2024 October

数学月刊

Mathematics Monthly

Author :
Guanzhong Yang(student)
Yan Wang(teacher)

Mathematics changes our lives

PREFACE

This month, we are going to talk about the following questions.

We'd like for you to define the squarilarity of two-dimensional shapes. The squarilarity should be a number between 0 and 1 quantifying how much a shape looks like a square. All squares should have squarilarity equal to 1, but aside from that restriction you can decide how to assign numbers to various shapes.

We welcome and are interested in seeing any ideas you have. It may be a good idea to start thinking of squarilarity definitions for simple shapes, and then expanding to more complicated shapes. Ultimately, for whatever definition(s) you come up with, we'd like for you to do the following.

(a) Specify the class of shapes for which your definition applies.

(b) State your definition of squarilarity for the shapes you're considering.

(c) Verify that the squarilarity is always between 0 and 1, and verify that the squarilarity of any square is 1.

(d) Explore the implications of your definition! Compute the squarilarity of a few shapes, and make observations of patterns you notice. Is your definition of squarilarity a good one?

This is the admission question from 2024 ROSS Program. If you have other brilliant ideas, email to anmiciuangray@163.com **for surprising rewards!**

Introduction

If we want to find the most suitable method to compare the similarity between a square and an arbitrary shape, we need to take all information that the arbitrary shape gives to us into consideration. Thus, we need to answer a question in advance: how many points on a shape do we need to fully describe an arbitrary shape?

Understandably, the answer is all the points on the shape.

But, obviously, it is impossible for us to take every point as a variable into the calculation, because there are infinite points! So, instead, we consider these points on the arbitrary shape as a whole: we use a continuous function to describe them.

So, how should we derive such a function? Well, for a shape on which all lines passing through its center of gravity only pass through two points on the edges, if we take its center of gravity as the origin, then we can express the points on the shape by using polar coordinates. In this way, we can get a function to express the shape.

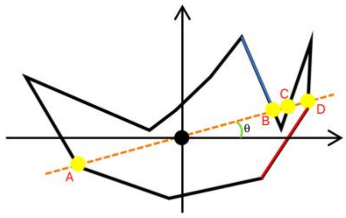
After having the function representing the arbitrary shape on which all lines passing through its center of gravity only pass through two points on the edges, we can use the similar method to derive the function representing a square, with its center of gravity lying on the origin.

Then, by using correlation coefficient, we will have a basic understanding of the similarity of a square and an arbitrary shape on which all lines passing through its center of gravity only pass through two points on the edges.

Some Auxiliary Results

Lemma 4.1 For an arbitrary shape with its center of gravity lying on the origin, if there exists a line, passing the center of gravity of the shape and more than two points on the edges of the shape, or one point on the edges of the shape, we cannot use a continuous function in a polar coordinate system to express it.

Proof.



graph 4.1

We can prove it through a general example.

In this situation, $f(\theta')$ has 4 possible outcome all together: A, B, C or D.

In order to be a continuous function, understandably, the outcome cannot be A.

But, $f(\theta' + \Delta\theta)$ is the blue line and $f(\theta' - \Delta\theta)$ is the red line, where $\Delta\theta \to 0$ and $\Delta\theta > 0$, meaning that no matter $f(\theta')$ is B, C or D, we cannot derive a continuous function.

Similarly, we can prove that if there exists a line, passing the center of gravity of the shape and one point on the edges, then the shape cannot be expressed by a continuous function in a polar coordinate system.

Lemma 4.2 For $f(\theta)$, a periodical function with period of 2π, if we know the coordinates of all its points but not its exact expression, then we can use Fourier Transformation to express it.

$$f(\theta) = \frac{a_0}{2} + \sum_{n=1}^{\infty}(a_n\cos n\theta + b_n\sin n\theta)$$

where

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(\theta)\cos n\theta \, d\theta,$$
$$b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(\theta)\sin n\theta \, d\theta,$$

where n = 1,2,3,⋯.

