Белорусский национальный технический университет, Минск

ИНФОРМАЦИОННОЕ МОДЕЛИРОВАНИЕ ДАННЫХ ПРОИЗВОДСТВЕННЫХ ПРОЦЕССОВ КАК ЗАДАЧА АППРОКСИМАЦИИ ФУНКЦИИ СВЯЗИ

В. М. Романчак, П. С. Серенков

The algorithm of information modeling as a basis of mechanism of acquisition and analysis of manufacturing processes quality data of a typical industrial enterprise was defined. Requirements and constraints to the model of process quality indexes as function of influencing factors in real process operation were determined. Universal method of approximation of passively registered quality data as an algorithm of approximation by superposition of wavelets in the form of delta-shaped functions was offered.

Математическую модель аппроксимации функции связи представим в виде:

$$\varphi = f(x_1, x_2 ... x_m, a_1, a_2, ... a_n) \tag{1}$$

где φ - оцениваемый показатель результативности продукции и (или) процесса (показатель качества), x_1, x_2 ...x_m – факторы, влияющие на показатель качества, f - аналитическая функция связи, a_1, a_2, ...a_n – параметры аналитической функции связи, подлежащие определению.

С позиций задач информационного моделирования производственных процессов мы предлагаем следующую классификацию моделей:

- модель с фиксированным количеством параметров;
- модель с нефиксированным количеством параметров.

К первой группе относятся математические модели, у которых число параметров **n** фиксировано и не зависит от количества **s** случаев (зарегистрированных наборов значений факторов x_{1s}, x_{2s}, ... x_{ms}).

Типичными представителями моделей с фиксированным количеством параметров являются классические параметрические регрессионные модели, ряды с фиксированным числом слагаемых и т.п.

Ко второй группе относятся математические модели, у которых число параметров **n** не фиксировано и определяется количеством **s** случаев (зарегистрированных наборов значений факторов x_{1s}, x_{2s}, ... x_{ms}).

Типичными представителями моделей с нефиксированным количеством параметров являются оценки типа Парзена - Розенблата, оценивание по методу потенциалов, а также полиномы Бернштейна [4 - 6].

Модели с фиксированным количеством параметров, не всегда позволяют адекватно описать реальные производственные процессы. Их применение для построения аналитической функции связи неизбежно вносит методическую ошибку при аппроксимации входных данных. Сущность методической ошибки заключается в том, что фактическую, априорно неизвестную зависимость оцениваемого показателя качества от влияющих факторов инженер – аналитик представляет в виде заранее выбранной им модели конкретного типа. Например, если фактическую нелинейную зависимость аппроксимировать линейной регрессионной моделью, появляется методическая ошибка аппроксимации, причем значение ее может быть соизмеримо не только со случайными составляющими ошибки, но и с самим значением показателя качества.

Для решения сложных задач анализа данных и принятия решений в условиях существенной априорной неопределенности наиболее перспективными, на наш взгляд, являются модели с нефиксированным количеством параметров. Использование такого рода моделей для решения данного класса задач позволяет свести к минимуму методическую ошибку аппроксимации.

Один и тот же алгоритм математического моделирования функции связи должен обеспечивать решение двух задач:

1. приближенное восстановление функции - поиск гладких кривых или поверхностей, проходящих через множество заданных точек или линий;
2. сглаживание функции.

Можно считать, что показателем качества, разрабатываемой математической модели аппроксимации функции связи является комплексная характеристика, которая показывает, насколько данная модель пригодна для решения поставленных задач с учетом конкретных ограничений, и насколько точно она решает эти задачи.

Для того, чтобы обеспечить точность решения задач, модель должна удовлетворять ряду требований (принципов). Руководствуясь практическим опытом, мы предлагаем оценить показатель качества разрабатываемой модели аппроксимации функции связи соблюдением трех принципов:

1. принцип несмещенности (асимптотической несмещенности);
2. принцип состоятельности;
3. принцип интерполяционной устойчивости алгоритма порядка **k**.

Для анализа соответствия различных классов аналитических моделей сформулированным принципам и ограничениям, были рассмотрены модели:

- аппроксимация алгебраическими многочленами [7];
- регрессионное оценивание типа Парзена – Розенблата [5];
- полиномы Бернштейна [6];
- метод потенциалов (неортогональных разложений) [4];

Во всех случаях имели место несоответствие моделей определенным выше принципам и ограничениям предметной области.

