(\psi_k)(\boldsymbol{v}_i) = \psi_k(f(\boldsymbol{v}_i)) = a_{ki}$$

であり、一方、

$$f^*(\psi_k)(\boldsymbol{v}_i) = \sum_{j=1}^{n} b_{jk}\phi_j(\boldsymbol{v}_i) = \sum_{j=1}^{n} b_{ji}\delta_{ij} = b_{ki}$$

でもあるから、$b_{ki} = a_{ki}$ が成り立つ。すなわち、

$$B = A^\top$$

である。 ■

例：縦ベクトルと横ベクトルによる線形写像

[Todo 55: まとめ直す]

A を $m \times n$ 型行列とする

縦ベクトルに A を左からかけることによって定まる線形写像を次のように表す

$$f_A \colon \mathbb{R}^n \to \mathbb{R}^m \ (\boldsymbol{v} \mapsto A\boldsymbol{v})$$

これと対照的に、横ベクトルに右から A をかけることによって定まる次の線形写像を**転置写像**と呼ぶ

$$f_A^* \colon {}^t\mathbb{R}^m \to {}^t\mathbb{R}^n \ (\phi \mapsto \phi A)$$

横ベクトル $\phi A \in {}^t\mathbb{R}^n$ は、次の合成写像の表現行列である

$$\mathbb{R}^n \xrightarrow{f_A} \mathbb{R}^m \xrightarrow{\phi} \mathbb{R}$$

転置写像と自然なペアリング　A を $m \times n$ 型行列とし、$\phi \in {}^t\mathbb{R}^m$, $\boldsymbol{v} \in \mathbb{R}^n$ に対して、

$$\langle f_A^*(\phi), \boldsymbol{v} \rangle = \langle \phi, f_A(\boldsymbol{v}) \rangle$$

証明

$$\begin{aligned}
\langle f_A^*(\phi), \boldsymbol{v} \rangle &= (\phi A)(\boldsymbol{v}) \\
&= \phi(A\boldsymbol{v}) \\
&= \phi(f_A(v)) \\
&= \langle \phi, f_A(\boldsymbol{v}) \rangle
\end{aligned}$$

より、目的の等式が得られる　■

転置写像と座標関数　A を $m \times n$ 型行列とし、$y_1, \ldots, y_m \in {}^t\mathbb{R}^m$ を ${}^t\mathbb{R}^m$ 上の座標関数とするとき、

$$f_A^*(y_i) = \sum_{j=1}^n a_{ij} x_j \quad (1 \leq i \leq m)$$

証明

行ベクトルとしての観点から見ると、$y_i = {}^t\boldsymbol{e}_i$ として、

$$f_A^*(y_i) = f_A^*({}^t\boldsymbol{e}_i) = {}^t\boldsymbol{e}_i A = \begin{pmatrix} a_{i1} & \cdots & a_{in} \end{pmatrix}$$

これは双対基底 $x_j = {}^t\boldsymbol{e}_j$ を用いて、

$$\begin{aligned} f_A^*(y_i) &= \begin{pmatrix} a_{i1} & \cdots & a_{in} \end{pmatrix} \\ &= \sum_{j=1}^n a_{ij} {}^t\boldsymbol{e}_j \\ &= \sum_{j=1}^n a_{ij} x_j \end{aligned}$$

とも書ける ■

転置写像の表現行列　A を $m \times n$ 型行列とするとき、基底 $\{y_1, \ldots y_m\}$, $\{x_1, \ldots, x_n\}$ に関する f_A^* の表現行列は ${}^tA$ である

証明

[Todo 56: よくわからない]

表現行列は、基底 $\{y_i\}$ の各元が、写像を通してどのような線形結合で $\{x_j\}$ に写されるかを記述したものである

すなわち、写像 f_A^* の表現行列を求めることは、

$$f_A^*(y_i) = \sum_{j=1}^n a_{ij} x_j \quad (1 \leq i \leq m)$$

において、係数 a_{ij} を行列に並べることである

ここで、$f_A^* \colon {}^t\mathbb{R}^m \to {}^t\mathbb{R}^n$ において、

- 定義域の基底は $\{y_1, \ldots, y_m\} \subset {}^t\mathbb{R}^m$

- 値域の基底は $\{x_1, \ldots, x_n\} \subset {}^t\mathbb{R}^n$

先ほど示した等式

$$f_A^*(y_i) = \sum_{j=1}^{n} a_{ij} x_j \quad (1 \leq i \leq m)$$

より、表現行列の第 i 列が、$f_A^*(y_i)$ の係数ベクトル

$$\begin{pmatrix} a_{i1} & \cdots & a_{in} \end{pmatrix}$$

を転置して縦ベクトルにしたものになる

第 27 章

部分空間の双対性

行空間と核空間の直交

$A = (a_{ij})$ を $m \times n$ 行列とする。
このとき、A の第 i 行ベクトル $\begin{pmatrix} a_{i1} & \cdots & a_{in} \end{pmatrix} \in {}^t\mathbb{R}^n$ は、次の線形汎関数 ϕ_i と同一視できる。

$$\phi_i = a_{i1}x_1 + \cdots + a_{in}x_n$$

この行ベクトル $\phi_i \in {}^t\mathbb{R}^n$ が張る空間 $\langle \phi_1, \ldots, \phi_n \rangle \subset {}^t\mathbb{R}^n$ を、A の行空間と呼び、Row A と書く。

ここで、x_j はベクトル $\boldsymbol{v} \in \mathbb{R}^n$ に対して、j 番目の成分を返す座標関数である。そこで、

$$\begin{aligned} \phi_i(\boldsymbol{v}) &= a_{i1}x_1(\boldsymbol{v}) + \cdots + a_{in}x_n(\boldsymbol{v}) \\ &= a_{i1}v_1 + \cdots + a_{in}v_n \end{aligned}$$

とみると、ϕ_i は $\boldsymbol{v}$ に作用して、A の第 i 行ベクトルと $\boldsymbol{v}$ の内積を返すことがわかる。

すると、$\phi_i(\boldsymbol{v})$ を縦に並べたものは、$A\boldsymbol{v}$ に一致する。

$$A\boldsymbol{v} = \begin{pmatrix} a_{11}v_1 + \cdots + a_{1n}v_n \\ \vdots \\ a_{m1}v_1 + \cdots + a_{mn}v_n \end{pmatrix} = \begin{pmatrix} \phi_1(\boldsymbol{v}) \\ \vdots \\ \phi_m(\boldsymbol{v}) \end{pmatrix}$$

このとき、$A\boldsymbol{v} = \boldsymbol{o}$ となる場合は、

$$\phi_i(\boldsymbol{v}) = 0 \quad (i = 1, \ldots, m)$$

が成り立つことになる。
すなわち、A のすべての行ベクトルに対して、$\boldsymbol{v}$ との内積が 0 になる。

このことから、A の行空間に属するベクトルと、$A\boldsymbol{v} = \boldsymbol{o}$ の解空間 $\mathrm{Ker}\ A$ に属するベクトル $\boldsymbol{v}$ は、互いに直交することがわかる。
よって、次の関係が成り立つ。

$$\mathrm{Ker}\ A = (\mathrm{Row}\ A)^{\perp}$$

また、$(\mathrm{Row}\ A)^{\perp}$ の直交補空間は $\mathrm{Row}\ A$ に一致することから、両辺の直交補空間をとると、次も成り立つ。

$$(\mathrm{Ker}\ A)^{\perp} = \mathrm{Row}\ A$$

⚓ **核空間と行空間の直交関係**　A の核空間と、A の行空間（行ベクトルが張る空間）は、直交補空間の関係にある。

$$\mathrm{Ker}\ A = (\mathrm{Row}\ A)^{\perp}$$
$$(\mathrm{Ker}\ A)^{\perp} = \mathrm{Row}\ A$$

この定理は、核空間と像空間との関係として言い換えることもできる。

A の像空間 $\mathrm{Im}\ A$ は、A の列ベクトルが張る空間（列空間）であった。
A を転置すると行と列が入れ替わるので、$A^{\top}$ の行空間は A の列空間に対応する。

よって、定理を次のように書き換えることができる。

⚓ **核空間と転置行列の像空間の直交関係**　A の核空間と、$A^{\top}$ の像空間は、直交補空間の関係にある。

$$\mathrm{Ker}\ A = (\mathrm{Im}\ A^{\top})^{\perp}$$
$$(\mathrm{Ker}\ A)^{\perp} = \mathrm{Im}\ A^{\top}$$

直交補空間から零化空間へ

さて、$\mathrm{Ker}\, A = (\mathrm{Row}\, A)^{\perp}$ という直交補空間の関係を導くにあたって、ここでは次のような議論を行った。

1. 横ベクトルと同一視できる線形汎関数を考える
2. 線形汎関数に縦ベクトルを作用させたものを内積とみなす
3. 内積が 0 になることから直交補空間の関係を導く

つまり、ここでは内積が定められている空間（計量空間）で議論を行ったわけだが、内積を考えずに、線形汎関数の集合（双対空間）だけで議論を行うこともできる。

直交補空間の概念を内積を使わずに拡張し、一般の線形空間上で定義したものが、次に述べる零化空間である。

零化空間

V を線形空間とし、その部分空間 $W \subset V$ を考える。
V の双対空間 V^*（線形汎関数の集合）の中で、「W の元に作用させると 0 になる」ような線形汎関数を集めた集合を零化空間（`annihilator`）という。

零化空間　V を n 次元の線形空間とする。W を V の部分空間とするとき、

$$W^{\perp} = \{\phi \in V^* \mid \forall \boldsymbol{w} \in W, \langle \phi, \boldsymbol{w} \rangle = 0\}$$

を W の零化空間という。

$\phi \in V^*$ が W のすべてのベクトル $\boldsymbol{w}$ に対して $\phi(\boldsymbol{w}) = 0$ となるとき、その ϕ は W を「全滅させてしまう（`annihilate`）」という意味で、零化空間は `annihilator` と呼ばれる。

零化空間は V^* の部分空間

V^* の中から、$\phi(\boldsymbol{w}) = 0$ を満たす元 $\phi \in V^*$ を取り出した集合が $W^\perp$ であるので、$W^\perp$ は V^* の部分空間である。

零化空間の双対空間への包含関係　V を n 次元の線形空間とし、W を V の部分空間とする。

このとき、W の零化空間 $W^\perp$ は、V の双対空間 V^* の部分空間である。

証明

[Todo 57:]

第 28 章

双線形形式

直積集合

直積集合とは、2 つの集合からそれぞれ要素を取り出してつくったペアをすべて集めた集合である。

$$
\begin{array}{c|cccc}
 & b_1 & b_2 & \cdots & b_n \\
\hline
a_1 & (a_1, b_1) & (a_1, b_2) & \cdots & (a_1, b_n) \\
a_2 & (a_2, b_1) & (a_2, b_2) & \cdots & (a_2, b_n) \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
a_m & (a_m, b_1) & (a_m, b_2) & \cdots & (a_m, b_n)
\end{array}
$$

ただし、ペアには順序があり、たとえば、(a, b) と (b, a) は異なるものとみなす。このような順序を考慮したペアを順序対という。

直積集合は、順序対の集合である。

直積集合　2 つの集合 A, B に対して、A と B の直積集合は次のように定義される。

$$A \times B = \{(a, b) \mid a \in A, b \in B\}$$

直積集合の例：座標平面 $\mathbb{R}^2$

たとえば、2 次元平面内の各点は、2 つの実数の組 (x, y) で表すことができる。

このとき、x 座標と y 座標はそれぞれ実数の集合 $\mathbb{R}$ の要素であり、平面上の点 (x, y) を集めたものが、直積集合 $\mathbb{R} \times \mathbb{R}$ となる。

この $\mathbb{R} \times \mathbb{R}$ を、$\mathbb{R}^2$ と表記することが多い。

2 次元平面を $\mathbb{R}^2$ と表記していたのは、このような直積集合の考え方が背景にある。

内積と双線形形式

$\mathbb{R}^n$ 上の内積は、2 つのベクトル $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^n$ のペアから、スカラー値 $\mathbb{R}$ を返す関数として捉えることができる。

このように内積を写像に見立てて、この写像を b とおくと、

$$b \colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$$

と表すことができる。

さらに、$\mathbb{R}^n$ 上の内積は次のような双線形性を満たすものだった。

i. $(\boldsymbol{u}_1 + \boldsymbol{u}_2, \boldsymbol{v}) = (\boldsymbol{u}_1, \boldsymbol{v}) + (\boldsymbol{u}_2, \boldsymbol{v})$
ii. $(\boldsymbol{u}, \boldsymbol{v}_1 + \boldsymbol{v}_2) = (\boldsymbol{u}, \boldsymbol{v}_1) + (\boldsymbol{u}, \boldsymbol{v}_2)$
iii. $(c\boldsymbol{u}, \boldsymbol{v}) = (\boldsymbol{u}, c\boldsymbol{v}) = c(\boldsymbol{u}, \boldsymbol{v})$

双線形性とは、2 つの引数それぞれに対して線形性があるという性質である。

線形性をもつ写像を線形写像として特別視したように、双線形性をもつ写像について考えてみよう。

双線形形式　U, V を線型空間とする。直積集合 $U \times V$ から $\mathbb{R}$ への写像 b が次の条件を満たすとき、b は $U \times V$ 上の双線形形式（`bilinear form`）であるという。

i. $b(\boldsymbol{u}_1 + \boldsymbol{u}_2, \boldsymbol{v}) = b(\boldsymbol{u}_1, \boldsymbol{v}) + b(\boldsymbol{u}_2, \boldsymbol{v})$

ii. $b(\boldsymbol{u}, \boldsymbol{v}_1 + \boldsymbol{v}_2) = b(\boldsymbol{u}, \boldsymbol{v}_1) + b(\boldsymbol{u}, \boldsymbol{v}_2)$

iii. $b(c\boldsymbol{u}, \boldsymbol{v}) = b(\boldsymbol{u}, c\boldsymbol{v}) = cb(\boldsymbol{u}, \boldsymbol{v})$

