УДК 519.216.3

©2002 г. О.Н.Граничин, канд. физ.–мат. наук
(Санкт–Петербургский государственный университет)

НЕМИНИМАКСНАЯ ФИЛЬТРАЦИЯ ПРИ НЕИЗВЕСТНЫХ ОГРАНИЧЕННЫХ ПОМЕХАХ В НАБЛЮДЕНИЯХ

В новой постановке рассматривается задача о предсказании значений случайного процесса, порождаемого белым шумом, пропущенным через линейный фильтр. Предполагается, что наблюдения производятся с неизвестными ограниченными помехами, а коэффициенты преобразования сигнала в канале наблюдения являются известными случайными величинами. Предлагается рандомизированный вариант линейного предсказывающего фильтра, дающий при определенных условиях лучшее качество фильтрации по сравнению с минимаксными алгоритмами.

1. Введение

Под фильтрацией понимаются алгоритмы обработки реализаций случайных процессов, направленные на подавление помех, зашумляющих (обычно аддитивно) полезный сигнал. В основе теории оптимальной фильтрации лежит метод Винера–Колмогорова [1] и его рекуррентные модификации, известные под общим названием фильтра Калмана–Бьюси [2].

При разработке алгоритмов фильтрации в большинстве математических исследований помехам в наблюдениях приписываются какие–либо полезные статистические характеристики. Наиболее часто предполагается, например, центрированность и независимость помех. Если это предположение было сделано без достаточного обоснования, то эффективность практического использования такого типа алгоритмов будет невысокой. Так обстоят дела, например, в условиях возможного противодействия "противника", в частности, если помеха определяется детерминированной (неслучайной) неизвестной функцией (противник глушит сигнал).

В ситуации произвольных ограниченных помех возможна минимаксная постановка задачи об оптимальной фильтрации (см., например, [3]), но получающиеся в её рамках решения при высоком уровне помех наблюдения малоинформативны.

Ниже изучается новая постановка задачи о предсказании значений случайного процесса, порождаемого белым шумом, пропущенным через линейный фильтр. Предполагается, что наблюдения производятся с неизвестными ограниченными помехами, а коэффициенты преобразования сигнала в канале наблюдения являются известными случайными независимыми величинами. Смешанная постановка задачи исследовалась и в работе [4] с той разницей, что в ней рассматривалась проблема фильтрации процесса, порождаемого неизвестным ограниченным шумом и наблюдаемого на фоне статистических помех. Данная работа является заключительной в серии из трёх [5–6], посвященных рандомизированным алгоритмам оценивания и оптимизации при произвольных помехах.

В следующем пункте формулируется точная математическая постановка задачи. Далее выводится основной результат, характеризующий свойства предсказывающих оценок предлагаемого нового рандомизированного фильтра. В конце работы качество оценивания рандомизированных и обычных алгоритмов фильтрации сравнивается на основе компьютерного моделирования процессов при различных последовательностях помех.

2. Постановка задачи о предсказании сигнала, наблюдаемого на фоне произвольных ограниченных помех

Ограничимся рассмотрением следующей постановки задачи: наблюдается скалярный сигнал, удовлетворяющий уравнению

(1) $$y_n = \varphi_n^{\mathrm{T}} \theta_n^\star + v_n,\ n = 0, 1, \ldots,$$

т.е. представляющий собой смесь преобразованного векторного сигнала $\{\theta_n^\star\}$, $\theta_n^\star \in \mathbb{R}^r$ и *неизвестной, но ограниченной* детерминированной помехи наблюдения $\{v_n\}$:

(2) $$|v_n| \le C_v < \infty.$$

В уравнении (1) $y_n \in \mathbb{R}^1$ — наблюдения в момент времени n, φ_n — r-мерный вектор, известный в момент времени n. Будем считать, что векторный сигнал $\{\theta_n^\star\}$ порождается устойчивым линейным фильтром

$$\theta_{n+1}^\star = \mathrm{A}\theta_n^\star + w_{n+1},\ \theta_0^\star \in \mathbb{R}^r, \tag{3}$$