Proof.

It will be helpful if we firstly gain a basic insight towards Fourier Transformation.

Fourier Transformation is based on a very simply idea: every periodical function can be 'expressed' by sums of several sine waves and cosine waves.

Express this idea in a mathematical way:

$$f(\theta) = \frac{a_0}{2} + \sum_{n=1}^{\infty}(a_n\cos n\theta + b_n\sin n\theta)$$

where $f(\theta)$ is a periodical function with period of 2π.

Given that a_n, b_n are all constants and the trigonometric functions are orthogonal on $[-\pi, \pi]$, we can get

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(\theta)\cos n\theta \, d\theta,$$
$$b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(\theta)\sin n\theta \, d\theta,$$

where n = 1,2,3,⋯.

In the real situation, when we want to find an expression for a function, the coordinates of all points of which are known, but the exact expression of which is not, the most obvious method is to use Fourier Transformation.

By knowing the coordinates of every point, we can find the value of a_n and b_n through mass calculation, since as long as we have enough values of $f(\theta)\cos n\theta$ and $f(\theta)\sin n\theta$ for different values of θ and n, we can approximately find out the areas below the curve.

It might be impossible for human to do that---but we have computers! I have to admit, if we do want to express an arbitrary shape on which all lines passing through its center of gravity only pass through two points on the edges by using a function, we, most of the time, need a computer to help us.

Lemma 4.3 In a polar coordinate system, for an arbitrary shape on which all lines passing through its point on the edges only pass through two points on the edges, using any point on the edge as the origin and then find a function f(θ), which can express the shape, then the line passing though the center of gravity is $\theta = \theta'$, where

$$\int_{\theta'}^{\pi+\theta'} (f(\theta))^2 \, d\theta = \int_{\pi+\theta'}^{2\pi+\theta'} (f(\theta))^2 \, d\theta.$$

Proof.

If the 'density' of the shape is consistent, then a line passing through the center of gravity will divide the shape into two parts with equal areas.

Thus, we firstly find out the area between $\theta = \theta_1$ and $\theta = \theta_2$, for f(θ), in a polar coordinate system.

$$A = \frac{1}{2}\sum_{i=0}^{n_2} (f(n_2\Delta\theta))^2 \, \Delta\theta - \frac{1}{2}\sum_{i=0}^{n_1} (f(n_1\Delta\theta))^2 \, \Delta\theta$$

where $n_1\Delta\theta = \theta_1$, $n_2\Delta\theta = \theta_2$ and $\Delta\theta \to 0$.

Rearrange

$$A = \frac{1}{2}\int_{\theta_1}^{\theta_2} (f(\theta))^2 \, d\theta.$$

Thus, using any point on the edge as the origin and then find a function f(θ), which can express the shape, then the line passing though the center of gravity is $\theta = \theta'$, where

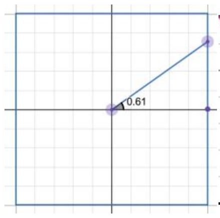
$$\int_{\theta'}^{\pi+\theta'} (f(\theta))^2 \, d\theta = \int_{\pi+\theta'}^{2\pi+\theta'} (f(\theta))^2 \, d\theta.$$

We repeat this operation twice to determine a point, which is the center of gravity, based on two non parallel lines.