例：行列による双線形形式

行列による双線形形式の構成　A を $m \times n$ 型行列とするとき、$\boldsymbol{u} \in \mathbb{R}^m$, $\boldsymbol{v} \in \mathbb{R}^n$ に対して

$$b(\boldsymbol{u}, \boldsymbol{v}) = \boldsymbol{u}^T A \boldsymbol{v}$$

により $\mathbb{R}^m \times \mathbb{R}^n$ 上の双線形形式が得られる。

証明

和に対する双線形性（i）

行列の和に対して転置を分配できることを用いて、

$$\begin{aligned} b(\boldsymbol{u}_1 + \boldsymbol{u}_2, \boldsymbol{v}) &= (\boldsymbol{u}_1 + \boldsymbol{u}_2)^T A \boldsymbol{v} \\ &= \boldsymbol{u}_1^T A \boldsymbol{v} + \boldsymbol{u}_2^T A \boldsymbol{v} \\ &= b(\boldsymbol{u}_1, \boldsymbol{v}) + b(\boldsymbol{u}_2, \boldsymbol{v}) \end{aligned}$$

和に対する双線形性（ii）

(i) と同様に、

$$\begin{aligned} b(\boldsymbol{u}, \boldsymbol{v}_1 + \boldsymbol{v}_2) &= \boldsymbol{u}^T A (\boldsymbol{v}_1 + \boldsymbol{v}_2) \\ &= \boldsymbol{u}^T A \boldsymbol{v}_1 + \boldsymbol{u}^T A \boldsymbol{v}_2 \\ &= b(\boldsymbol{u}, \boldsymbol{v}_1) + b(\boldsymbol{u}, \boldsymbol{v}_2) \end{aligned}$$

スカラー倍に対する双線形性（iii）

行列の積に対する転置の性質と、スカラー（1×1 型行列）を転置しても変わらないことを用いて、

$$\begin{aligned} b(c\boldsymbol{u}, \boldsymbol{v}) &= (c\boldsymbol{u})^T A\boldsymbol{v} \\ &= c(\boldsymbol{u}^T A\boldsymbol{v}) \\ &= cb(\boldsymbol{u}, \boldsymbol{v}) \end{aligned}$$

以上より、b は $\mathbb{R}^m \times \mathbb{R}^n$ 上の双線形形式である。 ■

特に、$m = n$ で $A = E$ の場合、

$$b(\boldsymbol{u}, \boldsymbol{v}) = \boldsymbol{u}^T \boldsymbol{v}$$

となり、$\mathbb{R}^n$ 上の内積と一致する。

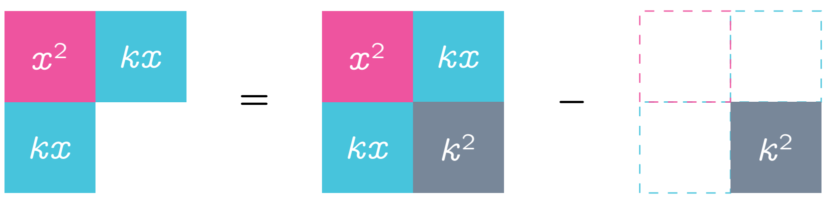
第 29 章

二次形式

二次式の平方完成

乗法公式 $(x+k)^2 = x^2 + 2kx + k^2$ を利用した次の形を、二次式の平方完成という

$$(x+k)^2 - k^2 = x^2 - 2kx$$



斉次二次式と行列

2 つの文字 x, y の斉次二次式は、一般に次のように表される

$$ax^2 + 2bxy + cy^2 \quad (a, b, c \neq 0)$$

この式は、次のように行列の積として表すことができる

$$\begin{aligned} ax^2 + 2bxy + cy^2 &= ax^2 + byx + bxy + cy^2 \\ &= (ax + by)x + (bx + cy)y \\ &= \begin{pmatrix} ax + by & bx + cy \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \\ &= \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} a & b \\ b & c \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \end{aligned}$$

すなわち、

$$A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}, \boldsymbol{x} = \begin{pmatrix} x \\ y \end{pmatrix}$$

とおくと、

$$ax^2 + 2bxy + cy^2 = {}^t\boldsymbol{x}A\boldsymbol{x}$$

ここで、A は実対称行列になっている

このような斉次二次式を一般化したものが、n 個の文字 $x_1, \ldots, x_n$ についての二次形式である

二次形式

n 個の変数 $x_1, \ldots, x_n$ の斉次二次式を二次形式という

各項の係数を a_{ij} とすると、一般の二次形式（n 変数斉次二次式）は次のように書くことができる

$$Q(\boldsymbol{x}) = \sum_{i=1}^{n} a_{ii}x_i^2 + 2\sum_{i<j} a_{ij}x_ix_j$$

ここで、各変数は可変、すなわち $x_ix_j = x_jx_i$ であるので、$i \neq j$ の場合は、$i < j$ を満たす項だけの和として書き、それを 2 倍している

あえて展開して書くと、次のようになる

$$Q(\boldsymbol{x}) = \sum_{i=1}^{n} a_{ii}x_{ii}x_{ii} + \sum_{i<j} a_{ij}x_ix_j + \sum_{i<j} a_{ji}x_jx_i$$

$i < j$ においては $x_i x_j = x_j x_i$ であり、その係数についても $a_{ij} = a_{ji}$ が成り立つので、行列 $A = (a_{ij})$ は対称行列である

$$a_{ij} = \begin{cases} a_{ii} & (i = j) \\ a_{ij} = a_{ji} & (i < j) \end{cases}$$

このように a_{ij} を定めた上で、$\sum$ を 1 つにまとめることができる

$$Q(\boldsymbol{x}) = \sum_{i,j=1}^{n} a_{ij} x_i x_j$$

二次形式の係数行列　二次形式は対称行列 $A = (a_{ij})$ によって、次のように表される

$$Q(\boldsymbol{x}) = \sum_{i,j=1}^{n} a_{ij} x_i x_j$$

このとき、A を二次形式 $Q(\boldsymbol{x})$ の係数行列という

i が A の行番号、j が列番号であるので、x_i は横ベクトル、x_j は縦ベクトルの成分である

$$\begin{aligned} Q(\boldsymbol{x}) &= \sum_{i,j=1}^{n} x_i a_{ij} x_j \\ &= \begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix} A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \end{aligned}$$

そこで、$\boldsymbol{x}$ を縦ベクトルとみるとき、二次形式 $Q(\boldsymbol{x})$ とその係数行列は次のような関係にある

$$Q(\boldsymbol{x}) = {}^t\boldsymbol{x} A \boldsymbol{x}$$

この関係を用いて、任意の対称行列 A から二次形式を作ることができる

$Q(\boldsymbol{x})$ から A を作り、A から $Q(\boldsymbol{x})$ を作ることができるので、n 変数の二次形式 $Q(\boldsymbol{x})$ と n 次の対称行列 A は対応し、さらにこの対応は一対一である

実二次形式の標準化

A が実対称行列であることから、A は適当な直交行列 P を用いて対角化できる

$$B = P^{-1}AP = \begin{pmatrix} \alpha_1 & & \\ & \ddots & \\ & & \alpha_n \end{pmatrix}$$

与えられた二次形式 $Q(\boldsymbol{x}) = {}^t\boldsymbol{x}A\boldsymbol{x}$ に対して、$\boldsymbol{y} = P^{-1}\boldsymbol{x}$、すなわち $\boldsymbol{x} = P\boldsymbol{y}$ という変数の変換を行うと、

$$\begin{aligned} {}^t\boldsymbol{x}A\boldsymbol{x} &= {}^t(P\boldsymbol{y})A(P\boldsymbol{y}) \\ &= {}^t\boldsymbol{y}{}^tPAP\boldsymbol{y} \\ &= {}^t\boldsymbol{y}(P^{-1}AP)\boldsymbol{y} \\ &= {}^t\boldsymbol{y}B\boldsymbol{y} \end{aligned}$$

直交行列の定義 ${}^tP = P^{-1}$

となるので、変数 $\boldsymbol{y}$ に関する係数行列は $B = P^{-1}AP$ である

B の形から、実際に ${}^t\boldsymbol{y}B\boldsymbol{y}$ を計算してみると、

$$\begin{aligned} {}^t\boldsymbol{y}B\boldsymbol{y} &= \begin{pmatrix} y_1 & \cdots & y_n \end{pmatrix} \begin{pmatrix} \alpha_1 & & \\ & \ddots & \\ & & \alpha_n \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \\ &= \alpha_1 y_1^2 + \cdots + \alpha_n y_n^2 \end{aligned}$$

となり、交叉項 y_iy_j $(i \neq j)$ が現れない形に書き換わったことがわかる

このような交叉項のない形を実二次形式の標準形という

実二次形式の直交対角化と標準形　実二次形式 $Q(\boldsymbol{x}) = {}^t\boldsymbol{x}A\boldsymbol{x}$ に対して、A を対角化する直交行列 P による座標変換 $\boldsymbol{x} = P\boldsymbol{y}$ を行えば、

$$Q(\boldsymbol{x}) = \alpha_1 y_1^2 + \cdots + \alpha_n y_n^2$$

という、変数 $\boldsymbol{y}$ に関する交叉項のない形（実二次形式の標準形）にできる

ここで、$\alpha_1, \ldots, \alpha_n$ は重複を含めて A の固有値と一致する

第 30 章

固有値とスペクトル分解

対称行列の固有ベクトルの正規直交化

n 次対称行列 A の固有値を $\lambda_1, ..., \lambda_n$ とする
対称行列の場合、異なる固有値に対する固有ベクトルは互いに直交する

n 個の固有値に重複するものがあると、重複する固有値に対しては対応する固有ベクトルも重複することになる
しかし、それらの任意の線型結合も同じ固有値に対する固有ベクトルとなるので、それらをグラム・シュミットの直交化法によって、互いに直交するように選ぶことができる

この結果、対称行列の固有ベクトル $\{\boldsymbol{u_1}, ..., \boldsymbol{u_n}\}$ を正規直交系となるように選ぶことができ、これは $\mathbb{R}^n$ の正規直交基底となる

対称行列のスペクトル分解

固有値と固有ベクトルの関係式

$$A\boldsymbol{u}_i = \lambda_i \boldsymbol{u}_i \quad (\boldsymbol{u}_i \neq \boldsymbol{0})$$

は、A が正規直交基底 $\{\boldsymbol{u_1}, \ldots, \boldsymbol{u_n}\}$ をそれぞれ $\lambda_1\boldsymbol{u_1}, \ldots, \lambda_n\boldsymbol{u_n}$ に写像することを意味する

よって、正規直交基底による表現行列の展開より、A は次のように表される

$$A = \lambda_1\boldsymbol{u}_1\boldsymbol{u}_1^\top + \cdots + \lambda_n\boldsymbol{u}_n\boldsymbol{u}_n^\top$$

このように、対称行列は、その固有値と固有ベクトルによって表すことができ、この式をスペクトル分解あるいは固有値分解と呼ぶ

対称行列のスペクトル分解　n 次対称行列 A は、その固有値と固有ベクトルによって表すことができる

$$A = \sum_{i=1}^{n} \lambda_i\boldsymbol{u}_i\boldsymbol{u}_i^\top$$

ここで、λ_i は固有値、$\boldsymbol{u}_i$ は対応する固有ベクトルの正規直交系である

各 $\boldsymbol{u}_i\boldsymbol{u}_i^\top$ は、各固有ベクトル $\boldsymbol{u}_i$ の方向（A の主軸）への射影行列である
よって、スペクトル分解とは、

> A を各主軸方向への射影行列の線形結合で表す

ものである

このことから、対称行列による空間の変換は、

1. 各点を主軸方向に射影する
2. それを固有値倍する
3. それらをすべての主軸にわたって足し合わせる

という操作の結果と解釈することができる

単位行列のスペクトル分解

単位行列 E も対称行列の一種である

単位行列の固有値はすべて 1 であるので、単位行列のスペクトル分解は次のように表される

$$E = \sum_{i=1}^{n} \boldsymbol{u}_i \boldsymbol{u}_i^\top$$

対称行列のランクと固有値

n 次対称行列 A の列の任意の線形結合

$$c_1\boldsymbol{a}_1 + \cdots + c_n\boldsymbol{a}_n = \begin{pmatrix} \boldsymbol{a}_1 & \cdots & \boldsymbol{a}_n \end{pmatrix} \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} = A\boldsymbol{c}$$

を考える

A の n 個の固有値のうち、0 でないものの個数を r とすれば、A のスペクトル分解の式において、$\lambda_{r+1}, \ldots, \lambda_n = 0$ とおいて、

$$\begin{aligned} A\boldsymbol{c} &= \lambda_1 \boldsymbol{u}_1 \boldsymbol{u}_1^\top \boldsymbol{c} + \cdots + \lambda_r \boldsymbol{u}_r \boldsymbol{u}_r^\top \boldsymbol{c} \\ &= \lambda_1 (\boldsymbol{u}_1^\top \boldsymbol{c}) \boldsymbol{u}_1 + \cdots + \lambda_r (\boldsymbol{u}_r^\top \boldsymbol{c}) \boldsymbol{u}_r \end{aligned}$$

すなわち、A の列の任意の線形結合は、互いに直交する $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_r$ の線形結合で書ける

互いに直交するベクトルは線型独立であることから、

- $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ の張る部分空間（線形結合の集合）の次元は r である
- n 本の列のうち、r 本しか線型独立ではない

ということがいえる

行列 A の n 本の列のうち、線型独立なものの個数を A のランクあるいは階数というので、次のことがいえる

⚓ **対称行列における階数と非零固有値の個数**　A を対称行列とするとき、$\mathrm{rank}(A)$ は、A の非零の固有値の個数に等しい

A は対称行列であるから、行についても同じことがいえる

スペクトル分解による対称行列の対角化

スペクトル分解の式を用いることで、対称行列の対角化について簡潔に議論できるようになる

対称行列 A のスペクトル分解の式

$$A = \lambda_1 \boldsymbol{u}_1 \boldsymbol{u}_1^\top + \cdots + \lambda_n \boldsymbol{u}_n \boldsymbol{u}_n^\top$$

は、次のように書き換えられる

$$\begin{aligned} A &= \begin{pmatrix} \lambda_1 \boldsymbol{u}_1 & \cdots & \lambda_n \boldsymbol{u}_n \end{pmatrix} \begin{pmatrix} \boldsymbol{u}_1^\top \\ \vdots \\ \boldsymbol{u}_n^\top \end{pmatrix} \\ &= \begin{pmatrix} \boldsymbol{u}_1 & \cdots & \boldsymbol{u}_n \end{pmatrix} \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} \begin{pmatrix} \boldsymbol{u}_1^\top \\ \vdots \\ \boldsymbol{u}_n^\top \end{pmatrix} \\ &= U \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} U^\top \end{aligned}$$

