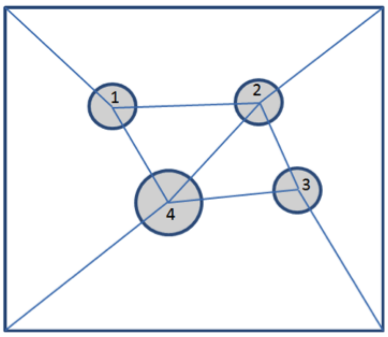
Faculty of Arts and Sciences
Department of Computer Science
CMPS 351 – Numerical Optimization
Assignment 11 – Due Apr 8, 2019

1. Planar Disk Contact. We generalize the 1D contact problem we saw earlier by considering a set of n rigid disks in the plane with radii r_i. The disks are constrained to lie in the unit square $[0\ 1] \times [0\ 1]$, and their centers are connected to each other by elastic springs with spring constants k_j. In addition, every disk's center is connected to one or more corners of the square by springs with given spring constants. The disks cannot interpenetrate each other. We seek to find the equilibrium configuration of the disks. The contact between the disks in the equilibrium configuration is not known a priori and has to be found as part of the solution.

The solution is obtained by minimizing the potential energy of the system defined as the sum of the energies of the springs. We assume that these energies are quadratic in the lengths Δ_j of the springs ($\frac{1}{2}k_j\Delta_j^2$).

- Determine the appropriate variables and constraints and formulate the problem as an appropriate optimization problem with a quadratic objective and quadratic constraints. Is the problem convex?
- Write the Lagrangian of the problem. What is the significance of the multipliers?
- Solve the problem illustrated below with the following data: $n = 4, r_1 = r_2 = r_3 = 0.1$, $r_4 = 0.15$, springs connected to disk 2 have $k = 10$, all others have $k = 1$



2. Trust Region Constraint. In many nonconvex problems, a trust region constraint is added to the (sub)problems to promote convergence to a local minimum. The trust region constraint ensures that the step size is bounded at every iteration. In this exercise, we consider an L_2 trust region constraint for an unconstrained minimization problem. The following problem is solved at the k-th iteration to find the corresponding step p.

$$\begin{aligned} \text{minimize} \quad & p^tHp + 2g^tp \\ \text{subject to} \quad & p^tp \leq \Delta^2 \end{aligned}$$

This problem has a global solution which may be computed exactly via the dual.

- Write the Lagrangian and derive the dual problem

$$\begin{aligned} \text{maximize} \quad & -g^t(H + \lambda I)^{-1}g - \lambda\Delta^2 \\ \text{subject to} \quad & H + \lambda I \succeq 0 \end{aligned}$$

- Express this dual in the computationally oriented form:

$$\begin{aligned} \text{maximize} \quad & -\textstyle\sum_{i=1}^{n}(q_i^tg)^2/(\lambda_i + \lambda) - \lambda\Delta^2 \\ \text{subject to} \quad & \lambda \geq -\lambda_{\min}(H) \end{aligned}$$

  where λ_i, q_i are the eigenvalues and eigenvectors of H.

- Describe in pseudo-code how to solve this problem and find the optimal minimizer p. (Refer to section 4.3 of reference 1).

3. Gauss-Newton Method for a PDE-Constrained problem. PDE-constrained optimization problems are optimization problems with partial differential equations as constraints. One common strategy for solving such problems is to discretize the PDEs and represent them as sets of algebraic equations, and then apply appropriate optimization algorithms to find the solution. In this exercise, we consider such a problem: an inverse parameter identification problem where the corresponding forward problem is Poisson's equation. (No implementation is needed in this exercise).

Consider the 1D Poisson's equation

$$\frac{d}{dx}\left(c(x)\frac{du}{dx}\right) = w(x)$$

on the spatial region $0 \leq x \leq 1$ with boundary conditions $u(0) = 0$ and $u(1) = 0$. $c(x)$ is a spatially-varying parameter. In the *forward* problem, we are given $c(x)$ and the goal is to find $u(x)$. This may be readily done by discretizing the spatial domain into n intervals/cells/elements. Let the vector $c_{n\times 1}$ represent the element parameters (c_i is a constant parameter for every interval). The goal is to find the values of u (known as the state variables) at grid points. If the domain is discretized using a uniform grid of spacing h, the differential equation may be replaced by the following finite difference algebraic equation at every interior grid point:

$$\frac{1}{h^2}\left[c_i u_{i+1} - (c_i + c_{i-1})u_i + c_{i-1}u_{i-1}\right] = w_i$$

Together with the boundary conditions, the solution can be determined uniquely by solving a set of linear equations $A(c)u = w$ for a given set of parameters c and right hand side w.

In the *inverse* problem, we are given w and an observed subset d of the state variables u (measured data) and the goal is to recover the coefficients c. The problem may be written as the following (nonconvex) optimization problem:

$$\begin{array}{ll} \text{minimize} & \frac{1}{2}||Qu - d||^2 + \frac{\alpha}{2}||c - c^0||^2 \\ \text{subject to} & A(c) = w \end{array}$$

where Q is a rectangular "picking" matrix that extracts from u the values that are to be compared to measurements, d is a given set of measurements, c^0 is a parameter distribution near which the solution is known to lie, and α is a regularization parameter. This formulation of the problem is known as a simultaneous formulation, because we seek to find both the parameters c and the corresponding state variables u simultaneously (both are optimizations variables).

- Explain briefly, in your own words, the significance of each of the terms in the optimization problem above.
- Write the Lagrangian $L(u, c, \nu)$ of the problem and the KKT optimality conditions.
- Derive the following Gauss-Newton iteration for solving the KKT equations. (Hint: Remember that Gauss-Newton is obtained by discarding from the Hessian matrix the terms involving second derivatives of the individual elements in the residual vectors)

$$\begin{bmatrix} Q^tQ & 0 & A^t \\ 0 & \alpha I & G^t \\ A & G & 0 \end{bmatrix} \begin{bmatrix} \Delta u \\ \Delta c \\ \Delta \nu \end{bmatrix} = - \begin{bmatrix} L_u \\ L_c \\ L_\nu \end{bmatrix}$$

  What is the structure of the matrix G?