本征正交分解

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

- 本征正交分解（Proper Orthogonal Decomposition）
  - 直接分解方法
  - 奇异值分解
- 降阶模型算法
- 降阶模型实例

解的近似

- 常微分方程(ODE)

$$\frac{d\boldsymbol{u}}{dt} = f(\boldsymbol{u}(t), t, x)$$

其中 $\boldsymbol{u}(t) \in R^N$是状态变量

$\boldsymbol{u}(0) = \boldsymbol{u}_0$是初始条件

- 输出

$$y(t) = g(\boldsymbol{u}(t), t)$$

解的近似

- 降阶模型

$$\frac{d}{dt}\left(V\boldsymbol{q}(\mathrm{t})\right) = f(V\boldsymbol{q}(\mathrm{t}), t, x) + r(t)$$

由于基底V并不依赖于时间

$$V\frac{d}{dt}\left(\boldsymbol{q}(\mathrm{t})\right) = f(V\boldsymbol{q}(\mathrm{t}), t, x) + r(t)$$

$$y(t) \approx g(V\boldsymbol{q}(\mathrm{t}), t)$$

N个微分方程组成的一组方程，其中涉及k个未知函数

$$\boldsymbol{q}(\mathrm{t}) = [q_1(t), q_2(t) \cdots\cdots q_k(t)]$$

超定系统 $k < N$

本征正交分解方法

- 本征正交分解

  固定初始条件 $u_0 \in R^N$

  给定$0 \leq t \leq T$的解u_t，本征正交分解方法寻求一个固定秩 k 的正交投影 $\Pi_{V,V}$ ，以最小化积分投影误差

$$J(\Pi_{V,V}) = \int_0^T \left\| u(t) - \Pi_{V,V} u(t) \right\|_2^2 dt = \int_0^T \left\| \mathcal{E}_{V^\perp}(t) \right\|_2^2 dt$$

本征正交分解方法

本征正交分解

定义对称正定矩阵$\hat{K} \in R^{N\times N}$

$$\hat{K} = \int_0^{\mathrm{T}} u(t)u(t)^T dt$$

$\hat{K}$的特征值为$\hat{\lambda}_1 \geq \hat{\lambda}_2 \geq \cdots \geq \hat{\lambda}_N \geq 0$, 对应的特征向量为$\hat{\phi}_i \in R^N$，满足

$$\hat{K}\hat{\phi}_i = \hat{\lambda}_i\hat{\phi}_i，i = 1,2,\cdots,N$$

假设$\hat{\lambda}_k > \hat{\lambda}_{k+1}$，那么使$J\left(\Pi_{V,V}\right)$达到最小值的子空间$\mathcal{V} = \mathrm{range}(\hat{V})$，是矩阵$\hat{K}$关于特征值$\hat{\lambda}_1$，$\hat{\lambda}_2$，$\cdots$，$\hat{\lambda}_k$的不变子空间。

本征正交分解方法

- 求解特征值问题

  一般情况，求解特征值问题$\hat{K}\hat{\phi}_i = \hat{\lambda}_i\hat{\phi}_i$在计算上难以处理
  - 矩阵$\hat{K}$的维度N通常很大
  - 矩阵$\hat{K}$通常是稠密矩阵

  状态数据通常以离散的向量形式(snapshot)提供

$$\{u(t_i)\}_{i=1}^{N_{snap}}$$

  在这种情况下$\int_0^{\mathrm{T}} u(t)u(t)^T dt$可以利用数值积分规则进行近似

$$\int_0^T u(t)u(t)^T dt \approx \sum_{i=1}^{N_{snap}} \alpha_i u(t_i)u(t_i)^T$$

本征正交分解方法

- 求解特征值问题

定义快照(snapshot)矩阵$S \in R^{N \times N_{snap}}$

$$S = \left[\sqrt{\alpha_1}u(t_1); \sqrt{\alpha_2}u(t_2); \cdots\cdots \sqrt{\alpha_{N_{snap}}}u\left(t_{N_{snap}}\right)\right]$$

那么

$$K = SS^T$$

目标：近似K的前k个特征值和特征向量。

本征正交分解方法

- 求解特征值问题

  定义$R = S^T S \in R^{N_{snap} \times N_{snap}}$，假设$Rank(R) = r$，求解特征值问题有

$$R\psi_i = \lambda_i \psi_i \qquad i = 1,2,\cdots r$$

那么$K = SS^T$的前r个特征向量满足

$$\phi_i = \frac{1}{\sqrt{\lambda_i}} S\psi_i$$

复杂度：$\mathcal{O}(N_{snap}^2 N)$

条件数（condition number）：$\mathrm{cond}(R) \approx \mathrm{cond}(S)^2$

奇异值分解

➢ 奇异值分解

给定矩阵 $A \in R^{N\times M}$，存在两个正交矩阵，$U \in R^{N\times N}(UU^T = I)$，$V \in R^{M\times M}(VV^T = I)$，使得

$$A = U\Sigma V^T$$

其中 $\Sigma \in R^{M\times N} = \mathrm{diag}\{\sigma_1, \sigma_2, \cdots, \sigma_{\min\{M,N\}}\}$ 是对角矩阵，满足

