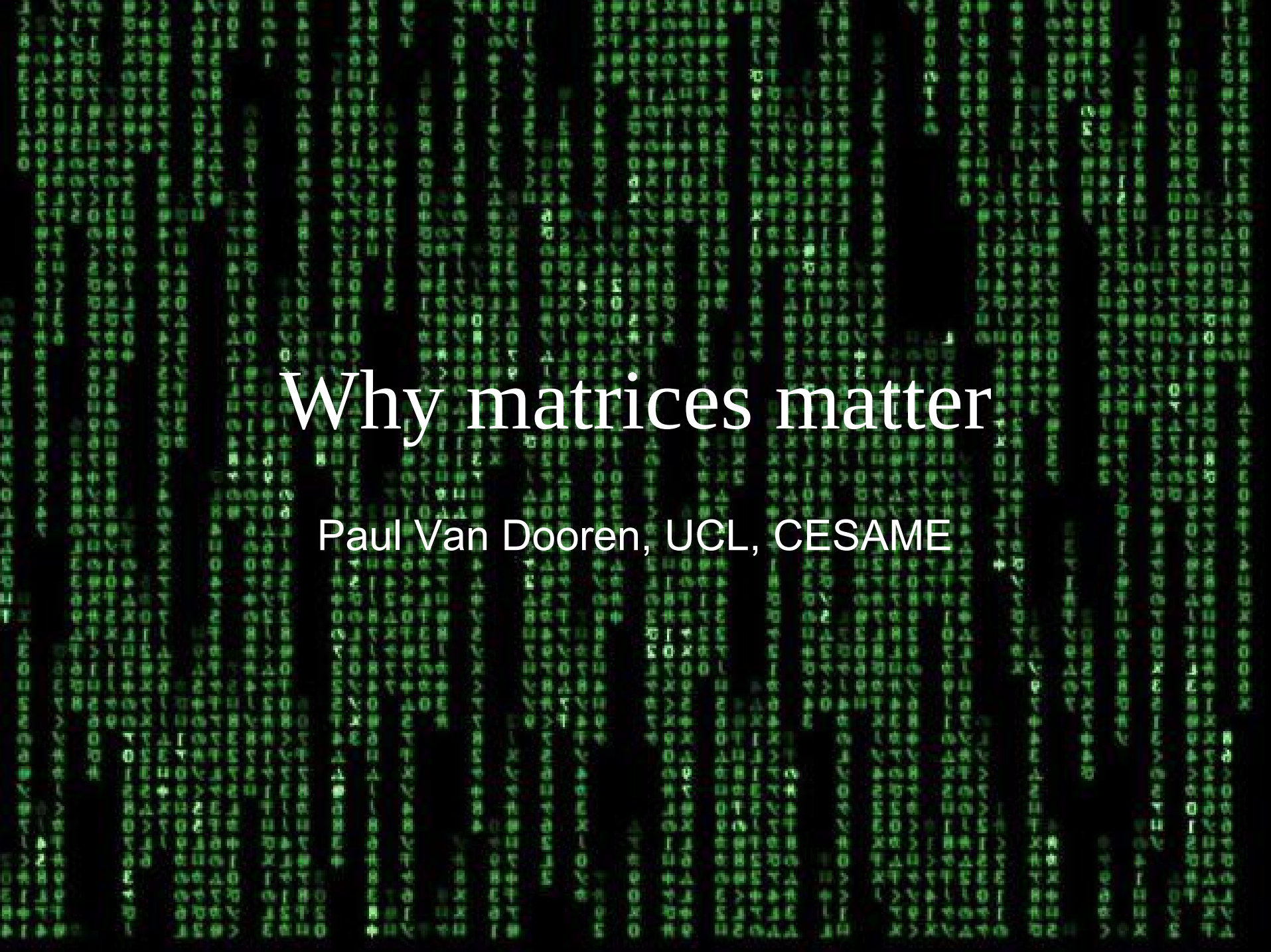


Why matrices matter

Paul Van Dooren, UCL, CESAME

Where are matrices coming from ?

ma·trix (mā'trĭks) *n.*, *pl.*, ma·tri·ces (mā'trĭ-sēz')

