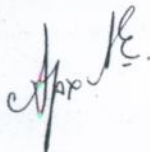


НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
"КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ"

На правах рукопису

УДК 681.5.015



АРХИПОВ Олександр Євгенійович

БУТСТРЕП-АНАЛІЗ ЯКОСТІ РОЗВ'ЯЗАННЯ ЗАДАЧ ІДЕНТИФІКАЦІЇ

05.13.03 - Системи і процеси управління

А в т о р е ф е р а т
дисертації на здобуття наукового ступеня
доктора технічних наук

Київ - 1995

81.5

ЛННБ України ім.В.Стефаника

Дисертація



00360432 (I)

Робота виконана на кафедрі
Національного технічного університету України "КПІ"

Науковий консультант: Доктор технічних наук, професор
Сильвестров А.М.

Офіційні опоненти: Доктор технічних наук, професор
Костяк В.І.

Доктор технічних наук, професор
Блохін Л.М.

Доктор технічних наук, старший науковий
співробітник Степашко В.С.

Провідна установа: Київський університет імені Т.Г. Шевченко

Захист відбудеться "___" _____ 1996 р. о ___ годині на
засіданні спеціалізованої ради Д 01.02.08 в Національному
технічному університеті України "Київський політехнічний
інститут" за адресою: 252056, Київ-56, пр. Перемоги 36, ауд.56-14

З дисертацією можна ознайомитися у бібліотеці НТУ України,
"КПІ".

Автореферат розісланий "___" _____ 1996 р.

Учений секретар
спеціалізованої ради
д.т.н., проф.

В.Д. Романенко

ЛННБ ім. В. Стефаника
АН України

АНОТАЦІЯ

Мета роботи - побудова методології дослідження якості ідентифікації, зокрема аналізу та кількісного оцінювання якості з подальшим використанням отриманих результатів для підвищення якості розв'язання задач ретроспективної ідентифікації лінійних та нелінійних моделей стаціонарних статичних і динамічних об'єктів за даними, що представлені вибірками малого та середнього об'єму з довільним характером зашумлення в умовах апріорної невизначенності властивостей вихідних даних.

Для досягнення поставленої мети в роботі вирішено такі задачі:

- виконано аналіз та дослідження ряду поширених методів та показників, що застосовуються для підтвердження і перевірки моделей (результатів ідентифікації);

- побудовано методологію оцінювання якості ідентифікації моделей довільної складності та структури при використанні довільної процедури структурно-параметричної ідентифікації;

- розроблено та досліджено систему показників якості параметричної та структурно-параметричної ідентифікації;

- запропоновано методи варіювання вихідних даних, що забезпечують випадково-природний характер варіації та статистичну подібність варіюваних даних;

- розроблено технологію оцінювання якості параметричної та структурно-параметричної ідентифікації;

- розроблено методи підвищення точності оцінок параметрів при використанні одного та групи методів параметричної ідентифікації;

- розроблено методи оцінювання та корегування зміщень в оцінках параметрів;

- запропоновано методи селекції найкращої моделі на множині моделей-претендентів;

- виконано розробку та дослідження методів згладжування часових рядів в умовах невизначенності характеристик вихідних даних, зокрема адаптивних методів згладжування;

- досліджено застосування згладжування часових рядів для підвищення якості ідентифікації моделей динамічних об'єктів.

Автор захищає:

- методологів дослідження якості ідентифікації моделей довірливої структури та складності;
- систему показників якості ідентифікації;
- методи варіювання вихідних даних та нові методи генерації реалізацій стаціонарних випадкових процесів і векторних випадкових величин;
- методи підвищення якості ідентифікації шляхом корегування змієнь оцінок параметрів та зменшення рівня розсіювання оцінок;
- методи вибору найкращої моделі з числа моделей-претендентів;
- технологію оцінювання якості розв'язання задач структурно-параметричної ідентифікації включно з принципами систематизації показників якості та методів варіювання даних;
- нові методи згладжування даних та способи їх використання з метов поліпшення якості ідентифікації;
- принципи побудови системи сертифікації згладжуваних фільтрів та отримані результати сертифікації певних типів мінімальних фільтрів;
- результати досліджень та приклади практичного застосування наведених у дисертації методів аналізу та підвищення якості ідентифікації.

ЗАГАЛЬНА ХАРАКТЕРИСТИКА РОБОТИ

Актуальність проблеми. У процедурі ідентифікації традиційно виділяють три етапи: вибір структури моделі, оцінювання її параметрів та перевірка (підтвердження) моделі. Найбільшим теоретичним та практичним наробком характеризується другий етап. Безперервно зростає і так вже величезна кількість публікацій, де пропонуються нові методи та алгоритми параметричної ідентифікації, досліджуються особливості їх застосування в окремих виключних випадках, розробляються нові та модифікуються відомі оптимізаційні схеми, деталізуються чи змінюються існуючі постановки задач параметричної ідентифікації і т.п. При цьому відходять на другий план або просто не обговорюються питання оптимізації структури моделі чи вибору найкращої моделі на множині моделей-претендентів, перевірки якісних показників побудованої моделі, підтвердження можливостей її практичного

застосування, тобто саме ті аспекти ідентифікації, які становлять зміст її першого та третього етапів.

Причин, що призвели до такої ситуації, декілька. Так здавалося б найбільш об'єктивний спосіб підтвердження моделі шляхом її перевірки в умовах практичного застосування має ряд вад. Орієнтація на конкретне застосування обумовлює розробку вузьких, індивідуальних методів перевірки споживчих якостей моделі, кожен з яких має свою індивідуальну систему показників якості. Пріоритет в цьому випадку належить емпірико-інтуїтивному підходу, бо якщо навіть в принципі й можливі формалізація і строго математичне вирішення задачі підтвердження моделі, трудомістність цієї розробки звичайно набагато вища за цінність отриманого результату. Крім того, цей результат буде дещо несвоєчасним, враховуючи, що практичну перевірку моделі від моменту її ідентифікації відділяє досить значний часовий інтервал.

Зручні й більш поширені оперативні методи перевірки (підтвердження) моделей. Звичайно це аналітичні статистико-ймовірнісні методи, розроблені для моделей певного класу в умовах певної системи обмежень і найчастіше орієнтовані на випадок асимптотики. В задачах реальної ідентифікації, що розв'язуються за вибірками вихідних даних малого та середнього об'єму в умовах апріорної невизначеності щодо властивостей цих даних, зокрема шумів вимірювання, застосування таких методів та отримані з їх допомогою висновки мають формально-поверховий характер. Додамо, що на етапі вибору структури моделі звичайно потрібне порівняльне оцінювання якості моделей різних класів, структур та складності, що виходить за межі застосування вищезгаданих методів.

Крім того, звичайно вважається, що той чи інший метод структурно-параметричної ідентифікації застосовано виходом із знання реальних умов та особливостей конкретної задачі ідентифікації, тобто орієнтовано на оптимальний результат. Фактична ж інформація, яку має у своєму розпорядженні дослідник для визначення способу ідентифікації, носить дуже приблизний, недостатній характер, тому обов'язковою є апостеріорна перевірка та підтвердження результатів ідентифікації, загальна оцінка якості ідентифікації в цілому.

Тому актуальна розробка загального підходу, що забезпечує оперативний аналіз результатів, отриманих при розв'язанні задачі ідентифікації, зокрема, дає змогу дослідити якість ідентифікацій моделей довільної структури й складності, виконати їх порівняльний аналіз та селекцію кращої з них за умов, звичайних для практичної ідентифікації. Важливим і актуальним є ще один аспект дослідження результатів ідентифікації, а саме: використання отриманих висновків для поліпшення якості ідентифікації. Відомий приклад цього - вже згадувана вище селекція моделі за результатами аналізу якості кожної з моделей-претендентів. Значно менш вивчені питання оптимізації вибору методу параметричної ідентифікації, уточнення оцінки параметра на множині його оцінок, кожна з котрих обчислена окремим методом параметричної ідентифікації і т.п. До останнього часу цим задачам приділялася незначна увага, але лавиноподібне збільшення інформації, що стосується методів та алгоритмів оцінювання, робить питання розв'язання цих задач дуже актуальним.

Методи дослідження. Для розв'язання сформульованих задач в роботі використано методи теорії ідентифікації, математичного та імітаційного моделювання, теорії ймовірностей та математичної статистики.

Наукова новизна. Автором одержано та викладено в дисертаційній роботі такі наукові результати:

- розроблено методологію оцінювання якості ідентифікації моделей довільної складності та структури;

- побудовано систему кількісних показників якості ідентифікації, до якої входять:

- а) показники якості параметричної ідентифікації, що базуються на обчисленні рівня стійкості оцінок параметрів моделі;

- б) комбіновані показники якості структурно-параметричної ідентифікації, в яких враховано як стійкісні, так і апроксимаційні характеристики рішення задачі ідентифікації;

- розроблено методи моделювання реалізації стаціонарних випадкових процесів та векторних величин при непараметричній формі опису їх характеристик (законів розподілу, кореляційних чи

коваріаційних залежностей);

- запропоновано методи корегування зміщень в оцінках, обчислених за даними, що містять похибки в незалежних змінних;

- розроблено метод підвищення якості оцінювання параметрів за результатами обробки групи оцінок кожного параметра, отриманих застосуванням ряду окремих методів параметричної ідентифікації;

- розроблено методи селекції найкращої моделі на множині моделей-претендентів довільної складності та структури за комбінованими показниками якості ідентифікації;

- отримано метод адаптивного згладжування даних, у якому реалізується адаптація структури та параметрів лінійного фільтру до локальних характеристик послідовності, що згладжується;

- розроблено метод рекурентного згладжування даних лінійним фільтром з несиметричною двобічно експоненційно спадаючою ваговою функцією;

- побудовано систему сертифікації згладжувачих фільтрів, основа якої - розроблені для кожного типу фільтра моделі, що узгоджують якість згладжування з впливовими на неї факторами;

- розроблено технологію дослідження задачі ідентифікації, що включає:

- а) формалізацію процедури дослідження, поділ її на основні етапи та операції;

- б) систематизації на множині кваліфікаційних ознак та їх станів показників якості ідентифікації та засобів варіювання даними.

Практична цінність. Один з головних напрямків практичного застосування результатів дисертаційної роботи - аналіз якості моделей, побудованих за даними, отриманими в процесі натурних випробувань динамічних об'єктів, особливість яких - значна вартість, велике енергетично-паливне споживання, обмеження на час та кількість експериментів, що іноді мають принципово одиничний характер. В цих умовах об'єктивне дослідження та аналіз якості ідентифікації моделей дозволяють значно підвищити ефективність наслідків моделювання, рівень довіри до результатів, отриманих з допомогою моделей, що врешті-решт призводить до скорочення загального часу натурних випробувань, зменшення фінансових та паливно-енергетичних витрат на їх забезпечення.

