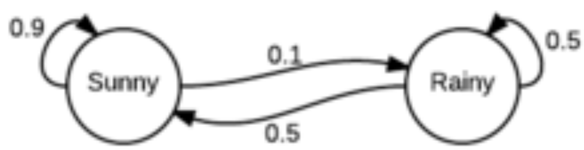
A very simple weather model

The probabilities of weather conditions (modeled as either rainy or sunny), given the weather on the preceding day, can be represented by a transition matrix:

$$P = \begin{bmatrix} 0.9 & 0.1 \\ 0.5 & 0.5 \end{bmatrix}$$

The matrix P represents the weather model in which a sunny day is 90% likely to be followed by another sunny day, and a rainy day is 50% likely to be followed by another rainy day. The columns can be labeled "sunny" and "rainy", and the rows can be labeled in the same order.

$(P)_{ij}$ is the probability that, if a given day is of type i, it will be followed by a day of type j. Notice that the rows of P sum to 1: this is because P is a stochastic matrix.



State diagram representing P

Predicting the weather

The weather on day 0 is known to be sunny. This is represented by a vector in which the "sunny" entry is 100%, and the "rainy" entry is 0%:

$$\mathbf{x}^{(0)} = \begin{bmatrix} 1 & 0 \end{bmatrix}$$

The weather on day 1 can be predicted by:

$$\mathbf{x}^{(1)} = \mathbf{x}^{(0)} P = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} 0.9 & 0.1 \\ 0.5 & 0.5 \end{bmatrix} = \begin{bmatrix} 0.9 & 0.1 \end{bmatrix}$$

(Note: in the above calculation, 1 * 0.9 + 0 * 0.5 = 0.9 and 1 * 0.1 + 0 * 0.5 = 0.1)

Thus, there is a 90% chance that day 1 will also be sunny.

The weather on day 2 can be predicted in the same way:

$$\mathbf{x}^{(2)} = \mathbf{x}^{(1)} P = \mathbf{x}^{(0)} P^2 = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} 0.9 & 0.1 \\ 0.5 & 0.5 \end{bmatrix}^2 = \begin{bmatrix} 0.86 & 0.14 \end{bmatrix}$$

or

$$\mathbf{x}^{(2)} = \mathbf{x}^{(1)} P = \begin{bmatrix} 0.9 & 0.1 \end{bmatrix} \begin{bmatrix} 0.9 & 0.1 \\ 0.5 & 0.5 \end{bmatrix} = \begin{bmatrix} 0.86 & 0.14 \end{bmatrix}$$

General rules for day *n* are:

$$\mathbf{x}^{(n)} = \mathbf{x}^{(n-1)} P$$
$$\mathbf{x}^{(n)} = \mathbf{x}^{(0)} P^n$$

Steady state of the weather

In this example, predictions for the weather on more distant days are increasingly inaccurate and tend towards a *steady state vector*. This vector represents the probabilities of sunny and rainy weather on all days, and is independent of the initial weather.

The steady state vector is defined as:

$$\mathbf{q} = \lim_{n \to \infty} \mathbf{x}^{(n)}$$

q converges to a strictly positive vector only if *P* is a regular transition matrix (that is, there is at least one P^n with all non-zero entries).

Since the **q** is independent from initial conditions, it must be unchanged when transformed by *P*. This makes it an eigenvector (with eigenvalue 1), and means it can be derived from *P*. For the weather example:

$$\begin{aligned} P &= \begin{bmatrix} 0.9 & 0.1 \\ 0.5 & 0.5 \end{bmatrix} \\ \mathbf{q}P &= \mathbf{q} && (\mathbf{q} \text{ is unchanged by } P.) \\ &= \mathbf{q}I \\ \mathbf{q}(P - I) &= \mathbf{0} \\ &= \mathbf{q}\left(\begin{bmatrix} 0.9 & 0.1 \\ 0.5 & 0.5 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right) \\ &= \mathbf{q} \begin{bmatrix} -0.1 & 0.1 \\ 0.5 & -0.5 \end{bmatrix} \end{aligned}$$

$$\begin{bmatrix} q_1 & q_2 \end{bmatrix} \begin{bmatrix} -0.1 & 0.1 \\ 0.5 & -0.5 \end{bmatrix} = \begin{bmatrix} 0 & 0 \end{bmatrix}$$

So $-0.1q_1 + 0.5q_2 = 0$ and since they are a probability vector we know that $q_1 + q_2 = 1.$

Solving this pair of simultaneous equations gives the steady state distribution:

$$\begin{bmatrix} q_1 & q_2 \end{bmatrix} = \begin{bmatrix} 0.833 & 0.167 \end{bmatrix}$$

In conclusion, in the long term, about 83.3% of days are sunny and 16.7% of days are rainy.

The power method

Instead of solving the equations, we can use the following power method to approximate the steady state vector **q**. As in the above example, suppose the initial vector for day 0 $\mathbf{X}^{(0)}$ = [1, 0], then:

$\mathbf{X}^{(1)} = \mathbf{X}^{(0)} P$ = [0.9, 0.1]

$\mathbf{X}^{(2)} = \mathbf{X}^{(1)} P$ = [0.86, 0.14]

$\mathbf{X}^{(3)} = \mathbf{X}^{(2)} P$ = [0.844, 0.156]

$\mathbf{X}^{(4)} = \mathbf{X}^{(3)} P$ = [0.8376, 0.1624]

$\mathbf{X}^{(5)} = \mathbf{X}^{(4)} P$ = [0.83504, 0.16496]

$\mathbf{X}^{(6)} = \mathbf{X}^{(5)} P$ = [0.834016, 0.165984]

$\mathbf{X}^{(7)} = \mathbf{X}^{(6)} P$ = [0.8336064, 0.1663936]

$\mathbf{X}^{(8)} = \mathbf{X}^{(7)} P$ = [0.83344256, 0.16655744]

…

Observe that the vector converges as the computation goes on. We can stop this process if the vector doesn't change much in its values (within a threshold) comparing to the last iteration. For example, because $\mathbf{X}^{(8)} \approx \mathbf{X}^{(7)}$, we can stop at $\mathbf{X}^{(8)}$ and consider the steady vector $\mathbf{q} \approx \mathbf{X}^{(8)}$.

Acknowledgement: The above description modifies and extends content from the "Examples of Markov chains" available at http://en.wikipedia.org/wiki/Examples_of_Markov_chains.