А К А Д Е М И Я Н А У К С С С Р
СИБИРСКОЕ ОТДЕЛЕНИЕ
А В Т О М Е Т Р И Я

№ 3 1965

УДК 62—503

В. А. ВИТТИХ, А. Н. ГИНЗБУРГ
(Новосибирск)

ОПТИМАЛЬНАЯ ДИСКРЕТИЗАЦИЯ ИЗМЕРИТЕЛЬНЫХ СИГНАЛОВ*

Рассматривается задача оптимальной дискретизации непрерывных измерительных сигналов, которая сводится к минимизации некоторого функционала ошибки, и ее решение обычным методом классического анализа и методом динамического программирования.

Уменьшение объема измерительной информации — актуальнейшая проблема современной телеметрии. В последнее время появилось много работ, посвященных исследованию одного из эффективных методов сжатия информации — обобщенной адаптивной дискретизации как с экстраполяцией, так и с интерполяцией сигнала [1]. Однако до сих пор алгоритм оптимальной дискретизации не был найден.

Интерполяционные алгоритмы адаптивной дискретизации позволяют получать бо́льшие коэффициенты сжатия, чем экстраполирующие, и требуют задержки сигнала. Причем, чем больше допустимая задержка, тем бо́льшие коэффициенты сжатия могут быть получены. Это находится в соответствии с известным положением теории связи [2]: введение достаточно длительных задержек в передатчике и приемнике позволяет приблизиться к оптимальному (в смысле Шеннона) кодированию, а в ряде случаев существенно снизить мощность, требуемую для передачи сигнала [3].

В статье решаются задачи оптимальной (в принятом смысле) дискретизации сигналов, которые сводятся к минимизации некоторого функционала ошибки. Предполагается, что сигнал известен полностью, и это позволяет рассматривать его как детерминированную функцию времени.

* * *

Обозначим через $G_M = G_M[\alpha, \beta]$ пространство функций $g(t)$, ограниченных на интервале $[\alpha, \beta]$ и имеющих на нем ограниченную $(n+1)$-ю производную

$$\left| g^{(n+1)}(t) \right| \leqslant M_{(n+1)}.$$

Задача оптимальной дискретизации заключается в минимизации числа некоторых обобщенных координат, которые являются дискретным изображением заданной функции $g(t) \in G_M$ и по которым последняя

* Материал доложен на VI Всесоюзной конференции по автоматическому контролю и методам электрических измерений в сентябре 1964 г. в Новосибирске.

может быть восстановлена с заданной точностью (в смысле введенной в пространстве G_M метрики).

Эффективность дискретизации может быть оценена по отношению к ε-энтропии H_ε [4], по существу, представляющей собой минимальное число двоичных символов, которым функция $g(t)$ может быть описана с точностью до величины ε. Другими словами, задача оптимальной дискретизации заключается в отыскании минимума H_ε (основная задача).

Указанная задача решается методом последовательных приближений. На каждом шаге задается некоторое число обобщенных координат q и определяется соответствующий минимум ошибки ε_q. Процесс продолжается до тех пор, пока ошибка ε_q не станет равной допустимой. Таким образом, решение основной задачи можно свести к нахождению минимума ε_q при заданном q. В дальнейшем мы будем рассматривать круг вопросов, связанных с решением этой задачи.

Пусть

$$g(t) \simeq \varphi(t) = \begin{cases} \sum\limits_{i=0}^{m} a_{i1}\psi_{i1}(t), & \alpha = u_0 \leqslant t \leqslant u_1; \\ \cdots\cdots\cdots\cdots & \\ \sum\limits_{k=0}^{n} a_{kj}\psi_{kj}(t), & u_{j-1} < t \leqslant u_j; \\ \cdots\cdots\cdots\cdots & \\ \sum\limits_{l=0}^{p} a_{lN+1}\psi_{lN+1}(t), & u_N < t \leqslant u_{N+1} = \beta, \end{cases} \tag{1}$$

т. е. на интервале $[\alpha, \beta]$ функция $g(t)$ представляется набором координат:

$$\{u_j;\ \psi_j;\ a_{i1},\ \ldots,\ a_{kj},\ \ldots,\ a_{lN+1}\};$$
$$i = 0,\ 1, \ldots, m;$$
$$k = 0,\ 1, \ldots, n;$$
$$\cdots\cdots$$
$$l = 0,\ 1, \ldots, p;$$
$$j = 0,\ 1, \ldots, N+1.$$

Рассмотрим функционал

$$I = \int_{\alpha}^{\beta} [g(t) - \varphi(t)]^2\, dt, \tag{2}$$

который представляет собой квадрат расстояния между элементами $g(t)$ и $\varphi(t)$ пространства L_2 и может быть использован для оценки степени близости этих функций. При заданном числе координат необходимо найти такие их значения, которые обеспечивают минимум функционала (2). Однако решение этой задачи в самом общем случае, когда приходится варьировать всеми координатами, оказывается чрезвычайно сложным. Поэтому мы рассмотрим лишь два наиболее важных частных случая.

