Instutut Informatyki Uniwersytetu Wrocławskiego Licencjacki Projekt Programistyczny Prowadzący: Stanisław Lewanowicz

Maciej Pacut

Znajdowanie zer wielomianu w postaci Beziera metodą przycinania sześciennego

1. Wstęp

Niniejsza praca prezentuje efektywną metodę znajdowania zer wielomianu w postaci Beziera. Przed zaprezentowaniem metody przycinania sześciennego zbudujemy narzędzia i terminologię potrzebną do łatwego i jednoznacznego przedstawienia metody. Opisane będą podejścia do numerycznego znajdowania zer funkcji. Zaprezentowane będą metody znajdowania zer specyficznej rodziny funkcji jakimi są wielomiany wraz z własnościami, które umożliwiają projektowanie specyficznych dla wielomianów metod znajdowania zer. Przedstawione zostaną wielomiany w postaci Beziera wraz z ich zastosowaniami. Zaprezentujemy specyficzne własności wielomianów w tej postaci. Dokładniej opisane będzie zadanie aproksymacji wielomianowej.

2. Numeryczne znajdowanie zer

2.1. Znajdowanie jednego zera

W niniejszej pracy będziemy rozpatrywać zbiór metod znajdowania zer o pewnych cechach wspólnych. Działanie tych metod polega na zawężaniu danego przedziału do momentu osiągnięcia zadanej dokładności lub wykluczenia istnienia zera w danym przedziale.

Metoda znajdowania jednego zera funkcji dostaje argumenty:

- 1. Pewien opis funkcji f, który umożliwia obliczenie wartości tej funkcji dla dowolnego argumentu.
- 2. Przedział początkowy P_0 , czyli przedział w którym szukamy zera funkcji f; metoda nie znajdzie żadnego zera spoza tego przedziału.
- 3. Żądaną dokładność ε znalezienia zera funkcji f.

Metoda numerycznego znajdowania jednego zera funkcji f generuje ciąg przedziałów $\langle P_0, P_1, ..., P_n \rangle$ o własnościach:

1. Każdy przedział tego ciągu jest podprzedziałem swojego poprzednika:

$$P_k \subset P_{k-1} \qquad (1 \leqslant k \leqslant n)$$

2. Jeśli pewien przedział zawiera zero funkcji f to następnik tego przedziału także zawiera to zero:

$$x_0 \in P_{k-1} \Rightarrow x_0 \in P_k \qquad (1 \leqslant k \leqslant n; f(x_0) = 0)$$

3. P_n jest jedynym przedziałem krótszym niż 2ε .

Jeśli w przedziałe P_0 znaleziono zero, to wynikiem działania metody jest środek przedziału P_n . Skoro zero znajduje się w P_n i jego długość jest mniejsza

niż 2ε to udało się zlokalizować zero funkcji f z błędem mniejszym niż ε . Jeśli nie znaleziono zera, P_n jest przedziałem pustym.

Metody znajdowania zer zdefiniowane powyższym schematem różnią się sposobem konstrukcji ciągu $\langle P_0, P_1, ..., P_n \rangle$. Do kompletnego opisu metody wystarczy zdefiniować, w jaki sposób mając przedział P_k konstruujemy przedział P_{k+1} .

Przykładem metody znajdowania jednego zera funkcji jest bisekcja. W bisekcji przedział P_{k+1} otrzymuje się dzieląc P_k na pół i jako P_{k+1} wybierając tę połowę, w której jest możliwość wystąpienia zera. W przypadku, gdy w żadnej połowie nie ma możliwości wystąpienia zera, $P_{k+1} = P_n$ i P_{k+1} jest przedziałem pustym i metoda kończy działanie. W bisekcji należy określić realizację operacji określającej możliwość wystąpienia zera w danym przedziałe (w danej połowie przedziału P_k). Przykładową taką operacją jest sprawdzenie, czy funkcja f ma w końcach sprawdzanego przedziału różne znaki.

! rysunek bisekcji

2.2. Znajdowanie wielu zer

Metod znajdowania jednego zera nie stosuje się do funkcji mających więcej niż jedno zero w przedziale P_0 . Przykładowo, metoda bisekcji ma działanie niezdefiniowane w przypadku, gdy w obu połowach przedziału P_k jest możliwość wystąpienia zera.

