非线性数据同化问题

黄政宇
北京大学北京国际数学研究中心
北京大学国际机器学习研究中心

数据同化问题

- 非线性数据同化问题

演化方程: $x_{n+1} = \mathcal{F}(x_n) + \omega_{n+1}$

观测方程: $y_{n+1} = \mathcal{H}(x_{n+1}) + \eta_{n+1}$

- 假设

高斯演化噪音: $\rho_\omega = \mathcal{N}(x; 0, \Sigma_\omega)$

高斯观测噪音: $\rho_\eta = \mathcal{N}(x; 0, \Sigma_\eta)$

高斯先验分布: $\rho_{\text{prior}}(x) = \mathcal{N}(x; r_0, \Sigma_0)$

数据同化问题

- 优化方法（weak constraint 4D variational data assimilation, w4D-Var）

$$\mathrm{argmin}_X \frac{1}{2} \| C_0^{-\frac{1}{2}} (x_0 - m_0) \|^2$$

$$+ \frac{1}{2} \sum_{n=0}^{N-1} \| \Sigma_\omega^{-\frac{1}{2}} (x_{n+1} - \mathcal{F}(x_n)) \|^2$$

$$+ \frac{1}{2} \sum_{n=0}^{N-1} \| \Sigma_\eta^{-\frac{1}{2}} (y_{n+1} - \mathcal{H}(x_{n+1})) \|^2$$

$\mathrm{X} = \{x_0, x_1, \cdots, x_N\}$

平滑问题的最大后验估计

数据同化问题

- 优化方法（4D variational data assimilation, 4D-Var）

$$\mathrm{argmin}_X \frac{1}{2} \| C_0^{-\frac{1}{2}} (x_0 - m_0) \|^2$$

$$+\frac{1}{2}\sum_{n=0}^{N-1} \| \Sigma_\eta^{-\frac{1}{2}} (y_{n+1} - \mathcal{H}(x_{n+1})) \|^2$$

Constraints: $x_{n+1} = \mathcal{F}(x_n)$

$\mathrm{X} = \{x_0, x_1, \cdots, x_N\}$

演化噪音相对于观测噪音更小

数据同化问题

- 优化方法（3D variational data assimilation, 3D-Var）

$$x_0 = m_0$$

对 $n = 0,1,\cdots,N-1$

$$\mathrm{argmin}_{x_{n+1}} \frac{1}{2} \| \Sigma_\omega^{-\frac{1}{2}} (x_{n+1} - \mathcal{F}(x_n)) \|^2 + \frac{1}{2} \| \Sigma_\eta^{-\frac{1}{2}} (y_{n+1} - \mathcal{H}(x_{n+1})) \|^2$$

当 $\mathcal{H}(x) = Hx$, 我们有

$$x_{n+1} = (I - KH)\mathcal{F}(x_n) + Ky_{n+1}$$
$$S = H\Sigma_\omega H^T + \Sigma_\eta$$
$$K = \Sigma_\omega H^T S^{-1}$$

滤波问题的最大后验估计

Lorenz63系统

- 常微分方程系统

$$\frac{dx_1}{dt} = \sigma(x_2 - x_1)$$
$$\frac{dx_2}{dt} = x_1(r - x_3) - x_2$$
$$\frac{dx_3}{dt} = x_1x_2 - \beta x_3$$

