Discrete Optimization

Elizabeth Davenport, Megan Johnston, Calli Nguyen, Shuowen Wei
Non-Linear Optimization
April 18, 2013

1. An general introduction to discrete optimization:

What's discrete optimization? Well, discrete optimization is a very important part of optimization that mainly in applied mathematics and computer science area. As opposed to continuous optimization, the variables used in the some or all of the mathematical problems are restricted to discrete values, such as the integers. Two notable branches of discrete optimization are:

- combinatorial optimization, which refers to problems on graphs, matroids and other discrete structures
- integer programming[1]

Combinatorial optimization is a topic that consists of finding an optimal object from a finite set of objects[2].

An integer programming problem is a mathematical optimization in which some or all of the variables are restricted to be integers[3].

Some examples in discrete optimization are as follows:

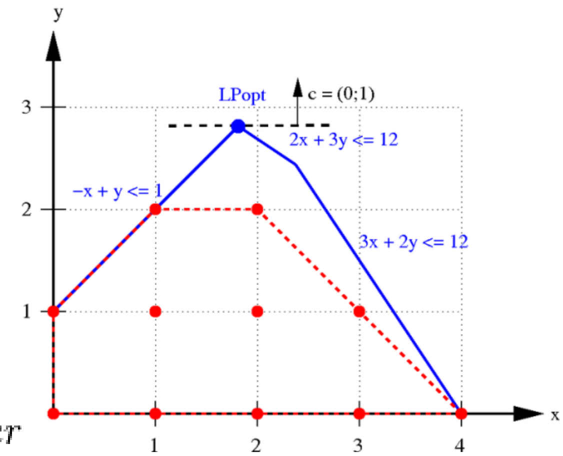
- Integer linear programming
- Set cover problem
- Knapsack problem
- Graph theory
  - Minimum spanning tree
  - Vertex cover problem
  - Traveling salesman problem (Hamiltonian circuit)
  - Shortest path problem
- Scheduling problem
  - Maximum flow problem

An integer linear programming problem can be typically expressed as follows[3]:

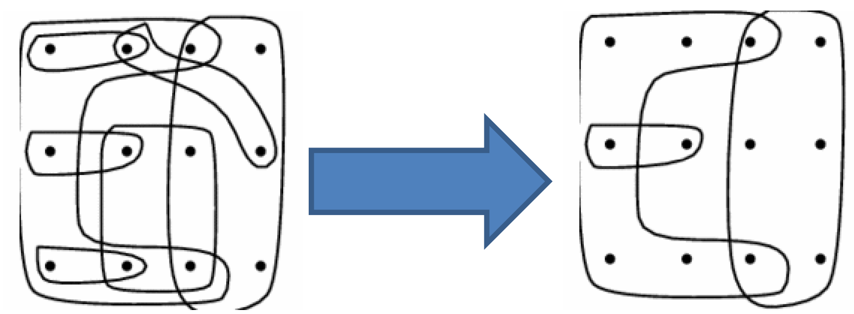
$$\begin{aligned} &\text{maximize} && c^T x \\ &\text{subject to} && Ax \le b \\ & && x \ge 0 \\ &\text{and} && x \text{ integer} \end{aligned}$$

An example of ILP:

$$\max y$$
$$-x + y \leq 1$$
$$3x + 2y \leq 12$$
$$2x + 3y \leq 12$$
$$x, y \geq 0, integer$$

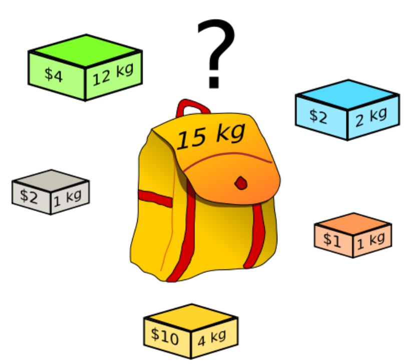


The set cover problems is a NP-hard problem, this problem is state as: given a set of elements "{1, 2, … , n}" (called the universe) and a set S of m sets with their union equals the whole universe {1, 2, …, n}, then the set cover problem is to find the smallest subset of the union of which contains all elements in the universe.



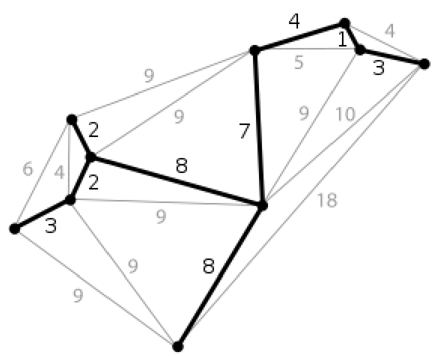
While we should point out that the exact set cover problem is NP-complete, and the two graphs above are not exact set cover[4].

