КЛАСТЕРИЗАЦИЯ ФАЗОВОЙ ТРАЕКТОРИИ ДИНАМИЧЕСКОЙ СИСТЕМЫ

В.В. Стрижов, В.В. Шакин

В этой работе описан алгоритм для качественного анализа фазовых траекторий динамических систем.

Введение

Рассмотрим автономную динамическую систему, описываемую системой *n* обыкновенных дифференциальных уравнений

$$\boldsymbol{x}'=\mathrm{F}(\boldsymbol{x})$$

с начальными условиями $\boldsymbol{x}(0)=\boldsymbol{x}_0$, где $\boldsymbol{x}=\boldsymbol{x}(\mathrm{t})$, $\mathrm{t}\in[0,\mathrm{T}]$, $\boldsymbol{x}\in R^n$, где R^n – фазовое пространство системы. Годограф вектор-функции $\boldsymbol{x}=\boldsymbol{x}(t)$ есть линия в фазовом пространстве, называемая фазовой траекторией динамической системы.

Для решения некоторых задач по классификации состояний системы требуется представлять фазовую траекторию в виде последовательности кластеров. Разбиение траектории на кластеры зависит от формы траектории и от задаваемых требований к искомой кластеризации, например, от размерности подпространства $R^r\subseteq R^n$, в котором находятся кластеры траектории, и от размеров параллелепипеда или эллипса, содержащих кластер и лежащих в подпространстве R^r.[1]

Постановка задачи

В качестве тестового примера возьмем фазовую траекторию системы с аттрактором Лоренца. Данная фазовая траектория соответствует решению системы обыкновенных дифференциальных уравнений

$$\begin{aligned}
x_1'(t) &= -3\,(x_1(t) - x_2(t))\\
x_2'(t) &= -x_1(t)\,x_3(t) + 26.5\,x_1(t) - x_2(t)\\
x_3'(t) &= x_1(t)\,x_2(t) - x_3(t)
\end{aligned}$$

при начальных условиях $x_1(0) = x_3(0) = 0$, $x_2(0) = 1$.
Полученная в пространстве R^3 фазовая траектория была линейно отображена в R^n с аддитивным наложением некоррелированного гауссова шума с дисперсией σ^2.

Значения функции $x=x(t)$ представлены в виде последовательности фазовых векторов $\boldsymbol{x_1}, \boldsymbol{x_2}, \ldots \boldsymbol{x_m}$, где $\boldsymbol{x}=\langle x_1, x_2, \ldots x_n\rangle\in R^n$ в моменты времени $t = 0,\ \tau, 2\tau, 3\tau, \ldots, (m-1)\tau$.
Пусть

$$A = (x_{i,j})_{i,j=1}^{m,n}$$

– матрица размерности $m\times n$, n - количество параметров, или размерность пространства, в котором находится вектор $\boldsymbol{x}$; m - количество отсчетов времени, или количество векторов.

Пусть S^p – образ A при линейном отображении f

$$f: (A-B)\rightarrow S^p\ ,$$

где:
A – конечное упорядоченное множество векторов $A = \{\boldsymbol{x_i} : \boldsymbol{x_i}\in R^n, i=1\ldots m\}$,
S^p – конечное упорядоченное множество векторов $S^p = \{\boldsymbol{s_i} : \boldsymbol{s_i}\in R^p, i=1\ldots m\}$.
$B = (b_{i,j})_{i,j=1}^{m,n}$ – матрица невязок, приближающая S^p к $f^{-1}(A)$.

При линейном отображении f каждый вектор $\boldsymbol{x_i}$ отображается соответственно в вектор $\boldsymbol{s_i}$, $i = 1\ldots m$.

Пусть множество S^p состоит из упорядоченных подмножеств S^p_ξ ,

$$S^p = \{\, S^p_\xi : S^p_\xi \subseteq R^{k\times p},\ \xi = 1\ldots k\}$$

и выполняются условия $\bigcup_{\xi=1}^{k} S_{\xi}^{p} = S^{p}$ и $\bigcap_{\xi=1}^{k} S_{\xi}^{p} = \varnothing$. Множество S^p будем называть множеством кластеров фазовой траектории *A*, представленных в пространстве R^p.