Zmodyfikujmy powyższy schemat znajdowania jednego zera. Zamiast generowania ciągu przedziałów, metoda szukania wielu zer generuje drzewo przedziałów. Drzewo przedziałów ma własności:

- 1. Korzeniem jest przedział początkowy P_0 .
- 2. Każdy potomek jest podprzedziałem swojego rodzica.
- 3. Przedziały mające wspólnego rodzica nie mają części wspólnych.
- 4. Każde zero funkcji f znajdujące się w pewnym przedziale znajduje się w także w pewnym jego przedziale potomnym.
- 5. Każdy liść Ljest jedynym przedziałem na ścieżce od korzenia do L,który ma długość mniejszą niż 2ε

Wynikiem działania metody jest zbiór środków niepustych przedziałów znajdujących się w liściach drzewa przedziałów.

!! Rysunek drzewa przedziałów w bisekcji przy dwóch podprzedziałach, rysunek drzewa przedziałów zdegenerowanego do listy !! Podpis do rysunku: W wypadku 1 zera drzewo moze sie degenerowac do listy, ktora mozna traktowac jako ciąg z poprzedniego schematu.

Przedstawiony został w powyższym schemacie pewien zbiór metod znajdowania wielu zer funkcji f. Metody w tym zbiorze różnią się realizacją funkcji generującej drzewo przedziałów.

Przykładem metody szukania wielu zer funkcji jest modyfikacja opisanej w poprzednim podrozdziale metody bisekcji. W przypadku gdy w obu połowach

dzielonego przedziału jest możliwość wystąpienia zera, obie połowy są określane jako potomkowie przedziału.

2.3. Szybkość metody znajdowania zer

Szybkość znajdowania zer jest jedną z miar jakości metody. Możemy do pomiaru szybkości użyć komputera. Można zaprogramować metodę szukania zer w pewnym ustalonym języku programowania, za pomocą ustalonego kompilatora i uruchomić ją na komputerze z ustalonym procesorem, ustaloną hierarchią pamięci, na ustalonym systemie operacyjnym, na którym pracują równolegle ustalone programy i podać na wejście programowi ustaloną funkcję, której zer szukamy i ustaloną dokładność. Czas, który upłynął od podania danych do otrzymania wyniku można traktować jako szybkość działania metody. To podejście ma pewne wady, takie jak testowanie metody na skończonej liczbie funkcji oraz brak możliwości powtórzenia eksperymentu w dokładnie tych samych warunkach. Jednak jest to pomiar łatwy do wykonania i dający pewne wskazówki co do szybkości działania metody.

Zamiast mierzenia czasu możemy zmierzyć liczbę operacji zmiennoprzecinkowych, które wykonano znajdując zera konkretnej funkcji. To podejście wyklucza wiele czynników, biorących udział w poprzednim pomiarze. Przykładowo, pomijane są realizacje operacji przydziału i zwalniania pamięci, wczytywania danych i wypisywania wyników. Zyskujemy możliwość porównywania metod znajdowania zer przez uruchamianie ich na tych samych danych, lecz na innych komputerach. Jednakże, wciąż jesteśmy ograniczeni do testowania na skończonej liczbie funkcji.

Zostały wymyślone lepsze metody pomiaru. Możemy rozpatrzyć zależności między kolejnymi przedziałami dla dowolnej funkcji, analizując operację znajdowania następnika przedziału. Posługujemy się pojęciem rzędu zbieżności, czyli miarą tego, ile krótszy jest następny wygenerowany przedział od poprzedniego. Oznaczając długość k-tego przedziału w ciągu jako e_k metoda ma rząd zbieżności p, gdy:

$$e_k^p \sim e_{k+1}$$

Biorąc pod uwagę, że:

- przedziały nie mogą mieć ujemnej długości
- zależność może się stabilizować dopiero od pewnego kroku
- analizowana metoda powinna być zbieżna

metoda ma rzad zbieżności 1, gdy:

$$\lim_{k \to \inf} \frac{e_{k+1}}{e_k} = u \qquad \quad 0 < u < 1$$

oraz rząd zbieżności p dla p > 1, gdy:

$$\lim_{k \to \inf} \frac{e_{k+1}}{e_k^p} = u \qquad u > 0$$

Przykładowo bisekcja ma rząd zbieżności 1, gdyż następny przedział jest połową poprzedniego:

$$\begin{cases} e_{k+1} = \frac{1}{2}e_k \\ \lim_{k \to \inf} \frac{e_{k+1}}{e_k} = \frac{1}{2} \end{cases}$$
 (1)

