

数值解析・最適化工学特論

反復解法

反復解法による解の計算

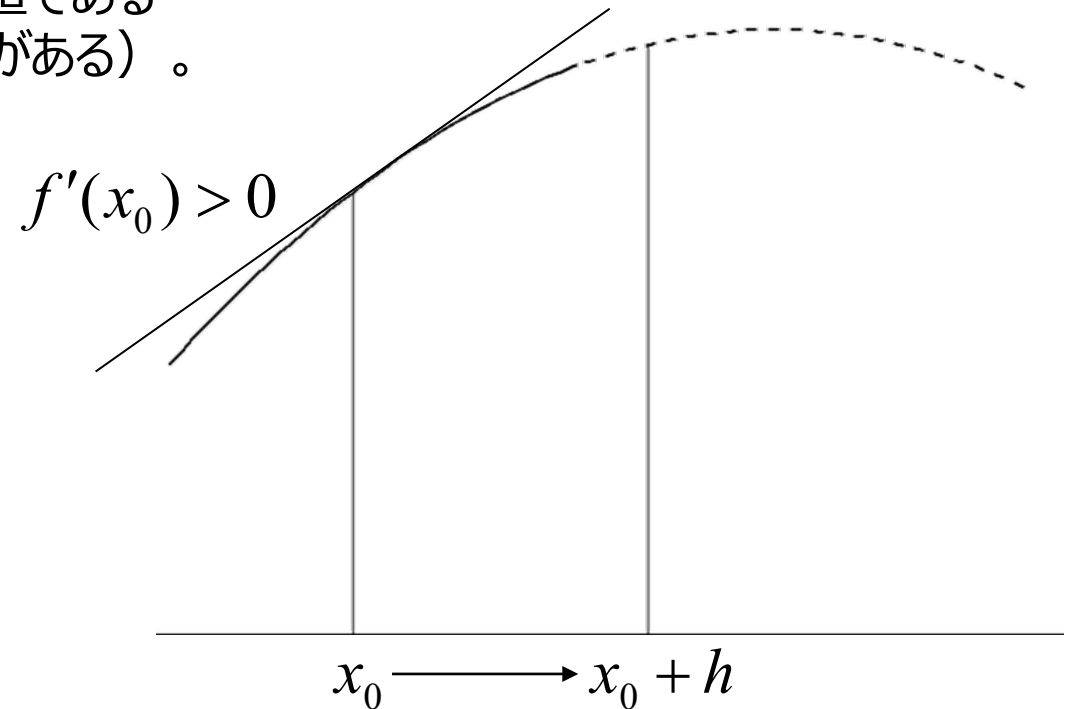
- 最小二乗法のように解を解析的に計算できない問題では、解の初期値を与えて新たな解を計算することを解が収束するまで繰り返す反復解法を用いることが多い。
- 反復解法には次のような方法が存在する。
 - ◆ 勾配法
 - ◆ ニュートン法
 - ◆ ガウス・ニュートン法
 - ◆ レーベンバーグ・マーカート法

勾配法の原理

- 初期値を与えて、**勾配方向**に向かって解を探索する方法
- 1回の反復で勾配方向にどれだけ移動するかによって解が収束するまでの反復回数が増える
 - ◆ 次の反復により解の位置を更新して、位置が変化しなくなるか、もしくは勾配が0になるまで繰り返す。