$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_{\min\{M,N\}} \geq 0$

那么 $\{\sigma_i\}_{i=1}^{\min\{M,N\}}$ 是矩阵 A 的奇异值，而 U 和 V 的列分别是 A 的左奇异向量和右奇异向量

$U = [u_1\ u_2 \cdots\cdots\ u_N]$ $\qquad$ $V = [v_1\ v_2\ \cdots\cdots v_N]$

奇异值分解

- 奇异值分解的性质

  $\mathrm{Rank}(A)=r$，其中 r 是最小非零奇异值的序号。

  定义非零奇异值 $\Sigma_r=\mathrm{diag}\{\sigma_1,\sigma_2,\cdots,\sigma_r\}$的左、右奇异向量

  $$U_r=[u_1\ u_2\ \cdots\cdots u_r],\quad V_r=[v_1\ v_2\ \cdots\cdots v_r]$$

  以及

  $$U_{N-r}=[u_{r+1}\ u_{r+2}\cdots\cdots u_N],\quad V_{N-r}=[v_{r+1}\ v_{r+2}\cdots\cdots v_N]$$

  那么

  $$A=\mathrm{U}_r\Sigma_\mathrm{r}V_r^T \qquad A=\sum_{i=1}^{r}\sigma_i u_i v_i^T$$

  $$\mathrm{range}(A)=\mathrm{range}(U_r),\quad \mathrm{range}(A^T)=\mathrm{range}(V_r)$$
  $$\ker(A)=\mathrm{range}(U_{N-r}),\quad \ker(A^T)=\mathrm{range}(V_{N-r})$$

奇异值分解

- 奇异值分解的性质

给定

$$A = U_r \Sigma_r V_r^T \qquad A = \sum_{i=1}^{r} \sigma_i u_i v_i^T$$

我们能得到AA^T和A^TA的特征值分解

$$AA^T = U_r \Sigma_r U_r^T,\quad A^TA = V_r \Sigma_r V_r^T$$

它们的非零特征值为$\sigma_1^2, \sigma_2^2, \cdots, \sigma_r^2$

对应于AA^T的特征向量为$U_r = [u_1\ u_2\ \cdots\cdots\ u_r]$

对应于A^TA的特征向量为$V_r = [v_1\ v_2\ \cdots\cdots\ v_r]$

奇异值分解

奇异值分解(Schmidt-Eckart-Young-Mirsky)

给点矩阵$A \in R^{N\times M}$，以及$k < r = \mathrm{rank}(A)$

$$\min_{\mathrm{rank}(X)=k} \|A - X\|_2 = \sigma_{k+1}(A),\ \sigma_k(A) > \sigma_{k+1}(A)$$

$$\min_{\mathrm{rank}(X)=k} \|A - X\|_F = \sum \sigma_i(A)\,,$$

其中$X = \sum_{i=1}^{r} \sigma_i u_i v_i^T$

本征正交分解方法

- 奇异值分解的性质

  给定快照(snapshot)矩阵$S \in R^{N \times N_{snap}}$的奇异值分解
  $$S = U_r \Sigma_r V_r^T$$

  那么$\widehat{K}$的特征值分解满足
  $$\widehat{K} = SS^T = U_r \Sigma_r^2 U_r^T$$

  正交投影：V由U_r的前k列组成

本堂课大纲

- 本征正交分解（Proper Orthogonal Decomposition）
  - 直接分解方法
  - 奇异值分解
- 降阶模型算法
- 降阶模型实例

算法：降阶模型维度选取

- 维度选取

$\mathcal{E}_{POD}(k)$代表由前k个本征正交基向量捕获的快照的能量

$$\mathcal{E}_{POD}(k)=\frac{\sum_{i=1}^{k}\sigma_i^2(S)}{\sum_{i=1}^{r}\sigma_i^2(S)}$$

