Aproksymacja funkcji

Plan wykładu

1. Problem aproksymacji, normy, rodzaje aproksymacji

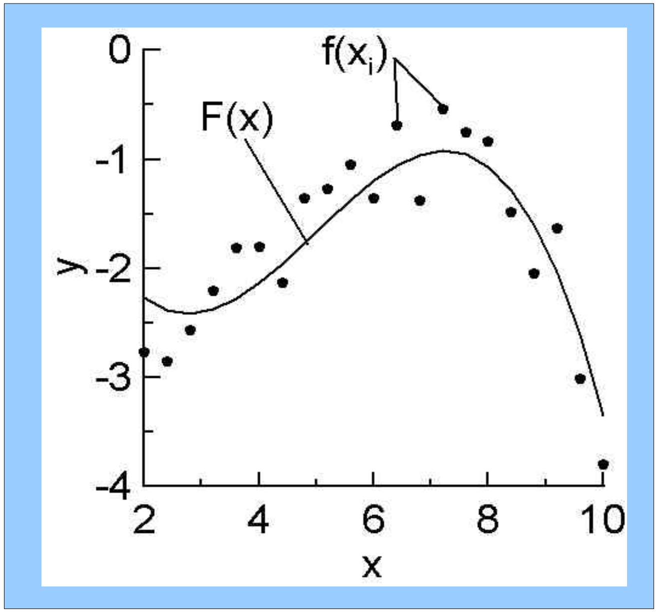
2. Aproksymacja średniokwadratowa

    a) w bazie jednomianów
    b) w bazie wielomianów ortogonalnych
    c) w bazie funkcji trygonometrycznych
    d) w bazie funkcji sklejanych

3. Aproksymacja Padego

Aproksymacja liniowa funkcji f(x) polega na wyznaczeniu współczynników $a_0, a_1, a_2, \ldots, a_m$ funkcji aproksymującej:

$$F(x) = a_0\varphi_0(x) + a_1\varphi_1(x) + \ldots + a_m\varphi_m(x)$$



gdzie: $\varphi_i(x)$ - są funkcjami bazowymi (m+1) wymiarowej podprzestrzeni liniowej X_{m+1} ($X_{m+1} \in X$)

Żądamy aby funkcja F(x) spełniała warunek

$$\|f(x) - F(x)\| = minimum$$

Wybór podprzestrzeni i bazy zależy od rodzaju problemu:

- podprzestrzeń funkcji trygonometrycznych z bazą:

$$1,\ sin(x),\ cos(x),\ sin(2x),\ cos(2x), \ldots,\ sin(kx),\ cos(kx)$$

- podprzestrzeń wielomianów stopnia m z bazą:

$$1,\ x,\ x^2,\ \ldots, x^m$$

- podprzestrzeń funkcji, których o własnościach ściśle związanych z własnościami rozważanego problemu, np.:

$$exp\left(a_0 + a_1x + a_2x^2\right)$$

Przykłady norm stosowanych w aproksymacji

- norma Czebyszewa

$$\|f(x) - F(x)\| = \sup_{[a,b]} |f(x) - F(x)|$$

- norma L_2

$$\|f(x) - F(x)\| = \left(\int_a^b |f(x) - F(x)|^2 dx \right)^{\frac{1}{2}}$$

- norma L_2 z wagą

$$\|f(x) - F(x)\| = \left(\int_a^b w(x) |f(x) - F(x)|^2 dx \right)^{\frac{1}{2}}$$

gdzie: w(x) jest nieujemną ciągłą funkcją wagową

- jeśli funkcja f(x) jest określona na dyskretnym zbiorze punktów wówczas norma L_2 z wagą przyjmuje postać:

$$\|f(x) - F(x)\| = \left(\sum_{i=0}^{n} w(x_i) \left[f(x_i) - F(x_i) \right]^2 \right)^{\frac{1}{2}}$$

Aproksymacja średniokwadratowa

Dla funkcji ciągłej f(x) określonej w przedziale [a,b] poszukujemy minimum wartości całki:

$$\|F(x) - f(x)\| = \int_a^b w(x)[F(x) - f(x)]^2 dx$$

lub sumy gdy funkcja jest określona na dyskretnym zbiorze n+1 punktów (**metoda najmniejszych kwdratów**):

$$\|F(x) - f(x)\| = \sum_{i=0}^{n} w(x_i)[F(x_i) - f(x_i)]^2$$

$$w(x_i) \geq 0 \qquad i = 0, 1, 2, \ldots, n$$

Aproksymacja jednostajna

Dla funkcji f(x)określonej w przedziale [a,b] poszukujemy F(x) dającej najmniejsze maksimum różnicy między nimi w całym przedziale:

$$\|F(x) - f(x)\| = \sup_{x \in [a,b]} |F(x) - f(x)|$$

Tw. 1 (Weierstrassa)

