Lecture 21

NONNEGATIVE MATRICES

(and applications to PageRank)

* PSET 3 is out! $\boxed{\text{Due 10/30 9PM.}}$

* Checkups (revamped!)

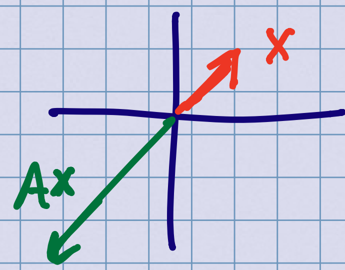
* Midterm Re-dos (signup sheet)

* OFFICE HOURS (P.) Wed 10/28 11AM

Eigenvalues / Eigenvectors

$A \in \mathbb{R}^{n\times n}$ (square)

$$\overset{n\times n}{\underset{\text{matrix}}{A}}\;\overset{n\times 1}{x} = \overset{1\times 1}{\underset{\text{scalar (eigenvalue)}}{\lambda}}\;\overset{n\times 1}{x}$$



DIAGONALIZATION: (sometimes)

$$A = \underset{\text{eigenvectors}}{T}\,\overset{\text{eigenvalues}}{D}\,T^{-1}$$

Symmetric Matrices

$$A = A^T$$

- All eigenvalues are real
- Full basis of orthogonal eigenvectors

$$\Rightarrow \quad A = \overset{\text{eigenvectors}}{T}\,\underset{\text{eigenvalues}}{D}\,T^T$$

(always diagonalizable)

T is orthogonal ($TT^T = T^TT = I$)

NON NEGATIVE MATRICES

A square matrix where all entries are non-negative numbers

$$\boxed{A_{ij} \geq 0 \qquad \forall ij}$$

EX: $\begin{bmatrix} 1 & 3 & 2 \\ 0 & 4 & 3 \\ 5 & 7 & 1 \end{bmatrix}$ ✓ $\begin{bmatrix} 0 & 1 \\ 10 & 2 \end{bmatrix}$ ✓ $\begin{bmatrix} 2 & -3 & 4 \\ 5 & 1 & 2 \\ 0 & 3 & -1 \end{bmatrix}$ ✗

IF A, B ARE NONNEGATIVE MATRICES, ARE THESE ALSO NONNEGATIVE?

- $A + B$ ✓ $a_{ij} + b_{ij}$
- $A \cdot B$ ✓ $\sum_{k=1}^{n} a_{ik} b_{kj}$
- Ax ✓ (x is a nonneg. vector)
- A^{-1} ✗

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

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

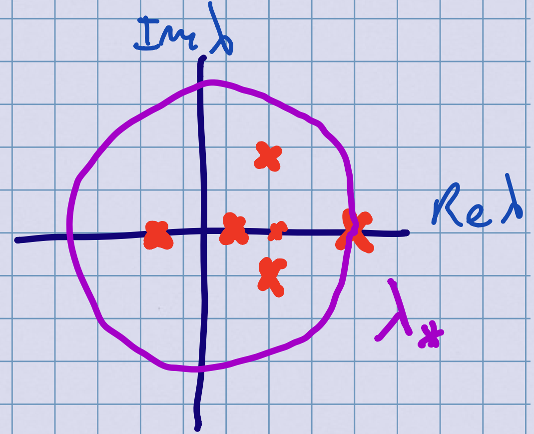
WHAT ABOUT EIGENVALUES / EIGENVECTORS?

SOMETHING VERY INTERESTING HAPPENS!

ASSUME $A_{ij} > 0$ (strictly positive).

THEN, THERE EXISTS A REAL, POSITIVE EIGENVALUE $\lambda_* > 0$ SUCH THAT:

- λ^* HAS MULTIPLICITY 1 (SINGLE EIGENVALUE).



- ALL OTHER EIGENVALUES HAVE SMALLER ABSOLUTE VALUE $|\lambda_i| < \lambda_*$
- THE EIGENVECTOR V_* ASSOCIATED w/ λ_* IS ALSO POSITIVE $(V_*)_i > 0$

EXAMPLE:

(Lecture 17)

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix}$$ all positive entries

$$\begin{cases} \lambda_1 = 4 \\ \lambda_2 = -1 \end{cases}$$

(SANITY CHECK!

TRACE? DETERMINANT?)

