

METODY UWZGLĘDNIANIA RYZYKA W ZADANIACH OPTYMALIZACJI REGUŁ DECYZYJNYCH

Andrzej Karbowski

Instytut Automatyki i Informatyki Stosowanej
Politechniki Warszawskiej
ul. Nowowiejska 15/19, 00-665 Warszawa

Streszczenie

Praca dotyczy stochastycznych zadań sterowania optymalnego ze sprzężeniem zwrotnym. Jak wiadomo, rozwiązywanie optymalne w takich zadaniach, gdy zakłócenia etapowe są niczależne, zaś kryterium oceny jest wartością oczekiwana sumy wskaźników etapowych, jest obliczane przy pomocy algorytmu stochastycznego programowania dynamicznego. Jednak przy takim podejściu nie jest brane pod uwagę ryzyko związane z kształtem funkcji gęstości rozkładu wskaźnika jakości. Można je uwzględnić na dwa sposoby: modyfikując wskaźnik jakości - licząc jego tzw. użyteczność, lub uwzględniając dodatkowe informacje o rozkładzie wskaźnika, na przykład wariancję. Obydwa te podejścia, wraz z metodami obliczeniowymi, są w pracy przedstawione oraz porównane.

Słowa kluczowe

podejmowanie decyzji, ryzyko, sterowanie optymalne, stochastyczne programowanie dynamiczne, funkcja użyteczności, kryteria typu wartość oczekiwana - wariancja,

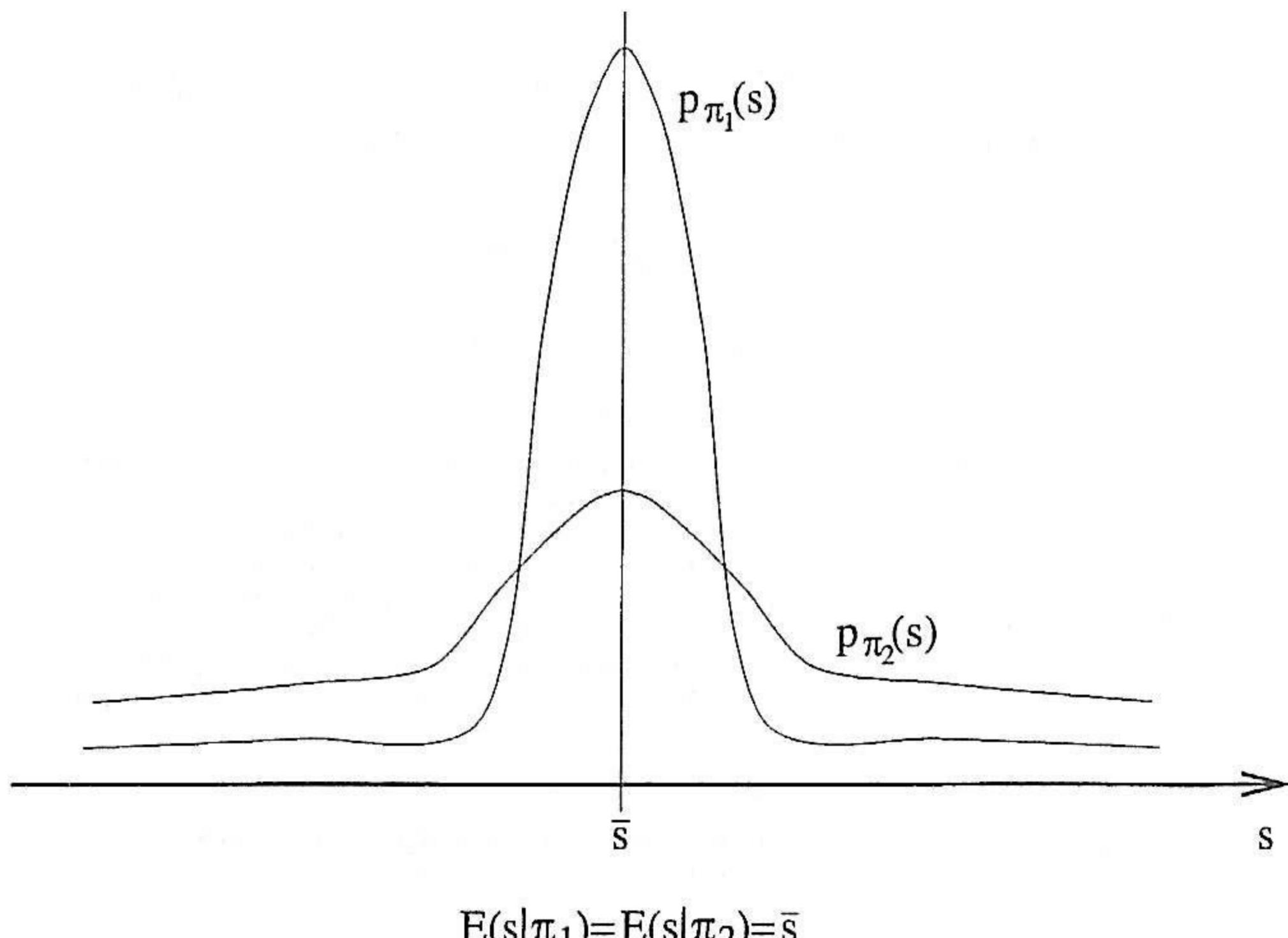
I. WSTĘP

Wiele problemów decyzyjnych dotyczy systemów działających w warunkach niepewności typu losowego, przy czym ocena jakości decyzji jest dokonywana na podstawie obserwacji zachowania się tych systemów w pewnym skończonym czasie. Z sytuacją taką mamy do czynienia na przykład w zadaniach optymalnego inwestowania (wyboru portfela akcji i obligacji), sterowania ruchem w sieciach teleinformatycznych (np. typu ATM), określania polityki zrzutów ze zbiorników retencyjnych w czasie powodzi lub suszy. Dysponując probabilistycznym opisem niepewności, czyli tzw. wejściem niesterowanych lub zakłóceń, optymalizuje się najczęściej wartość oczekiwana pewnych wskaźników jakości opisujących, w zależności od kontekstu, zyski lub koszty.

Jeszcze kilka lat temu cały materiał na temat podejmowania decyzji dotyczących procesów przebiegających w warunkach niepewności losowej, czyli sterowania stochastycznymi układami dynamicznymi, koncentrował się na tym jednym kryterium - wartości oczekiwanej całkowitego zysku lub kosztu. Jednak taki wskaźnik typu całkowitego zysku lub kosztu, jako funkcja losowych zakłóceń, jest przecież zmienną losową charakteryzującą się pewną funkcją gęstości rozkładu prawdopodobieństwa, dla której wartość oczekiwana - czyli pierwszy moment, jest tylko jedną, niewystarczającą charakterystyką.

Rozpatrzmy na przykład dwie funkcje gęstości rozkładu pewnego wskaźnika s odpowiadające dwóm politykom decyzyjnym π_1, π_2 , przedstawione na Rys. 1. Z rysunku tego widać, że mimo że obydwie polityki prowadzą do tej samej wartości oczekiwanej \bar{s} , to ich funkcje gęstości rozkładu prawdopodobieństwa są różne. Przede wszystkim polityka π_1 jest mniej ryzykowna, zapewnia zaistnienie kosztów/zysków bardziej przewidywalnych, o małym rozrzucie. Nie grozi ona częstym występowaniem bardzo słabych wyników. Polityka π_2 prowadząca do płaskiej funkcji gęstości rozkładu jest natomiast bardziej ryzykowna, wyniki otrzymywane w kolejnych eksperymentach będą bardziej odbiegały od siebie.

Ze względu na to, powinniśmy porównywać kompletne rozkłady związane z różnymi politykami, a nie tylko ich wartości oczekiwane.



Rys. 1. Przykład dwóch rozkładów wskaźnika jakości o tej samej wartości oczekiwanej równej \bar{s} różniących się stopniem ryzyka

W niniejszej pracy przedstawimy dwa podstawowe podejścia wykorzystywane do oceny reguł i polityk decyzyjnych uwzględniające element ryzyka: oparte na wykorzystaniu funkcji użyteczności oraz dwukryterialne. W tym ostatnim poszukuje się rozwiązań optymalnych w sensie Parety dla dwóch kryteriów jednocześnie: wartości oczekiwanej oraz wariancji.

II. SFORMUŁOWANIE ZADANIA

A. Model systemu.

Przymajemy, że mamy dany model obiektu sterowanego w postaci dyskretnego równania stanu:

$$x_{k+1} = f_k(x_k, u_k, w_k), \quad k = 0, 1, \dots, N-1 \quad (1)$$

gdzie:

- k - numer etapu; przyjmujemy że etap ten obejmuje przedział czasu $[t_k, t_{k+1})$
- x_k - stan układu na początku etapu k , czyli w chwili t_k
- u_k - sterowanie na etapie k
- w_k - zakłócenie na etapie k
- N - horyzont sterowania

Zakładamy, że stan układu x_k należy do przestrzeni stanów S , sterowanie u_k do przestrzeni sterowań C , a zakłócenie jest elementem przestrzeni zakłóceń D . Sterowanie u_k może przybierać wartości z niepustego podzbioru $U_k(x_k)$ przestrzeni C , który w ogólnym przypadku może zależeć od bieżącego stanu x_k ($u_k \in U_k(x_k)$ dla wszystkich $x_k \in S$ i k).

Poszukujemy optymalnej polityki, czyli ciągu:

$$\pi = \{\mu_0(\cdot), \mu_1(\cdot), \dots, \mu_{N-1}(\cdot)\} \quad (2)$$

reguł decyzyjnych (inaczej praw sterowania lub reguł sterowania)

$$u_k = \mu_k(x_k), \quad k = 0, 1, \dots, N-1 \quad (3)$$

gdzie każda reguła $\mu_k(\cdot), k = 0, 1, \dots, N-1$ jest funkcją z przestrzeni stanów w przestrzeń sterowań.

A.1 Zakłócenia.

Przymajemy że zakłócenia w_k mają charakter losowy, przy czym znane są ich rozkłady prawdopodobieństwa na kolejnych etapach

$$P_k(w_k|x_k, u_k), \quad k = 0, 1, \dots, N-1 \quad (4)$$

określone na zbiorach $W_k(x_k, u_k) \subseteq D$. Zakładamy zatem, że zakłócenia na kolejnych etapach są niezależne od siebie, ale mogą zależeć od bieżącego stanu oraz bieżącego sterowania.

Założenie to może budzić pewne wątpliwości: co to za niepewność która zależy od stanu układu? Przecież zawsze przyjmowało się, że zakłócenia to wejścia na które nie mamy żadnego wpływu!

