Równanie Bellmana

*Wojciech KRYŃSKI**

Rozwiązanie zadania M 1202.
Wykażemy indukcyjnie, że dla każdej liczby naturalnej $n \geqslant 3$ istnieje n różnych liczb, z których każda jest dzielnikiem sumy $n-1$ pozostałych liczb.

Dla $n = 3$ liczby 1, 2, 3 spełniają powyższy warunek. Przyjmijmy więc, że spośród różnych liczb całkowitych dodatnich $a_1, a_2, \ldots, a_n$ każda jest dzielnikiem sumy pozostałych liczb. Warunek ten oznacza, że dla każdego $i = 1, 2, \ldots, n$, liczba a_i jest dzielnikiem liczby $a_1 + a_2 + \ldots + a_n$. Rozpatrzmy $n + 1$ liczb:

$$(*) \qquad a_1, a_2, \ldots, a_n, a_1+a_2+\ldots+a_n\,.$$

Wykorzystując założenie indukcyjne wnioskujemy, że każda z tych liczb jest dzielnikiem sumy tych liczb, czyli liczby $2(a_1 + a_2 + \ldots + a_n)$. Zatem każda spośród liczb $(*)$ jest dzielnikiem sumy pozostałych liczb. Dowód indukcyjny został zakończony.

Wyobraźmy sobie Robinsona, który po katastrofie okrętu dostaje się na bezludna wyspę. Wśród wyrzuconych na brzeg przedmiotów znajduje skrzynię z ziarnem. Po kilku próbach udaje mu się opanować technologię wypieku chleba. Jednak ilość ziarna jest ograniczona, a Robinson zdaje sobie sprawę, że prawdopodobnie na wyspie pozostanie do końca życia. Dlatego postanawia, że co roku będzie dzielił zboże na dwie części. Jedną będzie siał, a drugą konsumował. Pragnie zrobić to w sposób dający mu największą korzyść.

Problem Robinsona jest jednym z typowych problemów występujących w ekonomii. Dane są zasoby, którymi należy dysponować w taki sposób, aby uzyskać jak największy zysk. Innymi słowy, należy zmaksymalizować wartość pewnej funkcji. Podstawową trudność stanowi długi okres, w jakim rozważany jest problem. Liczba zmiennych jest na tyle duża, że standardowe kryteria na maksimum funkcji, na przykład zerowanie się pochodnych, stają się nieefektywne. Aby lepiej zrozumieć problem, wróćmy do Robinsona.

Załóżmy, że z każdego zasianego kilograma zboża po roku można zebrać dwa kilogramy plonów. Dla uproszczenia pomijamy wszystkie czynniki zewnętrzne, takie jak możliwe susze i inne klęski. Oznaczmy przez $x(t)$ ilość zboża, jaką posiada Robinson w roku t, a $u(t)$ niech będzie ilością zboża konsumowaną w tym roku. Zatem

$$x(t+1) = 2(x(t) - u(t))$$

i oczywiście spełnione jest ograniczenie $u(t) \in [0, x(t)]$. Ponadto przyjmujemy, że w znalezionej skrzyni było 100 kg zboża, czyli

$$x(0) = 100$$

(Robinson zaczyna liczyć lata od momentu przybycia na wyspę). Zakładamy też, że korzyść h, jaką czerpie Robinson z konsumowanego ziarna, zależy logarytmicznie od jego ilości. Dokładniej

$$h(u) = \ln(u).$$

W większości rzeczywistych przypadków nie istnieje ograniczenie na t, czyli horyzont czasowy jest nieskończony. Dodatkowo wprowadza się pewne czynniki, które deprecjonują zyski w odległej przyszłości. Zakłada się, że bardziej istotny jest zysk osiągany najszybciej (długoterminowe kalkulacje są też mniej pewne z powodu wielu nieprzewidywalnych czynników).

Aby uprościć problem, zakładamy, że Robinson szacuje, iż będzie żył jeszcze 40 lat. Zatem naszym celem jest znalezienie takich $u(0), \ldots, u(40)$, że suma

$$h(u(0)) + h(u(1)) + \cdots + h(u(40))$$

jest maksymalna.

