Analysis of the convergence rate for the cyclic projection algorithm applied to semi-algebraic convex sets

Liangjin Yao

University of Newcastle

June 3rd, 2013

The University of Newcastle
liangjin.yao@newcastle.edu.au
Joint work with Jon Borwein and Guoyin Li

Outline

1. **Motivation**
2. **Notation and auxiliary results**
   - Notation
3. **References**

Motivation

Given finitely many closed convex sets $C_1, C_2, \cdots, C_m$ in $\mathbb{R}^n$ with $\bigcap_{i=1}^m C_i \neq \varnothing$. Let $x_0 \in \mathbb{R}^n$. The sequence of *cyclic projections*, $(x_k)_{k\in\mathbb{N}}$, is defined by

$$x_1 := P_1 x_0, x_2 := P_2 x_1, \cdots, x_m := P_m x_{m-1}, x_{m+1} := P_1 x_m \ldots,$$

where P_i denotes the Euclidean projection to the set C_i.

Bregman showed that

$$x_k \longrightarrow x_\infty \in \bigcap_{i=1}^m C_i.$$

In this talk, we study

$$\text{how fast} \quad (x_k)_{k\in\mathbb{N}} \ \text{converges a point in} \bigcap_{i=1}^m C_i.$$

Notation and definitions

Throughout this talk,

- $\mathbb{R}^n$ is a *Euclidean space* with the norm $\|\cdot\|$ and the inner product $\langle\cdot,\cdot\rangle$.
- Let $C \subseteq \mathbb{R}^n$. The *interior* of C is $\operatorname{int} C$ and $\overline{C}$ is the *norm closure* of C.
- $\mathbb{B}(x_0,\delta) := \big\{x \in \mathbb{R}^n \mid \|x - x_0\| \leq \delta\big\}$.
- $[\alpha]_+ := \max\{\alpha, 0\}$.
- The *distance function* to the set C is $\operatorname{dist}(x, C) := \inf_{c\in C} \|x - c\|$.
- The *projector operator* to the set C is $P_C(x) := \{c \in C \mid \|y - c\| = \operatorname{dist}(x, C)\}$.

Notation and definitions

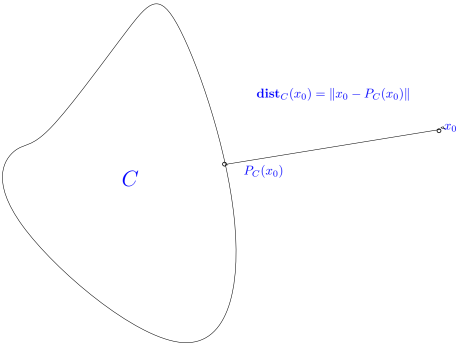

Figure : The projection of the point x_0 to the set C.

- $\beta(n) := \binom{n}{[n/2]}$ is the *central binomial coefficient* with respect to n.

Notation and definitions

- $f : \mathbb{R}^n \to \mathbb{R}$ is a *polynomial* if

$$f(x) := \sum_{0 \le |\alpha| \le r} \lambda_\alpha x^\alpha,$$

where $\lambda_\alpha \in \mathbb{R}$, $x = (x_1, \cdots, x_n)$, $x^\alpha := x_1^{\alpha_1} \cdots x_n^{\alpha_n}$, $\alpha_i \in \mathbb{N} \cup \{0\}$, and $|\alpha| := \sum_{j=1}^n \alpha_j$. The corresponding constant r is called the *degree* of f.

$$\boxed{\text{A polynomial } f \text{ is constant on } [x, y] \Rightarrow f \text{ is constant on } \operatorname{aff}[x, y].}$$

- We say that C is a *semi-algebraic convex set* if

$$\boxed{C = \{x \in \mathbb{R}^n \mid g_i(x) \le 0, i = 1, \cdots, N\},}$$

where g_i is convex polynomial for every $i = 1, \ldots, N$.

Notation and definitions

- From now on we assume that

$$m \in \mathbb{N}, \gamma_i \in \mathbb{N}, \; i = 1, \ldots, m$$

$g_{i,1}, g_{i,2}, \cdots, g_{i,\gamma_i}$ are convex polynomials on $\mathbb{R}^n$

$$C_i := \left\{ x \in \mathbb{R}^n \mid g_{i,1}(x) \leq 0, g_{i,2}(x) \leq 0, \cdots, g_{i,\gamma_i}(x) \leq 0 \right\}$$

$$P_i := P_{C_i}$$

$$C := \bigcap_{i=1}^{m} C_i \neq \varnothing.$$

Auxiliary results

Fact 1. (Kollár, 1999)
Let g_i be polynomials on $\mathbb{R}^n$ with degree $\leq d$ for every $i = 1, \cdots, m$. Let $g(x) := \max_{1 \leq i \leq m} g_i(x)$. Suppose that there exists $\varepsilon_0 > 0$ such that $g(x) > 0$ for all $x \in \mathbb{B}(0, \varepsilon_0) \backslash \{0\}$. Then there exist constants $c, \varepsilon > 0$ such that

$$\boxed{\|x\| \leq c\, g(x)^{\frac{1}{\beta(n-1)d^n}}, \quad \forall \|x\| \leq \varepsilon.}$$

