doi:10.3969/j.issn.1001-4616.2017.01.002

一类半正定变分不等式的随机下降算法

徐海文[1],孙黎明[2]

(1.中国民用航空飞行学院计算机学院,四川 广汉 618307)
(2.南京审计大学理学院,江苏 南京 211815)

[摘要] 校正投影收缩算法的下降量证明中多次使用了放大不等式,因此本文利用满足固定均值的随机数适当扩张步长,得到了一类半正定变分不等式问题的随机下降算法. 在适当的假设条件下,利用马尔可夫不等式和依概率收敛的性质,给出了随机下降算法的依概率收敛性证明. 通过一系列的数值试验验证了随机下降算法的有效性,并且表明了合理选择随机数的均值和方差可以提高随机下降算法的计算效率.

[关键词] 半正定变分不等式问题,校正投影收缩算法,随机下降算法,依概率收敛
[中图分类号]O221 [文献标志码]A [文章编号]1001-4616(2017)01-0006-07

The Stochastic Descent Algorithm for a Kind of Semidefinite Variational Inequality Problem

Xu Haiwen[1],Sun Liming[2]

(1.College of Computer Science and Technology,Civil Aviation Flight University of China,Guanghan 618307,China)
(2.College of Science,Nanjing Audit University,Nanjing 211815,China)

Abstract:The amplification inequality is used for many times in the proof of drop function of correction projection and contraction algorithm,so we propose the stochastic descent algorithm for a class of semidifinite variational inequality problem through the random steplength extension with the random number series satisfying the Gaussian distribution or Uniform distribution and these random number series have a fixed mean. Subsequently,the probability convergence of stochastic descent algorithm is provided by the properties of Markov's inequality and probability convergence under some suitable conditions. Finally,some numerical experiments show the effectiveness and efficiency of the stochastic descent algorithm,and reasonable selecting mean and variance of random number can improve the efficiency of the algorithm.

Key words:semidefinite variational inequality problem,correction projection and contraction algorithm,stochastic descent algorithm,probability convergence

20 世纪 90 年代,Alizadeh,Kojima,Shinsoh 和 Hara 等人将系统控制论和组合优化的一些问题模型总结为半正定互补问题[1-2],即 SDCP(F,S_+^n)[3]:找到 $\boldsymbol{X}\in S_+^n$满足

$$\boldsymbol{X}\in S_+^n,F(\boldsymbol{X})\in S_+^n,\boldsymbol{X}\bullet F(\boldsymbol{X})=0. \tag{1}$$

这里假定 $S_+^n=\{\boldsymbol{X}\in S^n \mid \boldsymbol{X}\text{ 是半正定阵}\}$,$S^n=\{\boldsymbol{X}\in\mathbf{R}^{n\times n},\boldsymbol{X}^{\mathrm{T}}=\boldsymbol{X}\}$ 是一个实对称矩阵集,F 是 S^n 到 S^n 的映射,Frobenius 范数和 $\bullet$ 分别是 S^n 中的范数和内积,即 $\|\boldsymbol{X}\|=(\boldsymbol{X}\bullet\boldsymbol{X})^{1/2},=\Big(\sum\limits_{i=1,j=1}^{n}\boldsymbol{X}_{ij}^2\Big)^{1/2}$,$\boldsymbol{X}\bullet\boldsymbol{Y}=tr(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{Y})$,$\forall\boldsymbol{X},\boldsymbol{Y}\in S^n$,$tr(\bullet)$是矩阵的迹. Gowda 等人指出半正定互补问题 SDCP(F,S_+^n)等价于一类特殊的变分不等式问题[4],记为半正定变分不等式 SDVI(F,S_+^n),即找到矩阵 $\boldsymbol{X}^*\in S_+^n$满足

$$-F(\boldsymbol{X}^*)\in N_{S_+^n}(\boldsymbol{X}^*), \tag{2}$$

式中,$N_{S_+^n}(\boldsymbol{A})=:\begin{cases}\boldsymbol{A}:\{(\boldsymbol{Y}-\boldsymbol{X})\bullet\boldsymbol{A}\leqslant 0,\forall\boldsymbol{Y}\in S_+^n\} & \text{如果 }\boldsymbol{X}\in S_+^n,\\ \varnothing & \text{否则.}\end{cases}$

收稿日期:2016-10-16.
基金项目:国家自然科学基金(U1233105).
通讯联系人:徐海文,博士生,副教授,研究方向:最优化理论与算法. E-mail:xuhaiwen_dream@163.com

由此半正定变分不等式问题作为一个有力工具应用在数学规划[5]、工业和经济[6]、系统与控制和交通[7]等方面取得了广泛的应用,更多应用见文献[8-9]. 很多学者在求解半正定互补问题时做了大量工作,主要有内点算法[3,10-11]、非内点算法[12-13]和光滑牛顿算法[14]等. 而投影收缩算法[15-19]是一类求解变分不等式的高效算法,文献[20]借鉴上述投影收缩算法的思想,得到了问题(2)的校正投影收缩算法,数值试验表明该算法是有效的,并且对一类没有显示函数的半正定变分不等式问题具有较好的优势. 但在校正投影收缩算法的下降量收敛证明中多次使用了放大不等式,并没有完全利用下降量,因此本文利用符合固定均值的随机数,适当地随机扩张校正投影收缩算法的步长,从而得到了半正定变分不等式问题的随机下降算法,并且在合适的假设条件下,给出了依概率收敛性证明.数值试验验证了随机下降算法的可行性和有效性,并表明了适当选择均值和方差可以提高计算效率.

