最急降下法

宮澤　彬

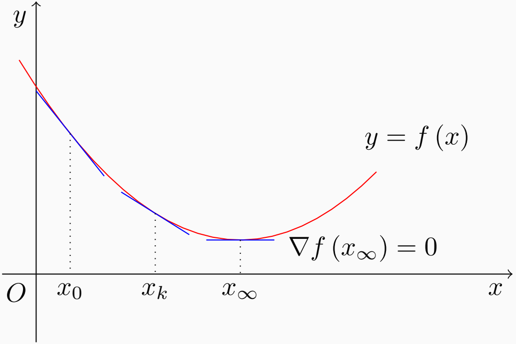
総合研究大学院大学 博士前期 1 年

miyazawa-a@nii.ac.jp

July 13, 2015

最急降下法

関数の停留点（特に極小点）を，解析的に求めること難しい場合にどのように求めればよいか．接線の傾きが負である点から，0 に近づく方向に移動していけばよさそうである．

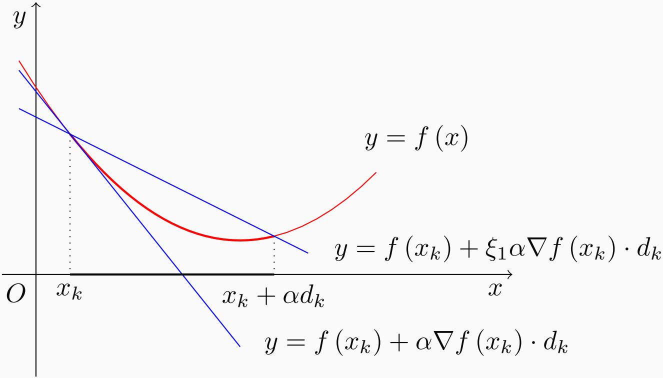


Armijo 条件

$0 < \xi_1 < 1$ であるような定数 ξ_1 に対して，

$$f\left(x_k + \alpha d_k\right) \leq f\left(x_k\right) + \xi_1 \alpha \nabla f\left(x_k\right) \cdot d_k$$

を満たす $\alpha > 0$ を選ぶ． この条件を **Armijo 条件** [1] という．



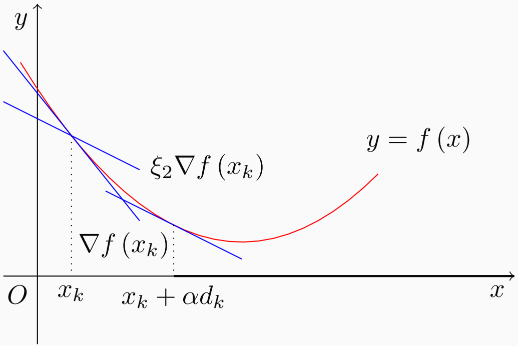
1 スペイン語読みをするならばおそらく/ar'mixo/.

Wolfe 条件

$0 < \xi_1 < \xi_2 < 1$ であるような ξ_1, ξ_2 に対して

$$\xi_2 \nabla f(x_k) \cdot d_k \leq \nabla f(x_k + \alpha d_k) \cdot d_k$$

を満たす $\alpha > 0$ を選ぶ．この条件を**曲率条件** (curvature condition) と呼ぶ．この条件と Armijo 条件を合わせて **Wolfe 条件**と呼ぶ．



Zoutendijk 条件

定理 目的関数 $f(x)$ は下に有界で，かつ，初期点 x_0 における準位集合 $\{x\,;\,f(x) \leq f(x_0)\}$ におけるを含む開集合 U において連続的微分可能であるとする．また勾配 $\nabla f(x)$ は U で Lipschitz 連続であるとする．すなわち，ある正定数 L が存在して，任意の $x,\,y \in U$ に対して

$$\|\nabla f(x) - \nabla f(y)\| \leq L\,\|x - y\|$$

が成り立つとする．

このとき $x_{k+1} = x_k + \alpha_k d_k$ を以下の条件を満たすようにとる．

- 各 α_k が Wolfe 条件を満たす．
- 各 d_k が降下方向である．すなわち $\nabla f(x_k) \cdot d_k < 0$ を満たす．

すると点列 $(x_k)_k$ について

$$\sum_{k=0}^{\infty} \left(\frac{\nabla f(x_k) \cdot d_k}{\|d_k\|} \right)^2 < \infty$$

が成り立つ．

Zoutendijk 条件

証明 曲率条件と $x_{k+1} = x_k + \alpha_k d_k$ から

$$\xi_2 \nabla f\left(x_k\right) \cdot d_k \leq \nabla f\left(x_{k+1}\right) \cdot d_k$$
$$\left(\xi_2 - 1\right) \nabla f\left(x_k\right) \cdot d_k \leq \left(\nabla f\left(x_{k+1}\right) - \nabla f\left(x_k\right)\right) \cdot d_k$$

