1 連立方程式と掃きだし法（春学期の続き）

春学期の続きとして，連立方程式を考える．皆さん，期末試験では連立方程式の解き方で苦戦していたので，もう一度，復習を兼ねてやります．ノートは春学期に配ったので，要点のみ再録します．この節では連立方程式

$$A\boldsymbol{x} = \boldsymbol{b} \tag{1.0.1}$$

を考える．ここで A は $m \times n$ 行列，$\boldsymbol{x}$ と $\boldsymbol{b}$ は

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ x_n \end{bmatrix}, \qquad \boldsymbol{b} = \begin{bmatrix} b_1 \\ b_2 \\ \cdot \\ \cdot \\ b_m \end{bmatrix} \tag{1.0.2}$$

というベクトルで，$x_1 \sim x_n$ が未知数である．

教科書と僕のノートでは m, n の役割が逆である．このノートを作ってから気づいたのだが，下手に直すと間違いそうなので，このままにしておく．

1.1 掃きだし法（復習）

この節の内容は全学期の補講でカバーしたが，期末のできが悪かったので，少しだけ復習する．
この節では

- 「掃きだし法」を使って連立一次方程式が解けるようになること．
- 一次方程式系の解の様子には３つの可能性があることを理解すること．
  - 解が全く存在しない（不能）
  - 解が存在し，一意に定まる
  - 解が無数にたくさん存在する（不定）

を理解することが目的である．

「掃き出し法」とは原理的には（中学以来の）変数を消去して連立方程式を解く方法だが，ある程度間単にできるように整理したもので，以下の３つの操作の繰り返しからなる：

(0) ２つの方程式の順序を入れ替える．
(a) １つの方程式に，別の方程式の定数倍を加える．
(b) １つの方程式にゼロでない数をかける．

この３つの操作のそれぞれについて，操作の前と後では，方程式の解の集合は変わらない（不変である）．つまり，これらは方程式系に対する同値変形になっているわけで，この３つの同値変形をくり返して，方程式をわかりやすい形に変形するのが目的である（具体例は黒板で）．

(行列との関係)
掃き出し法での変形をよく見ると，未知数をいちいち x, y, z と書かなくても，その係数のつくる「拡大係数行列」$(A, \boldsymbol{b})$ だけ取り出して，同様の計算をやれば良い．この行列に対する操作は，以下の３つである（これも黒板で）．

(0) ２つの行を入れ替える．
(a) １つの行に，別の行の定数倍を加える．
(b) １つの行にゼロでない数をかける．

全学期にはあまり強調しなかったが，上の３つの操作を**行列に対する行の基本変形**という．

では，これから一次方程式系には３つの場合があることを例を使って学習しよう．典型的な例では解が存在して一意に定まる．

しかし，そうでない例もある．以下が一例である：

$$(1)\quad \begin{cases} x - y + z = 4 \\ 2x - 2y + z = 6 \\ -x + y + 2z = 2 \end{cases} \qquad (2)\quad \begin{cases} x - y + z = 2 \\ 2x - 2y + z = 0 \\ -x + y + 2z = 4 \end{cases} \tag{1.1.1}$$

この場合（解き方は各自やってみること），掃きだし法で解いた結果は

$$(1)\quad \begin{cases} x - y = 2 \\ z = 2 \\ 0 = 0 \end{cases} \qquad (2)\quad \begin{cases} x - y = -2 \\ z = 4 \\ 0 = -6 \end{cases} \tag{1.1.2}$$

となる．(1) の方は，$z = 2$ かつ，$x = y+2$ なら何でも良い．つまり，t を任意の実数として，$x = t+2, y = t, z = 2$ が解なのである．この場合，解は無数にあるわけだ．

一方，(2) の場合は一番下の式が矛盾している．x, y, z をどのようにとっても，この３つを満たすことはできない．つまり，もともとの (2) の解は存在しないのだ．

以上を多少強引にまとめると，連立一次方程式系の解については，以下の３つの可能性があることがわかる：

(a) 解が存在し，一意的に定まる（上の例 0 のように）

(b) 解が無数に存在する（上の例 (1) のように）— 連立方程式系は「不定」であるという．

(c) 解が全く存在しない（上の例 (2) のように）— 連立方程式系は「不能」であるという．

与えられた方程式系がこの３つのどれであるかは，一般には解いてみないとわからないが[1]，以下でもう少し考える．未知数の数を n，方程式の数を m とすると，$m = n$ なら (a)，$m > n$ なら (c)，$m < n$ なら (b) と言いたくなるが，これは一般には正しくないから注意のこと．（各自，反例を考えてみよう．）

行列の階数の話に入る前に，今までの宿題の一つを片づけておこう．

($\mathbb{R}^m$ において，$m+1$ 本以上のベクトルが一次従属であることの初等的証明)

ベクトルが n 本あるとする（$n > m$）．これらが一次独立か従属かを判定するには，方程式

$$x_1\,\boldsymbol{a}_1 + x_2\,\boldsymbol{a}_2 + x_3\,\boldsymbol{a}_3 + \ldots + x_n\,\boldsymbol{a}_n = \mathbf{0} \tag{1.1.3}$$

を $x_1, x_2, \ldots, x_n$ について解き，解が「すべてゼロ」に限るかどうかを見れば良かった（定理??）．我々は一次従属だと言いたいのだから，これがゼロでない解を持つ，と言いたい．

そこで，この方程式を掃きだし法で解く．この節の基本操作を繰り返し，できるだけ簡単な形になるように頑張るのである．ここで「簡単な形」というのは，(??) のような階段状のものを指す．（黒板で説明するように，いつでもこの階段状の形には持っていける．）具体的には

$$\begin{cases} x_1 + a'_{12}x_2 + a'_{13}x_3 + a'_{14}x_4 + \cdots + a'_{1n}x_n = 0 \\ x_2 + a'_{23}x_3 + a'_{24}x_4 + \cdots + a'_{2n}x_n = 0 \\ x_4 + \cdots + a'_{2n}x_n = 0 \\ \cdots = \cdots \\ x_\ell + \cdots = 0 \end{cases} \tag{1.1.4}$$

のような形になっている．（上では３行目が x_4 から始まっているが，そうとは限らない．だけど，このように階段状になるのは間違いない．）

[1] ただし，斉次の方程式の場合はいつでも「すべてゼロ」の解があるから，(c) の可能性はない

さて，階段状になれば，どのような解があるかは明らかになる．つまり，下の方から順次解いていけばよい．このとき，一番下の式が２つ以上の x_i を含んでいればこれで証明終わりである．と言うのも，そのような式は必ず,「すべてがゼロ」とは限らない解を持ち，これを上のそれぞれの方程式に代入して解けば，ゼロでない解が得られるからである．

不幸にして一番下の式が

$$x_n = 0 \tag{1.1.5}$$

となっていれば，ここでは話がすまない．これを上のところにすべて代入し，x_n をなくした式を改めて解く．下から２番目の式が $x_{n-1} = 0$ でなければオシマイ．もし x_{n-1} ならもう一つ上を見る．こうやって上っていくが，方程式の数が未知数の数より多いから，絶対にどこかでゼロ以外の解が入ってくるはずである.（このところはすぐ後で，行列の「階数」と関連させてもう一度扱う.） □

1.2 行列の基本変形と行列の階数

この節の内容は教科書の 2.2 節である．

では，上に挙げた３つの可能性がどのようにして出てくるのか，連立方程式の一般論と併せて考えよう．また，上ででてきた「行列の基本変形」についてももう少し考える．

まず，行列の「階数」の概念を思い出そう．春学期の命題 4.5.3 は以下のようになっていた：

命題 4.5.3 $m \times n$ 行列 A の階数 rank A は，A の n 個の列ベクトル $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n$ の中の一次独立なベクトルの最大数に等しい．

実は行列の階数は「行の基本変形」を用いて計算できるのである．行の基本変形とは

(0) ２つの行を入れ替える．
(a) １つの行に，別の行の定数倍を加える．
(b) １つの行にゼロでない数をかける．

であった．この３つの操作を繰り返して，変形後の行列を階段状にした場合，ゼロでない数の入った行の数がその行列の階数になる．つまり，この操作をやることで，行列に入っている独立な行の数が，基本変形をやることでわかるのである．

なぜこのような操作で階数がわかるのか，は以下の定理 1.2.3 からわかる．その前に，基本変形の性質を列挙する．

定理 1.2.1 行列 A に対する行の基本変形は，ある正則行列を A の左からかけることで表現できる．

（証明）教科書 p.104 にある通り．具体的に，かけるべき行列を書き下してやればよい． □

定理 1.2.2 行の基本変形は，可逆である．つまり，行列 A に行の基本変形をほどこして行列 B になったとすると，B に（別の）基本変形を施して A に戻ることができる．

（証明）教科書 p.106 にある．しかしこんなことをやらなくても，具体的にどのようにすれば戻れるか，考えてみればすぐにわかる． □

以上の準備の元に,「行列の階数の求め方」を基礎づけられる．

定理 1.2.3 行の基本変形により，行列の階数は不変である．

（証明）教科書 p.106 参照． □

基本変形によって階数が変わらないのだから，行列の階数は（基本変形を繰り返した後で）計算しやすい形の行列にしてから計算すれば良い．ところが，階段状になった行列の階数が，その対角線上のゼロでない成分の数に等しいことはすぐにわかる．従って，この節の最初に述べた階数の計算法が正当化される．

1.3 連立方程式の解の構造

この節の内容は教科書の 2.1 節である．

最後に，上に述べた行列の階数と連立方程式の階の構造の関係についてまとめておく.（以下の内容を必要以上に恐れる必要はない．要点は既に，掃き出し法による具体的な解法で尽きている.）

定理 1.3.1 連立方程式 $A\boldsymbol{x}=\boldsymbol{b}$ が解を持つ必要十分条件は $\mathrm{rank}\ A = \mathrm{rank}\ (A,\boldsymbol{b})$ が成り立つことである．

（証明）教科書の p.99 にあるが，以下のように考えても良い．

（解があればランクが等しいことの証明）もともと，連立方程式 $A\boldsymbol{x}=\boldsymbol{b}$ とは，A の列ベクトルを $\boldsymbol{a}_1,\boldsymbol{a}_2,\ldots,\boldsymbol{a}_n$ と書いたとき，

$$\boldsymbol{x}_1\boldsymbol{a}_1+\boldsymbol{x}_2\boldsymbol{a}_2+\ldots+\boldsymbol{x}_n\boldsymbol{a}_n=\boldsymbol{b} \tag{1.3.1}$$

という関係だった．これは $\boldsymbol{a}_1 \sim \boldsymbol{a}_n$ の線型結合によって $\boldsymbol{b}$ が表せる，特に $\boldsymbol{b}$ は $\boldsymbol{a}_1 \sim \boldsymbol{a}_n$ とは独立でない，ということだ．つまり，$\boldsymbol{a}_1 \sim \boldsymbol{a}_n$ と $\boldsymbol{b}$ を併せたものの中の一次独立なベクトルの最大数は，$\boldsymbol{b}$ があってもなくても変わらない．よって，前学期の命題 4.5.3 を思い出すと，A と $(A,\boldsymbol{b})$ のランクは等しい．

（ランクが等しいならば解が存在することの証明）前学期の命題 4.5.3 を思い出すと，ランクが等しいならば「$\boldsymbol{a}_1 \sim \boldsymbol{a}_n$ と $\boldsymbol{b}$ を併せたものの中の一次独立なベクトルの最大数」は「$\boldsymbol{a}_1 \sim \boldsymbol{a}_n$ の中の一次独立なベクトルの最大数」に等しい．ということは,（$\boldsymbol{b}$ があってなくても最大数が変わらないのだから）$\boldsymbol{b}$ は $\boldsymbol{a}_1 \sim \boldsymbol{a}_n$ の線型結合で書けているということだ（もし書けていないとどうなるかを考えよ）．これは (1.3.1) に他ならない． □

定理 1.3.2 A を $m\times n$ 行列とし，n 個の未知変数に対する斉次連立方程式 $A\boldsymbol{x}=\boldsymbol{0}$ を考える．この方程式の解の自由度（解に現れる任意パラメーターの数）は $n-\mathrm{rank}\ A$ に等しい．

（証明）A の表す線型変換を L_A と書くと，次元定理から，

$$\dim(\mathrm{Im}\ L_A)+\dim(\ker L_A)=n \tag{1.3.2}$$

である．ここで $\dim(\ker L_A)$ は $A\boldsymbol{x}=\boldsymbol{0}$ の解の自由度に他ならない．また，$\dim(\mathrm{Im}\ L_A)$ は行列 A の階数そのものである． □

1.4 逆行列の求め方

この章の最後に,「逆行列」の簡単な計算法を眺めておこう．定義を思い出すと，$n \times n$ 行列 A に対して（I_n は $n \times n$ の単位行列）

$$AB = BA = I_n \tag{1.4.1}$$

となる $n \times n$ 行列 B を A の逆行列と言ったのだった．

まず，以下の２つの定理を証明する．以下では（これまでと同じく）$n \times n$ 行列 A の表す線型写像を L_A とかく．

定理 1.4.1 A を $n \times n$ 行列とする．A の階数が n ならば，A は正則行列である．

（証明）A の定める線型写像 L_A を考える．L_A の階数が n ということは，L_A の像空間が $\mathbb{R}^n$ 全体ということである．すなわち，任意の $\boldsymbol{y} \in \mathbb{R}^n$ に対して，$L_A(\boldsymbol{x}) = \boldsymbol{y}$ となる $\boldsymbol{x} \in \mathbb{R}^n$ を探してくることができる．

さらに，このような $\boldsymbol{x}$ は（$\boldsymbol{y}$ を決めれば）一意に決まる．なぜなら，もし $\boldsymbol{x}, \boldsymbol{x}'$ の２つが $L_A(\boldsymbol{x}) = L_A(\boldsymbol{x}') = \boldsymbol{y}$ を満たしていれば，$L_A(\boldsymbol{x} - \boldsymbol{x}') = \boldsymbol{0}$ となるが，次元定理から，$\dim(\ker L_A) = n - \mathrm{rank}\ A = 0$ であるから，$\boldsymbol{x} - \boldsymbol{x}' = \boldsymbol{0}$ となるからである．

