УДК 621.391:519.2

А. С. ГАВРИШ

СПОСОБЫ ЧИСЛЕННОГО РЕШЕНИЯ СИСТЕМ УРАВНЕНИЙ МАКСИМИЗАЦИИ ПОЛИНОМА ПРИ НЕОДИНАКОВО РАСПРЕДЕЛЕННЫХ ВЫБОРОЧНЫХ ЗНАЧЕНИЯХ

В настоящее время в статистической радиотехнике разработаны и широко используются методы синтеза измерителей различного назначения в предположении, что принимается полезный сигнал в аддитивной смеси с гауссовскими помехами. Однако гауссовские модели являются математической идеализацией реальных помех и не всегда адекватны на практике. В связи с этим интенсивно разрабатываются новые способы описания негауссовских помех, которые близки к реальным, и новые методы измерения параметров сигналов, принимаемых на фоне негауссовских помех.

До последнего времени при описании негауссовских случайных величин последовательностью моментов или кумулянтов и при наличии выборки независимых неодинаково распределенных значений единственным методом нахождения оценок параметров оставался метод наименьших квадратов. В работе [1] для нахождения оценок параметров случайной величины предложен метод максимизации полинома, основанный на использовании стохастических полиномов. Показано, что оценки, найденные методом максимизации полинома, могут быть значительно эффективнее соответствующих оценок, найденных методом наименьших квадратов. Улучшение точностных характеристик измерительных устройств достигается только при нелинейной обработке выборочных данных вследствие учета и оптимального использования тонкой структуры негауссовских помех в виде конечной последовательности кумулянтов высших порядков. Однако при степени полинома $s \geq 2$ аналитическое решение системы уравнений максимизации полинома удается получить достаточно редко [1]. В связи с этим для нахождения искомой оценки векторного параметра необходимо использовать численные методы решения системы уравнений максимизации полинома. Совершено очевидно, что для решения данной задачи могут использоваться различные приближенные методы.

Последние годы характеризуются стремительным развитием вычислительной техники и программного обеспечения. На рынке программных продуктов появилось огромное количество готовых прикладных пакетов программ, реализующих разработанные алгоритмы. В данном случае для решения нелинейных систем уравнений максимизации полинома при неодинаково распределенных выборочных значениях можно воспользоваться пакетом Mathcad.

Для решения систем уравнений в Mathcad использован итерационный метод Левенберга-Маркардта, являющийся квазиньютоновским методом (разновидностью градиентного метода). Предлагаемый метод уже программно реализован, и кроме того, подробно описан в научной литературе [2, 3], поэтому изложение его сути и особенностей в данной работе представляется излишним.

Отметим, что для большей эффективности метода Левенберга-Маркардта для решения систем уравнений максимизации полинома могут применяться его различные модификации. В общем случае Mathcad позволяет находить решение системы нелинейных трансцендентных уравнений, максимальное число которых равно пятидесяти.

Очевидно, что Mathcad предназначен для решения широкого круга задач в различных областях, поэтому в нем реализованы методы, имеющие широкую область применения. Последнее замечание указывает, в частности на то, что использование пакета Mathcad не всегда приводит к желаемым результатам и должно обязательно осуществляться с оглядкой на требования, предъявляемые к точности оценок. В результате в ряде случаев, когда Mathcad приводит к неточным или вообще ошибочным результатам, для реализации разработанных алгоритмов нахождения оценок предлагается использовать оригинальные авторские программы.

Отметим, что по своей сути метод максимизации полинома аналогичен методу максимального правдоподобия [1], поэтому для решения систем уравнений максимизации полинома целесообразно пользоваться теми же численными методами, которые используются для нахождения оценок параметров методом максимального правдоподобия [4, 5]. Подобный подход к выбору численных методов позволит обеспечить простоту вычислений и требуемую точность приближенного решения. В работе [1] рассматривается возможность использования численных методов для нахождения оценки как скалярного, так и векторного параметров методом максимизации полинома при одинаково распределенных выборочных значениях. Далее будут предложены различные варианты итеративных и рекуррентных процедур для совместного решения уравнений максимизации полинома при неодинаково распределенных выборочных значениях.

