Nonlinear Least Squares Estimation

Jun Zhang

SAMSRI

May 13, 2010
Outline

1. An Example
2. Newton’s Method and Gauss-Newton Method
3. Nelder-Mead Method
Outline

1. An Example
2. Newton’s Method and Gauss-Newton Method
3. Nelder-Mead Method
Suppose we want to study the effects of certain medicine on a patient, we draw a blood sample at time $t_j$ and measure the concentration $y_j$ for each sample. We have $n$ pairs of $(t_j, y_j)$. 
Based on previous expert experiences, we find that the following function \( \phi(x; t) \) may be a good fit for the relationship

\[
\phi(x; t) = x_1 + x_2 t + x_3 e^{-x_4 t}
\]
Figure:  *Relationship between concentration of medicine and time*
To estimate the parameter vector $x = [x_1, \ldots, x_4]$, we write the difference between the predicted values and observed values:

$$\min_x f(x) = \frac{1}{2} \sum_{j=1}^{n} [\phi(x; t_j) - y_j]^2$$
Figure:  *Relationship between concentration of medicine and time*
True Function:
\[ \phi(x; t) = 0.1 - 0.2t + 2.1e^{-1.3t} \]

Fitted Function:
\[ \phi(x; t) = -0.0183 - 0.1918t + 2.0976e^{-1.0064t} \]
If the objective function $f(x)$ has the form

$$f(x) = \frac{1}{2} \sum_{j=1}^{n} r_j^2(x) = \frac{1}{2} ||r(x)||^2$$

where $r_j$ is a smooth function from $R^m$ to $R$, and $r_j$ is called residual. We assume $n > m$. If the residuals $r_j(x)$’s are nonlinear function of $x$, then the problem is a Nonlinear Least Squares Estimation (NLSE) problem.
Outline

1. An Example
2. Newton’s Method and Gauss-Newton Method
3. Nelder-Mead Method
We have seen an example of Nonlinear Least Squares Estimation (NLSE).

We will introduce how to use Gauss-Newton Method to find the solution for an NLSE problem.

To understand Gauss-Newton Method, we need to understand Newton’s Method first.
Newton’s Method: Univariate Case

Target Function:

\[ f(x) = \frac{1}{3}x^3 - 4x \]

Derivative of Target Function:

\[ f'(x) = x^2 - 4 \]

Problem: how to find the extrema of the target function iteratively?
Newton’s Method: Univariate Case

The graph shows a function $f(x)$ plotted against $x$, with the function values ranging from $-300$ to $300$. The x-axis ranges from $-10$ to $10$. The curve indicates a typical behavior of a function where the rate of change increases as $x$ moves away from zero.
Newton’s Method: Univariate Case

The condition for extrema:

\[ f'(x) = 0 \]

We can use Newton’s Method to find the roots of the derivative.
Newton’s Method: Univariate Case

Figure: The slope of tangent line at $x_k$ is $f''(x_k)$. We need to find $x_{k+1}$ where the tangent line and $x$ axis intersect.
Newton’s Method: Univariate Case

When we know two points \((x_k, f'(x_k))\) and \((x_{k+1}, 0)\) on the tangent line, we can express the slope \(f''(x_k)\) as

\[
f''(x_k) = \frac{0 - f'(x_k)}{x_{k+1} - x_k}
\]

and we have the iterative formula

\[
x_{k+1} = x_k - \frac{f'(x_k)}{f''(x_k)}
\]

Note here \(f'(x_{k+1}) \neq 0\). But what if \(x_k\) is very close to the true root?
Answer: at the neighborhood of the true root, the tangent line is a good approximation of the original derivative function $f'(x)$. In other words, $f'(x_{k+1}) \approx 0$ and $x_{k+1}$ could be a good approximation of the true root.
Root-finding by Newton-Raphson Method: $x^2 - 4 = 0$
Root-finding by Newton-Raphson Method: \( x^2 - 4 = 0 \)
Root-finding by Newton-Raphson Method: $x^2 - 4 = 0$
Root-finding by Newton-Raphson Method: $x^2 - 4 = 0$
Root-finding by Newton-Raphson Method: $x^2 - 4 = 0$
The condition of extrema is

$$\nabla f(x_{k+1}) = 0$$

With Multivariate Taylor expansion we obtain

$$\nabla f(x_{k+1}) \approx \nabla f(x_k) + (x_{k+1} - x_k) \nabla^2 f(x_k)$$
Newton’s Method Multivariate Case

By simply setting the derivative $\nabla f(x_{k+1})$ to 0, we obtain the multivariate Newton’s Method iterative formula:

$$\nabla f(x_k) + (x_{k+1} - x_k)\nabla^2 f(x_k) = 0$$

$$x_{k+1} = x_k - (\nabla^2 f(x_k))^{-1}\nabla f(x_k)$$

Jun Zhang  Nonlinear Least Squares Estimation
Gauss-Newton Method

\[
\min_x f(x) = \frac{1}{2} \| r(x) \|^2
\]

\[
\nabla^2 f(x_k) \approx \nabla r(x_k)^T \nabla r(x_k)
\]

\[
\nabla f(x_k) = (\nabla r(x_k))^T r(x_k)
\]

\[
 x_{k+1} = x_k - \nabla^2 f(x_k)^{-1} \nabla f(x_k)
\]

\[
\approx x_k - (\nabla r(x_k)^T \nabla r(x_k))^{-1} (\nabla r(x_k))^T r(x_k)
\]
The Approximation of Hessian matrix doesn’t require additional calculations.

In practice, many problems have small residuals and the approximation is accurate and leads to fast convergence.
Outline

1. An Example
2. Newton’s Method and Gauss-Newton Method
3. Nelder-Mead Method
A Simplex in Two Dimensions

- Evaluate function at vertices

- Note:
  - The highest (worst) point
  - The next highest point
  - The lowest (best) point

- Intuition:
  - Move away from high point, towards low point
Summary: The Simplex Method

Original Simplex

reflection

reflection and expansion

contraction

multiple contraction
Nelder-Mead Method

Direction for Optimization

- Line through worst point and average of other points
- Average of all points, excluding worst point
Nelder-Mead Method

Reflection

This is the default new trial point
Nelder-Mead Method

Reflection and Expansion

If reflection results in new minimum...

Move further along minimization direction
Nelder-Mead Method

Contraction (One Dimension)

Try a smaller step

If \( x' \) is still the worst point...
Nelder-Mead Method

**Contraction ...**

"passing through the eye of a needle"

If a simple contraction doesn't improve things, then try moving all points towards the current minimum.
Thank You!