Control Lab
www.a-lab.ee

TALLINN UNIVERSITY OF TECHNOLOGY

ISS0031 Modeling and Identification

Lecture 4: Simplex Method. Nelder-Mead Method. Penalty Functions.

Aleksei Tepljakov, Ph.D.

September 26, 2017

Solutions to a Linear Programming Problem

Recall the two forms of the linear programming problem:

$$\begin{aligned} z &= CX \to \max(\min) \\ AX &\leqslant B \\ X &\geqslant 0 \end{aligned}$$

(standard form)

or

$$\begin{aligned} z &= CX \to \max(\min) \\ AX &= B \\ X &\geqslant 0 \end{aligned}$$

(canonical form)

Solutions to a Linear Programming Problem

Recall the two forms of the linear programming problem:

$$
\begin{aligned}
z &= CX \to \max(\min) \\
AX &\leqslant B \\
X &\geqslant 0
\end{aligned}
$$

(standard form)

or

$$
\begin{aligned}
z &= CX \to \max(\min) \\
AX &= B \\
X &\geqslant 0
\end{aligned}
$$

(canonical form)

Definition 1. *A* ***basic solution*** *is a solution obtained by fixing enough variables to be equal to zero, so that the equality constraints have a unique solution.*

Solutions to a Linear Programming Problem

Recall the two forms of the linear programming problem:

$$\begin{aligned} z &= CX \to \max(\min) \\ AX &\leqslant B \\ X &\geqslant 0 \end{aligned} \qquad \text{or} \qquad \begin{aligned} z &= CX \to \max(\min) \\ AX &= B \\ X &\geqslant 0 \end{aligned}$$

(standard form) (canonical form)

Definition 1. *A **basic solution** is a solution obtained by fixing enough variables to be equal to zero, so that the equality constraints have a unique solution.*

Definition 2. *A feasible solution to the linear programming problem which is also the basic solution is called the **basic feasible solution**. Basic feasible solutions are of two types: **degenerate**—if the value of at least one basic variable is zero,—and **non-degenerate**—if all values of basic variables are non-zero and positive.*

Simplex Method: Introductory Example

Solve the following linear programming problem.

$$
\begin{aligned}
z = 3x_1 + 4x_2 &\to \max \\
2x_1 + 4x_2 &\leqslant 120 \\
2x_1 + 2x_2 &\leqslant 80 \\
x_1 \geqslant 0, x_2 &\geqslant 0.
\end{aligned}
$$

We shall solve this problem using the Simplex method.

Simplex Method: Initial Steps

Step 1: Standard form of a maximum problem

A linear programming problem in which the objective function is to be maximized is referred to as a maximum linear programming problem. Such problems are said to be in standard form if the following conditions are satisfied:

- All the variables are nonnegative;

Simplex Method: Initial Steps

Step 1: Standard form of a maximum problem

A linear programming problem in which the objective function is to be maximized is referred to as a maximum linear programming problem. Such problems are said to be in standard form if the following conditions are satisfied:

- All the variables are nonnegative;
- All other constraints are written as a linear expression, i.e., less than or equal to a positive constant.

Simplex Method: Initial Steps

Step 1: Standard form of a maximum problem

A linear programming problem in which the objective function is to be maximized is referred to as a maximum linear programming problem. Such problems are said to be in standard form if the following conditions are satisfied:

- All the variables are nonnegative;
- All other constraints are written as a linear expression, i.e., less than or equal to a positive constant.

Conclusion: The given problem is in standard form.

Simplex Method: Initial Steps (continued)

Step 2: Slack variables and initial simplex table

In order to solve the maximum problem by means of the simplex method, we need to do the following first:

- Introduce slack variables;

Simplex Method: Initial Steps (continued)

Step 2: Slack variables and initial simplex table

In order to solve the maximum problem by means of the simplex method, we need to do the following first:

- Introduce slack variables;
- Construct the initial simplex table.

Simplex Method: Initial Steps (continued)

Step 2: Slack variables and initial simplex table

In order to solve the maximum problem by means of the simplex method, we need to do the following first:

- Introduce slack variables;
- Construct the initial simplex table.

In this problem we have the constraints as linear expressions less than or equal to some positive constants. That means there is a slack between the left and right sides of the inequalities. In order to take up the slack between the left and right sides of the problem we introduce slack variables.