ここで、

$$U = \begin{pmatrix} \boldsymbol{u}_1 & \cdots & \boldsymbol{u}_n \end{pmatrix}$$

は、列が正規直交系をなすことから直交行列である

そして、U が直交行列であれば、その転置 $U^\top$ も直交行列である

それゆえ、直交行列の行も正規直交系をなす

A の式の両辺に左から $U^\top$、右から U をかけると、

$$U^\top A U = U^\top U \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} U^\top U$$

直交行列の定義 $U^\top U = E$ より、

$$U^\top A U = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$$

として、対称行列 A は、直交行列 U によって対角化できることがわかる

対称行列の逆行列

[Todo 58:]

第 31 章

特異値と特異値分解

正定値行列と半正定値行列

固有値がすべて 0 以上になる対称行列は、応用上さまざまな場面で現れる

- 半正定値行列：すべての固有値が非負（正または零）である対称行列
- 正定値行列：すべての固有値が正である対称行列

正定値行列　A をエルミート行列（対称行列）とし、任意のベクトル $\boldsymbol{x} \in \mathbb{C}^n$（$\boldsymbol{x} \in \mathbb{R}^n$）に対して、

$$(A\boldsymbol{x}, \boldsymbol{x}) > 0$$

が成り立つとき、A は正定値行列であるという

正定値性と固有値の正実性　エルミート行列 A が正定値行列であることと、A のすべての固有値が正の実数であることは同値である

証明

正定値行列 $\Longrightarrow$ 固有値が正

A の固有値を λ、対応する固有ベクトルを $\boldsymbol{x}$ とすると、

$$A\boldsymbol{x} = \lambda\boldsymbol{x}$$

両辺で $\boldsymbol{x}$ との内積をとると、

$$(A\boldsymbol{x}, \boldsymbol{x}) = \lambda(\boldsymbol{x}, \boldsymbol{x}) = \lambda\|\boldsymbol{x}\|^2$$

A が正定値行列であることから、$(A\boldsymbol{x}, \boldsymbol{x}) > 0$ が成り立ち、

$$\lambda\|\boldsymbol{x}\|^2 > 0$$

ここで、固有ベクトルは零ベクトルではないので、$\|\boldsymbol{x}\|^2 > 0$ である

よって、$\lambda\|\boldsymbol{x}\|^2 > 0$ の両辺を $\|\boldsymbol{x}\|^2$ で割ることにより、

$$\lambda > 0$$

が得られる ■

固有値が正 $\Longrightarrow$ 正定値行列

A の固有値を $\lambda_1, \ldots, \lambda_n > 0$ とする

A はエルミート行列であることから、ユニタリ行列 U を用いて次のように対角化できる

$$A = UDU^{-1} = UDU^* = U \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} U^*$$

随伴による内積の表現より、

$$(A\boldsymbol{x}, \boldsymbol{x}) = \boldsymbol{x}^* A\boldsymbol{x} = \boldsymbol{x}^* UDU^*\boldsymbol{x}$$

ここで、$\boldsymbol{y} = U^*\boldsymbol{x}$ とおくと、

$$\boldsymbol{y}^* = (U^*\boldsymbol{x})^* = \boldsymbol{x}^* U$$

となるので、次のように書き換えられる

$$(A\boldsymbol{x}, \boldsymbol{x}) = \boldsymbol{y}^* D\boldsymbol{y} = (D\boldsymbol{y}, \boldsymbol{y})$$

左辺の内積を計算すると、

$$\begin{aligned}(D\boldsymbol{y}, \boldsymbol{y}) &= \begin{pmatrix} \lambda_1 y_1 & & \\ & \ddots & \\ & & \lambda_n y_n \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \\ &= \lambda_1 |y_1|^2 + \cdots + \lambda_n |y_n|^2\end{aligned}$$

$\lambda_1, \ldots, \lambda_n > 0$ であることから、すべての項が正になるので、

$$(A\boldsymbol{x}, \boldsymbol{x}) = (D\boldsymbol{y}, \boldsymbol{y}) > 0$$

よって、A は正定値行列である ■

半正定値行列は、正定値行列の条件に等号を含むようにしたものである

半正定値行列　A をエルミート行列（対称行列）とし、任意のベクトル $\boldsymbol{x} \in \mathbb{C}^n$（$\boldsymbol{x} \in \mathbb{R}^n$）に対して、

$$(A\boldsymbol{x}, \boldsymbol{x}) \geq 0$$

が成り立つとき、A は半正定値行列であるという

半正定値性と固有値の非負実性　エルミート行列 A が半正定値行列であることと、A のすべての固有値が非負の実数であることは同値である

対称行列を構成する行列積

スペクトル分解は対称行列に対するものだったが、これを任意の長方行列に拡張したものが特異値分解である

対称行列から任意の行列へ議論を拡張するにあたって、次の定理が重要となる

⚓ **自身の随伴行列との積で構成されるエルミート行列**　A を任意の複素行列（長方行列）とするとき、A^*A および AA^* はエルミート行列である

証明

積をエルミート行列にすると順序が入れ替わることに注意して、

$$(A^*A)^* = A^*(A^*)^* = A^*A$$

よって、A^*A はエルミート行列である

同様に、

$$(AA^*)^* = (A^*)^*A^* = AA^*$$

よって、AA^* もエルミート行列である　■

A を実行列とすれば、次が成り立つ

⚓ **自身の転置行列との積で構成される対称行列**　A を任意の実行列（長方行列）とするとき、$A^\top A$ および $AA^\top$ は対称行列である

A^*A および AA^* という形の行列には、さらに固有値に関する重要な性質がある

⚓ **自身の随伴行列との積で構成される半正値行列**　任意の行列 A に対して、AA^* および A^*A はともに半正値行列である

証明

エルミート行列 AA^* の固有ベクトルを $\boldsymbol{u}$ とし、その固有値を $\lambda \in \mathbb{C}$ とすると、

$$AA^*\boldsymbol{u} = \lambda\boldsymbol{u}$$

両辺で $\boldsymbol{u}$ との内積をとると、

$$(\boldsymbol{u}, AA^*\boldsymbol{u}) = \lambda(\boldsymbol{u}, \boldsymbol{u}) = \lambda\|\boldsymbol{u}\|^2$$

この左辺は、随伴公式を用いて、

$$\begin{aligned} (\boldsymbol{u}, AA^*\boldsymbol{u}) &= (\boldsymbol{u}, A(A^*\boldsymbol{u})) \\ &= (A^*\boldsymbol{u}, A^*\boldsymbol{u}) \\ &= \|A^*\boldsymbol{u}\|^2 \geq 0 \end{aligned}$$

外側の A に
随伴公式を適用

となるので、

$$\|A^*\boldsymbol{u}\|^2 = \lambda\|\boldsymbol{u}\|^2 \geq 0$$

ここで、固有ベクトルは零ベクトルではないので、$\|\boldsymbol{u}\|^2 > 0$ である

よって、$\lambda\|\boldsymbol{u}\|^2 \geq 0$ の両辺を $\|\boldsymbol{u}\|^2$ で割ることにより、

$$\lambda \geq 0$$

が得られる

A^*A についても同様に、

$$(\boldsymbol{u}, A^*A\boldsymbol{u}) = (A\boldsymbol{u}, A\boldsymbol{u}) = \|A\boldsymbol{u}\|^2 \geq 0$$

から、$\lambda \geq 0$ が得られる ■

特異値と左右特異ベクトルの対応関係 A を O でない任意の行列とするとき、$A^\top A$ と $AA^\top$ は共通の正の固有値 σ^2 を持ち、それぞれの固有ベクトル $\boldsymbol{u}, \boldsymbol{v}$ は次の関係を満たす

$$A\boldsymbol{v} = \sigma\boldsymbol{u}, \quad A^\top\boldsymbol{u} = \sigma\boldsymbol{v}$$

証明

$AA^\top$ の固有値が σ^2 と仮定した場合

$AA^\top$ の固有値が非負の固有値 σ^2 を持ち、対応する固有ベクトルが $\boldsymbol{u}$ であるとすると、

$$AA^\top \boldsymbol{u} = \sigma^2 \boldsymbol{u}$$

この両辺に左から $A^\top$ をかけて、

$$A^\top AA^\top \boldsymbol{u} = A^\top \sigma^2 \boldsymbol{u}$$

ここで、$\boldsymbol{v} = \dfrac{A^\top \boldsymbol{u}}{\sigma}$ とおくと、$A^\top \boldsymbol{u} = \sigma \boldsymbol{v}$ となるので、

$$\begin{aligned} A^\top A\sigma \boldsymbol{v} &= \sigma^3 \boldsymbol{v} \\ A^\top A\boldsymbol{v} &= \sigma^2 \boldsymbol{v} \end{aligned}$$

よって、σ^2 は $A^\top A$ の固有値でもあり、対応する固有ベクトル $\boldsymbol{v}$ は

$$A^\top \boldsymbol{u} = \sigma \boldsymbol{v}$$

を満たす ■

$A^\top A$ の固有値が σ^2 と仮定した場合

$A^\top A$ の固有値が非負の固有値 σ^2 を持ち、対応する固有ベクトルが $\boldsymbol{v}$ であるとすると、

$$A^\top A\boldsymbol{v} = \sigma^2 \boldsymbol{v}$$

この両辺に左から A をかけて、

$$AA^\top A\boldsymbol{v} = A\sigma^2 \boldsymbol{v}$$

ここで、$\boldsymbol{u} = \dfrac{A\boldsymbol{v}}{\sigma}$ とおくと、$A\boldsymbol{v} = \sigma \boldsymbol{u}$ となるので、

$$\begin{aligned} AA^\top \sigma \boldsymbol{u} &= \sigma^3 \boldsymbol{u} \\ AA^\top \boldsymbol{u} &= \sigma^2 \boldsymbol{u} \end{aligned}$$

よって、σ^2 は $AA^\top$ の固有値でもあり、対応する固有ベクトル $\boldsymbol{u}$ は

$$A\boldsymbol{v} = \sigma \boldsymbol{u}$$

を満たす ■

特異値と特異ベクトル

スペクトル分解の拡張である特異値分解では、任意の行列がその特異値と特異ベクトルによって表せる

特異値と特異ベクトル　零行列ではない任意の $m \times n$ 行列 A に対して、

$$A\boldsymbol{v} = \sigma\boldsymbol{u}, \quad A^\top\boldsymbol{u} = \sigma\boldsymbol{v}$$

となる正の数 σ を特異値と呼び、

- 左特異ベクトル：m 次元ベクトル $\boldsymbol{u}\,(\neq \boldsymbol{0})$
- 右特異ベクトル：n 次元ベクトル $\boldsymbol{v}\,(\neq \boldsymbol{0})$

を合わせて特異ベクトルと呼ぶ

特異ベクトルと固有ベクトルの関係

特異値と特異ベクトルの関係式

$$A\boldsymbol{v} = \sigma\boldsymbol{u}, \quad A^\top\boldsymbol{u} = \sigma\boldsymbol{v}$$

において、第 1 式の両辺に $A^\top$ を左からかけると、

$$\begin{aligned} A^\top A\boldsymbol{v} &= \sigma A^\top\boldsymbol{u} \\ &= \sigma^2\boldsymbol{v} \end{aligned} \quad \text{第 2 式を代入}$$

また、第 2 式の両辺に A を左からかけると、

$$\begin{aligned} AA^\top\boldsymbol{u} &= \sigma A\boldsymbol{v} \\ &= \sigma^2\boldsymbol{u} \end{aligned} \quad \text{第 1 式を代入}$$

得られた結果をまとめると、

$$AA^\top\boldsymbol{u} = \sigma^2\boldsymbol{u}, \quad A^\top A\boldsymbol{v} = \sigma^2\boldsymbol{v}$$

ここで、A は任意の長方行列だが、$AA^\top$ と $A^\top A$ は対称行列となる

すなわち、

- 左特異ベクトル $\boldsymbol{u}$ は m 次対称行列 $AA^\top$ の固有ベクトル
- 右特異ベクトル $\boldsymbol{v}$ は n 次対称行列 $A^\top A$ の固有ベクトル

であり、特異値の 2 乗 σ^2 は $AA^\top, A^\top A$ 共通の固有値である

特異ベクトルの正規直交化

A の特異値を $\sigma_1 \geq \cdots \geq \sigma_r > 0$ とする

ここで、重複があってもよい

対応する r 本の左特異ベクトル $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_r$ と r 本の右特異ベクトル $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_r$ は、どちらも対称行列の固有ベクトルであるから、それぞれを正規直交系に選ぶことができる

特異値分解

k 本の左特異ベクトルの正規直交系 $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k$ を拡張して、$\mathbb{R}^m$ の正規直交基底 $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k, \boldsymbol{u}_{k+1}, \ldots, \boldsymbol{u}_m$ が定義できる。

同様に、k 本の右特異ベクトルの正規直交系 $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k$ を拡張して、$\mathbb{R}^n$ の正規直交基底 $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k, \boldsymbol{v}_{k+1}, \ldots, \boldsymbol{v}_n$ が定義できる。

$\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n$ と $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ はそれぞれ $AA^\top$ と $A^\top A$ の固有ベクトルであり、これらに対応する共通の固有値を $\lambda_1, \ldots, \lambda_n$ とおく。

$AA^\top$ および $A^\top A$ は半正定値行列であるので、その固有値はすべて零か正の数である。

また、$AA^\top$ および $A^\top A$ は対称行列であり、対称行列の階数 r は非零の固有値の個数に等しい。

n 個の固有値のうち、r 個ある正の固有値は特異値の条件を満たすので、

- $\lambda_1, \ldots, \lambda_r$ は特異値（正の固有値）$\sigma_1, \ldots, \sigma_r$
- $\lambda_{r+1}, \ldots, \lambda_n$ は零の固有値

