Algorithmic and Geometric Aspects of Alternating Linear Minimization (Part 1)

Sebastian Pokutta

joint work with: Gábor Braun and Robert Weismantel

Technische Universität Berlin
and
Zuse Institute Berlin

pokutta@math.tu-berlin.de
@spokutta

PolyU - Department of Applied Mathematics

February 2nd, 2026 · Hong Kong

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 couple of new algorithms.

(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 $P = \text{conv}(U)$ and $Q = \text{conv}(W)$ 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.

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:

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

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:

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

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:

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

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

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:

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

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

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

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

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

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

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,

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

3) divide by T, then

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

as distances are non-increasing. $\square$

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 $x_1, y_1 \dots, x_T, y_T \in \mathbb{R}^n$ 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)

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

Alternating Linear Minimizations

[Braun et al., 2022]

After close inspection and some meditation, 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 $P \times Q$.

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., 2025]

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., 2025]

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.)

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 $x^1, \dots \in P_0 \times \dots \times P_{k-1}$

1: **for** $t = 0$ **to** $\dots$ **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

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

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

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

Alternating Linear Minimization algorithm

Alternating Linear Minimizations

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

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 ... do
    u_t ← argmin_{x∈P} ⟨x_t − y_t, x⟩
    x_{t+1} ← x_t + 2/(t+2) · (u_t − x_t)
    v_t ← argmin_{y∈Q} ⟨y_t − x_{t+1}, y⟩
    y_{t+1} ← y_t + 2/(t+2) · (v_t − y_t)
```

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

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 steps 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

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

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

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

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

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

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:

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

Alternating Linear Minimization yields:

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

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

[Braun et al., 2022]

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 $U \subseteq \mathrm{vert}(P)$, $V \subseteq \mathrm{vert}(Q)$ with $\mathrm{dist}(\mathrm{conv}(U), \mathrm{conv}(V)) < \varepsilon_{PQ}$, it holds $\mathrm{conv}(U) \cap \mathrm{conv}(V) \neq \emptyset$.

