

Работа 3 Исследование абсолютно-оптимальных рекуррентных алгоритмов

Цель: исследование эффективности использования абсолютно-оптимальных рекуррентных алгоритмов оценивания параметров линейных регрессионных объектов при различных статистических свойствах шумов измерений по сравнению с оптимальными алгоритмами; точность и скорость сходимости оценок к истинным значениям параметров в зависимости от статистических характеристик шумов измерений.

3.1. Основные теоретические сведения

Как было отмечено ранее, использование рекуррентных алгоритмов идентификации параметров систем (идентификация в реальном масштабе времени) позволяет существенно снизить объем хранимой информации и время вычислений оцениваемых параметров.

Функциональная структурная схема, соответствующая рекуррентному процессу идентификации, изображена на рис. 3.1.

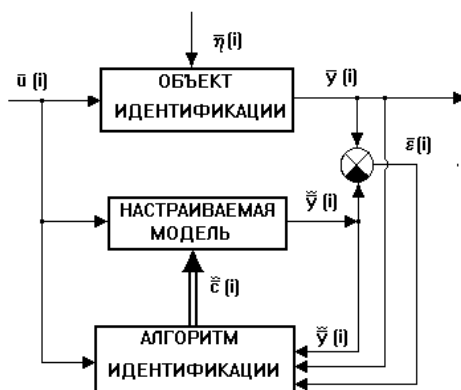


Рис. 3.1. Функциональная схема рекуррентного процесса идентификации

В этом случае информация, снимаемая в каждый момент времени с входов $\bar{u}(i)$ и выходов $\bar{y}(i)$ объекта, а также невязка

$\bar{e}(i) = \bar{y}(i) - \tilde{y}(i)$ между выходом объекта $\bar{y}(i)$ и выходом модели $\tilde{y}(i)$, используется для настройки параметров модели.

Основная рекуррентная последовательность алгоритмов построена по принципу коррекции параметров модели на основе новой полученной информации:

$$\hat{c}_{i+1} = \hat{c}_i - \Gamma_{i+1} \left(\Phi(F(\psi(u(i+1), \tilde{c}) - y(i+1))) \Big|_{\tilde{c} = \hat{c}_i} \right). \quad (3.1)$$

В предыдущей лабораторной работе изучена рекуррентная форма метода наименьших квадратов, которая была получена простым преобразованием обычной формы метода наименьших квадратов для линейных регрессионных моделей. Однако на практике часто возникает необходимость использовать рекуррентные соотношения для оценки параметров нелинейных объектов при произвольной функции потерь и различных статистических характеристиках шумов, действующих на объект идентификации.

Как известно [1], в основе всех рекуррентных методов идентификации лежит метод стохастической аппроксимации, разработанный в 1951 г Робинсоном и Монро [2] для решения скалярных стохастических уравнений вида

$$\psi(c, \eta(i)) - y(i) = 0, \quad i = 1, 2, \dots, \quad (3.2)$$

где c — искомый параметр; $\eta(i)$ — случайная последовательность с характеристиками

$$M\{\eta(i)\} = 0, \quad \text{cov}\{\eta(i), \eta(j)\} = \sigma_\eta^2 \delta_k(i - j).$$

Согласно этому методу решение уравнения (3.1) формируется в виде рекуррентной последовательности:

$$\hat{c}_{i+1} = \hat{c}_i - \gamma_{i+1} (\psi(\hat{c}_i) - y(i+1)), \quad (3.3)$$

причем оценка корня стохастического уравнения будет состоятельна, если на коэффициент γ_i наложены три условия:

1) $\gamma_i > 0$ при монотонно возрастающей $\psi(\hat{c})$, $\gamma_i < 0$ при монотонно убывающей $\psi(\hat{c})$;

$$2) \sum_{k=1}^{\infty} \gamma_k^2 < \infty;$$

$$3) \sum_{k=1}^{\infty} \gamma_k = \infty.$$

Возможным вариантом задания последовательности γ_i , удовлетворяющей приведенным выше условиям, является последовательность вида

$$\gamma_i = \frac{b}{i}. \quad (3.4)$$

При решении задач идентификации, по сути дела, приходится решать многомерное стохастическое уравнение

$$y(i) - \psi(\bar{u}(i), \bar{c}) + \eta(i) = 0, \quad (3.5)$$

$$M\{\eta(i)\} = 0, \quad \text{cov}\{\eta(i), \eta(j)\} = \sigma_\eta^2 \delta_k(i - j),$$

очевидно, найти однозначное решение многомерного параметра по одному уравнению не представляется возможным.

Рассмотрим некоторую четную функцию $F(\psi(\bar{u}(i), \tilde{c}) + \eta(i) - y(i))$ — аналог функции потерь, которая достигает минимума при значениях $\tilde{c} = \bar{c}$.

Тогда \bar{c} является корнем стохастического векторного уравнения:

$$\left. \frac{\partial F(\psi(\bar{u}(i), \tilde{c}) + \eta(i) - y(i))}{\partial \tilde{c}} \right|_{\tilde{c} = \bar{c}} = \bar{0}. \quad (3.6)$$

Принципиально векторное уравнение (3.6) ничем не отличается от стохастического скалярного уравнения (3.2), причем, по крайней мере для линейных систем, уравнение (3.6) — монотонно возрастающее относительно параметра \tilde{c} .

По аналогии с решением скалярных уравнений составим рекуррентную последовательность для решения векторного уравнения (3.6):

$$\hat{c}_{i+1} = \hat{c}_i - \Gamma_{i+1} \left(\left. \frac{\partial F(\psi(u(i+1), \tilde{c}) - y(i+1))}{\partial \tilde{c}} \right|_{\tilde{c} = \hat{c}_i} \right). \quad (3.7)$$

Так как выражение $\psi(u(i+1), \tilde{c}) - y(i+1)$ представляет собой невязку $\varepsilon(i+1, \tilde{c})$ между выходом модели и объекта, то можно записать:

$$\hat{c}_{i+1} = \hat{c}_i - \Gamma_{i+1} \left(\frac{\partial F(\varepsilon(i+1, \tilde{c}))}{\partial \tilde{c}} \bigg|_{\tilde{c} = \hat{c}_i} \right), \quad (3.8)$$

где Γ_{i+1} — матрица коэффициентов усиления, диагональные элементы которой удовлетворяют условиям:

- 1) $\gamma_{j,i+1} > 0, j = \overline{1, m};$
- 2) $\sum_{i=1}^{\infty} \gamma_{j,i+1}^2 < \infty, j = \overline{1, m};$
- 3) $\sum_{i=1}^{\infty} \gamma_{j,i+1} = \infty, j = \overline{1, m}.$

Одним из возможных видов матричной последовательности Γ_{i+1} , удовлетворяющей приведенным выше условиям, является матричный ряд

$$\Gamma_{i+1} = \frac{B}{i+1}, \quad (3.9)$$

где B — некоторая положительно определенная матрица.

