実対称行列の直交行列による対角化

新居俊作

2023 年 4 月 13 日

一般に対称行列は直交行列で対角化できるが、これを線形写像の表現行列の座標変換の下での変化と捉えることは線形代数の主要な概念が登場し、良い復習になる。そこで本講ではこの見方に従って、直交行列の固有値が互いに異なる場合に限って証明する。

1 線形写像

中学高、高等学校の数学で習った関数の中で最も簡単なものを考えるとそれは一次関数

$$f(x) = ax, \qquad x \in \mathbb{R} \tag{1.1}$$

(ただし $a \in \mathbb{R}$ は定数) だろう。

次に $\boldsymbol{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$ に対し $\mathbb{R}^n$ の他の点を対応させる規則 (このような規則を**写像**とよぶ) で最も簡単なものは何か考えてみる。

(一変数) 関数と同様に

$$f(\boldsymbol{x}) = a\boldsymbol{x} \tag{1.2}$$

(ただし $a \in \mathbb{R}$ は定数) は確かに非常に簡単だが、これは単にベクトルを a 倍するだけなのでさすがに単純すぎて特に分析するべきことはない。

次に簡単なのは、 $f(\boldsymbol{x}) = \begin{pmatrix} f_1(\boldsymbol{x}) \\ \vdots \\ f_n(\boldsymbol{x}) \end{pmatrix} \in \mathbb{R}^n$ の各成分 $f_i(\boldsymbol{x})$ が x_j の一次

式で表される場合、すなわち

$$\begin{cases} f_1(\boldsymbol{x}) = \displaystyle\sum_{k=1}^{n} a_{1,k} x_k \\ \qquad\vdots \\ f_n(\boldsymbol{x}) = \displaystyle\sum_{k=1}^{n} a_{n,k} x_k \end{cases} \tag{1.3}$$

(ただし $a_{i,j} \in \mathbb{R},\ i,j = 1,\ldots,n$ は定数) となっている場合だろう。このような写像を**線形写像**とよぶ。

行列を用いると (1.3) は次のように表される。

$$f(\boldsymbol{x}) = A\boldsymbol{x}, \qquad A = \begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & a_{n,n} \end{pmatrix} \tag{1.4}$$

線形写像は次の性質を持つ：

定理 1 (線型性)**.**

$c \in \mathbb{R},\ \boldsymbol{x},\ \boldsymbol{x}' \in \mathbb{R}^n$ とする。写像 $f(\boldsymbol{x})$ が (1.3) の形、つまり (1.4) の形をしているならば次が成り立つ：

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

(2) $f(\boldsymbol{x} + \boldsymbol{x}') = f(\boldsymbol{x}) + f(\boldsymbol{x}')$

これは (1.3) または (1.4) から簡単に確かめられる。重要なのは次が成り立つことである。

定理 2.

$\mathbb{R}^n$ の元に $\mathbb{R}^n$ の元を対応させる写像 f が線型性をもつ、すなわち定理 1 の (1)、(2) をみたすならば、f は (1.3) の形つまり (1.4) の形に表される。

証明.

先ず $\boldsymbol{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ \\ \\ \vdots \\ 0 \end{pmatrix}, \ldots,$ $\boldsymbol{e}_k = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \leftarrow k$ 番目 $, \ldots,$ $\boldsymbol{e}_n = \begin{pmatrix} 0 \\ \vdots \\ \\ \\ \vdots \\ 0 \\ 1 \end{pmatrix}$ とおくと

$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \sum_{k=1}^{n} x_k \boldsymbol{e}_k$ である。更に $\boldsymbol{f}_k = f(\boldsymbol{e}_k)$ $(k = 1, \ldots, n)$ とおく。

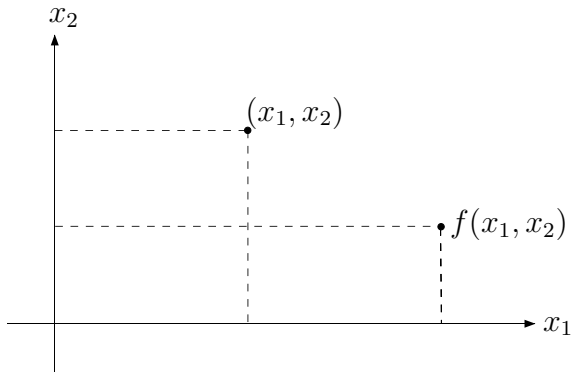
$f(\boldsymbol{x}) = f\left(\sum_{k}^{n} x_k \boldsymbol{e}_k\right)$ は線型性を用いると次のように表される：

$$
\begin{aligned}
f(\boldsymbol{x}) &= f\left(\sum_{k}^{n} x_k \boldsymbol{e}_k\right) \\
&= f(x_1 \boldsymbol{e}_1) + \cdots + f(x_k \boldsymbol{e}_k) + \cdots + f(x_n \boldsymbol{e}_n) \\
&= x_k f(\boldsymbol{e}_1) + \cdots + x_k f(\boldsymbol{e}_k) + \cdots + x_n f(\boldsymbol{e}_n) \\
&= \sum_{k=1}^{n} x_k \boldsymbol{f}_k \\
&= \begin{pmatrix} \boldsymbol{f}_1 \cdots \boldsymbol{f}_k \cdots \boldsymbol{f}_n \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_k \\ \vdots \\ x_n \end{pmatrix}
\end{aligned}
\tag{1.5}
$$