Jeżeli funkcja f(x) jest ciągła na skończonym przedziale [a,b], to dla każdego ε dodatniego można dobrać takie n, że jest możliwe utworzenie wielomianu $P_n(x)$ stopnia n (n=n(ε)), który spełnia nierówność:

$$\|f(x) - P_n(x)\| \leq \varepsilon$$

Z twierdzenia powyższego wynika, że **zawsze** można znaleźć wielomian o dowolnie małym odchyleniu od funkcji f(x).

Tw. 2 (Weierstrassa)

Jeżeli funkcja f(x) jest funkcją ciągłą na R i okresową o okresie 2π to dla każdego *ε* dodatniego istnieje wielomian trygonometryczny

$$\begin{aligned} S_n(x) &= a_0 + \sum_{k=1}^{n} (a_k cos(kx) + b_k sin(kx)) \\ n &= n(\varepsilon) \end{aligned}$$

spełniający dla wszystkich x nierówność

$$|f(x) - S_n(x)| < \varepsilon$$

Metoda aproksymacji średniokwadratowej

Dysponując układem funkcji bazowych w podprzestrzeni X_n:

$$\varphi_i(x), \quad i = 0, 1, \ldots, m$$

szukamy wielomianu F(x) będącego najlepszym przybliżeniem średniokwadratowym funkcji f(x) na zbiorze $X=(x_j)$:

$$F(x) = \sum_{i=0}^{m} a_i \varphi_i(x)$$

Dla F(x) liczymy normę L_2

$$\begin{aligned} & H(a_0, a_1, \ldots, a_m) = \\ & = \sum_{j=0}^{n} w(x_j) \left[f(x_j) - \sum_{i=0}^{m} a_i \varphi_i(x_j) \right]^2 \\ & = \sum_{j=0}^{n} w(x_j) R_j^2 \end{aligned}$$

gdzie: R_j jest odchyleniem w punkcie x_j

Szukamy minimum funkcji H (wielu zmiennych) ze względu na współczynniki $a_0, a_1, \ldots$

$$\frac{\partial H}{\partial a_k} = 0, \quad k = 0, 1, \ldots, m$$

Warunek ten generuje m+1 równań liniowych z m+1 niewiadomymi:

$$\begin{aligned} \frac{\partial H}{\partial a_k} &= -2 \sum_{j=0}^{n} w(x_j) \left[f(x_j) - \sum_{i=0}^{m} a_i \varphi_i(x_j) \right] \varphi_k(x_j) = 0 \\ k &= 0, 1, 2, \ldots, m \end{aligned}$$

Powyższy układ równań zwany jest **układem normalnym**. Ponieważ funkcje bazowe są liniowo niezależne, istnieje więc dokładnie jedno rozwiązanie minimalizujące wartość H. Układ równań można zapisać w postaci macierzowej (**zakładamy $\omega(x)=1$**):

Uwaga:
- a) Macierz D może nie być kwadratowa np. w tzw. **regresji liniowej** baza jest dwuelementowa {1,x}, a węzłów może być dowolna ilość
- b) $D^T D$ jest macierzą kwadratową i symetryczną o rozmiarach (m+1)x(m+1)

$$D^T D \boldsymbol{a} = D^T \boldsymbol{f}$$

$$D = \begin{bmatrix} \varphi_0(x_0) & \ldots & \varphi_m(x_0) \\ \varphi_0(x_1) & \ldots & \varphi_m(x_1) \\ \ldots & \ldots & \ldots \\ \varphi_0(x_n) & \ldots & \varphi_m(x_n) \end{bmatrix} \qquad \boldsymbol{a} = \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_m \end{bmatrix} \qquad \boldsymbol{f} = \begin{bmatrix} f(x_0) \\ f(x_1) \\ \vdots \\ f(x_n) \end{bmatrix}$$

