

# Aproksymacja profilu wysokościowego

Jan Kaczerski 193237

---

## 1. Wprowadzanie

Celem projektu jest zbadanie przydatności dwóch metod aproksymacji funkcji:

1. Metoda aproksymacji Lagrange'a
2. Metodę wykorzystującą funkcje sklepane trzeciego stopnia

Metody te zbadane zostały przeze mnie na podstawie dwóch geograficznych tras:

1. Trasa równomiernie rosnąca – dalej nazywana trasą rosnącą
2. Przekrój Wielkiego Kanionu – trasa o zróżnicowanym charakterze, dynamicznie zmieniająca wysokości

Dane wykorzystane w poniższej pracy pochodzą z portalu eNauzanie.

## 2. Implementacja

Algorytmy zaimplementowane zostały w języku C++ z użyciem biblioteki `cmath` do wykonywania obliczeń oraz `sstream` do szybszego wczytywania plików `csv`. Ze względu na wykorzystanie macierzy, jako podstawę pod to zadanie użyłem swojego kodu z projektu nr 2.

Do wykonania wykresów użyłem biblioteki open-source `matplotlibcpp`.

## 3. Podstawy teoretyczne

### 3.1 Metoda Lagrange'a

Pierwszą badaną przeze mnie metodą aproksymacji jest metoda Lagrange'a. Charakteryzuje się ona przede wszystkim prostotą implementacji oraz niską złożonością pamięciową – nie ma potrzeby tworzyć ogromnych macierzy jak w przypadku sklepania funkcji kubicznych. Wadami owej metody jest pojawianie się błędu Runge'go, czyli pogorszenia jakości aproksymacji na brzegach przedziału wraz ze wzrostem węzłów interpolacji.

Wartość aproksymowanej funkcji według metody Lagrange'a otrzymuje się według następujących wzorów:

$$F(x) = \sum_{i=1}^{n+1} y_i \phi_i(x) \qquad \phi_i(x) = \prod_{j=1, j \neq i}^{n+1} \frac{(x - x_j)}{(x_i - x_j)}$$

Podchodząc do problemu ze strony informatycznej, widać, że aby znaleźć wartość zinterpolowanej funkcji wystarczy zaimplementować dwie pętle, skąd wynika prostota implementacji.

### 3.2 Metoda wykorzystująca funkcje sklepane trzeciego stopnia

Metoda ta pozwala rozwiązać problem występowania błędu Runge'go, jednak jest ona o wiele bardziej skomplikowana w implementacji, a wymagane do jej działania macierze zajmują znaczącą ilość miejsca w pamięci. Implementacja metody polega na obliczeniu parametrów wielomianu trzeciego stopnia w taki sposób, że pomiędzy każdymi dwoma węzłami otrzymujemy inne parametry do funkcji. Zastosowanie trzeciego stopnia wielomianu pozwala kontrolować zakrzywienie funkcji oraz, poprzez kontrolę drugiej pochodnej, minimalizuje ilość punktów przegięcia.

Chcąc poprowadzić funkcję trzeciego stopnia pomiędzy dwoma punktami, otrzymujemy poniższe wzory:

$$\begin{aligned}a_0 + b_0(x_0 - x_0) + c_0(x_0 - x_0)^2 + d_0(x_0 - x_0)^3 &= f(x_0) \\x_{i+1} - x_i &= h \\a_0 + b_0h + c_0h^2 + d_0h^3 &= f(x_1)\end{aligned}$$

Natomiast dodając do tego kontrolę nad kształtem funkcji, czyli przyrównanie pochodnych w celu zachowania tempa wzrostu oraz podwójnych pochodnych w celu ograniczenia niepoprawnych oscylacji wokół oryginalnej funkcji (zmniejszenie ilości punktów przegięcia), otrzymujemy:

$$\begin{aligned}S'_0(x_1) = S'_1(x_1) \text{ w } x_1 &\Rightarrow b_0 + 2c_0h + 3d_0h^2 = b_1 \\S''_0(x_1) = S''_1(x_1) \text{ w } x_1 &\Rightarrow 2c_0 + 6d_0h = 2c_1\end{aligned}$$

Mając więc cztery równania, możemy odnaleźć cztery parametry a, b, c, d między każdymi dwoma punktami. W miejscu, gdy nie będzie już następnej funkcji dodajemy dwie znane zależności:

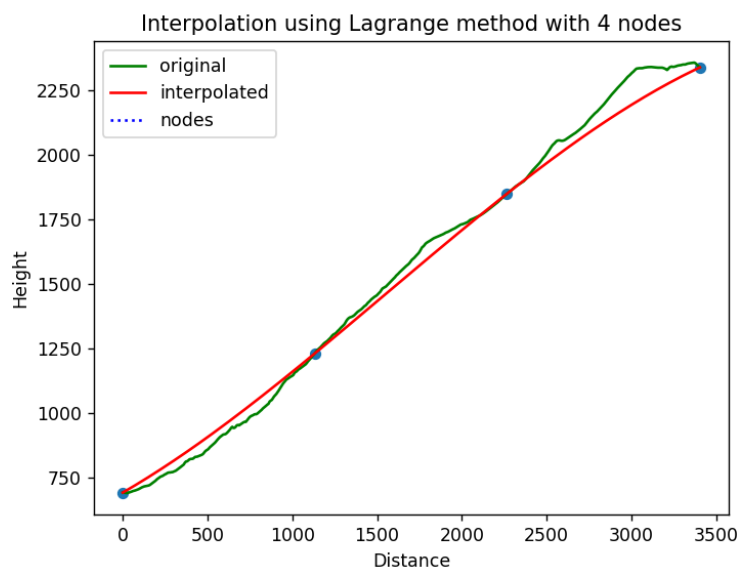
$$\begin{aligned}S''_0(x_0) = 0 \text{ ; } S''_1(x_2) = 0 \text{ pochodnej: } &\Rightarrow \\c_0 = 0 \\2c_1 + 6d_1h = 0\end{aligned}$$

