УДК 550.34.01:519.2

ОЦЕНИВАНИЕ ВЕКТОРА КАЖУЩЕЙСЯ МЕДЛЕННОСТИ ПЛОСКОЙ ВОЛНЫ ПО ДАННЫМ ТРЕХКОМПОНЕНТНОЙ СЕЙСМИЧЕСКОЙ ГРУППЫ: СТАТИСТИЧЕСКАЯ ЗАДАЧА С МЕШАЮЩИМИ ПАРАМЕТРАМИ

А.Ф. Кушнир

Международный институт теории прогноза землетрясений и математической геофизики Российской академии наук

С помощью вычислительных алгоритмов оцениваются параметры распространения плоской сейсмической волны по наблюдениям трехкомпонентной малоапертурной сейсмической группы. Задача трактуется как статистическая с мешающими параметрами в связи с наличием существенно коррелированных (часто – когерентных) помех, воздействующих на датчики группы, и отсутствием полной априорной информации о временно́й форме сейсмической волны. Для волновой формы используются две модели: модель стационарного случайного процесса со спектральной плотностью, зависящей от конечномерного мешающего параметра, и модель полностью неизвестной наблюдателю последовательности, обладающей временно́й автокорреляционной функцией (также неизвестной). Для каждой из моделей строятся состоятельные и асимптотически нормальные алгоритмы оценивания, отвечающие асимптотическим условиям оптимальности. Свойства оценок исследуются при малом и большом отношениях сигнал/шум.

ESTIMATION OF APPARENT SLOWNESS VECTOR FOR A PLANE WAVE USING DATA FROM A THREE-COMPONENT SEISMIC ARRAY: A STATISTICAL PROBLEM INVOLVING NUISANCE PARAMETERS

A.F. Kushnir

International Institute of Earthquake Prediction Theory and Mathematical Geophysics, Russian Academy of Sciences

This paper is concerned with computational algorithms for estimating the propagation parameters of a plane seismic using data from a small-aperture three-component seismic array. This problem is treated as a statistical one involving nuisance parameters arising from a significantly correlated (and frequency coherent) noise that affects the seismometers and an absence of complete prior information on the waveform to be analyzed. The wave is represented by two models; the one assumes a stationary random process with a spectral density that is a function of a finite-dimensional nuisance parameter, while the other is assumed to be a time series that is completely unknown tothe observer and which has a time averaged autocorrelation function of time (which is unknown as well). For each of these models we construct consistent and asymptotically normal algorithms for estimatin the propagation parameters of a wave, the algorithms obeying the asymptotic requirements. The estimators are compared for large and small signal/noise ratios.

1. МАТЕМАТИЧЕСКАЯ МОДЕЛЬ НАБЛЮДЕНИЙ В ВИДЕ СЛУЧАЙНОГО ВРЕМЕННО́ГО РЯДА С НЕИЗВЕСТНЫМИ ИНФОРМАТИВНЫМИ И МЕШАЮЩИМИ ПАРАМЕТРАМИ

Опыт эксплуатации современных автоматизированных систем регионального сейсмического мониторинга показывает, что точность лоцирования очагов сейсмических событий может быть существенно повышена при использовании информации о направлении прихода сейсмических волн на датчики сейсмической сети. В принципе, если измерены азимут прихода и моменты вступления P- и S-волн (точнее, каких-либо двух сейсмических фаз P- и S-типа), то при известной региональной модели Земли возможно лоцирование очага по данным одного сейсмического датчика. Ошибка лоцирования при этом, естественно, зависит от точности оценки азимута и моментов вступления фаз и от правильности идентификации фаз. Одним из наиболее надежных источников информации для идентификации фаз является величина лучевого параметра сейсмической фазы – кажущейся медленности. Азимут α, кажущаяся медленность $|\mathbf{p}|$ и угол выхода волны β однозначно определяются вектором кажущихся медленностей $\mathbf{p} = (p_x, p_y)^{\mathrm{T}}$, который является наиболее удобной характеристикой направления прихода волны:

$$\mathrm{tg}\alpha = \frac{p_x}{p_y}, \quad |\mathbf{p}| = \sqrt{p_x + p_y}, \quad \sin\beta = v|\mathbf{p}|,$$

где v – скорость сейсмической волны у дневной поверхности.

Из сказанного следует, что оценка вектора $\mathbf{p}$ дает весьма существенную информацию для решения основной задачи сейсмического мониторинга – лоцирования очагов событий.

Оценка вектора $\mathbf{p}$ может быть получена методами поляризационного анализа записей трехкомпонентных сейсмометров. Однако существенно более высокая точность оценивания $\mathbf{p}$ обеспечивается при применении малоапертурных сейсмических групп. Важно, что с помощью группы оценку кажущейся медленности сейсмической фазы можно осуществлять, не зная величины v – фазовой скорости волны в приповерхностном слое земной коры. Указанные преимущества малоапертурных групп являются весьма существенными при мониторинге слабой сейсмичности в регионах с ответственными инженерными сооружениями (плотинами, нефтяными морскими платформами, атомными электростанциями), а также для контроля подземных ядерных испытаний. При решении задач мониторинга единичная малоапертурная группа может успешно конкурировать с локальной или даже региональной сетью трехкомпонентных станций.

В настоящей работе оценивание вектора кажущейся медленности плоской волны по данным трехкомпонентной малоапертурной сейсмической группы трактуется как задача статистического оценивания параметров многомерных стохастических временны́х рядов. Этот подход существенно отличается от традиционного, согласно которому оценивание направления прихода плоской волны рассматривается как задача пространственного спектрального анализа [1, 2].

Исключение составляет работа [3], где дан сравнительный анализ алгоритмов оценивания с помощью малоапертурной группы однокомпонентных датчиков и одиночного трехкомпонентного сейсмометра с точки зрения достижения статистических границ качества оценивания. Эта работа, однако, носит преимущественно экспериментальный характер и не содержит выводов теоретических формул для оценок направления по данным группы. Исследуется лишь одна оценка, и не обсуждаются условия, при которых ее применение оправдано. И хотя указывается, что эта оценка не оптимальна, т.е. ее качество не достигает статистической границы, ей не предлагается никакой альтернативы. Следует отметить также, что в работе [3] принято существенное ограничение относительно сейсмических помех, на фоне которых наблюдается анализируемая сейсмическая волна: помеха считается белым шумом, некоррелированным для различных датчиков группы.

На практике эти помехи часто являются когерентными: поверхностными помеховыми волнами, порождаемыми либо прибоем на морских побережьях, либо промышленными источниками, или объемными волнами других сейсмических фаз данного события, перекрывающихся по времени с анализируемой волной (например, волны Pn и Pg, имеющие разные кажущиеся скорости, перекрываются во времени на определенных расстояниях от источника). Наконец определенный интерес представляет ситуация искусственного сокрытия подземного ядерного взрыва, который произведен в момент прихода волны от достаточно сильного землетрясения. Здесь помехой для объемной волны от взрыва выступает объемная волна от землетрясения, имеющая, вообще говоря, другое направление распространения. Анализ оценивания направления плоской волны в условиях коррелированных (когерентных) помех, таким образом, имеет существенное практическое значение.

Математическая модель сигналов и помех, которая используется в дальнейшем, во временно́й области имеет вид

$$\mathbf{x}(t) = \mathbf{u}(t) + \boldsymbol{\xi}(t) = \mathbf{h}(t, \mathbf{p}, v) * s(t) + \boldsymbol{\xi}(t), \quad [0, T], \tag{1.1}$$

где $\mathbf{x}(t) = (\mathbf{x}_1^{\mathrm{T}}(t), \ldots, \mathbf{x}_M^{\mathrm{T}}(t))^{\mathrm{T}}$ – многоканальная запись наблюдений на M датчиках группы, $\mathbf{x}_l(t)$ – трехканальная запись l-го датчика группы, $\mathbf{u}(t) = (\mathbf{u}_1^{\mathrm{T}}, \ldots, \mathbf{u}_M^{\mathrm{T}})^{\mathrm{T}}$ – сигналы, порождаемые анализируемой сейсмической волной, $\boldsymbol{\xi}(t) = (\xi_1(t), \ldots, \xi_M(t))^{\mathrm{T}}$ – помехи, воздействующие на датчики группы, $s(t)$ – форма колебаний сейсмической волны на центральном датчике группы, $\mathbf{h}(t, \mathbf{p}, v)$ – векторная импульсная переходная характеристика среды, определяющая преобразования колебаний волны при ее распространении от центрального датчика к другим датчикам группы.

После дискретизации наблюдений $\mathbf{x}(t)$ с частотой дискретизации $f_{\text{д}} > 2f_{\text{в}}$ (где $f_{\text{в}}$ – верхняя граничная частота колебаний волны) модель наблюдений (1.1) можно представить в частотной области дискретного преобразования Фурье

$$\mathbf{x}(f) = \mathbf{h}(f, \mathbf{p}, v)s(f) + \boldsymbol{\xi}(f), \quad f \in [0, f_{\text{д}}/2], \tag{1.2}$$

где $\mathbf{h}(f, \mathbf{p}, v)$ – векторная частотная характеристика среды (преобразование Фурье от $\mathbf{h}(t, \mathbf{p}, v)$).

