随机热力学（书版）笔记重制-Day6：大偏差原理

#StochasticThermaldynamics

可以非常简单地说，一个随机变量到底是几阶矩有界就决定了它向均值集中的速度。一种非常特殊的情形是各阶矩都有界的情况，因为此时有所谓大偏差原理，这是一种用于检查随机变量在尾部的收敛速度的定理。考虑随机变量 X，其累积量（中心矩）生成函数定义为：

$$\varphi(\lambda) = \ln \mathbb{E} \exp(\lambda X)$$

大偏差的一个简单推导：

$$\begin{aligned}
\mathbb{P}(X > t) &= \mathbb{P}(\exp(X) > \exp(t)) \\
&= \mathbb{P}(\exp(\lambda X) > \exp(\lambda t)) \\
&\leq \frac{1}{\exp(\lambda t)} \mathbb{E}(\exp(\lambda X)) \\
&= \frac{1}{\exp(\lambda t)} \exp(\varphi(\lambda)) \\
&= \exp(\varphi(\lambda) - \lambda t)
\end{aligned}$$

上面的式子对任意 $\lambda > 0$ 都是成立的，我们应当使用其中最紧的 Bound：

$$\mathbb{P}(X > t) \leq \exp[-\sup_{\lambda>0}(\lambda t - \varphi(\lambda))] = \exp(-I(t))$$

这里的 $I(t)$ 称为速率函数，是累积量生成函数的勒让德变换。实际上，尾部分布是高斯的随机变量非常有趣，如果数据从这样的分布中 iid 采样，你可以推出统计学习模型的一个正比于 $\sqrt{\log|\mathcal{H}|}$ 的泛化界（$|\mathcal{H}|$ 是假设空间的大小）。如果你有一个随机变量 $Y = \sum_{i=1}^N X_i$，仿照上述推导给出 Y 的尾部分布衰减速率：

$$\mathbb{P}(Y > t) \leq \exp(-NI(t))$$

下文中我们将会研究一些可观测量的时间平均，它们的尾部分布以 $\exp(-TI(t))$ 的速度衰减（T 为测量的时间）。

Currents , Traffics and Other Observables

随机热力学中的宏观热力学量大概可以分成两类，一类是与在一个态上停留的时间有关系的，记停留在 y 态上的时间：

$$\tau_y = \int_{t_0}^{t_f} \mathrm{d}t \delta_{y,x(t)}$$

这样的量可以写成：

$$\mathcal{A}[x] = \sum_y \alpha_y \tau_y$$

另一类当然是与在两个态之间的跳跃次数有关的：记两个态之间的跳跃次数：

$$n_{xx'}[x] = \sum_{k=1}^{n} \delta_{xx_k} \delta_{x'x_{k-1}}$$

这样的可观测量记为：

$$J[x] = \sum_{x \neq x} n_{xx'}[x] d_{xx'}$$

按照我们上面的定义，这些可观测量（例如一个非平衡稳态系统的熵产生）是会随着时间增长的。我们现在要研究这些变量如何偏离均值，所以我们希望把均值控制到不随着时间增长。为此，我们将停留在各态上的时间换成经验频率：$f_y = \dfrac{\tau_y}{T}$，并且将与跳跃相关的量全都除以 T。

现在考虑在小的驱动下稍稍偏离精致平衡的系统，但是系统仍然位于非平衡稳态。我们希望考虑与跳跃次数相关的变量 j 的分布。首先，对于处在精致平衡的系统，j 的分布必然关于 0 对称，这对其 CGF 有要求：

$$\varphi^{(j)}(q) = \varphi^{(j)}(-q)$$

而偏离精致平衡时，$\varphi^{(j)}(q)$ 不再是偶函数，可以将其写成：

$$\varphi^{(j)}(q) = \varphi^{+}(q) + \delta \cdot \varphi^{-}(q)$$

的形式。将奇函数、偶函数部分全部展开到一阶：

$$\varphi^{+}(q) \approx \varphi_2 q^2 \quad \varphi^{-}(q) \approx \varphi_1 q$$

通过勒让德变换求解速率函数，得到：

$$q^\star \approx \frac{j - \delta\varphi_1}{2\varphi_2} \Rightarrow I(j) \approx \frac{(j - \delta\varphi_1)^2}{4\varphi_2}$$

