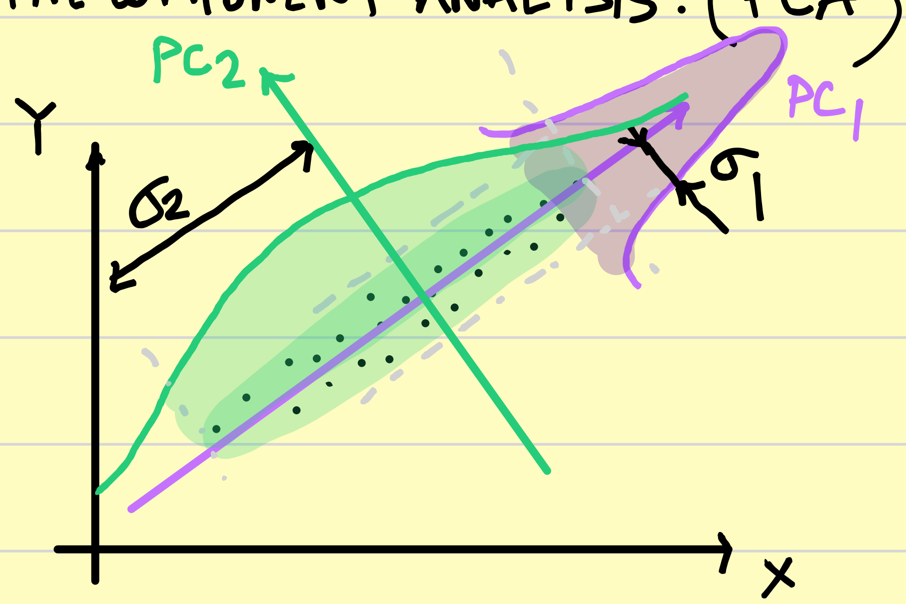
20250416_AKIB2_Python

HAUPTKOMPONENTEN ANALYSE. Dimensionalitätsreduktionsalgorithmen

PRINCIPAL COMPONENT ANALYSIS. (PCA)



Graphische Intuition

$\sigma_1 << \sigma_2$

n: Anzahl Datensätze

- Der Vektor → beschreibt die Daten am besten.
- Warum? Weil die Variabilität (Std Abweichung) der Daten in dieser Richtung minimal ist.
- Diese Richtung in der die Variabilität minimal ist, erklärt die Daten am besten und heißt: HAUPTKOMPONENTE 1. (PC_1).
- Senkrecht zu PC_1 ist PC_2. PC_2 erklärt wenigere Variabilität (höhere Std Abweichung) der Daten. →

$\sigma_1 << \sigma_2$ (!)

- Die volle Variabilität von einem „n" dimensionalen System wird mit „n" Hauptkomponenten dargestellt.

> Def. Die Hauptkomponenten eines Systems sind die EIGEVEKTOREN der KOVARIANZMATRIX.

KOVARIANZ MATRIX von einem System mit n VARIABLEN wird wie folgt definiert:

$$\underset{\text{Mat}}{\text{KOV}}\left[X_1, X_2, \dots, X_n\right] = \begin{bmatrix} \text{VAR}[X_1] & \text{KOV}[X_1, X_2] & \cdots & \text{KOV}[X_1, X_n] \\ \text{KOV}[X_2, X_1] & \text{VAR}[X_2] & \cdots & \text{KOV}[X_2, X_n] \\ \vdots & \vdots & & \\ \text{KOV}[X_n, X_1] & \text{KOV}[X_n, X_2] & \cdots & \text{VAR}[X_n] \end{bmatrix}$$

$i = 1, 2, \dots, n$

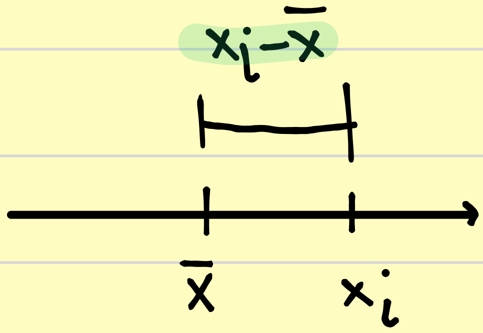
$$\text{VAR}[x_i] = \frac{\sum (x_i - \bar{x}_i)^2}{n-1} \qquad \text{KOV}[x_i, x_j] = \frac{\sum (x_i - \bar{x}_i)(x_j - \bar{x}_j)}{n-1}$$

