Model Predictive Control for Linear and Hybrid Systems. Basics on Optimization

Francesco Borrelli

Department of Mechanical Engineering,
University of California at Berkeley,
USA

fborrelli@berkeley.edu

February 3, 2011

References

From my Book:

- Main Concepts **(Chapter 2)**
- Optimality Conditions: Lagrange duality theory and KKT Conditions **(Chapter 3)**
- Polyhedra, Polytopes and Simplices **(Chapter 4)**
- Linear and Quadratic Programming **(Chapter 5)**

Note: The following notes have been extracted from

- Stephen Boyd's lecture notes for his course on convex optimization. The full notes can be downloaded from `http://www.stanford.edu/~boyd`
- "Nonlinear Programming Theory and Algorithms" by Bazaraa, Sherali, Shetty
- "LMIs in Control" by Scherer and Weiland
- "Lectures on Polytopes" by Ziegler

Outline

1. Main concepts
   - Optimization problems
   - Continuous problems
   - Integer and Mixed-Integer Problems
   - Convexity
2. Optimality conditions: Lagrange duality theory and KKT conditions
3. Polyhedra, polytopes and simplices
4. Linear and quadratic programming

Abstract optimization problems

- Optimization problems are abundant in economical daily life (wherever a decision problem is)
- General abstract features of problem formulation:
  - Decision set Z
  - **Constraints** on decision and subset $S \subseteq Z$ of **feasible** decisions.
  - Assign to each decision a **cost** $f(x) \in \mathbb{R}$.
  - Goal is to **minimize** the cost by choosing a suitable decision.
- Concrete features of problem formulation:
  - Z is real vector space: Continuous problem.
  - Z is discrete set: Discrete or combinatorial problem.

Concrete Optimization Problems

Generally formulated as

$$\begin{array}{ll} \inf_z & f(z) \\ \text{subj. to} & z \in S \subseteq Z \end{array} \tag{1}$$

- The **vector** z collects the decision variables
- Z is the **domain** of the decision variables,
- $S \subseteq Z$ is the set of *feasible* or *admissible* decisions.
- The function $f : Z \to \mathbb{R}$ assigns to each decision z a *cost* $f(z) \in \mathbb{R}$.

Shorter form of problem (1)

$$\inf_{z \in S \subseteq Z} f(z) \tag{2}$$

Problem (2) is a called *nonlinear mathematical program* or simply *nonlinear program.*

Concrete Optimization Problems

Solving problem (2) means

- Compute the least possible cost J^* (called the **optimal value**)

$$J^* \triangleq \inf_{z \in S} f(z) \tag{3}$$

  J^* is the greatest lower bound of $f(z)$ over the set S:

$$f(z) \geq J^*, \ \forall z \in S$$

  AND

  1. $$\exists \bar{z} \in S : f(\bar{z}) = J^*$$

  OR

  2. $$\forall \varepsilon > 0 \ \exists z \in S | \ f(z) \leq J^* + \varepsilon$$

- Compute the **optimizer**, $z^* \in S$ with$f(z^*) = J^*$. If z^* exists, then rewrite (3) as

$$J^* = \min_{z \in S} f(z) \tag{4}$$

Concrete Optimization Problems

Consider the Nonlinear Program (NLP)

$$J^* = \min_{z \in S} f(z)$$

Notation:

- If $J^* = -\infty$ the problem is unbounded below.
- If the set S is empty then the problem is said to be infeasible (we set $J^* = +\infty$).
- If $S = Z$ the problem is said to be unconstrained.
- The set of all optimal solutions is denoted by

$$\mathrm{argmin}_{z \in S} f(z) \triangleq \{z \in S : f(z) = J^*\}$$

Continuous problems

Consider the problem

$$
\begin{array}{lll}
\inf_z & f(z) & \\
\text{subj. to} & g_i(z) \leq 0 & \text{for } i = 1, \ldots, m \\
 & h_i(z) = 0 & \text{for } i = 1, \ldots, p \\
 & z \in Z &
\end{array}
\tag{5}
$$

- The **domain** Z of the problem (5) is a subset of $\mathbb{R}^s$ (the finite-dimensional Euclidian vector-space), defined as:

$$
Z = \{z \in \mathbb{R}^s \ : z \in \text{dom } f, \ z \in \text{dom } g_i, \ i = 1, \ldots, m, \\
z \in \text{dom } h_i, \ i = 1, \ldots, p\}
$$

- A point $\bar{z} \in \mathbb{R}^s$ is **feasible** for problem (5) if: $\bar{z} \in Z$ and $g_i(\bar{z}) \leq 0$ for $i = 1, \ldots, m$, $h_i(\bar{z}) = 0$ for $i = 1, \ldots, p$
- The **set of feasible vectors** is

$$
S = \{z \in \mathbb{R}^s : z \in Z, \ g_i(z) \leq 0, \ i = 1, \ldots, m, \ \ h_i(z) = 0, \\
i = 1, \ldots, p\}.
$$

Local and Global Optimizer

- Let J^* be the optimal value of problem (5). A **global optimizer**, if it exists, is a feasible vector z^* with $f(z^*) = J^*$.
- A feasible point $\bar{z}$ is a **local optimizer** for problem (5) if there exists an $R > 0$ such that

$$
\begin{array}{rll}
f(\bar{z}) = & \inf_z & f(z) \\
& \text{subj. to} & g_i(z) \leq 0 \quad \text{for } i = 1, \ldots, m \\
& & h_i(z) = 0 \quad \text{for } i = 1, \ldots, p \\
& & \|z - \bar{z}\| \leq R \\
& & z \in Z
\end{array}
\tag{6}
$$

Active, Inactive and Redundant Constraints

Consider the problem

$$\begin{array}{lll} \inf_z & f(z) & \\ \text{subj. to} & g_i(z) \leq 0 & \text{for } i = 1, \ldots, m \\ & h_i(z) = 0 & \text{for } i = 1, \ldots, p \\ & z \in Z & \end{array}$$

- The i-th inequality constraint $g_i(z) \leq 0$ is **active** at $\bar{z}$ if $g_i(\bar{z}) = 0$, otherwise is **inactive**
- Equality constraints are always active for all feasible points.
- Removing a **redundant constraint** does not change the feasible set S, this implies that removing a redundant constraint from the optimization problem does not change its solution.

Problem Description

- The functions f, g_i and h_i can be available in **analytical form** or can be described through an **oracle model** (also called "black box" or "subroutine" model).
- In an oracle model f, g_i and h_i are not known explicitly but can be evaluated by querying the oracle. Often the oracle consists of subroutines which, called with the augment z, return $f(z)$, $g_i(z)$ and $h_i(z)$ and their gradients $\nabla f(z)$, $\nabla g_i(z)$, $\nabla h_i(z)$.

Integer and Mixed-Integer Problems

- If the decision set Z in the optimization problem is finite, then the optimization problem is called **combinatorial** or **discrete**.
- If $Z \subseteq \{0,1\}^s$, then the problem is said to be **integer**.
- If Z is a subset of the Cartesian product of an integer set and a real Euclidian space, i.e., $Z \subseteq \{[z_c, z_b] : z_c \in \mathbb{R}^{s_c}, z_b \in \{0,1\}^{s_b}\}$, then the problem is said to be **mixed-integer**.

The standard formulation of a **mixed-integer non-linear program** is

$$
\begin{array}{ll}
\inf_{[z_c,z_b]} & f(z_c, z_b) \\
\text{subj. to} & g_i(z_c, z_b) \leq 0 \quad \text{for } i = 1, \ldots, m \\
& h_i(z_c, z_b) = 0 \quad \text{for } i = 1, \ldots, p \\
& z_c \in \mathbb{R}^{s_c},\ z_b \in \{0,1\}^{s_b} \\
& [z_c, z_b] \in Z
\end{array}
\tag{7}
$$

Convexity

A set $S \in \mathbb{R}^s$ is **convex** if

$$\lambda z_1 + (1-\lambda) z_2 \in S \text{ for all } z_1, z_2 \in S, \lambda \in [0,1].$$

A function $f : S \to \mathbb{R}$ is convex if S is convex and

$$\begin{aligned} f(\lambda z_1 + (1-\lambda) z_2) \quad &\leq \quad \lambda f(z_1) + (1-\lambda) f(z_2) \\ &\quad\quad \text{for all } z_1, z_2 \in S, \lambda \in [0,1]. \end{aligned}$$

A function $f : S \to \mathbb{R}$ is **strictly convex** if S is convex and

$$\begin{aligned} f(\lambda z_1 + (1-\lambda) z_2) \quad &< \quad \lambda f(z_1) + (1-\lambda) f(z_2) \\ &\quad\quad \text{for all } z_1, z_2 \in S, \lambda \in (0,1). \end{aligned}$$

A function $f : S \to \mathbb{R}$ is **concave** if S is convex and $-f$ is convex.

