

Минимаксная линейная рекуррентная фильтрация неопределенно-стохастических последовательностей по интегральному критерию¹

А.Р.Панков, Г.Б.Миллер

Московский Авиационный Институт (Технический Университет), Москва, Россия

Поступила в редколлегию 06.08.2001

Аннотация—Рассматривается проблема рекуррентной фильтрации случайного процесса в разностной стохастической линейной системе со среднеквадратическим критерием качества. Особенность постановки задачи состоит в том, что моментные характеристики 1-го и 2-го порядков шумов динамики и ошибок наблюдений известны лишь с точностью до принадлежности некоторым множествам допустимых характеристик, заданным априорно. Для решения задачи предложен общий подход, основанный на построении минимаксных оценок параметров стохастических систем. Для построения минимаксного фильтра используется решение двойственной задачи выпуклого программирования. Предложен численный итерационный алгоритм определения параметров минимаксного фильтра, приведены результаты численного компьютерного моделирования.

1. ВВЕДЕНИЕ

Для практического использования алгоритма оптимальной рекуррентной фильтрации Калмана критическим моментом является наличие полной априорной информации о моментных характеристиках 1-го и 2-го порядков возмущающих процессов в уравнениях, описывающих динамику процесса и измерительную систему [1],[2]. Обычно на практике такая информация оказывается недоступной, что снижает достоверность получаемых оценок или делает их построение вообще невозможным. Для преодоления указанной трудности были начаты исследования по разработке минимаксных вариантов алгоритма калмановской фильтрации, обладающих свойством робастности к отклонению истинных моментных характеристик от расчетных [3]–[5]. Значительное число работ посвящено минимаксной нестохастической фильтрации Калмана, где возмущения предполагаются детерминированными неопределенными ограниченными последовательностями [6]. Для стохастических моделей в основном развивалась теория апостериорного минимаксного оценивания [7]. Для моделей частного вида теория минимаксной априорной фильтрации изучалась в работах [8]–[11], причем наиболее общие результаты относятся к стационарным системам [9], а также к моделям со специальными ограничениями на неопределенные возмущения [11]. В данной работе с помощью методов идентификации обобщенных неопределенно-стохастических регрессионных моделей [12],[13] при весьма общих предположениях о распределениях возмущений получены алгоритмы рекуррентной минимаксной фильтрации как для случая чисто стохастической модели наблюдения, так и при наличии совместных возмущений неопределенного и стохастического типа. Существенное внимание уделено решению проблемы численного расчета параметров алгоритма минимаксной фильтрации, для чего предложен универсальный метод, основанный на итерационном решении двойственной задачи оптимизации. Показано, что при весьма общих предположениях указанный метод является сходящимся. Предлагаемые в работе теоретические методы протестированы численно на ряде модельных примеров, результаты которых также обсуждаются в работе.

¹ Работа выполнена при поддержке гранта Российского фонда фундаментальных исследований, грант № 99-01-01088.

2. ОПТИМАЛЬНАЯ СТОХАСТИЧЕСКАЯ ФИЛЬТРАЦИЯ ПО ИНТЕГРАЛЬНОМУ КРИТЕРИЮ

Далее в работе используются следующие основные обозначения и сокращения: $E_P[x]$ – среднее случайного вектора x по распределению P ; $\text{cov}(x, y) = E_P[(x - m_x)(y - m_y)^*]$ – ковариация случайных векторов x, y ; $\text{col}(x_1, \dots, x_n)$ – колонка из векторов (матриц) $x_1, \dots, x_n$; $\text{tr}[A]$ – след матрицы A ; A^+ – псевдообратная к A матрица; $\text{co}[V]$ – выпуклая замкнутая оболочка множества точек V ; $\arg\min_{x \in X} f(x)$ – множество точек глобального минимума функции $f(x)$ на множестве X ; $A \geq B$, если матрица $A - B$ неотрицательно определена; $A > 0$ – матрица A положительно определена; $\text{diag}(A_1, \dots, A_n)$ – блочно-диагональная матрица; A^* – транспонированная матрица A ; $A \otimes B$ – кронекерово произведение матриц A и B ; $\ker[A]$ – ядро оператора A .

Рассмотрим классическую модель наблюдения Калмана [1],[2] с дискретным временем

$$\begin{cases} y_n = a_n y_{n-1} + b_n \varepsilon_n, & y_0 = \gamma, \\ z_n = c_n y_n + d_n w_n, & n = 1, \dots, N, \end{cases} \quad (1)$$

где $y_n \in \mathbb{R}^p$ – фазовый вектор системы; $z_n \in \mathbb{R}^q$ – вектор наблюдения в момент $n = 1, \dots, N$, где N – общее количество моментов наблюдений; $\gamma \in \mathbb{R}^p$ – случайный центрированный вектор начальных условий с ковариацией $\text{cov}(\gamma, \gamma) = S_\gamma$; ε_n и w_n – центрированные стационарные в широком смысле векторные белые шумы с $\text{cov}(\varepsilon_n, \varepsilon_n) = S_\varepsilon$ и $\text{cov}(w_n, w_n) = S_w$. Далее всюду предполагается, что $\gamma, \{\varepsilon_n\}, \{w_n\}$ независимы в совокупности. Матрицы $\{a_n, b_n, c_n, d_n\}$ считаются неслучайными и известными.

Предположим, что по наблюдениям $\{z_n\}$ подлежит оцениванию процесс $x_n = \varphi_n y_n$, $n = 1, \dots, N$, где $\{\varphi_n\}$ – произвольная последовательность заданных неслучайных матриц, $x_n \in \mathbb{R}^r$. Обозначим $Z_n = \text{col}(z_1, \dots, z_n)$, $f_n = (f_{n1}, f_{n2}, \dots, f_{nn})$, где f_{nk} – произвольная матрица размера $(r \times q)$.

Случайный вектор

$$\tilde{x}_n = f_n Z_n, \quad n = 1, \dots, N, \quad (2)$$

называется линейной оценкой фильтрации для x_n по наблюдениям Z_n .

Если дополнить f_n нулевыми блоками до матрицы F_n размера $(r \times (q \cdot N))$, то оценку (2) можно представить в виде

$$\tilde{x}_n = F_n Z_N, \quad n = 1, \dots, N.$$

Обозначим $F = \text{col}(F_1, \dots, F_N)$, а совокупность всех неупреждающих операторов оценивания (стратегий) F рассмотренного вида, т.е. операторов фильтрации, будем далее обозначать через $\mathcal{F}$. Очевидно, что множество $\mathcal{F}$ выпукло, замкнуто, но в общем случае неограничено.

Для определения точности оценки $\{\tilde{x}_n\}$ процесса $\{x_n\}$ на всем интервале наблюдения будем использовать интегральный среднеквадратический критерий

$$\mathbf{J}(F, P_\rho) = E_{P_\rho} \left[\sum_{n=1}^N (x_n - \tilde{x}_n)^* (x_n - \tilde{x}_n) \right], \quad (3)$$

где P_ρ – закон распределения вектора $\rho = \text{col}(\gamma, \varepsilon_1, \dots, \varepsilon_N, w_1, \dots, w_N)$, объединяющего все случайные векторы, входящие в (1).

Предположим, что ковариации S_γ, S_ε и S_w известны. В этом случае P_ρ определяется лишь с точностью до принадлежности классу распределений $\mathcal{P}_0$, удовлетворяющему всем сделанным выше предположениям относительно ρ .

Пусть $\hat{F} = \text{col}(\hat{F}_1, \dots, \hat{F}_N)$ – оптимальный по критерию (3) на классе $\mathcal{P}_0$ оператор оценивания из $\mathcal{F}$:

$$\mathbf{J}(\hat{F}, P_\rho) \leq \mathbf{J}(F, P_\rho) \quad \forall F \in \mathcal{F}, \quad \forall P_\rho \in \mathcal{P}_0.$$

Такой оператор существует и может быть вычислен с помощью рекуррентных уравнений фильтра Калмана [1],[2]:

$$\hat{x}_n = \varphi_n \hat{y}_n, \quad n = 1, \dots, N, \quad (4)$$

где

$$\begin{cases} \hat{y}_n = a_n \hat{y}_{n-1} + k_n (z_n - c_n a_n \hat{y}_{n-1}), \quad \hat{y}_0 = 0, \\ k_n = \bar{P}_n c_n^* (c_n \bar{P}_n c_n^* + d_n S_w d_n^*)^{-1}; \end{cases} \quad (5)$$

$$\begin{cases} \bar{P}_n = a_n \bar{P}_{n-1} a_n^* + b_n S_\varepsilon b_n^*, \quad P_0 = S_\gamma, \\ P_n = (I - k_n c_n) \bar{P}_n. \end{cases} \quad (6)$$

Действительно, оценка $\hat{x}_n$ в силу (4)–(6) удовлетворяет условию неупреждаемости $\hat{x}_n = \hat{F}_n Z_N$, $\forall n \leq N$, поэтому оператор $\hat{F} = \text{col}(\hat{F}_1, \dots, \hat{F}_N) \in \mathcal{F}$. Известно также [1],[2], что $E_{P_\rho}[\hat{y}_n - y_n] = 0$ $\forall P_\rho \in \mathcal{P}_0$, а

$$P_n = \text{cov}(y_n - \hat{y}_n, y_n - \hat{y}_n) \leq E_{P_\rho}[(y_n - \tilde{y}_n)(y_n - \tilde{y}_n)^*] = \tilde{P}_n$$

для любой оценки $\tilde{y}_n = \tilde{F}_n Z_N$, где $\tilde{F} = \text{col}(\tilde{F}_1, \dots, \tilde{F}_N) \in \mathcal{F}$.

