运筹学与控制论

DOI:10.11721/cqnuj20220608

求解凸可行问题的非精确变样本采样投影算法*

杨振平[1]，赵　勇[2]

(1. 嘉应学院 数学学院,广东 梅州 514015; 2. 重庆交通大学 数学与统计学院，重庆 400074)

摘要:【目的】构造求解大规模凸可行问题的有效算法,以克服现有算法要求投影运算具有显式表达式或者可以求得精确投影的局限。【方法】借助非精确近似技术和变样本采样技术,提出求解凸可行问题的非精确变样本采样投影算法。【结果】在样本增长率和非精确参数满足一定的条件下,证明了算法依概率 1 的收敛性。然后在样本增长率分别为几何增长和多项式增长的条件下,分析了算法的收敛率和计算复杂度。特别地,当样本率呈几何增长时,算法具有线性收敛率。【结论】数值实验结果验证了算法的有效性。

关键词:凸可行问题;非精确;变样本采样;收敛率;复杂度

中图分类号:O221.5　　**文献标志码**:A　　**文章编号**:1672-6693(2022)06-0009-11

本文考虑如下形式的凸可行问题:求解 $\boldsymbol{x}\in\mathbf{R}^n$,使得:

$$\boldsymbol{x}\in\mathcal{X}, \tag{1}$$

其中:$\mathcal{X}\subseteq\mathbf{R}^n$ 为非空闭凸集。在实际应用中,集合 $\mathcal{X}$ 有时可表示为有限个简单非空闭凸集 $X_1,\cdots,X_m$ 的交集(这里"简单"指某个点到 X_i 的投影运算具有显示表达或容易求得某个点到 X_i 的近似投影),即 $\mathcal{X}=\bigcap\limits_{i=1}^{m}X_i$。然而,有时 $\mathcal{X}$ 不能表示为简单闭凸集的交集,但可利用可数或不可数随机近似集合$\{X_\xi\}$的交集精确或近似表示 $\mathcal{X}$,即 $\mathcal{X}=\bigcap\limits_{\xi\in\Xi}\mathcal{X}_\xi$ 或 $\mathcal{X}\subseteq\bigcap\limits_{\xi\in\Xi}\mathcal{X}_\xi$,其中:$\xi$ 为定义在概率空间(Ξ,P)中的随机变量。故可将 $\mathcal{X}$ 松弛为 $\bigcap\limits_{\xi\in\Xi}\mathcal{X}_\xi$。进而将问题(1)松弛为凸可行问题:求解 $\boldsymbol{x}\in\mathbf{R}^n$,使得:

$$\boldsymbol{x}\in\bigcap_{\xi\in\Xi}\mathcal{X}_\xi。 \tag{2}$$

近些年来,凸可行问题因在放射治疗、磁共振成像、小波去噪、数据压缩、神经网络、机器学习与最优控制领域的广泛应用[1-8]而备受关注,逐渐成为优化领域中重要且富有挑战性的热点研究课题之一。为有效求解凸可行问题,学者们提出了诸多高效的投影型算法,如次梯度投影法、均值投影法、循环投影法、最大距离投影法、序列投影法、交替条件梯度法等[9-15]。考虑到问题(2)的可行集是大规模简单集合的交集,甚至可能是由不可数随机近似集合的交集,故上述算法均不能有效求解此类凸可行问题。为进一步提高算法的有效性,学者们借助随机近似技术,提出了一系列求解大规模凸可行问题的随机投影型算法,例如小批量随机次梯度投影法[16]、随机块外推投影法[17]、小批量外推投影法[18]等。然而,上述算法均要求随机选取固定样本量的随机样本用于算法迭代,样本量的大小将影响算法的计算效率:当样本量太小时算法收敛慢甚至不收敛;当样本量太大时则需要支付较昂贵的计算成本。而且上述算法要求简单集合 X_i 的投影运算具有显示表达或可求得精确投影,而在实际计算中往往难以求得精确投影点。有鉴于此,本文将构造求解大规模凸可行问题的数值算法以克服上述困难。

受文献[18-20]的启发,本文结合变样本采样技术、非精确近似技术以及外推投影算法,提出求解凸可行问题的非精确变样本采样投影算法,该算法在某种程度上可以有效降低计算成本。在样本增长率和非精确参数满足一定的条件下,本文讨论了该算法收敛性,并进一步,分别在样本增长率为几何增长和多项式增长的条件下分析了算法的收敛率和复杂度。最后,本文通过数值实验结果验证了算法的有效性。

* **收稿日期**:2022-03-30　**修回日期**:2022-05-30　**网络出版时间**:2022-12-08 14:35

资助项目:国家自然科学基金(No. 12101262;No. 12001072);广东省基础与应用基础研究基金 (No. 2022A1515010263)

第一作者简介:杨振平,男,博士,研究方向为随机优化,E-mail:yangzhenping1026@163.com;**通信作者**:赵勇,男,副教授,博士,E-mail:zhaoyongty@126.com

网络出版地址:https://kns.cnki.net/kcms/detail//50.1165.N.20221207.1559.008.html

1 算法

本节首先给出求解凸可行问题(2)的非精确变样本采样投影算法。然后,在适当的假设条件下证明算法的渐近收敛性。具体算法步骤如下。

算法 1(非精确变样本采样投影算法) 步骤0,选取初始点 $\boldsymbol{x}^0\in\mathbf{R}^n$,非负序列$\{\alpha_k\}$和$\{\delta_k\}$,给定样本量序列$\{N_k\}$。令 $k:=0$。

步骤1,从 Ξ 中随机选取独立同分布样本 $\xi_1^k,\cdots,\xi_{N_k}^k\sim P$。

步骤2,求得 $\boldsymbol{x}^{k+1}$ 满足 $\|\boldsymbol{x}^{k+1}-\boldsymbol{y}^k\|\leqslant\delta_k$,其中:$\boldsymbol{y}^k:=\boldsymbol{x}^k-\alpha_k\left(\boldsymbol{x}^k-\frac{1}{N_k}\sum_{i=1}^{N_k}\Pi_{X_{\xi_i^k}}[\boldsymbol{x}^k]\right)$,$\|\cdot\|$为欧氏范数,$\Pi_{X_{\xi_i^k}}[\boldsymbol{x}^k]$表示 $\boldsymbol{x}^k$ 到 $X_{\xi_i^k}$ 的正交投影运算。令 $k:=k+1$,返回步骤1。

注 1 若 $\delta_k\equiv0$ 且 $N_k\equiv N$,则算法1退化为文献[18]中的算法RPM。考虑到投影运算在很多情形下难以找到精确解,故本文考虑 $\delta_k>0$ 的情形。此外,考虑到在第 k 次迭代时,算法需要计算 N_k 次投影计算。因此,为降低计算成本,可利用计算机并行技术减少计算量。

由文献[18]中定理3.3知,求解凸可行问题(2)可等价转化为求解如下优化问题:

$$\min_{\boldsymbol{x}\in\mathbf{R}^n}F(\boldsymbol{x}):=\frac{1}{2}\mathbb{E}[\|\boldsymbol{x}-\Pi_{X_\xi}[\boldsymbol{x}]\|^2],\tag{3}$$

其中:$\mathbb{E}[\cdot]$表示数学期望。记 $\mathcal{Y}$ 为问题(3)的解集,并作如下假设。

假设 1 $\mathcal{X}=\mathcal{Y}$。

根据文献[18]中引理3.5可知,如果存在 $\kappa>0$ 使得,对所有的 $\boldsymbol{x}\in\mathbf{R}^n$,均有 $\mathrm{dist}^2(\boldsymbol{x},\mathcal{X})\leqslant\kappa\,\mathbb{E}[\mathrm{dist}^2(\boldsymbol{x},\mathcal{X}_\xi)]$(随机线性正则条件),其中:$\mathrm{dist}(\boldsymbol{a},\mathcal{C})$表示 $\boldsymbol{a}$ 到集合 $\mathcal{C}$ 的距离,则 $\mathcal{X}=\mathcal{Y}$。根据文献[18],由假设1,有 $\mathcal{X}=\mathcal{Y}=\arg\min_{\boldsymbol{x}}F(\boldsymbol{x})$,最优值 $F^*=0$ 以及 $F(\boldsymbol{x})=\frac{1}{2}\mathbb{E}[\mathrm{dist}^2(\boldsymbol{x},\mathcal{X}_\xi)]$。因此,随机线性正则条件可等价表示为如下二阶增长性条件:

$$F(\boldsymbol{x})-F^*\geqslant\frac{1}{2\kappa}\|\boldsymbol{x}-\Pi_{\mathcal{X}}[\boldsymbol{x}]\|^2,\ \forall\,\boldsymbol{x}\in\mathbf{R}^n。\tag{4}$$

令 $\gamma\geqslant0$ 为满足下列不等式的最小常数:

$$\|\mathbb{E}[\boldsymbol{x}-\Pi_{X_\xi}[\boldsymbol{x}]]\|^2\leqslant\gamma\,\mathbb{E}[\|\boldsymbol{x}-\Pi_{\mathcal{X}_\xi}[\boldsymbol{x}]\|^2],\ \forall\,\boldsymbol{x}\in\mathbf{R}^n。\tag{5}$$