Zresztą, najczęściej jest tak, że wymienione wyżej charakterystyki etapowe zakłóceń nie zależą ani od stanu, ani od sterowania, czyli:

$$P_k(w_k|x_k, u_k) = P_k(w_k) \quad (5)$$

i

$$W_k(x_k, u_k) = W_k \quad (6)$$

Inaczej jest już jednak, gdy zakłócenia w kolejnych etapach zależą od siebie, na przykład są procesem Markowa. Tak jest w przypadku systemów wodnych: dopływy do zbiorników w kolejnych okresach są ze sobą związane, nie zmieniają się całkowicie przypadkowo - dzięki temu można je przewidywać.

Tego typu zakłócenia można z reguły sprowadzić do przypadku zakłóceń niezależnych, stosując technikę rozszerzenia przestrzeni stanów.

Podobnie można postąpić, gdy rozważamy zakłócenia typu scenariuszowego, czyli pęk J trajektorii zakłóceń

$$\bar{w}_{[0,N-1]}^j = \{\bar{w}_0^j, \bar{w}_1^j, \bar{w}_2^j, \dots, \bar{w}_{N-1}^j\}, \quad j = 1, \dots, J \quad (7)$$

z przypisanymi do nich prawdopodobieństwami wystąpienia p^j .

Sposoby sprowadzenie obydwu wspomnianych modeli zakłóceń, oraz tzw. procesów ARMA, popularnych szczególnie w zastosowaniach ekonomicznych, do postaci standardowej (z zakłóceniami niezależnymi na kolejnych etapach) przedstawione są w Dodatku.

Pojawiające się w związku ze sprowadzaniem zadania do postaci standardowej dodatkowe współrzędne wektora stanu są podstawowym uzasadnieniem obecności wektora stanu wśród argumentów etapowych charakterystyk zakłóceń (4).

Na koniec warto podać jeszcze jeden ważny przykład wykorzystania zależności etapowego rozkładu zakłóceń od bieżącego stanu i sterowania. W przypadku, gdy rozpatrywany układ znajdujący się pod wpływem zakłóceń losowych jest dyskretny z natury i

może przyjmować tylko skończoną liczbę stanów, wygodnie jest go opisać jako sterowany łańcuch Markowa, czyli określić prawdopodobieństwa przejść pomiędzy poszczególnymi stanami układu

$$p_k^{ij}(u) = P(x_{k+1} = j | x_k = i, u_k = u) \quad (8)$$

Jest to prawdopodobieństwo tego, że następnym stanem układu x_{k+1} będzie j , gdy stanem bieżącym x_k jest i oraz zastosowano sterowanie u . Model ten można sprowadzić do równania (1) przyjmując następujące równanie stanu:

$$x_{k+1} = w_k \quad (9)$$

gdzie rozkład prawdopodobieństwa zakłócenia w_k określony jest następująco:

$$P(w_k = j | x_k = i, u_k = u) = p_k^{ij}(u) \quad (10)$$

B. Kryterium oceny.

B.1 Wskaźniki Bolzy, Lagrange'a i Mayera.

Najprostszym i jednocześnie najbardziej popularnym sposobem liczenia wskaźnika jakości jest suma kosztów (lub zysków) etapowych, czyli

$$\varphi(\pi, x_0, w) = \sum_{k=0}^{N-1} g_k(x_k, u_k, w_k) + g_N(x_N) \quad (11)$$

gdzie w jest ciągiem zakłóceń $\{w_0, w_1, \dots, w_{N-1}\}$. Wskaźnik taki nosi nazwę wskaźnika jakości typu Bolzy.

Są dwa szczególne przypadki wskaźnika Bolzy:

- gdy nie występuje funkcja kosztu (zysku) końcowego, czyli

$$g_N(\cdot) \equiv 0 \quad (12)$$

Mówimy w tym przypadku o wskaźniku typu Lagrange'a

- gdy nie występują funkcje kosztów (zysków) etapowych, czyli

$$g_k(\cdot, \cdot, \cdot) \equiv 0 \quad \forall k = 0, 1, \dots, N-1 \quad (13)$$

Mówimy wówczas o wskaźniku typu Mayera.

B.2 Kryterium typu wartości oczekiwanej

Jak wspomniano, przy zakłóceniach losowych wartość wskaźnika jakości jest zmienną losową, w zadaniach optymalnej syntezy bierze się najczęściej pod uwagę jej podstawową charakterystykę - pierwszy moment, czyli wartość oczekiwana. Wartość tego kryterium, oznaczmy je jako J , będzie zależna od polityki decyzyjnej π oraz warunku początkowego x_0 , czyli

$$J(\pi, x_0) = \underset{w_0, \dots, w_{N-1}}{\mathbf{E}} \varphi(\pi, x_0, w) \quad (14)$$

III. NIEKTÓRE WŁASNOŚCI KRYTERIÓW TYPU WARTOŚCI OCZEKIWANEJ

A. Dlaczego poszukujemy optymalnych reguł, a nie trajektorii?

Rozważmy następujące zadanie sterowania optymalnego:

Przykład 1: Mamy dany niestacjonarny skalarny układ dynamiczny opisany na pierwszych dwóch etapach równaniami:

$$x_1 = u_0 + w_0 \quad (15)$$

$$x_2 = x_1 + u_1 \quad (16)$$

Skalar w_0 jest losowym zakłóceniem mogącym przyjmować wartości $+1$ i -1 z prawdopodobieństwem $\frac{1}{2}$. Należy tak dobrać sterowania u_0 i u_1 , by zminimalizować oczekiwana wartość modułu stanu x_2 , czyli:

$$E|x_2| \quad (17)$$

Jest to więc zadanie ze wskaźnikiem typu Mayera. Przeprowadzimy najpierw optymalizację trajektorii (programu) sterowania. Mamy:

$$E\{|x_2|\} = E\{|u_0 + u_1 + w_0|\} = \frac{1}{2}(|u_0 + u_1 + 1| + |u_0 + u_1 - 1|) \quad (18)$$

Łatwo sprawdzić, że optymalną wartością wskaźnika jakości (17) jest:

$$\min_{u_0, u_1} E\{|x_2|\} = 1 \quad (19)$$

i że optymalnymi wartościami sterowań u_0 i u_1 są te dla których

$$-1 \leq u_0 + u_1 \leq 1 \quad (20)$$

Rozpatrzmy teraz zadanie w którym optymalizujemy polityki sterowania, czyli sytuację, gdy u_0 i u_1 są wybierane sekwencyjnie, tak że x_1 jest znane podczas wybierania u_1 . Wówczas oczywiście optymalna polityka jest następująca:

$$\hat{\mu}_1(x_1) = -x_1 \quad (21)$$

$$\hat{\mu}_0(x_0) = \text{dowolne} \quad (22)$$

Dzięki takiej polityce optymalny koszt jest zmniejszony do 0, czyli:

$$J(\hat{\pi}, x_0) = 0 \quad (23)$$

Zauważmy, że jeżeli nie ma niepewności, tzn. gdy $w_0 = \bar{w}_0$ z prawdopodobieństwem 1, to optymalny koszt jest zerowy, niezależnie od tego, czy x_1 jest znane przed wyborem u_1 , czy nie. Jest to fakt ogólny i intuicyjnie oczywisty, że sprzężenie zwrotne nie pomaga w przypadku deterministycznym, tzn., gdy wszystko jest przewidywalne ze 100% pewnością. Z tego powodu nie warto wówczas optymalizować polityk sterowania, co jest zawsze zajęciem trudniejszym niż optymalizacja trajektorii sterowania. Do rozwiązania tego ostatniego zadania można na przykład wykorzystać bardzo efektywne gradientowe metody programowania matematycznego.

B. Metoda rozwiązania

W zadaniu Bolzy w przypadku poszukiwania optymalnej polityki sterowania, czyli ciągu reguł decyzyjnych, optymalizowane kryterium ma postać:

$$J(\pi, x_0) = \underset{w_0, \dots, w_{N-1}}{\mathbb{E}} \varphi(\pi, x_0, w) = \underset{w_0, \dots, w_{N-1}}{\mathbb{E}} \left\{ \sum_{k=0}^{N-1} g_k(x_k, u_k, w_k) + g_N(x_N) \right\} \quad (24)$$

Optymalną politykę można wyznaczyć przy pomocy algorytmu programowania dynamicznego o postaci [1]:

1. W zadaniach minimalizacji kosztu całkowitego

$$J_k(x_k) = \min_{u_k \in U_k(x_k)} \underset{w_k}{\mathbb{E}} [g_k(x_k, u_k, w_k) + J_{k+1}(f_k(x_k, u_k, w_k))] \quad (25)$$

2. W zadaniach maksymalizacji zysku całkowitego

$$J_k(x_k) = \max_{u_k \in U_k(x_k)} \underset{w_k}{\mathbb{E}} [g_k(x_k, u_k, w_k) + J_{k+1}(f_k(x_k, u_k, w_k))] \quad (26)$$

gdzie $J_k(\cdot)$ jest funkcją optymalnego kosztu/zysku dojścia (tzw. funkcją Bellmana), dla $k = N-1, N-2, \dots, 1$, przyjmując $J_N(\cdot) \equiv g_N(\cdot)$.

B.1 Przykłady ułomności kryterium typu wartości oczekiwanej

Posługiwanie się tylko jednym kryterium - wartością oczekiwanaą wskaźnika jakości, często może prowadzić do nieco lekkomyślnych, a czasami wręcz niebezpiecznych decyzji. Rozpatrzmy dwa następujące zadania decyzyjne:

Przykład 2: Inwestor rozważa czy włączyć się, czy nie do gry, która trwa już k_0 etapów. Polega ona na tym, że wpłaca się sumę 2^{k_0} (przyjmijmy że złotych) i czeka na wynik losowania - rzutu monetą. Inwestor dostaje 2^i złotych, gdzie i liczba wylosowanych pod rząd orłów od jego wejścia do gry.

