H-求解线性约束凸优化问题的预测-校正型广义邻近点算法

古稀之后的新收获–II   maths.nju.edu.cn/~hebma   何炳生

1 线性约束的凸优化问题及其变分不等式的 PPA 算法

我们把线性约束凸优化问题的 Lagrange 函数的鞍点归结为求变分不等式

$$w^* \in \Omega, \quad \theta(u) - \theta(u^*) + (w - w^*)^T F(w^*) \ge 0, \quad \forall\, w \in \Omega \tag{1.1}$$

的解. 为求解这个变分不等式, 我们有如下的邻近点算法 (PPA)

邻近点算法(PPA) 第 k-步迭代从给定的核心变量 v^k 开始, 求得新的迭代点 w^{k+1}, 使得

$$w^{k+1} \in \Omega, \ \ \theta(u) - \theta(u^{k+1}) + (w - w^{k+1})^T F(w^{k+1}) \ge (v - v^{k+1})^T H(v^k - v^{k+1}), \ \forall\, w \in \Omega, \tag{1.2}$$

成立. 其中 H 是对称正定矩阵. 交替方向法中 $u=(x,y), w=(x,y,\lambda), v=(y,\lambda)$. 有的情形中 $v=w$.

将 (1.2) 中任意的 $w \in \Omega$ 设为 w^*, 就马上得到 $(v^{k+1} - v^*)^T H (v^k - v^{k+1}) \ge 0$.

设 $a=(v^k - v^*), b=(v^{k+1}-v^*)$, 利用 $\boxed{b^T H(a-b) \ge 0 \ \Rightarrow\ \|b\|_H^2 \le \|a\|_H^2 - \|a-b\|_H^2}$ 有

$$\text{(PPA 收敛性质)} \qquad \|v^{k+1} - v^*\|_H^2 \le \|v^k - v^*\|_H^2 - \|v^k - v^{k+1}\|_H^2, \quad \forall v^* \in \mathcal{V}^*. \tag{1.3}$$

2 预测-校正算法的统一框架及其收敛性

除了 PPA, 我们还有求解变分不等式 (1.1) 的预测-校正的算法统一框架.

[预测] 第 k 步迭代从给定的核心变量 v^k 开始, 求得预测点 $\tilde{w}^k$, 使得

$$\tilde{w}^k \in \Omega, \ \ \theta(u) - \theta(\tilde{u}^k) + (w - \tilde{w}^k)^T F(\tilde{w}^k) \ge (v - \tilde{v}^k)^T Q (v^k - \tilde{v}^k), \quad \forall\, w \in \Omega \tag{2.1}$$

成立. 其中矩阵 $Q^T + Q$ 正定.

[校正] 根据预测得到的 $\tilde{v}^k$, 给出核心变量 v 的新迭代点 v^{k+1} 的公式为

$$v^{k+1} = v^k - M(v^k - \tilde{v}^k). \tag{2.2}$$

我们称 (2.1) 中的 Q 为预测矩阵, (2.2) 中的 M 为校正矩阵.

收敛性条件 对给定的预测矩阵 Q, 要求设计的校正矩阵 M 满足如下条件:

$$\text{存在正定矩阵}\ \ H \succ 0 \ \ \text{使得}\ \ HM = Q, \ \ \text{并且}\ \ G := Q^T + Q - M^T H M \succ 0. \tag{2.3}$$

在《从好不容易凑出一个算法到并不费劲构造一簇算法》的短篇的定理 G 中, 我们证明了收缩性质

$$\text{(算法收敛性质)} \qquad \|v^{k+1} - v^*\|_H^2 \le \|v^k - v^*\|_H^2 - \|v^k - \tilde{v}^k\|_G^2, \quad \forall v^* \in \mathcal{V}^*. \tag{2.4}$$

预测-校正方法 (2.1)-(2.2) 和邻近点算法 (1.2) 之间的关系: 若 (2.1) 中的矩阵 Q 对称, 就记其为 H, 并记其中的 $\tilde{w}^k$ 为 w^{k+1}, 便得到 (1.2). 这也相当于在 (2.2) 中取 $M = I$, (2.3) 中的 G 就等于 H. 由于 $\tilde{v}^k = v^{k+1}$, $G = H$, 预测-校正算法的收敛性质 (2.4) 就成了邻近点算法的收敛性质 (1.3).

