Lecture 2: Nonnegative Matrix Factorization

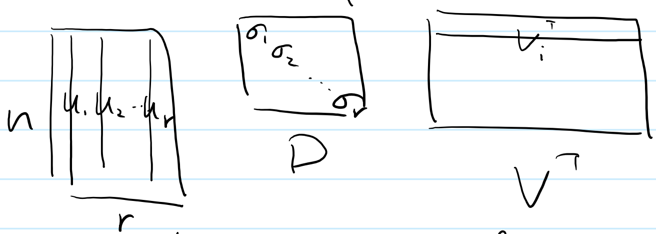
Wednesday, August 31, 2016 9:51 AM

Outline:

- Nonnegative Matrix Factorization
- Geometric Interpretation
- New algorithm for "separable" NMF

- Recall: Singular Value Decomposition (SVD, PCA)

$$M = UDV^T = \sum_{i=1}^{r} \sigma_i u_i v_i^T$$



$\{u_i\}$ orthonormal $\qquad\qquad$ $\{v_i\}$ orthonormal

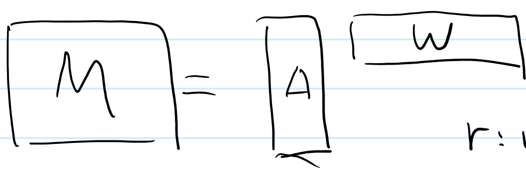
$\|u_i\| = 1, \langle u_i, u_j \rangle = 0$

- Ambiguity $\qquad M = UV^T = (UR)(VR)^T$

  if $RR^T = I$

- nonnegative matrix factorization

$$M = AW = \sum_{i=1}^{r} a_i w_i^T$$



$\forall i,j \quad A_{i,j} \geq 0$

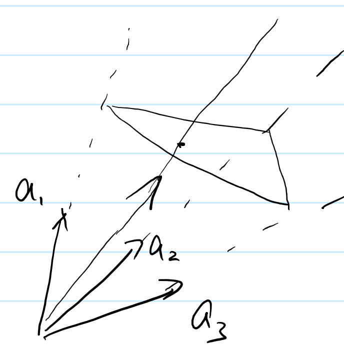
$\qquad\quad W_{i,j} \geq 0$

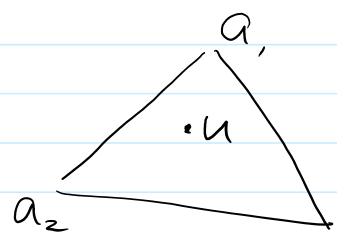
r: nonnegative rank

- Why nonnegative
  - geometric interpretation (Probabilistic int.)

- Def. conical hull: for vectors $v_1, v_2, \cdots, v_r \in \mathbb{R}^n$

  $\text{cone}\{v_1, \cdots, v_r\} = \{u \mid u = \sum_{i=1}^{r} c_i v_i, \ c_i \geq 0\}$

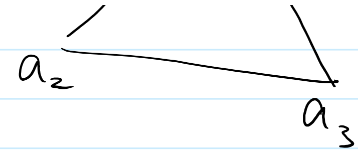




Convex hull $\text{conv}\{v_1, \cdots, v_r\} = \{u \mid u = \sum_{i=1}^{r} c_i v_i, \ c_i \geq 0, \ \sum_{i=1}^{r} c_i = 1\}$

u is a "mixture" of v_i

u is a "mixture" of v_i $\qquad \sum_{i=1}^{r} c_i = 1$



NMF $\Longleftrightarrow$ Given points $m_1, \cdots, m_n \in \mathbb{R}^m$, find nonnegative vectors $a_1, a_2, \cdots, a_r \in \mathbb{R}^m$ such that $\forall i$ $m_i \in \text{cone}\{a_j\}$