根据Jensen不等式可知,$\gamma\leqslant1$。对于特定的集合和分布 P,γ 可严格小于1。由文献[18]的定理4.1,有如下结论成立。

引理 1 假设随机线性正则条件成立,则对任意 $\boldsymbol{x}\in\mathbf{R}^n$ 均有:$(1-\gamma)\|\boldsymbol{x}-\boldsymbol{x}^*\|^2\leqslant\langle\Pi[\boldsymbol{x}]-\Pi[\boldsymbol{x}^*],\boldsymbol{x}-\boldsymbol{x}^*\rangle\leqslant\left(1-\frac{1}{\kappa}\right)\|\boldsymbol{x}-\boldsymbol{x}^*\|^2$,其中:$\Pi[\boldsymbol{x}]=\mathbb{E}[\Pi_{\mathcal{X}_\xi}[\boldsymbol{x}]]$和 $\boldsymbol{x}^*=\Pi_{\mathcal{X}}[\boldsymbol{x}]$。

接下来给出随机近似算法收敛性分析的基本工具,即Robbins-Siegmund收敛性结论。

引理 2[21] 设$\{\mathcal{F}_k\}$是 σ 代数中的递增序列且$\{y^k\}$,$\{u^k\}$,$\{a^k\}$,$\{b^k\}$都是关于$\{\mathcal{F}_k\}$的非负随机变量。如果在依概率1意义下有 $\sum a^k<\infty$,$\sum b^k<\infty$,且对所有 $k\geqslant0$ 均有$\mathbb{E}[y^{k+1}|\mathcal{F}_k]\leqslant(1+a^k)y^k-u^k+b^k$,则在依概率1意义下,$\{y^k\}$收敛且 $\sum u^k<\infty$。

为了研究算法的收敛性,需作如下假设。

假设 2 非负序列$\{\delta_k\}$和$\{N_k\}$满足 $\sum_{k=0}^{\infty}N_k^{-1}<\infty$ 和 $\sum_{k=0}^{\infty}(1+N_k)\delta_k^2<\infty$。

受文献[18]中定理6.2的启发,根据假设1、2和(5)式,有如下收敛性结果。

定理 1 假设 $\mathcal{X}$ 为非空闭凸集。令$\{\boldsymbol{x}^k\}_{k\geqslant0}$ 为算法1产生的迭代序列且 $0<\alpha_k<2$,则对所有的 $k\geqslant0$ 和 $\boldsymbol{x}^*\in\mathcal{X}$ 均有:

$$\mathbb{E}[\|\boldsymbol{x}^{k+1}-\boldsymbol{x}^*\|^2|\boldsymbol{x}^k]\leqslant\left(1+\frac{1}{N_k}\right)\|\boldsymbol{x}^k-\boldsymbol{x}^*\|^2-2(2\alpha_k-\alpha_k^2)F(\boldsymbol{x}^k)+(1+N_k)\delta_k^2。\tag{6}$$

若假设1、2成立,且存在 $0<\vartheta<1$ 使得 $\vartheta\leqslant\alpha_k\leqslant2-\vartheta$,则序列$\{\boldsymbol{x}^k\}$依概率1收敛到 $\mathcal{X}$ 的一个随机点,并有 $\lim_{k\to\infty}F(\boldsymbol{x}^k)=0$ 依概率1成立。

证明　任取 $\boldsymbol{x}^*\in\mathcal{X}$，则根据 Young 不等式和算法 1，有：

$$\|\boldsymbol{x}^{k+1}-\boldsymbol{x}^*\|^2\leqslant\left(1+\frac{1}{N_k}\right)\|\boldsymbol{y}^k-\boldsymbol{x}^*\|^2+(1+N_k)\|\boldsymbol{x}^{k+1}-\boldsymbol{y}^k\|^2\leqslant\left(1+\frac{1}{N_k}\right)\|\boldsymbol{y}^k-\boldsymbol{x}^*\|^2+(1+N_k)\delta_k^2。\quad(7)$$

为简便起见，定义 $\Pi_i[\cdot]=\Pi_{\mathcal{X}_{\xi_i^k}}[\cdot]$。对于上式右端的第一项，有：

$$\|\boldsymbol{y}^k-\boldsymbol{x}^*\|^2=\left\|\boldsymbol{x}^k-\boldsymbol{x}^*-\alpha_k\left(\boldsymbol{x}^k-\frac{1}{N_k}\sum_{i=1}^{N_k}\Pi_i[\boldsymbol{x}^k]\right)\right\|^2=\left\|\boldsymbol{x}^k-\boldsymbol{x}^*-\frac{\alpha_k}{N_k}\sum_{i=1}^{N_k}(\boldsymbol{x}^k-\Pi_i[\boldsymbol{x}^k])\right\|^2=$$

$$\|\boldsymbol{x}^k-\boldsymbol{x}^*\|^2-\frac{2\alpha_k}{N_k}\sum_{i=1}^{N_k}\langle\boldsymbol{x}^k-\boldsymbol{x}^*,\boldsymbol{x}^k-\Pi_i[\boldsymbol{x}^k]\rangle+\frac{\alpha_k^2}{N_k^2}\left\|\sum_{i=1}^{N_k}(\boldsymbol{x}^k-\Pi_i[\boldsymbol{x}^k])\right\|^2,\quad(8)$$

其中：$\langle\cdot,\cdot\rangle$ 表示内积运算。由于对所有 $\boldsymbol{x}^*\in\mathcal{X}\subseteq\mathcal{X}_{\xi_i^k}$ 均有 $\langle\Pi_i[\boldsymbol{x}^k]-\boldsymbol{x}^*,\boldsymbol{x}^k-\Pi_i[\boldsymbol{x}^k]\rangle\geqslant0$，因此，$\langle\boldsymbol{x}^k-\boldsymbol{x}^*,\boldsymbol{x}^k-\Pi_i[\boldsymbol{x}^k]\rangle=\langle\boldsymbol{x}^k-\Pi_i[\boldsymbol{x}^k],\boldsymbol{x}^k-\Pi_i[\boldsymbol{x}^k]\rangle+\langle\Pi_i[\boldsymbol{x}^k]-\boldsymbol{x}^*,\boldsymbol{x}^k-\Pi_i[\boldsymbol{x}^k]\rangle\geqslant\|\boldsymbol{x}^k-\Pi_i[\boldsymbol{x}^k]\|^2$。结合(8)式可得：

$$\|\boldsymbol{y}^k-\boldsymbol{x}^*\|^2\leqslant\|\boldsymbol{x}^k-\boldsymbol{x}^*\|^2-\frac{2\alpha_k}{N_k}\sum_{i=1}^{N_k}\|\boldsymbol{x}^k-\Pi_i[\boldsymbol{x}^k]\|^2+\frac{\alpha_k^2}{N_k^2}\left\|\sum_{i=1}^{N_k}(\boldsymbol{x}^k-\Pi_i[\boldsymbol{x}^k])\right\|^2=$$

$$\frac{\alpha_k^2}{N_k^2}\Big(\sum_{i=1}^{N_k}\|\boldsymbol{x}^k-\Pi_i[\boldsymbol{x}^k]\|^2+\sum_{i\neq j;i,j=1}^{N_k}\langle\boldsymbol{x}^k-\Pi_i[\boldsymbol{x}^k],\boldsymbol{x}^k-\Pi_j[\boldsymbol{x}^k]\rangle\Big)+\|\boldsymbol{x}^k-\boldsymbol{x}^*\|^2-\frac{2\alpha_k}{N_k}\sum_{i=1}^{N_k}\|\boldsymbol{x}^k-\Pi_i[\boldsymbol{x}^k]\|^2。\quad(9)$$

根据 F 的定义，可得：$F(\boldsymbol{x}^k)=\frac{1}{2}\mathbb{E}[\|\boldsymbol{x}^k-\Pi_i[\boldsymbol{x}^k]\|^2\mid\boldsymbol{x}^k]=\frac{1}{2}\mathbb{E}[\|\boldsymbol{x}^k-\Pi_{\mathcal{X}_\xi}[\boldsymbol{x}^k]\|^2\mid\boldsymbol{x}^k]$，对(9)式两边在 $\boldsymbol{x}^k$ 的条件下取条件期望，并结合 $\Pi_i[\boldsymbol{x}^k]$ 和 $\Pi_{\mathcal{X}_{\xi_j^k}}[\boldsymbol{x}^k]$ $(i\neq j)$ 的无关性，可得：

$$\mathbb{E}[\|\boldsymbol{y}^k-\boldsymbol{x}^*\|^2\mid\boldsymbol{x}^k]\leqslant$$

$$\frac{\alpha_k^2}{N_k^2}\sum_{i\neq j}\langle\mathbb{E}[\boldsymbol{x}^k-\Pi_i[\boldsymbol{x}^k]\mid\boldsymbol{x}^k],\mathbb{E}[\boldsymbol{x}^k-\Pi_j[\boldsymbol{x}^k]\mid\boldsymbol{x}^k]\rangle+\|\boldsymbol{x}^k-\boldsymbol{x}^*\|^2-4\alpha_kF(\boldsymbol{x}^k)+\frac{2\alpha_k^2}{N_k}F(\boldsymbol{x}^k)=$$

