

I. ПРОГНОЗ ЗЕМЛЕТРЯСЕНИЙ И МОДЕЛИ СЕЙСМИЧЕСКОГО ПРОЦЕССА

УДК 550.341

Г.М.Молчан

ОПТИМАЛЬНЫЕ СТРАТЕГИИ В ПРОГНОЗЕ ЗЕМЛЕТРЯСЕНИЙ

G.M.Molchan

OPTIMAL STRATEGIES IN EARTHQUAKE PREDICTION

Prediction of stationary sequence of rare events (large earthquakes) is considered. The prediction uses any information that is related in a stationary way to the sequence. The goal of prediction is stated as minimization of a loss function which is a function of two long-run errors: the rate of failures-to-predict (in percent to total number of strong earthquakes) and the total time of alarms (in percent to time-period considered). We describe the optimal prediction strategies for any convex loss function. Stability of prediction is analyzed within the framework of this loss-function approach. The results provide refinement of some earthquake prediction algorithms.

Введение

В [1] акцентировано внимание на целевом подходе к прогнозу землетрясений. Он важен как на исследовательской стадии, так и на стадии принятия решений. Без этого невозможны точная постановка задачи прогноза, сравнение методов прогноза и их оптимальное использование. Цель прогноза формулировалась как задача минимизации некоторой функции потерь $v = f(\hat{n}, \hat{t})$, зависящей от двух долговременных ошибок: доли непредсказанных событий \hat{n} и доли времени тревоги \hat{t} . Были описаны v -оптимальные стратегии прогноза в некоторых частных случаях. Ограничения касались выбора функций потерь и особенно текущей информации, используемой для прогноза. В настоящей работе эти ограничения сняты.

Основной результат работы формулируется вполне естественно: для любой выпуклой функции потерь оптимальная стратегия прогноза в точке сводится к объявлению тревоги всякий раз, когда "функция риска" $r(t)$ (см. ниже) превышает некоторый порог. В важном частном случае (γ -линейная функция) порог определяется наклоном изолиний γ . Порог зависит от места, если функция потерь неоднородна по пространству. Функция риска хорошо известна как оптимальный предиктор числа событий в малом пространственно-временном объеме. Оказалось, что она оптимальна и в принципиально ином типе прогноза, связанном с минимизацией долговременных потерь γ .

Полученный результат можно рассматривать как обоснование и уточнение некоторых методов прогноза, основанных на байесовском принципе комбинирования предвестников (см., например, [2]). Такие методы сводятся к картированию апостериорного риска в изолиниях, оставляя в стороне вопрос о выборе уровня изолинии для прогноза либо решая его чисто эмпирически. Разумеется, обоснование не относится к практике расчета апостериорного риска. Обычно такие расчеты основаны на волевых решениях о корреляционных связях элементарных предвестников и поэтому требуют самостоятельного анализа.

В работе дополнительно рассмотрен вопрос устойчивости или robustности прогноза, правда, в очень ограниченной форме. Тем не менее полученный вывод важен для иллюстрации того факта, что целевой подход заметно стабилизирует задачу прогноза. Задача заведомо неустойчива, если, следуя Вир-Джонсу, ее понимать как задачу оценки функции риска в целом [3]. Ниже рассмотрен пример прогноза, основанный только на информации о времени до последнего сильного события. В этом случае $r(t)$ аналогична функции риска из анализа данных типа времени жизни; она определяется распределением времени между сильными событиями и очень чувствительна к поведению хвостовой части этого распределения, обычно мало изученного. Тем не менее ошибки и структура прогноза, минимизирующего, например, $\gamma = \max(\hat{n}, \hat{t})$, удивительно стабильны для целого класса распределений, включающего логнормальное, гамма- и Вейбулла-распределения с одинаковыми первыми двумя моментами. Указанные типы распределений являются типичными конкурирующими моделями в исследованиях по сейсмическому риску [4, 5].

Ниже используются следующие обозначения: E – символ математического ожидания, $P(\cdot | \xi)$ – условная вероятность и $E(\cdot | \xi)$ – условное математическое ожидание относительно случайного объекта ξ .

1. Постановка задачи

Рассмотрим поток сильных событий $\{t_i\}$ в точке (области) g с конечной интенсивностью $\lambda > 0$; его определяет считающая мера $dN(t)$, $(N(t))$ – число событий в интервалах $(0, t)$.

Пусть $\xi(t)$ – информация, используемая в момент t для прогноза $\{t_i\}$. Обычно она включает данные о состояниях предвестников и /или геофизических полей в некотором интервале времени $(t-a, t-b)$, доступные наблюдателю к моменту t , а также каталог последних сильных событий за период $(t-c, t)$ и т.п. Функциональная природа $\xi(t)$ для дальнейшего не очень существенна. Чтобы охватить реальную ситуацию достаточно считать $\xi(t)$ случайным процессом со значениями в некотором (сепарабельном) метрическом пространстве H .

В прогнозе $\{t_i\}$ информация $\xi(t)$ используется для выработки решения (стратегии), $\pi(t)$, объявлять ($\pi = 1$) или не объявлять ($\pi = 0$) тревогу в интервале $(t, t+\delta)$, $\delta \ll \lambda^{-1}$. Событие t_i считается предсказанным, если оно накрывается интервалом времени тревоги. Основными показателями качества прогноза событий из интервала $(0, t)$ являются число непредсказанных событий $n_\pi(t)$ и суммарная продолжительность тревоги $\tau_\pi(t)$. Очевидно они связаны с π и dN стохастическими соотношениями:

$$dn_\pi(t) = (1-\pi(t))dN(t), \quad n_\pi(0) = 0, \quad (1)$$

$$d\tau_\pi(t) = \pi(t)dt, \quad \tau_\pi(0) = 0, \quad (2)$$

где $df = f(t+\delta) - f(t)$ – "прогрессивный" дифференциал, $\delta > 0$. Для строгого анализа (1), (2) необходимо накладывать на π определенные условия регулярности. Класс $\{\pi\}$ не должен быть узким, поскольку наша цель – описать оптимальные стратегии π . Чтобы не затруднить изложение чисто формальными трудностями, удобно перейти к дискретному времени с шагом δ , т.е. t пробегает решетку $\delta Z^1 = \{\delta k, k = 0, \pm 1, \pm 2, \dots\}$.

