

# СОВМЕСТНОЕ ОЦЕНИВАНИЕ ПАРАМЕТРОВ И СОСТОЯНИЯ НЕЛИНЕЙНЫХ ДИНАМИЧЕСКИХ СИСТЕМ (ЧАСТЬ 1: ЛС МОДЕЛИ)

**О.Ю. Копысов**

*Декартов Научен Център*

Болгария, 9002, гр. Варна, ж.к. Чайка, бл.185, ап.13

E-mail: iainstitute@inbox.ru

**Ключевые слова:** модели линейной структуры (ЛСМодели), пространство параметров и реализаций ЛСМоделей, сингулярное разложение (SVD), вспомогательная ЛСМодель, динамические нелинейные системы.

**Аннотация:** Для некоторого класса нелинейных систем, названных моделями линейной структуры (ЛСМоделями), предлагается аналитический метод перехода к линейной вспомогательной модели, в которой вектор параметров расширяется начальными состояниями, и новая задача идентификации параметров и начальных состояний становится линейной.

## 1. Введение

Хорошо известные методы совместного оценивания состояния и параметров (квазилинеаризации, инвариантного погружения, расширенный фильтр Калмана и другие им подобные) расширяют вектор состояния системы, включая в него уравнения для параметров. Такая задача совместного оценивания параметров и состояния носит нелинейный характер даже для линейных систем управления.

Для некоторого класса нелинейных систем, названных моделями линейной структуры (ЛСМоделями), предлагается аналитический метод перехода к линейной вспомогательной модели, в которой вектор параметров расширяется начальными состояниями, и новая задача идентификации параметров и начальных состояний становится линейной.

## 2. Модели линейной структуры

Если физические объекты рассматривать как набор элементов (например: путь, скорость, ускорение) объединенных балансовыми отношениями, законами сохранения и другими, то математические модели таких объектов хорошо укладываются в алгебраическое понятие линейной зависимости элементов линейного пространства.

Однако имеется небольшое отличие. Если физические законы не зависят от выбора системы отсчета, то свойство линейной зависимости зависит от выбора начала координат.

Например, два вектора  $a = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$  и  $b = \begin{bmatrix} 5 \\ 1 \end{bmatrix}$  линейно независимы  $\begin{vmatrix} 1 & 5 \\ 3 & 1 \end{vmatrix} = -14$ . Однако с переносом начала координат в точку (3 2) получим  $\bar{a} = \begin{bmatrix} 1 \\ 3 \end{bmatrix} - \begin{bmatrix} 3 \\ 2 \end{bmatrix} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$ ,  $\bar{b} = \begin{bmatrix} 5 \\ 1 \end{bmatrix} - \begin{bmatrix} 3 \\ 2 \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$ , и тогда  $\begin{vmatrix} -2 & 2 \\ 1 & -1 \end{vmatrix} = 0$  и вектора линейно зависимы. Смена начала координат не меняет расположения точек в линейном пространстве, но меняет длины векторов и углы между ними. В дальнейшем постараемся учесть эти расхождения.

Класс моделей линейной структуры [1, 2] запишем формулой вида:

$$(1) \quad \sum_{i=1}^n \alpha_i f_i(t) = \alpha_0, \text{ для любых } t \text{ принадлежащих множеству } \mathbf{t},$$

где  $\alpha_0, \alpha_1, \dots, \alpha_n$  – параметры, элементы поля  $A$ ;  ${}^e f_i(t) = \{f_i(t) : t \in \mathbf{t}\}$  – элемент линейного пространства  ${}_A F$  над полем  $A$ .

С другой стороны, уравнение (1) в системе координат  $f_1, \dots, f_n$  является уравнением гиперплоскости, содержащей все точки  ${}^r f(t) = \{f_1(t) \dots f_n(t)\}$ , которые называются реализациями ЛСМодели (1), а гиперплоскость называется модельной плоскостью. Очень важно, что модельная плоскость содержит все возможные реализации во всех возможных экспериментах, проводимых с ЛСМоделью с заданными параметрами. Вектор параметров  $\alpha = [\alpha_1 \dots \alpha_n]'$  является нормалью к модельной плоскости. Совокупность всех векторов  $\alpha$  называется общим решением задачи идентификации и является подпространством пространства реализаций ортогональным образующему подпространству модельной плоскости. Таким образом, рассмотренное пространство реализаций содержит и вектора реализаций и вектора параметров, поэтому его следует называть пространством параметров и реализаций (ППР), в котором оси имеют двойное название  $f_1 \setminus \alpha_1, \dots, f_n \setminus \alpha_n$ .

