

Тема: Прямая задача теории погрешности

Суть - оценивание погрешности произвольной функции нескольких переменных через заданные погрешности аргументов.

Пусть $u=f(x_1, x_2, \dots, x_n)$ – дифференцируемая функция,

Δx_i - погрешности аргументов.

Тогда абсолютная погрешность функции

$$\Delta u = \left| f(x_1 + \Delta x_1, x_2 + \Delta x_2, \dots, x_n + \Delta x_n) - f(x_1, x_2, \dots, x_n) \right|$$

Тогда можно записать:

$$|\Delta u| \approx \left| df(x_1, x_2, \dots, x_n) \right| = \left| \sum_{i=1}^n \frac{\partial f}{\partial x_i} \Delta x_i \right| \leq \sum_{i=1}^n \left| \frac{\partial f}{\partial x_i} \right| \cdot |\Delta x_i| \quad (1)$$

Отсюда для предельной абсолютной погрешности функции получаем:

$$\Delta_u = \sum_{i=1}^n \left| \frac{\partial u}{\partial x_i} \right| \cdot \Delta_{x_i}$$

Оценка относительной погрешности u : // обе части неравенства (1) делим на u

$$\delta \leq \sum_{i=1}^n \left| \frac{\frac{\partial f}{\partial x_i}}{u} \right| \cdot |\Delta x_i| = \sum_{i=1}^n \left| \frac{\partial}{\partial x_i} \cdot \ln f(x_1, \dots, x_n) \right| |\Delta x_i|$$

Следовательно, предельная относительная погрешность функции u :

$$\delta_u = \sum_{i=1}^n \left| \frac{\partial}{\partial x_i} \cdot \ln(u) \right| \cdot \Delta_{x_i}$$

Прямая задача теории погрешности. Пример.

Найти предельную абсолютную и относительную погрешности объёма шара

$$V = \frac{1}{6}\pi d^3 \quad , \text{ если диаметр } d = 3,7 \text{ см } \pm 0,05 \text{ см, а } \pi \approx 3,14.$$

Решение. Рассматривая π и d как переменные величины, вычислим частные производные

$$\frac{\partial V}{\partial \pi} = \frac{1}{6}d^3 = 8,44 ;$$

$$\frac{\partial V}{\partial d} = \frac{1}{2}\pi d^2 = 21,5 .$$

Тогда предельная погрешность объёма:

$$\begin{aligned} \Delta v &= \left| \frac{\partial V}{\partial \pi} \right| |\Delta \pi| + \left| \frac{\partial V}{\partial d} \right| |\Delta d| = \\ &= 8,44 \cdot 0,0016 + 21,5 \cdot 0,05 = 1,088 \text{ см}^3 \approx 1,1 \text{ см}^3 \end{aligned}$$

$$\text{Поэтому} \quad V = \frac{1}{6}\pi d^3 \approx 26,51 \text{ см}^3 \pm 1,1 \text{ см}^3 .$$

Предельная относительная погрешность

$$\text{объёма:} \quad \delta_v = \frac{1,088 \text{ см}^3}{26,51 \text{ см}^3} = 0,041 \approx 4\% .$$

Обратная задача теории погрешности

Суть - оценивание погрешностей аргументов произвольной функции нескольких переменных, обеспечивающих заданную или меньшую погрешность функции.

В такой постановке задача некорректна, так как допускает множество разных решений. Нужны дополнительные предположения.

Принцип равных влияний. Все частные дифференциалы

$$\frac{\partial f}{\partial x_i} \Delta x_i \quad (i = 1, 2, \dots, n)$$

вносят одинаковый вклад в образование общей абсолютной погрешности функции.

Пусть величина абсолютной погрешности функции Δ_u задана. Воспользуемся формулой предыдущего раздела:

$$\Delta_u = \sum_{i=1}^n \left| \frac{\partial u}{\partial x_i} \right| \cdot \Delta_{x_i}$$

Согласно принципу равных влияний:

$$\left| \frac{\partial u}{\partial x_1} \right| \Delta_{x_1} = \left| \frac{\partial u}{\partial x_2} \right| \Delta_{x_2} = \dots = \left| \frac{\partial u}{\partial x_n} \right| \Delta_{x_n} = \frac{\Delta_u}{n}$$

Отсюда получаем:

$$\Delta_{x_i} = \frac{\Delta_u}{n \cdot \left| \frac{\partial u}{\partial x_i} \right|}$$

(2)

Обратная задача теории погрешности. Пример.