Simplex Method: Initial Steps (continued)

For the given problem we introduce slack variables s_1 and s_2, such that $s_1, s_2 \geqslant 0$, and

$$
\begin{aligned}
2x_1 + 4x_2 + s_1 &= 120 \\
2x_1 + 2x_2 + s_2 &= 80.
\end{aligned}
$$

Simplex Method: Initial Steps (continued)

For the given problem we introduce slack variables s_1 and s_2, such that $s_1, s_2 \geqslant 0$, and

$$\begin{aligned} 2x_1 + 4x_2 + s_1 &= 120 \\ 2x_1 + 2x_2 + s_2 &= 80. \end{aligned}$$

The objective function z can be rewritten as $z - 3x_1 - 4x_2 = 0$. Thus, we now have

$$\begin{aligned} z - 3x_1 - 4x_2 &= 0 \\ 2x_1 + 4x_2 + s_1 &= 120 \\ 2x_1 + 2x_2 + s_2 &= 80. \end{aligned}$$

Simplex Method: Initial Steps (continued)

For the given problem we introduce slack variables s_1 and s_2, such that $s_1, s_2 \geqslant 0$, and

$$
\begin{aligned}
2x_1 + 4x_2 + s_1 &= 120 \\
2x_1 + 2x_2 + s_2 &= 80.
\end{aligned}
$$

The objective function z can be rewritten as $z - 3x_1 - 4x_2 = 0$. Thus, we now have

$$
\begin{aligned}
z - 3x_1 - 4x_2 &= 0 \\
2x_1 + 4x_2 + s_1 &= 120 \\
2x_1 + 2x_2 + s_2 &= 80.
\end{aligned}
$$

The goal is to find the particular solution (x_1, x_2, s_1, s_2, z) that maximizes z.

Simplex Method: Initial Simplex Table

We now constuct the initial simplex table.

x_1	x_2	s_1	s_2	z	b	variables
2	4	1	0	0	120	s_1
2	2	0	1	0	80	s_2
-3	-4	0	0	1	0	z

The coefficients of the objective function are arranged in the bottom row, which is called the **objective row**.

From this point on, the simplex method consists of pivoting from one table to another until the optimal solution is found.

Simplex Method: Pivoting

Pivoting: To pivot a matrix about a given element, called the *pivot element*, is to apply row operations, so that the pivot element is replaced by 1 and all other entries in the same column (called *pivot column*) become 0.

Simplex Method: Pivoting

Pivoting: To pivot a matrix about a given element, called the *pivot element*, is to apply row operations, so that the pivot element is replaced by 1 and all other entries in the same column (called *pivot column*) become 0.

Pivot element: The pivot element for the Simplex method is found using the following rules:

Simplex Method: Pivoting

Pivoting: To pivot a matrix about a given element, called the *pivot element*, is to apply row operations, so that the pivot element is replaced by 1 and all other entries in the same column (called *pivot column*) become 0.

Pivot element: The pivot element for the Simplex method is found using the following rules:

- The pivot column is selected by locating the most negative entry in the objective row. If all the entries in this column are negative, the problem is unbounded and there is no solution.

Simplex Method: Pivoting

Pivoting: To pivot a matrix about a given element, called the *pivot element*, is to apply row operations, so that the pivot element is replaced by 1 and all other entries in the same column (called *pivot column*) become 0.

Pivot element: The pivot element for the Simplex method is found using the following rules:

- The pivot column is selected by locating the most negative entry in the objective row. If all the entries in this column are negative, the problem is unbounded and there is no solution.
- Divide each entry in the last column by the corresponding entry from the same row in the pivot column. Ignore the rows in which the pivot column entry is less than or equal to 0. The row in which the smallest positive ratio is obtained is the pivot row.

The pivot element is the entry at the intersection of the pivot row and the pivot column.

Simplex Method: Locating the Pivot Element

x_1	x_2	s_1	s_2	z	b	variables	ratio
2	4	1	0	0	120	s_1	
2	2	0	1	0	80	s_2	
-3	-4	0	0	1	0	z	

- Inspect the objective row.

Simplex Method: Locating the Pivot Element

x_1	x_2	s_1	s_2	z	b	variables	ratio
2	4	1	0	0	120	s_1	
2	2	0	1	0	80	s_2	
-3	-4	0	0	1	0	z	

- Inspect the objective row.
- The most negative entry is -4 in the objective row, hence the 2nd column is the pivot column.

Simplex Method: Locating the Pivot Element

x_1	x_2	s_1	s_2	z	b	variables	ratio
2	4	1	0	0	120	s_1	$120/4 = 30$
2	2	0	1	0	80	s_2	$80/2 = 40$
-3	-4	0	0	1	0	z	

- Inspect the objective row.
- The most negative entry is -4 in the objective row, hence the 2$^{\text{nd}}$ column is the pivot column.
- Compute the ratios by dividing each entry in the 6$^{\text{th}}$ column by the corresponding entry in the 2$^{\text{nd}}$ column.