Мы будем использовать два предположения относительно $s(t)$ (временно́й функции сейсмической волновой фазы), существенные для оценивания направления сейсмической волны:

а) $s(t)$ – есть реализация гауссовского регулярного стационарного временно́го ряда с нулевым средним и спектральной плотностью $\nu_s(f)$,

б) $s(t)$ – есть неизвестная детерминированная последовательность.

Хотя предположение а) может показаться искусственным, с точки зрения сейсмологов-практиков, в действительности оно означает, что при синтезе и анализе алгоритмов оценивания мы ограничиваемся учетом только усредненного энергетического спектра сигнала, игнорируя всякую информацию о фазовом его спектре. Это оправдано, так как на практике последняя информация, как правило, отсутствует. Поэтому предположения о гауссовости и стационарности временно́й функции волновой фазы являются чисто методическим приемом, позволяющим использовать аппарат статистического анализа временны́х рядов.

В сейсмологической практике направление прихода сейсмической волны, как правило, определяют в условиях, когда временна́я форма волны $s(t)$ и даже ее энергетический спектр $\nu_s(f)$ неизвестны и изменяются от случая к случаю в широких пределах. При этом весьма существенно, что сигнал $s(t)$ является широкополосным, и замена его распространенной моделью Берлаге

$$s(t) = A \exp[-\alpha t] \cos(2\pi f_0 t + \varphi)$$

(где f_0 – центральная частота спектра волны) является неоправданным упрощением, тем более что f_0 и α также неизвестны априори. В связи с этим при оценивании информативных параметров, т.е. вектора кажущейся медленности $\mathbf{p} = (p_x, p_y)$, по наблюдениям, удовлетворяющим модели (1.1), (1.2), приходится вводить в рассмотрение мешающие параметры, связанные с отсутствием априорной информации о временно́й форме сигнальной волны, а также (при наличии трехкомпонентных датчиков в малоапертурной группе) о фазовой скорости v сейсмической волны.

Для модели а) случайного гауссовского сигнала $s(t)$ будем предполагать, что его энергетический спектр $\nu_s(f)$ известен с точностью до q мешающих параметров – вектора $\mathbf{c} = (c_k,\ k \in \overline{1,q})$, т.е. $\nu_s(f) = \varphi(f, \mathbf{c})$. Простейшей моделью такого рода является линейная модель:

$$\nu_s(f) = \varphi(f, \mathbf{c}) = \sum_{k=1}^{q} c_k \varphi_k(f), \tag{1.3}$$

где $\varphi_k(f)$, например, – типичные спектры мощности сейсмических сигналов в различных частотных полосах: $\varphi_k(f) = \varphi(f - f_k)$, f_k – центральные частоты этих полос.

Поскольку сигнал $s(t)$ – широкополосный, число q мешающих параметров в задаче оценивания кажущейся медленности достаточно мало.

Для модели б), где волновая форма сейсмической фазы рассматривается в виде детерминированной последовательности $s(t)$, естественно полагать, что все элементы этой последовательности неизвестны, т.е. в задаче оценивания информативных параметров p_x, p_y имеется N скалярных мешающих параметров $s_k = s(k/f_д)$, число которых равно числу векторных наблюдений $\mathbf{x}_k = \mathbf{x}(k/f_д)$, $k \in \overline{1,N}$, $N = T/f_д$.

Таким образом, введенная в настоящем разделе статистическая модель наблюдений (1.1), (1.2) позволяет рассматривать задачу определения кажущихся медленностей p_x, p_y плоской сейсмической волны по данным группы трехкомпонентных датчиков как задачу статистического оценивания с мешающими параметрами. Общность постановки задачи связана еще с тем, что наблюдения $\mathbf{x}_k$ предполагаются коррелированными как по отсчетам k, $k \in \overline{1,N}$, так и по координатам x_{kn}, x_{kl}, $n, l = \overline{1,M}$ векторов $\mathbf{x}_k$. Ограничением для рассматриваемой модели наблюдений является то, что при выводе и анализе оптимальных алгоритмов оценивания мы будем предполагать, что матричная спектральная плотность помех $\mathbf{F}(f)$ известна. Это ограничение может быть оправдано тем, что помехи $\boldsymbol{\xi}(t)$ во многих случаях доступны наблюдению до прихода сигнальной волны $\mathbf{u}(t)$. Спектральная плотность помех при этом может быть оценена специальной процедурой адаптации [4], и оценка $\hat{\mathbf{F}}(f)$ далее использована в оптимальных алгоритмах оценивания вектора кажущейся медленности.

В связи с неполной информацией о спектре помех важное значение приобретает вопрос об устойчивости статистически оптимальных алгоритмов оценки вектора кажущейся медленности к отклонениям матричного спектра $\tilde{\mathbf{F}}(f)$ реальных помех от предполагаемого спектра $\mathbf{F}(f)$, используемого в алгоритмах определения вектора $\mathbf{p}$. Этот вопрос в общем виде рассмотрен в работе [5], где даны методы оценки устойчивости и рекомендации по построению алгоритмов оценивания робастных к спектру помех.

Наконец, в случае когда оценивание спектральной матрицы помех $\mathbf{F}(f)$ по предварительным наблюдениям невозможно, остается путь расширения пространства мешающих параметров задачи за счет параметризации матрицы $\mathbf{F}(f)$ новыми мешающими параметрами помех $\boldsymbol{\vartheta}$: $\mathbf{F}(f) = \mathbf{F}(f, \boldsymbol{\vartheta})$, $\boldsymbol{\vartheta} = (\vartheta_1, \ldots, \vartheta_r)^{\mathrm{T}}$ [5].

В настоящей работе для простоты изложения будем рассматривать класс оценок $\mathbb{K}$ кажущейся медленности, обладающих моментами второго порядка. Такие оценки практически целесообразны в рассматриваемой задаче, так как гарантируют отсутствие случайных аномально больших выбросов при оценивании. В качестве критерия точности произвольной оценки кажущейся медленности $\tilde{\mathbf{p}}_N = (\tilde{\mathbf{p}}_{xN}, \tilde{\mathbf{p}}_{yN})^{\mathrm{T}} \in \mathbb{K}$ будем использовать ее асимптотическую ковариационную матрицу

$$\boldsymbol{\Psi}(\mathbf{p}, \boldsymbol{\theta}|\tilde{\mathbf{p}}) = \overline{\lim_{N\to\infty}} N \operatorname{cov}_{p\theta}(\tilde{\mathbf{p}}_N), \tag{1.4}$$

где $\operatorname{cov}_{p\theta}(\tilde{\mathbf{p}}_N) = E_{p\theta}\{(\tilde{\mathbf{p}}_N - \mathbf{p})(\tilde{\mathbf{p}}_N - \mathbf{p})^{\mathrm{T}}\}$, математическое ожидание берется при истинных значениях параметров $\mathbf{p}$ и $\boldsymbol{\theta}$; $\boldsymbol{\theta}$ – совокупность всех мешающих параметров задачи.

При конечных N эта матрица определяет ковариацию оценки $\tilde{\mathbf{p}}_N$ по формуле

$$\operatorname{cov}_{p\theta}(\tilde{\mathbf{p}}_N) = \boldsymbol{\Psi}(\mathbf{p}, \boldsymbol{\theta}|\tilde{\mathbf{p}})/N + O(1/N^2),$$

где матрица $O(\alpha)$ такова, что $\|O(\alpha)\|/\alpha < C$ при $\alpha \to 0$, $\|A\|$ – евклидова норма матрицы: $\|A\|^2 = \sum_{i,j=1}^{M} a_{ij}^2$.

Поскольку сравнение оценок удобнее осуществлять, используя какую-либо числовую меру их качества, введем риск оценки $r(\boldsymbol{\Psi})$ в виде числовой функции от ее асимптотической ковариационной матрицы (1.4), ограничившись классом положительных неубывающих функций над полуупорядоченным множеством неотрицательно определенных матриц, растущих не быстрее, чем экспоненциально,с ростом евклидовой нормы матрицы. Пример такой функции риска есть $r(\boldsymbol{\Psi}) = \mathrm{tr}(\boldsymbol{\Psi})$ – сумма асимптотических дисперсий оценки $\tilde{\mathbf{p}}_N$.