下面给出矩阵函数及其单调性的定义.

定义 1[21] 对于 $\boldsymbol{X}\in S^n$,则 $F(\boldsymbol{X})=\boldsymbol{P}\mathrm{diag}[f(\lambda_1),f(\lambda_2),\cdots,f(\lambda_n)]\boldsymbol{P}^{\mathrm{T}}$.

这里 $\boldsymbol{X}$ 的特征值 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 是实的且谱分解格式为 $\boldsymbol{X}=\boldsymbol{P}\mathrm{diag}[\lambda_1,\lambda_2,\cdots,\lambda_n]\boldsymbol{P}^{\mathrm{T}}$,$\boldsymbol{P}\in W=\{\boldsymbol{A}\in\mathbf{R}^{n\times n}:\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{A}^{-1}\}$,$\mathrm{diag}[\lambda_1,\lambda_2,\cdots,\lambda_n]$是 $n\times n$ 对角阵,第 i 个对角元是 λ_i,f 是 $\mathbf{R}$ 到 $\mathbf{R}$ 上的一个映射.

定义 2 函数 $F:S^n\to S^n$ 是单调的,如果 $\forall\boldsymbol{X},\boldsymbol{Y}\in S^n$, 有.

$$[F(\boldsymbol{X})-F(\boldsymbol{Y})]\bullet(\boldsymbol{X}-\boldsymbol{Y})\geqslant 0.$$

记 $\boldsymbol{X}\in S^n$ 在 K 上的投影为 $P_K(\boldsymbol{X})$,且

$$P_K(\boldsymbol{X})=\arg\min\{\|\boldsymbol{X}-\boldsymbol{Y}\|\mid\forall\boldsymbol{Y}\in K\},\ \forall\ \boldsymbol{X}\in S^n.$$

定义 3 依概率收敛[22]. 设 $\boldsymbol{X}_1,\boldsymbol{X}_2,\cdots,\boldsymbol{X}_k,\cdots$是一列随机变量,如果对$\forall\delta>0$,恒有

$$\lim_{k\to\infty}P\{|\boldsymbol{X}_k-\boldsymbol{X}|\geqslant\delta\}=0.$$

则称$\{\boldsymbol{X}_k\}$依概率收敛到 $\boldsymbol{X}$,记作 $\boldsymbol{X}_k\xrightarrow{P}\boldsymbol{X}$.

引理 1 Markov 不等式[22]:对于$\forall\ \delta>0$ 和 $t>0$,有

$$P(|\boldsymbol{X}|\geqslant\delta)\leqslant\frac{(E|\boldsymbol{X}|)^t}{\delta^t}.$$

引理 2[22] 设 $\boldsymbol{X}_k\xrightarrow{P}\boldsymbol{X}$,则必有子序列 $\boldsymbol{X}_{n_k}\to\boldsymbol{X}a.s.$ 几乎必然地.

为了书写方便,文章中的后面部分记 $K=S^n_+$,$o(\boldsymbol{X},\beta)=\boldsymbol{X}-P_K[\boldsymbol{X}-\beta F(\boldsymbol{X})]$, $\forall\beta>0$.并且假定 Ω 为 SDVI(F,K)的解集且非空,$F(\boldsymbol{X})$为 S^n 上的单调函数,$0<\beta_l\leqslant\inf\limits_k\{\beta_k\}$. 随后将给出一些基本性质,这些性质和引理在算法收敛性的证明过程中起到了重要作用.

引理 3 $\boldsymbol{X}^*\in K$ 是 SDVI(F,K)的解当且仅当 $o(\boldsymbol{X}^*,\beta)=0$, $\forall\beta>0$.

由于 $K=S^n_+$是 S^n 中的非空闭凸子集,根据投影、内积和范数的定义,可以得到下面的引理.

引理 4 任意 $\boldsymbol{X}\in K,\boldsymbol{Y}\in S^n$, 有 $\|\boldsymbol{X}-P_K(\boldsymbol{Y})\|^2\leqslant\|\boldsymbol{X}-\boldsymbol{Y}\|^2-\|\boldsymbol{Y}-P_K(\boldsymbol{Y})\|^2$.