Обозначим $Q_n = \varphi_n^* \varphi_n$, тогда при каждом $n = 1, 2, \dots, N$ выполнено $E_{P_\rho}[(x_n - \hat{x}_n)] = 0$ и

$$E_{P_\rho}[(x_n - \hat{x}_n)^*(x_n - \hat{x}_n)] = E_{P_\rho}[(y_n - \hat{y}_n)^* Q_n (y_n - \hat{y}_n)] = \text{tr}[Q_n P_n] \leq \text{tr}[Q_n \tilde{P}_n],$$

так как $\text{tr}[Q_n \tilde{P}_n] - \text{tr}[Q_n P_n] = \text{tr}[Q_n^{\frac{1}{2}} (\tilde{P}_n - P_n) Q_n^{\frac{1}{2}}] \geq 0$ в силу $\tilde{P}_n - P_n \geq 0$ и $Q_n \geq 0$.

Таким образом, для $\forall F \in \mathcal{F}$ и $\forall P_\rho \in \mathcal{P}_0$ выполнено

$$\mathbf{J}(\hat{F}, P_\rho) = \sum_{n=1}^N E_{P_\rho}[(y_n - \hat{y}_n)^* Q_n (y_n - \hat{y}_n)] \leq \sum_{n=1}^N E_{P_\rho}[(y_n - \tilde{y}_n)^* Q_n (y_n - \tilde{y}_n)] = \mathbf{J}(\tilde{F}, P_\rho).$$

Заметим, что стратегия $\hat{F}$, определяемая (4)–(6), на классе распределений $\mathcal{P}_0$ является уравнивающей [14], т.е.

$$\mathbf{J}(\hat{F}, P_\rho) = \mathbf{J}(\hat{F}, P'_\rho) = \sum_{n=1}^N \text{tr}[Q_n P_n], \quad \forall P_\rho, P'_\rho \in \mathcal{P}_0.$$

Поэтому $\hat{F}$ является также и минимаксной стратегией из $\mathcal{F}$ на классе распределений $\mathcal{P}_0$:

$$\hat{F} \in \underset{F \in \mathcal{F}}{\text{argmin}} \sup_{P_\rho \in \mathcal{P}_0} \mathbf{J}(F, P_\rho).$$

Более того, пара $(\hat{F}, \hat{P}_\rho)$, где $\hat{P}_\rho$ – гауссовское распределение из $\mathcal{P}_0$, является седловой точкой критерия $\mathbf{J}(F, P_\rho)$ на $\mathcal{F}' \times \mathcal{P}_0$, где $\mathcal{F}'$ – класс всех измеримых неупреждающих оценивателей (как линейных, так и нелинейных):

$$\mathbf{J}(\hat{F}, P_\rho) \leq \mathbf{J}(\hat{F}, \hat{P}_\rho) \leq \mathbf{J}(F, \hat{P}_\rho), \quad \forall F \in \mathcal{F}', \quad P_\rho \in \mathcal{P}_0.$$

Итак, алгоритм (4)–(6) обеспечивает оптимальную фильтрацию процесса $\{x_n\}$ по интегральному критерию $\mathbf{J}(F, P_\rho)$ на $\mathcal{F}' \times \mathcal{P}_0$.

3. МИНИМАКСНАЯ СТОХАСТИЧЕСКАЯ ФИЛЬТРАЦИЯ ПО ИНТЕГРАЛЬНОМУ КРИТЕРИЮ

Предположим теперь, что ковариации S_γ , S_ε и S_w возмущений в модели (1) известны лишь с точностью до принадлежности некоторым заданным множествам неопределенности:

$$S_\gamma \in D_\gamma, \quad S_\varepsilon \in D_\varepsilon, \quad S_w \in D_w. \quad (7)$$

Обозначим через $K = \text{diag}(S_\gamma, S_\varepsilon, S_w)$ и будем полагать, что $\mathcal{K}$ — множество всех матриц K , удовлетворяющих (7). Тогда из (4)–(6) следует, что при каждом $K \in \mathcal{K}$

$$P_n = P_n(K), \quad \hat{F}_n = \hat{F}_n(K), \quad n = 1, \dots, N,$$

поэтому

$$\mathbf{J}(\hat{F}, P_\rho) = I(K) = \sum_{n=1}^N \text{tr}[Q_n P_n(K)] \quad (8)$$

Обозначим через $\mathcal{P}$ класс всех распределений P_ρ вектора ρ таких, что $E_{P_\rho}[\rho] = 0$ и $R_\rho(K) = \text{cov}(\rho, \rho) = \text{diag}(S_\gamma, I_N \otimes S_\varepsilon, I_N \otimes S_w)$ такова, что $K \in \mathcal{K}$.

Определение. Оператор $\hat{F}$ оценки фильтрации $\{\hat{x}_n\}$ процесса $\{x_n\}$ называется минимаксным по критерию $\mathbf{J}(F, P_\rho)$ на классе $\mathcal{F} \times \mathcal{P}$, если

$$\hat{F} \in \underset{F \in \mathcal{F}}{\text{argmin}} \sup_{P_\rho \in \mathcal{P}} \mathbf{J}(F, P_\rho).$$

Покажем, что оператор $\hat{F}$ при определенных условиях существует и задается уравнениями (4)–(6), если надлежащим образом выбрана матрица $K \in \mathcal{K}$. Для этого воспользуемся результатами по минимаксному оцениванию в обобщенных неопределенно-стохастических моделях линейной регрессии [12],[13].

Преобразуем модель (1) к регрессионному виду. При каждом $n = 1, \dots, N$ выполнено

$$y_n = \Phi^\gamma \gamma + \Phi_n^\varepsilon E_n, \quad (9)$$

где $E_n = \text{col}(\varepsilon_1, \dots, \varepsilon_n)$, $\Phi^\gamma = \prod_{k=0}^{N-1} a_{N-k}$, а

$$\Phi_n^\varepsilon = [a_n a_{n-1} \dots a_2 b_1, a_n a_{n-1} \dots a_3 b_2, \dots, a_n b_{n-1}, b_n].$$

Подставляя (9) в (1), получаем:

$$z_n = c_n \Phi^\gamma \gamma + c_n \Phi_n^\varepsilon E_n + d_n w_n = \Psi^\gamma \gamma + \Psi_n^\varepsilon E_n + d_n w_n,$$

где $\Psi^\gamma = c_n \Phi^\gamma$, $\Psi_n^\varepsilon = c_n \Phi_n^\varepsilon$.

Таким образом, $y_n = \Phi_n \rho$ и $z_n = \Psi_n \rho$, где

$$\Phi_n = [\Phi^\gamma \quad \Phi_n^\varepsilon \quad O_{n+1}^{y,\varepsilon} \dots O_N^{y,\varepsilon} \quad O_1^{y,w} \dots O_N^{y,w}],$$

$O_i^{y,\varepsilon}$, $O_i^{y,w}$, $i = 1, \dots, N$ — нулевые матрицы размера $p \times p$ и $p \times q$ соответственно. Аналогично,

$$\Psi_n = [\Psi^\gamma \quad \Psi_n^\varepsilon \quad O_{n+1}^{z,\varepsilon} \dots O_N^{z,\varepsilon} \quad O_1^{z,w} \dots O_{n-1}^{z,w} \quad d_n \quad O_{n-1}^{z,w} \dots O_N^{z,w}].$$

Пусть теперь $X_n = \text{col}(x_1, \dots, x_n)$, $\Phi = \text{col}(\varphi_1 \Phi_1, \dots, \varphi_N \Phi_N)$, $\Psi = \text{col}(\Psi_1, \dots, \Psi_N)$, тогда в этих обозначениях модель наблюдения (1) принимает вид

$$\begin{cases} X_N = \Phi \rho \\ Z_N = \Psi \rho, \end{cases} \quad (10)$$

и, следовательно, является частным случаем модели, рассмотренной в [12].