$$\| $$

$$\text{KOV}[x_j, x_i] = \frac{\sum (x_j - \bar{x}_j)(x_i - \bar{x}_i)}{n-1}$$

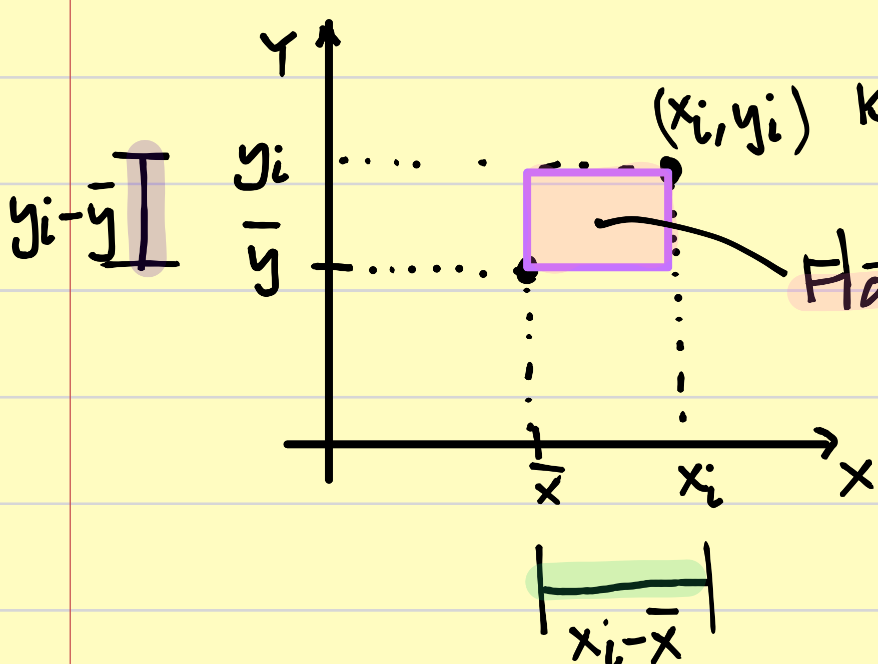
$$\boxed{\underset{\text{Mat}}{\text{KOV}}\left[X_1, X_2, \dots, X_n\right] = \begin{bmatrix} \text{VAR}[X_1] & \text{KOV}[X_1, X_2] & \cdots & \text{KOV}[X_1, X_n] \\ \text{KOV}[X_1, X_2] & \text{VAR}[X_2] & & \text{KOV}[X_2, X_n] \\ & & & \\ \text{KOV}[X_1, X_n] & \text{KOV}[X_2, X_n] & & \text{VAR}[X_n] \end{bmatrix}}$$

Graphische Interpretation Varianz & Kovarianz



$$VAR[X] = \frac{\sum (x_i - \bar{x})^2}{n-1}$$

· Die Varianz ist die Summe der Abstände zum Mittelwert in Quadrat durch $(n-1)$.



$$KOV[X,Y] = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{n-1}$$

$$\text{Fläche} = (x_i - \bar{x})(y_i - \bar{y})$$

· Die Kovarianz ist die Summe der Flächen der Vierechen zum Mittelwert durch $(n-1)$.

Beispiel: Eine Firma steuert sich mit einem 2. Dimensionalen Kennzahlen System, DLZ & Kosten pro Stück. Bitte ermitteln Sie die Kovarianzmatrix des Systems.

DLZ [Tage]: $\{17, 19, 21, 27\} \mathrel{\hat{=}} x$

€/Stück [€/St]: $\{7, 11, 12, 15\} \mathrel{\hat{=}} y$

$$\text{Kov Mat.}[X,Y] = \begin{bmatrix} VAR[X] & KOV[X,Y] \\ KOV[X,Y] & VAR[Y] \end{bmatrix} = \begin{bmatrix} 18'67 & 13'33 \\ 13'33 & 10'92 \end{bmatrix}$$

$$\cdot \ VAR[X] = \frac{\sum (x_i - \bar{x})^2}{n-1} = \frac{(17-21)^2 + (19-21)^2 + (21-21)^2 + (27-21)^2}{4-1} = 18'67$$

