

董云达的代表作系列一

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

$$\alpha_{k+1}^2 \leq \alpha_k^2 - \beta_k \gamma_k, \quad k = 0, 1, \dots,$$

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：

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

$$\alpha_k \leq \varepsilon_k \leq \alpha_0, \quad \lim_{k \rightarrow +\infty} \varepsilon_k = \lim_{k \rightarrow +\infty} \alpha_k.$$

Brezis 与菲尔兹奖获得者 Lions 证明邻近点方法收敛率与 $1/k$ 的同阶。董氏引理进一步将其提升至高阶无穷小。截至2024年末，该成果已被下载超**4000次**！



值得指出的是，利用董氏引理可以证明 Abel-Dini 定理。

董云达的代表作系列二

假设 f 在欧氏空间中是二次连续可微的凸函数。考虑无约束极小化问题 $\min f(x)$ 。相应的拟 Newton 法为

$$x^+ = x + \alpha d, \quad d := -W \nabla f(x),$$

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

$$\begin{aligned} f(x + \alpha d) &\leq f(x) + c_1 \alpha \nabla f(x)^T d, \\ \nabla f(x + \alpha d)^T d &\geq c_2 \nabla f(x)^T d, \end{aligned}$$

其中 $0 < c_1 < c_2 < 1$ 。

在 2007-2010 年间，董云达提出了董氏条件

$$c_2 \nabla f(x)^T d \leq \nabla f(x + \alpha d)^T d \leq c_1 \nabla f(x)^T d, \quad 0 < c_1 < c_2 < 1.$$

对于董氏条件，2020年德国卡尔斯鲁厄理工学院（德国版MIT）的学者在计算力学顶刊 Comput Method Appl M 上予以好评：“与 Wolfe 条件形成鲜明对比的是，董氏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

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

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

董云达的代表作系列三

在 2003 年博士论文的第四章，董云达在 Hilbert 空间中考虑了下面的问题

$$0 \in Mx + q + B(x),$$

其中 M 是有界、线性和单调的， q 是一个向量， B 是极大单调的，并且提出了下面的方法

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

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

where

$$\begin{aligned} r(x) &= x - (I + B)^{-1}(x - Mx - q), \\ \gamma_k &= \|(I + M^T)r(x^k)\|^{-2}\|r(x^k)\|^2. \end{aligned} \quad (4.10)$$

更多细节参考 <https://ydong2024.github.io/downloads/dongthesis.pdf>

该结果发表在 Appl Math Letters (2005)，2024年获排名第一的数学顶刊 Acta Numerica 的好评。

[//doi.org/10.1017/S0962492923000065](https://doi.org/10.1017/S0962492923000065) Published online by Cambridge University Press

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

$$\begin{aligned} &\text{for } n = 0, 1, \dots \\ &\quad \begin{cases} a_n = J_{\gamma_n A}(x_n - \gamma_n L^* y_n^*) \\ l_n = Lx_n \\ b_n = J_{\sigma_n B}(l_n + \sigma_n y_n^*) \\ t_n = b_n - La_n \\ t_n^* = \gamma_n^{-1}(x_n - a_n) + \sigma_n^{-1}L^*(l_n - b_n) \\ \tau_n = \|t_n\|^2 + \|t_n^*\|^2 \\ \text{if } \tau_n > 0 \\ \quad \theta_n = \lambda_n(\gamma_n^{-1}\|x_n - a_n\|^2 + \sigma_n^{-1}\|l_n - b_n\|^2)/\tau_n \\ \text{else } \theta_n = 0 \\ x_{n+1} = x_n - \theta_n t_n^* \\ y_{n+1}^* = y_n^* - \theta_n t_n. \end{cases} \end{aligned} \quad (9.11)$$

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.

