Alternating Linear Minimization: Revisiting von Neumann's alternating projections

Sebastian Pokutta

joint work with: Gábor Braun and Robert Weismantel

Technische Universität Berlin
and
Zuse Institute Berlin

pokutta@math.tu-berlin.de
@spokutta

Halfway to Twelfty

July 4th, 2023 · Paris, France

Berlin Mathematics Research Center

What is this talk about?

Introduction

Given P, Q compact convex sets,
does there exist $x \in P \cap Q$?

What is this talk about?
Introduction

Given P, Q compact convex sets,
does there exist $x \in P \cap Q$?

Why? At the core of many algorithms. Allows for optimization via binary search.

What is this talk about?

Introduction

Given P, Q compact convex sets,
does there exist $x \in P \cap Q$?

Why? At the core of many algorithms. Allows for optimization via binary search.

Today: von Neumann's approach and a new algorithm.

(Hyperlinked) References are not exhaustive; check references contained therein.

Some trivial insights...

Polytopes: H-representation and V-representation

Some trivial insights...

Example. (H-representation)

Let $P = \{x \mid A_P x \leq b_P\}$ and $Q = \{x \mid A_Q x \leq b_Q\}$ be polytopes. Then $x \in P \cap Q$?

Polytopes: H-representation and V-representation

Some trivial insights...

Example. (H-representation)

Let $P = \{x \mid A_P x \leq b_P\}$ and $Q = \{x \mid A_Q x \leq b_Q\}$ be polytopes. Then $x \in P \cap Q$?

Solution: Linear programming! Check feasibility of

$$P \cap Q = \{x \mid A_P x \leq b_P, A_Q x \leq b_Q\}.$$

Polytopes: H-representation and V-representation

Some trivial insights...

Example. (H-representation)

Let $P = \{x \mid A_P x \leq b_P\}$ and $Q = \{x \mid A_Q x \leq b_Q\}$ be polytopes. Then $x \in P \cap Q$?

Solution: Linear programming! Check feasibility of

$$P \cap Q = \{x \mid A_P x \leq b_P, A_Q x \leq b_Q\}.$$

Example. (V-representation)

Let $P = \operatorname{conv}(U)$ and $Q = \operatorname{conv}(W)$ be polytopes. Then $x \in P \cap Q$?

Polytopes: H-representation and V-representation

Some trivial insights...

Example. (H-representation)

Let $P = \{x \mid A_P x \leq b_P\}$ and $Q = \{x \mid A_Q x \leq b_Q\}$ be polytopes. Then $x \in P \cap Q$?

Solution: Linear programming! Check feasibility of

$$P \cap Q = \{x \mid A_P x \leq b_P, A_Q x \leq b_Q\}.$$

Example. (V-representation)

Let $P = \operatorname{conv}(U)$ and $Q = \operatorname{conv}(W)$ be polytopes. Then $x \in P \cap Q$?

Solution: Linear programming! Check feasibility of

$$\left\{(\lambda, \kappa) : \sum_{u \in U} \lambda_u u = \sum_{w \in W} \kappa_w w, \sum_{u \in U} \lambda_u = \sum_{w \in W} \kappa_w = 1, \lambda, \kappa \geq 0\right\}.$$

More general setup

Some trivial insights...

What if access to P and Q is only given implicitly?

More general setup

Some trivial insights...

What if access to P and Q is only given implicitly?

What if P and Q are more general, e.g., compact convex?

von Neumann's Alternating Projections

The algorithm
von Neumann's Alternating Projections

Let P and Q be compact convex sets. Π_P, Π_Q being the respective projectors.

The algorithm

von Neumann's Alternating Projections

Let P and Q be compact convex sets. Π_P, Π_Q being the respective projectors.

Algorithm von Neumann's Alternating Projections (POCS)

Input: Point $y_0 \in \mathbb{R}^n$, Π_P projector onto $P \subseteq \mathbb{R}^n$ and Π_Q projector onto $Q \subseteq \mathbb{R}^n$.
Output: Iterates $x_1, y_1 \ldots \in \mathbb{R}^n$

1: **for** $t = 0$ **to** $\ldots$ **do**
2: $\quad x_{t+1} \leftarrow \Pi_P(y_t)$
3: $\quad y_{t+1} \leftarrow \Pi_Q(x_{t+1})$

appeared in lecture notes first distributed in 1933; see reprint [von Neumann, 1949]

Convergence

von Neumann's Alternating Projections

Suppose $P \cap Q \neq \emptyset$ and let $u \in P \cap Q$. The binomial formula is your friend:

$$\|y_t - u\|^2$$

Convergence

von Neumann's Alternating Projections

Suppose $P \cap Q \neq \emptyset$ and let $u \in P \cap Q$. The binomial formula is your friend:

$$\|y_t - u\|^2 = \|y_t - x_{t+1} + x_{t+1} - u\|^2 = \|y_t - x_{t+1}\|^2 + \|x_{t+1} - u\|^2 - 2\langle x_{t+1} - y_t, x_{t+1} - u\rangle$$

Convergence

von Neumann's Alternating Projections

Suppose $P \cap Q \neq \emptyset$ and let $u \in P \cap Q$. The binomial formula is your friend:

$$\|y_t - u\|^2 = \|y_t - x_{t+1} + x_{t+1} - u\|^2 = \|y_t - x_{t+1}\|^2 + \|x_{t+1} - u\|^2 - 2 \underbrace{\langle x_{t+1} - y_t, x_{t+1} - u \rangle}_{\leq 0}$$

