

# СРАВНЕНИЕ РЕКУРРЕНТНЫХ И НЕРЕКУРРЕНТНЫХ СХЕМ ПОСТРОЕНИЯ АЛГОРИТМОВ, ОСНОВАННЫХ НА ЛИНЕАРИЗАЦИИ, ПРИ РЕШЕНИИ ОДНОГО КЛАССА ЗАДАЧ НЕЛИНЕЙНОГО ОЦЕНИВАНИЯ

**О.А. Степанов**

*АО «Концерн «ЦНИИ «Электроприбор»*  
Россия, 197046, Санкт-Петербург, Малая Посадская ул., 30  
*Университет ИТМО*  
Россия, 197101, Санкт-Петербург, Кронверкский пр., 49  
E-mail: soalax@mail.ru

**А.М. Исаев**

*АО «Концерн «ЦНИИ «Электроприбор»*  
Россия, 197046, Санкт-Петербург, Малая Посадская ул., 30  
*Университет ИТМО*  
Россия, 197101, Санкт-Петербург, Кронверкский пр., 49  
E-mail: itmo\_student@mail.ru

**Ю.А. Литвиненко**

*АО «Концерн «ЦНИИ «Электроприбор»*  
Россия, 197046, Санкт-Петербург, Малая Посадская ул., 30  
*Университет ИТМО*  
Россия, 197101, Санкт-Петербург, Кронверкский пр., 49  
E-mail: ya\_litvinenko@mail.ru

**Ключевые слова:** нелинейные задачи оценивания, рекуррентные алгоритмы, линейаризация, нерекуррентные алгоритмы, сопоставление, погрешность оценивания.

**Аннотация:** Применительно к одному классу задач нелинейного оценивания сопоставляются две схемы построения субоптимальных алгоритмов оценивания, синтезируемых с использованием байесовского подхода и основанных на линейаризации функций, описывающих поведение оцениваемого вектора состояний и модель измерений. Одна из них - традиционная, в которой оценка формируется рекуррентно относительно измерений, а другая - нерекуррентная, предполагающая использование всего имеющегося набора измерений одновременно. Обсуждаются достоинства и недостатки анализируемых алгоритмов и приводится иллюстрирующий их пример.

## 1. Введение

В последнее время в литературе, посвященной комплексной обработке измерительной информации, значительное внимание уделяется алгоритмам, основанным на использовании методов оптимизации на фактор-графах, называемых также алгоритмами фактор-граф оптимизации (Factor-graph optimization – FGO) [1-3]. Их суть заключается в том, что задача оценивания сводится к минимизации

функционала, представляющего собой квадратичную форму. Эти алгоритмы, как правило, сопоставляются с рекуррентным обобщенным фильтром Калмана (ОФК) и при этом обычно отмечается факт повышения точности оценивания, достигаемой при использовании FGO, по сравнению с ОФК.

Известно, что в рамках байесовского подхода может быть найдена оптимальная оценка, минимизирующая среднеквадратическую погрешность оценивания и представляющая собой математическое ожидание, соответствующее апостериорной плотности для вектора оцениваемых параметров. В связи с тем, что оба упомянутых алгоритма являются субоптимальными, их сопоставление желательно проводить с учетом возможности вычисления в рамках байесовского подхода потенциальной точности, достигаемой с использованием оптимального алгоритма. При этом важно учитывать следующие обстоятельства. Во-первых, FGO, также как и ОФК, основан на линеаризации, а во-вторых, в отличие от рекуррентного ОФК, при получении оценки FGO используется весь имеющийся набор измерений, т.е. алгоритм, по сути, является нереккуррентным. Таким образом, для понимания причин отличия в точности двух субоптимальных вариантов алгоритмов, основанных на линеаризации, целесообразно сопоставить две схемы их построения – рекуррентную и нереккуррентную. Вопросу сопоставления в рамках байесовского подхода этих схем применительно к одному классу задач нелинейного оценивания и посвящен предлагаемый доклад.