引理 5 任意 $\boldsymbol{X}\in K$ 和 $\tilde{\beta}\geqslant\beta_l>0$,有

$$\|o(\boldsymbol{X},\tilde{\beta})\|\geqslant\|o(\boldsymbol{X},\beta_l)\|.$$

1 半正定变分不等式问题的随机下降算法和收敛性定理

1.1 半正定变分不等式问题的随机下降算法(SDC 算法)

步骤 1 初始化 $\boldsymbol{X}=\boldsymbol{X}^0\in K$,stop $c=1$,$\beta_0=1$,$0<r<s<1$,$\boldsymbol{X}^0\in K$,$\underline{\tau}$,$\bar{\tau}$,$k=0$,容许度 ε;

步骤 2 计算 $\boldsymbol{Y}^k=P_K[\boldsymbol{X}^k-\beta_kF(\boldsymbol{X}^k)]$.

步骤 3 当 stop $c\leqslant\varepsilon$,停机;

如果

$$\delta_k:=\frac{\beta_k\|F(\boldsymbol{X}^k)-F(\boldsymbol{Y}^k)\|}{\|\boldsymbol{X}^k-\boldsymbol{Y}^k\|}\leqslant s,$$

那么依次计算

$$\text{stop }c=\|o(\boldsymbol{X}^k,\beta_k)\|=\|\boldsymbol{X}^k-\boldsymbol{Y}^k\|,$$

$$\alpha(\boldsymbol{X}^k,\beta_k)=o(\boldsymbol{X}^k,\beta_k)-\beta_k[F(\boldsymbol{X}^k)-F(\boldsymbol{Y}^k)],$$

$$\tau^k=\sigma(\underline{\tau},\bar{\tau}),$$

$$\eta_k:=\frac{1}{k}\sum_{i=1}^{k}\tau_i,$$

$$\xi_k=\eta_k\omega^k=\eta_k\frac{o(\boldsymbol{X}^k,\beta_k)\bullet\alpha(\boldsymbol{X}^k,\beta_k)}{\|\alpha(\boldsymbol{X}^k,\beta_k)\|^2},$$

$$\boldsymbol{X}^{k+1}=P_K[\boldsymbol{X}^k-\xi_k\beta_kF(\boldsymbol{Y}^k)],$$

$$\beta_k:=\begin{cases}1.5\beta_k & \text{如果 } \delta_k\leqslant r,\\ \beta_k & \text{否则.}\end{cases}$$

令 $\beta_{k+1}=\beta_k$ 和 $k=k+1$,返回到步 2. 否则转到步 4.

步骤 4　缩小 β_k, 即 $\beta_k=\dfrac{2\beta_k\min\{1,1/r_k\}}{3}$,返回到步 2.

注 1　$\sigma(\underline{\tau},\bar{\tau})$ 为服从区间$(\underline{\tau},\bar{\tau})$上的高斯分布或均匀分布产生的随机数且数学期望 $E(\tau^k)=\nu_0$.

注 2　利用 $\alpha(\boldsymbol{X}^k,\beta_k)$ 的定义,可知

$$\omega^k=\frac{o(\boldsymbol{X}^k,\beta_k)\bullet\alpha(\boldsymbol{X}^k,\beta_k)}{\|\alpha(\boldsymbol{X}^k,\beta_k)\|^2}\geqslant\frac{1-s}{1+s^2}>\frac{1-s}{2},\ \forall k>0.$$

SDC 算法保留了投影收缩算法的最优步长改进和 β_k 的自适应控制,增加了高斯分布或均匀分布随机数扩张步长,增加了步长选取的灵活性同时增加了 η_k 的期望选择控制,确保了算法的依概率收敛性.

1.2　收敛性定理

利用上面的引理 3 和 SDVI(F,K)的基本定义,容易得到下面的引理.

引理 6　假设$\{\boldsymbol{X}^k\}$为 SDC 算法产生的序列,$0<r<s<1$,则

$$(\boldsymbol{X}^{k+1}-\boldsymbol{Y}^k)\bullet\beta_kF(\boldsymbol{Y}^k)\geqslant(\boldsymbol{X}^{k+1}-\boldsymbol{Y}^k)\bullet\alpha(\boldsymbol{X}^k,\beta_k).$$

证明　由于 $\boldsymbol{Y}^k=P_K[\boldsymbol{X}^k-\beta_kF(\boldsymbol{X}^k)]=P_k\{\boldsymbol{X}^k-[\beta_kF(\boldsymbol{Y}^k)-\alpha(\boldsymbol{X}^k,\boldsymbol{Y}^k)]\}$,根据引理 3 和(2)知,

$$\boldsymbol{Y}^k\in K,(\boldsymbol{X}-\boldsymbol{Y}^k)\bullet[\beta_kF(\boldsymbol{Y}^k)-\alpha(\boldsymbol{X}^k,\beta_k)]\geqslant 0,\ \forall \boldsymbol{X}\in K.$$

令 $\boldsymbol{X}=\boldsymbol{X}^{k+1}$,可以得到

$$(\boldsymbol{X}-\boldsymbol{Y}^k)\bullet\beta_kF(\boldsymbol{Y}^k)\geqslant(\boldsymbol{X}^{k+1}-\boldsymbol{Y}^k)\bullet\alpha(\boldsymbol{X}^k,\beta_k).$$

由 $F(\boldsymbol{X})$的单调性,可以得到下面的引理.