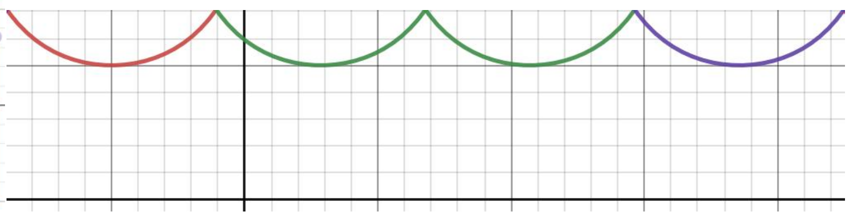
Lemma 4.4 The expression for a square with its center of gravity lying on the origin is

$$\begin{cases} r = \sec(-\theta + c) & \text{for } -\frac{\pi}{4} \le \theta \le \frac{\pi}{4} \\ r = \sec(\frac{\pi}{2} - \theta + c) & \text{for } \frac{\pi}{4} \le \theta \le \frac{3\pi}{4} \\ r = \sec(\pi - \theta + c) & \text{for } \frac{3\pi}{4} \le \theta \le \frac{5\pi}{4} \\ r = \sec(\frac{3\pi}{2} - \theta + c) & \text{for } \frac{5\pi}{4} \le \theta \le \frac{7\pi}{4} \end{cases}$$

where θ is the angle between line, connecting the point and the origin, and the positive direction of x-axis in the graph 4.1, and c is the distance of the transformation towards the positive direction of x-axis of the whole function, just like graph 4.3.



graph 4.2



graph 4.3 [in this graph, c = -1]

Proof.

$$r^2 = 1 + x^2$$

where

$$\begin{cases} x = \tan(-\theta) & \text{for } -\frac{\pi}{4} \le \theta \le \frac{\pi}{4} \\ x = \tan(\frac{\pi}{2} - \theta) & \text{for } \frac{\pi}{4} \le \theta \le \frac{3\pi}{4} \\ x = \tan(\pi - \theta) & \text{for } \frac{3\pi}{4} \le \theta \le \frac{5\pi}{4} \\ x = \tan(\frac{3\pi}{2} - \theta) & \text{for } \frac{5\pi}{4} \le \theta \le \frac{7\pi}{4}. \end{cases}$$

Recall that $\sec^2\theta = 1 + \tan^2\theta$.

The reason why we add a 'c' behind is that, sometimes, we need to rotate the square in graph 4.2 in order to find a square which is most resemble to the arbitrary shape on which all lines passing through its center of gravity only pass through two points on the edges.

Lemma 4.5 For a continuous function,

$$\mu = \frac{\int x f(x) dx}{\int f(x) dx} \text{ and } \sigma^2 = \frac{\int x^2 f(x) dx}{\int f(x) dx} - \mu^2$$

Proof.

For grouped data X,

$$\mu = \frac{\sum xf}{\sum f} \text{ and } \sigma^2 = \frac{\sum (x-\mu)^2 f}{\sum f} = \frac{\sum (x^2 - 2x\mu + \mu^2) f}{\sum f} = \frac{\sum x^2 f}{\sum f} - 2\mu \frac{\sum xf}{\sum f} + \mu^2 = \frac{\sum x^2 f}{\sum f} - \mu^2$$

Taking an integral can be seen as the sum of countless x_i, where $\Delta x \to 0$, thus

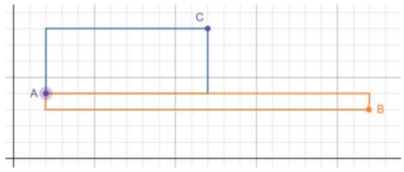
$$\mu = \frac{\int xf(x)dx}{\int f(x)dx} \text{ and } \sigma^2 = \frac{\int x^2 f(x)dx}{\int f(x)dx} - \mu^2$$

