

Układy równań nieliniowych

Ogólna teoria

- W przypadku układów równań nieliniowych możemy zastosować podobne podejście jak dla pojedynczych równań.

$$f(x) = 0, \quad f: \mathbb{R}^n \rightarrow \mathbb{R}^n$$

- Koncepcja opiera się o stworzenie modelu zmiany residuum w związku ze zmianą

$$x_{k+1} = x_k + p$$

- W związku z tym modelujemy $r(x_k + p)$

$$r(x_k + p) \approx r(x_k) + J_k p$$

Metoda Newtona

- Idea polega na znalezieniu takiego p , żeby $r(x_k + p) = 0$, tj.

$$J_k p = -r(x_k)$$

- J_k powinno być macierzą pochodnych cząstkowych (macierzą Jakobiego) w punkcie x_k

$$J_k = J(x_k) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_n} \end{bmatrix}$$

- Lub jej aproksymacją (inne metody)

Zbieżność

- Jeżeli jesteśmy blisko rozwiązania a funkcja f jest różniczkowalna, a pochodna spełnia warunek Lipshitz, to mamy zbieżność kwadratową

$$x_{k+1} - x^* = O(\|x_{k+1} - x^*\|^2)$$

Niedokładne metody Newtona

- Istnieje klasa metod, która wyznacza kolejny krok rozwiązania p jako

$$\|r_k + J_k p_k\| \leq \eta_k \|r_k\|$$

gdzie $\eta_k < \eta$, $\eta \in [0,1]$ to tzw. ciąg wymuszający

Metoda Broydena

- Metoda Broydena jest uogólnieniem metody siecznych
- Niech $s_k = x_{k+1} - x_k$, $y_k = r(x_{k+1}) - r(x_k)$
- Konstruujemy aproksymację macierzy Jakobiego B_k , która ma spełniać tzw. warunek siecznej

$$y_k = B_{k+1}s_k$$

Wzór Broydena

- Najpopularniejszą aproksymacją jest tzw. wzór Broydena

$$B_{k+1} = B_k + \frac{(y_k - B_k s_k) s_k^T}{s_k^T s_k}$$

Jest to najmniejsza możliwa zmiana w aproksymacji przy spełnieniu warunku siecznej.

Metody praktyczne – funkcja celu

- Metody Newtonowskie i quasi-Newtonowskie są zbieżne jedynie lokalnie.
- Aby uniknąć tego problemu, zaproponowano aby korygować długość kroku między iteracjami tj.

$$x_{k+1} = x_k + \varepsilon p$$

- Długość kroku, dobieramy tak aby poprawić wartość funkcji celu

Funkcja celu – suma kwadratów residuów

- Najpopularniejszą funkcją celu jest suma kwadratów residuów

$$m(x) = \frac{1}{2} \sum_{i=1}^n r_i^2(x)$$

- Z oczywistych względów pierwiastek równania jest minimum funkcji celu

Poszukiwanie kierunkowe

- Podsumowując, kolejne kroki rozwiązania znajdujemy zmniejszając wartość funkcji celu.
- Nie jest konieczne poszukiwanie dokładne minimum
- Wystarczy spełnienie tzw. warunków Wolfe'a, w uproszczeniu:
 - Nie robimy długich kroków, chyba że jest to bardzo uzasadnione
 - Jeżeli funkcja po naszym kroku dalej silnie maleje, to trzeba zrobić dłuższy krok

Problem

- Podobnie jak w jednowymiarowej metodzie Newtona problemem jest jak pochodna się zeruje.
- W tym przypadku, pochodna (gradient) funkcji celu wynosi:

$$\nabla m(x) = J(x_k)^T r(x_k)$$

- Oznacza to, że jeżeli J jest osobliwa, to dla niezerowych residuuów możliwe jest istnienie minimum funkcji celu.

Metody regionu zaufania (trust region)

- Aby uniknąć problemów ze złymi kierunkami i osobliwością zaproponowano metodę mocniej ugruntowaną w optymalizacji.
- Wyznaczamy następny krok nie na kierunku tylko wewnątrz pewnego zbioru (regionu zaufania), w którym jesteśmy pewni naszego modelu.
- Zbiór aktualizujemy z iteracji na iterację
- Jak zbliżamy się do rozwiązania przechodzi to w metodę Newtona, ale uniemożliwia ucieczkę.