Łatwo jest sprawdzić że wartość oczekiwana $E\varphi$ zysku inwestora włączającego się do gry w etapie k_0 jest równa:

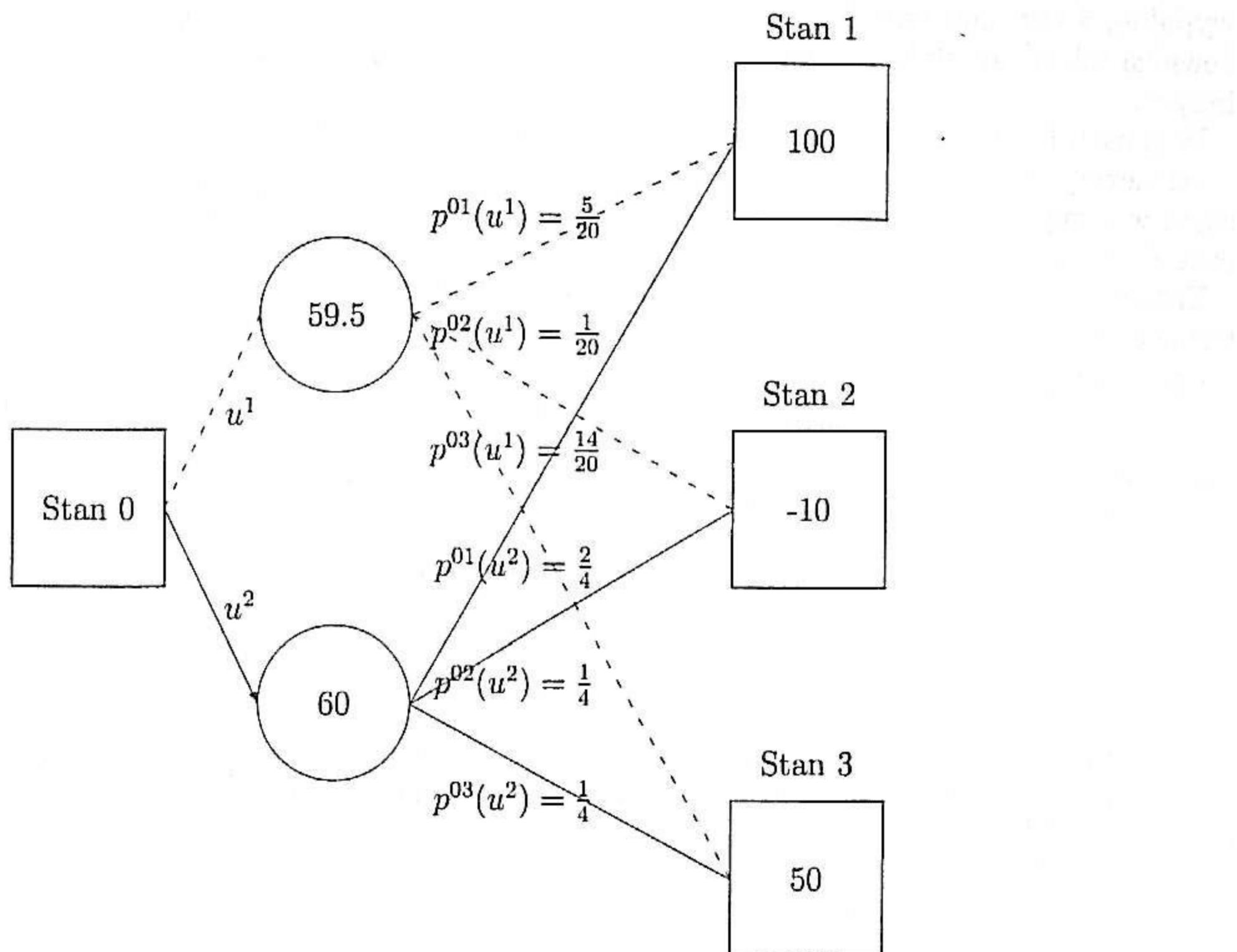
$$\begin{aligned} E\varphi &= \frac{1}{2} \cdot 1 + \frac{1}{2} \cdot \left\{ \frac{1}{2} \cdot 2 + \frac{1}{2} \cdot \left[\frac{1}{2} \cdot 4 + \frac{1}{2} \cdot \dots \right] \right\} - 2^{k_0} = \\ &= \sum_{k=k_0}^{\infty} \left(\frac{1}{2} \right)^{k-k_0+1} \cdot 2^{k-k_0} - 2^{k_0} = \infty \end{aligned} \quad (27)$$

czyli zawsze się opłaca wejść do gry, niczależnie od sumy 2^{k_0} jaką trzeba zapłacić na początku.

Sytuacja przedstawiona w tym przykładzie przypomina tzw. piramidy finansowe. Jak widać, gdyby się kierować jedynie kryterium wartości oczekiwanej, zawsze należałyby brać udział w takich przedsięwzięciach.

Przykład 3: Zadanie to zostało zaczerpnięte z pracy [5].

Weźmy pod uwagę następujące, jednoetapowe zadanie typu Mayera. Na etapie zerowym możemy zastosować jedno z dwóch sterowań u^1 lub u^2 . Pod wpływem sterowania proces może się znaleźć w jednym z trzech stanów, z którymi związane są pewne wartości zysku. Odpowiednie prawdopodobieństwa przejść przedstawione są na Rys. 2. Celem sterowania jest maksymalizacja zysku.



Rys. 2. Proste zadanie w którym jedno kryterium typu wartości oczekiwanej jest niewystarczające

Posługując się jednym kryterium oceny, w postaci wartości oczekiwanej wskaźnika jakości, dochodzimy do wniosku, że optymalne jest sterowanie u^2 , ponieważ prowadzi do większej wartości oczekiwanej zysku (równiej 60). Z tym sterowaniem wiąże się jednak spore ($\frac{1}{4}$) prawdopodobieństwo przejścia do stanu drugiego, w którym tracimy 10 jednostek. W przypadku sterowania u^1 , to prawdopodobieństwo jest pięciokrotnie mniejsze, a wartość oczekiwana zysku mniejsza tylko o 0.5 (czyli o 0.83%). Bardziej rozsądny wydaje się więc wybór u^1 .

IV. FUNKCJA UŻYTECZNOŚCI

Jest jeden przypadek, gdy porównywanie jedynie wartości oczekiwanej może zastąpić porównywanie całych rozkładów prawdopodobieństwa wskaźników jakości. Jest tak wtedy, gdy liczymy nie wartość oczekiwana pierwotnego wskaźnika jakości, ale pewnej jego nielinowej funkcji, tzw. funkcji użyteczności. Teoria użyteczności pochodzi od von Neumanna i Morgensterna. Tak jak w ich książce [7] będziemy nazywać rozkłady wskaźników jakości loteriami.

Podstawowym warunkiem istnienia funkcji użyteczności jest, by rozkłady prawdopodobieństwa wartości wskaźnika jakości odpowiadające różnym politykom dały się uporządkować liniowo, tzn. by wszystkie były ze sobą porównywane i można im było przypisać liczby rzeczywiste. Ponadto, w przestrzeni rozkładów musi być spełniony warunek zachowania porządku między rozkładami wskaźnika jakości, czyli loteriami, przy kombinacji

wypukłej z dowolną inną loterią oraz istnienia pośredniej loterii ocenianej przez użytkownika tak samo jak loteria uzyskana na drodze kombinacji wypukłej dowolnych dwóch innych.

W sposób formalny założenia te można przedstawić następująco:

Oznaczmy przestrzeń loterii jako \mathcal{P} ; relację określoną na tej przestrzeni jako " \preceq ", przy czym w przypadku gdy między dwiema loteriami P_1 i P_2 zachodzi jednocześnie $P_1 \preceq P_2$ oraz $P_1 \succeq P_2$ będziemy pisać w skrócie $P_1 \sim P_2$.

Założenie 1: Relacja " \preceq " jest relacją pełnego porządku.

Oznacza to iż:

- jest zwrotna, czyli

$$P \preceq P, \forall P \in \mathcal{P} \quad (28)$$

- jesteśmy zdolni określić nasze preferencje między dowolnymi dwiema loteriami $P_1, P_2 \in \mathcal{P}$
- jest przechodnia, czyli jeżeli

$$P_1 \preceq P_2 \quad \text{oraz} \quad P_2 \preceq P_3$$

to

$$P_1 \preceq P_3$$

Założenie 2: Jeśli między dwiema loteriami P_1 i P_2 występuje relacja $P_1 \preceq P_2$, to po policzeniu kombinacji wypukłej z dowolną loterią P , relacja między otrzymanymi loteriami będzie zachowana, czyli:

$$\alpha P_1 + (1 - \alpha)P \preceq \alpha P_2 + (1 - \alpha)P \quad (29)$$

Założenie 3: Dla dowolnych trzech loterii P_1, P_2, P_3 , takich że

$$P_1 \preceq P_2 \preceq P_3 \quad (30)$$

istnieje współczynnik kombinacji wypukłej α taki że:

$$\alpha P_1 + (1 - \alpha)P_3 \sim P_2 \quad (31)$$

O ile pierwsze i trzecie z powyższych założeń są dość oczywiste, to drugie warto jest bliżej wyjaśnić. Z założenia tego wynika, że jeżeli rozważamy dwie loterie:

$$P_1 = 8000 \text{ zł z prawd. 1} \quad (32)$$

$$P_2 = \begin{cases} 10000 \text{ zł} & \text{z prawd. 0.9} \\ 0 & \text{z prawd. 0.1} \end{cases}$$

i przyjmiemy że

$$P_1 \succeq P_2 \quad (33)$$

oraz uwzględnimy loterię

$$P = \begin{cases} 0 & \text{z prawd. 0.8} \\ 1000 \text{ zł} & \text{z prawd. 0.2} \end{cases}$$

to w bardziej skomplikowanej sytuacji, w której rozważamy loterie mieszane:

- P'_1 w której może być wylosowana loteria P_1 z prawdopodobieństwem α i loteria P z prawdopodobieństwem $1 - \alpha$, czyli:

$$P'_1 = \begin{cases} 8000 \text{ zł} & \text{z prawd. } \alpha \\ 1000 \text{ zł} & \text{z prawd. } 0.2 \cdot (1 - \alpha) \\ 0 & \text{z prawd. } 0.8 \cdot (1 - \alpha) \end{cases} \quad (34)$$

- P'_2 w której może być wylosowana loteria P_2 z prawdopodobieństwem α i loteria P z prawdopodobieństwem $1 - \alpha$, czyli:

$$P'_2 = \begin{cases} 10000 \text{ zł} & \text{z prawd. } 0.9 \cdot \alpha \\ 1000 \text{ zł} & \text{z prawd. } 0.2 \cdot (1 - \alpha) \\ 0 & \text{z prawd. } 0.7 \cdot (1 - \alpha) + 0.1 \end{cases} \quad (35)$$

to niezależnie od wartości α zawsze będzie zachowana preferencja:

$$P'_1 \succeq P'_2 \quad (36)$$

Można mieć zastrzeżenia co do słuszności relacji (36), bo przecież głównym powodem uznania loterii P_1 za lepszą niż P_2 było to, że dawała ona 8000 zł na pewno, podczas gdy w P_2 można było wprawdzie dostać nieco więcej bo 10000 zł, ale nie zawsze. W 10% przypadków nie dostałoby się nic.