従って，$\boldsymbol{y}$ に $\boldsymbol{x}$ を対応させる写像 $M : \mathbb{R}^n \to \mathbb{R}^n$ を定義することができるが，この M は L_A の逆写像になっている．実際，$M(\boldsymbol{y}) = \boldsymbol{x}$ の定義から

$$M(L_A(\boldsymbol{x})) = M(\boldsymbol{y}) = \boldsymbol{x} \tag{1.4.2}$$

および

$$L_A(M(\boldsymbol{y})) = L_A(\boldsymbol{x}) = \boldsymbol{y} \tag{1.4.3}$$

が成り立つからである.（ここでもちろん，上の $\boldsymbol{x}, \boldsymbol{y}$ は $\mathbb{R}^n$ の任意の元になりうることを使っている.）

さらに，この M は線型写像である．なぜなら，$\boldsymbol{y}_1 = L_A(\boldsymbol{x}_1), \boldsymbol{y}_2 = L_A(\boldsymbol{x}_2)$ とすると L_A の線型性から（任意のスカラー α, β に対し）$L_A(\alpha\boldsymbol{x}_1 + \beta\boldsymbol{x}_2) = \alpha\boldsymbol{y}_1 + \beta\boldsymbol{y}_2$ となる．したがって，M の定義から $M(\alpha\boldsymbol{y}_1 + \beta\boldsymbol{y}_2) = \alpha\boldsymbol{x}_1 + \beta\boldsymbol{x}_2$（線型性）がいえるからである．

以上から L_A の逆線型写像 M の存在が証明できた．そこでこの線型写像 M の表現写像を B と書くことにすると,（合成写像 $L_A M$ の表現行列は AB であることから）

$$AB = BA = I_n \tag{1.4.4}$$

がなりたつ．これは B が A の逆行列であることを主張している． □

定理 1.4.2 A, B を $n \times n$ 行列とすると，以下の３つは同値である．
(1) $AB = BA = I_n$
(2) $AB = I_n$
(3) $BA = I_n$

(1) は「A の逆行列が B」である定義そのものだから，上の定理はこの条件が (2) または (3) のどちらか一方で十分，ということを保証するものである．

（証明） (1) は (2) や (3) を含むから，(2) から (1) が出ることを示そう（(3) から (1) も同様）．つまり，$AB = I_n$ ならば A は正則で $B = A^{-1}$ であることを示せば良い．さて，$AB = I_n$ なら，教科書の定理 4.5.1 から，

$$\min(\mathrm{rank}\ A, \mathrm{rank}\ B) \geq \mathrm{rank}\ AB \geq \mathrm{rank}\ I_n = n \tag{1.4.5}$$

が成り立つ．つまり，A, B ともにその階数は n なのだ．そこで上の定理 1.4.1 から直ちに，A, B ともに正則行列であることがいえる．そこで A の逆行列を A^{-1} と書き，これを $AB = I_n$ の左からかけると

$$A^{-1} = A^{-1}AB = I_n B = B \tag{1.4.6}$$

が得られる．よって $BA = I_n$ である． □

逆行列の具体的計算法

定理 1.4.2 は要するに，A の逆行列 X を求めたければ $AX = I_n$ となる行列を求めれば良い，と主張している．となれば，逆行列を求めるのは簡単だ．X の列ベクトルを $\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3, \ldots, \boldsymbol{x}_n$ と書いてみると，$AX = I_n$ は

$$A\boldsymbol{x}_1 = \boldsymbol{e}_1, \quad A\boldsymbol{x}_2 = \boldsymbol{e}_2, \quad \ldots, \quad A\boldsymbol{x}_n = \boldsymbol{e}_n \tag{1.4.7}$$

ということだ．このそれぞれは n 個の変数に対する連立方程式であるが，この解き方は既に「掃き出し法」としてやっているから，それを n 回くりかえせば $\boldsymbol{x}_1$ から $\boldsymbol{x}_n$ が求まる．

実はもう少し，工夫できる．掃き出し法では拡大係数行列 $(A, \boldsymbol{e}_j)$ を基本変形するが（$j = 1, 2, \ldots, n$），係数行列 A はすべての j に共通である．そこで，$\boldsymbol{e}_j$ のところもまとめて並べてしまって，

$$(A, \boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3, \ldots, \boldsymbol{e}_n) = (A, I_n) \tag{1.4.8}$$

という（超）拡大係数行列を考えてみよう．これに行の基本変形をやって，左の A の部分が I_n になるまで変形すると，残った右側がちょうど求める X になる．つまり

$$(A, I_n) \quad \xrightarrow{\text{行の基本変形}} \quad (I_n, X) \tag{1.4.9}$$

これが A の逆行列 $X = A^{-1}$ の求め方である．

連立方程式の場合と同じく，逆行列を求めたら，検算すること！

2 行列式

この節では「行列式」について学ぶ．この節の内容は概念的には難しいものではない（ハズだ）し，教科書にも詳しく載っているので，レジュメは簡単なものになる 予定である．ただし，ある程度の計算練習をしておかないと試験の時に（また将来，物理学科で）困るだろうから，練習しておいてくださいね．

（本論にはいる前に：何のために行列式をやるのか？）

- 行列式にはいろいろな意味づけができるが，この講義にとって一番大事なのは，「正則行列＝行列式がゼロでない行列」と言うことだ．つまり，行列式を計算することで，その行列が 正則か正則でないか がわかる．
- 「行列が正則か正則でないか」の判定は後で「行列の固有値と固有ベクトル」を求める上で不可欠になる．
- と言うわけで，行列式は後々に重要になってくるので，その計算方法を知ることが重要なのである．

（少し進んだ注）

行列式の定義には大体，以下の 3 とおりがある．(以下では細かい用語などはわからなくても，大体の感じが伝わればよい.）これらは互いに同値で，どれかを定義にとって，残りの２つを証明することができる．

(a) 行列式の定義式を「置換」「互換」から陽に与える．
  - 長所：陽に定義式を与えているので安心できる．大方の教科書はこの方法であろう．
  - 短所：定義そのものが（置換・互換の定義から始めると）かなり長い．更に，「行列式の基本的性質」を導出した後では，このように苦労した定義はほとんど使わない．

(b) 「行列式の基本的性質」（定理 2.2.1) を満たすような多重線型汎関数として定義する．
  - 長所：一番大事な性質そのものを定義とするので，無駄がない．
  - 短所：そのような性質を満たすものが実際に存在するのか，存在するとして一意に定まるか（要するに定義がきちんとできているのか）の検証が案外と厄介で，結局，(a) の定義を導出することになる．

(c) $n \times n$ 行列の行列式を $(n-1) \times (n-1)$ 行列の行列式を用いて，帰納的に定義する．
  - 長所：2×2 行列の行列式は知っているから，定義そのものは明確.
  - 短所：帰納的に定義しているから，「行列式の基本的性質」などを導くのが厄介.

いろいろと考えたのだが，この講義では教科書通りに (a) の方針に従うことにする（この講義ノートの 2.1 節）．その後で，上の (b) の部分（講義ノートの 2.2 節），および (c) の部分（講義ノートの 2.3 節），の順番で進む．

2.1 行列式の定義

この節では行列式を定義する．2×2 または 3×3 行列の行列式は高校でもやったと思うが，この節では 4×4 以上でも成り立つ定義を与える．

2.1.1 置換と互換

しばらく定義が続くが，我慢して欲しい．教科書の該当部分は 6.1,6.2 節で，かなり細かく書いてある．しかし，ここまで細かくやると本筋を外れすぎる恐れもあるので，要点だけを簡単に述べる．

n 個の数字 $1, 2, 3, \ldots, n$ を重複なしで並べたものを 1〜n の**順列**と言うのは高校でやったはずだ．ここではこれを発展させる．

順列を表すために，上の行に 1〜n，下の行にその順列の並び方（数字）を並べて行列のように書いてみよう．例えば（$n = 5$），

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 3 & 4 & 5 \end{pmatrix}, \qquad \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 2 & 1 & 4 & 5 \end{pmatrix}, \qquad \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 4 & 3 & 2 & 1 \end{pmatrix} \tag{2.1.1}$$

（の下の行）はどれも 1〜5 の順列である．このように書くと，順列は上の行の 1〜n の数字のそれぞれに下の行の数字を対応させる変換（写像）と考えられる．そこで，この対応関係を**置換**と呼ぶ．置換には σ, τ のような記号

を用いることが多い．また，置換 σ による数字 j の行き先（変換先）を $\sigma(j)$ と書く．(2.1.1) の真ん中の例なら $\sigma(1)=3, \sigma(2)=2, \sigma(3)=1$ などとなる．

（いくつかの注意）

1．「置換」と言うときは上の対応関係（つまり，1 に対応するのはどの数字か，2 に対応するのはどの数字か,...）のみに注目する．つまり，一行目を並べ替えても，それに応じて２行目も並べ替えておけば，両者は同じ置換とみなす．例：

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 2 & 1 & 4 & 5 \end{pmatrix} = \begin{pmatrix} 5 & 4 & 3 & 2 & 1 \\ 5 & 4 & 1 & 2 & 3 \end{pmatrix} \tag{2.1.2}$$

2．(2.1.1) の最初の置換は，どの数字も変わっていない．これを 恒等置換 または 単位置換 と呼ぶ．教科書では単位置換を ϵ の記号で表している．

3．置換 σ の置き換えられた数字を元に戻す置換を σ の 逆置換 と呼び，σ^{-1} と書く．

4．２つの置換 σ, τ が与えられたとき，その 積 を，この２つの置換を続けて行ったものとして定義する．つまり，数字 i を数字 $\tau(\sigma(i))$ に変える変換を $\tau\sigma$ と書き（順序に注意），σ と τ の積と定義するのである．

5．上の積の定義によると，逆置換は

$$\sigma^{-1}\,\sigma = \sigma\,\sigma^{-1} = \epsilon \text{（単位置換）} \tag{2.1.3}$$

を満たすものである，と言える．

２つの数字だけ入れ替え，残りの数字は変えないような置換を 互換 と呼ぶ．互換は勿論，置換の一種であるし，互換を重ねて行ったものは置換になっている．大事なのはその逆も成り立つことである：

- 任意の置換は，適当な互換の積として表せる（教科所の定理 6.2.1）．
- 一つの置換を互換の積として表す表し方は一通りではない．しかし，互換が 偶数個 必要か 奇数個 必要かは置換によって一意に決まる（教科所の定理 6.2.2）．

以上を認めて：

定義 2.1.1 置換の符号 を，その置換を互換の積で表したときに

- 偶数個の互換が必要ならば +1（このような置換を 偶置換 と言う），
- 奇数個の互換が必要ならば −1（このような置換を 奇置換 と言う），

と定義する．置換 σ の符号を $\mathrm{sgn}(\sigma)$ と書く．

2.1.2 行列式の定義

これでいよいよ行列式を定義することができる（教科書 6.3 節）．

定義 2.1.2 $n \times n$ 行列 A の**行列式** $\det A$ を，

$$\det A \equiv \sum_{\sigma} \mathrm{sgn}(\sigma)\, a_{\sigma(1)1}\, a_{\sigma(2)2}\, \ldots\, a_{\sigma(n)n} = \sum_{\sigma} \mathrm{sgn}(\sigma)\, a_{1\sigma(1)}\, a_{2\sigma(2)}\, \ldots\, a_{n\sigma(n)} \tag{2.1.4}$$

によって定義する．上の和は 1〜n の置換全体についてとる．

注意：上の定義には式が二つ書いてあるから，この二つが同じものであることはもちろん，確かめないといけない．

（定義からすぐにわかる例；各自，確かめること！）

- 2×2 行列の行列式は高校でやったよね．上の定義が高校でやったものに合致することを確かめよう．

- 対角行列の行列式は簡単だ．A が $n \times n$ の対角行列の場合，$\det A = a_{11}a_{22}\ldots a_{nn}$ （対角成分の積）になる．
- 同じく，A が $n \times n$ の上半三角行列の場合，その行列式も $\det A = a_{11}a_{22}\ldots a_{nn}$ （対角成分の積）になる．

高校で 3×3 の行列式（たすきがけ，またはサラスの方法）をやった人も多いだろう．しかし，このような簡単な方法は 4×4 以上ではなりたたないので注意 — 例年，この点を強調するが，それでもテストでたすきがけをする人がいる．なお，この講義ではサラスの方法は学習しない（ので，知らなくても良い）．

やる気のある人への問題：
適当な 4×4 行列（例えば，今日返却したレポートの A）の行列式を，上の定義に基づいて（24 個の項の和を計算することで）求めてみよ．一度でもこれをやっておくと，以下で習う方法のありがたみがよくわかる．

2.2 行列式の性質と計算法

前節では行列式を定義したが，この定義では行列の次数が高くなると非常に大変でやってられない．($n \times n$ 行列なら $n!$ 個の置換があるわけで，そいつらについての和を計算するのは大変！）この節では行列式の満たす性質を調べることで，もっと簡単に行列式を計算することを考える．

2.2.1 行列式の性質

基本になるのは以下の定理である（教科書 6.4 節）．

定理 2.2.1 (教科書の定理 6.4.1, 6.4.2, 6.4.3) 以下の式では，対応する $\cdots$ のところは同じものと解釈する．（詳しくは講義で！）

- 行列の i 列と j 列を入れ替えると ($i \neq j$)，行列式の符号は変わる：

$$\det[\cdots, \boldsymbol{a}_i, \cdots, \boldsymbol{a}_j, \cdots] = -\det[\cdots, \boldsymbol{a}_j, \cdots, \boldsymbol{a}_i, \cdots] \tag{2.2.1}$$

- 行列の一つの列を c 倍（c はスカラー）すると，行列式の値は c 倍になる：

$$\det[\cdots, c\boldsymbol{a}_i, \cdots] = c\,\det[\cdots, \boldsymbol{a}_i, \cdots] \tag{2.2.2}$$

- 一つの列をたすと，行列式も和になる：

$$\det[\cdots, \boldsymbol{a}_i + \boldsymbol{b}_i, \cdots] = \det[\cdots, \boldsymbol{a}_i, \cdots] + \det[\cdots, \boldsymbol{b}_i, \cdots] \tag{2.2.3}$$

（注意）

- (2.2.2) について：たった一つの列を c 倍するだけで行列式全体が c 倍になる．行列全体を c 倍すると，$n \times n$ 行列の行列式は c^n 倍になる．**ここは間違いやすいから注意！**
- 上の定理は「列」について述べたが,「列」をすべて「行」に読み替えても定理は成立する．