3.2. Оптимальные рекуррентные алгоритмы

Очевидно, различные способы задания матрицы Γ_{i+1} и различные функции потерь $F(\varepsilon(i, \tilde{c}))$ будут давать различные скорости сходимости оценок к истинным значениям параметров.

В качестве меры скорости сходимости алгоритмов в теории оценивания принято использовать асимптотическую матрицу ковариаций ошибок оценки (АМКО) [1]:

$$V = \lim_{i \rightarrow \infty} i \operatorname{var} \{ \hat{c}_i - \bar{c} \} =$$

$$= \lim_{i \rightarrow \infty} iM \left\{ (\hat{\bar{c}}_i - \bar{c})(\hat{\bar{c}}_i - \bar{c})^T \right\}, \quad (3.10)$$

чем меньше АМКО, тем быстрее скорость сходимости алгоритма. Так как $\hat{\bar{c}}_i$ функционально зависит от вида функции потерь $F(\varepsilon)$ и от вида матрицы Γ_i , то и АМКО является функционалом $F(\varepsilon)$ и Γ_i .

В дальнейшем несколько упростим задачу и будем задавать Γ_i в виде

$$\Gamma_i = \frac{B}{i}, \quad (3.11)$$

причем матрица B , как указывалось ранее, — некоторая положительно определенная матрица.

Очевидно, АМКО зависит от выбора матрицы B и от вида функции потерь $F(\varepsilon)$.

В работе [1] показано, что при заданной функции потерь минимальное значение АМКО достигается при значениях

$$B^* = \frac{1}{\int_{-\infty}^{\infty} F''(\eta) f(\eta) d\eta} A^{-1}(\bar{c}, \sigma_{\eta}^2), \quad (3.12a)$$

или

$$B^* = \frac{1}{-\int_{-\infty}^{\infty} F'(\eta) f'(\eta) d\eta} A^{-1}(\bar{c}, \sigma_{\eta}^2). \quad (3.12б)$$

Такие алгоритмы в дальнейшем будем называть оптимальными. В этом случае АМКО будет иметь вид [1]

$$V(B^*) = \frac{\int_{-\infty}^{\infty} F'^2(\eta) f(\eta) d\eta}{\left[\int_{-\infty}^{\infty} F''(\eta) f(\eta) d\eta \right]^2} A^{-1}(\bar{c}, \sigma_{\eta}^2)$$

или

$$V(B^*) = \frac{\int_{-\infty}^{\infty} F'^2(\eta) f(\eta) d\eta}{\left[\int_{-\infty}^{\infty} F'(\eta) f'(\eta) d\eta \right]^2} A^{-1}(\bar{c}, \sigma_{\eta}^2).$$

Как видно, полученная АМКО является функционалом от функции потерь $F(\varepsilon)$ и плотности распределения $f(\eta)$. В дальнейшем АМКО, оптимальную относительно матрицы B , будем обозначать $V(F, f)$, т.е.

$$V(F, f) = \frac{\int_{-\infty}^{\infty} F'^2(\eta) f(\eta) d\eta}{\left[\int_{-\infty}^{\infty} F'(\eta) f'(\eta) d\eta \right]^2} A^{-1}(\bar{c}, \sigma_{\eta}^2) d\eta \quad (3.13 \text{ а})$$

или

$$V(F, f) = \frac{\int_{-\infty}^{\infty} F'^2(\eta) f(\eta) d\eta}{\left[\int_{-\infty}^{\infty} F''(\eta) f(\eta) d\eta \right]^2} A^{-1}(\bar{c}, \sigma_{\eta}^2) d\eta. \quad (3.13 \text{ б})$$

Подставляя матрицу B^* , определяемую по формуле (3.12 а) или (3.12 б), в рекуррентный алгоритм (3.8) и раскрывая частную производную, получим:

$$\begin{aligned} \hat{\tilde{c}}_i = \hat{\tilde{c}}_{i-1} + \frac{1}{i} \frac{1}{\int_{-\infty}^{\infty} F'(\eta) f'(\eta) d\eta} A^{-1} \left(\hat{\tilde{c}}_{i-1}, \sigma_{\eta}^2 \right) \frac{dF(\varepsilon)}{d\varepsilon} \Big|_{\varepsilon=y(i)-\psi(i, \hat{\tilde{c}}_{i-1})} \times \\ \times \frac{\partial \psi(i, \tilde{\tilde{c}})}{\partial \tilde{\tilde{c}}} \Big|_{\tilde{\tilde{c}}=\hat{\tilde{c}}_{i-1}} \end{aligned} \quad (3.14a)$$

или

$$\begin{aligned} \hat{\tilde{c}}_i = \hat{\tilde{c}}_{i-1} + \frac{1}{i} \frac{I}{\int_{-\infty}^{\infty} F''(\eta) f(\eta) d\eta} A^{-1} \left(\hat{\tilde{c}}_{i-1}, \sigma_{\eta}^2 \right) \frac{dF(\varepsilon)}{d\varepsilon} \Big|_{\varepsilon=y(i)-\psi(i, \hat{\tilde{c}}_{i-1})} \times \\ \times \frac{\partial \psi(i, \tilde{\tilde{c}})}{\partial \tilde{\tilde{c}}} \Big|_{\tilde{\tilde{c}}=\hat{\tilde{c}}_{i-1}} . \end{aligned} \quad (3.14б)$$

Полученные рекуррентные формулы являются наиболее общими и могут быть использованы для идентификации параметров нелинейного объекта при произвольной функции потерь $F(\varepsilon)$.

