提供：
Optical Solutions
株式会社オプティカルソリューションズ
〒101-0032
東京都千代田区岩本町 2-16-2（F ビル 6 階）
TEL：03-5833-1332　FAX：03-3865-3318
http://www.osc-japan.com/
e-mail:info@osc-japan.com

光学設計ノーツ　５２（ver.1.0）

最適化とは、ニュートンーラフソン法の応用について

株式会社タイコ
牛山善太

光学設計における最適化について前回から触れさせて戴いているが、今回は勾配法より効率の良い、ニュートンーラフソン（Newton-Raphson）法(あるいは単にニュートン法とも呼ばれる)について解説させていただきたい。関数化された対象を扱う場合の最適化手法としては大変重要なものであるが、この関数化の部分が、この手法をレンズ設計に持ち込むためのネックとなる。しかし、後述させていただくことになる減衰最小二乗法などの構造もそこから理解しやすくなる。

１．　１次元の場合のニュートンーラフソン法の応用

関数 f(x)が、１回微分 f'(x)のみならず２回微分 f''(x)が可能なものである場合、前回の勾配法よりも効率の良い最適化手法としてニュートンーラフソン法を応用したものがある。

点 x0 近傍の点 x_0＋△xでは、以下如くにのテーラー展開が可能である。

$$f\left(x_0+\Delta x\right)=f\left(x_0\right)+f'\left(x_0\right)\Delta x+\frac{1}{2}f''\left(x_0\right)\Delta x^2+\frac{1}{3!}f'''\left(x_0\right)\Delta x^3+\cdots \quad (1)$$

△ xは微小量であるので，(1)式の△xの 3 次以上の項を無視して(1)式の 2 次までの項をとった関数の最大値を考えるとして、(1)式を△x で微分して、その値が０となる f(x_0+△x)の極値を探すと、

$$f'(x_0)+f''(x_0)\Delta x=0 \tag{2}$$

の解、

$$\Delta x=-\frac{f'(x_0)}{f''(x_0)} \tag{3}$$

を得ればよい。従って x 座標としては

$$x=x_0-\frac{f'(x_0)}{f''(x_0)} \tag{4}$$

として、新たな地点が得られる。この座標を新たな x_0 として(3)式をまた、計算し、新たな x を得る。これを繰り返していくと、x_0 と x の間に大きな差がなくなる。そこが収束点となる。

２． ｎ次元の場合ニュートン－ラフソン法の応用

１次元からｎ次元になっても基本的な考え方は同じであるが、手法としては当然、複雑になるので以下に説明させていただく。今度は初期点を、

$$(x_{01},x_{02},\cdots x_{0n})$$

として、そして、その近傍の点、

$$(x_{01}+\Delta x_1,x_{02}+\Delta x_2,\cdots x_{0n}+\Delta x_n)$$

を考えるときに、(1)式と同様に、

$$f(x_{01}+\Delta x_1,x_{02}+\Delta x_2,\cdots x_{0n}+\Delta x_n)$$

$$=f_0+\sum_{i=1}^{n}\frac{\partial f(x_{01},\cdots x_{0n})}{\partial x_i}\Delta x_i$$

$$+\frac{1}{2}\left(\frac{\partial^2}{\partial x_1^2}\Delta x_1^2+\frac{\partial^2}{\partial x_1\partial x_2}\Delta x_1\Delta x_2+\cdots+\frac{\partial^2}{\partial x_n^2}\Delta x_n^2\right)f_0 \qquad (5)$$

上式の右辺第３項は以下の様に書ける。

$$\frac{1}{2}\left(\frac{\partial}{\partial x_1}\Delta x_1+\frac{\partial}{\partial x_2}\Delta x_2+\cdots+\frac{\partial}{\partial x_n}\Delta x_n\right)^2 f_0$$

ここでは、

$$f_0=f\left(x_{01},x_{02},\cdots x_{0n}\right) \qquad (6)$$

と表した。従って、△x_{i}で偏微分すれば(5)式の△x_{i}がかかっていない項は消えて、

$$\frac{\partial f_0}{\partial x_i}+\frac{1}{2}\left[\left(\frac{2\partial^2}{\partial x_i\partial x_1}\Delta x_i\Delta x_1+\cdots+\frac{\partial^2}{\partial {x_i}^2}\Delta x_i^2+\cdots+\frac{2\partial^2}{\partial x_i\partial x_n}\Delta x_i\,\Delta x_n\right)\right]' f_0$$

$$=\frac{\partial f_0}{\partial x_i}+\left[\left(\frac{\partial^2}{\partial x_i\partial x_1}\Delta x_1+\cdots+\frac{\partial^2}{\partial {x_i}^2}\Delta x_i+\cdots+\frac{\partial^2}{\partial x_i\partial x_n}\Delta x_n\right)\right] f_0 \qquad (7)$$

