

Aproksymacja średniokwadratowa funkcji ciągłej za pomocą wielomianów.

Często w praktyce programistycznej zachodzi potrzeba przybliżenia jakiejś ciągłej funkcji w zadanym przedziale inną funkcją, łatwiejszą do obliczenia przez komputer. Jedną z najłatwiejszych dla komputera funkcji jest wielomian, dlatego, że wymaga wykonywania jedynie operacji dodawania i mnożenia. Wielomian możemy sobie zdefiniować w sposób następujący:

$$W_N(x) = \sum_{i=0}^N a_i x^i \quad (1)$$

gdzie N to stopień wielomianu. Najbardziej znanymi wielomianami są dwumian (którego wykresem jest linia prosta) w postaci $y=a_1x+a_0$, oraz trójmian kwadratowy (którego wykresem jest parabola) w postaci $y=a_2x^2+a_1x+a_0$. W praktyce obliczeniowej przydają się także wielomiany wyższych rzędów.

Zagadnieniem najważniejszym przy zastępowaniu jakiejś funkcji przybliżeniem wielomianowym jest określenie współczynników tego wielomianu a_i w taki sposób, aby popełniany błąd był jak najmniejszy w zadanym przedziale. O tym jest właśnie ten tekst.

Zacznijmy od tego w jaki sposób będziemy oceniać to, jak dobrze jakiś wielomian (czy ogólnie jakaś funkcja przybliżająca) przybliży naszą oryginalną funkcję $f(x)$. Używanych jest kilka kryteriów, z których najczęstsze są trzy:

1. Maksymalna odchyłka powinna być jak najmniejsza. Z pozoru niezłe kryterium, problem tylko w tym, że za cenę zmniejszenia odchyłki maksymalnej możemy dostać funkcję, która np. w żadnym miejscu nie pokrywa się z oryginałem.
2. Suma (a ściślej całka) wartości bezwzględnej odchyłek w zadanym przedziale powinna być jak najmniejsza. To już lepiej, ale to kryterium (aproksymacja z jego użyciem jest nazywana **aproksymacją jednostajną**) przykładła zbyt dużą wagę do dużych odchyłek, nie zapewniając możliwie najlepszej jakości przybliżenia tam, gdzie odchyłka jest mniejsza. Dodatkowo funkcja wartości bezwzględnej ma nieciągłą pochodną co stwarza problem przy obliczeniach.
3. Suma (całka) kwadratów odchyłek w zadanym przedziale powinna być jak najmniejsza. To jest tak zwana **aproksymacja średniokwadratowa**. Ta daje dobre jakościowo przybliżenia no i nie ma problemów z nieciągłością pochodnej funkcji kwadratowej. Tą właśnie aproksymację zastosujemy.

Zanim się zabierzemy do pracy, należałoby formalnie zapisać nasze średniokwadratowe kryterium. Oznaczmy funkcję oryginalną jako $f(x)$, przybliżający wielomian będzie oznaczony jako $W(x)$. Przedział w którym przybliżamy będzie się rozciągał od u do v . W tym momencie możemy napisać funkcję błędu E :

$$E = \int_u^v (f(x) - W(x))^2 dx \quad (2)$$

Od czego zależy (albo ściślej funkcją jakich zmiennych jest) E ? Wbrew pozorom E nie jest bynajmniej funkcją x , zmienna ta po prostu pod całką przebiega od u do v . Błąd średniokwadratowy E zależy oczywiście od postaci funkcji W , a ponieważ W jest wielomianem, to błąd zależy od współczynników wielomianu. Wstawiając (1) do (2) i zapisując jawnie zmienne funkcji E otrzymamy:

$$E(a_0, a_1, \dots, a_N) = \int_u^v (f(x) - \sum_{i=0}^N a_i x^i)^2 dx \quad (3)$$