Operations preserving convexity

1. The intersection of an arbitrary number of convex sets is a convex set:

$$\text{if } S_n \text{ is convex } \forall n \in \mathbb{N}^+ \text{ then } \bigcap_{n\in\mathbb{N}^+} S_n \text{ is convex.}$$

   The empty set is convex.

2. The sub-level sets of a convex function f on S are convex:

$$\text{if } f(z) \text{ is convex } \text{ then } S_\alpha \triangleq \{z \in S : f(z) \leq \alpha\} \text{ is convex } \forall\alpha.$$

3. $f_1, \ldots, f_N$ convex $\Rightarrow \sum_{i=1}^{N} \alpha_i f_i$ is convex function $\forall\ \alpha_i \geq 0$, $i = 1, \ldots, N$.
4. The composition of a convex function $f(z)$ with an affine map $z = Ax + b$ generates a convex function $f(Ax + b)$ of x.
5. A linear function $f(z) = c'z + d$ is both convex and concave.
6. A quadratic function $f(z) = z'Qz + 2s'z + r$ is convex if and only if $Q \in \mathbb{R}^{s\times s}$ is positive semidefinite. Strictly convex if $Q \in \mathbb{R}^{s\times s}$ is positive definite.

Operations preserving convexity

- Suppose $f(x) = h(g(x)) = h(g_1(x), \dots, g_k(x))$ with $h : R^k \to R$, $g_i : R^s \to R$. Then,
  1. f is convex if h is convex, h is nondecreasing in each argument, and g_i are convex,
  2. f is convex if h is convex, h is nonincreasing in each argument, and g_i are concave,
  3. f is concave if h is concave, h is nondecreasing in each argument, and g_i are concave.
- The pointwise maximum of a set of convex functions is a convex function:

$$\begin{aligned}&f_1(z), \dots, f_k(z) \text{ convex functions} \Rightarrow \\ &f(z) = \max\{f_1(z), \dots, f_k(z)\} \text{ is a convex function.}\end{aligned}$$

Convex optimization problems

- Problem (5) is *convex* if the cost f is convex on Z and S is convex.
- In convex optimization problems local optimizers are also global optimizers.
- Non-convex optimization problems solved by iterating between the solutions of convex sub-problems.
- Convexity of the feasible set S difficult to prove except in special cases. (See operations preserving the convexity)
- Non-convex problems exist which can be transformed into convex problems through a change of variables and manipulations on cost and constraints.

Outline

1 Main concepts

2 Optimality conditions: Lagrange duality theory and KKT conditions
- Optimality Conditions
- Duality theory
- Certificate of optimality
- Complementarity slackness
- KKT conditions

3 Polyhedra, polytopes and simplices

4 Linear and quadratic programming

Optimality Conditions

Consider the problem

$$\begin{array}{lll} \inf_z & f(z) & \\ \text{subj. to} & g_i(z) \leq 0 & \text{for } i = 1, \ldots, m \\ & h_i(z) = 0 & \text{for } i = 1, \ldots, p \\ & z \in Z & \end{array}$$

- In general, an analytical solution does not exist.
- Solutions are usually computed by recursive algorithms which start from an initial guess z_0 and at step k generate a point z_k such that $\{f(z_k)\}_{k=0,1,2,\ldots}$ converges to J^*.
- These algorithms recursively use and/or solve analytical **conditions for optimality**

Optimality Conditions for **Unconstrained** Optimization Problems

Theorem (Necessary condition*)

$f : \mathbb{R}^s \rightarrow \mathbb{R}$ differentiable at $\bar{z}$. If there exists a vector $\mathbf{d}$ such that $\nabla f(\bar{z})'\mathbf{d} < 0$, then there exists a $\delta > 0$ such that $f(\bar{z} + \lambda \mathbf{d}) < f(\bar{z})$ for all $\lambda \in (0, \delta)$.

- The vector $\mathbf{d}$ in the theorem above is called **descent direction**.
- The direction of **steepest descent** $\mathbf{d}_s$ at $\bar{z}$ is defined as the normalized direction where $\nabla f(\bar{z})'\mathbf{d}_s < 0$ is minimized.
- The direction $\mathbf{d}_s$ of steepest descent is $\mathbf{d}_s = -\frac{\nabla f(\bar{z})}{\|\nabla f(\bar{z})\|}$.

Corollary

$f : \mathbb{R}^s \rightarrow \mathbb{R}$ is differentiable at $\bar{z}$. If $\bar{z}$ is a local minimizer, then $\nabla f(\bar{z}) = 0$.

Optimality Conditions for **Unconstrained** Optimization Problems

Theorem (Sufficient condition*)

Suppose that $f : \mathbb{R}^s \to \mathbb{R}$ *is twice differentiable at* $\bar{z}$. *If* $\nabla f(\bar{z}) = 0$ *and the Hessian of* $f(z)$ *at* $\bar{z}$ *is positive definite, then* $\bar{z}$ *is a local minimizer.*

Theorem (Necessary and sufficient condition*)

Suppose that $f : \mathbb{R}^s \to \mathbb{R}$ *is differentiable at* $\bar{z}$. *If* f *is convex, then* $\bar{z}$ *is a global minimizer if and only if* $\nabla f(\bar{z}) = 0$.

*Proofs available in Chapter 4 of M.S. Bazaraa, H.D. Sherali, and C.M. Shetty. Nonlinear Programming Theory and Algorithms. John Wiley & Sons, Inc., New York, second edition, 1993.

Duality Theory. The Lagrange Function

Consider the **primal** optimization problem

$$\begin{array}{lll} \inf_z & f(z) & \\ \text{subj. to} & g_i(z) \leq 0 & \text{for } i = 1, \ldots, m \\ & h_i(z) = 0 & \text{for } i = 1, \ldots, p \\ & z \in Z & \end{array}$$

Any feasible point: upper bound of the optimal value
Lagrange dual problem: lower bound on optimal value

- Construct **Lagrange function**

$$\begin{array}{rl} L(z,u,v) = & f(z) + u_1 g_1(z) + \ldots + u_m g_m(z) + \\ & + v_1 h_1(z) + \ldots + v_p h_p(z) \end{array}$$

- More compact

$$L(z,u,v) \triangleq f(z) + u' g(z) + v' h(z)$$

- u_i and v_i called **Lagrange multipliers** or dual variables
- objective is augmented with weighted sum of constraint functions

Duality Theory. The Lagrange Function

- Consider **Lagrange function**

$$L(z,u,v) \triangleq f(z) + u'g(z) + v'h(z)$$

Let $z \in S$ be feasible. For arbitrary vectors $u \geq 0$ and v trivially have

$$L(z,u,v) \leq f(z)$$

- After infimization we infer

$$\inf_{z \in Z} L(z,u,v) \leq \inf_{z \in Z,\ g(z) \leq 0,\ h(z)=0} f(z)$$

- Best lower bound: since $u \geq 0$ and v arbitrary

$$\sup_{(u,v),\ u \geq 0} \inf_{z \in Z} L(z,u,v) \leq \inf_{z \in Z,\ g(z) \leq 0,\ h(z)=0} f(z)$$

Duality Theory. The Dual Problem

Let

$$\Theta(u,v) \triangleq \inf_{z \in Z} L(z,u,v) \in [-\infty, +\infty] \tag{8}$$

Lagrangian dual problem

$$\sup_{(u,v),\ u \geq 0} \Theta(u,v) \tag{9}$$

Remarks

- The (8) (*Lagrangiang dual subproblem*) is an **unconstrained optimization problem**. Only points (u,v) with $\Theta(u,v) > -\infty$ are interesting
- $\Theta(u,v)$ always **concave** $\Rightarrow$ the (9) is concave, much easier to solve than the primal (non convex in general)
- **Weak duality** always holds:

$$\sup_{(u,v),\ u \geq 0} \Theta(u,v) \leq \inf_{z \in Z,\ g(z) \leq 0,\ h(z)=0} f(z)$$

Duality Theory. Duality Gap and Certificate of Optimality

Let:

$$d^* = \max_{(u,v),\ u \geq 0} \Theta(u, v)$$

$$J^* = \min_{z \in Z,\ g(z) \leq 0,\ h(z)=0} f(z)$$

then

- we always have $d^* \leq J^*$
- $d^* - J^*$ is called **optimal duality gap**
- **Strong duality** if $d^* = J^*$
- In case of strong duality u^* and v^* serve as **certificate of optimality**

Duality Theory. Constraint Qualifications: The Slater's Condition

When do we have
Dual optimal value = Primal optimal value?

