

# МАТЕМАТИЧЕСКИЙ АППАРАТ И ПРОГРАММНАЯ РЕАЛИЗАЦИЯ ПОСТРОЕНИЯ АДЕКВАТНОЙ МАТЕМАТИЧЕСКОЙ МОДЕЛИ

## MATHEMATICAL APPARATUS AND SOFTWARE IMPLEMENTATION OF THE CONSTRUCTION OF AN ADEQUATE MATHEMATICAL MODEL

V. Amosov  
A. Petrov

*Summary.* A technique is proposed for constructing an adequate mathematical model from an evaluation function. Characteristics are introduced that determine the degree of acceptability of the model: the acceptable discrepancy between the results of modeling and experiment, the required number of states of the object of study, thresholds of significance of independent variables.

A prototype of the software package for constructing an adequate mathematical model has been developed, which allows automating the process of its construction.

The proposed methodology and software package will allow shifting to the mathematical model with a predetermined discrepancy between the results of modeling and experiment.

Using the methodology and software package, it is possible to refine existing mathematical models to more acceptable ones.

*Keywords:* evaluation function, adequate mathematical model, construction methodology, degree of acceptability, software complex.

**Амосов Владимир Владимирович**

К.т.н., доцент, Санкт-Петербургский политехнический университет Петра Великого  
amosov\_vv@spbstu.ru

**Петров Александр Владимирович**

Старший преподаватель, Санкт-Петербургский политехнический университет Петра Великого  
petrov\_av@spbstu.ru

*Аннотация.* Предложена методика построения адекватной математической модели по оценочной функции. Введены характеристики, определяющие степень приемлемости модели: допустимое расхождение результатов моделирования и эксперимента, необходимое количество состояний объекта исследования, пороги значимостей независимых переменных.

Разработан прототип программного комплекса построения адекватной математической модели, позволяющий автоматизировать процесс её построения.

Предложенные методика и программный комплекс позволят перейти к математической модели с наперед заданным расхождением результатов моделирования и эксперимента.

С помощью методики и программного комплекса возможна доработка существующих математических моделей до более приемлемых.

*Ключевые слова:* оценочная функция, адекватная математическая модель, методика построения, степень приемлемости, программный комплекс.

## Введение

**В** основе научных исследований, как правило, лежит математическая модель, поэтому важно, чтобы её степень приемлемости позволяла получать удовлетворительный результат. Степень приемлемости определяется как единица, делённая на разность между получающимся и допустимым расхождениями моделирования и эксперимента, при этом заказчик исследования задаёт допустимое расхождение результатов моделирования и эксперимента

Объект исследования можно описать оценочной функцией, зависящей от зависимых и независимых переменных, причём под независимыми переменными понимаются случайные величины, для которых заданы интервалы (допуски) и которые распределены в них

по заданным законам распределения (равномерный, нормальный и другие).

Для перехода от оценочной функции к адекватной математической модели для каждой независимой переменной определяются её абсолютная и относительная релевантности, максимальный коэффициент влияния, затем определяются количество её представительных значений ( $NW_i$ ) и их распределение на интервале неопределённости, а также энтропия и значимость переменной.

Экспертами задаётся порог значимости и те независимые переменные, значимость которых меньше этого порога, задаются в математической модели константами (берётся среднее арифметическое по границам соответствующего интервала); те независимые переменные,

значимость которых выше порога, задаются в математической модели дискретным распределением представительных значений на интервале неопределённости этих независимых переменных.

Задача состоит в подборе для каждой независимой переменной одного или нескольких представительных значений так, чтобы полученный набор из представительных значений ( $NW_i$ ) или параметров модели для всех независимых переменных удовлетворял условию, при котором разность между получающимся и допустимым расхождением моделирования и эксперимента была минимальна, а степень приемлемости максимальна.

Таким образом, получаем параметризованную математическую модель, переменными в которой являются зависимые переменные оценочной функции.

При переходе от оценочной функции к математической модели степень её приемлемости можно повышать, изменяя необходимое количество состояний объекта исследования  $N_{необх}$  пороги значимостей каждой независимой переменной и допустимое расхождение результатов математической модели с данными эксперимента.

Из проведённого литературного обзора журналов, входящих в библиографические и реферативные базы данных, был найден ряд статей [1, 2, 3], в которых рассматриваются семейства оценочных функций, но ни в одной из них не говорится о предлагаемом переходе от оценочной функции к математической модели.

