

Министерство образования Республики Беларусь  
Учреждение образования  
Белорусский государственный университет  
информатики и радиоэлектроники

УДК 519.233.22

Будный  
Роман Игоревич

Идентификация стохастических систем

**АВТОРЕФЕРАТ**

на соискание степени магистра технических наук  
по специальности 1-40 80 02  
«Системный анализ, управление и обработка информации»

---

Научный руководитель  
Муха Владимир Степанович  
доктор технических наук, профессор

---

Минск 2018

## КРАТКОЕ ВВЕДЕНИЕ

Многие проблемы, решаемые в различных отраслях науки и техники, можно свести к задаче идентификации систем. Диагностика, управление, автоматизация принятия решений, распознавание образов — лишь некоторые практические цели решения данной задачи. Получение новых научных результатов в этой области приводит к развитию близких прикладных задач, таких как детектирование, прогнозирование и обучение.

Существующие исследования методов идентификации систем нельзя считать достаточно завершёнными: отсутствует обоснование используемых критериев, не приводится сравнительный численный анализ различных алгоритмов идентификации, отсутствуют рекомендации по применению того или иного алгоритма. Представляется целесообразной попытка независимого решения этих задач.

## ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

**Связь работы с научными исследованиями университета.** Диссертационная работа была выполнена в рамках госбюджетной научно-исследовательской работы ГБ №16-2016 кафедры информационных технологий автоматизированных систем БГУИР «Модели и методы оптимизации решений в современных информационно-аналитических и управляющих системах» на 2016–2020 годы.

**Цель и задачи исследования.** Предметом исследования работы является идентификация стохастических систем.

Целью работы является формирование рекомендаций по предпочтительному использованию некоторых методов аппроксимации статистических данных для идентификации стохастических систем на основе результатов численного сравнительного анализа их точности.

В задачи работы входят выполнение численного сравнительного анализа точности методов идентификации стохастических систем второго типа, описание условий предпочтительного использования сравниваемых методов для оценивания параметров систем.

**Новизна полученных результатов.** В работе получены следующие новые научные результаты:

- 1) предложена классификация стохастических объектов для целей их идентификации и сравнения;
- 2) разработан алгоритм метода симметричной аппроксимации;
- 3) получена эмпирическая зависимость для выбора более точного метода оценивания параметров линейных стохастических систем второго типа;
- 4) для решения задачи идентификации адаптированы нелинейный метод наименьших квадратов и метод рядов Тейлора;
- 5) описаны условия предпочтительного использования данных методов для идентификации нелинейных стохастических второго типа.

**Положения, выносимые на защиту.**

- 1) Классификация стохастических объектов, выделяющая полустохастические объекты и стохастические объекты первого и второго типов, позволяющая выявить проблемный для исследования класс стохастических объектов второго типа.
- 2) Алгоритм метода симметричной аппроксимации, позволяющий применять данный метод для идентификации систем с помощью ЭВМ.
- 3) Эмпирическая зависимость, предназначенная для выбора метода оценивания параметров линейных стохастических систем второго типа.
- 4) Условия предпочтительного использования нелинейного метода наименьших квадратов и метода рядов Тейлора для идентификации нелинейных стохастических второго типа.

**Апробация результатов диссертации.** Результаты исследования были представлены на следующих конференциях:

- 1) «51-ая научная конференции аспирантов, магистрантов и студентов» учреждения образования «Белорусский государственный университет информатики и радиоэлектроники», БГУИР, Минск, Беларусь, 13–17 апреля 2015 года.
- 2) «53-ая научная конференции аспирантов, магистрантов и студентов» учреждения образования «Белорусский государственный университет информатики и радиоэлектроники», БГУИР, Минск, Беларусь, 2–6 мая 2017 года.
- 3) Международная научная конференция «Информационные технологии и системы», БГУИР, Минск, Беларусь, 25 октября 2017 года.

**Опубликованность результатов исследования.** Результаты исследования были опубликованы в виде материалов указанных выше трех конференций.

**Структура и объем диссертации.** Диссертация 60 с., 18 рисунков (15 с.), 2 таблицы (2 с.), 2 приложения (10 с.), 34 источника.

Работа состоит из трех глав. В первой главе приведен краткий исторический обзор развития теории идентификации систем, а также математическая модель задачи идентификации в общем виде. Предложена классификация стохастических систем для целей их идентификации и сравнения. Рассмотрены классический и симметричный критерии идентификации систем. Выполнена постановка задачи исследования.

Вторая глава работы посвящена выявлению условий предпочтительного использования классического и симметричного критериев идентификации для оценивания параметров линейных стохастических объектов второго типа, а также прогнозирования наблюдений их выхода по наблюдениям входа. Приведены алгоритмы методов, основанных на указанных критериях идентификации. На основании результатов сравнительного численного анализа получена эмпирическая зависимость для выбора более точного метода оценивания параметров идентифицируемых систем.

В третьей главе рассмотрена задача идентификации нелинейных стохастических систем второго типа. Для решения данной задачи предложено использовать нелинейный метод наименьших квадратов и метод рядов Тейлора. Приведены алгоритмы данных методов. На основании результатов сравнительного численного анализа описаны условия предпочтительного использования методов для идентификации ряда нелинейных систем.

