Różniczkowanie numeryczne
Maciej Dragun

1) Wstęp

Najprostszym podejściem do tego jest to co już znamy i lubimy a mianowicie różniczkowanie symboliczne. Nie będę wchodził w szczegóły biorąc pod uwagę, że ultymatywnie metoda ta nie jest w duchu tego przedmiotu. Chodzi oczwiście o zrobienie czegoś takiego tylko na poziomie kodu

$$f(x)=3x^2 \Rightarrow f'(x)=6x \tag{1.1}$$

W matlabie służyłaby do tego funkcja diff()

Wspominam o tym z dwóch powodów.

1. Wypada o tym przynajmniej wspomnieć ze względu na oczywistość tego podejścia
2. Ta metoda ma fundamentalną wadę z której wynika potrzeba opracowania numerycznych metod różniczkowania – nie działa w momencie gdy różniczkowana funkcja nie daje się wyrazić symbolicznie w szczególności gdy jest zadana przez pary punktów (co oczywiście jest częstym przypadkiem)

2) Ilorazy różnicowe pierwszego stopnia

Pierwszą opcją jest bezpośrednie spojrzenie na definicję pochodnej. Mając bowiem

$$f'(x)=\lim_{h\to 0}\frac{f(x+h)-f(x)}{h} \tag{2.1}$$

Można przyjąć za przybliżenie pochodnej następujące wyrażenie określane mianem przedniego ilorazu różnicowego

$$D_{f1}(x,h)=\frac{f(x+h)-f(x)}{h} \tag{2.2}$$

Gdzie h jest krokiem. Dla funkcji zadanej przez pary punktów f(x+h) może być po prostu kolejnym pomiarem

Przyjmując z kolei x = x – h (tj. poprzedni punkt pomiarowy) otrzymujemy wzór wstecznego ilorazu różnicowego

$$D_{b1}(x,h)=\frac{f(x)-f(x-h)}{h} \tag{2.3}$$

Istotna różnica między obydwoma podejściami polega na użyciu kolejnego lub poprzedniego punktu pomiarowego

Możemy oszacować błąd obcięcia dla powyższych metod stosując wzór Taylora. Dla przypomnienia

$$f(x)=f(a)+\frac{x-a}{1!}f^{(1)}(a)+\frac{(x-a)^2}{2!}f^{(2)}(a)+\frac{(x-a)^3}{3!}f^{(3)}(a)+\ldots+\frac{(x-a)^n}{n!}f^{(n)}(a)+R_n(x,a) \tag{2.4}$$

Gdzie a jest dowolnym punktem z dziedziny f. Zakładamy, że dana funkcja jest n-krotnie różniczkowalna

Podstawiając za x := x + h, a = x otrzymujemy

$$f(x+h)=f(x)+hf^{(1)}(x)+\frac{h^2}{2}f^{(2)}(x)+\frac{h^3}{6}f^{(3)}(x)+\ldots \quad (2.5)$$

Przenosząc f(x) na drugą stronę i dzieląc wyrażenie przez h != 0 otrzymujemy

$$D_{f1}(x,h)=\frac{f(x+h)-f(x)}{h}=f^{(1)}(x)+\frac{h}{2}f^{(2)}(x)+\frac{h^2}{6}f^{(3)}(x)+\ldots \quad (2.6)$$

$$D_{f1}(x,h)=\frac{f(x+h)-f(x)}{h}=f'(x)+O(h) \quad (2.7)$$

Analogicznie dla wstecznego ilorazu różnicowego

$$D_{b1}(x,h)=\frac{f(x)-f(x-h)}{h}=f^{(1)}(x)-\frac{h}{2}f^{(2)}(x)+\frac{h^2}{6}f^{(3)}(x)-\ldots \quad (2.8)$$

$$D_{b1}(x,h)=\frac{f(x)-f(x-h)}{h}=f'(x)+O(h) \quad (2.9)$$

Widzimy zatem, że błąd obcięcia obydwu metod jest proporcjonalny do kroku. Należałoby również przeprowadzić analizę błędów zaokrąglenia aczkolwiek zostanie to omówione dalej. Istotne są tu wzory (2.6) i (2.8). Dodać je do siebie stronami

$$\frac{f(x+h)-f(x)}{h}+\frac{f(x)-f(x-h)}{h}=2f^{(1)}(x)+\frac{2h^2}{6}f^{(3)}(x)+\frac{2h^4}{5!}f^{(5)}(x)+\ldots \quad (2.10)$$

$$D_{c2}(x,h)=\frac{f(x+h)-f(x-h)}{2h}=f'(x)+O(h^2) \quad (2.11)$$

Powyższy wzór określamy mianem centralnego ilorazu różnicowego. Istotne są tu dwie obserwacje