Рассмотрим в качестве примера ЛСМодель вида:

$$\dot{x}(t) + 3x(t) - 2x^3(t) = -1, \quad {}^r f(t) = [\dot{x}(t) \ x(t) \ x^3(t)], \quad \alpha = [1 \ 3 \ -2], \quad \alpha_0 = -1.$$

Эта модель имеет три особые точки: устойчивый узел  $((1 - \sqrt[3]{3})/2, 0) \cong (-0.366, 0)$  обозначенный на рисунках кружочком и два неустойчивых  $((1 + \sqrt[3]{3})/2, 0) \cong (1.366, 0)$  и  $(-1, 0)$  обозначенных ромбиками. На рисунке 1 даны решения  $x(t)$  для пяти начальных условий: 1.36601, 1.36599, -0.366, -0.997, -1.003. На рисунке 2 фазовый портрет – кубическая функция  $\dot{x}(t) = 2x^3(t) - 3x(t) - 1$ . На рисунке 3 в ППР даны модельная плоскость (синяя решетка) и все возможные траектории (жирная синяя линия) на ней, а также общее решение задачи идентификации – прямая (красная) проходящая по вектору параметров  $\alpha$  (красная стрелка), перпендикулярно модельной плоскости.

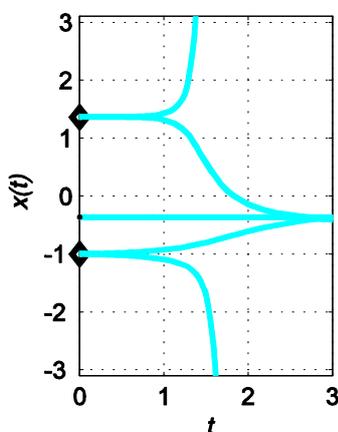


Рис. 1.

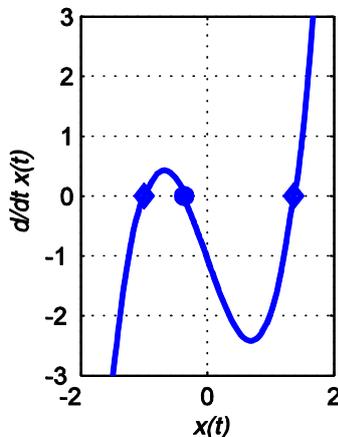


Рис. 2.

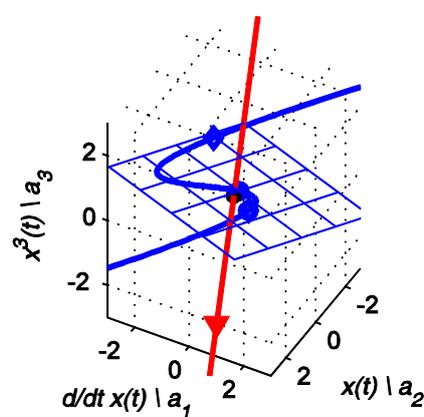


Рис. 3.

### 3. Разделение ППР (точные измерения)

Самым лучшим методом для разделения ППР на два подпространства является сингулярное разложение матрицы экспериментальных данных, однако метод работает только для уравнений вида  $F(\mathbf{t})\boldsymbol{\alpha}=\mathbf{0}$ .