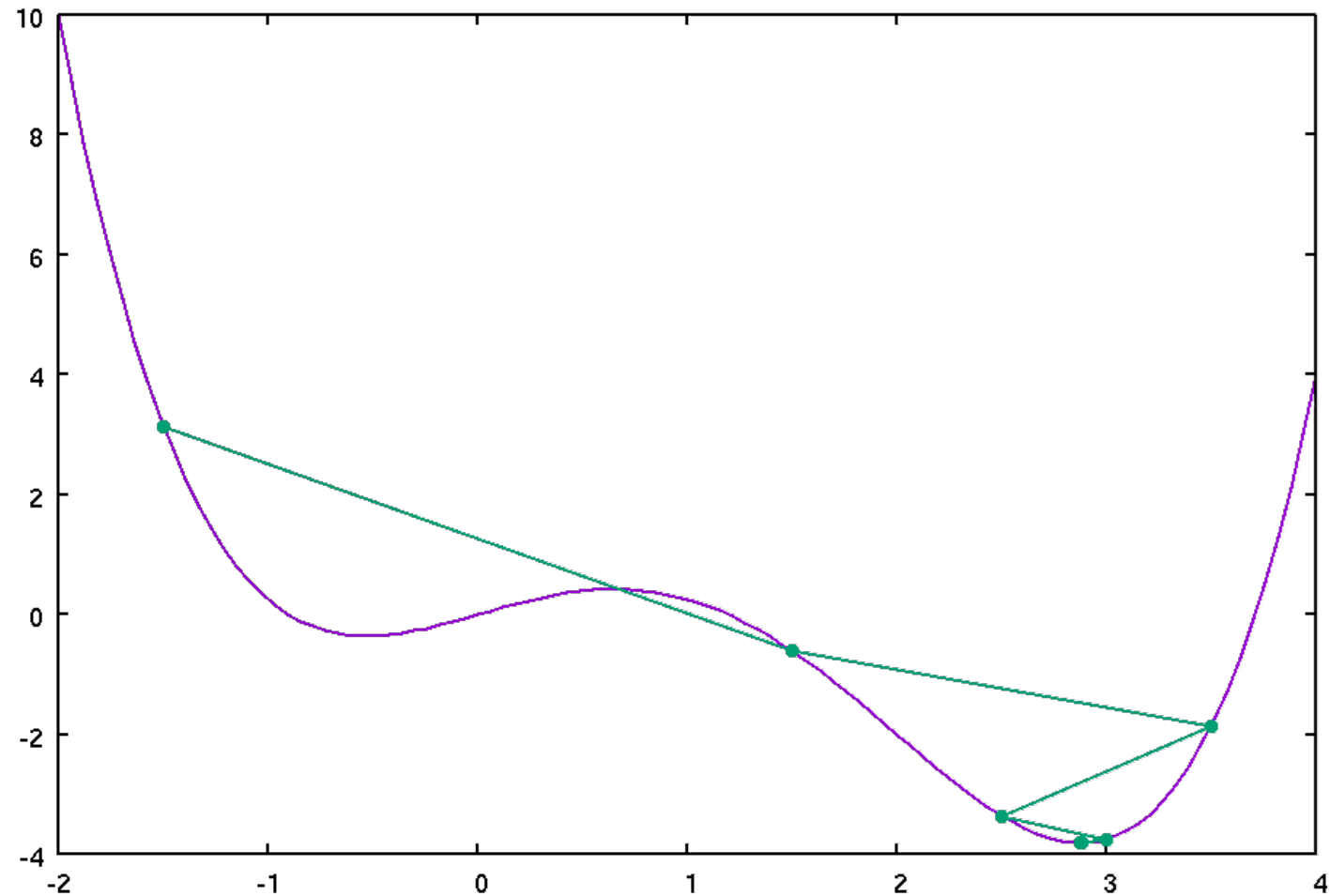
$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} - \alpha \nabla f(\mathbf{x}^{(k)})$$

- ◆ α は勾配方向にどれだけ移動するかを調整する数値である
(機械学習などの分野では**学習率**と呼ばれることがある)。
- 得られた解が**大域解**ではないこともある。
- **共役勾配法**や**確率的勾配法**などの関連手法が存在する



解の探索速度の工夫

- $\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} - \alpha \nabla f(\mathbf{x}^{(k)})$ の α の値を動的に変えることで探索速度を向上させることができる。
 - ◆ 勾配方向が変わらない間は α の値を増加させていく



ニュートン法

- $f(\bar{x} + \Delta x)$ をテーラー展開すると次式を得る。

$$f(\bar{x} + \Delta x) = f(\bar{x}) + f'(\bar{x})\Delta x + \frac{1}{2}f''(\bar{x})\Delta x^2 + \dots$$

- Δx の2次式

$$f_{\text{II}}(\bar{x} + \Delta x) = f(\bar{x}) + f'(\bar{x})\Delta x + \frac{1}{2}f''(\bar{x})\Delta x^2$$

