В.С.Степашко

Теоретические аспекты МГУА как метода индуктивного моделирования*

Исследуется задача построения (структурной идентификации) оптимальной модели в классе структур, линейных по параметрам, на основе заданной короткой выборкой данных. Выбор структуры модели с минимальной дисперсией ошибки прогнозирования, или помехоустойчивой модели, принят в качестве главной цели решения задачи. Исследуются особенности и закономерности построения оптимальной модели в зависимости от уровня шума и длины выборки, изучается эффективность в этой задаче внешних критериев МГУА.

Индуктивный подход к моделированию

Статья О.Г.Ивахненко [1] о МГУА, опубликованная 1968 г., положила начало научному направлению, которое получило название “индуктивная самоорганизация моделей по экспериментальным данным", или просто "индуктивное моделирование".

В этом принципиально новом подходе к моделированию с помощью ЭВМ вместо традиционного *дедуктивного* пути “от общих закономерностей функционирования объекта – к конкретной математической модели” применяется *индуктивный* способ “от конкретных данных наблюдений – к общей модели”: исследователь предоставляет выборку, выдвигает гипотезу о возможном классе моделей и задает критерий выбора наилучшей модели в этом классе. Далее работает компьютер, при этом появляется возможность минимизировать влияние субъективных факторов и получить модель как объективный результат.

МГУА как воплощение индуктивного подхода – это оригинальный метод построения моделей по экспериментальным данным в условиях неопределенности. Полученные по этому методу модели оптимальной сложности отображают неизвестные закономерности функционирования исследуемого объекта (процесса), информация о которых неявно содержится в выборке данных. В МГУА для построения моделей применяются принципы автоматической генерации вариантов, неокончательных решений и последовательной селекции лучших моделей по внешним критериям. Такие критерии основаны на делении выборки на части, при этом оценивание параметров и проверка качества моделей выполняются на разных подвыборках. Это разрешает обойтись без отягощающих априорных предположений, поскольку деление выборки разрешает автоматически (неявно) учесть разные виды априорной неопределенности. Эффективность метода подтверждена решением многочисленных реальных задач в экологии, гидрометеорологии, экономике, технике [2,3].

В [4] была установлена качественная аналогия между задачами моделирования по зашумленным данным и прохождения сигнала через канал с шумом. Оказалось, что задачу моделирования по данным наблюдений, или структурной идентификации сложных объектов, в случае стохастических предположений можно рассматривать как задачу выделения сигнала на

---

* Доложено на I Международной конференции по индуктивному моделированию, Львов, 20-25 мая 2002 г.

фоне помех. На основе этого вывода и идей теории помехоустойчивого приема сигналов Котельникова в [5] предложен новый подход к теоретическому анализу задач моделирования. Это разрешило начать построение теории помехоустойчивого моделирования, основные принципы которой изложены в [6]. Главный результат теории: сложность оптимальной прогнозирующей модели зависит от уровня шума в данных – чем он выше, тем более простой должна быть оптимальная модель. С того времени началось развитие теории МГУА как индуктивного метода извлечения информации и знаний из экспериментальных данных.

Постановка задачи

Рассматривается класс задач, которые сводятся к выбору наилучшей в определенном смысле модели объекта из множества генерированных моделей, линейных по параметрам.

Пусть задана выборка $W=(X \mid y)$ данных в виде $n\times m$ матрицы X и $n\times 1$ вектора y, что содержат информацию n наблюдений за m независимыми входами и одним выходом статического объекта. Матрица X считается детерминированной полного ранга $rank\ X = m$, а вектор y содержит шум:

$$y = \overset{o}{y} + \xi, \quad \overset{o}{y} = Ey = X\theta_0, \quad E\xi = 0, \quad E\xi\xi^T = \sigma^2 I_n \ , \tag{1}$$

где $\overset{o}{y}$ – точный (незашумленный) выход объекта, θ_0 – точный (неизвестный) вектор параметров объекта, ξ – вектор шума с независимыми, одинаково распределенными компонентами, E – символ математического ожидания по всем возможным реализациям вектора шума, I_n – единичная $n\times n$ матрица, σ^2 – неизвестная конечная дисперсия шума.

Считаем, что неизвестная истинная модель (зависимость между $\overset{o}{y}$ и X) принадлежит к классу моделей, линейных по параметрам. Задача состоит в выборе наилучшей по заданному критерию модели из некоторого их множества, сформированного определенным генератором моделей в указанном классе.

