Dyskretny wzór Itô

*Michał MIŚKIEWICZ**

* Wydział Matematyki, Informatyki i Mechaniki UW

Wśród twierdzeń analizy matematycznej jest jedno szczególnie ważne, czasem zwane *podstawowym*. Opiera się ono na tożsamości

$$(\star) \qquad a_N - a_0 = (a_N - a_{N-1}) + \ldots + (a_2 - a_1) + (a_1 - a_0),$$

czyli obserwacji, że całkowity przyrost (tutaj: ciągu) jest sumą lokalnych przyrostów. Twierdzenie to stwierdza równość

$$f(T) - f(0) = \int\limits_0^T f'(t)\,\mathrm{d}t \qquad \text{dla „dobrych" funkcji } f,$$

ale jeśli Czytelnik nie zna pojęć pochodnej f' oraz całki $\int \ldots \mathrm{d}t$, to nic nie szkodzi. Intuicyjnie, pierwsze odpowiada lokalnym przyrostom (tutaj: funkcji), a drugie – sumowaniu wielu składników. Twierdzenie to można zresztą wywnioskować z równości $(\star)$ (zgodnie z uwagą na marginesie).

Dla każdego $N \in \mathbb{N}$ zachodzi równość

$$f(T) - f(0) = \sum_{n=0}^{N-1} \frac{f(\frac{n+1}{N}T) - f(\frac{n}{N}T)}{T/N} \cdot \frac{T}{N},$$

będąca szczególnym przypadkiem $(\star)$. Dla odpowiednio regularnych funkcji f tezę podstawowego twierdzenia analizy otrzymujemy teraz w granicy $N \to \infty$.

W analizie *stochastycznej*, a więc dziale analizy zajmującym się procesami losowymi, znane jest podobne twierdzenie pod nazwą wzoru Itô. Wzór ten określa zmienność funkcji $f(W_t)$, w której liniowo rosnący argument $0 \leqslant t \leqslant T$ zastąpiono losowym procesem, konkretnie procesem Wienera W_t. Proces ten jest też znany – zwłaszcza w naukach stosowanych – jako ruch Browna; służy on do opisu chaotycznego ruchu drobnych pyłków zawieszonych w cieczy, jak również zmian cen akcji na giełdzie (zob. Δ^4_{83}, Δ^5_{83}). O dziwo, wzór Itô fundamentalnie różni się od poprzedniego wzoru:

$$f(W_T) - f(W_0) = \int\limits_0^T f'(W_t)\,\mathrm{d}W_t + \frac{1}{2}\int\limits_0^T f''(W_t)\,\mathrm{d}t,$$

a więc oprócz „sumy lokalnych przyrostów" (czyli pierwszej z całek) zawiera też poprawkę drugiego rzędu.

Użyta we wzorze obok całka $\int \ldots \mathrm{d}W_t$ to *całka Itô*. Tak jak poprzednio, znajomość tego pojęcia, jak również procesu Wienera, nie ma decydującego znaczenia dla lektury dalszej części tekstu.

Jak to możliwe, że całkowity przyrost okazał się różny od całki, która miała wyrażać sumę lokalnych przyrostów? Czy da się to naprawić? A może z tego dziwnego fenomenu wynika coś pożytecznego? Na te pytania postaram się odpowiedzieć, sięgając po model najprostszy z możliwych – dyskretny.

Krótko o pochodnych dyskretnych. Słowo *dyskretny* oznacza w tym kontekście, że zamiast funkcji określonych dla wszystkich liczb rzeczywistych będziemy tu rozważali funkcję f określoną jedynie w całkowitych wielokrotnościach pewnej ustalonej liczby $h > 0$, a więc $0, h, -h, 2h, -2h$ etc. O różnicy między światem ciągłym a światem dyskretnym warto przeczytać więcej w artykule *Punkt, odcinek, trójkąt...* z Δ^9_{76}.

