

# 针对非线性系统的 降阶模型

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



# 本堂课大纲

- 非线性动力系统
- 轨迹分段线性方法
- 离散经验插值法



# 非线性动力系统

## ➤ 非线性模型

$$\frac{d}{dt}\mathbf{u}(t) = A\mathbf{u}(t) + f(\mathbf{u}(t))$$

## ➤ 基于Galerkin 投影的降阶模型

$$\mathbf{u}(t) \approx V\mathbf{q}(t) \quad V \in R^{N \times k}$$
$$\frac{d\mathbf{q}(t)}{dt} = V^T A V \mathbf{q}(t) + V^T f(V\mathbf{q}(t))$$

$$V^T A V = A_r \in R^{k \times k}$$
$$\underbrace{V^T}_{k \times N} \underbrace{f(V\mathbf{q}(t))}_{N \times 1} : \text{复杂度 } O(2kN) \quad \text{☹}$$



# 非线性动力系统

## ➤ 降阶模型

需要额外的一层近似来确保简化后的非线性方程组

$$\frac{dq(t)}{dt} = V^T AVq(t) + V^T f(Vq(t))$$

的计算成本不会随着高维模型的维度 $N$ 而增长。

## ➤ 嵌套式的近似方法

- 状态近似方法 ( $u(t) \approx Vq(t)$ )
- 非线性函数近似
  - 线性化方法
  - 超级缩减方法 ( Hyper-reduction approaches )



# 本堂课大纲

- 非线性动力系统
- 轨迹分段线性方法
- 离散经验插值法



# 轨迹分段线性方法

➤ 轨迹分段线性方法(Trajectory Piecewise Linear Method)

$$\begin{aligned}\frac{d}{dt}\mathbf{u}(t) &= A\mathbf{u}(t) + f(\mathbf{u}(t)) \\ &\approx A\mathbf{u}(t) + f(\mathbf{u}_1) + \nabla f(\mathbf{u}_1)(\mathbf{u}(t) - \mathbf{u}_1)\end{aligned}$$

多点线性化：

$$\begin{aligned}\frac{d}{dt}\mathbf{u}(t) &= A\mathbf{u}(t) + f(\mathbf{u}(t)) \\ &\approx A\mathbf{u}(t) + \sum_{i=1}^s w_i(\mathbf{u}(t))(f(\mathbf{u}_i) + \nabla f(\mathbf{u}_i)(\mathbf{u}(t) - \mathbf{u}_i))\end{aligned}$$

$$\sum_{i=1}^s w_i(\mathbf{u}(t)) = 1$$



# 轨迹分段线性方法

## ➤ 基于Petrov-Galerkin投影的降阶模型

假设  $W^T V = I$

$$\frac{d}{dt} \mathbf{q}(t) = W^T A V \mathbf{q}(t) + \sum_{i=1}^s \tilde{w}_i(\mathbf{q}(t)) (W^T f(\mathbf{u}_i) + W^T \nabla f(\mathbf{u}_i) (V \mathbf{q}(t) - \mathbf{u}_i))$$

$$\sum_{i=1}^s \tilde{w}_i(\mathbf{q}(t)) = 1$$

$$A_r = W^T A V + \sum_{i=1}^s \tilde{w}_i(\mathbf{q}(t)) W^T \nabla f(\mathbf{u}_i) V$$



# 轨迹分段线性方法

## ➤ 线性化点的选取 $\{\mathbf{u}_i\}_{i=1}^S$

- 使用高维模型的轨迹来选择线性化的区域
- 使得选的点间隔足够远

## ➤ 减阶基的选取 $W, V$

- 本征正交基
- 平衡截断方法
- 矩匹配方法

## ➤ 权重的选取 $\{\tilde{w}_i(\mathbf{q}(t))\}_{i=1}^S$

-  $\mathbf{q}_i = (V^T V)^{-1} V^T \mathbf{u}_i$

-  $\tilde{w}_i(\mathbf{q}(t)) = \frac{\exp\left(-\frac{\beta d_i^2}{d^2}\right)}{\sum_{i=1}^S \exp\left(-\frac{\beta d_i^2}{d^2}\right)}$

$d_i = \|\mathbf{q}(t) - \mathbf{q}_i\|_2 \quad d = \min\{d_i\}$



# 轨迹分段线性方法

## ➤ 优点

- 降阶模型的成本不会随着高维模型的规模  $N$  增长。
- 降阶模型不需使用高维模型(比如计算  $f$ )。

## ➤ 缺点

- 选择好的线性化点至关重要，这一过程需要在训练阶段完成。
- 要求能够从高维模型的代码中提取雅可比矩阵。
- 有许多参数需要调整（例如线性化点的数量、权重等）。



# 本堂课大纲

- 非线性动力系统
- 轨迹分段线性方法
- 离散经验插值法



# 人脸重构

- 人脸数据集，基底  $V_f \in R^{N \times m'}$
- 新人脸的数据  $\tilde{f}$ ，但只是在一些像素上有数据 (掩码矩阵  $P \in R^{N \times m}$ )

