Computation of the Covariance Matrix

Analytics Vidhya
Learn everything about analytics

What is Covariance?

- Measure of strength of correlation between two features
- Mathematically, for two features X and Y with n observations

  (X_i, Y_i):

$$cov(X,Y) = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})$$

Analytics Vidhya
Learn everything about analytics

What is a Covariance Matrix?

	Y_1	Y_2	Y_3
Y_1			
Y_2			
Y_3			

Analytics Vidhya
Learn everything about analytics

What is a Covariance Matrix?

	Y_1	Y_2	Y_3
Y_1	5		
Y_2		8	
Y_3			1

Analytics Vidhya
Learn everything about analytics

What is a Covariance Matrix?

	Y_1	Y_2	Y_3
Y_1	5	-1	0
Y_2	-1	8	3
Y_3	0	3	1

Analytics Vidhya
Learn everything about analytics

Computation of the Covariance Matrix

- Consider a data matrix X with 3 features (A, B, C) and 2 rows

A	B	C
a_1	b_1	c_1
a_2	b_2	c_2

Analytics Vidhya
Learn everything about analytics

Computation of the Covariance Matrix

- The data is column standardised

$$\bar{A} = \bar{B} = \bar{C} = 0$$

Analytics Vidhya
Learn everything about analytics

Computation of the Covariance Matrix

- Covariance between two features A and B

$$cov(A,B) = \frac{1}{n}\sum_{i=1}^{n}(a_i - \bar{A})(b_i - \bar{B})$$

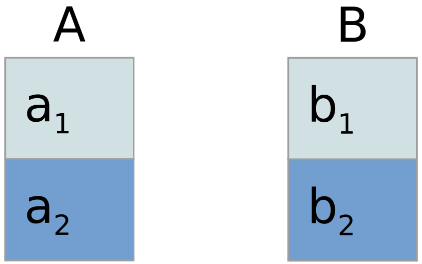
$$= \frac{1}{2}\sum_{i=1}^{2}(a_i - 0)(b_i - 0) = \frac{1}{2}\sum_{i=1}^{2}(a_i)(b_i)$$

$$= \frac{1}{2}\boxed{(a_1 b_1 + a_2 b_2)}$$

Analytics Vidhya
Learn everything about analytics

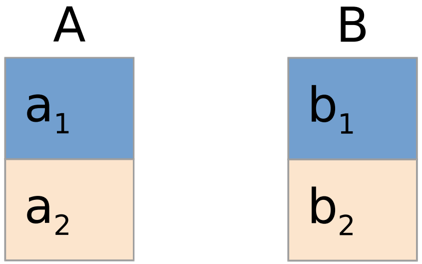
Computation of the Covariance Matrix

$$\underbrace{a_1 b_1} + a_2 b_2$$



Computation of the Covariance Matrix

$$a_1 b_1 + \underbrace{a_2 b_2}_{}$$



Analytics Vidhya
Learn everything about analytics

Computation of the Covariance Matrix

$$\underbrace{a_1 b_1 + a_2 b_2}_{} = A.B$$

A	B
a_1	b_1
a_2	b_2

$$\text{cov}(A, B) = A.B$$

Analytics Vidhya
Learn everything about analytics

Computation of the Covariance Matrix

cov =

	A	B	C
A	A.A	A.B	A.C
B	B.A	B.B	B.C
C	C.A	C.B	C.C

Analytics Vidhya
Learn everything about analytics

Computation of the Covariance Matrix

$$
\underbrace{\begin{array}{c|cc} & & \\ \hline A & a_1 & a_2 \\ & b_1 & b_2 \\ B & c_1 & c_3 \\ C & & \end{array}}_{X^T}
\quad
\underbrace{\begin{array}{ccc} A & B & C \\ \hline a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{array}}_{X}
\quad = \quad
$$

	A	B	C
A	A.A	A.B	A.C
B	B.A	B.B	B.C
C	C.A	C.B	C.C

cov

Analytics Vidhya
Learn everything about analytics

Computation of the Covariance Matrix

$$
\begin{array}{c}
\begin{array}{c c}
\begin{array}{c} A \\ B \\ C \end{array} &
\begin{bmatrix} a_1 & a_2 \\ b_1 & b_2 \\ c_1 & c_3 \end{bmatrix}
\end{array} \\
X^T
\end{array}
\quad
\begin{array}{c}
\begin{array}{ccc} A & B & C \end{array} \\
\begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{bmatrix} \\
X
\end{array}
\quad = \quad
$$

	A	B	C
A	A.A	A.B	A.C
B	B.A	B.B	B.C
C	C.A	C.B	C.C

cov

Analytics Vidhya
Learn everything about analytics

Computation of the Covariance Matrix

$$\begin{array}{c|cc} A & a_1 & a_2 \\ B & b_1 & b_2 \\ C & c_1 & c_3 \end{array} \quad \begin{array}{ccc} A & B & C \\ \hline a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{array} = \begin{array}{c|ccc} & A & B & C \\ \hline A & A.A & A.B & A.C \\ B & B.A & B.B & B.C \\ C & C.A & C.B & C.C \end{array}$$

X^T $\qquad$ X $\qquad$ cov

Analytics Vidhya
Learn everything about analytics

Computation of the Covariance Matrix

$$\text{cov} = X^T X$$

Analytics Vidhya
Learn everything about analytics

Thank You!

Analytics Vidhya
Learn everything about analytics