АКАДЕМИЯ НАУК СССР
СИБИРСКОЕ ОТДЕЛЕНИЕ
АВТОМЕТРИЯ
№ 4 1972

УДК 621.396.983+378.9

Б. М. ПУШНОЙ
(Новосибирск)

К ТЕОРИИ ИЗМЕРИТЕЛЬНЫХ СИСТЕМ С ПЕРИОДИЧЕСКИМИ ФУНКЦИЯМИ ПРЕОБРАЗОВАНИЯ

Среди измерительных систем особое место занимают такие, у которых функция преобразования $z=f(y)$, связывающая регистрируемые значения z с измеряемыми величинами y, является периодической. К таким системам относятся в первую очередь фазовые. Область практического использования фазовых систем в последнее время значительно расширилась в связи с развитием когерентной оптики.

Периодическая функция преобразования представима рядом Фурье

$$z(y) = \sum_{k=0}^{s} c_k \sin(k\omega y + \varphi_k), \tag{1}$$

причем диапазон измеряемой величины во много раз больше периода функции преобразования

$$\theta = y_{\max} - y_{\min} \gg \tau = \frac{2\pi}{\omega},$$

так что функция $y(z)$ неоднозначна. В свою очередь,

$$y = F(a_0, a_1, \ldots, a_{m-1}, x). \tag{2}$$

В большинстве случаев (2) может быть представлено в линейной форме

$$Y = XA. \tag{3}$$

Задача измерения обычно состоит в получении оценок $\hat{Y}$ или $\hat{A}$ по результатам регистрации $z^*(x)$ в n равноотстоящих точках интервала $T = x_n - x_1$ при наличии случайных погрешностей измерения ξ, которые можно считать аддитивными по отношению к z:

$$\xi_i = z_i^* - z_i \quad \left(i = \overline{1, n}\right).$$

Задачи такого типа обычно решаются итерационными процедурами, базирующимися на линейных методах параметрического оценивания. При этом приходится сталкиваться с двумя специфическими особенностями. Одна из них связана с наличием локальных минимумов функционала

$$\xi^2 = \sum_{i=1}^{n} \left(\hat{z}_i - z_i\right)^2, \tag{4}$$

из-за которых итерационная процедура может с некоторой вероятностью сходиться к неправильным оценкам. Другая связана с неоднозначностью функции $y(z)$ и с необходимостью получать однозначные оценки $\hat{Y}$. Получение однозначных оценок требует использования дополнительных результатов измерения, которые также содержат погрешности, так что на практике всегда приходится считаться с вероятностью неправильного «разрешения» многозначности. Таким образом, в системах с периодическими функциями преобразования могут возникать ошибки двух типов. Значения вероятностей этих ошибок определяют в каждом конкретном случае возможность практического использования этих систем. Приводимые ниже соотношения позволяют оценить вероятности ошибок и в какой-то мере охарактеризовать потенциальные возможности измерительных систем с периодическими функциями преобразования.

Можно показать, что, если n достаточно велико и функция $y(t)$ на интервале T изменяется в широких пределах, имеет место следующее приближенное равенство:

$$\sum_{i=1}^{n}\left[\sum_{k=1}^{s} c_k \sin(k\omega y_i + \varphi_k)\right]^2 \cong \frac{n}{2}\sum_{k=1}^{s} c_k^2.$$

Отсюда следует, что пространство параметров в евклидовом пространстве R_n близко к $(m+1)$-мерной сфере с радиусом

$$r = \sqrt{\frac{n}{2}\sum_{k=1}^{s} c_k^2}.$$

К некоторой точке сферы, соответствующей истинному значению параметров $a_0, \ldots, a_{m-1}$, приложен n-мерный вектор случайной погрешности Ξ с компонентами ξ_i. Оценки $\hat{a}_0, \ldots, \hat{a}_{m-1}$ определяются угловыми координатами вектора $\hat{A}$, представляющего собой сумму вектора истинных значений A и проекции Ξ_{m+1} вектора случайной погрешности на $(m+1)$-мерное подпространство, в котором располагается сфера. Если

$$\sigma_{\xi_i} \ll r,$$

погрешности оценок параметров будут связаны с вектором Ξ практически линейно. Отсюда непосредственно следует, что итерационные процедуры, основанные на линейных методах оценивания, дают асимптотически оптимальные оценки параметров.

