Topic 12A: Probability vectors and stochastic matrices

Def^n A vector $\underset{\sim}{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n$ is a probability vector if $0 \leq x_i \leq 1$ for all $1 \leq i \leq n$, and

$$x_1 + x_2 + \cdots + x_n = 1.$$

Examples

Some probability vectors:

$$\begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}, \quad \begin{bmatrix} \frac{1}{3} \\ \frac{1}{3} \\ \frac{1}{3} \end{bmatrix}, \quad \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

$$\begin{bmatrix} 1/3 \\ 2/3 \end{bmatrix}, \quad \begin{bmatrix} 0.2 \\ 0.8 \end{bmatrix}$$

The following are not probability vectors:

$$\begin{bmatrix} -0.5 \\ 1.5 \end{bmatrix}, \quad \begin{bmatrix} 1.5 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Def^n An $n \times n$ matrix P is a stochastic matrix if all its columns are probability vectors.

Note For each entry p_{ij} of P, if P is stochastic then $0 \le p_{ij} \le 1$.

Examples
The following are stochastic matrices:

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 0.3 & 0.4 \\ 0.7 & 0.6 \end{bmatrix}, \quad \begin{bmatrix} 0.5 & 1 & 0 \\ 0.25 & 0 & 1 \\ 0.25 & 0 & 0 \end{bmatrix}.$$

The following are not stochastic:

$$\begin{bmatrix} 0.8 & 0.5 \\ \underbrace{0.1}_{\text{not a probability vector}} & 0.5 \end{bmatrix}, \quad \begin{bmatrix} 0.7 & -0.1 \\ 0.3 & \underbrace{1.1}_{\text{not a probability vector}} \end{bmatrix}$$

Def^n An $m \times n$ matrix A is positive if all its entries satisfy $a_{ij} > 0$.

Def^n A stochastic matrix P is regular if P^m is positive, for some $m \ge 1$.

Examples

1.
$$P = \begin{bmatrix} 0.5 & 0 & 1 \\ 0.25 & 0 & 0 \\ 0.25 & 1 & 0 \end{bmatrix}.$$

P is stochastic, but not positive as some entries are 0.

$$P^2 = \begin{bmatrix} 0.5 & 1 & 0.5 \\ 0.125 & 0 & 0.25 \\ 0.375 & 0 & 0.25 \end{bmatrix}$$

check!

P^2 is stochastic but not positive.

$$P^3 = \begin{bmatrix} 0.625 & 0.5 & 0.5 \\ 0.125 & 0.25 & 0.125 \\ 0.25 & 0.25 & 0.375 \end{bmatrix}$$

check!

P^3 is stochastic and positive.

Since P^3 is positive, the matrix P is regular.

2.
$$Q = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

Q is stochastic, but not positive.

$$Q^2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2$$

I_2 is stochastic, but not positive.

$$Q^3 = Q^2 Q = I_2 Q = Q$$
$$Q^4 = (Q^2)^2 = (I_2)^2 = I_2$$

$$Q^m = \begin{cases} Q & \text{if } m \text{ is odd} \\ I & \text{if } m \text{ is even} \end{cases}$$

Since Q^m is never positive, the matrix Q is <u>not</u> regular.

<u>Theorem</u> Let P be an $n \times n$ matrix. Then:

(1) if P is stochastic, then P^m is stochastic

(2) if P is positive, then P^m is positive.

<u>Proof</u> See tutorials for (1).
For (2), exercise. $\square$

//

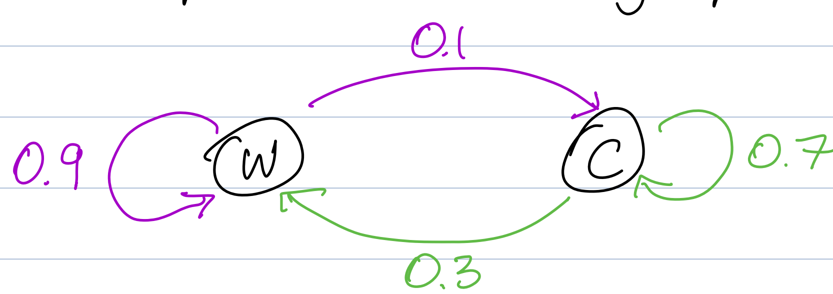
MATH1002 Linear Algebra

Topic 12B: Markov chains

Example Suppose there are 2 chains of grocery stores, W and C.
Each month:

- 90% of people who shop at W, continue to shop at W in the next month
- 10% of people who shop at W will switch to C
- 70% of people who shop at C will continue to shop at C
- 30% of people who shop at C will switch to W.