Дискретизируя время, отождествим события t_i из каждой ячейки, что физически естественно. Иначе говоря, будем считать, что

$$P\{N(t+\delta) - N(t) > 1\} = 0, \quad t \in \delta Z^1. \quad (3)$$

Для дальнейшего важны следующие допущения о структуре введенных процессов:

(A) $\{N(t) - N(t-\delta), \xi(t), \pi(t)\}$ – стационарный эргодический векторный процесс на вероятностном пространстве $(\Omega, \mathcal{U}, \mathcal{P}(d\omega))$;

(B) стратегия π принадлежит классу Π , т.е.,

$$\pi(t) = \begin{cases} 1 & \text{с вероятностью } p_t(\omega), \\ 0 & 1-p_t(\omega), \end{cases} \quad (4)$$

где $p_t(\omega)$ измеримо относительно $\xi(t)$ (если Π конечномерно, то $p_t(\omega) = \varphi(\xi(t))$, где φ – измеримая функция).

Класс Π достаточно широк и включает обычные нерандомизированные стратегии, у которых $p_t(\omega)$ принимает два значения: 1 и 0.

Величину вида $r(t) = P\{N(t+\delta)-N(t) > 0 | \xi(t)\}/\delta$ назовем функцией риска. (Если время непрерывно, под $r(t)$ надо понимать предел правой части по вероятности при $\delta \downarrow 0$, если он существует.)

Утверждение 1. В условиях (A), (B) с вероятностью 1 существуют следующие пределы при $t \rightarrow \infty$:

$$n_{\pi}(t)/N(t) \rightarrow \bar{n}_{\pi} = E(1-\pi(s)) r(s)/\lambda, \quad (5)$$

$$\tau_{\pi}(t)/t \rightarrow \bar{\tau}_{\pi} = E\pi(s), \quad (6)$$

где s – любое.

Доказательство. Утверждение 1 является следствием индивидуальной эргодической теоремы [6]. Действительно, пусть $t = n\delta$. Тогда

$$N(t)/t = n^{-1} \sum_{i=1}^n [N(t_i + \delta) - N(t_i)]/\delta \rightarrow E[N(s+\delta) - N(s)]/\delta = \lambda,$$

$$\tau_{\pi}(t)/t = n^{-1} \sum_{i=1}^n \pi(t_i) \rightarrow E\pi(s),$$

$$n_{\pi}(t)/t = (\delta n)^{-1} \sum_{i=1}^n (1 - \pi(t)) [N(t_i + \delta) - N(t_i)] \rightarrow$$

$$\rightarrow E(1 - \pi(s)) [N(s + \delta) - N(s)]/\delta = I.$$

Последний предел можно переписать, используя измеримость $\pi(t)$ относительно $\xi(t)$: $I = E(1 - \pi(t)) E(N(t+\delta) - N(t) | \xi(t)) \delta^{-1}$.

В силу (3) имеем

$$I = E(1 - \pi(t)) P(N(t+\delta) - N(t) > 0 | \xi(t)) \delta^{-1} = E(1 - \pi(t)) r(t).$$

Полученные соотношения доказывают (5), (6).

Пределы (5), (6) определяют две долговременные ошибки прогноза: долю непредсказанных событий – \bar{n} и долю времени тревоги – $\bar{\tau}$. Заметим, что всегда существуют стратегии с любым наперед заданным уровнем одной из ошибок \bar{n} или $\bar{\tau}$. Например, случайное объявление тревоги с частотой p приводит к ошибкам $\bar{n} = 1-p$ и $\bar{\tau}=p$. На плоскости $(\bar{n}, \bar{\tau})$ этим тривиальным стратегиям отвечают точки отрезка $\bar{n} + \bar{\tau} = 1$, $(\bar{n}, \bar{\tau}) > 0$; нетривиальным стратегиям отвечают остальные точки $(\bar{n}, \bar{\tau})$ треугольника с вершинами: $(0,0)$ – идеальная, но

пока недостижимая стратегия, $(0, 1)$ – стратегия "пессимиста", объявляющего постоянно тревогу, и $(1, 0)$ – стратегия "оптимиста", никогда не объявляющего тревогу. Очевидно, что только идеальная стратегия не хуже стратегий пессимиста и оптимиста сразу по двум ошибкам одновременно. Поэтому разные стратегии сравнимы лишь на базе одномерной характеристики. Если исходить из приоритета долговременных ошибок (\hat{n}, \hat{t}) , то основой для сравнения стратегий должна стать некоторая функция $v = f(\hat{n}, \hat{t})$ этих ошибок. Если v возрастает по каждому аргументу назовем ее *функцией потерь*.

Цель прогноза естественно ассоциировать с минимизацией потерь v . Стратегию π^* , на которой достигается минимум потерь: $v_{\pi^*} \leq v_{\pi}$, $\pi \in \Pi$, назовем *v*-оптимальной в классе Π .

В приложениях и на разных этапах исследований цели прогноза могут быть разными. Поэтому ниже от функций потерь требуется только условие

(С) $v = f(\hat{n}, \hat{t})$ – непрерывная, выпуклая функция, строго возрастающая по каждому аргументу.

2. Оптимальные стратегии

Утверждение 2. В условиях А, В, С *v*-оптимальная стратегия сводится к объявлению тревоги с вероятностями

$$p_t(\omega) = \begin{cases} 1, & r(t) > r_v, \\ p_v, & \text{если } r(t) = r_v, \\ 0, & r(t) < r_v. \end{cases} \quad (7)$$

