2.2 相似・対角化・スペクトル分解

固有値・固有ベクトルという概念を使って、行列を変形することを考えます。主成分分析は、このテクニックを使って計算しています。固有値はすでに先週説明しましたが、大切なところなので、その復習から入ります。

行列をベクトル的概念でとらえれために、行列の説明で最初にやった、100m先の角を左に曲がって200m行った先のベクトルの話に戻ります。この時道が３０度歪んでいたので、頭の中で直行している地図を、実際の位置に合わせるための変形として、行列を左からかけるという話です。

$$\begin{pmatrix} 1 & 0 \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{pmatrix}$$

というのがその時の変換に使った行列です。すでに使った図ですが、変換の内容を図示すると、次のよう描けます

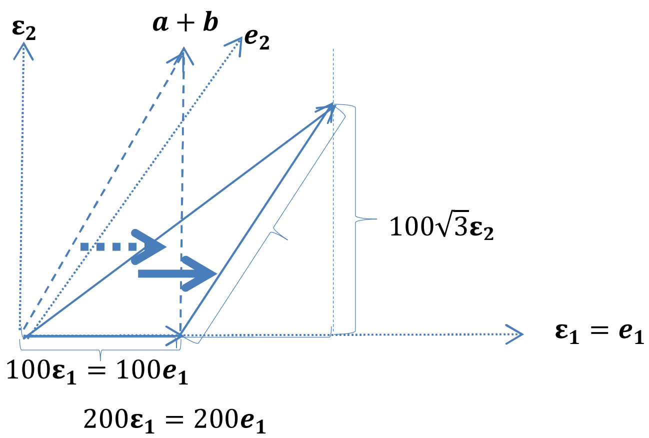

図 1.ベクトル空間の移し替え

この図の中で、座標軸の移動に伴って、ベクトル作る三角形は歪みましたが、破線で示した太い矢印は、実線のように移動して少し長さが変わったものの、矢印の方向は変わっていません。このように行列を掛けることによって、移動したり、大きくなっても方向が変わらないベクトルをその行列の固有ベクトルと言います。また、どのくらい大きくなるかという比率を固有値と言います。

固有値の説明にしばしば使われるのが、次のモナリザの絵です。何故モナリザの絵が使われるのかわかりませんが、昔からそうなっているので、慣習にしたがって、一応紹介します。

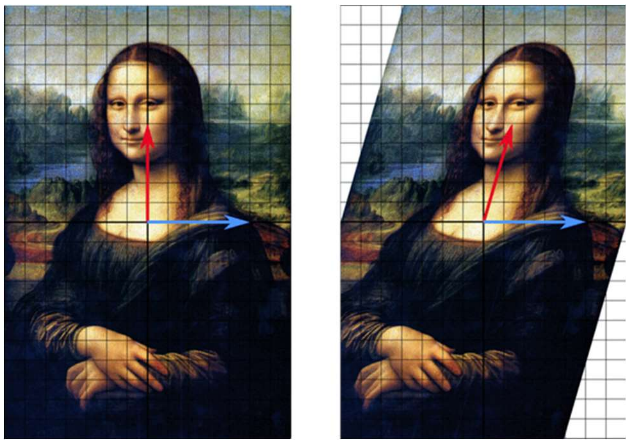

図 2.固有ベクトルの説明によく使われるモナリザの絵

図 2 は左のモナリゼの絵の原図に「平行四辺形西つぶすような」行列を掛けて、右側の絵のベクトル空間に投影した図なのですが。元の絵の赤いベクトルは方向が変わっていますが、青いベクトルは方向が変わっていません。このようベクトルを、投影に使った行列の固有 g ベクトルと言います。この場合は、長さも変わっていませんから、固有値は 1 です。

ベクトルの方向が変わらずに大きさだけが変わるということを式で表すと、

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ \vdots \\ x_n \end{pmatrix} = \lambda \begin{pmatrix} x_1 \\ \vdots \\ \vdots \\ x_n \end{pmatrix}$$

ということです。λ は固有値です。行列式を連立方程式で書くと

$$a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = \lambda x_1$$
$$a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = \lambda x_2$$
$$\vdots$$
$$a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = \lambda x_2$$

なので、左辺に移行して行列式の形で書くと

$$\begin{pmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - \lambda \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ \vdots \\ 0 \end{pmatrix}$$

ということです。

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

解けば固有値 λ が得られます。[1]これを固有方程式と言います。この式からも明らかなよ

うに n 次の行列には通常 n 個の固有値があります。もちろん重根があれば、その分重複ができますから、重根の数だけ固有値は少なくなります。
具体的にやってみます。
図に示した地図の例です

$$\begin{vmatrix} 1-\lambda & 0 \\ \frac{1}{2} & \frac{\sqrt{3}}{2}-\lambda \end{vmatrix} = 0$$

$$(1-\lambda)\left(\frac{\sqrt{3}}{2}-\lambda\right) = 0$$

$$\lambda = 1, \qquad \lambda = \frac{\sqrt{3}}{2}$$

これを使って、固有ベクトルを求めます。この場合はかなり特殊で、$\lambda = 1$を代入すると、

$$\begin{pmatrix} 0 & 0 \\ \frac{1}{2} & \frac{\sqrt{3}}{2}-1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\frac{1}{2}x_1 + \left(\frac{\sqrt{3}}{2}-1\right)x_2 = 0$$

になって、

$$x_1 + \left(\sqrt{3}-2\right)x_2 = 0$$

$$x_1 = (2-\sqrt{3})x_2$$

だから、

$$t\begin{pmatrix} 2-\sqrt{3} \\ 1 \end{pmatrix}$$

t は任意の実数

が固有値 1 に属する固有ベクトルです。
もう一つの固有値については

$$\begin{pmatrix} 1-\dfrac{\sqrt{3}}{2} & 0 \\ \dfrac{1}{2} & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$x_1 = 0$$

で y は任意の値をとりえるので、ベクトルとしての記述は

$$t\begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

t は任意の実数

となり、x_1軸に平行なベクトルが固有ベクトルの一つとなり、確かに実感と会います。
つまり、下図で示した２本の矢印と並行する矢印は、行列を掛けて変形しても、矢印の向きが変わらないのです。

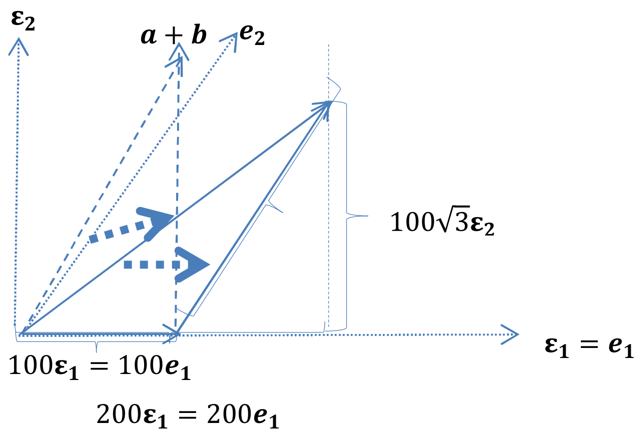

図 3.固有ベクトル

ここに挙げた例はかなり特殊なので
もう少し一般的な例を上げます。

$$A = \begin{pmatrix} 5 & -6 \\ 2 & -2 \end{pmatrix}$$

の固有値、固有ベクトルを求めます。

$$\begin{vmatrix} 5-\lambda & -1 \\ 1 & -2-\lambda \end{vmatrix}$$

$$= (5-\lambda)(-2-\lambda) + 12$$

$$= \lambda^2 - 3\lambda - 10 + 12$$

$$= (\lambda - 2)(\lambda - 1) = 0$$

$$\lambda = 2 \text{ or } \lambda = 1$$

$$\begin{pmatrix} 5-\lambda & -6 \\ 2 & -2-\lambda \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

に、$\lambda = 2$ を代入

$$\begin{pmatrix} 3 & -6 \\ 2 & -4 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$x_1 - 2x_2 = 0$$

固有値$\lambda = 2$に属する固有ベクトルは

$$\mathrm{t}\begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

$\lambda = 1$を代入

$$\begin{pmatrix} 4 & -6 \\ 2 & -3 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$2x_1 - 3x_2 = 0$$

固有値$\lambda = 1$に属する固有ベクトルは

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

となります。ここで見たように,固有値を求めるには

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

を解けばよいのです。これを一般化して書くと

$$|\boldsymbol{A} - \lambda\boldsymbol{E}| = 0$$

$\boldsymbol{E}$は単位行列

と書けます。これを行列 A の固有方程式といいます。

これに対して λ を t に変えたものを

$$|\boldsymbol{A} - t\boldsymbol{E}|$$

を固有多項式といいます。

$$\phi_A(t) = |\boldsymbol{A} - t\boldsymbol{E}|$$

と表すことが多いと思います。行列$\boldsymbol{A}$と行列$\boldsymbol{B}$ が相似であれば,

$$\phi_A(t) = \phi_B(t)$$

であり、行列$\boldsymbol{A}$と行列$\boldsymbol{B}$の固有値は一致します。

行列の相似

行列の相似という概念について説明します。

行列の相似は次のように定義されています。
行列 $\boldsymbol{C}$ と $\boldsymbol{D}$ が次のような関係にあるとき、行列 $\boldsymbol{C}$ と $\boldsymbol{D}$ を相似であるという。

$$\boldsymbol{C}=\boldsymbol{P^{-1}DP}$$

もっと簡単言えば、$\boldsymbol{P}$のような行列が存在すれば、$\boldsymbol{C}$と $\boldsymbol{D}$は相似だということです。
この定義の意味を３x３の行列を例にあげて説明します。ここでは、
D を対角成分以外の成分が０（対角行列）にしています。この説明では、対角行列でなくても良いのですが、実際、相似の対角行列を作るということがしばしば行われるので、対角行列の例を挙げています。

$$\boldsymbol{D}=\begin{pmatrix}3&0&0\\0&4&0\\0&0&1\end{pmatrix}$$
$$\boldsymbol{P}=\begin{pmatrix}1&2&1\\1&1&1\\1&1&2\end{pmatrix}$$

とすると、

$$|\boldsymbol{P}|=(2+2+1)-(1+4+1)=-1$$
$$\boldsymbol{P'}=\begin{pmatrix}1&1&1\\2&1&1\\1&1&2\end{pmatrix}$$
$$\boldsymbol{P^{-1}}=\frac{1}{|P|}\begin{pmatrix}\begin{vmatrix}1&1\\1&2\end{vmatrix}&-\begin{vmatrix}2&1\\1&2\end{vmatrix}&\begin{vmatrix}2&1\\1&1\end{vmatrix}\\-\begin{vmatrix}1&1\\1&2\end{vmatrix}&\begin{vmatrix}1&1\\1&2\end{vmatrix}&-\begin{vmatrix}1&1\\1&1\end{vmatrix}\\\begin{vmatrix}1&1\\1&1\end{vmatrix}&-\begin{vmatrix}1&1\\2&1\end{vmatrix}&\begin{vmatrix}1&1\\2&1\end{vmatrix}\end{pmatrix}$$
$$=\frac{1}{-1}\begin{pmatrix}1&-3&1\\-1&1&0\\0&1&-1\end{pmatrix}$$
$$=\begin{pmatrix}-1&3&-1\\1&-1&0\\0&-1&1\end{pmatrix}$$

念のために、ここで、$\boldsymbol{P\cdot P^{-1}=I}$を確認しましょう。

$$\boldsymbol{P\cdot P^{-1}}=\begin{pmatrix}1&2&1\\1&1&1\\1&1&2\end{pmatrix}\begin{pmatrix}-1&3&-1\\1&-1&0\\0&-1&1\end{pmatrix}$$
$$=\begin{pmatrix}-1+2+0&3-2-1&-1+0+1\\-1+1+0&3-1-1&-1+0+1\\-1+1+0&3-1-2&-1+0+2\end{pmatrix}$$
$$=\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}$$

以上で、

$$\begin{pmatrix}1&2&1\\1&1&1\\1&1&2\end{pmatrix}^{-1}=\begin{pmatrix}-1&3&-1\\1&-1&0\\0&-1&1\end{pmatrix}$$

が確認されました。

次に、$\boldsymbol{P^{-1}}DP$=Cとして

$\boldsymbol{D}$と $\boldsymbol{C}$が相似であることを確認します。

$$\boldsymbol{P^{-1}D}=\begin{pmatrix}-1&3&-1\\1&-1&0\\0&-1&1\end{pmatrix}\begin{pmatrix}3&0&0\\0&4&0\\0&0&1\end{pmatrix}$$

$$=\begin{pmatrix}-3&12&-1\\3&-4&0\\0&-4&1\end{pmatrix}$$

$$\boldsymbol{P^{-1}}\mathbf{DP}=\begin{pmatrix}-3&12&-1\\3&-4&0\\0&-4&1\end{pmatrix}\begin{pmatrix}1&2&1\\1&1&1\\1&1&2\end{pmatrix}$$

$$=\begin{pmatrix}-3+12-1&-6+12-1&-3+12-2\\3-4+0&6-4+0&3-4+0\\0-4+1&0-4+1&0-4+2\end{pmatrix}$$

$$=\begin{pmatrix}8&5&7\\-1&2&-1\\-3&-3&-2\end{pmatrix}$$

ということで、

$$\boldsymbol{C}=\begin{pmatrix}8&5&7\\-1&2&-1\\-3&-3&-2\end{pmatrix}$$

確かに $\boldsymbol{C}$と $\boldsymbol{D}$は相似ですが、そのことにどんな意味があるのかわかりません

ここで、$\boldsymbol{C}$および $\boldsymbol{D}$について、固有値を求めます。

$\boldsymbol{D}$については、固有値$\lambda = 4,3,1$

固有ベクトルは、固有値λ =4 について、$(0 \quad 1 \quad 0)'$

固有値λ =3 について、$(1 \quad 0 \quad 0)'$

固有値λ =1 について、$(0 \quad 0 \quad 1)'$

であることは自明

$\boldsymbol{C}$については、

$$\begin{vmatrix}(8-\lambda)&5&7\\-1&(2-\lambda)&-1\\-3&-3&(-2-\lambda)\end{vmatrix}=0$$

を解いて

$$\lambda^3-8\lambda^2+19\lambda-12=0$$

$$(\lambda-4)(\lambda-3)(\lambda-1)=0$$

$$\lambda=4,3,1$$

固有ベクトルは

$$\begin{pmatrix}8&5&7\\-1&2&-1\\-3&-3&-2\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=\lambda\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}$$

を解いて

$\lambda = 4$について、

$$8x_1+5x_2+7x_3=4x_1$$

$$-1x_1 + 2x_2 - x_3 = 4x_2$$
$$-3x_1 - 3x_2 - 2x_3 = 4x_3$$

の連立方程式を得る。

$$4x_1 + 5x_2 + 7x_3 = 0 \quad ①$$
$$-x_1 - 2x_2 - x_3 = 0 \quad ②$$
$$-3x_1 - 3x_2 - 6x_3 = 0 \quad ③$$

① ＋4 x②

$$-3x_2 + 3x_3 = 0$$
$$x_2 = x_3$$
$$-3x_1 - 9x_2 = 0$$
$$x_1 = -3x_2$$

固有ベクトルは

$$t\begin{pmatrix} 3 \\ -1 \\ -1 \end{pmatrix}$$

$\lambda = 3$について、

$$8x_1 + 5x_2 + 7x_3 = 3x_1$$
$$-1x_1 + 2x_2 - x_3 = 3x_2$$
$$-3x_1 - 3x_2 - 2x_3 = 3x_3$$

の連立方程式を得る。

$$5x_1 + 5x_2 + 7x_3 = 0$$
$$-x_1 - x_2 - x_3 = 0$$
$$-3x_1 - 3x_2 - 5x_3 = 0$$
$$x_3 = 0$$
$$x_1 = -x_2$$

固有ベクトルは

$$t\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$$

$\lambda = 1$について、

$$8x_1 + 5x_2 + 7x_3 = x_1$$
$$-1x_1 + 2x_2 - x_3 = x_2$$
$$-3x_1 - 3x_2 - 2x_3 = x_3$$

という連立方程式を得る。

$$7x_1 + 5x_2 + 7x_3 = 0$$
$$-x_1 + x_2 - x_3 = 0$$
$$-3x_1 - 3x_2 - 3x_3 = 0$$
$$x_2 = 0$$

$$x_1 = -x_3$$

固有ベクトルは

$$t\begin{pmatrix}1\\0\\-1\end{pmatrix}$$

以上で、それぞれの固有ベクトルが求まりました。さてここで、D の固有ベクトルをそれぞれ

$$\boldsymbol{P_1} = (1 \quad 0 \quad 0)$$
$$\boldsymbol{P_2} = (0 \quad 1 \quad 0)$$
$$\boldsymbol{P_3} = (0 \quad 0 \quad 1)$$

C の固有ベクトルをそれぞれ

$$\boldsymbol{P_1}' = (1 \quad -1 \quad 0)$$
$$\boldsymbol{P_2}' = (3 \quad -1 \quad -1)$$
$$\boldsymbol{P_3}' = (-1 \quad 0 \quad 1)$$

（$\boldsymbol{P_3}'$については$\mathrm{t} = -1$とし、他は$\mathrm{t} = 1$としました。数学的な理由はありません。後段の説明の流れを解りやすくするためです。）

としてそれぞれの固有ベクトル間の内積を考えると

$$\boldsymbol{P_1} \cdot \boldsymbol{P_2} = \boldsymbol{P_2} \cdot \boldsymbol{P_3} = \boldsymbol{P_3} \cdot \boldsymbol{P_1} = 0$$

であり、それぞれの固有ベクトルは互いに直交しています。しかし

$$\boldsymbol{P_1}' \cdot \boldsymbol{P_2}' \neq 0, \boldsymbol{P_2}' \cdot \boldsymbol{P_3}' \neq 0, \boldsymbol{P_3}' \cdot \boldsymbol{P_1}' \neq 0$$

であり、D の固有ベクトルは互いに直交していません。

つまり、固有ベクトルがつくる角度は、相似な行列間で等しくないということがわかります。

次に、固有値（固有ベクトル方向のベクトルの長さがどのくらいに拡大されるか、固有ベクトル上の２点の距離がどのくらい拡大されるか）について計算してみます。

上記の例に挙げた、

$$\boldsymbol{D} = \begin{pmatrix}3 & 0 & 0\\0 & 4 & 0\\0 & 0 & 1\end{pmatrix}$$

について

$$\boldsymbol{D}\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix} = \begin{pmatrix}3 & 0 & 0\\0 & 4 & 0\\0 & 0 & 1\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}$$

という変換を考えます。

この変換の固有値$\lambda = 3, \lambda = 4, \lambda = 1$であり、それぞれの固有値に対応する固有ベクトルが、$\begin{pmatrix}t\\0\\0\end{pmatrix}, \begin{pmatrix}0\\t\\0\end{pmatrix}, \begin{pmatrix}0\\0\\t\end{pmatrix}$

だということは、$x_1 - x_2 - x_3$の座標軸で表される空間上の 2 点 A:$\begin{pmatrix} x_{1A} \\ x_{2A} \\ x_{3A} \end{pmatrix}$、B:$\begin{pmatrix} x_{1B} \\ x_{2B} \\ x_{3B} \end{pmatrix}$間のベクトル$\overrightarrow{AB}$が固有ベクトルの条件を満たせば（簡単に言えば、ベクトルの方向軸が一致すれば）、返還後に移された 2 点 ***A'***:$\begin{pmatrix} x_{1A'} \\ x_{2A'} \\ x_{3A'} \end{pmatrix}$、***B'***:$\begin{pmatrix} x_{1B'} \\ x_{2B'} \\ x_{3B'} \end{pmatrix}$のベクトルの長さ（２点間の距離）

$\left|\overrightarrow{A'B'}\right|$は、$\left|\overrightarrow{AB}\right|$はλ倍となる。式で書くと以下の通り

$$\frac{\left|\overrightarrow{A'B'}\right|}{\left|\overrightarrow{AB}\right|} = \lambda$$

$A: (1 \quad 0 \quad 0), B: (2 \quad 0 \quad 0)$を考える。

$$\left|\overrightarrow{AB}\right| = 1$$

ベクトル$\overrightarrow{AB} = (2-1 \quad 0-0 \quad 0-0) = (1 \quad 0 \quad 0)$

であり、このベクトルは固有ベクトル$(t \quad 0 \quad 0)$です。

$$\overrightarrow{A'B'} = \mathrm{D}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

$$= \begin{pmatrix} 3 \\ 0 \\ 0 \end{pmatrix}$$

$$\left|\overrightarrow{A'B'}\right| = 3$$

$$\frac{\left|\overrightarrow{A'B'}\right|}{\left|\overrightarrow{AB}\right|} = 3$$

