Wstęp

Specyfika przewozu i przechowywania produktów szybko tracących wartość (w tym łatwo psujących się) polega na tym, że ich ilość lub wartość zmienia się w trakcie transportu. Występują różne przyczyny takiego stanu rzeczy, które można podzielić na dwie podstawowe grupy. Do pierwszej z nich należą te, które występują w sposób ciągły i zależą głównie od cech transportowanych dóbr. Najważniejszymi są właściwości fizyczne i chemiczne, zmiana wartości spowodowana upływem czasu oraz niezauważone wcześniej braki, które ujawniają się dopiero w momencie sprzedaży użytkownikowi końcowemu. Drugą grupę stanowią przyczyny występujące w sposób mniej regularny (choć ich regularność wzrasta wraz ze zwiększeniem rozmiaru działalności), spowodowane głównie działalnością człowieka. Do najważniejszych należą niewłaściwe warunki transportowania lub magazynowania oraz różnego rodzaju wypadki i przestępstwa (np. kradzieże lub uszkodzenia powstałe w wyniku czynów zabronionych). Warto zwrócić uwagę, że choć podział na dwie grupy jest istotny z punktu widzenia przeciwdziałania występowaniu różnych przyczyn, to nie ma wpływu na sposób modelowania matematycznego tego typu sytuacji. W dalszej części pracy wszystkie możliwe przyczyny będziemy więc traktować jednakowo.

Do dóbr szybko tracących wartość, omówionych w tej książce, należą: żywność (zwłaszcza świeże warzywa i owoce, produkty mleczne), produkty firm farmaceutycznych (zwłaszcza leki), substancje radioaktywne wykorzystywane w diagnostyce obrazowej, krew i produkty krwiopochodne, ale też odzież i inne towary, których wartość spada ze względu na przemijającą modę.

W modelach matematycznych ten spadek wartości odzwierciedlany jest poprzez mnożniki wprowadzone do warunków ograniczających opisujących przepływy (w ten sposób różne typowe zadania przepływów w sieciach stają się uogólnionymi zadaniami przepływów). Choć zmiana ta wpływa w niewielki sposób na postać samych modeli, to jest bardzo znacząca dla struktury rozwiązań odpowiednich zadań optymalizacyjnych. W miejscu lasów rozpinających pojawiają się bardziej skomplikowane struktury kombinatoryczne, co znacznie zwiększa poziom trudności zadań. Z tego względu istotne jest znalezienie efektywnych metod rozwiązywania jak najszerszej klasy zadań uogólnionych przepływów. Książka ta wychodzi naprzeciw tej potrzebie, dostarczając wielu algorytmów dla różnego typu nieliniowych i stochastycznych wersji uogólnionego zadania przepływu o minimalnym koszcie. Co istotne, w literaturze można znaleźć niewiele prac poświęconych nieliniowym i stochastycznym wersjom omawianych zadań (w szczególności metodom ich rozwiązywania).

Wobec powyższego głównym celem autora było przedstawienie modeli i efektywnych algorytmów dla nieliniowych i stochastycznych problemów przewozu dóbr szybko tracących wartość. Ze względu na niewielką liczbę prac na ten temat, osiągnięcie postawionego celu zapewnia nowatorski charakter niniejszego opracowania.

Mając na uwadze cel główny, można wyodrębnić trzy cele szczegółowe rozprawy: teoretyczny, empiryczny i aplikacyjny.

Celem teoretycznym był matematyczny opis problemu przewozów dóbr szybko tracących wartość i prezentacja algorytmów ich rozwiązywania. Cel ten został osiągnięty za pomocą prezentacji kilkunastu modeli matematycznych będących sformalizowanym zapisem różnych typów zadań związanych z przewozem takich dóbr. Zaprezentowano również liczne algorytmy rozwiązujące zadania o omawianej postaci. Osiągnięciu tego celu służyła również analiza zbieżności przedstawianych metod.

Celem empirycznym rozprawy była analiza efektywności przedstawionych algorytmów. Cel ten osiągnięto za pomocą licznych eksperymentów obliczeniowych, polegających na rozwiązaniu pewnej liczby losowo wygenerowanych zadań testowych. W eksperymentach wykorzystane zostały autorskie programy komputerowe napisane samodzielnie przez autora.

Celem aplikacyjnym było modelowanie konkretnych sytuacji decyzyjnych związanych z przewozami. Jego osiągnięciu posłużyło przedstawienie pewnej liczby rzeczywistych przykładów zaczerpniętych z literatury oraz zaprezentowanie przykładów obliczeniowych.

Celom podporządkowana została struktura rozprawy. W rozdziale 1 omówiono pokrótce możliwe przyczyny utraty wartości przez transportowane dobra, a także miejsce systemów wspomagania decyzji w systemach informacyjnych przedsiębiorstw oraz umiejscowienie algorytmów w tychże systemach. Dalsza część rozdziału została poświęcona towarom szybko tracącym wartość, trudnościom wiążącym się z optymalizacją ich przewozów i sposobom ich modelowania z użyciem uogólnionych przepływów. Przedstawiono też bardziej szczegółowo (w podrozdziale 1.5) zakres tematyczny książki. Cały rozdział 1 został oparty w większości na badaniach literaturowych przedmiotu.

Od tego miejsca większość treści książki stanowi już własny dorobek autora (poza oczywistymi, koniecznymi odwołaniami do literatury). W rozdziałach 2 i 3 omówiono modele oparte na sieciach jednowarstwowych i przedstawiono algorytmy rozwiązywania odpowiednich zadań optymalizacyjnych. W szczególności, w rozdziale 2 omówiono problemy z nieliniowymi (wypukłymi i niewypukłymi) kosztami oraz algorytmy ich rozwiązywania. W rozdziale 3 zaprezentowano z kolei jednowarstwowe problemy stochastyczne i opracowane dla nich algorytmy. W dalszej części omówiono bardziej zaawansowane modele. W rozdziale 4 modele i algorytmy dla sieci jednowarstwowych zostały stopniowo uogólnione na sieci dwuwarstwowe, wielowarstwowe i topologicznie uporządkowane sieci o dowolnym kształcie. W rozdziale 5 omówiono modele i algorytmy dla problemów wieloasortymentowych zarówno jednowarstwowych, jak i bazujących na sieciach logistycznych o dowolnej strukturze. Rozdział 6 zawiera modele i algorytmy dla problemów, w których oprócz kryterium minimalizacji kosztu lub oczekiwanego kosztu występują również dodatkowe kryteria – czasu albo ryzyka, mierzonego zmiennością oczekiwanego kosztu. Również w tym rozdziale omówiono problemy oparte zarówno na jednowarstwowych, jak i dowolnych sieciach logistycznych. W poszczególnych rozdziałach przedstawiono również wyniki eksperymentów obliczeniowych.

Rozdział 1

Problem badawczy

1.1. Wprowadzenie

Ilość transportowanych towarów stale wzrasta. Oczywistymi powodami są wzrost produkcji i konsumpcji. Istnieje jednak również inna przyczyna. Jeśli weźmiemy pod uwagę wielkość transportu w pewnym zakładanym okresie, ilość towarów zwiększa się także dlatego, że czas dostawy jest średnio coraz dłuższy. Jest to spowodowane przede wszystkim przez globalizację. Część firm lokalizuje swoją produkcję z dala od rynków zbytu, ponieważ pozwala to na optymalizację jej kosztów (dobrym przykładem są tu producenci elektroniki i ubrań, umieszczający swoje fabryki w Azji Południowo-Wschodniej, podczas gdy największymi rynkami zbytu są Ameryka Północna i Europa Zachodnia). Niektóre towary nie są dostępne w pewnych regionach (na przykład węgiel, gaz, ropa naftowa, metale i wiele rodzajów roślin). To sprawia, że łańcuchy logistyczne wydłużają się, bez względu na to, czy weźmiemy pod uwagę odległość geograficzną, czy czas dostawy.

Im dłuższe są łańcuchy logistyczne, tym większe jest ryzyko, że niektóre produkty utracą część lub całość swej wartości (na przykład na skutek uszkodzenia albo zmiany właściwości spowodowanej upływem czasu). Jest rzeczą ważną, aby przewidzieć takie sytuacje i uwzględnić je w planowaniu dystrybucji towarów. Naturalnym sposobem modelowania takich problemów jest użycie tzw. uogólnionych sieci (pojęcie to zostało szerzej wyjaśnione w podrozdziale 1.4).

W kolejnym podrozdziale omówiono bardziej szczegółowo tego typu straty występujące w łańcuchach logistycznych, a w ostatnim podrozdziale przedstawiono uogólnione przepływy, wraz z ich zastosowaniami, w szczególności w modelowaniu strat występujących podczas transportu.