Первый случай ($j=1$).
Соотношение (1) примет вид

$$g(t) \simeq \varphi(t) = \sum_{i=0}^{n} a_i \psi_i(t), \quad u_0 = \alpha \leqslant t \leqslant \beta = u_1,$$

и решение задачи сведется к выбору такой системы базисных функций $\{\psi_i(t)\}$, которая обеспечивала бы минимум функционала (2), а это, как известно [5], имеет место, когда в качестве базиса разложения выбрана система функций, ортогональных на отрезке $[\alpha, \beta]$, например полиномов Лежандра.

Заметим, что решение последней задачи может расматриваться как частный случай решения более общей задачи, а именно, когда $\varphi(t)$ обладает тем свойством, что $\varphi^{(n)}(t)$ удовлетворяет условию Липшица с константой γ ($n \geqslant 0$, $\gamma \geqslant 0$).

В [6] показано, что если $g(t)$ имеет непрерывную 1-ю производную, то для того, чтобы функция $\varphi(t)$ являлась решением задачи (т. е. чтобы обеспечивала минимум функционала (2)), необходимо, чтобы почти всюду на отрезке $[\alpha, \beta]$ выполнялось одно из соотношений:

$$\int_{\alpha}^{t} (\xi - t)^n [g(\xi) - \varphi(\xi)]\, d\xi = 0;$$

$$\varphi^{(n+1)}(t) = \gamma \operatorname{sign} \left\{ \int_{\alpha}^{t} (\xi - t)^n [g(\xi) - \varphi(\xi)]\, d\xi \right\}$$

и, кроме того, чтобы для любого многочлена $P(t)$ степени, меньшей или равной n, выполнялось условие

$$\int_{\alpha}^{\beta} P(t) [g(t) - \varphi(t)]\, dt = 0.$$

В частном случае, когда $\gamma = 0$, мы и приходим к задаче нахождения такого многочлена $\varphi(t)$ n-й степени, который на отрезке $\alpha \leqslant t \leqslant \beta$ имеет наименьшее квадратичное уклонение от заданной функции $g(t)$.

Однако условие $j=1$ в ряде случаев при решении основной задачи приводит к значительному числу членов разложения функции $g(t)$ на интервале $[\alpha, \beta]$ и появлению дополнительных трудностей, связанных с вычислением коэффициентов при полиномах высших степеней. По-видимому, целесообразнее отыскивать минимум функционала (2) при ограничениях, указанных ниже.

Второй случай ($n = \text{const}$, $\psi_j = \text{const}$, $j = 1, \ldots, N+1$).
Соотношение (1) примет вид

$$g(t) \simeq \varphi(t) = \begin{cases} \sum\limits_{i=0}^{n} a_{i1} \psi_i(t), & \alpha = u_0 \leqslant t \leqslant u_1; \\ \dots\dots\dots\dots\dots & \\ \sum\limits_{i=0}^{n} a_{ij} \psi_i(t), & u_j \leqslant t \leqslant u_{j+1}; \\ \dots\dots\dots\dots\dots & \\ \sum\limits_{i=0}^{n} a_{iN+1} \psi_i(t), & u_N \leqslant t \leqslant u_{N+1} = \beta. \end{cases} \qquad (3)$$

Задача сводится к минимизации функции

$$F(u_1, u_j, \ldots, u_N, a_{i1}, a_{i2}, \ldots, a_{ij}, \ldots, a_{iN+1}) = \\ = \sum_{j=1}^{N+1} \int_{u_{j-1}}^{u_j} \left[g(t) - \sum_{i=0}^{n} a_{ij} \psi_i(t) \right]^2 dt \tag{4}$$

по всем u_j и a_{ij}.

Воспользуемся обычным методом классического анализа. Возьмем частные производные от F по u_j и a_{ij} и, приравняв их нулю, будем решать систему из $[(N+1)(n+1)+N]$ уравнений для $[(N+1)(n+1)+N]$ неизвестных:

$$\begin{cases} \dfrac{\partial F}{\partial a_{ij}} = 0, & i = 0, 1, \ldots, n; \\ \dfrac{\partial F}{\partial u_j} = 0, & j = 1, 2, \ldots, N+1. \end{cases}$$

При большом числе неизвестных решение этой в общем случае нелинейной системы не может быть получено непосредственно. Поэтому обычно используются численные методы решения, например метод Ньютона [7].