1. Powyższy wzór zależny jest od trzech punktów pomiarowych (f(x + h) i f(x – h))
2. Błąd obcięcia tej metody jest lepszy

3) Ilorazy różnicowe wyższych rzędów i ekstrapolacja Richardsona

Mając powyższe wzory możemy poczynić pewne sztuczki otrzymując wzory precyzyjniejsze pod kątem błędu obcięcia

$$D_{f1}(x,h)=\frac{f(x+h)-f(x)}{h}=f^{(1)}(x)+\frac{h}{2}f^{(2)}(x)+\frac{h^2}{6}f^{(3)}(x)+\ldots \quad (3.1)$$

$$D_{f1}(x,2h)=\frac{f(x+2h)-f(x)}{2h}=f^{(1)}(x)+\frac{2h}{2}f^{(2)}(x)+\frac{(2h)^2}{6}f^{(3)}(x)+\ldots \quad (3.2)$$

$$2D_{f1}(x,h)-D_{f1}(x,2h)=2\frac{f(x+h)-f(x)}{h}-\frac{f(x+2h)-f(x)}{2h}=f'(x)-\frac{2h^2}{3!}f^{(3)}(x)+... \quad (3.3)$$

$$D_{f2}(x,h)=2D_{f1}(x,h)-D_{f1}(x,2h)=\frac{-f(x+2h)+4f(x+h)-3f(x)}{2h}=f'(x)+O(h^2) \quad (3.4)$$

Widzimy więc, że poprzez sprytną manipulację wzorem przedniego ilorazu różnicowego otrzymujemy wyrażenie o lepszym błędzie obcięcia aczkolwiek wymagające większej liczby pomiarów. Wzór ten określamy przednim ilorazem różnicowym drugiego rzędu

Analogicznie uzyskać można

$$D_{b2}(x,h)=2D_{b1}(x,h)-D_{b1}(x,2h)=\frac{3f(x)-4f(x-h)+f(x-2h)}{2h}=f'(x)+O(h^2) \quad (3.5)$$

A także

$$D_{c4}(x,h)=\frac{4D_{c2}(x,h)-D_{c2}(x,2h)}{3}=\frac{8f(x+h)-8f(x-h)-f(x+2h)+f(x-2h)}{12h}=f'(x)+O(h^4)$$
(3.6)

Powyższy proces można kontynuować odnajdując coraz to lepsze przybliżenia pochodnych. Ta sztuczka nie bierze się z nikąd – wynika ona z Ekstrapolacji Richardsona.

Ekstrapolacja Richardsona jest ogólną metodą iteracyjnego tworzenia lepszych aproksymacji które zależą od pewnego dodatniego kroku oraz ich błąd daje się wyrażać poprzez

$$E(A)=a_0h^{k_0}+a_1h^{k_1}+... \quad (3.7)$$

A dane rozwiązanie A jest wyrażone jako funkcja od kroku (h)

Wykorzystuje się ją w wielu obszarach obliczeń numerycznych np. w obliczeniach przybliżenia liczby pi lub jak niżej przy różniczkowaniu numerycznym

Bez wchodzenia w szczegóły i powiązanie z Ekstrapolacją Richardsona uogólnienie omawianej procedury wygląda następująco

$$D_{f,n+1}(x,h)=\frac{2^nD_{f,n}(x,h)-D_{f,n}(x,2h)}{2^n-1} \quad (3.8)$$

$$D_{b,n+1}(x,h)=\frac{2^nD_{b,n}(x,h)-D_{b,n}(x,2h)}{2^n-1} \quad (3.9)$$

$$D_{c,2(n+1)}(x,h)=\frac{2^{2n}D_{c,2n}(x,h)-D_{c,2n}(x,2h)}{2^{2n}-1} \quad (3.10)$$

