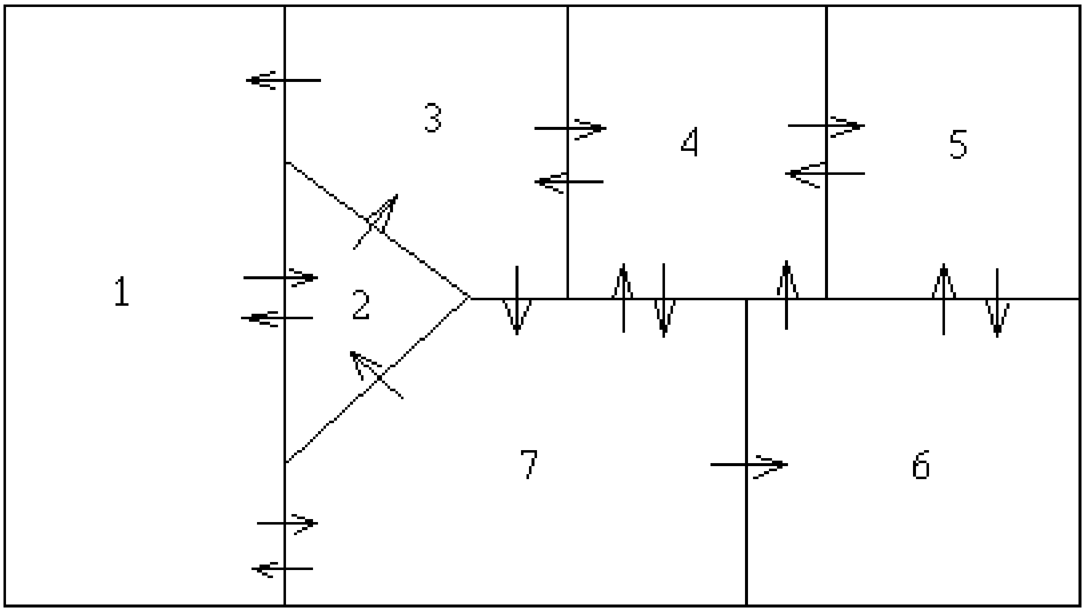
Math 235 **Summer 2001**

Project VII

Note: In this project you will find some hints on how you can use Mathematica. But you are welcome to use any other computer softwear if you prefer.

Consider the diagram of rooms below where arrows indicate doors between rooms which can be used only in the direction indicated; for example it is possible to go from room 1 to rooms 2 and 7, but no other rooms.



Suppose there was a cat in room 1 and every 3 minutes it considers wandering out one of the doors from its current room, to another room. Suppose that the probability of which door it goes through is independent of the door, and equal to the probability that it stays where it is. For example, if the cat is in room 1, it will stay where it is, go through the door to room 2, or go through the door to room 7 each with probability $\frac{1}{3}$ (there are 3 possibilities, and the sum of the probabilities must be equal to 1).

We wish to answer the following questions:

- What is the probability that a cat starting in room j will be in room i after k steps?
- Suppose that we put 100 cats in the rooms with some initial distribution. What will the distribution of cats be after a long time (after a large number of steps)?
- If a cat starts in room j and we observe a large number of steps, what proportion of time do we expect the cat to stay in room i?

These seem very difficult questions to answer, however with the aid of linear algebra they are relatively simple.

We start by creating a 7×7 (number of rooms) transition matrix P whose ij component p_{ij} is the probability that a cat in room j will move to room i. So for example, p_{21} is the probability that a cat in room 1 will move to room 2 which in our example is $\frac{1}{3}$. The entries p_{ii} are the probabilities that the cat will stay in room i.

1. Write the matrix P.

2. The probability that the cat moves from room j to toom i in 2 steps is the sum (over l) of the probabilities it moves from room j to room l and then from room l to room i, i.e.

$$p_{i1}p_{1j} + p_{i2}p_{2j} + p_{i3}p_{3j} + p_{i4}p_{4j} + p_{i5}p_{5j} + p_{i6}p_{6j} + p_{i7}p_{7j} = \sum_{l=1}^{7} p_{il}p_{lj}$$

   (a) Notice that this is just the ij component of the matrix P^2 and that it follows by an argument of induction that the probability of going from room j to room i in k steps is the ij component of P^k.

   (b) Compute P^4 and explain the meaning of the elements p_{33}, p_{35}, p_{53}.