Для произвольной линейной неупреждающей оценки $\tilde{X}_N = F Z_N$, $F \in \mathcal{F}$ вектора X_N интегральный критерий качества (3) имеет вид

$$\mathbf{J}(F, P_\rho) = E_{P_\rho} [\|X_N - F Z_N\|^2], \quad P_\rho \in \mathcal{P}.$$

Очевидно, $\mathbf{J}(F, P_\rho)$ можно представить как функцию от F и K , где $K \in \mathcal{K}$. Действительно,

$$\mathbf{J}(F, P_\rho) = E_{P_\rho} [\|\Phi \rho - F \Psi \rho\|^2] = \text{tr}[(\Phi - F \Psi) R_\rho(K) (\Phi - F \Psi)^*],$$

где $R_\rho(K) = \text{diag}[S_\gamma, I_N \otimes S_\varepsilon, I_N \otimes S_w]$, $K = \text{diag}(S_\gamma, S_\varepsilon, S_w) \in \mathcal{K}$.

Таким образом, при каждом $F \in \mathcal{F}$ и $K \in \mathcal{K}$

$$\mathbf{J}(F, P_\rho) = \text{tr}[(\Phi - F \Psi) R_\rho(K) (\Phi - F \Psi)^*] = J(F, K). \quad (11)$$

Переходя в силу (11) от $\mathbf{J}(F, P_\rho)$ к $J(F, K)$, задачу минимаксной оптимизации фильтра можно представить в окончательном виде:

$$\hat{F} \in \underset{F \in \mathcal{F}}{\text{argmin}} \sup_{K \in \mathcal{K}} J(F, K). \quad (12)$$

Двойственная по отношению к (12) задача оптимизации имеет вид:

$$\hat{K} \in \underset{K \in \mathcal{K}}{\text{argmax}} \inf_{F \in \mathcal{F}} J(F, K). \quad (13)$$

Так как при каждом $K \in \mathcal{K}$

$$\inf_{F \in \mathcal{F}} J(F, K) = \inf_{F \in \mathcal{F}} \mathbf{J}(F, P_\rho) = I(K),$$

где $I(K)$ определен в (8), то (13) имеет вид

$$\hat{K} \in \underset{K \in \mathcal{K}}{\text{argmax}} I(K), \quad (14)$$

где $I(K)$ называется двойственным функционалом, а задача (14) – двойственной задачей оптимизации.

Если множества неопределенности D_γ , D_ε и D_w – выпуклые компакты, то $R_\rho(K)$ в (11) также есть выпуклый компакт. Тогда из теоремы 3 работы [13] следует, что решение $\hat{K} = \text{diag}(\hat{S}_\gamma, \hat{S}_\varepsilon, \hat{S}_w)$ задачи (13) существует. При этом $\hat{K}$ имеет смысл “наихудшей” ковариации на множестве $\mathcal{K}$.

Рассмотрим оператор фильтрации $F(\hat{K})$ вида

$$F(\hat{K}) \in \underset{F \in \mathcal{F}}{\text{argmin}} J(F, \hat{K}), \quad (15)$$

т.е. оператор оптимальной фильтрации, построенный при условии $K = \hat{K}$. Соответствующая оценка вычисляется по формулам (4)–(6), где следует заменить S_γ , S_ε и S_w , соответственно, на $\hat{S}_\gamma$, $\hat{S}_\varepsilon$ и $\hat{S}_w$. Оператор $F(\hat{K})$ будет минимаксным, если

$$F(\hat{K}) \in \operatorname{argmin}_{F \in \mathcal{F}} \sup_{K \in \mathcal{K}} J(F, K). \quad (16)$$

В [12] показано, что для (15) будет выполнено (16) в следующих случаях:

а) множества D_γ , D_ε и D_w имеют максимальные элементы $\bar{S}_\gamma$, $\bar{S}_\varepsilon$ и $\bar{S}_w$, т.е. для $\forall S_\gamma \in D_\gamma$, $S_\varepsilon \in D_\varepsilon$ и $S_w \in D_w$ выполнены неравенства

$$S_\gamma \leq \bar{S}_\gamma, \quad S_\varepsilon \leq \bar{S}_\varepsilon \quad \text{и} \quad S_w \leq \bar{S}_w;$$

б) множество $\operatorname{argmin}_{F \in \mathcal{F}} \sup_{K \in \mathcal{K}} J(F, K) \neq \emptyset$, а $F(\hat{K})$ – единственная точка минимума $J(F, \hat{K})$ на $\mathcal{F}$.

Пусть $R_z(K) = \operatorname{cov}(Z_N, Z_N) = \Psi R_\rho(K) \Psi^*$ – ковариационная матрица вектора наблюдений в (10). Если $\mathcal{K}$ таково, что

$$\bigcap_{K \in \mathcal{K}} \ker[R_z(K)] = \{0\}, \quad (17)$$

функционал $\bar{J}(F) = \sup_{K \in \mathcal{K}} J(F, K)$ не имеет рецессивных направлений, т.е. $\bar{J}(F) \rightarrow \infty$ при $\|F\| \rightarrow \infty$, откуда следует (см. теорему 27.2 из [15]), что множество

$$\mathcal{F}_0 = \operatorname{argmin}_{F \in \mathcal{F}} \bar{J}(F) \neq \emptyset, \quad (18)$$

является замкнутым и ограниченным. Если к тому же $\hat{K}$ такова, что $R_z(\hat{K}) > 0$, то $F(\hat{K})$ – единственный оптимальный оператор фильтрации при $K = \hat{K}$, т.е.

$$F(\hat{K}) = \operatorname{argmin}_{F \in \mathcal{F}} J(F, \hat{K}).$$

Таким образом, из условия $R_z(\hat{K}) > 0$ следует (16), так как в этом случае (17) так же выполнено, т.е. имеет место (18).

Наконец, если $\hat{P}_\rho$ – гауссовское распределение с нулевым средним и ковариацией $R_\rho(\hat{K})$, то $F(\hat{K})$ – наилучший оператор фильтрации не только на $\mathcal{F}$, но также и на $\mathcal{F}'$ – классе произвольных измеримых неупреждающих операторов оценивания.

Сформулируем теперь полученные результаты в виде теоремы.

Теорема 1. Пусть множество неопределенности $\mathcal{K}$ – выпуклый компакт, тогда

1) решение $\hat{K} = \operatorname{diag}(\hat{S}_\gamma, \hat{S}_\varepsilon, \hat{S}_w)$ двойственной задачи

$$\hat{K} \in \operatorname{argmax}_{K \in \mathcal{K}} \sum_{n=1}^N \operatorname{tr}[Q_n P_n(K)]$$

существует.

2) Пусть выполнено любое из следующих условий:

а) множество $\mathcal{K}$ имеет максимальный элемент $K_{\max}$:

$$K \leq K_{\max} \quad \forall K \in \mathcal{K};$$

б) в точке $\hat{K}$ выполнено $R_z(\hat{K}) > 0$,

тогда минимаксная оценка $\{\hat{x}_n\}$ процесса $\{x_n\}$ по критерию $\mathbf{J}(F, P_\rho)$, где $F \in \mathcal{F}'$, $P_\rho \in \mathcal{P}$ определяется соотношениями

$$\begin{cases} \hat{x}_n = \varphi_n \hat{y}_n, & n = 1, \dots, N; \\ \hat{y}_n = a_n \hat{y}_{n-1} + \hat{k}_n (z_n - c_n a_n \hat{y}_{n-1}), & \hat{y}_0 = 0, \end{cases} \quad (19)$$

где

$$\hat{k}_n = \bar{P}_n c_n^* (c_n \bar{P}_n c_n^* + d_n \hat{S}_w d_n^*)^+, \quad (20)$$

$$\begin{cases} \bar{P}_n = a_n P_{n-1}(\hat{K}) a_n^* + b_n \hat{S}_\varepsilon b_n^*, & P_0(\hat{K}) = \hat{S}_\gamma, \\ P_n(\hat{K}) = (I - \hat{k}_n c_n) \bar{P}_n. \end{cases} \quad (21)$$

3) Для $\forall P_\rho \in \mathcal{P}$ выполнено

$$\mathbf{J}(F(\hat{K}), P_\rho) = E_{P_\rho} \left[\sum_{n=1}^N \|x_n - \hat{x}_n\|^2 \right] \leq J(F(\hat{K}), \hat{K}) = I(\hat{K}) = \sum_{n=1}^N \text{tr}[Q_n P_n(\hat{K})], \quad (22)$$

причем равенство в (22) достигается при $P_\rho = \hat{P}_\rho$, т.е. гарантированное значение критерия не улучшаемо.

4) Пара $(F(\hat{K}), \hat{P}_\rho)$ образует седловую точку критерия $\mathbf{J}(F, P_\rho)$ на $\mathcal{F}' \times \mathcal{P}$.

