Materiały wykładowe
(fragmenty)

Robert Susmaga

Instytut Informatyki

ul. Piotrowo 2

Poznań

<u>kontakt mail’owy</u>

Robert.Susmaga@CS.PUT.Poznan.PL

<u>kontakt osobisty</u>

Centrum Wykładowe, „blok informatyki”, pok. 7

Techniki optymalizacji

Cz. 1

Wyłączenie odpowiedzialności

Prezentowane materiały, będące dodatkiem pomocniczym do wykładów, z konieczności fragmentarycznym i niedopracowanym, należy wykorzystywać z pełną świadomością faktu, że mogą nie być pozbawione przypadkowych błędów, braków, wypaczeń i przeinaczeń :-)

Autor

…

Wymagane (i wykorzystywane) pojęcia algebraiczne

- Podstawowe operacje na wektorach i macierzach
  - (prawie) wszędzie ☺
- Różniczkowanie funkcji
  - metody „newtonowskie”
- Macierz nieujemnie/niedodatnio określona
  - metoda Levenberga-Marquarda, metoda MDS
- Wartości własne macierzy
  - metoda Levenberga-Marquarda, metoda MDS
- …

Wymagane (i wykorzystywane) pojęcia geometryczne

- Przestrzenie wielowymiarowe
  - (prawie) wszędzie ☺
- Interpretacja wektorów w przestrzeniach wielowymiarowych
  - (prawie) wszędzie ☺
- Interpretacja funkcji w przestrzeniach wielowymiarowych
  - (prawie) wszędzie ☺

…

Przedstawiane metody/rozwiązania

- Rodzaje przedstawianych metod:
  klasyczne dla dziedziny optymalizacji <u>ciągłej</u>
  - dawno zdefiniowane
  - dokładnie przebadane
    - to się jednak zmienia!
- Związki z metodami metaheurystycznymi:
  metody klasyczne mają zastosowania podrzędne
  - heurystyczne lub dokładne (w zależności od postaci wyników) metody lokalne
  - potencjalne wykorzystanie: generowanie kolejnych rozwiązań (jako lokalnych ekstremów) w metodach metaheurystycznych

Przedstawiane metody/rozwiązania

- Postać problemu
  - minimalizacja funkcji
    - co w przypadku maksymalizacji?
    - co w przypadku poszukiwania konkretnej wartości?
      - optymalizacja/aproksymacja
  - przy (ewentualnie istniejących) ograniczeniach
    - ograniczenia jednowymiarowe (zakres zmienności)
    - ograniczenia wielowymiarowe /robocze, właściwe/

Przedstawiane metody/rozwiązania

- Postaci funkcji celu i ograniczeń
  - (zawsze) dane analitycznie
    - wzór!
  - (zazwyczaj) dodatkowo uwarunkowane
    - ciągłe, gładkie, …

      (przykłady?)
      - nieraz konkretnie: liniowe, kwadratowe, …

Przedstawiane metody/rozwiązania

- Generowane rozwiązania
  - dokładne
    - w sensie dokładności maszynowej*
  - przybliżone

* temat dokładności maszynowej nie będzie bliżej na tym wykładzie przedstawiany

…

Optymalizacja w języku

- „Słownik wyrazów obcych PWN”

optymalizacja -cji, *ż*, *blm*
1. «organizowanie jakichś działań, procesów itp. w taki sposób, aby dały jak największe efekty przy jak najmniejszych nakładach»
2. *ekon.* «poszukiwanie za pomocą metod matematycznych najlepszego ze względu na wybrane kryterium (np. koszt lub zysk) rozwiązania danego zagadnienia gospodarczego, przy uwzględnieniu określonych ograniczeń»

optymalny
«najlepszy z możliwych w jakichś warunkach»
<fr. optimal>
!Niepoprawnie: Najbardziej optymalne, poprawnie: optymalne, rozwiązanie.

Optymalizacja w języku

- Komentarz: kryterium:
  - warunek („… muszą spełnione być pewne kryteria”)
  - trzecie znaczenie: funkcja celu („wielokryterialne PL”)
  - koszt lub zysk? („… kryterium (np. koszt lub zysk) ...”)

Optymalizacja w zastosowaniach

- Optymalizacja ma zastosowania w takich dziedzinach jak
  - fizyka
  - technika
  - chemia
  - inżynieria
  - informatyka
  - biologia
  - ekonomia
  - ...

Optymalizacja w życiu… Hodży Nasreddina :-)

- Kim jest/był Hodża Nasreddin?
  - bliskowschodni średniowieczny mędrzec-podróżnik, zajmujący się rozwiązywaniem zagadek, m.in. matematycznych i logicznych
  - jego rozmaite przygody są fabularyzowanymi zapisami bliskowschodnich mądrości ludowych
- Więcej o życiu Hodży Nasreddina m.in. w książce „Przygody Hodży Nasreddina” (a jeszcze więcej: w Internecie!)

Optymalizacja w życiu… Hodży Nasreddina :-)

- Problem majątku
  – dwaj bracia odziedziczyli majątek po ojcu, który przykazał im podzielić się nim sprawiedliwie, nie podał jednak konkretnie, które dobra mają przypaść w spadku któremu z braci
  – bracia natychmiast pokłócili się o majątek, ponieważ każdy z nich proponował inny podział pozostałych po ojcu dóbr na dwie części: każdy dzielił rzeczy w taki sposób, aby wartości obu części nie były równe, oczywiście przydzielając sobie część o większej wartości, a swemu bratu część o mniejszej wartości

Optymalizacja w życiu… Hodży Nasreddina :-)

- Jak można rozwiązać powyższy konflikt?
- Jak rozwiązał ten konflikt Hodża Nasreddin?

Optymalizacja w życiu… Hodży Nasreddina :-)

- Możliwe rozwiązanie konfliktu
  - jeden z braci dokonuje podziału dziedziczonych rzeczy na dwie części
  - drugi podejmuje decyzję to tym, która część przypadnie komu w udziale

Optymalizacja w życiu… Hodży Nasreddina :-)

- Interesujące cechy zaproponowanego rozwiązania
  - jeżeli dokonujący podziału podzieli dziedziczone dobra na dwie części o nierównej wartości, to naraża się na to, że (wskutek decyzji drugiego z braci) przypadnie mu w udziale część mniej wartościowa
  - dokonujący podziału powinien więc dążyć do tego, aby różnica wartości obu części spadku była jak najmniejsza, w rezultacie czego bracia zostaną sprawiedliwie obdzieleni spadkiem
    - w idealnym przypadku wartości obu części będą jednakowe! (choć taki podział może nie być możliwy do zrealizowania)

Optymalizacja w życiu… Hodży Nasreddina :-)

- Optymalizacyjny punkt widzenia tego problemu i jego rozwiązania
  – tzw. problem min-max
    - osoba dokonująca podziału wie, że jeżeli któraś z utworzonych przez nią części majątku będzie większej wartości, to osoba wybierająca na pewno przydzieli tę część sobie (wniosek: tworzenie jakiejkolwiek części o wartości większej od innych części nie jest korzystne!)
    - oznacza to, że osoba dokonująca podziału powinna minimalizować (min) wartość największej (max) z tworzonych części

  lub

  – tzw. problem max-min
    - osoba dokonująca podziału wie, że jeżeli któraś z utworzonych przez nią części majątku będzie mniejszej wartości, to osoba wybierająca na pewno przydzieli sobie inną część (wniosek: tworzenie części o wartości mniejszej od innych części nie jest korzystne!)
    - oznacza to, że osoba dokonująca podziału powinna maksymalizować (max) wartość najmniejszej (min) z tworzonych części

…

Problem ekstremów

- Problem ekstremów
  - czyli jak dla nietrywialnego problemu optymalizacji bez ograniczeń (czyli zadania przeszukiwania przestrzeni wielowymiarowych) skonstruować nietrywialne podejście rozwiązujące ten problem

…

Elementy ekstremalne w zbiorach uporządkowanych

- W zbiorze X uporządkowanym relacjami '$\leq$', '$\geq$', '<' oraz '>' można zdefiniować
  - element najmniejszy:
    jest nim $a \in X$ spełniający

    $\forall_{a \neq x \in X}\ a < x$
  - element największy:
    jest nim $b \in X$ spełniający

    $\forall_{b \neq x \in X}\ b > x$

Elementy ekstremalne w zbiorach uporządkowanych

- W zbiorze X uporządkowanym relacjami ‘$\leq$’, ‘$\geq$’, ‘<‘ oraz ‘>’ można zdefiniować
  - element minimalny:
    jest nim $c \in X$ spełniający

    $\forall_{x \in X}\ c \leq x \quad \Leftrightarrow \quad \neg(\exists_{x \in X}\ x < c)$

  - element maksymalny:
    jest nim $d \in X$ spełniający

    $\forall_{x \in X}\ d \geq x \quad \Leftrightarrow \quad \neg(\exists_{x \in X}\ x > c)$

Dygresja

- Quiz
  - czy istnieje element najmniejszy zbioru $\{a_1, a_2, a_3, a_4, a_5, a_6, a_7, a_8, a_9\}$, gdzie $a_1=3$, $a_2=9$, $a_3=2$, $a_4=7$, $a_5=1$, $a_6=9$, $a_7=4$, $a_8=4$, $a_9=5$?
    - tak!
    - jest nim: a_5 (równe 1)

Dygresja

- Quiz
  - czy istnieje element minimalny zbioru $\{a_1, a_2, a_3, a_4, a_5, a_6, a_7, a_8, a_9\}$, gdzie a_1=3, a_2 =9, a_3 =2, a_4 =7, a_5 =1, a_6 =9, a_7 =4, a_8 =4, a_9 =5?
    - tak!
    - jest nim: a_5 (równe 1)

Dygresja

- Quiz
  - czy istnieje element największy zbioru $\{a_1, a_2, a_3, a_4, a_5, a_6, a_7, a_8, a_9\}$, gdzie $a_1=3$, $a_2=9$, $a_3=2$, $a_4=7$, $a_5=1$, $a_6=9$, $a_7=4$, $a_8=4$, $a_9=5$?
    - nie!

Dygresja

- Quiz
  - czy istnieje element maksymalny zbioru $\{a_1, a_2, a_3, a_4, a_5, a_6, a_7, a_8, a_9\}$, gdzie $a_1=3$, $a_2=9$, $a_3=2$, $a_4=7$, $a_5=1$, $a_6=9$, $a_7=4$, $a_8=4$, $a_9=5$?
    - tak! (i to nie jeden, lecz dwa!)
    - są nimi: a_2 (równe 9) oraz a_6 (równe 9)

Elementy ekstremalne w zbiorach uporządkowanych

- Definicje elementu minimalnego/maksymalnego wykorzystuje się w przypadku problemu optymalizacji funkcji f(x)
  - zbiorem uporządkowanym jest przeciwdziedzina funkcji f(x) (która dla funkcji rzeczywistej stanowi podzbiór zbioru liczb rzeczywistych)
  - poszukiwany jest argument funkcji (czyli element jej dziedziny D), dla którego wartość tej funkcji jest <u>minimalna</u> (niekoniecznie najmniejsza), czyli $x^* \in D$ taki, że $\forall_{x \in D}\ f(x^*) \leq f(x)$
  - skrócony zapis powyższej zależności:
    $x^* = \mathrm{argmin}_{x \in D} f(x)$
    (x^* będzie dalej nazywany „rozwiązaniem (funkcji)")

    (czym „argmin" różni się od „min"?)

…

Pochodne funkcji w postaci wektorowo/macierzowej

- Pochodne prostych funkcji w postaci skalarnej (przypomnienie)
  - afinicznej (popularnie zwanej liniową)
    - $f(x) = ax + b$
    - $\partial f/\partial x = a$
    - $\partial^2 f/\partial x^2 = 0$
    - $\partial^3 f/\partial x^3 = 0$
    - …
  - kwadratowej:
    - $f(x) = ax^2 + bx + c$
    - $\partial f/\partial x = 2ax + b$
    - $\partial^2 f/\partial x^2 = 2a$
    - $\partial^3 f/\partial x^3 = 0$
    - $\partial^4 f/\partial x^4 = 0$
    - …

Pochodne funkcji w postaci wektorowo/macierzowej

- Pochodne prostych funkcji w postaci skalarnej (przypomnienie)
  - liniowej
    - $f(x) = ax$
    - $\partial f/\partial x = a$
    - $\partial^2 f/\partial x^2 = 0$
    - $\partial^3 f/\partial x^3 = 0$
    - …
  - (ściśle) kwadratowej:
    - $f(x) = ax^2$
    - $\partial f/\partial x = 2ax$
    - $\partial^2 f/\partial x^2 = 2a$
    - $\partial^3 f/\partial x^3 = 0$
    - $\partial^4 f/\partial x^4 = 0$
    - …

Pochodne funkcji w postaci wektorowo/macierzowej

- Pochodne prostych funkcji w postaci wektorowo/macierzowej
  - liniowej:
    - $f(\mathbf{x}) = \mathbf{a}^T\mathbf{x}$
    - $\partial f/\partial \mathbf{x} = \mathbf{a}$
  - formy kwadratowej:
    - $f(\mathbf{x}) = \mathbf{x}^T\mathbf{A}\mathbf{x}$
    - $\partial f/\partial \mathbf{x} = (\mathbf{A} + \mathbf{A}^T)\mathbf{x}$
    - $\partial^2 f/\partial \mathbf{x}^2 = \mathbf{A} + \mathbf{A}^T$

      w szczególnym przypadku, gdy $\mathbf{A}^T = \mathbf{A}$ (czyli macierz **A** jest symetryczna)
      - $\partial^2 f/\partial \mathbf{x}^2 = \mathbf{A} + \mathbf{A}^T = \mathbf{A} + \mathbf{A} = 2\mathbf{A}$ (jednocześnie $\partial^2 f/\partial \mathbf{x}^2 = 2\mathbf{A}^T$)

      w bardzo szczególnym przypadku, gdy $\mathbf{A} = [a]$ (czyli macierz **A** reprezentuje skalar)
      - $\partial^2 f/\partial \mathbf{x}^2 = [a] + [a]^T = [a] + [a] = 2[a]$ (czyli, właściwie, $\partial^2 f/\partial \mathbf{x}^2 = 2a$)

Pochodne funkcji w postaci wektorowo/macierzowej

- Pochodne prostych funkcji w postaci wektorowo/macierzowej
  - afinicznej:
    - $f(\mathbf{x}) = \mathbf{a}^T\mathbf{x} + b$
    - $\partial f/\partial \mathbf{x} = \mathbf{a}$
  - (pełnej) kwadratowej:
    - $f(\mathbf{x}) = \mathbf{x}^T\mathbf{A}\mathbf{x} + \mathbf{b}^T\mathbf{x} + c$
    - $\partial f/\partial \mathbf{x} = (\mathbf{A} + \mathbf{A}^T)\mathbf{x} + \mathbf{b}$
    - $\partial^2 f/\partial \mathbf{x}^2 = \mathbf{A} + \mathbf{A}^T$

Pochodne funkcji w postaci wektorowo/macierzowej

- Pochodne prostych funkcji w postaci wektorowo/macierzowej
  - $f(x) = ax^1$
    - $\partial f/\partial x = a$
  - $f(x) = ax^2$
    - $\partial f/\partial x = 2ax$
  - $f(x) = ax^3$
    - $\partial f/\partial x = 3ax^2$

  ...

  - $f(x) = a(x - x_0)^1$
    - $\partial f/\partial x = a$
  - $f(x) = a(x - x_0)^2$
    - $\partial f/\partial x = 2a(x - x_0)$
  - $f(x) = a(x - x_0)^3$
    - $\partial f/\partial x = 3a(x - x_0)^2$

  ...

Pochodne funkcji w postaci wektorowo/macierzowej

- Oznaczenie pochodnych wielokrotnych
  - $f^{(k)}(x)$ – oznaczenie k-tej pochodnej funkcji f(x)
  - w szczególności
    - $f^{(0)}(x) \equiv f(x)$ – funkcja
    - $f^{(1)}(x) \equiv f'(x)$ – jej pierwsza pochodna
    - $f^{(2)}(x) \equiv f''(x)$ – jej druga pochodna
    - …

…

Szereg Taylora

- „Generator” wzorów funkcji, np. $f(x) = e^x$
  - rozwinięcie nieskończone:

    $e^x = 1 + x + x^2/2! + x^3/3! + x^4/4! + x^5/5! + ...$

  - rozwinięcie skończone, sześcioelementowe:

    $e^x = 1 + x + x^2/2! + x^3/3! + x^4/4! + x^5/5!$

Szereg Taylora

- Dane jest wyrażenie $T(x) = \sum_{k=0..\infty} a_k(x - x_0)^k$
  - jest ono zależne od zmiennego argumentu x, ustalonej wartości x_0 oraz ustalonych wartości a_k (dla $k=0..\infty$)
  - wyrażenie to reprezentuje sumę nieskończonego ciągu o elementach $a_k(x - x_0)^k$ (dla $k=0..\infty$)
    - T(x) jest więc sumą nieskończonej liczby elementów
- Przyjmując, że $w^0 = 1$ dla wszystkich możliwych w (także dla $w = 0$), wyrażenie T(x) można przedstawić w postaci $T(x) = a_0 + \sum_{k=1..\infty} a_k(x - x_0)^k$
  - gdy dla wszystkich k większych od pewnego ustalonego n zachodzi $a_k(x - x_0)^k = 0$, wyrażenie T(x) można zapisać w postaci $T_n(x) = a_0 + \sum_{k=1..n} a_k(x - x_0)^k$
    - wyrażenie to reprezentuje wtedy sumę skończonego ciągu o elementach $a_k(x - x_0)^k$ (dla $k=0..n$)
    - $T_n(x)$ jest więc sumą skończonej liczby elementów, a konkretniej: wielomianem stopnia n od argumentu x

Szereg Taylora

- Dzięki podobieństwu do wielomianu, wyrażenie T(x) może być różniczkowane (operacja jest analogiczna do różniczkowania wielomianów)
- Tzn. jeżeli:

$$T(x) = a_0 + a_1(x - x_0) + a_2(x - x_0)^2 + a_3(x - x_0)^3 + a_4(x - x_0)^4 + \dots$$

- to:

$$T'(x) = 0 + a_1 + 2a_2(x - x_0) + 3a_3(x - x_0)^2 + 4a_4(x - x_0)^3 + \dots$$

$$T''(x) = 0 + 0 + 2a_2 + 3 \cdot 2a_3(x - x_0) + 4 \cdot 3a_4(x - x_0)^2 + \dots$$

$$T'''(x) = 0 + 0 + 0 + 3 \cdot 2a_3 + 4 \cdot 3 \cdot 2a_4(x - x_0) + \dots$$

…

Szereg Taylora

- Wyrażenie T(x) może zostać użyte do wyrażania (w przybliżony lub dokładny sposób) wartości pewnej funkcji f(x) (czyli funkcji zależnej od argumentu x)
  - wyrażanie to ma szanse powodzenia, gdy możliwe jest znalezienie wartości x_0 oraz a_k (dla $k=0..\infty$), które gwarantują f(x) = T(x) dla wszystkich (lub wybranych) x należących do dziedziny funkcji f(x)
    - mówimy wtedy, że dokonano rozwinięcia wartości funkcji f(x) w szereg Taylora
  - często operacji tej dokonuje się najpierw ustalając wartość x_0, a potem dopiero wartości a_k (dla $k=0..\infty$)
    - mówimy wtedy, że dokonano rozwinięcia wartości funkcji f(x) wokół wartości x_0
  - operacja ta jest szczególnie łatwa dla funkcji wielokrotnie różniczkowalnych

Szereg Taylora

- Niech dana będzie wielokrotnie różniczkowalna funkcja f(x), której wartość ma być wyrażona z użyciem T(x), a więc zakładamy istnienie a_i oraz x_0 takich, że:

  $$f(x) = a_0 + a_1(x - x_0) + a_2(x - x_0)^2 + a_3(x - x_0)^3 + a_4(x - x_0)^4 + \ldots$$

Szereg Taylora

- Wystarczy więc tylko znaleźć wartości a_i oraz x_0 i…

  rozwinięcie funkcji f(x) w szereg Taylora jest gotowe!