## 2. Постановка исследуемой задачи и описание алгоритмов

Рассматривается задача нелинейной дискретной фильтрации  $n$ -мерного случайного вектора, описываемого с помощью формирующего фильтра

$$(1) \quad x_k = f_k(x_{k-1}) + \Gamma_k w_k + u_k$$

по  $m$ -мерным измерениям следующего вида:

$$(2) \quad y_k = h_k(x_k) + v_k.$$

В этих соотношениях предполагается, что  $k$  – индекс дискретного времени;  $u_k = (u_{1k}, \dots, u_{nk})^T$  –  $n$ -мерный вектор известных входных сигналов (при использовании двойной индексации в подстрочных индексах здесь и далее первый индекс соответствует номеру компоненты, второй – времени);  $x_0$  –  $n$ -мерный случайный гауссовский вектор с функцией плотности распределения вероятности (ф.п.р.в.)  $p(x_0) = N(x_0; \bar{x}_0, P_0)$ ;  $w_k$  –  $n_w$ -мерный центрированный дискретный гауссовский белый шум, не зависящий от  $x_0$ , с известной матрицей ковариаций  $Q_k$  размерности  $n_w \times n_w$ ;  $\Gamma_k$  – матрица размерности  $n \times n_w$ ;  $v_k$  –  $m$ -мерный центрированный дискретный гауссовский белый шум, не зависящий от  $x_0$  и  $w_k$ , с ковариационной матрицей  $R_k$ ;  $f_k(\bullet)$ ,  $h_k(\bullet)$  – известные нелинейные  $n$ - и  $m$ -мерные вектор-функции, описывающие динамику для вектора состояний и измерения.

Суть решения задачи фильтрации в рамках байесовского подхода заключается в получении оптимальной в среднеквадратическом смысле оценки вектора состояний  $\hat{x}_k^{opt}(Y_k)$  по измерениям  $Y_k = (y_1^T, \dots, y_k^T)^T$  и соответствующей ей условной матрицы ковариаций погрешностей оценивания, характеризующей ее текущую (соответствующую конкретному набору измерений  $Y_k$ ) точность.

Известно, что оптимальная в среднеквадратическом смысле оценка  $\hat{x}_k^{opt}(Y_k)$  определяются следующим образом [4]:

$$(3) \quad \hat{x}_k^{opt}(Y_k) = \int x_k p(x_k/Y_k) dx_k,$$

где  $p(x_k/Y_k)$  – апостериорная (условная) к измерениям ф.п.р.в. (далее для простоты – «плотность»). Важно подчеркнуть, что оптимальная оценка минимизирует условную, и безусловную матрицы ковариаций погрешностей оценивания [4]. При этом элементы

на главной диагонали безусловной матрицы ковариаций представляют собой дисперсии погрешностей оптимальных оценок и характеризуют потенциальную точность оценивания соответствующих компонент вектора состояния [4]. В (3) предполагается, что  $p(x_k/Y_k)$  формируется с использованием известных рекуррентных соотношений для апостериорной ф.п.р.в. [4].

Наряду с рекуррентными алгоритмами искомая оценка вектора состояния в произвольный момент времени может быть получена путем ее нахождения в составе оценки для вектора  $X_k = (x_1^T, \dots, x_k^T)^T$ , включающего значения  $x_i$  для всех моментов времени  $i = \overline{1, k}$ , по векторному измерению  $Y_k = (y_1^T, \dots, y_k^T)^T$ . В этом случае задача формулируется так: оценить вектор  $X_k$  с априорной ф.п.р.в.  $p(X_k)$  по вектору измерений размерности  $m \times k$

$$(4) \quad Y_k = H_k(X_k) + V_k,$$

в которых  $H_k(X_k) = (h_1^T(x_1), \dots, h_k^T(x_k))^T$ ,  $V_k = (v_1^T, \dots, v_k^T)^T$  – составной вектор ошибок измерений с блочно-диагональной матрицей ковариаций  $\mathbf{R}_k$  с матрицами  $R_i$ ,  $i = \overline{1, k}$  на главной диагонали. Для получения искомой оценки в этом случае надо располагать апостериорной ф.п.р.в.  $p(X_k/Y_k)$ , задаваемой в виде