$$\tilde{f} = P^T f$$

- 重构

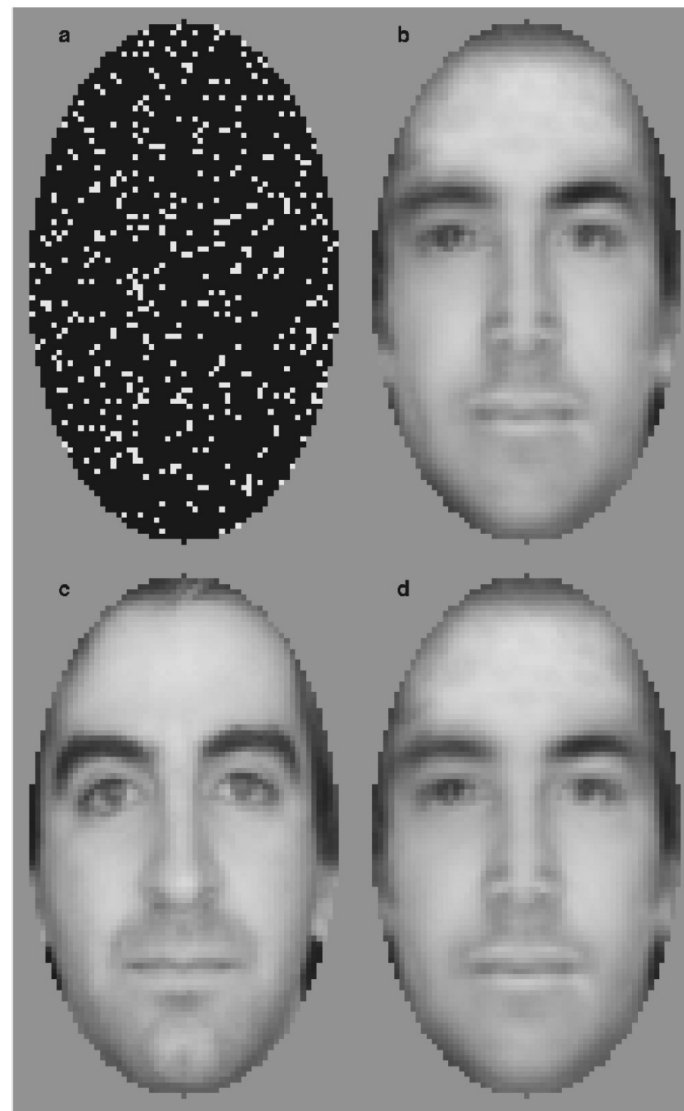
$$f = V_f \left( \tilde{V}_f^T \tilde{V}_f \right)^{-1} \tilde{V}_f^T \tilde{f} \quad \tilde{V}_f = P^T V_f$$

图a :  $\tilde{f}$

图c :  $f$

图b :  $V_f (V_f^T V_f)^{-1} V_f^T f$

图d :  $V_f \left( \tilde{V}_f \tilde{V}_f^T \right)^{-1} \tilde{V}_f^T \tilde{f}$





# 离散经验插值法

## ➤ 离散经验插值

如何近似  $\underbrace{V^T}_{k \times N} \underbrace{f(Vq(t))}_{N \times 1}$

复杂度： $O(2kN) \rightarrow O(2km)$

$$\underbrace{V^T}_{k \times N} \underbrace{f(Vq(t))}_{N \times 1} \approx \underbrace{V[i_1, i_2, \dots, i_m]^T}_{k \times m} \underbrace{f(Vq(t))[i_1, i_2, \dots, i_m]}_{m \times 1}$$

定义掩码矩阵： $P = [e_{i_1}, e_{i_2}, \dots, e_{i_m}] \in R^{N \times m}$

$$f(Vq(t))[i_1, i_2, \dots, i_m] = P^T f(Vq(t)) = \hat{f}(\hat{P}^T Vq(t)) \in R^m$$



# 离散经验插值法

## ➤ 离散经验插值

定义掩码矩阵： $P = [e_{i_1}, e_{i_2}, \dots, e_{i_m}] \in R^{N \times m}$

$$f(Vq(t))[i_1, i_2, \dots, i_m] = P^T f(Vq(t)) = \hat{f}(\hat{P}^T Vq(t)) \in R^m$$

如何选掩码矩阵 $P$ ？



# 离散经验插值法

## ➤ 离散经验插值

假设  $f$  也在一个低维空间中，选定基底：

$$V_f = [u_1 \ u_2 \ \cdots \ u_{m'}] \in R^{N \times m'} \quad f_r(t) \in R^{m'}$$

低维近似： $f(Vq(t)) \approx V_f f_r(t)$

给定  $P \in R^{N \times m}$  和任意  $f \in R^N$ ，通过  $P^T f \approx P^T V_f f_r$

我们可以重构：

$$f_r = (P^T V_f)^\dagger P^T f$$

选掩码矩阵  $P$ ，使得  $f - V_f f_r$  尽量小



# 离散经验插值法

## ➤ 离散经验插值

当  $m \geq m'$  :

$$\hat{f}(\cdot) = V_f (P^T V_f)^\dagger P^T f(\cdot) = \Pi_{V_f, P} f$$

最小二乘问题 :

$$f_r(t) = \operatorname{argmin}_{y_r} \|P^T f - P^T V_f y_r\|$$

奇异值分解 :  $P^T V_f = U \Sigma Z^T$

$$\Sigma = \operatorname{diag}(\sigma_1, \sigma_2, \dots, \sigma_r, 0, \dots, 0)$$

广义逆矩阵 :  $(P^T V_f)^\dagger = Z \Sigma^\dagger U^T$

$$\Sigma^\dagger = \operatorname{diag}\left(\frac{1}{\sigma_1}, \frac{1}{\sigma_2}, \dots, \frac{1}{\sigma_r}, 0, \dots, 0\right)$$



# 离散经验插值法

## ➤ 贪心算法

输入：  $V_f = [u_1, u_2, \dots, u_m] \in R^{N \times m}$

输出：  $[i_1, i_2, \dots, i_m] \in R^m$

$i_1 = \operatorname{argmax}\{|u_1|\}$

$V_f = [u_1], P = [e_{i_1}]$

For  $l = 2:k$  do

    求解  $c : (P^T V_f)c = P^T u_l$

$r = u_l - V_f c$

$i_l = \operatorname{argmax}\{|r|\}$

$V_f = [V_f \ u_l], P = [P \ e_{i_l}]$



# 离散经验插值法

## 收敛性

对任意向量  $f \in R^N$ ，和正交基底  $V_f \in R^{N \times m}$ ，使用离散经验插值法近似出来的

$$\hat{f} = V_f (P^T V_f)^{-1} P^T f$$

满足

$$\|f - \hat{f}\|_2 \leq C \varepsilon_{V_f}$$

其中  $\varepsilon_{V_f} = \|(I - V_f V_f^T) f\|_2$  是投影误差， $C = \|(P^T V_f)^{-1}\|_2$  有上界

$$C \leq (1 + \sqrt{2N})^{m-1} \|u_1\|_\infty^{-1}$$



# 离散经验插值法

## ➤ 练习

$$s(x, y; \mu) = \frac{1}{\sqrt{(x - \mu_1)^2 + (x - \mu_2)^2 + 0.1^2}}$$

$$(x, y) \in [0.1, 0.9]^2 \in R^2, (\mu_1, \mu_2) \in [-1.0, -0.01]^2 \in R^2$$

生成  $[0.1, 0.9]^2$  上的均匀网格 ( $n_x = n_y = 20$ )