Требуется:

1) представить матрицу *A* в виде:

$$A \to S^{p} + B$$

где

$S^{p} = \bigcup_{\xi=1}^{k} S_{\xi}^{p}$ – объединение *k* кластеров фазовой фазовой траектории *A,*

p – заданное число (p $\in$ Z), при соблюдении условия $p \le r$, где $r = \operatorname{rank}(A - B)$.

2) построить:

$$T = \bigcup_{p=1}^{r} \left(\bigcup_{\xi=1}^{k} S_{\xi}^{p} + B \right)$$

где

T – древовидный граф, включающий в себя все множества S^{p}, находящиеся в пространствах $R^1 \ldots R^r$, где r – rank(A - B).

Алгоритмическое решение

1. Нахождение размерности:

Находим фактическую размерность[2] матрицы (A-B), используя сингулярное разложение

$$A = U\, W\, V^{T}$$

Матрицы *U* и *V* являются ортогональными, матрица *W* содержит на своей диагонали убывающие по значению сингулярные числа. Выполняется условие

$$W = \operatorname{diag}(w_1 \le w_2 \le \ldots \le w_r \ldots \le w_n = 0).$$

Размерность матрицы (A-B) соответствует индексу *r* сингулярного числа w_r при выполнении условия

$$\frac{w_r^2}{m} \le \frac{\sigma^2}{m},$$

где σ^2 – наперед заданная величина.

Верны следующие утверждения:

1. При размерности пространства R^p: $p = r$, $r = \operatorname{rank}(A - B)$, фазовая траектория, описанная с помощью матрицы *A* будет представлена в виде только одного кластера $A = S_{1}^{p=r} + B$.
2. Любой кластер фазовой траектории, находящийся в пространстве размерности *p* , может быть представлен в виде объединения кластеров, находящихся в пространстве меньшей размерности *q*, причем $1 \le q \le p \le r$, $r = \operatorname{rank}(A-B)$.

2. Кластеризация:

Рассмотрим процесс кластеризации в пространстве R^1. В первой итерации каждый кластер S_{ξ}^{1} содержит в себе только один вектор $\boldsymbol{x}_i$, $\boldsymbol{x}_i \in A$, причем $\xi = i = 1 \ldots m$.

Дальнейшие итерации производятся следующим образом:

Пусть $S_\xi^1=\{\boldsymbol{x}_\zeta, \zeta=1\ldots l\}$. Левой и правой медианами кластера S_ξ^1 будем считать соответственно значения векторов $\boldsymbol{x}_{\frac{l}{2}}$ и $\boldsymbol{x}_{\frac{l+2}{2}}$, если l четно, или значение вектора $\boldsymbol{x}_{\frac{l+1}{2}}$, если l нечетно.

$$\mathrm{med}_l(S_\xi^1)=\boldsymbol{x}_{\frac{l}{2}},\ \mathrm{med}_r(S_\xi^1)=\boldsymbol{x}_{\frac{l+2}{2}}, \qquad \text{при } \mathrm{mod}_2(l)=0$$

$$\mathrm{med}_l(S_\xi^1)=\mathrm{med}_r(S_\xi^1)=\boldsymbol{x}_{\frac{l+1}{2}}, \qquad \text{при } \mathrm{mod}_2(l)=1$$

Для значений $\xi=2\ldots k\text{-}1$ рассматриваются кластеры $S_{\xi-1}^1, S_\xi^1, S_{\xi+1}^1$. Если расстояние (в заданной топологической метрике) между $\mathrm{med}_l(S_{\xi-1}^1)$ и $\mathrm{med}_r(S_\xi^1)$ больше, чем расстояние между $\mathrm{med}_r(S_\xi^1)$ и $\mathrm{med}_l(\boldsymbol{S}_{\xi+1}^1)$, то кластер S_ξ^1 присоединяется справа к кластеру $S_{\xi-1}^1$, в противном случае кластер S_ξ^1 присоединяется слева к кластеру $S_{\xi+1}^1$. Процесс присоединения кластера S_ξ^1 к правому или левому кластеру происходит только в том случае, если после присоединения размерность пространства, в котором существует данный кластер остается неизменной с точностью до следующего сингулярного числа w_{p+1}. Итеративный процесс присоединения кластеров считается завершенным, если никакой кластер S_ξ^1, $\xi = 1\ldots k$ нельзя присоединить к правому или левому соседнему кластеру.

