CHALMERS

UNIVERSITY OF TECHNOLOGY

SSY281 - MODEL PREDICTIVE CONTROL

NIKOLCE MURGOVSKI

Division of Systems and Control
Department of Electrical Engineering
Chalmers University of Technology
Gothenburg, Sweden

2021-02-09

Lecture 7: Optimisation basics and convexity

Goals for today:

- To formulate a general constrained optimisation problem
- To formulate necessary conditions for optimality
- To master the basics of convex sets and convex functions
- To formulate a standard convex optimisation problem
- To characterize a standard quadratic program (QP)

Lecture 7: Optimisation basics and convexity

Goals for today:

- To formulate a general constrained optimisation problem
- To formulate necessary conditions for optimality
- To master the basics of convex sets and convex functions
- To formulate a standard convex optimisation problem
- To characterize a standard quadratic program (QP)

Learning objectives:

- Understand and explain basic properties of the optimisation problem as an ingredient of MPC, in particular concepts like linear, quadratic and convex optimisation, optimality conditions, and feasibility

Constrained optimisation problem

A basic optimisation problem is formulated as

$$\begin{aligned} \text{minimize} \quad & f(x) \\ \text{subject to} \quad & g_i(x) \leq 0, \quad i = 1, \dots, m \\ & h_i(x) = 0, \quad i = 1, \dots, p \end{aligned} \tag{60}$$

where

$x = \{x_1, \dots, x_n\}$ are the optimisation or decision variables

$f : \mathbb{R}^n \to \mathbb{R}$ is the objective or cost function

$g_i : \mathbb{R}^n \to \mathbb{R},\ i = 1, \dots, m$ are inequality constraint functions

$h_i : \mathbb{R}^n \to \mathbb{R},\ i = 1, \dots, p$ are equality constraint functions.

Constrained optimisation problem

A basic optimisation problem is formulated as

$$\begin{aligned} \text{minimize} \quad & f(x) \\ \text{subject to} \quad & g_i(x) \leq 0, \quad i = 1, \dots, m \\ & h_i(x) = 0, \quad i = 1, \dots, p \end{aligned} \tag{60}$$

where

$x = \{x_1, \dots, x_n\}$ are the optimisation or decision variables

$f : \mathbb{R}^n \to \mathbb{R}$ is the objective or cost function

$g_i : \mathbb{R}^n \to \mathbb{R},\ i = 1, \dots, m$ are inequality constraint functions

$h_i : \mathbb{R}^n \to \mathbb{R},\ i = 1, \dots, p$ are equality constraint functions.

The *optimal solution* x^* has the smallest value of $f(\cdot)$ among all vectors x that belong to $\mathrm{dom}\, f$ (the *domain* of f, i.e. the subset of $\mathbb{R}^n$ where f is defined) and satisfy the constraints. The *optimal value* p^* is always defined,

$$p^* = \inf\{f(x) \mid g_i(x) \leq 0,\ i = 1, \dots, m; h_j(x) = 0,\ j = 1, \dots, p\}$$

$$p^* = \infty, \quad \text{if problem is infeasible}$$

$$p^* = -\infty, \quad \text{if problem is unbounded below.}$$

Examples of functions and their optima

- $f(x) = 1/x, \operatorname{dom} f = \mathbb{R}_{++}$(strictly positive reals) : $p^* = 0,$ no optimal solution
- $f(x) = -\log x, \operatorname{dom} f = \mathbb{R}_{++} : p^* = -\infty$
- $f(x) = x \log x, \operatorname{dom} f = \mathbb{R}_{++} : p^* = -1/e, x^* = 1/e$
- $f(x) = x^3 - 3x : p^* = -\infty,$ local optimum at $x = 1$.

Feasible directions

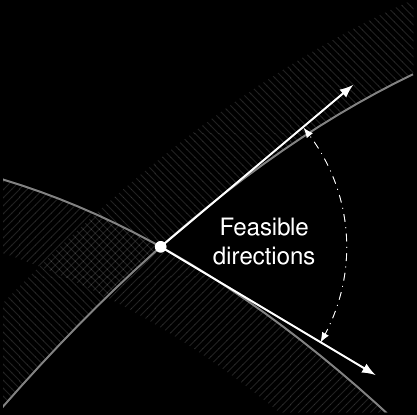

Figure 11: Feasible directions from the point of intersection of two nonlinear constraints.

Conditions for local optimality – unconstrained case

- **First order necessary condition**

$$x^* \text{ is a stationary point} \quad \Rightarrow \quad \nabla f(x^*) = 0. \tag{61}$$

- **Second order sufficient conditions**

$$\nabla f(x^*) = 0 \text{ and } \nabla^2 f(x^*) \succ 0 \quad \Rightarrow \quad x^* \text{ is a strict local minimum.} \tag{62}$$

First order necessary conditions – equality constraints

Consider the optimisation problem

$$\begin{aligned} \text{minimize} \quad & f(x) \\ \text{subject to} \quad & h(x) = 0. \end{aligned}$$

Assume x^* is a local minimum and that x^* is regular. Then there is a unique vector λ^* such that

$$\begin{aligned} \nabla f(x^*) + \nabla h(x^*)\lambda^* &= 0 \\ h(x^*) &= 0. \end{aligned}$$

This is a system of non-linear equations having $n + p$ equations for the $n + p$ unknowns (x and λ). The vector λ contains the *Lagrange multipliers* $\lambda_i,\ i = 1, \ldots, p$.