引理 7　假设$\{\boldsymbol{X}^k\}$为 SDC 算法产生的序列,则

$$(\boldsymbol{X}^{k+1}-\boldsymbol{X}^*)\bullet F(\boldsymbol{Y}^k)\geqslant(\boldsymbol{X}^{k+1}-\boldsymbol{Y}^k)\bullet F(\boldsymbol{Y}^k).\tag{3}$$

证明　由于 $\boldsymbol{X}^*\in K$,根据(2)知 $-F(\boldsymbol{X}^*)\in N_{S^n_+}(\boldsymbol{X}^*)$,所以$(\boldsymbol{Y}^k-\boldsymbol{X}^*)\bullet F(\boldsymbol{X}^*)\geqslant 0$. 又 $F(\boldsymbol{X})$是单调的,$\boldsymbol{X}^k\in K,\boldsymbol{Y}^k\in K$,所以

$$(\boldsymbol{Y}^k-\boldsymbol{X}^*)\bullet F(\boldsymbol{Y}^k)\geqslant(\boldsymbol{Y}^k-\boldsymbol{X}^*)\bullet F(\boldsymbol{X}^*)\geqslant 0.$$

进而得到

$$(\boldsymbol{X}^{k+1}-\boldsymbol{X}^*)\bullet F(\boldsymbol{Y}^k)-(\boldsymbol{X}^{k+1}-\boldsymbol{Y}^k)\bullet F(\boldsymbol{Y}^k)=(\boldsymbol{Y}^k-\boldsymbol{X}^*)\bullet F(\boldsymbol{Y}^k)\geqslant 0.$$

所以

$$(\boldsymbol{X}^{k+1}-\boldsymbol{X}^*)\bullet F(\boldsymbol{Y}^k)\geqslant(\boldsymbol{X}^{k+1}-\boldsymbol{Y}^k)\bullet F(\boldsymbol{Y}^k).$$

为了简化证明,令

$$\Theta(\boldsymbol{X}^k,\boldsymbol{X}^{k+1})=2\xi^k\beta_k(\boldsymbol{X}^{k+1}-\boldsymbol{X}^*)\bullet F(\boldsymbol{Y}^k)+\|\boldsymbol{X}^k-\boldsymbol{X}^{k+1}\|^2,$$

$$\Lambda(\boldsymbol{X}^k,\boldsymbol{X}^{k+1})=\xi^k(\boldsymbol{X}^k-\boldsymbol{X}^{k+1})\bullet\alpha(\boldsymbol{X}^k,\beta_k)-\xi^k(\boldsymbol{X}^k-\boldsymbol{Y}^k)\bullet\alpha(\boldsymbol{X}^k,\beta_k).$$

下面的引理为 SDC 算法的收敛性证明奠定了基础,说明了 SDC 算法的下降量.

引理 8　假设$\{\boldsymbol{X}^k\}$为 SDC 算法产生的序列,$0<r<s<1,\boldsymbol{X}^0\in K,0<\underline{\tau}\leqslant E(\tau^k)\leqslant\nu_0\leqslant\bar{\tau}<2$,则

$$\|\boldsymbol{X}^{k+1}-\boldsymbol{X}^*\|^2\leqslant\|\boldsymbol{X}^k-\boldsymbol{X}^*\|^2-\frac{1}{4}\eta_k(2-\eta_k)(1-s)\|o(\boldsymbol{X}^k,\beta_k)\|^2.$$

证明　由引理 4 知,

$$\|X^*-X^{k+1}\|^2\leqslant\|X^*-X^k+\xi^k\beta_kF(Y^k)\|^2-\|X^k-\xi^k\beta_kF(Y^k)-X^{k+1}\|^2=\|X^k-X^*\|^2-2\xi^k\beta_k(X^k-X^*)\bullet F(Y^k)-\|X^k-X^{k+1}\|^2+2\xi^k\beta_k(X^k-X^{k+1})\bullet F(Y^k)=\|X^k-X^*\|^2-\Theta(X^k,X^{k+1}).\tag{4}$$

