A (first?) foray into Linear Algebra

Day 1 – Fundamentals of vectors, matrices and their operations Motivated by searching and ranking

vector vector multiplication (inner product), cosine rule, vector norms
 matrix matrix-vector multiplication, matrix-vector systems, matrix-matrix multiplication, matrices as operators, orthogonal matrices, inverses and determinants, and a bit about eigenvalues and eigenvectors

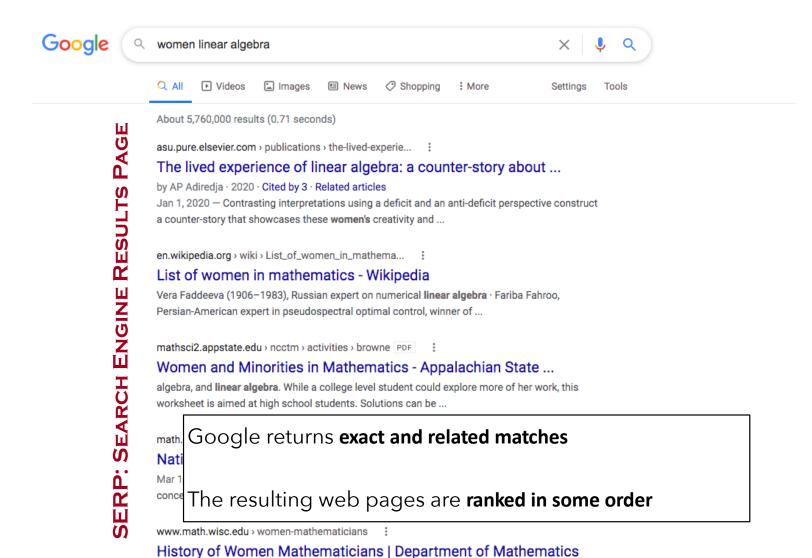
Day 2 – Diving deeper with as ultimate goal the SVD and PCA With compression and regression

matrix Singular Value Decomposition, least squares, normal equations, QR decomposition

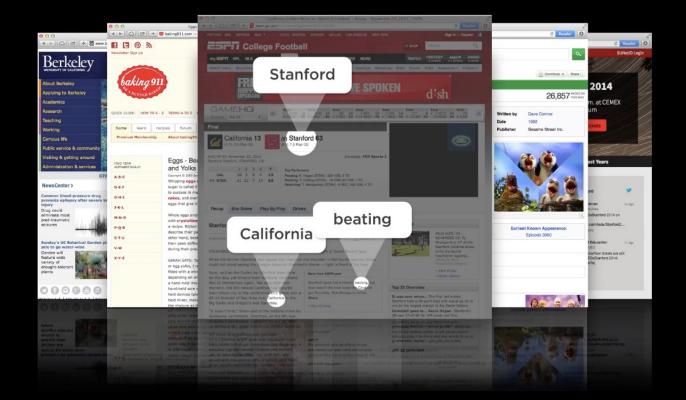
Approximate Schedule

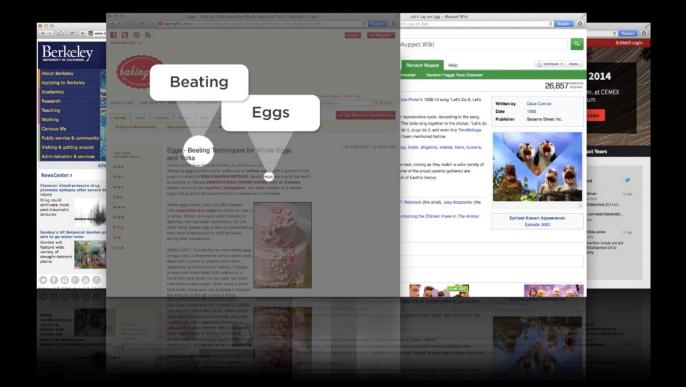
Day 1		Day 2	
8-8:15	Intros, breakouts	8-9	New material with Margot
8:15-9	New material with Margot	9-9:30	Practice with Nadim
9-9:30	Practice with Nadim	9:30-9:45	Break
9:30-9:45	Break	9:45-10:30	New material with Nadim
9:45-10:30	New material with Margot	10:30-10:50	Practice with Nadim
10:30-11	Practice with Nadim	10:50-11	Closing