これを一般化すれば

$A: (a \quad 0 \quad 0), B: (b \quad 0 \quad 0)$ (b>a)について。

$$\overrightarrow{AB} = (b-a \quad 0 \quad 0)$$

$$\left|\overrightarrow{AB}\right| = \mathrm{b} - \mathrm{a}$$

$$\overrightarrow{A'B'} = \mathbf{D}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} b-a \\ 0 \\ 0 \end{pmatrix}$$

$$= \begin{pmatrix} 3(b-a) \\ 0 \\ 0 \end{pmatrix}$$

$$\left|\overrightarrow{A'B'}\right| = 3(\mathrm{b} - \mathrm{a})$$

$$\frac{\left|\overrightarrow{A'B'}\right|}{\left|\overrightarrow{AB}\right|} = 3$$

固有値$\lambda = 4, \lambda = 1$と、それぞれの固有値に対応する固有ベクトル、$\begin{pmatrix} 0 \\ t \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ t \end{pmatrix}$についても、同様にベクトルの絶対値の変化率が固有値となることはわざわざ取り上げて計算するまで

もないでしょう。
次に、

$$\boldsymbol{C}=\begin{pmatrix}8 & 5 & 7\\ -1 & 2 & -1\\ -3 & -3 & -2\end{pmatrix}$$

について

$$\boldsymbol{C}\begin{pmatrix}x_1\\ x_2\\ x_3\end{pmatrix}=\begin{pmatrix}8 & 5 & 7\\ -1 & 2 & -1\\ -3 & -3 & -2\end{pmatrix}\begin{pmatrix}x_1\\ x_2\\ x_3\end{pmatrix}$$

という変換を考えます。
　この変換の固有値$\lambda=3,\lambda=4,\lambda=1$であり、それぞれの固有値に対応する固有ベクトル

は、$\begin{pmatrix}t\\ -t\\ 0\end{pmatrix},\begin{pmatrix}3t\\ -t\\ -t\end{pmatrix},\begin{pmatrix}t\\ 0\\ -t\end{pmatrix}$

まず、空間上で、$\overrightarrow{AB}=\begin{pmatrix}t\\ -t\\ 0\end{pmatrix}$となる。二点 A,B を考えます。
たとえば、点 A:$(1 \quad -1 \quad 0)$と点 A:$(2 \quad -2 \quad 0)$は

$$\overrightarrow{AB}=\begin{pmatrix}1\\ -1\\ 0\end{pmatrix}$$

なので、$\overrightarrow{AB}$は固有ベクトルです。
一般的には、点 A:$(a \quad -a \quad 0)$と点 A:$(b \quad -b \quad 0), b>a$について

$$\overrightarrow{AB}=\begin{pmatrix}b-a\\ a-b\\ 0\end{pmatrix}$$

となり、固有ベクトルです。

$$|\overrightarrow{AB}|=\sqrt{(b-a)^2+(a-b)^2+0^2}=\sqrt{2}(b-a)$$

変換後の点 A”,B”を考えると

$$\overrightarrow{A''B''}=\mathrm{C}\begin{pmatrix}b-a\\ a-b\\ 0\end{pmatrix}=\begin{pmatrix}8 & 5 & 7\\ -1 & 2 & -1\\ -3 & -3 & -2\end{pmatrix}\begin{pmatrix}b-a\\ a-b\\ 0\end{pmatrix}$$

$$=\begin{pmatrix}8(b-a)+5(a-b)\\ -(b-a)+2(a-b)\\ -3(b-a)-3(a-b)\end{pmatrix}$$

$$=\begin{pmatrix}3(b-a)\\ -3(b-a)\\ 0\end{pmatrix}$$

$$\left|\overrightarrow{A''B''}\right|=\sqrt{\{3(b-a)\}^2+\{-3(b-a)\}^2}=3\sqrt{2}(b-a)$$

$$\frac{\left|\overrightarrow{A''B''}\right|}{\left|\overrightarrow{AB}\right|}=3$$

固有値$\lambda=4,\lambda=1$と、それぞれの固有値に対応する固有ベクトル、$\begin{pmatrix}3t\\ -t\\ -t\end{pmatrix},\begin{pmatrix}t\\ 0\\ -t\end{pmatrix}$について

も、同様に計算できます。固有値　$\lambda = 4$、固有ベクトル$\begin{pmatrix} 3t \\ -t \\ -t \end{pmatrix}$については、若干計算が複雑になるので、念のために計算例を示します。

点 A、B 間のベクトルは

$$\overrightarrow{AB} = \begin{pmatrix} 3(b-a) \\ a-b \\ a-b \end{pmatrix}$$

$$\left|\overrightarrow{AB}\right| = \sqrt{\{3(b-a)\}^2 + (a-b)^2 + (a-b)^2} = \sqrt{11}(b-a)$$

$$\overrightarrow{A''B''} = \mathrm{C}\begin{pmatrix} 3(b-a) \\ a-b \\ a-b \end{pmatrix} = \begin{pmatrix} 8 & 5 & 7 \\ -1 & 2 & -1 \\ -3 & -3 & -2 \end{pmatrix}\begin{pmatrix} 3(b-a) \\ a-b \\ a-b \end{pmatrix}$$

$$= \begin{pmatrix} 8 \times \{3(b-a)\} + 5 \times (a-b) + 7 \times (a-b) \\ -1 \times \{3(b-a)\} + 2 \times (a-b) - 1 \times (a-b) \\ -3 \times \{3(b-a)\} - 3 \times (a-b) - 2 \times (a-b) \end{pmatrix}$$

$$= \begin{pmatrix} 12(b-a) \\ 4(a-b) \\ 4(a-b) \end{pmatrix}$$

$$\left|\overrightarrow{A''B''}\right| = \sqrt{\{12(b-a)\}^2 + \{4(a-b)\}^2 + \{4(a-b)\}^2}$$

$$= 4\sqrt{\{3(b-a)\}^2 + (a-b)^2 + (a-b)^2} = 4\sqrt{11}(b-a)$$

$$\frac{\left|\overrightarrow{A''B''}\right|}{\left|\overrightarrow{AB}\right|} = 4$$

つまり、C と D を変換と考えると、固有ベクトルは違っていますが、固有値は同じなのです。相似な行列はいくらでも考えられますが、これを使って、行列で表されたデータを、分析しやすい都合の良い行列に変換できるということです。これが、相似の実用的な意味です。どうせ変換するなら、直交的な空間に投影したいと思うでしょう。それが次に説明する対角化です。普通、幾何学では、ある空間に描かれる立体の構成要素の角度が同じで、長さが異なる場合に、相似だと言いますが、行列では、固有ベクトルの方向が違っていて、固有値（孫方向への拡大倍率）が同じ場合に相似です。

対角化

「行列の相似」という概念を説明したのは、それを用いた「対角化」という技術を紹介するためです。それを説明する前に、少し、相似の説明を変更しておきます。

筆者がここで「相似」について行った説明のしかたは、あまり一般的ではないのです（一般にある説明とは、説明の方向が逆です。）。そこで、良くある説明の形に変形します

$$\boldsymbol{C} = \boldsymbol{P}^{-1}\boldsymbol{D}\boldsymbol{P}$$

つまり、対角行列である$\boldsymbol{D}$の両辺に左から P、右から P^{-1} をかけると、何か行列ができると言っているだけです。多分それは、多くの場合、対角行列ではない。そんなに大したこと

は言っていません。
これを私たちが考えている、何かの行列の両辺に行列とその逆行列を掛けて、直交かされた行列$\boldsymbol{D}$を作りたいという式に変形します。
その場合、左から$\boldsymbol{P}$、右から$\boldsymbol{P}$をかけます

$$\boldsymbol{PCP^{-1}} = \boldsymbol{PP^{-1}DPP^{-1}}$$

$$\boldsymbol{PCP^{-1}} = \boldsymbol{D}$$

$(\mathbf{P^{-1}})^{-1}$=Pなので

$$\boldsymbol{P^{-1}} = \boldsymbol{Q}$$

とすると

$$\boldsymbol{P} = \boldsymbol{Q^{-1}}$$

$$\mathbf{Q^{-1}CQ} = \boldsymbol{D}$$

これが私たちのやりたいことです。これを対角化といいます。
つまり、「行列 $\boldsymbol{C}$ に相似の行列はいくらでも考えられ、無限個に存在しますが、適切な行列 Q を選び出して、ある行列 $\boldsymbol{C}$に、左から行列 $\boldsymbol{Q}$の逆行列と、右から行列 $\boldsymbol{Q}$をかけて変換すると、固有値を対角因子とする対角行列にする。」。これが対角化の作業です。これを理解するために、その行の原理である相似という概念を導入するという方が、普通の説明なのです。
次のような言い方で、対角化を言い表します。

$$= \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \lambda_p \end{pmatrix}$$

行列の次数としては n を使いたいのですが、n はデーターの数の意味に使うので、以下、行列の次数として p を使いますので、勘違いしないように気を付けてください。いずれにしても、このようになる $\boldsymbol{Q}$が存在するとき、行列$\boldsymbol{C}$は対角化できます。

ここで、$\boldsymbol{Q} = (\boldsymbol{Q_1} \quad \boldsymbol{Q_2} \quad \cdots \quad \boldsymbol{Q_p})$のように、変換に使う行列 $\boldsymbol{Q}$を、縦ベクトル$\boldsymbol{Q_i}$を横に並べたものだと考えます。
$\boldsymbol{Q^{-1}CQ} = \mathbf{D}$という式の両辺に、両辺に左から Q をかけると

$$\boldsymbol{Q}(\boldsymbol{Q^{-1}CQ}) = \boldsymbol{Q}\begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \lambda_p \end{pmatrix}$$

$$\boldsymbol{CQ} = (\boldsymbol{Q_1} \quad \boldsymbol{Q_2} \quad \cdots \quad \boldsymbol{Q_n})\begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \lambda_p \end{pmatrix}$$

$$(\boldsymbol{CQ_1} \quad \boldsymbol{CQ_2} \quad \cdots \quad \boldsymbol{CQ_n}) = (\lambda_1\boldsymbol{Q_1} \; \lambda_2\boldsymbol{Q_2} \cdots \lambda_n\boldsymbol{Q_p})$$

省略して書けば、

$$\mathbf{C}\boldsymbol{Q_i} = \lambda_i \boldsymbol{Q_i}$$

Q_iはベクトルで、固有値、固有ベクトルの定義は

$$\mathrm{A}\vec{x} = \lambda\vec{x}$$

なので、Q_iはλ_iに対応する$\mathbf{C}$の固有ベクトルです。これは重要な発見です。念のために確認します。

$$\boldsymbol{Q} = (\boldsymbol{Q_1} \quad \boldsymbol{Q_2} \quad \cdots \quad \boldsymbol{Q_p})$$

ですから、それぞれのベクトルを順番に並べた行列$\boldsymbol{Q}$ が、直交行列に変換する行列だということですね。試してみます。

$$\boldsymbol{C} = \begin{pmatrix} 8 & 5 & 7 \\ -1 & 2 & -1 \\ -3 & -3 & -2 \end{pmatrix}$$

という先ほど固有ベクトルと固有値を求める計算の例に使った行列を使います。この行列の固有ベクトルは

$$\begin{pmatrix} t \\ -t \\ 0 \end{pmatrix}, \begin{pmatrix} 3t \\ -t \\ -t \end{pmatrix}, \begin{pmatrix} t \\ 0 \\ -t \end{pmatrix}$$

でした。

１番目と３番目のベクトルでは$\mathrm{t} = -1$、最後のベクトルでは$\mathrm{t} = 1$とすれば

$$\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ -1 \\ -1 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$$

したがって、

$$\boldsymbol{Q} = \begin{pmatrix} -1 & 3 & -1 \\ 1 & -1 & 0 \\ 0 & -1 & 1 \end{pmatrix}$$

確かに、私たちが最初に与えた

$$\boldsymbol{P} = \begin{pmatrix} 1 & 2 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 2 \end{pmatrix}$$

の逆関数　$\boldsymbol{P^{-1}} = \begin{pmatrix} -1 & 3 & -1 \\ 1 & -1 & 0 \\ 0 & -1 & 1 \end{pmatrix}$そのものになっています。

いまさら、やってみるまでもありませんが

$$\begin{pmatrix} 1 & 2 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 2 \end{pmatrix}\begin{pmatrix} -1 & 3 & -1 \\ 1 & -1 & 0 \\ 0 & -1 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} -1+2+0 & 3-2-1 & -1+0+1 \\ -1+1+0 & 3-1-1 & -1+0+1 \\ -1+1+0 & 3-1-2 & -1+0+2 \end{pmatrix}$$

$$= \boldsymbol{I}$$

$\boldsymbol{I}$は単位行列の意味

だということは確認できます。確かにそうなっていますが、少し釈然としない、納得がい

かない部分がありませんか。
それぞれの固有値に属する固有ベクトルは

$$\begin{pmatrix} t \\ -t \\ 0 \end{pmatrix}, \begin{pmatrix} 3t \\ -t \\ -t \end{pmatrix}, \begin{pmatrix} t \\ 0 \\ -t \end{pmatrix}$$

で表されて、無限に存在します。 t は、それぞれの固有ベクトルについて勝手に決めて良いという意味ですね。それを使って、1 番目と 3 番目のベクトルでは$\mathrm{t} = -1$、最後のベクトルでは$\mathrm{t} = 1$として、もともとの $\boldsymbol{P}$ と同じになるようにしたのです。 t を適当に選べば、$\boldsymbol{P}$ になることは確かですが、どんな t を選んでも必ず、対角化させる行列になるのか、少し疑わしい気がしませんか、すくなくとも計算上どういうことになっているのか理解しないと不安で使えませんね。
ではやりましょう、この疑問は、固有ベクトルは、本来、次のように書くべきだったと言っているのですね。

$$\begin{pmatrix} -a \\ a \\ 0 \end{pmatrix}, \begin{pmatrix} 3b \\ -b \\ -b \end{pmatrix}, \begin{pmatrix} -c \\ 0 \\ c \end{pmatrix}$$

したがって、

$$\boldsymbol{Q} = \begin{pmatrix} -a & 3b & -c \\ a & -b & 0 \\ 0 & -b & c \end{pmatrix}$$

まさにおっしゃる通りなので、その通りに計算してみましょう

$$\boldsymbol{Q}' = \begin{pmatrix} -a & 3b & -c \\ a & -b & 0 \\ 0 & -b & c \end{pmatrix}' = \begin{pmatrix} -a & a & 0 \\ 3b & -b & -b \\ -c & 0 & c \end{pmatrix}$$

$$|\boldsymbol{Q}| = \begin{vmatrix} -a & 3b & -c \\ a & -b & 0 \\ 0 & -b & c \end{vmatrix} = abc + abc - 3abc = -abc$$

$$\boldsymbol{Q}^{-1} = \frac{-1}{abc} \begin{pmatrix} \begin{vmatrix} -b & -b \\ 0 & c \end{vmatrix} & -\begin{vmatrix} 3b & -b \\ -c & c \end{vmatrix} & \begin{vmatrix} 3b & -b \\ -c & 0 \end{vmatrix} \\ -\begin{vmatrix} a & 0 \\ 0 & c \end{vmatrix} & \begin{vmatrix} -a & 0 \\ -c & c \end{vmatrix} & -\begin{vmatrix} -a & a \\ -c & 0 \end{vmatrix} \\ \begin{vmatrix} a & 0 \\ -b & -b \end{vmatrix} & -\begin{vmatrix} -a & 0 \\ 3b & -b \end{vmatrix} & \begin{vmatrix} -a & a \\ 3b & -b \end{vmatrix} \end{pmatrix}$$

$$= \frac{-1}{abc} \begin{pmatrix} -bc & -2bc & -bc \\ -ac & -ac & -ac \\ -ab & -ab & -2ab \end{pmatrix}'$$

$$= \begin{pmatrix} \frac{1}{a} & \frac{2}{a} & \frac{1}{a} \\ \frac{1}{b} & \frac{1}{b} & \frac{1}{b} \\ \frac{1}{c} & \frac{1}{c} & \frac{2}{c} \end{pmatrix}$$

$\boldsymbol{Q}^{-1}\boldsymbol{C}\boldsymbol{Q}$を計算します。

$$\boldsymbol{Q}^{-1}\boldsymbol{C}\boldsymbol{Q} = \begin{pmatrix} \frac{1}{a} & \frac{2}{a} & \frac{1}{a} \\ \frac{1}{b} & \frac{1}{b} & \frac{1}{b} \\ \frac{1}{c} & \frac{1}{c} & \frac{2}{c} \end{pmatrix} \begin{pmatrix} 8 & 5 & 7 \\ -1 & 2 & -1 \\ -3 & -3 & -2 \end{pmatrix} \begin{pmatrix} a & 3b & c \\ -a & -b & 0 \\ 0 & -b & -c \end{pmatrix}$$

$$\boldsymbol{Q}^{-1}\boldsymbol{C} = \begin{pmatrix} \frac{1}{a} & \frac{2}{a} & \frac{1}{a} \\ \frac{1}{b} & \frac{1}{b} & \frac{1}{b} \\ \frac{1}{c} & \frac{1}{c} & \frac{2}{c} \end{pmatrix} \begin{pmatrix} 8 & 5 & 7 \\ -1 & 2 & -1 \\ -3 & -3 & -2 \end{pmatrix}$$

$$= \begin{pmatrix} \frac{1}{a}(8-2-3) & \frac{1}{a}(5+4-3) & \frac{1}{a}(7-2-2) \\ \frac{1}{b}(8-1-3) & \frac{1}{b}(5+2-3) & \frac{1}{b}(7-1-2) \\ \frac{1}{c}(8-1-6) & \frac{1}{c}(5+2-6) & \frac{1}{c}(7-1-4) \end{pmatrix}$$

$$= \begin{pmatrix} \frac{3}{a} & \frac{6}{a} & \frac{3}{a} \\ \frac{4}{b} & \frac{4}{b} & \frac{4}{b} \\ \frac{1}{c} & \frac{1}{c} & \frac{2}{c} \end{pmatrix}$$

$$(\boldsymbol{Q}^{-1}\boldsymbol{C})\boldsymbol{Q} = \begin{pmatrix} \frac{3}{a} & \frac{6}{a} & \frac{3}{a} \\ \frac{4}{b} & \frac{4}{b} & \frac{4}{b} \\ \frac{1}{c} & \frac{1}{c} & \frac{2}{c} \end{pmatrix} \begin{pmatrix} a & 3b & c \\ -a & -b & 0 \\ 0 & -b & -c \end{pmatrix}$$

$$= \begin{pmatrix} -3+6 & \frac{b}{a}(-9-6-3) & \frac{c}{a}(-3+0+3) \\ \frac{a}{b}(4-4+0) & (12-4-4) & \frac{c}{b}(4-4) \\ \frac{a}{c}(1-1+0) & \frac{b}{c}(3-1-2) & (1+0-2) \end{pmatrix}$$

$$= \begin{pmatrix} 3 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

このように具体的に計算してみるとわかるように

a, b, cは何を選んでも計算の過程で消えてしまうので、固有ベクトルであれば、何を選んでも、対角化するための行列になります。

p 個の固有値を持つ p 次の行列は対角化できることわかりました。その方法も理解できました。

この発見の応用例が多変量解析です。多変量解析は、多項目のデータをいくつかの項目に集約して表現し、現象を単純化して把握したいという目的に使われるのが普通だろうと思います。その方法として、直交化は大いに裕子だろうと想像がつきます。

その目的のために、まず最も単純に考えそうなことは、次のようなことでしょう。
多変量的なデータが集約されて、ある行列として書かれている。この行列を、いくつかの独立した（ということは直交した）ベクトル上の点を重ね合わせたものとしてとらえられないか。
言いたいことを鮮明にするために、例を上げます。

$$\mathrm{D}=\begin{pmatrix}3&0&0\\0&4&0\\0&0&1\end{pmatrix}$$

については、

$$\mathrm{D}=\begin{pmatrix}3&0&0\\0&4&0\\0&0&1\end{pmatrix}=3\begin{pmatrix}1&0&0\\0&0&0\\0&0&0\end{pmatrix}+4\begin{pmatrix}0&0&0\\0&1&0\\0&0&0\end{pmatrix}+1\begin{pmatrix}0&0&0\\0&0&0\\0&0&1\end{pmatrix}$$

のように、直交する３つのベクトルの和として表現できています。
つまり、左辺の行列の、１行１列の行列の３は、$3\times1+4\times0+1\times0$、２行２列の４は $3\times0+4\times1+1\times0$のように見ようと思えば、そう見えるというわけです。もっと言えば、D が一つのデーターだとして、そのデーターを構成する要素（特性）に下の三つの単位ベクトルの長さで表現できる特性があり、