$$\|\boldsymbol{x}^k-\boldsymbol{x}^*\|^2-4\alpha_kF(\boldsymbol{x}^k)+\frac{2\alpha_k^2}{N_k}F(\boldsymbol{x}^k)+\frac{\alpha_k^2(N_k^2-N_k)}{N_k^2}\|\mathbb{E}[\boldsymbol{x}^k-\Pi_{\mathcal{X}_\xi}[\boldsymbol{x}^k]\mid\boldsymbol{x}^k]\|^2\leqslant$$

$$\|\boldsymbol{x}^k-\boldsymbol{x}^*\|^2-4\alpha_kF(\boldsymbol{x}^k)+\frac{2\alpha_k^2}{N_k}F(\boldsymbol{x}^k)+\frac{\alpha_k^2(N_k-1)\gamma}{N_k}\mathbb{E}[\|\boldsymbol{x}^k-\Pi_{\mathcal{X}_\xi}[\boldsymbol{x}^k]\|^2\mid\boldsymbol{x}^k]=$$

$$\|\boldsymbol{x}^k-\boldsymbol{x}^*\|^2-4\alpha_kF(\boldsymbol{x}^k)+\frac{2\alpha_k^2}{N_k}F(\boldsymbol{x}^k)+\frac{2\alpha_k^2(N_k-1)\gamma}{N_k}F(\boldsymbol{x}^k)=$$

$$\|\boldsymbol{x}^k-\boldsymbol{x}^*\|^2-4\alpha_kF(\boldsymbol{x}^k)+2\alpha_k^2\gamma F(\boldsymbol{x}^k)+\frac{2\alpha_k^2(1-\gamma)}{N_k}F(\boldsymbol{x}^k)\leqslant$$

$$\|\boldsymbol{x}^k-\boldsymbol{x}^*\|^2-4\alpha_kF(\boldsymbol{x}^k)+2\alpha_k^2\gamma F(\boldsymbol{x}^k)+2\alpha_k^2(1-\gamma)F(\boldsymbol{x}^k)=\|\boldsymbol{x}^k-\boldsymbol{x}^*\|^2-2(2\alpha_k-\alpha_k^2)F(\boldsymbol{x}^k),\quad(10)$$

其中：利用(5)式可得第二个不等号，利用 $\gamma\leqslant1$ 和 $N_k\geqslant1$ 可得最后一个不等号。因此，由(7)～(10)式知，对所有的 $k\geqslant0$ 和 $\boldsymbol{x}^*\in\mathcal{X}$，有：

$$\mathbb{E}[\|\boldsymbol{x}^{k+1}-\boldsymbol{x}^*\|^2\mid\boldsymbol{x}^k]\leqslant\left(1+\frac{1}{N_k}\right)\|\boldsymbol{x}^k-\boldsymbol{x}^*\|^2-\left(1+\frac{1}{N_k}\right)2(2\alpha_k-\alpha_k^2)F(\boldsymbol{x}^k)+(1+N_k)\delta_k^2\leqslant$$

$$\left(1+\frac{1}{N_k}\right)\|\boldsymbol{x}^k-\boldsymbol{x}^*\|^2-2(2\alpha_k-\alpha_k^2)F(\boldsymbol{x}^k)+(1+N_k)\delta_k^2。\quad(11)$$

其中：由 $0<\alpha_k<2$ 和 $F(\boldsymbol{x}^k)\geqslant0$ 可得第二个不等号。因此，由 $\vartheta\leqslant\alpha_k\leqslant2-\vartheta$，可得 $2\alpha_k-\alpha_k^2\geqslant\vartheta^2>0$。根据 $\sum_{k=0}^{\infty}N_k^{-1}<\infty$，$\sum_{k=0}^{\infty}(1+N_k)\delta_k^2<\infty$ 以及引理 2，对任意的 $\boldsymbol{x}^*\in\mathcal{X}$，$\|\boldsymbol{x}^k-\boldsymbol{x}^*\|^2$ 均在依概率 1 意义下收敛且 $\sum_{k=0}^{\infty}F(\boldsymbol{x}^k)<\infty$ 在依概率 1 意义下成立。因此，序列 $\{\boldsymbol{x}^k\}$ 在依概率 1 意义下有界。于是，序列 $\{\boldsymbol{x}^k\}$ 存在极限点 $\overline{\boldsymbol{x}}^*$。由 $\sum_{k=0}^{\infty}F(\boldsymbol{x}^k)<\infty$ 在依概率 1 意义下成立，可知 $F(\boldsymbol{x}^k)\to0$ 在依概率 1 意义下成立。故根据 F 的连续性，对 $\{\boldsymbol{x}^k\}$ 的任意聚点 $\overline{\boldsymbol{x}}^*$，均有 $F(\overline{\boldsymbol{x}}^*)=0$ 在依概率 1 意义下成立。故 $\overline{\boldsymbol{x}}^*\in Y$ 在依概率 1 意义下成立。由假设 1 可得，$\{\boldsymbol{x}^k\}$ 至少存在一个子列收敛到 $\mathcal{X}$ 的一个随机点 $\overline{\boldsymbol{x}}^*$。　　证毕

2 收敛率与复杂度分析

本节将在不同样本增长率条件下,研究算法1的收敛率和复杂度。下面首先考虑 N_k 呈几何增长的情形,即 $N_k=\mathcal{N}\lceil\rho^{-(k+1)}\rceil$ 的情形,其中:$\lceil\cdot\rceil$ 表示向上取整,$\rho\in(0,1)$,$\mathcal{N}>0$。为此,引入如下结论。

引理 3[22] 设序列 $\{a_k\}$ 满足:$a_0\leqslant c_0$ 且 $a_{k+1}\leqslant qa_k+c_1\rho^{k+1}$,$\forall\, q,\rho\in(0,1)$,$k\geqslant 0$,则对任意的 $k\geqslant 0$,均有:

$$a_k\leqslant\begin{cases}\left(c_0+\dfrac{c_1}{\max\{q/\rho,\rho/q\}-1}\right)\max\{q,\rho\}^k, & q\neq\rho\\ \left(c_0+\dfrac{c_1}{\ln((\varphi/q)^e)}\right)\varphi^k, & q=\rho,\varphi\in(q,1)\end{cases}。$$

受文献[22]的定理2、文献[23]的定理3.2以及文献[24]的定理3.3的启发,算法1具有如下收敛率结果。

定理 2 令 $\{\boldsymbol{x}^k\}$ 为算法1产生的迭代序列。设随机线性正则条件、假设2以及如下条件成立:i) 存 $0<\vartheta<1$,使得 $\vartheta\leqslant\alpha_k\leqslant 2-\vartheta$;ii) $N_k=\mathcal{N}\lceil\rho^{-(k+1)}\rceil$,$\delta_k=1/N_k$,$\rho\in(0,1)$,$\mathcal{N}\geqslant 1$;iii) 若 $\vartheta^2>\kappa$,则 $\kappa/\vartheta^2<\mathcal{N}<\kappa/(\vartheta^2-\kappa)$,否则,$\kappa/\vartheta^2<\mathcal{N}$。则有:$\mathbb{E}[\mathrm{dist}^2(\boldsymbol{x}^{k+1},\mathcal{X})]\leqslant\begin{cases}\left(\mathbb{E}[\mathrm{dist}^2(\boldsymbol{x}^0,\mathcal{X})]+\dfrac{2}{\mathcal{N}(\max\{\eta/\rho,\rho/\eta\}-1)}\right)\max\{\eta,\rho\}^k, & \eta\neq\rho\\ \left(\mathbb{E}[\mathrm{dist}^2(\boldsymbol{x}^0,\mathcal{X})]+\dfrac{2}{\mathcal{N}(\ln((\varphi/\eta)^e)}\right)\varphi^k, & \eta=\rho\end{cases}$。其中:$\eta=\left(1-\left(\dfrac{\vartheta^2}{\kappa}-\dfrac{1}{\mathcal{N}}\right)\right)$,$\varphi\in(\eta,1)$。

证明 根据(6)式,随机线性正则条件,$\delta_k=\dfrac{1}{N_k}$ 以及 $F^*=0$,有:$\mathbb{E}[\|\boldsymbol{x}^{k+1}-\boldsymbol{x}^*\|^2\mid\boldsymbol{x}^k]\leqslant\left(1+\dfrac{1}{N_k}\right)\|\boldsymbol{x}^k-\boldsymbol{x}^*\|^2-2(2\alpha_k-\alpha_k^2)(F(\boldsymbol{x}^k)-F^*)+\dfrac{(1+N_k)}{N_k^2}\leqslant\left(1+\dfrac{1}{\mathcal{N}}\right)\|\boldsymbol{x}^k-\boldsymbol{x}^*\|^2-\dfrac{2\alpha_k-\alpha_k^2}{\kappa}\mathrm{dist}^2(\boldsymbol{x}^k,\mathcal{X})+\dfrac{2}{N_k}$,其中:由 $N_k\geqslant\mathcal{N}$ 和(4)可得第二个不等号。对上式取全期望,得:

$$\mathbb{E}[\|\boldsymbol{x}^{k+1}-\boldsymbol{x}^*\|^2]\leqslant\left(1+\frac{1}{N}\right)\mathbb{E}[\|\boldsymbol{x}^k-\boldsymbol{x}^*\|^2]-\frac{2\alpha_k-\alpha_k^2}{\kappa}\mathbb{E}[\mathrm{dist}^2(\boldsymbol{x}^k,\mathcal{X})]+\frac{2}{N_k}。\tag{12}$$

