Нелинейная интерполяция сглаживающим полиномом на отрезке

Якубовский Е.Г.
e-mail yakubovski@rambler.ru

Аппроксимируем функцию константой при равенстве нулю невязки, получаем квадратное уравнение относительно константы. Получаем две комплексные значения константы. Аппроксимация реализуется на отрезке, содержащем N точек интерполяции. Добавляем линейный член, смещенный на начало отрезка интерполяции. Добиваемся нулевой невязки, откуда определяем коэффициент при линейном члене из квадратного уравнения. При этом из аппроксимируемой функции вычитаем полученное приближение. Корень получается комплексный. Потом проделываем эту же операцию с квадратичным членом, и т.д. Получим аппроксимацию с нулевой невязкой. В случае комплексной аппроксимируемой функции, мнимую часть выбираем ближайшую к значению комплексной аппроксимируемой функции. В случае аппроксимации действительной функции строим значение действительной функции и выбираем из двух значений ближайшую точку к аппроксимируемой функции, таким образом определяем знак квадратного корня, определяющего неизвестный коэффициент. Я надеюсь получить аппроксимацию, не отклоняющуюся от функции в промежуточных точках, так как на каждом шаге получаю значение корня, соответствующее данной точке. Т.е. среднее действительное значение я определяю однозначно, а среднеквадратичное отклонение я выбираю из трех значений мнимой части, положительного, отрицательного или нуля. Но как оказалось данная аппроксимация плохая. Предложена идея аппроксимации с помощью приближения дельта функции, при 10 точках аппроксимации она дает удовлетворительный результат. Но на границе отрезка аппроксимации она отклоняется от аппроксимируемой функции.

Рассмотрим в общем случае комплексную функцию, заданную в N точках, и составим невязку относительно константы и приравняем ее нулю

$$\sum_{m=0}^{N} (f_m - \delta_0)^2 \exp[-(x_m - x_0)^2 / 2\sigma] = 0; n = 0$$

$$c\delta_0^2 - 2\alpha\delta_0 + \beta = 0; \delta_0 = (\alpha \pm i\sqrt{c\beta - \alpha^2}) / N;$$

$$c = \sum_{m=0}^{N} \exp[-(x_m - x_0)^2 / 2\sigma]; \sigma = 1 / N^2$$

$$\alpha = \sum_{m=0}^{N} f_m \exp[-(x_m - x_0)^2 / 2\sigma]; \beta = \sum_{m=0}^{N} f_m{}^2 \exp[-(x_m - x_0)^2 / 2\sigma]$$

Рассмотрим аппроксимацию виде полинома

$$\sum_{m=0}^{N} [f_m - \sum_{k=0}^{n-1} \delta_k (x_m - a)^k - \delta_n (x_m - a)^n]^2 \exp[-(x_m - x_n)^2 / 2\sigma]$$

Определим следующий коэффициент из уравнения

$$c\delta_n^2 - 2\alpha\delta_n + \beta = 0; \delta_n = (\alpha \pm i\sqrt{c\beta - \alpha^2}) / c; |\delta_n| = \sqrt{\beta / c}$$

$$\alpha^2 < c\beta; (\sum_{k=0}^{N} a_k b_k)^2 < \sum_{k=0}^{N} a_k^2 \sum_{k=0}^{N} b_k^2;$$

$$c = \sum_{m=0}^{N} [(x_m - a)^n]^2 \exp[-(x_m - x_n)^2 / 2\sigma]; n = 1,..., N$$

$$\alpha = \sum_{m=0}^{N} [f_m - \sum_{k=0}^{n-1} \delta_k (x_m - a)^k](x_m - a)^n \exp[-(x_m - x_n)^2 / 2\sigma];$$

$$\beta = \sum_{m=0}^{N} [f_m - \sum_{k=0}^{n-1} \delta_k (x_m - a)^k]^2 \exp[-(x_m - x_n)^2 / 2\sigma]$$

Алгоритм создан таким образом, что на каждом шаге n+1 определяется коэффициент с индексом n+1 по функции, зависящей от n+1 координаты.

При этом данная функция аппроксимируется комплексной функцией, с коэффициентами модуль которых убывает

$$f(x) = \sum_{k=0}^{N} \operatorname{Re}\delta_k (x - a)^k + i\sum_{k=0}^{N} \operatorname{Im}\delta_k (x - a)^k$$

Это аппроксимация комплексной функции. Для действительной функции справедливо

$$f(x) = \sum_{k=0}^{N} \operatorname{Re}\delta_k (x - a)^k + \sum_{k=0}^{N} \operatorname{Im}\delta_k (x - a)^k \sin\arg[\sum_{k=0}^{N} \operatorname{Re}\delta_k (x - a)^k + i\sum_{k=0}^{N} \operatorname{Im}\delta_k (x - a)^k]$$

Где действительная часть описывает среднее значение, а мнимая часть описывает среднеквадратичное отклонение Выбирается нулевое, положительное или отрицательное значение мнимой части коэффициента

$$\min_{\operatorname{Im}\delta_n} | F_n(x_n) - f_n | = \min_{\operatorname{Im}\delta_n} | \sum_{k=0}^{n} \operatorname{Re}\delta_k (x_n - a)^k + \sum_{k=0}^{n} \operatorname{Im}\delta_k (x_n - a)^k \sin\varphi_n - f_n |$$

$$\varphi_n = \arg[\sum_{k=0}^{n} \operatorname{Re}\delta_k (x_n - a)^k + i \sum_{k=0}^{n} \operatorname{Im}\delta_k (x_n - a)^k]$$

Я выбрал фазу, не зависящую от множителя у функции. Таким образом я надеюсь достичь минимального отклонения от прямой, соединяющей две соседние точки.