Наилучшую с точки зрения критерия качества $r(\boldsymbol{\Psi})$ оценку $\hat{\mathbf{p}}_N$ будем называть асимптотически эффективной (АЭ) в $\mathbb{K}$ и определять как оценку, для которой при любом $\delta > 0$ верно неравенство

$$\sup_{|\nu-\theta|<\delta,\ |\rho-p|<\delta} r(\boldsymbol{\Psi}(\boldsymbol{\rho},\boldsymbol{\nu}|\hat{\mathbf{p}})) \le \sup_{|\nu-\theta|<\delta,\ |\rho-p|<\delta} r(\boldsymbol{\Psi}(\boldsymbol{\rho},\boldsymbol{\nu}|\tilde{\mathbf{p}})), \tag{1.5}$$

где $\tilde{\mathbf{p}}_N$ – произвольная оценка параметра $\mathbf{p}$ из класса $\mathbb{K}$.

Выбор критерия качества оценок в виде (1.4), (1.5) связан с тем, что построение в явной форме оценок, наилучших в смысле какого-либо неасимптотического критерия, например эффективных (имеющих наименьшую дисперсию при заданном размере выборки N),– задача неразрешимая в теоретическом и практическом отношениях. В то же время при асимптотическом подходе возможно использование аналитического аппарата асимптотической теории оценивания [6], позволяющего провести синтез и анализ АЭ-оценок в рассматриваемой задаче до конца и получить явные алгоритмы оценивания и формулы для их асимптотической ковариации.

2. АСИМПТОТИЧЕСКИ ЭФФЕКТИВНЫЕ ОЦЕНКИ ПРИ СЛУЧАЙНОЙ ВОЛНОВОЙ ФОРМЕ СИГНАЛА

В предположении, что сигнал $s(t)$ в модели наблюдений (1.1), (1.2) есть реализация гауссовского стационарного процесса с нулевым средним, векторный временной ряд наблюдений $\mathbf{x}(t) = \mathbf{u}(t) + \boldsymbol{\xi}(t)$ представляет собой реализацию многомерного гауссовского стационарного процесса с нулевым средним, распределение которого полностью определяется его матричной спектральной плотностью $\mathbf{F}_x(f)$, $f \in [0, f_д/2]$, где $f_д$ – частота дискретизации $\mathbf{x}(t)$. При дополнительном естественном предположении о статистической независимости временно́й функции сигнальной волны $s(t)$ и помех $\boldsymbol{\xi}(t)$ имеем:

$$\mathbf{F}_x(f) = \mathbf{F}(f) + \nu_s(f)\,\mathbf{h}(f,\mathbf{p},v)\,\mathbf{h}^*(f,\mathbf{p},v) = \mathbf{F}(f) + \varphi(f,\mathbf{c})\,\mathbf{H}(f,\mathbf{p},v), \tag{2.1}$$

где $\mathbf{h}(f,\mathbf{p},v) = (\exp(-i2\pi f\mathbf{r}_l^{\mathrm{T}}\mathbf{p})\mathbf{b}(\mathbf{p},v),\ l = \overline{1,m})$ – $3m$-мерный вектор-столбец частотных характеристик путей распространения волны от первого к остальным датчикам группы; $\mathbf{b}(\mathbf{p},v)$ – трехмерный вектор поляризации данной волновой фазы [7]; $\mathbf{r}_l^{\mathrm{T}} = (r_{lx}, r_{ly})$ – координаты l-го датчика группы; $\varphi(f,\mathbf{c})$ – линейно параметризованная спектральная плотность сигнала, выражаемая формулой (1.3); $\mathbf{H}(f,\mathbf{p},v) = \mathbf{h}(f,\mathbf{p},v)\,\mathbf{h}^*(f,\mathbf{p},v)$ – $3m \times 3m$ матрица. Отметим, что матричная спектральная плотность наблюдений весьма просто зависит от информативных и мешающих параметров задачи.

Как показано в работах [6,5], при весьма слабых ограничениях на вероятностную модель наблюдений АЭ-оценка $\hat{\mathbf{p}}_N$ информативного параметра p получается совместно с оценкой $\hat{\boldsymbol{\theta}}_N^{\mathrm{T}} = (\hat{\mathbf{c}}_N^{\mathrm{T}}, \hat{v}_N^{\mathrm{T}})$ мешающего параметра $\boldsymbol{\theta} = (\mathbf{c}^{\mathrm{T}}, v^{\mathrm{T}})$ методом максимума правдоподобия:

$$(\hat{\mathbf{p}}_N, \hat{\boldsymbol{\theta}}_N) = \arg\max_{p,\theta}(L(\mathbf{X}_N, \mathbf{p}, \boldsymbol{\theta})),$$

где $L(\mathbf{X}_N, \mathbf{p}, \boldsymbol{\theta}) = \ln(W(\mathbf{X}_N, \mathbf{p}, \boldsymbol{\theta}))$, $W(\mathbf{X}_N, \mathbf{p}, \boldsymbol{\theta})$ – многомерная совместная плотность наблюдений $\mathbf{X}_N = (\mathbf{x}_1^{\mathrm{T}}, \mathbf{x}_2^{\mathrm{T}}, \ldots, \mathbf{x}_N^{\mathrm{T}})^{\mathrm{T}}$. При этом асимптотическая ковариационная матрица совместной оценки $\hat{\boldsymbol{\vartheta}} = (\hat{\mathbf{p}}_N^{\mathrm{T}}, \hat{\boldsymbol{\theta}}_N^{\mathrm{T}})^{\mathrm{T}}$ определяется пределом нормированной информационной матрицы Фишера:

$$\boldsymbol{\Psi}_\vartheta = \lim_{N\to\infty} N\, E_\vartheta\{(\hat{\boldsymbol{\vartheta}}_N - \boldsymbol{\vartheta})(\hat{\boldsymbol{\vartheta}}_N - \boldsymbol{\vartheta})^{\mathrm{T}}\} = \boldsymbol{\Phi}_\vartheta^{-1}, \tag{2.2}$$

где $\boldsymbol{\Phi}_\vartheta = \lim\limits_{N\to\infty}(1/N)\boldsymbol{\Phi}_{\vartheta N}$,

$$\boldsymbol{\Phi}_{\vartheta N} = [E_\vartheta\{(\delta L(X_N, \boldsymbol{\vartheta})/\delta\vartheta_k)(\delta L(X_N, \boldsymbol{\vartheta})/\delta\vartheta_l)\}, \quad k, l = \overline{1, q+3}]$$

– информационная матрица Фишера, E_ϑ – математическое ожидание при истинных значениях параметров $\boldsymbol{\vartheta} = (\mathbf{p}^{\mathrm{T}}, \boldsymbol{\theta}^{\mathrm{T}})^{\mathrm{T}}$.

В рассматриваемой задаче с мешающими параметрами матрица Фишера $\boldsymbol{\Phi}_{\vartheta N}$ и ее нормированный предел $\boldsymbol{\Phi}_\vartheta$ имеют блочную структуру:

$$\boldsymbol{\Phi}_\vartheta = \begin{bmatrix} \boldsymbol{\Phi}_{pp} & \boldsymbol{\Phi}_{p\theta} \\ \boldsymbol{\Phi}_{\theta p} & \boldsymbol{\Phi}_{\theta\theta} \end{bmatrix}.$$

Точно также блочную структуру имеет обратная матрица:

$$\boldsymbol{\Phi}_\vartheta^{-1} = \begin{bmatrix} \mathbf{Q}_{pp} & \mathbf{Q}_{p\theta} \\ \mathbf{Q}_{\theta p} & \mathbf{Q}_{\theta\theta} \end{bmatrix},$$

причем, согласно правилам обращения блочных матриц,

$$\mathbf{Q}_{pp} = \left[\boldsymbol{\Phi}_{pp} - \boldsymbol{\Phi}_{p\theta}\boldsymbol{\Phi}_{\theta\theta}^{-1}\boldsymbol{\Phi}_{\theta p}\right]^{-1}.$$

Из (2.2) следует, что асимптотическая матрица ковариаций АЭ-оценки информативных параметров $\hat{\mathbf{p}}_N = (\hat{p}_{xN}, \hat{p}_{yN})^{\mathrm{T}}$ удовлетворяет соотношению

$$\boldsymbol{\Psi}_p = \lim_{N\to\infty} N E_{p\theta}\{(\hat{\mathbf{p}}_N - \mathbf{p})(\hat{\mathbf{p}}_N - \mathbf{p})^{\mathrm{T}}\} = \mathbf{Q}_{pp} = \left[\boldsymbol{\Phi}_{pp} - \boldsymbol{\Phi}_{p\theta}\boldsymbol{\Phi}_{\theta\theta}^{-1}\boldsymbol{\Phi}_{\theta p}\right]^{-1}. \tag{2.3}$$

Ясно, что в ситуации, когда мешающие параметры отсутствуют (т.е. вектор $\boldsymbol{\theta} = (\mathbf{c}^{\mathrm{T}}, v)^{\mathrm{T}}$ – известен), асимптотическая ковариация информативных параметров

$$\boldsymbol{\Psi}_p = \boldsymbol{\Phi}_{pp}^{-1}. \tag{2.4}$$

В то же время, поскольку $\boldsymbol{\Phi}_{p\theta} = \boldsymbol{\Phi}_{\theta p}^{\mathrm{T}}$ и матрица $\boldsymbol{\Phi}_{\theta\theta}$ – положительно определенная, можно утверждать, что при неизвестных мешающих параметрах

$$\mathrm{tr}(\mathbf{\Psi}_p) = \mathrm{tr}\big(\big[\mathbf{\Phi}_{pp} - \mathbf{\Phi}_{p\theta}\mathbf{\Phi}_{\theta\theta}^{-1}\mathbf{\Phi}_{\theta p}\big]^{-1}\big) \geq \mathrm{tr}(\mathbf{\Phi}_{pp}^{-1}), \tag{2.5}$$

т.е. сумма асимптотических дисперсий АЭ-оценок информативных параметров в присутствии неизвестных мешающих параметров всегда больше, чем в отсутствие последних. Формулы (2.3) и (2.4) позволяют подсчитать проигрыш в качестве оценивания информативных параметров за счет появления в задаче априорной неопределенности, отражаемой введением мешающих параметров.