Knapsack problems is also a NP-hard problem, this problem is described as: given a set of items, each with a weight w and a value v, so then Knapsack problem is asked to determine the number of each item to collect so that the total weight is less than or equal to a given limit capability of the “ bag” and meanwhile the total value is as large as possible：

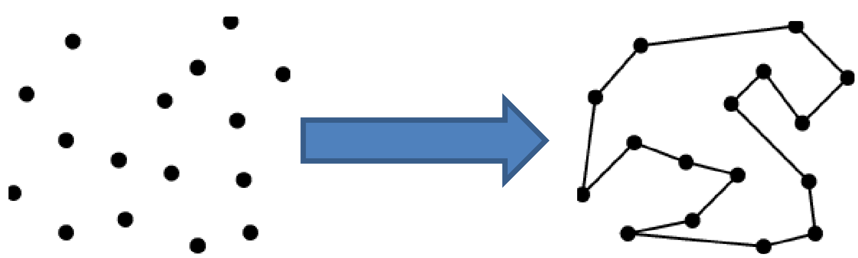


We also should point out that the 0-1 knapsack problem is NP-complete, we can prove that it is polynomial mapping reducible from exact set cover problem, which is also NP-complete as we mentioned above[5].

A minimum spanning tree (MST) problem is that, given a (usually undirected) graph, then to find a spanning tree hose weight is less than or equal to the weight of every other spanning tree. This problem can be solved by Kruskal's algorithm (which is a very famous greedy algorithm), with $O(n^2)$ steps where n is the number of vertices of the input graph. An example is as follows[6]:



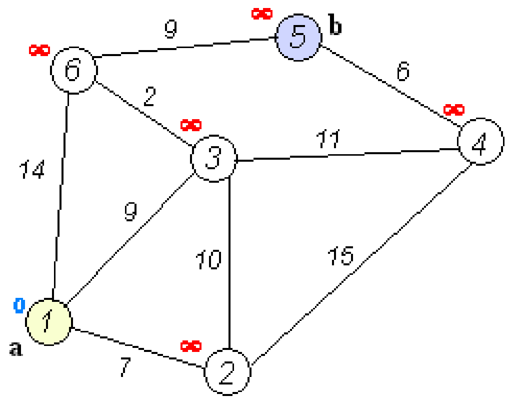
The traveling salesman problem is also NP-hard, which asks the following question: given a bunch of cities and the distances between each pair of cities, can the salesman come up with a shortest possible route that visits each city exactly once and finally returns to the origin city[4]?



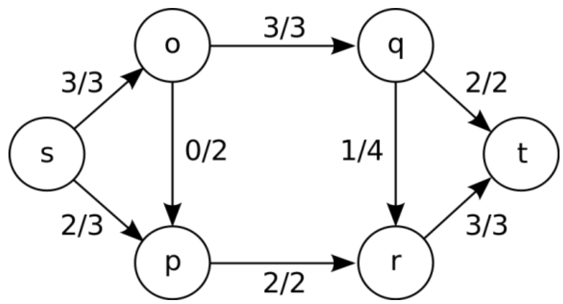
In graph theory, the shortest path problem in a (usually undirected) graph is that, to find a path between one or all pairs of vertices in the graph such that the sum of the values/distances of all the edges in that path is minimal:

- The single-source shortest path problem is to find the shortest paths from a source vertex v to all other vertices in the given graph.
- The single-destination shortest path problem is to find shortest paths from all vertices in the **directed** graph to a single destination vertex v. This can obviously be reduced to the single-source shortest path problem by reversing the arcs in the directed graph.
- The all-pairs shortest path problem is to find shortest paths between every pair of vertices v, v' in the graph.

Dijkstra's algorithm can solve the single-source shortest path problems in $O(n^2)$ steps where n is the number of vertices in the graph. The Floyd–Warshall algorithm can solve all pairs shortest paths in $O(n^3)$ steps where n is the number of vertices in the graph[7].



The maximum flow problem is to find a feasible flow through a single-source (s), single-sink (t) flow network that is maximal, and the max-flow min-cut theorem states that the maximum value of an s-t flow is equal to the minimum capacity over all s-t cuts[8]:



2. <u>Discrete Optimization in Radiation Therapy:</u>

Radiation therapy is the use of focused and targeted high-energy photons in the form of x-rays or γ-rays to treat cancer, most often soild tumors; it is part of the standard of care for many types of cancer along with surgery and chemo-therapy. There are many delivery methods for radiation therapy including:

- External Beam Radiation
  - External beam radiation therapy is the most traditional form of radiation therapy. The x-ray radiation is delivered to the patient from many axial angles. It is typically delivered in many small fractions (5 days a week for several weeks) in order to total the desired total dose
- Intensity Modulated Radiation Therapy (IMRT)
  - Similar to external beam radiation, intensity modulated radiation therapy (IMRT) delivered x-ray radiation from many axial angles, but also involved more complex collimation. The computer-controlled collimation allows for more precise shielding of the normal tissue and for high intensities to be delivered to the target volume (the tumor).

