Lecture 17

Last time

- Maximal monotone
- Douglas Rachford
- Augmented Lagrangian

Today

- Douglas-Rachford
- Concensus optimization

Douglas-Rachford

The method that we will see next dates back to the 1950's and came from ideas in differential equations and infinite dimensional problems.

Recall that given a monotone operator $A: E \rightrightarrows E$, we defined

$$R_A = (I + \alpha A)^{-1} \quad \text{and} \quad C_A = 2R_A - I.$$

Moreover, the fixed points of R_A and C_A correspond to the zeros of A, i.e.,

$\{x \mid 0 \in Ax\}.$

In turn, sometimes we are interested in finding a zero of a sum of operators maximal monotone operators, i.e., $x \in E$ s.t.

$$0 \in (A+B)x. \qquad (\ddot{\smile})$$

However, computing a resolvent of $A+B$ might be expensive.

Q: Is it possible to solve $(\ddot{\smile})$ using only R_A and R_B?

To answer this question Douglas–Rachford asked themselves the marvelous question of what are fixed points of $C_A \circ C_B$:

$$y \in C_A C_B y \iff y = 2R_A(2R_By - y) - (2R_By - y)$$

$$\iff R_By = R_A(2R_By - y)$$

$(\ast)$ $\quad x = R_By$

$$\iff x = R_A(2x - y)$$

$$\iff (2x - y) \in x + Ax$$

$$\iff (x - y) \in Ax.$$

Moreover, by definition

$$x = R_By \iff y - x \in Bx.$$

Proposition (M): Let $A, B : E \rightrightarrows E$ be maximal monotone operators. Then, the zeros of $A+B$ are exactly the images under R_B of fixed points $C_A \circ C_B$.

Proof: Note that

$$0 \in (A+B)x \iff \exists z \in Ax \text{ with } -z \in Bx$$
$$\iff \exists y \text{ s.t. } x-y \in Ax, \quad y-x \in Bx.$$

The rest of the argument follows by ($\lightning$). $\square$

Thus, we have transformed the problem into that of looking for a fixed point of $C_A \circ C_B$.
We could aim to apply the fixed point iteration directly

$$y_{k+1} \leftarrow C_A C_B y_k.$$

This method is known as Peaceman-Rachford (PR). As we know this iteration might fail to converge since $C_A C_B$ is merely nonexpansive (See lecture 15).

Instead Douglas - Rachford propo-
sed to use the averaged
update

$$y_{k+1} \leftarrow \frac{1}{2}\left(I + C_A \circ C_B\right) y_k,$$

which can be expanded as:

Douglas - Rachford Method (DRM)

Input: Initial y_0 and resolvents R_A, R_B.

Loop $k \geq 0$:

- $x_{k+1} \leftarrow R_B \, y_{k+1}$

Auxiliary statements

- $y' \leftarrow 2x_{k+1} - y_k$
- $\hat{x}_{k+1} \leftarrow R_A \, y'$
- $y'' \leftarrow 2\hat{x}_{k+1} - y' = 2\hat{x}_{k+1} - 2x_{k+1} + y_k$
- $y_{k+1} \leftarrow y_k + \hat{x}_{k+1} - x_{k+1} = \frac{1}{2}(y_k + y'')$

Remark: This algorithm does not "interact" with $A+B$ directly, instead it "splits" the computation into applications of R_A and R_B. $\dashv$

Theorem (♡) Assume that A, B are maximal monotone operators and $A+B$ has a zero. Then any sequence generated $\{y_k\}$ by DRM converges to some y^* and $x^* = R_B y^*$ satisfies $0 \in (A+B)x^*$.

Proof: This follows immediately from our KM iteration result and Proposition (11). $\square$

Concensus optimization

Let us see an important application of DRM. Suppose we are interested in minimizing

$$\inf_{x \in E} \sum_{i=1}^{k} f_i(x) \qquad (\diamond)$$

where each $f_i : E \to \mathbb{R} \cup \{+\infty\}$ is close, convex, and proper. Think of each f_i as being relatively simple in that we can compute their prox operator efficiently.

To split this problem we consider the formulation

$$\inf_{x_i \in E \ \forall i} f(x_1, \ldots, x_k) + \iota_L(x_1, \ldots, x_k)$$

where

$$f(x_1, \ldots, x_k) = \sum_{i=1}^{k} f_i(x_i) \quad \text{and}$$

$$L = \{(x_1, \ldots, x_k) \in E^k \mid x_1 = x_2 = \ldots = x_k\}.$$

Thus we want to solve

$$0 \in (\partial f + \partial \iota_L)(x_1, \dots, x_k).$$

Assuming $\bigcap_i \operatorname{int} \operatorname{dom} f_i \neq \emptyset$, we have that

$$\partial\left(\sum f_i\right)(x) = \sum \partial f_i(x).$$

Therefore,

$$x \text{ solves } (\Diamond) \iff 0 \in \partial\left(\sum f_i\right)(x)$$

$$\iff \forall i \in [k]\ \exists z_i \in \partial f_i(x)$$
$$\sum z_i = 0$$

$$\iff (z_1, \dots, z_k) \in \partial f(\underbrace{x, \dots, x}_{\in L})$$
$$\sum z_i = 0$$

Exercise

Claim: $-(z_1, \dots, z_k) \in \iota_L(x, \dots, x) = L^\perp$ iff $\sum_{i=1}^{k} z_i = 0.$ $\dashv$

$$\iff (z_1, \dots, z_k) \in \partial f(x, \dots, x)$$
and
$$-(z_1, \dots, z_k) \in \partial \iota_L(x, \dots, x)$$

Thus we could apply DRM with $A = \partial \iota_L$ and $B = \partial f$.
Notice that

$$R_B(y_1, \dots, y_K) = \begin{pmatrix} \mathrm{prox}_{\alpha f_1}(y_1) \\ \vdots \\ \mathrm{prox}_{\alpha f_K}(y_K) \end{pmatrix}$$

The beauty of splitting the variables.

$$R_A(y_1, \dots, y_K) = \mathrm{proj}_L(y_1, \dots, y_K)$$

Check!

$$= \frac{1}{K}\left(\sum_{i=1}^{K} y_i, \dots, \sum_{i=1}^{K} y_i\right).$$

For simplicity let $\bar{y} = \frac{1}{K}\sum_{i=1}^{K} y_i$.

Concensus Optimization via DR

Input: $y_0 \in E$ and maps $\mathrm{prox}_{\alpha f_i}$

Loop $k \geq 0$:

▷ $x_{k+1}^{(i)} \leftarrow \mathrm{prox}_{\alpha f_i}(y_k^{(i)}) \quad \forall i \in [K]$

$$\triangleright\ y^{(i)}_{k+1} \leftarrow y^{(i)}_k + 2\bar{x}_{k+1} - \bar{y}_k - x^{(i)}_{k+1}$$
$$\forall i \in [K].$$

By Theorem (♡) this algorithm generates a sequence $\bar{y}_k \to y^*$ s.t. $x^* = R_B y^*$ is a minimizer of (◇).

Remark: We can apply the same ideas to solve for $0 \in \sum_{i=1}^{K} A_i(x)$.