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

そのそれぞれの特性の値が、それぞれ３、４、１だったので、

$$\mathrm{D}=\begin{pmatrix}3&0&0\\0&4&0\\0&0&1\end{pmatrix}$$

と表現されていると考えるのです。至極当たり前のことですね。
確かに当たり前のことですが、対角行列では、行列が固有ベクトルが行列で表現される空間の座標軸と一致しているから、そういうことができるのです。下の行列では、なかなかそのような捉え方は難しいでしょう。
たとえば

$$\begin{pmatrix}8&5&7\\-1&2&-1\\-3&-3&-2\end{pmatrix}$$

では無理です。
行列をベクトルを並べたものと考えたときに、個々のベクトルを簡単にできますが、残念

ながら、行列全体を、「n 個のベクトルの実数倍の和」としてとらえることが、どんな行列でもできるわけではありません。対角化すれば、対角化された行列は直交ベクトルの和として表すことができますが、対角化された行列は、もちろん元の行列ではないのだから、今のころと、上記の議論に実用的な意味はありません。

しかし、ここでわれわれは、「複雑な物が重なりあってできた行列をいくつかの成分の和に分解することができたら、個々の要素について議論することができるのになあ。」というような、ぼんやりした願望を持つことができます。

対角化された行列を

$$\begin{pmatrix} -2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix} = -2\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + 1\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} + 3\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

このように書いてみると、

$$\boldsymbol{P^{-1}AP} = -2\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + 1\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} + 3\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

となったことがわかります。
ここで、両辺に左から$\boldsymbol{P}$、右から$\boldsymbol{P^{-1}}$をかけます。

$$\boldsymbol{PP^{-1}APP^{-1}} = \boldsymbol{P}\left(-2\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + 1\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} + 3\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}\right)\boldsymbol{P^{-1}}$$

$$\boldsymbol{PP^{-1}} = \boldsymbol{PP^{-1}} = \boldsymbol{I}$$

で、行列計算でも分配法則は成り立つから

$$\boldsymbol{A} = -2\boldsymbol{P}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}\boldsymbol{P^{-1}} + 1\boldsymbol{P}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}\boldsymbol{P^{-1}} + 3\boldsymbol{P}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}\boldsymbol{P^{-1}}$$

$$= -2\begin{pmatrix} 11 & -1 & 1 \\ 1 & 1 & 1 \\ -14 & 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}\boldsymbol{P^{-1}} + \begin{pmatrix} 11 & -1 & 1 \\ 1 & 1 & 1 \\ -14 & 1 & 1 \end{pmatrix}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}\boldsymbol{P^{-1}}$$

$$+3\begin{pmatrix} 11 & -1 & 1 \\ 1 & 1 & 1 \\ -14 & 1 & 1 \end{pmatrix}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}\boldsymbol{P^{-1}}$$

$$= -2\begin{pmatrix} 11 & 0 & 0 \\ 1 & 0 & 0 \\ -14 & 0 & 0 \end{pmatrix}\boldsymbol{P^{-1}} + \begin{pmatrix} 0 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{pmatrix}\boldsymbol{P^{-1}} + 3\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix}\boldsymbol{P^{-1}}$$

となるので、さらに計算を進めると

$$= -2 \times \frac{1}{30}\begin{pmatrix} 11 & 0 & 0 \\ 1 & 0 & 0 \\ -14 & 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 2 & -2 \\ -15 & 25 & -10 \\ 15 & 3 & 12 \end{pmatrix} + \frac{1}{30}\begin{pmatrix} 0 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & 2 & -2 \\ -15 & 25 & -10 \\ 15 & 3 & 12 \end{pmatrix}$$

$$+3 \times \frac{1}{30}\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 0 & 2 & -2 \\ -15 & 25 & -10 \\ 15 & 3 & 12 \end{pmatrix}$$

$$= -2 \times \frac{1}{30}\begin{pmatrix} 0 & 22 & -22 \\ 0 & 2 & -2 \\ 0 & -28 & 28 \end{pmatrix} + \frac{1}{30}\begin{pmatrix} 15 & -25 & 10 \\ -15 & 25 & -10 \\ -15 & 25 & -10 \end{pmatrix} + 3 \times \frac{1}{10}\begin{pmatrix} 15 & 3 & 12 \\ 15 & 3 & 12 \\ 15 & 3 & 12 \end{pmatrix}$$

$$= \begin{pmatrix} 0 + \frac{1}{2} + \frac{3}{2} & \frac{-22}{15} + \frac{-5}{6} + \frac{3}{10} & \frac{22}{15} + \frac{1}{3} + \frac{6}{5} \\ 0 - \frac{1}{2} + \frac{3}{2} & -\frac{2}{15} + \frac{5}{6} + \frac{3}{10} & \frac{2}{15} - \frac{1}{3} + \frac{6}{5} \\ 0 - \frac{1}{2} + \frac{3}{2} & \frac{28}{15} + \frac{5}{6} + \frac{3}{10} & -\frac{28}{15} - \frac{1}{3} + \frac{6}{5} \end{pmatrix}$$

$$= \begin{pmatrix} 2 & -2 & 3 \\ 1 & 1 & 1 \\ 1 & 3 & -1 \end{pmatrix}$$

となります。計算式の通り計算したら、正しい答えになったというだけのことです。式の変形に間違いがないことを計算して確認したに過ぎないと言われそうですが、見せたかったのは、

$$\boldsymbol{A} = -2\boldsymbol{P}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}\boldsymbol{P^{-1}} + 1\boldsymbol{P}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}\boldsymbol{P^{-1}} + 3\boldsymbol{P}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}\boldsymbol{P^{-1}}$$

というところです。

つまり、

$$\boldsymbol{PA}\mathbf{P^{-1}} = \boldsymbol{\Lambda} = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

のように対角化できるとき

$$\boldsymbol{A} = \lambda_1\boldsymbol{P}\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}\boldsymbol{P^{-1}} + \lambda_2\boldsymbol{P}\begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}\boldsymbol{P^{-1}} + \cdots + \lambda_n\boldsymbol{P}\begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}\boldsymbol{P^{-1}}$$

と表せるということです。

でもこれは、少し格好が悪いですね。代数的に計算を解りやすくするためにわざわざこのように書いたのですが、行列をベクトルを横に並べたものだと解釈したときに、$\boldsymbol{0}$ ベクトルに意味はありませんから、もちろん、

$$\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

のように書くべきです。ということで、

$$\boldsymbol{A} = \lambda_1\boldsymbol{P}\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}\boldsymbol{P^{-1}} + \lambda_2\boldsymbol{P}\begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}\boldsymbol{P^{-1}} + \cdots + \lambda_n\boldsymbol{P}\begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}\boldsymbol{P^{-1}}$$

ではなくて、

$$A=\lambda_1 P\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix}P^{-1}+\lambda_2 P\begin{pmatrix}0\\1\\\vdots\\0\end{pmatrix}P^{-1}+\cdots+\lambda_{\mathrm{n}} P\begin{pmatrix}0\\0\\\vdots\\1\end{pmatrix}P^{-1}$$

と書くのが、お作法というものです。

こうして書き換えたうえで、この作業が持っている幾何学的な意味を考えます。

普通の教科書では、一次独立がどうしたとか、線形結合がどうしたとか、基底がどうしたとか、難しそうな言葉をつかった解説があります。多分かなりわかりきった当たり前のことを言っているのだと思うのですが、私は独学だから、それらが正確にどんな意味なのかを知りません。そういう難しい言葉がつかった説明はできませんが、それらは、おそらく、次のようなことを言っているのだと思います。

$\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix}$、$\begin{pmatrix}0\\1\\\vdots\\0\end{pmatrix}$、…、$\begin{pmatrix}0\\0\\\vdots\\1\end{pmatrix}$というのは、$\boldsymbol{PA}\mathbf{P}^{-1}$という計算の結果、移し替えられた（写されたという表現のほうが良いかもしれません。射影という言葉がありますから。）行列が存在する空間を決めている座標系の一つ一つの座標軸で単位となっているベクトルのことですね。とりあえず、$\boldsymbol{A}$は基底となっている p 個のベクトルのスカラー倍を p 個並べたものという感じで、$\boldsymbol{P}$や$\mathbf{P}^{-1}$は移し替える(写す)操作という感じです。

$$\boldsymbol{PA}\mathbf{P}^{-1}=\begin{pmatrix}\lambda_1 & 0 & \cdots & 0\\0 & \lambda_2 & \cdots & 0\\\vdots & \vdots & \ddots & \vdots\\0 & 0 & \cdots & \lambda_p\end{pmatrix}$$

この移し替えは、相似なのだから、固有値が変わらずに固有ベクトル同士の角度が変わるだけで、その結果、「右辺は、①固有ベクトルがその空間の次元数の p だけあって、②その一つ一つの固有ベクトルが直交していて、③それらのベクトルの絶対値が元の行列の固有値の値になっている。そのようなベクトルを並べた行列になる。」ということが、この式の意味です。定義を良く知りませんが、基底という言葉は空間を決めている単位ベクトルの集まりのことだと思います。対角化の結果できた行列では、並べた p 個のベクトルはすべて互いに直交しています。これらのベクトルの単位長さのベクトル（単位ベクトル）に固有値を乗じたものが、対角化された行列です。この空間上のベクトル（あるいは点）は、すべて、基底をなすベクトルの実数倍をたし合わせて表すことができることが直感的にわかります。しかし、落ち着いて考えればベクトルの基底同士が直角になっていなくても、ｎ次元の空間上の点を、ｎ個のベクトルの和で表すことはできます。二次元平面で表すと図のようなことです。

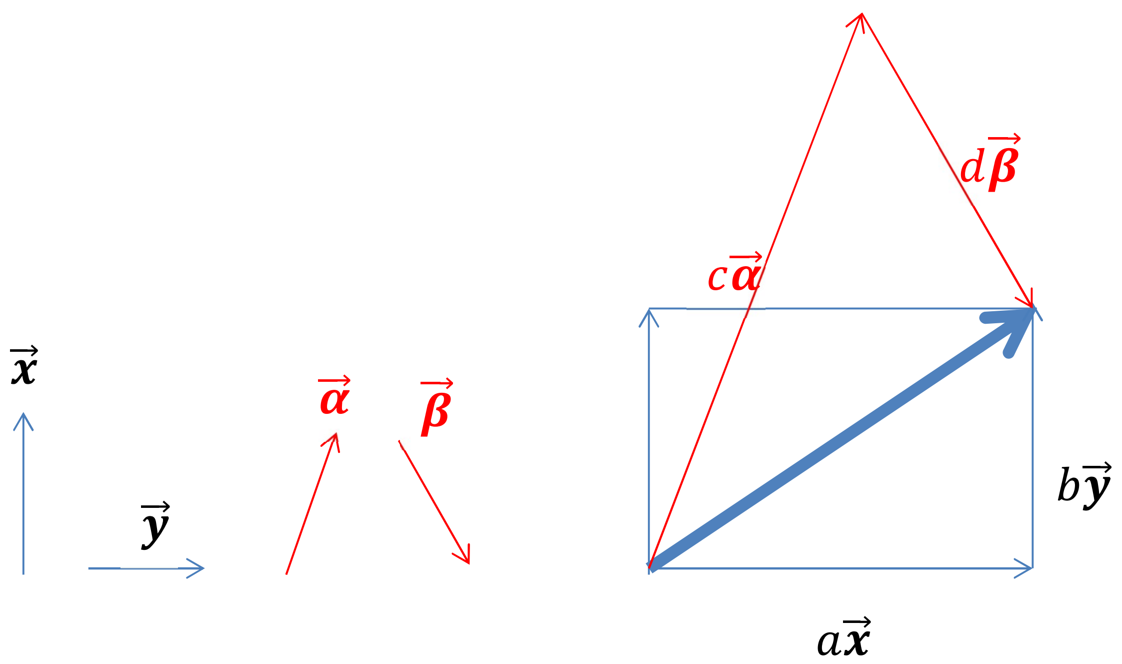

図 4.二次平面での対角化

数式で表現すれば、

$$\vec{\mathbf{A}} = a\vec{x} + b\vec{y} = c\vec{\alpha} + d\vec{\beta}$$

つまり、黒字の式で表すことも赤字の式で表すこともできるということです。$\vec{x}$と$\vec{y}$は何でもかまわないのですが、$\vec{x}$と$\vec{y}$が決まってしまえば、aとb、一通りの値しかとらないということです。このように、異なるベクトルをスカラー倍して足し合わせることを、線形結合と言います。今説明している平面は２次元の平面です。ベクトル同士が並行している場合は、この方法で表せるのは、そのベクトルのスカラー倍だけです。これではこの平面上の全ての点を表すことはできません。平面上の全ての点をこの方法で表そうとすると、必ず、並行していないベクトルが２本必要です。また、この平面上のすべてのベクトル（というか点）が、この方法で表現できることは直感的にわかるでしょう。こういうことを一次独立と言います。平面でなくて p 次空間でも p 個の一次独立のベクトルの組み合わせで基底をつくれば、すべてのベクトルを線形結合で表せることも理解できると思います。このように簡単なことが、数学の教科書では、とても難しい表現で書かれています。

そういう文句は別として、言いたかったことは、ベクトルを並べた行列$\boldsymbol{A}$に p 個の固有ベクトルがあれば、固有ベクトルの線形結合で、A が置かれている空間の全てのベクトルを表すことができるし、固有ベクトルでなくても、１次独立なベクトルを p 個用意すれば、すべてのベクトルを線形結合で表すことができるということです。

それはそれとして、

$$\boldsymbol{A} = \lambda_1 \boldsymbol{P} \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \boldsymbol{P^{-1}} + \lambda_2 \boldsymbol{P} \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} \boldsymbol{P^{-1}} + \cdots + \lambda_p \boldsymbol{P} \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} \boldsymbol{P^{-1}}$$

のところは、形が良くてすっきりしていて良いのですが、途中の計算は何やら恰好が良くありません。

$$\boldsymbol{A} = -2\boldsymbol{P} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \boldsymbol{P^{-1}} + 1\boldsymbol{P} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \boldsymbol{P^{-1}} + 3\boldsymbol{P} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \boldsymbol{P^{-1}}$$

$$= -2 \begin{pmatrix} 11 & -1 & 1 \\ 1 & 1 & 1 \\ -14 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \boldsymbol{P^{-1}} + \begin{pmatrix} 11 & -1 & 1 \\ 1 & 1 & 1 \\ -14 & 1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \boldsymbol{P^{-1}}$$

$$+ 3 \begin{pmatrix} 11 & -1 & 1 \\ 1 & 1 & 1 \\ -14 & 1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \boldsymbol{P^{-1}}$$

$$= -2 \begin{pmatrix} 11 & 0 & 0 \\ 1 & 0 & 0 \\ -14 & 0 & 0 \end{pmatrix} \boldsymbol{P^{-1}} + \begin{pmatrix} 0 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{pmatrix} \boldsymbol{P^{-1}} + 3 \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix} \boldsymbol{P^{-1}}$$

ここのところがカッコ悪いと思いませんか。

$$= -2 \begin{pmatrix} 11 & 0 & 0 \\ 1 & 0 & 0 \\ -14 & 0 & 0 \end{pmatrix} \boldsymbol{P^{-1}} + \begin{pmatrix} 0 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{pmatrix} \boldsymbol{P^{-1}} + 3 \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix} \boldsymbol{P^{-1}}$$

は

$$-2 \begin{pmatrix} 11 \\ 1 \\ -14 \end{pmatrix} \boldsymbol{P^{-1}} + \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} \boldsymbol{P^{-1}} + 3 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \boldsymbol{P^{-1}}$$

と書けるのだから、これはこれでスッキリしているから問題ないという感性もあるでしょうが、

$$\begin{pmatrix} 11 \\ 1 \\ -14 \end{pmatrix} \boldsymbol{P^{-1}}$$

のところで式の対称性が壊れています。

$\begin{pmatrix} 11 \\ 1 \\ -14 \end{pmatrix}$はベクトルを並べた行列$\boldsymbol{P} = \begin{pmatrix} 11 & -1 & 1 \\ 1 & 1 & 1 \\ -14 & 1 & 1 \end{pmatrix}$の一つのベクトルです。これに対して$\boldsymbol{P^{-1}}$は、$\boldsymbol{P}$の逆行列です。

せっかく$\lambda_1 \boldsymbol{P} \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \boldsymbol{P^{-1}}$のように、極めて美しいシンメトリーで始まったのに何か残念です。

だから、たとえば、

$$\begin{pmatrix} 11 \\ 1 \\ -14 \end{pmatrix} \begin{pmatrix} 11 \\ 1 \\ -14 \end{pmatrix}^{-1}$$

のように、対称性を取り戻せたらよいのにナーと思います。別に、感覚的にそう思うだけなので、別にどうということはないのですが、少し付き合ってください。しかし、これは

無理ですね。$\begin{pmatrix}11\\1\\-14\end{pmatrix}$は正方行列ではないのだから、$\begin{pmatrix}11\\1\\-14\end{pmatrix}^{-1}$なんてありえません。

$\begin{pmatrix}11\\1\\-14\end{pmatrix}^{T}$

ならばあります。添え字の T は転置行列の意味ですから。

$$\begin{pmatrix}11\\1\\-14\end{pmatrix}\begin{pmatrix}11\\1\\-14\end{pmatrix}^{T}=\begin{pmatrix}11\\1\\-14\end{pmatrix}\begin{pmatrix}11 & 1 & -14\end{pmatrix}=\begin{pmatrix}121 & 11 & -154\\11 & 1 & -14\\-154 & -14 & 196\end{pmatrix}$$

となります。これを行列Ｐにまで広げて

$$\boldsymbol{P}\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix}\boldsymbol{P}^{-1}=\boldsymbol{P}\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix}\boldsymbol{P}^{T}$$

つまり、

$$\boldsymbol{P}^{-1}=\boldsymbol{P}^{T}$$

となっていれば、願望に近づけないかナーと考えます。

行列計算では、式の並び方の順序は入れ替えられませんが、どこから先に計算しても構いません。そこで、

$$\boldsymbol{P}\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix}\boldsymbol{P}^{T}=\boldsymbol{P}\left(\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix}\boldsymbol{P}^{T}\right)$$

と考えて、カッコの中を先に計算します。

たとえば、

$$\boldsymbol{P}=\begin{pmatrix}a & d & g\\b & e & h\\c & f & i\end{pmatrix}$$

の場合、

$$\boldsymbol{P}^{T}=\begin{pmatrix}a & b & c\\d & e & f\\g & h & i\end{pmatrix}$$

$$\begin{pmatrix}1\\0\\0\end{pmatrix}\boldsymbol{P}^{T}=\begin{pmatrix}1\\0\\0\end{pmatrix}\begin{pmatrix}a & b & c\\d & e & f\\g & h & i\end{pmatrix}$$

と考えますが、こんな式はあり得ません。計算以前にむちゃくちゃです。ここはもうやけになって、

$$\begin{pmatrix}1\\0\\0\end{pmatrix}\boldsymbol{P}'=\begin{pmatrix}1 & 0 & 0\\0 & 0 & 0\\0 & 0 & 0\end{pmatrix}\boldsymbol{P}'=\begin{pmatrix}1 & 0 & 0\\0 & 0 & 0\\0 & 0 & 0\end{pmatrix}\begin{pmatrix}a & b & c\\d & e & f\\g & h & i\end{pmatrix}$$

と書いてしまいましょう。

$$\begin{pmatrix}1\\0\\0\end{pmatrix}\boldsymbol{P^T}=\begin{pmatrix}a&b&c\\0&0&0\\0&0&0\end{pmatrix}$$

ここでまた、この行列をベクトルを立てに並べたものだと考えて、０ベクトルなどは書いても意味がないと言い張って。

$$\begin{pmatrix}1\\0\\0\end{pmatrix}\boldsymbol{P^T}=\begin{pmatrix}a&b&c\end{pmatrix}$$

と強引に結論します。

ところで、

$$\boldsymbol{P}=\begin{pmatrix}a&0&0\\b&0&0\\c&0&0\end{pmatrix}$$

だったら、やはり、強引に

$$\boldsymbol{P}=\begin{pmatrix}a\\b\\c\end{pmatrix}$$

と書いて、全体としては、

$$\boldsymbol{A}=\lambda_1\begin{pmatrix}a\\b\\c\end{pmatrix}\begin{pmatrix}a\\b\\c\end{pmatrix}^T+\lambda_2\begin{pmatrix}d\\e\\f\end{pmatrix}\begin{pmatrix}d\\e\\f\end{pmatrix}^T+\lambda_3\begin{pmatrix}g\\h\\i\end{pmatrix}\begin{pmatrix}g\\h\\i\end{pmatrix}^T$$

となって、形式的な対称性が保たれます。

変なことを、考えましたが、ここで言いたいことは

「$\boldsymbol{A}$を対角化する行列$\boldsymbol{P}$があり、$\boldsymbol{P^{-1}}=\boldsymbol{P^T}$と書けるときには、$\boldsymbol{P}$の行ないし列をベクトルとして、

$$\boldsymbol{P}=\begin{pmatrix}\boldsymbol{p_1}&\boldsymbol{p_2}&\cdots&\boldsymbol{p_p}\end{pmatrix}$$