- In general, strong duality does not hold, even for convex primal problems.
- **Constraint qualifications**. Conditions on the constraint functions implying *strong duality for convex problems.*
- **Slater's condition** (A well known constraint qualification)
  Consider the primal problem. There exists $\hat{z} \in \mathbb{R}^s$ which belongs to the relative interior of the problem domain Z, which is feasible ($g(\hat{z}) \leq 0,\ h(\hat{z}) = 0$) and for which $g_j(\hat{z}) < 0$ for all j for which g_j is not an affine function.

Other constraint qualifications exist. Linear Independence, Cottle's, Zangwill's, Kuhn-Tucker's constraint qualifications.

Duality Theory. The Slater's Theorem

When do we have
Dual optimal value = Primal optimal value?

Theorem (Slater's theorem)

Consider the primal problem and its dual problem. If the primal problem is convex and Slater's condition holds then $d^* > -\infty$ and $d^* = J^*$.

- Then

$$\max_{(u,v),\ u\geq 0} \Theta(u,v) = \min_{z\in Z,\ g(z)\leq 0,\ h(z)=0} f(z)$$

- Slater's condition reduces to feasibility when all inequality constraints are linear. **Strong duality holds for convex QP and for LP (feasible)**

Certificate of Optimality

- Be z a feasible point. $f(z)$ is an **upper bound** on the cost, i.e., $J^* \leq f(z)$.
- Be (u, v) a dual feasible point, $\Theta(u, v)$ is an **lower bound** on the cost, i.e., $\Theta(u, v) \leq J^*$.
- If z and (u, v) are primal and dual feasible, respectively, then $\Theta(u, v) \leq J^* \leq f(z)$. Therefore z is ε-suboptimal, with ε equal to primal-dual gap, i.e., $\varepsilon = f(z) - \Theta(u, v)$.
- The optimal value of the primal (and dual) problems will lie in the same interval

  $$J^*, d^* \in [\Theta(u, v), f(z)].$$

  (u, v) is a *certificate* that proves the (sub)optimality of z.
- At iteration k, algorithms produce a primal feasible z_k and a dual feasible u_k, v_k with $f(z_k) - \Theta(u_k, v_k) \to 0$ as $k \to \infty$, hence at iteration k we **know** $J^* \in [\Theta(u_k, v_k), f(z_k)]$ (useful stopping criteria)

Complementary slackness

Suppose that z^*, u^*, v^* are primal, dual feasible with zero duality gap (hence, they are primal, dual optimal)

$$\begin{array}{rcl} f(z^*) & = & \Theta(u^*, v^*) \\ & = & \inf_z \left(f(z) + u^{*\prime} g(z) + v^{*\prime} h(z)\right) \\ & \leq & f(z^*) + u^{*\prime} g(z^*) + v^{*\prime} h(z^*) \end{array}$$

hence we have $\sum_{i=1}^{m} u_i^* g_i(z^*) = 0$ and so

$$u_i^* g_i(z^*) = 0,\ i = 1, \ldots, m$$

- called **complementary slackness** condition
- i-th constraint inactive at optimum $\Longrightarrow u_i = 0$
- $u_i^* > 0$ at optimum $\Longrightarrow i$-th constraint active at optimum

KKT optimality conditions

Suppose

- f, g_i, h_i are differentiable
- z^*, u^*, v^* are (primal, dual) optimal, *with zero duality gap*

by complementary slackness we have

$$\begin{aligned} f(z^*) + \sum_i u_i^* g_i(z^*) + \sum_j v_j^* h_j(z^*) = \\ \min_z \left(f(z) + \sum_i u_i^* g_i(z) + \sum_j v_j^* h_j(z) \right) \end{aligned} \tag{10}$$

i.e., z^* minimizes $L(z, u^*, v^*)$ therefore

$$\nabla f(z^*) + \sum_i u_i^* \nabla g_i(z^*) + \sum_j v_j^* \nabla h_j(z^*) = 0$$

KKT optimality conditions

z^*, (u^*, v^*) of an optimization problem, with differentiable cost and constraints and zero duality gap, have to satisfy the following conditions:

$$
\begin{aligned}
0 &= \nabla f(z^*) + \sum_{i=1}^{m} u_i^* \nabla g_i(z^*) + \sum_{j=1}^{p} v_j^* \nabla h_i(z^*), && \text{(11a)}\\
0 &= u_i^* g_i(z^*), \quad i = 1, \ldots, m && \text{(11b)}\\
0 &\leq u_i^*, \quad i = 1, \ldots, m && \text{(11c)}\\
0 &\geq g_i(z^*), \quad i = 1, \ldots, m && \text{(11d)}\\
0 &= h_j(z^*) \quad j = 1, \ldots, p && \text{(11e)}
\end{aligned}
$$

Conditions (11a)-(11e) are called the *Karush-Kuhn-Tucker* (KKT) conditions.

KKT optimality conditions

Consider the primal problem (5).

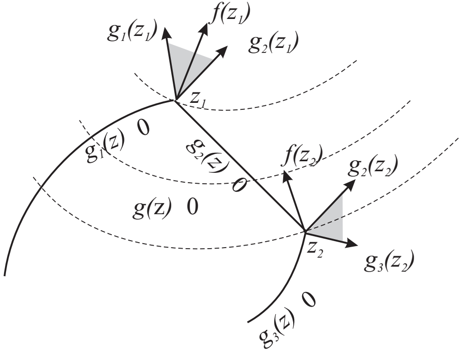
Theorem (Necessary and sufficient condition)

Suppose that problem (5) is convex and that cost and constraints f, g_i *and* h_i *are differentiable at a feasible* z^*. *If problem (5) satisfies Slaters condition then* z^* *is optimal if and only if there are* (u^*, v^*) *that, together with* z^*, *satisfy the KKT conditions.*

Theorem (Necessary and sufficient condition)

Let z^* *be a feasible solution and* $A = \{i : g_i(z^*) = 0\}$ *be the set of active constraints at* z^*. *Suppose that* f, g_i *are differentiable at* z^* $\forall i$ *and that* h_i *are continuously differentiable at* z^* $\forall i$. *Further, suppose that* $\nabla g_i(z^*)$ *for* $i \in A$ *and* $\nabla h_i(z^*)$ *for* $i = 1, \ldots, p$, *are linearly independent. If* z^*, (u^*, v^*) *are optimal, then they satisfy the KKT conditions. In addition, if problem (5) is convex, then* z^* *is optimal if and only if there are* (u^*, v^*) *that, together with* z^*, *satisfy the KKT conditions.*

KKT geometric interpretation



Rewrite the (11a), as

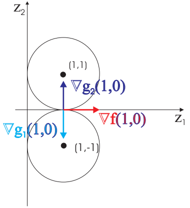
$$-\nabla f(z) = \sum_{i \in I} u_i \nabla g_i(z), \; u_i \geq 0,$$

i.e., the direction of cost steepest descent belongs to the convex cone spanned by ∇g_i's,

KKT conditions. Example

KKT conditions under only the convexity assumptions are not necessary Consider the problem:

$$\begin{array}{rl} \min & z_1 \\ \\ \text{subj. to} & (z_1-1)^2+(z_2-1)^2 \le 1 \\ & (z_1-1)^2+(z_2+1)^2 \le 1 \end{array} \tag{12}$$



