• CCBS. 2025 Fall. Lecture 5. Adaptation Bioreactions and Noise in bio.

⓪. Review of Last time.

Bioreaction I: Adaptation bioreactions. Cont'd.

① Integral feedback in biology as engineered systems.

② Incoherent feedforward reveals nonlinear structure of biology.

Noise in biological systems

③. A brief intro. — where it started.
— noise as unknown mechanism
— Estimate

④. Stochasticity in chemical reaction networks.
⇒. Chemical master equation. Distribution as Traj.

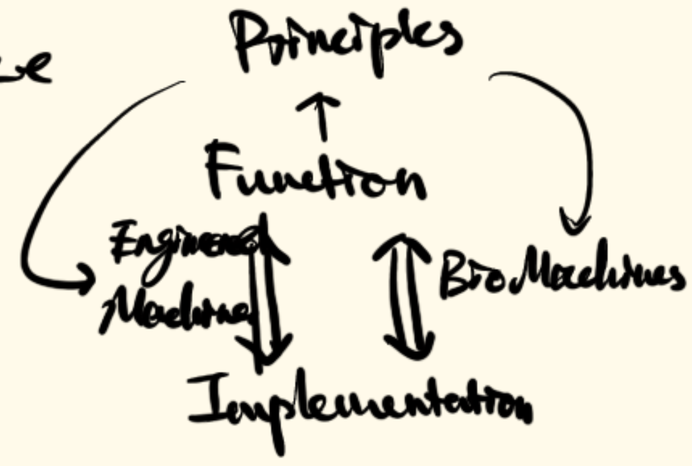
⑤. Gillespie algorithm.

⑥. Noise analysis at steady state.

⑦. Some stochastic phenomena. — bursty gene expression
— multimodality
— Distr. as. Trajectory.
— Ergodicity.

⑥. Last time.

- Biomachine. — a perspective we can take on understanding biological systems.



Treat it like a machine.

Biological systems perform various functions.

just like engineered systems. i.e. machines.

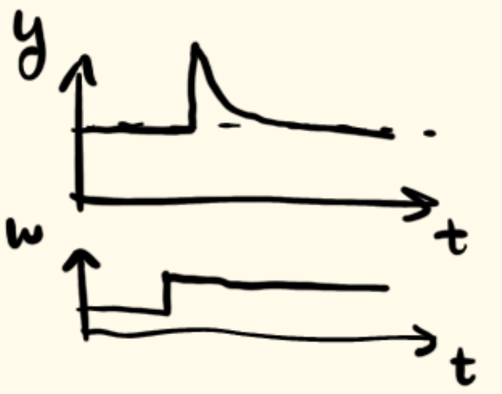
So principles on how to design machines with these functions can be used to understand why biological systems are designed.

But biological systems implement the same principles using different physical components. So there are comparisons and contrasts.

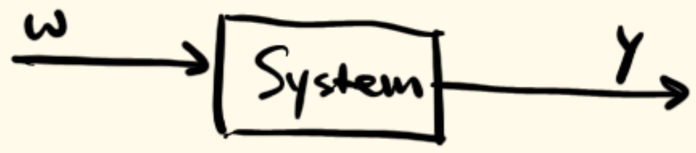
Even new theories and principles from biology back to engineering

- Adaptation — The function we started with.

Adaptation is a hallmark for biological systems. and autonomous machines we built.



- In control theory.

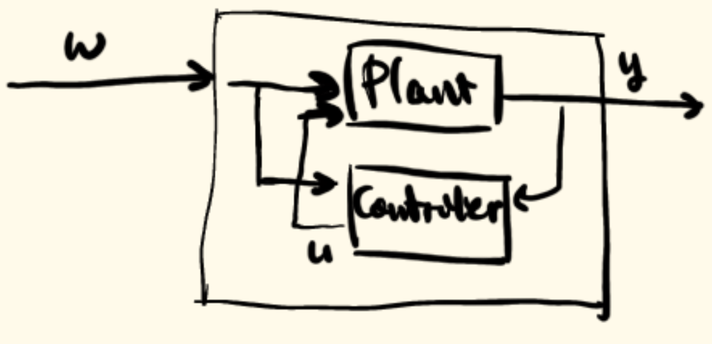


$$\dot{x} = f(x, w)$$
$$y = h(x, w)$$

Perfect adaptation means. $y \to 0$ as $t \to \infty$.

for all constant w.

- Controller design problem.

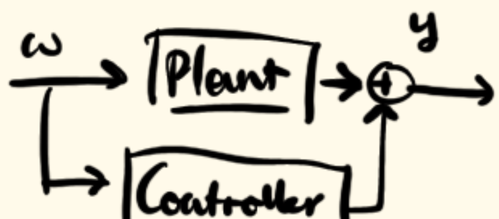


$$\begin{pmatrix} \dot{x}_k = f_k(x_k, w, y) \\ u = h_k(x_k, w, y) \end{pmatrix}.$$

$$\dot{x} = f(x, u, w)$$
$$y = h(x, w).$$

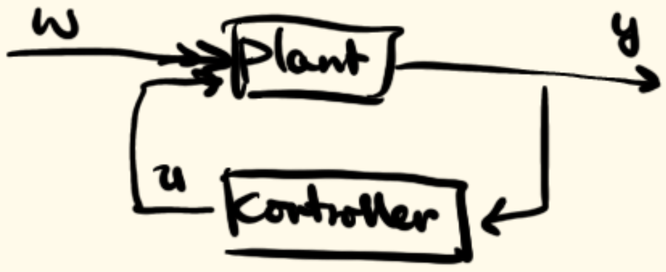
$w, y \to u$.

Feedforward



$w \to u$

Feedback



$y \to u$.

- From a simple example. we saw.

(1) Feedforward requires perfect matching of parameters. Can't do adaptation.

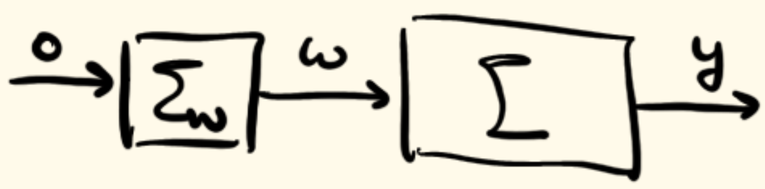
(2) Feedback uses the power of time, observes the past, act accordingly.