选取k使得

$$1-\mathcal{E}_{POD}(k)\leq\epsilon$$

比如$\epsilon=0.1\%$

算法：降阶模型维度选取

- 本征正交分解

  奇异值分解：

$$S \in R^{N \times N_{snap}}，S = U_r \Sigma_r V_r^T$$

  V由U_r的前k列组成

- 误差分析

$$\begin{aligned}\mathcal{E} &= \mathcal{E}_{V^\perp}(t) + \mathcal{E}_V(t) \\ &= (I - \Pi_{V,V})u(t) + V(V^T u(t) - q(t))\end{aligned}$$

$$\left\| \mathcal{E}_{V^\perp}(t_1)\ \mathcal{E}_{V^\perp}(t_2)\ \cdots\ \mathcal{E}_{V^\perp}\left(t_{N_{snap}}\right) \right\|_F = \sqrt{\sum_{i=k+1}^{r} \sigma_i^2(S)}$$

$$\frac{\left\| \mathcal{E}_{V^\perp}(t_1)\ \mathcal{E}_{V^\perp}(t_2)\ \cdots\ \mathcal{E}_{V^\perp}\left(t_{N_{snap}}\right) \right\|_F^2}{\left\| u(t_1)\ u(t_2)\ \cdots u\left(t_{N_{snap}}\right) \right\|_F^2} = 1 - \mathcal{E}_{POD}(k)$$

算法: 快照矩阵的选取

- 含参数不含时间的偏微分方程问题$f(u,\mu)=0$

  - 快照矩阵：

$$S=\mathrm{span}\{u(\mu_j)\}_j^{N_{snap}}$$

$$S=\mathrm{span}\{u(\mu_j),\nabla_\mu u(\mu_j)\dots\dots\}_j^{N_{snap}}$$

  - 参数选取：

用贪心法从集合D中选取后验误差最大

$$\mu_i=\mathrm{argmax}_{\mu\in D}\|\mathcal{E}(\mu)\|=\mathrm{argmax}_{\mu\in D}\|u(\mu)-Vq(\mu)\|$$

$$\mu_i=\mathrm{argmax}_{\mu\in D}\|r(\mu)\|=\mathrm{argmax}_{\mu\in D}\|f(Vq(\mu),\mu)\|$$

算法: 快照矩阵的选取

- 参数选取

  随机选取 μ_1

  求解高维模型$f(u(\mu_1),\mu_1)=0$

  构建快照矩阵，以及相应的基底V

  循环 $i=2,3,\cdots\cdots$

    求解降阶模型： $\mu_i=argmax_{\mu\in D}\|f(Vq(\mu),\mu)\|$

    求解高维模型： $f(u(\mu_i),\mu_i)=0$

    拓展快照矩阵，以及相应的基底V

算法

- 快照（snapshot）矩阵的选取
  - 含参数的含时问题

$$S = \mathrm{span}\{u(t_i;\mu_j)\}_{i,j}^{N_{snap}}$$

$$S = \mathrm{span}\{u(t_i;\mu_j), \nabla_\mu u(t_i;\mu_j) \ldots \ldots\}_j^{N_{snap}}$$

