C0: 5DA001: Non-linear Optimization

Introduction

Niclas Börlin Klas Markström

[1]Department of Computing Science, Umeå University, niclas.borlin@cs.umu.se, office C229
[2]Department of Mathematics and Mathematical Statistics, Umeå University, klas.markstrom@math.umu.se, office E328

5DA001 Non-linear Optimization

Motivation

- What is optimization?
  - To find the value of some parameter(s) that gives the optimal function value.
- What is non-linear?
  - Everything that is not linear.
  - A linear problem

    $$\min_x f(x) = ax - b,$$

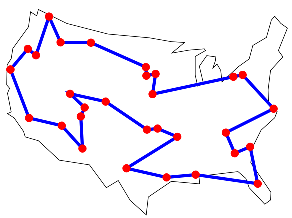
    does not have a finite minimum for $x \in \Re$.
  - Example: The non-linear problem

    $$\min_{x \in \Re} f(x) = x^2,$$

    has a minimum at $x = 0$.

Example 1: Travelling Salesman

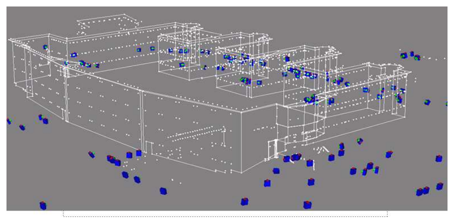
- *Given n cities, find the shortest route that visits each city exactly once.*



- Application: A robot moves its arm over a circuit board in order to place components at given locations. We want to mimize the total distance, and thereby time, the arm has to move in order to place all components on the board.

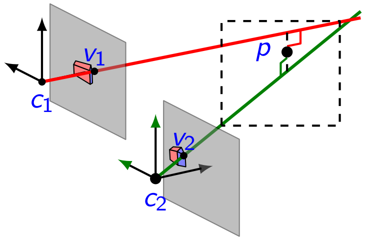
Example 2: Bundle Adjustment (photogrammetry/computer vision)

- *Given a number of measurements in multiple images, calculate the positions of the object points and cameras in 3D such that the points projected into the cameras are as close as possible to the measured points.*



Example 3: Triangulation

- *Given measurements in two images of the projection of a 3D point, determine the 3D point that is as close as possible to both rays in space.*



Example 4: The Dåvamyran co-generation plant

- The Dåvamyran co-generation plant has four boilers.
- Each boiler has
  - a fixed startup cost S_i,
  - a linear energy cost E_i, and
  - a maximum energy output h_i.
- *Given an expected energy demand d, find the cheapest way to meet the demand, i.e. which boilers should be operating and at what level?*

Important questions about the *problem*

- What do I want to miminize?
  - An objective function

$$f(x).$$

- What properties does my objective function have?
  - Continuous or discrete?
  - Differentiable?
- What parameters do I have?
- What restrictions do I have on my parameters?
  - What is the domain for the parameters? Are they discrete or continuous?
  - Are there any additional constraints on the parameters?
- Does the problem belong to any special class of problems?

Example 4: The Dåvamyran co-generation plant

- Parameters and domains:
  - b_i indicating if a boiler is operated or not (boolean).
  - e_i the energy output of each boiler (continuous).
- Objective function: Cost

$$f(x) = \sum_{b_i=1} S_i + E_i = \sum_{b_i=1} S_i + k_i e_i,$$

  where k_i is the unit energy cost for boiler i.
- Constraints:

$$\sum e_i \geq d,$$
$$0 \leq e_i \leq h_i.$$

- Class: Mixed-integer linear problem.

Example 3: Triangulation

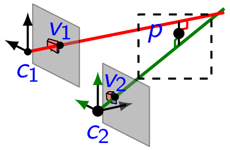
- Parameters and domains:
  - p position of object point (continuous $\Re^3$)
  - α_i position along each ray of point closest to p (continuous $\Re$)
- Objective function: Sum of squared point-ray distances

$$f(x) = \sum_i \|p - l_i(\alpha_i)\|^2,$$

  where

$$l_i(\alpha_i) = c_i + \alpha_i(v_i - c_i) = c_i + \alpha_i t_i.$$

- Constraints: None (or p should be in front of both cameras).
- Class: Continuous, unconstrained linear least squares.

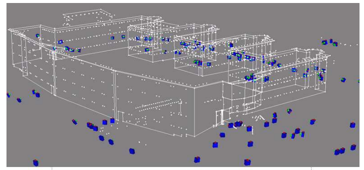


Example 2: Bundle adjustment

- Parameters and domains:
  - Object points p_i (continuous $\Re^3$)
  - Camera positions c_j (continuous $\Re^3$)
  - Camera rotations $a_j = [\omega_j, \phi_j, \kappa_j]$ (continuous $\Re$)
- Objective function: Sum of squared distances between projected object points p_i and measured image points q_j

$$f(x) = \sum_{(i,j) \in p_i \text{ visible in image } j} \|P(p_i, c_j, a_j) - q_j\|^2,$$

  where $P(p_i, c_j, a_j)$ is the Euclidean projection of point p_i into the camera at c_j and with orientation a_j.
- Constraints: None (or p_i should be in front of its cameras).
- Class: Continuous, unconstrained non-linear least squares.