Fact 2. (Li, 2010)
Let g be a convex polynomial on $\mathbb{R}^n$ with degree at most d. Let $S := \{x \mid g(x) \leq 0\}$ and $\bar{x} \in S$. Then there exist constants $c, \varepsilon > 0$ such that

$$\boxed{\operatorname{dist}(x, S) \leq c\, [g(x)]_+^{\frac{1}{(d-1)^n+1}}, \quad \forall \|x - \bar{x}\| \leq \varepsilon.}$$

Theorem 1. (Local error bounds for convex polynomial systems)
Let g_i be convex polynomials on $\mathbb{R}^n$ with degree at most d for every $i = 1, \cdots, m$. Let $S := \{x \in \mathbb{R}^n \mid g_i(x) \leq 0, i = 1, \cdots, m\}$ and $\bar{x} \in S$. Then there exist $c, \varepsilon > 0$ such that

$$\boxed{\mathrm{dist}(x, S) \leq c\Big(max_{1 \leq i \leq m}[g_i(x)]_+\Big)^{\tau} \quad \forall \|x - \overline{x}\| \leq \varepsilon,}$$

where $[\alpha]_+ := \max\{\alpha, 0\}$, $\tau := \max\left\{\frac{2}{(2d-1)^n+1}, \frac{1}{\beta(n-1)d^n}\right\}$, $\beta(n-1)$ is the central binomial coefficient with respect to $n-1$ which is given by $\binom{n-1}{[(n-1)/2]}$.

Theorem 2. (Hölderian regularity)
Let $\theta > 0$ and $K \subseteq \mathbb{R}^n$ be a compact set. Then there exists $c > 0$ such that

$$\boxed{\mathrm{dist}^{\theta}(x, C) \le c\left(\sum_{i=1}^{m} \mathrm{dist}^{\theta}(x, C_i)\right)^{\tau}, \quad \forall x \in K,}$$

where $\tau := \frac{1}{\min\left\{\frac{(2d-1)^{n}+1}{2}, \beta(n-1)d^{n}\right\}}$ and $\beta(n-1)$ is the central binomial coefficient with respect to $n-1$ which is given by $\binom{n-1}{[(n-1)/2]}$.

Theorem 3. (Cyclic convergence rate)
Suppose that $d > 1$. Let $x_0 \in \mathbb{R}^n$ and the sequence of cyclic projections, $(x_k)_{k\in\mathbb{N}}$, be defined by

$$x_1 := P_1 x_0,\ x_2 := P_2 x_1,\ \cdots,\ x_m := P_m x_{m-1},\ x_{m+1} := P_1 x_m \ldots.$$

Then x_k converges to $x_\infty \in C$, and there exists $M > 0$ such that

$$\boxed{\|x_k - x_\infty\| \leq M\frac{1}{k^\rho}, \quad \forall k \in \mathbb{N},}$$

where $\rho := \frac{1}{\min\left\{(2d-1)^n-1, 2\beta(n-1)d^n-2\right\}}$ and $\beta(n-1)$ is the central binomial coefficient with respect to $n-1$ which is given by $\binom{n-1}{[(n-1)/2]}$.

When $m = 2$, we can consider the general case where the intersection of these two sets is (possibly) empty.

We assume that

$$\boxed{\begin{array}{l} g_i, h_j \text{ are convex polynomials with degree at most } d \\ A := \{x \in \mathbb{R}^n \mid g_i(x) \leq 0, i = 1, \cdots, m\} \\ B := \{x \in \mathbb{R}^n \mid h_j(x) \leq 0, j = 1, \cdots, l\} \\ b_0 \in \mathbb{R}^n, \quad a_{k+1} := P_A b_k, \quad b_{k+1} := P_B a_{k+1}. \end{array}}$$

Theorem 4. (von Neumann alternating projection)
Assume $d > 1$. Then $a_k \longrightarrow \widetilde{a} \in A$ and $b_k \longrightarrow \widetilde{b} \in B$ with $\widetilde{b} - \widetilde{a} = v$ where $v := P_{B-A}0$. Moreover, there exists $M > 0$ such that

$$\boxed{\|a_k - \widetilde{a}\| \leq M\frac{1}{k^\rho} \quad \text{and} \quad \|b_k, -b\| \leq M\frac{1}{k^\rho}, \quad \forall k \in \mathbb{N},}$$

where $\rho := \frac{1}{\min\left\{(2d-1)^n-1, 2\beta(n-1)d^n-2\right\}}$ and $\beta(n-1)$ is the central binomial coefficient with respect to $n-1$ which is given by $\binom{n-1}{[(n-1)/2]}$.