Методика построения адекватной математической модели по оценочной функции

Оценочная функция задаётся зависимостью выходной переменной «у» от входных зависимых и независимых переменных:  $y = y(z_v, x_w)$ , где  $x_w = (x_1, \dots, x_m, \dots, x_L)$  — вектор независимых переменных,  $z_v = (z_1, \dots, z_k)$  — вектор зависимых переменных.

Здесь каждая зависимая переменная, например  $z_l$ , может быть задана несколькими значениями:  $z_{11} \dots z_{1k}$ .

Аналогично каждая независимая переменная, например  $x_m$ , может принимать ряд случайных значений  $x_{mj}, j = 1, \dots, n_m$  из заданного допустимого интервала или области неопределённости:  $\min_j x_{mj} \leq x_{mj} \leq \max_j x_{mj}$ , т.е. каждая независимая переменная задана допустимым интервалом.

Определим как независимая переменная  $x_m$  влияет на значение оценочной функции, при этом огра-

ничимся только одной зависимой переменной  $z_i$  (в дальнейшем будут учтены все зависимые переменные) [4].

При определении влияния независимой переменной  $x_m$  будем считать, что остальные независимые переменные принимают средние арифметические значения из своих интервалов  $x_{cp l}, l = 1, \dots, L, l \neq m$ .

Влияние области неопределённости независимой переменной  $X_m$  при заданном значении зависимой переменной  $z_i$  определяется абсолютной релевантностью  $Ra_{im}$ :

$$Ra_{im} = \frac{\max_j y(z_i, x_{cp1}, \dots, x_{mj}, \dots, x_{cpL})}{y(z_i, x_{cp1}, \dots, x_{cpm}, \dots, x_{cpL})} - \frac{\min_j y(z_i, x_{cp1}, \dots, x_{mj}, \dots, x_{cpL})}{y(z_i, x_{cp1}, \dots, x_{cpm}, \dots, x_{cpL})} \quad (1)$$

Далее для независимой переменной  $X_m$  и заданного значения зависимой переменной  $z_i$  определим коэффициент влияния или относительную релевантность:

$$Rb_{im} = \frac{Ra_{im}}{\sum_{i=1}^L Ra_{ii}} \quad (2)$$

После находим максимум коэффициента влияния независимой переменной  $X_m$ , перебирая все значения единственной зависимой переменной  $z_i$  и повторяя каждый раз вычисление коэффициента влияния:

$$R_m = \max_i Rb_{im} \quad (3)$$

Для учёта всех зависимых переменных  $z_1 \dots z_k$ , каждая из которых определяется несколькими значениями, по аналогии с примером для одной зависимой переменной перебирают все возможные наборы из значений разных зависимых переменных, для каждого из которых повторяют предыдущие вычисления и в конце определяют максимум коэффициента влияния независимой переменной  $X_m$ .

Аналогично находятся максимумы коэффициентов влияния всех независимых переменных  $(x_1, \dots, x_m, \dots, x_L)$ .

Значимость независимой переменной  $x_m$  описывается равенством

$$B_m = R_m \times H_m \quad (4)$$

где  $H_m$  — энтропия независимой переменной  $X_m$ , которая определяет её информативность. Для определения значимости независимой переменной  $X_m$  необходимо найти её энтропию.

Так как независимая переменная  $X_m$  непрерывно распределена на заданном допустимом интервале по заданной функции плотности вероятностей  $f_m(x)$ , то энтропия независимой переменной  $X_m$  вычисляется по формуле:

$$H_m(\Delta_m) = - \int_{-\infty}^{+\infty} f_m(x) \times \ln f_m(x) dx - \ln \Delta_m \quad (5)$$

здесь первое слагаемое описывает влияние на энтропию независимой переменной  $X_m$  со стороны вероятностного распределения, заданного функцией плотности вероятностей  $f_m(x)$ .

Второе слагаемое ( $-\ln \Delta_m$ ) учитывает зависимость от интервала дискретности ( $\Delta_m$ ).

Если в формулу энтропии  $H_m(\Delta_m)$  подставить функции плотности вероятностей  $f_m(x)$  равномерного и нормального законов распределения, то получим

для равномерного закона

$$H_{\text{равн}}(\Delta_m) = \ln((\max_j x_{mj} - \min_j x_{mj})/\Delta_m),$$

для нормального закона

$$H_{\text{норм}}(\Delta_m) = \ln \sqrt{\frac{2pey^2}{\Delta_m^2}},$$

здесь  $y$  — параметр нормального закона распределения.