Finalnie otrzymujemy  $4 \cdot (n-1)$  równań i  $4 \cdot (n-1)$  szukanych parametrów, co pozwala utworzyć jednoznaczne równanie macierzowe.

#### 4. Podstawowa analiza

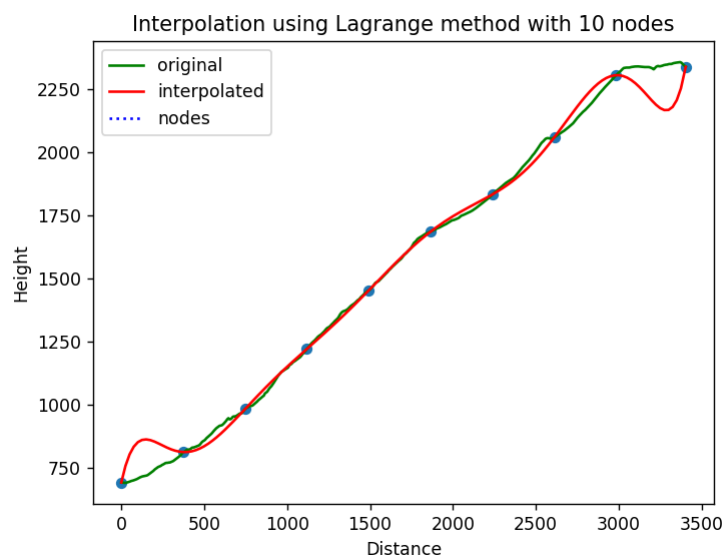
**4.1 Interpolacja wielomianowa trasy rosnącej** – badanie zależności jakości interpolacji od ilości węzłów

##### 4.1.1 Równomierne rozstawione 4 węzły



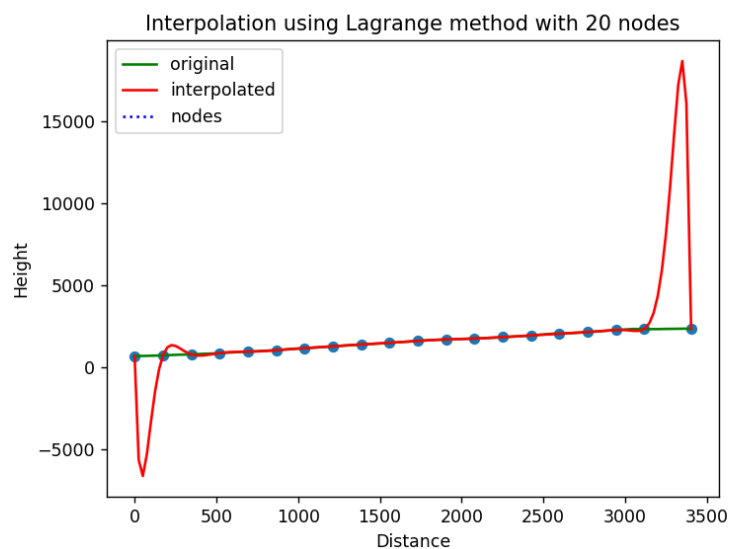
Jak widać na powyższym wykresie, interpolacja w podanym przykładzie jest zaskakująco dobra już dla czterech węzłów. Jest to jednak spowodowane jedynie stałym trendem wzrostu trasy.

#### 4.1.2 Równomierne rozstawienie 10 węzłów



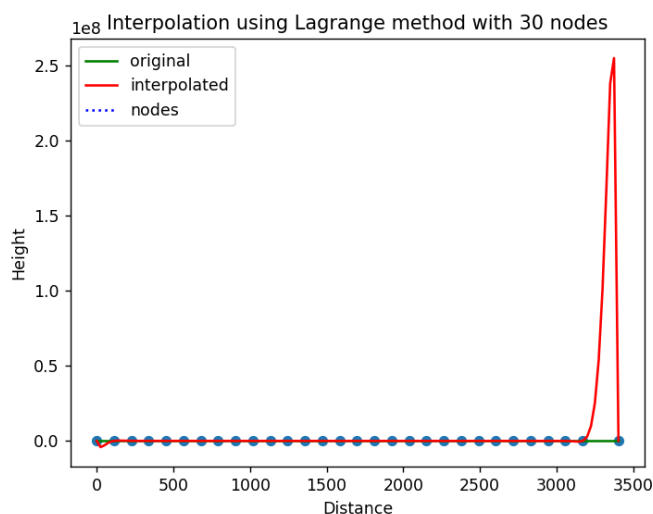
Jakość aproksymacji wyraźnie się poprawiła w środkowej części próbek danych, jednak na brzegach zauważyć można odchylenia. Są one jednak akceptowalne – nie mają zbyt dużych błędów, a błędy występują w skrajnych 10% przedziału z obu stron.

