УДК 550.834

© 2011

В. С. Мостовой, С. В. Мостовой

Оптимальные оценки нелинейных параметров в моделях сейсмоакустического мониторинга

(Представлено академиком НАН Украины В. И. Старостенко)

Запропоновано підхід розподілу лінійних і нелінійних параметрів моделі при оптимальному їх оцінюванні за допомогою критеріїв ризику, який дозволяє зменшити розмірність простору оцінюваних параметрів в техніці Монте-Карло при пошуку глобального мінімуму критеріїв ризику до розмірності простору лише нелінійних параметрів. Це особливо відчутно при оцінці моделей "старіння" при дослідженні геофізичних об'єктів, моделі яких саме і відрізняються великою розмірністю. Запропонований підхід до рішення проілюстрований обробкою і аналізом даних, отриманих в польових спостереженнях в режимі моніторингу.

Особенность динамического анализа различных систем в режиме мониторинга заключается в следующем: режим обработки наблюденных данных и принятия решения о состоянии объекта мониторинга должны быть выполнены в режиме, близком к реальному времени, в то время как системы часто описываются достаточно сложными моделями с большой размерностью вектора свободных параметров последних. Причем лишь часть этих параметров входит в модель линейно, а большая их часть входит нелинейно. В таких случаях существенную роль играет априорная изученность процессов, что позволяет упростить процесс оценки свободных параметров, при этом естественно строить такую процедуру оценивания параметров, которая базируется на бейесовских оценках и, тем самым, учитывает имеющуюся у исследователей априорную информацию. В результате мониторинга анализ изменения во времени множества свободных параметров позволяет говорить о состоянии объекта исследования и прогнозировать его поведение в будущем. Описанный подход представляется плодотворным в оценке состояния исторически ценных архитектурных памятников и их возможной реакцией на землетрясения и создания относительно оптимального уровня сопротивления сейсмическим событиям [1].

Математическая модель оценки параметров. Динамические системы моделируются суперпозицией периодических или квазипериодических процессов, которые и представляют предмет анализа, поскольку дают возможность прогнозировать по предыстории поведение системы в будущем.

Существенную роль в создании алгоритмов анализа играет тот факт, что в модели динамического процесса $M(\mathbf{\Lambda}, t)$ есть хотя бы часть свободных параметров, входящих в модель линейно. Это позволяет ускорить процесс обработки и анализа данных, получаемых в результате мониторинга. Модель процесса строим в виде многообразия [2], отображающего модель в точку N-мерного пространства R_N через свободные параметры модели, где N — количество этих параметров. $M(\mathbf{\Lambda}, t)$ кроме времени зависит еще и от множества свободных параметров $\mathbf{\Lambda}$. Результат мониторинга процесса $y(t, \mathbf{\Lambda}, \alpha)$ может быть представлен так:

$$y(t, \mathbf{\Lambda}, \alpha) = M(\mathbf{\Lambda}, t) + \alpha(t), \qquad t \in (0, T), \tag{1}$$

где множество параметров, отраженных в данной модели, упорядочены в прямоугольную матрицу свободных параметров, подлежащих оценке; $\alpha(t)$ — источник случайности, т.е. флуктуационные возмущения, которые не позволяют точно воспроизвести множество $\mathbf{\Lambda}$; T — область наблюдения во времени. Источником случайности кроме аддитивной помехи $\alpha(t)$ являются и флуктуации параметров модели.

Выбираем такую модель, в которой множество ее свободных параметров содержит подмножество, входящее в модель линейно. Упорядочив его в вектор $\mathbf{h} = \{h_k\}$, $k = \overline{1, n}$, и выделив из множества параметров $\mathbf{\Lambda}$ это подмножество, оставив в $\mathbf{\Lambda}$ лишь нелинейно входящие параметры, приходим к модели вида

$$M(\mathbf{\Lambda}, t) = \sum_{k=1}^{n_l} h_k \cdot \exp\{-\lambda_{k,1} t\} \sin(\lambda_{k,2} t - \lambda_{k,3}), \tag{2}$$

где $\mathbf{\Lambda}^{\langle k \rangle}$ — векторы-столбцы в матрице нелинейных параметров всего множества $\mathbf{\Lambda}$. Поскольку $\mathbf{\Lambda}^{\langle k \rangle}$ в $\mathbf{\Lambda}$, относящиеся к k-й подмодели $M_k(\mathbf{\Lambda}^{\langle k \rangle}, t)$, могут иметь различную размерность n_k, то выбрав $\max\limits_k(n_k) = n$ и заполнив в каждом векторе $\mathbf{\Lambda}^{\langle k \rangle}$ позиции, начиная с номера n_{k+1} до номера n нулями, мы получим прямоугольную матрицу нелинейно входящих в модель параметров $\mathbf{\Lambda}$ размерностью $n \times n_l$.

