

## Ćwiczenie 2: Metody redukcji wymiarowości – nieujemna faktoryzacja macierzy i dekompozycje tensorów

### Zadania:

1. Wygenerować faktory:  $\mathbf{A} = [a_{ij}] \in \mathbb{R}_+^{I \times J}$  i  $\mathbf{X} = [x_{jt}] \in \mathbb{R}_+^{J \times T}$ , gdzie  $a_{ij} = \max(0, \check{a}_{ij})$  i  $x_{jt} = \max(0, \check{x}_{jt})$  oraz  $\check{a}_{ij}, \check{x}_{jt} \sim N(0,1)$  (rozkład normalny). Wygeneruj syntetyczne obserwacje  $\mathbf{Y} = \mathbf{AX}$  dla  $I = 100, T = 1000, J = 10$ . Stosując wybrane algorytmy NMF (ALS, MUE, HALS) wyznacz estymowane faktory  $\hat{\mathbf{A}}$  i  $\hat{\mathbf{X}}$  oraz unormowany błąd residualny w funkcji iteracji naprzemiennych. Oceń jakość estymacji stosując miary MSE (ang. *Mean-Squared Error*) lub SIR (ang. *Signal-to-Interference Ratio*).
2. Wygenerować faktory:  $\mathbf{U}^{(1)} = [u_{i_1 j}^{(1)}] \in \mathbb{R}_+^{I_1 \times J}$ ,  $\mathbf{U}^{(2)} = [u_{i_2 j}^{(2)}] \in \mathbb{R}_+^{I_2 \times J}$ ,  $\mathbf{U}^{(3)} = [u_{i_3 j}^{(3)}] \in \mathbb{R}_+^{I_3 \times J}$ , gdzie  $u_{i_1 j}^{(1)} = \max(0, \check{u}_{i_1 j}^{(1)})$ ,  $u_{i_2 j}^{(2)} = \max(0, \check{u}_{i_2 j}^{(2)})$ ,  $u_{i_3 j}^{(3)} = \max(0, \check{u}_{i_3 j}^{(3)})$  oraz  $\check{u}_{i_1 j}^{(1)}, \check{u}_{i_2 j}^{(2)}, \check{u}_{i_3 j}^{(3)} \sim N(0,1)$  (rozkład normalny). Wygeneruj syntetyczne obserwacje  $\mathbf{Y}$  dla  $I_1 = 10, I_2 = 20, I_3 = 30, J = 5$ . Stosując wybrane algorytmy NTF (np. ALS) wyznacz estymowane faktory  $\hat{\mathbf{U}}^{(1)}, \hat{\mathbf{U}}^{(2)}$  i  $\hat{\mathbf{U}}^{(3)}$  oraz unormowany błąd residualny w funkcji iteracji naprzemiennych. Oceń jakość estymacji stosując miary MSE (ang. *Mean-Squared Error*) lub SIR (ang. *Signal-to-Interference Ratio*).
3. Obrazy twarzy z bazy ORL (lub podobnej) przedstaw za pomocą tensora  $\mathbf{Y} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$ , gdzie  $I_3$  jest liczbą obrazów. Rozdziel obrazy na zbiory trenujący i testujący według odpowiedniej zasady, np. 5-folds CV i utwórz odpowiednie tensory trenujący  $\mathbf{Y}_r$  i testujący  $\mathbf{Y}_t$ . Tensor trenujący poddaj dekompozycji CP (np. algorytmem ALS) oraz HOSVD dla  $J = 4, 10, 20, 30$ . Pogrupować obrazy stosując metodę k-średnich do faktora  $\hat{\mathbf{U}}^{(3)}$ . Badania przeprowadzić dla różnej liczby grup. Porównać dokładność grupowania z metodą PCA (z poprzedniego ćwiczenia). Następnie dokonaj projekcji obrazów z tensora  $\mathbf{Y}_t$  na podprzestrzeń cech generowaną czynnikami otrzymanymi z  $\mathbf{Y}_r$ . Dokonaj klasyfikacji obrazów w przestrzeni cech w  $\hat{\mathbf{U}}^{(3)}$  za pomocą klasyfikatora k-NN. Porównać efekty klasyfikacji różnymi metodami (np. PCA, CP, HOSVD).