Существенным недостатком данного алгоритма является необходимость рассчитывать, а затем обращать матрицу $A(\tilde{\tilde{c}}, \sigma_{\eta}^2)$. В работах [1], [7] приводится рекуррентная формула, позволяющая упростить вычисление этой матрицы. В результате рекуррентная формула (3.14б) распадается на две рекуррентные последовательности

$$\hat{\tilde{c}}_i = \hat{\tilde{c}}_{i-1} + \Gamma(i) F'(\varepsilon(i, \hat{\tilde{c}}(i-1))) \frac{\partial \psi(i, \tilde{\tilde{c}})}{\partial \tilde{\tilde{c}}} \Big|_{\tilde{\tilde{c}}=\hat{\tilde{c}}_{i-1}}, \quad (3.14в)$$

где $\Gamma(i)$ вычисляется на каждом шаге рекуррентного процесса по формуле

$$\Gamma(i) = \Gamma(i-1) -$$

$$\begin{aligned}
& \frac{\Gamma(i-1) \frac{\partial \psi(i, \tilde{c})}{\partial \tilde{c}} \Big|_{\tilde{c}=\hat{c}_{i-1}} \frac{\partial^T \psi(i, \tilde{c})}{\partial \tilde{c}} \Big|_{\tilde{c}=\hat{c}_{i-1}} \Gamma(i-1)}{\left(- \int_{-\infty}^{\infty} F'(\eta) f'(\eta) d\eta \right)^{-1} + \frac{\partial^T \psi(i, \tilde{c})}{\partial \tilde{c}} \Big|_{\tilde{c}=\hat{c}_{i-1}} \Gamma(i-1) \frac{\partial \psi(i, \tilde{c})}{\partial \tilde{c}} \Big|_{\tilde{c}=\hat{c}_{i-1}}} \quad (3.15a)
\end{aligned}$$

или

$$\begin{aligned}
& \Gamma(i) = \Gamma(i-1) - \\
& \frac{\Gamma(i-1) \frac{\partial \psi(i, \tilde{c})}{\partial \tilde{c}} \Big|_{\tilde{c}=\hat{c}_{i-1}} \frac{\partial^T \psi(i, \tilde{c})}{\partial \tilde{c}} \Big|_{\tilde{c}=\hat{c}_{i-1}} \Gamma(i-1)}{\left(\int_{-\infty}^{\infty} F''(\eta) f(\eta) d\eta \right)^{-1} + \frac{\partial^T \psi(i, \tilde{c})}{\partial \tilde{c}} \Big|_{\tilde{c}=\hat{c}_{i-1}} \Gamma(i-1) \frac{\partial \psi(i, \tilde{c})}{\partial \tilde{c}} \Big|_{\tilde{c}=\hat{c}_{i-1}}} . \quad (3.15.6)
\end{aligned}$$

В качестве начального приближения принимают матрицу

$$\Gamma(0) = \lambda I, \quad (3.16)$$

Где λ — большое число. Начальное приближение \hat{c}_0 может быть задано любым вектором соответствующей размерности.

3.3. Абсолютно-оптимальные рекуррентные алгоритмы

Как было указано ранее (см. формулу (3.13a)) АМКО, оптимальная по матрице В, имеет вид

$$V^*(F, f) = \frac{\int_{-\infty}^{\infty} F'^2(\eta) f(\eta) d\eta}{\left[\int_{-\infty}^{\infty} F'(\eta) f'(\eta) d\eta \right]^2} A^{-1}(\bar{c}, \sigma_{\eta}^2).$$

Будем считать, что плотность распределения ошибок измерений известна. Тогда [1] можно найти такую функцию потерь $F^*(\varepsilon)$, которая обеспечивает минимум АМКО:

$$F^*(\varepsilon) = -\ln f(\eta)|_{\eta=\varepsilon}. \quad (3.17)$$

Как видно, оптимальная функция потерь совпадает с функцией потерь, соответствующей методу максимума правдоподобия при аддитивной помехе.

Приведем примеры оптимальных функций потерь и их производных для различных плотностей распределения шумов.

1. Нормальная (гауссова) плотность распределения помех:

$$f(\eta) = \frac{1}{\sqrt{2\pi} \sigma_\eta} \exp\left[-\eta^2 / 2\sigma_\eta^2\right] = N(0, \sigma_\eta).$$

Функция потерь, согласно (3.17), будет иметь вид

$$F^*(\varepsilon) = \varepsilon^2 / 2\sigma_\eta^2 + \ln \sqrt{2\pi} \sigma_\eta, \quad F^{*'}(\varepsilon) = \varepsilon / \sigma_\eta^2. \quad (3.18)$$

Поскольку положение экстремума не зависит от постоянных слагаемых, то можно записать следующий оптимизируемый критерий:

$$J(\tilde{c}) = \sum_{i=1}^N (y(i) - y(i))^2,$$

или в векторной форме: $J(\tilde{c}) = (\bar{y} - \tilde{y})^T (\bar{y} - \tilde{y})$.

Таким образом, при нормальном законе распределения использование оптимальной функции потерь эквивалентно применению метода наименьших квадратов.

2. Экспоненциальная плотность распределения помех (плотность распределения Лапласа)

$$f(\eta) = \frac{1}{2\alpha} \exp[-|\eta|/\alpha] = L(0, \alpha).$$

Для функции потерь получаем:

$$F^*(\varepsilon) = |\varepsilon|/\alpha + \ln 2\alpha, \quad F^{*'}(\varepsilon) = \text{sign } \varepsilon / \alpha. \quad (3.19)$$

В данном случае получили модульную функцию потерь. Использование метода наименьших квадратов будет давать худшие асимптотические свойства оценок.

3. Дробная плотность распределения помех (Коши),

$$f(\eta) = \frac{I}{\pi\alpha} \frac{I}{(I + (\eta/\alpha)^2)} = C(0, \alpha).$$

В этом случае имеем:

$$F^*(\varepsilon) = \ln(\alpha^2 + \varepsilon^2) + \ln \pi / \alpha, \quad F'^*(\varepsilon) = 2\varepsilon / (\alpha^2 + \varepsilon^2). \quad (3.20)$$

Эти функции потерь и их производные приведены в табл. 3.1 [1].