Aproksymacja średniokwadratowa w bazie jednomianów

Jako bazę przyjmujemy ciąg jednomianów

$$1,\ x,\ x^2,\ \ldots,\ x^m$$

Warunek minimum przyjmuje postać:

$$\sum_{j=0}^{n}\left[f(x_j)-\sum_{i=0}^{m}a_i x_j^i\right]x_j^k=0 \qquad k=0,1,2,\ldots,m$$

po zmianie kolejności sumowania

$$\sum_{i=0}^{m}a_i\left(\sum_{j=0}^{n}x_j^{i+k}\right)=\sum_{j=0}^{n}f(x_j)x_j^k$$

i wprowadzeniu oznaczeń

$$g_{ik}=\sum_{j=0}^{n}x_j^{i+k} \qquad \rho_k=\sum_{j=0}^{n}f(x_j)x_j^k$$

otrzymujemy układ normalny:

$$\sum_{i=0}^{m}a_i g_{ik}=\rho_k \quad \longrightarrow \quad \boxed{G^T\boldsymbol{a}=\boldsymbol{\rho}}$$

Uwagi:

- jeżeli m=n wówczas funkcja aproksymująca pokrywa się z wielomianem interpolującym

- stopień wielomianu aproksymującego powienien być znacznie mniejszy od liczby węzłów x_k, aby „wygładzić” ewentualne błędy pomiarowe

- dla $m \geq 6$ macierz układu staje się źle uwarunkowana (pojedyncza precyzja), najprostszym remedium jest zastosowanie silniejszej arytmetyki (podwójna precyzja)

Aproksymacja średniokwadratowa w bazie wielomianów ortogonalnych

Def. Funkcje f(x) i g(x) nazywamy ortogonalnymi
na dyskretnym zbiorze punktów $x_1, x_2, \ldots, x_n$, jeśli.... .... funkcje f i g spełniają warunki:

$$\sum_{i=0}^{n} f(x_i) g(x_i) = 0$$

$$\sum_{i=0}^{n} [f(x_i)]^2 > 0 \qquad \sum_{i=0}^{n} [g(x_i)]^2 > 0$$

- w aproksymacji średniokwadratowej ciąg funkcyjny

$$\{\varphi_m(x)\} = \varphi_0(x), \varphi_1(x), \ldots, \varphi_m(x)$$

stanowi bazę ortogonalną dla węzłów aproksymacji $x_1, x_2, \ldots, x_n$, jeśli narzucimy dwa warunki:

$$\sum_{i=0}^{n} \varphi_j(x_i) \varphi_k(x_i) = 0, \quad j \neq k$$

oraz nie wszystkie węzły są zerami tych wielomianów

$$\sum_{i=0}^{n} \varphi_j^2(x_i) > 0$$

$$d_{jj} = \sum_{i=0}^{n} \varphi_j^2(x_i)$$

- wówczas macierz układu normalnego przy aproksymacji wielomianami ortogonalnymi jest macierzą diagonalną

$$D^T D = \begin{bmatrix} d_1 & 0 & \ldots & 0 \\ 0 & d_2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & d_n \end{bmatrix}$$

Wielomiany ortogonalne dla węzłów rozmieszczonych dowolnie

- kolejne wielomiany ortogonalne wyznaczamy rekurencyjnie
  tj. na podstawie znajmości postaci wielomianów niższych stopni:

$$\varphi_{j+1}(x) = (x - \alpha_{j+1})\varphi_j(x) - \beta_j \varphi_{j-1}(x) \qquad j = 0, 1, 2, \ldots$$

z warunkami „startowymi”

$$\varphi_{-1}(x) = 0 \qquad \varphi_0(x) = 1$$

Znajdźmy współczynnik α_{j+1}

Jak? - mnożymy przez $\varphi_j(x_i)$ i sumujemy po x_i wykorzystując ortogonalność funkcji

$$\sum_{i=0}^{n} \varphi_j(x_i)\varphi_{j+1}(x_i) = \sum_{i=0}^{n} \varphi_j(x_i)\left[(x_i - \alpha_{j+1})\varphi_j(x_i) - \beta_j \varphi_{j-1}(x_i)\right]$$

$$\sum_{i=0}^{n} \varphi_j(x_i)\varphi_{j+1}(x_i) = \sum_{i=0}^{n} x_i \varphi_j^2(x_i) - \sum_{i=0}^{n} \alpha_{j+1}\varphi_j^2(x_i) \Rightarrow \alpha_{j+1} = \frac{\sum_{i=0}^{n} x_i \varphi_j^2(x_i)}{\sum_{i=0}^{n} \varphi_j^2(x_i)}$$

