9 - Markov processes and Burt & Allison 1963

AGEC 642 - 2024

I. What is a Markov Chain?

A Markov chain is a discrete-time stochastic process in which the probability of moving from state x_t^i to x_{t+1}^j is written p_{ij} with $\sum_j p_{ij} = 1$. If the transition probabilities do not change over time, then the dynamic system is said to be stationary. A Markov transition matrix, P, for a system with 3 states, x^1, x^2 and x^3, therefore, might take the form:

$x_t \backslash x_{t+1}$	**1**	**2**	**3**
1	0.5	0.1	0.4
2	0.2	0.2	0.6
3	0.0	0.2	0.8

(1)

In this case, the probability of moving from state x^2 to x^1 in one period is 0.2, the probability of staying in state x^2 is 0.2, and the probability of moving from state x^2 to x^3 is 0.6.

Note that adding up the elements of any the i^{th} row gives the total probability of going to all states given that you have started in the i^{th} state. Since the set of states is comprehensive (i.e., it is a list of all the possible states) each row must sum to 1.

Now let's consider the probability of moving from state 1 to state 2 in two periods. This would be equal to

$$\underbrace{p_{11}\cdot p_{12}}_{\text{probability of going from 1 to 1 and then 1 to 2}} + \underbrace{p_{12}\cdot p_{22}}_{\text{probability of going from 1 to 2 and then 2 to 2}} + \underbrace{p_{13}\cdot p_{32}}_{\text{probability of going from 1 to 3 and then 3 to 2}} = 0.5\cdot 0.1 \;+\; 0.1\cdot 0.2 \;+\; 0.4\cdot 0.2 = 0.15,$$

i.e., the sum of the probability across all possible paths from 1 to 2 in two steps. Note that this equation is the 1st row times the 2nd column. Similarly, the probability of moving from state 2 to state 3 in two periods is equal to the 2nd row times the 3rd column. This is matrix multiplication. Hence, the complete probability transition matrix of moving from state i to state j in two periods can be found in the matrix $P^2=P\cdot P$, or

$x_t \backslash x_{t+2}$	**1**	**2**	**3**
1	0.27	0.15	0.58
2	0.14	0.18	0.68
3	0.04	0.20	0.76

(2)

In general, P^n is the n-step probability transition matrix with elements p_{ij}^n, each of which defines the probability of being in state j after n periods given that you started in state i. P^n is found by multiplying P by itself, n times.

In many cases, there exist a *limiting probability distribution*, defined as the Markov matrix $P^{\infty} = \lim_{n\to\infty} P^n$. If this limit exists, then the i,j^{th} element tells us the probability of being in state j in the distant future given that the sequence starts in state i.

Using the one-period Markov transition matrix defined above, we see that

	P^4 - After 4 periods		
$x_t \backslash x_{t+4}$	**1**	**2**	**3**
1	0.117	0.184	0.699
2	0.090	0.189	0.720
3	0.069	0.194	0.737

	P^{10} - After 10 periods		
$x_t \backslash x_{t+10}$	**1**	**2**	**3**
1	0.077	0.192	0.730
2	0.077	0.192	0.731
3	0.077	0.192	0.731

We see that after 4 iterations, P^4, the rows are beginning to converge; the chance of being in state 2 after 4 periods is between 18.4% and 19.4% depending on the state in which you start. After 10 periods, P^{10}, the rows are equal to 3 digits. Beyond 10 periods, the probability of being in each of the three states is the essentially same, regardless of the state in which we start. Looking out beyond period 10, there is a 19.2% chance that the variable will be in state 2, regardless of the state in which it started. Identifying the limiting probability distribution can be helpful for presenting the results of stochastic systems.

