Muti-Dimensional Optimization

Problem:

Given:
$f(x_1, x_2, ..., x_n)$

Find:
A point in n-dimensional space:
$\{X_1, X_2, ..., X_n\}$

such that
f takes on a minimum or maximum value

Muti-Dimensional Optimization

Local Optimization
Given a starting point,
find the nearest minimum

Global Optimization
Find the absolute minimum in the whole space
(very difficult)

Classical AI problem
Traveling Saleman Problem
(at least in one formulation)

Minimize:
f($state$) where f is an evaluation of the state
and the state can be represented as a vector of parameters

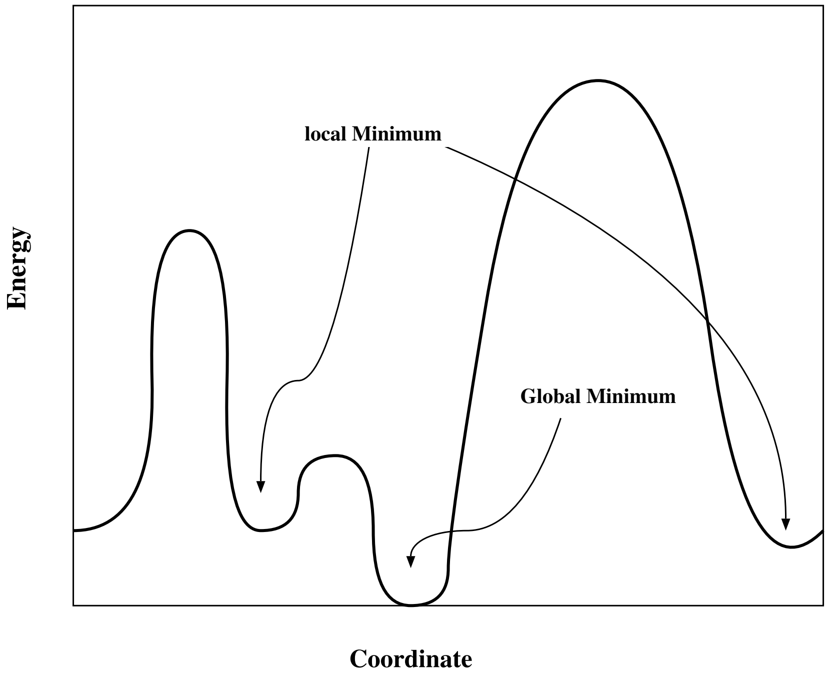
Why Numeric Method?

State Space Search

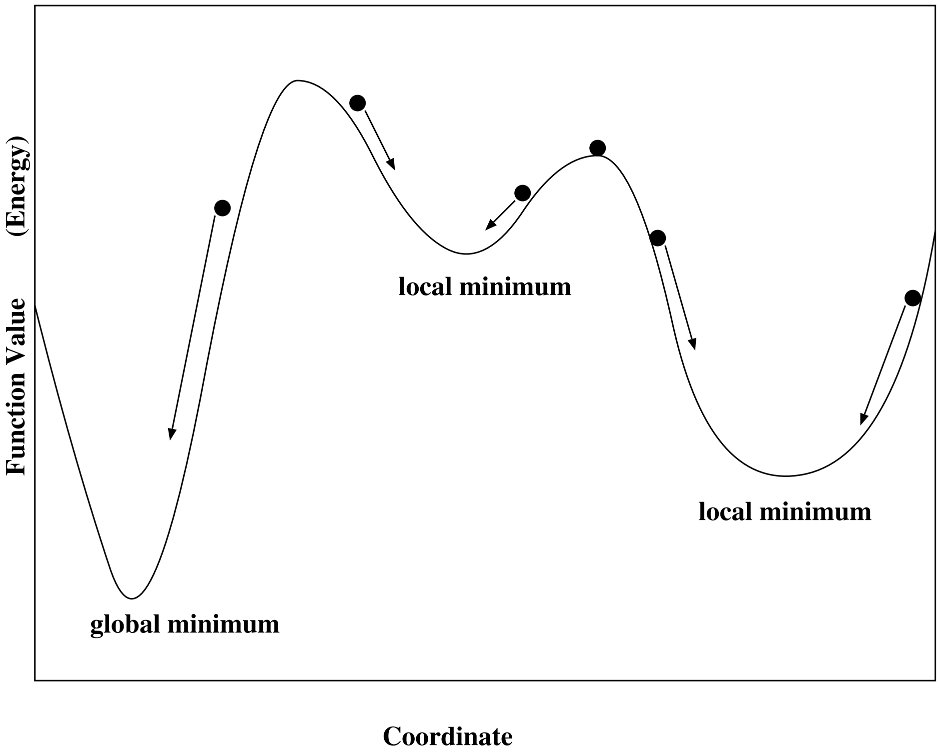
Minimum the cost function (heuristic)