4. АЛГОРИТМ ПОСТРОЕНИЯ МИНИМАКСНОГО ФИЛЬТРА

Из теоремы 1 следует, что если решение $\hat{K}$ двойственной задачи (14) найдено, то минимаксный оператор $\hat{F} = F(\hat{K})$ имеет аналитическое выражение (19)–(21). Поэтому для построения $\hat{F}$ достаточно решить аналитически или численно задачу (14).

Если $\exists K_{max} \in \mathcal{K}$ (см. п.2 теоремы 1), то

$$\bar{J}(F) = \sup_{K \in \mathcal{K}} J(F, K) = \sup_{K \in \mathcal{K}} \text{tr}[(\Phi - F\Psi)R_\rho(K)(\Phi - F\Psi)^*] = \text{tr}[(\Phi - F\Psi)R_\rho(K_{max})(\Phi - F\Psi)^*].$$

Поэтому $K_{max} \in \arg\max_{K \in \mathcal{K}} I(K)$, т.е. $\hat{F} = F(K_{max})$, где $F(K)$ – алгоритм фильтрации Калмана (19)–(21). Таким образом, в данном случае проблема минимаксной фильтрации имеет аналитическое решение (см. также [3]–[5]).

Если же $\mathcal{K}$ не содержит максимального элемента, то для нахождения $\hat{K}$ предлагается использовать итерационный алгоритм, описанный ниже.

Прежде всего преобразуем исходную задачу рекуррентной фильтрации к виду, удобному для вычисления значения критерия качества $J(F, K)$ при произвольном $K \in \mathcal{K}$ и заданном операторе оценивания F вида (19)–(21).

Из выражения (11) следует, что при каждом фиксированном $F \in \mathcal{F}$ функционал $J(F, K)$ линеен по $K \in \mathcal{K}$ и, следовательно, может быть представлен в виде

$$J(F, K) = \text{tr}[\Lambda_F K], \quad (23)$$

где Λ_F – некоторая неотрицательно определенная матрица, зависящая только от F . Предположим, что $K^{(s)}$ – приближение для $\hat{K}$, полученное на s – ой итерации. Обозначим через $F^{(s)}$ оператор

фильтра Калмана (4)–(6), вычисленный для $K = K^{(s)} = \text{diag}(S_\gamma^{(s)}, S_\varepsilon^{(s)}, S_w^{(s)})$. Рассмотрим алгоритм вычисления матрицы $\Lambda^{(s)} = \Lambda_{F^{(s)}}$ в (23), соответствующей $F^{(s)}$.

Для произвольного $K \in \mathcal{K}$ справедливо

$$J(F^{(s)}, K) = E_{P_\rho} \left[\sum_{n=1}^N \|x_n - x_n^{(s)}\|^2 \right] = \sum_{n=1}^N \text{tr}[Q_n R_\Delta^{(s)}(n)],$$

где $R_\Delta^{(s)}(n) = E_{P_\rho}[(y_n - y_n^{(s)})(y_n - y_n^{(s)})^*]$, $y_n^{(s)}$ – оценка для y_n вида $y_n^{(s)} = F_n^{(s)} Z_N$, $F^{(s)} = \text{col}(F_1^{(s)}, \dots, F_N^{(s)})$, $x_n^{(s)} = \varphi_n y_n^{(s)}$, P_ρ – распределение ρ , соответствующее K (т.е. истинное распределение ρ). Найдем разностное уравнение для вычисления $K_\Delta^{(s)}(n)$.

Пусть $\Delta_n^{(s)} = y_n - y_n^{(s)}$, тогда с учетом соотношения

$$y_n^{(s)} = a_n y_{n-1}^{(s)} + k_n^{(s)}(z_n - c_n a_n y_{n-1}^{(s)}),$$

где $\{k_n^{(s)}\}$ – известные матрицы усиления в алгоритме $F^{(s)}$ калмановской фильтрации, нетрудно получить уравнение:

$$\Delta_n^{(s)} = G_n^{(s)} \Delta_{n-1}^{(s)} + g_n^{(s)} \varepsilon_n + h_n^{(s)} w_n, \quad \Delta_0^{(s)} = \gamma, \quad (24)$$

где

$$G_n^{(s)} = (I - k_n^{(s)} c_n) a_n, \quad g_n^{(s)} = (I - k_n^{(s)} c_n) b_n \quad \text{и} \quad h_n^{(s)} = -k_n^{(s)} d_n.$$

Уравнение (24) является разностным стохастическим для процесса $\Delta_n^{(s)}$, поэтому $R_\Delta^{(s)}(n) = E_{P_\rho} \{ \Delta_n^{(s)} (\Delta_n^{(s)})^* \}$ удовлетворяет уравнению метода моментов [1]:

$$\begin{cases} R_\Delta^{(s)}(n) = G_n^{(s)} R_\Delta^{(s)}(n-1) (G_n^{(s)})^* + g_n^{(s)} S_\varepsilon (g_n^{(s)})^* + h_n^{(s)} S_w (h_n^{(s)})^*, \\ R_\Delta^{(s)}(0) = S_\gamma. \end{cases} \quad (25)$$

Функционал (23) можно представить в еще более простом виде, если обозначить $u = \text{col}(u_\gamma, u_\varepsilon, u_w)$, где u_γ , u_ε и u_w – векторы, составленные из всех элементов, соответственно, матриц S_γ , S_ε и S_w , лежащих на главных диагоналях и над ними. Например, если $S_\gamma = \{s_{ij}^\gamma\}$, то $u_\gamma = \text{col}(s_{11}^\gamma, \dots, s_{1p}^\gamma, s_{22}^\gamma, \dots, s_{2p}^\gamma, \dots, s_{pp}^\gamma)$. Таким образом, имеется взаимно-однозначное соответствие $K = T(u)$ между матрицами $K \in \mathcal{K}$ и векторами u , причем если компонента u_i вектора u соответствует элементу k_{lm} матрицы K , то $k_{lm} = k_{ml} = u_i$. Заметим также, что если $\mathcal{K}$ – выпуклый компакт, то $u \in \mathbb{U}$, где $\mathbb{U} = T^{-1}(\mathcal{K})$ также является выпуклым компактом.

Теперь можно записать

$$J(F^{(s)}, K) = \text{tr}[\Lambda^{(s)} K] = (\lambda^{(s)}, u) = f_s(u), \quad (26)$$

где $\Lambda^{(s)}$ однозначно выражается через $\lambda^{(s)}$.

Теперь нетрудно вычислить элементы $\{\lambda_i^{(s)}\}$ вектора $\lambda^{(s)}$, учитывая, что

$$\lambda_i^{(s)} = f_s(u_i) = J(F^{(s)}, K_i),$$

где вектор u_i является нулевым за исключением i -ой компоненты, которая равна 1, а $K_i = T(u_i)$. Теперь из (25), (26) следует, что для каждого $i = 1, 2, \dots, \dim(u)$ выполнено

$$\begin{cases} \lambda_i^{(s)} = J(F^{(s)}, K_i) = \sum_{n=1}^N \text{tr}[Q_n R_{\Delta}^i(n)], \\ R_{\Delta}^i(n) = G_n^{(s)} R_{\Delta}^i(n-1) (G_n^{(s)})^* + g_n^{(s)} S_{\varepsilon}^i (g_n^{(s)})^* + h_n^{(s)} S_w^i (h_n^{(s)})^*, \\ R_{\Delta}^i(0) = S_{\gamma}^i, \end{cases} \quad (27)$$

где $K_i = \text{diag}(S_{\gamma}^i, S_{\varepsilon}^i, S_w^i)$. Заметим, что формулы (27) реализуются весьма просто, так как коэффициенты Q_n , $G_n^{(s)}$, $g_n^{(s)}$ и $h_n^{(s)}$ не зависят от i , а матрицы S_{γ}^i , S_{ε}^i и S_w^i таковы, что какие-то две из них обязательно нулевые, а третья имеет не более двух отличных от нуля элементов.

Опишем теперь итерационный алгоритм решения двойственной задачи (14), позволяющий найти матрицу наилучших ковариаций $\hat{K}$.

- Алгоритм 1.** 1) Выбрать произвольно $K^{(0)} \in \mathcal{K}$ и положить $s = 0$.
 2) Вычислить $\lambda^{(s)}$ по формулам (27).
 3) Вычислить $I(K^{(s)}) = J(F^{(s)}, K^{(s)}) = (\lambda^{(s)}, u_s)$, где $u_s = T^{-1}(K^{(s)})$.
 4) Решить задачу обобщенного линейного программирования

$$\tilde{u}_s \in \underset{u \in \mathbb{U}}{\text{argmax}} (\lambda_s, u)$$

и определить соответствующую матрицу $\tilde{K}^{(s)} = T(\tilde{u}_s)$.