Lemma 4.6 For two continuous polar function, their correlation coefficient is

$$\frac{\left\{\dfrac{\int \theta[f_1(\theta) - \frac{\int \theta f_1(\theta)d\theta}{\int f_1(\theta)d\theta}][f_2(\theta) - \frac{\int \theta f_2(\theta)d\theta}{\int f_2(\theta)d\theta}]d\theta}{\int [f_1(\theta) - \frac{\int \theta f_1(\theta)d\theta}{\int f_1(\theta)d\theta}][f_2(\theta) - \frac{\int \theta f_2(\theta)d\theta}{\int f_2(\theta)d\theta}]d\theta}\right\}}{\left\{\left[\frac{\int \theta^2 f_1(\theta)d\theta}{\int f_1(\theta)d\theta} - \left(\frac{\int \theta f_1(\theta)d\theta}{\int f_1(\theta)d\theta}\right)^2\right]\left[\frac{\int \theta^2 f_2(\theta)d\theta}{\int f_2(\theta)d\theta} - \left(\frac{\int \theta f_2(\theta)d\theta}{\int f_2(\theta)d\theta}\right)^2\right]\right\}^{\frac{1}{2}}}.$$

Proof.

It might be helpful if we firstly have a basic understanding of how correlation coefficient is derived.



graph 4.4

In graph 4.4, we notice that (x_C - x_A) and (y_C - y_A) have the same signs(+&+ or -&-), meaning that they are positive correlation; (x_B - x_A) and (y_B - y_A) have the opposite signs(+&- or -&+), meaning that they are negative correlation.

If we calculate the 'area' of blue rectangle(result is positive) and the 'area' of orange rectangle(result is negative), then add them up, we can know the general relationship between x and y: are they positive correlation(final 'area' is positive) or negative(final 'area' is negative)?

You may think that this method is not rigorous enough for us to find out the general relationship between x and y, but as long as we have infinite different pairs of (x,y), we will definitely derive a more reasonable result.

Express our idea in a mathematical way:

$$E((x_i - x_j)(y_i - y_j)),$$

where i ≤ j.

But, understandably, it is too tough for us to calculate infinite different areas! So, we would like to simplify it:

$$E((x_i - \mu_X)(y_i - \mu_Y)),$$

it is much easier, right?

But now, we encounter another problem: what about if we want to compare the relationship between x and y, and the relationship between x and z?

Remember that the dispersion of data will influence the final 'area', but what we want is not the final 'area', instead, it is whether they are positive correlation or negative.

So we want to exclude the influence of dispersion of every data on the results:

$$\frac{E((x_i - \mu_X)(y_i - \mu_Y))}{\sigma_X \sigma_Y}.$$

This is the expression of correlation coefficient. In this way, there isn't an unit for correlation coefficient, meaning that we can compare the correlation coefficient of x and y, and the correlation coefficient of x and z.

For correlation coefficient, when X = kY + c, where k is positive,

$$\frac{E((x_i - \mu_X)(y_i - \mu_Y))}{\sigma_X \sigma_Y} = \frac{kE((x_i - \mu_X)(x_i - \mu_X))}{k\sigma_X \sigma_X} = 1,$$

when X = -kY + c, where k is positive

$$\frac{E((x_i - \mu_X)(y_i - \mu_Y))}{\sigma_X \sigma_Y} = \frac{-kE((x_i - \mu_X)(x_i - \mu_X))}{k\sigma_X \sigma_X} = -1,$$

and when there is not relationship between X and Y, if we find infinite different pairs of (x,y),

$$E((x_i - \mu_X)(y_i - \mu_Y)) = 0,$$

$$\frac{E((x_i - \mu_X)(y_i - \mu_Y))}{\sigma_X \sigma_Y} = 0.$$

In conclusion, the closer the value of the correlation coefficient is to 1, the stronger the relationship between X and Y is to be positive correlation; the closer the value of the correlation coefficient is to -1, the stronger the relationship between X and Y is to be negative correlation; the closer the value of the correlation coefficient is to 0, the more likely there is no relationship between X and Y.