Proof.

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

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

□

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

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: $\mathrm{conv}(U) \cap \mathrm{conv}(V) \neq \emptyset$.

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

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

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

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

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

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

How bad can it be?

Bounds on minimal distance

How bad can it be?

The minimal possible distance $\mathrm{dist}(P, Q)$ can be actually quite bad.

Bounds on minimal distance

How bad can it be?

The minimal possible distance $\operatorname{dist}(P, Q)$ can be actually quite bad.

[Deza et al., 2024]

Theorem

If P and Q are disjoint lattice (d, k)-polytopes, then

$$\frac{1}{(kd)^{2d}} \leq \operatorname{dist}(P, Q),$$

and for any large enough d, there exist two disjoint (d, k)-lattice polytopes P and Q such that

$$\operatorname{dist}(P, Q) \leq \frac{1}{\left(k\sqrt{d}\right)^{\sqrt{d}}}.$$

Bounds on minimal distance

How bad can it be?

The minimal possible distance $\mathrm{dist}(P, Q)$ can be actually quite bad.

[Deza et al., 2024]

Theorem

If P and Q are disjoint lattice (d, k)-polytopes, then

$$\frac{1}{(kd)^{2d}} \leq \mathrm{dist}(P, Q),$$

and for any large enough d, there exist two disjoint (d, k)-lattice polytopes P and Q such that

$$\mathrm{dist}(P, Q) \leq \frac{1}{\left(k\sqrt{d}\right)^{\sqrt{d}}}.$$

$\Rightarrow$ In case of disjoint polytopes running time can be as bad as

$$\Omega\left(\left(k\sqrt{d}\right)^{4\sqrt{d}}\right).$$

$\Rightarrow$ Bad news for our algorithms.

Can we do better?

Advanced FW algorithms over polytopes

Can we do better?

AFW, PCG, BCG, BPCG, etc. can solve

$$\min_{x \in P} f(x),$$

to accuracy ε in roughly

$$\mathcal{O}\left(\frac{LD^2}{\mu\delta^2}\log\frac{1}{\varepsilon}\right)$$

iterations, for f being L-smooth and μ-PL over a polytope P with pyramidal width δ.

Advanced FW algorithms over polytopes

Can we do better?

AFW, PCG, BCG, BPCG, etc. can solve

$$\min_{x \in P} f(x),$$

to accuracy ε in roughly

$$\mathcal{O}\left(\frac{LD^2}{\mu\delta^2} \log \frac{1}{\varepsilon}\right)$$

iterations, for f being L-smooth and μ-PL over a polytope P with pyramidal width δ.

Note. Exponentially better dependence on ε.

Advanced FW algorithms over polytopes

Can we do better?

AFW, PCG, BCG, BPCG, etc. can solve

$$\min_{x \in P} f(x),$$

to accuracy ε in roughly

$$\mathcal{O}\left(\frac{LD^2}{\mu\delta^2}\log\frac{1}{\varepsilon}\right)$$

iterations, for f being L-smooth and μ-PL over a polytope P with pyramidal width δ.

Note. Exponentially better dependence on ε.

Recall. Our problem can be formulated as

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

which is 1-smooth and 1-PL (basically "like" strong-convexity).

Pyramidal width Over Products

Can we do better?

With a little bit of geometric reasoning we can show: [Iommazzo et al., 2026]

Theorem (Pyramidal width of the product)

Let δ_P and δ_Q be the pyramidal widths of polytopes $P, Q \subseteq \mathbb{R}^n$. Then,

$$\delta_{P\times Q} = \sqrt{\frac{\delta_P^2 \delta_Q^2}{\delta_P^2 + \delta_Q^2}}.$$

Pyramidal width Over Products

Can we do better?

With a little bit of geometric reasoning we can show: [Iommazzo et al., 2026]

Theorem (Pyramidal width of the product)

Let δ_P and δ_Q be the pyramidal widths of polytopes $P, Q \subseteq \mathbb{R}^n$. Then,