従って $A = \begin{pmatrix} \boldsymbol{f}_1 \cdots \boldsymbol{f}_n \end{pmatrix}$ とおくと $f(\boldsymbol{x}) = A\boldsymbol{x}$ となる。 □

次に、線形写像の中で (1.2) の形のものの次に簡単なものを考える。
次の (x_1, x_2) を座標とする平面から平面への線形写像を考える：

$$
f(x_1, x_2) = (2x_1, \tfrac{1}{2}x_2) \tag{1.6}
$$

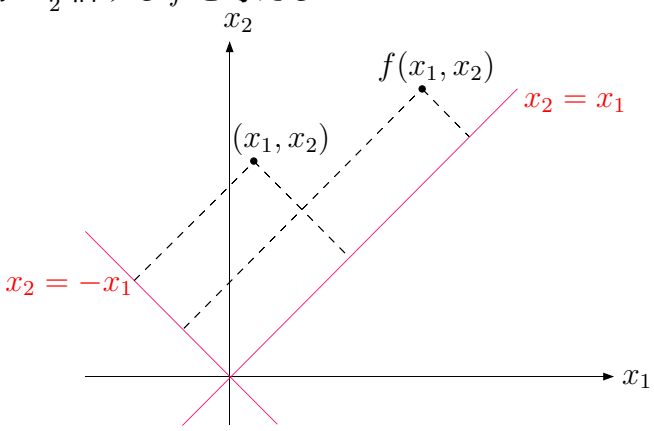


これは次のように表される：

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & \frac{1}{2} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \tag{1.7}$$

(ただし $f(x_1, x_2) = (y_1, y_2)$ とする。)

次に、(x_1, x_2) を、直線 $x_2 = x_1$ に沿う方向に 2 倍し、直線 $x_2 = -x_1$ に沿う方向に $\frac{1}{2}$ 倍する f を考える：



これは

$$f(x_1, x_2) = \left(\tfrac{5}{4}x_1 + \tfrac{3}{4}x_2, \tfrac{3}{4}x_1 + \tfrac{5}{4}x_2\right) \tag{1.8}$$

つまり

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} \frac{5}{4} & \frac{3}{4} \\ \frac{3}{4} & \frac{5}{4} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \tag{1.9}$$

と表される。

この写像は意味を考えると分かりやすいにも関わらず、(1.8) や (1.9) を見てもその意味は分かりづらい。そこで、もう少し分かりやすくこの写像を表すことを考える。

そのために定理2の証明と同様に考える。$\boldsymbol{e}_1 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$, $\boldsymbol{e}_2 = \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$ とすると、$\boldsymbol{e}_1$ は直線 $x_1 = x_2$ の単位方向ベクトルであり、$\boldsymbol{e}_2$ は直線 $x_2 = -x_1$ の単位方向ベクトルである。

$\boldsymbol{e}_1, \boldsymbol{e}_2$ は一次独立なので、任意の二次元ベクトルはこれらの一次結合で表される。すなわち、任意の $\boldsymbol{x} \in \mathbb{R}^2$ について $\boldsymbol{x} = \alpha_1 \boldsymbol{e}_1 + \alpha_2 \boldsymbol{e}_2$ となる $\alpha_1, \alpha_2 \in \mathbb{R}$ がただ一つ存在する。このとき

$$\begin{aligned} f(\boldsymbol{x}) &= f(\alpha_1 \boldsymbol{e}_1 + \alpha_2 \boldsymbol{e}_2) \\ &= f(\alpha_1 \boldsymbol{e}_1) + f(\alpha_2 \boldsymbol{e}_2) \\ &= \alpha_1 f(\boldsymbol{e}_1) + \alpha_2 f(\boldsymbol{e}_2) \end{aligned} \tag{1.10}$$

となるが、f は直線 $x_2 = x_1$ に沿う方向に 2 倍し、直線 $x_2 = -x_1$ に沿う方向に $\frac{1}{2}$ 倍する写像なので

$$f(\boldsymbol{e}_1) = 2\boldsymbol{e}_1, \quad f(\boldsymbol{e}_2) = \tfrac{1}{2}\boldsymbol{e}_2 \tag{1.11}$$

である。従って

$$f(\alpha_1 \boldsymbol{e}_1 + \alpha_2 \boldsymbol{e}_2) = 2\alpha_1 \boldsymbol{e}_1 + \tfrac{1}{2}\alpha_2 \boldsymbol{e}_2 \tag{1.12}$$