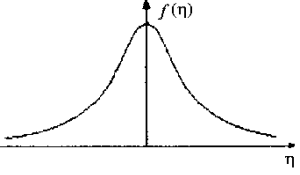
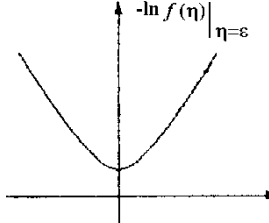
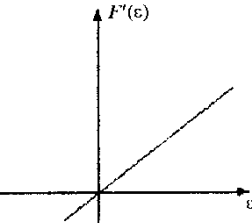
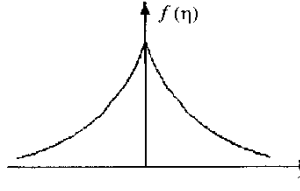
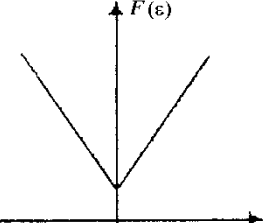
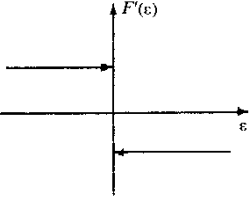
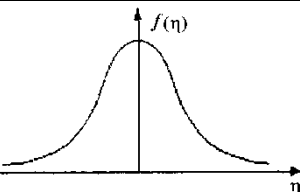
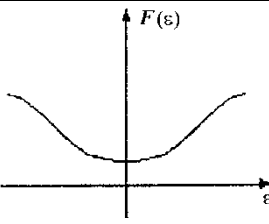
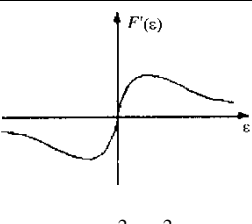
Найдем АМКО при оптимально выбранной функции потерь $F^*(\varepsilon) = -\ln f(\eta)|_{\eta=\varepsilon}$.

Подставив выражение (3.1 7) для оптимальной функции потерь в формулу для АМКО, получим:

$$\begin{aligned} V(F^*, f) &= \frac{\int_{-\infty}^{\infty} f(\eta) F'^{*2}(\eta) d\eta}{\left[\int_{-\infty}^{\infty} f'(\eta) F'^*(\eta) d\eta \right]^2} A^{-1}(\bar{c}, \sigma_{\eta}^2) = \\ &= \frac{\int_{-\infty}^{\infty} f(\eta) [(-\ln f(\eta))']^2 d\eta}{\left[\int_{-\infty}^{\infty} f'(\eta) (-\ln f(\eta))' d\eta \right]^2} A^{-1}(\bar{c}, \sigma_{\eta}^2) = \\ &= \frac{\int_{-\infty}^{\infty} f(\eta) [f'(\eta) / f(\eta)]^2 d\eta}{\int_{-\infty}^{\infty} f'(\eta) [-f'(\eta) / f(\eta)] d\eta} A^{-1}(\bar{c}, \sigma_{\eta}^2). \end{aligned}$$

Таблица 3.1

Оптимальные функции потерь и их производные

Плотность распределения $f(\eta)$	Функция потерь $F(\varepsilon)$	Производная функции потерь $F'(\varepsilon)$
 $N(0, \sigma_\eta^2) = \frac{1}{\sqrt{2\pi\sigma_\eta^2}} e^{-\eta^2/2\sigma_\eta^2}$	 $\varepsilon^2/2\sigma_\eta^2 + \ln\sqrt{2\pi\sigma_\eta^2}$	 $\varepsilon/\sigma_\eta^2$
 $L(0, s) = \frac{1}{2s} e^{-\frac{ \eta }{s}}$	 $ \varepsilon /s + \ln 2s$	 $1/s \operatorname{sign} \varepsilon$
 $c(0, s) = \frac{1}{\pi s} \cdot \frac{1}{\left(1 + \left(\frac{\eta}{s}\right)^2\right)}$	 $\ln(s^2 + \varepsilon^2) + \ln \frac{\pi}{s}$	 $2\varepsilon/(s^2 + \varepsilon^2)$

Приводя подобные члены, окончательно получаем:

$$V(F^*, f) = V(f) = \frac{A^{-1}(\bar{c}, \sigma_\eta^2)}{\int_{-\infty}^{\infty} [f'(\eta)]^2 / f(\eta) d\eta}. \quad (3.21)$$

Выражение, стоящее в знаменателе формулы (3.21), называется фишеровской информацией:

$$I_F(f(\eta)) = \int_{-\infty}^{\infty} [f'(\eta)]^2 / f(\eta) d\eta. \quad (3.22)$$

Очевидно, справедливо следующее неравенство:

$$V(F, f) \geq V(F^*, f), \quad (3.23)$$

где $F(\eta) \neq -\ln f(\eta)$, $F^*(\eta) = -\ln f(\eta)$. Или, раскрывая $V(F, f)$ и $V(F^*, f)$, получим:

$$\frac{-\int_{-\infty}^{\infty} F'^2(\eta) f(\eta) d\eta}{\left[\int_{-\infty}^{\infty} F'(\eta) f'(\eta) d\eta \right]^2} A^{-1}(\bar{c}, \sigma_\eta^2) \geq \frac{1}{\int_{-\infty}^{\infty} f'^2(\eta) / f(\eta) d\eta} A^{-1}(\bar{c}, \sigma_\eta^2). \quad (3.24)$$

Запишем абсолютно-оптимальный рекуррентный алгоритм для нелинейного объекта вида

$$y(i) = \psi(i, \tilde{c}) + \eta(i).$$

Для этого подставим в рекуррентные соотношения (3.15а), (3.14в) оптимальную функцию потерь (3.17), а именно:

$$F^*(\varepsilon) = -\ln f(\eta)|_{\eta=\varepsilon}.$$

Тогда получим:

$$\hat{c}(i) = \hat{c}(i-1) + \Gamma(i) \left(\frac{-1}{f(\eta)} f'(\eta) \right) \bigg|_{\eta=\varepsilon(i, \tilde{c}(i-1))} \frac{\partial \psi(i, \tilde{c})}{\partial \tilde{c}} \bigg|_{\tilde{c}=\tilde{c}(i-1)}, \quad (3.25)$$

$$\Gamma(i) = \Gamma(i-1) - \frac{\Gamma(i-1) \frac{\partial \psi(i, \tilde{c})}{\partial \tilde{c}} \bigg|_{\tilde{c}=\hat{c}(i-1)} \frac{\partial^T \psi(i, \tilde{c})}{\partial \tilde{c}} \bigg|_{\tilde{c}=\hat{c}(i-1)} \Gamma(i-1)}{(I_F)^{-1} + \frac{\partial^T \psi}{\partial \tilde{c}} \bigg|_{\tilde{c}=\hat{c}(i-1)} \Gamma(i-1) \frac{\partial \psi(i, \tilde{c})}{\partial \tilde{c}} \bigg|_{\tilde{c}=\hat{c}(i-1)}}, \quad (3.26)$$

$\hat{c}(0) = \hat{c}_0$, $\Gamma(0) = \lambda I$, $\lambda \gg 1$, \hat{c}_0 — любой вектор соответствующей размерности; I_F — фишеровская информация, рассчитываемая по формуле:

$$I_F = \int_{-\infty}^{\infty} \frac{(f'(\eta))^2}{f(\eta)} d\eta. \quad (3.27)$$

В табл. 3.2 приведены значения фишеровской информации для наиболее распространенных распределений.