の極値は次式を満たす。

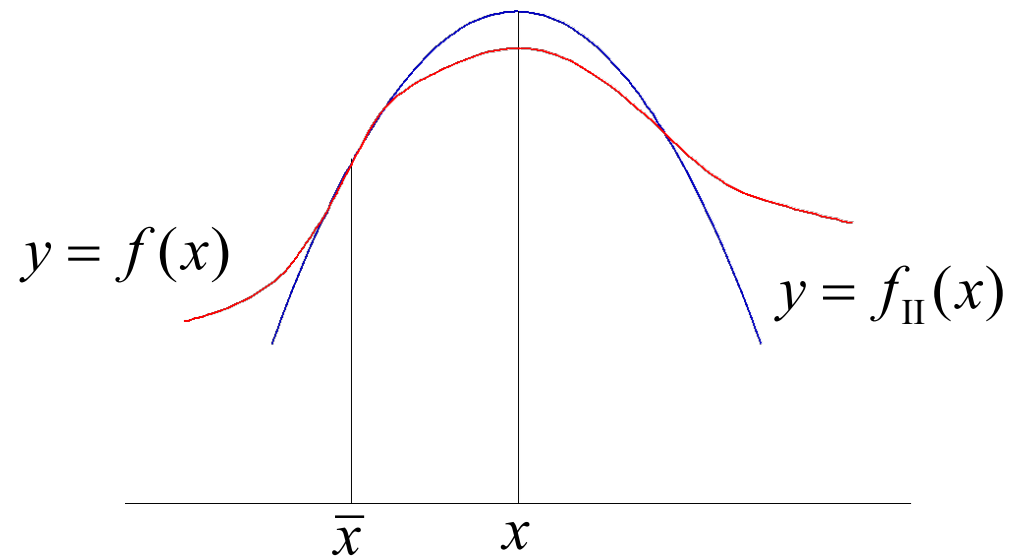
$$f'(\bar{x}) + f''(\bar{x})\Delta x = 0$$

- この解は $\Delta x = -\frac{f'(\bar{x})}{f''(\bar{x})}$ であり、 $x = \bar{x} - \frac{f'(\bar{x})}{f''(\bar{x})}$ は関数 $f(x)$ の極値のよりよい近似となる。

ニュートン法の幾何学的意味

- $$f_{\Pi}(x) = f(\bar{x}) + f'(\bar{x})(x - \bar{x}) + \frac{1}{2} f''(\bar{x})(x - \bar{x})^2$$

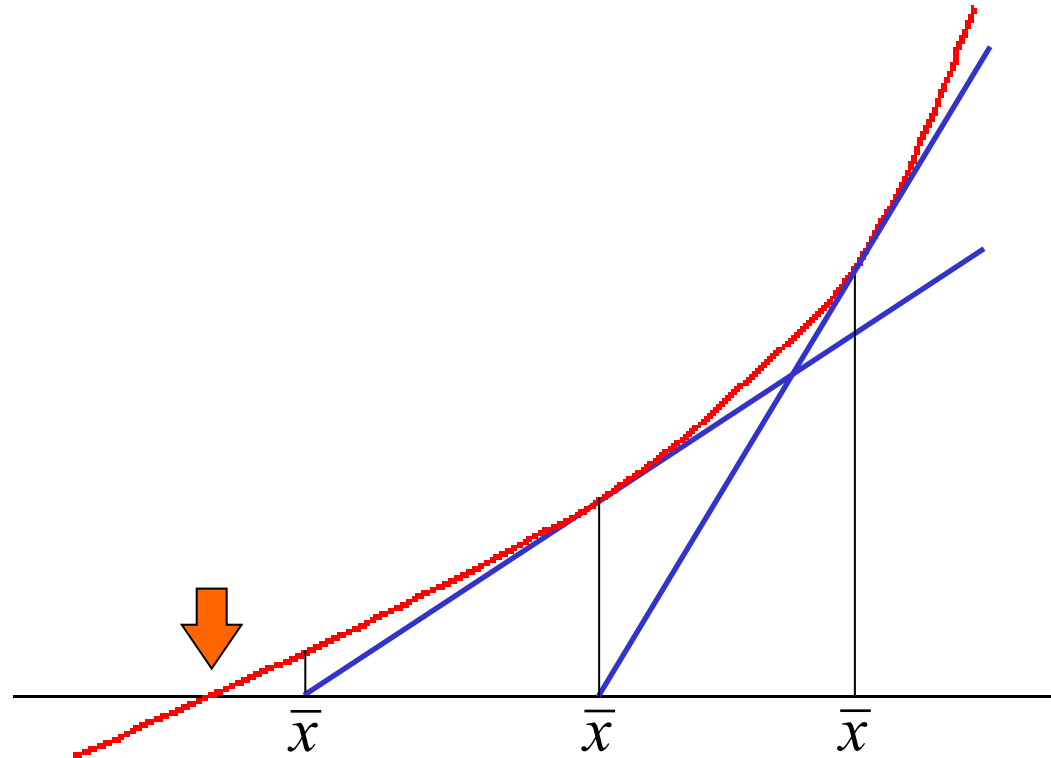
は、関数 $f(x)$ の \bar{x} における2次近似であり、もとの関数の極値の代わりに2次近似した関数の極値を計算することに相当する。