$n \times n$ 行列 A の行と列を入れ替えたもの，つまり (i, j) 成分が a_{ji} で与えられる行列を A の転置行列と言い，${}^tA$ と書く．(2.1.4) は以下を意味する．

定理 2.2.2 (教科書の定理 6.4.5) 転置行列の行列式は元の行列の行列式に等しい：

$$\det(A) = \det({}^tA) \tag{2.2.4}$$

（証明）要するに，(2.1.4) の右側の等号を証明すればよい．真ん中の $\sum_\sigma \mathrm{sgn}(\sigma)\, a_{\sigma(1)1}\, a_{\sigma(2)2}\, \ldots\, a_{\sigma(n)n}$ から出発し，$\sigma^{-1} = \tau$ （逆置換）として和を書き直す．σ についての和と τ についての和は同じ事だから，

$$\sum_\sigma \mathrm{sgn}(\sigma)\, a_{\sigma(1)1}\, a_{\sigma(2)2}\, \ldots\, a_{\sigma(n)n} = \sum_\tau \mathrm{sgn}(\tau^{-1})\, a_{1\tau(1)}\, a_{2\tau(2)}\, \ldots\, a_{n\tau(n)} \tag{2.2.5}$$

となる（ここで右辺では a_{ij} の積を並べ替えている）．$\mathrm{sgn}(\tau^{-1}) = \mathrm{sgn}(\tau)$ に注意すると，(2.1.4) の右辺が得られる． □

これらの性質を使うと，行列式を効率よく計算することができる．その際に以下の性質も使うので掲げておく（なぜ，以下の二つの性質が成り立つのかは，上の定理 2.2.1 から出る．各自で考えておくこと）．

- ２つの列が等しい行列の行列式はゼロである．
- 一つの列のスカラー倍を他の列に加えても，行列式の値は変わらない．

（上の２つは「列」を「行」と読み替えても成立する.）

行列の積の行列式についても簡単に触れておく．この定理は教科書では 6.7 節にあるが，この位置に持ってきた方がつながりが良い．

定理 2.2.3 (教科書の定理 6.7.1) $n \times n$ 行列 A, B の積について，以下の等式が成り立つ．

$$\det(AB) = \det(A)\, \det(B) \tag{2.2.6}$$

要するに，積の行列式は行列式の積なのだ．

系 2.2.4 行列 A が正則の時，

$$\det\bigl(A^{-1}\bigr) = \frac{1}{\det(A)} \tag{2.2.7}$$

つまり，逆行列の行列式はもとの行列の行列式の逆数である．

2.2.2 行列式の計算法

さて，行列式をどうやって計算するか，考えてみよう．大ざっぱな指針は，行（と列）の基本変形をくり返してできるだけ簡単な形（上半三角，または対角行列）に持っていくことである．ただし，連立方程式を解くのとは違って,「ある行をスカラー倍」したり「行を入れ替え」たりしたら，行列式の値が変わってくるから注意すること.（例を使って見せた方が速いので，講義で説明する.）

行列式の計算方法について．簡単にまとめると以下の通りである．

- 行の基本変形，または列の基本変形を用いて，行列式が簡単に計算できる形に変形する．その際，**行列式の値が変わってくる**かもしれないから注意する．
- 変形先としては，すぐ下に掲げてあるどれかの形を目指す．

今までのところで既に証明でき，かつよく使う行列式の性質を列挙しておこう.（いくつかは既に述べた.）簡単のため，教科書にもならって，行列式を $\det(A)$ の代わりに $|A|$ と書く．

（1．上半三角，または下半三角）

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = a_{11}\, a_{22} \cdots a_{nn} \tag{2.2.8}$$

（2．ある行が一つの成分以外，全部ゼロ）

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1,n-1} & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2,n-1} & a_{2n} \\ \vdots & \vdots & & \vdots & \\ a_{n-1,1} & a_{n-1,2} & \cdots & a_{n-1,n-1} & a_{n-1,n} \\ 0 & 0 & \cdots & 0 & a_{nn} \end{vmatrix} = a_{nn} \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1,n-1} \\ a_{21} & a_{22} & \cdots & a_{2,n-1} \\ \vdots & \vdots & & \vdots \\ a_{n-1,1} & a_{n-1,2} & \cdots & a_{n-1,n-1} \end{vmatrix} \tag{2.2.9}$$

（3．ある行が一つの成分以外，全部ゼロ）

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1,n-1} & 0 \\ a_{21} & a_{22} & \cdots & a_{2,n-1} & 0 \\ \vdots & \vdots & & \vdots & \\ a_{n-1,1} & a_{n-1,2} & \cdots & a_{n-1,n-1} & 0 \\ a_{n1} & a_{n2} & \cdots & a_{n,n-1} & a_{nn} \end{vmatrix} = a_{nn} \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1,n-1} \\ a_{21} & a_{22} & \cdots & a_{2,n-1} \\ \vdots & \vdots & & \vdots \\ a_{n-1,1} & a_{n-1,2} & \cdots & a_{n-1,n-1} \end{vmatrix} \tag{2.2.10}$$

以上はすべて，行列式の定義から導かれるものである．基本変形を行って行列式を計算する際，上のどれかの性質が使える形に変形することを目指す．

書くのが面倒だから，行列式の計算の具体例は黒板で説明する．各自，練習しておくように.（レポートとしても出題した.）

2.3 行列式の展開

（この小節の内容は教科書では 6.5 節）

さて，今までのところで一応，行列式は計算できるようになった．列と行，両方の基本変形を使えるので，連立方程式を解くより簡単である場合が多い．実際の計算法としてはこれで十分とも言える．しかし，理論的興味もあり，ここでは次数の高い行列の行列式を，より低い次数の行列の行列式に関連づける方法を学習する．

この節の主要な結論を書くために，まず,「余因子」を定義する．

> $n\times n$ 行列 A を考える．A の第 i 行と第 j 列を抜き取った行列を考える（これは $(n-1)\times(n-1)$ 行列）．この $(n-1)\times(n-1)$ 行列の行列式に $\underline{(-1)^{i+j}\text{ をかけた}}$ ものを $\underline{\text{行列 } A \text{ の } (i,j) \text{ 余因子}}$ と呼び，α_{ij} で表す．

（注意）教科書では α_{ij} から $(-1)^{i+j}$ を取り去ったもの（つまり，行列式そのもの）を $|A_{ij}|$ と書いており，余因子は 6.7 節まででてこない．これは多分，$(-1)^{i+j}$ の因子を忘れる学生が多いことに手を焼いた著者達の工夫であろう．しかし，この書き方は標準的ではないので将来，皆さんが混乱する可能性が高い上に，$(-1)^{i+j}$ のあるのとないのと両方が出てくるのはイヤだ．そこで敢えて従来通りの記号法を初めから採用した．(ただし，教科書との対応をつけるため，$|A_{ij}|$ の表式も書いておく.）なお，余因子に α を使うと a との区別が付きにくいのではないかと思うが，教科書がこうなっているので従うことにした．

> **定理 2.3.1 (教科書の定理 6.5.1)** $n\times n$ 行列 A について，以下の等式が成り立つ．
>
> $$\text{任意の } 1\le j\le n \text{ を固定した場合} \qquad \det A=\sum_{i=1}^{n} a_{ij}\,\alpha_{ij}=\sum_{i=1}^{n}(-1)^{i+j}\,a_{ij}\,|A_{ij}| \tag{2.3.1}$$
>
> $$\text{任意の } 1\le i\le n \text{ を固定した場合} \qquad \det A=\sum_{j=1}^{n} a_{ij}\,\alpha_{ij}=\sum_{j=1}^{n}(-1)^{i+j}\,a_{ij}\,|A_{ij}| \tag{2.3.2}$$
>
> 一つ目を第 j 列に関する余因子展開，２つ目を第 i 列に関する余因子展開，と呼ぶ．

しつこいが，α_{ij} は $(n-1)\times(n-1)$ 行列の行列式である．上の定理は $n\times n$ 行列の行列式を $(n-1)\times(n-1)$ 行列の行列式から帰納的に定義する式だとも思える．この意味で，この展開式はこの節の最初に述べた (c) の方法を与えているとも言える．

行列式の展開に関連しては，以下の定理が応用上も大事である：

> **定理 2.3.2 (教科書の定理 6.5.2)** $n\times n$ 行列 A について，以下の等式が成り立つ．
>
> $$\text{任意の } 1\le j,k\le n \text{ を固定した場合} \qquad \sum_{i=1}^{n} a_{ik}\,\alpha_{ij}=\delta_{kj}\,\det(A) \tag{2.3.3}$$
>
> $$\text{任意の } 1\le i,k\le n \text{ を固定した場合} \qquad \sum_{j=1}^{n} a_{kj}\,\alpha_{ij}=\delta_{ki}\,\det(A) \tag{2.3.4}$$
>
> ここで δ_{ij} はクロネッカーのデルタで，
>
> $$\delta_{ij}\equiv\begin{cases}1 & (i=j)\\ 0 & (i\ne j)\end{cases} \tag{2.3.5}$$
>
> という記号である．

註： 上の定理は，$k=j$ または $k=i$ の時には定理 2.3.1 で既に証明されている．この定理の新しさは，$k\ne j$ や $k\ne i$ の場合にある．

さて，上の定理を読み替えると以下の定理になる．ただし，この証明には「行列の積の行列式はそれぞれの行列の行列式の積である」こと（定理 2.2.3）を用いる．

定理 2.3.3 (教科書の定理 6.7.2) $n \times n$ 行列 A について，以下が成り立つ.

$$A \text{ が正則である} \qquad \Longleftrightarrow \qquad \det A \neq 0 \tag{2.3.6}$$

更に，A の逆行列の ij 成分 $[A^{-1}]_{ij}$ は（添え字の順序に注意）：

$$\left[A^{-1}\right]_{ij} = \frac{1}{\det(A)}\,\alpha_{ji} \tag{2.3.7}$$

で与えられる．つまり，A^{-1} は A の余因子行列を $\det(A)$ で割ったものになる —— しつこいけど，添字の順序に注意.

（略証）まず，定理 2.3.1 を

$$A\,A^{-1} = I_n \tag{2.3.8}$$

の両辺に用いると，

$$\det(A)\ \det(A^{-1}) = 1 \tag{2.3.9}$$

が得られる．逆行列 $A^[-1]$ が存在する限り，$\det(A^{-1})$ は無限大ではないから，上の式から $\det(A) \neq 0$ が結論できる．つまり，A が正則の場合は $\det(A) \neq 0$ なのである.

逆に，$\det(A) \neq 0$ であれば，(2.3.7) の右辺の行列は定義できる．更に，定理 2.3.2 を読み替えれば，この行列が確かに A の逆行列であること（つまり $AA^{-1} = I_n$ であること）がわかる．よって定理は証明された.

2.4 クラメールの公式

前節まででで行列式に関してやりたいことは大体，終わった．この節では「クラメールの公式」というものを紹介して，この章を締めくくろう.

連立一次方程式

$$A\boldsymbol{x} = \boldsymbol{b} \tag{2.4.1}$$

を考える．ここで $A = [\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n]$ は $n \times n$ 行列，$\boldsymbol{b}$ は与えられた n 項列ベクトル，$\boldsymbol{x}$ は未知数の作る n 項列ベクトルである．以下では特に，A が正則行列の場合を考える．この場合，以前にも注意したように

$$\boldsymbol{x} = A^{-1}\,\boldsymbol{b} \tag{2.4.2}$$

と解くことができる（ここまでは復習）.

この節ではこれを行列式を使って書く,「クラメールの公式」を考える.

定理 2.4.1 (クラメールの公式，教科書の定理 6.6.1)
A が正則の時，連立方程式 (2.4.1) の解 (2.4.2) は，以下の形に書ける：

$$x_j = \frac{1}{|A|}\ \det[\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_{j-1}, \boldsymbol{b}, \boldsymbol{a}_{j+1}, \ldots, \boldsymbol{a}_n] = \frac{\det[\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_{j-1}, \boldsymbol{b}, \boldsymbol{a}_{j+1}, \ldots, \boldsymbol{a}_n]}{\det[\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_{j-1}, \boldsymbol{a}_j, \boldsymbol{a}_{j+1}, \ldots, \boldsymbol{a}_n]} \tag{2.4.3}$$

要するに，x_j の表式は 2 つの行列式の比で与えられる．一つの行列式（分母）は，$\det(A)$ そのものである．一方，分子は分母にでてくる行列 A において，第 j 列を $\boldsymbol{b}$ で置き換えたものの行列式になっている.