Szereg Taylora

- Proces poszukiwania współczynników a_k
  - współczynnik a_0
    - równanie f(x) = T(x) ma postać
      $f(x) = a_0 + a_1(x - x_0) + a_2(x - x_0)^2 + a_3(x - x_0)^3 + a_4(x - x_0)^4 + \ldots$
    - niech $x = x_0$, wtedy
      $f(x_0) = a_0 + a_1(x_0 - x_0) + a_2(x_0 - x_0)^2 + a_3(x_0 - x_0)^3 + a_4(x_0 - x_0)^4 + \ldots$
      $f(x_0) = a_0 + a_1 \cdot 0 + a_2 \cdot 0^2 + a_3 \cdot 0^3 + a_4 \cdot 0^4 + \ldots$
      $f(x_0) = a_0$
    - a więc a_0 można ustalić obliczając $f(x_0)$

Szereg Taylora

- Proces poszukiwania współczynników a_k
  - współczynnik a_1
    - w rezultacie jednokrotnego zróżniczkowania (ze względu na x) obu stron równania f(x) = T(x) otrzymujemy
      $f'(x) = a_1 + 2a_2(x - x_0) + 3a_3(x - x_0)^2 + 4a_4(x - x_0)^3 + \ldots$
    - niech $x = x_0$, wtedy
      $f'(x_0) = a_1 + 2a_2(x_0 - x_0) + 3a_3(x_0 - x_0)^2 + 4a_4(x_0 - x_0)^3 + \ldots$
      $f'(x_0) = a_1 + 2a_2 \cdot 0^2 + 3a_3 \cdot 0^3 + 4a_4 \cdot 0^4 + \ldots$
      $f'(x_0) = a_1$
    - a więc a_1 można ustalić obliczając $f'(x_0)$

Szereg Taylora

- Proces poszukiwania współczynników a_k
  - współczynnik a_2
    - w rezultacie dwukrotnego zróżniczkowania (ze względu na x) obu stron równania f(x) = T(x) otrzymujemy
      $f''(x) = 2a_2 + 3 \cdot 2a_3(x - x_0) + 4 \cdot 3a_4(x - x_0)^2 + \dots$
    - niech $x = x_0$, wtedy
      $f''(x_0) = 2a_2 + 3 \cdot 2a_3(x_0 - x_0) + 4 \cdot 3a_4(x_0 - x_0)^2 + \dots$
      $f''(x_0) = 2a_2 + 3 \cdot 2a_3 \cdot 0 + 4 \cdot 3a_4 \cdot 0^2 + \dots$
      $f''(x_0) = 2a_2$
    - a więc a_2 można ustalić obliczając $f''(x_0)/2$

Szereg Taylora

- Proces poszukiwania współczynników a_k
  - współczynnik a_3
    - w rezultacie trzykrotnego zróżniczkowania (ze względu na x) obu stron równania f(x) = T(x) otrzymujemy
      $f'''(x) = 3 \cdot 2a_3 + 4 \cdot 3 \cdot 2a_4(x - x_0) + \ldots$
    - niech $x = x_0$, wtedy
      $f'''(x_0) = 3 \cdot 2a_3(x_0 - x_0) + 4 \cdot 3a_4(x_0 - x_0)^2 + \ldots$
      $f'''(x_0) = 3 \cdot 2a_3 \cdot 0 + 4 \cdot 3a_4 \cdot 0^2 + \ldots$
      $f'''(x_0) = 3 \cdot 2a_3$
    - a więc a_3 można ustalić obliczając $f'''(x_0)/(2 \cdot 3)$

Szereg Taylora

- Proces poszukiwania współczynników a_k
  - współczynnik a_4
    - w rezultacie czterokrotnego zróżniczkowania (ze względu na x) obu stron równania f(x) = T(x) otrzymujemy
      $f''''(x) = 4 \cdot 3 \cdot 2a_4 + \ldots$
    - niech $x = x_0$, wtedy
      $f''''(x_0) = 4 \cdot 3 \cdot 2a_4 + \ldots$
      $f''''(x_0) = 4 \cdot 3 \cdot 2a_4$
    - a więc a_4 można ustalić obliczając $f''''(x_0)/(2 \cdot 3 \cdot 4)$

Szereg Taylora

- Proces poszukiwania współczynników a_k
  - współczynnik a_k (w ogólności)
    - w rezultacie k-krotnego zróżniczkowania (ze względu na x) obu stron równania $f(x) = T(x)$ otrzymujemy $f^{(k)}(x) = k \cdot (k-1) \cdot \ldots \cdot 2 \cdot a_k + \ldots$
    - niech $x = x_0$, wtedy
      $f^{(k)}(x) = k \cdot (k-1) \cdot \ldots \cdot 2 \cdot a_k = k! \cdot a_k$
    - a więc a_k można ustalić obliczając $f^{(k)}(x_0)/(k!)$

Szereg Taylora

- Czyli dla wielokrotnie różniczkowalnej f(x) mamy
  $f(x) = f(x_0) + f'(x_0)(x - x_0) + f''(x_0)/2 \cdot (x - x_0)^2 + f'''(x_0)/3! \cdot (x - x_0)^3 + \ldots$
- Wykorzystując 0! = 1! = 1 mamy:

$$f(x) = \sum_{k=0..\infty} f^{(k)}(x_0)/k! \cdot (x - x_0)^k$$

(założenie: istnienie $f'(x_0)$, $f''(x_0)$, ..., $f^{(k)}(x_0)$)

Szereg Taylora

- Dzięki temu, że k! szybko rośnie, w wielu przypadkach, dla odpowiednio dużych k zachodzi

  $f^{(k)}(x_0)(x - x_0)^k << k!$

- Oczywiście wtedy:

  $f^{(k)}(x_0)(x - x_0)^k/k! << 1$ czy wręcz $f^{(k)}(x_0)(x - x_0)^k/k! \approx 0$

Szereg Taylora

- Dzięki temu możliwe jest „skrócenie” szeregu do kilku (np. n) początkowych elementów (czyli tych, dla których zakładamy, że $f^{(k)}(x_0)(x - x_0)^k/k! \approx 0$ nie zachodzi)

$$f(x) \approx \sum_{k=0..n} f^{(k)}(x_0)/k! \cdot (x - x_0)^k$$

- Gdy „skrócenie” jest niemożliwe, stosuje się zapis pozwalający na wyróżnienie tzw. reszty (oznaczenie R_{n+1})

$$f(x) = \sum_{k=0..\infty} f^{(k)}(x_0)/k! \cdot (x - x_0)^k = \sum_{k=0..n} f^{(k)}(x_0)/k! \cdot (x - x_0)^k + R_{n+1}$$

Szereg Taylora

- Szereg T(x) nosi nazwę szeregu Taylora
  - gdy $x_0 = 0$, szereg Taylora nazywa się szeregiem MacLaurina
- Wiele popularnych funkcji analitycznych posiada rozwinięcia w szereg Taylora (względnie MacLaurina)
  - wniosek: funkcje te dają się przedstawić w postaci „nieskończonego wielomianu”
    - zyski z powyższego dotyczą:
      - interpretacji
      - ewaluacji

Szereg Taylora

- Pytanie:
  - jeżeli (nawet skomplikowane) funkcje (np. e^x) posiadają rozwinięcia Taylora (czyli można je przedstawiać w postaci wielomianu), to może warto w ogóle zrezygnować z posługiwania się tymi funkcjami i po prostu wszędzie używać ich rozwinięć?
  - czyli np. przyjąć „raz na dobre”, że

$$f(x) \approx T_n(x) = \sum_{k=0..n} f^{(k)}(x_0)/k! \cdot (x - x_0)^k$$

(oczywiście pamiętając, że $T_n(x)$ jest jedynie przybliżeniem)

(ale o kontrolowalnej dokładności /wpływ parametru n/)

Szereg Taylora

- Odpowiedź:
  - mimo wszystko nie warto! :-) (wręcz: nie można!)
    - przyczyna: rozwinięcie jest lokalne
      (czyli: inne dla każdego x_0)

Szereg Taylora

- Przykład: sześcioelementowe $T_5(x)$ rozwinięcie funkcji e^x w szereg Taylora wokół wartości $x_0 = 0$
  - funkcja i jej pochodne:
    - $(e^x)^{(0)} = e^x$
    - $(e^x)^{(1)} = (e_x)' = e^x$
    - $(e^x)^{(2)} = (e_x)'' = e^x$
    - $(e^x)^{(3)} = (e_x)''' = e^x$
    - $(e^x)^{(4)} = (e_x)'''' = e^x$
    - $(e^x)^{(5)} = (e_x)''''' = e^x$
    - ...

Szereg Taylora

- Przykład: sześcioelementowe $T_5(x)$ rozwinięcie funkcji e^x w szereg Taylora wokół wartości $x_0 = 0$
  - współczynniki a_k
    - $a_0 = f^{(0)}(x_0)/(0!) = e^0/1 = 1/1 = 1$
    - $a_1 = f^{(1)}(x_0)/(1!) = e^0/1 = 1/1 = 1$
    - $a_2 = f^{(2)}(x_0)/(2!) = e^0/2 = 1/2$
    - $a_3 = f^{(3)}(x_0)/(3!) = e^0/6 = 1/6$
    - $a_4 = f^{(4)}(x_0)/(4!) = e^0/24 = 1/24$
    - $a_5 = f^{(5)}(x_0)/(5!) = e^0/120 = 1/120$
    - ...

Szereg Taylora

- Ostateczny wzór:
  - rozwinięcie nieskończone

$e^x = T(x) = \sum_{k=0..\infty} f^{(k)}(x_0)/k!\cdot(x - x_0)^k = 1 + x + x^2/2! + x^3/3! + x^4/4! + x^5/5! + ...$

  - rozwinięcie skończone, sześcioelementowe:

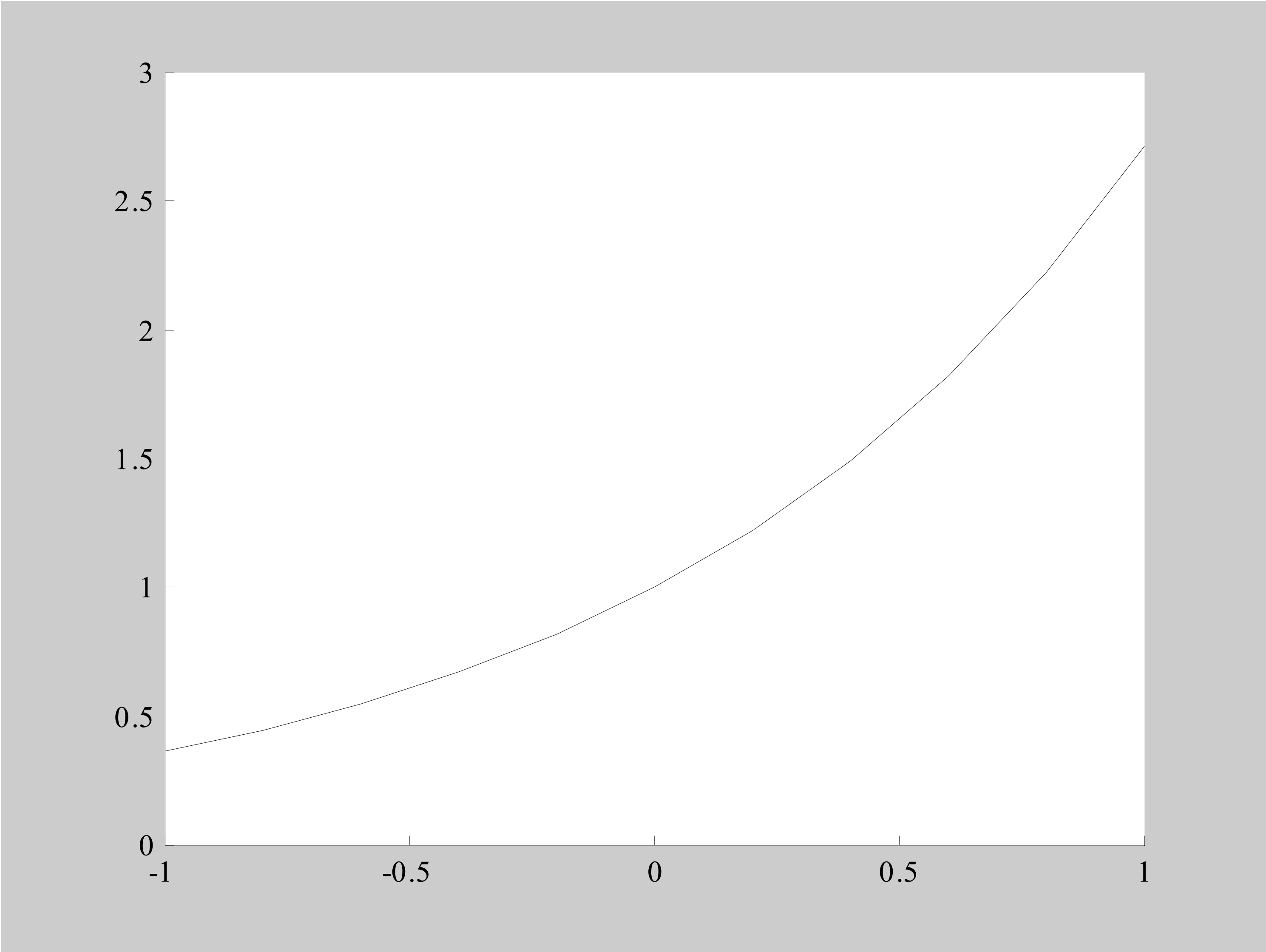
$e^x \approx T_5(x) = \sum_{k=0..5} f^{(k)}(x_0)/k!\cdot(x - x_0)^k = 1 + x + x^2/2! + x^3/3! + x^4/4! + x^5/5!$