But how can we use these to show the relationship between two functions? Well, imagine that for a given input θ', we will get a corresponding ($f_1(\theta')$,$f_2(\theta')$), where f_1 representing the first function and f_2 representing the second function. But working out the correlation coefficient of f_1 and f_2, we can have a basic understanding of the relationship between two functions:
The closer the value of the correlation coefficient is to 1, the stronger the relationship between f_1 and f_2 is to be positive correlation(larger the value of $f_1(\theta')$, larger the value of $f_2(\theta')$); the closer the value of the correlation coefficient is to -1, the stronger the relationship between X and Y is to be negative correlation(larger the value of $f_1(\theta')$, smaller the value of $f_2(\theta')$); the closer the value of the correlation coefficient is to 0, the more likely there is no relationship between X and Y .
That be said, when correlation coefficient is 1,

$$f_1(\theta) = kf_2(\theta) + c$$

where k is positive; when correlation coefficient is -1,

$$f_1(\theta) = -kf_2(\theta) + c$$

where k is negative.
when correlation coefficient is 0, two function have no relationship.

Now we will find the correlation coefficient for two continuous polar functions,
Recall that, for a continuous function,

$$\mu = \frac{\int xf(x)dx}{\int f(x)dx} \text{ and } \sigma^2 = \frac{\int x^2 f(x)dx}{\int f(x)dx} - \mu^2.$$

$$\frac{E((f_1(\theta) - \mu_{f_1})(f_2(\theta) - \mu_{f_2}))}{\sigma_{f_1}\sigma_{f_2}} = \frac{\frac{\int \theta(f_1(\theta) - \mu_{f_1})(f_2(\theta) - \mu_{f_2})d\theta}{\int (f_1(\theta) - \mu_{f_1})(f_2(\theta) - \mu_{f_2})d\theta}}{[(\frac{\int \theta^2 f_1(\theta)dx}{\int f_1(\theta)dx} - \mu_{f_1}{}^2)(\frac{\int \theta^2 f_2(\theta)dx}{\int f_2(\theta)dx} - \mu_{f_2}{}^2)]^{\frac{1}{2}}}$$