とする。

特異値が r 個あることから、左特異ベクトルと特異値の組の個数、右特異ベクトルと特異値の組の個数は、どちらも r であることがいえる。

$$k = r$$

以上の議論をまとめると、

$$AA^\top \boldsymbol{u}_i = \begin{cases} \sigma_i \boldsymbol{u}_i & (i = 1, \ldots, r) \\ \boldsymbol{o} & (i = r+1, \ldots, m) \end{cases}$$

$$A^\top A \boldsymbol{v}_i = \begin{cases} \sigma_i \boldsymbol{v} & (i = 1, \ldots, r) \\ \boldsymbol{o} & (i = r+1, \ldots, n) \end{cases}$$

ここで、$i = 1, \ldots, r$ の範囲に限っては、特異値と特異ベクトルの関係より、

$$A^\top \boldsymbol{u}_i = \sigma_i \boldsymbol{v}_i$$
$$A \boldsymbol{v}_i = \sigma_i \boldsymbol{u}_i$$

という形で書ける。

$i > r$ の場合についても同じ形で書くために、次の定理を示す。

行列積による零化

i. $AA^\top \boldsymbol{u} = \boldsymbol{o} \Longrightarrow A^\top \boldsymbol{u} = \boldsymbol{o}$

ii. $A^\top A \boldsymbol{v} = \boldsymbol{o} \Longrightarrow A \boldsymbol{v} = \boldsymbol{o}$

証明

(i) $AA^\top \boldsymbol{u} = \boldsymbol{o}$ について

$AA^\top \boldsymbol{u} = \boldsymbol{o}$ の両辺で $\boldsymbol{u}$ との内積をとって、

$$(\boldsymbol{u}, AA^\top \boldsymbol{u}) = 0$$

このとき、左辺は、

$$\begin{aligned} (\boldsymbol{u}, AA^\top \boldsymbol{u}) &= (\boldsymbol{u}, A(A^\top \boldsymbol{u})) \\ &= (A^\top \boldsymbol{u}, A^\top \boldsymbol{u}) \\ &= \|A^\top \boldsymbol{u}\|^2 \end{aligned}$$

外側の A に
随伴公式を適用

と変形できるので、

$$\|A^\top \boldsymbol{u}\|^2 = 0$$

が成り立つ。

ここで、内積の正値性

$$\|A^\top \boldsymbol{u}\|^2 = (A^\top \boldsymbol{u}, A^\top \boldsymbol{u}) \geq 0$$

において、等号が成立するのは、

$$A^\top \boldsymbol{u} = \boldsymbol{o}$$

の場合のみである。 ■

(ii) $A^\top A\boldsymbol{v} = \boldsymbol{o}$ について

$A^\top A\boldsymbol{v} = \boldsymbol{o}$ の両辺で $\boldsymbol{v}$ との内積をとって、

$$(\boldsymbol{v}, A^\top A\boldsymbol{v}) = 0$$

このとき、左辺は、

$$(\boldsymbol{v}, A^\top A\boldsymbol{v}) = (A\boldsymbol{v}, A\boldsymbol{v}) = \|A\boldsymbol{v}\|^2$$

と変形できるので、

$$\|A\boldsymbol{v}\|^2 = 0$$

が成り立つ。

ここで、内積の正値性

$$\|A\boldsymbol{v}\|^2 = (A\boldsymbol{v}, A\boldsymbol{v}) \geq 0$$

において、等号が成立するのは、

$$A\boldsymbol{v} = \boldsymbol{o}$$

の場合のみである。 ■

この定理を用いると、

$$A\boldsymbol{v}_i = \begin{cases} \sigma_i \boldsymbol{u}_i & (i = 1, \ldots, r) \\ \boldsymbol{o} & (i = r+1, \ldots, m) \end{cases}$$

$$A^\top \boldsymbol{u}_i = \begin{cases} \sigma_i \boldsymbol{v}_i & (i = 1, \ldots, r) \\ \boldsymbol{o} & (i = r+1, \ldots, n) \end{cases}$$

とまとめられる。

このことから、次に示すように、任意の行列は、その特異値と特異ベクトルによって表すことができる。これを特異値分解と呼ぶ。

A の特異値分解

A は $\mathbb{R}^n$ の正規直交基底 $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ をそれぞれ

$$\sigma_1 \boldsymbol{u}_1, \ldots, \sigma_r \boldsymbol{u}_r, \boldsymbol{o}, \ldots, \boldsymbol{o}$$

に写像するから、正規直交基底による表現行列の展開より、A は

$$A = \sigma_1 \boldsymbol{u}_1 \boldsymbol{v}_1^\top + \cdots + \sigma_r \boldsymbol{u}_r \boldsymbol{v}_r^\top \quad (\sigma_1 \geq \cdots \geq \sigma_r > 0)$$

と表すことができる。

$A^\top$ の特異値分解

同様に、$A^\top$ は $\mathbb{R}^m$ の正規直交基底 $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m\}$ をそれぞれ

$$\sigma_1 \boldsymbol{v}_1, \ldots, \sigma_r \boldsymbol{v}_r, \boldsymbol{o}, \ldots, \boldsymbol{o}$$

に写像するから、$A^\top$ は

$$A^\top = \sigma_1 \boldsymbol{v}_1 \boldsymbol{u}_1^\top + \cdots + \sigma_r \boldsymbol{v}_r \boldsymbol{u}_r^\top \quad (\sigma_1 \geq \cdots \geq \sigma_r > 0)$$

と表すことができる。

列空間と行空間の正規直交基底

ここでは、A の特異値の個数は A の階数に等しく、特異ベクトルは行空間と列空間の正規直交基底を成すことを示す。

列空間　行列 A の n 本の列の張る $\mathbb{R}^m$ の部分空間を A の**列空間**という。

行空間　行列 A の m 本の行の張る $\mathbb{R}^n$ の部分空間を A の**行空間**という。

列空間を $\mathcal{U}$、行空間を $\mathcal{V}$ と表記することにする。

列空間の正規直交基底

A の列 $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ の任意の線形結合を考える。

$$c_1\boldsymbol{a}_1 + \cdots + c_n\boldsymbol{a}_n = \begin{pmatrix} \boldsymbol{a}_1 & \cdots & \boldsymbol{a}_n \end{pmatrix} \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} = A\boldsymbol{c}$$

ここで、A の特異値分解の式

$$A = \sigma_1 \boldsymbol{u}_1 \boldsymbol{v}_1^\top + \cdots + \sigma_r \boldsymbol{u}_r \boldsymbol{v}_r^\top$$

の両辺に $\boldsymbol{c}$ を右からかけると、

$$\begin{aligned} A\boldsymbol{c} &= \sigma_1 \boldsymbol{u}_1 \boldsymbol{v}_1^\top \boldsymbol{c} + \cdots + \sigma_r \boldsymbol{u}_r \boldsymbol{v}_r^\top \boldsymbol{c} \\ &= \sigma_1 (\boldsymbol{v}_1, \boldsymbol{c}) \boldsymbol{u}_1 + \cdots + \sigma_r (\boldsymbol{v}_r, \boldsymbol{c}) \boldsymbol{u}_r \end{aligned}$$

すなわち、A の列の任意の線形結合 $A\boldsymbol{c}$ は、互いに直交する左特異ベクトル $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_r$ の線形結合で書ける。

よって、$\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ の張る列空間 $\mathcal{U}$ は、左特異ベクトル $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_r$ を正規直交基底とする r 次元部分空間である。

このことから、

r 本の列のみが線型独立である

ということもいえる。

行空間の正規直交基底

A の行は $A^\top$ の列であるので、同様の議論を $A^\top$ に対して行う。

$A^\top$ の特異値分解の式

$$A^\top = \sigma_1 \boldsymbol{v}_1 \boldsymbol{u}_1^\top + \cdots + \sigma_r \boldsymbol{v}_r \boldsymbol{u}_r^\top$$

の両辺に $\boldsymbol{c}$ を右からかけることで、

$$\begin{aligned} A^\top \boldsymbol{c} &= \sigma_1 \boldsymbol{v}_1 \boldsymbol{u}_1^\top \boldsymbol{c} + \cdots + \sigma_r \boldsymbol{v}_r \boldsymbol{u}_r^\top \boldsymbol{c} \\ &= \sigma_1 (\boldsymbol{u}_1, \boldsymbol{c}) \boldsymbol{v}_1 + \cdots + \sigma_r (\boldsymbol{u}_r, \boldsymbol{c}) \boldsymbol{v}_r \end{aligned}$$

となり、A の行の任意の線形結合 $A^\top \boldsymbol{c}$ は、互いに直交する右特異ベクトル $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_r$ の線形結合で書けることがわかる。

よって、A の行の張る行空間 $\mathcal{V}$ は、右特異ベクトル $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_r$ を正規直交基底とする r 次元部分空間である。

このことから、

r 本の行のみが線型独立である

ということもいえる。

特異値の個数と特異ベクトルによる基底

以上の議論から、次のことがわかる。

特異値の個数と階数　行列 A の階数 r は、A の特異値の個数に等しい。

特異ベクトルと列空間・行空間の正規直交基底　左特異ベクトル $\{\boldsymbol{u}_i\}_{i=1}^r$ と右特異ベクトル $\{\boldsymbol{v}_i\}_{i=1}^r$ は、それぞれ列空間 $\mathcal{U}$ と行空間 $\mathcal{V}$ の正規直交基底を成す。

列空間と行空間への射影

特異ベクトルが列空間・行空間の正規直交基底をなすことから、これらを用いて射影行列の展開式を考えることができる。

$\mathbb{R}^m$ の列空間 $\mathcal{U}$ への射影行列を $P_{\mathcal{U}}$、$\mathbb{R}^n$ の行空間 $\mathcal{V}$ への射影行列を $P_{\mathcal{V}}$ とすると、

$$
\begin{aligned}
P_{\mathcal{U}} &= \sum_{i=1}^{r} \boldsymbol{u}_i \boldsymbol{u}_i^\top \\
P_{\mathcal{V}} &= \sum_{i=1}^{r} \boldsymbol{v}_i \boldsymbol{v}_i^\top
\end{aligned}
$$

ここで、$\boldsymbol{u}_i$ は列空間 $\mathcal{U}$ の正規直交基底であることから、$\boldsymbol{u}_i \in \mathcal{U}$ である。

よって、部分空間への射影で議論したように、$P_{\mathcal{U}}$ は列空間 $\mathcal{U}$ の元をそのまま写すので、

$$
P_{\mathcal{U}} \boldsymbol{u}_i = \boldsymbol{u}_i
$$

このことから、A の特異値分解の式

$$
A = \sigma_1 \boldsymbol{u}_1 \boldsymbol{v}_1^\top + \cdots + \sigma_r \boldsymbol{u}_r \boldsymbol{v}_r^\top
$$

の両辺に左から $P_{\mathcal{U}}$ をかけても変化しないことが次のように導かれる。

$$
\begin{aligned}
P_{\mathcal{U}} A &= P_{\mathcal{U}} \left(\sigma_1 \boldsymbol{u}_1 \boldsymbol{v}_1^\top + \cdots + \sigma_r \boldsymbol{u}_r \boldsymbol{v}_r^\top \right) \\
&= \sigma_1 P_{\mathcal{U}} \boldsymbol{u}_1 \boldsymbol{v}_1^\top + \cdots + \sigma_r P_{\mathcal{U}} \boldsymbol{u}_r \boldsymbol{v}_r^\top \\
&= \sigma_1 \boldsymbol{u}_1 \boldsymbol{v}_1^\top + \cdots + \sigma_r \boldsymbol{u}_r \boldsymbol{v}_r^\top \\
&= A
\end{aligned}
$$

行についても同様に、$P_{\mathcal{V}} \boldsymbol{v}_i = \boldsymbol{v}_i$ であるから、$A^\top$ の特異値分解の式

$$
A^\top = \sigma_1 \boldsymbol{v}_1 \boldsymbol{u}_1^\top + \cdots + \sigma_r \boldsymbol{v}_r \boldsymbol{u}_r^\top
$$

の両辺に右から $P_{\mathcal{V}}$ をかけても変化しない。

$$
\begin{aligned}
A P_{\mathcal{V}} &= \left(\sigma_1 \boldsymbol{v}_1 \boldsymbol{u}_1^\top + \cdots + \sigma_r \boldsymbol{v}_r \boldsymbol{u}_r^\top \right) P_{\mathcal{V}} \\
&= \sigma_1 \boldsymbol{v}_1 P_{\mathcal{V}} \boldsymbol{u}_1^\top + \cdots + \sigma_r \boldsymbol{v}_r P_{\mathcal{V}} \boldsymbol{u}_r^\top \\
&= \sigma_1 \boldsymbol{v}_1 \boldsymbol{u}_1^\top + \cdots + \sigma_r \boldsymbol{v}_r \boldsymbol{u}_r^\top \\
&= A^\top
\end{aligned}
$$

以上をまとめると、次の式が成り立つ。

$$
\begin{aligned}
P_{\mathcal{U}} A &= A \\
A P_{\mathcal{V}} &= A^\top
\end{aligned}
$$

特異値分解の行列表記

特異値分解の式

$$
A = \sigma_1 \boldsymbol{u}_1 \boldsymbol{v}_1^\top + \cdots + \sigma_r \boldsymbol{u}_r \boldsymbol{v}_r^\top
$$

は、次のように行列で表すこともできる。

$$
\begin{aligned}
A &= \begin{pmatrix} \sigma_1 \boldsymbol{u}_1 & \cdots & \sigma_r \boldsymbol{u}_r \end{pmatrix} \begin{pmatrix} \boldsymbol{v}_1^\top \\ \vdots \\ \boldsymbol{v}_r^\top \end{pmatrix} \\
&= \begin{pmatrix} \boldsymbol{u}_1 & \cdots & \boldsymbol{u}_r \end{pmatrix} \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{pmatrix} \begin{pmatrix} \boldsymbol{v}_1^\top \\ \vdots \\ \boldsymbol{v}_r^\top \end{pmatrix}
\end{aligned}
$$

ここで、

$$
U_r = \begin{pmatrix} \boldsymbol{u}_1 & \cdots & \boldsymbol{u}_r \end{pmatrix}, \quad V_r = \begin{pmatrix} \boldsymbol{v}_1 & \cdots & \boldsymbol{v}_r \end{pmatrix},
$$
$$
\Sigma_r = \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{pmatrix}
$$