Рассмотрим численные методы решения системы уравнений максимизации полинома, в которых используется вся совокупность выборочных значений $\vec{x} = \{x_1, x_2, ..., x_n\}$. Согласно методу максимизации полинома, совместная оценка исследуемых параметров в общем случае находится из решения системы уравнений максимизации полинома:

$$\sum_{v=1}^{n}\sum_{i=1}^{s} k_{iv}^{(q)}(\vec{\vartheta})[x_v^i - m_{iv}(\vec{\vartheta})]\Big|_{\vec{\vartheta}=\hat{\vec{\vartheta}}} = 0,\ q = \overline{1,p}, \tag{1}$$

где $m_{iv}(\vec{\vartheta})$ – начальные моменты i-го порядка рассматриваемой случайной величины; в каждом q-ом уравнении (1) коэффициенты $k_{iv}^{(q)}(\vec{\vartheta})$, $i = \overline{1,s}$ находятся из решения системы линейных алгебраических уравнений

$$\sum_{j=1}^{s} k_{jv}^{(q)}(\vec{\vartheta})F_{(i,j)v}(\vec{\vartheta}) = \frac{\partial}{\partial \vartheta_q} m_{iv}(\vec{\vartheta}),\ i = \overline{1,s},\ v = \overline{1,n},\ q = \overline{1,p}.$$

В последнем выражении функции $F_{(i,j)v}(\vec{\vartheta})$ – центрированные коррелянты размером (i,j), которые связаны с начальными моментами соотношением

$$F_{(i,j)v}(\vec{\vartheta}) = m_{(i+j)v}(\vec{\vartheta}) - m_{iv}(\vec{\vartheta})m_{jv}(\vec{\vartheta}).$$

Предположим, что вектор $\breve{\vec{\vartheta}}_k$ является решением системы уравнений (1) при k-ой итерации и обозначим левую часть каждого уравнения системы (1) через $f_q(\vec{x}/\breve{\vec{\vartheta}}_k)$. Тогда для получения итеративной процедуры Ньютона-Рафсона необходимо разложить каждую функцию $f_q(\vec{x}/\breve{\vec{\vartheta}}_k)$ на $(k+1)$-м шаге в ряд Тейлора в окрестности значения вектора $\breve{\vec{\vartheta}}_k$, полученного на предыдущем k-м шаге итерации и ограничиваясь первыми дифференциалами:

$$f_q(\vec{x}/\breve{\vec{\vartheta}}_k) + \sum_{j=1}^{p}(\breve{\vartheta}_{k+1}^q - \breve{\vartheta}_k^q)\frac{\partial}{\partial \breve{\vartheta}_k^q} f_q(\vec{x}/\breve{\vec{\vartheta}}_k) = 0,\ q = \overline{1,p}, \tag{2}$$

где верхний индекс в $\breve{\vartheta}_k^q$ указывает на то, какая из компонент векторного параметра $\breve{\vec{\vartheta}}_k$ рассматривается, а нижний индекс означает шаг итерации.

Используя выражение (2), легко показать, что $(k+1)$-я итерация в векторной форме находится в соответствии с выражением

$$\breve{\vec{\vartheta}}_{k+1} = \breve{\vec{\vartheta}}_k - Y(\vec{x}/\breve{\vec{\vartheta}}_k)[Z(\vec{x}/\breve{\vec{\vartheta}}_k) - J_{sn}(\breve{\vec{\vartheta}}_k)]^{-1}, \tag{3}$$

где $Y(\vec{x}/\breve{\vec{\vartheta}}_k)$ – матрица-столбец размером p, составленная из элементов $f_q(\vec{x}/\breve{\vec{\vartheta}}_k)$; $J_{sn}(\breve{\vec{\vartheta}}_k)$ – матрица количества извлекаемой информации о векторном параметре $\breve{\vec{\vartheta}}_k$, элементы которой имеют вид:

$$J_{sn}^{(q,m)}(\vec{\vartheta}) = \sum_{v=1}^{n}\sum_{i=1}^{s}\sum_{j=1}^{s} k_{iv}^{(q)}(\vartheta_0)k_{jv}^{(m)}(\vartheta_0)F_{(i,j)v}(\vartheta_0) = \sum_{i=1}^{s}\sum_{v=1}^{n} k_{jv}^{(q)}(\vec{\vartheta})\frac{\partial}{\partial \breve{\vartheta}_k^m} m_{iv}(\vec{\vartheta}),\ m,q = \overline{1,p};$$

$Z(\vec{x}/\breve{\vec{\vartheta}}_k)$ – квадратная матрица размером p с элементами:

$$z_{q,j}(\vec{x}/\breve{\vec{\vartheta}}_k)=\sum_{i=1}^{s}\sum_{v=1}^{n}\frac{\partial}{\partial\breve{\vartheta}_k^j}k_{iv}^{(q)}(\breve{\vec{\vartheta}}_k)[x_v^i-m_{iv}(\breve{\vec{\vartheta}}_k)]. \qquad (4)$$

В качестве начального приближения $\breve{\vec{\vartheta}}_0$ целесообразно выбирать вектор, получая методом максимизации полинома при степени стохастического полинома $s=1$, или же произвольно с учетом имеющихся представлений о величине значений оцениваемых параметров.