従って、直線 $x_2 = x_1$ と $x_2 = -x_1$ を座標軸とする座標で考えると、$f(\boldsymbol{x}) = \beta_1 \boldsymbol{e}_1 + \beta_2 \boldsymbol{e}_2$ と書いて

$$\beta_1 \boldsymbol{e}_1 + \beta_2 \boldsymbol{e}_2 = 2\alpha_1 \boldsymbol{e}_1 + \tfrac{1}{2}\alpha_2 \boldsymbol{e}_2 \tag{1.13}$$

つまり

$$\begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & \frac{1}{2} \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} \tag{1.14}$$

となる。よって、この座標では f は対角行列で表され、非常に分かりやすくなる。

2 行列の対角化

この節では、与えられた線形写像が座標を変えることにより (1.14) のような形に表されるための条件を考える。

$\boldsymbol{x} \in \mathbb{R}^n$ に $f(\boldsymbol{x}) \in \mathbb{R}^n$ を対応させる線形写像が $n \times n$ 実行列 A を用いて

$$f(\boldsymbol{x}) = A\boldsymbol{x} \tag{2.1}$$

と表されているとする。

以下話を単純にするためにこの A について次を仮定する：

仮定 1.

A の固有値は全て実数で互いに異なるとする。(つまり A の固有方程式に重根は無いとする。)

A の固有値を $\lambda_1, \ldots, \lambda_n$ とし、$\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ をそれぞれに属する固有ベクトルとする。

補題 1.

$\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ は一次独立である。

証明.

$$\alpha_1 \boldsymbol{e}_1 + \cdots + \alpha_n \boldsymbol{e}_n = \boldsymbol{0} \tag{2.2}$$

とする。

両辺に左から A をかけると、$A\boldsymbol{e}_k = \lambda_k \boldsymbol{e}_k$ なので、

$$\begin{aligned} A\left(\alpha_1 \boldsymbol{e}_1 + \cdots + \alpha_n \boldsymbol{e}_n\right) &= \alpha_1 A\boldsymbol{e}_1 + \cdots + \alpha_n A\boldsymbol{e}_n \\ &= \alpha_1 \lambda_1 \boldsymbol{e}_1 + \cdots + \alpha_n \lambda_n \boldsymbol{e}_n \\ &= A\boldsymbol{0} = \boldsymbol{0} \end{aligned} \tag{2.3}$$

(2.3) から (2.2) の λ_1 倍を引くと：

$$\alpha_2(\lambda_2 - \lambda_1)\boldsymbol{e}_2 + \cdots + \alpha_n(\lambda_n - \lambda_1)\boldsymbol{e}_n = \boldsymbol{0} \tag{2.4}$$

さらに、これに左から A をかけたものから λ_2 をかけたものを引き、と繰り返すと

$$\alpha_n(\lambda_n - \lambda_1)(\lambda_n - \lambda_2)\cdots(\lambda_n - \lambda_{n-1})\boldsymbol{e}_n = \boldsymbol{0} \tag{2.5}$$

が得られる。

λ_k は互いに異なるので $\alpha_n = 0$。従って、(2.5) の一つ前の式は

$$\begin{aligned} &\alpha_{n-1}(\lambda_{n-1} - \lambda_1)\cdots(\lambda_{n-1} - \lambda_{n-2})\boldsymbol{e}_{n-1} + \alpha_n(\lambda_n - \lambda_1)\cdots(\lambda_n - \lambda_{n-2})\boldsymbol{e}_n \\ &= \alpha_{n-1}(\lambda_{n-1} - \lambda_1)\cdots(\lambda_{n-1} - \lambda_{n-2})\boldsymbol{e}_{n-1} \\ &= \boldsymbol{0} \end{aligned} \tag{2.6}$$

よって $\alpha_{n-1} = 0$。これを繰り返して、全ての k について $\alpha_k = 0$。 □

従って、$\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ は $\mathbb{R}^n$ の基底である。よって、任意の $\boldsymbol{x} \in \mathbb{R}^n$ は $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ の一次結合で一通りに表される。

$\boldsymbol{x} = \alpha_1 \boldsymbol{e}_1 + \cdots + \alpha_n \boldsymbol{e}_n$ とすると

$$\begin{aligned} f(\boldsymbol{x}) &= f(\alpha_1 \boldsymbol{e}_1 + \cdots + \alpha_n \boldsymbol{e}_n) \\ &= \alpha_1 f(\boldsymbol{e}_1) + \cdots + \alpha_n f(\boldsymbol{e}_n) \\ &= \alpha_1 \lambda_1 \boldsymbol{e}_1 + \cdots + \alpha_n \lambda_n \boldsymbol{e}_n \end{aligned} \tag{2.7}$$

つまり、$f(\boldsymbol{x}) = \beta_1 \boldsymbol{e}_1 + \cdots + \beta_n \boldsymbol{e}_n$ として $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ を基底とする座標で f を表すと

$$\begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} \tag{2.8}$$

と対角行列で表され、非常に分かりやすくなる。

この対角行列を元の A と関係づけるには $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ を元の座標で表せばよい。

