Autor: Miron Hunia

Prowadzący: Miron Hunia

Komputery kwantowe. Algorytm Grovera.

Szybki kurs algebry liniowej

Algebra liniowa bada wektory i funkcje liniowe. Jest kluczowa w opisie komputerów kwantowych

Definicja 1 (Ciało)**.** Ciałem nazywamy zbiór K, na którym możemy dodawać, odejmować, mnożyć i dzielić (nie przez 0) dowolne elementy, i te operacje spełniają arytmetyczne własności do których jesteśmy przyzwyczajeni.

Przykładami ciał są: $\mathbb{R}, \mathbb{Q}, \mathbb{C}, \mathbb{Z}_p$.
Przykładami zbiorów, które NIE są ciałami są: $\mathbb{Z}$ (bo nie można dzielić), $\mathbb{Z}_{10}$ (bo nie można dzielić przez 2 ani 5), $\mathbb{H}$ (bo mnożenie nie jest przemienne).

W sercu algebry liniowej jest poniższa definicja:

Definicja 2 (Przestrzeń liniowa)**.** Przestrzenią liniową nad ciałem K (nazywanym **ciałem skalarów**, jego elementy nazywamy **skalarami**) nazywamy zbiór V, którego elementy możemy dodawać, odejmować, oraz mnożyć przez skalary, i te operacje spełniają arytmetyczne własności do których jesteśmy przyzwyczajeni.

Przykładami przestrzeni liniowych są:

- każde ciało
- krotki liczb, np. (x, y, z).
- Wielomiany, np. $\mathbb{R}[x]$.
- $\mathbb{R}$ (jako przestrzeń liniowa nad $\mathbb{Q}$)
- Funkcje ciągłe $f : \mathbb{R} \to \mathbb{R}$

Elementy przestrzeni liniowej nazywamy *wektorami*. Na tym wykładzie wektory będę oznaczał w klamerkach: $|x\rangle$ (nazywa się to notacją Diraca).
Jeśli $n > 0$ i $a_1, \ldots, a_n$ to skalary, a $|x_1\rangle, \ldots, |x_n\rangle$ to wektory, to wyrażenie

$$a_1 |x_1\rangle + \cdots + a_n |x_n\rangle$$

nazywamy kombinacją liniową wektorów $|x_1\rangle, \ldots, |x_n\rangle$.
Zbiór wszystkich wektorów, które są kombinacjami liniowymi wektorów $|x_1\rangle, \ldots, |x_n\rangle$ nazywamy podprzestrzenią rozpiętą przez $|x_1\rangle, \ldots, |x_n\rangle$ i oznaczamy ją $\text{span}(|x_1\rangle, \ldots, |x_n\rangle)$.
Jeśli istnieje przedstawienie wektora 0 jako nietrywialna kombinacja liniowa (tj. taka, gdzie nie wszystkie skalary są zerami), to wektory nazywamy *liniowo zależnymi*. W przeciwnym przypadku nazywamy je *liniowo niezależnymi*.

Baza i wymiar

Niech V jest przestrzenią liniową. Zbiór wektorów $\mathcal{B} = \{|x_1\rangle, \ldots, |x_n\rangle\}$ nazywamy bazą przestrzeni V, jeśli są liniowo niezależne i $\text{span}(\mathcal{B}) = V$.

1. Baza zawsze istnieje.
2. Zazwyczaj dana przestrzeń liniowa ma wiele różnych baz.
3. Każda baza przestrzeni V ma tyle samo elementów.
4. Każdy element V można przedstawić jako kombinację liniową elementów bazy na dokładnie 1 sposób.

Rozmiar bazy nazywamy *wymiarem* V i oznaczamy $\dim V$.
Punkt 4. z powyższej listy daje nam jako wniosek, że dowolną przestrzeń liniową możemy sprowadzić poprzez wybór bazy do systemu współrzędnych $(a_1, \ldots, a_n) \in K^n$, gdzie $n = \dim V$.

