

最急降下法

宮澤 彬

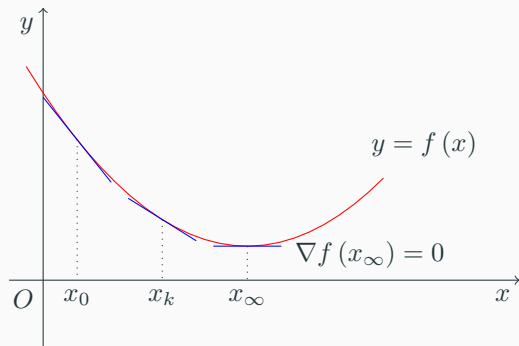
総合研究大学院大学 博士前期

miyazawa-a@nii.ac.jp

July 13, 2015
(modified: December 2, 2015)

最急降下法

関数の停留点（特に極小点）を，反復的な計算で求めるにはどうすればよいか．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．

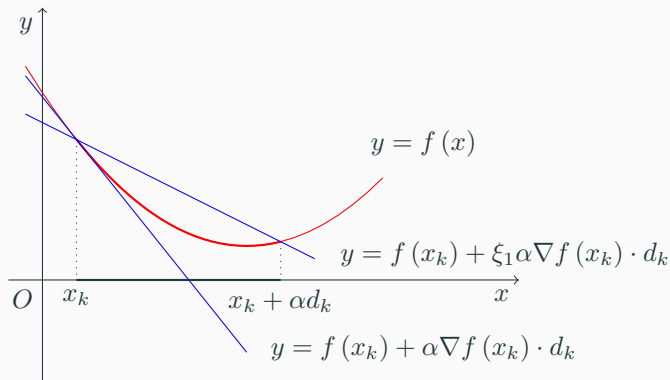


Armijo 条件

$0 < \xi_1 < 1$ であるような定数 ξ_1 に対して,

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

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



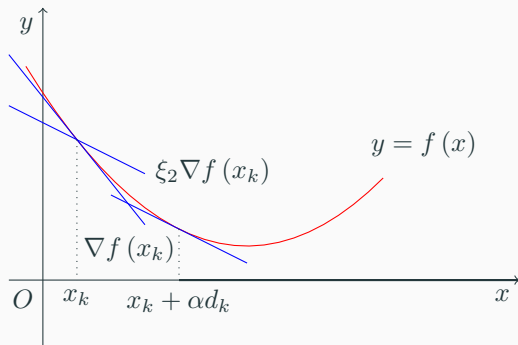
¹ スペイン語読みをするならばおそらく /ar'mixo/.

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

$$\begin{aligned}\xi_2 \nabla f(x_k) \cdot d_k &\leq \nabla f(x_{k+1}) \cdot d_k \\ (\xi_2 - 1) \nabla f(x_k) \cdot d_k &\leq (\nabla f(x_{k+1}) - \nabla f(x_k)) \cdot d_k\end{aligned}$$

が成り立つ. Lipschitz 条件より

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

が成り立つ. これらから

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

を得る.

得られた α_k を Armijo 条件に代入して

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

となる. ここで $k = 0$ から m までの和をとると

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

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

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

を得る。 ■

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 /'zautəndɪk/.

Zoutendijk 条件

Zoutendijk 条件が成り立つとする。このとき

$S := \sum_{k=0}^{\infty} (\nabla f(x_k) \cdot d_k)^2 / \|d_k\|^2$ 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

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

が成り立つ。ゆえに

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

となり、この級数は収束することが分かる。したがって

$$\frac{|\nabla f(x_k) \cdot d_k|}{\|d_k\|} \rightarrow 0 \quad (k \rightarrow \infty)$$

となる。

最急降下法の大域収束性

特に $d_k = -\nabla f(x_k)$ をとる. この d_k は $\nabla f(x_k) \cdot d_k = -\|\nabla f(x_k)\|^2 < 0$ 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,

$$\frac{|\nabla f(x_k) \cdot d_k|}{\|d_k\|} = \|\nabla f(x_k)\| \rightarrow 0 \quad (k \rightarrow \infty)$$

を満たす.

Cauchy-Schwarz の不等式における等号成立条件から, $\|d_k\|$ を固定して考えたとき, この d_k は $\nabla f(x_k) \cdot d_k$ 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 $d_k = -\nabla f(x_k)$ とする方法を**最急降下法** (steepest descent method) と呼ぶ.

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.