1.2. Przyczyny utraty wartości przez transportowane dobra

Podczas planowania dostaw bardzo użytecznym założeniem jest to, że ilości towarów wysyłanych z punktów zaopatrzenia są takie same jak te dostarczane do punktów docelowych. Założenie to upraszcza zarówno modele, jak i algorytmy i było powszechnie stosowane przez badaczy. Jednak nie zawsze jest ono realne – istnieje wiele powodów tego, że ilości towaru zmieniają się podczas przepływu przez łańcuch logistyczny.

Między innymi wyróżnić można następujące przyczyny:

- przyczyny występujące w sposób ciągły i zależące głównie od cech transportowanych towarów:
  - fizyczne i chemiczne właściwości transportowanych dóbr,
  - zmiana wartości transportowanych dóbr spowodowana upływem czasu,
  - niezarejestrowane wcześniej braki,
- przyczyny występujące w sposób mniej regularny i wynikające z działalności człowieka:
  - niewłaściwe warunki transportowania lub magazynowania,
  - wypadki,
  - przestępstwa.

Właściwości fizyczne i chemiczne transportowanych towarów mogą powodować zmiany w ich ilości. W tym kontekście warto wymienić żywność, w szczególności produkty rolne, na przykład owoce i warzywa. Jest to raczej oczywiste, że ich jakość zmniejsza się w czasie. Nie jest możliwe, aby utrzymać je w dobrym stanie przez dłuższy czas, tak więc straty są nieuniknione. Niektóre przykłady łatwo psujących się produktów rolnych, które były analizowane przez naukowców, można znaleźć w artykule przeglądowym Ahumady i Villalobosa [2009] (Czytelników zainteresowanych bardziej szczegółowymi informacjami odsyłamy do prac cytowanych w tym artykule). Rong, Akkerman i Grunow [2011] rozwinęli model opisujący zmiany w jakości świeżej żywności. Zaprezentowali następujące równanie różniczkowe przedstawiające zależność poziomu jakości od upływu czasu:

$$\frac{dq}{dt} = kq^{\omega}. \tag{1.1}$$

W powyższym wzorze q oznacza poziom jakości, t czas, a k współczynnik degradacji, zależny od warunków transportu i przechowywania. Autorzy zauważyli, że przeważnie $\omega \in \{0,1\}$, więc jakość jest liniową albo wykładniczą funkcją czasu t. Yu i Nagurney [2013] przyjęli, że relacja jest wykładnicza (podobne założenie przyjęte zostało w niedawno wydanej książce Nagurney i współautorów [2013]). Zanoni i Zavanella [2012] analizowali strategie decyzyjne dla zrównoważonych łańcuchów dostaw żywności. W szczególności zadali pytanie, czy lepiej jest mrozić produkty spożywcze, czy przewozić świeże. Zastosowali podaną wyżej relację między jakością i czasem, ale również wzór na współczynnik degradacji k (który, ich zdaniem, nie musi być stały w czasie). Jednak bez względu na to, jaka jest forma funkcji q, jeśli ktoś zna czas dostawy (lub przynajmniej jego wartość oczekiwaną), może przewidzieć, jaka będzie (oczekiwana) zmiana jakości (lub wartości) wybranego produktu po przesłaniu go przez dowolną część łańcucha dostaw.

Innym przykładem towarów nietrwałych są preparaty promieniotwórcze stosowane w służbie zdrowia, niezbędne do obrazowania medycznego. Materiały radioaktywne ulegają rozpadowi promieniotwórczemu, co powoduje zmiany w ich ilości. Przykład takiego łańcucha dostaw można znaleźć w artykule Nagurney i Nagurneya [2012] oraz w książce Nagurney i współautorów [2013] wspomnianej już powyżej. W obu pracach autorzy rozważają produkcję i transport molibdenu-99 (^{99}Mo) i technetu-99m ($^{99\mathrm{m}}$Tc).

Do nietrwałych produktów należą również krew i preparaty krwiopochodne. Problem został niedawno opisany przez Nagurney i Masoumiego [2012], Nagurney, Masoumiego i Yu [2012] oraz Nagurney i współautorów [2013]. Autorzy przedstawili dwukryterialny model uwzględniający koszt i ryzyko. Krew była już od dawna jednym z nietrwałych produktów najbardziej interesujących badaczy, jak stwierdził na przykład Nahmias [1982]. Prawdą jest, że z czysto fizycznego punktu widzenia ilość krwi lub produktów krwiopochodnych nie zmienia się z upływem czasu. Jednak te produkty mają ściśle określoną datę ważności. Niektóre z tych produktów, których termin przydatności upłynął, muszą być uznane za niemożliwe do wykorzystania, co może być ostatecznie interpretowane jako zmniejszenie ich ilości.

Jeszcze innym przykładem produktów, których ilość zmienia się ze względu na cechy fizyczne i chemiczne, są farmaceutyki. Ten problem był badany na przykład przez Masoumiego, Yu i Nagurney [2012], a ostatnio także przez Nagurney i współautorów [2013]. Autorzy opracowali uogólniony sieciowy model oligopolu w celu wykorzystania modelu Cournota dla branży producentów leków. Bardzo rzadko ilość leku lub innego produktu farmaceutycznego może rzeczywiście się zmieniać, ale podobnie jak w przypadku produktów krwiopochodnych, przekroczenie daty ważności przez pewną część produktów jest równoznaczne ze zmniejszeniem ich całkowitej ilości.

Rozważmy jeszcze jeden przykład. W tym wypadku jest jeszcze bardziej widoczne, że ilość towaru się nie zmienia, natomiast zmniejsza się udział produktów, które mogą być uważane za pełnowartościowe. Tę grupę produktów stanowi modna odzież, badana na przykład przez Nagurney i Yu [2011; 2012], jak również przez Choi i Chiu [2012]. Dziś łańcuchy logistyczne, w szczególności w przypadku firm międzynarodowych, są znacznie wydłużone. Przemysł tekstylny jest dobrym przykładem – w większości ubrania są produkowane w Azji Południowo-Wschodniej, a największe rynki znajdują się w Europie Zachodniej i Ameryce Północnej. To sprawia, że czas przepływu przez łańcuch logistyczny od producenta do klienta należy liczyć w miesiącach, a czasem cały proces od pozyskania surowca do zakupu gotowego produktu zajmuje nawet więcej niż rok. W przypadku modnych strojów oznacza to w szczególności, że nawet jeśli producent jest w stanie przewidzieć, jakiego rodzaju produkt będzie modny w kolejnym sezonie, to gdy sezon się skończy, produkt będzie musiał być sprzedany poniżej normalnej ceny (na wyprzedaży). Jest to konieczne, ponieważ najprawdopodobniej nie będzie można sprzedać tego samego produktu w kolejnym sezonie (istotną rolę odgrywa tu również ogólna tendencja do redukcji kosztów magazynowania). Po raz kolejny zauważamy – ilość dobra się w tym wypadku nie zmienia, ale jego wartość spada, dzięki czemu sytuacja może być traktowana jako równoważna.

Inną przyczyną strat są niewłaściwe warunki transportu i przechowywania. Taki problem może dotyczyć produktu każdego rodzaju. Żywność i farmaceutyki muszą być transportowane w określonej temperaturze i wilgotności. Te warunki są często niespełnione ze względu na tendencję do redukcji kosztów. Również transport delikatnych produktów, takich jak elektronika czy szkło, wymaga specjalnych warunków i ostrożnego obchodzenia się z nimi. Często warunki te nie są spełnione – redukcja kosztów powoduje nacisk na skrócenie czasu dostawy, czego jednym z efektów są uszkodzenia i straty podczas transportu.

Niekiedy do sprzedawców detalicznych są dostarczane produkty z wadami, które nie zostały wykryte wcześniej. Często takie produkty trafiają potem do klientów. Te uszkodzone produkty wracają do producentów w postaci zwrotów, które mogą być traktowane jako straty – w tym wypadku ilość transportowanych dóbr się nie zmienia, ale ilość produktów pełnowartościowych nie jest taka sama, jak zakładano przed rozpoczęciem transportu. Firmy starają się zmniejszyć liczbę odrzutów poprzez wprowadzenie procedur zarządzania jakością, ale niemożliwe jest uniknięcie wszystkich problemów w tej dziedzinie.