The central characteristic of P^{∞} is that if you multiply it by P, it does not change, i.e., $P^{\infty} = P^{\infty} \cdot P$.[1] Hence, using matrix algebra we find that $P^{\infty}(I-P)=0$. If a limiting probability distribution exists, each row of P^{∞} will be identical. Hence, if p^{∞} is one row of P^{∞}, then $p^{\infty}(I-P)=0$ defines a nonhomogeneous system of linear equations where the elements of p^{∞} are the unknowns. Alternatively, one can approximate P^{∞} by recursively multiplying P by itself many times. That is, starting with n=1, create $P_{n+1} = P_n \cdot P_n$ with P_1=P. Repeat this matrix multiplication process until the product does not change any more, i.e. $P_{n+1} - P_n \approx 0$.

Notice that the probabilities of the three rows in the matrix on the right above are essentially the same. When the limiting probability distribution has been found, these rows will be identical. Hence, the limiting probability distribution can be expressed as a vector, [0.077, 0.0192, 0.731], with each element indicating the probability of being in each state in the distant future.

Why might the limiting probability distribution not exist?
Suppose that the single-stage MTM is

$x_t \backslash x_{t+1}$	**1**	**2**	**3**
1	0	1	0
2	0	0	1
3	1	0	0

(3)

[1] It also follows, of course, that $P^{\infty} \cdot P^{\infty} = P^{\infty}$, so P^{∞} is an *idempotent* matrix.

In this case from state 1 we always move to state 2, from 2 to 3 and from 3 back to 1. If you start in state 1, then after 2 periods, you will be in state 3. At any time *t*, the probability is always either zero or one, and if it is one, the next period the probability is zero. There is no limiting probability distribution for systems that exhibit such cyclical behavior.

Periodicity: The periodicity of a state is the number of periods that one has to wait until there is a chance of returning to that state. For example, the periodicity of all states in matrix (1) is one, because there's a chance that all states will be revisited in one period while the periodicity of all states in matrix (3) is three because it will always take 3 periods to return to each state.

Ergodic or irreducible Markov Chains: A Markov chain is *irreducible* or *ergodic* if it is possible to move from all states to every other state. "A Markov chain is said to be ergodic if there exists a positive integer T_0 such that for all pairs of states *i,j* in the Markov chain, if it is started at time 0 in state *i* for all $t>T_0$, the probability of being in state *j* at time *t* is greater than 0."[2] Here are three matrices that are <u>*not*</u> irreducible:

0.2	0.8	0	0
0.6	.04	0	0
0	0	0.2	0.8
0	0	0.6	.04

0.5	0.5
0	1.0

0	0	0.2	0.8
0	0	0.6	.04
0.2	0.8	0	0
0.6	.04	0	0

In the first case, states 1 & 2 and states 3 & 4 can be reduced into two separate ergodic Markov transition matrices. If you start in state 1, you will never reach state 3, so it is not ergodic.
In the second case, state #2 is an <u>*absorbing state*</u> – once it is reached, it will stay there forever, so as $t\rightarrow\infty$ the probability of being in state 1 is zero.
In the third case the process is oscillating, if you start in state 1, you will be in states 3 or 4 in period 2, states 1 or 2 in period 3, 2 or 4 in period 4, and so on. So, it does not hold that for *all t* probability of being in any state is greater than zero.

Higher order Markov processes: In some applications, it is important to take into account "memory," i.e. how you got to a given state. Consider, for example, the question of adopting a farm management process -- 0 or 1. The probability of being in state 1 may depend not only on the state you were in last period, but on the state in the previous period. So, you may have a 2nd order Markov transition process like the following.

State in $t-2$ & $t-1$	$t-2$	$t-1$	*State in t*: 0	1
	0	0	0.9	0.1
	1	0	0.7	0.3
	0	1	0.3	0.7
	1	1	0.1	0.9

In this case, if you haven't adopted the practice in the previous two periods, there's only a 10% chance that you'll adopt it in *t*, while if you've adopted the practice in the previous

[2] nlp.stanford.edu/IR-book/html/htmledition/definition-1.html

two periods, there's a 90% chance that you'll stick with it. This is different from the situation in which you've been switching back and forth between adoption and non-adoption, in which case there's a 70% chance that you'll stick with the practice that you used last period.