В простейшем случае множество $\Im_m$ моделей, сравниваемых в процессе моделирования, содержит m “вложенных” структур, получаемых последовательным усложнением на один линейный член:

$$\Im_m : \quad \overset{o}{y} - X_s\hat{\theta}_s \ , s = \overline{1,m} \ , \tag{2}$$

где s – число оцениваемых параметров, X_s – подматрица из s произвольных (например, первых) столбцов матрицы X, $\hat{\theta}_s$ – соответствующая МНК-оценка по выборке W. Выражение (2) характеризует произвольный путь последовательного усложнения модели в случае полного перебора генерированных вариантов, тому рассмотрение лишь множества $\Im_m$ не влияет на общность анализа задачи.

Целью задачи моделирования будем считать выбор теоретически наиболее *помехоустойчивой модели* [5,6], то есть модели с минимальной дисперсией ошибки, определяемой выражением:

$$J(s) = \mathrm{E}\left\| \overset{o}{y} - \hat{y}_s \right\|^2 = E\left\| \overset{o}{y} - X_s \hat{\theta}_s \right\|^2 . \tag{3}$$

Этот “идеальный” критерий характеризует качество восстановления истинного сигнала $\overset{o}{y}$ по зашумленнному y с помощью оценки $\hat{y}_s$ и равняется суммарной *дисперсии ошибки восстановления* сигнала в n точках.

Разобьем, как это обычно делается в алгоритмах МГУА, выборку W на две не пересекающиеся подвыборки A и B такие, что : $W = (A^T B^T)^T$, $A = (X_A, y_A)$, $B = (X_B, y_B)$, $n = n_A + n_B$. По аналогии с (3), качество прогнозирования в точках X_B подвыборки B по модели сложности s, коэффициенты которой $\hat{\theta}_{As}$ оценены за МНК на подвыборке A, естественно характеризовать величиной ожидаемых потерь в области прогноза X_B , или величиной суммарной *дисперсии ошибки прогнозирования*:

$$J_B(s) \overset{\Delta}{=} J_B(s, X_A, X_B) = \mathrm{E}\left\| \overset{o}{y}_B - \hat{y}_{Bs} \right\|^2 = \mathrm{E}\left\| \overset{o}{y}_B - X_{Bs}\hat{\theta}_{As} \right\|^2 , \tag{4}$$

Очевидно, что в случае $X_B = X$ (4) превращается в (3). Укажем, что критерии (3) и (4) содержат точные (не измеренные) векторы выхода объекта, то есть они отображают “идеальную” цель моделирования. На практике надо применять некоторые их оценки, о чем будет идти речь дальше. Здесь же исследуются свойства критериев (3), (4) как функций сложности моделей и показателей неполноты информации.

Структуры моделей, которые имеют сложность (число оцениваемых параметров)

$$s^o = \arg \min_{s=1,m} J(s) , \quad s_B^o = \arg \min_{s=1,m} J_B(s) , \tag{5}$$

будем называть соответственно *J-оптимальной* (оптимальной *восстанавливающей*) и *J_B-оптимальной* (оптимальной *прогнозирующей*) структурами множества $\Im_m$.

Свойства дисперсии ошибки модели

Рассмотрим свойства $J(s)$ и $J_B(s)$ как функций дискретного аргумента s . Из (3) следует, что

$$J(s) = J^b(s) + J^v(s) = \left\| \overset{o}{y} - X_s \bar{\theta}_s \right\|^2 + \sigma^2 s , \tag{6}$$

где $\bar{\theta}_s = E[\hat{\theta}_s]$, индексом "$b$" (от “bias”) обозначена величина средних потерь $J^b(s)$ от смещения структуры модели, а индексом "v" (от ”variation”) – величина вариации функционала, или средние дополнительные потери $J^v(s)$ вследствие наличия шума. Компоненту $J^b(s)$ будем называть "*структурная* составляющая", а $J^v(s)$ – "*шумовая* составляющая" функционала $J(s)$.

В свою очередь, формула (5) приобретает вид [4]:

$$J_B(s) = J_B^b(s) + J_B^v(s) = \left\| \overset{o}{y}_B - X_{Bs}\bar{\theta}_{As} \right\|^2 + \sigma^2 tr\left[\left(X_s^T X_s \right)^{-1} X_{Bs}^T X_{Bs} \right], \qquad (7)$$

где $\bar{\theta}_{As} = \mathrm{E}[\hat{\theta}_{As}]$, а через $J_B^b(s)$ и $J_B^v(s)$ обозначены соответственно структурная и шумовая составляющие. Но характер их зависимости от s не очевиден, поэтому следует перейти к рекуррентным по s соотношениям.