- …
- *Anatomy.* The womb (uterus).
- *...*
- *Geology*. The solid matter in which a crystal is embedded.
- ...
- *Metallurgy.* The principal metal in an alloy, as the iron in steel.
- …
- *Printing*. A mold used in stereotyping. A metal plate used for casting typefaces.
- *…*
- *Computer Science*. The network of intersections in a computer.
- *…*
- *Mathematics*. A rectangular table of numeric or algebraic quantities subject to mathematical operations, with a formation of elements into columns and rows.
- …

[Middle English matrice, from Old French mātrīc-, from Latin, māter]

The early days …

Solving a system of equations by is often credited to Carl Friedrich Gauss (Gaussian elimination)

$$3x+2y+1z=39$$
$$2x+3y+1z=34$$
$$1x+2y+3z=26$$

$$3x+2y+1z=39$$
$$5y+1z=24$$
$$4y+8z=39$$

$$3x+2y+1z=39$$
$$5y+1z=24$$
$$36z=99$$

Gauss 1777-1855

Gauss worked on many other things :
least squares, statistics, heavenly bodies …

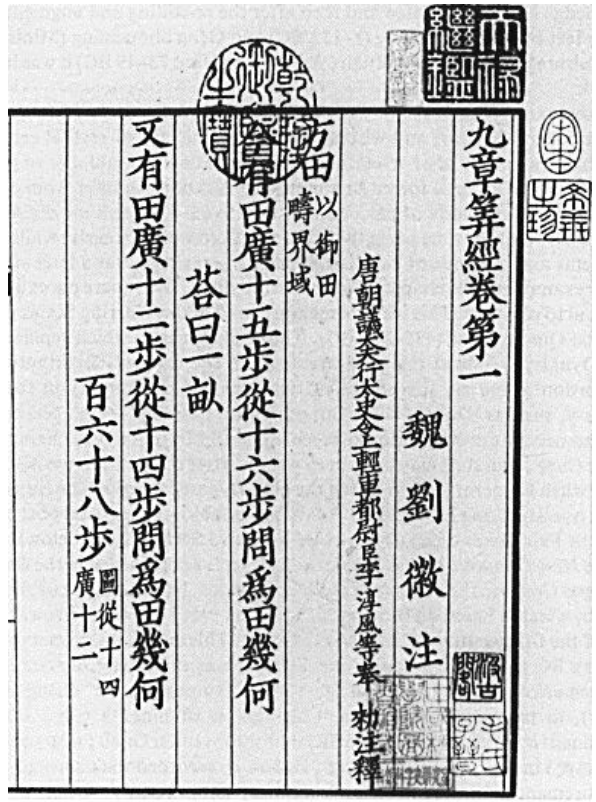
The early days …

But this was also known in 202 BC in
The Nine Chapters on the Mathematical Art

$$
\begin{aligned}
36z &= 99\\
1z+5y &= 24\\
1z+2y+3x &= 39
\end{aligned}
$$

$$
\begin{aligned}
8z+4y &= 39\\
1z+5y &= 24\\
1z+2y+3x &= 39
\end{aligned}
$$

$$
\begin{aligned}
3z+2y+1x &= 26\\
1z+3y+2x &= 34\\
1z+2y+3x &= 39
\end{aligned}
$$



This book also talked about taxation, square roots, areas, volumes and the Pythagorean theorem …

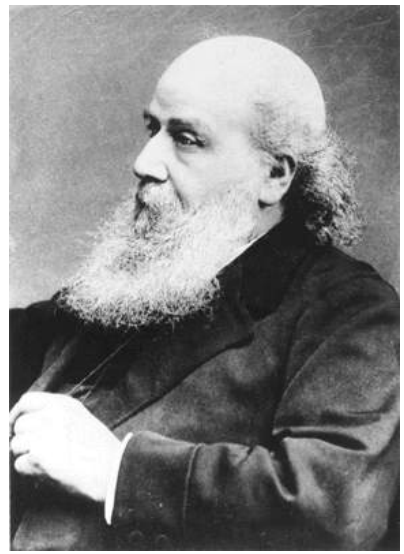
Positive quadratic forms

When is the form $f(x_1, ..., x_n) = \sum_i \sum_j s_{ij} x_i x_j$ >0 ?