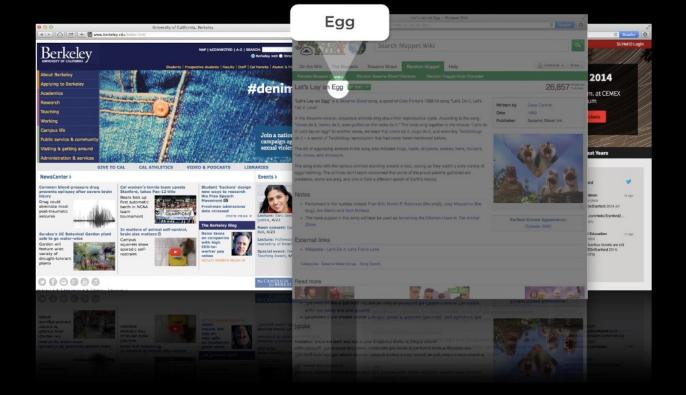
Slides after this are used throughout the short course

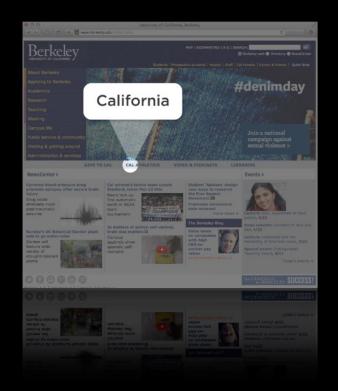


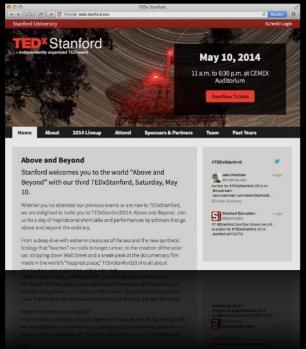
1886: The first American Woman to receive a PhD. in math (from Columbia ... The title of Dr.

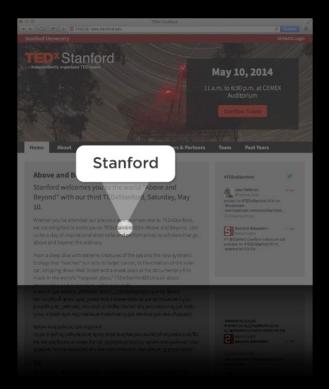










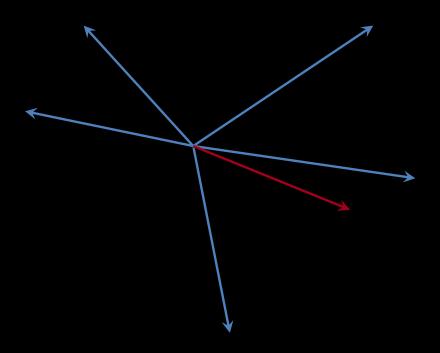


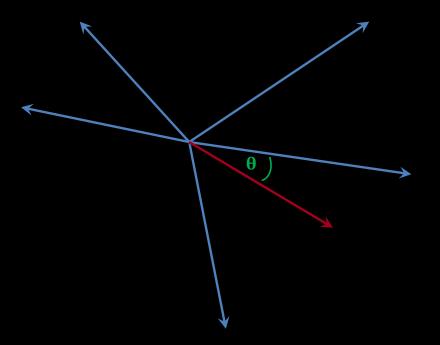
website words	Page 1	Pagez	Page 3	Page 4	Page 5
stanford	1	1	0	0	0
Beating	1	0	0	1	0
CAL	1	0	1	0	0
Eggs	0	0	0	1	1

How to search for the phrase "Stanford eggs"

