A Quasi-Newton Method for Solving a Spheres Intersection Problem

Oliver Kolossoski

Programa de Pós-Graduação em Matemática Aplicada,
PPGMA - Centro Politécnico, 81530-900 Curitiba, Pr, Brazil
email: OliverKolossoski@gmail.com

Luiz Carlos Matioli

Departamento de Matemática,
Av. Cel Francisco H dos Santos 81530-900 Curitiba, Pr, Brazil.
email: matioli@ufpr.br

Abstract

The geometrical problem of finding the intersection of m spheres in $\mathbb{R}^n$ has a lot of applications such as determining the structure of a protein when a set of distances between the atoms are known. This problem is also known as molecular geometrical problem. In this work we will provide a Newton based method for finding the solution of such problem. We begin by transforming the sphere intersection problem on an optimization one, then we show that the original Newton method can not be the best approach, because the Hessian may be not positive definite. We then find a Quasi-Newton scheme which aproximates the Hessian by a matrix B_k which we show that is always positive definite, along with other theoretical results, therefore providing a method for solving the sphere intersection problem.

Keywords Spheres Intersection, Quasi-Newton Methods, Positive Definiteness.

In the Molecular Geometrical Problem, or Spheres Intersection Problem, we aim to find the intersection of m spheres on $\mathbb{R}^n$, for applications, often n is considered as 3. In our work we shall be considering m spheres S_i with center c_i and radius d_i on $\mathbb{R}^n$, where for all i,

$$S_i = \{x \in \mathbb{R}^n; \quad ||x - c_i|| = d_i\}.$$

Given a vector in $\mathbb{R}^n - \{c_i\}$, the projection of x onto S_i is given by

$$P(x, S_i) = c_i + d_i \frac{(x - c_i)}{||x - c_i||}.$$

With this definition, our problem can be written in a minimization problem

$$\min_{x \in \mathbb{R}^n} f(x) = \sum_{j=1}^{m} ||x - P(x, S_j)||^2. \tag{1}$$

Which can be shown to be equivalent as minimizing the function

$$f(x) = \sum_{j=1}^{m} |||x - c_j|| - d_j|^2 \tag{2}$$

Therefore minimizing such function is the same as minimizing the differences of the distances from x_j to the centers c_j and the radius of the spheres d_j. In the case that the intersection exists, the expectation would to reach the value 0 on each term of the sum, giving us that $||x - c_j|| = d_j$, this is, x belongs to every sphere S_j.

Looking to the formula (2), a computation gives us that

$$\nabla f(x) = 2 \sum_{j=1}^{m} (x - c_j)(1 - \frac{d_j}{||x - c_j||}). \tag{3}$$

and

$$\nabla^2 f(x) = 2 \sum_{j=1}^{m} \frac{d_j (x - c_j)(x - c_j)^t}{||x - c_j||^3} + (1 - \frac{d_j}{||x - c_j||})I. \tag{4}$$

A typical Newton step at a point x^k to minimize $f(x)$ would give us the following:

ALGORITHM 1 *Given* x^0, $\epsilon > 0$, $k = 0$
While $||\nabla f(x^k)|| \geq \epsilon$
Solve $\nabla^2 f(x^k)p_k = -\nabla f(x^k)$ *for* p_k*;*
$x^{k+1} = x^k + p^k$*;*
$k = k + 1$*;*
End

However, in this case the direction p^k is not guaranteed to be descent, for the hessian given in (4) may not be positive definite, in the case that $||x - c_j|| < d_j$ for some indexes j's. We therefore purpose the following modification of the Hessian: At the step k let be $J = \{j \in \{1, \dots n\}; \quad [1 - \frac{d_j}{||x-c_j||}] > 0\}$. then instead of setting p_k by the natural means, solve $B_k p_k = -\nabla f(x^k)$, where B_k is the following alteration of (4):

$$B_k = 2\sum_{j=1}^{m} \frac{d_j(x - c_j)(x - c_j)^t}{||x - c_j||^3} + 2\sum_{j \in J}(1 - \frac{d_j}{||x - c_j||})I. \tag{5}$$

Consider then the quasi-Newton scheme given by algorithm 1, but with B_k instead of $\nabla^2 f(x^k)$ as a matrix of the system to solve. Then this method gives us a descent direction p_k at each step, because the matrix B_k is positive definite, under certain circunstances.
We derive also some relations between the norm of the gradient and the difference $f(x^{k+1}) - f(x^k)$, namely if p_k is the direction given by the quasi-newton algorithm, we get

$$f(x^{k+1}) - f(x^k) = \nabla^t f(x^k)d_k + O(||p_k||^2).$$

This ensures that at every step we will get a decayment proportional do the norm of the gradient, meaning that if the modified algorithm 1 has an accumulation point, then it has to be stationary.