が成り立つ．Lipschitz 条件より

$$\begin{aligned}
\left(\nabla f\left(x_{k+1}\right) - \nabla f\left(x_k\right)\right) \cdot d_k &\leq \left\|\nabla f\left(x_{k+1}\right) - \nabla f\left(x_k\right)\right\| \left\|d_k\right\| \\
&\leq L \left\|x_{k+1} - x_k\right\| \left\|d_k\right\| \\
&\leq \alpha_k L \left\|d_k\right\|^2
\end{aligned}$$

が成り立つ．これらから

$$\begin{aligned}
\alpha_k &\geq \frac{\left(\nabla f\left(x_{k+1}\right) - \nabla f\left(x_k\right)\right) \cdot d_k}{L \left\|d_k\right\|^2} \\
&\geq \frac{\xi_2 - 1}{L} \frac{\nabla f\left(x_k\right) \cdot d_k}{\left\|d_k\right\|^2}
\end{aligned}$$

を得る．

Zoutendijk 条件

得られた α_k を Armijo 条件に代入して

$$\begin{aligned} f\left(x_{k+1}\right) &\leq f\left(x_k\right) + \xi_1 \alpha_k \nabla f\left(x_k\right) \cdot d_k \\ &\leq f\left(x_k\right) - \frac{\xi_1\left(1-\xi_2\right)}{L} \frac{\left(\nabla f\left(x_k\right) \cdot d_k\right)^2}{\left\|d_k\right\|^2} \end{aligned}$$

となる．ここで $k=0$ から m までの和をとると

$$\begin{aligned} \sum_{k=0}^{m}\left(f\left(x_{k+1}\right) - f\left(x_k\right)\right) &\leq -\sum_{k=0}^{m} \frac{\xi_1\left(1-\xi_2\right)}{L} \frac{\left(\nabla f\left(x_k\right) \cdot d_k\right)^2}{\left\|d_k\right\|^2} \\ f\left(x_{m+1}\right) - f\left(x_0\right) &\leq -\frac{\xi_1\left(1-\xi_2\right)}{L} \sum_{k=0}^{m} \frac{\left(\nabla f\left(x_k\right) \cdot d_k\right)^2}{\left\|d_k\right\|^2} \end{aligned}$$

を得る．

Zoutendijk 条件

上式の右辺は m が増加するにつれて単調に減少する．また f は下に有界であると仮定していたので

$$\sum_{k=0}^{\infty} \frac{\left(\nabla f\left(x_k\right) \cdot d_k\right)^2}{\left\|d_k\right\|^2} < \infty \tag{Zoutendijk}$$

を得る． $\blacksquare$

上の (Zoutendijk) を **Zoutendijk 条件** [2] と呼ぶ．

2 オランダ語読みをするならばおそらく/'zautəndaɪk/.

Zoutendijk 条件

Zoutendijk 条件が成り立つとする．このとき
$S := \sum_{k=0}^{\infty} \left(\nabla f\left(x_k\right) \cdot d_k\right)^2 / \left\|d_k\right\|^2$ はある有限の値である．
Cauchy-Schwarz の不等式から，任意の自然数 m について

$$\left(\sum_{k=0}^{m} \frac{\left|\nabla f\left(x_k\right) \cdot d_k\right|}{\left\|d_k\right\|}\right)^2 \leq \sum_{k=0}^{m} \frac{\left(\nabla f\left(x_k\right) \cdot d_k\right)^2}{\left\|d_k\right\|^2} \leq S$$

が成り立つ．ゆえに

$$\sum_{k=0}^{\infty} \frac{\left|\nabla f\left(x_k\right) \cdot d_k\right|}{\left\|d_k\right\|} \leq \sqrt{S}$$

となり，この級数は収束することが分かる．したがって

$$\frac{\left|\nabla f\left(x_k\right) \cdot d_k\right|}{\left\|d_k\right\|} \to 0 \ \ (k \to \infty)$$

となる．

最急降下法の大域収束性

特に $d_k = -\nabla f\left(x_k\right)$ をとる．この d_k は $\nabla f\left(x_k\right) \cdot d_k = -\left\|f\left(x_k\right)\right\|^2 < 0$ を満たすので，降下方向である．さらに先に示した結果から，

$$\frac{\left|\nabla f\left(x_k\right) \cdot d_k\right|}{\left\|d_k\right\|} = \left\|f\left(x_k\right)\right\| \to 0 \ \ (k \to \infty)$$

を満たす．

Cauchy-Schwarz の不等式における等号成立条件から，$\left\|d_k\right\|$ を固定して考えたとき，この d_k は $\nabla f\left(x_k\right) \cdot d_k$ を最小にするものである．つまり最も急に減少させるものである．そのため $d_k = -\nabla f\left(x_k\right)$ とする方法を**最急降下法** (steepest descent method) と呼ぶ．

参考文献・おわりに

主に以下を参考にした.

- 矢部博, 新・工科系の数学「工学基礎　最適化とその応用」, 数理工学社, 2006.

また，このスライドのソースコードは https://github.com/pecorarista/documents にある.