Metody w Pythonie - `scipy.optimize.root`

- Parametr 'method'
- *hybr* - modyfikacja metody Powella (trust region)
- *lm* - nieliniowy problem najm. kw. Metodą Levenberga-Marquardta (zasadniczo trust region)
- *df-sane* - metoda spektralna bezgradientowa (zasadniczo interpolacja)
- *broyden1*, *broyden2*, *anderson*, *linearmixing*, *diagbroyden*, *excitingmixing*, *krylov* różne warianty metody Newtona

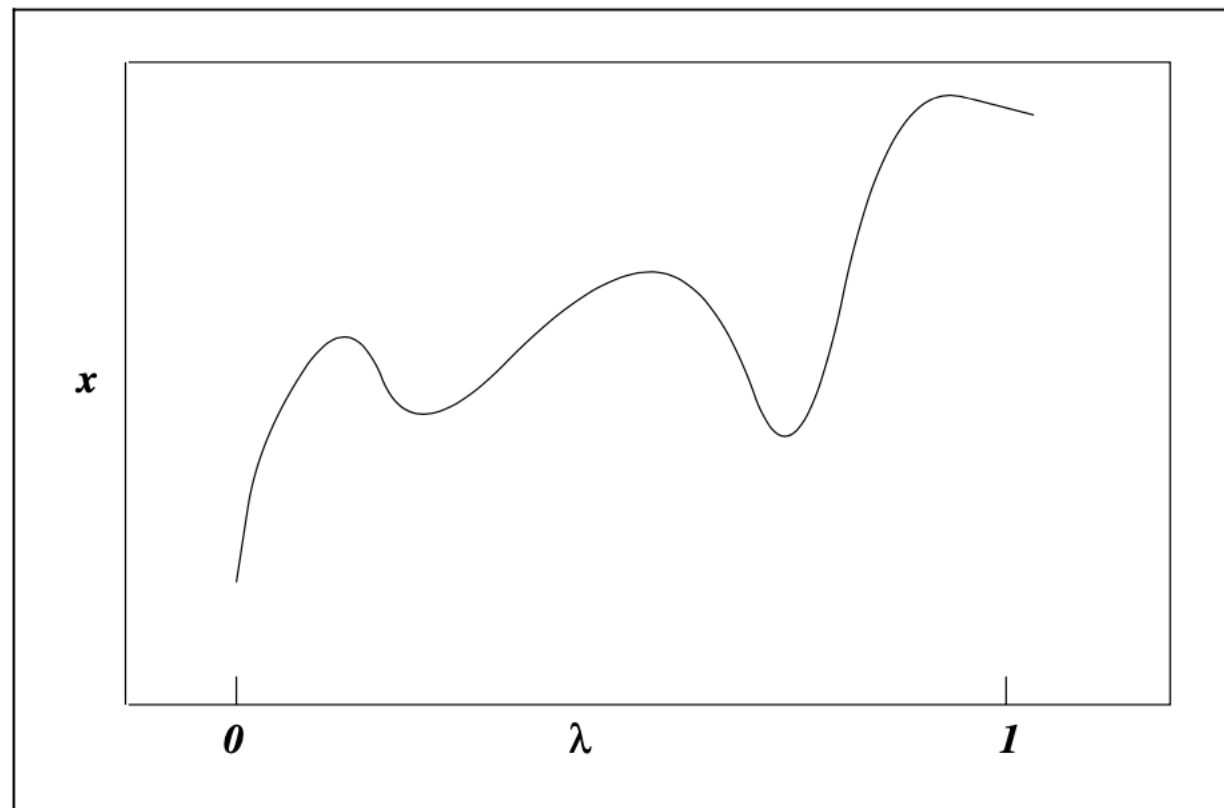
Alternatywa – homotopia/przedłużanie

- Pomysł polega na stopniowym przechodzeniu od problemu prostszego do trudniejszego

