Lecture 2: Markov Chains

John Sylvester   Nicolás Rivera   Luca Zanetti   Thomas Sauerwald

Lent 2019

UNIVERSITY OF CAMBRIDGE

Outline

Stochastic Process

Stopping and Hitting Times

Irreducibility and Stationarity

Periodicity and Convergence

Bonus: Gamblers Ruin

Stochastic Process

A *Stochastic Process* $X = \{X_t : t \in T\}$ is a collection of random variables indexed by time (often $T = \mathbb{N}$) and in this case $X = (X_i)_{i=0}^{\infty}$.

A vector $\mu = (\mu_i)_{i \in \mathcal{I}}$ is a *Probability Distribution* or *Probability Vector* on $\mathcal{I}$ if $\mu_i \in [0, 1]$ and

$$\sum_{i \in \mathcal{I}} \mu_i = 1.$$

Markov Chains

Markov Chain (Discrete Time and State, Time Homogeneous)

We say that $(X_i)_{i=0}^{\infty}$ is a *Markov Chain* on *State Space* $\mathcal{I}$ with *Initial Distribution* μ and *Transition Matrix* P if for all $t \geq 0$ and $i_0, \dots \in \mathcal{I}$,

- $\mathbf{P}[\, X_0 = i\,] = \mu_i$.
- The *Markov Property* holds:

$$\mathbf{P}\Big[\, X_{t+1} = i_{t+1} \Big| X_t = i_t, \dots, X_0 = i_0 \,\Big] = \mathbf{P}\Big[\, X_{t+1} = i_{t+1} \Big| X_t = i_t \,\Big] := P_{i_t, i_{t+1}}.$$

From the definition one can deduce that (check!)

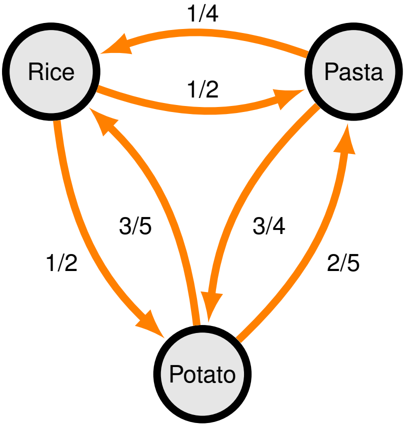
- $\mathbf{P}[\, X_{t+1} = i_{t+1}, X_t = i_t, \dots, X_1 = i_1, X_0 = i_0 \,] = \mu_{i_0} \cdot P_{i_0, i_1} \cdots P_{i_{t-1}, i_t} \cdot P_{i_t, i_{t+1}}$
- $\mathbf{P}[\, X_{t+m} = i \,] = \sum_{j \in \mathcal{I}} \mathbf{P}[\, X_{t+m} = i | X_t = j \,] \, \mathbf{P}[\, X_t = j \,]$

If the Markov Chain starts from as single state $i \in \mathcal{I}$ then we use the notation

$$\mathbf{P}_i[X_k = j] := \mathbf{P}[\, X_k = j | X_0 = i \,].$$

What does a Markov Chain Look Like?

Example : the carbohydrate served with lunch in the college cafeteria.



This has transition matrix:

$$
\begin{array}{c} \\ P = \end{array}
\begin{array}{c}
\begin{array}{ccc} \text{Rice} & \text{Pasta} & \text{Potato} \end{array} \\
\begin{bmatrix} 0 & 1/2 & 1/2 \\ 1/4 & 0 & 3/4 \\ 3/5 & 2/5 & 0 \end{bmatrix}
\end{array}
\begin{array}{l} \\ \text{Rice} \\ \text{Pasta} \\ \text{Potato} \end{array}
$$

Transition Matrices

The *Transition Matrix* P of a Markov chain (μ, P) on $\mathcal{I} = \{1, \ldots n\}$ is given by

$$P = \begin{pmatrix} P_{1,1} & \ldots & P_{1,n} \\ \vdots & \ddots & \vdots \\ P_{n,1} & \ldots & P_{n,n} \end{pmatrix}.$$

- $p_i(t)$: probability the chain is in state i at time t.
- $\vec{p}(t) = (p_0(t), p_1(t), \ldots, p_n(t))$: *State vector* at time t (Row vector).
- Multiplying $\vec{p}(t)$ by P corresponds to advancing the chain one step:

$$p_i(t+1) = \sum_{j \in \mathcal{I}} p_j(t) \cdot P_{j,i} \qquad \text{and thus} \qquad \vec{p}(t+1) = \vec{p}(t) \cdot P.$$

- The Markov Property and line above imply that for any $k, t \geq 0$

$$\vec{p}(t+k) = \vec{p}(t) \cdot P^k \qquad \text{and thus} \qquad P^k_{i,j} = \mathbf{P}[\, X_k = j | X_0 = i\,].$$

Thus $p_i(t) = (\mu P^t)_i$ and so $\vec{p}(t) = \mu P^t = ((\mu P)_1, (\mu P)_2, \ldots, (\mu P)_n)$.