(証明)

(2.4.2) は

$$\boldsymbol{b} = x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + \ldots + x_n\boldsymbol{a}_n \tag{2.4.4}$$

と書けることに注意すると分子の行列式は

$$\det[\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_{j-1}, \boldsymbol{b}, \boldsymbol{a}_{j+1}, \ldots, \boldsymbol{a}_n] = \det[\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_{j-1}, x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + \ldots + x_n\boldsymbol{a}_n, \boldsymbol{a}_{j+1}, \ldots, \boldsymbol{a}_n]$$

となる．この右辺において行列式の性質をつかってやると

$$
\begin{aligned}
&= x_1 \det[\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_{j-1}, \boldsymbol{a}_1, \boldsymbol{a}_{j+1}, \ldots, \boldsymbol{a}_n] + x_2 \det[\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_{j-1}, \boldsymbol{a}_2, \boldsymbol{a}_{j+1}, \ldots, \boldsymbol{a}_n] \\
&\quad + \ldots + x_n \det[\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_{j-1}, \boldsymbol{a}_n, \boldsymbol{a}_{j+1}, \ldots, \boldsymbol{a}_n]
\end{aligned}
$$

と変形できる．ところが，$\boldsymbol{a}_j$ が入っている項以外は（２つの列が比例するので）ゼロになり，結局，

$$
= x_j \det[\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_{j-1}, \boldsymbol{a}_j, \boldsymbol{a}_{j+1}, \ldots, \boldsymbol{a}_n] = x_j \ \det A \tag{2.4.5}
$$

となる．両辺を $\det A$ で割るとできあがり． □

実際に連立方程式を解く場合，クラメールの公式を使うことはあまり無い（計算量が膨大になってしまいがちだから —— 掃き出し法のほうが余程簡単である）．しかし，ある種の理論的な解析を行う場合，クラメールの公式が役立つことがある．

3 固有値と固有ベクトル

(物理からの動機付け)

これから，全く新しい概念「固有値と固有ベクトル」について学ぶ．この題材にはもちろん，以下に述べるような線型代数として重要な役割はある．しかし，物理学科の学生さんとしての皆さんには，それよりもまずは量子力学における固有値と固有ベクトルの役割の方が大切だろう．かいつまんでいうと，量子力学においては[2]

- すべての物理的状態はある線型空間のベクトルで書ける．
- すべての物理量（観測するということ）はその線型空間での適当な線型写像で表される．
- 物事を観測した結果（測定値）は，その観測量を表す線型写像の固有値になる（例：物理で多分一番重要な観測値 —— エネルギー —— はハミルトニアンと呼ばれる線型写像の固有値である）

という事情があるのだ.（これだけでは何のことかさっぱりわからんだろうから，近日中に希望者向けの補講を入れるつもりではあるが）ともかく大事なのは，物理屋にとって一番重要であるはずの観測量が（よく訳のわからない）線型写像の「固有値」というものになってしまうことだ．固有値（と固有ベクトル）の概念がよくわかっていないと，ただでさえわかりにくい量子力学を理解することはまず不可能になるだろう．という訳で，物理の学生さんである皆さんには,「固有値と固有ベクトル」を理解すべき重要な動機付けがあるのだ．

(線型代数独自の動機付け)

春学期に，線型写像について見た．特に「線型性」について学び，また，線型写像は（基底を定めることで）行列でかけることも見た．更に，線型写像を何回もやることは行列を何回もかけること，線型写像の逆写像は逆行列で表されること，も見た．

ところが，ある線型写像を何回もやった結果を求めるのは，一般に大変だ．表現行列をそれだけの回数，かける必要があるが，一般に行列の n 乗の計算は非常に大変だから．例えば，

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}, \qquad \boldsymbol{b} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

を考える．$A^n\boldsymbol{b}$ を（例えば $n = 1000$）計算したいと思った場合，闇雲にかけるのはなかなか大変だ．ところが,

$$\boldsymbol{b}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \qquad \boldsymbol{b}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

に対しては，$A^n\boldsymbol{b}_i$ は（$i = 1, 2$）簡単に計算できる．というのも，

$$A\boldsymbol{b}_1 = 3\boldsymbol{b}_1, \qquad A\boldsymbol{b}_2 = -\boldsymbol{b}_2$$

であって，A をかけた結果が自分自身（の定数倍）なのだ．だから

$$A^n\boldsymbol{b}_1 = 3^n\boldsymbol{b}_1, \qquad A^n\boldsymbol{b}_2 = (-1)^n\boldsymbol{b}_2$$

と簡単に計算できる．更に，この事実と，もともとの $\boldsymbol{b}$ を

$$\boldsymbol{b} = \frac{3}{2}\boldsymbol{b}_1 + \frac{1}{2}\boldsymbol{b}_2$$

と表せることから，$A^n\boldsymbol{b}$ も

$$A^n\boldsymbol{b} = \frac{3}{2}\,A^n\boldsymbol{b}_1 + \frac{1}{2}\,A^n\boldsymbol{b}_2 = \frac{3^{n+1}}{2}\boldsymbol{b}_1 + \frac{(-1)^n}{2}\boldsymbol{b}_2 = \frac{1}{2}\begin{bmatrix} 3^{n+1} + (-1)^n \\ 3^{n+1} - (-1)^n \end{bmatrix}$$

と簡単に求められる．当初，ほとんど不可能と思われた $A^n\boldsymbol{b}$ が計算できてしまった．

以上を振り返ると，重要なのは A をかけられても自分自身の定数倍になるような，そんなベクトル（上の例では $\boldsymbol{b}_1, \boldsymbol{b}_2$）だとわかる．このようなベクトルが見つかれば[3] これらをもとに上のように議論して，$A^n\boldsymbol{b}$ が計算できるわけだ．

[2]以下はかなりいい加減な書き方です．ホンのサワリだと思って読んでください

[3]実際にはこのようなベクトルが十分にたくさん，線型空間の基底をなすくらい，見つかれば

更に，ベクトル $\boldsymbol{x}$ を $A\boldsymbol{x}$ に変換する線型写像を一回だけ行う場合でも，その結果は教科書 p.152 の図 18, 図 19 にあるように，行列ごとに違ったものになるが，固有ベクトルはこれらの図でベクトルの向きが変わらない直線として現れる．このような事情から，線型写像や行列によるかけ算をより良く理解するためには，その固有値と固有ベクトルの情報が非常に有効である．これが線型代数において，固有値と固有ベクトルを学修する理由である．

> お断り：いままで，皆さんが複素数に慣れていないだろう事を考えて，スカラーは実数だとしてきた．また，ベクトルもその成分は実数だとしてきた．しかし，固有値や固有ベクトルを考える場合には，これらを複素数まで拡げておく方が見通しが良い．従って，以下で特に断らない限りは,「スカラーは複素数」「数ベクトルの成分も複素数」を許すものとする．

3.1 固有値と固有ベクトル

> **定義 3.1.1 (行列の固有値と固有ベクトル)** $n \times n$ 正方行列 A に対して，
>
> $$A\boldsymbol{v} = \alpha\,\boldsymbol{v} \tag{3.1.1}$$
>
> となるような複素数 α と <u>ゼロベクトルでない</u>n 項列ベクトル $\boldsymbol{v}$ とがある場合（$A\boldsymbol{v}$ は行列とベクトルのかけ算を表す），α を A の <u>固有値</u>，$\boldsymbol{v}$ を A の（α に対する）<u>固有ベクトル</u> と言う．

上の例では，A の固有値は 3 と 1，対応する固有ベクトルはそれぞれ $\boldsymbol{b}_1$ と $\boldsymbol{b}_2$ であった．

（注意）

- $\boldsymbol{v}$ が <u>ゼロベクトルでない</u> のは非常に重要である．と言うのは，任意のスカラー α に対して，$A\boldsymbol{0} = \alpha\boldsymbol{0}$ となってしまうが，これは全然面白くないので，$\boldsymbol{0}$ は固有ベクトルとはいわないから．$\boldsymbol{v} \neq \boldsymbol{0}$ となるためにこそ，α は特別の値を取る必要があるのだ．
- $\boldsymbol{v}$ が固有ベクトルなら，$k\boldsymbol{v}$ も固有ベクトルだ（ただし $k \neq 0$）．この意味で，固有ベクトルは最低限，定数倍だけの無限個の自由度がある．ただし，この自由度はショウモナイものだから，普通は気にしない（「固有ベクトルを求めよ」と言われても，k の自由度まで答える必要はない．それより大きな固有ベクトルの自由度については，後の「固有空間」のところで述べる．

上の定義を線型写像に拡大しておく．

> **定義 3.1.2 (線型写像の固有値と固有ベクトル)** 線型空間 X と線型写像 $f : X \to X$ が与えられているとき，
>
> $$f(\boldsymbol{v}) = \alpha\,\boldsymbol{v} \tag{3.1.2}$$
>
> となるような複素数 α と <u>ゼロベクトルでない</u> ベクトル $\boldsymbol{v} \in X$ とがある場合，α を f の <u>固有値</u>，$\boldsymbol{v}$ を A の（α に対する）<u>固有ベクトル</u> と言う．

固有ベクトルの効用は後で集中的にやる.（この節に入る前の例でも少し見た.）まずは，どのように固有値や固有ベクトルを求めたらよいかを考えよう．

（まず注意）春学期に,「線型写像はその表現行列を用いて書ける」ことを見た．これによると，線型写像の固有ベクトルを求めるには，まず，その表現行列の固有ベクトルを求め，それを焼き直せば良い．従って，以下では行列の固有ベクトルだけを扱うが，上の手順により線型写像の固有ベクトルも扱えることに注意してほしい．

そこで，定義 3.1.1 に戻る．固有ベクトルの定義の式は

$$(A - \alpha\, I_n)\,\boldsymbol{v} = \boldsymbol{0} \tag{3.1.3}$$

と書ける（I_n は $n\times n$ 単位行列）．これは n 連立方程式だが，今はこの方程式のゼロでない解 $\boldsymbol{v}$ を求めたい（$\boldsymbol{v}\neq\boldsymbol{0}$ が存在するような α を見つけたい）わけだ．ゼロでない解が存在するための条件はこれまでにやった．すなわち，

- 行列 $A-\alpha I_n$ が正則行列なら，$\boldsymbol{v}=\boldsymbol{0}$ の解しかない．
- 行列 $A-\alpha I_n$ が正則でないなら，$\boldsymbol{v}\neq\boldsymbol{0}$ な解が存在する

訳だ．つまり，固有値 α の満たすべき必要十分条件は，「行列 $A-\alpha I_n$ が正則でないこと」なのである．

さらに，「行列 A が正則でないことと，$\det A=0$ は同値である」ことも最近，学修した．そこで，上の固有値の必要十分条件は以下のように書き換えられる：

定理 3.1.3 (固有値の必要十分条件) 複素数 α が，$n\times n$ 行列 A の固有値になっているための必要十分条件は

$$\det(A-\alpha I_n)=0 \tag{3.1.4}$$

である．

これで固有値 α をすべて見つけることができる．α さえ求めれば，あとは連立方程式 (3.1.3) を解くことで，$\boldsymbol{v}$ も見つけられる．これで固有値と固有ベクトルを求めるプログラムが完成した．しつこくまとめると：

1. (3.1.4) を解いて，固有値 α を求める（一つとは限らない）
2. それぞれの α に対して (3.1.3) を解いて，$\boldsymbol{v}$ を求める（これも一つとは限らない）

と言うわけだ．計算手順としては簡単でしょ？

（注意）固有値を一つ見つけたとき，それに対する固有ベクトルは何通りもあり得る（ここでは定数倍の自由度以外の「何通りも」を言っている）．例えば，$A=I_n$（単位行列）の場合，固有値は 1 しかないが，固有ベクトルは $\boldsymbol{e}_i$（$i=1,2,\ldots,n$，標準基底）の n 個ある．この自由度についてはすぐ後で「固有空間」としてもっと学習する．

（用語の定義）変数 t の多項式としてみた $\Delta_A(t)=\det(tI_n-A)$ を行列 A の**固有多項式**という．(A が $n\times n$ 行列のとき，$\Delta_A(t)$ は t の n 次多項式で，かつ，t^n の係数は 1 である —— why?）また，固有値を求めるときに解くべき方程式 $\Delta_A(t)=0$ を行列 A の**固有方程式**という．

さて，t の n 次の多項式は一般に複素数の根 $\lambda_1,\lambda_2,\ldots,\lambda_n$ を用いて因数分解できる．$\Delta_A(t)$ の t^n の係数が 1 であることも考えると，

$$\Delta_A(t)=(t-\lambda_1)(t-\lambda_2)\cdots(t-\lambda_n) \tag{3.1.5}$$

と書けるはずだ．ただし，ここで λ_i のいくつかは互いに等しいかも知れない（重解，重根）．λ_1 に等しい λ_i が k 個あるとき，λ_1 は k-重根（λ_1 の重複度は k）という．各根の重複度をすべて加えると当然，n になっている．

（例１）$A=I_n$（$n\times n$ の単位行列）の場合．

$$\det(I_n-tI_n)=\det\bigl((1-t)I_n\bigr)=(1-t)^n\ \det(I_n)=(1-t)^n \tag{3.1.6}$$

なので，固有値は $t=1$ のみ（n 重根）．対応する $\boldsymbol{v}$ の方程式はなんと

$$O\boldsymbol{v}=\boldsymbol{0} \tag{3.1.7}$$

となってしまい（O はすべての成分がゼロの行列），全く $\boldsymbol{v}$ に関しての制限がつかない．(まあ，これは任意のベクトル $\boldsymbol{v}$ に対して $I_n\boldsymbol{v}=\boldsymbol{v}$ であることを思えばアタリマエだ）．という訳で，単位行列の固有ベクトルは「ゼロでない任意のベクトル」なのである．

（例２）A が $n\times n$ の対角行列で，対角成分が $\alpha_1,\alpha_2,\ldots,\alpha_n$ の時（簡単のため，$i\neq j$ なら $\alpha_i\neq\alpha_j$ とする）．

$$\det(A-tI_n)=(\alpha_1-t)(\alpha_2-t)\cdots(\alpha_n-t) \tag{3.1.8}$$

となるので，α_j（$j=1,2,\ldots,n$）が固有値（すべて単根）．α_j に対応する固有ベクトルは $\boldsymbol{e}_j$ の（ゼロでない）定数倍．

（例３）始めに例に挙げた 2×2 行列 A では，固有方程式は

$$0 = t^2 - 4t + 3 \tag{3.1.9}$$

となって，固有値は $1, 3$. それぞれの固有ベクトルを求めると，$\boldsymbol{b}_2, \boldsymbol{b}_1$ が得られる．

（これより複雑な例は，黒板で説明しよう.）

固有空間

α が $n \times n$ 行列 A の固有値であるとき，α に対する A の固有ベクトルに２本以上，独立なものが存在する場合もある．このような場合をうまく扱うため，

α に対する A の固有ベクトルの全体とゼロベクトル，つまり $A\boldsymbol{v} = \alpha\boldsymbol{v}$ なるベクトル $\boldsymbol{v}$ の全体

を W_α と書くことにする．W_α の中にはゼロベクトルも入っていることに注意（W_α の元は，このゼロベクトル以外は A の固有ベクトルである）．ゼロベクトルを W_α の元に含めたのは，いかに述べる定理が成り立つためである．

定理 3.1.4 (固有空間) うえで定義した W_α は，$\mathbb{R}^n$ の部分空間である．さらに，W_α の元に行列 A をかけた結果も W_α の元になっている．

上の定理の前半の性質から，W_α を「A の，固有値 α に対する 固有空間」と言う．また，定理の後半の性質（A をかけても部分空間に入ったまま）を「W_α は A の 不変部分空間 である」と言う．

（注）行列 A の固有値を α とすると，行列 A によって W_α の次元が１より大きいことも往々にして起こる．一番簡単な例は $A = I_n$（単位行列）の場合だ．この場合，固有値は１だけだが，固有ベクトルとしては ゼロでない任意のベクトルがとれた．このような事情から，問題としても「行列 A の固有値 α に対応する固有ベクトルは何？」という訊き方は，答え方が無限とおりありうるので，あまり格好よくない．そこでより簡単に「A の，固有値 α に対する 固有空間 は何？」と訊くことが多い．このように訊かれたら，上の W_α を答えれば良い.（また，W_α の基底を一組，答えても良い.）A が単位行列の場合，その固有空間は $\mathbb{R}^n$ 全体だ．

固有空間が不変部分空間になっていること，が以下で見る「行列の対角化」に本質的な役割を果たす．

（少し概念，書き方などについての補足）レポートの採点をしていて,「固有ベクトル」「固有空間」「固有空間の基底」の３つがかなりごちゃ混ぜになってると見受けられたので，注意しておきます.