KKT are necessary if **constraints qualification** is satisfied:

$\exists\ \hat{z} \in Z$ such that $g(\hat{z}) \le 0,\ h(\hat{z}) = 0$, $g_j(\hat{z}) < 0$ if g_j is not affine and $\hat{z} \in \text{int}Z$

Outline

1. Main concepts
2. Optimality conditions: Lagrange duality theory and KKT conditions
3. Polyhedra, polytopes and simplices
   - General Set Definitions and Operations
   - Polyhedra Definitions and Representations
   - Basic Operations on Polytopes
4. Linear and quadratic programming

General Set Definitions and Operations

- An **n-dimensional ball** $B(x_0, \rho)$ is the set $B(x_0, \rho) = \{x \in \mathbb{R}^n | \sqrt{\|x - x_0\|_2} \leq \rho\}$. x_0 and ρ are the center and the radius of the ball, respectively.
- **Affine sets** are sets described by the solutions of a system of linear equations:

$$F = \{x \in \mathbb{R}^n \ : \ Ax = b, \ \ \text{with} \ \ A \in \mathbb{R}^{m \times n}, \ b \in \mathbb{R}^m\}.$$

- The **affine combination** of $x_1, \ldots, x_k$ is defined as the point $\lambda_1 x_1 + \ldots + \lambda_k x_k$ where $\sum_{i=1}^{k} \lambda_i = 1$.
- The **affine hull** of $K \subseteq \mathbb{R}^n$ is the set of all affine combinations of points in K and it is

$$\text{aff}(K) = \{\lambda_1 x_1 + \ldots + \lambda_k x_k \mid x_i \in \ K, \ i = 1, \ldots, k, \ \sum_{i=1}^{k} \lambda_i = 1\}$$

General Set Definitions and Operations

- The **dimension** of an affine set is the dimension of the largest ball of radius $\rho > 0$ included in the set.
- The **convex combination** of $x_1, \ldots, x_k$ is defined as the point $\lambda_1 x_1 + \ldots + \lambda_k x_k$ where $\sum_{i=1}^{k} \lambda_i = 1$ and $\lambda_i \geq 0$, $i = 1, \ldots, k$.
- The **convex hull** of a set $K \subseteq \mathbb{R}^n$ is the set of all convex combinations of points in K and it is denoted as $\mathrm{conv}(K)$:

$$\mathrm{conv}(K) \triangleq \{\lambda_1 x_1 + \ldots + \lambda_k x_k \mid x_i \in K,\ \lambda_i \geq 0,\ i = 1, \ldots, k, \sum_{i=1}^{k} \lambda_i = 1\}.$$

- A **cone** spanned by a finite set of points $K = \{x_1, \ldots, x_k\}$ is defined as

$$\mathrm{cone}(K) = \{\sum_{i=1}^{k} \lambda_i x_i,\ \lambda_i \geq 0, i = 1, \ldots, k\}.$$

- The **Minkowski sum** of two sets $P, Q \subseteq \mathbb{R}^n$ is defined as

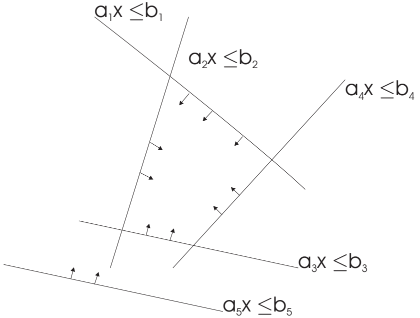
$$P \oplus Q \triangleq \{x + y | x \in P,\ y \in Q\}.$$

Polyhedra Definitions and Representations

An $\mathcal{H}$*-polyhedron* $\mathcal{P}$ in $\mathbb{R}^n$ denotes an intersection of a finite set of closed halfspaces in $\mathbb{R}^n$:

$$\mathcal{P} = \{x \in \mathbb{R}^n : \ Ax \leq b\}$$

A two-dimensional $\mathcal{H}$-polyhedron



Inequalities which can be removed without changing the polyhedron are called *redundant*. The representation of an $\mathcal{H}$-polyhedron is *minimal* if it does not contain redundant inequalities.

Polyhedra Definitions and Representations

- A $\mathcal{V}$**-polyhedron** $\mathcal{P}$ in $\mathbb{R}^n$ denotes the Minkowski sum:

$$\mathcal{P} = \mathrm{conv}(V) \oplus \mathrm{cone}(Y)$$

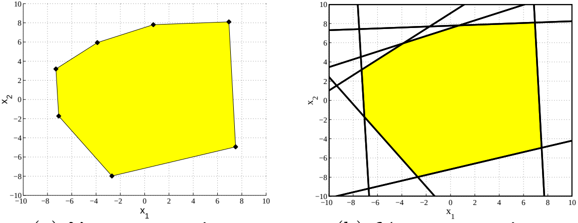
  for some $V = [V_1, \ldots, V_k] \in \mathbb{R}^{n \times k}$, $Y = [y_1, \ldots, y_{k'}] \in \mathbb{R}^{n \times k'}$.
- Any $\mathcal{H}$-polyhedron is a $\mathcal{V}$-polyhedron.
- An $\mathcal{H}$**-polytope** ($\mathcal{V}$**-polytope**) is a bounded $\mathcal{H}$-polyhedron ($\mathcal{V}$-polyhedron). Any $\mathcal{H}$-polytope is a $\mathcal{V}$-polytope
- The **dimension of a polytope (polyhedron)** $\mathcal{P}$ is the dimension of its affine hull and is denoted by $\dim(\mathcal{P})$.
- A polytope $\mathcal{P} \subset \mathbb{R}^n$, is **full-dimensional** if $\dim(\mathcal{P}) = n$ or, equivalently, if it is possible to fit a non-empty n-dimensional ball in $\mathcal{P}$
  Otherwise, we say that polytope $\mathcal{P}$ is **lower-dimensional**.
- If $\|P_i^x\|_2 = 1$, where P_i^x denotes the i-th row of a matrix P^x, we say that the polytope $\mathcal{P}$ is *normalized*.

Polyhedra Definitions and Representations

- A linear inequality $cz \leq c_0$ is said to be **valid** for $\mathcal{P}$ if it is satisfied for all points $z \in \mathcal{P}$.
- A **face** of $\mathcal{P}$ is any nonempty set of the form

$$\mathcal{F} = \mathcal{P} \cap \{z \in \mathbb{R}^s |\ cz = c_0\}$$

  where $cz \leq c_0$ is a *valid* inequality for $\mathcal{P}$.
- The faces of dimension 0,1, dim($\mathcal{P}$)-2 and dim($\mathcal{P}$)-1 are called **vertices**, **edges**, **ridges**, and **facets**, respectively.
- A d**-simplex** is a polytope of $\mathbb{R}^d$ with $d+1$ vertices.



(a) $\mathcal{V}$-representation. (b) $\mathcal{H}$-representation.

Polytopal Complexes

A set $\mathcal{C} \subseteq \mathbb{R}^n$ is called a **P-collection** (in $\mathbb{R}^n$) if it is a collection of a finite number of n-dimensional polytopes, i.e.

$$\mathcal{C} = \{\mathcal{C}_i\}_{i=1}^{N_C},$$

where $\mathcal{C}_i := \{x \in \mathbb{R}^n \ : \ C_i^x x \leq C_i^c\}$, $\dim(\mathcal{C}_i) = n$, $i = 1, \ldots, N_C$, with $N_C < \infty$.
The *underlying set* of a P-collection $\mathcal{C} = \{\mathcal{C}_i\}_{i=1}^{N_C}$ is the point set

$$\underline{\mathcal{C}} := \bigcup_{\mathcal{P} \in \mathcal{C}} \mathcal{P} = \bigcup_{i=1}^{N_C} \mathcal{C}_i.$$

Special Classes

- A collection of sets $\{\mathcal{C}_i\}_{i=1}^{N_C}$ is a **strict partition** of a set $\mathcal{C}$ if *(i)* $\bigcup_{i=1}^{N_C} \mathcal{C}_i = \mathcal{C}$ and *(ii)* $\mathcal{C}_i \cap \mathcal{C}_j = \emptyset$, $\forall i \neq j$.
- $\{\mathcal{C}_i\}_{i=1}^{N_C}$ is a **strict polyhedral partition** of a polyhedral set $\mathcal{C}$ if $\{\mathcal{C}_i\}_{i=1}^{N_C}$ is a strict partition of $\mathcal{C}$ and $\bar{\mathcal{C}}_i$ is a polyhedron for all i, where $\bar{\mathcal{C}}_i$ denotes the closure of the set $\mathcal{C}_i$
- A collection of sets $\{\mathcal{C}_i\}_{i=1}^{N_C}$ is a **partition** of a set $\mathcal{C}$ if *(i)* $\bigcup_{i=1}^{N_C} \mathcal{C}_i = \mathcal{C}$ and *(ii)* $(\mathcal{C}_i \backslash \partial \mathcal{C}_i) \cap (\mathcal{C}_j \backslash \partial \mathcal{C}_j) = \emptyset$, $\forall i \neq j$.