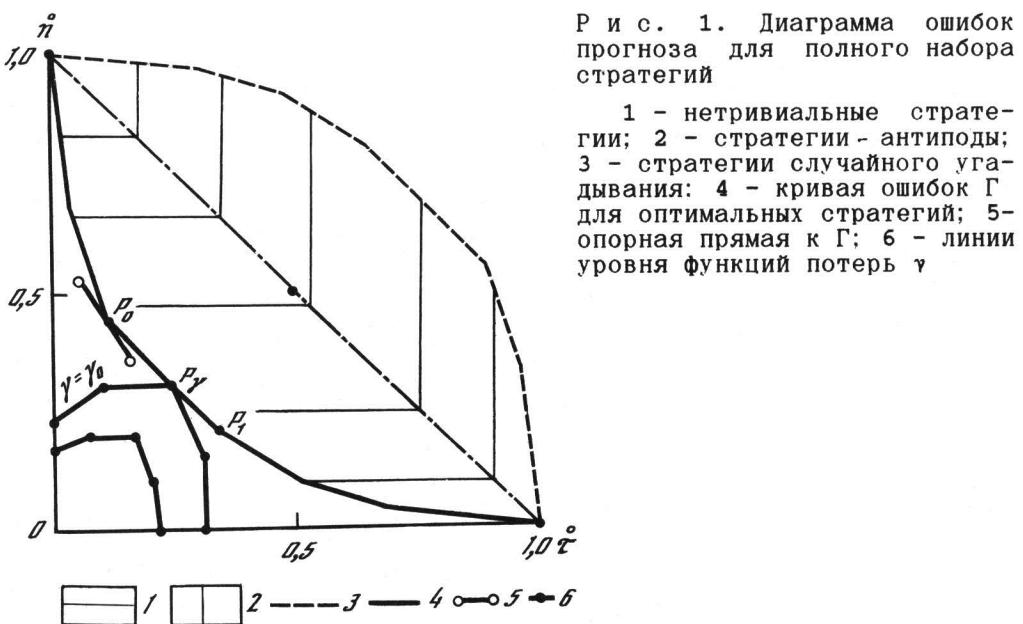
Для линейных потерь вида $v_1 = \alpha \lambda \hat{n} + \beta \hat{t}$ порог $r_v = \beta/\alpha$, т.е. не зависит от информации $\xi(t)$, а p_v – любое. В частности $p_v = 0$.

Иначе говоря, оптимальная стратегия в основном нерандомизирована. Рандомизация заведомо не нужна, если события $\{\omega : r(t, \omega) = c\}$ имеют нулевую вероятность. Случай, когда это не так, не столь экзотичны. Пусть $\xi(t)$ – время от последнего сильного события до момента t . Тогда $r(t) = F'(x)/(1 - F(x))$, если $\xi(t) = x$ и F – распределение времени между сильными событиями. Отсюда $r(t) = \text{const}$ для интервалов x , где $1 - F(x) = c \exp(-\beta x)$, $x \in \Delta$, т.е. $P\{r(t) = c\} = \int_{\Delta} dF(x) > 0$.

Функция v_1 определяет ожидаемые потери в единицу времени, связанные с прогнозом, если α – цена за пропуск события, а β – цена за единицу времени тревоги.

Доказательство. Пусть $\mathcal{P}_t(d\omega)$ – сужение вероятностной меры $\mathcal{P}(d\omega)$ на события. порожденные $\xi(t)$; удобно писать

Рис. 1. Диаграмма ошибок прогноза для полного набора стратегий



$\mathcal{P}_t(d\omega) = P\{\xi_t \in d\omega\}$. В силу (5), (6) и измеримости $\pi(t)$, $r(t)$ относительно ξ_t имеем

$$\dot{n}_\pi = \int_{\Omega} (1 - \pi(t, \omega)) r(t, \omega) \lambda^{-1} \mathcal{P}_t(d\omega) = \int_{\Omega} (1 - \pi(t, \omega)) \mathcal{P}_t^1(d\omega), \quad (8)$$

$$\dot{\tau}_\pi = \int_{\Omega} \pi(t, \omega) \mathcal{P}_t(d\omega), \quad (9)$$

где

$$\begin{aligned} \mathcal{P}_t^1(d\omega) &= r(t, \omega) \lambda^{-1} \mathcal{P}_t(d\omega) = \lambda^{-1} \mathcal{P}(N(\delta t) = 1 | \xi_t) \mathcal{P}_t(d\omega) = \\ &= \mathcal{P}\{\xi_t \in d\omega, N(\delta t) = 1\} / \mathcal{P}(N(\delta t) = 1) = \mathcal{P}\{\xi_t \in d\omega | N(\delta t) = 1\}. \end{aligned}$$

$$a \quad N(\delta t) = N(t+\delta) - N(t).$$

Если $\mathcal{P}_t(d\omega)$ – вероятностная мера на реализациях ξ_t , то $\mathcal{P}_t^1(d\omega)$ – условная вероятностная мера на таких реализациях ξ_t , которые предваряют сильные события в интервале $(t, t+\delta)$. Поэтому соотношения (8), (9) допускают очевидную интерпретацию. Имеется наблюдение ξ_t и две гипотезы о его распределении: $\mathcal{P}_t(d\omega)$ – гипотеза H_0 и $\mathcal{P}_t^1(d\omega)$ – гипотеза H_1 . Решающее правило π относит ξ_t к H_1 с вероятностью $p_t(\omega)$, при этом $(\dot{n}, \dot{\tau})$ определяют ошибки двух родов для такого решения. Если фиксировать ошибку одного рода, например $\dot{\tau}$, и минимизировать другую \dot{n} , то придем к задаче Неймана-Пирсона [7]. Ее решение имеет вид (4) с

$$p_t(\omega) = \begin{cases} 1, & L_t > c, \\ p, & \text{если} \\ 0, & L_t = c, \\ & L_t < c, \end{cases}$$

где $L_t = \mathcal{P}_t(d\omega)/\mathcal{P}_t(d\omega) = r(t) \lambda^{-1}$ – отношение правдоподобия, c и p – константы. Ниже показано, что в условиях (С) правило, минимизирующее потери $\gamma = f(\hat{n}, \hat{t})$, имеет тот же вид. Постоянные (c, p) зависят от γ и, вообще говоря, от t . Однако в условиях (А) меры $\mathcal{P}_t(d\omega)$ и $\mathcal{P}_t^1(d\omega)$ инвариантны относительно сдвига по времени. Поэтому, если параметры (c, p) оптимальны для γ при $t = s$, то они же будут оптимальными для любого t . Конкретизацию (c, p) для линейных потерь см. ниже.