と表せば、

$$\begin{aligned}\boldsymbol{A}&=\lambda_1\boldsymbol{p_1}\boldsymbol{p_1}^{\boldsymbol{T}}+\lambda_2\boldsymbol{p_2}\boldsymbol{p_2}^{\boldsymbol{T}}+\cdots+\lambda_p\boldsymbol{p_p}\boldsymbol{p_p}^{\boldsymbol{T}}\\&=\sum_{i=1}^{p}\lambda_i\boldsymbol{p_i}\,\boldsymbol{p_i}^{\boldsymbol{T}}\end{aligned}$$

表すことができる。」ということです。そうすると、

$$\boldsymbol{P^{-1}}=\boldsymbol{P^T}$$

が成り立つことがあるのか、あるとすればどれはどんな時なのかを考えなくてはなりません。

つまり、下記の様な式が、

$$\boldsymbol{x^T}\boldsymbol{A}\boldsymbol{x}=\boldsymbol{e}\boldsymbol{\Lambda}\boldsymbol{e^T}$$

$\boldsymbol{\Lambda}$は対角化された行列

$$\boldsymbol{\Lambda}=\begin{pmatrix}\lambda_1 & 0 & \cdots & 0\\ 0 & \lambda_2 & \cdots & 0\\ 0 & 0 & \ddots & 0\\ 0 & 0 & 0 & \lambda_p\end{pmatrix}$$

$\boldsymbol{x}$にかかわらず成り立つ（恒等的に成り立つ、未知数が何であっても成り立つ）ことがあるのか、あるのならばそれはどのような$\boldsymbol{A}$なのかということです。

とりあえず、恒等的に成り立つのだから$\boldsymbol{x}=\boldsymbol{I}$（単位行列）にしてみます

$$\boldsymbol{I}^{\boldsymbol{T}}=\boldsymbol{I}$$

$$\boldsymbol{IA}=\boldsymbol{AI}^{\boldsymbol{T}}=\boldsymbol{A}$$

だから

$$\boldsymbol{x}^{\boldsymbol{T}}\boldsymbol{Ax}=\boldsymbol{e\Lambda e}^{\boldsymbol{T}}$$

の式は

$$\boldsymbol{A}=\boldsymbol{e\Lambda e}^{\boldsymbol{T}}$$

となります。$\boldsymbol{\Lambda}$

ということなのですが、式の意味を分解して説明します。

$$\boldsymbol{\Lambda}=\begin{pmatrix}\lambda_1 & 0 & \cdots & 0\\ 0 & \lambda_2 & \cdots & 0\\ 0 & 0 & \ddots & 0\\ 0 & 0 & 0 & \lambda_p\end{pmatrix}$$

でしたね。

$$\boldsymbol{e\Lambda e}^{\boldsymbol{T}}=\begin{pmatrix}\boldsymbol{e_1}\\ \boldsymbol{e_2}\\ \vdots\\ \boldsymbol{e_p}\end{pmatrix}\begin{pmatrix}\lambda_1 & 0 & \cdots & 0\\ 0 & \lambda_2 & \cdots & 0\\ 0 & 0 & \ddots & 0\\ 0 & 0 & 0 & \lambda_p\end{pmatrix}\begin{pmatrix}\boldsymbol{e_1}^{\boldsymbol{T}} & \boldsymbol{e_2}^{\boldsymbol{T}} & \cdots & \boldsymbol{e_p}^{\boldsymbol{T}}\end{pmatrix}$$

つまり

$$\boldsymbol{A}=\lambda_1\boldsymbol{e_1}\boldsymbol{e_1}^{\boldsymbol{T}}+\lambda_2\boldsymbol{e_2}\boldsymbol{e_2}^{\boldsymbol{T}}+\cdots+\lambda_p\boldsymbol{e_p}\boldsymbol{e_p}^{\boldsymbol{T}}$$

こうならないかナーという願望です。言い換えると

$$\boldsymbol{P}^{\boldsymbol{T}}=\boldsymbol{P}^{-\boldsymbol{1}}$$

になる、行列を探すという話です。

$\boldsymbol{A}$から$\boldsymbol{\Lambda}$への対角化を可能にする変換である行列 $\boldsymbol{P}$とその逆行列 $\boldsymbol{P}^{-1}$があるとします。

$$\boldsymbol{PP}^{-\boldsymbol{1}}=\boldsymbol{I}$$

だから、式のどこに$\boldsymbol{PP}^{-\boldsymbol{1}}$を挿入しても式は変わりません。

$$\boldsymbol{x}^{\boldsymbol{T}}\boldsymbol{Ax}$$

$$=\boldsymbol{x}^{\boldsymbol{T}}\boldsymbol{PP}^{-\boldsymbol{1}}\boldsymbol{APP}^{-\boldsymbol{1}}\boldsymbol{x}$$

行列演算では式の順序は変えられない（交換法則が成り立たない）が、計算の順番は変えられるので

$$=\ (\boldsymbol{x}^{\boldsymbol{T}}\boldsymbol{P})(\boldsymbol{P}^{-\boldsymbol{1}}\boldsymbol{AP})(\boldsymbol{P}^{-\boldsymbol{1}}\boldsymbol{x})$$

という手順で計算することにします。（やっていることは、なんとか$(\boldsymbol{P}^{-\boldsymbol{1}}\boldsymbol{AP})=\boldsymbol{\Lambda}$という形を作るということです。）

結果

$$\boldsymbol{x'Ax}$$
$$= (\boldsymbol{x^T P})\boldsymbol{\Lambda}(\boldsymbol{P^{-1}x})$$

となりますが、さらに変形して

$$\boldsymbol{x'Ax}$$
$$= (\boldsymbol{p^T x})^{\boldsymbol{T}}\boldsymbol{\Lambda}(\boldsymbol{P^{-1}x})$$

$(\boldsymbol{x^T P}) \to (\boldsymbol{p^T x})^{\boldsymbol{T}}$というところは直感的に理解できますか、まあそんなもんだろうと直感的に理解する人と、きちんと確認したいという人と、個性の違いがありますね。確認します。

たとえば、$(a \quad b \quad c)$と$(\alpha \quad \beta \quad \gamma)$という行列があったとします。

$(a \quad b \quad c)'$=$\begin{pmatrix} a \\ b \\ c \end{pmatrix}$, $(\alpha \quad \beta \quad \gamma)' = \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix}$

$$(a \quad b \quad c)'(\alpha \quad \beta \quad \gamma) = \begin{pmatrix} a \\ b \\ c \end{pmatrix}(\alpha \quad \beta \quad \gamma) = \begin{pmatrix} a\alpha & a\beta & a\gamma \\ b\alpha & b\beta & b\gamma \\ c\alpha & c\beta & c\gamma \end{pmatrix}$$

なので、

$$\left((a \quad b \quad c)'(\alpha \quad \beta \quad \gamma)\right)' = \begin{pmatrix} a\alpha & a\beta & a\gamma \\ b\alpha & b\beta & b\gamma \\ c\alpha & c\beta & c\gamma \end{pmatrix}' = \begin{pmatrix} a\alpha & b\alpha & c\alpha \\ a\beta & b\beta & c\beta \\ a\gamma & b\gamma & c\gamma \end{pmatrix}$$

$$(\alpha \quad \beta \quad \gamma)'(a \quad b \quad c) = \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix}(a \quad b \quad c) = \begin{pmatrix} a\alpha & b\alpha & c\alpha \\ a\beta & b\beta & c\beta \\ a\gamma & b\gamma & c\gamma \end{pmatrix}$$

だから

$$\left((a \quad b \quad c)'(\alpha \quad \beta \quad \gamma)\right)' = (\alpha \quad \beta \quad \gamma)'(a \quad b \quad c)$$

証明というよりは、実際にやってみると感覚的にわかりますね。

もとにもどると、つまり、$\boldsymbol{x^T A x} = \boldsymbol{e \Lambda e^T}$が成り立つには

$$\boldsymbol{e} = (\boldsymbol{P^T x})^{\boldsymbol{T}} \qquad ①$$
$$\boldsymbol{e^T} = \boldsymbol{P^{-1}x} \qquad ②$$

が同時に成り立てば良いということになります。

① の式を変形します。左辺右辺の転置行列をとって

$$\boldsymbol{e^T} = ((\boldsymbol{P^T x})^{\boldsymbol{T}})^{\boldsymbol{T}} = \boldsymbol{P^T x}$$

となって、②に代入すれば

$$\boldsymbol{P^T x} = \boldsymbol{P^{-1} x}$$
$$\boldsymbol{P^T} = \boldsymbol{P^{-1}}$$

となって$\boldsymbol{P^T} = \boldsymbol{P^{-1}}$が、

$$\boldsymbol{A} = \lambda_1 \boldsymbol{e_1 e_1}^{\boldsymbol{T}} + \lambda_2 \boldsymbol{e_2 e_2}^{\boldsymbol{T}} + \cdots + \lambda_p \boldsymbol{e_p e_p}^{\boldsymbol{T}}$$

と書けるための条件だとわかります。
さて、それが成り立つ場合を考えましょう。

$$\boldsymbol{P^T} = \boldsymbol{P^{-1}}$$

のままでは考えにくいので、逆行列を外すために両辺に右から$\boldsymbol{P}$をかけます。

$$\boldsymbol{P^T P} = \boldsymbol{P^{-1} P} = \boldsymbol{I}$$

$$\boldsymbol{P^T P} = \boldsymbol{I}$$

同じことですが、この方が具体的に計算して考えやすいでしょう。しかし、それでもｎ次元で考えるのは複雑そうなので、まず、２次元で考えます。

$$\boldsymbol{P} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

とすると、

$$\boldsymbol{P^T P} = \boldsymbol{I}$$

$$\begin{pmatrix} a & c \\ b & d \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

ですから、以下の連立方程式が得られます。

$$a^2 + c^2 = 1 \qquad ①$$

$$b^2 + d^2 = 1 \qquad ②$$

$$ab + cd = 0 \qquad ③$$

未知数が４で式が３つですから、この連立方程式は解けませんが、よく見ると
$\boldsymbol{P}$を$\boldsymbol{P} = (\boldsymbol{P_1} \quad \boldsymbol{P_2})$のような、ベクトルの行と考えたとき、
すなわち

$$\boldsymbol{P_1} = \begin{pmatrix} a \\ c \end{pmatrix}$$

$$\boldsymbol{P_2} = \begin{pmatrix} b \\ d \end{pmatrix}$$

と考えたとき、①式は、$|\boldsymbol{P_1}|^2$ （つまりベクトル$\boldsymbol{P_1}$の長さの２乗)、式②は$|\boldsymbol{P_2}|^2$で、式③は、
$\boldsymbol{P_1} \cdot \boldsymbol{P_2}$（$\boldsymbol{P_1}$、$\boldsymbol{P_2}$の内積）になっています。
$|\boldsymbol{P_1}|^2 = 1$、$|\boldsymbol{P_2}|^2 = 1$ということは、$|\boldsymbol{P_1}| = 1$、$|\boldsymbol{P_2}| = 1$ですから、長さが１の単位ベクトルになっているベクトルをならべて、それらの内積が０、つまりそれらのベクトルが互いに直交しているということです。
行列が３次に場合についても

$$\begin{pmatrix} a & d & g \\ b & e & h \\ c & f & i \end{pmatrix}\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\begin{pmatrix} a^2+d^2+g^2 & ab+de+hg & ac+fd+ig \\ ba+de+hg & b^2+e^2+h^2 & cb+ef+ih \\ ac+fd+ig & cb+ef+ih & c^2+f^2+i^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

で

$$a^2+d^2+g^2=1$$
$$b^2+e^2+h^2=1$$
$$c^2+f^2+i^2=1$$
$$ab+de+hg=0$$
$$ac+fd+ig=0$$
$$cb+ef+ih=0$$

となり、構成する単位長さのベクトルが互いに直交している行列ならば、

$$\boldsymbol{P^T P} = \boldsymbol{I}$$

で

$$\boldsymbol{P^T} = \boldsymbol{P^{-1}}$$

で

$$\boldsymbol{A} = \lambda_1 \boldsymbol{e_1 e_1}^{\boldsymbol{T}} + \lambda_2 \boldsymbol{e_2 e_2}^{\boldsymbol{T}} + \cdots + \lambda_p \boldsymbol{e_p e_p}^{\boldsymbol{T}}$$

とかけます。ここまで来ると、何故、$\boldsymbol{e_i e_i}^{\boldsymbol{T}}$のような記号を使ったのか、その理由がわかるでしょう。$\boldsymbol{e}$は、単位ベクトルを表す時によく使われる記号です。単位ベクトルになることをあらかじめ予測して、$\boldsymbol{e}$にしておいたのです。また、$\boldsymbol{e_i}$はそれぞれのλ_1に属する固有ベクトルです。さて、この条件（p次の行列を、p個の因子をもつp個のベクトルを並べたものだと考えたとき、それらが単位ベクトルで、互いに直交している。）を満たす行列とはどんな行列かを示さなければなりません。それを「数学的」な解説とともに示そうとすると、例によって、確かに日本語ではあるが、その意味がまるでわからないという、数学の教科書によくある文章を書くことになります。今のところそんなことに意味があるように思えないので、我々がこれから対象にしようとしている、対称行列がその条件を満たすかどうかだけを考えます。ここでわざわざ取り上げるのだから、話の流れとして、当然、対称行列はこの条件に当てはまります。そのことだけ覚えておくだけでもよいのだけれど、一応、以下にその証明をします。なお、ここでは実数の範囲内での話をしています。もともと、複素数を含めて考えていないものとしてください。

証明する内容をかみ砕いておきます。今扱おうとしているのは、対角化するときに掛け合わせる行列のことです。途中の対角化の説明では、

$$\boldsymbol{Q^{-1}CQ} = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \lambda_p \end{pmatrix}$$

という計算に用いる$\boldsymbol{Q}$のことです。$\boldsymbol{Q}$は$\boldsymbol{C}$の固有ベクトルを並べたものです。ですから$\boldsymbol{Q}$を

構成するベクトルが直交するということは、$\boldsymbol{C}$の固有ベクトルが直交しているということです。また、固有ベクトルの長さは任意に決められるのだから、ベクトルの長さを１になるように定めればよいだけのことなので、証明すべきことは、$\boldsymbol{C}$の固有ベクトルが直交していることだけになります。

つまり、「対称行列 A が行列の次数 p の固有値を持つとき、任意に選んだ２つの固有ベクトルは互いに直交する。」を証明すればよいことになります。

任意の固有値をλ_1、λ_2、それぞれに対応する固有ベクトルを、$\boldsymbol{x_1}, \boldsymbol{x_2}$とします。

固有値、固有ベクトルの定義から

$$\boldsymbol{Ax_1} = \lambda_1 \boldsymbol{x_1} \qquad ①$$

$$\boldsymbol{Ax_2} = \lambda_2 \boldsymbol{x_2} \qquad ②$$

① 式の両辺を転置します。

$$(\boldsymbol{Ax_1})^{\boldsymbol{T}} = \lambda_1 \boldsymbol{x_1}^{\boldsymbol{T}}$$

$$\boldsymbol{x_1}^{\boldsymbol{T}} \boldsymbol{A}^{\boldsymbol{T}} = \lambda_1 \boldsymbol{x_1}^{\boldsymbol{T}} \qquad ①'$$

対称行列だから、

$$\boldsymbol{A}^{\boldsymbol{T}} = \boldsymbol{A}$$

これを①'に代入

$$\boldsymbol{x_1}^{\boldsymbol{T}} \boldsymbol{A} = \lambda_1 \boldsymbol{x_1}^{\boldsymbol{T}}$$

両辺に右から$\boldsymbol{x_2}$をかける。

$$\boldsymbol{x_1}^{\boldsymbol{T}} \boldsymbol{Ax_2} = \lambda_1 \boldsymbol{x_1}^{\boldsymbol{T}} \boldsymbol{x_2}$$

この左辺に②を代入

$$\boldsymbol{x_1}^{\boldsymbol{T}} \lambda_2 \boldsymbol{x_2} = \lambda_1 \boldsymbol{x_1}^{\boldsymbol{T}} \boldsymbol{x_2}$$

$$\lambda_2 \boldsymbol{x_1}^{\boldsymbol{T}} \boldsymbol{x_2} = \lambda_1 \boldsymbol{x_1}^{\boldsymbol{T}} \boldsymbol{x_2}$$

移項して

$$(\lambda_2 - \lambda_1) \boldsymbol{x_1}^{\boldsymbol{T}} \boldsymbol{x_2} = 0$$

これより

$$\lambda_2 - \lambda_1 = 0$$

または

$$\boldsymbol{x_1}^{\boldsymbol{T}} \boldsymbol{x_2} = 0$$

固有方程式が重根を持たなければ、

$$\lambda_2 - \lambda_1 \neq 0$$

したがって、

$$\boldsymbol{x_1}^{\boldsymbol{T}} \boldsymbol{x_2} = 0$$

$\boldsymbol{x_1}^{\boldsymbol{T}} \boldsymbol{x_2}$は$\boldsymbol{x_1}$と$\boldsymbol{x_2}$の内積になるのですから、$\boldsymbol{x_1}$と$\boldsymbol{x_2}$は直交します。

これで証明は終わりですが、重根を持つ場合は別に考えなければならないようです。同じ固有値を持つ違う固有ベクトル同士が直交しているという保証はありません。実はその場合には、グラム・シュミットの直交化法と呼ばれる、直交ベクトルを作る方法がありま

す。
簡単に説明すれば、たとえば、２重根の場合、重根となる固有ベクトルは一つの平面上のベクトルです。この平面は他の固有ベクトルと直交していますから、２つの固有ベクトルはそれぞれ、他の固有ベクトルとは間違いなく直交しています、そのうちの一つを選び出して、それとその平面上で直交するベクトルを求めれば良いのです。３重根や４重根、さらに m 重根も考えられますが、これらの重根に従属する固有ベクトルは、それぞれ３次空間、４次空間　m 次空間上に存在し、それらの空間は、他の固有ベクトルと直交しています。ですから、これも、グラム・シュミットの直交化法で対応できるでしょう。
具体的な計算方法は、ネット等で検索すれば出てくると思います（グラム・シュミットの直交化法)。

もっと問題になるのは、固有方程式が解をもたない場合です。一般の行列では、もちろん固有方程式が解をもたず、固有ベクトルが得られないことがあります。しかし、対称行列ではそのようなことは起こりません。証明しろと言われると困ります。実数解がないということは、虚根を持つということですね。行列を虚数まで拡張しなければなりません。再び得意のセリフを言います。「そんな難しいことを俺に訊いても知るわけないジャン。でも、そんなこと知る必要ないジャン。そうなる場合だけ考えればとりあえず問題ないジャン。問題があったら、そん時考えればいいジャン。」。

とにかく、対称行列では、

$$\boldsymbol{A} = \boldsymbol{\Lambda ee^T}$$

（$\boldsymbol{e}$は固有値の単位ベクトルを並べた行列、$\boldsymbol{\Lambda}$はラムダの大文字、固有値を対角とする対角行列の意味)
が成り立ちます
$\boldsymbol{\Lambda}$は固有値を対角とする対角行列だから

$$\boldsymbol{A} = \lambda_1 \boldsymbol{e_1 e_1}^{\boldsymbol{T}} + \lambda_2 \boldsymbol{e_2 e_2}^{\boldsymbol{T}} + \cdots + \lambda_p \boldsymbol{e_p e_p}^{\boldsymbol{T}}$$

それらしく、まとめて書けば

$$\boldsymbol{A} = \sum_{i=1}^{p} \lambda_i \boldsymbol{e_i e_i}^{\boldsymbol{T}}$$

（$\boldsymbol{e_i}$は $\boldsymbol{A}$ の固有値λ_iに対応する固有ベクトルの単位ベクトルで、任意の$\boldsymbol{e_i}$ $\boldsymbol{e_j}$は互いに直交する。)
このように対称行列を分解することをスペクトル分解と言います。教科書によっては、対称行列でない場合にもスペクトル分解と呼ぶことがあるようです。また、対称行列でない場合は、疑似的なスペクトル分解としている例もあります。
とにかく覚えておくことは、「対称行列は、固有ベクトルが直交していて。スペクトル分解できる。」ということです。

抽象的な議論が続いたので、理解のために、具体的な例でスペクトル分解をやってみます。

$$\boldsymbol{A}=\begin{pmatrix}3 & 1 & -1\\ 1 & 5 & -1\\ -1 & -1 & 3\end{pmatrix}$$

$\boldsymbol{A}$ は対称行列になっていますね。

まず、固有値を求めるために、固有方程式を作ります。

$$\begin{vmatrix}(3-\lambda) & 1 & -1\\ 1 & (5-\lambda) & -1\\ -1 & -1 & (3-\lambda)\end{vmatrix}=0$$