Markov processes of any order can be established, but in the DP solved in this class we will use only first-order processes, in which case the number of rows and number of columns are identical.

Although there are other interesting properties of Markov transition matrices, this will be sufficient for the analysis required here.

II. Why look at Burt & Allison 1963?

It is worth asking why we should even bother to look at a 1963 paper on computational methods. Certainly, given the incredible advances that have been made in both hardware and software, papers that discuss computational work 50 years old are of questionable value. Nonetheless, I can think of three basic reasons why looking at this paper is useful:

- The paper is a clear and simplistic formulation of a stochastic dynamic programming problem, making it ideal for pedagogical purposes;
- The paper presents a simple agricultural application that can easily be built on for more elaborate analysis;
- It was a watershed paper in Agricultural Economics and much of the dynamic programming in the field can be traced back to Oscar Burt and this paper.

III. Burt and Allison's farm management problem

The applied problem that the authors consider is the optimal fallow-planting decision by a farmer who faces stochastically changing soil moisture levels. Early in the paper the authors mistakenly call the previous period's decision the state variable, but the correct state variable is referred to in Table 1: the soil moisture content. The choice variable is whether to fallow the field (z_t=F) or to plant it with wheat (z_t=W).

Payoffs (i.e. the benefit function, $R(x_t, z_t)$) are dependent on the state and the decision,

State	**Soil moisture level (x_t)**	**Net return under Fallow $\pi(z_t=F)$**	**Net return with planting $\pi(z_t=W)$**
1	0-2	-2.33	4.52
2	2.1-4	-2.33	32.07
3	4.1-6	-2.33	36.26
4	6.1-8	-2.33	36.78
5	8.1 or more	-2.33	47.63

Depending on whether the decision is made to fallow or harvest, two distinct Markov transition matrices can be found in Table 1:

Choice-contingent Markov Transition Matrices

MTM - Fallow $P(F)$

	1	2	3	4	5
1	0	1/20	5/20	7/20	7/20
2	0	0	1/20	5/20	14/20
3	0	0	0	1/20	19/20
4	0	0	0	0	1
5	0	0	0	0	1

MTM - Wheat $P(W)$

	1	2	3	4	5
1	9/23	7/23	7/23	0	0
2	9/23	7/23	7/23	0	0
3	9/23	7/23	7/23	0	0
4	9/23	7/23	7/23	0	0
5	9/23	7/23	7/23	0	0

Given these values, the authors then solve for the optimal policy of an infinite-horizon optimization problem for the farmer seeking to maximize the net present value of his or her net returns using a 6% discount rate, β=1/1.06.

Hence, Burt & Allison's objective function would be

$$\max_{z_t=\{F,W\}} \sum_{t=0}^{\infty} \beta^t ER(x_t, z_t)$$

giving rise to the Bellman's equation

$$V(x_t) = \max_{z_t=\{F,W\}} R(x_t, z_t) + \beta E(x_t, z_t) V(x_{t+1})$$

where $E(x_t, z_t)$ is the expectation contingent on the current state and choice. Equivalently we could write the Bellman's equation using the MTMs:

$$V(x_t) = \max_{z_t=\{F,W\}} R(x_t, z_t) + \beta p_{z_t, x_t} V(x_{t+1}),$$

where p_{z_t, x_t} is the row of the choice-contingent MTM associated with current state x_t.

Maximum likelihood estimates of transition probabilities

Where might these MTMs have come from? They must be estimated. Burt and Allison estimate these probabilities using precipitation data, expert judgments, and assumptions. More generally, one can think about the problem of estimating each of the choice contingent transitions probabilities $p_{ij}(z)$ using observations of what has actually occurred.

