Discrete Random Variables

Allowable Values: $\bar{x} \in \mathcal{R}^n$

Markov Property:

$$\Pr(x_{t+1}|x_t, x_{t-1}, x_{t-2}, \cdots) = \Pr(x_{t+1}|x_t)$$

Markov Chains

State Space Vectors:

$$e_i (n \times 1)$$

$$e_{ij} = 1, j = i$$

$$e_{ij} = 0, j \neq i$$

Transition Matrix:

$$P_{ij} = \Pr(x_{t+1} = e_j | x_t = e_i)$$

$$\sum_{j=1}^{n} P_{ij} = 1, \text{for } i = 1, \ldots, n$$

Initial Probabilities:

$$\pi(n \times 1)$$

$$\pi_{0i} = \Pr(x_0 = e_i)$$

$$\sum_{i=0}^{n} \pi_{0i} = 1$$

Distributional Dynamics

Consider the initial distribution:

$$\pi_{0i} = \Pr(x_0 = e_i) \Leftrightarrow \pi_0 = \begin{bmatrix} \Pr(x_0 = e_1) \\ \Pr(x_0 = e_2) \\ \vdots \\ \Pr(x_0 = e_{n-1}) \\ \Pr(x_0 = e_n) \end{bmatrix}$$

Subsequent distributions:

$$\pi_1' = \Pr(x_1 = e_i) = \pi_0' P$$

$$\pi_2' = \Pr(x_2 = e_i) = \pi_0' P^2$$

$$\pi_t' = \Pr(x_t = e_i) = \pi_0' P^t$$

Stationary Distribution

$$\pi_{t+1} = \pi_t$$

$$\pi_{t+1}' = \pi_t' P$$

$$\pi'(I - P) = 0$$

$$(I' - P')\pi = 0$$

$$(I - P')\pi = 0$$

Stationary π is an eigenvector for unit eigenvalue of P'.

Normalized to $\sum \pi_i = 1$

Assigning Values to States

Define:

$$y_t = \bar{y}' x_t$$

Vector $\bar{y}$ is a list of the possible values that y may take.

Therefore, if $x_t = e_1$ and $\bar{y}' = [\ 5\ \ 1\ \ 3\ ]$, then

$$y_t = [\ 5\ \ 1\ \ 3\ ]\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = 5$$

Expectations

$$E(y_{t+1} | x_t = e_i) = \sum_j P_{ij} \bar{y}_j = (P\bar{y})_i$$

where $(P\bar{y})_i$ is the i'th element of $P\bar{y}$

Example

$$P = \begin{bmatrix} 1/2 & 1/2 \\ 1/3 & 2/3 \end{bmatrix} \text{ and } \bar{y} = \begin{bmatrix} 2 \\ 5 \end{bmatrix}$$

$$\begin{aligned} E(y_{t+1}|x_t = e_2) & \\ &= \left\{ \begin{bmatrix} 1/2 & 1/2 \\ 1/3 & 2/3 \end{bmatrix} \begin{bmatrix} 2 \\ 5 \end{bmatrix} \right\}_2 \\ &= \begin{bmatrix} (\frac{1}{2})(2) + (\frac{1}{2})(5) \\ (\frac{1}{3})(2) + (\frac{2}{3})(5) \end{bmatrix}_2 \\ &= (\frac{1}{3})(2) + (\frac{2}{3})(5) = 4 \end{aligned}$$

