1　正多角形の相関係数

2 次元の散布図に n 個のデータが正 n 角形の頂点となるように配置されているとき，相関係数は 0 であることを証明します。

1.1　定義

n 個のデータ $(x_0,\ y_0), (x_1,\ y_1), \cdots, (x_{n-1},\ y_{n-1})$ を，複素数 $z_k = x_k + iy_k\ (k = 0, 1, \cdots, n-1)$ で表します。複素数平面上で点 $z_0, z_1, \cdots, z_{n-1}$ はこの順に左回りに正 n 角形を頂点ををなしているとします。また，この正 n 角形の外心を複素数 c とし，外接円の半径を $r > 0$ とします。$z_0 - c$ の偏角は α とします。

1.2　$\sum_{k=0}^{n-1} \cos\frac{2\pi k}{n} = 0$，$\sum_{k=0}^{n-1} \sin\frac{2\pi k}{n} = 0$ であること

$\omega = \cos\frac{2\pi}{n} + i\sin\frac{2\pi}{n}$ とすると，

$$\omega^n = 1$$

$$\omega^n - 1 = 0$$

$$(\omega - 1)(\omega^{n-1} + \omega^{n-2} + \cdots + 1) = 0$$

$\omega \fallingdotseq 1$ なので，

$$\omega^{n-1} + \omega^{n-2} + \cdots + 1 = 0$$

$$\sum_{k=0}^{n-1} \omega^k = 0$$

$$\sum_{k=0}^{n-1} \left(\cos\frac{2\pi k}{n} + i\sin\frac{2\pi k}{n}\right) = 0$$

$$\sum_{k=0}^{n-1} \cos\frac{2\pi k}{n} + i\sum_{k=0}^{n-1} \sin\frac{2\pi k}{n} = 0$$

$$\sum_{k=0}^{n-1} \cos\frac{2\pi k}{n} = 0, \quad \sum_{k=0}^{n-1} \sin\frac{2\pi k}{n} = 0 \qquad \cdots\cdots\cdots\cdots\cdots\cdots ①$$

1.3　$\sum_{k=0}^{n-1} \cos\frac{4\pi k}{n} = 0$，$\sum_{k=0}^{n-1} \sin\frac{4\pi k}{n} = 0$ であること

$\omega = \cos\frac{4\pi}{n} + i\sin\frac{4\pi}{n}$ とおき直しても，上と同様に，

$$\omega^n = 1$$

$$\omega^n - 1 = 0$$

$$(\omega - 1)(\omega^{n-1} + \omega^{n-2} + \cdots + 1) = 0$$

$\omega \neq 1$ なので，

$$\omega^{n-1}+\omega^{n-2}+\cdots+1=0$$

$$\sum_{k=0}^{n-1}\omega^k=0$$

$$\sum_{k=0}^{n-1}\left(\cos\frac{4\pi k}{n}+i\sin\frac{4\pi k}{n}\right)=0$$

$$\sum_{k=0}^{n-1}\cos\frac{4\pi k}{n}+i\sum_{k=0}^{n-1}\sin\frac{4\pi k}{n}=0$$

$$\sum_{k=0}^{n-1}\cos\frac{4\pi k}{n}=0,\quad \sum_{k=0}^{n-1}\sin\frac{4\pi k}{n}=0 \qquad \cdots\cdots\cdots\cdots\cdots\cdots ②$$

1.4 $\frac{1}{n}\sum_{k=0}^{n-1}z_k=c$ であること

$z_k=c+r\left\{\cos\left(\alpha+\frac{2\pi k}{n}\right)+i\sin\left(\alpha+\frac{2\pi k}{n}\right)\right\}$ と表せるので，

$$\begin{aligned}\sum_{k=0}^{n-1}z_k&=\sum_{k=0}^{n-1}c+r\sum_{k=0}^{n-1}\left\{\cos\left(\alpha+\frac{2\pi k}{n}\right)+i\sin\left(\alpha+\frac{2\pi k}{n}\right)\right\}\\&=\sum_{k=0}^{n-1}c+r\sum_{k=0}^{n-1}\left\{\cos\alpha\cos\frac{2\pi k}{n}-\sin\alpha\sin\frac{2\pi k}{n}+i\sin\alpha\cos\frac{2\pi k}{n}+i\cos\alpha\sin\frac{2\pi k}{n}\right\}\\&=\sum_{k=0}^{n-1}c+r\cdot 0 \qquad (\text{① より})\\&=nc\end{aligned}$$