На практике чаще удобнее пользоваться итеративной процедурой, аналогичной процедуре накопления, впервые предложенной Фишером [4]. Легко показать, что вследствие усиленного закона больших чисел [6] при $n\to\infty$ каждый элемент $z_{q,j}(\vec{x}/\breve{\vec{\vartheta}}_k)$ вида (4) с вероятностью 1 стремится к нулю. Даже если объем выборки не достаточно велик, матрицей $Z(\vec{x}/\breve{\vec{\vartheta}}_k)$ можно пренебречь, поскольку она будет значительно меньше матрицы количества извлекаемой информации $J_{sn}(\breve{\vec{\vartheta}}_k)$. Тогда итеративную процедуру вида (3) можно представить в виде

$$\breve{\vec{\vartheta}}_{k+1}=\breve{\vec{\vartheta}}_k+Y(\vec{x}/\breve{\vec{\vartheta}}_k)J_{sn}^{-1}(\breve{\vec{\vartheta}}_k).$$

Данная процедура в вычислительном отношении значительно проще процедуры Ньютона-Рафсона, поскольку нет необходимости находить матрицу $Z(\vec{x}/\breve{\vec{\vartheta}}_k)$, зависящую от выборочных данных.

Основным недостатком итерационных методов является трудоемкость вычислительной процедуры, связанная с необходимостью обработки данных по всему объему выборки. Для устранения этого недостатка можно пользоваться рекуррентными методами решения системы уравнений максимизации полинома, в которых оценка векторного параметра вычисляется после каждого выборочного значения. Достоинством рекуррентных процедур является то, что воспользоваться результатами вычислений можно на любом из промежуточных шагов. Причем с каждым шагом, т. е. по мере получения выборочных данных, точность вычислений будет увеличиваться.

Пусть вектор $\breve{\vec{\vartheta}}^{(n)}$ является оценкой, получаемой из совместного решения уравнений (1) при объеме выборки n. Тогда при наблюдении совокупности $(n+1)$ выборочных значений оценка будет находиться из решения системы уравнений:

$$\sum_{i=1}^{s}\sum_{v=1}^{n+1}k_{iv}^{(q)}(\vec{\vartheta})[x_v^i-m_{iv}(\vec{\vartheta})]\Big|_{\vec{\vartheta}=\breve{\vec{\vartheta}}^{(n+1)}}=0,\ q=\overline{1,p}.$$

Для получения рекуррентной процедуры выражение для нахождения оценки удобно переписать в виде

$$f_n^{(q)}(x_1,\ldots,x_n/\vec{\vartheta})+f_{n+1}^{(q)}(x_{n+1}/\vec{\vartheta})=0,\ \ q=\overline{1,p}, \qquad (5)$$

где функция $f_n^{(q)}(x_1,\ldots,x_n/\vec{\vartheta})$ – уравнение максимизации полинома при объеме выборки n; $f_{n+1}^{(q)}(x_{n+1}/\vec{\vartheta})$ – уравнение максимизации полинома зависящее только от вновь полученного выборочного значения x_{n+1}.

Поскольку значение вектора $\breve{\vec{\vartheta}}^{(n+1)}$ незначительно отличается от $\breve{\vec{\vartheta}}^{(n)}$, то левую часть каждого уравнения системы (5) можно разложить в ряд Тейлора в окрестности значения вектора $\breve{\vec{\vartheta}}^{(n)}$, ограничившись первыми дифференциалами. Используя это разложение, можем получить рекуррентное соотношение для определения очередного значения оценки:

$$\breve{\vec{\vartheta}}^{(n+1)}=\breve{\vec{\vartheta}}^{(n)}-Y_{(n+1)}(\vec{x}_n/\breve{\vec{\vartheta}}^{(n)})[Z_n(\vec{x}_n/\breve{\vec{\vartheta}}^{(n)})+D_{n+1}(x_{n+1}/\breve{\vec{\vartheta}}^{(n)})-J_{s(n+1)}(\breve{\vec{\vartheta}}^{(n)})]^{-1}, \qquad (6)$$

