

УДК 629.7.083.03

ИНФОРМАЦИОННЫЙ МЕТОД ОПТИМИЗАЦИИ ПАРАМЕТРОВ ПРИ ДИАГНОСТИРОВАНИИ ГТД

Г.С. ЗОНТОВ

В данной статье предложен алгоритм метода оптимизации параметров для диагностирования ГТД с целью разграничения зон работоспособности. Акцент сделан на рациональное комбинирование методов математического анализа и статистики и разработку алгоритма, позволяющего оптимизировать применение математических методов посредством лонгитюдного сбора данных и машинного обучения.

Ключевые слова: авиационный ГТД, оптимизация параметров, алгоритм обработки данных.

Построение модели диагностирования является важным этапом разработки метода информационной оценки диагностических параметров для любой сложной технической системы, в частности, для авиационных ГТД. Фактически, данная модель закладывается в основу предлагаемого метода. В этой связи особое внимание уделяется ключевым этапам построения модели.

К числу основных этапов следует отнести: выбор входящих параметров (независимых переменных), подбор зависимостей и последующий анализ поведения модели в рамках тестового набора данных. Применительно к ГТД данные этапы имеют свои особенности, которые напрямую следуют из теоретических положений, более подробно рассматриваемых на каждом из этапов моделирования.

На этапе выбора параметров модели необходимо учитывать, что существуют внутренние показатели (характеристики) ГТД, а именно вектор комплектующих – *x-вектор* с размерностью p . На следующем этапе в качестве набора рассматриваемых переменных выбирается μ -мерный вектор входных воздействий z . Этот вектор представляет собой совокупность управляющих воздействий на ГТД, каждое из которых характеризуется своими параметрами.

Таким образом, строится многомерная матрица, учитывающая одновременно комплектующие и управляющие воздействия на ГТД. При этом на данном этапе моделирования матрица является двумерной с размерностью pz .

Далее необходимо выбрать χ -мерный вектор возмущающих воздействий f , который определяется условиями эксплуатации – это внешние воздействия, которые могут вводиться в модель бинарным или вероятностным способом. Отметим, что при этом матрица становится трехмерной с размерностью pzf . Данная матрица отражает независимые переменные модели, являющиеся входящими данными.

Результирующим набором переменных является m -мерный вектор выходных параметров u , в который полностью входит множество S конечных состояний объекта диагностирования (ОД), в данном случае - авиационного ГТД, подразделяемого на два непересекающихся подмножества: работоспособных ($S1$) и неработоспособных ($S2$) состояний.

Для построения регрессионной модели необходимо свести многомерную матрицу к двумерной. Для этого трансформируем ее следующим образом: по вертикали откладываются наблюдения; по горизонтали - исследуемые параметры.

Наблюдения, повторяющиеся в данной матрице, имеют больший вес, чем однократные. Однако подобного рода повторение возможно только в случае, если данные двух наблюдений являются полностью идентичными. Размерность подобной матрицы составит $N*(p_j*z_k*f_h)$, где j, k, h – представляют размерность используемых векторов, а N – количество наблюдений. Заме-

тим, что на практике требуется набор подобных матриц, соответствующих каждому параметру из совокупности y .

Указанные параметры могут быть представлены в матричной форме (рис. 1).

N наблюдений	p1	p2	pj	z1	z2	zk	f1	f2	fh
N наблюдений	p1	p2	pj	z1	z2	zk	f1	f2	fh
N наблюдений	p1	p2	pj	z1	z2	zk	f1	f2	fh
1									
2									
3									
4									
5									
...	6								
...	7								
...									

Рис. 1. Набор матриц для каждой совокупности параметров y

Перейдем к рассмотрению формы зависимости. Для определения формы зависимости необходимо протестировать взаимозависимости тестируемых параметров, а также проанализировать их влияние на независимые переменные. Затем сократить набор переменных и определить характер (степень) его влияния на результирующий показатель.