Przekształcenia liniowe

$\phi : V \to W$ (gdzie V, W to przestrzenie liniowe) nazywamy przekształceniem liniowym, jeśli respektuje kombinacje liniowe, czyli

$$\phi(a_1 |x_1\rangle + \cdots + a_n |x_n\rangle) = a_1\phi(|x_1\rangle) + \ldots a_n\phi(|x_n\rangle)$$

Przekształcenia liniowe tworzą przestrzeń liniową. Jej wymiar to $\dim V \cdot \dim W$. W takim razie przekształcenie liniowe możemy unikalnie zapisać w postaci tabelki nxm, nazywanej *macierzą*. Wektory możemy traktować jako macierze z jedną kolumną.
Złożenie przekształceń liniowych też jest przekształceniem liniowym i odpowiada mnożeniu macierzy.
Przekształcenia liniowe $\phi : V \to K$ (czyli takie, których zbiór wartości jest jednowymiarowy) nazywamy *funkcjonałami*. Funkcjonały możemy traktować jako macierze z jednym wierszem.

Prostopadłość i iloczyn hermitowski

Jeśli pracujemy nad ciałem $\mathbb{R}$, to iloczynem skalarnym wektorów $v = (a_1, \ldots, a_n)$ i $w = (b_1, \ldots, b_n)$ nazywamy $\langle v|w\rangle = a_1b_1 + \cdots + a_nb_n$. W szczególności wielkość $\langle v|v\rangle = |v|^2$ to długość wektora v.
Zauważmy, że w ten sposób wektor v zadaje funkcjonał, który posyła $w \mapsto \langle v|w\rangle$. Taki funkcjonał oznaczamy $\langle v|$ i mamy $\langle v| |w\rangle = \langle v|w\rangle$.
Jeśli pracujemy nad ciałem $\mathbb{C}$, to zamiast tego używamy iloczynu hermitowskiego, który jest zdefiniowany bardzo podobnie:

$$\langle v|w\rangle = a_1\overline{b_1} + \cdots + a_n\overline{b_n}$$

W szczególności $\langle w|v\rangle = \overline{\langle v|w\rangle}$.
Dzięki temu wielkość $|v| = \sqrt{\langle v|v\rangle}$ jest zawsze rzeczywista dodatnia, nawet gdy współrzędne wektorów są zespolone.
Wielkość $\langle v|w\rangle$ geometrycznie mierzy nam, jak mały jest kąt α pomiędzy v i w: $\langle v|w\rangle = |v| \cdot |w| \cos\alpha$. W szczególności jeśli v i w wskazują w tym samym kierunku, to $\langle v|w\rangle = |v| \cdot |w|$. Z drugiej strony, jeśli $|v\rangle$ i $|w\rangle$ są prostopadłe, to $\langle v|w\rangle = 0$. W abstrakcyjnych przestrzeniach liniowych w taki właśnie sposób definiuje się prostopadłość i ogólniej kąty pomiędzy wektorami.

Macierze unitarne

Przekształcenia, które zachowują iloczyn hermitowski nazywamy unitarnymi. Innymi słowy, macierz U jest unitarna jeśli dla dowolnych $|x\rangle, |y\rangle$ mamy $\langle x|y\rangle = \langle Ux|Uy\rangle$. Ten warunek możemy przepisać na bardziej kompaktowy: U jest unitarna wtedy i tylko wtedy, gdy

$$U^\dagger U = I$$

gdzie I to macierz przekształcenia identycznościowego i

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}^{\dagger} = \begin{pmatrix} \overline{a_{11}} & \cdots & \overline{a_{n1}} \\ \vdots & \ddots & \vdots \\ \overline{a_{1n}} & \cdots & \overline{a_{nn}} \end{pmatrix}$$

Przekształcenia unitarne zachowują kąty i długości wektorów, więc możemy o nich myśleć, jak o obrotach i symetriach (w odpowiednio więcej wymiarowej przestrzeni).

