最適化数学 3 章まとめ

3.1 勾配法

勾配法:制約条件がない場合に、関数の値を最大／最小化する方法 (最適化)の一つ

最大値を求める勾配法: 最大値に近いと思われる点x_0 を初期値とする

$f'(x_0)=0$ ならば終了(最大値があると仮定して)

そうでなければ 直線探索: 勾配∇f の方向の直線上で関数値が最大になる点を探索

$\mathbf{x}(t)=\mathbf{x}_0+t\nabla f_0$ と置いて、$F(t)=f(\mathbf{x}(t))$ を t で微分すると

$$\frac{dF}{dt}=\sum_{i=1}^{n}\frac{\partial f}{\partial x_i}\frac{dx_i}{dt}=\sum_{i=1}^{n}\frac{\partial f}{\partial x_i}\frac{\partial f_0}{\partial x_i}=(\nabla f,\nabla f_0) \qquad (3.1)$$

F が極値を取れば $(\nabla f,\nabla f_0)=0$

$\Rightarrow$ 直線探索の方向∇f_0と極値を取る点での勾配∇fが直交

定理 3.1 関数$f(x_1,...,x_n)$に対する勾配法の直線探索で定まる点では、その点を通る$f(x_1,...,x_n)$の等値面が探索直線に接する。したがって、次の直先探索の方向は、直前の探索方向と直交する

3.2 ニュートン法

ニュートン法: $f'(x)$と$f''(x)$が計算できる場合、勾配法よりも効率的な最適化手法

点$\bar{x}$における$f(x)$の **2** 次近似: $f_{II}(x)=f(\bar{x})+f'(\bar{x})(x-\bar{x})+\frac{1}{2}f''(\bar{x})(x-\bar{x})^2$

多変数の場合、点$(x_1,...,x_n)$の近くの点$(\bar{x}_1+\Delta x_1,...,\bar{x}_n+\Delta x_n)$での$f(x_1,...,x_n)$の値はテイラー展開により

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

3 次以上の項は無視し、Δx_iで偏微分し、0 とおくと

$$\frac{\partial\bar{f}}{\partial x_i}+\sum_{k=1}^{n}\frac{\partial^2\bar{f}}{\partial x_i\partial x_j}\Delta x_j=0 \qquad (3.18)$$

ここでヘッセ行列 $H=\begin{pmatrix}\frac{\partial^2 f}{\partial x_1^2} & \cdots & \frac{\partial^2 f}{\partial x_1\partial x_n}\\ \vdots & \ddots & \vdots\\ \frac{\partial^2 f}{\partial x_n\partial x_1} & \cdots & \frac{\partial^2 f}{\partial x_n^2}\end{pmatrix}$ (3.19) とし、$\bar{H}$を点$(\bar{x}_1,...,\bar{x}_n)$での値とすると

$$\bar{H}\Delta x=-\nabla\bar{f} \qquad (3.20)$$

が成り立つ。よって解$\mathbf{x}$のよりよい近似値は次で得られる:

$$x=\bar{x}-\bar{H}^{-1}\nabla\bar{f}$$

ニュートン法は、十分小さな値 δ に対し、$||\Delta\mathbf{x}||<\delta$ が成り立つまで上を繰り返す

長所:2次収束する

ニュートン法は、各ステップで関数を2次近似

短所: ヘッセ行列の計算には$\dfrac{\partial^2 f}{\partial x_i\partial x_j}$ の計算が必要、さらに悪いことにH^{-1}の計算も必要

このヘッセ行列の逆行列の計算を避けるための方法がいくつか提案されている

その一つが共役勾配法

3.3 共役勾配法

ニュートン法において K ステップ目の近似解を $(x^{(K)},y^{(K)})$とし、その点での fの値を$f^{(K)}$, ∇fの値を$\nabla f^{(K)}$, $\boldsymbol{H}$ の値を$H^{(K)}$と書くと、

$$f_{\mathrm{II}}(x) = f^{(K)} + (\nabla f^{(K)}, x - x^{(K)}) + \frac{1}{2}(x - x^{(K)}, H^{(K)}(x - x^{(K)})) \qquad (3.61)$$

この極値は次で定まる: $\nabla f_{II} = \nabla f^{(K)} + H^{(K)}(x - x^{(K)}) = 0$ (3.63)

ニュートン法では、これから　$\Delta x^{(K)} = -H^{(K)-1}\nabla f^{(K)}$ として求めた
共役勾配法では

$$m^{(K)} \propto \Delta x^{(K)} = x - x^{(K)}$$

というベクトル方向$m^{(K)}$ **(共役勾配)**　に進んでいくことを考える
(3.63)から　$H^{(K)} m^{(K)} \propto \nabla f^{(K)}$

等高線が円の場合、Hは単位行列Iの定数倍 $\Rightarrow m^{(K)} \propto \nabla f^{(K)}$
楕円の場合は方向が異なる　図 3.5 参照
$x^{(K)}$での等高線の接線ベクトルを$t^{(K)}$とすると、　$(\nabla f^{(K)}, t^{(K)}) = 0$
$\therefore (t^{(K)}, H^{(K)} m^{(K)}) = 0$ (3.65)
ここで　$m^{(K)} = \nabla f^{(K)} + \alpha^{(K)} t^{(K)}$ と書けるので (3,66)
（$m^{(K)}$は勾配$\nabla f^{(K)}$から接線方向$t^{(K)}$に少しずれている）
これを(3.65)に代入して

$$\alpha^{(K)} = -\frac{(t^{(K)}, H^{(K)}\nabla f^{(K)})}{(t^{(K)}, H^{(K)} t^{(K)})} \text{ となる } \quad (3.67)$$

これを用いて(3.66)で定まる共役勾配$m^{(K)}$の方向に直線探索すれば、
ヘッセ行列 H の逆行列を計算することが不要となる

実際には f_{II}ではなく関数fに対して直線探索を行う

点$x^{(K+1)}$で極値に達したとすれば、その点で探索直線は関数fの等高線に接する
すなわち、探索直線の方向が$x^{(K+1)}$での接線方向$t^{(K+1)}$になっている
この直線探索は共役勾配$m^{(K)}$の方向に行っているので、$t^{(K+1)} = m^{(K)}$
共役勾配法の反復公式:
初期値　$x^{(0)}$では$\alpha^{(0)} = 0$とし (つまり、勾配$\nabla f^{(0)}$ 方向に進む）

$$m^{(K)} = \nabla f^{(K)} + \alpha^{(K)} m^{(K-1)}$$

$$\alpha^{(K)} = -\frac{(m^{(K-1)}, H^{(K)}\nabla f^{(K)})}{(m^{(K-1)}, H^{(K)} m^{(K-1)})}$$

$$x^{(K+1)} = x^{(K)} + \beta^{(K)} m^{(K)}$$

関数 fが2次式ならば、　2回の反復で極値に到達

練習問題 3.5　関数 fが2次式のとき、上の式の $\beta^{(K)}$を求めよ（ヒント: 式(3.63)を用いる)

練習問題 3.6 共役勾配法を用いて $Ax = b$の解を求める方法を示せ。(ヒント:最小値を求める問題にする）