Отметим, что матрица $\mathbf{\Psi}_p(\mathbf{p}, \boldsymbol{\theta})$ в (2.3) определяет нижнюю границу точности для произвольных оценок $\tilde{\mathbf{p}}_N(\mathbf{X}_N)$ информативного параметра $\mathbf{p}$. Так, если $r(N\mathrm{cov}_{p\theta}(\tilde{\mathbf{p}}_N))$ – произвольная функция риска оценки $\tilde{\mathbf{p}}_N(\mathbf{X}_N)$, зависящая от ковариационной матрицы оценки и удовлетворяющая весьма слабым с точки зрения практики ограничениям [6], то для любой оценки $\tilde{\mathbf{p}}_N(\mathbf{X}_N)$ при любом $\delta > 0$ верно:

$$\sup_{|\rho-p|<\delta,\ |\nu-\theta|<\delta} \lim_{N\to\infty} r\big[N\mathrm{cov}_{\rho\nu}(\tilde{\mathbf{p}}_N)\big] \geq r[\mathbf{\Psi}_p(\mathbf{p}, \boldsymbol{\theta})], \quad \mathbf{p} \in \mathfrak{P}, \ \ \boldsymbol{\theta} \in \mathfrak{G}, \tag{2.6}$$

где $\mathfrak{P}$ и $\mathfrak{G}$ – ограниченные области в пространствах информативного и мешающего параметров.

Займемся построением алгоритма вычисления АЭ-оценок для параметров $\mathbf{p}$, $\boldsymbol{\theta}$ в модели наблюдений (1.1), (1.2) со случайным гауссовским стационарным сигналом. Логарифм гауссовской многомерной плотности распределения совокупности наблюдений $\mathbf{X}_N$ (функция правдоподобия) имеет вид

$$L(\mathbf{X}_N, \boldsymbol{\vartheta}) = -(MN/2)\ln(2\pi) - (1/2)\ln(\det[\mathbf{C}_N(\boldsymbol{\vartheta})]) - (1/2)\mathbf{X}_N^{\mathrm{T}}\mathbf{C}_N^{-1}(\boldsymbol{\vartheta})\mathbf{X}_N, \tag{2.7}$$

где $\mathbf{C} = [\mathbf{C}_{t-\rho},\ t, \rho = \overline{1, N}]$ – блочная теплицева матрица размера $NM \times NM$, составленная из блоков $\mathbf{C}_\tau = E\{\mathbf{x}_k\mathbf{x}_{k+\tau}^{\mathrm{T}}\} = E\{\mathbf{u}_k\mathbf{u}_{k+\tau}^{\mathrm{T}}\} + E\{\boldsymbol{\xi}_k\boldsymbol{\xi}_{k+\tau}^{\mathrm{T}}\}$.

Функция правдоподобия (2.7) зависит от параметров задачи $\boldsymbol{\vartheta} = (\mathbf{p}^{\mathrm{T}}, \boldsymbol{\theta}^{\mathrm{T}})^{\mathrm{T}}$ через значения элементов обратной матрицы $\mathbf{C}_N^{-1}(\boldsymbol{\vartheta})$ и ее детерминанта. Ясно, что построение каких-либо вычислительных процедур, связанных с максимизацией $L(\mathbf{X}_N, \boldsymbol{\vartheta})$ по параметрам ϑ, практически невозможно при сколько-нибудь значительных размерах $M \times N$. Однако при достаточно большом количестве наблюдений N возможна простая аппроксимация $L(\mathbf{X}_N, \boldsymbol{\vartheta})$ в частотной области, впервые указанная в работе [8].

Запишем модель наблюдений (1.1) в дискретной спектральной области:

$$\mathbf{x}_j = \mathbf{u}_j + \boldsymbol{\xi}_j, \qquad j \in \overline{1, N}, \tag{2.8}$$

где $\mathbf{x}_j \Leftrightarrow \mathbf{x}_k$, $\mathbf{u}_j \Leftrightarrow \mathbf{u}_k$, $\boldsymbol{\xi}_j \Leftrightarrow \boldsymbol{\xi}_k$ – дискретные конечные преобразования Фурье (ДКПФ) от $\mathbf{x}_k$, $\mathbf{u}_k$ и $\boldsymbol{\xi}_k$, соответственно; ДКПФ выражается формулой

$$\mathbf{a}_j = \frac{1}{\sqrt{N}}\sum_{k=1}^{N}\mathbf{b}_k e^{ik\lambda_j}, \quad \lambda_j = 2\pi f_j/f_{\text{д}}, \quad f_j = jf_{\text{д}}/N. \tag{2.9}$$

По теореме сдвига для преобразования Фурье, которая для ДКПФ верна в циклической форме, а в обычной формулировке – лишь асимптотически, можно записать

$$\mathbf{u}_j = \mathbf{h}_j(\mathbf{p}, v)s_j + \mathbf{O}_j(1/\sqrt{N}), \tag{2.10}$$

где $\mathbf{h}_j(\mathbf{p}, v) = (\exp(i2\pi f_j \mathbf{r}_k^{\mathrm{T}} \mathbf{p}) \mathbf{b}(\mathbf{p}, v),\ k \in \overline{1, M})$; s_j – скалярный дискретный спектр отсчетов временнóй функции сейсмической волны,

$$\max_{j \in 1,N} |\mathbf{O}_j(1/\sqrt{N})| \to 0.$$

ДКПФ шумов группы $\boldsymbol{\xi}_j$ есть совокупность случайных нормально распределенных векторов со следующими характеристиками [9]:

$$E\{\boldsymbol{\xi}_j\} = 0, \quad E\{\boldsymbol{\xi}_j \boldsymbol{\xi}_l^*\} = \delta_{jl} \mathbf{F}_\xi(f_j) + \mathbf{O}_{jl}(1/N), \tag{2.11}$$

где $\delta_{jl} = \begin{cases} 1 & j = l \\ 0 & j \neq l \end{cases}$ – символ Кронекера,

$$\max_{l,j \in 1,N} \|\mathbf{O}_{jl}(1/N)\| \to 0 \ \text{ при } \ N \to \infty.$$

Таким образом, значения ДКПФ сильно коррелированных по времени шумов группы слабо коррелированы между собой при достаточно больших N, т.е. ДКПФ является асимптотически декоррелирующим преобразованием.

Для случайной временнóй функции сейсмической волны на основании (2.8)–(2.11) можем заключить, что $\mathbf{x}_j$ есть совокупность асимптотически некоррелированных нормальных векторов с параметрами

$$E\{\mathbf{x_j}\} = 0, \quad E\{\mathbf{x}_j \mathbf{x}_l^*\} = \delta_{jl} \mathbf{F}_x(f_j) + \mathbf{O}_{jl}(1/N), \tag{2.12}$$

где $\max_{j,l} \|\mathbf{O}_{jl}(1/N)\| \to 0$ при $N \to \infty$.

Пренебрегая слабой статистической зависимостью значений $\mathbf{x}_j$ (т.е. отбрасывая в (2.12) члены $\mathbf{O}_{jl}(1/N)$), можно записать приближенное выражение правдоподобия для функции правдоподобия наблюдений (2.8) в частотной области:

$$L(\mathbf{X}_N, \boldsymbol{\vartheta}) = \ln W(\mathbf{x}_1, \ldots, \mathbf{x}_N; \boldsymbol{\vartheta}) \simeq -\frac{MN}{2} \ln 2\pi - \frac{1}{2} \sum_{j=1}^{N} \ln \det \mathbf{F}_{xj} -$$
$$- \frac{1}{2} \sum_{j=1}^{N} \mathbf{x}_j^* \mathbf{F}_{xj}^{-1} \mathbf{x}_j = L_f(\mathbf{X}_N, \boldsymbol{\vartheta}), \tag{2.13}$$

где $\mathbf{F}_{xj}^{-1} = \mathbf{F}_x^{-1}(f_j)$.