- We also saw, naturally, a PID controller emerge, can achieve perfect adaptation

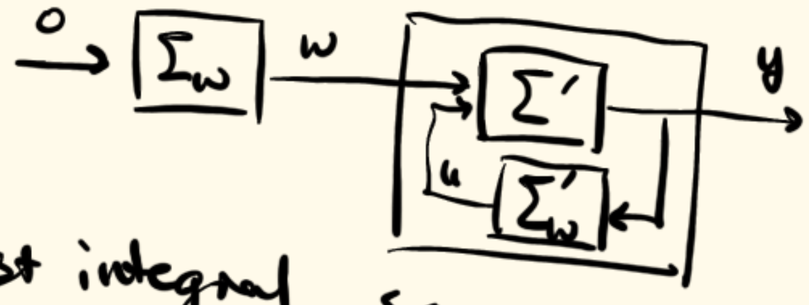
$$u = k_p y + k_I \int y + k_D \dot{y}.$$

- Which part of PID is responsible for adaptation?

Internal model principle.



Adapt to all w from Σ_w $\Leftrightarrow$ Internal model of Σ_w.

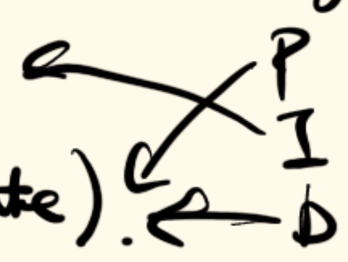


Perfect adaptation. Σ_w is just integral, so

Perfect adaptation $\Leftrightarrow$ Integral Feedback.

(+ stable steady state)

e.g.



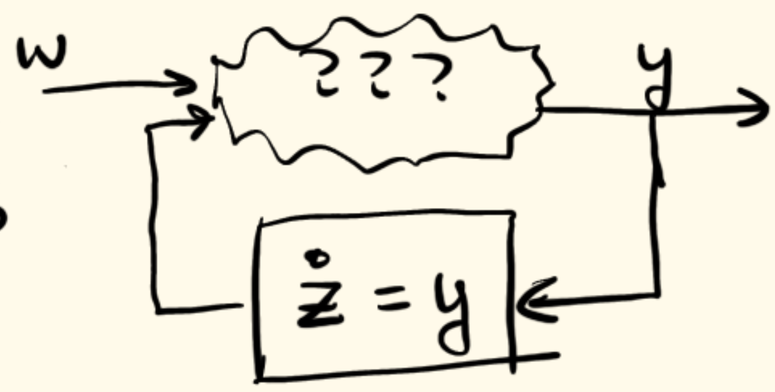
- Integral feedback is sufficient.

$u = k_I \int y \Rightarrow \dot{u} = k_I y \Rightarrow$ At s.s. $y = 0$.

But requires stability!

- This is powerful:

Adaptation regardless of the plant!



(as long as the system can reach steady state.)

Pretty much independent of everything!

(Look back at the simple example with only integral feedback.)

Observation prevents perfect adaptation.

- Today: Implementations of adaptation in bio vs electronic

And special strategies of bio!

① . Implementation of adaptation in engineered systems

- Recall our goal: use machine control principles to understand biology. (And in biology.)
- Control theory tells us that Perfect adaptation $\Leftrightarrow$ integral feedback

But how to implement integral feedback?

i.e. physically make it happen.

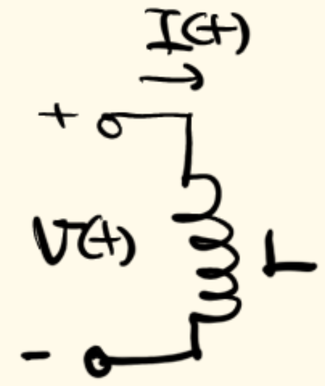
- Electrical circuits: variables are voltages V, currents I.

Resistors. 电阻 $V = RI$. $I = \frac{1}{R}V$.

This is proportional control.

Inductor. 电感. $V(t) = L\frac{dI(t)}{dt}$.

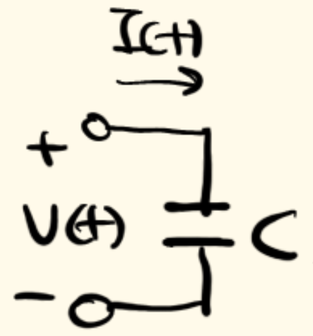
$$I(t) = I(0) + \frac{1}{L}\int_0^t V(\tau)d\tau$$



Integral or Derivative.

Capacitor. 电容. $V(t) = V(0) + \frac{1}{C}\int_0^t I(\tau)d\tau$

$$I(t) = C\frac{dV(t)}{dt}$$



Integral or Derivative.

So, PID control is very natural in electrical circuits!

- Exercise: show how PID can be done in mechanical systems.

(mass, spring, dampers...)

- But what about in biological reaction networks? (BRNs)

  — variables are concentrations, all positive. (rule [P])

  — dynamics are, by default, (in contrast to V, I. or position, velocity)
  constrained polynomials. (Rule [CP]).

$$\frac{dx}{dt} = \Gamma \Lambda_k x^{\alpha}.$$

Proportional : e.g. $u = k_p y = k_p (x_r - x)$.

often easy... $\frac{dx}{dt} = \underbrace{\mu - \nu x}_{u}$. $\mu = k_p x_r$. $\nu = k_p$.

in fact. any static regulation $u(y)$ works as proportional feedback... (upon linearization)

(another example: $y = \frac{x}{x_r}$ (more natural!)

$u = k_p y = \frac{k_p}{x_r} x$. $\Rightarrow \frac{dx}{dt} = \mu - \underbrace{\nu x}_{u}$. $\mu, \nu = \frac{k_p}{x_r}$)

Derivative : can often be done,
but not often used.
since biomolecular processes are noisy.
and derivative control amplifies noise...
(not used much in highly noisy scenarios...).

Integral — important for perfect adaptation. (if and only if).
but not naturally achieved!
because of rule [CP].

$$z = \int y \, dt = \int x_r - x \, dt.$$

$\Rightarrow \dot{z} = x_r - x$. z can't be a species due to rule [CP]!

— To achieve integral control in biology.

needs further sophistication!

Strategy 1: virtual variable.

e.g. $\dot{z}_1 = \mu - C$

$\dot{z}_2 = x - C$

This is called antithetic integral controller. Cell systems 2016. Nature 2019. Mustafa Khammash group.

$z_1 + z_2 \rightleftharpoons C \to \emptyset$.

$\Rightarrow$. Let $z = z_1 - z_2$. then $\dot{z} = \mu - x$.

(Recall. this is like the dual rail strategy in last lecture to circumvent rule [CP].

but the problem is that rule [PU]

makes it hard to directly use z

in other variable's dynamics.

e.g. $\dot{x} = kz - \gamma x = kz_1 - kz_2 - \gamma x$

this requires kz_1 and kz_2 have perfectly matching parameters. violate rule [PU]!