Для выполнения данного набора действий необходимо рассчитать корреляцию для каждого набора векторов (p, z, f) соответственно. Затем рассмотреть влияние каждого набора (по отдельности) на каждую переменную, входящую в множество y . И, с помощью t -критерия отсеять наименее значимые. Подобный способ позволяет подтвердить достоверность взаимосвязей независимых переменных, а именно тот факт, что они имеют схожую природу и одновременно учитывает их избыточность, в связи с тем, что при значении t -критерия менее 2 по модулю, влияние данного фактора можно считать несущественным или нестабильным. В результате, мы сократим количество влияющих переменных для каждой переменной из совокупности y .

Затем необходимо определить форму распределения каждой оставшейся независимой и зависимой переменной. Для этого необходимо построить график распределения, где на оси X будут расположены наблюдения, а на оси Y значения переменной для каждого наблюдения. Таким образом, мы сможем определить форму распределения для каждой переменной. И включить в итоговую модель регрессии с учетом данной форму, линеаризовав все переменные, включая зависимые.

Работоспособность и предсказательную силу модели необходимо протестировать с помощью процедуры кросс-валидации на сгенерированных данных. Поскольку для разных моделей двигателей, эксплуатируемых в различных условиях, наборы данных будут отличаться, необходимо и достаточно доказать работоспособность и эффективность модели на любом наборе данных, с учетом физических ограничений их вариативности. Для этого необходимо будет построить регрессию, с учетом линеаризации всех факторов. Проверить ее значимость и предсказательную силу с помощью F -статистики и R -квадрата. И вставить в конечную формулу методы линеаризации для каждой переменной.

Заметим, что в реальной практике модель будет считаться для конкретного набора данных. В связи с чем каждый этап данного алгоритма необходимо будет повторять сначала. В частности, для различных наборов данных избыточными могут оказаться различные переменные, а, кроме того, методы линеаризации для каждого вектора могут быть различными. Однако общий алгоритм окажется неизменным.

Кроме того, при машинном обучении и расчетах необходим именно алгоритм формирования модели, на основе которого и будут реализовываться конкретные решения с использованием основных функций математического моделирования в заданной последовательности. Предлагаемый алгоритм представлен на рис. 2.