#### 4.1.3 Równomiernie rozstawione 20 węzłów



Wraz z ilością węzłów postępuje efekt ulepszenia aproksymacji w środkowej części funkcji i pogorszenia na brzegach. Jest to zjawisko Runge'go, które zaczyna się nasilać.

#### 4.1.4 Równomiernie rozstawione 30 węzłów

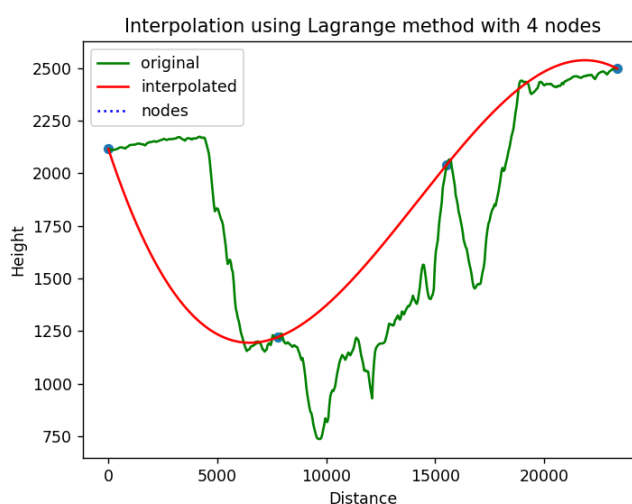


W podanym przykładzie zauważyć można efekt Runge'go w pełnej skali. Pomimo zwiększenia ilości węzłów i jak mogłoby się wydawać, poprawy jakości, aproksymacja wzrosła z maksymalnej wartości  $\sim 2300$  do  $\sim 2.5E8$ . Błąd ten jest ogromny, a aproksymacja przestaje być zdana do użytku w ponad 20% przedziału.

Wyniki badania w zadanej funkcji pozwalają z dużą dokładnością określić dane w około 80% przedziału, natomiast wartości przy brzegach nie nadają się do użytku.

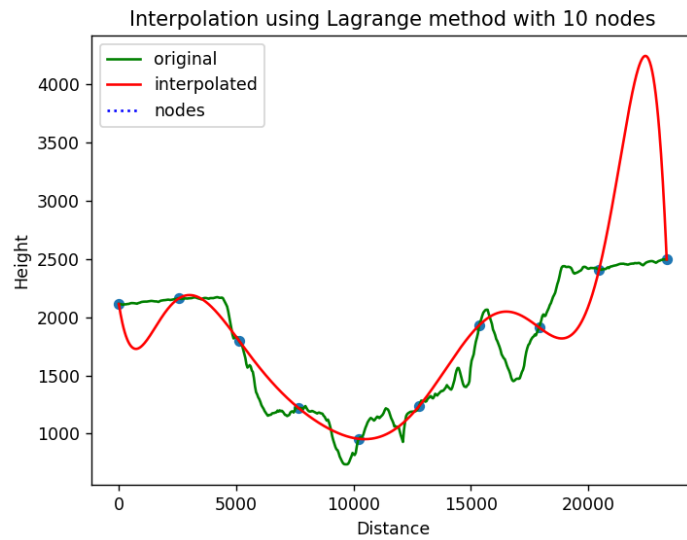
### 4.2 Interpolacja wielomianowa trasy Wielki Kanion – badanie zależności jakości interpolacji od ilości węzłów

#### 4.2.1 Równomiernie rozstawione 4 węzły



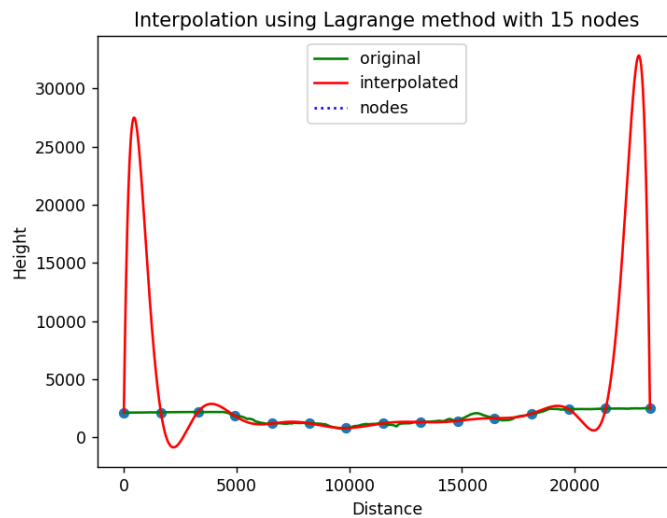
Trasa Wielkiego Kanionu ma dużo wzniesień i spadków. Tym razem aproksymacja z użyciem czterech węzłów nie jest w żadnym stopniu jakościowa, ponieważ funkcja ta nie rośnie (w przybliżeniu) jednostajnie.

