Nonlinear predictive control in robotics: optimization-based collision avoidance (OBCA), primal and dual approaches (polyhedral sets)

Antonio Sala

Lecture notes for aerospace/robotics engineers

Universitat Politècnica de València

Videos: http://personales.upv.es/asala/YT/V/**obca1pEN**.html, http://personales.upv.es/asala/YT/V/**obca2d1EN**.html, http://personales.upv.es/asala/YT/V/**obca3d2EN**.html, http://personales.upv.es/asala/YT/V/**obca4d3EN**.html, http://personales.upv.es/asala/YT/V/**obca5cEN**.html.



UNIVERSITAT POLITÈCNICA DE VALÈNCIA

Outline

Motivation:

Our goal is computing an "optimal" route for a given robot from a start to goal configuration given some initial guess waypoints and a cost index. The route must avoid collisions with (known) obstacles.

Objectives:

Understand the various options to incorporate collision avoidance into nonlinear model predictive control optimization.

Contents:

Point clouds, obstacles, safe corridors. Collision checks. Collision gradients. The log-exp smoothing of multiple constraints. Duality approaches guaranteeing point-to-polyhedron or polyhedron-to-polyhedron clearance.

UNIVERSITAT POLITÈCNICA DE VALÈNCIA

Collision avoidance, naive approach as "point cloud"

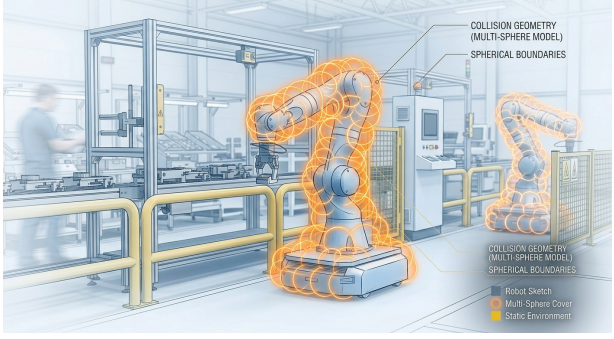
Robot state q will map to a series of "key" points $x_k(q)$ $k = 1, \ldots, n_p$, which must be outside obstacles and/or inside a safe corridor. They must be computed for a set of "close-enough" waypoints describing the whole trajectory $x_k(q_j)$.

- Center of a mobile robot/UAV, etc.
- A finite set of said "key" points

We may combine "obstacle avoidance" and "safe corridor" configuration constraints, if convenient (checking for "all" obstacles at "every" route waypoint).

UNIVERSITAT POLITÈCNICA DE VALÈNCIA

Obstacles should be augmented with a "clearance" so that if none of the points is in collision (if all are in the corridor), the whole robot isn't in collision (is in the safe corridor), say, a "spherical" cover.



Collision check

• If the safe corridor is $g(x) \leq 0$, We'll test $g(x(q)) \leq 0$ for all x in the key point collection. If it is defined by multiple constraints $\{g_i(x) \leq 0, i = 1, 2, \ldots, m\}$, we add them all to the optimizer constraints, equiv. $\max_i g_i(x(q)) \leq 0$.

• If an obstacle is $g(x) \leq 0$ with a single inequality, then we'll check $g(x(q)) > 0$. Say, outside an ellipse if $(x - c)^T Q(x - c) > 1$.

• If an obstacle is $\{g_i(x) \leq 0, i = 1, 2, \ldots, m\}$ then we need to check that at least one constraint does not hold: $\max_i g_i(x) \geq 0$.

***Polyhedral corridors/obstacles:** $Ax \leq b$, i.e., $Ax - b \leq 0$, as many constraints as polyhedron faces.

UNIVERSITAT POLITÈCNICA DE VALÈNCIA

Collision gradient

If a constraint is violated, in the 1st or 2nd case, the solver can compute a Jacobian $\frac{dg_i}{dx}\frac{dx}{dq}$ to obtain a step direction towards the constraint.

The third case is special: The "maximum" is not differentiable... if the point is "inside" the obstacle, we need to check which is the constraint closest to violation, and compute the step in that direction, but that is a "disjunction" (checking binary options) or a gradient discontinuity that may slow progress of large-scale solvers, which like "smooth" functions for descent based on local linear/quadratic constraint approximations.

UNIVERSITAT POLITÈCNICA DE VALÈNCIA

The simplest solution: smoothen the constraint and craft a single function, say

$$\tilde{g}(x) = \gamma^{-1} \cdot \log\left(\sum_i exp\left(\gamma g_i(x)\right)/m\right), \qquad \gamma > 0$$

so an obstacle with smoothed edges is $\tilde{g}(x) \leq 0$, and the level set contains the obstacle ($\tilde{g}$ would be zero if all $g_i = 0$, $i = 1, \ldots, m$); if we don't divide by m, we get an smoothed inner approximation).