董云达的代表作系列四

在 2003 年博士论文的第四章，董云达在 Hilbert 空间中考虑了下面的问题

$$0 \in F(x) + B(x), \quad (1)$$

其中 F 是连续单调的， B 是极大单调的。利用图(a)中的记号 $J_\lambda(x)$ ，提出了下面的方法

Proof. Since $J_\lambda(x) := (I + \lambda B)^{-1}(I + \lambda F)(x)$, we have

$$J_\lambda(x) + \lambda B(J_\lambda(x)) \ni x - \lambda F(x).$$

That is,

$$\lambda^{-1}(x - J_\lambda(x)) - F(x) \in B(J_\lambda(x)). \quad (4.3)$$

On the other hand, since x is a zero of $F \ni B + F$ then we have

$$-F(x) \in B(x) \quad (4.4)$$

27

Algorithm 4.2.4. Choose any starting point $x^0 \in \mathcal{H}$ and $\lambda_{-1} \in (0, +\infty)$. Also choose $\rho \in (0, 1)$ and $\beta \in (0, 1)$. For any given $k \geq 0$, computing (x^{k+1}, λ_k) from (x^k, λ_{k-1}) where λ_k is the largest $\lambda \in \{\lambda_{k-1}, \lambda_{k-1}\beta, \lambda_{k-1}\beta^2, \dots\}$ satisfying

$$\lambda \langle x^k - J_\lambda(x^k), F(x^k) - F(J_\lambda(x^k)) \rangle \leq (1 - \rho) \|x^k - J_\lambda(x^k)\|^2,$$

and let

$$x^{k+1} = x^k - \gamma_k(x^k - J_{\lambda_k}(x^k) - \lambda_k F(x^k) + \lambda_k F(J_{\lambda_k}(x^k))) \quad \forall k \geq 0,$$

(a)
(b)

更多细节参考 <https://ydong2024.github.io/downloads/dongthesis.pdf>

将 (1) 中的 B 换成 A ，则变为

$$0 \in F(x) + A(x). \quad (2)$$

同时，利用董云达论文 [Math Comput Simulat, 223, 86-107, 2024] 中的记号，将 $J_\lambda(x)$ 换成 $x(\alpha)$ 以及取 $\beta = 0.5$ ，我们将该方法重新叙述/摘录如下：

Choose $x^0 \in \text{dom}F \cap \text{dom}A$. Choose $\rho \in (0, 1)$ and $\alpha_{-1} > 0$. At k -th iteration, find the smallest j_k in $\alpha = \alpha_{k-1} 0.5^{j_k}$, $j = 0, 1, \dots$, such that $x^k(\alpha) = (I + \alpha A)^{-1}(x^k - \alpha F(x^k))$ satisfies

$$\alpha \langle x^k - x^k(\alpha), F(x^k) - F(x^k(\alpha)) \rangle \leq (1 - \rho) \|x^k - x^k(\alpha)\|^2.$$

Take $\alpha_k = \alpha_{k-1} 0.5^{j_k}$, and compute $\bar{x}^k = x^k(\alpha_k)$. Compute in order

$$d^k = x^k - \bar{x}^k - \alpha_k(F(x^k) - F(\bar{x}^k)), \quad \gamma_k = \langle x^k - \bar{x}^k, d^k \rangle / \|d^k\|^2.$$

Then choose $\theta_k \in (0, 2]$ and compute

$$x^{k+1} = x^k - \theta_k \gamma_k d^k.$$

由于不需要估计 Lipschitz 常数以及克服了 Tseng 方法的小步长现象，从而成为求解上述问题最实用的基本方法。2014年董云达与学生合作，证实了它在数值方面的明显优势。2024年这篇论文，还讨论了 F 仅为一致连续的情形，并给出方法弱收敛性的一个严格证明。在此情形下，所给出的 α 的选取方式具有原创性。

至于凸极小化问题 $\min f(x) + g(x)$ ，我们考虑其最优性条件

$$0 \in \nabla f(x) + \partial g(x),$$

并且在 (2) 中取 $F := \nabla f$ 以及 $A := \partial g$ 即可。

董云达的代表作系列五

董云达在论文 [Math Comput Simulat, 223, 86-107, 2024] 中考虑了下面的问题

$$0 \in F(x) + A(x) + Q^*B(Qx - q),$$

其中 F 是连续单调的, A, B 是极大单调的, Q 是线性的而 Q^* 是它的伴随, q 是一个向量, 并且提出了一个新的实用算法

In the case of F_i being Lipschitz continuous, we give

Algorithm 2.1. Our proposed splitting algorithm in Lipschitz continuity case

Step 0. For $i = 1, \dots, n$, choose $x_i^0 \in \mathcal{H}_i$, $u^0 \in \mathcal{G}$. Choose $\alpha_{i,-1} > 0$, $\rho \in (0, 1)$. Set $k := 0$.

