

УДК 621.391

ТОЧНОСТЬ ПРОЦЕДУР ФИЛЬТРАЦИИ, ЭКСТРАПОЛЯЦИИ И ИНТЕРПОЛЯЦИИ СЛУЧАЙНЫХ ПРОЦЕССОВ



[В.В. ПОПОВСКИЙ](#)

Харьковский национальный
университет радиоэлектроники



[И.В. СТРЕЛКОВСКАЯ](#)

Одесская национальная
академия связи им. А.С. Попова

Проаналізована точність оцінки фільтрації, екстраполяції та інтерполяції випадкових процесів. Показано, що точність екстраполяції знижується зі збільшенням інтервалу прогнозу. Стверджується, що екстраполяція має сенс лише на інтервал, що не перевищує інтервал кореляції. Показано, що за рахунок інтерполяції послідовності точкових оцінок точність відновлення реалізації випадкового процесу, що оцінюється, перевищує оцінки, отримані за допомогою фільтра Калмана-Б'юсі.

The accuracy of filtering estimation, extrapolation and interpolation of random processes has been analyzed. It is shown that the accuracy of extrapolation decreases with increasing of prediction interval. It is argued that the extrapolation is meaningful only within the interval, which does not exceed the correlation interval. It is shown that due to the interpolation of the sequence of the recursive point estimates, the accuracy of estimated realization recovery of the random process exceeds the estimates obtained by Kalman-Bucy.

Проанализирована точность оценки фильтрации, экстраполяции и интерполяции случайных процессов. Показано, что точность экстраполяции снижается с увеличением интервала прогноза. Утверждается, что экстраполяция имеет смысл лишь на интервал, не превышающий интервал корреляции. Показано, что за счет интерполяции последовательности точечных рекурсивных оценок точность восстановления оцениваемой реализации случайного процесса превышает точность оценок, полученных с помощью фильтра Калмана-Бьюси.

Введение

Общая постановка задачи. В современных телекоммуникационных технологических решениях используются различные процедуры стохастической оценки случайных процессов [1]. Примером тому являются процедура RED в алгоритме предотвращения перегрузки буфера очереди на маршрутизаторе сети, оценка времени кругового обращения RTT и др. На основе полученных оценок удастся построить алгоритмы управления [2]. Причем при выполнении условий теоремы о разделении алгоритмы управления оказываются детерминированными.

В зависимости от характера решаемых задач возникает необходимость получения оценок фильтрации, экстраполяции или интерполяции. Эффективность этих задач, используемых в вышеупомянутых процедурах, непосредственно зависит от точности получаемых оценок. Наиболее изученными и проанализированными являются оценки фильтрации [3]. Что же касается оценок экстраполяции и интерполяции, то мнение специалистов оказывается неоднозначным. Так, иногда вызывает сомнение утверждение о том, что за счет интерполяции можно улучшить точность оценки фильтрации. Попытаемся разобраться в этих задачах.

I. Повышение точности оценок случайных процессов

Фильтрация в теории случайных процессов определяется как неубывающее семейство σ -алгебр $\{F_t\}_{t \in T}$.

Это семейство,

$$F_t^x = \sigma\{x_s | S \leq t, s \in T\}, t \in T,$$

по сути, формирует интервальную на длительности T оценку случайного процесса $\hat{x}(t)$. Одной из разновидностей процедур, реализующих эту оценку, является фильтр Калмана-Бьюси (ФКБ):

$$\frac{d\hat{x}(t)}{dt} = F(t)\hat{x}(t) + V(t)N_v^{-1}H(t)[y(t) - H(t)\hat{x}(t)], \quad (1)$$

где $y(t) = H(t)x(t) + v(t)$ – процесс наблюдения, имеющий спектральную плотность мощности N_v ; $F(t)$ – матрица состояния, определяющая инерционность процесса $x(t)$.