これを c について解くと，

$$\begin{aligned}c&=\frac{1}{n}\sum_{k=0}^{n-1}z_k\\&=\frac{1}{n}\sum_{k=0}^{n-1}(x_k+iy_k)\\&=\frac{1}{n}\sum_{k=0}^{n-1}x_k+i\frac{1}{n}\sum_{k=0}^{n-1}y_k\\&=\overline{x}+i\overline{y} \qquad (\overline{x},\ \overline{y}\ \text{はそれぞれ}\ x_k,\ y_k\ \text{の平均})\end{aligned} \qquad \cdots\cdots\cdots\cdots\cdots\cdots ③$$

1.5 $\frac{1}{n}\sum_{k=0}^{n-1}f(z_k)=c$ であること

複素数 z に対して，関数 f を $f(z)=(z-c)^2+c$ と定めます。

$$\begin{aligned}f(z_k)&=f\left(c+r\left\{\cos\left(\alpha+\frac{2\pi k}{n}\right)+i\sin\left(\alpha+\frac{2\pi k}{n}\right)\right\}\right)\\&=r^2\left\{\cos\left(2\alpha+\frac{4\pi k}{n}\right)+i\sin\left(2\alpha+\frac{4\pi k}{n}\right)\right\}+c\end{aligned}$$

両辺に $\sum$ をつけると，

$$\sum_{k=0}^{n-1} f(z_k) = \sum_{k=0}^{n-1} r^2\left\{\cos\left(2\alpha + \frac{4\pi k}{n}\right) + i\sin\left(2\alpha + \frac{4\pi k}{n}\right)\right\} + \sum_{k=0}^{n-1} c$$

$$= r^2 \sum_{k=0}^{n-1}\left\{\cos 2\alpha \cos\frac{4\pi k}{n} - \sin 2\alpha \sin\frac{4\pi k}{n} + i\sin 2\alpha\cos\frac{4\pi k}{n} + i\cos 2\alpha \sin\frac{4\pi k}{n}\right\} + \sum_{k=0}^{n-1} c$$

$$= r^2 \cdot 0 + \sum_{k=0}^{n-1} c \qquad (\text{②より})$$

$$= nc$$

両辺の虚部をとって，

$$\mathrm{Im}\sum_{k=0}^{n-1} f(z_k) = \mathrm{Im}(nc)$$

$$= \mathrm{Im}(n\overline{x} + in\overline{y})$$

$$= n\overline{y} \qquad \cdots\cdots\cdots\cdots\cdots\cdots④$$

1.6　本番

再度 $f(z) = (z-c)^2 + c$ に戻ります。

$$f(z_k) = (z_k - c)^2 + c$$

$$= \left\{(x_k + iy_k) - (\overline{x} + i\overline{y})\right\}^2 + (\overline{x} + i\overline{y}) \qquad (\text{③より})$$

$$= \left\{(x_k - \overline{x}) + i(y_k - \overline{y})\right\}^2 + (\overline{x} + i\overline{y})$$

$$= (x_k - \overline{x})^2 - (y_k - \overline{y})^2 + 2i(x_k - \overline{x})(y_k - \overline{y}) + (\overline{x} + i\overline{y})$$

両辺の虚部をとって，

$$\mathrm{Im} f(z_k) = 2(x_k - \overline{x})(y_k - \overline{y}) + \overline{y}$$

$$(x_k - \overline{x})(y_k - \overline{y}) = \frac{1}{2}\left\{\mathrm{Im} f(z_k) - \overline{y}\right\} \qquad \cdots\cdots\cdots\cdots\cdots\cdots⑤$$

共分散 s_{xy} を求めると，

$$s_{xy} = \frac{1}{n}\sum_{k=0}^{n-1}(x_k - \overline{x})(y_k - \overline{y})$$

$$= \frac{1}{n}\sum_{k=0}^{n-1}\frac{1}{2}\left\{\mathrm{Im} f(z_k) - \overline{y}\right\} \qquad (\text{⑤より})$$

$$= \frac{1}{2n}\mathrm{Im}\sum_{k=0}^{n-1} f(z_k) - \frac{1}{2n}\sum_{k=0}^{n-1}\overline{y}$$

$$= \frac{1}{2n}\cdot n\overline{y} - \frac{1}{2n}\cdot n\overline{y} \qquad (\text{④より})$$

$$= 0$$

これで相関係数が 0 であることが示されました。

1.7　反省

やはり，z_k 全体を平行移動しても分散，共分散，相関係数は変わらないことをあらかじめ証明しておいた方がいいかも知れません。そうすれば $c = 0$ としてよいので，式が少し簡単になります。