

Wszystkie udostępniane materiały, skrypty, notatniki są wyłącznie przeznaczone do użytku prywatnego w celu łatwiejszego opanowania wiedzy.

Nie wolno ich rozprzestrzeniać, ani umieszczać w Internecie.

Optymalizacja

Optymalizacja wielowymiarowa bez ograniczeń

Metoda najszybszego wzrostu/spadku (ang. *Steepest Ascent/Descent Method*)

Naturalną strategią wspinaczki na wzgórze wydaje się określenie maksymalnego nachylenia z punktu widzenia pozycji wyjściowej, a następnie pójście w tym kierunku.

Niestety skuteczność takiego podejścia zależy od wyboru punktu startowego. Jeżeli nie będziemy na stoku prowadzącym bezpośrednio do maksimum, bardzo szybko odejdziemy od najbardziej stromego kierunku wznoszenia.

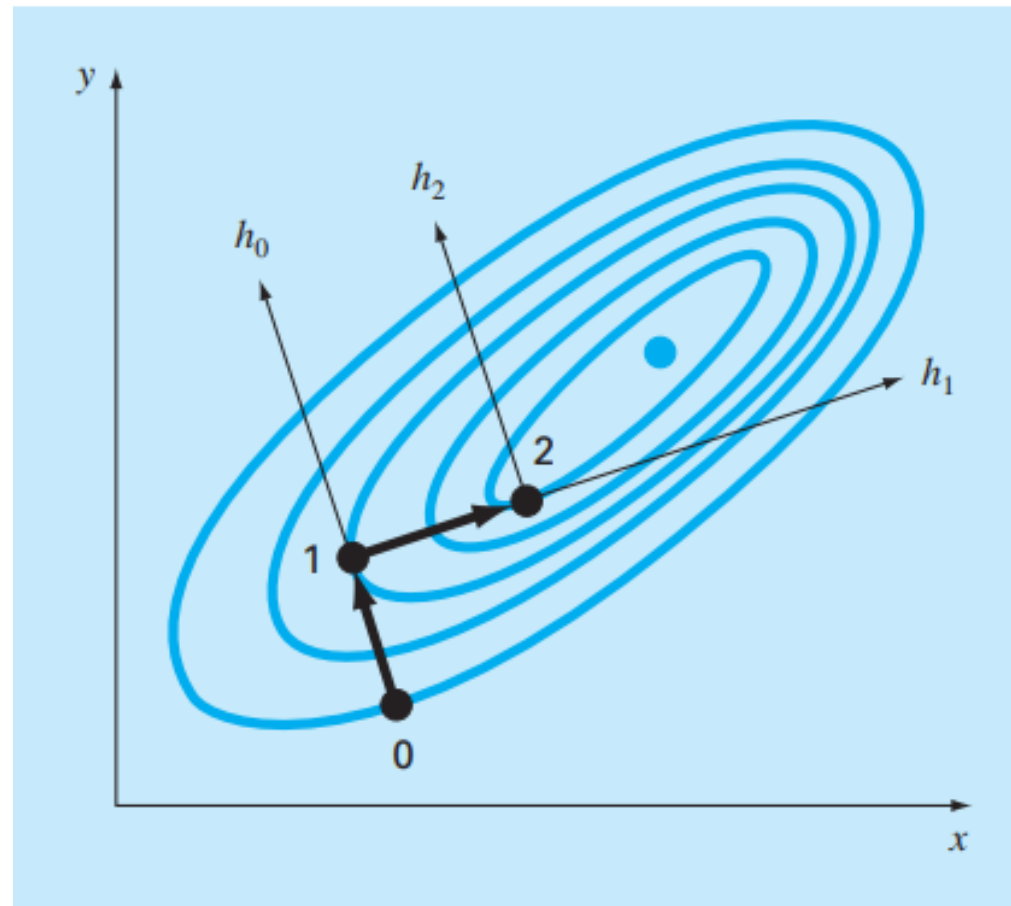
Rozwiązanie: idziemy niewielką odległość wzdłuż kierunku nachylenia. Następnie zatrzymujemy się i ponownie wyznaczamy gradient i idziemy w nowym kierunku. Powtarzamy proces. Niestety duże koszty obliczeniowe.

Rozwiązanie preferowane:

Przemieszczamy się po ustalonej ścieżce wzdłuż początkowego gradientu, aż $f(x, y)$ przestaje rosnąć. Ten punkt zatrzymania staje się punktem wyjścia, w którym ponownie wyznaczamy gradient. Proces powtarza się aż do osiągnięcia szczytu.