For example, suppose that the true probabilities are 20% chance of going to state 1 and 80% chance of going to state 2. If one observes 10 independent draws, and 4 of them go to state 1 and 6 go to state 2, then the probability of observing these draws is $0.2^4 \times 0.8^6$. Now suppose that we're trying to find the probabilities that are the best fit, i.e. the probabilities that maximize the likelihood of seeing the draws of 4 and 6. That is, we look for the probabilities p_1 and p_2 that maximize $L = p_1^{\ 4} \times p_2^{\ 6}$. Intuition suggests that the best guess for p_1 and p_2 would be $\frac{4}{10}$ and $\frac{6}{10}$. This is correct, but why?

Consider the general problem. Suppose that, starting at point *i*, a choice *z* has been observed k_i times and the number of times state *j*=1,…,*n* was reached is k_{ij} , so that $k_i = \sum_j k_{ij}$. For a given set of transition probabilities, p_{ij}, the likelihood of observing this data would be

$$L = \prod_j \left(p_{ij} \right)^{k_{ij}} .$$

To find the maximum likelihood probabilities, we maximize *L* (or equivalently here, *ln*(*L*)) subject to the constraint that $\sum_j p_{ij}\left(z\right) = 1$. Using the Lagrangian,

$$\mathcal{L} = \max_{\{p_{ij}\}} \sum_j k_{ij} \ln\left(p_{ij}\right) + \lambda\left(1 - \sum_j p_{ij}\right),$$

yields the FOCs $\dfrac{\partial \mathcal{L}}{\partial p_{ij}} = \dfrac{k_{ij}}{p_{ij}} - \lambda = 0 \;\; \forall j$, from which we find that

$\dfrac{k_{ij}}{p_{ij}} = \dfrac{k_{ik}}{p_{ik}} \;\; \forall j,k$ or $p_{ij} = \dfrac{k_{ij}}{k_{ik}} p_{ik} \;\; \forall j$. Now, using the constraint,

$1 = \sum_j p_{ij} = \sum_j \dfrac{k_{ij}}{k_{ik}} p_{ik} = \dfrac{p_{ik}}{k_{ik}} \sum_j k_{ij}$ so $p_{ik} = \dfrac{k_{ik}}{\sum_j k_{ij}} = \dfrac{k_{ik}}{k_i} \forall k$, i.e., the maximum likelihood estimates of the transition probabilities are simply the relative frequencies, just as we expected.

Hence, maximizing the likelihood function is intuitive and, in some cases, relatively straightforward.

There are a couple of important things to note about this very simplistic example. First, for dynamic problems the transition probabilities are state and choice contingent – here we assume that we observe $k_i(z)$ observations for choice *z*. If some choices are clearly "bad," it is likely that such choices are not likely to be observed very many times if at all. That is, the number of times you observe a state-choice combination is endogenous and that should be taken into account if your data are non-experimental. Second, if k_{ij}=0 for some choice, then the estimated probability p_{ij}=0. Hence, analysts strictly applying this approach would estimate that the probability of reaching a state is zero because it has never been reached in the limited number of observations available. In practice, therefore, the analyst will typically impose more structure on the problem, including possible prior probabilities leading to Bayesian estimation and taking into account that *z* is endogenous.

In any case, before advancing to solving the problem, we will assume that the analyst has obtained the choice-contingent MTMs needed to solve the problem.

IV. The solution algorithm

Burt and Allison employ a two-step solution algorithm that made sense when computers were slow, but I will explain here the *policy iteration* approach that is at the heart of their solution approach.