Convergence

von Neumann's Alternating Projections

Suppose $P \cap Q \neq \emptyset$ and let $u \in P \cap Q$. The binomial formula is your friend:

$$\|y_t - u\|^2 = \|y_t - x_{t+1} + x_{t+1} - u\|^2 = \|y_t - x_{t+1}\|^2 + \|x_{t+1} - u\|^2 - 2 \underbrace{\langle x_{t+1} - y_t, x_{t+1} - u \rangle}_{\leq 0}$$

$$\geq \|y_t - x_{t+1}\|^2 + \|x_{t+1} - u\|^2$$

Convergence

von Neumann's Alternating Projections

Suppose $P \cap Q \neq \emptyset$ and let $u \in P \cap Q$. The binomial formula is your friend:

$$\|y_t - u\|^2 = \|y_t - x_{t+1} + x_{t+1} - u\|^2 = \|y_t - x_{t+1}\|^2 + \|x_{t+1} - u\|^2 - 2 \underbrace{\langle x_{t+1} - y_t, x_{t+1} - u \rangle}_{\leq 0}$$

$$\geq \|y_t - x_{t+1}\|^2 + \|x_{t+1} - u\|^2 = \|y_t - x_{t+1}\|^2 + \|x_{t+1} - y_{t+1} + y_{t+1} - u\|^2$$

Convergence

von Neumann's Alternating Projections

Suppose $P \cap Q \neq \emptyset$ and let $u \in P \cap Q$. The binomial formula is your friend:

$$\|y_t - u\|^2 = \|y_t - x_{t+1} + x_{t+1} - u\|^2 = \|y_t - x_{t+1}\|^2 + \|x_{t+1} - u\|^2 - 2 \underbrace{\langle x_{t+1} - y_t, x_{t+1} - u \rangle}_{\leq 0}$$

$$\geq \|y_t - x_{t+1}\|^2 + \|x_{t+1} - u\|^2 = \|y_t - x_{t+1}\|^2 + \|x_{t+1} - y_{t+1} + y_{t+1} - u\|^2$$

$$= \|y_t - x_{t+1}\|^2 + \|x_{t+1} - y_{t+1}\|^2 + \|y_{t+1} - u\|^2 - 2 \langle y_{t+1} - x_{t+1}, y_{t+1} - u \rangle$$

Convergence

von Neumann's Alternating Projections

Suppose $P \cap Q \neq \emptyset$ and let $u \in P \cap Q$. The binomial formula is your friend:

$$\|y_t - u\|^2 = \|y_t - x_{t+1} + x_{t+1} - u\|^2 = \|y_t - x_{t+1}\|^2 + \|x_{t+1} - u\|^2 - 2 \underbrace{\langle x_{t+1} - y_t, x_{t+1} - u \rangle}_{\leq 0}$$

$$\geq \|y_t - x_{t+1}\|^2 + \|x_{t+1} - u\|^2 = \|y_t - x_{t+1}\|^2 + \|x_{t+1} - y_{t+1} + y_{t+1} - u\|^2$$

$$= \|y_t - x_{t+1}\|^2 + \|x_{t+1} - y_{t+1}\|^2 + \|y_{t+1} - u\|^2 - 2 \underbrace{\langle y_{t+1} - x_{t+1}, y_{t+1} - u \rangle}_{\leq 0}$$

$$\geq \|y_t - x_{t+1}\|^2 + \|x_{t+1} - y_{t+1}\|^2 + \|y_{t+1} - u\|^2.$$

Convergence

von Neumann's Alternating Projections

Suppose $P \cap Q \neq \emptyset$ and let $u \in P \cap Q$. The binomial formula is your friend:

$$\begin{aligned}
\|y_t - u\|^2 &= \|y_t - x_{t+1} + x_{t+1} - u\|^2 = \|y_t - x_{t+1}\|^2 + \|x_{t+1} - u\|^2 - 2 \underbrace{\langle x_{t+1} - y_t, x_{t+1} - u \rangle}_{\leq 0} \\
&\geq \|y_t - x_{t+1}\|^2 + \|x_{t+1} - u\|^2 = \|y_t - x_{t+1}\|^2 + \|x_{t+1} - y_{t+1} + y_{t+1} - u\|^2 \\
&= \|y_t - x_{t+1}\|^2 + \|x_{t+1} - y_{t+1}\|^2 + \|y_{t+1} - u\|^2 - 2 \underbrace{\langle y_{t+1} - x_{t+1}, y_{t+1} - u \rangle}_{\leq 0} \\
&\geq \|y_t - x_{t+1}\|^2 + \|x_{t+1} - y_{t+1}\|^2 + \|y_{t+1} - u\|^2.
\end{aligned}$$

Rearrange to

$$\|y_t - u\|^2 - \|y_{t+1} - u\|^2 \geq \|y_t - x_{t+1}\|^2 + \|x_{t+1} - y_{t+1}\|^2.$$