Wszystkie wymienione rodzaje możliwych strat są dość łatwe do przewidzenia, a ich rozkład jest zwykle w przybliżeniu jednostajny, co oznacza, że łatwo jest przewidzieć odsetek braków, nawet w krótkim okresie, w przypadku stosunkowo niewielkich dostaw. Ponadto odsetek braków jest zazwyczaj stosunkowo niewielki (do kilku procent), jeżeli czas dostawy nie przekracza standardowego. Dwie przyczyny strat omówione poniżej są jednak znacząco inne – nie są tak częste i ich rozkład w czasie jest zwykle niemożliwy do przewidzenia, ale jeżeli wystąpią, to zazwyczaj strata wynosi 100% – całą dostawę należy wymienić lub ponowić. Te dwie przyczyny to wypadki i przestępstwa.

Łańcuchy logistyczne stają się coraz dłuższe, w szczególności ze względu na internacjonalizację i globalizację handlu. Istnieje kilka rodzajów wypadków lub przestępstw, które mogą wystąpić. W dalszej części pominiemy takie losowe zdarzenia występujące w zakładach produkcyjnych, gdyż jesteśmy głównie zainteresowani zmianami ilości towarów w czasie transportu.

Na początku i na samym końcu łańcucha logistycznego towary są zwykle transportowane samochodami. Dziesiątki wypadków samochodowych występują niemal codziennie w prawie wszystkich krajach świata. Po wielu z nich cała dostawa jest tracona i musi być zastąpiona nową. Ponadto przyczyną strat na tym etapie może być przestępstwo. Najbardziej oczywista jest kradzież – część dostawy może być skradziona na przykład z parkingu, ale często zdarza się też, że cały samochód zostaje skradziony (lub nawet porwany razem z kierowcą).

Wypadki kolejowe są tak rzadkie, że mogą być ignorowane. Istnieją jednak przestępstwa, które mogą zostać popełnione podczas transportu kolejowego, w szczególności w przypadku przewozu towarów luzem – na przykład w mediach są niekiedy podawane informacje, że pewna ilość węgla została skradziona z pociągu towarowego albo że protestujący rolnicy wysypali produkty rolne z wagonów.

W globalnych łańcuchach logistycznych istnieje również inna część podróży niebezpieczna dla ładunku – transport morski. Produkty są transportowane między kontynentami w kontenerach. Może się zdarzyć, że pojemnik spada ze statku – niektóre szacunki mówią o nawet 10 000 pojemników, które wypadły za burtę kontenerowców, w wyniku wysokich fal, nieprawidłowego rozmieszczenia, wypadków pożarowych i działań piratów [Waters 2007, s. 70]. Piraci porywają nawet całe statki (szczególnie złą sławą cieszą się piraci somalijscy, operujący na Oceanie Indyjskim). Zazwyczaj zostają one wykupione, ale zwykle zajmuje to dużo czasu i w tym czasie co najmniej część transportowanych produktów traci swoją wartość.

Jak wspomniano wcześniej – w ostatnich omówionych przypadkach trudne (lub niemożliwe) jest przewidzenie zdarzeń, a nawet ich rozkładu. Jednakże w największych firmach te szacunki się łatwiejsze, ponieważ wypadki i przestępstwa występują tam częściej. Oznacza to, że niekiedy można, w pewnym przybliżeniu, przewidzieć skalę strat i uwzględnić tę wiedzę w procesie decyzyjnym.

Podsumujmy ten podrozdział krótkim opisem wyników badań jakościowych dotyczących produktów szybko tracących wartość, przeprowadzonych niedawno w Polsce. Pełne wyniki można znaleźć w pracy Anholcera i Kawy [2017]. Prawie wszyscy uczestnicy badania przyznali, że ich firmy zostały dotknięte problemem produktów szybko tracących wartość. Lista przyczyn była w każdym przypadku dość podobna. Jednym z ważniejszych problemów (choć niezbyt często występującym) są kradzieże. Czasami nie tylko produkty, ale nawet całe ciężarówki zostały skradzione. Właściciele firm transportowych zazwyczaj wykupują ubezpieczenie, ale obejmuje ono w takim przypadku przeważnie tylko wartość produktów i środków transportu. Utrata klienta z powodu opóźnienia w dostawie nie jest brana pod uwagę, podobnie jak straty w wizerunku firmy. Równie ważną, być może najczęstszą, przyczyną strat jest niewłaściwe traktowanie produktów podczas transportu. Polscy przewoźnicy mają tu do czynienia głównie z dwoma rodzajami sytuacji. Produkt może zostać uszkodzony, jeśli nie jest odpowiednio traktowany w czasie transportu (polega to na przykład na nieostrożnym obchodzeniu się z paczkami zawierającymi produkty wrażliwe na uszkodzenia), ale również w wyniku niewłaściwego zabezpieczenia (na przykład, jeśli ładunek nie jest dobrze umocowany, może wypaść ze statku lub ciężarówki).

Innym powodem są wypadki drogowe. To również obejmuje odpowiednie ubezpieczenie, ale nie pokrywa ono wszystkich rodzajów strat będących konsekwencją wypadku.

Straty wynikające z opóźnień (np. straty w świeżej żywności lub materiałach medycznych) występują rzadko, ponieważ przedsiębiorstwa transportowe mają świadomość, że w takim wypadku bardzo często cała dostawa nie nadaje się do użytku. Co ciekawe, takie sytuacje wydają się bardziej znaczącym problemem dla małych przedsiębiorstw, ze względu na ich mniejszą elastyczność oraz brak możliwości zrekompensowania strat. Jest to również duży problem dla przedsiębiorstw przewożących towary poza terytorium Unii Europejskiej, w szczególności w krajach Europy Wschodniej, gdyż istotna część opóźnień wynika z przedłużających się postojów na granicach (niekiedy dochodzących do 2–3 dób).

Istnieją również pewne bardzo rzadkie sytuacje, gdy dwa produkty, które nie powinny być przewożone razem, z jakiegoś powodu są. Jednym z przykładów podanych przez uczestników badania był transport oleju posiadającego specyficzny zapach wraz z mąką. W takim przypadku mąka może być czasami niezdatna do użytku. Respondenci wymienili również zbyt wysokie temperatury przewozu i nieużywanie chłodni. Utrata części ładunku zdarza się zwłaszcza w okresie zimowym, kiedy ze względu na niskie temperatury panujące na zewnątrz przewoźnik decyduje się wykorzystać tańszy, zwykły pojazd, zamiast chłodni. Niestety, czasem zdarza się nagła zmiana pogody (szczególnie wówczas, gdy ładunek transportowany jest na południe Europy) i przewożona żywność ulega zniszczeniu. Jak widać, przyczyny podawane przez uczestników badania są bardzo zbliżone do tych wymienionych wcześniej, przytaczanych w literaturze.

Do dóbr szybko tracących wartość należą w szczególności:

- żywność, w szczególności świeże mięso, warzywa i owoce, produkty mleczarskie,

– promieniotwórcze materiały medyczne,
– krew i produkty krwiopochodne,
– farmaceutyki,
– modna odzież.

Modelowanie i optymalizacja łańcuchów dostaw powyższych (i innych) dóbr tracących wartość wymaga wykorzystania odpowiedniej wiedzy eksperckiej oraz zastosowania właściwych narzędzi informatycznych i matematycznych. Zostały one omówione w kolejnych dwóch podrozdziałach.

1.3. Systemy informacyjne

Według Laudonów [Laudon i Laudon 2006] system informacyjny może być, technicznie rzecz ujmując, zdefiniowany jako zbiór współzależnych składowych, które zbierają (lub odtwarzają), przetwarzają, przechowują i udostępniają informacje w celu wspierania podejmowania decyzji i wspomagania kontroli w organizacji. W rzeczywistości może być wykorzystywany nie tylko do wspierania procesu decyzyjnego i koordynowania kontroli działalności organizacji, ale także do analizy problemów, pozwalającej zrozumieć złożone sytuacje i tworzyć nowe produkty i usługi.

Systemy informacyjne są miejscem, w którym przechowywane są informacje na temat członków organizacji, jak również jej produktów i istotnych elementów wewnątrz i na zewnątrz niej. Informacje te nie są tym samym co surowe dane. Te ostatnie to fakty, które stanowią pewne zdarzenia zachodzące w organizacji i w jej otoczeniu. Informacje to te same dane, ale po pewnym przetworzeniu, pozwalającym na ich wizualizację w formie użytecznej dla członków organizacji. Innymi słowy, są to dane, które zostały w jakiś sposób zorganizowane i wzbogacone o dodatkowe przydatne interpretacje i wizualizacje.

