

# 組合せ最適化を体系的に知って Python で実行してみよう

(株) 構造計画研究所 斉藤努

## 1 はじめに

これから、誰でも組合せ最適化が使えるようにご案内したいと思います。

組合せ最適化を使うコツは、全体像を理解することです。どんな要素があるのか、それらがどのような関係にあるのかを知ることにより、解きたい問題についてアプローチできるようになります。

**Python** を使うことにより、簡単に、様々な最適化ができます。実際のコードを見ながら、あとで確認しましょう。最初に、身近な最適化の例から見ていきます。

先日、あなたが実家に帰ると、お土産に野菜を持って帰るよう言われました。東京の野菜は高いので、「儲かった」と喜びましたが、量が多すぎます。せいぜい 5kg しか持って帰れないとします。(宅配便は使えないことにしてください) また、野菜は切ったりすると傷むので、そのまま持って帰ることにします。

あなたは「1kg あたりの販売価格の高いものから選んでいこう」と考えました。実は、これは(そこそこいい方法ですが)最適な方法ではありません。本当に最適な方法は、どうすればわかるでしょうか？

数理最適化を使えば、このような問題を解くことができます。

数理最適化では、数式を使って、問題を数理モデルで表します。先ほどの例は、

$$\begin{array}{ll} \text{最大化} & \sum_i p_i x_i \\ & \sum_i w_i x_i \leq 5 \\ & \forall x_i \in \{0, 1\} \end{array} \quad \begin{array}{l} p_i: \text{販売価格} \\ w_i: \text{重さ} \\ x_i: \text{持って帰るかどう} \end{array}$$

という数理モデルになります。数理最適化は、連続最適化と組合せ最適化に分けられます。

連続最適化: 連続要素のみ

組合せ最適化: 連続以外の離散要素を含む

今回は、組合せ最適化に焦点を当ててご説明します。

このように数理モデルを表すことを定式化するといいます。

## 2 定式化

### レッスン 1

組合せ最適化では、数理モデルを定式化する。

定式化をするには、3つの要素を決める必要があります。

1. 何を決めたいのか? 「持って帰る野菜を決めたいです」
2. どうなるとうれしいのか? 「持って帰る野菜の販売価格の合計が高くなるとうれしいです」
3. 守らないといけないことは? 「持って帰る野菜を 5kg 以下にします」

この3つをそれぞれ、変数、目的関数、制約条件とよびます。先ほどの例で見てみましょう。

$$\begin{aligned} \text{目的関数: } & \sum_i p_i x_i && \rightarrow \text{最大} \\ \text{制約条件: } & \sum_i w_i x_i \leq 5 \\ \text{変数: } & \forall x_i \in \{0, 1\} \end{aligned}$$

定式化できるようにするためには、慣れが必要です。今回は、いくつかの例を後で見ることになります。詳しく勉強する場合は、書籍「今日から使える!組合せ最適化」などを参考にしてください。

数理モデルから最適な答えを探すのは、(賢い人が作ってくれた) ツールを使います。このツールのことを最適化ソルバー、略してソルバーとよびます。つまり、問題を(誰でもわかるように) 表すだけで、ソルバーを使えば解(答えのこと)を出すことができます。

最近では、このソルバーが使いやすく、性能がたいへんよくなりました。みんなが、最適化が試せるようになってきているのです。

### 3 標準問題

先ほどの例の他に、どんな例があるでしょうか? あなたは実家に帰るときに、web で乗り換え駅を探索したかもしれません。

あなたの家の最寄り駅から、実家の最寄り駅までの最短路(あるいは最安路)を探したい。

これも最適化問題です。最適化問題は、いろいろなところにたくさんあります。でも、問題ごとに定式化するのは、大変ですよね。実は、別々の問題でも、同じ定式化になることがよくあります。そうすると、わざわざ定式化する必要がなくなり、便利です。世の中によくある問題を標準問題とよぶことにします。