Приняв, что $X_{s+1} = (X_s \mid x)$, $\bar{\theta}_{s+1} = (\bar{\theta}_s^T \bar{\vartheta}_{s+1})^T$, несложно получить рекуррентные выражения [3]:

$$J^b(s+1) = J^b(s) - \bar{\vartheta}_{s+1}^2 \beta_{s+1} = J^b(s) - \alpha_{s+1}^2 / \beta_{s+1},$$

$$J^v(s+1) = J^v(s) + \sigma^2,$$

или вообще

$$J(s+1) = J(s) - \bar{\vartheta}_{s+1}^2 \beta_{s+1} + \sigma^2, \qquad (8)$$

где введены такие обозначения:

$$\bar{\vartheta}_{s+1} = \alpha_{s+1} / \beta_{s+1}, \quad \alpha_{s+1} = x^T D_s \overset{o}{y}, \quad \beta_{s+1} = x^T D_s x, \qquad (9)$$

причем $D_s = I_n - X_s (X_s^T X_s)^{-1} X_s^T$ – идемпотентная матрица: $D_s^2 = D_s$, и $\beta_{s+1} > 0$ для всех $s = 0,1,...,m-1$. Видим, что $J^b(s)$ как функция целочисленного аргумента s строго монотонно убывает к нулю, а $J^v(s)$ линейно возрастает с увеличением s, поэтому $J(s)$ всегда имеет минимум в некоторой точке $s^o \in [1, m]$. В случае *увеличения* уровня (дисперсии) шума σ^2 значение s^o *уменьшается*, то есть оптимальная структура упрощается [5,6]. Эта закономерность помехоустойчивого моделирования отражена на рис.1.

Соответствующие формулы для $J_B^b(s)$ и $J_B^v(s)$ получен в [7] с учетом того, что $X_{A,s+1} = (X_{As} \mid x_A)$, $X_{B,s+1} = (X_{Bs} \mid x_B)$. Для $J_B^b(s+1)$ имеем соотношения:

$$J_B^b(s+1) = J_B^b(s) - 2\bar{\theta}_{A,s+1} a_B^T b_B + \bar{\vartheta}_{A,s+1}^2 b_B^T b_B, \qquad (10)$$

где введены обозначения:

$$a_B \overset{\Delta}{=} a_B(s) = \overset{o}{y}_B - X_{Bs}\bar{\theta}_{As}, \quad b_B \overset{\Delta}{=} b_B(s) = x_B - X_{Bs}(X_{As}^T X_{As})^{-1} X_{As}^T x_A. \qquad (11)$$

Рекуррентные выражения для $J_B^v(s)$:

$$J_B^v(s+1) = J_B^v(s) + \sigma^2 \, b_B^T b_B \big/ \beta_{A,s+1}, \qquad (12)$$

Тогда вообще для дисперсии ошибки прогнозирования имеем:

$$J_B(s+1) = J_B(s) - 2\bar{\vartheta}_{A,s+1} a_B^T b_B + \bar{\vartheta}_{A,s+1}^2 b_B^T b_B + \sigma^2 \frac{b_B^T b_B}{\beta_{A,s+1}}. \qquad (13)$$

Укажем, что в случае $X_B = X$ с (13) получаем (8).

Итак, в соответствии с (10), (12) составляющие $J_B^b(s)$ и $J_B^v(s)$ как функции s имеют такие свойства: структурная составляющая в общем случае немонотонно убывает от значения

$J_B^b(0) = \overset{o}{y}{}_B^T \overset{o}{y}_B$ к нулю, а шумовая строго монотонно возрастает. То есть функция $J_B(s) = J_B^b(s) + J_B^v(s)$, как и $J(s)$, имеет минимум (глобальный) внутри интервала изменения дискретного аргумента $s \in [1, m]$, и этот минимум отвечает тем меньшей оптимальной сложности s_B^o модели, чем большая дисперсия шума σ^2. Таким образом, поведение функции $J_B(s)$ качественно характеризует тот же рис.1.

Метод критических дисперсий

Графики на рис.1 дают наглядную качественную характеристику принципа помехоустойчивого моделирования, но не разрешают получить количественные оценки условий оптимальности той или иной структуры модели. Поэтому целесообразно перейти к значительно более информативному изображению зависимости дисперсии ошибки модели от дисперсии шума σ^2 для разных структур моделей (разных s). Такая зависимость, как видно из (6), (7), линейна, что упрощает построение соответствующих графиков.

