1997 ТОМ VII (XV)

УДК 681.5

АЛИЕВ Т.А., АМИРОВ З.А.

ВЫБОР ПАРАМЕТРА РЕГУЛЯРИЗАЦИИ ПРИ РЕШЕНИИ ЗАДАЧИ СТАТИСТИЧЕСКОЙ ИДЕНТИФИКАЦИИ

Введение. Известно, что решение задач, связанных со стохастическими прцессами (анализ прхождения через динамические системы, идендификация, оптимизация, адаптация, диагностика динамических систем), требует решения уравнений типа Винера- Хопфа. Определение статических зависимомтей между входными и выходными технологическими параметрами производится с помощью процедуры регрессионного анализа. В обоих случаях требуется определение корреляционных матриц по реализациям случайных прцессов. Считается, что реализации $g(t)$ случайных функций $x(t)$ являются стационарными эргодичными с нормальным законом распределения, а интервал наблюдения T выбран достаточно большим, поэтому вместо корреляционных функций $R_{xx}(\tau)$ можно пользоваться оценками $R_{gg}(\tau)$, вычисленными для $g(t)$.

Решение таких уравнений на практике сталкивается с проблемой некорректности, связанной с неустойчивостью результатов вычислений, когда небольшие погрешности при определении оценок корреляционных функций $R_{gg}(\tau)$ вызывают большие изменения в результатах, что приводит к искажению характеристики изучаемого объекта [1-4].

Для уменьшения влияния неточности исходных данных на результаты идендификации, разработаны методы, основанные на применении тестовых сигналов специального вида [4,5,6], разложении математических моделей объектов в ряд по ортогональным системам функций [5,6], сглаживание информационных сигналов [5,6], уравновешивание их погрешностей [7]. Несмотря на высокий теоретический уровень этих работ, опыт успешного практического применения их при построении математических моделей реальных объектов невелик.

Среди множества известных методов наиболее широкое распространение получили метод регуляризации и его модификации [8-13]. При всех своих бесспорных достоинствах, данный метод имеет существенный недостаток: невозможность практического выбора оптимального параметра регуляризации в силу его зависимости от неизвестных параметров. Ввиду этого нет полной гарантии того, что после применения метода регуляризации будут получены удовлетворительные результаты.

1. Постановка задачи. Предположим, что исследуемый технологический процесс описывается моделью вида

$$y = \beta_0 + \sum_{i=1}^{k} \beta_i x_i, \qquad (1.1)$$

где x_i - i-й входной сигнал, y - выходной сигнал, $\beta_i \left(i = \overline{0,k}\right)$- неизвестные коэффициенты математической модели.

Для определения неизвестных величин используется процедура регрессионного анализа, которая сводится к решению уравнения [11]

$$R_{xx} b = R_{xy} \qquad (1.2)$$

где $\beta = \left(\beta_0, \beta_1, \ldots, \beta_k\right)^T$;

$R_{xx} = X^T X$ - матрица, элементами которой являются значения авто- и взаимно- корреляционных функций между входными сигналами при временном сдвиге $\tau = 0$;

$R_{xy} = X^T y$ - вектор, элементами которого являются значения взаимнокорреляционных функций между входными и выходными сигналами при временном сдвиге $\tau = 0$;

$y = \left(y_1, \ldots y_n\right)^T$ - вектор, составленный по данным наблюдений выходного сигнала;

$X = \begin{pmatrix} 1 & x_{11} & \ldots & x_{k1} \\ 1 & x_{12} & \ldots & x_{x2} \\ . & . & . & . \\ 1 & x_{1n} & \ldots & x_{kn} \end{pmatrix}$ - матрица, составленная по данным наблюдений входных сигналов.

