

НОВЫЕ МЕТОДЫ НЕСМЕЩЕННОГО ОЦЕНИВАНИЯ МОДЕЛИ ГАММЕРШТЕЙНА

Аннотация: Рассматриваются новые методы несмещенного параметрического оценивания линейной и непараметрического оценивания нелинейной частей модели Гаммерштейна. Приводятся примеры решения задач для каждого из предлагаемых подходов.

Ключевые слова: аппроксимация, идентификация, модель Гаммерштейна, оптимизация, сходимость.

Введение

Подобно математике, где имеются понятия необходимых и достаточных условий (устойчивости, оптимальности и т.п.), в теории идентификации также можно сформулировать два вида условий [1]:

— необходимые – это задача аппроксимации с точностью до ε “поведения” объекта идентификации моделью произвольной структуры;

— достаточные, когда в дополнение к необходимым требуется выполнение условия близости структуры и параметров модели к структуре и параметрам реально существующего в объекте (в конечной пространственно-временной области) отображения входных переменных на переменные состояния и выходные переменные. Например, определение (диагностика) для заданного режима аэродинамических коэффициентов (АДК) летательных аппаратов (ЛА) по данным летных испытаний (ЛИ); определение балансировочных (статических) нелинейных зависимостей ЛА из неопределенной динамики ЛИ; определение нелинейных тарировочных зависимостей $f(x)$ первичных измерительных преобразователей, физика которых соответствует модели Гаммерштейна [2]

$$\sum \beta_i \frac{d^i y}{dt^i} = f[x(t)], \quad (1)$$

где n , β_i , $f(x)$ неизвестны, $x(t)$ и $y(t)$ измеряются с шумами на ограниченном нестационарностью, стоимостью и другими причинами интервале T времени.

Реальная ситуация, как правило, характеризуется недостаточной статистической представительностью выборки данных, неавтономностью, нелинейностью, нестационарностью объекта идентификации и, как следствие, приближенностью его модели. Стремление повысить адекватность локальных моделей путем сужения пространственно – временной области поведения объекта

приводит к возрастанию в измерениях соотношения “шум – сигнал”. Стремление учесть нелинейность и параметризовать ее резко расширяет размерность вектора β неизвестных коэффициентов. В обоих случаях имеет место некорректность [3] задачи идентификации.

Не претендуя на всеобщность, далее приведено несколько новых подходов, позволяющих в определенной мере обеспечить корректность задачи идентификации для реальных условий.

1. Метод несмещенного оценивания параметров модели при наличии шумов в измерениях входных и выходных сигналов

В ограниченной пространственно-временной области G нелинейная нестационарная динамика реального объекта

$$\dot{\mathbf{x}}^* = f^*(\mathbf{x}^*, t^*), \quad (2)$$

где звездочкой (*) обозначены точные значения, вектор-функция f представима (с точностью до ε^*) линейной моделью

$$\mathbf{y}^* = \mathbf{x}^* \cdot \beta^* + \varepsilon^*, \quad (3)$$

где $\mathbf{y}^* = \dot{\mathbf{x}}^*$, β^* – искомые коэффициенты.

Чем уже область G , тем точнее модель (3), но и тем больше соотношение “шум – сигнал” в измерениях

$$\mathbf{y} = \mathbf{y}^* + N_y; \quad \mathbf{x} = \mathbf{x}^* + N_x.$$

Известны методы конфлюэнтного анализа [4], статистически оптимального нелинейного оценивания (расширенный фильтр Калмана, квазилинеаризация и инвариантное погружение [5]). Однако этим методам необходима информация о вероятностных характеристиках помех и/или они не гарантируют сходимости релаксационного процесса приближения оценок β к β^* .

Простой способ несмещенного оценивания состоит в замене МНК-оценок [5]

$$\hat{\beta} = (\mathbf{x}^T \mathbf{x})^{-1} \mathbf{x}^T \mathbf{y} = A_1^{-1} B_1, \quad (4)$$

которые имеют смещение

$$\Delta \hat{\beta}' = \hat{\beta} - \beta^* = -(A^* + \Delta A_1)^{-1} \Delta A_1 \cdot \beta^*, \quad (5)$$

где $A^* = \mathbf{x}^{*T} \mathbf{x}^*$, $\Delta A_1 = A_1 - A^*$, СМНК - оценками [6]

$$\hat{\beta}_m = A_2^{-1} B_2, \quad (6)$$

полученными из условия минимума функционала

$$J(\varepsilon) = \varepsilon_0^T (\varepsilon_m + \varepsilon_{-m}) + (\varepsilon_m + \varepsilon_{-m})^T \varepsilon_0, \quad (7)$$

где $\varepsilon_0 = \mathbf{y}_0 - \mathbf{x}_0 \hat{\beta}_m$, $\varepsilon_{\pm m} = \mathbf{y}_{\pm m} - \mathbf{x}_{\pm m} \cdot \hat{\beta}_m$; индексы 0 и $\pm m$ означают нулевое или сдвинутый во времени на m отсчетов массив данных.