Komputer kwantowy

Uporawszy się z całą tą teorią, możemy zacząć gadać o modelu komputera kwantowego. Jest to model abstrakcyjny i pomija kwestie fizyczne związane z faktyczną budową komputera kwantowego.
W kontraście do komputera kwantowego, komputery których używamy na co dzień nazywamy komputerami klasycznymi. Te dwa byty są od siebie fundamentalnie różne - komputery kwantowe **nie są** po prostu komputerami z większą mocą obliczeniową. Ich możliwości i algorytmy są fundamentalnie inne.

Kubit

Kubit to kwantowy odpowiednik bita. Tak jak w komputerze klasycznym, kubit może przyjmować stany $|0\rangle, |1\rangle$. Jednak w przeciwieństwie do klasycznych bitów, kubit może być w stanie *superpozycji*, czyli kombinacji liniowej tych stanów bazowych.

$$|\psi\rangle = a_0 |0\rangle + a_1 |1\rangle$$

Jeśli kubit jest w stanie superpozycji i dokonujemy jego pomiaru, to nadal otrzymamy jeden ze stanów $|0\rangle$ lub $|1\rangle$, odpowiednio z prawdopodobieństwem $|a_0|^2$ lub $|a_1|^2$. Jak się okazuje, współrzędne a_0 i a_1 mogą być liczbami zespolonymi.
W takim razie stan kubitu jest opisywany poprzez dwuwymiarową przestrzeń liniową nad $\mathbb{C}$. Co więcej, ponieważ prawdopodobieństwa sumują się do 1, to musi zachodzić $|a_0|^2 + |a_1|^2 = 1$, czyli możliwe stany kubita są sferą w tej przestrzeni.
Jak wiemy z algebry liniowej, wybór bazy przestrzeni jest dość dowolny, więc możemy się spodziewać, że możemy dokonywać również pomiarów innych stanów, w rodzaju $|\xi\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$. Możemy uogólnić nasz wzór na prawdopodobieństwo:

Prawdopodobieństwo, że kubit po pomiarze będzie w stanie $\xi = \langle\xi|\psi\rangle \langle\psi|\xi\rangle$

Wiele kubitów

Na komputerze klasycznym, jeśli rejestr (czyli ciąg bitów) jest powiedzmy w stanach $|1\rangle, |0\rangle, |1\rangle$, to możemy powiedzieć, że stan rejestru to $|101\rangle$, który interpretujemy jako liczba binarna $|5\rangle$.
Na komputerze kwantowym, jeśli kubity są w stanach $|1\rangle, |0\rangle, \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$, to na intuicję, stan rejestru powinien wynosić $\frac{1}{\sqrt{2}}(|100\rangle + |101\rangle)$. Tak rzeczywiście jest, i tą intuicję możemy sformalizować używając iloczynu tensorowego $\otimes$. Stan rejestru składającego się z dwóch kubitów w stanach odpowiednio $|\psi\rangle$ i $|\xi\rangle$ to

$$|\psi\rangle \otimes |\xi\rangle = (a_0 |0\rangle + a_1 |1\rangle) \otimes (b_0 |0\rangle + b_1 |1\rangle) = a_0 b_0 |0\rangle \otimes |0\rangle + a_0 b_1 |0\rangle \otimes |1\rangle + a_1 b_0 |1\rangle \otimes |0\rangle + a_1 b_1 |0\rangle \otimes |1\rangle$$

Stany w rodzaju $|0\rangle \otimes |0\rangle$ zapisujemy skrótowo jako $|00\rangle$.
Przyjrzyjmy się teraz stanowi rejestru.

$$\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$$

Ten stan rejestru jest podejrzany, bowiem nie da się go wyfaktowyzować na iloczyn tensorowy stanów składowych kubitów (dlaczego?). Więcej o tym dalej.

