Douglas-Rachford splitting for nonconvex feasibility problems

Ting Kei Pong
Assistant Professor
Department of Applied Mathematics
The Hong Kong Polytechnic University
Hong Kong

The 3rd Workshop on Optimization and Risk Management
October 2014
(Joint work with Guoyin Li)

Feasibility Problem

- Given closed sets D_i, $i = 1, \ldots, m$, find a point

$$x \in \bigcap_{i=1}^{m} D_i.$$

- Example: Finding a solution of $Ax = b$ with $\|x\|_0 \leq r$.

Feasibility Problem

- Given closed sets D_i, $i = 1, \ldots, m$, find a point

$$x \in \bigcap_{i=1}^{m} D_i.$$

- Example: Finding a solution of $Ax = b$ with $\|x\|_0 \leq r$.
- The general problem can be reformulated as finding a point in

$$\{(x_1, \ldots, x_m) : \ x_1 = \cdots = x_m\} \cap (D_1 \times D_2 \times \cdots \times D_m).$$

- Only need to consider the intersection of a closed *convex* set C and a closed set D.

When D is convex

- Alternating projection:

$$x^{t+1} = P_D(P_C(x^t)).$$

- Douglas-Rachford (DR) splitting:

$$\begin{cases} y^{t+1} = \arg\min_{y \in C} \left\{ \|y - x^t\| \right\}, \\ z^{t+1} = \arg\min_{z \in D} \left\{ \|2y^{t+1} - x^t - z\| \right\}, \\ x^{t+1} = x^t + (z^{t+1} - y^{t+1}). \end{cases}$$

When D is convex

- Alternating projection:

$$x^{t+1} = P_D(P_C(x^t)).$$

- Douglas-Rachford (DR) splitting:

$$\begin{cases} y^{t+1} = \arg\min_{y \in C} \left\{ \|y - x^t\| \right\}, \\ z^{t+1} = \arg\min_{z \in D} \left\{ \|2y^{t+1} - x^t - z\| \right\}, \\ x^{t+1} = x^t + (z^{t+1} - y^{t+1}). \end{cases}$$

Empirically, DR splitting is usually faster.

When D is nonconvex

For the convergence of DR splitting:

- Mainly local convergence results.
- Require various regularity conditions on the sets.
- Local convergence for finding intersection of $Ax = b$ and $\|x\|_0 \leq r$. (Hesse, Luke, Neumann '13).

When D is nonconvex

For the convergence of DR splitting:

- Mainly local convergence results.
- Require various regularity conditions on the sets.
- Local convergence for finding intersection of $Ax = b$ and $\|x\|_0 \leq r$. (Hesse, Luke, Neumann '13).
- Global convergence shown for the intersection of a circle and a straight line in $\mathbb{R}^2$. (Artacho, Borwein '12)

Our approach

- DR splitting: ($\gamma > 0$)

$$\begin{cases} y^{t+1} = \operatorname*{arg\,min}_y \left\{ \frac{1}{2} d_C^2(y) + \frac{1}{2\gamma}\|y - x^t\|^2 \right\}, \\ z^{t+1} \in \operatorname*{Arg\,min}_{z \in D} \left\{ \|2y^{t+1} - x^t - z\|^2 \right\}, \\ x^{t+1} = x^t + (z^{t+1} - y^{t+1}). \end{cases}$$

- The y-update is $\frac{1}{1+\gamma}(x^t + \gamma P_C(x^t))$.

Our approach

- DR splitting: $(\gamma > 0)$

$$\begin{cases} y^{t+1} = \arg\min_{y} \left\{ \frac{1}{2} d_C^2(y) + \frac{1}{2\gamma} \|y - x^t\|^2 \right\}, \\ z^{t+1} \in \operatorname{Arg}\min_{z \in D} \left\{ \|2y^{t+1} - x^t - z\|^2 \right\}, \\ x^{t+1} = x^t + (z^{t+1} - y^{t+1}). \end{cases}$$

- The y-update is $\frac{1}{1+\gamma}(x^t + \gamma P_C(x^t))$.
- DR splitting applied to minimizing $\frac{1}{2} d_C^2 + \delta_D$.

Convergence result I

Fact 1 (Li, P '14): [Global convergence]
Suppose that $0 < \gamma < \sqrt{\frac{3}{2}} - 1$, and either C or D is compact.
Then $\{(y^t, z^t, x^t)\}$ is bounded, and any cluster point (y^*, z^*, x^*) satisfies $z^* = y^*$. Moreover, y^* is a stationary point of

$$\min_{u \in D} \frac{1}{2} d_C^2(u),$$

i.e., $0 \in y^* - P_C(y^*) + N_D(y^*)$.

Convergence result I

Fact 1 (Li, P '14): [Global convergence]
Suppose that $0 < \gamma < \sqrt{\frac{3}{2}} - 1$, and either C or D is compact.
Then $\{(y^t, z^t, x^t)\}$ is bounded, and any cluster point (y^*, z^*, x^*) satisfies $z^* = y^*$. Moreover, y^* is a stationary point of

$$\min_{u \in D} \frac{1}{2} d_C^2(u),$$

i.e., $0 \in y^* - P_C(y^*) + N_D(y^*)$.