Кривая ошибок $\Gamma = (\hat{n}, \hat{t})$. Ошибки $(\hat{n}, \hat{t})_\pi$, отвечающие всевозможным стратегиям π прогноза по данным ξ_t образуют множество G (см.рис.1). Оно имеет центр симметрии $(1/2, 1/2)$: на всякий прогноз есть его антипод, меняющий местами тревогу и нетревогу. Поэтому их ошибки симметричны относительно указанного центра. Множество G выпукло: если правила π_1 и π_2 применять случайно с частотами p и $(1-p)$, то новое правило приведет к усреднению ошибок с теми же весами. Установленная связь с проверкой гипотез и лемма Неймана-Пирсона [7] показывают, что ошибки для стратегий (4), (7) отвечают нижней границе Γ множества G . Несмотря на два параметра r_γ, p_γ , кривая Γ однопараметрическая. Кривая Γ – выпуклая (следствие выпуклости G); за вычетом линейных участков она отвечает чистым, нерандомизированным стратегиям, у которых $p_\gamma = 0$. На каждом линейном участке Γ параметром служит p_γ . Он определяет веса $(1-p_\gamma, p_\gamma)$, с которыми смешиваются стратегии, отвечающие концевым точкам линейного участка (см. рис.1) В свою очередь, любой линейный участок Γ связан с событиями $\{r(t) = c\}$ ненулевой вероятности. Это, очевидно, следует из (c, p) – параметрического представления Γ :

$$\hat{n} = \int_{r(t) < c} \mathcal{P}(d\omega) + (1-p) \int_{r(t) = c} \mathcal{P}(d\omega), \quad (10)$$

$$\hat{t} = \lambda^{-1} \int_{r(t) > c} r(t) \mathcal{P}(d\omega) + p \lambda^{-1} c \int_{r(t) = c} \mathcal{P}(d\omega). \quad (11)$$

Кривая ошибок Γ и целевая функция γ определяют γ -оптимальную стратегию. Действительно, из выпуклости изолиний γ и кривой Γ следует, что существует точка касания $P_\gamma = (\hat{n}, \hat{t})_\gamma$, изолинии $\gamma = \gamma_0$ с Γ (наименьший уровень γ_0 -изолинии, пересекающейся с Γ).

Тогда, учитывая (10, 11), порог $r_y = -\lambda \frac{dn}{d\tau}(P_y)$. Выпуклая кривая Г всегда дифференцируема за исключением возможно счетного числа точек. Если P_y принадлежит к последним, то $-\frac{dn}{d\tau}$ надо понимать как любую субпроизводную в этой точке (наклон любой опорной для Г прямой в точке P_y (см.рис.1)). Если точка P_y принадлежит линейному участку Г с концевыми точками $(\hat{n}, \hat{\tau})_i$, $i = 0, 1$, то, очевидно, $p_y = (\tau_y - \tau_0)/(\tau_1 - \tau_0)$, $\tau_0 < \tau_1$.

Отсюда вытекают все известные способы построения y -оптимальных методов проверки простых гипотез и важные примеры прогноза.

Пример 1. Пусть $y_1 = \alpha \hat{n} + \beta \hat{\tau}$. Тогда в точке касания P_y прямой $y = y_0$ с Г прямая $y = y_0$ является опорной для Г. Значит ее наклон $k = -\beta/\alpha$ можно взять за субпроизводную Г в P_y . Отсюда $r_y = \beta/\alpha$. Если P_y принадлежит линейному участку Г, тогда участок Г целиком принадлежит прямой $y = y_0$ (следствие касания). Поэтому из всех точек касания P_y можно выбрать концевую точку участка, где $p = 0$.

Пример 2. Пусть $y = \max(\alpha \hat{n}, \beta \hat{\tau})$, тогда области $\{\hat{n}, \hat{\tau}: y < c\} = \{\alpha \hat{n} < c, \beta \hat{\tau} < c\}$ образуют прямоугольники с вершинами на прямой $\alpha \hat{n} = \beta \hat{\tau}$. Поэтому точка касания P_y лежит на пересечении прямой $\alpha \hat{n} = \beta \hat{\tau}$ с Г, т.е. параметры (7) определяются уравнением $\alpha \hat{n} = \beta \hat{\tau}$.

Пример 3. В работах по прогнозу часто используется величина $(1-\hat{n})/\hat{\tau}$ либо как мера эффективности прогноза [8], либо как условная мера повышения вероятности сильного события предвестника с прогнозными характеристиками $(\hat{n}, \hat{\tau})$ [9]. Обратная величина $y = \hat{\tau}/(1-\hat{n})$ является функцией потерь. Ее линии уровня образуют пучок прямых $1-\hat{n} = k \hat{\tau}$ с центром в точке $P_0 = (1, 0) \in \Gamma$. Поэтому $P_0 = P_y$, т.е. y -оптимальной стратегией является тривиальная стратегия оптимиста, никогда не объявляющего тревогу. Эта стратегия обладает наивысшей эффективностью при любой используемой информации $\xi(t)$ и определяется величиной $|dn/d\tau(P_0)|$, $(\hat{n}, \hat{\tau}) \in \Gamma$.

Следовательно, величина $(1-\hat{n})/\hat{\tau}$ непригодна для целей сравнения различных методов прогноза. Однако она имеет смысл как мера повышения вероятности сильного события в текущем прогнозе. Действительно, в простейших ситуациях величина $\log((1-\hat{n})/\hat{\tau}) = \log P\{\text{событие}|\text{предвестник}\}/P\{\text{событие}\}$ совпадает с количеством информации Шеннона, которую получает наблюдатель о сильном событии при наблюдении элементарного предвестника.

Таким образом, кривая Г содержит всю информацию о потенциальных возможностях прогноза по данным $\xi(t)$. Она удобна тем, что не за-

висит от весьма проблематичной и разнообразной целевой функции прогноза. Практические методы прогноза тоже удобно представлять кривыми ошибок. Обычно метод зависит от некоторых внутренних параметров θ . Им отвечают ошибки $\{\hat{n}(\theta), \hat{t}(\theta)\}$. Нижняя граница выпуклой оболочки этих ошибок дает представление о силе метода.

3. Минимаксные стратегии

Стратегию, оптимальную относительно потерь $v = \max(\hat{n}, \hat{t})$, назовем минимаксной. Как показано выше, она определяется формулами (4), (7) и уравнением $\hat{n} = \hat{t}$ для ее параметров.

Минимаксная стратегия тесно связана с оптимальными стратегиями для линейных потерь $v_1 = \alpha \lambda \hat{n} + \beta \hat{t}$. Чтобы описать эту связь определим меру эффективности стратегии π при заданных потерях v .