Teraz czeka nas znalezienie minimum tej funkcji. Minimum błędu to najdokładniejsza aproksymacja. Minimum funkcji od razu kojarzy się nam z zerowaniem się w tym miejscu jej pochodnej. A co jeżeli znajdziemy maksimum? Ano nie znajdziemy, matematycznie ściśle wyjdzie to nam trochę później, na razie odrobinka geometrycznej intuicji. Nasza oryginalna funkcja w przedziale (u, v) to jakiś kawałek krzywej, wielomian W w tym samym przedziale to też kawałek krzywej. Jeżeli teraz byśmy chcieli zrobić anty-aproksymację (uzyskać maksymalny możliwy błąd) to nic nas nie powstrzyma przed zwiększaniem tego błędu w nieskończoność. Na „chłopski rozum” więc, maksimum nie znajdziemy. A czy musimy znaleźć w ogóle jakieś ekstremum (czyli minimum, bo maksimum właśnie wyeliminowaliśmy)? Trzymając się naszej geometrycznej intuicji – jeżeli będziemy wyginać nasz wielomian, zmieniając mu współczynniki to **jakaś** ich kombinacja musi dać najmniejszy możliwy błąd. I to będzie właśnie nasze poszukiwane minimum. Jeżeli ktoś nie wierzy w geometryczne intuicje to śpieszę go zawiadomić, że panowie **Weierstrass** i **Stone** udowodnili, że każdą funkcję ciągłą w pewnym przedziale da się przybliżyć w tym przedziale wielomianem (odpowiedniego stopnia rzecz jasna) z dowolną dokładnością. Jest to twierdzenie nazwane (co za niespodzianka...) **twierdzeniem Weierstrassa-Stone'a**.

Jak szukamy minimum? Patrzymy gdzie się zeruje pochodna funkcji. W przypadku funkcji wielu zmiennych (a taką jest E) poszukiwane minimum będzie się znajdowało w miejscu gdzie wyzerują się wszystkie pochodne cząstkowe. Musimy więc znaleźć wszystkie pochodne cząstkowe funkcji E . Wygląda groźnie? Nic podobnego. Najpierw sobie podniesiemy do kwadratu to nasze wyrażenie podcałkowe. Żeby było lepiej widać pewne rzeczy, rozpiszę to mnożenie w postaci tabelki.

	$F(x)$	$-a_0$	$-a_1 x$	$-a_2 x^2$	$-a_3 x^3 \dots$
$F(x)$	$F^2(x)$	$-a_0 F(x)$	$-a_1 x F(x)$	$-a_2 x^2 F(x)$	$-a_3 x^3 F(x)$
$-a_0$	$-a_0 F(x)$	a_0^2	$a_0 a_1 x$	$a_0 a_2 x^2$	$a_0 a_3 x^3$
$-a_1 x$	$-a_1 x F(x)$	$a_0 a_1 x$	$a_1^2 x^2$	$a_1 a_2 x^3$	$a_1 a_3 x^4$
$-a_2 x^2$	$-a_2 x^2 F(x)$	$a_0 a_2 x^2$	$a_1 a_2 x^3$	$a_2^2 x^4$	$a_2 a_3 x^5$
$-a_3 x^3 \dots$	$-a_3 x^3 F(x)$	$a_0 a_3 x^3$	$a_1 a_3 x^4$	$a_2 a_3 x^5$	$a_3^2 x^6$

Dostajemy całą masę wyrazów, oczywiście wszystko to pod całką. Jak wiadomo z teorii całkowania całkę sumy jakichś wyrażeń możemy zapisać jako sumę całek. I to wszystko potem całkować i liczyć N pochodnych cząstkowych? Na szczęście nie wszystko. W momencie, kiedy liczymy pochodną $\delta E / \delta a_i$ wszystkie wyrazy nie zawierające a_i nam się wyzerują, więc możemy je pominąć. Przykładowo w tabelce zaznaczyłem wszystkie wyrazy potrzebne do policzenia pochodnej cząstkowej $\delta E / \delta a_2$. Spróbujmy je zapisać w jakimś zwartym wyrażeniu, zwracając uwagę na to, że tabela jest symetryczna względem głównej przekątnej.

$$\frac{\delta E}{\delta a_2} = \int_u^v -2 a_2 x^2 F(x) dx + \int_u^v 2 a_0 a_2 x^2 dx + \int_u^v 2 a_1 a_2 x^3 dx + \int_u^v a_2^2 x^4 dx + \int_u^v 2 a_2 a_3 x^5 dx + \dots \quad (4)$$