Argument o całkowitej pewności wygranej znika, gdy rozważamy loterie P'_1 i P'_2 . Wówczas żadna wygrana nie jest pewna, a różnica w prawdopodobieństwie wygranej 8000 zł w loterii P'_1 i 10000 zł w loterii P'_2 na tyle mała, że większość wybrałaby loterię P'_2 , bo jeśli już musimy grać to lepiej o wyższą stawkę. Szczególnie dla małych prawdopodobieństw dużej wygranej, czyli wartości α .

Na marginesie warto zauważyć, że preferencja (33) jest sprzeczna z posługiwaniem się przy wyborze kryterium typu wartości oczekiwanej zysku.

Wróćmy jednak do zadania sterowania. Oznaczmy pierwotną funkcję oceny polityk sterowania na całym horyzoncie (nie zastanawiamy się tutaj jak jej wartość zależy od kosztów etapowych), jako ψ . Przy występujących zakłóciach losowych w , dla wektora stanu x i ustalonej polityki π wartość funkcji $\psi(\pi, x_0, w)$ jest zmienną losową, opisaną loterią (rozkładem) P .

Załóżmy że porównujemy ze sobą dwie polityki decyzyjne π_1 oraz π_2 , które prowadzą do loterii, odpowiednio P_1 i P_2 .

Definicja 1: Funkcją użyteczności nazywamy funkcję $U : \mathfrak{R} \mapsto \mathfrak{R}$, taką że

$$P_1 \preceq P_2 \iff E[U(\psi(\pi_1, x_0, w))] \leq E[U(\psi(\pi_2, x_0, w))] \quad (37)$$

Funkcja użyteczności tłumaczy zatem preferencje w przestrzeni loterii, czyli rozkładów wskaźników jakości, symbolizowane za pomocą relacji " \preceq " na zwykłe nierówności na osi liczbowej wyrażone przez relację " \leq ". Jest więc bardzo atrakcyjnym narzędziem do porównywania loterii otrzymywanych przy zastosowaniu różnych polityk sterowania.

Przedstawimy teraz twierdzenie mówiące o tym kiedy taka funkcja użyteczności istnieje.

Twierdzenie 1: Jeżeli spełnione są założenia (1)-(3), oraz funkcja $\psi(\dots, \dots)$ może przybierać wartości ze skończonego zbioru, to istnieje funkcja użyteczności o własności (37).

Twierdzenie powyższe jest dość naturalne i dlatego jego formalny dowód zostanie pominięty. Zamiast niego przedstawimy jedną z metod konstruowania takiej funkcji (na której zresztą można oprzeć dowód Twierdzenia 1).

Założyliśmy że funkcja $\psi(., ., .)$ może przybierać wartości ze skończonego zbioru, możemy więc oznaczyć kolejne rosnące wartości jako $\psi_1, \psi_2, \psi_3, \dots, \psi_K$. Mamy zatem:

$$\psi_1 < \psi_2 < \psi_3 < \dots < \psi_K \quad (38)$$

Przyjmiemy wstępnie że

$$U(\psi_1) = 0 \quad (39)$$

oraz

$$U(\psi_K) = 1 \quad (40)$$

Będziemy teraz subiektywnie nadawali użyteczność p_i wskaźnikowi ψ_i , czyli

$$U(\psi_i) = p_i \quad (41)$$

Wartość p_i ma interpretację takiego prawdopodobieństwa że uważamy za równoważne loterie

$$P_{1,K}^i = \begin{cases} \psi_1 & \text{z prawd. } 1 - p_i \\ \psi_2 & \text{z prawd. 0} \\ \vdots & \\ \psi_{K-1} & \text{z prawd. 0} \\ \psi_K & \text{z prawd. } p_i \end{cases} \quad (42)$$

oraz

$$P_i^i = \begin{cases} \psi_1 & \text{z prawd. 0} \\ \psi_2 & \text{z prawd. 0} \\ \vdots & \\ \psi_{i-1} & \text{z prawd. 0} \\ \psi_i & \text{z prawd. 1} \\ \psi_{i+1} & \text{z prawd. 0} \\ \vdots & \\ \psi_{K-1} & \text{z prawd. 0} \\ \psi_K & \text{z prawd. 0} \end{cases} \quad (43)$$

czyli

$$P_{1,K}^i \sim P_i^i \quad (44)$$

Oczywiście jest to tylko jeden ze sposobów uzyskania własnej funkcji użyteczności. Sposobów tych jest wiele, tak jak wiele jest funkcji użyteczności. Obowiązuje twierdzenie:

Twierdzenie 2: Funkcja użyteczności $U(.)$ jest jednoznaczna z dokładnością do transformacji afonicznej o dodatnim współczynniku kierunkowym, tzn. jeżeli $U^*(.)$ jest inną funkcją użyteczności, to istnieją scalary rzeczywiste $\alpha > 0, \beta$ takie że:

$$U^*(.) = \alpha \cdot U(.) + \beta \quad (45)$$

Twierdzenie to wynika wprost z liniowości operatora wartości oczekiwanej. Można je wykorzystać na przykład do skalowania funkcji użyteczności otrzymanej przy pomocy metody opisanej powyżej.

Analizując funkcję użyteczności można ocenić jaki stosunek do ryzyka ma decydent.

Definicja 2: Mówimy że decydent jest:

1. Niechętny ryzyku, gdy

- w przypadku maksymalizacji wskaźnika jakości

$$E[U(\psi(\pi, x_0, w))] \leq U[E(\psi(\pi, x_0, w))] \quad \forall \pi, \forall x_0 \quad (46)$$

- zaś w przypadku minimalizacji wskaźnika jakości

$$E[U(\psi(\pi, x_0, w))] \geq U[E(\psi(\pi, x_0, w))] \quad \forall \pi, \forall x_0 \quad (47)$$

2. Sklonny do ryzyka, gdy

- w przypadku maksymalizacji wskaźnika jakości

$$E[U(\psi(\pi, x_0, w))] \geq U[E(\psi(\pi, x_0, w))] \quad \forall \pi, \forall x_0 \quad (48)$$

- zaś w przypadku minimalizacji wskaźnika jakości

$$E[U(\psi(\pi, x_0, w))] \leq U[E(\psi(\pi, x_0, w))] \quad \forall \pi, \forall x_0 \quad (49)$$

3. Obojętny w stosunku do ryzyka, gdy

$$E[U(\psi(\pi, x_0, w))] = U[E(\psi(\pi, x_0, w))] \quad \forall \pi, \forall x_0 \quad (50)$$

Decydent jest więc niechętny ryzyku, gdy wolałby otrzymać wartość oczekiwana loterii niż wziąć w niej udział, a sklonny do ryzyka w przeciwnym wypadku. Można pokazać, że niechęć do ryzyka odpowiada wklęsłości funkcji użyteczności w zadaniach z maksymalizacją wskaźnika jakości i wypukłości w zadaniach z minimalizacją, a skłonność do ryzyka odwrotnie. Obojętność na ryzyko odpowiada oczywiście liniowości funkcji użyteczności.

W poprzednim rozdziale gdy rozważaliśmy kryteria:

$$E[\varphi(\pi, x_0, w)] \quad (51)$$

funkcja $\varphi(\cdot, \cdot, \cdot)$ była już ostatecznym wskaźnikiem, po zastosowaniu funkcji użyteczności U do pierwotnej funkcji ocen $\psi(\cdot, \cdot, \cdot)$, czyli

$$\varphi(\pi, x_0, w) = U(\psi(\pi, x_0, w)) \quad (52)$$

Jeśli funkcję użyteczności $U(\cdot)$ pomija się, czyli przyjmuje

$$U(\xi) = \xi \quad (53)$$

oraz

$$\varphi(\cdot, \cdot, \cdot) \equiv \psi(\cdot, \cdot, \cdot) \quad (54)$$

a więc nie modyfikuje się pierwotnego wskaźnika jakości (jest to bardzo częsty przypadek przy formułowaniu zadań optymalnej syntezy), to oznacza to, że decydent ma obojętny stosunek do ryzyka, nie ma ono dla niego żadnego znaczenia. Liczy się tylko wartość oczekiwana wskaźnika jakości.

Nawet z tej krótkiej prezentacji podstaw teorii użyteczności wynika, że już same założenia tej teorii budzą poważne wątpliwości. Ponadto, wyznaczenie funkcji użyteczności jest czynnością bardzo kłopotliwą, szczególnie gdy mamy do czynienia z dużymi zbiorami

możliwych wartości wskaźników jakości. Kłopotliwą rzeczą jest uporządkowanie loterii bazowych (43) - nadanie im skalarnych wartości liczbowych.

Funkcja użyteczności, jeśli uwzględniamy ryzyko, jest zawsze nieliniowa. Z punktu wiedzenia algorytmu numerycznego, który wykorzystamy do rozwiązania zadania oznacza to, iż w zadaniach ze wskaźnikiem Bolzy (czyli tam, gdzie koszt lub zysk łączny jest sumą kosztów lub zysków etapowych) nie będzie można skorzystać z dekompozycji etapowej opierającej się na liniowości operatora wartości oczekiwanej.

Funkcja ta jest definiowana w sposób subiektywny, czyli może się zmieniać w zależności od samopoczucia decydenta, sytuacji zewnętrznej, itp. Przy rozwiązywaniu tego samego zadania różni decydenci będą mieli różne funkcje użyteczności. Powstaje przy tym nienaturalna sekwencja:

- | | |
|--------------------------------------|--|
| 1) tworzenie modelu systemu | - czynność maksymalnie zbiektywizowana |
| 2) określenie funkcji użyteczności | - czynność oparta na intuicji z dużym i nieuchronnym udziałem czynnika subiektywnego |
| 3) rozwiązanie zadania optymalizacji | - czynność sformalizowana, i zalgorytmizowana, najczęściej wykonywana przez komputer |

Kluczowe znaczenie ma etap środkowy, który jest heurystyczno - subiektywny, choć na początku i na końcu wydaje się, że jeśli mamy dobry model i dobre metody obliczeniowe, to i wynik musi być dobry. Otóż nie będzie dobry, jeśli zastosujemy złą funkcję użyteczności. Jeśli ten etap pominiemy, to znaczy jeśli funkcja użyteczności będzie tożsamością, to będzie to oznaczało że ignorujemy ryzyko związane z różnymi politykami sterowania rozważanymi przez nas.