- 5) Вычислить

$$\delta_s = J(F^{(s)}, \tilde{K}^{(s)}) - J(F^{(s)}, K^{(s)}) = (\lambda^{(s)}, \tilde{u}_s) - I(K^{(s)}).$$

Если $\delta_s = 0$, то положить $\hat{K} = K^{(s)}$ и закончить итерации.

Если $\delta_s > 0$, то перейти к следующему шагу.

- 6) Найти максимум $I(K)$ по направлению $\{K^{(s)} \rightarrow \tilde{K}^{(s)}\}$, т.е. определить число τ_s из условия

$$\tau_s \in \underset{0 \leq \tau \leq \bar{\tau}_s}{\text{argmax}} I(K_{\tau}^{(s)}), \quad (28)$$

где $K_{\tau}^{(s)} = (1 - \tau)K^{(s)} + \tau\tilde{K}^{(s)}$, а $\bar{\tau}_s = \max\{\tau \geq 0 : K_{\tau}^{(s)} \in \mathcal{K}\}$.

- 7) Положить $K^{(s+1)} = K_{\tau_s}^{(s)}$, увеличить s на 1 и перейти к п. 2) алгоритма.

Докажем сходимость описанного итерационного процесса. Пусть $\Omega = \underset{K \in \mathcal{K}}{\text{argmax}} I(K)$ — множество всех точек глобального максимума двойственного функционала $I(K)$, $R_z(K) = \Psi R_{\rho}(K) \Psi^*$ — ковариация вектора наблюдений Z_N в (10), $d(K, \Omega) = \inf_{\tilde{K} \in \Omega} \|K - \tilde{K}\|$ — расстояние от $K \in \mathcal{K}$ до множества Ω .

Теорема 2. Пусть $\mathcal{K}$ выпукло, компактно и $R_z(K) > 0$ для $\forall K \in \mathcal{K}$. Тогда $d(K^{(s)}, \Omega) \rightarrow 0$ при $s \rightarrow \infty$, если $\{K^{(s)}\}$ — бесконечная последовательность. Если же $\hat{K} = K^{(s)}$ при $s < \infty$, то $\hat{K} \in \Omega$.

Доказательство. Рассмотрим многозначное алгоритмическое отображение $\mathcal{L} : \mathcal{K} \rightarrow \mathcal{K}$, где $K^{(s+1)} = \mathcal{L}(K^{(s)})$, $K^{(0)} \in \mathcal{K}$, определенное алгоритмом 1. В условиях теоремы $I(K) = J(F(K), K)$, где $F(K)$ — оператор фильтрации Калмана, который в нерекуррентном виде дает оценку для X_N по Z_N в модели (10), т.е. $\hat{X}_N = F(K)Z_N$. В силу условия $R_z(K) > 0$, компактности $\mathcal{K}$ и линейности $R_{\rho}(K)$

по K на $\mathcal{K}$ заключаем, что $F(K)$ непрерывно зависит от K , а множество $\mathcal{F}^0 = \{F(K) : K \in \mathcal{K}\}$ — компактно. Следовательно, по теореме 7.2 из [16] заключаем, что функционал

$$I(K) = \min_{F \in \mathcal{F}^0} J(F, K) \quad (29)$$

непрерывен по K на $\mathcal{K}$.

Покажем, что $K^{(s)} \in \Omega$ тогда и только тогда, когда $\delta_s = 0$. Функционал $I(K)$ есть точная нижняя грань на $\mathcal{F}^0$ линейных по K функций $J(F, K)$, поэтому $I(K)$ вогнут на $\mathcal{K}$. В силу этого условия $K^{(s)} \in \Omega$ равносильно

$$\sup_{K \in \mathcal{K}} I^{(t)}(K^{(s)})(K - K^{(s)}) = 0, \quad (30)$$

где $I^{(t)}(K^{(s)})(\Delta K)$ — производная $I(K)$ в точке $K^{(s)}$ по направлению ΔK .

В силу (29), компактности $\mathcal{K}$ и теоремы 3.5 из [17] заключаем, что

$$\begin{aligned} I^{(t)}(K^{(s)})(K - K^{(s)}) &= J(F^{(s)}, K - K^{(s)}) = J(F^{(s)}, K) - J(F^{(s)}, K^{(s)}) \\ &= J(F^{(s)}, K) - I(K^{(s)}) = (\lambda^{(s)}, u) - I(K^{(s)}). \end{aligned}$$

Поэтому условие (30) равносильно $\delta_s = 0$, т.е. $K^{(s)} \in \Omega$.

Если же $K^{(s)} \notin \Omega$, то $\delta_s > 0$. Из определения $K^{(s+1)}$ следует, что

$$I(K^{(s+1)}) = \max_{0 \leq \tau \leq \bar{\tau}_s} I((1 - \tau)K^{(s)} + \tau \tilde{K}^{(s)}) \geq I(K^{(s)}), \quad (31)$$

причем в силу вогнутости $I(K)$ равенство в (31) возможно лишь при $\delta_s = 0$, что следует из (30). Поэтому $I(K^{(s+1)}) > I(K^{(s)})$ если $\delta_s > 0$, (т.е. если $K^{(s)} \notin \Omega$).

Представим теперь отображение $\mathcal{L}$ в виде композиции алгоритмических отображений: $\mathcal{L} = \mathcal{L}_3 \cdot \mathcal{L}_2 \cdot \mathcal{L}_1$, где

- 1) $\mathcal{L}_1 : \mathcal{K} \rightarrow \mathcal{F}^0 \times \mathcal{K}$, где $(F^{(s)}, K^{(s)}) = \mathcal{L}_1(K^{(s)})$;
- 2) $\mathcal{L}_2 : \mathcal{F}^0 \times \mathcal{K} \rightarrow \mathcal{K} \times \mathcal{K}$, где паре $(F^{(s)}, K^{(s)})$ ставится в соответствие $(K^{(s)}, \tilde{K}^{(s)})$ (т.е. $\mathcal{L}_2$ задает направление для максимизации $I(K)$, а $\tilde{K}^{(s)}$ определено в п.4) алгоритма 1);
- 3) $\mathcal{L}_3 : \mathcal{K} \times \mathcal{K} \rightarrow \mathcal{K}$ и переводит пару $(K^{(s)}, \tilde{K}^{(s)})$ в $K^{(s+1)}$ по правилу $K^{(s+1)} = (1 - \tau_s)K^{(s)} + \tau_s \tilde{K}^{(s)}$, а τ_s находится из решения (28) в п. 6) алгоритма 1.

В силу предыдущих рассмотрений заключаем, что каждое из отображений $\mathcal{L}_1$, $\mathcal{L}_2$ и $\mathcal{L}_3$ определяется с помощью решения соответствующей задачи оптимизации непрерывных функций на компактах. Поэтому соответствующие отображения замкнуты (теорема 3.7 из [16]), а $\mathcal{L}$ замкнуто на $\mathcal{K}$ как композиция замкнутых отображений (следствие 4.2.1 из [16]).

Итак, отображение $\mathcal{L}$ обладает следующими свойствами:

- а) $\mathcal{L} : \mathcal{K} \rightarrow \mathcal{K}$, где $\mathcal{K}$ — компакт;
- б) если $\mathcal{L}(K^{(s)}) = K^{(s)}$, то $K^{(s)} \in \Omega$;
если $K^{(s+1)} = \mathcal{L}(K^{(s)}) \neq K^{(s)}$, то $I(K^{(s)}) < I(K^{(s+1)})$;
- в) $\mathcal{L}$ замкнуто в каждой точке $K \in \mathcal{K}$.

Теперь из Теоремы А и леммы 11.2 из [16] заключаем, что для последовательности $\{K^{(s)}\}$ вида $K^{(s)} = \mathcal{L}(K^{(s-1)})$, $\forall K^{(0)} \in \mathcal{K}$ справедливо $d(K^{(s)}, \Omega) \rightarrow 0$ при $s \rightarrow \infty$, причем если $K^{(s)} = \mathcal{L}(K^{(s)})$, то $K^{(s)} \in \Omega$. Теорема 2 доказана.

Замечание. Утверждение теоремы 2 означает, что могут иметь место три следующих случая:

- 1) алгоритм сходится за конечное число шагов;

2) если подпоследовательность $K^{(s_m)}$ сходится, то $\lim_{m \rightarrow \infty} K^{(s_m)} \in \Omega$, причем по крайней мере одна такая подпоследовательность существует. Если же такая подпоследовательность единственна, то $\lim_{s \rightarrow \infty} K^{(s)} \in \Omega$.

3) если Ω — одноточечное множество, и $\hat{K} = \operatorname{argmax}_{K \in \mathcal{K}} I(K)$, то $K^{(s)} \rightarrow \hat{K}$ при $s \rightarrow \infty$.