На главной диагонали матрицы F расположены элементы $\alpha_{ii} = \tau_{ii}^{-1}$, где τ – интервал корреляции соответствующего процесса, $V(t)$ – апостериорная дисперсия ошибки оценки $\tilde{x}(t) = (\hat{x}(t) - x(t))$, определяемая из уравнения Риккати:

$$\frac{dV(t)}{dt} = F^T(t)V(t) + V(t)F(t) - V(t)H^T(t)N_v^{-1}H(t)V(t) + G^T(t)N_\xi G(t). \quad (2)$$

Динамика состояния оцениваемой системы определяется дифференциальным уравнением:

$$\frac{dx(t)}{dt} = F(t)x(t) + G(t)\xi(t), \quad (3)$$

где $\xi(t)$ – гауссов белый шум генерации, некоррелированный с шумом $v(t)$ в канале наблюдения; $G(t)$ – коэффициент, определяющий уровень процесса $x(t)$.

Очевидно, оценка (1) на длительности интервала T является последовательностью точечных оценок, получаемых в каждый момент в течение времени $t \in T$. Апостериорная дисперсия (2) позволяет определить качество точечных оценок на выходе ФКБ как в переходный период, когда режим ФКБ устанавливается, так и в уже установившемся режиме, когда $dV(t)/dt \rightarrow 0$. С учетом этого из уравнения (2) имеем

$$V(\infty) = \frac{2N_\xi}{1 + \sqrt{1 + h^2}}, \quad (4)$$

где $h^2 = 2N_\xi / \alpha N_v$ – соотношение сигнал/шум в полосе частот $\Delta f_{эф} = \alpha = \tau_{кор}^{-1}$.

Из формулы (4) следует, что относительное значение апостериорной дисперсии $V(\infty)/N_\xi$ с увеличением h^2 асимптотически уменьшается, абсолютное же значение этой дисперсии возрастает пропорционально уровню оцениваемого процесса $x(t)$. Последнее утверждение часто упускается из виду при качественном анализе результатов ФКБ.

Рассмотренный вариант ФКБ (1) является аналоговым. Его обычно анализируют с методической точки зрения. Практическую значимость имеет дискретный вариант ФКБ, уравнение оценки которого:

$$\hat{x}(k+1) = F(k+1, k)\hat{x}(k) + K(k)[H(k)F(k+1, k)\hat{x}(k) - y(k)], \quad (5)$$

где

$$\begin{aligned} K(k) &= V(k)H^T(k)N_v^{-1}, \\ V(k) &= [I - K(k)H(k)]V(k, k-1); \\ V(k, k-1) &= F^T(k, k-1)V(k-1)F(k, k-1) + G^T(k)N_\xi G(k). \end{aligned} \quad (6)$$

В представленных уравнениях сохранены обозначения, совпадающие по смыслу с аналоговым алгоритмом. Значения коэффициентов, входящих в уравнение состояния:

$$\begin{aligned} F(k, k-1) &= e^{-\alpha\Delta t} \text{ – коэффициент состояния,} \\ G(k) &= \sqrt{1 - e^{-\alpha\Delta t}} \text{ – коэффициент генерации,} \end{aligned}$$

где $\Delta t = t_k - t_{k-1}$ – шаг дискретизации процесса $x(k)$.

Исследование качества работы дискретного ФКБ удобно вести с использованием имитационной модели в виде формирующего фильтра (3). При этом ошибка оценки $\tilde{x}(k) = \hat{x}(k) - x(k)$ используется для оценки выборочной дисперсии:

$$\sigma_{\tilde{x}}^2 = \frac{1}{n-1} \sum_{k=1}^n \tilde{x}^2. \quad (7)$$

Анализ показывает, что точность оценки $\hat{x}(k)$ возрастает (дисперсия (7) уменьшается) при следующих условиях:

- при улучшении сигнально-помеховой обстановки (с увеличением h^2);
- с уменьшением шага дискретизации Δt по отношению к интервалу корреляции $\tau_{кор}$;
- при приближении параметров выбранной модели к реальной сигнально-помеховой обстановке.

Экстраполяция – особый вид аппроксимации, при котором функция аппроксимируется вне заданного интервала. В теории оценивания экстраполированная оценка $\hat{x}(t + \tau_s)$ находится по результатам наблюдения $y(t) = H(t)x(t) + v(t)$, то есть экстраполяция – это прогноз оценки на интервале τ_s . Для обычно используемых в динамических случайных системах марковских моделей, функция корреляции которых:

$$R(\tau) = \sigma_x^2 e^{-\alpha|\tau|},$$

прогноз процесса $x(t)$ на интервале τ находится по формуле:

$$x(t + \tau) = e^{-\alpha\tau} x(t),$$

где предполагается, что данный процесс имеет нулевое среднее.