Пусть $v(\pi)$ — потери, связанные с использованием стратегии π ; $\pi_{1/2}^*$ — тривиальная стратегия случайного гадания, когда тревога и нетревога объявляются с вероятностью $1/2$. Отношение потерь $e_v(\pi) = v(\pi_{1/2}^*)/v(\pi)$ назовем эффективностью π относительно $\pi_{1/2}^*$.

Утверждение 3. Пусть π_{mm} — минимаксная стратегия K — класс линейных потерь. Тогда для любого $v_1 \in K$

$$e_{v_1}(\pi_{mm}) = \min_{v_1 \in K} \max_{\pi} e_{v_1}(\pi). \quad (12)$$

Иначе говоря, минимаксная стратегия v_1 — оптимальна в самом неблагоприятном случае (по эффективности) линейных потерь. Это обстоятельство интересно с точки зрения приложений. Дело в том что обычно отношение $\beta/\alpha\lambda$ оценивается плохо. Поэтому может быть оправдано применение наиболее осторожной минимаксной стратегии. Взамен можно ожидать определенную компенсацию в виде устойчивости стратегии и ожидаемых потерь. Некоторые результаты в этом направлении приведены ниже.

Доказательство. В точке $(\hat{n}, \hat{t}) \in \Gamma$, $\hat{n} = \hat{t}$ построим опорную прямую вида $\bar{v}_1 = v_0$, где $\bar{v}_1 = \bar{\alpha} \lambda \hat{n} + \bar{\beta} \hat{t}$. Тогда

$$e_{v_1}(\pi_{mm}) = 1/(2 \hat{n}_*) = \max_{\pi} e_{\bar{v}_1}(\pi).$$

С другой стороны, если (\hat{n}, \hat{t}) — ошибки v_1 -оптимальной стратегии, то

$$\max_{\pi} e_{v_1}(\pi) = 0,5 (\alpha \lambda + \beta) / (\alpha \lambda \hat{n}_* + \beta \hat{t}_*) \geq$$

$$\geq 0,5 (\alpha \lambda + \beta) / (\alpha \lambda \hat{n}_* + \beta \hat{t}_*) = 1/(2 \hat{n}_*) = e_{v_1}(\pi_{mm}).$$

Оба соотношения доказывают (12).

Устойчивость минимаксной стратегии. Рассмотрим специальный случай прогноза, когда текущая информация $\xi(t)$ состоит только из продолжительности затишья τ_- , т.е. времени, прошедшего от последнего сильного события. Прогноз такого рода рассматривался в [5] для тихоокеанского сейсмического пояса.

Пусть $F(x)$ – распределение времени τ между сильными событиями. Тогда функция риска для информации $\xi(t) = \tau_-$ есть

$$r(t) = F'(t)/(1 - F(t)), \quad t \geq 0. \quad (13)$$

если за 0 принять момент последнего события t_i в интервале $(-\infty, t)$. Правая часть равенства хорошо известна как функция риска в анализе данных типа времени жизни [10].

В инженерных исследованиях в качестве модели F обычно используют распределение Вейбулла:

$$\begin{aligned} F(x) &= 1 - \exp(-(\lambda x)^\alpha), \quad \alpha > 0, \\ \lambda &= \Gamma(1 + 1/\alpha)/m. \end{aligned} \quad (14)$$

Оно же является основным при анализе сейсмического риска в инженерной литературе [4], хотя в качестве возможной альтернативы допускается Гамма-распределение:

$$F(x) = c \int_0^x u^{\alpha-1} \exp(-\alpha u/m) du. \quad (15)$$

Однако оно не считается физически естественным. Для прогноза характеристических землетрясений в [5] используется логнормальная модель F :

$$F(x) = \Phi(\alpha^{-1} \ln(x/m) + \alpha/2) \quad (16)$$

где Φ – стандартное гауссовское распределение и $\alpha^2 = \text{Var}(\ln \tau)$.

Всюду выше m – среднее значение τ . Согласно [5] типичные значения коэффициента вариации $\rho = \sigma/m$ (σ^2 – дисперсия τ) принадлежат интервалу $0,25+0,6$. Эти оценки получены по историческим землетрясениям и палеоданным. Поэтому ниже мы будем ориентироваться на указанные значения $\rho < 1$.

Функция (13) очень чувствительна к выбору модели F . Однако в рамках целевого подхода прогноз оказывается значительно устойчивее. Это подтверждают следующие расчеты.

Поведение Вейбулла- и Гамма-распределений качественно похожи: при $\sigma/m < 1$ они унимодальны. Функция риска (13) возрастает; поэтому для любой функции потерь γ -оптимальные стратегии сводятся к объявлению тревоги, когда затишье τ_- превысит порог $k_\gamma m$. При $\sigma/m > 1$ плотности $F'(x)$ убывают. Убывает $r(x)$, поэтому γ -опти-

мальные стратегии должны объявлять тревогу сразу после сильного события и продолжать ее до порогового значения $\tau_- = k_y m$, (если тревога не снимается очередным событием). Из-за монотонности $F'(x)$ модели (14), (15) практически мало интересны при $\sigma/m > 1$.

Плотность логнормального распределения и ее функция риска универсальны при всех $\rho > 0$. Поэтому γ -оптимальная тревога объявляется только в конечном интервале: $k_y < \tau_- < \bar{k}_y$.

Пороги k_y минимаксной стратегии для моделей (14)–(16) даны на рис.2. Для распределений (14), (15) они практически совпадают при $\rho < 1$, различаясь не более чем на 0,01 или 1,5%. С ними практически совпадает нижний порог k_y для логнормального распределения: при $\rho < 0,6$ отклонение $|\Delta k_y| < 0,03$. Верхний порог k_y в том же диапазоне $\rho < 0,6$ велик, $k_y \in (3, 3, \infty)$ и вносит незначительный вклад в ошибки прогноза. Порог \bar{k}_y целиком зависит от проблематичного поведения модели F на ∞ . Поэтому для логнормальной модели естественно рассматривать также слегка загрубленную γ -оптимальную стратегию с одним нижним порогом тревоги k_y . (Величина k_y по-прежнему находится из условий $\hat{n} = \hat{t}$, $\hat{t} = \min$.) Порог \bar{k}_y хорошо согласован с k_y для (14), (15) во всем интервале $\rho < 1$: $|\Delta k_y| < 0,03$.

