Covariance

Mahdi Akbari Zarkesh
9612762638

INSTRUCTIONS

Covariance is a measure of how much two random variables vary together. It's similar to variance, but where variance tells you how a single variable varies, co variance tells you how two variables vary together.

The Covariance Formula

1. Cov(X,Y) = Σ E((X-μ)E(Y-ν)) / n-1 where:
2. X is a random variable
3. E(X) = μ is the expected value (the mean) of the random variable X and
4. E(Y) = ν is the expected value (the mean) of the random variable Y
5. n = the number of items in the data set

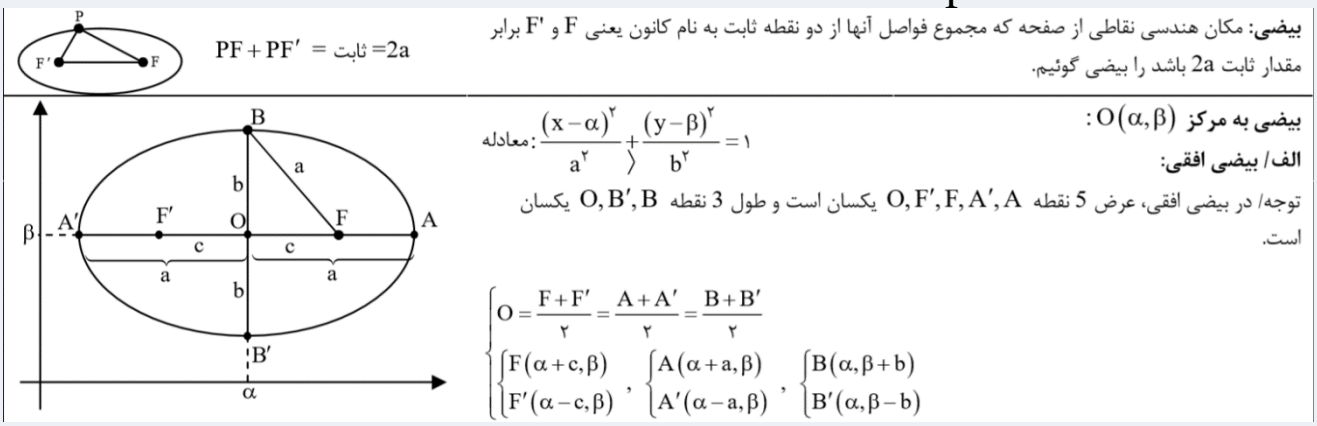
Advantages of the Correlation Coefficient

The Correlation Coefficient has several advantages over covariance for determining strengths of relationships:

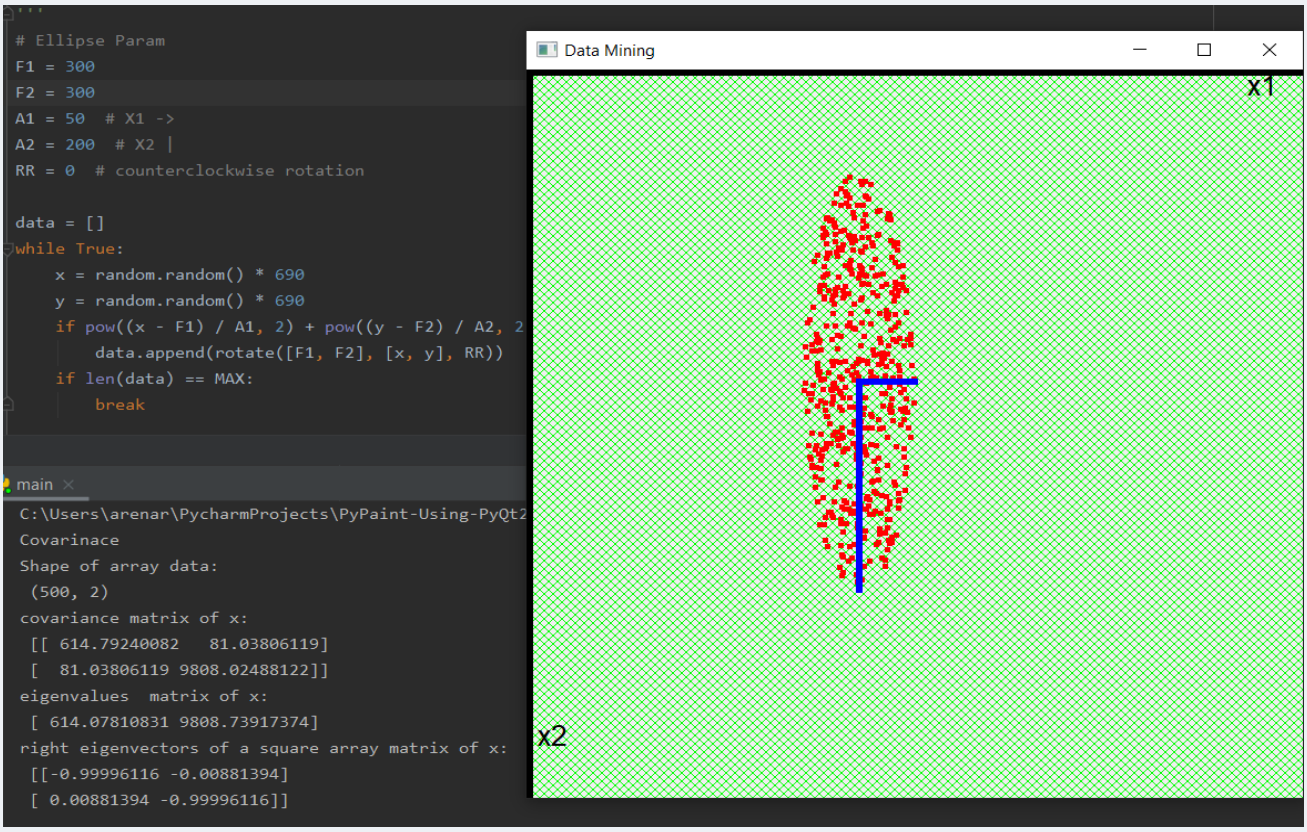
1. Covariance can take on practically any number while a correlation is limited: -1 to +1.
2. Because of it's numerical limitations, correlation is more useful for determining how strong the relationship is between the two variables.
3. Correlation does not have units. Covariance always has units
4. Correlation isn't affected by changes in the center (i.e. mean) or scale of the variables

now look at the our result:

In our test we use random data and fit them in ellipse.