  - 参数选取：

用贪心法从集合D中选取后验误差最大

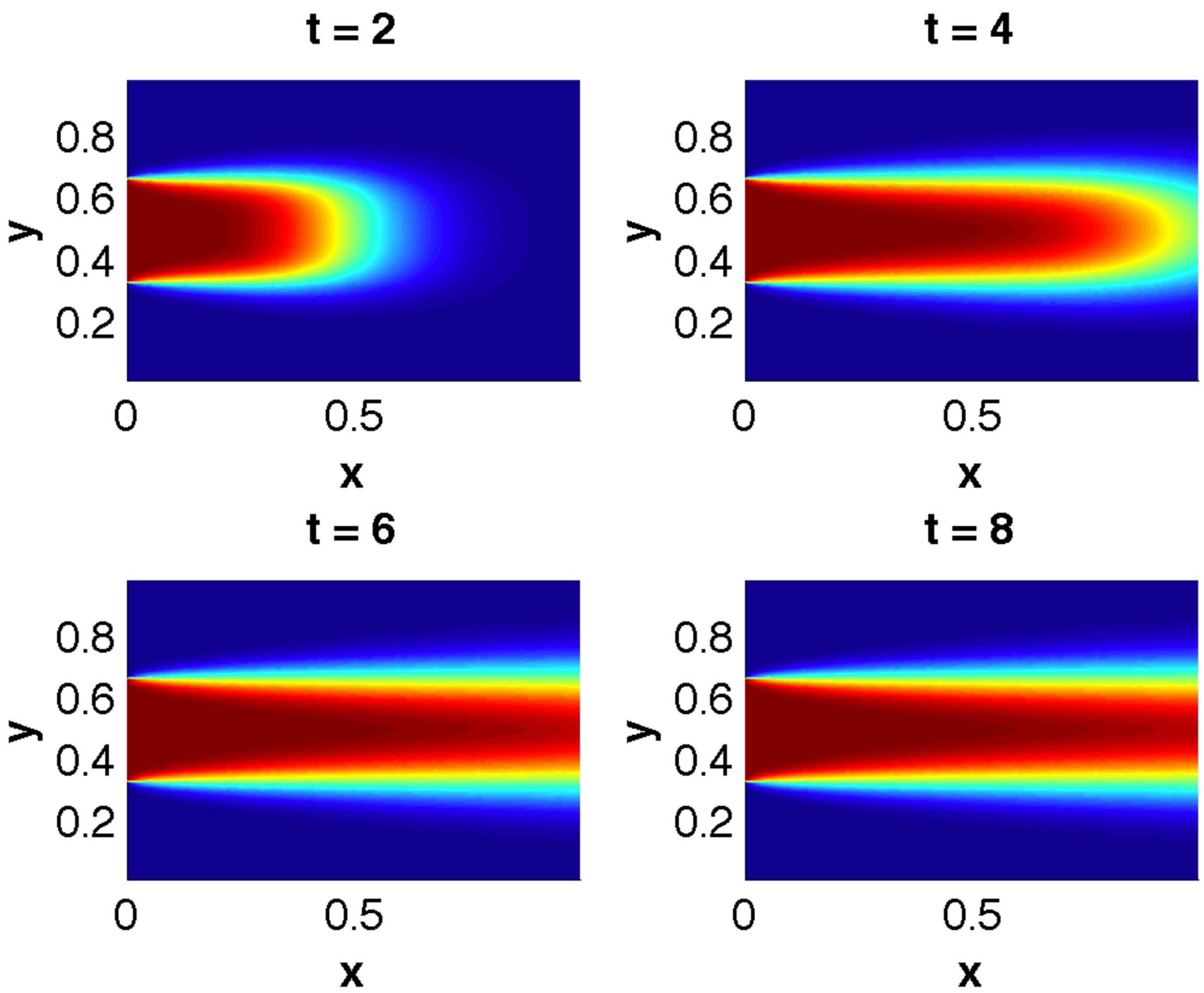
$$\mu_i = \mathrm{argmax}_{\mu\in D}\sqrt{\int_0^T \left\| V\frac{d}{dt}q(\mu) - f(Vq(\mu),\mu)\right\|^2 dt}$$

