

Name: Matouš Dzivjak
Date: 9.11.2018

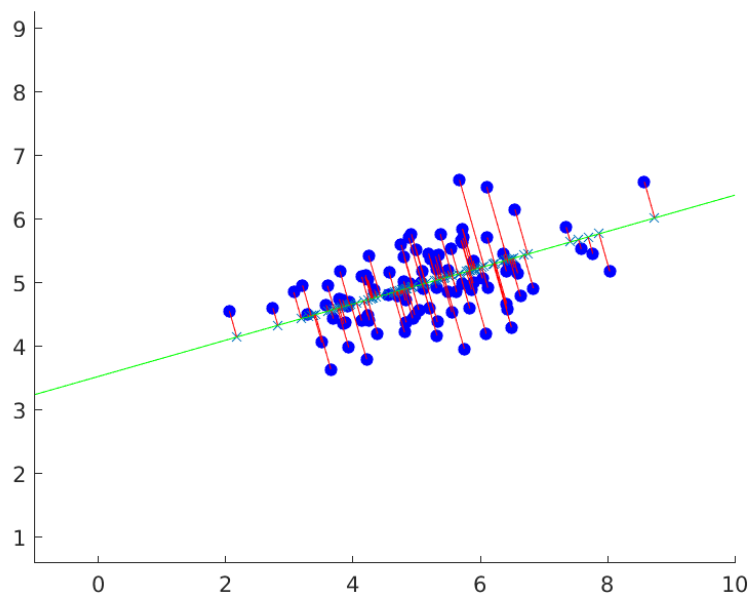
PCA

1 Zadání

https://gitlab.fel.cvut.cz/B0B330PT/public/blob/master/cviceni/03_pca/03_pca.pdf

2 Řešení 1. části

2.1



2.2

Součet čtverců kolmých vzdáleností je $err = 24.2419$

2.3

2.3.1

Nejdříve máme řešení reprezentovat ve formě $\{y \in \mathbb{R}^2 | y^T \cdot x = \alpha\}$
Kde

$$y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

$$\|y\| = 1$$

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Po roznásobení pro přímku dostáváme rovnici

$$y_1 \cdot x_1 + y_2 \cdot x_2 = \alpha$$

Po dosazení normálového vektoru přímky n (kolmý na směrový vektor přímky, za y_1 a y_2) a souřadnice těžiště bodů (za x_1 a x_2) dostáváme $\alpha = -3.3931$

$$y = \begin{bmatrix} -0.2740 \\ 0.9617 \end{bmatrix}$$

$$x = \begin{bmatrix} 5.1231 \\ 4.9876 \end{bmatrix}$$

2.3.2

V druhé části máme řešení reprezentovat ve formě: $\{y_0 + t \cdot s | t \in \mathbb{R}\}$. To však máme stíženo a omezeno podmínkou, že norma $\|y_0\|$ musí být minimální. Tedy musíme nalézt bod na přímce, který je nejbližší počátku (zatím co jinak bychom mohli použít například již nalezené těžiště bodů).

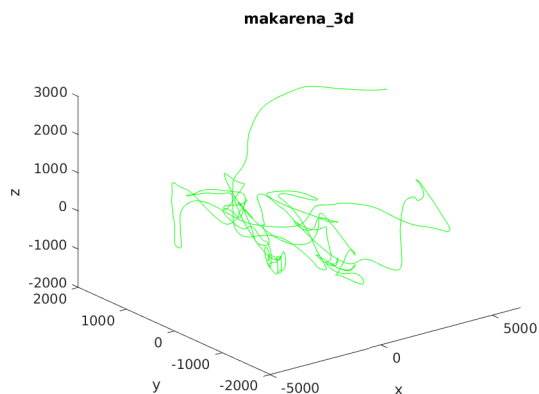
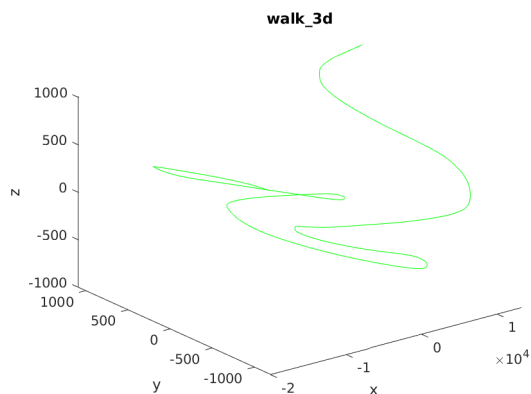
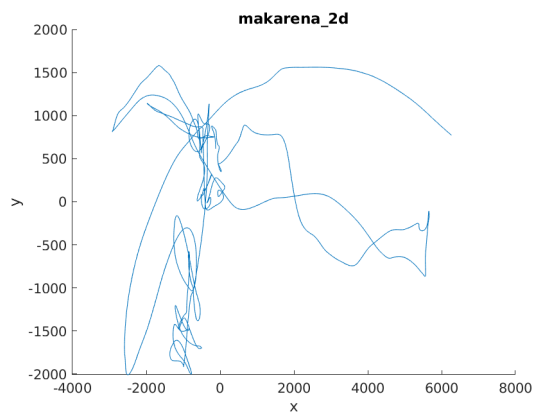
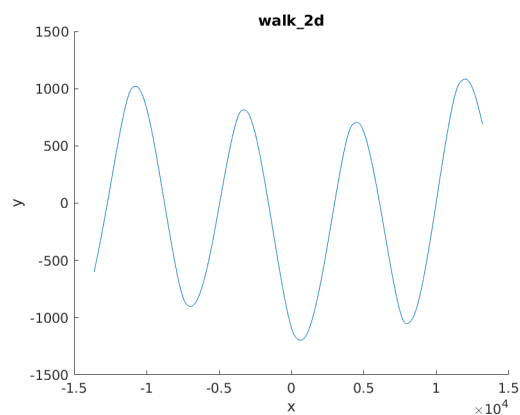
3 Řešení 2. části

3.1

Tabulka s optimálními hodnotami kritéria pro $r \in \{1, 2, 5, 10, 15\}$

r	$\sum_{i=1}^m \ \tilde{a}_i - a_i\ ^2$
1	$4.6166 \cdot 10^8$
2	$1.6925 \cdot 10^8$
5	$1.0453 \cdot 10^7$
10	$1.1982 \cdot 10^6$
15	$2.5626 \cdot 10^5$

3.2



3.3

Pokud se body pohybují pouze po přímce stačí nám pro dosažení nulové aproximační chyby podprostor dimenze 1. Tedy zredukujeme reprezentaci na vektor reálných čísel, která budou udávat offset na přímce, po které se body pohybují.

3.4

Máme počítat optimální chybu aproximace pro různé hodnoty $r \leq n$. Dostaneme n čísel, jaký mají vztah k láním číslům?