WDWR projekt

Jakub Ostrzołek

Zadanie 1

Założenia

Biorąc pod uwagę poniższe fakty:

- koszty pracy przedsiębiorstwa (magazynowania) są niezależne od zmiennej losowej R (dochodów jednostkowych produktów),
- liczba produkowanych produktów jest niezależna od zmiennej losowej R,

można uprościć złożone zadanie poszukiwania wartości oczekiwanej łącznych zysków z pracy przedsiębiorstwa. Zamiast tego wystarczy niezależnie obliczyć wartość oczekiwaną zmiennej losowej R i za pomocą przekształceń matematycznych uzyskać wartość oczekiwaną zysków. Dzięki temu nie ma potrzeby generowania próbek z rozkładu R i uśredniania wyniku końcowego metodą numeryczną, a zamiast tego można obliczyć analitycznie $\mathbb{E}(R)$ i użyć gotowej wartości przy rozwiązywaniu zadania optymalizacji.

$$\begin{aligned}
\mathbb{E}(z) &= \mathbb{E}(d - k) \\
&= \mathbb{E}(d) - k \\
&= \sum_{p \in P} \sum_{n \in N} \mathbb{E}(x_{pn} \cdot R_p) - k \\
&= \sum_{p \in P} \sum_{n \in N} x_{pn} \cdot \mathbb{E}(R_p) - k
\end{aligned}$$

gdzie:

- z – łączne zyski z pracy przedsiębiorstwa w rozważanym czasie [zł],
- d – łączne dochody z pracy przedsiębiorstwa w rozważanym czasie [zł],
- k – łączne koszty pracy przedsiębiorstwa w rozważanym czasie [zł],
- $x_{pn} \quad p \in P, n \in N$ – liczba sprzedanych produktów p w miesiącu n [szt],
- $R_p \quad p \in P$ – jednostkowy dochód za produkt p (zmienna losowa) [zł/szt],
- $P = \{P1, P2, P3, P4\}$ – produkty,
- $N = \{1, 2, 3\}$ – rozpatrywane miesiące.

Wyznaczanie średnich jednostkowych dochodów dla każdego z produktów

Rozkład t-Studenta jest ciągły, więc wartość oczekiwana na przedziale domkniętym jest taka sama jak wartość oczekiwana na przedziale otwartym.

$$R_1 \sim Tt_{(5;12)}(9,16;4)$$
$$R_2 \sim Tt_{(5;12)}(8,9;4)$$
$$R_3 \sim Tt_{(5;12)}(7,4;4)$$
$$R_4 \sim Tt_{(5;12)}(6,1;4)$$

$$\mathbb{E}(R_1) = 9 + 4 \cdot \frac{\Gamma(3/2) \cdot ((4 + (-1)^2)^{-3/2} - (4 + (\frac{3}{4})^2)^{-3/2}) \cdot 4^2}{2(F_4(\frac{3}{4}) - F_4(-1))\Gamma(2)\Gamma(1/2)} = 8,63$$

$$\mathbb{E}(R_2) = 8 + 3 \cdot \frac{\Gamma(3/2) \cdot ((4 + (-1)^2)^{-3/2} - (4 + (\frac{4}{3})^2)^{-3/2}) \cdot 4^2}{2(F_4(\frac{4}{3}) - F_4(-1))\Gamma(2)\Gamma(1/2)} = 8,30$$

$$\mathbb{E}(R_3) = 7 + 2 \cdot \frac{\Gamma(3/2) \cdot ((4 + (-1)^2)^{-3/2} - (4 + (\frac{5}{2})^2)^{-3/2}) \cdot 4^2}{2(F_4(\frac{5}{2}) - F_4(-1))\Gamma(2)\Gamma(1/2)} = 7,61$$

$$\mathbb{E}(R_4) = 6 + 1 \cdot \frac{\Gamma(3/2) \cdot ((4 + (-1)^2)^{-3/2} - (4 + (6)^2)^{-3/2}) \cdot 4^2}{2(F_4(6) - F_4(-1))\Gamma(2)\Gamma(1/2)} = 6,42$$

Model rozwiązania

Zbiory:

- $P = \{P1, P2, P3, P4\}$ – produkty
- $M = \{SZ, WV, WH, FR, TO\}$ – maszyny (odpowiednio: szlifierki, wiertarki pionowe, wiertarki poziome, frezarki, tokarki)
- $G = \{G1, G2\}$ – grupy produktów, z których tylko jedną można magazynować w danym miesiącu
- $GP = \{(G1, P1), (G1, P2), (G2, P3), (G2, P4)\}$ – przypisania produktów do grup
- $N = \{1, 2, 3\}$ – rozpatrywane miesiące