Eigenvalue $\lambda_1 = 4$ is positive

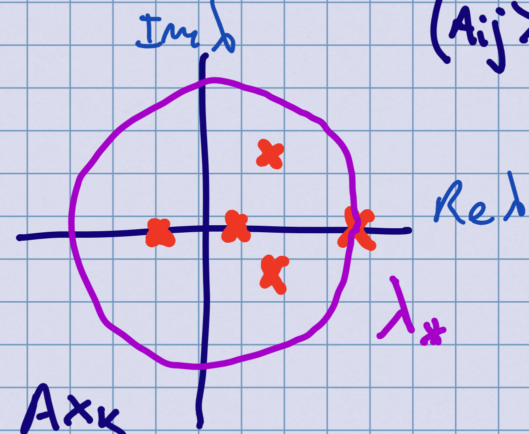
Eigenvector $\begin{pmatrix} 2 \\ 3 \end{pmatrix}$ has positive entries

(What if I use Julia, and it gives me the eigenvector $\begin{pmatrix} -4 \\ -6 \end{pmatrix}$?)

Why is this true?

Eigenvalues of A

$(A_{ij} > 0)$



- Recall the power method

$$x_{k+1} = \text{normalize}\left(A x_k\right) = \frac{A x_k}{\|A x_k\|}$$

For a "random" initial vector x_0

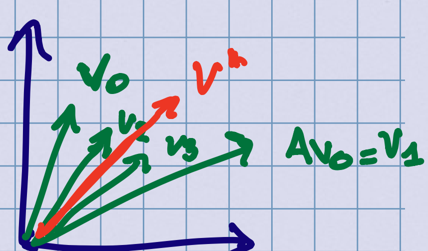
$\Rightarrow$ Notice that if $x_0 > 0$ (componentwise),

then $x_k > 0$ for all k.

$A_{ij} > 0$

$\Rightarrow$ x_k converges to eigenvector $v_* > 0$

$\Rightarrow$ The sequence $x_k^T A x_k \to \lambda_* > 0$



MARKOV MATRICES

(also "STOCHASTIC" MATRICES).

A nonnegative matrix (as before), with the additional property that the sum of each row (or each column) is equal to 1.

"row stochastic" or "column stochastic".

- Why do we care?

Because then (in either case) the matrix has $\lambda = 1$ as an eigenvalue, with the corresponding eigenvector nonnegative.

WHY?

ROW STOCHASTIC $\iff$ $A \begin{bmatrix} \vdots \\ \vdots \end{bmatrix} = 1 \cdot \begin{bmatrix} \vdots \\ \vdots \end{bmatrix}$ $\Rightarrow$ ($\lambda = 1$ is an eigenvalue),

and $\lambda(A) = \lambda(A^T)$

EXAMPLE:

$$A = \frac{1}{4}\begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix}$$ (column stochastic)

$$\Rightarrow \quad \lambda_* = 1 \qquad v_* = \begin{pmatrix} 2/5 \\ 3/5 \end{pmatrix}$$

normalized so all entries add up to 1.

Transpose

$$B = \frac{1}{4}\begin{bmatrix} 1 & 3 \\ 2 & 2 \end{bmatrix} \begin{matrix} =1 \\ =1 \end{matrix}$$ (row-stochastic)

$$\Rightarrow \quad \underline{\lambda_*} = 1 \qquad v_* = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

surprising?