#### レッスン 2

標準問題を理解する。

関連する標準問題を集めて、標準問題クラスという枠組みを作ります。標準問題クラスは、7つあります。

標準問題クラス	標準問題
グラフ・ネットワーク問題	最小全域木問題
	最短路問題
	最大流問題
	最小費用流問題
経路問題	運搬経路問題
	巡回セールスマン問題
集合被覆・分割問題	集合被覆問題
	集合分割問題
切出し・詰込み問題	ナップサック問題
	n 次元パッキング問題
割当・マッチング問題	一般化割当問題
	最大マッチング問題
配置問題	施設配置問題
スケジューリング問題	ジョブショップ問題
	スケジューリング問題

野菜の選び方はナップサック問題、乗り換え駅探索は、最短路問題といいます。標準問題は、よく研究もされているので、多くの場合、効率的な解法があります。あるいは、定式化がされているので、すぐ解くことができます。あとで、やってみましょう。

## 4 数理問題

最近、私がやっているコンテナの仕事のお話しをします。

世界中の人たちが、いろいろなものを安く買えるのはコンテナ輸送のおかげです。中国などで生産したものを日本やアメリカやヨーロッパに、大量に安く運べるからです。でも、空のコンテナが、どんどんたまります。また中国に戻さないといけません。いつ、どこからどこに戻すかを決めるのが、最小費用流問題になります。

ところが、最小費用流問題で表せない制約条件もあります。1つが、カボタージュとよばれるものです。カボタージュというのは、国内のみの輸送を自国業者のみに規制することです。

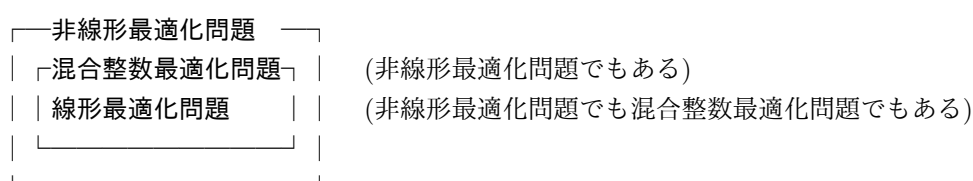
標準問題で対応できない場合は、数理モデルをそのまま扱います。数理モデルで表された問題を数理問題とよびます。標準問題は、数理問題として見ることもできます。

### レッスン 3

数理問題を理解する。

数理問題には、何があるのでしょうか？

数理問題の違いは、解法 (アルゴリズム) の違いによります。この解法の種類によって、解きやすさが全く違います。また、数理問題ごとにソルバーも異なることがあります。なるべく、簡単に解ける数理問題に持って行けるかが、重要になります。数理問題は、親子関係になります。親から見ていきましょう。



全ての最適化問題は、非線形最適化問題です。なぜ、わざわざ非線形最適化というのでしょうか？それは、線形最適化問題を非線形最適化として扱うのは非効率的だからです。非線形最適化という言葉には、非線形な目的関数か非線形な制約条件があることを暗黙に示唆しています。

### 線形最適化問題

数理問題で最も解きやすいのは、線形最適化問題です。最小費用流問題も線形最適化問題です。変数が連続変数で、目的関数と制約条件が線形 (1 次式) で表される問題です。これは、かなり大きな問題でも解けるようになってきました。

### 混合整数最適化問題

目的関数と制約条件は線形ですが、離散変数も許した問題を混合整数最適化問題とよびます。ナップサック問題や最短路問題も混合整数最適化問題です。

実務でよく現れる問題で、私の扱っている問題の多くは、これになります。難しい問題ですが、最近では、定式化や問題の規模次第で解けるようになってきました。

「混合整数最適化問題が解けるようになってきた」これこそが、最適化のしきいが下がってきた要因といえるでしょう。数理最適化初心者にとって、1つの目標が、簡単な混合整数最適化問題の定式化することになるでしょう。

ただし、実務で、混合整数最適化問題を扱うのは難しく、いろいろ工夫が必要になることもあります。

## 非線形最適化問題

非線形最適化は、さらに難しい問題です。しかし、小規模であれば、現状でも解けるようになってきました。また、非線形最適化の中でも非線形凸最適化は、大規模なものでも扱えますが、ここでは省略します。