本堂课大纲

- 本征正交分解（Proper Orthogonal Decomposition）
  - 直接分解方法
  - 奇异值分解
- 降阶模型算法
- 降阶模型实例

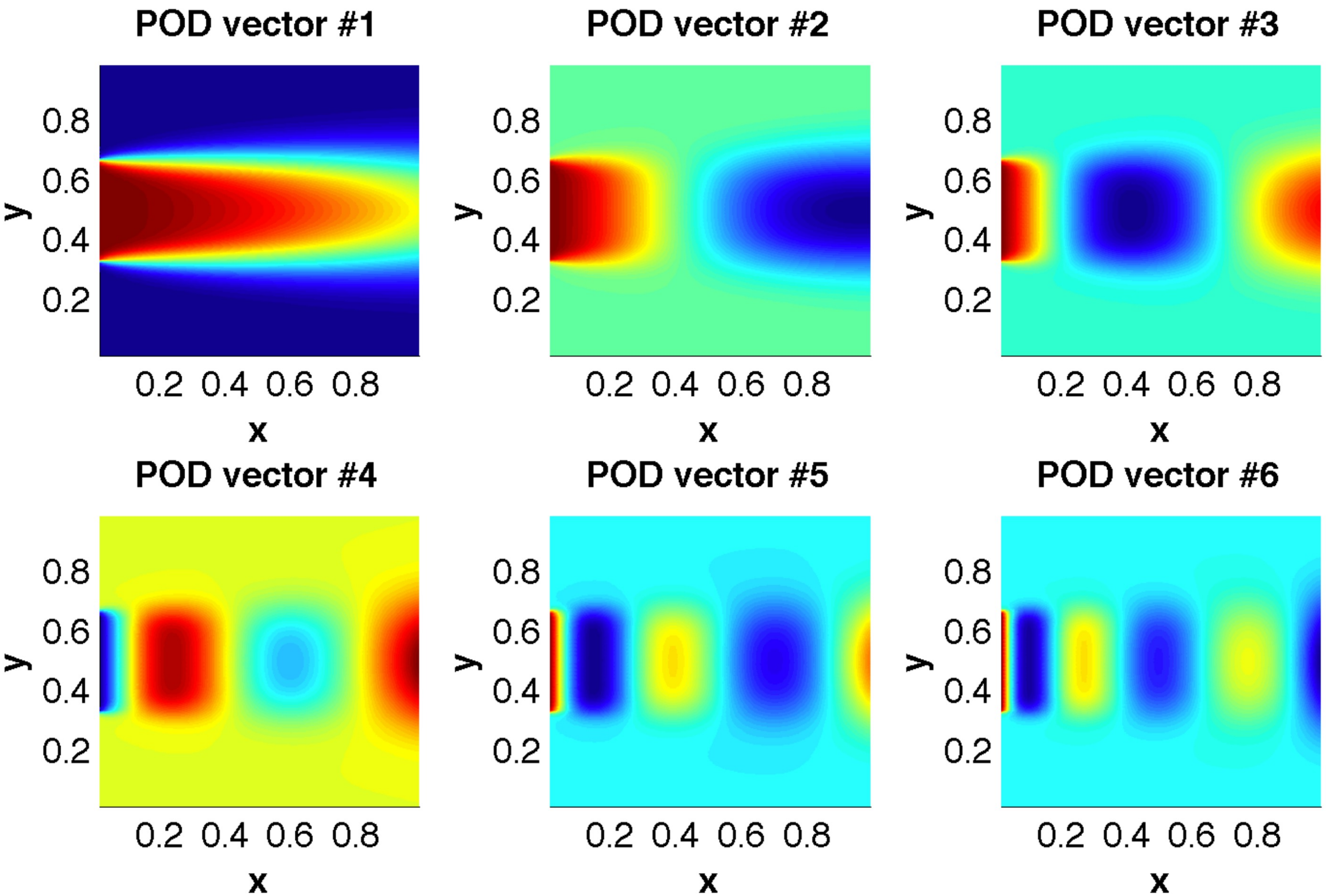
例子

➢ 对流扩散方程（$N = 5402$）



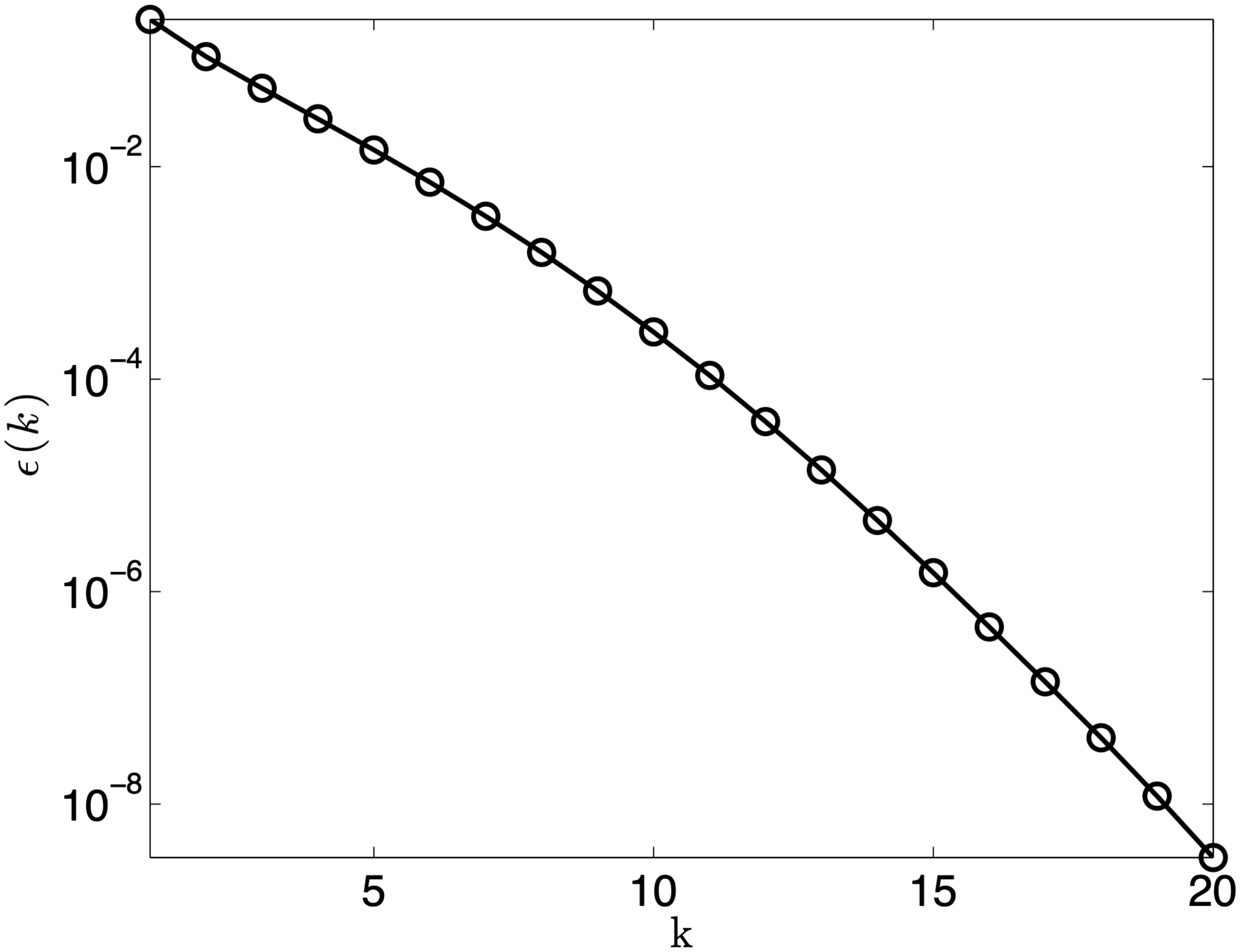
例子

- 特征向量


POD vector #1
POD vector #2
POD vector #3
POD vector #4
POD vector #5
POD vector #6
0.2 0.4 0.6 0.8
0.8
0.6
0.4
0.2
x
y