Учитывая, что измерения реальных сигналов искажаются аддитивными помехами с законом распределения, близким к нормальному и с математическими ожиданиями, близкими к нулю, уравнение *(1.2)* приводится к виду:

$$R_{gg} b = R_{g\eta} \qquad (1.3)$$

где $b = \left(b_0, b_1, \ldots, b_k\right)$- вектор оценок неизвестных коэффициентов;

$$R_{gg} = G^T G;\ R_{g\eta} = G^T \eta;\ \eta = \left(\eta_1, \ldots, \eta_n\right)^T;\ G = \begin{pmatrix} 1 & g_{11} & \ldots & g_{k1} \\ 1 & g_{12} & \ldots & g_{k2} \\ . & . & . & . \\ 1 & g_{1n} & \ldots & g_{kn} \end{pmatrix} \qquad (1.4)$$

$$g_{ij} = x_{ij} + \varepsilon_{ij} \quad \eta_j = y_j + \varepsilon_{yj} \quad i = \overline{1,k} \quad j = \overline{1,n} \qquad (1.5)$$

Чем больше число входных сигналов, тем больше вероятность плохой обусловленности матрицы R_{gg}, связанной с близостью одного или несколько ее собственных чисел к нулю, и следовательно, с близостью к нулю определителя матрицы R_{gg}. Так как для решения уравнения $R_{gg} b = R_{g\eta}$ (*(1.3)* необходимо вычислить матрицу R_{gg}^{-1}, элементы которой определяются как алгебраические дополнения элементов $R_{g_i g_j}(\mu)$ матрицы

R_{gg}, деленной на определитель матрицы R_{gg}, то небольшие ошибки в вычислениях могут вызвать существенные погрешности в вычисляемых элементах R_{gg}^{-1}, а следовательно, и в оценках b.

С увеличением общей дисперсии σ^2 помехи при плохой обусловленности матрицы R_{gg} резко увеличивается дисперсия оценок коэффициентов [11].

Для получения более устойчивых оценок при плохой обусловленности матрицы R_{gg} методом регуляризации матрица R_{gg} заменяется на матрицу $R_{gg}^r = R_{gg} + rE$, то есть все диагональные элементы увеличиваются на r - пара-метр регуляризации).

Определение параметра регуляризации, способного дать лучшие оценки, требует знания оценки дисперсии σ^2 и вектора неизвестных коэффициентов β. Подстановка вместо них неточных оценок делает выбор стохастическим и не гарантирует снижения квадратичной ошибки коэффициентов регрессии $\sum_{i=1}^{k}\left(b_i^r - \beta_i\right)^2$.

В данной работе ставится задача нахождения оценки дисперсии помехи σ^2 для определения эффективного параметра регуляризации r и вычисления регрессионных коэффициентов, обеспечивающих адекватность математической модели.

2. Алгоритм определения оценок дисперсий сигналов. Предположим, что g-центрированный дискретизированный стационарный случайный сигнал с нормальным законом распределения, состоящий из сигнала x и помехи ε с математическим ожиданием, близким к нулю. С учетом влияния помехи автокорреляционную функцию R_{gg} можно представить в виде:

$$R_{gg}(\mu) = R_{xx}(\mu) + \lambda(\mu) \qquad (2.1)$$

где $\lambda(\mu) = \frac{1}{n}\sum_{\nu=1}^{n}\left(x_\nu \varepsilon_{\nu+\mu} + \varepsilon_\nu x_{\nu+\mu} + \varepsilon_\nu \varepsilon_{\nu+\mu}\right)$ - погрешность корреляционной функции $R_{gg}(\mu)$.

Значения ε_ν и $\varepsilon_{\nu+\mu}$ поиехи коррелируют между собой, только если $\mu = 0$ [11]. Следовательно

$$\frac{1}{n}\sum_{\nu=1}^{n}\varepsilon_\nu \varepsilon_{\nu+\mu} \approx 0, \quad \mu \neq 0; \qquad \frac{1}{n}\sum_{\nu=1}^{n}\varepsilon_\nu^2 = \sigma^2 \qquad (2.2)$$

С учетом *(2.2)* равенство *(2.1)* можнозаписать в виде

$$R_{gg}(\mu = 0) = \frac{1}{n}\sum_{\nu=1}^{n}\left(x_\nu^2 + 2\varepsilon_\nu x_\nu\right) + \sigma^2$$

$$R_{gg}(\mu \neq 0) = \frac{1}{n}\sum_{\nu=1}^{n}\left(x_\nu x_{\nu+\mu} + x_\nu \varepsilon_{\nu+\mu} + \varepsilon_\nu x_{\nu+\mu}\right) \tag{2.3}$$