в котором A — известная матрица: $\|\mathrm{A}\| < 1$, а $\{w_n\}$ — центрированный независимый случайный процесс: w_{n+1} не зависит от σ-алгебры вероятностных событий $\mathcal{F}_n$, порожденной $\{\varphi_0, \ldots, \varphi_n, w_1, \ldots, w_n\}$, и его дисперсия ограничена некоторой постоянной σ_w^2:

$$\mathrm{E}\{w_n\} = 0,\ \mathrm{E}\{\|w_n\|^2\} \leq \sigma_w^2 < \infty. \tag{4}$$

(Здесь и далее: $\|\cdot\|$ — евклидова норма для векторов в $\mathbb{R}^r$; для матриц

$$\|\mathrm{A}\| = \sqrt{\lambda_{\max}(\mathrm{A}\mathrm{A}^{\mathrm{T}})},$$

$\lambda_{\max}(\cdot)$ и $\lambda_{\min}(\cdot)$ — наибольшее и наименьшее собственные значения.)

Задача фильтрации с прогнозом на один шаг состоит в нахождении оценки $\hat{\theta}_{n+1}$ значения процесса $\{\theta_n^\star\}$ в момент времени $n+1$ по наблюдениям $y_i, \varphi_i,\ i \leq n$. Качество фильтрации определяется средней величиной квадрата невязки

$$\mathrm{E}\{\|\hat{\theta}_{n+1} - \theta_{n+1}^\star\|^2\}.$$

Обычно считают, что в модели наблюдений векторы $\{\varphi_n\}$ определяются детерминированной последовательностью. В этой работе будем предполагать, что

(П) $\{\varphi_n\}$ — последовательность независимых одинаково распределенных случайных векторов с ограниченным известным средним значением: $\|\mathrm{E}\{\varphi_n\}\| = M_\varphi < \infty$. Для любого n центрированные случайные векторы $\Delta_n = \varphi_n - \mathrm{E}\{\varphi_n\}$ не зависят от σ-алгебры $\tilde{\mathcal{F}}_n$, порожденной случайными величинами $\{\varphi_0, \ldots, \varphi_{n-1}, w_1, \ldots, w_n\}$, имеют симметричную функцию распределения $\mathrm{P}(\cdot)$ ($\mathrm{P}(\Omega) = \mathrm{P}(-\Omega)$ для любого борелевского подмножества $\Omega \subset \mathbb{R}^r$) и

$$\mathrm{E}\{\Delta_n \Delta_n^{\mathrm{T}}\} = \mathrm{B} > 0,\ \|\mathrm{B}\| < \infty,\ \mathrm{E}\{\|\Delta_n\|^4\} = M_4 < \infty.$$

(В последней формуле матричное неравенство понимается в смысле квадратичных форм.)

Заметим, что в рассматриваемой постановке задачи отсутствуют считающееся обычными условия о центрированности помех наблюдения $\{v_n\}$ и об их независимости, в том или ином статистическом смысле. В работах [5,6] для оценивания неизменяющихся во времени неизвестных параметров при такого же типа помехах в наблюдениях была обоснована состоятельность оценок алгоритмов типа случайного поиска, достаточно простой структуры. Математическое обоснование состоятельности оценок при почти произвольных помехах в наблюдениях в существенной степени опиралось на знание вероятностных характеристик используемого в алгоритмах пробного возмущения, рандомизирующего правило построения очередной оценки.

3. Рандомизированный фильтр

Исследуем поведение оценок рандомизированного рекуррентного алгоритма

$$\hat{\theta}_{n+1} = \mathrm{A}\hat{\theta}_n - \alpha \mathrm{A}\Gamma\Delta_n(\varphi_n^{\mathrm{T}}\hat{\theta}_n - y_n), \quad \Delta_n = \varphi_n - \mathrm{E}\{\varphi_n\}, \tag{5}$$

где $n = 0, 1, \ldots$, $\alpha > 0$ — размер шага и Γ — положительно определенная симметричная матрица. Будем считать, что начальные данные $\hat{\theta}_0$ заданы произвольным неслучайным вектором из $\mathbb{R}^r$. В алгоритме (5) наблюдаемый в момент времени n, центрированный случайный вектор Δ_n играет ту же роль, что и вектор пробного одновременного возмущения в [6].