如果展开到更高阶就会看到非高斯的修正。

Large Deviation and Fluctuation Relations

现在将大偏差定理和之前推导的精细涨落关系联系起来。令 $j_s^{tot} = \dfrac{s^{tot}}{T}$，那么关于熵的精细涨落关系是：

$$\frac{p(j_s^{tot}, \lambda)}{p(-j_s^{tot}, \hat{\lambda})} = \exp\left(\frac{1}{k_B} T j_s^{tot}\right)$$

但是，在正向、逆向控制协议 $\lambda, \hat{\lambda}$ 不可区分（或者说使用恒定控制协议），且系统已经处于稳态（可以是非平衡稳态）的条件下，将 $\hat{x}$ 当作逆向路径计算的路径概率 $P_{\hat{x}}[\hat{\lambda}] = P_{\hat{x}|x_f}[\hat{\lambda}] p_{x_f}(t_f)$ 与将 $\hat{x}$ 当作正向路径计算的路径概率 $P_{\hat{x}}[\lambda] = P_{\hat{x}|x_0}[\lambda] p_{x_0}(t_0)$ 是相同的。所以我们可以直接提取出正向路径上熵产生 j_s^{tot} 和正向路径上熵产生 $-j_s^{tot}$ 的概率之比，将上面的精细涨落关系改写成：

$$\frac{p(j_s^{tot}, \lambda)}{p(-j_s^{tot}, \lambda)} = \exp\left(\frac{1}{k_B} T j_s^{tot}\right)$$

研究 j_s^{tot} 非常偏离均值的情形，此时用大偏差原理描写 $p(\cdot)$：

$$p(j_s^{tot}, \lambda) \approx \exp(-T I(j_s^{tot}))$$

这给出精细涨落关系的等价表述：

$$I(j_s^{tot}) - I(-j_s^{tot}) = -\frac{j_s^{tot}}{k_B}$$

也可以视作对 CGF 施加了限制：

$$\varphi^{j_s}(q) = \sup_{j_s}(j_s q - I(j_s)) = \sup_{j_s}\left(j_s\left(q + \frac{1}{k_B}\right) - I(-j_s)\right) = \psi^{j_s}\left(-q - \frac{1}{k_B}\right)$$

Fluctuation Theorem for Currents

首先我们说明：主方程显然可以用图表示。考虑有环图，在图上找到任意一颗生成树，设图有 V 个节点，E 条边，则生成树必然包含 $V-1$ 条边，剩余的 $C = E - (V-1)$ 条边称为弦。每次选择一条弦加入到生成树中，由该弦和生成树中若干条边必然组成一个环路，该环路称为图的一个基本环，利用所有的弦可构造图的所有基本环。（非平衡稳态时）所有与跃迁次数有关的热力学变量（如熵产生速率）均可在基本环上分解（这里以熵产生为例）：

$$\sigma = \sum_{\alpha=1}^{C} n_\alpha A_\alpha \frac{1}{T_e}$$

其中，n_α 是在第 α 个基本环上完成跳跃的圈数，A_α 称为“亲和势”（当然这里代表的是跳跃一圈产生多少熵）。仍然考虑处在非平衡稳态的系统，当然，熵产生可以写成能在基本环上分解的部分以及剩余部分。但是非平衡稳态的时候，沿着环路有恒定的概率流，所以只有能写到基本环上的熵产生要正比于运行时间。算一下大偏差原理要用到的 Scaled CGF（不难看出不乘这个 $\dfrac{1}{T}$ 的话这个量要正比于时间上升）：

$$\varphi^j(\vec{q}) = \lim_{T\to+\infty}\frac{1}{T}\ln\left\langle\exp\left(\sum_\alpha q_\alpha n_\alpha[x]\right)\right\rangle$$

$$\begin{aligned}
\varphi^j\left(q-\frac{A}{K_BT_e}\right) &= \lim_{T\to+\infty}\frac{1}{T}\ln\int\mathcal{D}xP_x\exp\left(\sum_\alpha\left(q_\alpha-\frac{A_\alpha}{k_BT_e}\right)n_\alpha[x]\right)\\
&\approx \lim_{T\to+\infty}\frac{1}{T}\ln\int\mathcal{D}xP_x\exp\left(\sum_\alpha q_\alpha n_\alpha[x]-\frac{1}{k_B}s^{tot}[x]\right)\\
&= \lim_{T\to+\infty}\frac{1}{T}\ln\int\mathcal{D}\hat{x}P_{\hat{x}}\exp\left(\sum_\alpha q_\alpha n_\alpha[x]\right)\\
&= \lim_{T\to+\infty}\frac{1}{T}\ln\int\mathcal{D}\hat{x}P_{\hat{x}}\exp\left(-\sum_\alpha q_\alpha n_\alpha[\hat{x}]\right)\\
&= \lim_{T\to+\infty}\frac{1}{T}\ln\int\mathcal{D}xP_x\exp\left(-\sum_\alpha q_\alpha n_\alpha[x]\right)\\
&= \varphi^j(-q)
\end{aligned}$$