3 统一框架中方法收敛性另一重要性质的证明

满足收敛条件 (2.3) 的统一框架算法不仅具有收敛性质 (2.4), 序列 $\{\|v^k - v^{k+1}\|_H\}$ 还是单调不增的.

定理H 求解变分不等式 (1.1), 设 $\{v^k\}, \{\tilde{w}^k\}$ 是由预测-校正框架 (2.1)-(2.2) 生成的. 如果收敛性条件 (2.3) 满足, 则有

$$\|v^{k+1} - v^{k+2}\|_H^2 \le \|v^k - v^{k+1}\|_H^2 - \|(v^k - \tilde{v}^k) - (v^{k+1} - \tilde{v}^{k+1})\|_G^2. \tag{3.1}$$

更简单地说, 有

$$\|v^{k+1} - v^{k+2}\|_H^2 \le \|v^k - v^{k+1}\|_H^2. \tag{3.2}$$

证明 根据 (2.1), 我们有

$$\theta(u)-\theta(\tilde{u}^k)+(w-\tilde{w}^k)^T F(\tilde{w}^k)\ge (v-\tilde{v}^k)^T Q(v^k-\tilde{v}^k),\ \forall w\in\Omega \tag{3.3}$$

和

$$\theta(u)-\theta(\tilde{u}^{k+1})+(w-\tilde{w}^{k+1})^T F(\tilde{w}^{k+1})\ge (v-\tilde{v}^{k+1})^T Q(v^{k+1}-\tilde{v}^{k+1}),\ \forall w\in\Omega. \tag{3.4}$$

将不等式 (3.3) 和 (3.4) 中任意的 $w\in\Omega$ 分别设为 $\tilde{w}^{k+1}$ 和 $\tilde{w}^k$, 得到

$$\theta(\tilde{u}^{k+1})-\theta(\tilde{u}^k)+(\tilde{w}^{k+1}-\tilde{w}^k)^T F(\tilde{w}^k)\ge(\tilde{v}^{k+1}-\tilde{v}^k)^T Q(v^k-\tilde{v}^k)$$

和

$$\theta(\tilde{u}^k)-\theta(\tilde{u}^{k+1})+(\tilde{w}^k-\tilde{w}^{k+1})^T F(\tilde{w}^{k+1})\ge(\tilde{v}^k-\tilde{v}^{k+1})^T Q(v^{k+1}-\tilde{v}^{k+1}).$$

将上述两个不等式相加并利用 $(\tilde{w}^k-\tilde{w}^{k+1})^T\big(F(\tilde{w}^k)-F(\tilde{w}^{k+1})\big)=0$, 就有

$$(\tilde{v}^k-\tilde{v}^{k+1})^T Q\{(v^k-\tilde{v}^k)-(v^{k+1}-\tilde{v}^{k+1})\}\ge 0. \tag{3.5}$$

在不等式 (3.5) 两边加上 $\{(v^k-\tilde{v}^k)-(v^{k+1}-\tilde{v}^{k+1})\}^T Q\{(v^k-\tilde{v}^k)-(v^{k+1}-\tilde{v}^{k+1})\}$, 得到

$$(v^k-v^{k+1})^T Q\{(v^k-\tilde{v}^k)-(v^{k+1}-\tilde{v}^{k+1})\}\ge\tfrac{1}{2}\|(v^k-\tilde{v}^k)-(v^{k+1}-\tilde{v}^{k+1})\|^2_{(Q^T+Q)}.$$

利用 $HM=Q$ 和 $M(v^k-\tilde{v}^k)=(v^k-v^{k+1})$ (见(2.2)), 就有

$$(v^k-v^{k+1})^T H\{(v^k-v^{k+1})-(v^{k+1}-v^{k+2})\}\ge\tfrac{1}{2}\|(v^k-\tilde{v}^k)-(v^{k+1}-\tilde{v}^{k+1})\|^2_{(Q^T+Q)}. \tag{3.6}$$