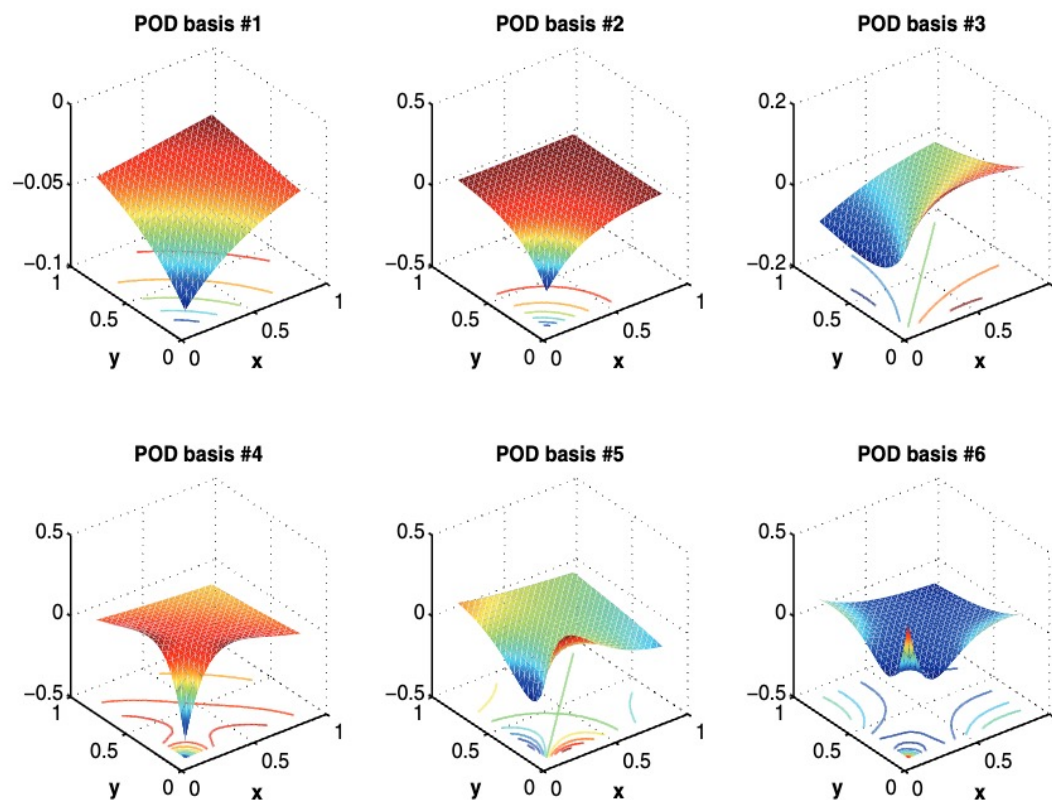
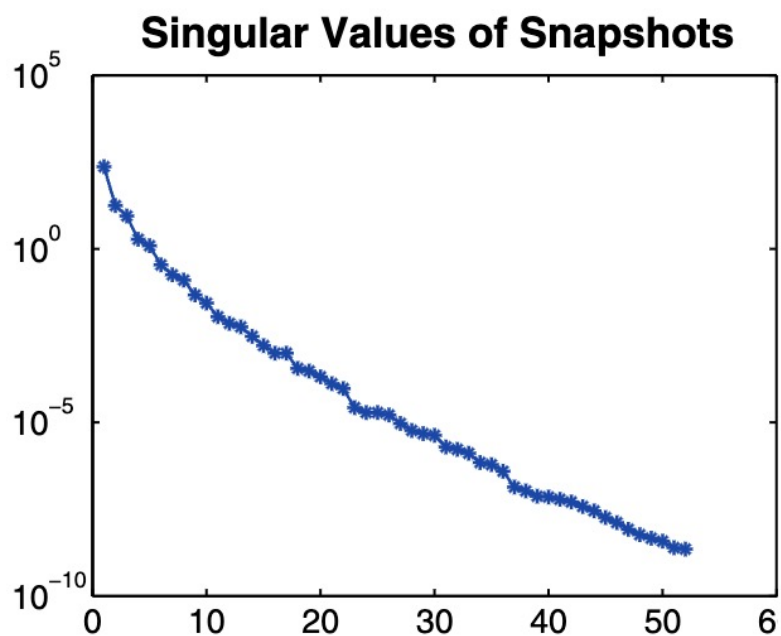
$$s(\mu) = [s(x_i, y_j, \mu)] \in R^{n_x \times n_y}$$