根据引理6和引理7,经过一些简单的化简可以得到

$$\Theta(X^k,X^{k+1})\geqslant 2\xi^k(X^{k+1}-Y^k)\bullet\beta_kF(Y^k)+\|X^k-X^{k+1}\|^2=2\xi^k(X^{k+1}-X^k)\bullet\beta_kF(Y^k)+2\xi^k(X^k-Y^k)\bullet\beta_kF(Y^k)+\|X^k-X^{k+1}\|^2=-2\Lambda(X^k,X^{k+1})+\|X^k-X^{k+1}\|^2.$$

对于$\Lambda(X^k,X^{k+1})$第一项,利用$A\bullet B=\frac{1}{2}\|A\|^2+\frac{1}{2}\|B\|^2-\frac{1}{2}\|A-B\|^2$,得到

$$\xi^k(X^k-X^{k+1})\bullet\alpha(X^k,\beta_k)=\frac{1}{2}\|X^k-X^{k+1}\|^2+\frac{1}{2}\|\xi^k\alpha(X^k,\beta_k)\|^2-\frac{1}{2}\|X^k-X^{k+1}-\xi^k\alpha(X^k,\beta_k)\|^2=\frac{1}{2}\|X^k-X^{k+1}\|^2+\frac{1}{2}(\xi^k)^2\|\alpha(X^k,\beta_k)\|^2-\frac{1}{2}\|X^k-X^{k+1}-\xi^k\alpha(X^k,\beta_k)\|^2.\tag{5}$$

对于$\Lambda(X^k,X^{k+1})$ 第二项,根据步长$\xi_k=\eta_k\omega^k=\eta_k\frac{o(X^k,\beta_k)\bullet\alpha(X^k,\beta_k)}{\|\alpha(X^k,\beta_k)\|^2}$,得到

$$-\xi^k(X^k-Y^k)\bullet\alpha(X^k,\beta_k)=-\eta_k(\omega^k)^2\|\alpha(X^k,\beta_k)\|^2.\tag{6}$$

根据式(5)和(6),得到

$$\Lambda(X^k,X^{k+1})=\frac{1}{2}\|X^k-X^{k+1}\|^2+\frac{1}{2}\eta_k^2(\omega_k)^2\|\alpha(X^k,\beta_k)\|^2-\frac{1}{2}\|X^k-X^{k+1}-\xi^k\alpha(X^k,\beta_k)\|^2-\eta_k(\omega^k)^2\|\alpha(X^k,\beta_k)\|^2.\tag{7}$$

由式(7),对$\Theta(X^k,X^{k+1})$化简,可以得到

$$\Theta(X^k,X^{k+1})\geqslant 2\eta_k(2-\eta_k)(\omega^k)^2\|\alpha(X^k,\beta_k)\|^2+\|X^k-X^{k+1}-\xi^k\alpha(X^k,\beta_k)\|^2.\tag{8}$$

利用SDC算法产生的条件

$$\delta_k:=\frac{\beta_k\|F(X^k)-F(Y^k)\|}{\|X^k-Y^k\|}\leqslant s<1,$$

容易得到

$$\|\alpha(X^k,\beta_k)\|^2=\|o(X^k,\beta_k)-\beta_k[F(X^k)-F(Y^k)]\|^2\geqslant(1-s)\|o(X^k,\beta_k)\|^2.\tag{9}$$

于是将式(8)代入到式(4)中,并且利用式(9)、注2,化简可以得到

$$\|X^{k+1}-X^*\|^2\leqslant\|X^k-X^*\|^2-\eta_k(2-\eta_k)(\omega^k)^2\|\alpha(X^k,\beta_k)\|^2-\|X^k-X^{k+1}-\xi^k\alpha(X^k,\beta_k)\|^2\leqslant\|X^k-X^*\|^2-\frac{1}{4}\eta_k(2-\eta_k)(1-s)^3\|o(X^k,\beta_k)\|^2\leqslant\|X^k-X^*\|^2-\frac{1}{4}\eta_k(2-\eta_k)(1-s)\|o(X^k,\beta_k)\|^2.$$

证毕.

由引理8容易得到下面的收敛性定理.

定理1 假设$\{X^k\}$为SDC算法产生的序列,$0<r<s<1$,$X^0\in K$,$0<\underline{\tau}\leqslant E(\tau^k)\leqslant\nu_0\leqslant\bar{\tau}<2$,则$\|X^k-X^*\|\xrightarrow{P}0$,$k\to\infty$.