First order necessary conditions – the KKT conditions

Consider the optimisation problem

$$\begin{aligned} \text{minimize} \quad & f(x) \\ \text{subject to} \quad & g(x) \le 0 \\ & h(x) = 0. \end{aligned} \tag{63}$$

Assume x^* is a local minimum and that x^* is regular. Then there are unique vectors μ^* and λ^* such that

$$\nabla f(x^*) + \nabla g(x^*)\mu^* + \nabla h(x^*)\lambda^* = 0 \tag{64a}$$

$$\mu^* \ge 0 \tag{64b}$$

$$g(x^*) \le 0, \quad h(x^*) = 0 \tag{64c}$$

$$\mu_i^* g_i(x^*) = 0, \quad i = 1, \dots, m. \tag{64d}$$

These conditions are referred to as the *KKT (Karush-Kuhn-Tucker) conditions*.

The Lagrangian and complementary slackness

The first condition (64a) can conveniently be expressed as a condition on the *Lagrangian* $\mathcal{L}$,

$$\nabla_x \mathcal{L}(x^*, \mu^*, \lambda^*) = 0, \qquad \mathcal{L}(x, \mu, \lambda) = f(x) + \mu^\top g(x) + \lambda^\top h(x). \tag{65}$$

The last of the KKT conditions, $\mu_i^* g_i(x^*) = 0$, are called the *complementary slackness* conditions. The implication of these is that if g_i is inactive at x^* then $\mu_i^* = 0$. Conversely, if g_i is active, then either $\mu_i > 0$ (the constraint is *strictly active*) or $\mu^* = 0$ (the constraint is not strictly active, i.e. it is *weakly active*).

Intuition of the KKT conditions

- $\nabla\mathcal{L}(x^*,\mu^*) = \nabla f(x^*) + \mu^*\nabla g(x^*) = 0.$
- $-\nabla f(x^*)$ is analogous to *gravitational force* at x^*.
- $-\mu^*\nabla g(x^*)$ is analogous to *reactive force* from the constraint.

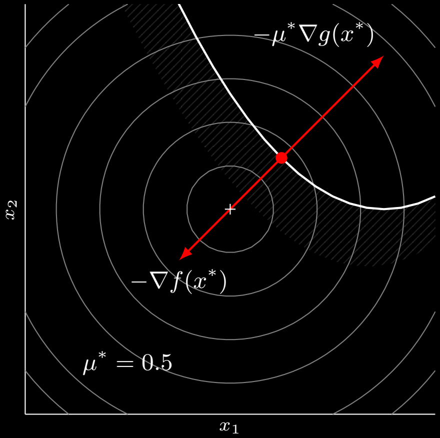

Figure 12: An illustration of KKT conditions for a system with two variables. The objective function f, depicted by contour lines, pulls the solution towards the middle, marked by a plus sign. The inequality constraint prevents reaching this solution. Instead, the optimum, i.e. the point of *minimum potential energy*, is achieved at the point marked by a filled circle.

Intuition of the KKT conditions

- $\nabla \mathcal{L}(x^*, \mu^*) = \nabla f(x^*) + \mu^* \nabla g(x^*) = 0.$
- $-\nabla f(x^*)$ is analogous to *gravitational force* at x^*.
- $-\mu^* \nabla g(x^*)$ is analogous to *reactive force* from the constraint.

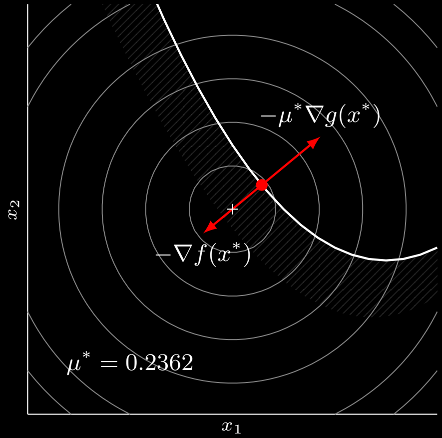

Figure 12: An illustration of KKT conditions for a system with two variables. The objective function f, depicted by contour lines, pulls the solution towards the middle, marked by a plus sign. The inequality constraint prevents reaching this solution. Instead, the optimum, i.e. the point of *minimum potential energy*, is achieved at the point marked by a filled circle.