Outline

Stochastic Process

Stopping and Hitting Times

Irreducibility and Stationarity

Periodicity and Convergence

Bonus: Gamblers Ruin

Stopping and Hitting Times

A non-negative integer random variable τ is a *Stopping Time* for $(X_i)_{i \geq 0}$ if for every $n \geq 0$ the event $\{\tau = n\}$ depends only on $X_0, \ldots, X_n$.

Example - College Carbs Stopping times:

✓ "We had Pasta yesterday"

× "We are having Rice next Thursday"

For two states $x, y \in \mathcal{I}$ we call $h_{x,y}$ the *Hitting Time* of y from x:

$$h_{x,y} := \mathbf{E}_x[\tau_y] = \mathbf{E}[\tau_y | X_0 = x] \quad \text{where } \tau_y = \inf\{t \geq 0 : X_t = y\}.$$

For $x \in \mathcal{I}$ the *First Return Time* $\mathbf{E}_x\left[\tau_x^+\right]$ of x is defined

$$\mathbf{E}_x\left[\tau_x^+\right] = \mathbf{E}\left[\tau_x^+ | X_0 = x\right] \quad \text{where } \tau_x^+ = \inf\{t \geq 1 : X_t = x\}.$$

Comments

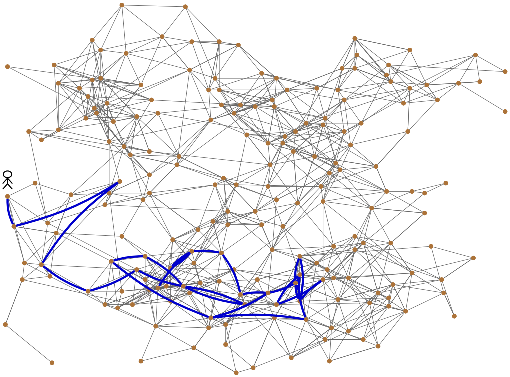
- Notice that $h_{x,x} = \mathbf{E}_x[\tau_x] = 0$ whereas $\mathbf{E}_x\left[\tau_x^+\right] \geq 1$.
- For any $y \neq x$, $h_{x,y} = \mathbf{E}_x\left[\tau_y^+\right]$.
- Hitting times are the solution to the set of linear equations:

$$\mathbf{E}_x\left[\tau_y^+\right] \overset{\text{Markov Prop.}}{=} 1 + \sum_{z \in \mathcal{I}} \mathbf{E}_z[\tau_y] \cdot P_{x,z} \qquad \forall x, y \in V.$$

Random Walks on Graphs

A *Simple Random Walk (SRW)* on a graph G is a Markov chain on $V(G)$ with

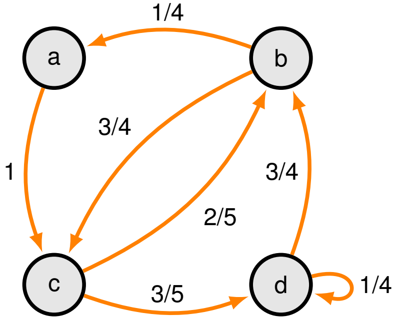
$$P_{ij} = \begin{cases} \frac{1}{d(i)} & \text{if } ij \in E \\ 0 & \text{if } ij \notin E \end{cases}.$$



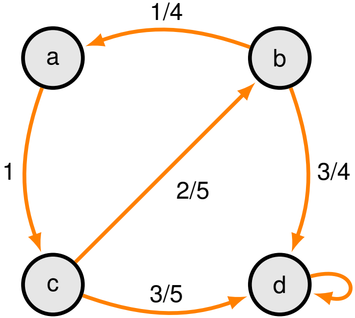
Outline

Irreducible Markov Chains

A Markov chain is *Irreducible* if for every pair of states $(i, j) \in \mathcal{I}^2$ there is an integer $m \geq 0$ such that $P_{i,j}^m > 0$.