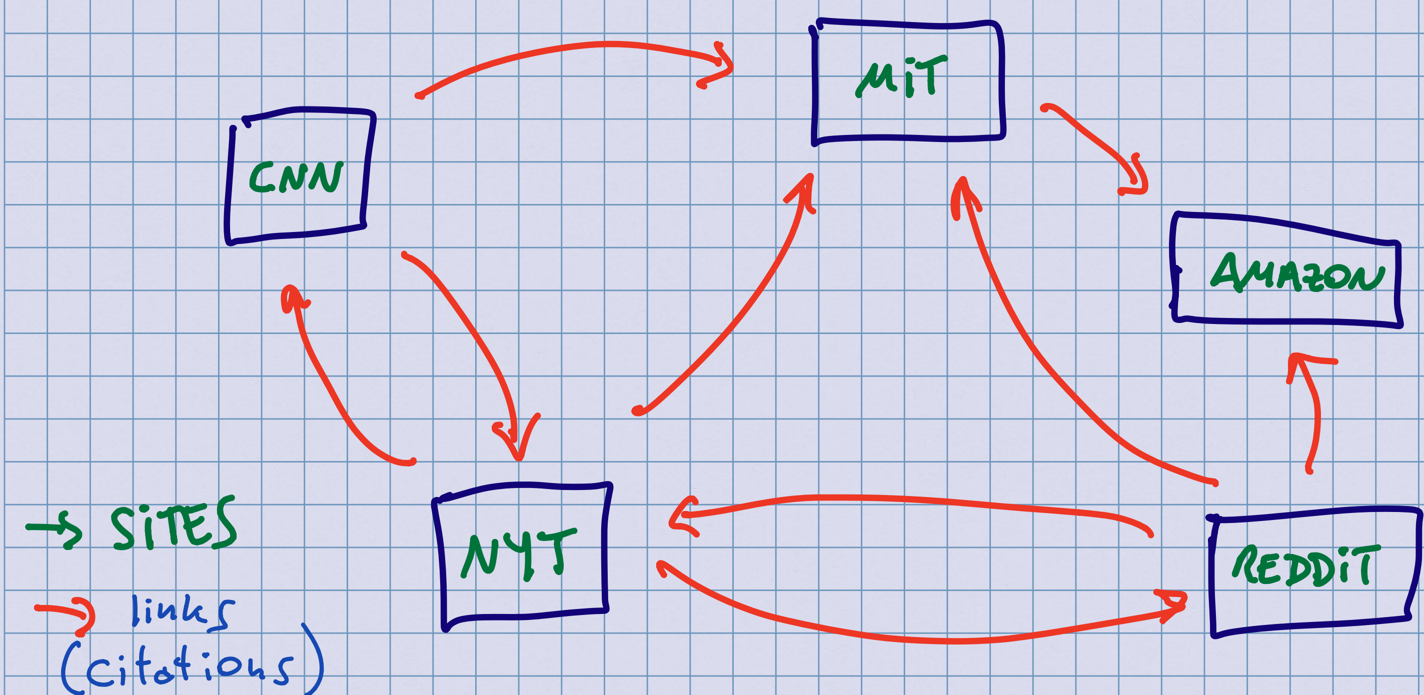
PAGE RANK :

How does a search engine (e.g. Google) rank websites?

What makes a site more "important" than others?

Many models (PageRank, HITS, etc...)

Here's a simple one:



(Def)

THE IMPORTANCE (RANK) OF A SITE DEPENDS ON HOW MANY (AND WHICH) IMPORTANT SITES POINT TO IT.

- RECURSIVE (OR "FIXED POINT" DEFINITION)

typo!

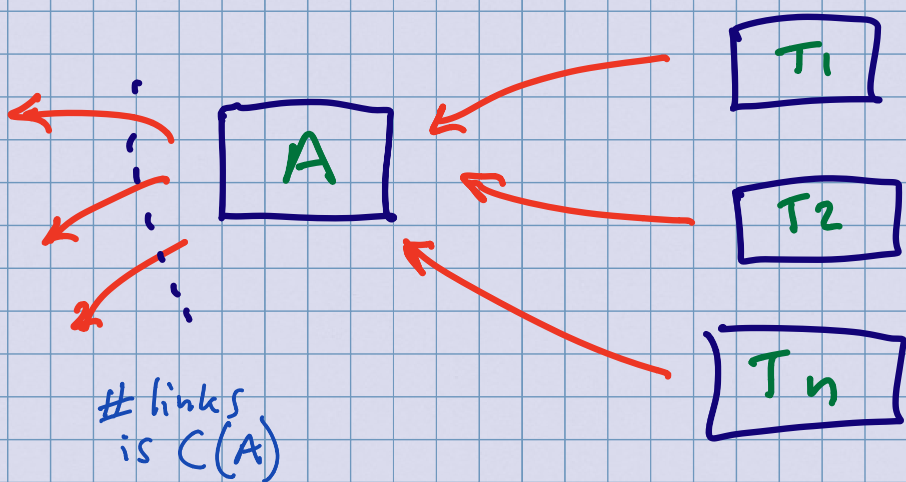
(from the Brin/Page paper)

We assume page A has pages T1...Tn which point to it (i.e., are citations). The parameter d is a damping factor which can be set between 0 and 1. We usually set d to 0.85. There are more details about d in the next section. Also C(A) is defined as the number of links going out of page A. The PageRank of a page A is given as follows:

$$PR(A) = \frac{(1-d)}{N} + d\,(PR(T1)/C(T1) + \ldots + PR(Tn)/C(Tn))$$

linear equations in PR

Note that the PageRanks form a probability distribution over web pages, so the sum of all web pages' PageRanks will be one.

