

Cvičení 2 – příklady u tabule

Příklad 1.: (př. 4.5.1. ze skript) Na pěti objektech byly zjišťovány hodnoty dvou znaků. Datový soubor je tvaru

$$\begin{pmatrix} 3 & 7 \\ 5 & 6 \\ 6 & 8 \\ 7 & 10 \\ 9 & 9 \end{pmatrix}$$
. Najděte vektor průměrů, výběrovou varianční matici a výběrovou korelační matici. (Pro lepší interpretaci lze říci, že se jedná o

výsledky testů pěti studentů ze dvou předmětů.)

Příklad 2.: (př. 4.5.2. ze skript) Na 24 objektech byly pozorovány znaky X_1 , X_2 a X_3 . Z datového souboru byla vypočtena výběrová varianční

matice $\mathbf{S} = \begin{pmatrix} 451,39 & 271,17 & 168,70 \\ 271,17 & 171,73 & 103,29 \\ 168,70 & 103,29 & 66,69 \end{pmatrix}$. Vlastní čísla získaná řešením rovnice $|\mathbf{S} - \lambda \mathbf{I}| = 0$ a jim odpovídající vlastní vektory jsou:

$l_1 = 680,411$, $l_2 = 6,5016$, $l_3 = 2,8573$, $\mathbf{v}_1 = (0,8126; 0,4955; 0,3068)^T$, $\mathbf{v}_2 = (0,5454; -0,8321; -0,1009)^T$, $\mathbf{v}_3 = (0,2053; 0,2493; -0,9464)^T$.

Vyjádřete hlavní komponenty a určete, kolik procent variability obsažené v matici \mathbf{S} každá z nich vyčerpává.

Příklad 3.: Pro údaje z příkladu 1. sestavte matici euklidovských vzdáleností.

Příklad 4.: (př. 4.5.3. ze skript) Pomocí matice euklidovských vzdáleností vypočtené v příkladě 3. proveďte shlukovou analýzu metodou nejvzdálenějšího souseda. Výsledek znázorněte dendrogramem.