这给出了对 $\varphi^j(q)$ 的一个约束关系，注意这是一个普适的恒等式，无论离平衡态多远都成立。这个约束关系可以用于推导昂萨格倒易关系：简便起见，记 $\tilde{A}_\alpha=\dfrac{A_\alpha}{K_BT_e}$，我们上面将 SCGF 表达成了 $\vec{q},\vec{A}$ 的函数，根据 SCGF 的定义，其前两阶导数给出两阶累积量：

$$\frac{\partial\varphi}{\partial q_\alpha}|_{\vec{q}=0} = \lim_{T\to+\infty}\frac{1}{T}\langle n_\alpha\rangle = \langle J_\alpha\rangle$$

$$\frac{\partial^2\varphi}{\partial q_\alpha\partial q_\beta}|_{\vec{q}=0} = \lim_{T\to+\infty}\frac{1}{T}(\langle n_\alpha n_\beta\rangle - \langle n_\alpha\rangle\langle n_\beta\rangle) := D_{\alpha\beta}$$

这里的 A_α 大概可以被视作一个驱动力，因为

$$\frac{P_{cw}}{P_{ccw}} = \frac{w_{1\to2}w_{2\to3}\cdots w_{n\to1}}{w_{1\to n}\cdots w_{3\to2}w_{2\to1}} = \exp\left(\frac{A_\alpha}{k_BT_e}\right)$$

所以 A_α 代表了一个基本环上正、逆概率之比，这个比例与外加驱动有关，只有在 $A_\alpha\neq 0$ 的时候系统才会沿着环转圈。根据昂萨格的说法，在近平衡态的时候，流与驱动力成正比关系：

$$\langle J_\alpha\rangle = \sum_\beta L_{\alpha\beta}\tilde{A}_\beta$$

将 Scaled CGF 展开到二阶：

$$\begin{aligned}
\varphi(\vec{q},\vec{A}) &\approx \sum_{\alpha,\beta}L_{\alpha\beta}\tilde{A}_\beta q_\alpha + \frac{1}{2}\sum_{\alpha\beta}D_{\alpha\beta}q_\alpha q_\beta\\
&= \sum_{\alpha,\beta}L_{\alpha\beta}\tilde{A}_\beta(-q_\alpha-\tilde{A}_\alpha) + \frac{1}{2}\sum_{\alpha\beta}D_{\alpha\beta}(-q_\alpha-\tilde{A}_\alpha)(-q_\beta-\tilde{A}_\beta)
\end{aligned}$$

对比两侧 $q_\alpha\tilde{A}_\beta$ 前的系数，得到：

$$D_{\alpha\beta} = 2L_{\alpha\beta} \Rightarrow L_{\alpha\beta} = L_{\beta\alpha}$$

我们从任意远离平衡态的情形恢复了近平衡态的倒易关系。

Tilting

从上文中我们知道，SCGF 包含了对于分布尾部的重要信息，但是有时候不好计算。一种计算技术称为“倾斜/Tilting”，先定义条件 SCGF：

$$\varphi_x^{(j)}(q,t) = \langle\exp(q\cdot T\cdot j)|x(t)=x\rangle = \int\mathcal{D}xP_x\delta_{x,x(t)}\exp(q\cdot j[x]\cdot T)$$

这里是只考虑了那些在 t 时刻运行到 x 处的路径的 SCGF，要还原标准是 SCGF 当然需要取平均：

$$\varphi^{(j)}(q) = \sum_x \varphi_x^{(j)}(q,t)$$

现在说明我们定义的这个条件 SCGF 反而是好算的。显式展开其表达式：

$$\begin{aligned}\varphi_x^{(j)}(q,t) &= \int \mathcal{D}x \delta_{x,x_n} \exp\left(q\sum_{k=1}^{n} d_{x_k,x_{k-1}}\right) \exp(-k_{x_n}^{out}(t-t_n)) k_{x_n,x_{n-1}} \cdots k_{x_1 x_0} \exp(\cdots) p_{x_0}(t_0) \\ &= \int \mathcal{D}\vec{x} \delta_{x,x_n} \exp(-k_{x_n}^{out}(t-t_n)) \exp(q d_{x_n,x_{n-1}}) k_{x_n,x_{n-1}} \cdots \exp(q d_{x_1,x_0}) k_{x_1 x_0} \exp(\cdots) p_{x_0}(t_0)\end{aligned}$$