とおくと、

$$
A = U_r \Sigma_r V_r^\top
$$

という、行列 A を 3 つの行列を用いて分解した式として表すことができる。

この式は、A の**簡約された特異値分解**と呼ばれる。

特異値分解のより一般的な形

先ほどの式が「簡約された」特異値分解と呼ばれるということは、簡約する前のより一般的な形も考えられるということである。

元々、特異値分解の式は、A が $\mathbb{R}^n$ の正規直交基底 $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ をそれぞれ

$$\sigma_1 \boldsymbol{u}_1, \ldots, \sigma_r \boldsymbol{u}_r, \boldsymbol{o}, \ldots, \boldsymbol{o}$$

に写像することから導かれた。

そこで、$r+1$ 番以降の項も省略せずに書くと、

$$\begin{aligned}
A &= \sigma_1 \boldsymbol{u}_1 \boldsymbol{v}_1^\top + \cdots + \sigma_r \boldsymbol{u}_r \boldsymbol{v}_r^\top + \boldsymbol{o}\boldsymbol{v}_{r+1}^\top + \cdots + \boldsymbol{o}\boldsymbol{v}_n^\top \\
&= \begin{pmatrix} \sigma_1 \boldsymbol{u}_1 & \cdots & \sigma_r \boldsymbol{u}_r & \boldsymbol{o} & \cdots & \boldsymbol{o} \end{pmatrix} \begin{pmatrix} \boldsymbol{v}_1^\top \\ \vdots \\ \boldsymbol{v}_r^\top \\ \boldsymbol{v}_{r+1}^\top \\ \vdots \\ \boldsymbol{v}_n^\top \end{pmatrix} \\
&= \begin{pmatrix} \boldsymbol{u}_1 & \cdots & \boldsymbol{u}_m \end{pmatrix} \left(\begin{array}{ccc|c} \sigma_1 & & & \\ & \ddots & & O \\ & & \sigma_r & \\ \hline & O & & O \end{array}\right) \begin{pmatrix} \boldsymbol{v}_1^\top \\ \vdots \\ \boldsymbol{v}_n^\top \end{pmatrix}
\end{aligned}$$

補足

この式変形は、ブロック行列の積の計算に基づいている。

たとえば、

$$A = \begin{pmatrix} A_{11} & A_{12} \end{pmatrix}, \quad B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}$$

とおけば、

$$AB = \begin{pmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \end{pmatrix}$$

のように計算できる。

そこで、

$$U_1 = \begin{pmatrix} \boldsymbol{u}_1 & \cdots & \boldsymbol{u}_r \end{pmatrix}, \quad U_2 = \begin{pmatrix} \boldsymbol{u}_{r+1} & \cdots & \boldsymbol{u}_m \end{pmatrix},$$
$$D = \mathrm{diag}(\sigma_1, \ldots, \sigma_r)$$

とおくと、

$$\begin{aligned}\begin{pmatrix} U_1 & U_2 \end{pmatrix}\begin{pmatrix} D & O \\ O & O \end{pmatrix} &= \begin{pmatrix} U_1D + U_2O & U_1O + U_2O \end{pmatrix} \\ &= \begin{pmatrix} U_1D & O \end{pmatrix} \\ &= \begin{pmatrix} \sigma_1\boldsymbol{u}_1 & \cdots & \sigma_r\boldsymbol{u}_r & \boldsymbol{o} & \cdots & \boldsymbol{o} \end{pmatrix}\end{aligned}$$

という式変形が確かめられる。

ここで、

$$U = \begin{pmatrix} \boldsymbol{u}_1 & \cdots & \boldsymbol{u}_m \end{pmatrix}, \quad V = \begin{pmatrix} \boldsymbol{v}_1 & \cdots & \boldsymbol{v}_n \end{pmatrix},$$

$$\Sigma = \begin{array}{c} \begin{array}{cc} \overleftrightarrow{\quad r \quad} & \overleftrightarrow{\ n-r\ } \end{array} \\ \left(\begin{array}{ccc|c} \sigma_1 & & & \\ & \ddots & & O \\ & & \sigma_r & \\ \hline & O & & O \end{array}\right)\begin{array}{l} \updownarrow r \\ \updownarrow m-r \end{array} \end{array}$$

とおくと、

$$A = U\Sigma V^\top$$

と表せる。この式を A の特異値分解と呼ぶ。

簡約された特異値分解は、特異値分解において U の余計な列と Σ の零行を省いたものだといえる。

第 32 章

行列ノルム

さまざまな距離

データ量の削減などを目的に行列の近似を行ったり、誤差のある測定値を成分とする行列を扱う際には、行列に関しても誤差を考える必要が生じる。
つまり、「もとの行列にどれくらい近いか？」などという、距離を考える必要がある。

距離はいわば長さを測る物差しであり、一定の基準（公理）を満たせば、さまざまな距離を定義することができる。そして、使い道に応じて最適な距離を選ぶことができる。

行列の距離を考える前に、一般的な距離について考えてみよう。

距離の公理

まず、距離を名乗るものはどれも次の性質を満たすように作る必要がある。

i. 2 点間の距離は負の数にはならない
ii. 2 点が同じ点 $\iff$ 2 点間の距離は 0
iii. 2 点間の距離は行きと帰りで変わらない
iv. 寄り道した方が必ず総距離は長い

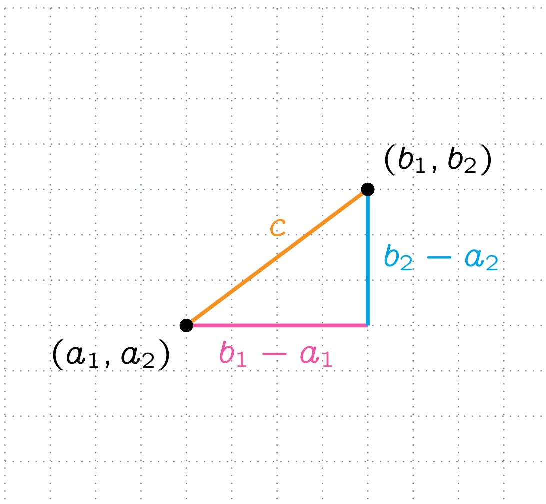
直観的に書いた上の 4 つの性質を数学的にまとめると、次のような公理になる。

距離の公理　集合 V 上の次の性質を満たす関数 $d\colon V \times V \to \mathbb{R}$ を距離と定める。ここで、$a, b, c \in V$ とする。

i. $d(a, b) \geq 0$（非負性）
ii. $d(a, b) = 0 \iff a = b$（同一性）
iii. $d(a, b) = d(b, a)$（対称性）
iv. $d(a, c) \leq d(a, b) + d(b, c)$（三角不等式）

例：ユークリッド距離とマンハッタン距離

状況に応じて最適な距離が異なることを、簡単な例で考えてみよう。



次の 2 パターンの条件下において、2 点 $A = (a_1, a_2)$ と $B = (b_1, b_2)$ の距離をどう定めるべきかを考える。

1. 斜めに移動できる
2. ジグザグにしか移動できない

斜めに移動できる場合は、単純に直線距離として考えることができるので、三平方の定理を用いて、

$$d(A, B) = c = \sqrt{(b_1 - a_1)^2 + (b_2 - a_2)^2}$$

のように計算すればよい。

このような距離はユークリッド距離と呼ばれる。

一方、ジグザグにしか移動できない場合は、実際に A から B へ移動するときの遠さを直線距離では表現できないので、次のように考えた方がよい。

$$d(A, B) = |b_1 - a_1| + |b_2 - a_2|$$

このような距離はマンハッタン距離と呼ばれる。

ノルムと距離の関係

ノルムは「長さ」を拡張した概念であり、ベクトルに対して定義された。

今、2 つのベクトル $\boldsymbol{a}, \boldsymbol{b}$ を

$$\boldsymbol{a} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}, \quad \boldsymbol{b} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$$

とおくと、先ほど示したユークリッド距離の式

$$d(A, B) = \sqrt{(b_1 - a_1)^2 + (b_2 - a_2)^2}$$

は内積を経由して次のように書き換えられる。

$$d(A, B) = \sqrt{(\boldsymbol{b} - \boldsymbol{a}, \boldsymbol{b} - \boldsymbol{a})} = \|\boldsymbol{b} - \boldsymbol{a}\|$$

このように、距離からノルムを定めることができ、逆にノルムから距離を定めることもできる。

上述の式は 2 次元平面上の点の間の距離を表しているが、ベクトルのノルムとしてユークリッド距離を考えれば、$\boldsymbol{n}$ 次元空間への拡張も容易になる。

$$\boldsymbol{a} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}, \quad \boldsymbol{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$

とおいたとしても、

$$d(A, B) = \sqrt{\sum_{i=1}^{n} (b_i - a_i)^2} = \sqrt{(\boldsymbol{b} - \boldsymbol{a}, \boldsymbol{b} - \boldsymbol{a})} = \|\boldsymbol{b} - \boldsymbol{a}\|$$

が成り立つ。

L_p ノルム

ユークリッド距離とマンハッタン距離は、次の L_p ノルムによって一般化される。

L_p ノルム　ベクトル $\boldsymbol{a} = (a_i)_{i=1}^n$ に対し、

$$\|\boldsymbol{a}\|_p = \left(\sum_{i=1}^n |a_i|^p \right)^{\frac{1}{p}}$$

はノルムの公理を満たし、これを L_p ノルムという。

L_p ノルムが距離として意味を持つのは、$p \geq 1$ のときである。

$p = 1$ の場合、

$$\|\boldsymbol{a}\|_1 = \sum_{i=1}^n |a_i|$$

となり、$\boldsymbol{a}$ を $\boldsymbol{b} - \boldsymbol{a}$ に置き換えれば、マンハッタン距離の式となる。

$p = 2$ の場合、

$$\|\boldsymbol{a}\|_2 = \left(\sum_{i=1}^n |a_i|^2 \right)^{\frac{1}{2}} = \sqrt{\sum_{i=1}^n a_i^2}$$

となり、$\boldsymbol{a}$ を $\boldsymbol{b} - \boldsymbol{a}$ に置き換えれば、ユークリッド距離の式となる。

行列の距離：フロベニウスノルム

[Todo 59:]

行列の距離：作用素ノルム

[Todo 60:]

特異値分解とベクトルのノルム

特異値分解はユニタリ変換であり、ユニタリ変換はベクトルのノルムを変えないことから、さまざまな有用な性質が導かれる。

⚓ **ユニタリ行列による座標変換**　$\mathcal{U} = \{\boldsymbol{u}_i\}_{i=1}^n$ を $\mathbb{C}^n$ の正規直交基底とし、ユニタリ行列 $U = (\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n)$ を定める。このとき、$\boldsymbol{v} \in \mathbb{C}^n$ の $\mathcal{U}$ に関する座標ベクトルは $U^*\boldsymbol{v}$ で与えられる。

証明

次式が成り立つことから、U は $\mathbb{C}^n$ の標準基底 $\{\boldsymbol{e}_i\}_{i=1}^n$ から $\mathcal{U}$ への変換行列とみなせる。

$$\begin{pmatrix} \boldsymbol{u}_1 & \cdots & \boldsymbol{u}_n \end{pmatrix} = \begin{pmatrix} \boldsymbol{e}_1 & \cdots & \boldsymbol{e}_n \end{pmatrix} U$$

座標ベクトルの変換則より、$\boldsymbol{v}$ の $\mathcal{U}$ に関する座標ベクトルを $\boldsymbol{c} \in \mathbb{C}^n$ とすると、

$$\boldsymbol{v} = U\boldsymbol{c}$$

両辺に左から U^* をかけると、

$$U^*\boldsymbol{v} = U^*U\boldsymbol{c} = \boldsymbol{c}$$

となり、たしかに $\boldsymbol{c}$ は $U^*\boldsymbol{v}$ で与えられることがわかる。 ■

A が特異値分解されていると、ベクトルのノルムの変化がわかりやすい。

⚓ **特異値分解に基づくノルムの展開表示**　行列 A がユニタリ行列 U, V を用いて $A = U\Sigma V^*$ と特異値分解されているとする。

V に対応する $\mathbb{C}^n$ の基底を $\mathcal{V}$ とし、$\boldsymbol{v} \in \mathbb{C}^n$ の $\mathcal{V}$ に関する座標ベクトルを $\boldsymbol{c} = (c_i) \in \mathbb{C}^n$ とすると、

$$\|A\boldsymbol{v}\|^2 = \sum_{i=1}^{r} \sigma_i^2 |c_i|^2$$

証明

ユニタリ行列による座標変換より、座標ベクトルは $\boldsymbol{c} = V^*\boldsymbol{v}$ で与えられる。

よって、

$$\|A\boldsymbol{v}\|^2 = \|U\Sigma V^*\boldsymbol{v}\|^2 = \|U\Sigma\boldsymbol{c}\|^2$$

ここで、左からユニタリ行列 U をかけてもノルムは変わらないので、

$$\|A\boldsymbol{v}\|^2 = \|\Sigma\boldsymbol{c}\|^2$$

Σ は対角成分に特異値 $\sigma_1, \ldots, \sigma_r$（$\sigma_i > 0$）を持ち、それ以外は 0 の $m \times n$ 行列であるから、

$$\Sigma\boldsymbol{c} = (\sigma_1 c_1, \ldots, \sigma_r c_r, 0, \ldots, 0)^T$$

よってそのノルムの二乗は、$\mathbb{C}^n$ 上の自分自身との内積と考えて、

$$\|\Sigma\boldsymbol{c}\|^2 = \sum_{i=1}^{r} |\sigma_i c_i|^2 = \sum_{i=1}^{r} \sigma_i^2 |c_i|^2$$

となり、目的の式が示された。 ■

最大特異値と作用素ノルム

A の最大特異値 σ_1 は、次の解釈をもつ。

⚓ **最大特異値と作用素ノルムの一致**　複素行列 A に対して、

$$\max_{\|\boldsymbol{v}\|=1} \|A\boldsymbol{v}\| = \sigma_1$$

証明

特異値分解に基づくノルムの展開表示と、$\sigma_i \leq \sigma_1$ を用いて、

$$\|A\boldsymbol{v}\|^2 = \sum_{i=1}^{r} \sigma_i^2 |c_i|^2 \leq \sum_{i=1}^{r} \sigma_1^2 |c_i|^2 = \sigma_1^2 \sum_{i=1}^{r} |c_i|^2$$

