Introduction to Nonlinear Programming (NLP)
This lecture was adapted from Thomas W. Reiland, North Carolina State University and from Daniel P. Loucks & Eelco van Beek, UNESCO-Delft Hydraulics

Dr. Arturo S. Leon, BSU (Spring 2010)
Nonlinear Programming (NLP)

- An NLP problem has a nonlinear objective function and/or one or more nonlinear constraints.
- The mathematics involved in solving NLPs is quite different than for LPs.
Possible Optimal Solutions to NLPs
(not occurring at corner points)

Feasible Region

linear objective, nonlinear constraints

Feasible Region

nonlinear objective, nonlinear constraints

Feasible Region

nonlinear objective, linear constraints

Feasible Region

nonlinear objective, linear constraints
Steepest Descent Method

\[ f(x_1, x_2) = x_1^2 + 5x_2^2 \]

Contours are shown below
The gradient at the point \((x_1^1, x_2^1)\) is

\[
\nabla f(x_1^1, x_2^1) = (2x_1^1, 10x_2^1)^T
\]

If we choose \(x_1^1 = 3.22, x_2^1 = 1.39\) as the starting point represented by the black dot on the figure, the black line shown in the figure represents the direction for a line search.

\[
-\nabla f(x_1^1, x_2^1) = (-6.44, -13.9)^T
\]

Contours represent from red \((f = 2)\) to blue \((f = 20)\).
Now, the question is how big should the step be along the direction of the gradient? **We want to find the minimum along the line before taking the next step.**

The minimum along the line corresponds to the point where the new direction is orthogonal to the original direction. The new point is \((x_1^2, x_2^2) = (2.47, -0.23)\) shown in blue.

**Steepest Descent**
Steepest Descent

By the third iteration we can see that from the point \((x_1^2, x_2^2)\) the new vector again misses the minimum, and here it seems that we could do better because we are close.

Steepest descent is usually used as the first technique in a minimization procedure, however, a robust strategy that improves the choice of the new direction will greatly enhance the efficiency of the search for the minimum.
The GRG Algorithm (Used in solver - Excel)

- GRG can also be used on LPs but is slower than the Simplex method.
- Solver uses the Generalized Reduced Gradient (GRG) algorithm to solve NLPs.
- The following discussion gives a general (but somewhat imprecise) idea of how GRG works.
An NLP Solution Strategy

Feasible Region

A (the starting point)

objective function level curves
Local vs. Global Optimal Solutions

Feasible Region

Local optimal solution

Local and global optimal solution
Comments About NLP Algorithms

- It is not always best to move in the direction producing the fastest rate of improvement in the objective.
- NLP algorithms can terminate at local optimal solutions.
- The starting point influences the local optimal solution obtained.
Comments About Starting Points

- The null starting point should be avoided.
- When possible, it is best to use starting values of approximately the same magnitude as the expected optimal values.
A Note About “Optimal” Solutions

When solving a NLP problem, Solver normally stops when the first of three numerical tests is satisfied, causing one of the following three completion messages to appear:

1) “Solver found a solution. All constraints and optimality conditions are satisfied.”

This means Solver found a local optimal solution, but does not guarantee that the solution is the global optimal solution.
A Note About “Optimal” Solutions

- When solving a NLP problem, Solver normally stops when the first of three numerical tests is satisfied, causing one of the following three completion messages to appear:

2) “Solver has converged to the current solution. All constraints are satisfied.”

This means the objective function value changed very slowly for the last few iterations.
A Note About “Optimal” Solutions

- When solving a NLP problem, Solver normally stops when the first of three numerical tests is satisfied, causing one of the following three completion messages to appear:

  3) “Solver cannot improve the current solution. All constraints are satisfied.”

This rare message means the your model is degenerate and the Solver is cycling. Degeneracy can often be eliminated by removing redundant constraints in a model.
Location Problems

- Many decision problems involve determining optimal locations for facilities or service centers. For example,
  - Manufacturing plants
  - Warehouse
  - Fire stations
  - Ambulance centers

- These problems usually involve distance measures in the objective and/or constraints.

The straight line (Euclidean) distance between two points \((X_1, Y_1)\) and \((X_2, Y_2)\) is:

\[
\text{Distance} = \sqrt{(X_1 - X_2)^2 + (Y_1 - Y_2)^2}
\]
A Location Problem: Rappaport Communications

- Rappaport Communications provides cellular phone service in several mid-western states.
- They want to expand to provide inter-city service between four cities in northern Ohio.
- A new communications tower must be built to handle these inter-city calls.
- The tower will have a 40 mile transmission radius.
Graph of the Tower Location Problem

Cleveland
x=5, y=45

Akron
x=12, y=21

Youngstown
x=52, y=21

Canton
x=17, y=5
Defining the Decision Variables

\[ X_1 = \text{location of the new tower with respect to the X-axis} \]

\[ Y_1 = \text{location of the new tower with respect to the Y-axis} \]
Defining the Objective Function

Minimize the total distance from the new tower to the existing towers

\[
\min \sqrt{(5 - X_1)^2 + (45 - Y_1)^2} + \sqrt{(12 - X_1)^2 + (21 - Y_1)^2} + \sqrt{(17 - X_1)^2 + (5 - Y_1)^2} + \sqrt{(52 - X_1)^2 + (21 - Y_1)^2}
\]
Defining the Constraints

- Cleveland
  \[\sqrt{(5 - X_1)^2 + (45 - Y_1)^2} \leq 40\]

- Akron
  \[\sqrt{(12 - X_1)^2 + (21 - Y_1)^2} \leq 40\]

- Canton
  \[\sqrt{(17 - X_1)^2 + (5 - Y_1)^2} \leq 40\]

- Youngstown
  \[\sqrt{(52 - X_1)^2 + (21 - Y_1)^2} \leq 40\]

Problem solved in class
Analyzing the Solution

- The optimal location of the “new tower” is in virtually the same location as the existing Akron tower.
- Maybe they should just upgrade the Akron tower.
- The maximum distance is 39.8 miles to Youngstown.
- This is pressing the 40 mile transmission radius.
Comments on Location Problems

- The optimal solution to a location problem may not work:
  - The land may not be for sale.
  - The land may not be zoned properly.
  - The “land” may be a lake.

- In such cases, the optimal solution is a good starting point in the search for suitable property.

- Constraints may be added to location problems to eliminate infeasible areas from consideration.
Another example (Taken from Daniel P. Loucks & Eelco van Beek)

Consider the system shown below where a reservoir is upstream of three demand sites along a river.
Example (Cont.)

The net benefits derived from each use depend on the reliable amounts of water allocated to each use. Letting $x_{it}$ be the allocation to use $i$ in period $t$, the net benefits for each period $t$ equal

1. $6x_{1t} - x_{1t}^2$
2. $7x_{2t} - 1.5 x_{2t}^2$
3. $8x_{3t} - 0.5 x_{3t}^2$

Assume the average inflows to the reservoir in each of four seasons of the year equal 10, 2, 8, 12 units per season and that the reservoir capacity is 5 volume units.

Find the optimal operating policy for this reservoir that maximizes the total (four season) allocation benefits for the users.  

*Problem solved in class*