其中，$d_{x_k,x_{k-1}}$ 代表从态 x_{k-1} 跳转到 x_k 对流的贡献。所以你可以认为 $\varphi_x^{(j)}(q,t)$ 被写成一个非归一化主方程的路径概率的形式，这个主方程是：

$$\frac{\mathrm{d}}{\mathrm{d}t}\varphi_x^{(j)}(q,t) = \sum_{x'} L_{xx'} \varphi_{x'}^{(j)}(q,t)$$

转移矩阵：

$$L_{xx'}(q) = \begin{cases} k_{xx'} \exp(q d_{xx'}) & x \neq x' \\ -k_x^{out} & x = x' \end{cases}$$

在时间很长的时候，上面这个线性方程组的解是被 $L_{xx'}(q)$ 的最大本征值控制的，所以你基本上只需要（利用各种数值方法）求出这个最大本征值就好了。

如果求本征值还是太繁琐了，那么可以尝试模拟以上非平衡主方程，有一种技术称为 Cloning。具体而言，引入以下中间量：

- 倾斜跃迁速率：$k_{xx'}(q) = k_{xx'} \exp(q d_{xx'})$
- 倾斜脱离率：$k_x^{out}(q) = \sum_{x \neq x'} k_{x'x}(q)$
- 停留时间分布：$\rho_x(t,q) = k_x^{out}(q) \exp(-k_x^{out}(q) t)$
- 权重：$Y_x(t,q) = \exp((k_x^{out}(q) - k_x^{out}(0))t)$
  利用这一些记号，可以重写上面的概率：

  $$\varphi_x^{(j)}(q,t) = \int \mathcal{D}x \delta_{x,x_n} \frac{1}{k_{x_n}^{out}(q)} \rho_{x_n}(t-t_n;q) Y_x(t-t_n;q) \frac{k_{x_n x_{n-1}}(q)}{k_{x_{n-1}}^{out}(q)} \cdots p_{x_0}(t_0)$$

  为了模拟这个系统，设定一个固定的时间步长 ΔT，并初始化$N \gg 1$ 份系统的克隆。时刻 t 时第 α 个系统由其状态 x^α 表示。在每一个时间切片 $[t, t+\Delta T]$ 内，反复进行以下步骤：
- **独立演化**：让所有 N 个克隆在 ΔT 时间内，根据修正后的转移速率 $k_{xx'}(q)$ 和逃逸率 $k_x^{out}(q)$，独立进行连续时间随机游走（例如使用标准的 Gillespie 算法）。
- **计算连续权重**：时间推进 ΔT 后，强制暂停所有克隆的演化。对每个克隆 α，计算其在这段 ΔT 时间内积累的真实权重：

  $$W^\alpha = \exp\left(\int_t^{t+\Delta T} \left(k_{x^\alpha(t')}^{out}(q) - k_{x^\alpha(t')}^{out}(0)\right) dt'\right)$$

  (注：如果在这段 ΔT 内克隆 α 发生了跳转并经历了多个状态，则该积分就是各个状态的停留时间 Δt_i 与对应适应度乘积的离散累加)
- **计算归一化因子**：计算当前群体在这 ΔT 内的平均权重：

  $$Z = \frac{1}{N}\sum_{\alpha=1}^{N} W^\alpha$$

- **全局重采样（克隆/删除）**：根据每个克隆的相对权重调整其数目：$$
  Y^{\alpha} = \lfloor \frac{W^{\alpha}}{Z} + \epsilon \rfloor

  其中$\epsilon \sim \mathcal{U}[0,1]$。若$y^\alpha = 0$，则该克隆被删除；若$y^\alpha = 1$，则保留一个；若$y^\alpha > 1$，则该克隆被复制为$y^\alpha$个。