#### 4.2.2 Równomiernie rozstawione 10 węzłów



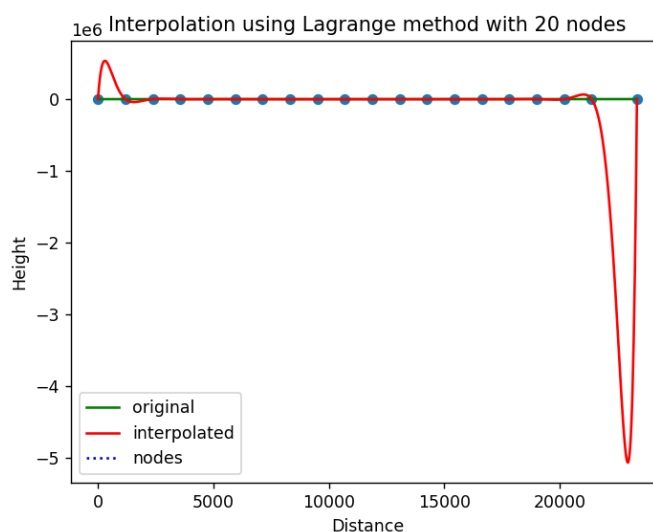
W powyższym przykładzie danych z Wielkiego Kanionu zaobserwować można znaczne polepszenie jakości interpolacji względem przykładu z punktu 4.2.1, jednak cały czas nie jest ona wiarygodna. Ponadto, zaobserwować można zaczynający się nasilać efekt Runge'go.

#### 4.2.3 Równomiernie rozstawione 15 węzłów

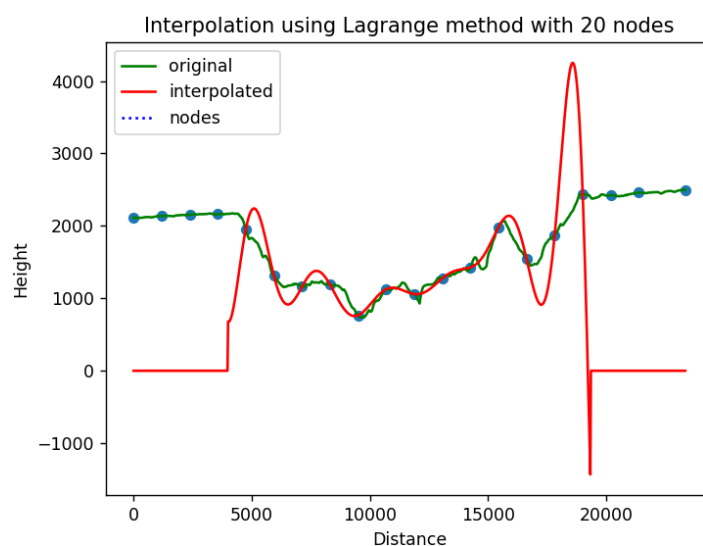


Na wykresie widać delikatne polepszenie jakości interpolacji w środkowej części danych. Błędy na skrajach są silne. Dla 15 węzłów nie otrzymujemy ani dokładnych wartości, ani braku błędu Runge'go.

#### 4.2.4 Równomiernie rozstawione 20 węzłów



Dla 20 węzłów maksymalna wartość bezwzględna aproksymowanej funkcji sięga  $5E6$ , gdzie w oryginalnej funkcji nie osiąga do 2500. Efekt Runge'go jest tu bardzo silnie zaznaczony, a końcówki przedziału nie nadają się do obliczeń.

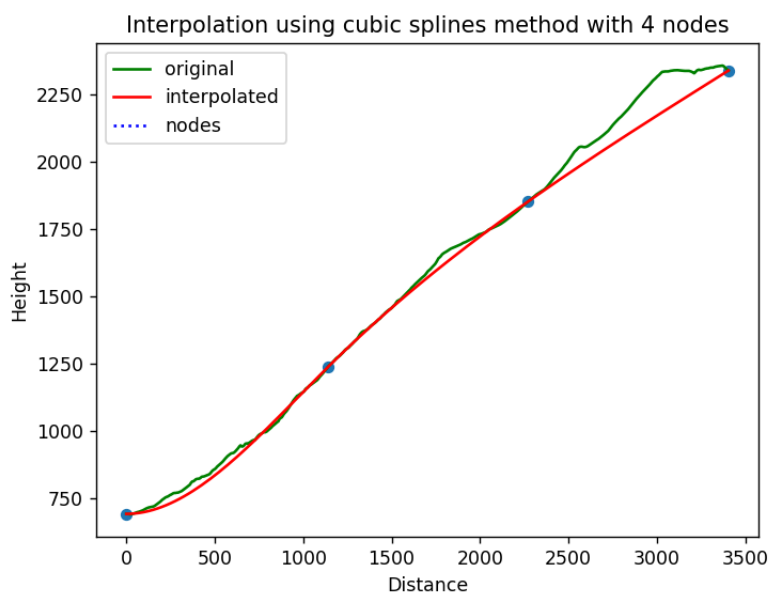


Po wyzerowaniu problematycznych wartości na skrajach przedziału, zauważyć można, że w środkowej części danych aproksymacja pierwszy raz zbliżyła się wystarczająco dobrze do funkcji, aby uznać ją za akceptowalną do dalszych obliczeń.

Przeprowadzone badania ukazują, że metoda wielomianowa jest użyteczna, gdy obliczenia prowadzone są z dużą ilością węzłów, jednak ponad 40% wykonanych obliczeń nie nadaje się do użytku.