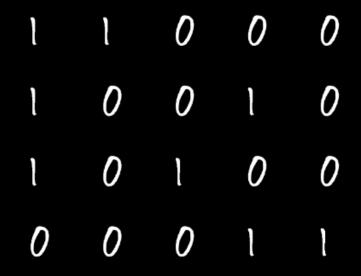
website words	Page 1	Page 2	Pages	Page 4	Page 5	
Stanford	1	1	0	0	0	1
Beating	1	0	0	1	0	0
CAL	1	0	1	0	0	0
Eggs	0	0	0	1	1	1

How would you determine which of the page vectors are closest?

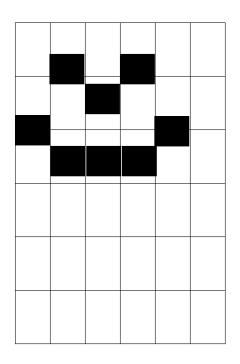




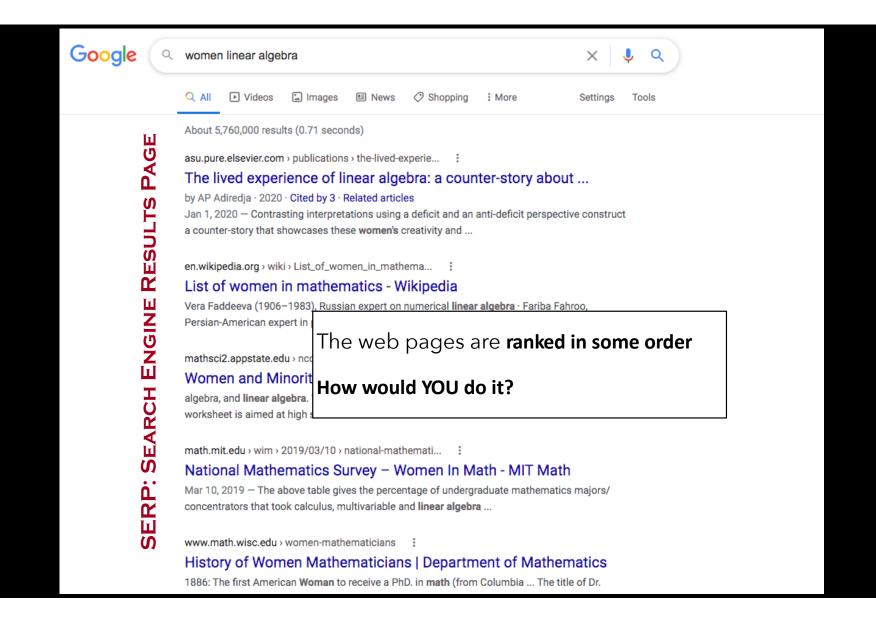
website words	Page 1	Page 2	Page 3	Page 4	Page 5
Stanford	1	1	0	0	0
Beating	1	0	0	1	0
CAL	1	0	1	0	0
Eggs	0	0	0	1	1



$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix}$$



0	0	0	0	0	0
0	1	0	1	0	0
0 0 0 1	0	1	0	0	0
		$\mathbf{\cap}$	\cap	1	
⁻ 1	U	U	U	Τ	U
				0	