Identycznie znajdujemy współczynnik
(mnożymy przez $\varphi_{j-1}(x_i)$ i sumujemy po x_i)

$$\beta_j = \frac{\sum_{i=0}^{n} x_i \varphi_{j-1}\varphi_j(x_i)}{\sum_{i=0}^{n} \varphi_{j-1}^2(x_i)}$$

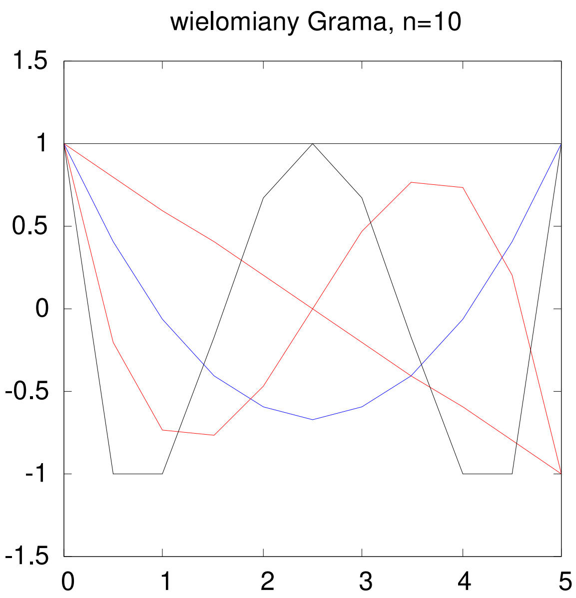
Wielomiany Grama dla 11 węzłów (n=10),h=0.5

$$F(x) = \sum_{k=0}^{m} b_k \varphi_k(x)$$

$$b_k = \frac{C_k}{S_k}$$

$$C_k = \sum_{i=1}^{n} y_i \varphi_k(x_i)$$

$$S_k = \sum_{i=0}^{n} \varphi_k^2(x_i)$$



Aproksymacja średniokwadratowa w bazie funkcji trygonometrycznych

- funkcje okresowe aproksymujemy przy użyciu funkcji trygonometrycznych, czyli w bazie

$$1,\ sin(x),\ cos(x),\ sin(2x),\ cos(2x),\ldots$$

- wielomian trygonometryczny o okresie 2π ma postać:

$$Q_m(x) = \frac{a_0}{2} + \sum_{j=1}^{m} \left[a_j cos(j\,x) + b_j sin(j\,x)\right]$$

- jeśli funkcja f(x) jest określona na dyskretnym zbiorze równoodległych punktów, a liczba punktów jest parzysta i wynosi 2n:

$$x_i = \frac{\pi i}{n}, \quad i = 0, 1, 2, \ldots, 2n-1$$

$$\sum_{i=0}^{2N-1} \sin(mx_i)\sin(kx_i) = \begin{cases} 0, & m \neq k \\ N, & m = k \neq 0 \\ 0, & m = k = 0 \end{cases}$$

$$\sum_{i=0}^{2N-1} \cos(mx_i)\cos(kx_i) = \begin{cases} 0, & m \neq k \\ N, & m = k \neq 0 \\ 2N, & m = k = 0 \end{cases}$$

$$\sum_{i=0}^{2N-1} \cos(mx_i)\sin(kx_i) = 0$$

m,k - dowolne

Szukamy wielomianu w postaci:

$$F(x) = \frac{1}{2}a_0 + \sum_{j=1}^{m} \left[a_j cos(jx) + b_j sin(jx)\right], \qquad m < N$$

Współczynniki a_j oraz b_j wyznacza się z warunku minimalizacji wyrażenia:

$$\sum_{i=0}^{2N-1} \left[f(x_i) - F(x_i)\right]^2 = min$$

co prowadzi do zależności na współczynniki

$$\mathrm{a}_0 = \tfrac{1}{2N} \sum_{i=0}^{2N-1} f(x_i)$$

$$a_{j>0} = \frac{1}{N} \sum_{i=0}^{2N-1} f(x_i) cos(jx_i) = \frac{1}{N} \sum_{i=0}^{2N-1} f(x_i) cos\frac{\pi ij}{N}$$

$$b_j = \frac{1}{N} \sum_{i=0}^{2N-1} f(x_i) sin(jx_i) = \frac{1}{N} \sum_{i=0}^{2N-1} f(x_i) sin\frac{\pi ij}{N}$$