Optymalizacja wielowymiarowa bez ograniczeń

Metoda najszybszego wzrostu/spadku (ang. *Steepest Ascent/Descent Method*)



Optymalizacja wielowymiarowa bez ograniczeń

Metoda najszybszego wzrostu/spadku (ang. *Steepest Ascent/Descent Method*)

Przykład: Znajdź maksimum funkcji $f(x, y) = 2xy + 3x - x^2 - 2y^2$ dla w.p. $x = -1, y = 1$

$$\frac{\partial f}{\partial x} = 2y + 2 - 2x = 0$$

$$\frac{\partial f}{\partial y} = 2x - 4y = 0$$

$$\frac{\partial^2 f}{\partial x^2} = -2$$

$$\frac{\partial^2 f}{\partial y^2} = -4$$

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x} = 2$$

Hesjan jest równy $|H| = -2(-4) - 2^2 = 4$

$|H| > 0$ i $\frac{\partial^2 f}{\partial x^2} < 0 \rightarrow$ maksimum

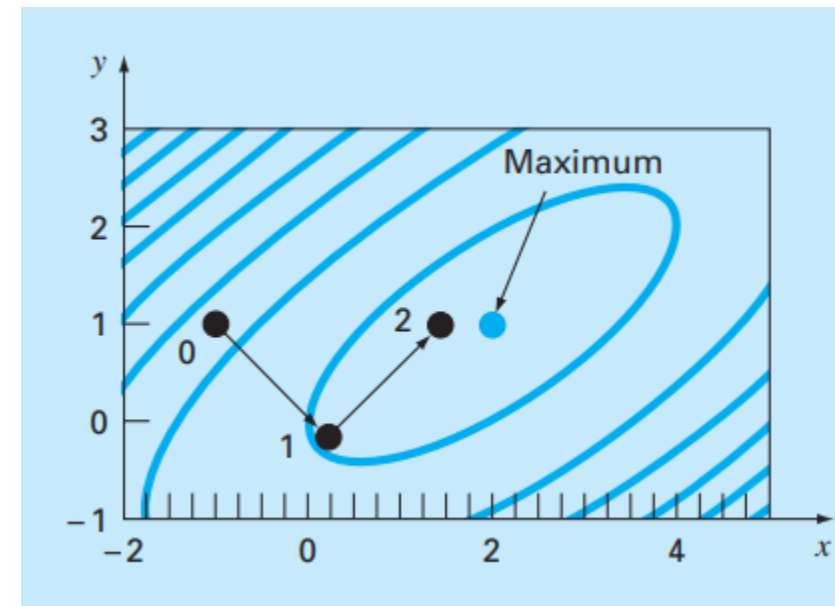
Wyznaczając krok początkowy: $g(h) = -180h^2 + 72h - 7$

Wyznaczając wprost maksimum ($h = h^*$):

$$g'(h^*) = 0$$

$$-360h^* + 72 = 0$$

$$h^* = 0.2$$



Optymalizacja wielowymiarowa bez ograniczeń

Metoda Newtona

Metoda Newtona dla funkcji jednej zmiennej może być uogólniona dla funkcji wielu zmiennych:

$$f(\mathbf{x}) = f(\mathbf{x}_i) + \nabla f^T(\mathbf{x}_i)(\mathbf{x} - \mathbf{x}_i) + \frac{1}{2}(\mathbf{x} - \mathbf{x}_i)^T H_i(\mathbf{x} - \mathbf{x}_i)$$

gdzie H_i jest macierzą Hessego (Hesjan). W minimum:

$$\frac{\partial f(\mathbf{x})}{\partial \mathbf{x}_j} = 0 \text{ dla } j = 1, 2, \dots, n$$

Zatem:

$$\nabla f = \nabla f(\mathbf{x}_i) + H_i(\mathbf{x} - \mathbf{x}_i) = 0$$

Jeżeli H jest macierzą nieosobliwą:

$$\mathbf{x}_{i+1} = \mathbf{x}_i - H_i^{-1} \nabla f$$

Optymalizacja wielowymiarowa bez ograniczeń

Metoda Marquardt

największy spadek gdy daleko od optimum, a gdy blisko metoda Newtona (modyfikacja diagonalu Hessianu)

Metody quasi-Newtonowskie (metody zmiennej metryki)

bazują na oszacowaniu bezpośredniej drogi do optimum w sposób podobny do metody Newtona. Jednak zamiast wykorzystywać macierz Hessego, wykorzystują aproksymację tej macierzy przy użyciu tylko pierwszych pochodnych cząstkowych.

Główne metody:

Davidon-Fletcher-Powell (DFP)

Broyden-Fletcher-Goldfarb-Shanno (BFGS)

Podsumowanie

$$x = m_1 B^{E_1}, \quad y = m_2 B^{E_2}$$

Błędy numeryczne

DOKŁADNOŚĆ W OBLICZENIACH NUMERYCZNYCH

Metody numeryczne są obarczone błędami.