これを解きます。

$$\begin{vmatrix}(3-\lambda) & 1 & -1\\ 1 & (5-\lambda) & -1\\ -1 & -1 & (3-\lambda)\end{vmatrix}=0$$

$$(3-\lambda)(5-\lambda)(3-\lambda)+1+1-(3-\lambda)-(5-\lambda)-(3-\lambda)=0$$

$$36-36\lambda+11\lambda^2-\lambda^3=0$$

$$(\lambda-6)(\lambda-3)(\lambda-2)=0$$

固有値 $\lambda_1=6,\lambda_2=3,\lambda_3=2$

固有ベクトルを求めます。

$\lambda_1=6$　に属する固有ベクトル

$$\begin{pmatrix}3 & 1 & -1\\ 1 & 5 & -1\\ -1 & -1 & 3\end{pmatrix}\begin{pmatrix}x_1\\ x_2\\ x_3\end{pmatrix}=6\begin{pmatrix}x_1\\ x_2\\ x_3\end{pmatrix}$$

$$3x_1+x_2-x_3=6x_1 \qquad ①$$

$$x_1+5x_2-x_3=6x_2 \qquad ②$$

$$-x_1-x_2+3x_3=6x_3 \qquad ③$$

$$-3x_1+x_2-x_3=0 \qquad ①'$$

$$x_1-x_2-x_3=0 \qquad ②'$$

$$-x_1-x_2-3x_3=0 \qquad ③'$$

$$-2x_1-2x_3=0 \qquad ①'+②'$$

$$-2x_2-4x_3=0 \qquad ②'+③$$

$$x_1=-x_3$$

$$x_2=-2x_3$$

したがって固有値 6 に属する固有ベクトルは$t\begin{pmatrix}1\\ 2\\ -1\end{pmatrix}$

$\lambda_1=3$　に属する固有ベクトル

$$\begin{pmatrix}3 & 1 & -1\\ 1 & 5 & -1\\ -1 & -1 & 3\end{pmatrix}\begin{pmatrix}x_1\\ x_2\\ x_3\end{pmatrix}=3\begin{pmatrix}x_1\\ x_2\\ x_3\end{pmatrix}$$

$3x_1+x_2-x_3=3x_1$ ①

$x_1+5x_2-x_3=3x_2$ ②

$-x_1-x_2+3x_3=3x_3$ ③

$x_2-x_3=0$ ①'

$x_1+2x_2-x_3=0$ ②'

$-x_1-x_2=0$ ③'

$$x_1=-x_2$$

$$x_2=x_3$$

したがって固有値 3 に属する固有ベクトルは$t\begin{pmatrix}1\\ -1\\ -1\end{pmatrix}$

$\lambda_1=2$　に属する固有ベクトル

$$\begin{pmatrix}3 & 1 & -1\\ 1 & 5 & -1\\ -1 & -1 & 3\end{pmatrix}\begin{pmatrix}x_1\\ x_2\\ x_3\end{pmatrix}=2\begin{pmatrix}x_1\\ x_2\\ x_3\end{pmatrix}$$

$3x_1+x_2-x_3=2x_1$ ①

$x_1+5x_2-x_3=2x_2$ ②

$-x_1-x_2+3x_3=2x_3$ ③

$x_1+x_2-x_3=0$ ①'

$x_1+3x_2-x_3=0$ ②'

$-x_1-x_2+x_3=0$ ③'

$-2x_2=0$ ①' − ②'

$$x_2=0$$

これを一に代入

$x_1-x_3=0$ ②' + ③

$$x_1=x_3$$

したがって固有 2 に属する固有ベクトルは$t\begin{pmatrix}1\\ 0\\ 1\end{pmatrix}$

ここまでの結果をまとめると、固有値$\lambda_1=6,\lambda_2=3,\lambda_3=2$に属する固有ベクトルは、それぞれ

$$t\begin{pmatrix}1\\ 2\\ -1\end{pmatrix}、t\begin{pmatrix}1\\ -1\\ -1\end{pmatrix}、t\begin{pmatrix}1\\ 0\\ 1\end{pmatrix}$$

それぞれの単位ベクトルは

$$\boldsymbol{e_1}=\begin{pmatrix}\frac{1}{\sqrt{6}}\\\frac{2}{\sqrt{6}}\\\frac{-1}{\sqrt{6}}\end{pmatrix},\ \ \boldsymbol{e_2}=\begin{pmatrix}\frac{1}{\sqrt{3}}\\\frac{-1}{\sqrt{3}}\\\frac{-1}{\sqrt{3}}\end{pmatrix},\boldsymbol{e_3}=\begin{pmatrix}\frac{1}{\sqrt{2}}\\0\\\frac{1}{\sqrt{2}}\end{pmatrix}$$

となるので、スペクトル分解の結果は

$$\boldsymbol{A}=6\boldsymbol{e_1}\boldsymbol{e_1}^{\boldsymbol{T}}+3\boldsymbol{e_2}\boldsymbol{e_2}^{\boldsymbol{T}}+2\boldsymbol{e_3}\boldsymbol{e_3}^{\boldsymbol{T}}$$

$$\boldsymbol{A}=6\begin{pmatrix}\frac{1}{\sqrt{6}}\\\frac{2}{\sqrt{6}}\\\frac{-1}{\sqrt{6}}\end{pmatrix}\begin{pmatrix}\frac{1}{\sqrt{6}}\\\frac{2}{\sqrt{6}}\\\frac{-1}{\sqrt{6}}\end{pmatrix}^T+3\begin{pmatrix}\frac{1}{\sqrt{3}}\\\frac{-1}{\sqrt{3}}\\\frac{-1}{\sqrt{3}}\end{pmatrix}\begin{pmatrix}\frac{1}{\sqrt{3}}\\\frac{-1}{\sqrt{3}}\\\frac{-1}{\sqrt{3}}\end{pmatrix}^T+2\begin{pmatrix}\frac{1}{\sqrt{2}}\\0\\\frac{1}{\sqrt{2}}\end{pmatrix}\begin{pmatrix}\frac{1}{\sqrt{2}}\\0\\\frac{1}{\sqrt{2}}\end{pmatrix}^T$$

となります。一応、確かにそのようになっているかどうか、計算して確かめてみましょう。

右辺の一項目

$$6\begin{pmatrix}\frac{1}{\sqrt{6}}\\\frac{2}{\sqrt{6}}\\\frac{-1}{\sqrt{6}}\end{pmatrix}\begin{pmatrix}\frac{1}{\sqrt{6}}\\\frac{2}{\sqrt{6}}\\\frac{-1}{\sqrt{6}}\end{pmatrix}^T=6\begin{pmatrix}\frac{1}{\sqrt{6}}\\\frac{2}{\sqrt{6}}\\\frac{-1}{\sqrt{6}}\end{pmatrix}\begin{pmatrix}\frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} & \frac{-1}{\sqrt{6}}\end{pmatrix}$$

$$=6\begin{pmatrix}\frac{1}{6} & \frac{1}{3} & -\frac{1}{6}\\\frac{1}{3} & \frac{4}{6} & -\frac{1}{3}\\-\frac{1}{6} & -\frac{1}{3} & \frac{1}{6}\end{pmatrix}=\begin{pmatrix}1 & 2 & -1\\2 & 4 & -2\\-1 & -2 & 1\end{pmatrix}$$

2 項目

$$3\begin{pmatrix}\frac{1}{\sqrt{3}}\\\frac{-1}{\sqrt{3}}\\\frac{-1}{\sqrt{3}}\end{pmatrix}\begin{pmatrix}\frac{1}{\sqrt{3}}\\\frac{-1}{\sqrt{3}}\\\frac{-1}{\sqrt{3}}\end{pmatrix}^T=3\begin{pmatrix}\frac{1}{\sqrt{3}}\\-\frac{1}{\sqrt{3}}\\-\frac{1}{\sqrt{3}}\end{pmatrix}\begin{pmatrix}\frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}}\end{pmatrix}$$

$$3\begin{pmatrix}\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3}\\-\frac{1}{3} & \frac{1}{3} & \frac{1}{3}\\-\frac{1}{3} & \frac{1}{3} & \frac{1}{3}\end{pmatrix}=\begin{pmatrix}1 & -1 & -1\\-1 & 1 & 1\\-1 & 1 & 1\end{pmatrix}$$

3 項目

$$2\begin{pmatrix}\frac{1}{\sqrt{2}}\\0\\\frac{1}{\sqrt{2}}\end{pmatrix}\begin{pmatrix}\frac{1}{\sqrt{2}}\\0\\\frac{1}{\sqrt{2}}\end{pmatrix}^T=2\begin{pmatrix}\frac{1}{\sqrt{2}}\\0\\\frac{1}{\sqrt{2}}\end{pmatrix}\begin{pmatrix}\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}}\end{pmatrix}$$

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

右辺は

$$\begin{pmatrix}1 & 2 & -1\\2 & 4 & -2\\-1 & -2 & 1\end{pmatrix}+\begin{pmatrix}1 & -1 & -1\\-1 & 1 & 1\\-1 & 1 & 1\end{pmatrix}+\begin{pmatrix}1 & 0 & 1\\0 & 0 & 0\\1 & 0 & 1\end{pmatrix}=\begin{pmatrix}3 & 1 & -1\\1 & 5 & -1\\-1 & -1 & 3\end{pmatrix}$$

確かに、スペクトル分解の公式どおりの結果になっています。

直交性の方も確認しておきますか。

$$\boldsymbol{e_1}^{\boldsymbol{T}}\boldsymbol{e_2}=\begin{pmatrix}\frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} & \frac{-1}{\sqrt{6}}\end{pmatrix}\begin{pmatrix}\frac{1}{\sqrt{3}}\\\frac{-1}{\sqrt{3}}\\\frac{-1}{\sqrt{3}}\end{pmatrix}=\frac{1}{\sqrt{6}}\frac{1}{\sqrt{3}}-\frac{2}{\sqrt{6}}\frac{1}{\sqrt{3}}+\frac{1}{\sqrt{6}}\frac{1}{\sqrt{3}}=0$$

$$\boldsymbol{e_2}^{\boldsymbol{T}}\boldsymbol{e_3}=\begin{pmatrix}\frac{1}{\sqrt{3}} & \frac{-1}{\sqrt{3}} & \frac{-1}{\sqrt{3}}\end{pmatrix}\begin{pmatrix}\frac{1}{\sqrt{2}}\\0\\\frac{1}{\sqrt{2}}\end{pmatrix}=\frac{1}{\sqrt{3}}\frac{1}{\sqrt{2}}-\frac{1}{\sqrt{3}}0-\frac{1}{\sqrt{3}}\frac{1}{\sqrt{2}}=0$$

$$\boldsymbol{e_2}^{\boldsymbol{T}}\boldsymbol{e_3}=\begin{pmatrix}\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}}\end{pmatrix}\begin{pmatrix}\frac{1}{\sqrt{6}}\\\frac{2}{\sqrt{6}}\\\frac{-1}{\sqrt{6}}\end{pmatrix}=\frac{1}{\sqrt{2}}\frac{1}{\sqrt{6}}-0\frac{2}{\sqrt{6}}-\frac{1}{\sqrt{2}}\frac{1}{\sqrt{6}}=0$$

確かに直交しています。

計算作業を通じて、スペクトル分解が具体的に理解できたと思います。

２次形式

何故、ここまでスペクトル分解について説明してきたのかというと、スペクトル分解というテクニックが多変量解析で使われるからです。

$$\begin{pmatrix} 3 & 1 & -1 \\ 1 & 5 & -1 \\ -1 & -1 & 3 \end{pmatrix} = 6\begin{pmatrix} \frac{1}{6} & \frac{1}{3} & -\frac{1}{6} \\ \frac{1}{3} & \frac{4}{6} & -\frac{1}{6} \\ -\frac{1}{6} & -\frac{1}{6} & \frac{1}{6} \end{pmatrix} + 3\begin{pmatrix} \frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} \\ -\frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ -\frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix} + 2\begin{pmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix}$$

というのが、直前にやったスペクトル分解の具体例ですが、よく見ると、対称行列を分解した結果、対称行列の固有値倍の総和になっています。以前、対称行列でない一般の行列について、

$$\boldsymbol{A} = \lambda_1 \boldsymbol{P}\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}\boldsymbol{P^{-1}} + \lambda_2 \boldsymbol{P}\begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}\boldsymbol{P^{-1}} + \cdots + \lambda_n \boldsymbol{P}\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}\boldsymbol{P^{-1}}$$

の、$\boldsymbol{P}\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}\boldsymbol{P^{-1}}$とは何か、という疑問を掲げて、一般には意味がなさそうだが、特殊な場合には意味を持つかもしれないと予測を述べておきました。

対称行列ならば、スペクトル分解によって、

$$\boldsymbol{A} = \lambda_1 \boldsymbol{e_1}\boldsymbol{e_1}^{\boldsymbol{T}} + \lambda_2 \boldsymbol{e_2}\boldsymbol{e_2}^{\boldsymbol{T}} + \cdots + \lambda_p \boldsymbol{e_p}\boldsymbol{e_p}^{\boldsymbol{T}}$$

のように書けることを示しました。つまり、

$$\lambda_i \boldsymbol{P}\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}\boldsymbol{P^{-1}} = \lambda_i \boldsymbol{e_i}\boldsymbol{e_i}^{\boldsymbol{T}}$$

ということです。$\boldsymbol{e_i}\boldsymbol{e_i}^{\boldsymbol{T}}$は対称行列になりますね。直感的にわからなければ、何か数字を入れた行列を作って実際に計算してみたらよいでしょう。対称行列は、対称行列の固有値倍の総和として書けるのです。こういうのはなんていうのかわかりません。空間幾何学的な相似とは違うので「フラクタクル構造」というのとは違うのでしょうね。しかし、対称行列が、その下のレベルの複数の対称行列から構成されるというのは、日本語でいえば、入れ子構造とか、ロシアのマトリューシュカ人形のようなイメージですね。

長々と説明しましたが、こんな結論ならば、最初から、対称行列の固有ベクトルは直交するということを説明し、これをスペクトル分解するとどうなるかと説明にした方がわかりやすくて、速かったでしょう。そう覚えてもらって構いません。しかし、その前に、相似とはどういうことなのか、相似を使って何が出来るのかを説明しておきたかったのです。

V-2-4. 二次形式

V-2-4-1. 対称行列と二次形式

つぎのような式を二次形式と呼びます。二次式だけで出来ているからです。

$$ax_1{}^2+bx_2{}^2+cx_3{}^2+dx_1x_2+ex_2x_3+fx_3x_1$$

そう考えると、次の形の式は一次形式ということになります。

$$\mathrm{a}x_1+bx_2+cx_3$$

１次形式を行列で表すと

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix}^T \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

例としたあげた２次形式（$ax_1{}^2+bx_2{}^2+cx_3{}^2+dx_1x_2+ex_2x_3+fx_3x_1$）を行列で表すと次のようになります。

$$\begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix} \begin{pmatrix} a & d & f \\ 0 & b & e \\ 0 & 0 & c \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

確かめてみます。

$$\begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix} \begin{pmatrix} a & d & f \\ 0 & b & e \\ 0 & 0 & c \end{pmatrix} = \begin{pmatrix} ax_1 & dx_1+bx_2+ & fx_1+ex_2+cx_3 \end{pmatrix}$$

$$\begin{pmatrix} ax_1 & dx_1+bx_2+ & fx_1+ex_2+cx_3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

$$=ax_1{}^2+bx_2{}^2+cx_3{}^2+dx_1x_2+ex_2x_3+fx_3x_1$$

確かに、２次形式を表すことができます。ですから、このように書いても間違いではありませんし、そちらの方が普通でしょう。

しかし、ここでは、次のように表したいのです。

$$\begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix} \begin{pmatrix} a & \frac{d}{2} & \frac{f}{2} \\ \frac{d}{2} & b & \frac{e}{2} \\ \frac{f}{2} & \frac{e}{2} & c \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

実際にそうなっていることを確かめます。

$$\begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix} \begin{pmatrix} a & \frac{d}{2} & \frac{f}{2} \\ \frac{d}{2} & b & \frac{e}{2} \\ \frac{f}{2} & \frac{e}{2} & c \end{pmatrix} = \left(ax_1+\frac{d}{2}x_2+\frac{f}{2}x_3 \quad \frac{d}{2}x_1+bx_2+\frac{e}{2}x_3 \quad \frac{f}{2}x_1+\frac{e}{2}x_2+cx_3 \right)$$

$$\left(ax_1+\frac{d}{2}x_2+\frac{f}{2}x_3 \quad \frac{d}{2}x_1+bx_2+\frac{e}{2}x_3 \quad \frac{f}{2}x_1+\frac{e}{2}x_2+cx_3 \right) \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

$$=\mathrm{a}x_1{}^2+\frac{d}{2}x_1x_2+\frac{f}{2}x_1x_3+\frac{d}{2}x_1x_2+bx_2{}^2+\frac{e}{2}x_1x_3+\frac{f}{2}x_1x_3+\frac{e}{2}x_2x_3+cx_3{}^2$$

$$=ax_1{}^2+bx_2{}^2+cx_3{}^2+dx_1x_2+ex_2x_3+fx_3x_1$$

つまり、わざわざ対称行列にしているのです。そこが重要なところです。

p次の対称行列を使うと、２次形式は次のように書けます。

$$(x_1 \quad x_2 \quad \cdots \quad x_p)\begin{pmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1p} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{i1} & \cdots & a_{ij} & \cdots & a_{ip} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{p1} & \cdots & a_{pj} & \cdots & a_{pp} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix}$$

ベクトル$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix}$を$\boldsymbol{x}$、$\begin{pmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1p} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{i1} & \cdots & a_{ij} & \cdots & a_{ip} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{p1} & \cdots & a_{pj} & \cdots & a_{pp} \end{pmatrix}$を$\boldsymbol{A}$とあらわすと、この式はもっと簡便化して次にように書くことができます。

$$\boldsymbol{x}^{\boldsymbol{T}}\boldsymbol{A}\boldsymbol{x}$$

V-2-4-2. 二次形式の対角化

二次形式が実数値を持つ時、2次形式は空間に軌跡を描きます。その形をn次元空間で、言葉で表現することは難しいのですが、例えば、2次元空間で、二次形式が描くきせきは、円、楕円、双曲線、放物線のような、二次曲線、円錐曲線です。

例えば、対称行列$\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$について考えます。

固有方程式

$$\begin{vmatrix} 1-\lambda & 2 \\ 2 & 1-\lambda \end{vmatrix} = 0$$

$$(\lambda - 1)^2 - 4 =$$

$$\lambda = 3, -1$$

固有ベクトルを求めます。

$$x_1 + 2x_2 = 3x_1$$

$$2x_1 + x_2 = 3x_2$$

$$x_1 = x_2$$

$$\boldsymbol{e}_1 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$$

$$x_1 + 2x_2 = -x_1$$

$$2x_1 + x_2 = -x_2$$

$$x_1 = -x_2$$

$$\boldsymbol{e}_2 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix}$$

対角化行列

$$\boldsymbol{P} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}$$

$$\boldsymbol{P}^{-1} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}$$

$$\begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}$$

$$\begin{pmatrix} \frac{1}{\sqrt{2}} + \sqrt{2} & \sqrt{2} + \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} - \sqrt{2} & \sqrt{2} - \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} \frac{1}{2} + 1 + 1 + \frac{1}{2} & \frac{1}{2} + 1 - 1 - \frac{1}{2} \\ \frac{1}{2} - 1 + 1 - \frac{1}{2} & \frac{1}{2} - 1 - 1 + \frac{1}{2} \end{pmatrix}$$

$$= \begin{pmatrix} 3 & 0 \\ 0 & -1 \end{pmatrix}$$

スペクトル分解

$$\lambda_1 \boldsymbol{e}_1 {\boldsymbol{e}_1}^T = 3 \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} = 3 \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

$$\lambda_2 \boldsymbol{e}_2 {\boldsymbol{e}_2}^T = -1 \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} = - \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

$$\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} = \lambda_1 \boldsymbol{e}_1 {\boldsymbol{e}_1}^T + \lambda_2 \boldsymbol{e}_2 {\boldsymbol{e}_2}^T = 3 \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \pm 1 \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}$$

$$= 3\begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} - \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

このスペクトル分解の結果から、反対に二次式を作ります。

$$\boldsymbol{x}^T \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \boldsymbol{x} = \lambda_1 \boldsymbol{x}^T \boldsymbol{e}_1 {\boldsymbol{e}_1}^T \boldsymbol{x} + \lambda_2 \boldsymbol{x}^T \boldsymbol{e}_2 {\boldsymbol{e}_2}^T \boldsymbol{x}$$