✓ irreducible



× not-irreducible (thus reducible)

Finite Hitting Theorem

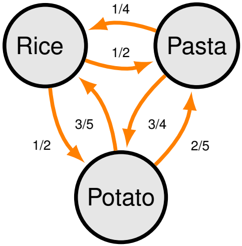
For any states x and y of a finite irreducible Markov chain $\mathbf{E}_x\left[\tau_y^+\right] < \infty$.

Stationary Distribution

A probability distribution $\pi = (\pi_1, \ldots, \pi_n)$ is the *Stationary Distribution* of a Markov chain if $\pi P = \pi$, i.e. π is a left eigenvector with eigenvalue 1.

College carbs example:

$$\underbrace{\left(\frac{4}{13}, \frac{4}{13}, \frac{5}{13}\right)}_{\pi} \cdot \underbrace{\begin{pmatrix} 0 & 1/2 & 1/2 \\ 1/4 & 0 & 3/4 \\ 3/5 & 2/5 & 0 \end{pmatrix}}_{P} = \underbrace{\left(\frac{4}{13}, \frac{4}{13}, \frac{5}{13}\right)}_{\pi}$$



A Markov chain reaches *Equilibrium* if $\vec{p}(t) = \pi$ for some t. If equilibrium is reached it *Persists*: If $\vec{p}(t) = \pi$ then $\vec{p}(t+k) = \pi$ for all $k \geq 0$ since

$$\vec{p}(t+1) = \vec{p}(t)P = \pi P = \pi = \vec{p}(t).$$

Existence of a Stationary Distribution

Existence and uniqueness of a positive stationary distribution

Let P be finite, irreducible M.C., then there is a unique probability distribution π on $\mathcal{I}$ such that $\pi = \pi P$ and $\pi_x = 1/\mathbf{E}_x\left[\tau_x^+\right] > 0, \forall x \in \mathcal{I}$.

Proof: [Existence] Fix $z \in \mathcal{I}$ and define $\mu_y = \sum_{t=0}^{\infty} \mathbf{P}_z\left[X_t = y, \tau_z^+ > t\right]$, this is the expected number of visits to y before returning to z. For any state y, we have $0 < \mu_y \leq \mathbf{E}_z\left[\tau_z^+\right] < \infty$ since P is irreducible. To show $\mu P = \mu$ we have

$$
\begin{aligned}
(\mu P)_y &= \sum_{x\in\mathcal{I}} \mu_x \cdot P_{x,y} = \sum_{x\in\mathcal{I}} \sum_{t=0}^{\infty} \mathbf{P}_z\left[X_t = x, \tau_z^+ > t\right] \cdot P_{x,y} \\
&= \sum_{x\in\mathcal{I}} \sum_{t=0}^{\infty} \mathbf{P}_z\left[X_t = x, X_{t+1} = y, \tau_z^+ > t\right] = \sum_{t=0}^{\infty} \mathbf{P}_z\left[X_{t+1} = y, \tau_z^+ > t\right] \\
&= \sum_{t=0}^{\infty} \mathbf{P}_z\left[X_{t+1} = y, \tau_z^+ > t+1\right] + \mathbf{P}_z\left[X_{t+1} = y, \tau_z^+ = t+1\right] \\
&= \mu_y - \overset{(a)}{\mathbf{P}_z\left[X_0 = y, \tau_z^+ > 0\right]} + \overset{(b)}{\sum_{t=0}^{\infty} \mathbf{P}_z\left[X_{t+1} = y, \tau_z^+ = t+1\right]} = \mu_y.
\end{aligned}
$$

Where (a) and (b) are 1 if $y = z$ and 0 otherwise so cancel. Divide μ though by $\sum_{x\in\mathcal{I}} \mu_x < \infty$ to turn it into a probability distribution π. □

Uniqueness of the Stationary Distribution

Existence and uniqueness of a positive stationary distribution

Let P be finite, irreducible M.C., then there is a unique probability distribution π on $\mathcal{I}$ such that $\pi = \pi P$ and $\pi_x = 1/\mathbf{E}_x\left[\tau_x^+\right] > 0$, $\forall x \in \mathcal{I}$.