- **维持总数**：由于随机取整，此时克隆总数 N' 可能略微偏离 N。通过随机均匀地复制或删除少量克隆，将群体总数精确调整回 N。

- **记录与循环**：记录当前的 Z，将时间更新为 $t+\Delta T$，并进入下一个时间切片的循环。
  最终的 SCGF 可以通过长时间模拟中记录的 Z 求出：

$$\varphi(q)=\lim_{M\to\infty}\frac{1}{M\Delta T}\sum_{m=1}^{M}\ln Z_m$$

这个模拟其实是在看：如果我不做维持总数这一步，那系统的数量要增长到多少？这个增长速度和非归一化主方程的转移矩阵的最大本征值有关。

Different Levels of Large Deviation Principle

我们需要处理不同层级的随机变量：

- Level 1 是关于瞬时状态，例如系统的能量 ϵ_x
- Level 2 是关于经验频率 f_x
- Level 3 是关于轨道的泛函，例如熵产生 $s[\vec{x}]$
  此时要使用一个结论：若随机变量 X,Y 间有函数关系 $y=f(x)$，则二者间的速率函数也有关系：

$$I^{(Y)}(y)=\inf_{x=f(y)}I^{(X)}(x)$$

  这意味着 Y 的速率函数尤其对应的最可能的 X 值决定。
  对马氏系统可以引入一个“Level 2.5”，也就是检查经验频率 f_x 和经验跳率的联合分布 $r_{xx'}=\frac{n_{xx'}}{T}$。考虑系统处于稳态，真的稳态分布是 p_x^{st}，真的跳率是 $k_{xx'}$，引入一个辅助过程，使得其跳率满足 $r_{xx'}=\tilde{k}_{xx'}f_{x'}$，那么 $\tilde{k}_{xx'}$ 就直接包含了两个可观测量的信息。直观来看， $\tilde{k}_{xx'}$ 将随着时间的延长逐渐变成 $k_{xx'}$，不难相信下面这个东西也要随着时间增长趋于稳定值：

$$\begin{aligned}\frac{P_x}{P_x^\star}&=\left(\prod_{i=1}^{N}\frac{k_{x_ix_{i-1}}}{\tilde{k}_{x_ix_{i-1}}}\right)\left(\prod_{i=0}^{n}\exp(-(k_{x_i}^{out}-\tilde{k}_{x_i}^{out})(t_{i+1}-t_i))\right)\\&=\sum_{xx'}\left(\frac{k_{xx'}}{\tilde{k}_{xx'}}\right)^{r_{xx'}}\exp\left(\sum_{x'}-(k_{x'}^{out}-\tilde{k}_{x'}^{out})f_{x'}T\right)\\&=\exp\left(-T\sum_{xx'}\left(\tilde{k}_{xx'}\ln\frac{\tilde{k}_{xx'}}{k_{xx'}}-\tilde{k}_{xx'}+k_{xx'}\right)f_{x'}\right)\end{aligned}$$

  从而，记按照真实主方程，在 t 时刻获得经验频率 f，经验跳率 $r_{xx'}=\tilde{k}_{xx'}f_{x'}$ 的概率是 $p(f,\tilde{k},t)$，那么：

$$p(f,\tilde{k},t)=p^\star(f,\tilde{k},t)\left\langle\frac{P_x}{P_x^\star}\right\rangle$$

  注意：这里的 $p^\star(f,\tilde{k},t)$ 是用经验跳率构造的主方程给出的概率，它很接近 1。这里对概率比取平均是在所有 $(f,\tilde{k})$ 路径上取平均。那么，至少以正确的数量级，我们有：

$$p(f,\tilde{k},t)\approx\exp(-TI^{(f,\tilde{k})}(f,\tilde{k})),\ \ I^{(f,\tilde{k})}(f,\tilde{k})=\sum_{xx'}\left(\tilde{k}_{xx'}\ln\frac{\tilde{k}_{xx'}}{k_{xx'}}-\tilde{k}_{xx'}+k_{xx'}\right)f_{x'}$$

  利用上面的结论你可以提取其中任意一个变量的速率函数，例如：

$$p(f,T)=\exp(-TI^{(f)}(f)),\ \ I^{(f)}(f)=\inf_{\tilde{k}}I^{(f,\tilde{k})}(f,\tilde{k})$$

  对于其他变量，例如 $j_{xx'}=\tilde{k}_{xx'}f_{x'}-\tilde{k}_{x'x}f_x$，它的速率函数可以用乘子法求出。