5. МИНИМАКСНАЯ ФИЛЬТРАЦИЯ НЕОПРЕДЕЛЕННО-СТОХАСТИЧЕСКОЙ ПОСЛЕДОВАТЕЛЬНОСТИ

Рассмотрим модель наблюдения (1) в предположении, что помимо стохастических возмущений в ней присутствуют также и неопределенные ограниченные детерминированные возмущения:

$$\begin{cases} y_n = a_n y_{n-1} + u_n + b_n \varepsilon_n, & y_0 = \gamma, \\ z_n = c_n y_n + v_n + d_n w_n, & n = 1, \dots, N, \end{cases} \quad (32)$$

где (u_n, v_n) — неопределенные детерминированные векторы. Обозначим $u = \operatorname{col}(u_1, \dots, u_N)$, $v = \operatorname{col}(v_1, \dots, v_N)$, тогда эквивалентная статическая модель, аналогичная (10), для (32) имеет вид

$$\begin{cases} X_N = \Phi \rho + \Phi^{(1)} u \\ Z_N = \Psi \rho + \Psi^{(1)} u + \Psi^{(2)} v, \end{cases}$$

где $\Phi^{(1)}$, $\Psi^{(1)}$ и $\Psi^{(2)}$ зависят лишь от $\{a_n, b_n\}$ и определяются совершенно аналогично матрицам Φ , Ψ .

Предположим, что $u \in \mathcal{U}$, $v \in \mathcal{V}$, где $\mathcal{U}, \mathcal{V}$ — компактные центрально-симметричные относительно нуля множества векторов соответствующих размеров.

В данной задаче интегральный критерий качества для $\forall F \in \mathcal{F}$ имеет вид:

$$\begin{aligned} H(F, K, u, v) = & \operatorname{tr}[(\Phi - F\Psi)R_\rho(K)(\Phi - F\Psi)^* \\ & + (\Psi^{(1)} - F\Phi^{(1)})uu^*(\Psi^{(1)} - F\Phi^{(1)})^* + F\Psi^{(2)}vv^*(F\Psi^{(2)})^*], \end{aligned} \quad (33)$$

где $K \in \mathcal{K}$, $u \in \mathcal{U}$, $v \in \mathcal{V}$.

Используя теорему 2 из [12] можно показать, что для построения минимаксного оператора $\hat{F}$ вместо $H(F, K, u, v)$ из (33) можно использовать критерий

$$\begin{aligned} J(F, K, U, V) = & \operatorname{tr}[(\Phi - F\Psi)R_\rho(K)(\Phi - F\Psi)^* \\ & + (\Psi^{(1)} - F\Phi^{(1)})U(\Psi^{(1)} - F\Phi^{(1)})^* + F\Psi^{(2)}V(F\Psi^{(2)})^*], \end{aligned} \quad (34)$$

где $U \in \Sigma_u = \operatorname{co}[uu^* : u \in \mathcal{U}]$, $V \in \Sigma_v = \operatorname{co}[vv^* : v \in \mathcal{V}]$.

Пусть $F(K, U, V)$ — оператор, минимизирующий $J(\cdot)$:

$$F(K, U, V) \in \operatorname{argmin}_{F \in \mathcal{F}} J(F, K, U, V),$$

тогда двойственный функционал, аналогичный (8), имеет вид

$$I(K, U, V) = J(F(K, U, V), K, U, V).$$

Если множества $\mathcal{K}$, $\mathcal{U}$, и $\mathcal{V}$ компактны, то двойственная задача

$$(\hat{K}, \hat{U}, \hat{V}) \in \operatorname{argmax}_{K \in \mathcal{K}, U \in \Sigma_u, V \in \Sigma_v} I(K, U, V) \quad (35)$$

разрешима, а соответствующий оператор $\hat{F} = F(\hat{K}, \hat{U}, \hat{V})$ при определенных условиях будет минимаксным:

$$\hat{F} \in \operatorname{argmin}_{F \in \mathcal{F}} \sup_{K, u, v} H(F, K, u, v).$$

Заметим, что в общем случае матрицы $U \in \Sigma_u$ и $V \in \Sigma_v$, в отличие от $R_\rho(K)$, не являются блочно-диагональными, поэтому рекуррентное вычисление оператора $\hat{F}$ невозможно. Однако для множеств $\mathcal{U}$ и $\mathcal{V}$ специального вида соответствующая структура оценки может быть обеспечена.

Пусть $\mathcal{U}$ и $\mathcal{V}$ содержат все векторы u и v , удовлетворяющие ограничениям:

$$\sum_{n=1}^N u_n^* W_n u_n \leq r_u, \quad \sum_{n=1}^N v_n^* T_n v_n \leq r_v, \quad (36)$$

где $W_n > 0$, $T_n > 0$ — весовые матрицы, а $r_u \geq 0$ и $r_v \geq 0$ — заданные числа. Тогда для u и v справедливо

$$u^* W u \leq r_u, \quad v^* T v \leq r_v,$$

где $W = I_N \otimes W_n$, $T = I_N \otimes T_n$. В этом случае (см. теорему 1 и замечание 4 из [12]) при $\forall K \in \mathcal{K}$ матрицы $\hat{U}$ и $\hat{V}$ находятся аналитически и имеют следующий вид:

$$\hat{U} = r_u (I_N \otimes W_n^{-1}), \quad \hat{V} = r_v (I_N \otimes T_n^{-1}).$$

Таким образом, $\hat{U}$ и $\hat{V}$ являются блочно-диагональными.

Пусть теперь $\hat{K}$ является решением оставшейся части двойственной задачи (35), т.е.

$$\hat{K} \in \operatorname{argmax}_{K \in \mathcal{K}} I(K, \hat{U}, \hat{V}). \quad (37)$$

Подставляя найденные $\hat{K}$, $\hat{U}$ и $\hat{V}$ в (34), видим, что $J(F, \hat{K}, \hat{U}, \hat{V})$ совпадает с критерием оптимизации для чисто стохастической задачи, имеющей вид (32), где $\{u_n\}$ и $\{v_n\}$ — центрированные белые шумы с ковариациями $\operatorname{cov}(u_n, u_n) = r_u W_n^{-1}$ и $\operatorname{cov}(v_n, v_n) = r_v T_n^{-1}$, не зависящие друг от друга и от γ , $\{\varepsilon_n\}$ и $\{w_n\}$.

Таким образом, оператор $\hat{F} = F(\hat{K}, \hat{U}, \hat{V})$ имеет рекуррентную структуру, а соответствующая оценка $\{\hat{x}_n\}$ определяется соотношениями (19), приведенными в теореме 1, где

$$\hat{k}_n = \bar{P}_n c_n^* (c_n \bar{P}_n c_n^* + r_v T_n^{-1} + d_n \hat{S}_w d_n^*)^+, \quad (38)$$

$$\begin{cases} \bar{P}_n = a_n \hat{P}_{n-1} a_n^* + r_u W_n^{-1} + b_n \hat{S}_\varepsilon b_n^*, & \hat{P}_0 = \hat{S}_\gamma, \\ \hat{P}_n = (I - \hat{k}_n c_n) \bar{P}_n, \end{cases} \quad (39)$$

где $\hat{K} = \operatorname{diag}(\hat{S}_\gamma, \hat{S}_\varepsilon, \hat{S}_w)$.

Приведем окончательный вид утверждения о минимаксности оператора $\hat{F} = F(\hat{K}, \hat{U}, \hat{V})$, которое, фактически, следует из теоремы 1.

Теорема 3. Пусть выполнено (36), а $\mathcal{K}$ — выпуклый компакт, тогда

1) решение двойственной задачи (37) существует;

2) пусть выполнено любое из следующих условий:

а) множество $\mathcal{K}$ имеет максимальный элемент $K_{\max}$;

б) матрица $R_z(\hat{K}, \hat{U}, \hat{V}) = \Psi R_\rho(\hat{K})\Psi^* + \Psi^{(1)}\hat{U}(\Psi^{(1)})^* + \Psi^{(2)}\hat{V}(\Psi^{(2)})^*$ положительно определена,

тогда минимаксная оценка $\{\hat{x}_n\}$ для $\{x_n\}$ по интегральному критерию (33) на $\mathcal{K} \times \mathcal{U} \times \mathcal{V}$ определяется рекуррентными соотношениями (19), (38), (39), причем для $\forall P_\rho \in \mathcal{P}$

$$E_{P_\rho}[\sum_{n=1}^N \|\hat{x}_n - x_n\|^2] \leq \sum_{n=1}^N \text{tr}[Q_n \hat{P}_n] = \hat{H},$$

где $\hat{H}$ — наилучшее гарантированное значение критерия.

Для нахождения $\hat{K}$ может использоваться алгоритм 1 с очевидными изменениями. Достаточным условием сходимости итераций является $R_z(K, \hat{U}, \hat{V}) > 0$ для $\forall K \in \mathcal{K}$.

