Inference Control and Driving of Natural Systems

MSci/MSc

Nick Jones

nick.jones@imperial.ac.uk

Mathematics for Reasoning with Genetic Circuits (better inference through chemistry)

In this lecture we will provide some of the mathematical tools that are employed in building genetic circuits. These circuits are intrinsically stochastic and will be used for inference about an external signal.
The material presented in this lecture will help provide the ground-work for the next two practical sessions. It also helps understand the following two papers: Libby et al [1] and Kobayashi et al [2] which you are expected to read.

What we will cover

Chemical reactions
The stochiometry matrix
Point process models
The well mixed limit
Gillespie algorithm
MWC model.

Reminder Poisson Point Processes

Poisson Point Processes are purely specified by a rate, λ, at which events happen per unit time. A point process specifies a sequence of times at which events occur.
A classic example is the arrivals of buses: a certain number of buses, λ, are expected to arrive per unit time.
The intervals, τ, between the events are distributed like $\lambda exp(-\lambda\tau)$ and each inter-event interval is independent of the next.

Reminder Poisson Point Processes II

You might have been introduced to radioactive decay $NucleusA \rightarrow NucleusB + alphaparticle$. There will be a rate at which any one nucleus decays, μ, and each decay is independent of any other, so if the population of type A nuclei is N_1 the rate at which events occur is $\lambda = \mu N_1$. After a decay event the new rate will be $\lambda = \mu(N_1 - 1)$: after each event, conditional on population size, we update the rate at which events occur.

Reminder Poisson Point Processes III

- If I had two elements one unstable, and one fairly stable, then the decay rate of one element will be faster than the other $\mu_1 > \mu_2$.
- A mix of N_1 atoms of the first type and N_2 of the second then the rate at which my Geiger counter (radiation detector) registers an event will be $\lambda = \mu_1 N_1 + \mu_2 N_2$.
- When an event occurs I know that the chance that the decay came from element 1 is $\mu_1 N_1/(\mu_1 N_1 + \mu_2 N_1)$.

Simulating Radioactive decay

Suppose we have N_1 elements of type 1, N_2 elements of type 2, and decay rates of μ_1, $\mu 2$ at time t.

1. The time of the next event, $t + \tau$, can be calculated by drawing τ from $\lambda exp(-\lambda\tau)$ with $\lambda = \mu_1 N_1 + \mu_2 N_2$.
2. When an event occurs I know that the chance that the decay came from element 1 is $p_1 = \mu_1 N_1/(\mu_1 N_1 + \mu_2 N_1)$ so draw a uniform random variable on [0,1] and if it exceeds p_1 then $N_2 \rightarrow N_2 - 1$ else $N_1 \rightarrow N_1 - 1$.
3. Update N_1 or N_2 as above and let $t \rightarrow t + \tau$ and repeat.

This is the intuition behind the Gillespie algorithm for exact stochastic simulation. We will first explain the connection to chemical circuits and then formalize the above.

Chemical Reactions I

We will consider the well-mixed scenario where we suppose that spatial effects are not relevant. The only thing we need to track about the chemical species are their amounts.
Chemical state vector: $S(t)$ where $S_i(t)$ returns the number of molecules of species i.

Chemical Reactions II

Propensity function: the probability of reaction j occurring in interval $[t, t+dt)$, $a_j(S(t))dt$ (it is thus a time dependent rate).
Reaction events are Poisson point processes.
$S_1 \xrightarrow{k_1} \phi$ decay $a_1(S(t)) = k_1 S_1$.
$S_2 \xrightarrow{k_2} S_1$ conversion $a_2(S(t)) = k_2 S_2$.
Or instead we could have two diffusing species that collide:
$S_1 + S_2 \xrightarrow{k_3} S_3$ reaction $a_3(S(t)) = k_3 S_1(t) S_2(t)$.
$S_1 + S_1 \xrightarrow{k_4} S_4$ bimolecular reaction
$a_3(S(t)) = k_4 S_1(t)(S_1(t) - 1)/2$ (from ${}^nC_2$).
Stoichiometry matrix: Has as many columns as there are distinct reactions and as many rows as there are distinct chemical species.
The stoichiometry matrix tells us how the numbers of species i are changed under reaction j.

