コスブラック P.167 式 9.2 の補遺

Kosugitti の研究ノート

2019/08/17

1　はじめに

拙著「言葉と数式で理解する多変量解析入門」(北大路書房) には，次のような式が出てきます[*1]。

$$\boldsymbol{A} = \lambda_1 \boldsymbol{x}_1 \boldsymbol{x}_1^T + \lambda_2 \boldsymbol{x}_2 \boldsymbol{x}_2^T + \cdots \lambda_m \boldsymbol{x}_m \boldsymbol{x}_m^T = \sum \lambda_i \boldsymbol{x}_i \boldsymbol{x}_i^T \tag{1}$$

これについて「固有ベクトルの大きさを適切に選んでやればできます」とだけしか書いてませんでした。本稿ではこれを詳しく説明しておきたいと思います。

これを説明 (証明) するためには，対称行列の固有値，固有ベクトルの性質を理解する必要があります。

2　準備

2.1　性質 1; 実対称行列の固有値は一般に実数になる

行列の固有値は実数で得られますが、これは固有方程式の解として得られるものです。そのため、行列の要素 (方程式の係数) が実数であっても、その解である固有値は一般に複素数で得られることになります。しかし、実対称行列の場合は、固有値がつねに実数で得られ、固有ベクトルも実数の範囲で得られるというありがたい性質があります。統計学では，分散共分散行列や相関行列，距離行列など，扱う正方行列が基本的に実対称行列ですので，固有値が複素数にならないのはありがたいことです。

話を始める前に，複素数ベクトルの内積について押さえておきましょう。

複素ベクトル $\boldsymbol{a}, \boldsymbol{b}$ の内積 $(\boldsymbol{a}, \boldsymbol{b})$ は，

$$(\boldsymbol{a}, \boldsymbol{b}) = \boldsymbol{a}^T \overline{\boldsymbol{b}} = \sum a_k \overline{b}_k = a_1 \overline{b}_1 + a_2 \overline{b}_2 + \cdots + a_n \overline{b}_n$$

で定義されます。ここで $\overline{\boldsymbol{b}}$ はベクトル $\boldsymbol{b}$ の共役複素数です。共役複素数とは，複素数 $a + bi$ に対する $a - bi$ のことを言います。

Proof. 実対称行列 $\boldsymbol{A}$ の固有値 λ に対する固有ベクトルを $\boldsymbol{x}$ とします。固有ベクトルの内積に固有値をかけたものを考えますと，

[*1] テキストでは転置をプライム ($\prime$) で表していましたが，本稿ではより見やすくするために T で表現しました。

内積はスカラーなので，一つにだけかければよくて，

$$\lambda(\boldsymbol{x}, \boldsymbol{x}) = (\lambda\boldsymbol{x}, \boldsymbol{x})$$

固有値の定義上，前の項は元行列になりますから

$$\begin{aligned} &= (\boldsymbol{A}\boldsymbol{x}, \boldsymbol{x}) \\ &= (\boldsymbol{A}\boldsymbol{x})^T\overline{\boldsymbol{x}} \end{aligned}$$

転置の計算の性質から，

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

$\boldsymbol{A}$ は実対称行列なので $\boldsymbol{A}^T = \boldsymbol{A}$，後ろのベクトルでくくり直して

$$= \boldsymbol{x}^T(\overline{\boldsymbol{A}\boldsymbol{x}})$$

これは複素数の内積の式なので

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

後ろの項は固有値の定義から

$$= (\boldsymbol{x}, \lambda\boldsymbol{x})$$

計算にあたっては $\boldsymbol{x}^T\overline{\lambda\boldsymbol{x}}$ になるので，$\overline{\lambda}$ だけ内積の式から取り出すと

$$= \overline{\lambda}(\boldsymbol{x}, \boldsymbol{x}).$$

したがって

$$\begin{aligned} \lambda(\boldsymbol{x}, \boldsymbol{x}) &= \overline{\lambda}(\boldsymbol{x}, \boldsymbol{x}) \\ (\lambda - \overline{\lambda})(\boldsymbol{x}, \boldsymbol{x}) &= 0. \end{aligned}$$