Например, при построении зависимостей для $J(\sigma^2 \mid s)$ на оси ординат графика (см. рис.2) следует сначала отложить ординаты структурной составляющей , а потом провести прямые линии, отвечающие разным s. Наклон этих прямых постепенно увеличивается от нулевого при s = 0 до максимального при s=m. Очевидно, что все прямые, отвечающие сложностям , будут проходить через начало координат. На рис.2 прямая, соответствующая истинной структуре , изображена утолщенной линией, а пунктирной – прямая для s = 4. Укажем, что случай s = 0 соответствует отсутствию модели, или модели с нулевыми коэффициентами, то есть использованию заданного вектора измерений y в качестве “модели”, поэтому на рис.1 все графики выходят из одной точки:

.

Графики на рис.2 дают принципиально новую информацию: положение любой прямой и точки их пересечения однозначно определяются свойствами моделированного объекта, точнее, данными выборки, и не зависят от характеристик шума. При этом огибающая снизу этих графиков (обозначенная пунктирными линиями со стрелками) является экстремалью критерия $J(s)$, то есть каждая точка на ней равняется минимальному значению критерия при соответствующей дисперсии шума. Это означает, что такая огибающая является линией “переключения” сложности оптимальных структур: в случае увеличения дисперсии шума (начиная с) сначала оптимальной будет истинная структура (здесь), после достижения первой точки переключения – структура на единицу меньшей сложности (s = 2), и т.д., до s = 0.

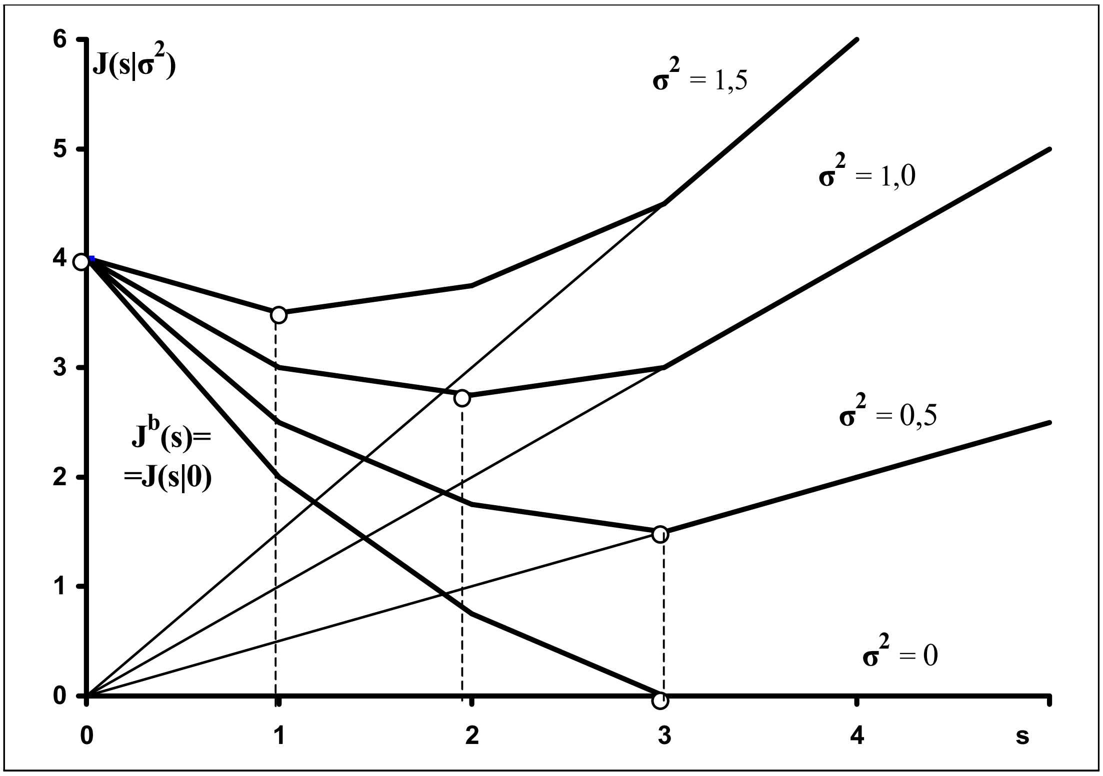

Рис.1. Дисперсия ошибки модели как функция сложности структуры ***s*** при различных дисперсиях шума (кружечками указаны положения минимумов критерия $J(s)$, истинная структура соответствует сложности $\boldsymbol{s}_0 = 3$, общее число регрессоров $m = 5$)