J.J.Sylvester introduced matrices to rewrite this as

$$f(\mathbf{x}) = \mathbf{x}^T S \mathbf{x},$$

$$S = \begin{bmatrix} s_{11} & \dots & s_{1n} \\ \vdots & & \vdots \\ s_{n1} & \dots & s_{nn} \end{bmatrix}, \ \mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$



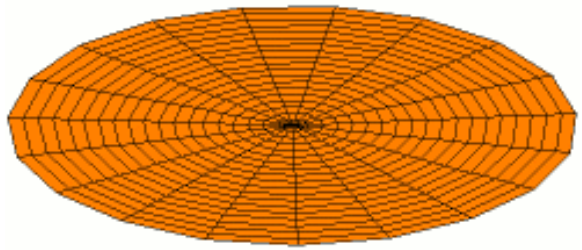
Sylvester 1814 -1897

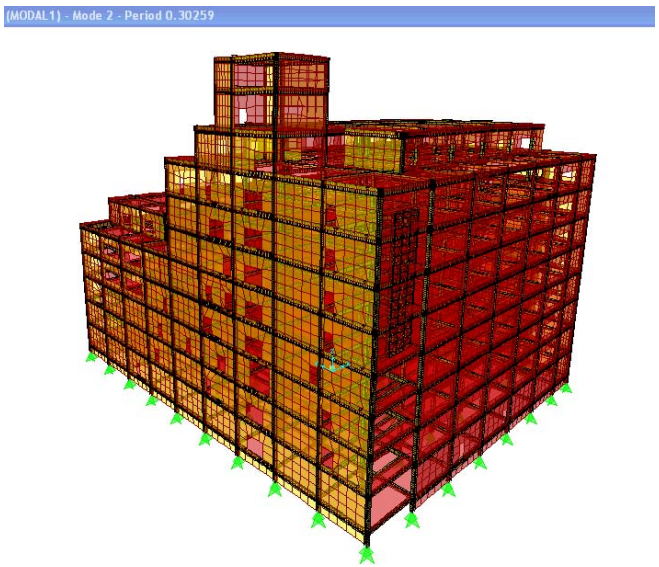
This also led to eigenvectors and eigenvalues

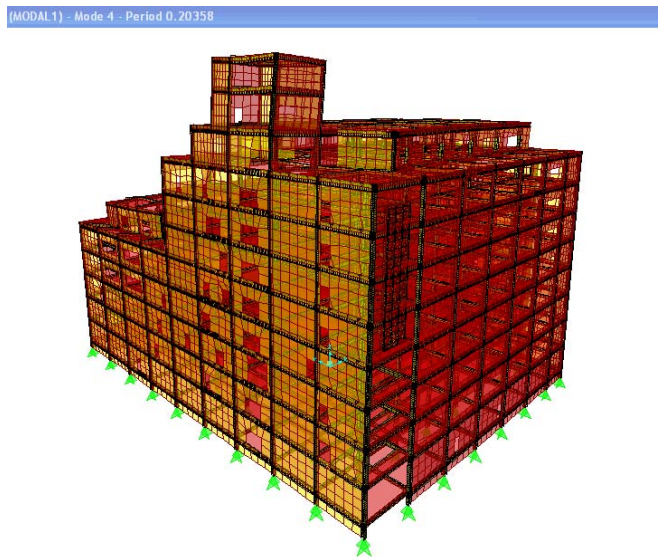
If $S\mathbf{x} = \lambda \mathbf{x}$ then $f(\mathbf{x}) = \lambda \mathbf{x}^T \mathbf{x} > 0 \Leftrightarrow \lambda > 0$

More eigenvectors and eigenvalues…

$S\mathbf{x} = \lambda\mathbf{x}$ shows up in vibrations

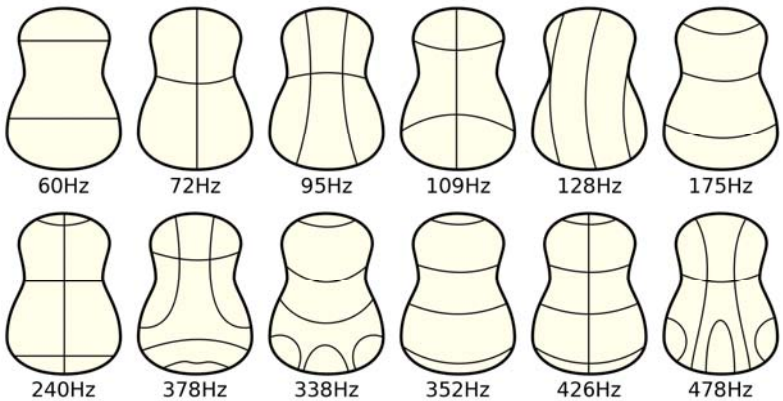






Eigenvectors “avant-la-lettre”