Крім того слід відзначити універсальний характер розробленої

в роботі методології аналізу якості ідентифікації, який забезпечується незалежністю аналізу якості від кінцевої мети розробки моделі, відмовою від аналітичних методів дослідження якості, пов'язаних з необхідністю виконання складних та громіздких математичних перетворень (що не гарантують вивідність остаточного результату), та заміною цих методів чисельним моделюванням. Наслідком такої універсальності запропонованої методології є її придатність для дослідження якості ідентифікації моделей практично довільної структури та складності за умов мінімальної апріорної інформації про характеристики даних, зокрема про статистичні властивості помилок вимірювань.

Виконана в роботі розробка технології дослідження якості ідентифікації включно з класифікацією показників якості та методів варіювання даних дозволяє значно спростити реалізацію процедури дослідження, що розширює коло потенційних користувачів та створює принципові можливості для автоматизації цієї процедури.

Реалізація результатів роботи. Результати дисертаційної роботи практично в повному обсязі використані у науково-дослідних розробках, що виконувались в межах найважливішої тематики відповідно до напрямку 08. Комплексної науково-технічної програми "Політ" (Мінавіапром СРСР) та за розпорядженням Президії НАН України на ряді підприємств: НВО "Молнія" (м. Москва), Таганрозький механічний завод ім. Г. Димитрова (м. Таганрог), Центральний НДІ автоматики та гідравліки (м. Москва), АНТК "Антонов" (м. Київ).

Частково отримані результати використані у розробках, що виконувались в інституті надтвердих матеріалів НАН України, а також у обробці даних ряду медико-біологічних досліджень.

У Додатку до дисертації наведені документи про впровадження результатів виконаної роботи.

Матеріали дисертації були використані в учбовому процесі при постановці лекційних курсів "Основи статистичного аналізу", "Теорія цілеорієнтованих систем первинної обробки даних".

Апробація роботи. Основні положення дисертації доповідались та обговорювались на III та V Всесоюзних симпозіумах "Методи теорії ідентифікації в задачах вимірювальної техніки і метрології", м. Новосибірськ, 1982, 1989 р.р., II Республіканській конференції по підвищенню надійності та довговічності машин і споруд, м. Дніпропетровськ, 1985 р., 7-й Уральській

науково-технічній конференції, м. Устинов, 1986 р., Всесоюзній науково-технічній конференції "Статистичні методи в теорії передачі та перетворення інформаційних сигналів", м. Львів, 1988 р., III-й польсько-радянській науково-технічній конференції "Комплексна автоматизація промисловості", м. Вроцлав, 1988 р., IX Magyar Repulástudományi Napok Budapest, 1988, 1-st Mini Conference of Vehicle System Dynamics Identification and Anomalies, Budapest, 1988, науково-технічній конференції "Проблеми комплексної автоматизації функціональних випробувань в машинобудуванні", м. Казань, 1988 р., Всесоюзній науково-технічній нараді "Автоматизація проектування та виробництва в електромашинобудуванні" м. Суздаль, 1989 р., науково-технічній конференції "Проблеми розвитку АСУ та інформаційних послуг у нових умовах господарювання", м. Душанбе, 1989 р., IV-й науково-технічній конференції "Проблеми комплексної автоматизації", м. Київ, 1990 р., I-й та III-й науково-технічних конференціях країн СНД "Контроль та управління в технічних системах", м. Вінниця, 1992, 1995 р.р., I-й Український конференції з автоматичного керування "Автоматика-94", м. Київ, 1994 р., науково-технічній конференції "Техніка і фізика електронних систем та пристроїв", м. Суми, 1995 р., I-й Міжнародній науково-технічній конференції "Математичне моделювання в електротехніці й електроенергетиці", м. Львів, 1995 р.

Публікації. По темі дисертації опубліковано 48 робіт, з них 3 учбових посібника для студентів вузів.

Структура та обсяг роботи. Дисертація складається із вступу, 6 розділів, заключення, викладених на сторінках друкованого тексту, списку літератури з 104 найменувань, додатку на сторінках, рисунків та таблиць.

Вступ містить обґрунтування актуальності теми, формулювання мети роботи, перелік задач та питань, що розглянуто в дисертації, характеристику отриманих теоретичних та практичних результатів.

В першому розділі розглянуто ряд понять, що використовуються при розв'язанні задач вибору та перевірки моделей, конкретизовано зміст терміну "якість ідентифікації", обґрунтовано доцільність та дієвість дослідження якості ідентифікації в умовах практики для оперативної перевірки результатів ідентифікації, виконано порівняльний аналіз поширених методів оцінювання якості

ідентифікації, зроблено постановку та деталізацію задачі дослідження якості ідентифікації.

Другий розділ містить розробку методології аналізу та оцінювання якості ідентифікації, побудову системи кількісних показників якості, дослідження їх властивостей, можливості та області використання.

Третій розділ - це розробка процедури дослідження якості ідентифікації статичних об'єктів, зокрема, розробка способів варіювання складу даних, методології побудови точносних характеристик оцінок параметрів, уточнення оцінок за результатами використання групи методів параметричної ідентифікації.

Зміст четвертого розділу - дослідження якості ідентифікації динамічних об'єктів, особливості дослідження при представленні вихідних даних часовими рядами, зокрема, можливість усунення зміщення оцінок параметрів, згладжування часових рядів з метою підвищення якості ідентифікації в умовах апіорної невизначеності властивостей вихідних даних шляхом використання адаптивних згладжувачих фільтрів, ефекту "перегладжування", розробка методології сертифікації алгоритмів та методів згладжування.

В п'ятому розділі виконано формалізацію процедури дослідження якості ідентифікації, виділено та систематизовано задачі, що становлять зміст окремих етапів дослідження, формалізовано вибір складу та послідовності використання розроблених прийомів та методів аналізу і підвищення якості ідентифікації.

Шостий розділ містить приклади дослідження та аналізу якості розв'язання реальних задач ідентифікації, що підтверджують можливості розробленої методології дослідження якості, надійність запропонованої системи показників якості, дієвість методів підвищення якості ідентифікації.

ОСНОВНИЙ ЗМІСТ РОБОТИ

В роботі розглянуто деякі поширені підходи до розв'язання задачі підтвердження та перевірки моделей, побудованих внаслідок ідентифікації об'єктів за експериментально отриманими даними. Конкретизовано зміст понять, що становлять основи зазначених підходів - якість моделі, точність моделі,

якість і точність ідентифікації, можливості кількісного оцінювання цих характеристик в умовах розв'язання реальних задач для оперативної перевірки надійності та вірогідності результатів ідентифікації, їх придатності для практичного застосування. З цих позицій обґрунтовано доцільність та дієвість використання для перевірки та підтвердження моделей методів, що базуються на дослідженні якості ідентифікації. При цьому якість ідентифікації визначається як комплексна характеристика загального рівня успішності розв'язання задач ідентифікації, яка спирається на інформацію, що відома або може бути отримана на час закінчення процедури ідентифікації в рамках дослідження її результатів, зокрема шляхом аналізу вірності вибору структури моделі, адекватності моделі вихідним даним (апроксимаційні властивості моделі), точності та стійкості оцінок параметрів моделі.

Комплексність, складність поняття якість ідентифікації обумовлює необхідність використання для її оцінювання ряду показників, зміст котрих, форма та способи обчислення суттєво залежать від особливостей постановки конкретних задач ідентифікації, зокрема від обсягу наявної апріорної інформації про структуру моделі. Так, в разі апріорно відомої структури якість моделі цілком визначається рівнем стійкості процедури параметричної ідентифікації, характеристиками розсіювання оцінок параметрів. При відсутності будь-якої інформації про структуру моделі головним чинником якості ідентифікації стає рівень адекватності моделі вихідним даним.

У цій ситуації актуальності набуває питання побудови системи показників якості ідентифікації. З метою його вирішення виконано порівняльний аналіз ряду поширених кількісних показників якості. Показано, що переважна більшість тих з них, які застосовуються на практиці, являють показник наближення вихідних даних модельними. Типовий і водночас найбільш поширений приклад такого показника - емпіричний ризик I_n , що є обчислений на обмеженім об'ємі вихідних даних $(y_i, x_{i1}, x_{i2}, \dots)$, $i = \overline{1, n}$, $n < \infty$ оцінок середнього ризику $I = M\{\theta[z - m(x, A)]\}$, де $M(\cdot)$ - символ математичного сподівання, $\theta[\cdot]$ - цільова функція, $m(x, A)$ - математична модель з вектором параметрів $A = [a_0, a_1, \dots]^T$, $x = [x_1, x_2, \dots]$ - вектор незалежних (вхідних) змінних, z - результат виміру залежної змінної y , що містить випадкову помилку E : $z = y + e$, $M(E) = 0$,

$D(E) = \sigma_e^2$ - дисперсія помилки. Середній ризик I є універсальною і об'єктивною характеристикою якості моделі, однак емпіричний ризик втрачає ці властивості: значення I_n суттєво залежить від обсягу вихідних даних, характеру та рівня їх зашумлення, складності моделі. Останній фактор є дуже важливим, бо задача визначення структури апроксимативних моделей, селекції моделей на множині моделей-претендентів довільної структури є найбільш складною та поширеною. У цих випадках застосовуються зкореговані показники емпіричного ризику, в основі яких лежать відомі з літератури співвідношення на зразок

$$I \approx M(I_n) + \text{tr}(\nabla_A^2 I \text{ cov } \tilde{A}), \quad (1)$$

де $\nabla_A^2 I$ - матриця других похідних середнього ризику I (матриця Гессе), $\text{cov } \tilde{A}$ - асимптотична матриця коваріацій вектора помилок оцінок параметрів моделі. Призначення другого додатку в (1) - корегування зміщення оцінки I_n , обумовленого спільним впливом згаданих вище факторів. Безпосередньо використати (1) для визначення I не можна, бо звичайно невідомі елементи матриці Гессе та матриці коваріацій, немає змоги реально обчислити математичне сподівання $M(I_n)$. Тому замість елементів правої частини (1) використовуються їх досить довільно знайдені оцінки, наприклад, замість $M(I_n)$ - єдине обчислене значення I_n . Наслідок цього - велике розсіювання можливих значень зкорегованого емпіричного ризику, їх неконтрольовані зміщення, низька надійність та ефективність застосування такого показника.

Наведений висновок досить типовий і для інших показників якості, відмінних од середнього ризику. В зв'язку з цим можливі два напрямки побудови системи показників якості ідентифікації:

- удосконалення та поліпшення існуючих кількісних показників;
- розробка принципово нового підходу до аналізу якості ідентифікації і формування на його базі нових додаткових чи альтернативних показників якості.