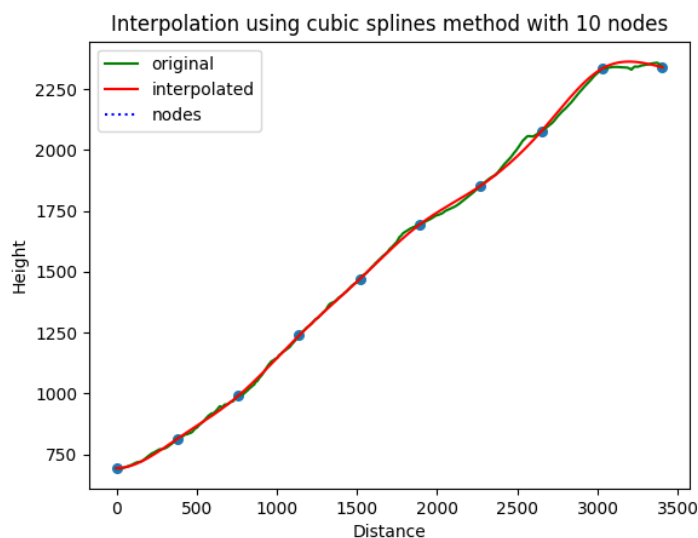
#### 4.3 Aproksymacja wykorzystująca funkcje sklepane dla trasy rosnącej – badanie zależności jakości interpolacji od ilości węzłów

##### 4.3.1 Równomierne rozmieszczenie 4 węzłów



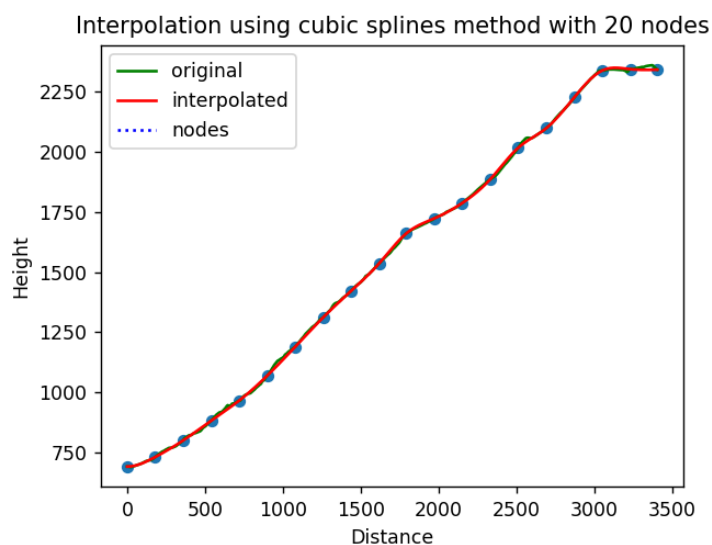
Równomierne rozstawienie 4 węzłów zwraca akceptowalne przybliżenie aproksymowanej funkcji. Należy jednak pamiętać o wnioskach z punktu 4.2, w których badania dowiodły, że jest to spowodowane jedynie przez przybliżoną jednostajność wzrostu funkcji.

##### 4.3.2 Równomierne rozmieszczenie 10 węzłów



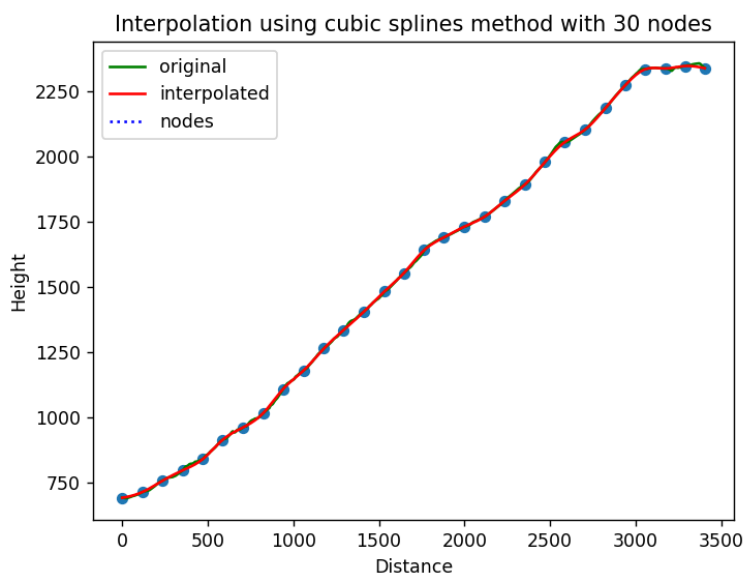
Wyniki interpolacji wykorzystującej 10 węzłów są wiarygodne i w badanym przypadku można ich używać do dalszych obliczeń. Należy również zwrócić uwagę na brak jakichkolwiek zapowiedzi efektu Runge'go.

#### 4.3.3 Równomierne rozmieszczenie 20 węzłów



Wartości wyliczone poprzez aproksymację praktycznie pokrywają się z oryginalną funkcją. Już w tym momencie można zauważyć, że metoda sklejanie funkcji jest znacznie dokładniejsza niż metoda wielomianowa, jednak należy pamiętać że rozpatrywany przypadek jest specyficzny, a interpolacja będzie zachowywać się inaczej przy dynamicznych funkcjach.