Таблица 3.2

Фишеровская информация

№	Распределение	Фишеровская информация
1	Нормальное	$(\sigma_\eta^2)^{-1}$
2	Лапласа	$(s^2)^{-1}$
3	Коши	$(2s^2)^{-1}$

В случае линейного регрессионно-авто-регрессионного (РАР) объекта $\frac{\partial \psi(i, \tilde{c})}{\partial \tilde{c}} = \bar{z}(i)$.

Тогда, подставляя последнее выражение в рекуррентные соотношения (3.25), (3.26), получим:

$$\hat{c}(i) = \hat{c}(i-1) + \Gamma(i) \left(\frac{-1}{f(\eta)} f'(\eta) \right) \bigg|_{\eta=\varepsilon(i, \tilde{c}(i-1))} \bar{z}(i), \quad (3.28)$$

$$\Gamma(i) = \Gamma(i-1) - \frac{\Gamma(i-1) \bar{z}(i) \bar{z}^T(i) \Gamma(i-1)}{(I_F)^{-1} + \bar{z}^T(i) \Gamma(i-1) \bar{z}(i)}, \quad (3.29)$$

$$\hat{c}(0) = \bar{c}_0, \quad \Gamma(0) = \lambda I, \quad \lambda \gg 1.$$

В заключение настоящего раздела запишем абсолютно-оптимальные рекуррентные алгоритмы для различных плотностей распределения, приведенных в табл. 3.1.

1. Нормальная плотность распределения помехи:

$$f(\eta) = \frac{1}{(2\pi\sigma_\eta^2)^{1/2}} \exp\left(-\frac{1}{2\sigma_\eta^2} \eta^2\right).$$

Введем обозначения: $H(i) = \Gamma(i) / \sigma_\eta^2$.

Очевидно, при нормальном законе распределения функция потерь для абсолютно-оптимального будет квадратичной, а именно:

$$F^*(\varepsilon) = -\ln f(\eta) \big|_{\eta=\varepsilon} \Rightarrow F^*(\varepsilon) = \frac{1}{2\sigma_\eta^2} \varepsilon^2.$$

Тогда в новых обозначениях и учитывая вид функции невязки, абсолютно-оптимальный рекуррентный алгоритм для нормального закона распределения ошибок измерений принимает вид

$$\hat{c}(i) = \hat{c}(i-1) + H(i) \bar{z}(i) \varepsilon \big|_{\varepsilon=(y(i) - \bar{z}^T(i) \hat{c}(i-1))}, \quad (3.30)$$

$$\hat{c}(0) = \bar{c}_0;$$

$$H(i) = H(i-1) - \frac{H(i-1) \bar{z}(i) \bar{z}^T(i) H(i-1)}{1 + \bar{z}^T(i) H(i-1) \bar{z}(i)}, \quad (3.31)$$

$$H(i) = \lambda I, \quad \lambda \gg 1.$$

Как видим, данный алгоритм имеет линейный относительно невязки вид, а потому получил название «линейный» рекуррентный алгоритм.

Сравнивая (3.30), (3.31) с рекуррентными соотношениями метода наименьших квадратов, можно сделать вывод, что рекуррентная форма метода наименьших квадратов полностью совпадает с абсолютно-оптимальным алгоритмом для нормального распределения помехи.

Алгоритм (3.31, 3.32) называется линейным. Как видим, линейный оптимальный алгоритм не зависит от дисперсии помехи, хотя, как нетрудно заметить, АМКО для этого распределения будет зависеть от дисперсии шума.

2. Плотность распределения помехи — Лапласа:

$$f(\eta) = \frac{1}{2s} e^{-\frac{|\eta|}{s}}.$$

Принимая во внимание вид плотности распределения Лапласа и соответствующую фишеровскую информацию (см. табл. 3.2), функция потерь и абсолютно-оптимальный рекуррентный алгоритм (3.28), (3.29) принимают вид:

$$\begin{aligned} F^*(\varepsilon) &= -\ln f(\eta)|_{\eta=\varepsilon} \Rightarrow F^*(\varepsilon) = |\varepsilon|; \\ \hat{\bar{c}}(i) &= \hat{\bar{c}}(i-1) + H(i)\bar{z}(i)\text{sign}(\varepsilon)|_{\varepsilon=(y(i)-\bar{z}^T(i)\hat{\bar{c}}(i-1))}; \\ \hat{\bar{c}}(0) &= \bar{c}_0; \end{aligned} \quad (3.32)$$

$$\begin{aligned} H(i) &= H(i-1) - \frac{H(i-1)\bar{z}(i)\bar{z}^T(i)H(i-1)}{s_\eta^2 + \bar{z}^T(i)H(i-1)\bar{z}(i)}, \\ H(0) &= \lambda I, \quad \lambda \gg 1, \quad H(i-1) = \frac{\Gamma(i-1)}{s}. \end{aligned} \quad (3.33)$$

Учитывая релейную зависимость от невязки, данный алгоритм получил название «релейный».