Intuition of the KKT conditions

- $\nabla \mathcal{L}(x^*, \mu^*) = \nabla f(x^*) + \mu^* \nabla g(x^*) = 0.$
- $-\nabla f(x^*)$ is analogous to *gravitational force* at x^*.
- $-\mu^* \nabla g(x^*)$ is analogous to *reactive force* from the constraint.

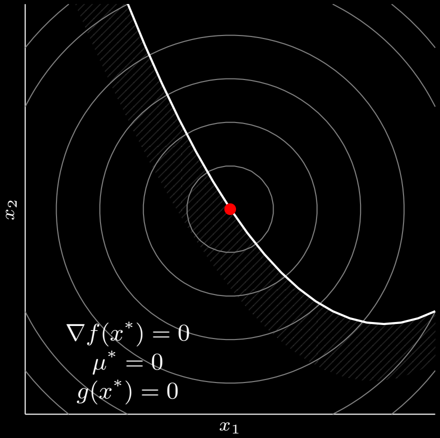

Figure 12: An illustration of KKT conditions for a system with two variables. The objective function f, depicted by contour lines, pulls the solution towards the middle, marked by a plus sign. The inequality constraint prevents reaching this solution. Instead, the optimum, i.e. the point of *minimum potential energy*, is achieved at the point marked by a filled circle.

Intuition of the KKT conditions

- $\nabla \mathcal{L}(x^*, \mu^*) = \nabla f(x^*) + \mu^* \nabla g(x^*) = 0.$
- $-\nabla f(x^*)$ is analogous to *gravitational force* at x^*.
- $-\mu^* \nabla g(x^*)$ is analogous to *reactive force* from the constraint.

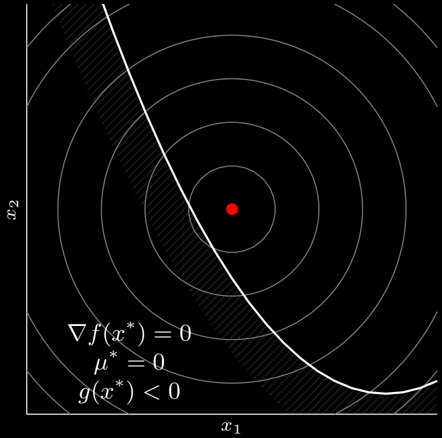

Figure 12: An illustration of KKT conditions for a system with two variables. The objective function f, depicted by contour lines, pulls the solution towards the middle, marked by a plus sign. The inequality constraint prevents reaching this solution. Instead, the optimum, i.e. the point of *minimum potential energy*, is achieved at the point marked by a filled circle.

Second order sufficient conditions

Consider the optimisation problem

$$\begin{aligned} \text{minimize} \quad & f(x) \\ \text{subject to} \quad & g(x) \leq 0 \\ & h(x) = 0. \end{aligned} \tag{66}$$

Assume x^* is regular and that x^*, μ^*, λ^* satisfy the KKT conditions with all active constraints being strictly active. Further assume that

$$d^\top \nabla_x^2 \mathcal{L}(x^*, \mu^*, \lambda^*) d > 0, \quad \text{for all } d \text{ such that } d^\top \begin{bmatrix} \nabla g_{\mathbb{A}} & \nabla h \end{bmatrix} = 0,$$

where $\nabla_x^2 \mathcal{L}$ is the Hessian of the Lagrangian. Then x^* is a local minimum.

Convex optimisation problem

$$
\begin{aligned}
\text{minimize} \quad & f(x) \\
\text{subject to} \quad & g_i(x) \leq 0, \quad i = 1, \ldots, m \\
& h_i(x) = 0, \quad i = 1, \ldots, p
\end{aligned}
$$

where the objective function f and constraint functions $\{g_i\}$ are convex, i.e.

$$f(\theta x_1 + (1-\theta)x_2) \leq \theta f(x_1) + (1-\theta) f(x_2), \quad 0 \leq \theta \leq 1$$

and the functions $\{h_i\}$ are *affine* (linear).

Affine sets

- A **line** through x_1 and x_2 are all points x,

$$x = \theta x_1 + (1 - \theta) x_2, \quad \theta \in \mathbb{R}.$$

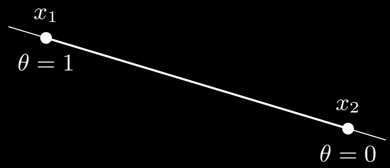

Figure 13: An illustration of an affine set in two dimensions.

- An **affine set** contains the line through any two distinct points in the set, see Figure 13. An example is the solution set of linear equations $\{x \mid Ax = b\}$.
  All affine sets can be described as solutions to a system of linear equations.