Dobór bazy funkcyjnej

- nierzadko zależy nam na dopasowaniu do danych pomiarowych określonej zależności funkcyjnej (np. wynikającej z zasady działania danego urządzenia pomiarowego lub spodziewanego rozwiązania z modelu teoretycznego)

- często stosuje się poniższe upraszczające formuły aproksymacyjne:

$$\begin{aligned} y &= ax^b + c \\ y &= e^{ax^2+bx+c} \\ y &= ax^2 + bx + c \\ y &= ax^b e^{cx} \end{aligned}$$

naszym zadaniem jest znalezienie współczynników: a, b, c, ...

Aproksymacja Padego

- funkcję aproksymowaną przybliżamy funkcją wymierną tj. ilorazu dwóch wielomianów

$$\boxed{R_{n,k}(x) = \frac{L_n(x)}{M_k(x)}}$$

gdzie: N=n+k

- zaletą powyższego przybliżenia (w problemie aproksymacji jednostajnej) są mniejsze błędy niż aproksymacja wielomianem stopnia N (otrzymanych np. z rozwinięć Taylora czy Maclaurina)

- zadanie polega na znalezieniu N+1 współczynników L_n oraz M_k

$$L_n(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n$$

$$M_k(x) = b_0 + b_1 x + b_2 x^2 + \ldots + b_k x^k, \qquad b_0 \neq 0$$

tak aby w x_0=0 funkcje aproksymowana i aproksymująca miały jak najwięcej równych pochodnych

Rozwijamy f(x) w szereg Maclaurina

$$f(x) = \sum_{i=0}^{\infty} c_i x^i$$

Liczymy **błąd aproksymacji** (w celu otrzymania zależności współczynniki a_i oraz b_i)

$$f(x) - \frac{L_n(x)}{M_k(x)} == \frac{\left(\sum_{i=0}^{\infty} c_i x^i\right)\left(\sum_{i=0}^{k} b_i x^i\right) - \sum_{i=0}^{n} a_i x^i}{\sum_{i=0}^{k} b_i x_i}$$

Wykorzystujemy warunki z ciągłością pochodnych w x=0

$$f^{(m)}(x)\Big|_{x=0} - R_{n,k}^{(m)}(x)\Big|_{x=0} = 0, \qquad m = 0, 1, 2, \ldots, k+n$$

Powyższy warunek będzie spełniony, gdy licznik zapiszemy jako

$$\left(\sum_{i=0}^{\infty} c_i x^i\right)\left(\sum_{i=0}^{k} b_i x^i\right) - \sum_{i=0}^{n} a_i x^i = \sum_{j=1}^{\infty} d_{N+j} x^{N+j} \qquad N = n + k$$

- dla warunku

$$\boxed{f(0) - R_{n,k} = 0}$$

dostajemy równanie

$$(b_0 + b_1 x + \ldots + b_k x^k)(c_0 + c_1 x + \ldots) = (a_0 + a_1 x + \ldots + a_n x^n)$$

z którego wydobywamy zależności

$$\begin{array}{lcl} a_0 & = & b_0 c_0 \\ a_1 & = & b_0 c_1 + b_1 c_0 \\ a_2 & = & b_0 c_2 + b_1 c_1 + b_2 c_0 \\ \ldots & \ldots & \ldots\ldots\ldots\ldots \end{array}$$

i ostatecznie wzór ogólny

$$a_r = \sum_{j=0}^{r} c_{r-j} b_j, \quad r = 0, 1, 2, \ldots, n$$

- wykorzystujemy też założenie o równości pochodnych (do rzędu n+k+1) co daje dodatkową zależność

$$\sum_{j=0}^{k} c_{n+k-s-j} b_j = 0, \quad s = 0, 1, 2, \ldots, k-1$$

Sposób postępowania:

1. Wyznaczamy współczynniki szeregu McLaurina.
   Numerycznie dokładnie – tylko przy użyciu liczb dualnych, ilorazy różnicowe są niedokładne.
   W niektórych przypadkach (rzadko) możliwe jest wykorzystanie wzoru analitycznego na pochodne.