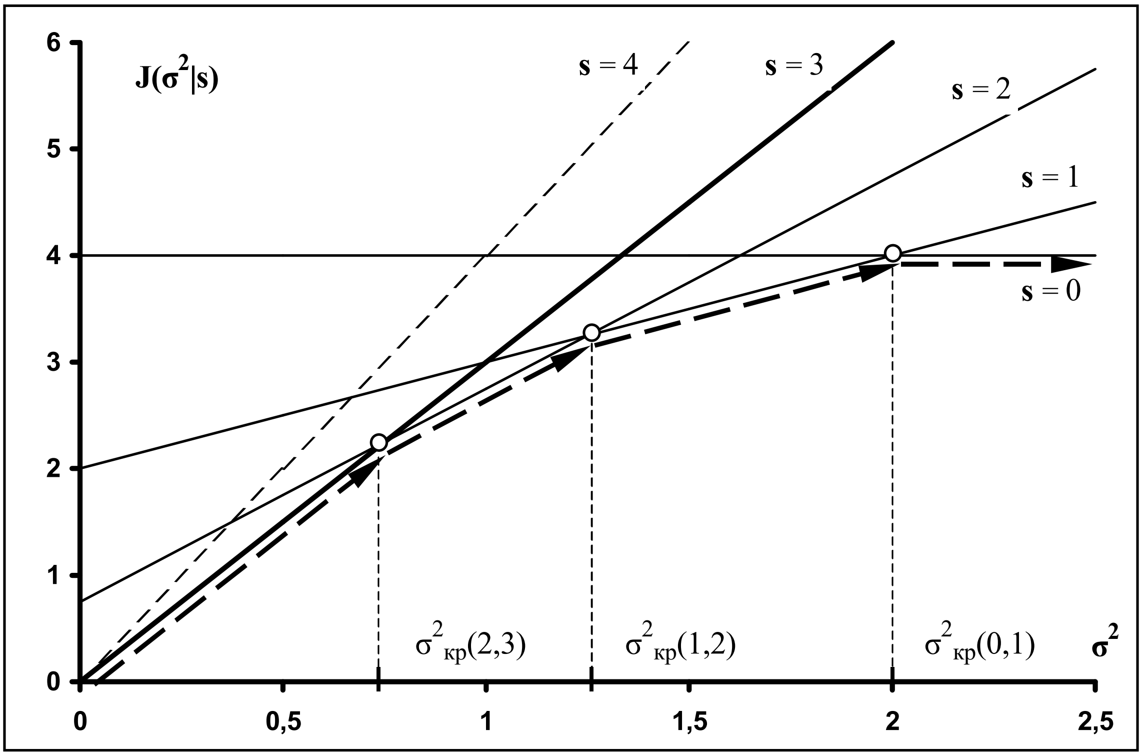

Рис.2. Дисперсия ошибки модели как функция дисперсии шума $\boldsymbol{\sigma}^2$ для структур разной сложности (кружочками указаны критические точки изменения структуры модели, которым соответствуют критические дисперсии $\sigma^2_{\text{кр}}(s,s+1)$; стрелками обозначена *экстремаль* критерия $J(s)$)

Поскольку точки переключения являются неподвижными *критическими* точками, характерными для заданной выборки, естественно ввести так называемые ***критические дисперсии*** [3] для обозначения координат этих точек. Как видно из рис.2, они определяются как решение относительно σ^2 уравнений

$$\sigma_{kp}^2(s,s+1)\ :\quad J(\sigma^2\mid s)=J(\sigma^2\mid s+1),\ s=1,2,\ldots,m-1, \tag{14}$$

которые отвечают необходимому условию экстремума функции дискретного аргумента. При этом понятно, что для заданной дисперсии шума оптимальной будет структура s, для которой выполняется неравенство

$$\sigma_{kp}^2(s-1,s)>\sigma^2\geq\sigma_{kp}^2(s,s+1)\ , \tag{15}$$

и наоборот – структура сложности s будет оптимальной для σ^2, которая будет удовлетворять этому неравенству.

Дисперсия ошибки прогнозирования $J_B(s)$ также линейно зависит от σ^2, поэтому рис.2 актуален и в этом случае, то есть можно означить критические дисперсии $\sigma_{Bkp}^2(s,s+1)$ из условия, подобной к (14). Эти критические дисперсии получаем из рекуррентных соотношений (8), (13) (в обозначениях (9), (11)):

$$\sigma_{kp}^2(s,s+1)=\overline{\vartheta}_{s+1}^2\beta_{s+1}, \tag{16}$$

$$\sigma_{Bkp}^2(s,s+1)=2\overline{\vartheta}_{A,s+1}\beta_{A,s+1}\frac{a_B^Tb_B}{b_B^Tb_B}-\overline{\vartheta}_{A,s+1}^2\beta_{A,s+1}. \tag{17}$$

Совокупность процедур по вычислению критических дисперсий, анализу их свойств и интерпретации полученных результатов в задаче оптимизации структур моделей называется ***методом критических дисперсий***, который лежит в основе теории помехоустойчивого моделирования, включая МГУА.