3. Плотность распределения помехи Коши:

$$f(\eta) = \frac{1}{\pi s} \cdot \frac{1}{\left(1 + \left(\frac{\eta}{s}\right)^2\right)}$$

В этом случае функция потерь будет иметь вид

$$F^*(\varepsilon) = -\ln f(\eta)|_{\eta=\varepsilon} \Rightarrow F^*(\varepsilon) = \ln(s^2 + \varepsilon^2) + \ln \frac{\pi}{s}.$$

Подставляя функцию потерь в абсолютно-оптимальный рекуррентный алгоритм (3.28), (3.29), соответствующий распределению Коши, получим «нелинейный» абсолютно оптимальный рекуррентный алгоритм:

$$c(i) = c(i-1) + 4H(i)\bar{z}(i) \frac{\varepsilon}{s^2 + [\varepsilon]^2} \Big|_{\varepsilon=y(i)-\bar{z}^T(i)\hat{c}(i-1)}, \quad (3.34)$$

$$\hat{c}(0) = \bar{c}_0;$$

$$H(i) = H(i-1) - \frac{H(i-1)\bar{z}(i)\bar{z}^T(i)H(i-1)}{s^2 + \bar{z}^T(i)H(i-1)\bar{z}(i)}, \quad (3.35)$$

$$H(0) = \mathcal{M}, \quad \lambda \gg 1, \quad H(i) = \frac{G(i)}{2}.$$

Приведенные абсолютно-оптимальные алгоритмы идентификации РАР-объектов с простой помехой, как было показано в предыдущем разделе, обладают предельно возможной скоростью сходимости. Их АМКО равна:

$$V(F, f) = \frac{1}{I_F(f)} A^{-1}(\bar{c}, \sigma_\eta^2), \quad (3.36)$$

где $I_F(f) = \int_{-\infty}^{\infty} \frac{f'^2(\eta)}{f(\eta)} d\eta$ — фишеровская информация;

$A(\bar{c}, \sigma_\eta^2) = M\{\bar{z}, \bar{z}^T\}$ — нормированная информационная матрица системы.

3.4. Пример использования абсолютно-оптимальных рекуррентных алгоритмов для идентификации параметров линейного регрессионного объекта

Для оценки эффективности абсолютно-оптимальных алгоритмов рассмотрим задачу идентификации параметров линейного регрессионного объекта вида

$$y(i) = 3 - 0,5u_1(i) + 0,0u_2(i) + 6,0u_3(i) - 2,0u_4(i) + \eta(i). \quad (3.37)$$

Пусть шум измерений $\eta(i)$ имеет распределение Коши:

$$f_c(\eta) = \frac{1}{\pi s} \cdot \frac{1}{(1 + (\eta/s)^2)}.$$

Тогда, в зависимости от того, насколько хорошо изучен «объект» идентификации, могут возникнуть следующие ситуации.

1. Объект изучен хорошо, правильно определена плотность распределения шума, т.е. принятая и реальная плотности распределения совпадают. Оптимальная функция потерь и ее производная будут иметь вид

$$F^*(\varepsilon) = -\ln f(\eta)|_{\eta=\varepsilon} = \ln(\pi/s) + \ln(s^2 + \varepsilon^2);$$

$$F^{*'}(\varepsilon) = 2\varepsilon/(s^2 + \varepsilon^2), \quad \varepsilon(i, \hat{c}(i-1)) = y(i) - \bar{z}^T(i) \cdot \hat{c}(i-1).$$

В результате рекуррентный алгоритм принимает вид

$$\hat{c}(i) = \hat{c}(i-1) + \Gamma(i) \frac{2(y(i) - \hat{b}_0(i-1) - \dots - \hat{b}_4(i-1)u_4(i))}{s^2 + (y(i) - \hat{b}_0(i-1) - \dots - \hat{b}_4(i-1)u_4(i))^2} \bar{z}(i),$$

$$\Gamma(i) = \Gamma(i-1) - \frac{\Gamma(i-1)\bar{z}(i)\bar{z}^T(i)\Gamma(i-1)}{2s^2 + \bar{z}^T(i)\Gamma(i-1)\bar{z}(i)}.$$

Асимптотическая ковариационная матрица ошибки оценки, соответствующая этому алгоритму при плотности распределения шума (плотности распределения Коши), будет:

$$V^*(F^*(\eta), f(\eta)) = \frac{A^{-1}(\bar{c}, \sigma_\eta^2)}{\int_{-\infty}^{\infty} [f'(\eta)]^2 / f(\eta) d\eta} = 2s^2 A^{-1}(\bar{c}, \sigma_\eta^2).$$

2. Объект изучен плохо. При этом разработчик, как правило, считает, что плотность распределения шума соответствует нормальному закону распределения, хотя на самом деле шум имеет распределение Коши. Таким образом, опираясь на ложную гипотезу о нормальном распределении шума, разработчик в качестве функции потерь выбирает квадратичную функцию:

$$F(\varepsilon) = \frac{\varepsilon^2}{2\sigma_\eta^2},$$

при этом ожидаемая АМКО будет

$$V^*(F(\eta) = -\ln(f_N(\eta)), f_N(\eta)) = \sigma_\eta^2 A^{-1}(\bar{c}, \sigma_\eta^2).$$

Однако так как на самом деле шум имеет распределение Коши, то АМКО будет иметь совсем другое значение:

$$V(F(\eta) = -\ln(f_N(\eta)), f_c(\eta)) = \frac{\int_{-\infty}^{\infty} F'^2(\eta) f_c(\eta) d\eta}{\left[\int_{-\infty}^{\infty} F''(\eta) f_c(\eta) d\eta \right]^2} A^{-1}(\bar{c}, \sigma_\eta^2) = \infty.$$

Таким образом, если истинное распределение шума – распределение Коши, а мы считаем, что распределение шума – нормальное, ошибки оценки будут очень велики, причем установить факт ошибки практически невозможно.

Для оценки эффективности использования абсолютно-оптимальных рекуррентных алгоритмов проводилось сравнение сглаженной ошибки оценки параметров линейного регрессионного объекта (3.37), рассчитанных с помощью нелинейного и линейного алгоритмов, при различных значениях параметра распределения Коши — s . Напомним, что сглаженная ошибка оценки вычисляется по формуле

$$\Delta_{\text{ск}}(i) = \frac{\sum_{j=0}^{10} \sum_{k=0}^n (\hat{b}_k(i-j) - b_k)}{10}; \quad i = 10, 11, 12, \dots$$

Было проведено три серии экспериментов для различных значений s по два эксперимента в каждой серии, соответствующие двум алгоритмам оценивания (линейного и нелинейного). Данные для моделирования приведены в табл. 3.3, «истинные» коэффициенты объекта для моделирования приведены в столбце 8 табл. 3.3.