$$s = 0, 1, 2, \ldots, k-1$$

$$\sum_{j=0}^{k} c_{n+k-s-j} b_j = 0$$

2. Tworzymy układ równań, którego rozwiązanie to współczynniki b_i

$$j = 0 \rightarrow b_0 = 1$$

$$\begin{bmatrix} c_{n-m+1} & c_{n-m+2} & \cdots & c_n \\ c_{n-m+2} & c_{n-m+3} & \cdots & c_{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ c_n & c_{n+1} & \cdots & c_{n+m-1} \end{bmatrix} \begin{bmatrix} b_m \\ b_{m-1} \\ \vdots \\ b_1 \end{bmatrix} = \begin{bmatrix} -c_{n+1} \\ -c_{n+2} \\ \vdots \\ -c_{n+m} \end{bmatrix}$$

3. Teraz możemy wyznaczyć kolejno współczynniki a_j

$$a_i = \sum_{j=0}^{i} c_{i-j} \cdot b_j, \quad i = 0, 1, \ldots, n$$

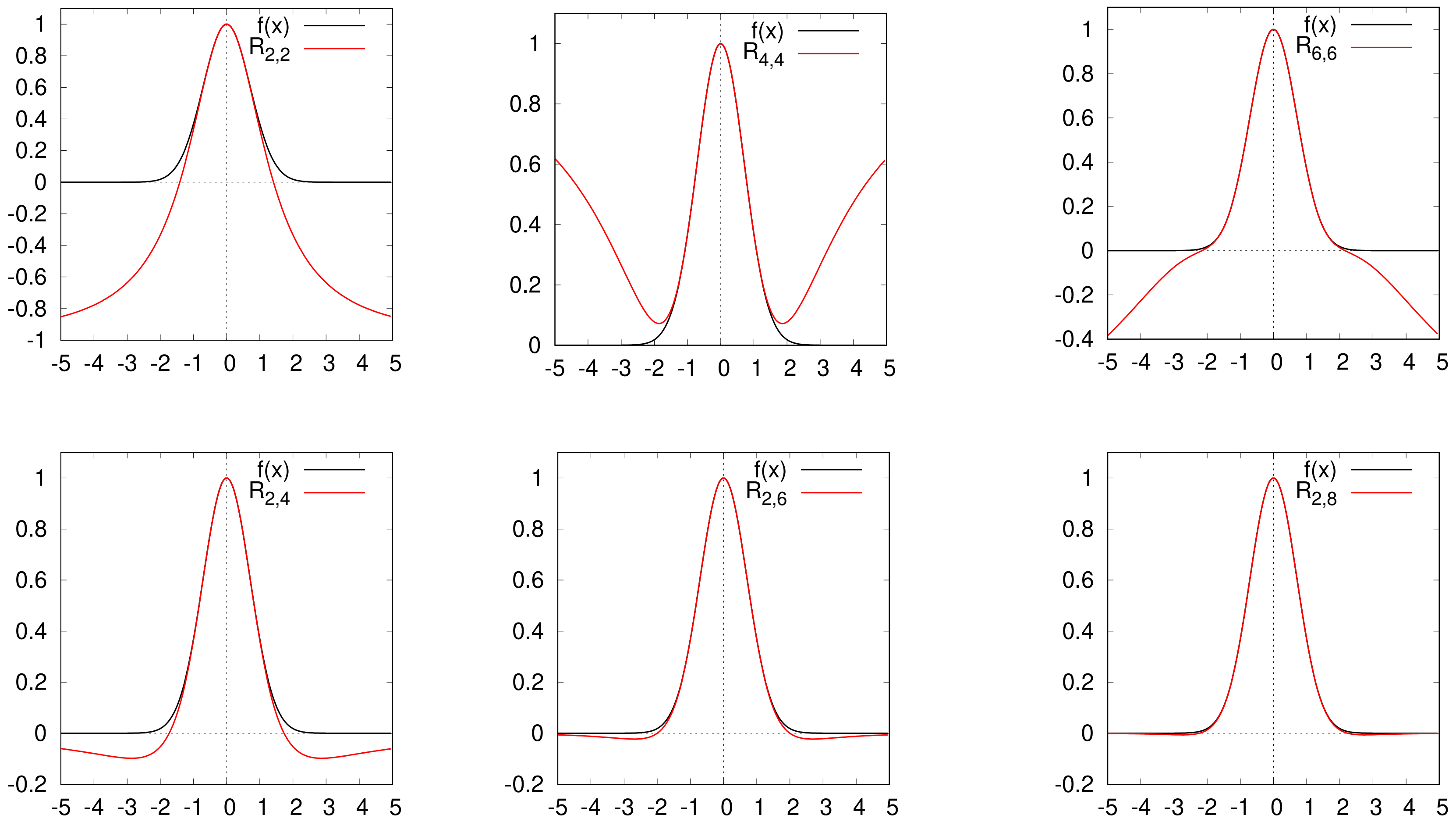
Przykład Aproksymacja Pade funkcji

$$f(x) = exp(-x^2), \quad x \in [-5, 5]$$

$$T_{12}\{f(x)\} \approx 1 - x^2 + \frac{x^4}{2} - \frac{x^6}{6} + \frac{x^8}{24} - \frac{x^{10}}{120} + \frac{x^{12}}{720}$$