Wyrzucimy teraz co się da przed całki, a wyrazy od drugiego do końca zapiszmy korzystając ze znaku sumy:

$$\frac{\delta E}{\delta a_2} = \frac{\delta}{\delta a_2} \left(-2 a_2 \int_u^v x^2 F(x) dx + 2 a_0 a_2 \int_u^v x^2 dx + 2 a_1 a_2 \int_u^v x^3 dx + a_2^2 \int_u^v x^4 dx + 2 a_2 a_3 \int_u^v x^5 dx + \dots \right) \quad (5)$$

$$\frac{\delta E}{\delta a_2} = \frac{\delta}{\delta a_2} \left(-2 a_2 \int_u^v x^2 F(x) dx + 2 a_2 \sum_{j=0}^1 a_j \int_u^v x^{j+2} dx + a_2^2 \int_u^v x^4 dx + 2 a_2 \sum_{j=3}^N a_j \int_u^v x^{j+2} dx \right) \quad (6)$$

Tutaj proszę zwrócić uwagę na specjalne traktowanie elementu tabeli, dla którego $j = 2$. Wszystkie inne elementy występują w tablicy dwa razy, ale ten tylko raz, dlatego nie mógł zostać wciągnięty pod znak sumy, a i pochodna z niego będzie liczona inaczej (bo a_2 występuje tu w kwadracie, a nie tak jak wszędzie w pierwszej potędze). Analogicznie moglibyśmy dla każdej pochodnej cząstkowej po a_i napisać:

$$\frac{\delta E}{\delta a_i} = \frac{\delta}{\delta a_i} \left(-2 a_i \int_u^v x^i F(x) dx + 2 a_i \sum_{j=0}^{i-1} a_j \int_u^v x^{i+j} dx + a_i^2 \int_u^v x^{2i} dx + 2 a_i \sum_{j=i+1}^N a_j \int_u^v x^{i+j} dx \right) \quad (7)$$

Zwróćmy teraz uwagę na fakt, że do policzenia pochodnej nie musimy wcale liczyć tych całek (a przynajmniej nie od razu), z punktu widzenia zmiennej a_i są one bowiem stałymi. No to bierzemy się za różniczkowanie:

$$\frac{\delta E}{\delta a_i} = -2 \int_u^v x^i F(x) dx + 2 \sum_{j=0}^{i-1} a_j \int_u^v x^{i+j} dx + 2 a_i \int_u^v x^{2i} dx + 2 \sum_{j=i+1}^N a_j \int_u^v x^{i+j} dx \quad (8)$$

Teraz okazuje się, że można już zakończyć odrębne traktowanie przypadku $i = j$. Brak dwójki, wynikający z tego, że ten wyraz występuje tylko raz, został nadrobiony przy liczeniu pochodnej (a_i występowała w kwadracie). Możemy więc teraz trzy ostatnie wyrazy wciągnąć pod jedną sumę:

$$\frac{\delta E}{\delta a_i} = -2 \int_u^v x^i F(x) dx + 2 \sum_{j=0}^N a_j \int_u^v x^{i+j} dx \quad (9)$$

Jeżeli teraz przyrównamy tę pochodną cząstkową do zera, to dostaniemy zwykłe równanie liniowe o N niewiadomych $a_0 \dots a_N$. Wbrew pozorom całki będą to zwykłe stałe, czyli współczynniki równania. Tak czy inaczej trzeba je jednak policzyć. Całki pod sumą są trywialne do policzenia. Najpierw korzystamy z:

$$\int x^n dx = \frac{x^{n+1}}{n+1} \quad (10)$$

Następnie z właściwości całki oznaczonej:

$$\text{jeżeli } \int f(x) dx = g(x) \text{ to } \int_a^b f(x) dx = g(b) - g(a) \quad (11)$$

Łatwość policzenia pierwszej całki natomiast zależy od postaci $F(x)$. Jeżeli jest łatwo to można policzyć analitycznie całkę nieoznaczoną i tak, jak poprzednio, skorzystać z (11). Jeżeli natomiast całka jest „wredna” to można spokojnie policzyć jej wartość jedną z metod numerycznych, oczywiście pamiętając o wymaganej precyzji. Po zaaplikowaniu (10) i (11) do (9) otrzymamy (pierwszą całkę zostawiłem nie ruszoną):

$$\frac{\delta E}{\delta a_i} = -2 \int_u^v x^i F(x) dx + 2 \sum_{j=0}^N a_j \frac{v^{i+j+1} - u^{i+j+1}}{i+j+1} = 0$$