Jest to oczywiście podejście niepraktyczne. W życiu stopień ryzyka jest ważnym kryterium przy podejmowaniu decyzji. Na przykład przy wyborze banku lub spółdzielni mieszkaniowej, którym się powierza swoje oszczędności, bierze się zawsze pod uwagę ich wiarygodność. Podobnie kupując bilet lotniczy bierze się pod uwagę nie tylko cenę ale i zaufanie do przewoźnika. Warto by było uwzględnić ten element również i przy wyznaczaniu optymalnych polityk.

Jeśli przedstawione powyżej podejście z występującą jawnie lub niejawnie funkcją użyteczności nie jest ani wygodne, ani naturalne, może lepiej byłoby mierzyć ryzyko wprost za pomocą dodatkowego kryterium, na przykład wariancji. Wydaje się, że większość decydentów postępuje w taki właśnie sposób. Warto jest też dostarczyć decydentowi narzędzia pozwalającego na łatwą zmianę preferencji i obserwację wywołanych nią rezultatów optymalizacji.

Metody i algorytmy umożliwiające stworzenie takiego właśnie narzędzia omówione są w dalszej części pracy.

V. UWZGLĘDNIANIE RYZYKA POPRZEZ DODATKOWE KRYTERIUM TYPU WARIANCJI

W wielu zadaniach warto rozważyć dwa kryteria jednocześnie:

- wartość oczekiwana kosztów lub zysków
- ich wariancję

i dokonać wyboru takiej ich kombinacji, przy której wartość pierwszego kryterium będzie zadowalająca, a przy tym wartość drugiego nie będzie zbyt duża [6]. W ten sposób dochodzimy do zadania wielokryterialnego sterowania optymalnego w układzie zamkniętym z niepewnością. Główną trudnością przy jego rozwiązyaniu będzie nieseparowalność wskaźnika typu wariancji, ponieważ przy jej liczeniu podnosi się do kwadratu wartość oczekiwanych kosztów (zysków). Nie można więc wprost skorzystać z separowalności etapowej oraz algorytmu programowania dynamicznego. Jest to jednak możliwe po dokonaniu pewnych transformacji zadania. Dwie z nich zostaną opisane poniżej.

A. Sformułowanie zadania.

Oznaczmy dla uproszczenia wskaźnik jakości (11) wyrażający koszt (zysk) całkowity jako φ_π , czyli:

$$\varphi_\pi = \sum_{k=0}^{N-1} g_k(x_k, \mu_k(x_k), w_k) + g_N(x_N) \quad (55)$$

Będziemy dla niego obliczać zarówno wartość oczekiwanych $E(\varphi_\pi)$ jak i wariancję

$$V(\varphi_\pi) = E[\varphi_\pi - E(\varphi_\pi)]^2 = E(\varphi_\pi^2) - E^2(\varphi_\pi) \quad (56)$$

i optymalizować je w sensie Parety względem polityk π , czyli poszukiwać zbioru:

$$\mathbb{D}_{--}^e \cup_{\pi \in \Pi} \left\{ \begin{bmatrix} E(\varphi_\pi) \\ V(\varphi_\pi) \end{bmatrix} \right\} \quad (57)$$

lub, gdy mamy do czynienia z zyskami a nie kosztami:

$$\mathbb{D}_{+-}^e \cup_{\pi \in \Pi} \left\{ \begin{bmatrix} E(\varphi_\pi) \\ V(\varphi_\pi) \end{bmatrix} \right\} \quad (58)$$

Jako Π oznaczony został zbiór dopuszczalnych polityk sterowania. Operator e_D symbolizuje poszukiwanie punktów efektywnych, optymalnych względem stożka D , czyli

$$e_D Q = \{q^* \in Q : (q^* + D) \cap Q = \emptyset\} \quad (59)$$

zaś \tilde{D}_{--} , \tilde{D}_{+-} to stożki związane z "kierunkami" optymalizacji po współrzędnych. Dokładniej:

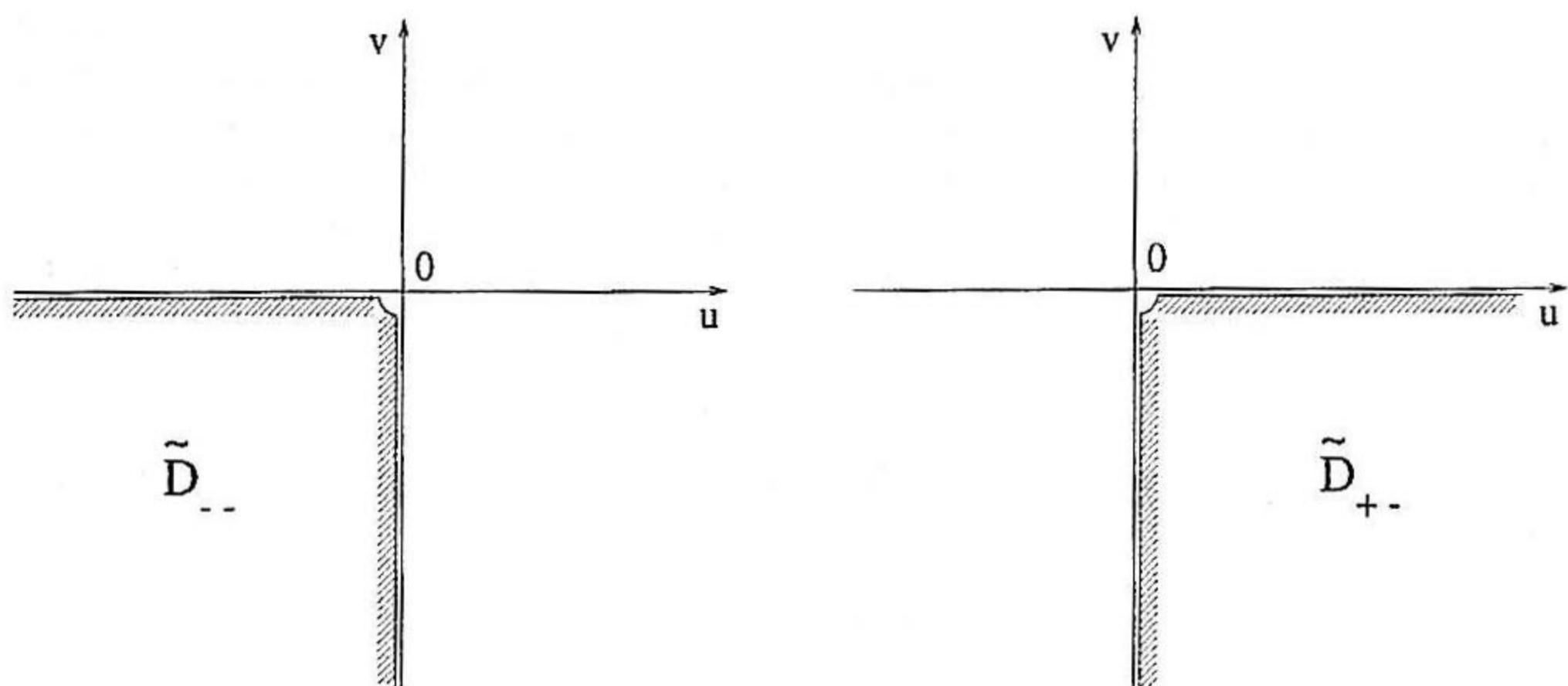
$$\tilde{D}_{--} = \{(u, v) \in \mathbb{R}^2 : u \leq 0, v \leq 0, u^2 + v^2 > 0\} \quad (60)$$

$$\tilde{D}_{+-} = \{(u, v) \in \mathbb{R}^2 : u \geq 0, v \leq 0, u^2 + v^2 > 0\} \quad (61)$$

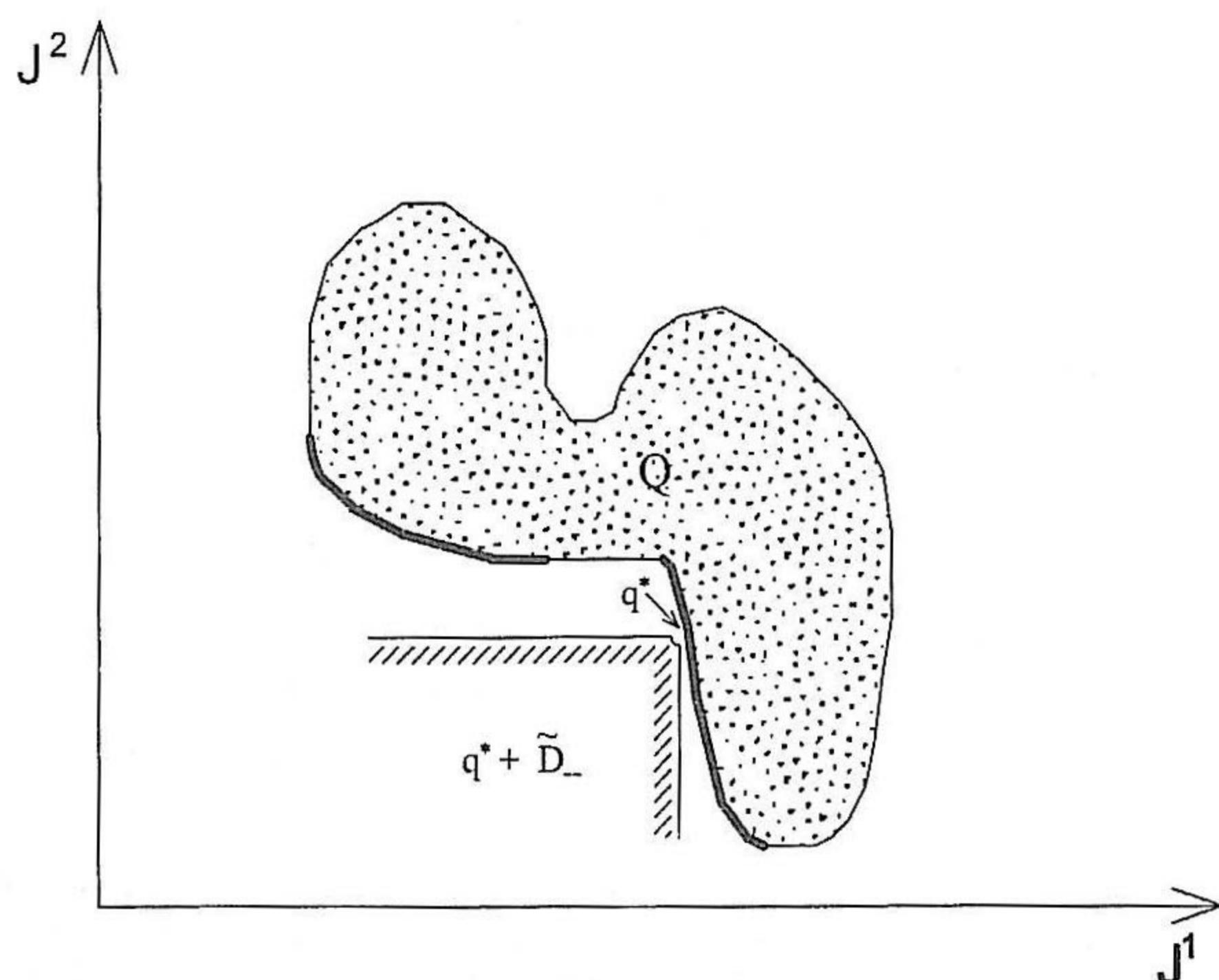
Stożki te są przedstawione na Rys. 3, odpowiadające im zbiory Parety zaś na Rys. 4, 5.