Przyjmijmy że:

x wartość ścisła pewnej wielkości (często niewiadoma)

\tilde{x} wartość przybliżona dla x

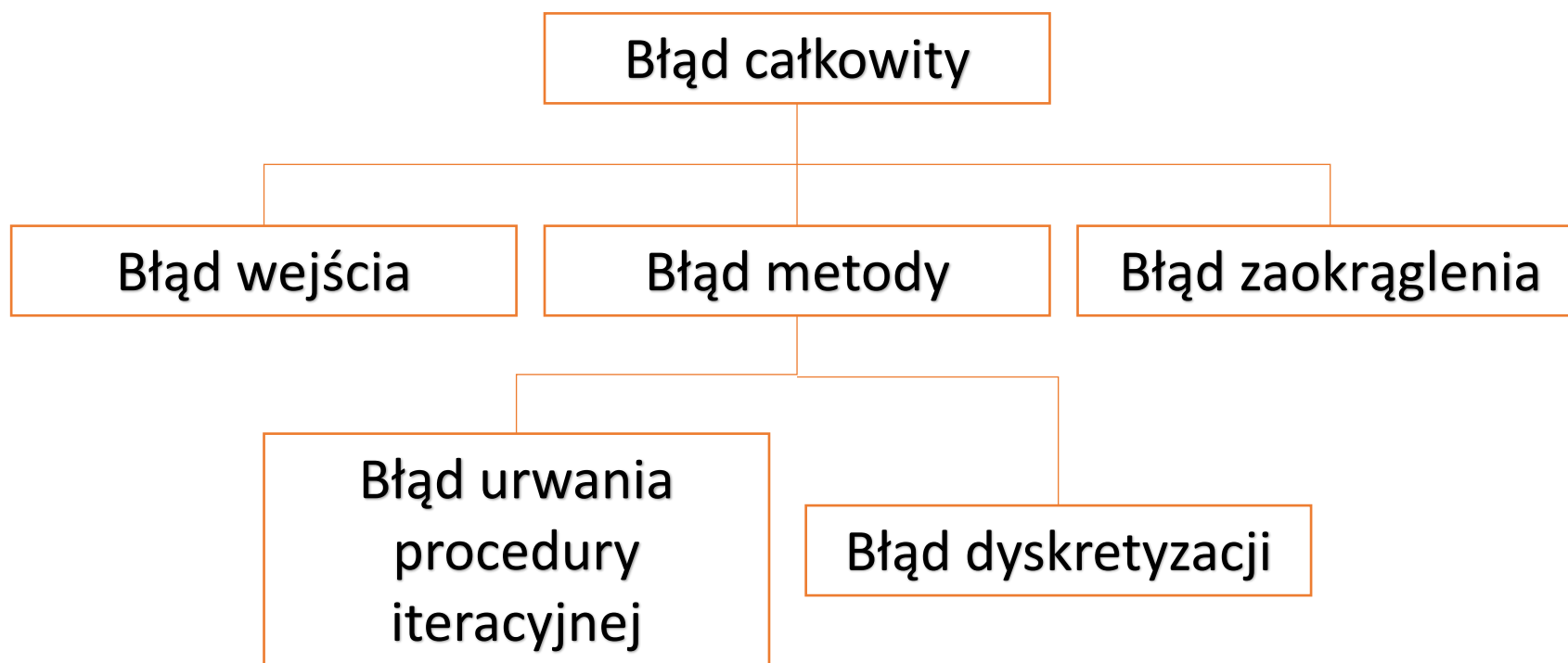
Błąd bezwzględny: $\epsilon(x) = |\Delta x| = |x - \tilde{x}|$

Błąd względny: $\epsilon_{rel}(x) = \left| \frac{\Delta x}{x} \right| = \left| \frac{x - \tilde{x}}{x} \right|$

Błędy numeryczne

DOKŁADNOŚĆ W OBLICZENIACH NUMERYCZNYCH

RODZAJE BŁĘDÓW



Miejsca zerowe funkcji

Metoda równego podziału

KROK 1: wybierz granice przedziału x_l i x_u gdzie następuje zmiana znaku. Poprawność wyboru można sprawdzić:
$$f(x_l)f(x_u) < 0$$