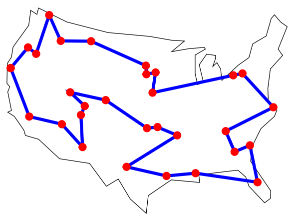


Example 1: Travelling Salesman

- Parameters and domains: x_i indicating the city number (discrete, $[1, n]$).
- Objective function: Sum of travelled distances

$$f(x) = \sum_i dist(p_{x_i}, p_{x_{i-1}}),$$

  where p_i is the position of city i and $dist(\cdot)$ is the distance between the city pair.
- Constraints: $x = [x_1, \ldots, x_n]$ must be a permutation of $1, \ldots, n$.
- Class: Discrete permutation problem.



Important questions about the *solution*

- Does a solution exist?
- Is the solution unique?
  - Are there other solutions "nearby"? (Local minimizer.)
  - Are there other solutions elsewhere? (Global minimizer.)
- What other properties does the solution have?
  - How do I determine if x is a solution?
  - What kind of solution do I need?
    - A global minimum.
    - A local mimimum.
    - Upper/lower bounds on the minimum.
    - A "good enough" value.
    - The best value given the available time.

Important questions about the *algorithm*

- Can the problem be solved by a direct method?
  - Examples:
    - SVD on rigid-body problem.
    - Triangulation: Rewrite problem as

$$\min_x \| \underbrace{\begin{pmatrix} I_3 & -t_1 & 0 \\ I_3 & 0 & -t_2 \end{pmatrix}}_{A} \underbrace{\begin{pmatrix} p \\ \alpha_1 \\ \alpha_2 \end{pmatrix}}_{x} - \underbrace{\begin{pmatrix} c_1 \\ c_2 \end{pmatrix}}_{b} \|^2.$$

    - The solution is given by the normal equations

$$A^T A x = A^T b.$$

- For iterative methods:
  - How do I start?
  - How do I continue?
  - How do I know when to stop?

Important questions about the *algorithm*

- How do I start?
  - Random start values.
  - Fixed start values, e.g. 0 or $[1, 2, \ldots, n]$.
  - Estimated start values (solve simplified problem or heuristics).
- How do I continue?
  - Random new value (stochastic algorithms).
  - Next value (brute force algorithms on discrete problems).
  - Improve on current value (descent algorithms).
- How do I know when to stop?
  - High enough probability of having found minimum (stochastic).
  - Have tried all values (brute force).
  - Update is small enough (iterative algorithms).
  - Objective function values increase in all direction from here (descent methods).
  - No time left (real-time algorithms).

Solving a Real-world Optimization Problem

- Transform a described problem into a mathematical problem.
  - Formulate an objective function.
  - Determine the metric (max error, avg error, min energy, min distance, min cost, etc.)
  - Determine what variables to estimate.
  - Formulate constraints.
- Classify the optimization problem.
- Determine which algorithm to use.
- Determine a starting approximation.
- Solve the optimization problem.
- Analyze the result.

Goals

- My personal goals:
  - I want you to feel comfortable with the word *non-linear*.
  - I want you to appreciate how you can rewrite a non-linear text problem into a mathematical formulation.
  - I want you to be able to avoid common formulation mistakes.
  - I want you to know the most common algorithms for solving these problems.
  - I want you to feel comfortable using constraints in your optimization.
  - I want you to be able to write *modular* optimization code.
  - I want you to feel that implementing Jacobians are *fun*![1]

[1] Oh, well, that's perhaps a bit over the top. I still want it, though.

Discussion time

- Discuss the four example problems. Do you understand
  - the formulation,
  - the objective function,
  - the parameters to be estimated,
  - etc.?
- For each of the example problems, answer the following questions:
  - Does a solution always exist?
  - If a solution exists, when is it unique?

Shortest path

- Find the shortest distance between two nodes a and b in a graph.
- A data package is to be sent from a webserver to a surfer on the internet. Find the fastest route for the data package through the web.

Max cut

- Given a graph G find a partition of its vertex set into two sets A and B such that as many edges as possible has one endpoint in A and one endpoint in B
- An antiferromagnet is a material where neighbouring atoms minimize their interaction energy if they have spins in opposite direction. Find a way of setting the spins to values A and B so that the total energy of the crystal is minimized. This is called a ground state of the the crystal.

Max Flow/Min Cut

- Given a graph and two nodes in the graph, and capacities for the edges, find the largest possible flow from one vertex to the other.
- An oil rig is connected to one, or several, refinieries by system of piplines. How much oil can be transported from the rig through the current pipeline system?

Structure Optimization

- Given a shape of a structure, minimize its weight while maintaining its strength.

GPS

- Given a number of measurements of the time delay of signals from several satellites, determine the position of the receiver.

Protein Folding

- A string of protein is constructed in a stretched out state. How can the protein be folded up in order to minimize its energy?

Optimal Log Cutting

- Given a tree log, a what lengths should it be cut to produce maximum revenue?