RANKING DEPENDS ON MANY FACTORS

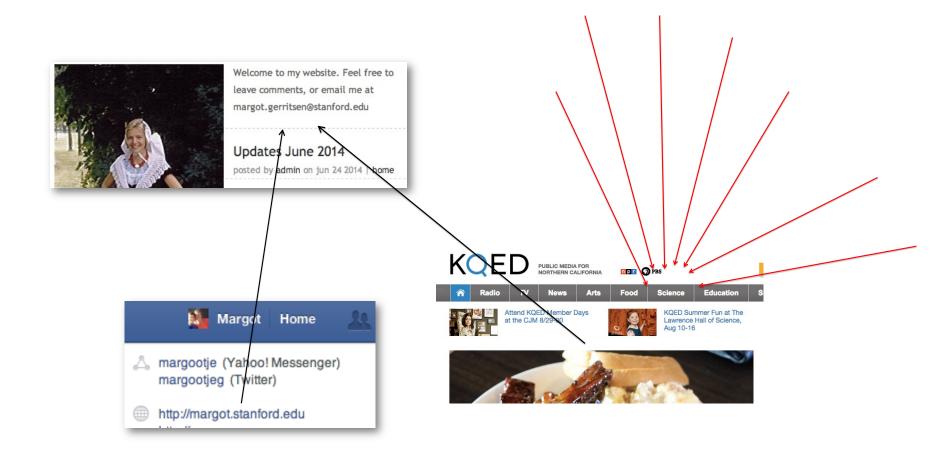
DIFFERENT SEARCH ENGINES USE DIFFERENT FACTOR/WEIGHTINGS

- The traditional page rank
 (measure of importance of a page)
- Number of visits
- Age
- Recent edits

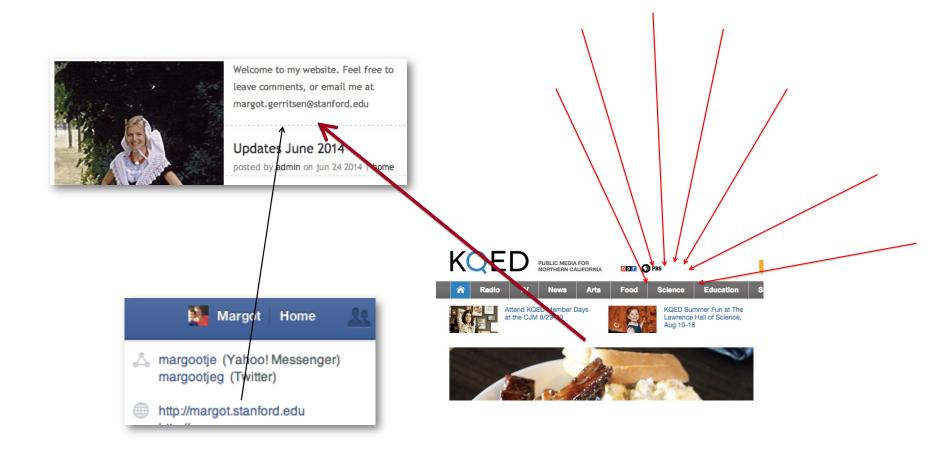
• • • • • • • • • • •

PAGE RANKING A LA 1990S

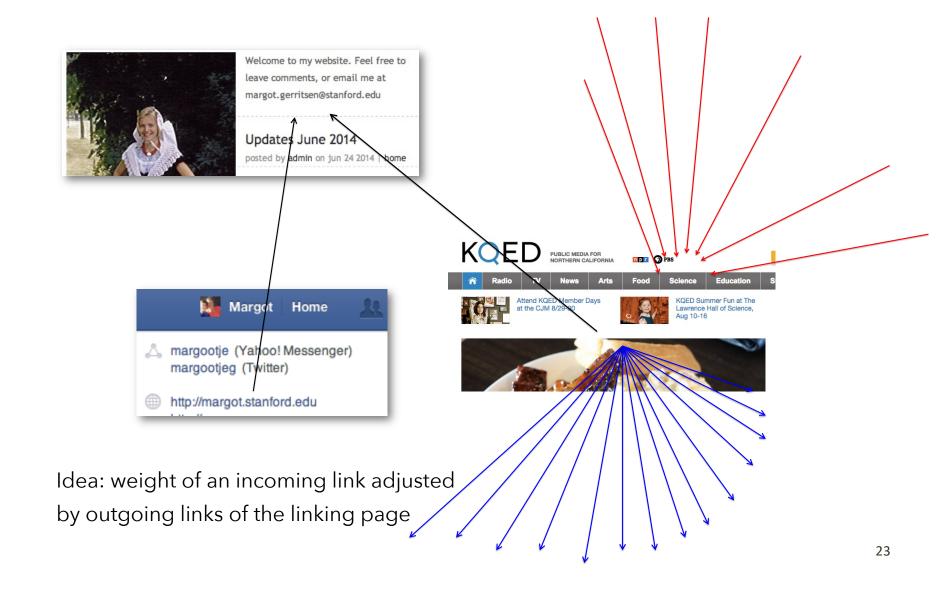
Lawrence Page, Sergey Brin, Rajeev Motwani, Terry Winograd "The PageRank Citation Ranking: Bringing Order to the Web" Technical Report, Stanford InfoLab, 1999