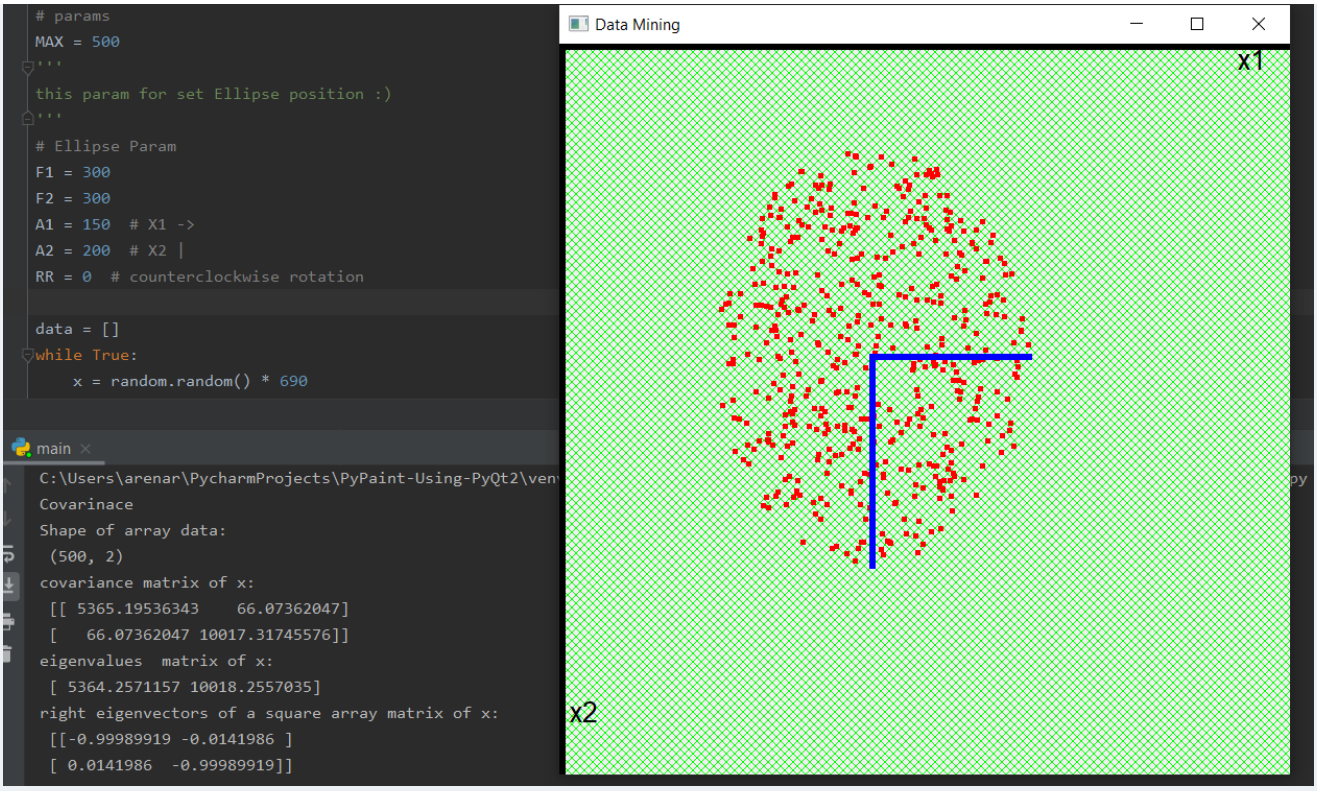


for first test we make ellipse stretched to measure eigenvalues of covariance:



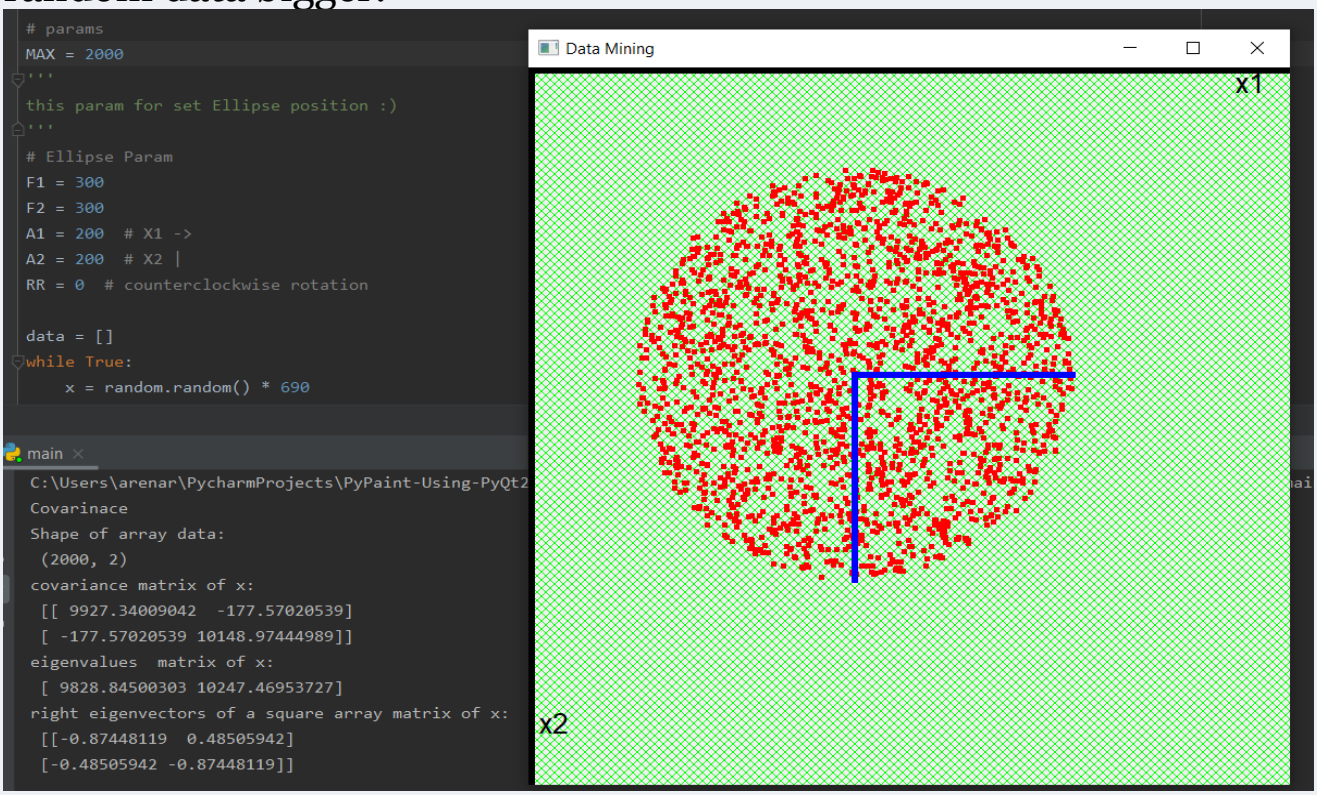
As you see eigenvalues of covariance of parameter X2 is bigger than X1 because of the data are more distribute on x2 more than x1.