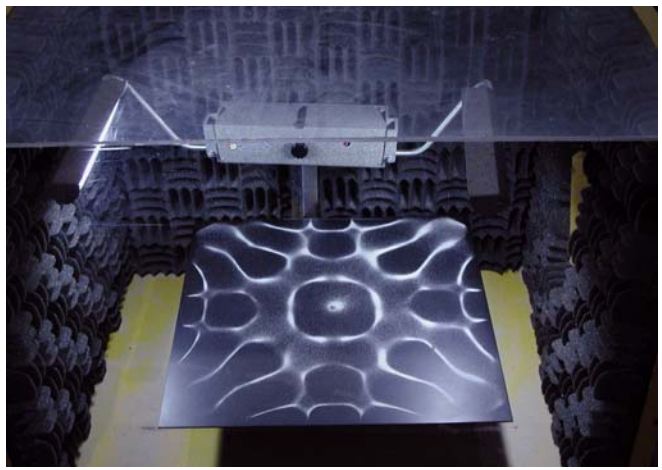
Ernest Chladni studied acoustics (e.g. of guitars)



Chladni called them nodal lines

Chladni 1756 -1827

How did he (and Hooke) observe these ?

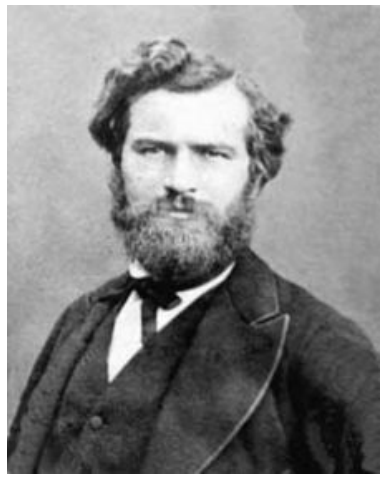


These are the nodal regions of the eigenvectors

Camille Jordan

C. Jordan gives the complete algebraic theory of the standard eigenvalue problem (also Kronecker)

$$AX = X \begin{bmatrix} \alpha & 1 & & & \\ & \alpha & 1 & & \\ & & \alpha & & \\ & & & \beta & 1 \\ & & & & \beta \end{bmatrix}$$



Jordan 1838-1922

This quickly led to matrix algebra (Cayley) and matrix equations (Sylvester) such as

$$AX - XB = C$$

$$AX^2 + BX + C = 0$$

and also to functions of matrices : $X(t) = exp^{At}$

All are very important for the study of dynamical systems

Where systems theory comes in

Rudolf Kalman and state-space models (1960)

$$\begin{aligned} x_{k+1} &= Ax_k + Bu_k \\ y_k &= Cx_k + Du_k \end{aligned}$$

$$\begin{aligned} \frac{d}{dt}x(t) &= Ax(t) + Bu(t) \\ y(t) &= Cx(t) + Du(t) \end{aligned}$$

Kalman 1930 -

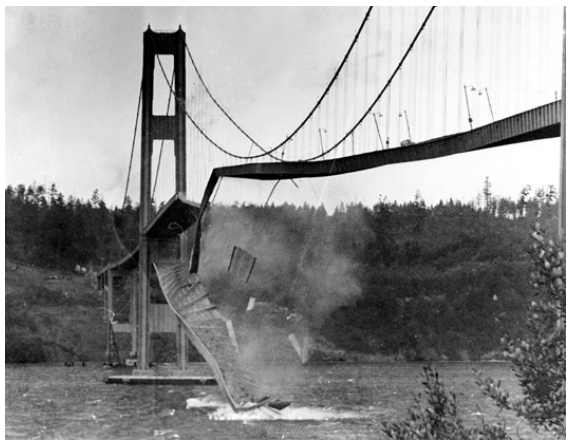
These models had a immense impact in several disciplines
- aeronautical, electrical and mechanical engineering
- signal processing, acoustics, statistics

