MATHEMATICS

How to construct a covariance matrix from a 2x2 data set

Asked 7 years, 6 months ago   Active 6 years, 5 months ago   Viewed 55k times

20

17

so if given a covariance matrix I can find the eigenvalues and move forward from there... but I seem to have trouble with the step before if I am given a data set and am told to create the covariance matrix. Looking at the notes I see the formula:

$$cov(x) = \frac{1}{n-1}\sum(x_i - \bar{x})(y_i - \bar{y})$$

I'm not too sure what to do with this formula and was hoping you can tell me how.

Data Set:

```
X1 | X2
---|---
3  | 7
2  | 4
```

matrices   correlation

Share Cite Follow

I saw some sources writing covariance matrix as a product of $X_1 X_2^T/(n-1)$, could anyone help explain why it works? – Logan Apr 19 '17 at 8:12

1 Answer

Active   Oldest   Votes

The *variance-covariance* matrix has the following structure:

$$\begin{bmatrix} var(x) & cov(x,y) \\ cov(x,y) & var(y) \end{bmatrix}$$

46

where $var(x) = \frac{1}{n-1}\sum(x_i - \bar{x})^2$ and $cov(x,y) = \frac{1}{n-1}\sum(x_i - \bar{x})(y_i - \bar{y})$.

for your data,

$$\bar{x} = \frac{(3+2)}{2} = \frac{5}{2}$$

$$\bar{y} = \frac{(7+4)}{2} = \frac{11}{2}$$

$$var(x) = (3 - \tfrac{5}{2})^2 + (2 - \tfrac{5}{2})^2$$

$$var(y) = (7 - \tfrac{11}{2})^2 + (4 - \tfrac{11}{2})^2$$

$$cov(x,y) = (3 - \tfrac{5}{2})(7 - \tfrac{11}{2}) + (2 - \tfrac{5}{2})(4 - \tfrac{11}{2})$$

so, all you need to do is calculate these values and put them in the right places in the matrix. Does that make sense?

Share Cite Follow

@Brad S. I have a slightly different problem. I also want to obtain a covariance matrix. I need it to use it as input for a generalized χ^2 minimization in order to fit a model when the errors from the data are correlated. In short, I have a dataset x_i of uncorrelated data with $i = 1..30$. But then I need to apply the following transformation to the data $y_i = x_{i-1} - 2x_i + x_{i+1}$. After that, the errors get correlated and I need to account for that by means of the covariance matrix given by the transformation. Could you please give me a hand or advise on how to get that matrix? – Stefano May 20 '18 at 15:43 ✎

2 I think variance formula should be 1/N not 1/N-1, am I missing something? – PeerNet Oct 13 '19 at 22:28

@PeerNet No, when looking at a sample of a population the correct variance is $\frac{1}{N-1}\sum(x_i - \bar{x})$. You must be confusing this with the formula for the population variance; there we divide by N only. – alexandre lopez Dec 1 '20 at 1:11