Пример. Радиус основания цилиндра $R \approx 2$ м. Высота цилиндра $H \approx 3$ м. С какими абсолютными погрешностями нужно определить R и H , чтобы его объем V можно было вычислить с точностью до 0.1 м^3 ?

Решение. Имеем $V = \pi R^2 H$ и $\Delta_V = 0,1 \text{ м}^3$.

Полагая $R = 2$ м; $H = 3$ м; $\pi = 3,14$; приближенно получим:

$$\begin{aligned}\frac{\partial V}{\partial \pi} &= R^2 H = 12 \\ \frac{\partial V}{\partial R} &= 2\pi R H = 37,7 \\ \frac{\partial V}{\partial H} &= \pi R^2 = 12,6.\end{aligned}$$

Отсюда, так как $n = 3$, то на основании формулы (2) будем иметь:

$$\begin{aligned}\Delta_\pi &= \frac{0,1}{3 \cdot 12} < 0,003; \\ \Delta_R &= \frac{0,1}{3 \cdot 37,7} < 0,001; \\ \Delta_H &= \frac{0,1}{3 \cdot 12,6} < 0,003.\end{aligned}$$

Тема: Вычисление значений функций

2. Вычисление значения алгебраического полинома.

Схема Горнера.

Рассмотрим полином

$$P(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0$$

Наша задача – найти значение этого полинома при $x = \xi$. То есть вычислить $P(\xi) = a_n \xi^n + a_{n-1} \xi^{n-1} + \dots + a_1 \xi + a_0$

Для рационального вычисления переписем $P(\xi)$ в виде

$$P(\xi) = (\dots (((a_n \xi + a_{n-1}) \xi + a_{n-2}) \xi + \dots + a_2) \xi + a_1) \xi + a_0$$

Отсюда легко сконструировать нужную последовательность вычислений, начиная с самых внутренних скобок:

$$b_0 = a_n;$$

$$b_1 = a_{n-1} + b_0 \xi;$$

$$b_2 = a_{n-2} + b_1 \xi;$$

.....

$$b_{n-1} = a_1 + b_{n-2} \xi;$$

$$b_n = a_0 + b_{n-1} \xi.$$

□ это искомое значение: $b_n = P(\xi)$

Вычисление значений функций. Схема Горнера.

Сравним эффективность предложенного варианта вычислений (схемы Горнера) с прямым вычислением «в лоб». В качестве критерия будем использовать количество операций умножения.

При вычислении по формуле нам потребуется выполнить $1+2+3+\dots+n=n(n+1)/2$ операций умножения. А по схеме Горнера только n операций.

Тема: Вычисление значений функций

Покажем, что числа $b_0, b_1, \dots, b_{n-1}$ являются коэффициентами полинома $Q(x)$, полученного в качестве частного при делении данного полинома $P(x)$ на двучлен $x - \xi$.

$$Q(x) = \beta_0 x^{n-1} + \beta_1 x^{n-2} + \dots + \beta_{n-1}$$

$$P(x) = (\beta_0 x^{n-1} + \beta_1 x^{n-2} + \dots + \beta_{n-1})(x - \xi) + \beta_n$$

$$P(x) = \beta_0 x^n + (\beta_1 - \beta_0 \xi)x^{n-1} + (\beta_2 - \beta_1 \xi)x^{n-2} + \dots + (\beta_{n-1} - \beta_{n-2} \xi)x + (\beta_n - \beta_{n-1} \xi)$$