And now we're making the ellipse a bit more regular with a radius of 4 to 3



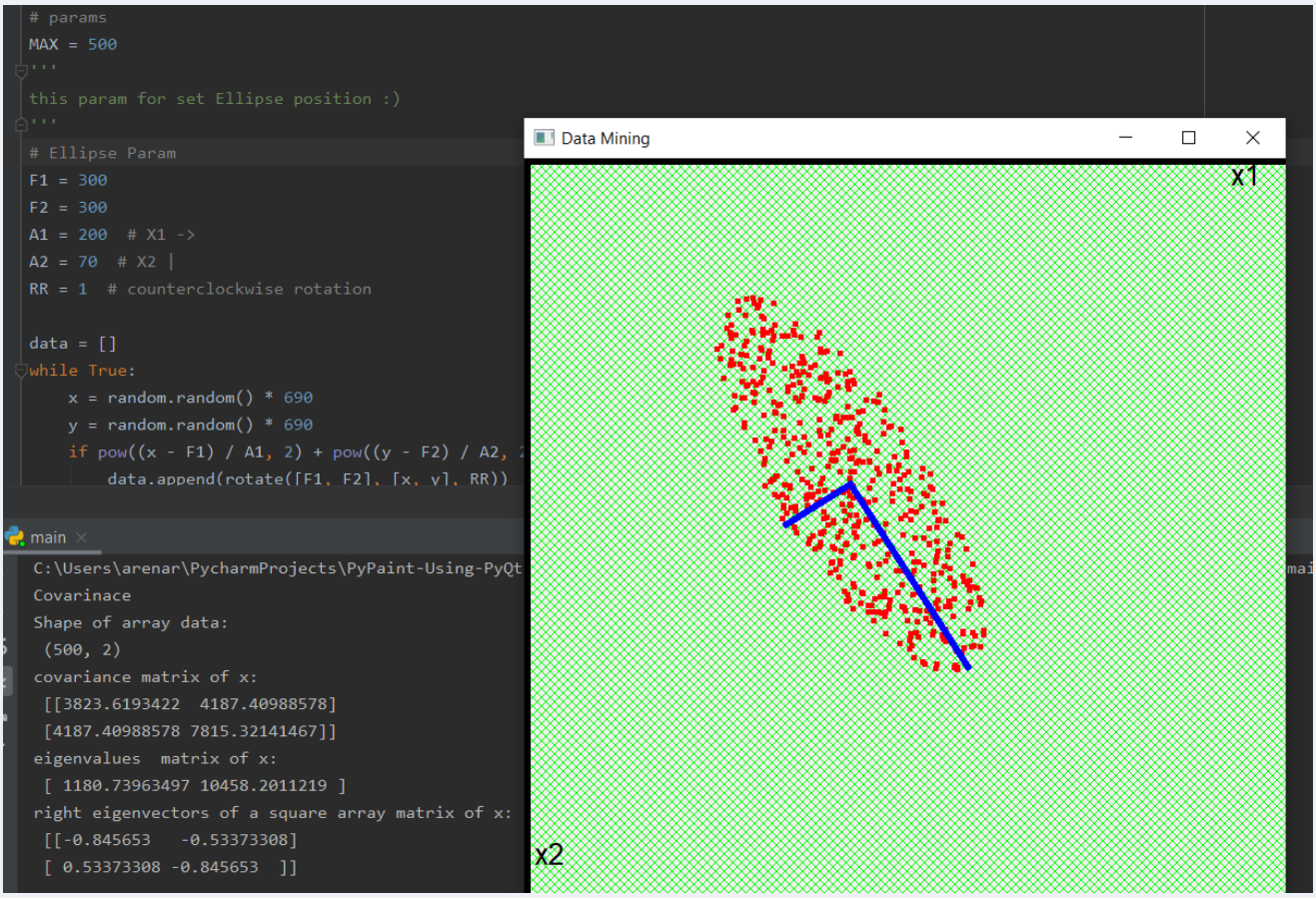
As we expected, the values got closer.