Подставив формулы (1) и (3) в алгоритм (5), для ошибки предсказания получаем

$$\hat{\theta}_{n+1}-\theta^\star_{n+1} = \mathrm{A}(\mathrm{I}-\alpha\Gamma\Delta_n\Delta_n^{\mathrm{T}})(\hat{\theta}_n-\theta^\star_n)-\alpha\mathrm{A}\Gamma\Delta_n(\mathrm{E}\{\varphi_n\}^{\mathrm{T}}(\hat{\theta}_n-\theta^\star_n)-v_n)-w_{n+1}.$$

Обозначим через $D_n := \|\hat{\theta}_{n+1}-\theta^\star_{n+1}\|^2$. При сделанных выше предположениях, в силу независимости случайных векторов Δ_n и w_{n+1}, усредняя условно по отношению к σ-алгебре $\mathcal{F}_n$, выводим, что

$$\mathrm{E}\{D_n|\mathcal{F}_n\} = \|\mathrm{A}(\mathrm{I}-\alpha\Gamma\Delta_n\Delta_n^{\mathrm{T}})(\hat{\theta}_n-\theta^\star_n)-\alpha\mathrm{A}\Gamma\Delta_n(\mathrm{E}\{\varphi_n\}^{\mathrm{T}}(\hat{\theta}_n-\theta^\star_n)-v_n)\|^2+\sigma_w^2.$$

Из–за симметричности распределения $\mathrm{P}(\cdot)$ имеем

$$(\hat{\theta}_n - \theta_n^\star)^{\mathrm{T}} \mathrm{E}\{(\mathrm{I} - \alpha \Delta_n \Delta_n^{\mathrm{T}} \Gamma) \mathrm{A}^{\mathrm{T}} \alpha \mathrm{A} \Gamma \Delta_n | \tilde{\mathcal{F}}_n\} (\mathrm{E}\{\varphi_n\}^{\mathrm{T}} (\hat{\theta}_n - \theta_n^\star) - v_n) = 0.$$

Следовательно, усреднив по отношению к σ-алгебре $\tilde{\mathcal{F}}_n$, можно заключить, что

$$\mathrm{E}\{D_n | \tilde{\mathcal{F}}_n\} \le (1 - 2\alpha \lambda_{\min}(\mathrm{B}\Gamma) + \alpha^2 \|\Gamma\|^2 M_4^4) \|\mathrm{A}\|^2 D_{n-1} +$$
$$+ \alpha^2 (\mathrm{E}\{\varphi_n\}^{\mathrm{T}} (\hat{\theta}_n - \theta_n^\star) - v_n)^2 \|\mathrm{A}\Gamma\|^2 \mathrm{Tr}[\mathrm{B}] + \sigma_w^2$$

($\mathrm{Tr}[\mathrm{B}]$ — след матрицы B).
Несложно убедиться, что при любом $\rho > 0$

$$2\mathrm{E}\{\varphi_n\}^{\mathrm{T}} (\hat{\theta}_n - \theta_n^\star) v_n \le \rho M_\varphi v_n^2 + \frac{M_\varphi}{\rho} D_{n-1}.$$

Далее, взяв безусловное математическое ожидание от обеих частей последней формулы, для среднего значения ошибки предсказания получаем при любом $\rho > 0$ оценку

$$\mathrm{E}\{D_n\} \le b(\alpha, \rho) \mathrm{E}\{D_{n-1}\} + \alpha^2 (1 + M_\varphi \rho) \|\mathrm{A}\Gamma\|^2 \mathrm{Tr}[\mathrm{B}] C_v^2 + \sigma_w^2,$$

где

$$(6)\ \ b(\alpha, \rho) = (1 - 2\alpha \lambda_{\min}(\mathrm{B}\Gamma) + \alpha^2 \|\Gamma\|^2 (M_4 + (M_\varphi + \frac{1}{\rho}) M_\varphi \mathrm{Tr}[\mathrm{B}])) \|\mathrm{A}\|^2.$$

4. Основной результат

Из последнего неравенства непосредственно следует заключение теоремы.