$$x_u = x_r$$

KROK 2: wyznacz miejsce zerowe x_r jako: $x_r = \frac{x_l + x_u}{2}$

$$x_l = x_r$$



KROK 3: Sprawdź znak wyrażenia:
$$f(x_l)f(x_r)$$



$$f(x_l)f(x_r) < 0$$

$$f(x_l)f(x_r) = 0$$

$$f(x_l)f(x_r) > 0$$



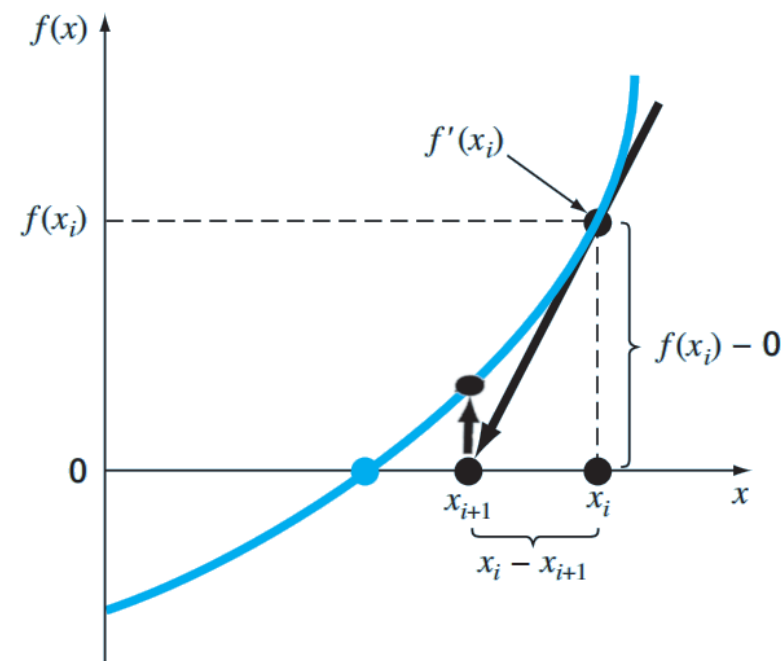
Metoda Newtona - Raphsona

- Załóżmy, że naszą wartością początkową jest x_i . Z definicji funkcji tangens możemy wyznaczyć współczynnik kierunkowy stycznej w x_i , który jest równocześnie pochodną funkcji w punkcie x_i

$$f'(x_i) = \frac{f(x_i) - 0}{x_i - x_{i+1}}$$

Przekształcając otrzymujemy:

$$x_{i+1} = x_i - \frac{f(x_i)}{f'(x_i)}$$



Różniczkowanie

Przypomnienie

Twierdzenie Taylora dla funkcji jednej zmiennej.

Jeżeli $y = f(x)$ jest funkcją ciągłą w przedziale $[a, a + h]$ i ma w tym przedziale ciągłe pochodne do rzędu $n - 1$ włącznie, przy czym wewnątrz tego przedziału istnieje pochodna rzędu n , to zachodzi:

$$\begin{aligned} f(a + h) = & f(a) + \frac{h}{1!} f'(a) + \frac{h^2}{2!} f''(a) + \dots \\ & + \frac{h^{n-1}}{(n-1)!} f^{(n-1)}(a) + \frac{h^n}{n!} f^n(a + \theta h) \end{aligned}$$

gdzie $0 < \theta < 1$. Wielkość h może przyjmować wartości dodatnie jak i ujemne.

RN – różnice centralne

Wyznaczmy $f'(x)$ z równania:

$$f(x+h) - f(x-h) = 2hf'(x) + \frac{h^3}{3}f'''(x) + \dots$$

$$f'(x) = \frac{f(x+h) - f(x-h)}{2h} - \frac{h^2}{6}f'''(x) - \dots$$

$$f'(x) = \frac{f(x+h) - f(x-h)}{2h} + \mathcal{O}(h^2)$$

Błąd odcięcia w notacji dużego O

Przybliżenie pierwszej pochodnej
RÓŻNICĄ CENTRALNĄ

Pochodne cząstkowe

Pochodne cząstkowe wyznaczamy analogicznie jak w przypadku funkcji jednej zmiennej. Przykładowo, dla $f(x, y)$ równomiernie próbkowanej, pierwsza pochodna cząstkowa może być przedstawiona z wykorzystaniem różnicy centralnej:

$$\frac{\partial f}{\partial x} = \frac{f(x + \Delta x, y) - f(x - \Delta x, y)}{2\Delta x}$$

$$\frac{\partial f}{\partial y} = \frac{f(x, y + \Delta y) - f(x, y - \Delta y)}{2\Delta y}$$

Wszystkie pozostałe, omówione metody można wykorzystać w analogiczny sposób.

Interpolacja i Aproksymacja

Interpolacja

Funkcja:

- Liniowa
- Kwadratowa
- Wielomianowa
- Splajnowa

Interpolacja Lagrange'a

Interpolacja Newtona

Interpolacja – Funkcje sklejane

Funkcje sklejane – splajny stopnia drugiego, ang. *Quadratic Splines*

Splajny stopnia drugiego mają ciągłe pierwsze pochodne w węzłach.