この3種類の例は後ほど見てみます。

## 5 ビジネスでの実例

標準問題と数理問題を知ることが、組合せ最適化を使う上で重要になります。

ビジネスにおける最適化問題も、標準問題としてとらえることができます。標準問題に落とし込むことで、ビジネスの関係者間でのコミュニケーションの助けにもなります。

いろいろな事例と標準問題の対応を知ることが、標準問題を見つけるための訓練になるでしょう。

事例	標準問題クラス	標準問題
空箱の輸送コスト最適化	グラフ・ネットワーク問題	最小費用流問題
ビークル間連携配送最適化	経路問題	運搬経路問題
船舶スケジューリング最適化	集合被覆・分割問題	集合被覆問題
店舗シフトスケジューリング	スケジューリング問題	スケジューリング問題
3次元パッキング最適化	切出し・詰込み問題	n次元パッキング問題
避難施設配置最適化	配置問題	施設配置問題
大規模データベース配置	割当・マッチング問題	一般化割当問題
バス仕業作成最適化	割当・マッチング問題	最大マッチング問題

標準問題を数理問題として、とらえることも重要です。数理問題の違いによって、解きやすさが変わってくるからです。ここであげている事例の数理問題は、一部を除き混合整数最適化問題になります。ただし、「空箱の輸送コスト最適化」は近似的に線形最適化問題として解いています。また、「大規模データベース配置」は非線形最適化問題としています。

### 空箱の輸送コスト最適化

#### 背景

物を箱に入れて輸送しているときに、需要地と供給地に偏りがあると、空箱が需要地に溜まっていきます。この空箱を供給地に戻さないといけません。コンテナやパレットやレンタカーなど、いろいろな分野で見られます。

#### 問題

グラフ上において、複数の需要点と複数の供給点があります。需要点から供給点へ、費用を最小にするフロー(流量)を求めます。

- 標準問題の最小費用流問題となります。解法としては、負閉路除去法などがあります。
- 事例では、数理問題の線形最適化問題として定式化し、汎用ソルバーで解いています。理由は、追加の制約があること。混合整数最適化問題とすると、計算時間がかかりすぎることにあります。整数性を無視しているため、0.5個のような解が出てことがあります。

### ビークル間連携配送最適化

#### 背景

物資の保管所から避難場所に物資を運びたい。輸送手段は、陸海空全て使うことができる。また、フェリーなどの定期便も利用できる。

## 問題

品物、輸送元、輸送先、数量、輸送期限からなる配送オーダーを満たす配送計画を求めます。配送計画では、いつ、何を、どうやって運ぶかを決めます。オーダーは1回で運んでもよいし、複数の輸送手段で連携して運んでもよい。

- 標準問題の運搬経路問題となります。配送計画問題ともよばれます。運搬経路問題は、多くのソフトウェアが売られていますが、局所探索法をベースにした解法が用いられることが多いです。
- 事例では、貪欲法的一种である挿入法を用いています。数理問題の混合整数最適化問題になりますが、問題が大きすぎるので汎用ソルバーで解くことはできません。

## 船舶スケジューリング最適化

### 背景

船を使って、生産工場から倉庫に物資を輸送している。倉庫からは、定期的にトラックで需要地に運んでいる。

### 問題

倉庫が空にならないように、生産工場から倉庫へ輸送する船のスケジュールを求めます。

- 船のスケジュールの候補を採用するしないを変数とすることにより、標準問題の集合被覆問題とすることができます。解法としては、分枝限定法、動的最適化などがあります。
- 事例では、混合整数最適化問題に定式化して汎用ソルバーで解いています。

## 店舗シフトスケジューリング

### 背景

全国の店舗では、店長が毎月末に翌月の社員の勤務スケジュールを作成しています。様々な制約があるため、作成には時間がかかっています。

### 問題

各店舗ごとに、社員の勤務スケジュールを作成します。勤務スケジュールでは、社員ごと日ごとのシフトを決定します。シフトは、日勤、休み、早番、遅番などがあります。制約としては、各社員の休みの希望や各日ごとの最低シフト数やシフトの禁止パターンなどがあります。

