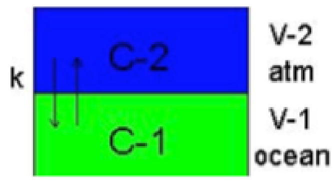
Eigenvalue problems

The eigenvalue problem is a classic problem of applied math that gives an elegant solution to the box model.

Consider a simple two-box model



$$\frac{dC_1}{dt} = -\frac{k}{h}C_1 + \frac{k}{h}C_2 \tag{1}$$

$$\frac{dC_2}{dt} = \frac{k}{h}C_1 - \frac{k}{h}C_2 \tag{2}$$

We can write tthis system elegantly as a matrix equation

$$\frac{d}{dt}\begin{bmatrix} C_1 \\ C_2 \end{bmatrix} = \begin{bmatrix} -\frac{k}{h} & \frac{k}{h} \\ \frac{k}{h} & -\frac{k}{h} \end{bmatrix} \begin{bmatrix} C_1 \\ C_2 \end{bmatrix} \tag{3}$$

$$\frac{d\mathbf{c}}{dt} = \mathbf{Kc} \tag{4}$$

where $\mathbf{c} = \begin{bmatrix} C_1 \\ C_2 \end{bmatrix}, \mathbf{K} = \begin{bmatrix} -\frac{k}{h} & \frac{k}{h} \\ \frac{k}{h} & -\frac{k}{h} \end{bmatrix}$

This points towards the solution to a system of differential equations. We are looking for a function with a derivative equal to a constant times the function itself. There is only one such function that has this property by construction: the exponential

Let's try an exponential in the equation:

$$\mathbf{c}(t) = \begin{bmatrix} e_1 \\ e_2 \end{bmatrix} exp(\lambda t) \tag{5}$$

$$\frac{d}{dt}\begin{bmatrix} C_1 \\ C_2 \end{bmatrix} = \begin{bmatrix} e_1 \\ e_2 \end{bmatrix} \lambda exp(\lambda t) = \begin{bmatrix} -\frac{k}{h} & \frac{k}{h} \\ \frac{k}{h} & -\frac{k}{h} \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \end{bmatrix} exp(\lambda t) \tag{6}$$

Write the equation above in matrix form,

$$\begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \end{bmatrix} exp(\lambda t) = \begin{bmatrix} -\frac{k}{h} & \frac{k}{h} \\ \frac{k}{h} & -\frac{k}{h} \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \end{bmatrix} exp(\lambda t) \tag{7}$$

$$\begin{bmatrix} -\frac{k}{h} - \lambda & \frac{k}{h} \\ \frac{k}{h} & -\frac{k}{h} - \lambda \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{8}$$

The above equation can also be written as:

$$(\mathbf{K} - \lambda \mathbf{I})\mathbf{e} = \mathbf{0} \tag{9}$$

where $\mathbf{I}$ is identity matrix, $\mathbf{e} = \begin{bmatrix} e_1 \\ e_2 \end{bmatrix}$

The Equation(9) holds if $\mathbf{e} = \mathbf{0}$ (the "trivial solution") or if $det((\mathbf{K} - \lambda\mathbf{I})) = 0$. Solve for $det((\mathbf{K} - \lambda\mathbf{I})) = 0$

$$(\frac{k}{h} + \lambda)^2 - (\frac{k}{h})^2 = 0 \tag{10}$$

$$\lambda_1 = 0, \lambda_2 = -\frac{2k}{h}$$

So there will be non-trivial solutions and we know the eigenvalues (λ_1, λ_2). Now solve for eigenvectors.

For $\lambda = 0$, the corresponding eigenvector is $[1, 1]^T$.

For $\lambda = -\frac{2k}{h}$, the corresponding eigenvector is $[1, -1]^T$. (Without normalization)

So the solution for this problem can be written as

$$\begin{bmatrix} C_1 \\ C_2 \end{bmatrix} = a_1 \begin{bmatrix} 1 \\ 1 \end{bmatrix} exp(0t) + a_2 \begin{bmatrix} 1 \\ -1 \end{bmatrix} exp(-\frac{2k}{h}t) \tag{11}$$

where a_1 and a_2 are constants and can be found using initial conditions $C_1(0) = 0, c_2(0) = C_0$:

$$a_1 = \frac{C_0}{2}, a_2 = -\frac{C_0}{2}$$

$$C_1 = \frac{C_0}{2} - \frac{C_0}{2} exp(-\frac{2k}{h}t)$$

$$C_2 = \frac{C_0}{2} + \frac{C_0}{2} exp(-\frac{2k}{h}t)$$

Markov Chain

Definition: A sequence of random variables taking values in a state space is called a Markov Chain if the probability of the next step only depends on the current state.

