

Programowanie w R - 04

jaworski@prz.edu.pl, e-rejwer@prz.edu.pl

L Zajęcia laboratoryjne – czynniki i macierze

Zadanie L.1 factor, levels, cut

1. Niech $x = c(1, 2, 3, 3, 5, 3, 2, 4, NA)$, jakie są poziomy czynnika x ?
2. Niech $z \leftarrow \text{factor}(c("p", "q", "p", "r", "q"))$ będzie czynnikiem, którego poziomy to: „p”, „q”, „r”. Zmień poziom „p” na „w”, co da nam czynnik z postaci: „w”, „q”, „w”, „r”, „q”.
3. Niech: $s1 \leftarrow \text{factor}(\text{sample}(\text{letters}, \text{size}=5, \text{replace}=\text{TRUE}))$, $s2 \leftarrow \text{factor}(\text{sample}(\text{letters}, \text{size}=5, \text{replace}=\text{TRUE}))$, połącz $s1$ i $s2$ w jeden czynnik.
4. Rozważ zmienną `Sepal.Length` w zestawie danych `iris`. Z danych określonych przez tę zmienną utwórz czynnik, który przypisze dane do pięciu poziomów odpowiadającym domyślnie określonym przedziałom:
(4.3, 5.02] (5.02, 5.74] (5.74, 6.46] (6.46, 7.18] (7.18, 7.9].
5. Niech $\text{fodp} \leftarrow \text{factor}(c("Tak", "Tak", "Nie wiem", "Nie", "Tak"))$. Wprowadźmy nowy poziom "Absolutnie nie".

Zadanie L.2 factor, sapply, attributes, levels, cut, droplevels

1. Zainstaluj (`install.packages()`) i załaduj (`library()`) pakiety `gapminder` oraz `tidyverse`. Sprawdź, które zmienne w pakiecie `gapminder` są typu `factor` i ile mają poziomów.
2. Dodaj kontynent `Antarctica`.
3. Na kontynencie (poziomie) `Antarctica` nie ma stałej populacji ludzkiej, zatem usuń ten poziom z czynnika.
4. Dodaj dwa nowe poziomy: zamiast "Americas", dodaj "South America", "North America". Kraje w wektorze $\text{sam.c} \leftarrow c("Argentina", "Bolivia", "Brazil", "Chile", "Colombia", "Ecuador", "Paraguay", "Peru", "Uruguay", "Venezuela")$ należy przypisać do "South America".
5. Ułóż poziomy czynnika `continent` w kolejności alfabetycznej.
6. Ponownie uporządkuj poziomy czynnika `continent`, aby pojawiały się w kolejności według całkowitej liczby ludności w 2007 r.

Zadanie L.3 matrix

Stwórz cztery wektory liczbowe, każdy o trzech elementach: v_1 będący permutacją liczb -1 i 1 , v_2 o wartościach $1, 2, 3$, v_3 zawierający zera oraz v_4 zawierający wartości losowe o rozkładzie normalnym. Następnie

1. stwórz macierz A_1 której wiersze będą kolejnymi wektorami; po utworzeniu macierzy dodaj nazwy wierszy zgodne z nazwami wektorów od których pochodzą oraz następujące nazwy kolumn: "no.1", "no.2", "no.3",
2. stwórz macierz A_2 której wiersze będą kolejnymi kolumnami oraz jednocześnie, w tej samej instrukcji zdefiniuj nazwy kolumn i wierszy analogicznie do powyższych,
3. oblicz iloczyny macierzy $A_3 = A_1 \cdot (A_1)^T$, $A_4 = (A_1)^T \cdot A_1$ – algebraiczne iloczyny, gdzie T oznacza transpozycję macierzy,
4. dla odmiany względem powyższego wykonaj działanie "*" na macierzach A_1 oraz $(A_2)^T$ (iloczyn odpowiadających sobie elementów),

