Geometry Optimization

C. David Sherrill

School of Chemistry and Biochemistry
Georgia Institute of Technology
Optimization

Many problems in computational chemistry (and scientific computing in general!) are optimization problems: i.e., *finding the “stationary points” where a multidimensional function has vanishing gradients.*

- The energy as a function of nuclear coordinates. Minima, transition states may be of interest.

- Error functions depending on parameters (e.g., fitting molecular mechanics methods).

- Variational electronic structure methods like Hartree-Fock, MCSCF, CI. Minima are ground states, saddle points are excited states. (Usually done as matrix diagonalization).
Optimization Methods

• Dumb method: step one variable until the function is a minimum in this dimension, then go on to other variables, repeat until no change. This won’t normally work for complicated functions/many variables.

• Most methods assume an analytic (i.e., directly calculated, not numerical) gradient is available.

• Some methods assume the Hessian (second derivative) or an approximation of it is also available.
Steepest Descent

- Step in gradient direction, do a “line minimization” until the minimum in that direction is reached; repeat.
- Simple approach guaranteed to find minimum but may be slow; suffers from zig-zags, each step partially spoils last.
Conjugate Gradient Methods

Improvement on Steepest Descent; attempts not to spoil previous progress

Do line minimizations in a direction which is a combination of the current gradient and the previous one

\[ d_i = -g_i + \beta_i d_{i-1} \]

The different conjugate-gradient methods provide different ways to choose \( \beta \); they involve dot products of current and previous gradients, e.g., Polak-Ribiere:

\[ \beta_{PR} = \frac{g_i^T (g_i - g_{i-1})}{g_{i-1}^T g_{i-1}} \]

Expand the energy function to second order around the current geometry $x_0$:

$$E(x) \approx E(x_0) + g^\dagger (x - x_0) + \frac{1}{2} (x - x_0)^\dagger H (x - x_0)$$

If we find the stationary point ($dE/dx$) of this approximate energy expression, we obtain

$$(x - x_0) = -H^{-1} g$$

This makes intuitive sense...think about a 1D example like Harmonic Oscillator. Need larger step if (i) gradient is larger, or (ii) force constant is smaller. Exact step if quadratic PES and exact gradient, Hessian.
Newton-Raphson with Diagonal Hessian

Suppose the Hessian matrix $\mathbf{H}$ (second derivative of electronic energy with respect to nuclear coordinates) is diagonalized to give normal modes and eigenvalues

$$\mathbf{H}u_i = \epsilon_i u_i. \quad (1)$$

The Newton-Raphson step in terms of the normal modes becomes

$$(\mathbf{x} - \mathbf{x}_0) = \sum_i \frac{-F_i}{\epsilon_i} u_i, \quad (2)$$

where $F_i = g^\dagger u_i$, i.e., the component of the gradient in each normal mode direction.
Minimize/Maximize

The Newton-Raphson step

$$(x - x_0) = \sum_i \frac{-F_i}{\epsilon_i} u_i,$$  \hspace{1cm} (3)

minimizes along directions with a positive Hessian eigenvalue $\epsilon_i$ and maximizes along directions with a negative eigenvalue. Finds minimum if all $\epsilon_i$ are positive ($H$ positive definite), otherwise a saddle point if some $\epsilon_i$ are negative. Big problems if the Hessian has the wrong “local structure.” Can sometimes solve by level shifting to make positive definite:

$$(x - x_0) = \sum_i \frac{-F_i}{\epsilon_i - \lambda} u_i.$$  \hspace{1cm} (4)
Obtaining the Hessian

- Can compute analytically if the program allows it, but can be inefficient (can take 10 times as long as energy alone); cost scales at least as $\mathcal{O}(N^3)$
- Can approximate using tables or lower level computations
- Can improve approximate Hessian using gradient information gained during optimization: Davidon-Fletcher-Powell (DFP), Broyden-Fletcher-Goldfarb-Shanno (BFGS), Powell.
- In Q-Chem, read a previously computed Hessian for geometry optimizations using `GEOM_OPT_HESSIAN READ`
Choice of Coordinates

- Z-matrix coordinates: old-style, inefficient
- Cartesian coordinates: even worse, un-natural, not like normal modes, more of them ($3N$ vs $3N - 6$)
- Internal coordinates: simple and symmetry-adapted
- Natural internal coordinates/redundant internal coordinates (Peter Pulay): similar to symmetry-adapted internals, commonly used now
- Q-Chem tries “delocalized internal coordinates” (similar to above) and switches to Cartesian if it gets confused
Multiple minima

- For larger molecules, more minima become possible

- Thermodynamic properties arise from Boltzmann distribution of all energetically accessible minima

- Can be hard to locate all relevant minima; search depends on initial guess geometries (and Hessians)

- Methods to help find additional minima: genetic algorithms, molecular dynamics, Monte Carlo methods, simulated annealing
Constrained Optimization

- Sometimes it’s useful to fix a geometrical parameter and perform the optimization subject to this constraint.

- Can rewrite function to be minimized as a Lagrange function which is the original function minus a Lagrange multiplier times a constraint function $g = 0$

\[
L(x_1, x_2, \cdots, x_N, \lambda) = f(x_1, x_2, \cdots, x_N) - \lambda g(x_1, x_2, \cdots, x_N)
\]

- Q-Chem has constrained optimization; not always straightforward.
Transition State Optimization

- Use an eigenvector following method: always go uphill in the direction of the lowest Hessian eigenvector
- Usually need a good guess for geometry \textit{and} Hessian
- Also need a theoretical method that doesn’t break down around the transition state region
- Many fancy algorithms exist but most never make it into standard program packages
Transition State Searching in Q-Chem

- Supply a good guess geometry for the TS
- Tell Q-Chem it’s a TS search by the keyword JOBTYPE TS
- You will probably need a better approximate Hessian. Tell Q-Chem to compute the Hessian first for your guess geometry (at the same or a lower level of theory) by doing a frequency calculation (JOBTYPE FREQ) and then do a optimization in the same input file; separate the two jobs with a separator line @@@. In the second input set JOBTYPE TS and GEOM_OPT_HESSIAN READ
Example of Q-Chem Linked Input

$molecule
[molecule specified here]
$end

$rem
JOBTYPE    FREQ
EXCHANGE   HF
BASIS      STO-3G
$end

@@@
$molecule
READ
$end

$rem
JOBTYPE TS
EXCHANGE HF
BASIS 6-31G*
GEOM_OPT_HESSIAN READ
$end