- normalization : conical hull $\longrightarrow$ convex hull

  hyperplane $\{x \mid \langle \vec{1}, x \rangle = 1\}$

  $M = AW \qquad m_i = \sum_{j=1}^{r} W_{j,i}\, a_j$

  normalize $M \qquad \overline{m_i} = \dfrac{m_i}{|m_i|_1}$

  normalize $A \qquad \overline{a_i} = \dfrac{a_i}{|a_i|_1}$

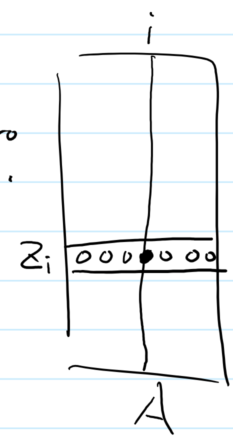
  $$\overline{m_i} = \sum_{j=1}^{r} W'_{j,i}\, \overline{a_j}$$

  $$\langle \vec{1}, \overline{m_i} \rangle = \langle \vec{1}, \sum_{j=1}^{r} W'_{j,i}\, \overline{a_j} \rangle$$

  $$1 = \sum_{j=1}^{r} W'_{j,i}$$

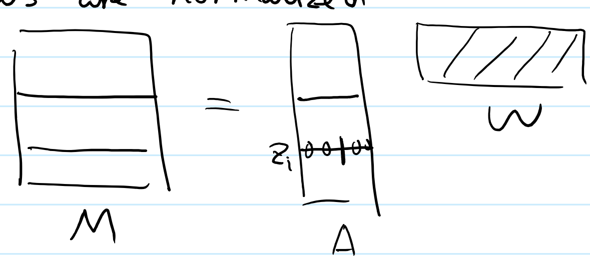
- Algorithm
  - NMF is NP-hard [Vavasis]
  - in practice: alternate between A, W
- Assumption: Separability

  $M = AW$ is separable, if for any $i = 1, 2, \cdots, r$ there is a row z_i such that $A_{z_i,i}$ is the only nonzero entry.



  $$A_{z_i,i} > 0, \quad \forall j \neq i \;\; A_{z_i,j} = 0$$

  if rows are normalized



  $M_{z_i,:}$ is a pure element, equal to $W_{i,:}$

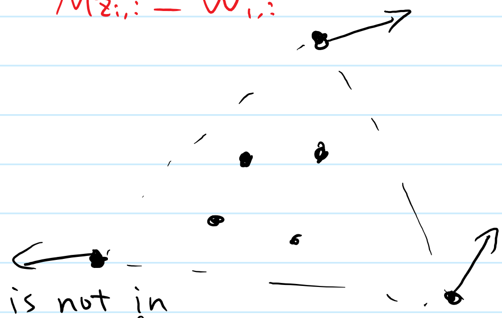
- Geometric interpretation of separable NMF

  Given points $M_{1,:}, M_{2,:}, \cdots, M_{n,:}$

there exists unknown $W_{i,:}$ such that $\forall i\ M_{i,:} \in \mathrm{conv}\{W_{j,:}\}$

there exists unknown z_i such that $M_{z_i,:} = W_{i,:}$

find $W_{i,:}$



alg will try to identify vertices

- Claim: $M_{i,:}$ is vertex $\Longleftrightarrow$ $M_{i,:}$ is not in convex hull of other rows

- Handling Noise if $M \approx AW$

$$\| M_{i,:} - (AW)_{i,:} \| \le \varepsilon$$

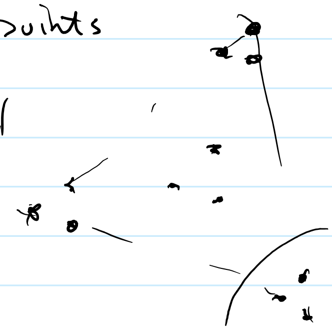
① non-vertex is no longer in convex hull of others

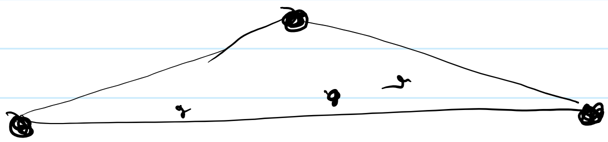
solution: they are still close

② vertex may be close to other points

solution: remove nearby rows

③ vertex may still be close to convex hull





Assumption: the rows of W are α-robust.

Def: $v_1, v_2, \cdots, v_r$ is α-robust, if $\forall i: \|v_i\| \le 1$, and

$\forall i$ v_i is not close to convex hull of others

$\forall u \in \mathrm{conv}\{v_1, v_2, \cdots, v_r\} \setminus \{v_i\}$

$\|u - v_i\| \ge \alpha$

Theorem: If rows of W are α robust, all rows of M are ε-close to AW, can find vertices that are $O\left(\frac{\varepsilon}{\alpha}\right)$ close.