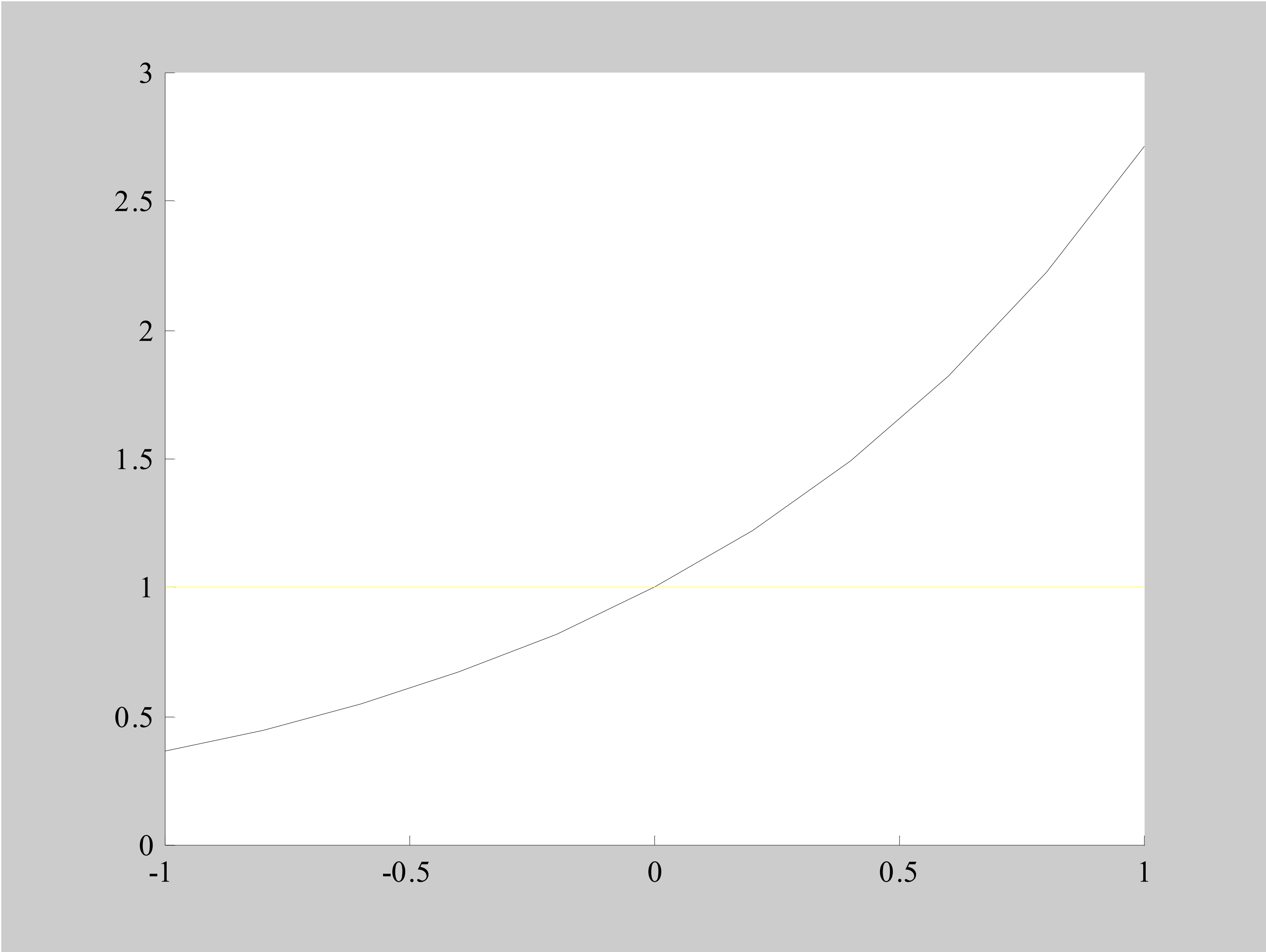
Szereg Taylora

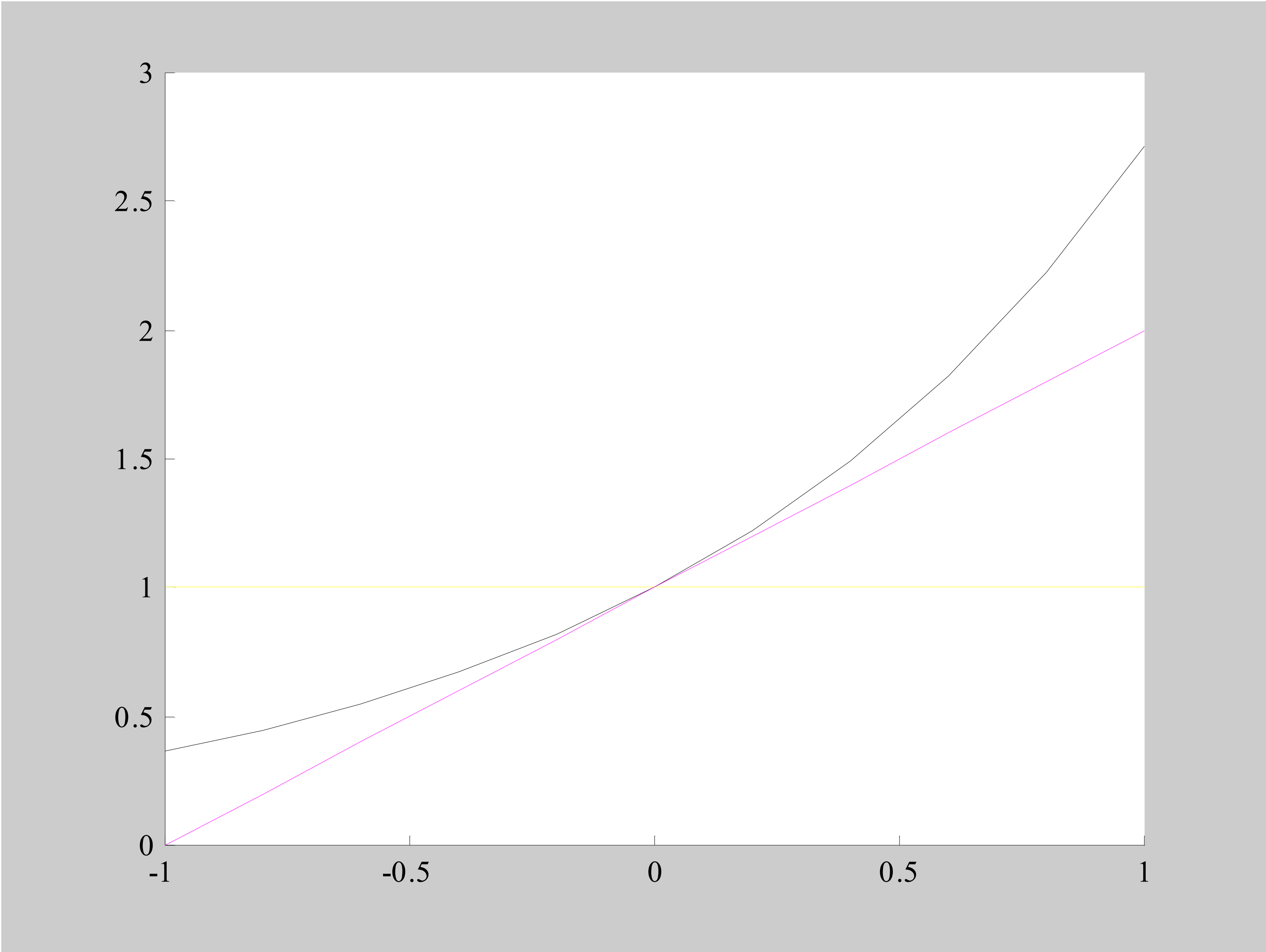
x	$T_5(x)$	e^x
....	....	....
0.000000000000000	1.000000000000000	1.000000000000000
....	....	....
2^{-4} = 0.062500000000000	1.064494458834330	1.064494458917860
2^{-3} = 0.125000000000000	1.133148447672526	1.133148453066826
2^{-2} = 0.250000000000000	1.284025065104167	1.284025416687741
2^{-1} = 0.500000000000000	1.648697916666667	1.648721270700128
2^{0} = 1.000000000000000	2.716666666666666	2.718281828459046
2^{+1} = 2.000000000000000	7.266666666666667	7.389056098930650
2^{+2} = 4.000000000000000	42.86666666666667	54.59815003314423
2^{+3} = 8.000000000000000	570.0666666666667	2980.957987041728
2^{+4} = 16.00000000000000	12296.46666666667	8886110.520507872
....	....	....

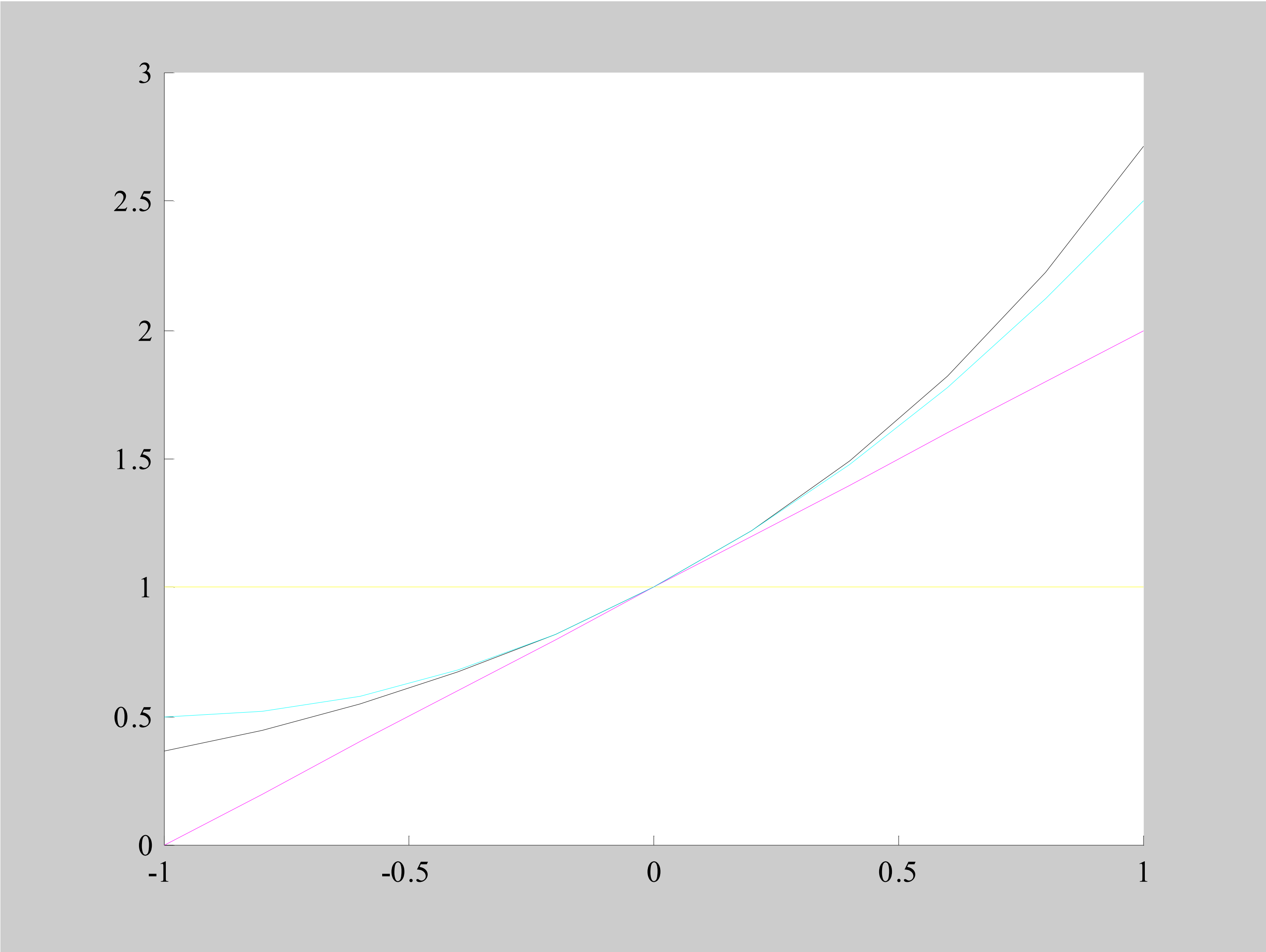
Szereg Taylora

- Kolory wykresów:
  - $f(x) = e^x$
  - $w_0(x) = 1$
  - $w_1(x) = 1 + x$
  - $w_2(x) = 1 + x + x^2/2!$
  - $w_3(x) = 1 + x + x^2/2! + x^3/3!$
  - $w_4(x) = 1 + x + x^2/2! + x^3/3! + x^4/4!$
  - $w_5(x) = 1 + x + x^2/2! + x^3/3! + x^4/4! + x^5/5!$