- Gamma-Knife Radiosurgery
  - An array of colbolt-60 radioactive sources is placed around the patient (typically the head for brain tumor treatment) and each source emits γ-rays by radioactive decay. Each source has the same target point, creating a sphere of radiation dose at that location. This allows for more targeted treatments, high intensities and this fewer fractionations.
- Brachytherapy
  - Brachytherapy involved implanting radioactive seeds into or near the tumor which then emit γ-rays with very short ranges as to only affect a small volume surrounding the seed. This technique is often used for prostate cancer.

The goal of all methods of radiation treatment is to deliver the prescribed dose to the target volume (includes the tumor and a small margin surrounding) in a uniform manner and minimize the dose delivered to the surrounding normal tissue. The types of surrounding tissue and their sensitivity to radiation are considered. For example, if a tumor is near the spinal cord, special attention is paid to minimize the dose delivered to the cord, more so than other types of surrounding normal tissue such as muscle, which can more readily regenerate.

The basic problem of optimizing the radiation dose can put stated in the following way:

$$\begin{aligned} Dose(i,j,k) &\leq U_{\mathcal{R}} + M * Exceed(i,j,k) \quad \forall (i,j,k) \in \mathcal{R} \\ \sum_{(i,j,k)\in\mathcal{R}} Exceed(i,j,k) &\leq \beta_{\mathcal{R}} * card(\mathcal{R}). \end{aligned}$$

Where Dose(i,j,k) is the dose to the voxel located at (i,j,k), U_R is the prescribed dose to the region R. M is a constant and Exceed is a binary variable. β_R is the fraction of voxels allowed to exceed the prescribed dose[9].

For the specific methods of radiation delivery, there are additional variables to optimize. In the case of IMRT, the beam angles, wedge orientations, beam intensity, and time of exposure in each orientation all need to be optimized, not all of which are discrete variables. Gamma knife requires that first the positions and the weight of the shots, two continuous variables be optimized with quassi-Newton methods, then the number of shots and the discrete collimator size be optimized with simulated annealing, another discrete optimization method[10]. Brachytherapy has additional variables including the number and location of seeds. The number of needles required to insert the seeds is another variable to minimize[9].

3. Discrete Optimization in Medical Imaging:

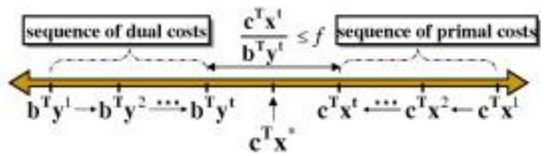
In order to compare two medical images in the same space they must be registered to each other, or to a standard template. This allows for the same point to correspond to the same anatomical feature in both images. Automated and accurate registrations are crucial for group comparisons when multiple images must be compared. Medical image registration is faced with four main obstacles: curse of dimensionality, curse of non-convexity, curse of non-linearity, and the curse of modularity[11]. The curse of dimensionality occurs when more parameters must be added to the system to increase the powerfulness of the model, but this also increases the complexity of the problem. This is often seen in registration procedures that require high dimensional warping. The curse of non-convexity occurs when there are more parameters input than constraints, or, in other words, it is an ill-posed problem. This often occurs when an image of many voxels is being related to a single bio-marker. The curse of non-linearity occurs because the parameters are related in a highly non-linear fashion. This is again seen in registration procedures with high dimensional warping. The curse of modularity occurs because solutions are found for very specific problems, rather than a broad application. Different imaging modalities, CT or MRI, usually need their own registration algorithm due to differences in the intensity values. MRI is specifically difficult because the intensity values are not consistent across scanners, where a CT image will have a standard unit of measure. This makes MRI registration especially difficult because it cannot be based on exact intensities, but on the differences, and geometry in the image.

These problems make image registration, and other imaging problems, a prime candidate for the implementation of optimization algorithms. Most image registration algorithms require an iterative approach. At each iteration the similarity between the two “registered” images is calculated, and the algorithm continues moving towards a better similarity measure, until it converges within some tolerance. Many optimization methods have been used for image registration in the past. These methods include: Powell’s (conjugate direction) method, Downhill Simplex, Levenberg-Marquardt, Newton-Raphson, stochastic search methods, gradient descent methods, and quasi-exhaustive search methods[12]. Recently, Glocker et al. published a registration algorithm using linear programming[13]. Linear programming is a branch of combinatorial optimization, within discrete optimization.