Step 1: The first step, presented in Table 2 of the paper, is to identify an array of *candidate* optimal choices. They do this by starting out by assuming that the value for any soil moisture content in the next period is zero, i.e., $V^0(x)=0$ for all x. In this case, since the payoffs from wheat are greater than the payoffs from fallowing in all states, it is never optimal to fallow, $z^1(x)=[W \quad W \quad W \quad W \; W]$. We will call this a ***candidate*** *optimal policy vector*. Let R^1 be the first *candidate optimal vector of the returns* at each point in the state space, i.e. $R^1(x)=[4.52 \quad 32.07 \quad 36.26 \quad 36.78 \quad 47.63]$. Finally, we also obtain the <u>*candidate*</u> *optimal Markov transition matrix*, P^1, which is constructed by taking the rows of the Markov transition matrix associated with the vector of optimal policies. Since $z^1(x)=W$ for all x, it is simply equal to MTM associated with planting, $P(W)$:

	1	**2**	**3**	**4**	**5**
1	9/23	7/23	7/23	0	0
2	9/23	7/23	7/23	0	0
3	9/23	7/23	7/23	0	0
4	9/23	7/23	7/23	0	0
5	9/23	7/23	7/23	0	0

Step 2: The next step is to use the candidate optimal policy identified in step 1 to obtain an estimate of the value function. The idea here is to calculate what the value function would be if we followed this policy forever. First, imagine that we followed it for just two periods. In this case, after obtaining the benefits as defined by $R^1(x)$, the transition matrix P^1 would determine the probability of where we ended up next period. Suppose that you start in state 2 at time t=1. Following the policy $z^1(x)$ you would earn 32.07 in the first period, but you wouldn't know with certainty where you will end up in period 2. That is, there's a 9/23 chance you'll be in state 1, a 7/23 chance in state 2, etc. The expected value of being in state 2 and following the candidate policy $z^1(x)$ for 2 periods is 32.07 + $\beta\cdot[(9/23\times4.52)+(7/23\times32.07)+(7/23\times36.26)+(0\times36.78)+(0\times47.63)]$. This is shown in the shaded cells below.

$$\begin{bmatrix} V^1(x^1) \\ V^1(x^2) \\ V^1(x^3) \\ V^1(x^4) \\ V^1(x^5) \end{bmatrix} = \begin{bmatrix} 4.52 \\ 32.07 \\ 36.26 \\ 36.78 \\ 47.63 \end{bmatrix} + \beta\times \begin{bmatrix} 9/23 & 7/23 & 7/23 & 0 & 0 \\ 9/23 & 7/23 & 7/23 & 0 & 0 \\ 9/23 & 7/23 & 7/23 & 0 & 0 \\ 9/23 & 7/23 & 7/23 & 0 & 0 \\ 9/23 & 7/23 & 7/23 & 0 & 0 \end{bmatrix} \times \begin{bmatrix} 4.52 \\ 32.07 \\ 36.26 \\ 36.78 \\ 47.63 \end{bmatrix}$$

Using matrix notation, we can write the value for the entire vector of states,

$$V^1 = R^1 + \beta P^1 R^1$$

where V^1 is our first estimate of the candidate value function assuming the candidate policy vector $z^1(x)$.

If we followed this policy for 3 periods, our estimate of the value function would be closer to the infinite-horizon outcome:

$$V^1 = R^1 + \beta P^1\left(R^1 + \beta P^1 R^1\right) = R^1 + \beta P^1 R^1 + \beta^2 P^1 P^1 R^1 .$$

Following this process for n+1 periods we would have

$$V^1 = R^1 + \beta P^1 R^1 + \beta^2\left(P^1\right)^2 R^1 + \ldots + \beta^n\left(P^1\right)^n R^1 .$$