Приведем общую характеристику связи основных свойств величин критических дисперсий с поведением оптимальных значений s^o, s_B^o, то есть с экстремальными свойствами критериев *J(s)*, *$J_B(s)$*: а) структурная составляющая критерия строго *монотонно* убывает, если для всех $s=\overline{1,m-1}$ критические дисперсии строго *положительны*; б) критерий имеет *единственный минимум* только тогда, когда все значения критических дисперсий строго *монотонно уменьшаются*:

$$\sigma_{kp}^2(0,1)>\sigma_{kp}^2(1,2)>\ldots>\sigma_{kp}^2(m-1,m)\ ; \tag{18}$$

в) при этом *глобальный* минимум критерия определяется условием (15); г) если (18) не выполняется, то (15) является условием *локального* минимума, и для определения оптимальной сложности необходимо найти минимальное значение дисперсии ошибки модели среди всех точек локального минимума.

Методика анализа реальных критериев выбора моделей

Эффективность заданного критерия $CR(s)$ будем анализировать с точки зрения того, как оптимальная сложность структуры модели, обозначенная как

$$s^*_{CR} = \arg\min_{s=\overline{1,m}} \mathrm{E}[CR(s)] \ , \tag{19}$$

оценивает неизвестное значение s^o_B, которое отвечает минимуму теоретического критерия.

*Определение***.** Критерий $CR(s)$ называется *оптимальным* для оценивания моделей с минимальной дисперсией ошибки, если $s^*_{CR} = s^o_B$ (несмещенное оценивание s^o_B), *адекватным* в случае $s^*_{CR} \le s^o_B$ (помехоустойчивое оценивание s^o_B) и *неадекватным,* если $s^*_{CR} > s^o_B$.

Учитывая это, эффективность заданного критерия целесообразно анализировать в такой последовательности: найти его математическое ожидание; проверить компромиссный характер структурной и шумовой составляющих полученного выражения; определить условия оптимальности или адекватности критерия. Последняя задача решается формализовано с помощью метода критических дисперсий.

Величина критической дисперсии $\sigma^2_{CR}(s,s+1)$ как теоретическая мера различимости моделей сложности s и s+1 по критерию $CR(s)$ равняется такому значению σ^2, которое является решением уравнения

$$\mathrm{E}[CR(s,\sigma^2)] = \mathrm{E}[CR(s+1,\sigma^2)] \ . \tag{20}$$

Это разрешает анализировать эффективность заданного $CR(s)$, основываясь на таком результате [8].

Утверждение. Критерий $CR(s)$ является оптимальным ли адекватным, если выполняются условия $\sigma^2_{CR}(s,s+1) = \sigma^2_{Bkp}(s,s+1)$ или $\sigma^2_{CR}(s,s+1) \le \sigma^2_{Bkp}(s,s+1)$ соответственно; иначе он является неадекватным.

Это утверждение дает необходимые и достаточные условия оптимальности или адекватности критериев.

Анализ эффективности критериев МГУА

Рассмотрим применение изложенной методики для анализа эффективности одного из основных критериев МГУА, а именно чаще всего применяемого критерия регулярности $AR(s)$:

$$AR(s) = \left\| y_B - \hat{y}_{Bs} \right\|^2 = \left\| y_B - X_{Bs}\hat{\theta}_{As} \right\|^2 \ . \tag{21}$$

По аналогии с (16), несложно вычислить математическое ожидание этого критерия:

$$\overline{AR}(s) = \mathrm{E}[AR(s)] = J_B(s) + \sigma^2 n_B \ , \tag{22}$$

причем здесь $J_B(s)$ равняется дисперсии ошибки прогнозирования (7).

Из этого сразу следует, что критерий $AR(s)$ обладает свойством помехоустойчивости (сдвига минимума при увеличении дисперсии шума σ^2) как необходимым свойством эффективного критерия. Теперь следует проверить оптимальность или адекватность исследуемого критерия регулярности.