ここで、$\|\boldsymbol{v}\| = 1$ ならば、$\boldsymbol{v} \in \mathbb{C}^n$ の $\boldsymbol{\mathcal{V}}$ に関する座標ベクトル $\boldsymbol{c}$ も $\|\boldsymbol{c}\| = 1$ となるので、

$$\|A\boldsymbol{v}\|^2 \leq \sigma_1^2 \sum_{i=1}^{r} |c_i|^2 = \sigma_1^2 \|\boldsymbol{c}\|^2 = \sigma_1^2$$

よって、$\|A\boldsymbol{v}\| \leq \sigma_1$ が任意の単位ベクトル $\boldsymbol{v}$ に対して成り立つ。

すなわち、

$$\max_{\|\boldsymbol{v}\|=1} \|A\boldsymbol{v}\| \leq \sigma_1$$

等号は、$c_1 = 1$ で他の成分 $c_2, \ldots, c_r$ が 0 のときに成り立つ。 ■

作用素ノルム　複素行列 A に対して、A の作用素ノルムを次のように定義する。

$$\|A\| := \max_{\|\boldsymbol{v}\|=1} \|A\boldsymbol{v}\|$$

第 33 章

低ランク近似

ランク 1 行列による圧縮

1 つのグレースケール画像を行列 $A \in \mathbb{R}^{m \times n}$ に見立てると、画像のファイルサイズは mn に比例する。単純に行列の成分を保存しようとすると、膨大な記憶容量が必要となる。

そこで、圧縮に向けた一つの考え方として、$m \times n$ 型行列 A を、

> いくつかの「縦ベクトルと横ベクトルの積」の和として近似的に表現する

ことを考える。

ランク 1 行列

ここで重要なのは次の事実である。

> 「縦ベクトルと横ベクトルの積」は階数が 1 の行列

たとえば、

$$\boldsymbol{u} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \quad \boldsymbol{v}^\top = \begin{pmatrix} 4 & 5 \end{pmatrix}$$

とすると、これらの積は次のように計算される。

$$\boldsymbol{u}\boldsymbol{v}^\top = \begin{pmatrix} 1 \cdot 4 & 1 \cdot 5 \\ 2 \cdot 4 & 2 \cdot 5 \\ 3 \cdot 4 & 3 \cdot 5 \end{pmatrix} = \begin{pmatrix} 4 & 5 \\ 8 & 10 \\ 12 & 15 \end{pmatrix}$$

各行は、最初の行 $(4, 5)$ のスカラー倍となっていることに注目しよう。

つまり、独立な行は $(4, 5)$ だけであり、他の行はこの行の線形結合で表現できる。

独立な行が 1 つしかないので、この行列の階数は 1 である。

このことは、一般的な成分表示で考えることもできる。

$$\boldsymbol{u} = \begin{pmatrix} u_1 \\ \vdots \\ u_m \end{pmatrix}, \quad \boldsymbol{v}^\top = \begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix}$$

とおくと、これらの積は次のように計算される。

$$\boldsymbol{u}\boldsymbol{v}^\top = \begin{pmatrix} u_1 v_1 & u_1 v_2 & \cdots & u_1 v_n \\ u_2 v_1 & u_2 v_2 & \cdots & u_2 v_n \\ \vdots & \vdots & \ddots & \vdots \\ u_m v_1 & u_m v_2 & \cdots & u_m v_n \end{pmatrix}$$

ここで、i 行目を取り出すと、

$$u_i \cdot \begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix} = u_i \boldsymbol{v}^\top$$

となっているので、すべての行は $\boldsymbol{v}^\top$ のスカラー倍で表現できることがわかる。

低ランク近似による圧縮

縦ベクトルと横ベクトルの積は階数 1 の行列となることから、もし k 個の和で近似するならば、近似した行列の階数は k 以下となる。

もし、k 個の和で良い近似になっているのであれば、

- m 次元ベクトル（縦ベクトル）を k 本
- n 次元ベクトル（横ベクトル）を k 本

保持すればよいことになる。

これで、mn 個の成分を保持する必要はなくなり、$k(m+n)$ 個の成分を保持すれば十分となる。

たとえば、$m = 1000, n = 1000$ の行列を $k = 50$ 個の和でうまく近似できるとしたら、

$$\frac{k(m+n)}{mn} = \frac{50(1000+1000)}{1000 \times 1000} = 0.1$$

より、ファイルサイズを 90% 削減できることがわかる。

このようなアイデアを、工学では圧縮といい、数学的には行列の低ランク近似という。

特異値分解による低ランク近似

さて、O でない任意の行列 A は、次のように特異値分解できた。

$$A = \sigma_1 \boldsymbol{u}_1 \boldsymbol{v}_1^\top + \cdots + \sigma_k \boldsymbol{u}_k \boldsymbol{v}_k^\top + \cdots + \sigma_r \boldsymbol{u}_r \boldsymbol{v}_r^\top$$

これは、「縦ベクトルと横ベクトルの積」の和の形になっている。

さらに、特異値 σ_i は大きい順に並んでいるので、ランク k の行列の中で、

$$A_k = \sigma_1 \boldsymbol{u}_1 \boldsymbol{v}_1^\top + \cdots + \sigma_k \boldsymbol{u}_k \boldsymbol{v}_k^\top$$

が A に最も近い近似となるのではないか？と予想できる。

ここで、「A との近さ」を測るためには、行列に関するノルムを考える必要がある。

第 34 章

一般化逆行列

ムーア・ペンローズの擬似逆行列

正方行列は、それが正則であれば逆行列を持つ。

これを O でない任意の長方行列に拡張するのが一般逆行列（擬似逆行列）である。

擬似逆行列の存在と一意性　O でない任意の $m \times n$ 型行列 A に対して、以下の 4 つの条件を満たす $n \times m$ 型行列 A^+ がただ一つ存在する。

i. $AA^+A = A$
ii. $A^+AA^+ = A^+$
iii. $(AA^+)^\top = AA^+$
iv. $(A^+A)^\top = A^+A$

この A^+ をムーア・ペンローズの擬似逆行列という。

存在性の証明

A の特異値分解を考える。

$$\Sigma_r = \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{pmatrix}, \quad \Sigma_r^{-1} = \begin{pmatrix} \dfrac{1}{\sigma_1} & & \\ & \ddots & \\ & & \dfrac{1}{\sigma_r} \end{pmatrix}$$

$$\Sigma = \begin{pmatrix} \Sigma_r & O \\ O & O \end{pmatrix}, \quad \Sigma^{-1} = \begin{pmatrix} \Sigma_r^{-1} & O \\ O & O \end{pmatrix}$$

とおくと、A は次のように特異値分解できる。

$$A = U\Sigma V^\top$$

ここで、

$$B = V\Sigma^{-1}U^\top$$

とおくと、B は次のように 4 つの条件を満たす。

(i) $AA^+A = A$

U, V はユニタリ行列であるから、$V^\top V = E$ および $U^\top U = E$ が成り立つ。

$$\begin{aligned} ABA &= (U\Sigma V^\top)(V\Sigma^{-1}U^\top)(U\Sigma V^\top) \\ &= U\Sigma(V^\top V)\Sigma^{-1}(U^\top U)\Sigma V^\top \\ &= U(\Sigma\Sigma^{-1})\Sigma V^\top \end{aligned}$$

ここで、

$$\Sigma\Sigma^{-1} = \begin{pmatrix} E_r & O \\ O & O \end{pmatrix}$$

であり、これに Σ をかけると、単位行列 E_r の部分が Σ_r に置き換わるだけとなるので、

$$(\Sigma\Sigma^{-1})\Sigma = \Sigma$$

が成り立つ。

この関係を用いると、

$$
\begin{aligned}
ABA &= U(\Sigma\Sigma^{-1})\Sigma V^\top \\
&= U\Sigma V^\top \\
&= A
\end{aligned}
$$

となり、条件 (i) が成り立つ。 ■

(ii) $A^+AA^+ = A^+$

同様に、

$$
\begin{aligned}
BAB &= (V\Sigma^{-1}U^\top)(U\Sigma V^\top)(V\Sigma^{-1}U^\top) \\
&= V\Sigma^{-1}(U^\top U)\Sigma(V^\top V)\Sigma^{-1}U^\top \\
&= V(\Sigma^{-1}\Sigma)\Sigma^{-1}U^\top \\
&= V\Sigma^{-1}U^\top \\
&= B
\end{aligned}
$$

となり、条件 (ii) が成り立つ。 ■

(iii) $(AA^+)^\top = AA^+$

まず AB を計算すると、

$$
\begin{aligned}
AB &= (U\Sigma V^\top)(V\Sigma^{-1}U^\top) \\
&= U\Sigma(V^\top V)\Sigma^{-1}U^\top \\
&= U(\Sigma\Sigma^{-1})U^\top
\end{aligned}
$$

よって、$(AB)^\top$ は、

$$
\begin{aligned}
(AB)^\top &= (U(\Sigma\Sigma^{-1})U^\top)^\top \\
&= U(\Sigma\Sigma^{-1})U^\top \\
&= AB
\end{aligned}
$$

となり、条件 (iii) が成り立つ。 ■

(iv) $(A^+A)^\top = A^+A$

まず BA を計算すると、

$$\begin{aligned} BA &= (V\Sigma^{-1}U^\top)(U\Sigma V^\top) \\ &= V\Sigma^{-1}(U^\top U)\Sigma(V^\top V) \\ &= V(\Sigma^{-1}\Sigma)V^\top \end{aligned}$$

よって、$(BA)^\top$ は、

$$\begin{aligned} (BA)^\top &= (V(\Sigma^{-1}\Sigma)V^\top)^\top \\ &= V(\Sigma^{-1}\Sigma)V^\top \\ &= BA \end{aligned}$$

となり、条件 (iv) が成り立つ。 ■

ムーア・ペンローズの擬似逆行列の構成

存在性の証明過程から、ムーア・ペンローズの擬似逆行列は次のように構成すればよいことがわかる。

ムーア・ペンローズの擬似逆行列　行列 A のムーア・ペンローズの擬似逆行列 A^+ は、次のように定義される。

$$A^+ = V\Sigma^{-1}U^\top = V\begin{pmatrix} \dfrac{1}{\sigma_1} & & \\ & \ddots & \\ & & \dfrac{1}{\sigma_r} \end{pmatrix}U^\top$$

ムーア・ペンローズの擬似逆行列の一意性

行列 A に対して A^+ が一意的に定まることは、次のように示される。

一意性の証明

B_1, B_2 がいずれもムーア・ペンローズの擬似逆行列の 4 つの条件

i. $ABA = A$
ii. $BAB = B$

iii. $(AB)^\top = AB$

iv. $(BA)^\top = BA$

を満たすとすると、

$$\begin{aligned} B_1 &\overset{\text{ii}}{=} B_1AB_1 \\ &\overset{\text{iv}}{=} (B_1A)^\top B_1 && B_1A = B_1 \\ &\overset{\text{ii}}{=} (B_1AB_2A)^\top B_1 && AB_2A = A \\ &\overset{\text{iii}}{=} (B_2A)^\top (B_1A)^\top B_1 && \text{積の転置の性質} \\ &\overset{\text{iv}}{=} (B_2A)(B_1A)B_1 \\ &\overset{\text{i}}{=} B_2AB_1 && AB_1A = A \end{aligned}$$

同様の計算により、$B_2 = B_2AB_1$ も得られる。

よって、$B_1 = B_2$ である。 ■

特異値分解の展開式による表記

O でない $m \times n$ 型行列 A が次のように特異値分解されているとする。

$$A = \sigma_1 \boldsymbol{u}_1 \boldsymbol{v}_1^\top + \cdots + \sigma_r \boldsymbol{u}_r \boldsymbol{v}_r^\top$$

このとき、ムーア・ペンローズの擬似逆行列 A^+ は次のように表される。

$$A^+ = \frac{\boldsymbol{v}_1 \boldsymbol{u}_1^\top}{\sigma_1} + \cdots + \frac{\boldsymbol{v}_r \boldsymbol{u}_r^\top}{\sigma_r}$$

σ_i が逆数に、$\boldsymbol{u}_i \boldsymbol{v}_i^\top$ が $\boldsymbol{v}_i \boldsymbol{u}_i^\top$ に置き換わっていることに注意しよう。

擬似逆行列と行空間・列空間への射影

逆行列は、もとの行列との積が単位行列となるものとして定義された。

$$A^{-1}A = AA^{-1} = E$$

一方、ムーア・ペンローズの擬似逆行列ともとの行列との積は、列および行の張る空間への射影行列となる。

$$A^+A = P_{\mathcal{V}}, \quad AA^+ = P_{\mathcal{U}}$$

ムーア・ペンローズ逆による行空間・列空間への射影　A^+ をムーア・ペンローズの擬似逆行列とするとき、A^+A は行空間 $\mathcal{V}$ への射影行列、AA^+ は列空間 $\mathcal{U}$ への射影行列となる。

証明

$A^+A = P_{\mathcal{V}}$

A と A^+ を正規直交化された特異ベクトルによる特異値分解で表すと、その積は、

$$
\begin{aligned}
A^+A &= \left(\sum_{i=1}^{r} \frac{\boldsymbol{v}_i \boldsymbol{u}_i^\top}{\sigma_i}\right)\left(\sum_{j=1}^{r} \sigma_j \boldsymbol{u}_j \boldsymbol{v}_j^\top\right) \\
&= \sum_{i=1}^{r}\sum_{j=1}^{r} \frac{\sigma_j}{\sigma_i} \boldsymbol{v}_i (\boldsymbol{u}_i^\top \boldsymbol{u}_j) \boldsymbol{v}_j^\top \\
&= \sum_{i=1}^{r}\sum_{j=1}^{r} \frac{\sigma_j}{\sigma_i} \delta_{ij} \boldsymbol{v}_i \boldsymbol{v}_j^\top \\
&= \sum_{i=1}^{r} \boldsymbol{v}_i \boldsymbol{v}_i^\top = P_{\mathcal{V}}
\end{aligned}
$$