Parametry:

- $n_m \quad m \in M$ – liczba dostępnych maszyn m [brak jednostki]
- $t_{mp} \quad m \in M, p \in P$ – jednostkowy czas produkcji produktu p na maszynie m [h/szt]
- $\mathbb{E}(R_p) \quad p \in P$ – średni jednostkowy dochód za produkt p [zł/szt] (wartości obliczone powyżej)
- $x_{pn}^{max} \quad p \in P, n \in N$ – maksymalna sprzedaż produktu p w miesiącu n [szt]

- $c^{mag} = 1$ – cena magazynowania jednostki produktu przez miesiąc [zł/szt]
- $m^{max} = 200$ – maksymalna liczba zmagazynowanych jednostek danego produktu na miesiąc [szt]
- m_p^{start} – liczba zmagazynowanych produktów p na start (na koniec grudnia) [szt]
- $h^{rob} = 24 \cdot 8 \cdot 2 = 348$ – liczba godzin roboczych w miesiącu [h]

Zmienne decyzyjne:

- $x_{pn} \quad p \in P, n \in N$ – liczba sprzedanych produktów p w miesiącu n [szt]
- $p_{pn} \quad p \in P, n \in N$ – liczba wyprodukowanych produktów p w miesiącu n [szt]
- $m_{pn} \quad p \in P, n \in (\{0\} \cup N)$ – liczba zmagazynowanych produktów p na koniec miesiąca n [szt]
- $u_{gn} \quad g \in G, n \in N$ – czy grupa produktów g jest magazynowana w miesiącu n (zmienna binarna: 0 – nie, 1 – tak)

Ograniczenia:

- $x_{pn} \geq 0 \quad \forall p \in P, n \in N$ – sprzedaż nieujemna
- $p_{pn} \geq 0 \quad \forall p \in P, n \in N$ – produkcja nieujemna
- $m_{pn} \geq 0 \quad \forall p \in P, n \in N$ – stan magazynu nieujemny
- $u_{gn} \in \{0, 1\} \quad \forall g \in G, n \in N$ – zmienna binarna
- $\sum\limits_{p \in P} p_{pn} \cdot t_{mp} \leq h^{rob} \cdot n_m \quad \forall m \in M, n \in N$ – łączny czas użycia maszyny m w miesiącu n nie przekracza liczby roboczych godzin
- $x_{pn} \leq x_{pn}^{max} \quad \forall p \in P, n \in N$ – sprzedaż produktu p nie przekracza rynkowego limitu na miesiąc n
- $m_{p0} = m_p^{start} \quad \forall p \in P$ – początkowy stan magazynu dla produktu p
- $\sum\limits_{g \in G} u_{gn} \leq 1 \quad \forall n \in N$ – w miesiącu n może być wybrana maksymalnie jedna grupa produktów g do magazynowania
- $\sum\limits_{\{p:(g,p) \in GP\}} m_{pn} \leq m^{max} \cdot u_{gn} \quad \forall g \in G, n \in N$ – produkt p należący do grupy g może być magazynowany maksymalnie w liczbie c^{max} szt, jeśli grupa g jest wybrana do magazynowania, lub w liczbie 0 szt w przeciwnym wypadku
- $p_{pn} + m_{p(n-1)} = x_{pn} + m_{pn} \quad \forall p \in P, n \in N$ – dla każdego miesiąca n i produktu p sztuki wyprodukowane i pozostałe w magazynach z poprzedniego miesiąca muszą zostać sprzedane lub powtórnie zmagazynowane

Cel:

- $\max \sum\limits_{n \in N} \sum\limits_{p \in P} (x_{pn} \cdot \mathbb{E}(R_p) - m_{pn} \cdot c^{mag})$ – maksymalizacja łącznego zysku, czyli różnicy dochodu ze sprzedaży produktów i wydatków na magazynowanie produktów na przestrzeni rozpatrywanych miesięcy (koszty magazynowania na miesiąc grudzień pominięte)

Wyniki działania modelu

Powyższy model został zaimplementowany w języku AMPL i uruchomiony przy użyciu solwera CPLEX. Implementacja znajduje się w plikach: `src/z1.{dat,mod,run}`. Poniżej wyniki działania.