Рассмотрим процесс кластеризации в пространстве R^n. Пусть имеется множество кластеров S^{p-1}, состоящее из упорядоченных подмножеств S_ξ^{p-1}. Тогда процесс присоединения кластера S_ξ^{p-1} к левому или правому соседнему кластеру происходит и завершается так же, как и для случая кластеризации в пространстве R^1. Размерность пространства, в котором находится кластер после его присоединения к левому или правому соседнему кластеру правому или не должна превышать p.

Программная реализация

Для решения поставленной задачи выполняем последовательность процедур. Процедуры 1-3 необходимы для генерации исходных данных задачи, процедуры 4 и 5 решают поставленную задачу.

1. Решаем систему обыкновенных дифференциальных уравнений Лоренца:

```
La=Evaluate[{x[t], y[t], z[t]} /.
       NDSolve[ {x'[t] == -3 (x[t] - y[t]),
              y'[t] == -x[t] z[t] + 26.5 x[t] - y[t],
              z'[t] == x[t] y[t] - z[t],
              x[0] == z[0] == 0, y[0] == 1},
              {x, y, z},
              {t, 0, 40},
              MaxSteps->3000 ]
];
```

2. Представляем полученную фазовую траекторию в виде матрицы:

```
tau = 0.01;
A=Table[First[La],{t,0,40,tau}];
```

3. Линейно отображаем матрицу в n-мерное пространство и накладываем шум.

```
n = Random[Integer,{4,10}];
C= RandomTable[{n},{n}];
B= MakeNoise[{n,Length[First[A]},sigma2];
A=FreeRotate[A,C] + B;
```

4. Находим фактическую размерность матрицы A:

```
{u,w,v}=SingularValues[A];
```

```
r=FindItemIndex[Min[w, w^2 > sigma2]]
r=3
```

5. Находим кластеры для $S^1 \ldots S^r$ и строим древовидный граф *T*:

```
T={};
For[i=1,i<=r,i++,
        T=Append[
                T,
                Clusters[A,i]
        ];
];
```

Результат работы программы находится в структурированном массиве *T*.

Полученные результаты

Проекция исходной фазовой траектории на плоскость представлена на рис.1:

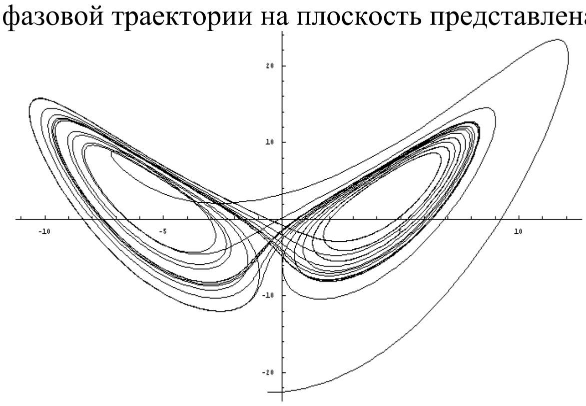

рис.1 Проекция исходной фазовой траектории на плоскость

Покажем, как выглядят кластеры, полученные в результате работы алгоритма. На рисунке 2 показаны некоторые кластеры данной фазовой траектории в пространстве R^2.