取 $\boldsymbol{x}^*=\Pi_X[\boldsymbol{x}^k]$,则有:$\mathbb{E}[\mathrm{dist}^2(\boldsymbol{x}^{k+1},\mathcal{X})]=\mathbb{E}[\|\boldsymbol{x}^{k+1}-\Pi_X[\boldsymbol{x}^{k+1}]\|^2]\leqslant\mathbb{E}[\|\boldsymbol{x}^{k+1}-\boldsymbol{x}^*\|^2]$。
结合(12)式,有:

$$\mathbb{E}[\mathrm{dist}^2(\boldsymbol{x}^{k+1},\mathcal{X})]\leqslant\left(1+\frac{1}{N}-\frac{2\alpha_k-\alpha_k^2}{\kappa}\right)\mathbb{E}[\mathrm{dist}^2(\boldsymbol{x}^k,\mathcal{X})]+\frac{2}{N_k}。\tag{13}$$

因为 $\vartheta\leqslant\alpha_k\leqslant 2-\vartheta$ 和 $0<\vartheta<1$,故有 $2\alpha_k-\alpha_k^2\geqslant\vartheta^2$。于是,由(13)式可得:

$$\mathbb{E}[\mathrm{dist}^2(\boldsymbol{x}^{k+1},\mathcal{X})]\leqslant\left(1-\left(\frac{\vartheta^2}{\kappa}-\frac{1}{\mathcal{N}}\right)\right)\mathbb{E}[\mathrm{dist}^2(\boldsymbol{x}^k,\mathcal{X})]+\frac{2}{N_k}。\tag{14}$$

由 $N_k=\mathcal{N}\lceil\rho^{-(k+1)}\rceil\geqslant\mathcal{N}\rho^{-(k+1)}$ 和(14)式可知,$\mathbb{E}[\mathrm{dist}^2(\boldsymbol{x}^{k+1},\mathcal{X})]\leqslant\eta\,\mathbb{E}[\mathrm{dist}^2(\boldsymbol{x}^k,\mathcal{X})]+\dfrac{2}{\mathcal{N}}\rho^{k+1}$。结合iii)和引理3,可得结论。 证毕

类似于文献[22]中定理2、文献[23]中定理3.3以及文献[24]中定理3.4的证明,可以得到算法1的迭代复杂度和oracle复杂度。

定理 3 令序列 $\{\boldsymbol{x}^k\}$ 为算法1产生的迭代序列。设定理2中的假设成立,则为达到 $\mathbb{E}[\mathrm{dist}^2(\boldsymbol{x}^k,\mathcal{X})]\leqslant\varepsilon$,算法的迭代复杂度上界 $\mathcal{K}(\varepsilon)$ 为:

$$\mathcal{K}(\varepsilon):=\begin{cases}\dfrac{1}{\ln(1/\eta)}\ln\left(\left(\mathbb{E}[\mathrm{dist}^2(\boldsymbol{x}^0,\mathcal{X})]+\dfrac{2\rho}{\mathcal{N}(\eta-\rho)}\right)\varepsilon^{-1}\right), & \rho<\eta<1\\ \dfrac{1}{\ln(1/\vartheta)}\ln\left(\left(\mathbb{E}[\mathrm{dist}^2(\boldsymbol{x}^0,\mathcal{X})]+\dfrac{2\rho}{\mathcal{N}\ln((\vartheta/\eta)^e)}\right)\varepsilon^{-1}\right), & \rho=\eta。\\ \dfrac{1}{\ln(1/\rho)}\ln\left(\left(\mathbb{E}[\mathrm{dist}^2(\boldsymbol{x}^0,\mathcal{X})]+\dfrac{2\eta}{\mathcal{N}(\rho-\eta)}\right)\varepsilon^{-1}\right), & \eta<\rho<1\end{cases}$$

而样本函数计算次数的上界 $\mathcal{M}(\varepsilon)$ 为:

$$\mathcal{M}(\varepsilon):=\begin{cases}\dfrac{\mathcal{N}}{\rho\ln(1/\rho)}\left(\left(\mathbb{E}[\mathrm{dist}^2(\boldsymbol{x}^0,\mathcal{X})]+\dfrac{2\rho}{\mathcal{N}(\eta-\rho)}\right)\varepsilon^{-1}\right)^{\frac{\ln(1/\rho)}{\ln(1/\eta)}}+\mathcal{N}\mathcal{K}(\varepsilon),\ \rho<\eta<1\\ \dfrac{\mathcal{N}}{\eta\ln(1/\eta)}\left(\left(\mathbb{E}[\mathrm{dist}^2(\boldsymbol{x}^0,\mathcal{X})]+\dfrac{2\rho}{\mathcal{N}\ln((\vartheta/\eta)^e)}\right)\varepsilon^{-1}\right)^{\frac{\ln(1/\eta)}{\ln(1/\vartheta)}}+\mathcal{N}\mathcal{K}(\varepsilon),\ \rho=\eta。\\ \dfrac{\mathcal{N}}{\rho\ln(1/\rho)}\left(\mathbb{E}[\mathrm{dist}^2(\boldsymbol{x}^0,\mathcal{X})]+\dfrac{2\eta}{\mathcal{N}(\rho-\eta)}\right)\varepsilon^{-1}+\mathcal{N}\mathcal{K}(\varepsilon),\ \eta<\rho<1\end{cases}$$

下面，考虑 N_k 呈多项式增长的情形，即 $N_k=\mathcal{N}\lceil(k+1)^s\rceil$，其中，$s>0$，$\mathcal{N}>1$。类似于文献[22]中性质 1 的证明，下列结论成立。

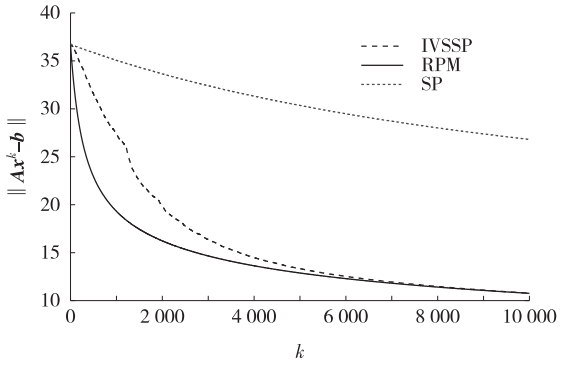
定理 4　令$\{\boldsymbol{x}^k\}$为算法 1 产生的迭代序列。设随机线性正则条件、假设 2 以及如下条件成立：i) $\vartheta\leqslant\alpha_k\leqslant2-\vartheta$，$0<\vartheta<1$；ii) $N_k=\mathcal{N}\lceil(k+1)^s\rceil$，$\delta_k=1/N_k$，$s>0$，$\mathcal{N}>1$；iii) 若 $\vartheta^2>\kappa$，则 $\kappa/\vartheta^2<\mathcal{N}<\kappa/(\vartheta^2-\kappa)$，否则，$\kappa/\vartheta^2<\mathcal{N}$。则有：$\mathbb{E}[\mathrm{dist}^2(\boldsymbol{x}^{k+1},\mathcal{X})]\leqslant\eta^{k+1}\left(\mathbb{E}[\mathrm{dist}^2(\boldsymbol{x}^0,\mathcal{X})]+\dfrac{2\mathrm{e}^{2s}\eta^{-1}}{\mathcal{N}(1-\eta)}\right)+\dfrac{2\eta^{-1}}{\mathcal{N}\ln(1/\eta)}(k+1)^{-s}$，其中：$\eta=\left(1-\left(\dfrac{\vartheta^2}{\kappa}-\dfrac{1}{\mathcal{N}}\right)\right)$。而且，为达到$\mathbb{E}[\mathrm{dist}^2(\boldsymbol{x}^k,\mathcal{X})]\leqslant\varepsilon$，算法的迭代复杂度和 oracle 复杂度分别为 $\mathcal{O}(s(1/\varepsilon)^{1/s})$和 $\mathcal{O}(\mathrm{e}^s s^s(1/\varepsilon)^{1+\frac{1}{s}})$。