Эти оценки при $\mu = 0, \Delta t, 2\Delta t$ и шаге дискретизации $\Delta t \leq \frac{f_c}{4 \div 6}$ становятся близкими величинами, и их разность оказывается соизмеримой с шагом квантования по уровню величины Δx, которая определяется по разрешающей способности измерительного прибора, например, для аналого-цифрового преобразователя (АЦП) она равна весу младшего разряда [3, 14, 15]. При стремлении времени наблюдения T к бесконечности и шага дискретизации Δt к нулю оценки $R_{xx}(\mu = 0), R_{xx}(\mu = \Delta t), R_{xx}(\mu = 2\Delta t)$, настолько близки между собой, что можно считать справедливым неравенство:

$$[R_{xx}(\mu = 0) - R_{xx}(\mu = \Delta t)] - [R_{xx}(\mu = \Delta t) - R_{xx}(\mu = 2\Delta t)] << \Delta x \tag{2.4}$$

В случае выполнения условия *(2.4)* при отсутствии корреляции между сигналом и помехой можно написать

$$\sigma^2(\varepsilon) \approx R_{gg}(\mu = 0) + R_{gg}(\mu = 2) - 2R_{gg}(\mu = 1) \tag{2.5}$$

В более общем случае, когда между сигналов и помехой имеется корреляция, необходимо учесть влияние помехи на положительные и отрицательные произведения отсчетов при вычислении оценок по выражению $\frac{1}{n}\sum g_\nu g_{\nu+\mu}$.

Принимая приближенные равенства

$$\frac{1}{n_{2+}}\sum_{g_\nu g_{\nu+2}>0}\left(x_\nu \varepsilon_\nu + \varepsilon_\nu x_\nu\right) \approx \frac{1}{n_{2+}}\sum_{g_\nu g_{\nu+2}>0}\left(x_\nu \varepsilon_{\nu+1} + \varepsilon_\nu x_{\nu+1}\right) \approx \frac{1}{n_{2+}}\sum_{g_\nu g_{\nu+2}>0}\left(x_\nu \varepsilon_{\nu+2} + \varepsilon_\nu x_{\nu+2}\right) \tag{2.6}$$

которые выполняются для реальных сигналов, а также учитывая равенство *(2.4)* и то, что

$$\frac{1}{n_{2+}}\sum_{g_\nu g_{\nu+2}} \varepsilon_\nu^2 \approx \frac{1}{n}\sum_{\nu=1}^{n} \varepsilon_\nu^2 \approx \sigma^2(\varepsilon) \tag{2.7}$$

можно записать следующее соотношени

$$\frac{1}{n_{2+}}\sum_{g_\nu g_{\nu+2}>0}\left(x_\nu^2 + 2x_\nu \varepsilon_\nu\right) - \frac{1}{n_{2+}}\sum_{g_\nu g_{\nu+2}>0}\left(x_\nu x_{\nu+1} + x_\nu \varepsilon_{\nu+1} + \varepsilon_\nu x_{\nu+1}\right) =$$

$$= \frac{1}{n_{2+}}\sum_{g_\nu g_{\nu+2}>0}\left(x_\nu x_{\nu+1} + x_\nu \varepsilon_{\nu+1} + \varepsilon_\nu x_{\nu+1}\right) - \frac{1}{n_{2+}}\sum_{g_\nu g_{\nu+2}>0}\left(x_\nu x_{\nu+2} + x_\nu \varepsilon_{\nu+2} + \varepsilon_\nu x_{\nu+2}\right) = \tag{2.8}$$

или, что то же самое

$$\sigma^2(\varepsilon) = \frac{1}{n_{2+}}\sum_{g_\nu g_{\nu+2}}\left(g_\nu^2 - 2g_\nu g_{\nu+1} + g_\nu g_{\nu+2}\right) \tag{2.9}$$

В таблицах приводятся результаты вычислительного эксперемента, демонстрирующего зависимость точности оценки дисперсии помехи от выбора шага дискретизации. В качестве примера брался сигнал в виде суммы синосоиды и экспоненты $x = 100 + 100 \cdot \exp(-t) + 40 \cdot \sin t$, к которому добавлялась помеха- нормально распределенная случайная величина с дисперсией и нулевым математическим ожиданием.