Można wyróżnić formalne i nieformalne systemy informacyjne. W dalszej części tej książki będą nas interesować wyłącznie systemy formalne, czyli takie, które opierają się na szeregu ściśle określonych i ustalonych definicji i procedur na wszystkich etapach korzystania z danych: gromadzenia, przetwarzania i użytkowania. Nieformalne systemy są również ważne dla każdej organizacji, ale nie są przedmiotem naszego zainteresowania. Ponadto autor skupił się na systemach informacyjnych wspomaganych komputerowo, tj. formalnych systemach używających sprzętu komputerowego i oprogramowania w celu gromadzenia danych oraz tworzenia i rozpowszechniania informacji. Nie jest zainteresowany ręcznymi systemami wykorzystującymi papier. Dla uproszczenia, od tej chwili będziemy używać pojęcia „systemy informacyjne” tylko do określenia formalnych systemów informacyjnych wspomaganych komputerowo.

rozwiązań dopuszczalnych, bo to zbytnio upraszczałoby zadania, upodabniając je pod względem trudności do ich liniowych odpowiedników). Istotne znaczenie dla struktury rozwiązania, a więc również dla efektywności algorytmu, mają jednak wartości mnożników. Dzieje się tak nawet w wypadku zadań programowania liniowego (więcej na ten temat można znaleźć na przykład w artykule Anholcera i Kawy [2012]), a tym bardziej w przypadku zagadnień rozpatrywanych w tej pracy. Z tego względu, przy ich doborze kierowano się dwiema zasadami. Po pierwsze, ich wartości nie mogły być zbyt bliskie 1, bo to niwelowałoby ich wpływ i uniemożliwiło przeanalizowanie działania algorytmów dedykowanych uogólnionym zadaniom przepływu. Po drugie, powinny być jednak stosunkowo wysokie, gdyż zarówno z analizy dostępnej literatury, jak i z przeprowadzonych badań empirycznych [Anholcer i Kawa 2017] wynika, że typowy odsetek traconych dóbr jest dość niewielki. Wyniki badań ankietowych przeprowadzonych w polskich firmach zajmujących się przewozami (jeszcze nieopublikowane) potwierdzają słuszność wyboru przedziałów dla mnożników.

2.4. Metoda rzutowania gradientu

Metoda rzutowania gradientu była jedną z pierwszych metod rozwiązywania ogólnych zadań optymalizacji nieliniowej z ograniczeniami. Została ona opublikowana przez J.B. Rosena [1960; 1961]. Zbieżność metody udowodnili Du i Zhang [1986; 1989]. Dokładniej, autorzy udowodnili zbieżność nieco zmodyfikowanego algorytmu. Istnieją przykłady pokazujące, że odwzorowanie algorytmiczne oryginalnej metody Rosena nie jest domknięte. Również odwzorowanie algorytmiczne metody przedstawionej przez Du i Zhanga nie jest domknięte, ale autorzy zdołali udowodnić zbieżność, analizując ε-sąsiedztwa rozwiązań niespełniających warunków KKT. Metody (w obu wersjach) skrótowo przedstawili i omówili także Bazaraa, Sherali i Shetty [2006, podrozdział 10.5]. Ponieważ wszystkie problemy, które analizujemy w tej książce, mają liniowe ograniczenia, jesteśmy zainteresowani tylko wersją algorytmu przedstawioną w pracy Rosena [1960] z późniejszymi modyfikacjami (zobacz także [Bazaraa, Sherali i Shetty 2006, podrozdział 10.5]).

Załóżmy, że dane jest zadanie programowania nieliniowego postaci:

$$\min f(x) \tag{2.46}$$

p.w.

$$Ax \leq b, \tag{2.47}$$

$$Ex = b. \tag{2.48}$$

Metoda rzutowania gradientu może być zapisana w następujący sposób (algorytm 2).

Algorytm 2. Metoda rzutowania gradientu dla zadań programowania wypukłego z liniowymi ograniczeniami

Krok 1: *Wyznaczenie rozwiązania początkowego*
Wybierz dodatnią stałą $c > 0$. Podstaw $k \leftarrow 1$ i znajdź pierwsze rozwiązanie dopuszczalne x. Przejdź do kroku 2.

Krok 2: *Wyznaczenie macierzy M*
Podziel A i b na dwie części A_1 i A_2 (odpowiednio b_1 i b_2) w taki sposób, że $A_1x = b_1$ i $A_2x < b_2$. Macierz M składa się z wszystkich wierszy E i A_1. Przejdź do kroku 3.

Krok 3: *Sprawdzenie optymalności*
Jeżeli M jest pusta i $\triangledown f(x) = 0$, to STOP, x jest rozwiązaniem optymalnym. Jeżeli macierz M jest pusta i $\triangledown f(x) \neq 0$, przejdź do kroku 4. Jeżeli M jest niepusta, przejdź do kroku 5.

Krok 4: *Wyznaczenie kierunku poprawy – wariant 1*
Podstaw

$$d \leftarrow -\triangledown f(x). \tag{2.49}$$

Przejdź do kroku 9.

Krok 5: *Wyznaczenie kierunku poprawy – wariant 2*
Wyznacz macierz rzutowania, korzystając ze wzoru

$$P = I - M^{\mathrm{T}}(MM^{\mathrm{T}})^{-1}M. \tag{2.50}$$

Podstaw

$$d \leftarrow -P\triangledown f(x). \tag{2.51}$$

Przejdź do kroku 6.

Krok 6: *Wyznaczenie wektora zmiennych dualnych*
Wyznacz wektor zmiennych dualnych, korzystając ze wzoru

$$w = -(MM^{\mathrm{T}})^{-1}M\triangledown f(x). \tag{2.52}$$

Można go rozdzielić na dwa wektory v i u takie, że v odpowiada macierzy E, a u macierzy A_1. Przejdź do kroku 7.

Krok 7: *Sprawdzenie optymalności*
Jeżeli $u \geq 0$ i $d = 0$, to STOP, x jest rozwiązaniem optymalnym. Jeżeli $u \geq 0$ i $d \neq 0$, przejdź do kroku 9. Jeżeli u ma ujemną składową u_j, przyjmij $u_h = \min_j\{u_j\}$. Jeżeli $||d|| \leq |u_h|c$, przejdź do kroku 8, w przeciwnym razie przejdź do kroku 9.

Krok 8: *Wyznaczenie kierunku poprawy – wariant 3*
Usuń z macierzy A_1 wiersz h. Przekształć odpowiednio macierz M. Wyznacz nową macierz rzutowania, wykorzystując wzór (2.50), i nowy kierunek poprawy, korzystając ze wzoru (2.51). Przejdź do kroku 9.

Krok 9: *Wyznaczenie maksymalnej długości kroku*
Niech

$$d' = A_2 d. \tag{2.53}$$

Jeżeli $d' \leq 0$, podstaw

$$\lambda_{\max} = \infty. \tag{2.54}$$

W przeciwnym razie, niech

$$b' = b_2 - A_2 x \tag{2.55}$$

oraz

$$\lambda_{\max} = \min_i \{b'_i / d'_i \mid d'_i > 0\}. \tag{2.56}$$

Przejdź do kroku 10.

Krok 10: *Wyznaczenie optymalnej długości kroku i nowego rozwiązania*
Niech λ' będzie rozwiązaniem optymalnym zadania przeszukiwania liniowego

$$\min f(x + \lambda d) \tag{2.57}$$

p.w.

$$0 \leq \lambda \leq \lambda_{\max}. \tag{2.58}$$

Podstaw

$$x \leftarrow x + \lambda' d \tag{2.59}$$

i wróć do kroku 2.

Bardzo czasochłonną częścią każdej iteracji jest wyznaczenie macierzy P (2.50), kierunku poprawy d (2.51) i wektora zmiennych dualnych w (2.52), niekiedy przeprowadzane więcej niż jeden raz. W przypadku wypukłego NGTP obliczenia mogą być znacznie uproszczone, ponieważ struktura macierzy M, MM^{T} i $(MM^{\mathrm{T}})^{-1}$ zależy przede wszystkim od rozmiaru problemu, a nie od wartości parametrów i zmiennych. Poniżej udowodnimy odpowiednie twierdzenia. Wykorzystamy przy tym pomysły podobne do przedstawionych wcześniej przez Anholcera [2005a], choć odpowiednie macierze i kroki algorytmu wyglądają inaczej, ze względu na inne uporządkowanie zmiennych, a przede wszystkim inną postać zadania. Od teraz zakładamy, że zmienne są uporządkowane leksykograficznie rosnąco względem indeksów. Rozważmy następujące sformułowanie NGTP:

$$\min f(x) = \sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij}(x_{ij}) + \sum_{j=1}^{n} f_j\Big(\sum_{i=1}^{m} r_{ij}x_{ij}\Big), \tag{2.60}$$

p.w.

$$\sum_{j=1}^{n} x_{ij} \le a_i, \qquad i = 1,\ldots,m, \tag{2.61}$$

$$-x_{ij} \le 0, \qquad i = 1,\ldots,m; j = 1,\ldots,n. \tag{2.62}$$

Porządkujemy warunki względem rosnącej wartości i, a dla ustalonego i pierwszy jest odpowiedni warunek z grupy (2.61), a następnie warunki z grupy (2.62) uporządkowane według rosnącej wartości j.