3 数值实验

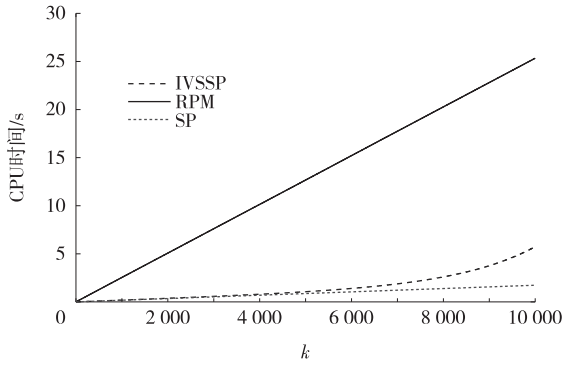
为验证算法的有效性和可比较性，本节将算法 1（记为 IVSSP）与文献[17]中取 $\tau=1$ 时的算法 SBP(记为 SP)和文献[18]中的算法 RPM 进行比较。在分别重复执行算法 20 次之后，比较算法平均误差变化和平均 CPU 时间(单位 s)变化。对于算法 IVSSP 和算法 RPM，取步长 $\alpha=1.9$。而对于算法 SP，则取步长 $\alpha=0.05$。所有的算法均在迭代次数达到 $K=10^4$ 时停止。为研究不同样本量增长方式对算法 IVSSP 的影响，本文考虑了 4 种几何增长和 4 种多项式增长，分别为 $N_k=\lceil0.999\ 43^{-(k+1)}\rceil$，$N_k=\lceil0.999\ 379^{-(k+1)}\rceil$，$N_k=\lceil0.999\ 361^{-(k+1)}\rceil$，$N_k=\lceil0.999\ 332^{-(k+1)}\rceil$，$N_k=\lceil(k+1)^{0.619}\rceil$，$N_k=\lceil(k+1)^{0.674}\rceil$，$N_k=\lceil(k+1)^{0.695}\rceil$和 $N_k=\lceil(k+1)^{0.726}\rceil$。上述样本量取值可使得 $N_K=300$、500、600、800。对于算法 RPM，则取 $N\equiv N_K$。程序运行环境为 Windows10 (64 bit) 系统 Intel(R) Core(TM) i9-10900 CPU @ 2.80 GHz，RAM：64 GB。MATLAB 版本为 2018b。

第一个例子选自文献[18]的例 1。考虑线性等式方程组 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$，其中，矩阵 $\boldsymbol{A}\in\mathbf{R}^{m\times n}$ 和向量 $\boldsymbol{b}\in\mathbf{R}^m$ 中的元素均在区间[−2，2]随机产生，取 $m=n=1\ 000$。从数值结果(图 1～4)可以发现，无论样本量是几何增长还是多项式增长，由算法 IVSSP 和算法 RPM 产生的误差$\|\boldsymbol{A}\boldsymbol{x}^k-\boldsymbol{b}\|$保持相同的减小趋势。同时，在达到最大迭代次数时，两个算法产生的误差几乎一致。然而，算法 IVSSP 所支付的计算成本比算法 RPM 低很多：当算法 IVSSP 的样本量呈几何增长时，算法 IVSSP 的 CPU 时间大概是 RPM 的$\dfrac{1}{5}$，即算法 IVSSP 在得到相同的误差的条件下，减少了大约 80% 的计算成本；而当算法 IVSSP 的样本量呈多项式增长时，虽然算法 IVSSP 的 CPU 时间要比几何增长时有所增加，但仍然比算法 RPM 的计算时间少很多，减少了大约 40% 的计算成本。因此，从计算成本而言，算法 IVSSP 优于算法 RPM。对于算法 SP，虽然它的 CPU 时间要优于算法 IVSSP，但是它的迭代误差要远大于算法 IVSSP。故从计算误差而言，算法 IVSSP 优于算法 SP。

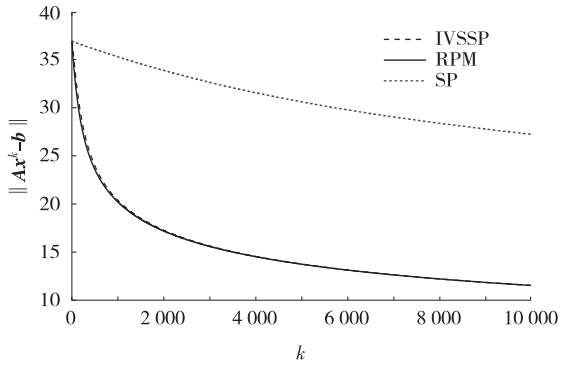
第二个例子选自文献[17]的例 1。考虑线性不等式方程组 $\boldsymbol{B}\boldsymbol{x}\leqslant\boldsymbol{b}$，其中矩阵 $\boldsymbol{B}=\lambda\boldsymbol{I}_m+\boldsymbol{C}\in\mathbf{R}^{m\times m}$ 和向量 $\boldsymbol{b}\in\mathbf{R}^m$，$\boldsymbol{I}_m$ 为 $m\times m$ 的单位矩阵。矩阵 $\boldsymbol{C}$ 是由数据集 CIFAR-10（该数据集详见 https://www.cs.toronto.edu/kriz/cifar.html）产生的 Gaussian 核矩阵，即给定训练集$\{\boldsymbol{z}_i,\boldsymbol{y}_i\}_{i=1}^m$，计算 $C_{ij}=\mathrm{e}^{-\zeta\|\boldsymbol{z}_i-\boldsymbol{z}_j\|^2}$，而$\boldsymbol{b}=\boldsymbol{y}$。在测试中，取 $\zeta=1$ 和 $\lambda=0.01$，并用$\|\max(0,\boldsymbol{B}\boldsymbol{x}^k-\boldsymbol{b})\|$来度量算法产生的误差。从数值结果(图 5～8)中可以看出，当算法 IVSSP 的样本量呈多项式增长时，由算法 IVSSP 和算法 RPM 产生的误差具有相同的下降趋势，而算法 IVSSP 所花费的时间要比算法 RPM 至少减少 11% 的计算时间。而且算法 IVSSP 减少的计算成本随最大样本量增加而增加：当 $N_K=800$ 时，算法 IVSSP 减少 23% 的 CPU 时间。当算法 IVSSP 的样本量呈几何增长时，虽然算法 IVSSP 产生的误差比算法 RPM 稍大，但是在算法终止时，两个算法产生的误差很接近。此时，算法 IVSSP 至少减少 30% 的 CPU 时间。从该角度而言，算法 IVSSP 优于算法 RPM。对于算法 SP，与第一个例子类似，虽然算法 SP 的 CPU 时间要少于算法 IVSSP，但是它的迭代误差要远劣于后者。故算法 IVSSP 优于算法 SP。



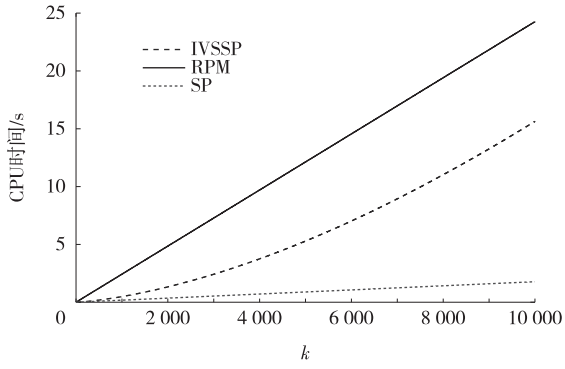
a 几何样本量误差变化趋势



b 几何样本量 CPU 时间



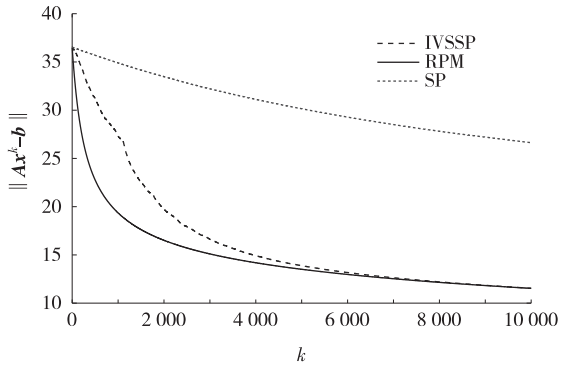
c 多项式样本量误差变化趋势



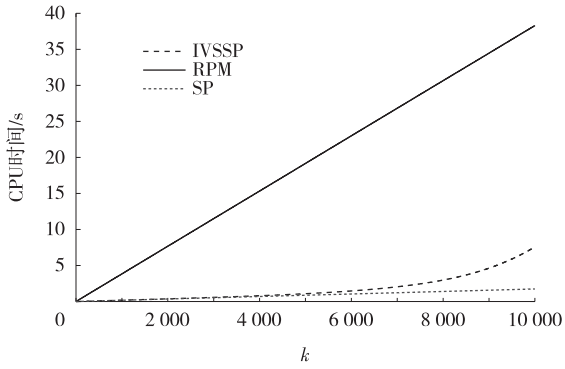
d 多项式样本量 CPU 时间

图 1 $\boldsymbol{Ax}=\boldsymbol{b}$, $N_K=300$

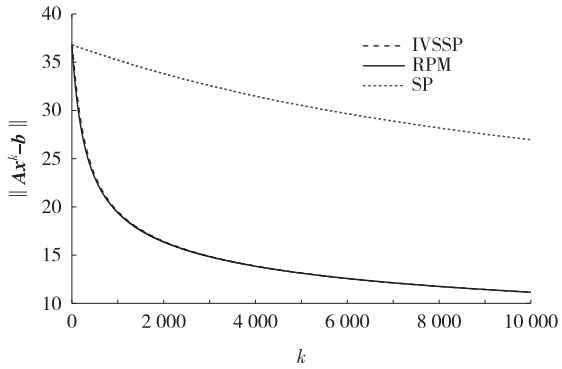
Fig. 1 $\boldsymbol{Ax}=\boldsymbol{b}$, $N_K=300$