Po przyrównaniu do zera otrzymujemy równanie liniowe z N niewiadomymi, pierwsza całka (ta z $F(x)$) jest wyrazem wolnym. Możemy napisać N takich równań (po jednym dla każdej pochodnej cząstkowej), w rezultacie dostaniemy liniowy układ N równań z N niewiadomymi. Rozwiązanie takiego układu dowolną metodą to już banał. Macierzowo taki układ możemy zapisać następująco:

$$M \cdot A = F$$

gdzie:

- A jest wektorem niewiadomych, $A = [a_0 \dots a_N]$.
- M jest macierzą współczynników o rozmiarze $N \times N$. Elementy tej macierzy dane są wzorem:

$$M_{ij} = \frac{v^{i+j+1} - u^{i+j+1}}{i+j+1}$$

- F jest wektorem wyrazów wolnych. Elementy tego wektora są dane wzorem:

$$F_i = \int_u^v x^i F(x) dx$$

Otrzymany układ równań rozwiązujemy dowolną metodą w wyniku otrzymując współczynniki wielomianu.

Przykład

Obrazy w formatach PGM i PPM (*Portable GrayMap* i *Portable PixMap*) używają funkcji korekcji gamma przy zapisie wartości pikseli w celu osiągnięcia większego zakresu dynamicznego kolorów. Przy dekodowaniu obrazów w tych formatach należy każdy piksel (a w przypadku PPM każdą z trzech składowych koloru piksela) poddać korekcji odwrotnej. Funkcja korekcji odwrotnej zdefiniowana jest następująco:

$$F(x) = \begin{cases} 0.230809661858x, & x \in \langle 0, 0.0779863366857 \rangle \\ \left(\frac{x+0.099}{1.099} \right)^{2.2}, & x \in \langle 0.0779863366857, 1 \rangle \end{cases}$$

Naszym zadaniem jest aproksymacja drugiego, potęgowego odcinka tej funkcji z dokładnością adekwatną dla obrazków 8-bitowych. Dopuszczalny błąd to pół przedziału kwantowania koloru. Dla zakresu $0 - 255$ będzie to $1/510 = 1.96078431373 \cdot 10^{-3}$. Wstępnie zaczniemy od trójmianu kwadratowego. Macierz M będzie wyglądać następująco:

$$m = \begin{bmatrix} v-u & \frac{v^2-u^2}{2} & \frac{v^3-u^3}{3} \\ \frac{v^2-u^2}{2} & \frac{v^3-u^3}{3} & \frac{v^4-u^4}{4} \\ \frac{v^3-u^3}{3} & \frac{v^4-u^4}{4} & \frac{v^5-u^5}{5} \end{bmatrix} = \begin{bmatrix} 0.922013663314 & 0.496959065645 & 0.333175232446 \\ 0.496959065645 & 0.333175232446 & 0.249990752718 \\ 0.333175232446 & 0.249990752718 & 0.199999423071 \end{bmatrix}$$

Wektor wyrazów wolnych natomiast ma poniższą postać:

$$F = \begin{bmatrix} \int_u^v F(x) dx \\ \int_u^v xF(x) dx \\ \int_u^v x^2 F(x) dx \end{bmatrix} = \begin{bmatrix} 0.34244195 \\ 0.25353567 \\ 0.20168857 \end{bmatrix}$$

Rozwiązując ten układ równań otrzymamy:

$$\begin{cases} a_0 = 0.01628181 \\ a_1 = 0.00593321 \\ a_2 = 0.97390593 \end{cases}$$