3. Znajdowanie zer wielomianów

3.1. Zera wielomianów niskiego stopnia

Do szukania zer wielomianów niskich stopni można użyć wzorów analitycznych. Wielomian stopnia pierwszego ma jeden pierwiastek, który można otrzymać za pomocą jednego dzielenia i zmiany znaku. Wielomian stopnia drugiego ma zero, jeden lub dwa pierwiastki, które można otrzymać wykonując proste operacje arytmetyczne i obliczając pierwiastek kwadratowy. Wielomian stopnia trzeciego ma jeden, dwa lub trzy pierwiastki, które można otrzymać za pomocą prostych działań arytmetycznych i obliczania pierwiastków drugiego i trzeciego stopnia.

Obliczenie pierwiastka n-tego stopnia z liczby a wymaga rozwiązania równania $x^n=a$. Dlatego mimo istnienia analitycznych wzorów na zera wielomianów stopnia 2 i 3, obliczenie ich wymaga zastosowania metod numerycznych. Do obliczenia pierwiastka można użyć wyżej opisanej bisekcji, jednak lepsze rezultaty daje skorzystanie z metody Newtona.

3.2. Zera wielomianów dowolnego stopnia

Do znalezienia zer wielomianu stopnia wyższego niż 5 należy użyć metod numerycznych. Tych zer nie da się zapisać za pomocą skończonej liczby prostych operacji arytmetycznych i pierwiastkowania. Znajdowanie zer wielomianów nie jest jednak tak ogólnym zadaniem jak znajdowanie zer funkcji ciągłych. Możemy skorzystać z wielu własności charakterystycznych dla wielomianów. Własności te, takie jak ciągłość, ograniczenie liczby zer rzeczywistych przez stopień wielomianu, łatwość obliczania pochodnej dają narzędzia do konstrukcji specyficznych dla wielomianów metod szukania zer.

4. Wielomiany w postaci Beziera

Zbiór wielomianów stopnia nie większego niż n jest przestrzenią liniową o wymiarze n. Każda przestrzeń liniowa ma nieskończoną liczbę baz. Jedną z baz wielomianów jest baza potęgowa $1, x, x^2, x^3, ..., x^n$. Inną bazą jest baza Bernsteina $\{B_0^n(x), ..., B_n^n(x)\}$, gdzie $B_i^n(x) = \binom{n}{k} x^i (1-x)^i$. Wielomian wyrażony w bazie Bernsteina nazywany jest wielomianem w postaci Beziera. Każdy wielomian

w bazie potęgowej $P(x) = \sum_{i=0}^n p_i x^i$ można przedstawić w bazie Bernsteina $P(x) = \sum_{i=0}^n b_i B_i^n(x)$. Zależność między wektorem $p = \langle p_0, ..., p_n \rangle$ a wektorem $b = \langle b_0, ..., b_n \rangle$ w formie $p[m_{ij}] = b$ nazywana jest macierzą konwersji albo macierzą przejścia.

Postać Beziera ma szereg innych własności niż postać potęgowa. W jednych zastosowaniach sprawdza się postać potęgowa, w innych postać Beziera, w innych sprawdzają się jeszcze inne bazy. Przykładowym zastosowaniem wielomianów w postaci Beziera jest grafika komputerowa. Zgodnie z twierdzeniem Weierstrassa (udowodnionym zresztą przy pomocy wielomianów w bazie Bernsteina) krzywa parametryczna opisana parą wielomianów może odwzorowywać dowolny kształt ciągły z dowolną dokładnością. Specjalne krzywe wielomianowe, których wielomiany są opisane w postaci Beziera umożliwiają łatwe kształtowanie krzywej.

Niech para wielomianów $X(t) = \sum_{i=0}^n x_i B_i^n(t)$ oraz $Y(t) = \sum_{i=0}^n y_i B_i^n(t)$ dla $0 \le t \le 1$ opisuje krzywą wielomianową. Punkty $\langle (x_0, y_0), ..., (x_n, y_n) \rangle$ nazywane są punktami kontrolnymi krzywej. Można zaobserwować wizualny związek między kształtem krzywej a położeniem punktów kontrolnych.