Convergence

von Neumann's Alternating Projections

Suppose $P \cap Q \neq \emptyset$ and let $u \in P \cap Q$. The binomial formula is your friend:

$$\begin{aligned}
\|y_t - u\|^2 &= \|y_t - x_{t+1} + x_{t+1} - u\|^2 = \|y_t - x_{t+1}\|^2 + \|x_{t+1} - u\|^2 - 2 \underbrace{\langle x_{t+1} - y_t, x_{t+1} - u \rangle}_{\leq 0} \\
&\geq \|y_t - x_{t+1}\|^2 + \|x_{t+1} - u\|^2 = \|y_t - x_{t+1}\|^2 + \|x_{t+1} - y_{t+1} + y_{t+1} - u\|^2 \\
&= \|y_t - x_{t+1}\|^2 + \|x_{t+1} - y_{t+1}\|^2 + \|y_{t+1} - u\|^2 - 2 \underbrace{\langle y_{t+1} - x_{t+1}, y_{t+1} - u \rangle}_{\leq 0} \\
&\geq \|y_t - x_{t+1}\|^2 + \|x_{t+1} - y_{t+1}\|^2 + \|y_{t+1} - u\|^2.
\end{aligned}$$

Rearrange to

$$\|y_t - u\|^2 - \|y_{t+1} - u\|^2 \geq \|y_t - x_{t+1}\|^2 + \|x_{t+1} - y_{t+1}\|^2.$$

Whenever you see something like this, it is checkmate in 3 moves...

Convergence

von Neumann's Alternating Projections

Starting from

$$\|y_t - u\|^2 - \|y_{t+1} - u\|^2 \geq \|y_t - x_{t+1}\|^2 + \|x_{t+1} - y_{t+1}\|^2.$$

Convergence

von Neumann's Alternating Projections

Starting from

$$\|y_t - u\|^2 - \|y_{t+1} - u\|^2 \geq \|y_t - x_{t+1}\|^2 + \|x_{t+1} - y_{t+1}\|^2.$$

1) Simply sum up

$$\sum_{t=0,\ldots,T-1} \left(\|y_t - u\|^2 - \|y_{t+1} - u\|^2\right) \geq \sum_{t=0,\ldots,T-1} \left(\|y_t - x_{t+1}\|^2 + \|x_{t+1} - y_{t+1}\|^2\right).$$

Convergence

von Neumann's Alternating Projections

Starting from

$$\|y_t - u\|^2 - \|y_{t+1} - u\|^2 \geq \|y_t - x_{t+1}\|^2 + \|x_{t+1} - y_{t+1}\|^2.$$

1) Simply sum up

$$\sum_{t=0,\ldots,T-1} \left(\|y_t - u\|^2 - \|y_{t+1} - u\|^2 \right) \geq \sum_{t=0,\ldots,T-1} \left(\|y_t - x_{t+1}\|^2 + \|x_{t+1} - y_{t+1}\|^2 \right).$$

2) which implies, via telescoping,

$$\|y_0 - u\|^2 \geq \sum_{t=0,\ldots,T-1} \left(\|y_t - x_{t+1}\|^2 + \|x_{t+1} - y_{t+1}\|^2 \right).$$

Convergence

von Neumann's Alternating Projections

Starting from

$$\|y_t - u\|^2 - \|y_{t+1} - u\|^2 \geq \|y_t - x_{t+1}\|^2 + \|x_{t+1} - y_{t+1}\|^2.$$

1) Simply sum up

$$\sum_{t=0,\dots,T-1} \left(\|y_t - u\|^2 - \|y_{t+1} - u\|^2\right) \geq \sum_{t=0,\dots,T-1} \left(\|y_t - x_{t+1}\|^2 + \|x_{t+1} - y_{t+1}\|^2\right).$$

2) which implies, via telescoping,

$$\|y_0 - u\|^2 \geq \sum_{t=0,\dots,T-1} \left(\|y_t - x_{t+1}\|^2 + \|x_{t+1} - y_{t+1}\|^2\right).$$

3) divide by T, then

$$\frac{\|y_0 - u\|^2}{T} \geq \frac{1}{T} \sum_{t=0,\dots,T-1} \left(\|y_t - x_{t+1}\|^2 + \|x_{t+1} - y_{t+1}\|^2\right)$$

Convergence
von Neumann's Alternating Projections

Starting from

$$\|y_t - u\|^2 - \|y_{t+1} - u\|^2 \geq \|y_t - x_{t+1}\|^2 + \|x_{t+1} - y_{t+1}\|^2.$$

1) Simply sum up

$$\sum_{t=0,\dots,T-1} \left(\|y_t - u\|^2 - \|y_{t+1} - u\|^2 \right) \geq \sum_{t=0,\dots,T-1} \left(\|y_t - x_{t+1}\|^2 + \|x_{t+1} - y_{t+1}\|^2 \right).$$

2) which implies, via telescoping,

$$\|y_0 - u\|^2 \geq \sum_{t=0,\dots,T-1} \left(\|y_t - x_{t+1}\|^2 + \|x_{t+1} - y_{t+1}\|^2 \right).$$

3) divide by T, then

$$\frac{\|y_0 - u\|^2}{T} \geq \frac{1}{T} \sum_{t=0,\dots,T-1} \left(\|y_t - x_{t+1}\|^2 + \|x_{t+1} - y_{t+1}\|^2 \right) \geq \|x_T - y_T\|^2,$$

as distances are non-increasing. □

Convergence

von Neumann's Alternating Projections

Proposition (von Neumann [1949] + minor perturbations)