となり、行空間 $\mathcal{V}$ への射影行列となる。 ■

$AA^+ = P_{\mathcal{U}}$

同様に、AA^+ を計算すると、

$$
\begin{aligned}
AA^+ &= \left(\sum_{i=1}^{r} \sigma_i \boldsymbol{u}_i \boldsymbol{v}_i^\top\right)\left(\sum_{j=1}^{r} \frac{\boldsymbol{v}_j \boldsymbol{u}_j^\top}{\sigma_j}\right) \\
&= \sum_{i=1}^{r}\sum_{j=1}^{r} \frac{\sigma_i}{\sigma_j} (\boldsymbol{u}_i^\top \boldsymbol{u}_j) \boldsymbol{v}_i \boldsymbol{v}_j^\top \\
&= \sum_{i=1}^{r}\sum_{j=1}^{r} \frac{\sigma_i}{\sigma_j} \delta_{ij} \boldsymbol{v}_i \boldsymbol{v}_j^\top \\
&= \sum_{i=1}^{r} \boldsymbol{v}_i \boldsymbol{v}_i^\top = P_{\mathcal{U}}
\end{aligned}
$$

となり、列空間 $\mathcal{U}$ への射影行列となる。 ■

A が正則行列の場合

A が正則行列の場合、A のムーア・ペンローズの擬似逆行列は、A の逆行列に一致する。

すなわち、A が正則の場合は、次が成り立つ。

$$A^{+}A = AA^{+} = E$$

この意味で、ムーア・ペンローズの擬似逆行列は逆行列の一般化とみなせる。

また、単位行列は全空間への射影行列であるので、ムーア・ペンローズの擬似逆行列ともとの行列との積が射影行列になることの特別な場合といえる。

正則行列に対するムーア・ペンローズ逆　A が正則行列であれば、ムーア・ペンローズの擬似逆行列 A^{+} は逆行列 A^{-1} に一致する。

証明

正則行列 A の逆行列を A^{-1} とすると、

$$A^{+}A = A^{-1}A = E$$
$$AA^{+} = AA^{-1} = E$$

また、$AA^{-1} = E$ と $A^{-1}A = E$ より、

$$(AA^{-1})^{\top} = E^{\top} = AA^{-1}$$
$$(A^{-1}A)^{\top} = E^{\top} = A^{-1}A$$

以上より、A^{-1} はムーア・ペンローズの擬似逆行列 A^{+} の定義を満たす。 ■

逆変換としての擬似逆行列

ムーア・ペンローズの擬似逆行列は、線形変換の逆変換という視点でも、逆行列の一般化となっている。

逆変換と恒等変換

逆行列 A^{-1} は、線形変換 A の逆変換を表すものだった。
逆変換は「元に戻す」操作である。

変換 A と逆変換 A^{-1} を合成すると、「なにもしない」という変換（恒等変換）が得られる。
A してからそれを A^{-1} で打ち消すのだから、結局なにもしなかったことになるのである。

このことを数式で表したものが、逆行列 A^{-1} の定義式である。

$$A^{-1}A = AA^{-1} = E$$

単位行列 E は、恒等変換を表す行列である。

列空間上の逆変換

$\boldsymbol{x}$ が列空間 $\mathcal{U}$ の元である場合は、列空間へ射影しても変わらないので、

$$P_{\mathcal{U}}\boldsymbol{x} = \boldsymbol{x}$$

が成り立つ。
つまり、列空間 $\mathcal{U}$ においては、$P_{\mathcal{U}}$ は「なにもしない」恒等変換を表す。

このことから、

$$AA^{+} = P_{\mathcal{U}}$$

という式は、

列空間 $\mathcal{U}$ において、A^{+} は A の逆変換を表す

と解釈できる。

行空間上の逆変換

同様に、$\boldsymbol{x}$ が行空間 $\mathcal{V}$ の元である場合は、行空間へ射影しても変わらないので、

$$P_{\mathcal{V}}\boldsymbol{x} = \boldsymbol{x}$$

が成り立つ。

つまり、行空間 $\mathcal{V}$ においては、$P_{\mathcal{V}}$ は「なにもしない」恒等変換を表す。

このことから、

$$A^{+}A = P_{\mathcal{V}}$$

という式は、

行空間 $\mathcal{V}$ において、A^{+} は A の逆変換を表す

と解釈できる。

第 35 章

最小二乗解と最小ノルム解

解けない線形方程式

$\boldsymbol{x}$ についての線形方程式を考える。

$$A\boldsymbol{x} = \boldsymbol{b}$$

ここで、A は $m \times n$ 型行列、$\boldsymbol{x}$ は n 次元ベクトル、$\boldsymbol{b}$ は m 次元ベクトルである。

$m = n$ で、A が正則であれば、その逆行列を用いて、一意な解を求めることができる。

$$\boldsymbol{x} = A^{-1}\boldsymbol{b}$$

しかし、この線型方程式が解けない場合を扱うこともある。

i. $m > n$（A が縦長）の場合、条件式の数が未知数の数より多く、方程式は不能（解なし）
ii. $m < n$（A が横長）の場合、条件式の数が未知数の数より少なく、方程式は不定（解が無数にある）

このような場合には、次のようなアプローチが考えられる。

i. 解が存在しないなら、なるべく近い解を探す
ii. 解が無数にあるなら、その中から適した解を探す

なにをもって「近い」「適した」とすべきかは状況によって異なるが、近い解は最小二乗解、適した解は最小ノルム解とすることが多い。

最小二乗解

$A\boldsymbol{x} = \boldsymbol{b}$ の解がないなら、「せめて $A\boldsymbol{x}$ が $\boldsymbol{b}$ にできるだけ近くなるような $\boldsymbol{x}$ を求めよう」というのが最小二乗解のアプローチである。

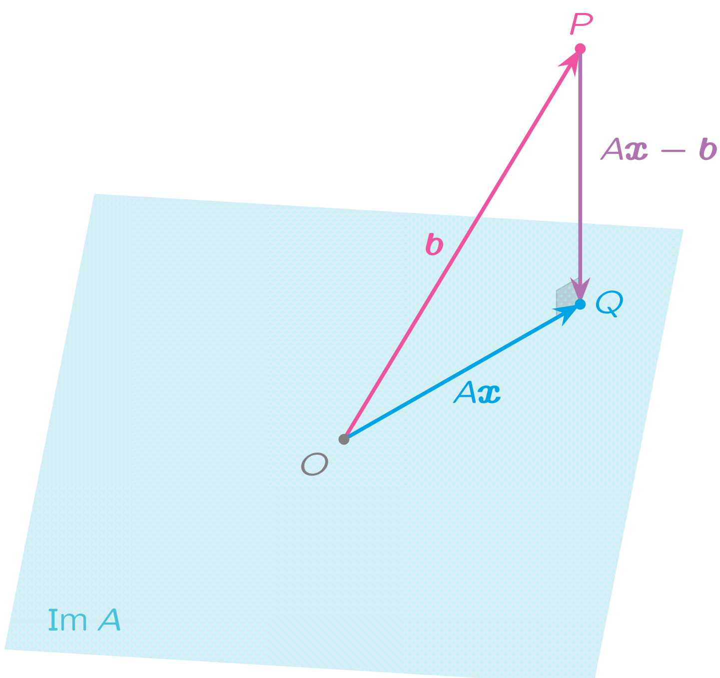
最小二乗解とは、次のような残差 J を最小化するような $\boldsymbol{x}$ のことである。

$$J = \|A\boldsymbol{x} - \boldsymbol{b}\|^2$$

残差の最小化と直交射影

$\boldsymbol{x}$ を動かすと、$A\boldsymbol{x}$ は $\mathrm{Im}\,A$ 上のさまざまなベクトルをとる。

$m > n$ の場合、$\boldsymbol{b}$ は $\boldsymbol{x}$ よりも高次元のベクトルなので、$\mathrm{Im}\,A$ からはみ出してしまう。



ここで、点 P から最短となる $\mathrm{Im}\,A$ 上の点は、P を $\mathrm{Im}\,A$ へ直交射影した点 Q である。

このことから、$\|A\boldsymbol{x} - \boldsymbol{b}\|$ が最小となるのは、

$\boldsymbol{b}$ を列空間 $\mathrm{Im}\,A$ へ直交射影したベクトルが $A\boldsymbol{x}$

になっているときだとわかる。

$\operatorname{Im} A$ は A の列空間（A の列ベクトルが張る空間）であるから、これまで通り $\mathcal{U}$ と表すことにしよう。
すると、最小二乗解 $\boldsymbol{x}$ が満たすべき条件は、次のように書ける。

$$A\boldsymbol{x} = P_{\mathcal{U}}\boldsymbol{b}$$

最小二乗解と列空間への直交射影の関係　A の列空間を $\mathcal{U}$ とし、$\mathcal{U}$ への直交射影行列を $P_{\mathcal{U}}$ とすると、線形方程式 $A\boldsymbol{x} = \boldsymbol{b}$ の最小二乗解は次の関係を満たす。

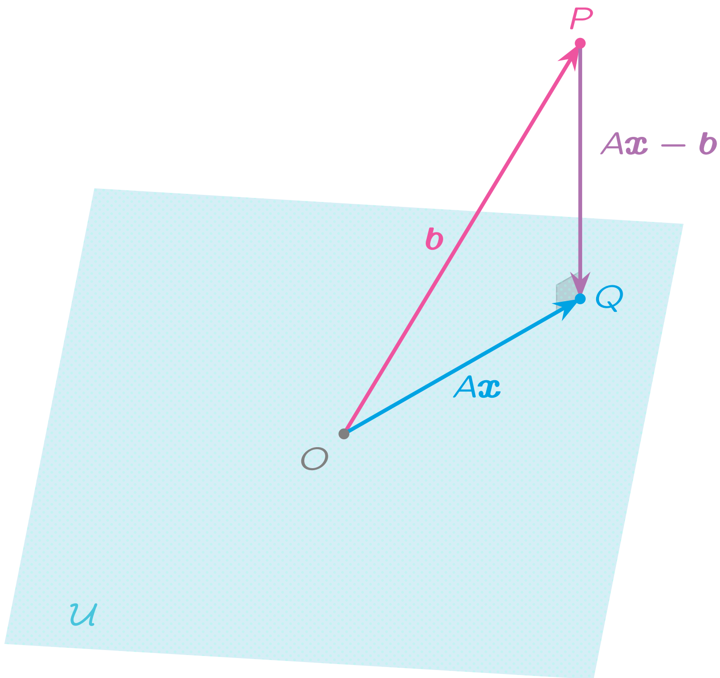
$$A\boldsymbol{x} = P_{\mathcal{U}}\boldsymbol{b}$$

列フルランクの場合の最小二乗解

$n = \operatorname{rank} A$ だと仮定する。階数は線型独立な列の最大個数なので、A の n 個の列ベクトルは線型独立である。

A の列ベクトルを $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ とすると、これらは $\mathcal{U}$ を張るベクトルであり、かつ線型独立であるので、$\mathcal{U}$ の基底を成す。
よって、$\mathcal{U}$ 上の任意のベクトルは、これらの線型結合で表される。

直交射影の定義より、$A\boldsymbol{x} - \boldsymbol{b}$ は $\mathcal{U}$ 上のすべてのベクトルに直交する。



$A\boldsymbol{x} - \boldsymbol{b}$ が $\mathcal{U}$ 上の任意のベクトルに直交するには、

$$\boldsymbol{a}_i^\top (A\boldsymbol{x} - \boldsymbol{b}) = 0 \quad (i = 1, \ldots, n)$$

が成り立てばよい ($\mathcal{U}$ 上の任意のベクトルは $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ の線型結合で表されるのだから)。

この条件を行列の形に書き直すと、

$$A^\top (A\boldsymbol{x} - \boldsymbol{b}) = \boldsymbol{o}$$

となる。

補足

$A^\top (A\boldsymbol{x} - \boldsymbol{b}) = \boldsymbol{o}$ は、$\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ すべてとの内積を並べた形である。

$$\begin{pmatrix} \boldsymbol{a}_1^\top \\ \vdots \\ \boldsymbol{a}_n^\top \end{pmatrix} (A\boldsymbol{x} - \boldsymbol{b}) = \begin{pmatrix} \boldsymbol{a}_1^\top (A\boldsymbol{x} - \boldsymbol{b}) \\ \vdots \\ \boldsymbol{a}_n^\top (A\boldsymbol{x} - \boldsymbol{b}) \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

これを展開すると、

$$A^\top A\boldsymbol{x} - A^\top \boldsymbol{b} = \boldsymbol{o}$$
$$A^\top A\boldsymbol{x} = A^\top \boldsymbol{b}$$

となるので、$A^\top A$ が正則であれば、$\boldsymbol{x} = \cdots$ の形に整理できる。

$n = \operatorname{rank} A$ という仮定のもとでは、次の定理が成り立つ。

線型独立な列ベクトルと自己共役積の正則性　$m \times n$ 型行列 A において、$m \geq n = \operatorname{rank} A$（列ベクトルが線型独立である）とき、$A^\top A$ は正則である。

証明

$A^\top A$ が正則であることと同値な条件として、$A^\top A$ の核空間の次元が 0 であることを示す。すなわち、次を示せばよい。

$$A^\top A\boldsymbol{x} = \boldsymbol{o} \Longrightarrow \boldsymbol{x} = \boldsymbol{o}$$

$A^\top A\boldsymbol{x} = \boldsymbol{o}$ の両辺に $\boldsymbol{x}^\top$ をかけると、

$$\begin{aligned} \boldsymbol{x}^\top A^\top A\boldsymbol{x} &= \boldsymbol{o} \\ (A\boldsymbol{x})^\top (A\boldsymbol{x}) &= 0 \\ \|A\boldsymbol{x}\|^2 &= 0 \end{aligned}$$

ノルムが 0 であるベクトルは零ベクトルのみであるから、

$$A\boldsymbol{x} = \boldsymbol{o}$$

ここで、$A\boldsymbol{x} = \boldsymbol{o}$ ということは、A の列ベクトルの線形結合として $\boldsymbol{o}$ になる $\boldsymbol{x}$ が存在することを意味する。

しかし、仮定より A の列ベクトルは線型独立なので、$\boldsymbol{x} = \boldsymbol{o}$ しか解がない。

よって、$A^\top A\boldsymbol{x} = \boldsymbol{o}$ の解空間である核空間 $\operatorname{Ker}(A^\top A)$ は、零ベクトルしか含まない。