#### 4.3.4 Równomierne rozmieszczenie 30 węzłów

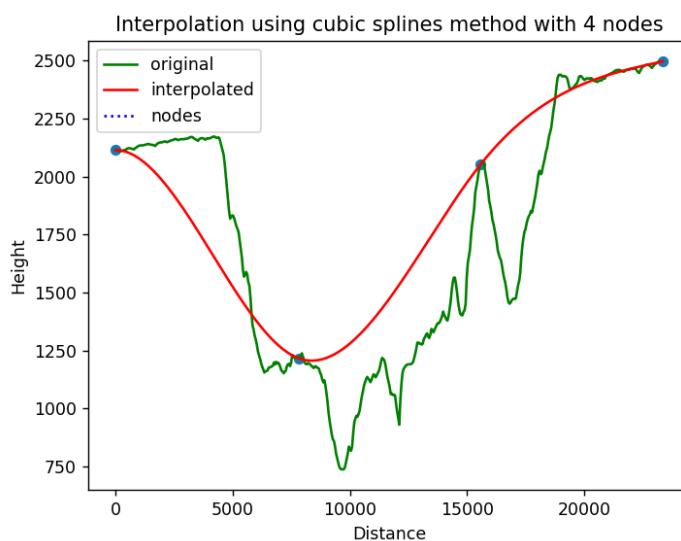


Dodanie dodatkowych 10 węzłów przybliżyło wyniki do oryginalnej funkcji, jednak już przy w 4.3.2 wyniki mogły być uznane za wiarygodne. Przypadek ten obrazuje jednak ogromną zaletę metody sklejanie funkcji, czyli brak efektu Runge'go przy aż 30 węzłach.



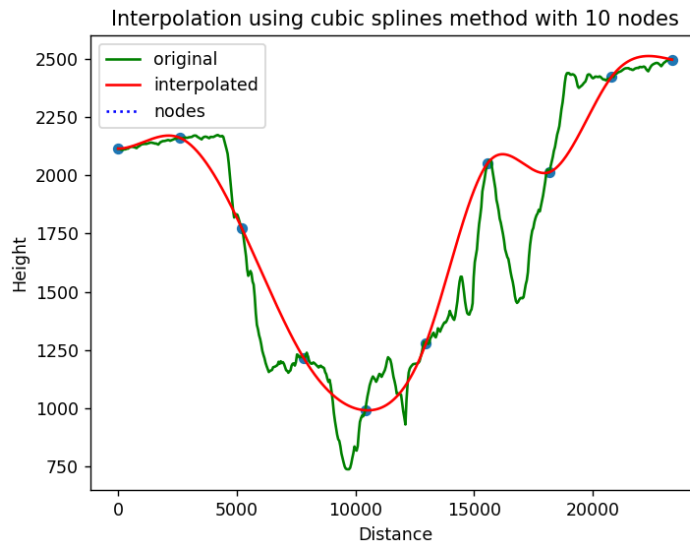
#### 4.4 Aproksymacja wykorzystująca funkcje sklejane dla trasy Wielkiego Kanionu – badanie zależności jakości interpolacji od ilości węzłów

##### 4.4.1 Równomierne rozmieszczenie 4 węzłów



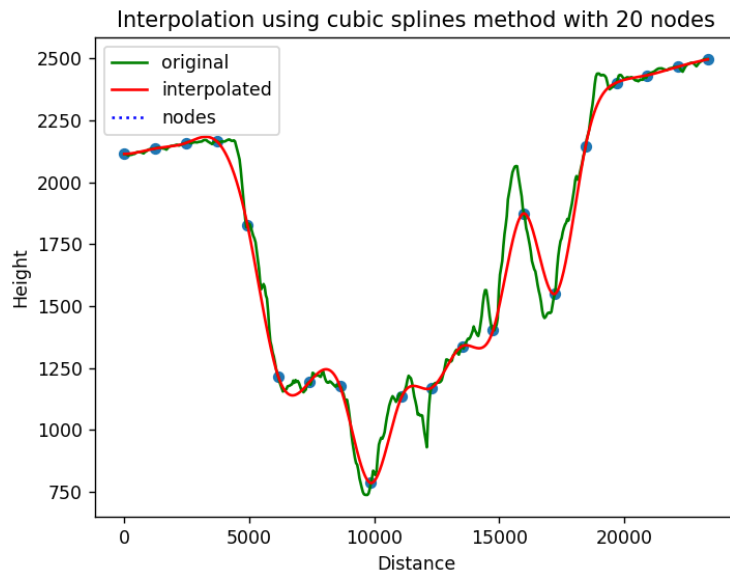
Przy użyciu czterech węzłów interpolacji uzyskane wyniki nie różnią się gradualnie od wyników z metody wielomianowej (4.2.1). Rezultaty dalej nie odwzorowują oryginalnej funkcji, a do ich uzyskania trzeba było użyć więcej (choć niewiele) zasobów komputera.

##### 4.4.2 Równomierne rozmieszczenie 10 węzłów



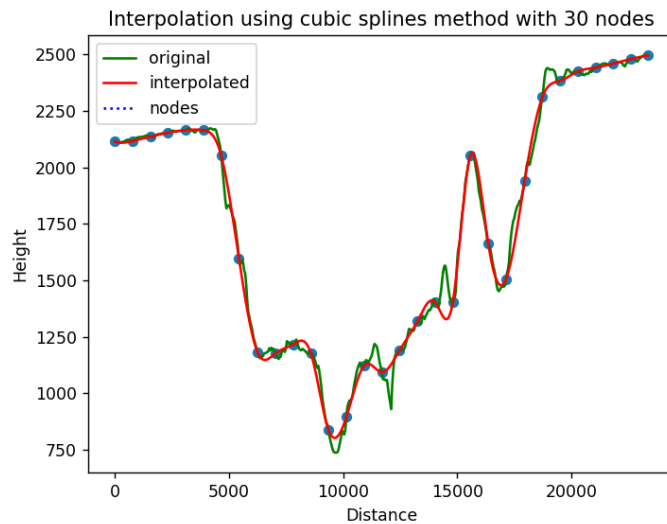
Aproksymacja funkcji przy użyciu 10 węzłów jest znacznie lepsza, jednak wyniki cały czas nie zahaczają o większość wartości oryginalnej funkcji. Należy jednak zwrócić uwagę na to, że przy użyciu 10 węzłów metoda wielomianowa zaczynała generować błędy związane z efektem Runge’go, które tu nie występują.