# 离散经验插值法

## 练习

随机采样 $\mu$ ，生成基底 $V_f$

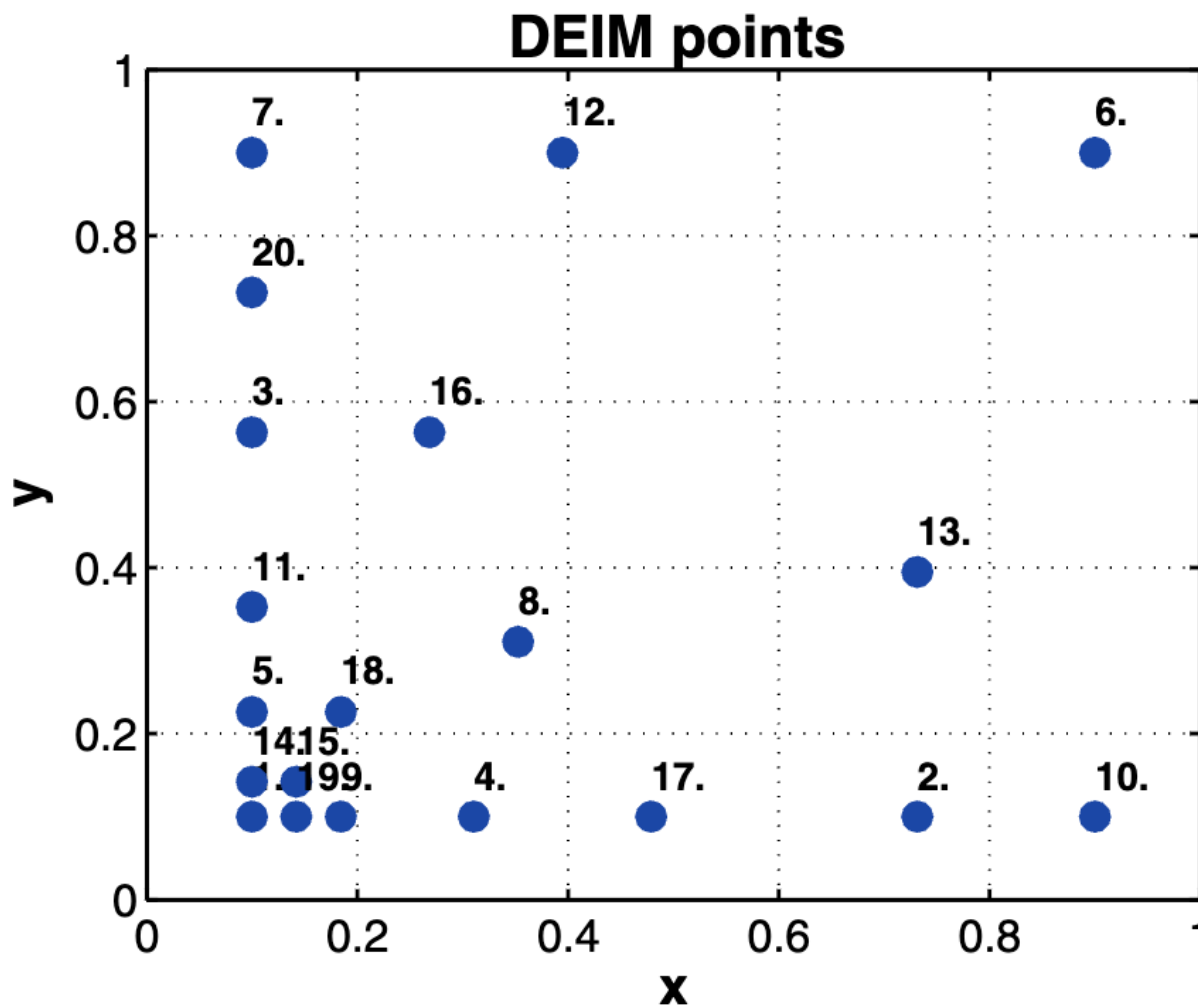




# 离散经验插值法

➤ 练习

选取掩码矩阵  $P$





# 离散经验插值法

## 练习

$$s(\mu) \approx V_f (P^T V_f)^{-1} P^T s(\mu)$$

Full dim= 400,  $[\mu_1, \mu_2] = [-0.05, -0.05]$

POD: dim = 6,  $L^2$  error:  $8.2e-3$

DEIM: dim = 6,  $L^2$  error:  $1.8e-2$

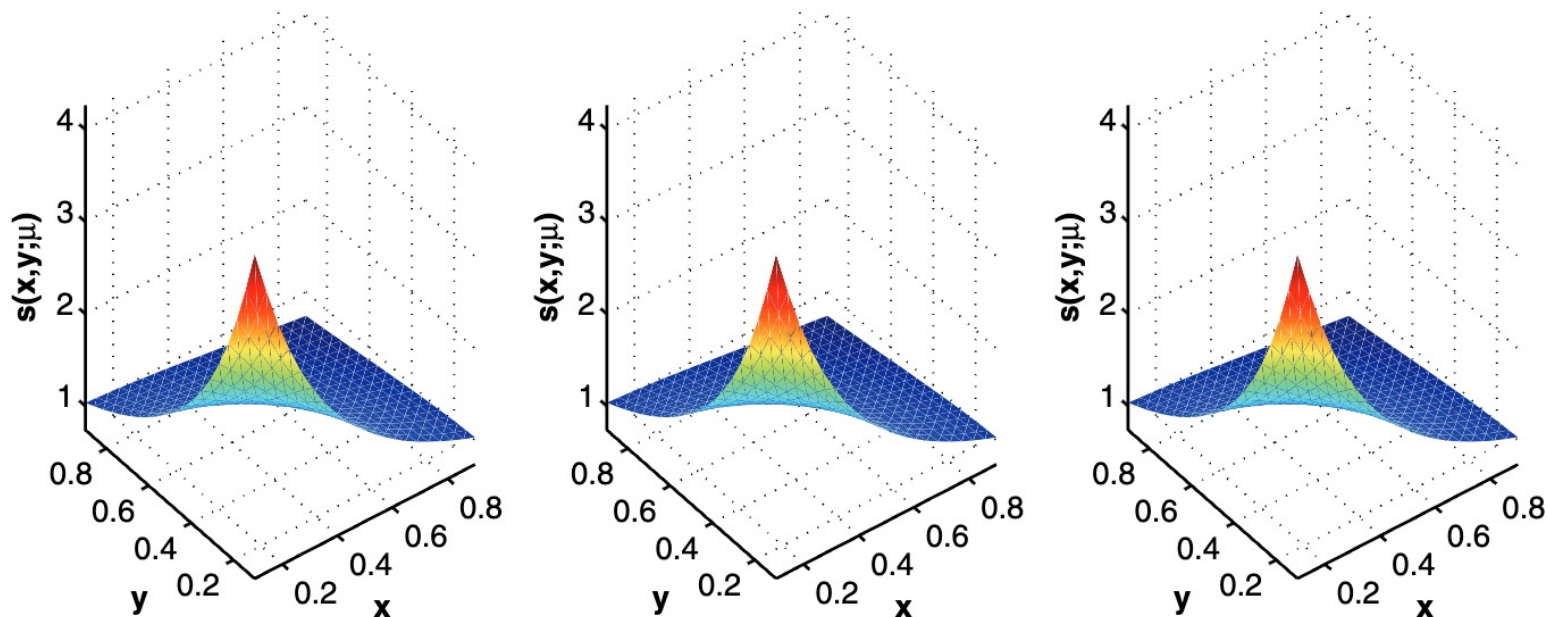


FIG. 3.7. Compare the original nonlinear function (3.39) of dimension 400 with the POD and DEIM approximations of dimension 6 at parameter  $\mu = (-0.05, -0.05)$ .