Klasyczna definicja pochodnej oparta jest na badaniu ilorazu różnicowego $\frac{f(x+d)-f(x)}{d}$ dla małych $d \neq 0$. U nas jednak d nie może być dowolnie małe. Najbliższy zeru jest wybór $d = \pm h$, co daje dwóch potencjalnych kandydatów na *dyskretną pochodną*: $\frac{f(x+h)-f(x)}{h}$ oraz $\frac{f(x)-f(x-h)}{h}$. Nie widać jednak, dlaczego któryś z kierunków prawo–lewo miałby być wyróżniony, dlatego przyjmiemy symetryczną uśrednioną definicję:

$$Df(x) := \frac{1}{2}\left(\frac{f(x+h)-f(x)}{h} + \frac{f(x)-f(x-h)}{h}\right) = \frac{f(x+h)-f(x-h)}{2h}.$$

Drugą pochodną natomiast wprowadza się jako pochodną pochodnej. W naszym przypadku będzie więc naturalne rozważenie ilorazu różnicowego ilorazów różnicowych:

$$D^2 f(x) := \frac{1}{h}\left(\frac{f(x+h)-f(x)}{h} - \frac{f(x)-f(x-h)}{h}\right) = \frac{f(x+h)+f(x-h)-2f(x)}{h^2}.$$

O błądzeniu losowym – dyskretnym krewnym W_t. Funkcję f zastosujemy do procesu losowego zwanego *błądzeniem losowym*. Nazwa ta dobrze oddaje, o co chodzi: startujemy z punktu 0 i wielokrotnie rzucamy monetą, zależnie od wyniku poruszając się w lewo lub w prawo. Przyjmiemy, że kolejne kroki mają długość h każdy – dzięki czemu zawsze trafiamy w punkt określoności f – i że następują one w odstępach czasowych h^2. Pozycję w chwili t oznaczymy przez S_t; rozkład prawdopodobieństwa S_t dla $t = 0, h^2, 2h^2$ ilustrują wykresy

Rozwiązanie zadania F 1082.
Powietrze spełnia równania dwuatomowego gazu doskonałego. Dla n moli takiego gazu wypełniającego objętość V pod ciśnieniem p, w temperaturze bezwzględnej T spełnione są zależności:

$$pV = nRT,$$
$$U = C_V nT,$$

przy czym dla gazu dwuatomowego molowe ciepło właściwe $C_V = 5R/2$ (ciepło właściwe w stałej objętości).
a) Jeżeli podczas ogrzewania ciśnienie gazu p było stałe (oczywiście także objętość pokoju V nie zmieniała się), to stała pozostawała również energia wewnętrzna U gazu zawartego w pokoju.
b) Zmieniała się za to liczba moli gazu:

$$n(T) = \frac{pV}{RT}.$$

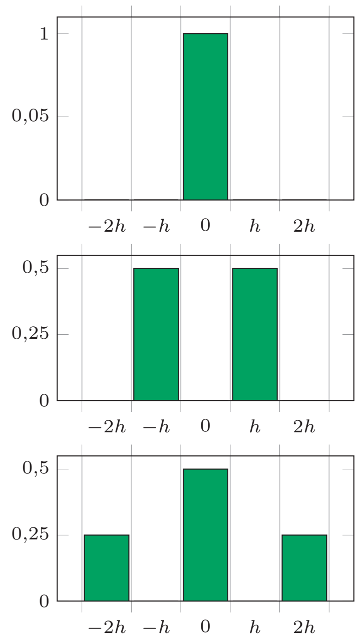
Liczba moli zmalała więc od $n_p = 2{,}643 \cdot 10^3$ do $n_k = 2{,}463 \cdot 10^3$.
c) Grzanie następowało pod stałym ciśnieniem p. Molowe ciepło właściwe gazu doskonałego pod stałym ciśnieniem $C_p = C_V + R$. Zwiększenie temperatury gazu w pokoju od T do $T + dT$ wymagało więc dostarczenia ciepła:

$$dQ = n(T)C_p dT = \frac{7}{2}R\frac{pV}{RT}dT.$$

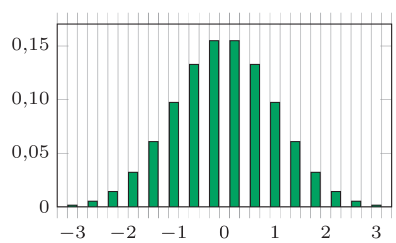
Podczas ogrzewania od $T_0 = 273{,}15$ K do $T_k = 293{,}15$ K grzejnik dostarczył ciepło:

$$Q = \frac{7pV}{2}\ln\left(\frac{T_k}{T_p}\right).$$

Po podstawieniu danych liczbowych otrzymujemy $Q = 1{,}484 \cdot 10^3$ kJ.
Założona równość ciśnień w pokoju i na zewnątrz odpowiada sytuacji typowej – okna i drzwi nie są na tyle szczelne, żeby mogła wystąpić różnica ciśnień z atmosferą na zewnątrz.