- 2次近似と極値計算を反復することで、もとの関数の極値を計算するのがニュートン法である。

ニュートン法による方程式の解の計算

- $f'(\bar{x}) \rightarrow f(\bar{x}), f''(\bar{x}) \rightarrow f'(\bar{x})$ と置き換えれば、ニュートン法により方程式の解を計算できる。



例題

- 関数 $f(x) = x^3 - 2x^2 + x + 3$ の $x = 2$ における2次近似を求めよ。

$$f_{\Pi}(x) = f(\bar{x}) + \underbrace{f'(\bar{x})}_{\downarrow} (x - \bar{x}) + \frac{1}{2} \underbrace{f''(\bar{x})}_{\downarrow} (x - \bar{x})^2$$

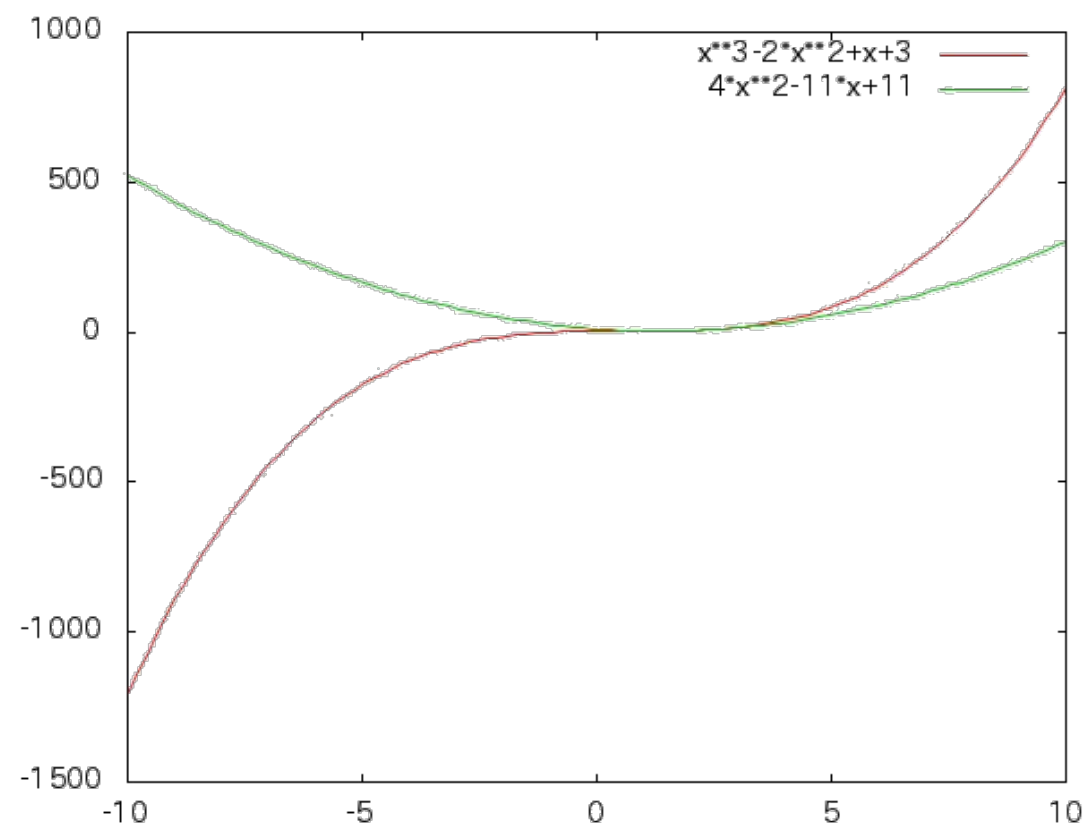
$$f'(x) = 3x^2 - 4x + 1 \quad f''(x) = 6x - 4$$