Т е о р е м а 1 *Пусть последовательности* $\{y_n\}$,$\{\varphi_n\}$,$\{v_n\}$,$\{\theta_n^\star\}$ *и* $\{w_n\}$ *связаны уравнениями (1) и (3),* $\alpha > 0$, Γ *— положительно определенная матрица и* $\hat{\theta}_0$ *— произвольный неслучайный вектор из* $\mathbb{R}^r$.

Если выполнены предположения **(П)**, *(2) и (4),* $\{v_n\}$ *— определяется неизвестной, но ограниченной детерминированной функцией,* $\{w_n\}$ *— центрированный независимый случайный процесс:* w_{n+1} *не зависит от* σ*-алгебры* $\mathcal{F}_n$,

тогда *для ошибок предсказания оценок* $\{\hat\theta_n\}$, *генерируемых по алгоритму (5), для любого* $\rho > 0$ *и достаточно малого* α *такого, что* $b(\alpha,\rho) < 1$, *выполняются неравенства*

$$\mathrm{E}\{\|\hat\theta_{n+1}-\theta^\star_{n+1}\|^2\} \le \frac{\sigma_w^2 + \alpha^2(1+M_\varphi\rho)\|\mathrm{A}\Gamma\|^2\mathrm{Tr}[\mathrm{B}]C_v^2}{1-b(\alpha,\rho)} + b(\alpha,\rho)^n\mathrm{E}\{\|\hat\theta_0-\theta^\star_0\|^2\},$$

где величина $b(\alpha,\rho)$ *определяется формулой (6).*

Результат теоремы 1 дает возможность изучить зависимость качества фильтрации от размера шага алгоритма α.

Пусть $\Gamma = \mathrm{B}^{-1}$, $\|\mathrm{A}\|^{-2} = 1 + \mathcal{O}(\alpha^3)$ и $\mathrm{E}\{\|\hat\theta_0 - \theta^\star_0\|^2\} = 0$. Обозначим через $c = \sigma_w^2/2$, $d(\rho) = (M_4 + (M_\varphi + 1/\rho)M_\varphi\mathrm{Tr}[\mathrm{B}])/(2\lambda^2_{\min}(\mathrm{B}))$. В этом случае для достаточно малых α из заключения теоремы 1 следует, что

$$\mathrm{E}\{\|\hat\theta_{n+1} - \theta^\star_{n+1}\|^2\} \le D(\alpha,\rho) + \mathcal{O}(\alpha^2),$$

где

$$(7)\qquad D(\alpha,\rho) = c\Big(\frac{1}{\alpha} + d(\rho) + \left(d(\rho)^2 + \frac{(1+M_\varphi\rho)\mathrm{Tr}[\mathrm{B}]C_v^2}{2c\lambda^2_{\min}(\mathrm{B})}\right)\alpha\Big).$$

Последнее выражение характеризует взаимное влияние возможности фильтрации и чувствительности к помехам. В случае $M_\varphi = 0$ при решении задачи фильтрации без прогноза похожий результат можно получить как следствие теоремы 4 из [7].

Оптимизируя по α и ρ выражение для $D(\alpha,\rho)$, находим

$$\alpha^\star = \left(d(\rho^\star)^2 + \frac{(1+M_\varphi\rho^\star)\mathrm{Tr}[\mathrm{B}]C_v^2}{2c\lambda^2_{\min}(\mathrm{B})}\right)^{-\frac{1}{2}},$$

где $\rho^\star$ — точка минимума функции

$$(8)\ \ \bar D(\rho) = c\left(d(\rho) + 2\sqrt{d(\rho)^2 + (1+M_\varphi\rho)\mathrm{Tr}[\mathrm{B}]C_v^2\lambda^{-2}_{\min}(\mathrm{B})/(2c)}\right).$$

Если $M_\varphi = 0$, то функция $D(\alpha,\rho)$ не зависит от ρ. В такой ситуации получаем, что

$$\alpha^\star = \frac{2\sigma_w\lambda^2_{\min}(\mathrm{B})}{\sqrt{M_4^2\sigma_w^2 + 4\lambda^2_{\min}(\mathrm{B})\mathrm{Tr}[\mathrm{B}]C_v^2}}$$