Functions on Polytopal Complexes

- A function $h(\theta) : \Theta \to \mathbb{R}^k$, where $\Theta \subseteq \mathbb{R}^s$, is **piecewise affine (PWA)** if there exists a strict partition $R_1,\ldots,R_N$ of Θ and $h(\theta) = H^i\theta + k^i$, $\forall\theta \in R_i$, $i = 1, \ldots, N$.
- A function $h(\theta) : \Theta \to \mathbb{R}^k$, where $\Theta \subseteq \mathbb{R}^s$, is **piecewise affine on polyhedra** (PPWA) if there exists a strict polyhedral partition $R_1,\ldots,R_N$ of Θ and $h(\theta) = H^i\theta + k^i$, $\forall\theta \in R_i$, $i = 1, \ldots, N$.
- A function $h(\theta) : \Theta \to \mathbb{R}$, where $\Theta \subseteq \mathbb{R}^s$, is **piecewise quadratic (PWQ)** if there exists a strict partition $R_1,\ldots,R_N$ of Θ and $h(\theta) = \theta' H^i\theta + k^i\theta + l^i$, $\forall\theta \in R_i$, $i = 1, \ldots, N$.
- A function $h(\theta) : \Theta \to \mathbb{R}$, where $\Theta \subseteq \mathbb{R}^s$, is **piecewise quadratic on polyhedra** (PPWQ) if there exists a strict polyhedral partition $R_1,\ldots,R_N$ of Θ and $h(\theta) = \theta' H^i\theta + k^i\theta + l^i$, $\forall\theta \in R_i$, $i = 1, \ldots, N$.

Basic Operations on Polytopes

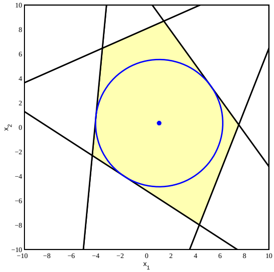
- **Convex Hull** of a set of points $V = \{V_i\}_{i=1}^{N_V}$, with $V_i \in \mathbb{R}^n$,

$$\mathrm{conv}(V) = \{x \in \mathbb{R}^n \ : x = \sum_{i=1}^{N_V} \alpha_i V_i,\ 0 \le \alpha_i \le 1,\ \sum_{i=1}^{N_V} \alpha_i = 1\}. \tag{13}$$

  Used to switch from a $\mathcal{V}$-representation of a polytope to an $\mathcal{H}$-representation.
- **Vertex Enumeration** of a polytope $\mathcal{P}$ given in $\mathcal{H}$-representation. (dual of the convex hull operation)

Basic Operations on Polytopes

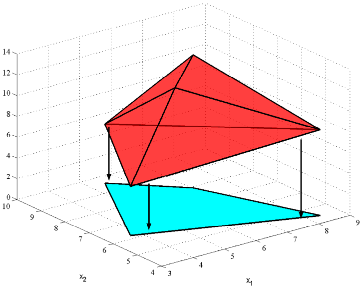
- **Polytope reduction** is the computation of the minimal representation of a polytope. A polytope $\mathcal{P} \subset \mathbb{R}^n$, $\mathcal{P} = \{x \in \mathbb{R}^n \ : \ P^x x \leq P^c\}$ is in a **minimal representation** if the removal of any row in $P^x x \leq P^c$ would change it (i.e., if there are no redundant constraints).
- The **Chebychev Ball** of a polytope $\mathcal{P} = \{x \in \mathbb{R}^n \mid P^x x \leq P^c\}$, with $P^x \in \mathbb{R}^{n_P \times n}$, $P^c \in \mathbb{R}^{n_P}$, corresponds to the largest radius ball $\mathcal{B}(x_c, R)$ with center x_c, such that $\mathcal{B}(x_c, R) \subset \mathcal{P}$.



Basic Operations on Polytopes

- **Projection** Given a polytope $\mathcal{P} = \{[x'y']' \in \mathbb{R}^{n+m} \ : \ P^x x + P^y y \leq P^c\} \subset \mathbb{R}^{n+m}$ the projection onto the x-space $\mathbb{R}^n$ is defined as

$$\text{proj}_x(\mathcal{P}) := \{x \in \mathbb{R}^n \mid \exists y \in \mathbb{R}^m : P^x x + P^y y \leq P^c\}.$$



Basic Operations on Polytopes

- **Set-Difference** The set-difference of two polytopes $\mathcal{Y}$ and $\mathcal{R}_0$

$$\mathcal{R} = \mathcal{Y} \setminus \mathcal{R}_0 := \{x \in \mathbb{R}^n \ : \ x \in \mathcal{Y}, x \notin \mathcal{R}_0\},$$

in general, can be a nonconvex and disconnected set and can be described as a P-collection $\mathcal{R} = \bigcup_{i=1}^{m} \mathcal{R}_i$, where $\mathcal{Y} = \bigcup_{i=1}^{m} \mathcal{R}_i \bigcup (\mathcal{R}_0 \bigcap \mathcal{Y})$.
The P-collection $\mathcal{R} = \bigcup_{i=1}^{m} \mathcal{R}_i$ can be computed by consecutively inverting the half-spaces defining $\mathcal{R}_0$ as described in the following

Theorem

Let $\mathcal{Y} \subseteq \mathbb{R}^n$ be a polyhedron, $\mathcal{R}_0 \triangleq \{x \in \mathbb{R}^n : Ax \leq b\}$, and $\bar{\mathcal{R}}_0 \triangleq \{x \in \mathcal{Y} : Ax \leq b\} = \mathcal{R}_0 \bigcap \mathcal{Y}$, where $b \in \mathbb{R}^m$, $\mathcal{R}_0 \neq \emptyset$ and $Ax \leq b$ is a minimal representation of $\mathcal{R}_0$. Also let

$$\mathcal{R}_i = \left\{ \begin{array}{ll} x \in \mathcal{Y} : & \begin{array}{l} A^i x > b^i \\ A^j x \leq b^j, \forall j < i \end{array} \end{array} \right\} i = 1, \ldots, m$$

Let $\mathcal{R} \triangleq \bigcup_{i=1}^{m} \mathcal{R}_i$. Then, $\mathcal{R}$ is a P-collection and $\{\bar{\mathcal{R}}_0, \mathcal{R}_1, \ldots, \mathcal{R}_m\}$ is a strict polyhedral partition of $\mathcal{Y}$.

Basic Operations on Polytopes

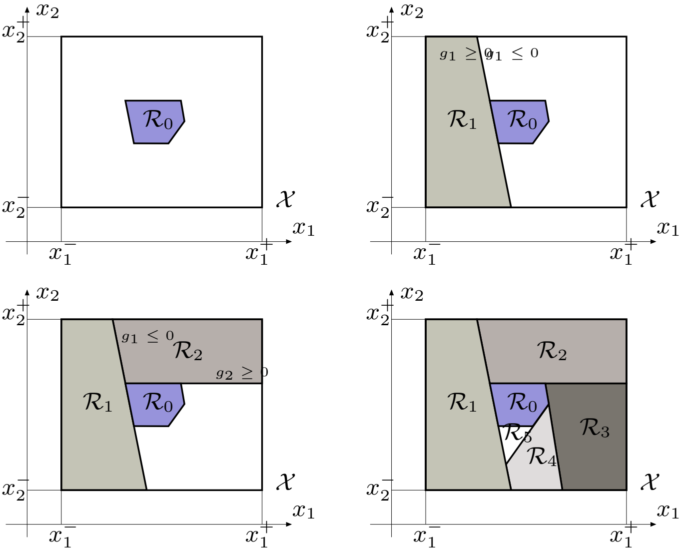


Figure: Two dimensional example: partition of the rest of the space $\mathcal{X} \backslash \mathcal{R}_0$.