Proof: [Uniqueness] Assume P has a stationary distribution μ and let $\mathbf{P}[\, X_0 = x\,] = \mu_x$. We shall show μ is uniquely determined

$$\mu_x \cdot \mathbf{E}_x\left[\tau_x^+\right] \stackrel{\text{Hw1}}{=} \mathbf{P}[\, X_0 = x\,] \cdot \sum_{t \geq 1} \mathbf{P}\left[\, \tau_x^+ \geq t \mid X_0 = x\,\right]$$

$$= \sum_{t \geq 1} \mathbf{P}\left[\, \tau_x^+ \geq t, X_0 = x\,\right]$$

$$= \mathbf{P}[\, X_0 = x\,] + \sum_{t \geq 2} \mathbf{P}[\, X_1 \neq x, \ldots, X_{t-1} \neq x\,] - \mathbf{P}[\, X_0 \neq x, \ldots, X_{t-1} \neq x\,]$$

$$\stackrel{(a)}{=} \mathbf{P}[\, X_0 = x\,] + \sum_{t \geq 2} \mathbf{P}[\, X_0 \neq x, \ldots, X_{t-2} \neq x\,] - \mathbf{P}[\, X_0 \neq x, \ldots, X_{t-1} \neq x\,]$$

$$\stackrel{(b)}{=} \mathbf{P}[\, X_0 = x\,] + \mathbf{P}[\, X_0 \neq x\,] - \lim_{t \to \infty} \mathbf{P}[\, X_0 \neq x, \ldots, X_{t-1} \neq x\,] \stackrel{(c)}{=} 1.$$

A sum S is Telescoping if

$$S = \sum_{i=0}^{n-1} a_i - a_{i+1} = a_0 - a_n.$$

Equality (a) follows as μ is stationary, equality (b) since the sum is telescoping and (c) by Markov's inequality and the Finite Hitting Theorem. □

Outline

Stochastic Process

Stopping and Hitting Times

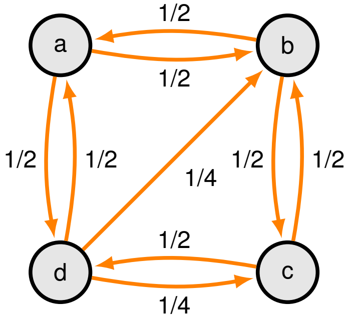
Irreducibility and Stationarity

Periodicity and Convergence

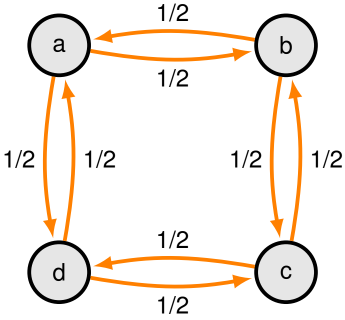
Bonus: Gamblers Ruin

Periodicity

- A Markov chain is *Aperiodic* if for all $x, y \in \mathcal{I}$, $\gcd\{t : P^t_{x,y} > 0\} = 1$.
- Otherwise we say it is *Periodic*.



✓ Aperiodic



× Periodic

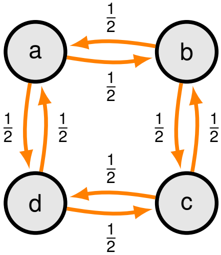
Lazy Random Walks and Periodicity

For some graphs G the simple random walk on G is periodic, as seen below. The *Lazy Random Walk (LRW)* on G given by $\widetilde{P} = (P + I)/2$,

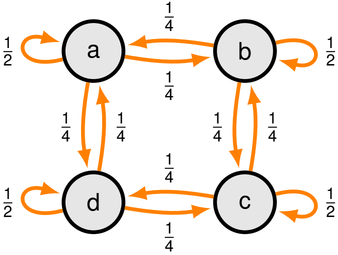
$$\widetilde{P}_{i,j} = \begin{cases} \frac{1}{2d(i)} & \text{if } ij \in E \\ \frac{1}{2} & \text{if } i = j \\ 0 & \text{Otherwise} \end{cases}.$$

P - SRW matrix
I - Identity matrix.

Fact: for any graph G the LRW on G is Aperiodic.



SRW on C_4, *Periodic*



LRW on C_4, *Aperiodic*

Convergence

Convergence Theorem

Let P be any finite, aperiodic, irreducible Markov chain with stationary distribution π. Then for any $i, j \in \mathcal{I}$