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

$$\lambda_1 \boldsymbol{x}^T \boldsymbol{e}_1 {\boldsymbol{e}_1}^T \boldsymbol{x} = 3(x_1 \quad x_2) \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 3\begin{pmatrix} \frac{x_1 + x_2}{2} & \frac{x_1 + x_2}{2} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

$$= 3\begin{pmatrix} \frac{x_1 + x_2}{2} & \frac{x_1 + x_2}{2} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \frac{3}{2}({x_1}^2 + 2x_1x_2 + {x_1}^2) = 3\left(\frac{x_1 + x_2}{\sqrt{2}}\right)^2$$

$$\lambda_2 \boldsymbol{x}^T \boldsymbol{e}_2 {\boldsymbol{e}_2}^T \boldsymbol{x} = -(x_1 \quad x_2) \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = -\begin{pmatrix} \frac{x_1 - x_2}{2} & \frac{-x_1 + x_2}{2} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

$$= -\frac{1}{2}({x_1}^2 - 2x_1x_2 + {x_1}^2) = -\left(\frac{x_2 - x_1}{\sqrt{2}}\right)^2$$

$$\therefore \quad \lambda_1 \boldsymbol{x}^T \boldsymbol{e}_1 {\boldsymbol{e}_1}^T \boldsymbol{x} + \lambda_2 \boldsymbol{x}^T \boldsymbol{e}_2 {\boldsymbol{e}_2}^T \boldsymbol{x} = 3\left(\frac{x_1 + x_2}{\sqrt{2}}\right)^2 - \left(\frac{x_2 - x_1}{\sqrt{2}}\right)^2 \qquad \text{i}$$

一方、もとの行列からそのまま計算すると、

$$\boldsymbol{x}^T \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \boldsymbol{x} = (x_1 \quad x_2) \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = (x_1 + 2x_2 \quad 2x_1 + x_2) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = {x_1}^2 + 4x_1x_2 + {x_2}^2 \quad \text{ii}$$

となりますが、確かに以下の計算変形が正しいということは、下の等式の右辺を計算してみればわかります。

$${x_1}^2 + 4x_1x_2 + {x_2}^2 = 3\left(\frac{x_1 + x_2}{\sqrt{2}}\right)^2 - \left(\frac{x_2 - x_1}{\sqrt{2}}\right)^2$$

もう少しわかりやすくするために、以下の変換を行います。

$$x_1 + x_2 = X_1$$

$$x_2 - x_1 = X_2$$

$${x_1}^2 + 4x_1x_2 + {x_2}^2 = \frac{{X_1}^2}{\left(\frac{\sqrt{2}}{\sqrt{3}}\right)^2} - \frac{{X_2}^2}{\sqrt{2}^2}$$

これが、たとえば、1 という実数値を持つならば、これは双曲線の式で、X_1、X_2は双曲線を描きます。

図 5 の赤い細い矢印は固有ベクトルです。固有ベクトルは直交しています。また、X_1、X_2

も固有ベクトルと並行しているので、固有ベクトルです。

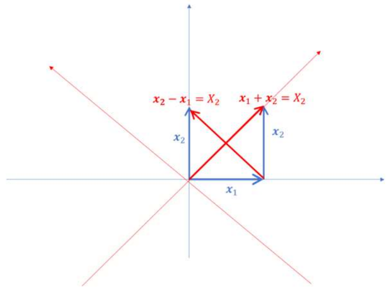

図 5 行列の固有ベクトルと、変換してできるベクトル

二次形式だけでは図形的な意味を持ちませんが、二次形式がある実数値をとるときには、その軌跡は特定の形を持ちます。その形は円錐曲線とよばれているもので、円錐を平面で切った時の断面の形状で、断面の向きによって形状が異なります。円錐曲線は準線と焦点からの距離の比が一定になる点の軌跡ですが、この比によって、円、楕円、放物線、双曲線に分けられます。双曲線は円錐曲線の一つですが、より実用的な的な定義は、２つの焦点からの距離の差が一定の点の集合です。

この定義に従って、双曲線を作ります。焦点を P,Q とします。

$$\mathrm{P}: (f \quad 0)$$

$$\mathrm{Q}: (f \quad 0)$$

$$|L\mathrm{Q}| - |L\mathrm{P}| = |2a| = \pm 2a$$

$$|L\mathrm{Q}| = \pm 2a + |L\mathrm{P}|$$

$$|LQ|^2 = 4a^2 \pm 4a|L\mathrm{P}| + |L\mathrm{P}|^2$$

$$|L\mathrm{P}| = \sqrt{(x-f)^2 + (y-0)^2}$$

$$|L\mathrm{Q}| = \sqrt{(x+f)^2 + (y-0)^2}$$

$$(x+f)^2 + y^2 = 4a^2 \pm 4a\sqrt{(x-f)^2 + y^2} + (x-f)^2 + y^2$$

$$4fx - 4a^2 = 4a\sqrt{(x-f)^2 + y^2}$$

$$fx - a^2 = a\sqrt{(x-f)^2 + y^2}$$

$$(fx - a^2)^2 = a^2(x-f)^2 + a^2y^2$$

$$f^2x^2 - 2c^2fx + a^2 = a^2x^2 - 2a^2fx + a^2f^2 + a^2y^2$$

$$(f^2 - a^2)x^2 - a^2y^2 = a^2f^2 - a^4$$

$$(f^2 - a^2)x^2 - a^2y^2 = a^2(f^2 - a^2)$$

$$\frac{x^2}{a^2} - \frac{y^2}{(f^2 - a^2)} = 1$$

$(f^2 - a^2) = b^2$とします。。

$$f^2 = a^2 + b^2$$

$$f = \pm\sqrt{a^2 + b^2}$$

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$$

これが、双曲線の一般式です。この式をさらに変形します。

$$\frac{x^2}{a^2}-1=\frac{y^2}{b^2}$$

$$1-\frac{a^2}{x^2}=\frac{a^2y^2}{b^2x^2}$$

$$\lim_{x\to\pm\infty}\left(1-\frac{a^2}{x^2}\right)=1$$

$$\lim_{x\to\pm\infty}\frac{a^2y^2}{b^2x^2}=1$$

$$\lim_{x\to\pm\infty}\frac{ay}{bx}=\pm1$$

この変形の意味は、xの増加に従ってxとyの比が一定の値に近づくということです。したがって、双曲線は$ay= bx$ あるいは $ay=-bx$ という直線に近づいていきます。この直線を漸近線といいます(グラフ A と B)。

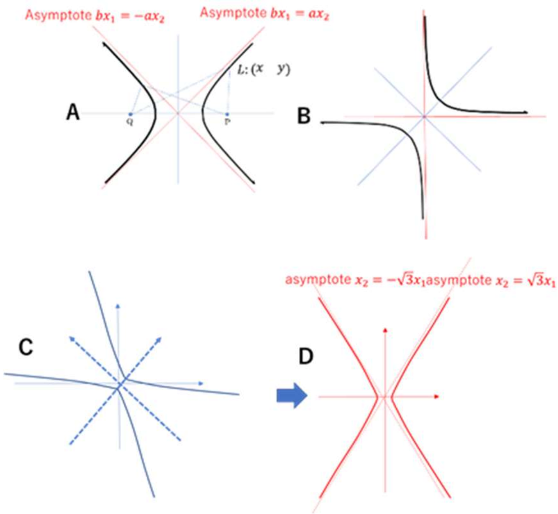

図 6. 双曲線と漸近線

図 6 のグラフ A は$x^2-y^2=1$のグラフです。これを反時計方向に $\frac{\pi}{4}$回転すると B のグラフが得られます。回転の式は以下の式です。

$$\begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}\cos\theta & \sin\theta\\-\sin\theta & \cos\theta\end{pmatrix}\begin{pmatrix}X\\Y\end{pmatrix}$$

$$\theta=\frac{\pi}{4}$$

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix}$$

$$x = \frac{1}{\sqrt{2}}(X+Y)$$

$$y = \frac{1}{\sqrt{2}}(-X+Y)$$

$$x^2 - y^2 = \frac{1}{2}((X+Y)^2 - (-X+Y)^2) = 2XY$$

$$2XY = 1$$

$$Y = \frac{1}{2X}$$

この式は反比例の式です。反比例の式も双曲線の式で、X 軸と Y 軸が漸近線になっています。上述の方法で双曲線の式を作るときは、X 線上に 2 つの焦点を置きました。しかし、一般的には焦点は X 線上にありません。これを回転して、X 線上に焦点を置く作業が直交化だと言えます。図 6 のグラフ C は以下の式の軌跡です。

$$\begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 1$$

ここから、直接、軌跡を描くことを考えます。

$${x_1}^2 + 4x_1x_2 + {x_2}^2 = 1$$
$${x_1}^2 + 4x_1x_2 + {x_2}^2 = 1$$
$$(x_2 + 2x_1)^2 - 3{x_1}^2 = 1$$
$$(x_2 + 2x_1)^2 = 1 + 3{x_1}^2$$
$$x_2 + 2x_1 = \pm\sqrt{1 + 3{x_1}^2}$$
$$x_2 = -2x_1 \pm \sqrt{1 + 3{x_1}^2}$$

$$\frac{x_2}{x_1} = -2 \pm \sqrt{\frac{1}{x^2} + 3}$$

$$\lim_{x_1 \to \pm\infty} \left(-2 \pm \sqrt{\frac{1}{x^2} + 3} \right) = -2 \pm \sqrt{3}$$

$$\lim_{x_1 \to \pm\infty} \frac{x_2}{x_1} = -2 \pm \sqrt{3}$$

したがって、漸近線は $x_2 = -\left(2 + \sqrt{3}\right)x_1$ と $x_2 = -\left(2 - \sqrt{3}\right)x_1$ です。
これを時計方向に$\frac{\pi}{4}$回転します。

$$\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 1 \\ -(2+\sqrt{3}) \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} + \frac{2+\sqrt{3}}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} - \frac{2+\sqrt{3}}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} \frac{3+\sqrt{3}}{\sqrt{2}} \\ \frac{-1-\sqrt{3}}{\sqrt{2}} \end{pmatrix}$$

$$\frac{X_1}{X_2} = \frac{3+\sqrt{3}}{-1-\sqrt{3}} = \frac{-(3+\sqrt{3})(\sqrt{3}-1)}{(\sqrt{3}+1)(\sqrt{3}-1)} = \frac{-(3-3-2\sqrt{3})}{3-1} = \sqrt{3}$$

$$\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 1 \\ -(2-\sqrt{3}) \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} + \frac{2-\sqrt{3}}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} - \frac{2-\sqrt{3}}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} \frac{3-\sqrt{3}}{\sqrt{2}} \\ \frac{-1+\sqrt{3}}{\sqrt{2}} \end{pmatrix}$$

$$\frac{X_1}{X_2} = \frac{3-\sqrt{3}}{-1+\sqrt{3}} = \frac{-(3-\sqrt{3})(\sqrt{3}+1)}{(\sqrt{3}-1)(\sqrt{3}+1)} = \frac{-(3-3+2\sqrt{3})}{3-1} = -\sqrt{3}$$

これから、この奇跡が円錐曲線になることは想像できますが、どんな形か想像するのは難しいかもしれません。次のように直交化すると、双曲線であることがわかります。

$$\begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} 3 & 0 \\ 0 & -1 \end{pmatrix}$$

$$\begin{pmatrix} X_1 & X_2 \end{pmatrix} \begin{pmatrix} 3 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} = 1$$

固有値に正の固有値と負の固有値がある場合、双曲線になります。

ごちゃごちゃ計算したので、要点だけをまとめます。図 6 の
A は双曲線の定義で、点 Q と P があって、その差が一定の点の軌跡です。この図は、点 P と Q を横軸上に置き、その中間点を原点にとっています。赤い線が漸近線($asymptotete$)です。これを$\frac{\pi}{2}$反時計回りに回転したのが、反比例の曲線です。例示した計算式では、${x_1}^2 + 4x_1x_2 + {x_2}^2 = 1$の軌跡は、C の様になっていますが、直交する固有ベクトルを、縦軸横軸にとると、つまり、x_1、x_2を固有ベクトルに投影すると、D のようになります。

次に$\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$について考えます。

対角化します。

$$\begin{vmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{vmatrix} = 0$$

$$(2-\lambda)^2 - 1 = 0$$

$$\lambda - 2 = \pm 1$$
$$\lambda = 3, 1$$
$$2x_1 + x_2 = 3x_1$$
$$x_1 + 2x_2 = 3x_2$$
$$x_1 = x_2$$
$$\boldsymbol{e}_1 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$$
$$2x_1 + x_2 = x_1$$
$$x_1 + 2x_2 = x_2$$
$$x_1 = -x_2$$
$$\boldsymbol{e}_2 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix}$$

したがって、対角化行列は

$$\boldsymbol{P} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}, \qquad \boldsymbol{P^{-1}} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}$$

となります。

$$\boldsymbol{P^{-1}} \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \boldsymbol{P}$$

$$= \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}$$

$$= \begin{pmatrix} \frac{2}{\sqrt{2}} + \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} + \frac{2}{\sqrt{2}} \\ \frac{2}{\sqrt{2}} - \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} - \frac{2}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} \frac{3}{\sqrt{2}} & \frac{3}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}$$

$$= \begin{pmatrix} \frac{3+3}{2} & \frac{3-3}{2} \\ \frac{1-1}{2} & \frac{1+1}{2} \end{pmatrix} = \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix}$$

スペクトル分解します。

$$\lambda_1 \boldsymbol{e}_1 {\boldsymbol{e}_1}^T = 3\begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}\begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} = 3\begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

$$\lambda_2 \boldsymbol{e}_2 {\boldsymbol{e}_2}^T = 1\begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix}\begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

$$\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} = \lambda_1 \boldsymbol{e}_1 {\boldsymbol{e}_1}^T + \lambda_2 \boldsymbol{e}_2 {\boldsymbol{e}_2}^T = 3\begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}\begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} + 1\begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix}\begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}$$

$$= 3\begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} + \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

この結果を使って、二次形式を作ります。

$$\boldsymbol{x}^T \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \boldsymbol{x} = \lambda_1 \boldsymbol{x}^T \boldsymbol{e}_1 {\boldsymbol{e}_1}^T \boldsymbol{x} + \lambda_2 \boldsymbol{x}^T \boldsymbol{e}_2 {\boldsymbol{e}_2}^T \boldsymbol{x}$$

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

$$\lambda_1 \boldsymbol{x}^T \boldsymbol{e}_1 {\boldsymbol{e}_1}^T \boldsymbol{x} = 3(x_1 \quad x_2)\begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 3\begin{pmatrix} \frac{x_1 + x_2}{2} & \frac{x_1 + x_2}{2} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

$$= 3\begin{pmatrix} \frac{x_1 + x_2}{2} & \frac{x_1 + x_2}{2} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \frac{3}{2}({x_1}^2 + 2x_1x_2 + {x_1}^2) = 3\left(\frac{x_1 + x_2}{\sqrt{2}}\right)^2$$

$$\lambda_2 \boldsymbol{x}^T \boldsymbol{e}_2 {\boldsymbol{e}_2}^T \boldsymbol{x} = (x_1 \quad x_2)\begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \frac{x_1 - x_2}{2} & \frac{-x_1 + x_2}{2} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

$$= \frac{1}{2}({x_1}^2 - 2x_1x_2 + {x_1}^2) = \left(\frac{x_2 - x_1}{\sqrt{2}}\right)^2$$

$$\therefore \quad \lambda_1 \boldsymbol{x}^T \boldsymbol{e}_1 {\boldsymbol{e}_1}^T \boldsymbol{x} + \lambda_2 \boldsymbol{x}^T \boldsymbol{e}_2 {\boldsymbol{e}_2}^T \boldsymbol{x} = 3\left(\frac{x_1 + x_2}{\sqrt{2}}\right)^2 + \left(\frac{x_2 - x_1}{\sqrt{2}}\right)^2 \qquad \text{i}$$

$$(x_1 \quad x_2)\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = (2x_1 + x_2 \quad x_1 + 2x_2)\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 2{x_1}^2 + 2x_1x_2 + 2{x_2}^2$$

$$= 3\left(\frac{x_1 + x_2}{\sqrt{2}}\right)^2 + \left(\frac{x_2 - x_1}{\sqrt{2}}\right)^2$$

$$x_1 + x_2 = X_1,\, x_2 - x_1 = X_2$$

$$(x_1 \quad x_2)\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \frac{{X_1}^2}{\left(\frac{\sqrt{2}}{\sqrt{3}}\right)^2} + \frac{{X_2}^2}{\sqrt{2}^2}$$

これは楕円の式です。楕円の式の一般的な形は次の通りです。

$$\frac{{x_1}^2}{a^2} + \frac{{x_2}^2}{b^2} = 1$$

楕円は２つの焦点からの距離の和が一定の点の軌跡です。この定義に従って楕円の式を作ります。

$$\mathrm{P}: (f \quad 0)$$

$$\mathrm{Q}: (f \quad 0)$$

$$|L\mathrm{Q}| + |L\mathrm{P}| = 2a$$

$$|L\mathrm{Q}|^2 + 2|L\mathrm{Q}||L\mathrm{P}| + |L\mathrm{P}|^2 = 4a^2$$

$$|L\mathrm{P}| = \sqrt{(x-f)^2 + (y-0)^2}$$

$$|L\mathrm{Q}| = \sqrt{(x+f)^2 + (y-0)^2}$$

$$(x+f)^2 + y^2 + 2\sqrt{(x-f)^2 + y^2}\sqrt{(x+f)^2 + y^2} + (x-f)^2 + y^2 = 4a^2$$

$$2x^2 + 2f^2 + 2y^2 + 2\sqrt{(x^2 - f^2)^2 + 2x^2y^2 + 2f^2y^2 + y^4} = 4a^2$$

$$\sqrt{(x^2 - f^2)^2 + 2x^2y^2 + 2f^2y^2 + y^4} = 2a^2 - f^2 - x^2 - y^2$$

$$x^4 - 2x^2f^2 + f^4 + 2x^2y^2 + 2f^2y^2 + y^4$$
$$= 4a^4 - 4a^2f^2 - 4a^2x^2 - 4a^2y^2 + f^4 + 2f^2x^2 + 2f^2y^2 + x^4 + 2x^2y^2 + y^4$$

$$-2x^2f^2 = 4a^4 - 4a^2f^2 - 4a^2x^2 - 4a^2y^2 + 2f^2x^2$$

$$4a^2x^2 + 4a^2y^2 - 4f^2x^2 = 4a^4 - 4a^2f^2$$

$$(a^2 - f^2)x^2 + a^2y^2 = a^2(a^2 - f^2)$$

$$\frac{x^2}{a^2} + \frac{y^2}{(f^2 - a^2)} = 1$$

$(f^2 - a^2) = b^2$とします。

$$f^2 = a^2 + b^2$$

$$f = \pm\sqrt{a^2 + b^2}$$

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$

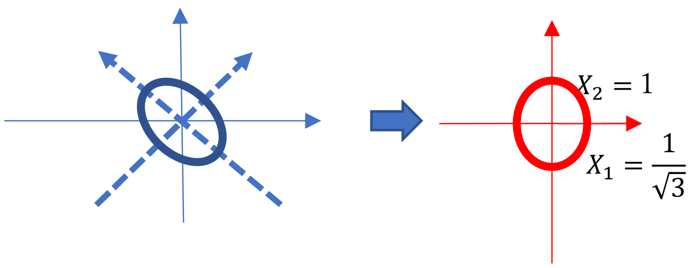

図 7.　正定値行列の直交化

この二次形式の形は楕円です。すべての固有値が正の場合、その行列を正定値だと表現します。この場合 2 次元なので楕円です。多次元の場合、すべての固有値が正の時、正定値と言います。正定値のもう一つ別の定義は、２次形式に表した式が、常に正であることです。第一の定義では、

$$\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

$$\lambda_1 = 3 > 0, \qquad \lambda_2 = 1 > 0$$

だから正定値です。第二の定義では、

$$(x_1 \quad x_2)\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = (2x_1 + x_2 \quad x_1 + 2x_2)\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 2{x_1}^2 + 2x_1x_2 + 2{x_2}^2 == 3\left(\frac{x_1 + x_2}{\sqrt{2}}\right)^2 + \left(\frac{x_2 - x_1}{\sqrt{2}}\right)^2 > 0$$

だから正定値です。第一の定義と第二の定義は実は同じことです。

この場合、２次元ですから、第二の定義でも正定値であることがわかりますが、多次元になると、第二の定義で正定値であることを証明するのは難しいでしょう。多次元の場合、正定置であれば、その次元の空間の中で、楕円球のラグビーボールのように空間的に閉じた形になります。二次形式の対角化とは、空間中に斜めに角度を持っておかれたラグビーボールを回転して、空間軸とラグビーボールの軸を一致させることだとも言えるでしょう。

V-2-4-3. スペクトル分解と二次形式

二次形式のスペクトル分解は、二次形式の対角化の応用です。$\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x} = \lambda_1\boldsymbol{x}^T\boldsymbol{e}_1{\boldsymbol{e}_1}^T\boldsymbol{x} + \lambda_2\boldsymbol{x}^T\boldsymbol{e}_2{\boldsymbol{e}_2}^T\boldsymbol{x} + \cdots \lambda_p\boldsymbol{x}^T\boldsymbol{e}_p{\boldsymbol{e}_p}^T\boldsymbol{x}$