Using the notation of transition probabilities to define the probability of going from state x to state y as $T(x|y)$, we can write this mathematically:

$$T(x_n|x_{n-1}, x_{n-1}..., x_1) = T(x_n|x_{n-1})$$

Some Jargon about Markov Chain

1. Homogeneous

A chain is homogeneous at step t if the transition probabilities are independent of t. Thus the evolution of the Markov chain only depends on the previous state with a fixed transition matrix.

2. Irreducible

Every state is accessible in a finite number of steps from another state. That is, there are no absorbing states. In other words, one eventually gets everywhere in the chain.

Consider as an example surfing the web. We do want to reach all parts of the web so we don't want to be trapped into an subset.

3. Recurrent

States visited repeatedly are recurrent: positive recurrent if time-to-return is bounded and null recurrent otherwise. Harris recurrent if all states are visited infinitely often as $t \to \infty$

4. Aperiodic

There are no deterministic loops. This would be bad in our web example as well as we would be stuck in a loop at some pages.

5. Stationarity

We can finally give a formal definition of stationarity. A stationary Markov chain produces the same *marginal* distribution when multiplied by the transition matrix.

That is

$$sT = s$$

or

$$\sum_i s_i T_{ij} = s_j$$

A irreducible (goes everywhere) and aperiodic (no cycles) markov chain will converge to a stationary markov chain. It is the marginal distribution of this chain that we want to sample from, and which we do in metropolis (and for that matter, in simulated annealing).

As we can see above, to find stationary distribution, we need to solve an eigenvalue problem.

A sufficient, but not necessary, condition to ensure that $s(x)$ is the desired stationary distribution is the already seen reversibility condition, also known as detailed balance:

$$s(x)T(y|x) = s(y)T(x|y)$$

If one sums both sides over x

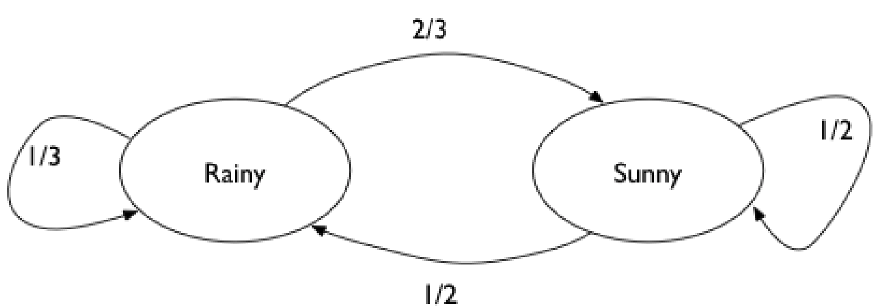
$$\int dxs(x)t(y|x) = s(y)\int dxT(x|y)$$

which gives us back the stationarity condition from above.

Thus we want to design us samplers which satisfy detailed balance.

Simple example of Markov Chain

We look at a very simple Markov Chain, with only two states: Rainy or Sunny. This chain is aperiodic and irreducible, so it has a stationary distribution. We find it by looking at powers of the transition matrix.



The stationary distribution can be solved for. Assume that it is

$$s = [p, 1-p]$$

. Then:

$$sT = s$$

gives us

$$p \times (1/3) + (1 - p) \times 1/2 = p$$

and thus $p = 3/7$

Applied to Box Models

The two boxes are two states in the Markov chain, and the behavior of each molecule is exactly like the example of weather prediction problem we discussed before.

Codes from lecture notes:

```
##Two box mass transfer model as a Markov Chain

First, plot the two box model analytical solution

k.h=.01                    ##k/h = 0.01 s^-1 (unit^-1)
 tt=0:1000                            #time vector
 c1.c2=100*exp(-2*tt*k.h)     #solution (c1 — c2)
 c1=(100 + c1.c2)/2
 c2=100-c1
 plot(tt,c1,type="l",col="red",ylim=c(0,100),lwd=3)
 lines(tt,c2,type="l",col="blue",lwd=3,lty=2)

## Markov chain
 n1=100;n2=0        #  number of particles, initial (arbitrary)
N1=n1;N2=n2                # Nj accumulates solutions
for(i in 1:1000){
if(n1==0)dn1=0 else dn1=sum(runif(n1)<k.h)
if(n2==0)dn2=0 else dn2=sum(runif(n2)<k.h)
##runif:  uniform distribution random numbers (Markov)
n1= n1-dn1+dn2       ## new value in box 1
n2= n2-dn2+dn1                ## in box 2 NB particle conserv
N1=c(N1,n1)
N2=c(N2,n2)
}
##add points to the plot, for the Markov chain problem
points(tt,N1,col="red",pch=16,cex=.6)
points(tt,N2,col="blue",pch=1,cex=.6)
legend("topright",legend=c("Box 1","Box 2"),col=c("red","blue"),
pch=c(16,1),lty=c(1,2),text.col=c("red","blue"))
## save the figure
dev.copy(png,"Fig.Markov.2box.png");dev.off()
```