Prosta strategia rozwiązywania tego typu problemów została opracowana przez Richarda Bellmana w latach pięćdziesiątych dwudziestego wieku. Nazywa się ona *programowaniem dynamicznym* i opiera się na *zasadzie optymalności*, która w przypadku Robinsona przybiera następującą postać:

Jeśli szacowana długość życia Robinsona wynosi 40 *lat, a ciąg* $u(0), u(1), \ldots, u(40)$ *jest optymalny dla warunku początkowego* $x(0)$*, to dla każdego* $t \in \{1, \ldots, 40\}$ *ciąg* $u(t), u(t+1), \ldots, u(40)$ *jest optymalny dla warunku początkowego* $x(t)$ *i szacowanej długości życia* $40 - t$.

Jest to konsekwencja następującego rozumowania. Gdyby ciąg $u(t), \ldots, u(40)$ nie był optymalny dla warunku początkowego $x(t)$, to istniałby lepszy ciąg $v(t), \ldots, v(40)$. Wówczas ciąg $u(0), \ldots, u(t-1), v(t), \ldots, v(40)$ byłby lepszy od ciągu $u(0), \ldots, u(40)$, o którym zakładaliśmy, że jest optymalny. Czyli otrzymalibyśmy sprzeczność.

Z zasady optymalności wynika, że problem Robinsona można rozwiązywać „od końca" (Czytelnik proszony jest o zatrzymanie się w tym miejscu, jeśli nie jest pewny, że rozumie zasadę optymalności; jest ona naprawdę prosta). Oznaczmy przez $V_t(x)$ maksymalną korzyść, jaką Robinson może osiągnąć przez $40 - t$ lat z warunkiem początkowym x. Zatem

$$V_t(x) = \max_{u(t)\in[0,x(t)],\ldots,u(40)\in[0,x(40)]} h(u(t)) + \cdots + h(u(40)),$$

gdzie $x(t) = x$. W szczególności obliczając $V_0(100)$, otrzymujemy maksymalną korzyść Robinsona w wyjściowym problemie.

*Instytut Matematyczny PAN

Bardziej ogólnie równanie przybiera postać
$$V_t(x) = \max_{u\in U} V_{t+1}(f(x,u,t)) + h(x,u,t),$$
gdzie U to pewien zbiór ograniczeń, f to funkcja przejścia, a h to funkcja korzyści. Obie funkcje mogą zależeć od zasobów x, sterowania u i czasu t.

Bezpośrednio z zasady optymalności dostajemy *równanie Bellmana*
$$V_t(x) = \max_{u(t)\in[0,x]} V_{t+1}(2(x-u(t))) + h(u(t)),$$
gdzie, tak jak wcześniej, wartość $2(x-u(t))$ jest ilością zboża w roku $t+1$, o ile w roku t było x zboża. Zakładamy dodatkowo, że
$$V_{41}(x) = 0,$$
co oznacza, że po śmierci Robinson nie czerpie już żadnych korzyści ze zboża. Zauważmy też, że $u(t)$, które maksymalizują kolejne $V_{t+1}(2(x-u)) + h(u)$, dają szukaną strategię optymalną.

Przejdźmy teraz do rachunków. Mając V_{41}, chcemy obliczyć V_{40}, korzystając z równania Bellmana. Dostajemy
$$V_{40}(x) = \max_{u(40)\in[0,x]} V_{41}(2(x-u(40))) + h(u(40)) = \max_{u(40)\in[0,x]} \ln(u(40)).$$
Logarytm jest funkcją rosnącą, więc maksimum będzie na końcu przedziału. Mamy zatem $V_{40}(x) = \ln(x)$ oraz $u(40) = x(40)$. Wynika stąd, że przed śmiercią Robinson powinien spożyć całe zboże.

Obliczamy następnie V_{39}, ponownie stosując równanie Bellmana.
$$V_{39}(x) = \max_{u(39)\in[0,x]} \ln(2(x-u(39))) + \ln(u(39)) = \max_{u(39)\in[0,x]} \ln(2(x-u(39))u(39)).$$
Jak wcześniej, korzystamy z faktu, że logarytm jest funkcją rosnącą i wnioskujemy, że szukane $u(39)$ będzie punktem maksimum funkcji kwadratowej
$$\varphi(u) = 2(x-u)u$$
na przedziale $[0,x]$. Bez trudu obliczmy, że jest ono przyjmowane dla $u = \frac{1}{2}x$. Zatem $V_{39}(x) = \ln\left(\frac{1}{2}x^2\right)$ oraz $u(39) = \frac{1}{2}x(39)$. Proces ten możemy kontynuować. Jako ćwiczenie pozostawiamy indukcyjny dowód, że dla $t \leqslant 39$
$$V_t(x) = \ln\left(\frac{2^{\binom{41-t}{2}}}{(41-t)^{40-t}}x^{41-t}\right)$$
oraz $u(t) = \frac{1}{41-t}x(t)$.