When and what for *next* reaction?

Consider $p(j, \tau | s(t), t) d\tau$ the probability that the next event that occurs is reaction j in interval $[t + \tau, t + \tau + d\tau)$. If we can draw from this distribution then, given an initial $s(t = 0)$ (or distribution over it) we have a process that allows us to simulate successive chemical reactions. Given reaction j is drawn we can update the vector $s(t) \to s(t + \tau)$ appropriately (just as we did with the nuclear decays).

Our objective in the following is to find a simple expression for $p(j, \tau | s(t), t)$.

We've used the notation that $S(t)$ represents a random variable and $s(t)$ a sample from it.

Deriving the Stochastic Simulation Algorithm

This is after Higham [3] but also see [4] and [5].
Define $P_0(\tau|s,t)$ as the probability that no reaction takes place in the time interval $[t, t+\tau)$. If we have M reactions:
$P_0(\tau+\delta\tau|s,t) = P_0(\tau|s,t)(1-\sum_{j=1}^{M} a_j(s)d\tau)$.
Rearranging we have

$$\frac{P_0(\tau+\delta\tau|s,t)-P_0(\tau|s,t)}{d\tau} = -a_{sum}(s)P_0(\tau|s,t)$$

where $a_{sum}(s) \equiv \sum_{j=1}^{M} a_j(s)$. Noting $P_0(0|s,t) = 1$ (because nothing has happened at t) we can let $d\tau \to 0$ and solve the corresponding ODE and obtain the solution:
$P_0(\tau|s,t) = exp(-a_{sum}(s)\tau)$.

Deriving the Stochastic Simulation Algorithm II

Given $P_0(\tau|s,t) = exp(-a_{sum}(s)\tau)$ and recalling that $p(j,\tau|s(t),t)$ is the joint probability on when and what the next reaction is we can write:

$p(j,\tau|s(t),t)d\tau = P_0(\tau|s,t)a_j(s)d\tau$

The probability that nothing happens up to time $t+\tau$ and then event j occurs in the interval $[t+\tau, t+\tau+d\tau)$. From the above we have:

$p(j,\tau|s(t),t) = exp(-a_{sum}(s)\tau)a_j(s) =$

$\frac{a_j(s)}{a_{sum}(s)}a_{sum}(s)exp(-a_{sum}(s)\tau) =$ prob event $j\times$ prob event at tau.

Stochastic Simulation Algorithm

We found that: $p(j, \tau | s(t), t) = exp(-a_{sum}(s)\tau)a_j(s) = \frac{a_j(s)}{a_{sum}(s)} a_{sum}(s)exp(-a_{sum}(s)\tau) =$ prob event $j \times$ prob event at tau.
The distribution neatly factorizes (no surprise since we're dealing with independent events).
So to simulate the next chemical event we draw τ from $a_{sum}(s)exp(-a_{sum}(s)\tau)$ and to find out what it is we use the probability $\frac{a_j(s)}{a_{sum}(s)}$. This is called the Gillespie algorithm [5]. This approach exactly what we found from our intuitive treatment of nuclear decay.

Gillespie Algorithm

We found that: $p(j, \tau | s(t), t) = exp(-a_{sum}(s)\tau)a_j(s) = \frac{a_j(s)}{a_{sum}(s)} a_{sum}(s)exp(-a_{sum}(s)\tau) =$ prob event $j \times$ prob event at tau.

1. Given $s(t)$ and the set $a_j(s(t)) \forall j$ find $a_{sum}(s) \equiv \sum_{j=1}^{M} a_j(s)$.
2. Find the time $t + \tau$ of the next event by drawing from $a_{sum}(s)exp(-a_{sum}(s)\tau)$.
3. Find the identity of the event by drawing from the discrete distribution where the probability of the j^{th} reaction is $\frac{a_j(s)}{a_{sum}(s)}$
4. Update $s(t)$ to $s'(t + \tau)$ by looking up the effect of reaction j in the stoichiometry matrix. Repeat.