Можно показать [4], что среднеквадратическая ошибка прогноза для процесса (3) по результатам оценки (1) определяется из выражения:

$$V(t, \tau_3) = R(\tau)V(t) + \sigma_x^2[1 - R(\tau)], \quad (8)$$

где σ_x^2 – дисперсия случайного процесса $x(t)$, в терминах рассматриваемых здесь задач $\sigma_x^2 = N_\xi$; $V(t)$ – определяется из (2).

Анализ дисперсии прогноза (8) показывает, что $V(t, \tau_3)$ изменяется от величины $V(t)$ (при $\tau_3 \rightarrow 0$) до σ_x^2 (при $\tau_3 > \tau_{кор}$). Следовательно, точность экстраполированной оценки снижается с увеличением интервала прогноза. Таким образом, прогнозировать за пределы интервала корреляции не имеет смысла, ибо значение апостериорной дисперсии $V(t, \tau_3)$ возрастает и стремится к априорной σ_x^2 .

Интерполяция – это разновидность аппроксимации последовательности, полученной на определенном интервале оценок некоей функцией с целью нахождения более точного отображения исходного оцениваемого процесса.

На практике оцениваемый процесс $x(k)$ часто отображают, соединяя линиями промежутки между соседними оценками $\hat{x}(k)$ и $\hat{x}(k-1)$. Очевидно, что такая линейная аппроксимация является не единственно возможной и не наиболее точной. Из теории аппроксимации известно [6], что с ростом порядка аппроксимации функции в виде последовательности точек точность аппроксимации возрастает. Таким образом, попытаемся получить более точное восстановление реализаций оцениваемого случайного процесса $x(t)$ методами интерполяции на интервале наблюдения по сравнению с оптимальными оценками ФКБ. В данной постановке задачи нет ни нон-сенса ни парадокса. Следует обратить внимание на то, что с помощью ФКБ получают последовательность точечных оценок

$$\vec{x}^T = \hat{x}(k), \hat{x}(k+1), \hat{x}(k+2) \dots \quad (9)$$

Решая задачу интерполяции последовательности (16) на определенном временном интервале получают уже интервальную оценку. Методов интерполяции имеется большое множество. Рассмотрим задачу определения точности интерполяции последовательности (9), используя сплайн-функции третьего порядка.

Используя в качестве узловых точек моменты отсчетов $t_k, k=1 \dots N$, построим кубический сплайн. В качестве граничных условий выберем условие непрерывности первой производной в узлах сплайна.

Пусть в узлах сетки $\Delta: a=t_0 < t_1 < \dots < t_N = b$ заданы значения некоторой функции $f(t)$ и ее производной $f'(t)$:

$$f_k = f(t_k), f'_k = f'(t_k), k = 0, 1, \dots, N.$$

Согласно [6], кубическим интерполяционным сплайном дефекта 2 (эрмитовым кубическим сплайном) будем называть функцию $S_{3,2}(f, t) \equiv S_{3,2}(t)$, удовлетворяющую условиям:

- 1) на каждом из промежутков $[t_k, t_{k+1}]$

$$S_{3,2}(t) = a_{k0} + a_{k1}(t - t_k) + a_{k2}(t - t_k)^2 + a_{k3}(t - t_k)^3;$$

$$2) S_{3,2}(t_k) = f_k; \quad S'_{3,2}(t_k) = f'_k; \quad k = 0, 1, \dots, N.$$

Очевидно, что $S_{3,2}(t) \in C^1[a, b]$, где $C^1[a, b]$ – класс функций, имеющих на $[a, b]$ непрерывную производную первого порядка. Производная второго порядка эрмита кубического сплайна, вообще говоря, разрывна в узлах сетки Δ . В целях сокращения записи мы будем опускать индекс, указывающий дефект сплайна.