Simplex Method: Locating the Pivot Element

x_1	x_2	s_1	s_2	z	b	variables	ratio
2	**4**	1	0	0	120	s_1	$120/4 = 30$
2	2	0	1	0	80	s_2	$80/2 = 40$
-3	-4	0	0	1	0	z	

- Inspect the objective row.
- The most negative entry is -4 in the objective row, hence the 2$^{\text{nd}}$ column is the pivot column.
- Compute the ratios by dividing each entry in the 6$^{\text{th}}$ column by the corresponding entry in the 2$^{\text{nd}}$ column.
- The smallest ratio is 30, so the element $a_{12} = 4$ is the pivot element.

Simplex Method: Row Operations and New Table

We now divide R_1 (the first row) by 4 and then apply the operations $R_2 = R_2 - 2R_1$ and $R_3 = R_3 + 4R_1$. The new table becomes

x_1	x_2	s_1	s_2	z	b	variables	ratio
$\frac{1}{2}$	1	$\frac{1}{4}$	0	0	30	x_2	
1	0	$-\frac{1}{2}$	1	0	20	s_2	
-1	0	1	0	1	120	z	

Simplex Method: Row Operations and New Table

We now divide R_1 (the first row) by 4 and then apply the operations $R_2 = R_2 - 2R_1$ and $R_3 = R_3 + 4R_1$. The new table becomes

x_1	x_2	s_1	s_2	z	b	variables	ratio
$\frac{1}{2}$	1	$\frac{1}{4}$	0	0	30	x_2	
1	0	$-\frac{1}{2}$	1	0	20	s_2	
-1	0	1	0	1	120	z	

- The most negative entry in the objective row is -4, hence the 1^{st} column is the pivot column.

Simplex Method: Row Operations and New Table

We now divide R_1 (the first row) by 4 and then apply the operations $R_2 = R_2 - 2R_1$ and $R_3 = R_3 + 4R_1$. The new table becomes

x_1	x_2	s_1	s_2	z	b	variables	ratio
$\frac{1}{2}$	1	$\frac{1}{4}$	0	0	30	x_2	60
1	0	$-\frac{1}{2}$	1	0	20	s_2	20
-1	0	1	0	1	120	z	

- The most negative entry in the objective row is -4, hence the 1^{st} column is the pivot column.
- Compute the ratios by dividing each entry in the 6^{th} column by the corresponding entry in the 1^{st} column.

Simplex Method: Row Operations and New Table

We now divide R_1 (the first row) by 4 and then apply the operations $R_2 = R_2 - 2R_1$ and $R_3 = R_3 + 4R_1$. The new table becomes

x_1	x_2	s_1	s_2	z	b	variables	ratio
$\frac{1}{2}$	1	$\frac{1}{4}$	0	0	30	x_2	60
1	0	$-\frac{1}{2}$	1	0	20	s_2	20
-1	0	1	0	1	120	z	

- The most negative entry in the objective row is -4, hence the 1st column is the pivot column.
- Compute the ratios by dividing each entry in the 6th column by the corresponding entry in the 1st column.
- The smallest ratio is 20, so the element $a_{21} = 1$ is the pivot element.

Simplex Method: Row Operations and Final Table

We apply the operations $R_1 = R_1 - \frac{1}{2}R_2$ and $R_3 = R_3 + R_2$. The new table becomes

x_1	x_2	s_1	s_2	z	b	variables
0	1	$\frac{1}{2}$	$-\frac{1}{2}$	0	20	x_2
1	0	$-\frac{1}{2}$	1	0	20	x_1
0	0	$\frac{1}{2}$	1	1	140	z

Simplex Method: Row Operations and Final Table

We apply the operations $R_1 = R_1 - \frac{1}{2}R_2$ and $R_3 = R_3 + R_2$. The new table becomes

x_1	x_2	s_1	s_2	z	b	variables
0	1	$\frac{1}{2}$	$-\frac{1}{2}$	0	20	x_2
1	0	$-\frac{1}{2}$	1	0	20	x_1
0	0	$\frac{1}{2}$	1	1	140	z

There are no negative entries in the objective row—the optimal solution has been found. Therefore, $z_{max}(20, 20) = 140$.