A wealth of (old and new) matrix theoretic results and applications
- new canonical forms for 2, 4 and several matrices
- linear and quadratic matrix equations and inequalities, …

Special matrix equations

In the context of linear control systems

$$AX + XA^T + BB^T = 0 \quad \text{(stability)}$$



Controllable system

$$\dot{x}(t) = Ax(t) + Bu(t)$$

is stable iff $X \succ 0$

$$AX + XA^T - XRX + Q = 0$$

Optimal control of the above dynamical system

Aleksandr Lyapunov

1857-1918

Count Riccati

1676-1754

Structured matrix problems

Hankel matrices

$$H = \begin{bmatrix} a & b & c & d & e \\ b & c & d & e & f \\ c & d & e & f & g \\ d & e & f & g & h \\ e & f & g & h & i \end{bmatrix}$$

Hankel 1839-1873

Realization
Model reduction
(approximation)
Input-output map
(convolution)

Toeplitz matrices

$$T = \begin{bmatrix} a & b & c & d & e \\ f & a & b & c & d \\ g & f & a & b & c \\ h & g & f & a & b \\ i & h & g & f & a \end{bmatrix}$$

Toeplitz 1881-1940

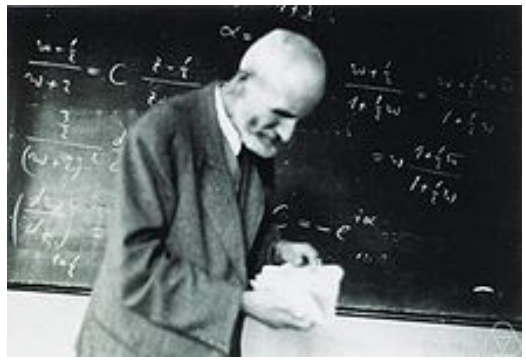
Realization
Correlation
Spectral density
and factorization

Strong links with
function theory

Non-negative matrices

Perron-Frobenius theorem

Matrices with positive entries
have an eigenvalue-eigenvector pair
with positive entries as well



Oscar Perron 1880-1975

$$M > 0 \quad \Rightarrow \quad \exists x > 0, \rho > 0, Mx = \rho x$$

F.G. Frobenius 1849-1917

Applications
- Markov processes, statistics (probabilities)
- Graphs and networks (links or edges)
- Datamining (occurrences of items)
- Google (reputation of a webpage)
- …

Voting matrices

f ?

	Smith	Popov
	o1	o2
$f_1 = 1$	4.2	3.4
$f_2 = 1$	4.5	3.3
$f_3 = 1$	2.8	4.9

$= \boldsymbol{X}$ (votes)

r ?

$$r = \boldsymbol{X}^T \frac{f}{\|f\|_1}$$

$r_1 = 3.8$ $r_2 = 3.9$

Voting matrices

$$f_i = d - \sum_{j=1}^{m} (X_{ij} - r_j)^2$$

Belief divergence
= Variance

f ?

	Smith (o1)	Popov (o2)
$f_1 = 4.6$	4.2	3.4
$f_2 = 4.2$	4.5	3.3
$f_3 = 3$	2.8	4.9

= ***X*** (votes)