Другий розділ дисертації містить виклад основних засад підходу до оцінювання якості ідентифікації рівнем стійкості результатів ідентифікації. Суть підходу розглянемо на прикладі параметричної ідентифікації об'єкту, структура моделі якого відома. Функціонування об'єкту описується системою змінних $Y, X = [X_1, X_2, \dots, X_p]$, де Y - залежна змінна. З точки зору причинно-наслідкових зв'язків між змінними об'єкт можна задати

співвідношенням $\varphi: X \rightarrow Y$, де φ - невідомий оператор, що задає відображення елементів вектора вхідних змінних X в значення вихідної скалярної величини Y . Бважасмо апіорно відомим, що в діапазоні припустимих значень змінних Y, X співвідношення $\varphi: X \rightarrow Y$ точно апроксимується моделлю $Y = m(A^*, X)$, зная з точністю до вектору параметрів A^* . Задача параметричної ідентифікації - оцінити елементи вектора A^* за експериментально отриманою інформацією, представленю матрицею вихідних даних:

$$(z_i, u_i) = \begin{bmatrix} z_i & u_{i1} & u_{i2} & \dots & u_{ip} \\ z_n & u_{n1} & u_{n2} & \dots & u_{np} \end{bmatrix}, \quad (2)$$

де $z_i = y_i + e_i$, $u_{ij} = x_{ij} + \delta_{ij}$, $j = \overline{1, p}$, $e_i \in E$, $\delta_{ij} \in \Delta_j$, - випадкові помилки вимірювань, $M(E) = M(\Delta_j) = 0$, $D(E) = \sigma_E^2$, $D(\Delta_j) = \sigma_{\Delta_j}^2$.

Формально зміст процедури параметричної ідентифікації визначимо відображенням $\pi: (Z, U) \rightarrow A$, що діє із простору вихідних даних у простір A оцінок параметрів. Реалізувати це відображення можна застосовувавши той чи інший метод оцінювання параметрів. Виділимо групу методів $\pi_1, \pi_2, \dots$, застосування кожного з котрих в умовах відсутності помилок призводить до точного результату, тобто $\pi_g: (Y, X) \rightarrow A^*$. Однак обробка цими методами реальних даних, що містять помилки вимірювань, дає множину неспівпадаючих в загальному випадку одне з одним та відмінних від A^* оцінок $\{\hat{\lambda}_g: \hat{\lambda}_g = \pi_g(z_i, u_i)\}$, $g = 1, 2, \dots$. Рівень відмінностей залежить від ряду факторів, зокрема від особливостей методів параметричної ідентифікації, що застосовуються. Виникає необхідність визначення якості параметричної ідентифікації кожним з методів $\pi_1, \pi_2, \pi_3, \dots$ та селекції за наслідками цього аналізу найкращого методу оцінювання параметрів.

Припустимо, що є змога багаторазового проведення експеримента з досліджуванним об'єктом, внаслідок чого маємо послідовність виборок $\{z_i, u_i\}_q$, $q = \overline{1, T}$ вихідних даних. Кожне повторення експерименту призводить до певної варіації складу матриці вихідних даних. Тому обробка послідовності виборок вихідних даних довільно обраним методом π_g дає сукупність оцінок $\{\hat{\lambda}_q\}_q$, $q = \overline{1, T}$, розташованих в околі точки A^* простору оцінок параметрів моделі. Запровадимо в цьому просторі метрику $\rho(\cdot, \cdot)$. Визначимо процедуру ідентифікації як стітку, якщо для оцінок $\hat{\lambda}_{qg}$.

розрахованих методом параметричної ідентифікації π за матрицями вихідних даних $(z_i, u_i)_q$, виконується умова $\rho(\hat{A}_{qg}, \hat{A}^*) \leq \chi_1$, де χ_1 - заданий пороговий рівень. Враховуючи, що вектор $\hat{A}^*$ невідомий, для незмішених оцінок можна умову стійкості ідентифікації представити у вигляді:

$$\rho(\hat{A}_{qg}, \hat{A}_{rg}) \leq \chi_2, \quad q \neq r. \quad (3)$$

де $\hat{A}_{qg}, \hat{A}_{rg}$ - два довільні вектора оцінок параметрів моделі.

Очевидно, що якість ідентифікації тим вища, чим менше верхній рівень можливих значень $\rho(\cdot)$, тобто останній можна використати як показник якості параметричної ідентифікації. Однак для його практичного застосування треба визначити можливість кількісного оцінювання значень цього показника в умовах реальної ідентифікації. Аналітичний розрахунок значень $\rho(\cdot)$ може бути досить складним і не завжди здійсненим залежно від особливостей застосованого методу параметричної ідентифікації, наявності та характеру апіорної інформації про вихідні дані, зокрема про властивості випадкових помилок в цих даних. Більш сприятливим є застосування чисельного методу визначення рівня стійкості процедури ідентифікації шляхом варіювання вихідних даних та наступної обробки отриманої множини векторних оцінок $(\hat{A}_q)$, $q = \overline{1, T}$, кожна з котрих - результат ідентифікації параметрів за черговою варіюваною вибіркою.

Крім того, можливо введення показників якості ідентифікації, що відрізняються за своїм змістом від метрики $\rho(\cdot)$, однак дозволяють кількісно оцінити рівень стійкості ідентифікації, простіші за способом обчислення, мають скалярний характер і т.п.

Варіювання вихідних даних шляхом багаторазового повторення натурного експерименту приводить до різкого зростання витрат, до того ж не завжди можливе. Доцільно використання штучно отриманих виборок - сурогатів, що своїми властивостями імітують реалізації експериментально здобутих даних. Природньо-випадковий характер варіювання даних забезпечує спосіб управління вибіркою, застосований у бутстрепі - відносно новому напрямку статистичної обробки даних. Зміст цього способу: вибірка (2) вихідних даних розглядається як генеральна сукупність значень випадкового вектора $[Z, U_1, \dots, U_r]$, з якої випадковим вибиранням формуються так звані бутстреп-вибірки вихідних даних $(z_i^0, u_i^0)_q$, $q = \overline{1, T}$, що у статистичному розумінні - аналоги вихідних даних (2).

Застосовувавши до бутстреп-виборок певний метод ідентифікації, отримаємо множину векторних бутстреп-оцінок π_g , $(\hat{\alpha}_{gq}^\sigma; \hat{\lambda}_{gq}^\sigma = \pi_g(z_{iq}^\sigma, u_{iq}^\sigma))$, за якими можна визначити рівень стійкості параметричної ідентифікації методом π_g . Зокрема для кількісної характеристики стійкості оцінок кожного з параметрів α_j , $j = \overline{1, p}$ можна використати середній квадрат $V^2(\hat{\alpha}_j)$ розсіювання відповідних оцінок $(\hat{\alpha}_{gjq}^\sigma)$:

$$V_g^2(\hat{\alpha}_j) = \frac{1}{I-1} \sum_{q=1}^I (\hat{\alpha}_{gjq}^\sigma - \bar{\alpha}_{gj}^\sigma)^2, \quad j = \overline{1, p}, \quad (4)$$

де $\bar{\alpha}_{gj}^\sigma = \frac{1}{I} \sum_{q=1}^I \hat{\alpha}_{gjq}^\sigma$ - середнє значення бутстреп-оцінок j -ого параметру моделі. За аналогією з регресійним аналізом введемо показник:

$$t_{gj} = \frac{|\bar{\alpha}_{gj}^\sigma| \sqrt{I-1}}{V_g^2(\hat{\alpha}_j)}$$

При $I \geq 30$ закон розподілу середнього $\bar{\alpha}_{gj}^\sigma$ близький до нормального, тому можна вважати, що значення t_{gj} підпорядковані розподілу Стюдента і для оцінювання значущості коефіцієнту $\bar{\alpha}_{gj}^\sigma$ може бути використана стандартна процедура перевірки, відома з лінійного регресійного аналізу. Однак на відміну від нього, в даному випадку ця перевірка придатна для аналізу оцінок коефіцієнтів будь-якої моделі включно з нелінійною за параметрами.

При ідентифікації багатовимірних моделей інформація про стійкість оцінювання параметрів моделі g -им методом задається вектором $[V_g^2(\hat{\alpha}_1), \dots, V_g^2(\hat{\alpha}_p)]$ чи відповідно вектором $[t_{g1}, \dots, t_{gp}]$. Векторна форма показника стійкості ідентифікації ускладнює порівняльний аналіз якості ідентифікації різними методами. Більш зручний скалярний показник якості, в якому враховано інформацію про стійкість оцінювання окремих параметрів:

$$\bar{V}_g^2 = \frac{1}{p} \sum_{j=1}^p V_g^2(\hat{\alpha}_j) / (\bar{\alpha}_{gj}^\sigma)^2. \quad (5)$$

Нестійке оцінювання того чи іншого параметру веде до зростання відповідного доданку $V_g^2(\hat{\alpha}_j) / (\bar{\alpha}_{gj}^\sigma)^2$, а разом з цим і всього показника $\bar{V}_g^2$ в цілому. Значна величина показника $\bar{V}_g^2$ - свідомство некоректності задачі, що розв'язується.

Ще один скалярний показник стійкості ідентифікації можна отримати, якщо оцінювання стійкості вести за значеннями емпіричного ризику, що розраховуються на множині бутстреп-

вибірок. Припустимо, $\hat{\lambda}_{gq}$ - вектор оцінок параметрів, розрахований методом π_g за q -ю бутстреп-вибіркою. Обчислимо на g -ій бутстреп-вибірці емпіричний ризик моделі $m(\hat{\lambda}_{gq})$:

$$I_{ng}^{(q,r)} = \frac{1}{n} \sum_{i=1}^n [y_{ir} - m(x_{ir}, \hat{\lambda}_{gq})]^2.$$

Для множини $(I_{ng}^{(q,r)})$, $r = \overline{1, I}$, $g \neq q$ визначимо середнє значення

$$\bar{I}_{ng}^{(q)} = \frac{1}{I-1} \sum_{\substack{r=1 \\ r \neq q}}^I I_{ng}^{(q,r)}$$

і потім статистику:

$$W_{gq}^2(I_n) = \frac{1}{I-2} \sum_{\substack{r=1 \\ r \neq q}}^I [I_{ng}^{(q,r)} - \bar{I}_{ng}^{(q)}]^2.$$

Приймаючи до уваги, що q змінюється в діапазоні $q = \overline{1, I}$, стійкість ідентифікації можна оцінити показниками типу:

$$\bar{W}_g^2 = \frac{1}{I} \sum_{q=1}^I W_{gq}^2(I_n), \quad (6)$$

$$W_{g \max}^2 = \max_{q=1, I} (W_{gq}^2(I_n)), \quad (7)$$

або статистиком

$$W_{gq}^2(I_n) = \frac{1}{(I-1)-1} \sum_{g=1}^I \sum_{\substack{r=1 \\ r \neq q}}^I (I_{ng}^{(q,r)} - \bar{I}_{ng}^{(q)})^2, \quad (8)$$

$$\bar{I}_{ng} = \frac{1}{I(I-1)} \sum_{g=1}^I \sum_{\substack{r=1 \\ r \neq q}}^I I_{ng}^{(q,r)}.$$

В групі показників якості (стійкості), що спираються на оцінювання варіації емпіричного ризику, відсутній порівняно з показником V^2 момент штучного згортання вектора до скалярної форми.