１．まず，固有ベクトルとは，定義通り，$A\boldsymbol{x} = \alpha\boldsymbol{x}$ となる，**ゼロでない**ベクトルです．だから，正しい解答は，例えば

$$A \text{ の固有値 } 1 \text{ に対応する固有ベクトルは} \qquad k\begin{bmatrix}1\\0\\1\end{bmatrix} \qquad \text{ただし } k \neq 0 \tag{3.1.10}$$

とか,

$$B \text{ の固有値 } 2 \text{ に対応する固有ベクトルは} \qquad k\begin{bmatrix}1\\0\\1\end{bmatrix} + \ell\begin{bmatrix}1\\1\\0\end{bmatrix} \qquad \text{ただし } (k,\ell) \neq (0,0) \tag{3.1.11}$$

などとなります.

２．次に,（固有値 α に対する）固有空間とは，α に対する固有ベクトルの全体に，ゼロベクトルを追加したもの，です．だから，上の例なら

$$\text{の固有値 } 1 \text{ に対する固有空間は} \qquad W_1 = \left\{ k\begin{bmatrix}1\\0\\1\end{bmatrix} \;\middle|\; k \in \mathbb{C} \right\} \tag{3.1.12}$$

とか

$$B \text{ の固有値 } 2 \text{ に対する固有空間は} \qquad W_2 = \left\{ k\begin{bmatrix}1\\0\\1\end{bmatrix} + \ell\begin{bmatrix}1\\1\\0\end{bmatrix} \;\middle|\; k,\ell \in \mathbb{C} \right\} \tag{3.1.13}$$

となります.「k,ℓ がゼロでない」条件はないことに再度，注意.

３．最後に，固有ベクトルの基底とは，上の固有空間を張る，一次独立なベクトルのことですから，

$$\text{の固有値 } 1 \text{ に対する固有空間 } W_1 \text{ の基底は} \qquad \left\langle \begin{bmatrix}1\\0\\1\end{bmatrix} \right\rangle \tag{3.1.14}$$

とか

$$B \text{ の固有値 } 2 \text{ に対する固有空間 } W_2 \text{ の基底は} \qquad \left\langle \begin{bmatrix}1\\0\\1\end{bmatrix}, \begin{bmatrix}1\\1\\0\end{bmatrix} \right\rangle \tag{3.1.15}$$

となります.

——— では，本題に戻る ———

固有値と固有ベクトルの求め方については前回で話は終わっているが，今日はもう少し，$\det(\alpha I_n - A) = 0$ の解の重複度との関係を見ておきたい.

$\det(\alpha I_n - A) = 0$ は α の n 次式であるので，複素数の範囲では重複度も含めて n 個の解を持つ．これは

$$\det(\alpha I_n - A) = (\alpha - \alpha_1)^{n_1} (\alpha - \alpha_2)^{n_2} \ldots (\alpha - \alpha_k)^{n_k} \tag{3.1.16}$$

と因数分解されることを意味している（k は 1 以上 n 以下の整数，$n_1 + n_2 + \ldots + n_k = n$）．このとき，以下が成り立つ.

定理 3.1.5 (固有ベクトルについての補足) $n \times n$ 行列 A の固有値と固有ベクトルについては，以下の性質がなりたつ．

1. $\det(\alpha I_n - A) = 0$ の解のそれぞれに対して，少なくとも一つの固有ベクトルが存在する．すなわち，$\det(\alpha I_n - A) = 0$ の解は定義通りの意味で固有値である．
2. 異なる固有値に属する固有ベクトルは一次独立である．
3. ある固有値 α_i の重複度を n_i とするとき，α_i の固有空間の次元は n_i 以下である．

証明：

1. $\det(\alpha I_n - A) = 0$ の解の一つを α_1 とすると，行列 $\alpha_1 I_n - A$ は正則ではない．よって，$(\alpha_1 I_n - A)\boldsymbol{x} = \mathbf{0}$ のゼロベクトルでない解 $\boldsymbol{x}$ が存在する．（これは前回もやった．）

2. $\alpha_1, \alpha_2, \ldots, \alpha_\ell$ を異なる固有値，対応する固有ベクトルを $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_\ell$ としよう．

$$k_1\boldsymbol{x}_1 + k_2\boldsymbol{x}_2 + \ldots + k_\ell\boldsymbol{x}_\ell = \mathbf{0} \tag{3.1.17}$$

を解いたら $k_i \equiv 0 \quad (i = 1, 2, \ldots, \ell)$ となることを言えばよい．そのためには (3.1.17) の両辺に A を何回もかけるのだ．$A\boldsymbol{x}_i = \alpha\boldsymbol{x}_i$ を使うと，A を一回かけて

$$k_1\alpha_1\boldsymbol{x}_1 + k_2\alpha_2\boldsymbol{x}_2 + \ldots + k_\ell\alpha_\ell\boldsymbol{x}_\ell = \mathbf{0} \tag{3.1.18}$$

となる．２回，３回，とかけていくと，結局

$$k_1\alpha_1^m\boldsymbol{x}_1 + k_2\alpha_2^m\boldsymbol{x}_2 + \ldots + k_\ell\alpha_\ell^m\boldsymbol{x}_\ell = \mathbf{0} \tag{3.1.19}$$

が任意の非負の整数 m について成り立つことになる．これらの連立方程式の解は $k_1 = k_2 = \ldots = 0$ しかないことは，やってみればわかる．（正直,「やってみる」のはちと大変で，ℓ についての数学的帰納法を使うのが一番簡単だろう．）

3. この証明は「行列の三角化」を使うと簡単なので，後回しにする．（ここで何とか証明してみようと少し考えたのだが，結局は「行列の三角化」の証明を部分的に行うようなものしか思いつかなかった．） □

3.2 行列の対角化（まずは計算操作として）

前節までで，行列の固有値と固有ベクトルに関する基本的な事柄を学び，それらを計算できるようになった（はずである）．ここではその続きとして,「行列の対角化」を取り扱う．大して難しい概念ではないが，取っつきやすさを考えて２段階で行う．まずは，以下の単純な計算問題を考えよう．

質問 1： $n \times n$ 行列 A が与えられたとき，正則な $n \times n$ 行列 P を探してきて，

$$B = P^{-1}AP \quad \text{が対角行列に} \tag{3.2.1}$$

なるようにせよ．（A に対してこのような P, B を求めることを「A を対角化する」と言う．）

すぐ後で見るように，このような行列 P が存在しない場合もあるので，上の問いはあくまで「可能ならばそのような P を求めよ」と解釈すべきものだ．この問いの背後には「線型写像の表現行列が，基底の変換に際してどう振る舞うか」が隠れているが，それは後で考える事にし，まずは計算問題として上の問いを扱おう．

上の問いの動機付け：

計算問題としてこのような問いが出るのは，A^{100} などを計算したい場合である．もし，A が 10×10 の行列で，成分にゼロのものがないとすると，A^2 を計算するだけで大変だ．ましてや，A^{100} などはほとんど不可能．

ところが，上のような P, B があれば，A^{100} も簡単に計算できる．つまり，$A = PBP^{-1}$ であるから，

$$A^2 = (PBP^{-1})(PBP^{-1}) = PB^2P^{-1}, \quad A^3 = (PB^2P^{-1})(PBP^{-1}) = PB^3P^{-1}, \quad A^\ell = PB^\ell P^{-1} \tag{3.2.2}$$

が成り立つ（ℓ は正の整数）．ここで B が対角行列なら，B^ℓ も対角行列で，すぐに計算できる．つまり

$$(B)_{ij} = b_i\,\delta_{ij} \quad \text{なら，} \quad (B^\ell)_{ij} = (b_i)^\ell\,\delta_{ij} \tag{3.2.3}$$

なのだ（対角成分がそれぞれ ℓ 乗されるだけ）．だから，(3.2.2) の計算も，対角行列である B^ℓ の両側から P と P^{-1} をかければよいので，何乗であっても計算できる．

と言うわけで，A^{100} を計算したいような場合，問い 1 の P, B が見つかるか否かでは大変な違いになる．この意味で，行列の対角化は実用上も非常に大事なのである．

P, B の見つけ方

さて，与えられた A に対して，どのようにして P, B を見つけたらよいか，考えよう．天下りだが答えを言ってしまうと，以下のようになる．

定理 3.2.1 (対角化その１) 与えられた $n \times n$ 行列 A が対角化できる必要十分条件は，A が丁度 n 本の一次独立な固有ベクトルを持つことである．このとき，A の独立な固有ベクトルを $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$ とすると（対応する固有値は $\alpha_1, \alpha_2, \ldots, \alpha_n$），(3.2.1) の行列 P, B は以下のようになる：

$$P = (\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n), \qquad B \text{ は対角成分が } \alpha_1, \alpha_2, \ldots, \alpha_n \text{ である対角行列} \tag{3.2.4}$$

証明：

十分条件（n 本の独立な固有ベクトルを持つならば対角化可能）は単なる計算だから，ここから始める．

$$A\,\boldsymbol{v}_j = \alpha_j\,\boldsymbol{v}_j \qquad (j = 1, 2, \ldots, n) \tag{3.2.5}$$

である．定理通りに P を作って，（１）P は正則であること，（２）$PB = AP$ が成り立つこと，を示せばよい．

（１）P の各列は n 個の独立な縦ベクトルからできている．このような行列が正則であるのは，春学期に既に見た（n 本独立 $\Longleftrightarrow$ 階数が n $\Longleftrightarrow$ 正則）．

（２）$PB = AP$ について：単なる計算だ．P, B の定義 (3.2.4) から，

$$AP = (A\boldsymbol{v}_1, A\boldsymbol{v}_2, \ldots, A\boldsymbol{v}_n) = (\alpha_1\boldsymbol{v}_1, \alpha_2\boldsymbol{v}_2, \ldots, \alpha_n\boldsymbol{v}_n) = PB \tag{3.2.6}$$

が成り立って，両者は等しい．よって十分条件であることは証明された．

では必要条件（対角化可能なら n 本の独立な固有ベクトル）の証明．

正則行列 P と対角行列 B があって，$PB = AP$ だったとしよう．この両辺の j 列目は，行列のかけ算の定義（および，B が対角行列であること）を用いて

$$A\,\boldsymbol{v}_j = \alpha_j\,\boldsymbol{v}_j \tag{3.2.7}$$

になっている．これは $\boldsymbol{v}_j \neq \boldsymbol{0}$ ならば，$\boldsymbol{v}_j$ が A の固有ベクトルであることを主張している．ところで，P は正則行列だから，各列 $\boldsymbol{v}_j$ は独立であり，特に，ゼロベクトルではあり得ない．従って，$j = 1, 2, \ldots, n$ に関して，$\boldsymbol{v}_j$ は A の固有ベクトルであり，かつこれらは独立だ．つまり，A は独立な固有ベクトルを n 本もつ．これで必要条件も証明された． □

系 3.2.2

(i) 行列 A が対角化できる必要十分条件は，A の固有値を α_i，その重複度を n_i，対応する固有空間を W_i と書くときに（$i = 1, 2, \ldots, k$），

$$\dim(W_i) = n_i \qquad (i = 1, 2, \ldots, k) \tag{3.2.8}$$

が成り立つことである．

(ii) $n \times n$ 行列 A が n 個の異なる固有値を持てば，A は対角化可能である（対角化可能の十分条件；決して必要条件ではないことに十分に注意のこと）．

証明：

(i) $n \times n$ 行列 A の対角化可能の必要十分条件は A が n 本の一次独立な固有ベクトルを持つことであった（定理 3.2.1）．しかし，定理 3.1.5 の 3 から，各固有値 α_i は高々n_i 個の独立な固有ベクトルしか持てない．$\sum_i n_i = n$ であるから，独立な固有ベクトルが n 本あるためには，α_i が n_i 本の独立な固有ベクトルを持つことが必要である．

逆に，α_i が n_i 本の独立な固有ベクトルを持てば，これらを集めてくればちょうど $\sum_i n_i = n$ 本の独立な固有ベクトルがあることになって，十分でもある（ここで定理 3.1.5 の 2 で示した，異なる固有値に属する固有ベクトルは独立であることを用いた）．

(ii) 定理 3.1.5 の 1 から，A は n 本の固有ベクトルを持つ．しかし，定理 3.1.5 の 2 によって，この n 本は独立である．従って，定理 3.2.1 の条件を満たすので，A は対角化可能． □

上の定理や系は，はなはだ不完全である．定理の証明を見ればわかるように，定理の主張自身がほとんど「アタリマエ」な感じである．その上，どんな行列に対して「n 本の独立な固有ベクトルを持つ」のか，その判定条件が与えられていない．従って現時点では各行列について固有ベクトルをすべて求める以外，判定手段がない．これについてはもっと良い判定条件（ただし，十分条件）を，今学期の最後の方で与えるであろう．

3.3 行列の三角化

今まで行列が対角化できる必要十分条件を見てきたが，世の中には対角化できない行列も多数，存在する．例えば $\begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$．このような行列が対角化できないことは今までの必要十分条件に照らせばわかる（十分な数の独立な固有ベクトルがない！）．では，対角化できない行列は どのくらいまで対角に近く できるのだろうか？より正確には

対角化できない $n \times n$ 行列 A に対して正則行列 P を探し，$P^{-1}AP$ をできるだけ対角形に近くせよ

という問題を考えたい．勿論「できるだけ近く」と言うのは主観的な言葉であるが，解答を見れば納得してもらえると思う．この問いに対する完全な解答（Jordan 標準型）は次小節で与えるが，それはなかなか大変で，すべてを説明することはできないだろう．この小節ではそれより簡単，しかし不完全な解答を与える．