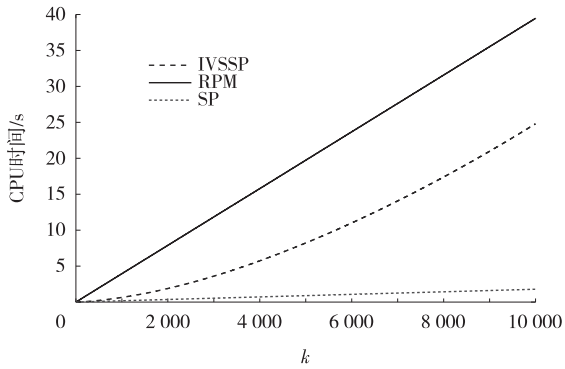
a 几何样本量误差变化趋势



b 几何样本量 CPU 时间



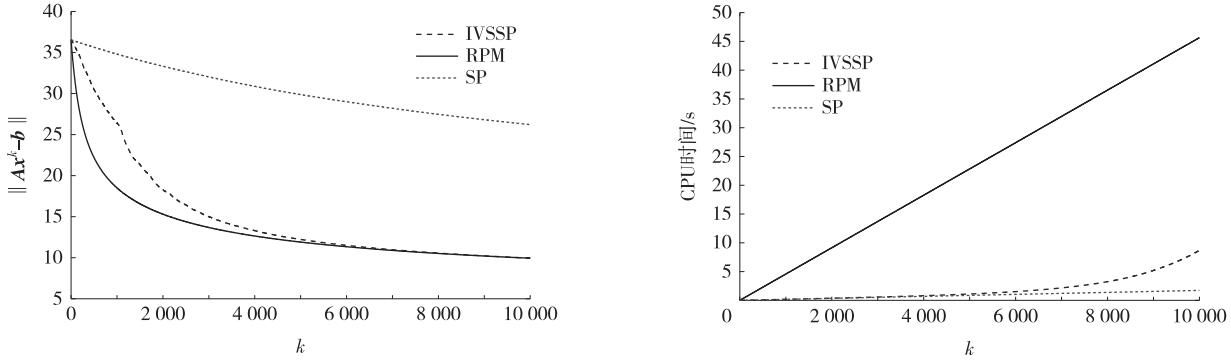
c 多项式样本量误差变化趋势



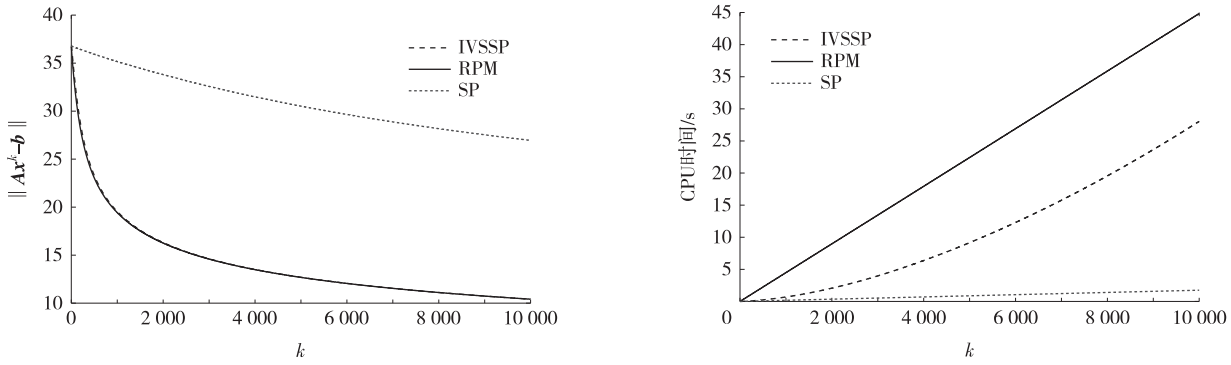
d 多项式样本量 CPU 时间

图 2 $\boldsymbol{Ax}=\boldsymbol{b}$, $N_K=500$

Fig. 2 $\boldsymbol{Ax}=\boldsymbol{b}$, $N_K=500$



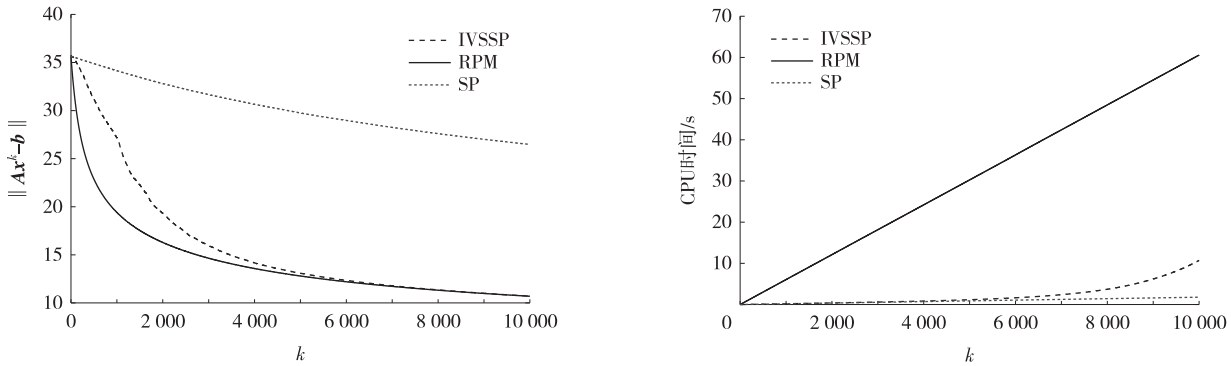
a　几何样本量误差变化趋势　　b　几何样本量 CPU 时间



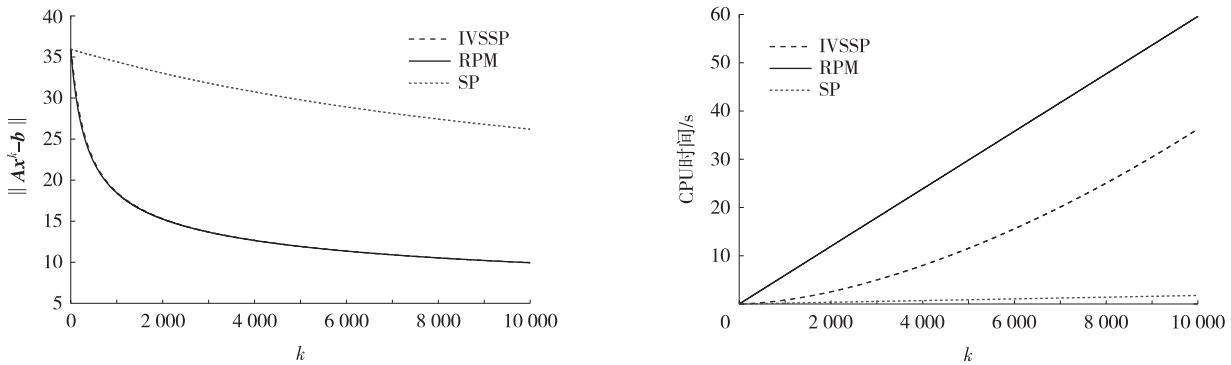
c　多项式样本量误差变化趋势　　d　多项式样本量 CPU 时间

图 3　$\boldsymbol{Ax}=\boldsymbol{b}$，$N_K=600$

Fig. 3　$\boldsymbol{Ax}=\boldsymbol{b}$，$N_K=600$



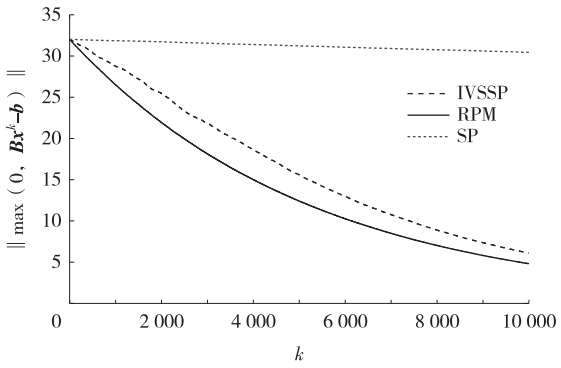
a　几何样本量误差变化趋势　　b　几何样本量 CPU 时间



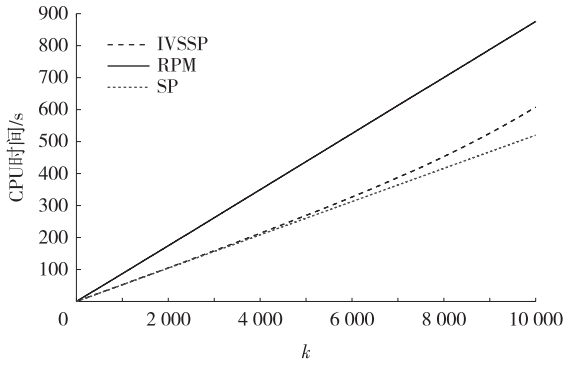
c　多项式样本量误差变化趋势　　d　多项式样本量 CPU 时间

图 4　$\boldsymbol{Ax}=\boldsymbol{b}$，$N_K=800$

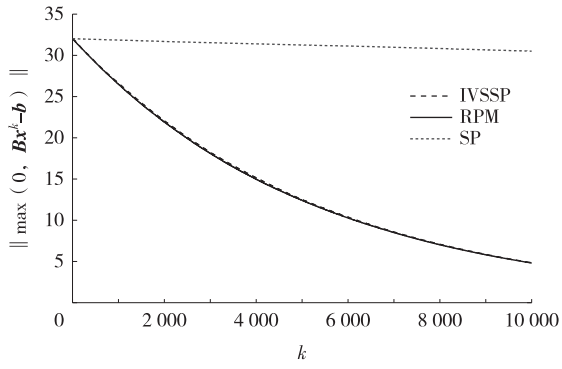
Fig. 4　$\boldsymbol{Ax}=\boldsymbol{b}$，$N_K=800$