P Zajęcia projektowe – czynniki i macierze

Zadanie P.1 factor, attributes, levels

1. Najpierw sprawdź, co konsola R podpowiada na temat pakietu `mtcars`, a następnie zobacz, jaki jest wynik, działania polecenia `attributes` dla `textttmtcars`.
2. Zdefiniuj zmienną o nazwie `numV` o wartościach 1, 2, 3, 4 i 5. Następnie, utwórz zmienną czynnikową o nazwie `factNumV` o takich samych wartościach. Porównaj atrybuty dla tych zmiennych.
3. Zdefiniuj zmienną o nazwie `charV` o wartościach A, B, C, C i C. Następnie, utwórz zmienną czynnikową `factCharV` o tych samych wartościach. Porównaj atrybuty dla tych zmiennych.
4. Zmodyfikuj pierwszy poziom `factCharV` z "A" na "a" oraz drugi poziom `factCharV` z "B" na "b". Porównaj zmienną i poziomy zmiennej czynnikowej `factCharV` przed i po modyfikacji.
5. Utwórz zmienną czynnikową `Plec` z wektora `c("M", "M", "F")`. Zmień poziomy płci na "Mężczyzna" i "Kobieta", odpowiednio dla "M" i "F". Spróbuj użyć funkcji czynnikowej, aby zrobić to w jednym kroku (w tym celu zapoznaj się z dokumentacją `factor`).
6. Utwórz zmienną czynnikową o wartościach `c("M", "M", "F", NA)` i wyklucz NA. Utwórz drugą zmienną czynnikową, wyłączając tylko "F" oraz trzecią zmienną, wyłączając zarówno NA, jak i "F". Porównaj strukturę tych trzech zmiennych czynnikowych.

Zadanie P.2 factor, sapply, attributes, levels, cut, droplevels

Wykonaj wszystkie poniższe ćwiczenia przy użyciu pakietu `forcats`. Staraj się również użyć pakietu `tidyverse` podobnie, jak w załączonych rozwiązaniach zadań laboratoryjnych.

1. Załaduj pakiet `forcats`. Znajduje się tam zestaw danych `gss_cat` (sprawdź `str(gss_cat)`). Dla kolumny `rincome` oblicz liczbę elementów (tej kolumny) dla każdego poziomu.
2. Zmień kolejność poziomów kolumny `rincome` zgodnie z kolejnością (order) występowania poziomów w zbiorze danych.
3. Zmień kolejność poziomów kolumny `rincome` zgodnie z częstotliwością (frequency) występowania poziomów w zbiorze danych.
4. Wszystkie wartości pojawiające się w kolumnie `rincome`, poza dziesięcioma najczęściej występującymi, połącz razem w kategorię "other answer".
5. Zamień poziomy "Not applicable", "Refused", "Don't know" i "No answer", na "Other". Zamień pozostałe poziomy (przedziały liczbowe) tak, aby miały nazwy "x to y".
6. Podziel poziomy (przedziały liczbowe) na grupy: "Below 10k", "10k – 25k", "Above 25k".
7. Usuń puste poziomy z kolumny `race`.

Zadanie P.3 cbind, rbind, upper.tri

Wygeneruj wektor v_5 składający się z 36 liczb ze zbioru $\{-2, 0, 1, 2, 4, 10\}$, następnie

1. utwórz macierze zbudowane z elementów wektora v_5 kolejno o liczbie kolumn równej 2, 3, 4, 8, 12, gdzie elementy będą umieszczane "kolumnami",
2. to co powyżej tylko że elementy umieszczanie powinny być wierszami,
3. utwórz macierz kwadratową A_3 z elementów wektora v_5 i następnie oblicz jej wyznacznik oraz macierz odwrotną,
4. utwórz macierz kwadratową A_4 o wymiarze 4×4 , będącą podmacierzą macierzy A_3 zawierającą kolumny od 2 do 5 oraz wiersze od 3 do 6,
5. wydobądź z macierzy A_3 macierz A_{tg} górną trójkątną zawierającą wartości na diagonalu,
6. wydobądź z macierzy A_3 elementy znajdujące się na diagonalu,
7. przypisz do A_4 macierz A_3 a następnie usuń z A_4 trzeci wiersz oraz piątą kolumnę,
8. stwórz macierz A_5 będą sklejeniem względem kolumn dwóch macierzy A_3 .

Zadanie P.4

Stwórz macierz A_k będącą trzydziestokrotnym sklejeniem macierzy `bbb <- matrix(1:841, ncol=29)` z samą sobą względem kolumn oraz analogiczną macierz A_w będącą sklejeniem względem wierszy (warto to zrobić bez wykorzystania `cbind` oraz `rbind`)

Zadanie P.5 `solve`

Wykorzystując macierzową postać układu oraz funkcję `solve` rozwiąż poniższy układ równań liniowych

$$\begin{cases} 3x + 12y &= 3 \\ x + 5y &= 3 \\ 2x + 4y + 2z &= 10 \end{cases}$$

oraz sprawdź czy iloczyn macierzy głównej oraz rozwiązania jest równy kolumnie wyrazów wolnych. Oblicz wartości własne macierzy głównej układu równań.