В матричной нотации уравнение (1) имеет вид:

$$(2) \quad F(\mathbf{t})\boldsymbol{\alpha} = \alpha_0(\mathbf{t}),$$

где матрицу  $F(\mathbf{t})$  будем называть матрицей экспериментальных данных (ЭД). Элемент модели  ${}^e\mathbf{f}_i(\mathbf{t})$  является  $i$ -ым столбцом матрицы  $F(\mathbf{t})$ , а реализация  ${}^r\mathbf{f}(\mathbf{t})$  –  $t$ -той строкой матрицы,  $\alpha_0(\mathbf{t})=\{\alpha_0: t \in \mathbf{t}\}$  – элемент-константа пространства  ${}_A\mathbb{F}$ .

Чтобы перейти к системе с  $\alpha_0=\mathbf{0}$  необходимо знать любую точку  ${}^r\mathbf{f}_0=\{f_{01} \dots f_{0n}\}$ , принадлежащую модельной плоскости, и вычесть ее из каждой реализации  ${}^r\mathbf{f}(\mathbf{t})-{}^r\mathbf{f}_0$ , что будет означать переход к новой системе координат с началом в точке  ${}^r\mathbf{f}_0$ .

Например, это может быть особая точка или средняя точка нескольких реализаций, что предпочтительнее, так как параметры нам неизвестны. Очевидно, что

$$\alpha_1 \sum_{i=1}^m f_1(t_i) + \dots + \alpha_n \sum_{i=1}^m f_n(t_i) = \sum_{i=1}^m \alpha_0,$$

отсюда

$$\alpha_1 \frac{1}{m} \sum_{i=1}^m f_1(t_i) + \dots + \alpha_n \frac{1}{m} \sum_{i=1}^m f_n(t_i) = \alpha_0,$$

другими словами, средняя точка  ${}^r\mathbf{f}_0 = \left[ \frac{1}{m} \sum_{i=1}^m f_1(t_i) \dots \frac{1}{m} \sum_{i=1}^m f_n(t_i) \right]$  принадлежит модельной плоскости.

Обозначим  $F$  приведенную матрицу ЭД  $F=F(\mathbf{t})-{}^r\mathbf{f}_0(\mathbf{t})$  с дискретными измерениями  $\mathbf{t}=\{t_1 \dots t_m\}$ , и пусть ее сингулярное разложение имеет вид:

$$(3) \quad F = USV^*; \text{ где } U=[\mathbf{u}_1 \dots \mathbf{u}_m], S=\text{diag}(s_1 \dots s_n), V=[\mathbf{v}_1 \dots \mathbf{v}_n];$$

${}^r\mathbf{f}_0(\mathbf{t})$  – матрица-константа, состоящая из строк  ${}^r\mathbf{f}_0=\{f_{01} \dots f_{0n}\}$ ;

$U$  и  $V$  ортогональные матрицы,  $*$  – операция комплексного сопряжения и транспонирования;  $S$  диагональная матрица, содержащая упорядоченные сингулярные числа  $s_1 \geq s_2 \geq \dots \geq s_n$  на диагонали.

Из (3) следует, что

$$F = \sum_{i=1}^n \mathbf{u}_i s_i \mathbf{v}_i^* = \sum_{i=1}^n F_i, \quad \text{где } F_i = \mathbf{u}_i s_i \mathbf{v}_i^* = F \mathbf{v}_i \mathbf{v}_i^*.$$

Столбцы  $\mathbf{u}_i$  матрицы  $U$  и  $\mathbf{v}_i$  матрицы  $V$  называются левыми и правыми сингулярными векторами матрицы  $F$ , при этом  $\mathbf{u}_i^* F = s_i \mathbf{v}_i^*$  и  $F \mathbf{v}_i = s_i \mathbf{u}_i$ .