Let P and Q be compact convex sets with $P \cap Q \neq \emptyset$ and let $x_1, y_1 \ldots, x_T, y_T \in \mathbb{R}^n$ be the sequence of iterates of von Neumann's algorithm. Then the iterates converge: $x_t \longrightarrow x$ and $y_t \longrightarrow y$ to some $x \in P$ and $y \in Q$ and

$$\|x_T - y_T\|^2 \leq \frac{1}{T} \sum_{t=0}^{T-1} \left(\|y_t - x_{t+1}\|^2 + \|x_{t+1} - y_{t+1}\|^2 \right) \leq \frac{\operatorname{dist}(y_0, P \cap Q)^2}{T}.$$

Projections are often expensive however...

von Neumann's Alternating Projections

What if access to P and Q is only given by Linear Minimization Oracles (LMOs)?
(e.g., via combinatorial algorithm like matching algorithm)

Projections are often expensive however...

von Neumann's Alternating Projections

What if access to P and Q is only given by Linear Minimization Oracles (LMOs)?
(e.g., via combinatorial algorithm like matching algorithm)

Quick reminder. Linear minimization is often cheaper than projection (basically quadratic programming).

Alternating Linear Minimizations

After close inspection and some meditation,

von Neumann's algorithm revisited

Alternating Linear Minimizations

After close inspection and some meditation, von Neumann's algorithm basically solves

$$\min_{(x,y)\in P\times Q} \|x-y\|^2,$$

i.e., we are minimizing the 2-norm over the product space $P \times Q$.

After close inspection and some meditation, von Neumann's algorithm basically solves

$$\min_{(x,y)\in P\times Q} \|x-y\|^2,$$

i.e., we are minimizing the 2-norm over the product space $P \times Q$.

In principle. Any Frank-Wolfe algorithm to solve the problem (only LMOs for P and Q).

[Braun et al., 2022]

von Neumann's algorithm revisited

Alternating Linear Minimizations

After close inspection and some meditation, von Neumann's algorithm basically solves

$$\min_{(x,y)\in P\times Q} \|x-y\|^2,$$

i.e., we are minimizing the 2-norm over the product space $P \times Q$.

In principle. Any Frank-Wolfe algorithm to solve the problem (only LMOs for P and Q).

[Braun et al., 2022]

However. We want von Neumann style algorithm with alternations.

(Note. Above formulation might hint that acceleration is unlikely to be possible as condition number is 1.)

The Cyclic Block-Coordinate Conditional Gradient algorithm

Luckily, [Beck et al., 2015] already thought about this...

Algorithm Cyclic Block-Coordinate Conditional Gradient algorithm [Beck et al., 2015]

Input: Points $x_i^0 \in P_i$, LMO for $P_i \subseteq \mathbb{R}^{n_i}$, $i = 0, \dots, k-1$ and $0 < \gamma_0, \dots, \gamma_t, \dots \leq 1$.
Output: Iterates $x^1, \dots \in P_0 \times \cdots \times P_{k-1}$

---

for $t = 0$ **to** $\dots$ **do**
    $i \leftarrow t \bmod k$
    $v^t \leftarrow \operatorname{argmin}_{x \in P_i} \left\langle \nabla_{P_i} f(x^t), x \right\rangle$
    $x^{t+1} \leftarrow x^t + \gamma_t (v^t - x_i^t)_{[i]}$

The Cyclic Block-Coordinate Conditional Gradient algorithm

Alternating Linear Minimizations

Luckily, [Beck et al., 2015] already thought about this...

Algorithm Cyclic Block-Coordinate Conditional Gradient algorithm [Beck et al., 2015]

Input: Points $x_i^0 \in P_i$, LMO for $P_i \subseteq \mathbb{R}^{n_i}$, $i = 0, \ldots, k-1$ and $0 < \gamma_0, \ldots, \gamma_t, \ldots \leq 1$.
Output: Iterates $x^1, \ldots \in P_0 \times \cdots \times P_{k-1}$

1: **for** $t = 0$ **to** $\ldots$ **do**
2: $\quad i \leftarrow t \bmod k$
3: $\quad v^t \leftarrow \operatorname{argmin}_{x \in P_i} \langle \nabla_{P_i} f(x^t), x \rangle$
4: $\quad x^{t+1} \leftarrow x^t + \gamma_t (v^t - x_i^t)_{[i]}$

Theorem (Convergence [Beck et al., 2015, cf Theorem 4.5])

Under standard assumptions

$$\textit{(primal)} \quad f(x^{kt}) - f(x^*) \leq \frac{2}{t+2} \left(\sum_{i=0}^{k-1} \frac{L_i D_i^2}{2} + 2LD \sum_{i=0}^{k-1} D_i \right),$$

$$\textit{(dual)} \quad \min_{1 \leq t \leq T} \max_{y \in P_0 \times \cdots \times P_{k-1}} \left\langle \nabla f(x^{kt}), x^{kt} - y \right\rangle \leq \frac{6.75}{T+2} \left(\sum_{i=0}^{k-1} \frac{L_i D_i^2}{2} + 2LD \sum_{i=0}^{k-1} D_i \right).$$