Целью анализа является построение оптимальных оценок $\widehat{\mathbf{\Lambda}}$ с помощью множества $\widetilde{\widehat{\mathbf{\Lambda}}} = \left\{\widehat{\mathbf{\Lambda}}_r(y(t, \alpha(t)))\right\}$ решающих правил. При выборе решающего правила в качестве критерия оптимальности на множестве из R решающих правил и наблюденных данных $y(t, \alpha(t))$ выбираем критерий среднего риска $R(\widehat{\mathbf{\Lambda}}_r(y(t, \alpha)))$ [3]:

$$R(\widehat{\mathbf{\Lambda}}_r(y(t, \alpha))) = \sum_r L(\widehat{\mathbf{\Lambda}}_r(y(t, \alpha), \mathbf{\Lambda})) \cdot P\left(\frac{\mathbf{\Lambda}}{y(t, \alpha)}\right). \tag{3}$$

Здесь $L(\widehat{\mathbf{\Lambda}}_r(y(t, \alpha), \mathbf{\Lambda}))$ — функция потерь, которая зависит от оценки свободных параметров, которые принимаются как решение по решающему правилу $\widehat{\mathbf{\Lambda}}_r(y(t, \alpha))$ с номером r, и истинного значения этих параметров $\mathbf{\Lambda}$. Вид функции $L(\widehat{\mathbf{\Lambda}}_r(y(t, \alpha), \mathbf{\Lambda}))$ с помощью теории не определяется, он выбирается из инженерных или интуитивных соображений, чем больше "соответствует истине" принятое решение, тем меньше должно быть значение функции потерь или хотя бы не больше. $\widetilde{\widehat{\mathbf{\Lambda}}}(y(t, \alpha))$ — множество из R решающих правил. В последнем выражении $P(\mathbf{\Lambda}/y(t, \alpha))$ — это условная вероятность значения нелинейных свободных параметров при условии реализации $y(t, \alpha)$.

Учитывая формулу Бейеса [4], средний риск с точностью до константы принимает вид:

$$R\left(\widehat{\mathbf{\Lambda}}_r(y(t, \alpha)) = C \cdot \sum_{\mathbf{\Lambda} \in \widetilde{\mathbf{\Lambda}}} L(\widehat{\mathbf{\Lambda}}_r(y(t, \alpha), \mathbf{\Lambda})) \cdot P\left(\frac{y(t, \alpha)}{\mathbf{\Lambda}}\right) \cdot P(\mathbf{\Lambda})\right), \tag{4}$$

где $C = (P(y(t, \alpha)))^{-1}$, величина, обратная полной вероятности $P(y(t, \alpha))$, зависит только от реализации $y(t, \alpha)$. Оптимальную оценку $\widehat{\mathbf{\Lambda}}^*$ параметров модели получаем минимизацией среднего риска на множестве решающих правил $\widehat{\mathbf{\Lambda}}_r(y(t, \alpha))$:

$$\widehat{\mathbf{\Lambda}}^* = \min_{\widehat{\mathbf{\Lambda}}_r(y(t, \alpha)) \in \widetilde{\widehat{\mathbf{\Lambda}}}} R(\widehat{\mathbf{\Lambda}}_r(y(t, \alpha))); \qquad r = \overline{1, R}. \tag{5}$$

Функция потерь в расчетах имела вид

$$L(\widehat{\mathbf{\Lambda}}_r(y(t,\alpha),\mathbf{\Lambda})) = \begin{cases} 0, & \text{если} \quad \|\widehat{\mathbf{\Lambda}}_r - \mathbf{\Lambda}\| \leqslant \Omega, \\ \dfrac{1}{\Omega}, & \text{если} \quad \|\widehat{\mathbf{\Lambda}}_r - \mathbf{\Lambda}\| \succ \Omega. \end{cases} \tag{6}$$

Если норма разности вектора оценок параметров модели не превосходит фиксированной величины Ω, то наблюдатель потерь не несет, в противном случае несет потерю $1/\Omega$.