Glocker et al.’s methods modeled the registration as a Markov Random Field (MRF), where a set of labels is associated with a set of deformations. The MRF is then optimized through a method known as Fast-PD[14]. This method uses the primal-dual schema. The primal-dual schema, first starts by seeking an optimal solution x* to an integer program (the primal problem), which is NP-hard. The problem is then relaxed to get a primal (minimization) and a dual (maximization) linear program. The pair of linear programs are shown below:

$$\begin{array}{llll} \text{Primal} : \min \mathbf{c}^{\mathrm{T}}\mathbf{x} & & \text{Dual} : \max \mathbf{b}^{\mathrm{T}}\mathbf{y} & \\ & \text{s.t. } \mathbf{A}\mathbf{x} = \mathbf{b},\ \mathbf{x} \geqslant \mathbf{0} & & \text{s.t. } \mathbf{A}^{\mathrm{T}}\mathbf{y} \leqslant \mathbf{c} \end{array}.$$

From this the main goal is to minimize gap between primal and dual. This is illustrated below:



Dual and Primal solutions make local improvements to each other until their final costs (b^Ty^t, c^Tx^t) are within a pre-set range

This method was tested on the automatic registration of individual brain MRI images to a template, and it was successful. Particularly, it addressed existing imaging issues such as the curse of non-convexity by implementing the primal-dual optimization technique.

4. Matlab Example

Matching problems are common problems in combinatorial optimization. In this example, we focus on the case when the underlying graph is bipartite. The maximum weight bipartite matching problem tries to pick out elements from A such that each row and column get only a single non-zero but the sum of all the chosen elements is as large as possible.

1. Construct a rectangular matrix A

```
A = rand(5,3)

A =

   0.4886    0.5468    0.6791

   0.5785    0.5211    0.3955

   0.2373    0.2316    0.3674

   0.4588    0.4889    0.9880

   0.9631    0.6241    0.0377
```

2. Determine the matching

```
[val mi mj] = bipartite_matching(A);
```

3. Rearrange the matrix based on the outputs of the previous problem

```
A1 = A(mi,mj)

A1 =

   0.5468    0.6791    0.4886

   0.4889    0.9880    0.4588

   0.6241    0.0377    0.9631
```

4. Sum the diagonal of the new matrix

```
ans = sum(diag(A1))

val

ans = 2.4979

val  = 2.4979
```

The answer should verify that the maximum weighted matching is obtained when using the bipartite_matching command

References:

[1] (2013). *Discrete Optimization*. Available: http://en.wikipedia.org/wiki/Discrete_optimization

[2] (2013). *Combinatorial Optimization* Available: http://en.wikipedia.org/wiki/Combinatorial_optimization

[3] (2013). *Integer Programming*. Available: http://en.wikipedia.org/wiki/Integer_programming

[4] S. Skiena. (2008). *Discrete Optimizatin Methods*. Available: http://www.cs.sunysb.edu/~algorith/implement/syslo/implement.shtml

[5] (2013). *Knapsack problem*. Available: http://en.wikipedia.org/wiki/Knapsack_problem

[6] (2013). *Minimum spanning tree*. Available: http://en.wikipedia.org/wiki/Minimum_spanning_tree

[7] (2013). *Shortest path problem*. Available: http://en.wikipedia.org/wiki/Shortest_path_problem

[8] (2013). *Maximum flow problem*. Available: http://en.wikipedia.org/wiki/Maximum_flow_problem

[9] M. C. Ferris*, et al.*, "Radiation Treatement Planning: Mixed Interger Programming Forulations and Approaches," in *Handbook on Modeling for Discrete OPtimization*, Appa*, et al.*, Eds., ed New York: Springer-Veriag, 2006.

[10] J. Lim, "Optimization in Radiation Treatment Planning," Doctor of Philosophy, Industrial Engineering, University of Wisconsin, Madison, WI, 2002.

[11] N. Paragios, "Graphical Models and Discrete Optimization in Biomedical Imaging: Theory and Applications," in *Seminars*, ed: Isaac Newton Institue for Mathematical Sciences, 2011.

[12] J. B. A. Maintz and M. A. Viergever, "An overview of medical image registration methods," *UU-CS,* 1998 1998.

[13] B. Glocker*, et al.*, "Dense image registration through MRFs and efficient linear programming," *Medical Image Analysis,* vol. 12, pp. 731-741, 2008/12// 2008.

[14] N. Komodakis and G. Tziritas, "Approximate labeling via graph cuts based on linear programming," *Pattern Analysis and Machine Intelligence, IEEE Transactions on,* vol. 29, pp. 1436-1453, 2007 2007.