Last time we showed

\[ x^* \in \text{argmin} \{ c^\top x : Ax = b, x \geq 0 \} \]

\[ (y^*, s^*) \in \text{argmax} \{ b^\top y : A^\top y + s = c, s \geq 0 \} \]

if, and only if

\[ Ax^* = b, A^\top y^* + s^* = c, c^\top x^* - b^\top y^* = 0, x^* \geq 0, s^* \geq 0, \]

i.e.,

\[
\begin{bmatrix}
A & 0 & 0 \\
0 & A^\top & 1 \\
c^\top & -b^\top & 0
\end{bmatrix}
\begin{bmatrix}
x^* \\
y^* \\
s^*
\end{bmatrix}
= \begin{bmatrix}
b \\
c \\
0
\end{bmatrix}
\]

and

\[
\begin{bmatrix}
x^* \\
0 \\
s^*
\end{bmatrix} \geq 0,
\]

\[
\begin{bmatrix}
x^* \\
y^* \\
s^*
\end{bmatrix} \in C_1 \cap C_2.
\]

- We showed that whenever \( C_1 \cap C_2 \neq \emptyset \) and \( z^0 \in \mathbb{R}^{2n+m} \), the sequence \( z^{k+1} = P_{C_2}P_{C_1}(z^k) \) converges to an element of \( C_1 \cap C_2 \).
- This is called the method of alternating projections (MAP).
- There are many algorithms for solving feasibility problems.
- Today we learn the Douglas-Rachford splitting (DRS) method. Example:
Unlike MAP, this method composes two reflections, not two projections.

DRS was developed in the 1950s, first for solving PDEs.

Later extended to feasibility problems and optimization problems.

Sometimes it is called the Alternating Directions Method of Multipliers, or ADMM.

In general DRS performs better than MAP, theoretically and practically.

Cost of applying MAP and DRS is exactly the same.

By KM iteration theorem DRS converges.

**Theorem.** Let $C_1, C_2 \subset \mathbb{R}^n$ be two closed convex sets. Given $z^0 \in \mathbb{R}^n$ the DRS sequence

$$(\forall k \in \mathbb{N}) \quad z^{k+1} = \frac{1}{2}(2P_{C_2} - I) \circ (2P_{C_1} - I)(z^k) + \frac{1}{2}z^k$$

$$= z^k - P_{C_1}(z^k) + P_{C_2}(2P_{C_1}(z^k) - z^k)$$

converges to a fixed point of the operator $(2P_{C_2} - I) \circ (2P_{C_1} - I)$, whenever one exists.

**Proof.** Last lecture we showed that $2P_{C_2} - I$ and $2P_{C_1} - I$ are nonexpansive. Thus, because it is the composition of nonexpansive mappings, the map $N = (2P_{C_2} - I) \circ (2P_{C_1} - I)$ is nonexpansive. Thus, as the DRS sequence is generated by iterating the mapping $N_{1/2}$, it must converge by the KM convergence theorem—provided that $\text{Fix}(N) \neq \emptyset$. In that case, its limit must be a fixed point of $N$. \qed
So the important question to answer is whether $(2P_{C_2} - I) \circ (2P_{C_1} - I)$ has a fixed point and what properties it might satisfy. First, let’s rewrite the DRS algorithm with intermediate variables.

\[
\begin{align*}
x_{C_1} &= P_{C_1}z \\
x_{C_2} &= P_{C_2}(2x_{C_1} - z) (= P_{C_2} \circ (2P_{C_1} - I)(z)) \\
z^+ &= z + (x_{C_2} - x_{C_1}) \\
     &= z - P_{C_1}z + P_{C_2}(2P_{C_1}z - z).
\end{align*}
\]

In pictures,

![Diagram](image)

**Theorem.** Fix$((2P_{C_2} - I) \circ (2P_{C_1} - I)) = \{x + g : x \in C_1 \cap C_2, g \in N_{C_1}(x) \cap (-N_{C_2}(x))\}$. Thus, if $z$ is a fixed point, then $P_{C_1}(z) \in C_1 \cap C_2$.

**Proof.** ($\subset$) Let $z$ be a fixed point. Then

\[
z = z^+ = z + (x_{C_2} - x_{C_1}).
\]

So $x_{C_1} = x_{C_2}$. Let $x := x_{C_1}$. Thus, $z - x \in N_{C_1}(x)$ and $x - z = (2x - z) - x \in N_{C_2}(x)$. Set $g = z - x$. Then $z = x + g$ with $g \in N_{C_1}(x) \cap (-N_{C_2}(x))$.  