$$P(x) = a_n x^n + a_{n-1} x^{n-1} + a_{n-2} x^{n-2} + \dots + a_1 x + a_0$$

$$\beta_0 = a_n,$$

$$\beta_1 - \beta_0 \xi = a_{n-1},$$

$$\beta_2 - \beta_1 \xi = a_{n-2},$$

$$\dots \dots \dots$$

$$\beta_{n-1} - \beta_{n-2} \xi = a_1,$$

$$\beta_n - \beta_{n-1} \xi = a_0.$$

$$\beta_0 = a_n = b_0,$$

$$\beta_1 = a_{n-1} + \beta_0 \xi = b_1$$

$$\beta_2 = a_{n-2} + \beta_1 \xi = b_2,$$

$$\dots \dots \dots$$

$$\beta_{n-1} = a_1 + \beta_{n-2} \xi = b_{n-1},$$

$$\beta_n = a_0 + \beta_{n-1} \xi = b_n,$$

$$b_0 = a_n,$$

$$b_1 = a_{n-1} + b_0 \xi,$$

$$b_2 = a_{n-2} + b_1 \xi,$$

$$b_3 = a_{n-3} + b_2 \xi,$$

$$\dots \dots \dots$$

$$b_{n-1} = a_1 + b_{n-2} \xi,$$

$$b_n = a_0 + b_{n-1} \xi,$$

Вычисление значений функций.

2. Приближенное суммирование числовых рядов

Определение. Числовой ряд $a_1 + a_2 + \dots + a_n + \dots$

называется сходящимся, если существует предел последовательности его частных сумм $S = \lim_{n \rightarrow \infty} S_n$, где $S_n = a_1 + a_2 + \dots + a_n$.

Задача. Найти сумму S сходящегося числового ряда $\sum_{i=1}^{\infty} a_i$ с нужной точностью ε (ε – абсолютная погрешность).

Решение. Сумма сходящегося ряда всегда представляется в виде:

$$S = S_n + R_n,$$

где S_n – n -ая частичная сумма, $R_n \rightarrow 0$ при $n \rightarrow \infty$
 R_n – остаток ряда, причем

Остаток ряда характеризует остаточную погрешность.

Источники погрешности при суммировании сходящегося ряда:

1. остаточная погрешность,
2. погрешность вычисления самих элементов ряда,
3. погрешность при округлении результата.

Вычисление значений функций.

2. Приближенное суммирование числовых рядов Процедура вычислений.

Выбираем три положительных числа $\varepsilon_1, \varepsilon_2, \varepsilon_3$, таких, что

$$\varepsilon_1 + \varepsilon_2 + \varepsilon_3 = \varepsilon$$

1) остаточная погрешность (ε_1):

Количество 'n' слагаемых в частичной сумме S_n выберем столь большим, чтобы остаточная погрешность

$$|R_n| < \varepsilon_1$$

2) погрешность вычисления самих элементов ряда (ε_2):

Каждое из слагаемых a_i частичной суммы вычисляем с предельной абсолютной погрешностью, не превышающей $\frac{\varepsilon_2}{n}$ и пусть $\tilde{a}_k$ - соответствующие приближенные значения членов ряда, и выполняется:

$$|\tilde{a}_k - a_k| \leq \frac{\varepsilon_2}{n}$$

Тогда для частичной суммы

$$\bar{S}_n = \sum_{k=1}^n \bar{a}_k$$

погрешность действий (суммирования) удовлетворяет неравенству

$$|S_n - \bar{S}_n| \leq \varepsilon_2$$

3) **Погрешность при округлении результата (ε_3).**

Результат $\bar{S}_n$ округлим до $\bar{\bar{S}}_n$ так, чтобы для погрешности округления результата выполнялось соотношение:

$$|\bar{S}_n - \bar{\bar{S}}_n| \leq \varepsilon_3$$

При выполнении такой процедуры суммирования ряда, полученный результат будет соответствовать искомой сумме с погрешностью не более ε .

В самом деле,

$$|S - \bar{\bar{S}}_n| \leq |S - S_n| + |S - \bar{S}_n| + |\bar{S}_n - \bar{\bar{S}}_n| \leq \varepsilon_1 + \varepsilon_2 + \varepsilon_3 = \varepsilon.$$