Теперь качественные соображения по поводу ошибки аппроксимации, справедливые для комплексной функции с комплексной аппроксимацией. Для действительной функции данные соображения не проходят. Разбиваем отрезок на N частей, отрезок приводим к значению 1, т.е. расстояние между точками 1/N. Влияние следующей точки на предыдущую при равенстве коэффициентов определяется фактором 1/N, и при N, стремящемся к бесконечности равно нулю. Но коэффициенты с ростом индекса убывают. Коэффициент $|\delta_n|$ надо разделить на среднее арифметическое коэффициентов $|\delta_n|$. Длина отрезка качественной экстраполяции x определяется из равенства $|\delta_n| x^n = 1; n = 1,...,N$, откуда имеем $x = \exp(\frac{\ln 1/|\delta_n|}{n})$. При убывающем коэффициенте $|\delta_n|$ меньше 1, с ростом количестве точек длина отрезка x больше 1, и при бесконечном количестве точек равна 1, т.е. соответствует длине отрезка аппроксимации. Но начальные значения $|\delta_n|$ могут быть большими, и тогда показатель экспоненты будет отрицательный и длина достоверной информации меньше 1. При убывании с ростом индекса коэффициента $|\delta_n|$ он может достигнуть 1, и тогда длина достоверной информации достигает 1. Но при коэффициенте $|\delta_n|$ меньше 1, длина достоверной информации больше 1. Чем меньшее значение имеет $|\delta_n|$, тем больше его вклад в достоверную информацию. Надо учитывать все значения $|\delta_n|$, т.е. все отрезки достоверной информации. Следовательно, время достоверного предсказания будущего, равно длине аппроксимации при самых благоприятных условиях. Но тут надо сказать, что аппроксимация должна включать все минимумы и максимумы аппроксимируемой функции, иначе выборка будет не достоверная. Если имеются скрытые максимумы и минимумы, то выборка также будет не достоверная.

Но честно говоря, чтобы обнаружить все подводные камни алгоритма, надо его запрограммировать и провести численные эксперименты. Пока это только перспективные теоретические идеи, которые могут рассыпаться при программировании и численных экспериментах.

Задачу лучше свести к комплексной, где мнимая часть у решения определяет ошибку действительной части. Тогда знак у мнимой части надо определять по действительной части параметра.

Но надо сказать, что алгоритм будет предсказывать будущее, если имеется связь между настоящим и будущим. Если этой связи нет, то алгоритм будет врать. Связь должна задаваться либо дифференциальным уравнением, либо общей функциональной зависимостью. Если просто просто задать значения функций, и по первой части этого задания, определять вторую, то ничего не получится, связи нет, и алгоритм врет.

Я произвел вычисления на алгоритмическом языке Mathcard. Изменил алгоритм, взяв весовую функцию у невязки, выделив точку аппроксимации. Взял отрезок длиной 1, разбил его на 10 частей и подсчитал с помощью полинома 10 степени. Графики не совпадают и приближенно не аппроксимируют кривую. Коэффициенты с ростом индекса не убывают. Значит предсказательной силы нет. Попробовал аппроксимацию с весом

$$f(x) = \sum_{m=0}^{M} f_m \frac{\exp[-(x_m - x)^2 / 2\sigma^2]}{\sqrt{2\pi}}; \sigma = 1/N; x \in [0,1]; x_m \in [0,1]$$

При 10 точках аппроксимации получается приближенная аппроксимация, при 100 точках графики аппроксимируемой функции и аппроксимирующей почти совпадают. Причем на границе отрезка точность аппроксимации меньшая. Для увеличения точности на границе надо использовать формулу

$$f(x) = \langle f \rangle + \sum_{m=0}^{M} [f_m - \langle f \rangle] \frac{\exp\{-(x_m - x)^2 / [2\sigma^2 \exp[k(x-0.5)^4]]\}}{\sqrt{2\pi}}; \langle f \rangle = \sum_{m=0}^{M} \frac{f_m}{M+1}$$

Но предсказательной силы она не имеет, функция вне отрезка затухает, вернее выходит на константу, равную среднему арифметическому

$$f(x) = \sum_{m=0}^{M} \frac{f_m}{M+1}$$

С вариациями ближе к границе. Коэффициент k определяется из условия

$$\int_0^1 \sum_{m=0}^{M} \frac{\exp\{-(x_m - x)^2 / [2\sigma^2 \exp[k(x-0.5)^4]]\}}{\sqrt{2\pi}} dx = 1$$

Значение коэффициента k=8.5 при M=10 точкам аппроксимации $\sum_{n=0}^{N} \sum_{m=0}^{10} \frac{\exp\{-(x_m - x_n)^2 / [2\sigma^2 \exp[k(x_n-0.5)^4]]\}}{101\sqrt{2\pi}} = 1$ и значение коэффициента k=1.2 при M=100 точки аппроксимации

$$\sum_{n=0}^{N} \sum_{m=0}^{M} \frac{\exp\{-(x_m - x_n)^2 / [2\sigma^2 \exp[k(x_n-0.5)^4]]\}}{101\sqrt{2\pi}} = 1; M = 100; N = 100 .$$

Количество точек аппроксимируемых функций, по которым берется среднее арифметическое в обоих случаях одинаковое большое, что соответствует интегралу от 0 до 1.