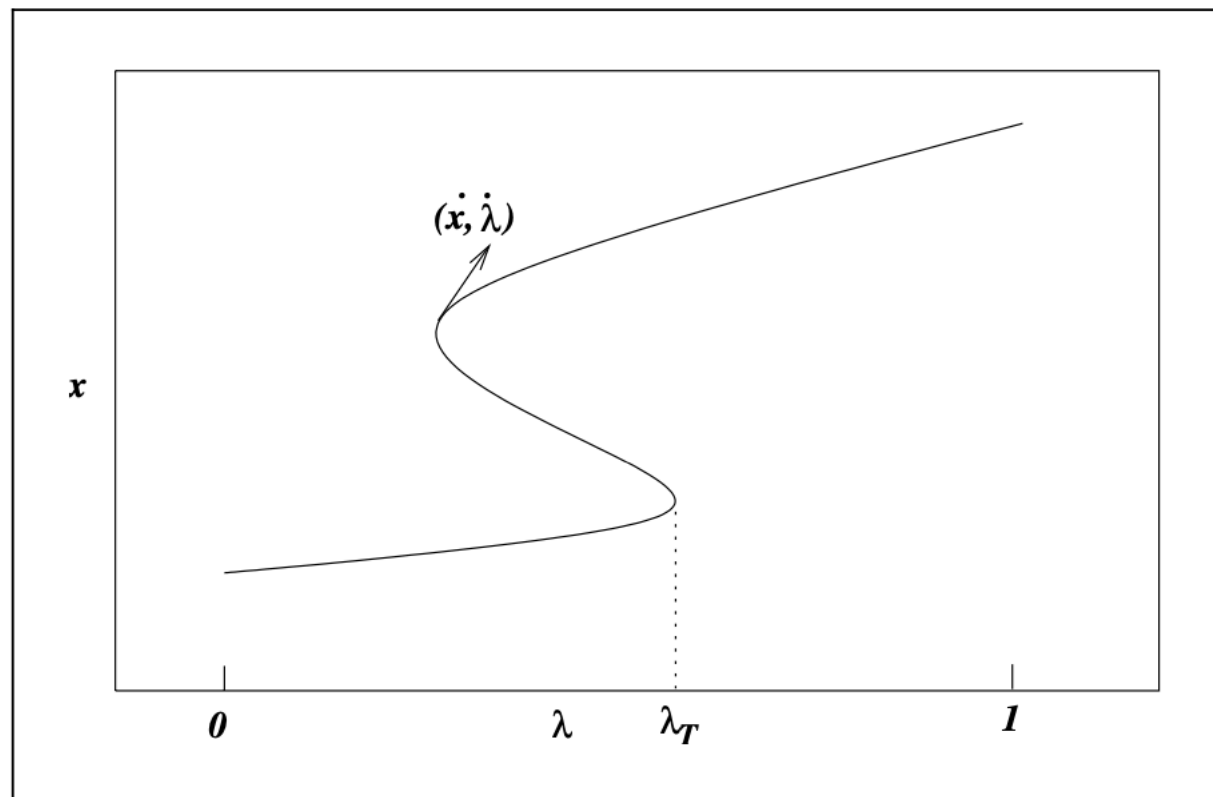
$$H(x, \lambda) = \lambda r(x) + (1 - \lambda)(x - a)$$

- $\lambda \in [0,1]$, zaczynamy od prostego problemu dla $\lambda = 0$ i stopniowo go zmieniając zmierzamy do rozwiązania