So. the best we can do is $\dot{x} = kz_1 - \gamma x$

for example.

This doesn't hurt Perfect Adaptation.

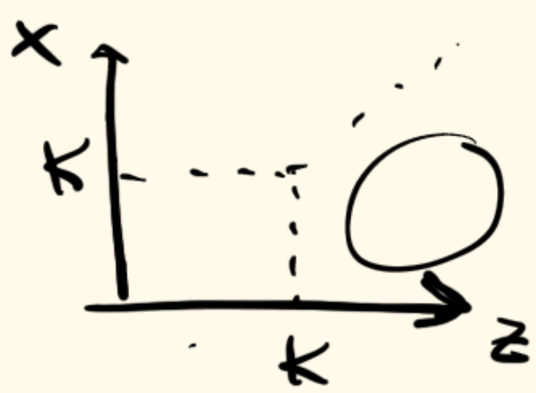
since the integral variable still exists.)

Strategy 2: constrain to certain regimes.

e.g. $\dot{z} = \mu - C$. $\quad x_f + z_f \overset{K}{\rightleftharpoons} C \to x_f$

In regime $C \sim x$.

we have $\dot{z} \approx \mu - x$.

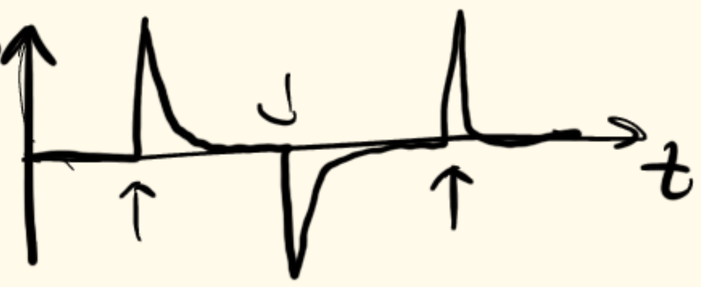


– Summary: because of rule [CP] and positivity,

BRN doesn't naturally implement integral feedback.

To achieve this. BRN uses more sophisticated (various)
strategies to indirectly achieve integral feedback.
or partially (regimes).

e.g. bacterial chemotaxis achieves perfect adaptation via sophisticated phosphorylation cascades!



HERE COMES THE TWIST..

② Incoherent Feedforward loops (IFFL) reveals nonlinear structure in biology.

- Previously. we took the machine perspective on biological systems.

In particular. since adaptation is a hallmark behavior in biology. we figured as an adaptation machine, biology must achieve adaptation, following similar design principles as engineered adaptation machines.

Internal model principle in control theory says. perfect adaptation ⇔ integral feedback.

And we figured due to physical constraints on biomolecular reactions, BRNs must achieve Perfect adaptation by implementing integral feedback

by more sophisticated strategies.

- But This is NOT an accurate picture of what's happening in reality!

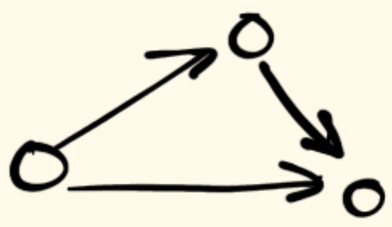
In 2002. Shen-Orr et.al. Nature. Genetics (Uri Alon) and Milo.et.al.2002.

looked at all 3-node network topologies

on transcription factors and operons in E.coli.
(115) (424)

And found network motif (i.e. more frequent than "random"...)

to be feedforward!



- But recall. perfect adaptation

can't be achieved by feedforward

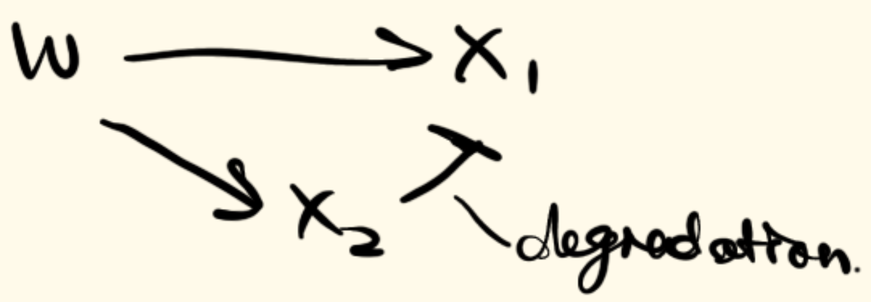
in engineered systems due to the requirement of

perfect matching of two processes' parameters!

How come biology is different?

what happened to rule PU?

- Example. (Snitker):



$$W \xrightarrow{k_1} W + X_1$$

$$W \xrightarrow{k_2} W + X_2$$

$$X_2 + X_1 \xrightarrow{\delta} X_2 \xrightarrow{\gamma} \phi$$

$$\frac{dx_1}{dt} = k_1 W - \delta x_1 x_2$$

$$\frac{dx_2}{dt} = k_2 W - \gamma x_2.$$

At steady state. $x_2^* = \frac{k_2}{\gamma} w$.

$$\Rightarrow . \; x_1^* = \frac{k_1 w}{\delta x_2} = \frac{k_1 w}{\delta \frac{k_2}{\gamma} w} = \frac{k_1 \gamma}{k_2 \delta}.$$

So. x_1^* is invariant to w! (Perfect adaptation)

And this invariance is independent of the parameters! No perfect matching required!

- The key lies in the exponents.

  i.e. the reaction order. ($v = k x^{\alpha}$)

  (which is α in mass action.
  but could be more complicated..
  with time-scale separation.)

(HW problem looks at this deeper.)

This is the nonlinear structure that biological systems can utilize, which is not present (or utilized) in engineered systems, since they are often linearized etc...

- But where is the integral variable?

$$z = k_2 x_1 - k_1 x_2.$$

$$\Rightarrow . \; \dot{z} = k_1 k_2 w - \delta k_2 x_1 x_2$$

$$- k_1 k_2 w + \gamma k_1 x_2$$

$$= x_2 (\gamma k_1 - \delta k_2 x_1)$$

New Topic : Noise in biological systems.

③ Intro.

• The biomachine perspective on biology we took

was a systems, engineering perspective.

It underlies the field of systems and synthetic biology

that started in 2000.