- 標準問題のスケジューリング問題となります。
- 事例では、混合整数最適化問題として定式化し、近似解法の制約計画ソルバーを用いて解いています。

## 3次元パッキング最適化

### 背景

航空貨物のほとんどは、パレット上にパッキングして、航空機内の特殊な形状に合わせて詰め込んでいる。

### 問題

様々な大きさの貨物をパレット上に効率よく詰め込む方法を求めます。一番下に置かないといけなとか上に積んではいけないとか様々な制約があります。

- 標準問題の  $n$  次元パッキング問題となります。解法としては、貪欲法などがあります。
- 事例では、2次元の貪欲法を3次元に拡張した解法を独自に開発しています。

## 避難施設配置最適化

### 背景

津波に備えて、避難ビルを整備することにします。避難ビルには、物資を保管するため費用がかかります。

### 問題

全ての避難者を収容可能として、費用を最小にする避難ビルを選択します。

- 標準問題の施設配置問題となります。解法としては、分枝限定法などがあります。
- 事例では、混合整数最適化問題として定式化し、汎用ソルバーで解いています。目的関数として、全ての避難者の避難完了時刻を最小にすると、非常に難しい問題となります。その場合は、近似的に避難完了時刻を制約にして解くことがあります。

## 大規模データベース配置

### 背景

コールセンターでは、全国から多くの問い合わせが来ます。問い合わせに答えるため個人データがデータベースに格納されています。個人データは、地方ごとにまとめますが、保存先のストレージは、複数あります。

### 問題

個人データの保存するストレージ先を選択します。そのときに、ストレージごとのアクセスが平準化されるようにします。

- 標準問題の一般化割当問題の変形となります。
- 事例では、数理問題の非線形最適化問題とし、局所探索法で解きました。平準化問題を混合整数最適化問題用の汎用ソルバーで解くのは効率が悪くなります。逆に、局所探索法は平準化問題に対して効率よく解けます。

## バス仕業作成最適化

### 背景

バスのダイヤを組合わせて仕業を作成したい。1つの仕業は1人のドライバーが対応する。労働基準法も満足しなければならない。

### 問題

バスのドライバーの人数を最小になる仕業を作成します。

- 標準問題の最大マッチング問題となります。解法としては、エドモンズ法などがあります。
- 事例では、混合整数最適化問題として定式化し、汎用ソルバーで解いています。

## 6 Python による実行例

### レッスン 4

実際にやってみよう。

Python のインストールは、多くのパッケージが同封されている Anaconda をおすすめします。(8 章参照) 追加のインストールは、最適化用に pulp が必要です。pulp は「pip install pulp」でできます。

Python は、2.7 でも 3.4 でもどちらでも同じプログラムで動きます。

## ナップサック問題

### ナップサック問題

ナップサックに、いくつかの荷物を詰めます。詰込む荷物の容量 (size) の和がナップサックの容量 (capacity) を超えないように、荷物の価値 (weight) の和を最大にします。

「pip install knapsack」として、ナップサック問題を解くモジュールをインストールします。Python2.7では、さらに「pip install backports.functools\_lru\_cache」も必要です。

Listing 1: ナップサック問題 (動的最適化)

```
import knapsack
size = [21, 11, 15, 9, 34, 25, 41, 52]
weight = [22, 12, 16, 10, 35, 26, 42, 53]
capacity = 100
knapsack.knapsack(size, weight).solve(capacity)
>>>
(105, [0, 1, 3, 4, 5])
```

選択された荷物の価値の総和 (105) と選択した荷物の順番 ([0, 1, 3, 4, 5]) が得られます。

## 最短路問題

### 最短路問題

グラフ (7 章参照) において、始点から終点までの経路の中で最も短い経路を探します。

Listing 2: 最短路問題

```
import networkx as nx
g = nx.fast_gnp_random_graph(8, 0.26, 1)
nx.dijkstra_path(g, 0, 2)
>>>
[0, 1, 6, 3, 5, 2]
```

8 個の点からなるランダムなグラフを作成し、ノード 0 からノード 2 への最短路となるノードのリスト ([0, 1, 6, 3, 5, 2]) が得られます。

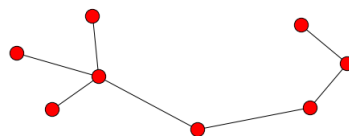


図 1: 最短路問題

## 最小費用流問題

### 最小費用流問題

グラフ (7 章参照) において、点ごとに需要量と供給量が与えられます。辺ごとにフローの上限である容量が決められています。供給点から需要点へ費用が最小となるフローを求めます。