Powyżej rozkład prawdopodobieństwa S_t dla $t = 0, h^2, 2h^2$. A poniżej rozkład S_1 dla $h = \frac{1}{5}$ (czyli po 25 krokach), zaskakująco podobny do wykresu $y = e^{-x^2}$.



Warto się dowiedzieć, czym są martyngały i jak pożyteczna może być wiedza, że dany proces jest martyngałem:

Paweł Hitczenko, *Martyngały*, Δ^6_{88}.
Rafał Sztencel, *Martyngały*, Δ^5_{07}.
Rafał Sztencel, *Kontra d'Alemberta na twierdzenie Dooba*, Δ^6_{07}.

Uwaga o niezależności $\mathbb{E}\, S_T^2$ od wyboru h (a dokładniej od N) sugeruje, że przy $N \to \infty$ można oczekiwać zbieżności S_T. I słusznie, ta zmienna losowa zbiega do W_T, czyli do położenia procesu Wienera po czasie T. Wybór średniej kwadratowej (zamiast np. $\mathbb{E}\,|S_T|^3$) ma tu znaczenie jedynie dla prostoty rozumowania.

obok. Wybór odstępu czasowego h^2 zamiast, powiedzmy, h wydaje się dziwny, ale są ku niemu dobre powody, do których jeszcze dojdziemy.

Zgodnie z obietnicą, zbadamy teraz zmienność $f(S_t)$. W przypływie naiwności moglibyśmy oczekiwać, że przyrost $f(S_{t+h^2}) - f(S_t)$ jest równy wielkości $(S_{t+h^2} - S_t) \cdot Df(S_t)$; sumując po różnych t, otrzymalibyśmy wtedy równość podobną do tej z dowodu podstawowego twierdzenia analizy. Jak jednak łatwo zauważyć, iloraz różnicowy $\frac{f(S_{t+h^2})-f(S_t)}{S_{t+h^2}-S_t}$ jest równy któremuś z ilorazów $I_\pm = \frac{f(S_t \pm h)-f(S_t)}{\pm h}$, jednak któremu – to jest losowe. Jako że $Df(S_t)$ jest średnią $\frac{I_+ + I_-}{2}$, to różnica między $\frac{f(S_{t+h^2})-f(S_t)}{S_{t+h^2}-S_t}$ a $Df(S_t)$ wynosi $\pm\frac{I_+ - I_-}{2}$, gdzie znak jest dobrany w zgodzie z $S_{t+h^2} - S_t$. Tym samym przekonaliśmy się, że zachodzi równość

$$(\star\star) \qquad f(S_{t+h^2}) - f(S_t) = (S_{t+h^2} - S_t) \cdot Df(S_t) + \frac{1}{2}(S_{t+h^2} - S_t)^2 \cdot D^2 f(S_t).$$

I w ten sposób otrzymaliśmy poprawkę drugiego rzędu, analogiczną do tej we wzorze Itô. Podobieństwo widać najlepiej, gdy przyjmiemy $T = Nh^2$ (dla pewnego naturalnego N) i w duchu $(\star)$ zsumujemy otrzymaną równość $(\star\star)$ dla wszystkich $0 \leqslant t < T$ będących wielokrotnościami h^2:

$$f(S_T) - f(S_0) = \sum_{n=0}^{N-1} (S_{(n+1)h^2} - S_{nh^2}) \cdot Df(S_{nh^2}) + \frac{h^2}{2} \sum_{n=0}^{N-1} D^2 f(S_{nh^2}).$$

Warto przy tym wspomnieć, że S_T stanowi przybliżenie procesu Wienera W_T, co oznacza, że w granicy $h \to 0$ rzeczywiście otrzymujemy – choć wyłącznie formalnie – wzór Itô.

Do czego ta poprawka się przydaje? Dyskretny wzór Itô w postaci podanej wyżej ma jedną kluczową cechę, która czyni go pożytecznym: pierwsza z sum po prawej stronie jest *martyngałem*. Żeby tę zaletę zrozumieć i należycie docenić, rzeczywiście należałoby posiadać pewną wiedzę o martyngałach, dlatego polecam Czytelnikowi sięgnąć po archiwalne numery *Delty* (szczegóły na marginesie).