Convex sets

- A **line segment** between x_1 and x_2 are all points x,

$$x = \theta x_1 + (1-\theta)x_2, \quad 0 \leq \theta \leq 1.$$

- A **convex set** contains the line segments between every two points in the set (see Figure 14), i.e.

$$x_1, x_2 \in \mathcal{S} \;\Rightarrow\; \theta x_1 + (1-\theta)x_2 \in \mathcal{S}, \quad 0 \leq \theta \leq 1.$$

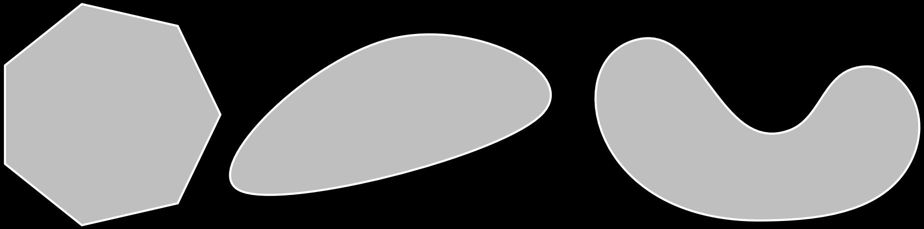

Figure 14: An example with two convex and one non-convex set.

Hyperplanes and half-spaces

- A **hyperplane** is a set of the form $\{x \mid a^\top x = b\}$.
- A **half-space** is a set of the form $\{x \mid a^\top x \leq b\}$.
  Hyperplanes are affine and convex; half-spaces are convex, see Figure 15.

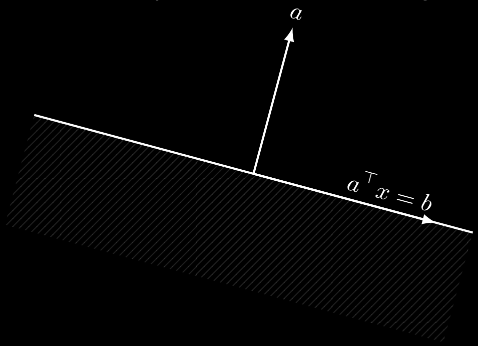

Figure 15: A hyperplane $a^\top x = b$ and a half-space depicted by the shaded region.

Polyhedra

- A **polyhedron** is the intersection of a finite number of half-spaces and hyperplanes or, equivalently, the solution set of a finite number of linear inequalities and equalities (see Figure 16), i.e.

$$
\begin{aligned}
Ax &\le b \\
Cx &= d.
\end{aligned}
$$

Polyhedra are convex sets.

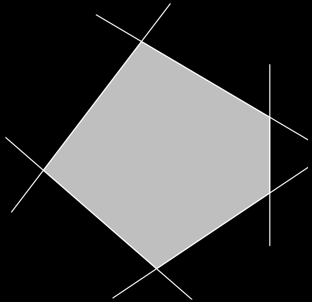

Figure 16: An example of a polyhedron, as the intersection of linear inequalities and equalities.

Convex functions

A function $f : \mathbb{R}^n \to \mathbb{R}$ is *convex* if $\operatorname{dom} f$ is convex and

$$f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta) f(y)$$

for all $x, y \in \operatorname{dom} f$ and $0 \leq \theta \leq 1$. An illustration is provided in Figure 17.

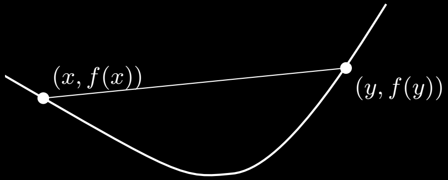

Figure 17: An example of a convex function.

Convex functions

A function $f : \mathbb{R}^n \to \mathbb{R}$ is *convex* if $\mathrm{dom}\, f$ is convex and

$$f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta) f(y)$$

for all $x, y \in \mathrm{dom}\, f$ and $0 \leq \theta \leq 1$. An illustration is provided in Figure 17.

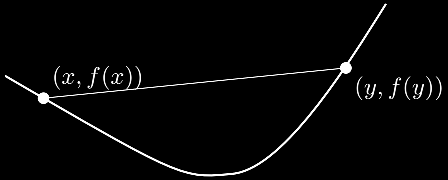

Figure 17: An example of a convex function.

Furthermore,

- f is *concave* if $-f$ is convex.
- f is *strictly convex* if

$$f(\theta x + (1-\theta)y) < \theta f(x) + (1-\theta) f(y)$$

for all $x, y \in \mathrm{dom}\, f$ and $0 < \theta < 1$.

Examples of convex and concave functions

Some examples of convex functions are

- affine: $a^\top x + b$;
- exponential: e^{ax};
- powers: $x^\alpha, \quad x > 0$, for $\alpha \geq 1$ or $\alpha \leq 0$.

Examples of convex and concave functions

Some examples of convex functions are

- affine: $a^\top x + b$;
- exponential: e^{ax};
- powers: $x^\alpha, \quad x > 0$, for $\alpha \geq 1$ or $\alpha \leq 0$.

Examples of concave functions are

- affine: $a^\top x + b$;
- logarithm: $\log x, \ x > 0$;
- powers: $x^\alpha, \quad x > 0$, for $0 \leq \alpha \leq 1$.

First and second order conditions

- Differentiable f with convex domain is convex if and only if

$$f(y) \geq f(x) + \nabla f(x)^\top (y - x) \quad \text{for all } x, y \in \operatorname{dom} f.$$

An illustration is provided in Figure 18.