With the completion of the human genome project,

we had a glance at "all components in a cell",

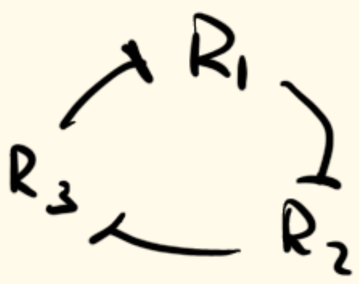
So we could start to reason about it as a system,

and engineer it like a machine. (Like an electrical circuit)

⇒. The phrase. toggle switch.

Repressilator. Elowitz 2000 Science. Was one of first work

pioneering this perspective. (reading in HW 1).



E.coli

Three genes' expression oscillates.

Limitations of theory for design at the time...

(ours based on regimes doesn't make Hill-type kinetics assumptions.)

But another big observation: Noise! between cells.

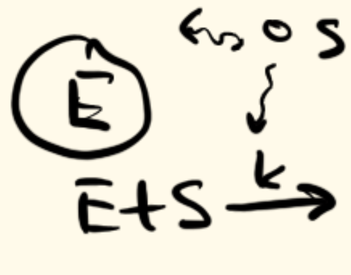
and within one cell.

under the microscope.

- But <u>where</u> does noise come from, and <u>when</u> does it matter?

— E.coli cells from the same clone. so not genetic difference.

⇒ Origin of noise is in <u>stochasticity of cellular processes</u>
Namely. chemical reactions.

— 

Reaction events happen by molecules encountering each other.

What's distr. for the # of reaction events in a time interval Δt ?

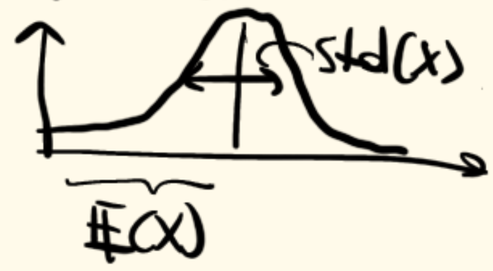
- Reactions happen when E and S hits
- Reaction rate is constant $a = kES$. unless E, S changes.
- The events are independent of each other.
they only depend on E and S 's conc. and k

⇒. Poisson distribution with parameter $\lambda = a\Delta t$, average # events in this interval

chemical reaction rate. $a = kES$.

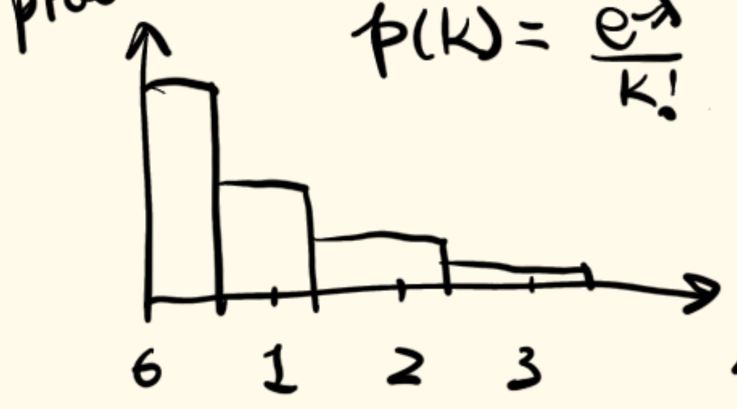
$X \sim \text{Poisson}(\lambda)$: $\mathbb{P}\{X = k\} = \frac{\lambda^k}{k!} e^{-\lambda}$

$\mathbb{E}(X) = \text{Var}(X) = \lambda$ ⇒. $CV = \frac{\text{Var}(X)}{\mathbb{E}(X)^2} = \frac{1}{\sqrt{\lambda}}$ coeff. of variation

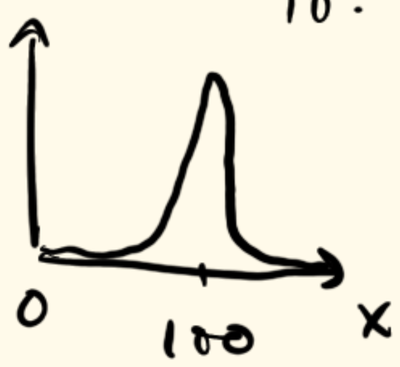


$\lambda = 1$. $CV = 1$.

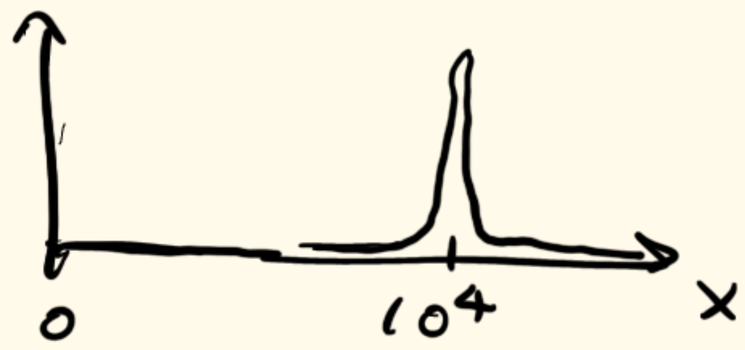
$p(k) = \frac{e^{-\lambda}}{k!}$



$\lambda = 100$. $CV = \frac{1}{10}$.



$\lambda = 10000$ $CV = \frac{1}{100}$.



Process noise becomes important when

events is small !

- Translates to molecule numbers?

events → production events.

in balance with dilution by growth. (i.e. no active degradation)

Then. if N_x number of molecule X in cell.

$\Rightarrow$ N_x produced every generation.

If one molecule X produced every event

$\Rightarrow$ N_x events.

$N_x \sim 10^2$ to 10^3 for proteins in E.coli CV ~ 0.1

$N_x \sim 10^5$ to 10^6 for metabolites in E.coli. CV ~ 10^{-3}.

So. noise typically don't matter. even in E.coli.

(Eukaryotes? Even less so,...)

(Active degradation? More events. less noise.)

- What's missing? Burstyness in gene expression.

We do see a big heterogeneity in single cells under the microscope!

Several proteins produced per event.

amplified by Tx-TL process.

gene → mRNA → protein.

If 100 proteins per event.

$\Rightarrow$ $N_{event} \sim \frac{N_x}{100} \sim$ 1 to 10 for proteins in E.coli.

So. noise start to matter.

- "Critical threshold" — burstyness $\approx$ typical # proteins in cells

So. low expression → noise dominates.

high expression → noise doesn't matter.

— Could be a strategy for bacteria.

e.g. bet hedging ... utilizing noise ..

yet noise's role / importance is tunable.

— Despite burstyness. noise doesn't play a big role in most biological processes!