17-3
Consider $z = x + g$ where $g \in N_{C_1}(x) \cap (-N_{C_2}(x))$. Then $(x + g) - x = g \in N_{C_1}(x)$ so $x = x_{C_1} = P_{C_1}(x + g)$. Also, $x_{C_2} = P_{C_2}(2x - (x + g)) = P_{C_2}(x - g)$ and $(x - g) - x = -g \in N_{C_2}(x)$ so $x = x_{C_2}$. Thus,

$$z^+ = z + (x_{C_2} - x_{C_1}) = z,$$

i.e., $z = x + g$ is a fixed point. \hfill \Box

The above theorem shows that, $\text{Fix}((2P_{C_2} - I) \circ (2P_{C_1} - I))$ is nonempty if and only if $C_1 \cap C_2 \neq \emptyset$, since we can always choose $g = 0$. In addition, limit points of the DRS algorithm recover elements $P_{C_1}(x) \in C_1 \cap C_2$.

In general, we can show that:

**Lemma.** For all fixed points $z^*$ we have

$$\forall k \in \mathbb{N} \quad \left\| x_{C_1}^{k+1} - x_{C_2}^{k+1} \right\| \leq \left\| x_{C_1}^{k} - x_{C_2}^{k} \right\|$$

$$\left\| x_{C_1}^{k} - x_{C_2}^{k} \right\| = o\left(\frac{\|z^0 - z^*\|}{\sqrt{k + 1}}\right)$$

which shows that

$$d_{C_2}(x_{C_1}^{k}) \leq \left\| x_{C_1}^{k} - x_{C_2}^{k} \right\| = o\left(\frac{\|z^0 - z^*\|}{\sqrt{k + 1}}\right),$$

$$d_{C_1}(x_{C_2}^{k}) \leq \left\| x_{C_1}^{k} - x_{C_2}^{k} \right\| = o\left(\frac{\|z^0 - z^*\|}{\sqrt{k + 1}}\right).$$

We can even improve this rate:

**Lemma.** Let $z^*$ be a fixed point of the DRS operator. For all $k \in \mathbb{N}$, let

$$\bar{x}_{C_1}^k = \frac{1}{k + 1} \sum_{i=0}^{k} x_{C_1}^i,$$

$$\bar{x}_{C_2}^k = \frac{1}{k + 1} \sum_{i=0}^{k} x_{C_2}^i.$$

Then $\bar{x}_{C_1}^k \in C_1, \bar{x}_{C_2}^k \in C_2$ and

$$d_{C_2}(\bar{x}_{C_1}^k) \leq \left\| \bar{x}_{C_1}^k - \bar{x}_{C_2}^k \right\| = O\left(\frac{\|z^0 - z^*\|}{k + 1}\right),$$

$$d_{C_1}(\bar{x}_{C_2}^k) \leq \left\| \bar{x}_{C_1}^k - \bar{x}_{C_2}^k \right\| = O\left(\frac{\|z^0 - z^*\|}{k + 1}\right).$$

17-4
Proof. Because \( z^{k+1} = N(z^k) \) for a nonexpansive operator \( N \), we have
\[
(\forall k \in \mathbb{N}) \quad \|z^{k+1} - z^*\| = \|N(z^k) - N(z^*)\| \leq \|z^k - z^*\| \leq \cdots \leq \|z^0 - z^*\|.
\]
Furthermore, because \( z^{k+1} = z^k + (x^{k}_{C_2} - x^{k}_{C_1}) \) we have
\[
\begin{align*}
    \|\bar{x}^k_{C_1} - \bar{x}^k_{C_2}\| &= \frac{1}{k+1} \left\| \sum_{i=0}^{k} (x^i_{C_1} - x^i_{C_2}) \right\| \\
    &= \frac{1}{k+1} \left\| \sum_{i=0}^{k} (z^i - z^{i+1}) \right\| \\
    &= \frac{1}{k+1} \|z^0 - z^{k+1}\| \\
    &\leq \frac{1}{k+1} \left( \|z^0 - z^*\| + \|z^* - z^{k+1}\| \right) \\
    &= \frac{2\|z^0 - z^*\|}{k+1}.
\end{align*}
\]
The distance inequalities follow because \( \bar{x}^k_{C_1}, \bar{x}^k_{C_2} \in C_1 \). \( \square \)

In general, the DRS method can converge arbitrarily slowly.

**Theorem (II).** For any \( h : \mathbb{R}_+ \to (0, 1) \) that is strictly decreasing to zero, there exist closed convex sets such that \( C_1 \cap C_2 = \{0\} \) and \( \|z^k - z^*\| \geq h(k)/e \).

However, for linear programs formulated as they are at the start of the lecture, the story is different.

**Theorem.** For the linear programming feasibility problem, the Douglas-Rachford algorithm satisfies
\[
(\forall k \in \mathbb{N}) \quad d_{C_1 \cap C_2}(z^{k+1}) \leq \delta d_{C_1 \cap C_2}(z^k)
\]
for some \( \delta \in (0, 1) \).

**References**