Co-operative effects

One way in which co-operative effects occur is when the binding of one molecule to an enzyme modulates the rate at which subsequent ones are bound. A classic example of this is the Monod Wyman Changeux Model (their '65 paper is worth a read).
We consider an enzyme having two states: an activated form B and a not-activated form N. I'll explain the notation below.

$$B_0 \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} N_0$$

We suppose that our enzyme binds a molecule s which then affects the binding of subsequent molecules in a manner *conditional* on whether it is in state N or B and depending on how many molecules have previously been bound. We suppose that the N, activated, state is faster at sequestering s than the non-activated state. We suppose that our enzyme (whether in N state or B) has two binding sites (The MWC-model is typically expressed for n modifications rather than 2).

Co-operative effects II

Toggle between activated and non-activated state: $B_0 \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} N_0$

Molecule s binding in the non-activated state:

$N_0 \underset{k_{-1}}{\overset{2sk_1}{\rightleftharpoons}} N_1$

$N_1 \underset{2k_{-1}}{\overset{sk_1}{\rightleftharpoons}} N_2$

and then the B state has a similar set of reactions:

$B_0 \underset{k_{-3}}{\overset{2sk_3}{\rightleftharpoons}} B_1$

$B_1 \underset{2k_{-3}}{\overset{sk_3}{\rightleftharpoons}} B_2$

We'll be looking at these reactions in the practical - but there's slightly different notation - this reflects the variability across texts and is worth knowing.

Co-operative effects III

Let's use n_i and b_i for $i \in \{1, 2, 3\}$ to denote amounts of molecule types.

Consider:

$N_0 \underset{k_{-1}}{\overset{2sk_1}{\rightleftharpoons}} N_1$

If our system is in equilibrium then the flux to the left has to equal the flux to the right:

$n_0 2sk_1 = n_1 k_{-1}$ or $n_1 = 2sK_1 n_0$ where $K_i = k_i / k_{-i}$.

We can repeat this for each of the reactions that we considered.

We can also write an expression for the fraction of sites on enzymes that are occupied with s molecules:

$$\frac{n_1 + 2n_2 + b_1 + 2b_2}{2(n_0 + n_1 + n_2 + b_0 + b_1 + b_2)}$$

Co-operative effects III

Inserting expressions like $n_1 = 2sK_1 n_0$ into the below

$$\frac{n_1 + 2n_2 + b_1 + 2b_2}{2(n_0 + n_1 + n_2 + b_0 + b_1 + b_2)}$$

we can obtain an expression for the proportion of bound sites that is only in terms of s:

$$\frac{sK_1^{-1}(1 + sK_1^{-1}) + K_2^{-1}[sK_3^{-1}(1 + sK_3^{-1})]}{(1 + sK_1^{-1})^2 + K_2^{-1}(1 + sK_3^{-1})^2}.$$

If we had had n binding sites (rather than two) then the expression becomes:

$$\frac{sK_1^{-1}(1 + sK_1^{-1})^{n-1} + K_2^{-1}[sK_3^{-1}(1 + sK_3^{-1})^{n-1}]}{(1 + sK_1^{-1})^n + K_2^{-1}(1 + sK_3^{-1})^n}.$$

The system can show a sensitive response to the amount of s. We will be exploiting this in the following.

What we've covered

Gentle introduction to Poisson point processes
Basics of chemical reactions
Stochastic simulation introduction
Co-operativity effects in chemistry

Bibliography

[1] E. Libby, T.J. Perkins, and P.S. Swain (2007) Noisy information processing through transcriptional regulation, PNAS 104, 7151.

[2] T.J Kobayashi and A. Kamimura (2011) Dynamics of intracellular information decoding, Phys. Biol. 8, 055007.

[3] D. Higham (2008). Modeling and Simulating Chemical Reactions, SIAM Review 50, 347.

[4] D. J. Wilkinson (2011) Stochastic Modelling for Systems Biology, Chapman and Hall/CRC 2nd Ed.

[5] D. T. Gillespie (1977) Exact Stochastic Simulation of Coupled Chemical Reactions, J. Phys. Chem. 81 2340.