Надлежащим выбором базиса на сфере оценки можно сделать независимыми. Тогда при некоррелированных компонентах Ξ и ортонормированном базисе имеет место следующее очевидное соотношение:

$$\sigma_{\hat{a}} \geqslant \sigma_{\xi}/\omega r. \qquad (5)$$

Получить оценки с меньшей дисперсией невозможно. Нетрудно заметить, что оценки $\hat{a}$ получаются смещенными, если на интервале T укладывается не целое число периодов. Если отношение σ_{ξ}/r близко к единице, оценки $\hat{a}$ будут статистически зависимыми при любом базисе на сфере, поскольку в оценку каждого параметра будет входить компонента вектора Ξ, направленная параллельно вектору A. При нормальном векторе Ξ совместное распределение оценок будет описываться m-мерным распределением угловых координат вектора с нормальными компонентами.

В этих условиях становится существенной вероятность получения ошибочных оценок. Эту вероятность можно приближенно оценить, полагая, что ошибка произойдет, если компонента Ξ, направленная от конца вектора A к его началу, превысит длину этого вектора, т. е.

$$P_{\text{ош}} \cong P\{\xi > r\}. \qquad (6)$$

Нетрудно видеть, что имеется возможность повышать достоверность оценок увеличением числа отсчетов на обрабатываемом интервале.

Применение итерационных алгоритмов оценивания связано с необходимостью задавать первое приближение. Естественно, вектор первого приближения $\hat{A}^*$ должен находиться в пределах полусферы, окружающей вектор A. В противном случае итерационный процесс может сойтись к локальному минимуму функционала (4). Действительно, если аргументы двух функций z, входящих в (4), существенно различны, эти функции практически ортогональны. Им соответствуют два взаимно ортогональных вектора, направленных из центра сферы к ее поверхности. Более подробный анализ функции $\varepsilon^2(\hat{a}_0, \ldots, \hat{a}_{m-1})$ показывает, что она равна нулю на счетном множестве точек с координатами

$$a_0 \pm k\tau,\ a_1,\ \ldots,\ a_{m-1} \quad (k = 0, 1, 2, \ldots).$$

В окрестностях этих точек ε монотонно возрастает, но нигде не превосходит

$$\sqrt{2n \sum_{k=1}^{s} c_k^2} = 2r.$$

При всех прочих значениях $\hat{a}$ функция колеблется около значения $r\sqrt{2}$, образуя множество локальных экстремумов.

Можно получить приближенное представление о форме и размерах области в линейном пространстве параметров a, в которой должно находиться первое приближение, обеспечивающее с вероятностью, определяемой выражением (6), выход итерационного процесса к главному минимуму. Нетрудно показать, что в окрестностях главных минимумов функция (4) близка к эллиптическому параболоиду, в частности, при ортонормированном полиноме (2) — к параболоиду вращения

$$\varepsilon^2 = r^2\omega^2 \sum_{j=0}^{m-1} a_j^2 .$$

Сечение этого параболоида гиперплоскостью $\varepsilon = r\sqrt{2}$ представляет собой m-мерную сферу радиуса $\sqrt{2}/\omega$. Этот радиус характеризует практические требования, предъявляемые к методам получения первого приближения. Если используются какие-либо беспоисковые методы, зная радиус области, можно оценить вероятность получения правильных оценок. При использовании поисковых методов, располагая априорными данными о диапазоне изменения параметров a_j, можно охарактеризовать сложность задачи поиска отношением объема пространства, в котором ведется поиск, к общему объему сфер:

$$\frac{\prod_{j=0}^{m-1} (a_{j\max} - a_{j\min})}{\dfrac{a_{0\max} - a_{0\min}}{\tau} V_c} = \frac{\tau \prod_{j=1}^{m-1} (a_{j\max} - a_{j\min})}{V_c}.$$

Это выражение показывает, что диапазон изменения a_0 независимо от априорных сведений об измеряемой функции следует задавать равным τ. Иногда необходимо получать однозначные результаты измерений. Это может быть достигнуто с помощью дополнительных измерительных систем. Представляет интерес случай, когда дополнительные системы также имеют периодические функции преобразования.