其中 $\sigma = 10$, $\beta = \frac{8}{3}, r = 10$ 或者 28。

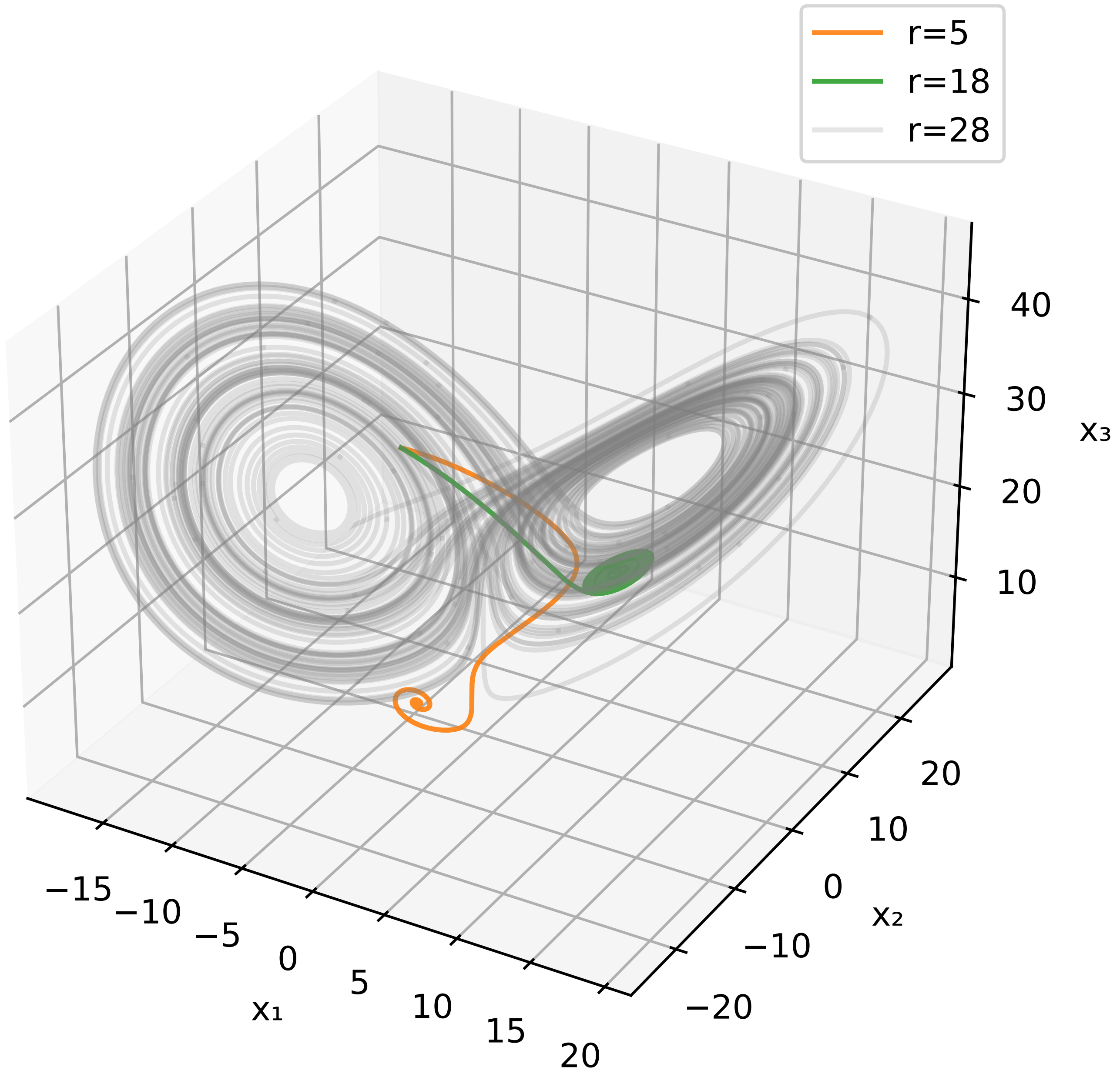
Edward Lorenz

其中$\sigma, r, \beta \in R^+$, 吸引子(平衡点) 包括 (0,0,0),
$(\pm\sqrt{\beta(r-1)}, \ \pm\sqrt{\beta(r-1)}, \ r-1)$。