Прежде всего, это ограничения, вызываемые реальными условиями осуществления производственных процессов и функционирования систем менеджмента качества, для которых решается данная задача:

- неупорядоченный (пассивный) механизм сбора и накапливания исходной информации, что приводит к неравномерности заполнения факторного пространства и пропускам входных данных;
- сложная иерархическая структура входных данных, привязанная к организационной структуре предприятия; представление данных как в явном так и неявном виде, как в количественной, так и в качественной форме;
- стохастичность и зашумленность - входных данных (характеристик продукции и параметров процессов), их низкая воспроизводимость;
- нестационарность входных данных во времени;
- недостаточный уровень подготовленности инженеров и менеджеров в области качества для проведения на постоянной основе квалифицированного информационного моделирования, недостаточные мотивация и профессиональный интерес к применению предлагаемых подходов для решения всего комплекса задач менеджмента качества.

Поиск модели аппроксимации функции связи показателя качества производственного процесса от влияющих факторов предлагается искать среди методов «точного» восстановления функции связи. К таковым, например, можно отнести, например, разложение в ряд Фурье или интегрального преобразования Фурье [8].

В свою очередь обозначим Vf, как изображение функции $f(t)$ при преобразовании Фурье:

$$Vf = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t)\, e^{-ixt}\, dt = \hat{f}(x) \tag{2}$$

В теории сигналов в последнее время активно развивается теория двухпараметрических интегральных преобразований или вейвлет - преобразований [8]. Идея состоит в разложении функции по двухпараметрическому базису.

Интегральным вейвлет - преобразованием функции $f(t)$ называется выражение:

$$W_{\psi}\ (f)(b,a)=\frac{1}{\sqrt{a}}\int_{-\infty}^{\infty} f(t)\ \overline{\psi(\frac{t-b}{a})}\ dt \tag{3}$$

Входящая в выражение функция $\psi\ (t)$ - называется вейвлетом и должна удовлетворять условию допустимости:

$$\int_{-\infty}^{\infty}\psi\ (t)dt=0 \tag{4}$$

При этом обеспечивается обратимость вейвлет - преобразования с помощью формулы:

$$f(t)=\frac{1}{C_{\psi}}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}W_{\psi}(f)(b,a)(\frac{1}{\sqrt{a}}\psi(\frac{t-b}{a}))\frac{db\,da}{a^2} \tag{5}$$

где $C_{\psi}=2\pi\int_{-\infty}^{\infty}\frac{\overline{\hat{\psi}(\omega)\hat{\psi}(\omega)}}{\omega}d\omega<\infty$, $\hat{\psi}(\omega)$ - преобразование Фурье функции $\psi\ (\omega)$.

Недостатком вейвлет – преобразования для определенного нами класса задач является тот факт, что функции, которые могут выступать в роли вейвлета, должны удовлетворять условию допустимости (4). Это значительно сужает класс функций, которые могут использоваться в качестве вейвлета и соответственно снижает функциональные возможности информационного моделирования производственных процессов.

Нами рассматриваются некоторые новые подходы к построению модели аппроксимации функции связи.

Предлагаемый подход предполагает, что аналитическая функция связи представлена вейвлет – преобразованием, причем в качестве вейвлета используются дельта – образные функции [4, 5, 7. 8].

Используя обозначения для интегрального вейвлет – преобразования, рассмотрим теперь аппроксимацию с помощью дельта – образных функций.

Пусть оцениваемая функция $f(t)$ непрерывна в точке b. Тогда ее значение $f(b)$ можно представить в виде:

$$f(b)=\int_{-\infty}^{\infty}f(t)\ \delta(t-b)\ dt \tag{6}$$

где $\delta(t)$ - дельта - образная функция.