Более точные оценки остаточных членов $\mathbf{O}_{jl}(1/N)$ в (2.11) и (2.12) позволяют записать следующее соотношение между функцией правдоподобия (2.7) во временнóй области и функцией правдоподобия (2.13) в частотной области [10]:

$$L(\mathbf{X}_N, \mathbf{p}, \boldsymbol{\theta}) = L_f(\mathbf{X}_N, \mathbf{p}, \boldsymbol{\theta}) + O(\sqrt{N}). \tag{2.14}$$

Видим, что "расстояние" между L и L_f с ростом N увеличивается. Однако, поскольку порядок величины L_f есть $O(N)$, можно считать, что L_f содержит "главную часть" точной функции правдоподобия наблюдений в рассматриваемой задаче.

Из приведенных рассуждений следует важный практический вывод: для нахождения АЭ-оценки совокупного вектора параметров задачи (вектора кажущейся медленности $\mathbf{p}$ и мешающих параметров $\boldsymbol{\theta}$) достаточно найти максимум приближенной функции правдоподобия наблюдений в частотной области

$$\hat{\boldsymbol{\vartheta}}_N = \arg\max_{\vartheta} L_f(\mathbf{X}_N, \boldsymbol{\vartheta}). \tag{2.15}$$

Следовательно, алгоритм АЭ-оценки может быть построен как процедура максимизации функционала

$$\Lambda(\mathbf{X}_N, \boldsymbol{\vartheta}) = \sum_{j=1}^{N} \ln\det \mathbf{F}_{xj}^{-1}(\boldsymbol{\vartheta}) - \sum_{j=1}^{N} \mathbf{x}_j^* \mathbf{F}_{xj}^{-1}(\boldsymbol{\vartheta}) \mathbf{x}_j, \tag{2.16}$$

где $\mathbf{F}_{xj}^{-1}(\boldsymbol{\vartheta}) = \mathbf{Q}_j(\mathbf{p}, \boldsymbol{\theta}) \mathbf{F}_j^{-1}$;

$$\mathbf{Q}_j(\mathbf{p}, \boldsymbol{\theta}) = \left[\mathbf{I} - \frac{\mathbf{F}_j^{-1} \mathbf{H}_j(\mathbf{p}, v)}{\varphi_j^{-1}(\mathbf{c}) + \operatorname{tr} \mathbf{F}_j^{-1} \mathbf{H}_j(\mathbf{p}, v)}\right], \tag{2.17}$$

индекс j означает зависимость соответствующих функций от f_j. Последняя формула для $\mathbf{F}_{xj}^{-1}(\boldsymbol{\vartheta})$ представляет собой известную формулу Бартлетта для обращения матриц вида (2.1) [11].

Можно преобразовать формулу (2.16) для функционала Λ к виду, который упрощает его вычисления в процессе итеративной максимизации и проясняет его "физический" смысл:

$$\Lambda(\mathbf{X}_N, \boldsymbol{\vartheta}) = C + \sum_{j=1}^{N} \ln\det \mathbf{Q}_j(\mathbf{p}, \boldsymbol{\theta}) + \sum_{j=1}^{N} |z_j(\mathbf{p}, \boldsymbol{\theta})|^2, \tag{2.18}$$

где $C = \sum_{j=1}^{N} \ln\det \mathbf{F}_j^{-1} + \sum_{j=1}^{N} \mathbf{x}_j^* \mathbf{F}_j^{-1} \mathbf{x}_j$

– член, не зависящий от параметров задачи $\boldsymbol{\vartheta}$;

$$z_j(\boldsymbol{\vartheta}) = \frac{\mathbf{h}_j^*(\mathbf{p}, v) \mathbf{F}_j^{-1} \mathbf{x}_j}{\sqrt{\varphi_j^{-1}(\mathbf{c}) + \mathbf{h}_j^*(\mathbf{p}, v) \mathbf{F}_j^{-1} \mathbf{h}_j(\mathbf{p}, v)}}$$

– выходной сигнал оптимального винеровского фильтра, который преобразует многоканальный сигнал группы в одноканальную трассу, максимизирует отношение сигнал/шум в этой трассе (без условия несмещенности оценки сигнала) и обеляет выходную помеху.

Обозначая

$$\mathbf{T}_j = \mathbf{F}_j^{-1} \mathbf{x}_j \mathbf{x}_j^*, \quad \boldsymbol{\Psi}_j(\boldsymbol{\vartheta}) = \mathbf{F}_j^{-1} \mathbf{H}_j(\boldsymbol{\vartheta}),$$
$$\psi_j(\boldsymbol{\vartheta}) = \varphi_j^{-1}(\mathbf{c}) + \operatorname{tr} \boldsymbol{\Psi}_j(\boldsymbol{\vartheta}),$$

получим расчетную формулу для функционала Λ:

$$\Lambda(\mathbf{X}_N, \boldsymbol{\vartheta}) = C + \sum_{j=1}^{N} \ln\det\left[\mathbf{I} - \frac{\boldsymbol{\Psi}_j(\boldsymbol{\vartheta})}{\psi_j(\boldsymbol{\vartheta})}\right] + \sum_{j=1}^{N} \frac{\operatorname{tr} \mathbf{T}_j \boldsymbol{\Psi}_j(\boldsymbol{\vartheta})}{\psi_j(\boldsymbol{\vartheta})}. \tag{2.19}$$

В процессе итеративной максимизации функционала Λ матрицы $\mathbf{T}_j$, зависящие от наблюдений, остаются неизменными, а пересчитываются только матрицы $\boldsymbol{\Psi}_j(\boldsymbol{\vartheta})$ и скаляры $\psi_j(\boldsymbol{\vartheta})$. Нахождение экстремума функционала (2.19) облегчается тем, что существуют простые явные выражения для его частных производных по параметрам задачи $\mathbf{p}$, $\mathbf{c}$ и v.

В рассматриваемой задаче оценивания вектора $\boldsymbol{\vartheta}$ нетрудно также найти аналитическое выражение для предела нормированной матрицы Фишера. Это позволяет, пользуясь формулой (2.3), аналитически вычислить асимптотическую ковариационную матрицу ошибок для оценки максимума правдоподобия вектора $\mathbf{p}$.

Анализ выражения (2.18) показывает, что при больших отношениях сигнал/шум, когда при всех j величинами $\varphi_j^{-1}(\mathbf{c})$ можно пренебречь по сравнению с $\mathbf{h}_j^*(\mathbf{p}, v)\mathbf{F}_j^{-1}\ \mathbf{h}_j(\mathbf{p}, v)$, зависимость функционала $\Lambda(\mathbf{X}_N, \mathbf{p}, \mathbf{c}, v)$ от неизвестного спектра сигнала $\varphi(f_j, \mathbf{c})$ ослабевает и при $\varphi^{-1} \to 0$ исчезает совсем. (Это не столь очевидно и требует внимательного исследования, поскольку матрица $\mathbf{Q}(\mathbf{p}, \mathbf{c}, v)$ в этом случае становится вырожденной и первое слагаемое в (2.18) неограниченно возрастает.) Таким образом, при большом отношении сигнал/шум АЭ-оценка вектора кажущейся скорости не намного отличается от оценки, максимизирующей функционал

$$\tilde{\Lambda}(\mathbf{X}_N, \mathbf{p}) = \sum_{j=1}^{N} |\tilde{z}_j(\mathbf{p}, v)|^2, \tag{2.20}$$

где $\quad \tilde{z}_j(\mathbf{p}, v) = \dfrac{\mathbf{h}_j^*(\mathbf{p}, v)\mathbf{F}_j^{-1}\mathbf{x}_j}{\sqrt{\mathbf{h}_j^*(\mathbf{p}, v)\mathbf{F}_j^{-1}\mathbf{h}_j(\mathbf{p}, v)}}.$

Иными словами, при больших отношениях сигнал/шум необходимо принимать во внимание лишь матричную спектральную плотность помех, в то время как информация о спектре сигнала оказывается несущественной. Усложнение статистически оптимальной процедуры оценивания кажущейся скорости за счет оценивания спектра сигнала оказывается в данном случае нецелесообразным.

Рассмотрим наконец частный случай функционала (2.18), когда и шум, и сигнал могут считаться белыми, а шум – некоррелированным по пространству: $\mathbf{F}(f) = \mathbf{I}\sigma^2$, $\varphi(\mathrm{c}) = \sigma_s^2$. Тогда

$$\Lambda(\mathbf{X}_N, \mathbf{p}, \sigma^2, \sigma_s^2, v) = C + \sum_{j=1}^{N} \ln\det \sigma^{-2}\left[\mathbf{I} - \frac{\mathbf{H}_j(\mathbf{p}, v)}{M + \sigma^2/\sigma_s^2}\right] -$$

$$-[\sigma^2(M + \sigma^2/\sigma_s^2)]^{-1} \sum_{j=1}^{N} |\mathbf{h}_j(\mathbf{p}, v)\mathbf{x}_j|^2.$$

При больших отношениях сигнал/шум ($\sigma^2/\sigma_s^2 \ll M$) статистически оптимальная оценка получается максимизацией функционала

$$\tilde{\Lambda}^w(\mathbf{X}_N, \mathbf{p}, v) = \sum_{j=1}^{N} |\mathbf{h}_j(\mathbf{p}, v)\mathbf{x}_j|^2$$

и совпадает с известной оценкой по методу широкополосного спектрально-временно́го анализа [1].

