CS/ECE/ME 532
Matrix Methods in Machine Learning

Welcome!

Assessment 2

- Thursday, February 29, in class.
  - 70-minutes long
  - Bring your wiscard
  - No calculators
  - **Formula Sheet**
    - You *must* bring one 8.5”x11” sheet of paper to the assessment
    - Write your name on the top right corner of the formula sheet.
    - The formula sheet must be *handwritten* (written by *you*)
    - You may use both sides of the paper.
    - You may use a tablet to compose the formula sheet (you must print the formula sheet)
    - You must turn in your formula sheet with the assessment (you can get it back afterward)
- How to prepare:
  - Revisit past exercises
    - Do the problems without looking the solutions!
  - Unit 2 Practice Problems
  - Attend class on Tuesday, February 27

Activity 10

Low rank decompositions:

taste vectors or patterns

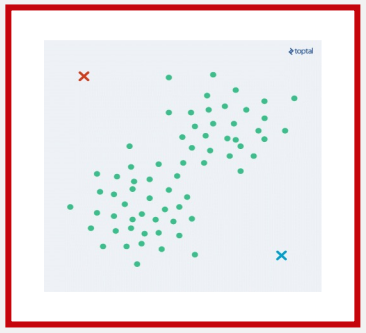
$$\boldsymbol{A} = \begin{bmatrix} 3 & 3 & 3 & -1 & -1 & -1 \\ 1 & 1 & 1 & -3 & -3 & -3 \\ 1 & 1 & 1 & -3 & -3 & -3 \\ 3 & 3 & 3 & -1 & -1 & -1 \end{bmatrix} \approx \boldsymbol{T}\boldsymbol{W}^T = \begin{bmatrix} t_{1,1} & t_{1,2} \\ t_{2,1} & t_{2,2} \\ t_{3,1} & t_{3,2} \\ t_{4,1} & t_{4,2} \end{bmatrix} \begin{bmatrix} w_{1,1} & \dots & w_{1,6} \\ w_{2,1} & \dots & w_{2,6} \end{bmatrix}$$

Before:
taste vectors were given or found using Gram-Schmidt

Unit 3:
find them in a more data-centric way: the SVD (or K-means)

Clustering with K-means:

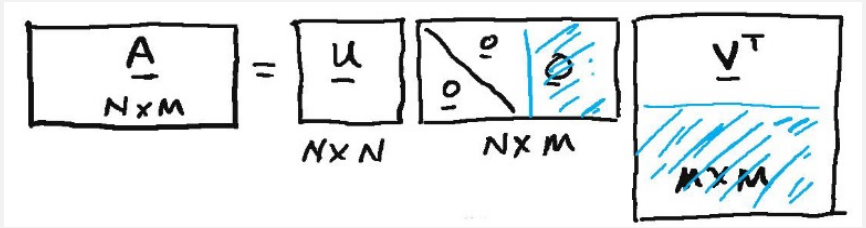
- columns of $\boldsymbol{A}$ are points in $\mathbb{R}^4$
- each column of $\boldsymbol{T}$ is a cluster center
- each column of $\boldsymbol{W}^T$ has a single 1 indicating cluster center



Numpy Tip
Picking columns of a matrix A:

```
cols = [3, 1, 2, 3]
A[:, cols]
```

The Singular Value Decomposition:



- Singular values $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_N \geq 0$
- "Importance" of patterns in **U**, **V** ranked by σ_i
- Optimal low rank approximation $\mathbf{A} \approx \sum_{i=1}^{r} \sigma_i \mathbf{u}_i \mathbf{v}_i^T$