- $r < 24.74$，系统不是混沌的，会收敛到一个吸引子
- $24.74 < r < 148.4$，系统变得混沌。

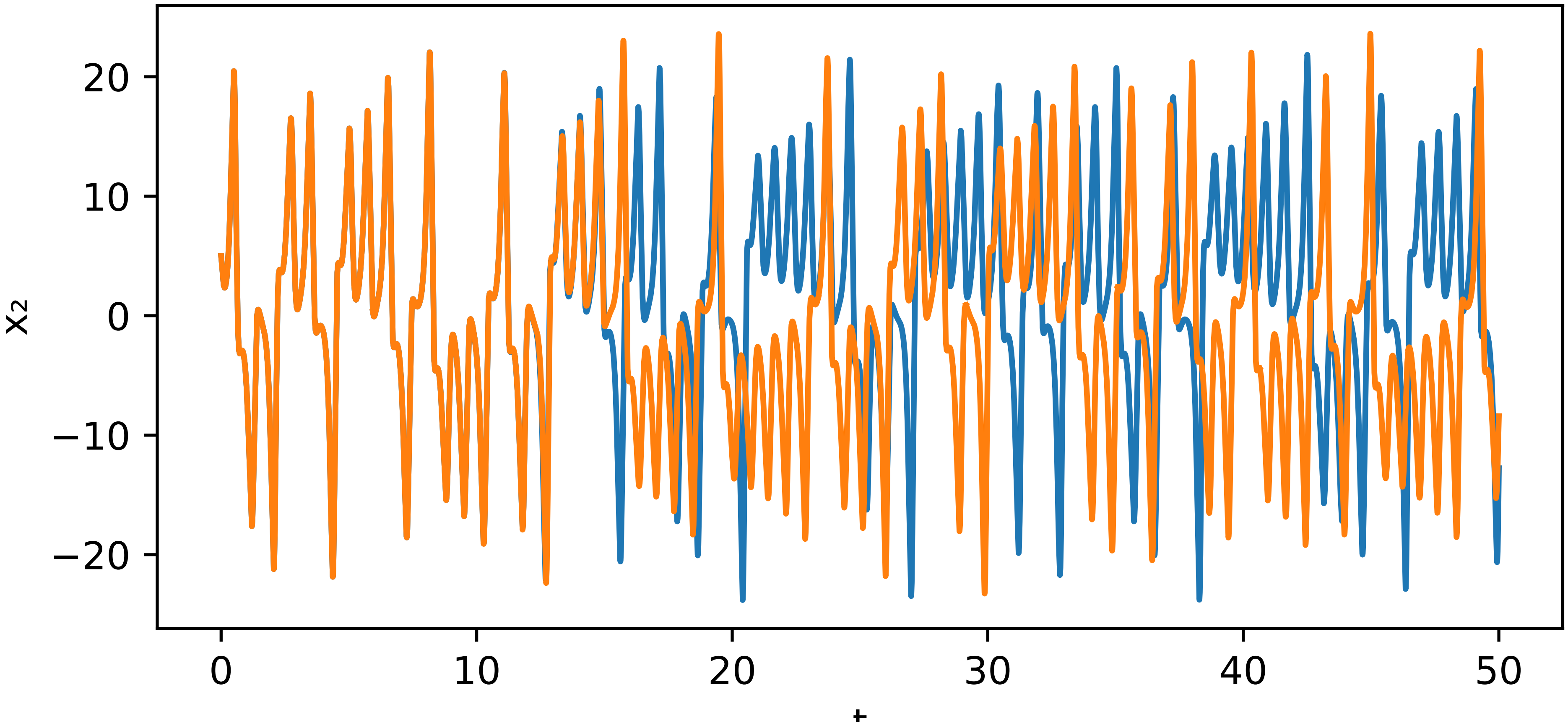
Lorenz63系统

- 混沌系统



Lorenz63系统

- 初值不连续依赖



$$\sigma = 10, r = 28, \beta = \frac{8}{3}$$

$$x(0) = [-8.0;\ 5.0;\ 25] \quad \delta x_2(0) = 10^{-6}$$

Lorenz63系统

- 导数“爆炸”

考虑 $\sigma = 10, \beta = \frac{8}{3}$，混沌Lorenz63系统，

$$\frac{dx_1}{dt} = \sigma(x_2 - x_1)$$
$$\frac{dx_2}{dt} = x_1(r - x_3) - x_2$$
$$\frac{dx_3}{dt} = x_1 x_2 - \beta x_3$$

考虑任意初值，比如[-8.0; 5.0; 25]，定义

$$\mathcal{G}(r) = \int_{30}^{50} x_3(t; r) dt$$

计算$\mathcal{G}(r)$，并用伴随求解器计算$\frac{d\mathcal{G}(r)}{dr}$。

Lorenz63系统

- 导数“爆炸”