Из таблицы 1 видно, что условие *(2.4)* реально выполняется для данного сигнала лишь при $\Delta t \le \pi/200$ (6 и 7 строки в 7 в столбце). Только при таком шаге дискретизации оценки дисперсии помехи (6 и7 строки 7 и 8 колонок в таблице 2) становятся близкими к действительной величине (приблизительно 49, 5). Остается добавить, что из-за наличия небольшой корреляции между сигналом и помехой оценка дисперсии помехи в последней колонке, полученная с помощью формулы *(2.9)*, в которой учитывается корреляция, выглядит лучше оценки в предпоследней колонке, полученной с помощью выражения *(2.5)*, где корреляция не учитывается.

3.Определение параметра регуляризации при помощи оценок дисперсии помехи сигналов. Алгоритм определения дисперсий помех случайных сигналов позволяет нам оценить общую дисперсию помехи в правой части уравнения *(1.1)*.

Во многих случаях для неизвестных коэффициентов выполняются соотношения

$$\beta_i = b_i^{(0)} - \Delta b_i \div b_i^{(0)} + \Delta b_i, \qquad (3.1)$$

где $b_i^{(0)}$- математическое ожидание априорной оценки β_i, а Δb_i- максимум отклонения параметра $b_i^{(0)}$ от β_i. Поэтому в качестве оценки общей дисперсии помехи можно принять величину

$$s^2 = \sigma^2(\varepsilon_y) + \sum_{i=1}^{k} \left(b_i^{(0)}\right)^2 \cdot \sigma(\varepsilon_i) \qquad (3.2)$$

Имея такую оценку, целесообразнее осуществить решение данной задачи при помощи итерационного улучшения параметра регуляризации. Начальное значение параметра регуляризации целесообразно выбирать по формуле [12]:

$$r^{(0)} = \frac{s^2}{s_\beta^2}, \quad s_\beta \approx \frac{\Delta b_{\max}}{3}, \quad \Delta b_{\max} = \max_{i=1,k} \Delta b_i \qquad (3.3)$$

Используя величину параметра регуляризации, производится очередной цикл итерации:

$$b^{(j)} = \left(R_{gg} + r^{(j-1)} \cdot E\right)^{-1} R_{g\eta} \qquad (3.4)$$

$$\left(s_{OCT}^{(j)}\right)^2 = \frac{\sum_{v=1}^{n}\left(y_v - \sum_{i=1}^{k} b_i^{(j)} x_{iv}\right)}{n-k}, \quad F^{(j)} = \frac{\left(s_{OCT}^{(j)}\right)^2}{s^2} \qquad (3.5)$$

$$r^{(j)} = \frac{s^2}{\left(\hat{\alpha}_{\max}^{(j)}\right)^2}, \quad \hat{\alpha}_{\max}^{(j)} = \max_{i=1,k}\left\{\left|\hat{\alpha}_i^{(j)}\right|\right\} \quad \hat{\alpha}^{(j)} = Vb^{(j)}, \quad j = 1,2,\ldots \qquad (3.6)$$

Этот процесс продолжается до тех пор, пока не будет выполнено условие адекватности $F^{(j)} \le F_{\acute{0}}$, $F_T = F(\alpha, n-k, k-1)$.

Вычисление по формуле *(3.6)* требует дополнительных вычислений. Поэтому вместо этой формулы лучше воспользоваться менее точной, но близкой оценкой через сами коэффициенты [13]:

$$r^{(j)} = \frac{s^2 k}{\sum_{i=1}^{k} \left(b_i^{(j)}\right)^2}. \qquad (3.7)$$

Таблица 1

Корреляционная функция неискаженного сигнала

Δt	$R(0)$	$R(1)$	$R(2)$	$R(0)-R(1)$	$R(1)-R(2)$	$\Delta R_1 - \Delta R_2$
$\pi/10$	1034.223	959.849	816.239	74.374	143.61	-69.236
$\pi/20$	1055.771	1025.767	978.616	30.004	47.151	-17.147
$\pi/40$	1067.061	1053.716	1036.195	13.345	17.521	-4.176
$\pi/50$	1069.361	1058.939	1045.863	10.422	13.076	-2.654
$\pi/100$	1074.002	1073.922	1071.339	0.08	2.583	-2.503
$\pi/200$	1076.343	1073.922	1071.339	2.421	2.583	-0.162
$\pi/400$	1077.519	1076.323	1075.087	1.196	1.236	-0.040