We can represent this graphically:



W, C are called states.

(we converted percentages to probabilities here).

Suppose that initially 100 people shop at W and 80 people shop at C.

How many people shop at each of W and C in the next month?

No. of people in state W in next month $= 0.9(100) + 0.3(80) = 90+24$
$= 114.$

No. of people in state C in next month $= 0.1(100) + 0.7(80) = 10+56$
$= 66.$

i.e.

$$0.9(100) + 0.3(80) = 114$$
$$0.1(100) + 0.7(80) = 66$$

i.e.

$$\begin{bmatrix} 0.9 & 0.3 \\ 0.1 & 0.7 \end{bmatrix} \begin{bmatrix} 100 \\ 80 \end{bmatrix} = \begin{bmatrix} 114 \\ 66 \end{bmatrix}$$

So if $\underset{\sim}{x}_0 = \begin{bmatrix} 100 \\ 80 \end{bmatrix}$ is the vector encoding the initial states, then

$\underset{\sim}{x}_1 = P\underset{\sim}{x}_0$ ~~for~~ $P = \begin{bmatrix} 0.9 & 0.3 \\ 0.1 & 0.7 \end{bmatrix}$

is the vector encoding the situation after 1 month.

Def^n A Markov chain is a process which changes over time. It consists of:

- a finite number of states
- at each step or point in time, the process can be in any one of the states
- at the next step, the process can stay in the same state or can switch to a different state
- the probability of moving from state i to state j is constant ie. at any step, the probability of moving from state i to state j doesn't depend on anything except for the current state i not on past history.

Notation

We use $1, 2, \ldots, n$ for the states.
We write $\underset{\sim}{x}_k \in \mathbb{R}^n$ for the vector whose i^{th} entry is the number of people/objects/animals etc in state i after k steps.

So $\underset{\sim}{x}_0$ is the initial data, called the initial state vector.

$\underset{\sim}{x}_1$ is the data after 1 step.
$\underset{\sim}{x}_2$ " " " " 2 steps
etc

We write P_{ij} for the probability of moving from state j to state i.
(note $0 \leq P_{ij} \leq 1$.)
We call P_{ij} the transition probability.

The transition matrix is the $n \times n$ matrix

$$P = [P_{ij}] = \begin{bmatrix} P_{11} & P_{12} & \cdots & P_{1n} \\ P_{21} & P_{22} & \cdots & P_{2n} \\ \vdots & \vdots & & \vdots \\ P_{n1} & P_{n2} & \cdots & P_{nn} \end{bmatrix}$$

This matrix has entries $0 \leq P_{ij} \leq 1$.
Add down column 1

$$\underbrace{P_{11}}_{\substack{\text{probability} \\ \text{of } (1) \circlearrowleft \\ \text{i.e. staying} \\ \text{at state } (1)}} + \underbrace{P_{21}}_{\substack{\text{probability} \\ \text{of} \\ (1) \to (2)}} + \cdots + \underbrace{P_{n1}}_{\substack{\text{probability} \\ \text{of} \\ (1) \to (n)}} = 1.$$

Similarly, any column's entries add up to 1.

So each column of P is a probability vector.

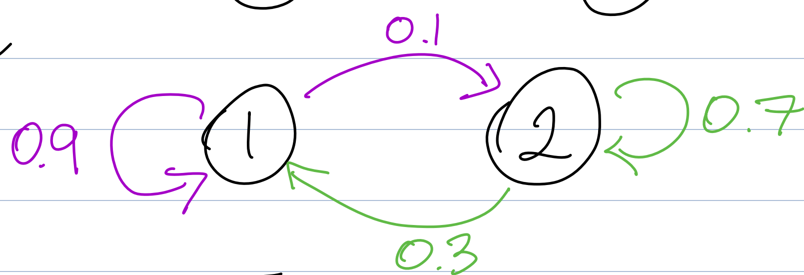
Hence P is a stochastic matrix.