3. АСИМПТОТИЧЕСКИ ЭФФЕКТИВНЫЕ ОЦЕНКИ ПРИ МАЛОМ ОТНОШЕНИИ СИГНАЛ/ШУМ

Рассмотренные в разд. 2 АЭ-оценки вектора кажущейся медленности являются наилучшими с точки зрения асимтотического критерия качества (1.4),(1.5). Однако нахождение экстремума функционала (2.18) численными методами является достаточно трудоемкой вычислительной задачей. Трудность ее решения усугубляется тем, что необходимо максимизировать (2.18) по $q+3$ параметрам, из которых $q+1$ являются мешающими, т.е., по существу, ненужными для основной проблемы. Их наличие в функционале (2.18) усложняет и замедляет итерационную процедуру оптимизации.

В ряде приложений существенный интерес представляет оценивание кажущейся медленности плоской сигнальной волны на фоне интенсивных когерентных помех. Как уже отмечалось, примерами такой ситуации могут служить задачи оценивания параметров слабой сейсмической фазы на фоне коды сильной предыдущей фазы или параметров волны от ядерного взрыва, маскируемого сильным землетрясением.

В предположении малого отношения сигнал/шум в модели наблюдения (1.1)–(1.3) функционал (2.18) для построения асимптотически эффективных оценок можно существенно упростить. Математически корректная постановка задачи оценивания кажущейся медленности при малом отношении сигнал/шум заключается в анализе следующей модели наблюдений:

$$\mathbf{x}_k = 1/(\sqrt[4]{N})\mathbf{u}_k + \xi_k; \quad \mathbf{x}_k = (x_{k1}, \ldots, x_{kM})^{\mathrm{T}}. \tag{3.1}$$

При этом дискретный энергетический спектр отсчетов $\mathbf{x}_k$, $k \in \overline{1, N}$ при случайном сигнале $s(t)$ (случайной временно́й форме волны):

$$\mathbf{F}_x(f_j) = \frac{1}{\sqrt{N}}\mathbf{H}(f_j, \mathbf{p}, v)\sum_{l=1}^{q} c_l\varphi_l(f_j) + \mathbf{F}(f_j). \tag{3.2}$$

Можно показать [4, 5], что при слабых ограничениях на $\mathbf{F}(f)$ функция правдоподобия (2.13) наблюдений (3.1), (3.2) имеет следующее асимптотическое представление:

$$L(\mathbf{X}_N|\mathbf{p}, \mathbf{c}/\sqrt{N}, v) = -L(\mathbf{X}_N, 0) + \mathbf{c}^{\mathrm{T}}\boldsymbol{\delta}(\mathbf{X}_N, \mathbf{p}, v) - \frac{1}{2}\mathbf{c}^{\mathrm{T}}\Gamma(\mathbf{p}, v)\mathbf{c} + \alpha(\mathbf{X}_N, \mathbf{p}, \mathbf{c}, v), \tag{3.3}$$

где $$\boldsymbol{\delta}(\mathbf{X}_N,\mathbf{p},v)=\frac{1}{\sqrt{N}}\sum_{j=1}^{N}\left[|\mathbf{h}_j^*(\mathbf{p},v)\mathbf{F}_j^{-1}\mathbf{x}_j|^2-\mathbf{h}_j^*(\mathbf{p},v)\mathbf{F}_j^{-1}\mathbf{h}_j(\mathbf{p},v)\right]\boldsymbol{\varphi}_j$$

представляет собой векторную асимптотически достаточную статистику для мешающих параметров $\mathbf{c}/\sqrt{N}$; $\boldsymbol{\varphi}_j=(\varphi_1(f_j),\ldots,\varphi_k(f_j))^{\mathrm{T}}$; матрица

$$\boldsymbol{\Gamma}_N(\mathbf{p},v)=\frac{1}{N}\sum_{j=1}^{N}\left[\mathbf{h}_j^*(\mathbf{p},v)\mathbf{F}_j^{-1}\mathbf{h}_j(\mathbf{p},v)\right]^2\boldsymbol{\varphi}_j\boldsymbol{\varphi}_j^{\mathrm{T}}$$

размерности $k\times k$ есть асимптотическое приближение для нормированной матрицы Фишера для мешающих параметров $\mathbf{c}/\sqrt{N}$ [5,6]; $\alpha(\mathbf{X}_N,\mathbf{p},\mathbf{c},v)$ – остаточный член разложения, сходящийся к нулю по вероятности как случайный процесс в пространстве $\mathbb{C}[\mathbb{Q}\times\mathbb{P}]$ непрерывных функций от $\boldsymbol{\theta}$, $\mathbf{p}$ с равномерной метрикой, $\mathbb{Q}$ – ограниченное множество значений мешающих параметров $\boldsymbol{\theta}=(\mathbf{c}^{\mathrm{T}},v)^{\mathrm{T}}$, $\mathbb{P}$ – ограниченное множество значений информативных парметров $\mathbf{p}$; $L(\mathbf{X}_N,0)$ – функция правдоподобия наблюдений $\mathbf{X}_N$ при нулевом значении параметра $\mathbf{c}$, т.е. функция правдоподобия чистого шума $\boldsymbol{\xi}_k$, $k=\overline{1,N}$. Поскольку этот член асимптотического разложения (3.3) не зависит от параметров задачи $\mathbf{p}$, $\mathbf{c}$, v, его можно не принимать в расчет при дальнейших рассуждениях.

Из (3.3) следует, что оценка максимума правдоподобия для модели наблюдений (3.1),(3.2) асимптотически имеет вид

$$\tilde{\boldsymbol{\vartheta}}_N=(\tilde{\mathbf{p}}_N^{\mathrm{T}},\tilde{\boldsymbol{\theta}}_N^{\mathrm{T}})^{\mathrm{T}}=\arg\max_{p,c,v}\left[\mathbf{c}^{\mathrm{T}}\boldsymbol{\delta}(\mathbf{X}_N,\mathbf{p},v)-\frac{1}{2}\mathbf{c}^{\mathrm{T}}\boldsymbol{\Gamma}_N(\mathbf{p},v)\mathbf{c}\right]. \qquad (3.4)$$

Оценка $\tilde{\boldsymbol{\vartheta}}_N$ (3.4) существенно проще общей АЭ-оценки максимума правдоподобия $\hat{\boldsymbol{\vartheta}}_N$ (2.14)–(2.19). Действительно, максимизируя (3.4) по $\mathbf{c}$ при фиксированных $\mathbf{p}$ и v, нетрудно получить

$$\tilde{\mathbf{c}}(\mathbf{p},v)=\arg\max_{c}\left[\mathbf{c}^{\mathrm{T}}\boldsymbol{\delta}(\mathbf{X}_N,\mathbf{p},v)\right]=\boldsymbol{\Gamma}_N^{-1}(\mathbf{p},v)\boldsymbol{\delta}(\mathbf{X}_N,\mathbf{p},v). \qquad (3.5)$$

Подставляя (3.5) в (3.4), получим

$$(\tilde{\mathbf{p}}_N,\tilde{v}_N)=\arg\max_{p,v}R(\mathbf{X}_N,\mathbf{p},v), \qquad (3.6)$$

где

$$R(\mathbf{X}_N,\mathbf{p},v)=\boldsymbol{\delta}^{\mathrm{T}}(\mathbf{X}_N,\mathbf{p},v)\boldsymbol{\Gamma}_N^{-1}(\mathbf{p},v)\boldsymbol{\delta}(\mathbf{X}_N,\mathbf{p},v). \qquad (3.7)$$

Таким образом, оценивание вектора кажущейся медленности $\mathbf{p}$ и фазовой скорости v сейсмической волны при малом отношении сигнал/шум сводится к максимизации функционала (3.7) по $\mathbf{p}$ и v, что существенно проще, чем максимизация функционала (2.18). Это связано как с меньшим количеством вычислений для нахождения значений $R(\mathbf{X}_N,\mathbf{p},v)$ и его производных по p_x,p_y и v, так и с тем, что оптимизацию необходимо производить всего по трем параметрам: p_x,p_y,v, вместо $q+3$ параметров $(\mathbf{c},\mathbf{p},v)$. Это, как правило, существенно увеличивает скорость сходимости итерационных процедур оптимизации. Максимизация функционала

(3.7) облегчается еще и тем, что существуют простые аналитические выражения для его частных производных первого и второго порядка по параметрам.