4) Analiza błędu ilorazów różnicowych

Oczywiście życie nie jest idealne i jak można się domyślić zarówno zmniejszania parametru h jak i ekstrapolacji nie powinno się kontynuować w nieskończoność. Owszem błędy obcięcia będą lepsze, ale co z błędami zaokrągleń? Rozważmy przedni iloraz różnicowy. Załóżmy, że pomiarów f(x) i f(x + h) dokonano z pewnymi błędami (odpowiednio e_0 i e_1)

$$D_{f1}(x,h)=\frac{f(x+h)+e_1-f(x)-e_0}{h}\overset{2.7}{=}f'(x)+\frac{e_1-e_0}{h}+O(h) \quad (4.1)$$

$$|D_{f1}(x,h)-f'(x)|=|\frac{e_1-e_0}{h}+O(h)|\leq|\frac{e_1-e_0}{h}|+|O(h)|\leq|2\frac{\epsilon}{h}|+|O(h)| \quad (4.2)$$

Tak więc górne ograniczenie całkowitego błędu składa się z dwóch składników

1. Błędu obcięcia O(h) wprost proporcjonalnego do kroku (to widać po przypomnieniu sobie faktu, że odcinamy wyrażenia szeregu Taylora zależne od parametru h)
2. Błędu zaokrągleń który jest odwrotnie proporcjonalny do kroku

Zauważyć więc można, że wraz z maleniem błędu obcięcia wzrasta błąd zaokrągleń i vice versa

Rozważmy teraz przedni iloraz różnicowy drugiego stopnia. Załóżmy, że pomiarów f(x), f(x + h) i f(x + 2h) dokonano z pewnymi błędami (odpowiednio e_0 e_1 i e_2)

$$D_{f2}(x,h)=\frac{-f(x+2h)-e_2+4f(x+h)+4e_1-3f(x)-3e_0}{2h}\overset{3.4}{=}f'(x)+\frac{-e_2+4e_1-3e_0}{2h}+O(h^2) \quad (4.3)$$

$$|D_{f2}(x,h)-f'(x)|=|\frac{-e_2+4e_1-3e_0}{2h}+O(h^2)|\leq|\frac{-e_2+4e_1-3e_0}{2h}|+|O(h^2)|\leq|4\frac{\epsilon}{h}|+|O(h^2)| \quad (4.4)$$

Obserwacje poczynione dla metody pierwszego rzędu pozostają prawdziwe. Dodatkowo warto zauważyć, że wraz z usprawnieniem błędu obcięcia pogorszył się błąd zaokrągleń (2 < 4)

Co ciekawe znając analityczną postać funkcji oraz jej pochodnych możliwe jest znalezienie optymalnego górnego ograniczenia h dla metody danego rzędu poprzez obliczenie minimum wyrażenia po prawej stronie nierówności. Warunki ku temu są oczywiście niepraktyczne więc sama czynność służyłaby dalszemu zgłębianiu teorii

Te rozważania prowadzą do następującego wniosku: dobór rzędu i kroku nie jest zadaniem trywialnym

5) Pochodne wyższego rzędu

Co w przypadku gdy chcemy wyciągnąć pochodne wyższego rzędu? Przyjrzyjmy się ponownie wzorom (3.1) i (3.2)

$$D_{f1}(x,h)=\frac{f(x+h)-f(x)}{h}=f^{(1)}(x)+\frac{h}{2}f^{(2)}(x)+\frac{h^2}{6}f^{(3)}(x)+\ldots \quad (3.1)$$

$$D_{f1}(x,2h)=\frac{f(x+2h)-f(x)}{2h}=f^{(1)}(x)+\frac{2h}{2}f^{(2)}(x)+\frac{(2h)^2}{6}f^{(3)}(x)+\dots \quad (3.2)$$

Odejmując od siebie te równania stronami celem eliminacji wyrazu $f^{(1)}(x)$ otrzymujemy

$$\frac{f(x+2h)-f(x)}{2h}-\frac{f(x+h)-f(x)}{h}=\frac{h}{2}f^{(2)}(x)+\frac{1}{2}h^2f^{(3)}(x)+\dots \quad (5.1)$$

$$D_{f1}^{(2)}=\frac{f(x+2h)-2f(x+h)+f(x)}{h^2}=f''(x)+O(h) \quad (5.2)$$

Analogicznie dla wstecznego ilorazu różnicowego

$$D_{b1}^{(2)}=\frac{f(x)-2f(x-h)+f(x-2h)}{h^2}=f''(x)+O(h) \quad (5.3)$$

Z kolei aby otrzymać centralny iloraz różnicowy

$$D_{f1}(x,h)=\frac{f(x+h)-f(x)}{h}=f^{(1)}(x)+\frac{h}{2}f^{(2)}(x)+\frac{h^2}{6}f^{(3)}(x)+\dots \quad (2.6)$$

$$D_{b1}(x,h)=\frac{f(x)-f(x-h)}{h}=f^{(1)}(x)-\frac{h}{2}f^{(2)}(x)+\frac{h^2}{6}f^{(3)}(x)-\dots \quad (2.8)$$