Autor: Miron Hunia

Prowadzący: Miron Hunia

Bramki kwantowe

W komputerze klasycznym operacje są wykonywane poprzez bramki logiczne, na przykład bramkę NOT, która posyła 0 na 1 i na odwrót. Jak byśmy się spodziewali, że bramka NOT zadziała na kubicie w stanie $\frac{1}{2}|0\rangle+\frac{\sqrt{3}}{2}|1\rangle$? Jeśli bramka się zaaplikuje do każdego stanu z osobna, to powinniśmy dostać $\frac{1}{2}|1\rangle+\frac{\sqrt{3}}{2}|0\rangle$. Tak faktycznie jest. Matematycznie możemy to ująć, że bramki kwantowe są przekształceniami liniowymi.
Dodatkowo wiemy, że bramka kwantowa U powinna zachowywać długość wektorów (bo stany kwantowe leżą na sferze jednostkowej). W takim razie U musi być przekształceniem unitarnym.
Uwaga. To oznacza w szczególności, że U musi być odwracalne, a to oznacza, że niektóre dość proste bramki znane z komputerów klasycznych, w rodzaju przypisanie bitu na 1, nie są możliwe na komputerze kwantowym. Ten problem rozwiązuje się przez dodanie do rejestru dodatkowych kubitów pomocniczych, i zamiast robić przypisanie, robi się $SWAP$ na dwóch kubitach.

Brama	Reprezentacja Macierzowa	Opis
X (NOT)	$\begin{pmatrix}0&1\\1&0\end{pmatrix}$	Zamienia stany $\|0\rangle$ i $\|1\rangle$
Z	$\begin{pmatrix}1&0\\0&-1\end{pmatrix}$	Zmienia fazę stanu $\|1\rangle$
H (Hadamard)	$\frac{1}{\sqrt{2}}\begin{pmatrix}1&1\\1&-1\end{pmatrix}$	Wprowadza równą superpozycję stanów $\|0\rangle$ i $\|1\rangle$
$S=\sqrt{Z}$	$\begin{pmatrix}1&0\\0&i\end{pmatrix}$	Pierwiastek kwadratowy z Z
$T=\sqrt{S}$	$\begin{pmatrix}1&0\\0&e^{i\pi/4}\end{pmatrix}$	Pierwiastek czwartego stopnia z Z
$CNOT$ (CX, Controlled Not)	$\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&0&1\\0&0&1&0\end{pmatrix}$	Kontrolowane NOT, działa na dwóch kubitach
$SWAP$	$\begin{pmatrix}1&0&0&0&0&0\\0&0&1&0&0&0\\0&1&0&0&0&0\\0&0&0&0&1&0\\0&0&0&1&0&0\\0&0&0&0&0&1\end{pmatrix}$	Zamienia dwa kubity
Toffoli ($CCNOT$)	$\begin{pmatrix}1&0&0&0&0&0&0&0\\0&1&0&0&0&0&0&0\\0&0&1&0&0&0&0&0\\0&0&0&1&0&0&0&0\\0&0&0&0&1&0&0&0\\0&0&0&0&0&1&0&0\\0&0&0&0&0&0&0&1\\0&0&0&0&0&0&1&0\end{pmatrix}$	Kontrolowany $CNOT$

Tablica 1: Podstawowe bramki kwantowe i ich reprezentacje macierzowe

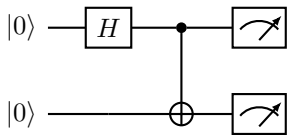
Bramek kwantowych też można brać iloczyn tensorowy. Na przykład bramka $H\otimes H=\frac{1}{2}\begin{pmatrix}1&1&1&1\\1&-1&1&-1\\1&1&-1&-1\\1&-1&-1&1\end{pmatrix}$ to bramka, której rezultat jest taki, jak zaaplikowanie H do pierwszego kubitu i H do drugiego kubitu.

Autor: Miron Hunia

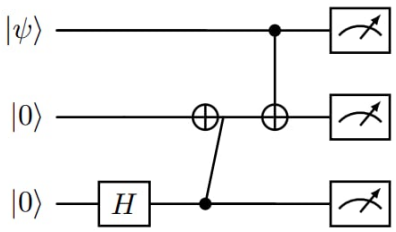
Prowadzący: Miron Hunia

Stany splątane i teleportacja kwantowa

Rozpatrzmy teraz następujący układ kwantowy.