# 离散经验插值法

## 练习

### 离散经验插值法的误差

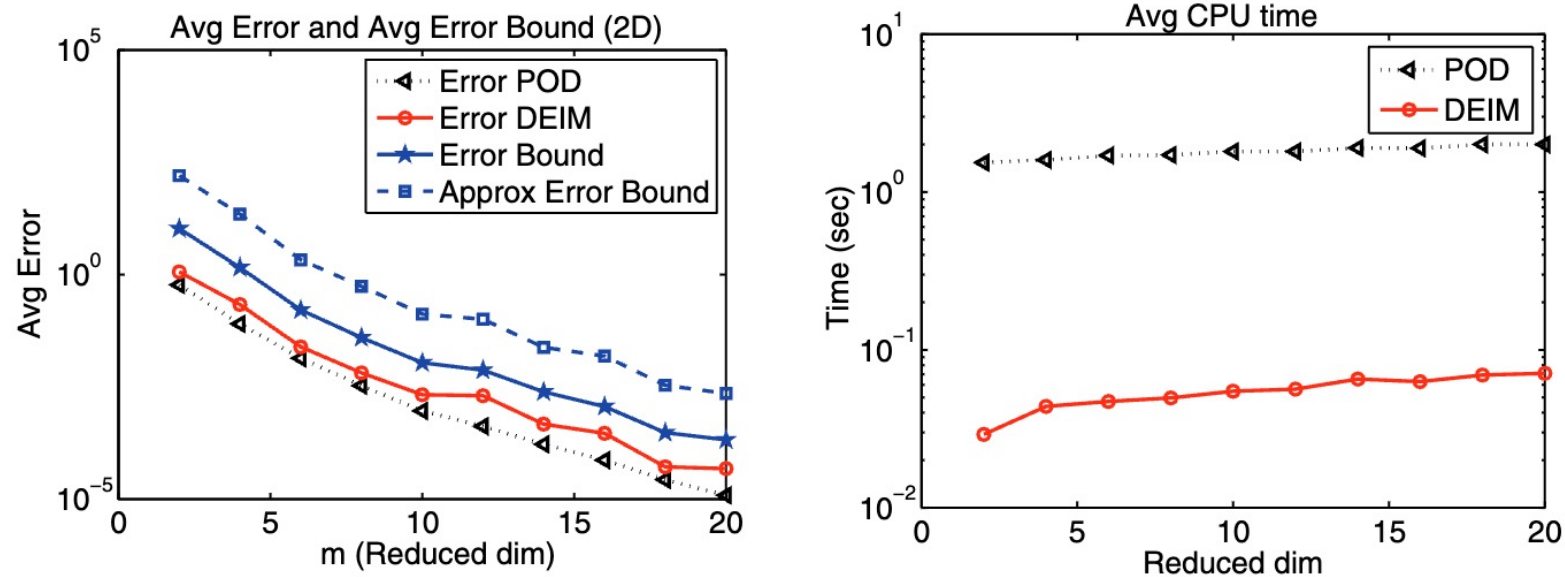


FIG. 3.8. Left: Average errors of POD and DEIM approximations for (3.39) with the average error bounds given in (3.34) and their approximations given in (3.35). Right: Average CPU time for evaluating the POD and DEIM approximations.



# 离散经验插值法

➤ 练习：非线性偏微分方程问题

$$\begin{aligned} -\nabla^2 u(x, y) + s(u; \mu) &= 100 \sin(2\pi x) \sin 2\pi y, & (x, y) \in \Omega \\ u(x, y) &= 0 & (x, y) \in \partial\Omega \end{aligned}$$