For noise to be important. need following factors. (typically)

- small cell size (e.g. not Eukaryotes)
- low flux. (not much degradation).
- low concentration. (e.g. genome.)
- large molecule size. (e.g. not metabolites)
- burstyness. (e.g. not metabolic reactions.)

CAUTION!

— Comment : Noise { Stochasticity (in processes)
Unknown (mechanisms).

We are talking about noise from process stochasticity

intrinsic randomness of underlying mechanisms. Like diffusion quantum mechanics.

here. not = a "wide spread in data" that also includes unknown mechanisms.

e.g. a deterministic process can look "noisy" in data. just because we don't know the mechanism.

(Like whether you bring umbrella to school. if you don't observe the weather.)

Noise in data is a completely different topic. because it's more about inference of the unknown mechanisms. not about stochasticity

④. Chemical Master Equation.

— Stochasticity in chemical reactions.

- Now we have some confidence on the role noise plays in biological systems. We would like a way to formally describe and analyze them.

- Consider a simple reaction system. (elementary)

$$\emptyset \xrightarrow{\mu} X \qquad X \xrightarrow{k} \emptyset$$

We could write this in net-change form.

$$X \overset{v_1}{\rightsquigarrow} X+1 \qquad v_1 = \mu.$$

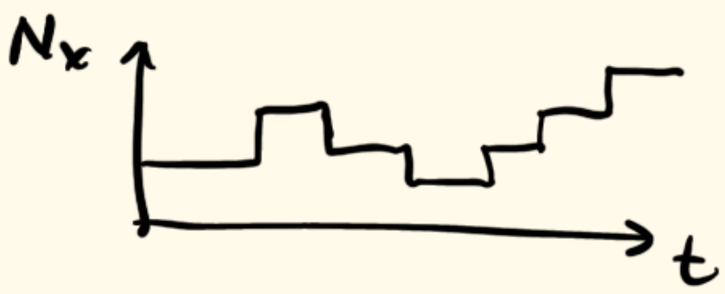
$$X \overset{v_2}{\rightsquigarrow} X-1 \qquad v_2 = kX$$

(Deterministic) rate eqn: $\frac{dX}{dt} = \mu - kX$

Stochastic case. N_X is discrete. $N_X = 0, 1, 2, \ldots$

it's a random variable that changes over time.

so, a stochastic process.



So its dynamics is described in terms of probabilities.

$$p_n(t) = \mathbb{P}\{ N_X(t) = n \}.$$

- What's the dynamics of $p_n(t)$? How does reactions change it?

$X \overset{v_2}{\rightsquigarrow} X-1$ says. $\mathbb{P}\{\text{reaction happen in } [t, t+dt)\}$

$v_2 = k N_X$

$= v_2(t)\, dt$

$= k N_X(t)\, dt.$

So. just consider this one reaction. It's.

$$P_n(t+dt) = P_n(t)\,\mathbb{P}\{\text{no reaction in } [t,t+dt)\} + P_{n+1}(t)\,\mathbb{P}\{\text{reaction in } [t,t+dt)\}$$

($dt \to 0$. so either one reaction or no reaction.)

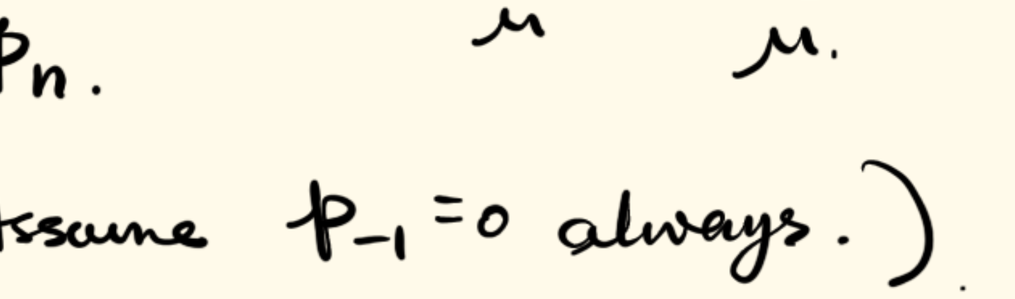
$$= P_n(t)(1 - k n\,dt) + P_{n+1}(t)\,k(n+1)\,dt$$

$$= P_n(t) + k(n+1)P_{n+1}(t)\,dt - kn\,P_n(t)\,dt$$

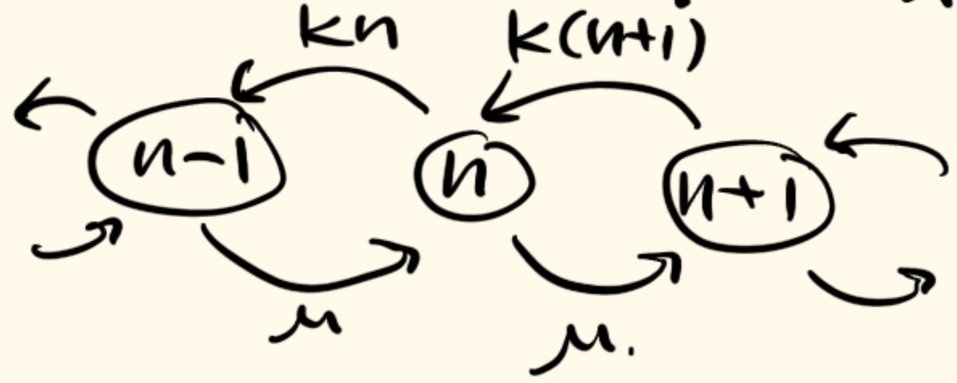
$$\Rightarrow \frac{P_n(t+dt) - P_n(t)}{dt} = k(n+1)P_{n+1}(t) - kn\,P_n(t).$$

$$\Rightarrow dt \to 0. \quad \frac{dP_n}{dt} = k(n+1)P_{n+1} - kn\,P_n.$$

$(n = 0, 1, 2, \ldots)$



— Now. we do the same for production. $\phi \xrightarrow{\mu} X$



$$\Rightarrow \frac{dP_n(t)}{dt} = k(n+1)P_{n+1} - kn\,P_n + \mu P_{n-1} - \mu P_n.$$

(for $n = 0, 1, 2, 3, \ldots$ Assume $P_{-1} = 0$ always.)

Chemical master equation. about distributions ...

...

Viewed as a dynamical system. this is infinite dimensional

But it has strong structure. so we can simulate/analyze accordingly.

⑤. Gillespie algorithm. (or. SSA — Stochastic simulation algorithm.)

