Stationary Distribution

- We'll show a surprising phenomenon: Markov chain ends in a unique distribution!
  (i.e. $\lim_{n\to\infty} M^n$ exists!)

- Counter-example? $M = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \Rightarrow M^2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} =: I$.
  $\Rightarrow$ Evolution: $M, I, M, I, \ldots \Rightarrow$ No limit!

$\rightarrow$ Issue with M is: It has no way to go from $1 \to 1$. (i.e. $M_{11} = 0$)
  " " M^2 is: $M_{12} = 0 \Rightarrow$ No way to move $1 \to 2$.

Let's exclude such examples:

Defn: Markov Chain $(\mathcal{M}, M)$ is regular if $\exists t$ s.t. $\forall\, i,j \in S : (M^t)_{ij} > 0$.

[I.e. M^t has all entries > 0. So every transition is 'allowed'.]

$\triangleright$ M^t has all entries $>0 \Rightarrow M^{t+1}$ has all entries >0.

Pf: • Note that in each row of M, some entry >0. (& others ≥ 0)

$\Rightarrow M\cdot(M^t)$ has every entry >0. $\square$

- For such M^t, M^{t+1} has a nice physical interpretation

its i-th-row $= (M^t)_{i*} \cdot M \Rightarrow$ So, it is like averaging the rows of M (without leaving any one out!). What happens if you do this many many times?

Theorem (Perron-Frobenius 1907): If M is the transition matrix of a regular Markov chain, then $\lim_{n\to\infty} M^n =: \bar{1} \cdot w^T$, where $\bar{1} :=$ column vector with all 1's & w is some prob. distribution.

↳ w is called the stationary distribution.

↳ For initial distribution μ: $\lim (\mu^T \cdot M^n) = \mu^T \cdot \bar{1} \cdot w^T = w^T$ is independent of μ! (Memorylessness?)

↳ Matrix $T \cdot w^T$ has rank $= 1$.

Pf. sketch: • We'll give an intuitive sketch.

Idea: Show that in the matrix action $M \cdot v_0 =: v_1$ the entries of v_1 'get closer' to each other, compared to those in v_0.

$\cdots\cdots \Longrightarrow M^n \cdot v_0$ is a scalar !

• Define: In v_0, max $=: M_0$ & min $=: m_0$.

In v_1, max $=: M_1$ & min $=: m_1$.

• Let M have min entry $=: \delta$.

(Wlog $\delta > 0$, else we work with M^t of positive entries.)

▷ $0 < \delta \leq 1/2$. [Pf: $\delta > 1/2 \Rightarrow$ a row-sum in M $\geq |S| \cdot \delta > |S|/2 \geq 1$. □]

• Consider the image-vector $v_1 = M \cdot v_0$:

each entry in v_1 is $\leq M_0 \cdot (1-\delta) + m_0 \cdot \delta$

$\geq 1/2$ $\leq 1/2$

[Why? Use row-sum in M is 1.]

each entry in v_1 is $\geq M_0 \cdot \delta + m_0 \cdot (1-\delta)$

[Why? $\delta \leq 1/2$ & row-sum in M is 1.]

$$\Rightarrow (M_1 - m_1) \leq M_0 \cdot (1-2\delta) + m_0 \cdot (2\delta - 1)$$
$$= (M_0 - m_0) \cdot (1-2\delta) < M_0 - m_0 \quad [\because \delta > 0]$$

▷ The gap in v_1 has <u>fallen</u> by the <u>fraction</u> $(1-2\delta)$.

$\Rightarrow \lim_{n\to\infty} v_n$ has equal entries.

$\Rightarrow v_n$ is a scalar $c_{v_0} \cdot \mathbb{1}$, for $c_{v_0} \in \mathbb{R}_{>0}$.

$\triangleright \lim_{n\to\infty} v_n = \lim_{n\to\infty} M^n \cdot v_0 = c_{v_0} \cdot \mathbb{1}$ --- (1)

- Vary v_0 as $\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\\\vdots\\0\end{pmatrix}, \ldots, \begin{pmatrix}0\\\vdots\\0\\1\end{pmatrix}$ $\rightarrow$ $|S|$ elementary vectors.

