Applied Math 9

Handout 5 for Markov Chains: What Does Multiplication on the Right Do?

If $\mathbf{v}$, a row vector, is a probability distribution on the points $\{1, 2, \ldots, m\}$, and if P is an $m \times m$ one step transition matrix, then we we can interpret $\mathbf{w} = \mathbf{v}P^n$ as follows. If X_n is a Markov chain with 1-step transition matrix P and distribution $\mathbf{v}$ at time zero (i.e., $v_i = P\{X_0 = i\}$), then $\mathbf{w}$ is the distribution of the chain at time n:

$$w_i = P\{X_n = i\}.$$

Thus each multiplication by P on the right corresponds to evolution of the distribution by one additional time step. Does multiplication of a vector by P on the *left* correspond to anything interesting? It turns out that the answer is yes.

Let $\mathbf{c}$ be an m dimensional column vector of costs. Suppose we will start the chain off in state i at time zero, run it for n time steps, and then pay the amount c_{X_n}. This is a *random* quantity. What is its expected value? By the definition of expected value, it is

$$E\left[c_{X_n} | X_0 = i\right] = \sum_{j=1}^{m} c_j P\{X_n = j | X_0 = i\}.$$

Since we started at i, we know that the distribution of X_n is given by the ith row of P^n, and the expected value is then the dot product of that row and $\mathbf{c}$. From the definition of matrix multiplication, the ith component of $P^n\mathbf{c}$ is the expected value we seek: $E\left[c_{X_n} | X_0 = i\right]$. In other words, $P^n\mathbf{c}$ is a vector of expected costs when the chain is run for n steps, and the value of the ith component corresponds to starting at state i at time zero.

Finally, observe that if we had an initial distribution $\mathbf{v}$ at time zero (rather than a particular state), then the expected cost at time zero would be

$$\sum_{i=1}^{m} v_i (P^n\mathbf{c})_i = \mathbf{v}P^n\mathbf{c}.$$

Example. Suppose that the cost vector is

$$\mathbf{c} = \begin{pmatrix} 5 \\ 1 \\ -2 \end{pmatrix},$$

and that the one step transition matrix is

$$P = \begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/3 & 0 & 2/3 \\ 1 & 0 & 0 \end{pmatrix}.$$

Suppose that we start in state 1 at time zero, and wish to compute $E\left[c_{X_1}|X_0 = i\right]$, i.e., one step into the future. We know that the distribution of X_1, given $X_0 = 1$, is given by

$$\begin{array}{ccc} 1 & \text{with probability} & 1/2 \\ 2 & \text{with probability} & 1/2 \\ 3 & \text{with probability} & 0 \end{array}.$$

Thus the expected value is just

$$E\left[c_{X_1}|X_0 = 1\right] = c_1 \cdot \frac{1}{2} + c_2 \cdot \frac{1}{2} + c_3 \cdot 0 = 5 \cdot \frac{1}{2} + 1 \cdot \frac{1}{2} = 3.$$

Note that this is indeed just the dot product of the first row of P with $\mathbf{c}$. Similarly,

$$E\left[c_{X_1}|X_0 = 2\right] = c_1 \cdot \frac{1}{3} + c_2 \cdot 0 + c_3 \cdot \frac{2}{3} = 5 \cdot \frac{1}{3} + -1 \cdot \frac{2}{3} = 1.$$

and

$$E\left[c_{X_1}|X_0 = 3\right] = c_1 \cdot 1 + c_2 \cdot 0 + c_3 \cdot 0 = 5 \cdot 1 = 5.$$

If we arrange these as a vector, then we get

$$\begin{pmatrix} 3 \\ 1 \\ 5 \end{pmatrix},$$

which is the same as

$$P\mathbf{c} = \begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/3 & 0 & 2/3 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 5 \\ 1 \\ -2 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \\ 5 \end{pmatrix}.$$