- Simulation first. because. (Reference: Erban, Chapman, Maini. 2007 Practical guide to stoch. sim. of reaction diffusion processes)

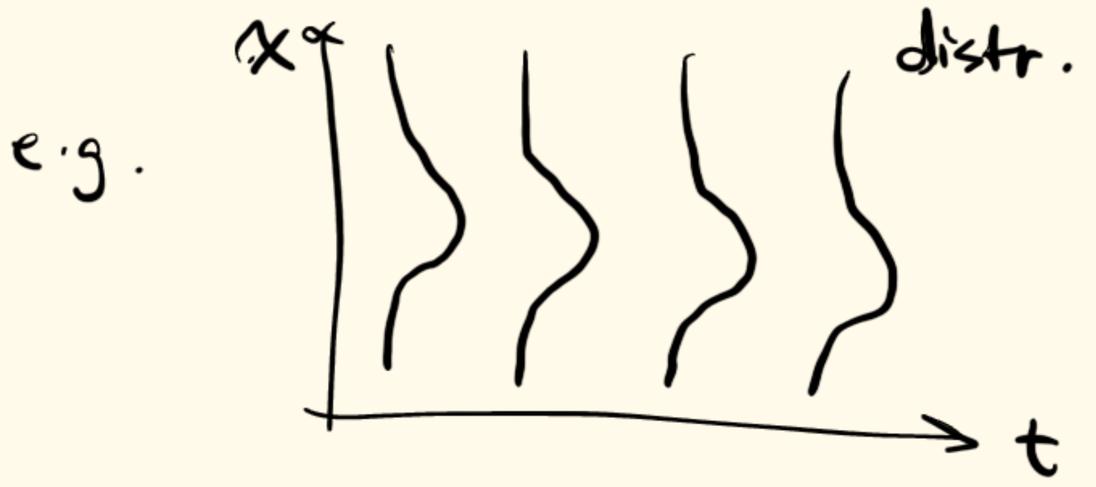
(1). Deterministic. already hard to analyze for general case.
This is much harder for stochastic case.

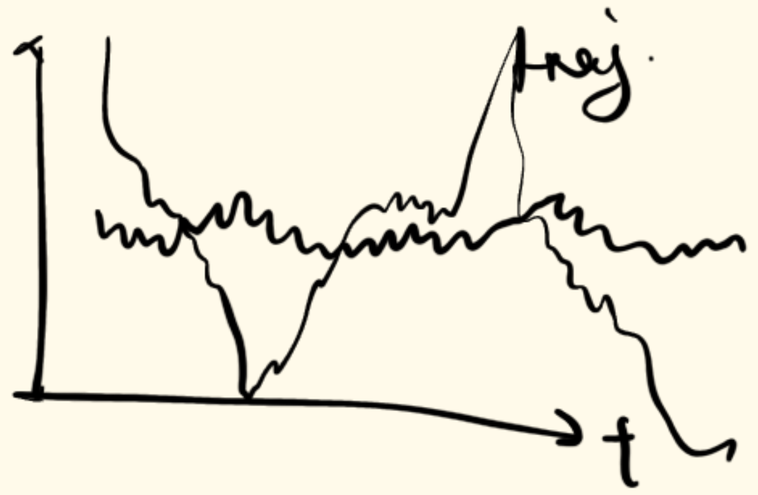
(2). Distribution vs Trajectory.

Exact analysis, even when doable, is often only possible for steady state distribution.
or. distribution dynamics.

But that's different from trajectory dynamics.

e.g.





(3). Usually analysis can be done after approximation. such as linearization.
Analysis can give the full picture. but approximate.
So. always helpful to check with simulations.

- Can't directly simulate the distribution. → inf. dim ODE. $P_n(t)$.
We can simulate trajectories. then distr. can be obtained from averages over lots of traj.

- Simplest idea. just like ODE sim. $\frac{dx}{dt} = f(x)$
(Euler integration) ⇒. $x(t+\Delta t) = x(t) + f(x(t))\Delta t$.

$X \xrightarrow{v_2} X-1$ $\quad v_2 = kx$

$$N_x(t+\Delta t) = \begin{cases} N_x(t) - 1 & \text{with prob } a_2\Delta t \\ & a_2 = k N_x(t) \\ N_x(t) & \text{with prob } 1 - a_2\Delta t. \end{cases}$$

But need Δt small to have a good approx...

Could be very costly.

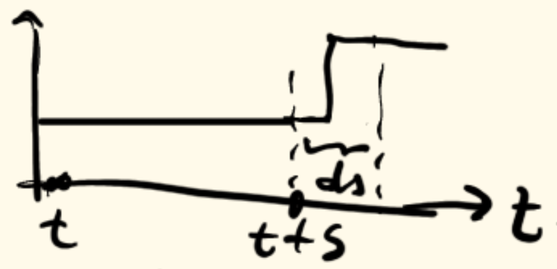
Could we simulate $N_x(t)$ exactly?

Yes. by transforming randomness from

coheather an event happens in an interval

to <u>when</u> does an event happen.

- Start at t, want τ s.t. $t+\tau$ is next reaction.



- Let $f(N_x(t), s)\,ds$ ← an infinitesimal time interval

pdf for τ.

$= \mathbb{P}\{$ $N_x(t)$ molecules at time t,
and the next reaction occurs in
time interval $[t+s, t+s+ds)$ $\}$

$g(N_x(t), s) = \mathbb{P}\{$ No reaction in interval $[t, t+s)\}$

Denote reaction with rate $v_\cdot(N_x)$.

(e.g. $v = kN_x$)

$\Rightarrow f(N_x(t), s)\,ds = g(N_x(t), s)\; v(N_x(t+s))\,ds$

$= N_x(t)$ since no reaction

$= g(N_x(t), s)\, v(N_x(t))\,ds.$

Let's solve for $g(N_x(t), s)$.

$g(N_x(t), \sigma+d\sigma) = \text{Prob}\{$ no reaction $[t, t+\sigma)$
and, no reaction $[t+\sigma, t+\sigma+d\sigma)$ $\}$

(memoryless or. independence) $= g(N_x(t), \sigma)\,[1 - v(N_x(t+\sigma))\,d\sigma]$.

$N_x(t+\sigma) = N_x(t)$

since no reactions occured.

$$\Rightarrow \quad \frac{g(N_x(t), \tau+d\tau) - g(N_x(t),\tau)}{d\tau} = -v(N_x(t))\, g(N_x(t),\tau)$$

$$\Rightarrow \quad d\tau \to 0. \qquad \frac{d}{d\tau} g(N_x(t),\tau) = -v(N_x(t))\, g(N_x(t),\tau)$$

$$\Rightarrow \quad g(N_x(t),\tau) = e^{-v(N_x(t))\tau}.$$

$$\Rightarrow \quad f(N_x(t), s)\,ds = v(N_x(t))\, e^{-v(N_x(t))s}\, ds$$

↳ pdf (prob density function) for exponential distr.