Listing 3: 最小費用流問題

```
import networkx as nx
g = nx.fast_gnp_random_graph(8, 0.2, 1, True)
g.node[1]['demand'] = -2 # 供給
g.node[7]['demand'] = 2 # 需要
g.adj[2][7]['capacity'] = 1 # 容量
result = nx.min_cost_flow(g).items()
for i, d in result:
    for j, f in d.items():
        if f: print((i, j), f)
>>>
(1, 2) 2
(2, 3) 1
(2, 7) 1
(3, 7) 1
```

8 個の点からなるランダムなグラフを作成します。ノード 1 から 2 の供給、ノード 7 に 2 の需要、ノード 2 からノード 7 への辺の容量を 1 にします。

最小費用流となるフローの辞書が得られます。

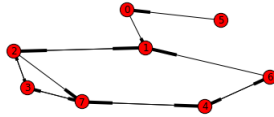


図 2: 最小費用流問題

## 最小費用流問題 (定式化)

最小費用流問題を定式化すると、次のようになります。

$$\begin{array}{l|l} \text{最小費用流問題} & \begin{array}{l} \text{最小化} \quad \sum_{e \in E} c_e x_e \\ \text{制約条件} \quad \sum_{e \in v_{out}} x_e - \sum_{e \in v_{in}} x_e = f(v) \quad \forall v \in V \\ 0 \leq x_e \leq u_e \quad \forall e \in E \end{array} \end{array}$$

ただし、 $c_e$  は辺の費用を、 $u_e$  は辺の容量を、 $f(v)$  は点の需要量 (正) または供給量 (負) を表しています。また、 $v_{out}$  は点  $v$  から出る辺の集合を、 $v_{in}$  は点  $v$  に入る辺の集合を表しています。

これを PuLP で表すと次のようになります。

Listing 4: 最小費用流問題 (定式化)

```
import networkx as nx
g = nx.fast_gnp_random_graph(8, 0.2, 1, True)
g.node[1]['demand'] = -2
g.node[7]['demand'] = 2
g.adj[2][7]['capacity'] = 1

from pulp import *
m = LpProblem()
x = {e:LpVariable('v%d'%i, 0, g.edge[e[0]][e[1]]\
```



```

        .get('capacity', None)) for i, e in enumerate(g.edges())}
m += lpSum(g.edge[e[0]][e[1]].get('weight', 1) * x[e] for e in g.edges())
for v in g.nodes():
    m += lpSum(x[e] for e in g.in_edges(v)) \
        - lpSum(x[e] for e in g.out_edges(v)) == g.node[v].get('demand', 0)
m.solve()
for e in g.edges():
    if value(x[e]): print(e, value(x[e]))
>>>
(1, 2) 2.0
(2, 3) 1.0
(2, 7) 1.0
(3, 7) 1.0

```

最小費用流問題は、数理問題の線形最適化問題になります。

## ナップサック問題 (定式化)

ナップサック問題を定式化すると、次のようになります。

$$\begin{array}{l|l}
 \text{ナップサック問題} & \begin{array}{l}
 \text{最大化} \quad \sum_{i \in N} w_i x_i \\
 \text{制約条件} \quad \sum_{i \in N} s_i x_i \leq c \\
 x_i \in \{0, 1\} \quad \forall i \in N
 \end{array}
 \end{array}$$

ただし、 $N$  は荷物の集合を、 $w_i$  は荷物の価値を、 $s_i$  は荷物の容量を、 $c$  はナップサックの容量を表しています。

これを PuLP で表すと次のようになります。

Listing 5: ナップサック問題 (定式化)

```

from pulp import *
size = [21, 11, 15, 9, 34, 25, 41, 52]
weight = [22, 12, 16, 10, 35, 26, 42, 53]
capacity = 100
m = LpProblem(sense=LpMaximize)
x = [LpVariable('v%d'%i, cat=LpBinary) for i in range(len(size))]
m += lpDot(weight, x)
m += lpDot(size, x) <= capacity
m.solve()
print(value(m.objective), [i for i in range(len(size)) if value(x[i]) > 0])
>>>
105.0 [0, 1, 3, 4, 5]