Jeśli jest dobrze



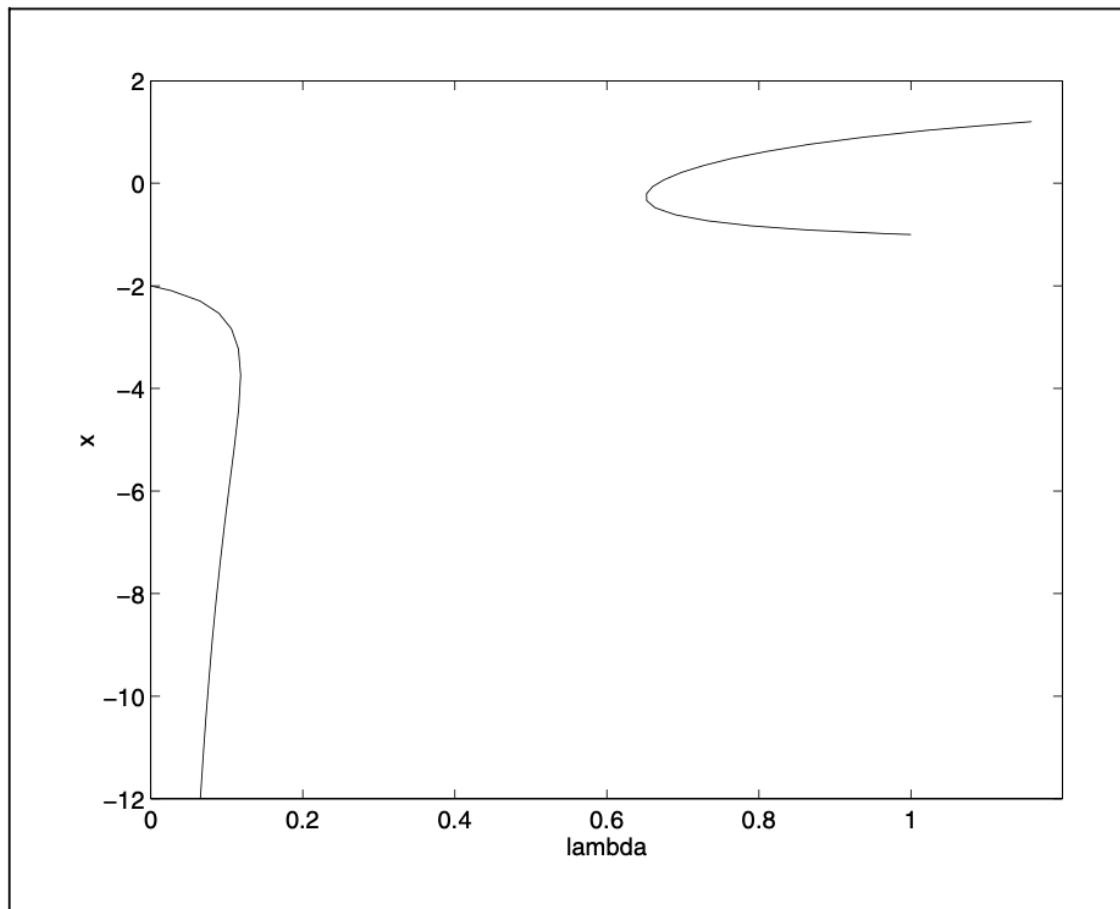
Czasem gorzej



W praktyce

- Wyznaczamy wektor styczny do homotopii
- W oparciu o niego konstruujemy albo równania różniczkowe albo algebraiczne, które trzymają nas dalej na ścieżce

Najgorzej



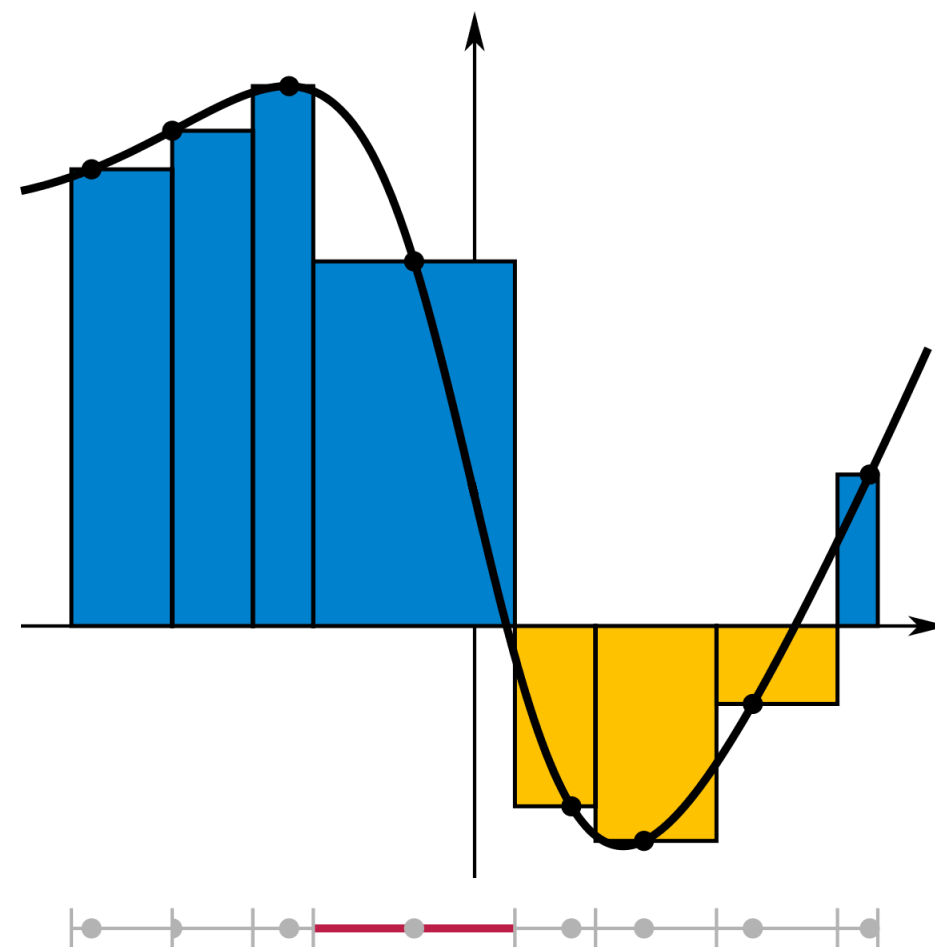
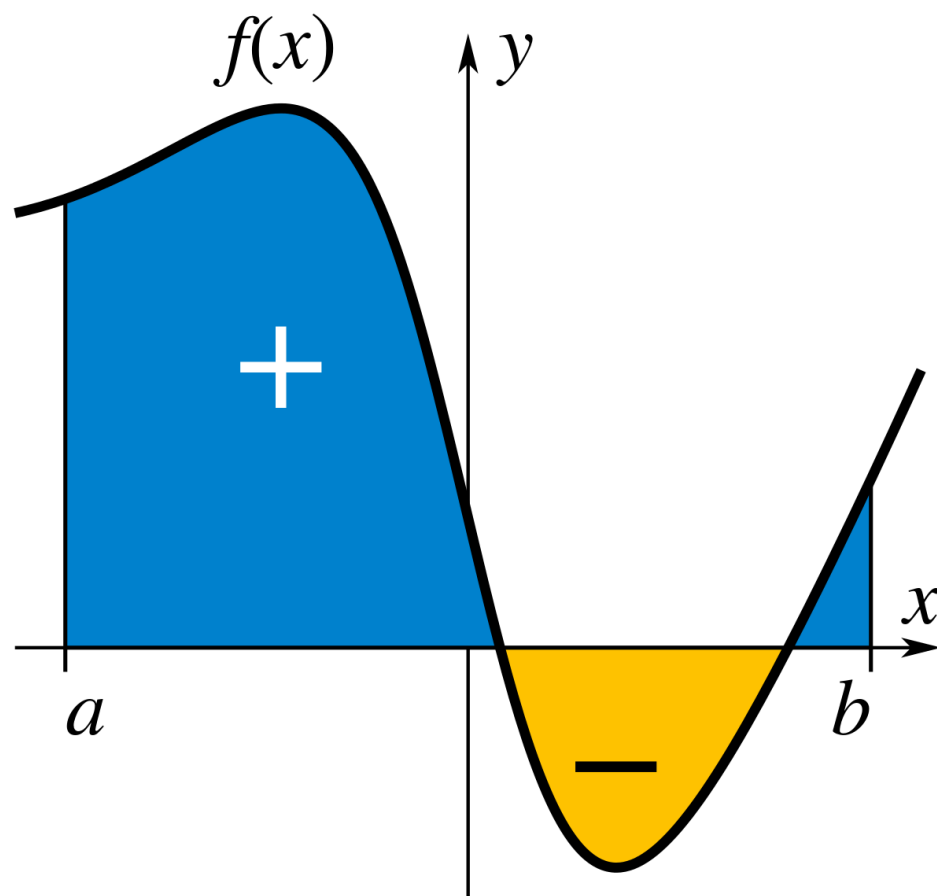
$$H(x, \lambda) = \lambda(x^2 - 1) + (1 - \lambda)(x + 2)$$