где $Y_{(n+1)}(x_{n+1}/\breve{\vec{\vartheta}}^{(n)})$ – матрица-столбец с элементами $f_{n+1}^{(q)}(x_{n+1}/\breve{\vec{\vartheta}}^{(n)})$; квадратные матрицы $Z_n(\vec{x}_n/\breve{\vec{\vartheta}}^{(n)})$ и $D_{n+1}(x_{n+1}/\breve{\vec{\vartheta}}^{(n)})$ соответственно имеют элементы:

$$z_{q,j}(\vec{x}_n/\breve{\vec{\vartheta}}^{(n)}) = \sum_{i=1}^{s}\sum_{v=1}^{n}\frac{\partial}{\partial\breve{\vartheta}_j^{(n)}}k_{iv}^{(q)}(\breve{\vec{\vartheta}}^{(n)})[x_v^i - m_{iv}(\breve{\vec{\vartheta}}^{(n)})],$$

$$d_{q,j}(x_{n+1}/\breve{\vec{\vartheta}}^{(n)}) = \sum_{i=1}^{s}\frac{\partial}{\partial\breve{\vartheta}_j^{(n)}}k_{i(n+1)}^{(q)}(\breve{\vec{\vartheta}}^{(n)})[x_{(n+1)}^i - m_{i(n+1)}(\breve{\vec{\vartheta}}^{(n)})].$$

Выражение (6) очень напоминает формулу (3) описывающую итеративную процедуру Ньютона-Рафсона. Однако, несмотря на относительное сходство формы, они существенно отличаются по содержанию. В сравниваемых выражениях матрицы $Y(\vec{x}/\breve{\vec{\vartheta}}_k)$ и $Y_{(n+1)}(x_{n+1}/\breve{\vec{\vartheta}}^{(n)})$ определяют величину поправки к оценке, полученной на предыдущем шаге. При этом в соотношении (3) поправка всегда зависит от всей совокупности наблюдаемых данных $\{x_1,...,x_n\}$, в то время как в выражении (6) соответствующая матрица зависит только от последнего значения x_{n+1}.

Существует целый ряд способов [7] упрощения рекуррентной процедуры, описываемой выражением (6). В вычислительном плане, для приближенного вычисления искомых оценок, удобно пользоваться процедурой, аналогичной процедуре Сакрисона [8]. При большом объеме наблюдаемых данных справедливо соотношение

$$Z_n(\vec{x}_n/\breve{\vec{\vartheta}}^{(n)}) + D_{n+1}(x_{n+1}/\breve{\vec{\vartheta}}^{(n)}) << J_{s(n+1)}(\breve{\vec{\vartheta}}^{(n)}),$$

следовательно, от выражения (6) можно перейти к рекуррентной процедуре вида

$$\breve{\vec{\vartheta}}^{(n+1)} = \breve{\vec{\vartheta}}^{(n)} + Y_{(n+1)}(\vec{x}_n/\breve{\vec{\vartheta}}^{(n)})J_{s(n+1)}^{-1}(\breve{\vec{\vartheta}}^{(n)}).$$

В последнем соотношении весовая матрица $J_{s(n+1)}(\breve{\vec{\vartheta}}^{(n)})$ не зависит от выборочных значений, что упрощает процесс нахождения оценки векторного параметра $\breve{\vec{\vartheta}}^{(n+1)}$.

Отметим, что все рассмотренные итеративные и рекуррентные процедуры нахождения оценок максимизации полинома являются приближенными.

Список литературы: **[1]** *Кунченко Ю.П., Лега Ю.Г.* Оценка параметров случайных величин методом максимизации полинома. К.: Наукова думка, 1992. – 180 с. **[2]** *Users Guide* to Minpack I, by Jorge J. More, Burton S. Garbow, and Kenneth E. Hillstrom, Argonne National Laboratory publication ANL–80-74. – 1980. **[3]** *Mathcad 6.0 Plus.* Финансовые, инженерные и научные расчеты в среде Windows 95: Перевод с англ. – М.: Филинъ, 1996. – 712 с. **[4]** *Закс Ш.* Теория статистических выводов.– М.: Мир, 1975. **[5]** *Репин В.Г., Тартаковский Г.П.* Статистический синтез при априорной неопределенности и адаптации информационных систем. – М.: Сов. радио, 1977. – 432 с. **[6]** *Гихман И.И., Скороходов А.В., Ядренко М.И.* Теория вероятностей и математическая статистика. – К.: Вища школа, 1979. – 408с. **[7]** *Невельсон М.Б., Хасьминский Р.З.* Стохастическая аппроксимация и рекуррентное оценивание. – М.: Наука, 1972. **[8]** *Sakrison D.T.* Efficient recursive estimation, application to estimating the parameters of covariance functions. // Internat. Jorn. Engineering Science. –1965. – Vol. 3. – N 4. – P. 461-483.

Черкасский инженерно-технологический институт

Поступила в редколлегию 9.03.2000