и
$$\bar{D}^\star = \frac{\sigma_w}{4\lambda^2_{\min}(\mathrm{B})}\left(M_4\sigma_w + 2\sqrt{M_4^2\sigma_w^2 + 4\lambda^2_{\min}(\mathrm{B})\mathrm{Tr}[\mathrm{B}]C_v^2}\right).$$
Пусть $r = 1$, $\{\varphi_n\}$ — скалярный бернуллиевский (равный $\pm\varphi$ с одинаковой вероятностью) независимый процесс и $\sigma_w^2 << C_v^2$, тогда

$$\alpha^\star \mathrm{A}\Gamma \approx \frac{\sigma_w}{|\varphi|C_v}.$$

Интересно заметить, что это значение приблизительно совпадает с предельной величиной калмановского коэффициента усиления для фильтра Калмана–Бьюси, если помехи наблюдения $\{v_n\}$ независимые и равны $\pm C_v$ с равной вероятностью.

5. Экспериментальные результаты

Рассмотрим пример применения рандомизированного алгоритма фильтрации (5) к решению задачи о предсказании значения скалярного сигнала $\{\theta_n^\star\}$, порождающегося устойчивым линейным фильтром

$$\theta^\star_{n+1} = 0,9999\theta^\star_n + w_{n+1},\ n = 0, 1, \ldots,\ \theta^\star_0 = 0,$$

где $\{w_n\}$ представляет собой реализацию независимых равномерно распределенных на интервале $[-\frac{1}{3}, \frac{1}{3}]$ случайных величин

$$\mathrm{E}\{w_n\} = 0,\ \mathrm{E}\{w_n^2\} = \frac{2}{81}.$$

Наблюдению в каждый момент времени доступны величины

$$y_n = \varphi_n\theta^\star_n + v_n,$$

представляющие собой смесь преобразованного сигнала $\{\theta^\star_n\}$ и *неизвестной, но ограниченной* детерминированной помехи $\{v_n\}$: $|v_n| \le 2$.

При компьютерном моделировании в качестве последовательности коэффициентов преобразования полезного сигнала в канале наблюдения $\{\varphi_n\}$ были выбраны случайные величины, порождаемые равномерным распределением на интервале $[0,5; 1,5]$. Процесс $\{\theta^\star_n\}$ наблюдался на интервале времени от 1 до 199. Качество предсказания определялось средней величиной невязки

$$\tilde{D}(\{\hat{\theta}_n\}) = \frac{1}{199}\sum_{n=1}^{199}\|\hat{\theta}_{n+1} - \theta^\star_{n+1}\|^2.$$

Оптимизация по ρ функции $\bar{D}(\rho)$ из (8) дает значение $\rho^\star = 0,269$. Подставив это значение в формулу для $\alpha^\star$, получаем, что в данном случае для алгоритма (5) оптимальное значение шага $\alpha^\star = 11,3808$, $\Gamma = 1/48$ и ошибка прогнозирования не превосходит $D(\alpha^\star,\rho^\star) = 1,3699 < C_v^2 = 4$.

Рис. 1 и 2 в типичных случаях показывают сравнительное поведение оценок предсказания, формируемых по трем алгоритмам: рандомизированному

$$\hat{\theta}_{n+1} = 0,9999(\hat{\theta}_n - 0,2371(\varphi_n - 1,0)(\varphi_n\hat{\theta}_n - y_n)), \tag{9}$$

упрощенному фильтру Калмана–Бьюси

$$\hat{\theta}_{n+1} = 0,9999(\hat{\theta}_n - 0,2371\varphi_n(\varphi_n\hat{\theta}_n - y_n)), \tag{10}$$

и фильтру Калмана–Бьюси

$$\hat{\theta}_{n+1} = 0,9999\hat{\theta}_n - k_n\varphi_n(\varphi_n\hat{\theta}_n - y_n), \tag{11}$$

$$k_n = \frac{0,9999\gamma_{n-1}}{\frac{16}{3} + \gamma_{n-1}\varphi_n^2},\ \gamma_n = \gamma_{n-1}0,9999^2 - \frac{\varphi_n^2\gamma_{n-1}^2}{\frac{16}{3} + \gamma_{n-1}\varphi_n^2} + \frac{2}{81},\ \gamma_0 = 0.$$