#### 4.4.3 Równomierne rozmieszczenie 20 węzłów



Wyniki są bliskie funkcji w większości miejsc, co oznacza, że aproksymację z użyciem sklejania funkcji można uznać za w dużej części wiarygodną przy węzłach. W otrzymanych wynikach dalej nie ma efektu Runge'go.

#### 4.4.4 Równomierne rozmieszczenie 30 węzłów

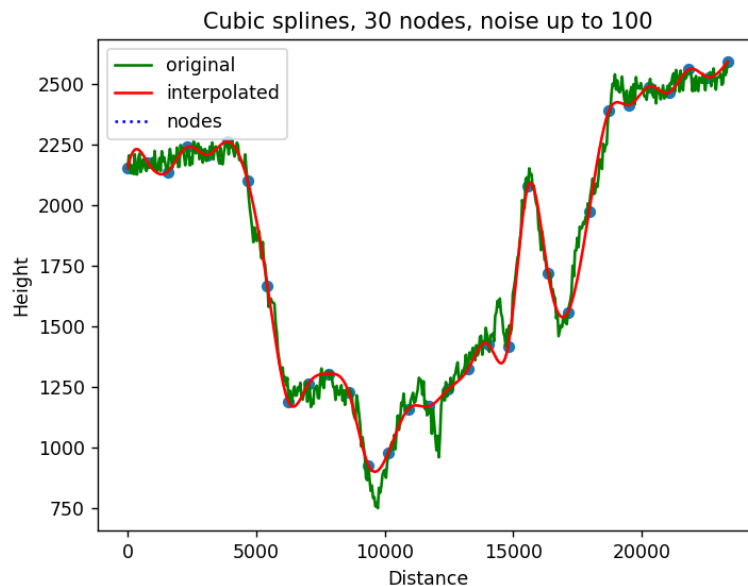


Przy 30 węzłach otrzymany rezultat zachowuje się prawie jak oryginalna funkcja. Jednocześnie nie ma śladu po efekcie Runge'go, co pokazuje jak dużą przewagę ma badana metoda nad metodą Lagrange'a.

## 5. Szczegółowa analiza

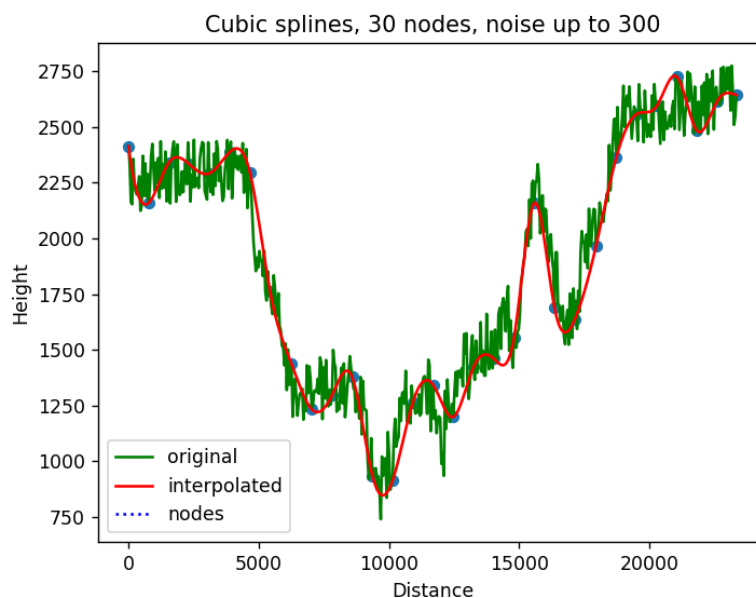
### 5.1 Zaszumienie danych

#### 5.1.1 Zaszumienie w granicach $\langle -100, 100 \rangle$



Zaszumienie sygnału wpływa na kształt funkcji, jednak cały czas aproksymacja wyraźnie podąża za trendem funkcji. Biorąc pod uwagę brak charakterystycznych dla szumów spadków i wzlotów w wynikowej funkcji, można zauważyć że metoda ta nadaje się do odzuszumiania sygnałów.