$$s(u; \mu) = \frac{\mu_1}{\mu_2} (e^{-\mu_2 u} - 1)$$

$$\Omega = [0, 1]^2 \in R^2, \quad (\mu_1, \mu_2) \in [0.01, 10]^2 \in R^2$$

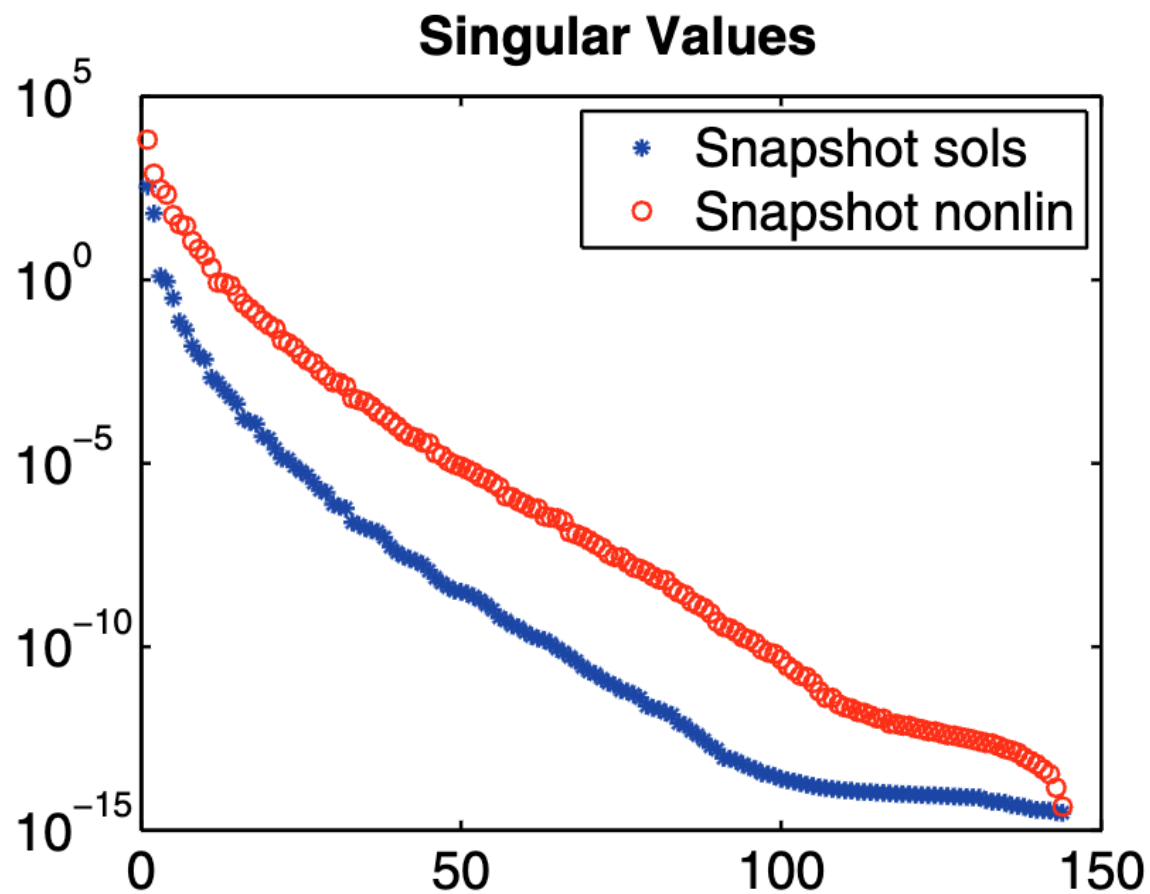
生成  $[0, 1]^2$  上的均匀网格 ( $n_x = n_y = 50$ )