Idea: The more incoming links, the higher the page rank



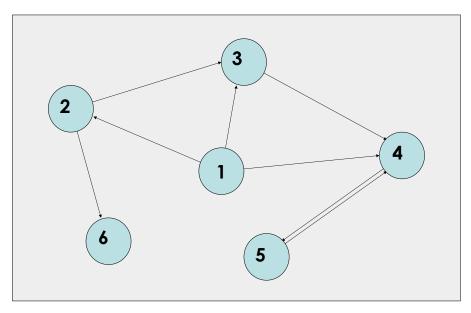
Idea: links coming in from an important page have more weight



RANK PROPAGATION

Rank of page j denoted by x_j

$$\mathbf{x}_1 = \mathbf{0}$$



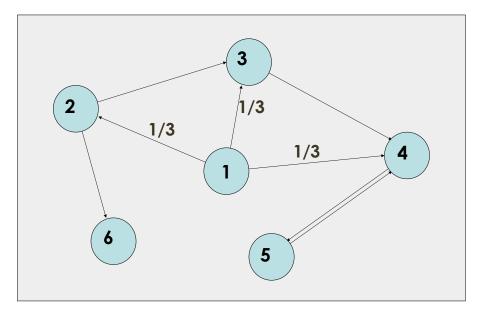
A very small internet "graph"

RANK PROPAGATION

Rank of page j denoted by x_j

$$x_1 = 0$$

$$x_2 = \frac{1}{3}x_1$$



A very small internet "graph"

RANK PROPAGATION

Rank of page j denoted by x_j

$$x_{1} = 0$$

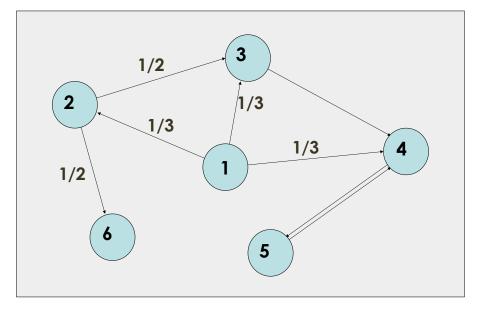
$$x_{2} = \frac{1}{3}x_{1}$$

$$x_{3} = \frac{1}{3}x_{1} + \frac{1}{2}x_{2}$$

$$x_{4} = \frac{1}{3}x_{1} + x_{3} + x_{5}$$

$$x_{5} = x_{4}$$

$$x_{6} = \frac{1}{2}x_{2}$$



A very small internet "graph"

$$\mathbf{x}_1 = \mathbf{0}$$

A COUPLED AND LINEAR

SYSTEM OF EQUATIONS

$$\mathbf{x}_2 = \frac{1}{3} \, \mathbf{x}_1$$

$$x_3 = \frac{1}{3}x_1 + \frac{1}{2}x_2$$

$$X_4 = \frac{1}{3}X_1 + X_3 + X_5$$

$$X_5 = X_4$$

$$X_6 = \frac{1}{2}X_2$$

IN MATRIX-VECTOR NOTATION

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1/3 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1/3 & 1/2 & 0 & 0 & 0 & 0 & 0 \\ 1/3 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix} \quad \text{or} \quad \mathbf{P}\mathbf{x} = \mathbf{x}$$

SVD FOR COMPRESSION ON MORE THAN A 5X5 MATRIX









256x256

keep 2 terms

keep 8 terms

keep 32 terms