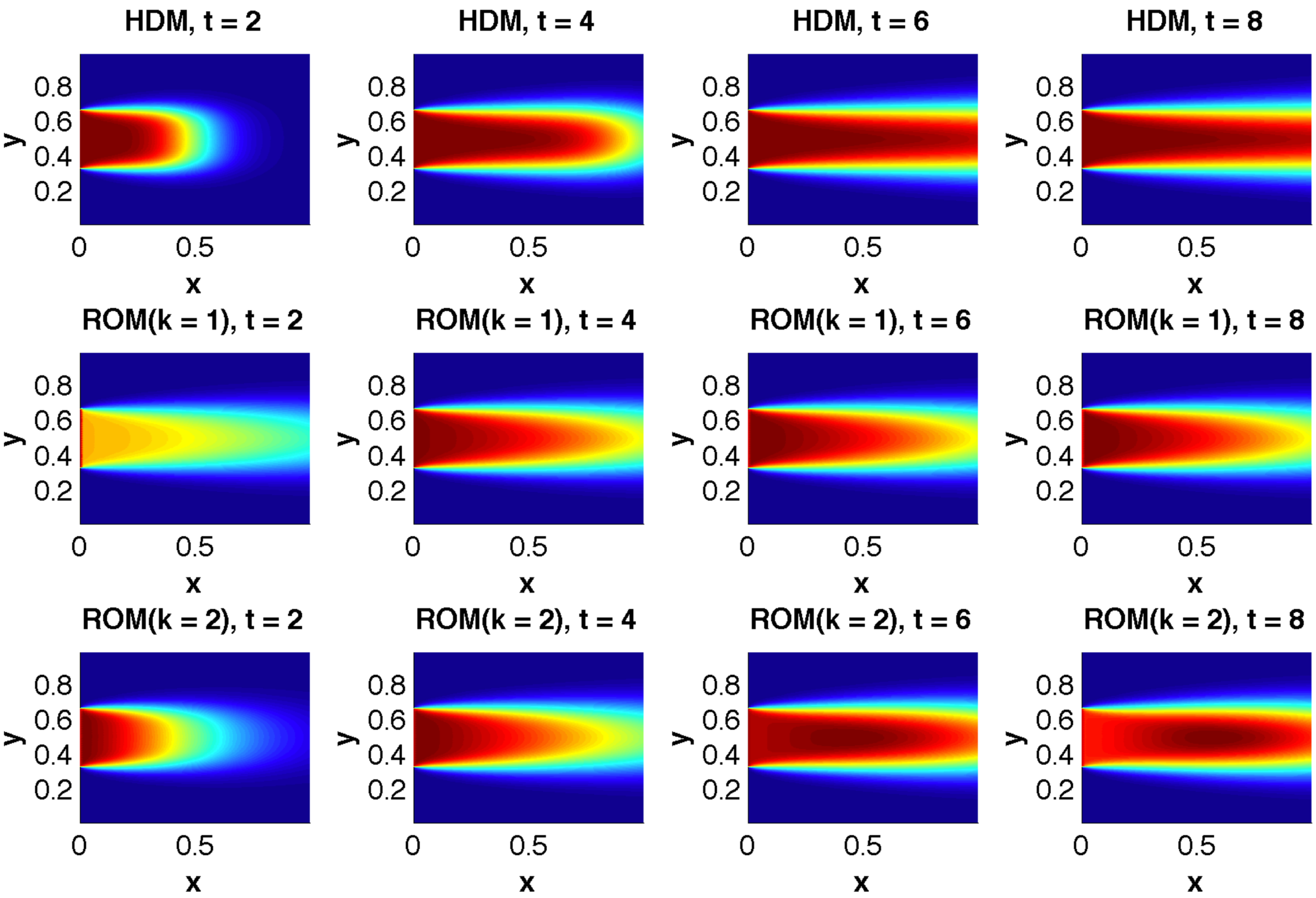
例子

- 投影误差$\mathcal{E}_{V^\perp}$



例子

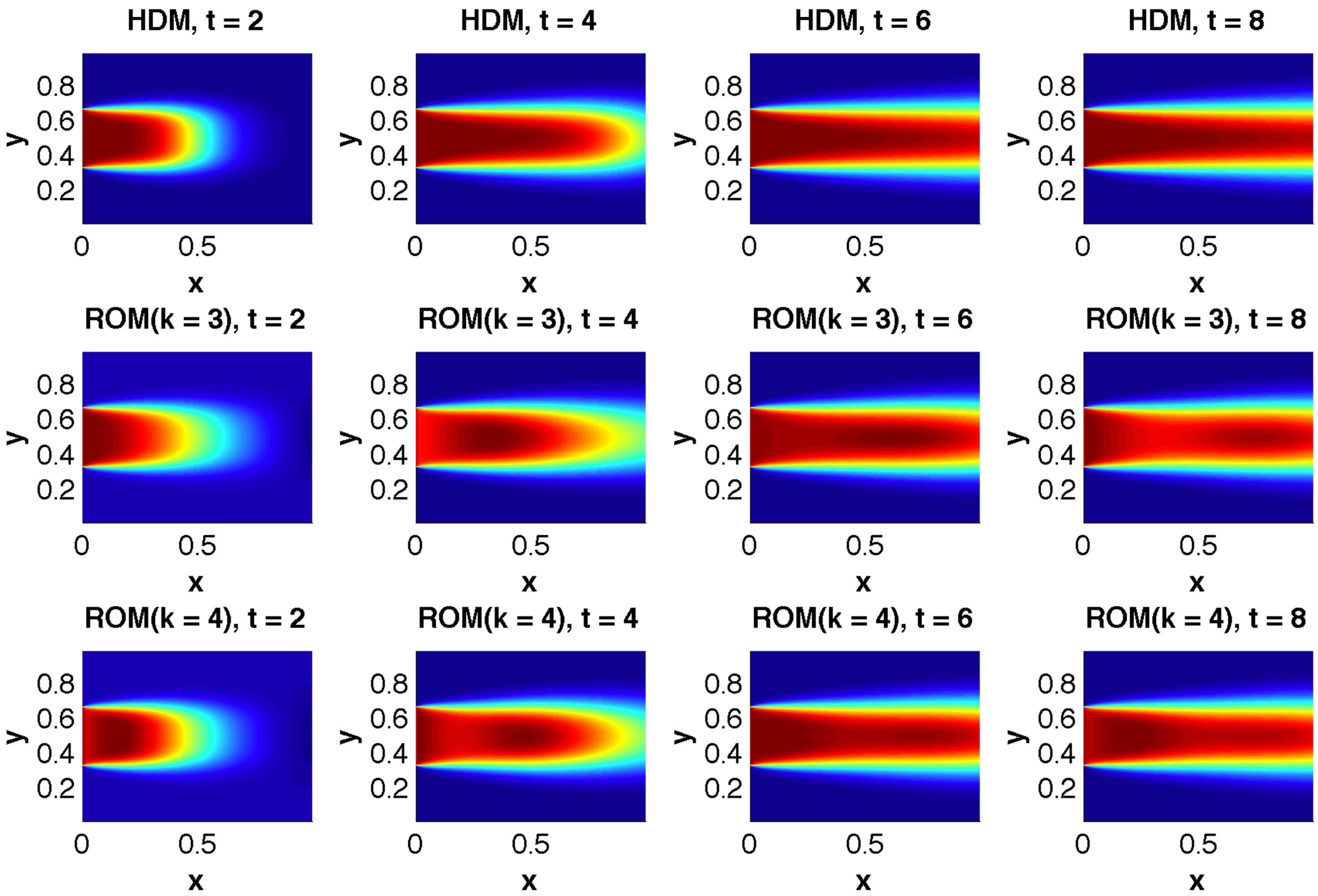
➢ 降阶模型的解



HDM, t = 2
HDM, t = 4
HDM, t = 6
HDM, t = 8
ROM(k = 1), t = 2
ROM(k = 1), t = 4
ROM(k = 1), t = 6
ROM(k = 1), t = 8
ROM(k = 2), t = 2
ROM(k = 2), t = 4
ROM(k = 2), t = 6
ROM(k = 2), t = 8
0.8
0.6
0.4
0.2
y
0
0.5
x