$\boldsymbol{x}^T\boldsymbol{e}_i$ という行列計算は線形です。$\boldsymbol{x}^T\boldsymbol{e}_i{\boldsymbol{e}_i}^T\boldsymbol{x}$ は線形同士をかけあわせた二次形式です。これは、次の式の変形からもわかります。

$$(x_1 \quad x_2)\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 3\left(\frac{x_1 + x_2}{\sqrt{2}}\right)^2 + \left(\frac{x_2 - x_1}{\sqrt{2}}\right)^2$$

また、$\boldsymbol{x}^T\boldsymbol{e}_i{\boldsymbol{e}_i}^T\boldsymbol{x}$は固有ベクトルの平方です。このことから、スペクトル分解とは、二次形式を固有ベクトルの平方の和として表すことだともいえるでしょう。

ここで、覚えておくべきことは、「対称行列は 2 次形式を作る行列で、その固有ベクトルは、２次形式化した式がしめす。円錐曲線の軸になっていて、互いに直交している。」ということで。ちなみに、分散共分散行列、相関行列は対称行列で 2 次形式を作ります。しかも、正定置ですから、もしそれが、実数の値を持つならば、そのイメージは楕円です。その行列が変換して投影するための写像の装置だとすれば、直交する超楕円のそれぞれの方向へ、引き延ばすという投影の仕方になります。

主成分分析

データは、もともとの実態を何かのベクトルへ映し出した写像として描かれています。主成分分析はその写像を固有ベクトルへの写像に作り変えることだと、数学的には言えます。それは、もとのデータの座標を含み固有ベクトルと直交する平面と原点への距離を計算することにほかなりません。

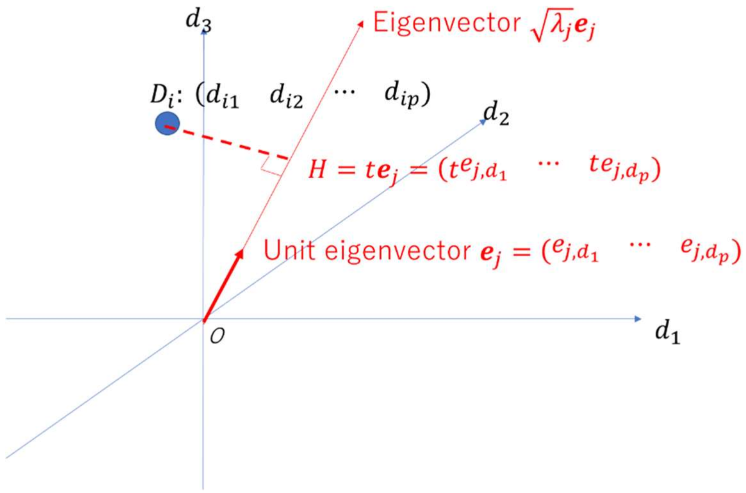

図 8.観測された変数の座標から固有ベクトルの座標への変換

図８にその座標変換のデータと固有ベクトルの関係を示しました。標本サイズがn、観測変数がp個として、データをD_i: $(d_{i,1}\quad \cdots\quad d_{i,p})$ $((i = 1, \cdots, n)$として、単位行列化した固有ベクトルを $\boldsymbol{e}_j$: $(e_{j,d_1}\quad \cdots\quad e_{j,d_p})$, とします。 Hは原点から平面に下した垂線の脚で、ベクトル$\overrightarrow{OH}$長さがtの固有ベクトルです。ここでは j 番目の固有ベクトル（j 番目の固有ベクトル）について示しました。図には長さが標準偏差の固有ベクトルも示しました($\sqrt{\lambda_j}\boldsymbol{e}_j$).

$$\overrightarrow{OH} = t\boldsymbol{e}_j = (te_{j,x_1}\quad \cdots\quad te_{j,x_p})$$

$$\overrightarrow{HD_\iota} = \left(d_{i,1} - te_{j,x_1}\quad \cdots\quad d_{i,p} - te_{j,x_p}\right)$$

$$\overrightarrow{HD_\iota} \perp \overrightarrow{OH}$$

$$\text{内積}\ \overrightarrow{HD_\iota} \cdot \overrightarrow{OH} = 0$$

$$\left(d_{i,1} - te_{j,x_1}\quad \cdots\quad d_{i,p} - te_{j,x_p}\right)\begin{pmatrix} te_{j,d_1} \\ \vdots \\ te_{j,d_p} \end{pmatrix} = td_{i,1}e_{j,d_1} - t^2 {e_{j,d_1}}^2 + \cdots + td_{i,p}e_{j,d_p} - t^2 {e_{j,d_p}}^2$$

$$= t\left(d_{i,1}e_{j,d_1} - t{e_{j,d_1}}^2 + \cdots + d_{i,p}e_{j,d_p} - t{e_{j,d_p}}^2\right)$$

$$= \ t\left(d_{i,1}e_{j,d_1} + \cdots + d_{i,p}e_{j,d_p} - t\left({e_{j,d_1}}^2 + \cdots + {e_{j,d_p}}^2\right)\right)$$

$$= \ t\left(d_{i,1}e_{j,d_1} + \cdots + d_{i,p}e_{j,d_p} - t\right)$$

$\because$ $\boldsymbol{e}_j$ は単位ベクトル

$${e_{j,d_1}}^2 + \cdots + {e_{j,d_p}}^2 = 1$$

$$t\left(t - d_{i,1}e_{j,d_1} + \cdots + d_{i,p}e_{j,d_p}\right) = 0$$

$$t = 0\ or\ t = d_{i,1}e_{j,d_1} + \cdots + d_{i,p}e_{j,d_p}$$

前提条件から

$$t \neq 0$$

したがって

$$t = d_{i,1}e_{j,d_1} + \cdots + d_{i,p}e_{j,d_p}$$

これがデータi の j 番目の主成分の主成分得点です。

$$\mathrm{PCS}_{i,j} = t_{ij} = d_{i,1}e_{j,d_1} + \cdots + d_{i,p}e_{j,d_p}$$

このような計算で以下の主成分得点表が作れます.

	主成分得点			
	PC1	PC2	⋯,	PCp
標本番号				
1	$\mathrm{PCS}_{1,1}$	$\mathrm{PCS}_{1,2}$	⋯	$\mathrm{PCS}_{1,p}$
2	$\mathrm{PCS}_{2,1}$	$\mathrm{PCS}_{2,2}$	⋯	$\mathrm{PCS}_{2,p}$
⋮	⋮	⋮	⋱	⋮
n	$\mathrm{PCS}_{n,1}$	$\mathrm{PCS}_{n,2}$	⋯	$\mathrm{PCS}_{n,p}$

これらをもとに２つあるいは３つの主成分を直交軸として分布図を書くことが出来ます。

VI-2-1-3. 分散共分散行列の対角化と主成分分析

二つの主成分分析の方法があります。データが同じでも、この二つの主成分分析の結果とその解釈は異なります。分散共分散行列を対角化するのが一つの方法で、もう一つは相関行列を対角化します。分散共分散行列も相関行列も対称行列で二次形式です。そのような行列の対角化によって得られた固有値が、上記の説明で使った固有値であることを示します。その過程で、二次形式の行列の空間幾何学と対称行列の性質を使います。行列が正定置であれば、二次形式の行列は、多次元空間において傾いた超楕円を表し、固有ベクトルは超楕円の軸を表しています。対称行列($\boldsymbol{P}$)の性質、その対角化行列の転置行列($\boldsymbol{P}^T$)が逆行列($\boldsymbol{P}^{-1}$)だということも使います。具体的には、次の式を使います。

一般的な対角化の式

$$\boldsymbol{P}^{-1}\boldsymbol{V}\boldsymbol{P} = \boldsymbol{\Lambda}$$

$\boldsymbol{V}$が二次形式ならば

$$\boldsymbol{P}^T = \boldsymbol{P}^{-1}$$

したがって$\boldsymbol{V}$の対角化は次の式になります。

$$\boldsymbol{P}^T\boldsymbol{V}\boldsymbol{P} = \boldsymbol{\Lambda}$$

そもそも、分散共分散行列は次のように作ります。

$$\boldsymbol{V} = \boldsymbol{D}\boldsymbol{D}^{\boldsymbol{T}}$$

$$\boldsymbol{P}^T\boldsymbol{V}\boldsymbol{P} = \boldsymbol{P}^T\boldsymbol{D}\boldsymbol{D}^{\boldsymbol{T}}\boldsymbol{P} = (\boldsymbol{P}^T\boldsymbol{D})(\boldsymbol{P}^T\boldsymbol{D})^{\boldsymbol{T}}$$

$$\boldsymbol{P}^T\boldsymbol{D} = \begin{pmatrix} \boldsymbol{e_1}^{\boldsymbol{T}} \\ \boldsymbol{e_2}^{\boldsymbol{T}} \\ \vdots \\ \boldsymbol{e_p}^{\boldsymbol{T}} \end{pmatrix} \begin{pmatrix} d_{11} & d_{21} & \cdots & d_{n1} \\ d_{12} & d_{22} & \cdots & d_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ d_{1p} & d_{2p} & \cdots & d_{np} \end{pmatrix}_{p\times n}$$

$$= \begin{pmatrix} e_{11} & \cdots & e_{j1} & \cdots & e_{p1} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ e_{1k} & \cdots & e_{jk} & \cdots & e_{pk} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ e_{1p} & \cdots & e_{jp} & \cdots & e_{pp} \end{pmatrix} \begin{pmatrix} d_{11} & d_{21} & \cdots & d_{n1} \\ d_{12} & d_{22} & \cdots & d_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ d_{1p} & d_{2p} & \cdots & d_{np} \end{pmatrix}_{p\times n}$$

$$= \begin{pmatrix} \sum_{j=1}^{p} e_{j1}d_{1j} & \sum_{j=1}^{p} e_{j1}d_{2j} & \cdots & \sum_{j=1}^{p} e_{j1}d_{nj} \\ \sum_{j=1}^{p} e_{j2}d_{1j} & \sum_{j=1}^{p} e_{j2}d_{2j} & \cdots & \sum_{j=1}^{p} e_{j2}d_{nj} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{j=1}^{p} e_{jp}d_{1j} & \sum_{j=1}^{p} e_{jp}d_{2j} & \cdots & \sum_{j=1}^{p} e_{jp}d_{nj} \end{pmatrix}_{p\times n}$$

$e_{jk}d_{ij}$ を $d_{i,j}e_{j,k}$ と書き換えます。

$d_{i,j}e_{k,d_1} + \cdots + d_{i,p}e_{k,d_p} = \sum_{j=1}^{p} e_{jk}d_{ij} = t_{ik} = \mathrm{PCS}_{i,k}$

（これは、固有ベクトルと直交する、点$\boldsymbol{d}$を含む超平面と原点の距離）

$$\boldsymbol{P}^T\boldsymbol{D} = \begin{pmatrix} t_{11} & t_{21} & \cdots & t_{n1} \\ t_{12} & t_{22} & \cdots & t_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ t_{1p} & t_{2p} & \cdots & t_{np} \end{pmatrix}_{p\times n} = (\boldsymbol{t_1} \quad \boldsymbol{t_2} \quad \cdots \quad \boldsymbol{t_n})$$

$$\boldsymbol{t_i} = \begin{pmatrix} t_{i1} \\ t_{i2} \\ \vdots \\ t_{ip} \end{pmatrix} = \begin{pmatrix} \mathrm{PCS}_{i,1} \\ \mathrm{PCS}_{i,2} \\ \vdots \\ \mathrm{PCS}_{i,p} \end{pmatrix}$$

$$\boldsymbol{D}^T\boldsymbol{P} = (\boldsymbol{P}^T\boldsymbol{D})^{\boldsymbol{T}} = \begin{pmatrix} t_{11} & t_{12} & \cdots & t_{1p} \\ t_{21} & t_{22} & \cdots & t_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ t_{n1} & t_{n2} & \cdots & t_{np} \end{pmatrix}_{n\times p} = \begin{pmatrix} \boldsymbol{t_1}^T \\ \boldsymbol{t_2}^T \\ \vdots \\ \boldsymbol{t_2}^T \end{pmatrix}$$

$$\boldsymbol{P}^T\boldsymbol{V}\boldsymbol{P} = \boldsymbol{P}^T\boldsymbol{D}\boldsymbol{D}^T\boldsymbol{P} = \begin{pmatrix} t_{11} & t_{21} & \cdots & t_{n1} \\ t_{12} & t_{22} & \cdots & t_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ t_{1p} & t_{2p} & \cdots & t_{np} \end{pmatrix}_{p\times n} \begin{pmatrix} t_{11} & t_{12} & \cdots & t_{1p} \\ t_{21} & t_{22} & \cdots & t_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ t_{n1} & t_{n2} & \cdots & t_{np} \end{pmatrix}_{n\times p}$$

$$
= \begin{pmatrix} \sum_{i=1}^{n} {t_{i1}}^2 & \sum_{i=1}^{n} t_{i1}t_{i2} & \cdots & \sum_{i=1}^{n} t_{i1}t_{1p} \\ \sum_{i=1}^{n} t_{i2}t_{i1} & \sum_{i=1}^{n} {t_{i2}}^2 & \cdots & \sum_{i=1}^{n} t_{i2}t_{1p} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^{n} t_{1p}t_{i1} & \sum_{i=1}^{n} t_{ip}t_{i2} & \cdots & \sum_{i=1}^{n} {t_{i2}}^2 \end{pmatrix}_{p\times p}
$$

二次形式の対称行列の固有ベクトルは互いに直交しています。したがって、その直交ベクトルに投影したベクトル同士も直交しています。

$$
\sum_{i=1}^{n} t_{ij}t_{ik} = \delta_{jk} \sum_{i=1}^{n} t_{ij}t_{ik}
$$

$$
\delta_{jk} = \{ \begin{matrix} 1\ (j = k) \\ 0\ (j \neq k) \end{matrix}
$$

δ_{jk} は クロネッカーのデルタ

$$
\boldsymbol{P}^T\boldsymbol{V}\boldsymbol{P} = \boldsymbol{P}^T\boldsymbol{D}\boldsymbol{D}^T\boldsymbol{P} == \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_p \end{pmatrix}_{p\times p}
$$

以上のように、λ_i は主成分iの平方和（固有ベクトル i 上のデータの平方和)です。

主成分分析を要約すると、分散共分散行列または相関行列の対角化で、これらの行列は、対称行列で正定置だから、固有ベクトルが直交していて、それぞれの固有ベクトル方向に固有値倍だけ引き延ばすような円環をする行列で、その行列の直交軸に、それ俺のデータを投影させる分析ということです、その直交化のための行列が、主成分負荷量で、投影されたベクトルが、主成分得点になるということです。

VI-2-1-4. 主成分分析の結果の記述.

主成分分析の意義の一つはデータの集約です。主成分分析の計算では、元データの変数の数と同じ数の主成分が出来ます。しかし、いくつかの主成分は、元のデータの変数の分散に比べて大きく、またある主成分は、元のデータの分散に比べて分散が小さいはずです。現象に対する影響の小さい主成分について考える必要はあまりないでしょう。そこで、まず、重要な主成分を選び出します。具体的には、全分散に占める割合の大きい分散を持つ主成分を選ぶのです。主成分の分散は固有値λです。全分散は、対角成分のトレース、つまり、対角因子の和です。

$$
V_{total} = \sum_{j=1}^{p} \lambda_j
$$

	寄与率	累積寄与率
PC1	$\frac{\lambda_1}{V_{total}}$	$\frac{\lambda_1}{V_{total}}$
PC2	$\frac{\lambda_2}{V_{total}}$	$\frac{\lambda_1+\lambda_2}{V_{total}}$
$\vdots$	$\vdots$	$\vdots$
PCp	$\frac{\lambda_p}{V_{total}}$	$\frac{\lambda_1+\lambda_2+\cdots+\lambda_p}{V_{total}}$

例えば、現象の 70%まで説明したいときには、累積寄与率 0.7 までに含まれる主成分だけを取り上げて、残りは意味のない変動として取り扱います。このやり方は単純ですが少し機械的すぎて、実際上困ります。たとえば、0.7 の周辺に、小さな分散の主成分がたくさん存在した場合、どうしたらよいか判断できません。よくある別の方法は、スクリー・プロットというやり方です。スクリー・プロットとは図 74 に示したように、固有値の大きいものから順番に折れ線グラフを書くという方法です。

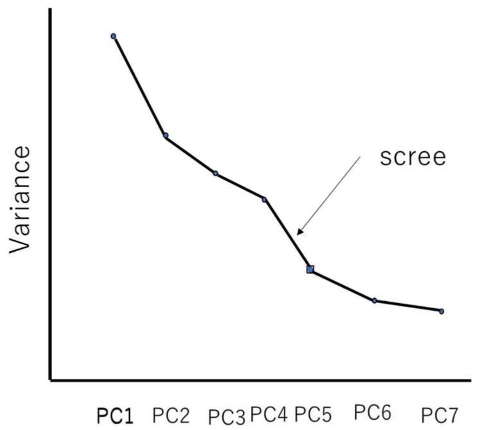

図 9. 主成分の分散のスクリー・プロット

第 4 主成分と第 5 主成分の間に大きな差があります。ここがスクリー（崖）になっています。この前後で考慮すべき主成分とそうでない主成分を分けて、第 5 主成分以下を切り捨てます。

この選択法は曖昧さを含んでいます。どこにスクリーが出来るかは場合によるからです。

主成分分析を相関行列から始めた場合には、もう少し数学的な方法が考えられます。

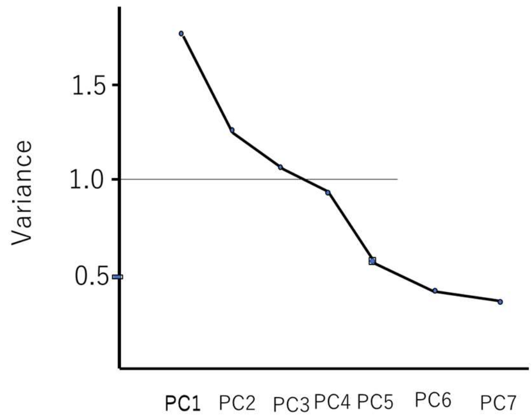

図 10. 相関行列の主成分分析のスクリー・プロット

相関行列の主成分分析では、すべての変数の分散が 1 です。相関行列の主成分分析は、すべての変数の分散を主成分に振り分けますが、その平均は 1 のままです。このことは、分散が 1 より大きい主成分は、分散を吸収する影響力の強い主成分です。反対に分散が 1 より小さい主成分は、盈虚力の小さな主成分です。ですから、分散が 1 より大きな主成分を選ぶというのも一つの選択です（図 10 参照）。

実際に公開されているソフトウエア―は結果の解釈のためにいくつかの機能が付け加えられています。最も一般的な主成分の解釈のための指標は主成分負荷量です。この指標は、主成分と実際に観測された変数の値の関係の強さを表しています

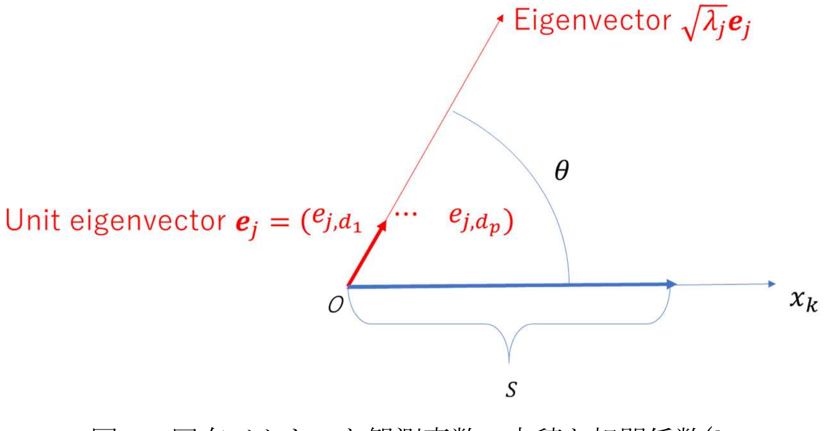

図 11. 固有ベクトルと観測変数の内積と相関係数$(k = 1, \cdots p)$

図 11 に固有ベクトルと元のデータ変数のベクトルを示しました。この図では、固有ベクトルの長さを固有値の平方根と S にしました。実際には、長さを$\sqrt{\lambda_j}$に固定する必要はありません。任意の実数で良いのです。ベクトルの関係性とは相関係数で相関係数は二つのベクトルがなす角度だからです。ここでことさら$\sqrt{\lambda_j}$ にしたのは、主成分負荷量と関係づけるためです。