```

ナップサック問題は、数理問題の混合整数最適化問題になります。

## ポートフォリオ最適化問題

### ポートフォリオ最適化問題

銘柄を購入する割合 ( $x_i$ ) を決めます。投資に対する利益の下限 ( $t$ /円) を条件として、リスク (分散 ( $v_i$ )) を最小化します。ただし、各銘柄間は独立とします。

ポートフォリオ最適化問題を定式化すると、次のようになります。

ポートフォリオ最適化問題

$$\begin{array}{ll} \text{最小化} & \sum_{i \in N} v_i x_i^2 \\ \text{制約条件} & \sum_{i \in N} p_i x_i \geq t \\ & \sum_{i \in N} x_i = 1 \\ & 0 \leq x_i \leq 1 \quad \forall i \in N \end{array}$$

ただし、 $N$  は銘柄の集合を、 $p_i$  は銘柄の 1 円当たりの利益を表しています。

これを PuLP で表すと次のようになります。

Listing 6: ポートフォリオ最適化問題

```
import numpy as np, scipy.optimize as so
p = [31, 86, 29, 73, 46, 39, 58] # 利益 / 円
v = [10, 60, 25, 50, 35, 30, 40] # 分散 / 円
t = 50 # 目標利益 / 円
so.fmin_slsqp(lambda x: sum(v*x*x), np.zeros(len(p)),
              eqcons=[lambda x: sum(x) - 1], ieqcons=[lambda x: sum(p*x) - t])
>>>
Optimization terminated successfully. (Exit mode 0)
Current function value: 4.50899167487
Iterations: 14
Function evaluations: 136
Gradient evaluations: 14
array([ 0.26829785,  0.13279566,  0.09965076,  0.1343941 ,  0.11783349,
        0.11506705,  0.13196109])
```

ポートフォリオ最適化問題は、数理問題の非線形最適化問題になります。

## 7 グラフ・ネットワーク

グラフ (graph) とは、点 (vertex) と点同士を結ぶ辺 (edge) で構成される構造です。現実の関係を抽象的に表現するためによく用いられます。

$G = (V, E)$  とは、「グラフ ( $G$ ) が点集合 ( $V$ ) と辺集合 ( $E$ ) で構成される」ことを表します。

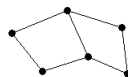


図 3: グラフの例

辺が一方通行のものを有向グラフ、そうでないものを無向グラフといいます。ある点から、別の点へ続く一連の辺を経路 (path) とといいます。経路の始点と終点が繋がっているものを閉路 (cycle) とといいます。

グラフ上の辺に何かしらの流れ (flow) を考えたものをネットワーク (network) とといいます。道路ネットワーク、通信ネットワークなどがあります。

## 8 ソフトのインストール

インストールが簡単にできる方法として、Anaconda を利用する方法をおすすめします。Anaconda のインストールは、「<http://continuum.io/downloads>」からインストーラをダウンロードして実行してください。

Anaconda には、Python 本体も含まれています。Anaconda 以外に必要なソフトは、数理最適化のモデラーである PuLP になります。PuLP のインストールは「`pip install pulp`」でできます。

また、PuLP 以外のパッケージについては、必要とするサンプルのところに掲載してあります。

## 9 余談

東ロボくんをご存じでしょうか？国立情報学研究所が中心となって、「ロボットは東大に入れるか」をテーマにした研究プロジェクトです。

実は、標準問題による最適化アプローチは、東ロボくんのアプローチと似ているところがあります。それは、千差万別の問題を類型問題としてとらえて解くというところです。

この試みを見ていると、日本語で問題を記述するだけで、自動で最適化して解けるようになる日が、いずれ来ると思われます。

しかし、それまでは自分で考えて、手を動かす必要があるでしょう。また、そうすることによって、徐々に、より複雑な問題への対処法を身につけることができると思います。