Исследование свойств оценки $(\tilde{\mathbf{p}}_N, \tilde{v}_N)$ (3.6) в рамках асимптотического подхода сводится к доказательству $\sqrt{N}$ -состоятельности этой оценки, т.е. ее сходимости по вероятности со скоростью $1/\sqrt{N}$ к $(\mathbf{p}_0, v_0)$ – истинному значению вектора кажущейся медленности и скорости волны сейсмической фазы, и нахождению асимптотической ковариационной матрицы оценки кажущейся медленности $\tilde{\mathbf{p}}_N$, т.е.

$$\lim_{N\to\infty} E_{\vartheta_0} N(\tilde{\mathbf{p}}_N - \mathbf{p}_0)(\tilde{\mathbf{p}}_N - \mathbf{p}_0)^{\mathrm{T}} = \mathbf{K}_{\tilde{p}}(\vartheta_0).$$

Сравнение $\mathbf{K}_{\tilde{p}}(\vartheta_0)$ с нижней границей (2.3) для матриц ковариации произвольных оценок позволяет определить возможный проигрыш в асимптотическом качестве оценивания вектора кажущейся медленности с помощью алгоритма (3.6) по сравнению с алгоритмом максимума правдоподобия (2.15). Указанное исследование асимптотических свойств упрщенных оценок (3.6) можно провести, используя методы, развитые в работах [5, 12].

4. ОЦЕНКА ВЕКТОРА КАЖУЩЕЙСЯ МЕДЛЕННОСТИ ПРИ ПОЛНОСТЬЮ НЕИЗВЕСТНОЙ ВРЕМЕННО́Й ФУНКЦИИ СИГНАЛЬНОЙ ВОЛНЫ

При полностью неизвестной временно́й форме волновой фазы $s(t)$ (разд. 1, модель б)) спектральные отсчеты s_j, $j = \overline{1, N}$, входящие в выражение (2.10) для дискретного спектра наблюдений, полностью неизвестны, и их следует рассматривать как мешающие параметры задачи. Если отбросить в (2.10) малые слагаемые $\mathbf{O}_j(1/\sqrt{N})$, то получим, что наблюдения в дискретной частотной области удовлетворяют нелинейной регрессионной модели с неизвестными "регрессорами" s_j. Поскольку число мешающих параметров s_j неограниченно возрастает с ростом числа наблюдений N, то возникает вопрос: имеются ли при этой постановке задачи $\sqrt{N}$-состоятельные оценки, для которых ошибки оценивания с ростом N стремятся к нулю и существуют предельные ковариационные матрицы вида (1.4). Этот вопрос рассматривался для классической статистической модели "функционального соотношения" в работах [13, 14]. Отметим, что исследуемая в настоящей работе модель наблюдений существенно отличается от модели "функционального соотношения", благодаря коррелированности помех и существованию временно́й ковариационной функции сигнала s_t:

$$r_\tau^{(s)} = \lim_{N\to\infty} (1/N) \sum_{t=\tau+1}^{N} s_t s_{t-\tau} \not\equiv 0 \quad \text{при} \quad \tau \not\equiv 0.$$

Ниже мы построим методом максимума правдоподобия пример оценки $\overset{\circ}{p}_N$, $\sqrt{N}$-состоятельность и аналитическое выражение для предельной ковариационной матрицы которой можно вывести из результатов работ [5, 12]. Из этого примера будет следовать, что в рассматриваемой постановке задачи существует, в принципе, целый класс $\mathbb{R}$ $\sqrt{N}$-состоятельных оценок. Следующим важным теоретическим вопросом является вопрос о достижимой нижней границе риска для оценок из класса $\mathbb{R}$, аналогичной границе (2.6). Ясно, что в данной постановке должно существовать другое выражение для правой части (2.6), отличное от функции риска $r(\boldsymbol{\Psi}_p(\mathbf{p}, \boldsymbol{\theta}))$. Действительно, вывод неравенства (2.6) был осуществлен в предположении, что число мешаюших параметров задачи конечно и существует предел нормированной информационной матрицы Фишера для всей совокупности параметров задачи (и, кроме того, выполняются определенные условия регулярности). Сформулированные выше теоретические вопросы исследованы и в значительной степени разрешены в работах [5, 12], где они рассматривались для общей задачи параметрической идентификации многомерной линейной системы по наблюдениям ее входных и выходных сигналов, искаженных шумами. Рассматриваемая здесь задача оценки направления прихода плоской волны по данным группы станций является частным случаем указанной задачи.

Вывод оценки максимума правдоподобия для вектора кажущейся медленности $\mathbf{p}$ и фазовой скорости v сейсмической волны при полностью неизвестной временно́й функции этой волны будем осуществлять на основе приближенного выражения для функции правдоподобия наблюдений в частотной области, аналогичного выражению (2.13). Поскольку s_j – неизвестные (но детерминированные) величины, а $\boldsymbol{\xi}_j$ – гауссовские векторы с моментами (2.11), то можно записать следующее выражение для функции правдоподобия:

$$\begin{aligned}
&L(\mathbf{X}_N|\mathbf{p}, v, \{s_j\}) = C - (1/2)\ln\det\mathbf{F} - \\
&-(1/2)\sum_{j=1}^{N}(\mathbf{x}_j - \mathbf{h}_j(\mathbf{p}, v)s_j)^*\mathbf{F}_j^{-1}(\mathbf{x}_j - \mathbf{h}_j(\mathbf{p}, v)s_j) + O(\sqrt{N}) = \\
&= l_f(\mathbf{X}_N|\mathbf{p}, v, \{s_j\}) + O(\sqrt{N}).
\end{aligned} \tag{4.1}$$

Оценкой максимума правдоподобия информационных параметров $\mathbf{p}, v$ (МП-оценкой) при полностью неизвестных значениях мешающих параметров s_j, $j \in \overline{1, N}$ будем называть оценку $(\overset{\circ}{\mathbf{p}}(\mathbf{X}_N), \overset{\circ}{v}(\mathbf{X}_N))$, являющуюся решением системы уравнений

$$\begin{aligned}
&\frac{\partial}{\partial \mathrm{Re}\, s_j} l_f(\mathbf{X}_N|\mathbf{p}, v, \{s_j\}) = 0; \quad \frac{\partial}{\partial \mathrm{Im}\, s_j} l_f(\mathbf{X}_N|\mathbf{p}, v, \{s_j\}) = 0; \\
&\frac{\partial}{\partial p_\alpha} l_f(\mathbf{X}_N|\mathbf{p}, v, \{s_j\}) = 0; \quad \frac{\partial}{\partial v} l_f(\mathbf{X}_N|\mathbf{p}, v, \{s_j\}) = 0, \quad j \in \overline{1, N}, \quad \alpha \in x, y,
\end{aligned} \tag{4.2}$$

где $\mathrm{Re}\, s_j$ и $\mathrm{Im}\, s_j$ – действительная и мнимая части спектральных отсчетов сигнала s_j. Представив положительно определенную эрмитову матрицу в виде: $\mathbf{F}_j^{-1} = \mathbf{F}_j^{-1/2}\mathbf{F}_j^{-1/2}$, запишем главный член (4.1) в форме

$$l_f(\mathbf{X}_N, \mathbf{p}, v) = C - (1/2)\sum_{j=1}^{N}\ln\det\mathbf{F}_j - (1/2)\sum_{j=1}^{N}|\mathbf{n}_j - \mathbf{d}_j(\mathbf{p}, v)s_j|^2, \tag{4.3}$$

где

$$\mathbf{n}_j = \mathbf{F}_j^{-1/2}\mathbf{x}_j; \quad \mathbf{d}_j(\mathbf{p}, v) = \mathbf{F}_j^{-1/2}\mathbf{h}_j(\mathbf{p}, v). \tag{4.4}$$

Первая подсистема уравнений (4.2) для точки максимума функционала (4.3) является линейной и, как легко проверить, имеет следующее аналитическое решение:

$$\overset{+}{s}_j(\mathbf{p}, v) = \frac{\mathbf{d}_j^*(\mathbf{p}, v)\mathbf{n}_j}{|\mathbf{d}_j(\mathbf{p}, v)|^2}, \quad j \in \overline{1, N}.$$

Подставляя $\overset{+}{s}_j(\mathbf{p}, v)$ во вторую подсистему из (4.2), получим, что МП-оценка вектора кажущейся медленности $\mathbf{p}$ и фазовой скорости волны v в рассматриваемой постановке задачи есть решение системы нелинейных уравнений

$$\begin{aligned} \rho_\alpha(\mathbf{X}_N, \mathbf{p}, v) &= \frac{\partial}{\partial p_\alpha}\left(\sum_{j=1}^{N} |\mathbf{\Pi}_j(\mathbf{p}, v)\mathbf{n}_j|^2\right) = 0, \quad \alpha \in \overline{x, y}, \\ \rho_v(\mathbf{X}_N, \mathbf{p}, v) &= \frac{\partial}{\partial v}\left(\sum_{j=1}^{N} |\mathbf{\Pi}_j(\mathbf{p}, v)\mathbf{n}_j|^2\right) = 0, \end{aligned} \tag{4.5}$$

где $\mathbf{\Pi}_j(\mathbf{p}, v) = \mathbf{I} - \mathbf{d}_j(\mathbf{p}, v)|\mathbf{d}_j(\mathbf{p}, v)|^{-2}\mathbf{d}_j^*(\mathbf{p}, v)$.