Пусть $\psi_i(t) = t^i$ и $n=1$, т. е. необходимо приблизить функцию $g(t)$ на интервале $[\alpha, \beta]$ функцией вида

$$y = \begin{cases} a_1 + b_1 t, & \alpha = u_0 \leqslant t \leqslant u_1; \\ a_2 + b_2 t, & u_1 \leqslant t \leqslant u_2; \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ a_{N+1} + b_{N+1} t, & u_N \leqslant t \leqslant u_{N+1} = \beta. \end{cases} \tag{5}$$

Таким образом, следует минимизировать

$$F(a_1, \ldots, a_{N+1}, b_1, \ldots, b_{N+1}, u_1, \ldots, u_N) = \\ = \sum_{j=1}^{N+1} \int_{u_{j-1}}^{u_j} [g(t) - a_j - b_j t]^2 dt. \tag{6}$$

Эта задача была поставлена Г. Стоуном [8]. Решение ее сводится к решению $3N+2$ уравнений с $3N+2$ неизвестными $a_1, \ldots, u_N$:

$$\frac{\partial F}{\partial a_p} = 0 = -2 \int_{u_{p-1}}^{u_p} g(t)\, dt + 2a_p (u_p - u_{p-1}) + b_p (u_p^2 - u_{p-1}^2),$$

$$p = 1, 2, \ldots, N+1; \tag{7а}$$

$$\frac{\partial F}{\partial b_p} = 0 = -2 \int_{u_{p-1}}^{u_p} tg(t)\, dt + a_p (u_p^2 - u_{p-1}^2) + \frac{2}{3} b_p (u_p^3 - u_{p-1}^3),$$

$$p = 1, 2, \ldots, N+1; \tag{7б}$$

$$\frac{\partial F}{\partial u_p} = 0 = 2g(u_p)(a_{p+1} - a_p) + 2u_p g(u_p)(b_{p+1} - b_p) -$$

$$- 2u_p(a_{p+1}b_{p+1} - a_p b_p) - u_p^2(b_{p+1}^2 - b_p^2) - a_{p+1}^2 + a_p^2,$$

$$p = 1, 2, \ldots, N. \qquad (7в)$$

Допустим, что величины $u_1, \ldots, u_N$ выбраны. Тогда, решая совместно (7а) и (7б), получим

$$a_p = \frac{3}{(u_p - u_{p-1})^3}\{^4/_3(u_p^2 + u_p u_{p-1} + u_{p-1}^2) L(u_p, u_{p-1}) -$$

$$- 2(u_p + u_{p-1}) J(u_p, u_{p-1})\}; \qquad (8а)$$

$$b_p = \frac{6}{(u_p - u_{p-1})^3}\{2J(u_p, u_{p-1}) - (u_p + u_{p-1}) L(u_p, u_{p-1})\}, \qquad (8б)$$

где

$$L(u_p, u_{p-1}) = \int\limits_{u_{p-1}}^{u_p} g(t)\,dt;$$

$$J(u_p, u_{p-1}) = \int\limits_{u_{p-1}}^{u_p} \mathrm{t}g(t)\,dt. \qquad (8в)$$

Если величины $u_1, \ldots, u_N$ не выбраны заранее, то, подставляя a_p и b_p из (8) в (7в), получим систему N нелинейных уравнений относительно $u_1, \ldots, u_N$.

Таким образом, даже в этом довольно простом случае классический метод решения оказывается практически неприемлемым в связи со значительными трудностями, возникающими при решении системы N нелинейных уравнений.

Задача минимизации функции (4) может быть успешно решена при использовании метода динамического программирования [9], который позволяет свести решение поставленной многомерной задачи к последовательности одномерных, а именно к решению N функциональных уравнений.

Пусть

$$f_N(\beta) = \min_{(a_{ij}, u_j)} F(u_1, \ldots, u_N\ a_{i1}, \ldots, a_{ij}, \ldots, a_{iN+1}). \qquad (9)$$

В этом случае функциональные уравнения примут вид:

$$f_1(\beta) = \min_{[a_{i1}, a_{i2}, u_1]} \left[\int\limits_{\alpha}^{u_1} \left[g(t) - \sum_{i=0}^{n} a_{i1}\psi_i(t)\right]^2 dt + \right.$$

$$\left. + \int\limits_{u_1}^{\beta} \left[g(t) - \sum_{i=0}^{n} a_{i2}\psi_i(t)\right]^2 dt; \right. \qquad (10)$$

$$f_N(\beta) = \min_{\alpha < u_N < \beta} \left\{\min_{[a_{iN+1}]} \int\limits_{u_N}^{\beta} \left[g(t) - \sum_{i=0}^{n} a_{iN+1}\psi_i(t)\right]^2 dt + f_{N-1}(u_N)\right\}. \qquad (11)$$

В качестве примера рассмотрим приближение $g(t)$ функций вида (5). Эта задача была поставлена Беллманом в [10] и получила дальнейшее развитие в [11] и [12].