Основная задача состоит в выборе таких соотношений между периодами функций преобразования, которые обеспечивают наибольшую вероятность правильного разрешения многозначности.

Будем рассматривать комплексную измерительную систему, в которую входит основная с периодом τ_0 и p дополнительных с различными периодами $\tau_1, \ldots, \tau_l, \ldots, \tau_p$, причем $\tau_l > \tau_0$. Для упрощения записи положим $\tau_0 = 1$.

Однозначная оценка параметра a_0 для каждой системы выражается так:

$$\hat{\hat{a}}_0 = a_{0l} + \hat{N}_l \tau_0 \quad \left(l = \overline{1, p}\right). \tag{7}$$

Коэффициенты N_l являются целочисленными. Если погрешности измерений отсутствуют, можно записать систему уравнений

$$\hat{a}_{00} - \hat{a}_{0l} = N_0 - N_l \tau_l \quad \left(l = \overline{1, p}\right), \tag{8}$$

которая содержит одно неизвестное N_0, так как между N_0 и всеми N_l имеется функциональная связь. Она обеспечивается тем, что периоды τ и начальные фазы φ_l периодических функций являются заданными. Связь между N_l выражается в параметрической форме равенствами (7), если считать $\hat{\hat{a}}_0$ независимой переменной и заметить, что значения a_{0l} всегда лежат в пределах $[0, \tau_l)$. Будем рассматривать правые части уравнений (8) как ортогональные координаты p-мерного пространства. Каждому значению $\hat{\hat{a}}_0$ будет соответствовать некоторая точка пространства. Поскольку правые части уравнений могут приобретать лишь дискретные значения, в пространстве будет определена таким образом некоторая решетчатая функция, каждой точке которой соответствует целочисленное значение N_0. Рассматривая (7) и (8), нетрудно заметить, что независимо от диапазона изменения $\hat{\hat{a}}_0$ все точки, соответствующие значениям N_0, относящимся к этому диапазону, лежат в пределах параллелотопа со сторонами τ_l $\left(l = \overline{1, p}\right)$. Отсюда непосредственно следует, что многозначность в принципе разрешима, если $\hat{\hat{a}}_0$ изменяется в конечном диапазоне и никакие две точки с различными значениями N_0 не совмещаются. Последнее условие обеспечивается надлежащим выбором τ_l.

Однако для достижения максимальной вероятности правильного разрешения необходимо соблюдение некоторых дополнительных условий. Рассмотрим точки параллелотопа, соответствующие левым частям уравнений (8). При отсутствии погрешностей измерения они будут совпадать с точками N_0. При наличии погрешностей задача разрешения многозначности состоит в нахождении точки решетчатой функции, которая в соответствии с некоторой мерой является ближайшей к точке с координатами

$$\hat{\alpha}_{0l} = \hat{a}_{00} - \hat{a}_{0l}. \tag{9}$$

Метод максимума правдоподобия приводит к мере

$$d^2 = \left(\hat{A}_0 - H\right)^T B^{-1} \left(\hat{A}_0 - H\right), \tag{10}$$