$$\bar{x} = \frac{17+19+21+27}{4} = 21$$

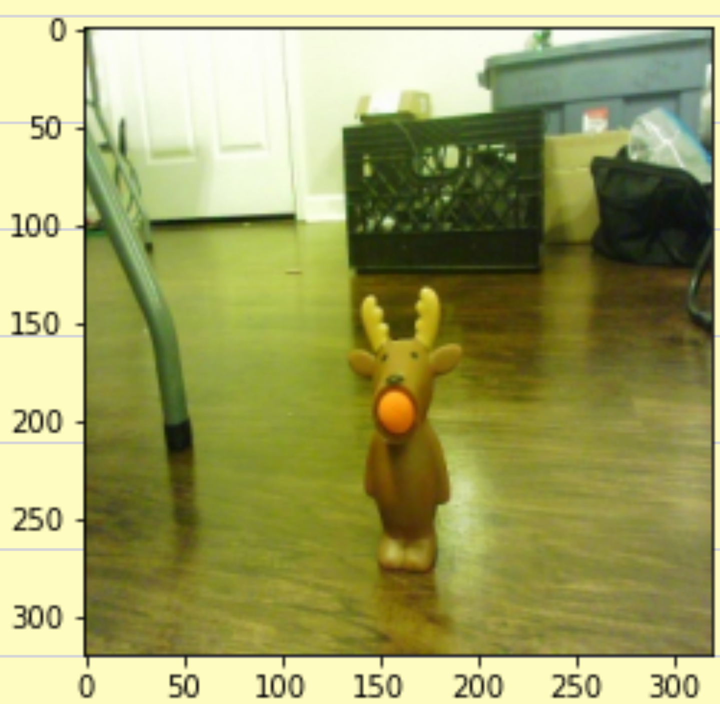
$$\cdot \ VAR[Y] = \frac{\sum (y_i - \bar{y})^2}{n-1} = \frac{(7-11'25)^2 + (11-11'25)^2 + (12-11'25)^2 + (15-11'25)^2}{4-1} = 10'92$$

$$\bar{Y} = \frac{7+11+12+15}{4} = 11'25$$

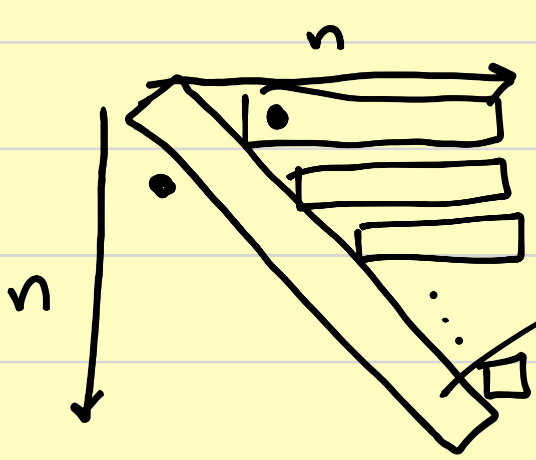
$$\cdot \ KOV[X,Y] = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{n-1} = \frac{(17-21)(7-11'25) + (19-21)(11-11'25) + (21-21)(12-11'25) + (27-21)(15-11'25)}{4-1} =$$

$$= \frac{17 + 0'5 + 0 + 22'5}{3} = 13'33$$

Bild Beispiel :



$325 \times 325 \rightarrow KOV\,[325 \times 325]$



n Elementen VAR

$$n + (n-1) + (n-2) + \cdots + 1 = ?$$

$$n = 3 \rightarrow 3+2+1 = 6$$

$$n = 4 \rightarrow 4+3+2+1 = 10$$

Eigenvektoren einer Matrix A ermitteln:
Eigenwerte

$A\vec{v} = \lambda \vec{v}$ $\vec{v}$: Eigenvektoren λ: Eigenwerte

· $\det[A - \lambda I] = 0$ I: Identitätsmatrix

$\downarrow$

λ ✓

$\downarrow$

$\vec{v}$ ✓

Beispiel: $A = \begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix}$ Eigenwerte bzw. Eigenvektoren

$$\det[A - \lambda I] = 0 \rightarrow \det\left[\begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right] = 0 \rightarrow$$

$$\rightarrow \det\left[\begin{bmatrix} 0-\lambda & 1 \\ -2 & -3-\lambda \end{bmatrix}\right] = 0 \rightarrow (-\lambda)(-3-\lambda) - 1(-2) = 0$$

$$\boxed{\det\begin{bmatrix} a & c \\ b & d \end{bmatrix} = a.d - c.b}$$

$$\rightarrow 3\lambda + \lambda^2 + 2 = 0 \rightarrow \lambda^2 + 3\lambda + 2 = 0$$

$$\lambda = \frac{-3 \pm \sqrt{9 - 4.2}}{2} = \begin{matrix} \nearrow \lambda_1 = -1 \\ \searrow \lambda_2 = -2 \end{matrix}$$ Eigenwerte

$$\boxed{ax^2 + bx + c = 0 \rightarrow x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}}$$

$$\boxed{\lambda_1 = -1} \rightarrow A\cdot\vec{v_1} = \lambda_1\cdot\vec{v_1} \rightarrow \begin{bmatrix}0&1\\-2&-3\end{bmatrix}\begin{bmatrix}v_{11}\\v_{12}\end{bmatrix} = (-1)\begin{bmatrix}v_{11}\\v_{12}\end{bmatrix}$$