Definition 3. *The variables corresponding to the columns of the identity matrix in the initial simplex table are called* ***basic variables*** *while the remaining variables are called* ***nonbasic*** *or* ***free variables****.*

Simplex Method: Definitions and Theorems

Definition 3. *The variables corresponding to the columns of the identity matrix in the initial simplex table are called* ***basic variables*** *while the remaining variables are called* ***nonbasic*** *or* ***free variables****.*

Theorem 1. *A basic feasible solution to a linear programming problem corresponds to an extreme point in the convex set of feasible solutions.*

Simplex Method: Definitions and Theorems

Definition 3. *The variables corresponding to the columns of the identity matrix in the initial simplex table are called* ***basic variables*** *while the remaining variables are called* ***nonbasic*** *or* ***free variables****.*

Theorem 1. *A basic feasible solution to a linear programming problem corresponds to an extreme point in the convex set of feasible solutions.*

Corollary 1. *Each extreme point corresponds to one or more basic feasible solutions. If one of the basic feasible solutions is non-degenerate, then an extreme point corresponds to it uniquely.*

Simplex Method: Solving Minimum Linear Programming Problems

Note that in general a minimum problem can be changed to a maximum problem by realizing that in order to minimize z we must maximize $-z$. That is, we multiply the objective by -1 and solve the resulting maximum problem using the Simplex method.

Simplex Method: Solving Minimum Linear Programming Problems

Note that in general a minimum problem can be changed to a maximum problem by realizing that in order to minimize z we must maximize $-z$. That is, we multiply the objective by -1 and solve the resulting maximum problem using the Simplex method.

Consider the following linear programming problem:

$$\begin{aligned} z = x_1 - 3x_2 + 2x_3 &\to \min \\ 3x_1 - x_2 + 2x_3 &\leqslant 7 \\ -2x_1 + 4x_2 &\leqslant 12 \\ -4x_1 + 3x_2 + 8x_3 &\leqslant 10 \\ x_1 \geqslant 0, x_2 \geqslant 0, x_3 &\geqslant 0 \end{aligned}$$

This is a problem of minimization, so we convert the objective function $z' = -x_1 + 3x_2 - 2x_3 \to \max$ and solve the obtained maximum problem.

Simplex Method: Solving Minimum Linear Programming Problems (continued)

After introducing the slack variables s_1, s_2, and s_3 the problem can be expressed as

$$\begin{aligned}
3x_1 - x_2 + 2x_3 + s_1 &= 7 \\
-2x_1 + 4x_2 + s_2 &= 12 \\
-4x_1 + 3x_2 + 8x_3 + s_3 &= 10 \\
x_1 \geqslant 0, x_2 \geqslant 0, x_3 &\geqslant 0 \\
s_1 \geqslant 0, s_2 \geqslant 0, s_3 &\geqslant 0
\end{aligned}$$

The objective function can be written as $z' + x_1 - 3x_2 + 2x_3 = 0$.

Simplex Method: Solving Minimum Linear Programming Problems (continued)

The initial simplex table for this system is

x_1	x_2	x_3	s_1	s_2	s_3	z'	b	variables	ratio
3	-1	2	1	0	0	0	7	s_1	
-2	4	0	0	1	0	0	12	s_2	
-4	3	8	0	0	1	0	10	s_3	
1	-3	2	0	0	0	1	0	z	

Simplex Method: Solving Minimum Linear Programming Problems (continued)

The initial simplex table for this system is

x_1	x_2	x_3	s_1	s_2	s_3	z'	b	variables	ratio
3	-1	2	1	0	0	0	7	s_1	
-2	4	0	0	1	0	0	12	s_2	$12/4 = 3$
-4	3	8	0	0	1	0	10	s_3	$\frac{10}{3}$
1	-3	2	0	0	0	1	0	z	

Simplex Method: Solving Minimum Linear Programming Problems (continued)

The initial simplex table for this system is

x_1	x_2	x_3	s_1	s_2	s_3	z'	b	variables	ratio
3	-1	2	1	0	0	0	7	s_1	
-2	4	0	0	1	0	0	12	s_2	$12/4=3$
-4	3	8	0	0	1	0	10	s_3	$\frac{10}{3}$
1	-3	2	0	0	0	1	0	z	

Row operations: Divide R_2 by 4 and apply $R_1 = R_1 + R_2$, $R_3 = R_3 - 3R_2$, and $R_4 = R_4 + 3R_2$.