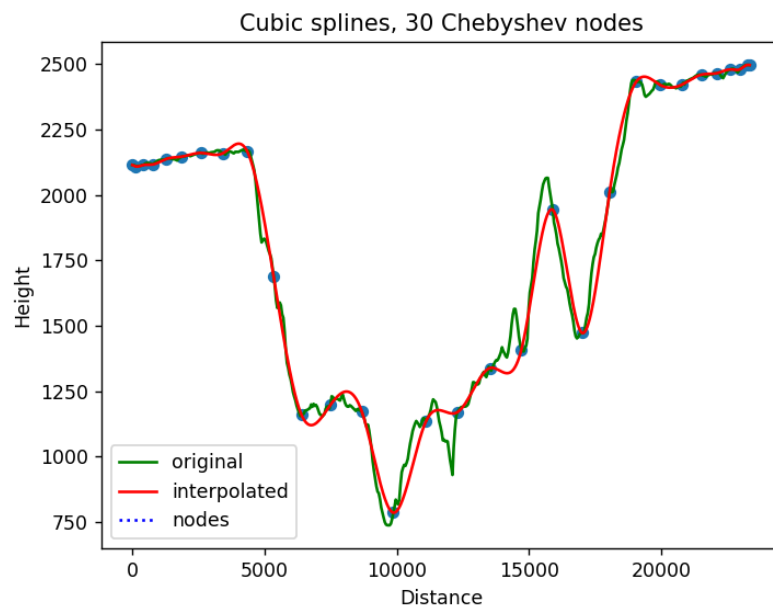
#### 5.1.2 Zaszumienie w granicach $\langle -300, 300 \rangle$



Przy zwiększeniu zaszumienia można zauważyć, że funkcja momentami zaczyna oscylować. Dzieje się tak z powodu wybierania równomiernych próbek, przez co raz można trafić w zaszumienie  $+300$ , a raz w zaszumienie  $-300$ . Tak umieszczone węzły są dobrze interpolowane, jednak wynikowy sygnał wygląda jakby był niedokładny.

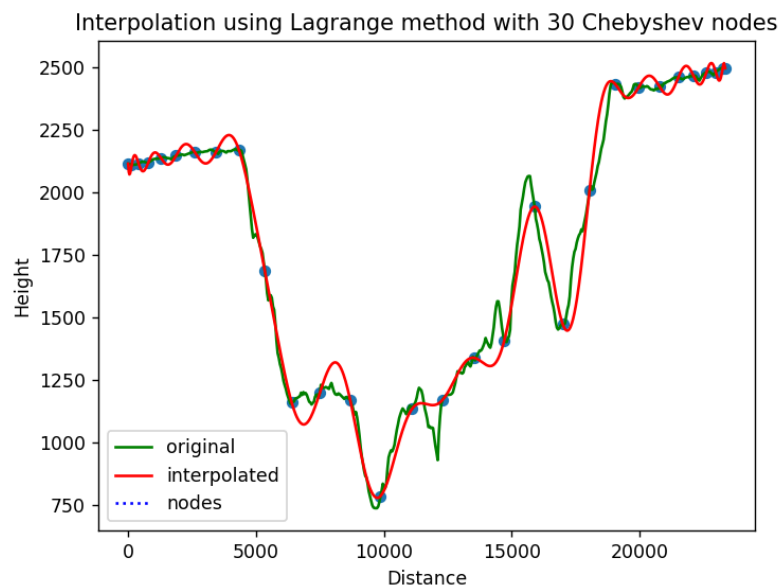
## 5.2 Wykorzystanie węzłów Chebysheva

### 5.2.1 W metodzie sklejania funkcji



Zastosowanie węzłów Chebysheva w metodzie sklejania funkcji nie zmienia drastycznie rezultatów aproksymacji. Porównując wykresy widać, że mniejsze zagęszczenie węzłów w środku skutkuje mniejszą dokładnością, jednak rezultaty te cały czas są akceptowalne.

### 5.2.2 W metodzie Lagrange'a



Zastosowanie węzłów Chebysheva pozwoliło zniwelować efekt Runge'go, co jest ogromnym ulepszeniem algorytmu. Wybranie tych węzłów obniża dokładność w środku próbki danych, jednak redukcja efektu Runge'go znacznie ulepsza algorytm.

## 6. Podsumowanie

Na podstawie powyższych badań i ich rezultatów wysnuć można poniższe wnioski:

- Interpolacja wielomianowa pomimo mniejszego zapotrzebowania na pamięć, nie daje akceptowalnych rezultatów na całej badanej dziedzinie.
- Interpolacja wielomianowa obarczona jest efektem Runge'go, przez co, aby zwiększyć dokładność pomiarów wymagane jest wykonanie dodatkowych obliczeń, których nie można użyć.
- Błąd Runge'go w interpolacji wielomianowej można poprawić poprzez zastosowanie w funkcji węzłów Chebysheva zamiast węzłów równomiernie rozłożonych.
- Interpolacja wykorzystująca sklejanie funkcji pozwala otrzymać lepsze rezultaty oraz wyeliminować błąd Runge'go.
- Wykorzystanie węzłów Chebysheva w metodzie sklejanie funkcji pogarsza jej rezultaty z powodu mniejszego zagęszczenia węzłów w środkowym przedziale dziedziny.
- Metoda sklejanie funkcji zwraca pewniejsze wyniki niż metoda wielomianowa.
- Niektóre funkcje mogą zostać dobrze przybliżone wykorzystując nawet tylko 4 węzły, jednak taki wynik może zostać uznany za poprawny jedynie mając dodatkowe informacje o przybliżanej funkcji.
- Funkcje aproksymacji nadają się do wygładzania zaszumionych sygnałów.
- Im więcej węzłów jest użyte do interpolacji, tym zwrócony wynik będzie dokładniejszy. Należy jednak pamiętać o ignorowaniu brzegów przedziału używając funkcji wielomianowej.
- Używając interpolacji wielomianowej błąd Runge'go pojawia się przy około 10 węzłach.