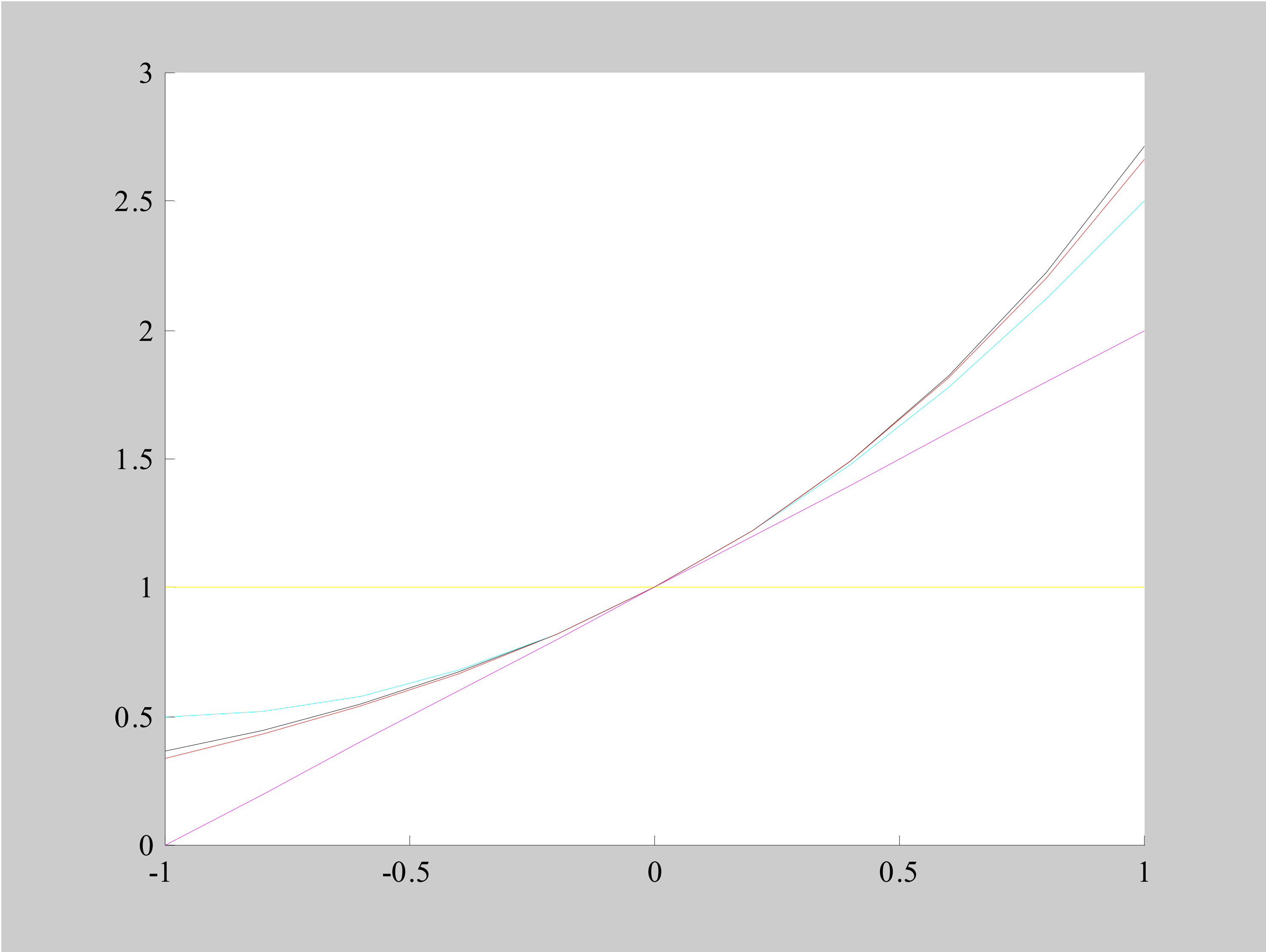

3
2.5
2
1.5
1
0.5
0
-1
-0.5
0
0.5
1

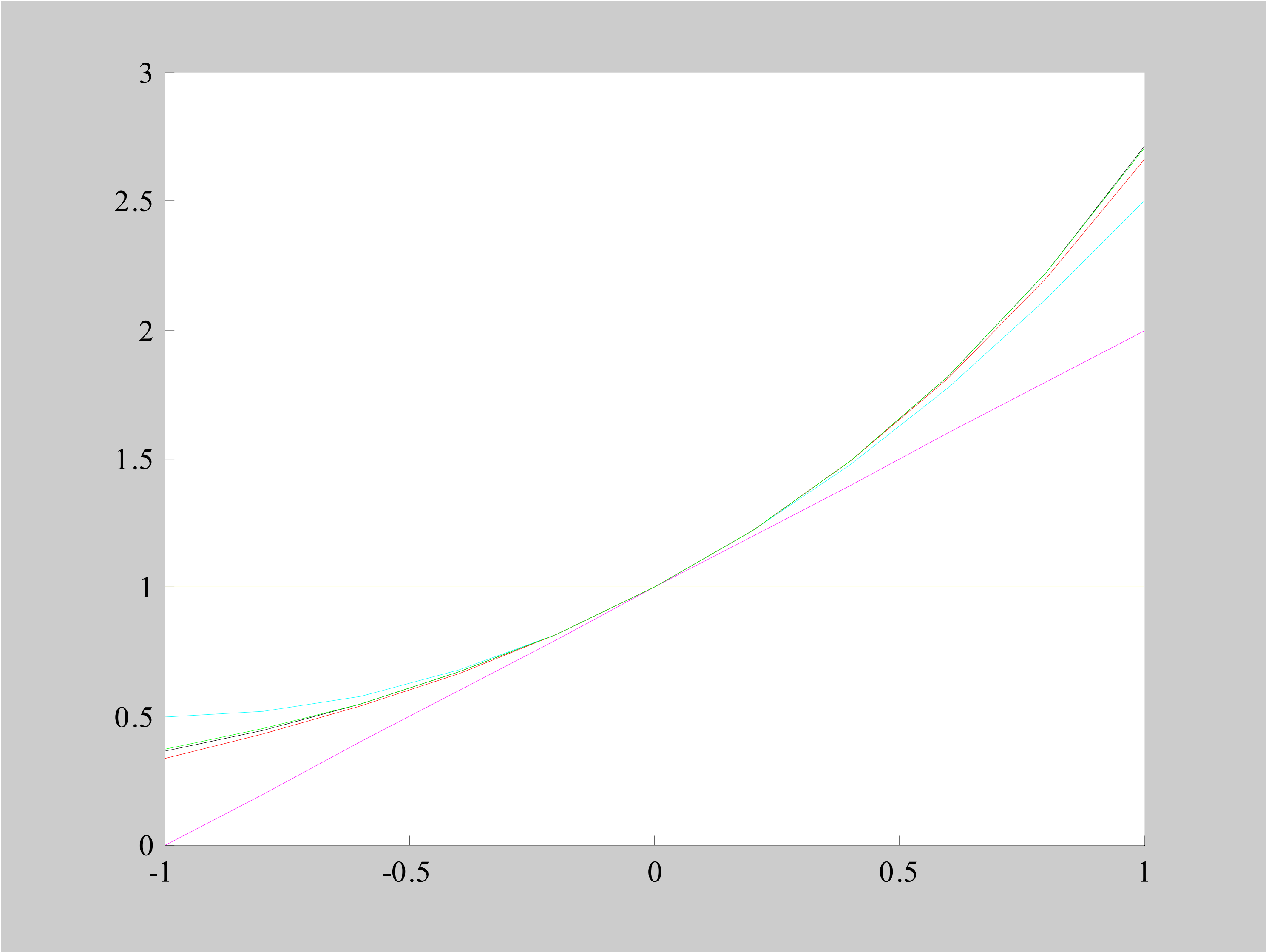

3
2.5
2
1.5
1
0.5
0
-1
-0.5
0
0.5
1

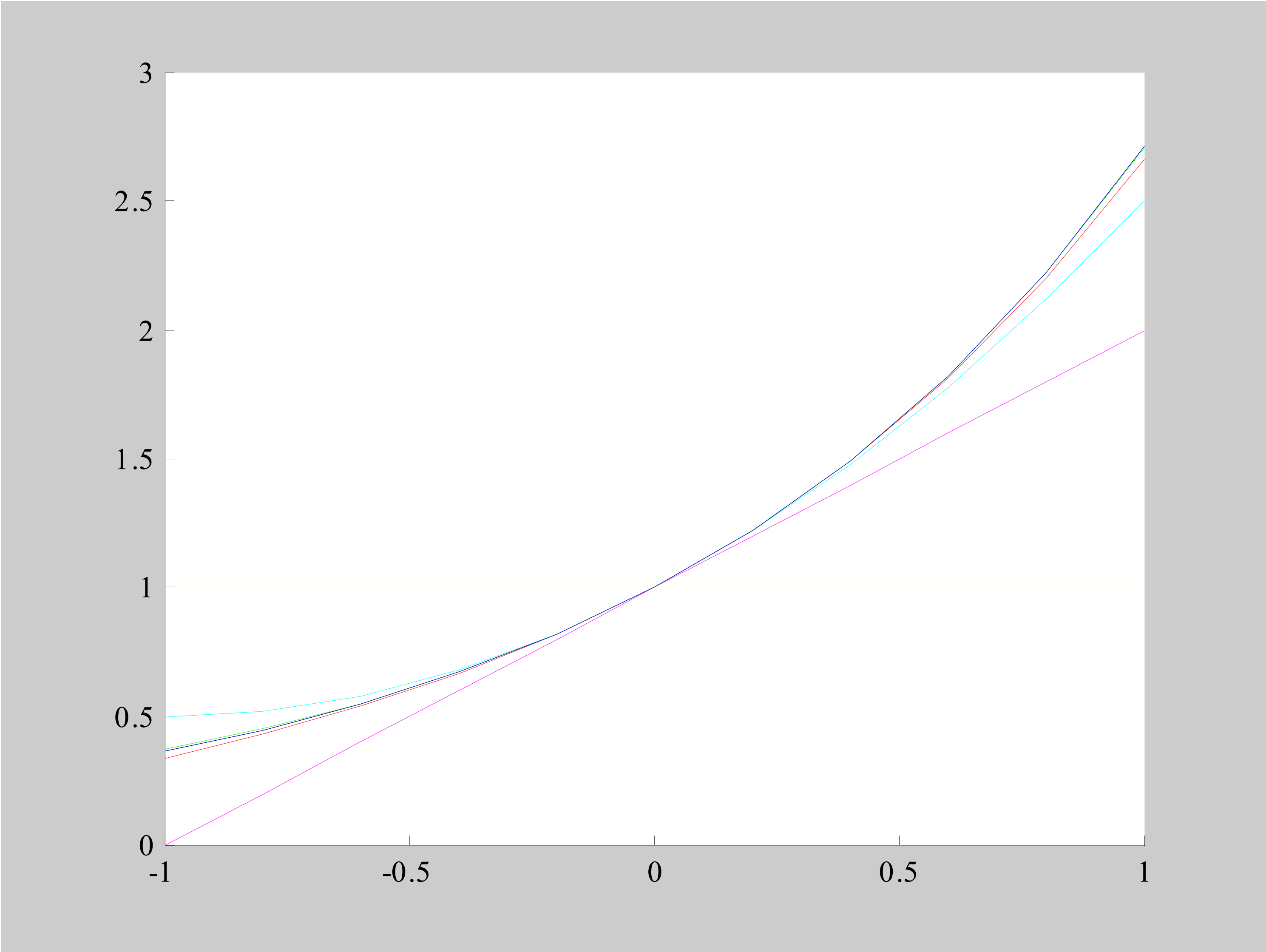

3
2.5
2
1.5
1
0.5
0
-1
-0.5
0
0.5
1



3
2.5
2
1.5
1
0.5
0
-1
-0.5
0
0.5
1


3
2.5
2
1.5
1
0.5
0
-1
-0.5
0
0.5
1


3
2.5
2
1.5
1
0.5
0
-1
-0.5
0
0.5
1



3
2.5
2
1.5
1
0.5
0
-1
-0.5
0
0.5
1

Szereg Taylora

- Inny przykład: dwunastoelementowe $T_{11}(x)$ rozwinięcie funkcji sin(x) w szereg Taylora wokół wartości $x_0 = 0$

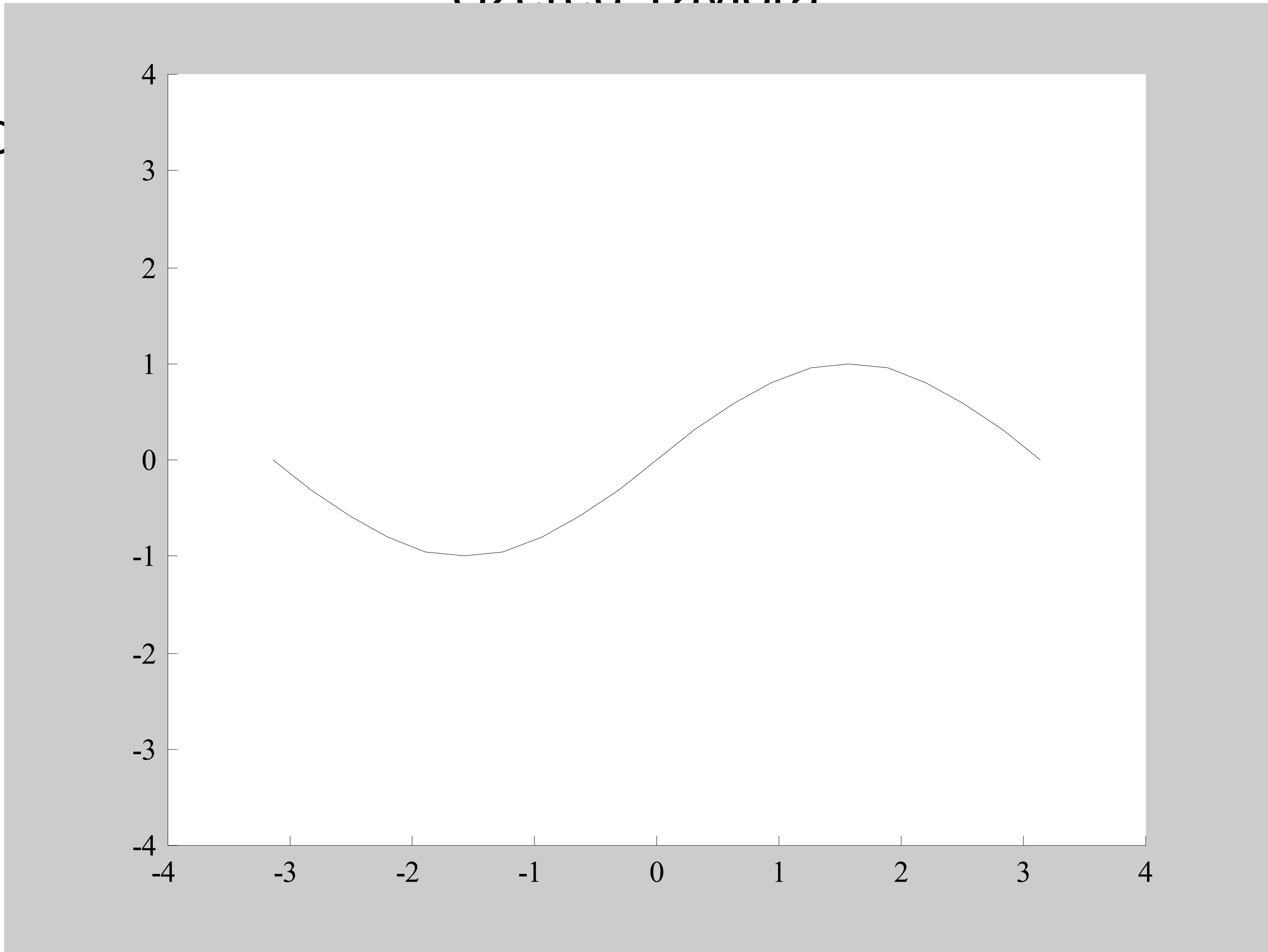
$$\sin(x) \approx T(x) = \sum_{k=0..\infty} f^{(k)}(x_0)/k! \cdot (x - x_0)^k =$$
$$= 0 + x + 0 - x^3/3! + 0 + x^5/5! + 0 - x^7/7! + 0 + x^9/9! + 0 - x^{11}/11! + ...$$

Szereg Taylora

- Kolory wykresów:
  - f(x) = sin(x)
  - $w_0(x) = 0$
  - $w_1(x) = x$
  - $w_2(x) = x$
  - $w_3(x) = x - x^3/3!$
  - $w_4(x) = x - x^3/3!$
  - $w_5(x) = x - x^3/3! + x^5/5!$
  - $w_6(x) = x - x^3/3! + x^5/5!$
  - $w_7(x) = x - x^3/3! + x^5/5! - x^7/7!$
  - $w_8(x) = x - x^3/3! + x^5/5! - x^7/7!$
  - $w_9(x) = x - x^3/3! + x^5/5! - x^7/7! + x^9/9!$
  - $w_{10}(x) = x - x^3/3! + x^5/5! - x^7/7! + x^9/9!$
  - $w_{11}(x) = x - x^3/3! + x^5/5! - x^7/7! + x^9/9! - x^{11}/11!$