Если сдвиг m больше времени корреляции помех в \mathbf{x} , но меньше критического $m_{кр}$ [6], когда матрица A_2 вырождена, то оценка (6) будет несмещенной. Чем более гладкой является $\mathbf{x}^*(t)$ и $N_x(t)$ ближе к “белому шуму”, тем лучше оценка (6). Если множество сдвигов m в пределах от 1 до $m_{кр}$ не единично, то с целью улучшения эффективных оценок можно использовать усредненную по m оптимально взвешенную функцией $\eta(m)$ оценку $\hat{\beta}$ интегрального СМНК:

$$\hat{\beta} = \left(\sum_{-m_{кр}}^{m_{кр}} \eta(m) \right)^{-1} \cdot \sum \eta(m) \cdot \hat{\beta}_m, \quad (8)$$

где выбором $\eta(m)$ достигается компромисс между смещением и дисперсией оценки (8). В условиях неопределенности характеристик помех вес $\eta(m)$ оптимизируется по внешнему (проверочному [7], главному [6]) критерию I .

Задачу синтеза можно параметризовать, задав симметричную относительно $m = 0$ финитную функцию такую, что $\eta(0) = \eta(\pm m_{кр}) = 0$. Например,

$$\eta(m) = (1 + |m|)^\theta \cdot \left(1 - \cos \frac{2\pi m}{m_{кр}} \right)^\gamma, \quad (9)$$

где θ и γ – оптимизируемые по I параметры.

$$(\theta^*, \gamma^*) = \arg \operatorname{extr}_{\theta, \gamma} I(\theta, \gamma). \quad (10)$$

Параметр γ влияет на ширину импульсов $\eta(\pm m)$, θ – на несимметрию (рис. 1).

В частности, при $\theta \rightarrow \infty$ оценка (8) равна (4), при $\gamma \rightarrow \infty$ оценка (8) равна (6).

Тестовый пример. Модель:

$$y^*(t) = \beta_1^* x_1^*(t) + \beta_2^* x_2^*(t) + \varepsilon^*(t), \quad \beta_1^* = \beta_2^* = 1;$$

$$R_{x_1^* x_2^*}(\tau) = e^{-k\tau}, \quad k = 0, 1; \quad i = 0, 1; \quad R_{x_i^* x_2^*}(\tau) = 0.5,$$

$\varepsilon^*(t)$ – белый шум с единичной дисперсией. Измерения $x_i(t) = x_i^*(t) + N_{x_i}(t)$, $i = 1, 2$; $y(t) = y^*(t) + N_y(t)$, где помехи $N_{x_i}(t)$ и $N_y(t)$ – взаимно некоррелированные шумы с соответствующими автокорреляциями $R_{x_i}(\tau) = e^{-\tau}$.

На рис. 2 приведены эллипсы рассеивания оценок (8) $\hat{\beta}_1, \hat{\beta}_2$, полученные для 100 выборок по 100 точек и соответствующих рис. 1 значениям θ и γ . На рис. 2а значения $\sigma_{\hat{\beta}}$ и $\Delta \hat{\beta}$ меньше средних, на рис. 2б – средние значения $\sigma_{\hat{\beta}}$ и $\Delta \hat{\beta}$; на рис. 2в – $\min \sigma_{\hat{\beta}}$ и $\max \Delta \hat{\beta}$;

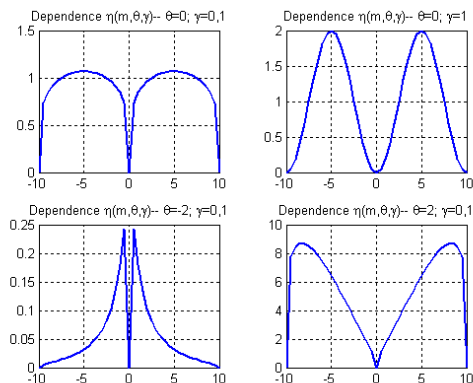


Рис. 1 – Зависимость $\eta(m, \theta, \gamma)$.