In like fashion:

$$E(y_{t+k}|x_t = e_i) = \sum_j P_{ij}^{(k)} \bar{y}_j = (P^k \bar{y})_i$$

where $P_{ij}^{(k)}$ denotes the (i,j) element of P^k

For the earlier example:

$$\begin{aligned} E(y_{t+2}|x_t = e_2) & \\ &= \left\{ \begin{bmatrix} 1/2 & 1/2 \\ 1/3 & 2/3 \end{bmatrix}^2 \begin{bmatrix} 2 \\ 5 \end{bmatrix} \right\}_2 \\ &= \left\{ \begin{bmatrix} \frac{5}{12} & \frac{7}{12} \\ \frac{7}{18} & \frac{11}{18} \end{bmatrix} \begin{bmatrix} 2 \\ 5 \end{bmatrix} \right\}_2 \\ &= \begin{bmatrix} \frac{15}{4} \\ \frac{23}{6} \end{bmatrix}_2 = \frac{23}{6} \end{aligned}$$

Application to discounted sums

$$\sum_{k=0}^{\infty} \beta^k E[y_{t+k}|x_t = e_i]$$

$$= \bar{y}_i + \beta E[y_{t+1}|x_t = e_i] + \beta^2 E[y_{t+2}|x_t = e_i] + \cdots$$

$$\bar{y}_i + \beta(P\bar{y})_i + \beta^2(P^2\bar{y})_i + \cdots$$

$$= [(I - \beta P)^{-1}]\bar{y}_i$$

Example:
Suppose $C = \{c_1, c_2, c_3\} = \{1,2,3\}$
Also suppose that $U(c) = \{4,6,7\}$:

$$U(1) = 4$$
$$U(2) = 6$$
$$U(3) = 7$$

In this case let:

$$\bar{y} = \begin{bmatrix} U(c_1) \\ U(c_2) \\ U(c_3) \end{bmatrix} = \begin{bmatrix} 4 \\ 6 \\ 7 \end{bmatrix}$$

Therefore:

$$U(c_t) = \bar{y}' x_t,$$

where x_t is a Markov chain governed by P :

$$P_{i,j} = \Pr(x_{t+1} = e_j | x_t = e_i)$$

Compute:

$$E \sum_{t=0}^{\infty} \beta^t U(c_t) = \sum_{t=0}^{\infty} \beta^t E[y_t]$$

Define:

$$v_i \equiv E\left[\sum_{t=0}^{\infty} \beta^t U(c_t) \middle| x_0 = \bar{e}_i \right]$$
$$= E\left[\sum_{t=0}^{\infty} \beta^t \bar{y}' x_t \middle| x_0 = \bar{e}_i \right]$$

From above:

$$v_i = [(I - \beta P)^{-1}] \bar{y}_i$$

Now define:

$$v = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = [(I - \beta P)^{-1}] \bar{y}$$

Finally:

$$V \equiv E\sum_{t=0}^{\infty}\beta^t U(c_t) = \pi_0' v$$

$$= \pi_0'[(I-\beta P)^{-1}]\begin{bmatrix} 4 \\ 6 \\ 7 \end{bmatrix}$$

Suppose:

$$\pi_0 = \begin{bmatrix} .2 \\ .4 \\ .4 \end{bmatrix} \text{ and } P = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{6} & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{6} & \frac{1}{2} \end{bmatrix}$$

For $\beta = 0.8$, $V = 29.385$.
Rows must sum to unity.

The Student Problem

Suppose we have three states, $X = \{sick, ok, great\}$
Transition Matrix: $P \Leftrightarrow P_{ij}$ =Prob$(x_{t+1} = \overline{x}_j | x_t = \overline{x}_i)$:

$$P = \begin{bmatrix} Prob(sick|sick) & Prob(ok|sick) & Prob(great|sick) \\ Prob(sick|ok) & Prob(ok|ok) & Prob(great|ok) \\ Prob(sick|great) & Prob(ok|great) & Prob(great|great) \end{bmatrix}$$

Suppose that transitions also depend on controls, u.
Let $U = \{TV, study\}$
Reward function, $r(x,u)$:

$$r(1,1) = r(sick,TV) = 1$$
$$r(1,2) = r(sick,study) = 0$$
$$r(2,1) = r(ok,TV) = 2$$
$$r(2,2) = r(ok,study) = 3$$
$$r(3,1) = r(great,TV) = 2$$
$$r(3,2) = r(great,study) = 5$$