В роботі виконано порівняльний аналіз показників якості параметричної ідентифікації, на системі тестів досліджено властивості цих показників. Показано, наприклад, що показник $\bar{V}^2$ характеризується найбільшою універсальністю та надійністю, зокрема лишається достатньо ефективним і в разі наявності помилок в значеннях незалежних змінних, коли інші показники з ростом рівня цих помилок стають неінформативними. Визначено, що за умов, коли варіація вихідних даних реалізується виключно варіацією значень помилки вимірювання E при постійній матриці плану експеримента $\{x_i\}$, тобто $(z_i, x_i)^\sigma = (y_i + e_i^\sigma, x_i)$, показник стійкості $V^2(\hat{\alpha}_j)$, $j = \overline{1, p}$ за своїм змістом співпадає з дисперсією

відповідної оцінки $D(\tilde{a}_j)$, зокрема знайденою за допомогою бутстрепу: $D^0(\tilde{a}_j) = V^2(\tilde{a}_j)$. При застосуванні інших способів варіювання даних остання рівність не виконується. Це свідчить, що умови утворення бутстреп-вибірок при аналізі якості ідентифікації значно менш жорсткі порівняно з традиційним бутстрепом. Тому маємо змогу розробити принципово нові методи побудови (генерації) варіюваних даних, які на відзнаку від традиційного бутстрепа доцільно назвати бутстреподібними. Цей же термін треба було б використовувати відносно методів і результатів подальшої обробки варіюваних даних, але вжитку набули більш зручні скорочення на зразок "бутстреп-генератор", "бутстреп-дослідження" та т.п., відсутні у власно бутстрепі.

У роботі також показано, що хоча загалом $D(\tilde{a}_j) \neq V^2(\tilde{a}_j)$, при великих значеннях n , зокрема у асимптотиці, та у випадку стійкої процедури ідентифікації співвідношення $D^0(\tilde{a}_j) \approx V^2(\tilde{a}_j)$ виконується незалежно од застосованого метода варіювання даних.

Наведені вище показники стійкості оцінювання являються об'єктивною характеристикою якості тільки параметричної ідентифікації, бо як показує аналіз їх властивостей, кількісні значення цих показників є зростаючою функцією рівня складності моделі, що робить неможливим їх застосування для селекції найкращої моделі на множині моделей-претендентів довільної структури. Емпіричному ризику I_n властива зворотня тенденція: його кількісні значення спадають з підвищенням рівня складності моделі. Поєднання цих двох типів показників дозволяє отримати комбінований показник якості ідентифікації, який дає змогу вибрати модель, високі апроксимаційні якості якої, що досягаються насамперед за рахунок достатньо складної структури моделі, забезпечуються разом із стійкою ідентифікацією параметрів.

В роботі розроблено та проаналізовано кілька типів комбінованих показників. Найуживаніший з них має вигляд:

$$Q_{1,j} = (1-\gamma) \bar{V}_j^2 + \gamma I_{n,j} / R_z^2, \quad j = \overline{1, K}, \quad (9)$$

де R_z - розмах вибірки $\{z_i\}$, k - загальна кількість моделей-претендентів $\{m_1, \dots, m_j, \dots, m_k\}$, γ - ваговий коефіцієнт, що визначає пріоритет вимог стійкості оцінювання параметрів або точності апроксимації моделлю вихідних даних. Кращою з моделей $m_1, \dots, m_k$ вважається та, що забезпечує мінімум показника $Q_{1,j}$.

Дуже суттєвим є правильний вибір значення вагового

коефіцієнту. Якщо моделі-претенденти мають приблизно однакові апроксимативні властивості, вибірка $\{I_{nj}\}$, $j = \overline{1, K}$ характеризується незначним розмахом $R_I = I_{n\max} - I_{n\min}$, де $I_{n\max} = \max_{j=\overline{1, K}} \{I_{nj}\}$, $I_{n\min} = \min_{j=\overline{1, K}} \{I_{nj}\}$, і пріоритетним чинником у виборі моделі стає стійкість оцінювання параметрів моделі, тобто $\gamma \rightarrow 0$, $(1-\gamma) \rightarrow 1$. Враховуючи, що у цій ситуації $I_{n\min} \rightarrow I_{n\max}$, $R_I \rightarrow 0$, можна коефіцієнт ваги визначити як $\gamma = R_I / I_{n\max} \rightarrow 0$. У протилежній ситуації, коли стійкість параметричної ідентифікації моделей приблизно однакова, а розсіювання показника I_n значне, маємо: $\gamma = R_I / I_{n\max} = (1 - I_{n\min} / I_{n\max}) \rightarrow 1$, $(1-\gamma) \rightarrow 0$, переважаючи роль у виборі моделі відіграє точність апроксимації моделлю вихідних даних, тобто емпіричний ризик I_{nj} .

Розглянутий спосіб визначення коефіцієнту γ не єдиний. Для цього можна, наприклад, використати співвідношення $\gamma = 1 - (\bar{V}_{\max}^2 - \bar{V}_{\min}^2) / \bar{V}_{\max}^2$, де $\bar{V}_{\max}^2$, $\bar{V}_{\min}^2$ - найбільший та найменший елемент вибірки $\{\bar{V}_j^2\}$, $j = \overline{1, K}$. Крім того, приймаючи до уваги, що розмах R_I чутливий до аномальних помилок (у даному разі це значення I_n , розраховані для аномально грубої моделі), коефіцієнт γ можна обчислити за формулою $\gamma = s \sqrt{D_I} / \bar{I}_n$, де D_I - вибіркова дисперсія, розрахована за цензурованою вибіркою $\{I_{nj}\}$, $\bar{I}_n$ - середнє значення для тієї ж вибірки, s - стала, $s \approx 1 + 2$. З тих же причин (чутливість розмаху до аномалій) при високому рівні шумів в залежній змінній доцільно у формулі (9) замінити розмах R_z на оцінку дисперсії $D(Z)$.

Загалом, залежно від способу поєднання окремих показників $\bar{V}^2$ і I_n , можна розробити цілий ряд комбінованих показників типу (9). Аналогічним шляхом можна будувати комбіновані показники якості ідентифікації на базі статистик I_n і W^2 . Наприклад, використавши стандартний прийом нормування порівнянням, отримаємо

$$Q_* = \frac{I_{nj}}{I_{n\max}} + \frac{\bar{W}_j^2}{\bar{W}_{\max}^2} \quad (10)$$

Незважаючи на можливість певного обґрунтування вибору вагових коефіцієнтів в формулах (9), (10), зокрема наведених вище міркувань з приводу вибору значення γ , за способом свого утворення

вирази (9),(10) - типові евристичні показники, придатність яких до практичного застосування визначається експериментальним шляхом, аналізом успішності їх використання при розв'язанні реальних задач, на тестових прикладах. Тому особливий інтерес представляє показник якості Q_2 , структура якого виводиться строго математичними засобами. У другому розділі роботи показано, що для лінійних моделей за умов незміщеного оцінювання вектора параметрів A має місце співвідношення:

$$I = M\{I_n(\hat{A})\} + \frac{1}{n} M\{(\hat{A} - A)^T G(\hat{A} - A)\}, \quad (11)$$

де G - інформаційна матриця, $G = [x_i]^T [x_i]$, $[x_i]$ - матриця регресорів (матриця плану), що входить у склад матриці вихідних даних (2). За своєю структурою вираз (11) аналогічний формулі (1), другий доданок характеризує зміщення емпіричного ризику як оцінки I . Безпосередньо використати (11) для визначення I не можна, бо невідомий вектор A , не має змоги реально обчислити математичне сподівання.

Докорінно змінити ситуацію дозволяє використання бутстреп-виборок. Обчислені на них оцінки $\hat{A}_1^{\sigma}, \dots, \hat{A}_I^{\sigma}$ дають змогу замінити A його оцінкою $\bar{A}^{\sigma} = \frac{1}{q} \sum_{q=1}^q \hat{A}_q^{\sigma}$, а математичне сподівання замінити усередненням на множині. В результаті отримаємо показник Q_2 , структура якого відповідає виразу (11):

$$Q_2 = \frac{1}{q} \sum_{q=1}^q I_{nq}^{\sigma} + \frac{1}{n(1-1)} \sum_{q=1}^q (\hat{A}_q^{\sigma} - \bar{A}^{\sigma})^T G_q^{\sigma} (\hat{A}_q^{\sigma} - \bar{A}^{\sigma}), \quad (12)$$

де статистики I_{nq}^{σ} обчислені на q -ій бутстреп-вибірці вихідних даних $(z_i^{\sigma}, x_i^{\sigma})_{i=1}^n$, яка характеризується інформаційною матрицею G_q^{σ} . За своїм змістом Q_2 з одного боку - оцінка середнього ризику I , що є незміщеною і має, за результатами тестових випробувань, суттєво низьке розсіювання, а з другого - типовий комбінований показник якості ідентифікації, перший доданок якого характеризує апроксимаційні якості моделі, а другий - рівень стійкості оцінювання її параметрів. Останній висновок стає очевидним, якщо припустити ортогональність незалежних змінних. В цьому випадку вираз (12) спрощується й приймає вигляд:

$$\begin{aligned} Q_2 &= \frac{1}{q} \sum_{q=1}^q I_{nq}^{\sigma} + \frac{1}{n(1-1)} \sum_{q=1}^q \left[\sum_{m=1}^p (\hat{a}_{mq}^{\sigma} - \bar{a}_m^{\sigma}) \sum_{l=1}^n (x_{lqm}^{\sigma})^2 \right] = \\ &= \frac{1}{q} \sum_{q=1}^q I_{nq}^{\sigma} + \sum_{m=1}^p V^2(\hat{a}_m^{\sigma}) \frac{1}{n} \sum_{l=1}^n (x_{lqm}^{\sigma})^2. \end{aligned} \quad (13)$$

Порівнюючи показники (9),(13) та приймаючи до уваги (5), бачимо, що вони відрізняються способом нормування, який

застосовано при згортанні вектора $[V^2(\tilde{a}_1), \dots, V^2(\tilde{a}_p)]$ до скаляра. Другий доданок в (13) (а в загальному випадку - в (12)) - це показник стійкості параметричної ідентифікації, який може використовуватися самостійно для визначення рівня якості оцінювання. Область застосування цього показника, як і Q_3 , в цілому - задачі ідентифікації лінійних моделей в умовах відсутності помилок в незалежних змінних, що значно вужче, ніж для V^2 , Q_1 (моделі довільної структури при довільному зашумленні даних) або W^2 , Q_4 (довільна структура але відсутність шумів в незалежних змінних). Однак у даному разі важлива математично доведена близькість Q_3 до оптимального показника якості ідентифікації - середнього ризику I , що дає змогу використати Q_3 як базу для порівняльного аналізу евристичних показників. Зокрема, показана в роботі подібність властивостей показників Q_1 , Q_4 та Q_3 в межах застосування останнього дозволяє припустити, зважаючи на незалежність процедури обчислення Q_1 , Q_4 від виду моделі та характеру вихідних даних, ефективність використання цих показників в загальному випадку.