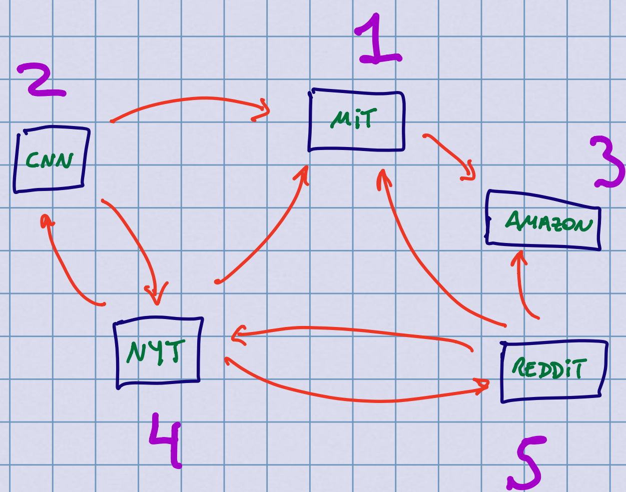


Writing it as a matrix... (linear system)

$$\begin{bmatrix} PR(T_1) \\ \vdots \\ \vdots \\ PR(T_n) \end{bmatrix} = \frac{1}{N} \begin{bmatrix} 1-d \\ 1-d \\ \vdots \\ 1-d \end{bmatrix} + d \cdot \begin{bmatrix} \begin{pmatrix} \text{WEIGHTED} \\ \text{ADJACENCY} \\ \text{MATRIX OF} \\ \text{GRAPH} \end{pmatrix} \end{bmatrix} \begin{bmatrix} PR(T_1) \\ \vdots \\ \vdots \\ PR(T_n) \end{bmatrix}$$

$$\begin{bmatrix} 0 & 1/2 & 0 & 1/3 & 1/3 \\ 0 & 0 & 0 & 1/3 & 0 \\ 1 & 0 & 0 & 0 & 1/3 \\ 0 & 1/2 & 0 & 0 & 1/3 \\ 0 & 0 & 1 & 1/3 & 0 \end{bmatrix}$$

(columns add up to 1)



- Parameter d chosen appropriately (mostly, to ensure matrix is strictly positive)

For $d = 0.85 \Rightarrow PR \simeq (0.19, 0.07, 0.28, 0.15, 0.31)$

- key practical difficulty: this is a <u>HUGE</u> problem! (~ 2 Billion websites)

As an eigenvalue problem:

$\left(\begin{array}{l} p \text{ is normalized} \\ \text{so } e^T p = 1, \\ \text{where } e = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \end{array} \right)$

From previous page:

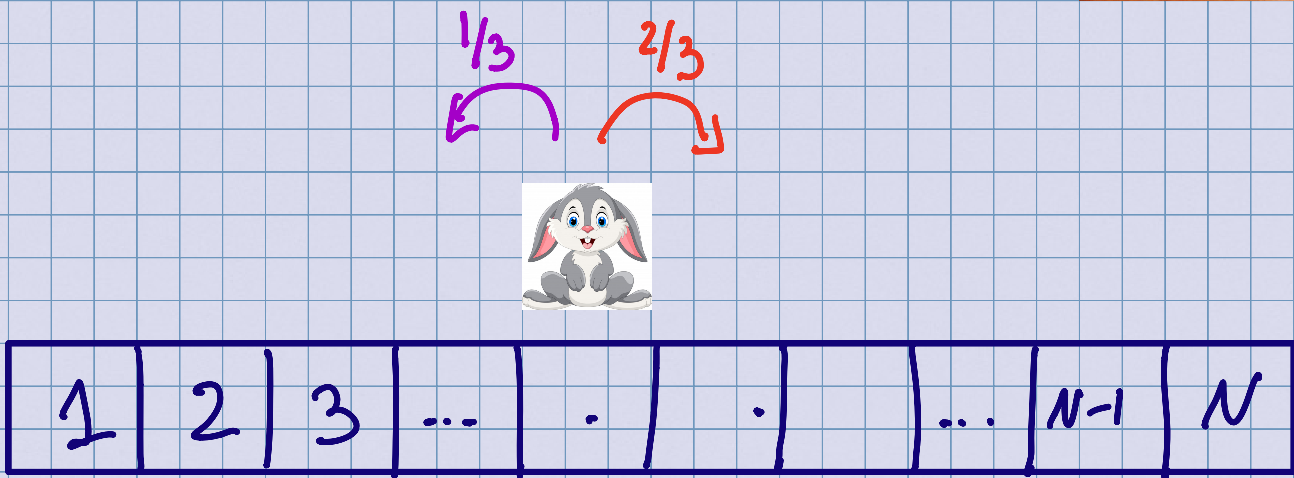
$$\begin{bmatrix} PR(T_1) \\ \vdots \\ \vdots \\ PR(T_n) \end{bmatrix} = \frac{1}{N} \begin{bmatrix} 1-d \\ 1-d \\ \vdots \\ 1-d \end{bmatrix} + d \cdot \begin{bmatrix} \begin{pmatrix} \text{WEIGHTED} \\ \text{ADJACENCY} \\ \text{MATRIX OF} \\ \text{GRAPH} \end{pmatrix} \end{bmatrix} \begin{bmatrix} PR(T_1) \\ \vdots \\ \vdots \\ PR(T_n) \end{bmatrix}$$