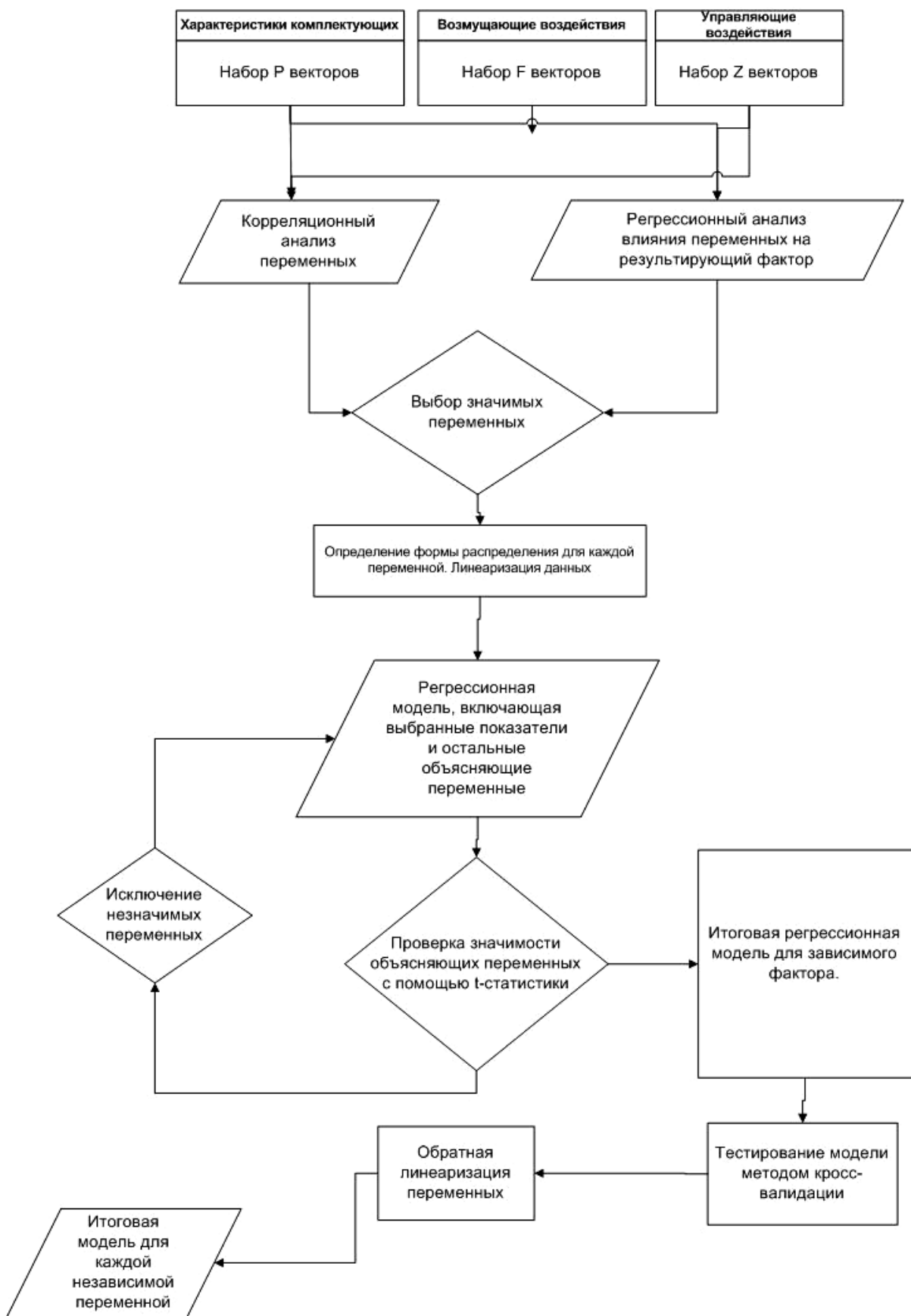


Рис. 2. Алгоритм диагностирования

Исходя из вышесказанного, карты диагностирования на основе информационного метода для авиационного ГТД должны строиться в четырехмерном пространстве, для осей p , z , f и y соответственно. Однако построение подобного рода пространства весьма затруднительно, в связи с чем его необходимо представить в качестве 3-х связанных двумерных графиков – карт диагностирования с осями p - y , z - y , f - y . Отметим, что, не смотря на вариацию только одного признака, на каждом графике должны быть представлены области работоспособных ($S1$) и неработоспособных ($S2$) состояний. Также необходимо указать, что векторы p , z , f , y на практике являются наборами векторов. Схематическое отображение столь сложной схемы не является разумным, в связи с чем автор считает, что необходимо использовать функцию из логистической модели, а именно уравнения (1) и (2) следующим образом.

Рассматривается логистическая регрессия, которая выражает статистическую связь в виде зависимости $P\{Y=1|X\}=f(X)$, т.е. прогнозируется вероятность события $\{Y = 1\}$, обусловленная значениями независимых переменных X^1, \dots, X^p . Геометрически суть задачи состоит в том, чтобы найти одну из возможных гиперплоскостей, которая бы в определенном смысле наилучшим образом разделяла бы две группы наблюдений (соответствующие 0 и 1) в пространстве регрессоров.