定理 3.3.1 (行列の三角化；教科書の定理 7.3.1) 任意の $n \times n$ 行列 A は,「三角化」できる．つまり適当に正則行列 P を選んで $P^{-1}AP$ を上半三角行列にできる．

言うまでもないが，対角行列は上半三角行列の一種であるから，上の定理は対角化可能な場合を（ショウモナイ形で）カバーしている．

この定理の証明は n についての帰納法で行うのが普通だが，なかなかややこしいので，講義では省略する（教科書の．ここではむしろ，この定理を認めて何が言えるか，その副産物に注目したい．

まず，$P^{-1}AP = B$ が上半三角の場合，その対角線上には A の固有値がその重複度の回数だけ出ていることに注意しよう．なぜなら：$\det(\alpha I_n - A) = \det(\alpha I_n - B)$ なので，$\det(\alpha I_n - A) = 0$ と $\det(\alpha I_n - B) = 0$ は同値であるから．(上半三角行列の行列式はその対角成分の積だったことを思い出そう）．

これを使って，定理 3.1.5 の 3 の証明ができる．α_i を重複度 n_i の固有値だとし，固有空間を求めるために

$$\bigl(\alpha_i I_n - A\bigr)\boldsymbol{x} = \boldsymbol{0} \tag{3.3.1}$$

を解いてみる．このまま解くのは難しいから，A を上半三角にする行列 P を使って，$\boldsymbol{y} = P^{-1}\boldsymbol{x}$ と変数変換してやると，上のは

$$\bigl(\alpha_i I_n - B\bigr)\,\boldsymbol{y} = \boldsymbol{0} \tag{3.3.2}$$

と同値になる．このような $\boldsymbol{y}$ の作る空間の次元を知りたいわけだ．

前期にやったことから，この次元は $n - \operatorname{rank}(\alpha_i I_n - B)$ であるとわかるので，行列 $\alpha_i I_n - B$ の階数を求めたい．さて，$\alpha_i I_n - B$ の対角線上には丁度 n_i 個だけのゼロが並んでおり，かつこの行列が上半三角であるから，この行列の階数は少なくとも $(n - n_i)$ だけはある．(基本変形してゼロにしていこうと思っても対角線上にある $(n - n_i)$ 個のゼロでない数を消すことはできないから.）従って，

$$\dim(W_i) = n - \operatorname{rank}(\alpha_i I_n - B) \le n_i \tag{3.3.3}$$

となる． □

さて，x の多項式 $f(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_m x^m$ と $n \times n$ 正方行列 A が与えられたとき，x のところに A を代入して

$$f(A) = a_0 I_n + a_1 A + a_2 A^2 + \ldots + a_m A^m$$

を定義することができる（I_n は $n \times n$ の単位行列）．上の三角化の定理を使って，以下のような性質を証明できる．これらは主に「理論的」な興味のものだが，将来役に立つかもしれない．

定理 3.3.2 (フロベニウスの定理) $n \times n$ 行列 A の固有値の一つが λ，対応する固有ベクトルが $\boldsymbol{p}$ のとき，$f(A)$ の固有値の一つは $f(\lambda)$，対応する固有ベクトルは $\boldsymbol{p}$ となる．

定理 3.3.3 (Cayley-Hamilton の定理；教科書の定理 7.3.2) $n \times n$ 行列 A の固有多項式を $\phi_A(\lambda) = \det(A - \lambda I_n)$ と書くと，$\phi_A(A) = 0$ である（右辺は $n \times n$ ゼロ行列）．

3.4 Jordan の標準型

$P^{-1}AP$ についての最後の小節である．前節での問いに対する完全な答えを与えるが，まず言葉の準備をする．

定義 3.4.1 (Jordan 細胞と Jordan 行列) 以下の形の $n\times n$ 行列を $J_n(\lambda)$ と書いて，(固有値 λ を持つ) n-次 Jordan 細胞とよぶ．見てのとおり，対角線上には λ，その一つ上には 1，残りはすべてゼロ，という形である．

$$J_n(\lambda)=\begin{pmatrix}\lambda & 1 & 0 & 0 & 0 & \cdots & 0 & 0\\ 0 & \lambda & 1 & 0 & 0 & \cdots & 0 & 0\\ 0 & 0 & \lambda & 1 & 0 & \cdots & 0 & 0\\ \cdot & \cdot & 0 & \lambda & 1 & 0 & \cdot & \cdot\\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot\\ \cdot & \cdot & \cdot & \cdot & 0 & \lambda & 1 & 0\\ 0 & 0 & 0 & 0 & \cdots & 0 & \lambda & 1\\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & \lambda\end{pmatrix} \tag{3.4.1}$$

対角線上にいろいろな次元と固有値をもつ Jordan 細胞が並んだ形の行列，つまり，

$$\begin{pmatrix}J_{n_1}(\lambda_1) & & & & & \\ & J_{n_2}(\lambda_2) & & & & \\ & & J_{n_3}(\lambda_3) & & & \\ & & & \cdot & & \\ & & & & \cdot & \\ & & & & & J_{n_k}(\lambda_k)\end{pmatrix} \tag{3.4.2}$$

の形の行列を Jordan 行列という（行列中，何も書いていないところは全部ゼロ）．

この定義の下で以下が成り立つ：

定理 3.4.2 (Jordan 標準形への変換；教科書の定理 7.4.3) 任意の $n\times n$ 行列 A は適当に正則行列 P を選ぶことにより，$P^{-1}AP$ を Jordan 行列の形にできる．結果として出てくる Jordan 行列は Jordan 細胞の並べ方を除いて一意に決まる．

つまり，すべての行列は Jordan 行列の形にまでは持っていけるわけだ．この定理の証明はなかなか大変だし，時間もないので，省略する．将来，ジョルダンの標準形という言葉が出てきたときに戸惑わないように，定理だけ紹介した．

4 内積空間（計量線型空間）

教科書の対応部分は８章である．教科書では実線型空間と複素線型空間を別々にあつかっているが，特に分けてやる必要もないだろうからまとめてやることにする．

4.1 内積の定義

天下りであるが，定義を与える．

定義 4.1.1 (内積 I) $\mathbb{C}^n$ のベクトル $\boldsymbol{x}, \boldsymbol{y}$ の**内積** $\boldsymbol{x} \cdot \boldsymbol{y} = (\boldsymbol{x}, \boldsymbol{y})$ を

$$(\boldsymbol{x}, \boldsymbol{y}) = \sum_{j=1}^{n} x_j \overline{y_j} \tag{4.1.1}$$

で定義する．ここで $\boldsymbol{x}$ の第 j 成分を x_j と書いた．また $\overline{y_j}$ は y_j の複素共役を表す．

上の定義での複素共役は，y_j が実数のときにはもちろん，必要ない．

（参考）上では具体的な内積の定義を与えたが，実は内積はもっと一般に与えることもできる．すなわち，与えられたベクトル空間に対して，異なる内積をいろいろ定義できるのである.（これらは目的に応じて使い分ければよいが，大抵は考えている問題によって，自然な内積の定義が決まってくる.）ともかく，その「一般の内積」の定義を書いておくと：

定義 4.1.2 (内積 II) 複素ベクトル空間 X が与えられたとき，その**内積**とは，以下を満たすような関数（２つのベクトル $\boldsymbol{x}, \boldsymbol{y}$ に対して複素数 $(\boldsymbol{x}, \boldsymbol{y})$ が決まるもの）である（$\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}$ は X の任意のベクトル）．

1. $(\boldsymbol{x}, \boldsymbol{y}) = \overline{(\boldsymbol{y}, \boldsymbol{x})}$ 　ここで $\overline{}$ は複素共役
2. $(\boldsymbol{x} + \boldsymbol{y}, \boldsymbol{z}) = (\boldsymbol{x}, \boldsymbol{z}) + (\boldsymbol{y}, \boldsymbol{z})$
3. $(k\boldsymbol{x}, \boldsymbol{y}) = k(\boldsymbol{x}, \boldsymbol{y})$ 　ここで k は任意の複素数
4. $(\boldsymbol{x}, \boldsymbol{x}) \geq 0$ であり，等号は $\boldsymbol{x} = 0$ の時に限る

定義 4.1.1 の内積が定義 4.1.2 の性質をすべて満たしていることはすぐに確かめられるだろう．定義 4.1.1 とは異なる内積（定義 4.1.2 の意味で）の簡単な例として，

$$(\boldsymbol{x}, \boldsymbol{y}) = x_1 \overline{y_1} + 2 \sum_{j=2}^{n} x_j \overline{y_j} \tag{4.1.2}$$

を挙げておく（第一成分とそれ以外で重みが異なる）．**なお，このように内積の定義はいろいろありうるが，この講義では以降，内積とは定義 4.1.1 のものに限定して考える．ただし，以下に展開する議論のほとんどは一般の内積（定義 4.1.2）に対しても成り立つ．興味のある人は各自，確かめてみてほしい．**
(参考終わり)

もう少し定義を：

定義 4.1.3 (ノルム，直交) 複素ベクトル空間 X とその上の内積が与えられたとき，

$$\|\boldsymbol{x}\| \equiv \sqrt{(\boldsymbol{x}, \boldsymbol{x})} \tag{4.1.3}$$

をベクトル $\boldsymbol{x}$ の長さ（**ノルム**）という．また，２つのベクトルの内積がゼロ，つまり

$$(\boldsymbol{x}, \boldsymbol{y}) = 0 \tag{4.1.4}$$

の時，**$\boldsymbol{x}$ と $\boldsymbol{y}$ は直交する**という．

ベクトルの内積やノルムは以下の性質を満たす.

定理 4.1.4 (内積，ノルムの性質) ベクトルの内積とノルムは以下を満たす（$k \in \mathbb{C}$，$\boldsymbol{x}, \boldsymbol{y} \in X$）.

- $\|k\,\boldsymbol{x}\| = |k|\,\|\boldsymbol{x}\|$
- $\|\boldsymbol{x}+\boldsymbol{y}\|^2 + \|\boldsymbol{x}-\boldsymbol{y}\|^2 = 2\bigl(\|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2\bigr)$
- $\bigl|(\boldsymbol{x}, \boldsymbol{y})\bigr| \le \|\boldsymbol{x}\|\,\|\boldsymbol{y}\|$ （Cauchy-Schwarz の不等式）
- $\|\boldsymbol{x}+\boldsymbol{y}\| \le \|\boldsymbol{x}\| + \|\boldsymbol{y}\|$ （三角不等式）

（角度としての内積）Schwartz の不等式から，$-1 \le \dfrac{(\boldsymbol{x},\boldsymbol{y})}{\|\boldsymbol{x}\|\,\|\boldsymbol{y}\|} \le 1$ が得られるので,

$$\cos\theta = \frac{(\boldsymbol{x},\boldsymbol{y})}{\|\boldsymbol{x}\|\,\|\boldsymbol{y}\|} \tag{4.1.5}$$

なる θ を一つ決めることができる（$0 \le \theta \le \pi$）．この θ をベクトル $\boldsymbol{x}, \boldsymbol{y}$ のなす角度と解釈すると，内積はベクトルの間の角度を与えるものと言える．ここで唐突に角度が出てきたと思う人は，定義 4.1.1 の内積の定義を 2 次元のベクトルに対して書いてみて，それが丁度 $\|\boldsymbol{x}\|\,\|\boldsymbol{y}\|\cos\theta$ と同じであることを確かめてみると，違和感が少なくなるであろう．$\mathbb{R}^3$ のベクトルに対しては，このような事情は今年の始めにやったね.

（参考）上では内積に基づいた「ノルム」を定義し，その性質の一例として定理 4.1.4 を示した．しかし，内積の定義されていない空間でもノルムを定義することがあり，この場合は上の定理 4.1.4 の性質（の一部）を満たすものをノルムとする．ノルムの一般の定義は以下のようなものである.

定義 4.1.5 (ノルムの一般の定義（参考までに）) 複素ベクトル空間 X が与えられたとき，その**ノルム**とは，以下を満たすような関数（1 つのベクトル $\boldsymbol{x}$ に対して非負の実数値 $\|\boldsymbol{x}\|$ が決まるもの）である（$\boldsymbol{x}, \boldsymbol{y}$ は X の任意のベクトル）.

1. $\|k\boldsymbol{x}\| = k\|\boldsymbol{x}\|$ ここで k は任意の複素数
2. $\|\boldsymbol{x}+\boldsymbol{y}\| \le \|\boldsymbol{x}\| + \|\boldsymbol{y}\|$ （三角不等式）
3. $\|\boldsymbol{x}\| \ge 0$ かつ，$\|\boldsymbol{x}\| = 0$ は $\boldsymbol{x} = 0$ と同値

(参考終わり)

4.2 正規直交基底（内積の効用 I）

正規直交基底とその効用について，簡単に触れる．ここでは複素ベクトル空間とその上の内積が与えられたものとしてすすむ.

定義 4.2.1 (正規直交系，正規直交基底)
(i) n 次元複素ベクトル空間 X の基底 $\langle \boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n \rangle$ が,「互いに直交して，かつ長さが 1」つまり

$$(\boldsymbol{u}_i, \boldsymbol{u}_j) = \begin{cases} 1 & (i = j) \\ 0 & (i \ne j) \end{cases} \tag{4.2.1}$$

を満たすとき，**正規直交基底**であるという.
(ii) より一般に，n 次元複素ベクトル空間 X のベクトルの組 $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_s\}$（$s \le n$）が

$$(\boldsymbol{u}_i, \boldsymbol{u}_j) = \begin{cases} 1 & (i = j) \\ 0 & (i \ne j) \end{cases} \tag{4.2.2}$$

を満たすとき，**正規直交系**であるという.