Если трактовать столбцы  $\mathbf{u}_i$  как ортонормированный базис в пространстве элементов ЛСМодели, то матрица  $U$  является матрицей перехода из базиса в базис. Аналогично, если трактовать столбцы  $\mathbf{v}_i$  как базис в пространстве параметров и реализаций, то матрица  $V$  будет матрицей перехода. Хорошо известно, что переходы в ортонормированные базисы не меняют взаимного расположения точек в линейном пространстве, а лишь изменяют их координаты. Матрица  $F$  в базисе  $\{\mathbf{u}_i\}$ , назовем ее  $F^u$ , а в базисе  $\{\mathbf{v}_i\}$  –  $F^v$ , как следует из (3), будет иметь вид:

$$F^u = U^* F = SV^* = \begin{bmatrix} s_1 \mathbf{v}_1^* \\ \vdots \\ s_n \mathbf{v}_n^* \end{bmatrix}, \quad F^v = FV = US = [s_1 \mathbf{u}_1 \dots s_n \mathbf{u}_n].$$

Пусть  $s_1 \geq s_2 \geq \dots \geq s_r > 0$ , а  $s_{r+1} = s_{r+2} = \dots = s_n = 0$ , тогда

$$s_i \mathbf{v}_i^* = 0 \text{ и } s_i \mathbf{u}_i = 0, \quad F \mathbf{v}_i = s_i \mathbf{u}_i = 0 \text{ и } F_i = \mathbf{u}_i s_i \mathbf{v}_i^* = 0, \quad i=r+1, \dots, n.$$

Отсюда также следует, что любая линейная комбинация векторов  $\mathbf{v}_i$  ( $i=r+1, \dots, n$ ) является решением уравнения  $F\boldsymbol{\alpha}=\mathbf{0}$ , так как  $F \sum_{i=r+1}^n b_i \mathbf{v}_i = \sum_{i=r+1}^n b_i F \mathbf{v}_i = \sum_{i=r+1}^n b_i s_i \mathbf{u}_i = 0$ . Кроме того следует, что ранг  $F$  равен  $r$  и

$$F = F_r = \sum_{i=1}^r \mathbf{u}_i s_i \mathbf{v}_i^* = U_r S_r V_r^*, \quad \text{где } U=[\mathbf{u}_1 \dots \mathbf{u}_r], S=\text{diag}(s_1 \dots s_r), V=[\mathbf{v}_1 \dots \mathbf{v}_r].$$

Другими словами, линейная оболочка  $\text{span}\{\mathbf{v}_1 \dots \mathbf{v}_r\}$  является модельной плоскостью в ППР, а ортогональная ей линейная оболочка  $\text{span}\{\mathbf{v}_{r+1} \dots \mathbf{v}_n\}$  является общим решением задачи идентификации.

Большим достоинством такого метода разделения ППР на два ортогональных подпространства является наличие великолепной процедуры вычисления сингулярного разложения любой матрицы. В пакете МАТЛАБ достаточно написать:  
 $\gg [U, S, V] = \text{svd}(F)$ .

#### 4. Разделение ППР (неточные измерения)

Очевидно, что практически невозможно получить точные нули в сингулярных числах, особенно в условиях неточных измерений и вычислений. Другими словами, матрицы экспериментальных данных почти всегда имеют полный ранг  $n$ . Однако, чрезвычайно важно для приближенных вычислений, что возмущения элементов матрицы приводят к меньшим или равным возмущениям ее сингулярных чисел. Пусть  $\tilde{F} = F + \Delta F$ , тогда  $|\tilde{s}_i - s_i| \leq \|\Delta F\|_2$ . А вот сингулярные вектора  $\mathbf{v}_{r+1} \dots \mathbf{v}_n$  неустойчивы к вычислениям, так как в случае неоднозначного решения  $r+1 < n$ , даже если само неоднозначное решение устойчиво, ортогональных базисов этого решения бесконечно много.

В 1901 году Карл Пирсон [3] решил задачу наилучшего прилегания системы точек к плоскости в смысле минимизации суммы квадратов длин перпендикуляров из точек на искомую плоскость и при этом он доказал принципиально важную теорему:

*Прямая линия (или плоскость) оптимальной подгонки для системы точек в пространстве любого порядка проходит через центроид (среднюю точку) системы.*

Для нас это означает, что центроид (средняя точка всех реализаций) принадлежит оптимальной модельной плоскости, как и при точных измерениях, он принадлежал модельной плоскости. Замечу, другой плоскости и другой центроид. Это решает проблему перевода начала координат на искомую оптимальную модельную плоскость.