Для поиска представительных значений ( $NW$ ) независимых переменных необходимо решить одну из задач теории оптимизации и принятия решений: представительные значения каждой независимой переменной следует выбрать так, чтобы переход от оценочной функции к математической модели был адекватным (точным), но, чтобы количество представительных значений в интересах уменьшения трудоёмкости было как можно меньше.

Рассмотрим один из способов решения этой задачи:

пусть  $L$  — число независимых переменных ( $x$ ), а  $N$  — число всех наборов (комбинаций) из представительных значений каждой независимой переменной, то есть  $N$  — это количество всех возможных состояний объекта исследования, тогда

$$N = \prod_{i=1}^L (NW_i) \quad (6)$$

Далее число представительных значений  $NW_m$  независимой переменной  $X_m$  полагают пропорциональным его значимости  $B_m$ :

$NW_m = P \times B_m$ , здесь  $P$  — коэффициент пропорциональности одинаковый для всех независимых переменных. Это первое допущение.

Если представить значимость через максимум коэффициента влияния независимой переменной  $X_m$  и её энтропию, то получим:

$$NW_m = P \times R_m \times H_m \quad (7)$$

Здесь энтропия  $H_m$  также зависит от числа представительных значений  $NW_m$  через  $\Delta_m$  — длину разбиения конечной области функции распределения независимой переменной  $X_m$ . Таким образом, получается уравнение с неизвестным  $NW_m$ , которое необходимо решить для всех независимых переменных.

Для его решения сначала ограничим функции распределения  $f_m(x)$  конечной областью ( $b$ ) и пусть  $\Delta$  для каждой независимой переменной постоянна:

$$b = \Delta \times NW \quad (8)$$

Из предположения о нормальном распределении (это второе допущение)  $b_{\text{норм}} = 6\sigma$  с другой стороны  $b = \Delta \times NW$ , поэтому

$$y = \frac{(\Delta \times NW)}{6}.$$

Если подставить выражение для  $y$  в формулу энтропии для нормального распределения получаем

$$H_{\text{норм}} = \ln \left( \frac{NW}{6} \times \sqrt{2pe} \right) = \ln(NW) + h_{\text{норм}}$$

здесь  $h_{\text{норм}}$  — дифференциальная энтропия для нормального закона распределения, которая равняется

$$h_{\text{норм}} = \frac{\sqrt{2pe}}{6} = -0.3728.$$

Аналогично рассчитывают дифференциальные энтропии для других законов распределения (Вейбулла, логарифмически нормального, равномерного).

Таким образом, для определения энтропий и значимостей всех независимых переменных осталось найти представительные значения ( $NW$ ) для каждой независимой переменной.

Для этого, зная дифференциальные энтропии законов распределения независимых переменных, эксперты задают необходимое число всех возможных состояний объекта исследования  $N_{\text{необх}}$ , которые следует учесть.

Задав  $N_{\text{необх}}$  — произведение количеств представительных значений всех независимых переменных, определяют их значения: сначала выбирают независимую переменную с наиболее низкой релевантностью, пусть это будет ( $x_m$ ).