У третьому розділі дисертації розглянуто застосування введеної вище системи показників для аналізу якості ідентифікації на прикладі розв'язання конкретних задач ідентифікації, а також питання підвищення якості ідентифікації за допомогою методів, для реалізації яких, як і в аналізі якості, використовуються способи управління (варіювання) складом вибірки вихідних даних.

Насамперед це дослідження можливостей традиційного бутстрепа для уточнення оцінок параметрів моделі та визначення характеристик їх розсіювання за умов використання для побудови бутстреп-вбірок нових методів, що застосовані для варіювання складу вихідних даних при обчисленні показників стійкості ідентифікації.

Один з таких методів - метод селектурних співвідношень, в якому бутстреп-вбірки будуться шляхом генерації даних за їх характеристиками, заданими у непараметричній формі. Якщо вихідні дані (2) - це реалізація випадкового вектора $[Z, U_1, \dots, U_p]$, кожен елемент якого має свій закон розподілу $[F_Z, F_1, \dots, F_p]$, а статистичні зв'язки елементів визначаються кореляційною матрицею $R = [\rho_{j,q}]$, $\rho_{j,j} = \text{cov}(Z, U_j) / \sigma_Z \sigma_{u_j}$, $\rho_{j,q} = \text{cov}(U_j, U_q) / \sigma_{u_j} \sigma_{u_q}$, $j, q = 1, p$, процедура методу селектурних співвідношень включає два етапи. На першому за знайденими з експериментальних даних

емпіричними законами $[F_z, F_1, \dots, F_p]$ генеруються некоррельовані послідовності $\{\tilde{z}_k\}, \{\tilde{u}_{k1}\}, \dots, \{\tilde{u}_{kp}\}, k = 1, 2, \dots$, з яких на другому етапі виділяються лише ті елементи, що відповідають певній системі селектурних співвідношень. Для змінної Z це перші n елементів, для U_i - це елементи послідовності $\{\tilde{u}_{ki}\}$, відносно яких виконуються умови:

$u_{i1}^\sigma \in [z_i^\sigma \tilde{\rho}_i, \sigma_{u1}/\sigma_z \pm \sigma_{u1} d_i^\pm \sqrt{1 - \tilde{\rho}_i^2}], i = \overline{1, n},$ (14)
де d_i^+, d_i^- - квантили розподілу F_1 . Для довільної змінної U_j бутстреп-вибірка складається з елементів, що відповідають системі умов:

$$\begin{cases} u_{ij}^\sigma \in [z_i^\sigma \tilde{\rho}_j, \sigma_{uj}/\sigma_z \pm \sigma_{uj} d_j^\pm \sqrt{1 - \tilde{\rho}_j^2}], \\ u_{ij}^\sigma \in [u_{iq}^\sigma \tilde{\rho}_{qj}, \sigma_{uj}/\sigma_{uq} \pm \sigma_{uj} d_j^\pm \sqrt{1 - \tilde{\rho}_{qj}^2}], \quad q < j, \end{cases} \quad (15)$$

$$i = \overline{1, n}, q = \overline{1, j-1}.$$

Значення $\tilde{\rho}_j, \tilde{\rho}_{qj}$, що входять в (14), (15) - елементи вибіркової кореляційної функції $\tilde{R} = [\tilde{\rho}_j, \tilde{\rho}_{qj}]$.

Крім методу селектурних співвідношень, в роботі розглянуто інші, більш прості за своєю алгоритмічною реалізацією, можливість застосування яких регламентовано вимогою $n \geq 40-50$. Це метод випадкового вибору (варіювання) рядків матриці вихідних даних, методи, що базуються на виділенні та варіюванні шумової складової вихідних даних при збереженні постійними інформативних значень цих даних, методи, пов'язані з розподілом вихідних даних на підвибірки меншого об'єму.

Виконано порівняльний аналіз розглянутих методів для побудови бутстреп-вибірки при дослідженні якості та поліпшенні результатів ідентифікації моделей. Показано, що при $n \geq 40-50$ спосіб утворення бутстреп-вибірок практично не впливає на наслідки бутстреп-досліджень.

Окрім аналізу можливостей застосування модифікованих засобів бутстреп-обробки даних в роботі розглянуто нові методи підвищення якості ідентифікації, що використовуються в рамках вже обраного інструментарію структурно-параметричної ідентифікації. Зокрема, в разі використання для оцінювання параметрів моделі кількох методів параметричної ідентифікації $\pi_1, \pi_2, \dots, \pi_k$, які дають сукупність незмішених векторних оцінок $\tilde{\lambda}_1, \tilde{\lambda}_2, \dots, \tilde{\lambda}_k$, запро-

поновано будувати оптимальні групові уточнені оцінки, що мають вигляд:

$$\bar{a}_j = \sum_{r=1}^k w_{jr} \tilde{a}_{jr}, \quad j = \overline{1, p}, \quad (16)$$

де $\tilde{a}_{jr}$ - елемент вектора оцінок $\tilde{A}_r$, $r = \overline{1, k}$. Якщо під оптимальністю оцінки (16) розуміється мінімум її можливої дисперсії $D(\bar{a}_j)$, то оцінка $\bar{a}_j$ - це, виходячи з геометричних уявлень, проекція вектора a_j (невідоме фактичне значення j -ого параметра моделі $m(A)$) на гіперплощину L_j , елементи якої визначаються виразом (16). В цьому випадку різниця $a_j - \bar{a}_j$ перпендикулярна площині L_j , тобто перпендикулярна до довільної прямої, яка лежить у гіперплощині L_j та проходить через основу перпендикуляру, зокрема до відрізків $\tilde{a}_{jr} - \bar{a}_j$, $r = \overline{1, k}$, що, приймаючи до уваги випадковий характер оцінок $\tilde{a}_{jr}$, $\bar{a}_j$, описується співвідношенням:

$$M((a_j - \bar{a}_j)(\tilde{a}_{jr} - \bar{a}_j)) = 0, \quad r = \overline{1, k}. \quad (17)$$

Оцінка $\bar{a}_j$, яка відповідає умові (17), має мінімальну дисперсію серед усіх оцінок, що лежать у гіперплощині L_j , $D(\bar{a}_j) = \sum_{r=1}^R \sum_{g=1}^k w_{jr} w_{jg} \text{cov}_{rg}^{(j)}$, де $\text{cov}_{rg}^{(j)} = \text{cov}(a_{jr}, a_{jg})$. Як показано в дисертаційній роботі, із співвідношення (17), підставивши в нього формулу (16), можна отримати $k-1$ незалежне рівняння, що разом із умовою $\sum_{r=1}^k w_{jr} = 1$ утворюють систему:

$$\begin{cases} \sum_{r=1}^k w_{jr} (\text{cov}_{rg}^{(j)} - \text{cov}_{rk}^{(j)}) = 0, & g = \overline{1, k-1}, \\ \sum_{r=1}^k w_{jr} - 1 = 0, \end{cases} \quad (18)$$

з якої обчислюються значення вагових коефіцієнтів w_{jr} , $r = \overline{1, n}$.

Для $n = 2$ із (18) маємо:

$$w_{j1} = \frac{\sigma_{j2}^2 - \text{cov}_{j2}}{\sigma_{j1}^2 + \sigma_{j2}^2 - 2\text{cov}_{j2}}, \quad w_{j2} = \frac{\sigma_{j1}^2 - \text{cov}_{j2}}{\sigma_{j1}^2 + \sigma_{j2}^2 - 2\text{cov}_{j2}}, \quad (19)$$

$$\text{де } \text{cov}_{j2} = \text{cov}_{j2}^{(j)} - \text{cov}_{j2}^{(j)}, \quad a_j = w_{j1} a_{j1} + w_{j2} a_{j2},$$

$$D(\bar{a}_j) = w_{j1}^2 \sigma_{j1}^2 + w_{j2}^2 \sigma_{j2}^2 + 2w_{j1} w_{j2} \text{cov}_{j2}, \quad \sigma_{jt}^2 = \text{cov}_{tt}^{(j)}, \quad t=1,2.$$

Елементи коваріаційних матриць $\{\text{cov}_{rg}^{(j)}\}$, $j = \overline{1, p}$ одержують, обробивши відповідними методами параметричної ідентифікації π_r , π_g бутстреп-вибірки $\{z_i^{\sigma}, x_i^{\sigma}\}_q$, $q = \overline{1, T}$ вихідних даних, після чого на отриманих множинах $\{A_{rg}^{\sigma}\}$, $\{A_{gq}^{\sigma}\}$ обчислюються вибіркові оцінки

$$\tilde{\text{cov}}_{rg}^{(j)} = \frac{1}{T-1} \sum_{q=1}^T (\tilde{a}_{jrg}^{\sigma} - \bar{a}_{jr})(\tilde{a}_{jgq}^{\sigma} - \bar{a}_{jg}).$$

У третьому розділі дисертації розглянуті також можливості

застосування бутстреп-характеристик якості ідентифікації, зокрема показників $\bar{V}^a$ і $\bar{I}_n$ для виділення суттєвих змінних моделі та пошуку її структури.