$\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ を成分で書いて $\boldsymbol{e}_1 = \begin{pmatrix} e_{1,1} \\ \vdots \\ e_{n,1} \end{pmatrix}, \ldots, \boldsymbol{e}_n = \begin{pmatrix} e_{1,n} \\ \vdots \\ e_{n,n} \end{pmatrix}$ とする。これを用いて (2.7) を書くと

$$\begin{aligned} f(\boldsymbol{x}) &= \alpha_1 \lambda_1 \boldsymbol{e}_1 + \cdots + \alpha_n \lambda_n \boldsymbol{e}_n \\ &= (\boldsymbol{e}_1 \boldsymbol{e}_2 \cdots \boldsymbol{e}_n) \begin{pmatrix} \lambda_1 \alpha_1 \\ \lambda_2 \alpha_2 \\ \vdots \\ \lambda_n \alpha_n \end{pmatrix} \\ &= \begin{pmatrix} e_{1,1} & \cdots & e_{1,n} \\ \vdots & \ddots & \vdots \\ e_{n,1} & \cdots & e_{n,n} \end{pmatrix} \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} \end{aligned} \tag{2.9}$$

だが、

$$
\begin{aligned}
f(\boldsymbol{x}) &= A\boldsymbol{x} \\
&= A\left(\alpha_1\boldsymbol{e}_1 + \cdots + \alpha_n\boldsymbol{e}_n\right) \\
&= A\left(\boldsymbol{e}_1 \cdots \boldsymbol{e}_n\right)\begin{pmatrix}\alpha_1\\ \vdots \\ \alpha_n\end{pmatrix} \\
&= A\begin{pmatrix} e_{1,1} & \cdots & e_{1,n} \\ \vdots & \ddots & \vdots \\ e_{n,1} & \cdots & e_{n,n}\end{pmatrix}\begin{pmatrix}\alpha_1\\ \vdots \\ \alpha_n\end{pmatrix}
\end{aligned}
\tag{2.10}
$$

なので

$$
A\begin{pmatrix} e_{1,1} & \cdots & e_{1,n} \\ \vdots & \ddots & \vdots \\ e_{n,1} & \cdots & e_{n,n}\end{pmatrix}\begin{pmatrix}\alpha_1\\ \vdots \\ \alpha_n\end{pmatrix} = \begin{pmatrix} e_{1,1} & \cdots & e_{1,n} \\ \vdots & \ddots & \vdots \\ e_{n,1} & \cdots & e_{n,n}\end{pmatrix}\begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n\end{pmatrix}\begin{pmatrix}\alpha_1\\ \vdots \\ \alpha_n\end{pmatrix}
\tag{2.11}
$$

となる。

補題 2.

A, B を $m \times n$ 行列とするとき、任意の $\boldsymbol{x} \in \mathbb{R}^n$ に対し $A\boldsymbol{x} = B\boldsymbol{x}$ が成り立つならば $A = B$ である。

証明.

A の各列を $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$、$B$ の各列を $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n$ として $A = (\boldsymbol{a}_1 \cdots \boldsymbol{a}_n)$、$B = (\boldsymbol{b}_1 \cdots \boldsymbol{b}_n)$ とする。

このとき $A\boldsymbol{x} = B\boldsymbol{x}$ に $\boldsymbol{x} = \begin{pmatrix}0\\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0\end{pmatrix} \leftarrow k$ 行目　を代入すると $\boldsymbol{a}_k = \boldsymbol{b}_k$ が得られる。従って $A = B$ である。 □

この補題を (2.11) に適用すると

$$A\begin{pmatrix} e_{1,1} & \cdots & e_{1,n} \\ \vdots & \ddots & \vdots \\ e_{n,1} & \cdots & e_{n,n} \end{pmatrix} = \begin{pmatrix} e_{1,1} & \cdots & e_{1,n} \\ \vdots & \ddots & \vdots \\ e_{n,1} & \cdots & e_{n,n} \end{pmatrix}\begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix} \tag{2.12}$$

である。この両辺に $\begin{pmatrix} e_{1,1} & \cdots & e_{1,n} \\ \vdots & \ddots & \vdots \\ e_{n,1} & \cdots & e_{n,n} \end{pmatrix}^{-1}$ を左からかけると

$$\begin{pmatrix} e_{1,1} & \cdots & e_{1,n} \\ \vdots & \ddots & \vdots \\ e_{n,1} & \cdots & e_{n,n} \end{pmatrix}^{-1} A \begin{pmatrix} e_{1,1} & \cdots & e_{1,n} \\ \vdots & \ddots & \vdots \\ e_{n,1} & \cdots & e_{n,n} \end{pmatrix} = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix} \tag{2.13}$$

となり、$\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ を基底とする座標で f を表した時の表現行列が、元の行列 A と基底 $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ を元の座標で表した時の成分で書かれた。

例 1.