!! rysunek wielomianu w postaci Beziera wraz z punktami kontrolnymi !! rysunek krzywej !! rysunek powierzchni

5. Znajdowanie zer wielomianu w postaci Beziera

Szukanie przecięć prostej z krzywą opisaną parą wielomianów jest zadaniem, które sprowadza się do znajdowania zer wielomianu. W przypadku, gdy krzywa jest opisana wielomianami w formie Beziera to zadanie sprowadza się do szukania zer wielomianu w formie Beziera.

Jednym ze sposobów rozwiązania tego zadania jest konwersja wielomianu z postaci Beziera do postaci potęgowej. Istnieją stabilne numerycznie i efektywne (liniowe w stosunku do stopnia wielomianu) algorytmy konwersji i szukania zer wielomianów w formie potęgowej. Jednakże, specyficzne własności wielomianów w postaci Beziera pozwalają na konstrukcję efektywnych i stabilnych numerycznie algorytmów operujących na wielomianach w postaci Beziera bez dodatkowego kroku, jakim jest konwersja baz.

Ważną własnością wielomianów w postaci Beziera jest własność otoczki wypukłej. W przedziale $\langle 0,1 \rangle$ wielomiany bazowe $\{B_0^n,B_1^n,...,B_n^n\}$ są nieujemne i sumują się do jedynki. Oznacza to, że zbiór wartości wielomianu w przedziale $\langle 0,1 \rangle$ jest ograniczony od dołu przez najmniejszy współczynnik wielomianu i ograniczony od góry przez największy współczynnik wielomianu.

Aby skorzystać z własności otoczki wypukłej dla pewnego przedziału P będącego podprzedziałem $\langle 0,1 \rangle$ należy przeskalować dziedzinę wielomianu tak, aby przedział P stał się nowym przedziałem $\langle 0,1 \rangle$. Aby to zrobić, moglibyśmy obliczyć wartość wielomianu w n+1 punktach przedziału P, następnie przeskalować te punkty do przedziału $\langle 0,1 \rangle$ i wykonać interpolację wielomianową na

tych punktach.

Jednakże istnieje prostszy sposób specyficzny dla wielomianów w postaci Beziera. Algorytm de Casteljau, który jest algorytmem obliczania wartości wielomianu P(x) w postaci Beziera dla argumentu t, poza wartością wielomianu w tym punkcie wyznacza także współczynniki wielomianów równych wielomianowi P(x) w przedziałach $\langle 0,t\rangle$ i $\langle t,1\rangle$. Wykonanie algorytmu de Casteljau w obu krańcach przedziału P daje współczynniki wielomianu z przeskalowanym przedziałem P do przedziału $\langle 0,1\rangle$.

Ekstrapolując własność otoczki wypukłej na dwa wymiary możemy zaprojektować metodę znajdowania zer wielomianu w postaci Beziera. W dwóch wymiarach wykres krzywej Beziera znajduje się w otoczce wypukłej punktów kontrolnych. Możemy skorzystać z tej własności dla wielomianów w postaci Beziera, traktując wielomian jako krzywą i wyznaczając otoczkę wypukłą tej krzywej.

!! rysunek otoczki wypulkej wielomianu

Zgodnie ze schematem metody znajdowania wielu zer funkcji, musimy jedynie zdefiniować operację znajdowania potomków przedziału w drzewie przedziałów. Niech dany będzie przedział P będący podprzedziałem $\langle 0,1 \rangle$. Potomkiem przedziału P jest przedział powstały przez znalezienie przecięcia otoczki wypukłej wielomianu w przedziale P z osią 0X. W przypadku, gdy znalezione przecięcie nie jest przynajmniej dwukrotnie krótsze od przedziału P, dzielimy przecięcie na dwie połowy i jako potomków przedziału określamy obie połowy.

Rząd zbieżności tej metody wynosi 2.

!! rysunek bezclip

6. Aproksymacja wielomianu wielomianem niższego stopnia

Aproksymacja to utworzenie uproszczonego modelu, który daje pewne informacje o oryginale i jest łatwiejszy do przetworzenia.

Na zadanie aproksymacji wielomianu wielomianem niższego stopnia można spojrzeć jak na zadanie geometryczne. Do ilustracji geometrycznej natury zadania aproksymacji użyjemy przykładu.