证明 由于$\tau_1,\tau_2,\cdots,\tau_k$是分别独立同分布的,则有

$$E(\eta_k)=E(\frac{1}{k}\sum_{i=1}^k\tau_i)=\frac{1}{k}E(\sum_{i=1}^k\tau_i)=\nu_0.\tag{10}$$

考察引理8的结论,根据式(10)和数学期望的性质,得到

$$E\|X^{k+1}-X^*\|^2\leqslant E\|X^k-X^*\|^2-\frac{1}{4}E\eta_k(2-E\eta_k)(1-s)E\|o(X^k,\beta_k)\|^2\leqslant E\|X^k-X^*\|^2-\frac{1}{4}\nu_0(2-\nu_0)(1-s)E\|o(X^k,\beta_k)\|^2\leqslant E\|X^k-X^*\|^2-\frac{1}{4}M_0E\|o(X^k,\beta_k)\|^2.$$

$$(\exists M_0,\text{s.t}\ \frac{1}{4}\nu_0(2-\nu_0)(1-s)\geqslant M_0>0.)\tag{11}$$

从而可以得到$\lim\limits_{k\to\infty}E\|o(X^k,\beta_k)\|=0$, 于是根据引理1知,

$$\forall\delta>0 \text{ 和 } t>0, P(\|o(\boldsymbol{X}^k,\beta_k)\|\geqslant\delta)\leqslant\frac{(E\|o(\boldsymbol{X}^k,\beta_k)\|)^t}{\delta^t}\to 0, k\to\infty.$$

则 $\lim\limits_{k\to\infty}P\{\|o(\boldsymbol{X}^k,\beta_k)\|\geqslant\delta\}=0$,所以有 $\|o(\boldsymbol{X}^k,\beta_k)\|\xrightarrow{P}0$,又根据引理 5,得到

$$\|o(\boldsymbol{X}^k,\beta_l)\|\xrightarrow{P}0. \tag{12}$$

根据式(11)知,$E\|\boldsymbol{X}^k-\boldsymbol{X}^*\|$ 有界,从而 $E\boldsymbol{X}^k$ 有界,所以有

$$\exists \text{子列 } k_m, s.t\ \|\boldsymbol{X}^{k_m}-\boldsymbol{X}^{\nabla}\|\xrightarrow{P}0, m\to\infty. \tag{13}$$

利用(12)和(13),可以得到

$$\|o(\boldsymbol{X}^{k_m},\beta_l)\|\xrightarrow{P}\|o(\boldsymbol{X}^{\nabla},\beta_l)\|=0.$$

根据引理 3,则 $\boldsymbol{X}^{\nabla}$ 是解,令 $\boldsymbol{X}^*=\boldsymbol{X}^{\nabla}$,则有

$$\|\boldsymbol{X}^k-\boldsymbol{X}^*\|\xrightarrow{P}0, k\to\infty.$$

证毕.

由引理 2,得到下面的推论.

推论 1　假设 $\{\boldsymbol{X}^k\}$ 为 SDC 算法产生的序列,$0<r<s<1$,$\boldsymbol{X}^0\in K$,$0<\underline{\tau}\leqslant E(\tau^k)=\nu_0\leqslant\bar{\tau}<2$,则必有子序列 $\boldsymbol{X}^{n_k}\to\boldsymbol{X}^*\ a.s.$ 几乎必然地.

2　数值试验

前两个数值算例测试随机下降算法的可行性和适用性,数据结果见表 1 和表 2. 第三个算例采用随机数生成的测试,表 3 为 SDC 算法和 CPC 算法的数值结果比较,表明了符合区间[1,2]和[1.4,2.4]的均匀分布随机数的计算效果更好一些,在 CPU 时间和迭代步上都取得了一定优势. 表 4 为由服从不同均值和不同方差的高斯分布产生随机数下的数值结果. 就本文算例而言,表明了适当选择期望和方差是可以提高 SDC 算法的计算效率. 同时可以得到 SDC 算法不仅对半正定单调变分不等式问题有效,而且对非线性半正定单调变分不等式问题也有效. 这里取 stop $c=\|\boldsymbol{X}^k-\boldsymbol{Y}^k\|_\infty\leqslant\varepsilon=10^{-3}$ 或 10^{-5} 停机. 试验中取 $\boldsymbol{A}^0=\boldsymbol{I}_5$,$\boldsymbol{I}_5$ 是单位阵.所有程序均在 Inter(R)2.27G 处理器,4G 内存联想 IBM 个人笔记本电脑上运行的,操作系统 Windows 7,Matlab 7.10.

测试问题 1(可行性测试)　半正定变分不等式问题中函数为 $F(\boldsymbol{A})=L_M(\boldsymbol{A})+\boldsymbol{Q}$,$L_M$ 是 lyapunov 变换,即 $L_M:S^n\to S^n$ 且 $L_M(\boldsymbol{A})=\boldsymbol{M}*\boldsymbol{A}+\boldsymbol{A}*\boldsymbol{M}^{\mathrm{T}}$,$\boldsymbol{M}\in\mathbf{R}^{n\times n}$.

$$\boldsymbol{M}=\begin{pmatrix}3&3&3&3&3\\3&6&9&12&15\\3&9&18&30&45\\3&12&30&60&105\\3&15&45&105&200\end{pmatrix},\boldsymbol{Q}=\begin{pmatrix}4&-2&&&\\-2&4&-2&&\\&-2&4&-2&\\&&-2&4&-2\\&&&-2&4\end{pmatrix},$$

可以验证 0 是精确解,表 1 是 SDC 算法的计算情况,SDC 算法在几步之内就给出了最优解.