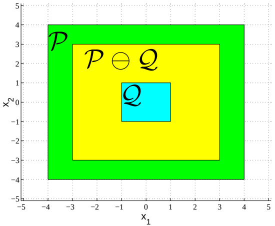
Basic Operations on Polytopes

- **Pontryagin Difference** The Pontryagin difference (also known as Minkowski difference) of two polytopes $\mathcal{P}$ and $\mathcal{Q}$ is a polytope

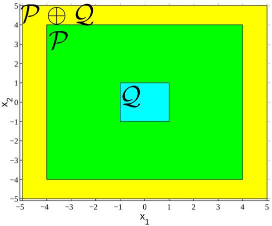
$$\mathcal{P} \ominus \mathcal{Q} := \{x \in \mathbb{R}^n \ : \ x + q \in \mathcal{P}, \ \forall q \in \mathcal{Q}\}.$$

- The **Minkowski sum** of two polytopes $\mathcal{P}$ and $\mathcal{Q}$ is a polytope

$$\mathcal{P} \oplus \mathcal{Q} := \{x \in \mathbb{R}^n \ : \exists y \in \mathcal{P}, \ \exists z \in \mathcal{Q}, \ x = y + z\}.$$



(a) Pontryagin difference $\mathcal{P} \ominus \mathcal{Q}$.



(b) Minkowski sum $\mathcal{P} \oplus \mathcal{Q}$.

Minkowski Sum of Polytopes

- The Minkowski sum is computationally expensive.
  Consider

$$P = \{y \in \mathbb{R}^n \mid P^y y \leq P^c\}, \quad \mathcal{Q} = \{z \in \mathbb{R}^n \mid Q^z z \leq Q^c\},$$

  it holds that

$$\begin{aligned}
W &= P \oplus Q \\
&= \Big\{x \in \mathbb{R}^n \mid \exists y \; P^y y \leq P^c, \; \exists z \; Q^z z \leq Q^c, \; y, z \in \mathbb{R}^n, \; x = y + z\Big\} \\
&= \Big\{x \in \mathbb{R}^n \mid \exists y \in \mathbb{R}^n, \text{ s.t. } P^y y \leq P^c, \; Q^z (x - y) \leq Q^c\Big\} \\
&= \left\{x \in \mathbb{R}^n \mid \exists y \in \mathbb{R}^n, \text{ s.t. } \begin{bmatrix} 0 & P^y \\ Q^z & -Q^z \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \leq \begin{bmatrix} P^c \\ Q^c \end{bmatrix} \right\} \\
&= \mathrm{proj}_x \left(\left\{ [x' y'] \in \mathbb{R}^{n+n} \mid \begin{bmatrix} 0 & P^y \\ Q^z & -Q^z \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \leq \begin{bmatrix} P^c \\ Q^c \end{bmatrix} \right\} \right).
\end{aligned}$$

Pontryagin Difference of Polytopes

- The Pontryagin difference is “not computationally expensive”.
- Consider

$$\mathcal{P} = \{y \in \mathbb{R}^n \text{ s.t. }, P^y y \leq P^b\}, \qquad \mathcal{Q} = \{z \in \mathbb{R}^n \text{s.t. } Q^z z \leq Q^b\},$$

  Then:

$$\mathcal{P} \ominus \mathcal{Q} = \{x \in \mathbb{R}^n \text{ s.t.} P^y x \leq P^b - H(P^y, \mathcal{Q})\}$$

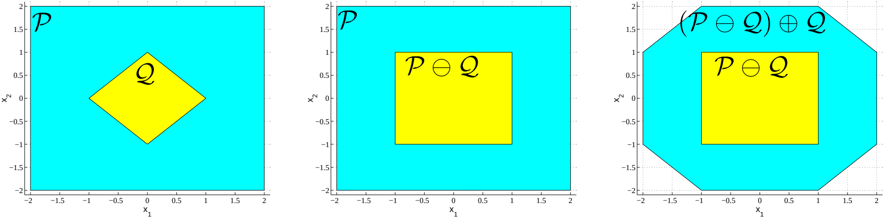
  where the i-th element of $H(P^y, \mathcal{Q})$ is

$$H_i(P^y, \mathcal{Q}) \triangleq \max_{x \in \mathcal{Q}} P_i^y x$$

  and P_i^y is the i-th row of the matrix P^y.
- For special cases (e.g. when $\mathcal{Q}$ is a hypercube), more efficient computational methods exist.

Basic Operations on Polytopes

Note that $(\mathcal{P} \ominus \mathcal{Q}) \oplus \mathcal{Q} \subseteq \mathcal{P}$.



(c) Two polytopes $\mathcal{P}$ and $\mathcal{Q}$.

(d) Polytope $\mathcal{P}$ and Pontryagin difference $\mathcal{P} \ominus \mathcal{Q}$.

(e) Polytope $\mathcal{P} \ominus \mathcal{Q}$ and the set $(\mathcal{P} \ominus \mathcal{Q}) \oplus \mathcal{Q}$.

Figure: Illustration that $(\mathcal{P} \ominus \mathcal{Q}) \oplus \mathcal{Q} \subseteq \mathcal{P}$.

Affine Mappings and Polyhedra

- Consider a polyhedron $\mathcal{P} = \{x \in \mathbb{R}^n \mid P^x x \le P^c\}$, with $P^x \in \mathbb{R}^{n_P \times n}$ and an affine mapping $f(z)$

$$f: \; z \in \mathbb{R}^n \mapsto Az + b, \; A \in \mathbb{R}^{n \times n}, \; b \in \mathbb{R}^n$$

- Define the composition of $\mathcal{P}$ and f as the following polyhedron

$$\mathcal{P} \circ f \triangleq \{z \in \mathbb{R}^m \mid P^x f(z) \le P^c\} = \{z \in \mathbb{R}^m \mid P^x A z \le P^c - P^x b\}$$

- Useful for backward-reachability

Affine Mappings and Polyhedra

- Consider a polyhedron $\mathcal{P} = \{x \in \mathbb{R}^n \mid P^x x \leq P^c\}$, with $P^x \in \mathbb{R}^{n_P \times n}$ and an affine mapping $f(z)$

$$f: \; z \in \mathbb{R}^n \mapsto Az + b, \; A \in \mathbb{R}^{n \times n}, \; b \in \mathbb{R}^n$$

- Define the composition of f and $\mathcal{P}$ as the following polyhedron

$$f \circ \mathcal{P} \triangleq \{y \in \mathbb{R}^{m_A} \mid y = Ax + b \; \forall x \in \mathbb{R}^n, \; P^x x \leq P^c\}$$

- The polyhedron $f \circ \mathcal{P}$ in can be computed as follows. Write $\mathcal{P}$ in $\mathcal{V}$-representation $\mathcal{P} = \text{conv}(V)$ and map the vertices $V = \{V_1, \ldots, V_k\}$ through the transformation f. Because the transformation is affine, the set $f \circ \mathcal{P}$ is the convex hull of the transformed vertices

$$f \circ \mathcal{P} = \text{conv}(F), \; F = \{AV_1 + b, \ldots, AV_k + b\}.$$

- If f is invertible $x = A^{-1}y - A^{-1}b$ and therefore

$$f \circ \mathcal{P} = \{y \in \mathbb{R}^{m_A} \mid P^x A^{-1} y \leq P^c + P^x A^{-1} b\}$$

- Useful for forward-reachability

Outline

1. Main concepts
2. Optimality conditions: Lagrange duality theory and KKT conditions
3. Polyhedra, polytopes and simplices
4. Linear and quadratic programming

Linear Programming

$$\begin{array}{ll} \inf_z & c'z \\ \text{subj. to} & Gz \leq W \end{array}$$

where $z \in \mathbb{R}^s$.