Зміст четвертого розділу - особливості дослідження якості ідентифікації динамічних систем. Вихідні дані у цьому випадку представлені сукупністю часових рядів: $\{z_i, u_{ij}\} = \{z(t_i), u_i(t_1), \dots, u_i(t_p)\}$, $i = \overline{1, n}$, $\tau = t_{i+1} - t_i = \text{const}$, що створює додаткові умови для виділення й аналізу шумової та корисної (інформативної) складових вихідних даних за рахунок використання операції згладжування, $\tilde{y}_i = \text{Sm}\{z_i\}$, $\tilde{x}_{ij} = \text{Sm}\{u_{ij}\}$, $\tilde{e}_i = z_i - \tilde{y}_i$, $\tilde{\delta}_{ij} = u_{ij} - \tilde{x}_{ij}$, де $\text{Sm}(\cdot)$ - оператор згладжування (згладжуючий фільтр), а значком $\sim$ позначені оцінки відповідних складових вихідних даних. Подальша побудова бутстреп-виборок здійснюється за формулами $z_i^\sigma = \tilde{y}_i + e_i^\sigma$, $u_{ij}^\sigma = \tilde{x}_{ij} + \delta_{ij}^\sigma$, тобто шляхом варіювання помилок вимірювання у вихідних даних. До утворених бутстреп-виборок можна застосувати усі розглянуті вище методи аналізу й поліпшення якості ідентифікації. Крім того, для дослідження якості ідентифікації динамічних систем запропоновано ряд суттєво нових методів підвищення якості.

Зокрема, це корегування методичних змішень в оцінках параметрів моделей. Розглянемо цю задачу на прикладі оцінювання параметрів лінійної регресійної моделі методом найменших квадратів (МНК) за вихідними даними, що мають випадкові помилки в значеннях незалежних змінних. Обчислений за цими даними вектор МНК-оцінок $\tilde{\lambda} = (\underline{u}^T \underline{u})^{-1} \underline{u}^T \underline{z}$, де $\underline{z}, \underline{u}$ - вибірки, що складають матрицю вихідних даних $\{z_i, u_{ij}\}$, $i = \overline{1, n}$, $j = \overline{1, p}$, має зміщення $B = M(\tilde{\lambda}) - A$. Як показано у роботі, складові вектора B можна визначити через математичне сподівання вектора $\tilde{\lambda}$: $B = (C - I)^{-1} \text{CM}(\tilde{\lambda})$, де $C = M((\underline{u}^T \underline{u})^{-1} \underline{u}^T \underline{\delta})$, $\underline{\delta} = \underline{u} - \underline{x}$ - матриця помилок вимірювання незалежних змінних. Використовуючи бутстреп, замінемо математичні сподівання відповідними середніми, обчисленими на множині бутстреп-виборок $\{u_i^\sigma\}_g, \{\delta_{ij}^\sigma\}_g$, $g = \overline{1, G}$ та бутстреп-оцінок $\lambda_1^\sigma, \dots, \lambda_p^\sigma$, тобто $\tilde{C} = \frac{1}{G} \sum_g (\underline{u}_g^{\sigma T} \underline{u}_g^\sigma)^{-1} \underline{u}_g^{\sigma T} \underline{\delta}_g^\sigma$, $\tilde{A}^\sigma = \frac{1}{G} \sum_g \lambda_g^\sigma$. Тоді $\tilde{B} = (\tilde{C} - I)^{-1} \tilde{C} \tilde{A}^\sigma$, різниця $\tilde{\lambda}_{\text{НС}} = \tilde{A}^\sigma - \tilde{B}$ дає незміщену оцінку вектора параметрів.

Наведений метод корегування змішень придатний тільки при роботі з лінійними за параметрами моделями. Для загального

випадку у дисертації розглянута евристична процедура корегування зміщень в оцінках параметрів моделі довільної структури. Вихідним етапом цієї процедури є генерація бутстреп-вибірок згідно з формулою:

$\{z_{1g}^{\sigma}, u_{1g}^{\sigma}, \dots, u_{p_g}^{\sigma}\} = \{y_{1g}^{\sigma} + \varepsilon_{1g}^{\sigma}, \tilde{x}_{1g}^{\sigma}, \dots, \tilde{x}_{p_g}^{\sigma} + \delta_{p_g}^{\sigma}\}, g = \overline{1, T}$,
де $y_{1g}^{\sigma} = m(A_{1g}, \tilde{x}_{1g}^{\sigma}, \dots, \tilde{x}_{p_g}^{\sigma})$ - розрахункові значення залежної змінної, обчислені за моделлю m з заданим вектором параметрів A_{1g} , причому вибір векторів $A_{1g}, \dots, A_{Tg}$ виконується таким чином, щоб сукупність $\{A_{1g}\}, g = \overline{1, T}$ перекривала окіл точки A - невідомого вектора дійсних значень параметрів моделі m .

За результатами оцінювання параметрів моделі на кожній з бутстреп-вибірок маємо сукупність оцінок $\{\tilde{A}_{1g}^{\sigma}\}, g = \overline{1, T}$ й відповідну сукупність пар $\{\tilde{A}_{1g}^{\sigma}, \tilde{B}_{1g}^{\sigma}\}$, де $\tilde{B}_{1g}^{\sigma} = \tilde{A}_{1g}^{\sigma} - A_{1g}$ - оцінка зміщень елементів вектора $\tilde{A}_{1g}^{\sigma}$. Побудова апроксимуючої залежності $B = \psi(\tilde{A})$ дозволяє оцінити зміщення вектора параметрів $\tilde{A}$, обчисленого за реальними вихідними даними $\{z_{1j}, u_{1j}\}$. При високому рівні зміщень доцільно знайти незміщену оцінку вектора параметрів $\tilde{A}_{\text{нз}} = \tilde{A} - \tilde{B} = \tilde{A} - \psi(\tilde{A})$.

Ефективність методики бутстреп-дослідження якості ідентифікації за даними, що представлені часовими рядами, визначається насамперед точністю розділення шумової та корисної складових вихідних даних, тобто якості процедури згладжування $Sm(\cdot)$. За умов мінімуму апріорної інформації про властивості шумової та корисної складових, їх нестационарності, оптимізація якості згладжування досягається використанням фільтрів з нестационарними імпульсними характеристиками (IX), вигляд та параметри яких адаптуються до поточних характеристик вихідних даних.

У роботі розглядається лінійний адаптивний фільтр із скінченю IX, згладжені оцінки на виході якого обчислюються за виразом

$$\tilde{y}_l = \sum_{j=-\mu}^k w_j z_{l+j}, \quad (20)$$

де $w_j, j = -\mu, \mu$ - IX фільтра. Показано, що за умов, коли корисний сигнал $y(t)$ на довільно розташовану відрізку часу, $t - t_1 \leq T, T = 2\mu_{\text{max}} \tau$, можна представити у вигляді $s(t) = \sum_{k=0}^p a_k t^k$, $p \leq p_{\text{max}}$, задача поточної оптимізації IX фільтра при обчисленні чергової оцінки $\tilde{y}_l$ еквівалентна задачі побудови локально-оптимальної моделі сигналу на інтервалі, середина якого припадає

на i -ий елемент оброблювальної послідовності вихідних даних. Локальна модель задається у вигляді $m(g, \alpha, p, \mu) = \sum_{k=0}^p \alpha_k P_k(g)$, $g = 0, 2\mu$, де $P_k(g)$ - ортогональні на множині $0, 1, \dots, 2\mu$ дискретні поліноми Чебишева. Двійка (p, μ) - параметри адаптації локальної моделі, які відповідно визначають порядок моделі та довжину $2\mu+1$ інтервала, на якому вона базується, і можуть змінюватися згідно характеру даних, зокрема поведінці складової $y(t)$ на інтервалі побудови локальної моделі. В цілому структурно-параметрична оптимізація моделі реалізується за критерієм

$$\tilde{\alpha}, p_{opt}, \mu_{opt} = \operatorname{argmin} \left(\frac{1}{2\mu-p-1} \sum_{j=-\mu}^{\mu} [z_{i+j} - \sum_{k=0}^p \alpha_k P_k(\mu+j)]^2 - \sigma_E^2 \right), \quad (21)$$

де σ_E^2 - дисперсія помилки вимірювання $E(t)$, тобто структурна оптимізація локальної моделі забезпечується шляхом адаптації параметрів p, μ . В роботі розглянуто особливості процедури оптимізації локальної моделі за виразом (21) та запропоновано простий алгоритм її практичної реалізації за підсиленням критерієм, відповідно якому згладжене значення $\tilde{y}_i = \sum_{k=0}^p \tilde{\alpha}_k P_k(\mu)$ обчислюється за моделлю, структура та параметри котрої визначено згідно з умовою

$$\frac{1}{2\mu-p-1} \sum_{\xi=0}^{\mu} \sum_{j=-\mu}^{\mu} (z_{i+j} - \tilde{y}_{i+j+\xi}) (z_{i+j+\xi} - \tilde{y}_{i+j+\xi}) \in [M + \gamma \sqrt{M}], \quad (22)$$

де $\tilde{y}_g = m(g, \tilde{\alpha}, p, \mu)$, $M = [1 - (p+1)/(2\mu+1)] \sigma_g^2$, $D = \sigma_g^2 ((2\mu-p-1)^2(2\mu+p+2) + 4\mu(p+1)^2 + (2\mu+1)(2\mu-p-1)) / (2\mu+1)^2 2\mu(2\mu-p)$, $\gamma \approx 2+3$.

Алгоритм адаптації локальної моделі до даних базується на почерговій перевірці умови (22) для моделей наростаючого порядку складності $p = 0, 1, \dots, p_{max}$ на інтервалі визначеної довжини, починаючи з найдовшого, що складається з $2\mu_{max}+1$ послідовних елементів. При невиконанні умови (22) для всіх p , включно з p_{max} , довжина інтервалу скорочується на 2 шляхом відкидання крайніх правого та лівого елементів. Для скороченого інтервалу повторюється цикл підбору порядку моделі. Так триває доки вперше не виконається умова (22). Використані у цьому випадку значення параметрів $p, \mu, \tilde{\alpha}$ близькі до оптимальних. Обчисливши за ними оцінку $\tilde{y}_i = m(g, \tilde{\alpha}, p, \mu)$, виконуємо зсув середини модельного інтервалу на одиницю вправо й повторюємо усю процедуру адаптації моделі знов для обчислення чергової поточної оцінки $\tilde{y}_{i+j}$.

Звичайно замість дисперсії σ_g^2 , що входить до складу

критеріїв (21), (22), використовують її оцінку $\hat{\sigma}_a^2$. Враховуючи нестационарний характер помилки E , доцільно реалізувати поточне уточнення оцінки дисперсії, застосовавши для цього простий рекурентний алгоритм, наприклад, метод експоненційного згладжування, корегуючи оцінку за рахунок інформації про середній квадрат поточної помилки S_a^2 апроксимації даних моделлю на черговому інтервалі обробки:

$$\hat{\sigma}_{a, l+1}^2 = \beta S_{al}^2 + (1-\beta) \hat{\sigma}_{al}^2,$$

де β - стала експоненційного фільтру.