$\Rightarrow$ Eqn.(1) gives: $\lim_{n\to\infty} M^n = [c_1 \cdot \mathbb{1}, c_2 \cdot \mathbb{1}, \ldots, c_{|S|} \cdot \mathbb{1}]$

$= \mathbb{1} \cdot (c_1, \ldots, c_{|S|}) =: \mathbb{1} \cdot \underline{w^T}$

[Recall: $\mu^T \cdot \mathbb{1} \cdot w^T = w^T$.] $\Rightarrow$ 'stationary distribution' $\square$

- Exercise 1: Why did we use the right-action $M \cdot v$, when left-action $p^T \cdot M$ is more natural?

- Exercise 2: M is doubly-stochastic $\Longrightarrow$ stationary-distribution is uniform distribution!

$\triangleright$ $M^n \longrightarrow \frac{1}{|S|} \cdot J$ [J is all 1 matrix.]

[Hint: Apply the Thm. on M & M^T.]

↳ Random walk in undirected graphs gives a doubly-stochastic process.

↳ visiting all vertices in the end!

Page Rank Algorithm

- eg. You want to rank pages on the Internet.
  First criterion: More links to webpage X means higher rank(X).

- Consider the Internet graph $G=(V,E)$:
  vertices $V :=$ set of pages $[n]$. ($n \approx 1$ billion or more!)
  edges E has (i,j) iff page-i links to page-j.

eg. $\begin{tikzcd} 0 & 1 \arrow[r] \arrow[rr, bend left] & 2 \arrow[r] & 3 \arrow[r] \arrow[ll, bend left] & 4 & 5 \end{tikzcd}$

Strategy 1: Label vertex-i by rank =

$$p_i := \sum_{j:(j,i)\in E} 1 \quad =: \text{in-deg}(i).$$

- This ranking seems to ignore Quality !?
- Improvements: (i) If j links to many, then give it less weight; and
  (ii) If j is less "important", then give it less weight.

Strategy 2: Define i's rank $p_i := \sum_{j:(j,i)\in E} p_j/m_j$,

where $m_j := \text{out-deg}(j) := \#\text{links from } j$.

- Qn: This is a cyclic definition. What are p_i's now? Does a ranking even exist?

- Relevant Qn: Could we turn this into a random walk on G, and hence into a Markov chain?

Assume:

(i) A person starts from a rnd. page in G, &

(ii) She moves from pg-j to pg-i with prob $= 1/m_j$.

▷ This defines a Markov chain, that should give the rk-vector $\bar{p}$ as its stationary distribution!

Qn: Is it regular? Does $\lim_{m\to\infty} M^m$ exist?

↳ In general, No! [Since an isolated page may exist.]

– So let's modify (ii) to:

(ii)' The web-surfer is allowed to "stray" to a random page with prob $=: p > 0$, or "follow" a link in the current-page.

Strategy 3: Thus,

$$M'_{ji} := \begin{cases} \underbrace{p \cdot \frac{1}{n}}_{\text{(stray)}} + \overbrace{(1-p)\cdot \frac{1}{m_j}}^{\text{(follow)}} , & \text{if } (j,i) \in E \\ p \cdot \frac{1}{n} , & \text{if } (j,i) \notin E . \end{cases}$$

▷ $M' = p \cdot J_n + (1-p) \cdot M$; where $J_n := \frac{1}{n} \cdot J$

↑ from Strategy-2 ↑ all-1 matrix

▷ M' is a regular, homog. Markov chain.

Pf: • M' has +ve entries ⇒ M' is regular.

• row-sum in M' = $p \cdot$ row-sum$(J_n) + (1-p) \cdot$ row-sum(M)

$= p \cdot 1 + (1-p) \cdot 1 = 1$.

□

Defn: $\lim_{m \to \infty} (M')^m =: \mathbb{1} \cdot w^T$, where w gives the rank of the n webpages.

Qn: How do you compute w, in practice?

Exercise: Develop an algo. using eigen-space(M').