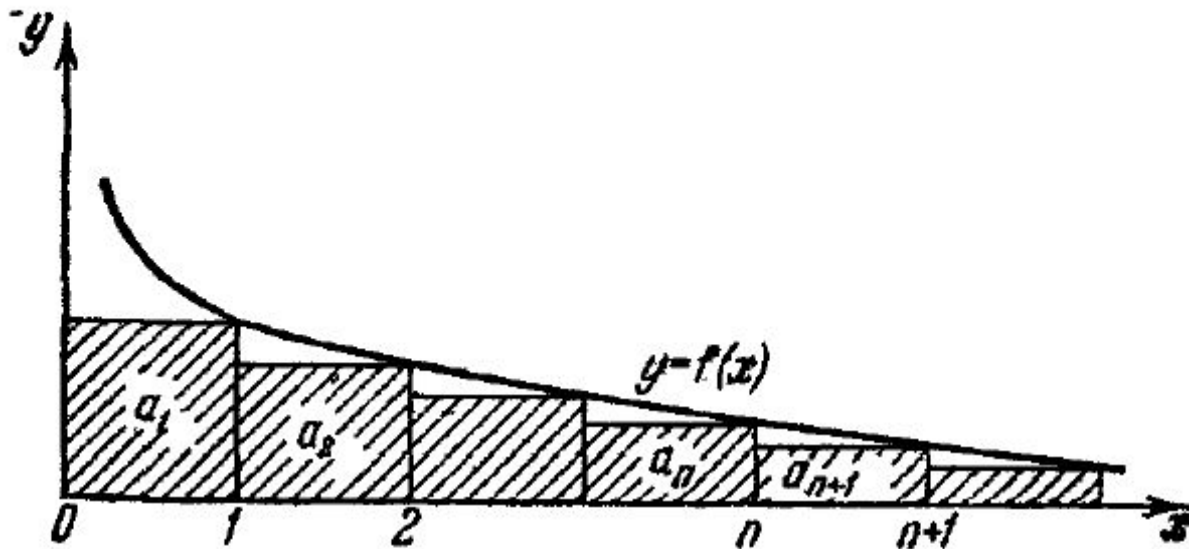
Разбиение числа ε на положительные слагаемые $\varepsilon_1, \varepsilon_2, \varepsilon_3$ еще всего производят следующим образом:

$$\varepsilon_1 = \frac{\varepsilon}{4}, \varepsilon_2 = \frac{\varepsilon}{4}, \varepsilon_3 = \frac{\varepsilon}{2}$$

Теорема 1. Если члены ряда представляют собой соответствующие значения положительной монотонно убывающей функции $f(x)$, то есть

$$a_n = f(n) \quad (n = 1, 2, \dots) \quad , \text{ тогда}$$

$$\int_1^{\infty} f(x) dx < R_n < \int_n^{\infty} f(x) dx$$



Теорема 2. Если ряд знакочередующийся и модули его членов монотонно убывают, тогда

$$|R_n| < |a_{n+1}|$$

Вычисление значений функций.

2. Приближенное суммирование числовых рядов

Пример.

Найти сумму ряда $S = \frac{1}{1^3} + \frac{1}{2^3} + \frac{1}{3^3} + \dots + \frac{1}{n^3} + \dots$ с точностью 0,004.

Решение. Округлять результат и слагаемые не будем $\varepsilon_2 = \varepsilon_3 = 0$.

Тогда погрешность вычисления суммы равна остаточной погрешности:

$$\varepsilon = \varepsilon_1 = 0,004$$

Члены ряда представляют собой значения монотонно убывающей функции

$$f(x) = \frac{1}{x^3}$$

Поэтому для n -го остатка ряда имеем оценку (по теореме 1):

$$R_n \leq \int_n^{\infty} \frac{dx}{x^3} = \frac{1}{2n^2}$$

Решая неравенство: $\frac{1}{2n^2} < \frac{1}{4000},$

$$n \geq \sqrt{2000} \approx 44,7.$$

Получим

Следовательно, $n=45$.

Вычисление значений функций.

3. Вычисление значений аналитических функций

Функция $f(x)$ аналитическая в точке $x = \xi$, если в окрестности этой точки она разлагается в ряд Тейлора, то есть ее можно представить в виде (1)

$$f(x) = f(\xi) + f'(\xi)(x - \xi) + \frac{f''(\xi)}{2!}(x - \xi)^2 + \dots + \frac{f^{(n)}(\xi)}{n!}(x - \xi)^n + \dots$$

Остаточный член разложения $R_n(x)$ представляет собой ошибку при замене функции отрезком ряда Тейлора:

$$R_n(x) = f(x) - \sum_{k=0}^n \frac{f^{(k)}(\xi)}{k!}(x - \xi)^k$$

Для остаточного члена существует несколько форм представления. Одна из них такая

$$R_n(x) = \frac{f^{(n+1)}(\xi + \theta(x - \xi))}{(n+1)!}(x - \xi)^{n+1}, \quad 0 < \theta < 1$$

Вычисление значений функций.

