Zlatniczki Ádám
adam.zlatniczki@cs.bme.hu

Kovariancia, korreláció

Szűk elméleti összefoglaló

- X és Y val. változók kovarianciáján a $Z = (X - \boldsymbol{E}X)(Y - \boldsymbol{E}Y)$ val. változó várható értékét értjük, tehát

$$cov(X,Y) = \boldsymbol{E}[(X - \boldsymbol{E}X)(Y - \boldsymbol{E}Y)]$$

- egy val. változó varianciája nem más, mint önmagával vett kovarianciája

$$cov(X,X) = \sigma^2 X$$

- korreláció:

$$R(X,Y) = \frac{cov(X,Y)}{\sigma X \sigma Y}$$

- tehát a korreláció a [-1,1] intervallumra normalizált kovariancia

Tételek:

- $cov(X,Y) = E(XY) - E(X)E(Y)$

- Ha X és Y függetlenek, akkor $cov(X,Y) = 0 \Rightarrow R(X,Y) = 0$, DE fordítva nem igaz!

- $\sigma^2(X + Y) = \sigma^2 X + \sigma^2 Y + 2cov(X,Y)$

- $\sigma^2(X - Y) = \sigma^2 X + \sigma^2 Y - 2cov(X,Y)$

- $\boldsymbol{\sigma^2\left(\sum_{i=1}^{p} a_i X_i\right) = \sum_{i=1}^{p}\sum_{j=1}^{p} a_i a_j cov\left(X_i, X_j\right)}$

- $\boldsymbol{cov(aX + bY, Z) = a \times cov(X,Z) + b \times cov(Y,Z)}$

- $cov(X,Y) \leq \sigma^2 X \times \sigma^2 Y$

- ha $R(X,Y) = \pm 1$, akkor $\exists a, b \in R{:}\ P(X = aY + b) = 1$

- ez igaz fordítva is; gyakorlatilag azt jelenti, hogy a korreláció azt mutatja meg mennyire erős lineáris összefüggés áll fenn a két változó között

- kovarianciamátrix:

$$\Sigma = \left(cov\left(X_i, X_j\right)\right), i = 1,2 \dots p, j = 1,2 \dots p$$

Zlatniczki Ádám
adam.zlatniczki@cs.bme.hu

Feladatok

1. Példa

Háromszor dobunk egy szabályos kockával. X a kapott hatosok száma, Y a kapott páros értékek száma. Adja meg X és Y együttes eloszlását, kovariancia mátrixát! Független X és Y?

Megoldás:

Az együttes eloszlást tipikusan táblázattal szoktuk megadni:

Y\X	0	1	2	3	Y perem
0	$\frac{3^3}{6^3}$	0	0	0	$\frac{27}{216}$
1	$\frac{(3^2 \times 2)\binom{3}{1}}{6^3}$	$\frac{(3^2 \times 1)\binom{3}{1}}{6^3}$	0	0	$\frac{81}{216}$
2	$\frac{(3 \times 2^2)\binom{3}{2}}{6^3}$	$\frac{(1 \times 2 \times 3)3!}{6^3}$	$\frac{(1 \times 1 \times 3)\binom{3}{1}}{6^3}$	0	$\frac{81}{216}$
3	$\frac{(3^0 \times 2^3)\binom{3}{3}}{6^3}$	$\frac{(1 \times 2^2)\binom{3}{1}}{6^3}$	$\frac{(1 \times 1 \times 2)\binom{3}{1}}{6^3}$	$\frac{(1 \times 1 \times 1)}{6^3}$	$\frac{27}{216}$
X perem	$\frac{125}{216}$	$\frac{75}{216}$	$\frac{15}{216}$	$\frac{1}{216}$	Szum: 1

X és Y nem függetlenek, mert pl $P(X = 3, Y = 0) \neq P(X = 3)P(Y = 0)$.

$$\Sigma = \begin{pmatrix} cov(X,X) & cov(X,Y) \\ cov(Y,X) & cov(Y,Y) \end{pmatrix}$$

$$cov(X,X) = \sigma^2 X = EX^2 - E^2X = \frac{2}{3} - \frac{1}{2^2} = \frac{5}{12}$$

$$EX = \sum_{i=0}^{3} i \times P(X = i) = \frac{1}{2}$$

$$EX^2 = \sum_{i=0}^{3} i^2 \times P(X = i) = \frac{2}{3}$$

$$EY = \frac{3}{2}$$

$$EY^2 = 3$$

$$cov(Y,Y) = \frac{3}{4}$$

$$cov(X,Y) = cov(Y,X) = E(XY) - E(X)E(Y)$$

$$E(XY) = \sum_{i=0}^{3} \sum_{j=0}^{3} i \times j \times P(X = i, Y = j) = \frac{1}{4}$$

Zlatniczki Ádám
adam.zlatniczki@cs.bme.hu

2. Példa

Legyen $X \in E(2)$. Határozza meg a $cov(X, X^2)$ számot!