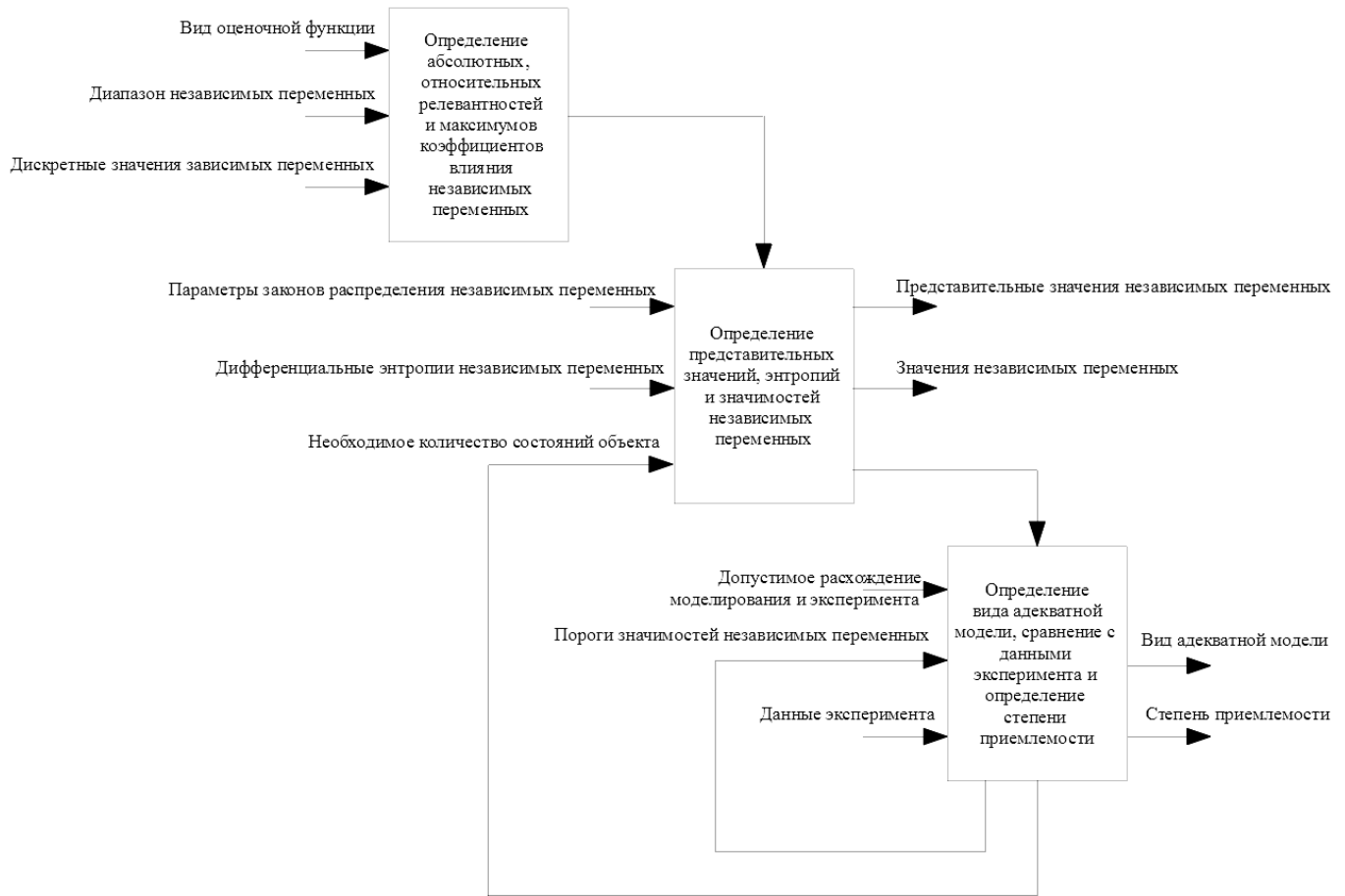


Рис. 1. Архитектура ПКПАММ

Для этой переменной заранее задают количество представительных значений  $NW_m$ .

Затем из выражения  $NW_m = P \times R_m \times H_m$  определяют  $P$ , дальше подставляют вместо  $(H_m)$  его выражение через дифференциальную энтропию  $H_m = \ln(NW_m) + h_m$  и из того, что коэффициент пропорциональности  $P$  одинаков для всех независимых переменных ( $P_m = P_l = P$ ), получают уравнение, из которого, задав количество представительных значений  $(NW_m)$  одной независимой переменной  $x_m$  можно найти представительные значения  $(NW_l)$  для остальных независимых переменных  $x_l$ .

Это уравнение выглядит следующим образом:

$$\frac{NW_m}{(\ln(NW_m) + h_m) \times R_m} = \frac{NW_l}{(\ln(NW_l) + h_l) \times R_l} \quad (9)$$

здесь  $NW_m$  — задано,  $h_m, R_m, h_l, R_l$  — определены, имеем единственное неизвестное — количество представительных значений  $(NW_l)$  для независимой переменной  $x_l$ . Каждый раз, перебирая все независимые переменные, решают это уравнение, меняя в нём дифференциальную энтропию  $(h)$  и максимум коэффициента влияния независимой пе-

ременной  $(R)$  на соответствующие той независимой переменной количество представительных значений которой определяется. В результате находят количество представительных значений  $(NW)$  для всех независимых переменных  $(x)$ . После этого количества представительных значений  $(NW)$  для всех независимых переменных  $(x)$  перемножают и результат сравнивают с заданным экспертами  $N_{необх}$  если результат меньше  $N_{необх}$  то количество представительных значений  $(NW_m)$  для независимой переменной  $(x_m)$  с наиболее низкой релевантностью увеличивают и повторяют вычисления количеств представительных значений  $(NW)$  для всех независимых переменных  $(x)$ , затем количества представительных значений  $(NW)$  для всех независимых переменных  $(x)$  снова перемножают и результат сравнивают с заданным экспертами  $N_{необх}$ ; если результат больше  $N_{необх}$  то возвращаются к предыдущей итерации и считают, что количества представительных значений  $(NW)$  всех независимых переменных  $(x)$  найдены.