Large values for the parameter γ make $\tilde{g}$ closer to the ideal $\max_i g_i$, but worsen the gradient conditioning (abrupt changes of gradient direction), which might impact solver step sizes and tolerance management.

Small values for γ make $\tilde{g}$ track more loosely the actual boundaries, so we lose accuracy and incur in conservatism/suboptimality... but we may use it to help with the needed "clearance" as robot is not a point.

©2027 A. Sala

UNIVERSITAT POLITÈCNICA DE VALÈNCIA

Advantages amd drawbacks

- [+] Applies to any obstacle defined by several differentiable constraints, be it convex or non-convex.
- [+] Allows to tune the clearance with the global γ or scaling each constraint.
- [-] Trajectory is "discretized" in a finite set of waypoints, inter-waypoint segments may not fulfill the constraints, so extra clearance is needed.
- [-] Actual value of the resulting clearance is unknown and changing depending on the direction of approximation, trial-and-error required. Some values of γ may give level sets $\tilde{g}(x) \leq 0$ which do not fully contain the obstacle.
- [-] Lack of precision for small γ: forbidden region much larger than actual obstacle.
- [-] Gradient is ill-conditioned, abruptly changing, for large γ.

©2027 A. Sala

UNIVERSITAT POLITÈCNICA DE VALÈNCIA

Collision avoidance, point out of convex polyhedron: DUALITY approach [Zhang21]

Consider the "base" problem

$$J^* = \min_t \|t\|, \quad \text{subject to} \quad A(x - t) \leq b$$

which will compute the minimal "translation" of a point x (constant for the moment being) so that it's inside $\{\xi : A\xi \leq b\}$. This is a QP optimization problem, so fast solutions exist.

Problem: if x is a trajectory decision variable, placing a constraint that "*the solution to the above QP is above a given* $d_{threshold}$", i.e., $J^* \geq d_{threshold}$ is awkward to solve... does it require nested optimization?

From the base problem,

$$J^* = \min_t \|t\|, \quad \text{subject to} \quad A(x-t) \leq b$$

Inserting KKT multipliers $\lambda \geq 0$, consider the "primal" problem:

$$J^p = \min_t \max_{\lambda \geq 0} \big(\underbrace{\|t\| + \lambda^T \cdot (A(x-t) - b)}_{\mathcal{L}} \big)$$

Lemma: If base problem is feasible, $J^* = J^p$, as $\max_{\lambda \geq 0} \lambda^T (A(x-t) - b)$ is zero if $A(x-t) \leq b$, and $+\infty$ if at least one constraint is violated.

UNIVERSITAT POLITÈCNICA DE VALÈNCIA

From primal $J^p = \min_t \max_{\lambda \geq 0} \left(\|t\| + \lambda^T \cdot ((-A)t + (Ax - b)) \right)$
Consider its game-theoretic dual problem:

$$J^d = \max_{\lambda \geq 0} \min_t \left(\underbrace{(Ax - b)^T \lambda + \|t\| - (\lambda^T A)t}_{\text{This is, too, } \mathcal{L}\text{, just grouping terms}} \right)$$

Minimax inequality (weak duality):

In a zero-sum game, exposing your strategy first is a disadvantage!

$J^d = \max_\lambda \min_t \mathcal{L} \leq J^p = \min_t \max_\lambda \mathcal{L}$.

Let $\lambda^* = \arg\max_\lambda (\min_t f(\lambda, t))$, and $t^* = \arg\min_t (\max_\lambda f(\lambda, t))$.
$\max_\lambda (\min_t f(\lambda, t)) = \min_t f(\lambda^*, t) \leq f(\lambda^*, t^*) \leq \max_\lambda f(\lambda, t^*) = \min_t (\max_\lambda f(\lambda, t))$

Solution of $\min t$**:** For fixed λ, if $\|\lambda^T A\| > 1$, the minimum is $-\infty$.
If $\|\lambda^T A\| \leq 1$ the minumum is zero.

Reformulation: Dual is $J^d = \max (Ax - b)^T \lambda$ subject to $\|A^T \lambda\| \leq 1$.

UNIVERSITAT POLITÈCNICA DE VALÈNCIA

Main result: point to obstacle separation

The distance of x to a polyhedral obstacle $A\xi \leq b$ is larger than $d_{threshold}$ if the following constraints in decision variables x and λ are feasible:

$$\lambda \geq 0\,, \qquad (Ax-b)^T\lambda \geq d_{threshold}\,, \qquad \|A^T\lambda\| \leq 1$$

So, for each waypoint x and obstacle, we add the auxiliary decision variables λ and the associated constraints above.