Simplex Method: Solving Minimum Linear Programming Problems (continued)

We thus obtain the 2nd table:

x_1	x_2	x_3	s_1	s_2	s_3	z'	b	variables	ratio
$\frac{5}{2}$	0	2	1	$\frac{1}{4}$	0	0	10	s_1	
$-\frac{1}{2}$	1	0	0	$\frac{1}{4}$	0	0	3	x_2	
$-\frac{5}{2}$	0	8	0	$-\frac{3}{4}$	1	0	1	s_3	
$-\frac{1}{2}$	0	2	0	$\frac{3}{4}$	0	1	9	z	

Simplex Method: Solving Minimum Linear Programming Problems (continued)

We thus obtain the 2nd table:

x_1	x_2	x_3	s_1	s_2	s_3	z'	b	variables	ratio
$\mathbf{\frac{5}{2}}$	0	2	1	$\frac{1}{4}$	0	0	10	s_1	4
$-\frac{1}{2}$	1	0	0	$\frac{1}{4}$	0	0	3	x_2	ignore
$-\frac{5}{2}$	0	8	0	$-\frac{3}{4}$	1	0	1	s_3	ignore
$-\frac{1}{2}$	0	2	0	$\frac{3}{4}$	0	1	9	z	

Simplex Method: Solving Minimum Linear Programming Problems (continued)

We thus obtain the 2nd table:

x_1	x_2	x_3	s_1	s_2	s_3	z'	b	variables	ratio
$\mathbf{\frac{5}{2}}$	0	2	1	$\frac{1}{4}$	0	0	10	s_1	4
$-\frac{1}{2}$	1	0	0	$\frac{1}{4}$	0	0	3	x_2	ignore
$-\frac{5}{2}$	0	8	0	$-\frac{3}{4}$	1	0	1	s_3	ignore
$-\frac{1}{2}$	0	2	0	$\frac{3}{4}$	0	1	9	z	

Row operations: Divide R_1 by $\frac{5}{2}$ and apply the operations $R_2 = R_2 + \frac{1}{2}R_1$, $R_3 = R_3 + \frac{5}{2}R_1$, and $R_4 = R_4 + \frac{1}{2}R_1$.

Simplex Method: Solving Minimum Linear Programming Problems (continued)

And then the 3rd table:

x_1	x_2	x_3	s_1	s_2	s_3	z'	b	variables
1	0	$\frac{4}{5}$	$\frac{2}{5}$	$\frac{1}{10}$	0	0	4	x_1
0	1	$\frac{2}{5}$	$\frac{1}{5}$	$\frac{3}{10}$	0	0	5	x_2
0	0	10	1	$-\frac{1}{2}$	1	0	11	s_3
0	0	$\frac{12}{5}$	$\frac{1}{5}$	$\frac{4}{5}$	0	1	11	z

Simplex Method: Solving Minimum Linear Programming Problems (continued)

And then the 3rd table:

x_1	x_2	x_3	s_1	s_2	s_3	z'	b	variables
1	0	$\frac{4}{5}$	$\frac{2}{5}$	$\frac{1}{10}$	0	0	4	x_1
0	1	$\frac{2}{5}$	$\frac{1}{5}$	$\frac{3}{10}$	0	0	5	x_2
0	0	10	1	$-\frac{1}{2}$	1	0	11	s_3
0	0	$\frac{12}{5}$	$\frac{1}{5}$	$\frac{4}{5}$	0	1	11	z

Simplex Method: Solving Minimum Linear Programming Problems (continued)

And then the 3rd table:

x_1	x_2	x_3	s_1	s_2	s_3	z'	b	variables
1	0	$\frac{4}{5}$	$\frac{2}{5}$	$\frac{1}{10}$	0	0	4	x_1
0	1	$\frac{2}{5}$	$\frac{1}{5}$	$\frac{3}{10}$	0	0	5	x_2
0	0	10	1	$-\frac{1}{2}$	1	0	11	s_3
0	0	$\frac{12}{5}$	$\frac{1}{5}$	$\frac{4}{5}$	0	1	11	z

In the above table we see that there are no negative entries in the objective row. Hence, the optimal solution is found. Therefore, $z'_{\max} = 11$ for $x_1 = 4, x_2 = 5, x_3 = 0, s_1 = 0, s_2 = 0, s_3 = 11$. Hence, the solution of the original problem is $z_{\min}(4, 5, 0) = -11$.