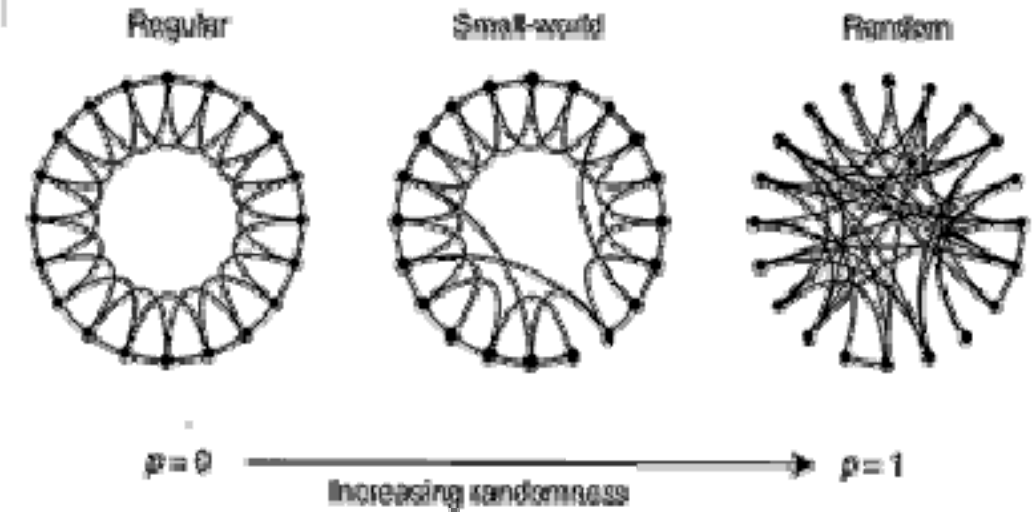
r ?

$$r = X^T \frac{f}{\|f\|_1}$$

$r_1 = 3.8$ $r_2 = 3.9$

Social networks

World with n citizens each with k “friends” (say n=20, k=4)



In a regular world most of my friends know each other

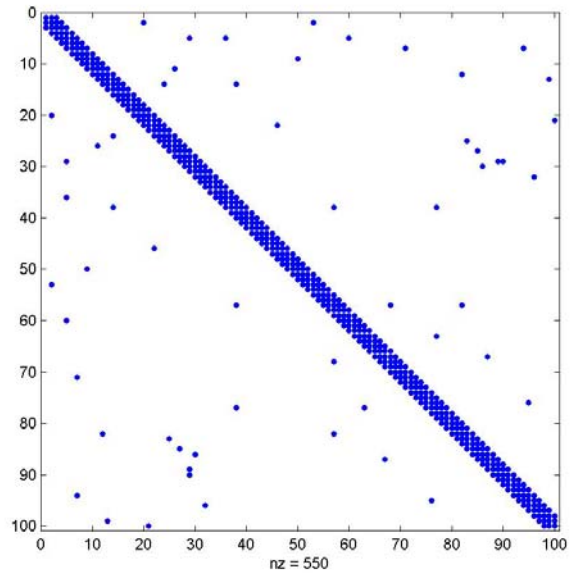
In a random world the “diameter” is small

Adjacency matrix A

Random connections with probability *p* quickly decrease diameter

Adjacency matrix A

(i,j)=1 ↔ *i and j* are "friends"

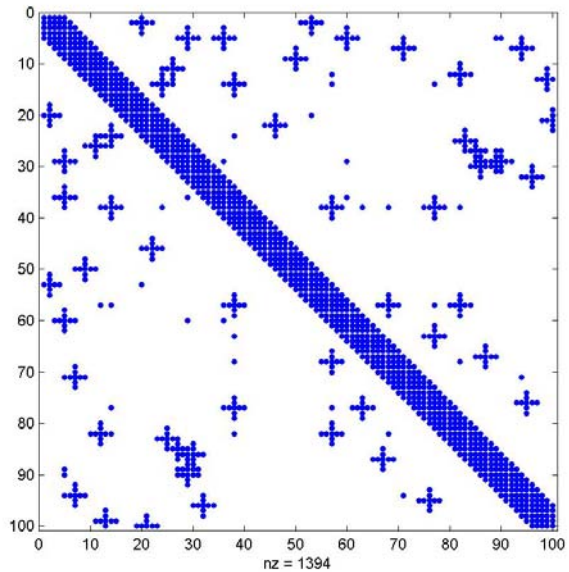


Powers of adjacency matrix

The k-th power of the adjacency matrix give the paths of length k

Adjacency matrix A^2

(i,j)=1 ↔ *i and j* are
"friends of my friends"