Slater conditions hold: we have "strong duality" so "if" is actually "**if and only if**"; see [Zhang21], [Boyd09].

Usage of the point-to-polyhedron conditions

$x(q)$ is a "spherical" robot (center) and we guarantee clearance. Refinement: several $x(q)$ are centers of spheres that cover the robot with some radius. The radius of each sphere is the d_{min} to add as constraint.

Conversely, we may "cover" an obstacle with spheres and the polyhedron be part of the robot.

UNIVERSITAT POLITÈCNICA DE VALÈNCIA

Why "dual"?

Both primal and dual can be solved via convex optimization for either t^* minimising $\|t\|$ or maximising $d_{threshold}$ to obtain d^* the distance between polyhedra. In path planning, this optimization is NOT the problem to solve: x is also a decision variable and the cost is not the distance, we just wish a clearance constraint $d_{threshold}$.

- Stating "Dual feasibility" guarantees the clearance.

- Stating "Primal feasibility" $A(x - t) \leq b$, $\|t\| \geq d_{threshold}$ is WRONG, as the feasible t ranges between minimum and maximum distance beween points in the sets. So x may be inside the polyhedra (collision) and exist a "large" t fulfilling the distance threshold.

- Stating "Primal feasibility" $A(x - t) \leq b$, $\|t\| \leq d_{threshold}$ is WRONG, as the problem is feasible with x inside the polyhedra $t = 0$.

UNIVERSITAT POLITÈCNICA DE VALÈNCIA

Advantages and drawbacks of dual approach

- [+] Certificate of clearance to the actual obstacle, not needing the ad-hoc tuning of γ in the log-sum-exp smoothing.
- [-] Trajectory is "discretized" in a finite set of waypoints, inter-waypoint segments may not fulfill the constraints, so extra clearance is needed.
- [-] For each waypoint, decision variables are robot configuration q (and its speed $\dot{q}$ if dynamics matters), control actions u, and one decision variable per face of each obstacles. If we have a lot of obstacles with a lot of faces, we may need to segment the path, and devise a "corridor" to limit the analysis to those inside the corridor.
- [-] Physical interpretation and initial guess of dual variables is unclear;
- [-] We need to cover the robot with spheres. If number of spheres is low, we may be conservative (they may span a large volume of free space), if number of spheres is high, the number of decision variables exacerbates.

Enhanced version: distance between polyhedra

The original problem is:

$$J^* = \min_{t,x} \|t\|, \quad \text{subject to} \quad A_1 x \leq b_1, \quad A_2(x-t) \leq b_2$$

easily solved via QP, but not usable as a direct trajectory constraint for $A_1 \equiv A_1(q)$, $b_1 \equiv b_1(q)$...

Adding Lagrange (KKT) multipliers we conform:

$$J^p = \min_{t,x} \max_{\lambda_1 \geq 0, \lambda_2 \geq 0} \left(\|t\| + \lambda_1^T (A_1 x - b_1) + \lambda_2^T (A_2(x-t) - b_2) \right)$$

Again $J^p = J^*$ if original problem is feasible, so we are solving the same problem.

 ©2027 A. Sala

The dual

The game-theoretic dual:

$$J^d = \max_{\lambda_1 \geq 0, \lambda_2 \geq 0} \min_{t,x} \left(\|t\| + \lambda_1^T (A_1 x - b_1) + \lambda_2^T (A_2 (x - t) - b_2) \right)$$

provides a lower bound for the optimal distance $J^d \leq J^p = J^*$, via minimax inequality. If we minimize over all primal variables t, x:

$$J^d = \max_{\lambda_1 \geq 0, \lambda_2 \geq 0} \min_{t,x} \left(\|t\| - t^T (A_2^T \lambda_2) + x^T (A_1^T \lambda_1 + A_2^T \lambda_2) - b_1^T \lambda_1 - b_2^T \lambda_2 \right)$$

Evaluating inner minimizations (blue) over t and x (known λ_1, λ_2):

- $\min_x x^T (A_1^T \lambda_1 + A_2^T \lambda_2) = -\infty$ unless $A_1^T \lambda_1 + A_2^T \lambda_2 = 0$, then zero.
- $\min_t (\|t\| - t^T A_2^T \lambda_2) = -\infty$ unless $\|A_2^T \lambda_2\| \leq 1$, then zero.

UNIVERSITAT POLITÈCNICA DE VALÈNCIA

The final "Dual Problem" Formulation:

$$
\begin{aligned}
J^d = \max_{\lambda_1 \geq 0, \lambda_2 \geq 0} \quad & -b_1^T \lambda_1 - b_2^T \lambda_2 \\
\text{subject to} \quad & A_1^T \lambda_1 + A_2^T \lambda_2 = 0 \\
& \|A_2^T \lambda_2\| \leq 1
\end{aligned}
$$

We also have **strong duality** ($J^* = J^d$), because the original problem is convex and Slater's condition (non empty interior) holds if obstacles and robot have "volume", so this problem computes the actual clearance between two polyhedra.