6. ПРИМЕРЫ

Пример 1. Пусть дана одномерная система

$$\begin{cases} y_n = a_n y_{n-1} + b_n \varepsilon_n \\ z_n = c_n y_n + d_n w_n \end{cases},$$

где a_n, b_n, c_n, d_n — числа, с ограничениями на дисперсию шумов вида $D_\varepsilon \leq D_\varepsilon^0$, $D_w \leq D_w^0$, где D_ε — дисперсия ε_n , D_w — дисперсия w_n , D_ε^0, D_w^0 — некоторые положительные числа. Распределение начального значения $y_0 = \gamma$ предполагается известным.

Очевидно, что в множестве допустимых дисперсий, заданном таким образом, существует максимальный элемент (пара D_ε^0, D_w^0), поэтому $\hat{D}_\varepsilon = D_\varepsilon^0$, $\hat{D}_w = D_w^0$.

Пример 2. Пусть матрицы системы имеют следующий вид:

$$a_n = \begin{pmatrix} 0.8 & 0.2 & -0.1 \\ 0.2 & 0.3 & 0.3 \\ 0.5 & 0.3 & 0.7 \end{pmatrix}, \quad b_n = \begin{pmatrix} 0.7 & 0.1 & 0.2 \\ 0.2 & 0.4 & 0.3 \\ 0.0 & 0.3 & 0.6 \end{pmatrix},$$

$$c_n = \begin{pmatrix} 0.5 & 0.4 & 0.1 \\ 0.3 & 0.5 & -0.1 \\ 0.1 & 0.4 & 0.6 \end{pmatrix}, \quad d_n = \begin{pmatrix} 0.7 & 0.2 & 0.1 \\ 0.2 & 0.1 & 0.7 \\ 0.3 & 0.0 & 0.6 \end{pmatrix}, \quad n = 1, \dots, N.$$

Множества допустимых ковариаций D_ε и D_w зададим следующим образом:

$$D_\varepsilon = \{S : \|S_\varepsilon - S\| < d\}, \quad D_w = \{S : S_w - d \cdot E \leq S \leq S_w + d \cdot E\}, \quad d = 0.1,$$

где $\|A\|^2 = (\sum_{i,j} A_{i,j}^2)$, E — матрица, состоящая из единиц и

$$S_\varepsilon = \begin{pmatrix} 0.5 & 0.0 & 0.0 \\ 0.0 & 0.7 & 0.0 \\ 0.0 & 0.0 & 0.6 \end{pmatrix}, \quad S_w = \begin{pmatrix} 0.9 & -0.4 & 0.2 \\ -0.4 & 0.7 & 0.1 \\ 0.2 & 0.1 & 0.6 \end{pmatrix}.$$

Распределение начального значения $y_0 = \gamma$ является стандартным гауссовским.

Возьмем в качестве начального приближения $S_\varepsilon^{(0)} = S_\varepsilon$, $S_w^{(0)} = S_w$ и воспользуемся тем, что для ограничений, представляющих собой шар, двойственная задача решается аналитически (см. пример 2.1 из [18]):

$$S_{\varepsilon}^{(1)} = S_{\varepsilon}^{(0)} + d \cdot \frac{\Lambda}{\|\Lambda\|},$$

где матрица Λ составлена из λ_i , соответствующих элементам матрицы S_{ε} . Решая задачу линейного программирования для оставшейся матрицы S_w , получим:

$$\hat{S}_{\varepsilon} = \begin{pmatrix} 0.597 & 0.012 & 0.009 \\ 0.012 & 0.715 & 0.001 \\ 0.009 & 0.001 & 0.601 \end{pmatrix}, \quad \hat{S}_w = \begin{pmatrix} 1.0 & -0.3 & 0.3 \\ -0.3 & 0.8 & 0.0 \\ 0.3 & 0.0 & 0.7 \end{pmatrix}.$$

Для данной системы решение было найдено на первом шаге итерационного процесса.

Пример 3. В этом примере рассматривается система, для которой итерационный процесс также сходится за конечное число шагов. Отличие от предыдущего примера состоит в том, что решение задачи находится не в угловой точке множества допустимых ковариаций, а в некоторой точке на границе.

Рассмотрим систему, заданную следующими матрицами:

$$a_n = \begin{pmatrix} 0.8 & 0.2 & -0.2 \\ -0.2 & 0.3 & 0.3 \\ 0.5 & -0.3 & 0.7 \end{pmatrix}, \quad b_n = \begin{pmatrix} 0.5 & 0.4 & 0.1 \\ 0.3 & 0.5 & -0.2 \\ -0.1 & 0.4 & 0.6 \end{pmatrix},$$

$$c_n = \begin{pmatrix} 0.7 & -0.1 & 0.2 \\ -0.2 & 0.4 & 0.3 \\ 0.1 & 0.3 & 0.6 \end{pmatrix}, \quad d_n = \begin{pmatrix} 0.7 & 0.2 & -0.1 \\ 0.2 & -0.1 & 0.7 \\ 0.3 & 0.1 & 0.6 \end{pmatrix}.$$

Ограничения на ковариационные матрицы зададим следующим образом:

$$D_{\varepsilon} = \{S : S_{\varepsilon} - d \cdot E \leq S \leq S_{\varepsilon} + d \cdot E\}, \quad D_w = \{S : S_w - d \cdot E \leq S \leq S_w + d \cdot E\}, \quad d = 0.1,$$

где

$$S_{\varepsilon} = \begin{pmatrix} 0.9 & -0.4 & 0.2 \\ -0.4 & 0.7 & 0.1 \\ 0.2 & 0.1 & 0.6 \end{pmatrix}, \quad S_w = \begin{pmatrix} 0.5 & 0.0 & 0.0 \\ 0.0 & 0.8 & 0.0 \\ 0.0 & 0.0 & 0.1 \end{pmatrix}.$$

Распределение начального значения y_0 здесь также является стандартным гауссовским.

Начальное приближение возьмем из предыдущего примера: $S_{\varepsilon}^{(0)} = S_{\varepsilon}$, $S_w^{(0)} = S_w$ (значение критерия для этих ковариационных матриц: $\frac{1}{N}J(F(S_{\varepsilon}^{(0)}, S_w^{(0)}), S_{\varepsilon}^{(0)}, S_w^{(0)}) = \frac{1}{N}I(S_{\varepsilon}^{(0)}, S_w^{(0)}) = 0.4377$, где N — время жизни системы).

На первом шаге были получены следующие матрицы:

$$S_{\varepsilon}^{(1)} = \begin{pmatrix} 1.0 & -0.3 & 0.1 \\ -0.3 & 0.8 & 0.0 \\ 0.1 & 0.0 & 0.7 \end{pmatrix}, \quad S_w^{(1)} = \begin{pmatrix} 0.6 & 0.1 & -0.1 \\ 0.1 & 0.9 & -0.1 \\ -0.1 & -0.1 & 0.2 \end{pmatrix}.$$

со значением критерия $\frac{1}{N}I(S_{\varepsilon}^{(1)}, S_w^{(1)}) = 0.6725$.

На следующем шаге:

$$\tilde{S}_{\varepsilon}^{(1)} = \begin{pmatrix} 1.0 & -0.3 & 0.1 \\ -0.3 & 0.8 & 0.0 \\ 0.1 & 0.0 & 0.7 \end{pmatrix}, \quad \tilde{S}_w^{(1)} = \begin{pmatrix} 0.6 & 0.1 & -0.1 \\ 0.1 & 0.9 & 0.1 \\ -0.1 & 0.1 & 0.2 \end{pmatrix}.$$

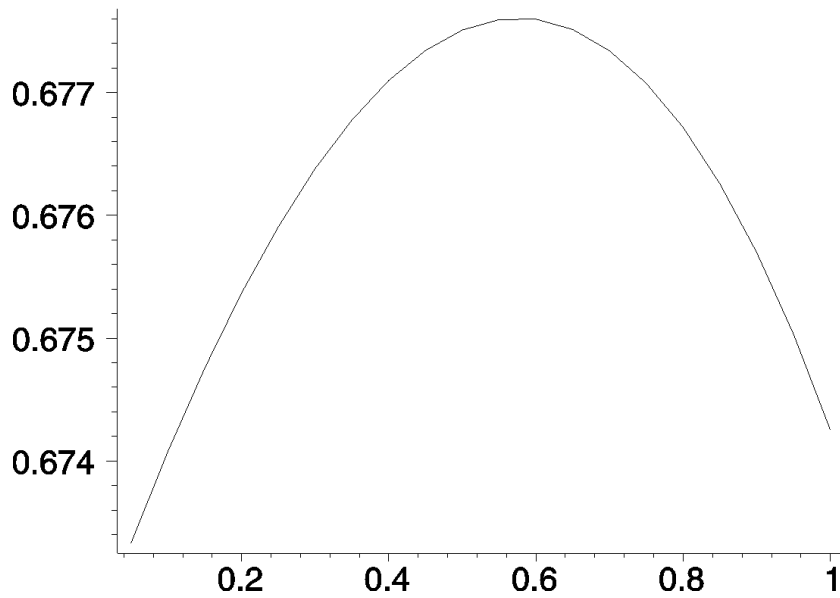


Рис. 1.