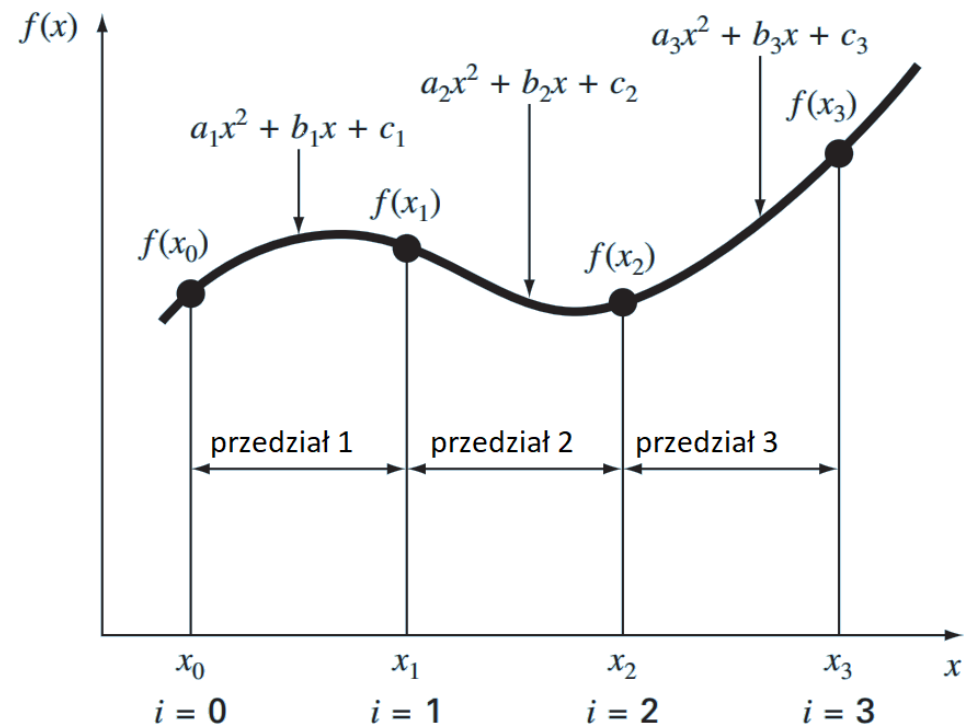
Zadaniem interpolacji splajnami stopnia drugiego jest wyznaczenie funkcji kwadratowych pomiędzy punktami danych:

$$f_i(x) = a_i x^2 + b_i x + c_i$$

Dla $n+1$ punktów ($i = 0, 1, \dots, n$),
Występuje n przedziałów i
w konsekwencji $3n$ niewiadomych
(a_i, b_i, c_i).



Potrzebujemy $3n$ równań, aby wyznaczyć niewiadome.



Interpolacja wielowymiarowa

Interpolacja biliniowa, dwuliniowa (ang. *Bilinear interpolation*)

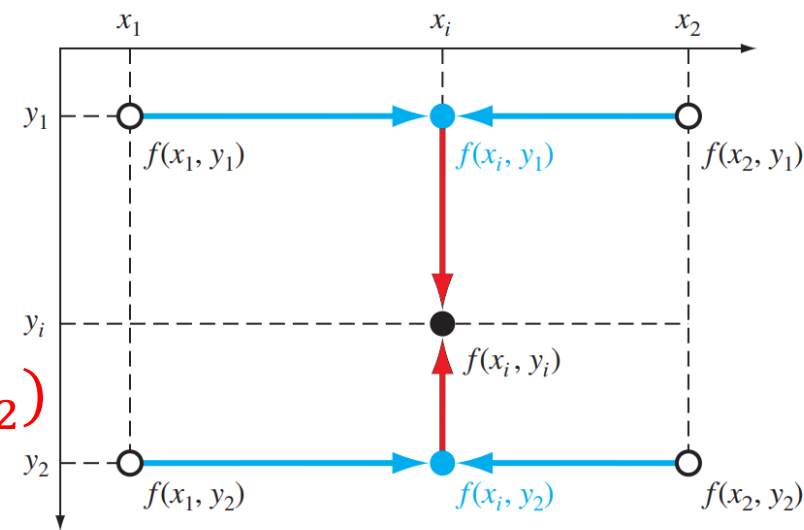
Dla ustalonej wartości y stosujemy jednowymiarową interpolację liniową w kierunku x . Bazując na formie Lagrange'a otrzymujemy:

$$f(x_i, y_1) = \frac{x_i - x_2}{x_1 - x_2} f(x_1, y_1) + \frac{x_i - x_1}{x_2 - x_1} f(x_2, y_1)$$

$$f(x_i, y_2) = \frac{x_i - x_2}{x_1 - x_2} f(x_1, y_2) + \frac{x_i - x_1}{x_2 - x_1} f(x_2, y_2)$$

Na bazie nowo wyznaczonych punktów liniowo interpolujemy w kierunku y :

$$f(x_i, y_i) = \frac{y_i - y_2}{y_1 - y_2} f(x_i, y_1) + \frac{y_i - y_1}{y_2 - y_1} f(x_i, y_2)$$