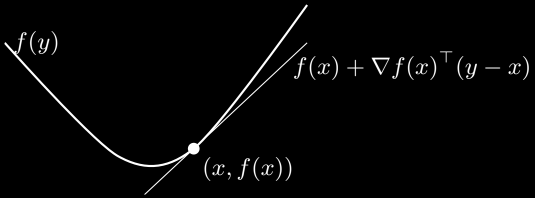

Figure 18: A convex function and its tangent.

First and second order conditions

- Differentiable f with convex domain is convex if and only if

$$f(y) \geq f(x) + \nabla f(x)^\top (y - x) \quad \text{for all } x, y \in \mathrm{dom}\, f.$$

An illustration is provided in Figure 18.

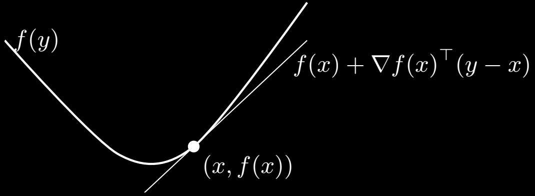

Figure 18: A convex function and its tangent.

- Twice differentiable f with convex domain is convex if and only if

$$\nabla^2 f(x) \succeq 0 \quad \text{for all } x \in \mathrm{dom}\, f.$$

Operations that preserve convexity

- The **intersection** of convex sets is a convex set.
- If f is **affine** ($f(x) = Ax + b$), then the image of a convex set under f is convex, i.e.

$$\mathcal{S} \text{ convex } \Rightarrow f(\mathcal{S}) \text{ convex.}$$

- If f is affine, then the **inverse image** of f is convex, i.e.

$$\mathcal{S} \text{ convex } \Rightarrow f^{-1}(\mathcal{S}) = \{x \mid f(x) \in \mathcal{S}\} \text{ convex.}$$

Examples include scaling, translation, projection.

Operations preserving convexity

- **Sub-level sets** $\mathcal{S}_\alpha$ of a convex function f are convex

$$\mathcal{S}_\alpha = \{x \in \mathrm{dom}\, f \mid f(x) \leq \alpha\}.$$

- **Nonnegative weighted sum** of convex functions is convex,

$$f_1, \ldots, f_N \text{ convex} \quad \Rightarrow \sum_{i=1}^{N} \alpha_i f_i \text{ convex, for all } \alpha_i \geq 0.$$

- **The composition with an affine function** is convex,

$$f \text{ convex} \quad \Rightarrow \quad f(Ax + b) \text{ convex.}$$

Convex optimisation problem

Standard form convex optimisation problem:

$$\begin{aligned} \text{minimize} \quad & f(x) \\ \text{subject to} \quad & g_i(x) \leq 0, \quad i = 1, \dots, m \\ & Ax = b, \quad \text{(affine equality constraints)} \end{aligned}$$

where f and $\{g_i\}$ are convex.

Remark

The feasible set of a convex optimisation problem is convex.

An example of a *convex* problem

Consider the following optimisation problem:

$$\begin{aligned} \text{minimize} \quad & f(x) = x_1^2 + x_2^2 \\ \text{subject to} \quad & g_1(x) = x_1/(1+x_2^2) \leq 0 \\ & h_1(x) = (x_1 + x_2)^2 = 0. \end{aligned}$$

It is not difficult to see that

- $f(x)$ is convex and the feasible set $\{(x_1, x_2) \mid x_1 = -x_2 \leq 0\}$ is convex, but ...
- the problem is not in the standard form, since g_1 is not convex and h_1 is not affine.

An example of a *convex* problem

Consider the following optimisation problem:

$$\begin{aligned} \text{minimize} \quad & f(x) = x_1^2 + x_2^2 \\ \text{subject to} \quad & g_1(x) = x_1/(1+x_2^2) \leq 0 \\ & h_1(x) = (x_1 + x_2)^2 = 0. \end{aligned}$$

It is not difficult to see that

- $f(x)$ is convex and the feasible set $\{(x_1, x_2) \mid x_1 = -x_2 \leq 0\}$ is convex, but ...
- the problem is not in the standard form, since g_1 is not convex and h_1 is not affine.

It is, however, in this case possible to transform the given optimisation problem into an equivalent, convex formulation:

$$\begin{aligned} \text{minimize} \quad & x_1^2 + x_2^2 \\ \text{subject to} \quad & x_1 \leq 0 \\ & x_1 + x_2 = 0. \end{aligned}$$

Optimality conditions for convex problems

Consider the convex optimisation problem

$$\begin{aligned} \text{minimize} \quad & f(x) \\ \text{subject to} \quad & g(x) \leq 0 \\ & h(x) = 0 \end{aligned} \tag{67}$$

where f and $\{g_i\}$ are convex and h is affine. Assume x^* is regular. Then x^* is globally optimal if and only if the KKT conditions are fulfilled for some $\mu^* \geq 0$, λ^*.