$$(5) \quad p(X_k/Y_k) = c p(X_k) p(Y_k/X_k).$$

Нетрудно убедиться, что в рассматриваемой задаче

$$(6) \quad p(X_k) p(Y_k/X_k) = c \exp \left\{ \frac{1}{2} J(X_k) \right\},$$

где  $c$  – нормирующий множитель,

$$(7) \quad -J(X_k) = x_0^T P_0^{-1} x_0 + \sum_{i=1}^k (x_i - f_i(x_{i-1}) - u_i)^T (\Gamma_i^T Q_i \Gamma_i)^{-1} (x_i - f_i(x_{i-1}) - u_i) + \sum_{i=1}^k (y_i - h_i(x_i))^T R_i^{-1} (y_i - h_i(x_i)),$$

$$(8) \quad \hat{x}_k^{opt}(Y_k) = \int x_k p(X_k/Y_k) dX_k.$$

Отметим, что  $J(X_k)$  совпадает с критерием, минимизируемым при построении алгоритмов FGO.

Ясно, что, если располагать точным описанием апостериорной ф.п.р.в., то оценки, вычисляемые согласно формулам (3) и (8), будут между собой совпадать. Вместе с тем, следует иметь в виду два обстоятельства. С одной стороны, при вычислении этих оценок на практике используются различного рода субоптимальные алгоритмы. В частности, алгоритмы, основанные на линеаризации и гауссовской аппроксимации апостериорной ф.п.р.в. на каждом шаге, такие, например, как ОФК. А с другой - тот факт, что вид плотности с увеличением числа измерений меняется. При решении прикладных задач нередко приходится сталкиваться с ситуацией, при которой уточнение искомого вектора состояний происходит после накопления определенного количества измерений. В этом случае при малом числе измерений апостериорная плотность, как правило, имеет негауссовский и многоэкстремальный характер. При увеличении числа измерений эта плотность становится одноэкстремальной и в ряде случаев гауссовской. Отмеченные особенности и создают предпосылки для различной эффективности рекуррентных и нерекуррентных алгоритмов, основанных на линеаризации и гауссовской аппроксимации апостериорной ф.п.р.в. при решении такого класса задач. Когда используется весь набор измерений, при получении оценок в каждый момент времени такой вид плотности будет иметь место для всех компонент вектора состояний. Именно это и происходит при использовании нерекуррентной схемы [4-6]. Проиллюстрируем сказанное на простом примере.

### 3. Пример

Рассмотрим частный случай сформулированной задачи фильтрации последовательности  $x_k$ , описываемой с помощью линейного формирующего фильтра:

$$(9) \quad x_k = \Phi x_{k-1} + \Gamma w_k + u,$$

где  $\Phi = e^{-\alpha \Delta t}$ ,  $\Gamma = \frac{\sqrt{2\sigma^2 \alpha}}{\alpha} (1 - e^{-\alpha \Delta t})$ ,  $\sigma^2$  – дисперсия процесса, а  $\alpha$  – величина, обратная интервалу корреляции  $\tau_k$ .

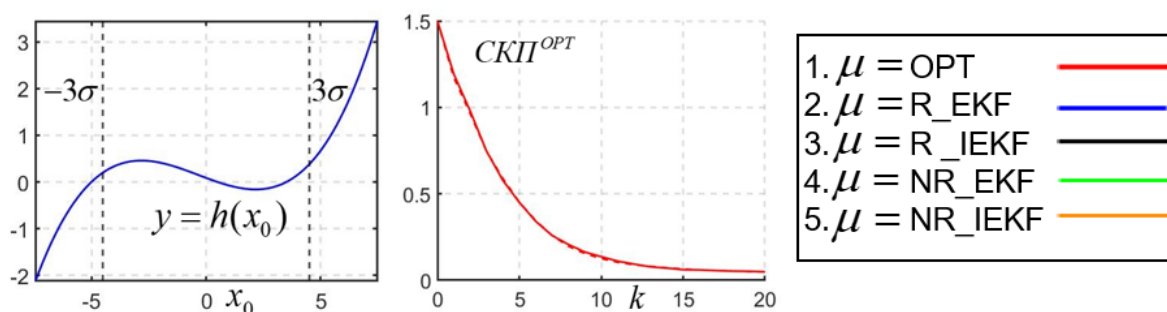
Нелинейные измерения (2) имеют вид:

$$(10) \quad y_k = h_1 + h_2 x_k + h_3 x_k^2 + h_4 x_k^3 + v_k,$$

где  $h_1, h_2, h_3, h_4$  – известные величины.