Note. Cyclic variant of stochastic BCFW [Lacoste-Julien et al., 2013]

Specializing Cyclic Block Coordinate Conditional Gradients [Beck et al., 2015]:

Alternating Linear Minimization algorithm

Alternating Linear Minimizations

Specializing Cyclic Block Coordinate Conditional Gradients [Beck et al., 2015]:

Algorithm Alternating Linear Minimizations (ALM)

Input: Points $x_0 \in P, y_0 \in Q$, LMO over $P, Q \subseteq \mathbb{R}^n$
Output: Iterates $x_1, y_1 \ldots \in \mathbb{R}^n$

for $t = 0$ **to** … **do**
  $u_t \leftarrow \text{argmin}_{x \in P} \langle x_t - y_t, x \rangle$
  $x_{t+1} \leftarrow x_t + \frac{2}{t+2} \cdot (u_t - x_t)$
  $v_t \leftarrow \text{argmin}_{y \in Q} \langle y_t - x_{t+1}, y \rangle$
  $y_{t+1} \leftarrow y_t + \frac{2}{t+2} \cdot (v_t - y_t)$

Alternating Linear Minimization algorithm

Specializing Cyclic Block Coordinate Conditional Gradients [Beck et al., 2015]:

Algorithm Alternating Linear Minimizations (ALM)

Input: Points $x_0 \in P$, $y_0 \in Q$, LMO over $P, Q \subseteq \mathbb{R}^n$
Output: Iterates $x_1, y_1 \ldots \in \mathbb{R}^n$

for $t = 0$ **to** $\ldots$ **do**
  $u_t \leftarrow \operatorname{argmin}_{x \in P} \langle x_t - y_t, x \rangle$
  $x_{t+1} \leftarrow x_t + \frac{2}{t+2} \cdot (u_t - x_t)$
  $v_t \leftarrow \operatorname{argmin}_{y \in Q} \langle y_t - x_{t+1}, y \rangle$
  $y_{t+1} \leftarrow y_t + \frac{2}{t+2} \cdot (v_t - y_t)$

Observe.

1. Trivial algorithm: von Neumann + Sliding = inexact projection via FW requiring around $O(1/t)$ FW step per iteration.
2. Here: Single(!!) Frank-Wolfe step on projection problem per iteration.

Convergence Guarantee

Alternating Linear Minimizations

Proposition (Intersection of two sets)

Let P and Q be compact convex sets. Then ALM generates iterates $z_t \doteq \frac{1}{2}(x_t + y_t)$*, such that*

$$\max\{\operatorname{dist}(z_t, P)^2, \operatorname{dist}(z_t, Q)^2\} \leq \frac{\|x_t - y_t\|^2}{4} \leq \frac{(1 + 2\sqrt{2})(D_P^2 + D_Q^2)}{t+2} + \frac{\operatorname{dist}(P, Q)^2}{4}$$

$$\min_{1 \leq t \leq T} \max_{x \in P, y \in Q} \|x_t - y_t\|^2 - \langle x_t - y_t, x - y \rangle \leq \frac{6.75(1 + 2\sqrt{2})}{T+2}(D_P^2 + D_Q^2).$$

Convergence Guarantee

Alternating Linear Minimizations

Proposition (Intersection of two sets)

Let P and Q be compact convex sets. Then ALM generates iterates $z_t \doteq \frac{1}{2}(x_t + y_t)$, such that

$$\max\{\operatorname{dist}(z_t, P)^2, \operatorname{dist}(z_t, Q)^2\} \leq \frac{\|x_t - y_t\|^2}{4} \leq \frac{(1 + 2\sqrt{2})(D_P^2 + D_Q^2)}{t+2} + \frac{\operatorname{dist}(P, Q)^2}{4}$$

$$\min_{1 \leq t \leq T} \max_{x \in P, y \in Q} \|x_t - y_t\|^2 - \langle x_t - y_t, x - y \rangle \leq \frac{6.75(1 + 2\sqrt{2})}{T+2}(D_P^2 + D_Q^2).$$

Note. *Rate is optimal, take $P = \Delta_n$ and $Q = \{0\} \Rightarrow$ standard lower bound for FW methods.*

Convergence Guarantee

Alternating Linear Minimizations

Proposition (Intersection of two sets)

Let P and Q be compact convex sets. Then ALM generates iterates $z_t \doteq \frac{1}{2}(x_t + y_t)$, such that

$$\max\{\operatorname{dist}(z_t, P)^2, \operatorname{dist}(z_t, Q)^2\} \leq \frac{\|x_t - y_t\|^2}{4} \leq \frac{(1 + 2\sqrt{2})(D_P^2 + D_Q^2)}{t+2} + \frac{\operatorname{dist}(P, Q)^2}{4}$$

$$\min_{1 \leq t \leq T} \max_{x \in P, y \in Q} \|x_t - y_t\|^2 - \langle x_t - y_t, x - y \rangle \leq \frac{6.75(1 + 2\sqrt{2})}{T+2}(D_P^2 + D_Q^2).$$

Note. *Rate is optimal, take $P = \Delta_n$ and $Q = \{0\} \Rightarrow$ standard lower bound for FW methods.*

Remark (Comparison to von Neumann's alternating projection algorithm)