Учитывая условия интерполяции, для вычисления коэффициентов $a_{k\alpha}, \alpha = 0, 1, 2, 3$, при каждом i имеем систему уравнений

$$S_3(t_k) = f_k; \quad S_3(t_{k+1}) = f_{k+1}; \quad S'_3(t_k) = f'_k; \quad S'_3(t_{k+1}) = f'_{k+1}.$$

Решив эту систему, получаем на $[t_k, t_{k+1}]$

$$S_3(t) = \varphi_1(p)f_k + \varphi_2(p)f_{k+1} + \varphi_3(p)\Delta t_k f'_k + \varphi_4(p)\Delta t_k f'_{k+1},$$

причем

$$\varphi_1(p) = (1 - p^2)(1 + 2p), \quad \varphi_2(p) = p^2(3 - 2p), \quad \varphi_3(p) = p(1 - p)^2, \quad \varphi_4(p) = -p^2(1 - p),$$

$$\Delta t_k = t_{k+1} - t_k, \quad p = (t - t_k) / \Delta t_k.$$

Для практического вычисления сплайна в точке $t \in [t_k, t_{k+1}]$ более выгодно с точки зрения количества выполняемых арифметических операций использование следующей формулы:

$$S_3(t) = f_k + (t - t_k)[f'_k + p(B + pA)],$$

где

$$A = -2(f_{k+1} - f_k) / \Delta t_k + (f'_k + f'_{k+1}),$$

$$B = -A + (f_{k+1} - f_k) / \Delta t_k - f'_k.$$

В большинстве случаев отсчеты берутся через одинаковые интервалы времени. Поэтому будем рассматривать только равномерную сетку Δ , т.е.

$$t_1 - t_0 = t_2 - t_1 = \dots = t_N - t_{N-1} = 1.$$

Для оценки апостериорной дисперсии ошибки оценки была использована известная формула:

$$\sigma_{\hat{x}}^2 = \frac{1}{n-1} \sum_{k=1}^n (\hat{x}(k) - x(k))^2, \quad n = 100.$$

Дисперсия ошибки интерполяции вычислялась на каждом из $k = 10000$ шагов по аналогичной формуле:

$$\sigma_{\hat{S}}^2 = \frac{1}{N-1} \sum_{k=1}^N (\hat{S}(k) - x_0(k))^2, \quad N = 10000,$$

где $\hat{S}(k)$ – значение сплайн-функции на k -м шаге.

По расчетным значениям построены графики. На рис. 1 а представлен график первоначальной выборки $x_0(k)$. Там же тонкой линией изображены оценки сплайн-аппроксимации $\hat{S}(k)$, линией с точками показаны оценки, полученные по процедуре ФКБ. Результаты получены при шаге дискретизации $\Delta t / \tau_{кор} = 1/10$.

На рис. 1 б представлены графики ошибок оценки $\tilde{x}(k)$ и $\Delta x(k)$, полученные с тем же шагом. На рис. 2 б представлены графики $\tilde{x}(k)$ и $\Delta x(k)$, полученные с шагом $\Delta t / \tau_{кор} = 1/2$.

Сопоставление графиков на рис. 1 и 2 показывает, что ошибки оценки при большем шаге (рис.2) заметно больше. Получены расчетные значения апостериорных дисперсий (6) и (7). Эти данные сведены в таблице.

Таблица. Расчетные значения апостериорных дисперсий (6) и (7)

Шаг дискретизации	Апостериорная дисперсия (по ФКБ) $\sigma_{\hat{x}}^2$	Апостериорная дисперсия при сплайн-аппроксимации $\sigma_{\hat{S}}^2$
1/10	4,6756	0,8551
1/2	7,7644	1,7638