Таблица 2

Корреляционная функция искаженного сигнала и оценки дисперсии помехи

Δt	$R(0)$	$R(1)$	$R(2)$	$R(0)-R(1)$	$R(1)-R(2)$	$\Delta R_1 - \Delta R_2$	$\sigma^2(\varepsilon)$
$\pi/10$	1140.985	1012.869	860.218	128.116	152.651	-24.535	(-20.775)
$\pi/20$	1208.877	1130.984	1979.229	77.893	51.755	26.138	(029.822)
$\pi/40$	1233.055	1162.994	1145.263	70.061	17.731	52.33	(073.872)
$\pi/50$	1216.603	1155.500	1145.028	60.503	10.472	50.031	(054.183)
$\pi/100$	1223.946	1172.666	1164.288	51.28	8.378	42.902	45.277
$\pi/200$	1227.086	1176.160	1173.534	50.926	2.626	48.3	50.432
$\pi/400$	1228.604	1179.909	1177.848	48.695	2.061	46.634	48.585

Литература

[1]. Цыпкин Я.З. *Основы информационной теории идентификации.*- М.:Наука, 1984.

[2]. Солодовников В.В., Бирюков В.Ф., Тумаркин В.И. *Принципы сложности в*

теории управления.- М.: Наука, 1977.
[3]. Алиев Т.А. *Эксперементальный анализ.*-М.: Машиностроение, 1991.
[4]. Бессонов А.А., Загашвили Ю.В., Маркелов А.С. *Методы и средства идентификации динамических объектов.*- Ленинград: Энергоатомиздат, 1989.
[5]. Дейч А.М. *Методы фиксации динамических объектов.*-М.: Энергия, 1974.
[6]. Эйкхофф П. *Основы фиксации систем управления.* – М.: Мир, 1975.
[7]. Алиев Т.А., Мусаева Н.Ф. *Статическая идентификация с уравновешиванием погрешностей.* Известия РАН. Теория и системы управления, 1995, №3.
[8]. Тихонов А.Н., Арсенин В.Я. *Методы решения некорректных задач.*-М.: Наука, 1974.
[9]. Тихонов А.Н., Гончаровский А.В., Степанов В.В., Ягола А.Г. *Регуляризирующие алгоритмы и априорная информация.* –М.: Наука, 1983.
[10]. Hoerl A., Kennard R. *Ridge regression: based estimation for nonorthogonal problems.*-Technometrics, v. 12, 1970.
[11]. Вучков И., Бояджиева Л., Солаков Е. *Прикладной линейный регриссионный анализ.*- М.: Финансы и статистика, 1987.
[12]. Goldstein M., Smith A. *Ridge- type estimators for regression analysis.*-J. Roy. Stat. Soc.- Ser B, v.36, 1974.
[13]. Hoerl A., Kennard R., Baldwin K. *Ridge regression: some simulations.*- Comm. Stat., v. 4, 1975
[14]. Цапенко М.П. *Измерительные информационные системы: Структуры и алгоритмы, системотехническое проектирование.*-М.: Энергоатомиздат,1985.
[15]. Куликов Е.И. *Методы измерения случайных процессов.* М.: Радио и связь, 1986.

Əliyev T.A., Əmirov Z.Ə.

STATİSTİK İDENTİFİKAİYA MƏSƏLƏSİNİN HƏLLİ ÜÇÜN REQULYARLAŞMA PARAMETRİNİN SEÇİLMƏSİ

Siqnal əngəli dispersiyası qiymətinin təyin edilməsi alqoritmi təsvir edilir. Tapılmış dispersiyanın köməyi ilə statistik identifikasiya məsələlərinin həlli üçün requlyarlaşma parametrinin seçilməsi iterasiya prosesi düzdəldilir.

Aliyev T.A., Amirov Z.A.

BRIDGE- REGRESSION PARAMETER CHOICE IN STATISTICAL IDENTIFICATION PROBLEM DECISION

Signal noise dispersion estimation determination algorithm is described. With help of bridge- regression parameter selection for statistical identification problems decision is constructed.