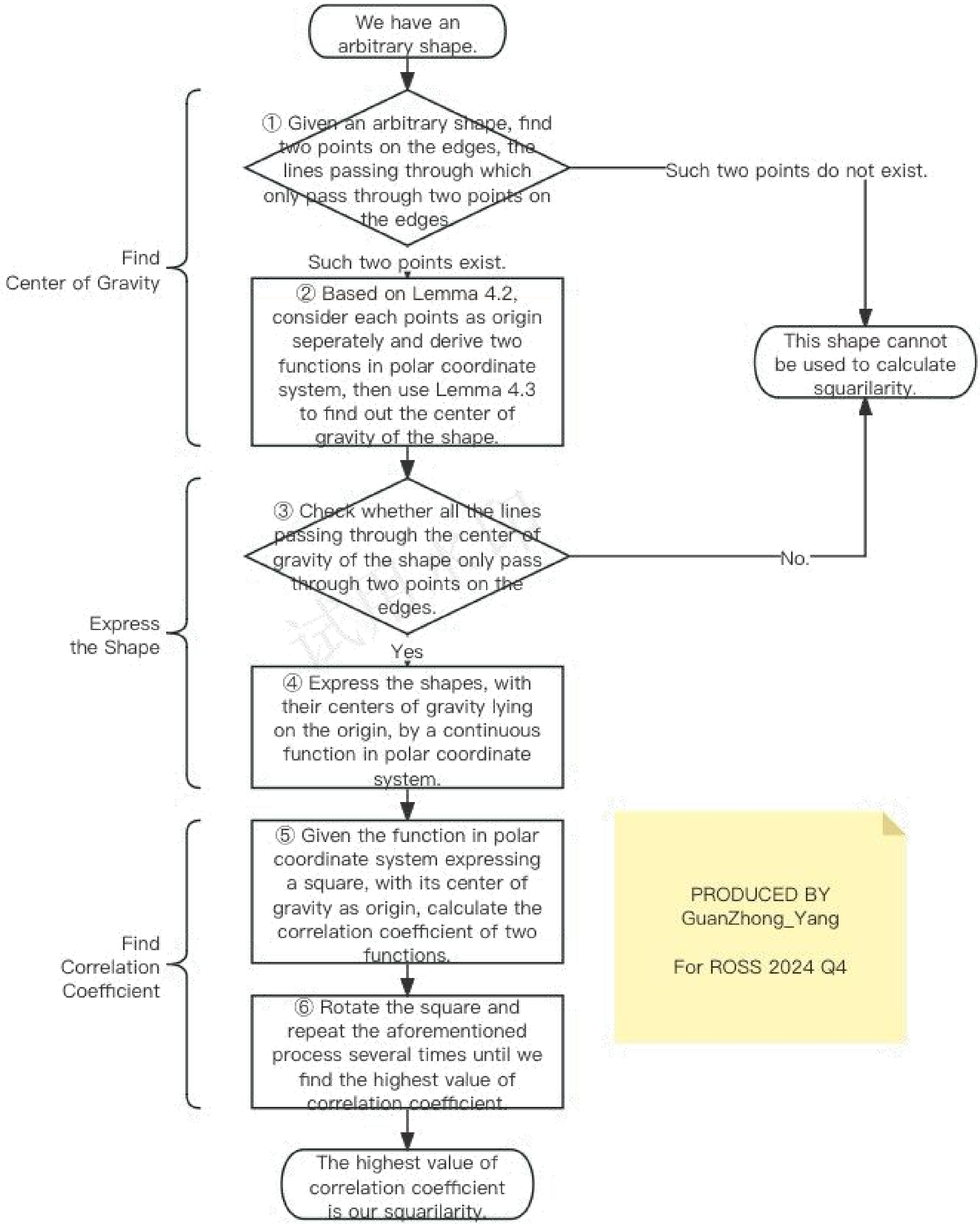
$$= \frac{\{\frac{\int \theta[f_1(\theta) - \frac{\int \theta f_1(\theta)d\theta}{\int f_1(\theta)d\theta}][f_2(\theta) - \frac{\int \theta f_2(\theta)d\theta}{\int f_2(\theta)d\theta}]d\theta}{\int [f_1(\theta) - \frac{\int \theta f_1(\theta)d\theta}{\int f_1(\theta)d\theta}][f_2(\theta) - \frac{\int \theta f_2(\theta)d\theta}{\int f_2(\theta)d\theta}]d\theta}\}}{\{[\frac{\int \theta^2 f_1(\theta)d\theta}{\int f_1(\theta)d\theta} - (\frac{\int \theta f_1(\theta)d\theta}{\int f_1(\theta)d\theta})^2][\frac{\int \theta^2 f_2(\theta)d\theta}{\int f_2(\theta)d\theta} - (\frac{\int \theta f_2(\theta)d\theta}{\int f_2(\theta)d\theta})^2]\}^{\frac{1}{2}}}.$$

Lemma 4.7 Given two continuous polar functions, one can be used to express a square with its center of gravity lying on the origin, and the other can be used to express an arbitrary shape on which all lines passing through its center of gravity only pass through two points on the edges and with its center of gravity lying on the origin. Assume that the arbitrary shape is fixed and the square can be rotated to find different correlation coefficient, then if we derive the correlation coefficient with a negative value in one situation, we can derive the correlation coefficient with a positive value in another.
Proof.
Recall that the closer the value of the correlation coefficient is to 1, the stronger the relationship between f_1 and f_2 is to be positive correlation(larger the value of $f_1(\theta')$, larger the value of $f_2(\theta')$); the closer the value of the correlation coefficient is to -1, the stronger the relationship between X and Y is to be negative correlation(larger the value of $f_1(\theta')$, smaller the value of $f_2(\theta')$).
Thus, if in one situation, the correlation coefficient is negative, meaning that larger the value of $f_1(\theta')$, smaller the value of $f_2(\theta')$, we can rotate the square, making the point with original highest value into smallest value and the point with original smallest value into larger value etc.
In this way, the general trend will be concerted into: larger the value of $f_1(\theta')$, larger the value of $f_2(\theta')$, meaning that the correlation coefficient is positive.