<u>Example</u> In the previous example,

let state 1 be W

" 2 " C

Then



and $\underset{\sim}{x}_0 = \begin{bmatrix} 100 \\ 80 \end{bmatrix}.$

The transition matrix is:

$$P = \begin{bmatrix} 0.9 & 0.3 \\ 0.1 & 0.7 \end{bmatrix}$$

P_{12} = prob. of $2 \to 1$ $= 0.3$

Thus after k months

$$\underset{\sim}{x}_k = P^k \underset{\sim}{x}_0 .$$

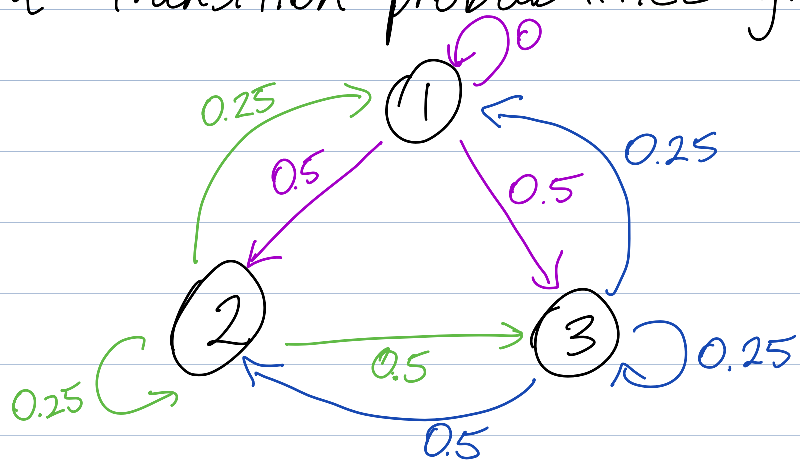
//

Topic 12C: Steady-state vectors

Example Suppose the weather in Sydney has 3 states:

(1) sunny
(2) cloudy
(3) rainy

with transition probabilities given by:



This has transition matrix

$$P = \begin{array}{c} \begin{array}{ccc} \text{State 1} & \text{State 2} & \text{State 3} \end{array} \\ \begin{bmatrix} 0 & 0.25 & 0.25 \\ 0.5 & 0.25 & 0.5 \\ 0.5 & 0.5 & 0.25 \end{bmatrix} \end{array}$$

Column j is probabilities on arrows starting at state j

Row i is probabilities on arrows ending at state i.

Suppose the weather is rainy (state 3) on day 0. So we have initial state vector

$$\underset{\sim}{x}_0 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

(This is a probability vector.)

What are the weather probabilities on day 2?

$$\underset{\sim}{x}_2 = P^2 \underset{\sim}{x}_0$$

$$= \begin{bmatrix} 0.1875 \\ 0.375 \\ 0.4375 \end{bmatrix}$$

So on day 2, the probability of

- sunny weather is 0.1875
- cloudy " " 0.375
- rainy " " 0.4375

What about in 7 days time?

$$\underset{\sim}{x}_7 = P^7 \underset{\sim}{x}_0 = \begin{bmatrix} 0.2000122070 \\ 0.4000244140 \\ 0.3999633789 \end{bmatrix}$$

In 15 days time?

$$\underset{\sim}{x}_{15} = P^{15}\underset{\sim}{x}_0 = \begin{bmatrix} 0.2000000002 \\ 0.4000000003 \\ 0.39999999994 \end{bmatrix}$$

It seems like $\underset{\sim}{x}_k \longrightarrow \begin{bmatrix} 0.2 \\ 0.4 \\ 0.4 \end{bmatrix}$ as $k \to \infty$.