Подставляем их и соответствующие дифференциальные энтропии  $(h)$  в формулу  $H = \ln(NW) + h$ , находим значимости всех независимых переменных и далее переходим от оценочной функции к адекватной математической модели.

Распределение самих представительных значений ( $NW$ ) всех независимых переменных ( $x$ ) на их областях неопределённости  $\min_j x_j \leq x_j \leq \max_j x_j$  ищутся из того, что вероятности этих значений должны соответствовать заданным функциям плотности вероятностей  $f(x)$  этих независимых переменных, то есть значения представительных значений ( $NW$ ) независимой переменной  $x_j$  должны соответствовать подинтервалам  $\Delta x_j$  области неопределённости  $x_j$ , на которых вероятности появления независимой переменной  $x_j$  одинаковы [4].

Далее для каждой независимой переменной подбирается одно или несколько представительных значений так, чтобы полученный набор из представительных значений ( $NW_i$ ) или параметров модели для всех независимых переменных удовлетворял условию, при котором разность между получающимся и допустимым расхождением моделирования и эксперимента была минимальна, а степень приемлемости максимальна.

Меняя входные данные  $N_{необх}$  и пороги значимостей, переопределяя вид адекватной модели, добиваются приемлемого совпадения результатов модели с данными эксперимента, при этом найденные входные данные определяют степень приемлемости математической модели.

Программный комплекс построения адекватной математической модели

Уровни и инструментарий проектирования программных комплексов описаны в работе [5].

На рисунке 1 представлена архитектура программного комплекса построения адекватной математической модели (ПКПАММ).

### Заключение

Предложены методика и программный комплекс, позволяющие перейти от оценочной функции к математической модели с наперёд заданным значением расхождения результатов моделирования и эксперимента.

Методика использует количественные характеристики теории принятия решений.

С помощью методики и программного комплекса возможна доработка существующих математических моделей до более приемлемых. В работе [6] описывается переход от математической модели к оценочной функции, но он не совпадает с предлагаемым, использующим количественные характеристики теории принятия решений.

Представлена архитектура программного комплекса построения адекватной математической модели, разработан его прототип, позволяющий автоматизировать процесс построения адекватной математической модели.

Прототип программного комплекса написан на языке программирования Python. Код программы: 4361 строк программного кода, 35 классов, 19 файлов. Комментарии: 1613 строк (37%)

### ЛИТЕРАТУРА

1. Fischer, A., Jiang, H. Merit functions for complementarity and related problems: A survey. Computational Optimization and Applications. 2000. 17 (2–3), pp. 159–182
2. Andreani, R., Júdice, J.J., Martínez, J.M., Patrício, J. On the natural merit function for solving complementarity problems. Mathematical Programming. 2011. 130 (1), pp. 211–223
3. Lu, L.-Y., Huang, Z.-H., Hu, S.-L. Properties of a family of merit functions and a merit function method for the NCP. Applied Mathematics. 2010. 25 (4), pp. 379–390
4. Мушик Э., Мюллер П. Методы принятия технических решений. М.: Мир, 1990. 208 с.
5. Черноуцкий И.Г., Амосов В. В., Петров А. В. Технологии проектирования при общепрофессиональной подготовке // Высокие интеллектуальные технологии и инновации в национальных исследовательских университетах: материалы Международной научно-методической конференции. 2014. Том 2, с. 126–129,
6. Koruba, Z., Krzysztofik, I. An algorithm for selecting optimal controls to determine the estimators of the coefficients of a mathematical model for the dynamics of a self-propelled anti-aircraft missile system. Proceedings of the Institution of Mechanical Engineers, Part K: Journal of Multi-body Dynamics. 2013. 227. DOI: 10.1177/1464419312455967

© Амосов Владимир Владимирович (amosov\_vv@spbstu.ru), Петров Александр Владимирович (petrov\_av@spbstu.ru).

Журнал «Современная наука: актуальные проблемы теории и практики»