Podsumowując, wnioskujemy, że optymalna strategia Robinsona, przy przyjętych założeniach, powinna wyglądać następująco. W pierwszym roku powinien przeznaczyć on na spożycie $\frac{1}{41}$ zboża, które posiada, w kolejnym roku $\frac{1}{40}$, potem $\frac{1}{39}$ etc. ad mortem... Zauważmy jednak, że w rozważanym modelu, jeśli wydłużymy czas życia Robinsona, to ilość wyhodowanego zboża będzie rosnąć do nieskończoności. Jest to, oczywiście, fizycznie niemożliwe. Rozsądnie jest więc wprowadzić ograniczenie na x w postaci
$$x(t+1) = \min\{2(x(t)-u(t)), x_{MAX}\}.$$
Stosując równanie Bellmana, można bez trudu pokazać, że dla takiego problemu optymalna strategia Robinsona, aż do osiągnięcia x_{MAX}, wyglada tak, jakby nie było ograniczeń. Natomiast, gdy krytyczna wartość zostanie osiągnięta, optymalna ilość zboża przeznaczonego do konsumpcji wynosi $\frac{1}{2}x_{MAX}$.

W dniach 26–27 września 2008 roku, na Wydziale Fizyki UAM w Poznaniu, pod patronatem Dziekana oraz Polskiego Towarzystwa Fizycznego i Polskiego Towarzystwa Astronomicznego, organizowany jest Ogólnopolski Festiwal **Nauki Przyrodnicze na Scenie 3**, który stanowi kontynuację bardzo udanych trzech poprzednich festiwali *Fizyka na Scenie*, dwóch festiwali *Nauki Przyrodnicze na Scenie* oraz trzech europejskich festiwali *Physics on Stage* i dwóch festiwali *Science on Stage*.

Występy zespołów bedą oceniane w trzech kategoriach:
- demonstracje zjawisk,
- działania artystyczne związane z fizyką, naukami przyrodniczymi oraz astronomią i naukami o przestrzeni kosmicznej (przedstawienia teatralne, fotografie, rysunki, formy przestrzenne, wiersze itp.),
- pokazy multimedialne z zakresu nauk przyrodniczych i nauk o przestrzeni kosmicznej.

Do Poznania zaproszone zostaną osoby i zespoły (w sumie około 200 osób), wybrane przez Krajowy Komitet Organizacyjny, które do 3 czerwca 2008 roku zgłoszą propozycje występów.

Pełną informację o festiwalu można znaleźć na stronie `http://main3.amu.edu.pl/ fizscena/`.

Przewodniczący KKO
Prof. Wojciech Nawrocik

Rozwiązanie zadania M 1203.

Niech $x = BD$, $y = CE$, $z = AF$ oraz $p = DC$, $q = EA$, $r = FB$. Oznaczmy przez $[\mathcal{F}]$ pole figury $\mathcal{F}$. Wówczas
$$[ABC] = \frac{(x+p)(y+q)(z+r)}{4R} \quad \text{oraz} \quad \frac{[AEF]}{[ABC]} = \frac{zq}{(y+q)(z+r)},$$
skąd uzyskujemy
$$[AEF] = \frac{zq(x+p)}{4R}.$$
Analogicznie dowodzimy równości
$$[BFD] = \frac{xr(y+q)}{4R} \quad \text{oraz} \quad [CDE] = \frac{yp(z+r)}{4R}.$$
Zatem
$$[DEF] = [ABC] - [AEF] - [BFD] - [CDE] =$$
$$= \frac{1}{4R}\Big((x+p)(y+q)(z+r) - zq(x+p) - xr(y+q) - yp(z+r)\Big) = \frac{1}{4R}(xyz+pqr),$$
co należało wykazać.