(1.8) つまり、行列表示すると (1.9) で表される線形写像を考える：

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} \frac{5}{4} & \frac{3}{4} \\ \frac{3}{4} & \frac{5}{4} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \tag{1.9}$$

$\boldsymbol{e}_1 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$, $\boldsymbol{e}_2 = \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$ とする。$A = \begin{pmatrix} \frac{5}{4} & \frac{3}{4} \\ \frac{3}{4} & \frac{5}{4} \end{pmatrix}$ と書くと
$A\boldsymbol{e}_1 = 2\boldsymbol{e}_1, A\boldsymbol{e}_2 = \frac{1}{2}\boldsymbol{e}_2$ である。
これを (2.13) の左辺に代入すると

$$\begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}^{-1} \begin{pmatrix} \frac{5}{4} & \frac{3}{4} \\ \frac{3}{4} & \frac{5}{4} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \tag{2.14}$$

となるが、これを計算すると確かに右辺

$$\begin{pmatrix} 2 & 0 \\ 0 & \frac{1}{2} \end{pmatrix} \tag{2.15}$$

になる。

3　対称行列の場合

(1.9) の表現行列 $A = \begin{pmatrix} \frac{5}{4} & \frac{3}{4} \\ \frac{3}{4} & \frac{5}{4} \end{pmatrix}$ は、対称行列という特別な形をしている：

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} \frac{5}{4} & \frac{3}{4} \\ \frac{3}{4} & \frac{5}{4} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \tag{1.9}$$

その結果として、A を対角化する座標変換は座標平面上の角度 $\frac{\pi}{4}$ の単なる回転である。以下ではそのようなことが一般的であることを解説する。

定義 1.
実行列 A が ${}^tA = A$ (つまり A の (i,j) 成分 $a_{i,j}$ は全て実数で $a_{i,j} = a_{j,i}$, $(i,j = 1,2,\ldots,n)$) を満たすとき、A は実対称行列であるという。

定義 2.
$\mathbb{R}^n$ のベクトル $\boldsymbol{x} = (x_1,\ldots,x_n), \boldsymbol{y} = (y_1,\ldots,y_n)$ について $(\boldsymbol{x},\boldsymbol{y}) := \sum_{i=1}^{n} x_i y_i$ をユークリッド内積または標準内積と呼ぶ。
$\mathbb{C}^n$ のベクトル $\boldsymbol{x} = (x_1,\ldots,x_n), \boldsymbol{y} = (y_1,\ldots,y_n)$ について $(\boldsymbol{x},\boldsymbol{y}) := \sum_{i=1}^{n} x_i \bar{y}_i$ をエルミート内積と呼ぶ。

性質 1 (内積の性質)**.** $\mathbb{R}^n$ または $\mathbb{C}^n$ のベクトル $\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}$ について

(1) $(\boldsymbol{x},\boldsymbol{x}) \geq 0$, かつ、 $(\boldsymbol{x},\boldsymbol{x}) = 0 \Leftrightarrow \boldsymbol{x} = \boldsymbol{0}$
($\boldsymbol{x} \in \mathbb{C}^n$ の場合でも $(\boldsymbol{x},\boldsymbol{x})$ は実数)

(2) $(\boldsymbol{x},\boldsymbol{y}) = (\boldsymbol{y},\boldsymbol{x})$ ($\boldsymbol{x},\boldsymbol{y} \in \mathbb{R}^n$ のとき)
$(\boldsymbol{x},\boldsymbol{y}) = \overline{(\boldsymbol{y},\boldsymbol{x})}$ ($\boldsymbol{x},\boldsymbol{y} \in \mathbb{C}^n$ のとき)

(3) $(\boldsymbol{x}+\boldsymbol{y},\boldsymbol{z}) = (\boldsymbol{x},\boldsymbol{z}) + (\boldsymbol{y},\boldsymbol{z})$

(4) $(\alpha\boldsymbol{x},\boldsymbol{y}) = (\boldsymbol{x},\alpha\boldsymbol{y}) = \alpha(\boldsymbol{x},\boldsymbol{y})$, ($\boldsymbol{x},\boldsymbol{y} \in \mathbb{R}^n$, $\alpha \in \mathbb{R}$ のとき)
$(\alpha\boldsymbol{x},\boldsymbol{y}) = \alpha(\boldsymbol{x},\boldsymbol{y})$, $(\boldsymbol{x},\alpha\boldsymbol{y}) = \bar{\alpha}(\boldsymbol{x},\boldsymbol{y})$, ($\boldsymbol{x},\boldsymbol{y} \in \mathbb{C}^n$, $\alpha \in \mathbb{C}$ のとき)

定義 3. $\mathbb{R}^n$ または $\mathbb{C}^n$ のベクトル $\boldsymbol{x}$ について

$$\|\boldsymbol{x}\| := \sqrt{(\boldsymbol{x},\boldsymbol{x})} \tag{3.1}$$

を $\boldsymbol{x}$ のノルムとよぶ。

定義 4. $\mathbb{R}^n$ または $\mathbb{C}^n$ のベクトル $\boldsymbol{x}, \boldsymbol{y}$ について

$$(\boldsymbol{x}, \boldsymbol{y}) = 0 \tag{3.2}$$

であるとき $\boldsymbol{x}$ と $\boldsymbol{y}$ は直交するという。

補題 3.
A を n 次実対称行列とするとき $\mathbb{R}^n$ または $\mathbb{C}^n$ のベクトル $\boldsymbol{x}, \boldsymbol{y}$ について $(A\boldsymbol{x}, \boldsymbol{y}) = (\boldsymbol{x}, A\boldsymbol{y})$ が成り立つ。

問題 1. 上の補題を証明せよ。

定理 3.