$$f(2) = 2^3 - 2 \cdot 2^2 + 2 + 3 = 5$$

$$f'(2) = 3 \cdot 2^2 - 4 \cdot 2 + 1 = 5$$

$$f''(2) = 6 \cdot 2 - 4 = 8$$

$$\begin{aligned} f_{\Pi}(x) &= 5 + 5(x - 2) + \frac{1}{2} \cdot 8(x - 2)^2 \\ &= 4x^2 - 11x + 11 \end{aligned}$$



多変数のニュートン法

- 点 $(\bar{x}_1, \dots, \bar{x}_n)$ の近傍の点 $(\bar{x}_1 + \Delta x_1, \dots, \bar{x}_n + \Delta x_n)$ で関数 $f(\bar{x}_1, \dots, \bar{x}_n)$ をテーラー展開すると次式を得る。

$$f(\bar{x}_1 + \Delta x_1, \dots, \bar{x}_n + \Delta x_n) = \bar{f} + \sum_{i=1}^n \frac{\partial \bar{f}}{\partial x_i} \Delta x_i + \frac{1}{2} \sum_{i,j=1}^n \frac{\partial^2 \bar{f}}{\partial x_i \partial x_j} \Delta x_i \Delta x_j + \dots$$

- この式の3次以降の項を省略した式を Δx_i で偏微分して0と置くと次式のようにになる。

$$\frac{\partial \bar{f}}{\partial x_i} + \sum_{j=1}^n \frac{\partial^2 \bar{f}}{\partial x_i \partial x_j} \Delta x_j = 0$$

- このことから、ニュートン法の反復式が次のように与えられる。

$$\begin{aligned} \bar{H} \Delta \mathbf{x} &= -\nabla \bar{f} \\ \Rightarrow \bar{H}(\mathbf{x} - \bar{\mathbf{x}}) &= -\nabla \bar{f} \\ \Rightarrow \mathbf{x} &= \bar{\mathbf{x}} - \bar{H}^{-1} \nabla \bar{f} \end{aligned} \quad \mathbf{H} = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \dots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \dots & \frac{\partial^2 f}{\partial x_n^2} \end{pmatrix}$$

ガウス・ニュートン法

- ニュートン法で使用するヘッセ行列を計算する際の2階微分を近似値に置き換えて計算を簡略化した方法

- ◆ $J = \frac{1}{2} \sum_{\alpha=1}^N \sum_{l=1}^r F_l(\mathbf{x}_\alpha, \mathbf{u})^2$ を \mathbf{u} の各成分 u_i で偏微分すると次のようになる。

$$\frac{\partial J}{\partial u_i} = \sum_{\alpha=1}^N \sum_{l=1}^r F_l(\mathbf{x}_\alpha, \mathbf{u}) \frac{\partial F_l(\mathbf{x}_\alpha, \mathbf{u})}{\partial u_i}$$

- ◆ さらに、 u_j で偏微分すると次のようになる。

$$\frac{\partial^2 J}{\partial u_i \partial u_j} = \sum_{\alpha=1}^N \sum_{l=1}^r \left(\frac{\partial F_l(\mathbf{x}_\alpha, \mathbf{u})}{\partial u_j} \frac{\partial F_l(\mathbf{x}_\alpha, \mathbf{u})}{\partial u_i} + F_l(\mathbf{x}_\alpha, \mathbf{u}) \frac{\partial^2 F_l(\mathbf{x}_\alpha, \mathbf{u})}{\partial u_i \partial u_j} \right)$$