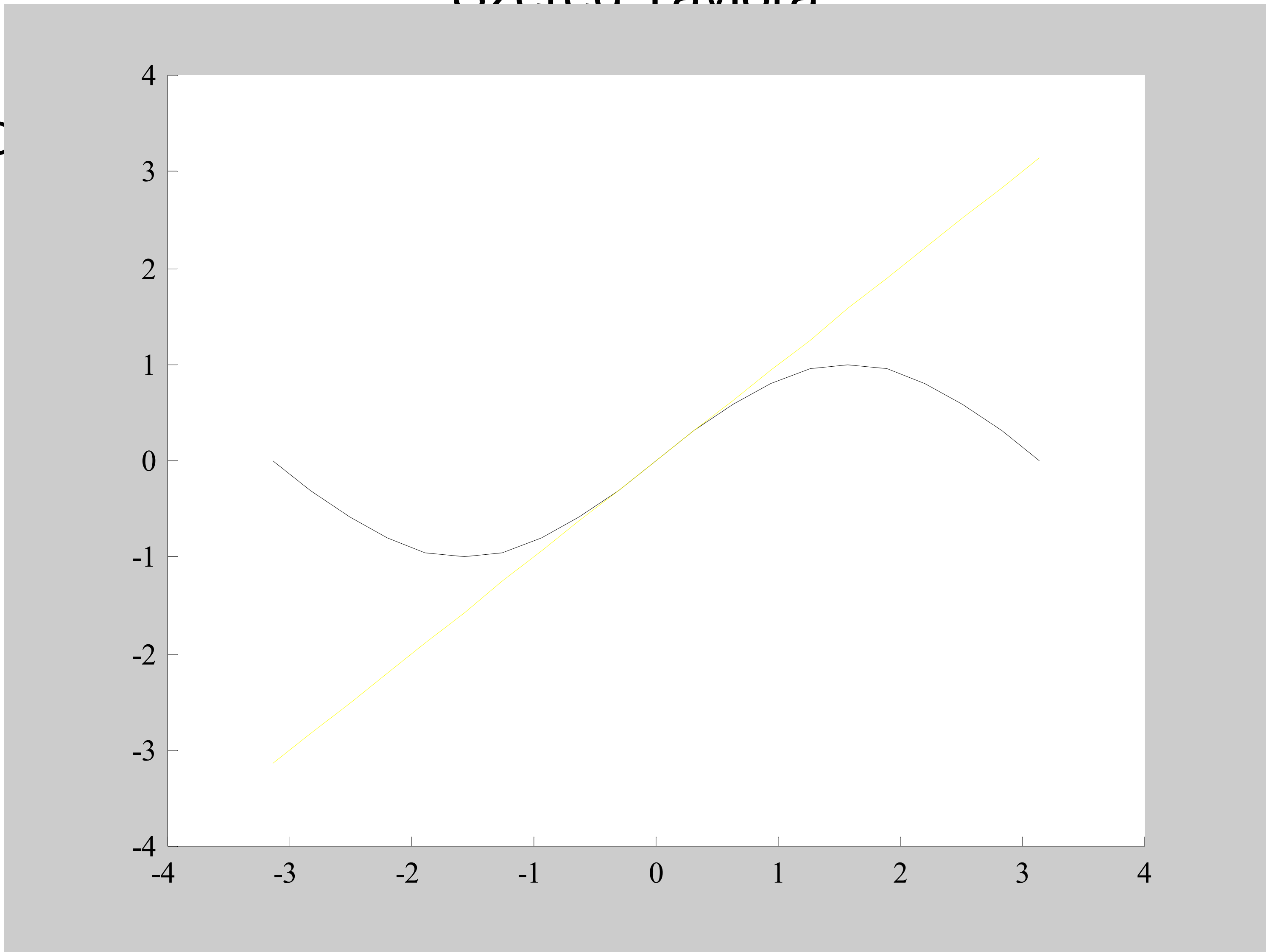
Szereg Taylora

• C



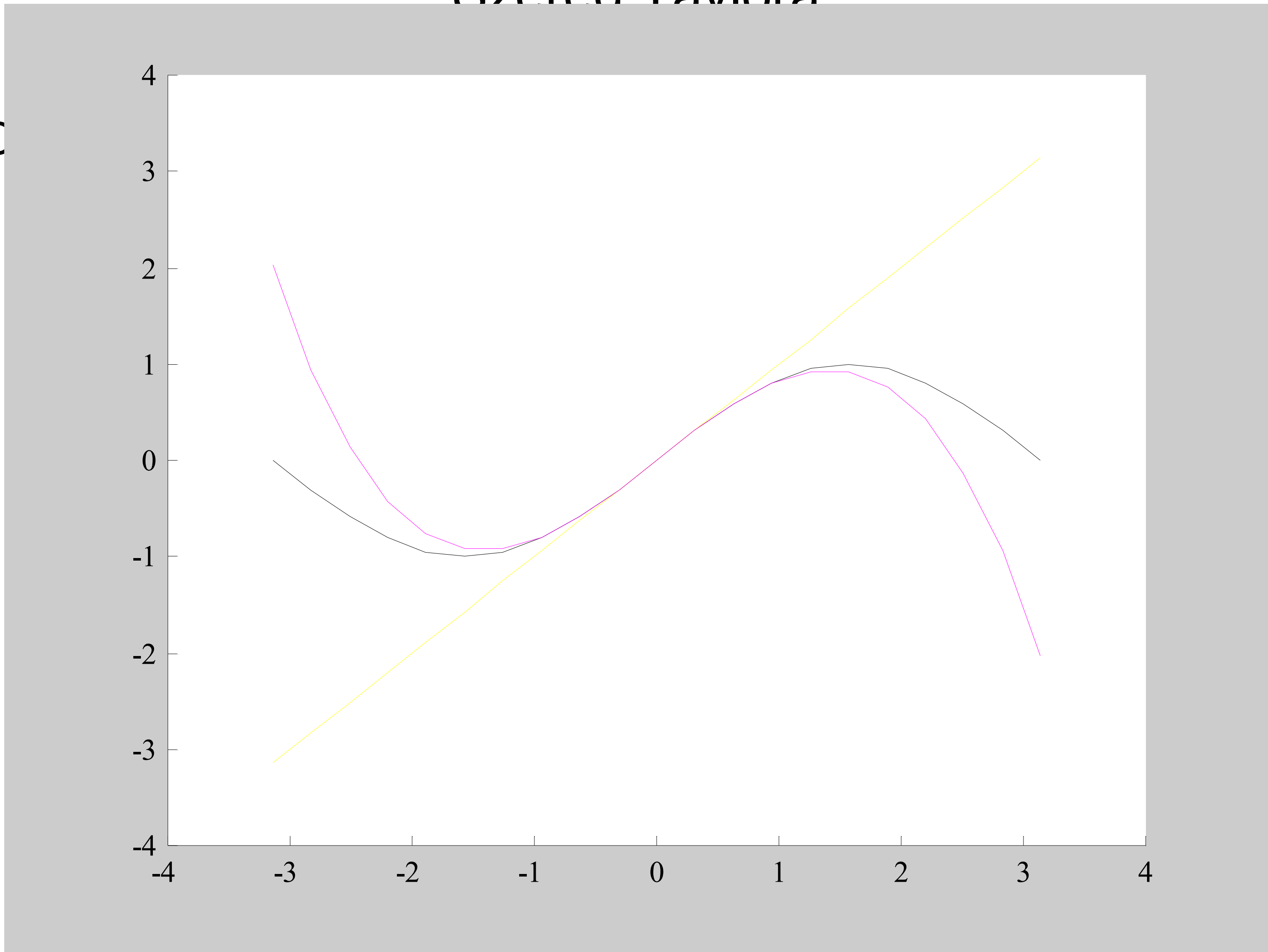
Szereg Taylora

• C



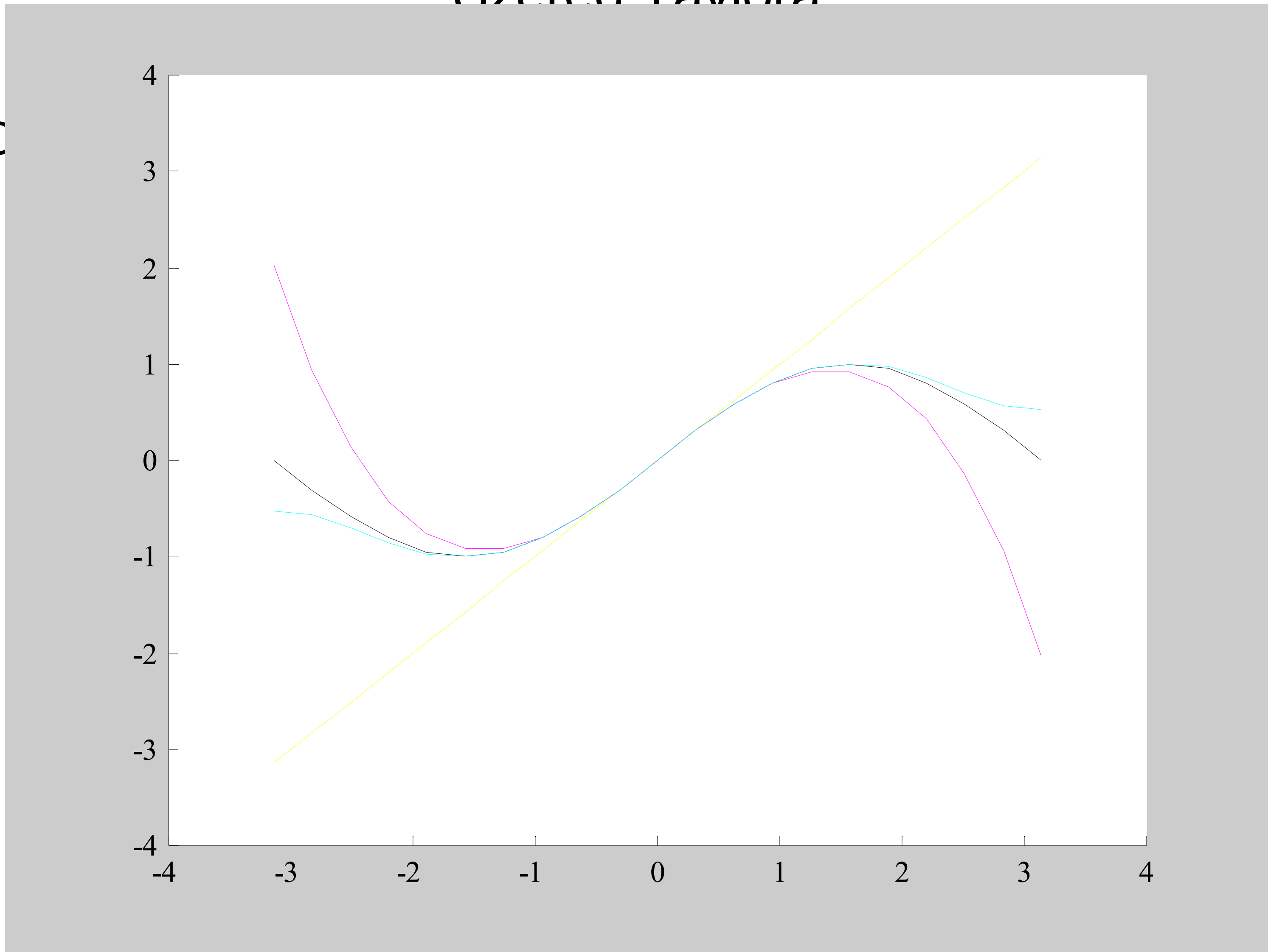
Szereg Taylora

- C



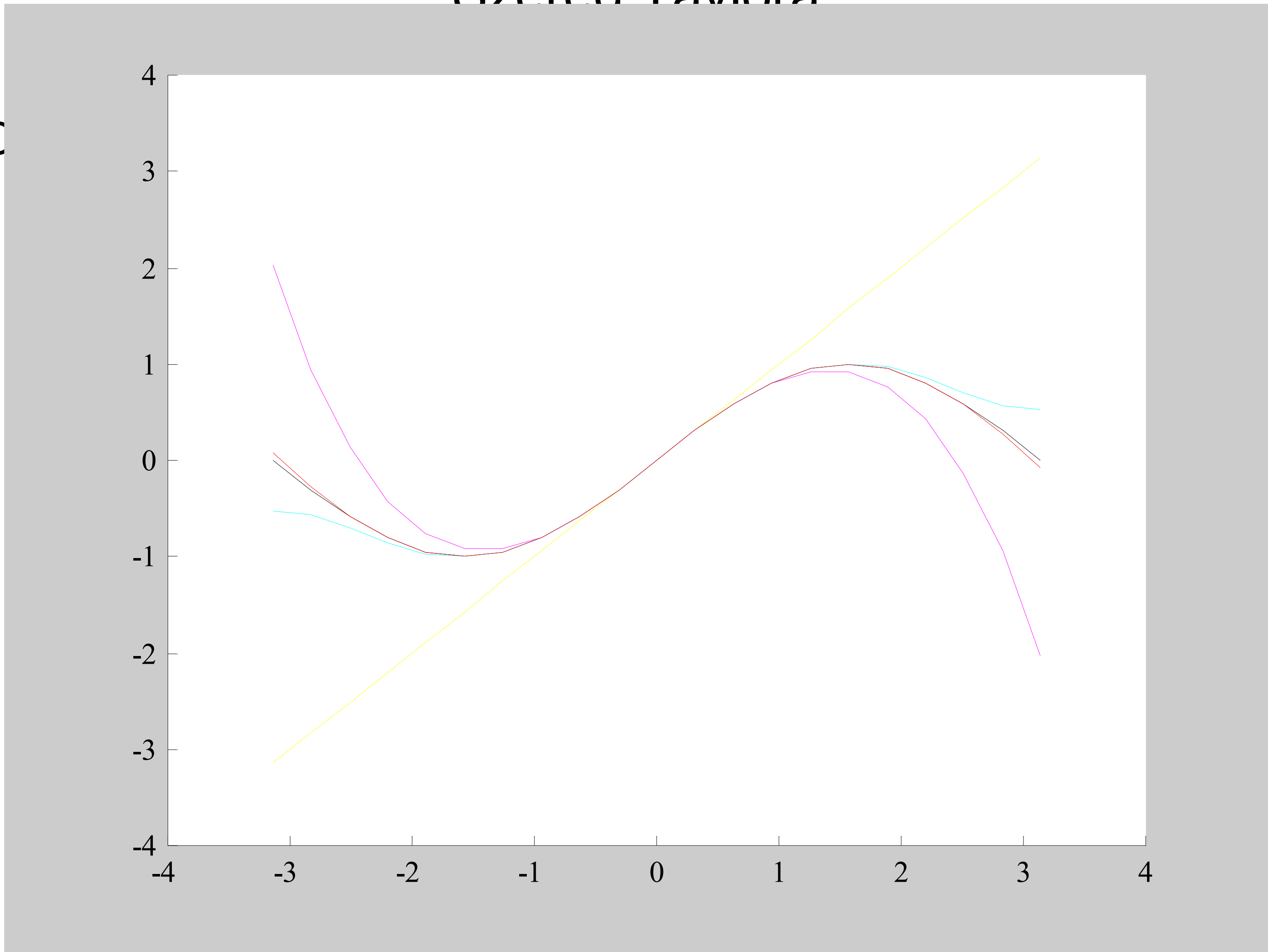
Szereg Taylora

- C



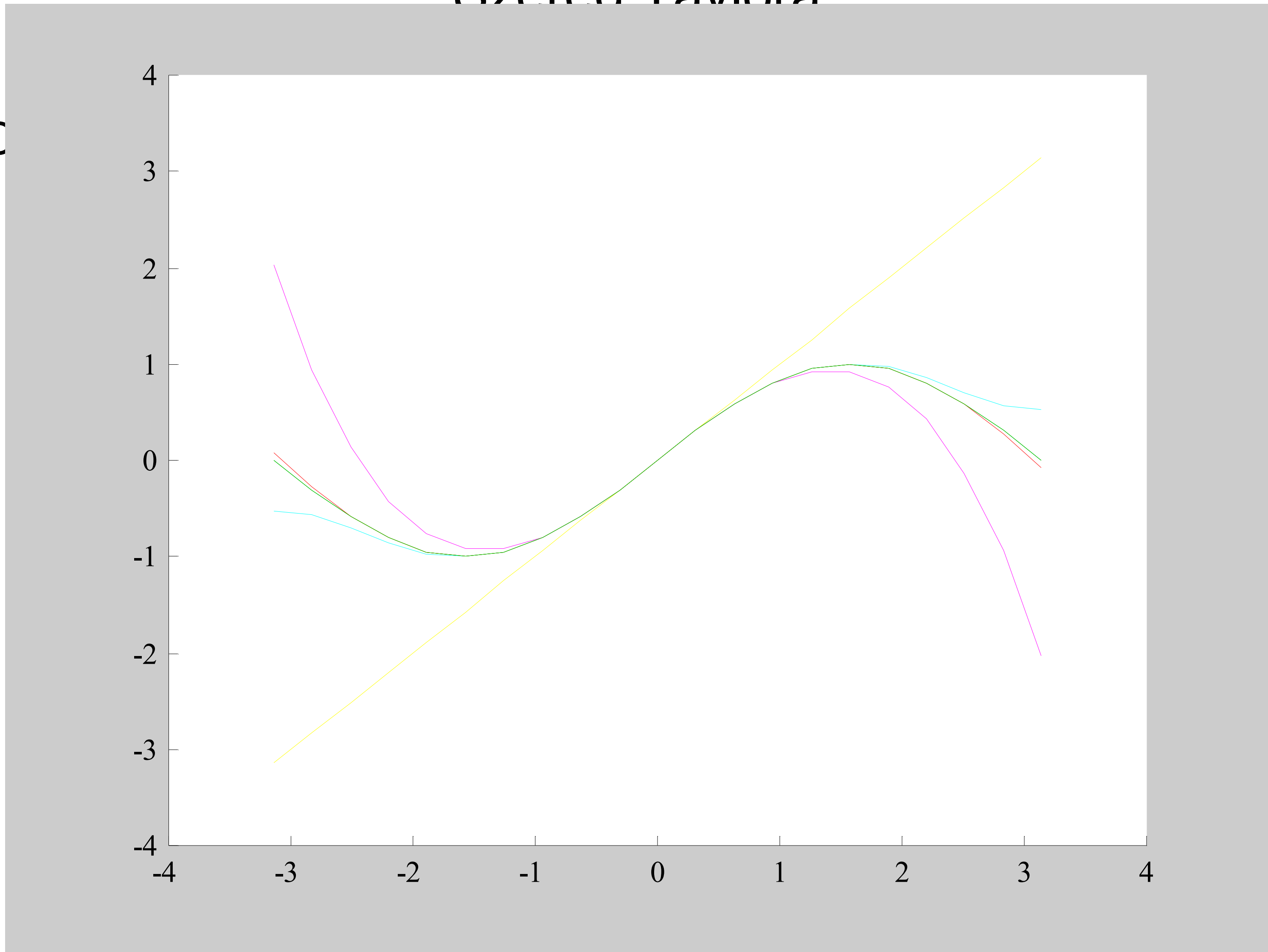
Szereg Taylora

- C



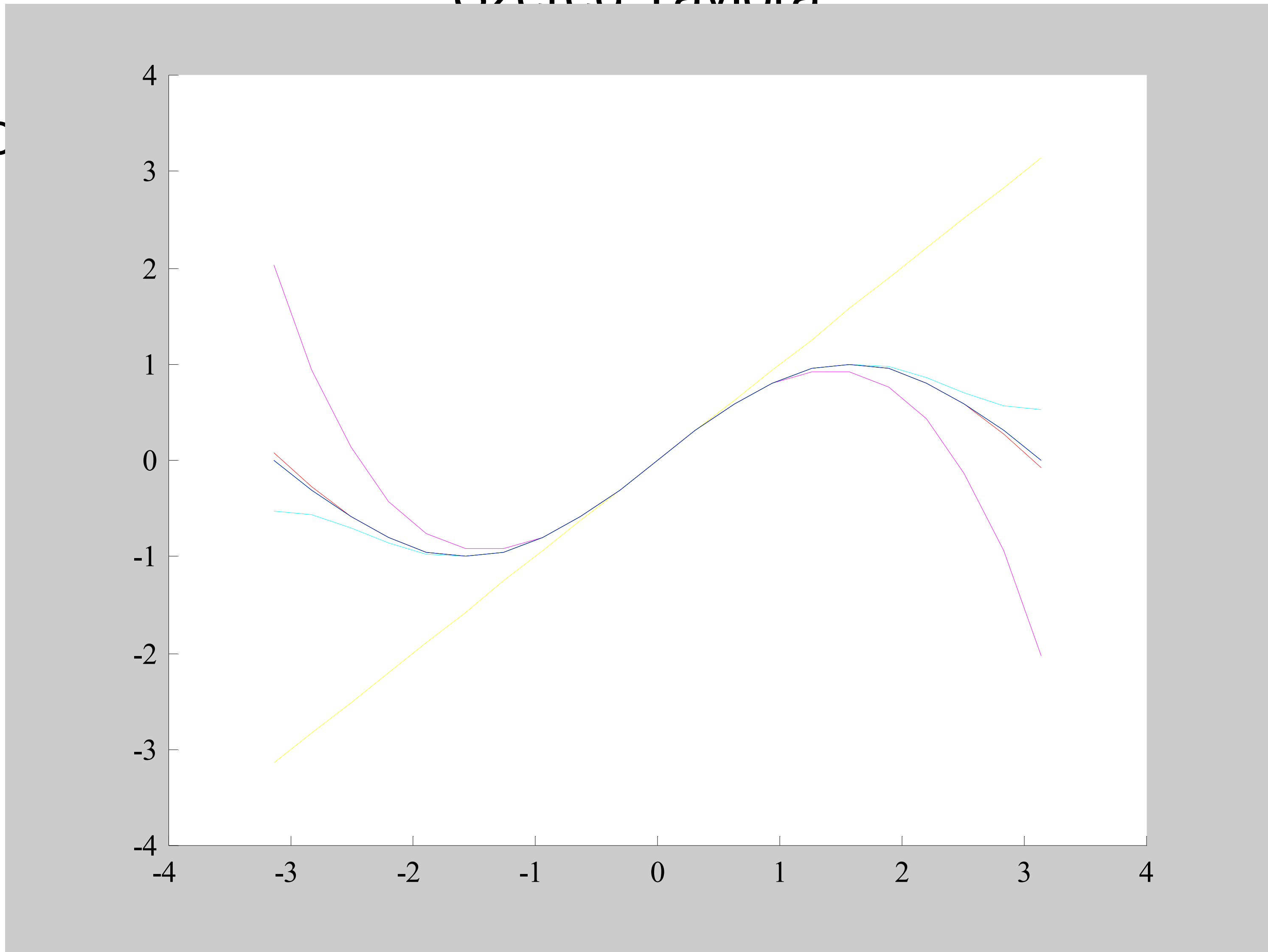
Szereg Taylora

• C



Szereg Taylora

- C



Szereg Taylora

- Jeszcze inny przykład: dwunastoelementowe $T_{11}(x)$ rozwinięcie funkcji cos(x) w szereg Taylora wokół wartości $x_0 = 0$

$\cos(x) \approx T(x) = \sum_{k=0..\infty} f^{(k)}(x_0)/k! \cdot (x - x_0)^k =$

$= \ldots$

Szereg Taylora

- Przybliżanie funkcji f(x) szeregiem Taylora
  - niech dane będą
    - ustalony obszar S
    - funkcja f(x) określona w obszarze S i posiadająca wszystkie pochodne określone w obszarze S
  - rozwinięcie T(x) funkcji f(x) w szereg Taylora wokół punktu $x_0 \in S$ dane jest następującym wzorem

$$T(x) = f^{(0)}(x_0)/(0!)\cdot(x-x_0)^0 + f^{(1)}(x_0)/(1!)\cdot(x-x_0)^1 + f^{(2)}(x_0)/(2!)\cdot(x-x_0)^2 + \ldots$$

  - uwaga: rozwinięcie może obejmować nieskończoną lub skończoną liczbę (niezerowych) składników (w przypadku liczby skończonej ostatni element szeregu jest innej postaci /i stanowi tzw. resztę/)

Szereg Taylora

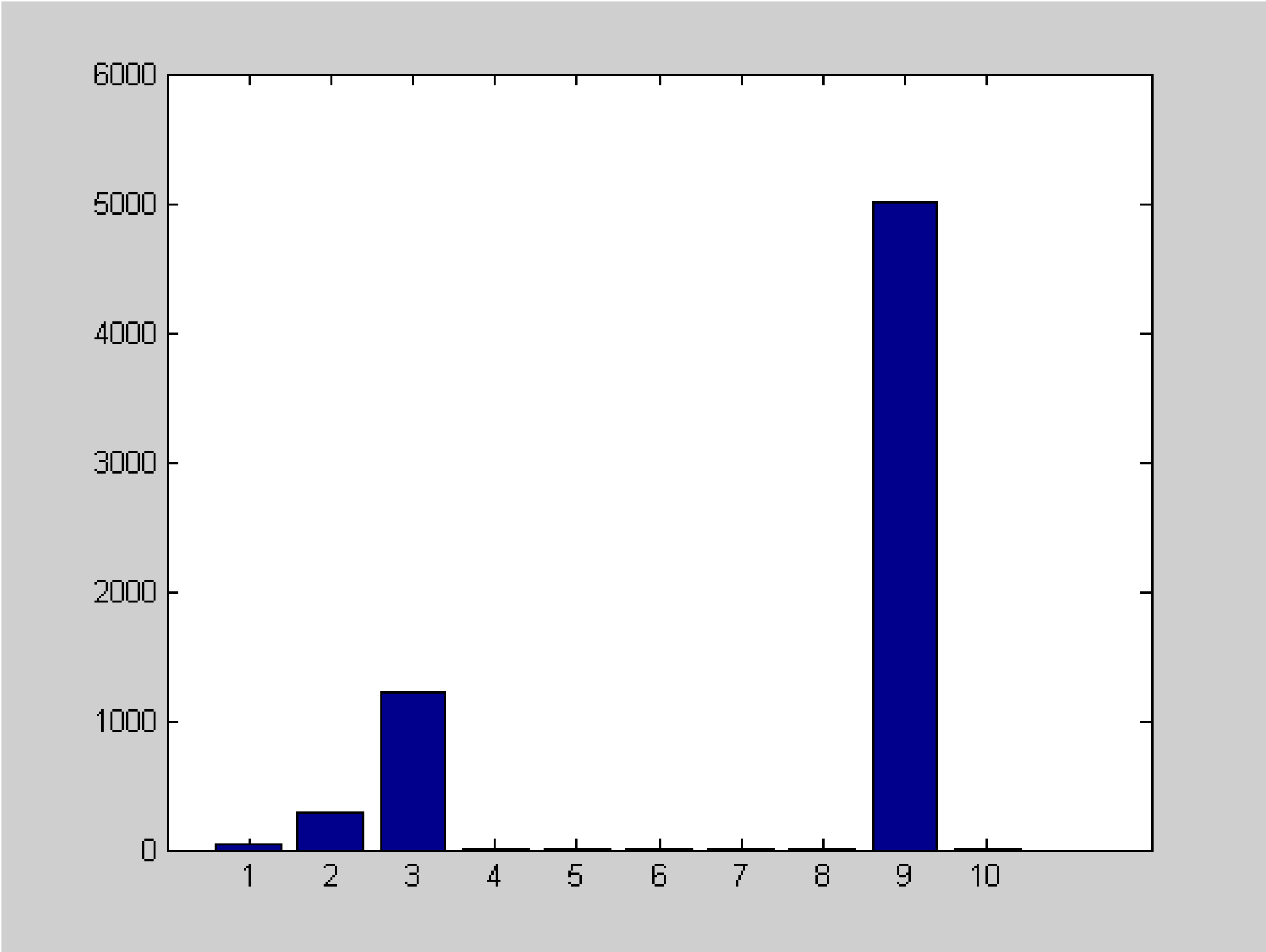
- Przybliżanie funkcji f(x) szeregiem Taylora, c.d.
  - aby otrzymać dokładniejsze przybliżenie wartości funkcji w punktach odległych od (zastosowanego) x_0 należy:
    - stosować wieloelementowe przybliżenia
    - użyć nowego x_0 do utworzenia nowego rozwinięcia

…

Dygresja

50,8944264922
295,2732174260
1219,6440273420
0,0000049136
0,8334066578
1,0792168463
0,0000042009
0,0000598881
5015,8814043346
0,0268475121