すなわち、$\dim \operatorname{Ker}(A^\top A) = 0$ であり、$A^\top A$ は正則である。 ■

$A^\top A$ が正則であることがわかったので、$\boldsymbol{x}$ は次のように求められる。

$$\boldsymbol{x} = (A^\top A)^{-1} A^\top \boldsymbol{b}$$

これが、$m \geq n = \operatorname{rank} A$ という条件のもとでの最小二乗解である。

⚓ **線型独立な列ベクトルに対する最小二乗解**　$m \times n$ 型行列 A において、$m \geq n = \operatorname{rank} A$（列ベクトルが線型独立である）とき、線形方程式 $A\boldsymbol{x} = \boldsymbol{b}$ の最小二乗解 $\boldsymbol{x}$ は次のように表される。

$$\boldsymbol{x} = (A^\top A)^{-1} A^\top \boldsymbol{b}$$

列空間への直交射影行列

$\boldsymbol{x} = (A^\top A)^{-1} A^\top \boldsymbol{b}$ の両辺に左から A をかけると、

$$A\boldsymbol{x} = A(A^\top A)^{-1} A^\top \boldsymbol{b}$$

この式を、最小二乗解と射影行列の関係 $A\boldsymbol{x} = P_{\mathcal{U}}\boldsymbol{b}$ と見比べると、

$$P_{\mathcal{U}} = A(A^\top A)^{-1} A^\top$$

であることがわかる。

⚓ **列空間への直交射影の行列表現**　$m \times n$ 型行列 A において、$m \geq n = \operatorname{rank} A$（列ベクトルが線型独立）のとき、$A$ の列空間への直交射影行列は次のように表される。

$$P_{\mathcal{U}} = A(A^\top A)^{-1} A^\top$$

最小ノルム解

$A\boldsymbol{x} = \boldsymbol{b}$ の解が無数にある場合は、「その中で有用な解を選ぼう」というアプローチをとる。このような場合、$\|\boldsymbol{x}\|$ が最小となる解を選ぶことが多い。

$\|\boldsymbol{x}\|$ が最小となる解を選ぶメリットはさまざまある。

- 解空間の中で「原点に最も近い」解
- 「変位が最小」など、物理学でも頻出する条件を満たす解

しかし、$\|\boldsymbol{x}\|$ は平方根を含むため、計算が複雑になりがちである。

そのため、2 乗して平方根を外したもの、すなわち $\|\boldsymbol{x}\|^2$ が最小となる解を考えることが多い。

$\|\boldsymbol{x}\|^2$ が最小となる解は、最小ノルム解と呼ばれる。

核空間と直交する空間への射影

解のパラメータ表示の再解釈でも考察したように、$A\boldsymbol{x} = \boldsymbol{b}$ の解は次のようにして構成することができる。

1. 1 つの解 $\boldsymbol{x}_0$ を見つける
2. $\mathrm{Ker}\, A$ の任意のベクトル $\boldsymbol{u}$ を持ってくれば、$\boldsymbol{x}_0 + \boldsymbol{u}$ も解である

補足：なんで $\boldsymbol{x}_0 + \boldsymbol{u}$ も解だといえるのか？

A によって $\boldsymbol{o}$ に写ってしまうような、つまり $A\boldsymbol{u} = \boldsymbol{o}$ となるような $\boldsymbol{u}$ すべての集合が $\mathrm{Ker}\, A$ である。

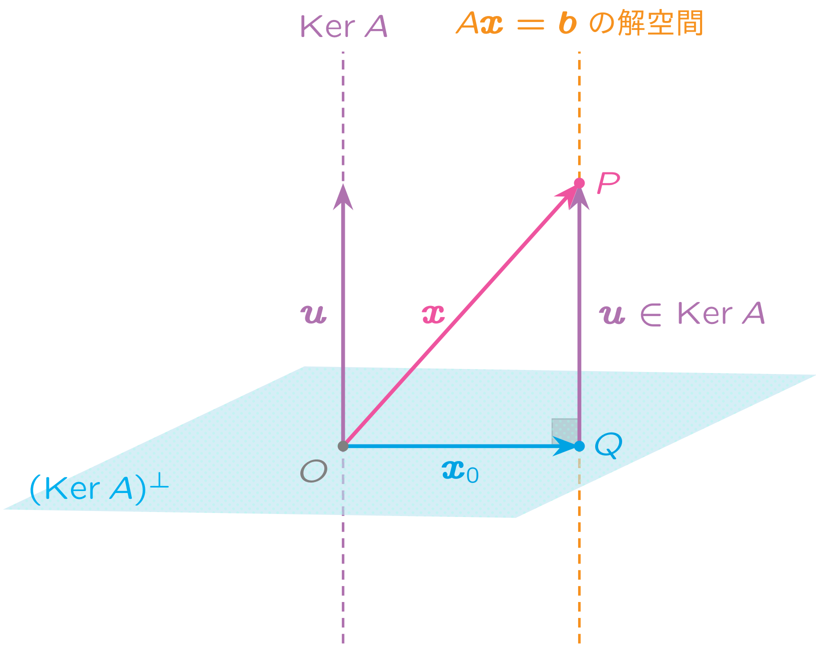
$A\boldsymbol{x}_0 = \boldsymbol{b}$ が成り立つことと合わせると、

$$A(\boldsymbol{x}_0 + \boldsymbol{u}) = A\boldsymbol{x}_0 + A\boldsymbol{u} = A\boldsymbol{x}_0 + \boldsymbol{o} = A\boldsymbol{x}_0 = \boldsymbol{b}$$

となり、たしかに $\boldsymbol{x} = \boldsymbol{x}_0 + \boldsymbol{u}$ も $A\boldsymbol{x} = \boldsymbol{b}$ を満たすことがわかる。

そのため、$\mathrm{Ker}\, A$ を $\boldsymbol{x}_0$ だけ平行移動したものが、$A\boldsymbol{x} = \boldsymbol{b}$ の解空間となる。

$$\text{解空間} = \{\boldsymbol{x}_0 + \boldsymbol{u} \mid \boldsymbol{u} \in \mathrm{Ker}\, A\}$$



さて、$A\boldsymbol{x} = \boldsymbol{b}$ の解

$$\boldsymbol{x} = \boldsymbol{x}_0 + \boldsymbol{u} \quad (\boldsymbol{u} \in \mathrm{Ker}\,A)$$

において、$\boldsymbol{x}$ のノルム（長さ）が最小となるのは、明らかに $\boldsymbol{u} = \boldsymbol{o}$ のときである。

つまり、$\boldsymbol{x} = \boldsymbol{x}_0$ が最小ノルム解となるが、$\boldsymbol{x}_0$ を使わずに最小ノルム解を表現したい。

図を見ると、$\boldsymbol{x}_0$ は、$\boldsymbol{x}$ を平面 $(\mathrm{Ker}\,A)^{\perp}$ へ直交射影したものであることがわかるので、

解 $\boldsymbol{x}$ のノルムを最小化するには、

$\boldsymbol{x}$ を Ker A と直交する空間に直交射影すればよい

ということがいえる。

ここで、核空間と行空間の直交関係より、「Ker A と直交する空間」は、A の行空間 $\mathrm{Im}\,A^{\top}$ である。

まとめると、$\|\boldsymbol{x}\|$ が最小となるのは、

$\boldsymbol{x}_0$ を行空間 $\mathrm{Im}\,A^{\top}$ へ直交射影したベクトルが $\boldsymbol{x}$

となっているときだとわかる。

行空間 $\operatorname{Im} A^\top$ をこれまで通り $\mathcal{V}$ と表すことにすると、最小ノルム解は次のように表現できる。

$$\boldsymbol{x} = P_{\mathcal{V}}\boldsymbol{x}_0 \quad (A\boldsymbol{x}_0 = \boldsymbol{b})$$

最小ノルム解と行空間への直交射影の関係　A の行空間を $\mathcal{V}$ とし、$\mathcal{V}$ への直交射影行列を $P_{\mathcal{V}}$ とすると、線形方程式 $A\boldsymbol{x} = \boldsymbol{b}$ の最小ノルム解は次の関係を満たす。

$$\boldsymbol{x} = P_{\mathcal{V}}\boldsymbol{x}_0 \quad (A\boldsymbol{x}_0 = \boldsymbol{b})$$

行空間への直交射影行列

転置すると行と列が入れ替わるため、A の行空間は、$A^\top$ の列空間と言い換えられる。

そのため、A の列空間への直交射影行列の式

$$P_{\mathcal{U}} = A(A^\top A)^{-1}A^\top$$

において、A を $A^\top$ に置き換えたものが、A の行空間への直交射影行列となる。

$$P_{\mathcal{V}} = A^\top(AA^\top)^{-1}A$$

行空間への直交射影の行列表現　$m \times n$ 型行列 A において、$n \geq m = \operatorname{rank} A$（行ベクトルが線型独立）のとき、$A$ の行空間への直交射影行列は次のように表される。

$$P_{\mathcal{V}} = A^\top(AA^\top)^{-1}A$$

行フルランクの場合の最小ノルム解

$\operatorname{rank} A = m$ という条件のもとでは、$P_{\mathcal{V}}$ を

$$P_{\mathcal{V}} = A^\top(AA^\top)^{-1}A$$

という式で表現できるので、これを $\boldsymbol{x} = P_{\mathcal{V}}\boldsymbol{x}_0$ という関係に代入すると、

$$\begin{aligned} \boldsymbol{x} &= A^\top(AA^\top)^{-1}A\boldsymbol{x}_0 \\ &= A^\top(AA^\top)^{-1}\boldsymbol{b} \end{aligned} \qquad A\boldsymbol{x}_0 = \boldsymbol{b}$$

という形で、最小ノルム解が求まる。

⚓ **線型独立な行ベクトルに対する最小ノルム解**　$m \times n$ 型行列 A において、$n \geq m = \operatorname{rank} A$（行ベクトルが線型独立である）とき、線形方程式 $A\boldsymbol{x} = \boldsymbol{b}$ の最小ノルム解 $\boldsymbol{x}$ は次のように表される。

$$\boldsymbol{x} = A^\top (AA^\top)^{-1} \boldsymbol{b}$$

擬似逆行列による解

ムーア・ペンローズの擬似逆行列と直交射影行列は、次のような関係で結ばれていた。

$$A^+ A = P_{\mathcal{V}}, \quad AA^+ = P_{\mathcal{U}}$$

これらの関係を用いると、最小二乗解や最小ノルム解などのすべての場合を包括した解の表現が得られる。

最小二乗解

最小二乗解と列空間への直交射影の関係より、最小二乗解は次の条件を満たすものだった。

$$A\boldsymbol{x} = P_{\mathcal{U}} \boldsymbol{b}$$

擬似逆行列と直交射影行列の関係 $AA^+ = P_{\mathcal{U}}$ を用いると、

$$A\boldsymbol{x} = AA^+ \boldsymbol{b}$$

この式から、$\boldsymbol{x} = A^+ \boldsymbol{b}$ が、$A\boldsymbol{x}$ を $\boldsymbol{b}$ に最も近づける近似解であることがわかる。

> A が正則でない場合を考えているので、両辺に A^{-1} をかけて $\boldsymbol{x}$ を得るという同値変形は成り立たない。それゆえ、$\boldsymbol{x} = A^+ \boldsymbol{b}$ は「近似解」でしかない。

よって、最小二乗解は、次のように表すことができる。

$$\boldsymbol{x} = A^+ \boldsymbol{b}$$

最小ノルム解

最小ノルム解と行空間への直交射影の関係より、最小ノルム解は次の条件を満たすものだった。

$$\boldsymbol{x} = P_{\mathcal{V}}\boldsymbol{x}_0$$

擬似逆行列と直交射影行列の関係 $A^+A = P_{\mathcal{V}}$ を用いると、

$$\boldsymbol{x} = A^+A\boldsymbol{x}_0 = A^+\boldsymbol{b}$$

よって、最小ノルム解は、次のように表すことができる。

$$\boldsymbol{x} = A^+\boldsymbol{b}$$

擬似逆行列はあらゆる場合の解を表せる

このように、最小二乗解も最小ノルム解も、ムーア・ペンローズの擬似逆行列を用いると同じ式で表現できる。
そして、ここでの議論は、射影による条件のみを前提としており、A の行や列の線型独立性（フルランクかどうか）に依存していない。

つまり、フルランクでない場合も、擬似逆行列 A^+ が求められれば、

$$\boldsymbol{x} = A^+\boldsymbol{b}$$

という形で、最小二乗解や最小ノルム解を求めることができる。

また、A が正則な場合は $A^+ = A^{-1}$ となるので、$\boldsymbol{x} = A^{-1}\boldsymbol{b}$ という逆行列を用いた解の表現も包括していることがわかる。

⚓ **擬似逆行列による線形方程式の解**　A のムーア・ペンローズ擬似逆行列 A^+ を用いると、線形方程式 $A\boldsymbol{x} = \boldsymbol{b}$ の解は次のように表される。

$$\boldsymbol{x} = A^+\boldsymbol{b}$$

第 36 章

空間の当てはめ

第 37 章

行列の因子分解

第 38 章

テンソル積

双線形写像

双線形形式の一般化として、双線形写像（**bilinear map**）を考える。

双線形写像とは、2 つの線形空間の直積から線形空間への写像で、各成分に対して線形となるもののことである。

双線形写像　V, W, U を線型空間とする。写像 $\Phi\colon V \times W \to U$ が次の条件を満たすとき、Φ を双線形写像という。

i. $\Phi(\boldsymbol{v}_1 + \boldsymbol{v}_2, \boldsymbol{w}) = \Phi(\boldsymbol{v}_1, \boldsymbol{w}) + \Phi(\boldsymbol{v}_2, \boldsymbol{w})$
ii. $\Phi(\boldsymbol{v}, \boldsymbol{w}_1 + \boldsymbol{w}_2) = \Phi(\boldsymbol{v}, \boldsymbol{w}_1) + \Phi(\boldsymbol{v}, \boldsymbol{w}_2)$
iii. $\Phi(c\boldsymbol{v}, \boldsymbol{w}) = \Phi(\boldsymbol{v}, c\boldsymbol{w}) = c\Phi(\boldsymbol{v}, \boldsymbol{w})$

ここで、$\boldsymbol{v}, \boldsymbol{v}_1, \boldsymbol{v}_2 \in V$，$\boldsymbol{w}, \boldsymbol{w}_1, \boldsymbol{w}_2 \in W$，$c \in \mathbb{R}$ である。

双線形写像は双線形形式の一般化

$U = \mathbb{R}$ とすると、Φ は双線形形式となる。

すなわち、双線形写像は双線形形式の一般化である。

Zebra Notes

Type	Number
todo	60
note	4
placeholder	2