Examples of convex optimisation problems

- **Linear programming (LP)**:

$$\begin{aligned} \text{minimize} \quad & c^\top x + d \\ \text{subject to} \quad & Gx \leq h \\ & Ax = b; \end{aligned}$$

Examples of convex optimisation problems

- **Linear programming (LP)**:

$$\begin{aligned} \text{minimize} \quad & c^\top x + d \\ \text{subject to} \quad & Gx \leq h \\ & Ax = b; \end{aligned}$$

- **Quadratic programming (QP)**:

$$\begin{aligned} \text{minimize} \quad & \frac{1}{2} x^\top Q x + p^\top x, \quad Q \succeq 0 \\ \text{subject to} \quad & Gx \leq h \\ & Ax = b. \end{aligned}$$

Examples of convex optimisation problems

- **Linear programming (LP)**:

$$\begin{array}{rl} \text{minimize} & c^\top x + d \\ \text{subject to} & Gx \leq h \\ & Ax = b; \end{array}$$

- **Quadratic programming (QP)**:

$$\begin{array}{rl} \text{minimize} & \frac{1}{2} x^\top Q x + p^\top x, \quad Q \succeq 0 \\ \text{subject to} & Gx \leq h \\ & Ax = b. \end{array}$$

In both cases, the feasible set is a polyhedron.

QP with inequality constraints only

Consider the QP problem

$$\begin{aligned} \text{minimize} \quad & \frac{1}{2}x^\top Qx + p^\top x, \quad Q \succeq 0 \\ \text{subject to} \quad & Gx \leq h. \end{aligned}$$

Assuming that G has full row rank, any point x is regular. Then global optimality is equivalent to the KKT conditions being fulfilled. Denoting the objective by $f(x)$, this can be stated in a simplified way as follows: the point x^* is optimal if and only if x^* is feasible (i.e. $Gx^* \leq h$) and

$$-\nabla f(x^*) = -(Qx^* + p) = \sum_{i \in \mathbb{A}} \mu_i G_i^\top, \quad \text{for some } \{\mu_i\} \text{ with } \mu_i \geq 0, \tag{68}$$

where G_i is the ith row of G and $\mathbb{A}$ is the active set.

Geometric interpretation of a quadratic program with inequality constraints

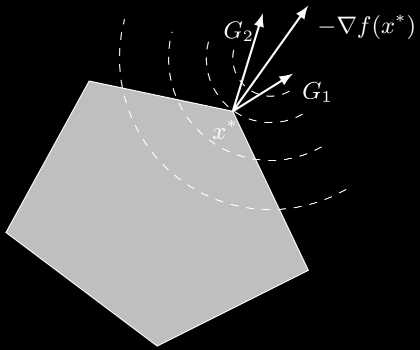

Figure 19: An illustration of a quadratic program with inequality constraints.

The Lagrangian

Consider the standard form problem (not necessarily convex)

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & g_i(x) \leq 0, \quad i = 1, \ldots, m \\ & h_i(x) = 0, \quad i = 1, \ldots, p \end{array}$$

where $x \in \mathcal{D} \subseteq \mathbb{R}^n$.

The Lagrangian

Consider the standard form problem (not necessarily convex)

$$\begin{array}{rll} \text{minimize} & f(x) & \\ \text{subject to} & g_i(x) \leq 0, & i = 1, \dots, m \\ & h_i(x) = 0, & i = 1, \dots, p \end{array}$$

where $x \in \mathcal{D} \subseteq \mathbb{R}^n$.
The *Lagrangian* $\mathcal{L} : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p \to \mathbb{R}$, with $\operatorname{dom} \mathcal{L} = \mathcal{D} \times \mathbb{R}^m \times \mathbb{R}^p$, is defined as

$$\mathcal{L}(x, \mu, \lambda) = f(x) + \sum_{i=1}^{m} \mu_i g_i(x) + \sum_{i=1}^{p} \lambda_i h_i(x) \equiv f(x) + \mu^\top g(x) + \lambda^\top h(x)$$

where

- μ_i is Lagrange multiplier associated with the constraint $g_i(x) \leq 0$;
- λ_i is Lagrange multiplier associated with the constraint $h_i(x) = 0$.

The Lagrangian

Consider the standard form problem (not necessarily convex)

$$
\begin{array}{ll}
\text{minimize} & f(x) \\
\text{subject to} & g_i(x) \leq 0, \quad i = 1, \dots, m \\
 & h_i(x) = 0, \quad i = 1, \dots, p
\end{array}
$$

where $x \in \mathcal{D} \subseteq \mathbb{R}^n$.
The *Lagrangian* $\mathcal{L} : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p \to \mathbb{R}$, with $\operatorname{dom} \mathcal{L} = \mathcal{D} \times \mathbb{R}^m \times \mathbb{R}^p$, is defined as

$$\mathcal{L}(x, \mu, \lambda) = f(x) + \sum_{i=1}^{m} \mu_i g_i(x) + \sum_{i=1}^{p} \lambda_i h_i(x) \equiv f(x) + \mu^\top g(x) + \lambda^\top h(x)$$

where

- μ_i is Lagrange multiplier associated with the constraint $g_i(x) \leq 0$;
- λ_i is Lagrange multiplier associated with the constraint $h_i(x) = 0$.