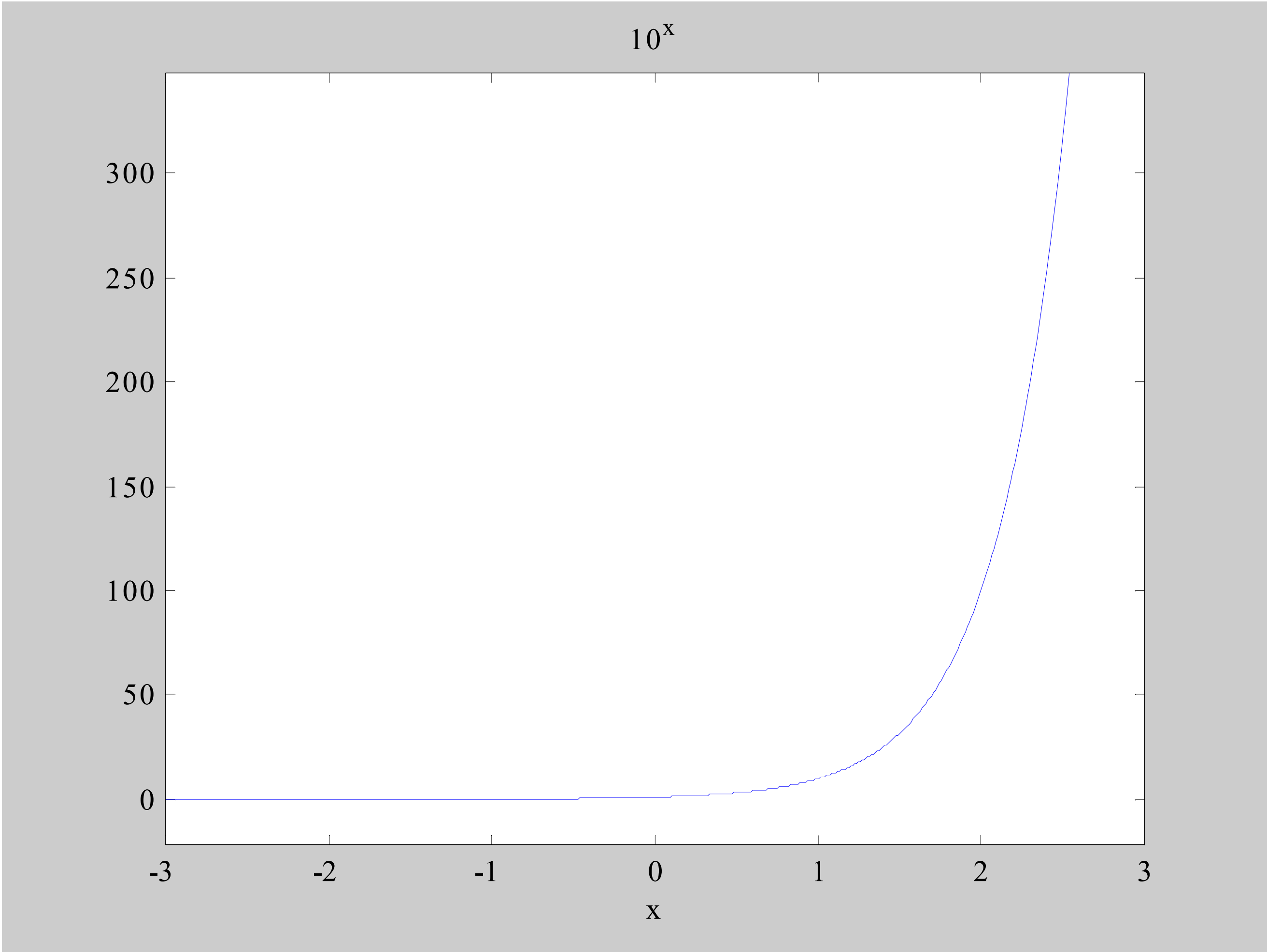
Dygresja



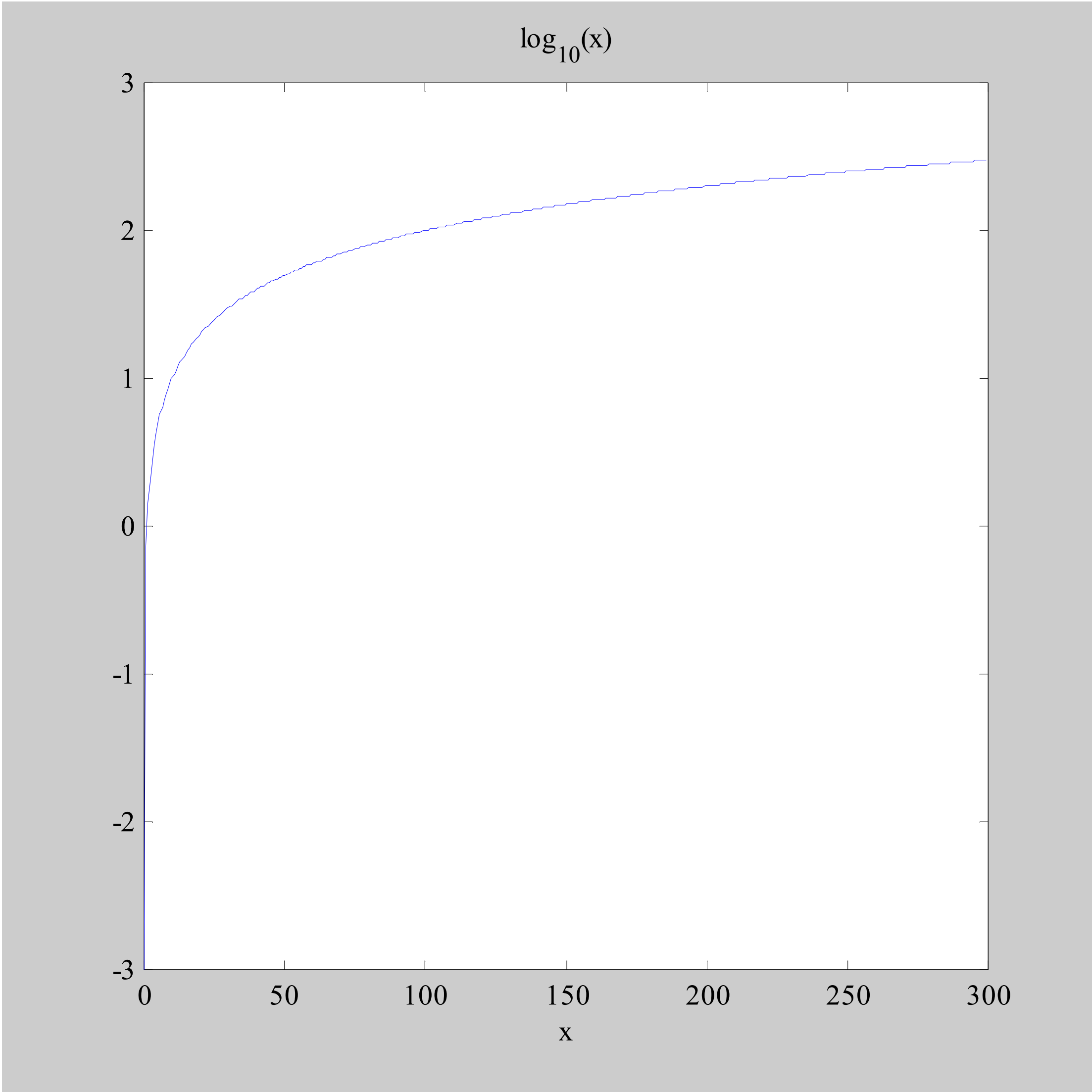
Dygresja

50,8944264922	50,894	50,89	50,9	...
295,2732174260	295,273	295,27	295,3	...
1219,6440273420	1219,644	1219,64	1219,6	...
0,0000049136	0,000	0,00	0,0	...
0,8334066578	0,833	0,83	0,8	...
1,0792168463	1,079	1,08	1,1	...
0,0000042009	0,000	0,00	0,0	...
0,0000598881	0,000	0,00	0,0	...
5015,8814043346	5015,881	5015,88	5015,9	...
0,0268475121	0,027	0,03	0,0	...

Dygresja



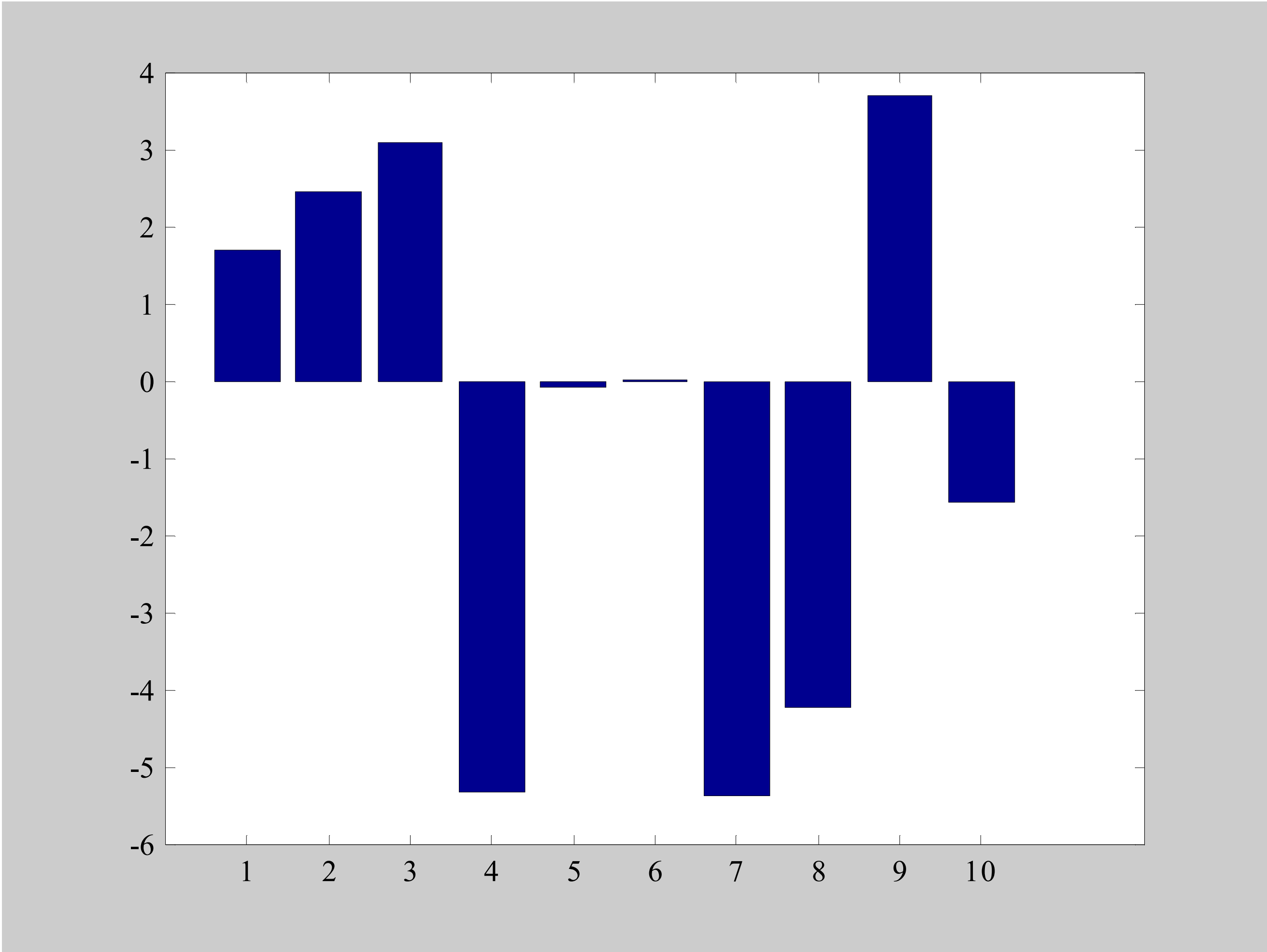
Dygresja



Dygresja

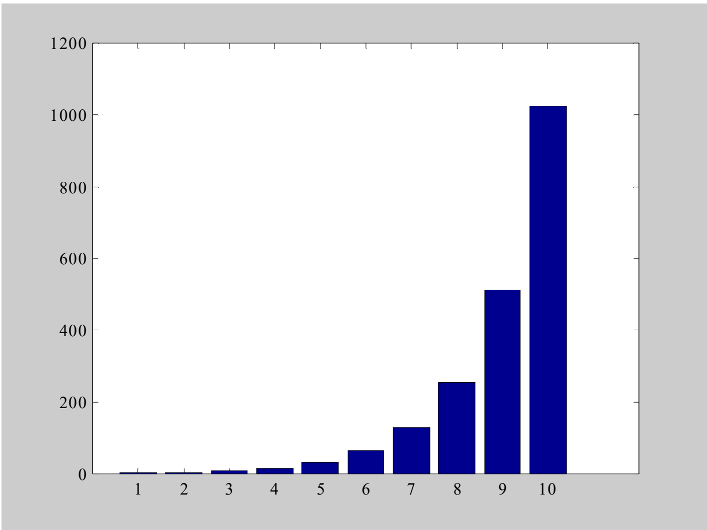
50,8944264922	1,707
295,2732174260	2,470
1219,6440273420	3,086
0,0000049136	-5,309
0,8334066578	-0,079
1,0792168463	0,033
0,0000042009	-5,377
0,0000598881	-4,223
5015,8814043346	3,700
0,0268475121	-1,571

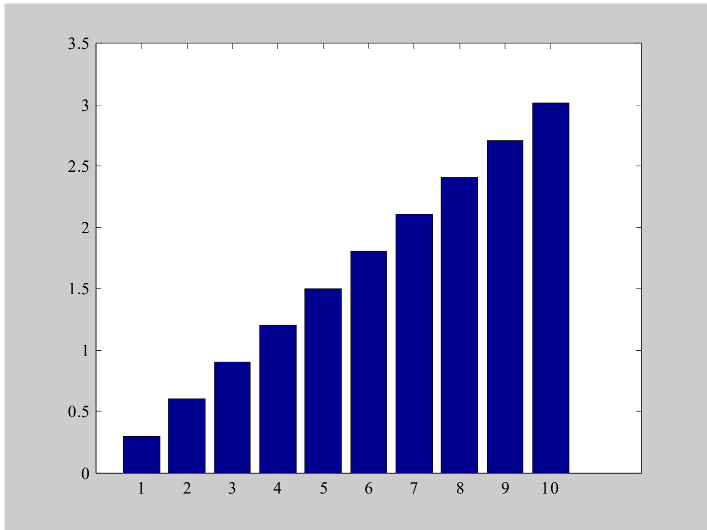
Dygresja



Dygresja

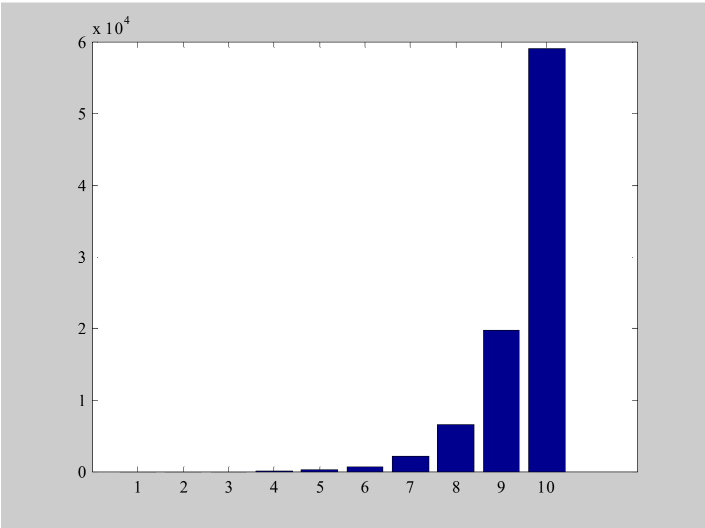
2	0,301029996
4	0,602059991
8	0,903089987
16	1,204119983
32	1,505149978
64	1,806179974
128	2,10720997
256	2,408239965
512	2,709269961
1024	3,010299957

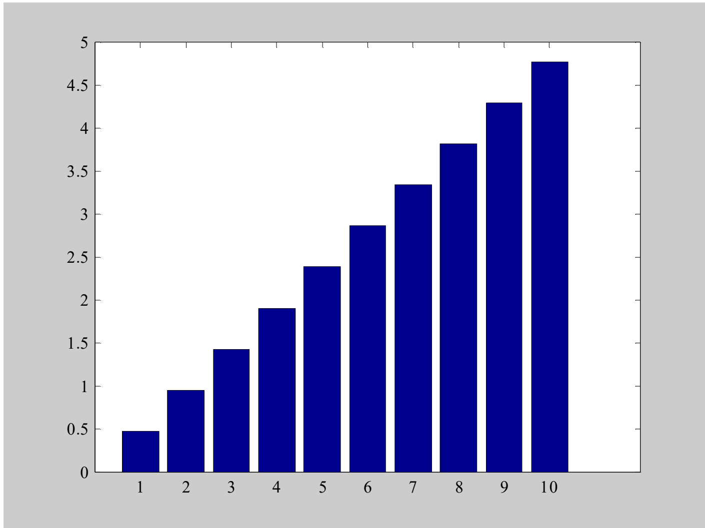




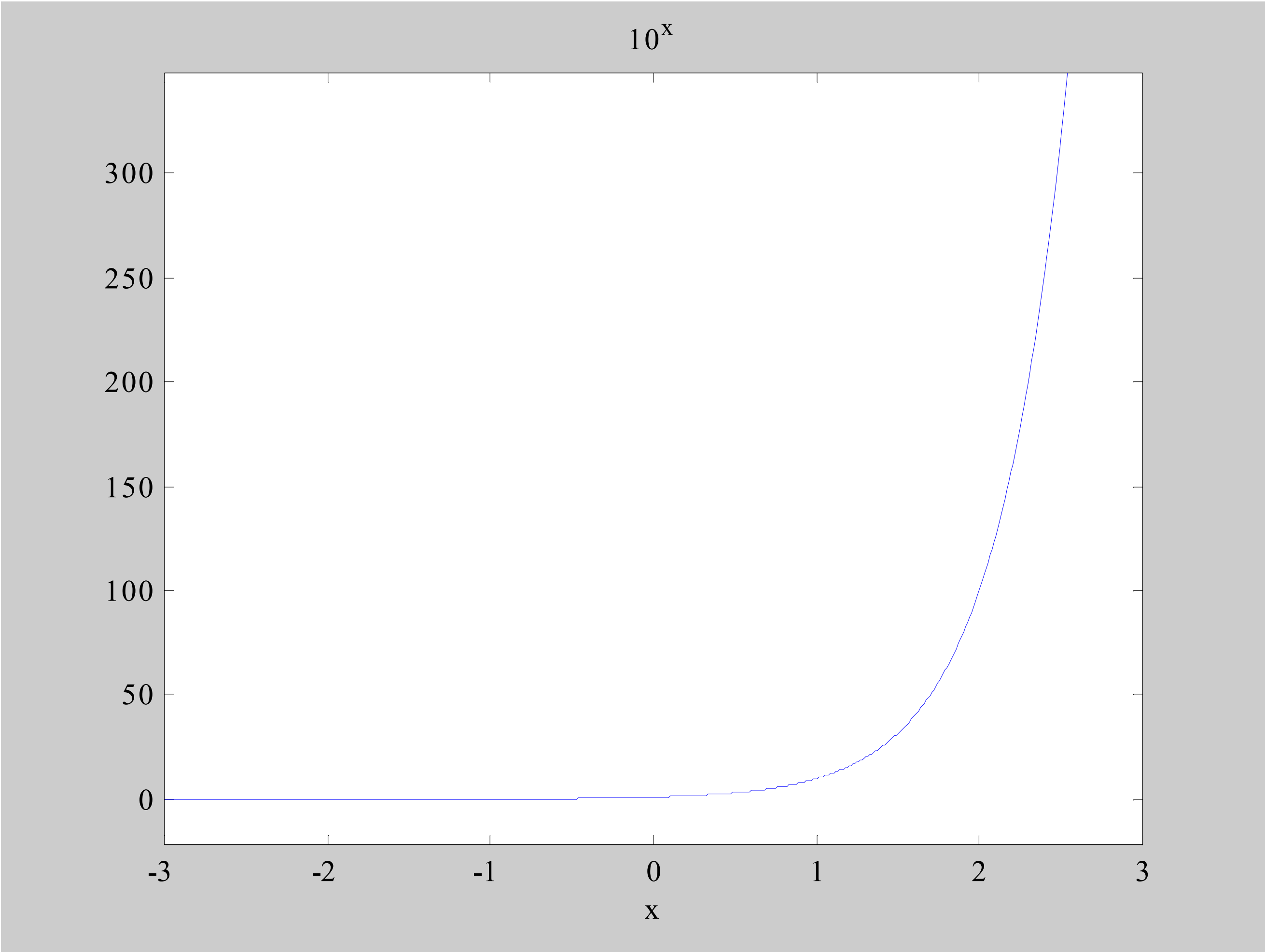
Dygresja

3	0,477121255
9	0,954242509
27	1,431363764
81	1,908485019
243	2,385606274
729	2,862727528
2187	3,339848783
6561	3,816970038
19683	4,294091292
59049	4,771212547