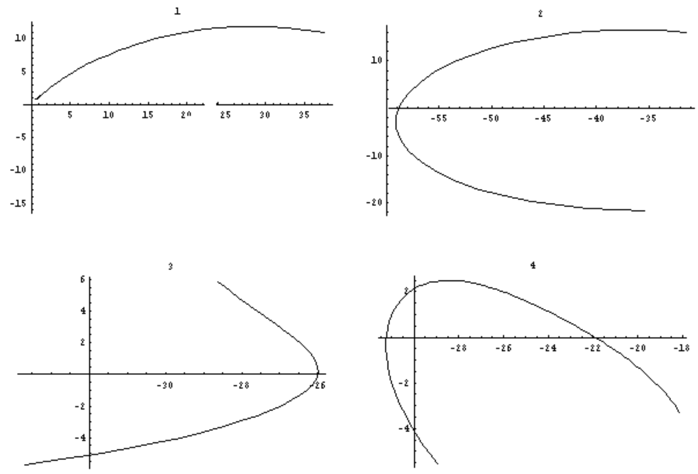

рис.2 Некоторые кластеры данной фазовой траектории

Для того, чтобы более наглядно показать взаимное расположение кластеров фазовой траектории группируем кластеры, находящиеся в одной плоскости с точностью до $\pm \sigma^2/2$. (на рис. 3 показана только одна группа).

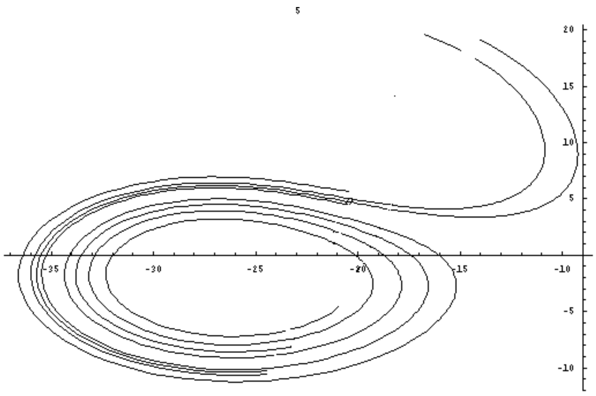

рис.3 Кластеры, лежащие в одной плоскости

Покажем древовидный граф T в графическом виде. Вершиной дерева является вся фазовая траектория, представленная в пространстве R^r размерности r. Следующий уровень состоит из узлов-кластеров в пространстве размерности r-1 и так далее до последнего уровня дерева, где фазовая траектория поделена на одномерные кластеры (см. пример дерева рис. 4).

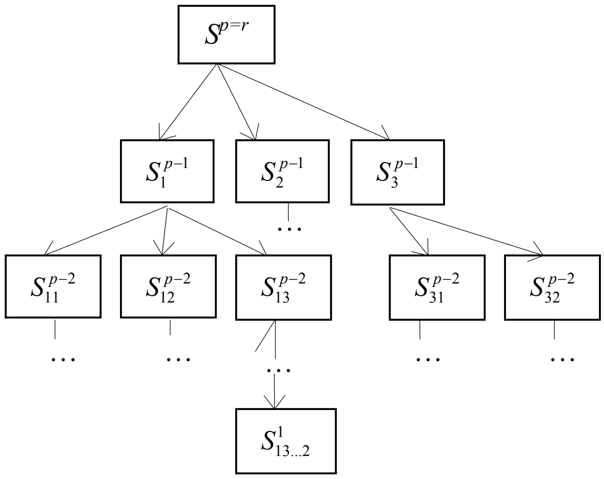

рис. 4 Пример древовидного графа

Заключение

Развит алгоритм, разделяющий данную фазовую траекторию на кластеры. Входными данными алгоритма является матрица, соответствующая данной фазовой траектории. Необязательными параметром могут быть размерность пространства, в котором должны находится кластеры и распределение шума, наложенного на фазовую траекторию. Результатом работы алгоритма является древовидный граф, в узлах которого находятся кластеры данной фазовой траектории. Алгоритм опробован на различных тестовых траекториях.

---

[1] Стрижов В.В., Шакин В.В. Программное обеспечение для исследования фазовых траекторий. МАТЕМАТИЧЕСКИЕ МЕТОДЫ РАСПОЗНАВАНИЯ ОБРАЗОВ (ММРО-9), Доклады IX всероссийской конференции. Научное издание – РАН, ВЦ, РФФИ, Москва, 1999. – С. 227-230.
[2] Классификация и снижение размерности /под ред. С.А. Айвазяна, М.: Финансы и статистика, 1989 – С. 332-419.