Homotopia - podsumowanie

- Metody homotopijne mogą być atrakcyjne pomimo swoich wad
- Przykładowo <http://www.hom4ps3.org> umożliwia rozwiązywanie równań wielomianowych

Całkowanie numeryczne

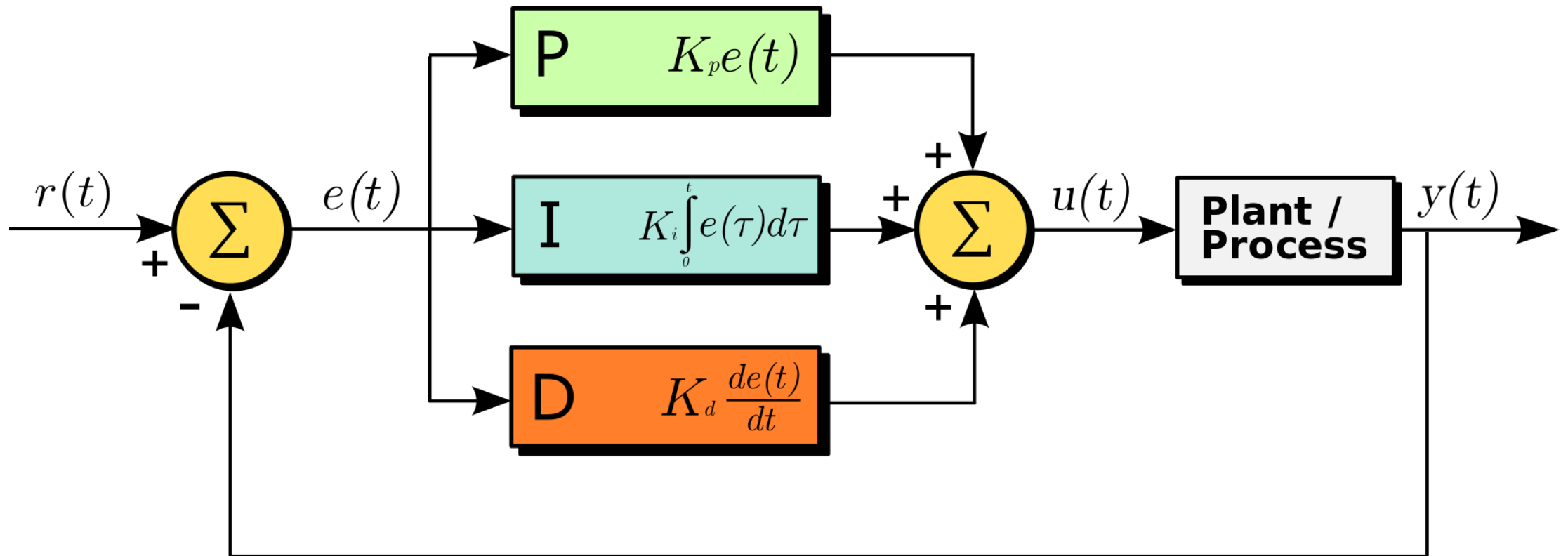
Praktyka to całki oznaczone



Trzy obszary całkowania numerycznego

- Całkowanie w czasie rzeczywistym
- Całkowanie funkcji jednej zmiennej
- Całkowanie funkcji wielu zmiennych