$\tau \sim Exp(v)$. pdf. $ve^{-v\tau}$

cdf. $1 - e^{-v\tau}$.

- We want τ. s.t. $t+\tau$ is time for next reaction.
  for a reaction with rate v.
  Then $\tau \in (0, +\infty)$ is a random number.

  $\tau \sim Exp(v)$.

  This is exact! e.g. $X \xrightarrow{v_2} X-1 \quad v_2 = kx$.

  At t. draw $\tau \sim Exp(kN_x(t))$.

  Then $N_x(t+\tau) = N_x(t) - 1$

- What if multiple reactions?.

  The reactions are independent. each with rate $v_1, \dots, v_m$.

  Let τ_0 be time $t+\tau_0$ that any reaction happens.

  $\Rightarrow \quad \tau_0 \sim Exp(v_1 + \dots + v_m)$.

  Which reaction? $\mathbb{P}\{\text{it's reaction } j\} = \frac{v_j}{v_1 + \dots + v_m}$

- This completes the Gillespie algorithm, or SSA.

  ~~Exact~~ simulation of stochastic trajectories by sampling event times.

⑥. Analysis of steady state distributions.

- Simulations can't give the full picture. over all parameters.

  ⇒ Analysis via moments. from CME.

  Mean. Var.

- Exact analysis. of Moments.

example. $X \xrightarrow{\mu} X+1$

$X \xrightarrow{k} X-1$

$$\frac{dp_n}{dt} = k(n+1)p_{n+1} + \mu p_{n-1} - (kn+\mu)p_n.$$

Mean: $M(t) = \sum_{n=0}^{\infty} n p_n(t)$

Variance: $V(t) = \sum_{n=0}^{\infty} (n - M(t))^2 p_n(t)$

$$\frac{dM}{dt} = \frac{d}{dt}\sum_{n=0}^{\infty} n p_n(t) = k\sum_{n=0}^{\infty} n(n+1)p_{n+1} + \mu\sum_{n=0}^{\infty} n p_{n-1}$$

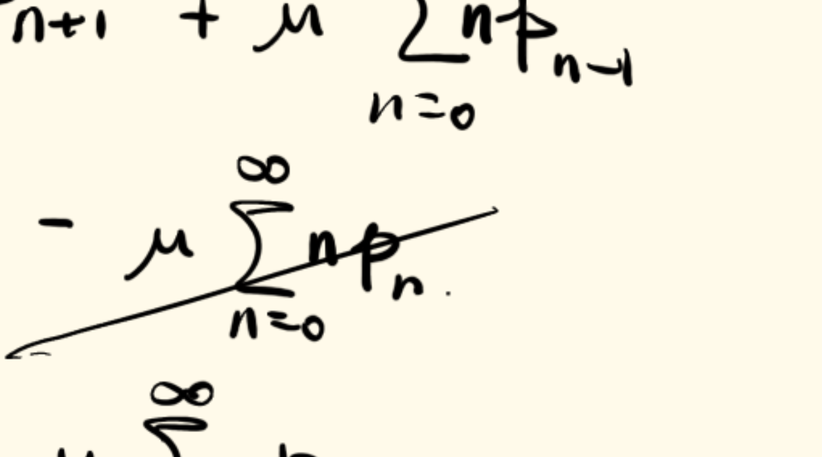
$k\sum_{n=0}^{\infty}(n-1)\,n\,p_n$ $\qquad$ $\mu\sum_{n=0}^{\infty}(n+1)\,p_n$

$$- k\sum_{n=0}^{\infty} n^2 p_n - \mu\sum_{n=0}^{\infty} n p_n.$$

$$= -k\underbrace{\sum_{n=0}^{\infty} n p_n}_{M} + \mu\underbrace{\sum_{n=0}^{\infty} p_n}_{1}$$

$$= \mu - kM.$$

(This is just like. $\frac{d}{dt}x = \mu - kx$. deterministic rate eqn. WARNING: Not always so.)



Similarly. observe. $\sum_{n=0}^{\infty} n^2 p_n = V + M^2$

$$\frac{d}{dt}(V+M^2) = \frac{d}{dt}\sum_{n=0}^{\infty} n^2 p_n = \underbrace{k\sum_{n=0}^{\infty} (n-1)^2 n\, p_n}_{k\sum_{n=0}^{\infty} n^2(n+1)\,p_{n+1}} + \underbrace{\mu\sum_{n=0}^{\infty}(n+1)^2 p_n}_{\mu\sum_{n=0}^{\infty} n^2 p_{n-1}}$$

$(n^2-2n+1)n$ $\qquad$ n^2+2n+1

$$- k\sum_{n=0}^{\infty} n^2 p_n - \mu\sum_{n=0}^{\infty} n^2 p_n$$

$$= \sum_{n=0}^{\infty}\left[k(-2n^2+n) + \mu(2n+1)\right]p_n$$

$$= -2k(V+M^2) + kM + 2\mu M + \mu.$$

$$\frac{d}{dt}(V+M^2) = \frac{dV}{dt} + 2M\frac{dM}{dt} = \frac{dV}{dt} + 2M(\mu - kM).$$

$$\Rightarrow \begin{cases} \frac{dV}{dt} = \mu + kM - 2kV. \\ \frac{dM}{dt} = \mu - kM. \end{cases}$$

At s.s. $M = \frac{\mu}{k}$. $V = \frac{\mu+kM}{2k} = \frac{\mu}{k}$.

So. Mean = Variance. (Poisson! In fact. it is...)

You can solve p_n at s.s. explicitly...

- But. this doesn't always work.

That the moments form a finite number of equations

is called Moment Closure.