For simplicity let us consider the case where $P \cap Q \neq \emptyset$.
After minor reformulation, von Neumann's alternating projection method yields:

$$\min_{t=0,\ldots,T-1} \max\{\operatorname{dist}(z_{t+1}, P)^2, \operatorname{dist}(z_{t+1}, Q)^2\} \leq \frac{\operatorname{dist}(y_0, P \cap Q)^2}{T}.$$

Alternating Linear Minimization yields:

$$\max\{\operatorname{dist}(z_T, P)^2, \operatorname{dist}(z_T, Q)^2\} \leq \frac{(1 + 2\sqrt{2})(D_P^2 + D_Q^2)}{T+2}.$$

Convergence for $P \cap Q = \emptyset$

Alternating Linear Minimizations

Corollary (Certifying that $P \cap Q = \emptyset$)

If some iterates of ALM satisfy

$$\|x_t - y_t\|^2 > \frac{4(1 + 2\sqrt{2})(D_P^2 + D_Q^2)}{t + 2}$$

then $P \cap Q = \emptyset$. If $P \cap Q = \emptyset$ then the above condition is satisfied after at most

$$\frac{8(1 + 2\sqrt{2})(D_P^2 + D_Q^2)}{\mathrm{dist}(P, Q)^2}$$

block-LMO calls.

Convergence for $P \cap Q = \emptyset$ without knowledge of D_P and D_Q

Corollary (Certifying $P \cap Q = \emptyset$ without knowledge of D_P and D_Q)

Then executing ALM, after at most

$$\frac{13.5(1+2\sqrt{2})(D_P^2 + D_Q^2)}{\operatorname{dist}(P,Q)^2}$$

block-LMO calls, some (of the already seen!) iteration t provides the following certificate for disjointness, which does not require explicit bounds on D_P and D_Q:

$$\min_{x \in P, y \in Q} \langle x_t - y_t, x - y \rangle > 0.$$

Moreover, this inequality is guaranteed to hold for every iteration

$$t > 4(1+2\sqrt{2})(D_P^2 + D_Q^2)(D_P + D_Q)^2 / \operatorname{dist}(P,Q)^4.$$

Convergence for $P \cap Q = \emptyset$ without knowledge of D_P and D_Q

Alternating Linear Minimizations

Corollary (Certifying $P \cap Q = \emptyset$ without knowledge of D_P and D_Q)

Then executing ALM, after at most

$$\frac{13.5(1 + 2\sqrt{2})(D_P^2 + D_Q^2)}{\operatorname{dist}(P, Q)^2}$$

block-LMO calls, some (of the already seen!) iteration t provides the following certificate for disjointness, which does not require explicit bounds on D_P and D_Q:

$$\min_{x \in P, y \in Q} \langle x_t - y_t, x - y \rangle > 0.$$

Moreover, this inequality is guaranteed to hold for every iteration

$$t > 4(1 + 2\sqrt{2})(D_P^2 + D_Q^2)(D_P + D_Q)^2 / \operatorname{dist}(P, Q)^4.$$

Note. Testing in each iteration would be inefficient (additional LMO call).

All done?

Alternating Linear Minimizations

All done? We have been cheating however...

Alternating Linear Minimizations

Both von Neumann's algorithm and ALM only approximately decide $x \in P \cap Q$!

All done? We have been cheating however...

Alternating Linear Minimizations

Both von Neumann's algorithm and ALM only approximately decide $x \in P \cap Q$!

For general compact convex sets this is as good as it gets but for polytopes?

Alternating Linear Minimizations for Polytopes

A simply observation

Alternating Linear Minimizations for Polytopes

Observation (Approximate-Exact Crossover)

Let $P, Q \subseteq \mathbb{R}^n$ *be polytopes. There exists* $\varepsilon_{PQ} > 0$*, so that for all* $U \subseteq \mathrm{vert}(P), V \subseteq \mathrm{vert}(Q)$ *with* $\mathrm{dist}(\mathrm{conv}(U), \mathrm{conv}(V)) < \varepsilon_{PQ}$*, it holds* $\mathrm{conv}(U) \cap \mathrm{conv}(V) \neq \emptyset$*.*

A simply observation

Alternating Linear Minimizations for Polytopes

Observation (Approximate-Exact Crossover)

Let $P, Q \subseteq \mathbb{R}^n$ be polytopes. There exists $\varepsilon_{PQ} > 0$, so that for all $U \subseteq \operatorname{vert}(P), V \subseteq \operatorname{vert}(Q)$ with $\operatorname{dist}(\operatorname{conv}(U), \operatorname{conv}(V)) < \varepsilon_{PQ}$, it holds $\operatorname{conv}(U) \cap \operatorname{conv}(V) \neq \emptyset$.

Proof.

Follows from the fact that polytopes having only a finite number of vertices:

$$\varepsilon_{PQ} := \min\{\operatorname{dist}(\operatorname{conv}(U), \operatorname{conv}(V)) : U \subseteq \operatorname{vert}(P), V \subseteq \operatorname{vert}(Q), \operatorname{conv}(U) \cap \operatorname{conv}(V) = \emptyset\}.$$

□

A simply observation

Alternating Linear Minimizations for Polytopes

Observation (Approximate-Exact Crossover)

Let $P, Q \subseteq \mathbb{R}^n$ be polytopes. There exists $\varepsilon_{PQ} > 0$, so that for all $U \subseteq \operatorname{vert}(P), V \subseteq \operatorname{vert}(Q)$ with $\operatorname{dist}(\operatorname{conv}(U), \operatorname{conv}(V)) < \varepsilon_{PQ}$, it holds $\operatorname{conv}(U) \cap \operatorname{conv}(V) \neq \emptyset$.