Solution

Our squarilarity is a way to compare the similarity between a square and an arbitrary shape. In order to ensure that we can derive two continuous functions in polar coordinate system, there must exist two points on the edges of shape, all lines passing through each of which only pass through two points on the edges, and all lines passing through the center of gravity of the shape only pass through two points on the edges.

The way to derive squarilarity is:



Based on Lemma 4.7, our squarilarity is always in the range of [0,1] because we can 'reserve' the general trend.

So now, I will calculate the squarilarity of a circle to show you that this method is workable. But to begin with, I have to admit that it is difficult for us to calculate the final result without the help of calculators(for example, I fail to deal with

$$\int \ln|\sec(-\theta)+\tan(-\theta)|d\theta.).$$

So, in this part, I get these answers through calculators.
If $f_1(\theta)$ is used to represent square and $f_2(\theta)$ is used to represent circle, then

$$\mu_{f_1(\theta)} \approx 2.35619449,$$

$$\sigma_{f_1(\theta)} \approx 1.813800785,$$

$$\mu_{f_2(\theta)} = \frac{3\pi}{4},$$

$$\sigma_{f_2(\theta)} \approx 1.813799364,$$

$$\text{squarilarity} \approx 0.6495187987.$$

You will definitely notice that I only find the squarilarity of one shape---but if we want to prove that this method is useful, we need to try more times and only in that way can we make an reasonable conclusion!
Well, I will firstly tell you why I just find the squarilarity of one shape: if I want to find the squarilarity of other shapes, I need a computer with super powerful computing power instead of a simple calculator. Because we during the calculation, we need to rotate the square for many times until we find a maximum value of correlation coefficient. For many shape, we need to repeat this operation for almost infinite times(actually, in real situation, we rotate a small angle for finite times and find a approximate value), whereas for a circle, we no longer need to do such operations because a circle has countless symmetry axes which pass through its center, meaning that rotating a square will not have an impact on the final result.
In conclusion, circle is the most suitable shape for us to calculate its squarilarity because we don't need to rotate it and find many different values of correlation coefficient; whereas for other shapes, we have to rotate them and find almost infinite different values of correlation coefficient so as to find the squarilarity, which is impossible if we only have calculator.
Then, I will show you why, as far as I'm concerned, this method is useful, which you can see in the afterword below.

Afterword

I have to say, broadening our horizons by reading books is always useful.
Thanks to my study of AP Statistics, at the first time I saw this question, I came up with the application of correlation coefficient immediately. Compared with other methods that I later came up with, like finding out the variance of $(f_1(\theta) - f_2(\theta))$, finding out the correlation coefficient of two functions has 2 biggest advantages:
① It takes as many features as possible into consideration, like μ_{f_1}, μ_{f_2}, σ_{f_1} and σ_{f_2}.
② The range of final result is in the range of [-1,1], and by finding out the largest value of correlation coefficient, the final final result is [0,1]. Instead of doing some meaningless operations so as to satisfy the question condition, our boundaries have their geometric meanings.
Thus, I would like to consider it as my squarilarity.

You may notice that, one of the key steps in calculating the squarilarity is to express an arbitrary shape on which all lines passing through its center of gravity only pass through two points on the edges and with its center of gravity lying on the origin by a continuous function in a polar coordinate system with the help of Fourier Transformation.
So, here's two question: where do I learn Fourier Transformation? And why don't we use Taylor expansion?
For the first question, this is due to my winter school experience. Thanks to my study of *Fitzwilliam College Online Winter School Programme*, when I had to submit my own work about a certain topic and did a presentation, I chose Fourier Transformation and gained a basic insight towards the application of Fourier Transformation. Eventually, I was awarded as one of the top two-performing students in this course. If you like, you can also see the video about my presentation at: https://meeting.tencent.com/v2/cloud-record/share?id=2507a607-d9f5-4c75-8207-f98e0ff53cc8&from=3&record_type=2 .