- Convex optimization problems.
- Other common forms:

$$\begin{array}{ll} \inf_z & c'z \\ \text{subj. to} & Gz \leq W \\ & G_{eq}z = W_{eq} \end{array}$$

or

$$\begin{array}{ll} \inf_z & c'z \\ \text{subj. to} & G_{eq}z = W_{eq} \\ & z \geq 0 \end{array}$$

- Always possible to convert one of the three forms into the other.

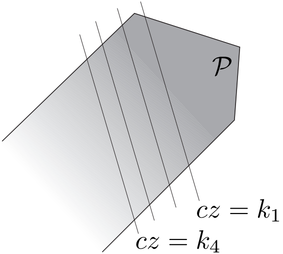
Graphical Interpretation and Solutions Properties

- Let $\mathcal{P}$ be the feasible set. $\mathcal{P}$ is a polyhedron.
- If $\mathcal{P}$ is empty, then the problem is infeasible.
- Denote by J^* the optimal value and by Z^* the set of optimizers $Z^* = \text{argmin}_{z \in \mathcal{P}} \; c'z$

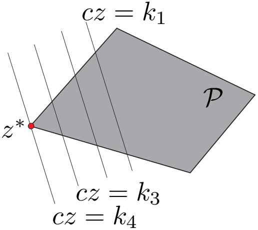
Case 1. The LP solution is unbounded, i.e., $J^* = -\infty$.

Case 2. The LP solution is bounded, i.e., $J^* > -\infty$ and the optimizer is unique. Z^* is a singleton.

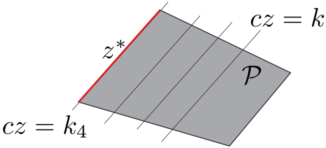
Case 3. The LP solution is bounded and there are multiple optima. Z^* is an uncountable subset of $\mathbb{R}^s$ which can be bounded or unbounded.



(a) Case 1



(b) Case 2



(c) Case 3

Dual of LP

Consider the LP

$$\begin{array}{ll} \inf_z & c'z \\ \text{subj. to} & Gz \leq W \end{array} \tag{14}$$

with $z \in \mathbb{R}^s$ and $G \in \mathbb{R}^{m \times s}$. The Lagrange function is

$$L(z, u) = c'z + u'(Gz - W).$$

The dual cost is

$$\Theta(u) = \inf_z L(z, u) = \inf_z (c' + u'G)z - u'W = \begin{cases} -u'W & \text{if} - G'u = c \\ -\infty & \text{if} - G'u \neq c \end{cases}$$

Since we are interested only in cases where Θ is finite,

$$\begin{array}{ll} \sup_u & -u'W \\ \text{subj. to} & -G'u = c \\ & u \geq 0 \end{array} \qquad \begin{array}{ll} \inf_u & W'u \\ \text{subj. to} & G'u = -c \\ & u \geq 0 \end{array}$$

KKT condition for LP

The KKT conditions for the LP (14) become

$$\begin{aligned} G'u &= -c, && \text{(15a)}\\ (G_j z - W_j)u_j &= 0, && \text{(15b)}\\ u &\geq 0, && \text{(15c)}\\ Gz &\leq W && \text{(15d)} \end{aligned}$$

which are: primal feasibility (15d), dual feasibility (15a), (15c) and slackness complementary conditions (15b).

Active Constraints and Degeneracies

Be $J \triangleq \{1, \dots, m\}$ the set of constraint indices and consider the sets

$$\begin{aligned} A(z) &\triangleq \{j \in J : G_j z = W_j\}, \text{ of active constraints at feasible z} \\ NA(z) &\triangleq \{j \in J : G_j z < W_j\}, \text{ of inactive constraints at feasible z.} \end{aligned}$$

C 1. $A(z^*)$ undefined, since z^* is undefined

C 2. The cardinality of $A(z^*)$ can be between s and m. Higher than s regardless whether $\mathcal{P}$ is minimal or not (see figure below)

C 3. If $\mathcal{P}$ is minimal, then the cardinality of $A(z^*)$ is $s - p$ with p the dimension of the optimal face at all points $z^* \in Z^*$ contained in the *relative interior* of the optimal face, otherwise it can be any number p with $1 < p \leq m$.

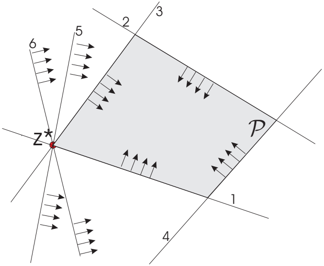


Figure: Primal Degeneracy in a Linear Program

Active Constraints and Degeneracies

Definition (Primal degenerate)

The LP is said to be primal degenerate if there exists a $z^* \in Z^*$ such that the number of active constraints at z^* is greater than the number of variables s.

Definition (Dual degenerate)

The LP is said to be dual degenerate if its dual problem is primal degenerate.

Homework. Show that if the primal problem has multiple optima, then the dual problem is primal degenerate (i.e., the primal problem is dual degenerate). The converse is not always true. Details in the book.

Convex Piecewise Linear Optimization

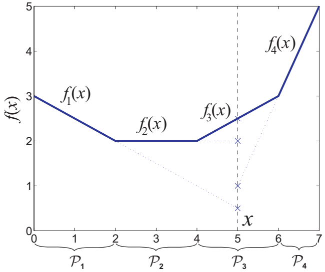
Consider

$$J^* = \begin{array}{ll} \min_z & J(z) \\ \text{subj. to} & Gz \leq W \end{array} \tag{16}$$

where the cost function has the form

$$J(z) = \max_{i=1,\dots,k} \{c_i z + d_i\} \tag{17}$$

where $c_i \in \mathbb{R}^s$ and $d_i \in \mathbb{R}$.



Convex Piecewise Linear Optimization

The cost function $J(z)$ in (17) is a convex PWA function. The optimization problem (16)-(17) can be solved by the following linear program:

$$\begin{array}{rll} J^* = & \min_{z,\varepsilon} & \varepsilon \\ & \text{subj. to} & Gz \leq W \\ & & c_i z + d_i \leq \varepsilon, \quad i = 1, \ldots, k \end{array}$$

Convex Piecewise Linear Optimization

Consider

$$
\begin{array}{rll}
J^* = & \min_z & J_1(z_1) + J_2(z_2) \\
& \text{subj. to} & G_1 z_1 + G_2 z_2 \leq W
\end{array}
\tag{18}
$$

where the cost function has the form

$$
\begin{array}{l}
J_1(z_1) = \max_{i=1,\ldots,k}\{c_i z_1 + d_i\} \\
J_2(z_2) = \max_{i=1,\ldots,j}\{m_i z_2 + n_i\}
\end{array}
\tag{19}
$$

The optimization problem (18)-(19) can be solved by the following linear program:

$$
\begin{array}{rll}
J^* = & \min_{z,\varepsilon_1,\varepsilon_2} & \varepsilon_1 + \varepsilon_2 \\
& \text{subj. to} & G_1 z_1 + G_2 z_2 \leq W \\
& & c_i z_1 + d_i \leq \varepsilon_1, \quad i = 1, \ldots, k \\
& & m_i z_2 + n_i \leq \varepsilon_2, \quad i = 1, \ldots, j
\end{array}
$$

Example

The optimization problem:

$$\begin{array}{rc} \min\limits_{z_1,z_2} f(x,u) = & \min |z_1+5| + |z_2-3| \\ \text{subject to} & 2.5 \leq z_1 \leq 5 \\ & -1 \leq z_2 \leq 1 \end{array}$$

can be solved in Matlab by using "v=linprog(f,A,b)" where $v = [\varepsilon_1^*, \varepsilon_2^*, z_1^*, z_2^*]$, $f = [1 \;\; 1 \;\; 0 \;\; 0]$,

$$A = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & -1 \\ -1 & 0 & 1 & 0 \\ -1 & 0 & -1 & 0 \\ 0 & -10 & 0 & 1 \\ 0 & -10 & 0 & -1 \end{bmatrix}, \;\; b = \begin{bmatrix} 5 \\ -2.5 \\ 1 \\ 1 \\ -5 \\ 5 \\ 3 \\ -3 \end{bmatrix}$$