$$s(\mu) = [s(x_i, y_j, \mu)] \in R^{n_x \times n_y}$$



# 离散经验插值法

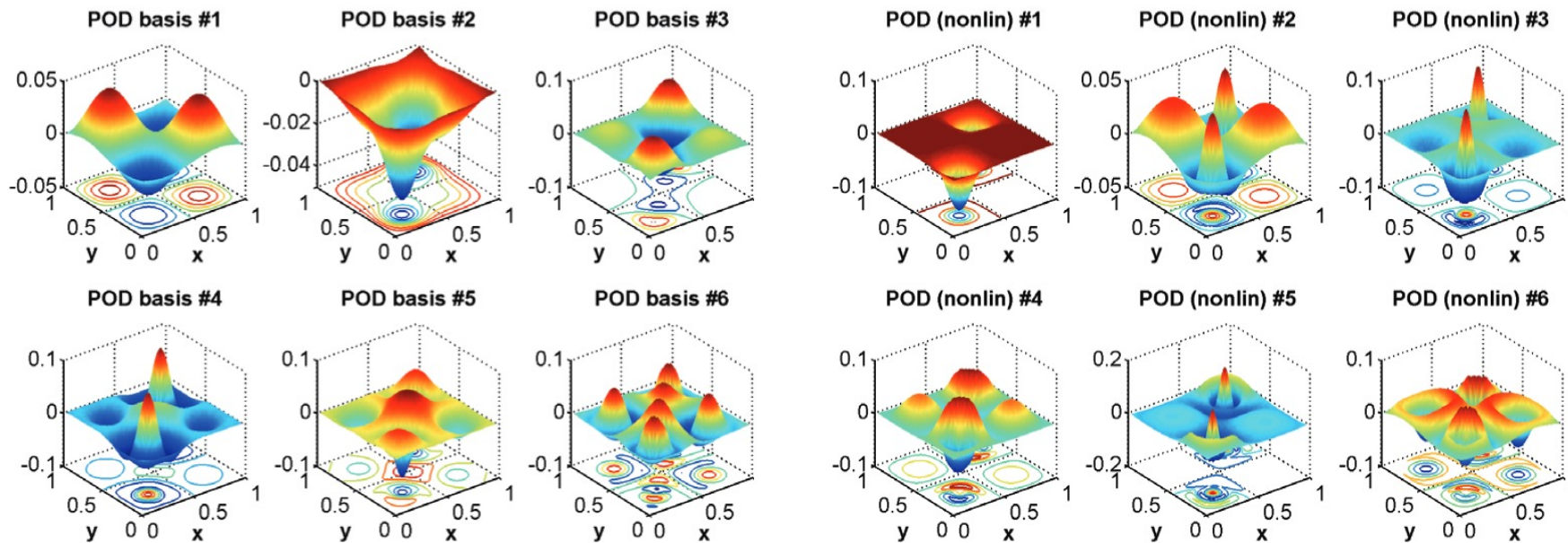
➤ 练习：非线性偏微分方程问题





# 离散经验插值法

## ➤ 练习：非线性偏微分方程问题





# 离散经验插值法

➤ 练习：非线性偏微分方程问题

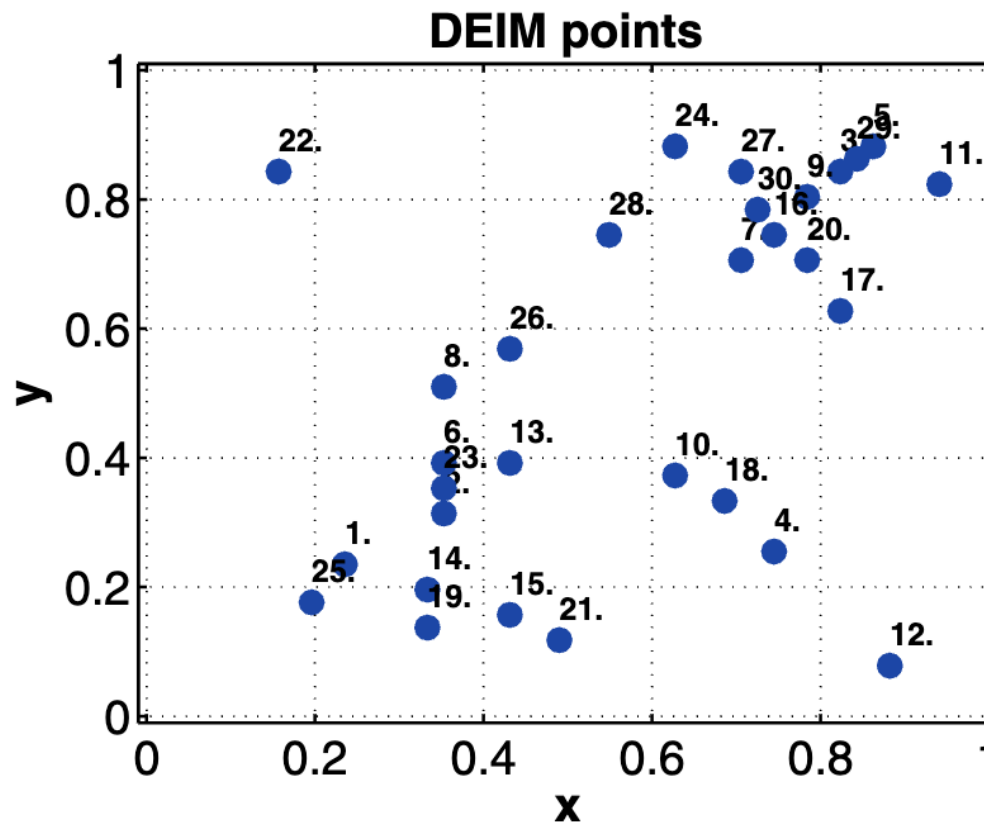
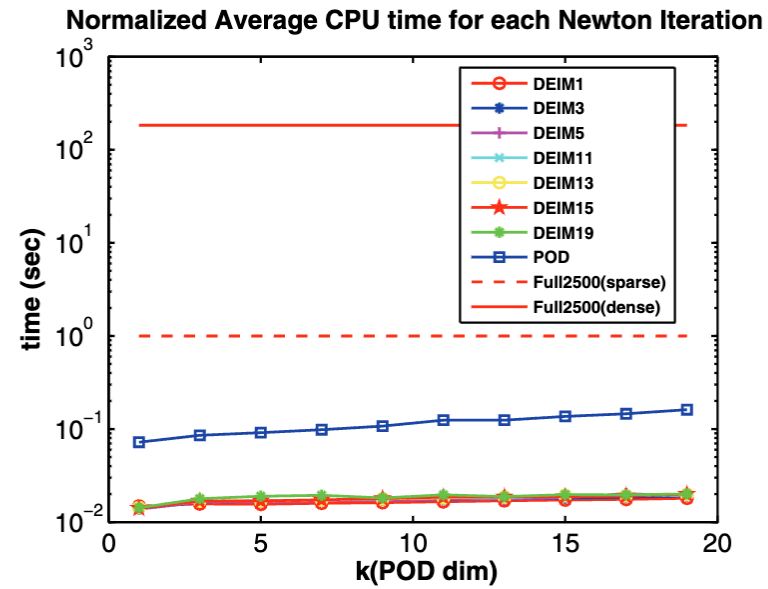
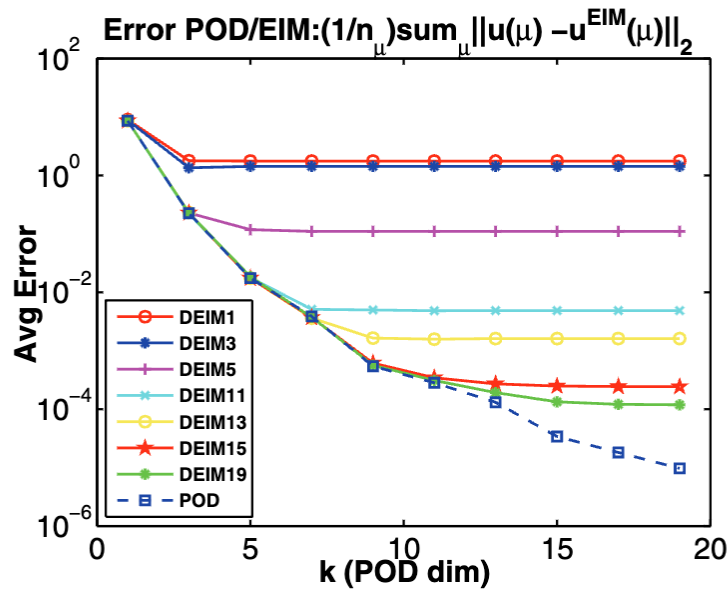
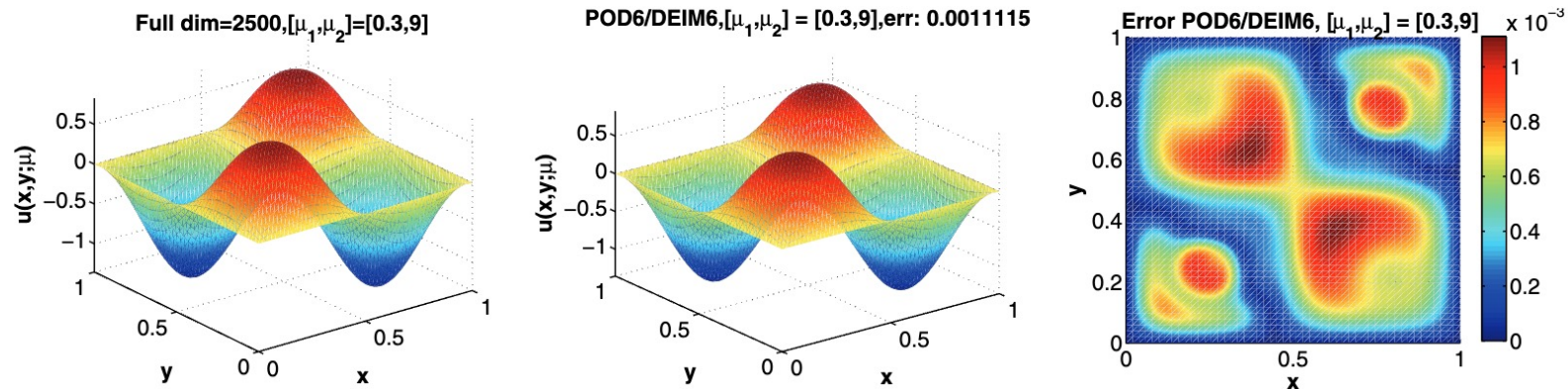