Funkcja jest parzysta, więc wielomiany w liczniku i w mianowniku $R_{n,k}$ będą miały niezerowe współczynniki tylko przy jednomianach o wykładnikach parzystych.

$$R_{6,6}(x) = \frac{1 - \frac{x^2}{2} + \frac{x^4}{10} - \frac{x^6}{120}}{1 + \frac{x^2}{2} + \frac{x^4}{10} + \frac{x^6}{120}}$$



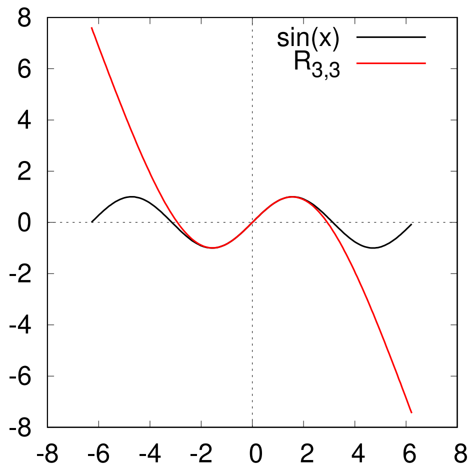
Przykład Aproksymacja Pade funkcji

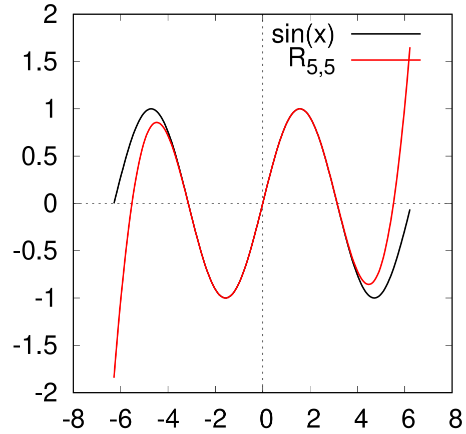
$$f(x) = \sin(x), \quad x \in [-2\pi,\, 2\pi]$$

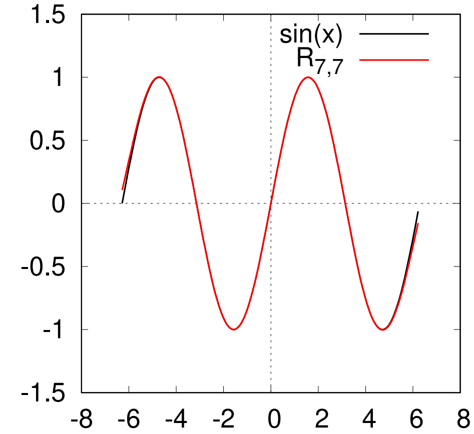
Funkcja aproksymowana jest nieparzysta – niezerowe współczynniki wielmianu L to te stojące przy jednomianach o wyładnikach nieparzystych.

$$R_{3,3}(x) = \frac{x - \frac{7}{60}\,x^3}{1 + \frac{x^2}{20}}$$

$$R_{5,5}(x) = \frac{x - \frac{53}{396}x^3 + \frac{551}{166320}x^5}{1 + \frac{13}{396}\,x^2 + \frac{5}{11088}\,x^4}$$







$$R_{7,7}(x) = \frac{x - \frac{29593}{207636}\,x^3 + \frac{34911}{7613320}\,x^5 - \frac{479249}{11511339840}\,x^7}{1 + \frac{1671}{69212}\,x^2 + \frac{97}{351384}\,x^4 + \frac{2623}{1644477120}\,x^6}$$