На жаль, наведені алгоритми адаптивного згладжування мають обмежену швидкодію, тому їх застосування для бутстреп-дослідження якості ідентифікації, що передбачає багаторазове повторення операцій згладжування на кожній новій бутстреп-реалізації даних, сприяє суттєвому подовженню терміну досліджень.

У ряді випадків, наприклад, при ідентифікації лінійної за параметрами моделі, поліпшення якості ідентифікації досягається застосуванням простих швидкодіючих лінійних згладжувачих фільтрів, що працюють в режимі "перегладжування" даних, чим забезпечується незміщенність оцінок параметрів моделі. В цьому разі можна використовувати, незважаючи на присутність операції згладжування, звичайну бутстреп-процедуру перевірки якості моделі. Слід, однак, мати на увазі, що "перегладжування" даних може кінець-кінцем призвести до появи мультиколінеарності й різко погіршити якість ідентифікації. Тому оптимальним фільтром при застосуванні прийому "перегладжування" слід вважати такий, що вносить мінімальні викривлення корисної (інформативної) складової вихідних даних при заданому рівні придушення шумової складової. Вибір фільтра та його параметрів за цих умов забезпечується через спеціально виконану сертифікацію ряду типів згладжувачих фільтрів.

Мета сертифікації - визначити якість згладжування фільтром вихідних даних за різних умов роботи: при різних рівнях шумової складової, різних частотно-часових властивостях даних и т.п.. Сукупна інформація про залежність якості згладжування від умов застосування фільтра інтегрується у так званій моделі фільтра, яка пов'язує кількісну оцінку якості (залежна змінна) із значеннями факторів, що впливають на неї. Як показник якості використовується статистика $Q = \sum_{i=1}^n (y_i - \tilde{y}_i)^2$, де $\tilde{y}_i = \sum_{j=-\mu_1}^{\mu_2} w_j z_{i,j}$ - згладжена оцінка на виході фільтра. Якщо шумова складова $E(t_i)$ -

некоррельована випадкова стаціонарна послідовність з $M(E(t_i))=0$, $D(E(t_i))=\sigma^2$, для статистики Q має місце співвідношення:

$$q = M(Q) = \frac{1}{n} \sum_{i=1}^n (y_i - \sum_{j=-\mu_1}^{\mu_2} w_j y_{i+j})^2 - \sigma^2 \sum_{j=-\mu_1}^{\mu_2} w_j^2 = q_1 + q_2, \quad (24)$$

де q_1 визначає рівень помилок, що їх вносить фільтр у корисний сигнал, q_2 - рівень остаточної випадкової помилки. Декомпозиція показника q загальної якості згладжування на дві компоненти q_1 , q_2 дає певні переваги при практичному використанні результатів сертифікації фільтрів. Крім того побудова часткових моделей для q_1 , q_2 значно простіша, ніж для загального показника q .

Власне процедура побудови моделей виконується на системі тестових корисних сигналів та сигналів-шумів, тобто на момент сертифікації фільтра відомі фактичні значення y_i , оцінки $\hat{y}_i$, уся можлива інформація про властивості корисної та шумової складових вихідних даних. При застосуванні вже побудованих моделей для вибору фільтра з метою оптимізації обробки певної реальної послідовності $\{z_i\}$ для кожного фільтра-претендента за відповідною моделлю обчислюються оцінки $\tilde{q}_1$, $\tilde{q}_2$. Для цього у моделі потрібно ввести інформацію про характер даних, що обробляються. Однак апіорна інформація про частотно-часові характеристики даних на практиці звичайно відсутня, оцінювання цих характеристик безпосередньо за самими даними - важка задача. Тому як незалежні фактори у моделях використано евристичні статистики, що мають просту, але досить стійку процедуру обчислення й разом з тим містять достатньо інформації про властивості вихідних даних. В роботі з цих міркувань застосовано ряд статистик, що будуються за даними ряду $\{z_i\}$, або похідного від нього ряду $\{r_i : r_i = z_{i+1} - z_i, i = \overline{1, n-1}\}$, які оцінюють рівень хаотичності вихідних даних (статистика F), їх мінливість, складність (статистика V), характер розподілу елементів ряду $\{r_i\}$ (статистика H), "довжину" часового ряду $\{y_i\}$ при умовно заданій відстані між його елементами (статистика L) та статистики B_1 , B_2 , залежні від інерційності тенденції розвитку процесу $y(t)$ та інші.

В цілому модель фільтра - типова регресійна модель на зразок

$$q_i = a_0 + a_1 B_1 + a_2 B_2 + a_3 H + a_4 L. \quad (25)$$

Крім наведених вище показників q_1 , q_2 доцільним при побудові моделей фільтрів видалось введення показника

$$d_i = \max_{(w_j), (y_i)} |y_i - \sum_{j=-\mu_1}^{\mu_2} w_j y_{i+j}|.$$

Значно спрощує практичне застосування результатів сертифікації використання розширених моделей, які містять у ранзі додаткових факторів внутрішні для цього класу параметри фільтрів (довжину IX $\mu_1 + \mu_2$, порядок p поліноміального фільтру, сталу експоненційного згладжування β і т.п.), наприклад:

$$q_1 = a_0 + a_1 \mu^2 + a_2 \mu B_1 + a_3 B_2 + a_4 H + a_5 \mu^2 L,$$

$$d_1 = a_0 + a_1 \mu + a_2 V + a_3 \mu^2 B_1 B_2.$$

В разі обробки даних з застосуванням режиму "перегладжування" оптимізація вибору фільтру за результатами сертифікації полягає в обчисленні для часового ряду $\{z_t\}$ оцінок кількісних значень факторів F, V, H, L, B_1, B_2 та визначенні типа і параметрів фільтра, що для заданої величини показника q_2 забезпечує мінімальні значення показників q_1, d_1 . Як показує практика та досвід застосування "перегладжування" для задач параметричної ідентифікації, прийнятливі результати мають місце, коли для обраного згладжувачого фільтра виконується умова: $q_1, d_1 \leq 0,1$. Перебільшення цього порогу призводить до різкого погіршення стійкості ідентифікації.

Як зазначалося при формулюванні мети дисертаційної роботи, в межах однієї загальної задачі - розробки методології дослідження якості ідентифікації, по суті розв'язуються дві окремі хоча й тісно поєднані між собою задачі - оцінювання якості ідентифікації та розробка методів її підвищення. Крім суто логічного причинно-наслідкового зв'язку ці дві задачі об'єднує спільна технологія їх розв'язання. Вона базується на бутстрепоподібних прийомах та способах обробки і аналізу даних, що застосовуються як для оцінювання якості ідентифікації, так і для підвищення якості результатів її розв'язання й складають єдину технологічну вісь роботи. Систематизація матеріалів, що стосуються використання бутстреп-методів на різних етапах дослідження якості ідентифікації, вивчення взаємозв'язків між цими етапами, взаємозалежності засобів та методів реалізації бутстреп-процедур - все це становить зміст п'ятого розділу дисертації, мета якого - побудова єдиної технології дослідження задач ідентифікації.

У розділі виконано формальний аналіз бутстреп-процедур дослідження якості ідентифікації, з'ясовано їх розподіл на головні (базові) та допоміжні етапи. Показано, що, незважаючи на чітко окреслене функціональне призначення кожного з етапів, їх

програмно-алгоритмічне наповнення може істотно змінюватися залежно од кінцевої мети дослідження якості, що в свою чергу визначається постановкою задачі ідентифікації, властивостями вихідних даних, засобами їх опису, типами моделей і т.п.. Необхідність врахування усіх цих обставин вимагає від дослідника досить високого рівня спеціальної підготовки й змушує коло можливих користувачів бутстреп-технологією. В певній мірі пом'якшує ситуацію систематизація показників якості ідентифікації та способів генерації бутстреп-даних, яку запропоновано в роботі і виконано на множині введених кваліфікаційних ознак та їх станів.

Згідно цієї систематизації, кожному показнику якості (способу генерації бутстреп-даних) відповідає певна кодова комбінація, позиції якої співвідносяться з введеними кваліфікаційними ознаками, а цифри на цих позиціях визначаються станом ознак, тобто кожна кодова комбінація - компактний запис, інформаційна модель, що характеризує умови розв'язання та особливості конкретної задачі ідентифікації. Оцінив для цієї задачі потрібну сукупність станів кваліфікаційних ознак, за складеними таким чином кодовими комбінаціями обираються тип показника якості ідентифікації та спосіб генерації бутстреп-даних (останній може бути синтезовано згідно його формального опису, зафіксованому у відповідній кодовій комбінації).

Ще одним наслідком розробки бутстреп-технології дослідження якості ідентифікації є стандартизація процедури та засобів цього дослідження, що забезпечує можливість його часткової чи повної автоматизації.

У шостому розділі роботи розглянуто застосування розробленої методології дослідження якості ідентифікації для розв'язання ряду практичних задач, а саме: побудови моделі руху літака у поздовжній площині за результатами його натурних випробувань, ідентифікації моделі динаміки давача перевантаження за даними стендових випробувань, побудови моделі площинного поля випромінювання, утвореного кількома випромінювачами, та ряду інших.

ОСНОВНІ РЕЗУЛЬТАТИ І ВИСНОВКИ

Підсумком дисертаційної роботи є побудова загального підходу до дослідження результатів розв'язання задач ретроспективної

ідентифікації лінійних та нелінійних моделей стаціонарних чи динамічних систем за вибірками їх довільно зашумлених вихідних даних при мінімальній апріорній інформації про властивості останніх, мета якого:

- отримати об'єктивний висновок про якість ідентифікації вибраними користувачем методами структурно-параметричної ідентифікації;

- підвищити якість ідентифікації за рахунок додаткової обробки вже одержаних результатів ідентифікації та наявних вихідних даних.

Головні наукові та практичні результати.

1. Виконано аналіз найбільш поширених показників якості ідентифікації. Показано, що у своїй більшості це різні варіанти середнього ризику, ефективність використання яких забезпечується лише у асимптотичі. При роботі з реальними вихідними даними із скінченим обсягом (особливо малим та середнім) показники цього типу мають значне розсіювання та зміння, що залежать від складності та вигляду моделі, рівня похибок в оброблюваних даних, методу оцінювання, який застосовується, і можуть привести до суперечливих або навіть взаємовиключних висновків.