$$p = \frac{(1-d)}{N} e + d\,G\,p$$

$$p = \underbrace{\left[\frac{(1-d)}{N} ee^T + d\,G \right]}_{\text{Makes every entry strictly positive}} p$$

$\left(\text{since } e^T p = 1\right)$

HOPPING RABBIT

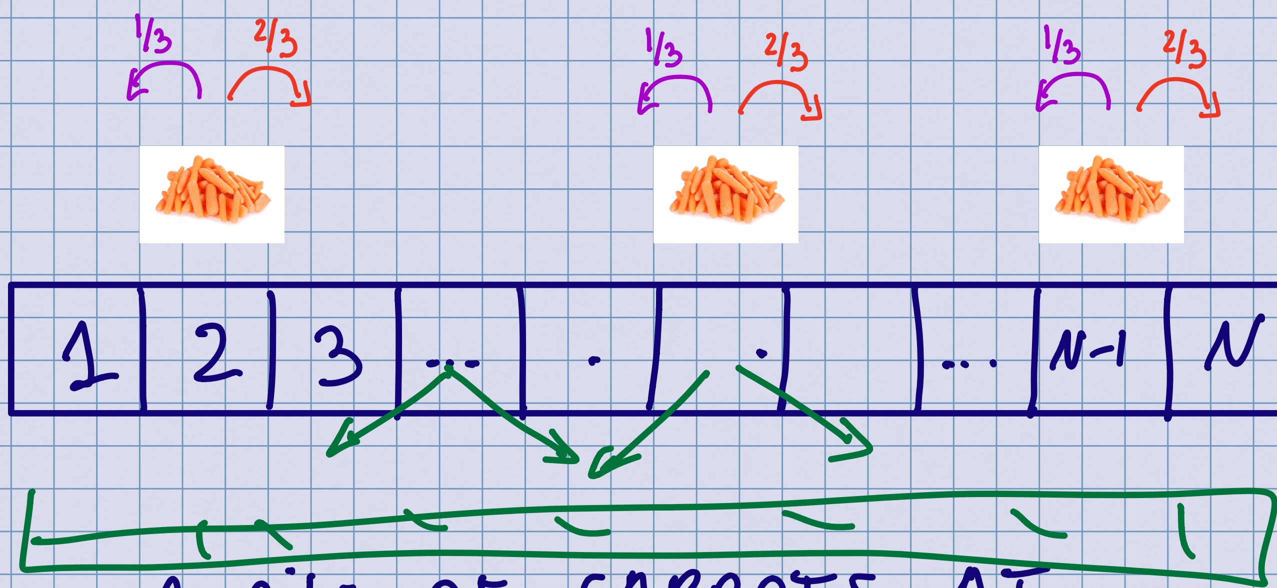


- AT EACH TIME STEP, THE BUNNY FLIPS A COIN (BIASED).
  - WITH PROB $1/3$ JUMPS TO THE LEFT
  - WITH PROB $2/3$ JUMPS TO THE RIGHT

- IF HE REACHES THE ENDS, BOUNCES BACK
  - IF AT BOX 1, AND JUMPS LEFT, REMAINS AT 1
  - IF AT BOX N, AND JUMPS RIGHT, REMAINS AT N

THE RABBIT HAS BEEN JUMPING ALL DAY LONG, AND YOU SHOW UP LATE ONE AFTERNOON TO VISIT HIM.