Равенство (6) можно приближенно записать в виде:

$$f(b\ )\approx\frac{1}{\sqrt{a}}W(b\ ,a) \tag{7}$$

где $$W_{\psi}\ (f)(b,a)=\frac{1}{\sqrt{a}}\int_{-\infty}^{\infty} f(t)\ \overline{\psi(\frac{t-b}{a})}\ dt \tag{8}$$

Функция $\psi(t)$ - дельта - образная функция, которая должна при этом удовлетворять условию:

$$\int_{-\infty}^{\infty}\psi(t)dt=1 \tag{9}$$

При этом $\psi(t)$ должна стремиться к $\delta(t)$ при $a\to\infty$.

Примечание. Метод дельта – образных функций нашел широкое применение при оценивании плотностей распределения и регрессионных зависимостей [4].

Введем определение сингулярного вейвлет - преобразования. Предположим, что $\psi \in L^2(R)$. Образуем из ψ двухпараметрическое семейство вейвлетов с помощью сдвигов и растяжений [8]:

$$\psi^{(a,b)}(x) = \frac{1}{\sqrt{a}}\psi(\frac{t-b}{a})dt \tag{10}$$

где $a, b \in R$, $a \neq 0$

Нормирование выберем так, чтобы $\left\|\psi^{(a,b)}\right\| = \left\|\psi\right\| = 1$

Непрерывное сингулярное вейвлет - преобразование на этом семействе вейвлетов мы предлагаем определить как

$$W_\psi\ (f - f(b))(b,a) = \frac{1}{\sqrt{a}}\int\limits_{-\infty}^{\infty}(f(t) - f(b))\ \overline{\psi(\frac{t-b}{a})}\ dt \tag{11}$$

В частном случае, когда для функции $\psi(t)$ выполняется условие допустимости (4) сингулярное интегральное вейвлет – преобразование совпадает с интегральным вейвлет - преобразованием:

$$W_\Psi\ (f - f(b))\,(b,a) = W_\Psi(f)\,(b,a) \tag{12}$$

<u>*ДОКАЖЕМ ТЕОРЕМУ*</u>. Пусть ψ- базисный сингулярный вейвлет, тогда

$$\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}W_\psi(f - f(b))\,(b,a)\,\overline{W_\psi\ g(b,a)}\,\frac{da}{a^2} = C_\psi\ (f,g)\ , \tag{13}$$

где $(f\ , g) = \int\limits_{-\infty}^{\infty}f(t)\ \overline{g(t)}dt$, $\tilde{N}_\psi = \int\limits_{-\infty}^{\infty}\frac{(\overline{\hat{\psi}(y)} - \overline{\hat{\psi}(0)})\ \hat{\psi}(y)\ dy}{y} < \infty$.

Действительно,

$$V(W_\psi f(b,a)) = \frac{a}{\sqrt{a}}\hat{f}(x)\,\overline{\hat{\psi}(ax)} \tag{14}$$

$$V(W_\psi g(b,a)) = \frac{a}{\sqrt{a}}\ \hat{g}(x)\ \overline{\hat{\psi}(ax)} \tag{15}$$

Кроме того

$$\begin{aligned}&\int\limits_{-\infty}^{\infty}W_\psi(f - f(b))\,(b,a)\,\overline{W_\psi\ g(b,a)}\,db = \\ &= \int\limits_{-\infty}^{\infty}W_\psi\ f(b,a)\,\overline{W_\psi(b,a)}\ db - \int\limits_{-\infty}^{\infty}\frac{a}{\sqrt{a}}\,f(b)\ \overline{\hat{\psi}(0)}\ \overline{W_\psi\ g(b,a)}\,db\end{aligned} \tag{16}$$

Применяя обобщенные равенства Парсеваля к каждому из интегралов в правой части равенства (16) с учетом (14), (15) получим

$$\int_{-\infty}^{\infty} W_\psi (f - f(b))(b,a)\overline{W_\psi g(b,a)}db =$$

$$= \frac{a^2}{a}\int_{-\infty}^{\infty} \hat{f}(x)\,\overline{\hat{g}(x)}\,(\overline{\hat{\psi}(ax)} - \overline{\hat{\psi}(0)})\;\hat{\psi}(ax)\,dx \quad . \qquad (17)$$