Simplex Method: Exercise

Solve the following linear programming problem by means of the Simplex method.

$$
\begin{aligned}
z = 4x_1 + 3x_2 &\to \max \\
3x_1 + x_2 &\leqslant 9 \\
-x_1 + x_2 &\leqslant 1 \\
x_1 + x_2 &\leqslant 6 \\
x_1 \geqslant 0, x_2 &\geqslant 0
\end{aligned}
$$

Simplex Method: Summary

1. Add slack variables to change the constraints into equations and write all variables to the left of the equal sign and constants to the right.
2. Write the objective function with all nonzero terms to the left of the equal sign and zero to the right. The variable to be maximized must be positive.
3. Set up the initial simplex table by creating an augmented matrix from the equations, placing the equation for the objective function last.
4. Determine a pivot element and use matrix row operations to convert the column containing the pivot element into a unit column.
5. If negative elements still exist in the objective row, repeat **Step 4**. If all elements in the objective row are positive, terminate the process.
6. When the final matrix has been obtained, determine the final basic solution. This will give the maximum value for the objective function and the values of the variables where this maximum occurs.

Simplex Method: Special Cases

- **Alternate optimal solutions**. The linear programming problem has alternative optimal solutions (multiple optimal solutions) if at least one of the coefficients of the nonbasic variable in the objective row equals to zero.

Simplex Method: Special Cases

- **Alternate optimal solutions**. The linear programming problem has alternative optimal solutions (multiple optimal solutions) if at least one of the coefficients of the nonbasic variable in the objective row equals to zero.
- **Degeneracy**. A situation, when a basic solution has a basic variable that is equal to zero. The basic solution is then said to be degenerate. This situation may lead to *cycling*—a sequence of pivots that goes through the same tables and repeats itself indefinitely. Fortunately, cycling is rare, though degeneracy is frequent.

Simplex Method: Special Cases

- **Alternate optimal solutions**. The linear programming problem has alternative optimal solutions (multiple optimal solutions) if at least one of the coefficients of the nonbasic variable in the objective row equals to zero.
- **Degeneracy**. A situation, when a basic solution has a basic variable that is equal to zero. The basic solution is then said to be degenerate. This situation may lead to *cycling*—a sequence of pivots that goes through the same tables and repeats itself indefinitely. Fortunately, cycling is rare, though degeneracy is frequent.
- **Unbounded optimum**. If an identified pivot column consists of only negative entries, the problem is unbounded.

Nelder-Mead Simplex Method

- The Nelder-Mead simplex method is used for solving unconstrained optimization problems of the form

$$\min_x F(x), \quad x \in \mathbb{R}^n.$$

- It is a direct search method, and is therefore well-suited to optimize a function whose derivatives are unknown or non-existent.
- The method typically produces significant improvement of a performance measure in industrial control applications within the first few iterations.
- It may be applied in cases, when cost function evaluation is relatively expensive.
- It is difficult to formulate stopping criteria for the method, since the iterates may enter a local minimum, and very little progress is made in terms of improving performance over the next iterations. Thus, a limit on the maximum number of iterations should be established.

Nelder-Mead Simplex Method: Summary

First, an initial simplex is constructed by determining $n+1$ vertices along with corresponding values of F. The kth iteration then consists of the following steps:

1. **Order.** Order the $n+1$ vertices, so that $F(x_1) \leqslant F(x_2) \leqslant \cdots \leqslant F(x_{n+1})$ is satisfied. Apply tie-breaking rules when necessary.
2. **Reflect.** Compute the reflection point $x_r = \bar{x} + \rho(\bar{x} - x_{n+1})$, where

$$\bar{x} = \sum_{i=1}^{n} x_i / n$$

   is the centroid of the n best vertices. Evaluate $F_r = F(x_r)$. If $F_1 \leqslant F_r < F_n$, set $x_{n+1} = x_r$ and terminate the iteration.

Nelder-Mead Simplex Method: Summary (continued)

3 **Expand.** If $F_r < F_1$, calculate the expansion point $x_e = \bar{x} + \chi(x_r - \bar{x})$, and evaluate $F_e = F(x_e)$. If $F_e < F_r$, set $x_{n+1} = x_e$ and terminate the iteration. Otherwise, set $x_{n+1} = x_r$ and terminate the iteration.

4 **Contract.** If $F_r \geqslant F_n$, perform a contraction between $\bar{x}$ and the better of x_{n+1} and x_r:

(a) **Contract outside.** If $F_n \leqslant F_r < F_{n+1}$, calculate $x_c = \bar{x} + \gamma(x_r - \bar{x})$ and evaluate $F_c = F(x_c)$. If $F_c \leqslant F_r$, set $x_{n+1} = x_c$ and terminate the operation. Otherwise, go to Step 5 (perform a shrink).

(b) **Contract inside.** If $F_r \geqslant F_{n+1}$, perform an inside contraction: calculate $x'_c = \bar{x} - \gamma(\bar{x} - x_{n+1})$ and evaluate $F'_c = F(x'_c)$. If $F'_c{<}F_{n+1}$ set $x_{n+1} = x'_c$ and terminate the iteration. Otherwise, go to Step 5 (perform a shrink).