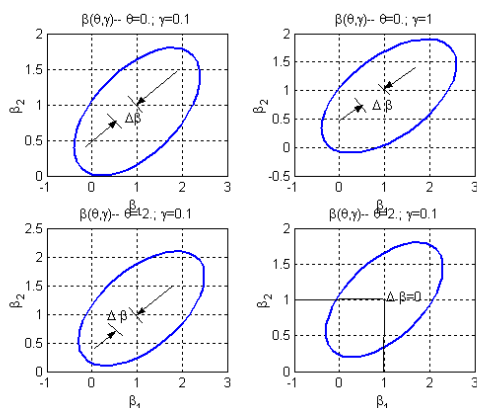


Рис. 2 – Зависимость $\hat{\beta}(\theta, \gamma)$.

на рис. 2г – $\max \sigma_{\hat{\beta}}$ и $\min \Delta \hat{\beta}$. Параметр $m_{кр}$ практически определяется из условия вырожденности матрицы A_2 для сглаженных простейшим фильтром значений $x_i(t)$.

2. Метод корректного оценивания в сколь угодно малой области G фазового пространства X^* вектора линейной составляющей (3) нелинейной модели (2)

Естественная (вследствие конечной мощности реальных процес-сов) гладкость отображения (2) позволяет представить его рядом Тейлора. Для i -й компоненты вектор-функции $y^*(t)$ в отклонениях от центра x_H^* области G , имеем:

$$\Delta y_i^*(k) = \left. \frac{\partial f_i^*}{\partial \mathbf{x}^*} \right|_{\mathbf{x}_H^*} \Delta \mathbf{x}^*(k) + \frac{1}{2} \Delta \mathbf{x}^*(k) \left. \frac{\partial^2 f_i^*}{\partial \mathbf{x}^* \partial \mathbf{x}^{*T}} \right|_{\mathbf{x}_H^*} \Delta \mathbf{x}^{*T}(k) + \dots \quad (11)$$

где k – дискреты времени t : $t_k = k\Delta t$. В векторно-матричном виде имеем:

$$\Delta \mathbf{y}_i^* = \Delta \mathbf{x}^* \beta_i + \Delta \mathbf{x}^* B_i \Delta \mathbf{y}_i^{*T} + \dots \quad (12)$$

$$\beta_i = [\beta_{i1}, \dots, \beta_{in}]^T; \quad i = \overline{1, n}, \quad k = \overline{1, M};$$

$$\Delta \mathbf{y}_i^* = \begin{bmatrix} \Delta y_i^*(1) \\ \dots \\ \Delta y_i^*(M) \end{bmatrix}; \quad B_i = \begin{bmatrix} b_{i11} & \dots & b_{i1n} \\ \dots & \dots & \dots \\ b_{in1} & \dots & b_{inn} \end{bmatrix};$$

$$\Delta X^* = \begin{bmatrix} \Delta x_1^*(1) & \dots & \Delta x_n^*(1) \\ \dots & \dots & \dots \\ \Delta x_1^*(M) & \dots & \Delta x_n^*(M) \end{bmatrix}.$$

Цель идентификации в данном случае состоит в получении достаточно точных оценок вектора β_i в (2) по зашумленным (как и в предыдущем методе) выборкам данных $Y(k)$, $X(k)$. Если область G сузить, растет отношение “шум – сигнал”, если расширить ее, растет размерность задачи (от n для линейной модели (3) до $n + n^2$ для квадратичной (12) и т.д.). И в первом и во втором случаях задача оценивания некорректна [3].

Основываясь на фундаментальном свойстве гладкости процессов, можно показать [1], что смещение $\Delta \beta_i$ вследствие приближенности модели (3) относительно точной модели (2) будет также гладкой (а для квадратичной модели – линейной) функцией показателя величины области G . Так как МНК-оценка $\hat{\beta}_i$ при условии квадратичной нелинейности (12) равняется

$$\hat{\beta}_i = Q \Delta Y_i = Q \Delta X \beta_i + Q (\Delta X \cdot B_i \Delta X^T) + \dots \quad (13)$$

где $Q = (\Delta X^T \Delta X)^{-1} \Delta X^T$.

