Optimization

or how the Computer finds the ‘best-fit’ Line(s)

Objectives

- Understand the process of optimization
- Understand some of the optimization methods (algorithms)
- Understand the Advantages and Disadvantages of some these methods

Optimization

- The objective is to Reduce WSS by making adjustment in the values of the Parameters of the Model
- Advantages described in terms of
  - Robustness
  - Speed
- Disadvantages
  - Lost
  - Cost

Optimization Algorithms

- Steepest Descent
- Gauss-Newton
- Marquardt
- Nelder Mead (Simplex)

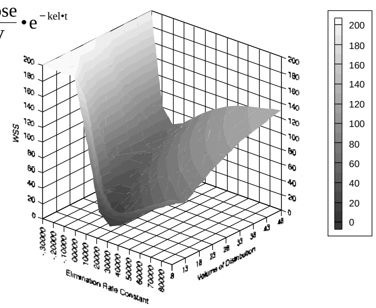
Optimization

- Move across the Sum of Squares surface to reach the Global Minimum

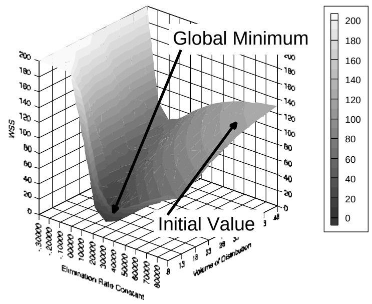
WSS = f (C_{obs}, t, Wt (the data), P, C (the model))

Sum of Squares Surface

$$Cp = \frac{Dose}{V} \bullet e^{-kel \bullet t}$$



Sum of Squares Surface



Steepest Descent Method

Pocket Creek and Hole-in-the-Wall Falls, Glacier Park, Montana from http:// pubcenter.com/silverking/9-198.shtml

Steepest Descent Method

$$P_{NEW} = P_{OLD} - h \bullet \frac{\partial WSS}{\partial P}$$

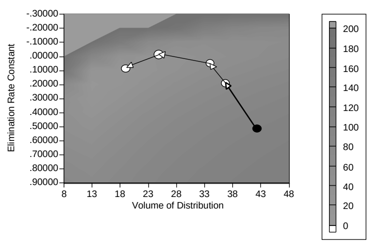
- Direction Based on Slope of WSS Surface
- Step Length based on Linear Search

Steepest Descent Method

- Always 'downhill'
- Avoids 'saddle points'
- Efficient further from the minimum

- Slower close to minimum
- Linear search may cause problems
- Might 'zigzag' down valleys

Steepest Descent Method



Gauss-Newton Method

- Single Parameter

$$P_{NEW} = P_{OLD} - \frac{\frac{dWSS}{dP}}{\frac{d^2WSS}{dP^2}}$$

- Multiple Parameters

$$P_{NEW} = P_{OLD} - \frac{\frac{dWSS}{dP}}{\frac{\partial^2 WSS}{\partial P \partial P}}$$

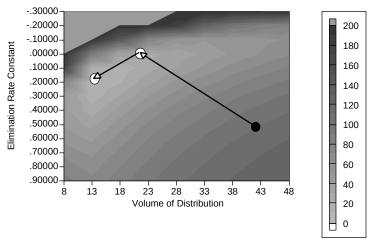
Gauss-Newton Method

- Numerical Differention
  dT Differentiation Step Size
  - for dWSS/dP type of calculations
- Iterative Process
  dC Convergence Criteria
  - Smallest change in WSS
  - Smallest change in P values

Gauss-Newton Method

- Relatively efficient (direction and step size determined)
- Works well near the minimum
- May become lost with poor initial estimates
  - Damping Gauss-Newton Method (start with Simplex)

Damping Gauss-Newton Method

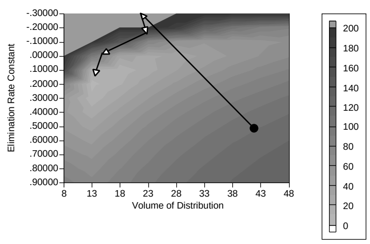


Marquardt Method

$$P_{NEW} = P_{OLD} - \frac{\frac{dWSS}{dP}}{\left(\left[\frac{\partial^2 WSS}{\partial P \partial P}\right] + \mu I\right)}$$

- μ term changes direction between Steepest Descent and Gauss-Newton during the iterative process
  - Moves from Steepest Descent automatically to the Gauss-Newton method to improve efficiency with difficult problems

Marquardt Method



Marquardt Method

Initial Estimate		Marquardt	Gauss-Newton
kel	**V**		
0.51	42	6	7
0.51	10	4	4
0.15	10	3	3
0.15	42	5	4

Nelder-Mead (Simplex) Method

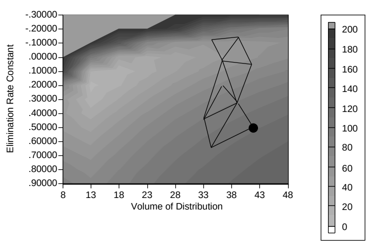
- Develop simplex (shape) with m+1 points
- Reflect worst point through centroid (center)
  - Best - reflect further
  - Good - repeat again
  - Worst - reflect closer
- Simplex moves over WSS surface and contacts around minimum

Simplex Method

- Relatively Robust
- Numerically less complicated

- Not very efficient for simple problems

Simplex Method



Grid Search Method

- Simply determine WSS at each point on the grid of parameter values
- May offer some protection against local minima
- Not very efficient especially with more parameters (with 3 parameters and 10 points per grid, 10 x 10 x 10 determinations required)

Optimization Methods

- Boomer
  - Gauss-Newton (Hartley), Damping Gauss-Newton, Marquardt, Simplex
- SAAM II
- WinNONLIN
  - Curve Stripping, Grid Search, Gauss-Newton (Hartley), Gauss-Newton (Levenberg), Nelder-Mead (Simplex)
- ADAPT II
  - Nelder-Mead (Simplex)