Логистическая регрессия выражает модель связи между откликом и переменными в виде формулы

$$P\{Y=1|X_1, \dots, X_p\} = \frac{e^{\hat{Y}}}{1+e^{\hat{Y}}}, \quad (1)$$

где переменная $\hat{Y} = \theta_0 + \theta_1 X_1 + \dots + \theta_p X_p$ называется логитом. Такая модель с бинарной зависимой переменной, по сути, является функцией логистического закона распределения

$$F(x) = \frac{e^{(x-a)/k}}{1+e^{(x-a)/k}}, \quad (2)$$

в которой в качестве аргумента используется линейная комбинация независимых переменных. Заметим, что при увеличении количества независимых векторов структура модели остается неизменной.

В результате, под действием различных векторов, входящих в множество p , z , f и y , возможно выявление принадлежности состояния авиационного двигателя к работоспособному, неработоспособному, а также выявить предаварийное состояние. Карта диагностирования в данном случае выглядит следующим образом (рис. 3).

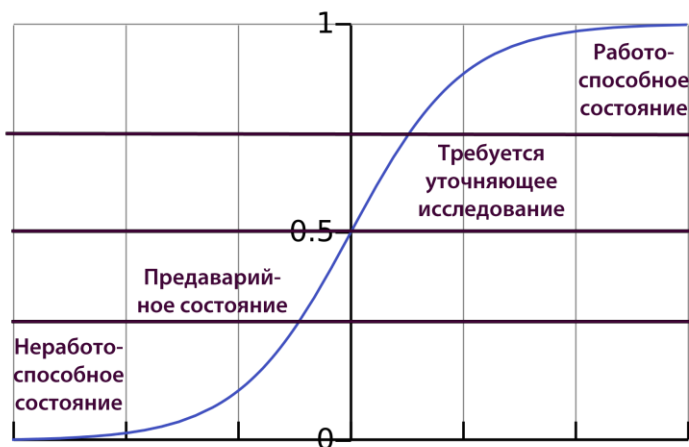


Рис. 3. Карта диагностирования

Отметим, что наиболее важной характеристикой для разграничения зон работоспособности является изгиб логистической кривой принадлежности состояния двигателя к работоспособному. Иными словами, смена состояния происходит при существенном изменении тангенса угла наклона касательной в точке к графику логистической регрессии.

Подобная предпосылка имеет под собой четкий математический критерий, проверка которого возможна в каждый момент времени, а, кроме того, зона различной работоспособности однозначно разделены между собой. Следует обратить внимание на зону, требующую уточняющего исследования. Это необходимо в связи с тем, что проведение дополнительных исследований двигателя и оперативное устранение недостатков может вернуть его в работоспособную зону, в то время как отсутствие подобных действий – вывести в предаварийное состояния.

Таким образом, обобщая все вышесказанное, автор предлагает следующую последовательность (алгоритм) расчетов для определения работоспособности авиационного ГТД:

1. Сведение всех векторов параметров p, z, f в единую таблицу, для каждого y из множества выходных параметров;
2. Упорядочивание и отсечение части множества с целью оптимизации затрат на измерения;
3. Расчет формулы регрессии для каждого y (рис. 2);
4. Сопоставление различных наборов y , с зонами работоспособности методом логистической регрессии (уравнения (1), (2), (рис. 3);
5. Сопоставление различных состояний (оцененных экспертным путем) с первоначальными данными по входным наборам векторов p, z, f ;
6. Машинное обучение системы, фиксирующее совпадение результатов расчетов, полученных в пунктах 4 и 5. В результате, формируется общий набор первоначальных параметров, загружаемый в логистическую регрессию и приводящий к итоговому результату (рис. 3), пропускающий промежуточные стадии исследования.

Отметим, что подобный подход применим только в долгосрочной перспективе и требует валидации на этапе внедрения, с последующими плановыми сравнениями результатов моделирования и расчетов. Важно подчеркнуть, что при изменении наблюдаемых параметров, методов измерения или иных внешних или внутренних факторов, реализация алгоритма должна повторяться в полном объеме.

