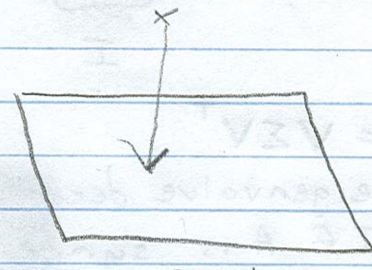
High dimensional data

$$\begin{array}{c c} & \text{25000 words} \\ \begin{array}{c}10^6\\ \text{papers}\end{array} & \begin{pmatrix} \quad & \quad & \quad \\ & & \\ & & \end{pmatrix} \end{array} \xrightarrow[\text{points to lower dimensional space}]{\text{project}}$$



Singular Value Decomposition

$$A = U \Sigma V^T$$

$$\Sigma \begin{pmatrix} \sigma_1 & & & 0 \\ & \sigma_2 & & \\ & & \ddots & \\ 0 & & & \sigma_n \end{pmatrix}$$

σ_1 max $\rightarrow$ $\sigma_1, \ldots \sigma_n$ monotonically decreasing

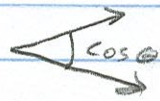
$$\Sigma^{(300)} \begin{pmatrix} \sigma_1 & & & & & & 0 \\ & \sigma_2 & & & & & \\ & & \ddots & & & & \\ & & & \sigma_{300} & & & \\ & & & & 0 & & \\ & & & & & \ddots & \\ 0 & & & & & & 0 \end{pmatrix}$$

$$A^{(300)} = U \Sigma^{(300)} V^T$$

$$\left| A - A^{(300)} \right|_F^2 = \min_{\substack{\text{rank } 300 \\ B\text{'s}}} \left| A - B \right|_F^2$$

$$\begin{array}{c c} & \text{words} \\ \text{papers} & \begin{pmatrix} \rule{3em}{0.4pt} \\ A \\ \quad \end{pmatrix} \end{array} \begin{array}{c c} & \text{papers} \\ \text{words} & \begin{pmatrix} \quad & \quad \\ & \\ & \end{pmatrix} \end{array} = \begin{array}{c c} & \text{papers} \\ \text{papers} & \begin{pmatrix} \text{Similarity} \\ \text{matrix} \end{pmatrix} \end{array}$$

normalize rows



→ elements 0 if no shared words, 1 if same paper.

$$AA^T = U\Sigma \underbrace{V^TV}_{I} \Sigma U^T = U\Sigma^2 U^T$$

$A = V\Sigma V^T$

eigenvalue decomposition

If A is symmetric then eigenvalues are real.

Real spectral theorem: If A is a real symmetric matrix then

1) eigenvalues are real
2) if eigenvalue has multiplicity k then there are k linearly independent eigenvectors associated with it.
3) A is diagonalizable

Proof (1):

$x^c \to$ transpose & conjugate of any complex number

$Ax = \lambda x$

$x^c A x = \lambda x^c x = \lambda$

$\lambda = x^c Ax = (x^c Ax)^{cc} = (x^c Ax)^c = \lambda^c$

Theorem $|A - A^{(k)}|_2 \leq \min_{\substack{B \text{ rank} \\ k \text{ or less}}} |A - B|_2$

Proof: Suppose $\exists B$ such that $|A-B|_2 < |A - A^{(k)}|_2$

Since B is of rank k or less, $\exists$ subspace W of dimension $n-k$ s.t. $\forall w \in W$ $Bw = 0$

For these w, $Aw = (A-B)w$

$|Aw|_2 = |(A-B)w|_2 \leq |A-B|_2 |w|_2 < |A-A^{(k)}|_2 |w|_2$

$$\begin{pmatrix} \sigma_1 & & & & & & \\ & \sigma_2 & & & & & \\ & & \ddots & & & & \\ & & & \sigma_k & & & \\ & & & & \sigma_{k+1} & & \\ & & & & & \ddots & \\ & & & & & & \sigma_n \end{pmatrix} \quad \sigma_1, \sigma_2 \ldots \sigma_k \to 0$$

$$\longrightarrow \quad |A - A^{(k)}|_2 \, |w|_2 = \sigma_{k+1} |w|_2$$

$$\boxed{\text{Hence } |Aw|_2 < \sigma_{k+1} |w|_2}$$

For the $k+1$ dimensional space spanned by first $k+1$ eigenvectors for any x from space

$$|Ax|_2 \geq \sigma_{k+1} \qquad \text{(will come back to this)}$$

$$x < \begin{pmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_{k+1} \\ | & | & & | \end{pmatrix}$$

$$x = c_1 v_1 + c_2 v_2 + \ldots$$

$$Ax = c_1 \sigma_1 v_1^T v_1 + \ldots$$

<u>Contradiction</u>: For vector in intersection of the two spaces we have a contradiction.

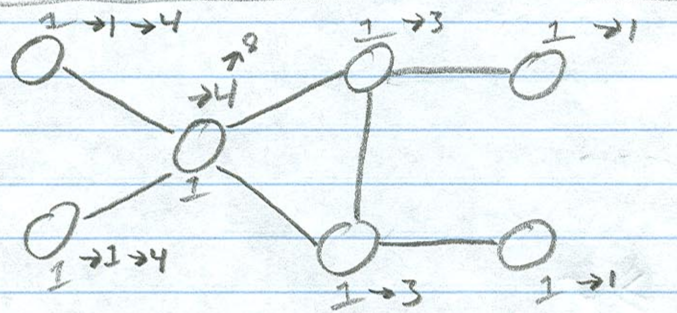
Let $v_1, v_2, \ldots, v_{k+1}$ be first $k+1$ eigen vectors.
Let $x = \sum c_i v_i$

$$|Ax|_2 = |V \Sigma V^T x|_2 = \left| V \Sigma \begin{pmatrix} c_1 \\ c_2 \\ c_{k+1} \\ 0 \\ \vdots \\ 0 \end{pmatrix} \right|_2$$

$$\boxed{\lambda \equiv \sigma}$$

$$= \left\| V \begin{pmatrix} \lambda_1 c_1 \\ \lambda_2 c_2 \\ \vdots \\ \lambda_{k+1} c_{k+1} \\ 0 \\ \vdots \\ 0 \end{pmatrix} \right\|_2 = \left\| \begin{matrix} | & | & & | \\ \lambda_1 c_1 v_1 & \lambda_2 c_2 v_2 & \cdots & \lambda_{k+1} c_{k+1} v_{k+1} \\ | & | & & | \end{matrix} \right\|_2 \geq \lambda_{k+2}$$

$$\left(\sum_{i=1}^{k+1} \lambda_i c_i v_i^T\right)\left(\sum_{i=1}^{k+1} \lambda_i c_i v_i\right) = \sum_{i=1}^{k+1} \lambda_i^2 c_i^2$$



$\to \infty$

normalize each iteration

importance

$$Ax = \lambda x$$