$$\delta_{P\times Q} = \sqrt{\frac{\delta_P^2 \delta_Q^2}{\delta_P^2 + \delta_Q^2}}.$$

Corollary (Useful lower bound for the pyramidal width of the product)

The pyramidal width of product polytope $P = \prod_{i\in[k]} P_i$ is at least

$$\delta_P = \Omega\left\{\frac{1}{\sqrt{k}} \min_{i\in[k]} \delta_{P_i}\right\}$$

with δ_{P_i} being pyramidal width of P_i; bound is essentially tight when one pyramidal width is much smaller than the others.

Putting it all together

Can we do better?

[Iommazzo et al., 2026]

Proposition (Faster exact variant)

Let P, Q be polytopes with diameters D_p and D_Q, respectively. Executing exact ALM variant with AFW, PCG, BCG, BPCG, etc. steps:

1. *If $P \cap Q \neq \emptyset$, then the algorithm returns $x \in P \cap Q$ in*

$$\mathcal{O}\left(\frac{D_P^2 D_Q^2}{\min\{\delta_P, \delta_Q\}^2} \log \frac{1}{\varepsilon_{PQ}}\right).$$

2. *If $P \cap Q = \emptyset$, then the algorithm certifies $P \cap Q = \emptyset$ in*

$$\mathcal{O}\left(\frac{D_P^2 D_Q^2}{\min\{\delta_P, \delta_Q\}^2} \log \frac{1}{\operatorname{dist}(P, Q)}\right).$$

Note. $D_P, D_Q, \delta_P, \delta_Q$ are translation invariant and only depend on P and Q, respectively.

Putting it all together

Can we do better?

[Iommazzo et al., 2026]

Proposition (Faster exact variant)

Let P, Q be polytopes with diameters D_p and D_Q, respectively. Executing exact ALM variant with AFW, PCG, BCG, BPCG, etc. steps:

1. *If $P \cap Q \neq \emptyset$, then the algorithm returns $x \in P \cap Q$ in*

$$\mathcal{O}\left(\frac{D_P^2 D_Q^2}{\min\{\delta_P, \delta_Q\}^2} \log \frac{1}{\varepsilon_{PQ}}\right).$$

2. *If $P \cap Q = \emptyset$, then the algorithm certifies $P \cap Q = \emptyset$ in*

$$\mathcal{O}\left(\frac{D_P^2 D_Q^2}{\min\{\delta_P, \delta_Q\}^2} \log \frac{1}{\operatorname{dist}(P, Q)}\right).$$

Note. $D_P, D_Q, \delta_P, \delta_Q$ are translation invariant and only depend on P and Q, respectively.
Worst-case example from before. Running time reduces to

$$\mathcal{O}\left(\frac{D_P^2 D_Q^2}{\min\{\delta_P, \delta_Q\}^2} \sqrt{d} \log k\sqrt{d}\right).$$

Outlook

Integrating LP solving into convex optimization is very powerful
Outlook

[Halbey et al., 2025]

In Entanglement Detection, Sliding, Splitting, etc. we encounter quadratic programs.

$\Rightarrow$ First-order optimality system is a linear program!

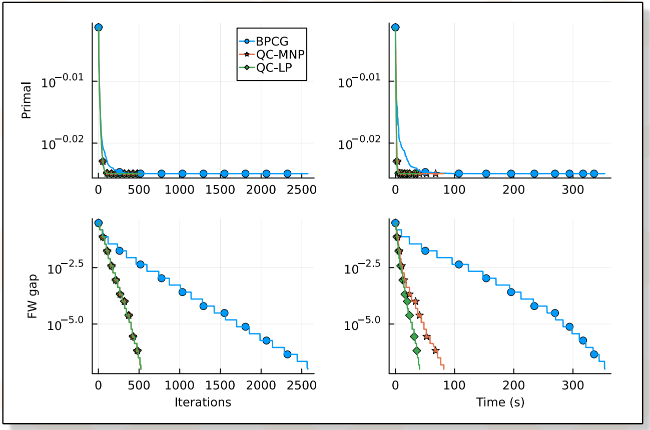
Integrating LP solving into convex optimization is very powerful

Outlook

[Halbey et al., 2025]

In Entanglement Detection, Sliding, Splitting, etc. we encounter quadratic programs.

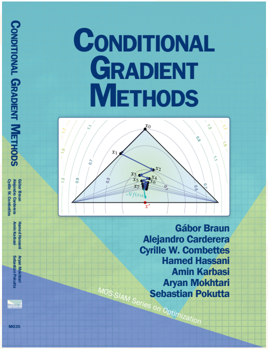
⇒ First-order optimality system is a linear program!



Integration of LP solving into convex optimization is very powerful.

If you want to learn more...

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`

MOS-SIAM Series on Optimization

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, S. Pokutta, and R. Weismantel. Alternating Linear Minimization: Revisiting von Neumann's alternating projections. *preprint*, 12 2022.

G. Braun, A. Carderera, C. W. Combettes, H. Hassani, A. Karbasi, A. Mokthari, and S. Pokutta. *Conditional Gradient Methods*. to appear in MOS-SIAM Series on Optimization, 1 2025.

A. Deza, S. Onn, S. Pokutta, and L. Pournin. Kissing polytopes. *to appear in SIAM Journal on Discrete Mathematics*, 7 2024.

J. Halbey, S. Rakotomandimby, M. Besançon, S. Designolle, and S. Pokutta. Efficient Quadratic Corrections for Frank-Wolfe Algorithms. *preprint*, 6 2025.

G. Iommazzo, D. Martínez-Rubio, F. Criado, E. Wirth, and S. Pokutta. Linear Convergence of the Frank-Wolfe Algorithm over Product Polytopes. *to appear in Proceedings of AISTATS*, 1 2026. doi: 10.48550/arXiv.2505.11259.

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.