Таблица 3.3

Данные для тестового моделирования

№	Распр Коши s	Нормальное распределение	σ_u^2	μ_u	Кол. изменений	$\Gamma(0)=\lambda I$	$\hat{b}_{ls}(0)$	\bar{b}
1	0,1	$\mu_\eta=0$ $\sigma_\eta^2=0.5$	50	1	190	$\lambda=0,1$	$\begin{bmatrix} 3,5 \\ -0,3 \\ 0,1 \\ 5,9 \\ -2,2 \end{bmatrix}$	$\begin{bmatrix} 3,0 \\ 0,5 \\ 0,0 \\ 6,0 \\ -2,2 \end{bmatrix}$
2	2							
3	10							

Результаты выполнения работы отображены на рис. 3.2, 3.3, 3.4.

Анализируя эти рисунки, можно сделать вывод о том, что использование абсолютно-оптимального «нелинейного» рекуррентного алгоритма обеспечивает существенно меньшую ошибку оценки по сравнению с «линейным» алгоритмом, особенно это заметно при больших значениях параметра распределения Коши s .

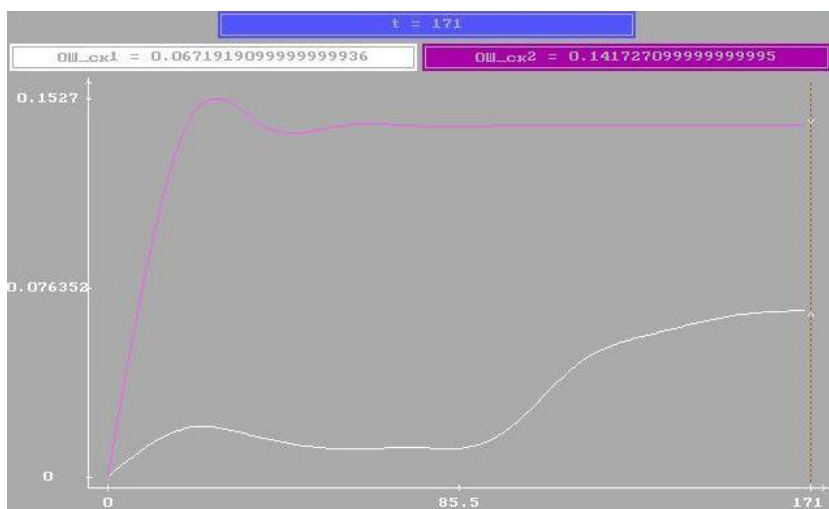


Рис.3.2. График зависимости сглаженной ошибки оценки от номера измерений при использовании линейного (white) и абсолютно-оптимального (magenta) алгоритмов при $s = 0,1$

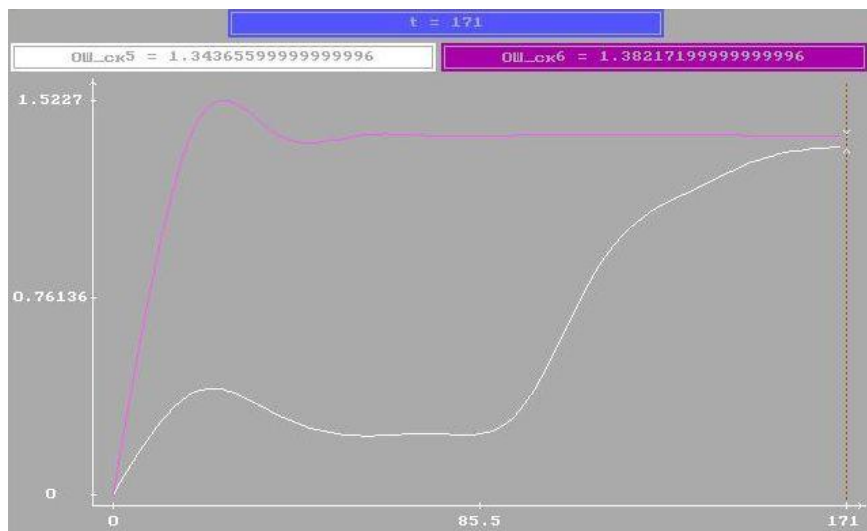


Рис. 3.3. График зависимости сглаженной ошибки оценки от номера измерений при использовании линейного (white) и абсолютно-оптимального (magenta) алгоритмов при $s = 2$

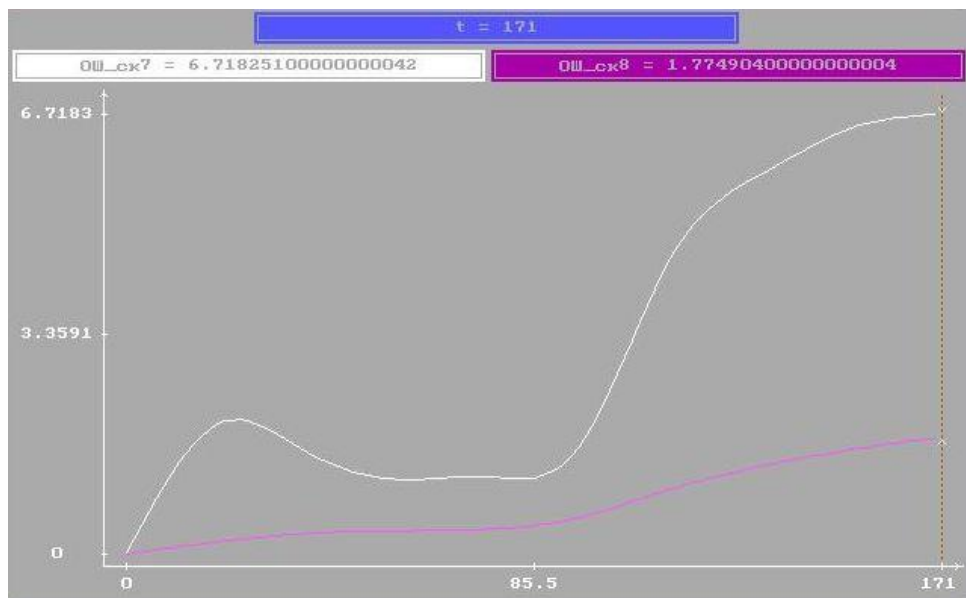


Рис. 3.4. График зависимости сглаженной ошибки оценки от номера измерений при использовании линейного (white) и абсолютно - оптимального (magenta) алгоритмов при $s = 10$

3.5. Описание интерфейса лабораторной работы

Интерфейс данной лабораторной работы при вводе данных о студенте и вводе данных для тестового моделирования объекта полностью совпадает с интерфейсом предыдущих лабораторных работ (см. рис. 1.2 – 1.14). Однако при выборе распределения шумов (см. рис. 1.9) необходимо задать распределение, соответствующее варианту работы. После моделирования работы объекта приступаем к формированию алгоритмов оценивания (рис. 3.5).