最后, 利用恒等式 $\|a\|_H^2-\|b\|_H^2=2a^T H(a-b)-\|a-b\|_H^2$ 和 (3.6), 我们得到

$$\begin{aligned}&\|v^k-v^{k+1}\|_H^2-\|v^{k+1}-v^{k+2}\|_H^2\\ &\quad= 2(v^k-v^{k+1})^T H\{(v^k-v^{k+1})-(v^{k+1}-v^{k+2})\}-\|(v^k-v^{k+1})-(v^{k+1}-v^{k+2})\|_H^2\\ &\quad\ge \|(v^k-\tilde{v}^k)-(v^{k+1}-\tilde{v}^{k+1})\|^2_{(Q^T+Q)}-\|(v^k-v^{k+1})-(v^{k+1}-v^{k+2})\|_H^2.\end{aligned} \tag{3.7}$$

对上式右端最后一项利用 $(v^k-v^{k+1})=M(v^k-\tilde{v}^k)$ (见(2.2)), 有

$$\|(v^k-v^{k+1})-(v^{k+1}-v^{k+2})\|_H^2=\|M(v^k-\tilde{v}^k)-M(v^{k+1}-\tilde{v}^{k+1})\|_H^2, \tag{3.8}$$

由 (3.7), (3.8), 并利用 $Q^T+Q-M^T HM=G$, 就得到

$$\|v^k-v^{k+1}\|_H^2-\|v^{k+1}-v^{k+2}\|_H^2 \quad\ge\quad \|(v^k-\tilde{v}^k)-(v^{k+1}-\tilde{v}^{k+1})\|_G^2.$$

这就是 (3.1), 证明完毕. □

定理 H 是对预测-校正的统一框架算法证明的. PPA 是统一框架算法的特例, 定理 H 的结论同样成立.

4 根据收敛条件的等价性得到的预测-校正的广义邻近点算法

满足收敛条件 (2.3) 的统一框架算法具有漂亮性质 (2.4) 和 (3.2). 如果统一框架中的某个算法能够使得 (2.4) 右端的 $\|v^k-\tilde{v}^k\|_G^2$ 等于 $\|v^k-v^{k+1}\|_H^2$, 那就得到更漂亮的 PPA 的性质 (1.3). 这是可以办到的!

当预测 (2.1) 中的预测矩阵 Q 满足 $Q^T+Q\succ 0$, 我们总可以取

$$D\succ 0,\quad G\succ 0 \qquad \text{使得} \qquad D+G=Q^T+Q. \tag{4.1}$$

然后令 $M^T HM=D$, 由矩阵方程组解得

$$\left\{\begin{array}{rcl} HM &=& Q,\\ M^T HM &=& D.\end{array}\right. \iff \left\{\begin{array}{rcl} HM &=& Q,\\ Q^T M &=& D.\end{array}\right. \iff \left\{\begin{array}{rcl} H &=& QD^{-1}Q^T,\\ M &=& Q^{-T}D.\end{array}\right. \tag{4.2}$$

得到满足收敛条件的 $M=Q^{-T}D$ 和 $H=QD^{-1}Q^T$. 然后去实行校正 (2.2), 收敛性质 (2.4) 仍然成立.

预测-校正的广义 PPA. 在 Q 非对称的预测-校正方法中, 如果在 (4.1) 中取一对特殊的 D 和 G, 使得

$$D=G=\tfrac{1}{2}(Q^T+Q) \tag{4.3}$$

由于 $D=G$, 收缩不等式 (2.4) 就成为

$$\|v^{k+1}-v^*\|_H^2\le\|v^k-v^*\|_H^2-\|v^k-\tilde{v}^k\|_D^2,\quad \forall v^*\in\mathcal{V}^*. \tag{4.4}$$

因为 $D=M^T HM$ (见 (4.2)), 利用 $M(v^k-\tilde{v}^k)=v^k-v^{k+1}$ (见 (2.2)), 不等式 (4.4) 就成了

$$\|v^{k+1}-v^*\|_H^2\le\|v^k-v^*\|_H^2-\|v^k-v^{k+1}\|_H^2,\quad \forall v^*\in\mathcal{V}^*. \tag{4.5}$$

这是跟 (1.3) 形式相同的收缩不等式, 因此我们称相应的算法为预测-校正的广义 PPA.

广义 PPA 的校正可以由 $Q^T(v^{k+1}-v^k)=\frac{1}{2}(Q^T+Q)(\tilde{v}^k-v^k)$ 完成. 序列 $\{v^k\}$ 具备性质 (3.2) 和 (4.5)!