Критическая дисперсия для критерия $AR(s)$ совпадает с выражением (17):

$$\sigma_{AR}^2(s,s+1) \equiv \sigma_{Bkp}^2(s,s+1) = 2\overline{\omega}_{A,s+1}\beta_{A,s+1}\frac{a_B^T b_B}{b_B^T b_B} - \overline{\omega}_{A,s+1}^2\beta_{A,s+1} , \qquad (23)$$

то есть критерий регулярности является *оптимальным* для построения прогнозирующей модели.

Аналогично метод критических дисперсий можно применить для аналитической проверки пригодности любого другого критерия МГУА в качестве вычислимой оценки теоретического критерия для задачи структурной идентификации прогнозирующих моделей. Итак, внешние критерии МГУА являются адекватными задачи построения моделей с минимальной дисперсией ошибки прогноза.

Прикладные задачи теории МГУА

Метод критических дисперсий может помочь в решении ряда актуальных задач теории МГУА: разбиение выборки, сравнительный анализ критериев, асимптотический анализ и т.п.. Рассмотрим коротко некоторые из полученных результатов в этих задачах.

Задача разбивки выборки. Назовем *оптимальным* [7] такое разбиение выборки W, при котором в результате идентификации по условию минимума дисперсии ошибки модели, коэффициенты которой вычисляются на A, на любой из подвыборок A и B будет получена одна и та же структура: $s_A^o = s_B^o$. Здесь s_A^o отвечает минимуму ошибки восстановление сигнала на A:

$$J_A(s) = \mathrm{E}\left\| \overset{o}{y}_A - X_{As}\hat{\theta}_{As} \right\|^2 , \quad s_A^o = \arg\min_{s=1,m} J_A(s) , \qquad (24)$$

а s_B^o – та же оптимальная структура, что и в (5). Как показано в [7], строгое решение такой задачи предусматривает планирование эксперимента для достижения пропорциональности информационных матриц:

$$\rho_B^2 X_A^T X_A = X_B^T X_B , \quad \rho_B^2 \neq 0, \qquad (25)$$

Тогда в условиях пассивного эксперимента задача нахождения наилучшего разбиения матрицы X на две подматрицы сводится к минимизации некоторой меры рассогласования левой и правой частей в (2).

Задача сравнительного анализа критериев. Можно показать, что при условии (25) выполняется равенство $\overline{\omega}_{A,s+1} = \overline{\omega}_{B,s+1} = \overline{\omega}_{s+1}$, которое дает основание для сравнения критериев МГУА в качестве оценки теоретического критерия (3). Например, для критерия регулярности (21) имеем

$$\sigma_{AR}^2(s,s+1) = \frac{1}{1+\rho_B^2}\sigma_{kp}^2(s,s+1) \ , \tag{26}$$

что свидетельствует про его адекватность за любых $\rho_B^2 \neq 0$. Симметричный критерий регулярности [6] также оказывается адекватным задаче восстановления сигнала на W, поскольку

$$\sigma_{AD}^2(s,s+1) = \frac{1}{\rho_B^2 + \frac{1}{\rho_B^2}}\sigma_{kp}^2(s,s+1) \ . \tag{27}$$

Что касается критериев несмещенности $CB(s)$ и вариативности $CV(s)$, которые принадлежат к группе критериев согласованности [6], то в условиях разбиения (25) они неработоспособны, поскольку вследствие $\overline{\omega}_{A,s+1} = \overline{\omega}_{B,s+1} = \overline{\omega}_{s+1}$ структурные составляющие этих критериев тождественно равны 0, и соответственно $\sigma_{CB}^2(\cdot) = \sigma_{CV}^2(\cdot) = 0$, то есть при этом они выбирают простейшие модели. Поэтому для их использования следует рекомендовать выбор разбиения по условию максимального рассогласования выражения (25), при этом критерии этой группы становятся адекватными. Совсем разное поведение критериев регулярности и несмещенности при условии (25) свидетельствует о целесообразности двухкритериального подхода [6] к задаче индуктивного моделирования.

Примеры применение метода критических дисперсий для аналитического сравнения критерия регулярности с другими известными критериями (Маллоуза C_p и Акаіке FPE) можно найти в [8].