Подставляя в (4.5) выражения (4.4) и учитывая легко проверяемое соотношение: $\mathbf{\Pi}_j\mathbf{\Pi}_j^* = \mathbf{\Pi}_j$, получаем окончательные уравнения для МП-оценки вектора кажущейся медленности и фазовой скорости волны:

$$\begin{aligned} \rho_\alpha(\mathbf{X}_N, \mathbf{p}, v) &= \sum_{j=1}^{N} \mathbf{x}_j \dot{\mathbf{A}}_{\alpha j}(\mathbf{p}, v)\mathbf{x}_j = 0, \quad \alpha \in \overline{x, y}, \\ \rho_v(\mathbf{X}_N, \mathbf{p}, v) &= \sum_{j=1}^{N} \mathbf{x}_j \dot{\mathbf{A}}_{vj}(\mathbf{p}, v)\mathbf{x}_j = 0, \end{aligned} \tag{4.6}$$

где

$$\dot{\mathbf{A}}_{\alpha j}(\mathbf{p}, v) = \frac{\partial}{\partial p_\alpha}\mathbf{A}(\mathbf{p}, v); \quad \dot{\mathbf{A}}_{vj}(\mathbf{p}, v) = \frac{\partial}{\partial v}\mathbf{A}(\mathbf{p}, v),$$

$$\mathbf{A}_j(\mathbf{p}, v) = \left(\mathbf{I} - \frac{\mathbf{F}_j^{-1}\mathbf{H}_j(\mathbf{p}, v)}{\mathrm{tr}\mathbf{F}_j^{-1}\mathbf{H}_j(\mathbf{p}, v)}\right)\mathbf{F}_j^{-1}.$$

Производя дифференцирование по координатам p_x, p_y вектора кажущейся медленности, получаем следующие расчетные формулы:

$$\dot{\mathbf{A}}_{\alpha j}(\mathbf{p}, v) = \frac{\mathbf{F}_j^{-1}}{\mathrm{tr}\mathbf{F}_j^{-1}\mathbf{H}_j(\mathbf{p}, v)}\left(\dot{\mathbf{H}}_{\alpha j}(\mathbf{p}, v) - \mathbf{H}_j(\mathbf{p}, v)\frac{\mathrm{tr}\mathbf{F}_j^{-1}\dot{\mathbf{H}}_{\alpha j}(\mathbf{p}, v)}{\mathrm{tr}\mathbf{F}_j^{-1}\mathbf{H}_j(\mathbf{p}, v)}\right),$$

где $\dot{\mathbf{H}}_{\alpha j}(\mathbf{p}, v) = \partial\mathbf{H}_j(\mathbf{p}, v)/\partial p_\alpha$.

Из сказанного выше следует важный вывод: МП-оценки вектора кажущейся медленности $\mathbf{p}$ и фазовой скорости v сейсмичской волны при полностью неизвестной временно́й форме этой волны минимизируют функционал

$$\Omega(\mathbf{X}_N, \mathbf{p}, v) = \sum_{j=1}^{N} \mathbf{x}_j^* \mathbf{A}(\mathbf{p}, v)\mathbf{x}. \tag{4.7}$$

Функционалу (4.7) можно придать наглядный геометрический смысл. Согласно (2.9), отсчеты сигнала $\mathbf{u}_j$ принадлежат (с точностью до $\mathbf{O}_j(1/\sqrt{N})$ одномерным подпространствам комплексного М-мерного евклидова пространства $\mathbb{C}^M$ значений $\mathbf{x}_j$. Эти подпространства (при каждом j – разные) определяются соотношением

$$\mathfrak{L}_j(\mathbf{p}, v) = \{\mathbf{h}_j(\mathbf{p}, v)n, \quad n \in \mathbb{C}^1\}.$$

Нетрудно показать, что функционал $\Omega(\mathbf{X}_j, \mathbf{p}, v)$ есть сумма квадратов расстояний наблюдений $\mathbf{x}_j$ до соответствующих подпространств $\mathfrak{L}_j(\mathbf{p}, v)$ в метрике пространства $\mathbb{C}^M$, определяемой (при каждом j) скалярными произведениями

$$(\mathbf{a}, \mathbf{b}) = \mathbf{a}^* \mathbf{F}_j^{-1} \mathbf{b}, \quad \mathbf{a}, \mathbf{b} \in \mathbb{C}^M.$$

Таким образом, метод максимума правдоподобия в рассматриваемой постановке задачи приводит к оценке вектора $(\mathbf{p}, v)$, являющейся обобшением известной статистической оценки ортогональной регрессии.

Элементарные преобразования функционала (4.7) показывают, что он приводит к тем же оценкам кажущейся медленности и фазовой скорости, что и функционал (2.20) для асимптотически оптимальной оценки при случайном гауссовском сигнале и большом отношении сигнал/шум. Действительно, из (4.6) и (4.7) следует, что

$$\Omega(\mathbf{X}_N, \mathbf{p}, v) = \sum_{j=1}^{N} \mathbf{x}_j^* \mathbf{F}_j^{-1} \mathbf{x}_j - \sum_{j=1}^{N} \frac{|\mathbf{h}_j^*(\mathbf{p}, v)\mathbf{F}_j^{-1}\mathbf{x}_j|^2}{\mathbf{h}_j^*(\mathbf{p}, v)\mathbf{F}_j^{-1}\mathbf{h}_j(\mathbf{p}, v)}. \tag{4.8}$$

Поскольку первое слагаемое в (4.8) не зависит от $\mathbf{p}, v$, минимизация функционала $\Omega(\mathbf{X}_N, \mathbf{p}, v)$ по $\mathbf{p}, v$ эквивалентна максимизации второго слагаемого, которое в точности совпадает с функционалом $\tilde{\mathbf{\Lambda}}(\mathbf{X}_N, \mathbf{p}, v)$ (2.20). Это важное совпадение есть проявление тесной взаимосвязи столь различных, на первый взгляд, моделей наблюдений, как модель гауссовского сигнала с неизвестным энергетическим спектром и модель сигнала с полностью неизвестой временно́й формой. Эта взаимосвязь обсуждена в работе [12].

ЛИТЕРАТУРА

1. *Mykkeltveit S., Ringdal F., Kvarna T., Alewine R.* Application of regional arrays in seismic verification research // Bull. Seism. Soc. Amer. 1990. Vol.80. P.1777-1800.
2. *Suteau-Henson A.* Estimation of azimuth and slowness from three-component and array stations // Bull. Seism. Soc. Amer. 1990. Vol.80. P.1987-1998.
3. *Harris D.B.* Comparison of the direction estimation performance of high-frequency seismic arrays and three-component stations // Bull. Seism. Soc. Amer. 1990. Vol.80. P.1951-1968.
4. *Кушнир А.Ф., Мостовой С.В.* Статистическая обработка геофизических полей. Киев: Наукова думка, 1990. 293 с.

5. *Кушнир А.Ф.* Параметрические методы статистического анализа геофизических временных рядов: Дис. ... докт. физ.-мат. наук. М.: Ин-т физики Земли РАН, 1989.
6. *Ибрагимов И.А. Хасьминский Р.З.* Асимптотическая теория оценивания. М.: Наука, 1979. 527 с.
7. *Гашин А.Н., Кушнир А.Ф., Яковлев А.П.* Выделение волновых форм сейсмических фаз по данным сейсмодеформометрической микрогруппы // Физика Земли. 1997. N2. С.30-51.
8. *Wittle P.* Some results in time series analysis // Skand. Aktuar. 1952. Vol.35. P.48-50.
9. *Бриллинджер Д.* Временные ряды. Обработка данных и теория. М.: Мир, 1980. 536 с.
10. *Джапаридзе К.О.* Оценка параметров и проверка гипотез в спектральном анализе стационарных временных рядов. Тбилиси: Тбилисский государственный университет, 1981. 261 с.
11. *Воеводин В.В., Кузнецов Ю.А.* Матрицы и вычисления. М.: Наука, 1984. 318 с.
12. *Кушнир А.Ф.* Идентификация линейных динамических систем как задача с мешающими параметрами // Современные методы интерпретации сейсмологических данных. М.: Наука, 1991. С.252-272. (Вычисл. сейсмология; Вып. 24).
13. *Кендал М., Стюарт А.* Статистические выводы и связи. М.: Наука, 1973. 899 с.
14. *Nusbaum M.* An asymptotic minimax resk for estimation of a linear functional relationship // J.Multivar. Anal. 1984. Vol.14, N3. P.300-314.