Wartość funkcji celu:

$$\sum_{n \in N} \sum_{p \in P} (x_{pn} \cdot \mathbb{E}(R_p) - m_{pn} \cdot c^{mag}) = 14531 \text{ [zł]}$$

- $x_{pn}, p_{pn}, m_{pn} \quad p \in P, n \in N$ – liczba sprzedanych, wyprodukowanych i zmagazynowanych produktów p w miesiącu n [szt]
  - $n = 1$ (styczeń)

p	x_{p1}	p_{p1}	m_{p1}
P1	200	200	0
P2	0	0	0
P3	50	100	0
P4	150	200	0

  - $n = 2$ (luty)

p	x_{p2}	p_{p2}	m_{p2}
P1	300	300	0
P2	100	100	0
P3	200	200	0
P4	200	200	0

  - $n = 3$ (marzec)

p	x_{p3}	p_{p3}	m_{p3}
P1	0	0	0
P2	300	300	0
P3	100	100	0
P4	200	200	0

- $u_{gn} \quad g \in G, n \in N$ – czy grupa produktów g jest magazynowana w miesiącu n (zmienna binarna: 0 – nie, 1 – tak)

$g \setminus n$	1	2	3
G1	0	0	0
G2	1	1	1

Wnioski z wyników

- Ograniczenia na maksymalny obrót produktem zostały w całości wykorzystane.
- Żadne z ograniczeń na maksymalny czas użycia maszyn nie miało znaczenia, rzeczywiste wykorzystanie maszyn było zawsze dużo mniejsze niż limit.
- Z poprzedniego punktu wynika, że magazynowanie było zbędne (nie licząc stanu magazynu na koniec grudnia). Sama produkcja wysyciła limit na obrót każdym z produktów, więc nie było sensu dopłacać za magazynowanie produktów.
- Koszty produkcji są zerowe (brak magazynowania; koszty materiałów nie są rozważane w zadaniu).

Zadanie 2

Założenia

Tym razem model będzie rozważał również miarę ryzyka przy generacji rozwiązań. Nie uda się zatem uniknąć generacji próbek (scenariuszy) z rozkładu zmiennej losowej R jak w przypadku pierwszego zadania, bo nie byłoby sposobu na obliczenie wartości miary ryzyka.

Rozwiązywane zadanie jest wielokryterialne, należy zatem zastosować jedną z metod generacji rozwiązań efektywnych dla zadań wielokryterialnych.

Początkowym moim pomysłem było podejście progowe (ograniczenie ryzyka do kolejnych konkretnych wartości i maksymalizowanie wartości oczekiwanej zysków). Niestety takie podejście, o ile w tym przypadku zdawało się dawać zadowalające wyniki, to nie gwarantuje generacji wyłącznie rozwiązań efektywnych. Wynika to z faktu, że nic nie stoi na przeszkodzie, żeby wygenerowany punkt nie był zdominowany przez inny punkt o takiej samej wartości oczekiwanej profitu, ale nieco mniejszym ryzyku.

W związku z powyższym, w końcowym rozwiązaniu zastosowałem zamiast tego metodę punktu referencyjnego. Ma ona tę wadę, że parametry końcowego rozwiązanie będą się z dużym prawdopodobieństwem nieco różnić od zadanych progów, jednak w zamian wygenerowane rozwiązanie będzie z pewnością rozwiązaniem efektywnym.

Do implementacji metody punktu referencyjnego użyłem „inżynierskiej” wersji maksymalizacji leksykograficznej. Parametr ρ reguluje istotność drugiego kryterium (sumy) w tej metodzie.

Model rozwiązania

Rdzeń modelu rozwiązania będzie identyczny jak w przypadku zadania 1. Zmiany będą dotyczyły:

- obliczania wartości oczekiwanej dochodów,
- dodania elementów związanych z ryzykiem,
- zmiany funkcji celu na taką zgodną z metodą punktu referencyjnego dla kryteriów: zysku i ryzyka.