Для практически интересного интервала $\rho = 0,25+0,6$ пороги k_y , \bar{k}_y вообще очень стабильны и равны $0,75 \pm 0,05$ (для равномерного распределения $k_y = 0,76$). Эти результаты свидетельствуют об устойчивости структуры минимаксной стратегии прогноза.

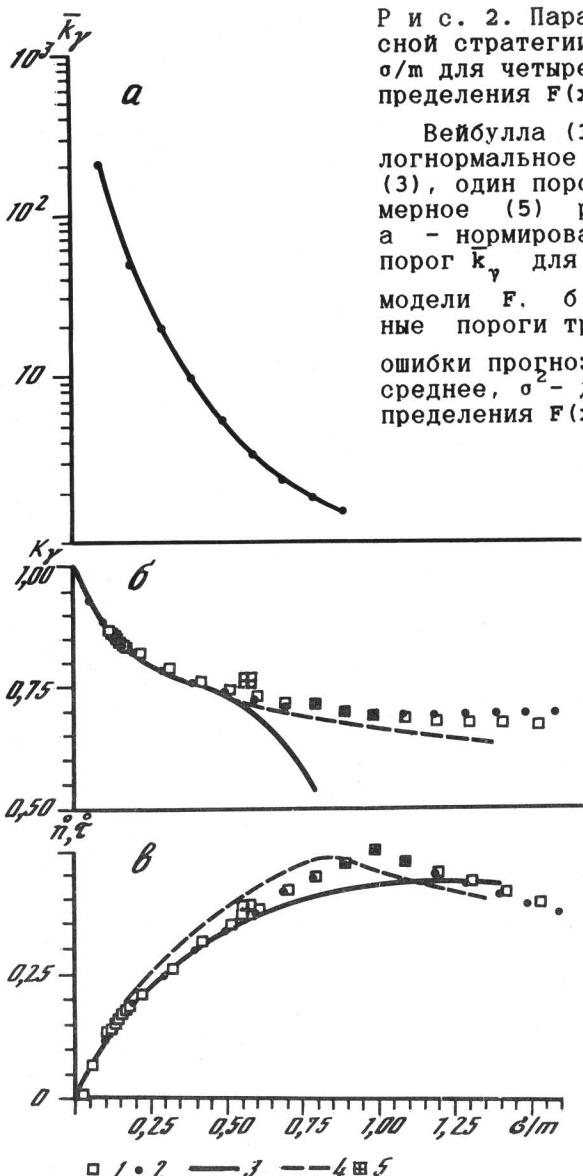
Второй тип устойчивости относится к результатам прогноза, т.е. к значениям потерь $\gamma = \max(\hat{n}, \hat{t})$ (см. рис.2). Значения γ для оптимальных стратегий с одним порогом тревоги, отличаются в пределах 0,02 или 3% при $\rho \leq 0,6$.

На рис.3 представлены кривые ошибок Γ для трех моделей F : (14), (15), (16). Кривые Γ хорошо согласованы для моделей (14), (16) в области $2\hat{n} > \hat{t}$. Это указывает на близость прогноза, его структуры и результатов, для моделей (14), (15) в широком классе функций потерь.

Обсуждение

1. Утверждение 2 дает представление об оптимальной структуре локального по пространству прогноза. Для естественных функций потерь локально-оптимальные прогнозы образуют оптимальный прогноз в целом по пространству.

Имеется принципиальная разница между прогнозными концепциями Колмогорова-Винера и рассмотренной выше. В первом случае прогноз



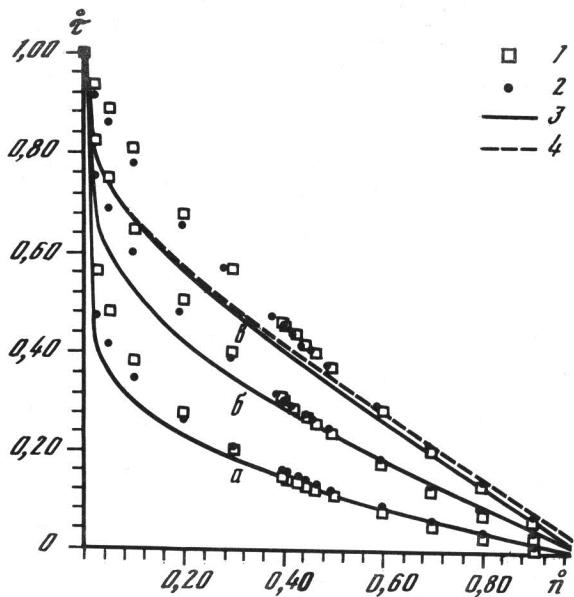
Р и с. 2. Параметры минимаксной стратегии как функции σ/m для четырех моделей распределения $F(x)$

Вейбулла (1), гамма (2), логнормальное (два порога (3), один порог (4)), равномерное (5) распределения; а - нормированный верхний порог k_y для логнормальной модели F . б - нормированные пороги тревог k_y , в - ошибки прогноза $\hat{n} = \bar{t}$; м - среднее, σ^2 - дисперсия распределения $F(x)$

сводится к наилучшей экстраполяции значений временного ряда (например, энергии землетрясений), а во втором - к наилучшей временной локализации особенностей этого ряда (например, его больших пиков). В первом случае можно говорить о прогнозе по вертикали, а во втором - о прогнозе по горизонтали. Функция риска $r(t)$ известна как локально оптимальный предиктор точечного процесса в первом смысле: $r(t)dt$ дает несмешенную оценку числа событий в интервале dt и минимизирует средний квадрат ошибки такого прогноза. Утверж-

Рис. 3. Кривые ошибок для распределений $F(x)$

Вейбулла (1), Гамма (2), логнормальное (два порога) (3), один порог (4) и для значений σ/μ : а - 0,25, б - 0,5, в - 0,75



дение 2 показывает, что $r(t)$ является главным компонентом оптимального прогноза и во втором смысле. Оказалось, что прогноз второго типа своеобразно связан с теорией проверки статистических гипотез, где $r(t)$ выступает как отношение правдоподобия двух простых гипотез. Этим объясняются многие параллели данной работы с указанной теорией.

2. Полученные результаты выходят за рамки чистой теории, поскольку оценка функций риска стала объектом сейсмологической практики. Имеется много попыток ее оценивания [9, 11-13] и даже карттирования (см., например, [2]). Эти попытки лежат в русле идей Утсу-Аки и сводятся к оценке "увеличения вероятности" $PG(\xi)$ сильного события при мониторинге серии предвестников $\xi = (A_1, \dots, A_n)$.