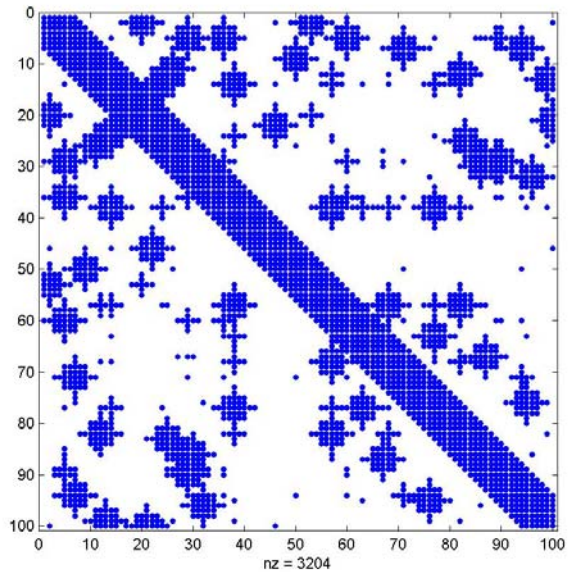


Powers of adjacency matrix

The k-th power of the adjacency matrix give the paths of length k

Adjacency matrix A^3

(i,j)=1 ↔ *i and j* are
"friends of friends of my friends"

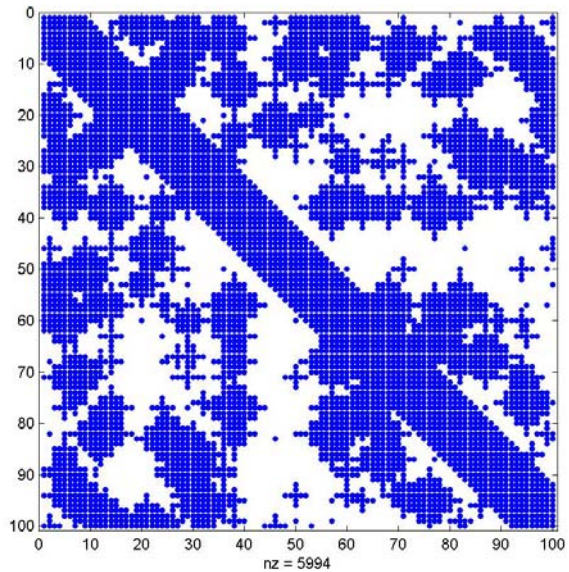


Powers of adjacency matrix

The k-th power of the adjacency matrix give the paths of length k

Adjacency matrix A^4

(i,j)=1 ↔ *i and j* are
"friends …. of my friends"

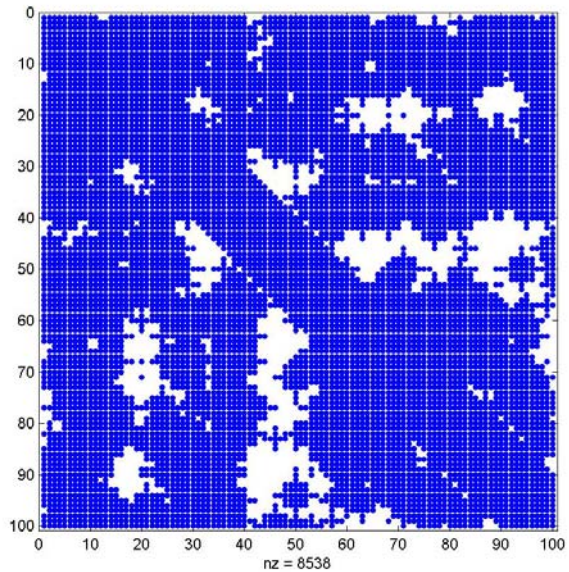


Powers of adjacency matrix

The k-th power of the adjacency matrix give the paths of length k

Adjacency matrix A^5

(i,j)=1 ↔ *i and j* are
"friends …. of my friends"

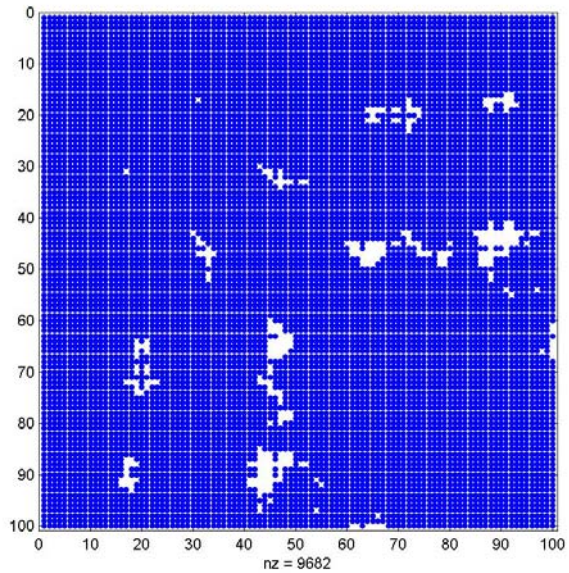


Powers of adjacency matrix

The k-th power of the adjacency matrix give the paths of length k

Adjacency matrix A^6

(i,j)=1 ↔ i and j are
"friends …. of my friends"