$$\lim_{t \to \infty} P^t_{j,i} = \pi_i.$$

- Proved : For finite irreducible Markov chains π exists, is unique and

$$\pi_x = \frac{1}{\mathbf{E}_x\left[\tau_x^+\right]} > 0.$$

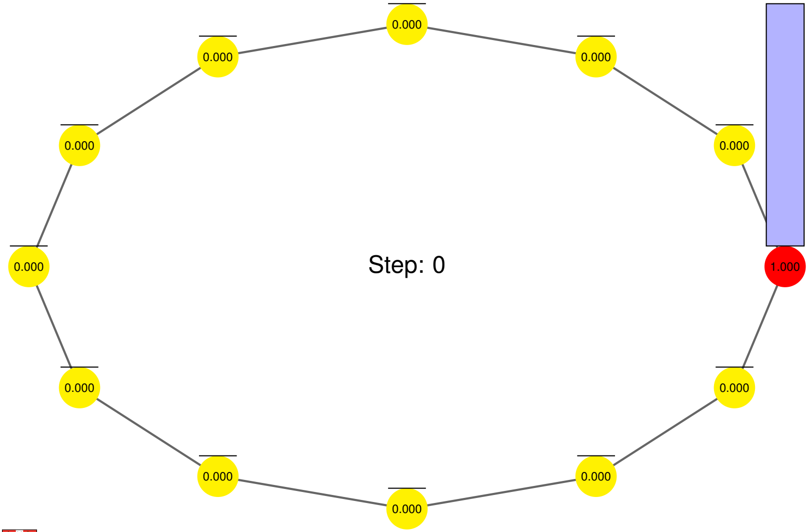
- If $P^t_{j,i}$ converges for all i, j we say the chain *Converges to Stationarity*.

Corollary

The Lazy random walk on any finite connected graph converges to stationarity.

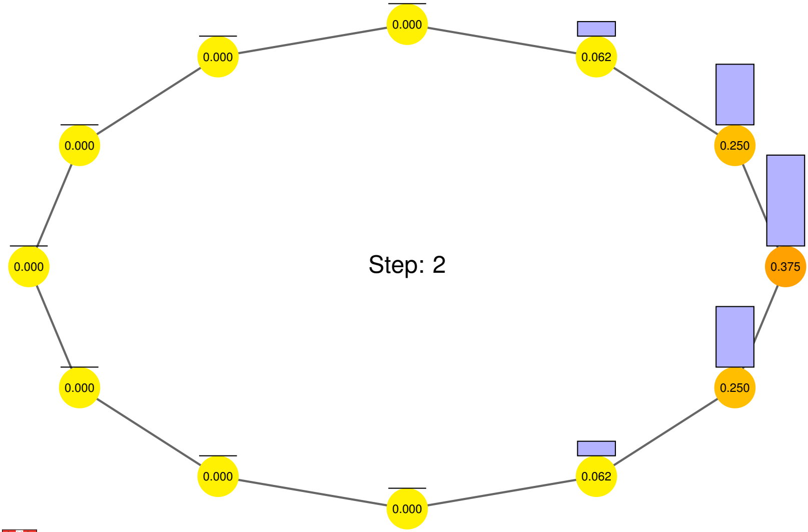
Convergence to Stationarity for the LRW on C_{12} from 0

At step t the value at vertex x is $P^t_{0,x}$.



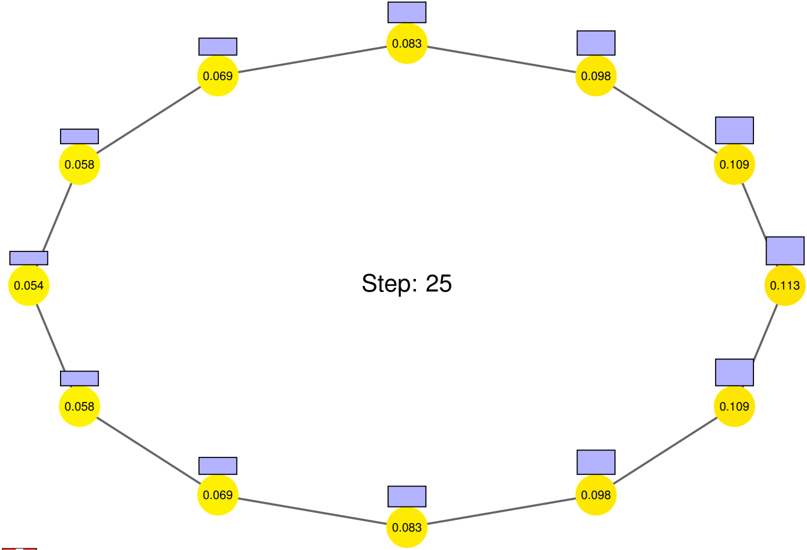
Convergence to Stationarity for the LRW on C_{12} from 0

At step t the value at vertex x is $P^t_{0,x}$.