Модель объекта исследования. Такой подход был применен авторами для обработки и анализа данных, полученных в полевых наблюдениях мониторинга объекта, спектральные характеристики которого лежат (0,1–8 Гц) в сейсмической области частот. Особенностью эксперимента было то обстоятельство, что за время мониторинга характеристики объекта менялись по определенной программе, и целью эксперимента было оценить эти изменения в пространстве параметров модели объекта. Размерность вектора параметров изменялась в зависимости от выбираемой модели в пределах от 12 до 20. Поскольку объект описывался множеством линейно и нелинейно входящих в модель параметров, а поиск глобального минимума критерия оптимальности оценок для каждого из решающих правил осуществлялся в зависимости от гипотезы об аддитивной помехе $\alpha(t_i)$. Эта помеха моделировалась, в общем случае, нестационарным гауссовым шумом. Решающее правило было следующим: принималась оценка, минимизирующая уклонение модели от наблюденных данных в норме Гильбертова пространства. Скалярное произведение для двух функций $\langle y(t), f(t)\rangle_r$ в их дискретном представлении выбиралось как для двух векторов таким образом:

$$\langle y(t), f(t)\rangle_r = k(\boldsymbol{\sigma}_r)\cdot\sum_{i,j} y(t_i) f(t_j) W_r(t_i,t_j); \qquad r=\overline{1,R}; \qquad i,j=\overline{1,M}. \tag{7}$$

В функции (7) — $W_r(t_i,t_j)$ матрица, обратная к матрице ковариаций аддитивного фона $\alpha(t)$ с параметрами $\boldsymbol{\sigma}_r$; $k(\boldsymbol{\sigma}_r)$ — коэффициент, зависящий от определителя матрицы ковариаций $\boldsymbol{\sigma}_r$ (где M — это размерность вектора дискретного представления функций на временном интервале T с интервалом квантования Δ так, что $M\Delta = T$); R — как уже отмечалось ранее, количество решающих привал.

Приведенный выше алгоритм в расчетах описывался двумя типами решающих правил. Одно из них было основано на предположении о высокочастотной помехе с фиксированным множеством значений мощности такого процесса. Такой процесс можно моделировать белым шумом [3], в этом случае ядро (7) будет:

$$k(\boldsymbol{\sigma}_r) W_r(t_i,t_j) = \frac{1}{\boldsymbol{\sigma}_r}\cdot\delta(t_i - t_j). \tag{8}$$

Норма уклонения одной функции от другой (норма разности) —

$$\sqrt{\langle y(t)-f(t), y(t)-f(t)\rangle_r} = \frac{1}{\boldsymbol{\sigma}_r}\sqrt{\sum_i (y(t_i)-f(t_i))^2}. \tag{9}$$

Второе решающее правило основано на гипотезе о низкочастотном стационарном нормальном шуме с фиксированным множеством параметров, определяющих этот шум [6]. Расчеты по данным мониторинга объекта показали, что гипотеза о высокочастотной помехе дает лучшее согласие модели объекта с наблюденными данными.

Модель объекта мониторинга. Гипотеза заключалась в том, что объект описываем линейной комбинацией из n_l осцилляторов, каждый из которых являет собой подмодель (случай (2)):

$$M_k(\mathbf{\Lambda}^{\langle k\rangle}, t) = \sum_{s=0}^{n_l-1} h_k \cdot \exp\{-\lambda_{k,s+1} \cdot t\} \sin(\lambda_{k,s+2} \cdot t - \lambda_{k,s+3}). \tag{10}$$

Модель (2) принимает вид

$$M(\mathbf{\Lambda}, t) = \sum_{k=1}^{n_l} h_k \cdot \sum_{s=0}^{n_l-1} \exp\{-\lambda_{k,1} \cdot t\} \sin(\lambda_{k,2} \cdot t - \lambda_{k,3}). \tag{11}$$

Относительно вектора $\mathbf{h}$ и наблюденных данных $y(t)$, при фиксированных значениях матрицы $\mathbf{\Lambda}$, используя метрику решающего правила с номером n, мы получаем систему линейных уравнений:

$$\Psi_r \mathbf{h}_r = \mathbf{I}_r; \qquad r = \overline{1, R}. \tag{12}$$

Здесь матрица с индексом r имеет элементы

$$\Psi_{k,s}^r = \frac{1}{\boldsymbol{\sigma}_r} \sum_{i=0}^{M} \sum_{k=1}^{n_l} M_k(\mathbf{\Lambda}^{\langle k\rangle}, t_i) \cdot M_s(\mathbf{\Lambda}^{\langle s\rangle}, t_i); \qquad s = \overline{0, n_l - 1}. \tag{13}$$

Вектор $\mathbf{I}_r$ правой части уравнения (12) вычислялся по формуле

$$\mathbf{I}_r = \{I_s^r\}; \qquad I_s^r = \frac{1}{\boldsymbol{\sigma}_r} \sum_{i=0}^{M} y(t_i) \cdot M_s(\mathbf{\Lambda}^{\langle s\rangle}, t_i). \tag{14}$$

Нужно отметить, что бейесовский подход к принятию решения в геофизике применяется широко. Достаточно отметить недавнюю работу [4].

