Optimization

Find the **optimum** (maximum or minimum) of an **objective function** f

Application examples:
Process Design: Find the operating conditions of temperature, pressure and catalyst concentration to maximize the yield of a chemical reactor.
Resource Allocation: For a production process, find the amounts of various inputs that minimizes the total manufacturing cost.

Objective Function

In general, the objective function depends on several variables: $f(x_1, x_2, \dots, x_n)$, e.g. the yield of the reactor may depend on temperature, pressure, and catalyst concentration.These variables are termed the **control variables**.

In most realistic optimization problems, the control variables are **constrained**,e.g. the reactor may be able to operate only within a specific temperature range,constrained by the thermodynamics of the reaction and reactor vessel safety considerations.

The objective function may be determined from fundamental considerations, e.g. by writing a model for the process based on conservation balances, or it may be determined empirically, e.g. by curve-fitting a function to process data.

Outline

- Unconstrained Optimization
- Constrained Optimization
- Linear Programming

A maximization problem can be converted into a minimization problem in terms of the negative of the objective function, and vice versa.

Minimum of a function of one variable

The necessary and sufficient condition for f to have a minimum at x=x* is:

$$h(x^*) = f'(x^*) = \left.\frac{df}{dx}\right|_{x=x^*} = 0$$

and

$$h'(x^*) = f''(x^*) = \left.\frac{d^2f}{dx^2}\right|_{x=x^*} > 0$$

e.g. f(x)=$x^2 - x \Rightarrow h(x) = 2x - 1$, $h'(x) = 2$
$\Rightarrow x^* = 1/2$ is root of h(x), and a minimum of f(x).
Note: The problem of minimization of the function f(x) has been transformed to that of finding the root of the function $h(x) = f'(x)$.
The root of a function h(x) is the value of x for which h(x)=0. Numerical methods may be necessary to find the root.

Bisection Method

Method: At each step, the search interval is reduced by a factor of 2.
Principle: A function has at least one root in an interval in which it changes sign.
Problem: Find root of h(x) in interval [a,b].
Procedure: Evaluate h(a), h(b) and h(c=(a+b)/2)
If h(a)h(c)$< 0 \Rightarrow x^* \in$[a,c]
If h(b)h(c)$< 0 \Rightarrow x^* \in$[c,b]
Continue procedure until convergence criteria is satisfied, e.g. h(x)$< 10^{-6}$
Advantages:
Simple to implement
Disadvantages:
Slow.
Not useful when the function depends on more than one variable.

Example 1

Find minimum of $f(x) = x^4/4 + 4x^3/3 - 10x + 1$ in the interval [1,2]$\Leftrightarrow$ find root of $h(x) = x^3 + 4x^2 - 10$ in the interval [1,2].

Step 1: h(1)=-5, h(2)=14, h(1.5)=2.375 $\Rightarrow x^* \in$[1,1.5]

Step 2: h(1.25)=-1.796 $x^* \in$[1.25,1.5]

Converges to $x^* = 1.365$ after twelve steps.

Newton-Raphson Method

Iterative Sequence:

$$x_{k+1} = x_k - \frac{h(x_k)}{h'(x_k)} = x_k - \frac{f'(x_k)}{f''(x_k)} \qquad k = 0, 1, \ldots$$

Requires an initial estimate x_0.

Advantages

Rapid convergence if initial estimate is good.

Disadvantages

Requires good initial estimate, or convergence can be slow or may even diverge.

Requires evaluation of derivatives.

Example 2

$h(x) = x^3 + 4x^2 - 10, x_0 = 1.5$

$$x_1 = x_0 - \frac{h(x_0)}{h'(x_0)} = 1.5 - \frac{h(1.5)}{h'(1.5)} = 1.5 - \frac{2.375}{18.75} = 1.3733$$

$$x_2 = x_1 - \frac{h(x_1)}{h'(x_1)} = 1.3733 - \frac{0.1343}{16.645} = 1.3653$$

Converges to $x^* = 1.365$ in 2 steps.