Смещение

$$\Delta \beta_i = (\Delta X^T \Delta X)^{-1} \Delta X^T \cdot (\Delta X \cdot B_i \Delta X^T) \quad (14)$$

является (вследствие усреднения произведений мгновенных значений нормы переменных) гладкой функцией отклонений $\|\Delta X\|$.

При $\|\Delta X\| \rightarrow 0$ $\hat{\beta}_i \rightarrow \beta_i$, $\Delta \beta_i \rightarrow 0$.

Тестовый пример. Точная модель

$$y(k) = \sum_{j=1}^3 x_j(k) + \sum_{j,q=1, j \geq q}^3 x_j(k) x_q(k) \quad (15)$$

с единичными коэффициентами для четырех выборок аппроксимировалась ее линейной частью при

$$x_1(k) = x_{\max}(l) \sin\left(\pi \frac{k-1}{M-1}\right), \quad x_2(k) = x_{\max}(l) \sin\left(2\pi \frac{k-1}{M-1}\right),$$

$$x_3(k) = x_{\max}(l) \cos\left(2\pi \frac{k-1}{M-1}\right), \quad k = \overline{1, M}, \quad M = 100, \quad i = \overline{1, 4}.$$

Сигналы зашумлены 10% помехой типа “белого шума”. Оценки $\hat{\beta}_j$ ($j = 1, 2, 3$) вычислялись по алгоритму (6) при $m = 1$.

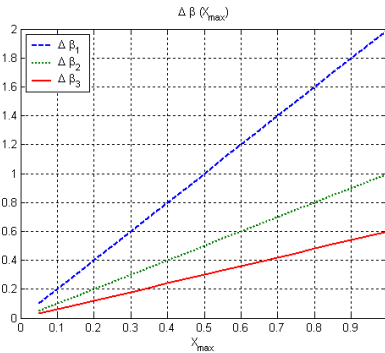


Рис. 3 – Зависимость $\Delta\beta(x_{\max})$.

На рис. 3 приведены линейные зависимости смещений $\Delta\beta_j$ от $x_{\max}(l)$, практически сходящиеся при нулевой амплитуде x_{\max} к нулю; соответственно оценки $\hat{\beta}_j$ сходятся к истинным β_j .

Попытка оценить все 9 коэффициентов модели (5) как по (4), так и по (6), не привела к желаемому результату вследствие вырожденности матриц A_1 и A_2 .

Реальный пример. На рис. 4 приведены графики семи режимов изменения руля высоты $\delta_B(t)$, угла атаки $\alpha(t)$ и угловой скорости $\omega_{Z_1}(t)$ в короткопериодическом продольном движении самолета М-17.

Полная модель зависимости $\dot{\omega}_{Z_1}(t)$ от $\delta_B(t)$, $\alpha(t)$ и $\omega_{Z_1}(t)$ подобна модели (15). В каждом из семи режимов определялись смещенные оценки коэффициентов $\hat{\beta}_j$, $j = 1, 2, 3$, по ним рассчитывался запас $\hat{\sigma}_n$ апериодической устойчивости по вертикальной перегрузке М-17 и аппроксимировался в зависимости от $\|\Delta\alpha\|$ линейной функцией (рис. 5)

$$\hat{\sigma}(\|\Delta\alpha\|) = 0,22 - 0,075 \|\Delta\alpha\|.$$

Прогнозное в $\|\Delta\alpha\| = 0$ значение $\hat{\sigma}_n(0)$ лежит в области истинных значений σ_n . Аналогично определяются все другие показатели и АДК ЛА.

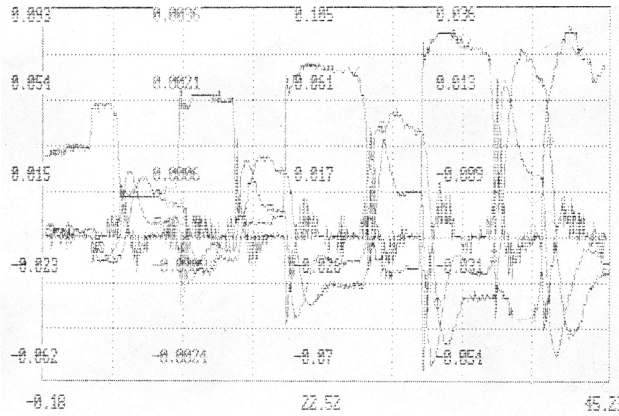


Рис. 4 – Графики режимов изменения руля высоты, угла атаки и угловой скорости.