Пусть e_i - индикатор предвестника A_i в интервале $(t-T, t)$, т.е. $e_i = 1$, если предвестник появился, и $e_i = 0$ в противном случае. Тогда величина $PG(\xi)$ есть не что иное, как нормированная величина риска

$$PG(\xi) = P(N(t+\delta) - N(t) > 0 | e_1, \dots, e_n) / P_E = r(t) / \lambda,$$

где $P_E = P(N(t+\delta) - N(t) > 0)$

(заметим, что в определении $PG(\xi)$ по Аки игнорируется факт случайности числа $v = \sum e_i$ появившихся предвестников).

Если $\{e_i\}$ условно независимы относительно случайной величины $N(\delta t) = N(t+\delta) - N(t)$, то $PG(\xi)$ можно вычислить по отдельным величинам $PG(e_i)$:

$$(PG(\xi)^{-1} - P_E)/(1 - P_E) = q \prod_i ((PG_i(1)^{-1} - P_E)/(1 - P_E)), \quad (17)$$

где величина $PG_i(1) = P(N(\delta t) > 0 | s_i = 1)/P_E$ – "увеличение вероятности" события по Аки при наблюдении i -го предвестника. Множитель q связан с ненаблюденными предвестниками:

$$q = \prod_{i : s_i = 0} ((PG_i(0)^{-1} - P_E)/(1 - P_E)), \quad i = m + 1, \dots, n,$$

где $PG_i(0) = P(N(\delta t) > 0 | s_i = 0)/P_E = (1 - PG_i(1)p_i)/(1 - p_i)$ и p_i – вероятность появления предвестника A_i в интервале $(t-T, t)$. Работы [2, 9, 11], базируются на формуле (17), причем в первых двух $q = 1$. Для истинных предвестников $PG(0) < 1$, поэтому величина $q > 1$. Следовательно, полагая $q = 1$, величина риска $r(t)$ завышается; причем по-разному в зависимости от реализованных предвестников. Для условно зависимых предвестников формула (17) также завышает риск, поэтому она пригодна для расчетов лишь с малым числом предвестников.

В работах китайских специалистов (см., например, [13]) сделаны попытки преодолеть эти трудности путем учета корреляций предвестников. К сожалению, этот вопрос решается пока на уровне экспертных оценок.

3. В [5] дан сравнительный анализ сейсмической опасности восьми участков G_i разлома Сан-Андреас. На момент 1988 г. район Паркфильда (P) признан наиболее опасным. Вывод основан на модели (16) и условных вероятностях

$$R = P \{ \text{событие } \in (1988 - 2010 \text{ гг.})/\tau_1 \},$$

где τ_1 – время от момента последнего события до 1988 г.

Это заключение анализировалось в [1] с позиций целевого подхода к прогнозу. Ситуация оказалась не столь простой. В рамках логнормальной модели (14) район P имеет величину $R > 0,9$ для $0 < \tau_1 < 800$ лет. В других районах $R < 0,5$ для любых τ_1 . Иначе говоря, район P оказывается наиболее опасным практически всегда. Вместе с тем согласно минимаксной стратегии, основанной на той же информации τ_1 к 1988 г., 5 из 8 районов должны находиться в состоянии постоянной "тревоги". С точки зрения линейных потерь $v = \alpha \dot{n} + \beta |G| \dot{\tau}_1$, $|G|$ – длина участка, район P имеет наименьшие потери при любых параметрах β, α . В указанных потерях "цена" за пропуск события на участке равна $\alpha \lambda_G^{-1}$, что естественно для задач страхования.

4. Выводы в разд.3 об устойчивости и возможном упрощении прогноза основаны на оценке коэффициента вариации $\sigma/m \leq 0,6$ для определения $F(t)$ интервала между сильными событиями. В недавних рабо-

так Я.Кагана (устное сообщение) эта оценка поставлена под сомнение; утверждается, что $\sigma/m \approx 1$. Один из доводов основан на экстраполяции данных $F'(t) \propto t^{-\alpha}$, полученных для "парных" землетрясений, т.е. для малых времен t , в область больших значений t . Поэтому оценка величины σ/m представляет важную статистическую задачу. В этой связи полезно напомнить результат автора из [1].

Пусть σ/m - накопленные напряжения в некоторой области. Пусть $\sigma(t)$ ограничено снизу (например, в некоторой шкале $\sigma(t) > 0$) и $\sigma(t+\Delta) = \max[\sigma(t) + a\Delta + b(w(t+\Delta) - w(t)), 0]$, $\sigma(0) = x$, $a \geq 0$, где $w(t)$ - стандартное броуновское движение, т.е. $Ew(t) = 0$ и $-Edw(t)dw(s) = \delta(t-s)dt ds$. Сильное событие происходит в момент t , когда $\sigma(t)$ достигает уровня σ_* . В этот момент происходит сброс напряжений до уровня x и система начинает функционировать заново по описанному правилу. В этой наивной модели гэпа распределение времени между событиями $F'(t)$ экспоненциально затухает, причем $\sigma/m \leq (2/3)^{-1/2} \approx 0,816$ при любых параметрах a, b, x и σ_* .

5. Цели в прогнозе априори многообразны и это создает значительные трудности при сравнении отдельных методов прогноза. Имеются попытки преодолеть их на пути "разумных" понятий эффективности прогноза (см. выше пример 3 в п. 2 и обсуждение вопроса в [1]). Однако выход состоит все же в том, чтобы характеризовать метод не одним числом u или точкой (\bar{n}, \bar{t}) , а целой кривой $\Gamma_A = \{\bar{n}_\theta, \bar{t}_\theta\}$, где θ - внутренние параметры метода. Заметим, что любой формализованный предвестник сильного события $M > M_0$ в регионе G содержит по крайней мере два параметра: размер аномалии (порог) и продолжительность тревоги. Обычно параметры конкретизируют путем подгонки, основываясь на чисто интуитивных представлениях о цели прогноза. Кривая Γ_A для пары (M_0, G) , а точнее нижняя граница ее выпуклой оболочки, является универсальной характеристикой прогностической способности метода. Замена Γ_A выпуклой кривой ясна из п.2. Кривые Γ_A допускают сравнение для различных методов и несут информацию о потерях при любых u .