Ten układ "wymnaża" się do bramki $(H \otimes I) \cdot CX = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & -1 \\ 1 & 0 & -1 & 0 \end{pmatrix}$. Jego pierwsza kolumna mówi nam, jaki stan dostaniemy na wyjściu, jeśli podamy mu na wejściu stan $|00\rangle$. Jak widać, będzie to $|\xi\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$. Chwila moment! Widzieliśmy już wcześniej taki stan i stwierdziliśmy, że nie da się go rozłożyć na iloczyn tensorowy stanów dwóch kubitów. Co to znaczy?
W takiej sytuacji stan kubitów pierwszego i drugiego nie jest określony indywidualnie, określony jest jedynie stan całego rejestru. W takich sytuacjach mówimy, że rejestr jest w stanie splątanym.
Zbadajmy teraz bramkę



Algorytmy kwantowe

Wyrocznia Deutsch-Jozsa

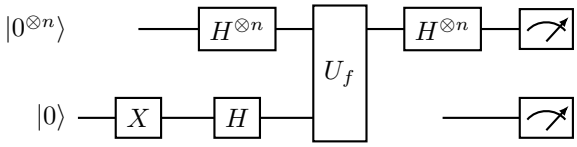
Przyjmijmy, że mamy daną funkcję $f : \{0,1\}^n \to \{0,1\}$, która spełnia jedną z dwóch własności:

1. $f(x) = 0$ dla każdego x lub $f(x) = 1$ dla każdego x (f jest stała)
2. Moc zbioru $\{x : f(x) = 0\}$ jest równa mocy zbioru $\{x : f(x) = 1\}$ (f jest zbalansowana)

Naszym celem jest stwierdzenie, czy f (podana na wejściu jako układ logiczny, w jakiejś czarnej skrzynce) jest stała, czy zbalansowana.
Na komputerze klasycznym oczywiście trzeba sprawdzić co najmniej $2^{n-1} + 1$ argumentów, żeby móc odpowiedzieć z pewnością. Okazuje się, że na komputerze kwantowym możemy uzyskać deterministyczne rozwiązanie, które aplikuje f jedynie raz. To daje wykładniczo lepszy rezultat!
Zdefiniujmy teraz problem bardziej formalnie, dla komputera kwantowego. Mamy podany operator U_f ("czarną skrzynkę") jako jakiś nieznany układ kwantowy, która zaaplikowanego do rejestru $|x\rangle |y\rangle$ zwraca $|x\rangle |y \oplus f(x)\rangle$, gdzie $\oplus$ oznacza xor bitowy (czyli dodawanie modulo 2 po współrzędnych).
Możemy wówczas użyć pomysłu, który w nazywany jest "phase kickback". Jeśli oznaczymy $\frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$ przez $|-\rangle$, to:

$$U_f |x\rangle |-\rangle = (-1)^{f(x)} |x\rangle |-\rangle$$

Wówczas zadanie jest rozwiązywane przez poniższy układ.



Przed zaaplikowaniem U_f rejestr jest w stanie $\frac{1}{\sqrt{2^n}}\sum_x |x\rangle \otimes |-\rangle$. Używając phase kickback, wiemy, że po zaaplikowaniu U_f stan wynosi

$$\frac{1}{\sqrt{2^n}}\left(\sum_x (-1)^{f(x)}|x\rangle\right)\otimes|-\rangle$$

Teraz możemy zignorować ostatni kubit. Aplikując jeszcze raz bramkę $H^{\otimes n}$ dostajemy

$$\frac{1}{2^n}\left(\sum_x\sum_z (-1)^{x\cdot z+f(x)}|z\rangle\right)$$

Skupiając się na samym stanie $|z\rangle = |0\rangle$ dostajemy, że współczynnik przy tym stanie wynosi

$$\frac{1}{2^n}\sum_x (-1)^{f(x)}$$

Ta suma wynosi 0 jeśli f jest zbalansowane. Jeśli f jest stałe, to ta suma wynosi 1 lub -1. Zatem jeśli zmierzony stan rejestru po zakończeniu algorytmu to 0 to wiemy, że f jest funkcją stałą, w przeciwnym razie możemy być pewni, że f jest stałe.