Proof.

Follows from the fact that polytopes having only a finite number of vertices:

$$\varepsilon_{PQ} := \min\{\operatorname{dist}(\operatorname{conv}(U), \operatorname{conv}(V)) : U \subseteq \operatorname{vert}(P), V \subseteq \operatorname{vert}(Q), \operatorname{conv}(U) \cap \operatorname{conv}(V) = \emptyset\}.$$

□

Of course we do not know ε_{PQ} ahead of time...

Another simple observation

Alternating Linear Minimizations for Polytopes

Observation (Recovery of $x \in P \cap Q$ by linear programming)

Assume x_t and y_t with $\|x_t - y_t\| < \varepsilon_{PQ}$ via ALM.

Another simple observation

Observation (Recovery of $x \in P \cap Q$ by linear programming)

Assume x_t and y_t with $\|x_t - y_t\| < \varepsilon_{PQ}$ via ALM.

Let $U \subseteq \mathrm{vert}(P)$ be all extreme points returned by the LMO for P throughout the execution of ALM and define $V \subseteq \mathrm{vert}(Q)$ accordingly. From Observation: $\mathrm{conv}(U) \cap \mathrm{conv}(V) \neq \emptyset$.

Another simple observation

Alternating Linear Minimizations for Polytopes

Observation (Recovery of $x \in P \cap Q$ by linear programming)

Assume x_t and y_t with $\|x_t - y_t\| < \varepsilon_{PQ}$ via ALM.

Let $U \subseteq \mathrm{vert}(P)$ be all extreme points returned by the LMO for P throughout the execution of ALM and define $V \subseteq \mathrm{vert}(Q)$ accordingly. From Observation: $\mathrm{conv}(U) \cap \mathrm{conv}(V) \neq \emptyset$.

Solve linear feasibility program

$$\sum_{u \in U} \lambda_u u = \sum_{v \in V} \kappa_u v$$

$$\sum_{u \in U} \lambda_u = 1, \sum_{v \in V} \kappa_u = 1$$

$$\lambda \geq 0, \kappa \geq 0,$$

to obtain

$$x := \sum_{u \in U} \lambda_u u = \sum_{v \in V} \kappa_u v \in P \cap Q.$$

An exact algorithm

Alternating Linear Minimizations for Polytopes

Algorithm Alternating Linear Minimizations (ALM) [exact version]

Input: Points $x_0 \in P, y_0 \in Q$, LMO over $P, Q \subseteq \mathbb{R}^n$
Output: Iterates $x_1, y_1 \ldots \in \mathbb{R}^n$

for $t = 0$ **to** ... **do**
  $u_t \leftarrow \operatorname{argmin}_{x \in P} \langle x_t - y_t, x \rangle$
  $x_{t+1} \leftarrow x_t + \frac{2}{t+2} \cdot (u_t - x_t)$
  $v_t \leftarrow \operatorname{argmin}_{y \in Q} \langle y_t - x_{t+1}, y \rangle$
  $y_{t+1} \leftarrow y_t + \frac{2}{t+2} \cdot (v_t - y_t)$
  if $t = 2^k$ for some k **then**
    if $\min_{x \in P, y \in Q} \langle x_{t+1} - y_{t+1}, x - y \rangle > 0$ **then**
      return "disjoint" and certificate $\langle x_{t+1} - y_{t+1}, x - y \rangle > 0$
    else
      Solve linear feasibility program.
      if feasible **then**
        return a solution $x \in P \cap Q$

An exact algorithm: Guarantees

Alternating Linear Minimizations for Polytopes

Basically we pay a factor of 2 in iterations for making exact.

Proposition (Exact variant)

Let P, Q be polytopes with diameters D_p and D_Q, respectively. Executing exact ALM variant:

1. *If $P \cap Q \neq \emptyset$, then after no more than*

$$\frac{16(1+2\sqrt{2})(D_P^2 + D_Q^2)}{\varepsilon_{PQ}^2}$$

   block-LMO calls, the algorithm returns $x \in P \cap Q$.

2. *If $P \cap Q = \emptyset$, then after no more than*

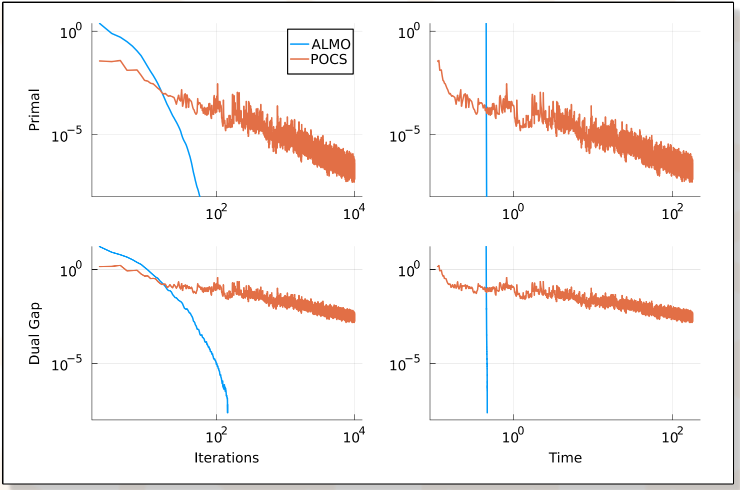
$$16(1+2\sqrt{2})(D_P^2 + D_Q^2)\frac{(D_P + D_Q)^2}{\operatorname{dist}(P,Q)^4}$$

   block-LMO calls the algorithm certifies $P \cap Q = \emptyset$.