Note that for $\mu \geq 0$ and any feasible x, we have $\mathcal{L}(x, \mu, \lambda) \leq f(x)$.

Lagrange dual function

Lagrange dual function $q : \mathbb{R}^m \times \mathbb{R}^p \to \mathbb{R}$,

$$q(\mu, \lambda) = \inf_{x \in \mathcal{D}} \mathcal{L}(x, \mu, \lambda) = \inf_{x \in \mathcal{D}} \left\{ f(x) + \mu^\top g(x) + \lambda^\top h(x) \right\}.$$

Properties:

- q is concave but may be $-\infty$ for some μ, λ;
- $q(\mu, \lambda) \leq p^*$ if $\mu \geq 0$ (p^* is the optimal value of the original problem).

The dual problem

The *Lagrange dual problem*

$$\begin{array}{ll} \text{maximize} & q(\mu, \lambda) \\ \text{subject to} & \mu \geq 0 \end{array}$$

- finds the best lower bound d^* on the primal optimal solution p^*,
- always is a convex, unconstrained problem,
- has *dual feasible* μ, λ if $\mu \geq 0$ and $(\mu, \lambda) \in \operatorname{dom} q$,
- always satisfies $d^* \leq p^*$ (*weak duality*).

The dual of an LP problem

Consider the standard linear program

$$\begin{aligned} \text{minimize} \quad & c^\top x \\ \text{subject to} \quad & Ax = b, \quad x \geq 0. \end{aligned}$$

The Lagrangian is given by

$$\mathcal{L}(x,\mu,\lambda) = c^\top x - \mu^\top x + \lambda^\top (Ax - b) = -b^\top \lambda + (c + A^\top \lambda - \mu)^\top x$$

The dual of an LP problem

Consider the standard linear program

$$\begin{aligned} \text{minimize} \quad & c^\top x \\ \text{subject to} \quad & Ax = b, \quad x \geq 0. \end{aligned}$$

The Lagrangian is given by

$$\mathcal{L}(x,\mu,\lambda) = c^\top x - \mu^\top x + \lambda^\top (Ax - b) = -b^\top \lambda + (c + A^\top \lambda - \mu)^\top x$$

The Lagrange dual function is then obtained by solving

$$q(\mu,\lambda) = \inf_x \mathcal{L}(x,\mu,\lambda) = \begin{cases} -b^\top \lambda, & A^\top \lambda - \mu + c = 0 \\ -\infty, & \text{otherwise.} \end{cases}$$

The dual of an LP problem

Consider the standard linear program

$$\begin{aligned} \text{minimize} \quad & c^\top x \\ \text{subject to} \quad & Ax = b, \quad x \geq 0. \end{aligned}$$

The Lagrangian is given by

$$\mathcal{L}(x, \mu, \lambda) = c^\top x - \mu^\top x + \lambda^\top (Ax - b) = -b^\top \lambda + (c + A^\top \lambda - \mu)^\top x$$

The Lagrange dual function is then obtained by solving

$$q(\mu, \lambda) = \inf_x \mathcal{L}(x, \mu, \lambda) = \begin{cases} -b^\top \lambda, & A^\top \lambda - \mu + c = 0 \\ -\infty, & \text{otherwise.} \end{cases}$$

It holds:

- q is linear on the affine domain $\{(\mu, \lambda) \mid A^\top \lambda - \mu + c = 0\}$, i.e. concave
- lower bound: $p^* \geq -b^\top \lambda$ if $A^\top \lambda + c \geq 0$.

The dual of an LP problem

Consider the standard linear program

$$\begin{aligned} \text{minimize} \quad & c^\top x \\ \text{subject to} \quad & Ax = b, \quad x \geq 0. \end{aligned}$$

The Lagrangian is given by

$$\mathcal{L}(x, \mu, \lambda) = c^\top x - \mu^\top x + \lambda^\top (Ax - b) = -b^\top \lambda + (c + A^\top \lambda - \mu)^\top x$$

The Lagrange dual function is then obtained by solving

$$q(\mu, \lambda) = \inf_x \mathcal{L}(x, \mu, \lambda) = \begin{cases} -b^\top \lambda, & A^\top \lambda - \mu + c = 0 \\ -\infty, & \text{otherwise.} \end{cases}$$

It holds:

- q is linear on the affine domain $\{(\mu, \lambda) \mid A^\top \lambda - \mu + c = 0\}$, i.e. concave
- lower bound: $p^* \geq -b^\top \lambda$ if $A^\top \lambda + c \geq 0$.