Now

$$P\begin{bmatrix} 0.2 \\ 0.4 \\ 0.4 \end{bmatrix} = \begin{bmatrix} 0 & 0.25 & 0.25 \\ 0.5 & 0.25 & 0.5 \\ 0.5 & 0.5 & 0.25 \end{bmatrix}\begin{bmatrix} 0.2 \\ 0.4 \\ 0.4 \end{bmatrix} = \begin{bmatrix} 0.2 \\ 0.4 \\ 0.4 \end{bmatrix}.$$

i.e. $P\underset{\sim}{x} = \underset{\sim}{x}$ for $\underset{\sim}{x} = \begin{bmatrix} 0.2 \\ 0.4 \\ 0.4 \end{bmatrix}$.

i.e. $P\underset{\sim}{x} = 1\underset{\sim}{x}$

i.e. $\underset{\sim}{x}$ is an eigenvector of P with eigenvalue 1.

Def^n Let P be the transition matrix of a Markov chain. A steady-state vector is any vector $\underset{\sim}{x}$ so that

$$P\underset{\sim}{x} = \underset{\sim}{x},$$

with non-negative entries summing to the total number of objects/people/animals etc in the Markov chain.

A steady-state probability vector is a probability vector $\underset{\sim}{x}$ so that
$$P\underset{\sim}{x} = \underset{\sim}{x}.$$

Example Above, $\underset{\sim}{x} = \begin{bmatrix} 0.2 \\ 0.4 \\ 0.4 \end{bmatrix}$ is a steady-state probability vector.

Note If $\underset{\sim}{x}$ is a steady-state vector or steady-state probability vector then

$$P\underset{\sim}{x} = \underset{\sim}{x}$$

so $P^2\underset{\sim}{x} = P\underset{\sim}{x} = \underset{\sim}{x}$

so $P^3\underset{\sim}{x} = P(P^2\underset{\sim}{x}) = P\underset{\sim}{x} = \underset{\sim}{x}$ etc

Thus

$$P^k\underset{\sim}{x} = \underset{\sim}{x} \quad \text{for all } k \geq 1.$$

Notation We sometimes write

SSV for steady-state vector

SSPV " " " probability vector.

Theorem If P is the transition matrix for a Markov chain, then 1 is an eigenvalue of P.

Corollary P has an 1-eigenvector $\underset{\sim}{y}$
ie, $P\underset{\sim}{y} = \underset{\sim}{y}$ and $\underset{\sim}{y} \neq \underset{\sim}{0}$.
We can rescale $\underset{\sim}{y}$ (ie multiply by some scalar) to get either a SSV or a SSPV.

Proof of Theorem
P is a stochastic matrix, so each column of P adds up to 1.
Let $\underset{\sim}{r} = \begin{bmatrix} 1 & 1 & \cdots & 1 \end{bmatrix}_{1\times n}$ be row vector with n 1s.
Then
$$\underset{\sim}{r} P = \begin{bmatrix} 1 & 1 & \cdots & 1 \end{bmatrix} \begin{bmatrix} P_{11} & P_{12} & \cdots & P_{1n} \\ P_{21} & P_{22} & & \vdots \\ \vdots & & \ddots & \vdots \\ P_{n1} & \cdots & \cdots & P_{nn} \end{bmatrix}$$
$$= \begin{bmatrix} P_{11}+P_{21}+\cdots+P_{n1} & \cdots & P_{1n}+P_{2n}+\cdots+P_{nn} \end{bmatrix}$$
ie. entries are column sums of P
$$= \begin{bmatrix} 1 & 1 & \cdots & 1 \end{bmatrix}$$
$$= \underset{\sim}{r}.$$
Thus
$$(\underset{\sim}{r} P)^T = \underset{\sim}{r}^T$$
so
$$P^T \underset{\sim}{r}^T = \underset{\sim}{r}^T$$
ie. $\underset{\sim}{r}^T$ is an eigenvector for P^T with corresp. eigenvalue 1.

Now P^T and P have the same eigenvalues. $\square$

To find a SSV or SSPV: Solve $[P-1I|\underset{\sim}{0}]$ then find a SSV or SSPV among the solutions, by taking a suitable scalar multiple.

Example Consider Markov chain with transition matrix

$$P = \begin{bmatrix} 0.9 & 0.3 \\ 0.1 & 0.7 \end{bmatrix}$$

Find the 1-eigenspace of P:

$$[P-I|\underset{\sim}{0}] = \left[\begin{array}{cc|c} -0.1 & 0.3 & 0 \\ 0.1 & -0.3 & 0 \end{array}\right]$$

$$\xrightarrow{R_2 \mapsto R_2+R_1} \left[\begin{array}{cc|c} -0.1 & 0.3 & 0 \\ 0 & 0 & 0 \end{array}\right]$$

Put $y=t$ then

$$-0.1x + 0.3t = 0$$

so $x = 3t$.