(1) 実対称行列の固有値は実数。

(2) 実対称行列の異なる固有値に属する固有ベクトルは直交する。

証明.

(1) $\lambda \in \mathbb{C}$ を A の固有値とし $\boldsymbol{x} \in \mathbb{C}^n$ を固有値 λ に属する固有ベクトルとする。

補題 3 より $(A\boldsymbol{x}, \boldsymbol{x}) = (\boldsymbol{x}, A\boldsymbol{x})$ だが、$A\boldsymbol{x} = \lambda\boldsymbol{x}$ なので $(\lambda\boldsymbol{x}, \boldsymbol{x}) = (\boldsymbol{x}, \lambda\boldsymbol{x})$ である。

定義 1(4) より $(\lambda\boldsymbol{x}, \boldsymbol{x}) = \lambda(\boldsymbol{x}, \boldsymbol{x}) = \lambda\|\boldsymbol{x}\|^2$ かつ $(\boldsymbol{x}, \lambda\boldsymbol{x}) = \bar{\lambda}(\boldsymbol{x}, \boldsymbol{x}) = \bar{\lambda}\|\boldsymbol{x}\|^2$

従って $\lambda = \bar{\lambda}$ であり λ は実数である。

(2) $\lambda, \mu \in \mathbb{R}$ を A の固有値とし $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ をそれぞれに属する固有ベクトルとする。 さらに $\lambda \neq \mu$ とする。

$A\boldsymbol{x} = \lambda\boldsymbol{x}, A\boldsymbol{y} = \mu\boldsymbol{y}$ を $(A\boldsymbol{x}, \boldsymbol{y}) = (\boldsymbol{x}, A\boldsymbol{y})$ に代入すると $\lambda(\boldsymbol{x}, \boldsymbol{y}) = \mu(\boldsymbol{x}, \boldsymbol{y})$ なので $(\boldsymbol{x}, \boldsymbol{y}) = 0$ である。

□

以下では次を仮定して話を単純にする：

仮定 2.

(1) A は実対称行列。

(2) A の固有値は互いに異なる。(すなわち A の固有方程式は重根を持たない。)

このとき、定理 3 の (1) より A の固有値は全て実数なので、前セクションの議論が全てそのまま適用できる。従って、A の固有値を $\lambda_1, \ldots, \lambda_n$ とし、$\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ をそれぞれに属する固有ベクトルとすると、(2.13) により

$$(\boldsymbol{e}_1 \cdots \boldsymbol{e}_n)^{-1} A\,(\boldsymbol{e}_1 \cdots \boldsymbol{e}_n) = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix} \tag{3.3}$$

となる。

定義 5. 実行列 A が ${}^tAA = A\,{}^tA = E$ すなわち $A^{-1} = {}^tA$ を満たすとき、A は直交行列であるという。

定理 4.
上の $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ が $\|\boldsymbol{e}_k\| = \cdots = \|\boldsymbol{e}_k\| = 1$ を満たすとすると $(\boldsymbol{e}_1 \cdots \boldsymbol{e}_n)$ は直交行列である。すなわち ${}^t(\boldsymbol{e}_1 \cdots \boldsymbol{e}_n)\,(\boldsymbol{e}_1 \cdots \boldsymbol{e}_n) = (\boldsymbol{e}_1 \cdots \boldsymbol{e}_n)\ {}^t(\boldsymbol{e}_1 \cdots \boldsymbol{e}_n) = E$ が成り立つ。

証明.

$$\begin{aligned} {}^t(\boldsymbol{e}_1 \cdots \boldsymbol{e}_n)\,(\boldsymbol{e}_1 \cdots \boldsymbol{e}_n) &= \begin{pmatrix} {}^t\boldsymbol{e}_1 \\ \vdots \\ {}^t\boldsymbol{e}_n \end{pmatrix} (\boldsymbol{e}_1 \cdots \boldsymbol{e}_n) \\ &= \begin{pmatrix} {}^t\boldsymbol{e}_1\boldsymbol{e}_1 & \cdots & {}^t\boldsymbol{e}_1\boldsymbol{e}_n \\ \vdots & \ddots & \vdots \\ {}^t\boldsymbol{e}_n\boldsymbol{e}_1 & \cdots & {}^t\boldsymbol{e}_n\boldsymbol{e}_n \end{pmatrix} \\ &= \begin{pmatrix} (\boldsymbol{e}_1, \boldsymbol{e}_1) & \cdots & (\boldsymbol{e}_1, \boldsymbol{e}_n) \\ \vdots & \ddots & \vdots \\ (\boldsymbol{e}_n, \boldsymbol{e}_1) & \cdots & (\boldsymbol{e}_n, \boldsymbol{e}_n) \end{pmatrix} \end{aligned} \tag{3.4}$$

