Stochastic Gene Expression Project: Homework 2
Due January 25th at 2:15 PM

Problem 1: Quasi steady state approximation warm-up.

We are interested in studying the stochastic dynamics of a gene which regulates itself. But to simplify the problem, we make two commonly used approximations. The first is to assume that there are so many proteins that protein concentration (the number of proteins in the cell divided by the cell's volume) is effectively a continuous variable; this allows us to work with a continuous stochastic process (using SDEs and the Fokker-Planck equation) rather than the discrete and hard to solve Chemical Master Equation.

The second thing is, since mRNA reaches equilibrium much more quickly than proteins do (the timescale $1/d_m \approx 10$ hours, where $1/d_p \approx 30$ hours; see here for more info), we assume that mRNA stays at its equilibrium concentration.

In this problem, you will work out the quasi steady state (QSS) approximation for a single unregulated gene. The reactions defining this system are:

$$g \xrightarrow{k_m} g + m$$
$$m \xrightarrow{k_p} m + p$$
$$m \xrightarrow{d_m} \varnothing$$
$$p \xrightarrow{d_p} \varnothing$$

(a) List each chemical species in this model (remember that g does not change, and so is a constant rather than a species). How many are there?

(b) Write out the SDEs corresponding to each species.

(c) What is the steady state value of m?

(d) Write out the new protein SDE.

Problem 2: Deriving the protein SDE for second order regulation.

In this problem, you will work out the protein SDE for a gene with second order regulation (the gene's protein can bind to the gene at most two times). The reactions defining this system are:

$$
\begin{aligned}
g^0 + p &\xrightarrow{b_{01}} g^1 \qquad & g^0 + p &\xleftarrow{b_{10}} g^1 \\
g^1 + p &\xrightarrow{b_{12}} g^2 \qquad & g^1 + p &\xleftarrow{b_{21}} g^2 \\
g^2 + p &\xrightarrow{b_{23}} g^3 \qquad & g^2 + p &\xleftarrow{b_{32}} g^3
\end{aligned}
$$

$$
\begin{aligned}
g^0 &\xrightarrow{k_0} g^0 + m \\
g^1 &\xrightarrow{k_1} g^1 + m \\
g^2 &\xrightarrow{k_2} g^2 + m
\end{aligned}
$$

$$
\begin{aligned}
m &\xrightarrow{k_p} m + p \\
m &\xrightarrow{d_m} \varnothing \\
p &\xrightarrow{d_p} \varnothing
\end{aligned}
$$

(a) List each chemical species in this model. How many are there?

(b) Write out the SDEs corresponding to m and p.

(c) What is the steady state value of m?

(d) Write out the new protein SDE.

(e) If all of the gene-protein binding reactions are at equilibrium, how are g^0 and g^1 related? How about g^1 and g^2?

(f) Use what you found in (e) to write g^1 and g^2 in terms of g^0 and p. You may want to define parameters $c_1 := (b_{01}/b_{10})$ and $c_2 := (b_{01}/b_{10})\,(b_{12}/b_{21})$ to ease notation.

(g) The total number of gene sites is fixed (let's call it G), so $g^0 + g^1 + g^2 = G$. Use this constraint to find an expression for g^0 in terms of p.

(h) Write out the final protein SDE.

Problem 3: Deriving the protein SDE for Nth order regulation.

This problem is the same as the previous one, but with Nth order regulation. The reactions defining this system are:

$$g^0 + p \xrightarrow{b_{01}} g^1 \qquad g^0 + p \xleftarrow{b_{10}} g^1$$
$$g^1 + p \xrightarrow{b_{12}} g^2 \qquad g^1 + p \xleftarrow{b_{21}} g^2$$
$$g^2 + p \xrightarrow{b_{23}} g^3 \qquad g^2 + p \xleftarrow{b_{32}} g^3$$
$$\cdots$$
$$g^{N-1} + p \xrightarrow{b_{N-1,N}} g^N \qquad g^{N-1} + p \xleftarrow{b_{N,N-1}} g^N$$
$$g^0 \xrightarrow{k_0} g^0 + m$$
$$g^1 \xrightarrow{k_1} g^1 + m$$
$$g^2 \xrightarrow{k_2} g^2 + m$$
$$\cdots$$
$$g^N \xrightarrow{k_N} g^N + m$$

$$m \xrightarrow{k_p} m + p$$
$$m \xrightarrow{d_m} \varnothing$$
$$p \xrightarrow{d_p} \varnothing$$

(a) List each chemical species in this model. How many are there?

(b) Write out the SDEs corresponding to m and p.

(c) What is the steady state value of m?

(d) Write out the new protein SDE.

(e) If all of the gene-protein binding reactions are at equilibrium, how are g^i and g^{i+1} related?

(f) Use what you found in (e) to write each g^i in terms of g^0 and p. You may want to define parameters c_i to ease notation.

(g) The total number of gene sites is fixed, so $g^0 + g^1 + \cdots + g^N = G$. Use this constraint to find an expression for g^0 in terms of p.

(h) Write out the final protein SDE.

Problem 4: Stochastic simulation in two dimensions.

Look at the file `bruteforce_solver_2D.py`, which is in the Github folder labeled `week_jan_21`. Given the 1D brute force solver code, and your 2D Euler-Maruyama code from last time, finish the code for the 2D brute force solver in that file.