Thus

$$E_1 = \left\{ t\begin{bmatrix} 3 \\ 1 \end{bmatrix} : t \in \mathbb{R} \right\}.$$

To find a SSV: total population is 180 people, so we need a vector in E_1 with non-negative entries adding up to 180.

Have $\begin{bmatrix} 3t \\ t \end{bmatrix} \in E_1$. Want $3t + t = 180$.

$$\Rightarrow 4t = 180$$
$$\Rightarrow t = 45.$$

Thus a SSV is

$$\underset{\sim}{x} = \begin{bmatrix} 135 \\ 45 \end{bmatrix}.$$

Check: $P\underset{\sim}{x} = \underset{\sim}{x}$.

To find a SSPV: want a vector in E_1 which is a probability vector.

Solve

$$3t + t = 1$$
$$\Rightarrow \quad 4t = 1$$
$$t = 0.25$$

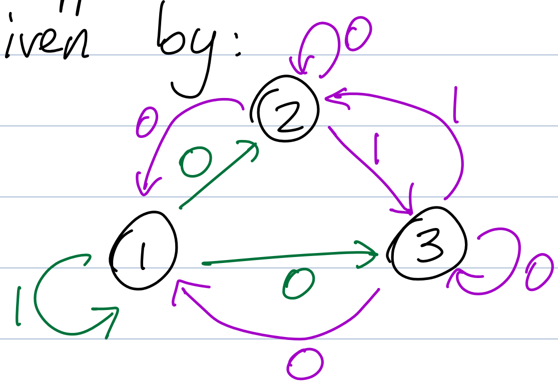
A SSPV is

$$\begin{bmatrix} 0.75 \\ 0.25 \end{bmatrix}.$$

//

Topic 12D: Regular Markov chains

Example Suppose we have a Markov chain given by:



for a population of 600 people.
Find SSVs and SSPVs.
i.e. vectors $\underset{\sim}{x}$ so that $P\underset{\sim}{x} = \underset{\sim}{x}$
for P the transition matrix
SSV: entries sum to 600
SSPV: " " " 1.

Transition matrix:

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

Find 1-eigenspace for P:

$$[P - I \mid \underset{\sim}{0}] = \left[\begin{array}{ccc|c} 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 1 & -1 & 0 \end{array}\right]$$

$$\xrightarrow{R_3 \mapsto R_3 + R_2} \left[\begin{array}{ccc|c} 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Put $z = t$ then $-y + t = 0$
so $y = t$.

Then $x = s$.

$$E_1 = \left\{ \begin{bmatrix} s \\ t \\ t \end{bmatrix} : s, t \in \mathbb{R} \right\}$$

$$= \left\{ s \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} : s, t \in \mathbb{R} \right\}$$

Now to find SSVs and SSPVs:

Two SSVs are: $\begin{bmatrix} 600 \\ 0 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 300 \\ 300 \end{bmatrix}$

(put $s = 600$, $t = 0$) (put $s = 0$, $t = 300$)

Two SSPVs are: $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 0.5 \\ 0.5 \end{bmatrix}$

(put $s = 1$, $t = 0$) (put $s = 0$ $t = 0.5$)

Here, we do not have unique SSVs (or SSPVs).

If the initial state vector is

$$\underset{\sim}{x}_0 = \begin{bmatrix} 600 \\ 0 \\ 0 \end{bmatrix}, \text{ then } \underset{\sim}{x}_k = \begin{bmatrix} 600 \\ 0 \\ 0 \end{bmatrix}.$$

and if $\underset{\sim}{x}_0 = \begin{bmatrix} 0 \\ 300 \\ 300 \end{bmatrix}$, then $\underset{\sim}{x}_k = \begin{bmatrix} 0 \\ 300 \\ 300 \end{bmatrix}$.

Recall: A stochastic matrix P is regular if P^k is positive for some $k \geq 1$.
i.e. all entries of P^k are > 0.