Megoldás:

$$cov(X, X^2) = E(X \times X^2) - E(X)E(X^2)$$

$$EX = \frac{1}{2}$$

$$EX^2 = \sigma^2 X + E^2 X = \frac{1}{2^2} + \frac{1^2}{2^2} = \frac{1}{2}$$

$$EX^3 = \int_0^\infty t^3 2e^{-2t} dt = \frac{3}{4}$$

$$cov(X, X^2) = \frac{3}{4} - \frac{1}{2} \times \frac{1}{2} = \frac{2}{4}$$

3. Példa

Bizonyítsa be, hogy ha X és Y azonos szórású valószínűségi változók, akkor X+Y és X-Y korrelálatlanok!

Megoldás:

A korreláció akkor 0, ha a kovariancia is az, tehát elég azt vizsgálni. $cov(X+Y, X-Y) = E[(X+Y)(X-Y)] - E(X+Y)E(X-Y)$

$$E[(X+Y)(X-Y)] = E(X^2 + Y^2) = EX^2 - EY^2$$

$$E(X+Y)E(X-Y) = (EX+EY)(EX-EY) = E^2X - E^2Y$$

$$cov(X+Y, X-Y) = EX^2 - EY^2 - E^2X + E^2Y$$

$$\sigma^2 X = EX^2 - E^2 X$$

$$\sigma^2 Y = EY^2 - E^2 Y$$

$$cov(X+Y, X-Y) = (EX^2 - E^2X) - (EY^2 - E^2Y) = \sigma^2 X - \sigma^2 Y$$

Mivel $\sigma^2 X = \sigma^2 Y$, így a kovariancia valóban 0.

4. Példa

Egy kalapban egy-egy cédulára fel vannak írva az 1,2,3 számjegyek. Egymás után, visszatevés nélkül húzunk kettőt. X az első, Y a második húzás eredménye. Adja meg $R(X, Y)$-t és döntse el, hogy függetlenek-e!

Megoldás:

Az együttes eloszlás táblázata:

Y\X	1	2	3	Y perem
1	0	1/6	1/6	1/3
2	1/6	0	1/6	1/3

3	1/6	1/6	0	1/3
X perem	1/3	1/3	1/3	Szum: 1

$$R(X,Y) = \frac{cov(X,Y)}{\sigma X \sigma Y}$$

$$cov(X,Y) = E(XY) - E(X)E(Y)$$

$$EX = \sum_{i=1}^{3} i \times P(X=i) = 2 = EY$$

$$E(XY) = \sum_{i=1}^{3}\sum_{j=1}^{3} i \times j \times P(X=i, Y=j) = \frac{11}{3}$$

$$cov(X,Y) = -\frac{1}{3}$$

$$\sigma^2 X = EX^2 - E^2 X$$

$$EX^2 = \sum_{i=1}^{3} i^2 \times P(X=i) = \frac{14}{3} = EY^2$$

$$\sigma X = \sqrt{\frac{14}{3} - 2^2} = \sqrt{\frac{2}{3}} = \sigma Y$$

$$R(X,Y) = \frac{-\frac{1}{3}}{\sqrt{\frac{2}{3}}\sqrt{\frac{2}{3}}} \neq 0$$

Tehát korreláltak, azaz nem lehetnek függetlenek.

5. Példa

Legyenek $X, Y \in N(0,1)$ függetlenek. $V = X + Y$ és $W = X - Y + 1$. Adja meg a $(V, W)^T$ vektor kovarianciamátrixát!

Megoldás:

$$\Sigma = \begin{pmatrix} cov(V,V) & cov(V,W) \\ cov(W,V) & cov(W,W) \end{pmatrix}$$

$$cov(V,V) = \sigma^2 V = \sigma^2(X+Y) = \sigma^2 X + \sigma^2 Y + cov(X,Y)$$

Mivel X és Y függetlenek, így $cov(X,Y) = 0$, tehát

$$cov(V,V) = 1 + 1 + 0 = 2$$

$$cov(W,W) = \sigma^2(W) = \sigma^2(X - Y + 1) = \sigma^2(X - Y) = \sigma^2 X + \sigma^2 Y - cov(X,Y) = 2$$

$$cov(V,W) = E(VW) - E(V)E(W) = E(VW) - 0 \times 0$$

$$E(VW) = E\big((X+Y)(X-Y+1)\big) = E(X^2 + X - Y^2 + Y)$$

$$EX^2 = \sigma^2 X + E^2 X = 1 + 0^2 = 1 = EY^2$$

$$E(VW) = 1 + 0 - 1 + 0 = 0$$

$$cov(V,W) = 0 = cov(W,V)$$

6. Példa

Legyen X,Y független valószínűségi változók, ahol EX=4, EY=0, $\sigma^2 X = 1$, $\sigma^2 Y = 2$. Határozza meg az alábbi mennyiségeket: $E(5X-6Y), \boldsymbol{E}XY, \sigma^2(5X-6Y+8), cov(5X,6Y)$!