だが、仮定より $(\boldsymbol{e}_k, \boldsymbol{e}_k) = \|\boldsymbol{e}_k\|^2 = 1$ $(k = 1, \ldots, n)$ であり、定理 3(2) より $(\boldsymbol{e}_k, \boldsymbol{e}_l) = 0$ $(k \neq l)$ なので、

$$\begin{pmatrix} (\boldsymbol{e}_1, \boldsymbol{e}_1) & \cdots & (\boldsymbol{e}_1, \boldsymbol{e}_n) \\ \vdots & \ddots & \vdots \\ (\boldsymbol{e}_n, \boldsymbol{e}_1) & \cdots & (\boldsymbol{e}_n, \boldsymbol{e}_n) \end{pmatrix} = E \tag{3.5}$$

が成り立つ。つまり ${}^t(\boldsymbol{e}_1 \cdots \boldsymbol{e}_n) = (\boldsymbol{e}_1 \cdots \boldsymbol{e}_n)^{-1}$ である。よって
$(\boldsymbol{e}_1 \cdots \boldsymbol{e}_n)\ {}^t(\boldsymbol{e}_1 \cdots \boldsymbol{e}_n) = E$ も成り立つ。 □

このことより、(3.3) の $(\boldsymbol{e}_1 \cdots \boldsymbol{e}_n)$ として直交行列を取ることができる。実際、具体例の (2.14) の $\begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$ は直交行列にとってある。

二つの平面ベクトル $\boldsymbol{e}_1, \boldsymbol{e}_2$ について、$\|\boldsymbol{e}_1\| = \|\boldsymbol{e}_2\| = 1$ かつ $\boldsymbol{e}_1$ と $\boldsymbol{e}_2$ は直交するとする。

先ず、$\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ を原点中心に適切な角度の回転をすると $\boldsymbol{e}_1$ に一致させることができる。このときとき $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ を同じ角だけ回転させるたものは $\boldsymbol{e}_1$ と直交するノルムが 1 のベクトルなので $\pm\boldsymbol{e}_2$ のどちらかと一致する。

すなわち、2 次の実対称行列を対角化するような座標は、元の座標を原点中心に回転させるか、回転後に $\boldsymbol{e}_1$ 方向の直線に対して折り返すことによって得られる。

具体例では $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ と $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ を $\frac{\pi}{4}$ だけ回転すると $\boldsymbol{e}_1 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$ と $\boldsymbol{e}_2 = \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$ に一致する。

3 次元以上でも事情は同じで、直交行列による座標変換は n 次元の回転と折り返しの合成によって得られる。(まず 3 次元について自分で考えてみよ。)

問題 2.

エルミート行列のユニタリー行列による対角化について、本講議論を再現せよ。ただし、複素正方行列 A は $A^* := {}^t\bar{A}$ として $A^* = A$ をみたす時エルミート行列であるといい、$A^*A = E$ をみたす時ユニタリー行列であるという。

注意 1.

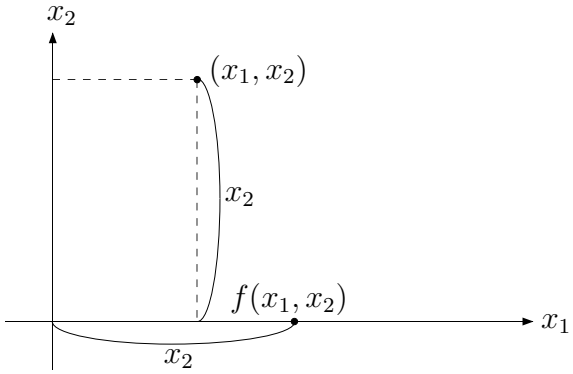
(1) 仮定 1 を満たさない正方行列は対角化できるとは限らない。一般にはジョルダン標準形に変換される。

(2) 実対称行列については、仮定 2 (2) を満たさなくても必ず直交行列で対角化される。詳細について [東電大資料] 等を見よ。

4　付録—ジョルダン標準形

以下では対角化できない線型写像について考える：

$$f(x_1, x_2) = (x_2, 0) \tag{4.1}$$



これは行列を使うと次の様に表される：

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \tag{4.2}$$

(ただし $f(x_1, x_2) = (y_1, y_2)$ とする。)

係数行列 $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ の固有値は 0 の重解なので、もし対角化できるのあれば

$$P^{-1} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} P = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \tag{4.3}$$

となる正則行列 P が存在するはずである。これはすなわち

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = P \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} P^{-1} \tag{4.4}$$