6. Что служит аналогом Γ_A для серии точечных или региональных объектов $\{g\} = G$? Ответ неоднозначен. Для примера рассмотрим суммарные линейные потери:

$$v = \int_G \alpha(g) \lambda(g) \bar{n}(g) dg + \int_G \bar{t}(g) \beta(g) dg = \bar{\alpha} \lambda \bar{n}(G) + \bar{\beta} |G| \bar{t}(G),$$

где $\beta(g)$ - потери от тревоги на единице пространства-времени; $|G|$ - площадь G и $\bar{n}(G)$; $\bar{t}(G)$ - аналог ошибок \bar{n}, \bar{t} . Полагая $\alpha(g) = \alpha$, $\beta(g) = \beta$, приходим к обычным усредненным показателям \bar{n}, \bar{t} :
2.Зак.I304

$$\hat{n}(G) = \int_G \lambda(g) \hat{n}(g) dg / \Lambda \approx$$

\approx (все пропуски цели)/(все сильные события);

$$\hat{\tau}(G) = \int_G \hat{\tau}(g) dg / |G| -$$

доля пространственно-временной тревоги, $\Lambda = \int \lambda(g) dg$.

Модель $\beta(g) = \beta$ приемлема на исследовательской стадии, если сильные события априори допускаются всюду в G . В противном случае величину $\hat{\tau}(G)$ можно сделать сколь угодно малой за счет присоединения асейсмичных участков. Поэтому выбор G для характеристизации метода должен быть хорошо обоснован. Это обстоятельство иногда преодолевается иначе. Положим $\beta(g) = c \hat{\lambda}(g)$, где $\hat{\lambda}(g)$ – хорошо сглаженная плотность эпицентров сильных землетрясений. Тогда для $\hat{\tau}(G)$ получаем альтернативный вариант:

$$\hat{\tau}(G) = \int_G \hat{\tau}(g) \hat{\lambda}(g) dg / \int_G \hat{\lambda}(g) dg ,$$

интересный на исследовательской стадии прогноза при отсутствии соглашения относительно выбора G .

7. Ознакомившись с настоящей работой, Е.Патэ-Корнелл обратила внимание автора на работу [14]. В ней Де Марэ рассматривает задачу прогноза точечного процесса, фиксируя ошибку \hat{n} и минимизируя долю ложных тревог \hat{n}_A . Случай $\hat{n}_A \neq 1$ отвечают краткосрочным прогнозам и соответствуют на кривой Γ вертикальному участку с $\hat{\tau} = 0$. Если такой участок невырожден, $\hat{n} \neq 1$, кривая $\Gamma_0 = (\hat{n}, \hat{n}_A)$ детализирует Γ при $\hat{\tau} = 0$ и является дополнительной характеристикой прогнозных стратегий с редкими или очень короткими тревогами. (Ошибка \hat{n}_A может быть заменена долей тревог на единицу времени.) В [14] осталось незамеченным, что критические множества для объявления тревог допускают в описанной ситуации существенные упрощения и могут быть по-прежнему выражены в терминах функции риска.

Настоящая работа является содержанием доклада, прочитанного автором на ассамблее IASPEI в августе 1989 г. в Стамбуле (Турция). Автор благодарен О.Е. Дмитриевой за проведенные расчеты.

Литература

1. Молчан Г.М. Стратегии в прогнозе сильных землетрясений // М.: Наука, 1990. С.3-27. (Вычисл. сейсмология; Вып.23).
2. Соболев Г.А., Челидзе Т.Л., Завьялов А.Д. и др. Карта ожидаемых землетрясений, основанная на комбинации предвестников (Инструкция). Фонды ИФЗ АН СССР, 1987. 45 с.
3. Vere-Jones D. Earthquake prediction - a statistician's view. // J.Phys. Earth. 1978. Vol.26. P.129-146.

4. Cornel C.A. Applicability of the Poisson earthquake - occurrence model. Sun Francisco. 1986. 66 p. (Prep./EPRI NP-4770; Project P.101-38).
5. Nishenko S.P. Circum - Pacific seismic potential 1989-1999. USGS, Open File Report 89-86. 1989. 126 p.
6. Биллингслий П. Эргодическая теория и информация. М: Мир, 1969. 236 с.
7. Леман Э. Проверка статистических гипотез. М.: Наука, 1964. 498 с.
8. Гусев А.А. Прогноз землетрясений. основанный на статистике сейсмичности // Сейсмичность и сейсмический прогноз. Свойства верхней мантии и их связь с вулканизмом на Камчатке. Новосибирск: Наука, 1974. С.109-119.
9. Aki K. A probabilistic synthesis of precursory phenomena // Maurice Ewing. Ser.4. 1981. P.566-574.
10. Кокс Д., Оукс Д. Анализ данных типа времени жизни // М.: Финансы и статистика, 1988. 190 с.
11. Grandori E., Grandori G. Some probabilistic aspects of earthquake prediction // Meccanica. 1983. Vol. 18. P.246-253.
12. Feng D., Gu J., Lin M. et all. Assessment of earthquake hazard by simultaneous of the statistical method and the method of fuzzy mathematics // PAGEOPH. 1984. Vol.122, N 6. P.982 - 997.
13. Zhuang K.Y., Wang W., Huang B-S. MYCIN inexact inference and its application to earthquake synthesis prediction // The 25-th General Assembly of IASPEI. Abstracts, Istanbul: Turkey. 1989. 697 p.
14. De Mare J. Optimal prediction of catastrophes with applications to Gaussian processes // Ann. Probab. 1980. Vol. 8, N 4. P.841-850.

УДК 550.341

Г.М.Молчан, О.Е.Дмитриева

ИДЕНТИФИКАЦИЯ АФТЕРШОКОВ: ОБЗОР И НОВЫЕ ПОДХОДЫ

G.M.Molchan, O.E.Dmitrieva

IDENTIFICATION OF AFTERSHOCKS: METHODS AND NEW APPROACH

The problem of aftershock identification in earthquake catalogs is studied. Some empirical methods are considered and quantitatively analysed. Game theory approach is applied to formulate the problem which allows for the whole class of optimal methods of aftershock identification. Each method is optimal dependently on the goals and gives the best trade off between the number of missed aftershocks and the number of failurely identified ones. Some illustrations of the new approach to aftershock identification problem are presented.