Taking the limit as $n\rightarrow\infty$, this sum can actually be written quite concisely. You probably have learned that $\sum_{t=1}^{\infty}\left(\frac{1}{1+r}\right)^t = {}^1\!/_r = \frac{\beta}{1-\beta}$.[3] This is a very useful formula: the present value of a stream of benefits of \$Y per year starting next year and discounted at the rate r is equal to \$Y/r. The analog to this for Markov Transition matrices is

$$\hat{V} = \lim_{n\rightarrow\infty} V_{T-n} = \lim_{n\rightarrow\infty}\sum_{t=0}^{n}\beta^t P^t R = \left(I-\beta P\right)^{-1} R ,$$

where I is the identity matrix and $\hat{V}$ is the vector of values that would result if a candidate policy that gives rise to P and R is followed ad infinitum. Hence, given the candidate payoff vector R^1 and candidate MTM P^1, V^1 can be found by solving a linear system of equations,

$$\left(I-\beta P^1\right)V^1 = R^1 \text{, or,} \tag{1}$$

$$V^1 = \left(I-\beta P^1\right)^{-1} R^1 . \tag{2}$$

This step is typically referred to as the *policy iteration* step, because it takes a candidate policy vector, which indicates a candidate MTM, P and a candidate payoff vector, R^1, and then finds the present value of the infinite stream of all future benefits assuming that that policy is iterated forever.

Solving this equation for the candidate policy presented above yields

$$V^1 = \begin{Bmatrix} 380.6 \\ 408.1 \\ 412.3 \\ 412.8 \\ 423.7 \end{Bmatrix}$$

Step 3: Our next step is to use the candidate value function to find a new candidate policy. To do this, we put V^1 on the right hand side of the Bellman's equation, and for

---

[3] The result is slightly different if payments start immediately: $\sum_{t=0}^{\infty}\beta^t = \frac{1}{1-\beta} = \frac{1+r}{r}$.

each state determine whether we are better off with a policy of F or W. For state 1, we compare two options:
if $z = \mathrm{F}$ then

$$\hat{V} = -2.33 + \beta\left(0 \cdot 380.6 + 1/20 \cdot 408.1 + 5/20 \cdot 412.3 + 7/20 \cdot 412.8 + 7/20 \cdot 423.7 \right) = 413.9$$

if $z = \mathrm{W}$ then

$$\hat{V} = 4.52 + \beta\left(9/23 \cdot 380.6 + 7/23 \cdot 408.1 + 7/23 \cdot 412.3 + 0 \cdot 412.8 + 0 \cdot 423.7 \right) = 403.1$$

That is, when soil moisture is at its lowest level, higher returns are obtained by fallowing than planting; the increase in the future payoff from fallowing more than makes up for the short term loss. In the remaining states W yields higher returns than F. Hence, our second candidate policy vector, z^2, candidate payoff vector, $R^2$and candidate Markov transition matrix, P^2, are:

	z^2	R^2	P^2				
1	F	-2.33	0	1/20	5/20	7/20	7/20
2	W	32.07	9/23	7/23	7/23	0	0
3	W	36.26	9/23	7/23	7/23	0	0
4	W	36.78	9/23	7/23	7/23	0	0
5	W	47.63	9/23	7/23	7/23	0	0

Again using (2), we can solve for a new candidate value function assuming that z^2 is followed ad infinitum. In this case we get $V^2 = \left[434.4 \;\; 454.8 \;\; 459.0 \;\; 459.5 \;\; 470.4\right]'$.

We repeat the process again, this time with the V^2 on the RHS of the Bellman's equation. In this case, it turns out that z^3 is the same as z^2. Hence, if we decided to follow the policy rule $[F\ W\ W\ W\ W]$ from t=2 onward, it would also be optimal to follow that same policy rule at t=1. This means that the candidate policy $z^2(x)$ is the optimal policy for an infinite horizon, $z^*(x)$.

As Burt and Allison point out, in fact soil moisture is not a discrete variable. Hence, this problem is actually not a true DD problem because although their choice variable is discrete – plant or fallow – the true state variable is continuous. Discretizing the state space as they did is only one of a number of ways to approximate the solution of a continuous state problem. We will discuss other methods to deal with continuous state variables in the next lecture.