Chcemy aproksymować pewien element przestrzeni wektorowej \Re^2 elementem pewnej prostej $S = \{(x, ax) \in \Re^2 : x \in \Re\}.$

!! rysunek takiej prostej

Bazą \Re^2 może być zbiór wektorów $\{\langle 0,1\rangle,\langle 1,0\rangle\}$ a bazą S może być $\psi=\{\langle 1,a+b\rangle\}$. Każdy element przestrzeni wektorowej jest kombinacją liniową elementów bazowych.

Mając element $f \in \Re^2$ chcemy znaleźć najbliższy do niego element $g \in S$. Wektor różnicy f-g jest najkrótszy spośród wszystkich wektorów $f-s, s \in S$. Geometrycznie, wektor f-g jest najkrótszy, gdy f-g jest prostopadły do S.

!! rysunek wektora błędu

Wektor jest prostopadły do przestrzeni wektorowej, gdy jest prostopadły do każdego wektora bazowego tej przestrzeni. Wektory są prostopadłe do siebie, gdy

ich iloczyn skalarny jest równy zero. Sma tylko jeden wektor bazowy, wystarczy więc rozwiązać równanie:

$$(\psi, f - g) = 0$$

Przedstawiając fig jako kombinacje wektorów bazowych odpowiednich przestrzeni wektorowych

$$f = \langle 0, 1 \rangle f_y + \langle 1, 0 \rangle f_x$$
$$g = g_{len} \psi$$

mamy:

$$(\psi, \langle 1, 0 \rangle f_x + \langle 0, 1 \rangle f_y - g_{len} \psi) = 0$$

Korzystajac z wlasnosci iloczynu skalarnego:

$$(\psi, \langle 1, 0 \rangle f_x + \langle 0, 1 \rangle f_y) = (\psi, g_{len}\psi)$$

$$(\psi, \langle 1, 0 \rangle f_x) + (\psi, \langle 0, 1 \rangle f_y) = (\psi, g_{len}\psi)$$

$$(\psi, \langle 1, 0 \rangle) f_x + (\psi, \langle 0, 1 \rangle) f_y = (\psi, \psi) g_{len}$$

$$g_{len} = (\psi, \langle 1, 0 \rangle) f_x + (\psi, \langle 0, 1 \rangle) f_y$$

Ostatecznie:

$$g = \psi((\psi, \langle 1, 0 \rangle) f_x + (\psi, \langle 0, 1 \rangle) f_y)$$

Do rozwiązania powyższego zadania wystarczy więc obliczyć iloczyny skalarne wektorów bazowych i wykonać proste operacje arytmetyczne.

Uogólniając powyższe zadanie aproksymujemy element f przestrzeni wektorowej L_n o bazie $\{\psi_0,...,\psi_n\}$ elementem g przestrzeni wektorowej L_m o bazie $\{\gamma_0,...,\gamma_m\}$, gdzie n>m. Wtedy warunki prostopadłości wektora błędu f-g do L_m :

$$(f - g, \gamma_k) = 0$$
 $(k = 0, 1, ..., m)$

gdzie:

$$f = \sum_{i=0}^{n} f_i \psi_i$$

$$g = \sum_{i=0}^{n} g_i \gamma_i$$

tworzą układ równań:

$$\sum_{i=0}^{n} (\psi_i, \gamma_k) f_i = \sum_{i=0}^{m} (\gamma_i, \gamma_k) g_i \qquad (k = 0, 1, ..., m)$$

Gdy baza przestrzeni L_m jest ortonormalna, czyli każdy wektor bazowy jest prostopadły do dowolnego innego wektora bazowego i długość każdego wektora wynosi 1, to układ redukuje sie do:

$$g_i = \sum_{i=0}^{n} (\psi_i, \gamma_k) f_i$$
 $(k = 0, 1, ..., m)$

Wielkości (ψ_i, γ_k) dla (i = 0, 1, ..., n; k = 0, 1, ..., m) można zapisać w macierzy. Po obliczeniu macierzy aproksymacji wektor aproksymujący otrzymuje się mnożąc wektor aproksymowany przez macierz aproksymacji.