2. Розроблено методологію оцінювання якості ідентифікації, в рамках якої

- визначено зміст поняття якості ідентифікації (якість розв'язання задачі ідентифікації);

- запропоновано й доведено можливість оцінювання якості ідентифікації рівнем стійкості результатів розв'язання задачі ідентифікації, який визначається величиною варіювання цих результатів при варіюванні вихідних даних;

- побудовано систему показників якості ідентифікації, що обіймає як кількісні оцінки стійкості ідентифікації, так і комбіновані показники, які одночасно враховують стійкісні та апроксимативні характеристики одержаних результатів ідентифікації;

- досліджено властивості введених показників якості, диференційовано їх можливості та межі застосування залежно од особливостей задачі ідентифікації, що розв'язується, визначено умови, за яких можлива математична інтерпретація ряду з цих показників, зокрема показників параметричної стійкості - як точносних характеристик оцінок параметрів, комбінованого показ-

ника - як скорегованої оцінки середнього ризику;

- розроблено принципи варіювання вихідних даних, що забезпечують природньо-випадковий характер варіювання й дозволяють реалізувати варіювання у формі, адекватній застосованому типу показника якості, особливостям поставленої задачі ідентифікації й властивостям вихідних даних;

- розроблено методи генерації реалізацій стаціонарного випадкового процесу й векторної випадкової величини за їх характеристик, поданих у непараметричній формі.

3. Розроблено методи підвищення якості розв'язання задач ідентифікації, які спираються на додаткову апостеріорну обробку результатів багаторазового відтворення на вибірках варіюваних даних розв'язання досліджуваної задачі ідентифікації, що дозволяє:

- зменшити рівень розсіювання оцінок параметрів, отриманих при застосуванні одного метода оцінювання;

- зменшити шляхом спеціальної інтегрованої обробки інформації рівень розсіювання оцінок параметрів, кожна з яких представлена кількома значеннями, обчисленими різними методами ідентифікації;

- визначити рівень і в разі необхідності виконати корегування зміщень оцінок параметрів у випадках ідентифікації моделей за даними, що мають похибки в значеннях незалежних змінних;

- виконати селекцію найкращої моделі на множині моделей-претендентів довільної структури і складності.

4. Реалізована адаптація розроблених загальних методів дослідження задачі ідентифікації для випадку ідентифікації динамічних об'єктів за даними, що подані у вигляді часових рядів. Для цього розроблено додаткові методи підвищення якості параметричної ідентифікації, в яких використано різні форми попереднього згладжування даних (локально-оптимальне згладжування, "перегладжування" даних при ідентифікації лінійних регресійних моделей).

Застосування цих методів на практиці за умов мінімуму апіорної інформації про характеристики нестационарного корисного сигналу та похибки вимірювань обумовило необхідність виконання:

- розробки системи згладжувачів фільтрів з поточною адаптацією їх структури та параметрів до властивостей вихідних даних, зокрема часових рядів, що згладжуються;

- сертифікації лінійних фільтрів, що застосовуються для "перегляджування" даних, яка реалізується з допомогою формуємих для фільтрів математичних моделей, що пов'язують кількісний показник якості згладжування із значеннями факторів, кожен з яких впливає на цю якість, такими як рівень мінливості часового ряду, його складність, максимальні значення його похідних, тощо.

5. Розроблено технологію дослідження розв'язання задач ідентифікації, що дозволяє з сукупності запропонованих у дисертації прийомів та засобів дослідження побудувати, згідно до конкретного варіанта задачі ідентифікації, технологічну послідовність операцій обробки даних, попередньо виконавши вибір показника (показників) якості ідентифікації, адекватного поставленій задачі й властивостям вихідних даних, узгодивши з ним спосіб варіювання даних, визначивши можливість, доцільність й порядок застосування тих чи інших методів підвищення якості ідентифікації.

Виконано формалізацію процедур вибору та узгодження елементів послідовності "мета ідентифікації - показник якості - спосіб варіювання даних - методи підвищення якості" шляхом систематизації показників якості і способів варіювання даними на множині введених кваліфікаційних ознак та їх станів, що дозволяє суттєво спростити реалізацію дослідження якості ідентифікації на практиці й забезпечити, хоча б частково, його автоматизацію проведення.

6. Практичне застосування розробленого підходу до дослідження якості розв'язання реальних задач ідентифікації підтвердило головні результати та висновки, що наведено в дисертації.

Основні результати дисертації опубліковано в роботах:

1. Архипов А. Е. Выбор, оценка качества и методы повышения точности аппроксимативных моделей в задачах идентификации. - К.: УМК ВО, 1992. - 56 с.
2. Архипов А. Е., Бимбас В. А., Синеглазов В. М. Оценивание статических и динамических характеристик воздушных судов по результатам стендовых испытаний. - К.: КИИГА, 1992. - 120 с.

Автору належить розробка методики оцінювання параметрів датчиків перевантаження та кутової швидкості, системи первинної обробки даних, бутстреп-аналізу результатів якості параметричної ідентифікації.

3. Синеглазов В. М., Архипов А. Е., Бимбас В. А. Стендовые испыта-

ния датчиков пилотажно-навигационных комплексов. - М.: Воздушный транспорт, 1993. - 176 с.

Автором запропоновано підхід до оцінки якості результатів ідентифікації шляхом багаторазового моделювання її розв'язання на псевдовиборках, засоби підвищення якості параметричної ідентифікації.

4. Архипов А.Е. Имитационные оценки точности решения задачи параметрической идентификации. Вестник Киев. политехн. ин-та, серия "Техническая кибернетика", N 13, 1989.
5. Архипов А.Е. Построение оптимальных групповых оценок. Сб. "Адаптивные системы автоматического управления". Киев. Техника, N 17, 1989, с.75-78.
6. Архипов А.Е. Система количественной классификации методов сглаживания случайных погрешностей данных // Вестн. Киев. политехн. ин-та. Техн. кибернетика. - 1992. - Вып. 15, с.46-52.
7. Архипов А.Е. Селекция наилучшей модели с помощью комбинированного критерия качества моделирования. Сб. "Адаптивные системы автоматического управления". - Киев: Техника, 1992, N19, с.46-49.
8. Архипов А.Е. Адаптивное сглаживание погрешностей наблюдения нестационарных случайных процессов цифровыми фильтрами. Сб. "Адаптивные системы автоматического управления". Киев: Техника. N 13, 1985. - с.102-108.
9. Архипов А.Е. Комбинированный алгоритм предварительной обработки данных. Вестник Киев. политехн. ин-та, серия "Техн. кибернетика", N 10, 1986, с.6-12.
10. Архипов А.Е. О моделировании некоторых типов случайных последовательностей // Вестн. Киев. политехн. ин-та. Техн. кибернетика. - 1988. - Вып. 12, с.39-44.
11. Silvestrov A.N., Arhipov A.E., Yudin A.Yu: Purposeful Multiple Adaptive System for Identification of Aircraft Aerodynamic Characteristics from Data of Natural Tests/Proceedings of the 1-st Mini Conference on Vehicle Sistem Dynamics Identification and Anomalies. - Budapest, 14-16 November, 1988, p.289-303.

Автору належить методика уточнення оцінок параметрів моделі при використанні ряду методів параметричної ідентифікації.

12. А.с. 1516960 СССР Ст 01 N 29/04. Устройство регистрации сигналов акустической эмиссии / Л.М.Циринская, С.Ф.Филоненко, А.Е.Архипов и др. - Оpubл. 23.10.89. Бл. N 39.

Автором розроблено алгоритм вилучення інформативної складової сигналу акустичної емісії.

13. Архипов А.Е. Принцип локальной оптимальности в задачах обработки данных. - Киев. политехн. ин-т, Киев, 1984. - 17 с. - Рукопись деп. в УкрНИИТИ 26.II.84, N 1936 Ук-84 Деп.
14. Архипов А.Е. Оценивание качества модели по уровню устойчивости результатов идентификации. - Киев. политехн. ин-т, Киев, 1995. - 7 с. - Рукопись деп. в УкрНИИТИ 13.04.95, N 821Ук-95Деп.
15. Архипов А.Е. Оценка качества линейных моделей с помощью комбинированного показателя. - Киев. политехн. ин-т, Киев, 1995. - 8 с. - Рукопись деп. в УкрНИИТИ 13.04.95, N 822Ук-95Деп.
16. Архипов О.Е. Селекція найкращої моделі з допомогою бутстреп-модельовання. / Тез. докл. I-ої Міжнародної науково-технічної конференції "Математичне моделювання в електротехніці й електроенергетиці", м. Львів, 19-22 вересня 1995р..

Handwritten signature: О.Е. Архипов

Архипов А.Е. Бутстреп-анализ качества решения задач идентификации. Рукопись. Диссертация на соискание ученой степени доктора технических наук по специальности 05.13.03 - Системы и процессы управления, Национальный технический университет Украины, Киев, 1995.

Защищается 16 научных работ, в которых рассмотрен подход к исследованию решения задачи идентификации с использованием бутстреп-методов. Цели исследования: 1) получение объективных заключений о качестве идентификации в реальной ситуации; 2) повышение качества идентификации путем дополнительной обработки исходных данных в рамках выбранных методов параметрической или структурно-параметрической идентификации. Решен комплекс задач, таких как разработка системы показателей качества идентификации для исследования пригодности моделей, выбор наилучшего варианта структуры модели, селекция информативных переменных, исключение смещений в оценках параметров для случая наличия погрешностей в независимых переменных, несмещенного оценивания посредством использования специальных типов сглаживающих фильтров и т.п.

Ключові слова: якість ідентифікації, селекція моделей, бутстреп, адаптивне згладжування, моделювання даних.

Arhipov A.E. Bootstrap analysis of solution quality for problems of identification. Manuscript. Doctor of Technical Sciences thesis in speciality - 05.13.03 - Systems and processes of control, National Technical Univesity of Ukraine, Kiev, 1995.

It is defended 16 scientific works in which have been seach the approach to studing of identification problem solution by using bootstrap - technology. Purposes of studing are: 1) to obtain objective conclusions about identification quality in realistic situation; 2) to rise identification quality by means of supplementary data processing in the framework of selected methods of parameter or structure - parameter identification. To attain this purposes have been solved the complex of problems such as: developing of the system of identification quality indicators for studing of the model validity, choice of the best variant of model structure, selection of the informative variables, excluding of the parameter estimates biases in the case when there are errors in regression variables, unbiased estimation by means of using special types of smoothing filters and another.

АВ 33.878

Підп. до друку 20.12.95. Формат 60×84¹/₁₆. Папір
друк. № 2. Друк офсетний. Умовн. друк. арк. 2,0
Умовн. фарбо-відб. 2,0. Облік-вид. арк. 2,0
Тираж 100. Зам. № 5-5366.

Фірма «ВІПОЛ».
252151, Київ, вул. Волинська, 60.