Запишем функциональные уравнения:

$$f_1(\beta) = \min_{(a_1, a_2, b_1, b_2, u_1)} \left\{ \int_{\alpha}^{u_1} [g(t) - a_1 - b_1 t]^2 dt + \int_{u_1}^{\beta} [g(t) - a_2 - b_2 t]^2 dt \right\}, \quad (12)$$

где $\alpha \leqslant u_1 \leqslant \beta$;

$$f_N(\beta) = \min_{\alpha < u_N < \beta} \left\{ \min_{[a_{N+1}, b_{N+1}]} \int_{u_N}^{\beta} [g(t) - a_{N+1} - b_{N+1} t]^2 dt + \right.$$
$$\left. + f_{N-1}(u_N) \right\}. \quad (13)$$

Численное решение этой задачи рассмотрим при дополнительном условии: прямая $a_j + b_j t$ совпадает с прямой $a_{j+1} + b_{j+1} t$ в точке u_j.

Обозначив через Y_j значение приближающей функции в точке u_j, можно записать:

$$y = Y_j + \frac{Y_j - Y_{j-1}}{u_j - u_{j-1}} (t - u_j), \quad u_{j-1} \leqslant t \leqslant u_j. \quad (14)$$

Теперь необходимо минимизировать функцию:

$$F(u_1, \ldots, u_N; Y_0, \ldots, Y_{N+1}) =$$
$$= \sum_{j=1}^{N+1} \int_{u_{j-1}}^{u_j} \left[g(t) - Y_j - \left(\frac{Y_j - Y_{j-1}}{u_j - u_{j-1}} \right) (t - u_j) \right]^2 dt \quad (15)$$

по всем u_j и Y_j.

Пусть

$$f_N(\beta) = \min_{[u_j, Y_j]} F(u_1, \ldots, u_N; Y_0, \ldots, Y_{N+1}); \quad (16)$$

$$h_N(\beta, \gamma) = \min_{[u_j, Y_j]} F(u_1, \ldots, u_N; Y_0, \ldots, Y_{N+1}; \gamma). \quad (17)$$

Другими словами, $h_N(\beta, \gamma)$ дает наилучшее приближение линейно-ломаной при дополнительном условии $Y_{N+1} = \gamma$ (введение γ необходимо для получения рекуррентных соотношений). Очевидно, что

$$f_N(\beta) = \min_{[\gamma]} h_N(\beta, \gamma). \quad (18)$$

Функциональные уравнения для h_N можно представить следующим образом:

$$h_0(\beta, \gamma) = \min_{[Y_0]} \left\{ \int_{\alpha}^{\beta} \left[g(t) - \gamma - \left(\frac{\gamma - Y_0}{\beta - \alpha} \right) (t - \beta) \right]^2 dt \right\}; \quad (19)$$

$$h_N(\beta, \gamma) = \min_{[u_N, Y_N]} \left\{ \int_{u_N}^{\beta} \left[g(t) - \gamma - \left(\frac{\gamma - Y_N}{\beta - u_N} \right) (t - \beta) \right]^2 dt + h_{N-1}(u_N, Y_N) \right\}. \quad (20)$$

Если ввести обозначения:

$$L(\beta, u_N) = \int_{u_N}^{\beta} g(t)\, dt;$$

$$J(\beta, u_N) = \int_{u_N}^{\beta} \mathrm{tg}(t)\, dt;$$

$$S(\beta, u_N) = \int_{u_N}^{\beta} [g(t)]^2\, dt, \quad (21)$$

то (19) и (20) можно записать так:

$$h_0(\beta, \gamma) = \min_{[Y_0]} \left[S(\beta, \alpha) + \frac{2(\alpha\gamma - \beta Y_0)}{\beta - \alpha} L(\beta, \alpha) - \right.$$
$$\left. - \frac{2(\gamma - Y_0)}{\beta - \alpha} J(\beta, \alpha) + \frac{4(\beta - \alpha)}{3} \left(\gamma^2 - \gamma Y_0 + Y_0^2 \right) \right]; \quad (22)$$

$$h_N(\beta, \gamma) = \min_{[u_N, Y_N]} \left[S(\beta, u_N) + \frac{2(u_N\gamma - \beta Y_N)}{\beta - u_N} L(\beta, u_N) - \right.$$
$$\left. - \frac{2(\gamma - Y_N)}{\beta - u_N} J(\beta, u_N) + \frac{4(\beta - u_N)}{3} \cdot \left(\gamma^2 + \gamma Y_N + Y_N^2 \right) + h_{N-1}(u_N, Y_N) \right]. \quad (23)$$