3. Вычисление значений аналитических функций

Если $f(\xi)$ известно и требуется найти значение $f(\xi + h)$, где h – «малая поправка», то формулу (1) выгодно записать в виде:

$$f(\xi + h) = f(\xi) + f'(\xi) \cdot h + \frac{f''(\xi)}{2!} \cdot h^2 + \dots + \frac{f^{(n)}(\xi)}{n!} \cdot h^n + R_n(h).$$

где

$$R_n(x) = \frac{f^{(n+1)}(\xi + \theta \cdot x)}{(n+1)!} h^{n+1}, \quad 0 < \theta < 1.$$

Вычисление экспоненты (e^x).

Известно из курса математического анализа, что экспоненциальная функция в окрестности точки 0 разлагается в ряд Тейлора вида:

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots + \frac{x^n}{n!} + \dots$$

Остаточный член этого ряда имеет вид: $R_n(x) = \frac{e^{\theta x}}{(n+1)!} x^{n+1} \quad 0 < \theta < 1.$

Непосредственно в таком виде использовать представление экспоненты неэффективно, так как x^{n+1} сильно растет, когда $|x| > 1$. Для достижения требуемой точности придется использовать много членов разложения (большая трудоемкость). Для обхода этих затруднений предлагается следующая процедура.

Представим аргумент x в виде: $x = E(x) + q,$

$E(x)$ – целая часть числа x ,

q – дробная часть ($|q| < 1$). Тогда

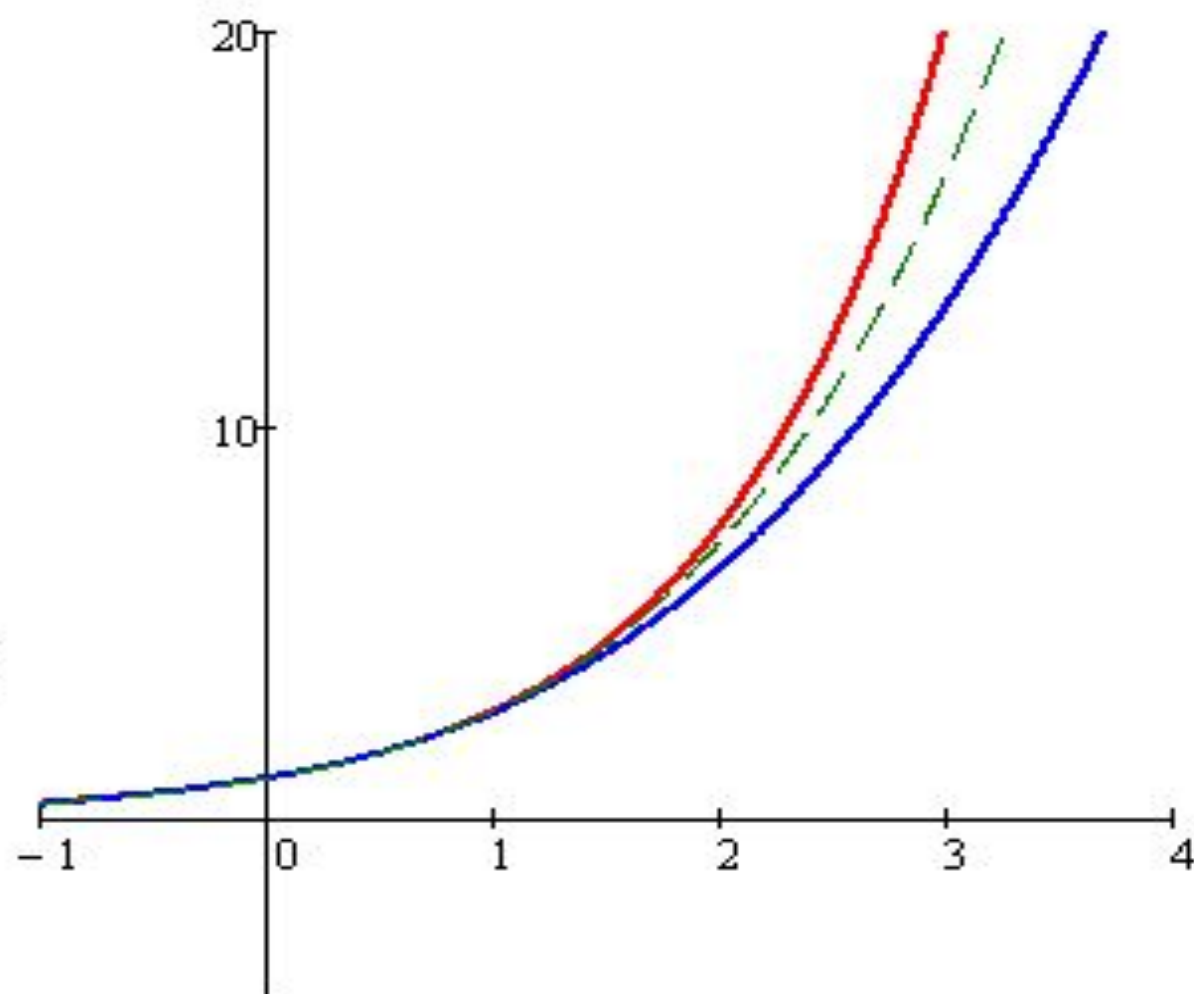
$$e^x = e^{E(x)} \cdot e^q$$

$$f(x) := e^x$$

$f(x)$

$$1 + x + \frac{x^2}{2} + \frac{x^3}{6}$$

$$1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \frac{x^4}{24}$$



$$f(x) \text{ series, } x = 0,9 \rightarrow 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \frac{x^4}{24} + \frac{x^5}{120} + \frac{x^6}{720} + \frac{x^7}{5040} + \frac{x^8}{40320}$$

Число $e = 2.71828182\dots$ известно с очень высокой точностью. Поэтому величину $e^{E(x)} = e^*e^*e^*\dots*e$ можно считать вычисляемой с какой угодно точностью. Второй сомножитель e^q будем вычислять, пользуясь разложением в ряд Тейлора. Очевидно, что при $|q| < 1$ сходимость ряда