B. Analiza zadania.

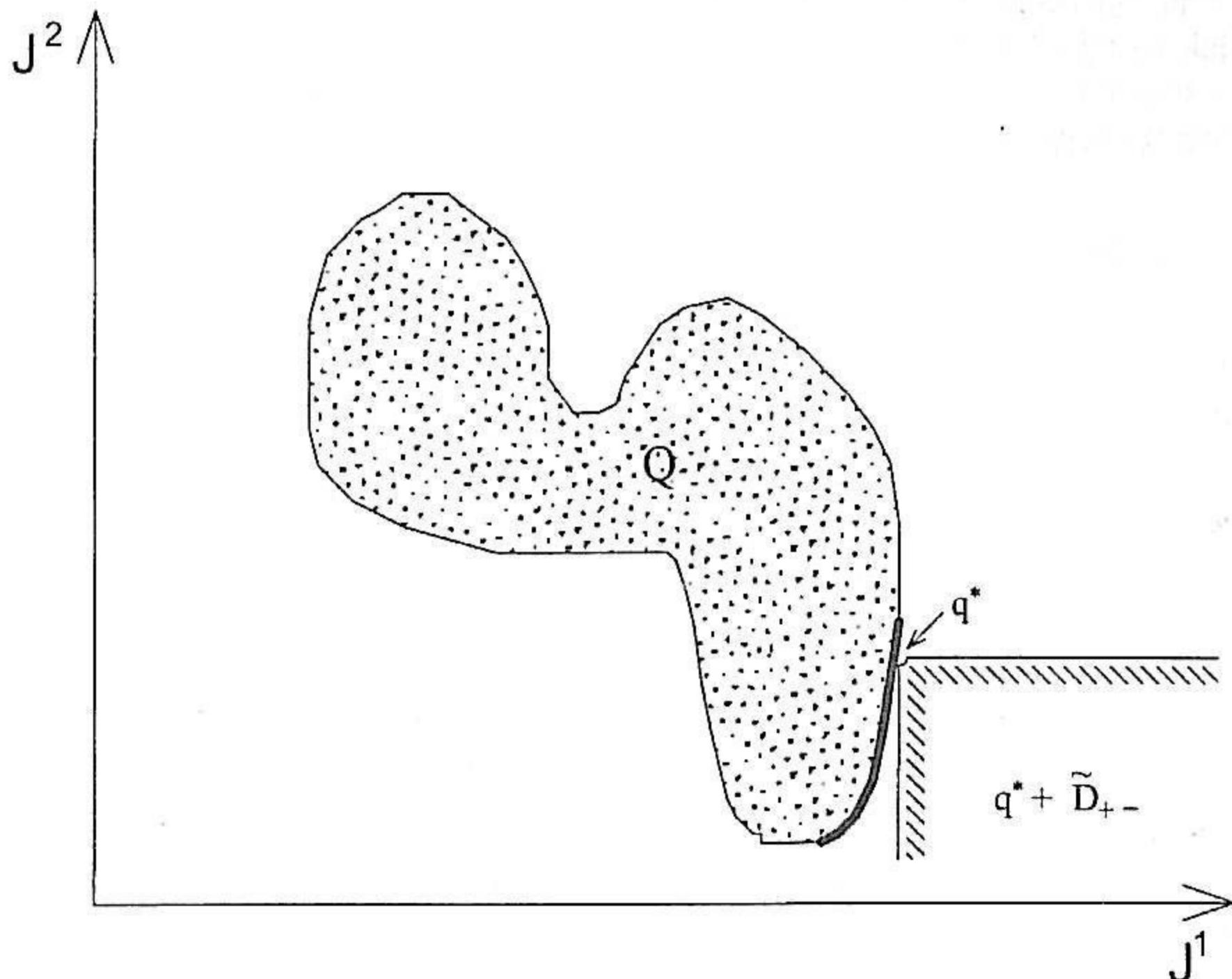
Zadanie przedstawione w poprzednim rozdziale, czyli (57) lub (58), gdzie operator V jest zdefiniowany przez (56), a funkcja φ_π przez (55), (2) z dynamiką opisaną przez równanie stanu (1), jest trudne do rozwiązania ze względu na występowanie kwadratu wartości oczekiwanej wskaźnika kosztów (zysków). Z tego powodu jest ono nieseparowalne w czasie, czyli niespełniony jest warunek konieczny zastosowania algorytmu programowania dynamicznego, który to algorytm, jak wcześniej stwierdziliśmy, jako jedyny zapewnia wyznaczenie optymalnej polityki sterowania w ogólnym zadaniu sterowania układem dynamicznym z niepewnością.



Rys. 3. Stożki kierunków poprawy w przypadku minimalizacji dwóch kryteriów oraz maksymalizacji pierwszego i minimalizacji drugiego.



Rys. 4. Zbiór punktów optymalnych w sensie Parety $e_{\tilde{D}_{--}} Q$ (zaznaczony grubą linią) w przypadku minimalizacji dwóch kryteriów.



Rys. 5. Zbiór punktów optymalnych w sensie Parety $e_{\tilde{D}_{+-}} Q$ (zaznaczony grubą linią) w przypadku maksymalizacji pierwszego i minimalizacji drugiego kryterium.

Pierwszym pomysłem jaki się tutaj nasuwa jest wprowadzenie dodatkowej współrzędnej stanu y_k , zliczającej dotychczasowe koszty (zyski). Będzie ona opisana równaniem:

$$y_{k+1} = y_k + g_k(x_k, u_k, w_k), \quad k = 0, 1, 2, \dots, N-1 \quad (62)$$

z warunkiem początkowym:

$$y_0 = 0 \quad (63)$$

To jednak nie wystarczy, bo nadal występuje kwadrat wartości oczekiwanej

$$E\{y_N + g_N(x_N)\} \quad (64)$$

Poniżej przedstawimy dwie metody, które pomogą rozwiązać to zadanie.

C. Metoda oparta na funkcji Lagrange'a.

Dla ustalenia uwagi przyjmijmy, że mamy do czynienia z zadaniem (58), czyli zależy nam na maksymalizacji wartości oczekiwanej wskaźnika jakości φ_π i minimalizacji jego wariancji.

Stosując skalaryzację poprzez ważenie [2] otrzymujemy funkcję:

$$S(\pi, v_1, v_2) = v_1 \cdot E(\varphi_\pi) - v_2 \cdot V(\varphi_\pi) \quad (65)$$

gdzie

$$v_1 > 0, v_2 \geq 0, \quad v_1 + v_2 = 1 \quad (66)$$

zaś $S(\pi, v_1, v_2)$ będzie maksymalizowane względem π przy wszystkich ograniczeniach zadania jak wyżej (stąd znak "-" przed $V(\varphi_\pi)$).

Oznaczmy teraz $\lambda = \frac{v_2}{v_1}$ (zakładamy że wariancja jest uwzględniana, czyli $v_1 \neq 0$). Zadanie z funkcją (65)

$$\max_{\pi \in \Pi} S(\pi, v_1, v_2) \quad (67)$$

jest równoważne zadaniu

$$\max_{\pi \in \Pi} \{E(\varphi_\pi) - \lambda \cdot V(\varphi_\pi)\}, \quad \lambda \geq 0 \quad (68)$$

Zdefiniujmy teraz funkcję parametru λ , określającą wynik rozwiązania tego zadania

$$H(\lambda) = \max_{\pi \in \Pi} \{E(\varphi_\pi) - \lambda \cdot V(\varphi_\pi)\} \quad (69)$$

Oczywiście, zgodnie z (55) i (62)-(63)

$$\varphi_\pi = y_N + g_N(x_N) \quad (70)$$

Z definicji wariancji mamy dla każdej liczby rzeczywistej a :

$$E[(\varphi_\pi - a)^2] = V(\varphi_\pi) + [E(\varphi_\pi) - a]^2 \quad (71)$$

czyli dla dowolnego a rzeczywistego, $E[(\varphi_\pi - a)^2]$ jest równe lub większe niż $V(\varphi_\pi)$.

Ponadto dla

$$a^* = E(\varphi_\pi) \quad (72)$$

mamy

$$E[(\varphi_\pi - a^*)^2] = V(\varphi_\pi) \quad (73)$$

Rozpatrzmy wobec tego następujące zadanie ze zmodyfikowaną funkcją Lagrange'a:

$$L(\lambda, a) = \max_{\pi \in \Pi} \{E(\varphi_\pi) - \lambda \cdot E[(\varphi_\pi - a)^2]\}, \quad \lambda \geq 0 \quad (74)$$

co można zapisać jako:

$$L(\lambda, a) = \max_{\pi \in \Pi} E\{\varphi_\pi - \lambda \cdot (\varphi_\pi - a)^2\}, \quad \lambda \geq 0 \quad (75)$$

Biorąc pod uwagę (70) otrzymujemy zatem klasyczne zadanie Mayera, które można rozwiązać przy pomocy standardowego algorytmu stochastycznego programowania dynamicznego.

Problem polega na tym, że występuje tutaj parametr $a \in \mathfrak{R}$, i że tak naprawdę, to zależy nam na znalezieniu rozwiązania zadania (69).

Okazuje się na szczęście, że zadania (75) i (69) są powiązane ze sobą, a związek między nimi określa następujące twierdzenie:

Twierdzenie 3:

$$H(\lambda) = \max_{-\infty \leq a \leq \infty} L(\lambda, a), \quad \forall \lambda \geq 0 \quad (76)$$

Dowód: Ponieważ λ jest dodatnie, z (74),(71),(69) wnioskujemy, że dla dowolnego $a \in \mathfrak{R}$, $H(\lambda)$ jest większe lub równe $L(\lambda, a)$. Jednakowoż, jeżeli $\hat{\pi} \in \Pi$ jest optymalną polityką w zadaniu (69), wówczas

$$H(\lambda) = E(\varphi_{\hat{\pi}}) - \lambda \cdot V(\varphi_{\hat{\pi}}) \quad (77)$$

i stąd dla $a^* = E(\varphi_{\hat{\pi}})$ mamy $H(\lambda) = L(\lambda, a^*)$, skąd bezpośrednio wynika równość (76). ■

A zatem, ażeby wyznaczyć $H(\lambda)$ musimy po prostu rozwiązywać zadanie (76), gdzie $L(\lambda, a)$ jest wyliczone przy pomocy stochastycznego programowania dynamicznego, wielokrotnie dla różnych wartości a , aż się znajdzie maksimum $L(\lambda, a)$.