Известно, что фильтр Калмана–Бьюси (11) дает оптимальные оценки в случае гауссовых независимых помех в наблюдении, метод (10) достаточно эффективен при центрированных независимых помехах. Поэтому при центрированных случайных помехах поведение оценок, генерируемых по алгоритмам (10),(11) достаточно хорошее, несмотря на высокий уровень помех наблюдения (см. рис. 1). В ситуациях с постоянной неизвестной помехой или при нулевой в среднем, но недостаточно "разнообразной", средние значения ошибок алгоритмов (10) и (11) соизмеримы с квадратом уровня помехи (см. рис. 2). В то же время средний уровень ошибок оценок, доставляемых рандомизированным алгоритмом, во всех ситуациях примерно одинаковый и в несколько раз лучше уровня помехи. В таблице 1 сведены итоговые результаты средних значений ошибок типичных компьютерных экспериментов.

СПИСОК ЛИТЕРАТУРЫ

1. *Колмогоров А.Н.*, "Интерполирование и экстраполирование стационарных случайных последовательностей" // Известия АН СССР. Сер. матем. 1941. No. 5. С. 3–14.

2. *Busy R.S., Kalman R.E.*, "New Results In Linear Filtering and Prediction Theory" // J.Basic.Eng. ASME. 1961. V. 83, No. 1. P. 95–108.

3. *Фомин В.Н.*, "Рекуррентное оценивание и адаптивная фильтрация." М.: Наука. 1984, 288 с.

4. *Digailova I.A., Kurzhanski A.B.*, "The State Estimation Problem Under Mixed Uncertainty" // In: Proc. of the 5-th IFAC Symposium NOLCOS'01. 2001. St. Petersburg. P.584.

5. *Граничин О.Н.*, "Оценивание параметров линейной регрессии при произвольных помехах" // АиТ. 2002. No 1. С. 30–41.

6. *Граничин О.Н.*, "Рандомизированные алгоритмы стохастической аппроксимации при произвольных помехах" // АиТ. 2002. No 2. С. 44–55.

7. *Guo L., Ljung L., Wang G-J.*, "Necessary and Sufficient Conditions for Stability of LMS" // IEEE Trans. Automat. Control. 1997. V. 42. No 6. P. 761–770.

Таблица 1.

	$\tilde{D}((9))$	$\tilde{D}((10))$	$\tilde{D}((11))$
$v_n = 4,0 * (rand() - 0,5)$	0,5309	0,1803	0,1256
$v_n = 0,1 * sin(n) + 1,9 * sign(50 - n \bmod 100)$	0,5700	2,8254	2,2640
$v_n = 2,0$	0,5954	3,1387	2,5335
$v_n = -2,0$	0,7826	3,4989	3,9582