となります。ここで、$\boldsymbol{x} \neq \boldsymbol{0}$ ですから、$(\boldsymbol{x}, \boldsymbol{x}) \neq 0$ より、$\lambda = \overline{\lambda}$ であることがわかります。ある値 λ とその共役複素数 $\overline{\lambda}$ が等しいことから、λ が実数であると言えます。また、このとき、固有ベクトルは実係数の同時連立 1 次方程式 $(\boldsymbol{A} - \lambda\boldsymbol{I})\boldsymbol{x} = \boldsymbol{0}$ の解であることから、実ベクトルの範囲で選べることになります。

□

参考:村上正康・佐藤恒雄・野澤宗平・稲葉尚志著 (1997)『教養の線形代数』四訂版　倍風館.

2.2　性質 2; 異なる固有値に対応する固有ベクトルは直交する

Proof. 実対称行列 $\boldsymbol{A}$ の異なる固有値 λ_1 と λ_2 に対する固有ベクトルをそれぞれ $\boldsymbol{x}$,$\boldsymbol{y}$ とします。

$$\lambda_1(\boldsymbol{x}, \boldsymbol{y}) = (\lambda_1\boldsymbol{x}, \boldsymbol{y})$$

内積はスカラーなので，一つにだけかければよく，前の項は固有値の定義ですから

$$= (\boldsymbol{Ax}, \boldsymbol{y})$$

内積の計算をしましょう

$$\begin{aligned} &= (\boldsymbol{Ax})^T \boldsymbol{y} \\ &= \boldsymbol{x}^T \boldsymbol{A}^T \boldsymbol{y} \end{aligned}$$

$\boldsymbol{A}$ は対称なので $\boldsymbol{A}^T = \boldsymbol{A}$ ですから，

$$\begin{aligned} &= \boldsymbol{x}^T (\boldsymbol{Ay}) \\ &= (\boldsymbol{x}, \boldsymbol{Ay}) \end{aligned}$$

$\boldsymbol{Ay} = \lambda_2 \boldsymbol{y}$ としましたから，

$$\begin{aligned} &= (\boldsymbol{x}, \lambda_2 \boldsymbol{y}) \\ &= \lambda_2 (\boldsymbol{x}, \boldsymbol{y}) \end{aligned}$$

したがって，

$$\begin{aligned} \lambda_1 (\boldsymbol{x}, \boldsymbol{y}) &= \lambda_2 (\boldsymbol{x}, \boldsymbol{y}) \\ (\lambda_1 - \lambda_2)(\boldsymbol{x}, \boldsymbol{y}) &= 0 \end{aligned}$$

ですから，$\lambda_1 = \lambda_2$ か，$(\boldsymbol{x}, \boldsymbol{y}) = 0$ です。最初に $\lambda_1 \neq \lambda_2$ と仮定したのですから，$(\boldsymbol{x}, \boldsymbol{y}) = 0$ であり，内積がゼロということはすなわち，この二つのベクトルが直交しているということです。 □

3　本題

Proof. 正方行列 $\boldsymbol{A}$ の固有値 $\lambda_1, \lambda_2, \cdots, \lambda_m$ を対角に並べた行列 $\boldsymbol{\Lambda}$ と，固有ベクトル $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_m$ を列にまとめた行列 $\boldsymbol{X}$ を考えます。

$$\boldsymbol{\Lambda} = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_m \end{pmatrix}, \boldsymbol{X} = \left(\begin{pmatrix} \\ \boldsymbol{x}_1 \\ \\ \end{pmatrix} \begin{pmatrix} \\ \boldsymbol{x}_2 \\ \\ \end{pmatrix} \cdots \begin{pmatrix} \\ \boldsymbol{x}_m \\ \\ \end{pmatrix} \right)$$