Jednak nawet bez tej wiedzy można znaleźć łatwe a interesujące zastosowanie. Ustalmy horyzont czasowy $T = Nh^2$ oraz funkcję f. Powiedzmy, że interesuje nas wartość oczekiwana $\mathbb{E}\, f(S_T)$, czyli średnia z możliwych wartości $f(S_T)$ (ważona prawdopodobieństwem otrzymania danej wartości). Zgodnie z dyskretnym wzorem Itô, dla jej znalezienia wystarczy obliczyć średnią obu sum po prawej stronie wzoru; uzasadnimy, że pierwsza z nich ma średnią równą zeru.

Zwróćmy uwagę na n-ty składnik: $(S_{(n+1)h^2} - S_{nh^2}) \cdot Df(S_{nh^2})$. Jeśli ustalimy wartość S_{nh^2}, to ustali się rownież $Df(S_{nh^2})$, ale różnica $S_{(n+1)h^2} - S_{nh^2}$ nadal pozostaje losowa – przyjmuje wartości $+h$ i $-h$ z prawdopodobieństwem $\frac{1}{2}$. To oznacza, że również badany składnik przyjmuje dwie przeciwne wartości z równym prawdopodobieństwem, a więc ma zerową średnią. Ponieważ nie zależy to od ustalonej wartości S_{nh^2}, więc zerową średnią otrzymamy również, biorąc pod uwagę losowość S_{nh^2}. Skoro tak jest dla każdego n, to samo możemy powiedzieć o całej sumie. W konsekwencji:

$$\mathbb{E}\, f(S_T) = f(0) + \frac{h^2}{2} \sum_{n=0}^{N-1} \mathbb{E}\, D^2 f(S_{nh^2}).$$

Dla przykładu rozważmy teraz funkcję $f(x) = x^2$. Łatwo się przekonać, że $D^2 f(x) = 2$ dla każdego x, a zatem każda ze średnich $\mathbb{E}\, D^2 f(S_{nh^2})$ też wynosi 2. Biorąc pod uwagę równość $h^2 = \frac{T}{N}$, uzyskujemy stąd $\mathbb{E}\, S_T^2 = T$, co samo w sobie jest interesującym wynikiem.

Ten przykład pokazuje nam jeszcze jedną rzecz. Gdybyśmy za odstęp między dwoma krokami przyjęli h zamiast h^2, to powyższy rachunek w miejsce T dałby $T \cdot h$. Okazuje się więc, że „nierówne" skalowanie w przestrzeni i czasie – w równaniach różniczkowych znane jako *skalowanie paraboliczne* – jest jedynym wyborem, jeśli chcemy, by wybór parametru h nie miał przełożenia na średnią kwadratową odległość po czasie T.

A czy niechcianej poprawki da się pozbyć? Da się! Żeby to zobaczyć, zauważmy najpierw, że dla prawdziwości wzoru $(\star\star)$ wystarczy, aby

S_t, S_{t+h^2} były liczbami różniącymi się o $\pm h$. Skoro tak, to prawdziwy jest również analogiczny wzór, w którym S_t i S_{t+h^2} zamienimy miejscami. Te dwa wzory – równość ($\star\star$) i jej odwróconą wersję – możemy odjąć stronami i podzielić przez 2, otrzymując

$$f(S_{t+h^2}) - f(S_t) = (S_{t+h^2} - S_t) \cdot \tfrac{Df(S_t)+Df(S_{t+h^2})}{2} + \frac{h^2}{4}(D^2 f(S_t) - D^2 f(S_{t+h^2})).$$

Jeśli teraz zsumujemy po $0 \leqslant t < T$ jak poprzednio, to feralna poprawka przyjmuje postać $\frac{h^2}{4}(D^2 f(S_0) - D^2 f(S_T))$. Co prawda nie znika zupełnie, ale jest dużo mniejsza niż poprzednio, bo zamiast N wyrazów są tylko 2. To oznacza, że w granicy $h \to 0$ otrzymujemy (znów: formalnie) wzór Itô, ale tym razem bez żadnej poprawki.

Czytelnik znający całkę Riemanna $\int_0^T g(t)\,\mathrm{d}t$ wie, że jest ona pozbawiona haczyka, o którym tu mowa. Całkę tę przybliżamy, dzieląc $[0, T]$ na drobne przedziały $[t_k, t_{k+1}]$ i sumując $a_k \cdot (t_{k+1} - t_k)$, przy czym za a_k możemy przyjąć wartość g w dowolnym punkcie z $[t_k, t_{k+1}]$, lub też średnią tych wartości.