ということだが、この左辺はゼロ行列ではないのに右辺はゼロ行列で矛盾する。

更に、一般に次の形の行列は対角化できない：

$$\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} \tag{4.5}$$

問題 3.
上の形の行列が対角化できないことを示せ。

$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ について考える。
先ず

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{4.6}$$

なので、$\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ は固有値 0 に属する固有ベクトルであり、$\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ は固有ベクトルではないが、固有値 0 に属する固有値ベクトルに写る。

同様のことが以下の仮定のもとで成り立つ。

仮定 3.
2次正方行列 A はスカラー行列ではなく、その固有方程式は重根をもつ。

この仮定の下で、λ を A の固有値とするとき、先ず、A はスカラー行列ではないので

$$A - \lambda E \neq O \tag{4.7}$$

だが、ハミルトン・ケーレーの定理より

$$(A - \lambda E)^2 = O \tag{4.8}$$

が成り立つ。

定理 5 (2次正方行列の場合のハミルトン・ケーレーの定理)**.**
$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ とし、その固有多項式

$$\det(A - \lambda E) = \lambda^2 - (a+d)\lambda + ad - bc \tag{4.9}$$

について、λ に形式的に A を代入したものはゼロ行列になる：

$$A^2 - (a+d)A + (ad - bc)E = O \tag{4.10}$$

問題 4.
2次正方行列の場合のハミルトン・ケーレーの定理を証明せよ。

(4.7) と (4.8) より 以下を満たすベクトル $\boldsymbol{u}, \boldsymbol{v}$ が存在する：

$$(A-\lambda E)\boldsymbol{u}=\boldsymbol{0}, \quad (A-\lambda E)\boldsymbol{v}=\boldsymbol{u} \tag{4.11}$$

すなわち

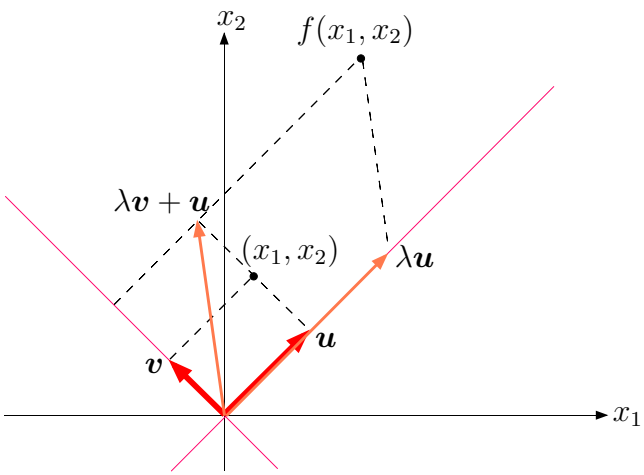
$$A\boldsymbol{u}=\lambda\boldsymbol{u}, \quad A\boldsymbol{v}=\lambda\boldsymbol{v}+\boldsymbol{u} \tag{4.12}$$

$\because (A-\lambda E)\neq O$ なので $(A-\lambda E)\boldsymbol{v}\neq\boldsymbol{0}$ である $\boldsymbol{v}$ が存在する。(すなわち $\boldsymbol{v}$ は固有ベクトルではない。)

$\boldsymbol{u}=(A-\lambda E)\boldsymbol{v}$ とおくと、$(A-\lambda E)^2=O$ より

$$(A-\lambda E)\boldsymbol{u}=(A-\lambda E)^2\boldsymbol{v}=\boldsymbol{0} \tag{4.13}$$

これはつまり、$\boldsymbol{u}, \boldsymbol{v}$ を座標軸の方向ベクトルとする座標系では、$\boldsymbol{u}$ 方向は λ 倍され $\boldsymbol{v}$ 方向は λ 倍されるだけではなく、元の大きさの分だけ $\boldsymbol{u}$ 方向にずれることを意味する。



従って、

$$A(\boldsymbol{u}\ \boldsymbol{v})=(A\boldsymbol{u}\ A\boldsymbol{v})=(\lambda\boldsymbol{u}\ \lambda\boldsymbol{v}+\boldsymbol{u})=(\boldsymbol{u}\ \boldsymbol{v})\begin{pmatrix}\lambda & 1\\ 0 & \lambda\end{pmatrix} \tag{4.14}$$

よって、 $P=(\boldsymbol{u}\ \boldsymbol{v})$ とおいて、上式全体に左から P^{-1} をかけることで

$$P^{-1}AP=\begin{pmatrix}\lambda & 1\\ 0 & \lambda\end{pmatrix} \tag{4.15}$$

となる。これをジョルダン標準形という。

注意 2.

2 次正方行列 A の固有方程式が重根を持ち、全ての 2 次元ベクトルが固有ベクトルの時は A はスカラー行列すなわち $A = \lambda E$ となり、最初から対角化されている。

参考文献

[東電大資料] 東京電機大学 システムデザイン工学部・未来科学部・工学部 数学系列 資料
「対称行列の直交行列による対角化」
https://www.cck.dendai.ac.jp/math/support/ch6-supp/対称行列の直交行列による対角化.pdf