V. What does it mean for the value & policy functions to be constant over time?

There is frequently some confusion about what is and what is not constant over time in an infinite horizon dynamic optimization problem. It is the functions that will be constant over time – not the policies themselves. For example, at the optimum of B&A's problem, over the first 4 periods, the value and policy functions would be constant as shown below.

x	t=1 $V(x)$	t=2 $V(x)$	t=3 $V(x)$	t=4 $V(x)$	t=1 $z^*(x)$	t=2 $z^*(x)$	t=3 $z^*(x)$	t=4 $z^*(x)$
1	434.4	434.4	434.4	434.4	F	F	F	F
2	454.8	454.8	454.8	454.8	W	W	W	W
3	459.0	459.0	459.0	459.0	W	W	W	W
4	459.5	459.5	459.5	459.5	W	W	W	W
5	470.4	470.4	470.4	470.4	W	W	W	W

However, in an actual realization of the problem, the state would change over time depending on the policy chosen and the random event distributed according to the MTM's $P(F)$ and $P(W)$. For example, we might see a path like the one presented below. Starting in period 1 in state 1 the optimal policy is to fallow. There are then four states in which we might end up next period, state 2 with probability 1/20, state 3 with probability 5/20, state 4 with probability 7/20 and state 5 with probability 7/20, but only one will actually occur. In the example below we assume that we have good luck and the soil moisture increases to state 5 in period 2. In period 2, since x=5, planting is optimal. In this case the soil moisture drops, say to state 2. The optimal policy in this state is W and there's a 9/23 chance that we will end up in state 1 in the following period, which we assume is what happens. Hence, in period 4 we are back in state 1 and fallowing would once again be optimal.

x	t=1 $V(x)$	t=2 $V(x)$	t=3 $V(x)$	t=4 $V(x)$	t=1 $z^*(x)$	t=2 $z^*(x)$	t=3 $z^*(x)$	t=4 $z^*(x)$
1	434.4	434.4	434.4	434.4	F	F	F	F
2	454.8	454.8	454.8	454.8	W	W	W	W
3	459.0	459.0	459.0	459.0	W	W	W	W
4	459.5	459.5	459.5	459.5	W	W	W	W
5	470.4	470.4	470.4	470.4	W	W	W	W
					R=-2.33	R=47.63	R=32.07	R=-2.33

Simulating an optimal policy path

In presenting your results, it is often necessary to use Monte Carlo simulation of a number of paths and present those results either graphically or numerically. To begin thinking about how one might actually simulate an optimal policy path, consider first the following thought experiment.

Suppose that there are 6 states of the world and, given your optimal choice in each of these states, the chance of ending up in each of these states is equal to 1/6. So, you could simulate this random path by rolling a die. Suppose you start in state 3. Then you roll your die and, if it comes up with a 1, you advance to state 1. If you roll a 5, you end up in state 5, and so on. In this way, rolling your die, time after time, you could obtain a random path that coincides with the optimal policy. Further, if you wanted to simulate a second random path, you could go back to period 1 and state 3 and start over again. In this way you could obtain the following simulated path.

	t=1	t=2	t=3	t=4	t=5	t=6	t=7	t=8	t=9	t=10
Simulation #1	x_t=3	x_t=5	x_t=4	x_t=2	x_t=5	x_t=1	x_t=3	x_t=2	x_t=2	x_t=4
Simulation #2	x_t=3	x_t=3	x_t=6	x_t=2	x_t=1	x_t=4	x_t=5	x_t=2	x_t=3	x_t=5

Now, how could we do this with a computer instead of dice? We could use a random number generator, which would generate a number between 0 and 1. If the value drawn is less than 1/6, this is treated as a 1 on the die, between 1/6 and 2/6 is a 2, and so on. In other words, we draw a random number, say ε, and then carry out the following logical test:

if $1/6 > \varepsilon$ then go to state 1,

  else if $2/6 > \varepsilon$ then go to state 2,

    else if $3/6 > \varepsilon$ then go to state 3, …

In this very simple case, the Markov Transition matrix would look like this.