Przy takim uporządkowaniu macierz E jest pusta, a macierz A ma postać

$$A = \left[\begin{array}{cccc|cccc|c|cccc}
1 & 1 & \ldots & 1 & 0 & 0 & \ldots & 0 & \ldots\ldots\ldots\ldots & 0 & 0 & \ldots & 0 \\
-1 & 0 & \ldots & 0 & 0 & 0 & \ldots & 0 & \ldots\ldots\ldots\ldots & 0 & 0 & \ldots & 0 \\
0 & -1 & \ldots & 0 & 0 & 0 & \ldots & 0 & \ldots\ldots\ldots\ldots & 0 & 0 & \ldots & 0 \\
\ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots\ldots\ldots\ldots & \ldots & \ldots & \ldots & \ldots \\
0 & 0 & \ldots & -1 & 0 & 0 & \ldots & 0 & \ldots\ldots\ldots\ldots & 0 & 0 & \ldots & 0 \\
\hline
0 & 0 & \ldots & 0 & 1 & 1 & \ldots & 1 & \ldots\ldots\ldots\ldots & 0 & 0 & \ldots & 0 \\
0 & 0 & \ldots & 0 & -1 & 0 & \ldots & 0 & \ldots\ldots\ldots\ldots & 0 & 0 & \ldots & 0 \\
0 & 0 & \ldots & 0 & 0 & -1 & \ldots & 0 & \ldots\ldots\ldots\ldots & 0 & 0 & \ldots & 0 \\
\ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots\ldots\ldots\ldots & \ldots & \ldots & \ldots & \ldots \\
0 & 0 & \ldots & 0 & 0 & 0 & \ldots & -1 & \ldots\ldots\ldots\ldots & 0 & 0 & \ldots & 0 \\
\hline
\ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots\ldots\ldots\ldots & \ldots & \ldots & \ldots & \ldots \\
\hline
0 & 0 & \ldots & 0 & 0 & 0 & \ldots & 0 & \ldots\ldots\ldots\ldots & 1 & 1 & \ldots & 1 \\
0 & 0 & \ldots & 0 & 0 & 0 & \ldots & 0 & \ldots\ldots\ldots\ldots & -1 & 0 & \ldots & 0 \\
0 & 0 & \ldots & 0 & 0 & 0 & \ldots & 0 & \ldots\ldots\ldots\ldots & 0 & -1 & \ldots & 0 \\
\ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots\ldots\ldots\ldots & \ldots & \ldots & \ldots & \ldots \\
0 & 0 & \ldots & 0 & 0 & 0 & \ldots & 0 & \ldots\ldots\ldots\ldots & 0 & 0 & \ldots & -1
\end{array}\right], \tag{2.63}$$

czyli może być zapisana jako

$$A = \left[\begin{array}{c|c|c|c|c} A_0 & \mathbf{0}_{n+1,n} & \mathbf{0}_{n+1,n} & \dots & \mathbf{0}_{n+1,n} \\ \hline \mathbf{0}_{n+1,n} & A_0 & \mathbf{0}_{n+1,n} & \dots & \mathbf{0}_{n+1,n} \\ \hline \mathbf{0}_{n+1,n} & \mathbf{0}_{n+1,n} & A_0 & \dots & \mathbf{0}_{n+1,n} \\ \hline \dots & \dots & \dots & \dots & \dots \\ \hline \mathbf{0}_{n+1,n} & \mathbf{0}_{n+1,n} & \mathbf{0}_{n+1,n} & \dots & A_0 \end{array}\right], \tag{2.64}$$

gdzie

$$A_0 = \left[\begin{array}{c} \mathbf{J}_{1,n} \\ \hline -\mathbf{I}_n \end{array}\right]. \tag{2.65}$$

Niech M_i będzie macierzą otrzymaną z i-tej kopii A_0 przez usunięcie wierszy odpowiadających niewiążącym warunkom. Dla każdego i niech

$$J_i = \{j \mid x_{ij} = 0\}, \tag{2.66}$$

oraz

$$n_i = |J_i|. \tag{2.67}$$

Jeżeli $\sum_{j=1}^{n} x_{ij} = a_i$, przyjmiemy $m_i = 1$, a w przeciwnym razie $m_i = 0$. Oznaczmy elementy zbioru J_i (rosnąco względem i) przez $j_i^1, j_i^2, \dots, j_i^{n_i}$. Każda z macierzy M_i ma $n_i + m_i$ wierszy i n kolumn. Jeżeli $\sum_{j=1}^{n} x_{ij} < a_i$ (czyli $m_i = 0$), macierz M_i ma n_i wierszy, a jej elementy są zdefiniowane za pomocą wzorów:

$$M_i[k,j] = -1, \qquad k = 1,\dots,n_i; j = j_i^k, \tag{2.68}$$

$$M_i[k,j] = 0, \quad k = 1,\dots,n_i; j = 1,\dots,n, j \neq j_i^k. \tag{2.69}$$

Jeżeli $\sum_{j=1}^{n} x_{ij} = a_i$ (czyli $m_i = 1$), macierz M_i ma $n_i + 1$ wierszy, przy czym pierwszy z nich ma postać $\mathbf{J}_{1,n}$, a pozostałe są takie same jak w poprzednim przypadku:

$$M_i[k,j] = 1, \qquad k = 1, \tag{2.70}$$

$$M_i[k,j] = -1, \qquad k = 2,\dots,n_i + 1; j = j_i^{k-1}, \tag{2.71}$$

$$M_i[k,j] = 0, \qquad k = 2,\dots,n_i + 1; j \neq j_i^{k-1}. \tag{2.72}$$

Łatwo zauważyć, że

$$M = \left[\begin{array}{c|c|c|c|c} M_1 & \mathbf{0}_{n_1+m_1,n} & \mathbf{0}_{n_1+m_1,n} & \ldots & \mathbf{0}_{n_1+m_1,n} \\ \hline \mathbf{0}_{n_2+m_2,n_1+m_1} & M_2 & \mathbf{0}_{n_2+m_2,n_3+m_3} & \ldots & \mathbf{0}_{n_2+m_2,n} \\ \hline \mathbf{0}_{n_3+m_3,n} & \mathbf{0}_{n_3+m_3,n} & M_3 & \ldots & \mathbf{0}_{n_3+m_3,n} \\ \hline \ldots & \ldots & \ldots & \ldots & \ldots \\ \hline \mathbf{0}_{n_m+m_m,n} & \mathbf{0}_{n_m+m_m,n} & \mathbf{0}_{n_m+m_m,n} & \ldots & M_m \end{array}\right], \tag{2.73}$$

oraz

$$MM^{\mathrm{T}} = \left[\begin{array}{c|c|c|c|c} B_1 & \mathbf{0}_{n_1+m_1,n_2+m_2} & \mathbf{0}_{n_1+m_1,n_3+m_3} & \ldots & \mathbf{0}_{n_1+m_1,n_m+m_m} \\ \hline \mathbf{0}_{n_2+m_2,n_1+m_1} & B_2 & \mathbf{0}_{n_2+m_2,n_3+m_3} & \ldots & \mathbf{0}_{n_2+m_2,n_m+m_m} \\ \hline \mathbf{0}_{n_3+m_3,n_1+m_1} & \mathbf{0}_{n_3+m_3,n_2+m_2} & B_3 & \ldots & \mathbf{0}_{n_3+m_3,n_m+m_m} \\ \hline \ldots & \ldots & \ldots & \ldots & \ldots \\ \hline \mathbf{0}_{n_m+m_m,n_1+m_1} & \mathbf{0}_{n_m+m_m,n_2+m_2} & \mathbf{0}_{n_m+m_m,n_3+m_3} & \ldots & B_m \end{array}\right], \tag{2.74}$$

gdzie dla każdego $i = 1, \ldots, m$, B_i jest macierzą kwadratową mającą $n_i + m_i$ wierszy i $n_i + m_i$ kolumn postaci

$$B_i = \left[\begin{array}{c|c} n\mathbf{I}_1 & -\mathbf{J}_{1,n_i} \\ \hline -\mathbf{J}^{\mathrm{T}}_{1,n_i} & \mathbf{I}_{n_i} \end{array}\right], \qquad \text{gdy } m_i = 1, \tag{2.75}$$

$$B_i = \mathbf{I}_{n_i}, \qquad \text{gdy } m_i = 0. \tag{2.76}$$

Twierdzimy teraz, że MM^{T} ma następującą postać:

$$(MM^{\mathrm{T}})^{-1} = \left[\begin{array}{c|c|c|c|c} C_1 & \mathbf{0}_{n_1+m_1,n_2+m_2} & \mathbf{0}_{n_1+m_1,n_3+m_3} & \ldots & \mathbf{0}_{n_1+m_1,n_m+m_m} \\ \hline \mathbf{0}_{n_2+m_2,n_1+m_1} & C_2 & \mathbf{0}_{n_2+m_2,n_3+m_3} & \ldots & \mathbf{0}_{n_2+m_2,n_m+m_m} \\ \hline \mathbf{0}_{n_3+m_3,n_1+m_1} & \mathbf{0}_{n_3+m_3,n_2+m_2} & C_3 & \ldots & \mathbf{0}_{n_3+m_3,n_m+m_m} \\ \hline \ldots & \ldots & \ldots & \ldots & \ldots \\ \hline \mathbf{0}_{n_m+m_m,n_1+m_1} & \mathbf{0}_{n_m+m_m,n_2+m_2} & \mathbf{0}_{n_m+m_m,n_3+m_3} & \ldots & C_m \end{array}\right], \tag{2.77}$$

gdzie dla każdego $i = 1, \ldots, m$, C_i jest kwadratową macierzą mającą $n_i + m_i$ wierszy i $n_i + m_i$ kolumn postaci

$$C_i = \left[\begin{array}{c|c} \frac{1}{n-n_i}\mathbf{I}_1 & -\frac{1}{n-n_i}\mathbf{J}_{1,n_i} \\ \hline -\frac{1}{n-n_i}\mathbf{J}^{\mathrm{T}}_{1,n_i} & \mathbf{I}_{n_i} + \frac{1}{n-n_i}\mathbf{J}_{n_i,n_i} \end{array}\right], \qquad \text{gdy } m_i = 1, \tag{2.78}$$

$$C_i = \mathbf{I}_{n_i}, \qquad \text{gdy } m_i = 0. \tag{2.79}$$

Rzeczywiście, łatwo zauważyć, że dla każdego $i = 1, \ldots, m$ mamy

$$B_i C_i = I_{n_i + m_i}, \tag{2.80}$$

czyli

$$C_i = B_i^{-1}. \tag{2.81}$$

Używając odpowiednich przekształceń, stwierdzamy, że iloczyn $(MM^{\mathrm{T}})^{-1}M$ jest postaci

$$(MM^{\mathrm{T}})^{-1}M = \left[\begin{array}{c|c|c|c|c} D_1 & \mathbf{0}_{n_1+m_1,n} & \mathbf{0}_{n_1+m_1,n} & \ldots & \mathbf{0}_{n_1+m_1,n} \\ \hline \mathbf{0}_{n_2+m_2,n} & D_2 & \mathbf{0}_{n_2+m_2,n} & \ldots & \mathbf{0}_{n_2+m_2,n} \\ \hline \mathbf{0}_{n_3+m_3,n} & \mathbf{0}_{n_3+m_3,n} & D_3 & \ldots & \mathbf{0}_{n_3+m_3,n} \\ \hline \ldots & \ldots & \ldots & \ldots & \ldots \\ \hline \mathbf{0}_{n_m+m_m,n} & \mathbf{0}_{n_m+m_m,n} & \mathbf{0}_{n_m+m_m,n} & \ldots & D_m \end{array}\right], \tag{2.82}$$

Każda z macierzy D_i ma $n_i + m_i$ wierszy i n kolumn. Jeżeli $m_i = 0$, to $D_i = M_i$, zgodnie ze wzorami (2.68)–(2.69). Jeśli $m_i = 1$, to elementy macierzy D_i są zdefiniowane za pomocą wzorów ($k = 1, \ldots, n_i + 1; j = 1, \ldots, n$):

$$D_i[k, j] = \frac{1}{n - n_i}, \qquad k = 1, \ldots, n_i + 1, j \notin J_i \tag{2.83}$$

$$D_i[k, j] = 0, \qquad k = 1, j \in J_i, \tag{2.84}$$

$$D_i[k, j] = -1, \qquad k = 2, \ldots, n_i + 1, j = j_i^{k-1}, \tag{2.85}$$

$$D_i[k, j] = 0, \qquad k = 2, \ldots, n_i + 1, j \in J_i, j \neq j_i^{k-1}. \tag{2.86}$$

Rzeczywiście, dla każdego i mamy $C_i B_i = D_i$. Dalsze przekształcenia pozwalają stwierdzić, że

$$M^{\mathrm{T}}(MM^{\mathrm{T}})^{-1}M = \left[\begin{array}{c|c|c|c|c} F_1 & \mathbf{0}_{n,n} & \mathbf{0}_{n,n} & \ldots & \mathbf{0}_{n,n} \\ \hline \mathbf{0}_{n,n} & F_2 & \mathbf{0}_{n,n} & \ldots & \mathbf{0}_{n,n} \\ \hline \mathbf{0}_{n_3+m_3,n} & \mathbf{0}_{n_3+m_3,n} & F_3 & \ldots & \mathbf{0}_{n,n} \\ \hline \ldots & \ldots & \ldots & \ldots & \ldots \\ \hline \mathbf{0}_{n,n} & \mathbf{0}_{n,n} & \mathbf{0}_{n,n} & \ldots & F_m \end{array}\right], \tag{2.87}$$

gdzie dla każdego $i = 1, \ldots, m$, $F_i = B_i^{\mathrm{T}} D_i$ jest macierzą kwadratową mającą n wierszy i n kolumn opisanych wzorami:

$$F_i[k,j] = \frac{1}{n-n_i}, \qquad k \notin J_i, j \notin J_i, \tag{2.88}$$
$$F_i[k,j] = 1, \qquad k \in J_i, j = k, \tag{2.89}$$
$$F_i[k,j] = 0, \qquad \text{w pozostałych przypadkach}, \tag{2.90}$$

gdy $m_i = 1$,

$$F_i[k,j] = 1, \qquad k \in J_i, j = k, \tag{2.91}$$
$$F_i[k,j] = 0, \qquad \text{w pozostałych przypadkach}, \tag{2.92}$$

gdy $m_i = 0$. Ostatecznie otrzymujemy

$$P = \mathbf{I}_{mn} - M^{\mathrm{T}}(MM^{\mathrm{T}})^{-1}M = \left[\begin{array}{c|c|c|c|c} P_1 & \mathbf{0}_{n,n} & \mathbf{0}_{n,n} & \ldots & \mathbf{0}_{n,n} \\ \hline \mathbf{0}_{n,n} & P_2 & \mathbf{0}_{n,n} & \ldots & \mathbf{0}_{n,n} \\ \hline \mathbf{0}_{n_3+m_3,n} & \mathbf{0}_{n_3+m_3,n} & P_3 & \ldots & \mathbf{0}_{n,n} \\ \hline \ldots & \ldots & \ldots & \ldots & \ldots \\ \hline \mathbf{0}_{n,n} & \mathbf{0}_{n,n} & \mathbf{0}_{n,n} & \ldots & P_m \end{array}\right], \tag{2.93}$$

gdzie dla każdego $i = 1,\ldots,m$, $P_i = \mathbf{I}_n - F_i$ jest macierzą kwadratową mającą n wierszy i n kolumn, opisaną wzorami:

$$P_i[k,j] = 1 - \frac{1}{n-n_i}, \qquad k \notin J_i, j = k, \tag{2.94}$$
$$P_i[k,j] = -\frac{1}{n-n_i}, \qquad k \notin J_i, j \notin J_i, j \neq k, \tag{2.95}$$
$$P_i[k,j] = 0, \qquad \text{w pozostałych przypadkach}, \tag{2.96}$$

gdy $m_i = 1$,

$$P_i[k,j] = 1, \qquad k \notin J_i, j = k, \tag{2.97}$$
$$P_i[k,j] = 0, \qquad \text{w pozostałych przypadkach}, \tag{2.98}$$

gdy $m_i = 0$.

Zauważmy, że usunięcie któregokolwiek z wierszy macierzy M może jedynie zmienić wartość n_i lub m_i dla pewnego i, więc również elementy macierzy B_i, C_i, D_i, F_i i P_i, jednak ogólne wzory nie ulegają zmianie. Może się też zdarzyć, że M_i będzie pusta dla pewnego i (czyli $n_i = m_i = 0$ i M_i będzie miała 0 wierszy). Wówczas również B_i, C_i i D_i będą puste, podczas gdy $F_i = \mathbf{0}_{n,n}$ oraz $P_i = \mathbf{I}_n$. Powyższe rozważania pozwalają nam sformułować następujące twierdzenie.