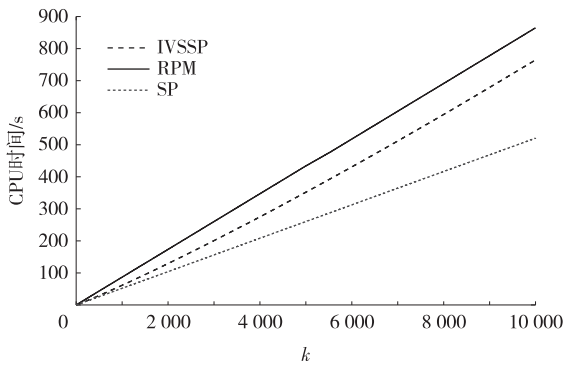
a 几何样本量误差变化趋势



b 几何样本量 CPU 时间



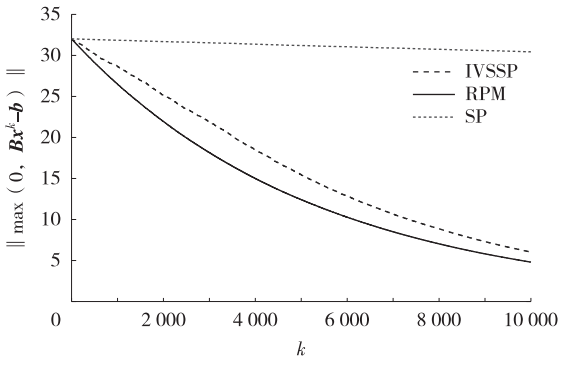
c 多项式样本量误差变化趋势



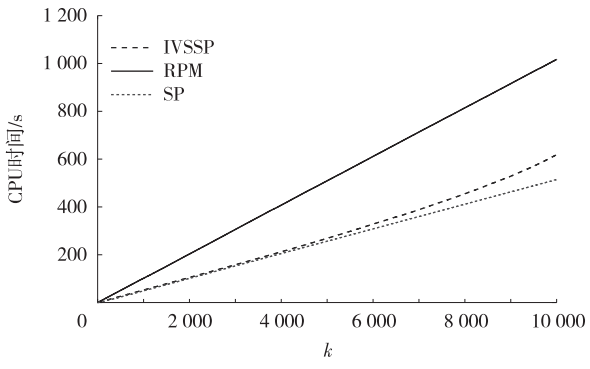
d 多项式样本量 CPU 时间

图 5 $\boldsymbol{Ax}=\boldsymbol{b}$, $N_K=300$

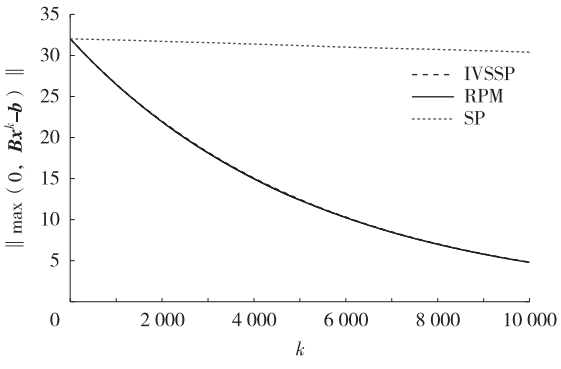
Fig. 5 $\boldsymbol{Ax}=\boldsymbol{b}$, $N_K=300$



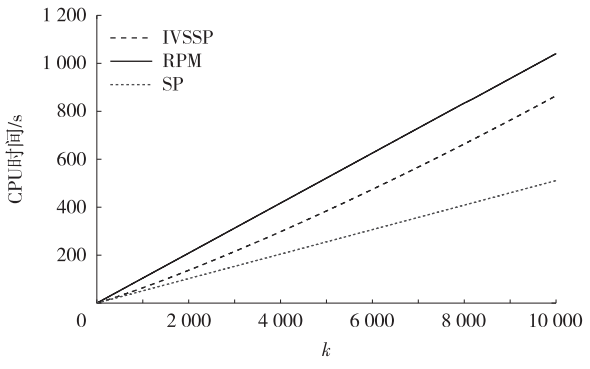
a 几何样本量误差变化趋势



b 几何样本量 CPU 时间



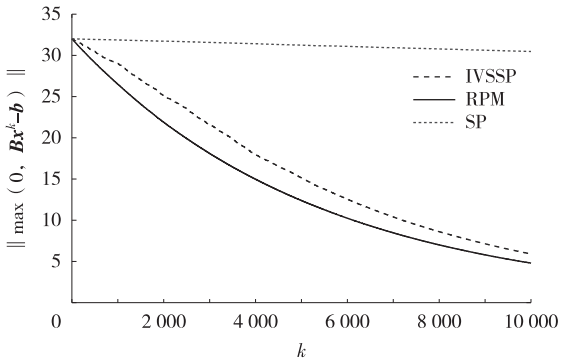
c 多项式样本量误差变化趋势



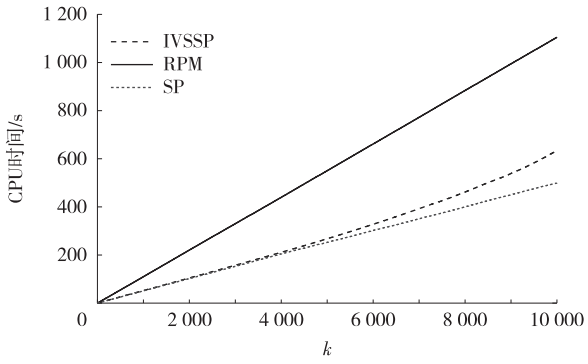
d 多项式样本量 CPU 时间

图 6 $\boldsymbol{Ax}=\boldsymbol{b}$, $N_K=500$

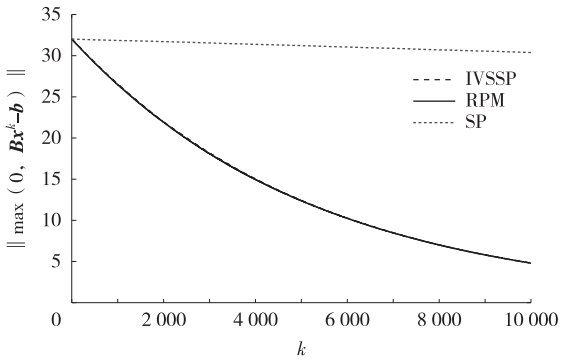
Fig. 6 $\boldsymbol{Ax}=\boldsymbol{b}$, $N_K=500$