And at last we make shape like circle and to get better result make random data bigger:

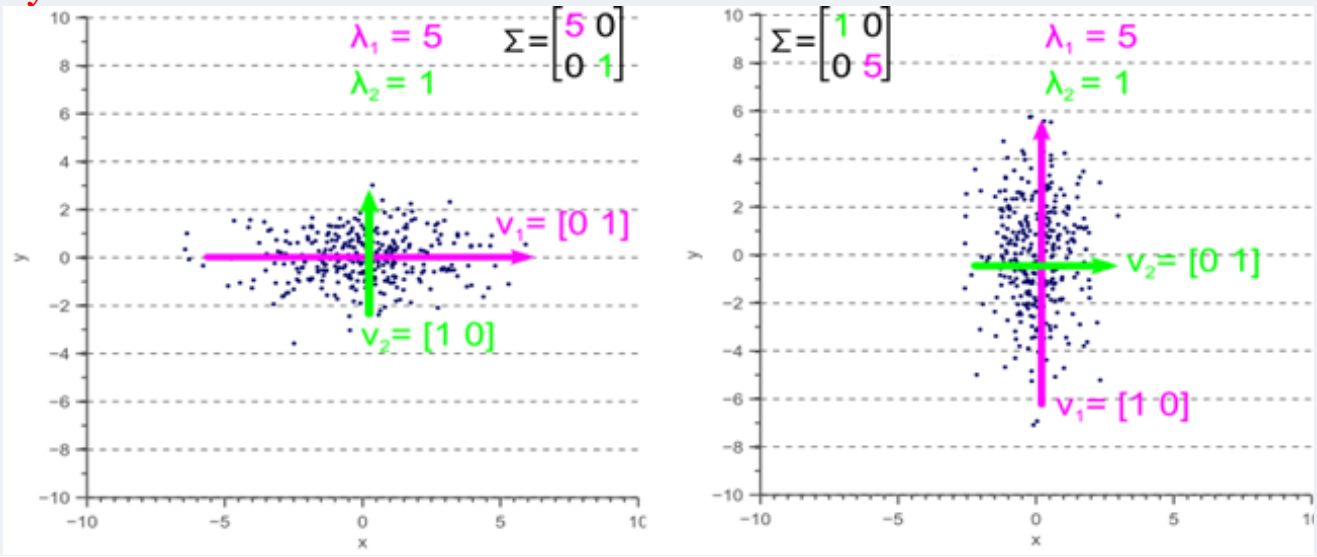


As we expected, the values got same.
because distribute data in X1 and X2 are same

now let's rotate ellipse:



As we can see, the Large to small radius ratio is 200/70 and we rotate a point counterclockwise by a given RR=1 around a given center of ellipse.After rotate the dependency of X1 and X2 are grows.
The covariance matrix defines the shape of the data. Diagonal spread is captured by the covariance, while axis-aligned spread is captured by the variance.



Thank you :)