FIG. 4.6. First 30 points selected by DEIM.



# 离散经验插值法

## 练习：非线性偏微分方程问题

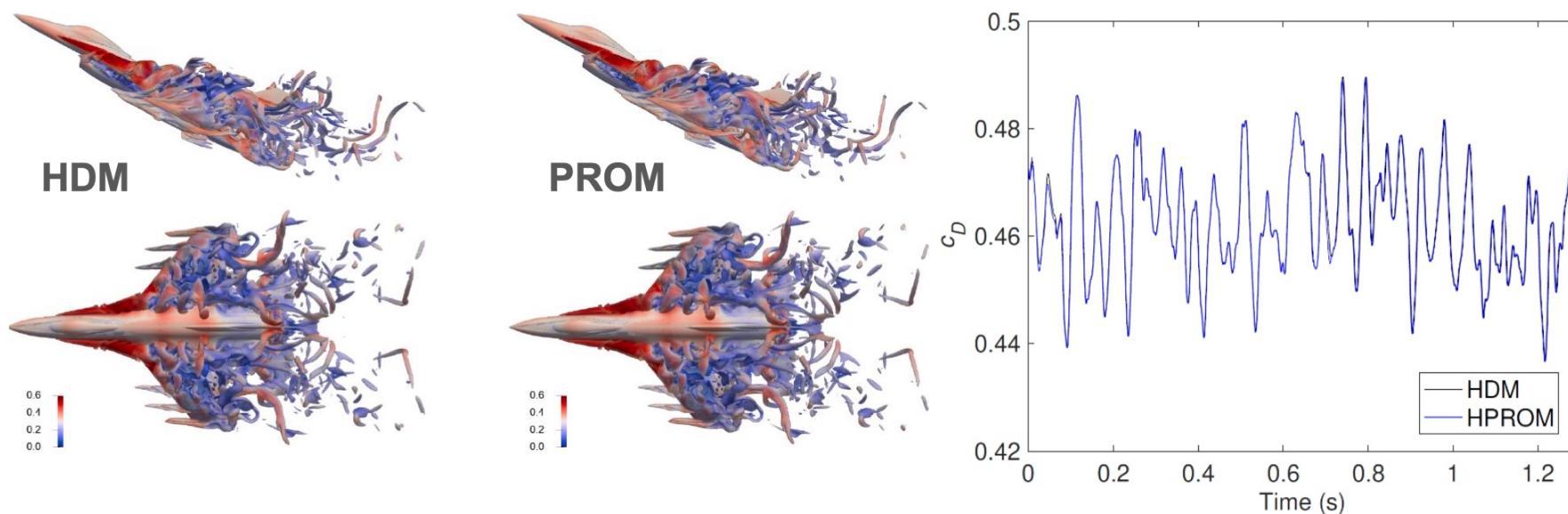




# 针对非线性系统的降阶模型

➤ F-16 (Re=182, 200, 000)

- 计算流体网格26, 919, 873格点
- 模拟仿真时间  $t = 1.29s$



高维计算模型：100h 3584CPU核

基于投影的降阶模型：5.8min, 32CPU核

- $10^3$ 倍提速（挂钟时间）
- $10^5$ 倍提速（CPU时间）



# 离散经验插值法

## ➤ 优点

- 降阶模型的成本不会随着高维模型的规模  $N$  增长。
- 对于偏离原始训练轨迹的测试仍具有较强的鲁棒性。
- 掩码矩阵  $P$  有物理对应。

## ➤ 缺点

- 降阶模型需要使用高维模型(比如计算  $f$ )

$$P^T f(V\mathbf{q}(t)) = \hat{f}(\hat{P}^T V\mathbf{q}(t))$$

写代码并不容易



# 扩展阅读

## ➤ 文献

Everson, Richard, and Lawrence Sirovich. "Karhunen – Loeve procedure for gappy data." *JOSA A* 12, no. 8 (1995): 1657-1664.

轨迹分段线性方法 : Rewieński, Michał, and Jacob White. "Model order reduction for nonlinear dynamical systems based on trajectory piecewise-linear approximations." *Linear algebra and its applications* 415, no. 2-3 (2006): 426-454.

离散经验插值法 : Chaturantabut, Saifon, and Danny C. Sorensen. "Nonlinear model reduction via discrete empirical interpolation." *SIAM Journal on Scientific Computing* 32, no. 5 (2010): 2737-2764.

针对非线性系统的降阶模型 : Carlberg, Kevin, Charbel Bou - Mosleh, and Charbel Farhat. "Efficient non - linear model reduction via a least - squares Petrov – Galerkin projection and compressive tensor approximations." *International Journal for numerical methods in engineering* 86, no. 2 (2011): 155-181.