Асимптотический анализ. Если ввести усредненные по числу наблюдений значения критериев, например, $\bar{J}(s) = \frac{1}{n}J(n,s)$, и соответствующую оптимальную сложность $\bar{s}^o$, то можно определить понятия *асимптотической несмещенности* оптимальной структуры через сходимость оптимальных значений критерия:

$$\bar{s}^o \to s_o \ : \ \lim_{n\to\infty} \bar{J}(\bar{s}^o) = \lim_{n\to\infty} \bar{J}(s_o) \ . \tag{28}$$

При этом можно доказать, что необходимым и достаточным условием асимптотической несмещенности структуры, оптимальной по заданному критерию, является расходимость критической дисперсии этого критерия:

$$\sigma_{kp}^2(n,s,s+1) \xrightarrow[n]{} \infty \ . \tag{29}$$

При выполнении традиционного для асимптотического анализа условия сильной регулярности регрессоров

$$\lim_{n\to\infty} \frac{1}{n} X_n^T X_n = \overline{H}, \tag{30}$$

где $\overline{H}$ - невырожденная конечная $m \times m$ матрица, критические дисперсии как для идеального критерия (3), так и для внешних критериев МГУА, являются расходящимися, то есть этом критерии асимптотически оптимальны.

Моделирование в неточном классе структур. На практике при моделировании сложных процессов и систем не всегда удается выбрать класс структур моделей, который содержит истинную структуру. При таких условиях задача моделирования решается как задача выбора наилучшего приближения к неизвестной истинной модели. Что изменяется при этом в постановке задачи и в полученных результатах?

Изложенная выше теория минимизации дисперсии ошибки восстановления (3) и прогнозирования (7) сигналов в целом остается действующей, как и формулы для внешних критериев типа (21) и выводы теории об упрощении оптимальной структуры при возрастании уровня шума. Основные коррективы при таких условиях сводятся к тому, что структурные составляющие критериев как функции сложности s на заданном классе структур нигде не равны нулю. Это означает, что на рис.1 кривая $J^b(s)$ не касается оси абсцисс, и соответственно на рис.2 ни одна из прямых не проходит через начало координат. При этом все комментарии к этим графикам качественно остаются такими же, но за счет фактического возрастания значений ординат структурной составной положение минимумов всех кривых на рис.1, а также критические точки на рис.2, смещаются влево. Таким образом, переход к задаче моделирования в неточном классе структур моделей вносит неопределенность, эквивалентную внесению определенного дополнительного уровня шума в условия задачи.

Укажем, что при неточном классе структур корректируется и условие оптимального разбиения выборки: в формулу (25) вместо матрицы X следует подставлять расширенную матрицу $W=(X \mid y)$, тогда все следствия этого условия квадратичной зависимости двух подвыборок остаются в силе.

Выводы

В рамках теории помехоустойчивого моделирования, начиная с работ [4-6], развиваются основы теории МГУА как индуктивного метода построения моделей с минимальной дисперсией ошибки восстановления и прогнозирования точного (незашумленного) сигнала. Аналитическим аппаратом этой теории является метод критических дисперсий, который позволяет детально исследовать закономерности изменения сложности оптимальных структур в зависимости от уровня шума и других показателей неполноты априорной информации, а также теоретически получать сравнительные оценки эффективности критериев МГУА и других методов. В частности, установлены условия оптимальности и адекватности для внешних критериев типа "перекрестного подтверждения", которые используются в МГУА и основаны на разбиении выборки. За счет разбиения эти критерии имеют свойство неявного (автоматического) согласования сложности оптимальной модели с уровнем неопределенности в данных наблюдений, что позволяет при моделировании обойтись без отягощающих априорных предположений, необходимых для применения других методов построения моделей по экспериментальным данным.

Литература

1. *Івахненко О.Г.* Метод групового урахування аргументів – конкурент методу стохастичної апроксимації // Автоматика. – 1968. – №3. – С.58-72.
2. *Ивахненко А.Г.* Долгосрочное прогнозирование и управление сложными системами. - Киев: Техніка, 1975. – 311 с.
3. *Ивахненко А.Г.* Индуктивные методы самоорганизации моделей сложных систем. – Киев: Наук. мысль, 1982. – 296 с.
4. *Ивахненко А.Г., Карпинский А.М.* Самоорганизация моделей на ЭВМ в терминах общей теории связи (теории информации) // Автоматика. – 1982. – № 4. – С.7-26.
5. *Степашко В.С.* Потенциальная помехоустойчивость моделирования по комбинаторному алгоритма МГУА без использования информации о помехах // Автоматика. – 1983. – № 3. – С.18-27.
6. *Ивахненко А.Г., Степашко В.С.* Помехоустойчивость моделирования. – Киев: Наук. думка, 1985. – 216 с.
7. *Степашко В.С.* Структурная идентификация прогнозирующих моделей в условиях планируемого эксперимента // Автоматика. – 1992.– №1. – С.26-35.
8. *Степашко В.С.* Анализ эффективности критериев структурной идентификации прогнозирующих моделей // Проблемы управления и информатики. – 1994. – № 3-4. – С.13-22.