Całkowanie w czasie rzeczywistym

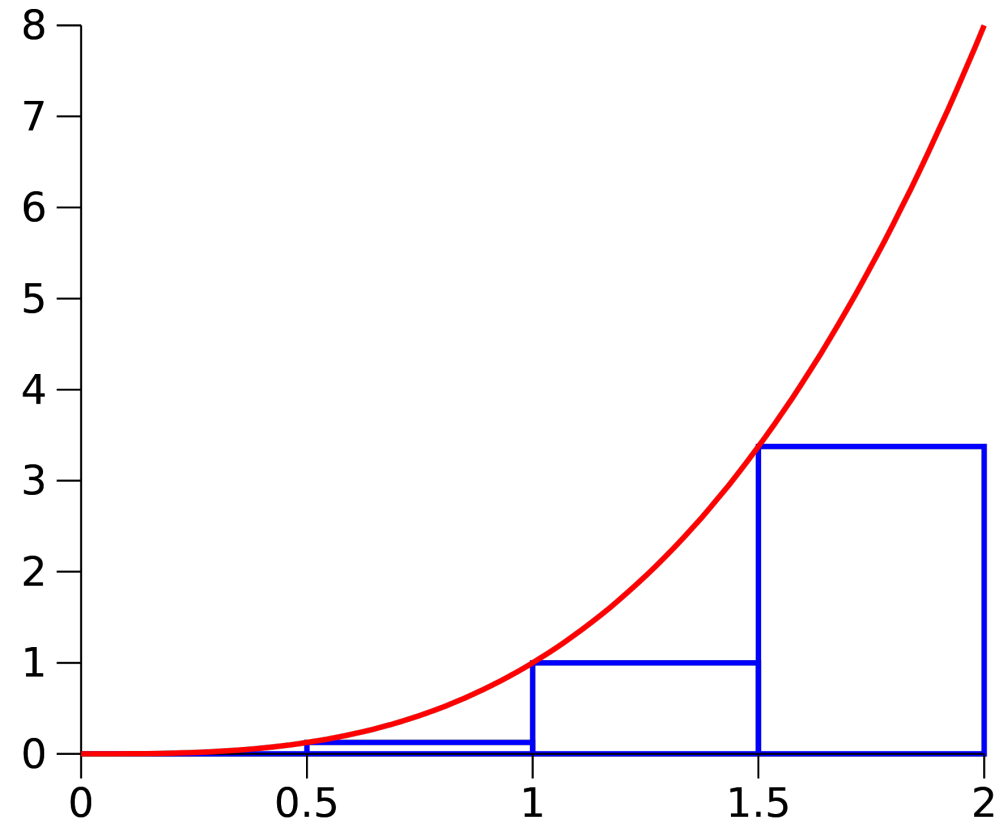


Praktyka całkowania on-line

- Otrzymujemy kolejne wartości funkcji w określonych odstępach czasu
- Nie możemy wybierać punktów
- Wartości funkcji są obarczone błędem pomiaru

Jak najprościej?

- Wzór prostokątów w przód, kwadratura Eulera, metoda prostokątów – wiele nazw na to samo
- Przybliżenie całki to poprzednia wartość pomnożona przez długość kroku między wartościami



Postać matematyczna

$$I(t_i) = I(t_{i-1}) + \int_{t_{i-1}}^{t_i} e(t)dt \approx \hat{I}(t_{i-1}) + e(t_{i-1}) \cdot \Delta t$$
$$t_i = t_{i-1} + \Delta t$$

Transmitancja

- Całkowanie w dziedzinie transformaty Laplace'a przedstawiamy jako

$$I(s) = \frac{1}{s} E(s)$$

- Do obliczeń dyskretnych w czasie używamy transformaty z , $z^{-1} = e^{-s\Delta t}$

$$\hat{I}(t_i) = \hat{I}(t_{i-1}) + e(t_{i-1}) \cdot \Delta t$$

$$\hat{I}(z) = z^{-1} \hat{I}(z) + \Delta t z^{-1} E(z)$$

$$\hat{I}(z) = \frac{\Delta t z^{-1}}{1 - z^{-1}} E(z) = \frac{\Delta t}{z - 1} E(z)$$