(ALTERNATIVE INTERPRETATION)

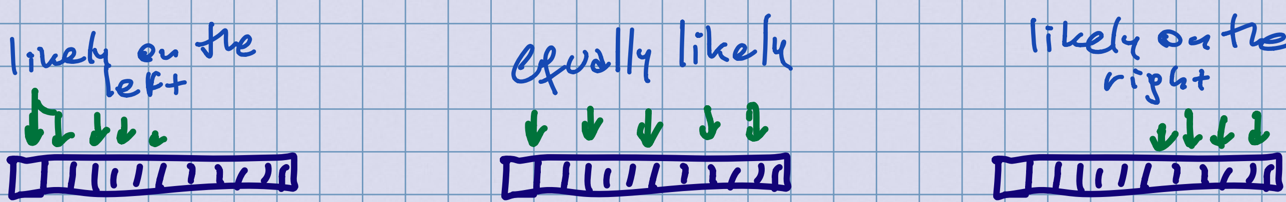


- A PILE OF CARROTS AT EVERY BOX
- EVERY TIMESTEP, WE SPLIT THEM.

[IF YOU DON'T LIKE RABBITS, YOU CAN THINK, E.G., ABOUT PARTICLES (ELECTRONS ON A LATTICE) UNDER AN EXTERNAL FIELD.]

(OR MANY OTHER THINGS)

Q: WHERE WOULD YOU EXPECT TO FIND THE RABBIT?
(e.g., left-side vs. right side?)

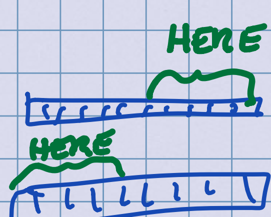


$\boxed{\text{SAY } N=10.}$

Q: IF I <u>BET</u> YOU WITH ODDS 1 TO 1 THAT THE RABBIT IS ON THE <u>RIGHT</u> HALF, WOULD YOU ACCEPT THE BET?

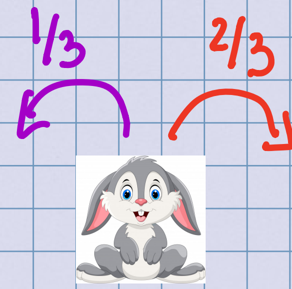
2 to 1?   10 to 1?   100 to 1?

(x to 1 :

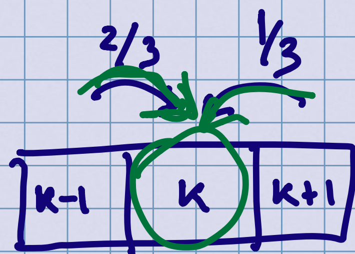


, YOU PAY ME 1
, I PAY YOU X)

HOW TO SOLVE THIS ?



Let p_k be the probability that
the rabbit is in box K.



Then :

$$\boxed{p_K = \frac{2}{3}\, p_{K-1} + \frac{1}{3}\, p_{K+1}}$$

(For $2 \le K \le N-1$)

$$\begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_k \\ \vdots \\ p_N \end{bmatrix} = \overbrace{\begin{bmatrix} 1/3 & 1/3 & & & & & \\ 2/3 & 0 & 1/3 & & & & \\ & 2/3 & 0 & 1/3 & & & \\ & & 2/3 & 0 & 1/3 & & \\ & & & \ddots & \ddots & \ddots & \\ & & & & 2/3 & 0 & 1/3 \\ & & & & & 2/3 & 2/3 \end{bmatrix}}^{A} \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_k \\ \vdots \\ p_N \end{bmatrix}$$