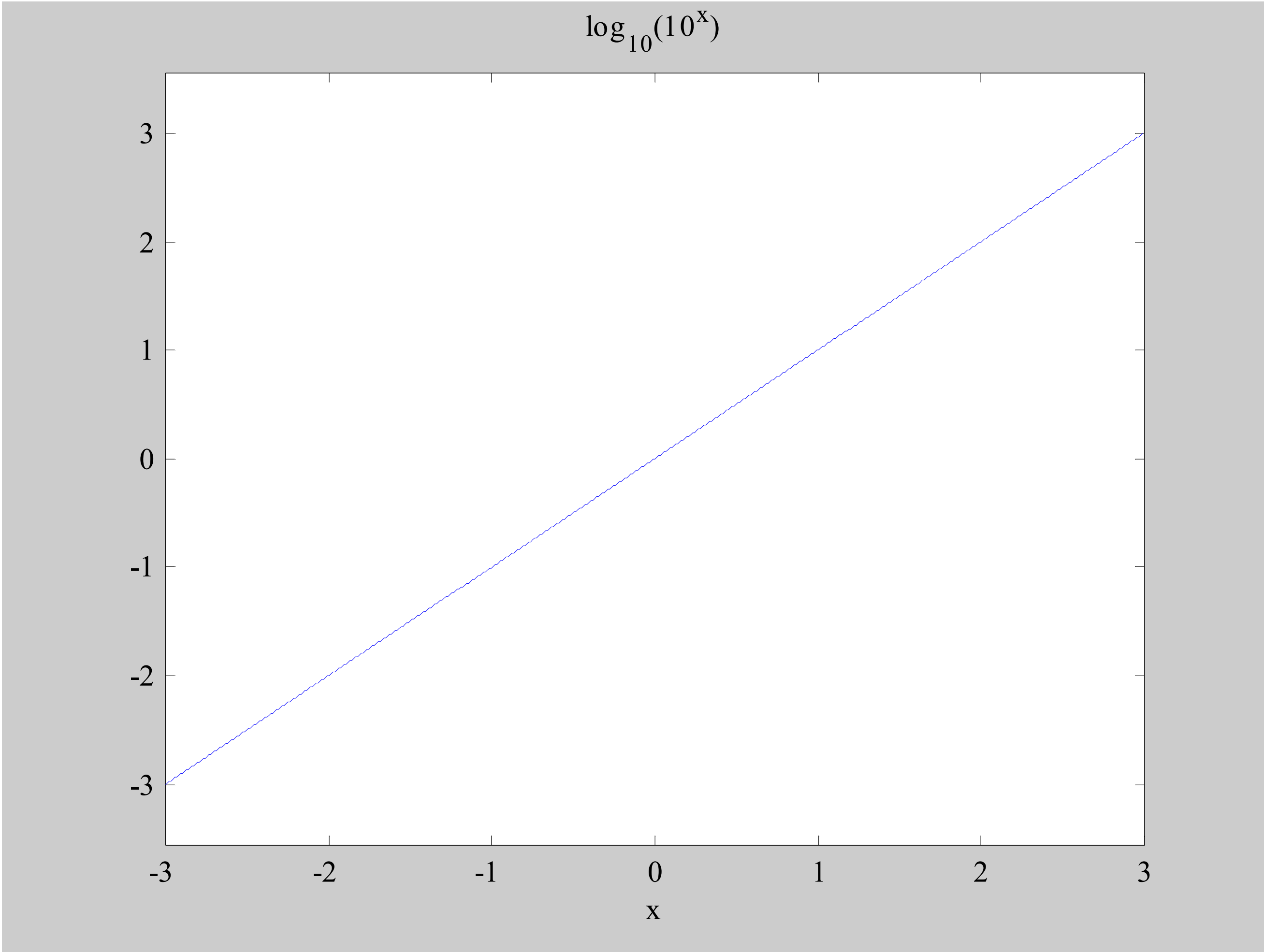




Dygresja



Dygresja



…

Współczynnik i rząd zbieżności ciągu

- Idea współczynnika i rzędu zbieżności ciągu <u>skalarów</u>

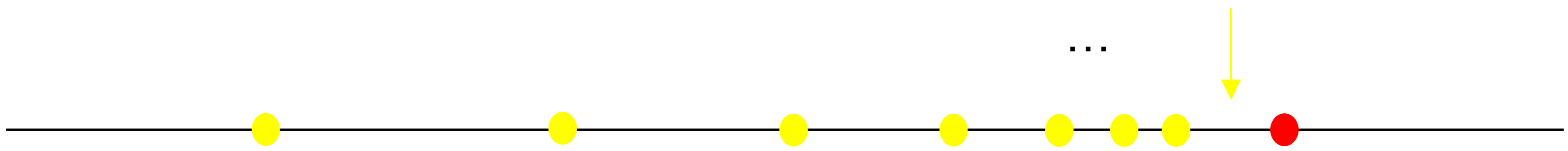
Współczynnik i rząd zbieżności ciągu

- Idea współczynnika i rzędu zbieżności ciągu <u>skalarów</u>
  - niech s_0, s_1, s_2, … będzie ciągiem skalarów zbieżnym do skalara $s = \lim_{k\to\infty} s_k$
  - niech $p \geq 1$ będzie maksymalną wartością, dla której istnieje granica

    $\beta = \lim_{k\to\infty}|s_{k+1} - s|/|s_k - s|^p$
  - wtedy
    - wartość p nazywamy rzędem zbieżności
    - wartość β nazywamy współczynnikiem zbieżności p-tego rzędu
  - jeżeli
    - $p = 1$ i $\beta \in (0,1)$, to ciąg ma zbieżność liniową
    - $p = 1$ i $\beta = 0$ lub $p > 1$, to ciąg ma zbieżność superliniową (w znaczeniu: lepszą od liniowej)

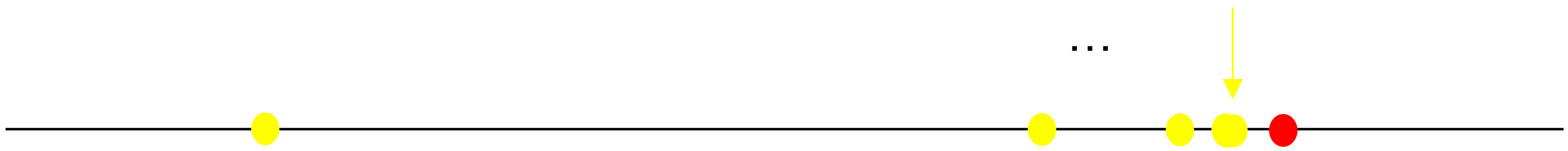
Współczynnik i rząd zbieżności ciągu

- Ilustracja (rząd pierwszy)



Współczynnik i rząd zbieżności ciągu

- Ilustracja (rząd drugi)



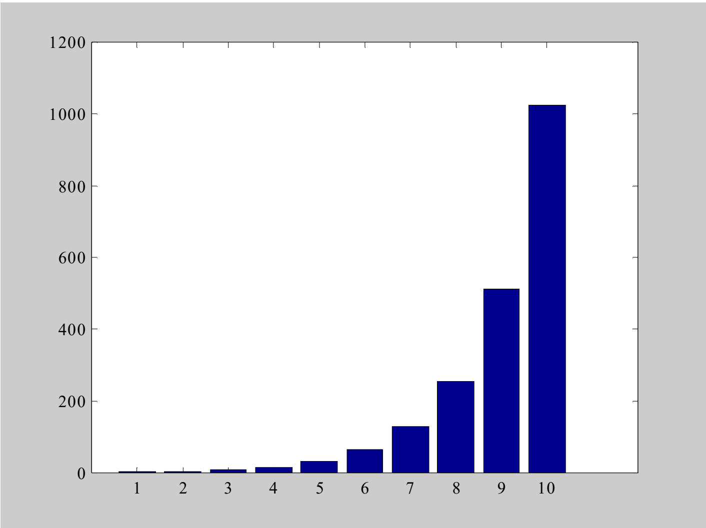
Współczynnik i rząd zbieżności ciągu

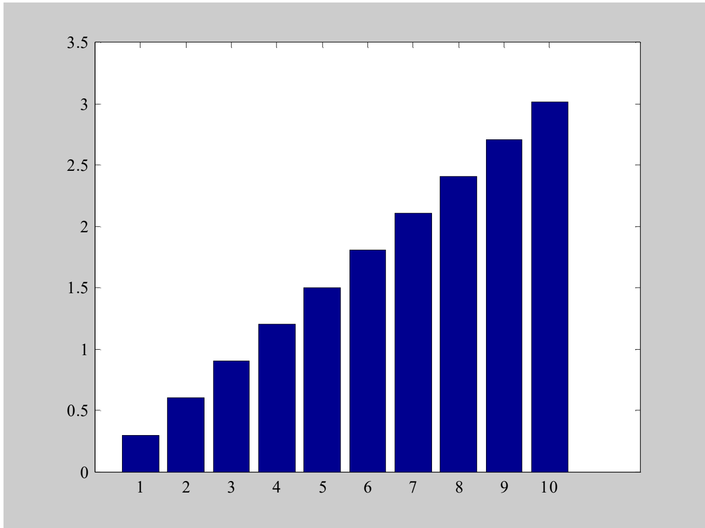
- Idea współczynnika...
  - po zastosowaniu oznaczenia $e_k = |s_{k+1} - s|$ (error), dla dużych n zachodzi $e_{k+1} = \beta(e_k)^p$

Współczynnik i rząd zbieżności ciągu

- Idea...
  - ciąg $e_{k+1} = \beta(e_k)^p$ dla
    $e_0 = 1$
    $\beta = 2$
    $p = 1$

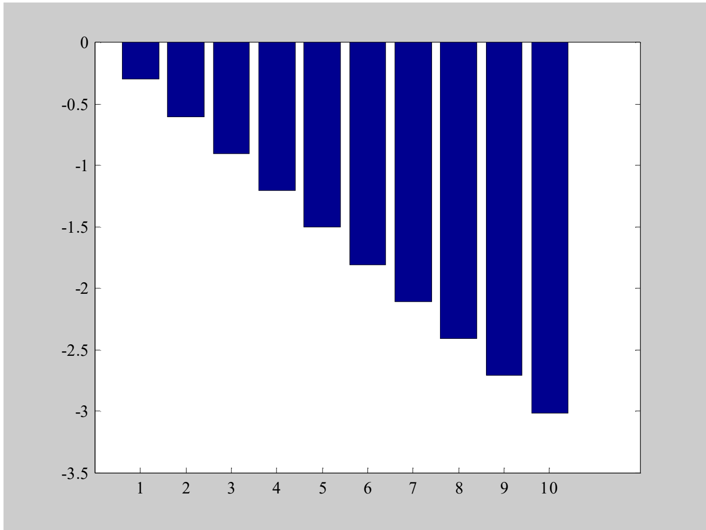
2	0,301029996
4	0,602059991
8	0,903089987
16	1,204119983
32	1,505149978
64	1,806179974
128	2,10720997
256	2,408239965
512	2,709269961
1024	3,010299957





Współczynnik i rząd zbieżności ciągu

- Idea...
  - ciąg $e_{k+1} = \beta(e_k)^p$
    dla
    $e_0 = 1$
    $\beta = 1/2$
    $p = 1$



Współczynnik i rząd zbieżności ciągu

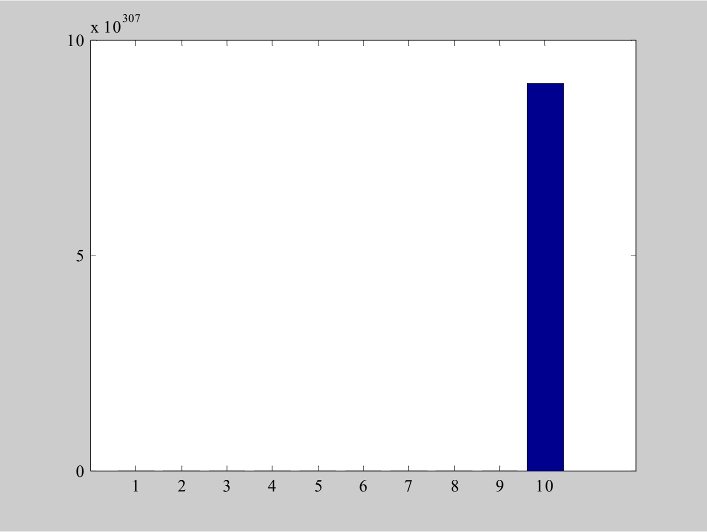
- Idea...
  - ciąg $e_{k+1} = \beta(e_k)^p$
    dla
    $e_0 = 1$
    $\beta = 2$
    $p = 2$

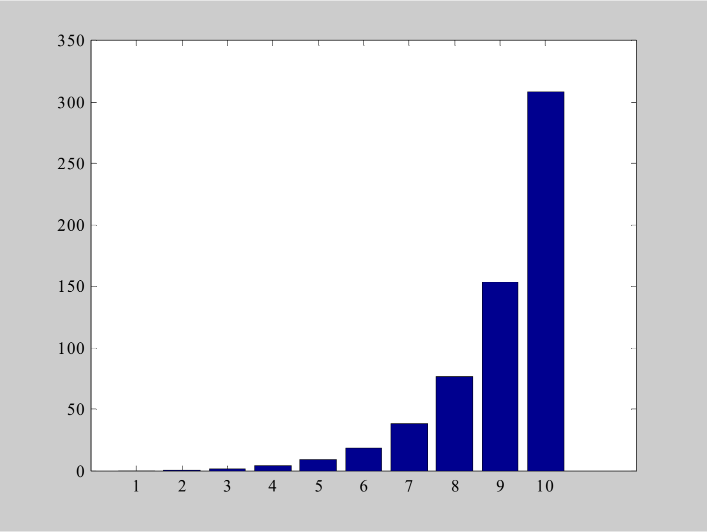
2.00
8.00
128.00
32768.00
2147483648.00
9223372036854775800.00
170141183460469230000000000000000000000.00
57896044618658098000000000000000000000000000000000000000000000000000000000000.00
6703903964971298000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000.0
000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000

Współczynnik i rząd zbieżności ciągu

- Idea...
  - ciąg $e_{k+1} = \beta(e_k)^p$ dla
    $e_0 = 1$
    $\beta = 2$
    $p = 2$

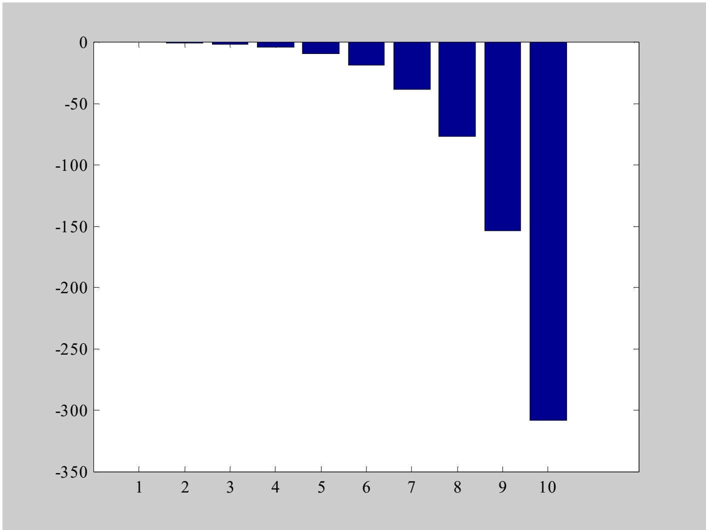
2	0,301029996
8	0,903089987
128	2,10720997
32768	4,515449935
2147483648	9,331929866
9,22E+18	18,96488973
1,70E+38	38,23080945
5,79E+76	76,76264889
6,70E+153	153,8263278
8,99E+307	307,9536856





Współczynnik i rząd zbieżności ciągu

- Idea...
  - ciąg $e_{k+1} = \beta(e_k)^p$
    dla
    $e_0 = 1$
    $\beta = 1/2$
    $p = 2$



Współczynnik i rząd zbieżności ciągu

- (Przykładowe) porównanie:

ciąg $e_{k+1} = \beta(e_k)^p$
dla
$e_0 = 1$
$\beta = 2$
$p = 1$
czyli
$e_{k+1} = 2(e_k)^1 = 2e_k$
wtedy
$\log(e_{k+1}) = \log(2e_k)$
$\quad = \log(2)+\log(e_k)$

ciąg $e_{k+1} = \beta(e_k)^p$
dla
$e_0 = 1$
$\beta = 2$
$p = 2$
czyli
$e_{k+1} = 2(e_k)^2$
wtedy
$\log(e_{k+1}) = \log(2(e_k)^2)$
$\quad = 2(\log(2)+\log(e_k))$