Remark 1. The intersection of these two sets A and B is (possibly) empty.

Example 1.
Let

$$\begin{aligned}
C_1 &:= \{(x,y) \in \mathbb{R}^2 \mid (x+1)^2 + y^2 - 1 \le 0\}\\
C_2 &:= \{(x,y) \in \mathbb{R}^2 \mid x + y - 1 \le 0\}\\
C_3 &:= \{(x,y) \in \mathbb{R}^2 \mid (x-1)^2 + y^2 - 1 \le 0\}\\
C_4 &:= \{(x,y) \in \mathbb{R}^2 \mid x + (y+2)^2 - 4 \le 0\}.
\end{aligned}$$

Take $x_0 \in \mathbb{R}^2$. Let $(x_k)_{k\in\mathbb{N}}$ be defined by

$$x_1 := P_1 x_0,\ x_2 := P_2 x_1,\ x_3 := P_3 x_2,\ x_4 := P_4 x_3,\ x_5 := P_1 x_4 \ldots$$

Then $\|x_k\| = O(\frac{1}{k^{\frac{1}{6}}})$.

Example 2.

Let $\alpha \geq 0$ and

$$A := (-1, 0) + \mathbb{B}(0, 1) \quad \text{and} \quad B := \big\{(x, y) \in \mathbb{R}^2 \mid -x + \alpha \leq 0\big\}.$$

Let $(a_k)_{k \in \mathbb{N}}$ and $(b_k)_{k \in \mathbb{N}}$ be defined by

$$b_0 \in \mathbb{R}^2, \quad a_{k+1} := P_A b_k, \quad b_{k+1} := P_B a_{k+1}.$$

Then for every $k \geq 2$

$$b_k = \Big(\alpha, \frac{t_1}{\sqrt{(1+\alpha)^{2(k-1)} + t_1^2 \sum_{i=0}^{k-2}((1+\alpha)^{2i}}}\Big)$$

$$a_{k+1} = \Big(-1 + \frac{\alpha + 1}{\sqrt{(\alpha+1)^2 + \frac{t_1^2}{(1+\alpha)^{2(k-1)} + t_1^2 \sum_{i=0}^{k-2}((1+\alpha)^{2i}}}},$$

$$\frac{t_1}{\sqrt{(1+\alpha)^{2k} + t_1^2 \sum_{i=0}^{k-1}((1+\alpha)^{2i}}}\Big).$$

Consequently,

If $\alpha = 0$, $a_k \longrightarrow 0$ and $b_k \longrightarrow (\alpha, 0)$ at the rate of $k^{-\frac{1}{2}}$

if $\alpha \neq 0$(then $A \cap B = \varnothing$),

$a_k \longrightarrow 0$ and $b_k \longrightarrow (\alpha, 0)$ at the rate of $(1+\alpha)^{-k}$.

Remark 2.
According to Theorem 4, we can only deduce that $(a_k)_{k\in\mathbb{N}}$ in Example 2 converges to $(0,0)$ and $(b_k)_{k\in\mathbb{N}}$ converge to $(\alpha, 0)$ at the rate of at least of $k^{-\frac{1}{6}}$.

Example 3.
Let

$$A := \{(x,y) \in \mathbb{R}^2 \mid (x+1)^2 + y^2 - 1 \le 0\}$$
$$B := \{(x,y) \in \mathbb{R}^2 \mid (x-1)^2 + y^2 - 1 \le 0\}.$$

Let $(x_0, y_0) \in \mathbb{R}^2$ and $(x_k, y_k)_{k\in\mathbb{N}}$ be defined by

$$(x_1, y_1) := P_A(x_0, y_0), (x_2, y_2) := P_B(x_1, y_1), (x_3, y_3) := P_A(x_2, y_2), \quad \cdots .$$

Figure 2 depicts the algorithm's trajectory with starting point $(0, 2)$.

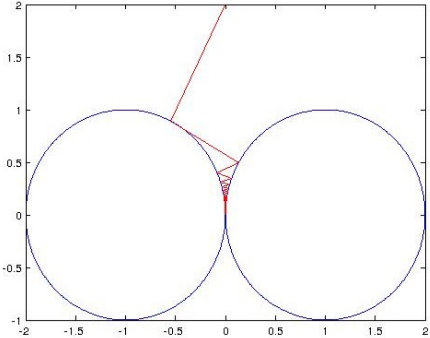

Figure : The iteration commencing at $(0, 2)$.