（例）$\mathbb{R}^n$ における基本ベクトル $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n$ は正規直交基底をなしている．

正規直交基底の良いところは，勝手なベクトル $\boldsymbol{x}$ をこの正規直交基底で**展開した係数が簡単に計算できる**ことである．少し思い出してみると，基底 $\langle \boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n \rangle$ で $\boldsymbol{x}$ を展開するとは

$$\boldsymbol{x} = x_1\boldsymbol{u}_1 + x_2\boldsymbol{u}_2 + \cdots + x_n\boldsymbol{u}_n \tag{4.2.3}$$

と書けるように係数 $x_1, x_2, \ldots, x_n$ を決めることだった．そして，今までこの問題に答えるには，上の方程式を連立方程式の形に書いて，一生懸命解くしかなかったのだ.（これは一般に非常に大変．試験の時など，４次元，かつ僕が係数を計算し易いようにしたものでも間違ったでしょう？）

ところが，$\underline{\langle \boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n \rangle \text{ が正規直交基底の場合に限り}}$ 大変簡単な計算法があるのだ：

定理 4.2.2 (正規直交基底での展開) 正規直交基底による展開 (4.2.3) の係数は以下で与えられる：

$$x_j = (\boldsymbol{x}, \boldsymbol{u}_j) \qquad (j = 1, 2, \ldots, n) \tag{4.2.4}$$

（証明）$\boldsymbol{x} = x_1\boldsymbol{u}_1 + x_2\boldsymbol{u}_2 + \ldots + x_n\boldsymbol{u}_n$ と書いておいて，x_j を求めてやろう（$j = 1, 2, \ldots, n$）．そのために，この表式と $\boldsymbol{u}_j$ の内積をとってやると，

$$(\boldsymbol{x}, \boldsymbol{u}_j) = \big(x_1\boldsymbol{u}_1 + x_2\boldsymbol{u}_2 + \ldots + x_n\boldsymbol{u}_n, \boldsymbol{u}_j\big) = x_1(\boldsymbol{u}_1, \boldsymbol{u}_j) + x_2(\boldsymbol{u}_2, \boldsymbol{u}_j) + \ldots + x_n(\boldsymbol{u}_n, \boldsymbol{u}_j) = x_j \tag{4.2.5}$$

となる —— $(\boldsymbol{u}_i, \boldsymbol{u}_j) = \delta_{ij}$ を思い出せ． □

これがどれほど簡単かは実際に 10 次元くらいの空間でやってみればわかる．従来の方法なら 10 連立方程式を解くしかない．しかし，正規直交基底ならちょこちょこと内積を計算するだけである．内積を導入する，一つの大きな理由はここにある．

ただし，この方法は正規直交基底の場合にのみ使える，ことはしつこく注意しておこう．

4.3 Gram-Schmidt の直交化法

このように便利な正規直交基底であるが，現実に得られる基底は正規直交基底になっていないことも多い．そのような場合，今持っている基底から正規直交基底を作る方法が存在し，Gram-Schmidt の直交化法とよばれている（教科書の p.178, p.183）．

要点だけ述べると，以下のようになる．

定理 4.3.1 (Gram-Schmidt の直交化法) $\{\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_s\}$ を，$\mathbb{C}^n$ の，一次独立な s 本のベクトルとする．このとき，

(i) 以下の手順によって，正規直交系 $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_s$ を作ることができる．

1. $\boldsymbol{u}_1 := \dfrac{\boldsymbol{a}_1}{\|\boldsymbol{a}_1\|}$ と定める．
2. $\boldsymbol{w}_2 := \boldsymbol{a}_2 - (\boldsymbol{a}_2, \boldsymbol{u}_1)\boldsymbol{u}_1$ を作り，$\boldsymbol{u}_2 := \dfrac{\boldsymbol{w}_2}{\|\boldsymbol{w}_2\|}$ と定める．
3. $\boldsymbol{w}_3 := \boldsymbol{a}_3 - (\boldsymbol{a}_3, \boldsymbol{u}_1)\boldsymbol{u}_1 - (\boldsymbol{a}_3, \boldsymbol{u}_2)\boldsymbol{u}_2$ を作り，$\boldsymbol{u}_3 := \dfrac{\boldsymbol{w}_3}{\|\boldsymbol{w}_3\|}$ と定める．
4. 以下，同様に進める．具体的には，$\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_{k-1}$ までを決めた時（$k \geq 4$），$\boldsymbol{w}_k := \boldsymbol{a}_k - \displaystyle\sum_{j=1}^{k-1} (\boldsymbol{a}_k, \boldsymbol{u}_j)\boldsymbol{u}_j$ を作り，$\boldsymbol{u}_k := \dfrac{\boldsymbol{w}_k}{\|\boldsymbol{w}_k\|}$ と定める．
5. 以上を $k \leq s$ のすべてに対して繰り返して作った $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_s\}$ が求める正規直交系である．

(ii) 更に，$\{\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_s\}$ の張る線型空間と，$\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_s\}$ の張る線型空間は同じである．

（注）上の定理で $s=n$ の場合，この手法により $\mathbb{C}^n$ の正規直交基底が作れる．

（略証）定理の (i) の部分については，このような $\{\boldsymbol{u}_1,\boldsymbol{u}_2,\ldots,\boldsymbol{u}_s\}$ を本当に作れること（つまり，上のプログラムを実行できること），および，出来た $\{\boldsymbol{u}_1,\boldsymbol{u}_2,\ldots,\boldsymbol{u}_s\}$ が正規直交系になっていること，を示す必要がある．

さて，このプログラムが実行できたとすれば，$\{\boldsymbol{u}_1,\boldsymbol{u}_2,\ldots,\boldsymbol{u}_s\}$ が正規直交系になっていることはほとんど自明だ．というのも，$\boldsymbol{u}_k:=\dfrac{\boldsymbol{w}_k}{\|\boldsymbol{w}_k\|}$ という定義から，$\|\boldsymbol{u}_k\|=1$ が成り立ち，$\boldsymbol{u}_k$ の長さは 1（正規性 O.K.）．

また，$\boldsymbol{w}_k:=\boldsymbol{a}_k-\sum_{j=1}^{k-1}(\boldsymbol{a}_k,\boldsymbol{u}_j)\boldsymbol{u}_j$ という定義から，$(\boldsymbol{w}_k,\boldsymbol{u}_j)=0$ がすべての $j<k$ に対して成り立つ．実際，$\boldsymbol{u}_1,\boldsymbol{u}_2,\ldots,\boldsymbol{u}_{k-1}$ が正規直交系なので，

$$(\boldsymbol{w}_k,\boldsymbol{u}_i)=(\boldsymbol{a}_k,\boldsymbol{u}_i)-\sum_{j=1}^{k-1}(\boldsymbol{a}_k,\boldsymbol{u}_j)\times(\boldsymbol{u}_j,\boldsymbol{u}_i)=(\boldsymbol{a}_k,\boldsymbol{u}_i)-(\boldsymbol{a}_k,\boldsymbol{u}_i)=0 \tag{4.3.1}$$

となるからである．$\boldsymbol{w}_k$ と $\boldsymbol{u}_k$ は比例しているので，$(\boldsymbol{u}_k,\boldsymbol{u}_j)=0$ がすべての $j<k$ に対して成り立つが，これがさらにすべての $k\le s$ で成り立つので，結局，直交性も O.K.

問題は上のプログラムが本当に実行できるか，である．プログラムが実行できないとすれば，それは分母の $\|\boldsymbol{w}_k\|$ がどこかの k でゼロになってしまう場合である．しかし，これは起こりえない．なぜなら，もし，どこかの k で $\|\boldsymbol{w}_k\|=0$ ということは，$\boldsymbol{w}_k=\boldsymbol{0}$，つまり

$$\boldsymbol{a}_k-\sum_{j=1}^{k-1}(\boldsymbol{a}_k,\boldsymbol{u}_j)\boldsymbol{u}_j=\boldsymbol{0} \qquad \text{すなわち} \quad \boldsymbol{a}_k=\sum_{j=1}^{k-1}(\boldsymbol{a}_k,\boldsymbol{u}_j)\boldsymbol{u}_j \tag{4.3.2}$$

を意味する．ところが，上の作り方から，$\boldsymbol{u}_j$ は $\boldsymbol{a}_1,\boldsymbol{a}_2,\ldots,\boldsymbol{a}_j$ の線型結合である．よって，(4.3.2) は，$\boldsymbol{a}_k$ が $\boldsymbol{a}_1,\boldsymbol{a}_2,\ldots,\boldsymbol{a}_{k-1}$ の線型結合であることを意味する．でもこれは $\{\boldsymbol{a}_1,\boldsymbol{a}_2,\ldots,\boldsymbol{a}_s\}$ が一次独立であったことに反する．

定理の (ii) の部分もほとんど自明だ．上で既に注意したように，$\boldsymbol{u}_j$ は $\boldsymbol{a}_1,\boldsymbol{a}_2,\ldots,\boldsymbol{a}_j$ の線型結合である．よって，$\boldsymbol{u}_1,\boldsymbol{u}_2,\ldots,\boldsymbol{u}_s$ の線型結合は，$\boldsymbol{a}_1,\boldsymbol{a}_2,\ldots,\boldsymbol{a}_j$ の線型結合でもある．よって，$\boldsymbol{u}_1,\boldsymbol{u}_2,\ldots,\boldsymbol{u}_s$ で張られる線型空間は，$\boldsymbol{a}_1,\boldsymbol{a}_2,\ldots,\boldsymbol{a}_j$ で張られる部分空間の部分集合（部分空間）のはずである．

ところが，上の作り方を逆にたどれば，$\boldsymbol{a}_j$ が $\boldsymbol{u}_1,\boldsymbol{u}_2,\ldots,\boldsymbol{u}_j$ の線型結合であることもわかる．つまり，上の議論での $\boldsymbol{a}$ と $\boldsymbol{u}$ の役割を入れ替えた議論が成立し，$\boldsymbol{a}_1,\boldsymbol{a}_2,\ldots,\boldsymbol{a}_j$ で張られる線型空間は，$\boldsymbol{u}_1,\boldsymbol{u}_2,\ldots,\boldsymbol{u}_s$ で張られる部分空間の部分集合（部分空間）のはずである．

お互いがお互いの部分空間になっていることが言えたので，両者は等しい． □

4.4 直交補空間

少し抽象的ではあるが，行列の対角化に関して，また将来進んだ話をするときにも重要なので，簡単に触れる．

定義 4.4.1 (直交補空間) W を $\mathbb{C}^n$ の部分空間とする．このとき，W のどの元とも直交するベクトルの全体

$$W^{\perp}:=\left\{\boldsymbol{x}\in\mathbb{C}^n \,\middle|\, \text{すべての } \boldsymbol{y}\in W \text{ に対して } (\boldsymbol{x},\boldsymbol{y})=0\right\} \tag{4.4.1}$$

を W の直交補空間 という．

（注）$W^{\perp}$ は実際に，$\mathbb{C}^n$ の部分空間になっている（各自で納得すること）．

定理 4.4.2 (直交補空間の性質) W を $\mathbb{C}^n$ の部分空間とすると，以下がなりたつ．
(i) $\dim W+\dim W^{\perp}=n$
(ii) $\left(W^{\perp}\right)^{\perp}=W$ つまり，$W^{\perp}$ の直交補空間は W である．

（証明）$W^{\perp}$ の基底を具体的に作ってしまうのが一番，簡単だ．

まず，W の基底を $\langle\boldsymbol{a}_1,\boldsymbol{a}_2,\ldots,\boldsymbol{a}_s\rangle$ とする．これに $(n-s)$ 本のベクトル $\boldsymbol{a}_{s+1},\boldsymbol{a}_{s+2},\ldots,\boldsymbol{a}_n$ を付け加えて $\mathbb{C}^n$ の基底を作ることができる．

さて，$\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n$ に Gram-Schmidt の直交化を適用して，$\mathbb{C}^n$ の基底 $\langle \boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n \rangle$ を作ろう．前節でやったことから，$\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_s$ の張る部分空間は $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_s$ の張る部分空間と同じである．つまり，$\langle \boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_s \rangle$ は W の基底になっている．

このことを利用して，$W^\perp$ の元がどのように書けるかを考えてみる．$W^\perp$ の任意の元 $\boldsymbol{x}$ は $\mathbb{C}^n$ の元でもあるから，

$$\boldsymbol{x} = c_1 \boldsymbol{u}_1 + c_2 \boldsymbol{u}_2 + \cdots + c_n \boldsymbol{u}_n \tag{4.4.2}$$

と線型結合で書けるはずだ．ところが，$\boldsymbol{x}$ は $\boldsymbol{u}_j$（$1 \le j \le s$）と直交しなければならない（なぜなら，$\boldsymbol{u}_j$ は W の元であって，$\boldsymbol{x}$ は W のすべての元と直交するから）．この条件は

$$0 = (\boldsymbol{x}, \boldsymbol{u}_j) = c_j \qquad (1 \le j \le s) \tag{4.4.3}$$

である．つまり，(4.4.2) は実は

$$\boldsymbol{x} = c_{s+1} \boldsymbol{u}_{s+1} + c_{s+2} \boldsymbol{u}_{s+2} + \cdots + c_n \boldsymbol{u}_n \tag{4.4.4}$$

と，$\boldsymbol{u}_{s+1}$ 以降の線型結合として書けるのだ．逆に，(4.4.4) の元は W の任意の元と直交する（各自でチェック）から，$W^\perp$ の基底の一つとして $\langle \boldsymbol{u}_{s+1}, \boldsymbol{u}_{s+2}, \ldots, \boldsymbol{u}_n \rangle$ がとれることがわかった．これは $(n-s)$ 個の独立なベクトルからなるので，$W^\perp$ の次元は $(n-s)$ とわかる —— (i) が証明できた．

また，$W^\perp$ の基底の一つとして $\langle \boldsymbol{u}_{s+1}, \boldsymbol{u}_{s+2}, \ldots, \boldsymbol{u}_n \rangle$ がとれる，というところから出発して $(W^\perp)^\perp$ の基底を構成すると，それが $\langle \boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_s \rangle$ になることは，上の議論を繰り返せばできる．でもこの基底は W の基底に他ならないから，$W = (W^\perp)^\perp$ が言える —— (ii) が証明された． □

4.5 直交行列とユニタリー行列

内積に関する章の締めくくりとして，２つの特殊な行列を定義し，それと内積の関係を考えておく．これは次の章で行列の対角化を求める際に，また登場するだろう．

定義 4.5.1 (転置行列とエルミート共役行列)　A を $n \times n$ 行列とする（成分は実数または複素数）
(i) 行列 A から新しい行列 ${}^tA$ を，その ij 成分が

$$\left({}^tA\right)_{ij} = A_{ji} \tag{4.5.1}$$

を満たすように作る．この行列 ${}^tA$ を A の**転置行列**という．
(ii) 行列 A から新しい行列 $A^\dagger$ を，その ij 成分が

$$\left(A^\dagger\right)_{ij} = \overline{A_{ji}} \tag{4.5.2}$$

を満たすように作る（$\overline{\quad}$ は複素共役）．この行列 $A^\dagger$ を A の**エルミート共役行列**という．

このような定義をする意味は以下の性質にある：

定理 4.5.2 (転置行列，エルミート共役行列と内積)
(i) n-次元実ベクトル空間 $\mathbb{R}^n$ において，定義 4.1.1 の内積を考える．このとき，任意の $n \times n$ 実行列 A（A の成分は実数）と $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ に対して

$$(A\boldsymbol{x}, \boldsymbol{y}) = (\boldsymbol{x}, {}^tA\boldsymbol{y}) \tag{4.5.3}$$