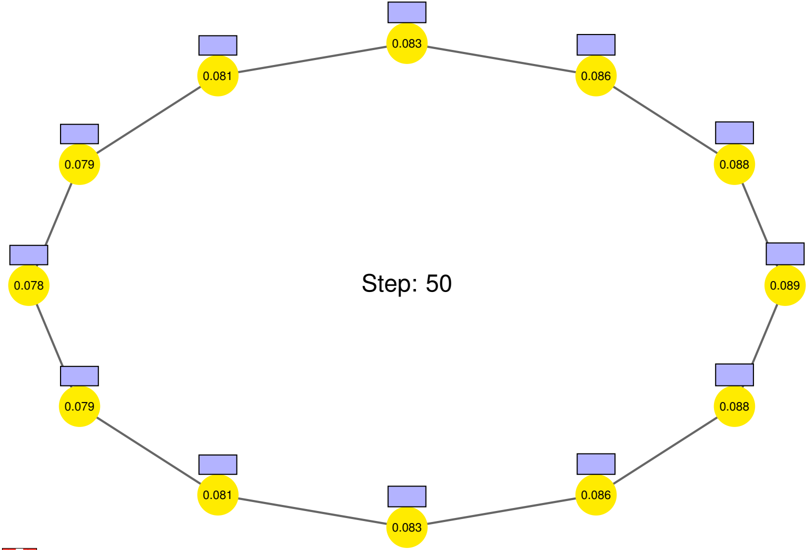
Convergence to Stationarity for the LRW on C_{12} from 0

At step t the value at vertex x is $P^t_{0,x}$.



Convergence to Stationarity for the LRW on C_{12} from 0

At step t the value at vertex x is $P^t_{0,x}$.



Outline

Gamblers Ruin

Gamblers Ruin

A gambler bets $\$1$ repeatedly on a biased coin ($\mathbf{P}[\,win\,] = a$, $\mathbf{P}[\,lose\,] = b = 1 - a$) until they either go broke or have $\$n$. What's more likely?

- Markov chain on $\{0, \ldots, n\}$ with $P_{i,i+1} = a$ and $P_{i,i-1} = b$ for each $1 \leq i \leq n$ and $P_{0,0} = P_{n,n} = 1$.
- Let X_t be the gambler's fortune at time t. Then, for any $S \subseteq \{0, \ldots, n\}$, $\tau_S = \inf\{t : X_t \in S\}$ is a stopping time.

Proposition

If the gambler starts with $\$s$, where $0 \leq s \leq n$, then

$$\mathbf{P}_s[\text{Gambler reaches } \$n \text{ before going broke}] = \begin{cases} \frac{1-\left(\frac{a}{b}\right)^s}{1-\left(\frac{a}{b}\right)^n} & \text{if } a \neq b \\ \frac{s}{n} & \text{if } a = b = 1/2 \end{cases}.$$

Gamblers Ruin

Proof: Let $\tau = \inf\{t \geq 0 : X_t \in \{0, n\}\}$ and $p_i = \mathbf{P}[X_\tau = n | X_0 = i]$. Then by the Law of total probability and the Markov property we have

$$p_i = ap_{i+1} + bp_{i-1}.$$

Using $1 = a + b$ and rearranging the above we have

$$p_{i+1} - p_i = \frac{b}{a}(p_i - p_{i-1}) = \cdots = \left(\frac{b}{a}\right)^i (p_1 - p_0) = \left(\frac{b}{a}\right)^i p_1. \tag{1}$$

Expressing $p_{i+1} = (p_{i+1} - p_i) + p_1$, writing it as a sum and applying (1) yields

$$p_{i+1} = p_1 + \sum_{k=1}^{i}(p_{k+1} - p_k) = p_1 + \sum_{k=1}^{i}\left(\frac{b}{a}\right)^k p_1 = \begin{cases} \frac{1-(b/a)^{i+1}}{1-b/a} p_1 & \text{if } a \neq b \\ (i+1)p_1 & \text{if } a = b \end{cases} \tag{2}$$

Since $p_n = 1$ we have the following from (2)

$$1 = p_n = \begin{cases} \frac{1-(b/a)^n}{1-b/a} p_1 & \text{if } a \neq b \\ np_1 & \text{if } a = b \end{cases} \text{ thus } \; p_1 = \begin{cases} \frac{1-b/a}{1-(b/a)^n} & \text{if } a \neq b \\ 1/n & \text{if } a = b, \end{cases}$$

inserting the expression for p_1 into (1) yields the result. □