表 1　可行性测试的计算结果

Table 1　Results of feasibility test

迭代步数	最优解	CPU 时间	误差
13	0	0.030 199	0

测试问题 2(非线性半正定变分不等式问题)　半正定变分不等式问题中函数为 $F(\boldsymbol{A})=\boldsymbol{M}*\boldsymbol{A}^2+\boldsymbol{A}^2*\boldsymbol{M}^{\mathrm{T}}+10\boldsymbol{Q}$,这里

$$\boldsymbol{M}=\begin{pmatrix}25&3&3&3&3\\3&20&9&12&15\\3&9&50&30&40\\3&12&30&120&10\\3&15&40&10&200\end{pmatrix},\boldsymbol{Q}=\begin{pmatrix}2&-1&&&\\-1&2&-1&&\\&-1&2&-1&\\&&-1&2&-1\\&&&-1&-8\end{pmatrix}.$$

表 2 是 SDC 算法的计算情况,$\boldsymbol{X}^*$ 为计算获得的近似最优解.

$$X^* = \begin{pmatrix} 0.000\,2 & 0.000\,6 & 0.001\,1 & -0.000\,8 & -0.008\,8 \\ 0.000\,6 & 0.002\,3 & 0.003\,9 & -0.002\,9 & -0.031\,9 \\ 0.001\,1 & 0.003\,9 & 0.006\,7 & -0.004\,9 & -0.054\,8 \\ -0.000\,8 & -0.002\,9 & -0.004\,9 & 0.003\,6 & 0.039\,9 \\ -0.008\,8 & -0.031\,9 & -0.054\,8 & 0.039\,9 & 0.444\,6 \end{pmatrix}.$$

表 2　非线性问题的计算结果

Table 2　Results for nonlinear SDVI

迭代步数	CPU 时间	误差
23	0.032 023	9.9e-004

测试问题 3(随机生成的例子)　半正定变分不等式问题中的函数为 $F(\boldsymbol{X})=\boldsymbol{M}*\boldsymbol{X}+\boldsymbol{X}*\boldsymbol{M}^{\mathrm{T}}+\boldsymbol{Q}$,$\boldsymbol{M}=\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$,$\boldsymbol{A}$ 为 $(-1,1)$ 上的随机数产生的 $n\times n$ 随机机矩阵,$\boldsymbol{Q}=\mathrm{diag}(\boldsymbol{q})$,$\boldsymbol{q}$ 为服从$(-5,5)$上均匀分布的随机数,表 3 为 SDC 算法的计算情况.

产生随机矩阵的 Matlab 代码如下:

```
A=zeros(n,n);t=0;for i=1:n for j=1:n t=mod(t*31 416+13 846,46 261);
A(i,j)=t*(2/46 261)-1;end;end;C=400*eye(n);A=C.*A;M=A'*A;
T=0;q=zeros(n,1);for j=1:n t=mod(t*45 278+13 846,46 219);q(j)=t;end;
Q=(q/46 219-1.0)*10;Q=diag(q);
```

表 3　SDC 算法和 CPC 算法的数值结果比较

Table 3　Comparison of numerical results between SDC algorithm and CPC algorithm

n	$r=0.3,s=0.9,\varepsilon=10^{-5}$ SDC Method								CPC Method	
	[0.5,1.5]		[1,2]		[1.4,2.4]		[1,2.8]			
	It	Tcpu	It	Tcpu	It	Tcpu	It	Tcpu	It	Tcpu
100	1 709	6.407	1 695	6.222	1 707	6.392	1 708	6.455	1 722	8.946
200	1 637	35.146	1 568	33.309	1 564	33.514	1 563	33.834	1 572	53.482
300	2 917	226.548	2 169	158.184	2 131	151.808	2 937	207.380	2 909	215.266
400	2 100	311.267	2 027	300.197	2 028	294.445	2 135	309.395	2 134	337.330
500	2 046	554.822	1 815	502.230	1 587	432.243	2 266	608.517	2 070	593.044
1 000	2 034	3 731.705	2 087	3 784.279	2 094	3 811.022	2 095	3 789.001	2 034	3 874.145

$[\underline{\tau},\bar{\tau}]$表示服从区间$[\underline{\tau},\bar{\tau}]$的均匀分布随机数.