   (c) Let $v = (0,0,0,0,1,0,0)^t$. Compute and explain the meaning of the vector $u = P^4 v$.

3. Let $v = (v_1, v_2, \ldots, v_7)^t$ be the initial distribution of cats (these must thus sum to 100). To compute how many cats we expect to be in room i after 1 step we sum (over j) the probability of moving from room j to room i times the number of cats currently in room j (i.e. the number of cats in j which will move from j to i), i.e.

$$\sum_{j=1}^{7} p_{ij}v_j$$

   (a) Notice that this is just the i^{th} component of the vector Pv. Again, induction shows that the i^{th} component of $P^k v$ is the expected number of cats in room i after k steps given the initial distribution v.

   (b) Choose four different v's and compute P^{10} for each choice.

   (c) Now compute P^{40} for each of your choices of v, and compare the answers. Comment on your observation.

   (d) Let $w = P^{40}v$ for one choice of v. Compute Pw, and explain the meaning of the vectors w and Pw, look at the difference $w - Pw$ and comment on what you see. What would happen if you let the system continue for 1000 more steps? Interpret the meaning of your answers in terms of the cats and rooms.

4. Let's assume that the cat starts in room 1. and let $v = (1,0,0,0,0,0,0)^t$. We know that the probability that the cat will be in room j after n steps is the j^{th} entry in $P^n v$. Then after n steps, the fraction of time the cat spends in room j is $\frac{1}{n}$ times the sum of the probabilities it will be in room j after each of 1, 2, 3, $\ldots$, n steps, that is

$$\frac{1}{n}(P^1 v + P^2 v + P^3 v + \ldots + P^n v) \qquad (*)$$

.

   (a) Compute the value of the above experssion $(*)$ for $n = 50$. You can use Mathematica to do this:

```
Clear[v,s];
v={ 1,0,0,0,0,0,0 } ;
s={ 0,0,0,0,0,0,0 } ;
Do[ s=s+MatrixPower[P,i].v, { i, 50 } ]
s/50 //N
```

   Notice how long it takes!

The reason it takes so long is that it has to keep computing P to a higher and higher power - at the 45^{th} step we compute P^{45} and then at the 46^{th} step we compute P^{46} which is the same as P^{45} times P. It would be more efficient to keep track of the power of P and at each stage just multiply the last result by P. Here is a program which will do this:

```
Clear[v,s,R];
v={ 1,0,0,0,0,0,0 } ;
s={ 0,0,0,0,0,0,0 } ;
R=IdentityMatrix[7];
Do[ R=P.R; s=s+R.v, { i, 50 } ]
s/50 //N
```

(Here R is successively multiplied by P and so is keeping track of higher and higher powers of P.)

(b) Compute the value of $(*)$ for $n = 150$. What proportion of the time does the cat spend in room 3 if it starts in room 1 and we watch for 150 steps? In which room is the cat most often? The least often?

This is what is required to get the exact answer for 150 steps. But if our question is more "after a *long* time, what proportion of the time does the cat spend in room 3?" then we can be more sophisticated. (We could just increase n but it will take longer and longer to compute.)

(c) Compute $w = P^{150}v$:

```
Clear[w];
w=MatrixPower[P,150].v //N
```

then compute Pw and compare the answers. You will find that $Pw = w$. This says that w is an eigenvector - what is the corresponding eigenvalue?

It is a fact that a Morkov matrix always has this as an eigenvalue, and all other eigenvalues are of modulus less than this.

(d) Check this claim by computing eigenvalues of P:

```
Eigenvalues[P] //N
```

and then computing the modulus of each eigenvalue.

From above, we have $P^n v = w$, at least approximately, for all $n > 150$. If we write out

$$\frac{1}{L}(P^1 v + P^2 v + \ldots + P^{150} v + P^{151} v + \ldots + P^L v)$$

for very large L (i.e. a long long time), then most of the terms will be $P^n = w$ (from $n = 150$ on). So if there are enough of these terms they will dominate the average, and the answer will be a lot like

$$\frac{1}{L}(L \cdot w) = w.$$

Thus the components of w give us the proportion of time the cat spends in each room, over a very long time. How long? Look at the answer you got for 150 steps and compare it to w - they are not too different.

(e) What proportion of the time does the cat spend in room 3 if it starts in room 1 and we watch for a very long time? In which room is the cat most often? The least often?