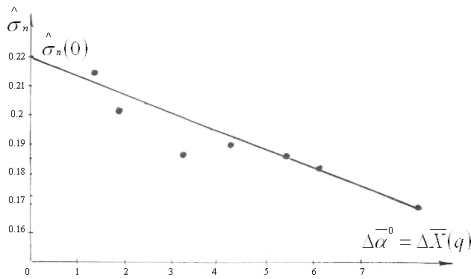


Рис. 5 –

3. Метод непараметрического оценивания нелинейной статической составляющей динамической модели

В полной модели (2) выделим управляющие переменные $U^*(t)$:

$$\dot{X}^* = f^*(X^*, U^*, t). \quad (16)$$

Для реально ограниченных скоростей $\dot{X}^*(t)$ в ограниченной области G изменения $X^*(t)$ с точностью до $\varepsilon^*(t)$ модель (16) представима в виде

$$\dot{X}^*(t) = AX^*(t) + f[U^*(t)] + \varepsilon^*(t) \quad (17)$$

или, относительно i -й составляющей $x_i^*(t)$ вектор-функции $X^*(t)$ в виде модели Гаммерштейна (1):

$$a_n \frac{d^n x_i^*}{dt^n} + a_{n-1} \frac{d^{n-1} x_i^*}{dt^{n-1}} + \dots + x_i^*(t) = f[U^*(t)], \quad (18)$$

где $a_j, j = \overline{1, n}$ и $f(U^*)$ неизвестны.

Задача состоит в определении статической нелинейной зависимости

$$x_i^*(t) = f(U^*) \quad (19)$$

при неизвестных a_j из произвольного динамического режима изменения $X^*(t)$ в G -области.

Произвольность режима не гарантирует линейной независимости переменных при a_j и, как следствие, невырожденности информационной матрицы A^* . Еще большая некорректность возникает в традиционном подходе при аппроксимации $f(U)$ степенным полиномом от U

$$\hat{f}(U) = \sum_{l=0}^m b_l U^l \quad (20)$$

или системой других квазиортогональных функций

$$\hat{f}(U) = \sum_{l=0}^m b_l \varphi_l(U), \quad (21)$$

когда из уравнения (18) в виде

$$\varepsilon(t) = x_i(t) - \sum_{l=0}^m b_l \varphi_l[U(t)] + \sum_{j=1}^n a_j \frac{d^j x_i(t)}{dt^j} \quad (22)$$

при неточных $x_i(t), U(t)$ и их производных необходимо определить $m + n$ неизвестных, где числа m и n также неизвестны. Пусть, к примеру, в (18) линейно независимы только первая и вторая производные от $x^*(t)$. Тогда для преодоления этой проблемы воспользуемся априори известным свойством гладкости зависимости $f(U)$. Будем оценивать коэффициенты a_j уравнения (18) не из условия минимума среднего квадрата ошибки (22), а из условия гладкости зависимости $f(U)$ [8]

$$(a_1, a_2) = \arg \min_{(a_1, a_2)} \sum_{k=1}^N \left(\frac{d^r x_{\text{ск}}(t_k)}{dU^r} \right)^2, \quad (23)$$

где r характеризует степень гладкости $f(U)$;

$$x_{\text{ск}}(t_k) = x(t_k) - a_1 \frac{dx(t_k)}{dt} - a_2 \frac{d^2 x(t_k)}{dt^2}$$

– скомпенсированное по динамике значение $x(t_k)$.

Практически в (23) вместо производных берутся конечные разности соответствующего порядка, которые вычисляются по предварительно сплаженным с помощью сплайнов и упорядоченным по возрастанию U массивам данных.

Тестовый пример. В качестве $f(U)$ взята зависимость

$$f(U) = 4U - 60 \sin(0.065U). \quad (24)$$

Входное воздействие $U(t)$ – последовательность ступенек:

$$U(t) = U_{\max} \left[-1 + \frac{1}{8} \sum_{k=1}^{16} 1(t - k\Delta t) \right].$$

На $x^*(t)$ наложен 20% “белый шум”. Результат совместного МНК-оценивания коэффициентов b_i и a_1, a_2 приведен на рис. 6.