Для полученного приближения значение критерия равно $\frac{1}{N}I(\tilde{S}_\varepsilon^{(1)}, \tilde{S}_w^{(1)}) = 0.6742$.

На рис. 1 приведен график зависимости величины $\frac{1}{N}I(S_\varepsilon^{(1)}(\tau), S_w^{(1)}(\tau))$, где $S_\varepsilon^{(1)}(\tau) = (1 - \tau)S_\varepsilon^1 + \tau\tilde{S}_\varepsilon^{(1)}$ и $S_w^{(1)}(\tau) = (1 - \tau)S_w^1 + \tau\tilde{S}_w^{(1)}$ от $\tau \in [0, 1]$ (верхняя граница для τ равна единице, так как $\tilde{S}_\varepsilon^{(1)}$ и $\tilde{S}_w^{(1)}$ лежат в угловой точке области допустимых ковариаций).

Таким образом, $\tau_1 = 0.5781$ и

$$S_\varepsilon^{(2)} = \begin{pmatrix} 1.0 & -0.3 & 0.1 \\ -0.3 & 0.8 & 0.0 \\ 0.1 & 0.0 & 0.7 \end{pmatrix}, \quad S_w^{(2)} = \begin{pmatrix} 0.600 & 0.100 & -0.100 \\ 0.100 & 0.900 & 0.016 \\ -0.100 & 0.016 & 0.200 \end{pmatrix}.$$

Далее, на следующем шаге было выполнено условие окончания итерационного процесса ($\delta_2 = 0$), и, окончательно, решение задачи: $\hat{S}_\varepsilon = S_\varepsilon^{(2)}$ и $\hat{S}_w = S_w^{(2)}$.

Пример 4. В этом примере рассматривается система, для которой итерационный процесс не сошелся за конечное число итераций, т.е. было найдено некоторое приближение для решения двойственной задачи.

$$a_n = \begin{pmatrix} 0.8 & 0.4 & 0.1 \\ 0.4 & 0.1 & -0.1 \\ -0.1 & 0.4 & 0.2 \end{pmatrix}, \quad b_n = \begin{pmatrix} -0.7 & 0.2 & -0.1 \\ 0.2 & 0.6 & 0.3 \\ 0.5 & 0.3 & 0.4 \end{pmatrix},$$

$$c_n = \begin{pmatrix} 0.2 & -0.2 & 0.1 \\ 0.2 & 0.8 & 0.5 \\ 0.2 & 0.0 & 0.4 \end{pmatrix}, \quad d_n = \begin{pmatrix} 0.8 & 0.2 & -0.2 \\ 0.2 & 0.3 & 0.3 \\ 0.5 & 0.3 & 0.7 \end{pmatrix}.$$

Ограничения, начальное условие и начальное приближение взяты из предыдущего примера. Значение критерия для начального приближения равно $\frac{1}{N}I(S_\varepsilon^{(0)}, S_w^{(0)}) = 1.2701$.

В результате после $s = 20$ итераций было достигнуто значение $\delta_s = 0.193 \cdot 10^{-4}$, соответствующие значения матриц

$$\hat{S}_\varepsilon = \begin{pmatrix} 1.000 & -0.300 & 0.300 \\ -0.300 & 0.800 & 0.153 \\ 0.300 & 0.000 & 0.700 \end{pmatrix}, \quad \hat{S}_w = \begin{pmatrix} 0.600 & 0.024 & -0.100 \\ 0.024 & 0.900 & -0.100 \\ -0.100 & -0.100 & 0.200 \end{pmatrix},$$

а значение критерия $\frac{1}{N}I(\hat{S}_\varepsilon, \hat{S}_w) = 1.5815$.

Пример 5. Пусть дана одномерная система с неопределенными неслучайными параметрами

$$\begin{cases} y_n = a_n y_{n-1} + b_n \varepsilon_n + u_n \\ z_n = c_n y_n + d_n w_n + v_n \end{cases},$$

с ограничениями на дисперсию шумов вида $D_\varepsilon \leq D_\varepsilon^0$, $D_w \leq D_w^0$ (D_ε – дисперсия ε_n , D_w – дисперсия w_n) и начальным условием $y_0 = 0$.

Ограничения на неслучайные параметры:

$$\sum_{n=1}^N u_n^2 \leq r_u, \quad \sum_{n=1}^N v_n^2 \leq r_v,$$

где $r_u > 0$, $r_v > 0$. Таким образом, $W_N = T_N = 1$. Очевидно также, что $\hat{S}_\varepsilon = D_\varepsilon^0$, $\hat{S}_w = D_w^0$ (см. пример 1).

Поэтому минимаксный фильтр имеет вид (19), где соотношения (38),(39) принимают следующую форму:

$$\hat{k}_n = \frac{\bar{P}_n c_n}{c_n^2 \bar{P}_n + r_v + d_n^2 D_w^0},$$

$$\begin{cases} \bar{P}_n = a_n^2 \hat{P}_{n-1} + r_u + b_n^2 D_\varepsilon^0, & \hat{P}_0 = 0, \\ \hat{P}_n = (1 - \hat{k}_n c_n) \bar{P}_n. \end{cases}$$

СПИСОК ЛИТЕРАТУРЫ

1. Острем К.Ю. *Введение в стохастическую теорию управления*. М.: Мир, 1973. (Åström K.J. *Introduction to Stochastic Control Theory*. New York: Academic Press, 1970.)
2. Григорьев Ф.Н., Кузнецов Н.А., Серебровский А.П. *Управление наблюдениями в автоматических системах*. М.: Наука, 1986.
3. Morris J.M. The Kalman Filter: A Robust Estimator for Some Classes of Linear Quadratic Problems. *IEEE Trans. Inform. Theory*, 1976, vol. IT-22, pp. 526–534.
4. Martin C.J., Mintz M. Robust Filtering and Prediction for Linear Systems with Uncertain Dynamics: A Game-Theoretic Approach. *IEEE Trans. Autom. Control*, 1983, vol. AC-28, pp. 888–896.
5. Кассам С.А., Пур Г.В. Робастные методы обработки сигналов. *ТИИЭР*, 1985, том 73, № 3, стр. 54–110.
6. Куржанский А.Б. *Управление и оценивание в условиях неопределенности*. М.: Наука, 1977.
7. Кац И.Я. *Оценки состояния управляемых многошаговых систем в условиях статистической неопределенности*. Свердловск: Изд-во Уральского гос. университета, 1990.
8. Голубев Г.А., Муравлев О.В., Писарев В.Ф. Линейная рекуррентная фильтрация динамических процессов с дискретным временем при частичной информации о возмущающих процессах. *Автоматика и телемеханика*, 1989, № 12, стр. 49–59.
9. Куркин О.М., Коробочкин Ю.Б., Шаталов С.А. *Минимаксная обработка информации*. М.: Энергоатомиздат, 1990.

10. Matasov A.I. *Estimators for Uncertain Dynamic Systems*. Dordrecht: Kluwer Academic Publ., 1998.
11. Pankov A.R., Borisov A.V. Optimal Filtering in Stochastic Discrete-Time Systems with Unknown Inputs. *IEEE Trans. Autom. Control*, 1994, vol. AC-39, pp. 2461–2464.
12. Панков А.Р., Семенихин К.В. Минимаксная идентификация неопределенно-стохастической линейной модели. *Автоматика и телемеханика*, 1998, № 11, стр. 158–171.
13. Панков А.Р., Семенихин К.В. Методы параметрической идентификации многомерных линейных моделей в условиях априорной неопределенности. *Автоматика и телемеханика*, 2000, № 5, стр. 76–92.
14. Боровков А.А. *Математическая статистика. Дополнительные главы*. М.: Наука, 1984.
15. Рокафеллар Р.Т. *Выпуклый анализ*. М.: Мир, 1973. (Rockafellar R.T. *Convex Analysis*. Princeton: Princeton Univ. Press, 1970.)
16. Зангвилл У. *Нелинейное программирование. Единый подход*. М.: Сов. радио, 1973. (Zangwill W.I. *Nonlinear Programming: A Unified Approach*. Englewood Cliffs: Prentice-Hall, 1969.)
17. Пшеничный Б.Н. *Необходимые условия экстремума*. М.: Наука, 1982.
18. Балакришнан А. *Введение в теорию оптимизации в гильбертовом пространстве*. М.: Мир, 1974. (Balakrishnan A.V. *Introduction to Optimization Theory in a Hilbert Space*. New York: Springer-Verlag, 1971.)

Статью представил к публикации член редколлегии Б.М.Миллер