Aproksymacja Eulera w przód

- Aproksymacja integratora

$$\frac{1}{s} \approx \frac{\Delta t}{z - 1}$$

- Lub analogicznie różniczki

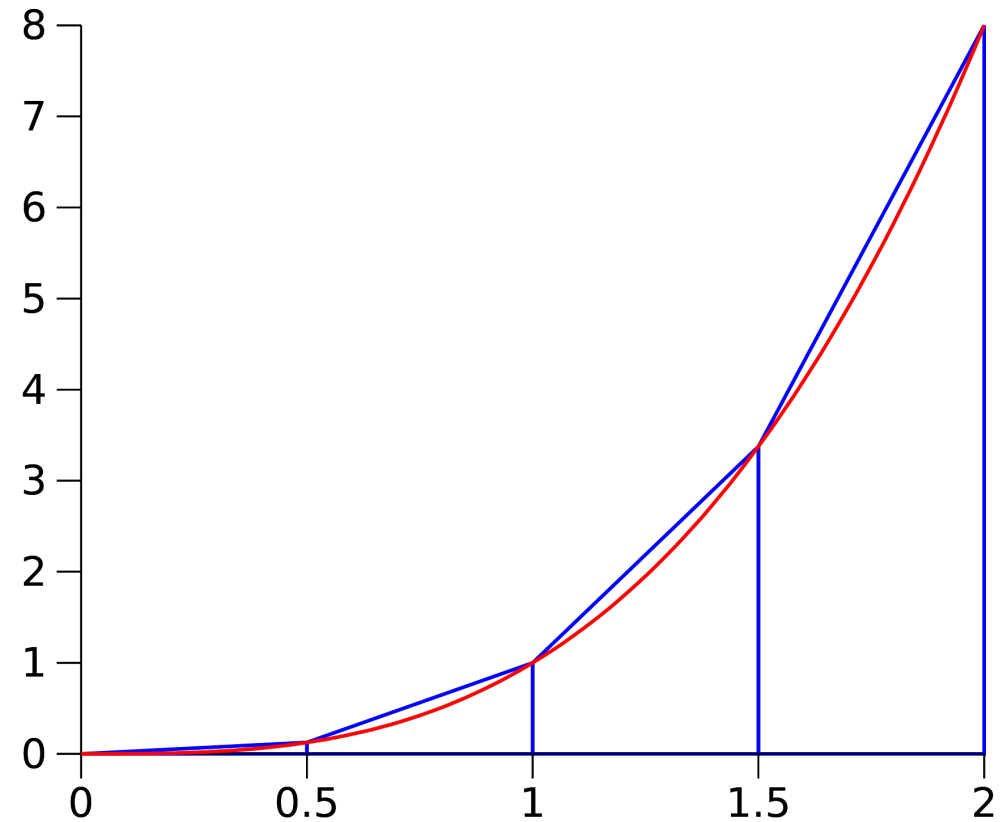
$$z \approx 1 + \Delta t \cdot s \iff s \approx \frac{1}{\Delta t} (1 - z)$$

Aproksymacja Eulera - zalety i wady

- Bardzo prosta do wyliczenia
- Da się przenieść na układy nieliniowe
- Nie gwarantuje zachowania stabilności
- Mało dokładna

Wzór trapezów

- Wykorzystujemy wartości zarówno poprzednią jak i obecną
- Podstawa bardzo popularnej aproksymacji Tustina



Postać matematyczna

$$I(t_i) = I(t_{i-1}) + \int_{t_{i-1}}^{t_i} e(t)dt \approx \hat{I}(t_{i-1}) + (e(t_i) + e(t_{i-1})) \cdot \frac{\Delta t}{2}$$
$$t_i = t_{i-1} + \Delta t$$

Transmitancja

- Zasada wyznaczania transmitancji jest analogiczna

$$\hat{I}(t_i) = \hat{I}(t_{i-1}) + (e(t_i) + e(t_{i-1})) \cdot \frac{\Delta t}{2}$$

$$\begin{aligned}\hat{I}(z) &= z^{-1}\hat{I}(z) + \frac{\Delta t}{2}(1 + z^{-1})E(z) \\ \hat{I}(z) &= \frac{\Delta t(1 + z^{-1})}{2(1 - z^{-1})}E(z) = \frac{\Delta t(z + 1)}{2(z - 1)}E(z)\end{aligned}$$

Aproksymacja Tustina

- Aproksymacja integratora

$$\frac{1}{s} \approx \frac{\Delta t(z + 1)}{2(z - 1)}$$

- Lub analogicznie różniczki

$$z \approx \frac{1 + s \Delta t/2}{1 - s \Delta t/2} \Leftrightarrow s \approx \frac{2}{\Delta t} \frac{z - 1}{z + 1}$$