For the second one, that is because these two functions are obviously periodical, and for periodical functions, expressing them using Fourier Transformation is a better choice, since Fourier Series is also periodical.

Apart from the aforementioned method of expressing a shape by a function, we can, actually, also express the shapes by functions like below.
For every straight line in a coordinate system, we can express them as

$$Ax + By + C = 0.$$

Then, if we want to express several lines together, whose expressions, respectively, are

$$\begin{gathered} A_1x + B_1y + C_1 = 0, \\ A_2x + B_2y + C_2 = 0, \\ \vdots \\ A_nx + B_ny + C_n = 0, \end{gathered}$$

we can express them as

$$(A_1x + B_1y + C_1)(A_2x + B_2y + C_2)\cdots(A_nx + B_ny + C_n) = 0.$$

But what about if the line, namely $A_ix + B_iy + C_i = 0$, is only in the range of $x_1 \le x \le x_2$?
In this way, we just need to 'add' some limitations:

$$\left(\sqrt{x - x_1}\sqrt{x_2 - x} + 1\right)(A_ix + B_iy + C_i) = 0.$$

Similarly, if we want to make y in the range of $y_1 \le y \le y_2$, then

$$\left(\sqrt{y - y_1}\sqrt{y_2 - y} + 1\right)(A_ix + B_iy + C_i) = 0.$$

You may wonder what I mean by the red part and the yellow part.
Well, the red part ensures the range of our graph, and the yellow part ensures that, when the result is 0, then the only possibility is

$$A_ix + B_iy + C_i = 0.$$

Apart from that, you can also express the line, namely $A_ix + B_iy + C_i = 0$, in the range of $x_1 \le x \le x_2$ by

$$(\big|\,|x - x_1| - |x - x_2|\,\big| - |x_1 - x_2|)(A_ix + B_iy + C_i) = 0,$$

because when x is not in the range of $x_1 < x < x_2$, then

$$\big|\,|x - x_1| - |x - x_2|\,\big| = |x_1 - x_2|,$$

making the parts of the line disappeared.
Similarly, you can also express the line, namely $A_ix + B_iy + C_i = 0$, in the range of $y_1 \le y \le y_2$ by

$$(\big|\,|y - y_1| - |y - y_2|\,\big| - |y_1 - y_2|)(A_ix + B_iy + C_i) = 0.$$

Well, if we consider the lines as 'pencil', then

$(\sqrt{x - x_1}\sqrt{x_2 - x} + 1)$ *[range eraser]*

$(\big|\,|y - y_1| - |y - y_2|\,\big| - |y_1 - y_2|)$ *[distance eraser]*

can be called as 'eraser's. In fact, there are many other kinds of 'eraser' such as

$(\sqrt{r^2 - [(x - x_{center})^2 + (y - y_{center})^2]} + 1)$ *[circle eraser]*

$(\sqrt{y - (ax^2 + bx + c)} + 1)$ *[binomial eraser].*

These 'eraser's plus 'pencil' are the most basic idea of expressing a shape without using polar coordinate system. Since every shape can be expressed by a polygon with infinite edges, we can express them by using the strategy above.
Whereas, obviously, there will be some problems when we use this strategy to express a shape:
① For one shape, we can find many different functions expressing it. So, it is impossible for us to use this method to find squarilarity.
② Not all shapes can be easily converted into polygons with infinite edges, unless we enlarge it to a sufficiently large size, which is almost impossible without the help of computers!
③ The calculation involved in this method is a disaster, I believe that even computers would not want to use this method to represent a graph and compare the correlation coefficients of two functions!
So all in all, compared with this method, our original way of conversion is quite efficient!