Step 1. For $i = 1, \dots, l$, choose appropriate $\alpha_{i,k} > 0$. For $i = l + 1, \dots, n$, find the smallest j_k in (19) such that

$$x_i^k(\alpha_i) = (I + \alpha_i A_i)^{-1}(x_i^k - \alpha_i(F_i(x_i^k) + Q_i^* u^k))$$

satisfies

$$\alpha_i \langle x_i^k - x_i^k(\alpha_i), F_i(x_i^k) - F_i(x_i^k(\alpha_i)) \rangle \leq (1 - \rho) \|x_i^k - x_i^k(\alpha_i)\|^2. \quad (20)$$

Take $\alpha_{i,k} = \alpha_{i,k-1} t^{j_k}$, and compute

$$\bar{x}_i^k = x_i^k(\alpha_{i,k}). \quad (21)$$

Choose $\beta_k > 0$ via (29) below and find \bar{u}^k such that

$$(\beta_k I + B^{-1})(\bar{u}^k) \ni \beta_k u^k + \sum_{i=1}^n Q_i \bar{x}_i^k - q. \quad (22)$$

If $\bar{x}_i^k = x_i^k$, $i = 1, \dots, n$, and $\bar{u}^k = u^k$, then stop. Otherwise go to Step 2.

89

Y. Dong

Mathematics and Computers in Simulation 223 (2024) 86–107

Step 2. For $i = 1, \dots, n$, compute

$$d_{x_i}^k = \alpha_{i,k}^{-1}(x_i^k - \bar{x}_i^k) - (F_i(x_i^k) - F_i(\bar{x}_i^k)) - Q_i^*(u^k - \bar{u}^k), \quad (23)$$

$$d_u^k = \beta_k(u^k - \bar{u}^k), \quad (24)$$

$$\gamma_k = \frac{\sum_{i=1}^n \langle x_i^k - \bar{x}_i^k, d_{x_i}^k \rangle + \langle u^k - \bar{u}^k, d_u^k \rangle}{\sum_{i=1}^n \|d_{x_i}^k\|^2 + \|d_u^k\|^2}. \quad (25)$$

Then choose $\theta_k \in (0, 2]$ and the new iterates are given by

$$x_i^{k+1} = x_i^k - \theta_k \gamma_k d_{x_i}^k, \quad i = 1, \dots, n, \quad (26)$$

$$u^{k+1} = u^k - \theta_k \gamma_k d_u^k. \quad (27)$$

Set $k := k + 1$, and go to Step 1.

它是“董云达的代表作系列四”中方法的一个拓展, 尤其适用于深度学习中的特征选取、高光谱解混等前沿问题。更多细节参考

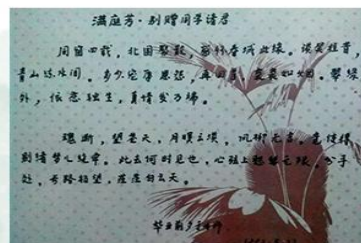
<https://ydong2024.github.io/downloads/journal2024.pdf>

满庭芳·别赠同学诸君

董云达

同窗四载，北国聚散，感怀春城此缘。谈笑往昔，青山流水间。多少宠辱恩怨，再回首，袅袅如烟。翠绿外，依恋独生，真情发万端。

魂断，望苍天，月暝云淡，风柳无言。竟使得别绪梦儿绕牵。此去何时见也，心弦上愁丝无限。分手处，长路极望，茫茫白云天。



毕业前夕于长师

1992.5.22

董云达的这首词，以其深情的笔触和细腻的描写，成功地勾勒出了八十年代大学生在毕业之际与同窗分别的复杂情感。

词的上半部分，作者以“同窗四载，北国聚散”开篇，简洁明了地交代了背景，同时也为全词奠定了感伤的基调。通过对青山流水、谈笑风生的往昔回忆，以及对宠辱恩怨的淡然态度，展现了对大学生活的深深眷恋和对同窗情谊的珍视。这种情感在“再回首，袅袅如烟”的描绘中得到了进一步的升华，使得读者能够深切地感受到作者内心的波澜与不舍。

词的下半部分，作者笔锋一转，以“魂断，望苍天”等词句，直白地抒发了离别之际的哀愁与无奈。同时，巧妙地借助月暝云淡、风柳无言等自然景物，将内心的情感与其融为一体，营造出一种凄美而深远的意境，使得离别的愁绪更加浓厚，令人动容。

总之，这首词情感真挚，语言优美，画面感强。不仅具有高度的艺术性，更富有鲜明的时代特色。