例子

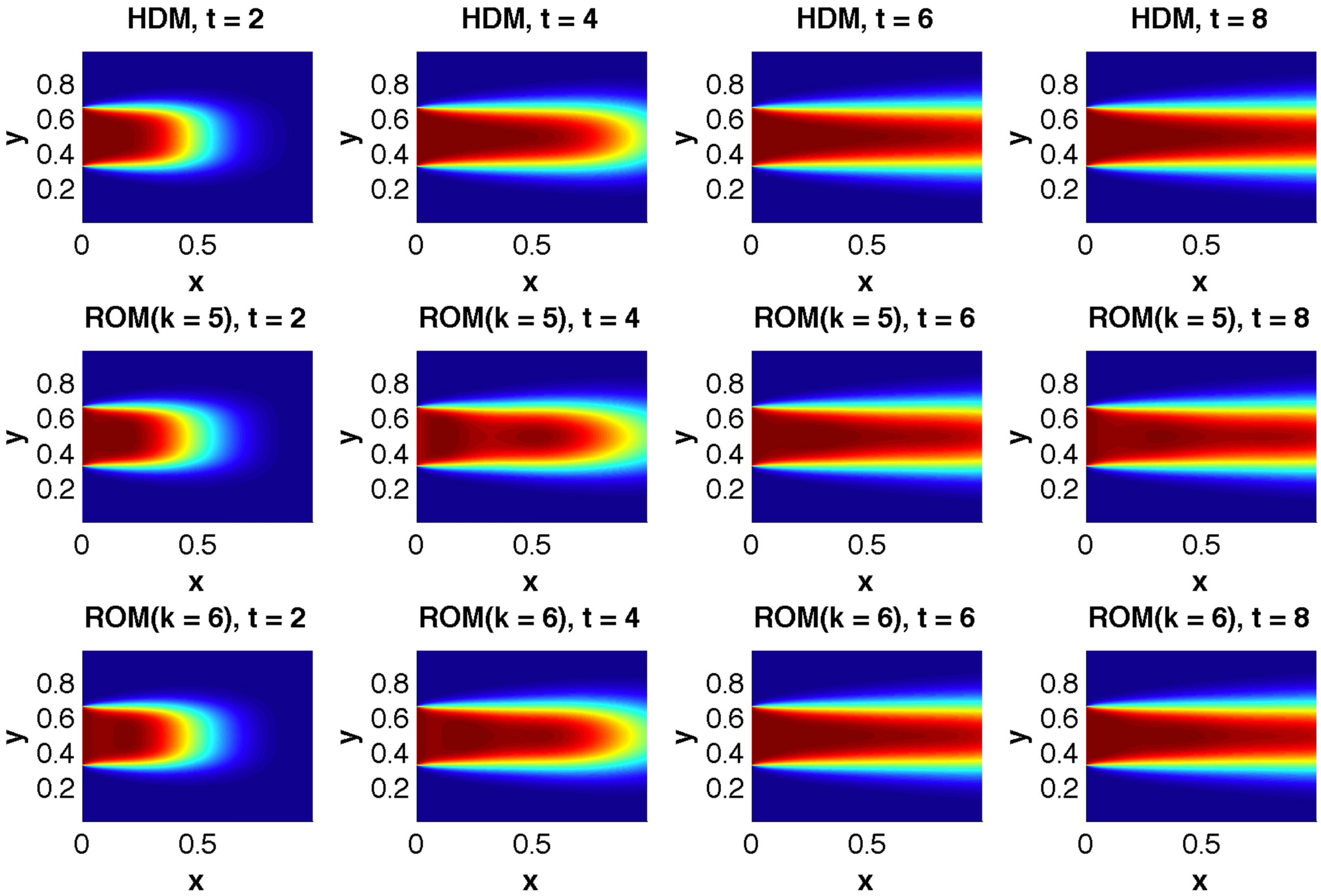
➢ 降阶模型的解



HDM, t = 2
HDM, t = 4
HDM, t = 6
HDM, t = 8
ROM(k = 3), t = 2
ROM(k = 3), t = 4
ROM(k = 3), t = 6
ROM(k = 3), t = 8
ROM(k = 4), t = 2
ROM(k = 4), t = 4
ROM(k = 4), t = 6
ROM(k = 4), t = 8
0.8
0.6
0.4
0.2
y
0
0.5
x

例子

➢ 降阶模型的解


HDM, t = 2
HDM, t = 4
HDM, t = 6
HDM, t = 8
ROM(k = 5), t = 2
ROM(k = 5), t = 4
ROM(k = 5), t = 6
ROM(k = 5), t = 8
ROM(k = 6), t = 2
ROM(k = 6), t = 4
ROM(k = 6), t = 6
ROM(k = 6), t = 8
y
x
0.8
0.6
0.4
0.2
0
0.5

例子

- 降阶模型误差$\mathcal{E}_V$和投影误差$\mathcal{E}_{V^\perp}$