Теорема Eckart-Young-Mirsky [4] в наших обозначениях утверждает:

*Пусть сингулярное разложение матрицы  $F$  имеет вид:  $F = \sum_{i=1}^n \mathbf{u}_i s_i \mathbf{v}_i'$  =  $\sum_{i=1}^n F_i$ .*

*Если  $r < n$  и  $\hat{F} = F_r = \sum_{i=1}^r \mathbf{u}_i s_i \mathbf{v}_i' = \sum_{i=1}^r F_i$ , то*

$$\min_{\text{rank}(D)=r} \|F - D\|_2 = \|F - \hat{F}\|_2 = s_{r+1},$$

$$\min_{\text{rank}(D)=r} \|F - D\|_F = \|F - \hat{F}\|_F = \sqrt{\sum_{i=r+1}^n s_i^2}.$$

Теорема говорит о том, что лучшая аппроксимация матрицы  $F$  ранга  $n$  матрицей  $\hat{F}$  заданного ранга  $r < n$  состоит в отсечении от суммы сингулярного разложения матрицы  $F = \sum_{i=1}^n \mathbf{u}_i s_i \mathbf{v}_i'$  последних  $n-r$  членов. Другими словами,  $F = \hat{F} + \Delta F$ ,  $\Delta F = \sum_{i=r+1}^n \mathbf{u}_i s_i \mathbf{v}_i'$ . При этом выбор любого ранга  $r$  ведет к оптимальному решению.

Теперь задача состоит в том, чтобы выбрать ранг  $r$ , который приведет нас еще и к правильному разделению ППР.

В [5] опубликована следующая теорема:

*Проблема Тотальных Наименьших Квадратов (Total Least Squares) состоит в поиске возмущения  $\Delta F$  имеющего минимальную норму Фробениуса  $\|\Delta F\|_F$ , так чтобы  $\text{rank}(F + \Delta F)$  был меньше ранга матрицы  $F$  и существовал вектор  $\alpha$ , такой что  $(F + \Delta F)\alpha = \mathbf{0}$ . Пусть  $F = USV'$  и если  $s_1 \geq \dots \geq s_r > s_{r+1} = \dots = s_n > 0$  то*

$$s_n = \min_{\text{rank}(F + \Delta F) < n} \|\Delta F\|_F.$$

*Кроме того, минимум достигается при  $\Delta F = -F\alpha\alpha'$ , где  $\alpha$  любой нормированный вектор из оболочки  $\text{span}\{\mathbf{v}_{r+1} \dots \mathbf{v}_n\}$ .*

Отсюда следует, что если  $s_{r+1}=s_n$ , а  $s_r > s_{r+1}$ , то оптимальная модельная плоскость имеет базис  $\mathbf{v}_1 \dots \mathbf{v}_r$ , а общее решение задачи идентификации базис  $\mathbf{v}_{r+1} \dots \mathbf{v}_n$ .

Надо учитывать, что на практике мы не можем рассчитывать на получение точного равенства последних сингулярных чисел, поэтому равенства  $s_{r+1}=\dots=s_n$  надо понимать приближенно.

Вспомним, что мы не располагаем полным вектором состояния и не можем сформировать матрицу экспериментальных данных  $F$ . Во второй части доклада, мы будем строить вспомогательную ЛСМодель, а сейчас для нашего примера прибегнем к численному дифференцированию по формулам центральных разностных производных:

$$\dot{x}(t_i) := (x(t_{i+1}) - x(t_{i-1})) / (t_{i+1} - t_{i-1}).$$

В расчетах, к точным значениям  $x(t_i)$  ( $x(0)=1.36599$ ) прибавлялись значения датчика шума нормального распределения с нулевым мат. ожиданием и заданной дисперсией и затем вычислялись  $\tilde{x}$  и  $\tilde{x}^3$ . По зашумленным данным: вычислялся центроид  $\mathbf{c}=[c_1 \ c_2 \ c_3]$ ,  $c_1 := \frac{1}{m} \sum_{i=1}^m \tilde{x}(t_i)$ ,  $c_2 := \frac{1}{m} \sum_{i=1}^m \tilde{x}^2(t_i)$ ,  $c_3 := \frac{1}{m} \sum_{i=1}^m \tilde{x}^3(t_i)$ ; формировалась матрица экспериментальных данных  $F := [\tilde{x}-c_1 \ \tilde{x}-c_2 \ \tilde{x}^3-c_3]$ ; вычислялось ее сингулярное разложение  $F=USV'$  и оценка  $\hat{F} := \sum_{i=1}^2 \mathbf{u}_i s_i \mathbf{v}_i' = [\hat{\mathbf{f}}_1 \ \hat{\mathbf{f}}_2 \ \hat{\mathbf{f}}_3]$ , по которой вычислялись оценки состояния:  $\hat{\mathbf{x}} := \hat{\mathbf{f}}_1 + c_1$ ,  $\hat{\mathbf{x}} := \hat{\mathbf{f}}_2 + c_2$ ,  $\hat{\mathbf{x}}^3 := \hat{\mathbf{f}}_3 + c_3$ .

Оценка параметров вычислялась как  $\hat{\alpha} := \mathbf{v}_3 / v_{13}$ , таким образом, во всех экспериментах  $\hat{\alpha}_1=1$ . В таблице даны сингулярные числа и оценки параметров для различных дисперсий шума. На рисунках – данные для дисперсии шума равной 0.02.

Дисперсия шума	$s_1$	$s_2$	$s_3$	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\alpha}_3$
<b>0</b>	12.740	5.8413	$2 \cdot 10^{-15}$	1	3	-2
<b>0.01</b>	7.5614	2.7368	0.20701	1	3.0643	-2.0400
<b>0.02</b>	7.5321	2.8626	0.32479	1	3.2683	-2.1180
<b>0.03</b>	7.4804	3.1899	0.42434	1	4.1104	-2.5348

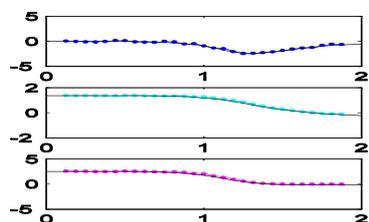


Рис. 4. Элементы ЛСМодели

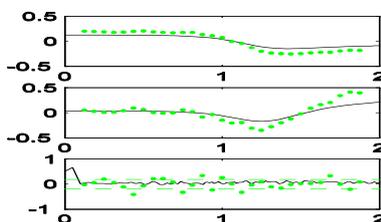


Рис. 5.  $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$

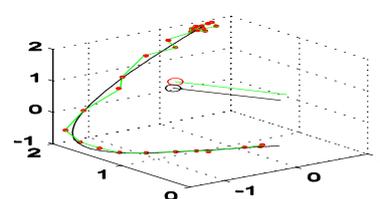


Рис. 6. ППР

На всех рисунках: черная линия – теоретические данные; точки – данные искаженные шумом; цветные линии – полученные оценки. На рис. 4 сверху вниз:  $\dot{x}(t)$ ,  $x(t)$ ,  $x^3(t)$ . На рис. 5 сверху вниз:  $\mathbf{u}_1$ ,  $\mathbf{u}_2$ ,  $\mathbf{u}_3$ . На рис. 6: траектории – линии, реализации – красные точки, центроиды – кружочки и параметры – прямые выходящие из центроидов.

## Список литературы

1. Копысов О.Ю. Elements of the Model Identification // IFAC-PapersOnLine, 2017. Vol. 50, No. 1. P. 2260-2265.
2. Копысов О.Ю. Идентификация посредством моделей линейной структуры .2-ое издание. Декартов Научен Центр, Варна, 2013. 425с. ISBN 978-954-92807-4-6.
3. Pearson K. On lines and planes of closed fit to system of points in space // Philosophical Magazine. Series 2. P. 559-572.
4. Van Huffel S., Vandewalle J. The Total Least-Squares Problem: Computational Aspects and Analysis. Philadelphia/PA, SIAM, 1991. 300 p.
5. Golub G., Van Loan C. An analysis of the Total Least-Squares problem // SIAM Journal on Numerical Analysis. 1980. Vol. 17, No. 6. P. 883-893.