Megoldás:

$$E(5X-6Y) = 5E(X) - 6E(Y) = 20 - 0 = 20$$

$$E(XY) \to mivel\ függetlenek \to E(X)E(Y) = 0$$

$$\sigma^2(5X-6Y+8) = \sigma^2(5X-6Y) \to függetlenek \to \sigma^2(5X) + \sigma^2(6Y) = 25\sigma^2 X + 36\sigma^2 Y = 25 + 72 = 97$$

$$cov(5X,6Y) = 30cov(X,Y) = 30 \times 0$$

7. Példa

Legyen $X \in N(-4{,}2), Y = 3X+1, Z = X^2 - 1$. Számolja ki $cov(Y,Z)$-t!

Megoldás:

$$cov(Y,Z) = E(YZ) - E(Y)E(Z)$$

$$E(Y) = E(3X+1) = 3EX + 1 = -11$$

$$EZ = E(X^2 - 1) = EX^2 - 1 = (\sigma^2 X + E^2 X) - 1 = (2 + 16) - 1 = 17$$

$$E(YZ) = E\big((3X+1)(X^2-1)\big) = E(3X^3 - 3X + X^2 - 1) = 3EX^3 - 3 \times (-4) + 18 - 1$$

$$EX^3 = \int_{-\infty}^{\infty} x^3 f_X(x)dx = \int_{-\infty}^{\infty} x^3 \varphi\left(\frac{x-EX}{\sigma X}\right)\frac{1}{\sigma X} = \int_{-\infty}^{\infty} x^3 \varphi\left(\frac{x+4}{2}\right)\frac{1}{2}\,dx = \int_{-\infty}^{\infty} (2v-4)^3 \varphi(v)\frac{1}{2}dv$$
$$= \int_{-\infty}^{\infty} (4v^3 - 24v^2 + 48v - 32)\varphi(v)dv = 4\boldsymbol{E}V^3 - 24\boldsymbol{E}V^2 + 48\boldsymbol{E}V - 32$$

Kihasználva, hogy $V \in N(0{,}1)$

$$EV^3 = EV = 0$$

$$EV^2 = \sigma^2 V + E^2 V = 1 + 0^2$$

$$EX^3 = -56$$

$$cov(Y,Z) = 3 \times (-56) + 29 - (-11) \times 17 = -168 + 29 + 187$$

Zlatniczki Ádám
adam.zlatniczki@cs.bme.hu

8. Példa

Legyen $X \in N(m, D), Y = 3X + 8, Z = 5 - 2X$. Számolja ki Y és Z korrelációs együtthatóját!

Megoldás:

$$cov(Y, Z) = cov(3X + 8, 5 - 2X) = cov(3X, -2X) = -6cov(X, X) = -6D^2$$

$$\sigma Y = 3D$$

$$\sigma Z = 2D$$

$$R(Y, Z) = -1$$

Másik megoldás:

$Y = \frac{3}{2}Z + \frac{31}{2} \rightarrow R(Y, Z) = -1$

9. Példa

Legyen $X \in U(0,2), Y = \cos X, Z = \sin X$. Mekkora $cov(Y, Z)$? Független-e Y és Z?

Megoldás:

$$EY = \int_0^2 \cos x \times f_X(x)dx = \int_0^2 \cos x \frac{1}{2} dx = \frac{\sin 2}{2}$$

$$EZ = \int_0^2 \sin x \times f_X(x)dx = \int_0^2 \sin x \frac{1}{2} dx = \frac{1 - \cos 2}{2}$$

$$E(YZ) = E(\sin X \cos X) = 0{,}5E(\sin 2X) = 0{,}5 \int_0^2 \sin 2x \frac{1}{2} dx = \frac{1 - \cos 4}{8}$$

$$cov(Y, Z) = \frac{1 - \cos 4}{8} - \frac{\sin 2}{2} \times \frac{1 - \cos 2}{2} \approx 0{,}216$$

cov(Y,Z) > 0, így biztosan nem függetlenek.