Def^n A Markov chain is regular if its transition matrix is regular.

Exercise The Markov chain above is not regular.

Theorem (Properties of regular Markov chains).
Let P be the transition matrix for a Markov chain. If P is regular:
(1) there is a unique SSV $\underset{\sim}{x}$.
(2) " " " " SSPV $\underset{\sim}{y}$.
(3) for every possible initial state vector $\underset{\sim}{x}_0$, we have $\underset{\sim}{x}_k \longrightarrow \underset{\sim}{x}$ the unique SSV as $k \to \infty$.

(Here, $\underset{\sim}{x}_k = P^k \underset{\sim}{x}_0$.)

(4) As $k \to \infty$,

$$P^k \longrightarrow L = \begin{bmatrix} | & | & & | \\ y & y & \cdots & y \\ | & | & & | \end{bmatrix}$$

where y is the unique SSPV.
We call L the long range transition matrix.

Proof Uses:

Theorem For any $n \times n$ transition matrix P, with eigenvalue λ

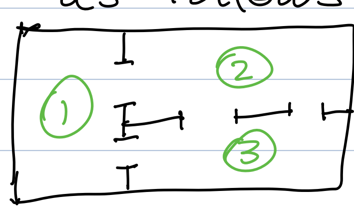
(1) $|\lambda| \leq 1$ ie. $-1 \leq \lambda \leq 1$

(2) if P is regular, then -1 is not an eigenvalue, so $|\lambda| < 1$.

Proof not an exercise.

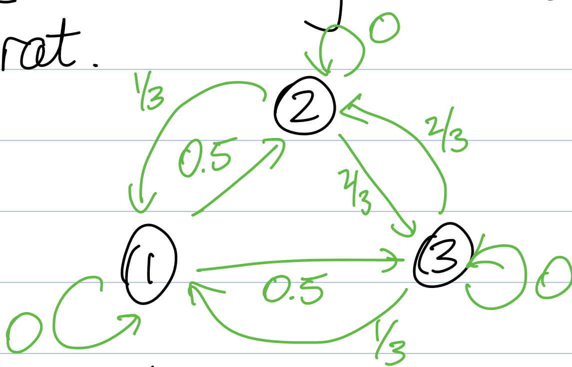
We would need to discuss convergence of matrices and vectors to give a rigourous proof of the main result. $\square$

Example A psychologist places a rat in a cage designed as follows:



The rat has been trained so that each time it hears a bell, it changes compartment, and is equally likely to choose each door.

Describe the long run behaviour of the rat.



Transition matrix:

$$P = \begin{bmatrix} 0 & 1/3 & 1/3 \\ 0.5 & 0 & 2/3 \\ 0.5 & 2/3 & 0 \end{bmatrix}$$ not positive

However P^2 is positive (check!) so P is regular.

To find SSV and SSPV:

$$[P - I \mid \underset{\sim}{0}] = \left[\begin{array}{ccc|c} -1 & 1/3 & 1/3 & 0 \\ 1/2 & -1 & 2/3 & 0 \\ 1/2 & 2/3 & -1 & 0 \end{array}\right]$$

$$\xrightarrow{\cdots} \left[\begin{array}{ccc|c} -3 & 1 & 1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Put $z = t$, then $-y + t = 0$ so $y = t$

Now $-3x + t + t = 0$

so $x = \frac{2}{3}t$.

$$E_1 = \left\{ t\begin{bmatrix} 2/3 \\ 1 \\ 1 \end{bmatrix} : t \in \mathbb{R} \right\}.$$

$$= \left\{ t\begin{bmatrix} 2 \\ 3 \\ 3 \end{bmatrix} : t \in \mathbb{R} \right\}$$

SSV: entries add to 1

$$2t + 3t + 3t = 1$$
$$8t = 1$$
$$t = \frac{1}{8}.$$

SSV: $\frac{1}{8}\begin{bmatrix} 2 \\ 3 \\ 3 \end{bmatrix} = \begin{bmatrix} 1/4 \\ 3/8 \\ 3/8 \end{bmatrix}$

this is also a SSPV.

In the long term, the rat is in
compartment 1 with probability 1/4
" 2 " " 3/8
" 3 " " 3/8.

//