Minimize f($\mathbf{x}$).

f is a scalar function of the control variables $\mathbf{x}=[x_1, x_2, \ldots, x_n]$

If f has a minimum at $\mathbf{x}=\mathbf{x}^* \Rightarrow \nabla f|_{\mathbf{x}=\mathbf{x}^*} = 0$.

Newton's Method

At minimum (using Taylor series expansion for ∇f):

$$\nabla f(\mathbf{x}^{k+1}) = \nabla f(\mathbf{x}^k) + \mathbf{H}(\mathbf{x}^k).\Delta\mathbf{x}^k = 0 \qquad \mathbf{H} = \nabla\nabla f$$

$$\Rightarrow \mathbf{H}(\mathbf{x}^k).\Delta\mathbf{x}^k = -\nabla f(\mathbf{x}^k)$$

$$\Rightarrow \Delta\mathbf{x}^k = \mathbf{x}^{k+1} - \mathbf{x}^k = -\mathbf{H}^{-1}(\mathbf{x}^k).\nabla f(\mathbf{x}^k)$$

$$f(x) = (x_1 - 2)^4 + (x_1 - 2)^2 x_2^2 + (x_2 + 1)^2 \qquad \mathbf{x}^0 = [1 \ 1]$$

$$\nabla f = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \end{bmatrix} = \begin{bmatrix} 4(x_1 - 2)^3 + 2(x_1 - 2)x_2^2 \\ 2(x_1 - 2)^2 x_2 + 2(x_2 + 1) \end{bmatrix}$$

$$\mathbf{H} = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} \end{bmatrix}$$

$$= \begin{bmatrix} 12(x_1 - 2)^2 + 2x_2^2 & 4(x_1 - 2)x_2 \\ 4(x_1 - 2)x_2 & 2(x_1 - 2)^2 + 2 \end{bmatrix}$$

$$\nabla f(\mathbf{x}^0) = \begin{bmatrix} -6 \\ +6 \end{bmatrix} \qquad \mathbf{x}^0 = [1 \ 1]$$

$$\mathbf{H}(\mathbf{x}^0) = \begin{bmatrix} 14 & -4 \\ -4 & 4 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} 14 & -4 \\ -4 & 4 \end{bmatrix} \begin{bmatrix} \Delta x_1 \\ \Delta x_2 \end{bmatrix} = \begin{bmatrix} +6 \\ -6 \end{bmatrix}$$

$$\Rightarrow \Delta x_1 = 0 \qquad \Delta x_2 = -1.5$$

#	x_1	x_2
0	+1.000	+1.000
1	+1.000	-0.500
2	+1.391	-0.696
3	+1.746	-0.949
4	+1.986	-1.048
5	+1.999	-1.000

Method of Steepest Descent

Minimize f($\mathbf{x}$).

$$\mathbf{x}^{k+1} = \mathbf{x}^k - t\nabla f(\mathbf{x}^k)$$

t is the minimum of g(t)=$f(\mathbf{x}^{k+1})$, or the root of g'(t).

Does not require evaluation of second derivatives.

Example

$$f(x) = (x_1 - 2)^4 + (x_1 - 2)^2 x_2^2 + (x_2 + 1)^2 \qquad \mathbf{x}^0 = [1 \;\; 1]$$

$$\nabla f = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \end{bmatrix} = \begin{bmatrix} 4(x_1 - 2)^3 + 2(x_1 - 2)x_2^2 \\ 2(x_1 - 2)^2 x_2 + 2(x_2 + 1) \end{bmatrix}$$

$$\mathbf{x}^1 = \mathbf{x}^0 - t\nabla f(\mathbf{x}^0) = \begin{bmatrix} 1 \\ 1 \end{bmatrix} - t \begin{bmatrix} -6 \\ +6 \end{bmatrix} = \begin{bmatrix} 1 + 6t \\ 1 - 6t \end{bmatrix}$$

$$g(t) = f(\mathbf{x}^1) = 2(1 - 6t)^4 + (2 - 6t)^2$$

Find the value of t for which g'(t)=0 by a root finding method,

$t \approx 0.25 \Rightarrow \mathbf{x}^1$ =[1+6(0.25) 1-6(0.25)]=[2.5, -0.5].

Continuing one finds the same solution as before