董云达的代表作系列一

假设 f 在欧氏空间中是二次连续可微的。考虑无约束极小化问题

$$\min \quad f(x).$$

相应的拟Newton法为

$$x^{k+1} = x^k - \alpha_k W_k \nabla f(x^k),$$

其中 W_k 由某个对称秩二校正公式自动生成，而 α_k 作为步长由下面 Wolfe 条件确定

$$f(x + \alpha d) \leq f(x) + c_1 \alpha \nabla f(x)^T d,$$
$$\nabla f(x + \alpha d)^T d \geq c_2 \nabla f(x)^T d,$$

其中 $0 < c_1 < c_2 < 1$.

在2007-2010年间，董云达提出了 Dong 条件

$$c_2 \nabla f(x)^T d \leq \nabla f(x + \alpha d)^T d \leq c_1 \nabla f(x)^T d,$$

其中 $0 < c_1 < c_2 < 1$.

国际小同行的引用与评议（原始材料）如下：“与 Wolfe 条件形成鲜明对比的是，Dong 条件仅仅依赖梯度的估值...”

10 *D. Wicht, M. Schneider and T. Böhlke / Computer Methods in Applied Mechanics and Engineering 358 (2020) 112611*

iteration reads

$$x^{k+1} = x^k - a^k \mathcal{H}^{-1}(x^k) \nabla f(x^k)$$

with the Hessian $\mathcal{H}$ of f and a damping factor $a^k \in (0, 1]$, cf. Chapter 9 in Boyd's book [60]. Instead of inverting $\mathcal{H}(x^k)$, the update can be performed by

$$x^{k+1} = x^k + a^k \Delta x$$

where Δx is an approximate solution of

$$\mathcal{H}(x^k) \Delta x = -\nabla f(x^k). \qquad (5.10)$$

The damping factor a^k is determined by a back-tracking procedure. In this paper, we use the stopping criteria of Dong [63]

$$c_2 \langle \nabla f(x^k), \Delta x \rangle_V \leq \langle \nabla f(x^k + a^k \Delta x), \Delta x \rangle_V \leq c_1 \langle \nabla f(x^k), \Delta x \rangle_V$$

with $0 < c_1 < c_2 < 1$. In contrast to the Wolfe conditions [64], Dong's criteria rely solely on gradient evaluations. This is beneficial, as evaluating f requires either the primal or dual condensed incremental potential, cf. Table 1, which is generally not available in FFT-based homogenization. Both w and w^* carry no physical meaning as they

注：Dong 条件下的拟Newton法 已成为优化领域的最基础性成果之一。

董云达的代表作系列二

Dong 引理 设 $\{\alpha_k\},\{\beta_k\},\{\gamma_k\}$ 是三个正序列。假设它们满足：

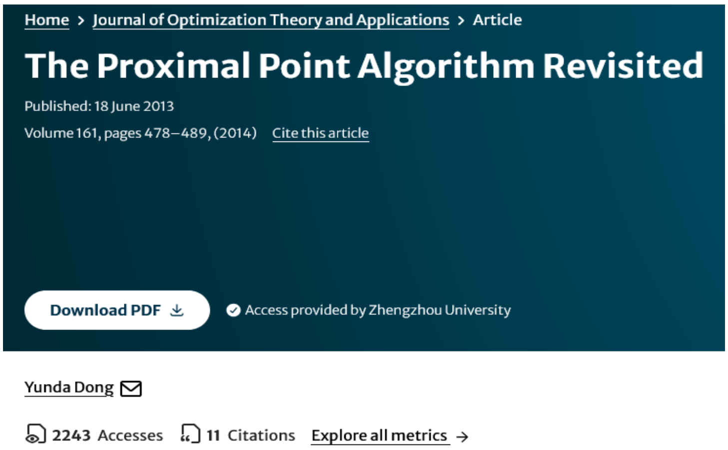
$$\alpha_{k+1}^2 \le \alpha_k^2 - \beta_k\gamma_k, \quad k = 0, 1, ..., \tag{1}$$