これで固有値と固有ベクトルの計算を一気に表現できます。すなわち，

$$\boldsymbol{AX} = \boldsymbol{X\Lambda}$$

です。

これを展開して，

$$\boldsymbol{A} = \boldsymbol{X\Lambda X}^{-1}$$

固有ベクトルは互いに独立ですから，$\boldsymbol{XX}^T = \boldsymbol{I}$ であり，ここから $\boldsymbol{X}^T = \boldsymbol{X}^{-1}$ なので

$$= \boldsymbol{X\Lambda X}^T$$

となります。 □

これを要素ごとに書き下したのが，

$$\boldsymbol{A} = \lambda_1 \boldsymbol{x}_1 \boldsymbol{x}_1^T + \lambda_2 \boldsymbol{x}_2 \boldsymbol{x}_2^T + \cdots \lambda_m \boldsymbol{x}_m \boldsymbol{x}_m^T$$

ということになります。

3.1　数値例

行列 $\boldsymbol{A} = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$ の固有値と固有ベクトルは，固有方程式 $|\boldsymbol{A} - \lambda \boldsymbol{I}|\boldsymbol{x} = 0$ より，

$$\begin{aligned} \begin{vmatrix} 1-\lambda & 2 \\ 2 & 1-\lambda \end{vmatrix} &= 0 \\ (1-\lambda)^2 - 4 &= 0 \\ \lambda^2 - 2\lambda - 3 &= 0 \\ (\lambda - 3)(\lambda + 1) &= 0 \\ \lambda &= 3, -1 \end{aligned}$$

$\lambda = 3$ のとき，固有ベクトルは

$$\boldsymbol{A}\boldsymbol{x}_1 = 3\boldsymbol{x}_1$$

より，$\boldsymbol{x}_1 = \begin{pmatrix} x \\ y \end{pmatrix}$ と置いて，

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

$$\begin{cases} x + 2y &= 3x \\ 2x + y &= 3y \end{cases}$$

より，$\boldsymbol{x}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ となります。正確には，ベクトルの二つの要素 (上の式では x, y) は $x = y$ であればなんでもよくて，$\boldsymbol{x}_1 = \begin{pmatrix} 3 \\ 3 \end{pmatrix}$ でも，$\boldsymbol{x}_1 = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$ でも構いません。

同様に，$\lambda_2 = -1$ に対応する固有ベクトルは

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

$$\begin{cases} x + 2y &= -x \\ 2x + y &= -y \end{cases}$$

より，$\boldsymbol{x}_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ となります。

さて，固有ベクトルの大きさはなんでも良い，と言いましたが，元の行列に復元するときは固有ベクトルの大きさを整えておく必要があります。ベクトルの大きさ＝ノルムは $\|\boldsymbol{x}\| = \sqrt{(\boldsymbol{x}, \boldsymbol{x})}$ なので，$\boldsymbol{x}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ と

すると，$\|\boldsymbol{x}_1\| = \sqrt{1+1}$ です。固有ベクトルのノルムを 1 に整えたいので，以後それぞれの固有ベクトルを $\boldsymbol{x}_1 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}, \boldsymbol{x}_2 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix}$ とします。

固有値と固有ベクトルから元の行列を復元するには，$\boldsymbol{A} = \lambda_1 \boldsymbol{x}_1 \boldsymbol{x}_1^T + \lambda_2 \boldsymbol{x}_2 \boldsymbol{x}_2^T$ ですから，

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

とできることになります。

因子分析の場合は，$\boldsymbol{R} = \boldsymbol{A}\boldsymbol{A}^T + \boldsymbol{D}^2$ とするわけですが，これはノルム 1 に整えられた固有ベクトル $\boldsymbol{x}_i$ を $\sqrt{\lambda_i}\boldsymbol{x}_i$ とすることで，$\boldsymbol{R} = \boldsymbol{X}\boldsymbol{X}^T$ の形 (固有値が固有ベクトルの中に溶け込んだイメージ) になっていると考え，そこから構築していったものと言えます。