$$p = A\,p$$

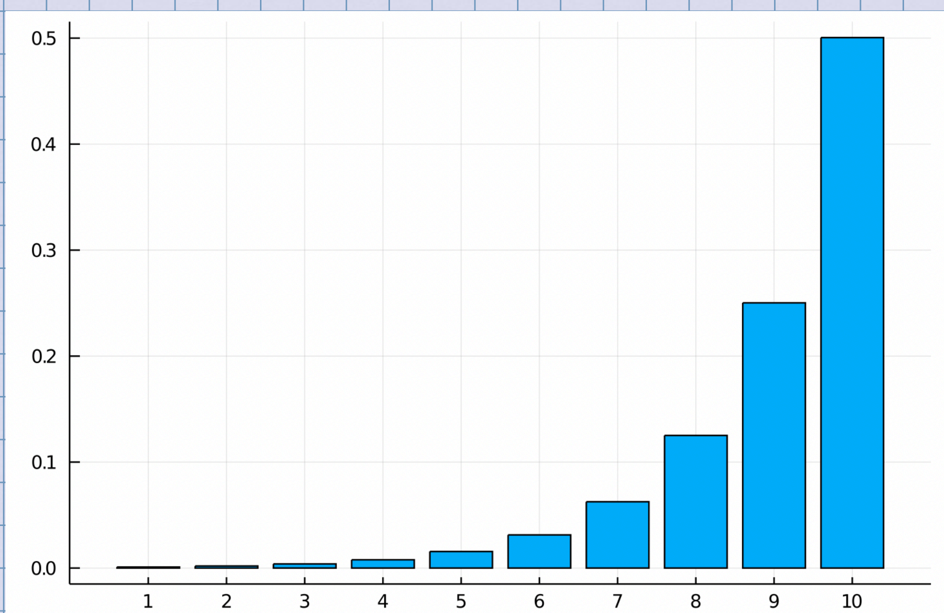
(and $\sum p_i = 1$)

- A is a nonnegative matrix
- Every column adds up to 1.

The vector p of probabilities is the solution of a linear system.

Alternatively, it is the eigenvector of A associated with the eigenvalue $\underline{1}$.

In our case $(N = 10)$, we have



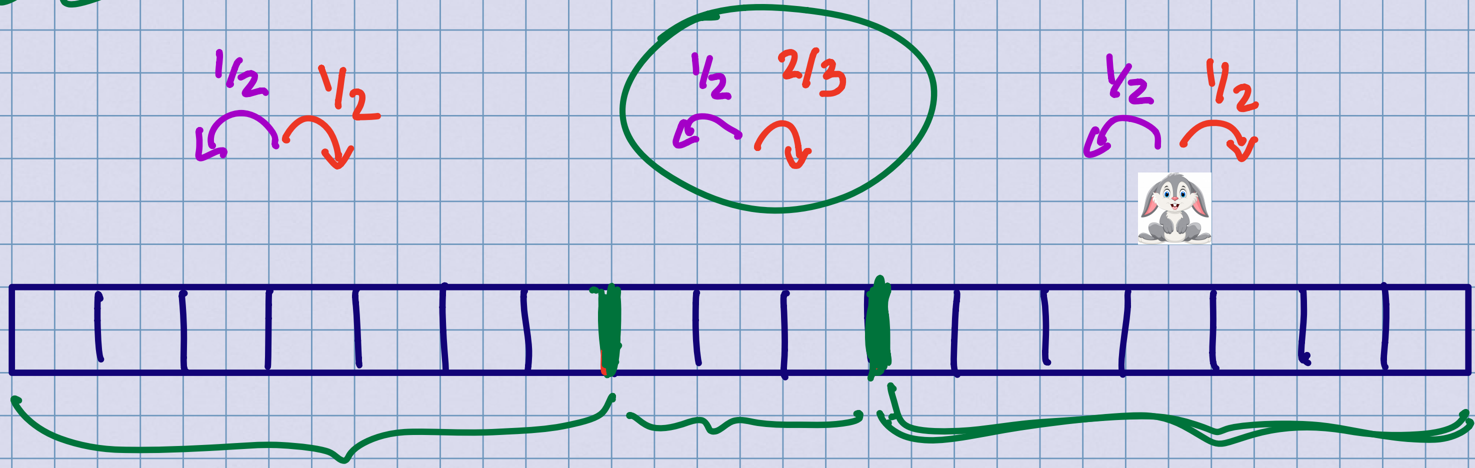
$\underbrace{\qquad\qquad}_{3\%}$ $\underbrace{\qquad\qquad}_{97\%}$

probabilities p_i (after solving the linear system)

$\Rightarrow$ "Fair odds" are $\sim 1:32$

CHALLENGE!

"TUNNEL EFFECT"



REGION 1	REGION 2	REGION 3
(EQUALLY LIKELY)	(BIASED TO RIGHT)	(EQUALLY LIKELY)

- WHAT DO YOU EXPECT WILL HAPPEN?
- HOW DOES THE SIZE OF THE MIDDLE REGION AFFECT THE RESULTS?

"MARKOV CHAINS"