前向欧拉方法：

$$x(t_{n+1};r) = x(t_n;r) + \Delta t f(x(t_n);r)$$

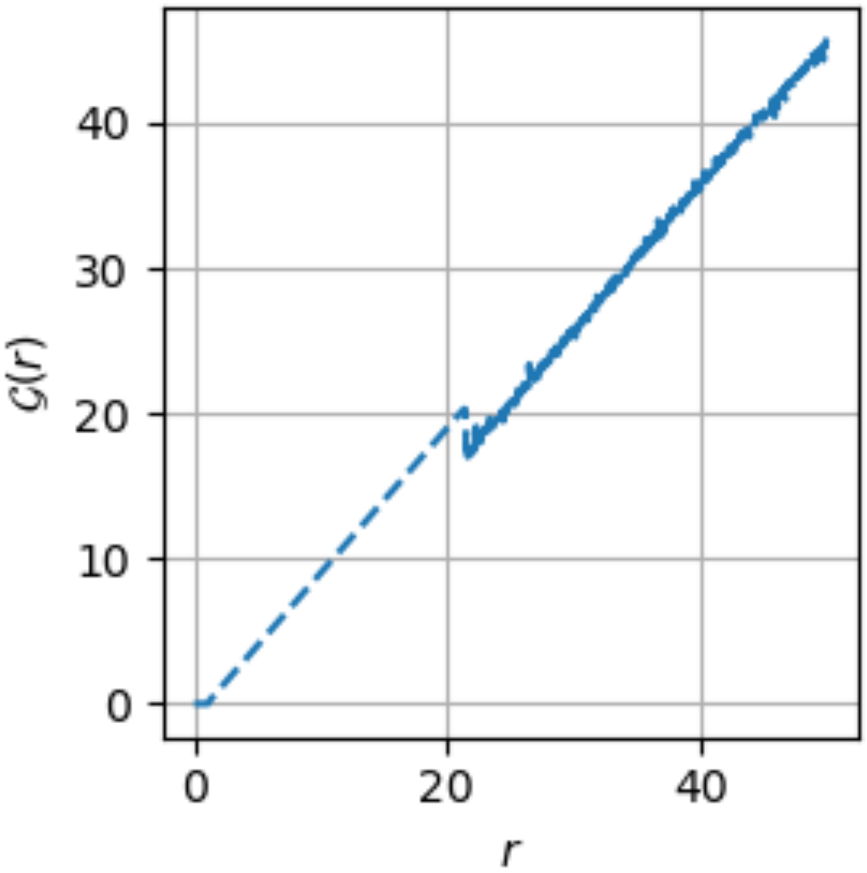
目标函数：

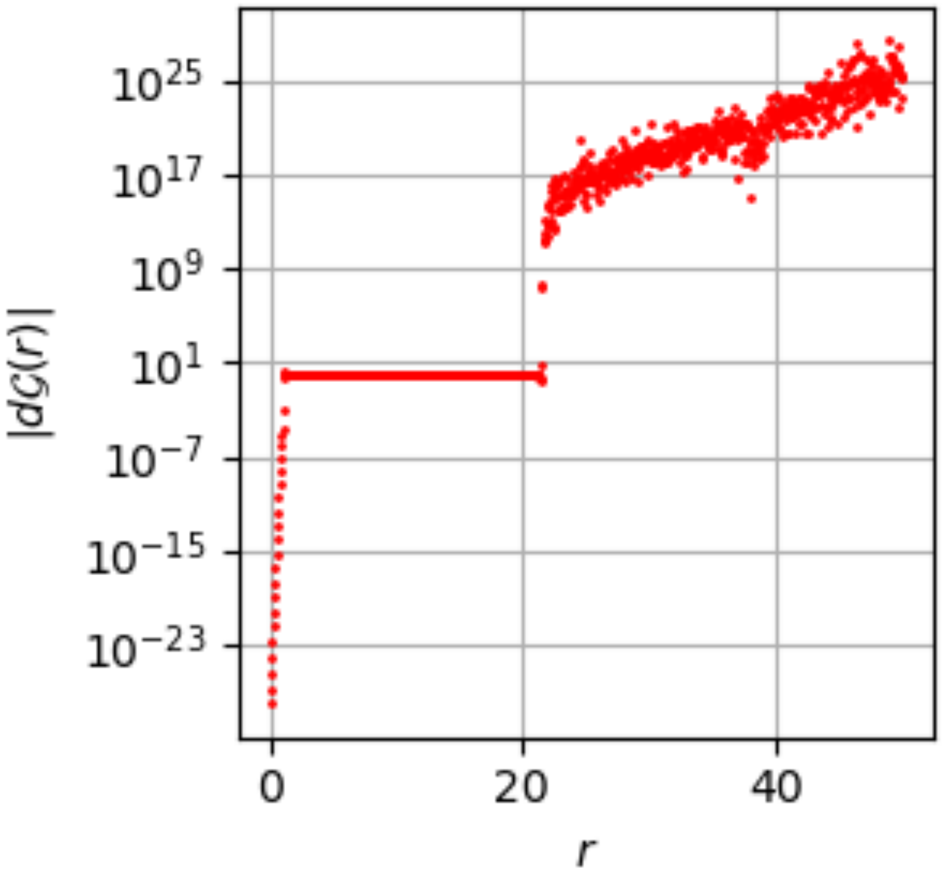
$$\mathcal{G}(r) = \frac{1}{N_t - N_s}\sum_{n=N_s}^{N_t-1} x_3(t_n;r)$$

计算$\mathcal{G}(r)$，并用伴随求解器计算$\frac{d\mathcal{G}(r)}{dr}$。

Lorenz63系统

➢ 导数“爆炸”





数据同化问题

- 贝叶斯方法