$$\rightarrow \left.\begin{aligned} 0\cdot v_{11}+1\cdot v_{12} &= -v_{11}\\ -2v_{11}-3v_{12} &= -v_{12}\end{aligned}\right| \triangleright \begin{aligned} v_{11}&=-v_{12}\\ v_{11}&=-v_{12}\end{aligned}$$

$$v_{11} = 1 \rightarrow v_{12} = -1 \rightarrow \vec{v_1} = \begin{bmatrix}1\\-1\end{bmatrix}$$ 1. Eigenvektor

$$\boxed{\lambda_2 = -2} \rightarrow A\cdot\vec{v_2} = \lambda_2\vec{v_2} \rightarrow \begin{bmatrix}0&1\\-2&-3\end{bmatrix}\begin{bmatrix}v_{21}\\v_{22}\end{bmatrix} = (-2)\begin{bmatrix}v_{21}\\v_{22}\end{bmatrix}$$

$$\rightarrow \left.\begin{aligned} 0\cdot v_{21}+1\cdot v_{22} &= -2v_{21}\\ -2v_{21}-3v_{22} &= -2v_{22}\end{aligned}\right| \triangleright \begin{aligned} v_{22}&=-2v_{21}\\ v_{22}&=-2v_{21}\end{aligned}$$

$$v_{21} = 1 \rightarrow v_{22} = -2 \rightarrow \vec{v_2} = \begin{bmatrix}1\\-2\end{bmatrix}$$ 2. Eigenvektor

Beispiel Kov. Matrix

$$\text{Kov Mat.}[X,Y] = \begin{bmatrix}VAR[X] & KOV[X,Y]\\ KOV[X,Y] & VAR[Y]\end{bmatrix} = \begin{bmatrix}18'67 & 13'33\\ 13'33 & 10'92\end{bmatrix}$$

$$\det[A-\lambda I]=0 \rightarrow \left|\begin{bmatrix}18'67 & 13'33\\ 13'33 & 10'92\end{bmatrix} - \begin{bmatrix}\lambda & 0\\ 0 & \lambda\end{bmatrix}\right| = 0 \rightarrow \begin{vmatrix}18'67-\lambda & 13'33\\ 13'33 & 10'92-\lambda\end{vmatrix} = 0$$

$$\rightarrow (18'67-\lambda)(10'92-\lambda) - 13'33^2 = 0 \rightarrow$$

$$\rightarrow \lambda^2 - 29'59\lambda + 26'187 = 0$$

$$\lambda = \frac{29'59 \pm \sqrt{875'6 - 104'75}}{2} = \frac{29'59 \pm 27'76}{2} = \begin{cases}\lambda_1 = 28'67\\ \lambda_2 = 0'915\end{cases}$$

$$\lambda_1 = 28'67 \to \begin{bmatrix} 18'67 & 13'33 \\ 13'33 & 10'92 \end{bmatrix} \begin{bmatrix} v_{11} \\ v_{12} \end{bmatrix} = 28'67 \begin{bmatrix} v_{11} \\ v_{12} \end{bmatrix} \to$$

$$\to \begin{matrix} 18'67\, v_{11} + 13'33 v_{12} = 28'67\, v_{11} \\ 13'33 v_{11} + 10'92 v_{12} = 28'67\, v_{12} \end{matrix} \to$$

$$\to 13'33 v_{12} = 10 v_{11} \to v_{11} = 1'333 v_{12}$$

$$v_{11} = 1 \to v_{12} = \frac{1}{1'333} = 0'75 \to \vec{v}_1 = \begin{bmatrix} 1 \\ 0'75 \end{bmatrix} \checkmark$$

$$\lambda_2 = 0'915 \to \begin{bmatrix} 18'67 & 13'33 \\ 13'33 & 10'92 \end{bmatrix} \begin{bmatrix} v_{21} \\ v_{22} \end{bmatrix} = 0'915 \begin{bmatrix} v_{21} \\ v_{22} \end{bmatrix} \to$$

$$\to \begin{matrix} 18'67 \cdot v_{21} + 13'33 v_{22} = 0'915\, v_{21} \\ 13'33 \cdot v_{21} + 10'92 v_{22} = 0'915\, v_{22} \end{matrix} \to$$

$$\to 17'755 v_{21} = -13'33 v_{22} \to v_{21} = -0'75 v_{22}$$

$$v_{22} = 1 \to v_{21} = -0'75 \to \vec{v}_2 = \begin{bmatrix} -0'75 \\ 1 \end{bmatrix} \checkmark$$