Zbiorem, w którym poszukujemy optymalnej wartości zmiennej a , nie musi być cała os liczbową \mathbb{R} , ale wystarczy rozpatrywać przedział

$$a = [a_{min}, a_{max}] \quad (78)$$

gdzie

$$a_{min} = \min_{\pi \in \Pi} E(\varphi_{\pi}) \quad (79)$$

$$a_{max} = \max_{\pi \in \Pi} E(\varphi_{\pi}) \quad (80)$$

D. Metoda wykorzystująca moment zwyczajny II rzędu.

W tej części będziemy tak jak poprzednio rozważać zadanie typu (58), z dodatkowym założeniem że:

$$E(\varphi_{\pi}) \geq 0 \quad (81)$$

Założenie to oznacza, że przy sprowadzaniu zadania typu (57) (z minimalizacją wartości oczekiwanej) do (58) (z maksymalizacją wartości oczekiwanej) nie wystarczy przemnożyć wyjściowy wskaźnik jakości przez -1 , ale trzeba go dodatkowo przesunąć o stałą wartość w stronę liczb dodatnich, czyli zadanie

$$\min E(\varphi_{\pi}) \quad (82)$$

zastępujemy zadaniem

$$\max E(-\varphi_{\pi} + c) \quad (83)$$

gdzie c jest tak dobrane, że

$$c - \varphi_{\pi} \geq 0, \quad \forall \pi \in \Pi \quad (84)$$

Zanurzymy teraz nasze zadanie ze wskaźnikiem (58) w ogólniejszej klasie zadań:

A:

$$\bigcup_{\tilde{D}_{+-}}^e \left\{ \begin{bmatrix} f(z) \\ g(z) - f^2(z) \end{bmatrix} \right\} \quad (85)$$

gdzie

$f(z), g(z)$ - ciągłe funkcje nieujemne

Z - zbiór zwarty

Będziemy rozważać również następujące zadanie:

B:

$$\bigcup_{\tilde{D}_{+-}}^e \left\{ \begin{bmatrix} f(z) \\ g(z) \end{bmatrix} \right\} \quad (86)$$

Rozwiązania powyższych zadań można scharakteryzować przy pomocy następujących twierdzeń:

Twierdzenie 4: Niech $[f(z^*), g(z^*) - f^2(z^*)]$ będzie punktem \tilde{D}_{+-} - optymalnym w zadaniu A ($[f(z^*), g(z^*)]$ w zadaniu B). Wówczas:

$$\begin{aligned} g(z^*) - f^2(z^*) &= \min_{\substack{z \in Z \\ f(z) = f(z^*)}} [g(z) - f^2(z)] \end{aligned} \quad (87)$$

Dowód: (Nie wprost)

Gdyby tak nie było, istniałby punkt $\tilde{z} \in Z$ taki że:

$$f(\tilde{z}) = f(z^*) \quad (88)$$

oraz

$$g(\tilde{z}) - f^2(\tilde{z}) < g(z^*) - f^2(z^*) \quad (89)$$

Ale wówczas otrzymalibyśmy z (89) i (88)

$$g(\tilde{z}) < g(z^*) \quad (90)$$

co oznaczałoby że z^* nie byłby \tilde{D}_{+-} optymalny w zadaniu A (B). ■

Twierdzenie 5: Założymy że funkcja

$$\begin{aligned} h(\epsilon) &= \min_{\substack{z \in Z \\ f(z) = \epsilon}} [g(z) - f^2(z)] \end{aligned} \quad (91)$$

gdzie $f(z) \geq 0$, $\forall z \in Z$ rośnie ściśle monotonicznie z ϵ .

Wówczas z^* jest \tilde{D}_{+-} optymalny w zadaniu A wtedy i tylko wtedy gdy z^* jest \tilde{D}_{+-} optymalny w zadaniu B.

Dowód: (nie wprost)

" \Rightarrow "

Założymy że istnieje punkt \tilde{z} lepszy niż z^* , tzn.

$$f(\tilde{z}) \geq f(z^*) \quad (92)$$

$$g(\tilde{z}) \leq g(z^*) \quad (93)$$

przy czym, przynajmniej jedna z powyższych nierówności jest ostra.

Ze względu na nieujemność funkcji f z (92) mamy

$$f^2(\tilde{z}) \geq f^2(z^*) \Rightarrow -f^2(\tilde{z}) \leq -f^2(z^*) \quad (94)$$

Dodając stronami (94) do (93) i uwzględniając fakt, iż przynajmniej jedna z nierówności (92), (93) jest ostra, otrzymujemy

$$g(\tilde{z}) - f^2(\tilde{z}) < g(z^*) - f^2(z^*) \quad (95)$$

co oznacza że z^* nie jest \tilde{D}_{+-} optymalny w zadaniu A.

" \Leftarrow "

Niech z^* będzie \tilde{D}_{+-} optymalny w zadaniu B. Oznacza to z Twierdzenia 4 iż $g(z^*) - f^2(z^*)$ jest minimalną wartością funkcji celu w zadaniu

$$\begin{array}{ll} \min & [g(z) - f^2(z)] \\ z \in Z & \\ f(z) = f(z^*) & \end{array} . \quad (96)$$

Z założenia o ścisłej monotoniczności wartości powyższej funkcji celu w punkcie optymalnym w zależności od $f(z^*)$ wynika, że nie może istnieć $\tilde{z} \in Z$, taki że

$$f(\tilde{z}) \geq f(z^*) \quad (97)$$

$$g(\tilde{z}) - f^2(\tilde{z}) \leq g(z^*) - f^2(z^*) \quad (98)$$

z przynajmniej jedną nierównością ostrą.

A zatem z^* jest \tilde{D}_{+-} optymalne w zadaniu A. ■

Z ostatniego twierdzenia wynika że zamiast rozwiązywać zadanie:

$$\mathop{\cup}\limits_{\tilde{D}_{+-}}^e \left\{ \begin{bmatrix} E(\varphi_\pi) \\ V(\varphi_\pi) \end{bmatrix} \right\} \quad (99)$$

można rozwiązać zadanie

$$\mathop{\cup}\limits_{\tilde{D}_{+-}}^e \left\{ \begin{bmatrix} E(\varphi_\pi) \\ E(\varphi_\pi^2) \end{bmatrix} \right\} \quad (100)$$

które jest dużo wygodniejsze, bo po zastosowaniu skalaryzacji poprzez ważenie współrzędnych otrzymujemy zadanie Bolzy (Mayera), czyli:

$$\max_{\pi \in \Pi} \{S(\pi, v_1, v_2) = v_1 \cdot E(\varphi_\pi) - v_2 \cdot E(\varphi_\pi^2) = E[v_1 \cdot \varphi_\pi - v_2 \cdot \varphi_\pi^2]\} \quad (101)$$

gdzie

$$v_1, v_2 \geq 0, v_1 + v_2 = 1,$$

które można rozwiązywać za pomocą standardowego algorytmu stochastycznego programowania dynamicznego.

Z metodą tą może być tylko jeden problem, związany z założeniem o monotoniczności funkcji $h(\epsilon)$ (91). Założenie to jest na pewno spełnione na przedziale będącym rzutem na przestrzeń wartości pierwszego kryterium zbioru Parety w zadaniu A. W przeciwnym bowiem razie, rzutowane punkty nie należałyby do zbioru rozwiązań efektywnych. Ale my nie znamy zbioru rozwiązań w zadaniu A, my go właśnie poszukujemy poprzez rozwiązanie zadania B. Nietrudno sobie wyobrazić sytuację, gdy punkt będący rozwiązaniem w B nie jest rozwiązaniem w A (wystarczy porównać ze sobą dwa punkty ze zbioru Parety w zadaniu B, ich współrzędne mogą być tak dobrane że jeden z nich będzie lepszy w zadaniu A). Ażeby te punkty wyeliminować, trzeba niestety wyznaczyć wszystkie rozwiązania zadania B (tzn. cały zbiór Parety), na przykład dobierając wszystkie kombinacje wag v_1, v_2 z przedziału $[0, 1]$. Potem należy policzyć dla argumentów (czyli polityk) wartości obydwu kryteriów w zadaniu A i narysować je. Te części krzywej, gdzie otrzymana funkcja rośnie monotonicznie będą stanowiły zbiór Parety w zadaniu A.

E. Wnioski

W niniejszej części pracy rozważaliśmy dwie metody prowadzące do etapowej separacji stochastycznych zadań dwukryterialnego sterowania optymalnego ze wskaźnikami typu wartość oczekiwana i wariancja funkcji kosztów (zysków) typu Bolzy.

Jak się okazało, obydwie metody wymagają dodatkowej zmiennej stanu zliczającej dotychczasowe koszty (zyski).

Zaletą drugiej metody (wykorzystującej moment zwyczajny II rzędu) jest to, iż wymaga ona jedynie pojedynczego rozwiązania stochastycznego zadania sterowania optymalnego (przy pomocy klasycznego algorytmu programowania dynamicznego) dla jednej kombinacji wag. Pierwsza metoda wymaga bowiem niezależnego rozwiązywania tego zadania w każdym kroku maksymalizacji funkcji Lagrange'a.

Z drugiej strony, pierwsza metoda dostarcza punktów które są Pareto - optymalne w zadaniu A, podczas gdy rozwiązania uzyskane przy pomocą metody drugiej stanowią jedynie dolny brzeg zbioru wyników w zadaniu A. Ze względu na to, w tym przypadku zbiór Parety musi być wyznaczony poprzez rozwiązanie zadań skalarnych dla wszystkich kombinacji wag i pewną procedurę uzupełniającą, np. przy pomocy graficznej reprezentacji funkcji $h(\epsilon)$.

Inną zaletą pierwszej metody jest łatwość jej uogólnienia na przypadek wielu par kryteriów wartość oczekiwana - wariancja, do czego druga metoda się nie nadaje.