Aproksymacja Tustina – zalety i wady

- Generalnie powszechnie stosowana
- Czasem może być trudna w obliczeniach analitycznych
- Nie przenosi się łatwo na obliczenia nieliniowe
- Gwarantuje stabilność
- Potencjalnie większe wymagania obliczeniowe

Kwadratury jednej zmiennej

- Numeryczne wyliczanie całki nazywamy kwadraturą
- Ogólnie chodzi o przedstawienie wartości całki na przedziale za pomocą ważonej sumy jej wartości w n punktach tj.

$$\int_{-1}^1 f(t) dt \approx \sum_{i=1}^n w_i f(t_i)$$

- O kwadraturze mówimy, że ma rząd wielomianowy m (zależny od n), jeżeli dla wielomianów stopnia m , tj. $f = p_m$ kwadratura jest równa dokładnie wartości całki.

Kwadratury interpolacyjne

- Postać wzoru na kwadraturę naturalnie kieruje nas w stronę interpolacji.
- Interpolujemy funkcję wielomianem i liczymy jego całkę
- Wyróżniamy kwadratury:
 - Newtona-Cotesa – interpolacja na węzłach równoodległych
 - Gaussa – wykorzystująca własności wielomianów ortogonalnych, w tym:
 - Gaussa-Legendre'a – kwadratura interpolacyjna na węzłach Legendre'a
 - Clenshawa-Curtisa – kwadratura interpolacyjna na węzłach Czebyszewa

Wielomiany Legendre'a

- Wielomiany ortogonalne dla iloczynu skalarnego

$$f \circ g = \int_{-1}^1 f \cdot g \, dt$$

- Wielomiany te spełniają relację

$$(k + 1)p_{k+1}(x) = (2k + 1)xp_k(x) - kp_{k-1}(x)$$

- Węzły Legendre'a to miejsca zerowe tych wielomianów

Podstawowe twierdzenie o kwadraturach interpolacyjnych

Dla każdego $n \geq 0$ kwadratura interpolacyjna na $n + 1$ węzłach ma rząd wielomianowy n . Kwadratura Gaussa-Legendre'a ma rząd wielomianowy $2n + 1$

Dowód:

Każdy wielomian stopnia $2n + 1$ możemy przedstawić jako $f(x) = P_{n+1}(x)q_n(x) + r_n(x)$

$$I = \int_{-1}^1 f(x) dx = \int_{-1}^1 P_{n+1}(x)q_n(x) dx + \int_{-1}^1 r_n(x) dx$$

Dowód cd.

- Z ortogonalności mamy

$$I = \int_{-1}^1 r_n(x) dx$$

- Ponieważ węzły Legendre'a to pierwiastki $P_{n+1}(x)$ to $f(x_k) = r_n(x_k)$
- Więc wartość kwadratury f będzie taka sama jak r_n , dla której jest dokładna z pierwszej części twierdzenia.

Podstawowe różnice

- Kwadratura Gaussa-Legendre'a jest najdokładniejsza dla wielomianów, ale jej węzły i wagi można wyliczyć jedynie numerycznie, co ma złożoność $O(n^3)$
- Kwadratura Clenshawa-Curtisa wykorzystuje aparat wielomianów Czebyszewa, więc złożoność jest dużo mniejsza $O(n \log(n))$
- Wszystkie kwadratury Gaussa mają w zasadzie niepowtarzalne węzły, więc podniesienie rzędu kwadratury wymaga wyliczenia wartości w nowych węzłach
- Kwadratury Newtona-Cotesa są używalne tylko dla niskich rzędów

Schemat wyliczania kwadratury CC

Kwadratura Clenshawa-Curtisa:

1. Wyliczamy wielomian interpolacyjny
2. Za pomocą FFT zamieniamy go na szereg Czebyszewa
3. Całki z poszczególnych wielomianów Czebyszewa znamy, więc obliczenie całki to suma iloczynów współczynników i odpowiadających im wartości całek.

Schemat wyliczania kwadratury GL

- Kwadraturę Gaussa-Legendre'a (a w zasadzie każdą Gaussa), czyli jej węzły i wagi, można wyliczyć rozwiązując pewien problem własny
- Klasyczna praca: G. H. Golub and J. H. Welsch, Calculation of Gauss quadrature rules, Math. Comp., 23 (1969), 221–230.

Kwadratury adaptacyjne

- Idea kwadratury adaptacyjnej polega na tym, aby liczyć kwadratury w podprzedziałach.
- Jednocześnie liczy się kwadratury dwóch różnych rzędów, które mają węzły na brzegach przedziału (i jak najwięcej wspólnych)
- Jak różnica między kwadraturami jest duża, to stosujemy każdą z kwadratur na podprzedziałach pomiędzy już wyliczonymi węzłami.