где H — вектор координат точки N_0; $\hat{A}_0$ — вектор оценок $\hat{\alpha}_{0l}$; B — корреляционная матрица оценок $\hat{\alpha}_{0l}$. При фиксированном d выражение (10) можно рассматривать как уравнение p-мерного эллипсоида с центром в точке N_0. Минимальная средняя вероятность ошибки будет достигаться в том случае, если точки N_0 будут располагаться в центрах равновеликих эллипсоидов (10), заполняющих параллелотоп плотнейшим образом без взаимного пересечения. Так как эллипсоиды ориентированы одинаково, они аффинным преобразованием переводятся в сферы, так что задача сводится к плотнейшей решетчатой укладке сфер. При решении практических задач придется считаться с тем, что в общем случае решетчатая функция N_0 может не совпадать с решеткой, соответствующей плотнейшей укладке эллипсоидов, что теоретический предел, основанный на плотнейшей укладке, может оказаться недостижимым. Однако в случае $p=2$ выбором τ_l можно реализовать два варианта плотнейшей укладки, в чем легко убедиться непосредственным построением.

Чтобы оценить предельные возможности комплексных систем с периодическими функциями преобразования, заметим прежде всего, что в соответствии с выражениями (7) и (8) расстояние между соседними точками решетчатой функции N_0 не может превышать $\sqrt{p}$, причем любые две такие точки лежат на прямой, параллельной главной биссектрисе p-мерного координатного угла, так что при любых τ_l расстояние между двумя противоположными точками эллипсоида не может быть больше $\sqrt{p}$. Поэтому диаметр d_c сферы, к которой приводится эллипсоид аффинным преобразованием G, будет, очевидно, равен длине вектора $W^*=GW$, где

$$W = \begin{vmatrix} 1 \\ 1 \\ \cdot \\ \cdot \\ \cdot \\ 1 \end{vmatrix}.$$

Таким образом,

$$d_c^2 = W^{*T}W^* = W^TG^TGW = W^TB^{-1}W = \sum b_{uv},$$

где B^{-1} — матрица, обратная корреляционной матрице оценок $\hat{\alpha}_{0l}$; $\sum b_{uv}$ — сумма всех ее элементов. Зная объем преобразованного параллелотопа

$$V^* = \prod_{l=1}^{p} (\tau_l) \det G,$$

нетрудно выразить максимальное количество эллипсоидов максимального размера, уложенных в параллелотопе:

$$N_{0\max} \leqslant \frac{\rho_p V^*}{V_c} = \frac{2^p \rho_p \prod_{l=1}^{p} (\tau_l) \det G}{\Omega_p \left(\sum b_{uv}\right)^{p/2}}, \tag{11}$$

где ρ_p — коэффициент заполнения пространства сферами при решетчатой плотнейшей укладке; V_c — объем сферы; Ω_p — угловая мера сферы. Если задан диапазон измерений, это выражение определяет инструментальную точность однозначного комплекса приборов с периодическими функциями преобразования, которую можно получить при минимальной вероятности ошибок разрешения многозначности. Повышение инструментальной точности при неизменном количестве дополнительных измерительных систем может быть достигнуто только за счет снижения вероятности правильного разрешения. Систему, характеристики которой удовлетворяют (11), следует считать оптимальной по надежности разрешения. Вероятность ошибки определяется вероятностью выхода вектора $\hat{A}_0$ из области, ограниченной гиперплоскостями, касательными к эллипсоиду в точках его соприкосновения с соседними. Для оптимальной системы можно указать простую верхнюю оценку вероятности ошибки

$$p_{\text{ош}} \leqslant N_{\text{c}} p \left\{ \xi > \frac{\sqrt{\Sigma b_{uv}}}{2} \right\}, \tag{12}$$

где N_{c} — количество эллипсоидов, соприкасающихся с данным; ξ — одномерная случайная компонента вектора $G\hat{A}_0$.

Анализ выражений (5), (9), (11) и (12) позволяет заключить, что вероятность ошибки зависит от соотношения между τ_l, причем минимум вероятности ошибки достигается в том случае, когда значения τ_l, соответствующие оптимальной решетчатой функции, мало различаются, т. е. когда параллелотоп близок к кубу.

Поступила в редакцию
15 марта 1972 г.