Metoda Najmniejszych Kwadratów – reprezentacja macierzowa

$$\mathbf{Y} = \mathbf{Z}\mathbf{A} + \mathbf{E}$$

$\mathbf{Y}^T = [y_1, \dots, y_n]$, wektor kolumnowy, zawierający wartości obserwacji

$\mathbf{E}^T = [e, \dots, e]$, wektor kolumnowy, zawierający reszt

$\mathbf{A}^T = [a_0, \dots, a_m]$, wektor kolumnowy, zawierający poszukiwane współczynniki
(rozmiar zależy od liczby współczynników, a nie próbek)

$$\mathbf{Z} = \begin{bmatrix} z_{01} & \cdots & z_{m1} \\ \vdots & \ddots & \vdots \\ z_{0n} & \cdots & z_{mn} \end{bmatrix}$$

gdzie macierz \mathbf{Z} zawiera wyliczone wartości funkcji bazowych dla zmierzonych argumentów.

$$\mathbf{A} = (\mathbf{Z}^T \mathbf{Z})^{-1} \mathbf{Z}^T \mathbf{Y}$$

Całkowanie

Całkowanie Numeryczne

Całkowanie numeryczne (kwadratura) jest z natury dokładniejszą procedurą w porównaniu z różniczkowaniem.

Całkowanie numeryczne $\int_a^b f(x)dx$ sprowadza się do aproksymacji sumą:

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

x_i - węzły całkowania, A_i - wagi, zależne od wybranej metody obliczania numerycznego całek.

Metody całkowania są wyprowadzane z wielomianowej interpolacji funkcji podcałkowej. Dlatego działają najlepiej, jeśli $f(x_i)$ można przybliżyć wielomianem.

CN – metoda trapezów

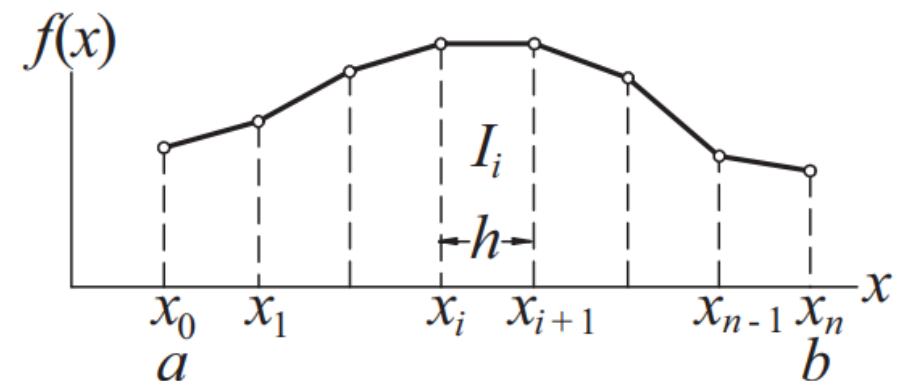
W praktyce, najczęściej metodę trapezów stosuje się odcinkami do całego sygnału.

Dla i – tego odcinka możemy zapisać:

$$I_i = (f(x_i) + f(x_{i+1})) \frac{h}{2}$$

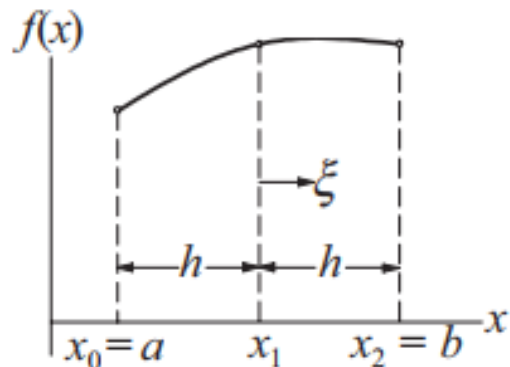
Czyli dla całego przedziału $[a, b]$ możemy zapisać:

$$I = \sum_{i=0}^{n-1} I_i = (f(x_0) + 2f(x_1) + 2f(x_2) + \dots + 2f(x_{n-1}) + f(x_n)) \frac{h}{2}$$



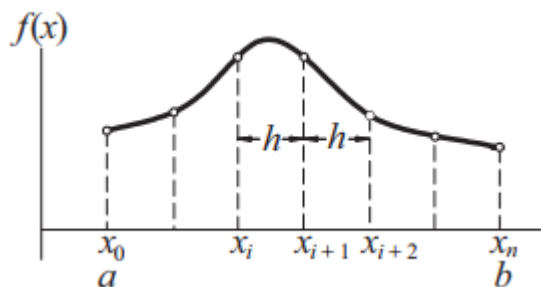
CN – metoda Simpsona

Metodę Simpsona opiera się na przybliżaniu funkcji całkowanej przez interpolację wielomianem drugiego stopnia i można ją otrzymać z wyrażenia Newtona-Cotesa dla $n = 2$



$$I = \left[f(a) + 4f\left(\frac{a+b}{2}\right) + f(b) \right] \frac{h}{3}$$

Przedział całkowania jest dzielony n parzystych przedziałów o szerokości $h = \frac{b-a}{n}$ każdy.
Dla dwóch sąsiednich przedziałów otrzymujemy:

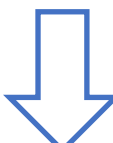



$$\int_{x_i}^{x_{i+2}} f(x) dx \approx \left(f(x_i) + 4f(x_{i+1}) + f(x_{i+2}) \right) \frac{h}{3}$$

Układy równań

Układy równań liniowych

$$\begin{aligned}a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,n}x_n &= b_1 \\a_{2,1}x_1 + a_{2,2}x_2 + \cdots + a_{2,n}x_n &= b_2 \\&\vdots \\a_{n,1}x_1 + a_{n,2}x_2 + \cdots + a_{n,n}x_n &= b_n\end{aligned}$$


$$\sum_{j=1}^n a_{i,j} x_j = b_i \quad (1 \leq i \leq n)$$


$$\begin{bmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & & \vdots \\ a_{i,1} & \cdots & a_{i,n} \\ \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,n} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_i \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_i \\ \vdots \\ b_n \end{bmatrix}$$



$$\mathbf{Ax} = \mathbf{b}$$

Układy równań liniowych

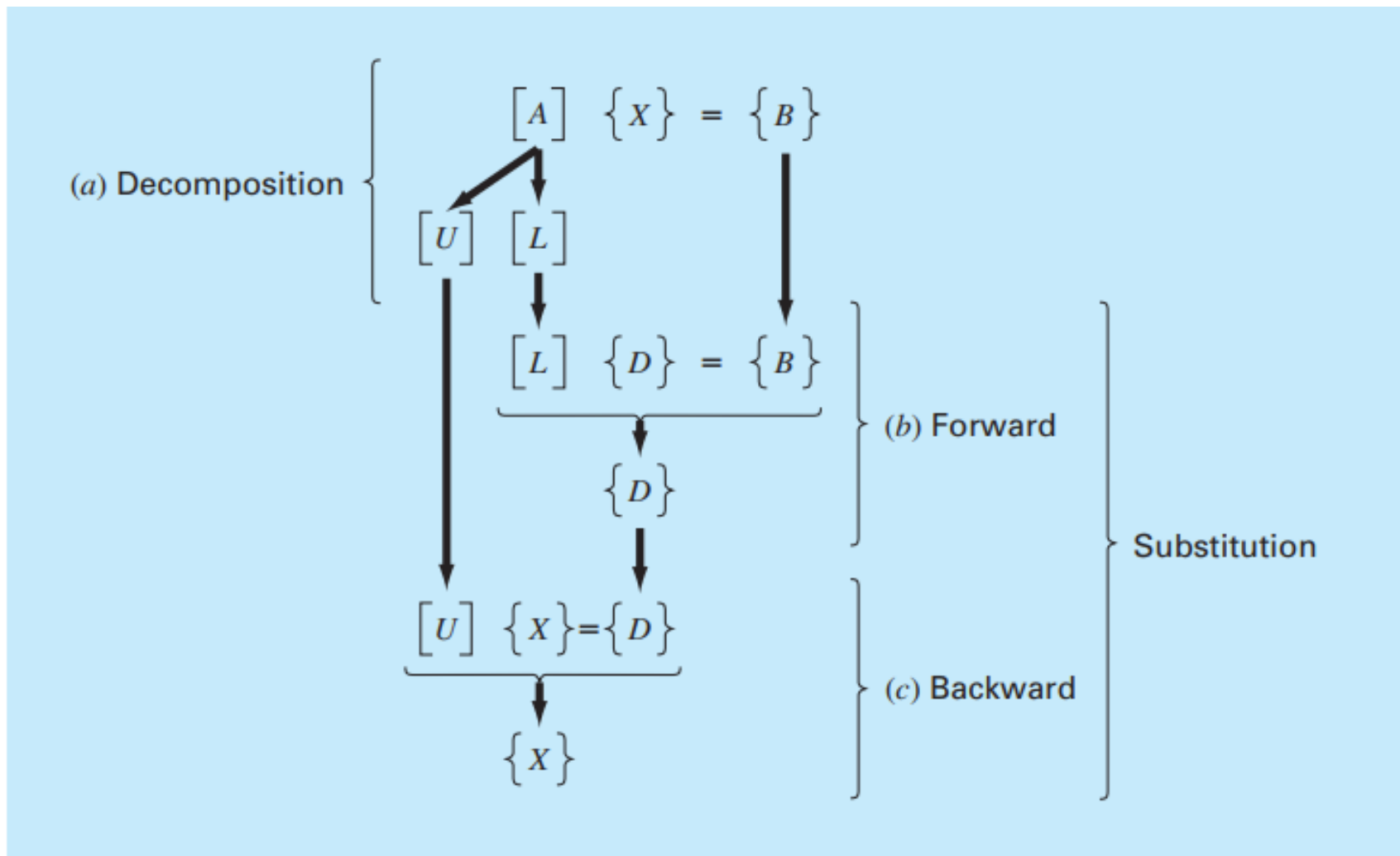
Eliminacja Gaussa

$$\begin{array}{c} \left[\begin{array}{ccc|c} a_{11} & a_{12} & a_{13} & b_1 \\ a_{21} & a_{22} & a_{23} & b_2 \\ a_{31} & a_{32} & a_{33} & b_3 \end{array} \right] \\ \Downarrow \\ \left[\begin{array}{ccc|c} a_{11} & a_{12} & a_{13} & b_1 \\ & a'_{22} & a'_{23} & b'_2 \\ & & a''_{33} & b''_3 \end{array} \right] \\ \Downarrow \\ \begin{array}{l} x_3 = b''_3/a''_{33} \\ x_2 = (b'_2 - a'_{23}x_3)/a'_{22} \\ x_1 = (b_1 - a_{12}x_2 - a_{13}x_3)/a_{11} \end{array} \end{array} \quad \begin{array}{l} \text{Forward} \\ \text{elimination} \\ \\ \text{Back} \\ \text{substitution} \end{array}$$

Eliminacja Gaussa – Jordana

$$\begin{array}{c} \left[\begin{array}{ccc|c} a_{11} & a_{12} & a_{13} & b_1 \\ a_{21} & a_{22} & a_{23} & b_2 \\ a_{31} & a_{32} & a_{33} & b_3 \end{array} \right] \\ \downarrow \\ \left[\begin{array}{ccc|c} 1 & 0 & 0 & b_1^{(n)} \\ 0 & 1 & 0 & b_2^{(n)} \\ 0 & 0 & 1 & b_3^{(n)} \end{array} \right] \\ \downarrow \\ \begin{array}{rcl} x_1 & & = b_1^{(n)} \\ & x_2 & = b_2^{(n)} \\ & & x_3 = b_3^{(n)} \end{array} \end{array}$$