Моделирование проводилось со следующими параметрами:  $\sigma = 1.5, \alpha = 0.1, r = 0.1, h_1 = 0.0875, h_2 = -0.1825, h_3 = 0.01, h_4 = 0.01, u = 1, L = 10000$ , где  $L$  – количество реализаций в методе статистических испытаний, интервал дискретизации  $\Delta t = 1$  с, время моделирования  $T = 20$  с. Далее рассматриваются следующие алгоритмы: фильтр частиц (ОПТ), классический рекуррентный и нерекуррентный ОФК (R\_EKF), (NR\_EKF) и итерационные алгоритмы, построенные с использованием рекуррентной и нерекуррентной схем. Сопоставление этих алгоритмов проводилось с точки зрения точности, состоятельности и вычислительной сложности с использованием методики, описанной в [7]. При этом в качестве базового алгоритма, относительно которого проводилось сопоставление по точности, использовался фильтр частиц [8,9], направленный на получение оптимальной оценки (3), с количеством частиц  $N=20000$ , а число итераций в итерационных фильтрах принималось равным 5.

На рис. 1, слева синей сплошной линией показан вид нелинейной функции (12) и чёрными пунктирными линиями обозначена априорная область  $\pm 3\sigma$ , посередине представлен график действительной среднеквадратической погрешности оценивания (СКП), соответствующей фильтру частиц, а справа расшифрованы цвета и обозначения, которые далее будут соответствовать анализируемым алгоритмам.



**Рис. 1.** Вид нелинейной функции измерений (12),  $СКП^{OPT}$  оценивания и анализируемые алгоритмы.

Из рис. 1 с очевидностью следует, что для рассматриваемого примера вид апостериорной плотности при малом числе измерений может быть многоэкстремальным. При увеличении же числа измерений погрешность оптимального оценивания снижается и плотность становится одноэкстремальной. На рис. 2, приводятся результаты расчёта коэффициентов [7], характеризующих точность  $\xi_k^\mu$  и состоятельность  $\zeta_k^\mu$ , а также вероятность попадания погрешностей оценивания в интервал трёх расчётных сигм.

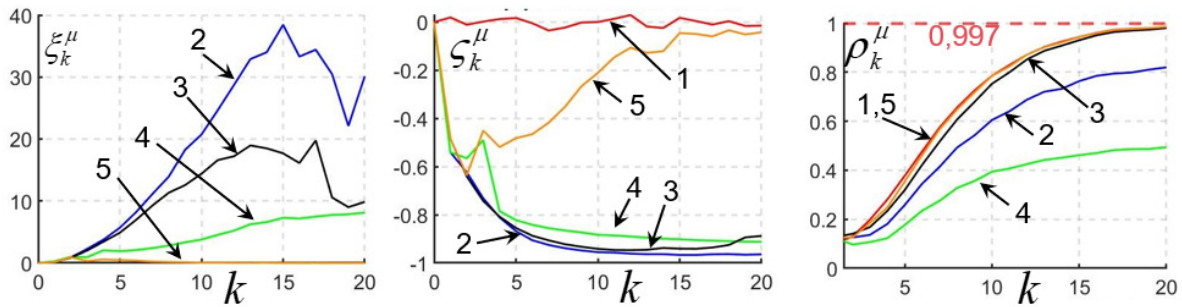


Рис. 2. Результаты расчёта коэффициентов, характеризующих точность и состоятельность алгоритмов, а также вероятности попадания в заданный интервал.

После обработки 13-го измерения, коэффициенты  $\xi_k^{NR\_IEKF}$  и  $\zeta_k^{NR\_IEKF}$  не превышают по модулю значения 0.1, то есть алгоритм NR\_IEKF удовлетворяет 10% уровню точности и состоятельности, тогда как остальные алгоритмы, использующие линейризацию, существенно ему проигрывают. Следует отметить, что вероятность попадания погрешностей оценивания в интервал трёх расчётных сигм в начальные моменты времени существенно отличается от значения  $P^G=0,997$ , соответствующего гауссовскому распределению, что подтверждает факт отличия вида апостериорной плотности в начальный момент времени от гауссовской. Из представленных результатов следует, что при использовании нерекуррентной схемы, итерационный фильтр Калмана при увеличении числа используемых измерений обеспечивает точность, близкую к потенциальной, хотя вырабатываемая в этом алгоритме расчетная характеристика точности и не всегда совпадает с ее действительным значением, особенно на начальном этапе оценивания.

