

# CS/ECE/ME 532

## Period 10

- Unit 2 Quiz – Tuesday next week (March 2<sup>nd</sup>), in class
  - 60 minutes
  - Unit 2 (including ridge regression)
  - no interaction with anyone besides instructors
  - must sit at your table, **video must be on**
- Unit 2 Integrative Summary assignment due Tuesday evening
- Unit 2 Practice Problems now open (end of week 5)

## Today – Intro to k-means and the SVD

Low rank decompositions:

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

taste vectors or patterns

*Before:*  
given taste vectors

*Unit 3:*  
find them in a more meaningful way  
the SVD (or k-means)

*K-means:*

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

*The SVD:*

$$\mathbf{A} = \begin{bmatrix} \mathbf{A} \\ N \times M \end{bmatrix} = \begin{bmatrix} \mathbf{U} \\ N \times N \end{bmatrix} \begin{bmatrix} \sigma_1 & & 0 \\ & \sigma_r & \\ 0 & & 0 \end{bmatrix} \begin{bmatrix} \mathbf{V}^T \\ K \times M \end{bmatrix}$$

Python

```
U, s, VT = np.linalg.svd(A, full_matrices=False)
print(U.shape, s.shape, VT.shape)
```

- Singular values  $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_N \geq 0$
- “Importance” of patterns in  $\mathbf{U}, \mathbf{V}$  ranked by  $\sigma_i$
- Optimal low rank approximation

$$\mathbf{A} \approx \sum_{i=1}^r \sigma_i \mathbf{u}_i \mathbf{v}_i^T$$