Zbiory:

- $S = \{1, 2, ..., 100\}$ – scenariusze

Parametry:

- $\mathbb{E}(R_p) \quad p \in P$ – parametr usunięty
- $R_{ps} \quad p \in P, s \in S$ – jednostkowy dochód za produkt p dla scenariusza s [zł/szt]
- ρ – istotność drugiego kryterium (sumy) w metodzie punktu referencyjnego [brak jednostki]
- a_r – wartość aspiracji dla ryzyka w metodzie punktu referencyjnego [zł]
- a_z – wartość aspiracji dla średnich zysków w metodzie punktu referencyjnego [zł]
- λ_r – istotność ryzyka w metodzie punktu referencyjnego [brak jednostki]
- λ_z – istotność średnich zysków w metodzie punktu referencyjnego [brak jednostki]

Zmienne:

- $r_s \quad s \in S$ – odchylenie zysków ze scenariusza s od średniej [zł]
- $r_s^+, r_s^- \quad s \in S$ – odpowiednio „nadmiar" i „niedobór" zysków ze scenariusza s w stosunku do średniej [zł]; służą do obliczenia wartości bezwzględnej we wzorze na odchylenie przeciętne
- $r^{śr}$ – wartość odchylenia przeciętnego zysków (miara ryzyka) [zł]
- $z_s \quad s \in S$ – łączny zysk dla scenariusza s [zł]
- $z^{śr}$ – wartość oczekiwana łącznego zysku [zł]
- f – minimum z wartości odchylenia: profitu lub ryzyka od aspiracji w metodzie punktu referencyjnego

Ograniczenia:

- $r_s^+, r_s^- \geq 0 \quad \forall s \in S$ – „nadmiary" i „niedobory" zysków są nieujemne
- $r_s = r_s^+ - r_s^-$ – odchylenie zysków ze scenariusza s od średniej składa się z „nadmiaru" i „niedoboru"
- $r_s = z^{śr} - z_s \quad \forall s \in S$ – obliczanie odchylenia zysków ze scenariusza s od średniej
- $r^{śr} = \sum\limits_{s \in S} (r_s^+ + r_s^-) \cdot \frac{1}{|S|}$ – obliczanie odchylenia przeciętnego; zakładam, że prawdopodobieństwa scenariuszy są jednakowe, dlatego we wzorze występuje dzielenie przez liczność scenariuszy

- $z_s = \sum_{n\in N}\sum_{p\in P}(x_{pn}\cdot R_{ps} - m_{pn}\cdot c^{mag}) \quad \forall s \in S$ – obliczanie łącznego zysku dla scenariusza s; analogicznie jak funkcja celu w zadaniu 1, jednak zamiast wartości oczekiwanej zmiennej losowej R jest wartość konkretnej próbki
- $z^{śr} = \sum_{s\in S} z_s \cdot \frac{1}{|S|}$ – obliczanie średniego łącznego zysku względem wszystkich scenariuszy
- $f \leq \lambda_z(z^{śr} - a_z)$ – wartość minimalna odchylenia w metodzie punktu referencyjnego jest mniejsza lub równa niż odchylenie średniego profitu od aspiracji
- $f \leq \lambda_r(a_r - r^{śr})$ – wartość minimalna odchylenia w metodzie punktu referencyjnego jest mniejsza lub równa niż odchylenie miary przeciętnego odchylenia od aspiracji; wartości w nawiasie są zamienione miejscami, bo przeciętne odchylenie jest minimalizowane

Cel:

- $lexmax\ (f, \lambda_z(z^{śr} - a_z) + \lambda_r(a_r - r^{śr}))$ – funkcja celu dla metody punktu referencyjnego: maksymalizacja w pierwszej kolejności minimum z odchyleń od aspiracji, a w drugiej kolejności sumy wszystkich odchyleń. Funkcja *lexmax* jest realizowana metodą „inżynierską” w następujący sposób:

$$\max\ f + \rho(\lambda_z(z^{śr} - a_z) + \lambda_r(a_r - r^{śr}))$$

Generacja scenariuszy

W celu otrzymania próbek zmiennej losowej R dla poszczególnych scenariuszy, znalazłem bibliotekę *Truncated Normal and Student's t-distribution toolbox* do programu *MatLab*, która pozwala na generację próbek z wielowymiarowego rozkładu t-Studenta z ograniczoną dziedziną. Kod generujący próbki znajduje się w pliku `src/generate_samples.m`, a wynik jego działania jest w `out/z2-samples.csv` (100 próbek).

Próbki należało również przekształcić do formatu `.dat` w celu odczytania przez AMPL, dlatego napisałem też skrypt `src/csv_to_dat_param.py`, który na podstawie pliku `.csv` generowanego z *MatLaba* tworzy plik `out/z2-samples.dat`.