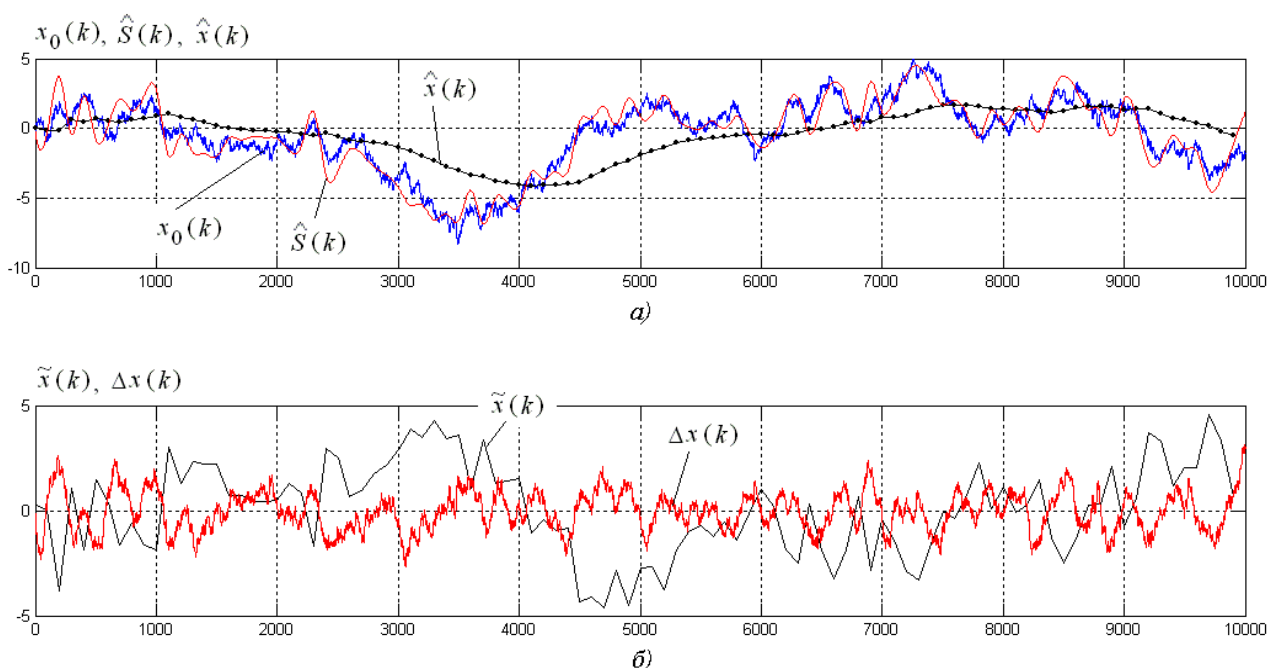


Рис. 1. Графики первоначальной выборки $x_0(k)$, оценки сплайн-аппроксимации $\hat{S}(k)$, оценки $\hat{x}(k)$, полученные по процедуре ФКБ (а) и ошибок оценки $\tilde{x}(k)$, $\Delta x(k)$ при шаге дискретизации $\Delta t / \tau_{кор} = 1/10$ (б)