We can make the implicit constraint $(\mu, \lambda) \in \{(\mu, \lambda) \mid A^\top \lambda - \mu + c = 0\}$ explicit when formulating the dual problem:

$$\begin{aligned} \text{maximize} \quad & -b^\top \lambda \\ \text{subject to} \quad & A^\top \lambda + c \geq 0. \end{aligned}$$

Weak and strong duality

Weak duality: $d^* \leq p^*$

- always holds (even for non-convex problems);
- gives lower bound for the original (*primal*) problem.

Strong duality: $d^* = p^*$

- does not hold in general;
- often holds for convex problems;
- conditions that guarantee this are called *constraint qualifications*.

Constraint qualifications

Strong duality holds for a convex problem

$$\begin{aligned} \text{minimize} \quad & f(x) \\ \text{subject to} \quad & g_i(x) \le 0, \quad i = 1, \dots, m \\ & Ax = b \end{aligned}$$

if any of the following conditions are fulfilled:

1. The gradients of the equality constraints and the active inequality constraints are linearly independent (LICQ);
2. The problem is strictly feasible, i.e. there exists some $\tilde{x} \in \operatorname{int}\mathcal{D}$ (the interior of $\mathcal{D}$) such that (Slater CQ)

$$g_i(\tilde{x}) < 0, \quad i = 1, \dots, m; \quad A\tilde{x} = b.$$

Quadratic programming

Consider the quadratic program (assuming P is positive definite, $P \succ 0$)

$$\begin{aligned} \text{minimize} \quad & x^\top P x \\ \text{subject to} \quad & Ax \le b. \end{aligned}$$

The dual function is

$$q(\mu) = \inf_x (x^\top P x + \mu^\top (Ax - b)) = -\frac{1}{4}\mu^\top A P^{-1} A^\top \mu - b^\top \mu.$$

Quadratic programming

Consider the quadratic program (assuming P is positive definite, $P \succ 0$)

$$\begin{array}{ll} \text{minimize} & x^\top P x \\ \text{subject to} & Ax \leq b. \end{array}$$

The dual function is

$$q(\mu) = \inf_x (x^\top P x + \mu^\top (Ax - b)) = -\frac{1}{4}\mu^\top A P^{-1} A^\top \mu - b^\top \mu.$$

Hence, the dual problem is defined as

$$\begin{array}{ll} \text{maximize} & -\frac{1}{4}\mu^\top A P^{-1} A^\top \mu - b^\top \mu \\ \text{subject to} & \mu \geq 0. \end{array}$$

Quadratic programming

Consider the quadratic program (assuming P is positive definite, $P \succ 0$)

$$\begin{aligned} \text{minimize} \quad & x^\top P x \\ \text{subject to} \quad & Ax \leq b. \end{aligned}$$

The dual function is

$$q(\mu) = \inf_x (x^\top P x + \mu^\top (Ax - b)) = -\frac{1}{4}\mu^\top A P^{-1} A^\top \mu - b^\top \mu.$$

Hence, the dual problem is defined as

$$\begin{aligned} \text{maximize} \quad & -\frac{1}{4}\mu^\top A P^{-1} A^\top \mu - b^\top \mu \\ \text{subject to} \quad & \mu \geq 0. \end{aligned}$$

It follows from Slater's condition that $p^* = d^*$ if $A\tilde{x} < b$ for some $\tilde{x}$. In fact, for convex quadratic programs, we *always have strong duality*, i.e. $p^* = d^*$.

Convex problems with strong duality

Consider the standard convex optimisation problem

$$\begin{aligned} \text{minimize} \quad & f(x) \\ \text{subject to} \quad & g_i(x) \leq 0, \quad i = 1, \ldots, m \\ & Ax = b \end{aligned}$$

where f and $\{g_i\}$ are convex.
Assume x^* is regular. Then the following statements are equivalent:

- x^* is a global optimum;
- there are μ^*, λ^* such that the KKT conditions hold.

References

[1] J.B. Rawlings, D.Q. Mayne, and M.M. Diehl. *Model Predictive Control: Theory, Computation, and Design, 2nd edition*. Nob Hill Publishing 2017.
Available online at `https://sites.engineering.ucsb.edu/~jbraw/mpc`

[2] G. Goodwin, M.M. Seron, and J.A. De Don·. *Constrained Control and Estimation*. Springer 2004.
Available online via Chalmers Library.

[3] F Borrelli, A. Bemporad, and M. Morari. *Predictive Control for Linear and Hybrid Systems*.
Available online at http://www.mpc.berkeley.edu/mpc-course-material

[4] J. Maciejowski. *Predictive Control with Constraints*. Prentice Hall 2002.

[5] S. Boyd and L. Vandenberghe. *Convex optimisation*. Cambridge University Press 2004.

[6] Diehl, M. *Real-Time Optimization for Large Scale Nonlinear Processes*. PhD thesis, University of Heidelberg, 2001.

AVANCEZ
1829
CHALMERS
UNIVERSITY OF TECHNOLOGY