が成り立つ．
(ii) n-次元複素ベクトル空間 $\mathbb{C}^n$ において，定義 4.1.1 の内積を考える．このとき，任意の $n \times n$ 複素行列 A と $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{C}^n$ に対して

$$(A\boldsymbol{x}, \boldsymbol{y}) = (\boldsymbol{x}, A^\dagger\boldsymbol{y}) \tag{4.5.4}$$

が成り立つ．

つまり，${}^tA$ や $A^\dagger$ は内積 $(A\boldsymbol{x},\boldsymbol{y})$ の A を左から右に移す際に自然に現れるものなのだ．

さて，ここで以下の問いを考えてみよう．

（問）$\mathbb{R}^n$ において，ベクトル $\boldsymbol{x},\boldsymbol{y}$ に実行列 A をかけたもの $A\boldsymbol{x}, A\boldsymbol{y}$ を考える．すべての $\boldsymbol{x},\boldsymbol{y}\in\mathbb{R}^n$ に対して，$(\boldsymbol{x},\boldsymbol{y}) = (A\boldsymbol{x},A\boldsymbol{y})$ となるような A —— つまり，A で変換してもベクトルの内積が変わらないような A は何か？また，同様の問いを複素ベクトル空間 $\mathbb{C}^n$ と複素行列に対して考えた答えは何か？

（答え）実ベクトル空間の場合から考える．ノルムの定義と上の定理から，

$$(A\boldsymbol{x}, A\boldsymbol{y}) = (\boldsymbol{x}, {}^tA\,A\,\boldsymbol{y}) \tag{4.5.5}$$

であるが，これがすべての $\boldsymbol{x},\boldsymbol{y}$ に対して $(\boldsymbol{x},\boldsymbol{y})$ に等しくなければならない．つまりすべての $\boldsymbol{x},\boldsymbol{y}\in\mathbb{R}^n$ に対して $(\boldsymbol{x},B\boldsymbol{y}) = (\boldsymbol{x},\boldsymbol{y})$ となるような行列 B は何か？ということだ．この答えは $B = I_n$ しかないことは容易にわかる（$\boldsymbol{x}=\boldsymbol{e}_1, \boldsymbol{y}=\boldsymbol{e}_2,\ldots$）などといろいろ代入してみるとよい）．つまり，${}^tA\,A = I_n$ となる行列 A が答えになっているのだ．複素の場合は同様に，$AA^\dagger = I_n$ となる A が答えだとわかる．

以上に基づいて，以下の定義を行おう．

定義 4.5.3 (直交行列とユニタリー行列)　A を $n\times n$ 行列とする（成分は実数または複素数）

(i) 行列 A が

$${}^tA\,A = A\,{}^tA = I_n \tag{4.5.6}$$

を満たす時，つまり $A^{-1} = {}^tA$ であるとき，A は**直交行列**という．

(ii) 行列 A が

$$A^\dagger\,A = A\,A^\dagger = I_n \tag{4.5.7}$$

を満たす時，つまり $A^{-1} = A^\dagger$ であるとき，A は**ユニタリー行列**という．

上で見たことも踏まえて，これらの行列と内積の関係をまとめると，以下のようになる．

定理 4.5.4 (直交行列の性質)　A は $n\times n$ 実行列とする．以下の 4 条件は同値である，

(i) A は実直交行列である．

(ii) 任意の $\boldsymbol{x},\boldsymbol{y}\in\mathbb{R}^n$ に対して $(A\boldsymbol{x},A\boldsymbol{y}) = (\boldsymbol{x},\boldsymbol{y})$.

(iii) 任意の $\boldsymbol{x}\in\mathbb{R}^n$ に対して $\|A\boldsymbol{x}\| = \|\boldsymbol{x}\|$.

(iV) A の列ベクトル $\boldsymbol{a}_1,\boldsymbol{a}_2,\ldots,\boldsymbol{a}_n$ は $\mathbb{R}^n$ の正規直交基底をなす．

定理 4.5.5 (ユニタリー行列の性質)　A は $n\times n$ 複素行列とする．以下の 4 条件は同値である，

(i) A はエルミート行列である．

(ii) 任意の $\boldsymbol{x},\boldsymbol{y}\in\mathbb{C}^n$ に対して $(A\boldsymbol{x},A\boldsymbol{y}) = (\boldsymbol{x},\boldsymbol{y})$.

(iii) 任意の $\boldsymbol{x}\in\mathbb{C}^n$ に対して $\|A\boldsymbol{x}\| = \|\boldsymbol{x}\|$.

(iV) A の列ベクトル $\boldsymbol{a}_1,\boldsymbol{a}_2,\ldots,\boldsymbol{a}_n$ は $\mathbb{C}^n$ の正規直交基底をなす．

5 正規行列の対角化

5.1 エルミート行列の対角化

この節は物理（特に量子力学）との関係でも重要なので，物理の学生さんは心して学修するように．

転置行列，エルミート共役行列の定義は既にやった．これを用いて以下の定義を行う：

定義 5.1.1 (対称行列とエルミート行列) 正方行列 A が，

- ${}^tA = A$ を満たす時，A は**対称行列**（symmetric matrix）という．
- $A^\dagger = A$ を満たす時，A は**エルミート行列**（hermitian matrix）という．

A の成分が実数の場合，A が対称行列ならば A はエルミート行列でもある．以下では主にエルミート行列の性質を証明するが，同様の証明はすべて，実対称行列についても成り立つことを注意しておく．

さて，定理 4.5.2 を思い出すと，A と $A^\dagger$ については

$$(\boldsymbol{x}, A\boldsymbol{y}) = (A^\dagger\boldsymbol{x}, \boldsymbol{y}) \qquad (\boldsymbol{x}, \boldsymbol{y} \in \mathbb{C}^n) \tag{5.1.1}$$

が成り立つのだった．従って，エルミート行列とは，

$$(\boldsymbol{x}, A\boldsymbol{y}) = (A\boldsymbol{x}, \boldsymbol{y}) \qquad (\boldsymbol{x}, \boldsymbol{y} \in \mathbb{C}^n) \tag{5.1.2}$$

が成り立つ行列ということになる．この性質から，以下の驚くべき性質が導かれる：

定理 5.1.2 (エルミート行列の固有値) A をエルミート行列または実対称行列とする．このとき，
(i) A の固有値は**すべて実数**である．
(ii) A の異なる固有値に対する固有ベクトル同士は**直交**する．

（証明）内積を使うと簡単だ．固有値 α に対する A の固有ベクトル（の一つ）を $\boldsymbol{a}$ としよう：$A\boldsymbol{a} = \alpha\boldsymbol{a}$．この両辺を，それぞれ $\boldsymbol{a}$ と内積をとると，

$$(A\boldsymbol{a}, \boldsymbol{a}) = (\alpha\boldsymbol{a}, \boldsymbol{a}) = \alpha(\boldsymbol{a}, \boldsymbol{a}) = \alpha \|\boldsymbol{a}\|^2 \tag{5.1.3}$$

となる．ところが左辺は，A がエルミート行列なので，(5.1.2) から

$$(A\boldsymbol{a}, \boldsymbol{a}) = (\boldsymbol{a}, A\boldsymbol{a}) = (\boldsymbol{a}, \alpha\boldsymbol{a}) = \overline{\alpha}(\boldsymbol{a}, \boldsymbol{a}) = \overline{\alpha} \|\boldsymbol{a}\|^2 \tag{5.1.4}$$

に等しい．両辺を引き算して

$$(\alpha - \overline{\alpha}) \|\boldsymbol{a}\|^2 = 0 \tag{5.1.5}$$

を得るが，$\boldsymbol{a}$ が固有ベクトルなので，$\|\boldsymbol{a}\| > 0$ である．よって，$\alpha - \overline{\alpha} = 0$ であり，α は実数と結論できる．

次に，$\alpha \neq \beta$ なる A の２つの固有値を持ってきて，対応する固有ベクトルをそれぞれ $\boldsymbol{a}, \boldsymbol{b}$ とする：

$$A\boldsymbol{a} = \alpha\boldsymbol{a}, \qquad A\boldsymbol{b} = \beta\boldsymbol{b} \tag{5.1.6}$$

一つ目の式と $\boldsymbol{b}$ の内積をとると

$$(A\boldsymbol{a}, \boldsymbol{b}) = (\alpha\boldsymbol{a}, \boldsymbol{b}) = \alpha(\boldsymbol{a}, \boldsymbol{b}) \tag{5.1.7}$$

であるが，やはり (5.1.2) から，

$$(A\boldsymbol{a}, \boldsymbol{b}) = (\boldsymbol{a}, A\boldsymbol{b}) = (\boldsymbol{a}, \beta\boldsymbol{b}) = \bar{\beta}(\boldsymbol{a}, \boldsymbol{b}) = \beta(\boldsymbol{a}, \boldsymbol{b}) \tag{5.1.8}$$

が成り立つ（最後のところでは固有値 β が実数であることを用いた）．２つの式を引き算して，

$$(\alpha - \beta)\,(\boldsymbol{a}, \boldsymbol{b}) = 0 \tag{5.1.9}$$

が得られる．これは $\alpha \neq \beta$ なら $(\boldsymbol{a}, \boldsymbol{b}) = 0$，つまり $\boldsymbol{a}$ と $\boldsymbol{b}$ は直交することを意味する． □

さてさて，エルミート行列や実対称行列には，更に次のような非常に良い性質がある．その前に

- $U^{-1} = U^{\dagger}$ である行列をユニタリー行列という
- $P^{-1} = {}^{t}P$ である行列 P を直交行列という

ことを思い出しておこう．

定理 5.1.3 (エルミート行列，実対称行列は対角化可能)

- エルミート行列は対角化できる．しかも，対角化に使う行列 P を「ユニタリー行列」にとることができる．
- 実対称行列は対角化できる．しかも，対角化に使う行列 P を「直交行列」にとることができる．

（証明）この定理の証明は少し面倒である上に，理解してもそれほど視野が広がるとは思われない．よって，後の定理 5.2.2 の証明と併せて（時間があれば）黒板で紹介するにとどめる．一応の証明は教科書の p.191 にある． □

行列がエルミート行列かどうかは**その形だけを見れば判定できる**．この意味で，うえの定理は実用上，非常に大事である（ただし，あくまで対角化可能の十分条件であることには注意）．

重要：期末試験では「与えられたエルミート行列を対角化するユニタリー行列を求めよ」というような問題を必ず出題するから，確実にできるようになっておくこと．この問題が解けることは，物理の学生さんには必須の技能であるから，今，できるようになろう．

与えられた行列 A を対角化する行列 P を求める方法は，既に習った（A の独立な固有ベクトルを並べれば良かった）．ここで新しいのは，その行列をユニタリー行列に取る必要があることだ．この部分は「Gram-Schmidt の直交化」で行うことができる（詳しくは黒板で）．

5.2 正規行列の対角化

さて，前節ではエルミート行列というものを特に取り上げ，その性質を調べた．特に，エルミート行列は絶対に対角化できる（しかもユニタリー行列で対角化できる）ことを注意した．しかし，エルミート行列でなくても，ユニタリー行列で対角化できる行列はいろいろある．その点をこの節では見て行く．まず定義：

定義 5.2.1 (正規行列) $n \times n$ の行列 A が $A\,A^{\dagger} = A^{\dagger}\,A$ を満たす時，A は**正規行列**（normal matrix）という．

（例）エルミート行列，ユニタリー行列は共に正規行列である（各自，定義をチェックすること）．
すると，以下の定理がなりたつ：

定理 5.2.2 (ユニタリー行列で対角化できる必要十分条件) 正方行列 A について，以下の a, b は同値である．

a. A は正規行列である．
b. A はユニタリー行列 U を使って対角化できる．つまり，$U^{-1}AU$ が対角行列になるようなユニタリー行列 U が存在する．

（証明）この定理の証明は大変である上に，理解してもそれほど視野が広がるとは思われない．よって（時間があれば）黒板で紹介するにとどめる．一応の証明は教科書の p.194 にある． □

この定理もまた，行列が対角化できる十分条件を比較的簡単に与えてくれるので，重宝である．特に，**エルミート行列やユニタリー行列は対角化できる**ことが保証された．ただし，世の中には，「ユニタリー行列では対角化できないが，もっと一般の行列 P を用いれば対角化できる」行列も存在することは再度強調しておこう．

「ユニタリー行列では対角化できないが，もっと一般の行列 P を用いれば対角化できる」行列の例

$$A = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} \tag{5.2.1}$$

この行列は正規行列ではない（${}^t\!A\,A \neq A\,{}^t\!A$）が，

$$P = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \qquad \text{によって} \qquad P^{-1}AP = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \tag{5.2.2}$$

と対角化される．

正規行列に関する一番重要な性質は上の定理でつきているが，少し補足的な性質を述べておこう．

定理 5.2.3 (正規行列と内積) $n \times n$ 正方行列 A について，以下の a, b は同値である．

a. A は正規行列である．
b. 任意の n 成分複素縦ベクトル $\boldsymbol{x}$ に対して，$\|A\boldsymbol{x}\| = \|A^\dagger \boldsymbol{x}\|$ が成り立つ．

定理 5.2.4 (正規行列と固有値) 正規行列 A の固有値 α に対する固有ベクトルを $\boldsymbol{a}$ とする（$A\boldsymbol{a} = \alpha\boldsymbol{a}$）と，$A^\dagger \boldsymbol{x} = \bar{\alpha}\boldsymbol{x}$ がなりたつ．つまり，α が A の固有値なら，$\bar{\alpha}$ は $A^\dagger$ の固有値になる．

定理 5.2.5 (正規行列と固有値) 正規行列 A の相異なる固有値に対する固有ベクトルは，互いに直交する．