Example 3 continued

Since $(x_k, y_k) \in \text{bdry}\, A \cup \text{bdry}\, B$, $r_k := \sqrt{x_k^2 + y_k^2}$ satisfies $r_k^2 = 2\alpha_k$, where $\alpha_k := |x_k|$. Hence

$$1 - \alpha_{k+1} = \frac{1 + \alpha_k}{\sqrt{1 + 4\alpha_k}}.$$

Linearizing the above equation, set $w_k := 4\alpha_k$ to obtain

$$w_{k+1} \approx w_k(1 - w_k) \quad \text{and then} \quad \frac{1}{w_{k+1}} - \frac{1}{w_k} = \frac{1}{1 - w_k}.$$

When summing and dividing by N, leads to

$$\begin{aligned}\lim_{N\to\infty} \frac{1}{Nw_N} &= \lim_{N\to\infty} \left(\frac{1}{Nw_N} - \frac{1}{Nw_0} \right) = \lim_{N\to\infty} \frac{1}{N} \sum_{k=0}^{N-1} \frac{1}{1 - w_k} \\ &= \lim_{N\to\infty} \frac{1}{1 - w_N} = 1 \quad (\text{since } (x_k, y_k) \longrightarrow 0).\end{aligned}$$

Then we have

$$\sqrt{x_k^2 + y_k^2} \sim \frac{1}{\sqrt{2k}}.$$

Hence $(x_k)_{k\in\mathbb{N}}$ and $(y_k)_{k\in\mathbb{N}}$ converge to 0 and at the rate of $k^{-\frac{1}{2}}$.

Remark 3.
According to Theorem 4, we can only deduce that $(x_k)_{k\in\mathbb{N}}$ and $(y_k)_{k\in\mathbb{N}}$ in Example 3 converge to $(0, 0)$ at the rate of at least of $k^{-\frac{1}{6}}$.

Example 4.
Let A, B be defined by

$$A := \{(x,y) \in \mathbb{R}^2 \mid x \le 0\} \quad \text{and} \quad B := \{(x,y) \in \mathbb{R}^2 \mid y^2 - x \le 0\}.$$

Let $b_0 \in \mathbb{R} \times \mathbb{R}_+$ with $\|b_0\| \le 1$, and $(a_k)_{k\in\mathbb{N}}$, $(b_k)_{k\in\mathbb{N}}$ be defined by

$$a_{k+1} := P_A b_k, \quad b_{k+1} := P_B a_{k+1}.$$

Then

$\boxed{(a_k)_{k\in\mathbb{N}} \text{ and } (b_k)_{k\in\mathbb{N}} \text{ converge to 0 at the rate of at exactly of } k^{-\frac{1}{2}}.}$

Remark 4.
According to Theorem 4, we can only deduce that $(a_k)_{k\in\mathbb{N}}$ and $(b_k)_{k\in\mathbb{N}}$ in Example 4 converge to $(0,0)$ at the rate of at least of $k^{-\frac{1}{6}}$.

Conclusion and questions

- Our explicit examples show that, in general, our estimate of the convergence rate of the cyclic projection algorithm will not be tight. It would be interesting to see how one can sharpen the estimate obtained in this talk and get a tight estimate for the cyclic projection algorithm.
- Can we extend the approach here to analyze the convergence rate of the Douglas-Rachford algorithm?

References

- H.H. Bauschke and P.L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, Springer, 2011.
- H.H. Bauschke and J.M. Borwein, “On the convergence of von Neumann’s alternating projection algorithm for two sets", *Set-Valued Analysis*, vol. 1, pp. 185–212, 1993.
- H.H. Bauschke and J.M. Borwein: “On projection algorithms for solving convex feasibility problems", *SIAM Review*, vol. 38, pp. 367-426, 1996.
- J.M. Borwein, G. Li, and L. Yao, “Analysis of the convergence rate for the cyclic projection algorithm applied to semi-algebraic convex sets”, submitted; `http://arxiv.org/abs/1304.7965v1`, April 2013.

References

- L.M. Brègman, Finding the common point of convex sets by the method of successive projection. *Doklady Akademii Nauk SSSR* , vol. 162, pp. 487–490, 1965.
- L.G. Gubin; B.T. Polyak, and E.V. Raik, The method of projections for finding the common point of convex sets, *USSR Computational Mathematics and Mathematical Physics*, vol. 7, pp. 1–24, 1967.
- J. Kollár, An effective Łojasiewicz inequality for real polynomials, *Periodica Mathematica Hungarica*, vol. 38, pp. 213–221, 1999.
- G. Li, "On the asymptotic well behaved functions and global error bound for convex polynomials", *SIAM Journal on Optimization*, vol. 20, pp. 1923–1943, 2010.

Thanks for your attention.