Nelder-Mead Simplex Method: Summary (continued)

5 **Shrink.** Define n new vertices from

$$x_i = x_1 + \sigma(x_i - x_1), \quad i = 2, \ldots, n+1,$$

and evaluate F at these points.

In the algorithm described above four scalar coefficients are used, i.e., the coefficients of reflection, expansion, contraction, and shrinkage, denoted by ρ, χ, γ, and σ, respectively. According to the original paper, these coefficients should satisfy

$$\rho > 0, \quad \chi > 1, \quad 0 < \gamma < 1, \quad 0 < \sigma < 1.$$

Nelder-Mead Simplex Method: Examples

- **Rosenbrock function**

$$f_R(x_1, x_2) = (1 - x_1)^2 + 100(x_2 - x_1^2)^2.$$

It has a global minimum at $f(1, 1) = 0$.

Nelder-Mead Simplex Method: Examples

- **Rosenbrock function**

$$f_R(x_1, x_2) = (1 - x_1)^2 + 100(x_2 - x_1^2)^2.$$

It has a global minimum at $f(1,1) = 0$.

- **Himmelblau's function**

$$f_H(x_1, x_2) = (x_1^2 + x_2 - 11)^2 + (x_1 + x_2^2 - 7)^2.$$

It has four local minima:

  - $f(3,2) = 0$,
  - $f(-2.805118, 3.131312) = 0$,
  - $f(-3.779310, -3.283186) = 0$,
  - $f(3.584428, -1.848126) = 0$.

General Nonlinear Optimization Problem

Recall that the single objective optimization problem can be written as

$$\min_{x \in \mathbb{R}^n} f(x) \quad \text{subject to} \begin{cases} c_i^{ne}(x) & = & 0, & i \in \mathscr{I}, \\ c_j^{ni}(x) & \leqslant & 0, & j \in \mathscr{E}, \\ x_k^L \leqslant x_k & \leqslant & x_k^U, & k \in \mathscr{K}, \end{cases} \tag{1}$$

where f, c_i^{ne} and c_j^{ni} are scalar-valued functions of the variables x, x_k^L and x_k^U are lower and upper scalar bounds, respectively, and $\mathscr{I}$,$\mathscr{E}$, and $\mathscr{K}$ are sets of indices.

Problems with Bounds and Constraints for Unconstrained Optimization Algorithms

- While certain optimization algorithms, such as the Nelder-Mead method, are designed to solve unconstrained problems unbounded in the search space, it is possible to introduce both variable search space bounds and constraints in terms of coordinate transformations and penalty functions.

Problems with Bounds and Constraints for Unconstrained Optimization Algorithms

- While certain optimization algorithms, such as the Nelder-Mead method, are designed to solve unconstrained problems unbounded in the search space, it is possible to introduce both variable search space bounds and constraints in terms of coordinate transformations and penalty functions.
- For search variable bounds coordinate transformations may be used. We use quadratic and trigonometric transformations.

Problems with Bounds and Constraints for Unconstrained Optimization Algorithms

- While certain optimization algorithms, such as the Nelder-Mead method, are designed to solve unconstrained problems unbounded in the search space, it is possible to introduce both variable search space bounds and constraints in terms of coordinate transformations and penalty functions.
- For search variable bounds coordinate transformations may be used. We use quadratic and trigonometric transformations.
- For introducing constraints we use penalizing terms (through the use of penalty functions) which are added to the cost function.

Problems with Bounds and Constraints: Coordinate Transformation

- Let x denote the vector of search variables of size $N \times 1$. For bound constraints, a coordinate transformation may be applied to each individual search variable. Let x_i^L and x_i^U denote the lower bound and upper bound on the ith search parameter, respectively, and denote by z the new search variable vector. Let $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ denote a coordinate transformation function, such that $x = \varphi(z)$ and $z = \varphi^{-1}(x)$.