- ◆ \mathbf{u} が解に近ければ $F_l(\mathbf{x}_\alpha, \mathbf{u}) \approx 0$ であるから、次のように近似できる。

$$\frac{\partial^2 J}{\partial u_i \partial u_j} \approx \sum_{\alpha=1}^N \sum_{l=1}^r \frac{\partial F_l(\mathbf{x}_\alpha, \mathbf{u})}{\partial u_j} \frac{\partial F_l(\mathbf{x}_\alpha, \mathbf{u})}{\partial u_i}$$

ガウス・ニュートン近似

- $\frac{\partial^2 J}{\partial u_i \partial u_j} \approx \sum_{\alpha=1}^N \sum_{l=1}^r \frac{\partial F_l(\mathbf{x}_\alpha, \mathbf{u})}{\partial u_j} \frac{\partial F_l(\mathbf{x}_\alpha, \mathbf{u})}{\partial u_i}$ をガウス・ニュートン近似と呼ぶ。
- $\nabla_{\mathbf{u}} J = \sum_{\alpha=1}^N \sum_{l=1}^r F_l(\mathbf{x}_\alpha, \mathbf{u}) \nabla_{\mathbf{u}} F_l(\mathbf{x}_\alpha, \mathbf{u})$ と $\mathbf{H}_{\mathbf{u}} = \sum_{\alpha=1}^N \sum_{l=1}^r (\nabla_{\mathbf{u}} F_l(\mathbf{x}_\alpha, \mathbf{u})) (\nabla_{\mathbf{u}} F_l(\mathbf{x}_\alpha, \mathbf{u}))^\top$ を用いたニュートン法をガウス・ニュートン法と呼ぶ。

$$\mathbf{u}^{(K+1)} = \mathbf{u}^{(K)} - \mathbf{H}_{\mathbf{u}}^{-1} \nabla_{\mathbf{u}} J$$

レーベンバーグ・マーカート法

- ニュートン法の解の更新速度を制御するパラメータを導入した方法

$$\boldsymbol{u}^{(K+1)} = \boldsymbol{u}^{(K)} - \boldsymbol{H}_{\boldsymbol{u}}^{-1} \nabla_{\boldsymbol{u}} J$$

+

$$\boldsymbol{u}^{(K+1)} = \boldsymbol{u}^{(K)} - \frac{1}{C} D[\boldsymbol{H}_{\boldsymbol{u}}^{(K)}]^{-1} \nabla_{\boldsymbol{u}} J^{(K)}$$

↓

$$\boldsymbol{u}^{(K+1)} = \boldsymbol{u}^{(K)} - \left(\boldsymbol{H}_{\boldsymbol{u}}^{(K)} + c D[\boldsymbol{H}_{\boldsymbol{u}}^{(K)}] \right)^{-1} \nabla_{\boldsymbol{u}} J^{(K)} \quad D[\cdot]: \text{対角成分の抽出}$$

レーベンバーグ・マーカート法のアルゴリズム

1. $c = \text{微小定数}(0.0001)$ とおく。
2. \mathbf{u} の初期値を与える
3. \mathbf{u} を用いて J の値を計算する
4. 勾配 $\nabla_{\mathbf{u}}J$ と（ガウス・ニュートン近似した）ヘッセ行列 $\mathbf{H}_{\mathbf{u}}$ を計算する
5. 次の連立1次方程式を解いて $\Delta\mathbf{u}$ を求める

$$(\mathbf{H}_{\mathbf{u}} + cD[\mathbf{H}_{\mathbf{u}}])^{-1}\Delta\mathbf{u} = -\nabla_{\mathbf{u}}J$$

6. 次のように \mathbf{u}' を計算する

$$\mathbf{u}' \leftarrow \mathbf{u} + \Delta\mathbf{u}$$

7. \mathbf{u}' を用いて J' を計算する
8. $J' > J$ なら $c \leftarrow 10c$ と更新してステップ5に戻る。そうでなければ、 $c \leftarrow \frac{c}{10}, J \leftarrow J', \mathbf{u} \leftarrow \mathbf{u}'$ とする
9. $\|\Delta\mathbf{u}\| < \delta$ であれば \mathbf{u} を返して終了する。そうでなければステップ4に戻る

特に覚えてほしいこと

- 勾配法
- ニュートン法
- ガウス・ニュートン法
- レーベンバーグ・マーカート法
- ガウス・ニュートン近似