## КРАТКОЕ СОДЕРЖАНИЕ РАБОТЫ

В **первой главе** сформулирована задача идентификации стохастических систем как задача оценивания параметров и состояния системы по результатам наблюдений над входными и выходными переменными, полученными в условиях её функционирования:

$$\begin{aligned}\bar{\eta} &= \psi(\bar{\theta}, \bar{\xi}), \\ \bar{x} &= \bar{\xi} + \bar{\varepsilon}_x, \\ \bar{y} &= \bar{\eta} + \bar{\varepsilon}_y,\end{aligned}$$

где  $\bar{\xi}, \bar{\eta}$  — векторы фактических значений входа и выхода объекта,  
 $\bar{\theta}$  — вектор фактических значений параметров,  
 $\psi$  — векторная функция регрессии,  
 $\bar{x}, \bar{y}$  — векторы наблюдаемых значений входной и выходной переменной,  
 $\bar{\varepsilon}_x, \bar{\varepsilon}_y$  — векторы независимых ошибок наблюдений входа и выхода.

Предложена следующая классификация стохастических систем.

1) Полустохастический объект — объект с детерминированным входом и случайным выходом.

2) Стохастический объект первого типа — объект со случайным входом и выходом. Возможно наличие ошибок в измерениях входных и выходных переменных.

3) Стохастический объектом второго типа — детерминированный объект с ошибками в измерениях входных и выходных переменных.

Рассмотрены классический и симметричный критерии идентификации стохастических систем, минимизирующие суммы квадратов вертикальных и перпендикулярных расстояний от наблюдений входных и выходных переменных до аппроксимирующей прямой или плоскости.

Выявлено, что области предпочтительного использования классического и симметричного критериев для идентификации линейных стохастических систем второго типа определены недостаточно четко, а также обоснована целесообразность описания условий предпочтительного использования методов идентификации нелинейных стохастических систем. Исходя из этого определены задачи исследования.

**Вторая глава** посвящена определению условий предпочтительного использования классического и симметричного критериев идентификации для оценивания параметров стохастических объектов второго типа, а также прогнозирования наблюдений их выхода по наблюдениям входа. Разработаны алгоритмы классического метода наименьших квадратов и метода симметричной аппроксимации, основанные на данных критериях. Для сравнения точности методов использована модель линейной скалярной системы:

$$\begin{aligned}h &= \alpha + \beta\xi, \\x &= \xi + \varepsilon_x, \\y &= h + \varepsilon_y,\end{aligned}$$

где  $\xi, h$  — фактические значения входной и выходной переменной,

$\alpha, \beta$  — фактические значения параметров объекта,

$x, y$  — наблюдаемые значения входной и выходной переменной,

$\varepsilon_x, \varepsilon_y$  — независимые ошибки наблюдений входной и выходной переменной, распределенные по нормальному закону:  $\varepsilon_x = N(0, \sigma_{\varepsilon_x}), \varepsilon_y = N(0, \sigma_{\varepsilon_y})$ .

Фактические значения входа системы  $\xi_i$  выбирались из равномерного в  $[0, 10]$  распределения. Для получения каждой оценки  $\hat{\alpha}, \hat{\beta}$  использовались результаты ста наблюдений  $(x_i, y_i), i = \overline{1, n}, n = 100$ . Сравнение точности оценивания параметров модели производилось на основании разности средних Евклидовых расстояний между точными значениями параметров  $\alpha, \beta$  и

их оценками, полученными методом наименьших квадратов ( $\hat{\alpha}_{\text{МНК}}, \hat{\beta}_{\text{МНК}}$ ) и методом симметричной аппроксимации ( $\hat{\alpha}_{\text{МСА}}, \hat{\beta}_{\text{МСА}}$ ):

$$d = \frac{1}{k} \sum_{j=1}^k \left( \sqrt{(\hat{\alpha}_{\text{МНК}_j} - \alpha)^2 + (\hat{\beta}_{\text{МНК}_j} - \beta)^2} - \sqrt{(\hat{\alpha}_{\text{МСА}_j} - \alpha)^2 + (\hat{\beta}_{\text{МСА}_j} - \beta)^2} \right),$$

где  $k$  — число оценок. Расчеты величины  $d$  производились в узлах сетки значений  $\sigma_{\varepsilon_x}, \sigma_{\varepsilon_y}$  в прямоугольнике  $[0, 2] \times [0, 2]$  с шагом 0,1. В каждом узле сетки вычислялось сто оценок ( $k = 100$ ).

Моделирование показало, что метод симметричной аппроксимации оценивает параметры линейных систем с большим коэффициентом усиления  $\beta$  более точно, чем классический метод наименьших квадратов. Для принятия решения о том, какой метод использовать более предпочтительно, предложено использовать следующую эмпирическую зависимость: «Если условие

$$\sigma_{\varepsilon_y} > (0,7 + |\beta|)\sigma_{\varepsilon_x}$$

выполняется, то метод наименьших квадратов оценивает параметры линейной стохастической системы второго типа более точно, чем метод симметричной аппроксимации. В противном случае метод симметричной аппроксимации позволяет получить оценки более высокой точности, чем метод наименьших квадратов».