Twierdzenie 2.8. *Załóżmy, że zmienne i ograniczenia w wypukłym NGTP zostały uporządkowane jak powyżej. Wtedy*

1. *Wzór (2.51) służący do wyznaczenia kierunku poprawy w krokach 5 i 8 algorytmu 2 przyjmuje uproszczoną postać:*

$$d_{ij} = -\frac{\partial f}{\partial x_{ij}}, \qquad m_i = 0, j \notin J_i, \tag{2.99}$$

$$d_{ij} = \left(\frac{1}{n-n_i} - 1\right)\frac{\partial f}{\partial x_{ij}} + \frac{1}{n-n_i}\sum_{k\in J_i, k\neq i}\frac{\partial f}{\partial x_{ik}}, \quad m_i = 1, j \in J_i, \tag{2.100}$$

$$d_{ij} = 0, \qquad \text{w pozostałych przypadkach} \tag{2.101}$$

(tutaj d_{ij} oznacza współrzędną wektora d odpowiadającą zmiennej x_{ij}).

2. *Wzór (2.52) służący do wyznaczenia wektora zmiennych dualnych w kroku 6 algorytmu 2 przyjmuje uproszczoną postać:*

$$w_{i0} = -\frac{1}{n-n_i}\sum_{k\notin J_i}\frac{\partial f}{\partial x_{ik}}, \qquad m_i = 1, \tag{2.102}$$

$$w_{ij} = \frac{\partial f}{\partial x_{ij}}, \qquad m_i = 0, j \in J_i, \tag{2.103}$$

$$w_{ij} = \frac{\partial f}{\partial x_{ij}} - \frac{1}{n-n_i}\sum_{k\notin J_i}\frac{\partial f}{\partial x_{ik}} = \frac{\partial f}{\partial x_{ij}} + w_{i0}, \qquad m_i = 1, j \in J_i \tag{2.104}$$

(tutaj $w_{ij}, j \in J_i$, jest współrzędną wektora w odpowiadającą zmiennej $x_{ij} = 0$, a w_{i0} odpowiada i-temu ograniczeniu nierównościowemu (2.61), o ile jest ono wiążące).

Uproszczone wzory (2.99)–(2.104) sprawiają, że algorytm jest istotnie szybszy i wymaga mniejszej ilości pamięci operacyjnej. Z tego powodu metoda ta może konkurować z algorytmem wyrównań przedstawionym w poprzednim podrozdziale. Niestety, metoda rzutowania gradientu dla wypukłego NGTP nie jest łatwo skalowalna (a więc nie można jej w prosty sposób uogólnić na zadania o bardziej złożonej strukturze), co czyni ją praktycznie bezużyteczną w przypadku zadań odpowiadających bardziej złożonym łańcuchom logistycznym, które rozpatrujemy w dalszych rozdziałach tej książki.

2.5. Rozwiązywanie zadań z niewypukłymi funkcjami celu

W dwóch poprzednich rozdziałach skupiliśmy się na problemach z wypukłymi kosztami. Jednak niekiedy konieczne jest rozważenie modelu z wklęsłą lub, bar-

Rozdział 6

Modele wielokryterialne

6.1. Wprowadzenie

Wszystkie modele rozpatrywane do tej pory miały jedną wspólną cechę: w każdym zadaniu optymalizacyjnym opartym na tych modelach występowała tylko jedna funkcja celu. W rzeczywistych zastosowaniach pojawiają się jednak czasem sytuacje, w których należy wziąć pod uwagę co najmniej dwa kryteria. Z taką sytuacją możemy mieć do czynienia również w przypadku zadań optymalizacji na sieciach (w szczególności w przypadku problemów zarządzania łańcuchami dostaw).

Chociaż głównym celem jest uzyskanie jednego rozwiązania zadania, jest bardzo ważne, aby znaleźć zbiór wszystkich rozwiązań optymalnych w sensie Pareta. Załóżmy, że mamy rozwiązać następujący problem optymalizacji wielokryterialnej:

$$\min\{f^1(x), f^2(x), \ldots, f^k(x)\}, \tag{6.1}$$

p.w.

$$x \in F, \tag{6.2}$$

gdzie F jest zbiorem rozwiązań dopuszczalnych. Rozwiązanie $y \in F$ jest optymalne w sensie Pareta, jeżeli nie ma innego rozwiązania $x \in F$, takiego że

$$f^i(x) \leq f^i(y), \qquad i = 1, \ldots, k, \tag{6.3}$$

$$\exists_{i \in \{1,\ldots,k\}} f^i(x) < f^i(y), \tag{6.4}$$

i jest słabo optymalne w sensie Pareta, jeżeli nie istnieje inne rozwiązanie $x \in F$, takie że

$$f^i(x) < f^i(y), \qquad\qquad i = 1, \ldots, k. \tag{6.5}$$

W literaturze można znaleźć wiele rodzajów warunków koniecznych i wystarczających optymalności w sensie Pareta. Na przykład Ben-Israel, Ben-Tal i Charnes [1977] przedstawili warunki konieczne i wystarczające dla problemów wielokryterialnych, w których wypukłe funkcje celu są minimalizowane na zbiorze wypukłym, a Bokrantz i Fredriksson [2014] określili warunki optymalności w sensie Pareta dla zadań wielokryterialnej optymalizacji odpornej. Różne warunki optymalności w optymalizacji wektorowej opublikowali Ivanov [2013], Li [1999] oraz Pop [2012]. Obszernym wstępem do teorii i metodologii nieliniowej optymalizacji wielokryterialnej jest praca Miettinena [1998], gdzie między innymi można również znaleźć różne warunki optymalności w sensie Pareta.

W tym rozdziale zajmiemy się dwoma modelami optymalizacji wielokryterialnej na sieciach, związanymi z łańcuchami dostaw towarów szybko tracących wartość. Optymalizacja nieliniowa na sieciach przyciągnęła nieco uwagi badaczy. Poniżej omówiono wybrane prace poruszające tę tematykę.

B. Gupta i R. Gupta [1983] przedstawili wariant metody sympleks dla wielokryterialnego programowania liniowego, który może być użyty (po pewnych oczywistych modyfikacjach) do rozwiązywania różnych problemów sieciowych. Z kolei Qiang i Nagurney [2012] analizowali wydajność łańcucha dostaw krytycznych w sytuacji występowania zakłóceń. Zdefiniowali kluczowe potrzeby, jak produkty i materiały, które są niezbędne dla zdrowia i życia ludzkiego. Przykłady obejmują żywność, wodę, leki i szczepionki. Do pomiaru wydajności autorzy zastosowali dwa kryteria, a mianowicie całkowite koszty i rozbieżność między popytem zaspokojonym a popytem rzeczywistym. Łańcuchy dostaw analizowane w tej pracy miały strukturę sieci czterowarstwowej. Inni autorzy – Raith i Ehrgott [2009] – rozwiązali dwukryterialne całkowitoliczbowe zadanie MCFP. Także inne rodzaje wielokryterialnych problemów sieciowych były obiektem zainteresowania badaczy. Na przykład Iori, Martello i Pretolani [2010] skoncentrowali się na wielokryterialnym problemie najkrótszej drogi (ang. *Shortest Path Problem*, SPP). Z kolei Zhang i Ong [2007] rozwiązali dwuryterialne uogólnione zadanie przydziału (ang. *Generalied Assignment Problem*, GAP). Ten wariant GAP można uznać za szczególną wersję GTP, w której każda dostawa musi być równa 0 lub 1 i podaż każdego źródła wynosi dokładnie 1. Przegąd algorytmów dla różnych wielokryterialnych zadań przepływu w sieci można znaleźć w artykule Hamachera, Pedersena i Ruziki [2007].

W kolejnym podrozdziale będziemy szczególnie zainteresowani problemem dwukryterialnym, w których jednym z kryteriów jest całkowity koszt, a drugim czas. Taki rodzaj problemów rozpatrywała między innymi Nagurney ze współautorami [Nagurney i Yu 2011; Nagurney i in. 2013]. Autorzy ci analizowali łańcuch dostaw modnej odzieży, gdzie sieć odpowiadająca łańcuchowi nie jest warstwowa. Aneja i Nair [1979] analizowali dwukryterialne TP, które odpowiada jednowarstwowemu łańcuchowi dostaw. Opracowali metodę znajdowania niezdominowanych punktów ekstremalnych w przestrzeni kryteriów. Ten sam problem był analizowany przez Nykowskiego [1986]. Z kolei Thirwani, Arora i Khanna [1997] rozwiązali dwukryterialne liniowe TP, gdzie jednym kryterium jest koszt całkowity ze stałą składową, a drugim czas. Ten sam problem analizowali w swojej pracy Basu, Pal i Kundu [1994]. Kolejny model dwukryterialny badali S. Puri i M.C. Puri [2006]. Autorzy rozpatrzyli liniowe TP z minimalizacją dwóch kryteriów – najkrótszego i najdłuższego czasu dostawy. Rozważali też zadanie z wklęsłą funkcją celu będącą sumą tych dwóch kryteriów. Kolejni autorzy – Basu i Acharya [2002] – rozwiązali uogólniony problem transportowy z kwadratową, ułamkową funkcją celu i zastosowali przedstawioną metodę do wyznaczenia krzywej czas–koszt. Na dwukryterialnym TP z kwadratową funkcją kosztów i kryterium czasu skupili się również Khurana i Arora [2011].