VI. PODSUMOWANIE

Mechaniczne posługiwanie się kryterium wartości oczekiwanej, jak pokazano powyżej, może często prowadzić do podjęcia nierożważnych decyzji. Bardziej racjonalne jest używanie tego kryterium w połączeniu bądź z funkcją użyteczności wskaźników jakości, bądź z dodatkowym kryterium - wariancją wskaźnika jakości. W obydwu przypadkach w zadaniach wieloetapowych traci się separowalność, odzyskanie jej wymaga wprowadzenia dodatkowej zmiennej stanu. Optymalizacja bezpośrednio wartości oczekiwanej wskaźnika jakości oznacza obojętny, czyli lekceważący stosunek do ryzyka.

Choć obydwa te podejścia są obecnie stosowane [4], bardziej naturalne wydaje się podejście wielokryterialne, ponieważ funkcja użyteczności jest czymś sztucznym, wprowadzającym zbyt wcześnie do rozwiązywanego zadania element subiektywny. Pomijając wątpliwości dotyczące podstawowych założeń tej metody, samo uzyskanie funkcji użyteczności jest sprawą dość kłopotliwą. Ponadto, w podejściu wielokryterialnym, szczególnie przy wygodnym, interakcyjnym interfejsie programu obliczeniowego, łatwo jest ocenić ile się traci na wartości oczekiwanej, gdy się nieco zmniejszy ryzyko. Dzięki temu podejmowane decyzje są oparte na głębszym zrozumieniu problemu.

Praca została wykonana w ramach grantu KBN Nr 8 T11A 011 15.

LITERATURA

- [1] Bertsekas, D.P., *Dynamic Programming and Optimal Control*, Athena Scientific, Belmont, Mass., 1995.
- [2] Galas, z., Nykowski, I i Z. Żółkiewski, *Programowanie wielokryterialne*, Państwowe Wydawnictwo Ekonomiczne, Warszawa, 1987.
- [3] Karbowksi, A., Methods of determination of the optimal control policies for multireservoir systems with expected value-variance criteria in Bolza problems, *Acta Geophysica Polonica*, vol. XLIV, no. 3, str.287-300, 1996.
- [4] Luenberger, D.G., *Investment Science*, Oxford University Press, New York - Oxford, 1998.
- [5] Magiera, P., *Optymalizacja reguł decyzyjnych w zadaniach z parą kryteriów oceny typu: wartość oczekiwana - wariancja*, Praca dyplomowa magisterska, Instytut Automatyki i Informatyki Stosowanej, Politechnika Warszawska, Warszawa, 1998.

- [6] Markowitz, H.M., *Mean-Variance Analysis in Portfolio Choice and Capital Markets*, Basil Blackwell, New York, 1987.
[7] von Neumann, J. i O. Morgenstern, *Theory of Games and Economic Behaviour*, Princeton University Press, Princeton, N.J., 1944.

DODATEK

Transformacja zakłóceń markowskich, typu ARMA oraz scenariuszowego do postaci standardowej (niezależnych zmiennych losowych)

A. Zakłócenia w postaci procesów Markowa.

Modele zakłóceń w postaci procesów Markowa są bardzo użyteczne ponieważ:

- uwzględniają zależności między zakłóceniemi na kolejnych etapach, w związku z czym opisują lepiej rzeczywistość niż idealizowane modele zakładające niezależność między kolejnymi zakłoceniami; na przykład dopływ do zbiornika retencyjnego nie może być modelowany jako proces stochastyczny z zakłoceniami niezależnymi na kolejnych etapach, ponieważ przepływy w rzece nie zmieniają się w sposób nagły
- dobrze się komponują z opisem układu dynamicznego w postaci równania stanu

Właściwością procesów Markowa która pozwala na to jest ich "skończona pamięć" poprzednich zakłóceń z punktu widzenia rozkładu zakłóceń bieżących, to znaczy:

$$P(w_k | w_{k-1}, w_{k-2}, \dots) = P(w_k | w_{k-1}, w_{k-2}, \dots, w_{k-L}) \quad (102)$$

Możemy dokonać przekształcenia tego modelu do modelu podstawowego (1), (4) wprowadzając sztuczne zmienne stanu y_1, \dots, y_L opisane bardzo prostymi równaniami:

$$y_{1_k} = w_{k-1}$$

$$y_{2_k} = w_{k-2} \quad (103)$$

⋮

$$y_{L_k} = w_{k-L}$$

Przyjmując nowy wektor stanu $\tilde{x} = \begin{bmatrix} x \\ y \end{bmatrix}$, zamiast zakłóceń zależnych na kolejnych etapach, ze (102) otrzymujemy zakłocenia, które zależą jedynie od stanu bieżącego \tilde{x}_k :

$$p_k(w_k | \tilde{x}_k) \quad (104)$$

czyli uzyskujemy już postać standardową (4).

Na przykład, w przypadku skalarnego łańcucha Markowa opisanego za pomocą macierzy przejść:

$$P_k = [p_k^{ij}] \quad (105)$$

gdzie

$$p_k^{ij} = P(w_k = w^j | w_{k-1} = w^i) \quad (106)$$

zaś w^i, w^j są, odpowiednio, i -tym i j -tym poziomem zakłocenia, otrzymamy:

$$y_k = w_{k-1} \quad (107)$$

$$\tilde{x}_k = \begin{bmatrix} x_k \\ y_k \end{bmatrix} = \begin{bmatrix} x_k \\ w_{k-1} \end{bmatrix} \quad (108)$$

z etapowym rozkładem prawdopodobieństwa zakłóceń:

$$p_k(w_k = w^j | \tilde{x}_{2_k} = w^i) = p_k^{ij} \quad (109)$$

B. Modele ARMA.

W modelach ARMA bieżąca wartość procesu w_k zależy od kombinacji liniowej pewnej liczby jego poprzednich wartości oraz pewnej liczby poprzednich zakłóceń ζ_l , $l < k$ o rozkładach niezależnych. Dokładniej obowiązuje równanie:

$$w_{k+1} = a_0 w_k + a_1 w_{k-1} + \dots + a_r w_{k-r} + b_0 \cdot \zeta_k + b_1 \cdot \zeta_{k-1} + \dots + b_s \cdot \zeta_{k-s} \quad (110)$$

czyli

$$w_{k+1} = \sum_{l=0}^r a_l w_{k-l} + \sum_{l=0}^s b_l \zeta_{k-l} \quad (111)$$

Transformacja do równania stanu (1) z opisem zakłóceń (4) polega na wprowadzeniu dodatkowych zmiennych y_1, y_2, \dots, y_{q+1} , gdzie $q = \max(r, s)$ opisanych równaniami:

$$y_{1_{k+1}} = a_q w_k + b_q \zeta_k \quad (112)$$

$$y_{2_{k+1}} = y_{1_k} + a_{q-1} w_k + b_{q-1} \zeta_k \quad (113)$$

⋮

$$y_{q+1_{k+1}} = y_{q_k} + a_0 w_k + b_0 \zeta_k \quad (114)$$

Nietrudno sprawdzić, że przyjmując $a_i = 0$, $r < i < q$, $b_i = 0$, $s < i < q$, otrzymujemy:

$$y_{q+1_{k+1}} \equiv w_{k+1} \quad (115)$$

czyli

$$y_{q+1_k} \equiv w_k \quad (116)$$

a więc zakłócenie w_k jest otrzymane jako wyjście układu dynamicznego:

$$y_{k+1} = \phi(y_k, \zeta_k) \quad (117)$$

A zatem przyjmując

$$\tilde{x} = \begin{bmatrix} x \\ y \end{bmatrix} \quad (118)$$

otrzymujemy model (1), (4).

C. Zakłócenia typu scenariuszowego

W wielu zadaniach sterowania można dość dokładnie przewidzieć kilka możliwych trajektorii zakłóceń na całym horyzoncie sterowania $[0, N-1]$. Innymi słowy, możemy uzyskać zbiór:

$$\bar{W} = \left\{ \{\bar{w}_0^m, \bar{w}_1^m, \bar{w}_2^m, \dots, \bar{w}_{N-1}^m\}, \quad m = 1, 2, \dots, M \right\} \quad (119)$$

M różnych scenariuszy \bar{w} przebiegu zakłóceń. Najczęściej mają one przypisane prawdopodobieństwa wystąpienia:

$$p^m, \quad m = 1, 2, \dots, M \quad (120)$$

Aczkolwiek przy takim modelu niepewności kolejne wartości zakłóceń są ze sobą ścisłe związane, możemy nadal stosować algorytm programowania dynamicznego, zakładający niezależność zakłóceń na kolejnych etapach. Przekształcenie do postaci (1), (4) polega

na rozszerzeniu wektora stanu o zmienną y zapamiętującą indeks realizowanego scenariusza oraz zdefiniowaniu zmiennej losowej ζ wykorzystywanej do wyboru scenariusza na początku horyzontu sterowania zgodnie z rozkładem (120).

Dokładniej, dodatkowe równanie stanu będzie miało postać:

$$y_{k+1} = \zeta_k \quad (121)$$

zaś rozszerzony wektor stanu \tilde{x} będzie miał postać

$$\tilde{x} = \begin{bmatrix} x \\ y \end{bmatrix} \quad (122)$$

Rozkłady etapowych zakłóceń (4) będą zdefiniowane jak niżej:

$$p_0(\zeta_0) = p_0(\zeta_0|\tilde{x}_0) = \begin{cases} 1 & \text{z prawdopodobieństwem } p^1 \\ 2 & \text{z prawdopodobieństwem } p^2 \\ \vdots & \\ m & \text{z prawdopodobieństwem } p^M \end{cases} \quad (123)$$

$$p_k(\zeta_k|\tilde{x}_k) = p_k(\zeta_k|y_k) = \begin{cases} 1 & \zeta_k = y_k \\ 0 & \text{w przeciwnym razie} \end{cases}, \quad k = 1, 2, 3, \dots, N-1 \quad (124)$$

Przyjęliśmy zatem, że realizowany scenariusz wybierany jest w sposób losowy na początku horyzontu sterowania. Na następnych etapach czynnik losowy nie ma już żadnego wpływu na przebieg zakłóceń.

Aby można było zastosować powyższą transformację do równania stanu (1) należy je zastąpić następującym:

$$x_{k+1} = \tilde{f}_k(x_k, u_k, \zeta_k) \quad (125)$$

z

$$\tilde{f}_k(x_k, u_k, \zeta_k) = f_k(x_k, u_k, \bar{w}_k^{\zeta_k}) \quad (126)$$

gdzie funkcja $f_k(., ., .)$ pochodzi z pierwotnego równania stanu (1) zaś $\bar{w}_k^{\zeta_k}$ jest zgodnie ze (119) wartością zakłócenia na etapie k -tym według scenariusza ζ_k -tego.