Умножив на $\frac{da}{a^2}$, получаем

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} W_\psi f(b,a)\,\overline{W_\psi g(b,a)}\frac{da}{a^2}\,db = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\hat{f}(x)\,\overline{\hat{g}(x)}\frac{(\overline{\hat{\psi}(ax)} - \overline{\hat{\psi}(0)})\,\hat{\psi}(ax))}{a}\,da\,dx =$$

$$= \int_{-\infty}^{\infty}\hat{f}(x)\overline{\hat{g}(x)}\int_{-\infty}^{\infty}\frac{(\overline{\hat{\psi}(y)} - \overline{\hat{\psi}(0)})\,\hat{\psi}(y)\,dy}{y}\,dx = C_\psi\int_{-\infty}^{\infty} f(x)\overline{g(x)}\,dx \qquad (18)$$

Теперь аналогично подходу, изложенному в [8], можно получить формулу обращения для сингулярного интегрального вейвлет – преобразования:

$$f(t) = \frac{1}{C_\psi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} W_\psi (f - f(b))\,(b,a)\,(\frac{1}{\sqrt{a}}\,\psi(\frac{t-b}{a}))\frac{db\,da}{a^2} \quad , \qquad (19)$$

где $\tilde{N}_\psi = 2\pi\int_{-\infty}^{\infty}\frac{(\overline{\hat{\psi}(y)} - \overline{\hat{\psi}(0)})\,\hat{\psi}(y)\,dy}{y} < \infty$, $\hat{\psi}(\omega)$ - преобразование Фурье функции $\psi(t)$.

Формула (19) показывает, что любую $f(t)$ из $L^2(R)$ можно с произвольной точностью аппроксимировать суперпозицией вейвлетов в виде дельта - образных функций.

Из формулы (13) следует равенство:

$$\int_{-\infty}^{\infty} f^2(t)\,dt = \frac{1}{\tilde{N}_\psi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} W_\psi (f - f(b))\,(b,a)\,\overline{W_\psi (f)\,(b,a)}\frac{db\,da}{a^2}\,, \qquad (20)$$

которое представляет собой аналог равенства Парсеваля. Последнее предполагает, что математическая модель функции связи в виде сингулярного интегрального преобразования обладает дополнительными качественными свойствами [8].

В дискретном случае можно воспользоваться следующим рядом:

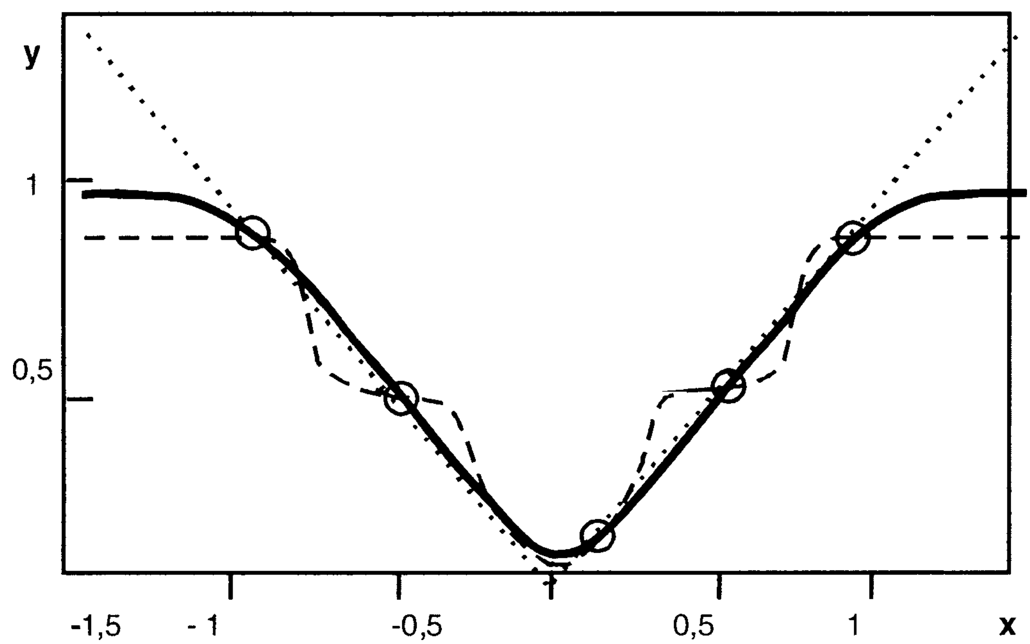
$$f(t) = \sum_{k=0}^{K}\frac{\sum_{m=1}^{M} g_k(b_m)\,\psi(\frac{t-b_m}{a_k})}{\sum_{m=1}^{M}\psi(\frac{t-b_m}{a_k})}\,, \qquad (21)$$

где $g_{k+1}(b_m) = W_\psi\,(g_k - g_k(b_m))(b_m, a_k)$, $g_0(b_m) = f(b_m)$, $a_k = \frac{a}{2^k}$.