従って、それぞれの i について、

$$\frac{\partial f_0}{\partial x_i}+\sum_{k=1}^{n}\frac{\partial^2 f_0}{\partial x_i\partial x_k}\Delta x_k=0 \qquad (8)$$

の時、極値をとる。

ところで、以下の様なヘッセ行列（Hessian matrix）と呼ばれるものがある。

$$H(f) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix} \tag{9}$$

これを利用して、(5)式の極値を得るための x_i についての総ての解は、行列を用いて、

$$\nabla \mathbf{f_0} + \mathbf{H}_0 \Delta \mathbf{x} = 0$$

$$\Delta \mathbf{x} = -\mathbf{H}_0{}^{-1} \nabla \mathbf{f_0} \tag{10}$$

と表現できる。従って、新しい x の組は

$$\mathbf{x} = \mathbf{x_0} - \mathbf{H}_0{}^{-1} \nabla \mathbf{f_0} \tag{11}$$

として、繰り返し計算により得られる。

因みに、

$$\nabla \mathbf{f_0} = \begin{pmatrix} \frac{\partial f_0}{\partial x_1} \\ \vdots \\ \frac{\partial f_0}{\partial x_n} \end{pmatrix} \tag{12}$$

$$\Delta \mathbf{x} = \begin{pmatrix} \Delta x_1 \\ \vdots \\ \Delta x_n \end{pmatrix} \quad (13$$

であり、

$$\mathbf{H}_0 = H(f_0)$$

である（(9)式参照）。(11)式を解くためには、右辺における逆行列計算が必要になるが、

$$\mathbf{H}_0 \Delta\mathbf{x} = -\nabla\mathbf{f_0} \tag{14}$$

として、連立方程式を解くと考えても良い。

ニュートン-ラフソン法は収束も非常に早く、原理的にも理解しやすく有用な手法であるが、エンジニアリング的な分野では、繰り返し実行されなければならない、ヘッセ行列の計算、つまり2次微分の計算に困難が発生する。この困難を克服するために、ヘッセ行列の代わりの2次微分を必要としない適当な行列Bを用いて、(14)式を

$$\mathbf{B}_0 \Delta\mathbf{x} = -\nabla\mathbf{f_0} \tag{15}$$

として、解を得ようというのが、準ニュートン法である。ここで、表記を、繰り返し計算のk回目の値と、k+1回目の値というように、繰り返し回数を用いたものに切り替えて、

$$\Delta\mathbf{x}_{k+1} = \mathbf{x}_{k+1} - \mathbf{x}_k \tag{16}$$

$$\mathbf{y}_k = \nabla\mathbf{f}_{k+1} - \nabla\mathbf{f}_k \tag{17}$$

とした時、代替え、近似行列Bは、

$$\mathbf{B}_{k+1} = \mathbf{B}_k - \frac{\mathbf{B}_k \Delta\mathrm{x}_k (\Delta\mathbf{x}_{k+1})^T \mathbf{B}_k}{(\Delta\mathbf{x}_{k+1})^T \mathbf{B}_k \Delta\mathbf{x}_{k+1}} + \frac{\mathbf{y}_k (\mathbf{y}_k)^T}{\mathbf{y}_k{}^T \Delta\mathbf{x}_{k+1}} \tag{18}$$

として得られる（DFP（Davidson-Fletcher-Powell）**法**）。右肩のTは転置ベクトルを表す。また、これとは別に、

$$\mathbf{B}_{k+1} = \mathbf{B}_k + \left(\frac{1 + (\Delta\mathbf{x}_{k+1})^T \mathbf{B}_k \Delta\mathbf{x}_{k+1}}{(\Delta\mathbf{x}_{k+1})^T \mathbf{y}_k} \right) \frac{\mathbf{y}_k (\mathbf{y}_k)^T}{\mathbf{y}_k{}^T \Delta\mathbf{x}_{k+1}} - \frac{\mathbf{y}_k (\Delta\mathbf{x}_{k+1})^T \mathbf{B}_k + \mathbf{B}_k \Delta\mathbf{x}_k \mathbf{y}_k^T}{(\Delta\mathbf{x}_{k+1})^T \mathbf{y}_k}$$

(19)

とする手法もある（BFGS (Broyden-Fletcher-Goldfarb-Shanno)法）。

３．参考文献

1)　大野豊、磯田和男監修：数値計算ハンドブック (オーム社、東京、1996)
2)　金谷健一：これなら分かる最適化数学 (共立出版、東京、2008)
3)　今野浩、山下浩：非線形計画法 (日科技連、東京、1978)
4)　牛山善太、草川徹：シミュレーション光学（東海大学出版会、東京、2003）