Applications: Stochastic Matrices and Markov Chains

Suppose we have a set of n *states* $X = \{X_1, \ldots, X_n\}$ that a given system can be in, and a *state vector* $\vec{v} = (v_1, \ldots, v_n)$ represents the distribution of things in these states. Examples include:

- $X_1, \ldots, X_n$ could literally be the 50 states, and $\vec{v}$ could be the population vector with v_1 the population of Alabama, v_2 the population of Alaska, etc.
- $X_1, \ldots, X_n$ could represent investment accounts and v_i could represent the number of dollars allocated to the ith account,
- $X_1, \ldots, X_n$ could represent positions and momenta of a set of moving molecules, with each $\vec{v}$ representing a particular configuration.

We can model the evolution of the system over time by finding a rule that describes how to go from the current state to the next state. The system may be *deterministic*, i.e. the next state may be determined by the current state, or it may be non-deterministic, i.e., the same current state may be followed by different states.

In the latter case, we can ask what is the *probability* of an element in state X_i transitioning to state X_j, let's call this number P_{ij}. The matrix of probabilities

$$T = \begin{bmatrix} P_{11} & P_{12} & \ldots & P_{1n} \\ P_{21} & P_{22} & \ldots & P_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ P_{n1} & P_{n2} & \ldots & P_{nn} \end{bmatrix}$$

is called the *transition matrix* of the system. If the current state is $\vec{v} = (v_1, \ldots, v_n)$ then an element in state X_i has a P_{ij} chance of moving to state X_j, so the next state vector $\vec{w} = (w_1, \ldots, w_n)$'s jth component gets a contribution of $P_{ij}v_i$ from state X_i. Summing over all the states, we see that $w_j = P_{1j}v_1 + P_{nj}v_2 + \ldots + P_{nj}v_n$, which we recognize as the dot product of the row vector $\vec{v}$ with the jth column of the transition matrix T. Thus, if $\vec{v}$ is the current state vector, then $\vec{w} = \vec{v}T$ if we write $\vec{v}, \vec{w}$ as row vectors or $\vec{w} = T^T\vec{v}$ if we write $\vec{v}, \vec{w}$ as column vectors.

Note also that if $X_1, \ldots, X_n$ are all of the possible states of the system, then an element of state X_i has to end up in one of these states, so $P_{i1} + P_{i2} + \ldots + P_{in}$ must equal 1. A matrix is *stochastic* if it has all nonnegative entries with every column summing to 1; every stochastic matrix is the transition matrix of a probabilistic system. In particular, if the transition probabilities are fixed, i.e. not dependent on the current state, then our system is called a *Markov chain.*

Example: A city has three districts, A, B and C and a stable total population of $300K$ samples every five years. Residents of district A tend to stay put, while residents of districts B and C tend to move more. The system has transition matrix

$$T = \begin{bmatrix} 0.8 & 0.2 & 0.3 \\ 0.1 & 0.5 & 0.3 \\ 0.1 & 0.3 & 0.4 \end{bmatrix}$$

Then if the current population distribution vector is $\vec{v} = (200K, 40K, 60K)$, the distribution vector for the next cycle will be

$$\begin{bmatrix} 200 & 40 & 60 \end{bmatrix} \begin{bmatrix} 0.8 & 0.2 & 0.3 \\ 0.1 & 0.5 & 0.3 \\ 0.1 & 0.3 & 0.4 \end{bmatrix} = \begin{bmatrix} 170K & 78K & 96K \end{bmatrix}.$$

After playing with these matrices a little, a question suggests itself: what kinds of population distributions are stable? Well, a stable state vector must satisfy $\vec{v} = T^T\vec{v}$; that is, it must be an eigenvector of T^T with eigenvalue $\lambda = 1$. In fact, we have

Theorem: (Perron-Frobenius) Every stochastic matrix has $\lambda = 1$ as an eigenvalue, and all other eigenvalues are less than 1 in absolute value. Moreover, if the P_{ij} are all strictly greater than zero, then there is a unique eigenvector $\vec{v}$ with eigenvalue $\lambda = 1$, i.e., there is a unique stable distribution vector. If some P_{ij}s equal zero, there may be multiple distinct stable distribution vectors.

Example:

$$\begin{bmatrix} 0.8-1 & 0.2 & 0.3 \\ 0.1 & 0.5-1 & 0.3 \\ 0.1 & 0.3 & 0.4-1 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}$$

so the eigenspace is spanned by $(1, 1, 1)$ and the stable distribution vector is $(100K, 100K, 100K)$.