Обработка наблюденных данных.Она осуществлялась по такому алгоритму: методом Монте-Карло, по априорным распределениям, выбрасывалась псевдослучайная точка в пространстве нелинейных параметров. Для этой точки для модели объекта (формула (10)) и наблюденных данных решением системы линейных уравнений (12) для линейных параметров вычислялась точка в подпространстве линейных параметров. Совокупность нелинейных и линейных свободных параметров модели, полученных таким образом, использовалась для оценки локального минимума в нелинейной задаче как начальная точка для алгоритма в методе Ливенберга–Маквардта [7]. Такая процедура выполнялась многократно, в результате чего вычислялась оценка минимума миниморума на множестве псевдослучайных точек. Такая оценка по вероятности сходится к глобальному минимуму. В 12-мерном пространстве свободных параметров модели объект отображался в точку с такими компонентами:

$$\lambda^T = \{7{,}313 \cdot 10^{-3}; 2{,}962; 0{,}698; 0{,}94; 0{,}114; 9{,}028; 0{,}452; -0{,}203; -0{,}231; 11{,}284;$$
$$6{,}069 \cdot 10^{-3}; 3{,}369\}.$$

Здесь нулевая, четвертая и восьмая компоненты — это показатель в экспоненте затухания соответственно первой второй и третьей гармоник. Первая, пятая и девятая компоненты — это круговая частота соответственно первой второй и третьей гармоник в радианах. Вторая, шестая и десятая компоненты — это амплитуды гармоник в относительных единицах. Третья, седьмая и одиннадцатая компоненты — это фазовый сдвиг в радианах.

Таким образом, предложен подход разделения линейных и нелинейных параметров модели, при оптимальном их оценивании с помощью критериев риска, что позволяет уменьшить размерность пространства оцениваемых параметров в технике Монте-Карло при поиске глобального минимума критериев риска до размерности пространства лишь нелинейных параметров. Это особенно ощутимо при оценке моделей с большой размерностью свободных параметров, что типично для задач геофизики. Обработка данных мониторинга объекта с меняющимися характеристиками по предложенному алгоритму позволила в пространстве свободных параметров модели оценить даже незначительные изменения в состоянии объекта, т. е. положении точки в пространстве его характеристик.

1. *Математическая* энциклопедия. Т. 3. – Москва: Сов. энциклопедия, 1977. – 742 с.
2. *Большаков И. А.* Статистические проблемы выделения потока сигналов из шума. – Москва: Сов. радио, 1969. – 453 с.
3. *Tassios T. P.* Seismic engineering of monuments // Bull. earthquake eng. – 2010. – **8**, No 6. – P. 1231–1265.
4. *Imoto M.* Hypocenter determination with prior information for clustering earthquakes // Geophys. J. Int. – 2010. – **182**, is. 3. – P. 1374–1382.
5. *Математическая* энциклопедия. Т. 1. – Москва: Сов. энциклопедия, 1977. – 402 с.
6. *Амиантов И. Н.* Избранные вопросы статистической теории связи. – Москва: Сов. радио, 1971. – 423 с.
7. *Marquardt D.* An Algorithm for Least-Squares Estimation of Nonlinear Parameters // SIAM J. Appl. Math. – 1963. – **11**. – P. 431–441.

Институт геофизики им. С. И. Субботина
НАН Украины, Киев

Поступило в редакцию 18.02.2011

V. S. Mostovyi, S. V. Mostovyi

Optimal estimations of nonlinear parameters in models of seismoacoustic monitoring

In the problem of optimal estimation of model parameters using risk criteria, we propose an approach to the separation of linear parameters from nonlinear ones. In the problem of finding a global minimum of risk criteria, our approach leads to a decrease of the dimension of the space of free variables up to the dimension of the space of nonlinear parameters. This allows one to obtain a simpler minimization problem, which can be solved more efficiently via Monte-Carlo methods. Such an improvement is very significant in the estimation of models of the object "aging" at the investigation of geophysical objects, models of which typically have high dimensionality. We illustrate the proposed method with the processing and analysis of data obtained during the field observations in the regime of monitoring.