Not closed if. e.g. $\mathbb{E}(X)$ depends on $\mathbb{E}(X^2)$ depends on $\mathbb{E}(X^3)$...

this happens when e.g. $X \xrightarrow{v} X-1 \quad v = x^2$

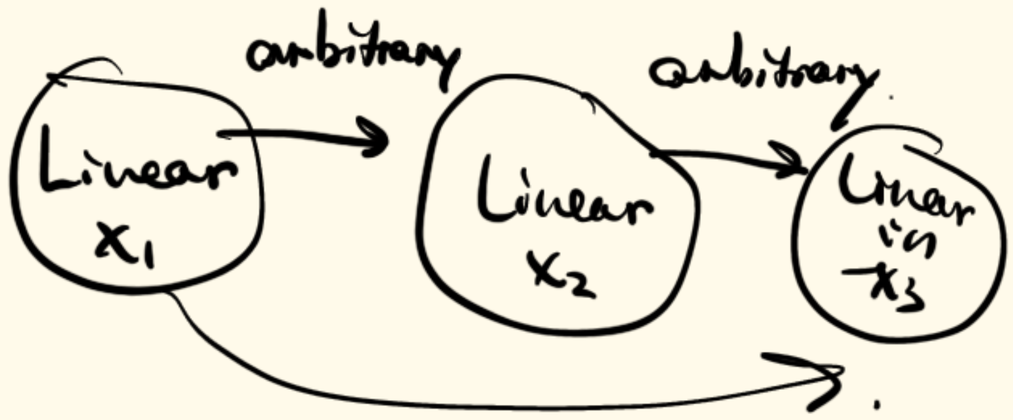
Cases with moment closure:

— Linear. i.e. 1st order or 0th order reactions.

v_j are all degree 1 polynomials of x.

e.g. c, $c+x_1$, x_2. but not x_1x_2, x_1^2.

— Feedforward structure



$$\dot{x}_1 = \mu - x_1$$
$$\dot{x}_2 = x_1^2 - x_2$$
$$\dot{x}_3 = x_1 x_2 - x_3$$

This excludes many interesting cases though..

- Linear Noise Analysis. (LNA).

— Just like using linearization to analyze nonlinear dynamical systems.

We can also do linearization for stochastic processes.

— Write reactions in net change form

$$X \xrightarrow{v_i(x)} X + \nu_i$$

x is a vector of species. molecular counts.

$i = 1, \dots, m$

$x = (x_j)$. $j = 1, \dots, n$.

(Might be too detailed. Could be better left as homework—?)

— $\frac{d}{dt} P(x,t) = \sum_i \left[v_i(x - \nu_i)\, P(x - \nu_i, t) - v_i(x)\, P(x,t) \right]$

Mean. $\frac{d}{dt}\langle x_j \rangle = \frac{d}{dt} \sum_x x_j P(x,t) = \left\langle \sum_i \nu_{ij}\, v_i(x) \right\rangle$

$$= \left\langle \sum_{i:\, \nu_{ij} > 0} \nu_{ij}\, v_i(x) \right\rangle - \left\langle \sum_{i:\, \nu_{ij} < 0} |\nu_{ij}|\, v_i(x) \right\rangle.$$

$$= \langle R_j^+(x) \rangle - \langle R_j^-(x) \rangle.$$

R_j^+: prod. flux. R_j^-: deg. flux

Covariance. $j, l = 1, \dots, n$.

$$\mathrm{Cov}(x_j, R_l^- - R_l^+) + \mathrm{Cov}(x_l, R_j^- - R_j^+) = \sum_i \nu_{ij}\nu_{il}\langle v_i\rangle$$

$$- \text{Recall.} \ \frac{1}{\tau_l}\frac{\mathrm{Cov}(x_j, R_l^- - R_l^+)}{\langle x_j\rangle\langle R_j\rangle} + \frac{1}{\tau_j}\frac{\mathrm{Cov}(x_l, R_j^- - R_j^+)}{\langle x_l\rangle\langle R_l\rangle}$$

(*).

$$= \frac{1}{\tau_j}\frac{\langle s_{l|j}\rangle}{\langle x_l\rangle} + \frac{1}{\tau_l}\frac{\langle s_{j|l}\rangle}{\langle x_j\rangle} = D_{jl} = \frac{\sum_i \nu_{ij}\nu_{il}\langle v_i\rangle}{\langle x_j\rangle\langle x_l\rangle}.$$

τ_j: average lifetime $\quad \tau_j = \dfrac{\langle x_j\rangle}{\langle R_j^{\pm}\rangle}$.

$\langle s_{j|l}\rangle$ stepsizes if $j = l$

$$\langle s_{j|j}\rangle = \sum_i |\nu_{ij}|\, p_{ij}, \quad p_{ij} = \frac{|\nu_{ij}|\langle v_i\rangle}{\sum_{i'} |\nu_{ij}|\langle v_i\rangle}$$

average step sizes of j conditioning on that l changes.

e.g. $x_1 \to x_1 + 1$, $x_1 \to x_1 - 10$. then. $\langle s_{1|1}\rangle = \dfrac{1-10}{2}$.

Co-step sizes if $j \neq l$.

$$\langle s_{l|j}\rangle = \sum_i p_{ij}\, |\nu_{il}|\, \mathrm{sgn}\{\nu_{ij}\nu_{ik}\}.$$

average change of x_l conditioning on x_j changes simultaneously in these reactions.

pos. or. neg.

= Coupling" of x_l, x_j in reactions.

e.g. $(x_1, x_2) \longrightarrow (x_1 - 1, x_2 + 1)$.

no other reactions coupling x_1, x_2 and no other prod. of x_2. then. $\langle s_{1|2}\rangle = -\frac{1}{2}$.

$$- \ \langle s_{l|j}\rangle\langle R_j\rangle = \langle s_{j|l}\rangle\langle R_1\rangle.$$

– The above are exact. How we make linearization.

- $R_j^{\pm}(x) \approx R_j^{\pm}(\langle x \rangle) + \sum_\ell \frac{\partial R_j^{\pm}(x)}{\partial x_\ell} \Delta x_\ell$.

- Define logarithmic gain. (prod-deg reaction order.)

$$H_{j\ell} = \frac{\partial R_j^{+}(\langle x \rangle)}{\partial x_\ell} \frac{\langle x_\ell \rangle}{\langle R_j^{\pm} \rangle} - \frac{\partial R_\ell^{-}(\langle x \rangle)}{\partial x_j} \frac{\langle x_j \rangle}{\langle R_\ell^{\pm} \rangle}$$

Then. ⭐ becomes.

$$M\eta + \eta M^T = D$$

where. $M_{ij} = -\frac{H_{ij}}{\tau_i}$.

$$D_{ij} = D_{ji} = \frac{1}{\tau_i}\frac{\langle s_{j|i} \rangle}{\langle x_j \rangle} + \frac{1}{\tau_j}\frac{\langle s_{i|j} \rangle}{\langle x_i \rangle}.$$

$$\eta_{ij} = \eta_{ji} = \frac{\mathrm{Cov}(x_i, x_j)}{\langle x_i \rangle \langle x_j \rangle}.$$

– Apply this to examples.

The one we started with.

Then more...