Сопоставление вычислительной сложности проводилось относительно наиболее простого – R\_EKF. В результате расчетов коэффициента вычислительной сложности  $T^\mu$  было показано, что ОПТ сложнее, чем R\_EKF примерно в 200 раз, тогда как NR\_IEKF лишь в 8 раз. Очевидно, вычислительная сложность алгоритмов, построенных с использованием нерекуррентной схемы, существенно возрастает при увеличении числа измерений. Вопрос сокращения объема вычислений, в том числе и за счет привлечения математического аппарата оптимизации на графах, заслуживает отдельного обсуждения.

## 4. Заключение

В рамках байесовского подхода проведено сопоставление основанных на линейризации алгоритмов, построенных с использованием рекуррентных и нерекуррентных схем и применяемых для решения задач, в которых уточнение искомого вектора состояний происходит после накопления определенного количества измерений. Показано, что итерационная нерекуррентная схема обработки при увеличении числа измерений обеспечивает точность, близкую к потенциальной. Таким образом, можно констатировать, что именно нерекуррентная схема и создает при решении такого класса задач нелинейного оценивания необходимые предпосылки для повышения точности субоптимальных алгоритмов, основанных на линейризации, которое и отмечается при сравнении ОФК и FGO.

Выявленные особенности создают хорошую основу для понимания возможных преимуществ при построении эффективных алгоритмов оценивания, основанных на комбинированном применении рекуррентных и нерекуррентных схем, реализующих итерационные ФК, и целесообразность привлечения методов FGO.

Исследование выполнено за счет гранта Российского научного фонда № 23-19-00626, <https://rscf.ru/project/23-19-00626>.

## Список литературы

1. Dunik J., Biswas S.K., Dempster A.G. Pany T., Closas P. State Estimation Methods in Navigation: Overview and Application // IEEE Aerospace and Electronic Systems Magazine. 2020. Vol. 35, No. 12. P. 16-31. doi: 10.1109/MAES.2020.3002001.
2. Weisong Wen, Tim Pfeifer, Xiwei Bai Li-Ta Hsu1. Factor graph optimization for GNSS/INS integration: A comparison with the extended Kalman filter // NAVIGATION. 2021. No. 68. P. 315-331.
3. Jun Dai, Songlin Liu, Xiangyang Hao, Zongbin Ren, Xiao Yang. UAV Localization Algorithm Based on Factor Graph Optimization in Complex Scenes // Sensors. 2022. Vol. 22.
4. Степанов О.А. Применение теории нелинейной фильтрации в задачах обработки навигационной информации. С.Пб: ГНЦ РФ ЦНИИ «Электроприбор», 1998. 370 с.
5. Васильев В.А., Литвиненко Ю.А., Степанов О.А., Торопов А.Б. Сравнение рекуррентных и нерекуррентных алгоритмов фильтрации при решении задачи навигации по геофизическим полям // Материалы. XIV Всероссийская мультиконференция по проблемам управления МКПУ-2021: в 4 т.. Т. 1. Ростов-на-Дону, 2021. С. 125-127.
6. Торопов А.Б. Алгоритмы фильтрации в задачах коррекции показаний морской навигационной системы с использованием нелинейных измерений. Диссертация на соискание учёной степени кандидата технических наук. Санкт-Петербург, 2012. 146 с.
7. Степанов О.А., Исаев А.М. Методика сравнительного анализа рекуррентных алгоритмов нелинейных фильтрации в задачах обработки навигационной информации на основе предсказательного моделирования. // Гироскопия и навигация. 2023. Т. 31. № 3 (122). С. 48-65. EDN: MVWKG6.
8. Chen Z. Bayesian Filtering: From Kalman Filters to Particle Filters, and Beyond. Adaptive Systems Lab., McMasterUniv., Hamilton, Canada, 2003.
9. Doucet A., Freitas N., Gordon N. Sequential Monte Carlo Methods in Practice. New York, NY, Springer, 2001, 590 p.