Dekompozycja LU



Wartości i wektory własne

Do tej pory mówiliśmy o układach równań niejednorodnych (czyli takich gdzie występował niezerowy wektor **B**):

$$\mathbf{AX} = \mathbf{B}$$

Jeżeli równania składające się na taki układ są liniowo niezależne (niezerowy wyznacznik), będą miały unikalne rozwiązanie.

W układy równań jednorodnych mają postać:

$$\mathbf{AX} = \mathbf{0}$$

Oczywiście takie układy mają rozwiązanie trywialne (wszystkie niewiadome równe zero), natomiast rozwiązanie nietrywialne również występuje, ale nie musi być unikalne.

Problem wartości własnych przedstawia się zwykle w formie ogólnej:

$$\begin{aligned}(a_{1,1} - \lambda)x_1 + a_{1,2}x_2 + \cdots + a_{1,n}x_n &= 0 \\ a_{2,1}x_1 + (a_{2,2} - \lambda)x_2 + \cdots + a_{2,n}x_n &= 0 \\ \vdots & \\ a_{n,1}x_1 + a_{n,2}x_2 + \cdots + (a_{n,n} - \lambda)x_n &= 0\end{aligned}$$

Wartości i wektory własne

$$\begin{aligned}(a_{1,1} - \lambda)x_1 + a_{1,2}x_2 + \cdots + a_{1,n}x_n &= 0 \\ a_{2,1}x_1 + (a_{2,2} - \lambda)x_2 + \cdots + a_{2,n}x_n &= 0 \\ \vdots & \\ a_{n,1}x_1 + a_{n,2}x_2 + \cdots + (a_{n,n} - \lambda)x_n &= 0\end{aligned}$$

λ – jest nieznanym parametrem zwanym wartością własną. Rozwiązanie \mathbf{X} takiego układu nosi nazwę wektora własnego.

W zwartej formie równanie można przedstawić jako:

$$(\mathbf{A} - \lambda\mathbf{I})\mathbf{X} = 0$$

Jednym ze sposobów wyznaczenia λ jest wykorzystanie faktu, iż wyznacznik $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$ dla nietrywialnego rozwiązania.

Wielomian, który powstaje z wyliczenia wyznacznika nosi nazwę wielomiany charakterystycznego, a pierwiastki są wartościami własnymi.