Рассмотрим численное решение поставленной задачи. Прежде всего выбирается сетка по γ и β, т. е. область возможных значений γ и интервал $[\alpha, \beta]$ разбиваются рядом равноотстоящих точек, число которых выбирается в зависимости от того, с какой точностью требуется получить решение.

Для различных значений β подсчитываем $h_0(\beta, \gamma)$ и соответствующие Y_0 по всем возможным γ. Для каждого γ уравнение (22) решается последовательными приближениями.

Получив $h_0(\beta, \gamma)$ h_1 и соответствующие u_1 и Y_1 по всей γ-сети вычисляем последовательными приближениями уравнения (23). Затем, используя (23), так же находим h_2, u_2, Y_2, h_3, u_3, Y_3 и т. д.

Из полученного множества значений h_N, используя (18), определяем f_N. Значение γ, минимизирующее $h_N(\beta, \gamma)$, есть искомое Y_{N+1}, а соответствующие u_N и Y_N находим из множеств h_N, u_N, Y_N. Аналогично «обратным ходом» определяем u_{N-1}, Y_{N-1}, u_{N-2}, Y_{N-2}, ..., u_1, Y_1 и, наконец, Y_0.

* * *

Таким образом, из всех изложенных методов оптимизации процесса адаптивной дискретизации наиболее приемлемым оказался, как и следовало ожидать, метод динамического программирования, дающий возможность получить численное решение поставленной задачи. Правда, его реализация требует применения универсальной цифровой вычислительной машины, и вряд ли у инженера-проектировщика возникнет мысль использовать ее на передающей стороне телеметрической системы. Алгоритмы обобщенной адаптивной дискретизации, которая является одним из наиболее эффективных методов уменьшения объема измерительных данных, должны реализовываться более простыми вычислительными средствами.

Однако оптимальное решение задачи необходимо как раз на стадии проектирования телеметрической системы, когда требуется выбрать наиболее эффективный, с точки зрения сжатия информации и экономических затрат, алгоритм дискретизации. Вот здесь-то инженеру и нужно знать, как далек от оптимального выбранный им алгоритм. И если такая оценка маскимально возможного коэффициента сжатия имеется, то поставленная задача может быть успешно решена после некоторых несложных инженерных экономических расчетов.

ЛИТЕРАТУРА

1. В. А. Виттих, А. Н. Гинзбург, Ю. П. Дробышев. Методы уменьшения объема измерительных сигналов.— Тезисы докладов и сообщений VI Всесоюзной конференции по автоматическому контролю и методам электрических измерений, Новосибирск, 1964.
2. К. Шеннон. Математическая теория связи.— В сб.: «Работы по теории информации и кибернетике». М.—Л., Изд-во иностр. лит., 1963.
3. И. М. Тепляков. Передача информации ортогональными сигналами.— Радиотехника, 1963, № 4.
4. А. Г. Витушкин. Оценка сложности задачи табулирования М., Физматгиз, 1959.
5. Б. П. Демидович, И. А. Марон, Э. З. Шувалова. Численные методы анализа. М., Физматгиз, 1963.
6. Л. С. Понтрягин, В. Г. Болтянский, Р. В. Гамкрелидзе, Е. Ф. Мищенко. Математическая теория оптимальных процессов. М., Физматгиз, 1961.
7. Б. П. Демидович, И. А. Марон. Основы вычислительной математики. М., Физматгиз, 1963.
8. H. Stone. Approximation of curves by line segments. Mathematics of computation, 1961, v. 15, № 73, p. 40—47.
9. Р. Беллман. Динамическое программирование. М.—Л., Изд-во иностр. лит., 1960.
10. R. Bellman. On the approximation of curves by line segments using dynamic programming. Communications of the association for computing machinery, 1961, v. 4, № 6, p. 284.
11. Brian Gluss. Further remarks on line segment curve-fitting using dynamic programming.— Communications of the association for computing machinery, 1962, v. 5, № 8, p. 441—443.
12. Brian Gluss. A line segment curve-fitting algorithm related to optimal encoding of information.— Information and control, 1962, v. 5, № 3, p. 261—267.

Поступила в редакцию
15 сентября 1964 г.