Note. We counted the resolution of one feasibility LP as one block-LMO.

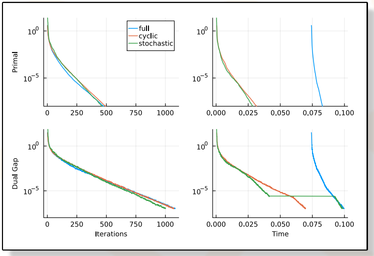
Some ~~cooked~~ preliminary computational results...

ALMO vs. POCS

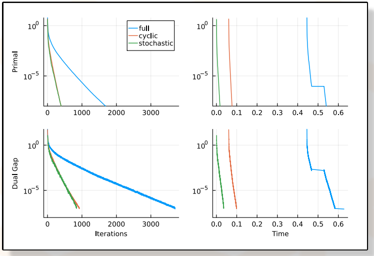


Intersection of two polytopes. Projection problem solved approximately via FW; relevant for time, irrelevant for iterations.

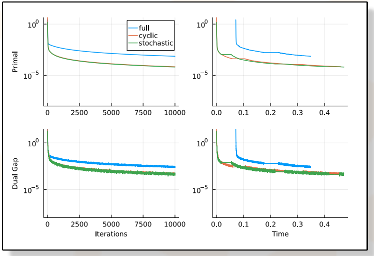
Full vs. stochastic vs. cyclic



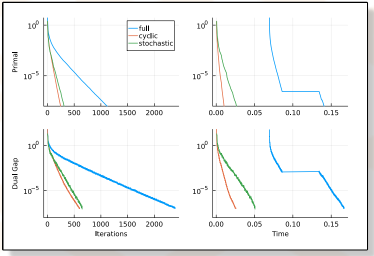
L_1-ball and random polytope (intersecting)



two random polytopes (non-intersecting)

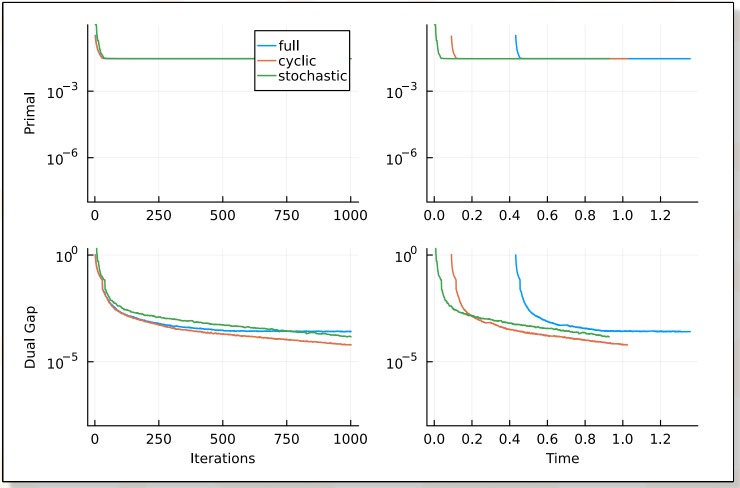


L_1-ball and random polytope (intersecting)



three random polytopes

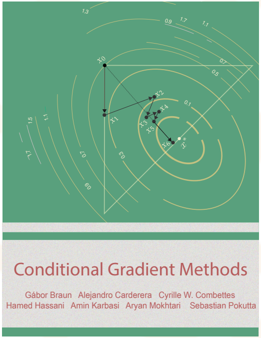
Solving SDPs



Intersection of SDP cone with polytope

Shameless plug...

Thank you!



Conditional Gradient Methods

Gábor Braun, Alejandro Carderera, Cyrille W Combettes, Hamed Hassani, Amin Karbasi, Aryan Mokhtari, and Sebastian Pokutta

`https://conditional-gradients.org/`
`https://arxiv.org/abs/2211.14103`

References I

A. Beck, E. Pauwels, and S. Sabach. The cyclic block conditional gradient method for convex optimization problems. *SIAM Journal on Optimization*, 25(4): 2024–2049, 2015. ISSN 1052-6234 (print); 1095-7189 (online). doi: 10.1137/15M1008397.

G. Braun, A. Carderera, C. W. Combettes, H. Hassani, A. Karbasi, A. Mokthari, and S. Pokutta. Conditional gradient methods, 8 2022.

S. Lacoste-Julien, M. Jaggi, M. Schmidt, and P. Pletscher. Block-coordinate frank-wolfe optimization for structural svms. In *International Conference on Machine Learning*, pages 53–61. PMLR, 2013.

J. von Neumann. On rings of operators. reduction theory. *Annals of Mathematics*, pages 401–485, 1949.