内積には次の二つの定義がありました。

$$\boldsymbol{V_1}\cdot\boldsymbol{V}_2=\lceil\boldsymbol{V}_1\rceil\lceil\boldsymbol{V_2}\rceil\cos\theta=(v_{1,1}\quad\cdots\quad v_{1,n})(v_{2,1}\quad\cdots\quad v_{2,n})^T$$
$$\boldsymbol{V_1}=(v_{1,1}\quad\cdots\quad v_{1,n}),\qquad \boldsymbol{V}_2=(v_{2,1}\quad\cdots\quad v_{2,n})$$

$\boldsymbol{V_1}$と$\boldsymbol{V}_2$ の相関係数は$\cos\theta$です。
図 11 に示した固有ベクトルと変数のベクトルの場合には内積は以下の通りです。

$$\sqrt{\lambda_j}S\cos\theta_{j,k}=\left(\sqrt{\lambda_j}e_{j,d_1}\quad\cdots\quad\sqrt{\lambda_j}e_{j,d_p}\right)\begin{pmatrix}0\\ \vdots\\ S\\ \vdots\\ 0\end{pmatrix}=\mathrm{S}\sqrt{\lambda_j}e_{j,d_k}$$

$$\cos\theta=e_{j,d_k}$$

相関係数は e_{j,d_j}だということになります。
主成分負荷量とは(PCL)、一つの変数の固有ベクトルに対する相関の大きさです。その大きさを、主成分の偏差の長さのベクトルに変数のベクトルを投影したときの、投影された部分の大きさだと考えると、次のような式になります。

$$\mathrm{PCL}_{j,k}=\sqrt{\lambda_j}r_{j,k}=\sqrt{\lambda_j}e_{j,d_k}$$

これを使って次のような表ができます。

	主成分負荷量			
	PC1	PC2	⋯,	PCp
偏差	$\sqrt{\lambda_1}$	$\sqrt{\lambda_2}$	⋯	$\sqrt{\lambda_p}$
変数				
変数 1	$\sqrt{\lambda_1}e_{1,d_1}$	$\sqrt{\lambda_2}e_{2,d_1}$	⋯	$\sqrt{\lambda_p}e_{p,d_1}$
変数 2	$\sqrt{\lambda_1}e_{1,d_2}$	$\sqrt{\lambda_2}e_{2,d_2}$	⋯	$\sqrt{\lambda_p}e_{p,d_2}$
⋮	⋮	⋮	⋱	⋮
変数 p	$\sqrt{\lambda_1}e_{1,d_p}$	$\sqrt{\lambda_2}e_{2,d_p}$	⋯	$\sqrt{\lambda_p}e_{p,d_p}$

これで、主成分と各変数の関係はわかりますが、変数間で$\sqrt{\lambda_i}$を比較することの意味については考える必要があります。変数に異なる最小単位で測られたデータが含まれていた場合、 $\sqrt{\lambda_i}$ は小さな単位で測られたデータで大きくなります。もう少し理論的に表現すると、変数の分散は変数間で異なります。おそらく、分散が大きく異なった主成分の間で、分散の大きさを比較しても意味がないでしょう。主成分分析には、分散共分散行列から固有値を求めるやり方と、相関行列から固有値を求めるやり方があります。二つのやり方で、異なる結果が出ます。分散共分散行列の主成分分析は標準化されていないデータの主成分分析です。相関行列の主成分分析は標準化されたデータの主成分分析です。分析の目的が違うのです。分散共分散行列の主成分分析でも、分散の違いが無視できないのであれば、次の表を作った方が良いかもしれません。

主成分と変数の相関行列

	PC1	PC2	$\cdots$,	PCp
Variable				
Variable 1	e_{1,d_1}	e_{2,d_1}	$\cdots$	e_{p,d_1}
Variable 2	e_{1,d_2}	e_{2,d_2}	$\cdots$	e_{p,d_2}
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
Variable n	e_{1,d_p}	e_{2,d_p}	$\cdots$	e_{p,d_p}

計算過程で示したように、e_{j,x_k}は相関係数です。

$$e_{j,d_k} = r_{jk}$$

この表を視覚化して表す一つの方法は、図 77 のような図を作ることです。図中の円は、PCa − PCb平面で切った超球の切断面です。半径は 1 です。

Vc は標準偏差の長さの変数 c ベクトルの、PCa − PCb平面への投影で、PCa − PCbの座標はPCa と PCbとの相関係数です。

$$\mathrm{V}c = (r_{ac} \quad r_{bc})$$

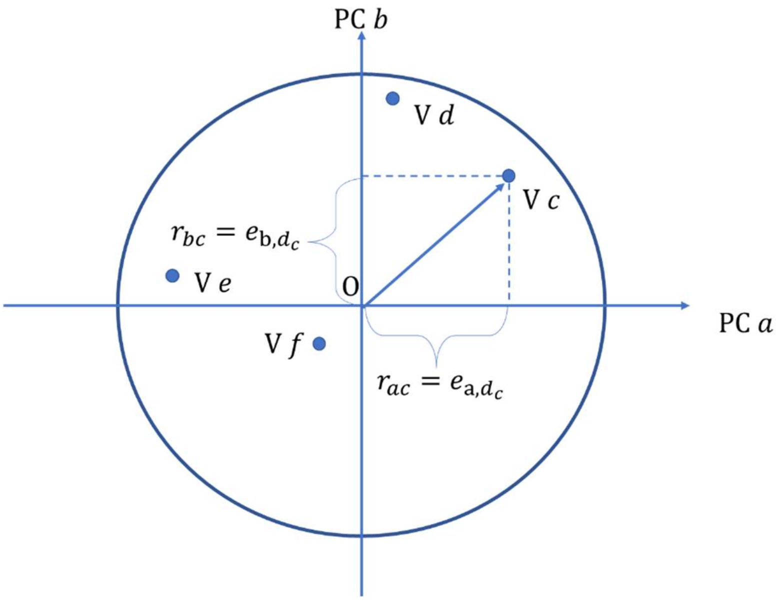

図 77. 二つの主成分の変数への貢献度の視覚的表現

$$\left|\overrightarrow{\mathrm{OV}c}\right|^2 = {r_{ac}}^2 + {r_{bc}}^2$$

$$\left|\overrightarrow{\mathrm{OV}c}\right| = \sqrt{{r_{ac}}^2 + {r_{bc}}^2}$$

$\left|\overrightarrow{\mathrm{OV}c}\right|^2$はPC$a$ と PCbの貢献度です。 Vcが、PCa と PCbで完全に説明できるのならば、

$$\left|\overrightarrow{\mathrm{OV}c}\right|^2 = {r_{ac}}^2 + {r_{bc}}^2 = 1$$

となり、ベクトル$\overrightarrow{\mathrm{OV}c}$の長さは、

$$\left|\overrightarrow{\mathrm{OV}c}\right| = \sqrt{{r_{ac}}^2 + {r_{bc}}^2} = 1$$

この図では、$\left|\overrightarrow{\mathrm{OV}c}\right|$の長さは、0.8 ぐらいでしょう。ですから、この変数の分散への主成分a 主成分bの寄与率は $0.8^2 = 0.64$ぐらいです。これは、Vcの分散の半分以上が主成分a と主成分bで説明できるということですから、無視できません。そこで、私たちは主成分aと主成分bは、Vcを説明する主要な主成分だと判定します。また、主成分a と主成分bの寄与率はほぼ同じぐらいです。Vdは円周に近いのですが、r_{ad}が小さいので、Vdは主として主成分bによって説明できると推定します。Veも円周に近いのですが、r_{ae}の値が負です。また、 r_{be}は小さな値です。ですから、この変数は主成分aに逆相関しています。Vfは原点に近づいています。このことは、Vfが主成分a と 主成分bにあまり関係がないことを示しています。 VfはPCa − PCbある角度をもって交差しているのです。したがって、このベクトルについては、別の平面に投影して関係性を考える必要があります。 最も簡単な、結果の解釈の方法は、二つの主成分を座標軸とする分布図に、変数のベクトルを投影してみることです。普通は、比較のために、変数のベクトルの長さを標準偏差にします。

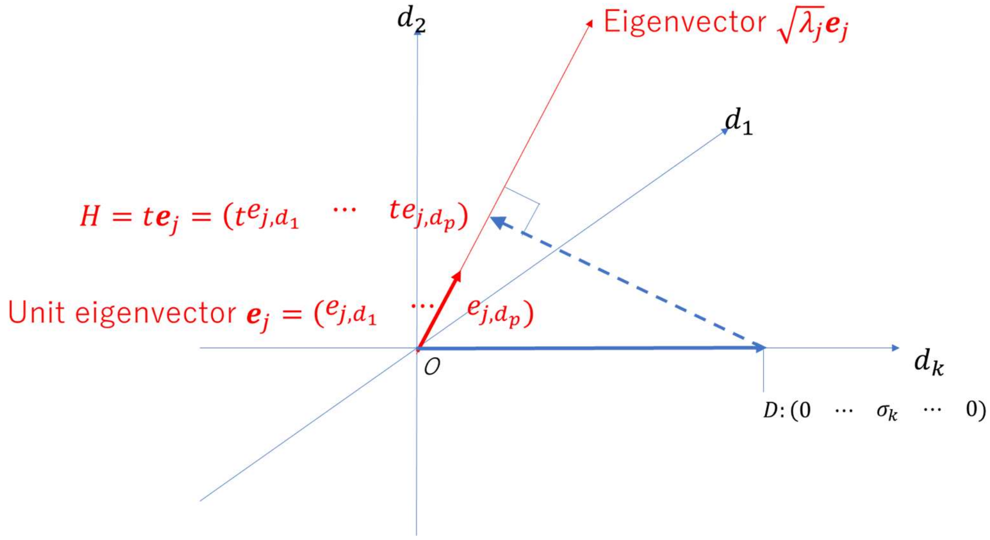

図 12. 変数ベクトルx_kの主成分ベクトルへの投影

変数ベクトルの投影の計算は主成分のベクトルへの座標変換と同じです。変数ベクトルの長さは標準偏差σ_kです。その座標は D=$(0 \quad \cdots \quad \sigma_k \quad \cdots \quad 0)$.

$$\overrightarrow{\mathrm{OH}} \perp \overrightarrow{\mathrm{DH}}$$

$$\overrightarrow{\mathrm{DH}} = (te_{j,d_1} \quad \cdots \quad te_{j,d_p}) - (0 \quad \cdots \quad \sigma_k \quad \cdots \quad 0) = \left(te_{j,d_1} \quad \cdots \quad te_{j,d_k} - \sigma_k \quad \cdots \quad te_{j,d_p}\right)$$

$$\overrightarrow{\mathrm{OH}} \cdot \overrightarrow{\mathrm{DH}} = 0$$

$$\boldsymbol{e}_j \cdot \overrightarrow{\mathrm{DH}} = 0$$

$$
\begin{pmatrix} e_{j,d_1} & \cdots & e_{j,d_p} \end{pmatrix}
\begin{pmatrix} te_{j,d_1} \\ \vdots \\ te_{j,d_k} - \sigma_k \\ \vdots \\ te_{j,d_p} \end{pmatrix} = 0
$$

$$
t\left({e_{j,d_1}}^2 + \cdots + {e_{j,d_p}}^2\right) - \sigma_k e_{j,d_k} = 0
$$

$$
t = \sigma_k e_{j,d_k}
$$

$$
\because {e_{j,x_1}}^2 + \cdots + {e_{j,x_p}}^2 = 1
$$

これらを使えば次の表が作れます。

	主成分					
変数	PC1	PC2	$\cdots$	PCk	$\cdots$	PCp
変数 1	$\sigma_1 e_{1,d_1}$	$\sigma_1 e_{2,d_1}$	$\cdots$	$\sigma_1 e_{k,d_1}$	$\cdots$	$\sigma_1 e_{p,d_1}$
変数 2	$\sigma_2 e_{1,d_2}$	$\sigma_2 e_{2,d_2}$	$\cdots$	$\sigma_2 e_{k,d_2}$	$\cdots$	$\sigma_2 e_{p,d_2}$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\ddots$	$\vdots$
変数 k	$\sigma_k e_{1,d_k}$	$\sigma_k e_{2,d_k}$	$\cdots$	$\sigma_k e_{k,d_k}$	$\cdots$	$\sigma_k e_{p,d_k}$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\ddots$	$\vdots$
変数 p	$\sigma_p e_{1,d_p}$	$\sigma_p e_{2,d_p}$	$\cdots$	$\sigma_p e_{k,d_p}$	$\cdots$	$\sigma_p e_{p,d_p}$

この表から、適当な２つの主成分を選び出して、二つの主成分のベクトルが作る平面上に、標準偏差の長さの変数ベクトルを投影することが出来ます。たとえば、PC1-PC2 平面に投影すると、それぞれの変数の PC1-PC2 の座標は以下のようになります。

Variable 1: $(\sigma_1 e_{1,d_1} \quad \sigma_1 e_{2,d_1})$

Variable 2: $(\sigma_2 e_{1,d_2} \quad \sigma_2 e_{2,d_2})$

$\vdots \qquad \vdots$

Variable k: $(\sigma_k e_{1,d_k} \quad \sigma_k e_{2,d_k})$

$\vdots \qquad \vdots$

Variable p: $(\sigma_p e_{1,d_p} \quad \sigma_p e_{2,d_p})$

これを、PC1-PC2 平面上のデータの分布に重ね合わせたのが図 79 です。

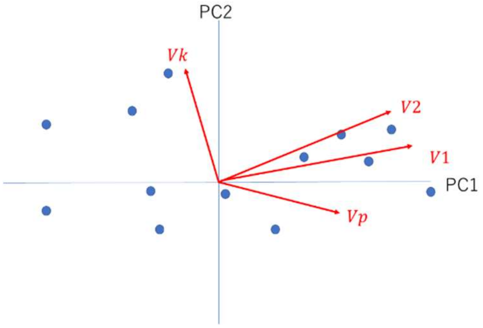

図 13. データの散布図と元の変数ベクトルの写像の重ね合わせの例

この操作は、ほぼ、対応分析（Correspondence analysis）と同じです。それはそれとして、変数 1 と変数 2 のベクトルの向きはほぼ主成分 1 と一致していて、十分な長さがあります。いつでもそうなるのではありませんが、主成分 1 は何かその現象や出来事の大きさにかかわる成分であることが多く、質的な特性、たとえば、差異、比率、不安定性、曖昧さのような質的な性質は第二主成分以降に来ることが多いようです。もし、主成分 1 が量的な特性を表しているのであれば、変数 1、変数 2 はともに大きさにかかわる変数のはずです。このような同一性を内的整合性あるいは内部的な一貫性といいます。内的一貫性は主成分分析ではあまり重要な意味を持たないのですが、因子分析、特に質問票を使った社会心理の調査の分析などでは重要な概念です。複数の質問に対する回答が高い相関性を持っているときには、それらの質問は、同じ内容のものを別の表現で訊いたことになるからです。これは、同じ内容を表す心理的傾向を違った角度から調べる必要がある心理学の分野では重要です。しかし、内的整合性は、他の分野では否定的にとらえられることがあります。同じ質問の繰り返しは無駄が多く、回答者に負担をかけて、質問票調査の質を低下させるからです。内的整合性は、クロンバックのα値（Cronbach's α value）で評価します。高いクロンバックのα値が得られた時、質問票をできるだけコンパクトにするために、次の質問票調査では、代表的な少数の質問に絞り込んだり、反対に、内的整合性を生かして、分析するときに変数を結合して、より感度の良い指標を作って分析します。

VI-2-1-5. 主成分分析の結果の解釈

主成分分析の数学的な意味は分かりやすいのですが、その結果の解釈はしばしば難しいことになります。主成分分析で何をしているかという説明は様々あるのですが、それは分析の目的によります。一番一般的な説明は、主成分分析は、データセットを構成している構成要素をいくつかの主要な構成要素にまとめているのだという説明です。専門知識があれば、主成分分析を使わなくても知識と経験で、すべて因子の中から代表的な因子を選び出すことが出来ます。現象の背景にある潜在的な因子を見つけることが主成分分析の目的だと説明することもできます。確かに、主成分分析によって、現象の背景にある仕組みを見つけ出せることがありますが、そういう運のよい例はまれです。主成分分析は、構成要素の軸（主成分）で結果を表現します。それぞれの軸は直交していて独立しています。その軸が示すものは時には何かの大きさだったり相違だったりしますが、その軸の意味は、しばしば、我々の日常言語では意味が解りません。結果の解釈は因子分析（Factor analysis：FA）の方が容易です。特に、斜交回転を使った因子分析は結果の解釈が自然に無理なくできます。斜交回転を使った因子分析のように、科学的な事前情報や日常生活の経験に合う分かりやすいベクトルに一致するように、独立性を無視して軸を回転したくなるのは自然なことです。似たような分析に見えますが、主成分分析と因子分析の目的は全く異なります。主成分分析は、現象の構造を理解する方法として重要です。主成分分析は、

多次元空間中のデータを変数間の相関を取り除いて別の空間に移し替えます。その本質はスペクトル分解です。
生物の分布は環境によって決まります。一方、環境の物理的要素は水温と塩分のように本来は独立しています。しかし、沿岸帯で塩分と水温を測定すると、実際には相関があります。たとえば、夏場に河口から沖に向けて測点を作って、塩分と水温を測定すると、水温は河口から離れるにしたがって減少し、塩分は増加します。その結果、塩分と水温には明瞭な逆相関の関係が生まれます。同様に、リン酸や硝酸濃度のような化学環境の測定項目についても相関が生まれます。生物は、そのような環境因子の組み合わせに適応しているのです。生物の分布を一つの水質項目という要素に還元して説明してもあまり意味はありません。もちろん、要素還元主義は生物の生理的な反応の理解には有効です。しかし、単純な要素還元主義では、行動や生物の進化メカニズムなど複雑なシステムの発見にはあまり有効ではないでしょう。複雑系を理解するためには、現象と関係がないかもしれない変数も含めて、網羅的なデータの蓄積が必要です。たとえば、何の事前情報もなく、沿岸における魚種の分布がどのように決まっているのかを知ろうとすれば、様々な場所で魚を捕獲し、魚種名、個体数、体重、体長など生物データと水深、塩分、水温、透明度。底質、サンプリング時間など、様々な物理・化学的データを集めるしかないでしょう。その中のいくつかの変数は相関し、別のものは逆相関し、残りは関係がないでしょう。確かに、このやり方は、よくデザインされているとは言えませんが、事前情報がなければ、こういう見通しのないやり方をとるしかありません。これは、仮説検証ではありません。科学とは仮説検証だと考える人もいます。これは誤りです。科学は、まず、無原則に積み重なったデータセットの中から、仮説を作らなくてはなりません。その目的のために、分析者は変数ごとに、データの分布を確かめるために頻度分布を作ります。ある分析者は、X-Y プロットを作って変数間の関係を見つけようとします。別の分析者は相関分析をします。最も一般的には、分散共分散行列、あるいは、相関行列を作って、潜在的な相関関係を見つけます。まだ、コンピュータが今のように普及していなかった頃のことです。筆者は、日本沿岸の養殖ノリの生産量を決める因子について分析しました。50 年以上にわたり、日本各地の海苔養殖場の単位面積当たりの生産量と価格のデータをあつめて、生産量と価格について、地域間の総あたりの行列をつくりました。そして、実際の地図上で、相関する地域を色で塗り分けたのです。これには２か月以上かかりました。その結果、ノリの生産にかかわる重要の要素を見つけることが出来ました。著者は、遠く離れた地域間で、同調的に変動する地域を見つけたのです。そして、それがなぜ同調するのか考えた結果、背景にある要因を見つけました。そのやり方は洗礼されていない稚拙なやり方です。もはやそんなやり方をする人はいないでしょう。その時点では、著者は主成分分析を知りませんでした。仮に知っていたとしても、コンピュータを持っていなかったので計算できなかったでしょう。しかし、やっていることは主成分分と同じです。今ならば、若い時の自分自身に、環境データを含めて、すべてのデータを入れて主成分分析をすることを薦めるでしょう。主

成分分析の最も重要な機能は、潜在的な背景の仕組みを含めて、データ間の関係を視覚化して見せることです。無知で若かった筆者は、背景の仕組みが見つかることを漠然と期待して、総当たりの相関行列を作りました。相関はいつでも因果関係を示しているとは限りませんし、遠く離れたノリ養殖漁場間に因果関係が存在する可能性はほとんどないでしょう。若かった筆者は、日本の沿岸環境とノリの生理学を知っていたので、遠く離れた養殖漁場に変動を作り出す潜在的な因子に気が付きました。漠然とした見通しで始めてたまたまうまくいったのです。つまり、その分野の情報が必要で、主成分分析だけでは結論に到達しないのです。しかし、このことは、主成分分析をする動機付けとしては十分でしょう。コンピュータを使えば、主成分分析は簡単だからです。描かれた絵の意味が一でもわかるわけではないにしても、主成分分析で全体像が描けます。まず、主成分分析をやってみるべきです。