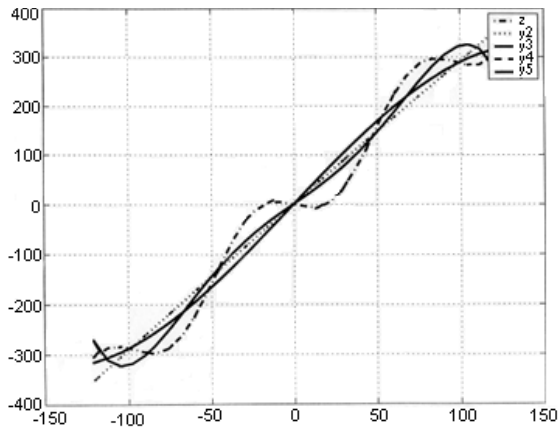


Рис. 6 –

Результат непараметрического оценивания нелинейности (24) из условия (23) гладкости $f(U)$ при $r = 1, 2, 3$ приведен на рис. 7.

Как следует из сопоставления графиков на рисунках непараметрическая оценка $\hat{f}(U)$ с коррекцией динамики из условия (23) существенно ближе к истинной $f(U)$, чем параметрическая МНК-оценка при совместном оценивании коэффициентов линейной и нелинейной части модели Гаммерштейна (18).

4. Метод восстановления многомерных нелинейных зависимостей по экспериментальным данным

Отсутствие информации о структуре нелинейной многомерной зависимости $y(x_1, \dots, x_n)$ не дает возможности применить методику оптимального планирования эксперимента. Как правило, снимаются частные сечения $y(x_i)$ при постоянных x_j ($j = \overline{1, n-1}$,

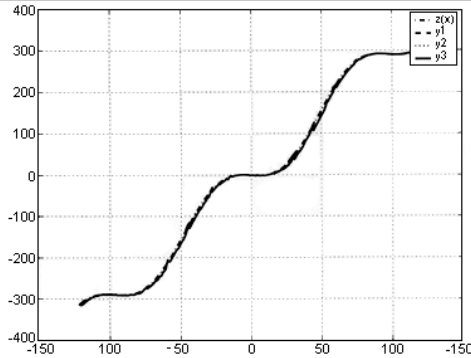


Рис. 7 –

$j \neq i$). Воспользовавшись свойством гладкости $y(x)$, представим эту зависимость кратным рядом Тейлора или его степенным эквивалентом:

$$y(x) = \beta_0 + \sum_{i=1}^n \beta_i x_i + \sum_{i=1}^n \sum_{j=1}^n \beta_{ij} x_i x_j + \dots \quad (25)$$

Для частного k -го сечения $y(x_k)$ при $x_i = \text{const}$, $i = \overline{1, n-1}$, $i \neq k$, из (25) получим одномерную зависимость

$$y(x_k) = \beta_{0k} + \beta_{1k} x_k + \beta_{2k} x_k^2 + \dots \quad (26)$$

Структура зависимости и параметры $\beta_{0k}, \beta_{1k}, \dots$ находится по МНК для (25) при различных, но фиксированных значениях остальных переменных. Далее коэффициенты частных моделей последовательно аппроксимируются как функции этих переменных.

Реальные примеры. В качестве примера исследовались высотно-скоростные характеристики турбовинтового авиадвигателя АИ-20М (зависимость тяги y от высоты x_1 , скорости x_2 и управляющего воздействия x_3).

Математическая модель (сплошные линии) получена из локальных моделей сечений $y(x_k)$ при $x_1 = c_i$, $x_2 = c_j$:

$$y(x_3) = \beta_0'' + \beta_1'' x_3 \quad (27)$$

путем МНК-аппроксимации ее коэффициентов квадратичной моделью по x_1 :

$$\beta_0'' = \beta_{01}' x_2 + \beta_{02}', \quad \beta_1'' = \beta_{11}' x_2 + \beta_{12}' \quad (28)$$

и последующей аппроксимации статически значимых коэффициентов моделей (28) квадратичной моделью по x_1 :