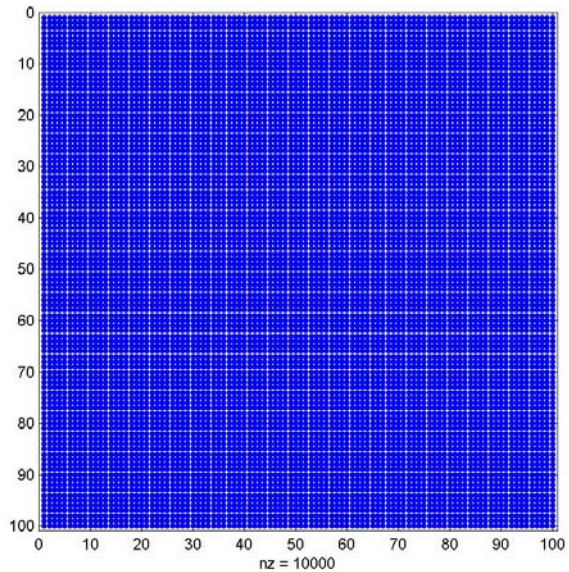


Powers of adjacency matrix

The k-th power of the adjacency matrix give the paths of length k

Adjacency matrix A^7

(i,j)=1 ↔ i and j are
"friends …. of my friends"



“Six degrees” of separation of a small word

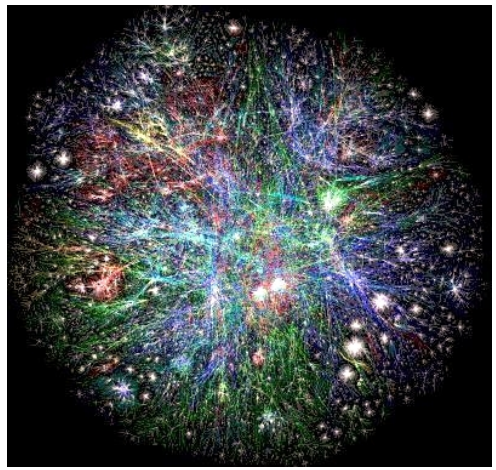
Movie world {co-actors}
Kevin Bacon distance

Bacon number	# of people
0	1
1	991
2	56479
3	82792
4	12183
5	1337
6	308
7	42
8	8

Scientific publications {coauthors}
Erdös number

Google matrix

The web carries tens of billions of pages
organized by tens of thousands of keywords



Each query selects relevant pages
from a low rank matrix approximation

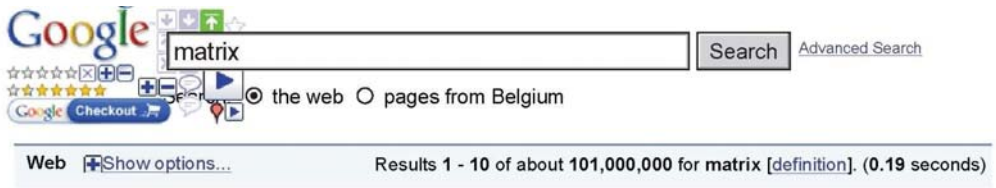
Selected pages are ordered according
to PageRank (computed for all pages)

PageRank x is an eigenvector of an NxN stochastic matrix ($N \sim 10^{10}$)
Largest eigenvector ever computed ?
Has to be adapted every month or so …

Is the steady state probability of a random walker being on a webpage
if he hops around from one webpage to another with equal probability

Perron-Frobenius vector of an appropriately defined Google matrix

Google search on MATRIX



Mathematics comes only in second place !

Next four lectures will do matrix reloading on

- Distance problems, spectra and pseudospectra
- Model reduction of linear systems
- Dominant feature extraction and structured matrices
- Networks and graphs