且序列 $\{\beta_k\}$ 是不可和的，序列 $\{\gamma_k\}$ 单调递减。那么，存在序列 $\{\varepsilon_k\}$ 使得下述关系成立：

$$\gamma_k \textstyle\sum_{i=0}^{k} \beta_i \le 2\alpha_0\varepsilon_k, \tag{2}$$

$$\alpha_k \le \varepsilon_k \le \alpha_0, \quad \lim_{k\to+\infty} \varepsilon_k = \lim_{k\to+\infty} \alpha_k. \tag{3}$$

利用Dong引理可以将邻近点算法的收敛率从$O(1/k)$改进为$o(1/k)$, 从而引起一时轰动。截止于2024年12月10日，已被下载四千多次!



注：邻近点算法的$O(1/k)$-收敛率是由Brezis和1994年Fields奖得主Lions给出的。

董云达的代表作系列三

在博士论文的第四章，董云达在Hilbert空间中考虑了下面的问题

$$0 \in F(x) + B(x),$$

其中F是连续单调的，B是极大单调的。在F是线性的情形下，提出了下面的方法

CHAPTER 4. NEW SPLITTING METHODS 29

Algorithm 4.2.1. *Let $F(x) = Mx + q$. Choose any starting point $x^0 \in \mathcal{H}$. For any given $k \geq 0$, computing x^{k+1} from x^k according to*

$$x^{k+1} = x^k - \gamma_k (I + M^T) r(x^k) \quad \forall k \geq 0. \tag{4.9}$$

where

$$r(x) = x - (I+B)^{-1}(x - Mx - q),$$

$$\gamma_k = \|(I + M^T) r(x^k)\|^{-2} \|r(x^k)\|^2. \tag{4.10}$$

国际小同行的引用与评议：该结果发表在Appl Math Letters (2005)，被引用在排名第一的数学项刊 Acta Numerica (2024)上

https://doi.org/10.1017/S0962492923000065 Published online by Cambridge University Press

THE GEOMETRY OF MONOTONE OPERATOR SPLITTING METHODS 595

sequence in $[\varepsilon, 2-\varepsilon]$, let $x_0 \in \mathcal{H}$, and let $y_0^* \in \mathcal{G}$. Iterate

$$
\begin{array}{l}
\text{for } n = 0, 1, \ldots \\
\left\lfloor
\begin{array}{l}
a_n = J_{\gamma_n A}(x_n - \gamma_n L^* y_n^*) \\
l_n = L x_n \\
b_n = J_{\sigma_n B}(l_n + \sigma_n y_n^*) \\
t_n = b_n - L a_n \\
t_n^* = \gamma_n^{-1}(x_n - a_n) + \sigma_n^{-1} L^*(l_n - b_n) \\
\tau_n = \|t_n\|^2 + \|t_n^*\|^2 \\
\text{if } \tau_n > 0 \\
\left\lfloor \theta_n = \lambda_n(\gamma_n^{-1}\|x_n - a_n\|^2 + \sigma_n^{-1}\|l_n - b_n\|^2)/\tau_n \right. \\
\text{else } \theta_n = 0 \\
x_{n+1} = x_n - \theta_n t_n^* \\
y_{n+1}^* = y_n^* - \theta_n t_n.
\end{array}
\right.
\end{array}
\tag{9.11}
$$

Then $(x_n)_{n\in\mathbb{N}}$ converges weakly to a point $x \in Z$ and $(y_n^*)_{n\in\mathbb{N}}$ converges weakly to a point $y^* \in Z^*$.

Remark 9.2. Here are notable instantiations of Proposition 9.1.

(i) The first instance of (9.11) in the literature seems to be that of Dong (2005), where $\mathcal{H}$ and $\mathcal{G}$ are Euclidean spaces, $A = 0$, and $(\forall n \in \mathbb{N})$ $\gamma_n = \sigma_n = 1$ and $\lambda_n = \lambda \in]0, 2[$. Convergence of the primal sequence $(x_n)_{n\in\mathbb{N}}$ was established by different means.