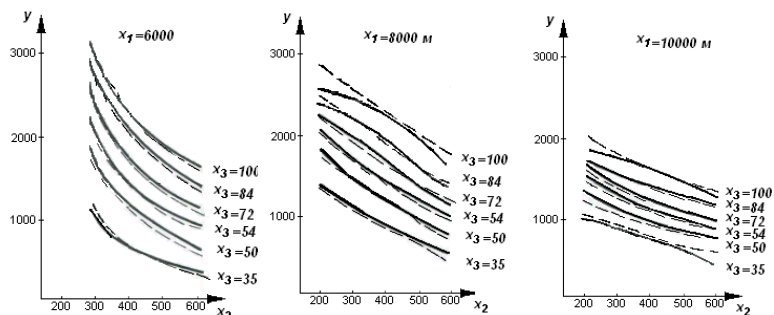


Рис. 8 –

$$\beta'_i = \beta_{i0} + \beta_{i1}x_1 + \beta_{i2}x_2, \quad i = 0, 1.$$

В результате, полная модель приобретает вид

$$y = (\beta_0 + \beta_{01}x_1 + \beta_{02}x_1^2) + (\beta_1 + \beta_{11}x_1 + \beta_{12}x_1^2)x_2 + (\beta_2 + \beta_{21}x_1 + \beta_{22}x_1^2)x_3 + (\beta_3 + \beta_{31}x_1 + \beta_{32}x_1^2)x_2x_3 + (\beta_4 + \beta_{41}x_1 + \beta_{42}x_1^2)x_2^2.$$

Практически совпали (рис. 9) продувочные в аэродинамической трубе данные с моделью для аэродинамической поправки y в функции угла атаки x_1 и положения закрылок x_2 самолета ТУ-144.

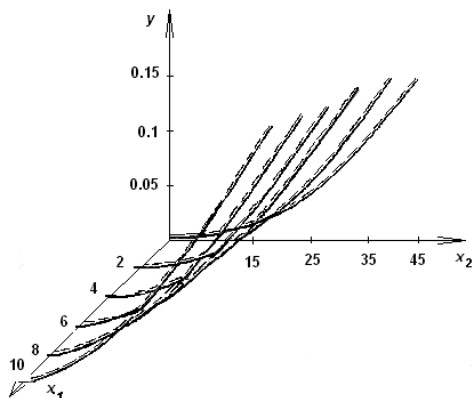


Рис. 9 –

Здесь из локальных моделей

$$y(x_2) = \beta'_{i1}x_2 + \beta'_{i2}x_2^2, \quad x_1 = \text{const}, \quad i = \overline{1, 6} \quad (29)$$

линейной по x_1 аппроксимацией коэффициентов модели (29)

$$\beta''_{i1}(x_1) = \beta_1 + \beta_2x_1; \quad \beta''_{i2}(x_1) = \beta_3 + \beta_4x_1$$

получена полная модель

$$y = \beta_1 x_2 + \beta_2 x_1 x_2 + \beta_3 x_2^2 + \beta_4 x_1 x_2^2.$$

Максимальная ошибка аппроксимации составила 2% от y_{\max} .

Заклучение

В статье на конкретных примерах показана эффективность применения предложенных методов для:

- определения аэродинамических коэффициентов ЛА;
- построения нелинейных балансировочных зависимостей из динамики движения ЛА;
- аналитического описания многомерных нелинейных зависимостей, представленных таблично (результатов продувки в аэродинамической трубе).

Дальнейшее развитие работы требует более строгой математической формулировки;

применения статистически оптимальных решений; создания современных программных продуктов для широкого применения представленных методов.

Библиографический список

1. Два альтернативных подхода к идентификации реальных объектов/А.Н. Сильвестров // Проблемы управления и информатики. –1996, №6.– С.54–65.
2. An iterative method for identification of nonlinear systems using a Hammerstein model./ K.S. Narendra, P.G. Gallman. // IEEE Trans. Automatic Control, vol. AC-11, p. 546, 1966.
3. Методы решения некорректных задач./ А.Н. Тихонов, В.Я. Арсенин // М: Наука, 1979.—286 с.
4. Дисперсионный анализ./Г. Шеффе// М: Физматгиз, 1963. – 625 с.
5. Идентификация систем/ Л. Льюнг.// Теория для пользователя. – М: Наука, 1991. – 432 с.
6. Идентификация и оптимизация автоматических систем/ А.Н. Сильвестров., П.И. Чинаев // М: Энергия, 1983. – 200 с.
7. Долгосрочное прогнозирование и управление сложными системами/А.Г. Ивахненко//К: Техніка, 1975. – 312 с.
8. Модели технологических процессов./ Г.Е. Пухов, Ц.С. Хатияшвили// К: Техніка, 1974. – 223 с.

Отримано 15.02.2013