Problems with Bounds and Constraints: Coordinate Transformation

- Let x denote the vector of search variables of size $N \times 1$. For bound constraints, a coordinate transformation may be applied to each individual search variable. Let x_i^L and x_i^U denote the lower bound and upper bound on the ith search parameter, respectively, and denote by z the new search variable vector. Let $\varphi : \mathbb{R} \to \mathbb{R}$ denote a coordinate transformation function, such that $x = \varphi(z)$ and $z = \varphi^{-1}(x)$.
- For problems with lower bounds we have

$$x_i = x_i^L + z_i^2,$$

from which it follows that $x_i \geqslant x_i^L$, since $z_i^2 \geqslant 0$. Initial estimates $z_{i,0}$ are obtained from original initial estimates $x_{i,0} \geqslant x_i^L$ as

$$z_{i,0} = \sqrt{x_{i,0} - x_i^L}.$$

Problems with Bounds and Constraints: Coordinate Transformation (continued)

- For problems with upper bounds we have

$$x_i = x_i^U - z_i^2,$$

from which it follows that $x_i \leqslant x_i^U$, since $-z_i^2 \leqslant 0$. Initial estimates $z_{i,0}$ are obtained from original initial estimates $x_{i,0} \leqslant x_i^U$ as

$$z_{i,0} = \sqrt{x_i^U - x_{i,0}}.$$

Problems with Bounds and Constraints: Coordinate Transformation (continued)

- For problems with lower and upper bounds, we have

$$x_i = x_i^L + (x_i^U - x_i^L)\frac{\sin(z_i)+1}{2},$$

from which it follows that $x_i^L \leqslant x_i \leqslant x_i^U$, since the values of the function $f(z_i) = (\sin(z_i)+1)/2$ are always inside of the interval $[0,1]$ and x_i is therefore bounded by $f(z_i) = 0 \Rightarrow x_i = x_i^L$, $f(z_i) = 1 \Rightarrow x_i = x_i^U$. Initial estimates $z_{i,0}$ are obtained from original initial estimates $x_i^L \leqslant x_{i,0} \leqslant x_i^U$ as

$$z_{i,0} = \Re\left(\arcsin\left(\frac{-2x_{i,0} + x_i^L + x_i^U}{x_i^L - x_i^U}\right)\right),$$

where $\Re(\cdot)$ denotes the real part.

Problems with Bounds and Constraints: Penalty Functions

To introduce constraints, a modification of the cost function

$$\kappa(\cdot) = \kappa^*,$$

where κ^* denotes the cost for the original optimization problem, is necessary. Let us define a function $f_{nz} : \mathbb{R} \to \mathbb{R}_+$ such that

$$f_{nz}(x) := \begin{cases} x, & x > 0 \\ 0, & x \leqslant 0. \end{cases}$$

Let us also define a *penalty function* $\rho : \mathbb{R} \to \mathbb{R}$ such that

$$\rho(x) := \begin{cases} \mathrm{e}^{\gamma} - 1 + x, & x > \gamma \\ \mathrm{e}^{x} - 1, & x \leqslant \gamma, \end{cases}$$

where $\gamma > 0$ is some predefined constant.

Problems with Bounds and Constraints: Penalizing Nonlinear Equality Constraints

For general nonlinear equality constraints of the form $c^{ne}(\cdot) \leqslant 0$, where $c^{ne} : \mathbb{R}^{q \times r} \to \mathbb{R}^{N_{ne} \times M_{ne}}$, we define the penalty function $\kappa^{ne} : \mathbb{R}^{N_{ne} \times M_{ne}} \to \mathbb{R}$ as

$$\kappa^{ne}(c^{ne}(\cdot)) := \rho(c^{ne}_{\Sigma}(\cdot)),$$

where

$$c^{ne}_{\Sigma}(\cdot) = \sum_{i=1}^{N_{ne}} \sum_{k=1}^{M_{ne}} f_{nz}(|c^{ne}_{i,k}(\cdot)|),$$

where $|\cdot|$ denotes the absolute value.

Problems with Bounds and Constraints: Penalizing Nonlinear Inequality Constraints

For general nonlinear inequality constraints of the form $c^{ni}(\cdot) \leqslant 0$, where $c^{ni} : \mathbb{R}^{q \times r} \to \mathbb{R}^{N_{ni} \times M_{ni}}$, we define the following penalty function $\kappa^{ni} : \mathbb{R}^{N_{ni} \times M_{ni}} \to \mathbb{R}$ as

$$\kappa^{ni}(c^{ni}(\cdot)) := \rho(c_{\Sigma}^{ni}(\cdot)),$$

where

$$c_{\Sigma}^{ni}(\cdot) = \sum_{i=1}^{N_{ni}} \sum_{k=1}^{M_{ni}} f_{nz}(c_{i,k}^{ni}(\cdot)).$$

Thus, taking both equality and inequality constraints into account, the new cost function is

$$\kappa(\cdot) = \kappa^{*} + \underbrace{\kappa^{ne} + \kappa^{ni}}_{\text{Penalizing terms}} .$$

Questions?

Thank you for your attention!