Для сравнения точности прогнозирования наблюдений выхода по наблюдениям входа использовалась разность средних Евклидовых расстояний между наблюдениями выхода модели и их оценками:

$$d = \frac{1}{k} \sum_{j=1}^k \left( \sqrt{\sum_{i=1}^n (\hat{\alpha}_{\text{МНК}_j} + \hat{\beta}_{\text{МНК}_j} x_{ij} - y_{ij})^2} - \sqrt{\sum_{i=1}^n (\hat{\alpha}_{\text{МСА}_j} + \hat{\beta}_{\text{МСА}_j} x_{ij} - y_{ij})^2} \right),$$

где  $k$  — число оценок.

На основании результатов исследования зависимости данной величины от значений коэффициента усиления модели и с. к. о. ошибок наблюдений рекомендовано использовать метод наименьших квадратов для прогнозирования наблюдений выхода систем по наблюдениям входа.

Предметом **третьей главы** является проблема идентификации нелинейных стохастических систем. Для решения задачи идентификации нелинейных стохастических систем второго типа предложено использовать нелинейный метод наименьших квадратов и метод рядов Тейлора. Приведены алгоритмы и

примеры использования программных реализаций данных методов. Для сравнения точности методов использована модель скалярной системы:

$$\begin{aligned}h &= \psi(\bar{\theta}, \xi), \\x &= \xi + \varepsilon_x, \\y &= h + \varepsilon_y,\end{aligned}$$

где  $\xi, h$  — фактические значения входной и выходной переменной,

$\psi$  — скалярно-векторная функция регрессии,

$\bar{\theta} = (\theta_1, \theta_2, \dots, \theta_m)$  — вектор фактических значений параметров объекта,

$x, y$  — наблюдаемые значения входной и выходной переменной,

$\varepsilon_x, \varepsilon_y$  — независимые ошибки измерений значений входной и выходной переменной, распределенные по нормальному закону:  $\varepsilon_x = N(0, \sigma_{\varepsilon_x}), \varepsilon_y = N(0, \sigma_{\varepsilon_y})$ . Данная модель была использована для генерации наблюдений входа и выхода системы, на основании которых были получены оценки её параметров методом наименьших квадратов и методом рядов Тейлора. Значения  $\xi_i$  выбирались из равномерного в  $[0, 10]$  распределения. Для получения каждой оценки  $\hat{\theta}$  использовались результаты ста наблюдений  $(x_i, y_i), i = \overline{1, n}, n = 100$ .

На примере систем с линейными, параболическими, синусоидальными, экспоненциальными и обратными функциями регрессии описана зависимость разности средних Евклидовых расстояний между точными значениями параметров модели и их оценками от с. к. о. ошибок наблюдений:

$$\begin{aligned}d(\sigma_{\varepsilon_x}, \sigma_{\varepsilon_y}) &= \frac{1}{k} \sum_{j=1}^k \sqrt{\sum_{i=1}^m (\hat{\theta}_{\text{МНК}_{ij}}(\sigma_{\varepsilon_x}, \sigma_{\varepsilon_y}) - \theta_{ij})^2} - \\ &- \frac{1}{k} \sum_{j=1}^k \sqrt{\sum_{i=1}^m (\hat{\theta}_{\text{МРТ}_{ij}}(\sigma_{\varepsilon_x}, \sigma_{\varepsilon_y}) - \theta_{ij})^2},\end{aligned}$$

где  $k$  — число оценок. Расчеты величины  $d$  производились в узлах сетки значений  $\sigma_{\varepsilon_x}, \sigma_{\varepsilon_y}$  в прямоугольнике  $[0, 2] \times [0, 2]$  с шагом 0,1. В каждом узле сетки вычислялось сто оценок ( $k = 100$ ).

Показано, что точность нелинейного метода наименьших квадратов существенным образом зависит от того, насколько «удачной» является используемая им опорная точка. В аналогичных условиях метод рядов Тейлора позволяет получать оценки параметров приемлемой точности без необходимости указания опорной точки.

В заключении перечислены новые научные результаты, полученные в работе.

## СПИСОК ОПУБЛИКОВАННЫХ РАБОТ

- [1— А.] Будный Р. И. Оценивание параметров нелинейной регрессионной зависимости / Р. И. Будный // Материалы 51-ой научной конференции аспирантов, магистрантов и студентов — Минск, 2015. — С. 55.
- [2— А.] Будный Р. И. Численный анализ методов линейной аппроксимации статистических данных / Р. И. Будный // Материалы 53-ой научной конференции аспирантов, магистрантов и студентов — Минск, 2017. — С. 68.
- [3— А.] Муха В. С. О линейной аппроксимации векторных статистических данных / В. С. Муха, Р. И. Будный // Информационные технологии и системы 2017 (ИТС 2017): материалы международной научной конференции, БГУИР, Минск, Беларусь, 25 октября 2017 г. — Минск, 2017. — С. 288–289.