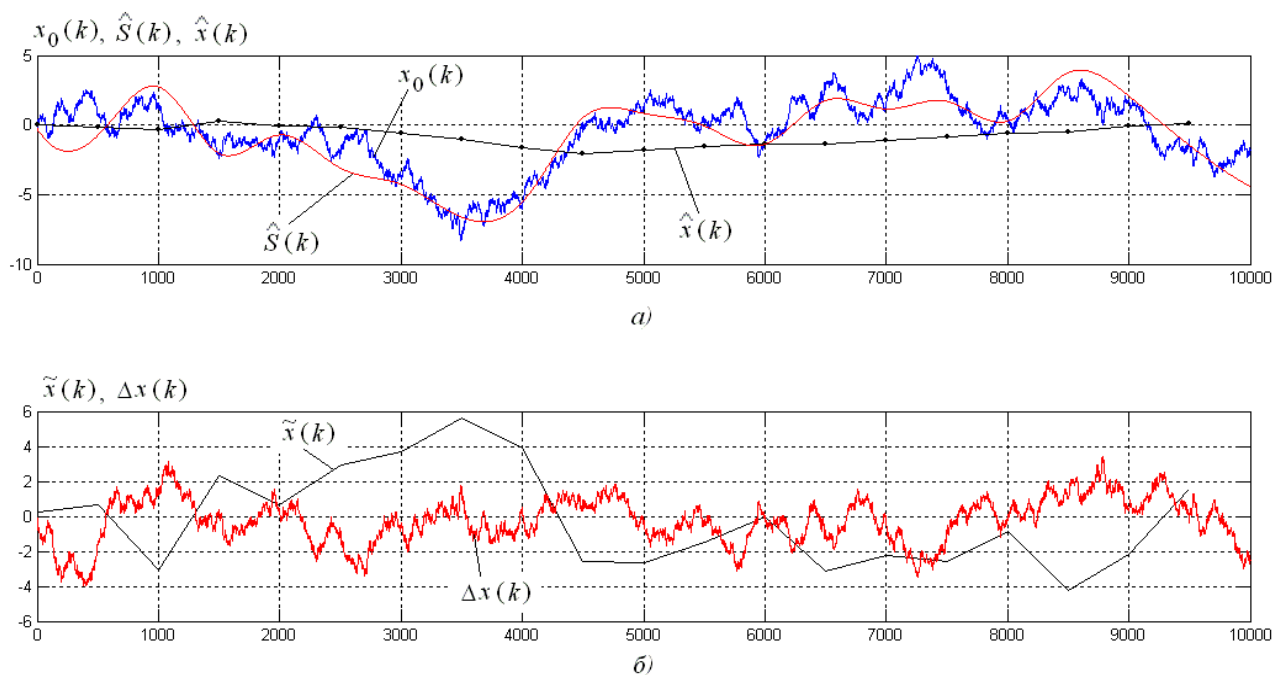


Рис. 2. Графики первоначальной выборки $x_0(k)$, оценки сплайн-аппроксимации $\hat{S}(k)$, оценки $\hat{x}(k)$, полученные по процедуре ФКБ (а) и ошибок оценки $\tilde{x}(k)$, $\Delta x(k)$ при шаге дискретизации $\Delta t / \tau_{кор} = 1/2$ (б)

Из таблицы видно, что при шаге дискретизации $\Delta t / \tau_{кор} = 1/10$ апостериорная дисперсия ошибки оценки, полученная по процедуре ФКБ, более чем в 5 раз превышает соответствующее значение при сплайн-аппроксимациях. Примерно такой же выигрыш имеет место при шаге $\Delta t / \tau_{кор} = 1/2$.

Выводы

1. Рекурсивные оценки фильтрации, получаемые при дискретном варианте ФКБ, зависят как от случайного процесса и шума наблюдения, так и от параметров и особенностей реализации фильтра. При сокращении шага дискретизации Δt по отношению к интервалу корреляции $\tau_{кор}$ точность оценок возрастает.

2. Точность экстраполяции для случайных процессов зависит от интервала корреляции этого процесса $\tau_{кор}$ и от интервала прогноза $\tau_э$. С увеличением $\tau_э$ точность экстраполяции уменьшается или же соответственно апостериорная дисперсия экстраполяции $V(t, \tau_э)$ увеличивается. При $\tau_э \approx 0$ величина $V(t, \tau_э)$ равняется апостериорной дисперсии оценки фильтрации $V(t)$, с увеличением $\tau_э$ дисперсия $V(t, \tau_э)$ возрастает и при $\tau_э > \tau_{кор}$ она достигает значений априорной дисперсии $\sigma_x^2 = N_\xi$. Таким образом, прогноз на время, превышающее интервал корреляции $\tau_{кор}$, не имеет смысла.

3. При интерполяции обычно получают восстанавливаемую функцию (в данном случае реализацию случайного процесса) на определенном интервале, как правило, по последовательности точек, где эта функция определена. В качестве таких

точек можно использовать последовательность получаемых оценок. Наиболее совершенной оптимальной процедурой рекурсивной оценки является ФКБ, позволяющий получать эту последовательность: $\hat{x}(k), \hat{x}(k+1), \hat{x}(k+2)...$ Каждая из этих оценок является точечной. При переходе к интервальной оценке путем интерполяции учитываются интервальные свойства оцениваемой функции. За счет этой дополнительной информации, не учитываемой ФКБ, удается получить дополнительный выигрыш в точности интервальной оценки по сравнению с точечной.

Список литературы:

1. Липцер Р.Ш., Ширяев А.Н. Статистика случайных процессов. – М.: Наука, 1974. – 696 с.
2. Сейдж Э., Мелс Дж. Теория оценивания и ее применение в связи и управлении – М.: Связь, 1976. – 496 с.
3. Поповский В.В., Олейник В.Ф. Математические основы управления и адаптации в телекоммуникационных системах. – Х.: СМІТ, 2011. – 362 с.
4. Родимов А.П., Поповский В.В. Статистическая теория поляризационно-временной обработки сигналов. – М.: Радио и связь, 1984. – 272 с.
5. Поповский В.В., Стрелковская И.В., Бухан Д.Ю. Анализ качества сплайн-аппроксимаций в результатах рекурсивной оценки состояния сетевых элементов и их режимов // Наукові записки УНДІЗ. – 2008. – №3 (5). – С. 29-33.
6. Завьялов Ю.С., Квасов Б.И., Мирошниченко В.Л. Методы сплайн-функций. – М.: Наука, 1980 – 352 с.