Odejmując od siebie te równania stronami celem eliminacji wyrazu $f^{(1)}(x)$ otrzymujemy

$$\frac{f(x+h)-f(x)}{h}-\frac{f(x)-f(x-h)}{h}=hf^{(2)}(x)+\frac{h^3}{12}f^{(4)}(x)+\dots \quad (5.4)$$

$$D_{c2}^{(2)}(x,h)=\frac{f(x+h)-2f(x)+f(x-h)}{h^2}=f^{(2)}(x)+O(h^2) \quad (5.5)$$

Ekstrapolacja Richardsona funkcjonuje analogicznie jak dla pochodnych pierwszego rzędu

6) Pochodne mieszane i alternatywne podejście do pochodnych wyższego rzędu

Poniżej znajduje się definicja pochdnej cząstkowej po zmiennej x_0

$$\frac{\partial f}{\partial x_0}=\lim_{h\to 0}\frac{f(x_0+h,x_1,\dots x_n)-f(x_0,x_1,\dots x_n)}{h} \quad (6.1)$$

Możemy więc przyjąć przedni iloraz różnicowy dla funkcji wielu zmiennych

$$D_{f1,x_0}(x_0,x_1,\dots x_n,h)=\frac{f(x_0+h,x_1,\dots x_n)-f(x_0,x_1,\dots x_n)}{h} \quad (6.2)$$

Łatwo zauważyć, że rozważania odnośnie pochodnych jednej zmiennej w funkcji jednej zmiennej utrzymują się dla pochodnej cząstkowej jednej zmiennej w funkcji wielu zmiennych

Co jednak jeżeli pożądamy pochodnej mieszanej? Zauważmy, że

$$\frac{\partial^2 f}{\partial x \partial y}=\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) \rightarrow D_{?,x}(D_{?,y}(x,y,h),h)=D_{?,x,y}(x,y,h) \qquad (6.3)$$

(zapis po prawej stronie nie jest idealny, ale pozostaje w konwencji)

Przykładowo dla przedniego ilorazu różnicowego

$$D_{f1,x}(D_{f1,y}(x,y,h))=\frac{D_{f1,y}(x+h,y,h)-D_{f1,y}(x,y,h)}{h}=\frac{\frac{f(x+h,y+h)-f(x+h,y)}{h}-\frac{f(x,y+h)-f(x,y)}{h}}{h}$$

(6.4)

Analogicznie wyprowadzać można inne pochodne mieszane

Zauważyć można, że to podejście jest zgodne z wcześniejszymi wyprowadzeniami pochodnych wyższego rzędu

$$D_{f1,x,x}(x,h)=D_{f1}^{(2)}=\frac{f(x+2h)-2f(x+h)+f(x)}{h^2} \qquad (6.5)$$

7) Różniczkowanie interpolacji/aproksymacji i ilorazy różnicowe dla niestałego kroku

Alternatywnym podejściem do stawianego problemu jest interpolacja lub aproksymacja na podstawie zadanych punktów i wyliczenie na podstawie otrzymanego wzoru pochodnej (dokładnej). Takie podejście nie tylko sprawdza się w boju ale i rozwiązuje problem braku równomiernie rozłożonych punktów, który jest częstym i poważnym w przypadku wszelkich danych pomiarowych.