При активации п.7 всплывает окно выбора рекуррентного алгоритма (рис. 3.6).

Далее для каждого выбранного алгоритма порядок выполнения и интерфейс аналогичны предыдущим работам.

Правила работы с программой графического оформления результатов также подробно изложены в описании работы 2.

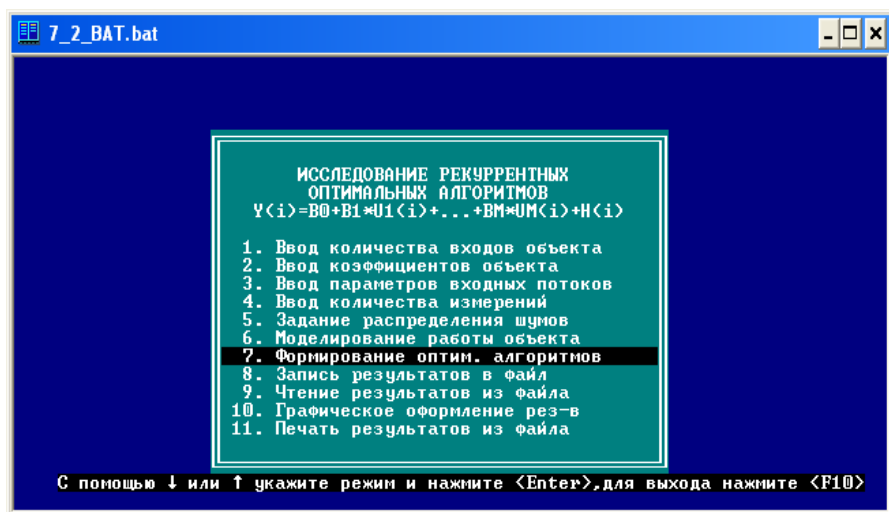


Рис. 3.5 Главное меню. Формирование оптимальных алгоритмов

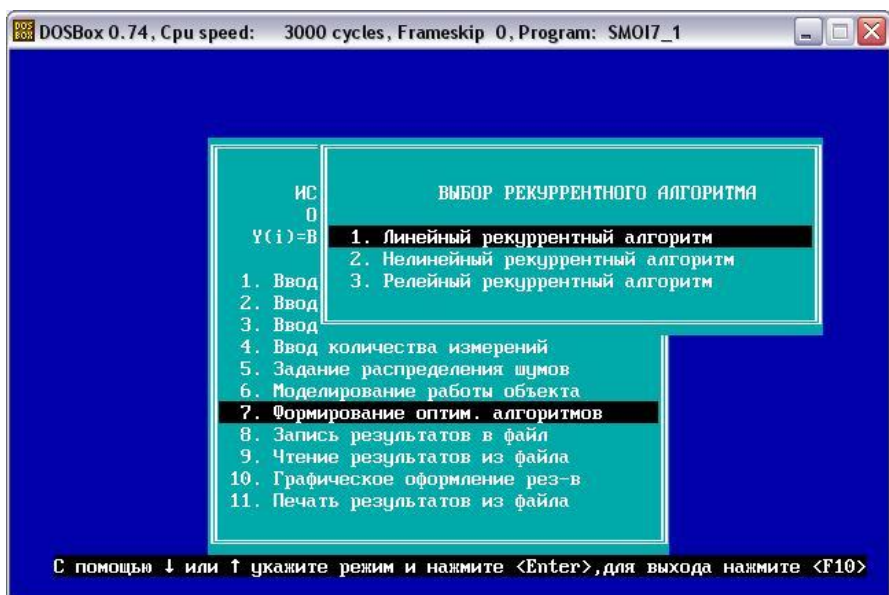


Рис. 3.6. Выбор рекуррентного алгоритма

3.6. Задание и порядок выполнения лабораторной работы

В лабораторной работе необходимо провести исследование эффективности использования абсолютно-оптимальных рекуррентных алгоритмов для оценки параметров линейного регрессионного объекта по сравнению с использованием «линейного» алгоритма при различных статистических характеристиках распределения вероятности шума измерений. В этой связи необходимо выполнить ряд действий.

1. Изучить основы теории абсолютно-оптимальных рекуррентных алгоритмов.

2. Записать абсолютно-оптимальный рекуррентный алгоритм и асимптотическую ковариационную матрицу ошибки оценки при условиях

а. Идентифицируемый процесс изучили хорошо, плотность распределения шума, принимаемая для расчетов, совпадает с «реальной» (задана вариантом);

б. Идентифицируемый процесс изучили плохо, плотность распределения шума, принимаемая для расчетов, имеет нормальный закон распределения и не совпадает с «реальной» (задана вариантом). Запишите ожидаемую асимптотическую ковариационную матрицу ошибки оценки.

3. Провести исследование полученных асимптотических ковариационных матриц ошибок оценки.

4. Повторить исследования (п. 2) для различных значений параметра s плотности распределения шума измерений.

5. Построить графики скользящей ошибки оценки в зависимости от номера измерений при различных значениях параметра распределения s , аналогично приведенным на рис. 3.2 – 3.4.

6. Оформить отчет по лабораторной работе.

3.7. Контрольные вопросы

1. Какая априорная информация необходима для реализации абсолютно-оптимального рекуррентного алгоритма?

2. В чем преимущество абсолютно-оптимальных рекуррентных алгоритмов по сравнению с оптимальными?

3. Запишите общий вид оптимального рекуррентного алгоритма.

4. Запишите асимптотическую ковариационную матрицу ошибки оценки для оптимального и абсолютно-оптимального рекуррентных алгоритмов.

5. Запишите функцию потерь, соответствующую абсолютно-оптимальному рекуррентному алгоритму.

6. Запишите абсолютно-оптимальный рекуррентный алгоритм при нормальном законе распределения ошибок измерений.