W wielu praktycznych zastosowaniach należy przyjąć, że parametry zadania nie są deterministyczne. Można ten problem ująć na różne sposoby. Przytoczmy kilka przykładów. Ojha, Mondal i Maiti [2011] analizowali rozmyte dwukryterialne TP z kryteriami czasu i kosztu, Gupta i Kumar [2012] rozwiązali wielokryterialne liniowe TP o rozmytych parametrach. Rozmyte dwukryterialne TP rozważali również Keshavarz i Khorram [2011]. Z kolei Hussein [1998] skoncentrował się na liniowym wielokryterialnym TP o posybilistycznych współczynnikach kosztów, a Li i Shi [2000] rozważali dynamiczne TP z wieloma kryteriami i wieloma poziomami ograniczeń. Ten sam problem był analizowany przez Shi [1995].

W tej książce jesteśmy zainteresowani w szczególności innym rodzajem niedeterministycznych problemów, a mianowicie tymi, których parametry są losowe. Wielokryterialne zadania z parametrami będącymi zmiennymi losowymi były badane przez wielu autorów. Na przykład Ben Abdelaziz i Masri [2010] przedstawili rozwiązanie dla stochastycznych problemów wielokryterialnych i zilustrowali je zadaniem planowania produkcji. Dwa lata później Muñoz i Ben Abdelaziz [2012] szukali rozwiązań zadowalających w kontekście programowania wielokryterialnego, a Ben Abdelaziz [2012] przedstawił przegląd różnych podejść do programowania wielokryterialnego z losowymi funkcjami celu i ograniczeniami oraz podał liczne przykłady, jednak nie omówił żadnych problemów sieciowych. W szczególności w obszarze zainteresowań wymienionych autorów nie znalazły się łańcuchy dostaw. Istnieją jednak również pewne artykuły, których autorzy zajmowali się tym tematem. Na przykład Mahapatra, Roy i Biswal [2013] studiowali STP z wielokrotnym wyborem, a Liu i Nagurney [2011] analizowali wpływ ryzyka i konkurencji na wydajność łańcuchów dostaw przedsiębiorstw zaangażowanych w działalność outsourcingową. Modele sieci łańcucha dostaw uwzględniały dwa kryteria: oczekiwany zysk i jego wariancję. Specjalnym rodzajem problemu sieciowego (z jednym źródłem i jednym ujściem) jest zadanie gazeciarza. Wariant stochastyczny tego problemu badali Choi i Chiu [2012]. Zdefiniowali oni model z dwoma kryteriami (średnią i wariancją) i zastosowali go do badania zrównoważonego handlu detalicznego na rynku modnej odzieży.

Jak już wspomniano, książka Miettinena [1998] jest dobrym źródłem wiedzy na temat nieliniowej optymalizacji wielokryterialnej. W szczególności można w niej znaleźć pewne użyteczne warunki optymalności w sensie Pareta. Niektórymi z nich będziemy szczególnie zainteresowani w dalszej części tego rozdziału.

Załóżmy, że zadanie (6.1)–(6.2) jest rozwiązywane metodą metakryterium, czyli po przypisaniu wagi $w_i \geq 0$ do każdego kryterium $f^i(x)$, $i = 1,\ldots,k$, $\sum_{i=1}^{k} w_i = 1$, rozwiązuje się następujące zadanie optymalizacji jednokryterialnej:

$$\min f(x) = \sum_{i=1}^{k} w_i f^i(x), \tag{6.6}$$

p.w.

$$x \in F. \tag{6.7}$$

Wówczas następujące twierdzenia są prawdziwe.

Twierdzenie 6.1 [Miettinen 1998, s. 78, twierdzenie 3.1.1]. *Rozwiązanie zadania (6.6)–(6.7) jest słabo optymalne w sensie Pareta.*

Twierdzenie 6.2 [Miettinen 1998, s. 78, twierdzenie 3.1.2]. *Rozwiązanie zadania (6.6)–(6.7) jest optymalne w sensie Pareta, gdy wszystkie wagi są dodatnie, czyli gdy $w_i > 0$ dla każdego $i = 1,\ldots,k$.*

Twierdzenie 6.3 [Miettinen 1998, s. 79, twierdzenie 3.1.3]. *Jedyne rozwiązanie zadania (6.6)–(6.7) jest optymalne w sensie Pareta.*

Odwrotne twierdzenia nie są prawdziwe w ogólnym przypadku. Jeżeli jednak jesteśmy zainteresowani zadaniami programowania wypukłego (czyli problemami, w których zarówno funkcja celu, jak i zbiór F są wypukłe), możemy wykorzystać następujący fakt.

Twierdzenie 6.4 [Miettinen 1998, s. 79, twierdzenie 3.1.4]. *Niech (6.1)–(6.2) będzie zadaniem programowania wypukłego. Jeżeli $x^\star \in F$ jest rozwiązaniem optymalnym w sensie Pareta, to istnieje wektor wag w ($w_i \geq 0, i = 1,\ldots,k, \sum_{i=1}^{k} w_i = 1$), taki że $x^\star$ jest rozwiązaniem zadania (6.6)–(6.7).*

Załóżmy teraz, że zadanie (6.1)–(6.2) jest rozwiązywane metodą jednego kryterium głównego, czyli że po przypisaniu pewnego ograniczenia ε_i do każdego kryterium $f^i(x)$, $i = 1,\dots,k$, następujące zadanie jest rozwiązywane dla każdego $l \in 1,\dots,k$:

$$\min f^l(x), \tag{6.8}$$

p.w.

$$f^i(x) \leq \varepsilon_i, \qquad i = 1,\dots,k, i \neq l, \tag{6.9}$$

$$x \in F. \tag{6.10}$$

Wówczas prawdziwe jest następujące twierdzenie.

Twierdzenie 6.5 [Miettinen 1998, s. 85, twierdzenie 3.2.2]. *Rozwiązanie $x^\star \in F$ jest optymalne w sensie Pareta wtedy i tylko wtedy gdy jest rozwiązaniem zadania (6.8)–(6.10) dla każdego $l = 1,\dots,k$, gdzie $\varepsilon_i = f^i(x^\star)$, $i = 1,\dots,k$, $i \neq l$.*

W kolejnym podrozdziale zajmiemy się dwoma modelami łańcucha dostaw (jednowarstwowym i o ogólnej strukturze), w których występować będą dwa kryteria: oczekiwany koszt i czas. W ostatnim podrozdziale kryteriami będą oczekiwany koszt i ryzyko.

6.2. Zadania z kryterium oczekiwanego kosztu i czasu

W tym podrozdziale rozważymy dwa modele, w których nie tylko oczekiwany koszt całkowity, ale również czas jest minimalizowany. Zaczniemy od najprostszego modelu, opartego na sieci jednowarstwowej. Ten fragment bazuje częściowo na artykule Anholcera [2013a]. Istotnym uogólnieniem jest to, że nie zakładamy, że funkcja gęstości popytu u odbiorcy j musi być dodatnia, czyli że $\varphi_j(x_j) > 0$ dla $j = 1,\dots,n$. Podstawowe fakty o losowym popycie i zasady zapisu zostały przedstawione w podrozdziale 3.1, nie będziemy ich więc powtarzać w tym miejscu.

Rozważmy jednowarstwowy łańcuch dostaw, jak w przypadku SGTP (por. podrozdział 3.1). Przyjmijmy, że poza kosztami jednostkowymi i mnożnikami do poszczególnych tras (łuków) przypisane są również czasy dostaw t_{ij} (dla każdego dostawcy i i odbiorcy j, $i = 1,\dots,m$; $j = 1,\dots,n$). Jeśli jakieś czasy (np. czasy załadunku lub wyładunku) są związane z dostawcami lub odbiorcami, dodajemy je do odpowiednich czasów dla łuków. Standardowo zakładamy, że dostawy mogą być wykonywane równocześnie, więc całkowity czas nie jest równy sumie, ale maksimum ze wszystkich czasów, którym odpowiadają realizowane dostawy. Omawiane zadanie będziemy określać mianem dwukryterialnego stochastycznego