Przykładowo rozpatrzmy interpolację wielomianem Lagrange'a dla nierównomiernie rozłożonych punktów (x_0, y_0), (x_1, y_1), (x_2, y_2) gdzie $x_0 < x_1 < x_2$, $x_1 = x_0 + h_1$, $x_2 = x_0 + h_2$

Dla przypomnienia punkty można interpolować wielomianem Lagrange'a w następujący sposób

$$P(x)=\sum_{j=1}^{n}\prod_{k=1,k\neq j}^{n} y_j \frac{x-x_k}{x_j-x_k}$$ gdzie n to liczba punktów (7.1)

Także dla rozpatrywanych danych mamy

$$P(x)=\frac{(x-x_1)(x-x_2)}{(x_0-x_1)(x_0-x_2)}y_0+\frac{(x-x_0)(x-x_2)}{(x_1-x_0)(x_1-x_2)}y_1+\frac{(x-x_0)(x-x_1)}{(x_2-x_0)(x_2-x_1)}y_2 \qquad (7.2)$$

$$P(x)=\frac{x^2-xx_2-xx_1+x_1x_2}{(x_0-x_1)(x_0-x_2)}y_0+\frac{x^2-xx_2-xx_0+x_0x_2}{(x_1-x_0)(x_1-x_2)}y_1+\frac{x^2-xx_0-xx_1+x_1x_0}{(x_2-x_0)(x_2-x_1)}y_2 \qquad (7.3)$$

$$P'(x)=\frac{2x-x_2-x_1}{(x_0-x_1)(x_0-x_2)}y_0+\frac{2x-x_2-x_0}{(x_1-x_0)(x_1-x_2)}y_1+\frac{2x-x_0-x_1}{(x_2-x_0)(x_2-x_1)}y_2 \qquad (7.4)$$

Podstawiając $x_1 = x_0 + h_1$, $x_2 = x_0 + h_2$ oraz dla $x = x_0$ otrzymujemy

$$P'(x_0)=\frac{-h_1-h_2}{h_1h_2}y_0+\frac{-h_2}{(h_1)(h_1-h_2)}y_1+\frac{-h_1}{(h_2)(h_2-h_1)}y_2 \quad (7.5)$$

Co ciekawe przyjmując $h_1 = h$ i $h_2 = 2h$ otrzymujemy przedni iloraz różnicowy drugiego stopnia

$$P'(x_0)=\frac{-3}{2h}y_0+\frac{2}{h}y_1+\frac{-1}{2h}y_2=\frac{-y_2+4y_1-3y_0}{2h} \quad (7.6)$$

Innymi słowy wzór (6.5) można określić mianem ilorazu różnicowego dla nierównomiernie rozłożonych punktów

Analogicznie możemy

- Wyprowadzić wsteczny iloraz różnicowy dla nierównomiernie rozłożonych punktów przeprowadzając interpolację wielomianem Lagrange'a dla punktów (x_{-2}, y_{-2}), (x_{-1}, y_{-1}), (x_0, y_0) gdzie $x_{-2} < x_{-1} < x_0$, $x_{-1} = x_0 - h_{-1}$, $x_{-2} = x_0 - h_{-2}$
- Wyprowadzić centralny iloraz różnicowy dla nierównomiernie rozłożonych punktów przeprowadzając interpolację wielomianem Lagrange'a dla punktów (x_{-1}, y_{-1}), (x_0, y_0), (x_1, y_1) gdzie $x_{-1} < x_0 < x_1$, $x_{-1} = x_0 - h_{-1}$, $x_1 = x_0 + h_1$

Drobne rozważania: być może możliwym byłoby opracowanie wzorów ekstrapolacji Richardsona dla nierównomiernie rozłożonych punktów. Przemawiają ku za tym dwie obserwacje:

1. Ograniczenie górne błędu interpolacji wielomianem Lagrange'a jest wyrażeniem spełniającym warunki dla ekstrapolacji Richardsona
2. Nic nie stoi na przeszkodzie aby wcześniej stosowane „sztuczki" zastosować przy nierównomiernym kroku. To podejście okazałoby się problematyczne w analizie błędu ze względu na jego wyrażenie jako $O(h_1, h_2, \ldots)$
   $D_{c2}(x,h_1,h_2)=D_{f1}(x,h_1)+D_{b1}(x,h_2)$
   $D_{f2}(x,h_1,h_2)=2D_{f1}(x,h_1)-D_{f1}(x,h_2)$

Ultymatywnie ilorazy różnicowe wyższych rzędów można otrzymać przez interpolację na większej liczbie punktów i odpowiednie ich uporządkowanie / obranie punktu x_0

Oczywiście interpolacja Lagrange'a nie jest jedyną opcją w kwestii interpolacji/aproksymacji