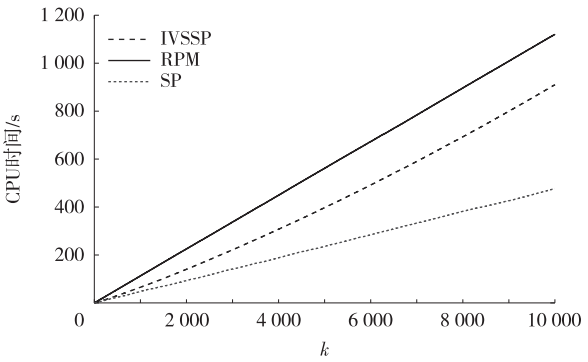
a　几何样本量误差变化趋势



b　几何样本量 CPU 时间



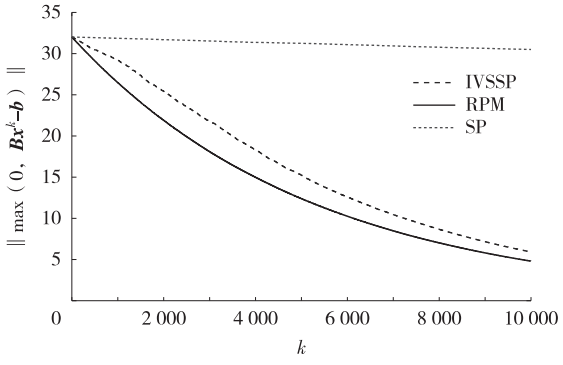
c　多项式样本量误差变化趋势



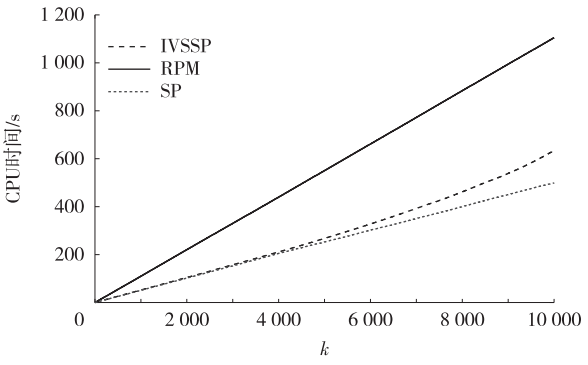
d　多项式样本量 CPU 时间

图 7　$\boldsymbol{Ax}=\boldsymbol{b}$，$N_K=600$

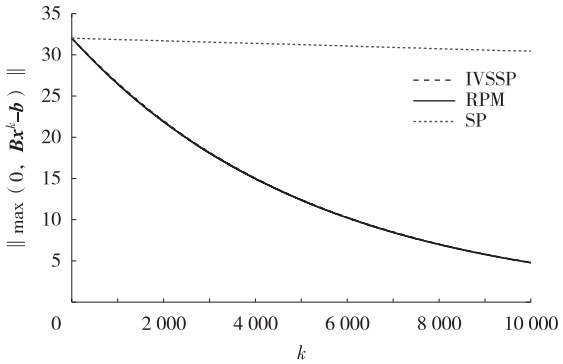
Fig. 7　$\boldsymbol{Ax}=\boldsymbol{b}$，$N_K=600$



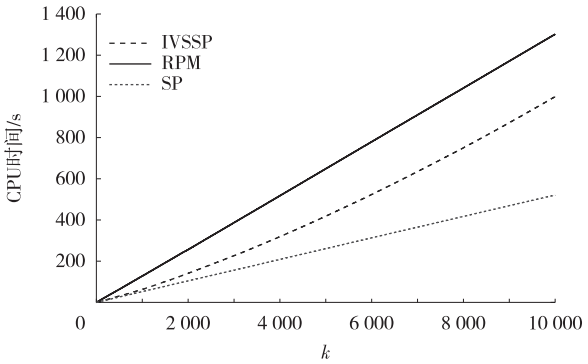
a　几何样本量误差变化趋势



b　几何样本量 CPU 时间



c　多项式样本量误差变化趋势



d　多项式样本量 CPU 时间

图 8　$\boldsymbol{Ax}=\boldsymbol{b}$，$N_K=800$

Fig. 8　$\boldsymbol{Ax}=\boldsymbol{b}$，$N_K=800$

4 结论

本文提出了求解大规模凸可行问题的非精确变样本采样投影算法。在适当的条件下,证明了算法的依概率1收敛性。分别在样本增长率为几何增长和多项式增长的条件下,研究了算法的收敛率和复杂度。特别地,在样本量呈几何增长时,算法达到线性收敛速度。最后,通过数值实验结果表明,所建立的算法是有效的且可比较的。

参考文献:

[1] HERMAN G T, CHEN W. A fast algorithm for solving a linear feasibility problem with application to intensity-modulated radiation therapy[J]. Linear Algebra and its Applications, 2008, 428(5/6): 1207-1217.

[2] SAMSONOV A A, KHOLMOVSKI E G, PARKER D L, et al. POCSENSE: POCS-based reconstruction for sensitivity encoded magnetic resonance imaging[J]. Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine, 2004, 52(6): 1397-1406.

[3] CHOI H, BARANIUK R G. Multiple wavelet basis image denoising using Besov ball projections[J]. IEEE Signal Processing Letters, 2004, 11(9): 717-720.

[4] SHARMA G. Set theoretic estimation for problems in subtractive color[J]. Color Research & Application, 2000, 25(5): 333-348.

[5] GU J, STARK H, YANG Y. Wide-band smart antenna design using vector space projection methods[J]. IEEE Transactions on Antennas and Propagation, 2004, 52(12): 3228-3236.

[6] BLATT D, HERO A O. Energy-based sensor network source localization via projection onto convex sets[J]. IEEE Transactions on Signal Processing, 2006, 54(9): 3614-3619.

[7] LIEWA W C, YAN H, LAW N F. POCS-based blocking artifacts suppression using a smoothness constraint set with explicit region modeling[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2005, 15(6): 795-800.

[8] PATRASCU A, NECOARA I. Nonasymptotic convergence of stochastic proximal point methods for constrained convex optimization[J]. The Journal of Machine Learning Research, 2017, 18(1): 7204-7245.

[9] BAUSCHKE H H, BORWEIN J M. On projection algorithms for solving convex feasibility problems[J]. SIAM review, 1996, 38(3): 367-426.

[10] ARAGON ARTECHO F J, CAMPOY R. A new projection method for finding the closest point in the intersection of convex sets[J]. Computational Optimization and Applications, 2018, 69(1): 99-132.

[11] BECK A, TEBOULLE M. Convergence rate analysis and error bounds for projection algorithms in convex feasibility problems [J]. Optimization Methods and Software, 2003, 18(4): 377-394.

[12] GOULD N I M. How good are projection methods for convex feasibility problems? [J]. Computational Optimization and Applications, 2008, 40(1): 1-12.

[13] CENSOR Y, CHEN W, COMBETTES P L, et al. On the effectiveness of projection methods for convex feasibility problems with linear inequality constraints[J]. Computational Optimization and Applications, 2012, 51(3): 1065-1088.

[14] DIAZ MILLAN R, FERREIRA O P, PRUDENTE L F. Alternating conditional gradient method for convex feasibility problems [J]. Computational Optimization and Applications, 2021, 80(1): 245-269.

[15] DANG Y, GAO Y. Non-monotonous accelerated parallel subgradient projection algorithm for convex feasibility problem[J]. Optimization, 2014, 63(4): 571-584.

[16] NECOARA I, NEDIC A. Minibatch stochastic subgradient-based projection algorithms for feasibility problems with convex inequalities[J]. Computational Optimization and Applications, 2021, 80(1): 121-152.

[17] NECOARA I. Stochastic block projection algorithms with extrapolation for convex feasibility problems[J]. Optimization Methods and Software, 2022: 1-31.

[18] NECOARA I, RICHTARIK P, PATRASCU A. Randomized projection methods for convex feasibility: Conditioning and convergence rates[J]. SIAM Journal on Optimization, 2019, 29(4): 2814-2852.

[19] IUSEM A N, JOFRE A, OLIVEIRA R I, et al. Extragradient method with variance reduction for stochastic variational inequalities[J]. SIAM Journal on Optimization, 2017, 27(2): 686-724.

[20] PATRASCU A, NECOARA I. On the convergence of inexact projection primal first-order methods for convex minimization

[J]. IEEE Transactions on Automatic Control,2018,63(10):3317-3329.

[21] ROBBINS H,SIEGUND D O. A convergence theorem for non-negative almost supermartingales and some applications[M]// RUSTAGI JS. Optimizing methods in statistics. NewYork:Academic Press,1971:233-257.

[22] LEI J L,SHANBHAG U V. Distributed variable sample-size gradient-response and best-response schemes for stochastic Nash games over graphs[J]. SIAM Journal on Optimization,2022,32(2):573-603.

[23] YANG Z P,LIN G H. Variance-based single-call proximal extragradient algorithms for stochastic mixed variational inequalities [J]. Journal of Optimization Theory and Applications,2021,190(2):393-427.

[24] YANG Z P, LIN G H. Two fast variance-reduced proximal gradient algorithms for SMVIPs-stochastic mixed variational Inequality Problems with suitable applications to stochastic network games and traffic assignment problems[J]. Journal of Computational and Applied Mathematics,2022,408:114132.

Operations Research and Cybernetics

Variable Sample Size Inexact Projection Algorithm for Convex Feasibility Problems

YANG Zhenping[1],ZHAO Yong[2]

(1. School of Mathematics,Jiaying University,Meizhou Guangdong 514015;

2. College of Mathematics and Statistics,Chongqing Jiaotong University, Chongqing 400074, China)

Abstract: [**Purposes**]An efficient algorithm for solving large-scale convex feasible problems is constructed to overcome existing algorithms that require explicit expression or accurate projection of the projection operations. [**Methods**]By virtue of the inexact approximation technology and variable sample-size sampling technology, an inexact variable sample-size sampling projection algorithm for solving convex feasible problems is proposed. [**Findings**]Under the conditions of the sample growth rate and inexact parameters, the asymptotic convergence of the algorithm is proved, and the convergence rate and complexity of the proposed algorithm are analyzed when the sample-size increases at a geometric rate or a polynomial rate. In particular, the linear convergence rate of the algorithm is achieved when the sample-size increases geometrically. [**Conclusions**]The effectiveness of the algorithm is verified by numerical experimental results.

Keywords: convex feasibility problem; inexact; variable sample-size; convergence rate; complexity

(责任编辑 陈 乔)