В частном случае, когда $k = 0$, ряд совпадает с оценкой Парзена - Розенблата .

Нами был реализован численный эксперимент по оцениванию комплексного показателя качества различных моделей аппроксимации, рассмотренных в данной статье. Сущность эксперимента заключалась в том, что в качестве объекта моделирования задавалась теоретическая функция, например, $y = |x|$, $y = \exp(x)$, $y = \sin(x)$, $y = x^2$. Для заданной теоретической функции связи генерировался случайным образом массив точек $(x_i, y_i \pm \delta_i)$, где δ_i – случайным образом заданная ошибка. При этом, $\delta_i \leq [\delta]$, x_i представляет собой неравномерный ряд значений аргументов.

Массив точек рассматривался как совокупность данных о результативности производственного процесса, зафиксированных в условиях его реализации, близких к реальным, по крайней мере, с позиций выявленных нами ограничений. Далее массив точек аппроксимировался различными методами, приведенными выше. На рис. 1 приведен фрагмент численного эксперимента по оцениванию комплексного показателя качества различных моделей аппроксимации функции связи $y = |x|$.



аппроксимация функции дискретным вейвлет – рядом;

оценка Парзена – Розенблата;

теоретическая функция связи $y = / x /$;

узлы интерполяции.

Рис.1. Аппроксимация теоретической функции связи $y = |x|$ различными методами для неравномерного распределения узловых точек.

Примечание. Вопросы, связанные с обоснованием оценки (21), выходят за рамки данной статьи.

Результаты численного эксперимента показали, что предлагаемый подход для данного круга модельных задач соответствует установленным требованиям:

- алгоритм аппроксимации суперпозицией вейвлетов в виде дельта - образных функций по несистематизированным исходным данным восстанавливает и сглаживает функцию связи, обеспечивая выполнение принципа интерполяционной устойчивости алгоритма не менее первого порядка;
- пригоден для решения поставленных задач менеджмента качества производственных процессов с учетом конкретных ограничений.

Заключение

Для реализации принципа менеджмента качества «управление, основанное на фактах», сформулирован алгоритм информационного моделирования как основа механизма сбора и анализа данных о качестве производственных процессов типового промышленного предприятия. Сформулированы требования и ограничения к модели функции связи показателей качества процесса с влияющими факторами для реальных условий его осуществления. На доказательной основе предложен универсальный метод аппроксимации пассивно регистрируемых данных о качестве как алгоритм аппроксимации суперпозицией вейвлетов в виде дельта - образных функций.

ЛИТЕРАТУРА

1. Окрепилов В.В. Управление качеством. М.: Экономика, 2000.

2. П.С. Серенков, В.Л. Соломахо К вопросу о методах и инструментах эффективного менеджмента качества. // Новости. Стандартизация и сертификация. 2002. №2. – С.57 – 60.

3. V.N. Koreshkov, P. S. Serenkov etc. Conception of Complex Methodology of “Cycled-Through” Quality Management. // Quality: The Way to Sustainability: Proceedings of 49th EOQ Annual Congress (25-27 April 2005, Antalya – Turkey) – 18 p.

4. Э.А. Надарая Непараметрическая оценка многомерной плотности вероятности и кривой регрессии. Тбилиси: ТГУ, 1983.

5. Wolfgang Hardle. Applied Nonparametric Regression. Berlin, 1994.

6. Farouki R.T. and Rajan V.T. Algorithms for polynomials in Bernstein form //Computer Aided Geometric Design. 1988. № 5. – P.1-26.

7. Калиткин Н.Н. Численные методы. М.:Наука, 1978.

8. И. Добеши. Десять лекций по вейвлетам. Москва–Ижевск. 2004