Linear gene networks

C.W. Seitz

March 18, 2022

1 Definitions

Consider a set of stochastic linear differential equations which govern the concentration of RNA x_i and protein y_i associated with a particular gene:

$$\dot{x_i} = \sum_j m_{ij} y_j - \alpha_i x_i + \sigma_x \eta_i^x$$
$$\dot{y_i} = r_i x_i - \beta_i y_i + \sigma_y \eta_i^y$$

where m_{ij} is the linear effect of a transcription factor j on gene i and therefore represents the rate of transcription. Similarly, r_i represents the rate of translation from RNA to protein. The constants α_i and β_i are the RNA degradation and protein degradation rates, respectively. Since these equations are stochastic, the gold-standard is to find a probability density $P_\theta(\boldsymbol{x}, \boldsymbol{y}; t)$. In higher dimensions, this is a very difficult problem and we may need to settle for finding the stationary distribution $P_\theta^0(\boldsymbol{x}, \boldsymbol{y})$, if it exists and finding its solution is tractable. To do this for the set of equations above, a fundamental assumption is that the rate of translation and protein degradation approximately balance and the timescale of of changes in protein concentration is very long. Intuitively, if y_j's are constant, growth or decay of the first term will eventually balance with the $-\alpha_i x_i$ term and the system will approach a fixed point.

If indeed changes in protein concentration are negligible over the course of evolution of the system i.e., $\dot{y_i} \approx 0$ we have that

$$y_i = \frac{r_i x_i - \sigma_x \eta_i^y}{\beta_i}$$

This assumption allows us to rewrite the system above as a single equation per gene

$$\begin{aligned}\dot{x_i} &= \sum_j m_{ij} \frac{r_j x_j - \sigma_x \eta_j^y}{\beta_j} - \alpha_i x_i + \sigma_x \eta_i^x \\ &= \sum_j \left(\gamma_{ij} x_j - \theta_{ij} \eta_j^y\right) - \alpha_i x_i + \sigma_x \eta_i^x\end{aligned}$$

If we additionally assume that $\sigma_x << r_i x_i$

$$\dot{x_i} = \sum_j \gamma_{ij} x_j - \alpha_i x_i + \sigma_x \eta_i^x$$

which is the multivariate Ornstein-Uhlenbeck process. This can then be written as an Ito stochastic differential equation

$$dx = -\Gamma x \, dt + \Sigma dW$$

For example, when $N = 3$ we have the following matrix Γ

$$\Gamma = \begin{bmatrix} \gamma_{11} - \alpha_1 & \gamma_{21} & \gamma_{31} \\ \gamma_{12} & \gamma_{22} - \alpha_2 & \gamma_{32} \\ \gamma_{13} & \gamma_{23} & \gamma_{33} - \alpha_3 \end{bmatrix}$$

The SDE given above corresponds to the Kramers-Moyal expansion (KME) of a transition density $T(x', t'|x, t)$

$$\frac{\partial P}{\partial t} = \sum_{n=1}^{\infty} \frac{1}{n!} \left(-\frac{\partial}{\partial V}\right)^n [M_n(V,t) P(V,t)] \tag{1}$$

where M_n is the nth moment of the transition density. In the diffusion approximation, the KME becomes the Fokker-Planck equation (Risken 1989)

$$\begin{aligned}\frac{\partial P}{\partial t} &= \left(-\sum_{i=1}^{N} \frac{\partial}{\partial x_i} S_i + \sum_{i,j=1}^{N} D_{ij} \frac{\partial^2}{\partial x_i \partial x_j}\right) P & (2) \\ &= \mathcal{L}_{FP} P & (3)\end{aligned}$$

S_i is the drift vector and D_{ij} the diffusion matrix. The second term is the diffusion term which is volatile over time. Incorporating the quantities $S = \Gamma x$ and $2D = \Sigma\Sigma^T$ into this form for the OU process,

$$\frac{\partial P}{\partial t} = \left(-\sum_{i=1}^{N} \frac{\partial}{\partial x_i} \Gamma x + \frac{1}{2} \sum_{i,j=1}^{N} (\Sigma\Sigma^T)_{ij} \frac{\partial^2}{\partial x_i \partial x_j} \right) P \tag{4}$$

$$\tag{5}$$

A distribution π is the stationary distribution (equilibrium distribution) of P if $\mathcal{L}_{FP}\pi = 0$. Such a system is said to obey *detailed balance*, in which the Fokker-Planck operator leaves the distribution invariant. Qualitatively, this means that, at equilibrium, the probability current out of an infinitesimal volume dV in the state space Ω is balanced by an equal and opposite current into dV.