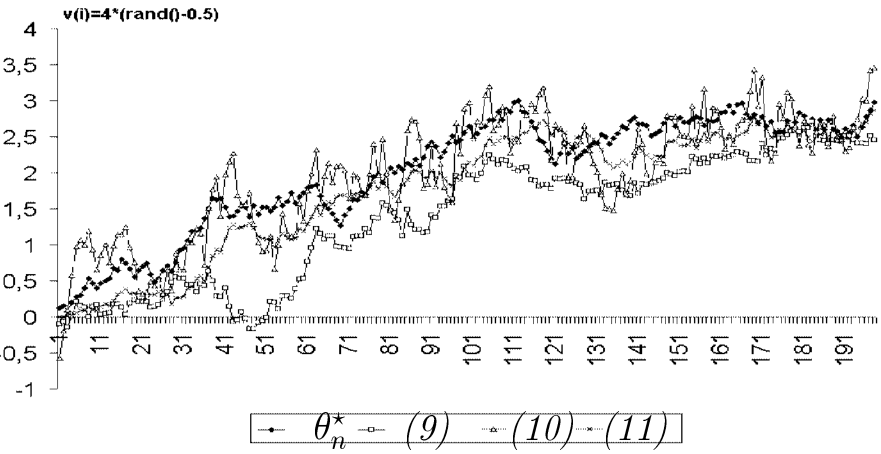

Рис. 1. Фильтрация при белошумных помехах

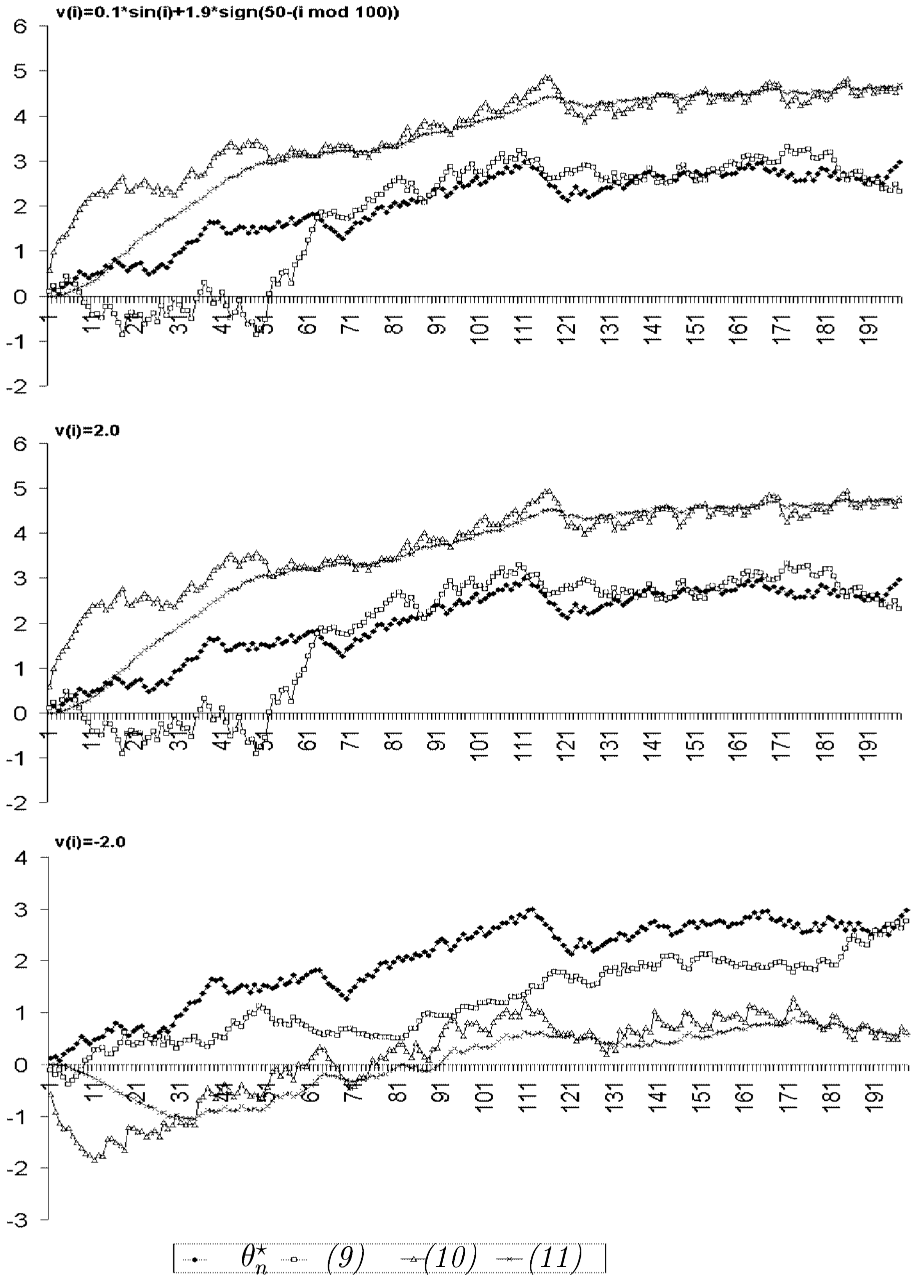

Рис. 2. Фильтрация при нерегулярных помехах