UNIVERSITAT POLITÈCNICA DE VALÈNCIA

Main result: Robot to Obstacle Separation

Two polyhedra ($A_1 x \leq b_1$, $A_2 x \leq b_2$) are separated by at least $d_{\text{threshold}}$ *if and only if* there exist dual vector variables λ_1, λ_2 with compatible sizes satisfying:

$$
\begin{aligned}
-b_1^T \lambda_1 - b_2^T \lambda_2 &\geq d_{\text{threshold}} \\
A_1^T \lambda_1 + A_2^T \lambda_2 &= 0 \\
\|A_2^T \lambda_2\| &\leq 1 \\
\lambda_1 \geq 0, \quad \lambda_2 &\geq 0
\end{aligned}
$$

This is a set of convex, differentiable, feasibility constraints where no "cartesian" point x appears. Actually, in a robotics trajectory, the robot polyhedron (say, $A_1 \xi \leq b_1$) will have $A_1(q)$ and $b_1(q)$, usually full of trigonometric roto-translation formulae.

UNIVERSITAT POLITÈCNICA DE VALÈNCIA

Why "dual"?

Both primal and dual can be solved with convex optimization software for either t^* minimising $\|t\|$ or maximising $d_{threshold}$ to obtand d^* the distance between polyhedra. But in path planning, this optimization is NOT the problem we wish to solve: x is also a decision variable and the cost is not the distance, we just wish a clearance constraint $d_{threshold}$.

- Stating "Dual feasibility" guarantees the clearance.
- Stating "Primal feasibility" $A_1 x \leq b_1$, $A_2(x - t) \leq b_2$, $\|t\| \geq d_{threshold}$ is WRONG, as the feasible t ranges between minimum and maximum distance beween points in the sets.
- Stating "Primal feasibility" $A_1 x \leq b_1$, $A_2(x - t) \leq b_2$, $\|t\| \leq d_{threshold}$ is WRONG, as the problem is feasible with colliding polyhedra (intersecting).

UNIVERSITAT POLITÈCNICA DE VALÈNCIA

Application

A_1 and b_1 will actually be a roto-translated polyhedron $A_1(q)$, $b_1(q)$. If the polyhedron is $A_0 x_{home} \leq b_0$ in a "home" base frame, and robot motion transform coordinates to 'world' frame as $x = R(q)x_{home} + t(q)$, then, we need to check $A_0 R(q)^{-1}(x - t(q)) \leq b_0$, so $A_1 = A_0 R(q)^{-1}$, $b_1 = b_0 + A_0 R(q)^{-1} t(q)$.

* Note that "polyhedral robot" **inside** workspace or safe-corridor polyhedron would be cast in primal vertex form, i.e., enumerating each of the vertices of the robot (with the suitable rotation+translation, of course from robot configuration q), decomposed as the union of convex polyhedra, say, and setting the condition that each of the vertices be inside the workspace/corridor.

 ©2027 A. Sala

Advantages and drawbacks

- [+] Certificate of clearance, without ad-hoc tuning, between two polyhedra.
- [-] Trajectory is "discretized" in a finite set of waypoints, inter-waypoint segments may not fulfill the constraints.
- [+] The approach can be generalized to a larger class of pairs of convex sets (such as LMI/SDP-representable sets).
- [-] We need to cover the robot and the obstacles with a collection of such sets (well, this is a usual step: there is a simplified collision geometry in many robot planning algorithms), simplicity versus conservatism trade-off.
- [-] The number of decision variables per waypoint is robot state $(q, \dot{q})$ plus one per face of robot collision geometry polyhedra, plus one per face of each active obstacle.
- [-] Intuition and drafting a sensible initial guess on dual variables is harder.

UNIVERSITAT POLITÈCNICA DE VALÈNCIA

Conclusions

- Optimal trajectories compatible with kinematics, dynamics and collision avoidance.
- **Primal:** Points inside a safe corridor $q(x) \leq 0$, or outside obstacles via a smoothed $\tilde{g}(x) \geq 0$ (considering some clearances), without additional decision variables.
- **Duality**-based Lagrange multiplier decision variables guarantee distances between points and polyhedra or between two polyhedra with no conservatism.
- If distance to obstacles is wished to be maximized as part of the cost index, it is easily done.
- Generalizations to distances between points, polyhedra, ellipsoids, or generic convex SDP-representable sets (LMI sets) in the seminal work [Zhang21] are addressed in complementary slides.

©2027 A. Sala

UNIVERSITAT POLITÈCNICA DE VALÈNCIA