	x_{t+1}=1	x_{t+1}=2	x_{t+1}=3	x_{t+1}=4	x_{t+1}=5	x_{t+1}=6
x_t=1	1/6	1/6	1/6	1/6	1/6	1/6
x_t=2	1/6	1/6	1/6	1/6	1/6	1/6
x_t=3	1/6	1/6	1/6	1/6	1/6	1/6
x_t=4	1/6	1/6	1/6	1/6	1/6	1/6
x_t=5	1/6	1/6	1/6	1/6	1/6	1/6
x_t=6	1/6	1/6	1/6	1/6	1/6	1/6

So, for any x_t, the <u>cumulative</u> probabilities would be as follows:

	$x_{t+1}\le 1$	$x_{t+1}\le 2$	$x_{t+1}\le 3$	$x_{t+1}\le 4$	$x_{t+1}\le 5$	$x_{t+1}\le 6$
x_t	1/6	2/6	3/6	4/6	5/6	6/6

We can now generalize the problem as follows. Suppose that an optimal Markov Transition matrix has been identified, say P, in which, $P(x_t,x_{t+1})$ is the probability of going from x_t to x_{t+1} given that the optimal choice is being made. In any period, the state variable x can take on n values, $x_1,...,x_n$. Using P, a randomly chosen policy path for T periods can be generated using the following algorithm:

Pseudo-code for stochastic policy simulation

Time loop

State loop

Probability check

Set $x_t = x_0$, a chosen value

For every period to be simulated, $t==1,\ldots T$

- Generate a uniform random number, $\varepsilon \in [0.1]$
- Set the variable CDF = 0
- for every possible value of
  - $x_{t+1}, x_i = x_1, x_2, \ldots, x_n$
  - CDF = CDF+$P(x_t, x_i)$
  - If CDF$\geq\varepsilon$, then
    - $x_{t+1} = x_i$
    - Exit State Loop
  - End if
- End of State loop
- Record x_t, $z_t^*(x_t)$, and x_{t+1}
- Set $x_t = x_{t+1}$

End of Time loop

Matlab code for stochastic policy simulation

```
% Initial value for the index
ix = 3;

for it = 1:nt
   % Store values
   spreadsheet(it, 1) = it;
   spreadsheet(it, 2) = xgrid(ix);

% Generate a random number between 0 and 1
   eps = rand;

% Figure out which cell was chosen using the
%Optimal Markov Transition Matrix, mtmStar.
% As soon as cdf>eps, we've found xt+1
   cdf = 0;
   for ix1 = 1:nx
      cdf = cdf + mtmStar(ix, ix1);
      if cdf > eps
         ix = ix1; % We'll go to ix1 next
         break;   % stop looping over ix1
      end
   end
end
```

note that this is intuitive, but not computationally efficient code.

In practice (and in your problem set) you would need to run many simulations to evaluate the distribution of your optimal paths over time.

It can also be useful to present the limiting probability distribution. In the Burt and Allison case, following the optimal policy leads to the following limiting probability distribution so that the most common state is state 3. Another way to interpret this is that in the long run fallowing will be optimal 28% of the time.

State	**1**	**2**	**3**	**4**	**5**
Limiting Probability	0.281	0.233	0.289	0.098	0.098

VI. References

Burt, O.R. and J.R. Allison. 1963. "Farm Management Decisions With Dynamic Programming." *Journal of Farm Economics* 45:121-37.

Rust, John. 1996. Numerical Dynamic Programming in Economics. In H. Amman, D. Kendrick and J. Rust (eds.), *Handbook of Computational Economics*. New York: North Holland.

VII. Readings for next class

Judd (1998) (Available here, must be logged into TAMU account to access).

- Read the section, "Limits of discretization methods" on p. 430, then look over the beginning of the section starting on p. 424.
- Read pages 434-440.