Gdzie tkwi haczyk? Otóż brzemienna w skutkach okazała się niewinna zamiana $Df(S_t)$ na $\frac{Df(S_t)+Df(S_{t+h^2})}{2}$. Zmienność S_t (i podobnie zmienność W_t) jest na tyle duża, że powstała różnica nie znika przy przejściu do granicy $h \to 0$. Otrzymany przez nas wzór zapisuje się zazwyczaj jako

$$f(W_T) - f(W_0) = \int_0^T f'(W_t) \circ \mathrm{d}W_t,$$

przez użycie symbolu $\circ$ sygnalizując, że mamy tu do czynienia z tak zwaną całką Stratonowicza, a nie całką Itô.

Skoro dla procesów stochastycznych mamy dwa różne pojęcia całki, to które wybrać? To pytanie pozostawiam już Czytelnikowi, bo za każdą z możliwości przemawiają pewne zalety, które tutaj zostały ledwo zarysowane.

Recenzja książki *Czy Wszechświat myśli? I inne ważne pytania nauki*

Czy Wszechświat myśli? bardziej niż książką o tym, co wiemy, jest książką o tym, czego nie wiemy, i być może dowiedzieć się nie jesteśmy w stanie. Sabine Hossenfelder w przystępny sposób pokazuje czytelnikowi dzisiejsze granice poznania oraz jasno określa, w którym momencie kończy się nauka, a zaczyna naukowy bełkot. Jednocześnie czyni to ze swadą i nie wdając się zbytnio w szczegóły, dzięki czemu odbiorcy nie dopada senność. Gdyby Jeremy Clarkson pisał o fizyce, z pewnością robiłby to właśnie tak.

Autorka nie uznaje żadnych świętości („naukowości”?) i nawet częściej niż bioenergoterapeutów czy wróżki punktuje uczonych, którzy zbyt daleko odpłynęli w regiony niemożliwej do zweryfikowania spekulacji (np. głosicieli koncepcji wieloświata). Podczas lektury miałem nieodparte skojarzenie z telewizyjnymi *Pogromcami mitów*: w każdym rozdziale pisarka bierze na warsztat jakiś „mit”, tj. pytanie (np. czy nasze zmarłe babcie wciąż żyją dzięki kwantom), po czym określa, czy w ogóle da się na nie odpowiedzieć, wykorzystując metody naukowe, a jeśli tak – to jak ta odpowiedź brzmi.

Dla kogo tę książkę przeznaczono? Dla fizyków chyba nie, za bardzo prześlizguje się po tematach, choć z pewnością może służyć za lekką lekturę wakacyjną. Dla kompletnych laików tym bardziej nie, gdyż z nadmierną dezynwolturą skacze z wątku na wątek, co osobie bez chociaż zgrubnego pojęcia o przedstawianych zagadnieniach raczej uniemożliwi połapanie się w nich. Dla amatorów, którzy może nie nabyli biegłości w całkach i różniczkach, ale za to nieobce im są nazwiska Feynmana, Gödla czy Hoyle'a – tak, ci powinni poczuć się nią usatysfakcjonowani.

Czy Wszechświat myśli? rozpoczyna się od cytatu z Carla Sagana*, mnie na zakończenie recenzji również przychodzi na myśl passus tego autora: *Zakładając, że będę czytał jedną książkę tygodniowo, w ciągu swojego życia przeczytam ich kilka tysięcy, co stanowi jedynie promil zbiorów dobrej biblioteki. Sztuka polega na tym, by wiedzieć, które przeczytać.* Książka Sabiny Hossenfelder raczej nie należy do pozycji „obowiązkowych” i nie sądzę, by za kilkadziesiąt lat dorobiła się takiego statusu, jak twórczość popularnonaukowa Sagana, Hawkinga czy Goulda. Niemniej istnieje bardzo dobry powód, aby ją przeczytać: dzięki niej macie szansę dowiedzieć się, ile jest „cukru w cukrze” w przebijających się do masowych mediów popularnonaukowych stwierdzeniach. I że czasem są one jedynie ubranym w naukowy język wyznaniem wiary.

Mateusz RAJKOWSKI

*„Moim zdaniem dalece bardziej rozsądne jest pojmowanie wszechświata takim, jaki jest w rzeczywistości, niż pozostawanie przy urojeniach, choćby przynosiły one satysfakcję i uspokojenie.” Carl Sagan, *Świat nawiedzany przez demony: nauka jako światło w mroku*, Poznań 1999, s. 26.