ЛИТЕРАТУРА

1. **Анатольев С.** *Непараметрическая регрессия* //Квантиль. 2009. Т. 7. С. 37-52.
2. **Вербик М.** *Путеводитель по современной эконометрике*; пер с англ. В.А. Банникова; под науч. ред. и предисл. С. А. Айвязана М.: *Научная книга*, 2008.
3. **Воронцов К. В.** *Машинное обучение: курс лекций* [Электронный ресурс] URL: [http://www. machinelearning. ru](http://www.machinelearning.ru). 2009.
4. **Канторович Г. Г.** *Лекции: Анализ временных рядов* //Экономический журнал ВШЭ. 2002. Т. 6. №. 1. С. 85-116.
5. **Терентьев А. Н., Бидюк П. И.** *Эвристический метод построения байесовских сетей* //Математические машины и системы. 2006. Т. 1. № 3.
6. **Зонтов Г.С.** *Разработка метода оценки информативности диагностических параметров авиационных ГТД в процессе технической эксплуатации* // Тез. докл. конф. «XLVIII Научные чтения памяти К.Э.Циалковского». Калуга.: Изд-во «Эйдос», 2013. С. 160-161.
7. **Машошин О.Ф.** *Диагностика авиационных газотурбинных двигателей с использованием информационного потенциала контролируемых параметров*: автореферат дис. ... д-ра техн. наук: 05.22.14. М., 2005. 233 с.

INFORMATION METHOD OF OPTIMIZATION PARAMETERS IN THE DIAGNOSIS OF GAS TURBINE ENGINES

Zontov G.S.

This article describes an algorithm parameter optimization method for the diagnosis of GTD in order to divide zones of efficiency. The autor focuses on the retional combination of methods of mathematical analysis and statistics and developing an algorithm that allows to optimize the use of mathematical methods by longitudinal data collection and machine learning.

Keywords: aviation turbine engine, optimization of parameters, data processing algorithm.

REFERENCES

1. **Anatolyev S.** *Neparametricheskaya regressiya*. Kvantil'. 2009. T. 7. Pp. 37-52. (In Russian).
2. **M. Verbeek** *Putevoditel' po sovremennoj jekonometrike*. Verbik Marno; per s angl. V.A. Bannikova; pod nauch. red. i predisl. S. A. Ajvjazana. M.: Scientific book. 2008. (In Russian).
3. **Vorontsov K.V.** Mashinnoe obuchenie. Kurs lektsiy [Elektronnyy resurs] URL: [http:// www. machinelearning. ru.](http://www.machinelearning.ru) - 2009. (In Russian).
4. **Kantorovich G.G.** Lekcii: Analiz vremennyh rjadov. *Economic Journal of Economics*. 2002. T. 6. №. 1. - Pp. 85-116. (In Russian).
5. **Terent'ev A.N., Bidyuk P.I.** Evristicheskiy metod postroeniya bayesovskikh setey. *Matematicheskie mashiny i sistemy*. 2006. T. 1. №. 3. (In Russian).
6. **Zontov G.S.** Razrabotka metoda otsenki informativnosti diagnosticheskikh parametrov aviatsionnyh GTD v protsesse tekhnicheskoy ekspluatatsii . Tez. dokl. konf. «XLVIII Nauchnye chteniya pamyati K.E.Tsiolkovskogo». Kaluga: Publishing. "Eidos", 2013. Pp. 160-161. (In Russian).
7. **Mashoshin O.F.** Diagnostika aviatsionnykh gazoturbinnnykh dvigateley s ispol'zovaniem informatsionnogo potentsiala kontroliruemykh parametrov: avtoreferat dissertatsii ... doktora tekhnicheskikh nauk: 05.22.14. Moscow, 2005. 233 p. (In Russian).

Сведения об авторе

Зонтов Георгий Сергеевич, 1990 г.р., окончил МГТУ ГА (2012), аспирант МГТУ ГА, инженер ПТО в ООО «АВКОМ Техник», автор 1 научной работы, область научных интересов – диагностика авиационной техники.