Solution is $v = [7.5 \;\; 0.2 \;\; 2.5 \;\; 1]$

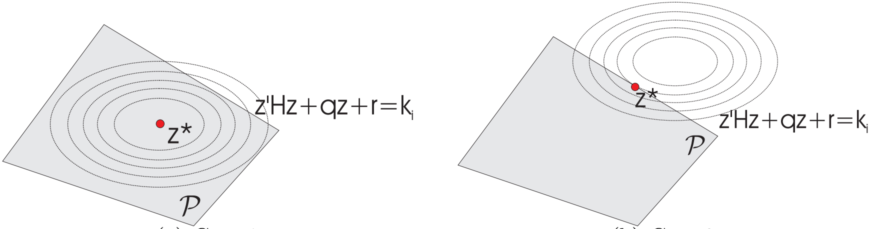
Quadratic Programming

$$\begin{array}{ll} \min_z & \frac{1}{2} z' H z + q' z + r \\ \text{subj. to} & Gz \leq W \end{array} \tag{20}$$

where $z \in \mathbb{R}^s$, $H = H' > 0 \in \mathbb{R}^{s \times s}$.

- Other QP forms often include equality and inequality constraints.
- Let $\mathcal{P}$ be the feasible set. Two cases can occur if $\mathcal{P}$ is not empty:

Case 1. The optimizer lies strictly inside the feasible polyhedron
Case 2. The optimizer lies on the boundary of the feasible polyhedron



(a) Case 1 (b) Case 2

Dual of QP

$$\begin{array}{ll} \min_z & \frac{1}{2}z'Hz + q'z \\ \text{subj. to} & Gz \leq W \end{array}$$

The Lagrange function is

$$L(z,u) = \{\frac{1}{2}z'Hz + q'z + u'(Gz - W)\}$$

The the dual cost is

$$\Theta(u) = \min_z \{\frac{1}{2}z'Hz + q'z + u'(Gz - W)\} \tag{21}$$

and the dual problem is

$$\max_{u \geq 0} \min_z \{\frac{1}{2}z'Hz + q'z + u'(Gz - W)\}.$$

For a given u the Lagrange function $\frac{1}{2}z'Hz + q'z + u'(Gz - W)$ is convex. Therefore it is necessary and sufficient for optimality that the gradient is zero

$$Hz + q + G'u = 0.$$

Dual of QP

From the previous equation we can derive $z = -H^{-1}(q + G'u)$ and substituting this in equation (21) we obtain:

$$\Theta(u) = -\frac{1}{2}u'(GH^{-1}G')u - u'(W + GH^{-1}q) - \frac{1}{2}q'H^{-1}q \tag{22}$$

By using (22) the dual problem can be rewritten as:

$$\begin{array}{ll} \min_u & \frac{1}{2}u'(GH^{-1}G')u + u'(W + GH^{-1}q) + \frac{1}{2}q'H^{-1}q \\ \text{subj. to} & u \geq 0 \end{array}$$

KKT condition for QP

Consider the QP (20), $\nabla f(z) = Hz + q$, $g_i(z) = G_i' z - W_i$, $\nabla g_i(z) = G_i$. The KKT conditions become

$$
\begin{aligned}
Hz + q + G'u &= 0 && \text{(23a)}\\
u_i(G_i' z - W_i) &= 0 && \text{(23b)}\\
u &\geq 0 && \text{(23c)}\\
Gz - W &\leq 0 && \text{(23d)}
\end{aligned}
$$

Active Constraints and Degeneracies

Let $J \triangleq \{1, \ldots, m\}$ be the set of constraint indices. Consider the set of active and inactive constraints at feasible z:

$$\begin{aligned} A(z) \quad &\triangleq \{j \in J : G_j z = W_j\} \\ NA(z) \quad &\triangleq \{j \in J : G_j z < W_j\}. \end{aligned}$$

- We have two cases:

Case 1. $A(z^*) = \{\emptyset\}$.
Case 2. $A(z^*)$ is a nonempty subset of $\{1, \ldots, m\}$.

- The QP is said to be primal degenerate if there exists a $z^* \in Z^*$ such that the number of active constraints at z^* is greater than the number of variables s.
- The QP is said to be dual degenerate if its dual problem is primal degenerate.

Constrained Least-Squares Problems

The problem of minimizing the convex quadratic function arises in many fields and has many names, e.g., linear regression or least-squares approximation.

$$\|Az - b\|_2^2 = z'A'Az - 2b'Az + b'b$$

The minimizer is

$$z^* = (A'A)^{-1}A'b \triangleq A^\dagger b$$

When linear inequality constraints are added, the problem is called constrained linear regression or *constrained least-squares*, and there is no longer a simple analytical solution. As an example we can consider regression with lower and upper bounds on the variables, i.e.,

$$\begin{array}{ll} \min_z & \|Az - b\|_2^2 \\ \text{subj. to} & l_i \le z_i \le u_i, \;\; i = 1, \ldots, n, \end{array}$$

which is a QP.

Example

Consider

$$\min_{z} \quad \|Az - b\|_2^2$$

where

$$A = \begin{bmatrix} 0.7513 & 0.5472 & 0.8143 \\ 0.2551 & 0.1386 & 0.2435 \\ 0.5060 & 0.1493 & 0.9293 \\ 0.6991 & 0.2575 & 0.3500 \\ 0.8909 & 0.8407 & 0.1966 \\ 0.9593 & 0.2543 & 0.2511 \end{bmatrix}, \quad b = \begin{bmatrix} 0.6160 \\ 0.4733 \\ 0.3517 \\ 0.8308 \\ 0.5853 \\ 0.5497 \end{bmatrix}$$

- Unconstrained Least-Squares: in Matlab "$z = A\backslash b$" or "z=quadprog($A'A$,$-b'A$)". $z^* = [0.7166, -0.0205, 0.1180]$,
- Assume $z_2 \geq 0$. Constrained Least-Squares in Matlab: "z=quadprog($A'A$,$-b'A$,[0 -1 0],0)". $z^* = [0.7045,\ 0,\ 0.1194]$.

Nonlinear Programming

Consider

$$\begin{array}{lll} \min_z & f(z) & \\ \text{subj. to} & g_i(z) \leq 0 & \text{for } i = 1, \ldots, m \\ & h_i(z) = 0 & \text{for } i = 1, \ldots, p \\ & z \in Z & \end{array}$$

- A variety of softwares exists
- In general, global optimality not guaranteed
- Solutions are usually computed by recursive algorithms which start from an initial guess z_0 and at step k generate a point z_k such that $\{f(z_k)\}_{k=0,1,2,\ldots}$ converges to J^*.
- These algorithms recursively use and/or solve analytical **conditions for optimality**
- In this class we will use "NPSOL"

Nonlinear Programming - NPSOL

- Possible syntax:

  [INFORM,ITER,ISTATE,C,CJAC,CLAMDA,OBJF,OBJGRAD,R,X] =
  npsol(X0,A,L,U,'funobj','funcon',OPTION);

- Solving

$$
\begin{array}{ll}
\min_z & \text{funobj}(z) \\
\text{subj. to} & L \leq \begin{bmatrix} z \\ Az \\ \text{funcon}(z) \end{bmatrix} \leq U
\end{array}
$$

- NPSOL Manual and Example on bSpace
- Note it is a mex-function

Nonlinear Programming - Matlab Optimization toolbox

- Possible syntax:

  [X,FVAL,EXITFLAG] =
  fmincon(funobj,X0,A,B,Aeq,Beq,LB,UB,NONLCON);

- Solving

$$
\begin{array}{ll}
\min_z & \text{funobj}(z) \\
\text{subj. to} & A\ z \leq B \\
& Aeq\ z = Beq \\
& LB \leq\ z \leq UB \\
& Ceq(z) = 0 \\
& C(z) \leq 0
\end{array}
$$

- The function NONLCON accepts z and returns the vectors $C(z)$ and $Ceq(z)$
- help "fmincon" in Matlab