- Clearly, if $d_C(y^*) = 0$, then y^* solves the feasibility problem.

Convergence result II

Fact 2 (Li, P '14): [Convergence of the whole sequence]
Suppose that $0 < \gamma < \sqrt{\frac{3}{2}} - 1$, C and D are semi-algebraic, and one of them is compact.
Then $\{(y^t, z^t, x^t)\}$ is bounded, and is convergent to some (y^*, z^*, x^*) satisfying $z^* = y^*$, with y^* being a stationary point of the problem $\min_{u \in D} \frac{1}{2} d_C^2(u)$. Furthermore,

$$\sum_{t=1}^{\infty} \|y^{t+1} - y^t\| < \infty.$$

Convergence result III

Fact 3 (Li, P '14): [Local convergence]
Let $C = \{x : Ax = b\}$ and D be a closed semi-algebraic set, $0 < \gamma < \sqrt{\frac{3}{2}} - 1$ and $\lim(y^t, z^t, x^t) = (y^*, z^*, x^*)$.
Suppose that $z^* \in C \cap D$ with

$$N_C(z^*) \cap -N_D(z^*) = \{0\}.$$

Then there exist $\eta \in (0, 1)$ and $\kappa > 0$ such that for all large t,

$$\mathrm{dist}\big(0, z^t - P_C(z^t) + N_D(z^t)\big) \leq \kappa\, \eta^t.$$

Convergence proof?

- KEY: Makes use of

$$\mathfrak{D}_\gamma(y,z,x) := \frac{1}{2}d_C^2(y) + \delta_D(z) + \frac{1}{2\gamma}\|x-y\|^2 - \frac{1}{2\gamma}\|x-z\|^2.$$

- Can show that for some $k_1, k_2 > 0$:

$$\mathfrak{D}_\gamma(y^t,z^t,x^t) - \mathfrak{D}_\gamma(y^{t+1},z^{t+1},x^{t+1}) \geq k_1\|y^{t+1}-y^t\|^2;$$
$$\mathrm{dist}(0,\partial\mathfrak{D}_\gamma(y^t,z^t,x^t)) \leq k_2\|y^{t+1}-y^t\|.$$

Numerical simulations

- Find a point in $Ax = b$ with $\|x\|_0 \leq r$.
- Consider random instances: generate an r-sparse vector $\tilde{x}$, an $m \times n$ matrix A, and set $b = A\tilde{x}$.
- Compare with alternating projection. Initialize both algorithms at $x^0 = 0$.
- Terminate when successive changes are less than 10^{-8}.

Numerical simulations

- Find a point in $Ax = b$ with $\|x\|_0 \leq r$.
- Consider random instances: generate an r-sparse vector $\tilde{x}$, an $m \times n$ matrix A, and set $b = A\tilde{x}$.
- Compare with alternating projection. Initialize both algorithms at $x^0 = 0$.
- Terminate when successive changes are less than 10^{-8}.
- For DR splitting, start with a $\gamma > \sqrt{\frac{3}{2}} - 1$, decrease γ if $\|y^{t+1} - y^t\|$ does not deteriorate quickly enough.

Numerical simulations

Over 50 trials for each m, n; sparsity is $\lceil \frac{m}{5} \rceil$; succ means $\mathrm{fval} < 10^{-12}$.

Data		DR: $fval = \frac{1}{2}d_C^2(z^t)$				Alt Proj: $fval = \frac{1}{2}d_C^2(x^t)$			
m	n	iter	$\mathrm{fval}_{\mathrm{max}}$	$\mathrm{fval}_{\mathrm{min}}$	succ	iter	$\mathrm{fval}_{\mathrm{max}}$	$\mathrm{fval}_{\mathrm{min}}$	succ
300	4000	600	3e-15	2e-16	50	872	4e-01	6e-14	3
300	5000	710	4e-15	4e-16	50	1068	3e-01	9e-14	3
300	6000	812	3e-15	2e-16	50	1252	3e-01	1e-13	1
400	4000	520	2e-15	3e-17	50	818	6e-01	8e-14	30
400	5000	579	3e-15	5e-16	50	946	4e-01	9e-14	12
400	6000	646	4e-15	6e-16	50	1108	3e-01	1e-13	4
500	4000	499	1e-16	1e-18	50	640	4e-01	6e-14	38
500	5000	519	1e-15	4e-17	50	846	4e-01	9e-14	37
500	6000	556	3e-15	3e-16	50	1071	5e-01	1e-13	22

Conclusion

- The DR splitting applied to $\min_{u \in D} \frac{1}{2} d_C^2(u)$, with either C or D being compact, can be shown to generate a bounded sequence that clusters at a stationary point.
- Under semi-algebraicity assumption, the whole sequence can be shown to be convergent.

Reference:
G. Li and T. K. Pong.
Douglas-Rachford splitting for nonconvex feasibility problems.
Available at *http://arxiv.org/abs/1409.8444*.

Thanks for coming! ☺