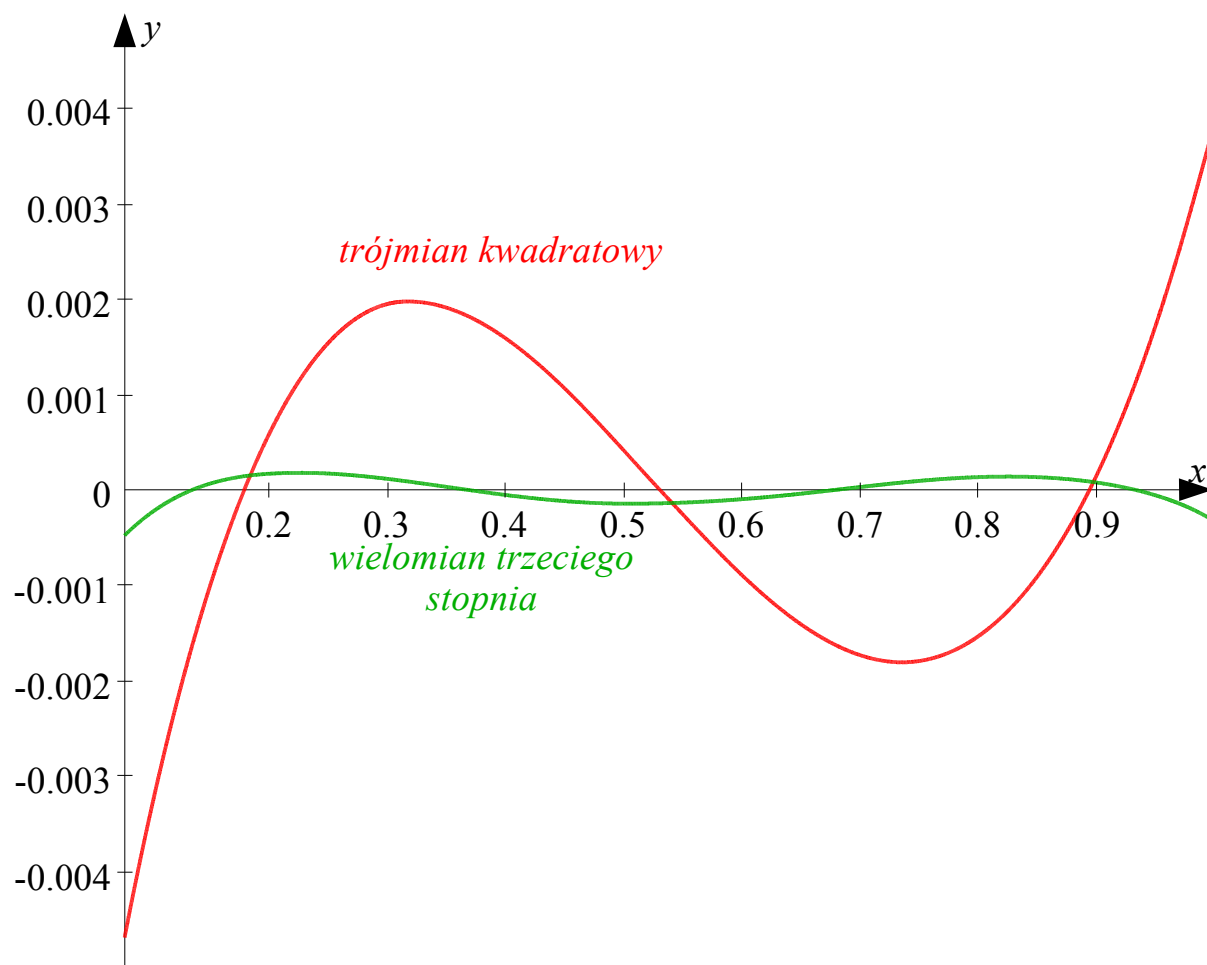
Tak więc ostatecznie aproksymujący trójmian ma postać:

$$y = 0,97390953x^2 + 0,00593321x + 0,01628181$$

Niestety okazuje się, że największa odchyłka jaką daje aproksymacja trójmianem, jest większa od dopuszczalnej. Wykres funkcji błędnej będącej różnicą funkcji oryginału $f(x)$ i aproksymującego wielomianu $W(x)$ znajduje się na następnej stronie. Największa odchyłka wynosi $4,7 \cdot 10^{-3}$ (na początku przedziału) i jest to zdecydowanie za dużo wobec dopuszczalnej $1,96 \cdot 10^{-3}$. Należy więc wziąć wielomian wyższego rzędu, a więc trzeciego. W opisany już sposób konstruujemy liniowy układ czterech równań z czterema niewiadomymi, którego postać tutaj pominę. Rozwiązując układ otrzymamy wielomian trzeciego rzędu:

$$y = 0,10738344 x^3 + 0,80027025 x^2 + 0,08582785 x + 0,00684798$$

Jego funkcję błędnej również znajdziemy na rysunku na następnej stronie, dla tego wielomianu maksymalna odchyłka wynosi $4,7 \cdot 10^{-4}$, a więc spełnia założenia zadania. Warto zauważyć, że aproksymacja wielomianem N -tego stopnia ma dokładnie N miejsc, w których funkcja aproksymująca pokrywa się z oryginałem, miejsca te są rozmieszczone symetrycznie względem środka przybliżanego przedziału.



mgr inż. Grzegorz Kraszewski
<krashan@teleinfo.pb.edu.pl>