Last period problem is obvious.
No consequences for actions taken.
If $x_t = sick$, watch TV and $r_t = 1$.
If $x_t = ok$, study and $r_t = 3$.
If $x_t = great$, study and $r_t = 5$.
Denote the last period value function as

$$v_0 : \begin{cases} v_0(sick) = 1 \\ v_0(ok) = 3 \\ v_0(great) = 5 \end{cases}$$

The optimal policy rule is:

$$h_0 : \begin{cases} h_0(sick) = 1 \\ h_0(ok) = 2 \\ h_0(great) = 2 \end{cases}$$

Although this problem is trivial, let's solve with Matlab.
Define:

$$r = \begin{bmatrix} 1 & 0 \\ 2 & 3 \\ 2 & 5 \end{bmatrix}$$

The Matlab max command is quite helpful.
Syntax:

[C,I]=max(A), where A is a matrix

Output:

C: Row vector listing the maximum entries in each column.

I: Row vector of the indices of the maximum entry

Example:

$$A = \begin{bmatrix} 4 & 2 & 1 \\ 7 & 8 & 5 \\ 9 & 4 & 10 \end{bmatrix}$$

[C,I]=max(A)

C=[9 8 10]

I=[3 2 3]

For this problem, let $R = r'$.

[v,h]=max(R) $\Rightarrow$ v=[1 3 5]
h=[1 2 2]

Meaning of h:

$h(1) = 1 \Rightarrow$ Watch TV when sick.

$h(2) = 2 \Rightarrow$ Study when OK.

$h(3) = 2 \Rightarrow$Study when great.

Meaning of v:

$v(1) = v(sick) = 1.$

$v(2) = v(ok) = 3.$

$v(3) = v(great) = 5.$

Next-to-the-last period problem is a little harder.

Consequences now matter.

Specify consequences with two transition matrices.

$P(TV)$: stochastic consequences of watching TV.

$P(study)$: stochastic consequences of studying.

Denote the next-to-last period value function as v_1.

Bellman equation for each initial state:

$$
\begin{aligned}
v_1(sick) &= \max\{r(sick,TV) + \beta E(v_0|sick,TV), r(sick,study) + \beta E(v_0|sick,study)\} \\
v_1(ok) &= \max\{r(ok,TV) + \beta E(v_0|ok,TV), r(ok,study) + \beta E(v_0|ok,study)\} \\
v_1(great) &= \max\{r(great,TV) + \beta E(v_0|great,TV), r(great,study) + \beta E(v_0|great,study)\}
\end{aligned}
$$

Expand:

$$
\begin{aligned}
v_1(sick) &= \max\{r(sick,TV) + \beta[P^{TV}(s|s)v_0(s) + P^{TV}(ok|s)v_0(ok) + P^{TV}(g|s)v_0(g)], \\
&\quad r(sick,study) + \beta[P^{Study}(s|s)v_0(s) + P^{Study}(ok|s)v_0(ok) + P^{Study}(g|s)v_0(g)]\} \\
v_1(ok) &= \max\{r(ok,TV) + \beta[P^{TV}(s|ok)v_0(s) + P^{TV}(ok|ok)v_0(ok) + P^{TV}(g|ok)v_0(g)], \\
&\quad r(ok,study) + \beta[P^{Study}(s|ok)v_0(s) + P^{Study}(ok|ok)v_0(ok) + P^{Study}(g|ok)v_0(g)]\} \\
v_1(great) &= \max\{r(great,TV) + \beta[P^{TV}(s|g)v_0(s) + P^{TV}(ok|g)v_0(ok) + P^{TV}(g|g)v_0(g)], \\
&\quad r(great,study) + \beta[P^{Study}(s|g)v_0(s) + P^{Study}(ok|g)v_0(ok) + P^{Study}(g|g)v_0(g)]\}
\end{aligned}
$$

