

Předmět: 3.2 Metody s latentními proměnnými a klasifikační metody
Přednášející: Prof. Ing. Oldřich Pytela, DrSc.

Zadání: Do příštího soustředění předložte v rámci své zkoušky z tohoto předmětu následující písemnou semestrální práci.

Otázka 1. Vypočtete algoritmem NIPALS 1. latentní proměnnou z matice A [řádek,sloupec]: $A[1,1]=1$, $A[2,1]=2$, $A[3,1]=3$, $A[1,2]=1$, $A[2,2]=2$, $A[3,2]=0$, $A[1,3]=6$, $A[2,3]=4$, $A[3,3]=2$. Matici před zpracováním standardizujte.

Otázka 2. S použitím vhodných kritérií určete nezbytný počet latentních proměnných, bylo-li z dat určeno: $PRESS(0)=S(0)=100$, $PRESS(1)=20$, $S(1)=10$, $PRESS(2)=3.5$, $S(2)=3.4$, $PRESS(3)=3.45$, $S(3)=3.39$.

Otázka 3. Odhadněte hodnotu chybějícího prvku $A[2,2]$, jestliže výpočtem z nekompletní matice byly určeny vektory p : 0.541 0.423 0.514 0.514 t : -1.340 -0.735 2.076

Otázka 4. Výpočtem metodou PCA byly určeny vektory p_1 : 0.012 0.458 -0.352 0.987 p_2 : 0.926 -0.238 0.872 -0.115 Vypočtete komunalitu a vyberte sloupec, který nejlépe charakterizuje celou matici.

Otázka 5. Vysvětlete, proč vysvětlená variabilita je při výpočtu metodou FA vždy nižší, než při výpočtu metodou PCA.

Otázka 6. Výpočtem metodou kanonických korelací bylo zjištěno: $0.297 X_1 + 0.298 X_2 + 0.050 X_3 + 0.256 X_4 = 0.493 Y_1 - 0.213 Y_2$ $r_1 = 0.830$
 $0.006 X_1 - 0.115 X_2 + 0.950 X_3 + 0.056 X_4 = 0.493 Y_1 + 0.213 Y_2$ $r_1 = 0.512$ Vypočtete skupinový korelační koeficient a interpretujte výsledky.

Otázka 7. Uveďte nějaký konkrétní příklad vhodný pro zpracování metodou metodou PLS.

Otázka 8. Jeden objekt je charakterizován metrickými znaky (2,10), druhý (3,8), třetí (4,9), čtvrtý (10,4) a pátý (11,5). Vypočtete matici vzdáleností v Euklidově metrice a dokumentujte výpočet shlukování některou z používaných metod. Výsledky interpretujte graficky.

Otázka 9. Popište slovně postup aplikace metod s latentními proměnnými nebo klasifikačních metod na nějakém konkrétním příkladu ze své praxe.

Výjimka: Místo výše zadaných otázek je možné v tomto předmětu 3.2 dle eventuelního zadání přednášejícího prof. Pytely ověřit Vaše úlohy již řešené softwarem STATISTICA a NCSS2007 rovněž v tomto předmětu jeho novým programem OPTIM.

Formát: Text úloh s tabulkami napište editorem Word, obrázky a výpočty pak s využitím softwaru OPTIM. Každá úloha bude vypracována dle formátu a stylu vzorových úloh v monografii *M. Meloun, J. Militký: Kompendium statistického zpracování experimentálních dat, Academia Praha 2006*, a bude vždy obsahovat: *Nadpis úlohy, Zadání, Data, Užitý program, Output, Obrázky, Komentované řešení, Závěr*. Důraz při hodnocení je kladen především na komentované řešení a vysvětlení postupu řešení, interpretaci výstupu užitého software. Na přiloženém CD, které je rovněž předmětem hodnocení a které současně s písemnou prací odevzdáte, budou výsledky na obalu a také na kotoučku CD evidovány následovně: jméno a příjmení, název úlohy, datum. Úlohy odevzdejte jednak vytištěné na papíru a jednak na CD ve dvou formách, *.doc a *.pdf včetně inputů a outputů použitých programů.