- W ogólności

$\log(e_{k+1}) = \log(\beta(e_k)^p) = p\cdot\log(\beta e_k) = p\cdot(\log(\beta)+\log(e_k))$

  - generuje przyrost wykładniczy dla $1 \neq p > 0$

Współczynnik i rząd zbieżności ciągu

- Idea współczynnika i rzędu zbieżności ciągu <u>wektorów</u>

Współczynnik i rząd zbieżności ciągu

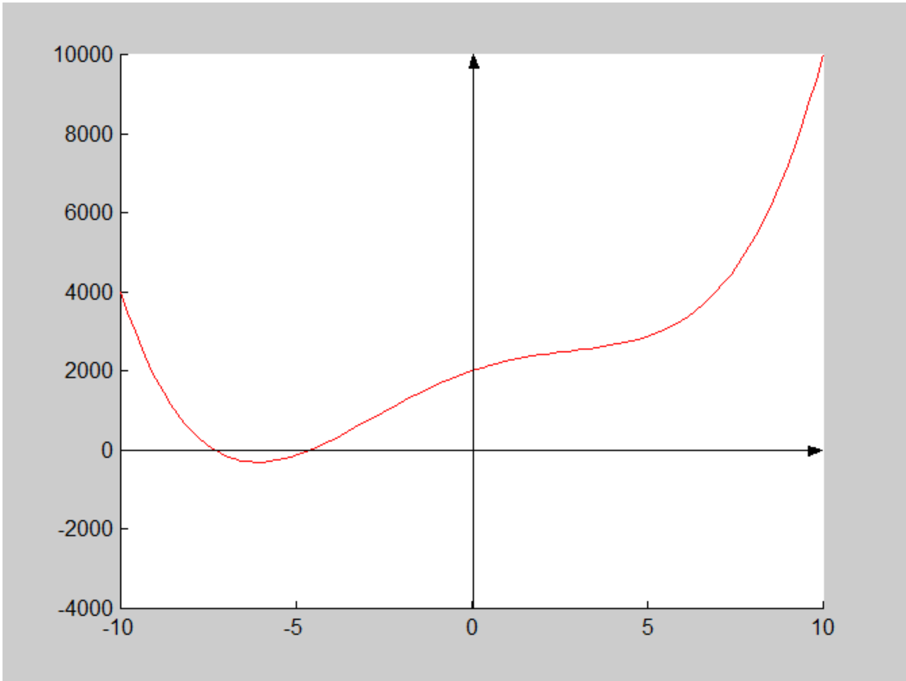
- Idea współczynnika i rzędu zbieżności ciągu <u>wektorów</u>
  - niech $\mathbf{w}_0$, $\mathbf{w}_1$, $\mathbf{w}_2$, … będzie ciągiem wektorów zbieżnym do wektora $\mathbf{w} = \lim_{k\to\infty}\mathbf{w}_k$
  - niech $p \geq 1$ będzie maksymalną wartością, dla której istnieje granica

    $\beta = \lim_{k\to\infty}||\mathbf{w}_{k+1} - \mathbf{w}||/||\mathbf{w}_k - \mathbf{w}||^p$,
  - wtedy
    - wartość p nazywamy rzędem zbieżności
    - wartość β nazywamy współczynnikiem zbieżności p-tego rzędu
  - jeżeli
    - $p = 1$ i $\beta \in (0,1)$, to ciąg ma zbieżność liniową
    - $p = 1$ i $\beta = 0$ lub $p > 1$, to ciąg ma zbieżność superliniową (w znaczeniu: lepszą od liniowej)

…

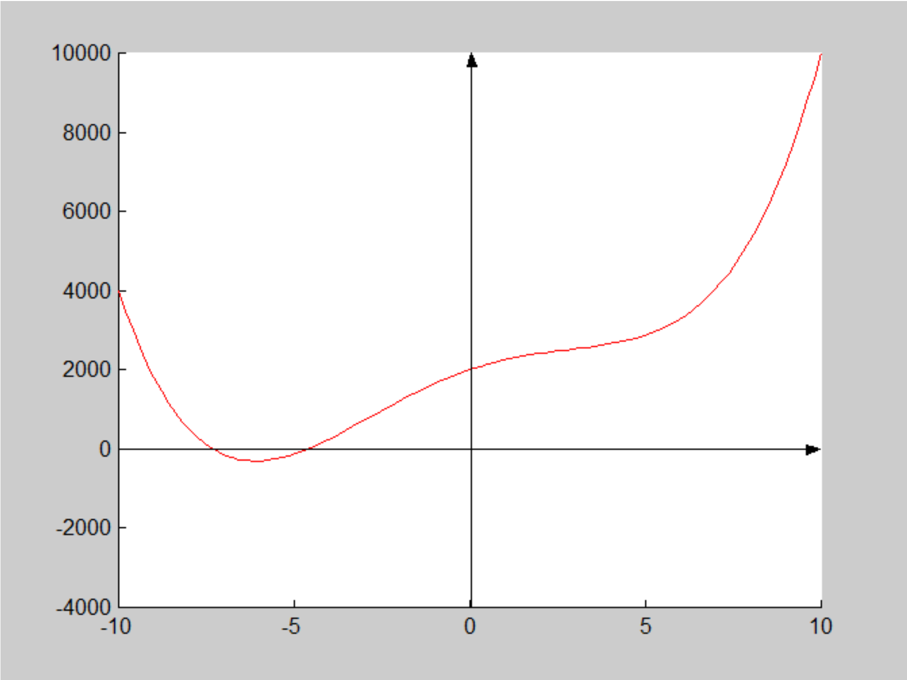
Aproksymacja i optymalizacja: ilustracja problemów

- Przykład funkcji: $f(x) = x^4 - 50x^2 + 300x + 2000$



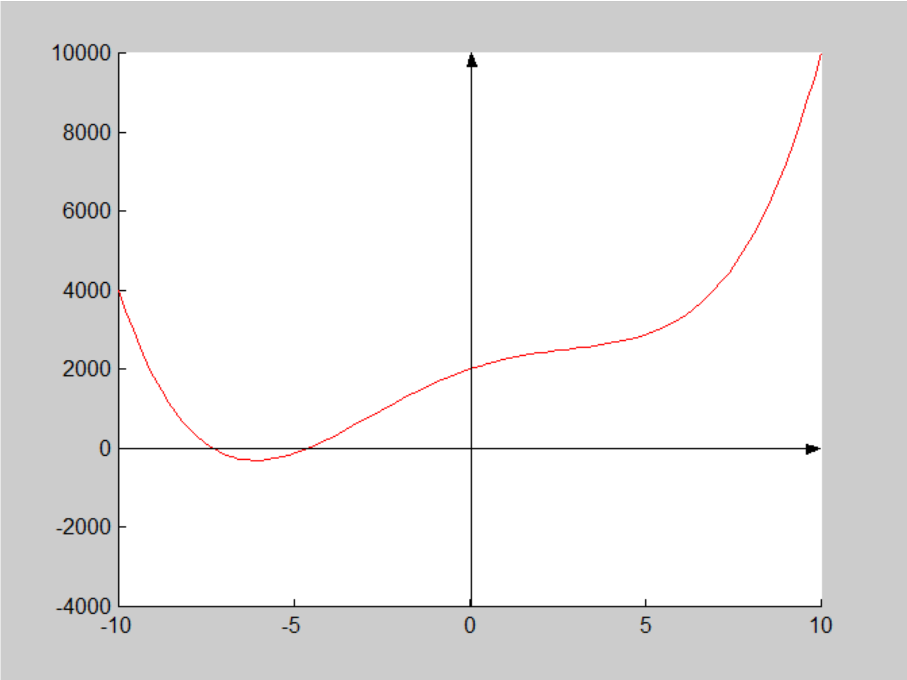
Aproksymacja i optymalizacja: ilustracja problemów

- Aproksymacja
  - problem istnienia rozwiązań (miejsc zerowych)
    - brak miejsc zerowych
    - asymptotyczne zbliżanie
    - miejsca zerowe poza granicami przedziału zmienności
    - ...



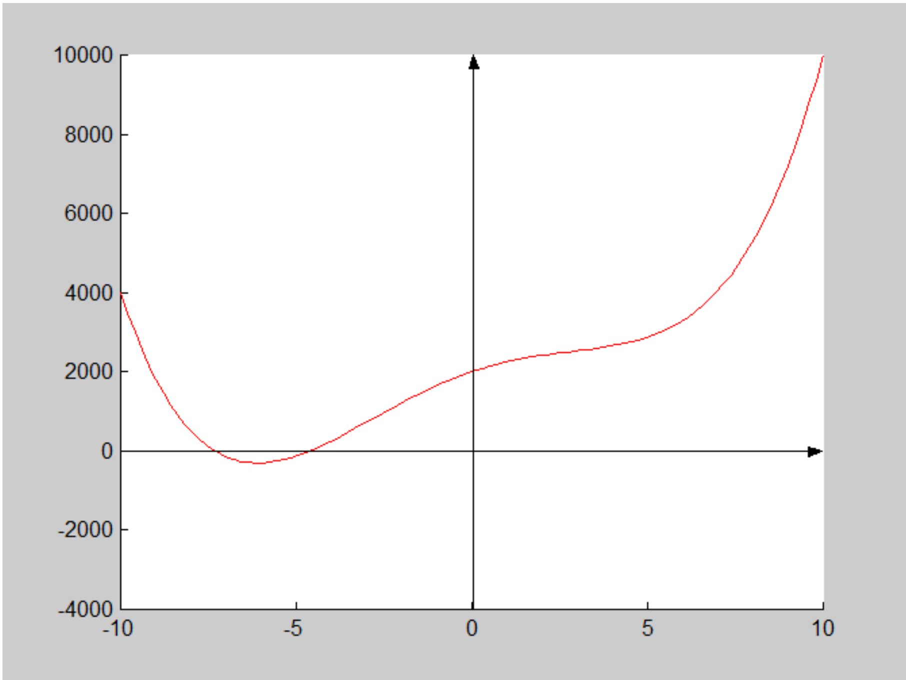
Aproksymacja i optymalizacja: ilustracja problemów

- Aproksymacja
  - problem jednoznaczności rozwiązań (miejsc zerowych)
    - policzalne liczby miejsc zerowych
    - niepoliczalne ilości miejsc zerowych
    - policzalne liczby niepoliczalnych ilości miejsc zerowych
    - ...



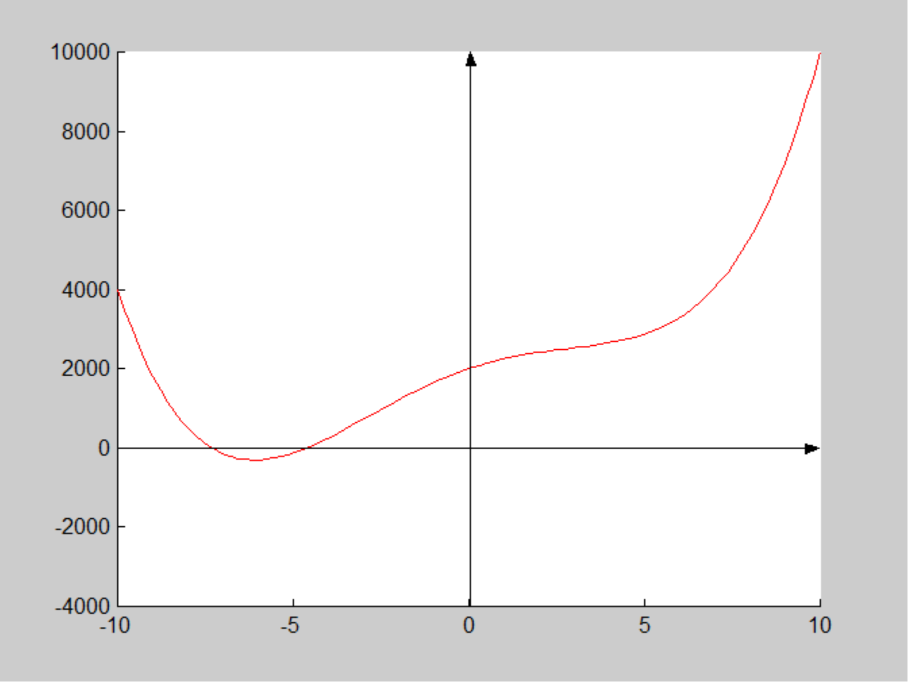
Aproksymacja i optymalizacja: ilustracja problemów

- Optymalizacja
  - problem istnienia rozwiązań
    - brak rozwiązań
    - asymptotyczne zbliżanie
    - rozwiązania poza granicami przedziału zmienności
    - ...



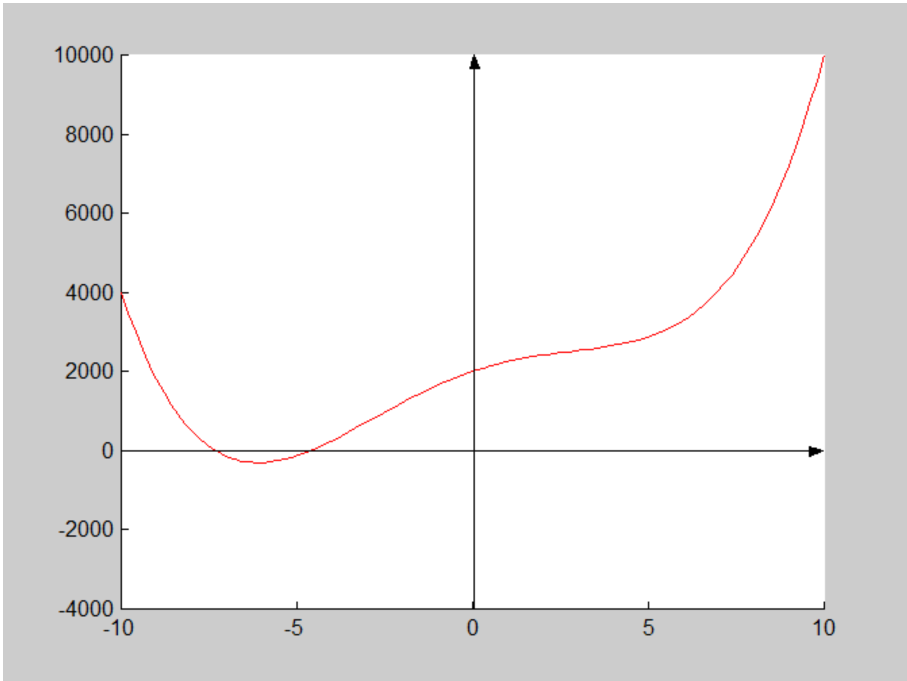
Aproksymacja i optymalizacja: ilustracja problemów

- Optymalizacja
  - problem jednoznaczności rozwiązań (minimów/maksimów)
    - policzalne liczby rozwiązań
    - niepoliczalne ilości rozwiązań
    - policzalne liczby niepoliczalnych ilości rozwiązań
    - ...



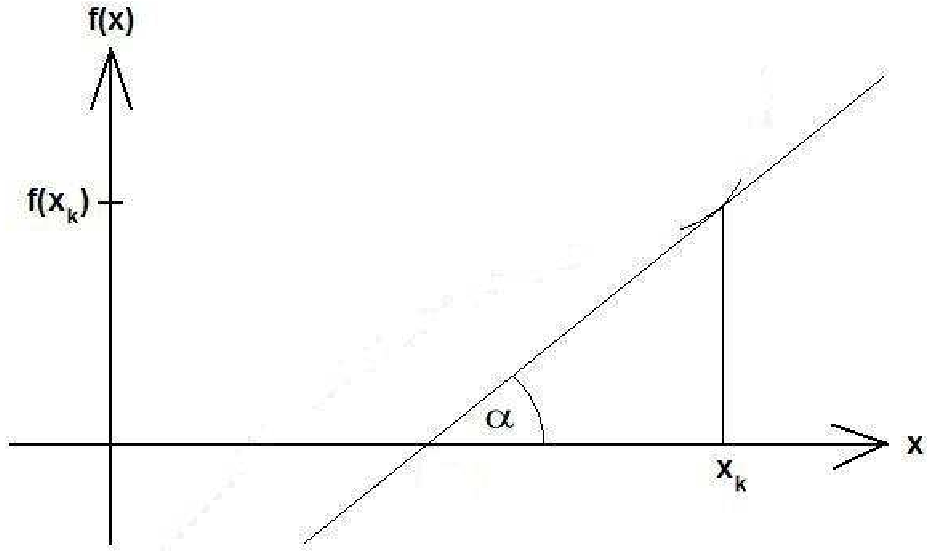
Aproksymacja i optymalizacja: ilustracja problemów

- Ograniczona ilość informacji w aproksymacji/optymalizacji



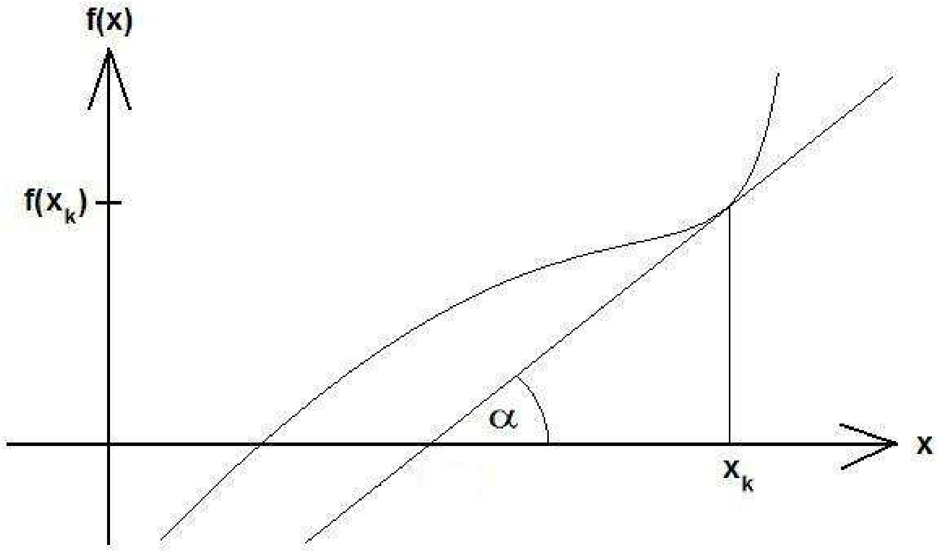
Aproksymacja i optymalizacja: ilustracja problemów

- Ograniczona ilość informacji w aproksymacji/optymalizacji
  - wartość funkcji i wartość jej (pierwszej) pochodnej



Aproksymacja i optymalizacja: ilustracja problemów

- Ograniczona ilość informacji w aproksymacji/optymalizacji
  - bez względu na ilość takich danych, to nie to samo, co informacja o całym „przebiegu” funkcji!



…