The states
correspond to
the points in the data space

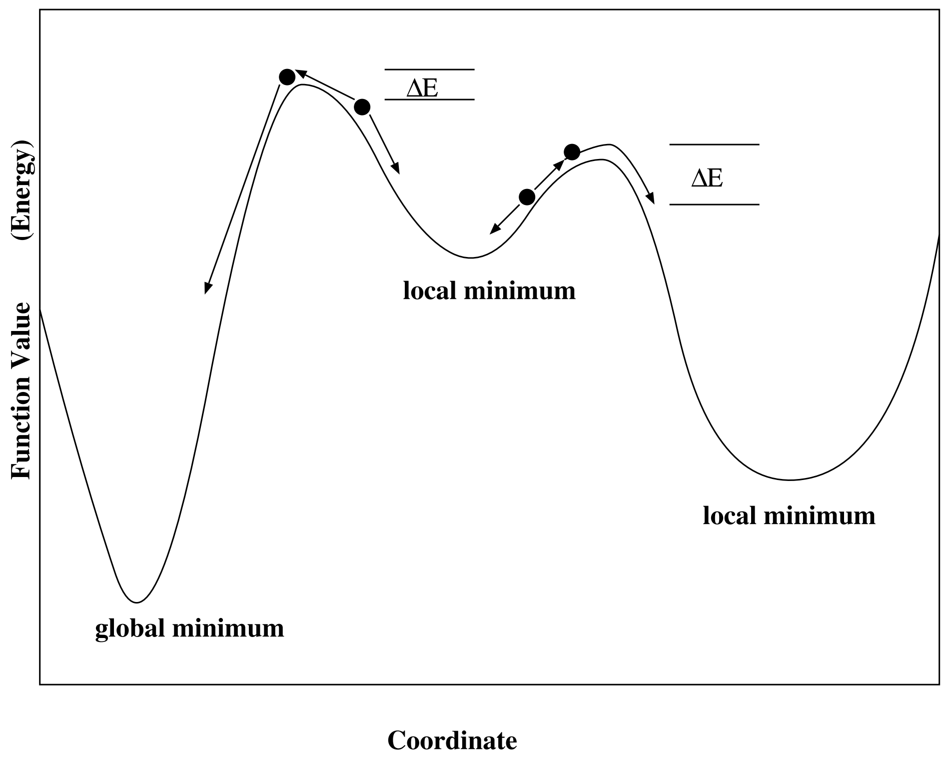
Global and Local Minimum



Gradient Methods: Local Minimum



Monte Carlo Methods: Global Minimum



Comparison

Steepest Decent
Simplex
Gradient Methods

From a given position
Find the direction to the nearest minimum
and move a toward it

Monte Carlo Methods
Simulated Annealing

Choose a point around the current point
Calculate the function value
Decide whether to accept the new point or not

Acceptance based on how much of a difference there is
and
a random factor

Direction Methods

Goal:
Find Nearest Local Minimum

Given a start point:
and
a direction: n

Find the minimum along the line:
$P_{min} = P + \lambda n$
A one dimensional optimization

The P_{min} is the next point

Direction Methods: Naive Approach

Pick as directions
the set of unit vectors
n_i

Find minimum along each in succession

Advantage:
Computationally Simple
(i.e. No gradient needed)

Disadvantage:
Can result in many very small steps
In the case of very narrow potentials

Steepest Descent: A Better Direction

Take as the direction
The local downhill gradient
$-\delta f(P)$

Each step is perpendicular to the last:

Thus, once again, in a narrow valley
a large number of small steps.
(Even in a quadratic Curve)

Conjugate Gradient

Pick a set of directions that are mutually orthogonal

This is done by setting up the following sequence
(Analogous to Gram-Schmidt bi-orthogonization)

$$g_{i+1} = g_i + \lambda A h_i$$

$$h_{i+1} = g_{i+1} + \gamma h_i$$

where

$$\lambda_i = \frac{g_i h_i}{h_i A h_i}$$

$$\gamma_i = \frac{g_{i+1} g_{i+1}}{g_i g_i}$$

Conjugate Gradient

If A is the Hessian
(the matrix of second derivatives)
For quadratic functions
Then convergence in N+1 steps

But we don't know A!!!

Fletcher-Reeves
Adding a line minimization step
Can converge without actually building A

Davidson Fletcher Powell

Called
Variable Metric
Quasi-Newton

Make Successive Approximations of the Hessian Matrix

General Idea
Can compute the Hessian from
N+1 finite difference calculations
using the gradient

Finds Local Minimum