Algorytm Grovera

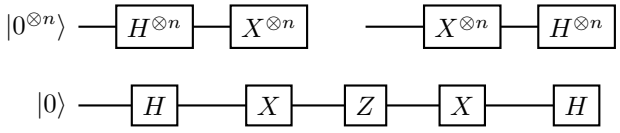
Gwóźdź programu. Tak samo jak algorytm Deutsch-Jozsy, algorytm Grovera jest algorytmem z wyrocznią. W przeciwieństwie do algorytmu Deutsch-Jozsy jednak, algorytm Grovera ma praktyczne zastosowania.
Sformułujmy problem. Mamy zaimplementowaną funkcję $f : \{0,1\}^n \to \{0,1\}$, do której dostęp mamy poprzez wyrocznię. Tym razem zakładamy, że $f(x)$ jest równe 1 w dokładnie jednym punkcie, nazwijmy go ω. Nasza wyrocznia jest określona przez $U_\omega|x\rangle = |x\rangle$ jeśli $x \neq \omega$ i $U_\omega|x\rangle = -|x\rangle$ jeśli $x = \omega$. Można to skrótowo zapisać jako

$$U_\omega|x\rangle = (-1)^{f(x)}|x\rangle$$

Algorytm Grovera opiera się na operatorze dyfuzji

$$U_s = 2|s\rangle\langle s| - I = \begin{pmatrix} \frac{1}{N} & \cdots & \frac{1}{N} \\ \vdots & \ddots & \vdots \\ \frac{1}{N} & \cdots & \frac{1}{N} \end{pmatrix}$$

gdzie $s = H^{\otimes n}|0^{\otimes n}\rangle$ jest wektorem w równej superpozycji wszystkich stanów. Możemy go zaimplementować jak poniżej.



Ten operator jest operatorem odbicia względem $|s\rangle$. W szczególności jeśli zainicjujemy rejestr w stanie $|s\rangle$, to wykonując operacje U_s i U_ω pozostaniemy w jednej płaszczyźnie, rozpiętej przez stany ortonormalne ω i $s' = \sqrt{\frac{N}{N-1}}(s - \frac{1}{\sqrt{N}}\omega)$. Wtedy U_ω jest operatorem odbicia względem s'. W takim razie ich złożenie U_sU_ω jest obrotem o kąt θ, który jest dwukrotnością kąta między s i s'. Możemy go policzyć iloczynem hermitowskim.

$$\cos\left(\frac{1}{2}\theta\right) = \langle s|s'\rangle = \langle\sqrt{\frac{N-1}{N}}s' + \frac{1}{\sqrt{N}}\omega|s'\rangle = \sqrt{\frac{N-1}{N}}\langle s'|s'\rangle + \langle\frac{1}{\sqrt{N}}\omega|s'\rangle = \sqrt{\frac{N-1}{N}} + 0 = \sqrt{\frac{N-1}{N}}$$

Stąd

$$\frac{\theta}{2} \approx \sin\left(\frac{\theta}{2}\right) = \sqrt{\frac{1}{N}}$$

Stąd widzimy, że aby wykonać obrót o $\frac{\pi}{2}$ wystarczy nam wykonać około $\frac{\pi\sqrt{N}}{4}$ obrotów.
Wylądujemy obok ω z dokładnością do $\frac{\omega}{2}$, co nam pozwala oszacować prawdopodobieństwo, że po pomiarze wyjdzie ω. Oznaczmy przez ψ nasz stan końcowy. Wówczas

$$\langle\psi|\omega\rangle \langle\omega|\psi\rangle \geq^2 \langle\psi|\psi\rangle \langle\omega|\omega\rangle \cos^2\left(\frac{\theta}{2}\right) \geq \left(1 - \frac{1}{N}\right)^2 \geq \frac{(N-1)^2}{N^2}$$