Do rozwiązania zadania wielomianowej aproksymacji średniokwadratowej jest wykorzystywany podobny mechanizm. Przestrzeń wielomianów Π_n w przedziale $\langle a,b \rangle$ jest przestrzenią wektorową o wymiarze n. Wystarczy zdefiniować iloczyn skalarny wielomianów:

$$(f,g) = \int_{a}^{b} p(x)f(x)g(x)dx$$

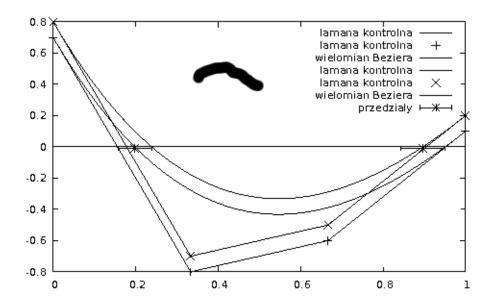
gdzie funkcja wagowa p(x) jest determinowana przez przedział $\langle a,b \rangle$. Tak zdefiniowany iloczyn skalarny spełnia własności iloczynu skalarnego takie jak przemienność czy liniowość. Rozumowania przeprowadzone dla aproksymacji wektorów mogą zostać powtórzone dla aproksymacji wielomianowej.

W niniejszej pracy szczególnie interesuje nas aproksymacja wielomianów w postaci Beziera innymi wielomianami w postaci Beziera. Jednakże wielomiany bazowe Bernsteina nie są ortonormalne według tak zdefiniowanego iloczynu skalarnego, przez co układ równań dla warunku prostopadłości nie może zostać zredukowany do mnożenia wektora współczynników wielomianu przez macierz aproksymacji.

Prostym sposobem rozwiązania tego problemu jest rozwiązanie układu równań jedną z metod rozwiązywania układów równań liniowych, np. metodą eliminacji Gaussa. Jest to metoda kosztowna, poza tym rozwiązywanie niektórych układów równań liniowych jest źle uwarunkowane.

Innym sposobem rozwiązania tego problemu jest szukanie wielomianu aproksymującego w bazie ortonormalnej i późniejsza konwersja do postaci Beziera. W przypadku zastosowania tej metody należy także przeprowadzić konwersję wielomianu aproksymowanego do tej samej bazy w celu obliczenia iloczynu skalarnego.

Istnieją lepsze sposoby rozwiązania tego problemu. Bezpośrednia zależność między współczynnikami aproksymowanego wielomianu w postaci Beziera a współczynnikami wielomianu aproksymującego została opracowana przez Juttlera. Ta zależność zapisana w formie macierzy pozwala na otrzymanie wielomianu aproksymującego za pomocą tylko jednego mnożenia wektora współczynników przez macierz. Efektywną metodą obliczania tej macierzy jest metoda Lewanowicza-Woźnego.



7. Przycinanie sześcienne

Przycinanie sześcienne to efektywna metoda znajdowania wielu zer wielomianu w postaci Beziera. Wykorzystuje ona wiele własności wielomianu w postaci Beziera opisanych w niniejszej pracy.

Opiszmy operację znajdowania potomków przedziału dla metody przycinania sześciennego.

Wejście:

Wielomian Q(x), którego zer szukamy

Przedział P, którego potomków szukamy

Algorytm:

 $R^3(x) \leftarrow$ aproksymacja Q(x) wielomianem stopnia trzeciego w przedziale P

 $R^n(x) \leftarrow \text{podniesienie stopnia wielomianu } R^3(x) \text{ do stopnia } n$

 $D(x) \leftarrow Q(x) - R^n(x)$

 $d \leftarrow$ górne ograniczenie maksymalnej wartości absolutnej wielomianu D(x) w przedziałe P

 $U(x) \leftarrow R^3(x) + d$

 $L(x) \leftarrow R^3(x) - d$

 $potomkowieP \leftarrow \text{podprzedziały } P$ w których U(x)jest dodatni a L(x)jest ujemny

Zauważmy, że wykres Q(x) w przedziale P, a w szczególności zera Q(x) w przedziale P zawierają się pomiędzy wykresami wielomianów L(x) i U(x).

Aby zastosować jedną macierz aproksymacji należy każdy aproksymowany przedział przeskalować do przedziału $\langle 0,1\rangle.$

Rząd zbieżności przycinania sześciennego wynosi 4.