Denote:

$$
v_0' = \begin{bmatrix} v_0(sick) \\ v_0(ok) \\ v_0(great) \end{bmatrix}
$$

Therefore:

$$
P^{TV}v_0' = \begin{bmatrix} P^{TV}(s|s)v_0(s) + P^{TV}(ok|s)v_0(ok) + P^{TV}(g|s)v_0(g) \\ P^{TV}(s|ok)v_0(s) + P^{TV}(ok|ok)v_0(ok) + P^{TV}(g|ok)v_0(g) \\ P^{TV}(s|g)v_0(s) + P^{TV}(ok|g)v_0(ok) + P^{TV}(g|g)v_0(g) \end{bmatrix}
$$

$$P^{Study}v_0' = \begin{bmatrix} P^{Study}(s|s)v_0(s) + P^{Study}(ok|s)v_0(ok) + P^{Study}(g|s)v_0(g) \\ P^{Study}(s|ok)v_0(s) + P^{Study}(ok|ok)v_0(ok) + P^{Study}(g|ok)v_0(g) \\ P^{Study}(s|g)v_0(s) + P^{Study}(ok|g)v_0(ok) + P^{Study}(g|g)v_0(g) \end{bmatrix}$$

Now compute:

$$R + \beta \begin{bmatrix} (P^{TV}v_0')' \\ (P^{Study}v_0')' \end{bmatrix}$$

Finally, perform the max command on this last expression to obtain: v_1 and h_1.

Infinite-horizon Student Problem

$$\underset{u}{Max} \sum_{t=0}^{\infty} \beta^t r(x,u)$$

subject to: Prob$(x_{t+1}|x_t) \sim P(u_t) \bullet x_t$

Bellman Equation:

$$V(x_t) = \underset{u_t}{Max} \ \{r(x_t,u_t) + \beta E(v(x_{t+1})|x_t,u_t\}$$

The Matlab notation equivalent is:

$$v' = \max_{TV,Study} \left\{ R + \beta \begin{bmatrix} (P^{TV}v')' \\ (P^{Study}v')' \end{bmatrix} \right\}$$

Iteration on the Value Function

Guess a function $(v^1)'$
Plug into RHS of Bellman.
Perform max and denote resulting LHS as $(v^2)'$
Now put $(v^2)'$ on RHS to derive LHS as $(v^3)'$
Continue until $(v^{k+1})' \approx (v^k)'$
Then the h output of the max command is the optimal policy.

Howard's Policy Improvement

Start with an assumed policy: $h^0 = [h_1^0, h_2^0, h_3^0]$.
Calculate the value of h^0 :

$$v^0(h^0) = \begin{bmatrix} r(sick, h^0(sick)) \\ r(ok, h^0(ok)) \\ r(great, h^0(great)) \end{bmatrix} + \beta P^{h^0} v^0(h^0)$$

P^{h^0} is composed of the appropriate rows of P^{TV} and P^{Study} for h^0.

$$Iv^0(h^0) = r^{h^0} + \beta P^{h^0} v^0(h^0)$$

$$v^0 = [I - \beta P^{h^0}]^{-1} r^{h^0}$$

Next, use:

$$\max \left\{ R + \beta \begin{bmatrix} (P^{TV}(v^0)')' \\ (P^{Study}(v^0)')' \end{bmatrix} \right\}$$

to obtain h^1.

Iterate until h converges.

Specific Example

$$P(TV) = \begin{bmatrix} 0.3 & 0.7 & 0 \\ 0.3 & 0.4 & 0.3 \\ 0.1 & 0.3 & 0.6 \end{bmatrix}$$

$$P(study) = \begin{bmatrix} 1.0 & 0 & 0 \\ 0.4 & 0.4 & 0.2 \\ 0.3 & 0.2 & 0.5 \end{bmatrix}$$