$$e^q = \sum_{n=0}^{\infty} \frac{q^n}{n!}$$

будет быстрой, то есть потребуется мало членов разложения.

Для величины остаточного члена можно получить более точную оценку. В самом деле, по определению остатка, можем записать

$$R_n(q) = \frac{q^{n+1}}{(n+1)!} + \frac{q^{n+2}}{(n+2)!} + \frac{q^{n+3}}{(n+3)!} + \dots =$$

$$\frac{q^{n+1}}{(n+1)!} \left[1 + \frac{q}{(n+2)} + \frac{q^2}{(n+2)(n+3)} + \dots \right] < \frac{q^{n+1}}{(n+1)!} \frac{1}{1 - \frac{q}{n+2}}$$

Если учесть, что $\frac{n+2}{n+1} < \frac{n+1}{n}$, то получим $R_n(q) < \frac{q^{n+1}}{(n+1)!} \frac{n+1}{n} = \frac{q^{n+1}}{n n!}$

Если обозначить $u_n = \frac{q^n}{n!}$, u_n - последний сохраненный член частичной суммы, тогда для остаточного члена можно записать

$$0 < R_n(q) < u_n \frac{q}{n}$$

Если нам задана остаточная погрешность (ε_r), тогда количество слагаемых частичной суммы ряда легко определить в процессе вычисления. Находить новые слагаемые и суммировать их можно прекратить, как только для очередного слагаемого выполниться условие $u_n \frac{q}{n} < \varepsilon_r$

Таким образом вычисление экспоненциальной функции сводится

$$e^x = u_0 + u_1 + u_2 + \dots + u_n, \text{ где } u_0 = 1, u_k = xu_{k-1}/k.$$

Предложенный ранее вариант определения количества слагаемых можно уточнить (дать более точные условия прекращения суммирования) если предположить, что выполняются следующие нежесткие условия, а именно: $n \geq 2$ $|x| > 0$. В этом случае

$$|R_n(x)| \leq R_n(|x|) < \frac{|x|^{n+1}}{(n+1)!} \frac{1}{1 - \frac{|x|}{n+2}}$$

С учетом принятого предположения, вторая дробь не может быть больше 2 Тогда,

$$R_n < \frac{2|x|^{n+1}}{(n+1)!} = \frac{2|x|}{n+1} \frac{|x|^n}{n!} < |u_n| \leq \varepsilon_r$$

Отсюда вывод, процесс суммирования можно прекратить, если последний вычисленный член u_n по модулю не превосходит ε_r , то есть $|u_n| < \varepsilon_r$.