表 4　SDC 算法在不同均值和方差下的数值结果比较

Table 4　Comparision of numerical results of SDC algorithm under different mean and variance

均值 方差 n	$r=0.3,s=0.9,\varepsilon=10^{-5}$											
	$E_{\eta k}=1.0$				$E_{\eta k}=1.5$				$E_{\eta k}=1.9$			
	0.5		1.0		0.5		1.0		0.5		1.0	
	It	Tcpu	It	Tcpu	It	Tcpu	It	Tcpu	It	Tcpu	It	Tcpu
100	1 690	6.176	1 698	6.233	1 737	6.357	1 686	6.192	1 738	6.323	1 744	6.413
200	1 637	35.146	1 637	35.065	1 568	33.534	1 582	33.950	1 562	32.679	1 562	32.465
300	2 228	178.761	2 147	311.544	2 169	154.171	2 089	147.810	2 163	152.666	2 162	167.096
400	2 032	293.623	2 027	300.197	2 027	317.591	2 028	300.938	2 135	315.760	2 025	300.930
500	2 112	690.666	2 079	556.000	1 786	481.725	2 084	568.532	2 079	566.975	1 806	500.972

这里的随机数由高斯分布产生.

3　结论

本文给出了半正定变分不等式的随机下降算法,并且在适当的条件下,给出了 SDC 算法的依概率收敛性证明. 通过一些非线性半正定不等式问题和随机生成半正定变分不等式问题的数值测试,说明了随机下降算法的可行性和有效性,并且通过选择适当的均值和方差,可以提高算法的计算效率.

[参考文献]

[1]　ALIZADEH F. Interior point methods in semidefinite programming with application to combinatorial optimization[J]. SIAM J Optim,1995(5):13-51.

[2] KOJIMA M,SHINDOH S,HARA S. Interior-point methods for the monotone semidefinite linear complementarity problems in Symmetric Matrices[J]. SIAM J Optim,1997(7):86-125.

[3] SHIDA M,SHINDOH S. Monotone semidefinite complementarity problems[R]. Tokyo:Tokyo Institute of Technology,1996.

[4] GOWDA S M,SONG Y G. On semidefinite complementarity problems[J]. Math Program,2000(88):575-587.

[5] DANTZIG G B,COTTLE R W. Positive(semidefinite)matrices and mathematical programming[J]. Report ORC,1963(13):63-18.

[6] FERRIS M C,PANG J S. Engineering and economic applications of complementarity problems[J]. SIAM Rev,1997(39):669-713.

[7] BOYD S,GHOUI L E,RERON E,et al. Linear matrix inequalities in system and control theory[M]. Siam Stud Appl Math, Philadelphia:SIAM,1994.

[8] HARKER P T,PANG J S. Finite-dimensional variational inequality and nonlinear complementartity problems:a survey of theory,algorithms and applications[J]. Math Program,1990(48):161-220.

[9] FACCHINEI F,PANG J S. Finite-dimensional variational inequalities and complementarity problems,Vol. Ⅰ and Ⅱ[M]. Springer Series in Operations Research. New York:Springer Verlag,2003.

[10] MONTEIRO R D C,ZANJACOMO P R. General interior-point maps and existence of weighted paths for nonlinear semidefinite complementarity problems[J]. Math Oper Res,2000,25(3):381-399.

[11] SIM C K,ZHAO G. Asymptotic behavior of Helmberg-Kojima-Monteiro(HKM)paths in interior-point methods for monotone semidefinite linear complementarity problems:general theory[J]. J Optimiz Theory App,2008(137):11-25.

[12] TSENG P. Merit functions for semidefinite complementarity problems[J]. Math Program,1998(83):159-185.

[13] CHEN X,TSENG P. Non-interior continuation methods for solving semidefinite complementarity problems[J]. Math Program, 2003,95(3):431-474.

[14] KANZOW C,NAGEL C. Semidefinite programs:new search directions,smoothing-type methods,and numerical results[J]. SIAM J Optim,2002,13(1):1-23.

[15] HE B S,LIAO L Z. Improvements of some projection methods for monotone nonlinear variational inequalities[J]. J Optimiz Theory App,2002,112(1):111-128.

[16] HE B S,XU M H. A general framework of contraction methods for monotone variational inequalities[J]. Pac J Optim,2008, 4(2):195-212.

[17] HE B S,TAO M,YUAN X M. Alternating direction method with Gaussian back substitution for separable convex programming[J]. SIAM J Optim,2012(22):313-340.

[18] HAN D R,YUAN X M. A note on the alternating direction method of multipliers[J]. J Optimiz Theory App,2012,155(1):227-238.

[19] CAI X J,HAN D R,Xu L L. An improved first-order primal-dual algorithm with a new correction step[J]. J Glonal Optim, 2013,57(4):1 419-1 428.

[20] 徐海文,张黔川,杨成,等. 一类半正定变分不等式的 CPC 算法[J]. 四川师范大学学报(自然科学版),2009,32(4):450-453.

[21] BHATIA R. Matrix Analysis[M]. New York:Springer-Verlag,1997:3-6.

[22] 林正炎,陆传荣,苏中根. 概率极限理论基础[M]. 北京:高等教育出版社,2003:16-17.

[责任编辑:陆炳新]